This is to certify that the
dissertation entitled

Computational Advances in Nanostructure Determination

presented by

Christopher Lyn Farrow

has been accepted towards fulfillment
of the requirements for the

Doctoral degree in Physics

Pm Duxbury
Major Professor's Signature

8/22/07
Date

COMPUTATIONAL ADVANCES IN NANOSTRUCTURE DETERMINATION

By

Christopher Lyn Farrow

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Physics and Astronomy

2007

ABSTRACT

COMPUTATIONAL ADVANCES IN NANOSTRUCTURE DETERMINATION

By

Christopher Lyn Farrow

The atomic pair distribution function (PDF) and extended x-ray absorption fine structure (EXAFS) techniques fill a hole in conventional crystallographic analysis, which resolves the average long-range structure of a material but inadequately determines deviations from the average. These techniques provide structural information on the sub-nanometer scale and are helping characterize modern materials. Despite their successes, PDF and EXAFS often fall short of adequately describing complex nanostructured materials. Parallel PDF and EXAFS refinement, or corefinement, is one attempt at extending the applicability of these techniques. Corefinement combines the best parts of PDF and EXAFS, the chemical-specific and short-range detail of EXAFS and the short and intermediate-range information from the PDF. New *ab initio* methods are also being employed to find structures from the PDF. These techniques use the bond length information encoded in the PDF to assemble structures without a model. On another front, new software has been developed to introduce the PDF method to a larger community. Broad awareness of the PDF technique will help drive its future development.

ACKNOWLEDGMENTS

I would like thank Professors Phil Duxbury and Simon Billinge for challenging me with hard problems and helping me see them through. I hope my future work reflects positively on your guidance.

Thanks go to Emil Božin and Pavol Juhas, who have been indispensable resources for PDF and programming knowledge. Thanks to Jiwu Liu, who has helped me become a better programmer, and is a constant reminder that I don't know as much about computers as I would like to. I would also like to thank Neil Aaronson for helping me get through my first year of graduate school.

I would like to acknowledge Michigan State University Department of Physics and Astronomy, Michigan State University College of Natural Sciences, the United States Department of Energy and the National Science Foundation for support throughout my graduate education.

Most importantly, I would like to give love and thanks to my family, who always make me feel like a success. Thank you, Dad, for always giving all that you could.

To my precious wife, Heidi. Thank you for being my best friend.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

Introduction

It takes scientists of all sorts to engineer, synthesize, probe and characterize a single chemical composition. Each of these disciplines utilizes specialized tools and techniques to perform its function. The focus of this thesis is the advancement of computational tools and techniques in the discipline of structure analysis.

The tools and techniques described herein are focused on two methods for local structure determination, the atomic pair distribution function (PDF) and the extended x-ray absorption fine-structure (EXAFS) method. Modern materials are complex and have crucial features on the nanometer scale. These materials are stretching the limits of traditional crystallographic techniques, and increasingly require local structure techniques for characterization. Crystallography, the analysis of diffraction patterns to specify the structure from which they result, requires long-range periodicity within a structure. It is not well suited for characterization of liquids and disordered solids. Even when long-range periodicity is present in a material, nanoscale deviations from the average periodic behavior are often missed.

Small structural features in a material are often associated with large changes in the material's properties. This is vividly demonstrated in the transition metal manganese perovskites [1]. These materials take on a variety of electronic and magnetic properties that can be tuned by intrinsic and extrinsic forces, such as chemical composition [2], temperature [3], applied magnetic field [4] and pressure [5]. The strong electron-lattice coupling in these materials closely binds their structure and function. These perovskites display "colossal magneto-resistance", a very large change in conductivity induced by an applied magnetic field [4]. These colossal changes in the physical properties are accompanied by local structural changes that challenge crystallographic detection. PDF [6] and EXAFS [7] have been pivotal in describing the nature of the local distortions in the perovskite structure that are induced by these forces. These techniques remain an important tool in the ongoing investigation of this system [8].

The minute alterations in structure that accompany functional changes in a material are easily overlooked in crystallographic analysis. This is the case with the $In_{1-x}Ga_xAs$ series of semiconductors. These materials are of interest because their conduction properties can be tuned by composition. The end members of this system, GaAs and InAs, each take on the zinc-blende structure; on average so do the alloys. The alloy closely follows Vegard's Law [9], that is, the lattice constant is the compositional average of that from the end members. Crystallography predicts a single nearest-neighbor bond length in the alloy that is commensurate with the lattice constant. However, local structure studies [10, 11] reveal that there are distinct Ga-As and In-As nearest-neighbor bond lengths that change with composition. The magnitude of this mismatch is shown in Figure 1.1.

Systems like these demonstrate the need for effective structural determination tools. The PDF and EXAFS methods are adequate tools for investigating nanostructure, but for various reasons, they do not always produce comparable solutions. This may be due to subtly erroneous data or insufficiencies of the techniques. The advancement of nanostructure determination requires new analysis strategies and software that can make the most out of these, and other techniques. One method

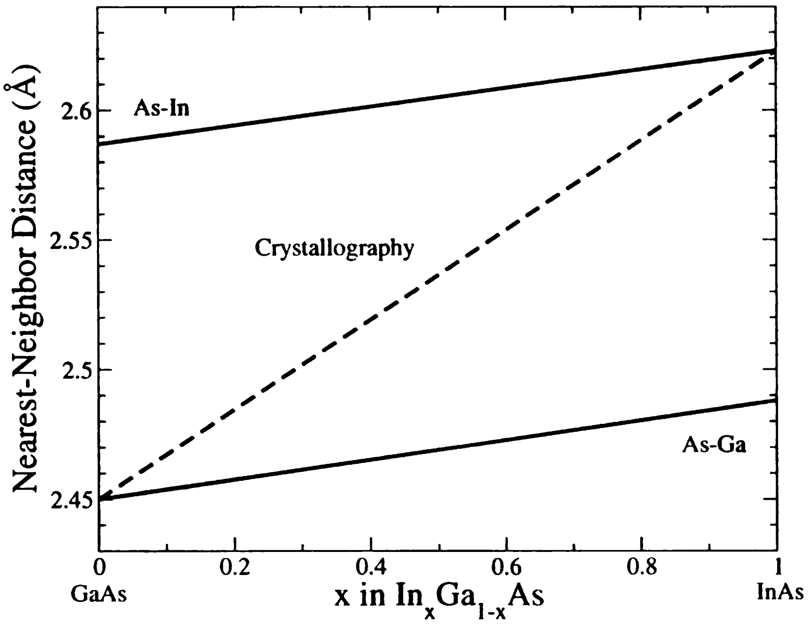


Figure 1.1: Nearest-neighbor distances measured by local-structure methods (top and bottom lines) and x-ray crystallography (dashed line) for the $In_xGa_{1-x}As$ series. Crystallography measures a single, average distance. This figure is reproduced from reference [10].

to assure consistency among structure solutions from different techniques is to derive a single solution from all the available data. Parallel refinement of experimental data, or corefinement, is presented Chapter 5. The application of this method to the $In_{1-x}Ga_xAs$ series is in Chapter 6. The specifics of the PDF and EXAFS techniques are given in Chapters 2 and 3. Further motivation for corefinement is given in Chapter 4.

Model-independent algorithms are another method used to construct a structure from experimental data. These methods are as detailed as the data they use, so PDF-based methods can achieve nanoscale resolution. The Liga algorithm [12] is one such algorithm. It uses the interatomic distances from PDF data to assemble a structure using geometrical principles. Inspired by Liga, Tribond is another distance-based algorithm that reconstructs randomly generated structures. These algorithms are described in Chapter 7.

Software is a vital tool in nanostructure investigations. Powerful software can remove much of the analysis burden from researchers and stimulate complex and detailed structural investigations. For years, the dominant tool for PDF analysis has been PDFFIT [13]. Recent collaborations, aimed at developing extensible and reusable diffraction analysis software, have lead to an update to this software. PDFFIT2 and PDFGUI [14] utilize modern software develpment techniques and represent the state of the art in PDF refinement. These programs are discussed in Chapter 8. Following this, Chapter 9 contains concluding remarks.

Chapter 2

The PDF method

2.1 Introduction

The PDF technique is a real-space, total-scattering diffraction technique for analyzing the local structure of materials. Traditional diffraction methods analyze the shape and location of Bragg peaks to reconstruct the three-dimensional atomic structure underlying a material. Such methods form the foundation of crystallography and have been widely successful. However, the existence of well defined Bragg peaks presupposes structural periodicity. Highly disordered, aperiodic, and "structurally challenged" [15] materials are therefore precluded from this form of analysis.

The PDF is the real-space Fourier transform of the structure factor, which is obtained from the coherent diffracted intensity measured in a standard x-ray or neutron scattering experiment. Whereas Bragg peaks in the diffraction pattern give information about the reciprocal lattice [16], the peaks in the PDF directly give the distance between atom pairs. In this way, the PDF gives the short- and intermediate-range structure of a material, and is complementary to the diffraction pattern. The PDF contains all of the information encoded within the diffraction pattern, the Bragg peaks and the diffuse scattering. Thus, PDF analysis is applicable to materials with few or diminished Bragg peaks.

The peaks in the PDF can be directly compared to the arrangement of atoms in a candidate structure. This forms the basis of the PDF technique. A candidate structure is used to calculate a theoretical PDF and this is then compared to the experimental PDF. The parameters which define the candidate structure can be adjusted in various ways to improve the agreement. The better the comparison between the experimental and theoretical PDF, the better the candidate structure. This procedure is very analogous to Rietveld fitting [17], and will be described in more detail in Section 2.4.

Before high-intensity sources were readily available, PDF analysis was applied almost exclusively to liquids and glasses [18]. The PDFs of such materials are relatively featureless beyond the first peak, so high resolution is not a strict prerequisite for investigation. In recent times, PDF analysis has been applied to many materials of scientific and technological interest, such as high-temperature superconducting cuprates [19, 20], and structurally related colossal magneto-resistant manganese perovskites [6, 21, 22, 23]. More recent PDF studies have investigated nanoparticle structure and function [24], crystals with nanostructured domains [25, 26, 27], saturated nanoporous materials [28, 29] and negative thermal expansion materials [30]. All of these materials have interesting and complex structural and functional relationships, and so are of interest to the EXAFS community as well.

The next section describes PDF data reduction. This is followed by a derivation and explanation of the theoretical PDF equation. PDF data analysis is described in the final section of this chapter.

2.2 PDF Data Reduction

In a diffraction experiment, a material sample is bombarded with an intense coherent x-ray or neutron beam. The wave interference of the scattered beam creates a diffraction pattern that is used to calculate the PDF. PDF and normal diffraction experiments differ slightly. PDF experiments must achieve a higher momentum transfer for high real-space resolution. These high momentum transfers (~ 40 Å$^-1$) are easily obtainable on modern high-intensity neutron and x-ray sources, such as the Advanced Photon Source and Intense Pulsed Neutron Source at Argonne National lab. PDF experiments are performed on powdered crystals, amorphous solids, gasses and liquids. What is measured in these experiments can be reduced to the intensity of scattering as a function of momentum transfer. There are standard methods in crystallography for getting the data into this form depending on how the intensity data was measured. This section describes how the PDF is obtained from the measured intensities.

The intensity measured in diffraction experiments is derived based on a few assumptions. Treating the incoming beam as a plane wave with wave number $\mathbf{k}$, the beam scattered from an atom at position $\mathbf{r_i}$ with wave number $\mathbf{k'}$ at an arbitrary point in space is given by

$$\Psi(\mathbf{r_i}) = e^{\mathrm{i}\mathbf{k'}\cdot\mathbf{r}} e^{\mathrm{i}(\mathbf{k'}-\mathbf{k})\cdot\mathbf{r_i}}. \tag{2.1}$$

This relationship is drawn in Figure 2.1. For simplicity, $\mathbf{k'} - \mathbf{k}$ is denoted by $\mathbf{Q}$, the momentum transfer. In the case of elastic scattering, when no energy is lost in the event, $|\mathbf{k}| = |\mathbf{k'}| = k$, $Q = 2k\sin(2\theta)$, where 2θ is the scattering angle between the incoming and outgoing waves. For elastic, single scattering, the intensity of the

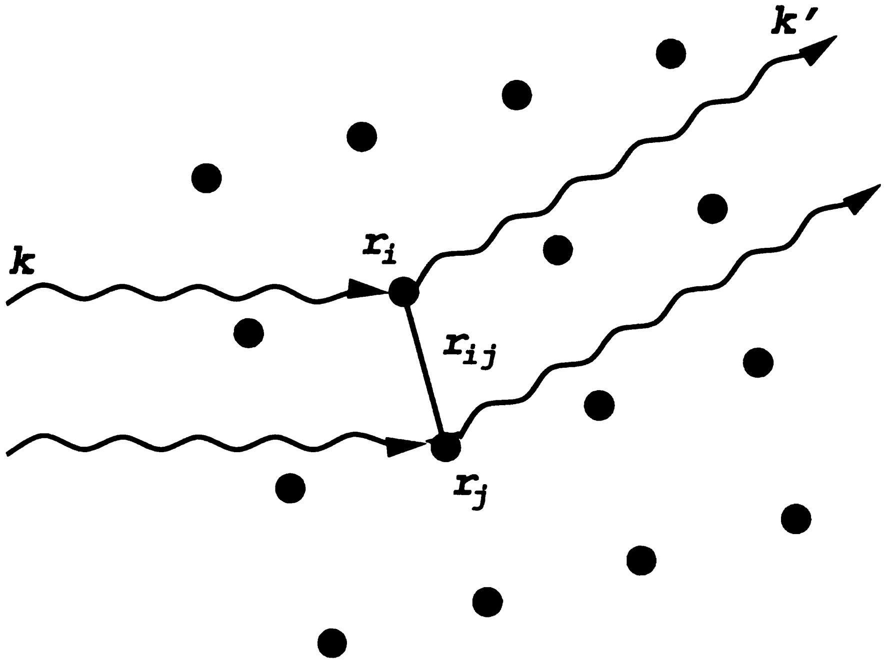


Figure 2.1: Diffraction of plane wave by a crystal.

scattered beam is given by

$$\begin{aligned} I(\mathbf{Q}) &= \left| \sum_i f_i(\mathbf{Q}) \Psi(\mathbf{r_i}) \right|^2 \\ &= \sum_{ij} f_i(\mathbf{Q}) f_j(\mathbf{Q}) e^{i\mathbf{Q}\cdot\mathbf{r_{ij}}}. \end{aligned} \tag{2.2}$$

Here, $f_i(Q)$ is the scattering amplitude from the atom at position i and $\mathbf{r_{ij}} = \mathbf{r_j} - \mathbf{r_i}$. This is the intensity from coherent scattering. Incoherent scattering is caused by multiple scattering and Compton scattering. Single scattering is almost assured by the high energy of the probing beam and through careful preparation of thin samples.

For a perfectly crystalline sample, the sum in Equation 2.2 is zero unless $\mathbf{Q} \cdot \mathbf{r_{ij}} = 2\pi l$, where l is an integer. In other words, Q must be a reciprocal lattice vector of the crystal lattice. When the lattice is not periodic on long length scales, there are no sharp diffraction peaks in the intensity. The structural information remains but is

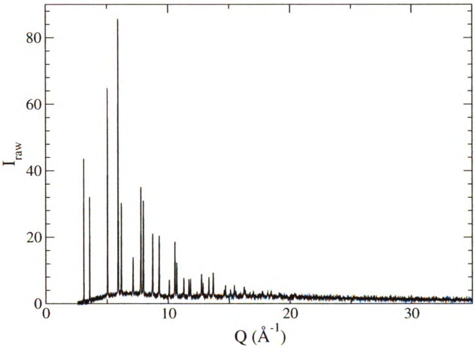


Figure 2.2: Raw neutron intensity data from fcc nickel at room temperature [31]. Data from a single detector is shown.

inaccessible without further processing. For powder or liquid diffraction the scatterers are uniformly isotropic. This effectively averages $I(\mathbf{Q})$ over all angles defined by $\mathbf{r_{ij}}$ and $\mathbf{Q}$. In this case, only the magnitude of these vectors is relevant, so the vector notation is dropped for the remainder of this chapter.

It is the structure and makeup of a material that determines its scattering cross section, and therefore its measured diffraction pattern. For an ideal periodic crystal, the peaks in the ideal diffraction pattern give the reciprocal lattice vectors of the underlying structure. In reality, the experiment is not perfect and the material being studied may not be crystalline. Experimental and structural effects obscure the peaks and cause them to broaden and flatten. The first step in data reduction is to remove the experimental effects from the scattering intensity. After that, the structural information can be extracted without any bias towards crystallinity. An example of raw neutron intensity data from fcc nickel is shown in Figure 2.2. This data was obtained from a single detector at room temperature [31].

The scattering intensity is reduced from the raw intensity by subtracting the effects of the ambient background and sample container. The background and sample container have their own diffraction signature that must be measured under the same conditions as the sample. The coherent scattering intensity is extracted from the raw intensity by applying theoretical corrections for multiple scattering and Compton scattering [18]. X-ray intensity data must also be corrected for polarization and absorption effects. The resulting intensity is reduced by a dispersion in the scattering factors and normalized. The structure factor is given by,

$$S(Q) = \frac{I_c(Q) - \sum_i c_i \left|f_i(Q)\right|^2}{\left|\sum_i c_i f_i(Q)\right|^2} + 1. \tag{2.3}$$

Here, $I_c(Q)$ is the normalized coherent scattering intensity. The sums run over atomic species in the material, c_i is the concentration of species i in the sample and $f_i(Q)$ the x-ray atomic form factor, which is tabulated in the crystallographic literature. In the case of neutron scattering, $f_i(Q)$ is replaced by the neutron scattering length, b_i, of atom i and the sum also runs over spin states. The structure factor is the structural information contained in the intensity pattern. The reduced structure factor, $Q(S(Q) - 1)$ is used to calculate the PDF. The reduced structure factor for fcc nickel at room temperature is shown in Figure 2.3.

The corrections mentioned above can be carried out without any adjustable parameters, but this procedure relies on approximations that can distort $S(Q)$. Once $S(Q)$ is obtained, another automated correction can be applied to force it to meet certain physical criteria [32]. This compensates for experimental effects and reduces

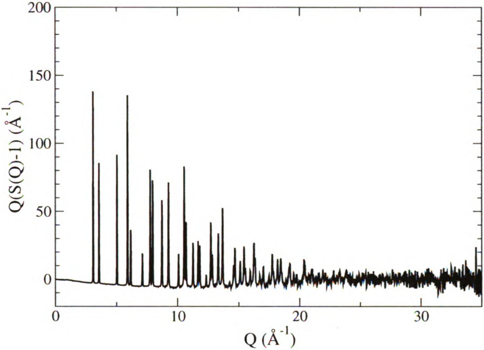


Figure 2.3: Scaled structure function for fcc nickel at room temperature from neutron intensity data [31]. The structure factor is scaled to match the integrand of Equation 2.4. This is data from a single detector, note the noise at high Q.

systematic error without subjective input from the scientist.

The PDF, denoted $G(r)$, is obtained from $S(Q)$ through a sine-Fourier transform of the reduced structure factor, which gives a real-space signal. The sine transform is appropriate for gasses, liquids, amorphous solids and powdered crystals, where the scatterers are isotropically oriented.

$$G(r) = \frac{2}{\pi}\int Q\,[S(Q)-1]\sin(Qr)\mathrm{d}Q. \tag{2.4}$$

The PDF of fcc nickel at room temperature is given in Figure 2.4. The figure shows a series of Gaussian-broadened peaks. The location of each peak relates directly to the

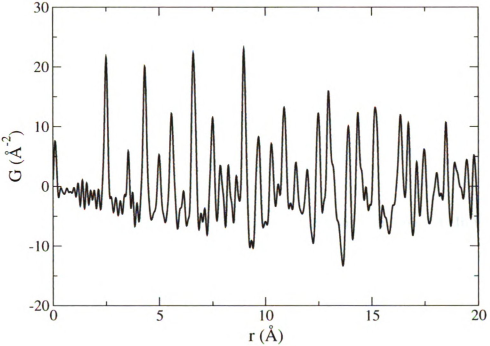


Figure 2.4: Pair distribution function of fcc nickel at room temperature from neutron diffraction data [31]. The peaks correspond to atomic pair distances within the structure. The small oscillations below 2.5 Å are due to experimental effects that were not corrected during data reduction.

distance between a given pair of atoms in the material and the height is proportional to the multiplicity, concentration, and scattering intensity of the pair. The broadening is due to the thermal and quantum-mechanical vibrations of the individual atoms as well as structural disorder.

2.3 The PDF Equation

The structure factor is derived from the single-scattering of plane waves from stationary point particles. It gives the correlation between atom pairs in a material as described in Equation 2.4. This equation does not take into account the thermal and quantum-mechanical zero-point vibrations within a material. This can be be treated with the dynamic structure factor $S(Q,\omega)$ [33, 16], which correlates the position of atoms over all time. The result of small vibrations is that the structure factor is modulated by a Gaussian envelope,

$$\begin{aligned} S(Q) &= e^{-2W} S_0(Q), \\ 2W &= \left\langle u^2 \right\rangle Q^2. \end{aligned} \tag{2.5}$$

Here, $S_0(Q)$ is the angle-averaged structure factor from Equation 2.4. In this equation, $\left\langle u^2 \right\rangle$ captures the atomic vibrations. It is the mean-square atomic displacement from equilibrium.

Carrying out the integration in Equation 2.4 with Equation 2.5 gives the PDF equation

$$\begin{aligned} G(r) &= 4\pi r \left[\rho(r) - \rho_0\right], \\ \rho(r) &= \frac{1}{4\pi r^2} \sum_{ij} A_{ij} \frac{1}{\sqrt{2\pi}\sigma_{ij}} \exp\left[-\frac{\left(r - r_{ij}\right)^2}{2\sigma_{ij}^2} \right], \\ A_{ij} &= \frac{c_i c_j f_i(0) f_j(0)}{\langle f \rangle^2}. \end{aligned} \tag{2.6}$$

Above, ρ_0 is the number density of the atomic structure. Each sum runs over all atoms: r_{ij} is the distance between atoms i and j, σ_{ij} is the correlated Debye-Waller factor and A_{ij} is the weighing factor of the atom pair. The weighing factor is dependent upon the concentration of each atom and its scattering amplitude at $Q = 0$. In the case of neutron scattering the scattering amplitudes are replaced by scattering lengths. The correlated Debye-Waller factor results from the Gaussian term in Equation 2.5. The correlated Debye-Waller factor and experimental corrections to

Equation 2.6 are treated in the next section.

2.4 PDF Analysis

Even though the PDF and structure factor of a material contain the same information, there are benefits to analyzing that information in real-space. The PDF uses the diffuse scattering component of the diffraction signal, which may not be adequately handled with reciprocal-space techniques. These contributions are spread out in reciprocal-space and can be the result of significant features in real-space. The indirect reconstruction of a material structure is necessitated by the *phase problem*. It is the nature of quantum-mechanical wave scattering that the scattering intensity is purely real (see Equation 2.2). The best one can do with the diffraction pattern is determine the relative distances between the scatterers. The solutions found are not guaranteed to be unique. This fundamental issue is not specific to PDF fitting. Given the information to work with, the PDF is its logical representation.

In order to fit Equation 2.6 to an experimental PDF, the equation parameters described above can be varied until a good fit is achieved. However, it is much more useful to relate these parameters to features of a three-dimensional structure so that the resulting set of fit parameters generate the underlying structure of the material. This is the approach taken by the PDF fitting program PDFFIT [13, 14]. In PDFFIT one or more sets of PDF data are simultaneously fit to a structural model consisting of one or more phases. A phase is a region that is structurally distinct from other such regions in the material. Physical parameters are constrained to guide the fitting toward meaningful and physically reasonable results. The three-dimensional structure used in PDF fitting is defined as a set of lattice parameters and a set of physical parameters for each atom: atomic coordinates, atomic occupancy, and anisotropic atomic displacement parameters. Additional parameters are used to

Table 2.1: Table of PDF fitting parameters.

Structural Parameters	
a, b, c, α, β, γ	lattice parameters
$x(i)$, $y(i)$, $z(i)$	lattice coordinates of atom i
$U11(i)$, $U22(i)$, $U33(i)$, $U12(i)$, $U13(i)$, $U23(i)$	anisotropic atomic displacement parameters of atom i
$occ(i)$	occupancy of atom i
Peak Shape Parameters	
δ	low temperature correlation term
ζ	high temperature correlation term
ϕ_0	static peak sharpening term
r_{cut}	cutoff applied to ϕ_0
σ_Q	resolution dampening term
α_Q	resolution peak broadening term
Scale Parameters	
Λ	Overall scale factor
λ	Weight of phase

account for experimental and dynamic effects. All PDF fitting parameters used in this investigation are displayed in Table 2.1. The notation used here closely follows that used in PDFFIT.

The parameters in Table 2.1 are used in calculating Equation 2.6. The lattice parameters and atomic coordinates define the inter-atomic distances. The occupancy defines the concentration of a particular atom species at a given atomic site. For example, two atomic species may occupy the same structural position in a substitutionally doped material. The peak shape parameters and the anisotropic displacement parameters are used to define the correlated Debye-Waller factor and to correct for experimental effects. The phase weight is used to define the concentration of a given phase in a multi-phase structural model. The overall scale factor scales the PDF fitting equation to the data. This factor is often necessary because the data reduction procedure cannot always correct for systematic errors in the data.

The correlated Debye-Waller factor found in the PDF equation is not the same as the Debye-Waller factor from normal scattering analysis. The PDF peaks represent pairs of atoms, and so the PDF Debye-Waller factor represents the mean-square variation in the distance between a pair of atoms. The motion of atoms in a material are usually correlated, especially at short distances. So, the correlated Debye-Waller factor is not simply the sum of the Debye-Waller factors of the individual atoms. The degree of vibrational correlation between atoms is well captured by a simple extension to the uncorrelated Debye-Waller factors of the atomic pair. The anisotropic atomic displacement parameters and atomic positions are used to compute the uncorrelated Debye-Waller factor,

$$\begin{aligned} \sigma_{ij}^{\prime 2} &= \sigma_i^2 + \sigma_j^2 \\ &= \hat{\boldsymbol{r}}_{ij} \cdot [\mathbb{U}(i) + \mathbb{U}(j)] \cdot \hat{\boldsymbol{r}}_{ij}. \end{aligned} \tag{2.7}$$

$\mathbb{U}(i)$ is the anisotropic displacement matrix for atom i, as defined by the anisotropic displacement parameters. The correlated Debye-Waller factor is a modification of this:

$$\begin{aligned} \sigma_{ij}^2 &= \sigma_{ij}^{\prime 2} \left(1 - \frac{\delta}{r_{ij}^2} - \frac{\zeta}{r_{ij}} + \alpha_Q^2 r_{ij}^2 \right) \phi(r_{ij}) \qquad (2.8) \\ \phi(r_{ij}) &= \begin{cases} \phi_0 & \text{for } r_{ij} < r_{cut} \\ 1 & \text{otherwise} \end{cases} \end{aligned}$$

where δ and ζ account for correlated atomic motion in the material. The $(1 - \delta/r_{ij}^2 - \zeta/r_{ij})$ term is derived from the correlated Debye model [34] with some simplifying approximations [35]. The correlated Debye model treats all phononic contributions equally without directional dependence. The phonons are limited by a finite maximum vibrational frequency, or equivalently the Debye temperature. The term quadratic in $1/r$ accounts for correlations at temperatures below the Debye temperature and the linear term accounts for correlations at high temperature. Due to the approximations of the correlated Debye model, much information is lost about the nature of the correlated motion, but it is sufficient for a wide range of materials [36]. For simplicity, the correlated Debye-Waller factor is usually referred to as just the Debye-Waller factor.

The α_Q and ϕ terms in Equation 2.8 are not part of the correlated Debye-Model. ϕ captures the effect of static disorder on the PDF. In materials with slight conformational disorder, the disorder will be apparent as multiple or shouldered peaks at low inter-atomic distances but will only broaden the peaks at larger distances. ϕ_0 simulates this by reducing the Debye-Waller factor at distances less than r_{cut} while allowing the displacement factors to capture the disordered behavior at larger distances. This term is included when static disorder is small and not included in the structural model.

Poor energy resolution can broaden or dampen the peaks in the PDF, and thus affect the determination of the anisotropic atomic displacement parameters [37]. α_Q from Equation 2.8 broadens the peaks and it simulates Q-dependent resolution loss. σ_Q simulates the dampening of the PDF peaks due to Q-independent instrument resolution. Either one or both of these terms can be applied during PDF fitting, depending on the experimental conditions [37]. There is also the effect of "ringing" in the PDF. As a result of the finite Q-range of any experiment, the integral in Equation 2.6 must be truncated at some Q_{max}. The resulting ripples can be as large as the noise in the PDF, so they are accounted for during PDF fitting. Truncating the integral in Equation 2.6 at Q_{max} is equivalent to convolving in real-space the ideal PDF of infinite Q-range with the sinc function $\sin(Q_{max}r)/r$.

The PDF fitting equation is given below with the experimental and Debye-Waller effects included.

$$G_{fit}(r) = \Lambda \frac{\sin(Q_{max} r)}{r} \exp(-\sigma_Q r^2) \sum_i \lambda_i G_i(r). \tag{2.9}$$

This equation includes the contribution of multiple phases, each weighted by its prevalence in the material, λ_i. Intra-phase contributions to the PDF are excluded. This is justified, provided that the extent of the fitting is small compared to the diameter of a phase domain or that the size and orientation of the different domains are uncorrelated.

Chapter 3

The EXAFS method

3.1 Introduction

X-ray absorption fine structure (XAFS), also referred to as x-ray absorption spectroscopy (XAS), is the minute modification of the x-ray absorption coefficient of a material due to the local environment of the absorbing atom [38, 39]. The x-ray absorption spectrum can be separated into two regimes. The x-ray absorption near-edge structure (XANES) is the oscillations in the absorption coefficient near ($\lesssim$ 30 eV) a core-level absorption edge. XANES is sensitive to the chemical bonding and oxidation state of the absorbing species. Extended x-ray absorption fine structure (EXAFS) [40, 41], extends approximately 1 keV beyond the XANES region. EXAFS is sensitive to the positions and species of atoms in the local environment of the absorbing atom.

EXAFS analysis can be performed in real or reciprocal space. The reciprocal space signal is a superposition of amplitude and phase modulated sine waves. The form of this modulation depends on the atoms in the neighborhood of the absorbing atom. When transformed to real space, this gives the positions of atoms in reference to a selected species in the sample. The absorbing atom is selected by the energy of the experiment. This chemical-specificity can only be obtained with other structural probes through expensive or difficult procedures, such as isotopic substitution for neutron PDF experiments [42, 43], anomalous x-ray scattering for x-ray PDF [44, 45] or with joint neutron and x-ray PDF studies [46].

With the intense x-rays provided at synchrotron light sources, EXAFS can probe the local environments of a wide range of elements to high resolution. Information about the first few nearest neighbors, conceptually organized into coordination shells around the absorber, is contained in the EXAFS signal. As a short-range probe, EXAFS is well suited for investigating liquids, glasses, disordered materials and nanostructures. Specialized experimental techniques make local structure accessible to EXAFS in unique and powerful ways. These will be discussed in the following sections.

The next section describes EXAFS data reduction. This is followed by a derivation and explanation of the theoretical EXAFS equation. EXAFS data analysis is described in the final section of this chapter.

3.2 EXAFS Data Reduction

Obtaining accurate information from the EXAFS requires careful and detailed processing. This section details the standard procedures for EXAFS data reduction [47]. The goal of data reduction is to isolate and normalize the oscillations in the absorption coefficient. Once this is obtained, it can be analyzed in real or reciprocal space.

EXAFS can be measured in a variety of experiments. In all of these experiments, a sample is placed in a high-intensity x-ray beam and some effect of the absorption is measured as a function of beam energy. The modulations in the absorption coefficient are caused by the atoms in the neighborhood of the photo-absorber. The electron that is emitted by the absorption event, the photo-electron, scatters from these neighboring atoms and alters the probability of absorption at the photo-absorber. The alteration

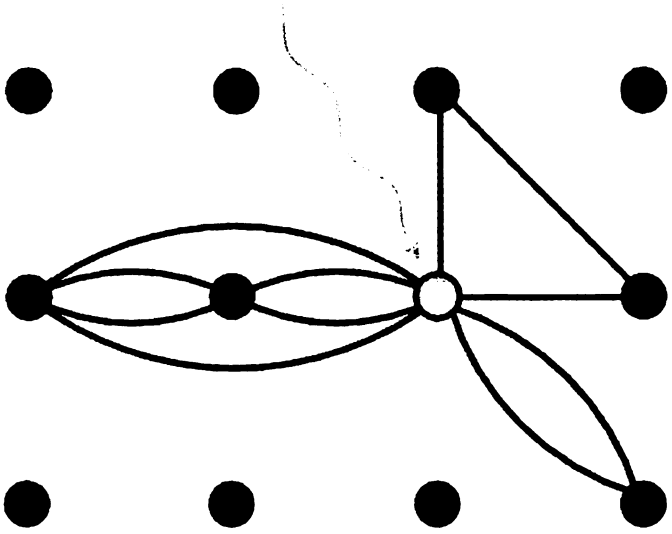


Figure 3.1: Possible paths taken by the photo-electron during x-ray absorption. The photo-absorber is indicated by the light circle. The photo-electron can scatter multiple times along any path.

specifically depends on how the photo-electron scatters around the material. Conceptually, the photo-electron follows a specific path that eventually leads back to the photo-absorber. This process is shown schematically in Figure 3.1. The relationship between this scattering and the EXAFS signal is described in the next section.

In a transmission-mode EXAFS experiment, the attenuation of the beam is measured directly by monitoring the intensity of the beam after it passes through the sample. The attenuation of the x-ray beam is due to the thickness and absorption coefficient, $\mu(E)$, of the sample:

$$\begin{aligned} I_f &= I_0 e^{-\mu x} \\ \mu x &= -\ln\left(\frac{I_0}{I_f}\right). \end{aligned} \tag{3.1}$$

Here, I_0 is the intensity of the incoming beam, I_f is the outgoing intensity and x is the sample thickness. As the beam energy approaches an atomic binding energy of a

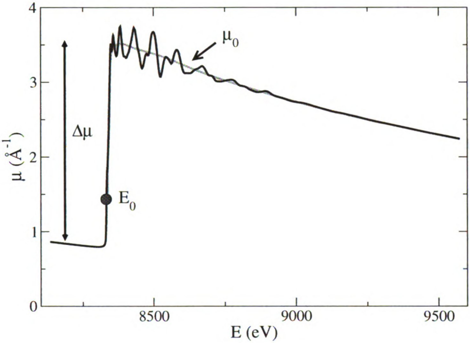


Figure 3.2: Absorption coefficient of fcc nickel at room temperature. The dark oscillatory line is $\mu(E)$. The isolated atom background function, μ_0, edge jump, $\Delta\mu$, and energy origin, E_0, are indicated.

given atomic species in the sample, the absorption coefficient shows a rapid increase and then a slow decay beyond the absorption edge. The absorption coefficient decays approximately as $\frac{\rho Z^4}{AE^3}$ after the absorption edge, where ρ is the material density, Z is the atomic number, A is the atomic mass and E is the x-ray energy. It is in this decaying region that the absorption fine structure is seen. A typical plot of μ against x-ray energy is shown in Figure 3.2.

The absorption of an x-ray photon by an atom leaves that atom in an excited state. The relaxation of the atom yields a photon, or less commonly multiple photons or electrons. In a fluorescence-mode EXAFS experiment, the intensity of the emitted photons is measured and indirectly related to the absorption coefficient. Electronyield EXAFS is similar, except that it measures the emission of Auger electrons. Relating the data from these indirect methods to the absorption coefficient requires more processing than for transmission-mode EXAFS. These pre-processing procedures are not presented here. However, once $\mu(E)$ is obtained, the data reduction is the same.

Data reduction begins by subtracting the embedded atom background, μ_0, from the absorption coefficient. The embedded atom background is the atomic absorption that would result if the photo-electron did not interact with these atoms. Unfortunately, this background cannot be calculated exactly, so it must be deduced from the data. This can be done by fitting the absorption coefficient with a smooth spline function [48]. The spline is tuned as to not add or remove oscillations from the signal. There is no way to know if the spline actually represents the background, only whether it meets this criterion. This step of the EXAFS data reduction is perhaps the hardest to control; beyond errors introduced by sample preparation and data collection [49], it is a major source of systematic error.

Once a smooth background is determined, it is subtracted from the absorption coefficient and normalized by the edge jump height, $\Delta\mu$ (see Figure 3.2). To obtain $\Delta\mu$, the pre-edge trend in the absorption coefficient is extrapolated to the extent of the signal, and the jump height is measured from the background to this line. $\Delta\mu$ is sometimes taken as the value of the jump at the edge, which is constant. The determination of $\Delta\mu$ is dependent on the quality of the data in the pre-edge region of the signal, or in the determination of the absorption edge; in either case it serves as a source of systematic error. The EXAFS signal, $\chi(k)$, is calculated from the absorption coefficient, background and edge jump, and is expressed in terms of k, the photo-electron wave number.

$$\chi(k) = \frac{\mu(k) - \mu_0(k)}{\Delta\mu(k)}. \tag{3.2}$$

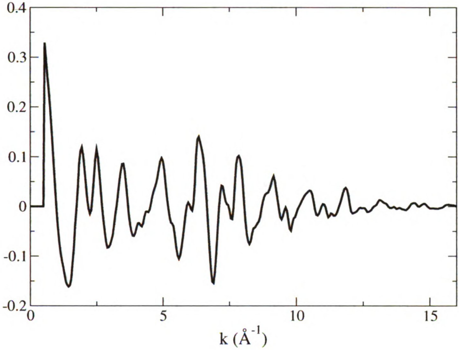


Figure 3.3: EXAFS signal from fcc nickel at room temperature.

$$k = \sqrt{\frac{2m_e(E - E_0)}{\hbar^2}}.$$

$E - E_0$ is the kinetic energy of the photo-electron after its ejection from the core level of the absorbing atom; E_0 is the binding energy of that level. The photo-electron wave number is a function of this energy and the electron mass, m_e. Atomic binding energies are well known, but the effects of neighboring atoms on atomic binding is dependent on local environment and must measured for each material. E_0 is often estimated as the point of inflection in the initial rise of the absorption coefficient or at some fraction of the initial edge jump (see Figure 3.2). It cannot be exactly determined during data reduction, and so is usually treated as an unknown during data analysis. $\chi(k)$ for fcc Ni at room temperature is shown in Figure 3.3.

After $\chi(k)$ is isolated, it can be Fourier transformed into real space.

$$\chi(R) = \int \chi(k)e^{\mathrm{i}2kR}\mathrm{d}k \tag{3.3}$$

The $2R$ in the exponent is there because the photo-electron makes a round-trip to the neighboring atoms. R is the half-path length. The Fourier transform in Equation 3.3 is complex. The amplitude of the resulting signal, $|\chi(R)|$, is a pair pseudo-distribution function. The amplitude, real and imaginary part of $\chi(R)$ for fcc nickel at room temperature is shown in Figure 3.4. The peaks in $|\chi(R)|$ roughly correspond the the location of coordination shells around the absorbing atom. The contribution to $\chi(k)$ from a specific scattering path is modulated by the scattering amplitude of the path. The Fourier transform of these contributions gives the peaks in $\chi(R)$. The first peak is from the first coordination shell. The second peak is due to the second coordination shell and multiple-scattering of the photo-electron between the first and second coordination shells. The scattering amplitude is complex and k-dependent, so the peaks in $\chi(R)$ are asymmetric and shifted to lower R than the atom-atom, or multi-atom distances they represent. The peaks have finite width due to both static and thermal disorder, that is, the Debye-Waller factor. Note the rapid decay in the signal with inter-atomic distance.

Once $\chi(k)$ has been extracted from the data, either it or $\chi(R)$ can be analyzed to determine the physical properties of the material. If the complex scattering amplitudes are known, the peaks in $\chi(R)$ can be used to extract atomic pair distances, coordination and Debye-Waller factors. This is done by fitting the EXAFS signal to a theoretical equation. A simple derivation of $\chi(k)$ is given in the next section. EXAFS analysis is described after that.

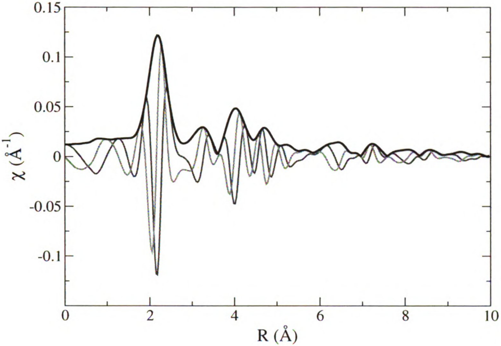


Figure 3.4: EXAFS signal from fcc nickel at room temperature. The thick, dark line is the amplitude of the real space signal, $|\chi(R)|$. The real and imaginary parts of the signal are given by the thinner dark and light lines, respectively. The amplitude at low R is due to long-wavelength oscillations that were either introduced or not removed during background subtraction.

3.3 EXAFS Theory

The physical process that leads to EXAFS is essentially the photo-electric effect. An energetic x-ray is absorbed by an atom, transferring enough energy to knock one of its electrons into an available state. Here it is assumed that the x-ray photon is absorbed completely and the photo-absorber only undergoes a single excitation. EXAFS is strongest when a core electron is excited, so only the excitation of the most tightly bound electron is considered here. This is EXAFS from the K-edge of the absorbing atom. EXAFS can be performed at a lower energy absorption edge, which corresponds to the excitation of less tightly bound electrons. The resulting EXAFS signal is approximately the same for the first few core levels [50]. The following derivation applies to transmission-mode EXAFS. Other experimental EXAFS techniques result in the same EXAFS equation, provided certain approximations hold. These are discussed further in the literature [49].

The x-ray absorption coefficient of a material is proportional to the rate at which the x-rays are absorbed, which is the rate at which the core electron excitation takes place. This is given by Fermi's Golden Rule,

$$\mu(E) \propto \frac{2\pi}{\hbar} \left| \langle i | \hat{H} | f \rangle \right|^2 \rho(E_f). \tag{3.4}$$

Here, $\langle i|$ is the initial state of the core level, $|f\rangle$ is the final state of the core-hole and photo-electron, $\hat{H}$ is the Hamiltonian operator and $\rho(E_f)$ is the density of final states. Implicit in this equation is the condition that $E_f = E - E_0$. That is, the kinetic energy of the photo-electron is the x-ray energy less the binding energy. It is assumed that the incoming photon has its electric field vector pointing along the **z** direction. The time-independent Hamiltonian can then be expressed as $\hat{H} = e\varepsilon\hat{z}$, where e is the electron charge and ε is the amplitude of the electric field. The time-independent Hamiltonian is valid provided that the electric field is constant over the range of interaction, which is generally the case [50]. This is referred to as the dipole approximation.

For tightly bound levels, the initial state is non-zero only in the vicinity of the core radius. The final state need only be considered in this region. For simplicity, this initial state is treated here as a Dirac delta function, $|i\rangle = \delta(\mathbf{r})$. Once ejected, the photo-electron radiates out from the photo-absorber as a spherical wave. As it spreads out it scatters from the neighboring atoms. This creates interference with the outgoing wave at the photo-absorber, which affects the final state and $\mu(E)$. As the x-ray energy changes, so does the wavelength of the photo-electron, thus altering the interference pattern. In essence, it is the photo-electron that probes the material.

$\chi(k)$ can be obtained from Equation 3.4 by expressing the final state as the sum of its outgoing and incoming parts. Some approximations and assumptions can simplify the derivation. First, it is assumed that the sample consists of randomly oriented crystals or amorphous regions and it has no preferred orientation. This assumption forgoes the need to keep track of the absolute trajectory of the photo-electron. Angular terms are omitted in the following derivation. Furthermore, all atoms are treated as points. This small atom approximation is valid in the regime of high photo-electron wave number and distant neighboring atoms [50]. The corrections to these approximations have received a thorough treatment, and can be found in the references [51, 52].

The outgoing wave is given by the Hankel function of the first kind, an irregular solution to the Schrödinger equation for angular momentum $l = 1$:

$$\Psi_o(kr) = h_1^+(kr) = \left(\frac{1}{kr} + \mathrm{i}\frac{1}{(kr)^2}\right) e^{\mathrm{i}kr}. \tag{3.5}$$

This form is valid in the vicinity of the origin. Outside the influence of the photo-absorber the outgoing wave acquires a phase shift due to the potential of the central atom [53]. This phase shift is denoted $\delta_c(k)+\pi/2$; the $\pi/2$ comes from the asymptotic form of the Hankel function. The incoming wave is determined by considering the interaction of the outgoing wave with the neighboring atoms. As mentioned above, EXAFS is due to the scattering of the ejected photo-electron from the atoms in its local environment. Single scattering is treated here, multiple-scattering is discussed later.

At a given scatterer, the scattered wave is subject to an amplitude change and phase shift [53]. This amplitude is denoted $f(k,\phi)e^{\mathrm{i}\delta(k)}$ for scattering angle ϕ. ϕ

is equal to π when the photo-electron is back-scattered, as it is during a single-scattering event. Note that $f(k, \phi)$ and $\delta(k)$ are specific to the species and location of the scattering atom, but the small atom approximation assures negligible distance dependence. The incoming wave can be put into the form

$$\Psi_i(kr) = F(kr)e^{i\delta_c(k)}. \tag{3.6}$$

This form is valid near the origin of the photo-absorber. The $\delta_c(k)$ phase shift is due the influence of the central atom potential on the incoming segment of the path. Near the scatterer, which is located at $\mathbf{R}$, $F(kr)$ is given by

$$F(kr) = h_1^+(kR)f(k,\pi)e^{i\delta(k)}\frac{e^{ik|\mathbf{r}-\mathbf{R}|}}{k\,|\mathbf{r}-\mathbf{R}|}e^{i(\delta_c(k)+\pi/2)}. \tag{3.7}$$

The outgoing central atom phase shift, $\delta_c(k) + \pi/2$, has been included explicitly. To put $F(kr)$ in a useful form, it must first be evaluated near the origin. Due to the dipole approximation, the $l = 1$ term must then be taken from the spherical harmonic expansion. The result of these operations gives

$$F(kr \to 0) = \Psi_o(kr)f(k,\pi)\frac{e^{2ikR}}{2k^2R^2}e^{i(\delta_c(k)+\pi/2)}. \tag{3.8}$$

Substituting this into $|f\rangle$ from Equation 3.4 gives

$$\mu(k) \propto \left| \int kr\ \delta(r)\Psi_o(kr)\left(1 + f(k,\pi)\frac{e^{2ikR}}{2k^2R^2}e^{i(2\delta_c(k)+\pi/2)}\right)\right|^2. \tag{3.9}$$

Writing this equation as $\mu_0(k)\,[1 + \chi(k)]$ reveals the relation of $\Psi_o(k)$ to the isolated atom background and $F(kr)$ to the EXAFS oscillations. The small atom approximation is equivalent to the condition that $kR \gg 1$. To lowest order in kR,

$$\begin{aligned}\chi(k) &= 2\Re(F(kr \to 0)e^{i\delta_c(k)}), \\ &= \frac{f(k,\phi)}{kR^2}\sin(2kR + \delta(k) + 2\delta_c(k)). \qquad (3.10)\end{aligned}$$

Equation 3.10 contains the basic features of an EXAFS signal. Modifications must be made to account for real EXAFS signals. In general, the photo-electron can scatter multiple times along various paths (see Figure 3.1). Each of these scattering paths contributes to the total EXAFS signal. The contribution from multiple-scattering paths can be cast in the same form as Equation 3.10. At every scatterer, the photo-electron is subject to an additional phase shift and altered amplitude, while the total path length will increase by the distance between the current and last scatterer. Thus, the form of Equation 3.10 is valid for multiple-scattering paths as well, where $f_P(k)$ and $\delta_P(k)$ represent the total amplitude change and phase shift for the path and R_P represents the half-length of the path. Note that $f_P(k)$ is dependent on the geometry of the scattering path. Both forward and back-scattering in a path can lead to a significant scattering amplitude, as can triangular scattering paths of large multiplicity [54]. A more thorough description of multiple-scattering EXAFS can be found in the references [52, 51, 55].

Other effects have been neglected in the derivation of Equation 3.10. The effects of inelastic scattering and the lifetime of the core-hole were not taken into account. The lifetime of the core-hole affects the transition rate in Equation 3.4. Inelastic scattering destroys the coherence of the photo-electron so that it causes no interference at the origin. These effects are considered phenomenologically in terms of a mean free path of the electron, $\lambda(k)$. This represents how far the photo-electron can travel in the material before these effects become significant. Related to the lifetime of the core-hole, the relaxation of the photo-absorber was not considered. The relaxation of the photo-absorber during the lifetime of the core-hole is accounted for by the passive electron amplitude reduction term, S_0^2.

Of more structural interest, thermal and static disorder in the material also affects the EXAFS signal. Both of these effects are captured in a series of peak broadening functions. Symmetrical peak broadening, as seen in Figure 3.4, is accounted for by applying a Gaussian envelope on each scattering path. Higher order, anharmonic terms account for asymmetrical broadening and are not detailed here. For more information, see reference [56]. The width of the Gaussian envelope term is the correlated Debye-Waller factor. This definition of the Debye-Waller factor includes the effects of small static disorder due to practical reasons; if the static disorder is not included explicitly during data analysis, it will be captured by in peak broadening term. The Gaussian peak-width comes from the second cumulant with respect to R of Equation 3.10. Proper determination of this cumulant includes curved-wave effects and introduces a small correction to the phase equal to $-4k\sigma^2/R$ [57].

Considering all of the effects above, the complete EXAFS equation is given by

$$\begin{aligned} \chi(k) &= S_0^2 \sum_P N_P \frac{f_P(k,\phi)}{kR_P^2} e^{-2k^2\sigma_P^2} e^{-2R_P/\lambda_P(k)} \\ &\quad \sin\left(2kR_P + 2\delta_c(k) + \delta_P(k) - \frac{4k\sigma_P^2}{R_P}\right). \end{aligned} \tag{3.11}$$

The sum runs over all scattering paths, P, taken by the photo-electron. N_P is the multiplicity of the path, which is defined by the geometry of the atomic arrangement. All other terms have been described above. Equation 3.11 is used during EXAFS data analysis, which is described in the next section.

3.4 EXAFS Analysis

EXAFS analysis does not necessarily involve EXAFS fitting. Qualitative information about a system can be obtained just by looking at the EXAFS signal. Quantitative analysis requires comparing an experimentally obtained EXAFS signal to an accepted mathematical representation of the signal, such as Equation 3.11. This section is concerned with quantitative analysis. Examples of other analysis methods can be found in the literature [38].

3.4.1 Limitations

Obtaining quantitative information from EXAFS data can be challenging. For a typical EXAFS investigation, one does not know the absorption energy precisely, and the amplitude of the EXAFS signal is prone to experimental effects [49]. These unknowns are highly correlated with parameters of more interest; the edge energy must be known to accurately measure the shell distances and the amplitude reduction factor must be known to determine the Debye-Waller factor and the coordination numbers. In turn, the coordination numbers and passive electron reduction factor are totally correlated. As a result, these unknowns must be provided by an outside source or be treated as variables during the analysis. Chemically and structurally similar materials can be used to obtain the amplitude reduction factor and edge energy, but this presumes the existence of a well known standard. Even then, the standard values are only approximate and subject to error if measured under different experimental conditions than the primary data. In some cases, special techniques can be used to decorrelate the amplitude factors [56, 58]. EXAFS signals obtained under different conditions, such as signals from different absorption edges or from different temperatures, can be used to constrain these parameters and reduce the correlations. Constraining parameters across multiple coordination shells can have the same effect.

In general, EXAFS analysis is even more difficult beyond the first coordination shell. Multiple scattering paths can become significant as soon as the second coordination shell is included. The scattering factors and phase shifts for multiple-scattering paths are not readily obtainable from standard materials, and so have to be computed theoretically. FEFF [59, 60] is one program for doing this. FEFF imparts its own limitations on EXAFS fitting. These are discussed in detail below. Properly constraining multiple-scattering paths can reduce correlations among the variables, but additional paths may introduce more fitting variables. As the fitting range increases, so does the likelihood that different paths will overlap in real space, which leads to more parameter intercorrelation. The number of parameters that can be extracted from an EXAFS signal is approximately $2 + 2\Delta R \Delta k/\pi$, where ΔR and Δk are the real and reciprocal space extent of the signal used in the fit, respectively [61]. When the number of significant scattering paths grows rapidly with the fitting range, as with disordered materials, they exhaust the information capacity of the signal without proper constraints. How these constraints are formulated varies from system to system.

3.4.2 EXAFS fitting

EXAFS analysis is commonly performed by adjusting the parameters in Equation 3.11 in order to fit an experimental EXAFS signal. The parameters that give the best fit are assumed to quantitatively describe the structure of the subject material. This procedure can be automated with an optimization algorithm that minimizes some measure of the discrepancy between the data and the fit. Minimization is discussed in Chapter 5. In order to use Equation 3.11 as a fitting function, f_P and δ_P, the path parameters, must be calculated for a every path needed by the fit. The range of the fit determines which paths are significant. Paths with weak signals are often omitted from the calculation.

The calculation of path parameters requires knowledge of the path, such as the atoms and distances involved. Of course, if that information were known entirely, there would be little need for data analysis. However, if the geometry of a path is known, the scattering angles and number of legs, then the amplitude factor and phase shift can be determined with only approximate knowledge of the path lengths. This information can be transferred to paths of the same geometry as long as the small atom approximation remains valid. If one wants to model a different cluster of atoms, then the amplitude and phase terms must be calculated again. This would not be a limitation except for the method by which path parameters are generated. The dominant programs for EXAFS fitting, such as IFEFFIT [62] and its graphical user interface ARTEMIS [63], use FEFF to calculate the path parameters. FEFF is not designed to be interfaced, but rather run as an application separate from the fitting procedure. The path parameters are generated beforehand, as opposed to at each iteration of the fitting algorithm. Because of this, the geometry of the scattering paths cannot be easily adjusted during EXAFS fitting. This is not a limitation in FEFF so much as it is a limitation of the path formalism. Allowing the path geometry to change during a fit would alter the path degeneracies unless symmetry were also preserved. New and removed paths would need to be handled by the fit routine, which greatly complicates the fitting.

Along with generating path parameters, FEFF calculates the central atom phase shift, $\delta_c(k)$, mean free path terms, $\lambda(k)$, and passive electron reduction factor, S_0^2, for a configuration of atoms. The $\lambda(k)$ from FEFF are considered good enough for EXAFS fitting, but the S_0^2 values are only approximate [59]. FEFF also has a path filter that eliminates paths with amplitude below a percentage of that of the most significant path. The path filter is critical since the number of scattering paths grows exponentially with the coordination shells.

EXAFS fitting performed in this Dissertation depends on FEFF paths and is

Table 3.1: Table of refinement parameters in EXAFS fitting.

A	The amplitude of the path
N	Degeneracy of the path
ΔE	The shift in the energy origin
ΔR	The shift in the half path-length
σ^2	The correlated Debye-Waller factor of the path
E_i	The imaginary energy shift (self-energy)
σ_3	The cubic anharmonic term
σ_4	The quartic anharmonic term

modeled after IFEFFIT. The EXAFS fitting parameters are summarized in Table 3.1. These terms exist for each path and relate to the parameters in Equation 3.11. They are adjusted by the fitting routine to find the best fit to Equation 3.11. A constraint mechanism exists for fixing, floating, or constraining these parameters to user-defined fitting variables. Physically meaningful relationships between the parameters can be captured in the constraints to reduce uncertainty and correlations among the fitting variables.

Some of the parameters in Table 3.1 deserve further discussion. The amplitude of the path, A, includes the passive electron reduction factor, S_0^2. It also compensates for systematic errors that manifest in the amplitude of the signal. The path degeneracy, N, is used primarily in first shell analysis. When the coordination of the first shell is not known, a single FEFF path can be fit with variable N. This has to be determined separately from the amplitude factor since they are entirely correlated. As mentioned above, the edge energy can only approximately determined during data reduction. ΔE represents the shift in the edge energy from that used during data reduction. It modifies E_0 in Equation 3.2, and thus alters the photo-electron wave number. Since FEFF generates the path parameters on a constant k-grid, these values must be interpolated to the altered k value during fitting. ΔR is the change in the path length from that used to generate the FEFF path. The self-energy term modifies the mean free path generated by FEFF. It compensates for the finite energy resolution of the experiment and for the fact that the core-level has a finite energy width. σ_3 and σ_4 are corrections to the symmetric peak broadening approximation. σ_3 manifests in the phase shift and σ_4 in the peak broadening term. These terms are used primarily in quantifying asymmetric static disorder.

Chapter 4

Comparison of PDF and EXAFS Techniques

4.1 Introduction

Despite the similarities between the PDF and real-space EXAFS signal, they do not, in practice, provide the same structural information. Both PDF and EXAFS techniques have practical limitations that affect the amount and quality of derived structural information. Some of these limitations are identified in Section 3.4. The way the structural information is obtained also differs between the techniques. PDF studies are usually interested in obtaining structural models of a material. Due the short range of the real-space EXAFS signal, it is often not possible to determine a complete structure from the data. Furthermore, there is no accepted way to relate the vibrations of individual atoms to the multiple-scattering Debye-Waller factors. When results from PDF and EXAFS analyses differ, it may to be due to noisy or insufficient data. However, the way in which the data is analyzed can also lead to differing results.

4.2 Limitations

Many perceived strengths and weaknesses of EXAFS and PDF are based on outdated information [64] and deserve some clarification. It has long been believed that EXAFS has better spatial resolution than conventional diffraction studies. This does not necessarily apply to PDF studies, although early PDF studies on liquids did not require high resolution. The smallest feature discernible in the PDF, the spatial resolution, is given by $\delta r_{PDF} \approx \pi/\Delta Q$, where ΔQ is the range of momentum transfer used to calculate the PDF. For EXAFS, this quantity is $\delta r_{EXAFS} \approx \pi/2\Delta k$, where Δk is the extent of the photo-electron wave number used to calculate the real-space EXAFS signal. EXAFS experiments are often performed out to 16 or 20 Å^{-1}, but sometimes less. Absorption edges on the high-energy side of the edge being investigated may reduce the maximum attainable k-range. PDF experiments can be performed up to and beyond $Q_{max} = 40$ Å^{-1}, for example at the neutron powder diffractometer NPD at the Manuel Lujan Neutron Scattering Center at Los Alamos National Laboratory [65], which gives comparable spatial resolution to EXAFS. It used to be that PDF data collection took much longer than EXAFS data collection, a comparison of hours as opposed to minutes. This is also no longer true. The RA-PDF method [66] uses high-intensity x-rays and an image plate detector to collect diffraction intensities suitable for PDF studies in a matter of seconds, albeit at Q_{max} less than 30 Å^{-1}.

Instrument resolution is a factor that affects the quality of PDF and EXAFS data. It modifies the shape of the peaks in the PDF, and can manifest in both the phase and amplitude of the EXAFS signal. Techniques for eliminating instrumental and intrinsic resolution effects are in current development for EXAFS [67, 68, 69, 70], but no single method has been widely adopted. Therefore, EXAFS data reduction and modeling software [62, 63] does not provide a means to compensate for experimental resolution. PDF software [13, 14] contains phenomenological terms [37] that account for certain types of resolution effects during modeling and generic experimental effects during data reduction [32]. However, recent work has shown that these resolution terms are not sufficient to compensate for effects from time-of-flight neutron diffraction experiments [71]. These time-of-flight resolution effects generally cause an overestimation of the PDF lattice constants and thermal parameters. This is usually compensated for by calibration with respect to a known standard. However, until resolution effects are adequately handled by PDF and EXAFS refinement software, the quantitative structural information from these techniques is less reliable than it could be.

A limitation that exists in PDF that does not apply to EXAFS is the limited accessibility of trace elements in a sample. Fluorescence-mode EXAFS detects the x-rays emitted by the absorbing atoms that accompany x-ray absorption instead of the attenuation of the x-ray beam. These are detected at angles outside of the x-ray beam, so that the weak signal is separated from the strong background intensity. Instruments can detect the fluoresced x-rays from atoms present at levels as low as 10 parts-per-million. Such a small concentration of atoms is invisible to PDF, as they contribute very little to PDF profile.

Both EXAFS and x-ray PDF studies may be hindered by limitations due to contrast. X-ray scattering intensities are dependent on atomic number, so elements that have atomic numbers within five of each other are difficult to distinguish with x-rays. In addition, light elements have weak x-ray absorption and scattering properties, so organic compounds are largely inaccessible to x-ray PDF and EXAFS. EXAFS is not completely blind to light atoms, though, as these atoms can be detected in multiple-scattering paths. Neutron PDF does not have these issues since neutron scattering power is not strongly influenced by atomic number. However, neutron PDF may suffer from peak cancellation due to the negative scattering length of some elements. This occurs in the nearest neighbor Cu-O/Ti-O peak of the neutron PDF of $CaCu_3Ti_4O_{12}$ due to the negative scattering length of Ti [72]. Even without cancellation, peak overlap is a common complication in PDF and EXAFS fitting. To the advantage of the PDF, the medium-range correlations help deconvolve the contribution of different atomic pairs under a single peak during analysis. However, EXAFS from multiple edges gives the best separation of this information. PDF techniques exist for separating the PDF into its contributions from different atoms, such as anomalous x-ray scattering [44, 45], isotopic substitution for neutron PDF [42, 43] and joint x-ray and neutron PDF [46]. Unfortunately, isotopic substitution is usually prohibitively expensive to perform and anomalous x-ray scattering is a very involved and difficult technique. Joint x-ray and neutron PDF is perhaps the most promising of these technique due to the ease with which it is performed and analyzed, but it is not widely used.

The most limiting feature of EXAFS is its short spatial range. Many interesting phenomena manifest in the medium-range structure, such as discommensurations in incommensurate charge density wave materials [73, 74] and structural phase transitions in transition metal manganese perovskites [75]. The limited distance information in the EXAFS signal precludes structure determination from EXAFS data alone for all but the simplest structures. Reverse Monte Carlo, which is described in the next chapter, has been used to model a complete structure from EXAFS data [76], but this approach has not been widely adopted. For this reason, approximations are often used to extract as much information from the EXAFS signal as possible. These methods are explored in the next section.

4.3 Structural Modeling

There are two general reasons why the structural information from PDF and EXAFS studies may differ. Foremost, errors introduced in the collection, reduction and processing of PDF and EXAFS data have an unpredictable influence on the derived results. Secondly, limitations of the technique, data or theory limit the modeling options and quality of the results. The second point is of critical interest in PDF and EXAFS corefinement.

The extent to which the structural information in the PDF and EXAFS signal is useful largely depends on how that information is modeled and extracted. For highly disordered or complex materials, the EXAFS and PDF signals are composed of many overlapping peaks. These peaks must be modeled as a distribution, since neither the EXAFS nor the PDF signal contains enough information to uniquely specify all the pair distances. For materials with less disorder, or even complex order, this information may still be modeled as a distribution at short distances if medium-range information is required to distinguish similar atom-sites. These medium-range correlations are present in the PDF, but not in the EXAFS. For these types of materials, comparison between EXAFS and PDF results is difficult because the PDF can make use of a structural model to capture the static disorder but the EXAFS may capture static disorder with the Debye-Waller factor. It is not always clear whether these systems have "weak" disorder, which is assumed in PDF modeling. Furthermore, failure to accommodate large disorder effects in the EXAFS equation can lead to improper measurements of the bond lengths [77].

Amorphous materials have received much attention from the EXAFS community. These materials often have skewed distributions that are poorly modeled as Gaussian disorder. The cumulant expansion approach has been developed to model the effects of asymmetric disorder on the EXAFS signal [56, 64, 78, 79, 80]. The cumulant expansion introduces corrections to the phase and amplitude of the EXAFS equation that model weak asymmetric disorder. Such corrections are not commonly applied to the PDF and are not available in PDF refinement software.

4.4 Study of $La_{1-x}Ca_xMnO_3$

To compare the compatibility of structural information from EXAFS and PDF techniques, a study was performed on the colossal magneto-resistance material $La_{1-x}Ca_xMnO_3$ [8]. $La_{1-x}Ca_xMnO_3$ undergoes a phase transition from a paramagnetic insulator (PI) to a ferromagnetic conductor (FC) as the temperature is reduced below about 260 K for $x \approx 0.30$ [7]. Accompanying the electronic and magnetic phase transition is a structural phase transition. Pure $LaMnO_3$ has a distorted perovskite structure containing a Jahn-Teller distorted oxygen octahedron surrounding a single manganese atom. There are three Mn-O bond lengths in the unit cell at about 1.92, 1.97, and 2.15 Å [81]. In the $0.20 \lesssim x \lesssim 0.48$ doping range, the same distortion of the Mn-O octahedra exists in the PI phase. In the FM phase the distortion is at least partially suppressed, leading to a narrower distribution of Mn-O bond lengths [82]. The first peak in the data, which is a negative peak in PDF due to the negative neutron scattering length of Mn, is due to the Mn-O bond length distribution. The resolution of the data is such that neither PDF nor EXAFS can uniquely resolve the separate bond lengths, except for the 310 K PDF data, which shows a pronounced shoulder on the high-r side the peak. The nearest neighbor Mn-O peak was analyzed to determine the mean Mn-O bond length and the Debye-Waller factor.

The EXAFS sample used in this study has doping concentration $x = 0.30$. EXAFS data collection was performed at the Mn K-edge and has a $k-$space range of 12 $Å^{-1}$ [83]. The EXAFS data covers 15 temperatures from 60 K to 416 K. The PDF data has doping concentration $x = 0.28$. The neutron PDF data has a $k-$space range of 35 $Å^{-1}$ [8]. It covers six temperatures from 10 K to 310 K.

For both data sets, the distribution of atoms in the nearest-neighbor triple-peak is modeled with a single Gaussian curve. The Debye-Waller factor therefore contains both the thermal component and a component due to the distribution of bond lengths in the peak. The EXAFS analysis follows work by Booth *et al.* [7]. The EXAFS

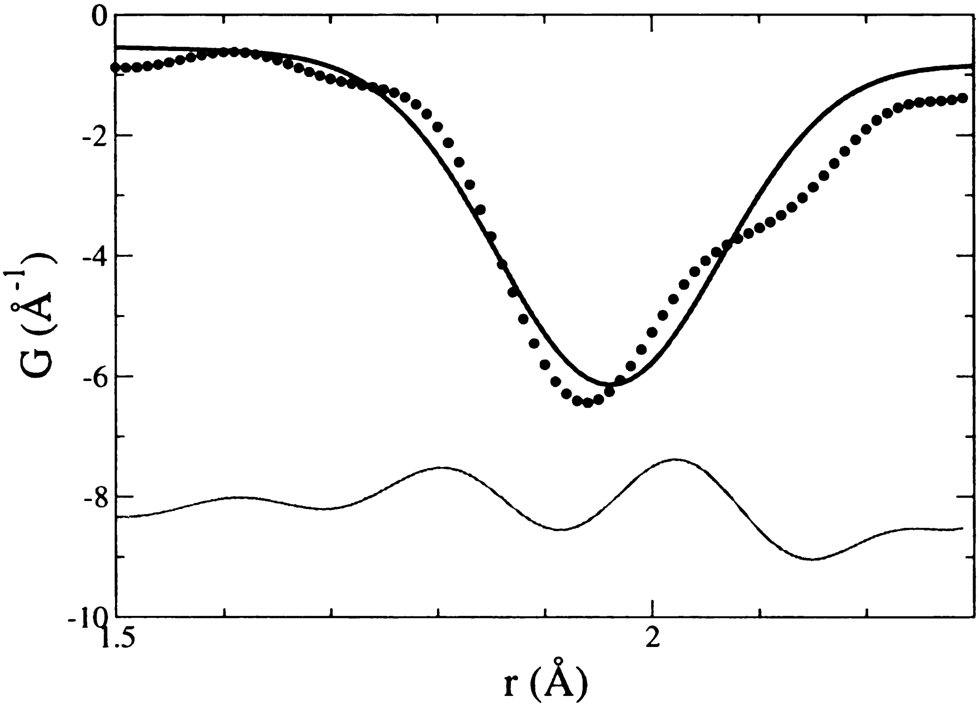


Figure 4.1: First peak PDF profile of $La_{0.7}Ca_{0.3}MnO_3$ at 310 K. The dots show the data and the thick line the fit. The thin line is the difference and is offset for clarity.

background was subtracted and fit to a first shell model using ATHENA and ARTEMIS, both part of the HORAE software package [62, 63]. The edge energy was determined as the inflection point in the absorption edge at 6554.5 eV and then shifted by -2 eV to get a better background fit. The EXAFS signal was fit to a single oxygen atom shell at a distance of 1.95 Å. The edge energy shift was fixed at zero, since it was already accounted for in the edge subtraction. This is believed to be a small approximation, but it is likely to affect the bond length measurements. The single shell model was fit with variable distance, amplitude, and Debye-Waller factor and the degeneracy of the peak was fixed at six. The first PDF peak was fit to the PDF equation as a single broad peak. The bond length, number density and Debye-Waller factor were varied in the fit. Two representative fits are shown in Figures 4.1 and 4.2.

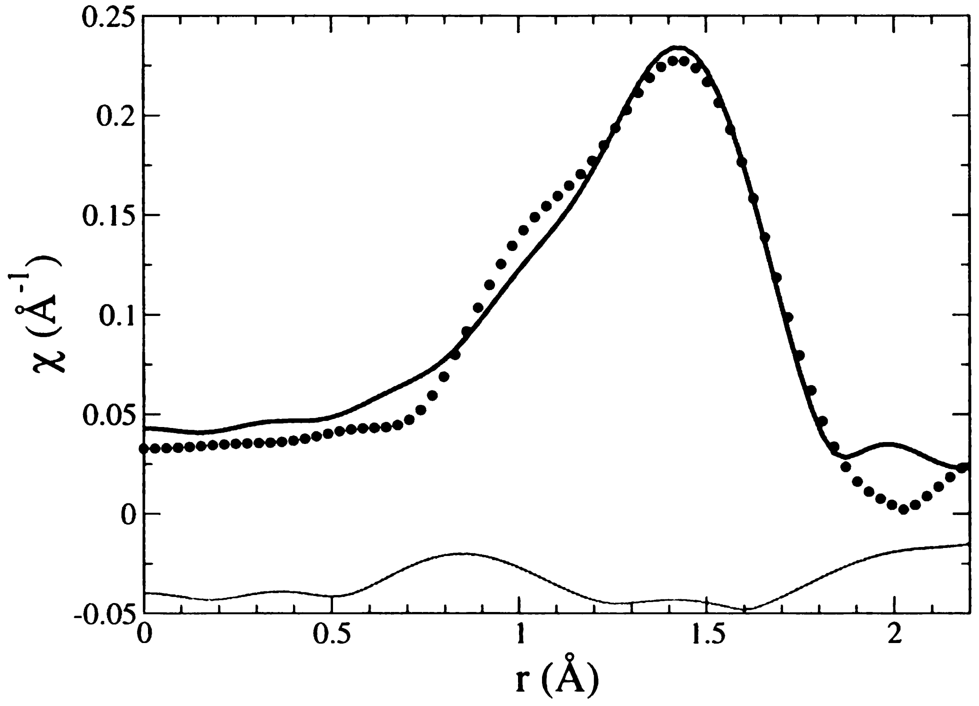


Figure 4.2: First peak of the EXAFS signal of $La_{0.7}Ca_{0.3}MnO_3$ at 310 K. The dots show the data and the thick line the fit. The thin line is the difference and is offset for clarity.

The comparison between the measured Mn-O bond lengths and Debye-Waller factors are shown in Figures 4.3 and 4.4. In each plot the error in the EXAFS data is shown. Two PDF fits, which differ by their fit ranges, give an indication of the uncertainty in the PDF results. The large discrepancy in these values at higher temperature is due to uncertainty in where the first peak ends and the second peak begins, as displayed in Figure 4.1. Figure 4.3 shows the reduction of the mean Mn-O bond length through the phase transition. The EXAFS results lie below those from the PDF data, although they are the same within the estimated uncertainty. The reason for the discrepancy is two-fold. First, the $1/r^2$ dependence in the EXAFS equation causes the longer bond lengths under the peak to be less emphasized in the fit [7]. This shifts the mean to a lower r-value. This is also true of the PDF results, but less pronounced because the PDF equation has only $1/r$ dependence. The second reason is due the time-of-flight instrument resolution effects discussed in the last section. When not compensated, these effects increase the bond lengths, and therefore all distances measured from the PDF. However, this effect is usually only significant at large distances and no standard data is available with which to quantify or determine its relevance in this case. The broadening in the Mn-O bond length distribution through the phase transition is shown in Figure 4.4. The agreement between the EXAFS and PDF fits is consistent, but the error bars indicate that the EXAFS fit is imprecise. Low resolution can contribute to peak broadening, but this would place the EXAFS Debye-Waller factors above the PDF values since the PDF data has better resolution.

The results of a full structural PDF analysis are also shown in Figure 4.3. The structure model of $La_{0.72}Ca_{0.28}MnO_3$ was fit using PDFgui [14], where the lattice constants, thermal parameters, and positions of the La/Ca and O atoms were refined within the Pbnm space group [22]. This fit yields three different Mn-O bond lengths, which are averaged in the figure. The results compare favorably with diffraction measurements on $La_{0.7}Ca_{0.3}MnO_3$ [84], although the PDF values are slightly larger. The actual transition point is obscured due to lack of data in the transition region. The trend in the full structural fit is contrary to that of the Gaussian fits (see Figure 4.3). That is, the average Mn-O bond lengths increase with temperature, as they should, rather than decrease. This contrary behavior is due to the exclusion of anharmonic effects in the single-shell EXAFS and PDF analyses [7]. As the distribution of bond lengths increases, the errors due to the Gaussian approximation become more severe, which manifests as a decrease in the average Mn-O bond length.

The purpose of this study was to highlight the difficulties in comparing structural quantities derived from PDF and EXAFS data. The use of the Gaussian model for

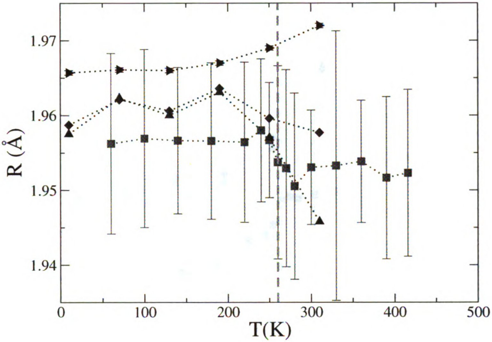


Figure 4.3: The first peak distance versus temperature for PDF and EXAFS data of $La_{1-x}Ca_xMnO_3$, for $x \approx 0.3$. The vertical dashed line indicates the phase transition temperature. The square symbols show the EXAFS results. The vertical bars give the estimated error on these values. The diamond and up-triangles show results from two different first-peak PDF fits. The right-triangles are results from a PDF refinement of the $La_{0.72}Ca_{0.28}MnO_3$ structure. The dotted lines are guides to the eye.

the Mn-O peak is the main source of this difficulty. The Gaussian model does predict the position of the phase transition, but it does not demonstrate the proper behavior of average Mn-O bond length with temperature because of the large structural component. This barrier cannot be overcome since a full structural EXAFS analysis is not possible with the given data. Resolution effects also influence the results. The errors caused by these issues is likely on the order of a few percent [71]. Theoretical development towards correcting these issues is slow, so other means to compensate for them are needed. The issues detailed in this chapter motivate the need for a

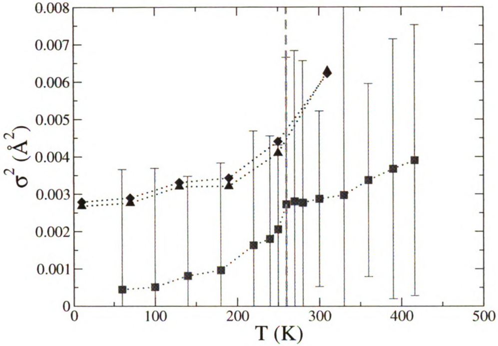


Figure 4.4: The first peak width versus temperature for PDF and EXAFS data of $La_{1-x}Ca_xMnO_3$, for $x \approx 0.3$. The vertical dashed line indicates the phase transition temperature. The square symbols show the EXAFS results. The vertical bars give the estimated error on these values. The diamond and triangles show results from two different PDF fits. The dotted lines are guides to the eye.

method of EXAFS and PDF data analysis that can utilize the strengths of the PDF and EXAFS techniques while mitigating their weaknesses.

Chapter 5

Corefinement

5.1 Introduction

The EXAFS and PDF methods are complementary structure determination techniques specialized for nanostructure investigations. Both techniques have been successfully applied to liquids, glasses, disordered and aperiodic materials, and crystalline materials with local distortions. These studies have elucidated many interesting physical phenomena, such as superconductivity, colossal magneto-resistance, thermoelectricity and negative thermal expansion. However, there is often disagreement within and between the EXAFS and PDF communities as to which explanation of a phenomenon is correct. Techniques that can reconcile the common discrepancies between EXAFS and PDF are of great value to both communities [85].

When carefully analyzed, EXAFS and PDF data provide complementary and contrasting information. Both techniques can be used to measure the distance between selected atom pairs in a material. The EXAFS signal gives the distance between a selected atomic species and its local neighbors to distances less than 10 Å [86]. The short range of EXAFS makes it less suitable for intermediate-range structure investigations, but the chemical specificity affords it a very detailed view of the local structure. Full structural refinement with EXAFS is often impossible due to the short range of the signal.

The PDF is a short- and intermediate-range probe. It provides the distances between all atom pairs within up to hundreds of angstroms of each other, regardless of their species. A common difficulty in structure modeling with the PDF is resolving overlapping peaks that result from multiple atomic pairs. Special techniques can be applied to resolve such peaks, such as anomalous x-ray scattering [44, 45], isotopic substitution for neutron PDF [42, 43] and joint x-ray and neutron PDF [46]. The contrasting PDFs obtained from these techniques are used to determine the profile corresponding to a selected set of atom species. The same contrast may be achieved by using the chemical-specific EXAFS signal to resolve short-ranged peaks in the PDF.

The goal of EXAFS and PDF corefinement is to determine reliable and unbiased structural solutions that make use of all the available data. Unbiased solutions are essential since no experimental technique can give ideal data without random and systematic error. Systematic errors are introduced during data collection, reduction and even analysis [47, 59]. By corefining multiple data, the effect of the systematic error from any one dataset is diminished.

The reliability of a structural solution is not only dependent on the errors in the data, but on the method used to obtain the solution. EXAFS and PDF refinement software commonly use downhill regression to fit a theoretical model to the experimental data [62, 13]. Interdependencies among the parameters in the fitting equation make finding a unique solution challenging. For example, the amplitude and phase parameters in the EXAFS equation are respectively highly correlated. This means that a change in one parameter can be compensated by a change in another with little effect on the quality of the fit. These correlations are not a property of the material, so they should be reduced in order to find the optimal solution. Overcoming this limitation in the way that EXAFS and PDF data are analyzed is a large part of the analysis effort. Known standards can be used to constrain or fix correlated parameters, as can additional data taken under different conditions, like EXAFS data from different absorption edges or temperatures. Increasing the range of a fit can introduce new parameters that can be constrained to reduce the overall correlation, but this can also complicate the analysis.

Techniques similar to corefinement currently exists in various forms. Reverse Monte Carlo (RMC) [87] frequently utilizes different types of structural data. RMC modeling refines a three-dimensional structure by randomly perturbing the atoms. A move is accepted if it increases the agreement between the model and experimental profile functions. Moves that decrease the agreement are accepted with a predefined probability. RMC has been used to model disordered materials simultaneously using EXAFS and PDF data [76], and has been recently applied to crystalline structures [88]. RMC modeling often employs outside knowledge, such as minimum distances between atoms, to restrain features of the structural model. RMC modeling produces the most disordered structures consistent with the data. Furthermore, RMC cannot distinguish between static and thermal disorder, so the latter is not included in the model. However, some thermal information can be derived from RMC solutions [89]. Another argument against RMC is the lack of parameter correlation information provided by the technique. This makes it difficult to fully assess the quality of its solutions.

Powder diffraction and EXAFS corefinement has been investigated extensively by Binsted and coworkers [90]. In that work, many of the pitfalls of corefinement were carefully studied, in particular, the treatment of disorder. The effects of correlated atomic motion that appear in the EXAFS signal also exists in the diffraction pattern, but is inaccessible with reciprocal-space refinement techniques. The solution presented by Binsted and coworkers was to not corefine the Debye-Waller factors for the first few coordination shells. The correlation effects die out with distance, and so can be safely ignored for the higher coordination shells. This approach devalues the corefinement by ignoring the most detailed and accurate information in the EXAFS signal. This problem does not exist in PDF and EXAFS corefinement since both of these models must take into account vibrational correlations. However, PDF and EXAFS corefinement must contend with the issue of Debye-Waller factors for multiple-scattering EXAFS paths.

The corefinement technique presented here is unique because it combines two complementary local-structure techniques that are often disjoint. The analysis presented in the following chapter takes full consideration of correlations in the fitting parameters. Like earlier work, this technique does not use all of the information available in the EXAFS signal, but it plays to the chemical-specific and short-range strengths of the technique.

5.2 Corefinement Technique

Corefinement uses established refinement protocols from the EXAFS refinement software, IFEFFIT [62] and ARTEMIS [63], and PDF refinement software, PDFFIT [13] and PDFGUI [14]. These programs use a least-squares downhill regression algorithm to minimize the χ^2 calculated from a theoretical fitting equation, experimental data and uncertainty in the data. Parameters of the fitting equation are constrained to variables that are automatically adjusted to minimize χ^2. In PDF and EXAFS corefinement, the structural variables common to both techniques, the nearest-neighbor bond lengths and Debye-Waller factors, are also constrained. This allows the PDF to benefit from the chemical specificity of the EXAFS signal, and in turn the EXAFS benefits from the intermediate-range information in the PDF.

The main complication with sharing information between parallel EXAFS and

PDF refinements is the treatment of multiple-scattering EXAFS paths. Like IFEFFIT, the corefinement algorithm uses FEFF [59, 60] to generate its theoretical standards. This approach to EXAFS fitting must handle changes in scattering path geometry with elaborate constraint schemes, or by regenerating the path information after each change. Although this latter approach is computationally feasible, it introduces logistical problems, such as how to handle the new paths that result from symmetry changes. There is an additional complication in comparing the Debye-Waller factors in the EXAFS and PDF equations. Equation 2.8 is obtained under the assumption of small disorder, so it applies equally well to the PDF and EXAFS Debye-Waller factors. There is no simple analogous formula for multiple-scattering EXAFS paths. Such a formula would likely introduce multi-body correlation parameters that would need to be determined.

The corefinement algorithm avoids these difficulties by only using the first shell of the real-space EXAFS signal. This approach is somewhat limiting, since it requires that the first shell can be cleanly isolated. It is not uncommon for coordination shells to blend together. In principle, information could be used from further shells, as long as the multiple-scattering paths are negligible. The exclusion of this further-neighbor information is acceptable since the real-space EXAFS signal decays rapidly after the first peak. Furthermore, EXAFS chemical information from other absorption edges is better suited for resolving peaks in the PDF than the structural information contained in the higher coordination shells.

5.3 Computational Details

The corefinement algorithm minimizes the combined χ^2 of the EXAFS and PDF components of the fit. χ^2 is a measure of error between data and a model, and for a single fit component is defined as

$$\chi^2 = \frac{N_{idp}}{N}\sum_{i=0}^{N}\frac{(y_i - f(x_i))^2}{\sigma_i^2}, \tag{5.1}$$

where the x_i is the i^{th} data sampling point, y_i and σ_i are the data and its uncertainty at that point, and f is the fitting equation. The N_{idp}/N term rescales χ^2 to represent the error over the number of independent data points as opposed to the number of sampled data points. The number of data points in the PDF or the real-space EXAFS signal, N, can be arbitrarily chosen when the real-space signal is Fourier transformed from reciprocal-space. N_{idp} is the theoretical limit on the number of fitting variables that can be uniquely specified in the fit. The rescaling affects the determination of uncertainties in the fit variables, which are derived from χ^2. The uncertainty in a variable, v, is given by

$$\sigma_v = \sqrt{2\left(\frac{\partial^2\chi^2}{\partial v^2}\right)^{-1}}. \tag{5.2}$$

This is measured for each variable after the fit has converged.

In EXAFS fitting, $N_{idp} \approx 2 + 2\Delta r\Delta k/\pi$, where Δr and Δk are the real-space and k-space extent of the fit [61]. For a first shell fit, this is typically around 10. In PDF fitting, $N_{idp} \approx \Delta r\Delta Q/\pi$, where $\Delta Q = Q_{max}$ is the Q-range of the data. This estimation of the number of independent points is not found in the PDF literature, but it can be derived in the same fashion as the EXAFS estimation, with only minor modification [61]. In the study presented in the next chapter, N_{idp}/N typically ranges from 1/2 to 1/20.

Fitting components from the same or multiple techniques can be refined in parallel by contributing to the χ^2 of the fit. In the corefinement routine, the total χ^2 is a weighted sum of the χ^2 values from the PDF and EXAFS components,

$$\chi^2_{co} = \sum_{i}^{PDF} w_i \chi_i^2 + \sum_{j}^{EXAFS} w_j \chi_j^2 \tag{5.3}$$

The sums run over the individual PDF and EXAFS components to the fit, such as PDF data taken with different radiation, or EXAFS data take from different edges. The weights w_i and w_j define the contribution of a given component to the overall fit. The larger its weighting, the more a component will influence the direction of the optimization algorithm as it attempts to minimize χ^2_{co}. An individual component of the fit may also hold more influence over the direction of the optimization by virtue of smaller uncertainty or larger number of independent points.

One measure of merit for a fit is the value of the reduced χ^2, denoted χ^2_ν. The reduced χ^2 is calculated as

$$\chi^2_\nu = \frac{\chi^2}{N_{idp} - \nu}, \tag{5.4}$$

where ν is the number of variables used in the fit. $N_{idp} - \nu$ represents the number of degrees of freedom in the fit. When data are corefined, N_{idp} represents the total number of independent data points from all data. An ideal fit has $\chi^2_\nu \approx 1$. As is the case in PDF and EXAFS fitting, even good corefinements can have χ^2_ν much larger than this. Accepted practice in the EXAFS community [91] is to rescale the calculated uncertainties to better reflect quality of the fit and to give a better estimate of uncertainty in the fit variables. For the corefinement algorithm, the uncertainties in the fit variables are multiplied by $\sqrt{\chi^2_\nu}$, which is equivalent to rescaling χ^2_ν to one. As a consequence of this rescaling, any global weighting of χ^2 does not affect the uncertainty calculations, so the weights in Equation 5.3 need not be normalized.

Corefinements are evaluated based on several measures. χ^2 is not a good measure of comparison between different fits since it is a function of the weighting ratios. A more proper indicator of goodness of fit is R_w, the root-mean-squared deviation of the fit from the data.

$$R_w = \sqrt{\frac{\sum_i (y_i - f(x_i))^2}{\sum_i y_i^2}}. \tag{5.5}$$

Since R_w is a relative measure of error, it is appropriate for comparing results using different models, or corefinements with different weightings. R_w for a corefinement is the weighted sum of the R_w values of its components.

In addition to the uncertainty in the fit variables, it is also important to evaluate the correlation between fitting variables. The linear statistical correlation between two variables, u and v, is defined as

$$\rho_{uv} = \frac{C_{uv}}{\sqrt{C_{uu} C_{vv}}}. \tag{5.6}$$

Here, C_{uv} is the covariance of variables u and v,

$$C_{uv} = 2 \left(\frac{\partial^2 \chi^2}{\partial u \partial v} \right)^{-1}. \tag{5.7}$$

Note that $C_{uu} = \sigma_u^2$, the square of the uncertainty. The linear statistical correlation gives an indication of how much a change in one variable affects the change in another. A large statistical correlation ($|\rho_{uv}| > 0.7$) indicates that two variables are strongly correlated and cannot be determined uniquely by the fit. Weak correlation ($|\rho_{uv}| < 0.3$) between variables is desirable, as it indicates that the effect of a change in one variable will manifest in χ^2, instead of in another variable.

Chapter 6

Corefinement Study of $In_xGa_{1-x}As$

6.1 Introduction

The semiconductor series GaAs, $In_{0.5}Ga_{0.5}As$, and InAs are used as test materials for the corefinement algorithm. The average structural properties of this system obey Vegard's law [9], that is, they follow a linear relationship between the end members. On average, each compound has the zinc-blende structure, where the Ga/In atoms randomly lie on one sub-lattice and the As atoms lie on the other. EXAFS and PDF studies [10, 92, 11, 93, 94] and force-field models [95, 96, 97, 98] show that the local structure of the alloyed compounds is distorted from the average. This distortion is clearly visible in the split nearest-neighbor peak in the PDF (Figure 6.1). Evidence suggests that the As sublattice in the alloys is significantly more distorted than the Ga/In sublattice. This forms the basis of the physical model used for the alloys.

6.2 Data collection

The PDF data were obtained from x-ray powder diffraction measurements performed at low temperature (10 K) at the Cornell High Energy Synchrotron Source. Specifics of sample preparation, data collection and data reduction have been published else-

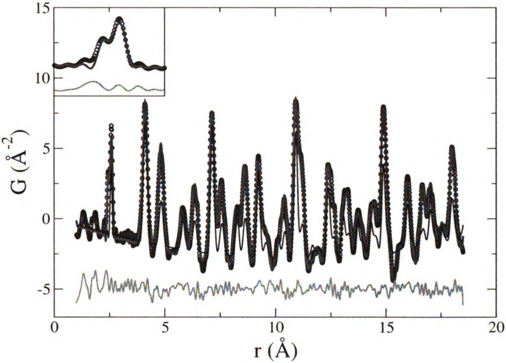


Figure 6.1: PDF of $In_{0.5}Ga_{0.5}As$ (circles), failed fit to the PDF (thin dark line), successful fit to the PDF (thick line) and the difference between the data and successful fit (thin line, shifted downward for clarity). The inset shows a close-up of the nearest-neighbor peak and the successful fit.

where [94]. The estimated error in the data was calculated from the statistical uncertainty in the detector counts and carried through the calculation of the PDF [99]. The average values for these uncertainties are given in Table 6.1.

The EXAFS data were collected by Moneeb Shatnawi and Wenduo Zhou of Michigan State University in transmission mode on beamline 5BM-DND at the Advanced Photon Source, using a Si $\langle 111 \rangle$ double crystal monochromator on the same samples as the PDF experiment. The samples were finely ground and brushed on to scotch tape. Layers of the tape were stacked to obtain optimal absorption at the edge, $\mu x \approx 1$. A displex was used to achieve a temperature of 19 ± 1 K. A minimum of

Table 6.1: Uncertainties PDF and EXAFS data.

Uncertainty in $G(r)$ (Å^{-2})	
GaAs	0.1064
$In_{0.5}Ga_{0.5}As$	0.1267
InAs	0.1405
Uncertainty in $\chi(k)$	
GaAs, As K-edge	0.0010
GaAs, Ga K-edge	0.0032
$In_{0.5}Ga_{0.5}As$, As K-edge	0.0012
$In_{0.5}Ga_{0.5}As$, Ga K-edge	0.0008
$In_{0.5}Ga_{0.5}As$, In K-edge	0.0003
InAs, As K-edge	0.0014
InAs, In K-edge	0.0005

three scans were performed on each K-edge of the samples. Scans were started at least 200 eV below each edge and proceeded in 0.2 eV steps beyond the edge. Post-edge scans continued out to at least 16 Å^{-1} in increments of 0.05 Å^{-1}. The detector count times were scaled to reduce noise, and at least 75000 detector counts were collected on the transmission end at each energy. ATHENA [63] was used to merge the scans and subtract the background using standard procedures. The first shells of the real-space EXAFS signals were isolated in real-space and then transformed to k-space for analysis. The uncertainty in EXAFS data was calculated as the root-mean-square of the real-space signal between 15 and 25 Å and converted to uncertainty in k-space using Parseval's theorem [91],

$$\epsilon_k = \epsilon_r \sqrt{\frac{\pi}{\delta k (k_{max} - k_{min})}}. \qquad (6.1)$$

Here, ϵ_r and ϵ_k are the uncertainties in real-space and k-space, respectively, δ_k is the spacing of points in k-space and $[k_{min}, k_{max}]$ is the extent of the Fourier transform from real-space to k-space. This procedure measures the noise in the signal in a region void of measurable structural features. These uncertainty values are summarized in Table 6.1.

6.3 Data analysis

The analyses of GaAs and InAs were carried out identically. For the PDF fits, the zinc-blende lattice constant, anisotropic atomic displacement parameters, low temperature Debye-Waller correlation term, r-independent experimental peak-broadening term and scale factor were allowed to vary. The displacement parameters were constrained to be isotropic. The EXAFS fits had variable nearest-neighbor bond lengths, Debye-Waller factors, edge energies and amplitudes. In the cases where data from multiple EXAFS edges were used, the different components of the fit were given equal χ^2-weights and the nearest-neighbor bond lengths and Debye-Waller factors were constrained from the different edges to be the same. The first-shell EXAFS fits were done in k-space with no windowing or k-weighting. The multi-shell fits were performed with ARTEMIS [63, 62] in real-space.

The corefinements of GaAs and InAs combined the fitting procedures from both the EXAFS and PDF fits. The nearest-neighbor bond lengths and correlated Debye-Waller factors from the EXAFS portion of the corefinements were constrained to be equal to those from the PDF portion. In this way, the PDF and EXAFS portions of the corefinements provide mutual feedback.

The structural model for $In_{0.5}Ga_{0.5}As$ is consistent with the work of Balzarotti and coworkers [96]. Assuming the Ga/In sublattice is fixed and ignoring interactions beyond the nearest-neighbors, As atoms exist in one of five local environments. These environments are differentiated by the number of nearest-neighbor Ga atoms and lead to eight unique nearest-neighbor bonds; four Ga-As distances and four In-As distances. The distortions in the structure result from the As atoms moving towards their Ga neighbors. It is not possible to model the different As environments in $In_{0.5}Ga_{0.5}As$ with a first-shell EXAFS fit since the contributions from the different environments to the nearest-neighbor EXAFS peak are irresolvable. Instead, in the EXAFS fits, the nearest-neighbor peaks are modeled as a single shell and the fits were performed like those for GaAs and InAs. In this case, the EXAFS Debye-Waller factor represents both the thermal and static disorder.

In the PDF fits, a 512 atom zinc-blende supercell was generated and the mixed sublattice was randomly occupied with Ga and In sites consistent with the stoichiometry. The As atoms within a given environment were allowed to move towards the center-of-mass of their neighboring Ga atoms. Thus, the As atoms with zero or four neighboring Ga atoms do not move and As atoms in the other environments move either in the $\langle 100 \rangle$ (for two Ga neighbors) or $\langle 111 \rangle$ (for one or three Ga neighbors) directions. The As atoms within a given environment were constrained to move by the same amount. The corefinements were performed like the PDF fits, except that the EXAFS contribution from each As environment was appropriately summed and compared to the EXAFS data to get the χ^2 value for the EXAFS portion of the fit. Like with the PDF fits, the static disorder is explicitly included in the structure model. Some variability in the results is expected to result from the non-ideal distribution of the As environments in the relatively small supercell [97], so multiple model alloys were used in the fits.

To avoid bias in the $In_{0.5}Ga_{0.5}As$ fits, the initial As displacements were randomly chosen for a series of trials. The results of two of these trials, one successful, one not, are shown in Figure 6.1. The poor fit is unphysical, as it overestimates the lattice constant while shifting the As atoms away from their Ga neighbors. To test the convergence of the corefinement algorithm, this poor fit was used as the starting point for several refinements.

To test the adequacy of the structural model, several PDF fits were performed where the Ga/In sublattice was allowed to relax after the As sublattice was relaxed and frozen into place. The Ga/In atoms were identified according to the environments of their neighboring As atoms and atoms within the same environment were constrained to move by the same amount towards the center of their neighboring

As atoms. Ideally, the relaxation procedure on both sublattices would be performed iteratively until a self-consistent answer is achieved. Unfortunately, after the first iteration, the number of local As environments leads to too many fitting variables to produce meaningful fitting results. Regardless, the results after the second relaxation stage will give insight into the quality of the structure model.

6.4 Corefinement Results

To illustrate the variability in the results from conventional PDF and EXAFS fitting, a series of fits was performed for GaAs, $In_{0.5}Ga_{0.5}As$ and InAs. The fits were configured as described in the previous section. Structural parameters of interest, namely the nearest-neighbor bond lengths and Debye-Waller factors, are compared in Figures 6.2 and 6.3. The figures indicate the EXAFS edges used in the EXAFS fits and corefinements. In the PDF fits and corefinements, the results are averaged over 7 to 12 model alloys. In each case, the uncertainty in the average values are less than or comparable to those produced by the fitting procedure. In each refinement the overall goodness of fit, R_w, was between 0.15 and 0.20.

The figures show excellent consistency among the corefinement results compared to the EXAFS results. These results were obtained by combining EXAFS and PDF data for each combination of EXAFS edges. The corefinement results from panels **a** and **b** of these figures agree well with the PDF and multi-edge, multi-shell EXAFS results. In addition, the PDF bond length values show good agreement with previous results [11]. It is apparent that consistent results can be obtained from PDF and EXAFS data with the proper processing. Note, however, that the single-shell EXAFS fits do not agree particularly well with each other or with the other fits. These single-shell fits are plagued by large correlations between the structural and non-structural variables of the fit. The largest correlations from these fits are displayed in Table 6.2.

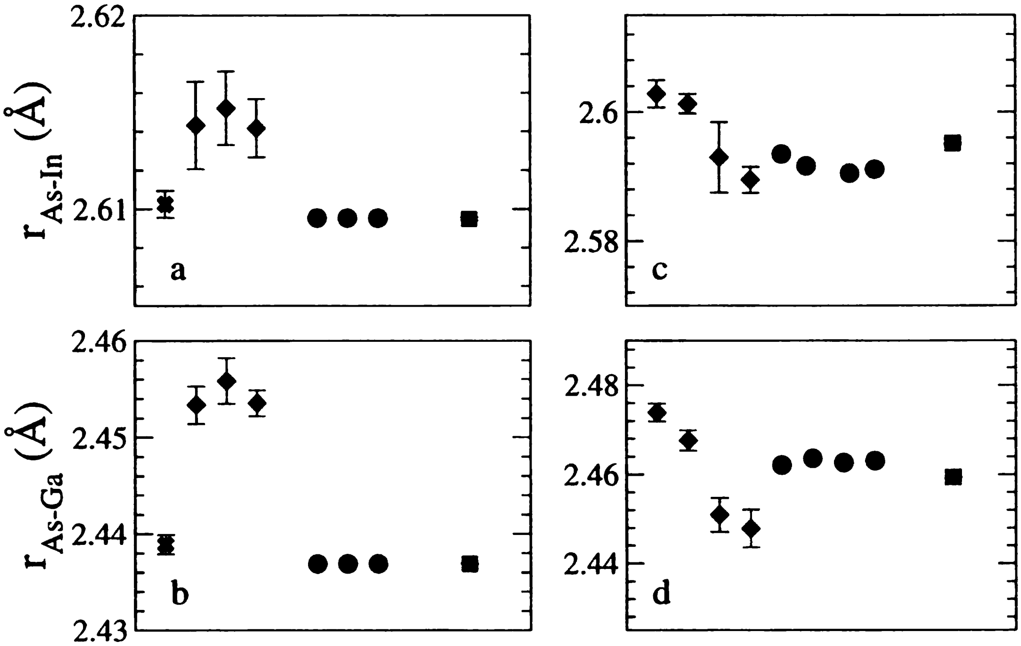


Figure 6.2: Variability in the nearest-neighbor bond lengths obtained from various EXAFS, PDF and equal-weight EXAFS and PDF corefinement trials. Uncertainty is indicated by the vertical bars. For the corefinement and PDF trials, the uncertainty in the values averaged over several model alloys is comparable to the uncertainty given by the least-squares fitting, which is smaller than the plot symbols. For the effects of weighting on the bond lengths, see Figure 6.4. (X) Results from multi-shell EXAFS fit to 5 Å. (♦) First shell EXAFS results. (●) Corefinement results averaged over several model alloys. (■) PDF results averaged over several model alloys. The EXAFS edges used in the trials are indicated on the horizontal axes. Panel a: In-As bond lengths from InAs. Panel b: Ga-As bond lengths from GaAs. Panel c: In-As bond lengths from $In_{0.5}Ga_{0.5}As$. Panel d: Ga-As bond lengths from $In_{0.5}Ga_{0.5}As$.

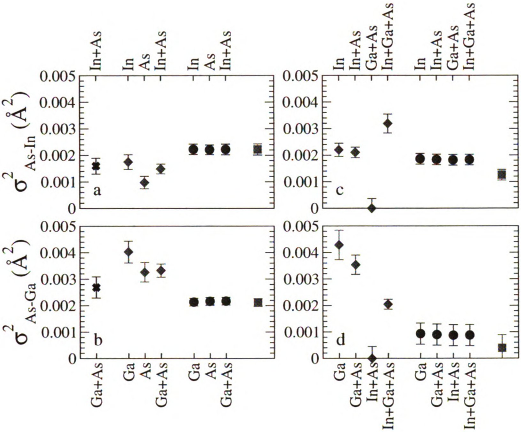


Figure 6.3: Variability in the nearest-neighbor Debye-Waller factors obtained from various EXAFS, PDF and equal-weight EXAFS and PDF corefinement trials. Uncertainty is indicated by the vertical bars. For the corefinement and PDF trials, the uncertainty in the Debye-Waller factors averaged over several model alloys is smaller than the uncertainty given by the least-squares fitting, which shown in the plot. For the effects of weighting on the Debye-Waller factors, see Figure 6.5. (×) Results from multi-shell EXAFS fit to 5 Å. (♦) First shell EXAFS results. (●) Corefinement results averaged over several model alloys. (■) PDF results averaged over several model alloys. The EXAFS edges used in the trials are indicated on the horizontal axes. Panel a: In-As Debye-Waller factors from InAs. Panel b: Ga-As Debye-Waller factors from GaAs. Panel c: In-As Debye-Waller factors from $In_{0.5}Ga_{0.5}As$. Panel d: Ga-As Debye-Waller factors from $In_{0.5}Ga_{0.5}As$.

Table 6.2: Statistical correlation in EXAFS fits of GaAs and InAs.[a]

GaAs refinements

Correlated variables		Ga+As [b]	Ga [c]	As [c]	Ga+As [c]
ΔR_{Ga-As}	ΔE_{Ga}	0.59	0.80	-	0.47
ΔR_{Ga-As}	ΔE_{As}	0.80	-	0.86	0.86
σ^2_{Ga-As}	A_{Ga}	0.63	0.79	-	0.44
σ^2_{Ga-As}	A_{As}	0.80	-	0.88	0.86
ΔE_{Ga}	ΔE_{As}	0.47	-	-	0.42
A_{Ga}	A_{As}	0.52	-	-	0.39

InAs refinements

Correlated variables		In+As [b]	In [c]	As [c]	In+As [c]
ΔR_{In-As}	ΔE_{In}	0.67	0.87	-	0.83
ΔR_{In-As}	ΔE_{As}	0.52	-	0.66	0.49
σ^2_{In-As}	A_{In}	0.74	0.89	-	0.86
σ^2_{In-As}	A_{As}	0.60	-	0.70	0.51
ΔE_{In}	ΔE_{As}	0.35	-	-	0.41
A_{In}	A_{As}	0.46	-	-	0.43

---

[a] Only correlations among first-shell parameters are included. Correlations less than 0.30 in magnitude are not reported.

[b] Multi-shell EXAFS refinement

[c] Single-shell EXAFS refinement

Strong correlations are expected from the underconstrained single-edge fits, but as the table shows, the combination of data from two edges does not reduce the correlations below 0.7.

Surprisingly, the multi-shell, multi-edge EXAFS fits do not have significantly lower correlations than the single-shell fits. These fits were performed out to the third coordination shell in the structure (≈ 5 Å), using data from both available absorption edges. For these samples, the EXAFS data alone is not sufficient to provide the structural parameters with the desired certainty.

The consistency among the corefinement results in Figures 6.2 and 6.3 suggests that combining the EXAFS and PDF data with a structural model produces an unambiguous structural solution. The correlations among the PDF and corefinement

Table 6.3: Statistical correlation in PDF fits and corefinements of GaAs and InAs.[a]

GaAs refinements

Correlated variables		PDF	Ga	As	Ga+As
Λ	σ_Q	0.68	0.68	0.68	0.68

InAs refinements

Correlated variables		PDF	In	As	In+As
Λ	σ_Q	0.70	0.70	0.69	0.70
A_{In}	ΔE_{In}	-	0.75	-	-
A_{As}	ΔE_{As}	-	-	0.44	-
ΔE_{In}	ΔE_{As}	-	-	-	0.46
A_{In}	A_{As}	-	-	-	0.80

---

[a] Only correlations among first-shell parameters are included. Correlations less than 0.30 in magnitude are not reported.

fit parameters are shown in Table 6.3. At first glance, the correlations may seem as strong as in the EXAFS fits. However, the correlations in Table 6.3 are all among non-structural parameters. These parameters result from experimental or physical effects that do not relate to the structural model. The strong PDF correlations can be eliminated by measuring σ_Q with a known standard, and do not cast doubt on the structure solution.

Table 6.3 appears to suggest that the PDF fits have few correlations that negatively affect the fit. This is not the case. GaAs and InAs have simple structural models with very few adjustable parameters. The model for $In_{0.5}Ga_{0.5}As$, on the other hand, is rather complicated. The As shift variables and the atomic displacement parameters from all three atom types determine the PDF peak shapes. These variables were occasionally correlated at strong levels. The different model alloys and EXAFS edges used in the corefinements had an unpredictable result on all the correlations. The EXAFS correlations were usually reduced, and some cross-correlation between the EXAFS and PDF parameters took place. In a few cases the PDF correlations increased slightly in the corefinements. Regardless, the average corefinement results for $In_{0.5}Ga_{0.5}As$ are much more consistent than the EXAFS results, and in the case of GaAs and InAs, agree well with the multi-shell EXAFS fit. Clearly, correlations among the fit variables are not the only indication of a reliable fit.

6.5 The effect of χ^2-weighting

Reliability of the $In_{0.5}Ga_{0.5}As$ fits is hard to judge based on goodness of fit and correlations alone. According to the R_w values, the corefinement combining PDF and EXAFS Ga K-edge data is the "best" fit, with $R_w = 0.15$. The other corefinements had R_w values between 0.18 and 0.20. The R_w factor is not a very helpful quality indicator when comparing corefinements that use different combinations of data. Data taken under different conditions are expected to have varying levels of systematic error. It is better to judge the fits based on the effects of this error on the structural parameters, rather than its effect on R_w.

One way to get a handle on the influence of the systematic errors on the corefinements is to vary the χ^2-weight of the PDF and EXAFS components. Without systematic errors, the optimal values of the structural parameters would be the same whether they came from PDF or EXAFS. In reality, the optimal values are a function of the error. In turn, the contribution of this error to the corefinement results is a function of the χ^2-weight of a component. By adjusting the χ^2-weights of the corefinement, the effects of the systematic errors can be seen in the measured quantities. Previous corefinement studies have not discussed in detail the effect of weighting on the results. This is an important issue, as the χ^2-weighting is not a property of the material. The most reliable solution will not depend on it.

The results of two weighting trials on $In_{0.5}Ga_{0.5}As$ are shown in Figures 6.4 and 6.5. In these trials, a single model alloy was used and the χ^2-weight of the EXAFS component was varied between zero and one. The weights for the PDF and EXAFS

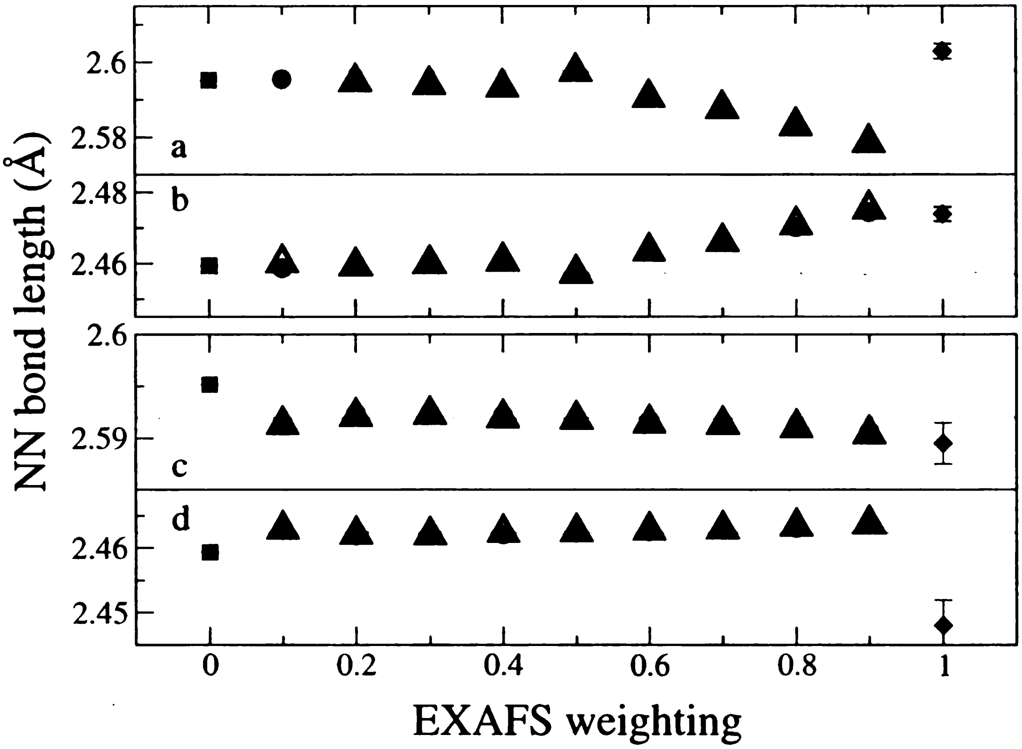


Figure 6.4: Consistency of corefined nearest-neighbor bond lengths for $In_{0.5}Ga_{0.5}As$ with respect to EXAFS weighting. The corefinement results are stable when EXAFS data from multiple edges is used. For the PDF and corefinement results, the uncertainty is smaller than the plot symbols. (■) PDF results. (♦) First shell EXAFS results. (●) Corefinement results. (△) PDF relaxation of Ga/In sublattice using corefinement results. The horizontal light-gray line goes through the 50% weighting point in each panel. Panel a: In-As bond lengths. Corefinements use Ga K-edge data. EXAFS result is from the In K-edge refinement and is shown for comparison. Note that the relaxation point at 0.1 weighting is out of frame. Panel b: Ga-As bond lengths. Corefinements use Ga K-edge data. EXAFS result is from the Ga K-edge refinement. Panel c: In-As bond lengths. Corefinements use In, Ga, and As K-edge data. EXAFS result is from the first-shell three-edge refinement. Panel d: Ga-As bond lengths. Corefinements use In, Ga, and As K-edge data. EXAFS result is from the first-shell three-edge refinement. Note the different scale in the panels.

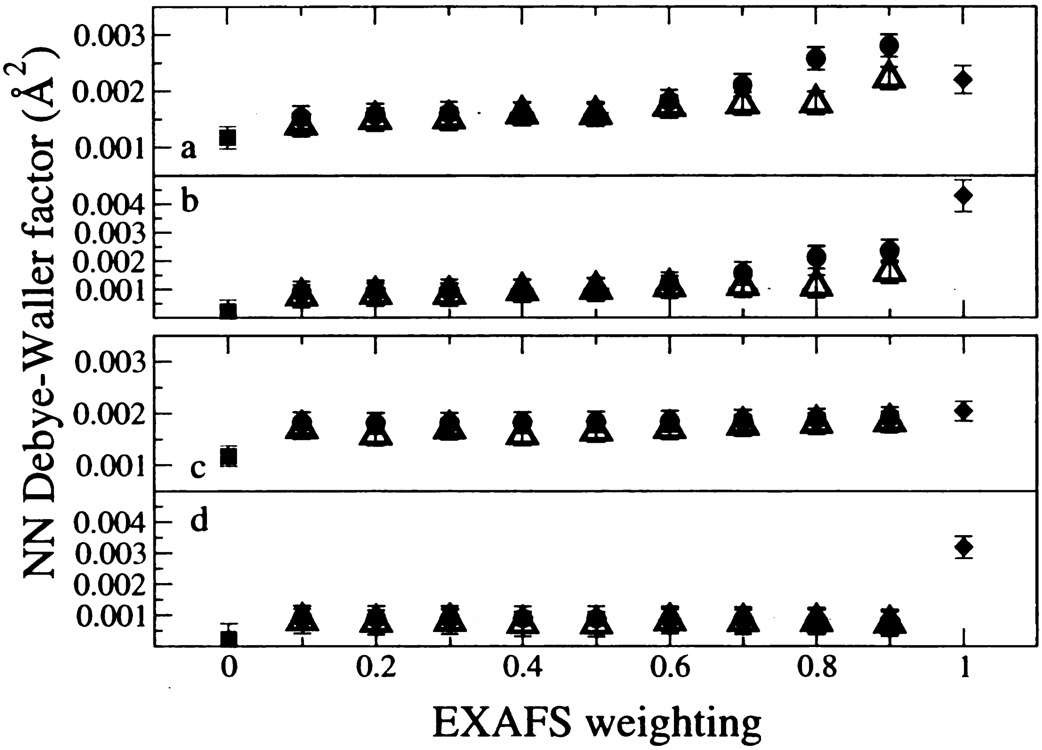


Figure 6.5: Consistency of corefined nearest-neighbor Debye-Waller factor for $In_{0.5}Ga_{0.5}As$ with respect to EXAFS weighting. The corefinement results are stable when EXAFS data from multiple edges is used. (■) PDF results. (♦) First shell EXAFS results. (●) Corefinement results. (△) PDF relaxation of Ga/In sublattice using corefinement results. The horizontal light-gray line goes through the 50% weighting point in each panel. Panel a: In-As Debye-Waller factors. Corefinements use Ga K-edge data. EXAFS result is from the In K-edge refinement and is shown for comparison. Panel b: Ga-As Debye-Waller factors. Corefinements use Ga K-edge data. EXAFS result is from the Ga K-edge refinement. Panel c: In-As Debye-Waller factors. Corefinements use In, Ga, and As K-edge data. EXAFS result is from the first-shell three-edge refinement. Panel d: Ga-As Debye-Waller factors. Corefinements use In, Ga, and As K-edge data. EXAFS result is from the first-shell three-edge refinement. Note the different scale in the panels.

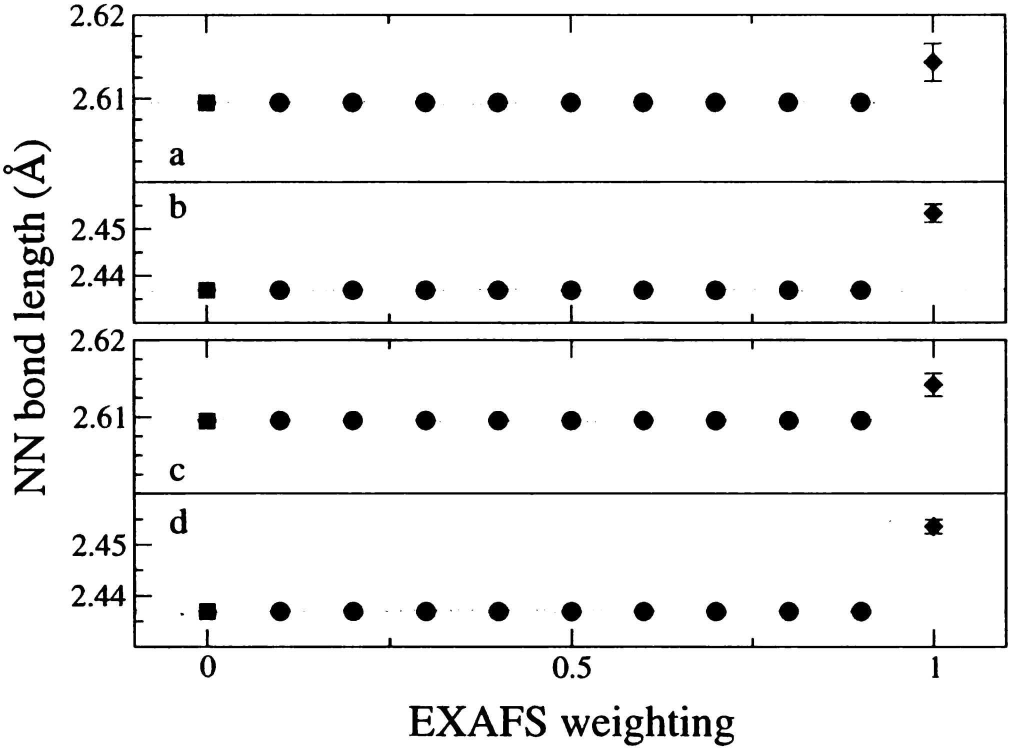


Figure 6.6: Consistency of corefined nearest-neighbor bond lengths for GaAs and InAs with respect to EXAFS weighting. For the PDF and corefinement results, the uncertainty is smaller than the plot symbols. (■) PDF results. (◆) First shell EXAFS results. (●) Corefinement results. The horizontal light-gray line goes through the 50% weighting point in each panel. Panel a: In-As bond lengths. EXAFS fits and corefinements use In K-edge data. Panel b: Ga-As bond lengths. EXAFS fits and corefinements use Ga K-edge data. Panel c: In-As bond lengths. EXAFS fits and corefinements use In and As K-edge data. Panel d: Ga-As bond lengths. EXAFS fits and corefinements use Ga and As K-edge data.

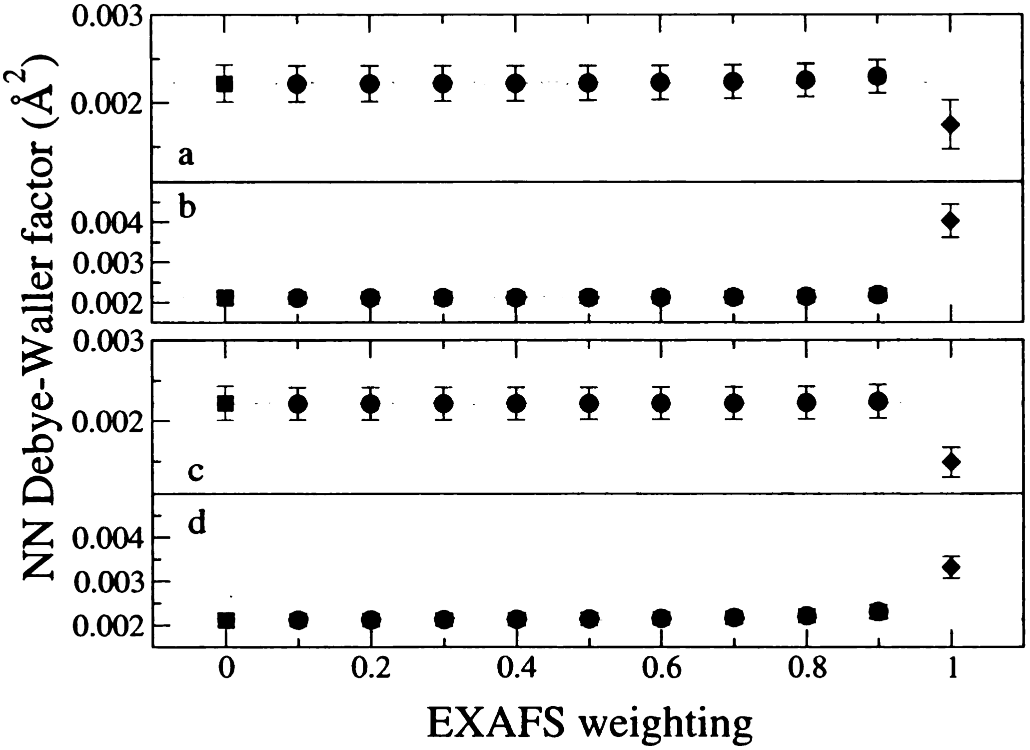


Figure 6.7: Consistency of corefined nearest-neighbor Debye-Waller factor for GaAs and InAs with respect to EXAFS weighting. (■) PDF results. (♦) First shell EXAFS results. (●) Corefinement results. The horizontal light-gray line goes through the 50% weighting point in each panel. Panel a: In-As Debye-Waller factors. EXAFS fits and corefinements use In K-edge data. Panel b: Ga-As Debye-Waller factors. EXAFS fits and corefinements use Ga K-edge data. Panel c: In-As Debye-Waller factors. EXAFS fits and corefinements use In and As K-edge data. Panel d: Ga-As Debye-Waller factors. EXAFS fits and corefinements use Ga and As K-edge data.

portions of the corefinement were forced to sum to one. Panels **a** and **b** in the figures show the results from the corefinements combining PDF and EXAFS Ga K-edge data. The results are strongly dependent on weighting. The deviation of the results from the equal-weight values gives an estimate of how the systematic errors affect the measurements. The series shows that the errors in the EXAFS data are not significant as long as the PDF portion of the corefinement is given the majority of the weight. Similar results were obtained for a PDF and In K-edge weighting series, which indicates that the single-edge corefinements are underconstrained.

Panels **c** and **d** of Figures 6.4 and 6.5 show the results from weighting trials that corefined PDF data with EXAFS data from the Ga, In and As K-edges. In these trials, the different EXAFS components were given equal shares of the EXAFS χ^2-weight. These results show only a weak dependence on the weighting. Furthermore, the corefinement values do not interpolate between the values from the independent fits. This is strong evidence that the systematic errors in the data have little effect on the results. The decrease in the In-As bond length from the PDF value is caused by the strong influence of the In K-edge data on the corefinement, which is a result of the low uncertainty in these data (see Table 6.1). The increase of the Ga-As bond length conserves the average nearest-neighbor bond length, which is dictated by the intermediate range of the PDF. This complex behavior demonstrates the interplay between the chemical-specific EXAFS data and intermediate-range PDF data. Such results could not have come about without refining both the PDF and EXAFS data.

Figures 6.6 and 6.7 show the bond length and Debye-Waller factor weighting series for GaAs and InAs. As can be seen in the figures, the results are consistent with weighting no matter how much EXAFS data is used. This indicates that the chemical information from the EXAFS data is not needed to resolve the peaks in the PDF, but the PDF and EXAFS data are indeed consistent. This fact could not be demonstrated with certainty, except through corefinement.

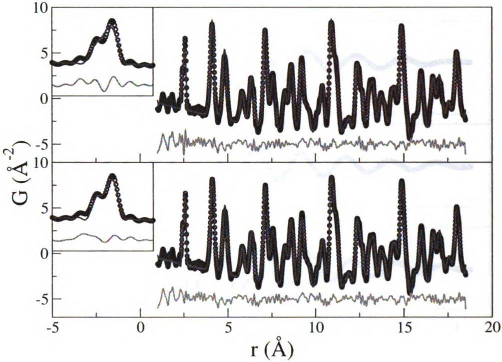


Figure 6.8: PDF of $In_{0.5}Ga_{0.5}As$ (circles), fit to PDF (thick line) and difference curve (thin line, shifted down for clarity) for a three-edge equal-weight corefinement. Top: Corefinement with nearest-neighbor inset. Bottom: PDF after relaxation of the Ga/In sublattice.

The PDF and EXAFS portions of the equal-weight three-edge corefinement are shown in Figures 6.8 and 6.9. The quality of the fits are typical of the series and of other model alloys. Note that the three-edge corefinement does not closely fit the first peak of the PDF. This is likely due to the large oscillations in the data that have resulted from an inadequate data reduction.

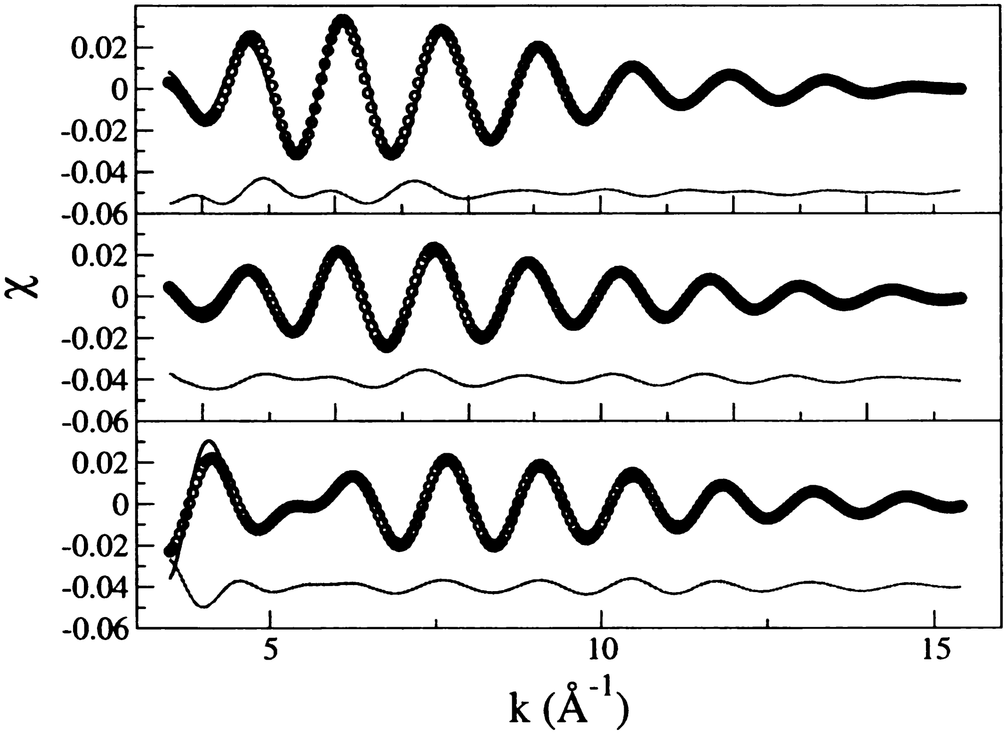


Figure 6.9: EXAFS signal of $In_{0.5}Ga_{0.5}As$ (circles), fit to EXAFS signal (thick line) and difference curve (thin line, shifted down for clarity) for a three-edge equal-weight corefinement. Top: Ga K-edge. Middle: In K-edge. Bottom: As K-edge.

Table 6.4: Nearest-neighbor bond lengths and Debye-Waller factors from corefinement.

Pair	r (Å)	σ^2 (Å^2)
	GaAs	
Ga-As	2.43690(6)	0.0022(1)
	InAs	
In-As	2.6095(1)	0.0022(2)
	$In_{0.5}Ga_{0.5}As$	
Ga-As	2.4616(2)	0.0018(2)
In-As	2.5927(2)	0.0009(4)

6.6 Discussion

The best bond length and Debye-Waller factors from the corefinements are displayed in Table 6.4. The bond length results are consistent with those reported by Mikkelsen and Boyce [10], with a nearly uniform bond contraction of 0.012(1) Å due to temperature. The large change in the Debye-Waller factors from the end compounds to the alloy is an indication that the structure model is unable to completely decorrelate the static and thermal disorder. Figures 6.4, 6.5 and 6.8 show results from the Ga/In relaxation of the corefinement results. It is apparent from Figure 6.5 that the Ga/In relaxation has little effect on the Debye-Waller factors, which suggests that As shifts have captured the effects of the static disorder. The large difference between the Ga-As and In-As Debye-Waller factors for $\mathrm{In}_{0.5}\mathrm{Ga}_{0.5}\mathrm{As}$ may suggest that the As shifts have overcompensated for the static disorder, but is not inconsistent with previous results [100]. Figure 6.4 indicates that the bond lengths change insignificantly upon relaxation of the Ga/In sublattice. In fact, the only apparent effect of the relaxation is the improved visual quality of the nearest-neighbor peak in the PDF (Figure 6.8).

The results of this corefinement study demonstrate that corefinement overcomes many of the shortcomings in traditional PDF and EXAFS analysis. In particular, the EXAFS data helps resolve PDF peaks beyond the first coordination shell. This, in turn, may help reduce the correlations that usually accompany EXAFS fits. The lack of significant correlation reduction in the $\mathrm{In}_{0.5}\mathrm{Ga}_{0.5}\mathrm{As}$ corefinements is due to the complexity of the structure model. The correlations in this case were strongly dependent on the model alloy and EXAFS data used as well as the χ^2-weights of the components. The stability of the $\mathrm{In}_{0.5}\mathrm{Ga}_{0.5}\mathrm{As}$ corefinements with respect to χ^2-weighting indicates that the correlations ultimately had little effect on the fits.

The GaAs and InAs fits show that the addition of PDF data to an EXAFS refinement has a similar effect as extending the EXAFS fit to higher coordination shells. This is of great benefit when the structure is complicated and a multi-shell EXAFS fit is infeasible, like in the case of $In_{0.5}Ga_{0.5}As$. Even when a multi-shell fit can be performed, it is likely that a combined first-shell EXAFS and PDF can achieve more consistent results out to much greater distances. This is better than a simple PDF fit, as the chemical information in the EXAFS data helps resolve ambiguous peaks in the PDF.

The corefinement results presented in the last section bring up a couple of issues. First among these is the quality of the corefinements. Based on goodness of fit measures and visual agreement, corefinements appear to be of poorer quality than the independent PDF and EXAFS fits. Likewise, for this system, corefinement does not improve predicted uncertainty over that already accessible from the PDF. The individual PDF and EXAFS fits have better agreement with the data than the corefinements because they can compensate for systematic errors in the data more effectively. However, this ultimately distorts the structural parameters in a way that is not reflected in the predicted uncertainties. The corefinements cannot compensate for errors in this way since errors from different components disparately affect the structural parameters. If the errors are uncorrelated, then they cannot all be accommodated and the corefinements appear to be of poorer quality. This specious flaw in the corefinement technique leads to more reliable structural parameters.

One way to evaluate how systematic errors affect the measured structural parameters is to vary the χ^2-weights of the components. If the results are stable with respect to weighting, then it is assumed that the errors compensate for one another and do not show up in the structural parameters. Assuming all the errors are small, they are more likely to cancel when more data is used. So, unsurprisingly, the three-edge corefinement series was stable with respect to weighting whereas the single-edge series was not.

The χ^2-weighting range that indicates the stability of a corefinement is not statically defined. By over-weighting a corefinement component, the results can be made to bias one source of data over the others. In the three-edge corefinement series, the actual EXAFS contribution to the total χ^2 was within 10% of the nominal EXAFS weighting. In the other weighting-series, the PDF component was dominant, but the EXAFS components were at least equally represented at the largest weightings. These weighting ranges are considered large enough to demonstrate stability. Note that the relative weights of the individual EXAFS components were not varied since this does not seem entirely appropriate. Weighting the EXAFS components equally emphasizes each component according to the uncertainty in the data (see Equation 5.1 and Table 6.1), which is meaningful since the fit discrepancies and data uncertainties are on the same scale.

6.7 Conclusion

Corefinement overcomes many of the limitations of PDF and EXAFS fitting, such as the lack of intermediate-range information in the EXAFS and overlapping peaks in the PDF. Corefinement may also reduce fitting correlations for simple structural models. Corefining data from multiple sources reduces the effects of systematic error on the measured structural quantities, which means that the corefined results may not be a simple interpolation between the independently obtained results. Furthermore, corefinement extends the modeling options available with EXAFS data. The effect of systematic errors on corefinement results can be tested by varying the contributions of the EXAFS and PDF data to the fit. Corefinements that are stable with respect to weighting are not greatly influenced by errors. This may be the only way to assess these effects.

Chapter 7

Ab Initio Structure Determination

7.1 Introduction

There are approaches to structure determination that do not involve structure refinement. Direct *ab initio* methods extract structural information from experimental data without a refining a model. In recent years, techniques from image processing [101] have been applied to single crystal [102] and powder [103] diffraction data. These diffraction imaging procedures reconstruct the missing phase information from diffraction intensities and result in a three-dimensional electron-density map of the material. Direct real-space methods treat structure determination as a geometrical problem. These methods can make use of the inter-atomic distances extracted a from PDF profile to assemble a structure consistent with the data. The Liga algorithm [12] is one implementation of this approach. Liga performs well on regular two- and three-dimensional monatomic structures. Another approach is implemented in the Tribond algorithm, which is designed to work on random structures. Section 7.2 describes the Liga algorithm. The Tribond algorithm is introduced and discussed in Section 7.3. Section 7.4 discusses the deficiencies of both of these algorithms and directions for future work.

7.2 Liga

The Liga algorithm assembles a complete structure by building up small populations, or divisions, of feasible sub-clusters of the structure. Each candidate cluster in a division has the same number of atoms and the divisions are differentiated by the number of atoms in the clusters. Inspired by European soccer leagues, like La Liga in Spain, clusters within a given division 'compete' to gain one or more atoms and move up to a new division. Meanwhile, losing clusters in a division lose atoms and move down to a lower division. Clusters are judged according to their agreement with the target distance list. The agreement is measured as the variance between the model and target distances, $\mathrm{var}(d) = \frac{1}{N_p}\sum_{k=1}^{N_p}(d_k^m - d_k^t)^2$, where N_p is the number of atom pairs in the cluster, d_k is the distance of pair k, and the suffixes m and t represent *model* and *target*, respectively. For a given 'tournament', the probability that a cluster is demoted down a division is proportional to its error and the probability that it is promoted to a new division is proportional to the inverse of its error. When a cluster is demoted, its atoms are judged according to their contribution to the total variance. The probability that an atom is removed is proportional to its contribution to the variance. Tournaments are held until a cluster of the desired size stands out as a clear winner.

Since the inter-atomic distances are derived from experimental data, they are subject to errors in magnitude and multiplicity. The multiplicities are found by integrating the PDF peak for a given atom pair. Liga has proven robust against errors in the bond length estimates. However, it does not perform well with strict, inaccurate multiplicities. For structures with highly degenerate sub-structures, relaxing the multiplicities of the measured bonds will often lead to a correct solution.

Liga performs well on regular structures because of the large multiplicity of identical sub-structures. This sub-structure degeneracy leads to a reduction in the number of possible clusters in a given division and increases the ratio of good to bad clusters.

When this degeneracy does not exist, Liga starts to perform like a brute-force approach. For example, Liga was tested on a Linux computer with a 2.4 GHz Pentium III processor to measure how quickly it assembles random structures. Liga was unable to assemble a random 10-atom structure within five days, but it reconstructed a carbon-60 buckyball in 30 seconds. More reconstruction times can be found in reference [12].

7.3 Tribond

The Tribond algorithm, like Liga, takes a geometrical approach to structure determination, but it is designed specifically for random structures. In its current incarnation, Tribond works with an ideal distance list generated from a random structure. It makes use of the fact that any overconstrained cluster of zero variance is very likely to be part of the random cluster. The strategy of the Tribond algorithm is to find such an overconstrained cluster and build the remainder of the structure around it.

7.3.1 Tribond in Two Dimensions

In two dimensions, the smallest overconstrained cluster contains four atoms and six bonds. It is overconstrained in the sense that removal of one its bonds leaves the cluster rigid, that is, the cluster cannot be sheared without stretching a bond. The goal of the two-dimensional Tribond algorithm is to construct such a cluster from distances taken from the target structure. This is realized by assembling edge-sharing triangles made from the distance list and verifying that the sixth distance, the bridging bond, appears in the distance list. This is displayed in Figure 7.1. Since the distances comprising a random structure are mostly unique and uncorrelated, it is very likely that this core cluster is part of the structure.

The core-finding portion of Tribond is performed by fixing a base triangle and

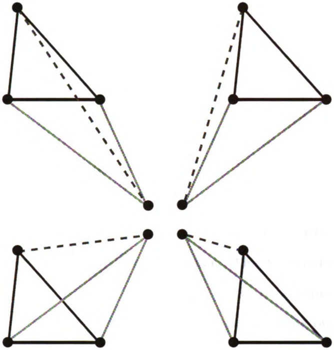


Figure 7.1: Possible clusters composed of two edge-sharing triangles. The dark horizontal line is the base bond and the dark triangle is the base triangle. The lighter triangle is the candidate. The broken line is the bridging bond.

searching through all other candidate triangles that share the same base bond until a bridging bond is found. If all candidate triangles are exhausted, then a new base triangle is selected and the procedure is continued. In practice, the distance list is sorted and the base bond is taken to be the shortest distance in the distance list. In addition, the feasible triangles are generated so that a base-candidate pair of triangles is never compared more than once. For a given pair of triangles, there are four relative orientations that must be checked. These are demonstrated in Figure 7.1. The points are positioned in space using triangulation. For a given bridging distance, a half-interval search is performed on the bond list to determine if the bridging bond is in the distance list.

Once the core is found, the second stage of the algorithm continues to cycle through the unchecked candidate triangles and records whenever a new bridge bond is found. At this stage, the base triangle is known to be part of the structure, so it does not need to be replaced. The structure is complete once all target distances are exhausted.

The core search is the most time consuming part of the algorithm since it must search over pairs of triangles. The base triangle is fixed once the core is found, which makes the second stage much faster than the first. In the worst case, the number of triangle pairs sharing a base bond grows like $\mathcal{O}\left(N^8\right)$, for an N-atom cluster. However, there are $\mathcal{O}\left(N^4\right)$ such pairs in the cluster, so on average Tribond finds the two-dimensional core in $\mathcal{O}\left(N^5 \log(N)\right)$ time. This is includes the search time of the bridging bond ($\mathcal{O}(N \log(N))$) in addition to the expectation time for finding the core ($\mathcal{O}\left(N^4\right)$). Tribond achieves faster times by searching over a restricted set of distances. The algorithm tries to construct the core with the smallest distances possible, based on the assumption that a 'small' core exists in the structure. The distance list is restricted to the first 50 distances, and then 100, and so on until a core is found. This procedure shows a considerable speedup for larger structures, but the amount of speedup has not been quantified.

An example of a successful reconstruction is shown in Figure 7.2. Note that the core is not necessarily the smallest subcluster of the structure. This is because Tribond constructs the core so that it contains the smallest bond in the structure. Running on a Linux computer with a 2.4 GHz Pentium III processor, Tribond was able to assemble this structure in about $0.25s$. On the same computer, Tribond can usually reconstruct other structures of this size in less than a second.

Tribond performs reasonably well on structures with hundreds of atoms and can reconstruct a 1000-atom cluster within a few hours on a desktop computer. In tests with randomly generated structures it was found that Tribond has around a 9% failure rate on 10-atom structures and around 19% on 100-atom structures. The failures are highly correlated with the presence of sites colinear with the base bond. Figure 7.3

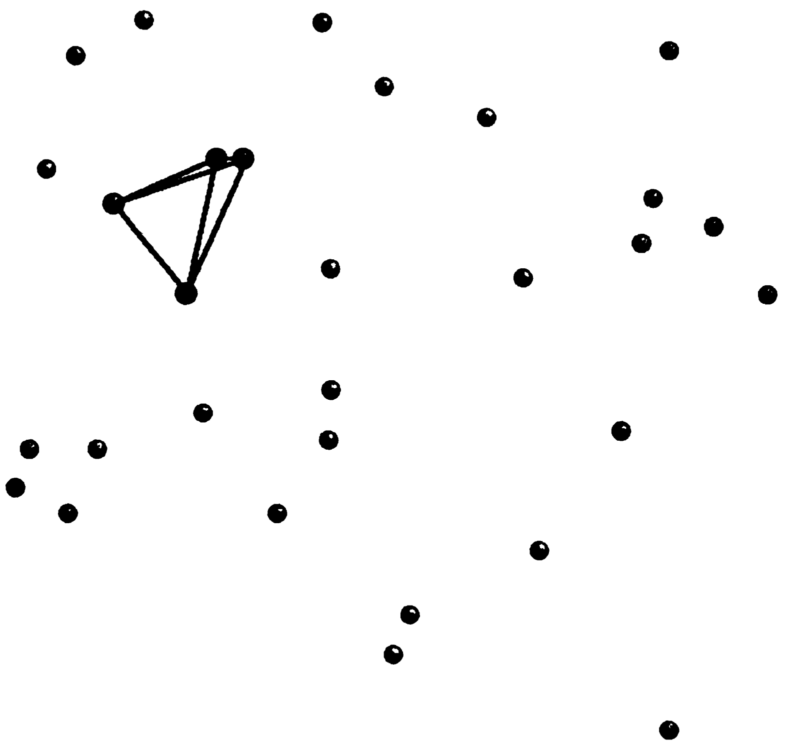


Figure 7.2: Random two-dimensional cluster with 30 sites. The core is indicated by the connecting bonds. Tribond assembled this structure from the distance list in less than a second.

shows a two-dimensional structure that Tribond failed to assemble. Due to the finite precision of the triangulation, the placement of sites can only be performed within a small tolerance. This affects the precision with which the bridging bond can be determined. This error cancels on average, but it accumulates in the placement of colinear sites, which causes Tribond to disregard valid candidate triangles.

7.3.2 Tribond in Three Dimensions

The three-dimensional Tribond algorithm works much like the two-dimensional version, except that it assembles triangle-sharing tetrahedra into five-atom clusters. Much as before, the tenth bond in the cluster is compared to distance list to determine whether or not the cluster is part of the structure. This three-dimensional

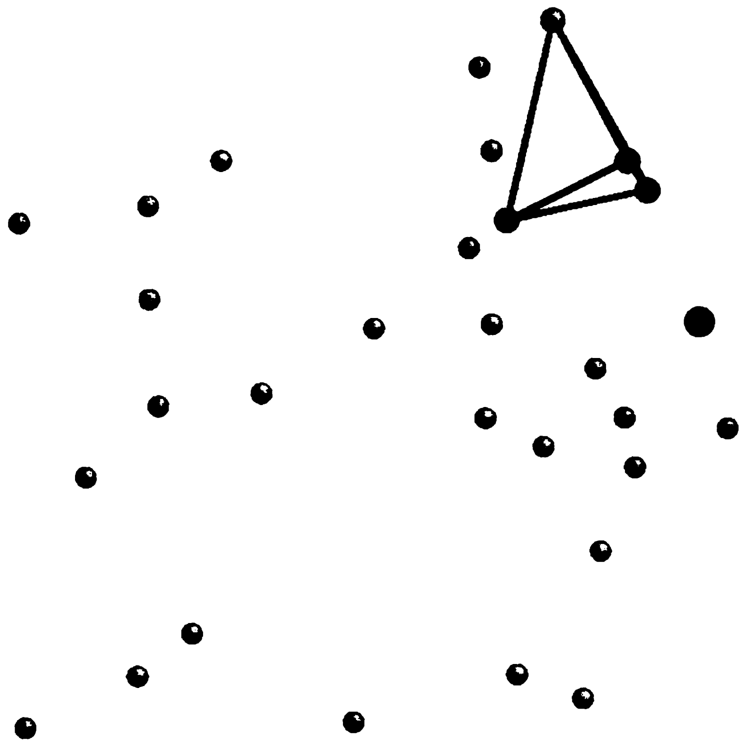


Figure 7.3: Random two-dimensional cluster with 30 sites and nearly colinear points. The core is indicated by the connecting bonds. Tribond could not place the large, dark site, and therefore failed to fully assemble the structure.

procedure must overcome some difficulties not found in the two-dimensional case. For example, even though the tetrahedral pair share a base triangle, that triangle cannot be fixed like the base bond in two dimensions. The number of tetrahedral pairs in three dimensions that share a side grows like $\mathcal{O}\left(N^{14}\right)$ with the number of sites in the structure. There is no way to know if the bonds in the five-atom cluster, including those in the base triangle, are correctly connected until the bridging bond is tested. Thus, the number of base triangles sharing a common bond must also be included in this number. This gives a total of $\mathcal{O}\left(N^{18}\right)$ possible combinations of tetrahedral pairs.

An additional complication comes from generating the tetrahedra that are used as construction units. The number of possible possible tetrahedra that share a common bond grows like $\mathcal{O}\left(N^{10}\right)$. This is ignoring physicality constraints like the applicability of the triangle inequality and positive volume. Regardless, even for a moderately sized structure, there are too many tetrahedra to store on most desktop computers. Thus, the tetrahedra must be constructed on demand rather than stored in memory. This imposes a small amount of computation overhead on the algorithm, as some tetrahedra may be constructed multiple times.

As in the two-dimensional case, the core-finding part of the algorithm is the most time consuming part. There are on the order of $\mathcal{O}\left(N^{10}\right)$ possible tetrahedra that use the base bond, so the average core-search time grows like $\mathcal{O}\left(N^9 \log(N)\right)$, including the search time for the bridging bond between each tetrahedral pair. Bond restriction works to speed this up, but not by the orders of magnitude necessary to make the algorithm generally applicable. It can take up to an hour to reconstruct a 10-atom cluster and several days to construct a 100-atom cluster.

7.4 Future Work

The Liga and Tribond algorithms so far only work on monatomic structures. Current research is under way to include chemical information into Liga. When a solution is found, it is likely to translate to the Tribond algorithms. The obvious fault of the three-dimensional Tribond algorithm is its lack of speed. Work is underway to speed up the algorithm. Speed schemes are focused on selecting a base tetrahedron and a bridging bond and doing an efficient search over the candidate tetrahedra based on some criteria. A suitable sorting criterion is yet to be identified. Since a list of tetrahedra is too large to store, this criterion will have to be incorporated into the tetrahedron generator.

The high failure rate in the Tribond algorithm is an issue that must be overcome. Since the failures are believed to be an issue with precision, arbitrary precision arithmetic could help overcome the problem. Unfortunately, this solution would likely slow the algorithm. The precision issues are most likely to appear after the core is already assembled, which means that any solution will be hard to test on the three-dimensional algorithm until it can be streamlined. On a related issue, Tribond has yet to be tested on noisy data. The current solution of comparing bonds, which is tolerant of small differences between compared values, may be easily adopted to deal with noise.

A complete *ab initio* geometrical structural solution will need to handle regular and random structures, and all things in between. Work has begun applying Tribond to regular structures by considering only the unique distances in the structure. Another approach to applying Tribond to regular structures is to perturb the distances until the structure seems "random enough". These strategies may make Tribond applicable to regular structures, but it is likely to be much more applicable to non-crystalline materials, such a proteins. A Liga-Tribond hybrid is a likely solution for general *ab initio* structure assembly.

Chapter 8

Advances in PDF analysis

8.1 Introduction

Although the PDF technique is gaining in popularity, it has a much smaller user community than EXAFS and traditional diffraction techniques. This is largely due to its relatively recent emergence as a tool for nanostructure determination. The technique is just now reaching its maturity and this is partly due to recent experimental [66] and analytical [71, 24, 104] advances in the technique. The technique is also being advanced with new software. Projects, such as the Distributed Data Analysis for Neutron Scattering Experiments (DANSE) [105] project, are driven by the idea that computational tools are essential to scientific advancement.

The DANSE project is a collaboration between several major universities and institutions to modernize data analysis software for neutron diffraction science. The goals of the project are to provide software that enables new and more sophisticated science to be performed with neutron scattering experiments, makes the analysis of data easier for all scientists and provides a robust software infrastructure that can be maintained in the future. The DANSE diffraction subproject is tasked with developing real- and reciprocal-space analysis tools.

The goal of the DANSE diffraction subproject is to provide software analysis tools for neutron diffraction neophytes and experts alike. This will include structure refinement applications as well as *ab initio* structure determination routines, like those discussed in the previous chapters. Future projects of the DANSE diffraction subproject will extend and improve upon corefinement, as discussed in Chapter 5, to include the other neutron scattering techniques. The tools developed by the diffraction subproject are collectively referred to as DIFFPY [106]. DIFFPY currently encompasses two applications, PDFFIT2 and PDFGUI, which are applications for PDF studies. These applications represent the state-of-the art in PDF analysis, and are described in the following section.

8.2 PDFfit2 and PDFgui

PDFFIT2 [14] is a major upgrade to PDFFIT [13], a widely-adopted application for structural refinement from PDF data. PDFGUI is a graphical user-interface (GUI) built to use PDFFIT2. These applications serve as a prototype for future projects of the DANSE diffraction group. They embody many of the principles of the DANSE project, including software reusability through object-oriented design and extensibility. The design of PDFFIT2 and PDFGUI was largely a team effort. I am responsible for the design of the GUI and macros. I also made minor contributions to the PDFFIT2 engine and the interpreter that translates GUI directives to the engine.

PDFFIT2 inherits many features from PDFFIT. PDFFIT is capable of fitting a theoretical three-dimensional structure to an experimentally determined PDF. It can simultaneously fit multiple structures, accounting for different structural phases in a material. PDFFIT has a constraint system that allows expressing structure variables as simple functions of fitted parameters. PDFFIT structure variables include lattice constants, data and phase scale factors, atomic site occupation, anisotropic displacement factors, and atomic vibrational correlations. PDFFIT has a built-in FORTRAN-style command language that understands simple `for` loops and some built in arithmetic functions.

The original PDFFIT was written in FORTRAN-77, which imposes some limitations on the program. For example, it uses fixed-size arrays for internal storage. This precludes the analysis of structures with large cells without modifying the code. Though the constraint system is powerful, it requires that a constraint equation be accompanied by its first derivative. This places the burden of determining the derivatives on the user, which can introduce errors that lead to instability in the convergence. Furthermore, the code is monolithic, not easily extensible and hard to integrate with external programs.

The primary focus of PDFFIT2 development was to remedy the limitations of PDFFIT while extending its functionality. The old PDFFIT engine has been completely rewritten in C++, and many bugs have been fixed. The new engine uses dynamic memory allocation so that the size of the structure or extent of the fit-range of the PDF is limited only by the physical memory available. The constraint system has also been upgraded. The program *automatically* computes the analytical derivatives of the constraints that are required by the minimization routine. This simplifies user input and reduces the possibility of errors. In addition, recent theoretical developments have been incorporated into the new engine, including nanoparticle form factors.

Instead of rewriting the PDFFIT command interpreter, which is used to define the fitting problem and to control and run the refinement, its functions are carried out using the Python language [107]. Python is a powerful, cross-platform, open-source interpreted[a] programming language that emphasizes object-oriented and modular design. PDFFIT2 scripts written in Python syntax take the place of PDFFIT macros

---

[a] Python code does not need to be compiled to run, similar to scripting languages like Perl.

and the Python interpreter can handle everything that the old interpreter could, and more. Using Python as an interpreter allows PDFFIT2 to be combined with and enhanced by other Python libraries. The PDFFIT2 engine can be used either directly with scripts or from the Python command line, or as part of larger and more complex software applications.

8.2.1 Design Principles

PDFGUI has been designed to provide users with an easy-to-use yet powerful interface for fitting structure models to PDF data. It makes use of an object oriented, component based architecture, which makes it highly extensible and maintainable. This allows for powerful usability features such as real-time plotting and remote execution of the fitting program whilst visualizing the results locally. PDFGUI is multi-threaded so that the work being done by the PDFFIT2 engine does not interfere with the tasks of the user interface.

Like PDFFIT2, PDFGUI is written in the Python programming language. Many Python libraries were leveraged in the construction of PDFGUI. Software codes written in a variety of programming languages can be bound into Python, which allows them to be used together. PDFGUI's interface is built using wxPython [108], the Python package for wxWidgets, a mature open-source cross-platform GUI library.

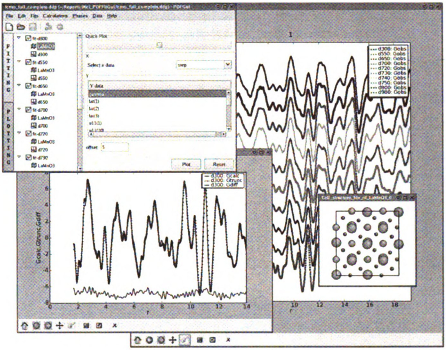
8.2.2 Capabilities

PDFGUI contains most of the functionality of PDFFIT2 along with additional enhancements for usability. PDFGUI can manage multiple fits at once. Each fit can have multiple experimental data sets and structure models. Fits in a sequence can call upon other fits for their starting parameters, and configured fits can be queued to run while the user is away. All the initial, final, and intermediate data are stored in a platform independent project file that can be loaded on any computer. All management tasks, such as fit creation, configuration, modification, and visualization, can be done through the graphical interface.

PDFGUI supports space group operations. Users can define an asymmetric unit and let PDFGUI expand it to a full cell with all symmetry related positions. PDFGUI can also generate symmetry constraints for atom positions and anisotropic displacement parameters. Users just need to specify the space group, and the program will identify equivalent sites and generate constraint equations for their coordinates and temperature factors to keep the structure consistent with the symmetry requirements. This can be done either for all atoms in the structure or for an arbitrary subset - for example when it is known that only a certain species shows a local distortion. The code for space group definitions was provided by the Python Macromolecular Library (mmLib) [109]. PDFGUI also supports supercell expansion of a normal unit cell.

PDFGUI uses the Matplotlib [110] Python package for two-dimensional plotting of data and results. Matplotlib has a friendly interface so the user can quickly and easily view the results of a fitting. PDFGUI lets users plot data from a series of fits and plot it against selected metadata (temperature, doping, etc.), plot the results of several fits in the same window, plot the PDF in real time as the fitting is running, plot the parameters or variables in real time as the refinement evolves, and save plots in common image formats or export the data to a text file. PDFGUI uses AtomEye [111] for three-dimensional visualization.

PDFGUI supports built-in macros for advanced fits. For example for a set of experimental data from one system at different temperatures or doping levels, PDFGUI can expand a template fit to a series of related fits. Another PDFGUI macro makes it easy to set up boxcar fits, where the same model is fit over different r-ranges of the PDF data. These macros are capable of reducing days of script configuration and data mining to a matter of minutes [112].



Select x data: step

offset: 5

Figure 8.1: The PDFGUI plot window, plots of the $LaMnO_3$ PDF fit at 300 K and the data at all temperatures, and the refined structure at 300 K.

8.2.3 Example

The capabilities of PDFGUI are demonstrated here on an example fitting of a temperature series of neutron PDF data from $LaMnO_3$. $LaMnO_3$ has a perovskite structure consisting of corner-shared MnO_6 octahedral units [113, 114]. At room temperature this material takes on orthorhombic symmetry (space-group *Pbnm*) where the Jahn-Teller distorted MnO_6 octahedra contain four short and two long Mn-O bonds and the Jahn-Teller long-bonds are ordered in space in a checker-board fashion [22]. As temperature is increased through 750 K, a structural phase transition occurs to a pseudo-cubic phase, without a change in space group [115]. Using neutron powder diffraction data collected at the NPDF diffractometer [116] at the Lujan Center at Los Alamos National Laboratory, PDFGUI was used to create a temperature series fit to investigate the orthorhombic to pseudo-cubic phase transition [75]. The average crystallographic structure for $LaMnO_3$ was used as a starting point for the fits.

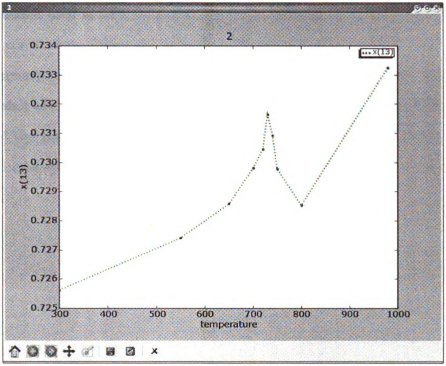


Figure 8.2: The full PDFGUI plot window showing the fractional x-coordinate of an axial oxygen versus temperature.

Each fit in the series was part of the same project and configured identically. The La and Mn atoms were constrained to have isotropic thermal factors as were the axial and polar oxygen atoms. Using the PDFFIT2 constraint mechanism allows one to explicitly include known correlations between the physical parameters of the fit and leads to less uncertainty in the resulting refinement. To speed up the convergence of the fits, the starting values used for a given temperature were taken from the converged values of the previous temperature in the series. Chaining fits together in such a way places the fitting parameters in the basin of attraction for the minimum in the parameter space. Figure 8.1 shows the fit organization and plot interface of PDFGUI, a stacked plot of all of the project data, a plot of the 300 K fit results, and the refined structure from the 300 K fit.

One of the goals of this series of fits was to track the orientation of the MnO_6 octahedra. A major strength of PDFGUI is the ability to plot any structure variable or a parameter used in a fit. This allows users to easily and creatively investigate the complex correlations among the fit parameters, without having to manually extract data from output files. An example of this can be seen in Figure 8.2. The figure shows a screen shot of a PDFGUI plot window displaying the refined fractional x-coordinate of one of the planar oxygens. The expected structural phase transition is clearly indicated by a spike in the value of this parameter at the transition temperature.

Chapter 9

Conclusion

This Dissertation describes new and improved analytical tools for nanostructure determination. It is hoped that the algorithms and software discussed herein will find a place in the future of this rapidly growing field. A unifying retrospective may help bring this possibility into focus.

PDF and EXAFS corefinement is hopefully the first of many endeavors to effectively combine data from different techniques. The corefinement scheme detailed in Chapter 5 is general enough to be applied to any structure analysis technique that has data and an objective function with structural parameters. The parallel refinement of data from many sources has the potential to drastically shorten the time between the discovery and structural characterization of new functional materials. The current rate limiting step in this procedure is the necessary consensus that must be shared by the experts in the PDF, EXAFS and crystallographic communities. Corefinement produces results that are consistent with the data from different techniques. This provides certainty that characteristics of a derived structural model are likely due to the model, rather than the errors in the data or deficiencies in the technique.

Liga and Tribond are two among many model-independent structure determination algorithms. Model-independent algorithms are important because they require little input from the investigator to determine a structure. In PDF and EXAFS analysis, for example, the investigator is likely to spend considerable time developing and testing a starting structure and constraint model. Specific starting guesses are usually not required by these algorithms. Unfortunately, Liga and Tribond do not give any insight into the vibrational or chemical properties of a material. In this sense, the structures produced by these algorithms are incomplete. These structures are well suited as starting models for structure refinement routines. Furthermore, since the algorithms are purely geometrical in nature, they are applicable to a wider range of problems than structure determination.

PDFFIT2 and PDFGUI represent a change in the way nanostructure analysis is being performed. For the hours to days that are spent at beam lines collecting sensitive data, days to weeks can be spent analyzing that data. National facilities have beamline scientists who aid users in data collection. Scientist are very fortunate if there are experts that are willing to help them learn an analysis program. Even though user facilities are on a much grander scale than analysis programs, when it comes to complexity and necessary resources, it is easy to argue that these factors affect the efficiency of data analysis. Time spent programming usability and extensibility features into these programs will eventually get paid back many times over by their users.

Unlike many analysis programs that are a labor of love for one or a few physicists, PDFFIT2 and PDFGUI have stable financial and community support. The design of these programs also sets them apart from other popular analysis software packages. Software development within the DANSE project is managed using tools and techniques from the commercial software industry. Modern software development techniques will help assure that the programs remain useful for many years. In addition, programs are open-source. Users are allowed and encouraged to modify and improve the software.

The programs and algorithms described in this Dissertation represent small steps towards complex modeling [85]. Complex modeling will go beyond corefinement by incorporating various experimental, theoretical and modeling techniques. For example, Liga or Tribond may generate initial structures for corefinement algorithms that are self-consistently fed into molecular dynamics simulations. Complex modeling is still many years off, as it requires tighter collaboration among all those who contribute to the structural characterization community, especially including computer scientists. Projects like DANSE foster such an environment of collaboration and allow small steps like these to be taken.

BIBLIOGRAPHY

Bibliography

[1] A. J. Millis. Lattice effects in magnetoresistive manganese perovskites. *Nature*, 392(6672):147–150, March 1998.

[2] P. G. Radaelli, D. E. Cox, M. Marezio, S-W. Cheong, P. E. Schiffer, and A. P Ramirez. Simultaneous structural magnetic, and electronic transitions in $La_{1-x}Ca_xMnO_3$ with $x = 0.25$ and 0.50. *Phys. Rev. Lett.*, 75(24):4488–4491, December 1995.

[3] P. G. Radaelli, G. Iannone, M. Marezio, H. Y. Hwang, S. W. Cheong, J. D. Jorgensen, and D. N. Argyriou. Structural effects on the magnetic and transport properties of perovskite $A_{1-x}A_x$'MnO_3 (x=0.25, 0.30). *Phys. Rev. B*, 56(13):8265–8276, 1997.

[4] S. Jin, T. H. Tiefel, M. McCormack, R. A. Fastnacht, R. Ramseh, and L. H. Chen. Thousandfold change in resistivity in magnetoresistive La-Ca-Mn-O films. *Science*, 264(5157):413–415, 1994.

[5] C. Meneghini, D. Levy, S. Mobilio, M. Ortolani, M. Nuñez Reguero, A. Kumar, and D. D. Sarma. High-pressure structure and electronic transport in hole-doped $La_{3/4}Ca_{1/4}MnO_3$ perovskites. *Phys. Rev. B*, 65(1):012111, 2002.

[6] S. J. L. Billinge, R. G. DiFrancesco, G. H. Kwei, J. J. Neumeier, and J. D. Thompson. Direct observation of lattice polaron formation in the local structure of $La_{1-x}Ca_xMnO_3$. *Phys. Rev. Lett.*, 77(4):715–718, July 1996.

[7] C. H. Booth, F. Bridges, G. H. Kwei, J. M. Lawrence, A. L. Cornelius, and J. J. Neumeier. Lattice effects in $La_{1-x}Ca_xMnO_3$ ($x = 0 \rightarrow 1$): Relationships between distortions, charge distribution, and magnetism. *Phys. Rev. B*, 57(17):10440–10454, 1998.

[8] E. S. Božin, M. Schmidt, A. J. DeConinck, G. Paglia, J. F. Mitchell, T. Chatterji, P. G. Radaelli, Th. Proffen, and S. J. L. Billinge. Understanding the insulating phase in colossal magnetoresistance manganites: Shortening of the Jahn-Teller long-bond across the phase diagram of $La_{1-x}Ca_xMnO_3$. *Phys. Rev. Lett.*, 98(13):137203, 2007.

[9] J C Woolley and B A Smith. Solid solution in A $^{III}B^{V}$ compounds. *Proc. Phys. Soc. London*, 72(2):214–223, 1958.

[10] J. C. Mikkelsen and J. B. Boyce. Atomic-scale structure of random solid solutions: Extended x-ray-absorption fine-structure study of $\mathrm{Ga}_{1-x}\mathrm{In}_x\mathrm{As}$. *Phys. Rev. Lett.*, 49(19):1412–1415, Nov 1982.

[11] V. Petkov, I.-K. Jeong, J. S. Chung, M. F. Thorpe, S. Kycia, and S. J. L. Billinge. High real-space resolution measurement of the local structure of $\mathrm{Ga}_{1-x}\mathrm{In}_x\mathrm{As}$ using x-ray diffraction. *Phys. Rev. Lett.*, 83(20):4089–4092, Nov 1999.

[12] P. Juhas, D. M. Cherba, P. M. Duxbury, W. F. Punch, and S. J. L. Billinge. Ab initio determination of solid-state nanostructure. *Nature*, 440:655–658, 2006.

[13] T. Proffen and S. J. L. Billinge. PDFFIT, a program for full profile structural refinement of the atomic pair distribution function. *J. Appl. Crystallogr.*, 32:572–575, 1999.

[14] C. L. Farrow, P. Juhas, J. W. Jiu, D. Bryndin, E. S. Božin, J. Bloch, Th. Proffen, and S. J. L. Billinge. PDFfit2 and PDFgui: Computer programs for studying nanostructure in crystals. *J. Phys.: Condens. Matter*, 19:335219, 2007.

[15] S. J. L. Billinge and M. G. Kanatzidis. Beyond crystallography: the study of disorder nanocrystallinity and crystallographically challenged materials. *Chem. Commun.*, 7:749–760, 2004.

[16] Neil W. Ashcroft and N. David Mermin. *Solid State Physics*. Brooks/Cole, 1976.

[17] H. M. Rietveld. A profile refinement method for nuclear and magnetic structures. *J. Appl. Cryst.*, 2:65–71, 1969.

[18] B. E. Warren. *X-ray diffraction*. Dover, New York, 1990.

[19] S. J. L. Billinge and T. Egami. Short-range atomic structure of $\mathrm{Nd}_{2-x}\mathrm{Ce}_x\mathrm{CuO}_{4-y}$ determined by real-space refinement of neutron-powder-diffraction data. *Phys. Rev. B*, 47(21):14386–14406, June 1993.

[20] S. J. L. Billinge, G. H. Kwei, and H. Takagi. Local octahedral tilts in $\mathrm{La}_{2-x}\mathrm{Ba}_x\mathrm{CuO}_4$: Evidence for a new structural length scale. *Phys. Rev. Lett.*, 72(14):2282–2285, April 1994.

[21] D. Louca, T. Egami, E. L. Brosha, H. Röder, and A. R. Bishop. Local Jahn-Teller distortion in $\mathrm{La}_{1-x}\mathrm{Sr}_x\mathrm{MnO}_3$ observed by pulsed neutron diffraction. *Phys. Rev. B*, 56(14):R8475–R8478, 1997.

[22] Th. Proffen, R. G. DiFrancesco, S. J. L. Billinge, E. L. Brosha, and G. H. Kwei. Measurement of the local Jahn-Teller distortion in $\mathrm{LaMnO}_{3.006}$. *Phys. Rev. B*, 60:9973, 1999.

[23] S. J. L. Billinge, Th. Proffen, V. Petkov, J. L. Sarrao, and S. Kycia. Evidence for charge localization in the ferromagnetic phase of $La_{1-x}Ca_xMnO_3$ from high real-space-resolution x-ray diffraction. *Phys. Rev. B*, 62(2):1203–1211, 2000.

[24] B. Gilbert, F. Huang, H. Zhang, G. A. Waychunas, and J. F. Banfield. Nanoparticles: Strained and stiff. *Science*, 305:651–654, 2004.

[25] H Lin, E. S. Božin, S. J. L. Billinge, Eric Quarez, and M. G. Kanatzidis. Nanoscale clusters in the high performance thermoelectric $AgPb_mSbTe_{m+2}$. *Phys. Rev. B.*, 72:174113, 2005.

[26] H. Lin, E. Božin, S. J. L. Billinge, J. Androulakis, C. H. Lin, and M. G. Kanatzidis. Phase separation and nanostructuring in the thermoelectric material $PbTe_{1-x}S_x$. *Phys. Rev. B*, 2007. submitted.

[27] S. Vensky, L. Kienle, R. E. Dinnebier, A. S. Masadeh, S. J. L. Billinge, and M. Jansen. The real structure of Na_3BiO_4 by electron microscopy, HR-XRD and PDF analysis. *Z. Kristallogr.*, 220:231–244, 2005.

[28] V. Petkov, S. J. L. Billinge, P. Larson, S. D. Mahanti, T. Vogt, K. K. Rangan, and M. G. Kanatzidis. Structure of nanocrystalline materials using atomic pair distribution function analysis: Study of $LiMoS_2$. *Phys. Rev. B*, 65(9):092105, 2002.

[29] S. J. L. Billinge, E. J. McKimmy, M. Shatnawi, H. J. Kim, V. Petkov, D. Wermeille, and T. J. Pinnavaia. Mercury binding sites in thiol-functionalized mesostructured silica. *J. Am. Chem. Soc.*, 127(23):8492–8498, June 2005.

[30] M. G. Tucker, A. L. Goodwin, M. T. Dove, D. A. Keen, S. A. Wells, and J. S. O. Evans. Negative thermal expansion in ZrW_2O_8: Mechanisms, rigid unit modes, and neutron total scattering. *Phys. Rev. Lett.*, 95(25):255501, December 2005.

[31] E. S. Božin. Unpublished data. 2007.

[32] P. F. Peterson, E. S. Božin, Th. Proffen, and S. J. L. Billinge. Improved measures of quality for atomic pair distribution functions. *J. Appl. Crystallogr.*, 36:53, 2003.

[33] M.F. Thorpe, V.A. Levashov, M. Lei, and S.J.L. Billinge. *From Semiconductors to Proteins: Beyond the Average Structure*, chapter Notes on the Analysis of Data for Pair Distribution Functions, pages 105-128. Kluwer Academic/Plenum Publishers, New York, 2002.

[34] G. Beni and P. M. Platzman. Temperature and polarization dependence of extended x-ray absorption fine-structure spectra. *Phys. Rev. B*, 14(4):1514–1518, 1976.

[35] I. K. Jeong, T. Proffen, F. Mohiuddin-Jacobs, and S. J. L. Billinge. Measuring correlated atomic motion using x-ray diffraction. *J. Phys. Chem. A*, 103(7):921–924, 1999.

[36] I. K. Jeong, R. H. Heffner, M. J. Graf, and S. J. L. Billinge. Lattice dynamics and correlated atomic motion from the atomic pair distribution function. *Phys. Rev. B*, 67(10):104301, 2003.

[37] B. H. Toby and T. Egami. Accuracy of pair distribution function analysis applied to crystalline and non-crystalline materials. *Acta. Cryst. A*, 48:336–345, 1992.

[38] D. C. Koningsberger and R. Prins, editors. *X-Ray Absorption: Principles, Applications, Techniques of EXAFS, SEXAFS AND XANES*, volume 92 of *Chemical Analysis: A Series of Monographs on Analytical Chemistry and its Applications*. Wiley (New York), 1988.

[39] J. J. Rehr and R. C. Albers. Theoretical approaches to x-ray absorption fine structure. *Rev. Mod. Phys.*, 72(3):621–654, July 2000.

[40] D. E. Sayers, E. A. Stern, and F. W. Lytle. New technique for investigating noncrystalline structures: Fourier analysis of extended x-ray absorption fine structure. *Phys. Rev. Lett.*, 27(18):1204, 1971.

[41] W. L. Schaich. Comment on theory of extended x-ray-absorption fine-structure. *Phys. Rev. B*, 8(8):4028–4032, 1973.

[42] P. H. Gaskell, M. C. Eckersley, A. C. Barnes, and P. Chieux. Medium-range order in the cation distribution of a calcium silicate glass. *Nature*, 350(6320):675–677, April 1991.

[43] L. Cormier, G. Calas, and P. H. Gaskell. Cationic environment in silicate glasses studied by neutron diffraction with isotopic substitution. *Chem. Geol.*, 174(1-3):349–363, April 2001.

[44] H. Stragier, J. O. Cross, J. J. Rehr, Larry B. Sorensen, C. E. Bouldin, and J. C. Woicik. Diffraction anomalous fine structure: A new x-ray structural technique. *Phys. Rev. Lett.*, 69(21):3064–3067, Nov 1992.

[45] M. G. Proietti, H. Renevier, J. L. Hodeau, J. García, J. F. Bérar, and P. Wolfers. Diffraction-anomalous-fine-structure spectroscopy applied to the study of III-V strained semiconductors. *Phys. Rev. B*, 59(8):5479–5492, Feb 1999.

[46] U. Hoppe, R. Kranold, A. Barz, D. Stachel, J. Neuefeind, and D. A. Keen. Combined neutron and x-ray scattering study of phosphate glasses. *J. Non-Cryst. Solids*, 293:158 168, November 2001.

[47] D. E. Sayers and B. A. Bunker. *X-Ray Absorption: Principles, Applications, Techniques of EXAFS, SEXAFS AND XANES*, volume 92 of *Chemical Analysis: A Series of Monographs on Analytical Chemistry and its Applications*, chapter 6, pages 211–253. Wiley (New York), 1988.

[48] J. W. Cook, Jr. and D. E. Sayers. Criteria for automatic x-ray absorption fine structure background removal. *J. Appl. Phys.*, 52(8):5024–5031, 1981.

[49] Steven M. Heald. *X-Ray Absorption: Principles, Applications, Techniques of EXAFS, SEXAFS AND XANES*, volume 92 of *Chemical Analysis: A Series of Monographs on Analytical Chemistry and its Applications*, chapter 3, pages 87–118. Wiley (New York), 1988.

[50] Edward A. Stern. *X-Ray Absorption: Principles, Applications, Techniques of EXAFS, SEXAFS AND XANES*, volume 92 of *Chemical Analysis: A Series of Monographs on Analytical Chemistry and its Applications*, chapter 1, pages 3–51. Wiley (New York), 1988.

[51] P. A. Lee and J. B. Pendry. Theory of extended x-ray absorption fine-structure. *Phys. Rev. B*, 11(8):2795–2811, 1975.

[52] C. A. Ashley and S. Doniach. Theory of extended x-ray absorption-edge fine-structure (EXAFS) in crystalline solids. *Phys. Rev. B*, 11(4):1279–1288, 1975.

[53] Albert Messiah. *Quantum Mechanics: Two Volumes Bound as One.* Dover, Mineola, New York, 1999.

[54] J. J. Rehr, R. C. Albers, and S. I. Zabinsky. High-order multiple-scattering calculations of x-ray-absorption fine structure. *Phys. Rev. Lett.*, 69(23):3397–3400, Dec 1992.

[55] A. V. Poiarkova and J. J. Rehr. Multiple-scattering x-ray-absorption fine-structure Debye-Waller factor calculations. *Phys. Rev. B*, 59(2):948–957, 1999.

[56] Grant Bunker. Application of the ratio method of EXAFS analysis to disordered systems. *Nucl. Instrum. Methods*, 207(3):437–444, April 1983.

[57] J. M. Tranquada and R. Ingalls. Extended x-ray-absorption fine-structure study of anharmonicity in CuBr. *Phys. Rev. B*, 28(6):3520–3528, Sep 1983.

[58] Joachim Stöhr. *X-Ray Absorption: Principles, Applications, Techniques of EXAFS, SEXAFS AND XANES*, volume 92 of *Chemical Analysis: A Series of Monographs on Analytical Chemistry and its Applications*, chapter 10, pages 443–571. Wiley (New York), 1988.

[59] S. I. Zabinsky, J. J. Rehr, A. Ankudinov, R. C. Albers, and M. J. Eller. Multiple-scattering calculations of x-ray-absorption spectra. *Phys. Rev. B*, 52(4):2995–3009, 1995.

[60] A. L. Ankudinov and J. J. Rehr. Relativistic calculations of spin-dependent x-ray-absorption spectra. *Phys. Rev. B*, 56:R1712 – R1715, 1997.

[61] Edward A. Stern. Number of relavent independent points in x-ray-absorption fine-structure spectra. *Phys. Rev. B*, 48(13):9825, October 1993.

[62] Matt Newville. IFEFFIT: interactive EXAFS analysis and FEFF fitting. *J. Synchrotron Rad.*, 8:322—324, 2001.

[63] B. Ravel and M. Newville. ATHENA, ARTEMIS, HEPHAESTUS: data analysis for x-ray absorption spectroscopy using IFEFFIT. *J. Synchrotron Rad.*, 12(4):537–541, July 2005.

[64] E. D. Crozier, J. J. Rehr, and R. Ingalls. *X-Ray Absorption: Principles, Applications, Techniques of EXAFS, SEXAFS AND XANES*, volume 92 of *Chemical Analysis: A Series of Monographs on Analytical Chemistry and its Applications*, chapter 9, pages 373 – 442. Wiley (New York), 1988.

[65] Th. Proffen, T. Egami, S. J. L. Billinge, A. K. Cheetham, D. Louca, and J. B. Parise. Building a high resolution total scattering powder diffractometer - upgrade of NPD at MLNSC. *Appl. Phys. A*, 74:s163–s165, 2002.

[66] P. J. Chupas, X. Qiu, J. C. Hanson, P. L. Lee, C. P. Grey, and S. J. L. Billinge. Rapid-acquisition pair distribution function (RA-PDF) analysis. *J. Appl. Crystallogr.*, 36:1342–1347, 2003.

[67] P. W. Loeffen, R. F. Pettifer, S. Müllender, M. A. van Veenendaal, J. Röhler, and D. S. Sivia. Deconvolution of lifetime broadening at rare-earth l_{III} edges compared to resonant inelastic x-ray scattering measurements. *Phys. Rev. B*, 54(21):14877–14880, Dec 1996.

[68] Adriano Filipponi. Deconvolution of the lifetime broadening from x-ray absorption spectra of atomic and molecular species. *J Phys B: At Mol Opt Phys*, 33(15):2835–2846, 2000.

[69] K V Klementev. Deconvolution problems in x-ray absorption fine structure spectroscopy. *J. Phys. D*, 34(15):2241–2247, 2001.

[70] T. T. Fister, G. T. Seidler, J. J. Rehr, J. J. Kas, W. T. Elam, J. O. Cross, and K. P. Nagle. Deconvolving instrumental and intrinsic broadening in core-shell x-ray spectroscopies. *Phys. Rev. B*, 75(17):174106, 2007.

[71] Il-Kyoung Jeong, M. J. Graf, and R. H. Heffner. Effects of bragg peak profiles and nanoparticle sizes on the real-space pair distribution function. *J. Appl. Crystallogr.*, 38(1):55–61, 2005.

[72] E. S. Božin, V. Petkov, P. W. Barnes, P. M. Woodward, T. Vogt, S. D. Mahanti, and S. J. L. Billinge. Temperature dependent total scattering structural study of $CaCu_3Ti_4O_{12}$. *J. Phys.: Condens. Matter*, 16(44):S5091–S5102, November 2004.

[73] C. Malliakas, S. J. L. Billinge, H. J. Kim, and M. G. Kanatzidis. Square nets of tellurium: Rare-earth dependent variation in the charge-density wave of $RETe_3$ (RE = rare-earth element). *J. AM. Chem. Soc.*, 127(18):6510–6511, 2005.

[74] H. Kim, C. Malliakas, A. Tomic, S. Tessmer, M. Kanatzidis, and S. Billinge. Local atomic structure and discommensurations in the charge density wave of $CeTe_3$. *Phys. Rev. Lett.*, 96(22):226401, June 2006.

[75] E. S. Božin, X. Qiu, M. Schmidt, G. Paglia, J. F. Mitchell, P. G. Radaelli, Th. Proffen, and S. J. L. Billinge. Local structural aspects of the orthorhombic to pseudo-cubic phase transformation in $La_{1-x}Ca_xMnO_3$. *Physica B*, 385-386:110–112, 2006.

[76] S. J. Gurman and R. L. McGreevy. Reverse monte-carlo simulation for the analysis of EXAFS data. *J. Phys.: Condens. Matter*, 2(48):9463–9473, December 1990.

[77] P. Eisenberger and George S. Brown. The study of disordered systems by EXAFS: Limitations. *Sol. Stat. Comm.*, 29(6):481–484, February 1979.

[78] A. I. Frenkel and J. J. Rehr. Thermal expansion and x-ray-absorption fine-structure cumulants. *Phys. Rev. B*, 48(1):585–588, Jul 1993.

[79] Nguyen Van Hung and J. J. Rehr. Anharmonic correlated Einstein-model Debye-Waller factors. *Phys. Rev. B*, 56(1):43–46, Jul 1997.

[80] G. Dalba, P. Fornasini, R. Grisenti, and J. Purans. Sensitivity of extended x-ray-absorption fine structure to thermal expansion. *Phys. Rev. Lett.*, 82(21):4240–4243, May 1999.

[81] P. Norby, I. G. Grogh Andersen, E. Krogh Andersen, and N. H. Andersen. The crystal structure of lanthanum manganate(III), $LaMnO_3$, at room temperature and at 1273K under N_2. *J. Solid State Chem.*, 119:191–196, 1995.

[82] H. Röder, Jun Zang, and A. R. Bishop. Lattice effects in the colossal-magnetorestance manganites. *Phys. Rev. Lett.*, 76(8):1356–1359, February 1996.

[83] F. Bridges. Unpublished data. 2005.

[84] P. G. Radaelli, M. Marezio, H. Y. Hwang, S-W. Cheong, and B. Batlogg. Charge localization by static and dynamic distortions of the MnO_6 octahedra in perovskite manganites. *Phys. Rev. B*, 54(13):8992–8995, Oct 1996.

[85] Simon J. L. Billinge and Igor Levin. The problem with determining atomic structure at the nanoscale. *Science*, 316(5824):561–565, 2007.

[86] E. Daryl Crozier. A review of the current status of XAFS spectroscopy. *Journal of Synchrotron Radiation*, 133(1-4):134–144, December 1997.

[87] R. L. McGreevy and L. Pusztai. Reverse monte carlo simulation: A new technique for the determination of disordered structures. *Mol. Simul.*, 1:359, 1988.

[88] Matthew G. Tucker, Martin T. Dove, and David A. Keen. Application of the reverse monte carlo method to crystalline materials. *J. Appl. Crystallogr.*, 34(5):630–638, Oct 2001.

[89] Andrew L. Goodwin, Matthew G. Tucker, Martin T. Dove, and David A. Keen. Phonons from powder diffraction: A quantitative model-independent evaluation. *Phys. Rev. Lett.*, 93(7):075502, 2004.

[90] N. Binsted, M. J. Pack, M. T. Weller, and J. Evans. Combined exafs and powder diffraction analysis. *Journal Of The American Chemical Society*, 118(42):10200–10210, October 1996.

[91] IXS Standards and Criteria Committee. Error reporting recommendations: A report of the standards and criteria committee. Technical report, International XAFS Society, July 26 2000.

[92] J. C. Mikkelsen and J. B. Boyce. Extended x-ray-absorption fine-structure study of $Ga_{1-x}In_xAs$ random solid solutions. *Phys. Rev. B*, 28(12):7130–7140, Dec 1983.

[93] V. Petkov, I.-K. Jeong, F. Mohiuddin-Jacobs, T. Proffen, S. J. L. Billinge, and W. Dmowski. Local structure of $In_{0.5}Ga_{0.5}As$ from joint high-resolution and differential pair distribution function analysis. *J. Appl. Phys.*, 88(2):665–672, July 2000.

[94] I.-K. Jeong, F. Mohiuddin-Jacobs, V. Petkov, S. J. L. Billinge, and S. Kycia. Local structure of $In_xGa_{1-x}As$ semiconductor alloys by high-energy synchrotron x-ray diffraction. *Phys. Rev. B*, 63(20):205202, Apr 2001.

[95] José Luís Martins and Alex Zunger. Bond lengths around isovalent impurities and in semiconductor solid solutions. *Phys. Rev. B*, 30(10):6217–6220, Nov 1984.

[96] A. Balzarotti, N. Motta, A. Kisiel, M. Zimnal-Starnawska, M. T. Czyyk, and M. Podgórny. Model of the local structure of random ternary alloys: Experiment versus theory. *Phys. Rev. B*, 31(12):7526–7539, Jun 1985.

[97] Matthias C. Schabel and José Luriaas Martins. Structural model for pseudobinary semiconductor alloys. *Phys. Rev. B*, 43(14):11873–11883, May 1991.

[98] J. S. Chung and M. F. Thorpe. Local atomic structure of semiconductor alloys using pair distribution functions. *Phys. Rev. B*, 55(3):1545–1553, 1997.

[99] Brian H. Toby and Simon J. L. Billinge. Determination of standard uncertainties in fits to pair distribution functions. *Acta Cryst. A*, 60(4):315–317, 2004.

[100] M. Podgorny, M. T. Czyzyk, A. Balzarotti, P. Letardi, N. Motta, A. Kisiel, and M. Zimnal-Starnawska. Crystallographic structure of ternary semiconducting alloys. *Solid State Communications*, 55(5):413–417, August 1985.

[101] J. R. Fienup. Phase retrieval algorithms: a comparison. *Applied Optics*, 21(15):2758–2769, 1982.

[102] J. M. Zuo, I. Vartanyants, M. Gao, R. Zhang, and L. A. Nagahara. Atomic resolution imaging of a carbon nanotube from diffraction intensities. *Science*, 300:1419–1421, 2003.

[103] Jinsong Wu, Kurt Leinenweber, John C. H. Spence, and Michael O'Keeffe. Ab initio phasing of x-ray powder diffraction patterns by charge flipping. *Nature*, 5:647–652, August 2006.

[104] Katsuaki Kodama, Satoshi Iikubo, Tomitsugu Taguchi, and Shin-ichi Shamoto. Finite size effects of nanoparticles on the atomic pair distribution functions. *Acta Crystallographica Section A*, 62(6):444–453, Nov 2006.

[105] DANSE. `http://wiki.cacr.caltech.edu/danse/index.php/Main_Page`.

[106] DIFFPY. `http://www.diffpy.org`.

[107] Python Software Foundation. `http://www.python.org`.

[108] wxPython. `http://www.wxpython.org`.

[109] J. Painter and E. A. Merritt. mmLib Python toolkit for manipulating annotated structural models of biological macromolecules. *Journal of Applied Crystallography*, 37:174–178, 2004.

[110] Matplotlib. `http://matplotlib.sourceforge.net`.

[111] J. Li. AtomEye: an efficient atomistic configuration viewer. *Model. Simul. Mater. Sc.*, 11(2):173–177, 2003.

[112] Emil Božin. Private Communication.

[113] Jacqueline B. A. A. Elemans, B. Van Laar, K. R. Van der Veen, and B. O. Loopstra. The crystallographic and magnetic structures of $La_{1-x}Ba_xMn_{1-x}Me_xO_3$ (Me = Mn or Ti). *J. Solid State Chem.*, 3(2):238–42, 1971.

[114] J. Rodriguez-Carvajal, M. Hennion, F. Moussa, A. H. Moudden, L. Pinsard, and A. Revcolevschi. Neutron-diffraction study of the Jahn-Teller transition in stoichiometric $LaMnO_3$. *Phys. Rev. B*, 57(6):R3189–R3192, 1998.

[115] Tapan Chatterji, Francois Fauth, Bachir Ouladdiaf, P. Mandal, and B. Ghosh. Volume collapse in $LaMnO_3$ caused by an orbital order-disorder transition. *Phys. Rev. B*, 68:052406, 2003.

[116] Th. Proffen and S. J. L. Billinge. Probing the local structure of doped manganites using the atomic pair distribution function. *Appl. Phys. A*, 74:1770, 2002.