



**HAL**  
open science

# The Advent of Quantum Crystallography: Form and Structure Factors from Quantum Mechanics for Advanced Structure Refinement and Wavefunction Fitting

Simon Grabowsky, Alessandro Genoni, Sajesh P Thomas, Dylan Jayatilaka

► **To cite this version:**

Simon Grabowsky, Alessandro Genoni, Sajesh P Thomas, Dylan Jayatilaka. The Advent of Quantum Crystallography: Form and Structure Factors from Quantum Mechanics for Advanced Structure Refinement and Wavefunction Fitting. *Structure and Bonding*, pp.1-80, 2020, 10.1007/430\_2020\_62 . hal-02972990

**HAL Id: hal-02972990**

**<https://hal.univ-lorraine.fr/hal-02972990>**

Submitted on 20 Oct 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# The Advent of Quantum Crystallography: Form and Structure Factors from Quantum Mechanics for Advanced Structure Refinement and Wavefunction Fitting

Simon Grabowsky,<sup>a</sup> Alessandro Genoni,<sup>b</sup> Sajesh P. Thomas,<sup>c</sup> Dylan Jayatilaka<sup>d</sup>

**Abstract:** X-ray diffraction experiments contain much more information than the information usually exploited for structure determination. In quantum crystallography, quantum mechanical wavefunctions are used to extract that information about bonding and properties from the measured X-ray structure factors. Here we show how quantum mechanically derived structure factors and atomic form factors are constructed to allow the improved description of the diffraction experiment. Subsequently, we discuss the basics and the applications of the advanced structure refinement method Hirshfeld Atom Refinement and of the X-ray constrained wavefunction fitting technique.

**Keywords:** Quantum Crystallography, Hirshfeld Atom Refinement, X-ray Constrained Wavefunction Fitting, Atomic Form Factors, Structure Factors

---

<sup>a</sup> University of Bern, Department of Chemistry and Biochemistry, Freiestrasse 3, 3012 Bern, Switzerland; [simon.grabowsky@dcb.unibe.ch](mailto:simon.grabowsky@dcb.unibe.ch)

<sup>b</sup> Université de Lorraine & CNRS, Laboratoire de Physique et Chimie Théoriques (LPCT), UMR CNRS 7019, 1 Boulevard Arago, F-57078 Metz, France; [Alessandro.Genoni@univ-lorraine.fr](mailto:Alessandro.Genoni@univ-lorraine.fr)

<sup>c</sup> Department of Chemistry and iNano, Aarhus University, Langelandsgade 140, Aarhus 8000, Denmark; [sajesh@inano.au.dk](mailto:sajesh@inano.au.dk)

<sup>d</sup> University of Western Australia, School of Molecular Science, Perth, WA, Australia; [dylan.jayatilaka@uwa.edu.au](mailto:dylan.jayatilaka@uwa.edu.au)

## Contents

1	Introduction.....	5
2.	Calculation of form factors and structure factors from wavefunctions .....	7
2.1	Electron densities and structure factors in traditional crystallography .....	7
2.2	Electron densities and structure factors in quantum mechanics.....	9
2.3	Atomic electron densities in quantum mechanics .....	14
2.4	Dealing directly with the two-center quantum mechanical electron densities.....	17
2.5	Fourier transforms of basis functions products .....	21
2.5.1	The McMurchie-Davidson approach.....	24
2.5.2	The Obara-Saika approach .....	26
3.	Introduction to Hirshfeld Atom Refinement.....	28
3.1	Ideas behind the technique of Hirshfeld Atom Refinement.....	28
3.2	The crystal asymmetric unit and its environment .....	32
3.3	A minimal HAR and HART.....	34
3.4	Periodic HAR .....	35
3.5	Relativistic HAR .....	36
3.6	HAR-ELMO.....	38
3.7	The lamaGOET interface for quantum crystallography.....	40
3.8	NoSpherA2 in Olex2.....	43
3.9	Multi-determinant HAR .....	44
3.10	HAR and powder diffraction.....	46
4.	Introduction to X-ray constrained wavefunction fitting .....	48
4.1	Basic assumptions of the XCW fitting technique .....	49
4.2	Different “flavors” of XCW fitting strategies .....	52
4.3	Current implementations of the different XCW approaches.....	61
4.4	Meaning of the differences between fitted and non-fitted wavefunctions .....	62
4.5	Open problems and future perspectives .....	65
5.	Chemical applications of quantum crystallography.....	66
5.1	Applications of HAR.....	68
5.1.1	Hydrogen atom treatment in HAR: accuracy and precision compared to results from neutron diffraction .....	68
5.1.2	Limitations of HAR.....	74
5.1.3	Some other recent applications of HAR.....	75
5.2	Opto-electronic properties from XCW fitting: polarizabilities, hyperpolarizabilities and refractive indices .....	79
5.3	Insights into chemical bonding from X-ray wavefunction refinement .....	80
5.4	Properties from X-ray constrained wavefunctions: current developments and future perspectives .....	88
	Acknowledgments.....	89
	References.....	90

## Abbreviations

ADPs	Anisotropic displacement parameters
AIXRD	Ab initio X-ray diffraction
AVID	Aarhus vacuum imaging-plate diffractometer
BCA	biscarbonyl[14]annulene
CC-CD	charge-concentrated-to-charge-depleted
CCSD	Coupled Cluster with single and double excitations
CE	CrystalExplorer
CIF	Crystallographic information file
CSD	Cambridge structural database
csu's	Combined standard uncertainties
DED	{4-bis(diethylamino)-methyliumphenyl} dicyanomethanide
DFT	Density functional theory
DKH2	Second-order Douglas-Kroll-Hess
ED	Electron density
ELF	Electron localization function
ELI	Electron localizability indicator
ELMOs	Extremely localized molecular orbitals
ELWFs	Extremely localized Wannier functions
GUI	Graphical user interface
HAR	Hirshfeld atom refinement
HARt	Hirshfeld atom refinement terminal
HCT	hydrochlorothiazide

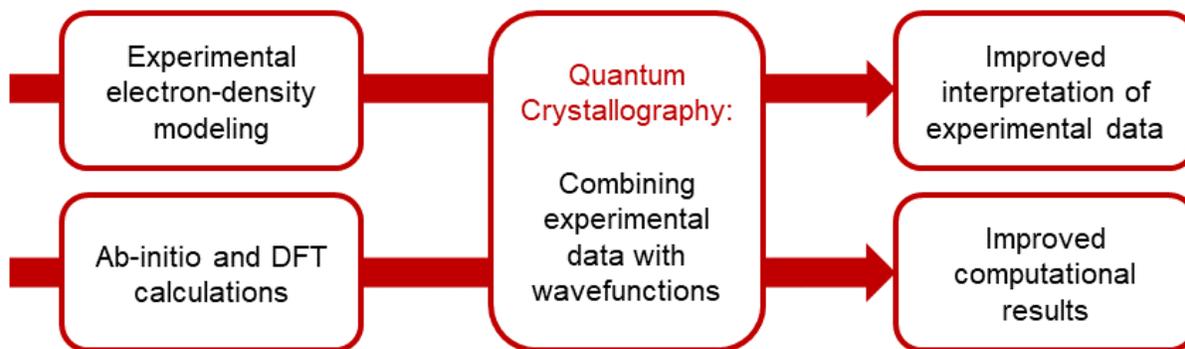
HF	Hartree-Fock
HRR	Horizontal recurrence relation
IAM	Independent atom model
IOTC	Infinite-order two-component
LOL	Localized orbital locator
MBADNP	3,5-dinitro-2-[1-phenyl-ethyl]-aminopyridine
MK	Michael & Koritsanszky
MLWFs	Maximally localized Wannier functions
MM	Multipole model
MOON	Molecular orbital occupation number
MP2	Second-order Møller-Plesset
NBOs	Natural bond orbitals
NRT	Natural resonance theory
PDF	Probability distribution function
QCT	Quantum chemical topology
QTAIM	Quantum theory of atoms in molecules
RGBIs	Roby-Gould bond indices
RHF	Restricted Hartree-Fock
RI	Resolution of identity
SDs	Standard deviations
SHADE	Simple hydrogen anisotropic displacement estimator
SPAnH	9-diphenylthiophosphoranyl anthracene
SPAnPS	9,10-bis-diphenyl-thiophosphoranyl anthracene·toluene
TAAM	Transferable aspherical atomic model

VRR	Vertical recurrence relation
XAO	X-ray atomic orbital
XC-ELMOs	X-ray constrained extremely localized molecular orbitals
XC-ELMO-VB	X-ray constrained extremely localized molecular orbital valence bond
XCSC	X-ray constrained spin-coupled
XCW	X-ray constrained wavefunction
XMO	X-ray molecular orbital
XWR	X-ray wavefunction refinement
ZTS	zinc [tris]thiourea sulphate

## 1 Introduction

Quantum crystallography is a field of crystallographic research that intimately combines theory with experiments (diffraction, spectroscopy, microscopy and others). It relates to phenomena in crystals that can only be explained or derived with the laws of quantum mechanics. This chapter deals primarily with the combination of diffraction data and quantum chemical calculations of various flavors. It shows, on the one hand, how quantum mechanical calculations can help in improving models of crystal structures from diffraction data, and, on the other hand, how experimental results can compensate for computational shortcomings (Fig. 1). Related techniques of multipole modeling for experimental electron-density analysis are discussed in a separate chapter of this book (Overgaard, details to be filled in by the editors).

The term quantum crystallography introduced in 1995 [1] and the methods related to it have seen a revival and reinvention in the last few years [2-7]. This development has led to the renaming of the International Union of Crystallography's "Commission of Charge, Spin and Momentum Densities" to "Commission of Quantum Crystallography" in 2017. A commonly accepted, strict definition of this field has not been given yet [3], but such a definition is probably neither necessary nor helpful for the development of the field.



**Fig. 1.** Quantum crystallography as a bridge between experiment and theory to obtain reliable chemical properties.

Following the Introduction (Section 1), the chapter is divided into four sections, which can be read independently of each other.

Section 2 is a thorough review of the theoretical foundations for calculating atomic form factors and structure factors from quantum mechanical wavefunctions. The emphasis is on the background behind the methods currently developed in our research groups, namely Hirshfeld atom refinement (HAR) and X-ray constrained wavefunction (XCW) fitting. Since atomic form factors are *always* calculated theoretically in crystallography, even for the traditional Independent Atom Model, this is also an introduction to the theory that underlies any routine structure determination or multipole modeling of the electron density.

Sections 3 and 4 deal with the conceptual bases, some details, the different implementations and the different flavors of HAR and XCW. Section 3 is dedicated to Hirshfeld Atom Refinement, while section 4 focuses on X-ray constrained wavefunction fitting. In both sections, the amalgamation of HAR and XCW fitting into the X-ray wavefunction refinement (XWR) approach is also briefly addressed.

Section 5 illustrates the usefulness of HAR, XCW fitting and XWR for answering chemical questions. It summarizes results (some of them published, some of them not) obtained by us and others.

## 2. Calculation of form factors and structure factors from wavefunctions

### 2.1 Electron densities and structure factors in traditional crystallography

It is known that in the kinematic scattering theory, which is based on the first Born approximation, the intensity of diffracted X-ray beams depends on the square of the magnitude of the structure factor  $|F_{\mathbf{h}}|^2$ , where the structure factor is the Fourier transform of the time and space averaged electron density in the unit cell ( $\bar{\rho}_{cell}$ ), namely:

$$F_{\mathbf{h}} = F[\bar{\rho}_{cell}] = \int d\mathbf{r} \bar{\rho}_{cell}(\mathbf{r}) e^{i2\pi(\mathbf{B}\mathbf{h})\cdot\mathbf{r}} = \int d\mathbf{r} \bar{\rho}_{cell}(\mathbf{r}) e^{i\mathbf{k}\cdot\mathbf{r}} = F_{\mathbf{k}} \quad (1)$$

Here  $F[\dots]$  represents a Fourier transform,  $\mathbf{k} = \mathbf{k}_o - \mathbf{k}_i$  is the X-ray scattering wavevector (i.e., the difference between the outgoing and incoming X-ray wavevectors), which is related to the triad of the Miller indices  $\mathbf{h}$  through the following relation:

$$\mathbf{k} = 2\pi\mathbf{B}\mathbf{h} \quad (2),$$

where  $\mathbf{B}$  is the reciprocal lattice matrix, namely the matrix whose columns are the reciprocal lattice vectors. Due to this relationship between vectors  $\mathbf{h}$  and  $\mathbf{k}$ , the structure factor  $F_{\mathbf{h}}$  may be labeled also as  $F_{\mathbf{k}}$ . For the sake of completeness and clarity, it is also worth pointing out that throughout the whole chapter, all the vectors in the direct space refer to an orthogonal Cartesian reference frame.

It is because of the relationship expressed by Eq. 1 that one often says loosely that *the X-ray experiment determines the electron density in the crystal*. Now, in crystal structure determination, it is common to write the static (i.e., unaveraged) unit-cell electron density  $\rho_{cell}$  as a sum of the different atomic contributions in the unit-cell:

$$\rho_{cell}(\mathbf{r}) = \sum_{A=1}^{N_{atoms}} \rho_A(\mathbf{r}) \quad (3),$$

where  $N_{atoms}$  is the number of atoms in the unit-cell and  $\rho_A(\mathbf{r})$  is the electron density of the generic atom  $A$  centered at position  $\mathbf{A} = (A_x, A_y, A_z)$  in the unit-cell. This atomic expansion is convenient because the averaged electron density in the unit-cell may be easily calculated as a sum of averaged atomic electron densities  $\bar{\rho}_A$ :

$$\bar{\rho}_{cell}(\mathbf{r}) = \sum_{A=1}^{N_{atoms}} \bar{\rho}_A(\mathbf{r}) \quad (4)$$

where  $\bar{\rho}_A$  is a convolution of the static atomic electron density  $\rho_A$  with the probability distribution function  $P_A$  for that atom, namely

$$\bar{\rho}_A(\mathbf{r}) = \rho_A(\mathbf{r}) \star P_A(\mathbf{u}) = \int d\mathbf{u} \rho_A(\mathbf{r} - \mathbf{u}) P_A(\mathbf{u}) \quad (5)$$

with  $\mathbf{u}$  as the displacement vector of atom  $A$  with respect to its equilibrium position  $\mathbf{A}$ .

The probability distribution function  $P_A$  is normally taken to be a three-dimensional Gaussian function, which is expressed in terms of a symmetric  $3 \times 3$  anisotropic displacement parameter (ADP) tensor  $\mathbf{U}_A$  as

$$P_A(\mathbf{u}) = (2\pi)^{-3/2} (\det(\mathbf{U}_A))^{-1/2} e^{-\frac{\mathbf{u}^T \mathbf{U}_A^{-1} \mathbf{u}}{2}} \quad (6)$$

All of this is essentially the Einstein model for independent atomic motion in a crystal [8], with the probability distribution function  $P_A$  that averages the static atomic electron density.

It is now straightforward to obtain the structure factor. It is simply given by the Fourier transform of Eq. 5, which, by exploiting the Fourier convolution theorem, is given by:

$$\bar{f}_A(\mathbf{k}) = F[\bar{\rho}_A(\mathbf{r})] = F[\rho_A(\mathbf{r}) \star P_A(\mathbf{u})] = T_A(\mathbf{k}) f_A(\mathbf{k}) \quad (7),$$

where  $T_A(\mathbf{k})$  is the temperature factor, namely the Fourier transform of the probability distribution function

$$T_A(\mathbf{k}) = \int d\mathbf{u} P_A(\mathbf{u}) e^{i\mathbf{k} \cdot \mathbf{u}} = e^{-\frac{\mathbf{k}^T \mathbf{U}_A \mathbf{k}}{2}} \quad (8),$$

and  $f_A(\mathbf{k})$  is the Fourier transform of the static electron density:

$$f_A(\mathbf{k}) = \int d\mathbf{r} \rho_A(\mathbf{r}) e^{i\mathbf{k} \cdot \mathbf{r}} \quad (9)$$

In most applications, the static atomic electron density  $\rho_A(\mathbf{r})$  is obtained from rather sophisticated quantum mechanical calculations, and then spherically averaged. The Fourier transforms of these spherical  $\rho_A$ 's are called "atomic form factors" and they only depend on the magnitude  $k$  of the scattering wavevector, not on its direction in  $k$ -space. Hence, tabulations are

easily constructed, and these are available in the International Tables of Crystallography; these quantum mechanically obtained atomic form factors are an essential ingredient in performing a standard X-ray structure refinement, where the atomic parameters (i.e., atomic positions  $\mathbf{A}$  and atomic displacement parameters  $\mathbf{U}_A$ ) are obtained by a least-squares fit against the experimentally measured  $|F_h|^2$ . This refinement procedure is called the independent atom model (IAM) and it is fundamental to X-ray structure refinement.

Given that *X-ray structure refinement depends so critically on theoretical atomic wavefunctions*, it is hard to really call the method purely “experimental” in nature. In any case, this refinement procedure works well, except for hydrogen atoms, whose X-ray refined positions are invariably found to be a little closer to the atoms they are bonded to than it might be expected when compared to definitive neutron diffraction or other measurements. This is because the electron density tends to be displaced somewhat toward the atom to which it is bonded.

If the X-ray data are very accurate, more advanced models are used, e.g. the multipole models, where the atomic density  $\rho_A(\mathbf{r})$  is assumed to be aspherical and parameters describing this asphericity are also refined to the X-ray data in addition to the positional and ADP parameters [9-13].

## ***2.2 Electron densities and structure factors in quantum mechanics***

In contrast to the atom-centered view of crystallography, the electron density in quantum mechanics derives from the square of a global wavefunction  $\Psi$ :

$$\rho(\mathbf{r}) = N \sum_{\xi=\alpha,\beta} \int d\mathbf{x}_2 \dots d\mathbf{x}_N |\Psi(\mathbf{r}, \xi, \mathbf{x}_2, \dots, \mathbf{x}_N)|^2 \quad (10)$$

Here, the squared modulus of  $\Psi$  is integrated over all the space-and-spin electron coordinates  $\mathbf{x}_i$ , except for the first one, and summed over the spin coordinate of the first electron. The result  $\rho(\mathbf{r})$  is the electron density of any one given electron, but because all electrons are identical, the total electron density is just  $N$  times the electron density of the first, where  $N$  is the number of electrons described by the wavefunction  $\Psi$ .

$\Psi$  itself is usually represented as a series expansion of atom-centered basis functions. Since the square of the wavefunction appears in Eq.10, it follows that the electron density is usually

expressed in terms of *one-* and *two-center* products of basis functions, as it will be shown in more details below.

In the simplest case of a wavefunction  $\Psi$  corresponding to a restricted Hartree-Fock (HF) wavefunction for a closed-shell system, the electron density can be written as

$$\rho(\mathbf{r}) = \sum_{i=1}^{N_{occ}} 2 |\phi_i(\mathbf{r})|^2 \quad (11)$$

where  $N_{occ}$  is the number of occupied Hartree-Fock molecular orbitals. For restricted open-shell wavefunctions with unpaired electrons, some of the occupation numbers in the equation above are equal to one (for each of the singly occupied molecular orbitals). Likewise, in the unrestricted approximation where the shapes of the spatial parts of the alpha and beta molecular orbitals are different, the occupation numbers are all equal to one. In the most general case, which also applies to post-Hartree-Fock correlated wavefunctions, the electron density can be expressed in terms of natural orbitals, which are the eigenvectors of the one-electron density matrix for the system under exam [14], namely  $\rho(\mathbf{r})$  is given by:

$$\rho(\mathbf{r}) = \sum_{i=1}^{N_{bf}} n_i |\phi_i(\mathbf{r})|^2 \quad (12)$$

with  $\phi_i$  as a generic spatial natural orbital depending on the type of wavefunction  $\Psi$  describing the system,  $n_i$  as its occupation number (always respecting the condition  $0 \leq n_i \leq 2$ ), and  $N_{bf}$  as the total number of spatial basis functions used in the calculation, namely the number of spatial basis functions used to approximate the natural orbitals and the wavefunction [14].

It should also be noted that the number of electrons in a real crystal is very large, and, while the number of electrons in the unit-cell of an infinite periodic crystal is finite, the wavefunction  $\Psi$  used to model the electron density in the unit-cell may not correspond to the unit-cell, or even to the symmetry-unique part of the unit-cell (i.e., the so-called asymmetric unit). Instead *the chosen wavefunction  $\Psi$  usually describes a fragment of the crystal that best models the asymmetric unit.* This implies that the wavefunction used should incorporate some of the effects of the environment on the asymmetric unit. In contrast, wavefunctions that are fully periodic *do*

incorporate all non-crystal-surface environmental effects, perhaps at the expense of limiting the accuracy at which electron correlation effects can be modeled.

In this chapter, we particularly focus on those modern methods of quantum crystallography that we are developing in our own research groups [Hirshfeld Atom Refinement [15-20] (HAR; see section 3) and X-ray constrained wavefunction fitting [21-27] (XCW; see section 4)]. In these methods, we generally consider the crystal as a collection of molecular units. For this reason, by exploiting the set  $\{\mathbf{Q}_j, \mathbf{q}_j\}_{j=1}^{N_m}$  of  $N_m$  unit-cell symmetry operations, the unit-cell electron density can be expressed only in terms of the averaged electron density  $\bar{\rho}_0$  of a reference molecular fragment in the unit-cell, namely

$$\bar{\rho}_{cell}(\mathbf{r}) = \sum_{j=1}^{N_m} \bar{\rho}_j(\mathbf{r}) = \sum_{j=1}^{N_m} \bar{\rho}_0(\mathbf{Q}_j^{-1}(\mathbf{r} - \mathbf{q}_j)) \quad (13),$$

The reference fragment is generally chosen as a contiguous asymmetric unit of the unit-cell. However, as we will see below, it can also be chosen corresponding to a larger subunit, for example when it is necessary to represent the crystal as a repetition of fragments consisting of a cluster of molecules because the intermolecular interactions in the cluster might cause significant changes in the asymmetric unit electron density. The treatment of symmetry for fragments of a crystal larger than the asymmetric unit is considered in more detail in section 2.4.

If we substitute Eq. 13 into Eq. 1 and if we exploit the fact that, for rotation matrices, their inverse corresponds to their transpose, we obtain:

$$F_{\mathbf{h}}^{calc} = \sum_{j=1}^{N_m} \int d\mathbf{r} \bar{\rho}_0(\mathbf{Q}_j^T(\mathbf{r} - \mathbf{q}_j)) e^{i2\pi(\mathbf{B}\mathbf{h})\cdot\mathbf{r}} \quad (14)$$

Introducing the change of coordinates

$$\mathbf{r}' = \mathbf{Q}_j^T(\mathbf{r} - \mathbf{q}_j) \quad (15),$$

implying also that

$$\mathbf{r} = \mathbf{Q}_j \mathbf{r}' + \mathbf{q}_j \quad (16),$$

Eq. 14 simply becomes

$$F_{\mathbf{h}}^{calc} = \sum_{j=1}^{N_m} e^{i2\pi \mathbf{q}_j \cdot (\mathbf{B}\mathbf{h})} \int d\mathbf{r}' \bar{\rho}_0(\mathbf{r}') e^{i2\pi (\mathbf{Q}_j^T \mathbf{B}\mathbf{h}) \cdot \mathbf{r}'} \quad (17),$$

where we have also exploited the following relation:

$$(\mathbf{B}\mathbf{h}) \cdot (\mathbf{Q}_j \mathbf{r}') = (\mathbf{Q}_j^T \mathbf{B}\mathbf{h}) \cdot \mathbf{r}' \quad (18).$$

For the sake of simplicity, in the remaining part of the text we will neglect the prime symbol for the position vector (namely, we will always consider  $\mathbf{r}' \equiv \mathbf{r}$ ), but it will be necessary to bear in mind that the new position vector  $\mathbf{r}$  is actually connected to the old one (namely, the one appearing in Eq. 14) through the change of coordinates described by Eqs. 15 and 16.

In Eq. 17, the electron density  $\bar{\rho}_0(\mathbf{r})$  is the averaged electron density of the chosen reference molecular fragment in the unit-cell. In other words, it is the convolution of the static electron density  $\rho_0(\mathbf{r})$  of the reference molecular fragment with a probability distribution function that takes into account the probabilities with which the different nuclei of the reference fragment move around their equilibrium positions. The static electron density is obtained by partial integration of the squared modulus of wavefunction  $\Psi_0$  chosen to describe the reference fragment in the unit-cell, as done in Eq. 10.

For practical quantum chemical calculations of (static) electron densities, the natural or molecular orbitals are usually represented as linear combinations of *basis functions*  $\{\chi_\alpha(\mathbf{r})\}_{\alpha=1}^{N_{bf}}$  centered on atomic nuclei, namely

$$\phi_i(\mathbf{r}) = \sum_{\alpha=1}^{N_{bf}} C_{\alpha i} \chi_\alpha(\mathbf{r}) \quad (19)$$

where the coefficients  $\{C_{\alpha i}\}$  are the natural or molecular orbital expansion coefficients.

Therefore, inserting Eq. 19 into Eq. 12, the static electron density of the reference fragment can be written like this:

$$\rho_0(\mathbf{r}) = \sum_{\alpha, \beta=1}^{N_{bf}} \chi_\alpha(\mathbf{r}) D_{\alpha\beta}^0 \chi_\beta(\mathbf{r}) \quad (20)$$

where  $D_{\alpha\beta}^0$ , which is the element  $(\alpha,\beta)$  of the one-particle density matrix  $\mathbf{D}^0$  of the reference molecular fragment, has the following expression:

$$D_{\alpha\beta}^0 = \sum_{i=1}^{N_{bf}} n_i C_{\alpha i} C_{\beta i} \quad (21)$$

The basis functions  $\chi_\alpha$  and  $\chi_\beta$  in Eq. 20 may be centered on the same or on different nuclear positions. Therefore, it follows that (static) electron densities obtained from quantum mechanical calculations on polyatomic systems generally contain two-center terms, as stated in the outset of this subsection.

Due to the *two-center form of the static electron density*, the averaged electron density for the reference fragment can be written as the convolution of the static electron density  $\rho_0$  given by Eq. 20 with a two-center probability distribution function  $P_{AB}(\mathbf{u}_A, \mathbf{u}_B)$  for the nuclei  $A$  and  $B$  on which basis functions  $\chi_\alpha$  and  $\chi_\beta$  are centered (with  $\mathbf{u}_A$  and  $\mathbf{u}_B$  as displacement vectors for nuclei  $A$  and  $B$ , respectively):

$$\bar{\rho}_0(\mathbf{r}) = \int d\mathbf{u}_A d\mathbf{u}_B \left[ \sum_{\alpha,\beta=1}^{N_{bf}} \chi_\alpha(\mathbf{r} - \mathbf{u}_A) D_{\alpha\beta}^0 \chi_\beta(\mathbf{r} - \mathbf{u}_B) \right] P_{AB}(\mathbf{u}_A, \mathbf{u}_B) \quad (22)$$

By analogy to  $P_A(\mathbf{u})$  (see Eq. 6),  $P_{AB}(\mathbf{u}_A, \mathbf{u}_B)$  can be taken as a six-dimensional Gaussian function, expressed in terms of a symmetric  $6 \times 6$  two-center ADP tensor  $\mathbf{U}$ .

$$\mathbf{U}(A, B) = \begin{bmatrix} \mathbf{U}_A & \mathbf{U}_{AB} \\ \mathbf{U}_{BA} & \mathbf{U}_B \end{bmatrix} = \begin{bmatrix} \mathbf{U}_A & \mathbf{U}_{AB} \\ \mathbf{U}_{AB}^T & \mathbf{U}_B \end{bmatrix} \quad (23)$$

This tensor would contain 21 independent parameters for each pair of atoms, 9 more than usual. Moreover, we would need to ensure that the diagonal, three dimensional one-atom marginals  $P_A(\mathbf{u}_A)$  (and consequently  $\mathbf{U}_A$ ) for all the six dimensional  $P_{AB}(\mathbf{u}_A, \mathbf{u}_B)$  referring to the same atom  $A$  are the same (the marginal of a probability distribution function (PDF) is the probability distribution function obtained by integrating over some or all the variables in the original PDF).

Nevertheless, the off-diagonal terms  $\mathbf{U}_{AB}$  in the two-center probability distribution functions  $P_{AB}(\mathbf{u}_A, \mathbf{u}_B)$  are unfortunately not available. To overcome this drawback, two different approaches are usually adopted. They will be outlined in the next two subsections.

### 2.3 Atomic electron densities in quantum mechanics

The first strategy to overcome the non-availability of the non-diagonal terms  $U_{AB}$  in Eq. 23 (see the end of the previous subsection) is to split the total two-center expansion of the electron density for the chosen reference fragment into atomic densities  $\rho_A(\mathbf{r})$ . Within this first strategy there are also several possibilities. In our work we have used the Hirshfeld atomic partitioning scheme [28, 29] to obtain static atomic electron densities (see Hirshfeld Atom Refinement discussed in section 3). Hirshfeld defined an atomic weight function  $w_A$  to “mask out” an atomic density from a given arbitrary electron density as follows:

$$w_A(\mathbf{r}) = \frac{\rho_A^o(\mathbf{r})}{\sum_{B=1}^{N_{atoms}} \rho_B^o(\mathbf{r})} \quad (24)$$

Hirshfeld’s weight function is a ratio of the spherical atomic density for the atom under consideration ( $\rho_A^o$ ) divided by the “promolecule” or “procrystal” density [30], which is the sum of the spherical atomic densities in the molecule or crystal. Hirshfeld’s static atomic electron density is then given by

$$\rho_A(\mathbf{r}) = w_A(\mathbf{r}) \rho_0(\mathbf{r}) \quad (25)$$

where  $\rho_0(\mathbf{r})$  is again the electron density of the reference molecular unit (see Eq. 20).

Hirshfeld’s formula works from the idea that each atom “owns” a part of the total electron density  $\rho_0(\mathbf{r})$  according to how much that atom’s spherical electron density  $\rho_A^o(\mathbf{r})$  contributes to the promolecule density, a kind of “stockholder” partitioning. The spherical atomic densities  $\rho_A^o(\mathbf{r})$  may be obtained from tabulations of the radial atomic densities or calculated from tabulations of the Clementi-Roetti [31] or Koga *et al.* [32] atomic wavefunctions. However, we generally use spherically averaged atomic electron densities resulting from atomic calculations performed at the same level of theory used to obtain the system wavefunction and its density  $\rho_0(\mathbf{r})$ . These atomic wavefunctions are normally calculated and stored once, at the start of any X-ray structure refinement. These spherically averaged atomic electron densities are afterwards used to obtain the aspherical Hirshfeld atomic densities  $\rho_A(\mathbf{r})$  (see Eq. 25), which are finally exploited to compute structure factors (see Eq. 9) by numerical integration using standard Becke grid techniques widely used in quantum chemistry [33].

An alternative used to obtain atomic electron densities from quantum mechanical wavefunctions is simply to fit the molecular or crystal density to an atom-centered expansion. In quantum chemistry this is a standard technique, which goes by the name of “resolution of identity” (RI) method [34]. Special RI “atomic fitting” basis sets are even available and optimized to work with “normal” basis sets [34, 35]. This technique has not been used yet in quantum crystallography, but it may be much quicker to use such an approach, since it does not involve the numerical integrations required when using the Hirshfeld atomic densities described above. Yet another method that has become quite popular is to represent aspherical atomic densities with a multipole expansion. For example, in the transferable aspherical atomic model (TAAM), aspherical atomic form factors are obtained by fitting to theoretical static structure factors generated from the electron density of the unit-cell resulting from the wavefunction of the whole system [36, 37].

In the framework of the strategy discussed in this section, it is important to bear in mind that distribution of atoms in a crystal also possesses space-group symmetry that needs to be taken into account. For this purpose, we assume that a particular atom  $A$  of the asymmetric unit is obtained  $N_m$  times under the application of the  $N_m$  symmetry operations  $\{\mathbf{Q}_j, \mathbf{q}_j\}_{j=1}^{N_m}$  of the unit-cell, as atoms  $A_1, A_2, \dots, A_{N_m}$ . Since the static electron density of each atom  $A_k$  is  $\rho_{A_k}(\mathbf{r} - \mathbf{A}_k)$  and it is located at position  $\mathbf{A}_k$ , the static electron density in the unit-cell is:

$$\rho_{cell}(\mathbf{r}) = \sum_{A=1}^{N_a} s_A^{-1} \left( \sum_{k=1}^{N_m} \rho_{A_k}(\mathbf{r} - \mathbf{A}_k) \right) \quad (26)$$

where it is important to note that  $N_a$  is the number of atoms for the chosen asymmetric unit and, consequently, unlike Eq. 3, here the sum runs only over the atoms of the asymmetric unit.

The aspherical atomic densities  $\rho_{A_k}$  are assumed to be centered at the origin. The factor  $s_A$  is the site symmetry factor, i.e. the number of times an atom  $A$  is mapped into itself by the crystal symmetry operations, and, therefore, the factor  $s_A^{-1}$  avoids double counting of electron density contributions. Now, let us suppose that the crystal symmetry operation is represented by a rotation matrix  $\mathbf{Q}_k$  and a translation or “centering” vector  $\mathbf{q}_k$ . Then, by symmetry, the electron density of atom  $A$  at a generic point  $\mathbf{r}$  in the direct space will be equal to the electron density of atom  $A_K$  at a symmetry-rotated and translated point  $\mathbf{Q}_k \mathbf{r} + \mathbf{q}_k$ , i.e.

$$\rho_A(\mathbf{r} - \mathbf{A}) = \rho_{A_k}(\mathbf{Q}_k \mathbf{r} + \mathbf{q}_k - \mathbf{A}_k) \quad (27)$$

where, always applying the rotation matrix  $\mathbf{Q}_k$  and the translation vector  $\mathbf{q}_k$ , we have

$$\mathbf{A}_k = \mathbf{Q}_k \mathbf{A} + \mathbf{q}_k \quad (28),$$

which is the position of atom  $A_k$ , namely the symmetry-rotated and translated position of atom  $A$ .

Now we introduce the following change of coordinates:

$$\mathbf{t} = \mathbf{Q}_k \mathbf{r} + \mathbf{q}_k \quad (29)$$

$$\mathbf{r} = \mathbf{Q}_k^T (\mathbf{t} - \mathbf{q}_k) \quad (30),$$

where in the last equation we have exploited again the fact that, for rotation matrices, their inverse corresponds to their transpose.

Introducing Eqs. 29 and 30 into Eq. 27, we obtain:

$$\rho_A(\mathbf{Q}_k^T (\mathbf{t} - \mathbf{q}_k) - \mathbf{A}) = \rho_{A_k}(\mathbf{t} - \mathbf{A}_k) \quad (31)$$

If in Eq. 26 we rename variable  $\mathbf{r}$  as  $\mathbf{t}$  and we exploit Eq. 31, we have:

$$\rho_{cell}(\mathbf{t}) = \sum_{A=1}^{N_a} S_A^{-1} \left( \sum_{k=1}^{N_m} \rho_{A_k}(\mathbf{t} - \mathbf{A}_k) \right) = \sum_{A=1}^{N_a} S_A^{-1} \left( \sum_{k=1}^{N_m} \rho_A(\mathbf{Q}_k^T (\mathbf{t} - \mathbf{q}_k) - \mathbf{A}) \right) \quad (32),$$

with, again,  $N_a$  as the number of atoms for the chosen asymmetric unit and  $N_m$  as the number of unit-cell symmetry operations. In other words, in this way we were able to express the unit-cell electron density only in terms of the atomic electron densities associated with the symmetry-unique atoms in the asymmetric unit.

Now we can calculate the structure factors for the static unit-cell electron density from this expression and, after some simple algebra, the result is

$$F_k^{static} = \sum_{A=1}^{N_a} S_A^{-1} \sum_{k=1}^{N_m} e^{i(\mathbf{Q}_k^T \mathbf{k}) \cdot \mathbf{A}} e^{i\mathbf{k} \cdot \mathbf{q}_k} \int d\mathbf{v} \rho_A(\mathbf{v}) e^{i(\mathbf{Q}_k^T \mathbf{k}) \cdot \mathbf{v}} \quad (33)$$

The last integral is the static structure factor evaluated at a symmetry-rotated scattering vector. If one wants the dynamic structure factor, the Einstein convolution trick allows us to immediately write:

$$F_{\mathbf{k}} = \sum_{A=1}^{N_a} s_A^{-1} \sum_{k=1}^{N_m} e^{i(\mathbf{Q}_k^T \mathbf{k}) \cdot \mathbf{A}} e^{i\mathbf{k} \cdot \mathbf{q}_k} \bar{f}_A(\mathbf{Q}_k^T \mathbf{k}) \quad (34)$$

where the aspherical static atomic structure factor has been replaced by the corresponding dynamic atomic structure factor. It is worth noting the similarity of this result with Eq. 17, which is defined with respect to a unique molecular unit rather than a set of unique atoms in the asymmetric unit. In both cases one requires the Fourier transforms at a set of symmetry-transformed scattering vectors to be evaluated. The reason for this can be traced to the fact that such transformations are required to generate the electron density in the unit-cell.

#### 2.4 Dealing directly with the two-center quantum mechanical electron densities

The second strategy to overcome the problem associated with the presence of the two-center probability distribution function in Eq. 22 consists in approximating  $P_{AB}(\mathbf{u}_A, \mathbf{u}_B)$  by using the parameters which describe the one-atom probability distribution functions, namely, the usual Anisotropic Displacement Parameters (ADPs).

In fact, introducing Eq. 22 into Eq. 17 and exploiting again the Fourier convolution theorem, it is possible to see that the expression for the structure factor becomes:

$$F_{\mathbf{h}} = \sum_{\alpha, \beta=1}^{N_{bf}} D_{\alpha\beta}^0 I_{\beta\alpha}^{\mathbf{h}} = \text{Tr}[\mathbf{D}^0 \mathbf{I}^{\mathbf{h}}] \quad (35)$$

with  $\mathbf{I}^{\mathbf{h}}$  as the matrix of the thermally averaged Fourier-transform integrals of the basis function products, namely the matrix having elements

$$I_{\alpha\beta}^{\mathbf{h}} = \sum_{j=1}^{N_m} e^{i2\pi \mathbf{q}_j \cdot (\mathbf{B}\mathbf{h})} T_{A\alpha B\beta}(\mathbf{Q}_j^T \mathbf{B}\mathbf{h}) \int d\mathbf{r} \chi_{\alpha}(\mathbf{r}) \chi_{\beta}(\mathbf{r}) e^{i2\pi(\mathbf{Q}_j^T \mathbf{B}\mathbf{h}) \cdot \mathbf{r}} \quad (36)$$

where  $T_{A\alpha B\beta}(\mathbf{Q}_j^T \mathbf{B}\mathbf{h})$  is formally the Fourier transform of the two-center probability distribution function for the nuclei  $A$  and  $B$  on which basis functions  $\chi_{\alpha}$  and  $\chi_{\beta}$  are centered. For the

moment, for this factor, we keep indicated the dependence on the indices  $\alpha$  and  $\beta$  characterizing the basis functions involved in the Fourier integral of Eq. 36. Just below, we will see that, for some methods approximating the factor  $T$ , this dependence can be neglected.

The thermal factor  $T_{A\alpha B\beta}(\mathbf{Q}_j^T \mathbf{B} \mathbf{h})$  can be approximated using four different approaches of increasing sophistication, namely those proposed by Coppens [38], Stewart [39], Tanaka [40] and Michael & Koritsanszky [41]. Following Coppens, the two-index thermal factor is simply the average of the thermal factors for the nuclei  $A$  and  $B$  on which basis functions  $\chi_\alpha$  and  $\chi_\beta$  are centered. In this case  $T_{A\alpha B\beta}$  does not depend on the particular types of basis functions centered on the two nuclei and, for this reason,  $T_{A\alpha B\beta}$  can be simply expressed as  $T_{AB}$  and, in particular, it assumes this form:

$$T_{AB}(\mathbf{h}') = \frac{\exp[-2\pi^2 \mathbf{h}' \cdot \mathbf{U}_A \mathbf{h}'] + \exp[-2\pi^2 \mathbf{h}' \cdot \mathbf{U}_B \mathbf{h}']}{2} \quad (37)$$

where  $\mathbf{h}'$  corresponds to  $\mathbf{Q}_j^T \mathbf{B} \mathbf{h}$  in Eq. 36 and  $\mathbf{U}_A$  and  $\mathbf{U}_B$  are the matrices of the ADPs for nuclei  $A$  and  $B$ , respectively. According to the Stewart model, the thermal factor has this form:

$$T_{AB}(\mathbf{h}') = \exp[-2\pi^2 \tau_{AB} \mathbf{h}' \cdot (\mathbf{U}_A + \mathbf{U}_B) \mathbf{h}'] \quad (38)$$

with  $\tau_{AB}$  equal to 0.5 if the motion of the nuclei  $A$  and  $B$  is correlated (distance between the nuclei lower than 2.5 bohr) or equal to 0.25 if the motion of the two nuclei is uncorrelated (distance between the nuclei greater than or equal to 2.5 bohr).  $\mathbf{U}_A$  and  $\mathbf{U}_B$  have the same meaning seen for Eq. 37. Also, in this case,  $T_{AB}$  does not depend on the types of basis functions on nuclei  $A$  and  $B$ .

Unlike the previous two strategies, in the remaining two approaches, the thermal factor also depends on the basis functions involved in the Fourier transform integral of Eq. 36 and, for this reason,  $T_{A\alpha B\beta}$  cannot be simplified. In particular for the Tanaka approach, we have:

$$T_{A\alpha B\beta}(\mathbf{h}') = \exp[-2\pi^2 \tau_{\alpha\beta} \mathbf{h}' \cdot (\tau_\alpha \mathbf{U}_A + \tau_\beta \mathbf{U}_B) \mathbf{h}'] \quad (39)$$

where  $\tau_\alpha$  and  $\tau_\beta$  are the exponents of the (primitive) Gaussian basis functions  $\chi_\alpha$  and  $\chi_\beta$  centered on nuclei  $A$  and  $B$ , respectively, and  $\tau_{\alpha\beta} = 1/(\tau_\alpha + \tau_\beta)$ , while  $\mathbf{U}_A$  and  $\mathbf{U}_B$  are again the ADPs for  $A$  and  $B$ .

Finally, in the more recent Michael & Koritsanszky (MK) strategy, the thermal factor is given by the following expression:

$$T_{A\alpha B\beta}(\mathbf{h}') = \exp[-2\pi^2 \tau_{\alpha\beta} \mathbf{h}' \cdot (\tau_{\alpha}^2 \mathbf{U}_{AA} + \tau_{\alpha}\tau_{\beta}(\mathbf{U}_{AB} + \mathbf{U}_{BA}) + \tau_{\beta}^2 \mathbf{U}_{BB}) \mathbf{h}'] \quad (40)$$

where  $\tau_{\alpha}$  and  $\tau_{\beta}$  are always the exponents of the (primitive) Gaussian basis functions  $\chi_{\alpha}$  and  $\chi_{\beta}$  centered on nuclei  $A$  and  $B$ ,  $\tau_{\alpha\beta} = 1/(\tau_{\alpha} + \tau_{\beta})^2$ . Furthermore, in this case,  $\mathbf{U}_{AA}$  and  $\mathbf{U}_{BB}$  are the diagonal blocks (associated with nuclei  $A$  and  $B$ , respectively) of the  $6 \times 6$  two-center displacement parameter tensor  $\mathbf{U}(A, B)$  (see Eq. 23), while  $\mathbf{U}_{AB}$  and  $\mathbf{U}_{BA}$  are the non-diagonal blocks of  $\mathbf{U}(A, B)$ . Because of the lack of the quantities  $\mathbf{U}_{AB}$  and  $\mathbf{U}_{BA}$ , the MK formula has never been tested in quantum crystallography refinements, but, from the theoretical point of view, it seems to be the best justified since, at least in principle, it also takes into account the correlation between the vibrations of nuclei  $A$  and  $B$  through the off-diagonal blocks of tensor  $\mathbf{U}(A, B)$ . When these blocks are unknown (as occurs in most of the situations), they are simply set to zero. Note that the Tanaka and MK methods work only with primitive Gaussian basis functions of quantum chemistry, and not with contracted ones (see subsection 2.5 for more details on the distinction between primitive and contracted Gaussian basis functions).

Now, as indicated above, we also consider how symmetry may be treated in Eqs. 22 and 35 when a two-center expansion of the electron density (see Eq. 20) is explicitly used and when the (chosen) reference molecular fragment does not correspond to the asymmetric unit of the unit-cell. As already explained, this is usually done when intermolecular interactions (e.g., hydrogen bonds) are not negligible and perturb the molecular electron density of the asymmetric unit. In this case the reference fragment chosen to model the electron density in the crystal is called *oversampled repeating unit*. Its electron density can be expressed like this

$$\tilde{\rho}_0(\mathbf{r}) = \sum_A \frac{\rho_A(\mathbf{r})}{n_A} \quad (41)$$

where, in this case, the densities  $\rho_A(\mathbf{r})$  are the atomic densities constituting the oversampled unit. It is this electron density that gives the structure factors when it is Fourier transformed. As we will explain below, in this case,  $\rho_A(\mathbf{r})$  is not the Hirshfeld atomic density for atom  $A$ .  $n_A$  is a factor that avoids the “double counting” of atoms in the unit-cell and it is defined as the number of times atom  $A$  is mapped into itself (or to a symmetry equivalent atom) by applying the unit-

cell symmetry operations. Note that  $n_A$  is not equal to the site symmetry factor  $s_A$  and that  $\sum_A \frac{1}{n_A} = N_a$ , with  $N_a$  as the number of atoms in the asymmetric unit (as seen in Eqs. 26, 32, 33 and 34). For example, if the oversampled unit was twice the asymmetric unit and there were no atoms on special positions, then  $n_A$  would be equal to two; anyway, in general, this factor depends on exactly which atoms are chosen for the oversampled fragment.

In order to obtain an atomic density  $\rho_A(\mathbf{r})$  from the two-center expansion of the electron density of the reference fragment, it is possible to use a basis function-based partitioning scheme, similar to the Tanaka or MK procedures used for the temperature factor. Thus, we have:

$$\rho_A(\mathbf{r}) = \sum_{\mu, \nu=1}^{N_{bf}} \chi_\mu(\mathbf{r}) D_{\mu\nu}^0 \chi_\nu(\mathbf{r}) f_{\mu\nu}^A \quad (42)$$

where  $D_{\mu\nu}^0$  is the density matrix coming from the wavefunction  $\Psi_0$  for the chosen oversampled fragment (see Eq. 20).

Comparing Eq. 42 to Eq. 20, it is easy to see that the factors  $f_{\mu\nu}^A$  partitions the density into atomic pieces. In other words,  $f_{\mu\nu}^A$  is the proportion of the true density  $\chi_\mu(\mathbf{r}) D_{\mu\nu}^0 \chi_\nu(\mathbf{r})$  assigned to atom  $A$ . Therefore, substituting Eq. 42 into Eq. 41, we obtain the density matrix associated with the electron density  $\tilde{\rho}_0(\mathbf{r})$  of the oversampled repeating unit, namely:

$$\tilde{D}_{\mu\nu}^0 = \sum_A n_A^{-1} D_{\mu\nu}^0 f_{\mu\nu}^A \quad (43)$$

The factor  $f_{\mu\nu}^A$  can be defined in different ways, but since it is always

$$\rho_0(\mathbf{r}) = \sum_A \rho_A(\mathbf{r}) \quad (44),$$

where the sum is over all the atoms in the oversampled fragment, we must have:

$$\sum_A f_{\mu\nu}^A = 1 \quad (45).$$

To satisfy the above equation, we might choose:

$$f_{\mu\nu}^A = \begin{cases} 1 & \text{if } (\chi_\mu \in A) \text{ and } (\chi_\nu \in A) \\ 1/2 & \text{if } ((\chi_\mu \in A) \text{ and } (\chi_\nu \notin A)) \text{ or } ((\chi_\mu \notin A) \text{ and } (\chi_\nu \in A)) \\ 0 & \text{if } (\chi_\mu \notin A) \text{ and } (\chi_\nu \notin A) \end{cases} \quad (46)$$

This is the Mulliken partitioning scheme, which is equivalent to partitioning the electron density associated with two basis functions on different centers equally between the pair of atoms concerned. More sophisticated partitioning schemes have also been proposed, as for example a Tanaka-like scheme based on the exponents of the primitive Gaussian basis functions [22].

Therefore, when we have an oversampled repeating unit, the structure factors can be expressed like this:

$$F_h = Tr[\tilde{\mathbf{D}}^0 \mathbf{I}^h] = Tr[\mathbf{D}^0 \tilde{\mathbf{I}}^h] \quad (47),$$

where in the second equality we have transferred the oversampling and partition factors into the Fourier transform integrals, namely:

$$\tilde{I}_{\mu\nu}^h = \sum_A n_A^{-1} I_{\mu\nu}^h f_{\mu\nu}^A \quad (48)$$

From Eqs. 43 and 48 it is clear that, to operatively exploit these equations, it is simply necessary: *i)* to determine the atom to which each basis function belongs; *ii)* to evaluate the oversampling factor  $n_A$  for each atom, which is obtained by establishing how many times the atom is mapped into itself under the unit-cell symmetry operations. If these factors are known, the modified density matrix  $\tilde{\mathbf{D}}^0$  or the matrix of the modified Fourier transform integrals  $\tilde{\mathbf{I}}^h$  used in Eq. 47 can be easily obtained from the unmodified matrices  $\mathbf{D}^0$  and  $\mathbf{I}^h$ . Compared to the Hirshfeld atom space-based partitioning, the essential difference in this method is that a basis function partitioning is used instead.

## 2.5 Fourier transforms of basis functions products

Now, let us reconsider Eq. 36 and, in particular let us focus on the Fourier transform integral of the product between basis functions  $\chi_\alpha$  and  $\chi_\beta$ :

$$I = \int d\mathbf{r} \chi_\alpha(\mathbf{r}; \mathbf{A}, L_\alpha) \chi_\beta(\mathbf{r}; \mathbf{B}, L_\beta) e^{i\mathbf{K}\cdot\mathbf{r}} \quad (49)$$

where, for the sake of simplicity, we have set

$$\mathbf{K} = 2\pi \mathbf{Q}_j^T \mathbf{B} \mathbf{h} \quad (50)$$

and where, for the involved basis functions, we have also explicitly indicated the nuclear positions on which they are centered ( $\mathbf{A} = (A_x, A_y, A_z)$  and  $\mathbf{B} = (B_x, B_y, B_z)$ ) and their total angular momentum ( $L_\alpha$  and  $L_\beta$ ). Note that vector  $\mathbf{K}$  defined in Eq. 50 is slightly different from vector  $\mathbf{k}$  in Eq. 1. They are identical only when matrix  $\mathbf{Q}_j$  is the identity matrix.

As is well known, basis functions used in quantum chemistry are generally *contracted Cartesian Gaussian basis functions*, namely, they are linear combinations of the so-called *Cartesian Gaussian primitive functions* centered on the same nuclei and having the same angular momentum. Therefore, basis functions  $\chi_\alpha$  and  $\chi_\beta$  in Eq. 49 can be expanded like this:

$$\chi_\alpha(\mathbf{r}; \mathbf{A}, L_\alpha) = \sum_{i=1}^{M_\alpha} N_{i\alpha} c_{i\alpha} g_{i\alpha}(\mathbf{r}; \mathbf{A}, L_\alpha) \quad (51)$$

and

$$\chi_\beta(\mathbf{r}; \mathbf{B}, L_\beta) = \sum_{j=1}^{M_\beta} N_{j\beta} c_{j\beta} g_{j\beta}(\mathbf{r}; \mathbf{B}, L_\beta) \quad (52)$$

with  $M_\alpha$  and  $M_\beta$  as the degrees of contraction,  $N_{i\alpha}$  and  $N_{j\beta}$  as the normalization constants for the primitive functions,  $c_{i\alpha}$  and  $c_{j\beta}$  as the contraction coefficients, and  $g_{i\alpha}(\mathbf{r}; \mathbf{A}, L_\alpha)$  and  $g_{j\beta}(\mathbf{r}; \mathbf{B}, L_\beta)$  as un-normalized Cartesian Gaussian primitive functions. Moreover, in Eqs. 51 and 52, both for the contracted basis functions  $\chi_\alpha$  and  $\chi_\beta$  and for the primitive functions  $g_{i\alpha}$  and  $g_{j\beta}$ , we have explicitly shown the nuclear positions  $\mathbf{A}$  and  $\mathbf{B}$  on which they are centered and their total angular momenta  $L_\alpha$  and  $L_\beta$ . Therefore, exploiting Eqs. 51 and 52, the integral given by Eq. 49 becomes:

$$\begin{aligned} I &= \sum_{i=1}^{M_\alpha} \sum_{j=1}^{M_\beta} N_{i\alpha} N_{j\beta} c_{i\alpha} c_{j\beta} \int d\mathbf{r} g_{i\alpha}(\mathbf{r}; \mathbf{A}, L_\alpha) g_{j\beta}(\mathbf{r}; \mathbf{B}, L_\beta) e^{i\mathbf{K}\cdot\mathbf{r}} = \\ &= \sum_{i=1}^{M_\alpha} \sum_{j=1}^{M_\beta} N_{i\alpha} N_{j\beta} c_{i\alpha} c_{j\beta} I_{i\alpha}^{j\beta} \end{aligned} \quad (53)$$

The generic un-normalized Cartesian Gaussian primitive functions can be actually written as:

$$g_{i\alpha}(\mathbf{r}; \mathbf{A}, L_\alpha) = (x - A_x)^{a_x} (y - A_y)^{a_y} (z - A_z)^{a_z} e^{-\alpha_i[(x-A_x)^2 + (y-A_y)^2 + (z-A_z)^2]} \quad (54)$$

and

$$g_{j\beta}(\mathbf{r}; \mathbf{B}, L_\beta) = (x - B_x)^{b_x} (y - B_y)^{b_y} (z - B_z)^{b_z} e^{-\beta_j[(x-B_x)^2 + (y-B_y)^2 + (z-B_z)^2]} \quad (55)$$

The former is an un-normalized Cartesian Gaussian function with exponent  $\alpha_i$ , centered on nucleus  $\mathbf{A} = (A_x, A_y, A_z)$ , with angular momentum indices  $\mathbf{a} = (a_x, a_y, a_z)$  and total angular momentum  $L_\alpha = a_x + a_y + a_z$ . The latter is an un-normalized Cartesian Gaussian function, with exponent  $\beta_j$ , centered on nucleus  $\mathbf{B} = (B_x, B_y, B_z)$ , with angular momentum indices  $\mathbf{b} = (b_x, b_y, b_z)$  and total angular momentum  $L_\beta = b_x + b_y + b_z$ . Note that exponents  $\alpha_i$  and  $\beta_j$  are analogous to exponents  $\tau_\alpha$  and  $\tau_\beta$  seen in Eqs. 39 and 40 for the expressions of the thermal factors proposed by Tanaka and Michael & Koritsanszky.

In the remaining part of this section, our focus will be on the computation of the integral  $I_{i\alpha}^{j\beta}$  appearing in Eq. 53:

$$I_{i\alpha}^{j\beta} = \int d\mathbf{r} g_{i\alpha}(\mathbf{r}; \mathbf{A}, L_\alpha) g_{j\beta}(\mathbf{r}; \mathbf{B}, L_\beta) e^{i\mathbf{K}\cdot\mathbf{r}} \quad (56)$$

Over the years, several techniques have been introduced to accomplish this task, from the pioneering works by McWeeny [42], Chandler and Spackman [43], and Barua and Weyrich [44], to the more recent efforts by Kirrander and coworkers, who took into account these kinds of integrals in their *ab initio* X-ray diffraction (AIXRD) method [45-47]. In the next two subsections, we will describe two efficient strategies of computing the integrals represented by Eq. 56. They are the approaches currently implemented for the methods of quantum crystallography: the first one (subsection 2.5.1) is the extension [48] of the McMurchie-Davidson method [49] and it is the technique currently implemented in the quantum crystallographic software *Tonto* [50] for performing Hirshfeld Atom Refinements and X-ray constrained wavefunction fittings. The second one (subsection 2.5.2) is the extension [51] of the Obara & Saika [52-54] and Head-Gordon & Pople [55] recurrence relations and it is the approach implemented and used for the computation of structure factors in the X-ray constrained ELMO

(XC-ELMO) [56-59], X-ray constrained ELMO-Valence-Bond (XC-ELMO-VB) [60, 61] and X-ray constrained Spin-Coupled (XCSC) [62, 63] techniques.

### 2.5.1 The McMurchie-Davidson approach

One way to evaluate these  $I_{i\alpha}^{j\beta}$  integrals is to use a modification [48] of the McMurchie-Davidson procedure. The reason why this method is useful in this context is that the primitive Gaussian basis function product may be decomposed into a sum of Hermite functions and the Fourier transform of a Hermite function is very easy to calculate.

Exploiting Eqs. 54 and 55, the integral  $I_{i\alpha}^{j\beta}$  can be rewritten as the product of three individual one-dimensional Fourier transforms:

$$I_{i\alpha}^{j\beta} = \int dx e^{iK_x(x-A_x)^{a_x}} e^{-\alpha_i(x-A_x)^2} (x-B_x)^{b_x} e^{-\beta_j(x-B_x)^2} \\ \int dy e^{iK_y y} (y-A_y)^y e^{-\alpha_i(y-A_y)^2} (y-B_y)^{b_y} e^{-\beta_j(y-B_y)^2} \\ \int dz e^{iK_z z} (z-A_z)^{a_z} e^{-\alpha_i(z-A_z)^2} (z-B_z)^{b_z} e^{-\beta_j(z-B_z)^2} \quad (57),$$

which, in short notation, becomes:

$$I_{i\alpha}^{j\beta} = \left( I_{i\alpha}^{j\beta} \right)_x \left( I_{i\alpha}^{j\beta} \right)_y \left( I_{i\alpha}^{j\beta} \right)_z \quad (58)$$

The McMurchie-Davidson idea is to write the integrand of each factor in terms of Hermite Gaussian functions  $\Lambda_t$ . For example, considering the variable  $x$ , we have:

$$(x-A_x)^{a_x} e^{-\alpha_i(x-A_x)^2} (x-B_x)^{b_x} e^{-\beta_j(x-B_x)^2} = \sum_{t_x=0}^{a_x+b_x} E_{t_x}^{(a_x, b_x)} \Lambda_{t_x}(x, \gamma, P_x) \quad (59),$$

where

$$\Lambda_{t_x}(x, \gamma, P_x) = \left( \frac{\partial}{\partial x} \right)^{t_x} e^{-\gamma(x-P_x)^2} \quad (60)$$

with

$$P_x = \frac{\alpha_i A_x + \beta_j B_x}{\gamma} \quad (61)$$

and

$$\gamma = \alpha_i + \beta_j \quad (62)$$

The reason why this expansion works is that the product of the two Gaussian functions includes polynomials of degree  $a_x + b_x$ , and the Hermite Gaussian functions form a basis for these polynomials. The problem is to evaluate the expansion coefficients  $E_{t_x}^{(a_x, b_x)}$ . A recurrence relation for these coefficients can be worked out from the recurrence relation obeyed by the Hermite polynomials:

$$\Lambda_{t_x+1} = -2 [\gamma \Lambda_{t_x-1} + (x - P_x) \Lambda_{t_x}] \quad (63)$$

or

$$-(x - A_x) \Lambda_{t_x} = t_x \Lambda_{t_x+1} + \frac{\beta}{\gamma} (A_x - B_x) \Lambda_{t_x} + \frac{1}{2\gamma} \Lambda_{t_x-1} \quad (64)$$

From this it is possible to deduce the required recurrence relation, which is

$$-E_{t_x}^{(a_x+1, b_x)} = (t_x + 1) E_{t_x+1}^{(a_x, b_x)} + \frac{\beta}{\gamma} (A_x - B_x) E_{t_x}^{(a_x, b_x)} + \frac{1}{2\gamma} E_{t_x-1}^{(a_x, b_x)} \quad (65)$$

with

$$E_0^{(0,0)} = e^{\frac{\alpha_i \beta_j}{\gamma} (A_x - B_x)^2} \quad (66)$$

The recurrence relation for the incremented  $b_x$  index can be obtained through a similar procedure, but a simpler procedure is to use a “transfer relation” whereby the polynomial part of one Gaussian is transformed to the polynomial part of the other, namely:

$$E_{t_x}^{(a_x, b_x+1)} = E_{t_x}^{(a_x+1, b_x)} + (A_x - B_x) E_{t_x}^{(a_x, b_x)} \quad (67)$$

Recurrence relations like this are discussed and used extensively in the following subsection. All that remains is to evaluate the Fourier transform of the Hermite Gaussian functions. Using standard results from Fourier analysis we have

$$F \left[ \left( \frac{\partial}{\partial x} \right)^{t_x} e^{-\gamma(x-P_x)^2} \right] = (-iK_x)^{t_x} \left( \frac{\pi}{\gamma} \right)^{\frac{1}{2}} e^{iK_x P_x} e^{-\frac{K_x^2}{\gamma}} \quad (68)$$

Therefore, for the one-dimensional Fourier integral, we have:

$$\left(I_{i\alpha}^{j\beta}\right)_x = \left(\frac{\pi}{\gamma}\right)^{\frac{1}{2}} e^{iK_x P_x} e^{-\frac{K_x^2}{\gamma}} \sum_{t_x=0}^{a_x+b_x} (-iK_x)^{t_x} E_{t_x}^{(a_x, b_x)} \quad (69)$$

The final result is a product of three terms like this:

$$I_{i\alpha}^{j\beta} = \left(\frac{\pi}{\gamma}\right)^{\frac{3}{2}} e^{i\mathbf{K}\cdot(i\mathbf{P}-\mathbf{K}/4\gamma)} \prod_{n=x,y,z} \sum_{t_n=0}^{a_n+b_n} (-iK_n)^{t_n} E_{t_n}^{(a_n, b_n)} \quad (70)$$

where,

$$\mathbf{P} = (P_x, P_y, P_z) = \frac{\alpha_i \mathbf{A} + \beta_j \mathbf{B}}{\gamma} \quad (71)$$

### 2.5.2 The Obara-Saika approach

In this subsection, we will show recurrence relations for the efficient computations of the integrals defined by Eq. 56. These recurrence relations have been obtained by extending the Obara-Saika [52-54] and Head-Gordon-Pople [55] recurrence relations, which were originally introduced in quantum chemistry for the evaluation of one- and two-electron integrals when Cartesian Gaussians are used as basis functions.

Defining

$$E_{i\alpha}^{j\beta} = e^{-\frac{\alpha_i \beta_j}{\alpha_i + \beta_j} (\mathbf{A}-\mathbf{B})^2} \quad (72)$$

and bearing in mind the definition of  $\mathbf{P}$  given by Eq. 71, if we substitute Eqs. 54 and 55 into Eq. 56, the integral  $I_{i\alpha}^{j\beta}$  can be also rewritten like this:

$$\begin{aligned} I_{i\alpha}^{j\beta} = E_{i\alpha}^{j\beta} \int dx e^{iK_x x} (x - A_x)^{a_x} (x - B_x)^{b_x} e^{-(\alpha_i + \beta_j)(x - P_x)^2} \\ \int dy e^{iK_y y} (y - A_y)^{a_y} (y - B_y)^{b_y} e^{-(\alpha_i + \beta_j)(y - P_y)^2} \\ \int dz e^{iK_z z} (z - A_z)^{a_z} (z - B_z)^{b_z} e^{-(\alpha_i + \beta_j)(z - P_z)^2} \quad (73), \end{aligned}$$

which can be expressed in short notation as:

$$I_{i\alpha}^{j\beta}(\mathbf{a}|\mathbf{b}) = I_{i\alpha}^{j\beta}(a_x, a_y, a_z | b_x, b_y, b_z) = E_{i\alpha}^{j\beta} I_x(a_x | b_x) I_y(a_y | b_y) I_z(a_z | b_z) \quad (74)$$

Let us consider the factor  $I_x(a_x|b_x)$ :

$$I_x(a_x|b_x) = \int dx e^{iK_x x} (x - A_x)^{a_x} (x - B_x)^{b_x} e^{-(\alpha_i + \beta_j)(x - P_x)^2} \quad (75).$$

Through a simple algebraic derivation [51], it is possible to show that:

$$\begin{aligned} I_x(a_x + 1|b_x) &= \left( P_x - A_x + \frac{iK_x}{2(\alpha_i + \beta_j)} \right) I_x(a_x|b_x) + \frac{a_x}{2(\alpha_i + \beta_j)} I_x(a_x - 1|b_x) \\ &+ \frac{b_x}{2(\alpha_i + \beta_j)} I_x(a_x|b_x - 1) \end{aligned} \quad (76),$$

which is the vertical recurrence relation (VRR) for the “bra part” of the Fourier transform integral.

In analogous way, it is also possible to show that:

$$\begin{aligned} I_x(a_x|b_x + 1) &= \left( P_x - B_x + \frac{iK_x}{2(\alpha_i + \beta_j)} \right) I_x(a_x|b_x) + \frac{a_x}{2(\alpha_i + \beta_j)} I_x(a_x - 1|b_x) \\ &+ \frac{b_x}{2(\alpha_i + \beta_j)} I_x(a_x|b_x - 1) \end{aligned} \quad (77),$$

which is the vertical recurrence relation (VRR) for the “ket part” of the Fourier transform integral.

Finally, comparing Eqs. 76 to 77, one can also obtain:

$$I_x(a_x|b_x + 1) = I_x(a_x + 1|b_x) + (A_x - B_x) I_x(a_x|b_x) \quad (78),$$

which is the horizontal recurrence relation (HRR). Note that this is essentially the same relation as expressed in Eq. 67.

It is worth noting that the starting point of these recurrence relations is the following integral involving only s-type Cartesian Gaussian functions (namely, Cartesian Gaussian functions with total angular momentum equal to zero), which is easy to compute:

$$I_x(0|0) = \int dx e^{iK_x x} e^{-(\alpha_i + \beta_j)(x - P_x)^2} = \sqrt{\frac{\pi}{\alpha_i + \beta_j}} e^{-\frac{K_x^2}{4(\alpha_i + \beta_j)}} e^{iK_x P_x} \quad (79)$$

The recurrence relations obtained above for the  $I_x(a_x|b_x)$  integral, can be also obtained for the integrals  $I_y(a_y|b_y)$  and  $I_z(a_z|b_z)$ . Therefore, Eqs. 76-78 can be generalized like this:

$$I_{ia}^{jb}(\mathbf{a} + \mathbf{1}_w|\mathbf{b}) = \left( P_w - A_w + \frac{iK_w}{2(\alpha_i + \beta_j)} \right) I_{ia}^{jb}(\mathbf{a}|\mathbf{b}) + \frac{a_w}{2(\alpha_i + \beta_j)} I_{ia}^{jb}(\mathbf{a} - \mathbf{1}_w|\mathbf{b}) + \frac{b_w}{2(\alpha_i + \beta_j)} I_{ia}^{jb}(\mathbf{a}|\mathbf{b} - \mathbf{1}_w) \quad (80)$$

$$I_{ia}^{jb}(\mathbf{a}|\mathbf{b} + \mathbf{1}_w) = \left( P_w - B_w + \frac{iK_w}{2(\alpha_i + \beta_j)} \right) I_{ia}^{jb}(\mathbf{a}|\mathbf{b}) + \frac{a_w}{2(\alpha_i + \beta_j)} I_{ia}^{jb}(\mathbf{a} - \mathbf{1}_w|\mathbf{b}) + \frac{b_w}{2(\alpha_i + \beta_j)} I_{ia}^{jb}(\mathbf{a}|\mathbf{b} - \mathbf{1}_w) \quad (81)$$

$$I_{ia}^{jb}(\mathbf{a}|\mathbf{b} + \mathbf{1}_w) = I_{ia}^{jb}(\mathbf{a} + \mathbf{1}_w|\mathbf{b}) + (A_w - B_w) I_{ia}^{jb}(\mathbf{a}|\mathbf{b}) \quad (82)$$

where  $w = x, y$  or  $z$  and  $\mathbf{1}_w = (\delta_{wx}, \delta_{wy}, \delta_{wz})$  with  $\delta$  as the usual Kronecker's delta, and where the starting point for all the recurrence relations is the integral

$$I_{ia}^{jb}(\mathbf{0}|\mathbf{0}) = E_{ia}^{jb} I_x(0|0) I_y(0|0) I_z(0|0) = E_{ia}^{jb} \left( \frac{\pi}{\alpha_i + \beta_j} \right)^{\frac{3}{2}} e^{-\frac{K^2}{4(\alpha_i + \beta_j)}} e^{i\mathbf{K}\cdot\mathbf{P}} \quad (83)$$

Finally, it is worth noting that, since Eq. 82 depends only on the nuclear centers of the primitives, it can be also directly exploited to evaluate integrals between contracted basis functions (see Eq. 53):

$$I[\mathbf{a}|\mathbf{b} + \mathbf{1}_j] = I[\mathbf{a} + \mathbf{1}_j|\mathbf{b}] + (A_j - B_j) I[\mathbf{a}|\mathbf{b}] \quad (84)$$

where the square brackets indicate that the integrals involve contracted Cartesian Gaussian basis functions. The use of this equation over contracted functions is the key to reduce the computational cost for highly contracted basis functions.

### 3. Introduction to Hirshfeld Atom Refinement

#### 3.1 Ideas behind the technique of Hirshfeld Atom Refinement

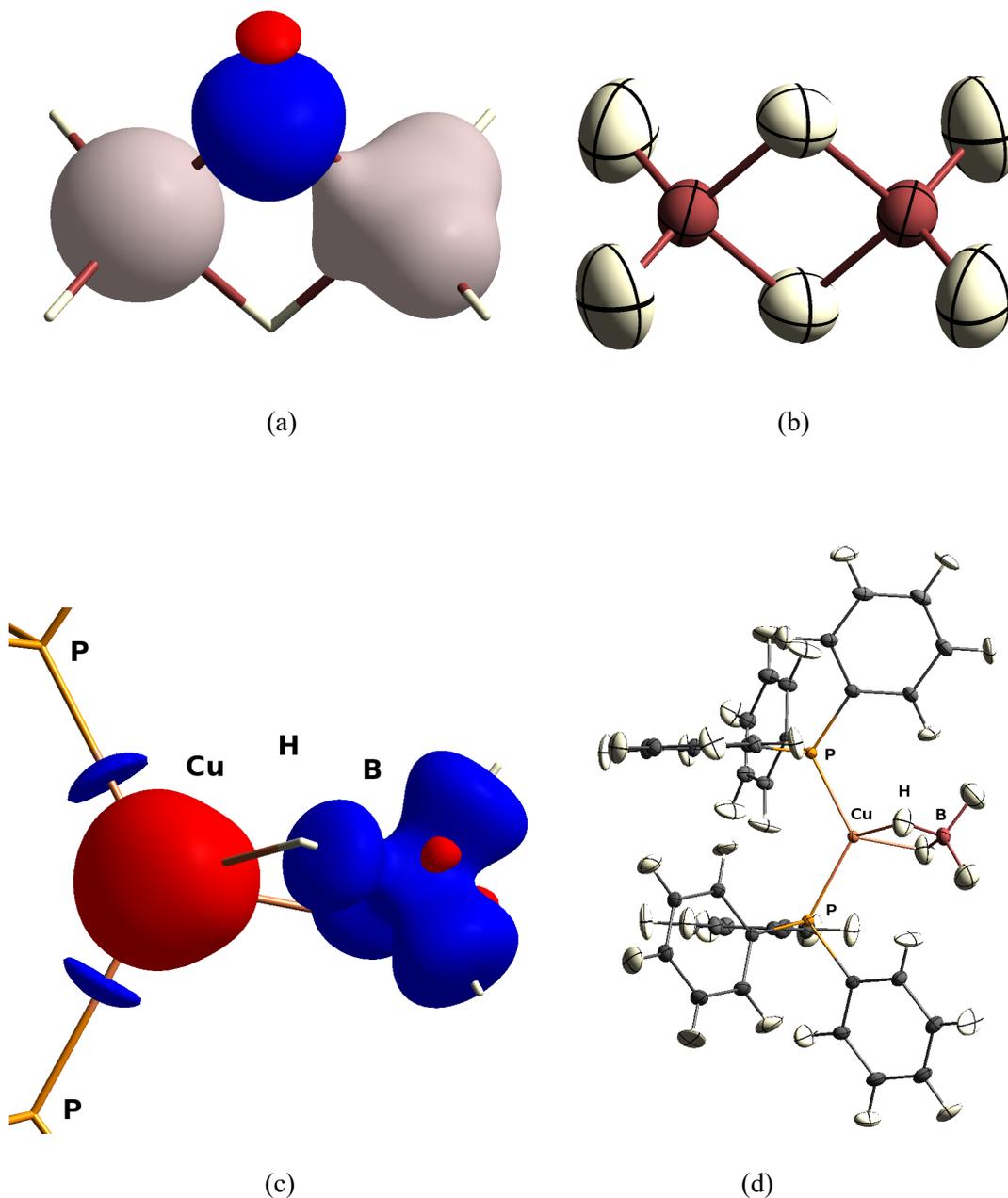
Atoms in molecules are not spherical but are deformed to a smaller or larger extent due to their interactions with neighboring atoms. Standard X-ray structure determination does not account for this effect as it uses the spherical form factors of the Independent Atom Model (IAM). Several ways to deal with this problem have been presented in Section 2. Among them, Hirshfeld Atom Refinement (HAR) is a simple and reliable method to account for the fact that X-rays and

electrons are not scattered from spherical atomic electron densities in molecules (subsection 2.3) [15, 16]. The method is based on the Hirshfeld “stockholder” partitioning technique [28, 29] which has been described in subsection 2.3 and which provides aspherical atomic electron densities  $\rho_A(\mathbf{r})$  given by Eq. 25.

The Fourier transform  $f_A$  of  $\rho_A$  is the non-spherical Hirshfeld atom form factor. It replaces the usual spherical atomic form factor in the least-squares refinement of the atomic coordinates and displacement parameters. Hence, Hirshfeld Atom Refinement is defined as the quantum-crystallographic structure refinement technique where the non-spherical atomic form factors are calculated as Fourier transforms of Hirshfeld atoms according to Hirshfeld’s stockholder partitioning.

Fig. 2(a) shows the shape of the boron Hirshfeld atom in the  $B_2H_6$  molecule in an isosurface representation of electron density. Its non-spherical nature is obvious, especially when compared to the spherical isosurface with the same electron-density value. The difference between promolecule and non-spherical atom density is called *deformation Hirshfeld density*, and it is the signature of deformation by chemical bonding. The blue isosurfaces in Figs 2(a) and 2(c) show the build-up of electron density due to the deformation into covalent bonding regions.

The deformation of electron density into the bonding region is especially consequential for hydrogen atoms, which unlike all other atoms in the periodic table have no spherical, highly contracted core electrons that scatter to high diffraction angles. Hydrogen atoms have only a single valence electron that is deformed into the direction of the covalently bonded interaction partner. Any improved description of hydrogen atom parameters therefore depends on reliable descriptions of the non-spherical electron densities and the corresponding atomic form factors in their respective chemical environments. Once this information is included, hydrogen atom positions and displacement parameters can be refined from the X-ray data accurately and precisely in HAR, see Figs 2(b) and 2(d) (see also subsection 5.1.1) [19].



**Fig. 2.** Visualization of Hirshfeld atoms and deformation Hirshfeld densities.

(a) Diborane  $B_2H_6$  [64], isosurfaces of the boron electron densities used to calculate the atomic form factors in the IAM (left, grey,  $\rho=0.08 \text{ e}/\text{\AA}^3$ ) and in HAR (right, grey,  $\rho=0.08 \text{ e}/\text{\AA}^3$ ). The difference between the Hirshfeld density and the spherical atom density (deformation Hirshfeld density) is depicted for one of the bridging H-atoms ( $\rho=0.006 \text{ e}/\text{\AA}^3$ ). Blue = positive deformation Hirshfeld density

(aspherical atom density > spherical atom density), red = negative deformation Hirshfeld density (aspherical atom density < spherical atom density).

(b) HAR-based molecular structure of diborane with anisotropic displacement parameters (ADPs) for all atoms at a 50% probability level.

(c) Deformation Hirshfeld densities for copper and boron atoms in (tetrahydroborato)bis(triphenylphosphine)copper(I) ( $\rho=0.006 \text{ e}/\text{\AA}^3$ ) [65]. Color code: see (a).

(d) HAR-based molecular structure of the compound in (c) with ADPs for all atoms at a 50% probability level.

Reprinted under the CC BY-NC license from reference [19].

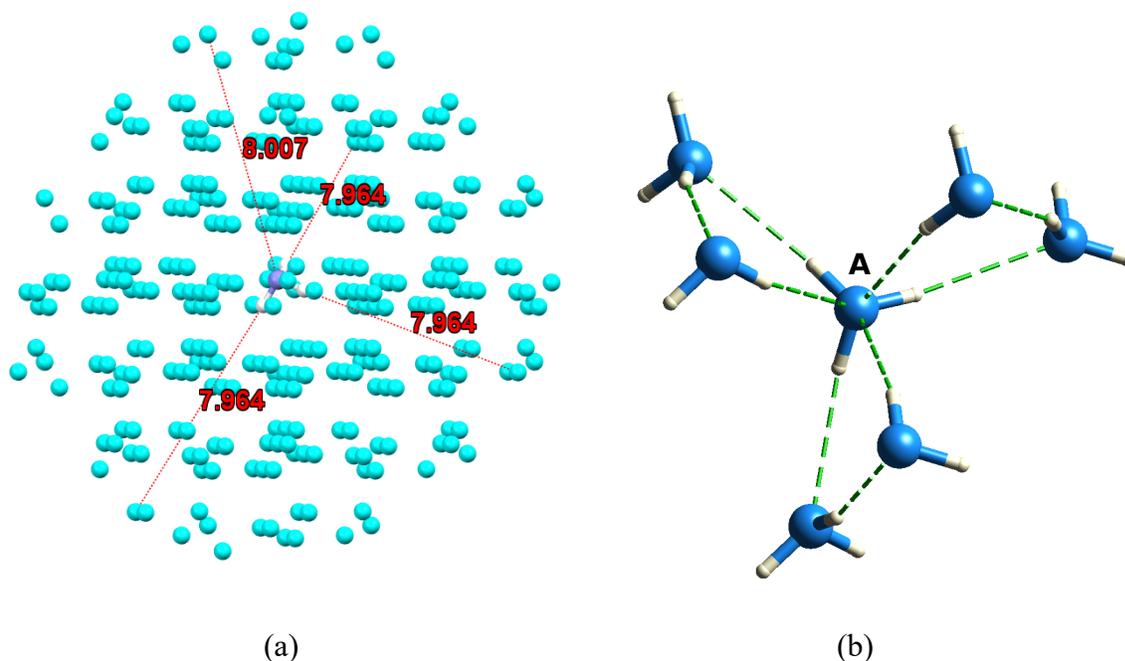
HAR was introduced by Jayatilaka and Dittrich in 2008 [15]. In the original version, a *rigid-atom approximation* was used, which means that the Hirshfeld atoms were determined once and then held fixed during the structural least-squares refinement. This is basically the equivalent of a normal structure refinement, with the only difference that the list of generalized spherical atomic form factors as deposited in the International Tables of Crystallography is replaced with a list of tailor-made non-spherical form factors for the atoms constituting the specific molecule under consideration. In the original version, this new list is never stored, but always recalculated on the fly. Already within the rigid-atom approximation, Jayatilaka and Dittrich showed that it is possible to obtain hydrogen-atom positions that are in agreement with positions from neutron diffraction for urea and benzene [15].

In a subsequent, iterative and automated version of HAR, the rigid-atom approximation was dropped and replaced by a two-step procedure [16]. The first step is called electron-density step (*ED step*) because here the ED is calculated quantum mechanically and subdivided into Hirshfeld atoms, which are afterwards Fourier transformed to obtain non-spherical atomic form factors. The second step is called *structural refinement step* and consists of several regular least-squares refinement cycles. The sequence of an ED step and a structural refinement step constitutes a *HAR iteration*. HAR iterations are repeated until convergence is reached, i.e. until parameter shifts divided by their standard uncertainties are less than 0.01 [16]. In this implementation of HAR, the ED step may be based on the HF or DFT methods with BLYP and B3LYP [66] functionals. A wide variety of all-electron basis sets is available, and new basis sets

can be adopted and added easily. The software inside which HAR was developed is *Tonto* (<https://github.com/dylan-jayatilaka/tonto>) [50]. The native version of *Tonto* controlled by an input file is still the most used program for performing HARs.

### ***3.2 The crystal asymmetric unit and its environment***

In the simplest implementation of HAR, *Tonto* calculates the quantum mechanical electron density of an isolated molecule or reference cluster. However, for strong intermolecular interactions such as hydrogen bonds, the crystal field must be accounted for. For this purpose, an implicit simulation of the environment by using point charges that perturb the molecular wavefunction is normally carried out. Within the ED step, atomic charges and point dipole moments are calculated from the Hirshfeld atoms and placed at symmetry-generated positions around the central molecular or reference unit. Molecules with at least one atom within a certain radius from any atom in the reference unit are included. Normally a radius of 8 Å is chosen (Fig. 3 (a)), but the radius is adjustable. A new ED is calculated in the presence of the simulated crystal field, new Hirshfeld atoms and Hirshfeld charges and moments are generated, and this procedure is iterated until convergence in the atomic charges, dipole moments and in the molecular energy have been reached. However, for species that are not unambiguously molecular or where the crystallographic symmetry prevents completion of neutral spherical clusters, the option to build an environment with complete molecules can be overwritten and refinements allowed to proceed with a simulated environment that might be charged overall.



**Fig. 3.** a) Cluster of point charges and point dipoles (light blue) within a radius of  $8\text{\AA}$  around a central ammonia molecule (dark blue). The wavefunction for the central ammonia molecule perturbed by the environment is used for the Hirshfeld atom partitioning. b) Explicit cluster of molecules. The wavefunction is calculated for the entire cluster, but the Hirshfeld atom partitioning is only performed for the symmetry-independent atoms of the central molecule A [25]. Only a third of the ammonia molecule is symmetry-independent, so only one nitrogen and one hydrogen atom form factor is calculated, and only the parameters of these two atoms in the asymmetric unit are refined.

Alternatively, a molecular cluster can be defined, and the wavefunction calculated for the entire cluster (Fig. 3(b)). This introduces the effects of intermolecular interactions explicitly into the electron-density distribution of the central molecule. Hirshfeld atoms and corresponding atomic form factors for use in the refinements are calculated only for the symmetry-independent atoms in the central molecule or reference unit (see caption of Fig. 3, where the example of crystalline ammonia is discussed). For systems containing strong hydrogen bonds with covalent contributions, it is mandatory to use such an explicit cluster, if accurate parameters for the affected hydrogen atoms are to be determined [20]. For salts and covalent network compounds, the choice of explicit clusters becomes even more essential and critical, since space-group symmetry may not allow to build a cluster that maintains the symmetry of the reference unit and

is electrically neutral, or may require covalent bonds to be broken. In addition, the software *Tonto* does not impose symmetry restrictions on the Hirshfeld densities of atoms in special positions but deletes linear dependencies; higher-symmetry space groups are thus especially problematic in the original HAR approach. For such systems, calculations with periodic boundary conditions can be a solution. Subsection 3.4 describes recent developments addressing this problem.

### 3.3 A minimal HAR and HART

To facilitate interfacing HAR to other refinement software such as *Olex2* [67] and to improve user-friendliness, the HAR routines were extracted from the software *Tonto*, so that they compile into a HAR standalone program called *HART* (HAR terminal). *HART* is available from <https://github.com/dylan-jayatilaka/tonto> or as a refinement tool in *Olex2* [20]. On a Linux or Windows console or on a terminal, *HART* requires a crystallographic information file (CIF) as input and reflection data either embedded in the CIF or as a separate file: *hart [options] <name-of.cif>*. The output after the refinement consists of a modified CIF.

If no further options are specified, a *minimal HAR* will be performed: minimal default settings are loaded in order to ensure a meaningful, yet quick refinement. The minimal settings imply a restricted Hartree-Fock calculation with the def2-SVP basis set without simulation of the environment. Beyond a minimal HAR, a full set of options is available from *hart -help*. They include the choice of the quantum-chemical method (HF or BLYP) and basis set (STO-3G – only for testing; def2-SVP and cc-pVDZ – adequate; def2-TZVP and cc-pVTZ – excellent; def2-TZVPP and cc-pVQZ – benchmark), the cluster radius for the implicit simulation of the crystal environment, the treatment of the hydrogen-atom displacement parameters, and details about anomalous dispersion correction as well as extinction. Note that calculation of structure factors requires all-electron basis sets which include core electrons. Thus, for heavier elements, the quantum mechanical calculations can become time-consuming and unstable. Hence, *ab initio* wavefunctions are not always robust, and relativistic effects cannot be treated in *HART* (see subsection 3.5 for relativistic HARs). For these reasons, refinement of molecules containing transition metals and of molecules with elements heavier than Kr is disallowed (see subsection 3.8 for alternative options).

In *HARt*, the wavefunction will be calculated for the block of atomic coordinates specified in the CIF. For a molecular compound with  $Z'=1$  (one formula unit in the asymmetric unit), this convention works well. However, whenever  $Z'<1$  or whenever an explicit cluster of molecules is needed in the case of strong intermolecular interactions, network compounds or salts, the block of atomic coordinates and their symmetry-related ADPs must be expanded manually (see the ammonia example in Figure 3 (b)). This can be achieved more easily through the *Olex2* interface to *HARt* [20]. In *Olex2*, the cluster of molecules can be constructed on the screen and then handed over to *HARt* internally. The same options and default settings as for the terminal use of *HARt* are available as drop-down menus in the *Tools* section of *Olex2*. Several jobs can be run at the same time, *Olex2* remains functional in the meantime and can even be closed without terminating the *HARt* job.

### **3.4 Periodic HAR**

A periodic HAR is the best way to deal with network compounds and salts. It minimizes the introduction of artefacts into the atomic form factors resulting from the termination implied in an explicit, finite cluster of molecules. Periodic adaptations of HAR are computationally intensive, though, and are thus the opposite of a minimal HAR – that is supposed to give a quick, but meaningful approximation of the electron density in a molecular crystal. Here, a periodic wavefunction is calculated and then transformed to make Hirshfeld's stockholder partitioning possible. The program *Crystal* [68] constructs periodic wavefunctions built from atomic basis functions but attempts to extract Hirshfeld atoms from them are largely untested and unpublished.

A wavefunction for only the symmetry-unique part of an infinite system is best obtained with a localized approach. Stoll and coworkers have developed such methods. The localized Wannier orbitals are obtained either by imposing an orthogonality constraint between the local and infinite system [69] or by optimizing the Wannier orbitals with the extremely localized molecular orbital (ELMO) method [70]. We call the latter “extremely localized Wannier functions” (ELWFs) to distinguish them from the related class of “maximally localized Wannier functions” (MLWFs) [71]. However, the ELWF method is unpublished and has not been implemented in software yet.

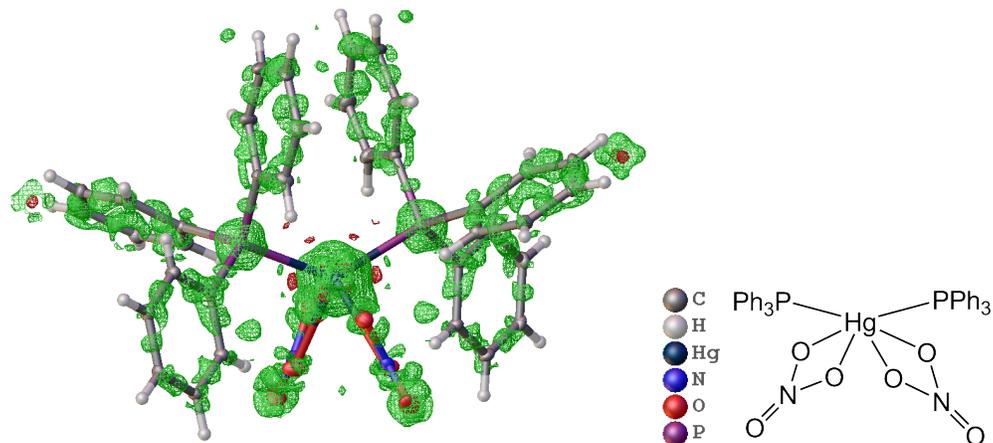
A periodic procedure resembling HAR, but without refinement of coordinates and ADPs, has been tested and published by Wall [18]. It is implemented in the software *Lunus* [72] through an interface to *VASP* [73]. A plane-wave periodic-boundary wavefunction was computed for the molecule urea, the total electron density derived from it, Hirshfeld’s stockholder partitioning applied and the Hirshfeld atoms convoluted with ADPs. In this way, X-ray structure factors were calculated. The resulting electron-density model was compared with electron densities obtained using other techniques such as multipole modeling. Wall concluded [18] that “ $2F_o - F_c$  maps and static charge densities that are distinct from the multipole model, but that nevertheless agree comparably with the experimental data” can be obtained by HAR and that HAR “can yield accurate charge densities that are consistent simultaneously with theory and experiment.”

### 3.5 Relativistic HAR

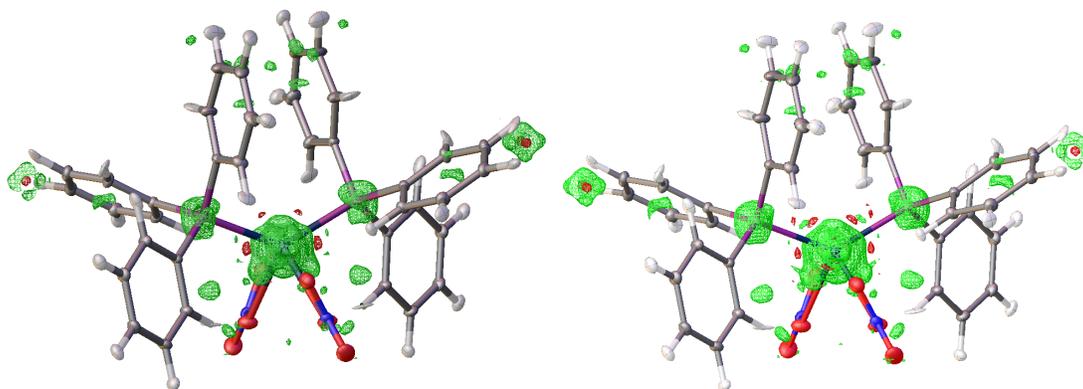
The refinement of structural models using HAR for compounds containing heavy elements may be error-prone or unstable, as mentioned in subsection 3.3. All-electron basis sets and relativistic methods must be combined, which increases the computational effort dramatically. Nevertheless, the Douglas-Kroll-Hess (DKH2) and infinite-order two-component (IOTC) methods of treating relativistic effects, which include scalar, spin-orbit and picture-change error, were implemented into *Tonto* by Bučinský [74]. The influence of the relativistic extensions on the structure factors was investigated and discussed [75, 76], as is also mentioned in subsection 4.3. Relativistic HARs at IOTC and BLYP-IOTC levels were reported for the molecules  $\text{HgPh}_2$  and  $\text{BiPh}_3$  [75, 76].

Here, in Fig. 4, we show first results of a B3LYP-IOTC HAR of the molecule  $(\text{Ph}_3\text{P})_2\text{Hg}(\text{NO}_3)_2$ , as described in the supporting information of the recent paper about the HAR-ELMO method [77], which will also be discussed below (see subsection 3.6). The refinement was carried out using *Tonto*. The regular HAR at HF non-relativistic level accounts completely for all bonding electron-density features not modeled in IAM (compare Fig. 4(a) with 4(b)). Two symmetry-independent regions of satellite or ghost peaks remain (the molecule lies on a mirror plane). In addition, unmodelled core features for both the P and Hg atoms remain. The positive residual electron density at the Hg core was lowered at its maximum only marginally from 2.1 for the IAM to  $1.7 \text{ e } \text{Å}^{-3}$  for the non-relativistic HF model, i.e. only ca.  $0.4 \text{ e } \text{Å}^{-3}$  are due to the effect of

the asphericity of the atomic core electron-density distribution. Introducing the high-level B3LYP-IOTC relativistic treatment (including not only scalar, but also spin-orbit effects) reduces the maximum residual electron density at the Hg core by another  $0.6 \text{ e } \text{\AA}^{-3}$  (Fig. 4(c)). However, all other residual features remain essentially the same, and the absolute value at the Hg atom is still above  $1.0 \text{ e } \text{\AA}^{-3}$ . This prevents any meaningful interpretation of the core electron-density distribution. Moreover, the overall time of the refinement increases from 11 hours (HAR) to 8 days and 19 hours (relativistic HAR) on a single CPU of a desktop computer. This example indicates that an accurate crystallographic treatment of the electron density near heavy elements is presently beyond reach. It probably requires a major community effort in the future, including revisiting data reduction strategies such as Bragg peak integration and absorption correction.



(a) IAM,  $-0.843/+2.072 \text{ e } \text{\AA}^{-3}$



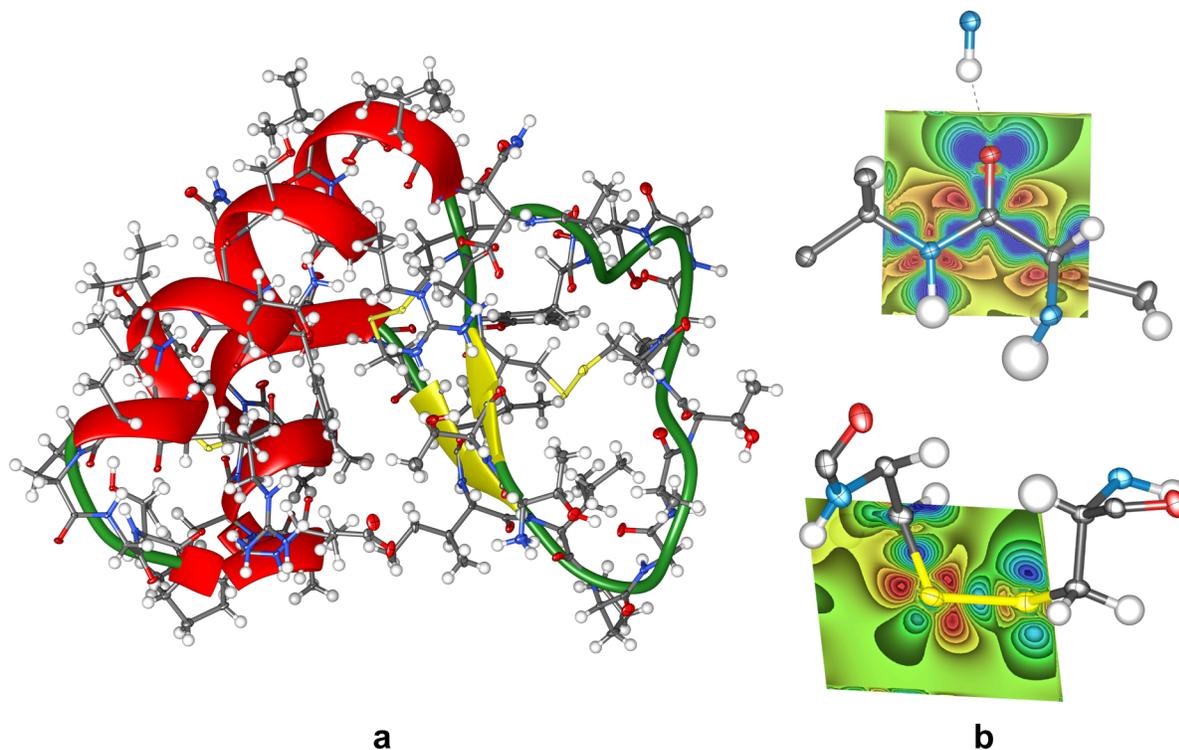
(b) HAR, HF/DZP,  $-0.675/+1.672 \text{ e } \text{\AA}^{-3}$  (c) HAR at B3LYP-IOTC/DZP,  $-0.805/+1.149 \text{ e } \text{\AA}^{-3}$

**Fig. 4.** Molecular structures and residual electron density maps of  $(\text{Ph}_3\text{P})_2\text{Hg}(\text{NO}_3)_2$  after refinement with (a) the independent atom model (IAM), (b) regular HAR, and (c) relativistic HAR. The minimum and maximum residual electron density peaks are also given. These are located at the Hg atom. Isovalue =  $\pm 0.2 \text{ e } \text{\AA}^{-3}$ , green=positive, red=negative. ADPs at 50% probability level.

### 3.6 HAR-ELMO

It has been shown that the atomic positions and displacement parameters determined with HAR for elements of period 2 and for hydrogen atoms are as accurate and precise as those from neutron diffraction (see subsection 5.1) [19]. The method should therefore be applicable to biological macromolecules as well. However, the problem with applying HAR to biological macromolecules is their size and the concomitant need of computer time. In 2016, Wall concluded that [18] “increasing the speed of calculations would enable the wider adaptation of HAR by adapting existing small-molecule crystallographic workflows [...] for obtaining high-resolution charge density models of [...] macromolecular crystals.”

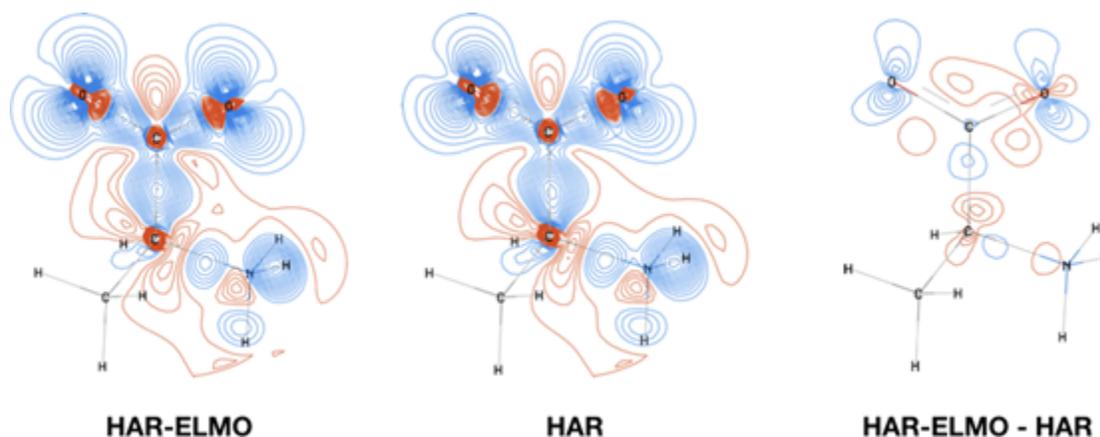
We have developed a method called HAR-ELMO that alleviates this problem [77]. The acronym ELMO stands for extremely localized molecular orbitals [70], a method that allowed a database of molecular orbitals strictly localized on atoms, bonds or functional groups to be constructed [78]. The database software *ELMOdb* can construct full wavefunctions of macromolecules within seconds using the deposited ELMOs, which are then used in the ED step in HAR within the software *Tonto*. The software *lamaGOET* (see subsection 3.7) acts as the interface. This procedure exploits the fact that ELMOs have been shown to be as transferable as multipole populations to construct electron-density maps and to calculate non-spherical atomic form factors for interpreting the measured structure factors [79].



**Fig. 5.** (a) Refined protein structure of crambin at  $d=0.54$  Å, not showing the disordered regions. (b) Deformation density maps of crambin in a peptide region (C=O in Leu25) and a disulfide bond (between Cys4 and Cys32). Contour interval:  $0.05$  eÅ<sup>-3</sup>, blue=positive, red=negative, green=zero. For (a) and (b), non-H ADPs at 50% probability. For (a), all H atoms are drawn with fixed spheres of  $0.3$  Å radius. For (b) isotropic H-atom displacement parameters are also drawn at 50% probability. Reprinted with permission from reference [77]. Copyright (2019) American Chemical Society.

Fig. 5 (a) shows the HAR-ELMO-refined structure of the protein crambin using some constraints, but no restraints. Fig. 5 (b) shows deformation electron-density features in regions of a peptide and a disulfide bond, exemplifying that bonding and lone-pair electron-density distributions are captured accurately for crambin. R-values obtained with HAR-ELMO are virtually identical with those obtained by regular HAR, as shown, e.g., for a fibril-forming segment of the human prion protein [77]. Note that the time of the refinement was reduced from 1 day and 7 hours (HAR) to 23 minutes (HAR-ELMO) on a single desktop-computer CPU. Normal HARs of proteins are prohibitive: for crambin, even the HAR-ELMO refinement took about 10 days. But this investment in time and computer resources for macromolecular

biochemistry will be worthwhile because the information value of HAR-ELMO structures and wavefunctions allows insights not available otherwise. For example, protonation states will be accessible because of the reliable localization of hydrogen atoms [77], oxidation states in metalloproteins will be deducible because of the quantum chemical wavefunction on which the HAR-ELMO method is based (compare reference [80]), and intermolecular interaction energies can be calculated [77].



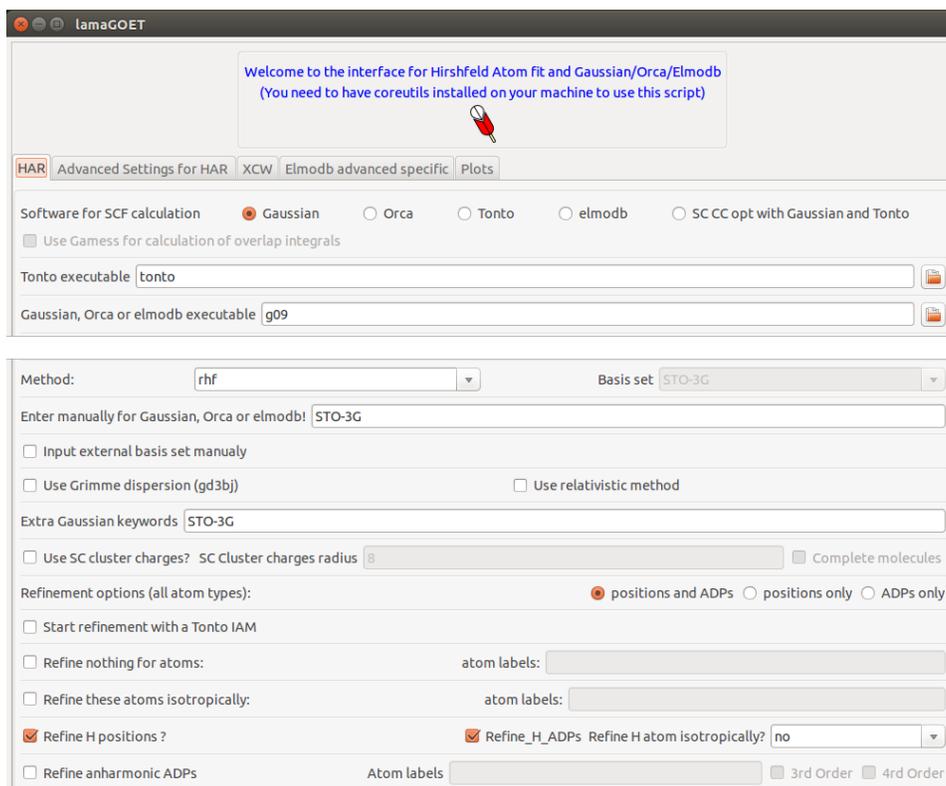
**Fig. 6.** Deformation density maps for *L*-Ala in the plane of the carboxylic group for HAR-ELMO and HAR refinements at 23 K. The difference between the HAR-ELMO and HAR deformation density maps is also shown. Contour levels: from 0.05 to 1.00 e Å<sup>-3</sup> with a step of 0.05 e Å<sup>-3</sup>. Blue (positive) and red (negative).

The performance of HAR and HAR-ELMO are gauged here by comparing respective deformation density maps of *L*-Ala (Fig. 6) based on data measured at 23K [81]. Visually, the deformation electron density distributions resulting from HAR and HAR-ELMO refinements look identical, but the HAR-ELMO minus HAR map (Fig. 6) reveals systematic differences in the oxygen lone pairs and the C-C bonds. It will be investigated in upcoming studies to what extent this effect of the localization scheme will influence derived properties.

### 3.7 The lamaGOET interface for quantum crystallography

There have been attempts in the past to use a graphical user interface (GUI) to facilitate preparing input files for *Tonto*, e.g. in an earlier version of *Olex2* or in *WinGX* [82]. However,

the flexibility of *Tonto* for HAR, but also for X-ray constrained wavefunction fitting (see Section 4), has never been fully exploited or made easily accessible. Therefore, the software *lamaGOET* was developed as a general interface for quantum crystallography (Fig. 7) [83]. It not only automatically deals with *Tonto* input files and monitors *Tonto* processes, but it also allows calculating wavefunctions with Gaussian and Orca or constructing them from the ELMO database (see subsection 3.6), and then handing them to *Tonto*. This idea was later also exploited in the *NoSpherA2* modification of the *Olex2* software (see subsection 3.8).

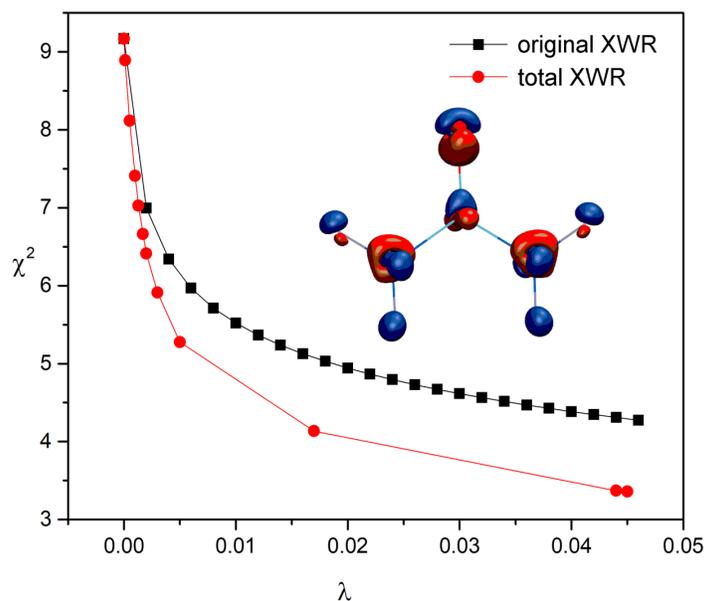


**Fig. 7.** The *lamaGOET* GUI combines input preparation for and control of HAR, HAR-ELMO and X-ray constrained wavefunction fitting in *Tonto* with different software packages such as *Gaussian*, *Orca* and *ELMOdb*. There are many more options than shown in the example above. (*lamaGOET* was designed and written by L. A. Malaspina in the Grabowsky group, University of Bremen.)

The interfacing of different programs such as *Tonto*, *Gaussian*, *Orca* and *ELMOdb* calls for special features of *lamaGOET* that exceed the capabilities of *Tonto*. A much wider variety of quantum-chemical methods and basis sets are available for HAR, basically all methods and basis sets available in Gaussian and Orca. Empirical dispersion energy corrections (Grimme’s GD3BJ

[84]) or relativistic extensions (e.g. DKH2 or IOTC) may be added to the quantum-chemical method, e.g. any DFT variety. Self-consistent Hirshfeld atomic point charges and point dipoles as calculated with *Tonto* (see subsection 3.2) can be extracted and used for wavefunction calculations in *Gaussian* or *Tonto*. Advanced settings for HAR include hydrogen-atom treatment, normalization of bond lengths, convergence criteria etc. *lamaGOET* reads *ELMOdb* output and uses it as input for HAR-ELMO treatments in *Tonto*; it also allows loading of tailor-made ELMOs for special residues or functional groups or ELMO wavefunctions for entire molecules.

*lamaGOET* controls the level of theory and degree of perturbation ( $\lambda$ -step, for details see section 4) used in X-ray constrained wavefunction fitting. It ensures that an XCW fitting following a regular HAR is performed at the same level of theory. We call this combination of HAR and XCW fitting X-ray wavefunction refinement (XWR; see also discussion in subsection 4.1) [85, 86]. The wavefunction resulting from XCW fitting may now be used for Hirshfeld atom partitioning and subsequent structure refinement again. If the procedure of XCW fitting and HAR is iterated, we call it *total XWR*. It leads to both improved geometries and ADPs as well as to a wavefunction consistent with the experimental data. The method is unpublished, but a first result for the molecule urea is shown in Fig. 8. Total XWR is one of the most advanced options in *lamaGOET*, not yet available in the current distribution.

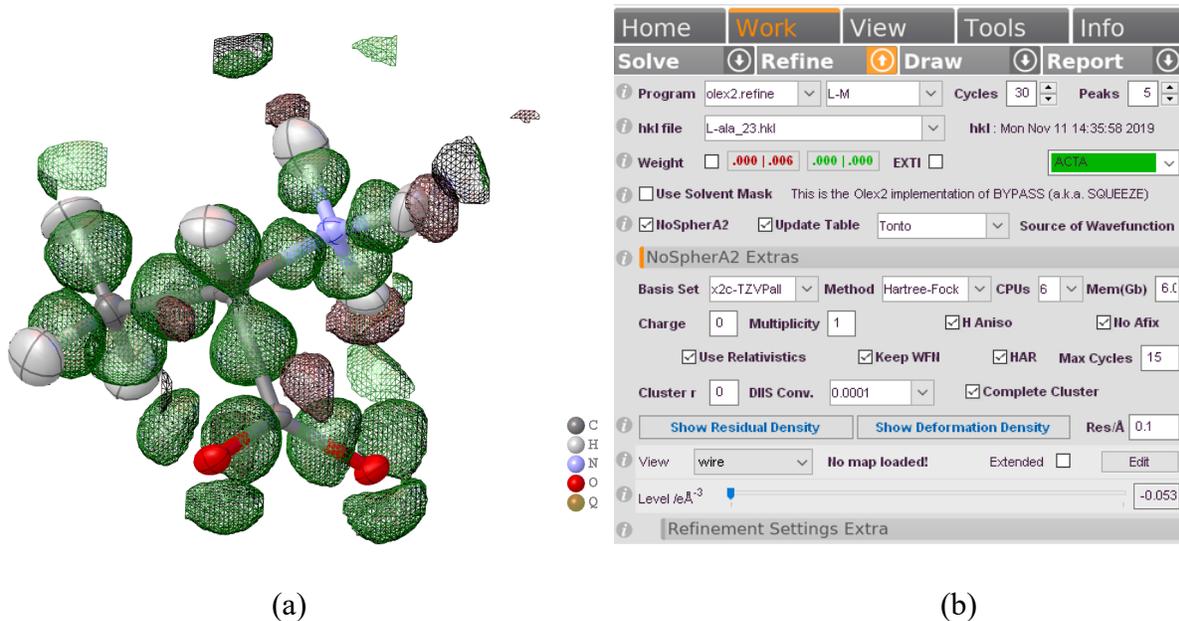


**Fig. 8.** Improvement of the fit measured in terms of  $\chi^2$  between *total XWR* and original XWR for urea. For an explanation of XCW fitting,  $\chi^2$  and  $\lambda$ , see section 4. The inset shows the difference electron density

between both XWR variants, i.e. original minus *total XWR*; isovalue =  $0.034 \text{ e } \text{\AA}^{-3}$ , blue = positive, red = negative. Unpublished results produced by R. Pal in the Grabowsky group of the University of Bremen.

### 3.8 NoSpherA2 in Olex2

In the original implementation of HAR within *Tonto* or *HARt* atomic form factors are not stored but recalculated on the fly after each refinement cycle to account for geometry changes. This makes it impossible to interface HAR with any other software optimized for crystallographic structure refinement, such as, e.g., the software *olex2.refine*. Recently, *olex2.refine* has been extended to use tabulated atomic form factors of any kind in regular structural least-squares [87], which allows to interface it with external quantum mechanical packages without the need of changing the code. The file format is called “tsc”, and requires for all symmetry-independent structure factors a list of the Miller indices and the respective form factors of the  $N_a$  atoms in the asymmetric unit:  $\{h k l f_i(h k l)\}_{i=1}^{N_a}$ . This input format is sufficiently general to allow storage of every flavor of atomic form factors, non-spherical externally calculated ones, e.g. Hirshfeld atomic form factors, as well as conventional spherical ones. With the *tsc* input, the full functionality (restraints, constraints, etc.) of the refinement engine *olex2.refine* is available to HAR.



**Fig. 9.** The NoSpherA2 GUI in Olex2 (designed and programmed by F. Kleemiss with the help of H. Puschmann, O. Dolomanov and M. Bodensteiner in a collaboration of OlexSys, Durham, and the Grabowsky group, University of Bern).

(a) Deformation density map of *L*-alanine after HAR in NoSpherA2, isovalue  $\pm 0.15 \text{ e } \text{\AA}^{-3}$ , green=positive, red=negative. ADPs at 50% probability.

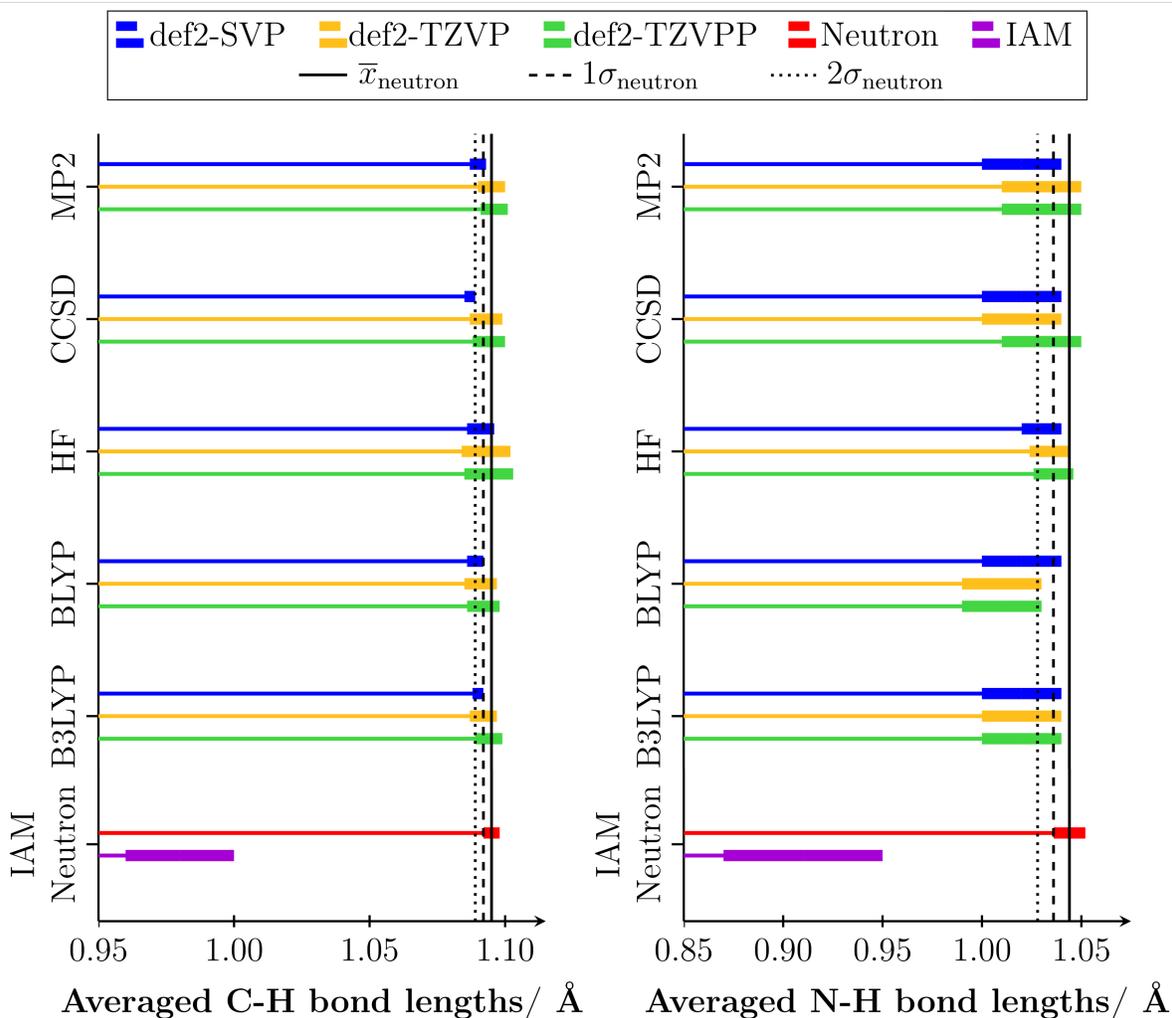
(b) Refinement options for NoSpherA2 in its current version.

The graphical-user interface for the non-spherical refinements based on the tsc format is called NoSpherA2 and presented in Fig. 9. At present, only Hirshfeld atom form factors are implemented as calculated with the program *Tonto*, based on wavefunctions obtained through *Gaussian09*, *Orca* or *Tonto*. A *rigid-atom approximation* as discussed in subsection 3.1 is used for HAR in the current version of NoSpherA2. The *ED* and *refinement steps* can be iterated manually (by ticking the “Update Table” box, Fig. 9(b)) until the desired convergence has been reached. A fully automated iterative version of HAR will be implemented soon. Figure 9(a) shows deformation density of a manually iterated HAR based on a tsc file calculated with *Tonto*, for the molecule *L*-alanine (23 K measurement by Destro *et al.* [81]). Bonds and lone pairs are clearly visible, representing the non-spherical chemical information included in the model that is used for the structural least-squares refinement. Anisotropic displacement parameters of hydrogen atoms and bond distances involving hydrogen atoms are accurately determined.

### **3.9 Multi-determinant HAR**

Despite the concept of HAR having been described in 2008, it has so far only been exploited in combination with the Hartree-Fock (HF) or density functional theory (DFT) methods (particularly, BLYP and B3LYP functionals), but never in combination with explicitly correlated post-Hartree-Fock techniques of quantum chemistry. As is well known, the latter approaches rely on a more sophisticated wavefunction *ansatz* consisting in the expansion of the wavefunction in terms of many Slater determinants. For this reason, at least in principle, the post-HF strategies provide better electron densities that could be potentially exploited in Hirshfeld atom refinements to obtain better structural parameters.

Some of us have started to explore the coupling of HAR with the MP2 (second-order Møller-Plesset) and CCSD (Coupled Cluster with single and double excitations) methods [88]. We have investigated whether results obtained with MP2 and CCSD are significantly different (and better) compared to those resulting from HARs carried out at the HF, BLYP and B3LYP levels. We used the amino acid *L*-alanine, for which we considered i) the X-ray diffraction data collected by Destro *et al.* [81] at 23 K to a resolution of  $1.08 \text{ \AA}^{-1}$  and ii) reference neutron diffraction measurements collected at 23 K to a resolution of  $1.00 \text{ \AA}^{-1}$  [77]. We used the basis-sets def2-SVP, def2-TZVP and def2-TZVPP in combination with the five quantum mechanical methods mentioned above to also study the effects of basis-sets in structural refinements involving correlated methods. Comparisons show (Figure 10) that the use of post-Hartree-Fock strategies within HAR does not lead to atomic positions and ADPs significantly different from those resulting from standard Hartree-Fock or DFT-based HARs. In relation to experimental standard uncertainties, the differences in structural parameters between the different types of refinements are insignificant, at least for *L*-alanine. It seems that for organic molecules nothing can be gained from going beyond the basic HAR protocol described in subsection 3.3. A proper choice of the basis set is more important than the choice of the quantum chemical technique.



**Fig. 10.** Averaged C-H and N-H bond lengths in *L*-alanine as obtained from IAM, neutron and Hirshfeld atom refinements. As indicated in the text, the HARs were performed at different levels of theory (Hartree-Fock, BLYP, B3LYP, MP2 and CCSD) in combination with three different basis-sets (def2-SVP, def2-TZVP and def2-TZVPP)

### 3.10 HAR and powder diffraction

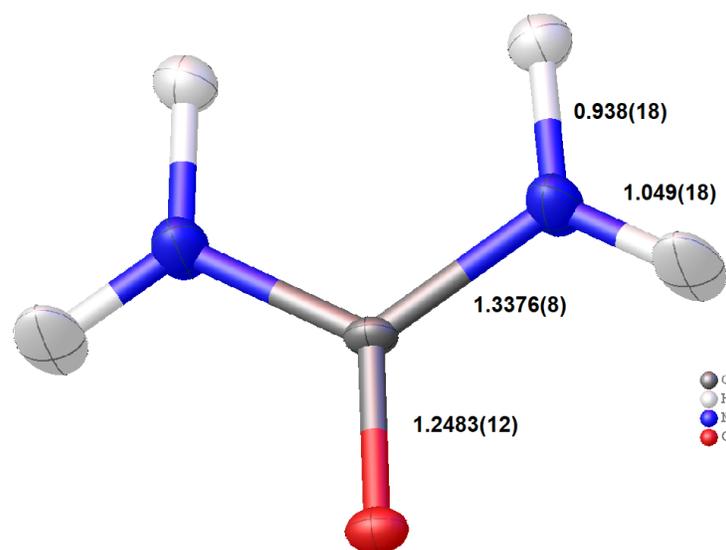
For many materials, structure elucidation by powder X-ray diffraction is the only viable option because it proves difficult to grow large crystals of many materials of technical interest. However, *ab initio* structure determination and three-dimensional structure refinement from powder data are often a challenge [89]:

i) Much of the information on the three-dimensional location of each reciprocal lattice point is lost upon projection onto a single dimension. As a consequence, indexing and geometric reconstruction of the reciprocal lattice is difficult.

ii) Overlap of diffraction lines aggravates the determination of integrated intensities. Rietveld structure refinement [90], although a powerful method that deals with the overlap problem, suffers from the shortcoming that intensity attribution depends to some extent on the model of the crystal structure.

Problem i) can be alleviated by improving the quality of the powder data, i.e. to gather information to as high scattering angles as possible to optimize angular resolution and to reduce background. The Aarhus Vacuum Imaging-Plate Diffractometer (AVID) [91] combined with synchrotron radiation delivers powder data of exceptional quality and to ultra-high resolution. Using powder data collected with this technique, highly accurate experimental electron density information for silicon and diamond were obtained including description of core deformation [92, 93]. More recently, a multipole model of the electron-density distribution of a small molecule, urea, could be refined from powder diffraction data [94].

Improvements for problem ii) can be expected from working with a better model of the crystal structure such as the use of HAR with its non-spherical atomic form factors. Once accurate structure factors for any stage of the structural model are obtained with HAR, an improved separation of lines during Rietveld refinement should be possible without adding extra parameters into the procedure. This method is currently being developed. Fig. 11 shows the results of the first successful HAR of the urea molecule based on AVID powder data. Note that here the intensities have still been extracted using a crystal structure model based on the routine spherical form factors used in IAM, and then HAR was performed once using these intensities. In this preliminary attempt, hydrogen atom ADPs and N-H bond distances could be freely refined and converged without problems, but they are not as accurate as for HAR single-crystal studies yet (compared to N-H = 1.022(3) and 1.018(3) Å from neutron diffraction at 12K [95]). We expect further improvements once the iterative technique is implemented.



**Fig. 11:** Hirshfeld atom refined structure of urea based on powder diffraction data with refined bond distances in Å and anisotropic displacement parameters at the 50% probability level (half of the molecule is symmetry-independent); courtesy of B. Svane and B. B. Iversen, Aarhus University.

#### 4. Introduction to X-ray constrained wavefunction fitting

In quantum mechanics, wavefunction and density matrix are fundamental entities that intrinsically contain all the information about a system. For this reason, the determination of wavefunctions and density matrices from experimental data (and especially from crystallographic measurements) has always been a tantalizing perspective.

The pioneering studies in this field date back to the 1960s when Clinton, Massa and collaborators proposed strategies to extract  $N$ -representable one-electron density matrices from X-ray diffraction data [96-100]. These studies were the starting points for the development of other techniques to obtain “experimental” wavefunctions and density matrices, such as Tanaka’s X-ray Atomic Orbital [40] (XAO) and X-ray Molecular Orbital [101] (XMO) strategies, the Molecular Orbital Occupation Number [102, 103] (MOON) method and all the approaches aiming at reconstructing the diagonal and the off-diagonal parts of the one-electron density matrices by

simultaneously exploiting X-ray diffraction data, magnetic structure factors and inelastic Compton scattering measurements [104-107].

All the methods mentioned in the previous paragraph simply fit wavefunctions or one-particle density matrices to crystallographic data corresponding to an experimental electron density. However, the Gilbert corollary [108] to Coleman's theorem [109] implies that in principle there is an infinite number of density matrices (and consequently wavefunctions) that are compatible with a given electron distribution. Therefore, the simple fitting of density matrices to electron density-related data does not guarantee to obtain the real "experimental" wavefunction or density matrix for the system under consideration, even when rigorous  $N$ -representability constraints are introduced in the computational procedure.

A way to overcome this drawback was offered by Henderson and Zimmermann [110]. According to them, among all the wavefunctions that reproduce a given electron density, we should select the one that minimizes the energy of the system. This idea is at the basis of the X-ray constrained wavefunction (XCW) fitting approach proposed by Jayatilaka in 1998 [21]. Nowadays, XCW can be considered as the most popular strategy to obtain wavefunctions/density matrices from experimental X-ray diffraction measurements. The method was originally developed in the framework of the Restricted Hartree-Fock (RHF) formalism [21-27]. Over the years, it has been gradually extended to other traditional strategies of quantum chemistry, ranging from Density Functional Theory [27] to the multi-determinant spin-coupled approach [62, 63].

A description of the XCW fitting method is the main topic of this section. After reviewing the basic assumptions of the technique (subsection 4.1), we will present the different forms with which the Jayatilaka strategy has been proposed (subsection 4.2) and implemented (subsection 4.3). We will then briefly describe the studies aimed at testing the capabilities of the X-ray constrained wavefunction technique in extracting correlation, polarization and relativistic effects from the electron density (subsection 4.4). Finally, we will discuss open problems and future perspectives of the approach (subsection 4.5).

#### ***4.1 Basic assumptions of the XCW fitting technique***

Applications of the XCW method normally pertain to molecular crystals and consider them as collections of non-interacting molecular units described by wavefunctions that are formally

identical. With this hypothesis and the additional assumption that all the molecular units also correspond to symmetry-unique portions of the crystal unit-cell, the electron density of the unit-cell can be expressed as a sum of molecular electron distributions  $\rho_j(\mathbf{r})$  related to the reference density  $\bar{\rho}_0(\mathbf{r})$  through the application of the unit-cell symmetry operations (see Eq. 13 in subsection 2.2).

This unit cell density is exact only if  $\bar{\rho}_0(\mathbf{r})$  does not result from a computation on the isolated reference crystal-unit. In the XCW fitting techniques, this is fulfilled by searching the wavefunction  $\Psi_0$  and the associated electron density  $\bar{\rho}_0(\mathbf{r})$  that not only minimize the energy of the reference crystal-unit, but that also reproduce experimental structure factor amplitudes within the limit of their experimental uncertainties. This condition implies that the effects of the surrounding crystal field are taken into account in an effective way, even if the X-ray constrained wavefunction computations are isolated-molecule calculations.

The XCW techniques determine the wavefunction  $\Psi_0$  of the reference crystal-unit that minimizes this functional:

$$J[\Psi_0] = \frac{\langle \Psi_0 | \hat{\mathcal{H}}_0 | \Psi_0 \rangle}{\langle \Psi_0 | \Psi_0 \rangle} + \lambda (\chi^2[\Psi_0] - \Delta) \quad (85),$$

the first term represents the energy of the system (with  $\hat{\mathcal{H}}_0$  as the Hamiltonian operator for the reference unit); the second term accounts for the experimental constraints imposed by the experimental X-ray diffraction data;  $\lambda$  is an external multiplier that is adjusted during the calculations and determines the strength of the experimental constraints;  $\Delta$  is the desired agreement between theoretical and observed structure factor amplitudes (usually set equal to 1), which is given by  $\chi^2$ :

$$\chi^2 = \frac{1}{N_r - N_p} \sum_{\mathbf{h}} \frac{(\eta |F_{\mathbf{h}}^{calc}| - |F_{\mathbf{h}}^{exp}|)^2}{\sigma_{\mathbf{h}}^2} \quad (86)$$

$N_r$  is the number of experimental data,  $N_p$  is the number of adjustable parameters,  $\mathbf{h}$  is the triad of Miller indices for the reflection under exam,  $\sigma_{\mathbf{h}}$  is the experimental uncertainty corresponding to the experimental structure factor amplitude  $|F_{\mathbf{h}}^{exp}|$  and  $\eta$  is an overall scale-factor, which multiplies each computed structure factor amplitude  $|F_{\mathbf{h}}^{calc}|$  and is obtained by minimizing  $\chi^2$ .

In all types of XCW computations, the calculated structure factors  $F_h^{calc}$  appearing in the expression of the statistical agreement  $\chi^2$  are obtained by applying Eqs. 35 and 36. In particular, since the analytical form of the time and space averaged electron densities associated with X-ray constrained wavefunctions is the one given by Eq. 22 (i.e. characterized by two-center products of basis functions), the treatment of thermal smearing requires one of the approximations introduced in subsections 2.3 and 2.4 to deal with the lack of thermal factors describing interatomic correlations. For every thermal motion approximation, the thermal factors depend parametrically on the symmetric matrices of the ADPs associated with the different atoms. Therefore, just like the nuclear positions, the ADPs are input parameters that are never optimized during the X-ray constrained wavefunction fitting calculations.

In this regard, it is worth mentioning a useful relationship between the XCW method and the Hirshfeld Atom Refinement (see also subsection 3.7). In a recent study, the XCW and HAR techniques were successfully combined to the so-called X-ray wavefunction refinement (XWR) method [85, 86]. In XWR, HAR represents the first step and deals with the refinement of atomic positions and thermal parameters (i.e., ADPs). These atomic parameters are kept fixed during the subsequent XWC fitting computation. The latter optimizes only the parameters associated with the electronic structure, i.e. the coefficients of the molecular orbitals, to obtain optimal agreement with the measured structure factors.

Validation tests have been performed on a set of high-resolution X-ray diffraction datasets for four amino acids and six tri-peptides. They have shown that the new XWR method is able to provide refined structures and electron densities that are in better agreement with the measured diffraction patterns than those from the multipole model refinements, and also agree with benchmarking periodic DFT calculations. Compared to multipole model refinements, the new strategy led to significant improvements in the description of the electron density distributions in the regions around polar bonds and hydrogen atoms [86].

As already mentioned in subsection 3.7, this coupling of HAR and XCW fitting is iterated in the so-called *total XWR* strategy. In preliminary investigations *total XWR* has led both to better structural parameters and to a better fit to the experimental X-ray structure factor magnitudes.

## 4.2 Different “flavors” of XCW fitting strategies

As indicated in the previous subsection, the XCW fitting technique currently exists in different forms that depend on the choice of the functional form (i.e., the *ansatz*) for the wavefunction  $\Psi_0$  associated with the reference crystal-unit. In this subsection, we will review different “flavors” of the XCW approach currently available. For each of them, we will describe the corresponding wavefunction *ansatz* and point out, when possible, the advantages and the disadvantages of that choice.

As mentioned above, the original XCW method [21] was developed within the Restricted Hartree-Fock (RHF) technique based on the following wavefunction *ansatz*:

$$\Psi_0 = \frac{1}{\sqrt{(2N)!}} \hat{A} \left[ \prod_{i=1}^N \varphi_i \bar{\varphi}_i \right] \quad (87),$$

where  $\hat{A}$  is the usual antisymmetrizer,  $\varphi_i$  is a spin orbital with spatial part  $\varphi_i$  and spin part  $\alpha$  and  $\bar{\varphi}_i$  a spin orbital with spatial part  $\varphi_i$  and spin part  $\beta$ . Starting from Eq. 87, it is possible to show that the minimization of the functional given by Eq. 85 is equivalent to the resolution of the Hartree-Fock-Jayatilaka equations:

$$\hat{F}^{exp} \varphi_i = \epsilon_i \varphi_i \quad (88).$$

$\hat{F}^{exp}$  is a modified Fock operator that can be expressed like this:

$$\hat{F}^{exp} = \hat{F} + \lambda \sum_{\mathbf{h}} K_{\mathbf{h}} \operatorname{Re}\{F_{\mathbf{h}}^{calc}\} \hat{I}_{\mathbf{h},R} + \lambda \sum_{\mathbf{h}} K_{\mathbf{h}} \operatorname{Im}\{F_{\mathbf{h}}^{calc}\} \hat{I}_{\mathbf{h},C} \quad (89)$$

where  $\hat{F}$  is the usual Fock operator of quantum chemistry,  $\hat{I}_{\mathbf{h},R}$  and  $\hat{I}_{\mathbf{h},C}$  are the real and imaginary parts, respectively, of the structure factor operator  $\hat{I}_{\mathbf{h}}$ , which is defined as follows

$$\hat{I}_{\mathbf{h}} = \sum_{j=1}^{N_m} e^{i2\pi(\mathbf{Q}_j \mathbf{r} + \mathbf{q}_j) \cdot (\mathbf{B}\mathbf{h})} = \hat{I}_{\mathbf{h},R} + i \hat{I}_{\mathbf{h},C} \quad (90).$$

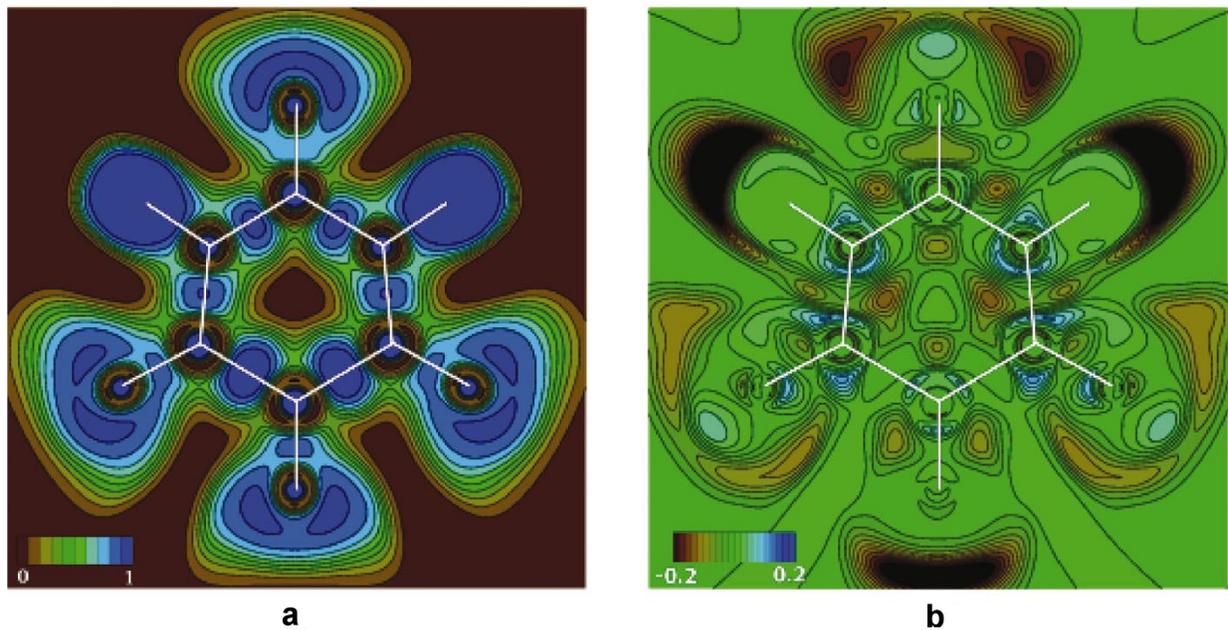
$K_{\mathbf{h}}$  is a multiplicative factor:

$$K_{\mathbf{h}} = \frac{2 \eta}{N_r - N_p} \frac{\eta |F_{\mathbf{h}}^{calc}| - |F_{\mathbf{h}}^{exp}|}{\sigma_{\mathbf{h}}^2 |F_{\mathbf{h}}^{calc}|} \quad (91).$$

For the sake of completeness, and in analogy with section 2 of this chapter,  $\mathbf{B}$  still represents the reciprocal lattice matrix, while  $\{\mathbf{Q}_j, \mathbf{q}_j\}_{j=1}^{N_m}$  is the set of the  $N_m$  symmetry operations of the unit-cell.

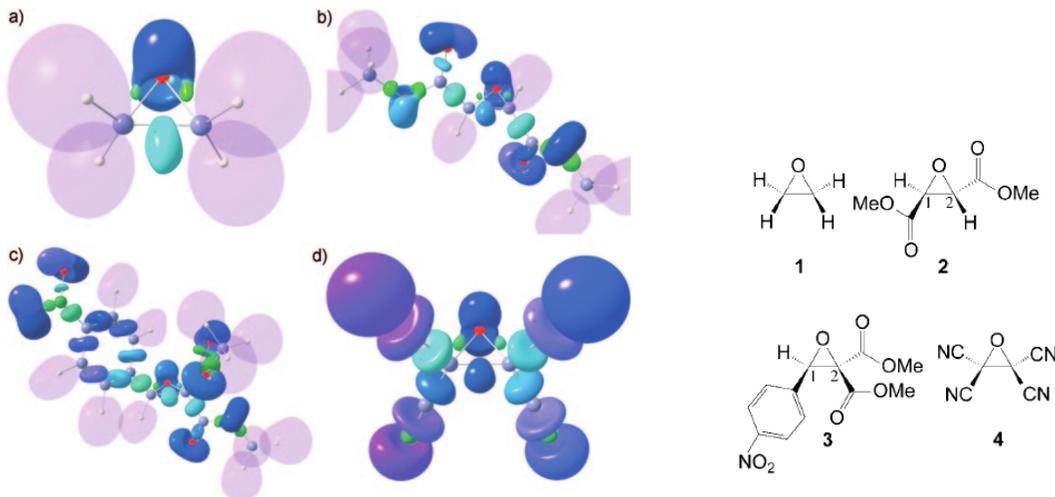
The XC-RHF approach has been straightforwardly extended to the unrestricted case [111], to the formalism of density functional theory [27] (BLYP and B3LYP functionals) and to relativistic methods (particularly, the second-order Douglas-Kroll-Hess [111] and the Infinite-Order Two-Component [75] approaches). The Hartree-Fock, DFT and relativistic versions of the X-ray constrained wavefunction approach are currently implemented in the software *Tonto* [50] and are, at the moment, undoubtedly the most widely used types of XCW calculations in quantum crystallography. “Experimental” XC-HF and XC-DFT wavefunctions/electron densities have been analyzed to obtain useful (and traditional) chemical information with Quantum Chemical Topology (QCT) strategies (such as, the traditional Quantum Theory of Atoms in Molecules [112] (QTAIM), the Electron Localization Function [113] (ELF), the Electron Localizability Indicator [114, 115] (ELI) and the Localized Orbital Locator [116] (LOL)), the Natural Bond Orbitals [117] (NBO) method and the Natural Resonance Theory [118] (NRT) [66, 85, 119-122].

As a first example, Fig. 12a shows a plot of the ELF associated with the X-ray constrained RHF wavefunction (Dunning DZP basis-set) of the alloxan molecule as obtained by Jayatilaka and Grimwood [119]. Fig. 12b, shows the differences between the ELF corresponding to the “experimental” wavefunction and the ELF resulting from the “gas-phase” Hartree-Fock calculation. It indicates the capability of the X-ray constrained wavefunction technique in taking into account the perturbations introduced by the crystal environment on the isolated molecules (see subsection 4.4 for more details).



**Fig. 12.** a) Plot of the ELF for the X-ray constrained Hartree-Fock wavefunction of alloxan (contours at 0.1 increments); b) plot of the difference between the ELF for the X-ray constrained Hartree-Fock wavefunction and the isolated-molecule Hartree-Fock wavefunction of alloxan (contours at 0.02 increments). [Reproduced with permission of the International Union of Crystallography](#) from reference [119].

Fig. 13 displays the ELI-D [115] localization domain representations for a series of epoxide compounds obtained by Grabowsky and coworkers through XC-RHF/cc-pVDZ computations. These plots provide insight into the electronic structure of the epoxide rings [120]. In agreement with the Förster-Coulson-Moffitt model, the ELI-D plots show externally bent bonds and, furthermore, they allowed a clear rationalization of substituent and crystal effects on the series of acceptor-substituted epoxide derivatives.



**Fig. 13.** Experimental ELI-D localization domain representations for the series of epoxide compounds studied by Grabowsky and collaborators [120], with valence basins for H atoms in transparent mode. Color code makes domains belonging to different basins distinguishable from each other; MolIso [123] graphics. Reproduced from reference [120] with permission of John Wiley & Sons, Inc.

In order to directly extract X-ray constrained wavefunctions close to the traditional chemical perception without the *a posteriori* application of chemical bonding analyses, several XCW fitting strategies based on the concept of Extremely Localized Molecular Orbitals (ELMOs) [70, 124-129] have also been proposed. The first one was the X-ray constrained ELMO (XC-ELMO) technique [56-59], which requires the preliminary definition of a localization scheme that subdivides the system under examination into subunits (i.e., atoms, bonds and functional groups) that may overlap. The fragmentation automatically defines the ELMOs for each subunit as linear combinations of only those atomic basis functions centered on the atoms that belong to the fragments. In other words, for the  $\gamma$ -th ELMO of the  $i$ -th fragment, we can write:

$$\varphi_{i\gamma}(\mathbf{r}) = \sum_{\mu=1}^{M_i} C_{i\mu,i\gamma} \chi_{i\mu}(\mathbf{r}) \quad (92)$$

with  $M_i$  as the total number of basis functions in the local basis-set  $\beta_i = \{\chi_{i\mu}(\mathbf{r})\}_{\mu=1}^{M_i}$  of the subunit. Consequently, the *ansatz* for the wavefunction of the reference crystal-unit becomes:

$$\Psi_0 = \frac{1}{\sqrt{(2N)! \det[\mathbf{S}]}} \hat{A} \left[ \prod_{i=1}^f \prod_{\gamma=1}^{n_i} \varphi_{i\gamma} \bar{\varphi}_{i\gamma} \right] \quad (93)$$

where  $\det[\mathbf{S}]$  is the determinant of the overlap-matrix of the occupied ELMOs (which accounts for the intrinsic non-orthogonality of ELMOs associated with different fragments),  $f$  is the total number of fragments,  $n_i$  is the number of occupied ELMOs for the  $i$ -th subunit.

Using the wavefunction *ansatz* given by Eq. 93, the minimization of the Jayatilaka functional (see Eq. 85) reduces to solving self-consistently the following modified Hartree-Fock equations for each fragment:

$$\hat{F}^{i,exp} \varphi_{i\gamma} = \epsilon_{i\gamma} \varphi_{i\gamma} \quad (94)$$

where the modified ‘‘experimental’’ Fock operator for the  $i$ -th subunit is:

$$\begin{aligned} \hat{F}^{i,exp} = & (1 - \hat{\rho} + \hat{\rho}_i^\dagger) \hat{F} (1 - \hat{\rho} + \hat{\rho}_i) + \lambda \sum_h K_h \operatorname{Re}\{F_h^{calc}\} (1 - \hat{\rho} + \hat{\rho}_i^\dagger) \hat{I}_{h,R} (1 - \hat{\rho} + \hat{\rho}_i) \\ & + \lambda \sum_h K_h \operatorname{Im}\{F_h^{calc}\} (1 - \hat{\rho} + \hat{\rho}_i^\dagger) \hat{I}_{h,C} (1 - \hat{\rho} + \hat{\rho}_i) \quad (95), \end{aligned}$$

with  $\hat{\rho}$  as the global density operator (which depends on all the occupied ELMOs of the system),  $\hat{\rho}_i$  as the density operator for fragment  $i$  (which depends only on the occupied ELMOs of the subunit),  $\hat{\rho}_i^\dagger$  as the adjoint of  $\hat{\rho}_i$ , and  $K_h$  as a multiplicative factor already defined by Eq. 91.

The XC-ELMO method allows the extraction of X-ray constrained molecular orbitals strictly localized on small molecular units and, consequently, close to the traditional chemical perception, which generally considers molecules as built from atoms, bonds or molecular fragments such as functional groups. In addition, the X-ray constrained wavefunction fitting approaches capture correlation and crystal field effects on the electron density (see subsection 4.4). Thus the XC-ELMO strategy might be used in the future also to extended the recently constructed ELMO libraries [78, 79, 130] by including extremely localized molecular orbitals that intrinsically take the crystal environment into account.

The more recent X-ray constrained ELMO Valence Bond (XC-ELMO-VB) technique is an alternative strategy [60, 61]. It was introduced to determine the weights of resonance structures in molecular systems from X-ray diffraction. In this case, the *ansatz* for the wavefunction of the reference crystal unit is

$$\Psi_0 = \sum_i C_i \Psi_{ELMO,i} \quad (96)$$

where the functions  $\{\Psi_{ELMO,i}\}$  are ELMO wavefunctions resulting from preceding unconstrained ELMO calculations based on localization schemes corresponding to the different resonance structures.

While the pre-computed ELMO wavefunctions appearing in Eq. 96 are kept frozen, the coefficients  $\{C_i\}$  are obtained through minimization of the following functional:

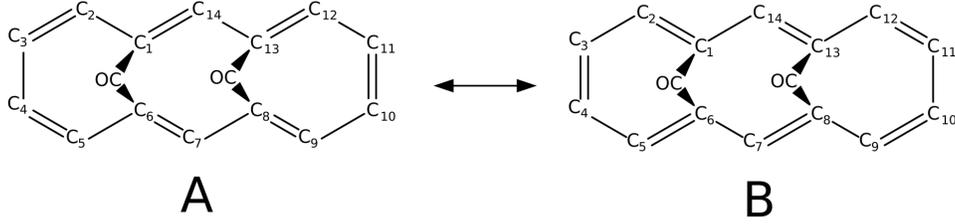
$$J[\mathbf{C}] = E[\mathbf{C}] + \lambda (\chi^2[\mathbf{C}] - \Delta) \quad (97)$$

where  $E$  is the energy associated with the XC-ELMO-VB wavefunction and  $\mathbf{C}$  indicates the functional dependence of  $J$ ,  $E$  and  $\chi^2$  on the coefficient  $\{C_i\}$ . Note that the coefficients  $\{C_i\}$  are not the weights of the resonance structures since the Slater determinants  $\{\Psi_{ELMO,i}\}$  are non-orthogonal. The weights are obtained by determining the Chirgwin-Coulson coefficients [131], which, in this case, are defined as:

$$K_i = |C_i|^2 + \sum_{j \neq i} C_i C_j S_{ij} \quad (98)$$

where  $S_{ij}$  is the overlap integral between  $\Psi_{ELMO,i}$  and  $\Psi_{ELMO,j}$ . The XC-ELMO-VB method is already sufficiently well developed for practical applications and can be considered as the first example of a multi-determinant XCW approach.

After some preliminary test calculations [60], the XC-ELMO-VB strategy has been fruitfully exploited to interpret the results of a recent X-ray charge density study on the syn-1,6:8-13-biscarbonyl[14]annulene (BCA) molecule [132]. At ambient pressure, resonance structures A and B of BCA (see Fig. 14) have been found to be almost equivalent (i.e., they have almost the same Chirgwin-Coulson coefficients), at high pressures resonance structure A becomes clearly predominant [61]. This represents a confirmation of the partial suppression of aromaticity (see also subsection 5.3).



**Fig. 14.** Resonance structures A and B of BCA. Hydrogen atoms are omitted for the sake of clarity.

It is worth bearing in mind that the XC-ELMO-VB technique relies on a wavefunction *ansatz* consisting of a linear combination of pre-determined Slater determinants that remain unchanged during the computations. To overcome this limitation, the new X-ray constrained spin-coupled (XCSC) approach has been introduced [62, 63]. As the name suggests, this strategy is an extension of the X-ray constrained wavefunction fitting technique making use of the spin-coupled method [133, 134]. The latter is a Valence Bond approach that *i*) provides more insights into the spatial arrangements of the electron clouds around atoms from the spin-coupled orbitals; and *ii*) allows to determine the relative importance of resonance structures from the weights obtained for all the possible spin-coupling modes.

In the XCSC case, the *ansatz* for the wavefunction  $\Psi_0$  associated with the reference crystal unit has the analytical form of a spin-coupled wavefunction for a system of  $N$  electrons in the spin-state  $(S, M)$ :

$$\begin{aligned} \Psi_0 &= \mathcal{N} \sum_{k=1}^{f_S^{Nv}} c_{S,k} \psi_{S,M;k}^N = \\ &= \mathcal{N} \sum_{k=1}^{f_S^{Nv}} c_{S,k} \hat{A} [\phi_1^c \bar{\phi}_1^c \dots \phi_j^c \bar{\phi}_j^c \dots \phi_{N_1}^c \bar{\phi}_{N_1}^c \varphi_1 \varphi_2 \dots \varphi_{N_v} \Theta_{S,M;k}^{Nv}] \quad (99) \end{aligned}$$

where  $\mathcal{N}$  is the overall normalization constant for the wavefunction,  $S$  and  $M$  are the quantum numbers associated with the spin operators  $\hat{S}^2$  and  $\hat{S}_z$ , respectively,  $\hat{A}$  is the usual antisymmetrizer and  $\{c_{S,k}\}$  the spin-coupling coefficients of the  $f_S^{Nv}$  linearly independent spin-coupled structures  $\{\psi_{S,M;k}^N\}$  corresponding to the linearly independent spin-eigenfunctions  $\{\Theta_{S,M;k}^{Nv}\}$ . As in any unconstrained spin-coupled method, the orbitals appearing in Eq. 99 can be

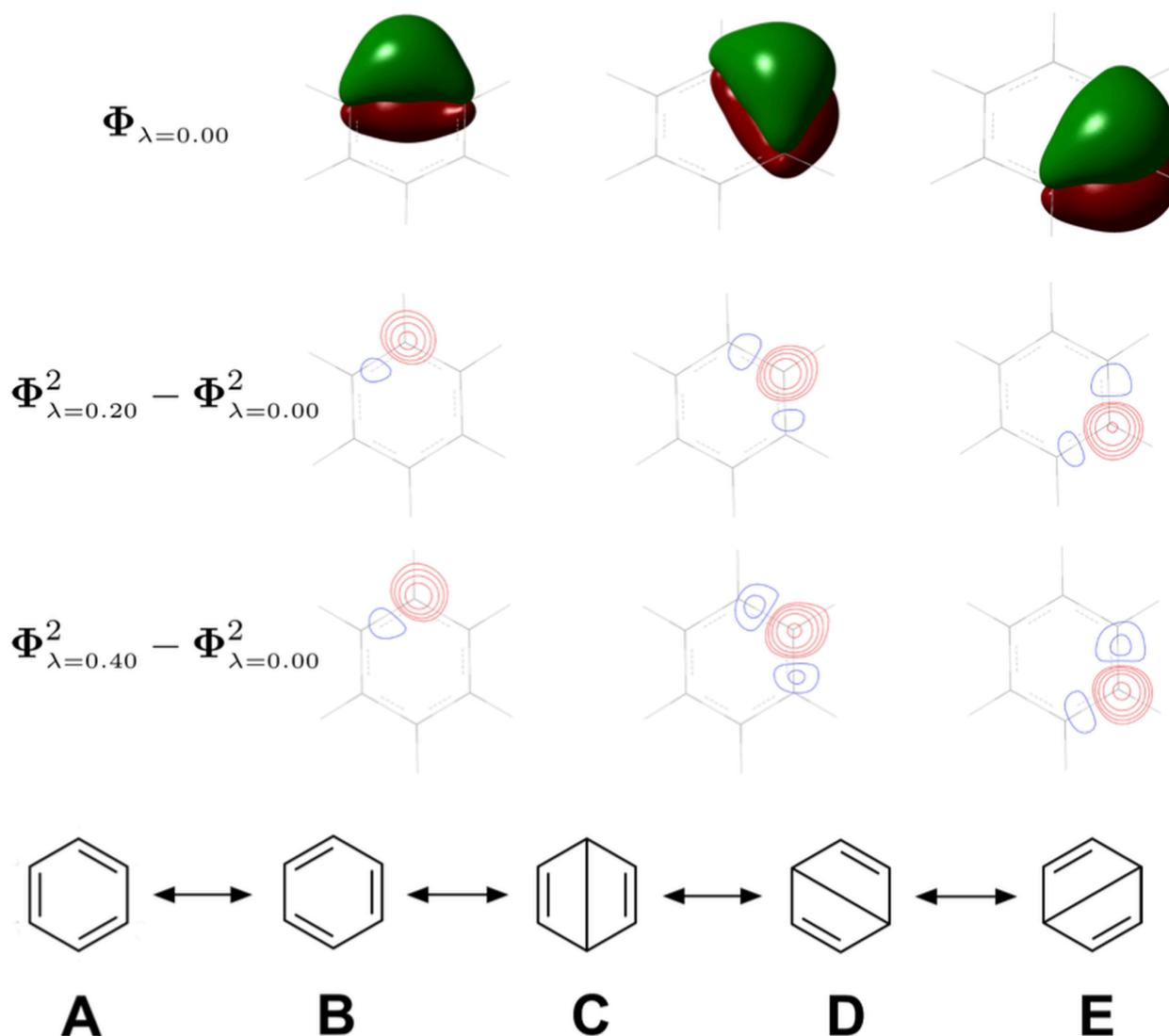
subdivided into two groups: *i*) the frozen, doubly occupied “core orbitals”  $\{\phi_i^c\}_{i=1}^{N_1}$ , which can be obtained through preliminary unconstrained Hartree-Fock or X-ray constrained Hartree-Fock calculations and which describe the subset of  $2N_1$  core (or inactive) electrons; *ii*) the “spin-coupled orbitals”  $\{\varphi_i\}_{i=1}^{N_v}$ , which describe the  $N_v$  valence (or active) electrons and which are the only ones to be optimized, thereby limiting the computational cost of the XCSC method.

Given the analytical form of the new wavefunction *ansatz* (see Eq. 99), the new XCSC method essentially consists in determining the sets of spin-coupled orbitals  $\{\varphi_i\}_{i=1}^{N_v}$  and spin-coupling coefficients  $\{c_{S,k}\}$  that minimize the Jayatilaka functional shown in Eq. 85. In this case, again, the spin coupling coefficients  $\{c_{S,k}\}$  do not directly provide the weights of the spin-coupled structures  $\{\psi_{S,M;k}^N\}$  due to the non-orthogonality of the latter. To obtain the weights, it is necessary to resort to the Chirgwin-Coulson coefficients, which, in this situation, can be expressed as

$$w_{S,k} = |c_{S,k}|^2 + \sum_{j \neq k} c_{S,k} c_{S,j} S_{kj} \quad (100),$$

where  $S_{kj}$  is the overlap integral between spin-coupled structures  $\psi_{S,M;k}^N$  and  $\psi_{S,M;j}^N$ .

The X-ray constrained spin-coupled technique was applied to X-ray diffraction data of benzene and the results compared to gas-phase results [62]. The differences in the spin-coupled orbitals (see Fig. 15), the resonance structure weights (see Table 1) and the electron density distributions were found to be non-negligible [62]. Once again, this revealed the potential of the XCSC strategy for capturing crystal field effects on the electronic structure of molecules. Further investigations have also shown that the use of XC-RHF molecular orbitals to describe the “inactive” electrons leads to better descriptions of the electronic structure and, consequently, of the electron density [63]. This is definitely the procedure to follow in any XCSC calculations and, for this reason, the new X-ray constrained spin-coupled technique can be undoubtedly defined as the first X-ray constrained post-Hartree-Fock method.



**Fig. 15.** First row: three-dimensional plots of the unconstrained spin-coupled orbitals of benzene ( $0.06 \text{ e/bohr}^3$  isovalues). Second row: two-dimensional plots of the differences between the square moduli of the X-ray constrained ( $\lambda = 0.20$ ) and unconstrained Spin-Coupled orbitals. Third row: two-dimensional plots of the differences between the square moduli of the X-ray constrained ( $\lambda = 0.40$ ) and unconstrained Spin-Coupled orbitals. Fourth row: spin-coupled resonance structures of benzene (A-B: Kekulé resonance structures; C-E: Dewar resonance structures). For each two-dimensional plot, the contours are drawn at  $\pm 1 \cdot 10^{-3} \text{ e/bohr}^3$  and at  $\pm 2, 4, 8 \cdot 10^n \text{ e/bohr}^3$  (with  $n$  as an integer ranging from -3 to 0) in a plane parallel to and  $0.5 \text{ \AA}$  above the one defined by the carbon atom on which the orbital is mainly localized and the two carbon atoms bonded to it. Red and blue contours indicate positive and negative values, respectively. Reproduced from Figs 1 and 3 of reference [62] with permission of John Wiley & Sons, Inc.

**Table 1.** Chirgwin-Coulson weights (in %) of the benzene spin-coupled resonance structures, as obtained from unconstrained ( $\lambda = 0.0$ ) and X-ray constrained Spin-Coupled calculations (A and B: Kekulé resonance structures; C, D, and E: Dewar resonance structures; see fourth row of Fig. 15). Reprinted from reference [62] with permission of John Wiley & Sons, Inc.

Resonance Structure	Chirgwin-Coulson weights		
	$\lambda = 0.0$	$\lambda = 0.20$	$\lambda = 0.40$
A	40.45	39.13	38.22
B	40.45	39.17	38.26
C	6.34	5.90	5.63
D	6.23	6.79	6.76
E	6.53	9.01	11.12

### 4.3 Current implementations of the different XCW approaches

The most popular software to perform X-ray constrained wavefunction fitting analysis is certainly the program *Tonto* [50], already introduced in subsection 3.1. In fact, this program was initially conceived for the implementation of the X-ray constrained wavefunction strategy and only later it was extended to perform Hirshfeld Atom Refinements. Nowadays, *Tonto* can be considered as a complete object-oriented system for computational chemistry and quantum crystallography. Apart from other functionalities, it performs X-ray constrained wavefunction fitting at the Hartree-Fock (restricted and unrestricted) and DFT (BLYP functional) levels of theory with a large variety of traditional basis sets of quantum chemistry. The relativistic extensions of the XCW technique at Hartree-Fock and DFT levels (second-order Douglas-Kroll-Hess and Infinite-Order Two-Component approaches) are also implemented in *Tonto*. As already discussed in subsection 3.7, the *Tonto* software has recently been interfaced with the graphical user interface *lamaGOET* [83] that allows preparing *stdin* input files for XCW computations.

The other types of X-ray constrained wavefunction fitting techniques discussed in the previous subsection (i.e., XC-ELMO, XC-ELMO-VB and XCSC methods) are not implemented in *Tonto*, but in separate stand-alone programs that were originally developed to perform the analogous unconstrained quantum chemistry calculations. The XC-ELMO method has been implemented by modifying a development version of the GAMESS-UK quantum chemistry package [135], in which Fornili and coworkers [124] originally introduced the Extremely Localized Molecular Orbital (ELMO) approach [70]. The XC-ELMO strategy has been introduced in GAMESS-UK by programming both the self-consistent resolution of Eq. 94 and a quasi-Newton procedure that enables to determine the coefficients of the ELMOs that directly minimize the functional given by Eq. 85. The input files for the XC-ELMO calculations are completely analogous to the ones for ELMO computations, with the only exception that some extra crystallographic information must be provided (unit-cell dimensions, symmetry operations, anisotropic displacement parameters and, obviously, the X-ray structure factors amplitudes to be used as external constraints).

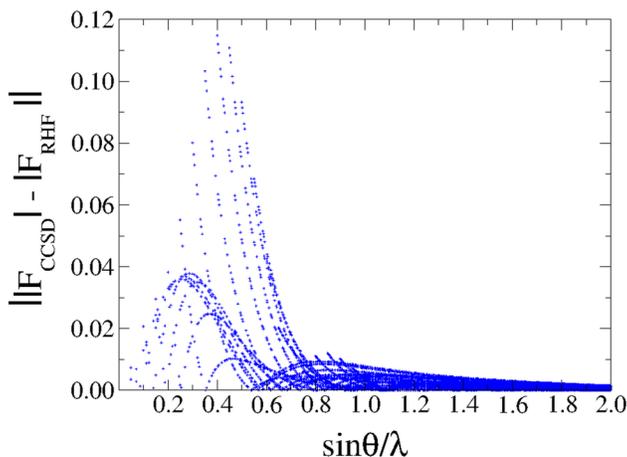
The more recent X-ray constrained spin-coupled technique has also been implemented by modifying an existing spin-coupled program [136]. A Newton-Raphson procedure based on the so-called “quadratic hill climbing” algorithm [137] simultaneously determines the spin-coupled orbitals  $\{\varphi_i\}_{i=1}^{N_v}$  and the spin-coupling coefficients  $\{c_{S,k}\}$  (see Eq. 99) that minimize the usual Jayatilaka functional (see again Eq. 85). Also, in this case, the input files for the XCSC calculations are practically identical to those for the unconstrained spin-coupled computations, except for the necessary crystallographic information.

#### ***4.4 Meaning of the differences between fitted and non-fitted wavefunctions***

Although the X-ray constrained wavefunction fitting calculations are formally single-molecule computations, the external X-ray diffraction constraints reveal effects that cannot be seen in corresponding gas-phase calculations. In other words, the XCW computations capture correlation, polarization and relativistic effects. Several studies that have proven this will be briefly reviewed.

First of all, also prompted by a preliminary study conducted by Jayatilaka *et al.* [138], Genoni and coworkers carried out X-ray constrained Hartree-Fock calculations by exploiting structure

factors resulting from highly-correlated (e.g., Coupled Cluster) quantum chemistry computations for a series of molecules in gas-phase (nitrogen molecule, cyanide anion, water, urea, benzene, glycine) to understand if the XCW fitting method really enables to capture correlation effects on the electron density [139]. The obtained XCW electron densities were compared to an uncorrelated (i.e., Hartree-Fock) and to the initial correlated (i.e., Coupled Cluster) electron distributions. It was observed that correlation effects are increasingly captured if: *i*) the external multiplier  $\lambda$  becomes quite large (i.e., larger than the values that it assumes in usual XCW calculations with experimental data as constraints); *ii*) the high-angle reflections, which contain less information on the electron correlation (see Fig. 16), are not considered. In fact, since the high-angle reflections are much more numerous than the low-angle ones, the former dominate in the Jayatilaka functional (Eq. 85) and reduce the influence of the low-angle data, namely of those data with the largest CCSD-HF differences (see again Fig. 16). In other words, by including the high-angle structure factors, the XCW fitting calculation would be biased towards the unconstrained Hartree-Fock wavefunction [139].

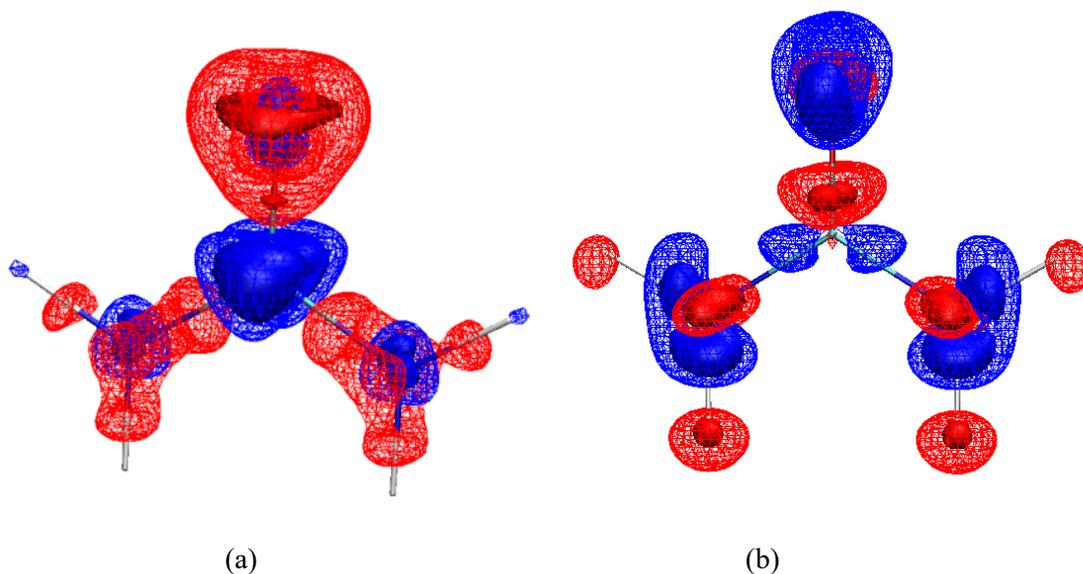


**Fig. 16.** Absolute differences between the gas-phase CCSD and the gas-phase RHF structure factor amplitudes of the nitrogen molecule (basis-set 6-311++G(2d,2p)) as a function of the resolution  $\sin\theta/\lambda$ .

The capability of the Jayatilaka method to capture correlation and polarization effects due to the crystal environment was also demonstrated in a more recent study [140]. The density difference plots in Fig. 17 illustrate the effects of both electron correlation (a) and polarization (b) for the molecule urea. Correlation and polarization effects were extracted by XCW fitting on theoretical structure factors obtained through Coupled Cluster calculations at CCSD level and by means of

periodic boundary computations, respectively. The difference between the non-fitted isolated HF wavefunction and the one fitted against CCSD or periodic-boundary structure factors represents the pure correlation or polarization effect. The shapes of the effects are different in Figs 17 (a) and (b), but the effect is clearly present in the core as well as in the valence regions of the atoms. 0.27 electrons are redistributed due to electron correlation, of which 94.3% can be recovered by XCW fitting. 0.25 electrons are redistributed due to polarization, of which 94.0% can be recovered. This implies that electron correlation and polarization can be fitted nearly completely if a high enough  $\lambda$  value is chosen. However, for both effects, the results are strongly resolution dependent as explained above (Figure 16) for the effect of electron correlation.

The problem of relativistic effects is non-negligible for compounds containing heavy elements. Bučinský *et al.* [75] performed a detailed study on biphenyl mercury ( $\text{HgPh}_2$ ) and tri-phenyl bismuth ( $\text{BiPh}_3$ ). They found that a non-relativistic XC-Hartree-Fock method captures both relativistic and correlation effects to some extent. Relativistic effects in the outer core and the valence shell around the heavy element are at least as important as correlation effects. This was corroborated by a detailed comparison of theoretical structure factors of the model compound  $\text{Hg}(\text{CN})_2$  calculated at different levels of theory [76].



**Fig. 17.** Difference electron densities of urea at the HF/pob-TZVP level of theory at  $0.7 \text{ \AA}^{-1}$  resolution: (a) XCW densities at  $\lambda_{\text{HF}} = 3.2$  (max. value) and  $\lambda_{\text{HF}} = 0.0$  fitted against structure factors calculated with the CCSD method (signifying *electron correlation*); (b) the XCW

densities at  $\lambda_{\text{HF}} = 1.6$  (max. value) and  $\lambda_{\text{HF}} = 0.0$  fitted against structure factors from a periodic-boundary calculation at the same level of theory (signifying *polarization/crystal effects*). Isosurface values: +0.005/+0.0025 a.u. (blue) and -0.005/-0.0025 a.u. (red).

#### ***4.5 Open problems and future perspectives***

It may seem that the X-ray constrained wavefunction strategy is a mature technique, routinely used in modern quantum crystallography studies. However, there are still open problems and thus room for improvements.

The most important criticism of the Jayatilaka approach concerns the extent to which the experimental measurements affect the refined wavefunctions and electron densities or, on the contrary, the extent to which the underlying quantum chemistry Hamiltonian influences the final results. Although it is difficult to provide a conclusive answer to this question, the work to date shows that X-ray constrained wavefunction techniques capture perturbations of the wavefunction by the crystal environment and by other effects not included in the electronic Hamiltonian of the quantum mechanical theory used (e.g. correlation and relativistic effects). Some examples of these studies will be discussed also in section 5.

The problem discussed above boils down to the determination of the correct value for the external multiplier  $\lambda$ , which controls the relative importance of the theoretical and experimental information in the XCW computations. As indicated in the seminal papers of the technique [21, 22], one should ideally choose the value of  $\lambda$  for which  $\chi^2 = 1$ . In that situation, the calculated structure factor amplitudes would be, on average, within one standard deviation of the corresponding experimental values. However, reaching statistical agreement is not always possible, since it is only meaningful if systematic measurement errors can be adequately corrected and the experimental standard uncertainties adequately estimated. Several alternative criteria have been proposed: from halting the XCW procedure when the calculations stop converging [141], or when the  $R_w(F)$  value approaches the one of a corresponding multipole model refinement [142], or when a minimum in the curve of free  $\chi^2$  over  $\lambda$  is reached [85], to empirical rules [57] based on the incremental ratio of  $\chi^2$  with respect to  $\lambda$  and on the variation of the X-ray constrained energy compared to the unconstrained value. Along with the XCSC

method, a criterion based on the second derivative  $\frac{\partial \chi^2}{\partial \lambda^2}$  was also adopted, with the suggestion of stopping the calculations when a clear inflection point is observed in the curve representing the statistical agreement as a function of  $\lambda$  [63]. Notwithstanding all these efforts, the choice of the correct  $\lambda$  value in the XCW calculations remains an open problem that deserves more detailed investigations. A clear answer to this problem may well increase the popularity of the technique and make it a routine tool for charge density studies.

Future opportunities for the XCW approach are plentiful. The way indicated by the recent X-ray constrained spin-coupled method [62, 63] may be pursued further by devising new multi-determinant (post XC-Hartree-Fock) strategies able to provide more accurate fittings of the experimental diffraction measurements. The XCW machinery may be extended to other types of experimental data leading to new joint-refinement techniques, similar to those already introduced in the multipolar methods [143-145] or those based on a wavefunction model [107]. For example, the recently developed XCSC technique could be extended by simultaneously exploiting X-ray diffraction and polarized neutron diffraction data as constraints. The resulting spin densities would be quantum mechanically rigorous and could be analyzed by means of proper quantum chemical topology techniques [146-148]. Moreover, the present X-ray constrained wavefunction techniques mainly work for molecular crystals but cannot be successfully applied to extended systems (e.g., Metal Organic Frameworks). It will therefore be necessary to develop XCW strategies based on periodic *ab initio* calculations (see subsection 4.4). Finally, given the capability of the XCW approach in capturing electron correlation effects on the electron density, we also envisage using the Jayatilaka method for studying experimental exchange-correlation potentials and developing new exchange-correlation functionals for density functional theory [140].

## 5. Chemical applications of quantum crystallography

For studies of molecular structure, chemists usually rely on routine X-ray crystallographic refinement based on the Independent Atom Model (IAM). As described in subsection 2.1, the IAM assumes that, for all atoms, all electrons are distributed spherically around their nuclei. This model can neither account for electronic redistributions due to chemical bonding nor localize

hydrogen atoms precisely and accurately. However, both of these points are required for a meaningful chemical analysis of a crystal structure.

A–H bond lengths are typically 0.1 Å too short (A is any atom) because chemical bonding shifts the entire electron density associated with the H atom, i.e. its single valence electron, towards the A–H bonding region. Since X-rays are diffracted by the electrons, not the nuclei, shortened A–H bond lengths result as artifacts.

In contrast, neutrons are diffracted by the atomic nuclei. Therefore, neutron diffraction does not suffer from the above shortcomings and is thus the most reliable and widely used technique for determining hydrogen atom positions in crystals accurately. However, single crystal neutron diffraction experiments require large crystal sizes ( $\sim 1 \text{ mm}^3$ ) compared to routine X-ray experiments ( $\sim 100 \text{ }\mu\text{m}^3$ ). In addition, the few neutron diffraction facilities available worldwide usually ask for specific beam time proposals, often a time-consuming and expensive endeavour. Of the one million or more crystal structures deposited in the CSD only around 0.2% are based on neutron diffraction experiments.

A good deal of the shortcomings associated with X-ray and neutron diffraction can be overcome by using quantum crystallographic methods. For locating hydrogen atoms, Hirshfeld atom refinement (HAR) is especially useful; it requires only medium resolution laboratory X-ray data (0.6–0.8 Å) for reliable structural refinements. Subsection 5.1.1 summarizes the HAR method and presents many examples of accurate and precise localization of hydrogen atoms.

The deformation of atomic electron densities due to chemical bonding has been studied extensively with the so-called charge density multipole model (MM) presented in more detail [in chapter XXX of this book \(Overgaard, details to be completed by the editors\)](#). Multipole modeling is a powerful tool to probe chemical bonding features of the electron density. It has nevertheless some shortcomings: (i) MM produces an electron density, not a wavefunction; hence, much of the chemical information inherent in the electronic wavefunction, the corresponding density matrix or the orbitals is lost; (ii) the MM refinement is based on representing the electron density with (angular) spherical harmonic and radial exponential functions; the latter are often insufficiently flexible and therefore the partitioning of electron density in the intermolecular space may not be accurate.

To overcome these shortcomings, methods such as X-ray wavefunction refinement (XWR) or X-ray constrained wavefunction (XCW) fitting have been developed. In essence, these methods find what has been called an “experimental wavefunction”, a term that may sound like an oxymoron. It seems justified though, because an “experimental wavefunction” minimizes total energy and simultaneously provides a charge density - its square - which best reproduces the experimental diffraction data. These methods are described and illustrated in section 4 of the present chapter.

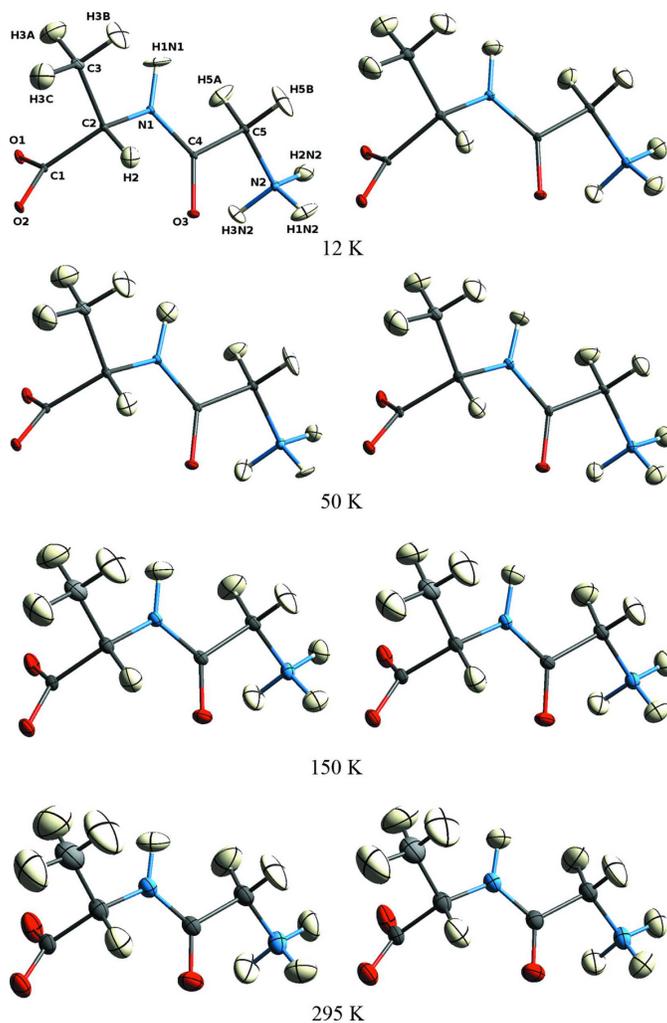
Given an experimental wavefunction and the appropriate operators, the electron density distribution and derived physical properties can be calculated. The wavefunction and the electron density contain polarization, electron correlation and relativistic effects (see subsection 4.4). Derived properties include fundamental physical quantities such as experimental exchange-correlation potentials, polarizabilities and susceptibilities, refractive indices etc. More specifically, some descriptors of chemical bonding are only accessible from wavefunctions. Two examples are: (i) Bond indices, e.g. Roby-Gould bond indices (RGBIs) [149], and (ii) electron localization indices, e.g. the Electron Localizability Indicators (ELI) [114, 115]. The Roby-Gould bond index measures the degree of covalent and ionic bonding by applying the Roby operator  $R$  and ionic operator  $I$  to the X-ray constrained wavefunctions. The sharing of electronic population between a pair of atoms estimates covalency, and the transfer of electron population indicates ionicity. The total bond index or effective bond order is the Pythagorean addition of ionic and covalent indices. Similarly, the descriptor ELI indicates the probability of same spin electron pairing in space; a large value of ELI-D (a variant of ELI) corresponds to a low pair density [115]. Note that Lewis formulae refer to opposite-spin electron pairs and ELI is defined via a same spin-spin electron pairing, but it has been discussed that, at least for closed-shell systems, these two views coincide [121]. Examples outlined later in this chapter exploit X-ray constrained wavefunctions to study chemical bonding situations using bonding descriptors such as RGBI and ELI-D, among others.

## ***5.1 Applications of HAR***

### ***5.1.1 Hydrogen atom treatment in HAR: accuracy and precision compared to results from neutron diffraction***

As also explained in subsection 3.1, the HAR procedure starts from the atomic coordinates obtained through an IAM refinement, calculates a wavefunction and its electron density, and separates the density into non-spherical atomic electron densities. These are Fourier transformed into non-spherical atomic form factors and introduced into a crystallographic least-squares calculation to reproduce the experimental X-ray diffraction data. With an improved set of coordinates, the procedure is iterated to self-consistency.

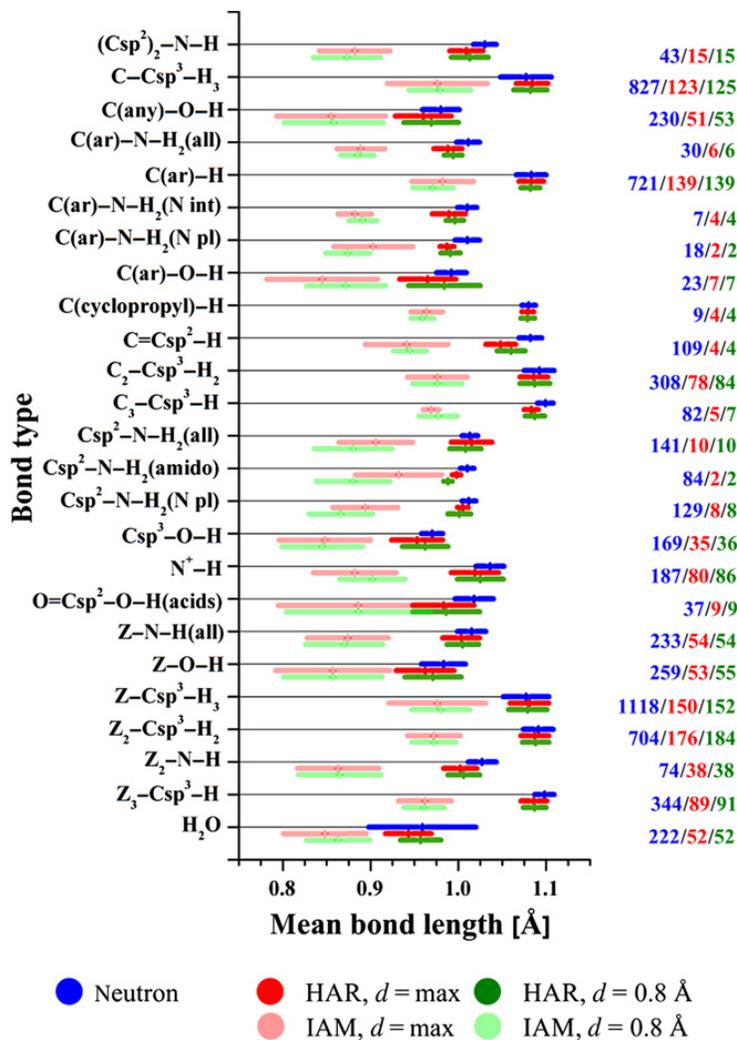
If HAR is supposed to substitute neutron diffraction experiments in the future, it is important to assess the accuracy and precision of the atomic coordinates and anisotropic displacement parameters (ADPs) by comparing the HAR results to those from neutron diffraction data. The method was tested with crystalline ammonia and the variable temperature X-ray diffraction data of the dipeptide Gly-L-Ala measured at 12, 50, 100, 150, 220 and 295 K (see Fig. 18) [16].



**Fig. 18.** Projections of the Gly-L-Ala molecule as obtained from X-ray data after HAR at the BLYP/cc-pVTZ level (left column) and from neutron data (right column). ADPs are shown at the 50% probability level. [Reproduced with permission of the International Union of Crystallography](#) from reference [16].

Starting from IAM structures, positions and anisotropic displacement parameters (ADPs) for all atoms including hydrogen atoms were freely refined without constraints or restraints, but using electron densities from Hartree–Fock and BLYP density functional calculations. The molecular geometries and ADPs for the hydrogen atoms from these refinements were indistinguishable from the neutron measurements. Their precision, estimated as the average neutron - HAR differences, was better than 0.009 Å for A–H bond lengths at 150 K or below (agreement with neutron results within 2 combined standard uncertainties (csu’s)). For the subset of C–H bonds alone, the agreement was 0.004 Å at all temperatures. BLYP density functionals produced better results than HF calculations, although BLYP functionals still resulted in shorter bond lengths compared to neutron values, whereas HF calculations resulted in longer ones. ADPs determined from HAR were found to be of the same size and orientation as those from neutron diffraction. The mean absolute differences in the H-atom ADPs (at 150 K and below) were between 0.004 and 0.006 Å<sup>2</sup>.

A more extensive study [19] has further demonstrated the reliability of determining hydrogen atom positions using X-ray diffraction and HAR. A–H bond lengths in 81 crystal structures of organic molecules were compared with those from neutron measurements. Fig. 19 summarizes the comparisons for several classes of A–H bonds.



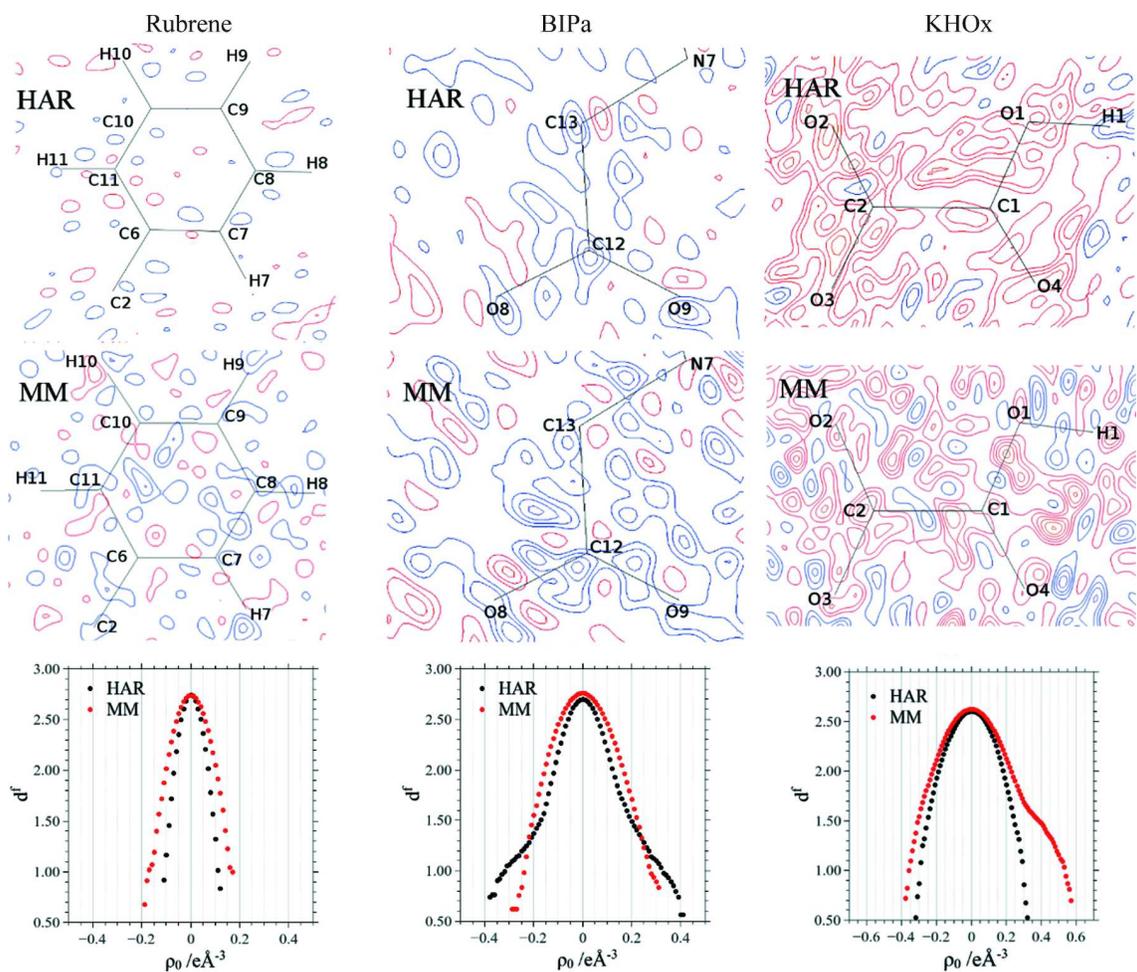
**Fig. 19.** Comparison of A–H bond lengths from HAR and neutron diffraction experiments. For each experiment type and 24 classes of A–H bonds the average A–H distances and their sample standard deviations (SDs) are given at restricted ( $d = 0.8 \text{ \AA}$ ) and full resolution ( $d = \max$ ). Neutron values were taken from Allen and Bruno [150]. The nature of the atom A, the hybridization of C-atoms and the number of atoms Z of any kind bonded to A are also given. The average O–H bond length was obtained from entries in the Cambridge Structural Database (CSD) containing water of crystallization. For all bond types the numbers of observations for each experimental method are given. Reprinted under the CC BY-NC license from reference [19].

HAR and neutron values show excellent agreement, whereas the A–H bond lengths from X-ray IAM are shorter by  $\sim 0.12 \text{ \AA}$  on average as is well known. The accuracy of the HAR results is generally higher by up to an order of magnitude. The reliability of HAR was also demonstrated

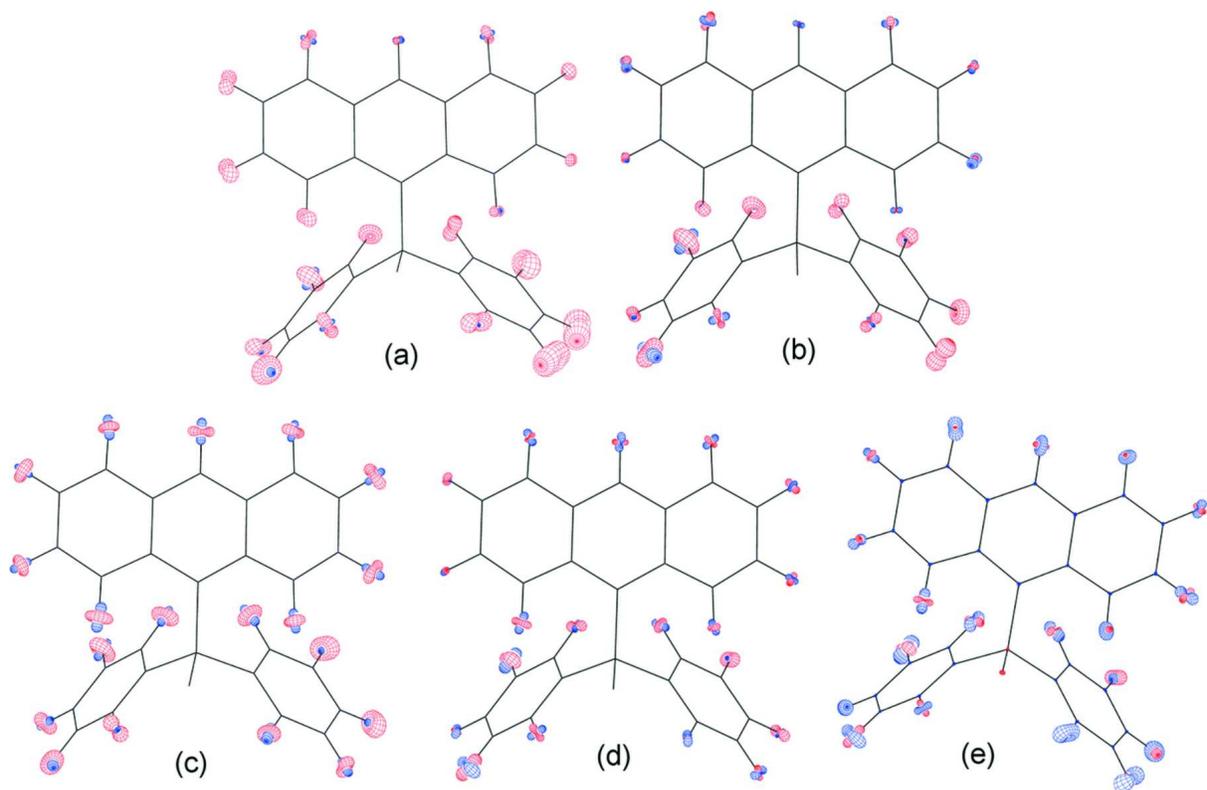
for a selection of inorganic compounds, including structures with hydrogen atoms that are bridging or bonded to transition metals. Such hydrogen atoms can usually not be located with IAMs based on X-ray data. The accuracy in hydrogen atom positions from HAR also ensures accuracy in intermolecular distances to hydrogen. This makes HAR a useful tool in solid-state supramolecular chemistry research.

In a further study, the applicability of the HAR method to a number of very different chemical situations was tested [20]: the aromatic hydrocarbon rubrene (orthorhombic 5,6,11,12-tetraphenyltetracene), a cocrystal of zwitterionic betaine, imidazolium cations and picrate anions (BIPa), and the salt potassium hydrogen oxalate (KHOx). HAR results were compared with accurate neutron diffraction geometries and the conventional multipole X-ray charge density model (MM). The non-hydrogen HAR-ADPs were in excellent agreement with the neutron-based values and were superior to IAM-ADPs. While the non-hydrogen HAR-ADPs were found to be as accurate and precise as the MM-ADPs, the residual electron density plots from HAR showed fewer features than the MM residual densities (Fig. 20). Similarly, the fractal dimension plots of the residual densities were better, i.e. closer to parabolic, for HAR. The fact that the purely theoretical electron density from HAR is a better model than a refined MM is remarkable and challenges the current notion that MM derived models represent the highest accuracy electron densities obtainable from experiment. However, this trend needs to be verified by more studies.

Stalke *et al.* [151] investigated various approximation methods for modelling positions and anisotropic displacement parameters (ADPs) for hydrogen atoms in 9-diphenylthiophosphoranyl anthracene (SPANH) and 9,10-bis-diphenylthiophosphoranyl anthracene·toluene (SPANPS). Empirical estimates of hydrogen ADPs were obtained from the SHADE-server [152] using the software APD-Toolkit [153] based on the invariom database [154]; refined values are from HAR (incorporated as the *HARt* GUI in *Olex2*) and both sets were compared with the neutron diffraction data. HAR was found to produce C–H bond lengths closest to neutron values. H-atom ADPs from HAR were larger than those from all other models, the multipole model being closest to neutron data (see Fig. 21 for X-N PEANUT plots [155] showing differences of ADPs between different models).



**Fig. 20.** Residual density plots for rubrene, BIPa and KHOx of the HAR models (top row) and the MMs (middle row). Blue denotes positive and red negative; the contour interval is  $0.05 \text{ e}\text{\AA}^{-3}$ . The bottom row shows Henn–Meindl fractal dimension plots based on the complete unit-cell electron density. [Reproduced with permission of the International Union of Crystallography](#) from reference [20].



**Fig. 21.** PEANUT plots (X–N) of the second molecule in the asymmetric unit of SPAnH. SHADE (a), APD-Toolkit (b), Aniso (no Q0) (c), Aniso+Q0 (d), HART (e). Red denotes a negative difference, indicating larger neutron ADPs, blue denotes a positive difference, indicating larger X-ray ADPs. For accentuation, differences are multiplied by a factor of 2. [Reproduced with permission of the International Union of Crystallography](#) from reference [151].

### 5.1.2 Limitations of HAR

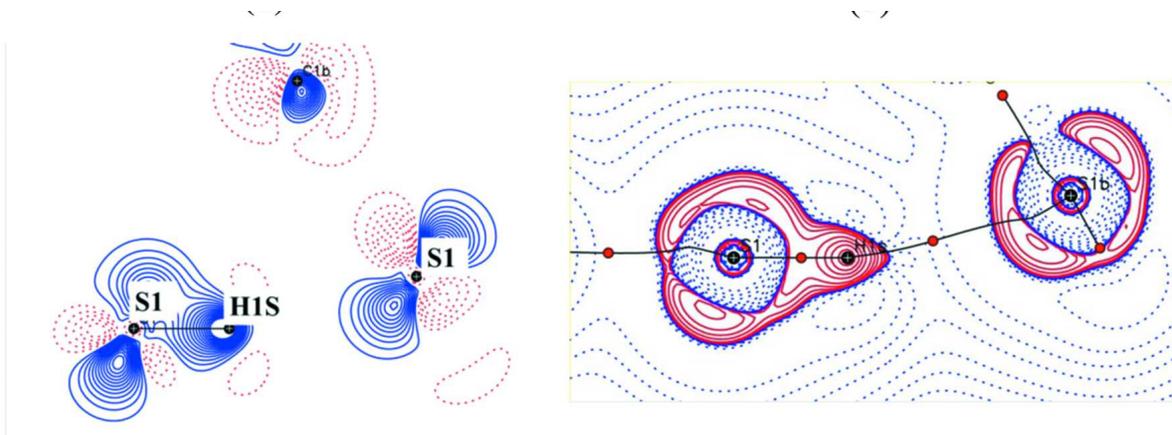
The application of HAR requires significant computational resources. In practice it is limited to crystal structures of relatively small molecules, without heavy elements and showing no crystallographic disorder. As described in subsection 3.6, HAR has been recently extended to biological macromolecules exploiting its HAR-ELMO extension [77], but extended structures (coordination polymers or metal-organic frameworks, for example) can currently not be treated satisfactorily. A more sophisticated statistical analysis by Grabowsky *et al.* showed with a two-sided Welch test that, despite the apparent agreement of HAR, bond lengths involving hydrogen atoms are not equal to results from neutron diffraction within a 95% confidence interval [19]. In almost all cases, the HAR values were found to be systematically smaller, with an average

difference between HAR and neutron diffraction-derived bond lengths of around 0.01 to 0.02 Å. It is also to be noted that there are systematic differences between the non-H atom ADPs from HAR and neutron diffraction found in the 2014 report by Capelli *et al.* [16]. Although the accuracy of A–H bond length determination is stable and rather independent of data quality and resolution, the same cannot be said for the determination of hydrogen atom ADPs; they can become non-positive definite or skewed even if the A–H distance is accurate. In such cases, implementation of certain restraints is necessary.

### ***5.1.3 Some other recent applications of HAR.***

HAR is a relatively new quantum crystallography technique. Its full potential to provide accurate structures is waiting to be realized by the general chemistry community. In this section, we outline applications of HAR dealing with S–H···S and hydrogen maleate O–H···O hydrogen bonds, a determination of chirality at an RR’N<sub>H</sub> center, an analysis of anisotropic atomic displacement parameters (ADPs) of methyl hydrogen atoms, i.e. their contribution to crystal entropy and thus to the crystal cohesive free energy.

In a study of the very weak S–H···S hydrogen bond, the hydrogen atom was located with HAR [156], since no accurate reference neutron distances are available for S–H, Se–H, etc. Based on the HAR derived hydrogen positions the electron density of the S–H···S hydrogen bond region was modelled using MM (Fig. 22) and the hydrogen bond interaction energy estimated to be ~10 kJ/mol, i.e. in the range of weak hydrogen bonds.

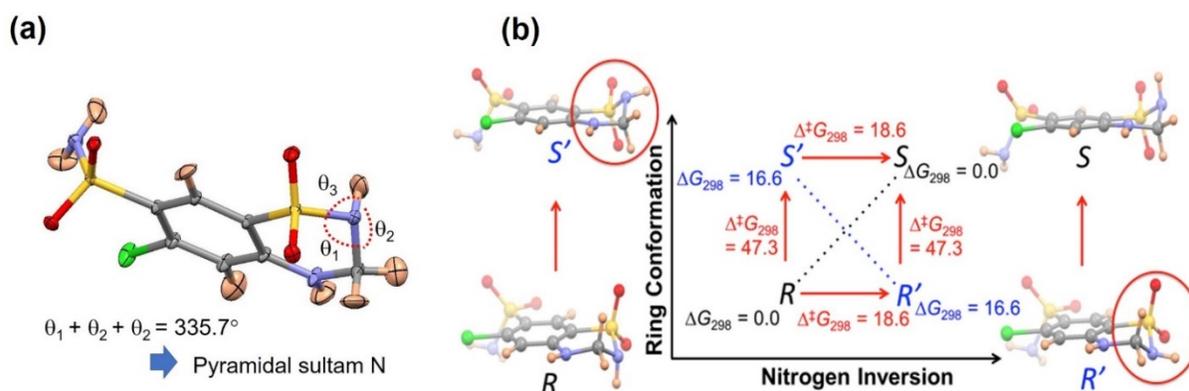


**Fig. 22.** Deformation density and two-dimensional Laplacian plot of the S–H···S hydrogen bond region in 2-mercaptobenzoic acid (experimental MM). Blue represents charge concentration and red represents charge depletion; contours are drawn at intervals of  $0.05 \text{ e}\text{\AA}^{-3}$ . The Laplacian is plotted on logarithmic contours. [Reproduced with permission of the International Union of Crystallography](#) from reference [156].

The influence of the crystal environment on the proton position in the intramolecular O···H···O hydrogen bond regions was studied by neutron diffraction in a series of hydrogen maleate salts [157]. The “bridging hydrogen atom” position correlates well with the O···O distance, the crystalline environment, and the degree of asymmetry of intermolecular hydrogen bonding interactions, but does not depend on temperature. This finding allows to accurately estimate the H atom position in hydrogen maleate salts from O···O distances in X-ray IAM models. In an extension of this study the location of the bridging hydrogen atom positions (usually found to be unreliable in X-ray models) were obtained using HAR and found to be in good agreement with the neutron diffraction results [158].

Crystals of the diuretic drug hydrochlorothiazide (HCT), a sulfonamide, show an unusual enantiomorphic twin crystal growth and the crystal structure indicates a subtle chirality phenomenon at the nitrogen atom (Fig. 23) [159]. The focus of this study was the geometry at the N atom and hence the location of the proton attached to it. N atoms attached to sulfur in sulfonamides usually exhibit planar coordination, owing to a resonance effect. Cyclic sulfonamides (sultams) are also expected to exhibit planar geometry around the sulfonamide N atom. However, in some examples these N atoms prefer a pyramidal geometry. In HCT the

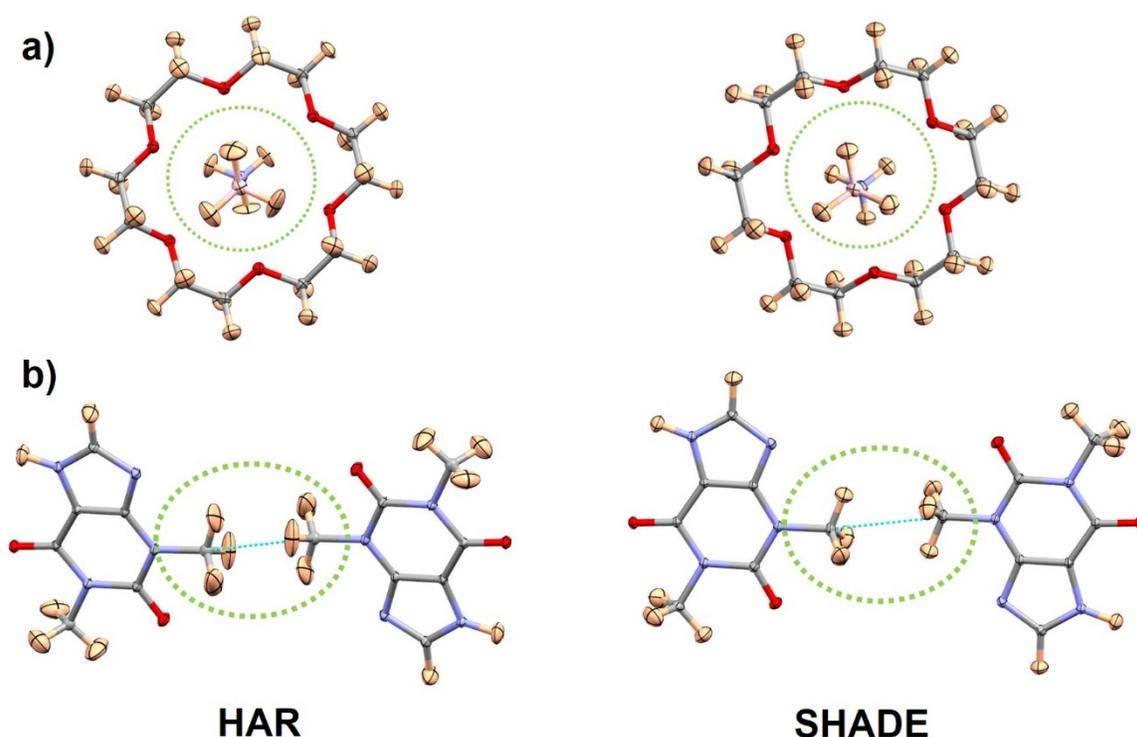
approximate location of the H atom at nitrogen from an IAM refinement indicated a pyramidal geometry, but the simultaneous presence of lone pair electron density near the  $sp^3$  N atom made the H-position ambiguous. HAR confirmed the pyramidal geometry and the approximate IAM position unambiguously (Fig. 23, unpublished results) [159]. The pyramidal geometry around the sultam N atom was initially found from HAR and was confirmed by the structural model from the neutron diffraction data [159]. Although HAR is a new technique, it is now sufficiently well tested for scientific referees in the crystallography/structural chemistry community to begin trusting its benefits. Fig. 23 also shows a computational analysis of the pathways for interconversion between the S and R enantiomers because circular dichroism spectra indicated racemic interconversion in solution. The calculations showed that conformational chirality is maintained even in the intermediate transition structure of this molecule.



**Fig. 23.** (a) Pyramidal geometry around the sultam N atom in hydrochlorothiazide, as confirmed by the accurate H atom positions from HAR refinement (HAR-ORTEP with 50% probability); (b) Gibbs free energies for the two symmetric interconversion pathways between the S and R enantiomers of HCT. B3LYP/6-31G(2df,p) optimized structures are shown with the relative G4(MP2) standard Gibbs free energies ( $\text{kJ mol}^{-1}$ ) for the local minima and transition structures [159].

Recent results demonstrate that HAR can provide anisotropic ADPs for hydrogen atoms that are realistic since they correspond to common modes of thermal motion in crystal structures. Fig. 24a shows the host-guest motif in the crystal structure of the ammonia-borane:18crown6 complex [160]. The hydrogen ADPs of the  $\text{NH}_3\text{BH}_3$  molecule indicate overall and/or intramolecular oscillatory motion. In this case the widely used SHADE model [152] cannot capture such motion, because it derives the hydrogen ADPs by combining contributions from a

rigid-body model of the non-hydrogen atoms with those from a database of bond stretch and angle bend motions of the hydrogen atom. Moreover, the B–H and N–H bond lengths in the host-guest complex obtained from HAR agree with the neutron diffraction values for the pristine ammonia-borane crystal structure. Modelling the host-guest electron density with the HAR derived hydrogen coordinates and ADPs using ultra-high-resolution X-ray synchrotron data resulted in a good quality MM charge density model.



**Fig. 24.** (a)  $\text{NH}_3\text{BH}_3$  in its host-guest crystalline complex with 18-crown-6. The hydrogen ADPs from HAR indicate roto-oscillations of the  $\text{NH}_3\text{BH}_3$  molecule in contrast to the ADPs from a corresponding SHADE analysis. Unpublished work [160]. (b) Roto-oscillatory vibrations of the methyl groups in a co-crystal of caffeine manifested in HAR-ADPs. The entropy contribution arising from these modes stabilizes the methyl...methyl hydrophobic interaction motif [161].

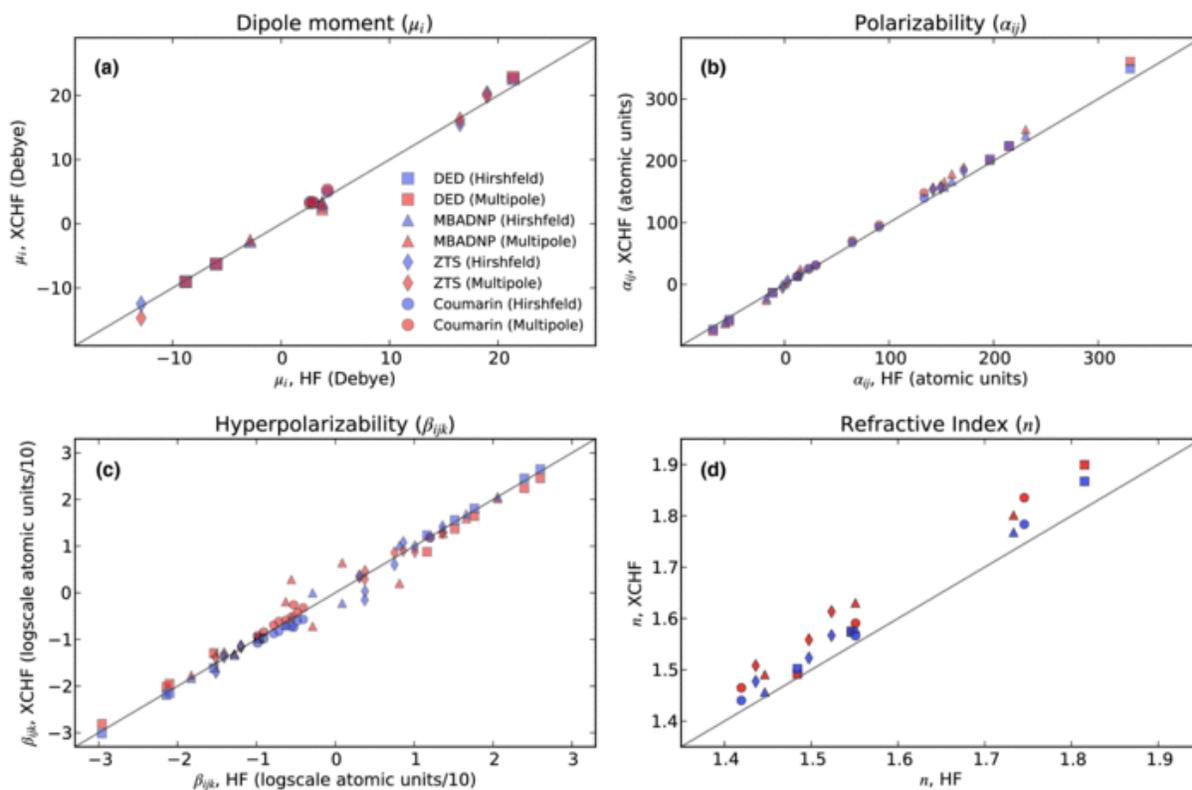
Rotatory thermal motions of methyl groups are known from neutron diffraction studies (e.g. in aspirin and paracetamol). Such oscillations have also been observed for a methyl...methyl hydrophobic interaction (Fig. 24b) [161]. Using HAR derived ADPs, methyl oscillation frequencies were derived by thermal motion analysis, and their entropy contribution estimated to

be around 15-16 J mol<sup>-1</sup> K<sup>-1</sup>, equivalent to free energy contributions of 4.4 to 4.8 kJ mol<sup>-1</sup> at room temperature. Thus, the entropy term  $TS \sim 5$  kJ mol<sup>-1</sup> lowers the crystal free energy and stabilizes these short hydrophobic contacts. Although small, contributions of solid-state entropies related to ADPs can be important in the context of the relative stability of polymorphs. Previous works by Aree *et al.* [162] and Madsen *et al.* [163] have demonstrated the quantification of solid-state entropies from ADPs obtained from good quality X-ray data.

## ***5.2 Opto-electronic properties from XCW fitting: polarizabilities, hyperpolarizabilities and refractive indices***

X-ray constrained wavefunction (XCW) fitting is an independent procedure, not related to HAR *per se*. In the HAR procedure, the positions of the atoms are refined against the experimental data such that the associated theoretical wavefunction and thus the theoretical electron density optimally fit the X-ray diffraction data; in XCW fitting not only the atomic positions, but also variables of the wavefunction (specifically, the molecular orbital coefficients) are varied to reach this goal (see section 4 for more details). From the resulting “experimental” wavefunction molecular properties become accessible, e.g. opto-electronic properties such as polarizabilities, hyper-polarizabilities, refractive indices, etc. Note that the one particle density obtained from the usual multipole refinement electron density model is insufficient to determine such response properties [142, 164]. Dipole polarizabilities, linear bulk susceptibilities,  $\chi^{(1)}$ , and refractive indices for benzene, urea, and 2-methyl-4-nitroaniline were calculated (via simplified sum-over-states expressions) from an XCW. Independent, high-level calculations clearly show that the XCW results include the effects of intermolecular interactions and electron correlation, especially for urea where the effects on the polarizability are known to be quite large [141]. Dipole moments, molecular polarizabilities, hyper-polarizabilities and refractive indices were computed from XCW for a series of crystals exhibiting non-linear optical properties: coumarin (1-benzopropan 2-one), DED ({4-bis (diethylamino)- methylum]phenyl}dicyanomethanide), MBADNP (3,5-dinitro-2-[1-phenyl-ethyl]-aminopyridine) and ZTS (zinc [tris]thiourea sulphate). The calculated XCW properties agreed with the Hartree–Fock level computational results within an average statistical deviation of 20% (see Fig. 25). In a later report on the optical properties of ZTS it was shown that the properties derived from a multipole electron density model are

different from those obtained by *ab initio* computations or X-ray wavefunction fitting by up to two orders of magnitudes [165]. The better agreement of XCW derived values could be understood as resulting from the perturbation of an isolated-molecule wavefunction by crystal field effects.



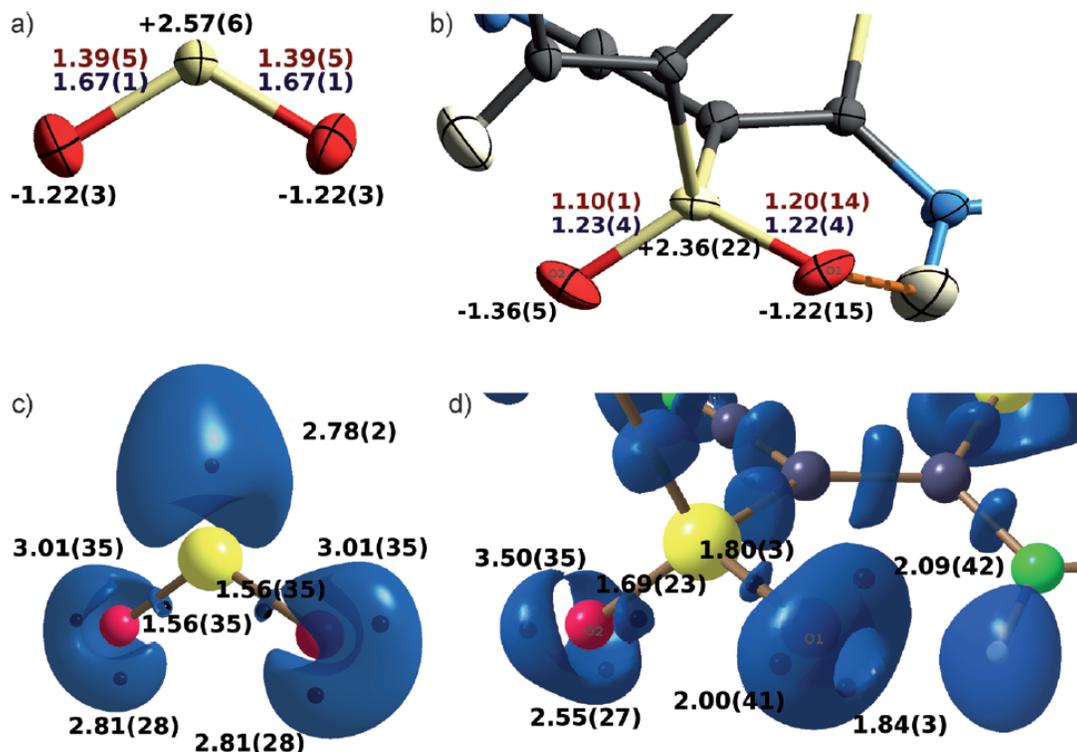
**Fig. 25.** Comparison of optical properties calculated at the Hartree-Fock level with those calculated from XCWs. Two slightly different geometries and sets of atomic displacement parameters were used for the XCW fit, one based on a multipole refinement, the other on a Hirshfeld atom refinement model. Reprinted from reference [141] with the permission of AIP Publishing.

### 5.3 Insights into chemical bonding from X-ray wavefunction refinement

Given a wavefunction that not only fulfils a variational energy minimization criterion but also shows an optimal fit to Bragg diffraction data, it becomes possible to analyse details of chemical bonding that include effects of the crystal field. The examples in this section address concepts such as bond order, covalent and ionic bonding, hypervalency, push-pull effects, donor-acceptor

interactions and aromaticity. The wavefunctions discussed below were obtained with a so-called X-ray wavefunction refinement (XWR). This procedure is a combination of HAR and XCW fitting (see also subsections 3.7 and 4.1).

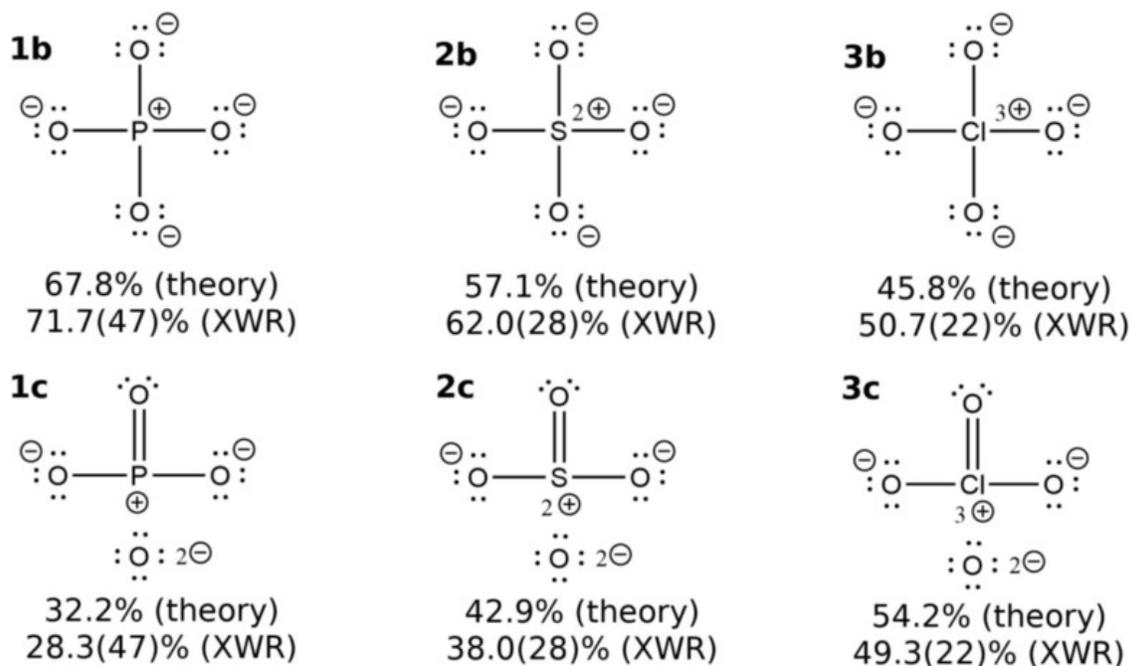
In its simplest form, the bonding notion of hypervalency is expressed by Lewis structures with atoms sharing more than four valence electron pairs with their neighbours, thus formally breaking the octet rule. Although hybridization schemes involving d orbitals ( $sp^3d$  or  $sp^3d^2$  hybrids) are proposed to describe such hypervalent bonding, the large promotion energies required for the formation of these hybrid orbitals makes such a picture unlikely for period 3 elements. In 2012, Grabowsky *et al.* analyzed the putative hypervalency in cryocrystallized sulfur dioxide  $SO_2$  and in the sulfonyl group of an organic molecule [85]. The bond indices of 1.5 and 1.2 in  $SO_2$  and the sulfonyl group derived from an X-ray wavefunction refinement showed that ionic and multi-center bonding predominate in the S–O bonds. These results ruled out the need for a hypervalent structure and showed that the electrostatic forces associated with the ionicity were the reason for S–O bond shortening in these molecules, and not the presence of multiple covalent bonds such as S=O (Fig. 26).



**Fig. 26.** Experimentally derived atoms-in-molecules charges (black, in e) and experimentally derived bond indices (first line/dark red: delocalization index; second line/dark blue: RGBI for a) SO<sub>2</sub> and b) the sulfonyl group; experimentally derived ELI valence basin populations (in e) for c) SO<sub>2</sub> and d) the sulfonyl group, black spheres below the transparent ELI isosurfaces indicate the positions of the basin attractors (that is, maxima); isovalue in (c): ELI=1.475, isovalue in (d): ELI=1.445 around the S-O bond attractors and ELI=1.550 around the lone pair attractors. Figure reproduced from reference [85] with permission of John/Wiley & Sons, Inc.

In a thematic continuation of this work, the concept of hypervalency was examined for the isoelectronic anions phosphate (PO<sub>4</sub><sup>3-</sup>), sulfate (SO<sub>4</sub><sup>2-</sup>) and perchlorate (ClO<sub>4</sub><sup>-</sup>) [166]. X-ray wavefunction refinement applied to high-resolution X-ray diffraction data of representative crystal structures (MgNH<sub>4</sub>PO<sub>4</sub>·6H<sub>2</sub>O, Li<sub>2</sub>SO<sub>4</sub>·H<sub>2</sub>O, and KClO<sub>4</sub>), were analyzed for the percentage contributions of various resonance structures. The P-O and S-O bonds were found to be highly polarized covalent bonds, and hypervalency was ruled out for phosphorus and sulfur atoms (Fig. 27). In the case of ClO<sub>4</sub><sup>-</sup>, a kind of hypervalent chlorine atom was indicated by the

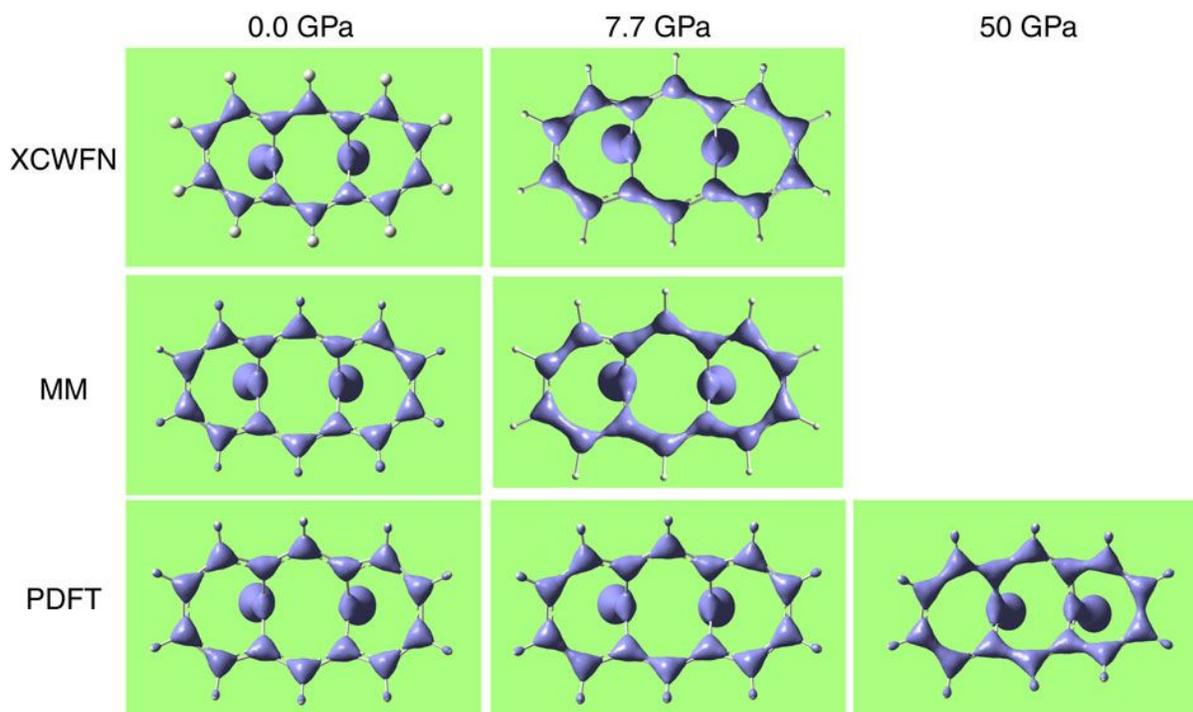
bonding indicators derived from both the X-ray and the pure quantum chemical wavefunctions. However, it is not due to d-orbital participation, but rather to pronounced hyperconjugation of the oxygen lone pairs with the  $\sigma^*(\text{Cl-O})$  orbitals.



**Fig. 27.** Lewis structures and corresponding weights obtained from natural resonance theory applied to the theoretical or the experimentally constrained wavefunction after X-ray wavefunction refinement (XWR). For Lewis structures 1c, 2c and 3c only one of the equivalent Lewis structures is depicted, the corresponding weights are a sum of all equivalent resonance forms. Figures reproduced with permission of John/Wiley & Sons, Inc. from reference [166].

The compound syn-1,6:8,13-Biscarbonyl[14]annulene (BCA) shows aromatic stabilization at ambient conditions which can be depicted by two equivalent Lewis structures showing alternating single and double bonds. At a pressure of 7.7 GPa one of the resonance structures is clearly favoured and aromatic stabilization decreases [132]. This conclusion follows from the molecular geometries as well as from the experimental electron density distribution modelled using MM and X-ray constrained wavefunctions (Fig. 28). As already mentioned in subsection 4.2, these results were also confirmed by XC-ELMO-VB calculations [61]. The resolution and completeness of high-pressure X-ray datasets are often limited by experimental constraints. Incomplete data may lead to chemically unrealistic results in MM for high pressure electron

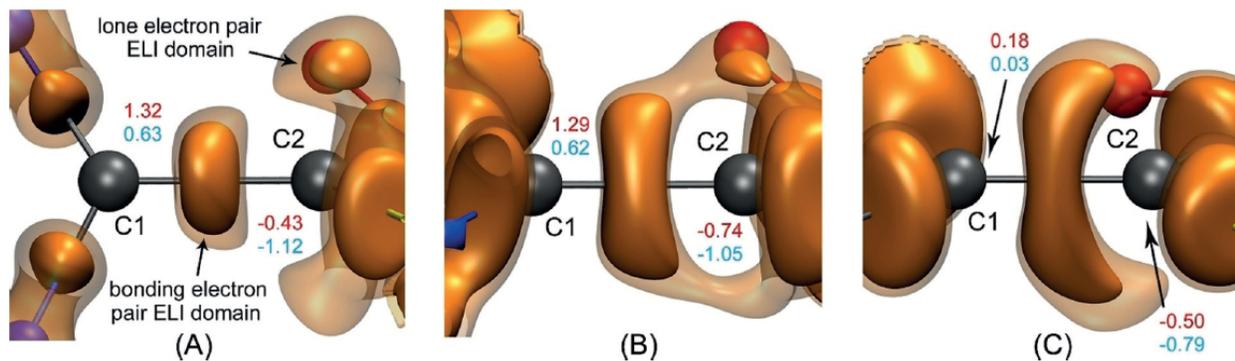
density studies. In XWR studies, such artefacts are likely to be (largely) compensated by the quantum chemical information implied in the X-ray constrained wavefunction. It may thus be argued that for high pressure investigations XWR is superior to MM refinements.



**Fig. 28.** Electron density distribution of BCA. Plots are shown at various pressure and from various sources: PDFT are periodic DFT calculations at the B3LYP level of theory; XCWFN are X-ray constrained wavefunctions computed at a Hartree–Fock level, but constrained against the experimentally measured diffraction intensities; MM is the electron density derived from a multipolar expansion, with population coefficients refined against experimentally measured intensities. Figure reproduced with permission of SpringerNature from reference [132].

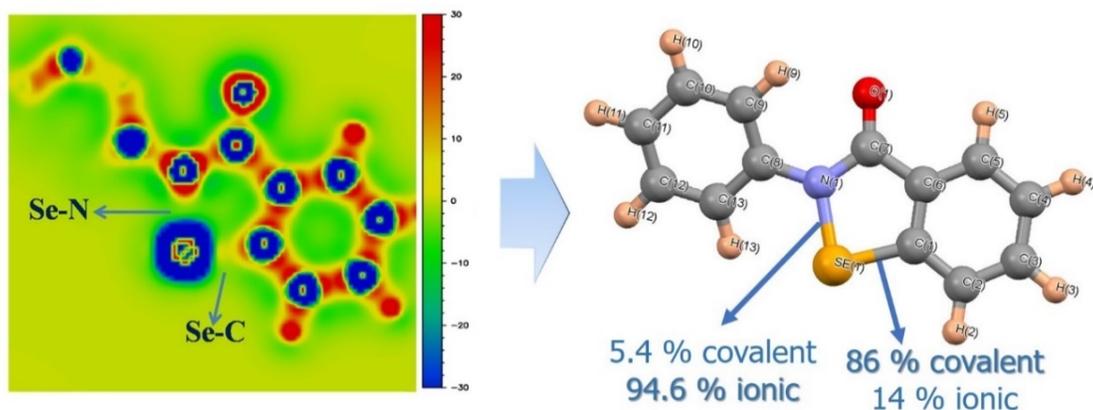
Yanai *et al.* studied chemical bonding in polarised ethylenes, so-called push–pull ethylenes, with both twisted and planar partial “C=C” bond [167]. Covalent and ionic bond indices and Electron Localizability Indicators derived from X-ray wavefunction refinement (XWR) revealed a dominant contribution of the charge-separated resonance structure in the twisted compound (Fig. 29). The nearly planar compound showed the expected  $\pi$ -bonding character, albeit with a contribution of the charge-separated resonance structure. The percentage of various resonance

contributions were extracted from the X-ray wavefunctions with the help of natural bond orbital (NBO) theory and natural resonance theory (NRT). The results from the Hartree-Fock based X-ray wavefunction agreed with the results from MP2 calculations, but not with those from regular Hartree-Fock calculations.



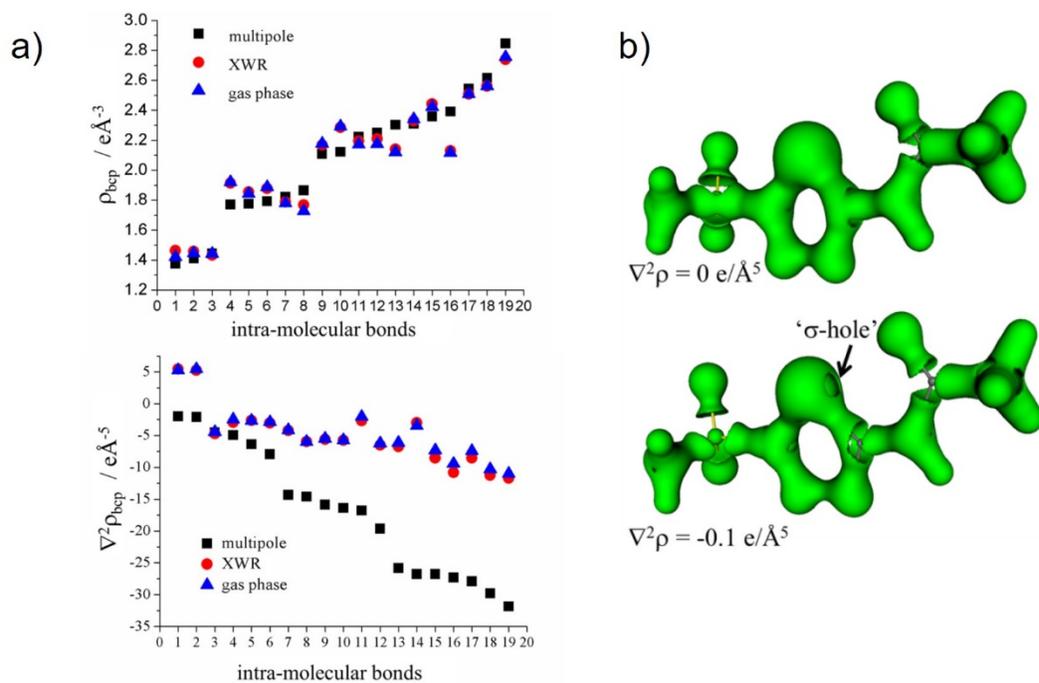
**Fig. 29.** Isosurfaces of the XWR-derived ELI for A) the twisted form, B) the planar form, and C) a reference conjugated C=C double bond. Isovalue=1.45 (transparent), 1.60/1.52 (solid, A/[B and C]). XWR derived atomic charges are given in e (red, QTAIM charge; blue, NPA charge). Figure reproduced with permission of John/Wiley & Sons, Inc. from reference [167].

The organoselenium anti-depressant candidate ebselen exhibits unusually short intermolecular chalcogen bonding interactions in its polymorphs [168] and strikingly different electron density features for intramolecular Se–N and Se–C bonds. In a MM charge density analysis of this compound the Laplacian of the electron density indicated that the Se–N bond might be much more polarised than the Se–C bond although both these bonds are perceived to be covalent in nature. In an extension of this study [169], a Roby Gould bond order analysis of the quantum crystallographic XWR wavefunction indicated that the Se–N bond is indeed highly polarised, being predominantly ionic in nature with 94.6% ionicity, whereas the Se–C bond is covalent with 86% covalency (Fig. 30). These results support the proposal of an antioxidant activity of this class of compounds which involves Se–N bond cleavage during the drug action. Given the highly polar Se–N single bond, such a mechanism is indeed possible in a polar solvent/supramolecular environment.

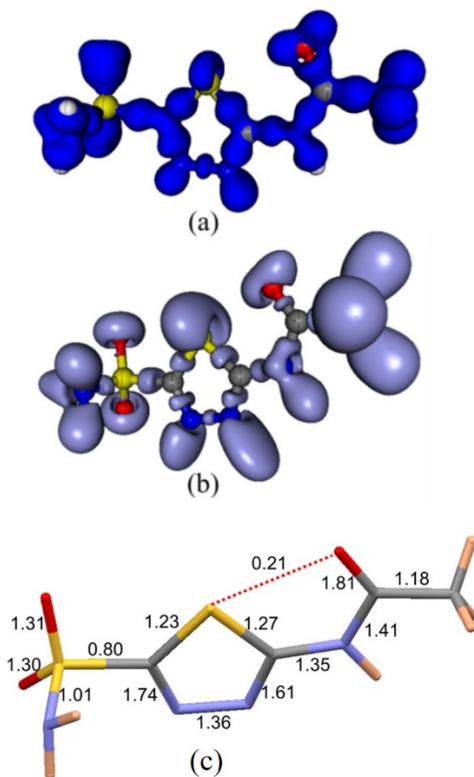


**Fig. 30.** (Left) The Laplacian of electron density in the C–Se–N molecular plane of ebselen obtained from MM indicating the difference in the nature of Se–C and Se–N bonds, and (right) percentage ionicity and covalency associated with these bonds estimated from an X-ray constrained wavefunction [169].

In another application of XWR, Row *et al.* investigated the features of the diuretic drug acetazolamide that shows intramolecular S $\cdots$ O chalcogen bonding [170]. The bond order of the S $\cdots$ O interactions as estimated from the X-ray constrained wavefunction was 0.21, a value comparable to those obtained for intermolecular S $\cdots$ O interactions in the molecular complexes of the drug riluzole. The features of the XWR deformation electron density were found to be quite similar to those from MM. The features of ELI derived from the X-ray constrained wavefunction near the S and O atoms showed the lone pair density on the O atom oriented to a region of less electron pairing on the S atom, which is consistent with the features from 3D Laplacian and deformation density maps showing charge-concentrated-to-charge-depleted (CC-CD) nature (Figs. 31b and 32). On the zero-Laplacian isosurface around the S atom  $\sigma$ -hole features were observed, which substantiated the S $\cdots$ O chalcogen bonding. The orientations of this pair of  $\sigma$ -holes matched with those obtained from the ESP maps. Importantly, this study brought out the difference in electron density topological features at different bond critical points obtained from XWR as compared to multipole model (Fig. 31a).



**Fig. 31.** a) Comparison of electron density,  $\rho_{\text{bcp}}$ , and its Laplacian,  $\nabla^2 \rho_{\text{bcp}}$ , obtained from multipole model, XWR, and a gas phase single point calculation. b) Laplacian isosurfaces obtained from XWR, plotted for  $\nabla^2 \rho = 0 \text{ e}\text{\AA}^{-5}$  and  $\nabla^2 \rho = -0.1 \text{ e}\text{\AA}^{-5}$ . Reproduced from reference [170].



**Fig. 32.** a) static 3D deformation density map ( $0.03 \text{ e } \text{\AA}^{-3}$ ) obtained from XWR, b) the ELI-D surface for ELI-D = 1.45, and c) bond orders of different bonds in acetazolamide, including the S...O chalcogen bond. Reproduced from reference [170].

#### ***5.4 Properties from X-ray constrained wavefunctions: current developments and future perspectives***

Quantum crystallography is a nascent field and its full potential in terms of chemical applications is yet to be explored. At present, most studies based on X-ray constrained wavefunctions are focused on extracting *intramolecular* chemical properties, but there is a genuine interest in deriving electronic descriptors for *intermolecular* interactions in crystals. In a recent report, Alhameedi *et al.* have classified and characterised unusual types of weak intermolecular interactions in crystals such as halogen bonding, chalcogen bonding etc. with bond order estimates derived from Roby-Gould bond indices [171]. Although this study was based on gas phase calculations on molecular dimers, it can be expected that bond order values for intermolecular interactions in crystals will be of significant interest because X-ray constrained

wavefunctions inherently represent crystal field effects. Current developments in this direction also include attempts to estimate intermolecular interaction energies from X-ray constrained wavefunctions using the model-energy CE-B3LYP scheme in the *CrystalExplorer* (CE) software [172].

Extracting electronic properties from X-ray constrained wavefunctions is another interesting research field. Jayatilaka derived the ionization energy of beryllium from the fitted wavefunction in the first report on XCW [21]. A value of 4.9 eV was derived from the wavefunction fitted to the diffraction data and the BLYP functional (comprising the Becke exchange functional and a correlation functional). This result agrees well with the known experimental value of 4.98 eV. We are currently extending this approach to molecular functional materials such as hybrid organic-inorganic semiconductors with the aim of extracting ionization energies and HOMO-LUMO gaps from X-ray constrained wavefunctions.

A procedure similar to HAR, but using periodic wavefunctions, has recently been reported by Wall (see subsection 3.4) [18]. This method will have to be developed further also to obtain *periodic X-ray constrained wavefunctions* and corresponding properties such as electronic band gaps of crystalline materials. Future developments of quantum crystallography in these directions can be expected to lead to useful insights into the electronic properties in chemistry and materials sciences.

## **Acknowledgments.**

The authors gratefully thank Professor Hans-Beat Bürgi for the careful and detailed revision of the whole text, which allowed to significantly improve the readability and the clarity of the chapter. The authors also acknowledge Dr. Horst Puschmann for the careful proofreading of section 3. The authors also thank all their colleagues and co-workers who are not co-authors of the present chapter, but who produced unpublished pictures and results reported here: Professor Bo B. Iversen, Bjarke Svane, Lorraine A. Malaspina, Dr. Rumpa Pal, Dr. Emanuel Hupf, Florian Kleemiss and Erna K. Wieduwilt.

A. G. acknowledges the French Research Agency (ANR) for financial support of the Young Researcher Project *QuMacroRef* through Grant No. ANR-17-CE29-0005. S. G. thanks the

German Research Foundation (DFG) for support within the Emmy Noether scheme GR 4451/1-1 and the project GR 4451/2-1. S.P.T. acknowledges European Union funding for a Marie Skłodowska-Curie Individual Fellowship (grant number 798633) and thanks Professor Bo B. Iversen and Professor Mark A. Spackman for the permission to discuss some unpublished results.

## References

- <sup>1</sup> Massa L, Huang L, Karle J (1995) Quantum crystallography and the use of kernel projector matrices. *Int J Quantum Chem, Quantum Chem Symp* 29: 372-384.
- <sup>2</sup> Grabowsky S, Genoni A, Bürgi H-B (2017) Quantum crystallography. *Chem Sci* 8: 4159-4176.
- <sup>3</sup> Genoni A, Bučinský L, Claiser N, Contreras-García J, Dittrich B, Dominiak PM, Espinosa E, Gatti C, Giannozzi P, Gillet J-M, Jayatilaka D, Macchi P, Madsen AØ, Massa LJ, Matta CF, Merz KM Jr, Nakashima PNH, Ott H, Ryde U, Schwarz K, Sierka M, Grabowsky S (2018) Quantum Crystallography: Current Developments and Future Perspectives. *Chem Eur J* 24: 10881-10905.
- <sup>4</sup> Tsirelson V (2017) Early days of quantum crystallography: A personal account. *J Comput Chem* 39: 1029-1037.
- <sup>5</sup> Massa L, Matta CF (2017) Quantum crystallography: A perspective. *J Comput Chem* 39: 1021-1028.
- <sup>6</sup> Matta CF (2017) A path through Quantum Crystallography. *Struct Chem* 28: 1279–1283.
- <sup>7</sup> Korlyukov AA, Nelyubina YV (2019) Quantum chemical methods in charge density studies from X-ray diffraction data. *Russ Chem Rev* 88: 677-716.
- <sup>8</sup> Einstein A (1907) Die Plancksche Theorie der Strahlung und die Theorie der spezifischen Wärme. *Ann Physik* 327: 180-190
- <sup>9</sup> Dawson B (1967) A general structure factor formalism for interpreting accurate X-ray and neutron diffraction data. *Proc R Soc Lond A* 298: 255–263.
- <sup>10</sup> Kurki-Suonio (1968) On the information about deformations of the atoms in X-ray diffraction data. *Acta Crystallogr, Sect A* 24: 379–390.
- <sup>11</sup> Hirshfeld FL (1971) Difference densities by least-squares refinement: fumaramic acid. *Acta Crystallogr, Sect B* 27: 769–781.

- <sup>12</sup> Stewart RF (1976) Electron population analysis with rigid pseudoatoms *Acta Crystallogr, Sect A* 32: 565–574.
- <sup>13</sup> Hansen NK, Coppens P (1978) Testing aspherical atom refinements on small-molecule data sets. *Acta Crystallogr, Sect A* 34: 909-921.
- <sup>14</sup> Davidson ER (1976) *Reduced Density Matrices in Quantum Chemistry*. Academic Press, New York (USA).
- <sup>15</sup> Jayatilaka D, Dittrich B (2008) X-ray structure refinement using aspherical atomic density functions obtained from quantum mechanical calculations. *Acta Crystallogr, Sect A* 64: 383-393.
- <sup>16</sup> Capelli SC, Bürgi H-B, Dittrich B, Grabowsky S, Jayatilaka D (2014) Hirshfeld Atom Refinement. *IUCrJ* 1: 361-379.
- <sup>17</sup> Woińska M, Jayatilaka D, Spackman MA, Edwards AJ, Dominiak PM, Woźniak K, Nishibori E, Sugimoto K, Grabowsky S (2014) Hirshfeld atom refinement for modeling strong hydrogen bonds. *Acta Crystallogr, Sect A* 70: 483-498.
- <sup>18</sup> Wall ME (2016) Quantum crystallographic charge density of urea. *IUCrJ* 3: 237-246.
- <sup>19</sup> Woińska M, Grabowsky S, Dominiak PM, Woźniak K, Jayatilaka D (2016) Hydrogen atoms can be located accurately and precisely by x-ray crystallography. *Sci Adv* 2: e1600192.
- <sup>20</sup> Fugel M, Jayatilaka D, Hupf E, Overgaard J, Hathwar VR, Macchi P, Turner J, Howard JAK, Dolomanov OV, Puschmann H, Iversen BB, Bürgi H-B, Grabowsky S (2018) Probing the accuracy and precision of Hirshfeld atom refinement with HART interfaced with Olex2. *IUCrJ* 5: 32-44.
- <sup>21</sup> Jayatilaka D (1998) Wave Function for Beryllium from X-Ray Diffraction Data. *Phys Rev Lett* 80: 798-801.
- <sup>22</sup> Jayatilaka D, Grimwood DJ (2001) Wavefunctions Derived from Experiment. I. Motivation and Theory. *Acta Crystallogr, Sect A* 57: 76–86.
- <sup>23</sup> Grimwood DJ, Jayatilaka D (2001) Wavefunctions Derived from Experiment. II. A Wavefunction for Oxalic Acid Dihydrate. *Acta Crystallogr, Sect A* 57: 87–100.
- <sup>24</sup> Bytheway I, Grimwood DJ, Jayatilaka D (2002) Wavefunctions Derived from Experiment. III. Topological Analysis of Crystal Fragments. *Acta Crystallogr, Sect A* 58: 232–243.
- <sup>25</sup> Bytheway I, Grimwood DJ, Figgis BN, Chandler GS, Jayatiaka D (2002) Wavefunctions Derived from Experiment. IV. Investigation of the Crystal Environment of Ammonia. *Acta Crystallogr, Sect A* 58: 244–251.

- <sup>26</sup> Grimwood DJ, Bytheway I, Jayatilaka D (2003) Wavefunctions Derived from Experiment. V. Investigation of Electron Densities, Electrostatic Potentials, and Electron Localization Functions for Noncentrosymmetric Crystals. *J Comput Chem* 24: 470–483.
- <sup>27</sup> Jayatilaka D (2012) Using Wave functions to Get More Information Out of Diffraction Experiments. In: Gatti C, Macchi P (eds) *Modern Charge-Density Analysis*, Springer Netherlands, Dordrecht, p 213-257.
- <sup>28</sup> Hirshfeld FL (1977) XVII. Spatial Partitioning of Charge Density. *Isr J Chem* 16: 198-201.
- <sup>29</sup> Hirshfeld FL (1977) Bonded-Atom Fragments for Describing Molecular Charge Densities. *Theor Chim Acta* 44: 129-138.
- <sup>30</sup> Spackman MA, Maslen EN (1986) Chemical properties form the promolecule. *J Phys Chem* 90: 2020-2027.
- <sup>31</sup> Clementi E, Roetti C (1974) Roothaan-Hartree-Fock atomic wave functions: Basis functions and their coefficients for ground and certain excited states of neutral and ionized atoms,  $Z \leq 54$ . *At Data Nucl Data Tables* 14: 177-478.
- <sup>32</sup> Koga T, Omura M, Teuya H, Thakkar AJ (1993) Improved Roothaan-Hartree-Fock wavefunctions for isoelectronic series of the atoms He to Ne. *J Phys B* 28: 3113-3121.
- <sup>33</sup> Becke AD (1988) A multicenter numerical integration scheme for polyatomic molecules. *J Chem Phys* 88: 2547-2553.
- <sup>34</sup> Weigend, F (2006) Accurate Coulomb-fitting basis sets for H to Rn. *Phys Chem Chem Phys* 8: 1057-1065.
- <sup>35</sup> Aquilante F, Gagliardi L, Pedersen TB, Lindh R (2009) Atomic Cholesky decompositions: A route to unbiased auxiliary basis sets for density fitting approximation with tunable accuracy and efficiency. *J Chem Phys* 130: 154107.
- <sup>36</sup> Koritsanszky T, Volkov A, Coppens P (2002) Aspherical-atom scattering factors from molecular wave functions. 1. Transferability and conformation dependence of atomic electron densities of peptides within the multipole formalism. *Acta Crystallogr, Sect A* 58: 464-472.
- <sup>37</sup> Dominiak PM, Volkov A, Li X, Messerschmidt M, Coppens P (2007) A Theoretical Databank of Transferable Aspherical Atoms and Its Application to Electrostatic Interaction Energy Calculations of Macromolecules. *J Chem Theory Comput* 3: 232-247.

- <sup>38</sup> Coppens P, Willoughby TV, Csonka LN (1971) Electron population analysis of accurate diffraction data. I. Formalisms and restrictions. *Acta Crystallogr, Sect A* 27: 248–256.
- <sup>39</sup> Stewart RF (1969) Generalized X-ray scattering factors. *J Chem Phys* 51: 4569–4577.
- <sup>40</sup> Tanaka K (1988) X-ray analysis of wavefunctions by the least-squares method incorporating orthonormality. I. General formalism. *Acta Crystallogr, Sect A* 44: 1002–1008.
- <sup>41</sup> Michael JR, Koritsanszky T (2015) Validation of convolution approximation to the thermal-average electron density. *J Math Chem* 53: 250-259.
- <sup>42</sup> McWeeny R (1953) X-ray scattering by aggregates of bonded atoms. III. The bond scattering factor: simple methods of approximation in the general case. *Acta Crystallogr, Sect A* 6: 631-637.
- <sup>43</sup> Chandler GS, Spackman MA (1978) Fourier transforms of Gaussian orbital products. *Acta Crystallogr, Sect A* 6: 631-637.
- <sup>44</sup> Barua M, Weyrich W (1986) Efficient evaluation of X-ray Scattering Integrals of Cartesian Gaussian-Type Functions. *Acta Crystallogr, Sect A* 42: 257-261.
- <sup>45</sup> Northey Y, Zotev N, Kirrander A (2014) Ab initio calculation of molecular diffraction. *J Chem Theory Comput* 10: 4911-4920.
- <sup>46</sup> Moreno Carrascosa A, Yong H, Crittenden DL, Weber P, Kirrander A (2019) Ab-initio calculation of total x-ray scattering from molecules. *J Chem Theory Comput* 15: 2836-2846.
- <sup>47</sup> Northey T, Kirrander A (2019) Ab Initio Fragment Method for Calculating Molecular X-ray Diffraction. *J Phys Chem A* 123: 3395-3406.
- <sup>48</sup> Jayatilaka D (1994) Fourier transform of property densities with Gaussain functions. *Chem Phys Lett* 230: 228-230.
- <sup>49</sup> McMurchie LE, Davidson ER (1978) One- and two- electron integrals over cartesian gaussian functions. *J Comput Phys* 26: 218-231.
- <sup>50</sup> Jayatilaka D, Grimwood DJ (2003) Tonto: A Fortran Based Object-Oriented System for Quantum Chemistry and Crystallography. In: Sliot PMA, Abramson D, Bogdanov AV, Dongarra JJ, Zomaya AY, Gorbachev YE (eds) *Computational Science - ICCS 2003*, Springer-Verlag, Berlin & Heidelberg, p 142-151.
- <sup>51</sup> Genoni A (2020) On the use of the Obara-Saika recurrence relations for the computation of structure factors in quantum crystallography. *Acta Crystallogr, Sect A* 76: 172-179.

- <sup>52</sup> Obara S, Saika A (1986) Efficient recursive computation of molecular integrals over Cartesian Gaussian functions. *J Chem Phys* 84: 3963-3974.
- <sup>53</sup> Obara S, Saika A (1988) General recurrence formulas for molecular integrals over Cartesian Gaussian functions. *J Chem Phys* 89: 1540-1559.
- <sup>54</sup> Ahlrichs R (2006) A simple algebraic derivation of the Obara-Saika scheme for general two-electron interaction potentials. *Phys Chem Chem Phys* 8: 3072-3077.
- <sup>55</sup> Head-Gordon M, Pople JA (1988) A method for two-electron Gaussian integral and integral derivative evaluation using recurrence relations. *J Chem Phys* 89: 5777-5786.
- <sup>56</sup> Genoni A (2013) Molecular Orbitals Strictly Localized on Small Molecular Fragments from X-ray Diffraction Data. *J Phys Chem Lett* 4: 1093-1099.
- <sup>57</sup> Genoni A (2013) X-ray Constrained Extremely Localized Molecular Orbitals: Theory and Critical Assessment of the New Technique. *J Chem Theory Comput* 9: 3004-3019.
- <sup>58</sup> Dos Santos LHR, Genoni A, Macchi P (2014) Unconstrained and X-ray constrained extremely localized molecular orbitals: analysis of the reconstructed electron density. *Acta Crystallogr, Sect A* 70: 532-551.
- <sup>59</sup> Genoni A, Meyer B (2016) X-Ray Constrained Wave Functions: Fundamentals and Effects of the Molecular Orbitals Localization. *Adv Quantum Chem* 73: 333-362.
- <sup>60</sup> Genoni A (2017) A first-prototype multi-determinant X-ray constrained wavefunction approach: the X-ray constrained extremely localized molecular orbital-valence bond method. *Acta Crystallogr, Sect A* 73: 312-316.
- <sup>61</sup> Casati N, Genoni A, Meyer B, Krawczuk, A, Macchi, P (2017) Exploring charge density analysis in crystals at high pressure: data collection, data analysis and advanced modelling. *Acta Crystallogr, Sect B* 73: 584-597.
- <sup>62</sup> Genoni A, Franchini D, Pieraccini S, Sironi M (2018) X-ray constrained Spin-Coupled wavefunction: a new tool to extract chemical information from X-ray diffraction data. *Chem Eur J* 24: 15507-15511.
- <sup>63</sup> Genoni A, Macetti G, Franchini D, Pieraccini S, Sironi M (2019) X-ray constrained spin-coupled technique: theoretical details and further assessment of the method, *Acta Crystallogr, Sect A* 75: 778-797.
- <sup>64</sup> Hübschle CB, Messerschmidt M, Lentz D, Luger P (2004) Neubestimmung der Ladungsdichte und topologische Analyse von  $\beta$ -Diboran bei 94K. *Z Anorg Allg Chem* 630: 1313–1316.

- <sup>65</sup> Moncol J, Gembicky M, Coppens P (2005) (Tetrahydroborato)bis(triphenylphosphine)copper(I): A redetermination at 90 K. *Acta Cryst. E* 61: m242–m243.
- <sup>66</sup> Fugel M, Kleemiss F, Malaspina LA, Pal R, Spackman PR, Jayatilaka D, Grabowsky S (2018) Investigating the Resonance in Nitric Acid and the Nitrate Anion Based on a Modern Bonding Analysis. *Aust J Chem* 71: 227–237.
- <sup>67</sup> Dolomanov OV, Bourhis LJ, Gildea RJ, Howard, JAK, Puschmann H (2009) OLEX2: a complete structure solution, refinement and analysis program. *J Appl Cryst* 42: 339–341.
- <sup>68</sup> Dovesi R, Erba A, Orlando R, Zicovich-Wilson CM, Civalleri B, Maschio L, Rerat M, Casassa S, Baima J, Salustro S, Kirtman B (2018) Quantum-mechanical condensed matter simulations with CRYSTAL. *WIREs Comput Mol Sci* 8: e1360.
- <sup>69</sup> Shukla A, Dolg M, Stoll H (1998) Wannier-function-based ab initio Hartree-Fock approach extended to polymers: Applications to the LiH chain and trans-polyacetylene. *Phys Rev B* 58: 4325-4334.
- <sup>70</sup> Stoll H, Wagenblast H, Preuß H (1980) On the use of local basis sets for localized molecular orbitals. *Theoret Chim Acta* 57: 169-178.
- <sup>71</sup> Marzari N, Mostofi AA, Yates JR, Souza I, Vanderbilt D (2012) Maximally localized Wannier functions: Theory and applications. *Rev Mod Phys* 84: 1419-1475.
- <sup>72</sup> Wall M. E. (2009) Methods and software for diffuse X-ray scattering from protein crystals. *Methods Mol Biol* 544: 269–279.
- <sup>73</sup> Kresse G, Furthmüller J. (1996) Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys Rev B* 54: 11169–11186.
- <sup>74</sup> Bučinský L, Biskupič S, Jayatilaka D (2012) Study of the picture change error at the 2nd order Douglas Kroll Hess level of theory. Electron and spin density and structure factors of the Bis[bis(methoxycarbimido) aminato] copper (II) complex. *Chem Phys* 395: 44–53.
- <sup>75</sup> Bučinský L, Jayatilaka D, Grabowsky S (2016) Importance of Relativistic Effects and Electron Correlation in Structure Factors and Electron Density of Diphenyl Mercury and Triphenyl Bismuth. *J Phys Chem A* 120: 6650-6669.
- <sup>76</sup> Bučinský L, Jayatilaka D, Grabowsky S (2019) Relativistic Quantum Crystallography of Diphenyl and Dicyano Mercury. Theoretical Structure Factors and Hirshfeld Atom Refinement. *Acta Cryst. A* 75: 705–717.

- <sup>77</sup> Malaspina LA, Wieduwilt EK, Bergmann J, Kleemiss F, Meyer B, Ruiz-López MF, Pal R, Hupf E, Beckmann J, Piltz RO, Edwards AJ, Grabowsky S, Genoni A (2019) Fast and Accurate Quantum Crystallography: from Small to Large, from Light to Heavy. *J Phys Chem Lett* 10: 6973–6982.
- <sup>78</sup> Meyer B, Genoni A (2018) Libraries of Extremely Localized Molecular Orbitals. 3. Construction and Preliminary Assessment of the New Databanks. *J Phys Chem A* 122: 8965–8981.
- <sup>79</sup> Meyer B, Guillot B, Ruiz-Lopez MF, Jelsch C, Genoni A (2016) Libraries of Extremely Localized Molecular Orbitals. 2. Comparison with the Pseudoatoms Transferability. *J Chem Theory Comput* 12: 1068–1081.
- <sup>80</sup> Arias-Olivares D, Wieduwilt EK, Contreras-García J, Genoni A (2019) NCI-ELMO: A New Method To Quickly and Accurately Detect Noncovalent Interactions in Biosystems. *J Chem Theory Comput* 15: 6456-6470.
- <sup>81</sup> Destro R, Marsh RE, Bianchi R (1988) A low-temperature (23 K) study of L-alanine. *J Phys Chem* 92: 966–973.
- <sup>82</sup> Farrugia LJ (1999) WinGX suite for small-molecule single-crystal crystallography. *J Appl Cryst* 32: 837-838.
- <sup>83</sup> Malaspina LA, Genoni A, Jayatilaka D, Turner MJ, Sugimoto K, Nishibori E, Grabowsky S (2019) lamaGOET: an interface for quantum crystallography. In preparation for *J Appl Cryst*.
- <sup>84</sup> Grimme S (2006) Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J Comput Chem* 27: 1787–1799.
- <sup>85</sup> Grabowsky S, Luger P, Buschmann J, Schneider T, Schirmeister T, Sobolev AN, Jayatilaka D (2012) The Significance of Ionic Bonding in Sulfur Dioxide: Bond Orders from X-ray Diffraction Data. *Angew Chem, Int Ed* 51: 6776-6779.
- <sup>86</sup> Wońska M, Jayatilaka D, Dittrich B, Flaig R, Luger P, Woźniak K, Dominiak PM, Grabowsky S (2017) Validation of X-ray Wavefunction Refinement. *ChemPhysChem* 18: 3334–3351.
- <sup>87</sup> Midgley L, Bourhis LJ, Dolomanov O, Peyerimhoff N, Puschmann H (2019) Crystallographic Refinement using Non-Spherical Form Factors in olex2.refine. Under revision in *J Appl Cryst*.
- <sup>88</sup> Wieduwilt EK, Macetti G, Malaspina LA, Jayatilaka D, Grabowsky S, Genoni A (2020) Post-Hartree-Fock methods for Hirshfeld atom refinement: are they necessary? Investigation of a strongly hydrogen-bonded molecular crystal. *J Mol Struct* 1209: 127934.

- <sup>89</sup> Langford JI, Louër D (1996) Powder Diffraction. Rep Prog Phys 59: 131-234.
- <sup>90</sup> Rietveld HM (1967) Line profiles of neutron powder-diffraction peaks for structure refinement. Acta Cryst. 22: 151-152.
- <sup>91</sup> Tolborg K, Jørgensen MRV, Christensen S, Kasai H, Becker J, Walter P, Dippel A-C, Als-Nielsen J, Iversen BB (2017) Accurate charge densities from powder X-ray diffraction—a new version of the Aarhus vacuum imaging-plate diffractometer. Acta Cryst. 73: 521-530.
- <sup>92</sup> Bindzus N, Straasø T, Wahlberg N, Becker J, Bjerg L, Lock N, Dippel A-C, Iversen BB (2014) Experimental determination of core electron deformation in diamond. Acta Cryst. A 70: 39-48.
- <sup>93</sup> Wahlberg N, Bindzus N, Bjerg L, Becker J, Dippel A-C, Iversen BB (2016) Synchrotron powder diffraction of silicon: high-quality structure factors and electron density. Acta Cryst. A 72: 28-35.
- <sup>94</sup> Svane B, Tolborg K, Jørgensen LR, Roelsgaard M, Jørgensen MRV, Iversen BB (2019) Multipole electron densities and atomic displacement parameters in urea from accurate powder X-ray diffraction. Acta Cryst. A 75: 600-609.
- <sup>95</sup> Swaminathan S, Craven BM, McMullan RK (1984) The Crystal Structure and Molecular Thermal Motion of Urea at 12, 60 and 123 K from Neutron Diffraction. Acta Cryst. B 40: 300-306.
- <sup>96</sup> Clinton WL, Massa LJ (1972) Determination of the electron density matrix from x-ray diffraction data. Phys Rev Lett 29: 1363-1366.
- <sup>97</sup> Clinton WL, Frishberg CA, Massa LJ, Oldfield PL (1973) Methods for obtaining an electron density matrix from X-ray diffraction data. Int J Quantum Chem: Quantum Chem Symp 7: 505-514..
- <sup>98</sup> Frishberg C, Massa LJ (1981) Idempotent density matrices for correlated systems from x-ray-diffraction structure factors. Phys Rev B 24: 7018-7024.
- <sup>99</sup> Goldberg MJ, Massa LJ (1983) Quantum model of coherent X-ray diffraction: Extension to Bloch orbitals. Int J Quantum Chem 24: 113-126.
- <sup>100</sup> Massa L, Goldberg M, Frishberg C, Boehme RF, La Placa SJ (1985) Wavefunctions Derived by Quantum Modeling of the Electron Density from Coherent X-Ray Diffraction: Beryllium Metal. Phys Rev Lett 55: 622-625.
- <sup>101</sup> Tanaka K (2018) X-ray molecular orbital analysis. I. Quantum mechanical and crystallographic framework. Acta Crystallogr, Sect A 74: 345-356.

- <sup>102</sup> Hibbs DE, Howard ST, Huke JP, Waller MP (2005) A new orbital-based model for the analysis of experimental molecular charge densities: an application to (Z)-N-methyl-C-phenylnitrone. *Phys Chem Chem Phys* 7: 1772-1778.
- <sup>103</sup> Waller MP, Howard ST, Platts JA, Piltz RO, Willock DJ, Hibbs DE (2006) Novel Properties form Experimental Charge Densities: An Application to the Zwitterionic Neurotransmitter Taurine. *Chem Eur J* 12: 7603-7614.
- <sup>104</sup> Gillet J-M, Becker PJ, Cortona P (2001) Joint refinement of a local wave-function model from Compton and Bragg scattering data. *Phys Rev B* 63: 235115.
- <sup>105</sup> Gillet, J-M, Becker PJ (2004) Position and momentum densities. Complementarity at work: refining a quantum model from different data sets. *J Phys Chem Sol* 65: 2017-2023.
- <sup>106</sup> Gillet J-M (2007) Determination of a one-electron reduced density matrix using a coupled pseudo-atom model and a set of complementary scattering data. *Acta Crystallogr, Sect A* 63: 234-238.
- <sup>107</sup> Guedidda S, Yan Z, Kibalin I, Voufack AB, Claiser N, Souhassou M, Lecomte C, Gillon B, Gillet J-M (2018) Joint refinement model for the spin resolved one-electron reduced density matrix of YTiO<sub>3</sub> using magnetic structure factors and magnetic Compton profiles data. *J Chem Phys* 148: 164106.
- <sup>108</sup> Gilbert TL (1975) Hohenberg-Kohn Theorem for Nonlocal Potentials. *Phys Rev B* 12: 2111-2120.
- <sup>109</sup> Coleman AJ (1963) Structure of Fermion Density Matrices. *Rev Mod Phys* 36: 668-686.
- <sup>110</sup> Henderson GA, Zimmermann RK (1976) One-Electron Properties as Variational Parameters. *J Chem Phys* 65: 619-622.
- <sup>111</sup> Hudák M, Jayatilaka D, Peraínova L, Biskupic S, Kozísek J, Bučinský L (2010) X-ray Constrained Unrestricted Hartree–Fock and Douglas–Kroll–Hess Wavefunctions. *Acta Crystallogr, Sect A* 66: 78–92.
- <sup>112</sup> Bader RFW (1990) *Atoms in Molecules: A Quantum Theory*. Oxford University Press, Oxford, U.K.
- <sup>113</sup> Becke AD, Edgecombe KE (1990) A simple measure of electron localization in atomic and molecular systems. *J Chem Phys* 92: 5397-5403.
- <sup>114</sup> Kohout M (2004) A measure of electron localizability. *Int J Quantum Chem* 97: 651-658.
- <sup>115</sup> Wagner FR, Kohout M, Grin Y (2008) Direct space decomposition of ELI-D: Interplay of charge density and pair-volume function for different bonding situations. *J Phys Chem A* 112: 9814-9828.
- <sup>116</sup> Schmider HL, Becke AD (2000) Chemical content of the kinetic energy density. *J Mol Struct (Theochem)* 527: 51-61.

- <sup>117</sup> Weinhold F, Landis CR (2001) Natural Bond Orbitals and Extensions of Localized Bonding Concepts. *Chem Educ Res Pract* 2: 91-104.
- <sup>118</sup> Glendening ED, Weinhold F (1998) Natural Resonance theory. I. General Formalism. *J Comput Chem* 19: 593-609.
- <sup>119</sup> Jayatilaka D, Grimwood, D (2004) Electron localization functions obtained from X-ray constrained Hartree-Fock wavefunctions for molecular crystals of ammonia, urea and alloxan. *Acta Crystallogr, Sect A* 60: 111-119.
- <sup>120</sup> Grabowsky S, Jayatilaka D, Mebs S, Luger P (2010) The Electron Localizability Indicator from X-Ray Diffraction Data – A First Application to a Series of Epoxide Derivatives. *Chem Eur J* 16: 12818-12821.
- <sup>121</sup> Grabowsky S, Weber M, Jayatilaka D, Chen Y-S, Grabowski MT, Brehme R, Hesse M, Schirmeister T, Luger P (2011) Reactivity Differences between  $\alpha,\beta$ -Unsaturated Carbonyls and Hydrazones Investigated by Experimental and Theoretical Electron Density and Electron Localizability Analyses. *J Phys Chem A* 115: 12715-12732.
- <sup>122</sup> Fugel M, Beckmann J, Jayatilaka D, Gibbs GV, Grabowsky S (2018) A Variety of Bond Analysis Methods, One Answer? An Investigation of the Element-Oxygen Bond of Hydroxides  $H_nXOH$ . *Chem Eur J* 24: 6248-6261.
- <sup>123</sup> Hübschle CB, Luger, P (2006) *MolIso* – a program for colour-mapped iso-surfaces. *J Appl Crystallogr* 39: 901–904.
- <sup>124</sup> Fornili A, Sironi M, Raimondi M (2003) Determination of extremely localized molecular orbitals and their application to quantum mechanics/molecular mechanics methods and to the study of intramolecular hydrogen bonding. *J Mol Struct (THEOCHEM)* 632: 157–172.
- <sup>125</sup> Genoni A, Sironi M (2004) A Novel Approach to Relax Extremely Localized Molecular Orbitals: the Extremely Localized Molecular Orbital-Valence Bond Method. *Theor Chem Acc* 112: 254-262.
- <sup>126</sup> Genoni A, Fornili A, Sironi M (2005) Optimal Virtual Orbitals to Relax Wavefunctions Built Up with Transferred Extremely Localized Molecular Orbitals. *J Comput Chem* 26: 827-835.
- <sup>127</sup> Genoni A, Ghitti M, Pieraccini S, Sironi M (2005) A novel extremely localized molecular orbitals based technique for the one-electron density matrix computation. *Chem Phys Lett* 415: 256-260.
- <sup>128</sup> Sironi M, Genoni A, Civera M, Pieraccini S, Ghitti M (2007) Extremely Localized Molecular Orbitals: Theory and Applications. *Theor Chem Acc* 117: 685-698.

- <sup>129</sup> Sironi M, Ghitti M, Genoni A, Saladino G, Pieraccini S (2009) DENPOL: A new program to determine electron densities of polypeptides using extremely localized molecular orbitals. *J Mol Struct (THEOCHEM)* 898: 8-16.
- <sup>130</sup> Meyer B, Guillot B, Ruiz-Lopez MF, Genoni A (2016) Libraries of Extremely Localized Molecular Orbitals. 1. Model Molecules Approximation and Molecular Orbitals Transferability. *J Chem Theory Comput* 12: 1052-1067.
- <sup>131</sup> Chirgwin BH, Coulson CA (1950) The electronic structure of conjugated systems. IV. *Proc R Soc London A* 201: 196-209.
- <sup>132</sup> Casati N, Kleppe A, Jephcoat A, Macchi P (2016) Putting pressure on aromaticity along with in situ experimental electron density of a molecular crystal. *Nat Commun* 7: 10901.
- <sup>133</sup> Cooper DL, Gerratt J, Raimondi M (1991) Applications of Spin-Coupled Valence Bond Theory. *Chem Rev* 91: 929-964.
- <sup>134</sup> Cooper DL, Gerratt J, Raimondi M (1986) The electronic structure of the benzene molecule. *Nature* 323: 699-701.
- <sup>135</sup> Guest MF, Bush IJ, van Dam HJJ, Sherwood P, Thomas JMH, van Lenthe JH, Havenith RWA, Kendrick J (2005) The GAMESS-UK Electronic Structure Package: Algorithms, Developments and Applications. *Mol Phys* 103: 719-747.
- <sup>136</sup> Cooper DL, Gerratt J, Raimondi M, Sironi M, Thorsteinsson T (1993) Expansion of the spin-coupled wavefunction in Slater determinants. *Theor Chim Acta* 85: 261-270.
- <sup>137</sup> Goldfield S, Quandt R, Trotter F (1966) Maximization by quadratic hill climbing. *Econometrica* 34: 541-551.
- <sup>138</sup> Bytheway I, Chandler G, Figgis B, Jayatilaka D (2007) Are intramolecular dynamic electron correlation effects detectable in X-ray diffraction experiments on molecular crystals? *Acta Crystallogr, Sect A* 63: 135-145.
- <sup>139</sup> Genoni A, Dos Santos LHR, Meyer B, Macchi P (2017) Can X-ray constrained Hartree-Fock wavefunctions retrieve electron correlation? *IUCrJ* 4: 136-146.
- <sup>140</sup> Hupf E, Kleemiss F, Borrmann T, Pal R, Krzeszczakowska JM, Woińska M, Dominiak P, Genoni A, Bučinský L, Bürgi H-B, Jayatilaka D, Grabowsky S (2019) The Effects of Electron Correlation and Polarization Extracted from Experimental X-ray Diffraction Data. In preparation.

- <sup>141</sup> Hickstein DD, Cole JM, Turner MJ, Jayatilaka D (2013) Modeling Electron Density Distributions from X-Ray Diffraction to Derive Optical Properties: Constrained Wavefunction Versus Multipole Refinement. *J Chem Phys* 139: 064108.
- <sup>142</sup> Whitten AE, Jayatilaka D, Spackman M (2006) Effective Molecular Polarizabilities and Crystal Refractive Indices Estimated from X-Ray Diffraction Data. *J Chem Phys* 125: 174505.
- <sup>143</sup> Deutsch M, Claiser N, Pillet S, Chumakov Y, Becker P, Gillet J-M, Gillon B, Lecomte C, Souhassou M (2012) Experimental determination of spin-dependent electron density by joint refinement of X-ray and polarized neutron diffraction data. *Acta Crystallogr, Sect A* 68: 675–686.
- <sup>144</sup> Deutsch M, Gillon B, Claiser N, Gillet J-M, Lecomte C, Souhassou M (2014) First spin-resolved electron distributions in crystals from combined polarized neutron and X-ray diffraction experiments. *IUCrJ* 1: 194–199.
- <sup>145</sup> Voufack AB, Claiser N, Lecomte C, Pillet S, Pontillon Y, Gillon B, Yan Z, Gillet J-M, Marazzi M, Genoni A, Souhassou M (2017) When combined X-ray and polarized neutron diffraction data challenge high-level calculations: spin-resolved electron density of an organic radical. *Acta Crystallogr, Sect B* 73: 544-549.
- <sup>146</sup> Gatti C, Orlando AM, Lo Presti L (2015) Insights on Spin Polarization through the Spin Density Source Function. *Chem Sci* 6: 3845-3852.
- <sup>147</sup> Gatti C, Macetti G, Lo Presti L (2017) Insights on spin delocalization and spin polarization mechanisms in crystals of azido copper(II) dinuclear complexes through the electron spin density Source Function. *Acta Crystallogr, Sect B* 73: 565-583.
- <sup>148</sup> Macetti G, Lo Presti L, Gatti C (2018) Spin density accuracy and distribution in azido Cu(II) complexes: a source function analysis. *J Comput Chem* 39: 587-603.
- <sup>149</sup> Gould MD, Taylor C, Wolff SK, Chandler GS, Jayatilaka D (2008) A definition for the covalent and ionic bond index in a molecule. *Theor Chem Acc* 119: 275-290.
- <sup>150</sup> Allen FH, Bruno IJ (2010) Bond lengths in organic and metal-organic compounds revisited: X-H bond lengths from neutron diffraction data. *Acta Crystallogr, Sect B* 66: 380-386.
- <sup>151</sup> Köhler C, Lübben J, Krause L, Hoffmann C, Herbst-Irmer R, Stalke D (2019) Comparison of different strategies for modelling hydrogen atoms in charge density analyses. *Acta Crystallogr, Sect B* 75: 434-441.
- <sup>152</sup> Madsen AØ (2006) SHADE web server for estimation of hydrogen anisotropic displacement parameters. *J Appl Crystallogr* 39: 757-758.

- <sup>153</sup> Lübben J, Bourhis LJ, Dittrich B (2015). Estimating temperature-dependent anisotropic hydrogen displacements with the Invariom database and a new segmented rigid-body analysis program. *J Appl Crystallogr* 48: 1785-1793.
- <sup>154</sup> Dittrich B, Hübschle CB, Pröpper K, Dietrich F, Stolper T, Holstein JJ (2013) The Generalized Invariom Database (GID). *Acta Crystallogr, Sect B* 69: 91-104.
- <sup>155</sup> Hummel W, Hauser J, Bürgi H-B (1990) PEANUT: computer graphics program to represent atomic displacement parameters. *J Mol Graph* 8: 214–220.
- <sup>156</sup> Pavan MS, Sarkar S, Guru Row TN (2017) Exploring the rare S—H···S hydrogen bond using charge density analysis in isomers of mercaptobenzoic acid. *Acta Crystallogr, Sect B* 73: 626-633.
- <sup>157</sup> Malaspina LA, Edwards AJ, Woińska M, Jayatilaka D, Turner MJ, Price JR, Herbst-Irmer R, Sugimoto K, Nishibori E, Grabowsky S (2017) Predicting the position of the hydrogen atom in the short intramolecular hydrogen bond of the hydrogen maleate anion from geometric correlations. *Cryst Growth Des* 17: 3812-3825.
- <sup>158</sup> Malaspina LA, Hoser AA, Edwards AJ, Woińska M, Turner MJ, Price JR, Sugimoto K, Nishibori E, Bürgi H-B, Jayatilaka D, Grabowsky S (2019) Obtaining hydrogen atom parameters from X-ray data: Hirshfeld atom refinement of hydrogen atoms in bridging positions. In preparation for *Cryst Eng Comm*.
- <sup>159</sup> Thomas SP, Grosjean A, Flematti GR, Karton A, Sobolev AN, Edwards AJ, Piltz RO, Iversen BB, Koutsantonis GA, Spackman MA (2019) Investigation of an unusual crystal habit of hydrochlorothiazide reveals large polar enantiopure domains and a possible crystal nucleation mechanism. *Angew Chem Int Ed* 58: 10255-10259.
- <sup>160</sup> Thomas SP, Malaspina LA, Nishibori E, Grabowsky S, Koutsantonis GA, Iversen BB, Spackman MA (unpublished results).
- <sup>161</sup> Sarkar S, Thomas SP, Lokeswara Rao P, Edwards A, Grosjean A, Ramanathan KV, Guru Row TN (2019) Experimental insights into the electronic nature, spectral features, and role of entropy in short CH<sub>3</sub>···CH<sub>3</sub> hydrophobic interactions. *J Phys Chem Lett* 10: 7224-7229.
- <sup>162</sup> Aree T, Bürgi H-B, Chernyshov D, Törnroos KW (2014) Dynamics and thermodynamics of crystalline polymorphs. 3.  $\gamma$ -glycine, analysis of variable-temperature atomic displacement parameters, and comparison of polymorph stabilities. *J Phys Chem A* 118: 9951-9959.
- <sup>163</sup> Kofoed PM, Hoser AA, Diness F, Capelli SC, Madsen AO (2019) X-ray diffraction data as a source of the vibrational free-energy contribution in polymorphic systems. *IUCrJ* 6: 558-571.

- <sup>164</sup> Jayatilaka D, Munshi P, Turner MJ, Howard JAK, Spackman MA (2009) Refractive indices for molecular crystals from the response of X-ray constrained Hartree–Fock wavefunctions. *Phys Chem Chem Phys* 11: 7209-7218.
- <sup>165</sup> Cole JM, Hickstein DD (2013) Molecular origins of nonlinear optical activity in zinc tris(thiourea)sulfate revealed by high-resolution x-ray diffraction data and ab initio calculations. *Phys Rev B* 88: 184105.
- <sup>166</sup> Fugel M, Malaspina LA, Pal R, Thomas SP, Shi MW, Spackman MA, Sugimoto K, Grabowsky S (2019) Revisiting a historical concept by using quantum crystallography: Are phosphate, sulfate and perchlorate anions hypervalent? *Chem. Eur. J.* 25: 6523-6532.
- <sup>167</sup> Yanai H, Suzuki T, Kleemiss F, Fukaya H, Dobashi Y, Malaspina LA, Grabowsky S, Matsumoto T (2019) Chemical bonding in polarised push–pull ethylenes. *Angew Chem, Int. Ed.* 58: 8839-8844.
- <sup>168</sup> Thomas SP, Satheeshkumar K, Mugesh G, Guru Row TN (2015) Unusually short chalcogen bonds involving organoselenium: Insights into the Se–N bond cleavage mechanism of the antioxidant ebselen and analogues. *Chem Eur J* 21: 6793-6800.
- <sup>169</sup> Thomas SP, Malaspina LA, Alhameedi K, Mugesh G, Jayatilaka D, Guru Row TN, Grabowsky S (unpublished results).
- <sup>170</sup> Thomas SP, Jayatilaka D, Guru Row TN (2015) S···O chalcogen bonding in sulfa drugs: insights from multipole charge density and X-ray wavefunction of acetazolamide. *Phys Chem Chem Phys* 17: 25411-25420.
- <sup>171</sup> Alhameedi K, Karton A, Jayatilaka D, Thomas SP (2018) Bond orders for intermolecular interactions in crystals: charge transfer, ionicity and the effect on intramolecular bonds. *IUCrJ* 5: 635-646.
- <sup>172</sup> Turner MJ, Grabowsky S, Jayatilaka D, Spackman MA (2014) Accurate and efficient model energies for exploring intermolecular interactions in molecular crystals. *J Phys Chem Lett* 5: 4249-4255.