

SOLVING ATOMIC STRUCTURE USING STATISTICAL MECHANICAL SEARCHES ON  
X-RAY SCATTERING DERIVED POTENTIAL ENERGY SURFACES

by

Christopher James Wright

Bachelor of Science  
Brown University 2014

---

Submitted in Partial Fulfillment of the Requirements

for the Degree of Masters of Science in

Chemical Engineering

College of Engineering and Computing

University of South Carolina

2016

Accepted by:

Xiao-Dong Zhou, Major Professor

Thomas Vogt, Committee Member

Mark Uline, Committee Member

Jochen Lauterbach, Committee Member

Lacy Ford, Vice Provost and Dean of Graduate Studies

© Copyright by Christopher James Wright, 2016  
All Rights Reserved.

## DEDICATION

## ACKNOWLEDGMENTS

## ABSTRACT

## TABLE OF CONTENTS

DEDICATION . . . . .	iii
ACKNOWLEDGMENTS . . . . .	iv
ABSTRACT . . . . .	v
LIST OF TABLES . . . . .	ix
LIST OF FIGURES . . . . .	x
TODO LIST . . . . .	1
CHAPTER 1 STATISTICAL MECHANICAL ENSEMBLES AND POTENTIAL ENERGY SURFACES . . . . .	5
1.1 Introduction . . . . .	5
1.2 Potential Energy Surfaces . . . . .	5
1.3 Ensembles . . . . .	7
CHAPTER 2 ATOMIC PAIR DISTRIBUTION FUNCTION: THEORY AND COMPUTATION . . . . .	12
2.1 Theory . . . . .	12
2.2 Computation . . . . .	15
CHAPTER 3 BENCHMARKING . . . . .	20

3.1 PDF . . . . .	20
3.2 PDF with ADPs . . . . .	29
CHAPTER 4 X-RAY TOTAL SCATTERING DATA ACQUISITION AND PROCESSING . . . . .	35
4.1 Introduction . . . . .	35
4.2 Data Storage and Management . . . . .	35
CHAPTER 5 ANNEALING AND AGGREGATION OF 2NM AU NANOPARTICLES . . . . .	52
5.1 Experiments . . . . .	52
5.2 Data Processing . . . . .	52
5.3 Data Analysis . . . . .	52
5.4 Simulation . . . . .	52
5.5 Structural Analysis . . . . .	52
5.6 Conclusions . . . . .	52
CHAPTER 6 PHASE CHANGES AND ANNEALING DYNAMICS OF $\text{Pr}_2\text{NiO}_4$ AND ITS DERIVATIVES . . . . .	53
6.1 Experiments . . . . .	53
6.2 Data Processing . . . . .	53
6.3 Data Analysis . . . . .	53
6.4 Simulation . . . . .	72
6.5 Conclusions . . . . .	72
CHAPTER 7 CONCLUSION . . . . .	74

BIBLIOGRAPHY . . . . .	75
------------------------	----

## LIST OF TABLES

## LIST OF FIGURES

Figure 2.1 Comparison of the CPU and GPU chip architectures . . . . .	16
Figure 3.1 Au <sub>55</sub> PDF fitting of DFT-optimized <i>c</i> -Au <sub>55</sub> . (a) the final structural solution ( $Rw=0.3\%$ ) with bond lengths color-coded by step of 0.05Å, (b) the target PDF(blue dots) overlaid with the PDF of the final structure (solid red lines) with the difference in green lines offset below, ?? the radial bond distribution, and (d) bond angle distribution. . . . .	24
Figure 3.2 Au <sub>55</sub> PDF fitting of surface-disordered Au <sub>55</sub> . a) the target structure, b) the final structural solution ( $Rw=0.6\%$ ), c) the comparison of PDFs, d) the CN distribution with the number of atoms in either the core or the surface, e) the radial bond distribution, and f) the bond angle distribution. . . . .	30
Figure 3.3 Similar to figure 3.2 for DFT-optimized amorphous Au <sub>55</sub> . . . . .	31
Figure 3.4 Similar to figure 3.2 for Au <sub>102</sub> as in DFT-optimized Au <sub>102</sub> MBA <sub>44</sub> cluster. . . . .	32
Figure 3.5 Similar to Fig. 3.6 with Marks decahedron as the starting structure.	33
Figure 3.6 . . . . .	34
Figure 4.1 Database Loading Workflow. Data is loaded from various sources, including images and text files, into the FileStore and Meta-dataStore databases. Data is then retrieved from the databases using the databroker. . . . .	36
Figure 4.2 Scattering onto a flat detector . . . . .	37
Figure 4.3 $Q$ resolution as a function of $Q$ . . . . .	38
Figure 4.4 Number of pixels as a function of $Q$ , binned at the $Q$ resolution of the detector. . . . .	39

Figure 4.5 Generated dead/hot pixel masks for a detector with 100 bad pixels. a) the standard even bin mask and b) the $Q$ resolution binned mask. The bad pixels are noted with open circles, masked pixels are noted with closed circles. . . . .	41
Figure 4.6 Generated dead/hot pixel masks for a detector with 300 bad pixels. a) the standard even bin mask and b) the $Q$ resolution binned mask. The bad pixels are noted with open circles, masked pixels are noted with closed circles. . . . .	41
Figure 4.7 Generated dead/hot pixel masks for a detector with 500 bad pixels. a) the standard even bin mask and b) the $Q$ resolution binned mask. The bad pixels are noted with open circles, masked pixels are noted with closed circles. . . . .	42
Figure 4.8 Generated dead/hot pixel masks for a detector with 1000 bad pixels. a) the standard even bin mask and b) the $Q$ resolution binned mask. The bad pixels are noted with open circles, masked pixels are noted with closed circles. . . . .	42
Figure 4.9 Generated beamstop holder masks for a beamstop holder with 10% transmittance. a) the raw image, b) the masked image, c) and the missed pixels . . . . .	43
Figure 4.10 Generated beamstop holder masks for a beamstop holder with 30% transmittance. a) the raw image, b) the masked image, c) and the missed pixels . . . . .	43
Figure 4.11 Generated beamstop holder masks for a beamstop holder with 50% transmittance. a) the raw image, b) the masked image, c) and the missed pixels . . . . .	43
Figure 4.12 Generated beamstop holder masks for a beamstop holder with 90% transmittance. a) the raw image, b) the masked image, c) and the missed pixels . . . . .	44
Figure 4.13 Generated beamstop holder masks which is rotated away from verticle . . . . .	44
Figure 4.14 Masked experimental data. a) the raw image, b) the mask . . . . .	45
Figure 4.15 Masked experimental data with Pt single crystal signal. a) the raw image, b) the mask . . . . .	45

Figure 4.16 Masked experimental data with Pt single crystal signal using figure's 4.14 as a starting mask. a) the raw image, b) the mask . . . . . 46

Figure 4.17 Masking, average, and standard deviation of an example x-ray total scattering measurement. This image was produced with no mask. a) the image, b) the mask, c) the mean and median values, d) the standard deviation (normalized to the median), e) a closeup of the  $28 \text{ \AA}^{-1}$ to  $31 \text{ \AA}^{-1}Q$  range for the mean and median, f)  $28 \text{ \AA}^{-1}$ to  $31 \text{ \AA}^{-1}Q$  range for the standard deviation . . . . . 49

Figure 4.18 Masking, average, and standard deviation of an example x-ray total scattering measurement. This image was produced with only an edge mask. a) the image, b) the mask, c) the mean and median values, d) the standard deviation (normalized to the median), e) a closeup of the  $28 \text{ \AA}^{-1}$ to  $31 \text{ \AA}^{-1}Q$  range for the mean and median, f)  $28 \text{ \AA}^{-1}$ to  $31 \text{ \AA}^{-1}Q$  range for the standard deviation . . . . . 50

Figure 4.19 Masking, average, and standard deviation of an example x-ray total scattering measurement. This image was produced combining an edge mask and the automatically generated mask. a) the image, b) the mask, c) the mean and median values, d) the standard deviation (normalized to the median), e) a closeup of the  $28 \text{ \AA}^{-1}$ to  $31 \text{ \AA}^{-1}Q$  range for the mean and median, f)  $28 \text{ \AA}^{-1}$ to  $31 \text{ \AA}^{-1}Q$  range for the standard deviation . . . . . 51

## TODO LIST

2	Why is atomistic engineering important . . . . .	2
3	Barriers to atomistic engineering . . . . .	2
4	How are we going to attack this problem . . . . .	2
5	Fix the figures so that they don't give the full caption . . . . .	4
6	This entire section needs some rewriting to distinguish this from the paper .	20
7	Put this somewhere . . . . .	24
8	masking parameters . . . . .	53
9	integration parameters . . . . .	53
10	PDF parameters . . . . .	53

11

## INTRODUCTION

12

### Why is atomistic engineering important

13     Engineering materials and chemicals on the atomic scale has been a goal for the  
14    chemistry, physics, materials science, and chemical engineering fields long before the  
15    advent of nanomaterials. Realizing this goal could lead to durable fuel cell catalysts,  
16    more bioavailable pharmaceuticals, and radiation damage resistant spacecraft shielding.

17

### Barriers to atomistic engineering

18     Before we can even think of making atomistically exact structures, durable struc-  
19    tures, or structures which change in reproducible ways, we need to know the atomic  
20    structure exactly.

21

### How are we going to attack this problem

22     This work addresses these issues by developing a methodology for solving the  
23    structure of nanomaterials by matching experimental x-ray scattering data with sim-  
24    ulated atomic structures.

25     Chapter 1 develops the statistical mechanical system used to match the theoretical  
26    structure. §1.2 focuses on the development of potential energy surfaces, including  
27    potential energy and force equations, which have minima where experimental results  
28    and simulated structures agree the most. §1.3 will discuss statistical mechanical  
29    ensembles which are used to search for minima on the potential energy surface.

30     Chapter 2 will discuss the mathematical and computational development of the  
31    atomic pair distribution function (PDF). §2.2 will focus on the rapid graphical pro-  
32    cessing unit based calculation of the PDF and its gradients.

33     Chapter 3 will discuss the benchmarking of the the combined statistical mechan-

34 ical optimizer and PDF calculation systems against a series of theoretical nanoparti-  
35 cles, focusing on understanding limitations of the method and structure reproduction.

36 Chapter 4 will focus on the aquesition of experimental data, its management, and  
37 processing. §4.2, 4.2, and 4.2 will discuss the derivation of the  $Q$  resolution function,  
38 the automated masking of 2D area detectors for x-ray total scattering measurements  
39 using the previously derrived  $Q$  resolution, and the impact of different averaging  
40 methods and masks on azimuthal integration, respecitvly.

41

Fix the figures so that they don't give the full caption

42

# CHAPTER 1

43

## STATISTICAL MECHANICAL ENSEMBLES AND POTENTIAL ENERGY SURFACES

45 1.1 INTRODUCTION

46 The approach taken in this work for solving the atomic structures of materials is  
47 one of optimization. The positional variables of the system are optimized so as to  
48 minimize the value of a potential energy surface (PES). The

49 1.2 POTENTIAL ENERGY SURFACES

50 A PES simply describes the potential energy of the system as a function of all its  
51 relevant coordinates in phase space, essentially providing a mapping  $\mathbb{R}^n \rightarrow \mathbb{R}$ . Usually  
52 these coordinates are the positions of the atoms  $q$  and their conjugate momenta  $p$ .  
53 Note that there could be more variables associated with the system, for instance the  
54 magnetic moments of the atoms could play a role in describing the system. In this  
55 magnetic system there would be positional variables for the atomwise spin vectors  
56 and their "momenta". Application of the term "momenta" might seem odd here, as  
57 the magnetic spin does not have a mass or a velocity. However, since the magnetic  
58 "position" is defined on the PES we need to describe its conjugate variable to properly  
59 formulate Hamiltonian dynamics and the kinetic portion of the PES.

60 **Experimentally Derived Potential Energy Surfaces**

61 Generally PESs are obtained from purely computational experiments including: ab-  
62 initio DFT, classical approximations via the embedded atom method, or even param-  
63 eter driven models with experimentally fitted parameters. However, one can derive  
64 a PES from an experiment which describes how well the model reproduces the ex-  
65 perimental data. In this case one needs a theoretical and computational framework  
66 mapping the atomistic variables of the simulation to the same space of the data ob-  
67 tained from the experiment. This allows the experiment to be compared directly  
68 against the predicted data via an experimentally derived PES.

69 **Potentials**

70 For an experiment which produces 1D data, like powder diffraction, EXAFS or XPS,  
71 the implemented potentials are:

$$\chi^2 = \sum_{a=a_{\min}}^{a_{\max}} (A_{\text{obs}} - \alpha A_{\text{calc}})^2 \quad (1.1)$$

$$Rw = \sqrt{\frac{\sum_{a=a_{\min}}^{a_{\max}} (A_{\text{obs}} - \alpha A_{\text{calc}})^2}{\sum_{a=a_{\min}}^{a_{\max}} A_{\text{obs}}^2}} \quad (1.2)$$

$$\chi_{\text{INVERT}}^2 = \frac{1}{N} \sum_j \sum_r [A_{\text{obs}}(r) - \alpha A_{j,\text{calc}}(r)]^2 \quad (1.3)$$

$$\alpha = \frac{\sum_{a=a_{\min}}^{a_{\max}} A_{\text{obs}} A_{\text{calc}}}{\sum_{a=a_{\min}}^{a_{\max}} A_{\text{calc}}^2} = \frac{\vec{A}_{\text{obs}} \cdot \vec{A}_{\text{calc}}}{|\vec{A}_{\text{calc}}|^2} \quad (1.4)$$

75 where  $A_{\text{calc}}$  and  $A_{\text{obs}}$  are the calculated and observed 1D experimental data and  $A_{j,\text{calc}}$   
76 is the calculated data for a single atom interacting with the other atoms of the system.  
77 Note that  $A_{\text{calc}}$  has a dependence on  $q$ , the positions of the system.

78 **Forces**

$$\vec{\nabla} \chi^2 = -2 \sum_{a=a_{\min}}^{a_{\max}} (\alpha \frac{\partial A_{\text{calc}}}{\partial \gamma_{i,w}} + A_{\text{calc}} \frac{\partial \alpha}{\partial \gamma_{i,w}})(A_{\text{obs}} - \alpha A_{\text{calc}}) \quad (1.5)$$

79

$$\vec{\nabla}Rw = \frac{Rw}{\chi^2} \sum_{a=a_{\min}}^{a_{\max}} (\alpha \frac{\partial A_{\text{calc}}}{\partial \gamma_{i,w}} + A_{\text{calc}} \frac{\partial \alpha}{\partial \gamma_{i,w}})(\alpha A_{\text{calc}} - (A_{\text{obs}})) \quad (1.6)$$

80

$$\frac{\partial \alpha}{\partial \gamma_{i,w}} = \frac{(\sum_{a=a_{\min}}^{a_{\max}} A_{\text{obs}} \frac{\partial A_{\text{calc}}}{\partial \gamma_{i,w}} - 2\alpha \sum_{a=a_{\min}}^{a_{\max}} A_{\text{calc}} \frac{\partial A_{\text{calc}}}{\partial \gamma_{i,w}})}{\sum_{a=a_{\min}}^{a_{\max}} A_{\text{calc}}^2} \quad (1.7)$$

81

$$\vec{\nabla}\chi^2_{\text{INVERT}} = \frac{-2}{N} \sum_{a=a_{\min}}^{a_{\max}} \sum_j (\alpha \frac{\partial A_{j,\text{calc}}}{\partial \gamma_{i,w}} + A_{j,\text{calc}} \frac{\partial \alpha}{\partial \gamma_{i,w}})(A_{\text{obs}} - \alpha A_{j,\text{calc}}) \quad (1.8)$$

82 where  $\gamma_{i,w}$  is the  $i$ th arbitrary positional variable in the  $w$ th direction. The concept  
 83 of an "arbitrary positional variable" might seem a bit cumbersome but it allows us  
 84 to define the forces for any atomic parameter which can be represented as a vector  
 85 in 3-space. This comes in handy when trying to define the forces acting on variables  
 86 like anisotropic displacement parameters or atomic magnetic spins.

87 DISCUSS INVERT A BUNCH. ALSO COMPARE RW AND CHI\*\*2, POTEN-  
 88 TIALY WITH A FIGURE.

89 1.3 ENSEMBLES

90 While PESs describe which atomic configurations are the most desirable and how  
 91 the atoms would like to get there, the ensemble describes how the atoms move on  
 92 the PES. The abstraction of the PES from the ensemble is an important one, as it  
 93 allows for the reuse and exchange of both PESs and ensembles for a wide array of  
 94 problems. Statistical mechanical ensembles can be described in two ways, analytically  
 95 and stochastically. For long simulation times and fine enough numerical or analytical  
 96 integration these two descriptions should be identical. In either case one starts by  
 97 defining the Hamiltonian  $\mathcal{H}$  as the total energy of the system. Thus, the Hamiltonian  
 98 is described as the sum of the potential  $U(q)$  and kinetic  $K(p)$  energies, where  $q$  is  
 99 the positions of the atoms and  $p$  is their momenta

$$\mathcal{H}(q, p) = U(q) + K(p) \quad (1.9)$$

100 where  $K(p) = \frac{1}{2} \sum_i \frac{p_i^2}{m_i}$  and  $i$  denotes the  $i$ th particle. Analytically one generally defines  
 101 a partition function, which describes the sum of probabilities over all potential atomic  
 102 states.

$$\Xi = \sum_i P_i(q, p)$$

103 where  $P_i$  is the probability of the  $i$ th state and is a function of the total energy of  
 104 that state. This partition function can then be used to obtain the probability of any  
 105 specific state.

## 106 Hamiltonian Monte Carlo

In order to model dynamics we need to describe the motion of the particles in our system, thus:

$$\frac{dq_i}{dt} = \frac{\partial \mathcal{H}}{\partial p_i} = p_i \quad (1.10)$$

$$\frac{dp_i}{dt} = -\frac{\partial \mathcal{H}}{\partial q_i} = -\vec{\nabla}U \quad (1.11)$$

Using these equations we can derive the position and momentum vectors at any point in time using the leap-frog algorithm:

$$p_i(t + \delta t/2) = p_i(t) - \frac{\delta t}{2} \frac{\partial}{\partial q_i} U(q(t)) \quad (1.12)$$

$$q_i(t + \delta t) = q_i(t) + \delta t * p_i(t + \delta t/2) \quad (1.13)$$

$$p_i(t + \delta t) = p_i(t + \delta t/2) - \frac{\delta t}{2} \frac{\partial}{\partial q_i} U(q(t + \delta t)) \quad (1.14)$$

107 Note that  $\frac{\partial}{\partial q_i}$  is the gradient with respect to  $q$  where  $i$  denotes the  $i$ th atom being  
 108 moved. Using this notation the gradient is

$$\vec{\nabla}U = \begin{bmatrix} \frac{\partial U}{\partial q_{0,x}} & \frac{\partial U}{\partial q_{0,y}} & \frac{\partial U}{\partial q_{0,z}} \\ \vdots & \frac{\partial U}{\partial q_{i,w}} & \vdots \\ \frac{\partial U}{\partial q_{n,x}} & \frac{\partial U}{\partial q_{n,y}} & \frac{\partial U}{\partial q_{n,z}} \end{bmatrix} = \begin{bmatrix} \vec{\mathcal{F}}_0 \\ \vdots \\ \vec{\mathcal{F}}_i \\ \vdots \\ \vec{\mathcal{F}}_n \end{bmatrix} \quad (1.15)$$

109 where  $\frac{\partial}{\partial q_{i,w}}$  is the derivative with respect to  $q$  where  $w$  denotes direction of the deriva-  
 110 tive ( $x$ ,  $y$ , or  $z$ ),  $n$  is the number of atoms and  $U$  is the potential which depends on  
 111  $q$ , and  $\vec{F}_i$  is the "force" on the  $i$ th atom.

## 112 No-U-Turn-Sampling

## 113 Grand Canonical Ensemble

### 114 Ensemble description

115 In the Grand Canonical Ensemble (GCE) two sets of variables are allowed to change,  
 116 the atomic positions and the total number of atoms and their associated identities.  
 117 These two variables are controlled by temperature and chemical potential. The par-  
 118 tition function is

$$\Xi = e^{-\beta(\mathcal{H}+\mu)} \quad (1.16)$$

119 This is translated into a Monte Carlo system, producing Grand Canonical Monte  
 120 Carlo (GCMC).

### 121 Grand Canonical Monte Carlo

122 While the probabilities for atomic motion are the same as in the Canonical Ensemble,  
 123 the addition or removal of an atom have their own probabilities. For the addition of  
 124 an atom the probability is formally:

$$\min[1, \frac{V}{(N+1)\Lambda(T)^3} e^{-\beta\Delta U + \beta\mu}] \quad (1.17)$$

125 Similarly the removal of an atom has the probability:

$$\min[1, \frac{(N)\Lambda(T)^3}{V} e^{-\beta\Delta U - \beta\mu}] \quad (1.18)$$

126 However, both of these equations depend of the overall simulation volume and the  
 127 thermal wavelength, which is undesirable as these are not really properties that we

128 are of interest to these simulations. Thus, we roll them into the definition of the  
129 chemical potential, essentially setting the base chemical potential to counteract these  
130 effects. This makes certain that our simulation does not change if we change the  
131 overall cell volume. A GCMC move consists of creating a new atomic configuration,  
132 where an atom has been added or removed, and checking the above criteria. However,  
133 previous results have shown that this method is computationally expensive in dense  
134 liquids, and exceedingly expensive in solid materials. The long simulation times  
135 are due to the random nature of the atomic additions or removals which produce:  
136 over-tightly packed atoms, atoms in the middle of nowhere, or unphysical vacancies.  
137 These configurations are rejected by the GCMC criteria but their probability of being  
138 sampled is much higher than configurations which are lower in energy, since the  
139 number of incorrect ways to add/remove atoms is much larger than the correct ways.  
140 Thus we have implemented methods for biasing the atomic addition positions and  
141 the atomic removals toward configurations which are more likely to be accepted.

## 142 GCMC biasing

143 The first method is to remove some of the excess options from the probability pool.  
144 Initially the insertion positions are calculated at random using a random number gen-  
145 erator and scaled to the size of the simulation cell. This produces probabilities which  
146 have floating point level precision, which is effectively infinite. While this produces  
147 a potentially infinite number of ways to create energetically favorable configurations,  
148 the infinite ways to produce bad configurations is much larger. Thus we can limit this  
149 by moving to voxels. In this case atoms are added to the center of voxels which have  
150 a pre-set resolution, limiting our total number of valid addition points. While this  
151 could produce some problems with ergodicity, we avoid this by allowing the atoms to  
152 translate throughout the system. Each voxel has a probability of being tried:

$$P_{i,j,k} = \frac{xyz}{abc} \quad (1.19)$$

153 where  $x, y, z$  and  $a, b, c$  are the resolutions and cell side lengths in the cardinal di-  
154 rections, respectively. While this does help to limit the total probability space it  
155 does not tell us which voxels are likely to lead to better configurations, leading to  
156 many rejected atomic additions. To combat this issue we can weigh the individual  
157 voxels, giving more probability to voxels which show promise and less to those with  
158 less likelihood to be accepted.

159 The approach most likely to yield success would be to measure the change in  
160 potential energy associated with the addition of an atom at the center of the voxel  
161 where the probability of a voxel to be tried is:

$$P_{i,j,k} = \frac{e^{\beta\Delta U_{i,j,k}}}{\sum_{i,j,k} e^{\beta\Delta U_{i,j,k}}} \quad (1.20)$$

162 where  $\Delta U_{i,j,k}$  is the change in energy. However, calculating  $\Delta U_{i,j,k}$  can be particu-  
163 larly expensive, especially when calculating scattering from atomic positions. The  
164 computational expense can be mitigated by using a cheaper potential, if only for the  
165 evaluation of the voxel energy, as previously shown. Similar to previous work we can  
166 use the Lennard Jones potential to approximate the addition potential.

167

## CHAPTER 2

168

# ATOMIC PAIR DISTRIBUTION FUNCTION: THEORY AND COMPUTATION

170 2.1 THEORY

171 To properly understand the PDF and its limitations we need to derive its mathemat-  
172 ics. The PDF has been previously derived many times so it is not rederived here.  
173 This discussion of the PDF and its gradients use the notation of Farrow and Billinge.  
174 [?]

175 **Derivation**

176 **Analytical Gradients**

177 Many optimization algorithms and simulations methodologies, including HMC, re-  
178 quire not only the potential energy of a given configuration but also the forces acting  
179 on that configuration. These forces are described by the gradient of potential energy  
180 of the system which in turn requires the gradient of the PDF. As previously shown the  
181 PDF is the Fourier Transform of the Debye equation. Since the Fourier Transform is  
182 expressed as an integral we can exchange the order of the gradient and the integral,  
183 allowing us to calculate the analytical gradient of the Debye equation and FFT the  
184 resulting function. The Debye equation, with a Debye-Waller vibrational correction  
185 is

$$F(Q) = \frac{1}{N\langle f \rangle^2} \sum_{j \neq i} f_i^*(Q) f_j(Q) \exp(-\frac{1}{2}\sigma_{ij}^2 Q^2) \frac{\sin(Qr_{ij})}{r_{ij}} \quad (2.1)$$

186 where

$$\sigma_{ij}^2 = (\vec{u}_{ij} * \hat{d}_{ij})^2 \quad (2.2)$$

$$\vec{u}_{ij} = \vec{u}_i - \vec{u}_j \quad (2.3)$$

$$\hat{d}_{ij} = \frac{\vec{d}_{ij}}{r_{ij}} \quad (2.4)$$

$$r_{ij} = \|\vec{d}_{ij}\| \quad (2.5)$$

$$\vec{d}_{ij} = \begin{bmatrix} q_{ix} - q_{jx} \\ q_{iy} - q_{jy} \\ q_{iz} - q_{jz} \end{bmatrix} \quad (2.6)$$

187 where  $Q$  is the scatter vector,  $f_i$  is atomic scattering factor of the  $i$ th atom, and  $r_{ij}$   
 188 is the distance between atoms  $i$  and  $j$  and has  $q$  dependence. For simplicities sake  
 189 we will break up  $F(Q)$  so that

$$F(Q) = \alpha \sum_{j \neq i} \beta_{ij} \tau_{ij} \Omega_{ij} \quad (2.7)$$

190 where

$$\alpha = \frac{1}{N\langle f \rangle^2} \quad (2.8)$$

$$\beta_{ij} = f_i^*(Q) f_j(Q) \quad (2.9)$$

$$\tau_{ij} = \exp(-\frac{1}{2} \sigma_{ij}^2 Q^2) \quad (2.10)$$

$$\Omega_{ij} = \frac{\sin(Qr_{ij})}{r_{ij}} \quad (2.11)$$

191 The derivatives are as follows:

$$\frac{\partial}{\partial q_{i,w}} F(Q) = \alpha \sum_j \beta_{ij} \left( \frac{\partial \tau_{ij}}{\partial q_{i,w}} \Omega_{ij} + \tau_{ij} \frac{\partial \Omega_{ij}}{\partial q_{i,w}} \right) \quad (2.12)$$

192 where

$$\frac{\partial \Omega_{ij}}{\partial q_{i,w}} = \frac{Q \cos(Qr_{ij}) - \Omega_{ij}}{r_{ij}^2} (q_{i,w} - q_{j,w}) \quad (2.13)$$

$$\frac{\partial \tau_{ij}}{\partial q_{i,w}} = \frac{\sigma_{ij} Q^2 \tau_{ij}}{r_{ij}^3} ((q_{i,w} - q_{j,w}) \sigma_{ij} - (u_{i,w} - u_{j,w}) r_{ij}^2) \quad (2.14)$$

193 Since  $\vec{u}_{ij}$  is a variable as well, we need the derivative with respect to it as well.

194 Thus

$$\frac{\partial}{\partial u_{i,w}} F(Q) = \alpha \sum_j \beta_{ij} \frac{\partial \tau_{ij}}{\partial u_{i,w}} \Omega_{ij} \frac{\partial \tau_{ij}}{\partial u_{i,w}} = -\frac{\sigma_{ij} Q^2 \tau_{ij}}{r_{ij}} (q_{i,w} - q_{j,w}) \quad (2.15)$$

## 195 Without ADPs

196 Without ADPs the equations simplify down to

$$F(Q) = \frac{1}{N\langle f \rangle^2} \sum_{j \neq i} f_i^*(Q) f_j(Q) \frac{\sin(Qr_{ij})}{r_{ij}} \quad (2.16)$$

197 and

$$\frac{\partial}{\partial q_{i,w}} F(Q) = \alpha \sum_j \beta_{ij} \frac{\partial \Omega_{ij}}{\partial q_{i,w}} \quad (2.17)$$

198 use of these equations, when ADPs are not appropriate (like at cyrogenic tempera-  
199 tures), greatly speeds up the computaiton.

## 200 Periodic Boundary Conditions

201 Periodic boundary conditions can be helpful when simulating extended solids or large  
202 nanoparticles. In this case all the non-crystallinity is contained within the simulation  
203 box and the box is repeated to create the longer distance peaks observed in the PDF.

204 To perform this we can break up the Debye equation into two main parts, the part  
205 that describes the interatomic distances within the simulation box and those between  
206 boxes. Neglecting the thermal motion portion:

$$F(Q) = \frac{1}{N\langle f \rangle^2} \left( \sum_{j \neq i} f_i^*(Q) f_j(Q) \frac{\sin(Qr_{ij})}{r_{ij}} + \sum_{i,j} f_i^*(Q) f_j(Q) \frac{\sin(QR_{ij})}{R_{ij}} \right) \quad (2.18)$$

207 where

$$R = |\vec{r} + \vec{u}| \quad (2.19)$$

$$\vec{u} = \gamma_1 * \vec{a} + \gamma_2 * \vec{b} + \gamma_3 * \vec{c} \quad (2.20)$$

208

209 2.2 COMPUTATION

210 Simply deriving the equations for the PDF is not enough. The many body nature of  
211 the PDF equation make analytical solution of the structure from the PDF impossible.  
212 Thus, the PDF must be computed from a structural candidates and compared against  
213 experimental results to evaluate the reliability of the model.

214 **HPC and GPUs**

215 To properly solve the structure of materials the PDF will need to be computed many  
216 times and checked against experimental results. This requires computation of the  
217 PDF, potentially over many atoms. Calculating these PDFs requires a fast, highly  
218 parallelized, computational framework.

219 **GPUs and Parallelization**

220 Computing the PDF is an embarrassingly parallel problem. The basic procedure is  
221 to calculate the reduced structure factor  $F(Q)$  for each atom pair and momentum  
222 transfer vector, sum over all the atom pairs, and Fourier transform the structure to  
223 the PDF. The first part of this procedure is perfectly parallelizable, as each atom pair is  
224 separate from the others. The summation over all the atomic reduced structure factors  
225 can be parallelized via distributed summing. Lastly the FFT can be parallelized using  
226 existing parallel FFT algorithms.

227 GPUs are particularly well suited to the task of computing PDFs. GPU chip  
228 architecture is designed to perform many tasks simultaneously by having potentially  
229 thousands of cores.

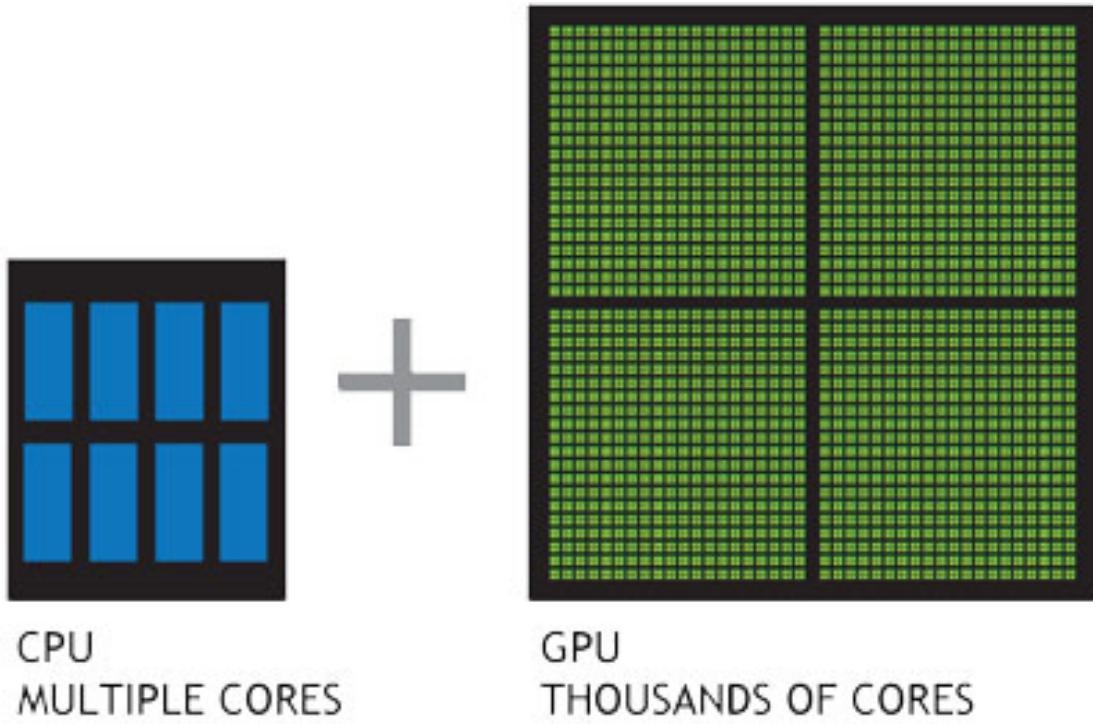


Figure 2.1: Comparison of the CPU and GPU chip architectures

### 230 Map from $ij$ space to $k$ space

231 The above equations, although formally correct, are very inefficient.  $F(Q)$  and its  
 232 gradient are indexed over all the atoms twice, however there are symmetries that  
 233 allow us to only compute over the atom pairs essentially mapping from an  $n \times n$  space,  
 $ij$  space, to a  $\frac{n(n-1)}{2}$  space,  $k$  space. For  $F(Q)$  we apply the following mapping where

$$\begin{array}{ccc}
 E & \xrightarrow{\psi} & E' \\
 \downarrow \phi & & \xrightarrow{\Sigma} Z \\
 B & \xrightarrow{\psi'} & B' \\
 & & \nearrow \Sigma'
 \end{array}$$

234

235  $E$  denotes the atomic coordinates in  $ij$  space,  $E'$  denotes  $F(Q)$  before the summation  
 236 in  $ij$  space,  $B$  denotes the atomic pairs in  $k$  space,  $B'$  denotes  $F(Q)$  in  $k$  space, and  
 237  $Z$  denotes the final summed  $F(Q)$ . For the operators,  $\phi$  denotes the mapping from

238  $ij$  space to  $k$  space  $k = j + i * \frac{i-1}{2}$ ,  $\psi$  and  $\psi'$  denote the  $F(Q)$  operation in  $ij$  and  $k$   
 239 space, respectivly.  $\Sigma$  denotes the sum over all the atoms.

240 To properly define  $\Sigma'$  we must establish whether  $F(Q)$  is an even function. We  
 241 can accomplish this by examining each of the portions of  $F(Q)$ ,  $\alpha, \beta, \tau, \Omega$ .  $\Omega$  is even,  
 242 since  $r_{ij}$  is the interatomic distance, which is the same despite a flip of indicies,  $Q$   
 243 does not depend on the atomic indicies, and since  $Qr_{ij}$  is even so is  $\sin Qr_{ij}$ . Thus,  
 244  $\Omega$  is even. Providing similar analysis to  $\tau$  we can see that while  $\vec{u}_{ij}$  is odd, so is  
 245 the unit displacement vector between the two atoms, thus the two odds cancel out.  
 246 Intuitivly this makes sense, since the  $F(Q)$  equation is fundamentally interested in the  
 247 interatomic distances which is even. Thus, switching atom indicies does not change  
 248  $F(Q)$ . Due to the even nature of the  $F(Q)$  operator the  $\Sigma'$  operator sums over all the  
 249 atom pairs, and multiplies by two to reflect the double counting of the  $\Sigma$  operator.

For the gradient a similar mapping is used:

$$\begin{array}{ccccc} E & \xrightarrow{\psi} & E' & \xrightarrow{\Sigma} & Z \\ \phi \downarrow & & & \nearrow \tilde{\phi}\Sigma & \\ B & \xrightarrow{\psi'} & B' & & \end{array}$$

250

251 In this mapping, however, we use the  $\tilde{\phi}\Sigma$  operator. This operator simultaniously  
 252 performs a reverse mapping from  $k$  to  $ij$  space, and a summation with the correct  
 253 symmetry. In this case the  $\psi$  and  $\psi'$  operators, which denote the  $\vec{\nabla}F(Q)$  operator  
 254 in  $ij$  and  $k$  space, are antisymmetric. Intuitivly this makes sense as an extension of  
 255 Newton's Second Law, since each particle's interation is felt oppositely by its partner.

## 256 GPU Memory Allocation

257 While GPUs are very fast computational engines they tend to be memory bound.  
 258 While a gradient array for a 10nm Au nanoparticle, consisting of 31,000 atoms and

259 half a billion unique distances, occupies 1.5 TB of memory a single GPU's RAM  
 260 allotment varies from 4GB on a NVIDIA GTX970 to 24 GB on a NVIDIA Tesla K80.  
 261 Thus, it is important to determine exactly how many atoms can fit on a GPU of  
 262 arbitrary size as a function of the number of atoms and the  $Q$  range. The memory  
 263 required per array is:

$$q[=]3n \quad (2.21)$$

$$d[=]3k \quad (2.22)$$

$$r[=]k \quad (2.23)$$

$$scatter[=]nQ \quad (2.24)$$

$$normalization[=]kQ \quad (2.25)$$

$$\omega[=]kQ \quad (2.26)$$

$$F_k(Q)[=]kQ \quad (2.27)$$

$$Sum[=]kQ \quad (2.28)$$

$$Sum2[=]kQ \quad (2.29)$$

$$F(Q)[=]Q \quad (2.30)$$

264 where  $n$  is the number of atoms,  $k$  is the number of unique distances,  $Q$  is the scatter  
 265 vector. Each of the above arrays are used in the computation and thus must be able  
 266 to be held in memory. Thus the number of atom pairs that can fit on a GPU with  
 267  $am$  bytes of available memory is:

$$k_{perGPU} = \frac{1}{16Q + 16} (-4Qn - 4Q + am - 12n) \quad (2.31)$$

268 If ADPs are included in the calculation, then the following arrays are also added to  
 269 the memory allocation:

$$adps = 3n \quad (2.32)$$

$$\sigma = k \quad (2.33)$$

$$\tau = kQ \quad (2.34)$$

270 Thus the pair allotment is:

$$k_{perGPU} = \frac{-4Qn - 4Q + am - 24n}{20Q + 20} \quad (2.35)$$

271 For the Gradient we need to calculate  $F(Q)$  and its gradient, so the total memory  
272 overhead is equal to the previously mentioned arrays plus:

$$g_{omega} = 3kQ \quad (2.36)$$

$$g_fq = 3kQ \quad (2.37)$$

$$rtn = 3nQ \quad (2.38)$$

273 Thus the gradient allotment is:

$$\vec{\nabla}k_{perGPU} = \frac{-16Qn + am - 12n}{32Q + 16} \quad (2.39)$$

274 For the gradient with ADPs the ADP gradient array is:

$$g_tau = 3kQ \quad (2.40)$$

275 Thus the allocation is:

$$\vec{\nabla}k_{perGPU} = \frac{-16Qn + am - 24n}{48Q + 20} \quad (2.41)$$

276 These equations were solved by sympy as their validity is very important to the overall  
277 reliability of the software. If the GPU is overallocated then the system may crash or  
278 return meaningless results.

279

## CHAPTER 3

280

### BENCHMARKING

281

This entire section needs some rewriting to distinguish this from the paper

282

The NUTS-HMC system was tested on a series of nanoparticle (NP) benchmarks.

283

The purpose of these benchmarks is to test the ability of the NUTS-HMC system to reproduce the target PDF and its associated structure. Systems were chosen for their size, crystallinity, and interfacial differences.

286

#### 3.1 PDF

287

The formation of NPs with both crystallographic and non-crystallographic structures [?] and with different chemical patterns [?] are well documented. For simplicity, we chose monometallic Au clusters as benchmarks and considered two groups of structures with different size and degrees of structural disorder in order to assess the reliability and efficiency of our HMC method for solving atomic structures from PDFs. The first group consists of  $\text{Au}_{55}$  clusters with different degrees of disorder, including a crystalline cluster structure in  $O_h$  (Octahedral) symmetry, a structure with a disordered surface, and an amorphous structure. The second group consists of the crystallographically solved  $\text{Au}_{102}$  structure as in the  $\text{Au}_{102}\text{MBA}_{44}$  nanocrystals [?, ?]. We used optimized structures from the Density Functional Theory (DFT) as target structures and generated the corresponding PDF,  $G_{\text{obs}}$ , according to

$$G_{\text{obs}} = \frac{2}{\pi} \int_{Q_{\min}}^{Q_{\max}} Q[S_{\text{obs}}(Q) - 1] \sin(Qr) dQ \quad (3.1)$$

298 where  $S_{\text{obs}}$  is the target structure's structure factor. Since all the target structures  
299 were optimized by DFT at zero Kelvin the target and model PDF profiles were  
300 calculated at zero temperature, with no atomic displacement parameters (ADPs).  
301 However, ADPs would have a considerable impact on the calculation of the PDF,  
302 especially for nanoparticles at non-zero temperatures.

303 Spin-polarized DFT calculations were carried out using the Vienna ab initio sim-  
304 ulation package (VASP) [?, ?] within the Perdew-Burke-Ernzerhof (PBE) exchange-  
305 correlation functional [?]. The projected augmented wave method [?] and a kinetic  
306 energy cutoff of 400 eV were used. Structural optimization was performed until the  
307 total energy and ionic forces were converged to  $10^{-6}$  eV and 10 meV/Å, respec-  
308 tively. The amorphous Au<sub>55</sub> structures were generated by simulated annealing using  
309 the classical embedded atom method potential [?]. Different annealing temperatures  
310 between 1200 K and 1670 K (bulk melting temperature of Au) were used and the  
311 thermally equilibrated structures were cooled down to 300 K before minimization at  
312 0 K. Further optimization using DFT leads to total energies that vary within 1-2  
313 eV among different amorphous structures and the lowest energy one was used as the  
314 target structure. The target structure of Au<sub>102</sub> was taken as the Au<sub>102</sub> core of the  
315 DFT-optimized Au<sub>102</sub>MBA<sub>44</sub> cluster [?].

316 All systems were refined using a PES which consists of a linear combination of  
317  $Rw$ , the repulsive and attractive thresholded spring potentials. The total potential  
318 energy in the Hamiltonian in Eq. (1.9) is expressed as:

$$U(q) = U_{Rw}(q) + U_{\text{spring}}(q, R_{\min}) + U_{\text{spring}}(q, R_{\max}) \quad (3.2)$$

319 The thresholded spring potentials are based on those previously proposed on by Pe-  
320 terson [?], i.e.  $U_{\text{spring}}(q, r_t) = \frac{\kappa}{2} \sum_{i,j} (r_{i,j} - r_t)^2$  for all atomic distance  $r_{i,j}$  outside the  
321 bounds of the spring threshold  $r_t$ . The resulting restoring forces on the out-of-bound  
322 atoms bring the system back within the bounds of the PDF,  $R_{\min}$  and  $R_{\max}$ , and  
323 therefore preventing the system from exploding or collapsing. Otherwise, incorrect

324 refinements may result by having atomic pair distances out of the PDF bounds.  $\kappa$  is  
325 the spring constant in eV/Å and the  $Rw$  potential is converted from unitless to eV  
326 via multiplication by a conversion factor  $\lambda$ .

327 Whereas the choice of the absolute values of  $\lambda$  and  $\kappa$  is somewhat arbitrary, their  
328 relative values are important in determining which term in Eq. (3.2) dominates the  
329 PES, especially when considering the effect of the simulation temperature. Generally,  
330 the ratio between the total potential energy and the temperature determines how  
331 much random motion will dominate the dynamics; a lower ratio implies that random  
332 motion will play a large role in the dynamics. The ratio between  $\lambda$  and  $\kappa$  of each  
333 spring describes how far the PDF can push the system below or above the bounds set  
334 by the spring potentials. Heuristically, too stiff a spring forbids the system to access  
335 new configurations, e.g. high energy “transition states” which may involve shorter  
336 bonds or a larger system size. Conversely, too small a spring constant makes it slower  
337 for the system to snap back within bounds and may lead to an explosion or implosion  
338 of the system, leaving the dynamics to drift aimlessly.

## 339 Model Parameters

340 Unless otherwise stated, the PDFs of the target and starting structures were generated  
341 using Eqn. (3.1) with a step of  $\delta R = .01$  Å,  $Q_{\min} = 0.1$  Å<sup>-1</sup>,  $Q_{\max} = 25.0$  Å<sup>-1</sup>.  $R_{\min}$   
342 and  $R_{\max}$  correspond to the first minimum before the first PDF peak and that after  
343 the last PDF peak, respectively, which ensure that the full meaningful region of the  
344 PDF is modeled. For each of the simulations, the Q resolution was calculated by

$$\delta Q = \frac{\pi}{R_{\max} + \frac{12\pi}{Q_{\max}}} \quad (3.3)$$

345 The HMC simulation was run with  $N = 300$  iterations, a target acceptance rate  
346 of 0.65, and an average starting momentum for each NUTS iteration of 10 eVfs/Å.  
347 Both repulsive and attractive spring potentials are used with  $\kappa = 200$  eV/Å and

348 thresholds matching  $R_{\max}$  and  $R_{\min}$  of the PDF, respectively.  $\lambda = 300$  eV was used  
349 as conversion factor for  $Rw$ . Each simulation was run with a pair of Nvidia GTX970  
350 graphics cards, with one card partially occupied with desktop visualization.

351 **Au55: surface relaxed**

352 We first test our algorithm by solving the crystalline Au<sub>55</sub> (*c*-Au<sub>55</sub>) cluster structure  
353 from its PDF. The starting structure is taken as the bulk-cut cuboctahedron of Au<sub>55</sub>  
354 with a uniform bond length of 2.89 Å. Due to finite-size and surface effects, the DFT-  
355 relaxed cluster structure shows a distinctively different bond length distribution as a  
356 function of the bond's distance to the cluster center of mass, and therefore is difficult  
357 to model with a small box approach which assumes an identical unit cell throughout  
358 the whole system.

359 **Run Parameters**

360  $R_{\min}$  and  $R_{\max}$  for this simulation were 2.45 Å and 11.4 Å, respectively, with  $\delta Q =$   
361 0.24 Å<sup>-1</sup>. The simulation ran for approximately 34 minutes, over a total of ~40  
362 thousand configurations. The results are shown in Fig. 3.1.

363 The PDF, radial bond distribution, and bond angle distribution show good agree-  
364 ment between the target and final fitted structures, with a  $Rw$  of 0.3% whereas  $Rw$   
365 of the starting structure is as high as 44.8%. DFT calculations yield a total energy of  
366 the final structure very close to that of the target structure (within a few meV). The  
367 success in the fitting is largely attributed to the factor that the target structure is  
368 only locally (and mildly) disturbed from its bulk-like counterpart and therefore there  
369 is no need to overcome any high PES barriers to reach the correct solution. As shown  
370 below, the situation is rather different for much more disordered target structures.  
371 Interestingly, the small-box solution using PDFgui[?] yields a rather large  $Rw$  of 43%,

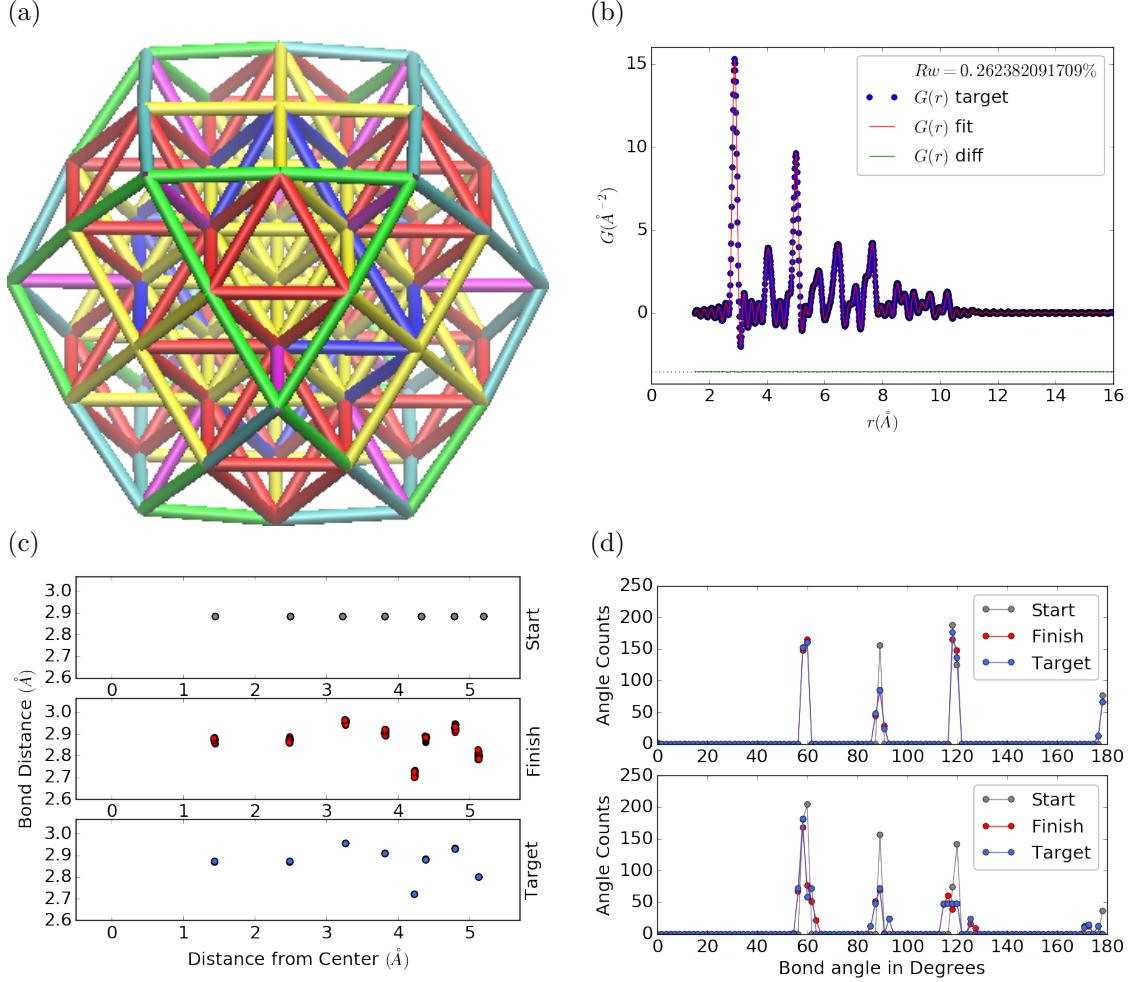


Figure 3.1:  $\text{Au}_{55}$  PDF fitting of DFT-optimized  $c\text{-}\text{Au}_{55}$ . (a) the final structural solution ( $Rw=0.3\%$ ) with bond lengths color-coded by step of  $0.05\text{\AA}$ , (b) the target PDF(blue dots) overlaid with the PDF of the final structure (solid red lines) with the difference in green lines offset below, ?? the radial bond distribution, and (d) bond angle distribution.

372 due to the failure to fit the surface contracted atoms with a unit cell. The PDF fits  
 373 of the starting structure and small-box solution are shown

374 Put this somewhere

375 .

376 **Au55: surface disordered**

377 In addition to surface relaxation, the structure of a cluster or nanoparticle is often  
378 disrupted by the presence of defects and/or ligand bound to the surface. To mimic  
379 such surface disorders, we took the DFT-optimized *c*-Au<sub>55</sub> structure from case I as  
380 the starting structure and randomly displaced the surface atoms with a normal distri-  
381 bution of  $\sigma = 0.2 \text{ \AA}$ . All atoms are allowed to move in the HMC simulation, including  
382 the originally undisturbed core, which is a Au<sub>13</sub> cluster with  $O_h$  symmetry.

383  $R_{\min}$  and  $R_{\max}$  for this simulation were 1.95 Å and 12.18 Å, respectively, with  
384  $\delta Q = 0.23 \text{ \AA}^{-1}$ . The simulation ran for approximately 3.6 hours, over a total of  $\sim 270$   
385 thousand configurations. The results of the simulation are shown in Fig. ??.

386 Overall, good agreement is found between PDFs of the target structure and the  
387 final structural solution, even out to larger  $r$ , with an  $Rw = 0.6\%$  starting from an  
388  $Rw = 50.4\%$  (see Fig. S2). The radial bond distribution and angle distribution  
389 show reasonably good agreement, but with lower degree of crystallinity in the final  
390 structure compared to the target structure. The discrepancy is most obvious in  
391 the core: despite the identical core structure in the starting and target structures,  
392 the core atoms were displaced in the HMC simulations in order to achieve a “best”  
393 solution. This is because PDF measures the global average of interatomic distances  
394 between each atomic pair and does not contain direct information about the locality  
395 of such pairs, e.g. on the surface or cores. If such information is obtained a priori, for  
396 example, from theoretical prediction or other experimental measurements, the core  
397 structure can then be fixed and excluded from HMC dynamics.

398 Similar discrepancies are found in the CN distribution. Since the initial displace-  
399 ments of the surface atoms are relatively mild, the interatomic connectivities remain  
400 more or less the same and therefore the target structure has an identical CN distri-  
401 bution to the starting (unperturbed) structure. This is, however, not the case for  
402 the final fitted structure, which shows discernible differences, especially at the low

403 and high CN numbers. This is partly caused by the displacement of the core atoms  
404 mentioned above, and partly by the lack of CN constraints for the PDF fitting, which  
405 has been previously demonstrated in the case of  $\alpha$ -Si [?]. Additional experimental  
406 data, e.g. from EXAFS or NMR, may help to steer the simulations towards better  
407 agreement in both PDF and CN distribution.

## 408 Au55: amorphous

409 Next, we turn to the case in which the entire cluster structure is disordered. We used  
410 a DFT-optimized amorphous  $Au_{55}$  ( $a$ - $Au_{55}$ ) as the target structure, and the DFT-  
411 relaxed  $c$ - $Au_{55}$  cluster from Case I as the starting structure. The total energy of  
412  $a$ - $Au_{55}$  was computed to be *lower* than that of  $c$ - $Au_{55}$  by as large as 2.9 eV, consistent  
413 with the 3.0 eV found in previous DFT work [?].

414  $R_{\min}$  and  $R_{\max}$  for this simulation were 2.6 Å and 11.26 Å, respectively, with  
415  $\delta Q = 0.25 \text{ \AA}^{-1}$ . The simulation ran for approximately an hour, over a total of  $\sim 87$   
416 thousand configurations. The results of the simulation are shown in Fig. ??.

417 Our PDF fitting yielded a final structure of  $Rw$  of 1.7%, whereas that of the  
418 initial structure is as high as 76.1% (see Fig. S3 ), due to the drastically different  
419 atomic structure of the crystalline and amorphous  $Au_{55}$  clusters. Overall reasonable  
420 agreement in PDF, bond angle distribution, and radial bond distance distribution  
421 was found, and the wide spread of the bond lengths was qualitatively reproduced.  
422 However, the mismatch in CNs is problematic, partly due to the lack of information  
423 and/or constraints on the CNs. The total energy of the final structure is computed to  
424 be  $\sim 6$  eV higher than that of the target structure and the difference is substantially  
425 larger than the variation among different amorphous structures computed by DFT  
426 ( $\Delta E_{\text{tot}} \sim \pm 1\text{-}2$  eV). Such a fitting result, despite the rather small  $Rw$ , clearly  
427 indicates the importance of complementary informations and/or constraints necessary  
428 for reliably solving disordered NP structures from PDF.

429 **Au102: triple phase**

430 Our final benchmark is Au<sub>102</sub>, whose structure was initially solved by Jadzinsky and  
431 co workers using x-ray crystallography [?] and further confirmed by DFT studies [?].  
432 The Au<sub>102</sub> structure consists of three main parts, a 49-atom Marks decahedron core,  
433 two  $C_5$  caps consisting of 20 atoms each, and 13 equatorial atoms. Unlike previous  
434 cases, the multi-symmetry nature of the structure, i.e. each part has its own distinct  
435 symmetry, poses a challenge for PDF-based solution of the structure. This is because  
436 of the atomically centralized nature of the PDF, in which each atom “sees” a density  
437 of other atoms surrounding it and has a strong tendency towards becoming the center  
438 of the main symmetry group. Such tendency may lead to a solution where some of  
439 the correct atomic symmetries are discarded in favor of the core symmetry.

440 **Starting from fcc structure**

441 The starting structure was generated by a spherical cut of the fcc bulk lattice, with  
442 two surface atoms removed to conserve the total number of Au atoms.

443  $R_{\min}$  and  $R_{\max}$  for this simulation were 2.7 Å and 16. Å, respectively, with  $\delta Q =$   
444 0.18 Å<sup>-1</sup>. The simulation ran for approximately two hours, over a total of ~82  
445 thousand configurations. The results of the simulation are shown in Fig. 3.6.

446 The initial structure of an fcc bulk-cut cluster, had a starting  $Rw$  of 77.6% (see Fig.  
447 S4), whereas the final structure has a  $Rw$  as low as 8.1%. The disagreement between  
448 the final and target PDFs shows that the majority of the error is in the high  $R$  region,  
449 which is related to the long range distances between the core, caps, and equatorial  
450 atoms. The agreement for other structural metrics is less satisfactory. The bond  
451 angle distribution for core atoms in the final structure has a poor correlation with  
452 those in the target structure, with much broader peak widths. This is likely caused  
453 by the high kinetic barrier to change from one high-symmetry core structure (fcc)  
454 to another (Marks Decahedron). In contrast, the bond angle distribution for surface

455 atoms, which are of lower symmetry than the core, show a much better agreement.  
456 This is due to the preference of Monte Carlo techniques for higher entropy, and thus  
457 lower symmetry, structures. Similarly, the radial bond distance does not show the  
458 correct clustering of bond lengths as expected from an ordered structure, indicating  
459 the amorphous nature of our fit. Finally, the CN distribution shows the largest  
460 discrepancy at CN=12, again due to the amorphous nature of the fit. Overall, the  
461 structural metrics beyond the PDF indicate the poor agreement between the final  
462 and target structures. A higher simulation temperature, potentially combined with  
463 CN or bond length aware potentials (such as DFT or EXAFS derived PESs) may  
464 help to resolve this discrepancy.

465 **Marks decahedron**

466 The starting structure, a Marks Decahedron, was generated by the ASE Cluster tool  
467 with 2 atoms on the [100] face normal to the 5-fold axis, 3 atoms on the [100] plane  
468 parallel to the 5-fold axis and, 1 atom deep Marks reentrance. This produced a  
469 structure with 101 atoms which was extended by one more Au atom to fill out the  
470  $\text{Au}_{102}$  structure.

471  $R$  bounds and Q resolution were the same as the previous case. The simulation  
472 ran for approximately 2.5 hours over a total of  $\sim$ 90 thousand configurations. The  
473 results of the simulation are shown in Fig. ??.

474 The starting structure of Marks decahedron ( $Rw=56.6\%$ , see Fig. S5) yielded  
475 a better structural solution, with a final  $Rw$  of 3.3%. However, the discrepancies at  
476 high  $R$  remains as in the previous case. By examining the final structure, we can see  
477 that these high  $R$  errors are due to a lack of the two 20-atom caps and 13 equatorial  
478 atoms. Similarly, the radial bond distance distribution displays a diffusive behavior  
479 unlike the bond length clustering in the target structure. Compared to the previous  
480 case, the agreement in the CN and bond angle distributions are improved, with the

481 latter capturing nearly all peaks in the target structure with the exception of the 110  
482 bond angle. Relatively large discrepancies are found in the CN distribution at the  
483 low and high ends.

484 **Au147**

485 **3.2 PDF WITH ADPs**

486 **ADP 50**

487 1. Basic 50% larger magnitude

488 2. Random addition to APDs

489 3. Janus ADPs

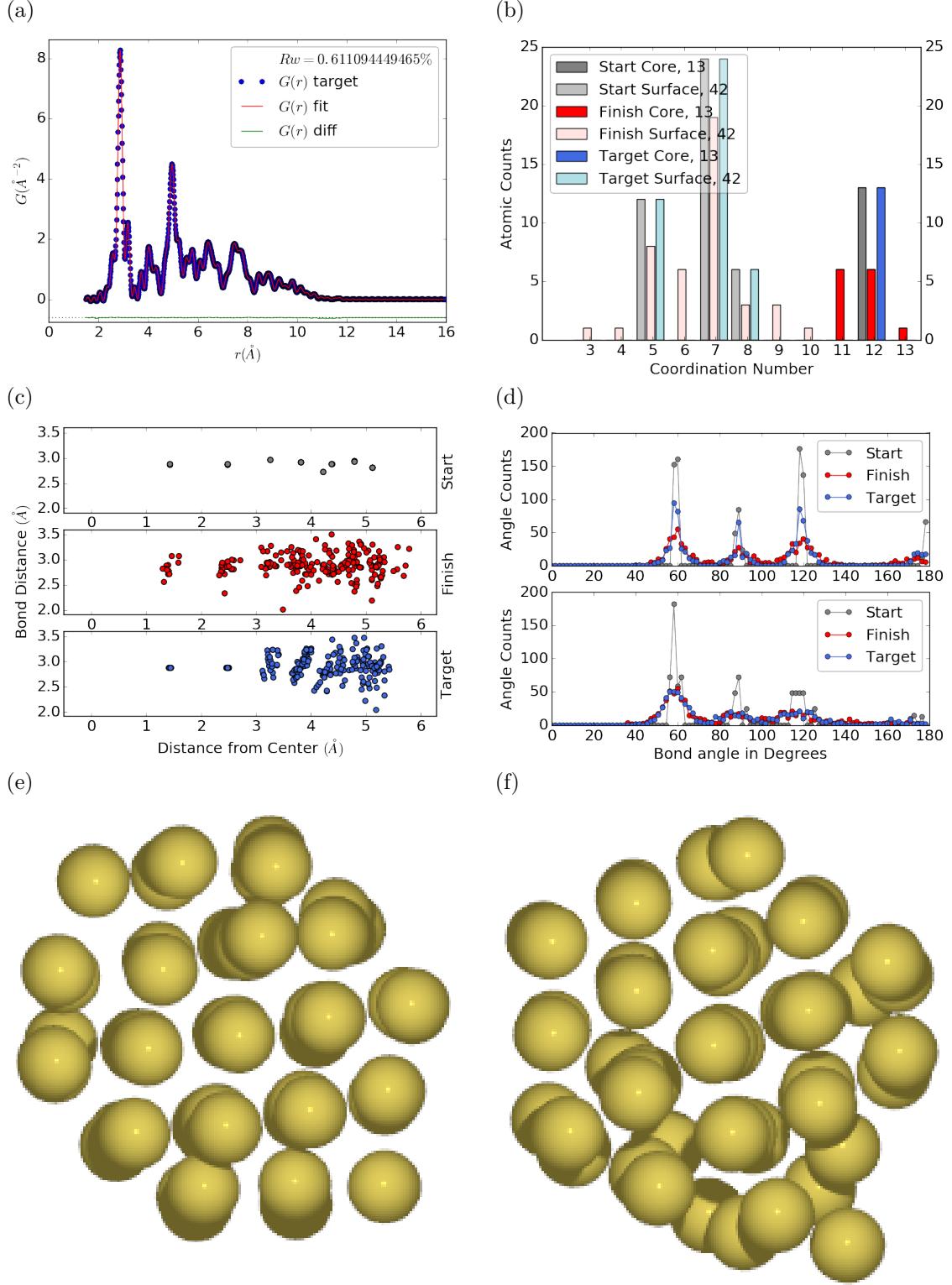


Figure 3.2:  $\text{Au}_{55}$  PDF fitting of surface-disordered  $\text{Au}_{55}$ . a) the target structure, b) the final structural solution ( $R_w=0.6\%$ ), c) the comparison of PDFs, d) the CN distribution with the number of atoms in either the core or the surface, e) the radial bond distribution, and f) the bond angle distribution.

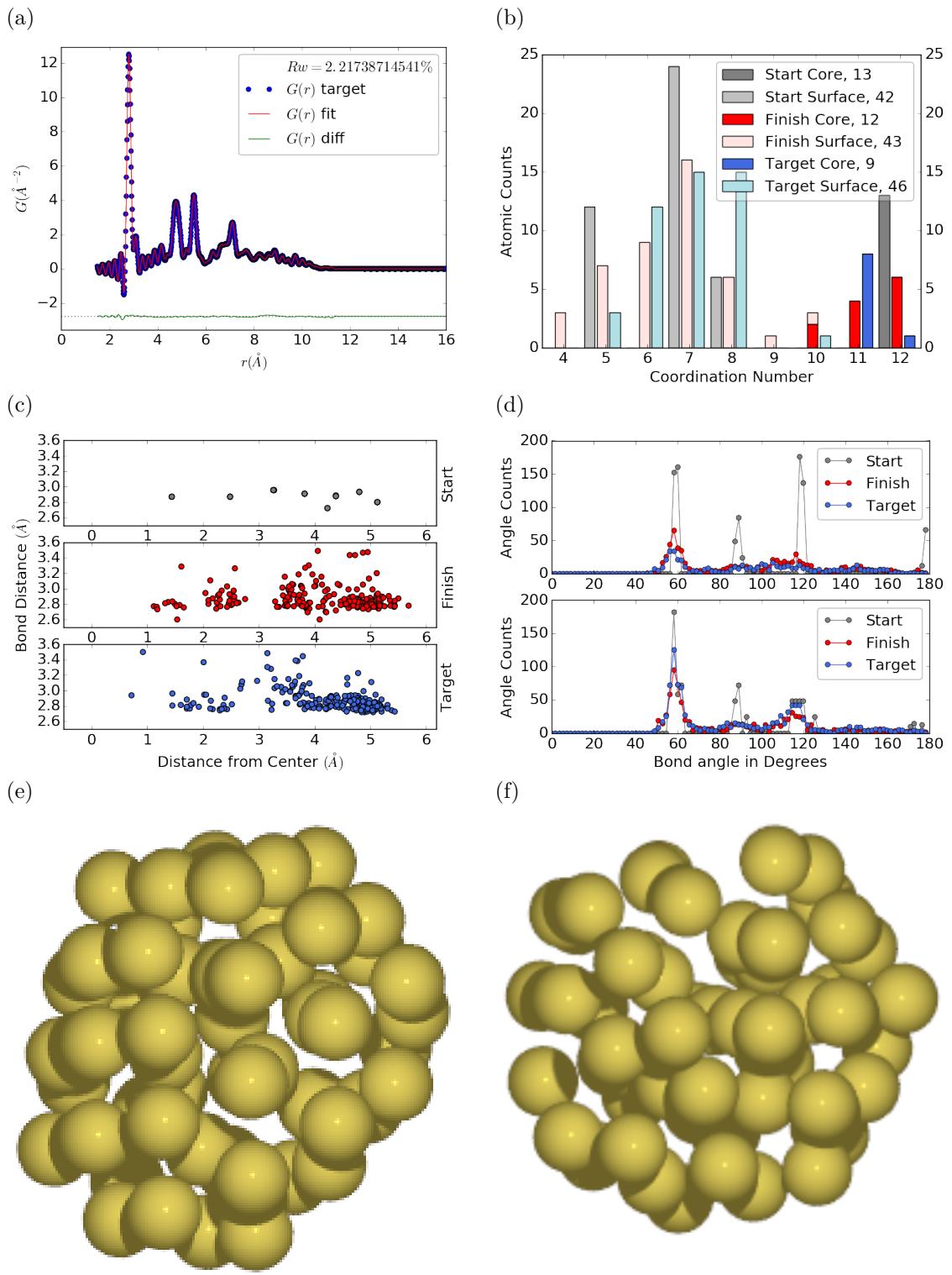


Figure 3.3: Similar to figure 3.2 for DFT-optimized amorphous  $\text{Au}_{55}$ .

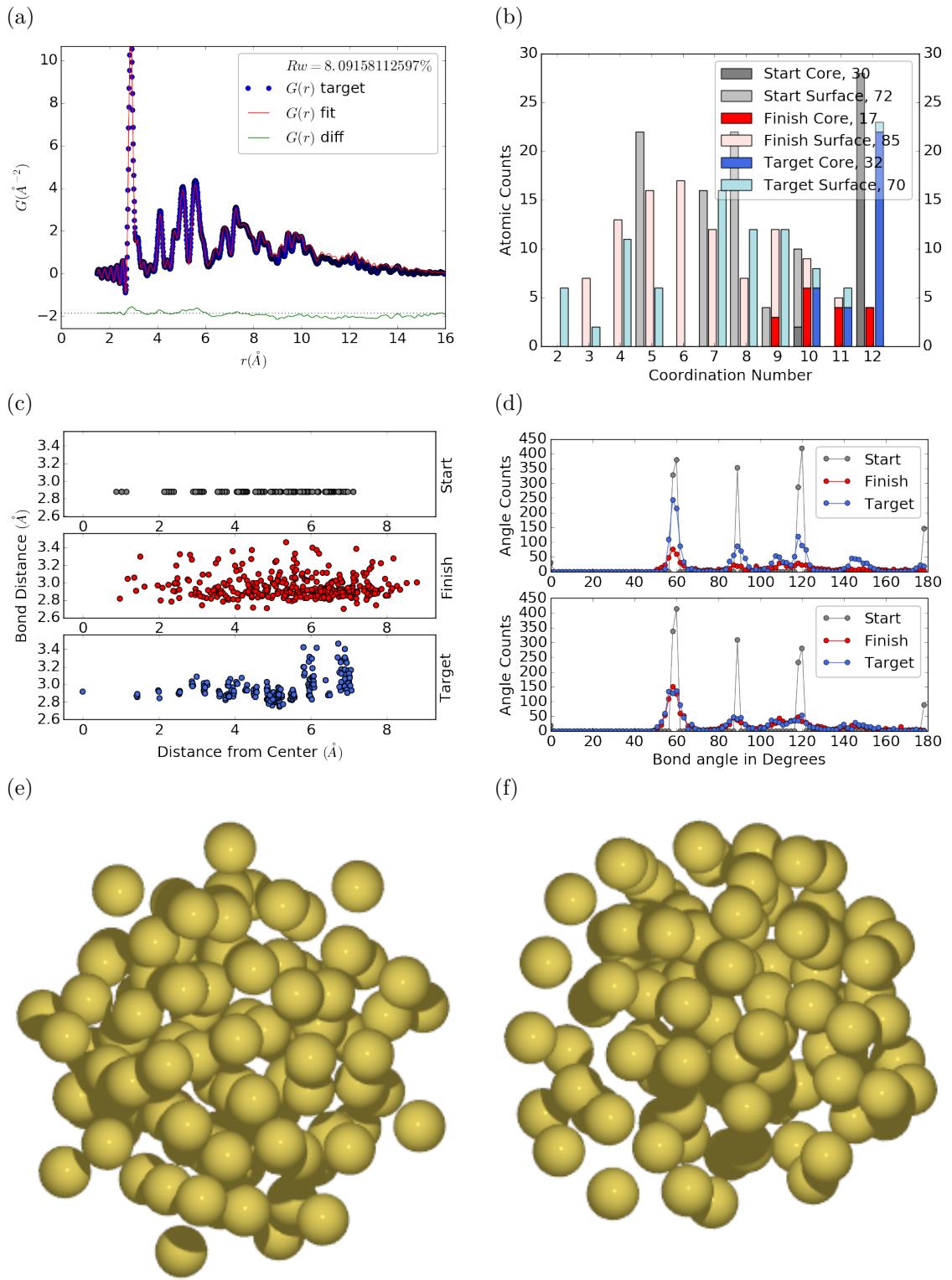


Figure 3.4: Similar to figure 3.2 for  $\text{Au}_{102}$  as in DFT-optimized  $\text{Au}_{102}\text{MBA}_{44}$  cluster.

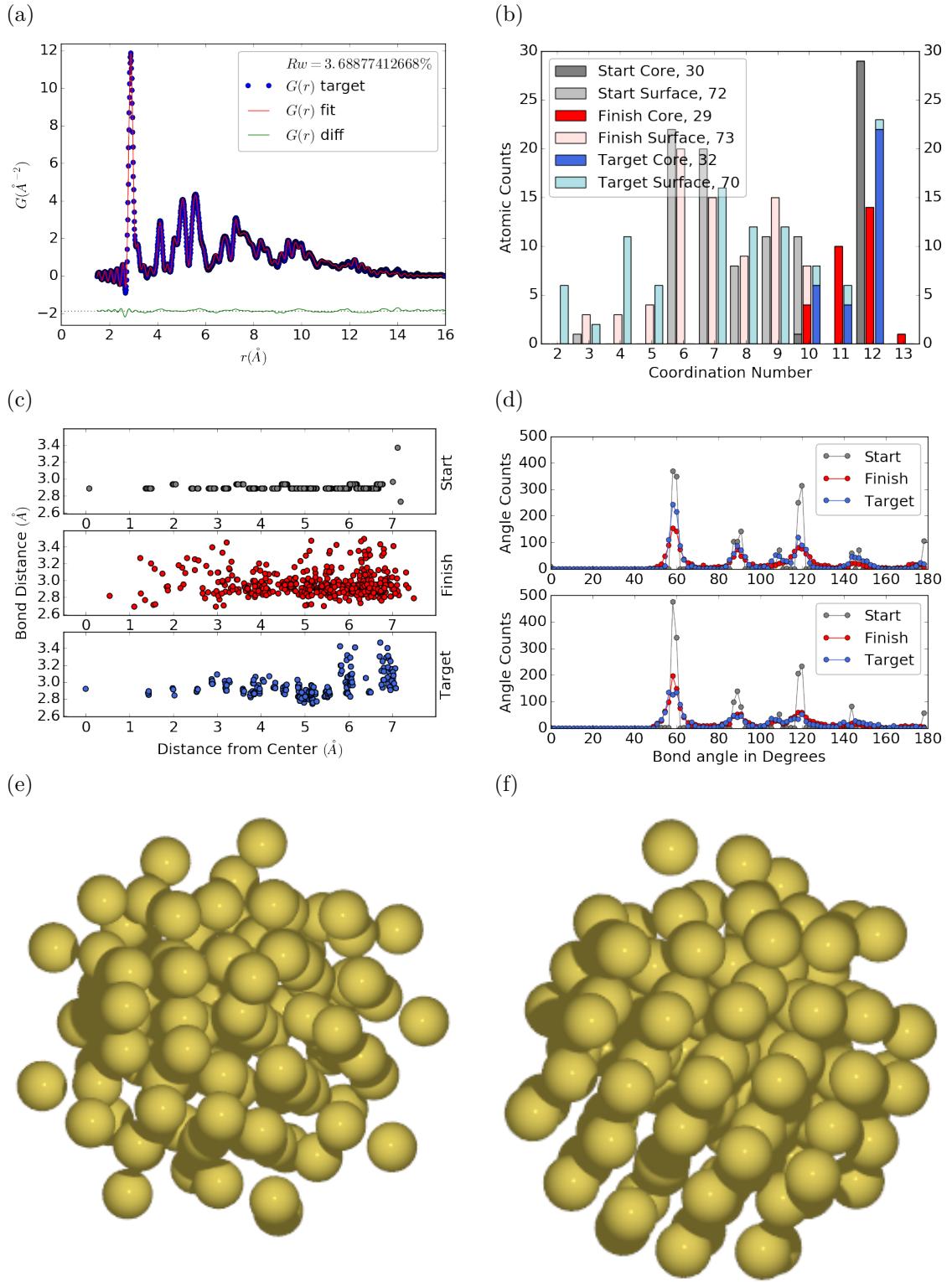


Figure 3.5: Similar to Fig. 3.6 with Marks decahedron as the starting structure.

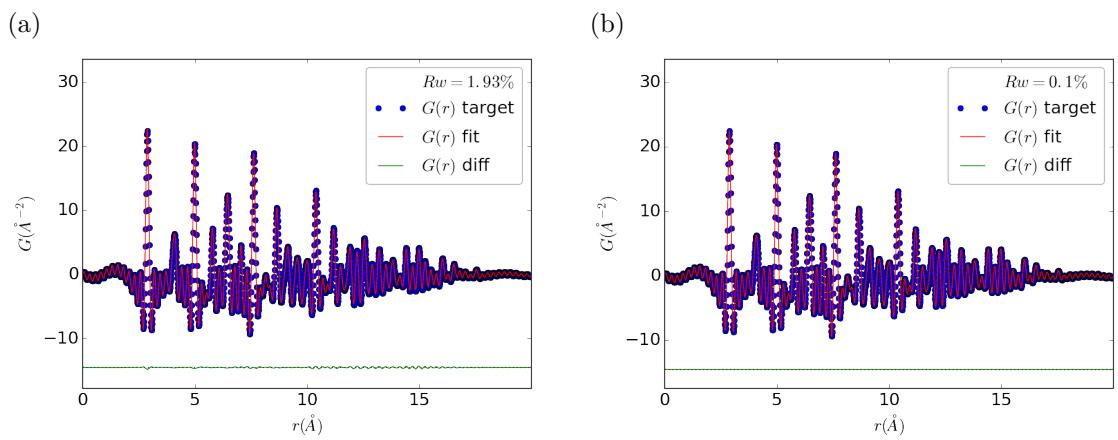


Figure 3.6

490

## CHAPTER 4

491

# X-RAY TOTAL SCATTERING DATA ACQUISITION AND PROCESSING

493

## 4.1 INTRODUCTION

494 X-ray total scattering experiments are generally performed at synchrotron light sources,  
495 as only these sources can provide the needed flux, energy, and high momentum trans-  
496 fer vectors needed to obtain reliable PDFs. [3, 4] Despite the need for a dedicated  
497 facility to perform the total scattering experiments, the experiments themselves are  
498 fairly forgiving, allowing for reactive gaseous environments, experiment temperatures  
499 ranging from 2 K to 1200 K, and even electrochemical cycling. [2, 8, 9] The rapid  
500 PDF data acquisition associated with 2D area detectors creates a data management  
501 problem, as 96 hours of beamtime could result in almost 10,000 images which need  
502 to be associated with the experimental conditions and detector metadata. [3] Finally,  
503 all this data needs to be processed by masking bad pixels and regions, integrating  
504 azimuthally, and converting the scattering data to the PDF. [6, 5, 10, 7, 1]

505

## 4.2 DATA STORAGE AND MANAGEMENT

506 Processing the raw pixel intensities to the PDF is very important as we are extracting  
507 most of our interesting information out of very high  $Q$  data. This data relies on good  
508 statistics and sound background subtraction. Talk about papers from Billinge Group  
509 with thin film PDF and dilute NP solutions. Diagram of the overall data processing  
510 workflow. Discuss the NSLS-II data stack.

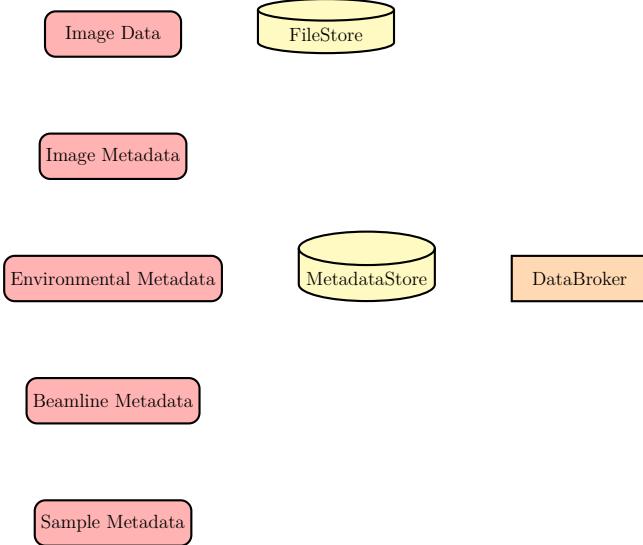


Figure 4.1: Database Loading Workflow. Data is loaded from various sources, including images and text files, into the FileStore and MetadataStore databases. Data is then retrieved from the databases using the databroker.

## 511 MetadataStore Side Loading

512 Design of sidewinder-spec for loading the data into metadatastore. Most of the design  
 513 considerations went into the loaders, which are different for each experiment.

## 514 Detector $Q$ resolution

515 To properly azimuthally integrate the images taken from the detector the  $Q$  resolution  
 516 of the pixels must be calculated. Integrating using even bins will cause pixels which  
 517 are not on the same ring to be binned together, causing the incorrect value of  $I(Q)$   
 518 to be obtained and a larger standard deviation in the integrated data. To properly  
 519 calculate the  $Q$  resolution the resolution of each of the pixels in  $2\theta$  must be calculated.  
 520 Figure 4.2 shows the scattering of x-rays onto a flat image plate detector. In this  
 521 diagram the bottom of the  $n$ th pixel is  $B$  while the top is  $B'$ . The resolution of this  
 522 pixel in  $2\theta$  is  $\angle BAC - \angle B'AC$ . Thus the resolution, calculated from the distances is

$$\Delta 2\theta = \arctan \frac{b}{d} - \arctan \frac{t}{d} \quad (4.1)$$

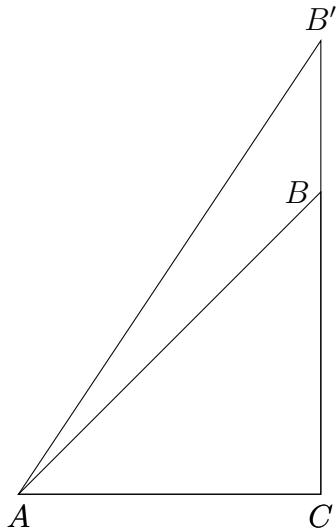


Figure 4.2: Scattering onto a flat detector

523 where  $d$  is the sample to detector distance,  $b$  is the distance to the bottom of a pixel,  
 524 and  $t$  is the distance to the top of that pixel. Note that these distances need to have  
 525 been corrected for detector tilt and rotation. Thus the resolution of a pixel in  $Q$  is

$$\Delta Q = \frac{4\pi(\sin \arctan \frac{b}{d} - \sin \arctan \frac{t}{d})}{\lambda} \quad (4.2)$$

526 where  $\lambda$  is the x-ray wavelength.

527 For a Perkin Elmer image plate, like the one used at the NSLS-II's XPD and the  
 528 APS's 11-ID-B, the resolution function is shown in 4.3. For the same detector the  
 529 number of pixels per  $Q$  is shown in 4.4

## 530 Automated Mask Generation

### 531 Introduction

532 Detector masking is an important part of any x-ray scattering workflow as dead/hot  
 533 pixels, streak errors, and beamstop associated features can be averaged into the data  
 534 changing the signal and its statistical significance. While some features, like the  
 535 beamstop holder, can be easily observed and masked by hand other are much more

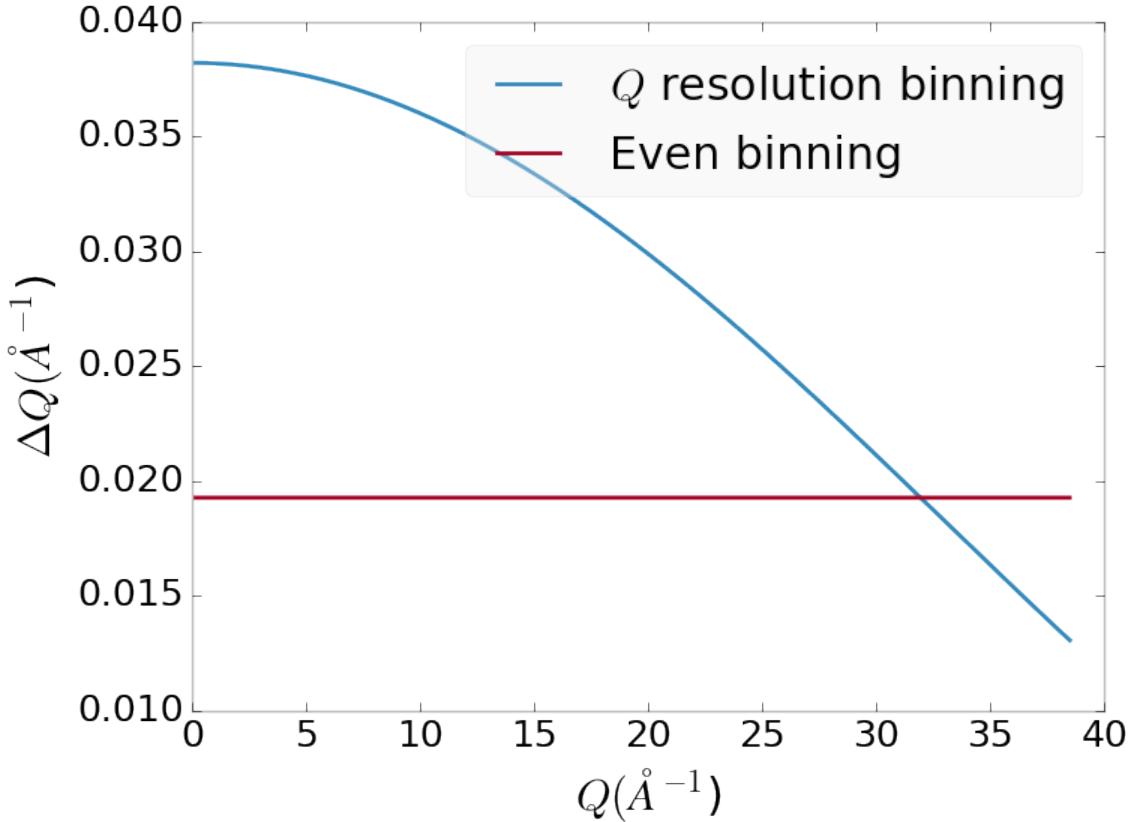


Figure 4.3:  $Q$  resolution as a function of  $Q$ .

536 difficult to observe even on large computer monitors. Additionally, while dead/hot  
 537 pixels and streaks are usually static the hot pixels associated with textured or sin-  
 538 gle crystal scattering or cosmic rays are not. Thus, coming up with an automated  
 539 method for finding such erroneous pixels is important, especially as high flux diffrac-  
 540 tion beamlines can generate data very quickly.

541 While this problem can be quite complex in the most general case, we can use the  
 542 annular symmetry of the powder scattering pattern to our advantage, by comparing  
 543 a pixel against pixels in the same ring. Since non-textured powder scattering should  
 544 produce the same pixel intensity for a given ring we can mask any pixels which are  $\alpha$   
 545 standard deviations away from the mean. This method relies on the aforementioned  
 546 pixel binning algorithm, as using miss sized bins will cause some pixels which should  
 547 be in separate rings to be put together, and others which should be in the same ring

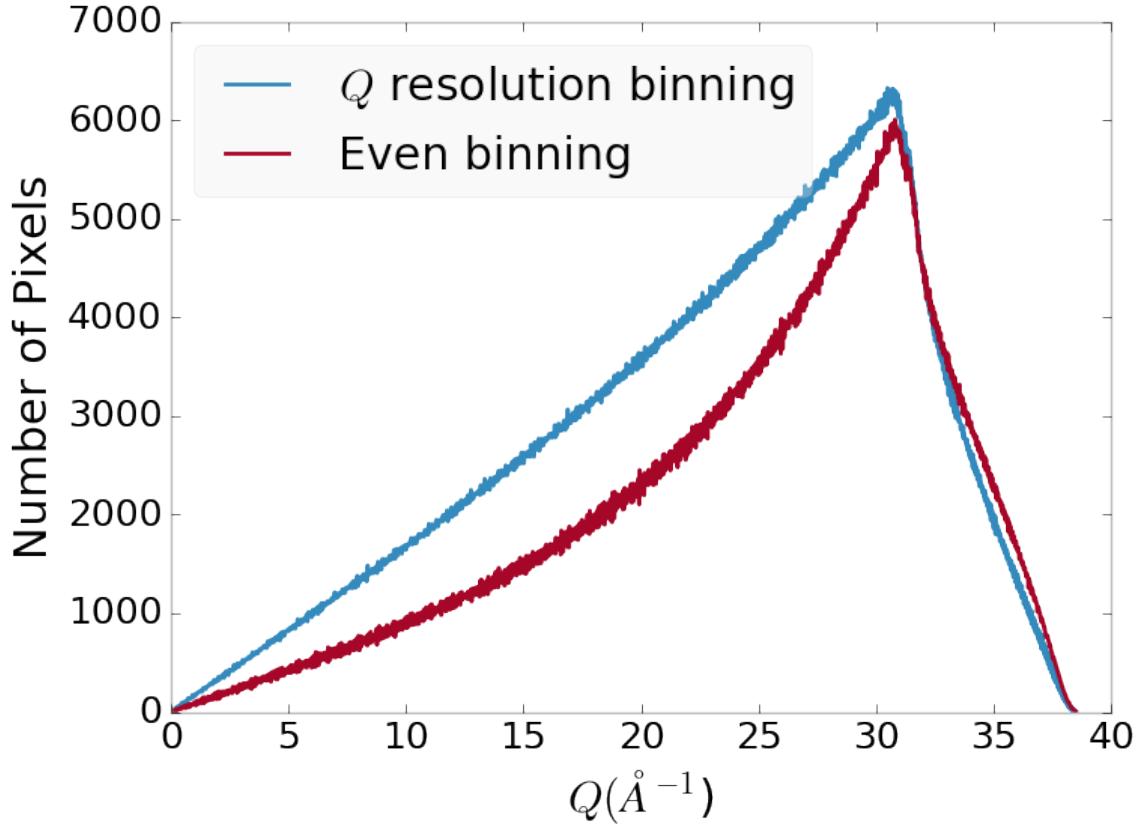


Figure 4.4: Number of pixels as a function of  $Q$ , binned at the  $Q$  resolution of the detector.

548 to be separated. In that case the masking algorithm will overestimate the number of  
 549 pixels to be masked due to the additional statistical variation in the sample.

## 550 Algorithm Design

551 The masking algorithm procedure takes in the image and a description of the pixel  
 552 positions in either distance from the point of incidence or in  $Q$ . The image is then  
 553 integrated twice, producing both the mean  $I(Q)$  and the standard deviation of each  
 554  $I(Q)$  ring. The mask is created by comparing the pixel values against each ring's  
 555 standard deviation and threshold  $\alpha$ . Note that the threshold can be a function of  
 556 distance from the point of incidence or  $Q$ .

557 **Test Cases**

558 To study the effectiveness of the masking we ran the algorithm against both simulated  
559 experimental data. In the case of the simulated data four systems were created: 1)  
560 dead/hot pixels with varying numbers of defective pixels, 2) beamstop holder with  
561 varying beamstop holder transmittance, 3) rotated beamstop holder with varying  
562 beamstop holder transmittance, and 4) beamstop holder with dead/hot pixels. The  
563 base scattering was produced by

$$I = 100 \cos(50r)^2 + 150 \quad (4.3)$$

564 where  $r$  is a pixel's distance from the beam point of incidence. The positions of  
565 the dead/hot pixels were chosen at random as was the dead or hot nature of the  
566 defect. Dead pixels had values from 0 to 10, while hot pixels had values from 200  
567 to 255. The beamstop was positioned at the vertical center of the detector with an  
568 initial width of 60 pixels and final width of 120 pixels. The hight of the beamstop  
569 was 1024 pixels. The beamstop was calculated to attenuate the x-ray scattering  
570 signal at various transmittance, as various beamstop holder materials have different  
571 transmittance. Two version of the masking algorithm were run for each test case, one  
572 using the standard even bin sizes for the integration step, and one where the bin sizes  
573 are tuned to the pixel  $Q$  resolution as discussed in 4.2.

574 **Results and Discussion**

575 Figures 4.5-4.12 show the results of the masking algorithm on simulated images. The  
576 dead/hot pixel masking shows the importance of using the  $Q$  resolution based bin sizes  
577 as the even bin based mask have a tendency to over mask the image, removing pixels  
578 which contain valuable signal. This overmasking is caused by pixels being improperly  
579 associated with one another by the even bins. Figure 4.5 indicates that the masking  
580 algorithm, with the proper binning, masks the image perfectly, with no missed bad

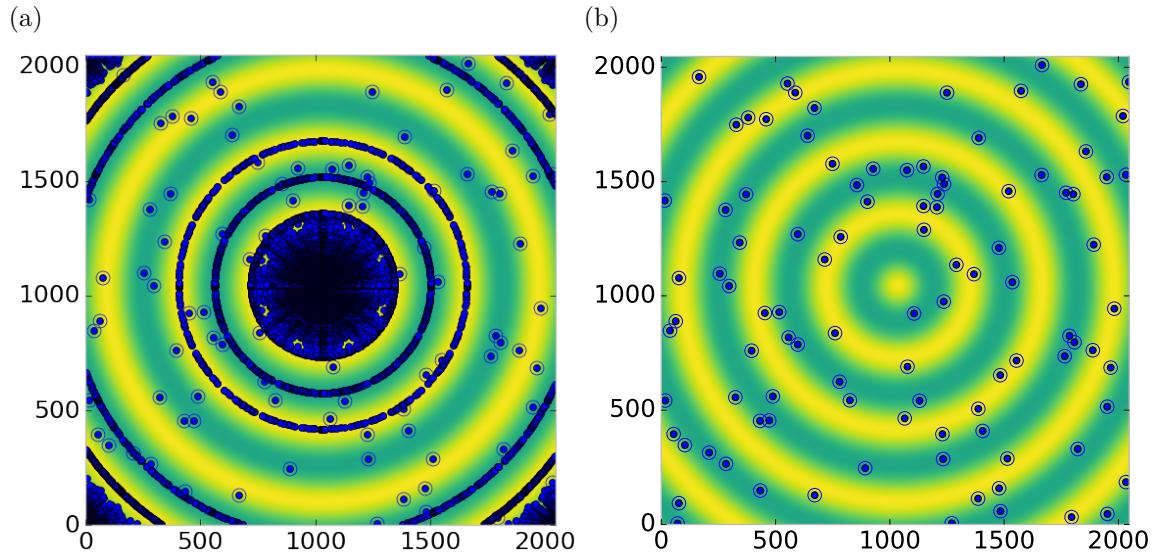


Figure 4.5: Generated dead/hot pixel masks for a detector with 100 bad pixels. a) the standard even bin mask and b) the  $Q$  resolution binned mask. The bad pixels are noted with open circles, masked pixels are noted with closed circles.

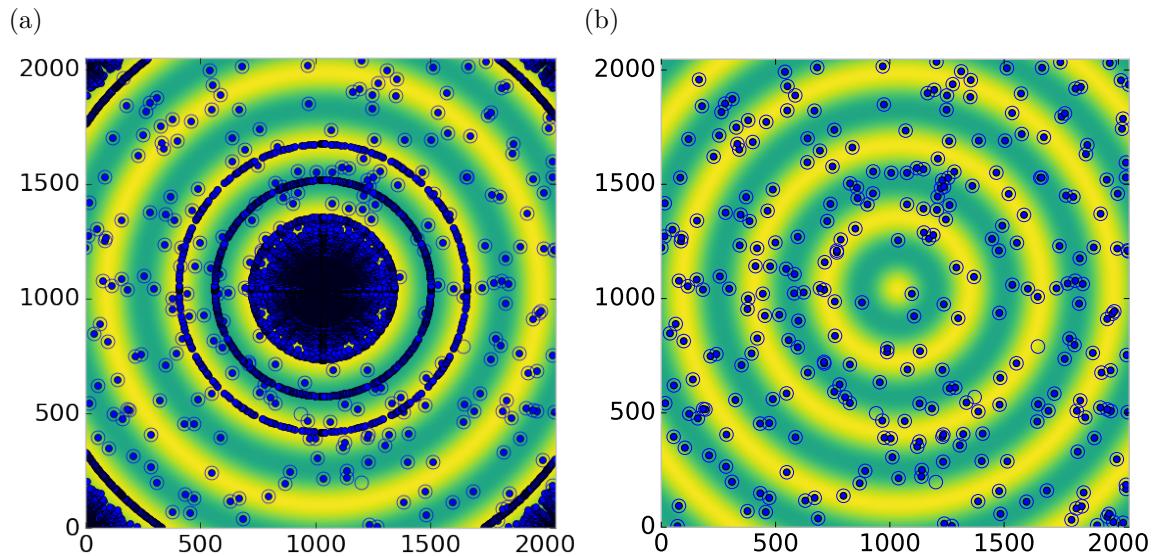


Figure 4.6: Generated dead/hot pixel masks for a detector with 300 bad pixels. a) the standard even bin mask and b) the  $Q$  resolution binned mask. The bad pixels are noted with open circles, masked pixels are noted with closed circles.

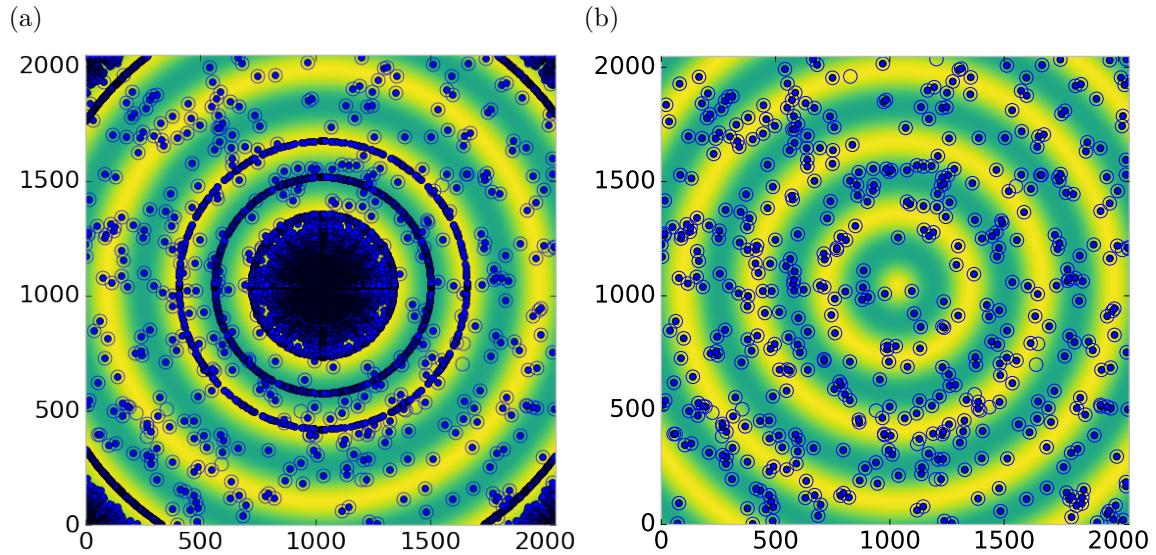


Figure 4.7: Generated dead/hot pixel masks for a detector with 500 bad pixels. a) the standard even bin mask and b) the  $Q$  resolution binned mask. The bad pixels are noted with open circles, masked pixels are noted with closed circles.

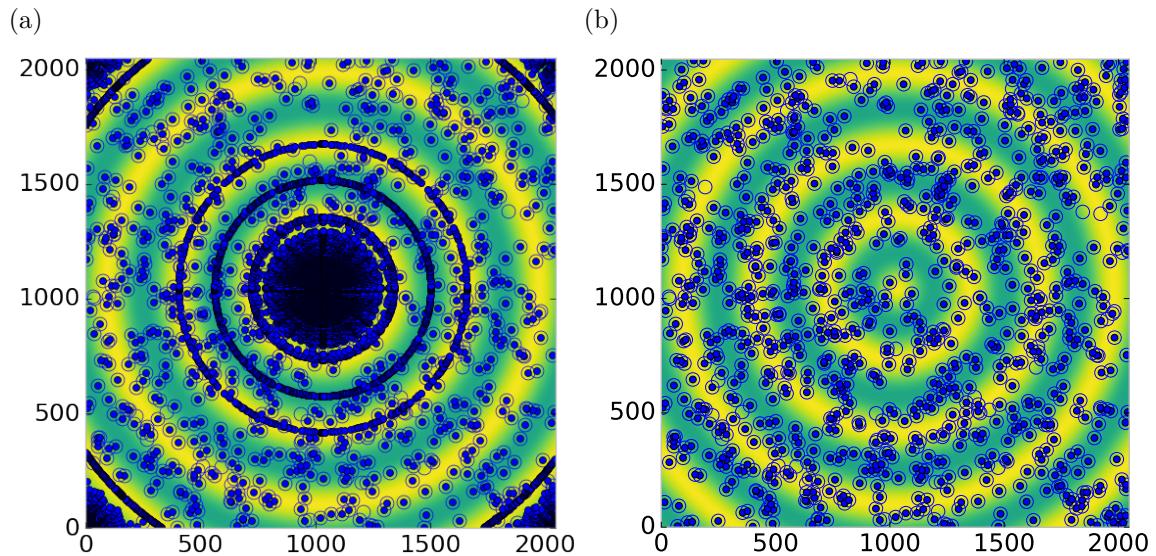


Figure 4.8: Generated dead/hot pixel masks for a detector with 1000 bad pixels. a) the standard even bin mask and b) the  $Q$  resolution binned mask. The bad pixels are noted with open circles, masked pixels are noted with closed circles.

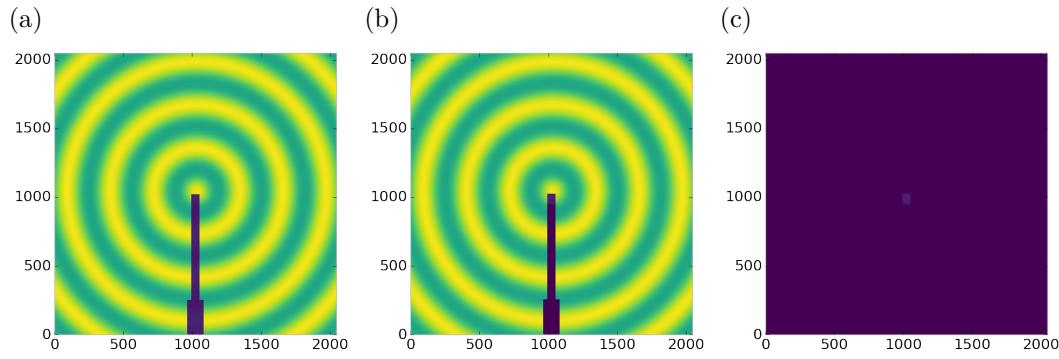


Figure 4.9: Generated beamstop holder masks for a beamstop holder with 10% transmittance. a) the raw image, b) the masked image, c) and the missed pixels

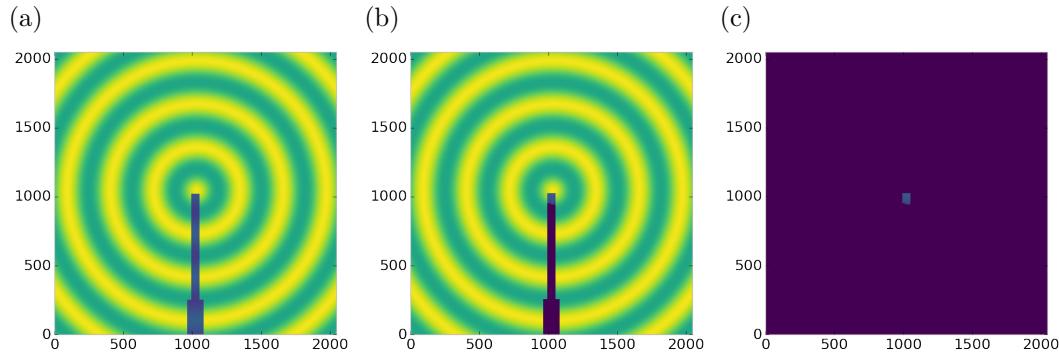


Figure 4.10: Generated beamstop holder masks for a beamstop holder with 30% transmittance. a) the raw image, b) the masked image, c) and the missed pixels

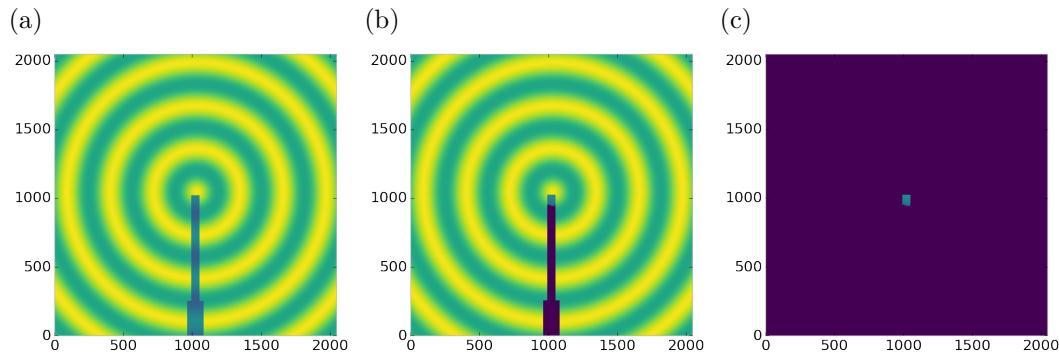


Figure 4.11: Generated beamstop holder masks for a beamstop holder with 50% transmittance. a) the raw image, b) the masked image, c) and the missed pixels

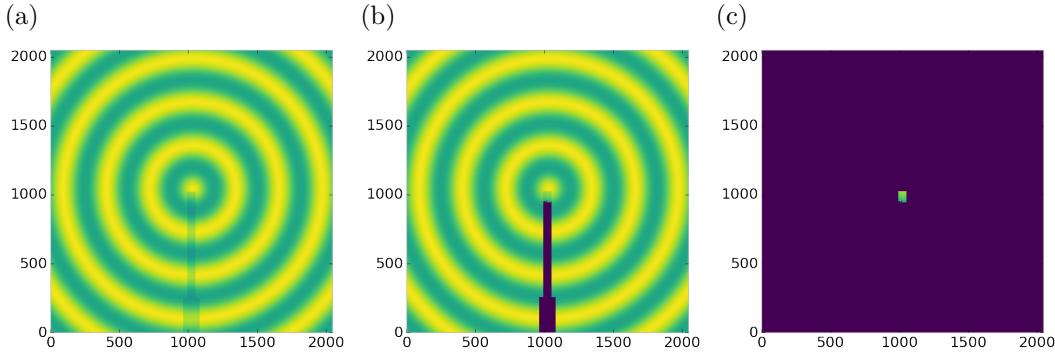


Figure 4.12: Generated beamstop holder masks for a beamstop holder with 90% transmittance. a) the raw image, b) the masked image, c) and the missed pixels

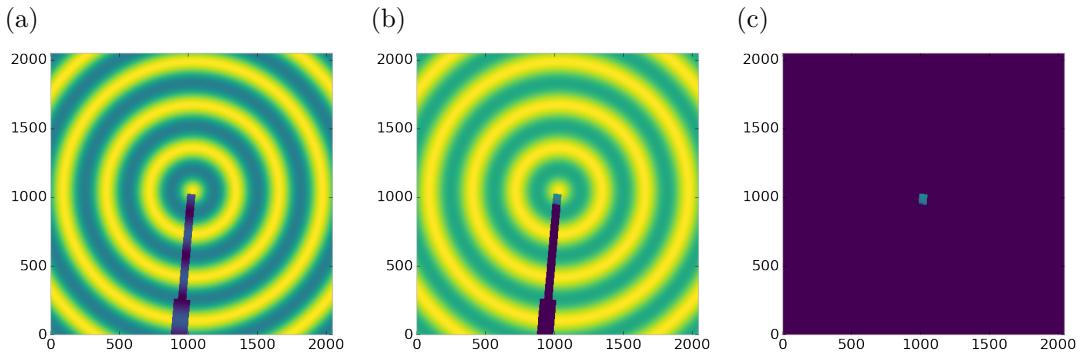


Figure 4.13: Generated beamstop holder masks which is rotated away from verticle

581 pixels or good pixels masked. This is not the case in figures 4.6 - 4.8 as we can see  
 582 pixels which should have been masked but were not. Despite these missed pixels no  
 583 pixels were improperly masked in any of the well binned images. These test cases  
 584 are actually more difficult than experimental data, as the dynamic range of most  
 585 detector causes the dead/hot pixels and single crystal/textue peaks to be orders of  
 586 magnitude away from the desired signal.

587 The beamstop holder masks shown in figures 4.9 - 4.12, which were all run with  
 588 the  $Q$  resolution binning show similar results across the transmittance range, missing  
 589 only a small part of the beamstop holder near the point of incidence. Near this point  
 590 the beamstop holder becomes a statistically significant part of the total number of  
 591 pixels in a given ring, thus it can not be masked out using a statistical search of the  
 592 rings. For most PDF and XRD studies this small area can be masked automaticlly

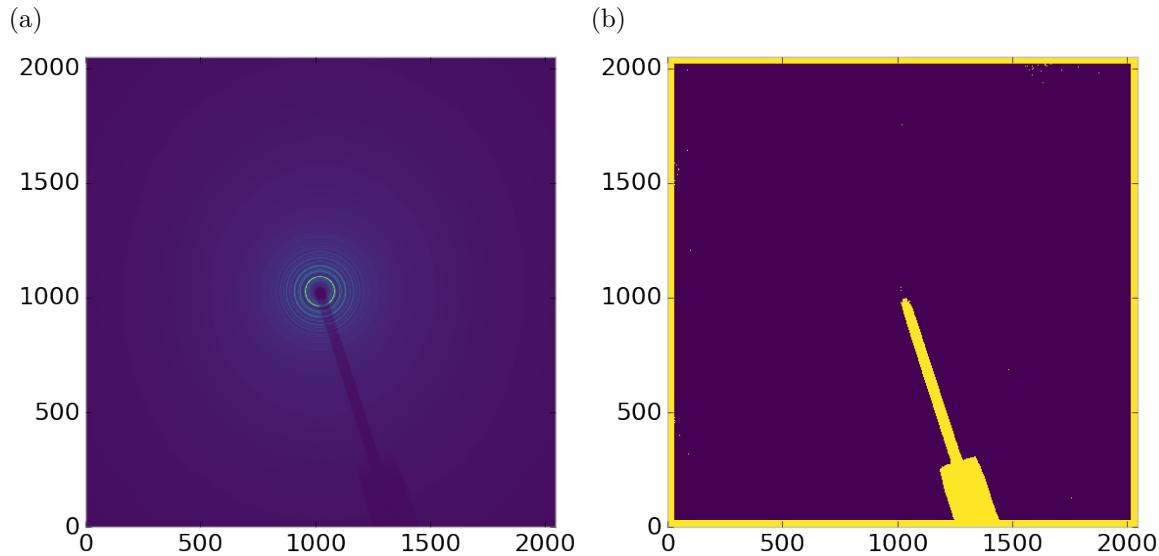


Figure 4.14: Masked experimental data. a) the raw image, b) the mask

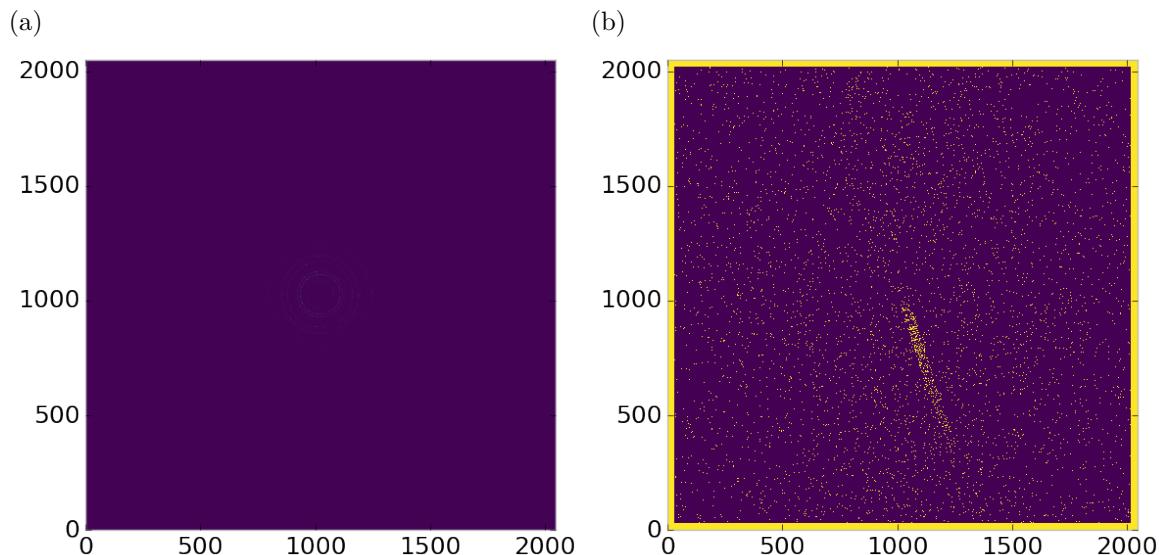


Figure 4.15: Masked experimental data with Pt single crystal signal. a) the raw image, b) the mask

593 by masking all the pixels who's distance from the point of incidence is smaller than a  
 594 given radius  $r$ , or can be negelected outright as the area is not used in the analysis or  
 595 refinement. Similar results were produced for beamstop holders which were rotated  
 596 away from the vericle position, as shown in figure 4.13

597 Working with actual experimental data, obtained at the Advanced Photon Source

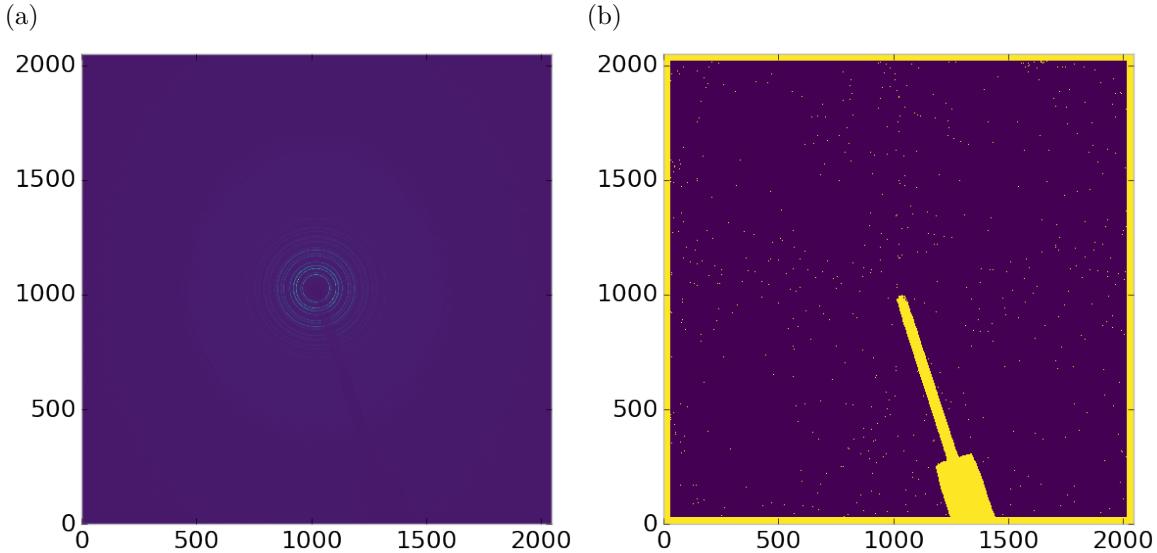


Figure 4.16: Masked experimental data with Pt single crystal signal using figure's 4.14 as a starting mask. a) the raw image, b) the mask

beamline 11-ID-B, shows the difficulty of masking images which have low phonon counts. While the masking of experimental data taken with longer exposures, consisting of 250 .2 second shots, shown in figure 4.14 provides very sharp edges to the beamstop holder, and very little extra masking beyond the occasional dead pixel, this is not the case for the single crystal data. The single crystal data is more problematic because of its short exposure time and low flux, with 500 frame at a .1 second exposure and having shrunk the beam size. The low flux is to prevent the very strong single crystal peaks from damaging the detector. However, this causes the image to be less statistically viable then ideal, causing problems with the mask as seen in figure 4.15. This can be alleviated to some degree by using the previously generated mask as a starting mask for the single crystal image, as shown in 4.16. While the masking algorithm still produces many diffuse masked pixels, they are far fewer, this may be due to the removal of the beamstop which could have contributed to the large standard deviation in figure 4.15.

612 **Conclusions**

613 In this section the masking algorithm, which relies on both  $Q$  resolution based binning  
614 and a statistical approach to azimuthal symmetry, was developed. The focus of  
615 this algorithm was to remove many unwanted detector features associated with pixel  
616 defect, beamstop holder associated scattering attenuation, and single crystal/texture  
617 peaks. Simulated data was used to evaluate the beamstop holder and dead/hot pixel  
618 masking capacity, while experimental data was used to check for single crystal and  
619 texture based masking.  $Q$  resolution based binning was shown to be very important to  
620 avoid overmasking. The ability of the mask writer to mask images is somewhat limited  
621 by the overall statistical image quality, although some deficiencies can be obtained by  
622 using previously generated masks as starting points. This masking algorithm is now  
623 in use in the data processing workflow and will be available in scikit-beam soon.

624 **Automated Image Azimuthal Integration**

625 Using the  $Q$  resolution binning and masking developed in sections 4.2 and 4.2 the  
626 images can be properly integrated. Generally, images are integrated by taking the  
627 mean value of the pixels in a ring. However, other statistical measures of the average  
628 value can be used, like the median.

629 Figures 4.17-4.19 show the importance of masking and the choice of average func-  
630 tion. All the figures were produced using the same dataset, 50 °C  $\text{Pr}_2\text{NiO}_4$  taken at  
631 the APS's 11-ID-B on a Perkin Elmer area detector. The automatic masking alpha  
632 was 3 standard deviations from the mean. While it is difficult to observe the changes  
633 the mask causes in the full  $I(Q)$  plot (subfigures a) and b)), the standard deviation  
634 plots show the effect of bad pixels on the data (subfigure c)). Subfigure c) for figures  
635 4.17-4.19 shows that removal of the beamstop holder lowers the low  $Q$  standard de-  
636 viation from around .1 to almost .01 out to 15  $\text{\AA}^{-1}$ . The high  $Q$  subfigures d) and f)  
637 in figures 4.17-4.19 show the “kink” effect of the detector edge and beamstop holder,

638 where there is a dip in the  $I(Q)$  scattering when the rings include the edge of the  
639 detector. This effect seems to be due to both errors in the edge pixel intensity and the  
640 beamstop holder as masking of the edges only seems to provide only partial removal  
641 of the issue. It is important to note that while integration using the mean of the  
642 ring has issues with only the edge mask, as evidenced by the change in slope in 4.18  
643 d) around  $29.5 \text{ \AA}^{-1}$ , the median integration does not include this error. Ideally the  
644 detector would have a normal distribution of pixel intensity for a given ring, which  
645 would imply an equivalency between the mean and median  $I(Q)$  values. Despite the  
646 closeness of the mean and median once the final mask has been created, it seems that  
647 the median is more reliable, as it was less effected by the beamstop holder in figure  
648 4.18. Thus, for subsequent integrations discussed in this work the median is used to  
649 avoid any defective features that the masking algorithm may have missed.

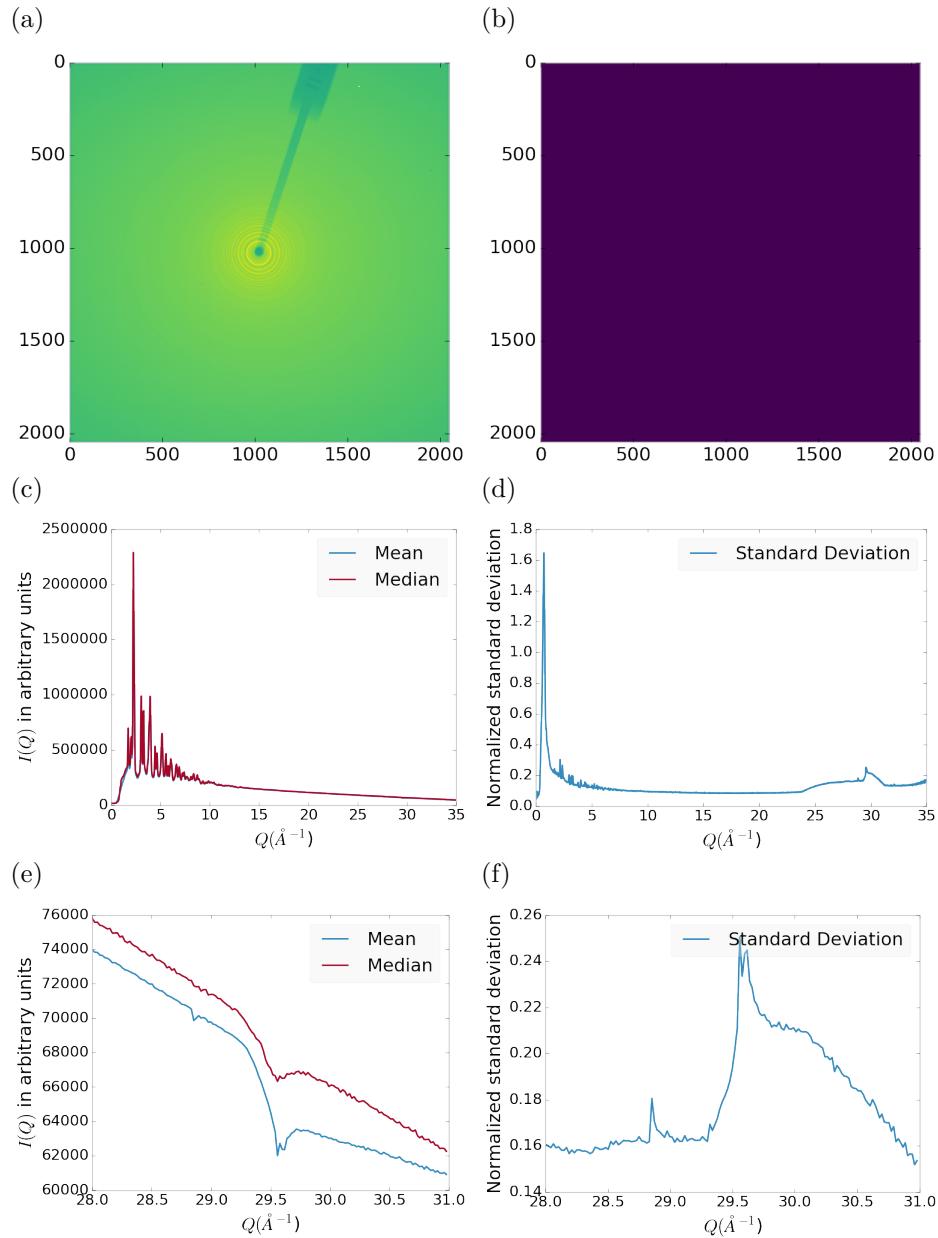


Figure 4.17: Masking, average, and standard deviation of an example x-ray total scattering measurement. This image was produced with no mask. a) the image, b) the mask, c) the mean and median values, d) the standard deviation (normalized to the median), e) a closeup of the  $28 \text{ \AA}^{-1}$  to  $31 \text{ \AA}^{-1}$   $Q$  range for the mean and median, f)  $28 \text{ \AA}^{-1}$  to  $31 \text{ \AA}^{-1}$   $Q$  range for the standard deviation

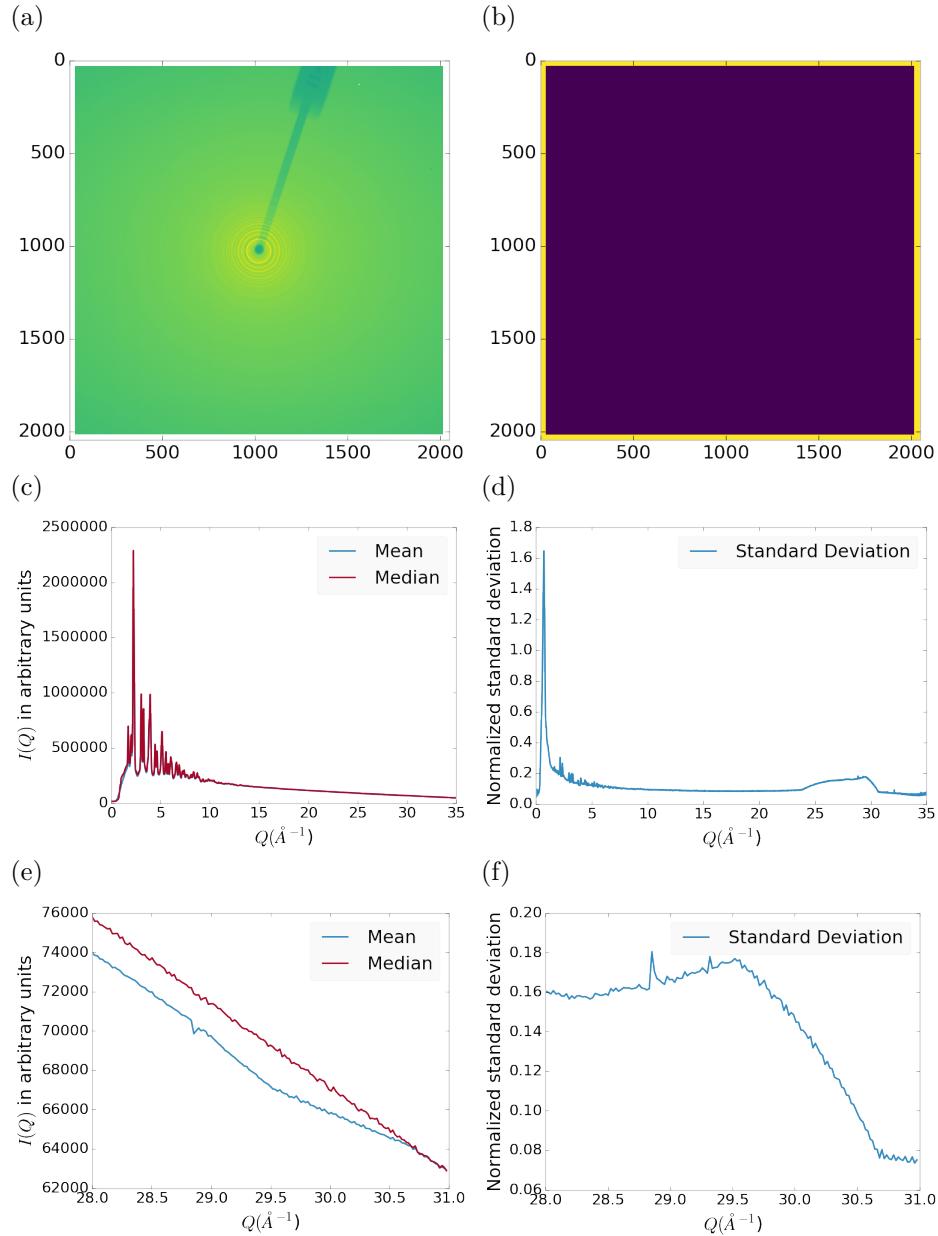


Figure 4.18: Masking, average, and standard deviation of an example x-ray total scattering measurement. This image was produced with only an edge mask. a) the image, b) the mask, c) the mean and median values, d) the standard deviation (normalized to the median), e) a closeup of the  $28 \text{ \AA}^{-1}$  to  $31 \text{ \AA}^{-1}$   $Q$  range for the mean and median, f)  $28 \text{ \AA}^{-1}$  to  $31 \text{ \AA}^{-1}$   $Q$  range for the standard deviation

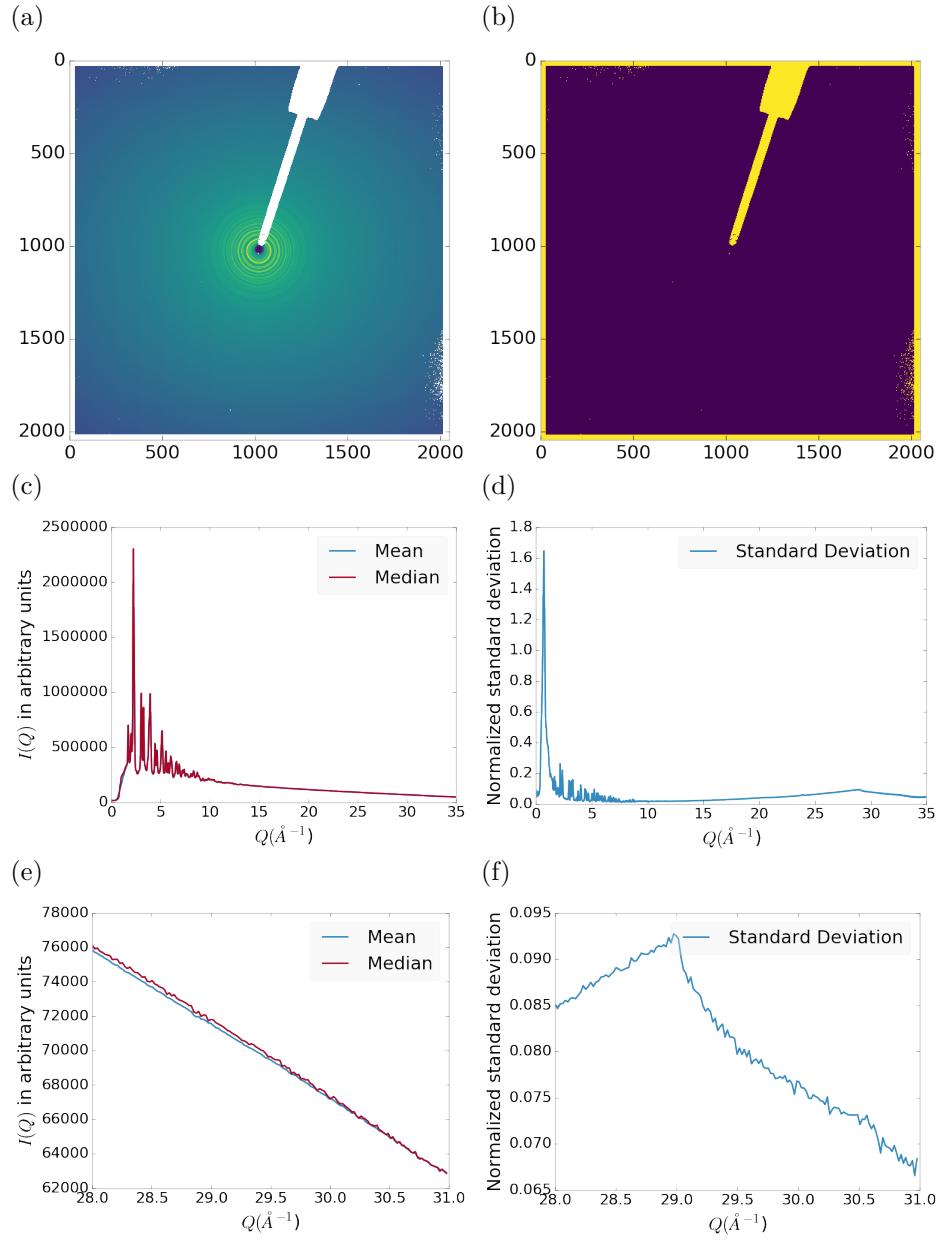


Figure 4.19: Masking, average, and standard deviation of an example x-ray total scattering measurement. This image was produced combining an edge mask and the automatically generated mask. a) the image, b) the mask, c) the mean and median values, d) the standard deviation (normalized to the median), e) a closeup of the  $28 \text{ \AA}^{-1}$  to  $31 \text{ \AA}^{-1}$   $Q$  range for the mean and median, f)  $28 \text{ \AA}^{-1}$  to  $31 \text{ \AA}^{-1}$   $Q$  range for the standard deviation

650

## CHAPTER 5

651

### ANNEALING AND AGGREGATION OF 2NM

652

#### AU NANOPARTICLES

653 5.1 EXPERIMENTS

654 NP Synthesis

655 X-ray Total Scattering Measurements

656 5.2 DATA PROCESSING

657 5.3 DATA ANALYSIS

658 5.4 SIMULATION

659 5.5 STRUCTURAL ANALYSIS

660 5.6 CONCLUSIONS

661

## CHAPTER 6

662

### PHASE CHANGES AND ANNEALING DYNAMICS OF

663

### $\text{Pr}_2\text{NiO}_4$ AND ITS DERIVATIVES

664 6.1 EXPERIMENTS

665  **$\text{Pr}_2\text{NiO}_4$  Synthesis**

666 **X-ray Measurements**

667 X-ray total scattering and x-ray powder diffraction experiments were performed at  
668 the APS's 11-ID-B beamline. An x-ray energy of XXX keV, YYY Å was used. The  
669 detector was moved between a 20cm and a 95 cm sample to detector distance to mea-  
670 sure the x-ray total scattering and x-ray diffraction patterns. Various PNO samples  
671 were annealed on the beamline during x-ray measurement.

672 6.2 DATA PROCESSING

673

masking parameters

674

integration parameters

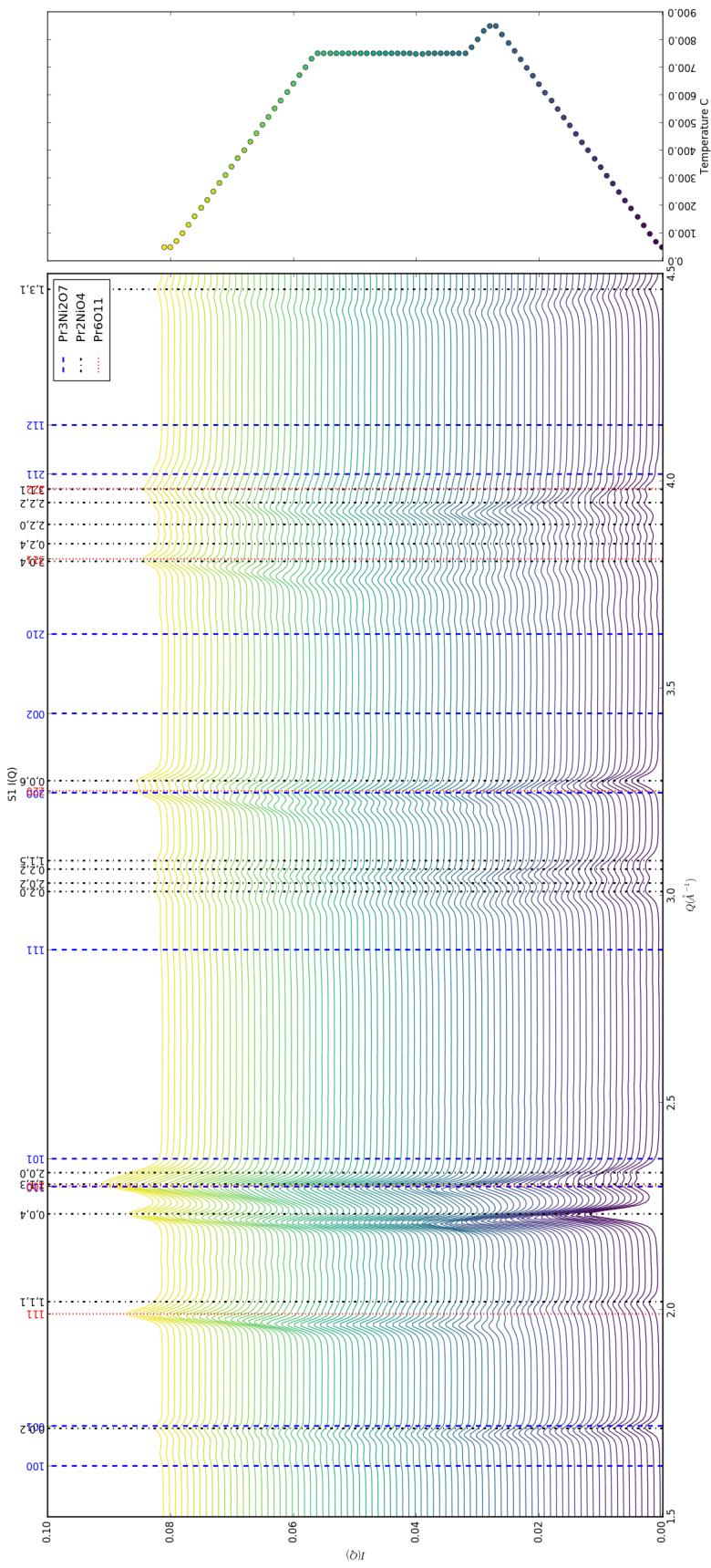
675

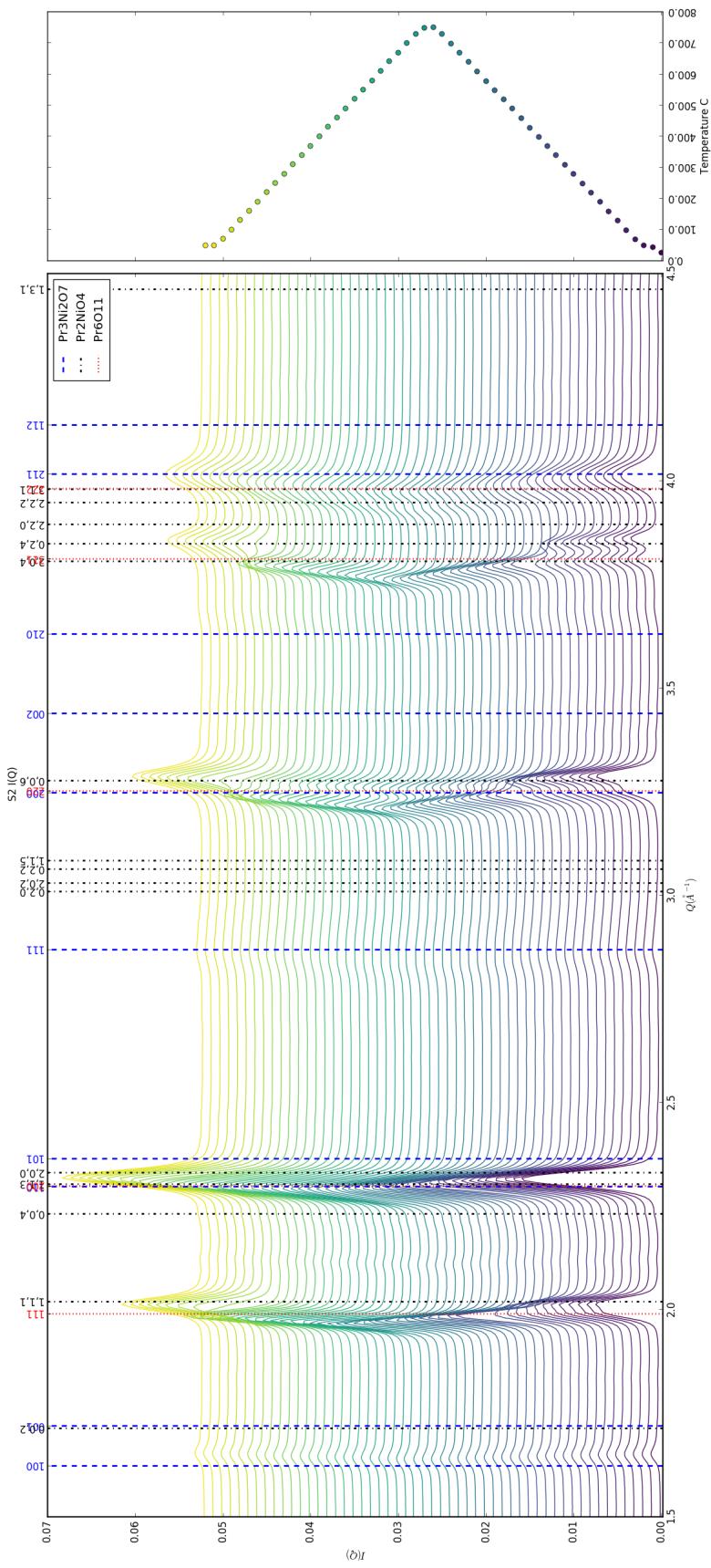
PDF parameters

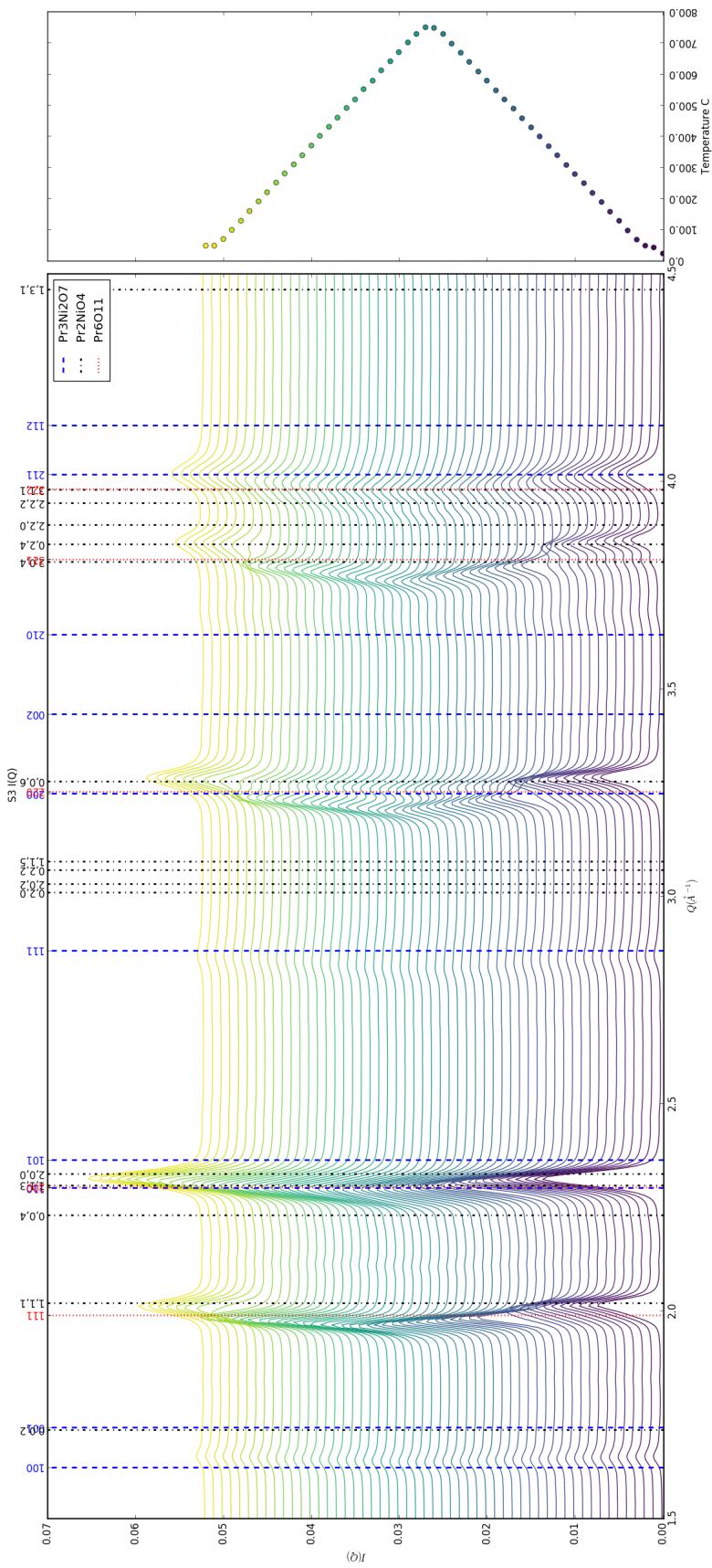
676 6.3 DATA ANALYSIS

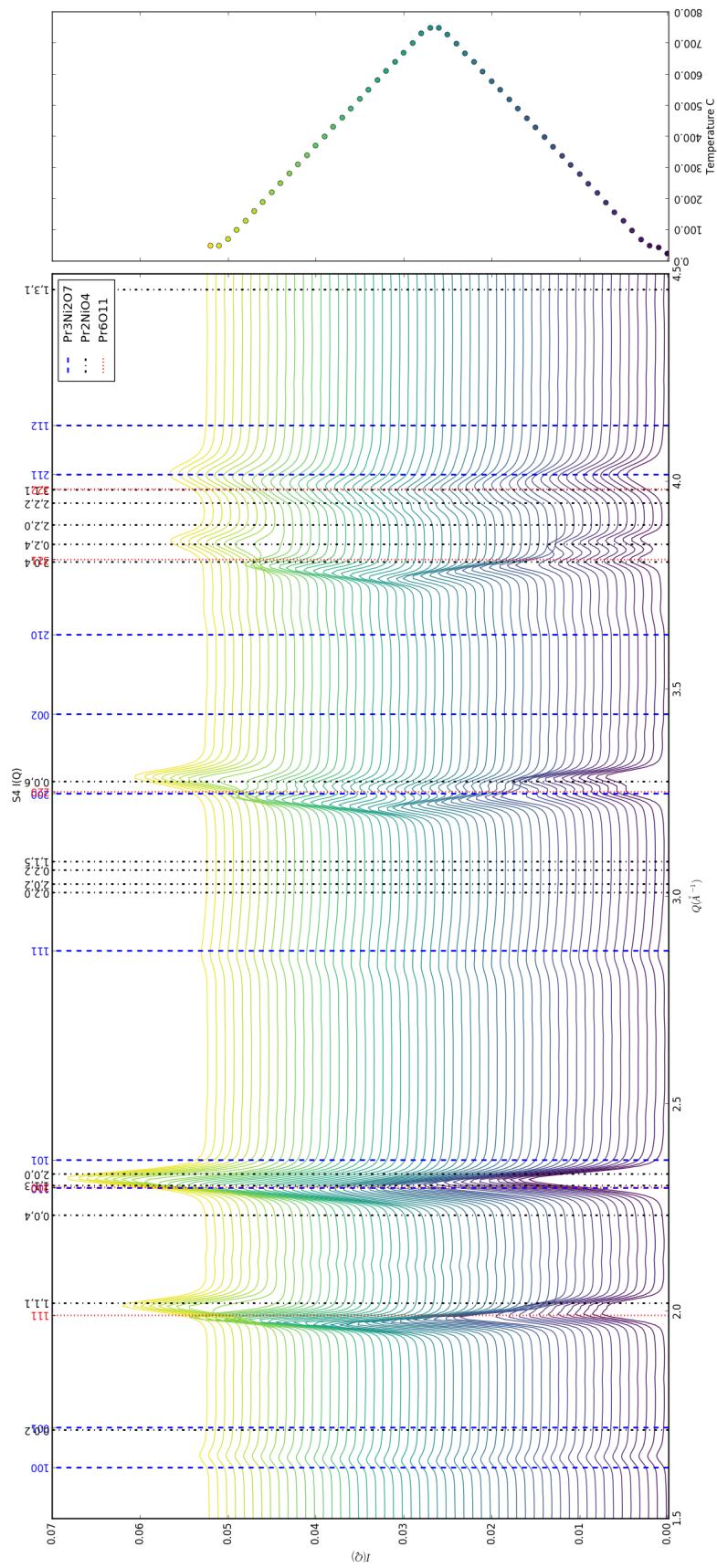
677 **Intra Sample Comparison**

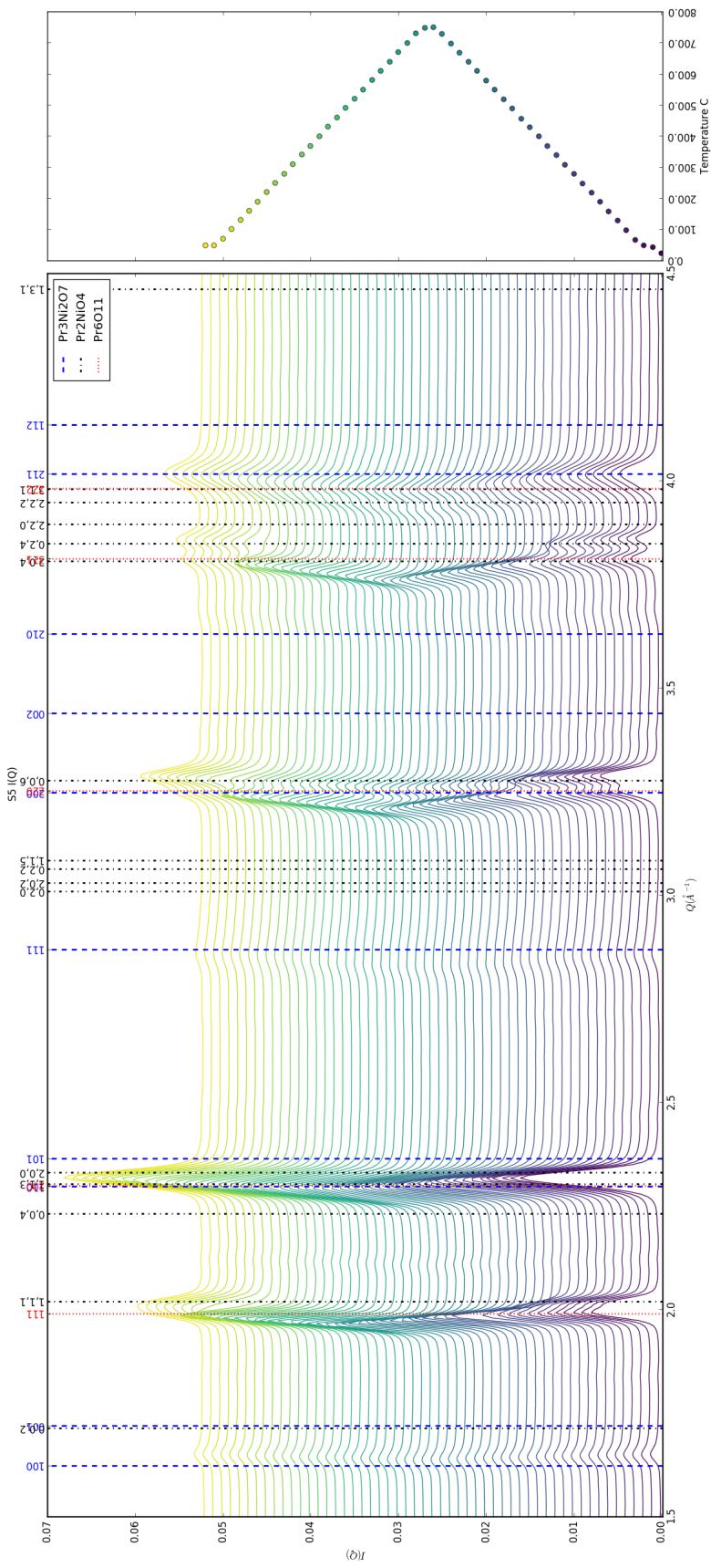
678 Changes in S1 but very little in S2-5.



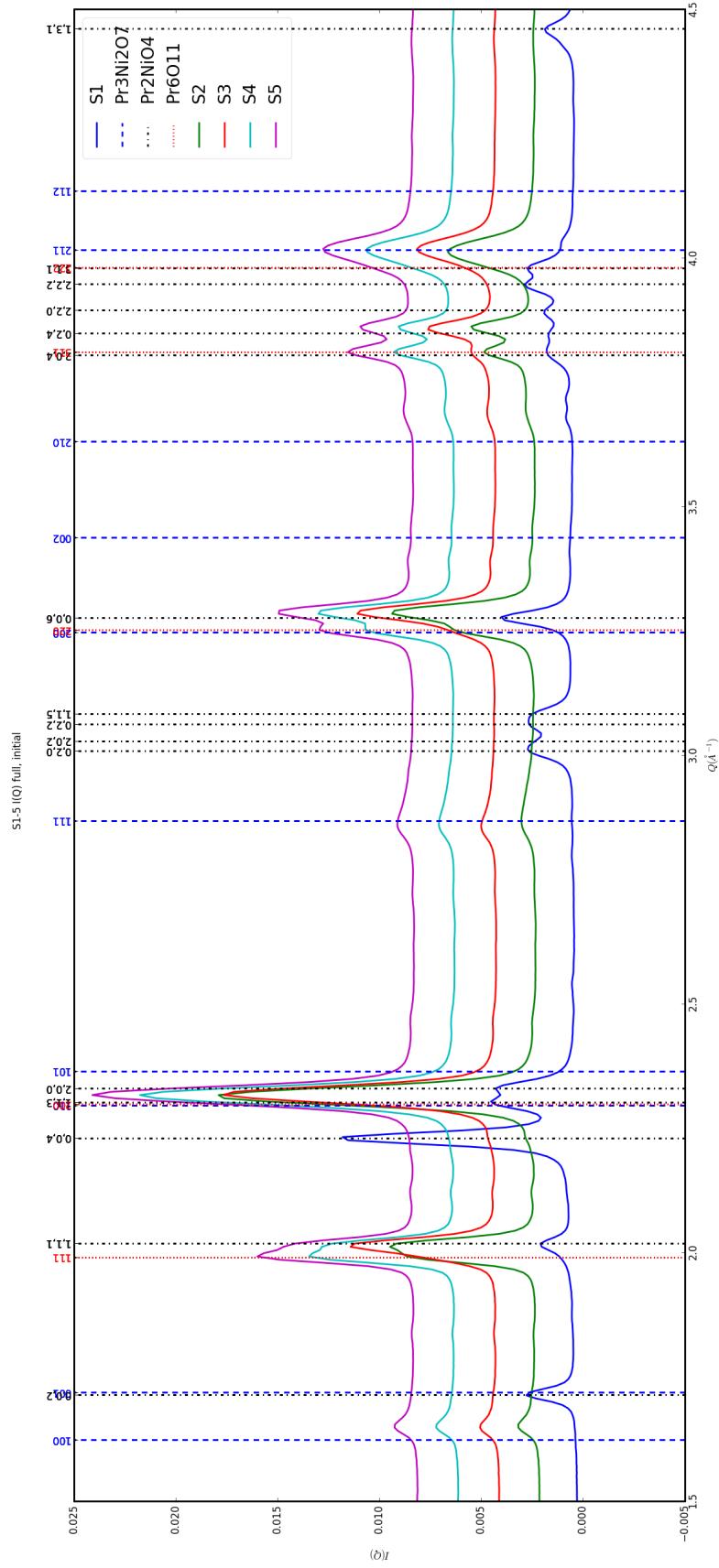


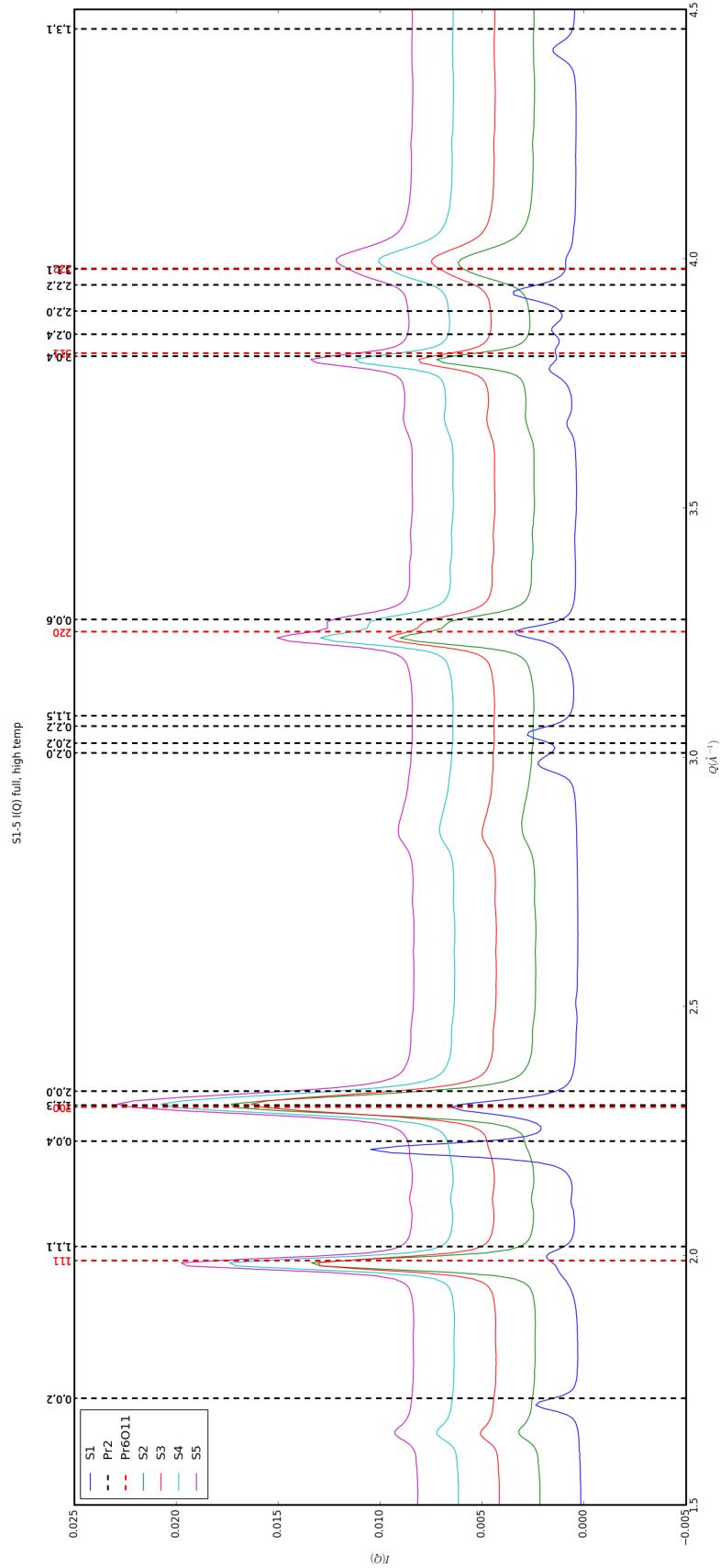


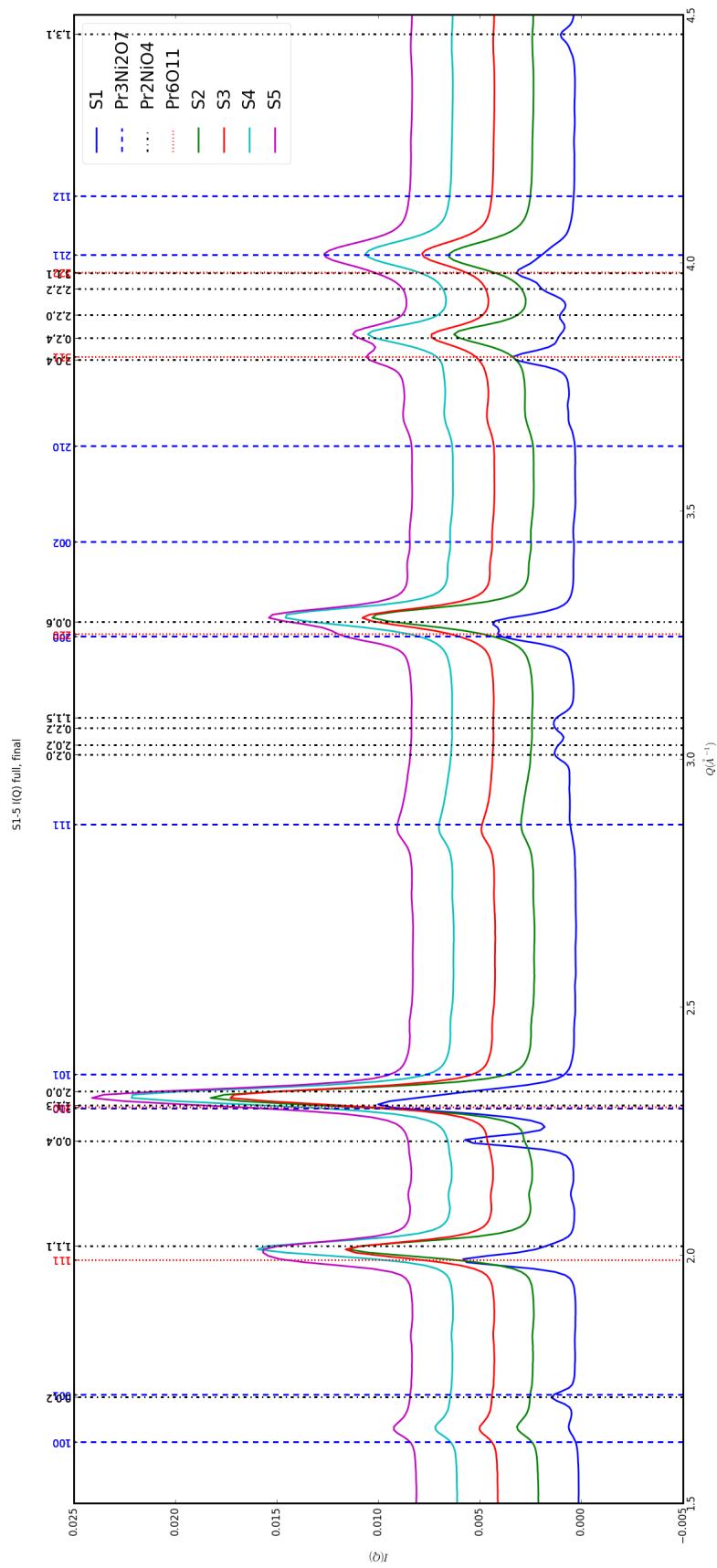


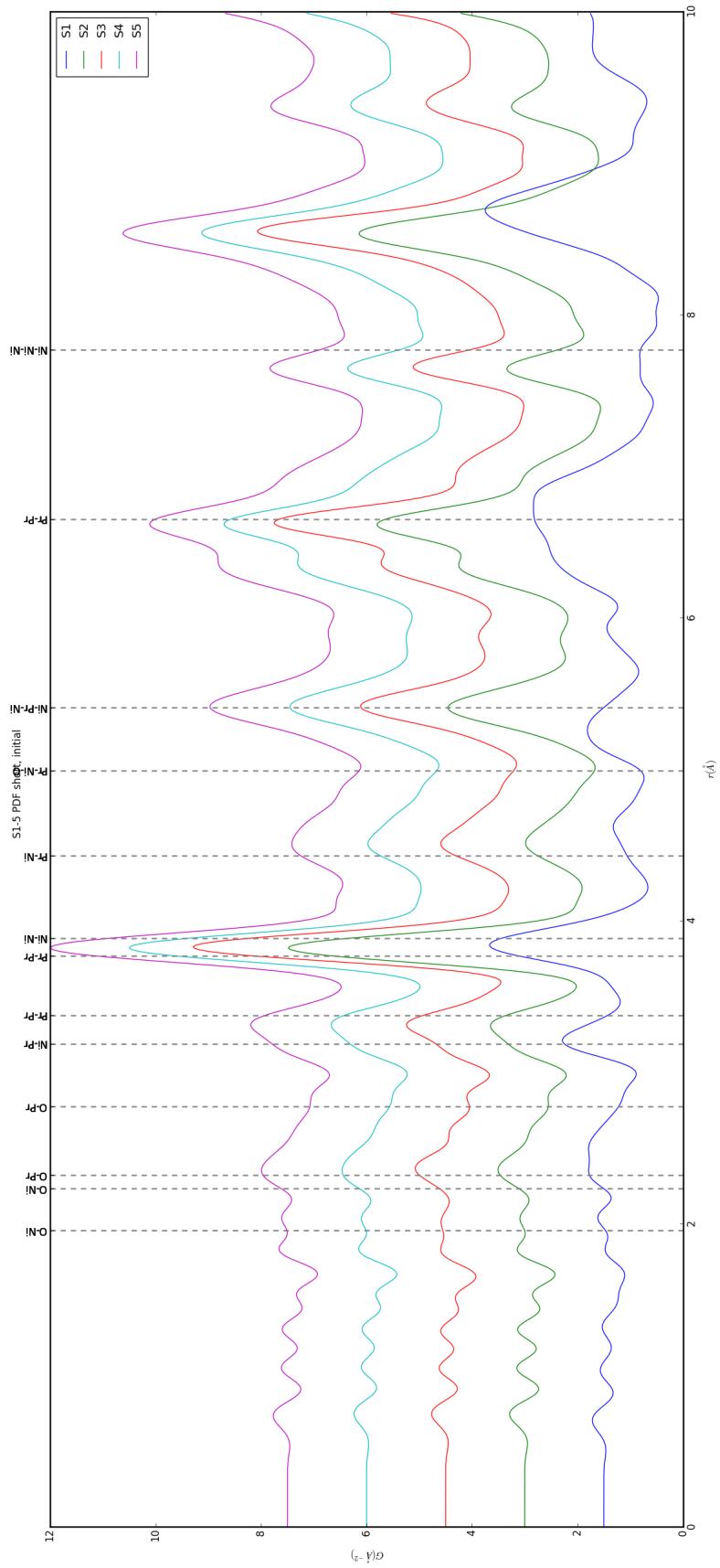


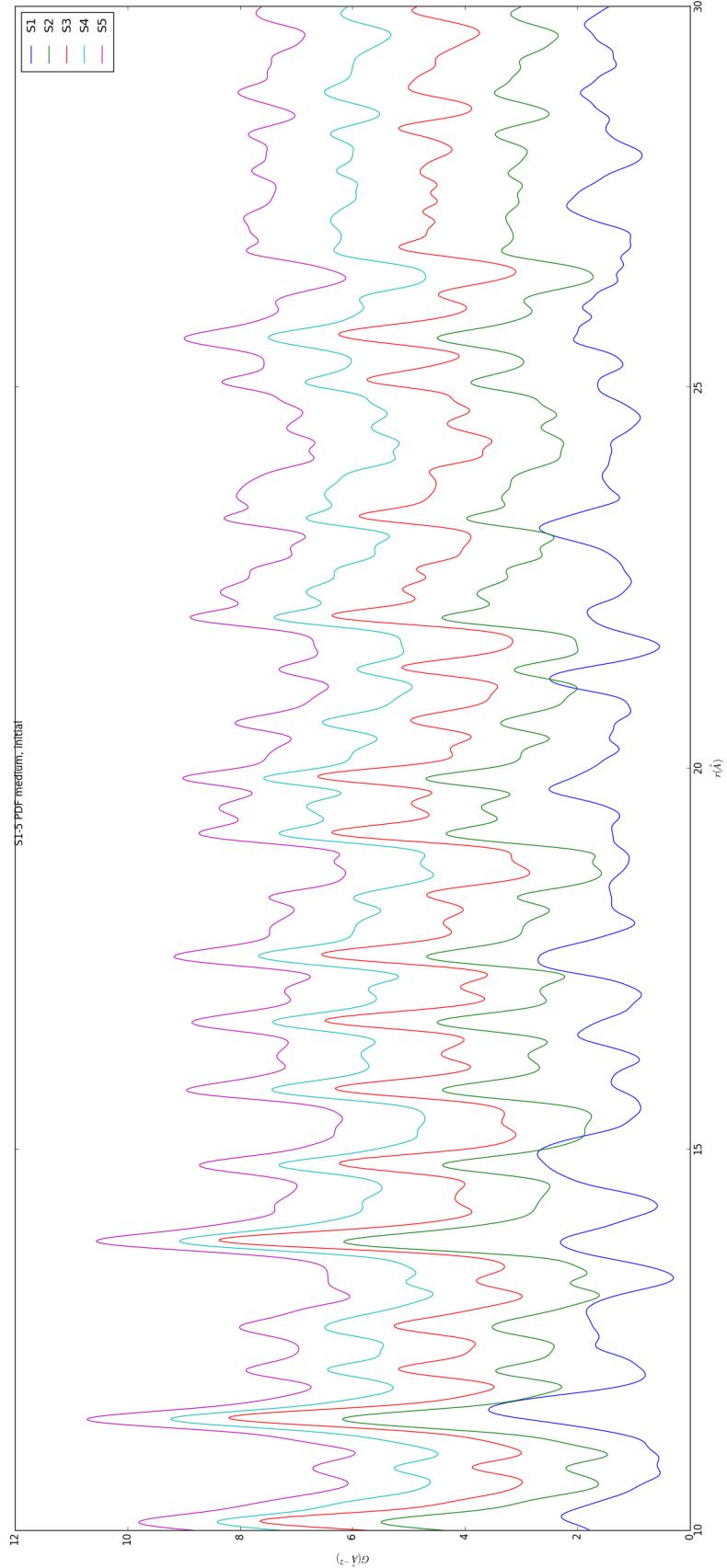
679 **Inter Sample Comparison**

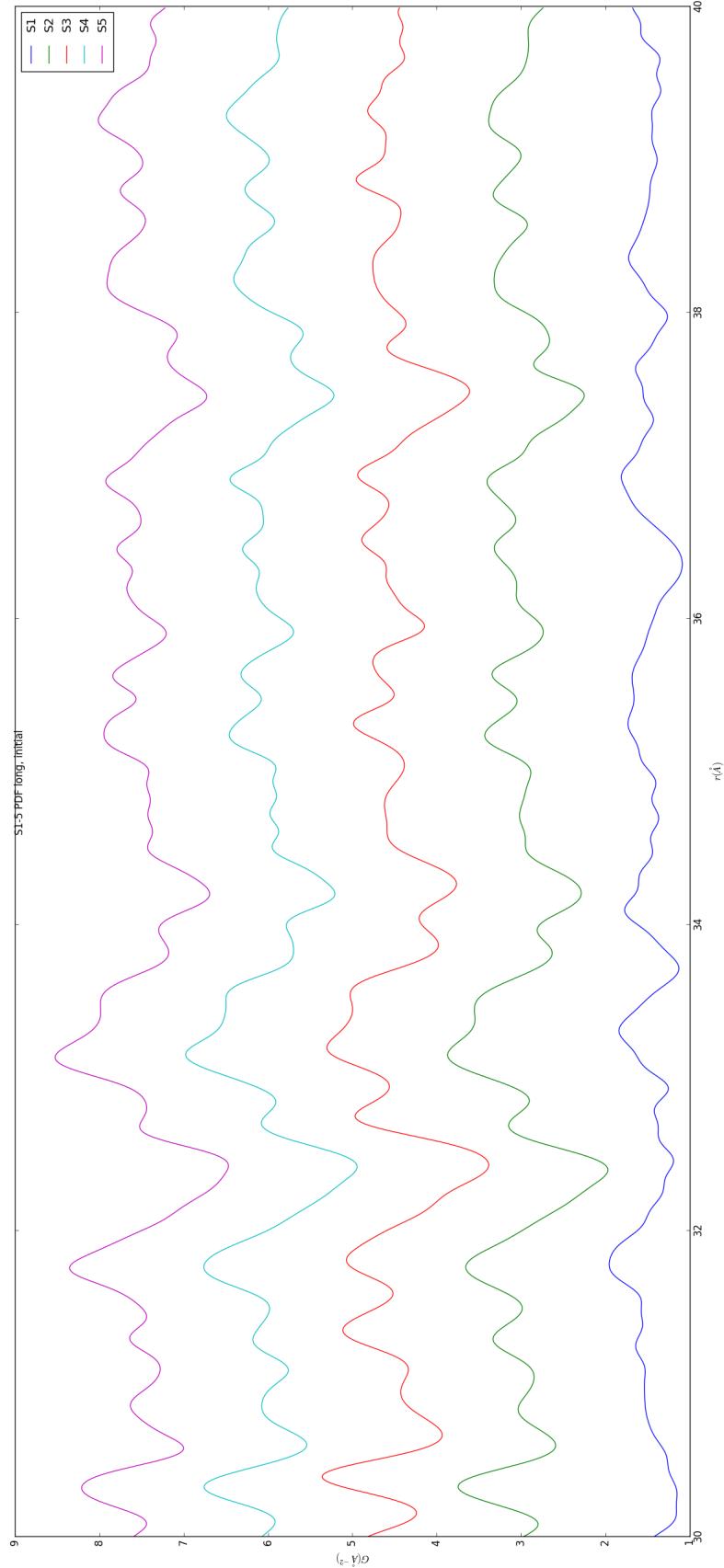


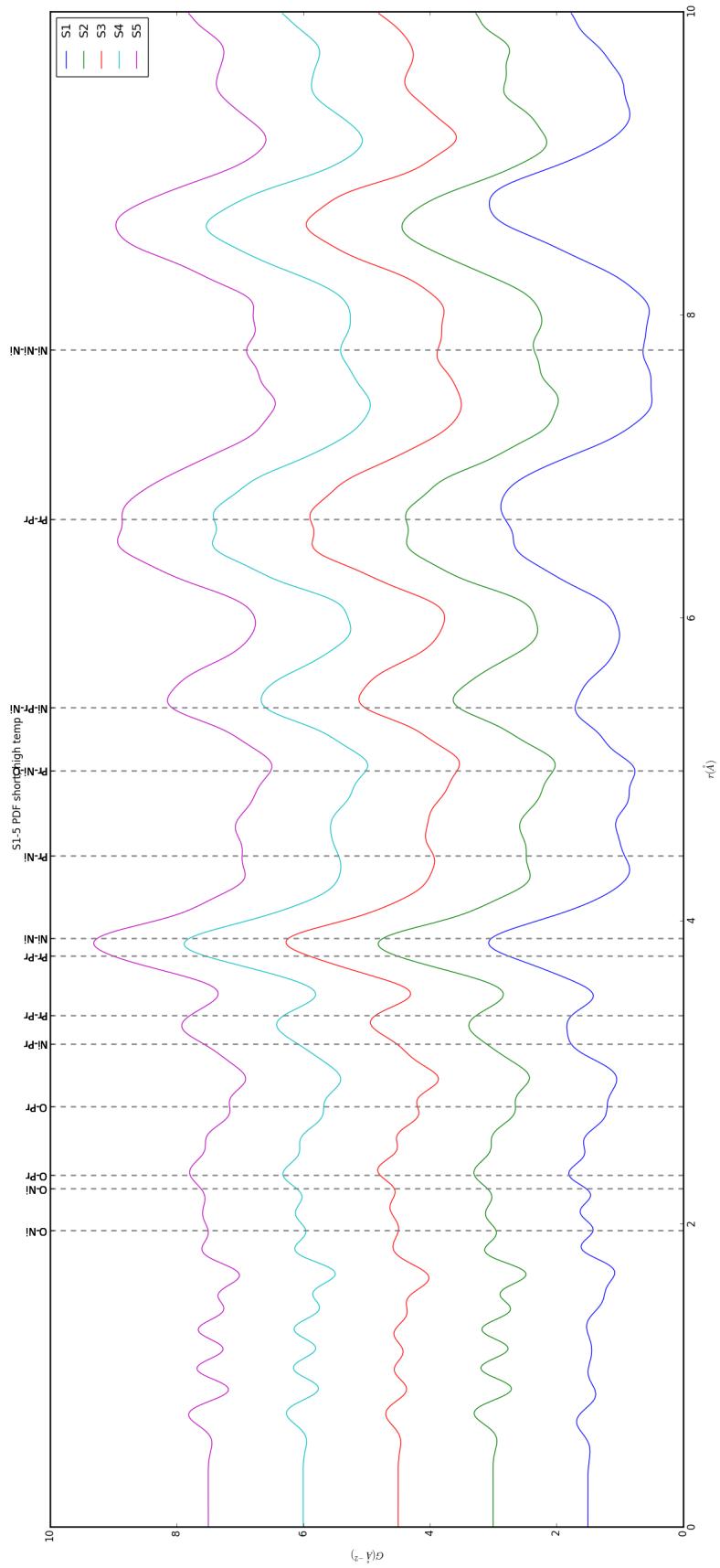


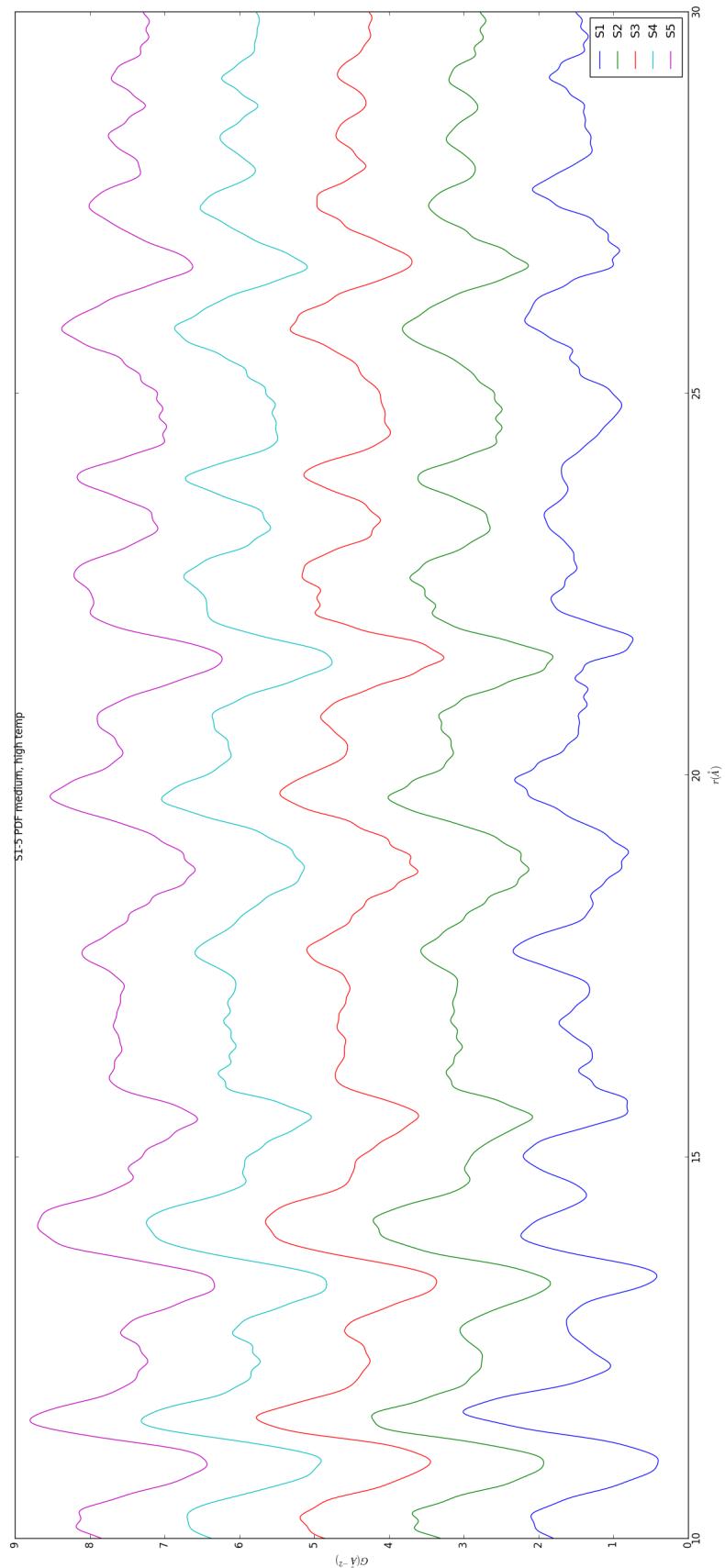


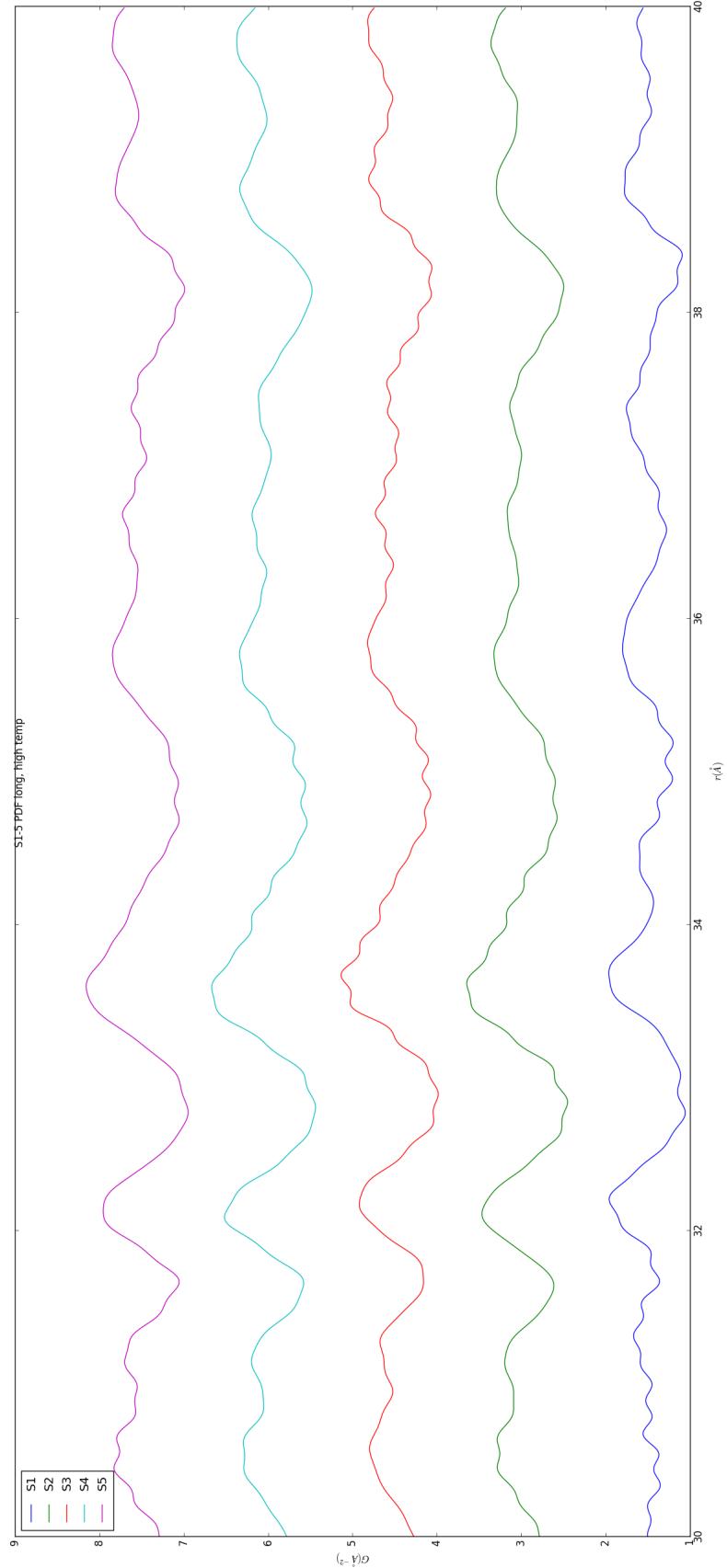


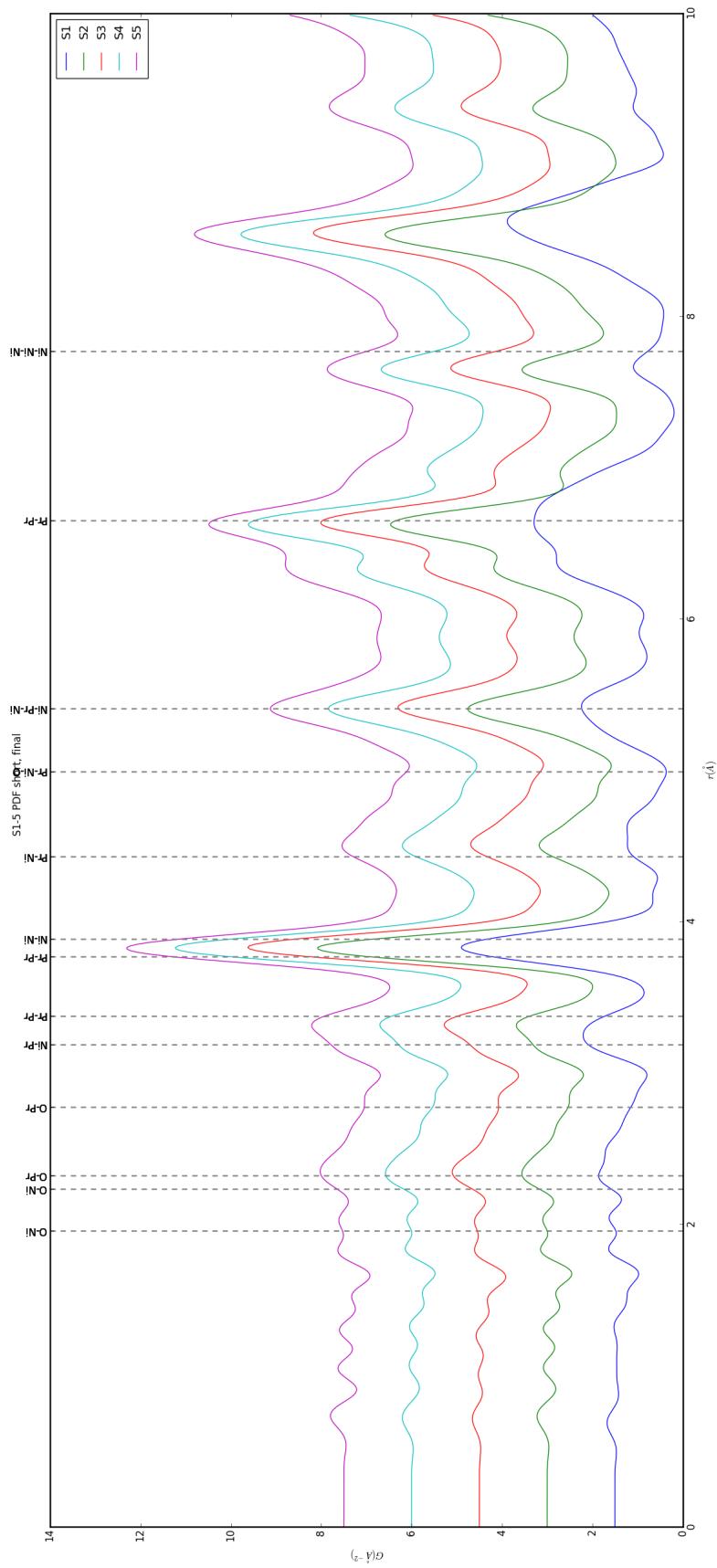


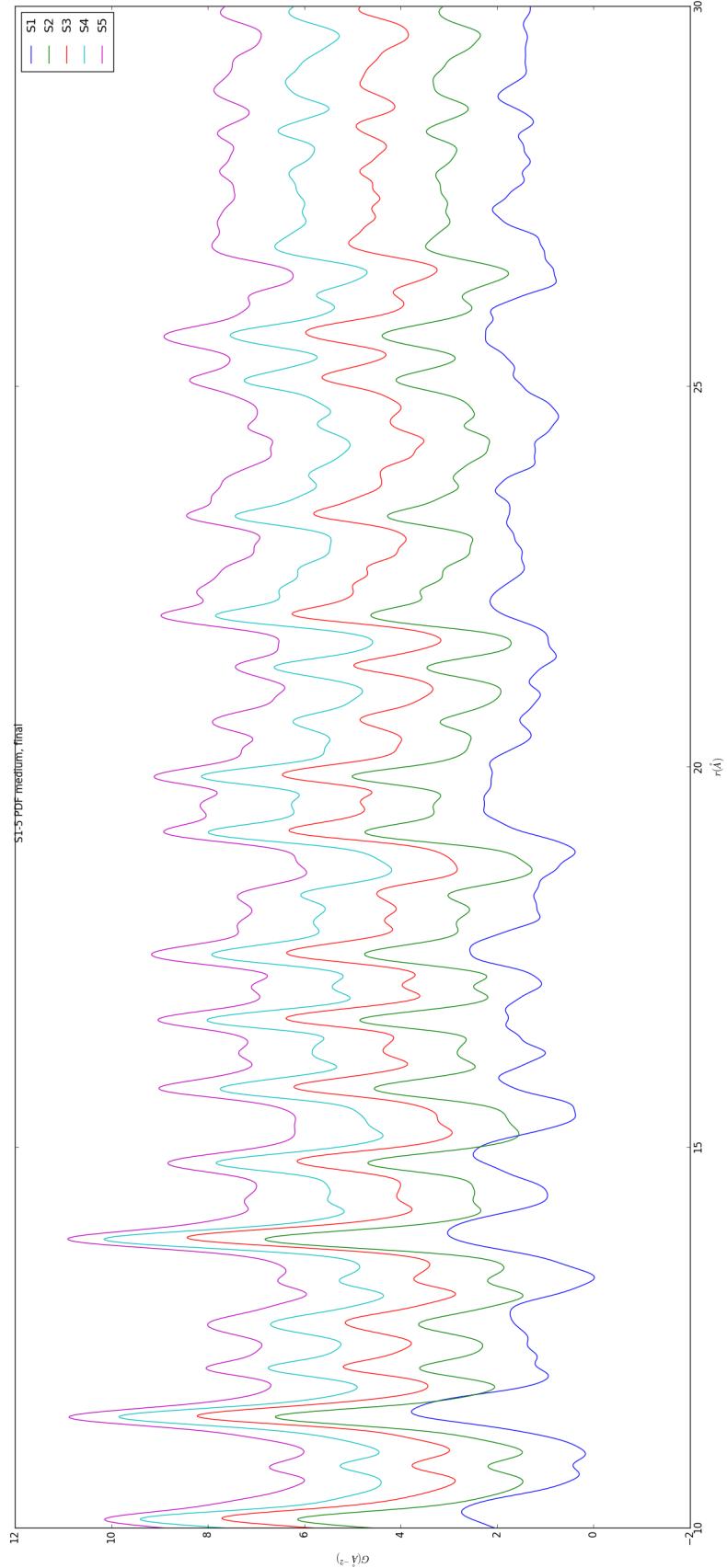


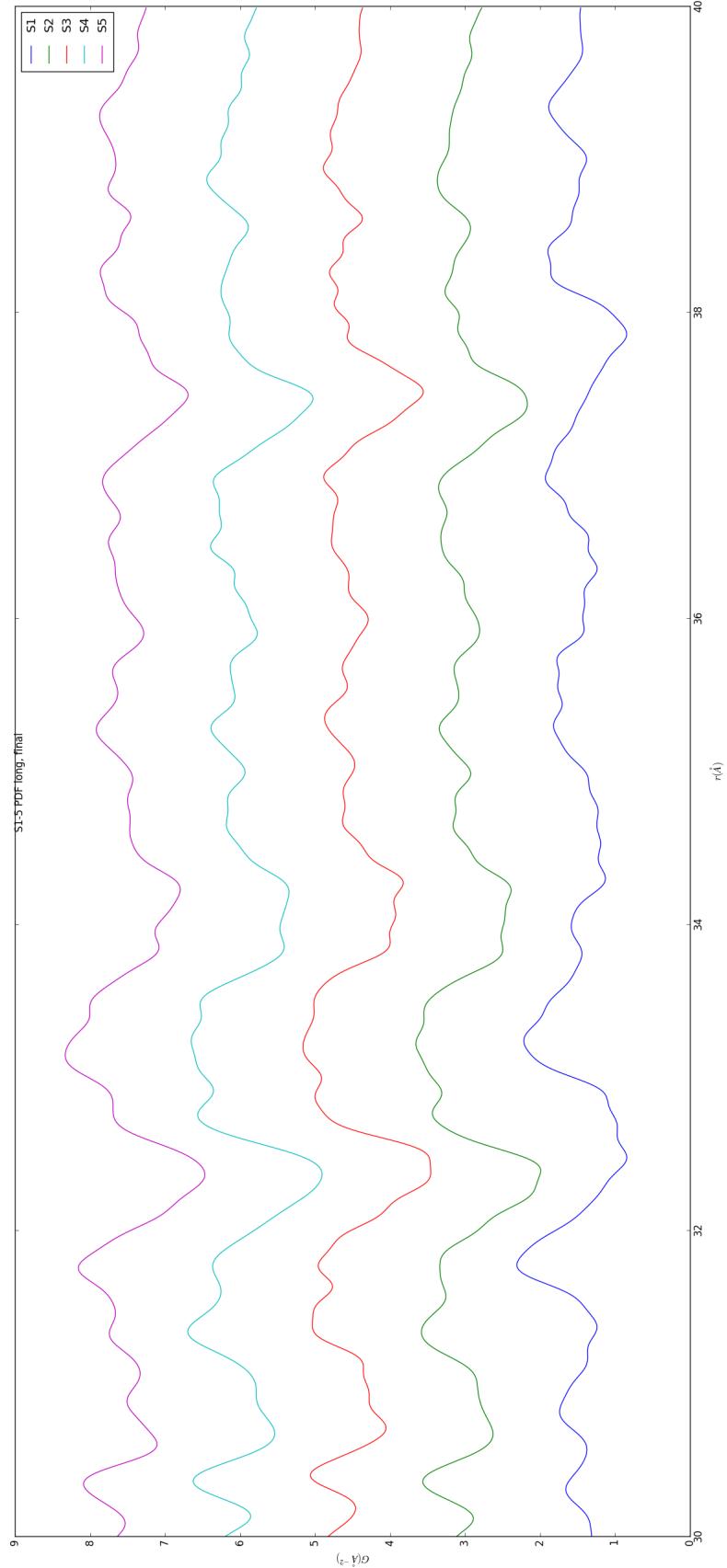












680        The PDF for S1 at operating temperature looks like S2-5 at the same temperature.

681    6.4    SIMULATION

682    Simulations have not been run yet on these PNO samples. Solving the structures of  
683    these samples is expected to be more difficult than the NP benchmarks previously  
684    solved. The difficulty of these simulations is due to:

- 685        1. The PDF's insensitivity to the oxygen positions, due to the poor x-ray scattering  
686              off the very electorn poor oxygens.
- 687        2. The large difference in mass between the oxygen and other atoms, causing the  
688              dynamics of the simulation to be governed by oxygen motion, nessecitating long  
689              simulation times to obtain movement of the other atoms.
- 690        3. The large parameter space caused by potential defects and degradation prod-  
691              ucts. Without knowing that the starting phase is pure, it is difficult to even  
692              produce starting structures, since the simulation will need to explore all the  
693              potential defect/degenerated structures.

694    6.5    CONCLUSIONS

695    X-ray total scattering and x-ray powder diffraction data was obtained on  $\text{Pr}_2\text{NiO}_4$   
696    powder samples annealed for various lengths of time. In-situ studies on the beamline  
697    were performed to understand how the structure of each of these powders changes  
698    at operating temperatures. The data was processed with the previously discussed  $Q$   
699    binning, masking, and integration methodology. The PDF results show very little  
700    change in the structure for the as synthesized sample. However, the PDFs show a  
701    large change in the previously annealed samples. These changes seem to reporduce  
702    PDFs similar to the as-synthesized PNO at operating temperatures. This would seem  
703    to imply that the source of the anamolus PNO phase/power density relationship may

704 be due to the adoption of an active structure upon heating which is universal despite  
705 the amount of thermal degradation observed at room temperature. In contrast to the  
706 PDF results, the XRD results seem to show signifigant changes in the PNO structure,  
707 both with ex-situ and in-situ annealing. The XRDs show the degradation of the PNO  
708 into various phases, potentially including  $\text{Pr}_2\text{O}_{11}$ , and higher ordered Pr based phases.  
709 The discrepancy between these two results is quite interesting as it seems that the  
710 XRD and PDF results are contradictory. Turbostratic diplacements between the  
711 layers may be a cause of the PDF/XRD disagreement, as these changes would cause  
712 very little change in the local structure observed in the PDF, while causing large  
713 changes in the XRD.

714

## CHAPTER 7

715

## CONCLUSION

716

## BIBLIOGRAPHY

- 717 [1] Simon J L Billinge and Takeshi Egami, *Underneath the Bragg Peaks: Structural  
718 Analysis of complex Materials*, vol. Volume 16, Pergamon, 2012.
- 719 [2] Peter J. Chupas, Karena W. Chapman, Charles Kurtz, Jonathan C. Hanson,  
720 Peter L. Lee, and Clare P. Grey, *A versatile sample-environment cell for non-  
721 ambient X-ray scattering experiments*, Journal of Applied Crystallography **41**  
722 (2008), no. 4, 822–824.
- 723 [3] PJ Chupas, X Qiu, JC Hanson, PL Lee, CP Grey, and SJL Billinge, *Rapid-  
724 acquisition pair distribution function (RA-PDF) analysis*, Journal of Applied  
725 Crystallography **36** (2003), 1342–1347.
- 726 [4] Timur Dykhne, Ryan Taylor, Alastair Florence, and Simon J L Billinge, *Data  
727 requirements for the reliable use of atomic pair distribution functions in amor-  
728 phous pharmaceutical fingerprinting.*, Pharmaceutical research **28** (2011), no. 5,  
729 1041–8.
- 730 [5] P. Juhás, T. Davis, C. Farrow, and S. Billinge, *PDFgetX3 : a  
731 rapid and highly automatable program for processing powder diffraction data into  
732 total scattering pair distribution functions*, Journal of Applied Crystallography  
733 **46** (2013), no. 2, 560–566 (en).
- 734 [6] Jérôme Kieffer and Dimitrios Karkoulis, *PyFAI, a versatile library for azimuthal  
735 regrouping*, Journal of Physics: Conference Series **425** (2013), 202012.
- 736 [7] B R Pauw, *Corrigendum: Everything SAXS: small-angle scattering pattern col-  
737 lection and correction (2013 J. Phys.: Condens. Matter 25 383201)*, Journal of  
738 Physics: Condensed Matter **26** (2014), no. 23, 239501.
- 739 [8] Valeri Petkov, Shiyao Shan, Peter Chupas, Jun Yin, Lefu Yang, Jin Luo, and  
740 Chuan-Jian Zhong, *Noble-transition metal nanoparticle breathing in a reactive  
741 gas atmosphere.*, Nanoscale **5** (2013), no. 16, 7379–87.
- 742 [9] Erin L. Redmond, Brian P. Setzler, Pavol Juhas, Simon J. L. Billinge, and  
743 Thomas F. Fuller, *In-Situ Monitoring of Particle Growth at PEMFC Cathode*

744        *under Accelerated Cycling Conditions*, Electrochemical and Solid-State Letters  
745        **15** (2012), no. 5, B72 (en).

746 [10] X. Yang, P. Juhás, and S. J L Billinge, *On the estimation of statistical uncertain-  
747        ties on powder diffraction and small-angle scattering data from two-dimensional  
748        X-ray detectors*, Journal of Applied Crystallography **47** (2014), no. 4, 1273–1283.