



Doctoral Thesis

The three-dimensional pair distribution function method for analyzing single crystal diffuse scattering - theory, software development and application

Author(s):

Simonov, Arkadiy

Publication Date:

2014

Permanent Link:

<https://doi.org/10.3929/ethz-a-010266429> →

Rights / License:

[In Copyright - Non-Commercial Use Permitted](#) →

This page was generated automatically upon download from the [ETH Zurich Research Collection](#). For more information please consult the [Terms of use](#).

DISS. ETH NO. 21974

***THE THREE-DIMENSIONAL PAIR DISTRIBUTION FUNCTION
METHOD FOR ANALYZING SINGLE CRYSTAL DIFFUSE
SCATTERING –THEORY, SOFTWARE DEVELOPMENT AND
APPLICATION.***

A thesis submitted to attain the degree of
DOCTOR OF SCIENCES of ETH ZURICH

presented by

ARKADIY SIMONOV

Dipl. phys. Lomonosov Moscow State University, Russia

born on 15.09.1986

citizen of Russia

accepted on the recommendation of

Prof. Dr. Walter Steurer, ETH Zürich
Prof. Dr. Reinhard Neder, University of Erlangen
Prof. Dr. Bernd Schönfeld, ETH Zürich
Dr. Thomas Weber, ETH Zürich

2014

Contents

Contents	2
Abstract	5
Zusammenfassung	7
1 Introduction	9
1.1 Authors declaration	10
2 Theory of diffuse scattering analysis	11
2.1 Diffuse scattering and the Three-Dimensional Pair Distribution Function	11
2.2 Classification of disorder	12
2.3 Methods of diffuse scattering analysis	12
3 3D-ΔPDF refinement method	17
3.1 Fundamentals of 3D- Δ PDF analysis	17
3.2 The program YELL	28
4 Examples	37
4.1 Complex metallic alloy $hP386\text{-Al}_{57.4}\text{Cu}_{3.5}\text{Ta}_{39.0}$	37
4.2 $\text{Ge}_4\text{Bi}_2\text{Te}_7$	40
4.3 Tris-t-butyl tricarboxamide	49
4.4 PbTe	72
5 Conclusions and outlook	77
5.1 Comparison of the 3D- Δ PDF approach with other diffuse scattering methods	77
5.2 Summary and outlook	79
A Script for generating YELL input file for $\text{Ge}_4\text{Bi}_2\text{Te}_7$	81
B YELL reference manual	85
Bibliography	131
Acknowledgments	137

CONTENTS

3

Curriculum Vitae

139

Abstract

In order to understand the properties of materials the determination of their structure is essential. If a material can be grown as a well-ordered single crystal, its structure can be fully determined with conventional crystallographic methods based on Bragg scattering. Such methods provide the average structure, a mathematical model in which the crystal consists of a perfectly periodic lattice of identical unit cells. However, many important materials depart from this ideal. Due to disorder, their real structure locally deviates from the average structure. The structure of such materials can be investigated using diffuse scattering, a weak structured signal which appears in addition to Bragg peaks in the diffraction images. Although a number of methods have been developed over the years to analyze diffuse scattering from single crystals, all have their limitations. Contrary to single crystals, diffuse scattering from powder samples can be routinely interpreted using the Pair Distribution Function (PDF) analysis. In this approach the total scattering from a powder sample is Fourier transformed to extract information about the distribution of pairs of atoms in the real structure.

This work continues the efforts of Miroslav Kobas and Philippe Schaub, former PhD students from the Laboratory of Crystallography at ETH Zürich, who have shown the three dimensional extension of powder PDF, the 3D-PDF, and especially three dimensional *difference* pair distribution function (3D- Δ PDF) can be used for qualitative and quantitative interpretation of single crystal disorder. The aim of the project described in this thesis was to extend this result and develop methods and tools for routine 3D- Δ PDF analysis. It was shown that without loss of generality the short range order can be described using only three basic types of correlations: substitutional correlation, which describes dependencies between occupancies of disordered sites, size-effect, which describes the relaxation around disordered sites, and atomic displacement correlations. The numerical values of such correlations can be obtained through least squares refinements. The 3D- Δ PDF approach can efficiently be applied not only on the level of atoms, but also on the level of molecules.

In the course of this project the program YELL for performing 3D- Δ PDF refinements was designed and implemented using C++ programming language. The program is thoroughly documented and released for Mac and Windows operating systems. YELL is a free software and its source code is available under the GPL license.

The 3D- Δ PDF approach was tested on diffuse scattering from four crystals. In the first case of a complex intermetallic compound hP386-Al_{57.4}Cu_{3.5}Ta_{39.0} the quality of experiment did not allow to extract quantitative diffuse scattering profiles. However, the model of disorder derived from the average structure could qualitatively reproduce observed diffuse scattering. The diffuse scattering from Ge₄Bi₂Te₇, tris-*t*-butyl-1,3,5-benzene tricarboxamide and PbTe could be successfully analyzed and short range order models could be quantitatively refined with YELL. Furthermore, the tris-*t*-

butyl-1,3,5-benzene tricarboxamide example was used to assess the reliability of single crystal diffuse scattering analysis. It was shown that 3D- Δ PDF can provide excellent accuracy which is mostly determined by the quality of data reduction, in particular compensation for experimental systematic errors coming from background scattering, the resolution function and geometrical distortions.

Zusammenfassung

Für das Verständnis der Eigenschaften von Materialien ist die Kenntnis ihrer Strukturen von entscheidender Bedeutung. Für den Fall, dass Materialien als wohlgeordnete Einkristalle verfügbar sind, können deren Strukturen vollständig mit konventionellen kristallographischen Methoden anhand der Braggstreuung bestimmt werden. Die Ergebnisse solcher Untersuchungen ist die mittlere Struktur, also ein mathematisches Modell, das den Kristall als ein perfektes periodisches Gitter identischer Einheitszellen beschreibt. Viele wichtige Materialien weichen jedoch von diesem Idealbild ab. Aufgrund von Fehlordnung ist deren reale Struktur auf lokaler Ebene nicht mit der mittleren Struktur identisch. Die strukturellen Eigenschaften solcher Materialien können mit Hilfe der diffusen Streuung, welches ein schwaches strukturiertes Signal ist, das zusätzlich zu Braggreflexen in einem Streubild beobachtet werden kann, untersucht werden. Zwar wurden im Laufe der Jahre eine Vielzahl von Methoden zur Untersuchung diffuser Streuung von Einkristallen entwickelt, jedoch sind alle mit Einschränkungen verbunden. Im Gegensatz zu Einkristallen wird die diffuse Streuung von Pulverproben routinemässig mit der Paarverteilungsmethode (*Pair Distribution Function*, PDF) analysiert. Bei diesem Ansatz wird die totale Streuung einer Pulverprobe fouriertransformiert, um die Information über die Verteilung der atomaren Paare in der Realstruktur zu erhalten.

Diese Dissertation führt die Arbeiten von früheren Doktoranden am Labor für Kristallographie an der ETH Zürich, Miroslav Kobas und Philippe Schaub, fort. Sie konnten zeigen, dass die Erweiterung der Pulver PDF auf drei Dimensionen (3D-PDF) und insbesondere die dreidimensionale Differenzpaarverteilungsfunktion (3D- Δ PDF) geeignet sind um qualitative und quantitative Untersuchungen von Fehlordnung in Einkristallen durchzuführen. Die Absicht des in dieser Dissertation beschriebenen Projekts war es die bisherigen Ergebnisse zu erweitern und neue Methoden und Werkzeuge für den Routinegebrauch der 3D- Δ PDF zu entwickeln. Es konnte gezeigt werden, dass es möglich ist lokale Ordnung ohne Verlust der Allgemeinheit durch nur drei fundamentale Korrelationstypen zu beschreiben, nämlich der substitutionellen Korrelation, welche die Abhängigkeit zwischen den Besetzungen verschiedener fehlgeordneter Lagen beschreibt, den Grösseneffekt, der die Relaxationen um die fehlgeordneten Lagen repräsentiert und Korrelationen zwischen den Verschiebungen von Atomen. Die Quantifizierung der Nahordnungskorrelationen kann beispielsweise über die Methode der kleinsten Fehlerquadrate erfolgen. Die 3D- Δ PDF Methode kann über die atomare Ebene hinaus auch effizient für die Untersuchung fehlgeordneter Molekularstrukturen eingesetzt werden.

Im Verlaufe dieses Projektes wurde das Computerprogramm YELL zur Ausführung von 3D- Δ PDF Verfeinerungen entwickelt und in der Programmiersprache C++ implementiert. Das Programm ist umfangreich dokumentiert und für Mac und Windows Betriebssysteme verfügbar. YELL ist ein freies Programm und steht mit seinem Programmcode unter der GPL Lizenz der Allgemeinheit zur Verfügung.

Der 3D- Δ PDF Ansatz wurde anhand der diffusen Streuung von vier Kristallen getestet. Im Fall der komplexen intermetallischen Verbindung hP386- $\text{Al}_{57,4}\text{Cu}_{3,5}\text{Ta}_{39,0}$ war es aufgrund der Qualität der experimentellen Daten nicht möglich quantitative diffuse Streuprofile zu extrahieren. Die beobachtete diffuse Streuung konnte jedoch mit einem aus der Interpretation der mittleren Struktur abgeleiteten Fehlordnungsmodell qualitativ reproduziert werden. Die diffuse Streuung von $\text{Ge}_4\text{Bi}_2\text{Te}_7$, Tris-t-butyl-1,3,5-benzol Tricarboxamid und PbTe konnten ebenfalls erfolgreich untersucht und deren Nahordnungsparameter mit YELL quantitativ verfeinert werden. Darüber hinaus wurde am Beispiel von Tris-t-butyl-1,3,5-benzol Tricarboxamids die Zuverlässigkeit der Ergebnisse von Einkristallrealstrukturanalysen basierend auf diffuser Streuung diskutiert. Es konnte gezeigt werden, dass mit der 3D- Δ PDF Methode eine exzellente Genauigkeit erzielt werden kann, die im Wesentlichen von der Qualität der Datenreduktion, insbesondere der Korrektur systematischer experimenteller Fehler durch Hintergrundstreuung, Auflösungsfunktion und geometrischer Verzerrungen, bestimmt wird.

Chapter 1

Introduction

In order to understand the properties of materials it is very important to know their crystal structures. If a material can be grown as a well-ordered single crystal, its structure can be fully determined with conventional crystallographic methods based on Bragg scattering. However, many crystals with interesting properties are disordered. For such crystals Bragg crystallography provides only the average structure – a projection of the whole structure into a single unit cell.

The real structure of a crystal contains significantly more information than the average structure. Thermal displacements of all atoms in a crystal are not independent and typically reflect the energy distribution of phonons which govern crystal dynamics. The composition of unit cells of substitutionally disordered crystals is also correlated with the composition of the neighboring unit cells. The careful analysis of subtle details of the real structure proved invaluable in explaining important properties such as ferroelectricity [1], ion-conductivity [2], giant magneto-resistivity [3], catalytic activity [4], anomalous lattice vibrations [5], behavior of host-guest systems [6], phasonic flips in quasicrystals [7] and dynamics in proteins [8] to name just a few.

Information about the real structure can be extracted from the diffuse scattering of a diffraction experiment. Diffuse scattering is one of the best techniques for probing the real structure of materials, because it is sensitive to both static and dynamic ordering. Thanks to the availability of powerful software [9] and the ease of data collection, diffuse scattering analysis has been widely adopted for investigating short range order of powders and nano materials. Such analysis typically utilizes the Pair Distribution Function, a Fourier transform of the total scattering from the sample. A number of approaches summarized later in section 2.3 also exists for single crystals, but due to various limitations, no simple and fast routine method is currently available.

The aim of this thesis was to generalize the successful powder Pair Distribution Function method to three dimensional single crystal data and to develop the program YELL, which could perform such refinements. The program was implemented and successfully tested on diffuse scattering data from the inorganic compounds $\text{Ge}_4\text{Bi}_2\text{Te}_7$ and PbTe , the complex metallic alloy $\text{hP386-Al}_{57.4}\text{Cu}_{3.5}\text{Ta}_{39.0}$ and the organic compound tris-*t*-butyl-1,3,5-benzene tricarboxamide.

The thesis is organized as follows. The chapter 2 provides the overview of diffuse scattering analysis methods. The chapter 3 which includes the papers [10] and [11], describes the theory of the 3D- Δ PDF refinement. The first paper introduces the 3D- Δ PDF formalism and provides equations to calculate diffuse scattering from individual interatomic pairs. The second paper expands the for-

malism to pairs of molecules and describes the implementation of the program YELL for performing diffuse scattering refinements. The chapter 4 presents the 3D- Δ PDF analysis of diffuse scattering of four single crystals: hP386-Al_{57.4}Cu_{3.5}Ta_{39.0}, Ge₄Bi₂Te₇, tris-t-butyl-1,3,5-benzene tricarboxamide and PbTe. In addition, the section about diffuse scattering from tris-t-butyl-1,3,5-benzene tricarboxamide also assesses the accuracy of 3D- Δ PDF refinements. The last chapter provides the comparison between 3D- Δ PDF and other methods for diffuse scattering analysis.

1.1 Authors declaration

The development of 3D- Δ PDF least squares refinement and the theory described in [10] (section 3.1), the implementation of the program YELL, and diffuse scattering analysis of the Tris-t-butyltricarboxamide described in section 4.3 were performed by myself. Interpretation of diffuse scattering from hP386-Al_{57.4}Cu_{3.5}Ta_{39.0}, Ge₄Bi₂Te₇ and PbTe was done in close collaboration with Julia Dshemuchadse and Sandro Bigler, Philipp Urban and Oliver Oeckler, and Thomas Weber respectively. My contribution in these projects can be summarized as follows:

hP386-Al_{57.4}Cu_{3.5}Ta_{39.0} I proposed and executed the calculation of the inter-structure difference electron densities map. I also performed diffuse scattering reconstructions and the calculated the model diffuse scattering. Crystal synthesis, data collection, average structure solutions and crystal chemical interpretations were done by Sandro Bigler and Julia Dshemuchadse.

Ge₄Bi₂Te₇ I performed diffuse scattering data reduction, developed and implemented the input file generator and helped with model building and refinement. Philipp Urban and Oliver Oeckler prepared the crystal, collected diffuse scattering, refined the average structure, and finalized the diffuse scattering refinement.

PbTe I helped to develop the model and to prepare the YELL input file. All other work was done by Thomas Weber.

Chapter 2

Theory of diffuse scattering analysis

2.1 Diffuse scattering and the Three-Dimensional Pair Distribution Function

Diffuse scattering is all scattering from a single crystal which is not Bragg scattering. Typically diffuse scattering shows up in the shape of lines, planes or diffuse clouds between or underneath Bragg peaks.

In kinematic approximation, the coherent scattering from a single coherently scattering crystallite can be expressed as the Fourier transform of its electron density $\rho(\mathbf{r})$ [12]:

$$F(\mathbf{h}) = \text{FT}[\rho(\mathbf{r})]$$

Most of the diffraction experiments are performed using single crystals, which contain many incoherently scattering mosaic blocks. The x-ray detector captures the average scattering intensity from such mosaic blocks:

$$I(\mathbf{h}) = \langle F(\mathbf{h})F^*(\mathbf{h}) \rangle = \text{FT}[\langle \rho(\mathbf{r}) \star \rho(\mathbf{r}) \rangle]$$

here the brackets $\langle \rangle$ denote averaging over exposure time and scattering mosaic blocks, the star \star denotes cross-correlation and the term $\rho(\mathbf{r}) \star \rho(\mathbf{r})$ is the autocorrelation function of the electron density.

The structure factor of the Bragg peaks is calculated as the Fourier transform of the average structure:

$$F_{Bragg}(hkl) = \text{FT}[\langle \rho(\mathbf{r}) \rangle]$$

The diffuse scattering is the difference between total and Bragg scattering:

$$\begin{aligned} I_{diffuse}(\mathbf{h}) &= I(\mathbf{h}) - I_{Bragg}(\mathbf{h}) = I(\mathbf{h}) - F_{Bragg}(\mathbf{h})F_{Bragg}^*(\mathbf{h}) \\ &= \text{FT}[\langle \rho(\mathbf{r}) \star \rho(\mathbf{r}) \rangle] - \text{FT}[\langle \rho(\mathbf{r}) \rangle \star \langle \rho(\mathbf{r}) \rangle] = \\ &= \text{FT}[P_{tot}(\mathbf{r}) - P(\mathbf{r})] = \\ &= \text{FT}[P_{\Delta}(\mathbf{r})] \end{aligned}$$

The autocorrelation of the average structure $P(\mathbf{r}) = \langle \rho(\mathbf{r}) \rangle \star \langle \rho(\mathbf{r}) \rangle$ is called Patterson function, the average autocorrelation of the real structure $P_{tot}(\mathbf{r}) = \langle \rho(\mathbf{r}) \star \rho(\mathbf{r}) \rangle$ is called Pair Distribution Function, the difference between the two $P_{\Delta}(\mathbf{r}) = P_{tot}(\mathbf{r}) - P(\mathbf{r})$ is called Difference Pair Distribution Function, or 3D- Δ PDF.

The Patterson function contains the information about interatomic vectors in the average structure. For each pair of atoms i and j the Patterson function will contain a peak at the interatomic vector $\mathbf{r}_{ij} = \mathbf{r}_j - \mathbf{r}_i$. The pair distribution function describes the interatomic vectors in the real structure. Consequently, the 3D- Δ PDF describes order in the crystal which is not captured by the average structure. 3D- Δ PDF can take both positive and negative values. Positive values mean that corresponding interatomic vectors appear more frequently in the real structure than in the average structure, while negative values mean the opposite.

2.2 Classification of disorder

It is useful to classify disorder according to its dimensionality. If the real structure of a disordered crystal is well ordered along two dimensions, and disordered only along one dimension, the crystal it is said to have *1D disorder*. The crystal is said to have *2D disorder* if its real structure is ordered along one dimension and disordered along two dimensions. In the general case where real structure is not periodic along any dimension, it is said to have *3D disorder* [13].

The dimensionality of disorder can be easily deduced from the shape of diffuse scattering. If the real structure is well ordered along some dimension with a periodicity vector \mathbf{p} , the Fourier transform of its electron density (and thus $I(\mathbf{s})$) will be concentrated in a set of sharp features that meet the condition $(\mathbf{p}\mathbf{s}) = n$ where n is integer¹. Diffuse scattering from 1D disorder is present as sharp streaks, diffuse scattering from 2D disorder forms layers, and diffuse scattering from 3D disorder is present in as broad features (Fig. 2.2.1).

One more commonly used distinction is between static and dynamic disorder. Disorder is called *static* if the atomic configuration of the crystal does not change in the course of the experiment, otherwise it is called *dynamic*. The distinction corresponds to the formalism used for describing disorder. Dynamic disorder may be described by lattice dynamics which is usually expressed in terms of phonons. Static disorder implies substitutional disorder or static displacements and typically involves discrete (present/absent) variables in the model, possibly along with continuous variables describing relaxation. The border between static and dynamic disorder is sometimes vague since the (pseudo-) elastic diffuse x-ray experiments do not deliver direct information about the static or dynamic origin of disorder. Inelastic x-ray or neutron experiments are required to resolve this issue.

2.3 Methods of diffuse scattering analysis

There exist a number of approaches for diffuse scattering investigations, which in the following are categorized in five groups.

¹In the case of modulated structures the sharp features could be indexed by several modulation vectors.

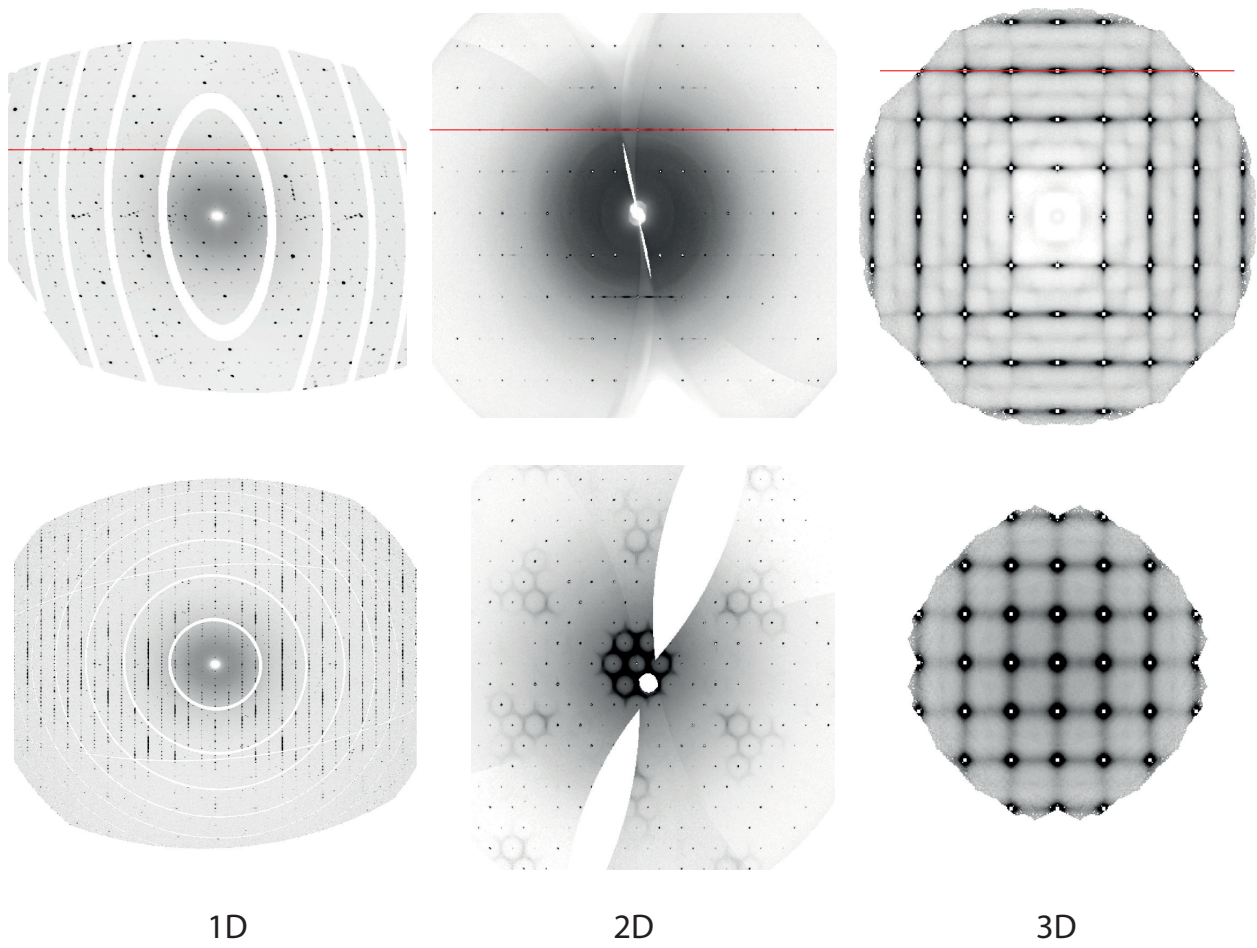


Figure 2.2.1: Examples of diffuse scattering of different dimensionality. The red line marks the crosssections of the layers presented in the lower row. The 1D, 2D and 3D diffuse scattering come from hP386-Al_{57.4}Cu_{3.5}Ta_{39.0} (section 4.1), tris-t-butyltricarboxamide (section 4.3) and PbTe (section 4.4) respectively.

Direct modeling

Currently, the most widely used approach for diffuse scattering analysis employs computer simulation of disordered crystals. First, a user defines a model of disorder in terms of direct space rules. Then a representative sample of the crystal such that it fulfills the rules is equilibrated in the computer. Finally, the resulting structure is Fourier transformed in order to obtain the model diffuse scattering, which is then compared to the experimental data. The parameters of the model can be adjusted to obtain the best fit between the model and the experiment.

Typically the modeling employs Monte-Carlo (MC) simulations. The adjustment of the parameters is performed using least squares refinement [14] or, more frequently, using sophisticated global minima finding algorithms like differential evolution [15] or swarm minimization [16], which are stable in the presence of sampling noise in diffuse scattering caused by the limited size of the calculated model.

The direct modeling approach is very flexible and allows to investigate any kind of disorder. The approach has two drawbacks, however. First, it requires one to guess a model that correctly describes the disorder at hand. Construction of such models is based on the shape and distribution of diffuse scattering and requires expert knowledge and significant experience from the user. Second, the refinements are computationally very demanding. Though they can be easily parallelized on a supercomputer [16], refinements sometimes require weeks of wall-clock time to converge.

There exist a number of packages which can perform direct modeling. The most widely used is the program DISCUS [17].

Reverse Modeling

Another approach, called Reverse Monte-Carlo (RMC) simulation, is similar to MC, but utilizes experimental diffuse scattering instead of structure related rules for building the model crystal. The algorithm starts with a random model which represents the average structure of a crystal and introduces random changes to the model trying to minimize the difference between calculated and experimental diffuse scattering.

This approach is flexible and converges much quicker than MC refinements. However, RMC models are significantly overparameterized and prone to fitting experimental artifacts along with true diffuse scattering signal. To avoid this, RMC models typically include additional constraints that prevent simulation of unphysical crystals.

RMC simulations can be performed e.g. in the programs RMCProfile [18], RMC++ [19] or DISCUS [17].

Normal mode analysis

Diffuse scattering from lattice vibrations (thermal diffuse scattering) can be calculated on the basis of phonon dispersion curves [20], which can be obtained from ab-initio calculations or inelastic neutron or x-ray scattering. The lowest part of acoustic branches, which is responsible for the majority of thermal diffuse scattering could also be obtained from mechanical properties of the crystal [1].

Whether it is possible to reconstruct dispersion curves from diffuse scattering alone, without the requirement of complicated ab-initio modeling, remains an open research question. Despite the very impressive study by Holt et. al. [21] which showed the possibility to refine the interatomic

force constants directly from diffuse scattering in silicon, this method was not generalized to other systems.

Analytical methods

Historically, diffuse scattering was analyzed by solving equations describing diffuse scattering analytically. Though relatively complicated to derive, they provide closed equations that describe diffuse scattering. A number of simple examples are presented in the text book by Cowley [12], a more recent example can be found in [22].

In the case of 1D disorder, a very general approach, which allows to calculate diffuse scattering from any stacking fault model is derived by Treacy [23]. In this approach, the stacking sequence is described as a Markov chain. The user provides the structure of all possible layers, the vectors which describe possible stacking pairs of these layers and the corresponding stacking probabilities. Then, a numerical algorithm can be used to calculate the diffuse scattering from such a model. Such calculations can be performed in the program DIFFaX [23], least square refinement against powder scattering can be performed in program DIFFaX Plus [24].

In the case of 2D and 3D disorder, a different approach may be employed. The real structure of the crystal is separated in two parts: the average structure and the local deviations from the average structure. The average structure gives rise to Bragg peaks, while the autocorrelation of the deviations from the average structure are responsible for diffuse scattering. Since in disordered crystals correlations decay quite quickly, it is possible to express diffuse scattering as finite series over various short range order correlations. A very extensive research showing the form of diffuse scattering from many different point and extended defects can be found in the book by Krivoglaz [25].

A very similar approach is utilized by the Warren-Cowley formalism [26] which is usually employed to describe disordered solid solutions with simple average structures. This formalism introduces two types of short-range order parameters: correlation in occupancies of disordered sites and the so-called size-effect, the relaxation of the neighboring atoms around mixed positions. Diffuse scattering resulting from such correlations can be calculated using the equations derived by using the first terms of Taylor expansions. The approach is computationally very efficient and is well suited for performing refinements (an example can be found in [27]), however no general purpose software for performing such calculations is currently available.

Powder pair distribution function

If single crystal experiments are not possible the local structure of a sample can be investigated using the powder pair distribution function (1D-PDF) method. The 1D-PDF is similar to Rietveld refinement, but considers not only Bragg scattering but also diffuse scattering. The refinement is typically performed in real space against the radial distribution function, i.e. the Fourier transform of total powder scattering.

1D-PDF is a well established [28] method, which shares all the advantages and disadvantages of powder methods. On the one side, it can be applied to a wide range of samples including crystalline, amorphous and liquid materials. The data collection is fast and simple and can be used for high throughput experiments. On the other side, powder data provides access only to the spherical

projection of the pair distribution function. Thus, the signals from pairs with similar distances, but different directions of interatomic vectors overlap. The amount of overlapped peaks quickly increases as $|\mathbf{r}|^2$ making interatomic vectors beyond $\sim 10\text{\AA}$ hard to interpret.

1D-PDF refinements can be performed e.g. in the program PDFFIT [9].

Chapter 3

3D- Δ PDF refinement method

3.1 Fundamentals of 3D- Δ PDF analysis

The current paper introduces the 3D- Δ PDF formalism on the level of individual interatomic pairs. It shows how to describe short range ordering in terms of three correlations: substitutional correlation, size effect and atomic displacements correlation. The fingerprints of short range correlations are presented in both PDF and diffuse scattering space.

The printed version of article contains a small error in equation (9): it should read $f_n(\mathbf{h})$ instead of $f_n^*(\mathbf{h})$. The correct version is thus:

$$I_{dif}(\mathbf{h}) = \sum_{\mathbf{R}_{uvw}} \sum_{mn}^{cryst\ cell} \{ p_{uvw}^{mn} \exp(-\mathbf{h}^T \beta_{uvw}^{mn} \mathbf{h}) \cos[2\pi \mathbf{h}^T (\mathbf{R}_{uvw} + \mathbf{r}_{mn} + \bar{\mathbf{u}}_{uvw}^{mn})] - c_m c_n \exp[-\mathbf{h}^T (\beta_m^{aver} + \beta_n^{aver}) \mathbf{h}] \cos[2\pi \mathbf{h}^T (\mathbf{R}_{uvw} + \mathbf{r}_{mn})] \} f_m(\mathbf{h}) f_n(\mathbf{h})$$

The above equation is applicable for real atomic form factors at wavelengths far from the absorption edges and for molecular form factors of centro-symmetric molecules. In cases where the imaginary components of form factors cannot be neglected, the equation (8) should be used:

$$I_{dif}(\mathbf{h}) = \sum_{\mathbf{R}_{uvw}} \sum_{mn}^{cryst\ cell} [p_{uvw}^{mn} \exp(-\mathbf{h}^T \beta_{uvw}^{mn} \mathbf{h}) \exp(2\pi i \mathbf{h} \bar{\mathbf{u}}_{uvw}^{mn}) - c_m c_n \exp[-\mathbf{h}^T (\beta_m^{aver} + \beta_n^{aver}) \mathbf{h}] \times f_m^*(\mathbf{h}) f_n(\mathbf{h}) \exp[2\pi i (\mathbf{R}_{uvw} + \mathbf{r}_{mn}) \mathbf{h}]]$$

Note that equations (7) and (8) also contain a small error. The terms $f_m f_n^*(\mathbf{h})$ should read $f_m^* f_n(\mathbf{h})$.

The three-dimensional pair distribution function analysis of disordered single crystals: basic concepts

Thomas Weber* and Arkadiy Simonov

Laboratory of Crystallography, ETH Zurich Wolfgang-Pauli-Str. 10, 8093 Zurich, Switzerland

Received December 9, 2011; accepted January 23, 2012

*3D-Pair distribution function /
Pair distribution function analysis / Disorder /
Diffuse scattering*

Abstract. Theory and principles of the three-dimensional pair distribution function analysis of disordered single crystals are introduced. The mathematical framework is presented and the appearance of pair distribution function patterns is discussed on the examples of some basic disorder models. It is further demonstrated how pair distribution function maps are affected by typical experimental problems. Approaches for a better understanding of such effects and strategies for a proper handling of artifacts in diffuse scattering experiments are proposed.

1. Introduction

In the recent years the pair distribution function analysis of powder samples (*powder PDF*) has become a popular tool for investigating disordered structures. The powder PDF is the Fourier transform of the total X-ray or neutron powder diffraction pattern of a sample and provides a direct measure for the real interatomic distances $|\mathbf{r}|$ in a material. Moderate experimental requirements, straightforward data evaluation procedures and availability of powerful software [6, 2] make the powder PDF a perfect tool for routine investigations of local structures, provided that the problem at hand does not exceed a certain complexity. Limitations are mainly due to the powder specific angular projection of the PDF densities, which makes interatomic vectors of similar length indistinguishable even if their spatial orientations differ significantly. Furthermore, the frequency of powder PDF peaks increases with $|\mathbf{r}|^2$ and therefore problems quickly become unmanageable if the interatomic vectors of interest are getting long.

PDFs from single crystals (*3D-PDF*) may be calculated as the Fourier transform either of the total single crystal diffraction pattern (*total 3D-PDF*) or of the diffuse scattering alone (*3D- Δ PDF*). In general, 3D-PDF investigations are experimentally more challenging than powder experiments, because high quality three-dimensional diffraction data sets need to be collected. On the other hand most of

the problems inherent to the powder PDF method may be overcome: the full 3D information about interatomic vectors is preserved and the peak frequency is approximately constant as a function of $|\mathbf{r}|$. In the case of 3D- Δ PDFs the number of significant peaks per volume even decreases with increasing distances. 3D-PDF methods were successfully applied to a series of disorder problems that are difficult to tackle with traditional methods like Monte Carlo modeling. In particular disorder in quasicrystals was studied with 3D-PDF methods. Investigations cover phononic and phasonic disorder in an $\text{Al}_{70}\text{Co}_{12}\text{Ni}_{18}$ decagonal quasicrystal [3, 4], as well as temperature dependent studies of structural disorder in the same compound [14]. The atomic structure of disordered clusters in the decagonal compound $\text{Al}_{65}\text{Cu}_{20}\text{Co}_{15}$ was recently identified and refined in [9]. 3D-PDF methods are, however, by no means restricted to quasiperiodic compounds, but may equally well be applied to disordered periodic structures [1, 8].

This paper is intended to introduce the basic concepts and properties of the total 3D-PDF and the 3D- Δ PDF methods. Application to real world examples is beyond the scope of this paper and will be presented in forthcoming publications.

2. Theory

2.1 Definition

In the following we will present the mathematical framework for the 3D-PDF theory. A similar approach was proposed *e.g.* in [16], which, however, focused on the description of diffuse scattering intensities. Here we will introduce a slightly different notation that more emphasizes total 3D-PDF and 3D- Δ PDF aspects.

In analogy to the powder PDF method the total 3D-PDF is defined as the Fourier transform of the total scattering from a single crystal:

$$P_{\text{tot}}(\mathbf{x}) = \text{FT}\langle I_{\text{tot}}(\mathbf{h}) \rangle. \quad (1)$$

$P_{\text{tot}}(\mathbf{x})$ is also called the *autocorrelation function* of the real crystal, $I_{\text{tot}}(\mathbf{h})$ are the properly corrected total scattering intensities from a single crystal diffraction experiment, FT stands for the Fourier transformation and $\langle \rangle$ denotes intensity averaging over time and over all coherently scat-

* Correspondence author (e-mail: thomas.weber@mat.ethz.ch)

tering volumes in a sample. In the following all vectors are expressed in dimensionless fractional lattice units, if not stated otherwise. For sake of simplicity we will ignore special experimental effects like anomalous, inelastic or multiple scattering and, without loss of generality, we will only use notations related to X-ray diffraction. The autocorrelation function of a crystal may be written as

$$P_{\text{tot}}(\mathbf{x}) = \sum_{MN}^{\text{cryst}} \varrho_M(\mathbf{x}) * \varrho_N(-\mathbf{x}) * \langle \delta(\mathbf{x} - \mathbf{x}_{MN}) \rangle. \quad (2)$$

The double summation runs over all atoms in the crystal, $\varrho(\mathbf{x})$ is the electron density of an atom at the origin of the crystal's coordinate system, \mathbf{x}_{MN} is the vector between the atoms M and N , $\delta(\mathbf{x})$ is the Dirac function and the asterisk symbolizes the convolution operation. This general expression is valid for crystalline and non-crystalline materials. Crystals are built of unit cells and have an average structure, so that we can write

$$P_{\text{tot}}(\mathbf{x}) = \sum_{MN}^{\text{cryst}} \varrho_M(\mathbf{x}) * \varrho_N(-\mathbf{x}) * \delta(\mathbf{x} - \mathbf{R}_{uvw} - \mathbf{r}_{MN}) * \langle \delta(\mathbf{x} - \mathbf{u}_{MN}) \rangle, \quad (3)$$

where \mathbf{R}_{uvw} is the lattice vector between the unit cells occupied by the atoms M and N , and \mathbf{u}_{MN} is the difference between the real vector \mathbf{x}_{MN} and the average structure related vector $(\mathbf{R}_{uvw} + \mathbf{r}_{MN})$. In a next step the double summation is no longer performed over all atoms in the crystal, but over all atomic sites¹ m and n within the average unit cell. We further average over all interatomic vectors connecting atoms that are occupying the sites m and n in unit cells separated by a same lattice vector \mathbf{R}_{uvw} . The expression for the total 3D-PDF then becomes

$$P_{\text{tot}}(\mathbf{x}) = \sum_{\mathbf{R}_{uvw}}^{\text{cryst cell}} \sum_{mn} p_{uvw}^{mn} \varrho_m(\mathbf{x}) * \varrho_n(-\mathbf{x}) * \delta(\mathbf{x} - \mathbf{R}_{uvw} - \mathbf{r}_{mn}) * \langle \delta(\mathbf{x} - \mathbf{u}_{uvw}^{mn}) \rangle, \quad (4)$$

where p_{uvw}^{mn} is the joint probability to simultaneously find an atom at site m in one unit cell and another atom at site n in a unit cell separated by \mathbf{R}_{uvw} . The term $\langle \delta(\mathbf{x} - \mathbf{u}_{uvw}^{mn}) \rangle$ is the probability density distribution of the random variable \mathbf{u}_{uvw}^{mn} , which represents the difference between the real and the average distance of atoms occupying sites m and n that are uvw unit cells apart. For a chemical understanding it is often more convenient to express the pair correlations in terms of conditional probabilities, which are defined as $p_{uvw}^{mn} = c_m p_{uvw}(n | m) = c_n p_{\bar{u}\bar{v}\bar{w}}(m | n)$, where c_m and c_n are the average site occupation factors of atomic sites m and n , respectively. The term $p_{uvw}(n | m)$ denotes the conditional probability that a site n is occupied, provided that the uvw unit cells separated site m is also occupied. The definition of $p_{uvw}(m | n)$ is analogous.

¹ An atomic site is defined as a placeholder for a specific element that occupies this position with a certain probability. The atom belonging to this site may rest on its ideal position or it may be displaced following a distribution function. It is important to note that in the case of substitutional disorder an atomic site needs to be defined for each of the elements that share an average position. The total number of atomic sites per unit cell therefore equals to the length of the atom list required for defining an average structure, if the symmetry of the crystal is represented in space group $P1$.

The Fourier transform of the Bragg scattering is well-known as the *Patterson function*, which is the autocorrelation function of the average structure of a crystal. It can be described by averaging Eq. (4) over all lattice vectors. As a consequence, each atomic site m will be filled with a probability c_m , even if the results would lead to unphysical interatomic vectors. The joint occupational probability of every pair of average atoms is equal to the product of the occupancies of the individual atoms, *i.e.* $p_{uvw}^{mn} = c_m c_n$ and the distributions \mathbf{u}_{uvw}^{mn} will split into two independent distributions of \mathbf{u}_m and \mathbf{u}_n . The full formula for describing the Patterson function is

$$P_{\text{Pat}}(\mathbf{x}) = \sum_{\mathbf{R}_{uvw}}^{\text{cryst cell}} \sum_{mn} c_m c_n \varrho_m(\mathbf{x}) * \varrho_n(-\mathbf{x}) * \delta(\mathbf{x} - \mathbf{R}_{uvw} - \mathbf{r}_{mn}) * \langle \delta(\mathbf{x} - \mathbf{u}_m) \rangle * \langle \delta(\mathbf{x} - \mathbf{u}_n) \rangle. \quad (5)$$

Typically, the average structure is well known before the real structure gets investigated. For a better understanding of local order phenomena it is therefore favorable to focus on the deviations from the average structure, *i.e.* to analyze $P_{\Delta}(\mathbf{x}) = P_{\text{tot}}(\mathbf{x}) - P_{\text{Pat}}(\mathbf{x})$. In reciprocal space this quantity is represented by the diffuse scattering $I_{\text{dif}}(\mathbf{h}) = I_{\text{tot}}(\mathbf{h}) - I_{\text{Bragg}}(\mathbf{h})$. The difference between Eqs. (4) and (5) yields:

$$P_{\Delta}(\mathbf{x}) = \sum_{\mathbf{R}_{uvw}}^{\text{cryst cell}} \sum_{mn} [p_{uvw}^{mn} \langle \delta(\mathbf{x} - \mathbf{u}_{uvw}^{mn}) \rangle - c_m c_n \langle \delta(\mathbf{x} - \mathbf{u}_m) \rangle * \langle \delta(\mathbf{x} - \mathbf{u}_n) \rangle] * \varrho_m(\mathbf{x}) * \varrho_n(-\mathbf{x}) * \delta(\mathbf{x} - \mathbf{R}_{uvw} - \mathbf{r}_{mn}). \quad (6)$$

The corresponding diffuse scattering intensity is the Fourier transform of Eq. (6):

$$I_{\text{dif}}(\mathbf{h}) = \sum_{\mathbf{R}_{uvw}}^{\text{cryst cell}} \sum_{mn} [p_{uvw}^{mn} \langle \exp(2\pi i \mathbf{h} \mathbf{u}_{uvw}^{mn}) \rangle - c_m c_n \langle \exp(2\pi i \mathbf{h} \mathbf{u}_m) \rangle \langle \exp(2\pi i \mathbf{h} \mathbf{u}_n) \rangle] f_m(\mathbf{h}) f_n^*(\mathbf{h}) \exp[2\pi i \mathbf{h} (\mathbf{R}_{uvw} + \mathbf{r}_{mn})]. \quad (7)$$

In cases where the expressions in angle brackets may be approximated by Gaussians we can write $\langle \exp(2\pi i \mathbf{h} \mathbf{u}) \rangle = \exp[-2\pi^2 \langle (\mathbf{h} \mathbf{u})^2 \rangle] = \exp(-\mathbf{h}^T \beta \mathbf{h})$ (see [10]), where β is a matrix representing dimensionless atomic displacement parameters (ADPs). The components of β are defined as $\beta_{ij} = 2\pi^2 a_i^* a_j^* U_{ij}$, where U_{ij} are the frequently used ADPs in units of $(\text{length})^2$ and a_i^* and a_j^* are the lengths of the i -th and j -th reciprocal space vectors. Care has to be taken, however, if substitutional and displacive disorder are correlated. In such cases the distribution of \mathbf{u}_{uvw}^{mn} may have a non-zero average and we obtain $\langle \exp(2\pi i \mathbf{h} \mathbf{u}_{uvw}^{mn}) \rangle = \exp(-\mathbf{h}^T \beta_{uvw}^{mn} \mathbf{h}) \exp(2\pi i \mathbf{h} \bar{\mathbf{u}}_{uvw}^{mn})$. The variable $\bar{\mathbf{u}}_{uvw}^{mn}$ indicates presence or absence of a so-called size effect distortion, depending on whether it is zero or not (see Section 3.3 for an example). In the harmonic approximation we can write:

$$I_{\text{dif}}(\mathbf{x}) = \sum_{\mathbf{R}_{uvw}}^{\text{cryst cell}} \sum_{mn} [p_{uvw}^{mn} \exp(-\mathbf{h}^T \beta_{uvw}^{mn} \mathbf{h}) \exp(2\pi i \mathbf{h} \bar{\mathbf{u}}_{uvw}^{mn}) - c_m c_n \exp(-\mathbf{h}^T (\beta_m^{\text{aver}} + \beta_n^{\text{aver}}) \mathbf{h})] * f_m(\mathbf{h}) f_n^*(\mathbf{h}) \exp[2\pi i \mathbf{h} (\mathbf{R}_{uvw} + \mathbf{r}_{mn})]. \quad (8)$$

where the matrix β_{uvw}^{mn} describes the joint probability displacement densities of atoms occupying sites m and n that are separated by an average vector ($\mathbf{R}_{uvw} + \mathbf{r}_{mn}$), while β_m^{aver} and β_n^{aver} are the space and time averaged ADPs of the sites m and n , respectively. After considering centrosymmetry and applying some simple mathematical manipulations Eq. (8) may be expressed without complex exponential terms as

$$I_{\text{dif}}(\mathbf{x}) = \sum_{\mathbf{R}_{uvw}}^{\text{cryst}} \sum_{mn}^{\text{cell}} \{P_{uvw}^{mn} \exp(-\mathbf{h}^T \beta_{uvw}^{mn} \mathbf{h}) \cos[2\pi \mathbf{h}(\mathbf{R}_{uvw} + \mathbf{r}_{mn} + \bar{\mathbf{u}}_{uvw}^{mn})] - c_m c_n \exp[-\mathbf{h}^T (\beta_m^{\text{aver}} + \beta_n^{\text{aver}}) \mathbf{h}] \cos[2\pi \mathbf{h}(\mathbf{R}_{uvw} + \mathbf{r}_{mn})]\} f_m(\mathbf{h}) f_n^*(\mathbf{h}). \quad (9)$$

2.2 Properties

Both, the total 3D-PDF and the 3D- Δ PDF are not periodic in the presence of disorder. The symmetry must therefore not be described by a space group, but by a Laue point group having the same symmetry as the total diffraction pattern or the diffuse scattering, respectively. Laue symmetry of Bragg and diffuse scattering are often identical, but they may differ in special cases [18, 13].

The 3D- Δ PDF provides information about features of the real structure that are not represented by the average structure. Positive 3D- Δ PDF values mean that the probability of finding scattering densities separated by the corresponding vector is higher than in the average structure and negative values indicate lower probabilities.

Using the 3D- Δ PDF for analyzing disorder instead of the total 3D-PDF is advantageous for several reasons.

- The number of peaks in PDF space gets smaller, because only interatomic vectors having different real and average structure properties contribute to the 3D- Δ PDF. This excludes all atoms that are fully ordered to a good approximation as well as atomic pairs that are completely uncorrelated, *e.g.* because they are separated by large distances [8].
- With the elimination of Bragg intensities one also removes any associated systematic and statistical errors at the same time. This is expected to significantly enhance the quality of the local structure information, since the integral *errors* of Bragg intensities are often of a same magnitude or even stronger than the integral *intensities* of diffuse scattering.
- The higher contrast of 3D- Δ PDF maps allows qualitative and semi-quantitative conclusions about the underlying disorder problem just by a visual inspection of the patterns [8, 14].

An important feature, which may be directly extracted from the 3D- Δ PDF pattern, is the correlation length of local order. It is measured from the longest significant 3D- Δ PDF vectors and provides the information how far and along which direction a local structure property may affect its environment.

The magnitudes of 3D- Δ PDF peaks essentially depend on the strength of pair correlations, on the product of the scattering power of the corresponding atomic pairs, on the contrast between the average and the real structure and on the multiplicity of the interatomic vectors. In general, the highest multiplicity is found for 3D-PDF peaks separated by lattice vectors, as the multiplicity of such vectors equals to the number of atomic sites per unit cell.

A further feature of the 3D- Δ PDF is that its origin peak may be directly calculated from the known average structure. It represents all interatomic vectors with length zero, *i.e.* $m = n$ and $\mathbf{R}_{uvw} = \mathbf{r}_{mn} = \mathbf{0}$. Since any atom is perfectly ordered with respect to itself we can write: $P_{000}^{mm} = c_m P_{000}(m | m) = c_m$ and $\mathbf{u}_{000}^{mm} = \mathbf{0}$. The origin peak may thus be calculated as:

$$P_{\Delta}^{\text{origin}}(\mathbf{x}) = \sum_m^{\text{cell}} [c_m \delta(\mathbf{x}) - c_m^2 \langle \delta(\mathbf{x} - \mathbf{u}_m) \rangle * \langle \delta(\mathbf{x} - \mathbf{u}_m) \rangle] * \varrho_m(\mathbf{x}) * \varrho_m(-\mathbf{x}), \quad (10)$$

i.e. all quantities required for calculating the 3D- Δ PDF origin peak are available from the average structure. This property provides a unique link between the average structure and the 3D- Δ PDF map, which was lost after diffuse and Bragg scattering were separated. An important application is the determination of a proper scale factor for diffuse scattering and $P_{\Delta}(\mathbf{x})$ maps of a model (see below).

From Eqs. (6) and (8) we know that the 3D- Δ PDF and its corresponding diffuse scattering may be described by the same set of parameters and therefore both, reciprocal space and PDF space oriented modeling, are likewise feasible. According to Parseval's law the following identity holds in the case of discrete Fourier transforms:

$$\sum_{i=0}^{K-1} |P_{\Delta}^{\text{obs}}[i] - P_{\Delta}^{\text{calc}}[i]|^2 = 1/K \sum_{j=0}^{K-1} |I_{\text{dif}}^{\text{obs}}[j] - I_{\text{dif}}^{\text{calc}}[j]|^2, \quad \text{where}$$

K equals to the number of grid points. Apart from this constant factor the unweighted χ^2 values of the model are the same in PDF and reciprocal space and therefore least-squares refinements are expected to lead to exactly the same results independently of the reference space. A major difference, however, is found in the ability for masking or weighting specific features. Experimental artifacts like beam-stop shadow, parasitic scattering, saturations etc. are easily masked in reciprocal space and statistical errors are also more easily considered if refinements are done against diffuse scattering data. Refinements in PDF space on the other hand allow selective modeling of structural features. It was shown in [9] that this feature may significantly reduce the complexity of a given problem.

3. Fingerprints of local order phenomena

In the following we will demonstrate the impact of some basic disorder models on the appearance of single crystal PDF maps. For sake of simplicity the example structure will be two-dimensional. Correspondingly, the resulting PDFs will be called 2D-PDF or 2D- Δ PDF. The findings and discussions, however, may be straightforwardly transferred to three-dimensional space.

3.1 Displacive disorder

The model structure is a chess board like arrangement of atoms A and B in the (non-standard) plane group $c4mm$. The scattering power of B is assumed to be twice as strong as the one of A and the average composition of the compound is AB . For our numerical calculations we used the atomic form factors of Al and Fe, respectively. The lattice constant a is defined to be 5 \AA . For the displacive disorder model we assume absence of any substitutional disorder and all displacements are defined to be harmonic and isotropic with $U_{\text{iso}} = 0.05 \text{ \AA}^2$, *i.e.* $\beta_{ij} \approx 0.0395$ if $i = j$, and zero in all other cases. The 2D- Δ PDF map is most easily obtained as the Fourier transform of the diffuse scattering calculated from Eq. (8), which in the case of pure displacive disorder in a two-dimensional structure simplifies to:

$$I_{\text{dif}}(\mathbf{h}) = \sum_{\mathbf{R}_{uv}} \sum_{mn}^{\text{cryst cell}} [\exp(-\mathbf{h}^T \beta_{uv}^{mn} \mathbf{h}) - \exp(-\mathbf{h}^T 2\beta^{\text{aver}} \mathbf{h})] f_m(\mathbf{h}) f_n^*(\mathbf{h}) \cos[2\pi\mathbf{h}(\mathbf{R}_{uv} + \mathbf{r}_{mn})]. \quad (11)$$

In the first example all displacements are independent of each other, corresponding to the real structure that has the highest entropy compatible with the Bragg intensities. In such a case we obtain $\beta_{uv}^{mn} = 2\beta^{\text{aver}}$, except for $\beta_{00}^{mm} = \beta_{00}^{nn} = \mathbf{0}$, because a real atom is never displaced with respect to itself. Consequently, Eq. (11) further simplifies to

$$I_{\text{dif}}(\mathbf{h}) = \sum_m^{\text{cell}} [1 - \exp(-\mathbf{h}^T 2\beta^{\text{aver}} \mathbf{h})] |f_m(\mathbf{h})|^2. \quad (12)$$

The diffuse scattering and the corresponding PDF maps are shown in the top row of Fig. 1. Intensities are zero at the origin of reciprocal space and increase continuously with increasing distance. The decrease of the intensities at large vectors is caused by the atomic form factor. In the case of neutron experiments, where the atomic scattering power is constant as a function of reciprocal space coordinates, the diffuse scattering would converge to a maximum value at large scattering angles. The 2D- Δ PDF is obtained as the Fourier transform of the diffuse scattering and we observe a relatively narrow signal at the origin, which has a so-called Mexican hat shape: high densities at the origin are enclosed by a ring of negative densities. The integral of the 2D- Δ PDF is zero as it is true for any other strictly displacive disorder model. At larger distances there are no signals visible because the displacements are uncorrelated and cannot be distinguished from the average correlations. The Mexican hat profile of the peak can be explained by the fact that the distance of an atom to itself is always zero. Therefore, a distance of zero is more likely to be observed in the real than in the average structure. On the other hand any distance of an atom to itself having a length larger than zero is never present in a real structure, giving rise to negative Δ PDF intensities around the positive peak at the origin. The observation that the positive peak is broader than a δ -function, as it would be expected from our argumentation, is due to the fact that electron densities of the atoms have non-zero finite widths, which broaden the PDF signals.

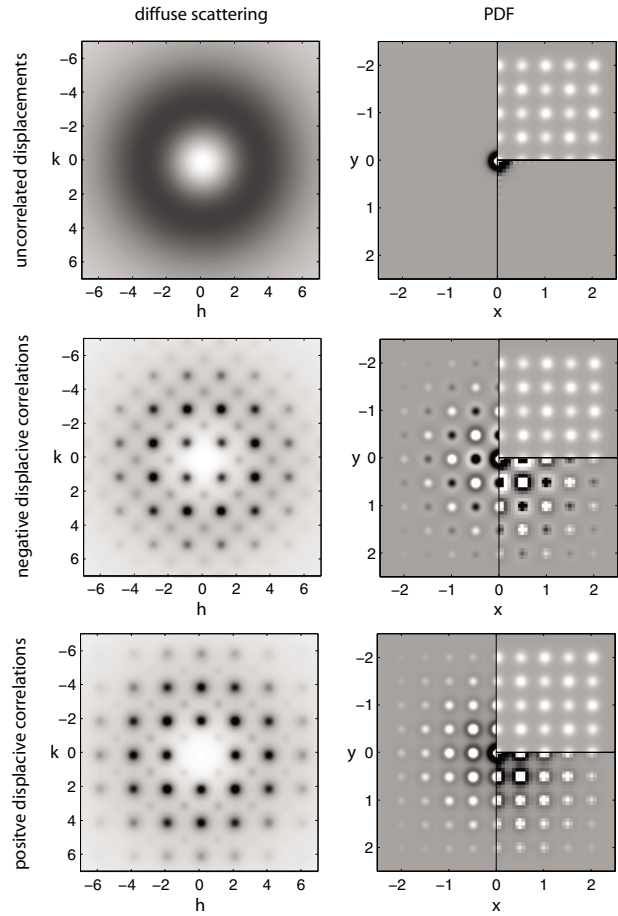


Fig. 1. Diffuse scattering and 2D-(Δ)PDF patterns of the displacive disorder models described in the text. Representations of diffuse and PDF maps follow conventional grey-scale schemes. In the case of diffuse scattering zero intensities are white, while strong intensities are shown in black. In the PDF space negative densities are shown in black, positive values are white and the zero level is in medium grey. On the left hand side of the PDF maps 2D- Δ PDFs are shown in very high resolution, while the lower right quarter shows the 2D- Δ PDFs in resolutions corresponding to the visible diffuse scattering in the left column, *i.e.* in a real space resolution of about 0.7 \AA . The upper right inset shows the total 2D-PDF pattern, *i.e.* the Fourier transform of Bragg and diffuse scattering. The total 2D-PDF and the 2D- Δ PDF color schemes are on different scales. In reality the total 2D-PDF intensities are by far stronger than the 2D- Δ PDFs. The Bragg reflections are not shown in the diffraction patterns. According to the c -centered lattice ($a = 5 \text{ \AA}$) they would be visible at $h + k = 2n$.

In a next step we allow negative correlations between the displacements, *i.e.* next neighboring atoms tend to move along anti-parallel directions. The diffuse intensity of this model is described by Eq. (11). In our example the values of β_{uv}^{mn} are defined by the equation

$$\beta_{uv}^{mn} = 2\beta^{\text{aver}} [1 - \sigma_{mn} \exp(-|\mathbf{R}_{uv} + \mathbf{r}_{mn}|)], \quad (13)$$

where $\sigma_{mn} = +1$ if m and n are occupied by same elements, otherwise it equals -1 . Any other parameters required in Eq. (11) are available from the average structure.

The resulting diffuse scattering and PDF patterns are shown in the middle row of Fig. 1. The diffuse scattering condenses around the Bragg positions and the envelope of the strong diffuse maxima is very similar to the diffuse scattering from the uncorrelated model. In 2D- Δ PDF

space the origin peak is exactly the same as for uncorrelated displacements, because this peak is exclusively defined by the average structure. Additional peaks are found at positions corresponding to the average interatomic vectors. As a consequence of the negative correlations of next neighboring atoms, different elements tend to move along opposite directions, while same elements prefer in-phase displacements. Consequently, real vectors between same elements have a high probability to be as long as the average vectors, leading to positive Mexican hat profiles. Different elements tend to avoid the average interatomic distances. Therefore Δ PDF intensities are negative at such positions and the surroundings are positive, *i.e.* the corresponding peaks look like negative Mexican hats. Because of decreasing correlations, real and average displacements of atomic pairs assimilate with increasing distances and the magnitudes of the Δ PDF peaks decrease accordingly.

The model for positive displacive correlations is analogously defined as the negative correlation model, but σ_{mn} is +1 in all cases. The results are shown in the lower row of Fig. 1. The diffuse scattering pattern is similar to the negative correlation pattern, however, the weak diffuse peaks in the former case become strong and *vice versa*. In 2D- Δ PDF space all peaks are forming positive Mexican hat patterns, because all atomic pairs with shorter distances than the correlation length tend to move in phase. Any other properties are the same as before.

In theory observation of a Mexican hat profile is a clear indication of displacive disorder. In practice, however, truncation effects may also generate similar features, if the experimental diffuse scattering considered for calculating PDF maps has a circular or a spherical envelope. Careful examinations of the origin of Mexican hat style features is therefore required.

3.2 Substitutional disorder

The average structure of the substitutional disorder model is defined such that each site of the previous example is mixed occupied by 0.5A and 0.5B atoms. Atoms are resting at their average positions so that all displacement vectors \mathbf{u} are zero. As a consequence of disorder the symmetry independent Wyckoff positions of the 5 Å structure can no longer be distinguished and the lattice transforms from a *c*-centered to a primitive cubic lattice with $a = 2.5$ Å. The plane group symmetry becomes $p4mm$. The 2D- Δ PDF space of such a model can be described as:

$$P_{\Delta}(\mathbf{x}) = \sum_{\mathbf{R}_{uv}}^{\text{cryst cell}} \sum_{mn} (p_{uv}^{mn} - 0.25) \varrho_m(\mathbf{x}) * \varrho_n(-\mathbf{x}) * \delta(\mathbf{x} - \mathbf{R}_{uv} - \mathbf{r}_{mn}), \quad (14)$$

and the diffuse scattering expression becomes:

$$I_{\text{dif}}(\mathbf{h}) = \sum_{\mathbf{R}_{uv}}^{\text{cryst cell}} \sum_{mn} (p_{uv}^{mn} - 0.25) f_m(\mathbf{h}) f_n^*(\mathbf{h}) \times \cos[2\pi\mathbf{h}(\mathbf{R}_{uv} + \mathbf{r}_{mn})]. \quad (15)$$

In the case of randomly distributed A and B atoms any pairs connected by a vector longer than zero have the joint occupational probabilities $p_{uv}^{AB} = p_{uv}^{BA} = p_{uv}^{AA} = p_{uv}^{BB} = c_{ACB} = 0.25$

and thus all of the corresponding terms in the summations of Eqs. (14) and (15) become zero. The conditional probabilities for the zeroth neighbor are $p_{00}(A|A) = p_{00}(B|B) = 1$ and $p_{00}(B|A) = p_{00}(A|B) = 0$. Consequently, p_{00}^{mn} equals to 0.5 for $m = n$ and to zero otherwise. This simplifies Eq. (14) to

$$\begin{aligned} P_{\Delta}(\mathbf{x}) &= 0.25[\varrho_A(\mathbf{x}) * \varrho_A(-\mathbf{x}) + \varrho_B(\mathbf{x}) * \varrho_B(-\mathbf{x}) \\ &\quad - \varrho_B(\mathbf{x}) * \varrho_A(-\mathbf{x}) - \varrho_A(\mathbf{x}) * \varrho_B(-\mathbf{x})] \\ &= 0.25[\varrho_A(\mathbf{x}) - \varrho_B(\mathbf{x})] * [\varrho_A(-\mathbf{x}) - \varrho_B(-\mathbf{x})], \end{aligned} \quad (16)$$

i.e. there is a single peak at the origin of PDF space, which is just the autocorrelation function of the difference between the electron densities of atoms A and B. The diffuse scattering obtained from such a model is proportional to the squared difference between the atomic form factors: $I_{\text{dif}}(\mathbf{h}) = 0.25|f_A(\mathbf{h}) - f_B(\mathbf{h})|^2$. Diffuse scattering and PDF maps of this model are shown in the top row of Fig. 2.

In a next step occupational short-range order is introduced. In our first model, chess board like arrangements

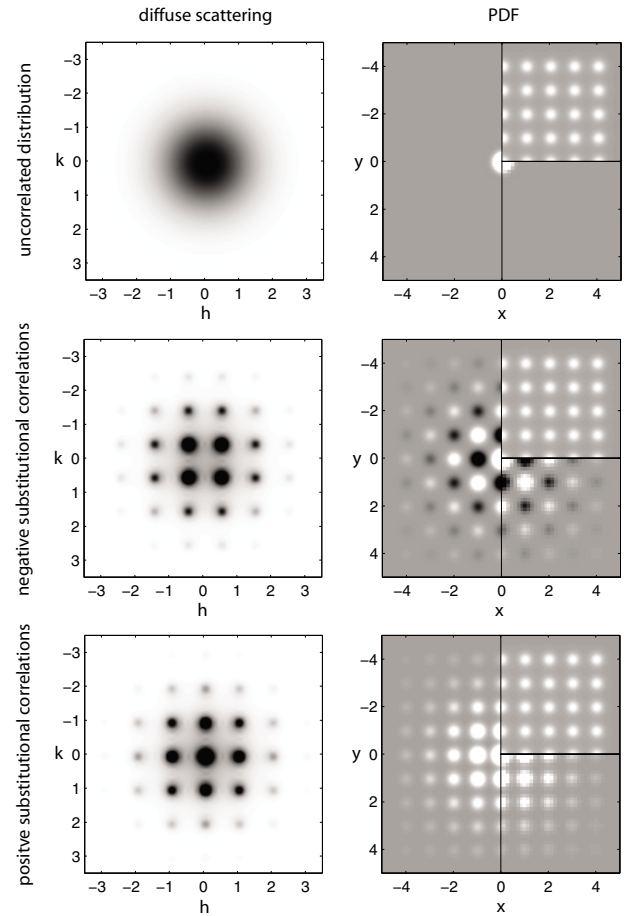


Fig. 2. Diffuse scattering and 2D-(Δ)PDF patterns of the substitutional disorder models described in the text. The grey-scale schemes and the meanings of sections in the PDF maps are as defined in Fig. 1. In reciprocal space the Bragg reflections are not shown. According to the *p*-lattice ($a = 2.5$ Å) they would be visible at integer h, k positions.

are formed on a local scale, *i.e.* there is a preference for hetero-atomic pairs AB or BA in the case that interatomic vectors $\mathbf{r} = u\mathbf{a} + v\mathbf{b}$ fulfill the condition $u + v = \text{odd}$ (based on the 2.5 \AA sized average unit cell), otherwise homo-atomic pairs AA or BB are more likely. The longer the interatomic vectors become the weaker the pair correlations get. For very long vectors the joint probabilities converge to $p_{uv}^{AA} \approx p_{uv}^{AB} \approx p_{uv}^{BA} \approx p_{uv}^{BB} \approx 0.25$. For any given distance $d = \sqrt{u^2 + v^2}$ we obtain the relations $p_d(A|B) = p_d(B|A) = 1 - p_d(A|A) = 1 - p_d(B|B)$. In our model, an isotropic correlation field is defined as

$$p_d(A|A) = 0.5[1 + \sigma_{uv} \exp(-d)], \quad (17)$$

where $\sigma_{uv} = +1$ if $u + v = \text{even}$, otherwise $\sigma_{uv} = -1$. The corresponding diffuse scattering and PDF patterns are seen in the middle row of Fig. 2. The diffuse scattering pattern shows broad maxima centered at positions $h, k = n/2$ ($n = \text{odd}$), *i.e.* at positions where the Bragg reflections are extinct due to the transformation of the 5 \AA sized c -lattice of the displacive disorder model discussed above to the 2.5 \AA p -lattice in the present example. This behavior can be easily understood as the structure consists of domains having a *local* chess board structure similar to the displacive disorder model. The 2D- Δ PDF pattern, however, allows a more direct interpretation of the experimental results. Positive peaks are visible at positions $u + v = \text{even}$, while peaks are negative at $u + v = \text{odd}$, *i.e.* the sign of the 2D- Δ PDF peaks directly indicates the properties of σ_{uv} . Again, the origin peak is the same as in the example of uncorrelated disorder. The profiles of the Δ PDF peaks are constant, but as a consequence of decreasing correlations the integral densities of the peaks decrease with increasing distances. The decay of the densities directly reflects the properties of the exponential term in Eq. (17).

The bottom row of Fig. 2 shows the case of positive correlations where the atoms prefer being next to a same element. The definition of the disorder model is the same as for negative correlations, however, σ_{uv} equals to $+1$ in all cases. Looking at the diffraction pattern we find that the diffuse peaks have moved to the Bragg positions, because the translation vector in the finite ordered domains of the real structure is the same as the average periodicity, *i.e.* 2.5 \AA . In 2D- Δ PDF space the information obtained is again more directly related to the real structure properties: all peaks are positive, what directly reflects the definition of σ_{uv} . Any other properties like the width of the diffuse peaks and the decay of pair correlations are equivalent to the negative correlation patterns.

3.3 Size effect distortions

Displacive disorder stimulated by substitutional or occupational disorder is called a *size effect*. The corresponding model discussed in this section is defined as follows: A and B atoms are randomly distributed among the sites of the 2.5 \AA structure and the average displacements are equivalent to the displacive disorder model defined above. In addition the local interatomic distances are depending

on the neighboring elements. For this model Eq. (8) may be rewritten as

$$I_{\text{dif}}(\mathbf{x}) = \sum_{\mathbf{R}_w} \sum_{mn}^{\text{cryst cell}} [p_{uv}^{mn} \exp(-\mathbf{h}^T \beta_{uv}^{mn} \mathbf{h}) \exp(2\pi i \mathbf{h} \bar{\mathbf{u}}_{uv}^{mn}) - 0.25 \exp(-\mathbf{h}^T 2\beta^{\text{aver}} \mathbf{h})] f_m(\mathbf{h}) f_n^*(\mathbf{h}) \exp[2\pi i \mathbf{h} (\mathbf{R}_w + \mathbf{r}_{mn})]. \quad (18)$$

AB distances are the same as in the average structure, while AA distances are shorter and BB pairs are further apart. In addition, we assume that the size effect induced shifts are much smaller than the independent displacements, which may be coming *e.g.* from thermal vibrations. This assumption allows us to further accept a harmonic approximation for the average displacements. From the definition of our model it follows that the disorder parameters p_{uv}^{mn} and β_{uv}^{mn} are exactly the same as in the cases of uncorrelated disorder discussed above, *i.e.* $p_{00}^{AA} = p_{00}^{BB} = 0.5$, $p_{00}^{AB} = p_{00}^{BA} = 0.0$ and $p_{uv}^{mn} = 0.25$ for any non-zero vector. The joint displacement probabilities follow the relation $\beta_{uv}^{mn} = 2\beta^{\text{aver}}$, except for $\beta_{00}^{mn} = \beta_{00}^{nn} = \mathbf{0}$. The only free model parameter is $\bar{\mathbf{u}}_{uv}^{mn}$. It is defined such that AA pairs neighboring along a main axis, are shifted by 0.01 \AA along this direction and BB distances are 0.01 \AA shorter. The size effect shifts decrease by a half for each additional step, *i.e.* it is 0.005 \AA for the second next neighbor, 0.0025 \AA for the third, and so on. For sake of simplicity we assume that size effect displacements are only effective for elements neighboring along $\langle 100 \rangle$ directions.

The resulting diffuse scattering and PDF maps are seen in Fig. 3. The diffuse scattering shows asymmetries with respect to integer h, k values, which is a very typical pattern for size-effect distortions [18]. The 2D- Δ PDF pattern straightforwardly reflects the disorder model. Correlations are only visible along the main axes. Similar to reciprocal space the size effect is seen in strong asymmetries along radial directions. For the present example the interatomic vectors point to positive values if the distances are larger than the average distances and negative for shorter separations. This observations can be directly interpreted such that distances between pairs of strong scatterers are longer than between pairs of weak scatterers.

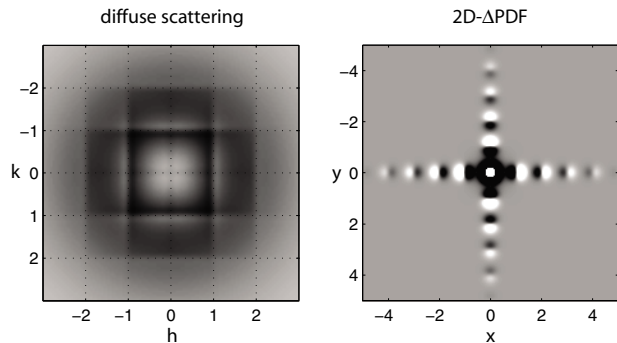


Fig. 3. Diffuse scattering and 2D- Δ PDF patterns of the size effect disorder model described in the text. The grey-scale schemes are as defined in Fig. 1. Bragg reflections are not shown, but would be visible at integer h, k positions according to the p -lattice ($a = 2.5 \text{ \AA}$). The dotted grid in the diffuse scattering pattern is at integer h and k values to demonstrate the asymmetry of the diffuse intensities with respect to the integer Bragg positions.

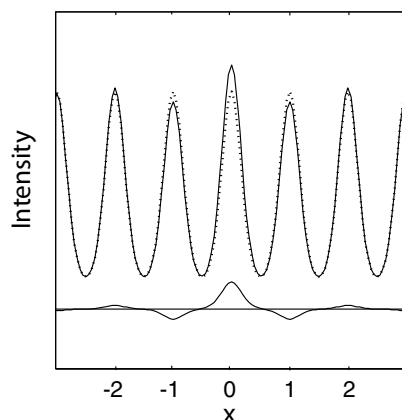


Fig. 4. Profiles from the intensities of the total 2D-PDF (upper full line), of the 2D- Δ PDF (lower full line) and of the Patterson function (dotted line) corresponding to the negative substitutional correlation model shown in Fig. 2 middle row. The scans are taken at $y = 0$. All PDF densities are on the same scale. The horizontal line represents the zero density level. It is clearly seen that the Patterson function, which was calculated as the difference between the total 2D-PDF and the 2D- Δ PDF, heavily dominates the total PDF.

3.4 Full PDF vs. Δ PDF modeling

The total 2D-PDF patterns of the displacive and substitutional disorder models discussed above are seen in the upper right quarters of the PDF maps shown in Figs. 1 and 2. Differences between the models are hardly visible. The reason for the similarity of the patterns is that the total 3D-PDF is heavily dominated by the Patterson function so that grey-scale representations do not allow any identification of particular disorder models. To demonstrate the relations between the full PDF, the Δ PDF and the Patterson function more clearly one-dimensional scans are shown in Fig. 4. The data are based on the negatively correlated substitutional disorder model discussed above, which was selected because it shows the strongest contrast between the real and the average structure. The dominance of the Patterson function is clearly visible and nicely demonstrates the advantage of the Δ PDF over the total PDF method. This example also emphasizes a significant advantage of the single crystal PDF method over the powder PDF, where a proper separation of Bragg and diffuse scattering is by far more complicated than in single crystal diffraction patterns.

4. Pathology and remedies

In the following chapter we will discuss a few typical experimental and modeling problems associated with the investigation of disorder and we will propose approaches for overcoming them. The examples are not exhaustive and the problems are not specific for the PDF method. The PDF space, however, is very well suited for a comprehensive understanding of the impact of corresponding errors on the real structure properties derived from any diffuse scattering studies, including Monte Carlo simulations.

4.1 Reciprocal space resolution

The experimental point-spread function in a diffuse scattering experiment is a convolution of various factors coming from the instrument or from the sample. In the following any experimental influence on Bragg reflection profiles will be called a resolution effect. With this definition Bragg peak profiles become a direct measure of the resolution function. Note that heavily strained samples and paracrystals are excluded from our discussion, because the Bragg peaks are broadened by disorder and therefore the Bragg peak profiles do not exclusively represent the experimental resolution function. Note also that the determination of the average structure is usually not affected by the resolution function, because it does not influence the integral Bragg peak intensities. Resolution function determining factors include *e.g.* beam divergence, spectral bandwidth of the beam, vibrations of the crystal or the instrument, cross-section of the beam with the sample, point spread function of the detector, step width in the data collection scans, artifacts in reciprocal space reconstructions and mosaicity of the sample. The complexity of the problem makes straightforward solutions, as *e.g.* routinely done in Rietveld powder refinements, difficult. To demonstrate the impact of the various effects on the single crystal PDF properties we roughly divided the resolution effects into three categories.

First, a function is considered that broadens Bragg peaks isotropically and uniformly. To a first approximation, most of the effects mentioned above belong to this class. Mathematically, this effect can be described as a convolution of a perfect diffraction pattern with a constant profile function. According to the convolution theorem the effect in 3D-PDF space is a multiplication of the true PDF with the Fourier transform of the resolution function. If we assume a Gaussian as a peak-shape function then its Fourier transform is also a Gaussian. The larger a PDF vector gets the more attenuated become the observed PDF densities as a consequence of the resolution effect. Profiles of the PDF peaks are (practically) not affected. The effect is shown in Fig. 5, second row. Typical half widths of Bragg reflections measured in synchrotron radiation experiments are in the order of 10^{-3} \AA^{-1} (reciprocal space units are defined as $d^* = 2 \sin(\theta)/\lambda$). In such cases the true PDF intensities get damped by about 3.5% at a distance of 100 \AA and by 50% at about 440 \AA [15], *i.e.* only long vectors are seriously affected. In typical in-house experiments the reciprocal space resolution is in the order of 10^{-2} \AA^{-1} and thus significant effects are expected at about one order of magnitude shorter distances. Resolution effects can be corrected in PDF space by dividing the observed PDF maps with the Fourier transform of the resolution function. In general, no corrections are required if the diffuse scattering features are by far broader than the resolution function.

The other two effects considered are causing broadenings of the peaks along radial or angular directions. The resolution function is not constant, but increases linearly with increasing distances from the origin of reciprocal space. Radial resolution effects are usually coming from the spectral width of the primary beam, while angular

broadenings may be caused *e.g.* by the mosaicity of the sample, by vibrations or by the data collection step width. As seen in the third and fourth row of Fig. 5 the effects are analogous in reciprocal and PDF space. In the case of radially shaped resolution functions Bragg's law transforms the constant monochromator's bandwidth $\Delta\lambda/\lambda$ to a constant $\Delta d/d$ in PDF space leading to the radial broadenings that are increasing with increasing distances. The angular resolution effects are also directly transformed from one space to the other, since rotations of the coordinate systems are covariant in real and reciprocal space.

There is no easy correction for such angular or radial resolution effects, but the resolution function may be simulated and applied to PDF or diffuse scattering model maps by using numerical methods. If not considered in

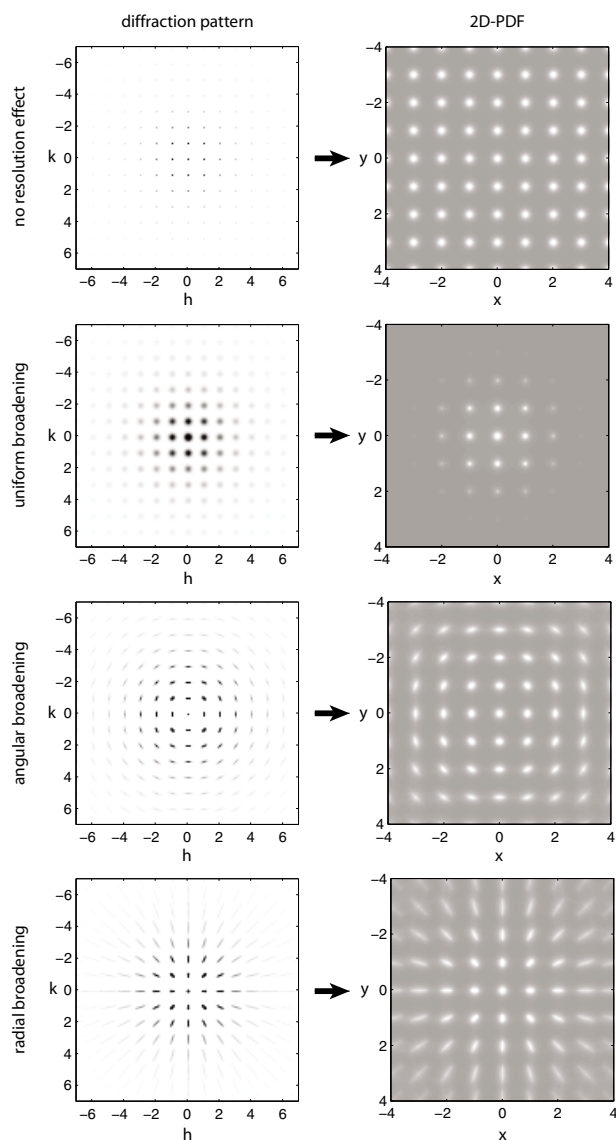


Fig. 5. Impact of reciprocal space resolution effects on the PDF pattern as described in the text. The top row shows a reference model of a diffraction experiment that is free of resolution effects. For the sake of clarity the examples are showing the resolution effects for Bragg scattering and for the Patterson function only, but the results may be straightforwardly transferred to diffuse scattering experiments. The grey-scale schemes are as defined in Fig 1.

diffuse scattering studies local atomic displacements will get overestimated independent of the modeling technique applied. Details will be discussed in forthcoming papers.

4.2 Separation of Bragg and diffuse data

In many cases separation of Bragg and diffuse scattering is straightforward, because the signals are found in different layers or lines. In other cases diffuse scattering is by far broader than the Bragg profiles and interpolation of the diffuse intensities beneath the Bragg scattering provides a reasonable approximation to the real diffuse intensities. More difficult situations are found if narrow diffuse maxima are exactly at or very close to the Bragg positions. In such cases the true diffuse scattering can hardly be reconstructed without providing additional information like theoretical profile functions. In the following we will propose three possibilities for solving such problems.

First, a total 3D-PDF study would overcome the necessity of separating Bragg and diffuse data. However, this approach requires perfectly measured diffuse and Bragg data at the same time – a task that is very challenging but feasible [12]. Second, despite the fact that interpolations may provide only rough approximations to the true diffuse scattering, the diffuse scattering beneath the Bragg peaks may be reconstructed using the so-called punch-and-fill method as discussed in [3, 4]. The method uses a simple interpolation algorithm for filling the gaps after elimination of Bragg peaks. It was shown that the resulting 3D- Δ PDF patterns are reliable if only very short vectors are considered, however, long PDF vectors may get heavily biased by this approximation. Finally, optimization of a 3D- Δ PDF model could be done in reciprocal space by applying zero weights to the Bragg scattering affected regions as it is typically done in Monte Carlo refinements (*e.g.* [17, 7, 11]). The consequences for the results obtained are very similar to the punch-and-fill method. In the case that the masked diffuse peaks are very narrow their major Fourier components correspond to long vectors in PDF space. By suppressing such experimental results, the information about long 3D- Δ PDF vectors may no longer be well defined. Independent of the modeling technique used, any analysis of long distance correlations may therefore become unreliable, unless a correct model for the decay of correlations is enforced.

4.3 Scale factor determination

The determination of a proper scale factor for a 3D- Δ PDF model is not trivial after Bragg and diffuse scattering was separated. In particular in the case of dominant substitutional disorder, diffuse scattering and 3D- Δ PDF maps linearly depend on the pair correlation parameters p_{uvw}^{nm} , *i.e.* they fully correlate with the scale factor. A proper determination of the scale factor is therefore not possible without providing additional information. Fortunately, the origin peak of the 3D- Δ PDF map provides such a constraint. As mentioned above the origin peak is independent of the disorder model and it can be calculated from the disorder properties known from the average structure like ADPs and site occupation factors. The scale factor is easily ob-

tained as the ratio of the experimental and calculated origin peaks. Any further pair correlation coefficients only need to be determined relative to the origin peak and thus the linear correlation between the scale factor the 3D- Δ PDF gets resolved.

There are, however, some practical problems that may bias a proper determination of the scale factor. It is a property of the Fourier transformation that the short 3D- Δ PDF vectors defining the origin peak are heavily affected by broad scattering intensities. A careful handling of background scattering is therefore essential as typical sources of background like electronic noise of the detector, air, fluorescence or Compton scattering as well as scattering from the crystal holder are smooth signals in most cases. A separation of background and diffuse scattering is easily done by interpolation, if the diffuse intensities are fully condensed in features that are narrow at least along one dimension like diffuse streaks or layers. In cases, where parts of the diffuse signal are very broad along all directions, diffuse scattering and background scattering are difficult to distinguish, and, as a consequence, the scale factor would get biased by the background. This problem can be minimized experimentally by using optimized setups and modern detectors like the PILATUS [5], which is not only free of intrinsic noise, but also allows suppression of fluorescence scattering. From a modeling point of view the problem can only be tackled, if additional constraints for pair correlations beyond the zeroth neighbor may be applied, as *e.g.* in the case of molecules, where a series of conditional occupational probabilities are often known to equal to unity. Such additional information allows extending the volume in PDF space that may be predicted from the average structure alone and make the scale factor determination significantly less sensitive to experimental artifacts.

5. Conclusions

Single crystal based PDF analysis is a straightforward method for analyzing disorder. Local order properties may be directly accessed through the Fourier transform of the full or the diffuse diffraction pattern. In the latter case the local order features are visible with higher contrast and they are less affected by experimental problems. A phase problem, as it is well-known for the determination of the average structure, does not exist, because the PDFs are directly calculated as the Fourier transform of the scattering intensities. Many local structure properties of a disordered material can be directly extracted from the single crystal PDF patterns just by a visual inspection. The 3D- (Δ) PDF patterns provides direct information about the type of disorder (*e.g.* displacive or substitutional disorder) present in a structure and about the local properties of the atoms involved, what is certainly the major advantage of PDF methods over reciprocal space oriented investigations. The formulas for calculating diffuse scattering or PDF maps from a PDF disorder model are perfectly suited for least-squares refinements. A computer program for performing such investigations is currently under development.

Powder and single crystal PDF methods are complementary. The major advantages of powder diffraction methods are that large single crystals are not required as well as the relatively simple and very fast experiments, which allow performing studies that cannot be done with single crystals. Single crystal PDF analysis on the other hand is to be preferred if complex local order needs to be investigated or if interatomic correlations become very long, even though collection of high quality diffuse scattering is much more complicated than in the case of powder experiments. For complex problems the by far better resolution of PDF peaks in three-dimensional space allows investigations of disorder problems that cannot be analyzed with powder diffraction techniques. The ability for elimination of average structure features and the possibility for selectively modeling parts of the disordered structure as demonstrated in [9], extends the tool-box for handling even extremely complex problems. Such advanced modeling techniques are much easier applied to single crystal PDFs than to powder diffraction data.

References

- [1] M. H. Chao, K. D. M. Harris, B. M. Kariuki, C. L. Bauer, B. M. Foxman: Characterization of intermolecular interactions in a disordered solid via a one-dimensional Patterson synthesis. *J. Phys. Chem. B* **2002**, *106*, 4032.
- [2] C. L. Farrow, P. Juhas, J. W. Liu, D. Bryndin, E. S. Bozin, J. Bloch, T. Proffen, S. J. L. Billinge: PDFfit2 and PDFgui: computer programs for studying nanostructure in crystals. *J. Phys.: Condens. Matter* **2007**, *19*, 335219.
- [3] M. Kobas, T. Weber, W. Steurer: Structural disorder in the decagonal Al-Co-Ni. I. Patterson analysis of diffuse X-ray scattering data. *Phys. Rev. B* **2005**, *71*, 224205.
- [4] M. Kobas, T. Weber, W. Steurer: Structural disorder in the decagonal Al-Co-Ni. II. Modelling. *Phys. Rev. B* **2005**, *71*, 224206.
- [5] P. Kraft, A. Bergamaschi, Ch. Broennimann, R. Dinapoli, E. F. Eikenberry, B. Henrich, I. Johnson, A. Mozzanica, C. M. Schlepueetz, P. R. Willmott, B. Schmitt: Performance of single-photon-counting pilatus detector modules. *J. Synchrotron Rad.* **2009**, *16*, 368.
- [6] T. Proffen, S. J. L. Billinge: PDFFIT, a program for full profile structural refinement of the atomic pair distribution function. *J. Appl. Crystallogr.* **1999**, *32*, 572.
- [7] T. Proffen, T. R. Welberry: Analysis of diffuse scattering of single crystals using Monte Carlo methods. *Phase Transit.* **1998**, *67*, 373.
- [8] P. Schaub, T. Weber, W. Steurer: Exploring local disorder in single crystals by means of the three-dimensional pair distribution function. *Philosoph. Mag.* **2007**, *87*, 2781.
- [9] P. Schaub, T. Weber, W. Steurer: Analysis and modelling of structural disorder by the use of the three-dimensional pair distribution function method exemplified by the disordered twofold superstructure of decagonal Al-Cu-Co. *J. Appl. Crystallogr.* **2011**, *44*, 134.
- [10] K. N. Trueblood, H. B. Bürgi, H. Burzlaff, J. D. Dunitz, C. M. Gramaccioli, H. H. Schulz, U. Shmueli, S. C. Abrahams: Atomic displacement parameter nomenclature – report of a subcommittee on atomic displacement parameter nomenclature. *Acta Crystallogr.* **1996**, *A52*, 770.
- [11] T. Weber, H. B. Bürgi: Determination and refinement of disordered crystal structures using evolutionary algorithms in combination with Monte Carlo methods. *Acta Crystallogr.* **2002**, *A58*, 526.
- [12] T. Weber, S. Deloudi, M. Kobas, Y. Yokoyama, A. Inoue, W. Steurer: Reciprocal-space imaging of a real quasicrystal. a feasibility study with PILATUS 6M. *J. Appl. Crystallogr.* **2008**, *41*, 669.

- [13] T. Weber, M. A. Estermann, H. B. Bürgi: Structural complexity of a polar perhydrotriphenylene inclusion compound brought to light by synchrotron radiation. *Acta Crystallogr.* **2001**, B57, 579.
- [14] T. Weber, M. Kobas, W. Steurer: The disordered 8 Ångstrom superstructure of a decagonal Al₇₀Co₁₂Ni₁₈ quasicrystal. *Ferroelectrics* **2004**, 305, 213.
- [15] T. Weber, W. Steurer: Structural Disorder in Quasicrystals. In: *Diffuse Scattering and the Fundamental Properties of Materials* (Eds. R. I. Barabash, G. E. Ice, P. E. A. Turchi), pp. 239–258.
- [16] T. R. Welberry, B. D. Butler: Diffuse X-ray scattering from disordered crystals. *Chem. Rev.* **1995**, 95, 2369.
- [17] T. R. Welberry, T. Proffen, M. Bown. Analysis of single-crystal diffuse X-ray scattering via automatic refinement of a Monte Carlo model. *Acta Crystallogr.* **1998**, A54, 661.
- [18] T. R. Welberry, *Diffuse X-ray Scattering and Models of Disorder*, Oxford University Press, New York **2004**.

3.2 The program YELL

The following paper [11] describes the implementation of 3D- Δ PDF method in the program YELL. The first part shows how to apply the three fundamental correlations (size effect, substitutional and displacement correlations) on the level of molecules and how to obtain their numerical values using the least squares refinement procedure.

The second part discusses the syntax of YELL input file. The average `UnitCell`¹ of the crystal is described in terms of `Variants` and constituent atomic groups. Displacements of atoms and molecules are defined in terms of `Modes`. Short range order parameters include `SubstitutionalCorrelation`, `SizeEffect` and `ADPCorrelation`.

A very important implementation detail, the fast FFT-based algorithm for diffuse scattering calculation is described in chapter 4.1. The appendix A presents the example input file on the basis of a Warren-Cowley model of Fe_{46.5}Ni_{53.5} refined by Jiang et. al. [29].

¹Here and throughout the thesis `typewriter` font indicates the YELL commands, key-words or input filenames.

***Yell*: a computer program for diffuse scattering analysis via three-dimensional delta pair distribution function refinement**

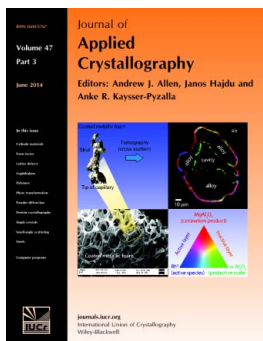
A. Simonov, T. Weber and W. Steurer

J. Appl. Cryst. (2014). **47**, 1146–1152

Copyright © International Union of Crystallography

Author(s) of this paper may load this reprint on their own web site or institutional repository provided that this cover page is retained. Republication of this article or its storage in electronic databases other than as specified above is not permitted without prior permission in writing from the IUCr.

For further information see <http://journals.iucr.org/services/authorrights.html>



Many research topics in condensed matter research, materials science and the life sciences make use of crystallographic methods to study crystalline and non-crystalline matter with neutrons, X-rays and electrons. Articles published in the *Journal of Applied Crystallography* focus on these methods and their use in identifying structural and diffusion-controlled phase transformations, structure-property relationships, structural changes of defects, interfaces and surfaces, *etc.* Developments of instrumentation and crystallographic apparatus, theory and interpretation, numerical analysis and other related subjects are also covered. The journal is the primary place where crystallographic computer program information is published.

Crystallography Journals Online is available from journals.iucr.org

Yell: a computer program for diffuse scattering analysis via three-dimensional delta pair distribution function refinement

A. Simonov, T. Weber* and W. Steurer

Laboratory of Crystallography, Department of Materials, ETH Zurich, Vladimir-Prelog-Weg 5, CH-8093 Zurich, Switzerland. Correspondence e-mail: thomas.weber@mat.ethz.ch

Received 6 February 2014

Accepted 16 April 2014

Yell, a program for routine refinement of disorder models against single-crystal diffuse scattering data, is presented. The analysis is based on the three-dimensional delta pair distribution function (3D- Δ PDF) method, which provides direct access to interatomic correlations in real crystal structures. Substitutional, displacive and size-effect disorder models are covered. The input file format supports flexible usage of arithmetic expressions for constraining dependent parameter values. The program is designed to be run on desktop computers. By using an efficient fast-Fourier-transform-based diffuse scattering calculation algorithm, full least-square refinements of medium complexity disorder models may be performed within minutes or hours, even if the experimental diffuse scattering is represented by large and fine-sampled reciprocal space volumes. The program is written in C++ and the source code is distributed under the GPL licence. Binary distributions are currently available for Mac and Windows operating systems.

© 2014 International Union of Crystallography

1. Introduction

Knowing the local structure of disordered crystals is frequently crucial for understanding their chemical and physical properties. Examples include giant magnetoresistivity (Adams *et al.*, 2000), specific catalytic activity (Simonovic & Armbruster, 2004), atomic diffusion paths in ion conductors (Ali *et al.*, 2008) and the behavior of host-guest systems (Weber *et al.*, 2000). In diffraction experiments the information about local order is found in the diffuse scattering. With recent advances in X-ray sources and detector technology (Henrich *et al.*, 2009; Weber *et al.*, 2008), the collection of high-quality diffuse scattering data sets from disordered crystals has now become relatively easy and fast. The availability of powerful software for analyzing diffuse scattering, however, is strongly lagging behind the current experimental possibilities.

In the past two decades direct space modeling techniques have become popular tools for studying disorder in single crystals. The methods include molecular dynamics (MD), Monte Carlo (MC) and reverse Monte Carlo simulations, all of which simulate small volumes of crystal structures that are supposed to be representative of the real crystal. Such simulations can be performed with programs like *DISCUS* (Proffen & Neder, 1997), *ZODS* (Chodkiewicz *et al.*, 2013), *ZMC* (Goossens *et al.*, 2011) or *Moldy* (Refson, 2000). The techniques are powerful but computationally very expensive, since the computer models may comprise hundreds of thousands of atoms, which need to be equilibrated and Fourier transformed. Although the use of supercomputers allows significant acceleration by parallelization (Michels-Clark *et al.*, 2013), full refinements of medium complexity problems may easily take days or even weeks of wall-clock time.

The powder-diffraction-based pair distribution function (PDF) method follows a different approach for modeling real structure properties of a disordered crystal. Instead of simulating a real crystal structure, it extracts information about the pair correlations directly

from the Fourier transform of the powder diffraction pattern, which reduces the computational demands significantly. In combination with powerful software (Proffen & Billinge, 1999), it has been successfully applied to a broad range of problems (Billinge, 2008). The PDF method inherits advantages and disadvantages of the powder diffraction method. It is experimentally much simpler and faster than single-crystal investigations but is limited regarding the complexity that can be analyzed. As a consequence of the spherical projection of diffraction intensity, only the norm and not the direction of interatomic vectors is directly represented in the experimental data.

Recently, the three-dimensional PDF (3D-PDF) method was introduced to extend the PDF approach to single-crystal applications (Weber & Simonov, 2012). Key features of the 3D-PDF method are preservation of the full angular information about interatomic vectors and a significantly smaller degree of overlapping of PDF signals. Both help to avoid numerical correlations between the model parameters and allow easier separation of diffuse and Bragg scattering compared to powder diffraction. The Fourier transform of the diffuse scattering alone is called the 3D- Δ PDF (Schaub *et al.*, 2007). It provides information about features of the real structure that are not represented on average. The idea of describing single-crystal diffuse diffraction patterns directly in terms of pair correlations is well established (Cowley, 1950; Welberry & Butler, 1995), but so far general software for quantitative routine investigations of disorder in single crystals has not been available. The 3D- Δ PDF modeling program *Yell*, which is introduced in this paper, is intended to fill this gap.

2. The theory of the 3D- Δ PDF method

The theory of the 3D- Δ PDF method was described in detail by Weber & Simonov (2012). Some basic concepts are briefly sketched in

the following. The 3D- Δ PDF method is based on the separation of diffuse and Bragg scattering:

$$3\text{D-}\Delta\text{PDF} = \text{FT}(I_{\text{diff}}) = \text{FT}(I_{\text{total}} - I_{\text{Bragg}}), \quad (1)$$

where FT symbolizes the Fourier transform operation. Once the average structure has been determined from the Bragg intensities, diffuse scattering may be calculated as the sum of signals from all atomic pairs in the crystal:

$$I(\mathbf{h})_{\text{diff}} = \sum_{\mathbf{R}_{uvw}} \sum_{mn}^{\text{crystal cell}} \left\{ p_{mn}^{uvw} \exp(-\mathbf{h}^T \boldsymbol{\beta}_{mn}^{uvw} \mathbf{h}) \exp[2\pi i \mathbf{h}(\mathbf{R}_{uvw} + \mathbf{r}_{mn} + \mathbf{u}_{mn}^{uvw})] - c_m c_n \exp[-\mathbf{h}^T (\boldsymbol{\beta}_m + \boldsymbol{\beta}_n) \mathbf{h}] \exp[2\pi i \mathbf{h}(\mathbf{R}_{uvw} + \mathbf{r}_{mn})] \right\} f_m^*(\mathbf{h}) f_n(\mathbf{h}), \quad (2)$$

where $\mathbf{r}_{mn} = \mathbf{r}_m - \mathbf{r}_n$ is a vector between the average positions of atoms m and n within a unit cell, c_n and c_m are the average atomic occupancies, $\boldsymbol{\beta}_m$ and $\boldsymbol{\beta}_n$ are the matrices of the average atomic displacement parameters, $f_m(\mathbf{h})$ and $f_n(\mathbf{h})$ are the atomic form factors, and \mathbf{R}_{uvw} is the lattice vector between lattice points that are (u, v, w) unit cells apart. All variables with sub- and superscript notation $[\substack{u \\ v \\ w}]$ represent short-range-order parameters, which are not known from the average structure but need to be quantified for describing the short-range order in the crystal: p_{mn}^{uvw} is the joint probability of finding the atoms m and n separated by the vector $\mathbf{R}_{uvw} + \mathbf{r}_{mn}$, $\boldsymbol{\beta}_{mn}^{uvw}$ is a matrix that describes the joint atomic displacement of atom n as seen from atom m , and \mathbf{u}_{mn}^{uvw} is the size effect parameter, which quantifies the difference between the real and the average interatomic vector between atoms m and n .

3. Definition and refinement of disorder models in Yell

The following section is intended to give an overview of how disorder models are defined in *Yell*. A comprehensive description of all features is found in the manual and a commented example of a simple but complete input file is given in Appendix A.

3.1. Description of the average structure

The average crystal structure is expressed in a hierarchical manner: the crystal consists of topological sites, which can be occupied by chemical units, *i.e.* individual atoms or groups of atoms. Groups of atoms represent ensembles, like molecules or clusters. Exclusive occupancies of topological sites, as present in the case of substitutional disorder, are expressed as variants, *e.g.*

```
UnitCell[
  GdFeVoid_site=Variant [
    (p=1/3)
    Gd 1 0 0 0 0.035 0.035 0.035 0 0 0

    (p=1/3)
    iron_molecule = [
      Fe 1 0 0 0.2845 0.035 0.035 0.035 0 0 0
      Fe 1 0 0 -0.2845 0.035 0.035 0.035 0 0 0
    ]

    (p=1/3)
    Void
  ]
]
```

In this example GdFeVoid_site is occupied either by a gadolinium atom, by a structural building block of two iron atoms, which is aliased as iron_molecule, or by a void. On average, each chemical unit is present with the same probability of $p = 1/3$. The sum of all probabilities of the constituents of a Variant must be equal to 1. If a site is under-occupied, the void must be listed explicitly. An atom is

defined by its element symbol followed by a multiplicity factor, three fractional coordinates, and an isotropic displacement parameter or, as shown in the example above, six anisotropic displacements parameters (ADPs).

Another example is the variant Fe3Si_site, which will be used later on. It defines a site that is occupied by a mixture of iron (83%) and silicon (17%):

```
Fe3Si_site = Variant[
  (p=0.87)
  Fe3 = Fe 1 0.5 0 0.5 0.04 0.04 0.04 0 0 0

  (p=0.13)
  Si 1 0.5 0 0.5 0.04 0.04 0.04 0 0 0
]
```

3.2. Definition of chemical unit motions

Motions of atoms or groups of atoms are defined analogously to normal mode analysis by

$$\Delta \mathbf{r}_i = \sum_j \mathbf{v}_{ij} \xi_j, \quad (3)$$

where $\Delta \mathbf{r}_i$ is the displacement vector of an atom i , ξ_j is an abstract variable describing the amplitude of some collective motion of the atoms and \mathbf{v}_{ij} is the vector of the displacement of the atom i associated with variable ξ_j . The index j counts all abstract variables. Such formalism can describe any motion of atomic groups including translations, rotations, expansion/contraction and bending in linear approximation. In *Yell* such collective motions are called modes. Before being used, modes need to be defined as shown in the following example of translational and rotational motions of the iron_molecule:

```
Modes[
  iron_molecule_mode_x = TranslationalMode(iron_molecule,x)
  iron_molecule_mode_ry = RotationalMode(iron_molecule,
                                          0,1,0, 0,0,0)
]
```

where iron_molecule_mode_x is an alias for a collective motion of the iron_molecule along vector \mathbf{x} , and the alias iron_molecule_mode_ry represents a linear approximation of the rotation of the iron_molecule around an axis parallel to \mathbf{y} passing through the center of the molecule with the unit-cell coordinates $(0, 0, 0)$.

3.3. Substitutional correlation

Substitutional correlation means that the occupancy of one site depends on the occupancy of another site. To specify substitutional correlations in *Yell*, two variants and a matrix of joint probabilities need to be defined. If the two variants consist of j and k atomic groups then the joint probabilities of a specific pair are expressed by a $j \times k$ correlation matrix. For example, the correlation between the variants Fe3Si_site (two constituents) and GdFeVoid_site (three constituents) is described by a 2×3 matrix:

```
SubstitutionalCorrelation(Fe3Si_site,GdFeVoid_site,
                          p_fe_gd,p_si_gd,
                          p_fe_fe2,p_si_fe2,
                          p_fe_v ,p_si_v)
```

The average structure imposes a set of restrictions on the joint probabilities. For example, the sum of all joint probabilities involving gadolinium $p_{\text{fe_gd}} + p_{\text{si_gd}}$ must be equal to the average occupancy of gadolinium, *i.e.* 1/3. In general, the average structure requires the sum of each column to be equal to the occupancy of the corresponding element from the first variant and the sum of each row to be equal to the occupancy of each element of the second variant. These constraints make the last row and last column of the prob-

ability matrix dependent. Given the upper-left independent part of the correlation matrix, *Yell* can calculate the dependent part. Thus, the previous example can be expressed in a nonredundant form as follows:

```
SubstitutionalCorrelation(Fe3Si_site, GdFeVoid_site, p_fe_gd,
                        p_fe_im)
```

The short version is especially convenient for the most common case of binary disorder, where instead of four values the user needs to provide only one.

3.4. Atomic displacement correlations

The second disorder type that can be refined by *Yell* is the correlation of atomic displacements. These displacements are present in all solids and typically manifest themselves in thermal diffuse scattering, but they may also be of static origin.

Yell assumes that average displacements are harmonic and that real structure displacement correlations follow a Gaussian distribution. The correlations are defined between pairs of modes. For example, if the displacement along the *a* axis of the two iron molecules is correlated, it can be expressed as follows:

```
(1, 0, 0)
ADPCorrelation(iron_molecule_mode_x,
               iron_molecule_mode_x, cov)
```

where (1, 0, 0) indicates that the correlation is between two iron molecules that are one unit cell apart along **a**, *iron_molecule_mode_x* is the translational mode of an average iron molecule and *cov* is the covariance of the two modes expressed in units of Å².

The correlations could be more complicated; for example, a screw correlation between two iron molecules would be defined as

```
(1, 0, 0)
ADPCorrelation(iron_molecule_mode_x,
               iron_molecule_mode_ry, cov)
```

where *iron_molecule_mode_x* and *iron_molecule_mode_ry* are the translational and rotational modes of the iron molecule, and *cov* is expressed in units of rad Å.

3.5. Size effect

The third type of correlation that may be analyzed with *Yell* is the so-called size effect. It is observed if chemical disorder on one site induces a displacement of a chemical unit on another site. *Yell* supports any kind of size effect where the displacements can be described by linear combinations of modes, including motions along any directions and rotations.

The following example defines the relaxation of one iron molecule along the *a* axis in the presence of another iron molecule:

```
SizeEffect(iron_molecule, iron_molecule_mode_x, amp1)
```

where *amp1* is the numerical variable that describes the amplitude of the displacement in Å.

3.6. Formulas

Yell supports definition of variables as arithmetic expression of other variables. Expressions can include the operations + − ∗ /, parentheses (), and the functions *cos*, *sin*, *exp*, *log*, *sqrt*, *abs*, *mod* and *pow*. The use of the expressions is very powerful for reducing the number of parameters that need to be refined. The following example shows how a set of correlation parameters may be constrained to follow an exponential decay, thus reducing the total number of

independent variables from six (*i.e.* *c0*...*c5*) to two (*i.e.* *c0*, *decay_constant*):

```
c0=5;
decay_constant=1.2;
c1=c0*exp(-decay_constant*1);
c2=c0*exp(-decay_constant*2);
c3=c0*exp(-decay_constant*3);
c4=c0*exp(-decay_constant*4);
c5=c0*exp(-decay_constant*5);
```

Advanced programming tools like loops, subroutines and if statements are not yet implemented in *Yell*.

3.7. Symmetry

Yell supports both space-group symmetry operators for expanding the asymmetric unit of the average structure and Laue group symmetry for the 3D-ΔPDF, which is defined by the diffraction symmetry of the diffuse scattering pattern. The automatic restriction of symmetry-related correlation coefficients and automatic determination of multiplicities of special positions are not yet implemented in the current version of the software and have to be defined by the user.

3.8. Refinement

Yell provides both modeling and refinement options. In modeling mode diffuse scattering and the 3D-ΔPDF are calculated from the disorder model, while the refinement mode allows optimization of parameter values through least-squares minimization.

The objective function to be minimized is

$$R_w = \frac{\sum [w(I_{\text{diff}}^{\text{obs}} - I_{\text{diff}}^{\text{calc}})]^2}{\sum (wI_{\text{diff}}^{\text{obs}})^2}, \quad (4)$$

where $I_{\text{diff}}^{\text{obs}}$ and $I_{\text{diff}}^{\text{calc}}$ are the observed and calculated diffuse scattering intensities, and *w* are the refinement weights of each data point. Note that the refinement is done against diffuse scattering intensity and not against the 3D-ΔPDF. In principle, refinements can be performed equally well against real or reciprocal space data, since Plancherel's theorem guarantees that unweighted *R* factors are independent of the choice of the space. Application of refinement weights, however, makes the difference. Despite the fact that weighting in PDF space is occasionally useful for filtering signals (Schaub *et al.*, 2011), *Yell* refines against the diffraction intensities, because this approach is better at accounting for experimental errors or missing data. Also note that, unlike powder PDF programs, *Yell* does not consider Bragg scattering for reasons discussed by Weber & Simonov (2012).

Similarly to Bragg peaks, the diffuse signals are broadened by experimental factors like beam divergence, wavelength dispersion, the quality of a sample or the rotation angle per frame during data collection. These factors are collectively called the resolution function of the experiment. The impact of the resolution function on the results of a local structure determination can be ignored if the diffuse scattering is significantly broader than the profiles of the Bragg peaks, but may become significant if the diffuse signals are relatively narrow (Weber & Simonov, 2012). *Yell* provides a simple approximation to account for resolution effects. It is assumed that the resolution function is uniform in reciprocal space, *i.e.* that each diffraction feature is convoluted with the same kernel, independent of its reciprocal space position. Using the convolution theorem, such a resolution function can be efficiently applied by multiplying the 3D-ΔPDF signals with the Fourier transform of the resolution function. The approximate shape of the resolution function may be estimated from profiles of unsaturated Bragg peaks.

Despite the fact that *Yell*'s least-squares refinements naturally deliver the standard uncertainties (s.u.s) of refined parameters, we recommend to use them with caution. The estimates take into

account the effect of random errors in experimental data under the assumptions that systematic errors are negligible. In typical diffuse scattering experiments this assumption is not true. Also, the number of experimental measurements, *i.e.* the number of pixels used in the refinement, often exceeds the number of model parameters by many orders of magnitude. Such effects may lead to artificially low s.u. values for the refined parameters. We recommend to use the reported s.u. values only for identifying strong numerical correlations or parameters that are weakly represented by the experimental data, but not as a measure of the accuracy of the refined parameters. The extraction of reliable s.u. values from 3D- Δ PDF refinements will be the subject of further investigations.

The experimental information is provided to *Yell* through several hdf5 formatted files (<http://www.hdfgroup.org/HDF5>) containing diffuse scattering intensity and, optionally, refinement weights, a measured pixel mask and resolution correction estimates. The background-corrected diffuse scattering data set is expected to be prepared by the user on a regular grid in the crystallographic coordinate system. Diffuse scattering beneath Bragg reflections is expected to be interpolated, *e.g.* using the punch-and-fill method (Kobas *et al.*, 2005), or masked by using zero weights. After the refinement, *Yell* outputs the calculated diffuse scattering, the observed and experimental 3D- Δ PDF, and a so-called delta-delta PDF or 3D- Δ^2 PDF, which is the difference between the observed and experimental 3D- Δ PDFs.

4. *Yell* internals

4.1. Fast diffuse scattering calculation algorithm

Yell provides the choice of two algorithms for calculating the diffuse scattering intensities from a disorder model. One option is to use equation (2), which is precise but very slow. As an alternative, *Yell* implements an algorithm for the fast calculation of diffuse scattering, which is very similar to the one used by Terwilliger (2011). It is a variation of the fast Fourier transform (FFT)-based method of the structure factor calculation of Sayre (1951) and Ten Eyck (1977).

In our algorithm (see Fig. 1) the 3D- Δ PDF map is constructed by summing up small bunches of single atomic pair 3D- Δ PDF contributions. The interatomic vector $\mathbf{r}_{mn}^{uvw} = \mathbf{R}_{uvw} + \mathbf{r}_{mn}$ of a pair m and n is split into its grid component $\mathbf{r}_{mn,g}^{uvw}$, which fits to the pixel representation of the PDF map, and a residual $\Delta\mathbf{r}_{mn,g}^{uvw} = \mathbf{r}_{mn}^{uvw} - \mathbf{r}_{mn,g}^{uvw}$. Then, a sample box is assigned to $\mathbf{r}_{mn,g}^{uvw}$. The sample box volume is typically some orders of magnitude smaller than the full PDF map but large enough to cover all significant contributions from the correlation function of the electron densities of the atoms m and n , *i.e.* a few Ångström along each dimension. With decreasing sample box size the algorithm becomes faster but less precise. The content of the sample box is calculated from the structure model and the atomic form factors *via* FFT. The intermediate reciprocal space intensities are multiplied by a phase factor $\exp(2\pi i \mathbf{h} \Delta\mathbf{r}_{mn,g}^{uvw})$ to account for the shift $\Delta\mathbf{r}_{mn,g}^{uvw}$. The content and position of a single sample box is therefore calculated as

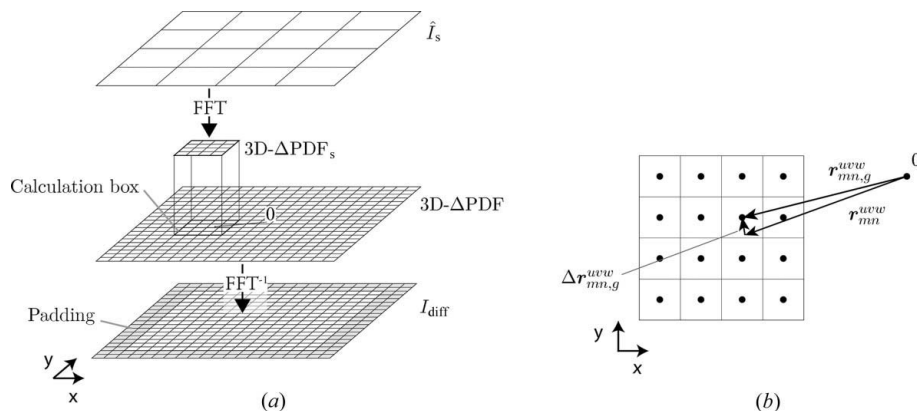


Figure 1

(a) A sketch of our FFT-based diffuse scattering calculation algorithm. The complex scattering signal \hat{I}_s from an interatomic pair is sampled on a rough grid. Then, using FFT, it is transformed into the 3D- Δ PDF_s box, which is added to the full 3D- Δ PDF with an offset of $\mathbf{r}_{mn,g}^{uvw}$ relative to the origin of the 3D- Δ PDF map. The loop is repeated for each interatomic pair. In the last stage the 3D- Δ PDF is transformed into the diffuse scattering map I_{diff} by FFT. (b) Each interatomic vector \mathbf{r}_{mn}^{uvw} is split into two components: the vector to the closest grid point $\mathbf{r}_{mn,g}^{uvw}$ and a difference $\Delta\mathbf{r}_{mn,g}^{uvw}$. The calculation box is constructed around $\mathbf{r}_{mn,g}^{uvw}$. Owing to the requirements of the FFT, the box has an even number of reflections along each dimension. The central pixel thus appears to be offset from the actual center of the box by half a pixel.

$$3D-\Delta PDF_s(\mathbf{x} - \mathbf{r}_{mn,g}^{uvw}) = \text{FFT} \left(\left\{ p_{mn}^{uvw} \exp(-\mathbf{h}^T \boldsymbol{\beta}_{mn}^{uvw} \mathbf{h}) \exp(2\pi i \mathbf{h} \mathbf{u}_{mn}^{uvw}) - c_m c_n \exp[-\mathbf{h}^T (\boldsymbol{\beta}_m + \boldsymbol{\beta}_n) \mathbf{h}] \right\} \exp(2\pi i \mathbf{h} \Delta\mathbf{r}_{mn,g}^{uvw}) f_m^*(\mathbf{h}) f_n(\mathbf{h}) \right). \quad (5)$$

This procedure is repeated for each atomic pair and the results are summed up in the full PDF map. Finally, the diffuse intensities are obtained as the inverse FFT of the PDF map.

Sampling in the limited volume of direct space yields a significant speedup in the computation of 3D- Δ PDF as compared to the direct application of equation (2) but introduces approximation errors, which affect both 3D- Δ PDF intensity and truncation ripples (typical for low-resolution experiments) as shown in Fig. 2.

The approximation error has a distinctive wave-like shape (see Fig. 3) in reciprocal space. The error typically increases towards large

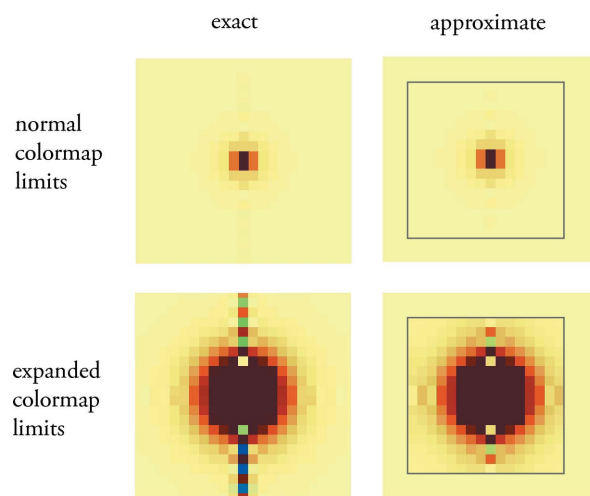


Figure 2

The PDF signal from a pair of platinum atoms calculated with the exact algorithm (left) and the approximation (right); the lower images are shown in a higher contrast to better demonstrate the effect of the approximation. The gray contour outlines a calculation box.

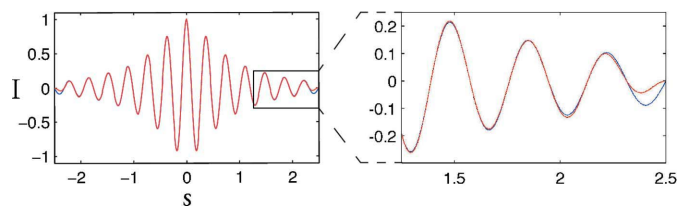


Figure 3
Comparison between exact (blue) and approximate (red) diffuse scattering profiles calculated for the contribution of one atomic pair.

[h]. In order to decrease the error, the approximate diffuse scattering can be calculated on a padded grid, which is then cropped.

Yell allows the user to specify both the number of sampling points and the padding in order to control the accuracy and speed of the diffuse scattering calculation.

Using appropriate settings, our algorithm can provide the approximation with negligible errors. In structure factor calculations they are usually chosen such that the error of this approximation is smaller than 0.5%. For such settings, our algorithm is typically 20–30 times faster than the direct summation.

4.2. Performance

The program is designed to be used on desktop computers. The execution time of the program is roughly proportional to $N_{\text{pix}}N_{\text{pairs}}N_{\text{params}}$, where N_{pix} is the number of pixels in the diffuse scattering map, N_{pairs} is the number of interatomic pairs and N_{params} is the number of refined parameters. Note that the number of pairs is proportional to the square of the number of atoms per chemical unit, which may grow quickly with the number of atoms in the unit cell. Full refinements of relatively simple problems like the one presented in Appendix A can be performed within seconds or minutes using the fast algorithm; medium complexity disorder models may take hours on a modern desktop computer.

The requirements for the operating memory may quickly increase with increasing complexity. Since *Yell* uses a full-matrix refinement, it requires the full Jacobian to be kept in the operating memory, and thus the memory consumption scales with $N_{\text{pix}}N_{\text{params}}$. Typical memory requirements per parameter are of the order of <1 MB for one-dimensional diffuse scattering streaks, 10 MB for two-dimensional diffuse planes, 100 MB for three-dimensional diffuse scattering reconstructed on a rough grid and 1 GB for three-dimensional diffuse scattering on a fine grid.

4.3. Implementation details

Yell is written in the C++ language. The program is a free software with open-source code distributed under the GPLv2 license. It uses several libraries: cctbx (Grosse-Kunstleve *et al.*, 2002) for general crystallographic calculations, boost.spirit for parsing the text input file, hdf5 (<http://www.hdfgroup.org/HDF5>) for binary input files and levmar (Lourakis, 2004) for the minimization using the Levenberg–Marquardt least-squares algorithm.

5. Conclusions

Yell is a program for diffuse scattering interpretation *via* 3D- Δ PDF refinement. It allows the user to explore details of local ordering in single crystals in terms of pairwise correlations without the necessity of explicitly constructing the full crystal model like in MC and MD approaches. This makes the refinements computationally much more

efficient. The fast diffuse scattering calculation algorithm allows disorder problems of medium complexity to be refined in a matter of minutes or hours. The program has been successfully applied to a number of problems (Dshemuchadse *et al.*, 2013; Urban *et al.*, 2014; Simonov *et al.*, 2014; Weber *et al.*, 2014).

Yell is a free software; its source code, documentation and binary executables for Windows and Mac OS X are available at <https://github.com/YellProgram/Yell>.

APPENDIX A

Example of a *Yell* input file

In the following section a commented example of a complete *Yell* model file is presented. It is based on a hypothetical disordered metallic phase with composition $\text{Fe}_{0.5}\square_{0.5}$. On average the structure is cubic closed packed. The structure is short-range ordered: the occupancies of neighboring sites are correlated and each iron atom introduces a relaxation field around it, *i.e.* size effects are observed.

The example is inspired by the disordered alloy $\text{Fe}_{46.5}\text{Ni}_{53.5}$ (Jiang *et al.*, 1996). In order to simplify the input file, the composition was

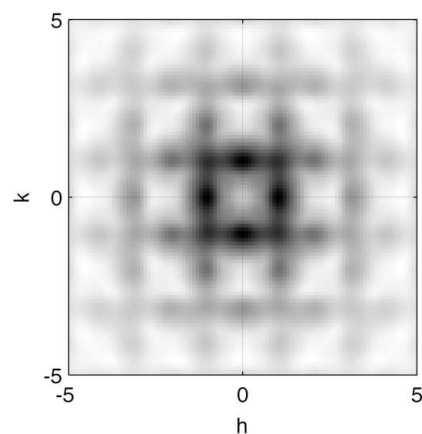


Figure 4
Diffuse scattering $hk0$ section from the example iron–void structure.

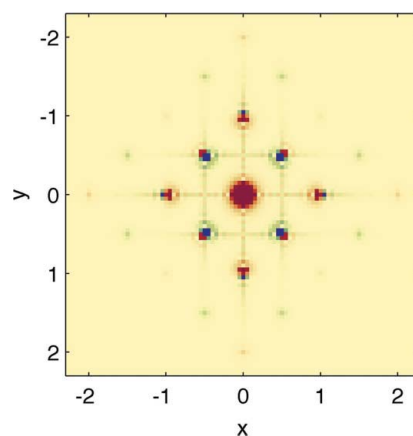


Figure 5
The $xy0$ section of the 3D- Δ PDF from the iron–void example structure. Positive 3D- Δ PDF densities are in red and negative densities are in blue. The strong asymmetries in the signals arise from the assumption that the distance between iron atoms separated by the average vector $(1/2, 1/2, 0)$ is longer than on average, while those separated by $(1, 0, 0)$ show a shorter interatomic distance. See Weber & Simonov (2012) for a more comprehensive discussion about the qualitative interpretation of 3D- Δ PDFs.

changed to 50/50, the Ni atom was replaced by the void and the correlations were restricted to seven neighboring shells; numerical values of correlations were left intact.

The input file starts with a preamble specifying the unit-cell dimensions, the grid for diffuse scattering sampling and the Laue symmetry of the crystal:

```
Cell 3.581 3.581 3.581 90 90 90
DiffuseScatteringGrid -10 -10 -10 0.1 0.1 0.1 200 200 200
LaueSymmetry m-3m
```

In the next step the settings of the program are specified. The calculation is set to a (slow) direct summation, and the refinement is turned off in order to merely calculate diffuse scattering:

```
CalculationMethod exact
Refine false
```

The next section specifies the short-range-order parameters. Their numerical values are usually determined in the course of a 3D- Δ PDF refinement. In the given example the values are taken from Tables II and III of Jiang *et al.* (1996). The meaning of each variable will be specified later.

```
Scale 1
RefinableVariables [
dp110=-0.0192;
dp200=0.0162;
dp211=-5.5000e-04;
dp220=9.2500e-04;
dp310=-0.0025;
dp222=9.2500e-04;
dp400=0.0018;

dx110=0.0211;
dx200=-0.0228;
]
```

The keyword `RefinableVariables` and the following opening and closing square brackets are not required here but would define the variables to be optimized in a least-squares refinement.

Then, the average structure is defined:

```
UnitCell
[
  FeVoid = Variant[
    (p=1/2)
    Fe = Fe 1 0 0 0 0.002

    (p=1/2)
    Void
  ]
]
```

The `Variant` defines a site that is occupied by an iron atom with probability $p = 1/2$ or by a void with the same probability. The `Variant` is associated with an alias `FeVoid` for later reference. The iron atom is defined by the line `Fe = Fe 1 0 0 0 0.002`, listing alias, atomic type, multiplicity, x , y and z coordinates, and an isotropic atomic displacement parameter U_{iso} .

The next section defines the possible displacements of the Fe atom along the x , y and z directions, aliased as `Fe_trans_x`, `Fe_trans_y` and `Fe_trans_z`, respectively:

```
Modes [
  Fe_trans_x = TranslationalMode(Fe, x)
  Fe_trans_y = TranslationalMode(Fe, y)
  Fe_trans_z = TranslationalMode(Fe, z)
]
```

The rest of the file defines short-range correlations of the model. The first entry is the zeroth neighbor, which is special because it represents a correlation of an atom with itself. Such correlation coefficients are to be extracted from the average structure and fixed before the refinement in order to obtain a correct scale factor, which is required to determine all the other correlation coefficients

correctly. In the present example, the probability and atomic displacement parameters of the Fe atom are assigned *via* `SubstitutionalCorrelation` and `ADPCorrelation`, respectively:

```
Correlations
[
  [(0,0,0)
  Multiplicity 1

  SubstitutionalCorrelation(FeVoid, FeVoid, 1/2)

  ADPCorrelation(Fe_trans_x, Fe_trans_x, 0.002)
  ADPCorrelation(Fe_trans_y, Fe_trans_y, 0.002)
  ADPCorrelation(Fe_trans_z, Fe_trans_z, 0.002)
]
```

Here `Multiplicity 1` sets the multiplicity of the interatomic pair with respect to the Laue group of the crystal.

The correlations in the first and the second shell of each iron atom include the substitutional correlation and size effect:

```
[(1/2,1/2,0)
Multiplicity 12

SubstitutionalCorrelation(FeVoid, FeVoid, 0.25+dp110)
SizeEffect(Fe, Fe_trans_x, dx110)
SizeEffect(Fe, Fe_trans_y, dx110)
]

[(1,0,0)
Multiplicity 6

SubstitutionalCorrelation(FeVoid, FeVoid, 0.25+dp200)
SizeEffect(Fe, Fe_trans_x, dx200)
]
```

The values $0.25 + dp110$ and $0.25 + dp200$ are the joint probabilities of finding two iron atoms at the same time separated from each other by vectors $(1/2, 1/2, 0)$ and $(1, 0, 0)$. In the case of a random distribution of iron and voids the joint probability would be $c_{\text{iron}}c_{\text{void}} = 0.5^2 = 0.25$. The variables `dp110` and `dp200` therefore represent the differences between joint probabilities in the real structure and average structure (Patterson probabilities). The values `dx110` and `dx200` contain the displacements due to the size effect.

The following section introduces the rest of the short-range-order parameters up to the seventh nearest neighbor. For the long-distance correlations only substitutional correlations are considered and size effects are ignored. The final square bracket closes the `Correlations` section:

```
[(1,1/2,1/2)
Multiplicity 24
SubstitutionalCorrelation(FeVoid, FeVoid, 0.25+dp211)
]

[(1,1,0)
Multiplicity 12
SubstitutionalCorrelation(FeVoid, FeVoid, 0.25+dp220)
]

[(3/2,1/2,0)
Multiplicity 24
SubstitutionalCorrelation(FeVoid, FeVoid, 0.25+dp310)
]

[(1,1,1)
Multiplicity 8
SubstitutionalCorrelation(FeVoid, FeVoid, 0.25+dp222)
]

[(2,0,0)
Multiplicity 6
SubstitutionalCorrelation(FeVoid, FeVoid, 0.25+dp400)
]
] #End of Correlations
```

Note that only the pair correlations within the asymmetric unit of the Laue group $m\bar{3}m$ need to be provided by the user. Any other pair

correlations are intrinsically computed by applying the Laue symmetry.

The diffuse scattering calculated from this model is presented in Fig. 4 and the 3D- Δ PDF is presented in Fig. 5.

We would like to express gratitude to the authors of the software libraries *levmar* and especially *ctbx* for making them public; this made development of *Yell* much easier. We would also like to express gratitude to Dmitry Logvinovich for help in preparation of the manuscript and testing of the program. This work was supported by the Swiss National Science Foundation grants 200021_121759 and 200020_140389.

References

- Adams, C. P., Lynn, J. W., Mukovskii, Y. M., Arsenov, A. A. & Shulyatev, D. A. (2000). *Phys. Rev. Lett.* **85**, 3954–3957.
- Ali, R., Yashima, M., Matsushita, Y., Yoshioka, H., Ohoyama, K. & Izumi, F. (2008). *Chem. Mater.* **20**, 5203–5208.
- Billinge, S. J. L. (2008). *J. Solid State Chem.* **181**, 1695–1700.
- Chodkiewicz, M., Weber, T., Ahrenberg, L. & Bürgi, H.-B. (2013). *ZODS – Zurich Oak Ridge Disorder Simulation*. University of Zurich, Switzerland.
- Cowley, J. M. F. (1950). *Phys. Rev.* **77**, 669–675.
- Dshemuchadse, J., Bigler, S., Simonov, A., Weber, T. & Steurer, W. (2013). *Acta Cryst.* **B69**, 238–248.
- Goossens, D. J., Heerdegen, A. P., Chan, E. J. & Welberry, T. R. (2011). *Metal. Mater. Trans. A*, **42**, 23–31.
- Grosse-Kunstleve, R. W., Sauter, N. K., Moriarty, N. W. & Adams, P. D. (2002). *J. Appl. Cryst.* **35**, 126–136.
- Henrich, B., Bergamaschi, A., Broennimann, C., Dinapoli, R., Eikenberry, E. F., Johnson, I., Kobas, M., Kraft, P., Mozzanica, A. & Schmitt, B. (2009). *Nucl. Instrum. Methods Phys. Res. Sect. A*, **607**, 247–249.
- Jiang, X., Ice, G. E., Sparks, C. J., Robertson, L. & Zschack, P. (1996). *Phys. Rev. B*, **54**, 3211–3226.
- Kobas, M., Weber, T. & Steurer, W. (2005). *Phys. Rev. B*, **71**, 224205.
- Lourakis, M. I. A. (2004). *Levmar: Levenberg-Marquardt nonlinear least squares algorithms in C/C++*, <http://www.ics.forth.gr/~lourakis/levmar/>.
- Michels-Clark, T. M., Lynch, V. E., Hoffmann, C. M., Hauser, J., Weber, T., Harrison, R. & Bürgi, H. B. (2013). *J. Appl. Cryst.* **46**, 1616–1625.
- Proffen, T. & Billinge, S. J. L. (1999). *J. Appl. Cryst.* **32**, 572–575.
- Proffen, Th. & Neder, R. B. (1997). *J. Appl. Cryst.* **30**, 171–175.
- Refson, K. (2000). *Comput. Phys. Commun.* **126**, 310–329.
- Sayre, D. (1951). *Acta Cryst.* **4**, 362–367.
- Schaub, P., Weber, T. & Steurer, W. (2007). *Philos. Mag.* **87**, 2781–2787.
- Schaub, P., Weber, T. & Steurer, W. (2011). *J. Appl. Cryst.* **44**, 134–149.
- Simonovic, P. & Armbruster, T. (2004). *Am. Mineral.* **89**, 421–431.
- Simonov, A., Weber, T. & Steurer, W. (2014). *J. Appl. Cryst.* In preparation.
- Terwilliger, T. (2011). Personal communication.
- Ten Eyck, L. F. (1977). *Acta Cryst.* **A33**, 486–492.
- Urban, P., Simonov, A., Weber, T. & Oeckler, O. (2014). *J. Appl. Cryst.* Submitted.
- Weber, Th., Boysen, H. & Frey, F. (2000). *Acta Cryst.* **B56**, 132–141.
- Weber, T., Deloudi, S., Kobas, M., Yokoyama, Y., Inoue, A. & Steurer, W. (2008). *J. Appl. Cryst.* **41**, 669–674.
- Weber, T. & Simonov, A. (2012). *Z. Kristallogr.* **227**, 238–247.
- Weber, T., Simonov, A., Malliakas, C. D. & Kanatzidis, M. G. (2014). In preparation.
- Welberry, T. R. & Butler, B. D. (1995). *Chem. Rev.* **95**, 2369–2403.

Chapter 4

Examples

4.1 Complex metallic alloy $hP386\text{-Al}_{57.4}\text{Cu}_{3.5}\text{Ta}_{39.0}$

A new hexagonal complex metallic phase in the system Al-Cu-Ta was discovered in our laboratory. The phase has a refined composition $\text{Al}_{57.4}\text{Cu}_{3.5}\text{Ta}_{39.0}$, very similar to several other complex cubic phases: $cF444\text{-Al}_{63.6}\text{Ta}_{36.4}$, $cF(5928 - 20)\text{-Al}_{56.6}\text{Cu}_{3.9}\text{Ta}_{39.5}$ and $cF(23256 - 122)\text{-Al}_{55.4}\text{Cu}_{5.4}\text{Ta}_{39.2}$. The newly discovered phase shows a peculiar behavior. The quenched sample contained two types of crystals with nearly identical unit cells and average structures: one with and another without diffuse scattering. The task was to investigate the origin of diffuse scattering.

The full result of this investigation is presented in [30]. The current chapter summarizes the part which considers diffuse scattering.

Experiment and data preparation

The diffuse scattering was measured using an in-house diffractometer with a Mo rotating anode x-ray source, a Ge monochromator and a MAR300 image plate detector. The exposure time was 1000 seconds and $\Delta\phi = 0.5^\circ$ per frame. The crystal was indexed in the program XDS [31]. The reconstruction of reciprocal space was performed using the program Xcavate [32]. The background was estimated as an average of the intensities slightly above and below the reconstructed layers.

Average structure

The crystal without diffuse scattering has the space group $P6_3/mmc$ ($\mathbf{a}=13.512(2)\text{\AA}$, $\mathbf{c}=39.022(5)\text{\AA}$) and a relatively complex structure with 39 independent positions (Fig 4.1.1a). Similarly to many inter-metallic compounds, the crystal contains a significant degree of disorder: 1 out of 16 Ta positions and 3 out of 14 Al positions are split, one Al position is deficient, all 5 copper atoms are found on mixed Al/Cu positions occupying them to an extent between 5% to 39%. The crystal structure can be understood as a close-packing of two Frank-Kasper clusters F_{69}^{32} (blue) and F_{74}^{39} (red) with radii of approximately 7\AA (Fig 4.1.1c). The notation F_v^f stands for a cluster with a fullerene-like outer shell having f faces and v vertices [33].

The crystal with diffuse scattering has the same unit cell within the experimental uncertainty and almost identical average structure with slightly higher atomic displacement parameters than

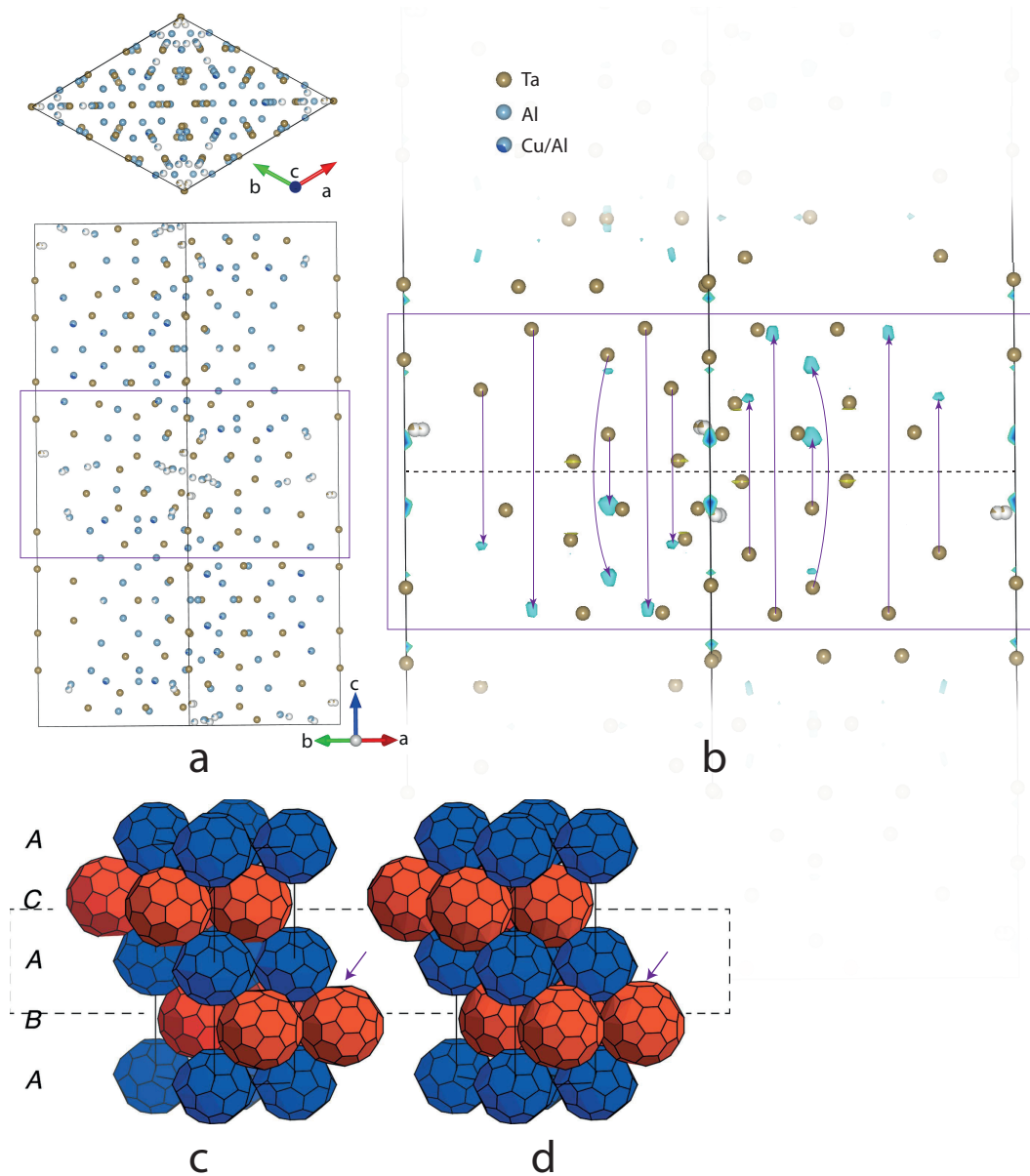


Figure 4.1.1: Structure of $hP386\text{-Al}_{57.4}\text{Cu}_{3.5}\text{Ta}_{39.0}$. **a)** The average structure of the crystal without diffuse scattering. **b)** The difference of the structures with and without diffuse scattering. Blue blobs show the electron density maxima in the latter crystal which correspond to the mirror image of Ta positions in a mirror plane $x, y, 1 - z$ (dashed line). The correspondence is marked by the arrows. In the disorder model all the atoms in the surrounded slab are inverted. **c)** Cluster description of the average structure consisting of two Frank-Kasper clusters F_{69}^{32} (blue) and F_{74}^{39} (red). **d)** The structure of a disordered inlet consisting of a mirror image of F_{69}^{32} cluster (blue) and two F_{76}^{40} clusters (red). The outer shells of F_{74}^{39} and F_{76}^{40} are almost identical except for one additional face (marked with arrow). The images **c** and **d** are taken from [30].

that of the crystal without diffuse scattering. After refinement the difference electron density of this crystal contained several additional maxima. Unfortunately, the crystal selected for diffuse scattering analysis had a large volume and suffered from absorption, thus electron density maps were hard to interpret.

The difference between the two crystals could be understood better from a difference map constructed from coefficients $(|F_{disord}(\mathbf{h})| - |F_{ord}(\mathbf{h})|)e^{i\phi_{ord}(\mathbf{h})}$ where $F_{disord}(\mathbf{h})$ are the experimental form factors of the crystal with diffuse scattering, $F_{ord}(\mathbf{h})$ and $\phi_{ord}(\mathbf{h})$ are the experimental form factors and the refined phases of the crystal without diffuse scattering. Such a map depicts the difference between the average structures of two crystals. The map revealed a set of negative minima close to some Ta positions and a set of positive maxima (blue surfaces in Fig 4.1.1b) related to the Ta positions by a non-crystallographic mirror plane $x, y, 1 - z$. The differences related to Cu and Al atoms were not visible. Overall, the crystal with diffuse scattering contains additional disorder in a form of an “inlet” around the $x, y, 0.5$ plane.

Ordering model

The diffuse scattering is present in the form of sharp streaks parallel to \mathbf{c}^* underneath Bragg peaks (Fig 4.1.2). Due to the broad Bragg peaks and the very large lattice constant $c = 39.022\text{\AA}$, it was impossible to extract the isolated diffuse scattering profiles and to calculate the 3D- Δ PDF map. However, interpretation of the two average structures was sufficient to describe local order.

Based on crystal chemical consideration, it was possible to find a model which describes the presence of additional electron maxima and produces a geometrically plausible structure. It was assumed that with a probability of approx. 5%, a slab around the plane $x, y, 0.5$ is flipped by a mirror plane m_z . As a result, the F_{69}^{32} (blue) cluster transforms to a mirror image of itself, while the cluster F_{74}^{39} (red) transforms into the cluster F_{76}^{40} found in all three related cubic Al-(Cu)-Ta compounds (Fig 4.1.1d). Further, it was assumed that the disordered slab does not disturb the average crystal structure. Since the flipped slabs are rare in the structure, their distribution was assumed to be uncorrelated along \mathbf{c} direction. Diffuse profiles contained no indication of any correlation between the slabs.

In terms of YELL, the model is very simple. A single slab is separated into a **Variant** with two options: flipped and not flipped slabs. The **Correlations** section contains a single entry with a zero-neighbor correlation. The diffuse scattering from this model was very similar to the experimentally observed one, except the model diffuse scattering decreased much slower with increasing \mathbf{c}^* . In order to compensate for this, the atomic displacement parameters U^{33} of all disordered atoms were increased by 0.02\AA^2 .

The final model nicely reproduces the experimental diffuse scattering (Fig 4.1.2). Such correspondence may be considered satisfying for an essentially zero-parameter model, where all of the structural information (apart from additive constant U^{33}) was derived from the chemical interpretation of the average structures and verified by diffuse scattering.

Discussions

This chapter showed that even in cases when diffuse scattering cannot be quantitatively extracted, YELL can be used to confirm a disorder model. Such analysis requires only a visual comparison of

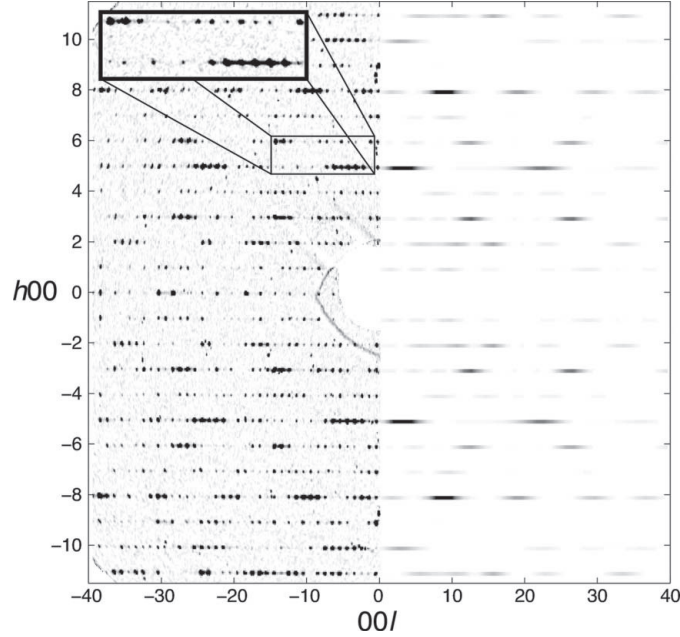


Figure 4.1.2: Experimental (left) and model (right) diffuse scattering from disordered $hP386\text{-Al}_{57.4}\text{Cu}_{3.5}\text{Ta}_{39.0}$. The inset shows the enlarged section containing diffuse scattering between the Bragg peaks. Despite the fact that Bragg peaks are closely spaced, the presence of diffuse scattering is evident. The image is taken from [30].

model and experimental diffuse scattering. In the current case, it allowed to support the model of unusual one dimensional disorder in the form of rare flips in the average structure matrix.

4.2 $\text{Ge}_4\text{Bi}_2\text{Te}_7$

Germanium bismuth tellurides $(\text{GeTe})_n\text{Bi}_2\text{Te}_3$ received much attention due to their thermoelectric properties. The system contains several thermodynamically stable low temperature phases. All of them crystallize in a tetradymite-like layered structures composed of hexagonal anion and cation layers, which are stacked in an fcc-close packed fashion to form slabs. The slabs are separated by van der Waals gaps. At high temperatures such crystals transform into cubic NaCl-type structures. The cations and vacancies become randomly distributed, while the anion positions stay exclusively occupied by Te. Quenching the high temperature phase yields a metastable cubic phase with pronounced vacancy ordering, which plays an important role in defining thermal conductivity and thus the thermoelectric figure of merit and can be potentially optimized by varying the crystal composition and thermal history of the sample. A reliable method for determining the local structure will be vital to controllably produce samples with the optimal thermoelectric properties.

The full study will be published in [34]. The current chapter presents only the diffuse scattering analysis.

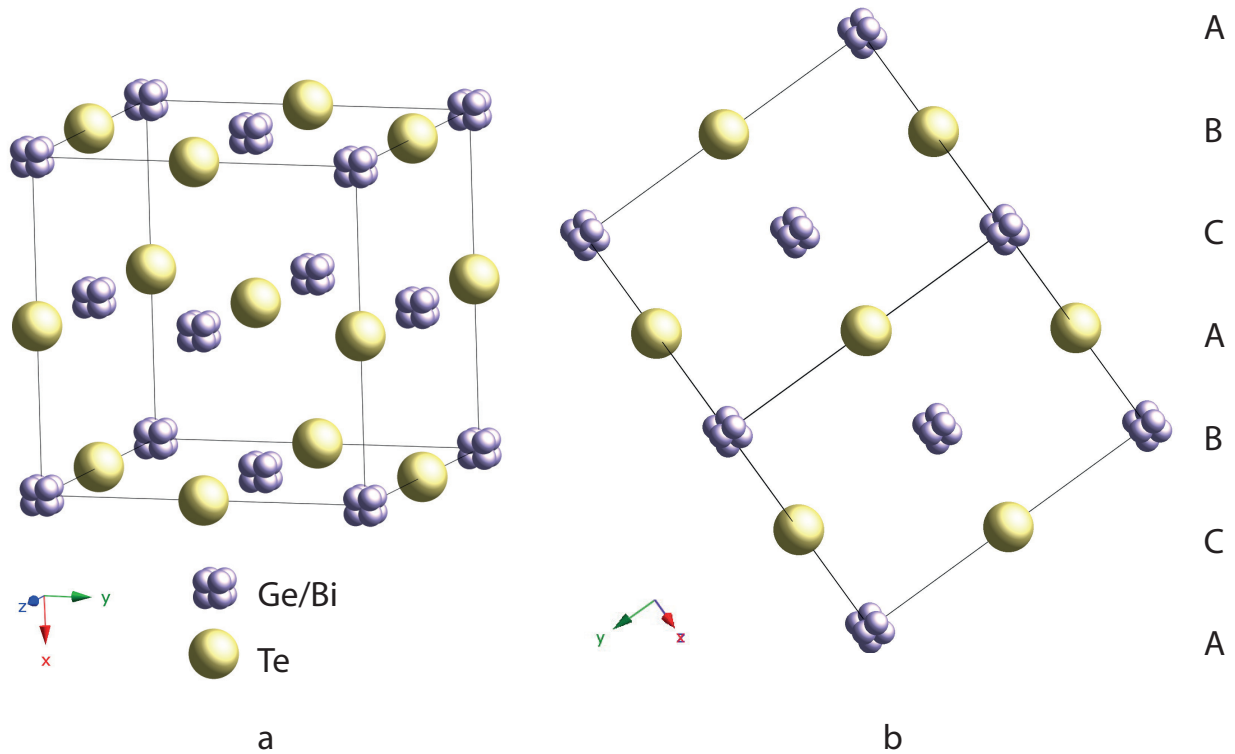


Figure 4.2.1: Average structure of $Ge_4Bi_2Te_7$ along two different projections. **a)** Arbitrary projection. **b)** Illustration of the sequence of layer stacking.

Average structure

The quenched $Ge_4Bi_2Te_7$ crystal has a NaCl type average structure (Fig 4.2.1). The cation position is substitutionally disordered and may be occupied by Ge, Bi or void. In addition, it is split along $\langle xxx \rangle$ directions with a displacement of 0.24\AA . The anion position is fully occupied by Te and does not show any resolvable splitting. However, its atomic displacement parameter ($U_{iso}(Te) = 0.027\text{\AA}^2$) is slightly higher than expected for a heavy atom at room temperature.

Diffuse scattering

The diffuse scattering from $Ge_4Bi_2Te_7$ is present in the form of sharp streaks along $\langle 111 \rangle_c$ directions¹ (Fig 4.2.2). The streaks have a strong intensity modulation, with a peak at position $(h-\delta, k-\delta, l-\delta)$ where $\delta \approx 1/7$, and a flat tail. A very convenient feature of the diffuse scattering is that it decreases to zero close to the Bragg peaks, which makes it easy to eliminate the Bragg reflections. Within the experimental resolution the lateral width of the streaks is comparable to the width of the Bragg reflections, i.e. the order along directions perpendicular to the streaks exceeds the coherence length of the experiment. Therefore, the structure can be well approximated as being built from slabs that

¹The index c corresponds to a cubic unit cell setting.

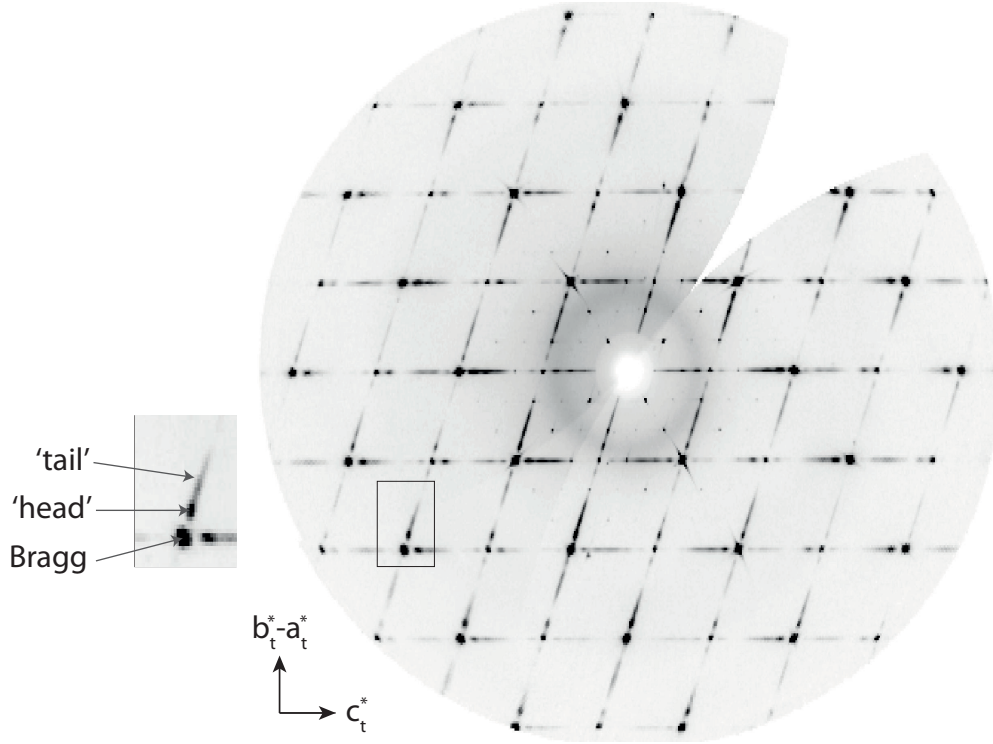


Figure 4.2.2: Diffuse scattering layer $h0l_t$ of $\text{Ge}_4\text{Bi}_2\text{Te}_7$.

are long-range ordered along two dimensions and short-range ordered along the $\langle 111 \rangle_c$ direction. The symmetry of diffuse scattering pattern is $m\bar{3}m$, what is equal to the Laue symmetry of the average structure.

Data preparation

Diffuse scattering was measured at the Materials Science Beam-line ID11 at the synchrotron ESRF (Grenoble, France) using a FReLoN2K CCD detector. The beam energy was set to 26.000 keV (0.47686 Å). The experiment contained a single run of 360° φ rotation with $\Delta\varphi = 0.1^\circ$ per frame.

Diffuse scattering was reconstructed using a home-written script. Corrections for air absorption and polarization were applied. The reconstructions were done using a trigonal basis a_t, b_t, c_t which is related to the cubic one by the following equations: $a_t = b_c/2 - c_c/2$, $b_t = a_c/2 + c_c/2$ and $c_t = a_c + b_c + c_c$. Three reconstructions were performed for the equivalent orientations of the c_t axis along $[111]$, $[\bar{1}\bar{1}\bar{1}]$, $[1\bar{1}\bar{1}]$. The reconstructed volume comprised $501 \times 501 \times 501$ pixels (pixel sizes $[1/37, 1/37, 1/56]$ in fractional coordinates with respect to the trigonal basis). The intensity of each streak was integrated by summing squares of 3×3 pixels along a_t and b_t directions. The background was estimated as the median of the pixels around the integrated area and subtracted. It was assumed that the different set of streaks come from different domains each of which contain disorder along one of the c_t direction and is periodic in the other two directions. Thus, the streaks were averaged in the Laue group $m\bar{3}m$. Outliers were rejected following a modified version of the method proposed

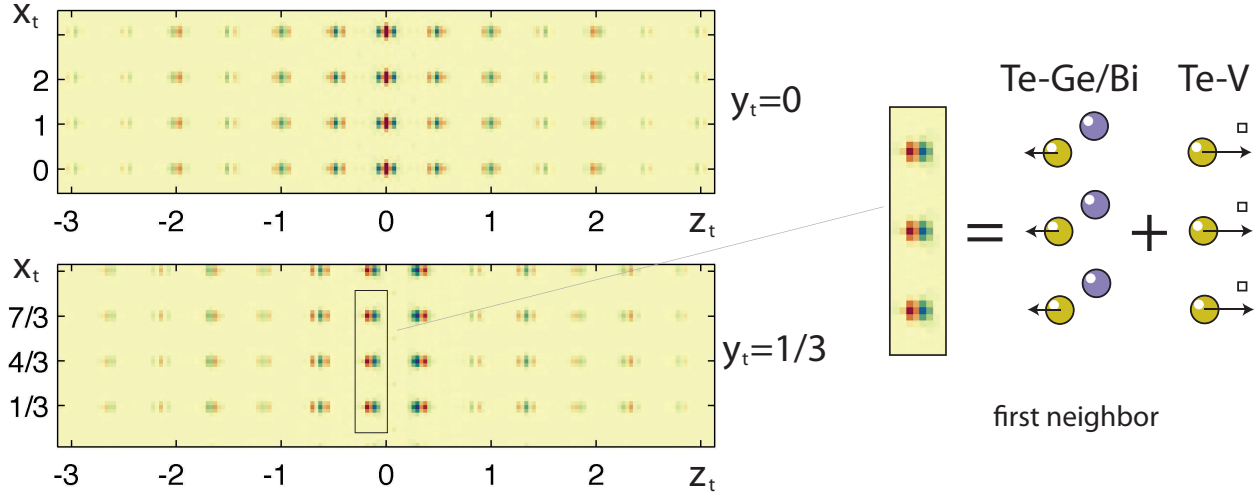


Figure 4.2.3: Experimental 3D- Δ PDF of $Ge_4Bi_2Te_7$. The inset on the right shows the expanded version of the Δ PDF signal of the first nearest neighbor which contains signature of size effect. The image can be understood as follows: when a certain layer is vacant, the layer closest to it moves towards the void, when the layer is occupied by Ge/Bi, the nearest neighbor moves away from it.

by Blessing [35]. The voxel with intensity I_i was considered as outlier if $|I_i - I_{median}| > 2t$, where I_{median} is the median of all symmetry equivalent intensities, $t = median(|I_i - I_{median}|)[n/(n-1)]^{1/2}$ and n is the number of averaged voxels. Bragg peaks as well as $\lambda/3$ and $\lambda/6$ artifacts were deleted and the intensities beneath them were approximated by linear interpolation in order to improve the quality of 3D- Δ PDF maps. Interpolated regions were zero weighted in the course of the refinement. The resolution function of the diffraction experiment was obtained from the profiles of unsaturated Bragg reflections. It was approximated by a Gaussian with a constant half-width of 0.0037 \AA^{-1} and its Fourier transform was used during refinement through `reciprocal_space_multipliers.h5` [11]. Absorption correction was not applied to the diffuse scattering.

3D- Δ PDF analysis

3D- Δ PDF map indicates the presence of a size effect with displacement vectors along the \mathbf{c}_t direction (Fig 4.2.3). First and second neighbors show “positive-negative” signals, which are characteristic for size effect with heavier interatomic pairs being separated stronger than the lighter pairs. The further neighbors contain more complicated signals.

It was decided not to use a `SizeEffect` model because the average structure contains split atoms. In such case the refinement of `SubstitutionalCorrelations` between different alternative atomic positions effectively simulates size effect and provides better accuracy and more flexibility.

Model of ordering

It is convenient to describe the ordering of the crystal in a trigonal setting (Fig 4.2.1b). In such a cell, the average crystal structure consists of hexagonal layers arranged in an ABCABC sequence.

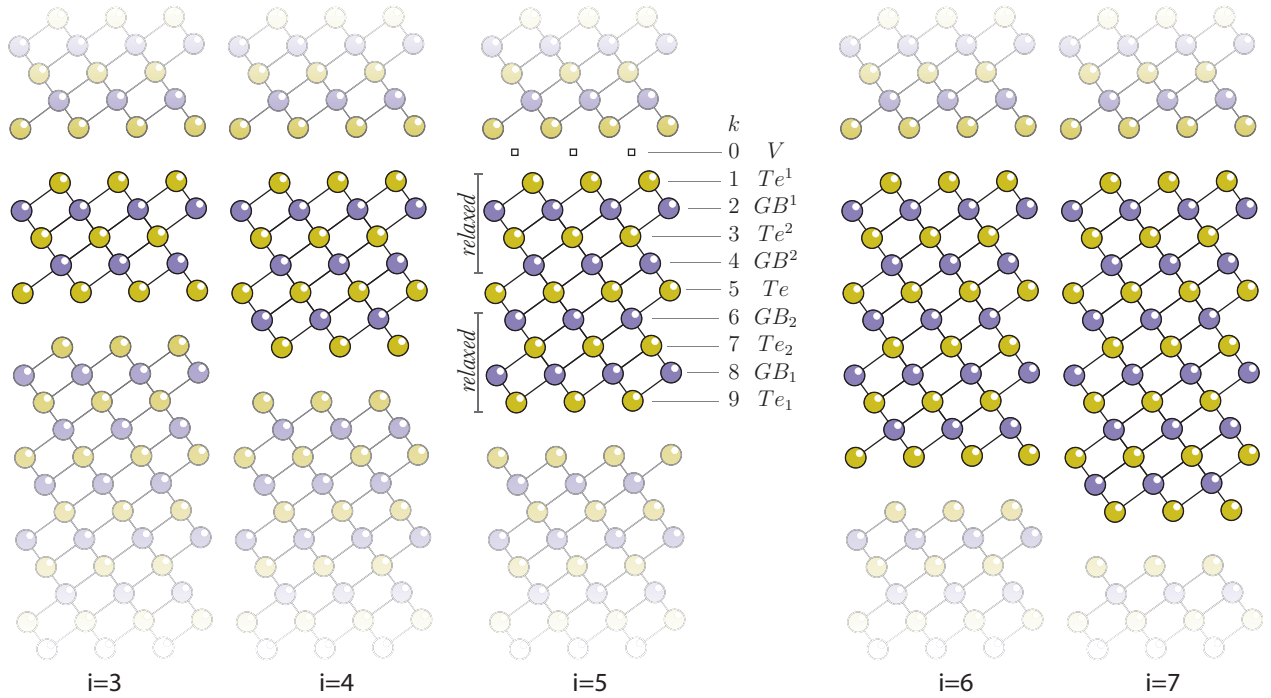


Figure 4.2.4: Examples of slabs which can be present in $\text{Ge}_4\text{Bi}_2\text{Te}_7$.

With a certain probability, a full cation layer is replaced by a void layer periodic along \mathbf{a}_t and \mathbf{b}_t directions, what was also observed in many ordered structures of the same system (e.g. $\text{Bi}_2\text{Ge}_3\text{Te}_6$ [36] 658931-ICSD).

The disorder model may be described as follows. The crystal consists of ordered slabs of different sizes. The slabs are terminated on both sides by Te layers and are separated by void layers (Fig 4.2.4), the so-called van der Waals gaps. The two anion layers closest to the void (labeled Te_1 and Te_2 if they are above and Te^1 and Te^2 if they are below the void layer) are allowed to relax along \mathbf{c}_t direction. The two cation layers closest to the void (labeled GB_1/GB^1 and GB_2/GB^2) may also relax along \mathbf{c}_t , and, in addition, their composition may differ from the bulk. We will call p_i the probability to find a slab with i anion layers. It is assumed that slabs are randomly stacked, i.e. that the probability to find a certain slab does not depend on the size of the preceding slab.

Based on the 3D- Δ PDF map it was assumed that slabs with $i > 15$ do not exist. In the course of refinement it turned out that the probability of finding layers with $i = 3$ was refining to a small negative number. Thus, it was assumed that such slab does not exist either.

Totally, the final model comprises 18 parameters to be refined: a scale factor, eleven probabilities p_4, \dots, p_{14} (p_{15} is obtained from the condition $\sum p_i = 1$), two parameters for defining ratio Bi:Gd in the layers GB_1/GB^1 and GB_2/GB^2 , and four displacements for layers Te_1/Te^1 , Te_2/Te^2 , GB_1/GB^1 and GB_2/GB^2 .

Definition of the stacking model in YELL

Average structure

The average structure of the crystal consists of two **Variants**. The first variant describes the Te position and contains five alternatives: bulk Te, Te_1 , Te^1 , Te_2 and Te^2 . The second variant describes the cation position and contains six alternatives: void (V), bulk GB, GB_1 , GB^1 , GB_2 and GB^2 .

The average occupancies of each of the alternatives depend on the slab probabilities p_i . For example, the occupancy of the void can be defined as $c(V) = \frac{N(V)}{N(Cat)+N(V)}$ where $N(V)$ is the number of void layers in the structure and $N(Cat)$ is the number of cation layers in the structure. Each slab contains i cation layers, and one void layer. Thus, $N(Cat) + N(V) = N_s \sum_i ip_i$ where N_s is the number of slabs in the crystal, while $N(V) = N_s$. Thus:

$$c(V) = \frac{N(V)}{N(Cat) + N(V)} = \frac{N_s}{N_s \sum_i ip_i} = \frac{1}{\sum_i ip_i}$$

In order to input this model in YELL, similar calculations need to be performed for each of the alternatives. The calculations are straightforward but tedious. In order to avoid the risk of introducing errors, the joint probabilities were calculated as a special case of substitutional correlations using equation (4.2.5) described below.

Correlations

$Ge_4Bi_2Te_7$ contains only **SubstitutionalCorrelations**. Let us denote the joint probability of finding a layer A and a layer B separated by k layers as $P^k(A, B)$. Here, the index k counts both cation and anion layers (Fig. 4.2.4), $A, B \in \{V, GB, GB_1, GB^1, GB_2, GB^2, Te, Te_1, Te^1, Te_2, Te^2\}$. The probabilities $P^k(A, B)$ are too complex to be derived by hand. They were calculated using the following recursive procedure.

Step 1. Calculate the structure composition after a van der Waals gap.

Assume that a certain position $k = 0$ in the crystal is occupied by a void layer. Independent from the size of the slab that follows the void layer, the layer at $k = 1$ will be occupied by Te^1 , and the layer at $k = 2$ will be occupied by GB^1 , as both of the layers appear in each slab. The third layer depends on the slab which follows the void. If it is the slab $i = 3$ than the third layer is occupied by Te , otherwise it is occupied by Te^2 . Thus, the third layer can be described as a linear combination $p_3Te + (1 - p_3)Te^2$. In general, the composition of layers $k = 0, \dots, 5$ can be calculated using a linear combination of all slabs taken with appropriate probabilities:

$$\begin{bmatrix} V \\ Te^1 \\ GB^1 \\ p_3Te + (1 - p_3)Te^2 \\ p_3GB_1 + p_4GB + (1 - p_3 - p_4)GB^2 \\ p_3Te_1 + p_4Te_2 + (1 - p_3 - p_4)Te \\ - \end{bmatrix} = p_3 \begin{bmatrix} V \\ Te^1 \\ GB^1 \\ Te \\ GB_1 \\ Te_1 \\ - \end{bmatrix} + p_4 \begin{bmatrix} V \\ Te^1 \\ GB^1 \\ Te^2 \\ GB \\ Te_2 \\ \dots \end{bmatrix} + p_5 \begin{bmatrix} V \\ Te^1 \\ GB^1 \\ Te^2 \\ GB^2 \\ Te \\ \dots \end{bmatrix} + \dots \quad (4.2.1)$$

The same calculation is depicted graphically at Fig. 4.2.5.

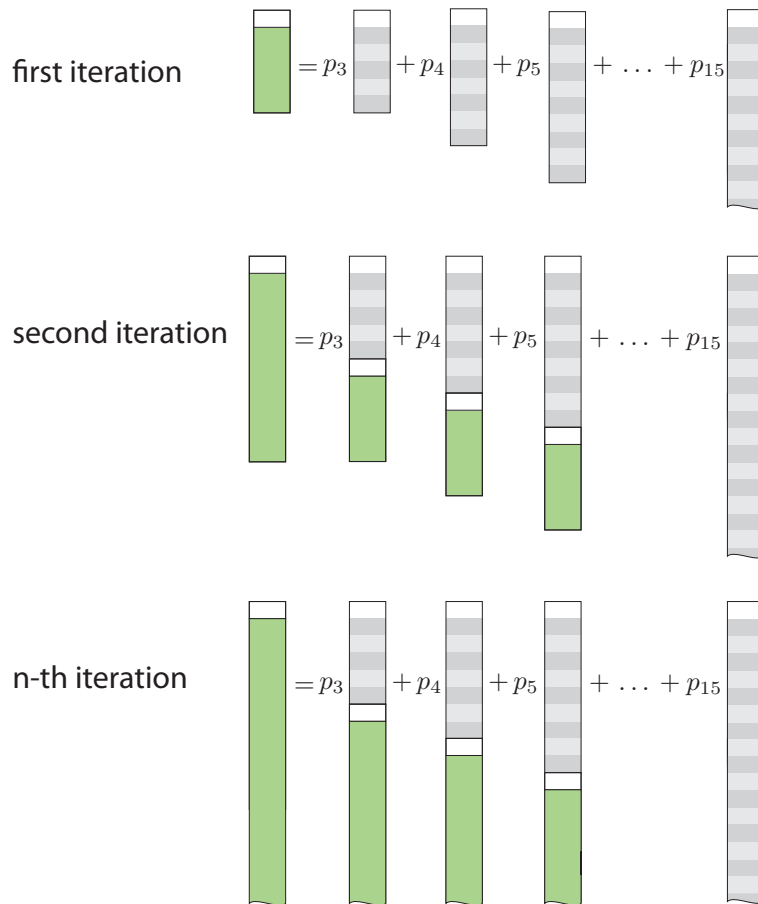


Figure 4.2.5: Schematic representation of an algorithm for calculating the composition of layers after void $P^k(A|V$ at 0) which is marked green. In the first iteration the composition of the first six layers is calculated as a linear combination of slabs i taken with probabilities p_i . Second iteration utilizes results of the first iteration to calculate composition of six more layers. Each consequent iteration uses the result of the previous iteration in order to calculate six more layers.

The layer $k = 6$ can not be determined from eq. (4.2.1) because the slab $i = 3$ contains only 6 layers with $k = 0...5$. However, notice that the layer that follows the slab $i = 3$ is a void layer. The void layer is followed by Te^1 then GB^1 and so on as described by left side of eq. (4.2.1). Thus, the result of the first iteration (marked green in Fig. 4.2.5) can be attached to the ends of each slab and the same calculation can be repeated in order to obtain layers for $k = 6, \dots, 11$. Then, the result of the second iteration can be used in the third iteration to obtain layers for $k = 12, \dots, 17$. In such a way it is possible to calculate the composition at any distance k from a void layer.

More formally, we call $P^k(A|V \text{ at } 0)$ a conditional probability to find a layer A at a position k given that the position $k = 0$ is occupied by V . It can be calculated as follows:

$$P^k(A|V \text{ at } 0) = \sum_i p_i C_i^k(A) \quad (4.2.2)$$

where $C_i^k(A)$ is equal to the probability to find a layer A at a position k given that the slab of size i starts at position $k = 0$. In turn, $C_i^k(A)$ can be calculated using the following equation:

$$C_i^k(A) = \begin{cases} 1 & \text{if } k < 2i \text{ and position } k \text{ in slab } i \text{ is occupied by a layer } A \\ 0 & \text{if } k < 2i \text{ and position } k \text{ in slab } i \text{ is occupied by a layer } \neq A \\ P^{k-2i}(A|V \text{ at } 0) & \text{if } k \geq 2i \end{cases} \quad (4.2.3)$$

Step 2. Calculate pairwise probabilities

Pairwise probabilities $P^k(A, B)$ are calculated by summing probability $C_i^j(A)$ to find layer A at each position j in each slab i multiplied by probability $C_i^{j+k}(B)$ to find layer B at position $j + k$:

$$P^k(A, B) = c(V) \sum_i p_i \sum_j C_i^j(A) C_i^{j+k}(B) \quad (4.2.4)$$

Average occupancies of each layer A can be calculated as a special case of eq. (4.2.4)

$$c(A) = P^0(A, A) \quad (4.2.5)$$

The algorithm described above was implemented using symbolic computation toolbox in Matlab and was then transformed into YELL input file using arithmetic expressions.

Results

The final refinement with YELL yielded $R_w = 0.25$. The R-factor R_w is defined as

$$R_w = \frac{\sum w(I_{exp} - I_{model})^2}{\sum w I_{exp}^2}$$

where w is weight of the experimental data, I_{exp} and I_{model} are the intensities of experimental and model diffuse scattering. Fig. 4.2.6 shows the comparison between experimental and model 3D- Δ PDF. Refined probabilities p_i are depicted in Fig. 4.2.7a. The distribution of the thickness of the slabs has a relatively sharp peak at $i = 6$ and a long tail. The probability for the two largest layers p_{14} and p_{15} is slightly bigger than p_{13} . It is expected that the probabilities should monotonically decay,

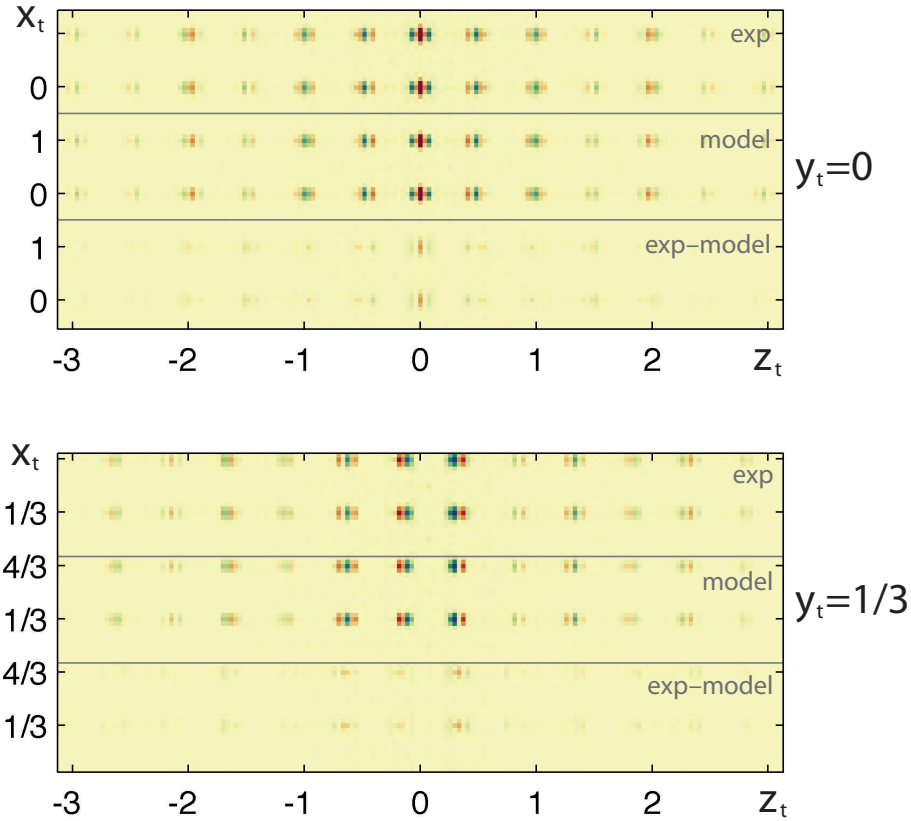


Figure 4.2.6: Comparison between experimental and model 3D- Δ PDF in $\text{Ge}_4\text{Bi}_2\text{Te}_7$.

thus the increase in the probabilities of the last two members may be attributed to the limitations of the model, which assumes that the distribution of the blocks is independent or to systematic errors in the data.

The relaxation of the four layers closest to the van der Waals gap is presented at Fig. 4.2.7b. The Te^1 and GB^1 move towards each other and detach from the bulk. In order to support the layer dimerization, the composition of GB^1 layer is enriched with the trivalent Bi. The distance between Te^2 and GB^2 is also decreased, though the effect is less pronounced. In order to achieve this, the Te^2 layer actually moves *away* from the void.

During refinement the concentration of void layers was not fixed and was defined by the refined probabilities p_i . The final value $p(V) = 0.134$, however, is in good agreement with the value obtained for the average structure 0.143. The total composition of the crystal converged to $\text{Ge}_{0.61}\text{Bi}_{0.26}\text{Te}$ which slightly deviates from the average structure composition $\text{Ge}_{0.56}\text{Bi}_{0.29}\text{Te}$. Refined interatomic distances compare reasonably well to that of an ordered phase $\text{Ge}_2\text{Bi}_2\text{Te}_5$ [37] with similar structural properties (Fig 4.2.7).

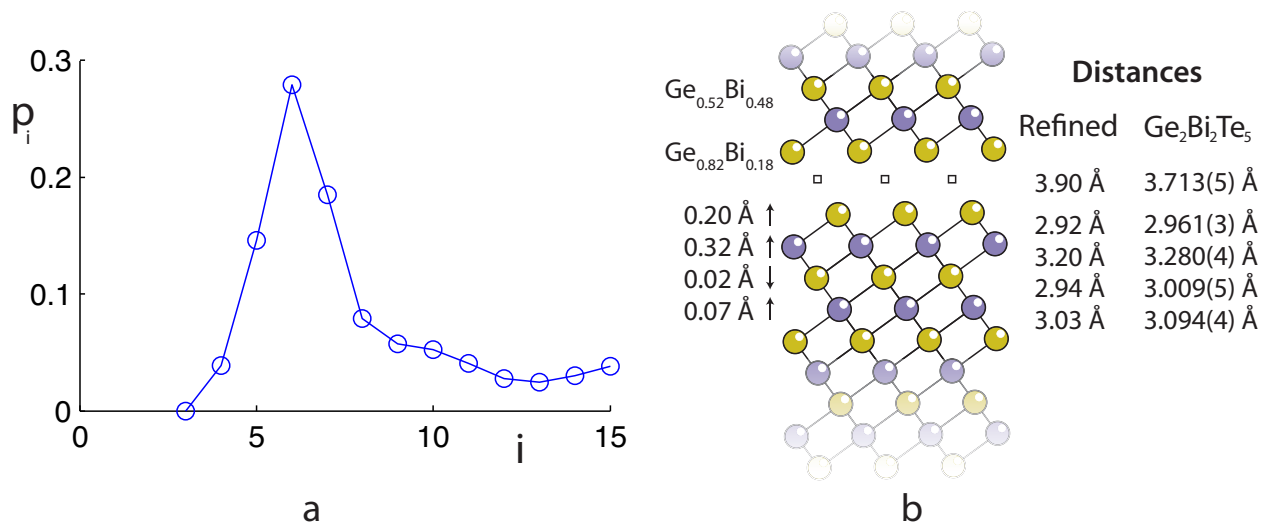


Figure 4.2.7: Refined frequencies of slabs (a) and interatomic distances (b). The interatomic distances for $\text{Ge}_2\text{Bi}_2\text{Te}_5$ are taken from [37].

Discussion

It has been shown that the distribution of slabs of different sizes can be reliably investigated. Disorder in the current crystal has a dominating effect on phonon scattering and thus determines thermal conductivity and the thermoelectric figure of merit. It would be interesting to follow evolution of disorder and thermoelectric efficiency of this and similar systems as a function of composition and temperature treatment.

The described algorithm for transforming a stacking model into a YELL input file could be implemented using 157 lines of Matlab code². The algorithm can be trivially generalized for cases when the probability of each next slab depends on the previous slab (i.e. stacking model is described by a Markov chain). Using the described algorithm it is possible to write a general program which will help preparing YELL input for any stacking fault model.

4.3 Tris-t-butyl tricarboxamide

In the following paper investigation of 2D disorder from the organic crystal of tris-t-butyl tricarboxamide is presented. In addition, the influence of various systematic errors coming from experiment and data treatment is carefully analyzed. It is shown that 3D- Δ PDF refinement provides reliable short range order parameters on the level of 0.04% for probabilities and 0.01Å for size effect parameters if systematic errors are carefully taken into account. The correlations for interatomic vectors of up to 100Å were successfully refined.

The paper is close to its final version and is intended to be submitted to the Journal of Applied Crystallography.

²For full code see appendix A

Experimental uncertainties of Three Dimensional Pair Distribution Function investigations exemplified on the diffuse scattering from a tris-*t*-butyl-1,3,5-benzene tricarboxamide single crystal

ARKADIY SIMONOV, THOMAS WEBER AND WALTER STEURER

*Laboratory of Crystallography, Department of Materials, ETH Zurich,
Vladimir-Prelog-Weg 5, CH-8093 Zurich, Switzerland.*

Abstract

Diffuse scattering from a substitutionally disordered tris-*t*-butyl-1,3,5-benzene tricarboxamide single crystal is analyzed with the Three-Dimensional Difference Pair Distribution Function (3D- Δ PDF) method. The real structure of the crystal is shown to consist of infinite polar molecular stacks along the *c* axis, which are laterally packed in a hexagonal fashion. The orientation of the stacks is disordered, but neighboring stacks strongly prefer anti-parallel arrangements. Quantitative orientational pair correlation coefficients are determined for all lateral pairs separated by less than 100Å. A careful analysis of the factors influencing the accuracy of the 3D- Δ PDF refinement is presented. It is shown, that the effect of statistical errors is small compared to systematic errors coming from diffraction geometry distortions, reciprocal space resolution or incompletely corrected background. Various strategies for identifying and decreasing systematic errors are discussed. The impact of the systematic errors on the uncertainty of the results is not specific for 3D- Δ PDF investigations, but also relevant for other quantitative diffuse scattering modeling techniques.

1. Introduction

With the recent advances in X-ray area detector technology, high quality diffuse scattering has become routinely available. This opens new perspectives for investigating the local structure of disordered crystals and provides valuable information about a broad spectrum of phenomena ranging from catalytic activity (Simoncic & Armbruster, 2004) to giant magneto resistance (Adams *et al.*, 2000). Several methods are currently used for diffuse scattering analysis, the most popular being the powder pair distribution function (PDF) method (Billinge, 2008), as well as single crystal Monte-Carlo (MC) (Butler & Welberry, 1992) and reverse Monte-Carlo (RMC) (Nield *et al.*, 1995) simulations.

Recently, we have introduced a new method called Three-Dimensional Difference Pair Distribution Function (3D- Δ PDF) modeling (Weber & Simonov, 2012) and have presented the program YELL, which can perform such refinements (Simonov *et al.*, 2014). The 3D- Δ PDF method provides direct access to local structure properties in terms of pairwise interatomic or intermolecular correlations. It gives a unified description to all types of disorder including substitutional, displacive and size effect phenomena.

Here we present a case study showing that the 3D- Δ PDF refinement method can reliably provide quantitative information about local order. Various factors influencing the uncertainty of the refined correlation coefficients are carefully analyzed. The factors are not specific for 3D- Δ PDF refinements, but relevant for any quantitative refinement method based on single crystal diffuse scattering including MC and RMC.

A single crystal of N,N',N''-tris-t-butyl-1,3,5-benzene tricarboxamide was selected as a model system. The material acts as a nucleating agent and an effective clarifier in the production of isotactic polypropylene (Blomenhofer *et al.*, 2005). It belongs to a family of substituted 1,3,5-benzene tricarboxamides, which also received attention

for their ability to form supramolecular polymers, i.e. polymers stabilized by weak hydrogen bonds and π -stacking between monomers. The filaments may show liquid crystalline behavior (Matsunaga *et al.*, 1986) and induce a gelation in a variety of liquids (Hanabusa *et al.*, 1997). Single filaments have a macroscopic dipole moment, which plays an important role in the process of self-assembly (Albuquerque *et al.*, 2013).

The average crystal structure, which has a hexagonal lattice ($a=14.114(2)\text{\AA}$, $c=6.930(1)\text{\AA}$) with space group $P6_3/m$, is reported in (Kristiansen *et al.*, 2009). The molecules are stacked on top of each other along \mathbf{c} forming infinite columnar stacks with rod group $p6_3$ (Fig. 1). Within the stacks the molecules are connected by a helical net of hydrogen bonds. Laterally, the columns are packed in a two-dimensional hexagonal closed fashion. Stacks appear in different chirality, which, in the average structure, are related by a mirror plane perpendicular to the \mathbf{c} axis, such that interpenetrating “up” and “down” orientations are present with equal probability.

2. Theory

The 3D- Δ PDF is the difference between the autocorrelation function of the real crystal $P_{tot}(\mathbf{r})$ and the autocorrelation function of the average structure (= Patterson function) $P_{average}(\mathbf{r})$. It can be obtained as the Fourier transform (FT) of the diffuse scattering:

$$3D-\Delta PDF(\mathbf{r}) = P_{tot}(\mathbf{r}) - P_{average}(\mathbf{r}) = FT[I_{diffuse}(\mathbf{h})]$$

The 3D- Δ PDF provides information about the local order in a crystal in terms of the distribution of interatomic vectors. 3D- Δ PDF values are positive if the probability for finding an atomic pair separated by a vector \mathbf{r} is higher than expected from the average structure and negative otherwise.

In the case that any displacive correlations apart from size effects are neglected

and the crystal consists of two building blocks A and B per unit cell, which mutually replace each other with the same average probability of 0.5, equation (9) in (Weber & Simonov, 2012) may be simplified to:

$$I_{diffuse}(\mathbf{h}) = \sum_{uvw} \sum_{i,j=\{A,B\}} \left[p_{uvw}^{ij} \exp(2\pi i \mathbf{h}(\mathbf{r}_{uvw} + \mathbf{u}_{uvw}^{ij})) - \frac{1}{4} \exp(2\pi i \mathbf{h} \mathbf{r}_{uvw}) \right] F_j(\mathbf{h}) F_i^*(\mathbf{h}) \quad (1)$$

where i and j may be either A or B , F_i are the molecular form factors, i.e. the Fourier transforms of the corresponding building blocks, p_{uvw}^{ij} is the joint probability for finding a molecule i in one unit cell and simultaneously a molecule j separated by a lattice vector uvw , and \mathbf{u}_{uvw}^{ij} is a size effect parameter, which is equal to the difference between the real and the average length of the corresponding interatomic vector. Not all of the size effect parameters are independent. To be consistent with the average structure they must follow the equation

$$\sum_{ij} p_{uvw}^{ij} \mathbf{u}_{uvw}^{ij} = 0$$

In the case of binary disorder it is convenient to replace the joint probability parameters p_{uvw}^{ij} by correlation coefficients c_{uvw} defined as:

$$c_{uvw} = p_{uvw}^{AA} + p_{uvw}^{BB} - p_{uvw}^{AB} - p_{uvw}^{BA}$$

Such coefficients may take values from -1 to 1. A value of -1 means that the corresponding pairs of unit cells are always occupied by different constituents, while a value of 1 indicates that they are always the same. By definition $c_{000} = 1$, because a molecule is never disordered with respect to itself.

3. Experiment and data processing

The experimental data was taken from Schaub *et al.* (2007). The experiment was performed using a Mar345 detector at the Swiss-Norwegian beam lines at the synchrotron

ESRF, Grenoble, France. The data was collected in a single run with a total rotation of 256.5° around the ϕ -axis ($\Delta\phi = 0.25^\circ/\text{frame}$) and a wavelength $\lambda = 0.75\text{\AA}$. The distance from the crystal to the detector was 180mm.

Diffuse scattering, which appears in layers, was reconstructed with the program Xcavate (Estermann & Steurer, 1998) based on the orientation matrix determined by the program XDS (Kabsch, 2010). The intensity of each layer was integrated along the \mathbf{c} direction to account for resolution effects. The background was estimated as the average of the intensities slightly above and below the diffuse scattering layers and finally subtracted. Bragg peaks were eliminated with the punch-and-fill method (Kobas *et al.*, 2005), i.e. pixels affected by Bragg scattering were set to the average intensity of the surrounding pixels. The diffuse scattering intensity was averaged in the Laue group $6/m$. Since Xcavate reconstructions are mapped in cartesian reciprocal space coordinates, averaging is only possible in the common subgroup $2/m$. Therefore reconstructions were performed in 3 equivalent orientations around the \mathbf{c} axis to allow averaging over the full Laue symmetry. The details of the experiment are summarized in the Table 1.

Table 1. *Characteristics of the experimental diffuse scattering.*

Experimental resolution	1.10Å
Number of reconstructed layers	13, $l=-6\dots6$
Diffuse scattering pixel size within the layers	$0.0015\text{\AA}^{-1} \times 0.0015\text{\AA}^{-1}$
Dimensions of reconstructed array	$1213 \times 1213 \times 13 \approx 1.9 \times 10^7$
Number of symmetry independent voxels	891'871
Average redundancy	8.1
$R_{int} = \sqrt{\frac{\sum (I_i - \langle I \rangle)^2}{\sum I_i^2}}$	18%
$\langle \frac{I}{\sigma_r} \rangle$	1.96

Estimated uncertainties of measured pixels σ_{exp} were calculated as the square-root of the variance of symmetry equivalent pixel values. As a consequence of the small redundancy the estimates of σ_{exp} showed significant noise (Fig. 2). This posed the risk that data points whose equivalent pixels by accident have a small variance, would get unjustifiably strong weights in the refinement. In order to avoid such situations

and to decrease the noise level in the σ_{exp} maps, a new set of uncertainties σ_r was determined. It was assumed that data points having similar intensities, background contributions and positions in reciprocal space are likely to have similar uncertainties as well. The standard uncertainties of individual pixels were therefore determined using a linear regression of σ_{exp} on a constant term C , the length of the scattering vector h (which affects the experimental resolution), the square root of total scattering intensity $\sqrt{I} = \sqrt{I_{background} + I_{diffuse}}$ and diffuse scattering $I_{diffuse}$ (which both affect count statistics). The model for estimating σ_r is thus:

$$\sigma_r = C + \alpha_b \sqrt{I} + \alpha_d I_{diffuse} + \alpha_s h$$

Regression was performed separately for each layer. As a consequence, the estimates σ_r became much smoother than σ_{exp} . On the other hand, uncertainties of pixels affected by singular experimental artifacts became underestimated. In order to down-weight such regions the estimates σ_r were changed to σ_{exp} if $\sigma_{exp} > 3\sigma_r$.

4. Diffuse scattering and local structure determination

Diffuse scattering is present as sharp layers perpendicular to \mathbf{c}^* at integer l positions. Within the layers, the diffuse scattering is found close to the Brillouin zone boundaries forming a honeycomb-like pattern (Fig. 2). The $hk0$ layer shows no diffuse scattering.

The sharpness of diffuse scattering along the \mathbf{c}^* direction means that the real structure is well ordered along \mathbf{c} and disordered only along \mathbf{a} and \mathbf{b} directions, which agrees well with the assumption that the molecules form infinite ordered stacks. The fact that the $hk0$ layer contains no visible diffuse scattering means that the projection of the structure along the \mathbf{c} axis is not disordered. This is consistent with the average structure feature that “up” and “down” stacks are related by a mirror plane perpendicular to \mathbf{c} , making them indistinguishable and thus perfectly ordered in the projection along

c.

The honeycomb shape of the diffuse scattering is a typical signature of a frustrated triangular lattice (see e.g. Welberry *et al.*, 2011). Such frustration arises when nearest neighbors prefer antiparallel orientations, but the odd-numbered triangular lattice does not allow to make all of the contacts heterogeneous.

The $xy0$ cross section from the experimental 3D- Δ PDF is depicted in Fig. 3. The picture consists of flower-like patterns forming a hexagonal lattice. Each of the flowers represents an up-down correlation between a pair of molecular stacks. A positive (red) signal at a lattice point \mathbf{r} means that a pair of stacks separated by such a vector tend to be in parallel orientation, while negative (green-blue) patterns indicate preference for antiparallel arrangements. The densities of the patterns are proportional to the corresponding correlation coefficients $c_{u\mathbf{v}0} = c_{\mathbf{v}u}$. The pattern in the center of 3D- Δ PDF space represents the zeroth neighbor, i.e. the correlation of a molecule with itself, which, by definition, has a perfect positive correlation $c_{00} = 1$. The first neighbors show negative correlations and the further shells alternate between positive and negative correlations with decreasing strength.

The numerical values of the correlation coefficients were obtained via least squares refinement against the diffuse scattering. The minimization criterion was

$$R_w = \left[\frac{\sum ((I_{exp} - I_{model})/\sigma_r)^2}{\sum (I_{exp}/\sigma_r)^2} \right]^{\frac{1}{2}}$$

The diffuse scattering was calculated using equation (1). The form factors F_{up} and F_{down} were computed in an early version of the program YELL, which did not yet allow least-squares refinements. The minimization was therefore done with the program Matlab (2011b). The final model covers correlation coefficients for the first seven shells and size effect parameters for the first three shells resulting in a total of 28 correlation coefficients and 12 size effect parameters.

From a statistical perspective, all refined values are highly significant. Under the

assumption that the uncertainty of refined variables is determined only by measurement statistics, the standard uncertainties were in the order of 0.0003 for substitution correlations and 0.0002Å for the size effect. These values define the lower bound of the refinement uncertainty and can only be achieved when the effects of all systematic errors are perfectly accounted for.

The refined correlation coefficients were corrected for reciprocal space resolution effects. In the case of an ideal experiment Bragg reflections would be measured as δ -functions and the Fourier transform of reconstructions showing only sharp Bragg peaks would be perfectly periodic. In a real experiment, however, the Bragg reflections are broadened by resolution effects. The profiles of unsaturated Bragg peaks are therefore a measure for the resolution (= point spread) function of the experiment. As a consequence the Patterson densities get more and more attenuated and blurred the further away they are from the origin of PDF space (Weber & Simonov, 2012). Since the experimental resolution function affects Bragg and diffuse scattering in the same way, an empirical description of the impact of the resolution function on the Patterson densities allows correcting resolution effects in the 3D- Δ PDF space. In order to do this, the reconstructed hki (where i is equal to integer) layers were cleaned from diffuse scattering and from 20¹ over-saturated Bragg peaks leaving 348 symmetry independent Bragg peaks. Then, the layers were Fourier transformed to obtain the (partial) Patterson functions (see Fig. 4). The fact that only a subset of the Bragg reflections was available does not significantly affect the determination of the resolution function in PDF space, which is measured as the deviation from an ideal periodic Patterson map. Attenuation of the Patterson peaks was obtained by summing up intensities of 5x5 pixels around lattice points of the Patterson grid to compensate blur effects. Fig. 4b shows the attenuation depending on the distance in PDF space. The distribution was approximated

¹ Saturated Bragg reflections: 400, 610, 620, 420, 630, 730, 540, 640, 740, 650, 101, 201, 111, 211, 541, 002, 432, 542, 303, 214

as the product of the Fourier transforms of a Lorentzian and a Gaussian. It accurately fits the tail of the distribution, but shows some deviation at very short distances. The fitted value for the neighbor (0,0,0) may be up to 10% higher than the one used in our correction. The consequence is that the scale is underestimated what may affect the numerical values of the refined correlation coefficients by about the same magnitude. For a discussion about the impact of the scale factor determination on the accuracy of 3D- Δ PDF investigations see (Weber & Simonov, 2012). The effect of blurring on the results of our refinements is slightly smaller. The width of the biggest blurring kernel does not exceed 1.1Å, leading to underestimated correlation coefficients of about 6%, i.e. blurring has an opposite effect on the results compared to the underestimated scale factor. Since our incomplete resolution function correction is expected to be the most significant systematic error in our refinement we estimate that the systematic deviations from the correct results do not exceed 10% of the refined parameter value. The refined correlation coefficients c_{uv} , the attenuation factors A_{uv} and the corrected correlation coefficients $c_{uv}^{corr} = c_{uv}/A_{uv}$ are listed in the Table 2. The standard uncertainties of corrected correlation coefficients, which take into account both, random and systematic errors, were calculated as $\sigma_{uv}^{corr} = \sqrt{\sigma_{uv}^2 + (\sigma_{sys}^r c_{uv})^2}/A_{uv}$, where σ_{uv} are the s.u.'s from the variance-covariance matrix of the least-squares refinement before resolution effects were corrected for and $\sigma_{sys}^r = 0.1$ is the relative systematic error after correction. A geometrical representation of the position of the c_{uv}^{corr} parameters values in PDF space is shown in Fig. 5.

Table 2. Refined correlation coefficients c_{uv} for pairs of molecular stacks separated by a vector $(uv0)$. The numbers in parentheses in the third column give the s.u.'s as obtained from variance-covariance matrix of the least-squares refinement and the fourth column indicates the relative errors. The fifth column reports the attenuation factor A_{uv} due to resolution effects and the last column shows the corrected pair correlation coefficients.

u	v	c_{uv}	s.u./ c_{uv}	A_{uv}	c_{uv}^{corr}
1	0	-0.2471(3)	0.0014	0.9030	-0.27(3)
2	1	0.0954(3)	0.0029	0.8333	0.114(11)
2	0	0.0968(3)	0.0029	0.8081	0.120(12)
3	1	-0.0413(3)	0.0067	0.7486	-0.055(6)
2	-1	-0.0510(3)	0.0054	0.7486	-0.068(7)
3	0	-0.0307(3)	0.0089	0.7167	-0.043(4)
4	2	0.0299(3)	0.0091	0.6758	0.044(4)
4	1	0.0154(3)	0.0177	0.6636	0.023(2)
3	-1	0.0230(3)	0.0120	0.6636	0.035(3)
4	0	0.0140(3)	0.0195	0.6300	0.022(2)
5	2	-0.0134(3)	0.0204	0.6001	-0.022(2)
3	-2	-0.0158(3)	0.0173	0.6001	-0.026(3)
5	1	-0.0064(3)	0.0423	0.5819	-0.0111(12)
4	-1	-0.0108(3)	0.0254	0.5819	-0.0185(19)
5	0	-0.0077(3)	0.0354	0.5487	-0.0140(15)
6	3	0.0072(3)	0.0378	0.5335	0.0135(14)
6	2	0.0063(3)	0.0435	0.5262	0.0119(13)
4	-2	0.0088(3)	0.0310	0.5262	0.0167(18)
6	1	0.0030(3)	0.0913	0.5053	0.0059(8)
5	-1	0.0051(3)	0.0538	0.5053	0.0101(11)
6	0	0.0052(3)	0.0523	0.4737	0.0110(12)
7	3	-0.0045(3)	0.0611	0.4678	-0.0095(11)
4	-3	-0.0054(3)	0.0503	0.4678	-0.0116(13)
7	2	-0.0028(3)	0.0977	0.4563	-0.0061(9)
5	-2	-0.0050(3)	0.0545	0.4563	-0.0110(13)
7	1	-0.0021(3)	0.1307	0.4347	-0.0048(8)
6	-1	-0.0033(3)	0.0818	0.4347	-0.0077(10)
7	0	-0.0034(3)	0.0800	0.4052	-0.0084(11)

The refined values of the size effect parameters, which describe the difference in the contact lengths of homo- and heterochiral columns are shown in Table 3. Only the x and y components were refined since no indication of size effect along z direction was visible in 3D- Δ PDF maps. The size effect parameters may heavily correlate with the lattice parameters and other geometrical factors used for reciprocal space reconstructions. In order to estimate the impact of systematic errors in the determination of the lattice constants on the uncertainty of the size effect parameters, we performed another refinement where the crystal lattice was refined together with the size effect param-

eters. As a result the orientation of the lattice basis vectors rotated by 0.05° and the lattice constant a decreased by 0.022\AA from 14.114\AA to 14.092\AA . At the same time the size effect parameters were reduced to about a half of their initial values (Table 3), i.e. the size effect significantly correlates with systematic errors in the determination of the lattice constants. Since diffuse and Bragg scattering are equally affected by systematic errors in the reconstructions, we determined the lattice constant from the centroids of Bragg peaks in the reconstruction in order to obtain internally consistent measures. The estimated unit cell rotated by 0.0057° and the lattice constant a was equal to 14.119\AA , which is by 0.005\AA larger than the value $14.114(2)\text{\AA}$ determined by XDS. Note, that the standard uncertainty of 0.002\AA determined by XDS does not take into account the errors that are introduced by the reconstruction procedure, making the difference $14.119\text{\AA} - 14.114\text{\AA} = 0.005\text{\AA}$ a better estimate for the extent of systematic geometrical errors.

The size effect parameters, whose s.u.'s were recalculated using the estimated uncertainties of the lattice base vectors, are presented in the last two columns of the Table 3. The observed size effect is very small. The absolute value for the displacement of the neighbor (1,0) is equal to $0.0114(14)\text{\AA}$. Given our simple size effect model, which assumes that the molecular stack displaces as a whole, such a small value should be considered of low significance, because it might compensate e.g. relaxations of the side chains of the molecules. The size effect for larger distance pairs are even smaller and their relative errors are bigger than for the (1,0) neighbor. It should be noted that the R-factor is almost insensitive the size effect and becomes only by 0.3% (14.5% versus 14.2%) worse when the relaxation is removed from the model. However the presence of the relaxation is clearly visible near the lattice points in the so-called delta-delta PDF or 3D- Δ^2 PDF, which is the difference between experimental and model 3D- Δ PDF.

Table 3. Refined size effect parameters for first six neighbors, \mathbf{u}_i is relaxation of “up”-”up” and “down”-“down” pairs, \mathbf{u}^{lat} is the relaxation in the model when the lattice constants are refined along with size effect, the values of \mathbf{u}_i^* are equal to \mathbf{u}_i , but their standard uncertainties take into account the systematic errors induced by the reciprocal space reconstruction.

x	y	$\mathbf{u}_x, \text{Å}$	$\mathbf{u}_y, \text{Å}$	$\mathbf{u}_x^{lat}, \text{Å}$	$\mathbf{u}_y^{lat}, \text{Å}$	$\mathbf{u}_x^*, \text{Å}$	$\mathbf{u}_y^*, \text{Å}$
1	0	0.0053(3)	-0.0078(3)	0.0009(4)	-0.0037(4)	0.0053(13)	-0.0078(5)
2	1	-0.00634(17)	0.00097(17)	-0.0028(2)	0.0011(2)	-0.0063(10)	0.0010(6)
2	0	-0.00240(17)	0.00100(17)	-0.0000(2)	-0.0011(2)	-0.0024(10)	0.0010(4)
3	1	0.0039(2)	0.0011(2)	0.0015(2)	0.0015(2)	0.0039(7)	0.0011(3)
2	-1	-0.0005(2)	-0.0032(2)	-0.0012(2)	-0.0005(2)	-0.0005(6)	-0.0032(4)
3	0	0.00122(19)	-0.00063(19)	-0.0001(2)	0.0004(2)	0.0012(5)	-0.0006(2)

5. Uncertainties of 3D- Δ PDF refinements

The lowest achievable standard uncertainties defined by random errors are very small (about 0.0002Å for size effects and 0.0003 for substitutional correlations). However, due to systematic errors, such small uncertainty cannot be achieved in practice. Some of the major systematic experimental errors are listed in the following. The errors may be better seen and understood in 3D- Δ PDF space, but they are by no means specific for 3D- Δ PDF investigations. The effects may equally bias the results from other quantitative modeling techniques relying on experimental diffuse scattering, including Monte Carlo simulations.

Instrumental resolution function. Due to the fact that the diffuse scattering is convolved with the experimental resolution function, the 3D- Δ PDF signals become attenuated and blurred. As seen in Fig. 4, even in the case of synchrotron experiments the resolution function may have significant impact on the accuracy of refined parameters. The correlation coefficient of the first neighbor ($r \approx 14\text{Å}$) is already attenuated by 10% and the correlation of the longest refined neighbor ($r = 99\text{Å}$) are weakened by 60%.

Estimating and correcting the three-dimensional resolution effects in single crystal experiments is by far more complicated than it is in the case of powder PDF investi-

gations. However, an isotropic approximation to the resolution function either during refinement (as it is now implemented in the program YELL) or after refinement (as it is done in this paper) may easily be applied. Such corrections may be omitted only if the profiles of the diffuse signals are by far broader than the Bragg peaks.

Background coming from air and Compton scattering, sample holder, glue, sample surface or from detector noise typically manifests itself as a smooth intensity which slowly varies in reciprocal space. If not subtracted, it is seen as a narrow peak in the center of 3D- Δ PDF space, which may lead to an overestimated scale factor and underestimated correlation coefficients (Weber & Simonov, 2012). A careful correction of background scattering is therefore essential for obtaining accurate results.

In the present case the background beneath the diffuse scattering could be reliably determined from the background above or below the sharp diffuse layers. Furthermore, presence of medium or big molecules with well-known geometry, as it is the case in the present study, allows compensating some of the systematic errors introduced by non-perfect background correction (Weber & Simonov, 2012).

Geometrical errors in the determination of position and shape of detector, beam and rotation axis directions may lead to slightly distorted reconstructions, which in turn may bias the results of diffuse scattering refinements. In the present case, the size effect with a refined magnitude of about 0.01\AA for the next neighbor was on the limit of detection. A similar observation could have been caused by a systematic error of about 0.02\AA in determination of the crystal lattice.

Random errors. During the experiment, the intensity of diffuse scattering is measured with a certain random error mainly due to counting statistics. However, since the total number of significant data points (frequently $\gg 10^6$) typically exceeds the number of refined parameters ($10^1 - 10^3$) by many orders of magnitude, the total influence of random errors is usually smaller than that of systematic errors.

In the current experiment the average I/σ_r for diffuse scattering was only 1.96 and not more than 9% of diffuse scattering pixels were observed with $I > 2\sigma_r$. However, with approximately 0.9 million symmetry independent pixels, or 22000 measurements per refined parameter, all correlation coefficients could be refined with high statistical precision. For most orientational correlation coefficients the random errors were significantly smaller than the systematic errors estimated as described above. Even for the smallest refined correlation coefficient $c_{71} = -0.0048(8)$ only about half of the uncertainty is coming from random errors. This correlation coefficient means that the corresponding pairs of stacks have similar orientation in 49.76(4)% of all cases, i.e. even very small deviations from a purely random correlation of 50% could be determined with a very small uncertainty.

We find, that for typical diffuse scattering experiments, the statistical errors may be misleading. They tend to be severely smaller than systematic errors and might give a false impression of the reliability of refined parameters. They should only be considered after all systematic errors are carefully accounted for.

6. Results and discussion

6.1. Structural results

The 3D- Δ PDF refinement converged to a very low diffuse scattering R-factor of 14.2%. Both, calculated 3D- Δ PDF (Fig. 3) and diffuse scattering (Fig. 2) are visually indistinguishable from experimental ones, suggesting that the most significant local order features are covered by our model. Neighboring molecular stacks have a strong tendency for anti-parallel orientations with a correlation coefficient of -0.27(3), which is close to theoretical minimum of -1/3 achievable on a hexagonal lattice. The correlation coefficients within the same shells are very similar (see Fig. 5). Slightly stronger correlations along $\langle 210 \rangle$ directions compared to $\langle 100 \rangle$ indicate that the structure has

a small preference to locally arrange in a lamellar pattern where stacks of similar orientations form lines along $\langle 100 \rangle$ type directions. The example of such local arrangement is highlighted in grey in the Fig. 6. The same arrangement is reported for the related ordered compound 1,3,5-tris(2,2-dimethylpropionylamino)benzene (Schmidt & Wittmann, 2012).

Neighboring stacks of different or same chirality show surprisingly small relaxations of $\approx 0.01\text{\AA}$. However, our model assumes that the molecules are relaxing as a whole and we cannot completely rule out the possibility of intramolecular distortions, which might compensate local differences in homo- and heterochiral contacts. Significant displacive correlations (“ADP correlations”) independent from size effect distortions could not be identified. Corresponding diffuse scattering is typically much weaker than diffuse scattering from substitutional correlations as investigated in this paper and may have easily been overseen in the cut-out regions between the strong diffuse honeycomb layers. If it would appear as thermal diffuse scattering beneath the Bragg peaks it may have been eliminated together with the Bragg peaks.

6.2. Methodological conclusions

It could be shown that 3D- Δ PDF refinements are capable of providing quantitative information up to a distance of at least 100\AA . The method is very sensitive to even weak substitutional correlations: we could reliably quantify preferred homo- or heterochiral correlations even if deviations from a purely random distribution was less than 1%. The uncertainty of the refined size effect amplitudes is in the order of 0.01\AA .

For obtaining accurate short range order parameters special attention has to be drawn to systematic errors. The most significant errors are caused by the experimental background and by insufficient resolution function corrections. In many cases diffuse scattering is present as sharp streaks or layers, which allow reliable background sub-

traction by interpolating the background close to the diffuse signals. In the case of broad diffuse scattering special care should be taken to measure the background experimentally, e.g. by recording some frames under the same experimental conditions as during data collection, but without the crystal. Reciprocal space resolution is currently the most challenging correction and to the best of our knowledge has so far never been applied accurately to single crystal diffuse scattering experiments. A simple approximation of a uniform resolution function, however, can be applied efficiently either during the refinement as it is now possible in the program YELL (Simonov *et al.*, 2014) or after the refinement following the procedure described in the current paper.

Estimating reliable standard uncertainties of refined parameters remains an open question. In conventional Bragg crystallography they are estimated on the basis of random errors of measured intensities. With the currently available experimental techniques, data reduction programs and modeling possibilities the contribution of random errors seem to be a minor bottle-neck for obtaining accurate results compared to systematic errors and therefore standard uncertainties determined from the least-squares variance-covariance matrix tend to be significantly underestimated.

7. Acknowledgements

This work was supported by the Swiss National Science Foundation grants 200021_121759 and 200020_140389. The authors would like to thank Prof. H.W. Schmidt (University of Bayreuth) for providing the sample, P. Pattison for assistance with the data collection, the Swiss-Norwegian Beamline for providing access to synchrotron radiation and Ch. Baerlocher, ETH Zurich, for critical comments.

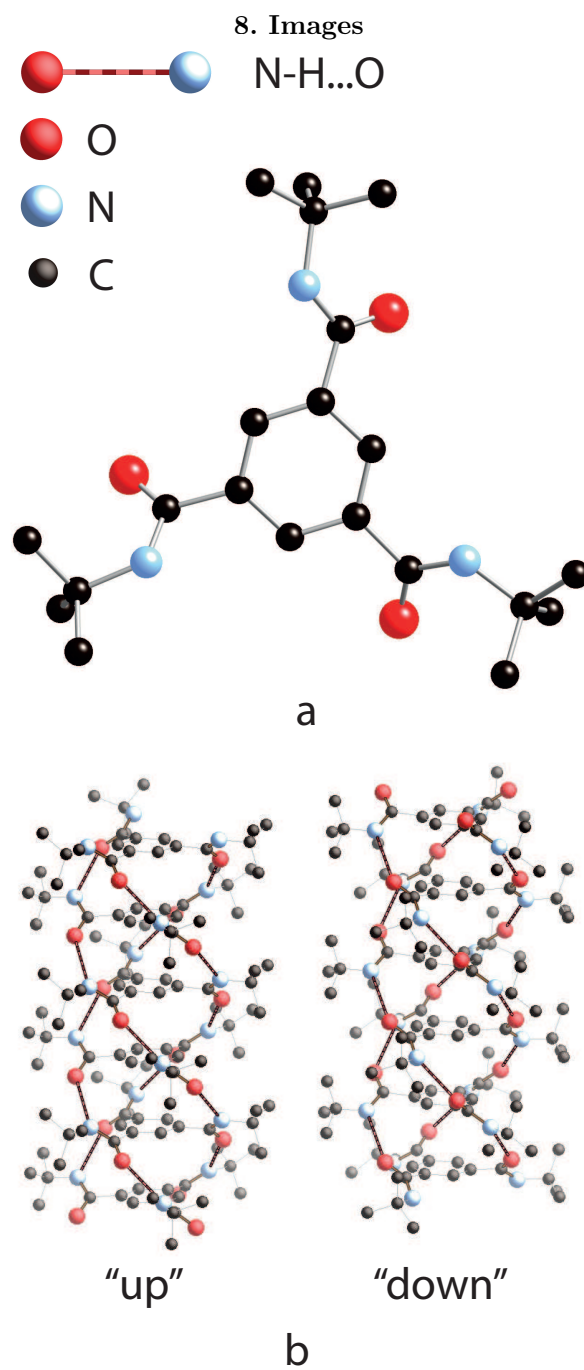


Fig. 1. a) The tris-t-butyl-1,3,5-benzene tricarboxamide molecule. b) Molecular stacks in the two different orientations. The helical system of hydrogen bonds is highlighted using black and red-black lines for covalent and hydrogen bonds respectively.

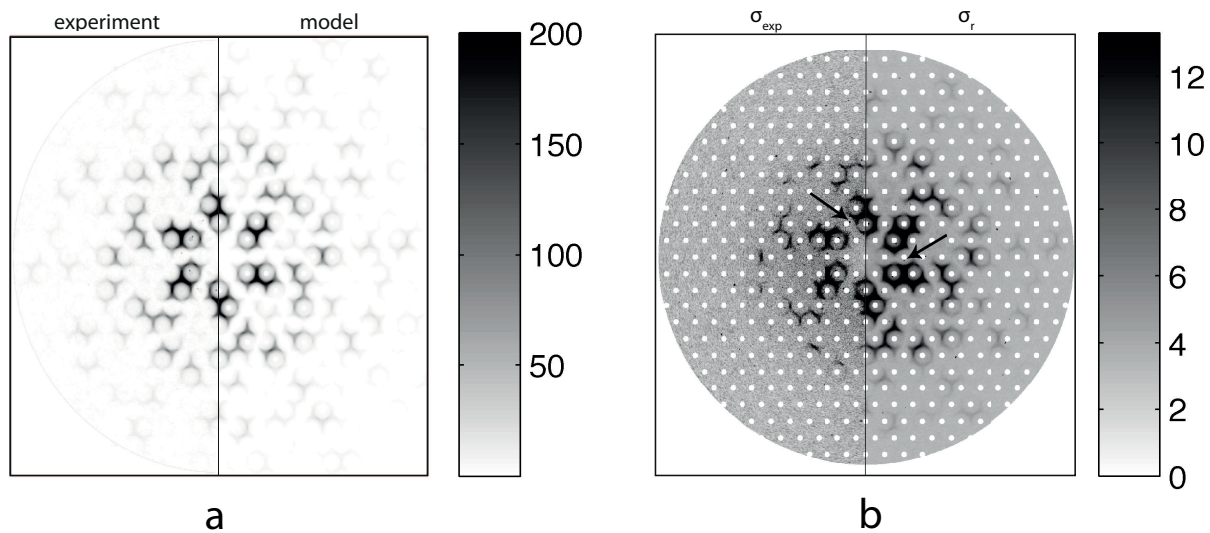


Fig. 2. a) Comparison of experimental and model diffuse scattering in the $hk1$ layer. Background and Bragg peaks are removed from the experimental data. b) Left: standard uncertainties of measured pixels estimated from symmetry equivalents σ_{exp} . Right: standard uncertainties σ_r after regression as described in the text. The arrows mark the outliers which are coming from from incomplete Bragg peak subtraction.

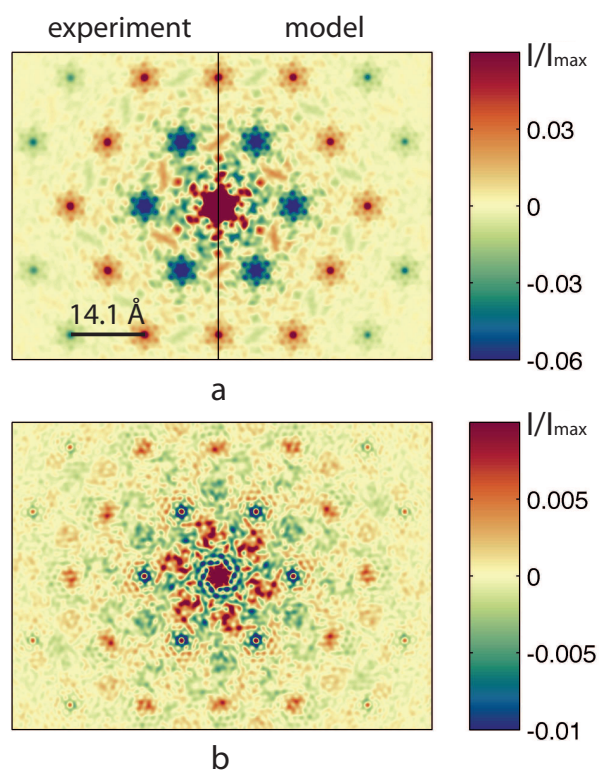


Fig. 3. a) Sections from observed and calculated 3D- ΔPDF maps at $z = 0$. b) Difference between observed and calculated 3D- ΔPDF maps (3D- $\Delta^2 PDF$) at $z = 0$. Note the different scales of the color bars. A bigger section of the same maps can be found in supplementary materials.

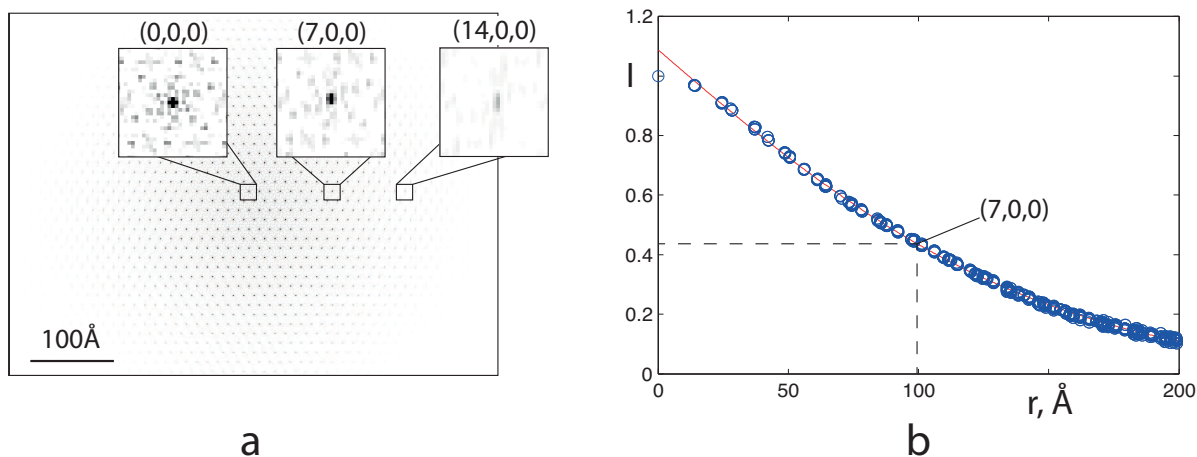


Fig. 4. a) Fourier transform of the reciprocal space reconstruction using only unsaturated Bragg peaks. Because of the reciprocal space resolution function the (partial) Patterson function is not periodic, but attenuated and blurred at long distances. b) Attenuation of Patterson densities due to resolution function effects. Blue dots correspond to integrated Patterson intensity around lattice points in $xy0$ cross-section of Patterson function. The red line corresponds to the fitted resolution function used for correcting the data. All intermolecular vectors up to $(7,0,0)$ at $\approx 100\text{\AA}$ were refined.

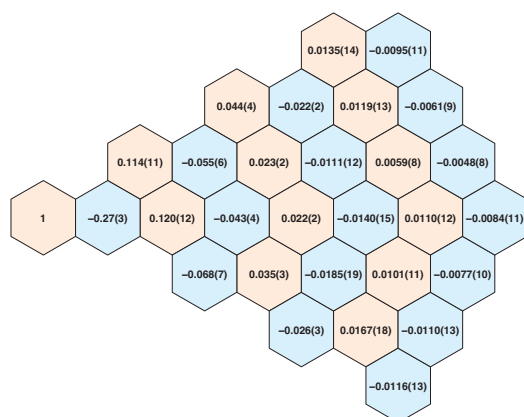


Fig. 5. Correlation coefficients c_{uv}^{corr} arranged according to their position in the 3D- Δ PDF map. Note the alternation of positive and negative correlation coefficients in subsequent shells. Only the asymmetric cone is presented.

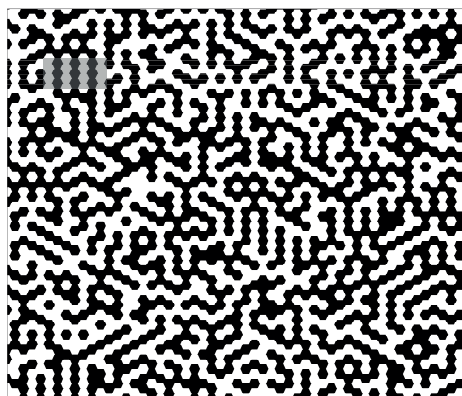


Fig. 6. A possible realization of the tris-butyl-tricarboxamide real structure that is consistent with the PDF model from this study. Black and white hexagons represent stacks oriented "up" and "down" respectively. Note, that the local structure tends to have lamellar arrangements. An example is highlighted in grey. The image was produced using a modified Reverse Monte Carlo (Nield *et al.*, 1995) procedure. Instead of diffuse scattering, the RMC simulation was done against the refined 3D- Δ PDF pair correlation coefficients.

References

- Adams, C., Lynn, J., Mukovskii, Y., Arsenov, A. & Shulyatev, D. (2000). *Physical review letters*, **85**(18), 3954.
- Albuquerque, R. Q., Timme, A., Kress, R., Senker, J. & Schmidt, H.-W. (2013). *Chemistry (Weinheim an der Bergstrasse, Germany)*, **19**(5), 1647–57.
URL: <http://www.ncbi.nlm.nih.gov/pubmed/23239528>
- Billinge, S. J. (2008). *Journal of Solid State Chemistry*, **181**(7), 1695–1700.
URL: <http://linkinghub.elsevier.com/retrieve/pii/S0022459608003460>
- Blomenhofer, M., Ganzleben, S., Stoll, K., Maeder, D. & Hoffmann, K. (2005). *Macromolecules*, **38**, 3688–3695.
URL: <http://pubs.acs.org/doi/abs/10.1021/ma0473317>
- Butler, B. & Welberry, T. (1992). *Journal of applied crystallography*, **25**(3), 391–399.
- Estermann, M. a. & Steurer, W. (1998). *Phase Transitions*, **67**(1), 165–195.
URL: <http://www.tandfonline.com/doi/abs/10.1080/01411599808219193>
- Hanabusa, K., Koto, C., Kimura, M., Shirai, H. & Kakehi, A. (1997). *Chemistry Letters*, **26**(5), 429–430.
URL: <http://japanlinkcenter.org/JST.JSTAGE/cl/1997.429?from=Google>
- Kabsch, W. (2010). *Acta crystallographica. Section D, Biological crystallography*, **66**, 125–32.
URL: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2815665&tool=pmcentrez&rendertype=abstract>
- Kobas, M., Weber, T. & Steurer, W. (2005). *Physical Review B*, **71**(22), 224205.
URL: <http://link.aps.org/doi/10.1103/PhysRevB.71.224205>
- Kristiansen, M., Smith, P., Chanzy, H., Baerlocher, C., Gramlich, V., McCusker, L., Weber, T., Pattison, P., Blomenhofer, M. & Schmidt, H.-W. (2009). *Crystal Growth and Design*, **9**(6), 2556–2558.
- Matlab, (2011b). The MathWorks, Inc., Natick, Massachusetts, United States.

- Matsunaga, Y., Nakayasu, Y., Sakai, S. & Yonenaga, M. (1986). *Molecular Crystals and Liquid Crystals*, **141**(3-4), 327–333.
URL: <http://www.tandfonline.com/doi/abs/10.1080/00268948608079619>
- Nield, V., Keen, D. & McGreevy, R. (1995). *Acta Crystallographica Section A: Foundations of Crystallography*, **51**(5), 763–771.
- Schaub, P., Weber, T. & Steurer, W. (2007). *Philosophical Magazine*, **87**(18-21), 2781–2787.
URL: <http://www.tandfonline.com/doi/abs/10.1080/14786430701367971>
- Schmidt, M. & Wittmann, J. (2012). *Crystal Growth*, **12**, 2543–2551.
URL: <http://pubs.acs.org/doi/abs/10.1021/cg300151s>
- Simoncic, P. & Armbruster, T. (2004). *American Mineralogist*, **89**(2-3), 421–431.
- Simonov, A., Weber, T. & Steurer, W. (2014). *Journal of Applied Crystallography*, **47**(3), 1146–1152.
- Weber, T. & Simonov, A. (2012). *Zeitschrift für Kristallographie*, **227**(5), 238–247.
- Welberry, T., Heerdegen, A., Goldstone, D. & Taylor, I. A. (2011). *Acta Crystallographica Section B: Structural Science*, **67**(6), 516–524.

4.4 PbTe

PbTe received much attention as a thermoelectric compound at temperatures from 400K to 800K. It has a very low thermal conductivity, which is usually not observed in simple NaCl-type structures. The mechanism of such peculiar behavior is a subject of debate. Using powder PDF analysis Božin et al. [38] observed an anomalous temperature dependent broadening of the nearest neighbor Pb-Te distance distribution. Based on this observation, they proposed that at 300K the Pb atoms move by approximately 0.2\AA from high symmetry position (4a), lowering the local symmetry of the crystal. Such displacements form electrostatic dipoles, which arrange locally in polar nano domains.

On the other hand, recent studies using extended x-ray absorption fine structure spectroscopy (EXAFS) showed that lead atoms occupy the high symmetry position without any observable off-centering [39]. The authors suggest that the low temperature conductivity is caused by an unusual structure dynamics, which becomes strongly anharmonic already at 300K. This conclusions were also supported by inelastic neutron scattering measurements [40] and ab-initio calculations [41].

The objective of the current study was to check whether 3D- Δ PDF analysis will support the model of static off-centering or the model of anharmonic displacements. The final conclusions will be published in a separate paper [42], here the part of study which reflects the process of 3D- Δ PDF analysis is presented.

Experiment and data preparation

An irregularly shaped fragment with an average diameter of about $42\mu\text{m}$ was used for the x-ray experiments. The measurement was performed at the X06SA beam-line at SLS, Villigen, Switzerland, equipped with a PILATUS 6M detector. The detector energy threshold for accepting x-rays photons was set to 16 keV (energy of the primary beam: 17.5 keV) to suppress fluorescence scattering as much as possible. The datasets were recorded at room temperature, 250K, 200K, 150K and 125K, however only the room temperature results are described in the current chapter. For more information see [42]. In addition 100 frames were collected under the same conditions as in the diffuse scattering measurements, but without sample and sample holder. The frames were averaged and taken as a model for background scattering.

The orientation matrix was determined using the program XDS [31]. The reconstructions were performed with Xcavate [32]. Datasets were corrected for polarization and air absorption effects. The datasets were averaged in $m\bar{3}m$ symmetry using the outlier rejection following the modified procedure described by Blessing [35]. A pixel was considered an outlier if its intensity I_i was different from median intensity of its symmetry equivalents by more than $|I_i - I_{median}| > 6t$, where $t = \text{median}(|I_i - I_{median}|) \sqrt{[n/(n-1)]}$ and n is the number of averaged pixels.

The background determined by empty frame measurements was subtracted from the diffuse data. It turned out that the background was slightly larger when the sample was out of the beam. The reason for this is that the sample absorbs the primary beam leading to reduced air scattering after the beam has passed the sample. The scale factor for absorption was manually adjusted to 0.9 in order to avoid extended negative intensity regions. In the final step, the datasets were corrected for sample absorption using the spherical absorption correction approximation.

Reconstructed Bragg peaks were only about 1 pixel sharp, but in order to make sure that any traces of Bragg peak tails were eliminated, the volumes of $3 \times 3 \times 3$ pixels around Bragg positions were

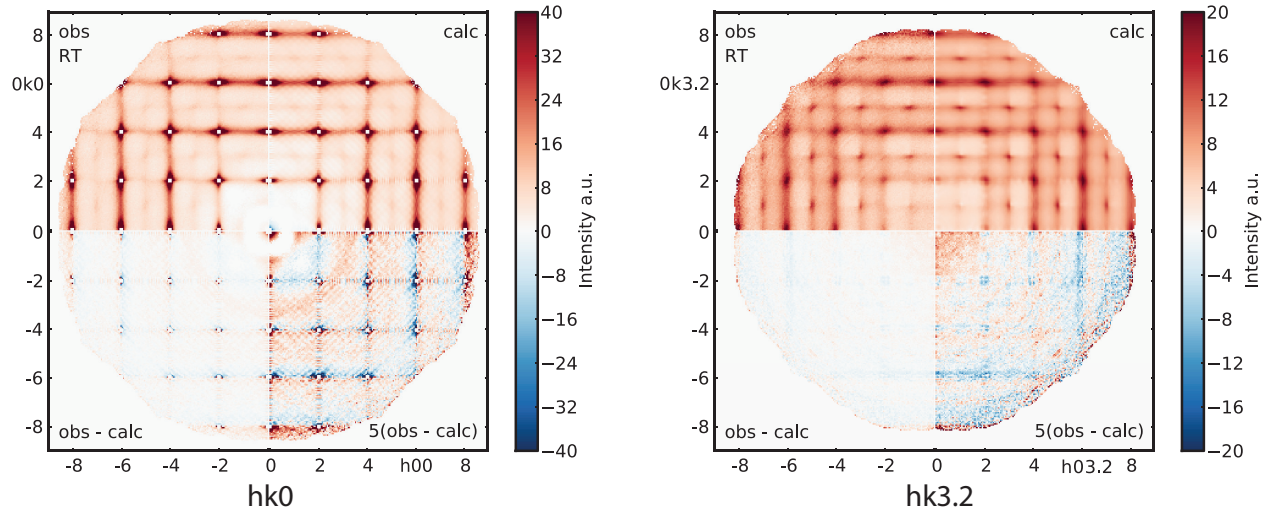


Figure 4.4.1: Observed, calculated and difference intensities of diffuse scattering of PbTe at room temperature at selected sections.

removed. The regions were not filled using the punch-and-fill method, but were kept zero and were excluded from refinement. The final dataset contained $360 \times 360 \times 360$ pixels.

Average structure

As already noted, PbTe has a NaCl-type structure. The atomic displacement parameters of Pb ($\sim 0.0260(2)\text{\AA}^2$ at 293K) are very high for such a heavy atom. This behavior might be interpreted as a sign of static disorder, but a careful examination of the average structure revealed that in the current case Pb atoms more likely occupy the high symmetry position than split position along $\langle x00 \rangle$, $\langle xx0 \rangle$ or $\langle xxx \rangle$ directions. For further details about the average structure description see [42].

Diffuse scattering and the 3D- Δ PDF

Diffuse scattering is present as wide three-dimensional bands oriented perpendicular to the crystal axes (Fig. 4.4.1). The symmetry of diffuse scattering is $m\bar{3}m$, i.e. the same as for the average structure. At different temperatures the diffuse scattering shows essentially the same appearance, only the overall amount of diffuse scattering decreases relative to the integral Bragg intensities when lowering the temperature.

The 3D- Δ PDF shows the typical signature of atomic displacement correlations (Fig 4.4.2). A signal with characteristic negative-positive-negative profile is found at each interatomic vector position [43]. This means that displacements of both heterogeneous Pb-Te as well as homogeneous Pb-Pb or Te-Te pairs are correlated.

In the final 3D- Δ PDF model all significant displacement correlations were refined. The pairs Te-Te and Pb-Pb perfectly overlap and could not be refined separately. The displacement correlation

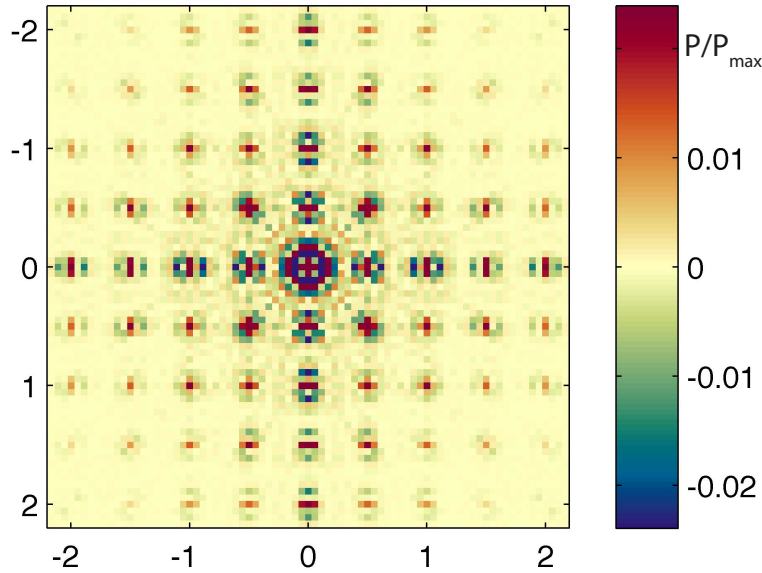


Figure 4.4.2: 3D- Δ PDF of PbTe, xy0 cross-section.

coefficients for such pairs were constraint to be equal. This is a justified assumption because the 3D- Δ PDF signals are dominated by Pb-Pb pairs since Pb has a stronger scattering power and also larger atomic displacement parameters.

Totally 716 pairs in the region $x \leq 10$, $y, z \leq 3$, in the asymmetric cone $0 \leq z \leq y \leq x$ were refined. If treated independently, the correlations would comprise ~ 3500 parameters. Full matrix Levenberg-Marquardt algorithm implemented in YELL requires all Jacobians to be simultaneously stored in the operating memory. Thus, the memory requirement scales as the number of pixels times the number of parameters. In the current case approx. 1000GB would be required to perform a full refinement—a quantity far beyond our computational capacities. It was found, that homo- and hetero-atomic correlations decay exponentially along $\langle 100 \rangle$ directions. This assumption failed only for very short interatomic vectors. The final model was as follows. The correlation matrices $Corr_{uvw}^{mn}$ were refined independently for interatomic vectors $(0.5\ 0\ 0)$, $(1\ 0\ 0)$, $(1.5\ 0\ 0)$, $(2\ 0\ 0)$, $(2.5\ 0\ 0)$ and $(3\ 0\ 0)$. The correlation coefficients of other pairs were set according to the relation $Corr_{uvw}^{mn} = a_{vw}^{mn} \odot \exp(-b_{vw}^{mn} u)$ where $Corr_{uvw}^{mn}$ is the correlation tensor a_{vw}^{mn} and b_{vw}^{mn} are symmetric 3×3 matrices, the symbol \odot represents element-wise multiplication, (u, v, w) are the average interatomic vectors, the indices mn represent either Pb-Pb or Pb-Te pairs. The final model contained 363 independent parameters. Despite the number of parameters was reduced significantly, memory limitations still did not allow to refine all parameters simultaneously. The least-squares refinements were therefore done in blocks of about 30 parameters and repeated until no further improvements could be observed. Single least-squares runs took about one to two hours on a modern desktop computer. The full refinement took about two days.

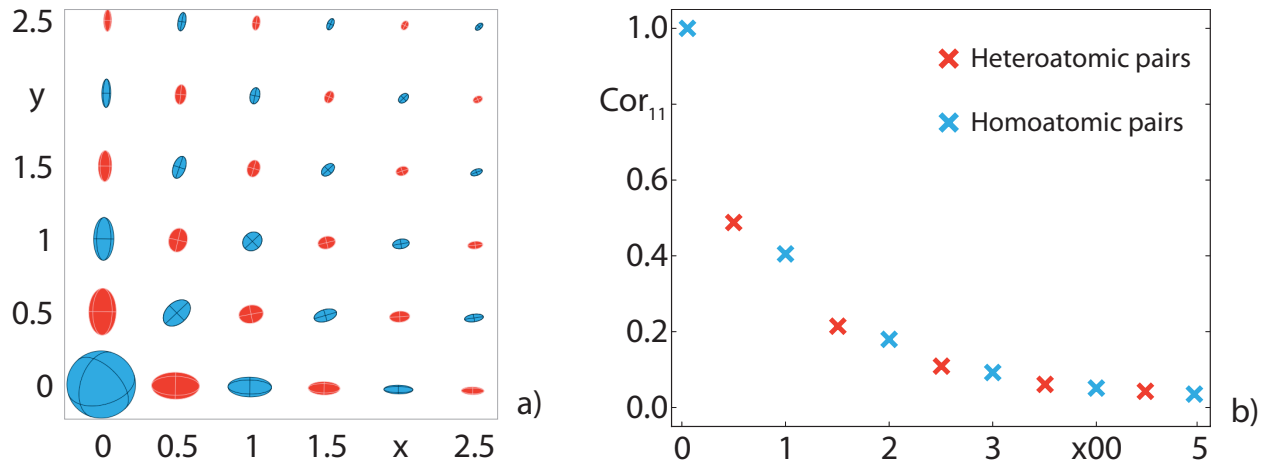


Figure 4.4.3: a) Refined atomic displacement correlation coefficients represented as ellipsoids. The correlation for the zero neighbor is equal to 1 by definition, and decrease for further neighbors. Note that the joint atomic displacement parameters have the inverse behavior - they are equal to zero for the zero neighbor and slowly increase for the further neighbors. b) Correlations Cor_{11} for first ten neighbors along $x00$ direction.

Results

Refinement against the room temperature dataset converged with $R=0.14$. The residual 3D- Δ^2 PDF maps contained obvious indication for anharmonicity of all the pairs along $\langle 100 \rangle$ direction (Fig 4.4.4). This is in a very good agreement with the results of the EXAFS study [39].

The most significant pair correlations are shown as ellipsoids in Fig 4.4.3a. In agreement with the qualitative interpretation it is seen that the correlations are strongest along $\langle 100 \rangle$ directions with dominating longitudinal correlations. Interestingly, the homo-atomic pairs generally show stronger correlations than hetero-atomic pairs (Fig 4.4.3b). The structural interpretation of the correlation coefficients is beyond the scope of the current chapter and will be published elsewhere.

Discussion

The current work supports the assumption that the average structure of PbTe is best described by the Pb atom occupying the high symmetry position. Diffuse x-ray scattering can be well modeled by assuming only the correlated atomic displacements. This contradicts the static displacement model of Božin et al. [38] and supports a model expressed in [39, 40, 41] where diffuse scattering has a purely dynamic origin. The most plausible reason for the low thermal conductivity of PbTe is thus the anharmonic dynamics of the crystal.

It has been shown, that 3D- Δ PDF method can be used to probe the crystal dynamics. It provides very detailed and reliable information about strength and direction of atomic displacement correlations. Though it does not provide phonon dispersions in reciprocal k -space, 3D- Δ PDF allows to investigate the dynamic phenomena in real space, which might prove useful for complicated examples with strongly anharmonic or localized vibrations.

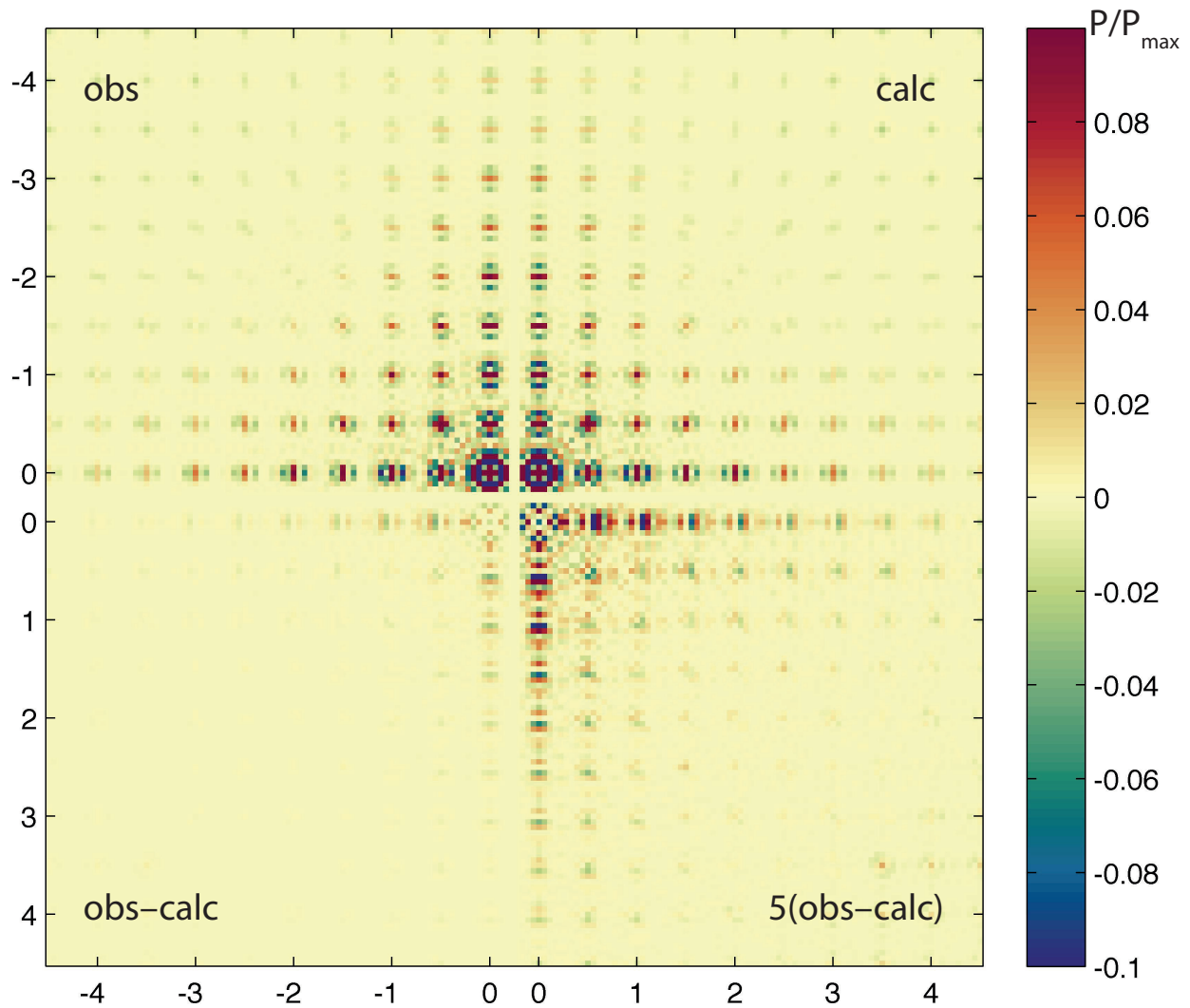


Figure 4.4.4: Observed, calculated and difference 3D- Δ PDF densities of the xy_0 layer at room temperature. The relatively strong difference densities at $x, y > 3$ in the enhanced difference density quadrants represent densities outside the volume covered by the model. The negative densities at large distances are coming from artifacts due to removing the Bragg peaks.

Chapter 5

Conclusions and outlook

5.1 Comparison of the 3D- Δ PDF approach with other diffuse scattering methods

Direct modeling

YELL refinements, similarly to average structure refinements, may result in non-physical correlation models, which are not compatible with any real structure model. For example, in unfavorable cases, joint probabilities can be refined to negative values or joint atomic displacement parameters can converge to non-positively definite matrices. There is also a possibility of more subtle errors, which are significantly harder to spot. For example, short range order parameters, even though meaningful individually, could be impossible to be realized collectively. The models obtained from MC simulations can be trivially examined for presence of any unphysical behavior and rarely contain such problems.

On the other hand, MC refinements have two serious issues. First, MC models are hard to construct. A user is required to guess a model that is flexible enough to describe the disorder at hand. If a model covers real structure properties only partially, it is often very hard to identify incomplete or wrong parts of the model. Comparison of the model and experimental diffuse scattering may in principle provide such information, but an expert knowledge is required to extract it. Thus, in practice, identification of a proper real structure model is usually depending on inspiration and many cycles of trials-and-errors. Second, MC refinements are computationally very slow. It is not uncommon that refinements take weeks to converge [16].

3D- Δ PDF refinements do not require a full representation of a crystal. Only the average structure and a list of correlations are needed for describing local order. Diffuse scattering calculations from a 3D- Δ PDF model is considerably faster than from MC models, which allows very quick tests of various models. 3D- Δ PDF maps are very easy to interpret, since correlation from each interatomic pair produces a predictable pattern in a well localized part of the map. The model can naturally be developed step by step starting from the strongest correlations. Additional correlations are added gradually only after their effect becomes clearly visible in the so-called delta-delta pair distribution function (3D- Δ^2 PDF), the difference between experimental and calculated 3D- Δ PDF.

Overall, the 3D- Δ PDF can be considered as a complementary approach to MC modeling. In the

cases where MC simulations are desired, it can be used prior to MC refinements. It allows a quick understanding of the origin of diffuse scattering and provides important information about the types of correlations needed for successful MC simulations.

Normal mode analysis

The PbTe example demonstrates that the 3D- Δ PDF method is capable of refining diffuse scattering of a purely dynamic origin. As a result, it provides a list of atomic displacement correlation parameters. Given the phonon dispersion curves of a crystal are available, the ADP correlation coefficients can be calculated using eq. (25) in [20]. Thus, the model can be tested not only against diffuse scattering, but also against the results of 3D- Δ PDF refinements.

Analytical methods

The 3D- Δ PDF can be considered as an extension of Warren-Cowley formalism. Our approach works not only for single atoms, but also for molecules. In addition to substitutional correlations and size effects, our approach adds atomic displacement correlation as an additional short range order parameter. Diffuse scattering calculated by YELL is very similar to diffuse scattering calculated from Warren-Cowley analytical formulas except the latter uses the first term of Taylor series with respect to size effect parameters. For large values, the results of the analytical linear approximation might be inconsistent with direct calculation. In such cases the program YELL provides more accurate results. Since currently there is no specialized program which can perform Warren-Cowley analysis, the program YELL may be used for such purposes.

Powder pair distribution function

The 3D-PDF is a natural extension of the radial distribution function which is used in powder 1D-PDF refinements. However, the availability of three dimensional diffuse scattering allows several important improvements. First, in three dimensional space it is much easier to separate Bragg and diffuse scattering. Bragg scattering can be analyzed with traditional methods to yield accurate average structures, while diffuse scattering is used separately to obtain the short range order correlations using the average structure as a constraint. Second, signals from pairs with interatomic vectors having similar length but different orientations do not overlap in three dimensions. As a consequence, the number of overlapping peaks does not increase with increasing interatomic vectors. In fact, in the 3D- Δ PDF approach the opposite is true. Since correlations decay with increasing interatomic vectors, the total amount of 3D- Δ PDF signals actually *decreases* at high $|\mathbf{r}|$ (see e.g. Fig. 4.4.4). It allows refinement of correlations to very large distances only limited by the experimental resolution function. In our experience interatomic vectors of up to 100Å may be refined easily and reliably if high quality synchrotron data are available.

From experimental perspective, collecting diffuse scattering from powders is easier and less time consuming. The technique is well suited for non-ambient conditions, phase transitions and multi-phase samples. In addition, 1D-PDF is applicable not only to crystalline materials, but also to nano-materials and liquids.

5.2 Summary and outlook

In this thesis the three dimensional difference pair distribution function (3D- Δ PDF) method is introduced. The method allows investigating diffuse scattering from single crystals and gives access to information about the real crystal structure. The program YELL for performing 3D- Δ PDF refinements was implemented and released. Owing to efficient least square minimization algorithm and the fast FFT-based method for diffuse scattering calculation, large datasets can be refined in a matter of minutes or hours.

The approach was successfully applied to four disordered crystal structures. The crystals covered a broad range of systems: organic, inorganic and metallic materials. They showed an almost complete set of diffuse scattering types, including one-, two- and three-dimensional diffuse scattering; static and dynamic diffuse scattering; diffuse scattering mostly governed by substitutional correlation, size effect and atomic displacement correlations. Despite the wide variety of systems, all of them were described using the same simple and coherent formalism.

The tris-*t*-butyl-1,3,5-benzene tricarboxamide example demonstrates that 3D- Δ PDF refinements provide reliable numerical values of short range order parameters with small estimated errors. The accuracy of the refined parameters is mostly determined by the systematic errors introduced during data reduction and can be further improved by careful experimental design and data preparation if needed.

In order to make future diffuse scattering analysis simpler, several tools have to be implemented. The most important issue is reciprocal space reconstruction. Software suites for Bragg peak integration are typically designed to visually inspect the quality of the experiment and are inadequate for quantitative diffuse scattering analysis. The program Xcavate [31], which is frequently used in diffuse scattering analysis, provides high quality reconstructions, but works only for restricted experimental geometries and in orthogonal reciprocal space coordinates. Development of a reconstruction program for user-friendly high quality reciprocal space reconstructions is essential for a broad adaptation of diffuse scattering analysis.

Another area for which more advanced tools are desirable is data reduction, which includes Bragg and background subtraction, background correction and symmetry averaging. Diffuse scattering from different crystals varies dramatically and so do methods for data reduction. It is unlikely that a general tool that would fit to all demands can be created. However, several simple procedures for outlier rejections, integration of diffuse streaks and planes, background estimation from crystals with 1D and 2D disorder and from experimentally measured frames are relatively easy to implement and can help providing datasets of a sufficient quality.

An unexpected finding in the current work is the importance of a proper determination of the scale coefficients. When refined from diffuse scattering alone, the scale coefficient is prone to errors due to incorrectly subtracted background. Errors in the scale coefficient almost perfectly correlate with a monotonic increase or decrease of the values of all short range order parameters and are very hard to identify. This problem is not specific to 3D- Δ PDF refinements, but affects any method which utilizes single crystal diffuse scattering. A simple way to overcome this problem is to use the scale coefficient from the average structure refinement. In order to do this straightforwardly it is crucial, that the programs for Bragg peak integration and reciprocal space reconstruction will be implemented in a consistent way and will deliver common scale coefficient for diffuse and Bragg intensity. This will allow to translate the reliable scale coefficients from the conventional

crystallographic programs like Shelx [44] or Jana [45] to diffuse scattering model and fix it without refinement, thus removing the influence of systematic errors.

The program YELL can also be improved in a number of ways. The speed of refinement can be accelerated for an important special case of purely substitutional molecular disorder by pre-calculating molecular form factors. The memory requirements can be decreased by utilizing the Laue symmetry and also by taking less memory demanding minimization algorithms for example the conjugate gradient descent [46]. The support for average structure (space group) and real structure (layer, rod and point groups) symmetry may be improved. This implies automatic calculation of multiplicity, selection of the symmetry independent part of the correlations and restriction of short range order parameters. Also a tool for visualization of diffuse scattering, the 3D- Δ PDF, and the refined parameters is currently missing.

Overall, the 3D- Δ PDF method has been developed to a stage at which it can be successfully used by non-experts in the field of diffuse scattering. The method is fast and intuitive and can be applied to a wide range of disorder problems.

Appendix A

Script for generating YELL input file for $\text{Ge}_4\text{Bi}_2\text{Te}_7$

The following script was used to calculate marginal average probabilities of various layers and generate `SubstitutionalCorrelations` for the first 35 neighbors in the $\text{Ge}_4\text{Bi}_2\text{Te}_7$ covered in the chapter 4.2. The script is written in Matlab and relies on symbolic computation toolbox.

```
%% Set up the stacking model.
%Define different stacking blocks.
%The one letter code is used to describe each stack, the correspondence
%of this letters with notation in the article is as follows:
%V - V
%u - Te^1
%U - GB^1
%p - Te^2
%P - GB^2
%N - GB
%n - Te
%O - GB_2
%o - Te_2
%D - GB_1
%d - Te_1
stacks={'VuUnDd', 'VuUpNoDd', 'VuUpPnOoDd', 'VuUpPnNnOoDd', 'VuUpPnNnNnOoDd', 'VuUpPnNnNnNnOoDd', '
VuUpPnNnNnNnNnOoDd', 'VuUpPnNnNnNnNnNnOoDd', 'VuUpPnNnNnNnNnNnNnOoDd', 'VuUpPnNnNnNnNnNnNnNnOoDd
', 'VuUpPnNnNnNnNnNnNnNnNnOoDd', 'VuUpPnNnNnNnNnNnNnNnNnNnOoDd', 'VuUpPnNnNnNnNnNnNnNnNnNnNnOoDd
'};
%setup probabilities
syms p2 p3 p4 p5 p6 p7 p8 p9 p10 p11 p12 p13 p14 real
%Enforce the sum of probabilities to be equal to 1
p15 = solve('p3+p4+p5+p6+p7+p8+p9+p10+p11+p12+p13+p14+p15=1', 'p15');
%Define probabilities of stacks
probabilities = [p3 p4 p5 p6 p7 p8 p9 p10 p11 p12 p13 p14 p15];
%% calculate marginal probabilities
possible_layers = 'UPDOVNupdon';
```

```

l2ind=@(l)(find(possible_layers==l,1));
marginal_probabilities=sym(zeros(1,length(possible_layers)));
for i=1:length(stacks)
    stack=stacks{i};
    probability=probabilities(i);
    for ii=1:length(stack)
        marginal_probabilities(l2ind(stack(ii))) = probability+marginal_probabilities(l2ind(stack(
            ii)));
    end
end
marginal_probabilities=marginal_probabilities/sum(marginal_probabilities)*2;
fprintf('Please paste the following probabilities in your model:\n')
marginal_probabilities=sym_matrix_to_string(marginal_probabilities);
for i=1:length(marginal_probabilities)
    fprintf('p(%s)=%s\n',possible_layers(i),marginal_probabilities{i})
end
%% Calculate layers after void
%Parameter which defines how far to calculate the model
last_neighbor_to_calculate = 35;
after_void_probabilities=sym([]);
valid_till=0;
for i=1:length(stacks)
    sizes(i)=length(stacks{i});
end
possible_layers = 'UPDOVNupdon';
l2ind=@(l)(find(possible_layers==l,1));
% Calculate distribution of different layers after void.
while(valid_till<last_neighbor_to_calculate)
    valid_till = size(after_void_probabilities,1) + min(sizes);
    res = sym(zeros(valid_till,length(possible_layers)));
    for i=1:length(stacks)
        stack=stacks{i};
        p=probabilities(i);
        for j=1:min(length(stack),valid_till)
            res(j,l2ind(stack(j)))=p+res(j,l2ind(stack(j)));
        end
        if(j<size(res,1))
            res(j+1:end,:)=res(j+1:end,:)+p*after_void_probabilities(1:size(res,1)-j,:);
        end
    end
    after_void_probabilities=res;
    valid_till
end
%% calculate and save definition of correlations to file correlations.txt
fid = fopen('correlations.txt','w');
resulting_correlation_matrix={};
%fid=1; %output on screed for debug
for neig=0:35;
    fprintf('calculating neighbor number %i\n',neig);

```

```

%this block calculates variable pair_probabilities which contains
%symbolic equations of pair probabilities

pair_probabilities=sym(zeros(length(possible_layers)));

for i=1:length(stacks)
    stack=stacks{i};
    len=length(stack);

    probability=probabilities(i);

    probs=sym(zeros(len,length(possible_layers)));
    for ii=1:len
        probs(ii,l2ind(stack(ii)))=1;
    end
    probs=cat(1,probs,after_void_probabilities);
    for ii=1:len
        pair_probabilities=pair_probabilities+probability*probs(ii,:)*probs(ii+neig,:);
    end
end

% Setting a couple of variables for displacements
disps={'0,0','1/3,2/3','-1/3,1/3'};
disps=disps([1 3 2]);
disp = disps{mod(neig,3)+1};
gap = '          ';

% Depending on which neighbor it is, even or uneven from the
% pair_probabilities extract the variable pp or two variables -
% pp_tete and pp. Variable pp is equal to joint probabilities between main layers
% in even case and main-to te layer in uneven case pp_tete is the pair
% between Te layers, only defined for even neighbors
if(mod(neig,2)==0)
    pp_tete = pair_probabilities(7:11,7:11)';
    string_pp_tete=sym_matrix_to_string(pp_tete/sum(pp_tete(:)));
    pp=pair_probabilities(1:6,1:6)';
else
    pp=pair_probabilities(1:6,7:11)';
end

% Normalize pp to sum up to one, simplify so that expression is as
% short as possible and convert from symbolic representation to a cell
% array containing strings
pp=pp/sum(sum(pp));
pp=simplify(pp);
string_pp=sym_matrix_to_string(pp);

%Generate output

```

```

if(mod(neig,2)==0)
  if neig==0
    multiplicity=0.5;
  else
    multiplicity=1;
  end
  %Gb-Gb kind of pairs
  fprintf(fid,['#Void-Void %i\n (%s,%i/6)\n multiplicity %g\n substitutional_correlation(
    Var,Var,%s,%s,%s,%s,%s,\n%s %s,%s,%s,%s,%s,\n%s %s,%s,%s,%s,%s,\n%s %s,%s,%s,%s,%s,\n%
    s %s,%s,%s,%s,%s)\n substitutional_correlation(Var2,Var2,%s,%s,%s,%s,\n%s %s,%s,%s,%s,%s
    ,\n%s %s,%s,%s,%s,\n%s %s,%s,%s,%s)\n]\n',...
    neig,disp,neig,multiplicity,string_pp{1,1},string_pp{1,2},string_pp{1,3},string_pp
    {1,4},string_pp{1,5},gap,...
    string_pp{2,1},string_pp{2,2},string_pp{2,3},string_pp{2,4},string_pp
    {2,5},gap,...
    string_pp{3,1},string_pp{3,2},string_pp{3,3},string_pp{3,4},string_pp
    {3,5},gap,...
    string_pp{4,1},string_pp{4,2},string_pp{4,3},string_pp{4,4},string_pp
    {4,5},gap,...
    string_pp{5,1},string_pp{5,2},string_pp{5,3},string_pp{5,4},string_pp
    {5,5},...
    string_pp_tete{1,1},string_pp_tete{1,2},string_pp_tete{1,3},
    string_pp_tete{1,4},gap,...
    string_pp_tete{2,1},string_pp_tete{2,2},string_pp_tete{2,3},
    string_pp_tete{2,4},gap,...
    string_pp_tete{3,1},string_pp_tete{3,2},string_pp_tete{3,3},
    string_pp_tete{3,4},gap,...
    string_pp_tete{4,1},string_pp_tete{4,2},string_pp_tete{4,3},
    string_pp_tete{4,4});
else
  % the generator to GB-Te kind of pairs.
  fprintf(fid,['#Void-Te %i\n (%s,%i/6)\n multiplicity 2\n substitutional_correlation(Var
    ,Var2,%s,%s,%s,%s,%s,\n%s %s,%s,%s,%s,%s,\n%s %s,%s,%s,%s,%s,\n%s %s,%s,%s,%s,%s)\n]\n
    ',...
    neig,disp,neig,string_pp{1,1},string_pp{1,2},string_pp{1,3},string_pp{1,4},string_pp
    {1,5},gap,...
    string_pp{2,1},string_pp{2,2},string_pp{2,3},string_pp{2,4},string_pp
    {2,5},gap,...
    string_pp{3,1},string_pp{3,2},string_pp{3,3},string_pp{3,4},string_pp
    {3,5},gap,...
    string_pp{4,1},string_pp{4,2},string_pp{4,3},string_pp{4,4},string_pp
    {4,5});
end
end
fclose(fid);
fprintf('everything is calculated, check file correlations.txt\n')

```

Appendix B

YELL reference manual

This appendix presents the YELL online manual. A version of manual with hyperlinks is available in the html format from the program website <https://github.com/YellProgram/Yell> or by the following short link: <http://goo.gl/BBho1l>.

Yell Reference

Refers to the program version Yell 1.0
Latest manual update: October 21st, 2014

Yell Reference

Installation

How to run Yell

Input files

Yell data format for arrays

Allowed array input files

model.txt

Output files

Reciprocal space datasets

Δ PDF space datasets

Keywords reference

File format

Comments

Assignments

Formulas and variables

Arithmetic expressions

Variables

Refinable variables

Special functions that may be used in expressions

Preamble

Cell

DiffuseScatteringGrid

LaueSymmetry

Scale

RefinableParameters

RecalculateAverage

DumpPairs

PrintCovarianceMatrix

CalculationMethod

Parameters controlling approximate diffuse scattering calculation

FFTGridSize

FFTGridPadding

PeriodicBoundaries

ReportPairsOutsideCalculatedPDF

Parameters controlling the least square refinement

Refine

MaxNumberOfIterations

- MinimizerTau
- MinimizerThresholds
- MinimizerDiff

Unit cell

- Groups of atoms
- Atoms
- Naming the entities of the unit cell
- Symmetry

Modes

- TranslationalMode
- RotationalMode

Correlations

- Multiplicity
- Substitutional correlation
 - Example:
 - Extended example
 - Formal description
 - Neutral joint probabilities
 - Zero - neighbor joint probabilities

ADPCorrelation

- Example
- Example with correlation matrices
- Formal definition
- Neutral correlation
- Zero-neighbor correlation

SizeEffect

- Example
- Neutral SizeEffect
- Zero-neighbor correlation
- Formal definition

Epilogue

- Print

How to calculate multiplicity

- Definition
- First way to determine multiplicity
 - Example

Second way to determine multiplicity

- Example

Combinatorial symmetry

- Independent cones for all Laue groups

Equation for diffuse scattering calculation

References

Weber, T., & Simonov, A. (2012). The three-dimensional pair distribution function analysis of disordered single crystals: basic concepts. *Zeitschrift für Kristallographie*, 227(5), 238-247. [researchgate](#)

Schomaker, V., & Trueblood, K. N. (1968). On the rigid-body motion of molecules in crystals. *Acta Crystallographica Section B: Structural Crystallography and Crystal Chemistry*, 24(1), 63-76. [link](#)

Cyvin, S. J. (1968). Molecular vibrations and mean square amplitudes. *Universitets Forl.* Chapter 3

Simonov, A., Weber, T., Steurer, W., Yell - a computer program for diffuse scattering analysis via 3D- Δ PDF refinement. in. prep.

Installation

Yell is distributed as a self-contained executable. Just copy it in a folder where the shell can find it. Instructions for building the binary from the source files will be added later.

How to run Yell

Yell is a simple console application. In order to run it, open a terminal, navigate to the folder with the `model.txt` file and type

```
yell.exe
```

(on Windows) or

```
yell
```

(on Mac).

Input files

Yell data format for arrays

All input arrays should be presented in [hdf5](#) formatted files. The hdf5 format is commonly used in scientific applications to hold large data arrays. It is not editable by hand, but the libraries for preparing and manipulating the files are available for most of the software platforms including Matlab, Python, Java, C++ and R.

Each Yell-formatted hdf5 files should have the following structure: the root / of the file is expected to have attribute 'filetype' equal to 'Yell 1.0' and the dataset should be located at the path '/data' within the file. The array should have [row-major order](#). Though, in principle, hdf5 files may contain several datasets, Yell only expects one dataset per file.

Here is code to save file in this format in Matlab:

```
function write4yell(filename, dataset)
    dataset=permute(dataset, [3 2 1]); % Matlab uses column-major order,
    thus we need to permute dimensions
    h5create(filename, '/data', size(dataset));
    h5write(filename, '/data', double(dataset));
    h5writeatt(filename, '/', 'filetype', 'Yell 1.0');
```

Allowed array input files

The following files are used

- **experiment.h5** File containing the experimental diffuse scattering (in reciprocal space). Required for refinement.
- **weights.h5** File containing the least-squares weights (in reciprocal space). Optional. If not present unit weights are used.
- **reciprocal_space_multiplier.h5** File containing the array that will be multiplied to the calculated diffuse scattering (in reciprocal space). Useful for zeroing-out the model intensities in places where experiment was not measured. Optional.
- **pdf_multiplier.h5** File containing the array that will be multiplied to the calculated PDF. Useful to simulate resolution function effects. Optional.

All arrays must have the same shape along all dimensions

The following indicates at which stages the arrays are used:

1. Diffuse scattering is calculated.
2. If available, PDF multipliers are applied in PDF space
3. If available, reciprocal space multipliers are applied in reciprocal space

4. The resulting dataset is called 'model intensities'

During a refinement $weights * (I_{experiment} - I_{model})$ is used for minimization.

No input arrays files are required if the refinement option is switched off (see below).

model.txt

This file contains the definition of the local structure model. For description see [reference](#).

Output files

The output files are also in hdf5 format. The following shows a Matlab function which reads in the data:

```
function res=read_from_yell(filename)
    res = h5read(filename, '/data');
    res = permute(res,[3 2 1]); #Transform from row-major order
```

Reciprocal space datasets

- **model.h5** Diffuse scattering calculated from the model. If a refinement is performed, `model.h5` is calculated from the final parameters; refined intensities are scaled to the experimental intensities.
- **average.h5** Full scattering calculated from average pairs. Only takes pairs touched by correlations as defined in `model.txt`, i.e. broad Bragg peaks and truncation ripples are possible. Scaled to the experimental intensities.
- **full.h5** Full scattering from real structure pairs. Only takes those pairs which were touched by correlations. Scaled to the experimental intensities.

Δ PDF space datasets

All of these datasets are calculated only if [diffuse scattering grid](#) allows FFT. The datasets are scaled to be in e^2/A^3 , if the [multiplicity](#) is set correctly.

- **delta-pdf.h5** Modeled Δ PDF.
- **exp-delta-pdf.h5** Experimental Δ PDF, i.e. the normalized Fourier transform of the diffuse intensities provided by experiment.h5.
- **delta-delta-pdf.h5** Δ^2 PDF = Δ PDF_{exp} - Δ PDF_{model}.

Keywords reference

File format

The file `model.txt` is a plain text file which describes the crystal model. The file contains the following sections:

- **preamble** defines parameters which influence the calculation
- **UnitCell** section defines the average crystal structure
- **Modes** section defines the possible motions of the atoms or molecules
- **Correlations** section defines the short-range ordering in the crystal
- in **epilogue** the user may request Yell to print various expressions

The input file contains square brackets [] for grouping various elements, assignments for providing aliases to the various parts of the structure and also variables and arithmetic expressions in order to express constraints.

The parser in Yell is quite flexible considering white spaces. Space, tabulation, newline character and comment are all considered whitespace. In most parts of the input files the whitespace characters are not required, but user can add them to improve the readability.

Comments

Comments in Yell start with a sharp sign # and go until the end of the line.

```
#this is a comment
Cell 1 1 1 90 90 90 #this is also comment
```

Assignments

The parts of the structure - atoms, groups of atoms, variants and modes can be given aliases for the later reference in the model. The aliases are expressed as assignments

```
Cu1_alias = Cu 1 0 0 0 0.01
```

Formulas and variables

Arithmetic expressions

The parser of Yell allows the arithmetic expressions in the place of almost any number in the input file. The expressions will be calculated during diffuse scattering calculation according to usual mathematical rules. Thus the following line

```
Cell 6-1 1+2*2 4 9*(15-5) 3*3*2*5 120
```

is equivalent to

```
Cell 5 5 4 90 90 120
```

Expressions can contain the operators `+` `-` `*` `/`, brackets `()`, and special mathematical functions `exp` `log` `sin` `cos` `sqrt` `abs` `mod` and `pow`. The whitespace characters are **not** allowed within formulas and will invoke an error.

Variables

User can define variables to use in arithmetic expressions. The definition of variables is similar to many programming languages:

```
a=1;
b=12*3;
```

Variable definition should end with semicolon ; . The whitespace characters are **not** allowed in the variable definition, e.g. the expression `b = 12 * 3;` is illegal.

Valid names for variables start with upper or lower case letters, names may also contain numbers and underscore. The names are case sensitive. Here are the examples:

```
A=1;
shift_along_111=1;
```

The variables can then be used in any arithmetic expression. In any case *float* algorithms are used.

```
a=12.1;
c=a/2;
alpha=90;
gamma=60;
```

```
Cell a a c alpha alpha gamma*2
```

Variables can be defined only in the “top level” of the input file, i.e. in preamble, epilogue and at the top levels of the definitions of `UnitCell`, `Modes` and `Correlations`. Here is an example which explicitly describes where the variables can and cannot be defined

```
a=1;
Cell 10 10 10 90 90 90
b=1; #Here is ok
DiffuseScatteringGrid -10 -10 -10 0.5 0.5 0.5 20 20 20
UnitCell
[
  c=1; #Here is ok
  Var = Variant
  [
    #but not here
    (p=0.5) #this is ok because it is a keyword expression and not an
assignment to a variable 'p'
    [
      #not here
      Ag 1 0 0 0 0.02
    ]
    (p=0.5)
    Void
  ]
  d=1; #Here is ok
]
Modes
[
```

```

e=1; #Here is ok
]
Correlations
[
  f=1; #Here is ok
  [(0,0,0)
   #not here
   SubstitutionalCorrelation(Var,Var,0.5)
  ]
]
g=1; #Here is ok
Print "g=" g

```

Refinable variables

User can also define a variable as a *refinable variable*. Such variables will be used in the least squares refinement procedure. Their value will be optimized in order to fit experimental data. Other variables, which were not defined as refinable, but depend on refinable variables in the course of refinement will always be updated, based on the actual value of refinable variables they depend upon.

Refinable variables should be refined in the preamble of the file, with `RefinableVariables` keyword. Example:

```

RefinableVariables
[
  a=10; #This variable will be changed in the course of a refinement
  ]
  b=a*2; #This variable will change its value if the value of variable 'a' is
  modified during the refinement.

```

Refinable variables can **not** be initialized by expressions:

```

RefinableVariables
[
  a=10/2; #Error
  ]

```

Special functions that may be used in expressions

`exp(x)` = e^x

`log(x)` Natural logarithm. Invokes error, if $x \leq 0$.

`sin(x)` sine

```
cos(x) cosine
sqrt(x) =  $\sqrt{x}$ . Invokes error if  $x < 0$ .
abs(x) =  $|x|$ 
mod(x,y) =  $x \bmod y$ 
pow(x,y) =  $x^y$ 
```

Preamble

The first part of `model.txt` contains general settings of the calculation, like definition of the unit cell, grid for diffuse scattering calculation, point group, definition of refinable variables, the calculation method and several parameters which can affect the calculation and refinement process.

The mandatory fields are `Cell`, `DiffuseScatteringGrid` and `PointGroup`, other keywords can be omitted. In such cases the program will use default values.

Cell

Mandatory field.

Defines the unit cell of the crystal.

Format: Cell a b c α β γ .

Units: Ångströms and degrees.

Example:

```
Cell 5.406 5.406 5.406 90 90 90
```

DiffuseScatteringGrid

Mandatory field.

Defines the grid on which the diffuse scattering should be calculated. In refinements experimental data and weights are to be provided with the same grid definition.

The format is the following:

```
DiffuseScatteringGrid lower_limit_x lower_limit_y lower_limit_z
                        step_size_x step_size_y step_size_z
                        number_of_pixels_x number_of_pixels_y
                        number_of_pixels_z
```


where `lower_limits` are the minimal `h`, `k` and `l` indices of the grid, `step_sizes` are the distances between lattice points in units of `h`, `k` and `l`, and `number_of_pixels` are the total number of pixels along each direction.

The main axes of the grid can only be defined along the main axes of the unit cell. If one needs to calculate the diffuse scattering grid along special directions, like, for example `111`, one has to transform the unit cell accordingly.

Yell uses a Fast Fourier Transform (FFT) algorithm for switching between reciprocal and PDF space. This adds two constraints on the diffuse scattering grid:

1. number of pixels along each dimension must be even
2. the origin of reciprocal space must be in the pixel with coordinates `(number_of_pixels_x/2+1, number_of_pixels_y/2+1, number_of_pixels_z/2+1)`

The conditions are a bit unusual for anyone who did not have experience with the FFT before. The arrays with uneven number of pixels and the central pixel in the center seems more natural. In order to use FFT such “natural” datasets should be stripped of the last planes along `x`, `y` and `z` directions.

The FFT algorithm allows to calculate cross-sections through the center of reciprocal space. Such cross-sections in reciprocal space correspond to the projection of the whole structure to a plane, or a line in PDF space. If the cross-section is calculated the number of pixels is allowed to be equal to 1 along some axis, then the `step_size` along the corresponding dimension is ignored, the `lower_limit` should be equal to 0.

Examples:

```
DiffuseScatteringGrid -5 -5 -5 0.1 0.1 0.1 50 50 50 #Usual three-  
dimensional case  
DiffuseScatteringGrid -10 -10 0 0.1 0.1 1 20 20 1 #hk0 section of  
diffuse scattering
```

If the grid is not consistent with the above mentioned rules, diffuse scattering can be calculated using the [direct](#) calculation algorithm and refinement can still be performed. However, the FFT algorithm is required if you want to

- apply **pdf_multipliers.h5**
- use [fast diffuse scattering calculation algorithm](#)
- obtain **delta-pdf.h5**, **exp-delta-pdf.h5** and **delta-delta-pdf.h5**.

The following example only works with ‘exact’ calculation method:

```
DiffuseScatteringGrid 0 0 0 0.2 0.2 0.2 30 30 30 #Only works with  
'exact' calculation method
```

LaueSymmetry

Mandatory field.

Possible values: `m-3m` `m-3` `6/mmm` `6/m` `4/mmm` `4/m` `-3:R` `-3:H` `-3m:H` `-3m:R` `mmm` `2/m`
`2/m:b` `-1`

Defines the Laue group of the crystal.

Laue groups which have both rhombohedral and hexagonal settings are noted by the `:R` or `:H` letters. The group `2/m` also have two standard settings, with unique `c` and with unique `b`. They are called `2/m` and `2/m:b` respectively. Non-crystallographic symmetry, e.g. five-fold rotation axes, are not supported by the program, but can be provided explicitly by the definition of correlations (see below).

Scale

default value: 1

Scaling coefficient between model and experiment; always refined.

RefinableParameters

default: empty `[]`

Registers variables in the square brackets to be used during refinement. The variables cannot be initialized with expressions.

Example:

```
RefinableVariables
[
  a=1;
  b=12;
]
```

The scale coefficient, which is also refined, should not be defined here, since it has a special [keyword](#).

RecalculateAverage

possible values: `true` `false`

default value: `true`

Controls whether the average PDF of the structure should be recalculated during refinement. If set to

`false` the average PDF is only calculated in the beginning of refinement and kept unchanged throughout the refinement.

In cases, when non of the refined variables may influence the average interatomic vectors, because the average structure is very well known, the recalculation can be turned off. This speeds up the refinement by the factor of two. It is safe to do so, when refinable variables do not appear in sections `UnitCell`, `Modes`, and in `Correlations` in the brackets `(1,0,0)` and in `Multiplicity`.

DumpPairs

possible values: `true` `false`

default: `false`

When turned on, prints all the interatomic pairs used to calculate diffuse scattering in the program output.

PrintCovarianceMatrix

possible values: `true`, `false`

default: `false`

When turned on, prints the full refinement covariance matrix in the program output.

CalculationMethod

possible values: `exact` `approximate`

default value: `exact`

Selects the diffuse scattering calculation method. Yell currently has two calculation methods:

- `exact` uses direct sum over all pairs, as shown in [this formula](#). Calculation is done in reciprocal space.
- `approximate` goes through real space. This works much quicker because for each interatomic pair it only calculates a certain block where the signal is significant (controlled by flag `FFTGridSize`) and ignores most of the PDF space, where the signal is almost zero. Details are described in ([Simonov et al. in. prep.](#))

The approximate algorithm provides a significant speedup, but introduces errors. It is advised to properly set up the `approximate` algorithm in the first stages of refinement, but always check the results with `exact` refinement before publishing.

Parameters controlling approximate diffuse scattering calculation

FFTGridSize

expects: three numbers

default: 16 16 16

Controls the size of parallelepiped in pixel units to calculate the PDF signal of a pair. Bigger values give better accuracy, smaller provide more speed. The values should not be bigger than the size of the dataset along the same dimension.

FFTGridPadding

expects: three numbers

default: 0 0 0

Defines the padding in the approximate algorithm.

Prior to PDF calculation, Yell extends reciprocal space by the selected number of pixels; after the calculation the padding is cut.

Frequently, the FFT approximation method has the strongest errors close to the edges of reciprocal space. The padding expands the reciprocal space putting the errors outside the region of interest.

Details are described in ([Simonov et al. in. prep.](#)).

PeriodicBoundaries

expects: three booleans

default: true true true

During the `approximate` diffuse scattering calculation allows not to periodically repeat the Δ PDF signals which lie outside the calculated PDF grid along specified directions.

Why this might be interesting

If the correlations are very long in PDF space, one is forced to reconstruct diffuse scattering on a very fine grid, and use a huge array for Δ PDF refinement. However, sometimes the values of the correlations with long correlation vectors are not interesting. In such case there is a trick one can use to reduce the amount of memory Yell requires for refinement.

Experimental data is prepared in a special way. From the reconstruction on the fine grid one calculates the Δ PDF. Then cuts a central part of the Δ PDF, containing the region of interest. And finally back-transforms the part of the Δ PDF back to reciprocal space.

This procedure introduces strong truncation ripples in reciprocal space, but preserves the Δ PDF. The Δ PDF can then be refined in Yell with the FFT method. The obtained correlations will be accurate, though the diffuse scattering will be modeled poorly.

In such case interatomic pairs which lie close to the edge of calculated Δ PDF region. The part of the Δ PDF signal from such pairs will lie inside the calculated region, and part of the signal will lie outside the calculated region. The *direct* calculation method and the *approximate* algorithm with default settings will wrap the Δ PDF signals which lie outside the calculated region into the calculated area. Turning periodic boundaries off along these dimensions avoids this problem.

Thus, whenever the above mentioned method of preparing diffuse scattering is used, the periodic boundaries should always be turned off along the dimensions along which the Δ PDF map was cut.

ReportPairsOutsideCalculatedPDF

possible values: `true`, `false`

default: `false`

Reports all the pairs which lie outside calculated Δ PDF map.

If an interatomic pair fall outside the calculated Δ PDF map, its signal will be periodically wrapped inside the Δ PDF map. For diffuse scattering comprising layers and streaks this is a valid behavior. However, if the diffuse scattering is not broad along some dimension, but certain pairs fall outside the calculated Δ PDF map along that dimension, this is an indication that diffuse scattering should be reconstructed on finer grid.

During the refinement with a [specially prepared dataset](#), with some [PeriodicBoundaries](#) turned off, the signal from most of the interatomic pairs which lie outside the calculated region will be silently discarded. Turning on this option allows to find unexpected behavior caused by this feature.

Parameters controlling the least square refinement

Refine

Possible values: `true`, `false`

Default: `true`

Specifies whether the program should refine diffuse scattering. If set to `false`, the program will just calculate diffuse scattering from a given model.

MaxNumberOfIterations

possible values: integer
default: 1000

Defines the maximum number of iterations the refinement algorithm is allowed to run.

MinimizerTau

default: 1E-03

Defines the initial Levenberg-Marquandt damping parameter. Note that this value works differently, than the `damping` in standard crystallographic packages. For more information see this [wikipedia article](#) and also somewhat minimalistic [manual](#) of the levmar library.

MinimizerThresholds

possible values: three numbers
default: 1E-17 1E-17 1E-17

Defines the thresholds to detect convergence of least squares procedure. The thresholds are for the size of the gradient, minimization step size, the third criterion works if the experimental and diffuse scattering are the same (say, for synthetic data). For more info see `\epsilon_1` `\epsilon_2` and `\epsilon_3` in [levmar documentation](#).

MinimizerDiff

possible values: one number
default: 1E-06

Yell calculates derivatives by finite difference method. This value provides the increment that is used for the calculation of derivatives.

Unit cell

This section of the input file defines the contents of the crystal's average unit cell.

In Yell, the average structure is defined in a hierarchical manner: the unit cell consists of topological sites, which are occupied by atoms or groups of atoms. The topological sites are called `Variants` and can be disordered (occupied with a certain probability by more than one group of atoms). Groups of atoms represent ensembles, like molecules or clusters.

Example:

```

UnitCell[
  GdFeVoid_site=Variant [
    (p=1/3)
    Gd 1 0 0 0 0.035 0.035 0.035 0 0 0

    (p=1/3)
    iron_molecule = [
      Fe 1 0 0 0.2845 0.035 0.035 0.035 0 0 0
      Fe 1 0 0 -0.2845 0.035 0.035 0.035 0 0 0
    ]

    (p=1/3)
    Void
  ]
]

```

In this example `GdFeVoid_site` is either occupied by a gadolinium atom, by a structural building block of two iron atoms, which is aliased as `iron_molecule`, or by a `Void`. On average, each chemical unit is present with the same probability of $p=1/3$.

Groups of atoms

Atoms can be grouped into one entity. These entities represent atoms which can not appear without each other, usually molecules or clusters.

The groups are defined without keyword, just with square brackets `[]`.

Example:

```

Fe2_molecule = [
  Fe 1 0 0 0.2845 0 0 0 0 0 0
  Fe 1 0 0 -0.2845 0 0 0 0 0 0
]

```

The grouping is recurrent, each group may contain not only atoms, but also other groups.

Example:

```

Fe2_molecule = [

  Fe 1 0 0 0.2845 0 0 0 0 0 0

  lower_iron = [
    Fe 1 0 0 -0.2845 0 0 0 0 0 0
  ]
]

```

Atoms

Atoms can be defined in two formats with both isotropic and anisotropic ADP:

```
AtomType mult x y z Uiso
AtomType mult x y z U11 U22 U33 U12 U13 U23
```

`x`, `y` and `z` are fractional coordinates, `Uiso` and `Uij` are atomic displacement parameters in Å². The multiplier `mult` can be thought of as a site multiplicity in the average structure. The occupancy of each atoms is defined as `mult*p` where `p` is the probability of the atomic group the atom belongs to (defined in the construction `(p=...)`) and `mult` is the multiplier. In the vast majority of the cases `mult` should be equal to 1. It is only different if the atomic group contains disorder which is not disentangled into different entries of `Variant`.

Example:

```
Fe 1 0 0 -0.2845 0 0 0 0 0 0
```

Naming the entities of the unit cell

Variants, atomic groups and atoms can be given names for symmetry expansion and use in `Correlations` and `Modes`. Example:

```
Variant1 = Variant[
  (p=1)
  Iron_molecule = [
    Fe1 = Fe 1 0 0 0.2845 adp_perp adp_perp adp_z 0 0 0
    Fe2 = Fe 1 0 0 -0.2845 adp_perp adp_perp adp_z 0 0 0
  ]
]
```

Symmetry

It is possible to apply symmetry elements to atoms and groups of atoms. The syntax is the following:

```
atomic_group*Symmetry(x,y,z)
```

Example for mirroring an atom with respect to the xy0 plane:


```
Fe_atom = Fe 1 0 0 0.2845 0.01
symmetry_equivalent_Fe = Fe_atom*Symmetry(x,y,-z)
```

Unlike in the average crystal structure, the symmetry elements are sensitive to integer translations. In cases when the center of the molecule is not in $(0, 0, 0)$ user should provide the appropriate translations. Here is example for the same iron molecule, with centered at $(0, 0, 1)$

```
Fe_atom = Fe 1 0 0 0.7155 adp_perp adp_perp adp_z 0 0 0
symmetry_equivalent_Fe = Fe_atom*Symmetry(x,y,2-z) #note 2-z here
# resulting coordinates are 0 0 1.2845
```

Modes

Modes is a way to define motions of atoms and molecules in Yell. All the motions are linear meaning that a displacement of each atom Δr is expressed as a linear combination of translation along all the modes:

$$\Delta r_i = \sum_n M_i^n \xi_n$$

where M_i^n is a set of vectors attached to each atom in a molecule, ξ_n is the amplitude of a displacement of the molecule along mode n .

Currently Yell supports two types of modes: translations and linear approximation of rotations.

TranslationalMode

Defines movements of atomic groups as a whole; expressed in Å.

Format: `TranslationalMode(atomic_group,axis)`

or `TranslationalMode(atom,axis)`

where `atomic_group` is a name to a group of atoms, and `axis` could be either `x`, `y` or `z`.

Example:

```
manganese_x = TranslationalMode(manganese,x)
```

RotationalMode

Defines linear approximation to rotation of a group of atoms along arbitrary axis; expressed in radians.

Format: `RotationalMode(atoimc_group, axis_x, axis_y, axis_z, center_x, center_y, center_z)`

where `axis_x` `axis_y` and `axis_z` define the rotation axis direction and `center_x` `center_y` `center_z` define a point on the rotation axis.

A displacement vector for each atom is defined as

$$\mathbf{r}_{axis} \times \frac{\mathbf{r}_{atom} - \mathbf{r}_{center}}{|\mathbf{r}_{atom} - \mathbf{r}_{center}|}$$

Correlations

In this section, short-range order pairwise correlations are defined. There are three basic types of correlations: substitutional correlations, displacement correlations and size effect (see [Weber & Simonov, 2012](#)). The combination of these correlations may describe any disorder in a disordered crystal.

In Yell, correlations which belong to the same group of atoms should be combined together in groups using square brackets. Example:

```
Correlations [  
  [(1,0,0)  
  Multiplicity 2  
  SubstitutionalCorrelation(CuVoid, AuVoid, 0.25+dp)  
  SizeEffect(Cu, Au_x, dx)  
  ADPCorrelation(Cu_x, Au_x, dUx)  
  ]  
]
```

Each group starts with a vector (u, v, w) , defining the lattice vector between the correlated atoms or molecules. It corresponds to the $\mathbf{R}_{u,v,w}$ in ([Weber & Simonov, 2012](#)). The elements of the vector are not required to be integer, and could for example be multiples of 0.5 (in case of C centering) and even irrational (e.g. in case of quasicrystals). The corresponding inter-atomic vector is calculated as $\mathbf{R}_{u,v,w} + \mathbf{r}_{Au} - \mathbf{r}_{Cu}$, where \mathbf{r}_{Cu} and \mathbf{r}_{Au} are the atomic coordinates of the copper and gold atoms defined within `CuVoid` and `AuVoid` variants (not shown in the example).

Multiplicity

Format: Multiplicity m
where m is number.

Defines the multiplicity of the interatomic or intermolecular pair with respect to the Laue group.

Effect: multiplies m with the multiplicity of each atomic pair in current correlation group. Does not have a letter dedicated to it in (Weber & Simonov, 2012), but is equivalent in effect by multiplication m to both joint probability p_{uvw}^{mn} and Patterson multiplicity $c_n c_m$.

In Yell, the Laue group is applied to the diffuse scattering automatically. However, Yell currently cannot automatically calculate the multiplicity of each correlation set. It is expected that the user manually provides the multiplicity for it.

The multiplicity m of a group should be equal to the number of times such correlation group appears in the Δ PDF space. The correlations in the center of Δ PDF space usually get multiplicity 1, the correlations on general positions get multiplicity that is equal to the order of the crystal Laue group.

For an example see the model for **FeVoid**.

Warning: The definition of the multiplicity has to be done with great care. Wrongly defined multiplicities may be perfectly compensated by over- or underestimated structure correlation parameters leading to errors that are difficult to recognize!

For detailed instructions on how to apply multiplicity see [this section](#).

Substitutional correlation

Substitutional correlation means that the occupancies of two Variants in the structure depend on each other.

Effect: sets the joint probabilities of atomic pairs to the ones specified as arguments. Affects p_{uvw}^{mn} as stated in (Weber & Simonov, 2012).

Format: SubstitutionalCorrelation(variant_A, variant_B, p11, p12, ..., pnm)

short format:

SubstitutionalCorrelation(variant_A, variant_B, p11, p12, ..., pn_minus_1m_minus_1)

where p_{ij} are the joint probability coefficients to find both blocks A_i and B_j present at the same time, the $p_{n-1,m-1}$ is the joint probability $p_{n-1,m-1}$.

Example:

```
SubstitutionalCorrelation(CuAu, CuAu, 0.25+x)
```

Extended example

For the unit cell:

```
UnitCell [  
  AuCu = Variant [  
    (p=1/3)  
    Au 1 0 0 0 0  
    (p=2/3)  
    Cu 1 0 0 0 0  
  ]  
]  
Correlations [  
  [(0,0,0)  
  SubstitutionalCorrelation(AuCu,AuCu,1/3) #corresponds to the matrix 1/3 0  
                                     #                                     0 2/3  
]  
]
```

This will produce the following pairs (differences are typed in red):

	m	p	x	y	z	Uxx	...	Uyz	\bar{p}	\bar{x}	\bar{y}	\bar{z}	\bar{U}_{xx}	...	\bar{U}_{yz}
Au Au	1	0.333333	0	0	0	0	0	0	0	0.111111	0	0	0	0	0
Cu Au	1	0	0	0	0	0	0	0	0	0.222222	0	0	0	0	0
Au Cu	1	0	0	0	0	0	0	0	0	0.222222	0	0	0	0	0
Cu Cu	1	0.666666	0	0	0	0	0	0	0	0.444444	0	0	0	0	0

Here the m marks pair multiplicity, p is joint probability, x y z are interatomic vector coordinates, U_{xx} ... U_{yz} are the components of the joint ADP tensors, the letters with overbar \bar{p} \bar{x} \bar{y} \bar{z} \bar{U}_{xx} ... \bar{U}_{yz} relate to the average probability, interatomic vector and ADP tensor.

Formal description

Assume that

```
variant_A = Variant [  
  (p=pA1)  
  A1  
  (p=pA2)  
  A2  
  ...  
  (p=pAn)  
  An  
]  
variant_B = Variant [  
  (p=pB1)  
  B1  
  (p=pB2)  
  B2  
  ...  
]
```

```
(p=pBm)
Bm
]
```

Where A_i and B_j are some chemical units, or void. Then, the full matrix of joint probabilities has size $m \times n$ and the form

$$P = \begin{bmatrix} p_{A_1B_1} & p_{A_2B_1} & \cdots & p_{A_nB_1} \\ p_{A_1B_2} & p_{A_2B_2} & \cdots & p_{A_nB_2} \\ \cdots & \cdots & \cdots & \cdots \\ p_{A_1B_m} & p_{A_2B_m} & \cdots & p_{A_nB_m} \end{bmatrix}$$

This will be expressed in yell in the following way:

```
SubstitutionalCorrelation(variant_A, variant_B, pA1B1, pA2B1, ..., pAnB1,
                           pA1B2, pA2B2, ..., pAnB2,
                           ...,
                           pA1Bm, pA2Bm, ..., pAnBm)
```

Since possibilities A_i are all exclusive and $\sum_i p_{A_i} = 1$ there are $m + n - 1$ independent constraints on $p_{A_iB_j}$ which constrain the sum of pair probabilities to probabilities of single chemical units:

$$\sum_i p_{A_iB_j} = p_{B_j}$$

$$\sum_j p_{A_iB_j} = p_{A_i}$$

The constraints make it possible to calculate the last row and last column of the matrix P , resulting in

$$P = \begin{bmatrix} p_{A_1B_1} & \cdots & p_{A_{n-1}B_1} & p_{B_1} - \sum_{i=1}^{n-1} p_{A_iB_1} \\ \cdots & \cdots & \cdots & \cdots \\ p_{A_1B_{m-1}} & \cdots & p_{A_{n-1}B_{m-1}} & p_{B_{m-1}} - \sum_{i=1}^{n-1} p_{A_iB_{m-1}} \\ p_{A_1} - \sum_{j=1}^{m-1} p_{A_1B_j} & \cdots & p_{A_{n-1}} - \sum_{j=1}^{m-1} p_{A_{n-1}B_j} & p_{A_nB_m} \end{bmatrix}$$

where

$$p_{A_nB_m} = 1 + \sum_{i,j=1}^{n-1,m-1} p_{A_iB_j} - \sum p_{A_i} - \sum p_{B_j}$$

Yell can automatically calculate the last row and the last column, given the upper left $(m - 1) \times (n - 1)$ independent part of joint probability matrix. Thus the above definition can be equally well expressed in the short form:

```
SubstitutionalCorrelation(variant_A, variant_B, pA1B1, pA2B1, ..., pAn-1B1,
                          pA1B2, pA2B2, ..., pAn-1B2,
                          ...
                          pA1Bm-1, pA2Bm-1, ..., pAnBm-1)
```

The short form is in general preferred, because it is a non-redundant representation of the structure model.

Neutral joint probabilities

If the occupancies of two variants are independent, e.g. because their interatomic distance is larger than the correlation length of the corresponding local order phenomenon, the joint probabilities are equal to the product of the occupancies $P_{AiBj}^{Patterson} = p_{Ai}p_{Bj}$. Thus the following probability matrix will produce no diffuse scattering and no signals in PDF space (and can be omitted):

$$P_{neutral} = \begin{bmatrix} p_{A1}p_{B1} & p_{A2}p_{B1} & \dots & p_{An}p_{B1} \\ p_{A1}p_{B2} & p_{A2}p_{B2} & \dots & p_{An}p_{B2} \\ \dots & \dots & \dots & \dots \\ p_{A1}p_{Bm} & p_{A2}p_{Bm} & \dots & p_{An}p_{Bm} \end{bmatrix}$$

Zero - neighbor joint probabilities

For the zero neighbor correlation, i.e. the correlation of a Variant with itself, the diagonal elements of the joint probability matrix should be equal to the average occupancy of the corresponding elements, and off-diagonal elements should be equal to zero:

$$P_{zero} = \begin{bmatrix} p_{A1} & 0 & \dots & 0 \\ 0 & p_{A2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & p_{An} \end{bmatrix}$$

Or in the Yell (long) form:

```
[ (0, 0, 0)
  SubstitutionalCorrelation(variant_A, variant_A, pA1, 0, ..., 0,
                            0, pA2, ..., 0,
                            ...
                            0, 0, ..., pAn) ]
```

It is important to note that the zero-neighbor correlation **must** be defined for all Variants showing occupational disorder and that the joint probabilities should not be refined but be the same as the occupation probability (p=...) defined in the UnitCell statement.

In practice the zero neighbor correlation is frequently over- or underestimated, because incomplete background subtraction or elimination of broad diffuse scattering with the background (=

overcorrection of background) may influence the zero neighbor correlation definition and thus the correct determination of the scale factor. As shown in (Weber & Simonov, 2012), such errors may lead to significant systematic deviations in the determination of other pair correlation parameters. Careful background determination is therefore indispensable for a high quality local structure refinement.

ADPCorrelation

This correlation appears when displacements of two atoms or molecules are not independent.

Effect: changes β_{uvw}^{mn} as stated in (Weber & Simonov, 2012)

Format: `ADPCorrelation(Mode1, Mode2, cov)`

where `cov` is a covariance of the displacements along the two modes $cov = \langle \xi_1 \xi_2 \rangle$; `cov` is expressed in `units1 * units2` where `unitsi` are the units of the two modes (e.g. Å for translational modes and rad for rotational ones).

The correlations of atomic displacements typically manifest themselves as thermal diffuse scattering (TDS), but the correlations could also be of static origin.

Yell assumes that displacements of all atoms in the crystal are jointly Gaussian and the distribution of the displacement differences $\Delta r_{ik} - \Delta r_{jl}$ is also Gaussian. The correlations of the displacements are expressed in terms of covariances of collective displacement of the blocks along `Modes`. Analogously to independent substitutional correlations ADPCorrelations do not need to be defined, if `Modes` are not correlated with any other mode. In fact, most of the internal modes of rigid molecules are not expected to be correlated.

Note that some ADP correlations, though symmetrically independent, can produce the same signal in PDF space. One very practical example is the covariances $\langle x_1 y_2 \rangle$ and $\langle y_1 x_2 \rangle$ produce exactly the same effect in PDF space, and should therefore be constrained.

The ADP correlations are subject of symmetry constraints by the symmetry of interatomic or intermolecular pairs. The constraints to the correlations of translational modes are equivalent to the site-symmetry ADP constraints in the average structure. The constraints involving rotational modes are equivalent to the TLS constraints (Schomaker & Trueblood 1968).

Example

```
#Cu structure (fcc)
#the unit cell has Fm-3m symmetry (225)
LaueSymmetry m-3m
UnitCell [
```

```

Variant [
  (p=1)
  Cu = Cu 1 0 0 0 0.01
]]
Modes [
  Cu_x = TranslationalMode(Cu,x)
]
Correlations[
  ...
  [(1,0,0)
  ADPCorrelation(Cu_x,Cu_x,0.001)
  ADPCorrelation(Cu_y,Cu_y,0.0001)
  ADPCorrelation(Cu_z,Cu_z,0.0001)]
  ...
]

```

Note that the yy and zz correlations are the same, while xx is different according to symmetry of the pair. Let us assume that the symmetry of this pair is described by the point group $4/m\bar{3}m$ with the 4-fold axis along **a** of the cubic crystal. This restricts the covariance matrix to the following form:

$$\begin{bmatrix} cov_{xx} & & \\ & cov_{yy} & \\ & & cov_{yy} \end{bmatrix}$$

Example with correlation matrices

The covariance matrices can be recalculated in correlation matrices.

$$corr_{ij}^{mn} = \frac{cov_{ij}^{mn}}{\sqrt{U_{ii}^{aver,m} U_{jj}^{aver,n}}}$$

where indices $i, j = 1, 2, 3$ mark tensor components, m and n count atoms, $corr_{ij}^{mn}$ is a correlation matrix, cov_{ij}^{mn} is a covariance matrix (in \AA^2), the $U_{ii}^{aver,m}$ and $U_{jj}^{aver,n}$ are the average ADP tensors of atoms m and n (in \AA^2).

In Yell correlation matrix can be calculated in the epilogue using the `Print` command. The above mentioned example becomes:

```

#Cu structure (fcc)
#the unit cell has Fm-3m symmetry (225)
LaueSymmetry m-3m
UnitCell [
  Uiso=0.01; #define a variable for Uiso
  Variant [
    (p=1)
    Cu = Cu 1 0 0 0 U
  ]
]

```



```

]]
Modes [
  Cu_x = TranslationalMode(Cu,x)
]
Correlations[
  ...
  cov_11=0.001;
  cov_22=0.002;
  [(1,0,0)
  ADPCorrelation(Cu_x,Cu_x,cov_11)
  ADPCorrelation(Cu_y,Cu_y,cov_22)
  ADPCorrelation(Cu_z,Cu_z,cov_22)]
  ...
]
Print "corr_11=" cov_11/Uiso " corr_22=" cov_22/Uiso " corr_33="
cov_22/Uiso

```

Will print:

```
Requested output: corr_11=0.1 corr_22=0.01 corr_3=0.01
```

Formal definition

The theory is somewhat similar to the normal mode analysis. The notation here is analogous to [\(Cyvin 1968\)](#).

1. Denote the equilibrium configuration of N atoms in a molecule α as

$$\mathbf{R}_{\alpha i} = \{X_{\alpha i}, Y_{\alpha i}, Z_{\alpha i}\} = \{R_{\alpha i1}, R_{\alpha i2}, R_{\alpha i3}\}$$

for $i=1,2,\dots,N$. Then the displacements are introduced as deviations from equilibrium:

$$u_{\alpha ij} = R_{\alpha ij} - \langle R_{\alpha ij} \rangle$$

2. Introduce another set of coordinates ξ_{α} which we identify as a set of internal coordinates and introduce a set of matrices M that transform internal coordinates into “external” atomic displacements:

$$u_{\alpha ij} = \sum_k M_{\alpha ij}^k \xi_{\alpha k}$$

If ξ has $3N$ independent coordinates and there exist an inverse transformation

$$\xi_{\alpha i} = \sum_{jk} M_{\alpha i}^{-1jk} u_{\alpha jk}$$

both representations are equivalent.

In Yell, the matrices $M_{\alpha ij}^k$ are called *Modes*, the variables $\xi_{\alpha i}$ are referred to as *the amplitudes of displacement along corresponding modes*.

3. Recall, that the elements of the ADP matrix of an individual atom k in the molecule α is defined as follows:

$$U_{\alpha k}^{ij} = \langle \mathbf{u}_{\alpha ki} \mathbf{u}_{\alpha kj} \rangle$$

The paper (Weber & Simonov, 2012) uses the notation $\beta_{\alpha k}^{ij}$; it is equivalent notation since $\beta_{\alpha k}^{ij} = 2\pi^2 a_i^* a_j^* U_{\alpha k}^{ij}$, where a_i^* and a_j^* are the lengths of the corresponding reciprocal lattice vectors.

4. A covariance between the displacements of two atoms belonging to two different molecules can be expressed in terms of covariances of the modes:

$$\langle \mathbf{u}_{\alpha ij} \mathbf{u}_{\beta kl} \rangle = \langle \sum_m M_{\alpha ij}^m \xi_{\alpha m} \sum_n M_{\beta kl}^n \xi_{\beta n} \rangle = \sum_{mn} M_{\alpha ij}^m M_{\beta kl}^n \langle \xi_{\alpha m} \xi_{\beta n} \rangle$$

5. The joint atomic displacement parameter matrix of one atom as seen from another atom can be expressed as follows:

$$\begin{aligned} U_{\alpha k, \beta l}^{ij} &= \langle (\mathbf{u}_{\alpha ki} - \mathbf{u}_{\beta li})(\mathbf{u}_{\alpha kj} - \mathbf{u}_{\beta lj}) \rangle = \\ &= \langle \mathbf{u}_{\alpha ki} \mathbf{u}_{\alpha kj} \rangle + \langle \mathbf{u}_{\beta li} \mathbf{u}_{\beta lj} \rangle - \langle \mathbf{u}_{\beta li} \mathbf{u}_{\alpha kj} \rangle - \langle \mathbf{u}_{\alpha ki} \mathbf{u}_{\beta lj} \rangle = \\ &= U_{\alpha k}^{ij} + U_{\beta l}^{ij} - \sum_{mn} (M_{\alpha ij}^m M_{\beta kl}^n + M_{\alpha ij}^n M_{\beta kl}^m) \langle \xi_{\alpha m} \xi_{\beta n} \rangle \end{aligned}$$

Note that in order to define the ADP correlations it is not necessary to construct the full set of internal modes. It is significant to know the ADPs of each of the atoms from the average structure and only the covariances of the modes which are correlated.

Neutral correlation

If the displacements of two molecules are independent, all the covariances are equal to zero and don't need to be listed explicitly.

Zero-neighbor correlation

In theory, the covariance in the zero neighbor is equal to the square of the amplitude of the mode. For example, for single atom, the ADP correlations of the zero neighbor should be identical with its ADP parameter:

```
adp_cu=0.01;
```

```
UnitCell [
```

```
Variant [
```

```

    (p=1)
    Cu = Cu 1 0 0 0 adp_cu
  ]]
  Modes [
    Cu_x = TranslationalMode(Cu,x)
    Cu_y = TranslationalMode(Cu,y)
    Cu_z = TranslationalMode(Cu,z)
  ]
  Correlations [
    [(0,0,0)
    ADPCorrelation(Cu_x,Cu_x,adp_cu)
    ADPCorrelation(Cu_y,Cu_y,adp_cu)
    ADPCorrelation(Cu_z,Cu_z,adp_cu)]
  ]

```

Similar to substitutional correlations zero neighbor ADP correlations may be heavily affected by over- or under-corrected experimental background.

Note that in substitutionally disordered crystals ADP correlations are frequently not seen because corresponding diffuse scattering is usually very weak. In such crystals zero-neighbor ADP correlation need only be defined if there is a clear indication to ADP correlations or size-effect correlations.

SizeEffect

Size effect occurs when a substitutional disorder in a `Variant` is correlated with a static displacement of another chemical unit.

Effect: changes the u_{www}^{mn} as stated in [\(Weber & Simonov, 2012\)](#)

Format: `SizeEffect(AtomicGroup1, Mode2, amp)`

or `SizeEffect(Mode1, AtomicGroup2, amp)`

The meaning of the `SizeEffect(AtomicGroup1, Mode2, amp)` : in the presence of `AtomicGroup1` the `AtomicGroup2` has an average displacement along the `Mode2` by the amount defined by the amplitude `amp` (expressed in the units of the `Mode2` e.g. Å for translational mode, rad for rotational)

Example

```

UnitCell [
  Variant [
    (p=0.99)
    Na = Na 1 0 0 0 0.01
  ]

```

```

    (p=0.01)
    Void
  ]]
  Modes [
    Na_x = TranslationalMode(Na,x)
  ]
  Correlations [
    ...
    [(1,0,0)
     SizeEffect(Na,Na_x,0.02)
    ]
    ...
  ]
]

```

Neutral SizeEffect

When the correlation is absent, the size effect is equal to zero and is not required to be defined.

Zero-neighbor correlation

The size-effect for the zero-neighbor is equal to zero. However, it is important to add the zero neighbor from the ADPCorrelation, otherwise the model might produce negative intensities.

Example:

```

#A disordered Fe-Ni alloy
#Space group Fm-3m
PointGroup m-3m
UnitCell
[
ADP=0.0004;
  FeNi = Variant[
    (p=1/2)
    Fe = Fe 1 0 0 0 ADP

    (p=1/2)
    Ni = Ni 1 0 0 0 ADP
  ]
]

Modes[
  Fe_x = TranslationalMode(Fe,x)
  Ni_x = TranslationalMode(Ni,x)
]

```

```

#... same for y and z
]
Correlations [
  [(0,0,0)
    SubstitutionalCorrelation(FeNi,FeNi,1/2)
    #IMPORTANT: even though the crystal does not have ADP correlation,
    #zero neighbor ADP correlation should be input in presense of size
effect:
    ADPCorrelation(Fe_x,Fe_x,ADP)
    ADPCorrelation(Ni_x,Ni_x,ADP)
    #... same for y and z
  ]
#...
#
  [(1,0,0)
    SubstitutionalCorrelation(FeNi,FeNi,0.25*(1+a200))

    SizeEffect(Fe,Fe_x,se200_FeFe)
    #IMPORTANT: in addition to size effect add ADP correlation
    #with amplitude se^2/2
    ADPCorrelation(Fe_x,Fe_x,pow(se200_FeFe,2)/2)

    SizeEffect(Fe,Ni_x,se200_FeNi)
    ADPCorrelation(Fe_x,Ni_x,pow(se200_FeNi,2)/2)

    SizeEffect(Ni,Fe_x,se200_FeNi)
    ADPCorrelation(Ni_x,Fe_x,pow(se200_FeNi,2)/2)
    SizeEffect(Ni,Ni_x,se200_NiNi)
    ADPCorrelation(Ni_x,Ni_x,pow(se200_NiNi,2)/2)
  ]
]

```

Formal definition

In the presence of size-effect:

$$\bar{u}_i^{\alpha k, \beta l} = \sum_n M_{\beta l i}^n A_{\beta n} - \sum_n M_{\alpha k i}^n A_{\alpha n}$$

For description of the notation see [ADPCorrelation](#)

Epilogue

In the last part of the file the custom output from Yell can be requested. The output will be produced

after the refinement is finished. The output can contain any expressions whose values will be calculated from the refined variables.

Print

optional.

Arguments: a list of strings in quotes " " or arithmetic expressions

Format: `Print "string1" expr1 ...`

Prints the requested output in the terminal.

Example:

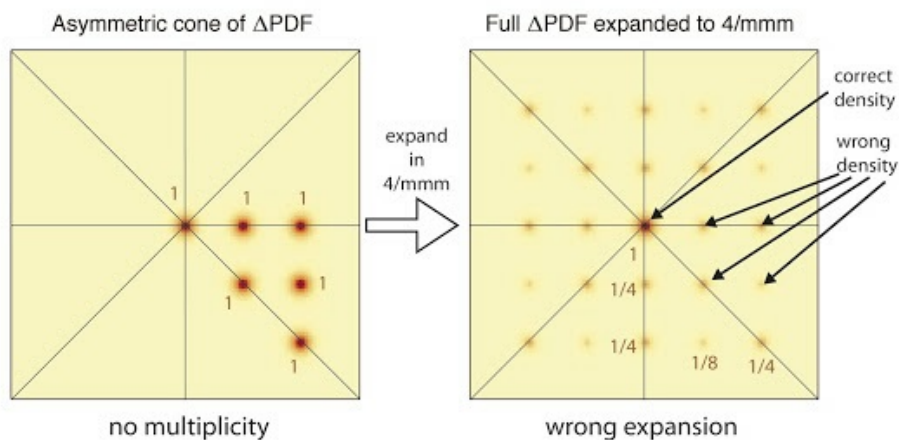
```
a=1;b=2;  
Print "The variable a=" a ", b=" b
```

This line produces the output `Requested output:The variable a=1, b=2`

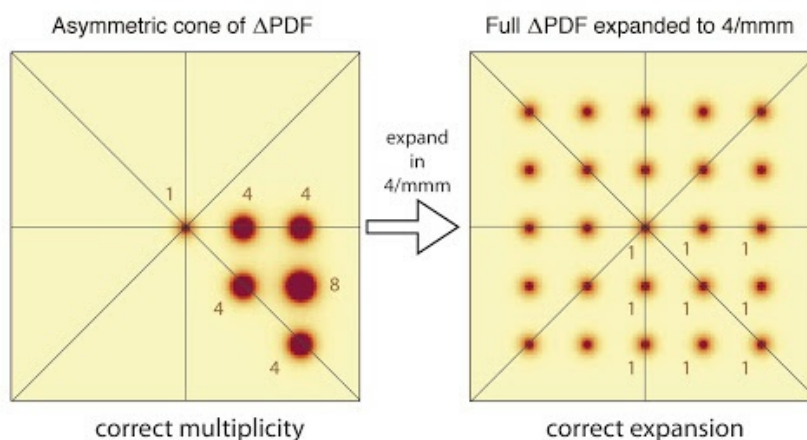
How to calculate multiplicity

Yell can apply the Laue symmetry to the calculated Δ PDF. This is a convenient feature since only the [independent cone](#) of correlations in Δ PDF space have to be defined.

Yell does not distinguish the interatomic pairs on special and average positions. Thus, the multiplicity of each interatomic (intermolecular) pair **must be provided manually**. If multiplicity is not provided, all the pairs except in the center of Δ PDF will get too little density:



If the multiplicity is provided, the Laue symmetry is applied properly:



Definition

By *multiplicity* of an interatomic (intermolecular) pair we mean the number of interatomic (intermolecular) pairs which are symmetry related to the current pair.

By symmetry we mean two distinct symmetries: the crystal space group symmetry and [combinatorial symmetry](#) which relates pair (A,B) to (B,A).

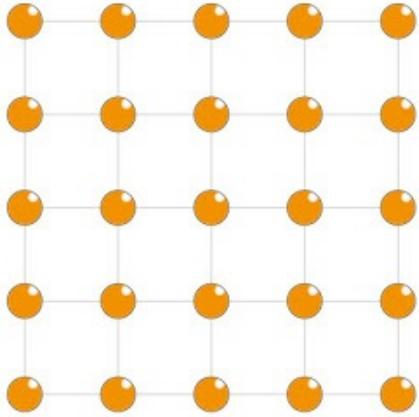
First way to determine multiplicity

Multiplicity can be determined by counting the symmetry equivalent pairs.

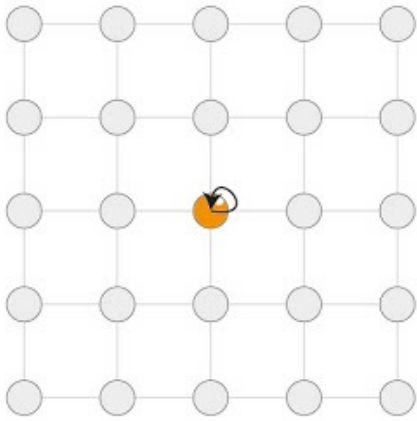
Example

This example is performed in two dimensions, but generalization to three dimensions is straightforward.

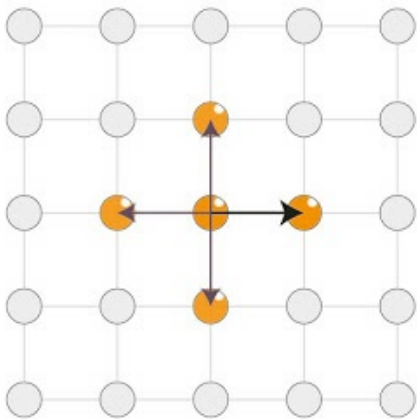
Assume a simple two-dimensional square crystal with a space group $p4mm$ and one atom in the unit cell:



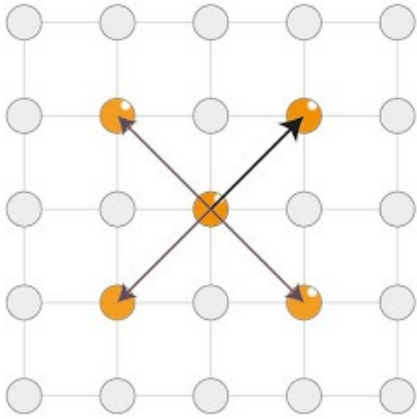
The zeroth neighbor connects an atom with itself. There is 1 such pair:



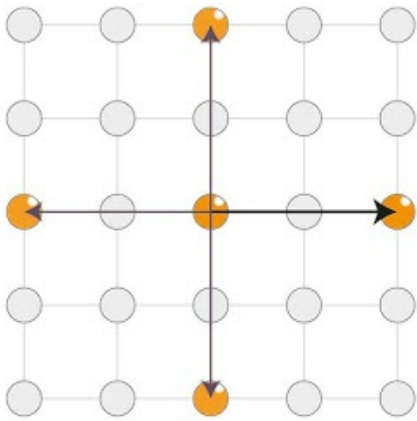
The first neighbor with interatomic vector $(1,0)$ has 4 symmetry equivalents:



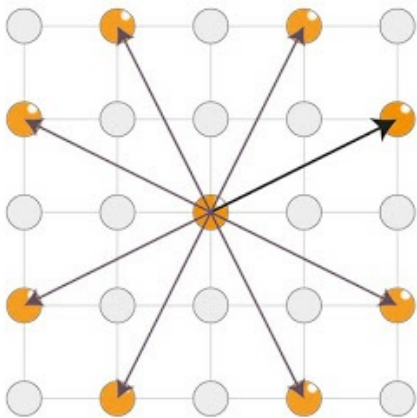
The neighbor $(1,1)$ also has 4 symmetry equivalents:



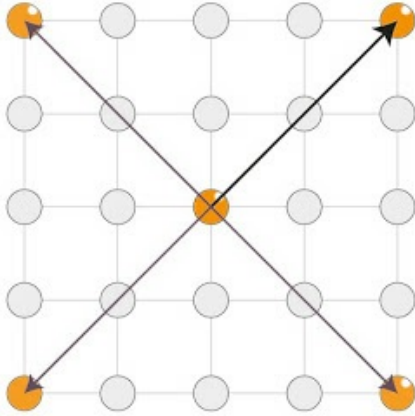
also the neighbor (2,0):



The neighbor (2,1) has 8 equivalents:



The neighbor (2,2) has again 4 equivalents:



The list of all pairs (without actual correlations) in Yell format will be the following:

```
Correlations [
  [(0,0,0)
  Multiplicity 1
  ...]

  [(1,0,0)
  Multiplicity 4
  ...]

  [(1,1,0)
  Multiplicity 4
  ...]

  [(2,0,0)
  Multiplicity 4
  ...]

  [(2,1,0)
  Multiplicity 8
  ...]

  [(2,2,0)
  Multiplicity 4
  ...]
]
```

Second way to determine multiplicity

Sometimes it is complicated to count the number of symmetry related pairs or there is a demand to calculate the multiplicity by an alternative approach to verify the results. In such cases, the multiplicity can be calculated from the internal symmetry of a single pair.

Assume:

- p is a pair
- $|G|$ is the number of symmetry elements in the crystal space group (space group order)
- $|Int(p)|$ is the number of symmetry elements in the internal symmetry of a pair (internal symmetry order)

Then, number of symmetry equivalent pairs is equal to:

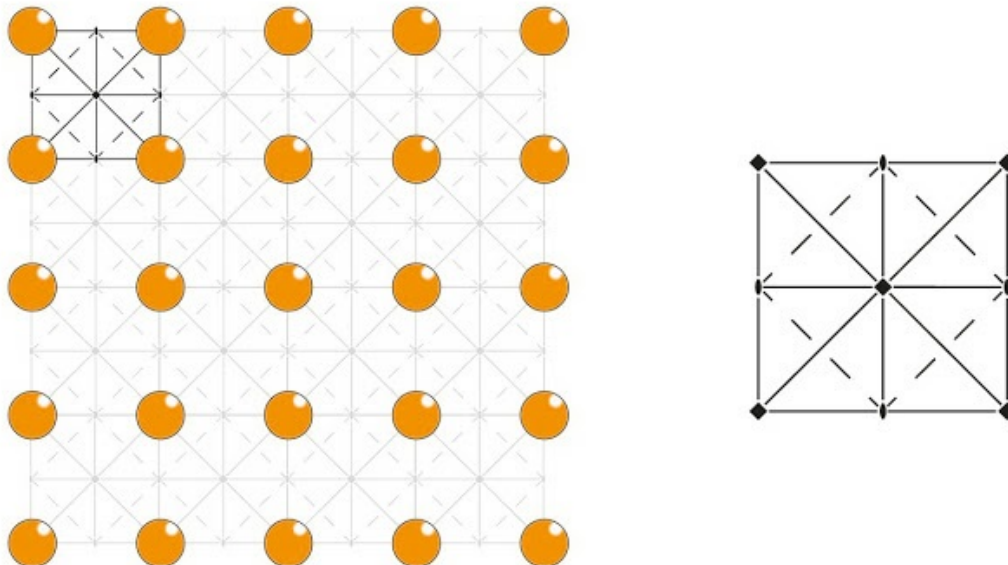
$$N_p = 2 \frac{|G|}{|Int(p)|} \text{ if the pair connects different atoms (molecules) and}$$

$$N_p = \frac{|G|}{|Int(p)|} \text{ for zeroth neighbor.}$$

The above mentioned formula is a consequence of the [orbit-stabilizer theorem](#). The factor 2 appears in the formula due to [combinatorial symmetry](#).

Example

Again, assume a simple square crystal. The crystal has a plane group $p4mm$ (No. 11 in International Tables of Crystallography):



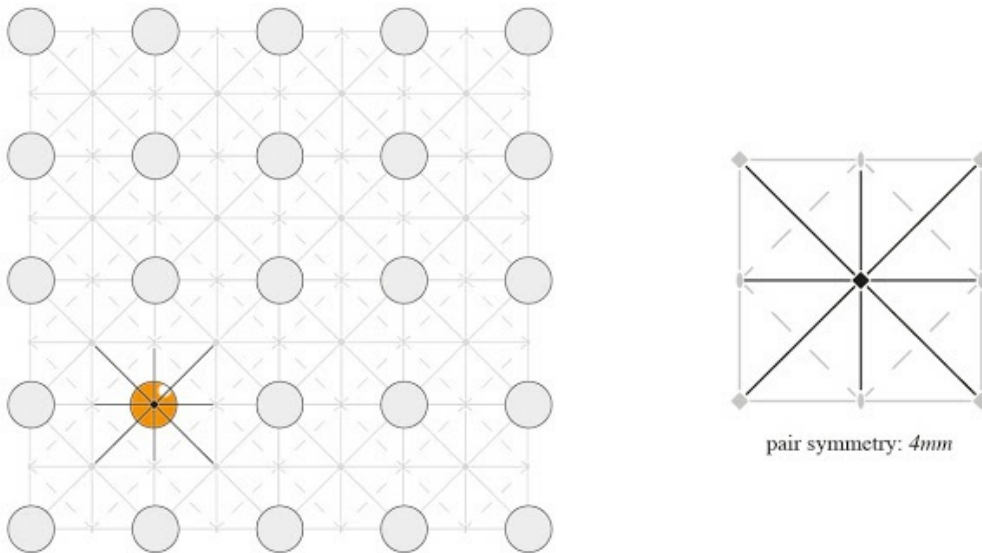
The plane group $p4mm$ has following symmetry operations:

$$(1) x, y \quad (2) \bar{x}, \bar{y} \quad (3) \bar{y}, x \quad (4) y, \bar{x}$$

$$(5) \bar{x}, y \quad (6) x, \bar{y} \quad (7) y, x \quad (8) \bar{y}, \bar{x}$$

Totally there are 8 operations, thus $|G| = 8$.

The zeroth neighbor has internal symmetry $4mm$:

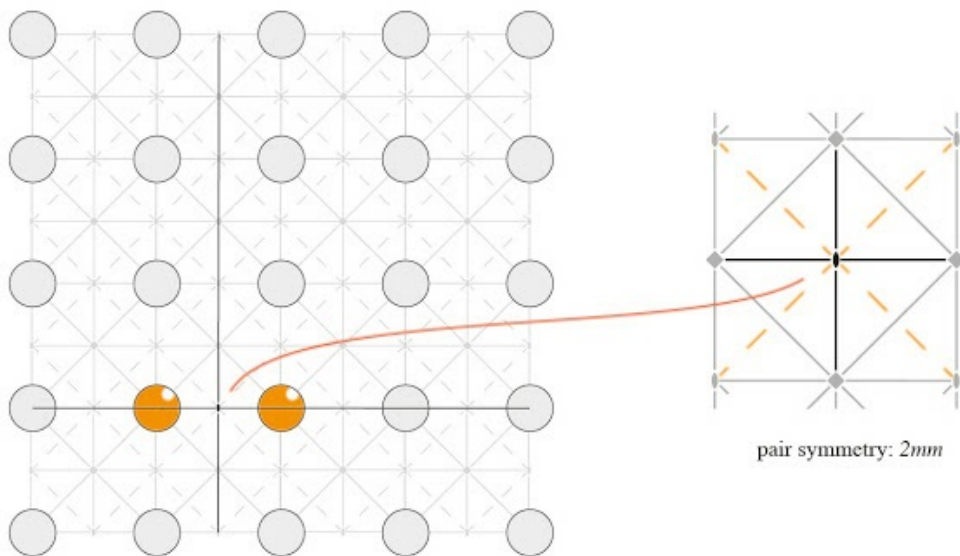


The order of internal symmetry group $4mm$ is 8 (group elements: $1, 4^1, 4^2, 4^3, m_x, m_y, m_{xy}, m_{yx}$); this pair is a zeroth neighbor:

$$N_{(0,0)} = \frac{|G|}{|Int(p_{(0,0)})|} = \frac{8}{8} = 1$$

Hence there is 1 such pair.

The neighbor (1,0) has internal symmetry $2mm$:



On the image the glide planes which are marked orange belong to the space group of the crystal, but

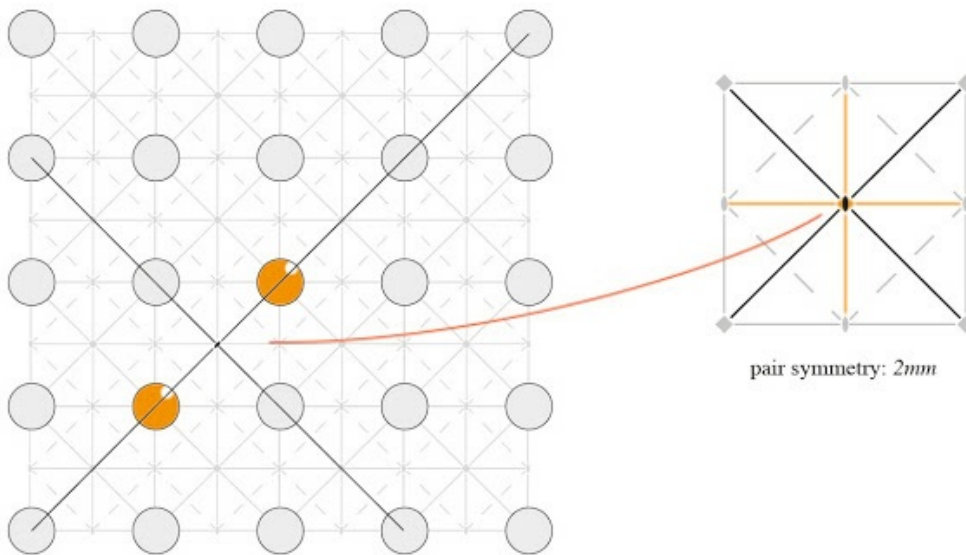
not to the internal symmetry of the pair.

The order of $2mm$ is 4 (group elements: $1, 2, m_x, m_y$), thus the number of such pairs is:

$$N_{(1,0)} = 2 \frac{|G|}{|Int(p_{(1,0)})|} = 2 \frac{8}{4} = 4$$

Note that glide plane operations do not count in this case, because the internal symmetry of the pair has to be described exclusively by point symmetry operations. In the case of periodic layers or rods, which are disordered along the other dimensions (corresponding to diffuse streaks or layers, respectively), glide or screw operations might also be considered along the periodic directions (see corresponding layer or rod groups).

The neighbor (1,1) also has internal symmetry $2mm$

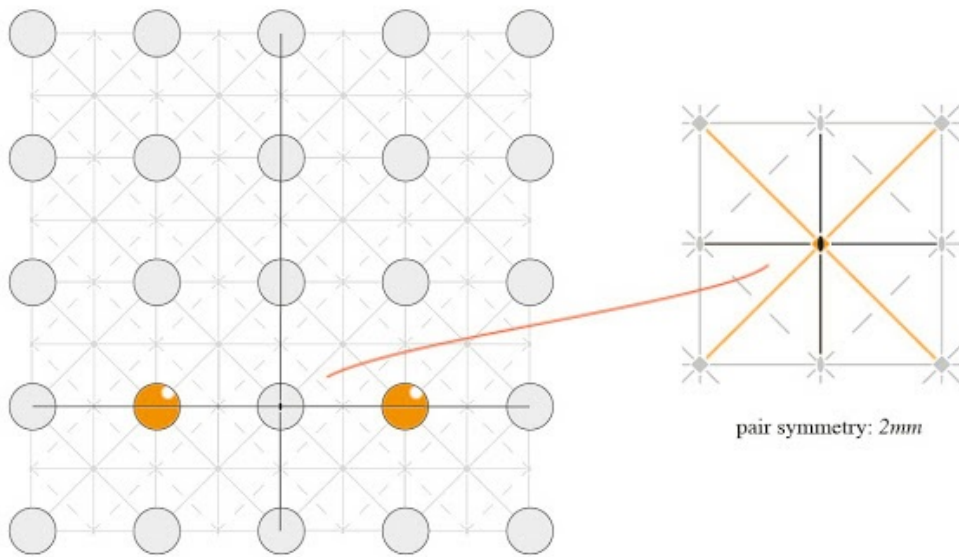


The average crystal has two additional planes and a four fold axis (marked orange) passing through coordinates (0.5,0.5), which do not completely belong to the internal symmetry of the pair. They are not counted. The remaining group elements are: $1, 4^2 \equiv 2, m_{xy}, m_{yx}$

The number of (1,1) pairs is therefore:

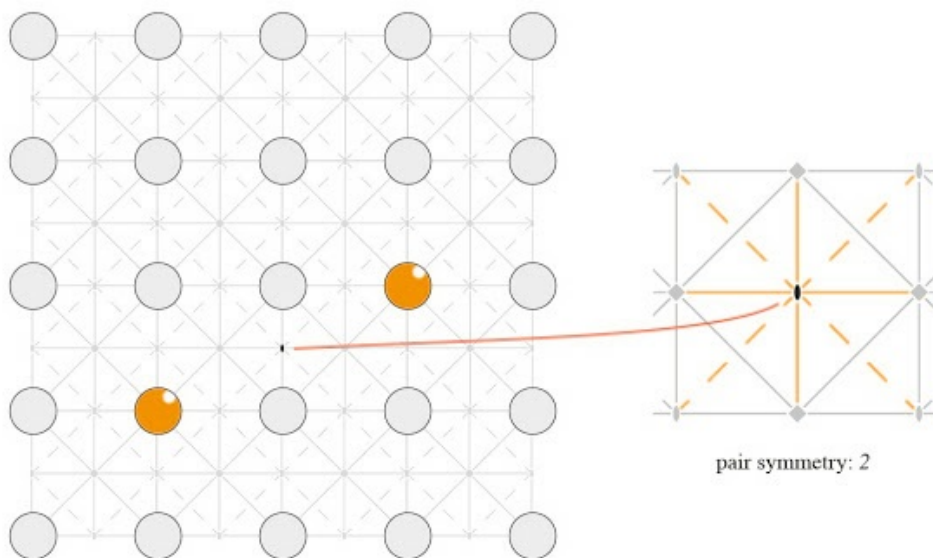
$$N_{(1,1)} = 2 \frac{|G|}{|Int(p_{(1,1)})|} = 2 \frac{8}{4} = 4$$

Neighbor (2,0) again has the symmetry $2mm$



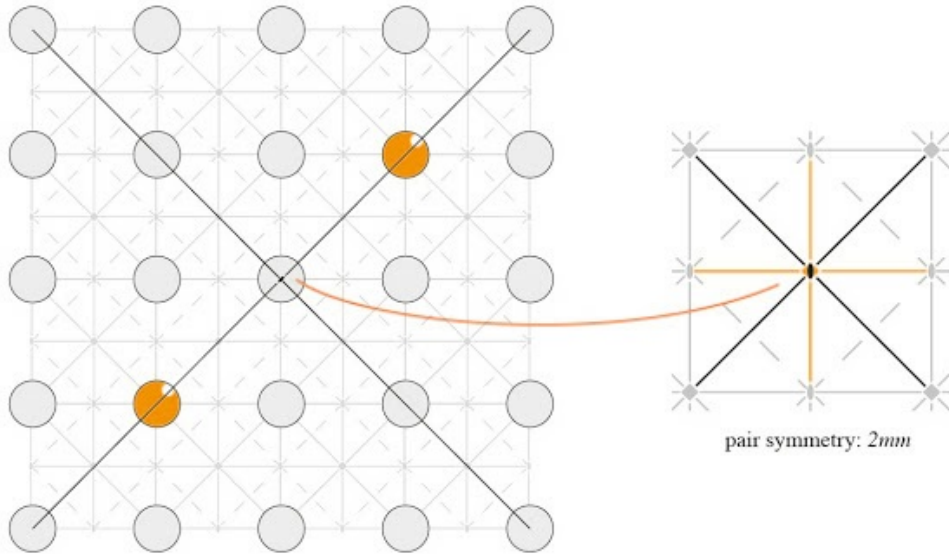
and $N_{(2,0)} = 2 * 8 / 4 = 4$

Neighbor (2,1) has the symmetry 2 (group elements: 1, 2):



Thus the number of such pairs is $N_{(2,0)} = 2 * 8 / 2 = 8$

Neighbor (2,2) has symmetry $2mm$ (group elements: 1, 2, m_{xy} , m_{yx}):



And thus there are 4 such pairs.

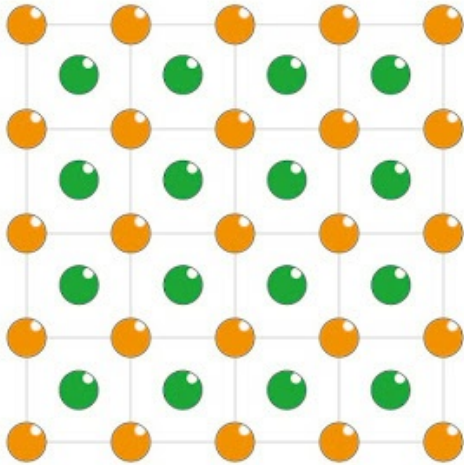
From symmetry analysis it is clearly seen:

- The zeroth neighbor has symmetry $4mm$ and only 1 equivalent.
- Neighbors $(x, 0)$ and (x, x) have symmetry $2mm$ and thus 4 equivalents.
- All the other neighbors (x, y) have symmetry 2 and thus 8 equivalents.

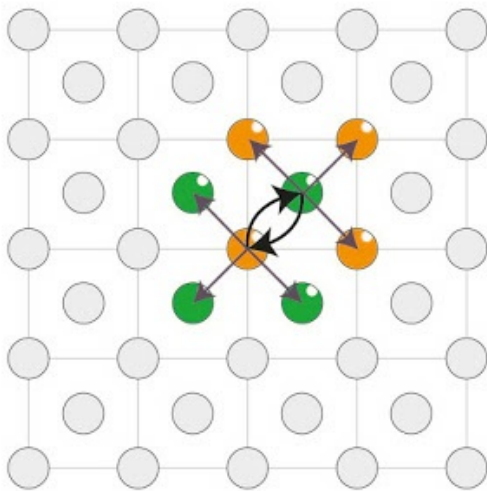
Combinatorial symmetry

By *combinatorial equivalent* of an (ordered) pair (A,B) we mean the pair (B,A). In PDF, combinatorial pairs are obtained by inverting interatomic vectors.

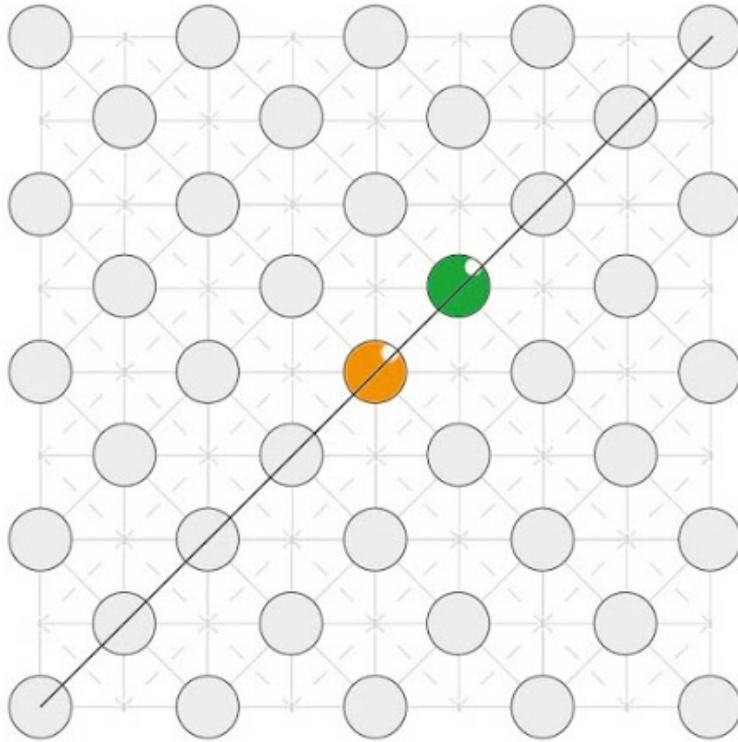
Assume we have a planar NaCl-type structure:



The pair Na-Cl with interatomic vector $(0.5, 0.5)$ has a combinatorially equivalent pair Cl-Na with vector $(-0.5, -0.5)$. Thus the multiplicity of such pair is 8:



When the multiplicity is calculated using [the second method](#), combinatorially symmetric pairs are automatically counted. The pair Na-Cl $(0.5, 0.5)$ has the internal symmetry m :



The multiplicity is equal to $N_{NaCl} = 2 * 8/2 = 8$

Independent cones for all Laue groups

The following table summarizes the independent parts of each Laue group:

Laue group	Group order	Independent cone conditions
$m\bar{3}m$	48	$x \geq y \geq z \geq 0$
$m\bar{3}$	24	$x \geq z, y \geq z, z \geq 0$
$6/mmm$	24	$x \geq 2y \geq 0, z \geq 0$
$6/m$	12	$x \geq y \geq 0, z \geq 0$
$\bar{3}m : H$	12	$x \geq y \geq 0, z \geq 0$
$\bar{3}m : R$	12	$z \geq y \geq x, x + y + z \geq 0$
$\bar{3} : H$	6	$x \geq 0, y \geq 0, z \geq 0$
$\bar{3} : R$	6	$x \geq y, x \geq z, x + y + z \geq 0$
$4/mmm$	16	$z \geq 0, x \geq y \geq 0$
$4/m$	8	$x \geq 0, y \geq 0, z \geq 0$
mmm	8	$x \geq 0, y \geq 0, z \geq 0$
$2/m$	4	$z \geq 0, y \geq 0$
$2/m : b$	4	$z \geq 0, y \geq 0$
$\bar{1}$	2	$z \geq 0$

Equation for diffuse scattering calculation

Diffuse scattering of a disordered crystal can be calculated by the equation (8) from [Weber & Simonov 2012](#):

$$I_{dif}(\mathbf{h}) = \sum_{\mathbf{R}_{uvw}} \sum_{mn}^{cryst\ cell} \{ p_{mn}^{uvw} \exp(-\mathbf{h}^T \beta_{uvw}^{mn} \mathbf{h}) \cos[2\pi \mathbf{h}(\mathbf{R}_{uvw} + \mathbf{r}_{mn} + \bar{\mathbf{u}}_{uvw}^{mn})] - c_m c_n \exp(-\mathbf{h}^T (\beta_m^{aver} + \beta_n^{aver}) \mathbf{h}) \cos[2\pi \mathbf{h}(\mathbf{R}_{uvw} + \mathbf{r}_{mn})] \} \cdot f_m(\mathbf{h}) f_n(\mathbf{h})$$

Here indices m and n go over all the atoms in a unit cell, \mathbf{R}_{uvw} over all lattice vectors; p_{mn}^{uvw} is a joint probability to find atom n in a unit cell and atom m in another unit cell separated by \mathbf{R}_{uvw} , β_{uvw}^{mn} is a joint ADP matrix expressed in fractional units, $\mathbf{r}_{mn} = \mathbf{r}_n - \mathbf{r}_m$ is the average interatomic vector between atoms m and n , $\bar{\mathbf{u}}_{uvw}^{mn}$ is a size effect parameter, c_m and c_n are the average occupancies, and β_m^{aver} and β_n^{aver} are the average ADP parameters, $f_m(\mathbf{h})$ and $f_n(\mathbf{h})$ are the atomic form-factors.

References

Weber, T., & Simonov, A. (2012). The three-dimensional pair distribution function analysis of disordered single crystals: basic concepts. *Zeitschrift für Kristallographie*, 227(5), 238-247. [researchgate](#)

Schomaker, V., & Trueblood, K. N. (1968). On the rigid-body motion of molecules in crystals. *Acta Crystallographica Section B: Structural Crystallography and Crystal Chemistry*, 24(1), 63-76. [link](#)

Cyvin, S. J. (1968). *Molecular vibrations and mean square amplitudes*. Universitets Forl.. Chapter 3

Simonov, A., Weber, T., Steurer, W., Yell - a computer program for diffuse scattering analysis via 3D- Δ PDF refinement. in. prep.

Bibliography

- [1] A. Bosak, D. Chernyshov, S. Vakhrushev, and M. Krisch, “A new model of correlated disorder in relaxor ferroelectrics,” *arXiv preprint arXiv:1101.0490*, 2011.
- [2] A. Roushoun, Y. Masatomo, M. Yoshitaka, Y. Hideki, O. Kenji, and I. Fujio, “Diffusion path of oxide ions in an apatite-type ionic conductor $\text{La}_{9.69}(\text{Si}_{5.70}\text{Mg}_{0.30})\text{O}_{26.24}$,” *Chem. Mater.*, vol. 20, no. 16, pp. 5203–5208, 2008.
- [3] C. Adams, J. Lynn, Y. Mukovskii, A. Arsenov, and D. Shulyatev, “Charge ordering and polaron formation in the magnetoresistive oxide $\text{La}_{0.7}\text{Ca}_{0.3}\text{MnO}_3$,” *Phys. Rev. Lett.*, vol. 85, no. 18, p. 3954, 2000.
- [4] P. Simoncic and T. Armbruster, “Peculiarity and defect structure of the natural and synthetic zeolite mordenite: a single-crystal x-ray study,” *Am. Mineral.*, vol. 89, no. 2-3, pp. 421–431, 2004.
- [5] B. Wehinger, D. Chernyshov, M. Krisch, S. Bulat, V. Ezhov, and A. Bosak, “Diffuse scattering in ice Ih,” *arXiv preprint arXiv:1402.2159*, 2014.
- [6] T. Weber, H. Boysen, and F. Frey, “Longitudinal positional ordering of n-alkane molecules in urea inclusion compounds,” *Acta. Crystallogr. Section B: Structural Science*, vol. 56, no. 1, pp. 132–141, 2000.

- [7] A. Letoublon, M. De Boissieu, M. Boudard, L. Mancini, J. Gastaldi, B. Hennion, R. Caudron, and R. Bellissent, “Phason elastic constants of the icosahedral Al-Pd-Mn phase derived from diffuse scattering measurements,” *Phil. Mag. Lett.*, vol. 81, pp. 273–283, Apr. 2001.
- [8] P. Faure, A. Micu, D. Perahia, J. Doucet, J. Smith, and J. Benoit, “Correlated intramolecular motions and diffuse x-ray scattering in lysozyme,” *Nat. Struct. Mol. Biol.*, vol. 1, pp. 124–128, 1994.
- [9] T. Proffen and S. J. L. Billinge, “PDFFIT, a program for full profile structural refinement of the atomic pair distribution function,” *J. Appl. Crystallogr.*, vol. 32, pp. 572–575, June 1999.
- [10] T. Weber and A. Simonov, “The three-dimensional pair distribution function analysis of disordered single crystals: basic concepts,” *Z. Kristallogr.*, vol. 227, pp. 238–247, May 2012.
- [11] A. Simonov, T. Weber, and W. Steurer, “Yell - a computer program for diffuse scattering analysis via 3D- Δ PDF refinement,” *Journal of Applied Crystallography*, accepted.
- [12] J. Cowley, *Diffraction Physics*. North-Holland, Amsterdam, 1975.
- [13] W. Zachariasen, “A general theory of x-ray diffraction in crystals,” *Acta. Crystallogr.*, vol. 23, no. 4, pp. 558–564, 1967.
- [14] T. Welberry, T. Proffen, and M. Bown, “Analysis of single-crystal diffuse x-ray scattering via automatic refinement of a monte carlo model,” *Acta. Crystallogr. Section A: Foundations of Crystallography*, vol. 54, no. 5, pp. 661–674, 1998.
- [15] H. Bürgi, J. Hauser, T. Weber, and R. Neder, “Supramolecular architecture in a disordered perhydrotriphenylene inclusion compound from diffuse x-ray diffraction data,” *Cryst. Growth. Des.*, vol. 5, no. 6, pp. 2073–2083, 2005.
- [16] T. Michels-Clark, V. Lynch, C. Hoffmann, J. Hauser, T. Weber, R. Harrison, and H. Burgi, “Analyzing diffuse scattering with supercomputers,” *J. Appl. Crystallogr.*, vol. 46, no. 6, pp. 1616–1625, 2013.

- [17] T. Proffen and R. B. Neder, "DISCUS : a program for diffuse scattering and defect-structure simulation," *J. Appl. Crystallogr.*, vol. 30, pp. 171–175, Apr. 1997.
- [18] M. G. Tucker, D. A. Keen, M. T. Dove, A. L. Goodwin, and Q. Hui, "RMCPProfile: reverse Monte Carlo for polycrystalline materials," *J. Phys.*, vol. 19, p. 335218, Aug. 2007.
- [19] G. Evrard and L. Pusztai, "Reverse Monte Carlo modelling of the structure of disordered materials with RMC++: a new implementation of the algorithm in C++," *J. Phys.*, vol. 17, pp. S1–S13, Feb. 2005.
- [20] R. Xu and T. C. Chiang, "Determination of phonon dispersion relations by X-ray thermal diffuse scattering," *Z. Kristallogr.*, vol. 220, pp. 1009–1016, Jan. 2005.
- [21] M. Holt, Z. Wu, H. Hong, P. Zschack, P. Jemian, J. Tischler, H. Chen, and T.-C. Chiang, "Determination of phonon dispersions from X-ray transmission scattering: The example of silicon," *Phys. Rev. Lett.*, vol. 83, p. 3317, Apr. 1999.
- [22] H.-B. Bürgi, M. Hostettler, H. Birkedal, and D. Schwarzenbach, "Stacking disorder: the hexagonal polymorph of tris(bicyclo[2.1.1]hexeno)benzene and related examples," *Z. Kristallogr.*, vol. 220, no. 12, pp. 1066–1075, 2005.
- [23] M. Treacy, J. Newsam, and M. Deem, "A general recursion method for calculating diffracted intensities from crystals containing planar faults," *P. Roy. Soc. Lond. A. Mat.*, vol. 433, no. 1889, pp. 499–520, 1991.
- [24] M. Leoni, A. F. Gualtieri, and N. Roveri, "Simultaneous refinement of structure and microstructure of layered materials," *J. Appl. Crystallogr.*, vol. 37, pp. 166–173, Jan. 2004.
- [25] M. A. Krivoglaz, *Theory of X-ray and thermal-neutron scattering by real crystals*. Plenum Press, 1969.
- [26] B. E. Warren, *X-ray Diffraction*. Courier Dover Publications, 1969.

- [27] B. Schönfeld, C. Sax, and A. Ruban, “Atomic ordering in Au-(42 to 50) at.% Pd: A diffuse scattering and first-principles investigation,” *Phys. Rev. B.*, vol. 85, p. 014204, Jan. 2012.
- [28] T. Proffen and H. Kim, “Advances in total scattering analysis,” *J. Mater. Chem.*, vol. 19, no. 29, p. 5078, 2009.
- [29] X. Jiang, G. Ice, C. Sparks, L. Robertson, and P. Zschack, “Local atomic order and individual pair displacements of $\text{Fe}_{46.5}\text{Ni}_{53.5}$ and $\text{Fe}_{22.5}\text{Ni}_{77.5}$ from diffuse x-ray scattering studies,” *Phys. Rev. B.*, vol. 54, no. 5, pp. 3211–3226, 1996.
- [30] J. Dshemuchadse, S. Bigler, A. Simonov, T. Weber, and W. Steurer, “A new complex intermetallic phase in the system Al-Cu-Ta with familiar clusters and packing principles,” *Acta Crystallogr. B.*, vol. 69, pp. 238–48, June 2013.
- [31] W. Kabsch, “Xds,” *Acta Crystallogr. D.*, vol. 66, pp. 125–32, Feb. 2010.
- [32] M. Estermann and W. Steurer, “Diffuse scattering data acquisition techniques,” *Phase Transit.*, vol. 67, pp. 165–195, Oct. 1998.
- [33] J. Dshemuchadse, D. Y. Jung, and W. Steurer, “Structural building principles of complex face-centered cubic intermetallics,” *Acta Crystallogr. B.*, vol. 67, pp. 269–92, Aug. 2011.
- [34] P. Urban, A. Simonov, T. Weber, and O. Oeckler, “Real structure of $\text{Ge}_4\text{Bi}_2\text{Te}_7$ refinement on diffuse scattering data with the 3D- Δ PDF method,” in prep.
- [35] R. Blessing, “Outlier Treatment in Data Merging,” *J. Appl. Crystallogr.*, vol. 30, pp. 421–426, Aug. 1997.
- [36] O. Karpinskii, L. Shelimova, M. Kretova, and G. Lubman, “X-ray Diffraction Study of the Layered Semiconductor $\text{Ge}_{3\pm\delta}\text{Bi}_2\text{Te}_6$,” *Inorg. Mater.*, vol. 30, no. 12, pp. 1406–1411, 1994.
- [37] T. Matsunaga, R. Kojima, N. Yamada, K. Kifune, Y. Kubota, and M. Takata, “Structures of stable and metastable $\text{Ge}_2\text{Bi}_2\text{Te}_5$, an intermetallic compound in a GeTe-Bi₂Te₃ pseudobinary system,” *Acta Crystallogr. B.*, vol. 63, pp. 346–52, June 2007.

- [38] E. S. Božin, C. D. Malliakas, P. Souvatzis, T. Proffen, N. Spaldin, M. G. Kanatzidis, and S. J. L. Billinge, “Entropically stabilized local dipole formation in lead chalcogenides,” *Science*, vol. 330, pp. 1660–3, Dec. 2010.
- [39] T. Keiber, F. Bridges, and B. C. Sales, “Lead is not off center in PbTe: The importance of r-space phase information in extended X-ray absorption fine structure spectroscopy,” *Phys. Rev. Lett.*, vol. 111, p. 095504, Aug. 2013.
- [40] O. Delaire, J. Ma, K. Marty, A. F. May, M. A. McGuire, M. Du, D. J. Singh, A. Podlesnyak, G. Ehlers, M. D. Lumsden, and B. C. Sales, “Giant Anharmonic Phonon Scattering in PbTe,” vol. 37831, 2011.
- [41] H. Leite Alves, A. Neto, L. Scolfaro, T. Myers, and P. Borges, “Lattice contribution to the high dielectric constant of PbTe,” *Phys. Rev. B.*, vol. 87, p. 115204, Mar. 2013.
- [42] T. Weber, A. Simonov, C. D. Malliakas, and M. G. Kanatzidis in prep.
- [43] T. Weber and A. Simonov, “The three-dimensional pair distribution function analysis of disordered single crystals: basic concepts,” *Z. Kristallogr.*, vol. 227, no. 5, pp. 238–247, 2012.
- [44] G. M. Sheldrick, “A short history of shelx,” *Acta. Cryst. A.*, vol. 64, no. 1, pp. 112–122, 2007.
- [45] V. Petříček, M. Dušek, and L. Palatinus, “Jana2006,” *The crystallographic computing system*, 2006.
- [46] M. R. Hestenes and E. Stiefel, *Methods of conjugate gradients for solving linear systems*, vol. 49. NBS, 1952.

Acknowledgments

First I want to thank Prof. Walter Steurer for supervising this project.

I would like to thank Dr. Thomas Weber for all his experience I could learn from and the freedom I enjoyed in deciding the direction of this research.

My very special thanks go to Dr. Dmitry Logvinovich for his tremendous help in writing numerous papers and manuscripts, very fruitful and interesting conversation about our work and for the faith in the success of the 3D- Δ PDF project. Dima, your support was priceless.

I would like to thank Philippe Urban, Prof. Oliver Oeckler, Dr. Julia Dshemuchadse for the fruitful collaboration and Dr. Dmitry Chernyshov for sharing diffuse scattering datasets.

Also thanks to all members of the LfK for the pleasant atmosphere during the last years.

I am very grateful to Prof. Reinhard Neder and Prof. Bernd Schönfeld for agreeing to become co-examiners.

This work was supported by the Swiss National Science Foundation.

Curriculum Vitae

Personal Details

Surname	Simonov
First Name	Arkadiy
Date of Birth	September 15, 1986
Sitizen of	Russia
Place of Birth	Moscow (Russia)
Martial status	Single

Education

03/2009-05/2014	Doctoral studies in Materials Science at ETH Zurich, Department of Materials PhD thesis: <i>The Three-Dimensional Pair Distribution Function method for analyzing single crystal diffuse scattering - theory, software development and application</i>
09/2003-01/2009	Physics studies, Lomonosov Moscow State University Diploma thesis: <i>Investigation of phase transitions in fast ionic conductors of the family $K_nH_m(SO_4)_{(n+m)/2}$</i>
09/1993-06/2003	Secondary school, Moscow, Russia