

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	x
LIST OF FIGURES	xi
TODO LIST	1
CHAPTER 1 STATISTICAL MECHANICAL ENSEMBLES AND POTENTIAL ENERGY SURFACES	5
1.1 Introduction	5
1.2 Potential Energy Surfaces	5
Experimentally Derived Potential Energy Surfaces	6
Potentials	6
Forces	7
1.3 Ensembles	8
Monte Carlo Modeling	9
Hamiltonian Monte Carlo	9
No-U-Turn-Sampling	10
Grand Canonical Ensemble	10
Ensemble description	10
Grand Canonical Monte Carlo	11
GCMC biasing	12

CHAPTER 2 ATOMIC PAIR DISTRIBUTION FUNCTION: THEORY AND COMPUTATION	14
2.1 Theory	14
Derivation	14
Analytical Gradients	14
Without ADPs	16
Periodic Boundary Conditions	16
2.2 Computation	17
HPC and GPUs	17
GPUs and Parallelization	17
Map from ij space to k space	18
GPU Memory Allocation	19
CHAPTER 3 BENCHMARKING	22
3.1 PDF	22
Model Parameters	24
Au55: surface relaxed	25
Run Parameters	25
Au55: surface disordered	27
Au55: amorphous	28
Au102: triple phase	29
Starting from fcc structure	29
Marks decahedron	30
Au147	31
3.2 PDF with ADPs	31
ADP 50	31
CHAPTER 4 X-RAY TOTAL SCATTERING DATA ACQUISITION AND PRO- CESSING	37
4.1 Introduction	37
4.2 Detector Q resolution	37
4.3 Automated Mask Generation	39
Introduction	39
Algorithm Design	41
Test Cases	41

Results and Discussion	42
Conclusions	48
4.4 Automated Image Azimuthal Integration	49
4.5 Conclusions	50
 CHAPTER 5 ANNEALING AND AGGREGATION OF 2NM AU NANOPARTICLES	54
5.1 Experiments	54
NP Synthesis	54
X-ray Total Scattering Measurements	54
5.2 Data Processing	54
5.3 Data Analysis	54
5.4 Simulation	54
5.5 Structural Analysis	54
5.6 Conclusions	54
 CHAPTER 6 PHASE CHANGES AND ANNEALING DYNAMICS OF Pr_2NiO_4 AND ITS DERIVATIVES	55
6.1 Experiments	55
Pr_2NiO_4 Synthesis	55
X-ray Measurements	55
6.2 Data Processing	55
6.3 Data Analysis	55
Intra Sample Comparison	55
Inter Sample Comparison	61
6.4 Simulation	74
6.5 Conclusions	74
 CHAPTER 7 CONCLUSION	76

BIBLIOGRAPHY	77
------------------------	----

LIST OF TABLES

LIST OF FIGURES

Figure 2.1 Comparison of the CPU and GPU chip architectures	18
Figure 3.1 Au ₅₅ PDF fitting of DFT-optimized <i>c</i> -Au ₅₅ . (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.	26
Figure 3.2 Au ₅₅ PDF fitting of surface-disordered Au ₅₅ . 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.	32
Figure 3.3 Similar to figure 3.2 for DFT-optimized amorphous Au ₅₅	33
Figure 3.4 Similar to figure 3.2 for Au ₁₀₂ as in DFT-optimized Au ₁₀₂ MBA ₄₄ cluster.	34
Figure 3.5 Similar to Fig. 3.6 with Marks decahedron as the starting structure.	35
Figure 3.6	36
Figure 4.1 Scattering onto a flat detector	38
Figure 4.2 Q resolution as a function of Q	39
Figure 4.3 Number of pixels as a function of Q , binned at the Q resolution of the detector.	40
Figure 4.4 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.	42

Figure 4.5 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.	43
Figure 4.6 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.	43
Figure 4.7 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.	44
Figure 4.8 Generated beamstop holder masks for a beamstop holder with 10% transmittance. a) the raw image, b) the masked image, c) and the missed pixels	44
Figure 4.9 Generated beamstop holder masks for a beamstop holder with 30% transmittance. a) the raw image, b) the masked image, c) and the missed pixels	45
Figure 4.10 Generated beamstop holder masks for a beamstop holder with 50% transmittance. a) the raw image, b) the masked image, c) and the missed pixels	45
Figure 4.11 Generated beamstop holder masks for a beamstop holder with 90% transmittance. a) the raw image, b) the masked image, c) and the missed pixels	45
Figure 4.12 Generated beamstop holder masks which is rotated away from verticle	46
Figure 4.13 Masked experimental data. a) the raw image, b) the mask	46
Figure 4.14 Masked experimental data with Pt single crystal signal. a) the raw image, b) the mask	47
Figure 4.15 Masked experimental data with Pt single crystal signal using figure's 4.13 as a starting mask. a) the raw image, b) the mask	47

Figure 4.16 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 \AA^{-1} to $31 \text{ \AA}^{-1}Q$ range for the mean and median, f) 28 \AA^{-1} to $31 \text{ \AA}^{-1}Q$ range for the standard deviation 51

Figure 4.17 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 \AA^{-1} to $31 \text{ \AA}^{-1}Q$ range for the mean and median, f) 28 \AA^{-1} to $31 \text{ \AA}^{-1}Q$ range for the standard deviation 52

Figure 4.18 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 \AA^{-1} to $31 \text{ \AA}^{-1}Q$ range for the mean and median, f) 28 \AA^{-1} to $31 \text{ \AA}^{-1}Q$ range for the standard deviation 53

1

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

TODO LIST

3	Fix the figures so that they don't give the full caption	1
4	Why is atomistic engineering important	3
5	Barriers to atomistic engineering	3
6	How are we going to attack this problem	3
7	This entire section needs some rewritting to distinguish this from the paper .	22
8	Put this somewhere	26
9	masking parameters	55
10	integration parameters	55
11	PDF parameters	55

12

INTRODUCTION

13

Why is atomistic engineering important

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

18

Barriers to atomistic engineering

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

22

How are we going to attack this problem

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

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

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

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

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

37 Chapter 4 will focus on the aquesition of experimental data, its management, and
38 processing. §4.2, 4.3, and 4.4 will discuss the derivation of the Q resolution function,
39 the automated masking of 2D area detectors for x-ray total scattering measurements
40 using the previously derrived Q resolution, and the impact of different averaging
41 methods and masks on azimuthal integration, respecitvly.

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 one
47 of optimization. The plan is to develop a potential energy surface (PES) which has
48 minima associated with atomic structures who's properties match the experimentally
49 observed properties. Thus, the various positional variables of the structure can be
50 solved by optimizing the structure against the PES. This approach is popular in the
51 PDF community for solving the structure of materials using both extensive large box
52 models and simpler small box models.

53 In this chapter we discuss the development of the various PESs used in the PDF
54 community for comparing theoretical and experimental PDFs. Special attention will
55 be paid to the gradients of the potential energy functions, as these are important
56 to some optimization techniques. Additionally, we also discuss the use of statistical
57 mechanical ensembles for finding minima on the PES.

58 1.2 POTENTIAL ENERGY SURFACES

59 A PES simply describes the potential energy of the system as a function of all its
60 relevant coordinates in phase space, essentially providing a mapping $\mathbb{R}^n \rightarrow \mathbb{R}$. Usually
61 these coordinates are the positions of the atoms q and their conjugate the momenta p .
62 Note that there could be more variables associated with the system, for instance the

63 magnetic moments of the atoms could play a role in describing the system. In this
 64 magnetic system there would be positional variables for the atomwise spin vectors
 65 and their "momenta". Application of the term "momenta" might seem odd here, as
 66 the magnetic spin does not have a mass or a velocity. However, since the magnetic
 67 "position" is defined on the PES we need to describe its conjugate variable to properly
 68 formulate Hamiltonian dynamics and the kinetic portion of the PES.

69 **Experimentally Derived Potential Energy Surfaces**

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

78 **Potentials**

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

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

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

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

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

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

87 The Rw and χ^2 potentials have been reported numerous times. [?] However, the
 88 INVERT potential is fairly new and aims to incorporate descriptions of the structural
 89 symmetry into the PES. [?, ?] In the case of the INVERT potential NMR or other
 90 symmetry sensitive data is used to describe the number of unique atomic coordina-
 91 tions. This is then used to describe the number of unique atomwise pair distribution
 92 functions, thus causing systems with more or less unique coordination environments
 93 to be higher in energy. This approach has been shown to be useful for C_{60} and other
 94 systems which are highly symmetric, creating a PES with an easier to find minima.
 95 [?, ?] However, many times this kind of data is unavailable when refining the struc-
 96 ture causing the potential to be less useful. Additionally, this potential introduces
 97 an element of user bias as the refiner must decide, based on some spectroscopic data,
 98 how many unique environments are in the material. This bias could be removed by
 99 using one of the other potentials with a method for simulating the observed spectra,
 100 allowing the computational system decide what structures properly reproduce all the
 101 observed data.

102 Forces

$$103 \quad \vec{\nabla} \chi^2 = -2 \sum_{a=a_{\min}}^{a_{\max}} (\alpha \frac{\partial A_{\text{calc}}}{\partial \gamma_{i,w}} + A_{\text{calc}} \frac{\partial \alpha}{\partial \gamma_{i,w}})(A_{\text{obs}} - \alpha A_{\text{calc}}) \quad (1.5)$$

$$104 \quad \vec{\nabla} Rw = \frac{Rw}{\chi^2} \sum_{a=a_{\min}}^{a_{\max}} (\alpha \frac{\partial A_{\text{calc}}}{\partial \gamma_{i,w}} + A_{\text{calc}} \frac{\partial \alpha}{\partial \gamma_{i,w}})(\alpha A_{\text{calc}} - (A_{\text{obs}})) \quad (1.6)$$

$$105 \quad \vec{\nabla} \chi_{\text{INVERT}}^2 = \frac{-2}{N} \sum_{a=a_{\min}}^{a_{\max}} \sum_j (\alpha \frac{\partial A_{j,\text{calc}}}{\partial \gamma_{i,w}} + A_{j,\text{calc}} \frac{\partial \alpha}{\partial \gamma_{i,w}})(A_{\text{obs}} - \alpha A_{j,\text{calc}}) \quad (1.7)$$

$$\frac{\partial \alpha}{\partial \gamma_{i,w}} = \frac{(\sum_{a=a_{\min}}^{a_{\max}} A_{\text{obs}} \frac{\partial A_{\text{calc}}}{\partial \gamma_{i,w}} - 2\alpha \sum_{a=a_{\min}}^{a_{\max}} A_{\text{calc}} \frac{\partial A_{\text{calc}}}{\partial \gamma_{i,w}})}{\sum_{a=a_{\min}}^{a_{\max}} A_{\text{calc}}^2} \quad (1.8)$$

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

111 ALSO COMPARE RW AND CHI**2, POTENTIALY WITH A FIGURE.

112 1.3 ENSEMBLES

113 While PESs describe which atomic configurations are the most desirable and how
114 the atoms would like to get there, the ensemble describes how the atoms move on
115 the PES. The abstraction of the PES from the ensemble is an important one, as it
116 allows for the reuse and exchange of both PESs and ensembles for a wide array of
117 problems. Statistical mechanical ensembles can be described in two ways, analytically
118 and stochastically. For long simulation times and fine enough numerical or analytical
119 integration these two descriptions should be identical.

120 In either case one starts by defining the Hamiltonian, \mathcal{H} , as the total energy of
121 the system. Thus, the Hamiltonian is described as the sum of the potential $U(q)$ and
122 kinetic $K(p)$ energies, where q is the positions of the atoms and p is their momenta

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

123 where $K(p) = \frac{1}{2} \sum_i \frac{p_i^2}{m_i}$ and i denotes the i th particle.

124 Analytically one generally defines a partition function, which describes the sum of
125 probabilities over all potential atomic states.

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

126 where P_i is the probability of the i th state and is a function of the total energy of
127 that state. This partition function can then be used to obtain the probability of any

128 specific state. The relationship of the probability of a state to the state's energy
 129 and other properties depends on the ensemble being used. For the microcanonical
 130 ensemble the probability of a state is:

$$P(q, p) = \frac{\delta(\mathcal{H}(q, p) - E)}{W} \quad (1.10)$$

131 where E is the energy of the system, W is the total number of states in the system,
 132 and δ is the Dirac Delta Function.

133 However, for the canonical ensemble the probability is:

$$P(q, p) = \exp\left(\frac{E - \mathcal{H}(q, p)}{k_b T}\right) \quad (1.11)$$

134 Monte Carlo Modeling

135 Monte Carlo can be used to simulate a statistical mechanical ensemble which can
 136 not be solved analytically. In most Monte Carlo systems the ensemble is simulated by
 137 randomly changing one of the system parameters and comparing the energy of the
 138 new system against the energy of the old system. If the energy of the new system is
 139 lower than the current energy then the new configuration is accepted. Otherwise the
 140 new system is rejected unless

$$\exp\left(\frac{-\Delta E}{E_T}\right) < u$$

141 where u is a random number $[0, 1]$ and E_T is the thermal energy characteristic to the
 142 system.

143 Hamiltonian Monte Carlo

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

$$\frac{dq_i}{dt} = \frac{\partial \mathcal{H}}{\partial p_i} = p_i \quad (1.12)$$

$$\frac{dp_i}{dt} = -\frac{\partial \mathcal{H}}{\partial q_i} = -\vec{\nabla} U \quad (1.13)$$

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

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

$$q_i(t + \delta t) = q_i(t) + \delta t * p_i(t + \delta t/2) \quad (1.15)$$

$$p_i(t + \delta t) = p_i(t + \delta t/2) - \frac{\delta t}{2} \frac{\partial}{\partial q_i} U(q(t + \delta t)) \quad (1.16)$$

¹⁴⁴ Note that $\frac{\partial}{\partial q_i}$ is the gradient with respect to q where i denotes the i th atom being
¹⁴⁵ moved. Using this notation the gradient is

$$\vec{\nabla} U = \begin{bmatrix} \frac{\partial U}{\partial q_{0,x}} & \frac{\partial U}{\partial q_{0,y}} & \frac{\partial U}{\partial q_{0,z}} \\ \vdots & \frac{\partial U}{\partial q_{i,w}} & \vdots \\ \frac{\partial U}{\partial q_{n,x}} & \frac{\partial U}{\partial q_{n,y}} & \frac{\partial U}{\partial q_{n,z}} \end{bmatrix} = \begin{bmatrix} \vec{\mathcal{F}}_0 \\ \vdots \\ \vec{\mathcal{F}}_i \\ \vdots \\ \vec{\mathcal{F}}_n \end{bmatrix} \quad (1.17)$$

¹⁴⁶ where $\frac{\partial}{\partial q_{i,w}}$ is the derivative with respect to q where w denotes direction of the derivative (x , y , or z), n is the number of atoms and U is the potential which depends on q , and $\vec{\mathcal{F}}_i$ is the "force" on the i th atom.

¹⁴⁹ No-U-Turn-Sampling

¹⁵⁰ Grand Canonical Ensemble

¹⁵¹ Ensemble description

¹⁵² In the Grand Canonical Ensemble (GCE) two sets of variables are allowed to change,
¹⁵³ the atomic positions and the total number of atoms and their associated identities.
¹⁵⁴ These two variables are controlled by temperature and chemical potential. The par-
¹⁵⁵ tition function is

$$\Xi = e^{-\beta(\mathcal{H}+\mu)} \quad (1.18)$$

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

158 **Grand Canonical Monte Carlo**

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

$$\min[1, \frac{V}{(N+1)\Lambda(T)^3} e^{-\beta\Delta U + \beta\mu}] \quad (1.19)$$

162 Similarly the removal of an atom has the probability:

$$\min[1, \frac{(N)\Lambda(T)^3}{V} e^{-\beta\Delta U - \beta\mu}] \quad (1.20)$$

163 However, both of these equations depend of the overall simulation volume and the
164 thermal wavelength, which is undesirable as these are not really properties that we
165 are of interest to these simulations. Thus, we roll them into the definition of the
166 chemical potential, essentially setting the base chemical potential to counteract these
167 effects. This makes certain that our simulation does not change if we change the
168 overall cell volume. A GCMC move consists of creating a new atomic configuration,
169 where an atom has been added or removed, and checking the above criteria. However,
170 previous results have shown that this method is computationally expensive in dense
171 liquids, and exceedingly expensive in solid materials. The long simulation times
172 are due to the random nature of the atomic additions or removals which produce:
173 over-tightly packed atoms, atoms in the middle of nowhere, or unphysical vacancies.
174 These configurations are rejected by the GCMC criteria but their probability of being
175 sampled is much higher than configurations which are lower in energy, since the
176 number of incorrect ways to add/remove atoms is much larger than the correct ways.
177 Thus we have implemented methods for biasing the atomic addition positions and
178 the atomic removals toward configurations which are more likely to be accepted.

179 **GCMC biasing**

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

$$P_{i,j,k} = \frac{xyz}{abc} \quad (1.21)$$

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

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

$$P_{i,j,k} = \frac{e^{\beta\Delta U_{i,j,k}}}{\sum_{i,j,k} e^{\beta\Delta U_{i,j,k}}} \quad (1.22)$$

199 where $\Delta U_{i,j,k}$ is the change in energy. However, calculating $\Delta U_{i,j,k}$ can be particu-
200 larly expensive, especially when calculating scattering from atomic positions. The
201 computational expense can be mitigated by using a cheaper potential, if only for the

202 evaluation of the voxel energy, as previously shown. Similar to previous work we can
203 use the Lennard Jones potential to approximate the addition potential.

204

CHAPTER 2

205

ATOMIC PAIR DISTRIBUTION FUNCTION: THEORY AND COMPUTATION

207 2.1 THEORY

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

212 Derivation

213 Analytical Gradients

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

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

223 where

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

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

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

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

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

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

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

227 where

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

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

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

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

228 The derivatives are as follows:

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

229 where

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

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

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

231 Thus

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

232 Without ADPs

233 Without ADPs the equations simplify down to

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

234 and

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

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

237 Periodic Boundary Conditions

238 Periodic boundary conditions can be helpful when simulating extended solids or large

239 nanoparticles. In this case all the non-crystallinity is contained within the simulation

240 box and the box is repeated to create the longer distance peaks observed in the PDF.

241 To perform this we can break up the Debye equation into two main parts, the part

242 that describes the interatomic distances within the simulation box and those between

243 boxes. Neglecting the thermal motion portion:

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

244 where

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

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

245

246 2.2 COMPUTATION

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

251 **HPC and GPUs**

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

256 **GPUs and Parallelization**

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

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

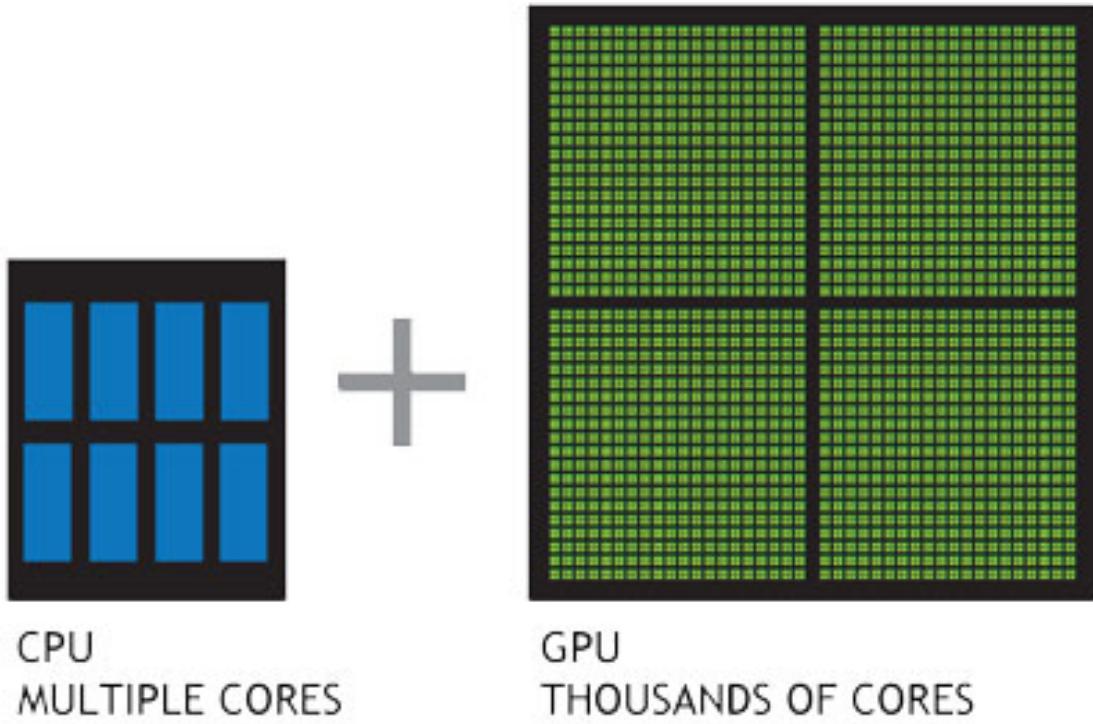


Figure 2.1: Comparison of the CPU and GPU chip architectures

267 Map from ij space to k space

268 The above equations, although formally correct, are very inefficient. $F(Q)$ and its
 269 gradient are indexed over all the atoms twice, however there are symmetries that
 270 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

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

271

272 E denotes the atomic coordinates in ij space, E' denotes $F(Q)$ before the summation
 273 in ij space, B denotes the atomic pairs in k space, B' denotes $F(Q)$ in k space, and
 274 Z denotes the final summed $F(Q)$. For the operators, ϕ denotes the mapping from

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

277 To properly define Σ' we must establish whether $F(Q)$ is an even function. We
 278 can accomplish this by examining each of the portions of $F(Q)$, $\alpha, \beta, \tau, \Omega$. Ω is even,
 279 since r_{ij} is the interatomic distance, which is the same despite a flip of indicies, Q
 280 does not depend on the atomic indicies, and since Qr_{ij} is even so is $\sin Qr_{ij}$. Thus,
 281 Ω is even. Providing similar analysis to τ we can see that while \vec{u}_{ij} is odd, so is
 282 the unit displacement vector between the two atoms, thus the two odds cancel out.
 283 Intuitivly this makes sense, since the $F(Q)$ equation is fundamentally interested in the
 284 interatomic distances which is even. Thus, switching atom indicies does not change
 285 $F(Q)$. Due to the even nature of the $F(Q)$ operator the Σ' operator sums over all the
 286 atom pairs, and multiplies by two to reflect the double counting of the Σ operator.

For the gradient a similar mapping is used:

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

287

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

293 GPU Memory Allocation

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

307 Thus the pair allotment is:

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

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

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

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

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

310 Thus the gradient allotment is:

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

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

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

312 Thus the allocation is:

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

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

316

CHAPTER 3

317

BENCHMARKING

318

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

319

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

320

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.

323

3.1 PDF

324

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 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 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_{obs} , according to

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

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

340 Spin-polarized DFT calculations were carried out using the Vienna ab initio sim-
341 ulation package (VASP) [?, ?] within the Perdew-Burke-Ernzerhof (PBE) exchange-
342 correlation functional [?]. The projected augmented wave method [?] and a kinetic
343 energy cutoff of 400 eV were used. Structural optimization was performed until the
344 total energy and ionic forces were converged to 10^{-6} eV and 10 meV/Å, respec-
345 tively. The amorphous Au₅₅ structures were generated by simulated annealing using
346 the classical embedded atom method potential [?]. Different annealing temperatures
347 between 1200 K and 1670 K (bulk melting temperature of Au) were used and the
348 thermally equilibrated structures were cooled down to 300 K before minimization at
349 0 K. Further optimization using DFT leads to total energies that vary within 1-2
350 eV among different amorphous structures and the lowest energy one was used as the
351 target structure. The target structure of Au₁₀₂ was taken as the Au₁₀₂ core of the
352 DFT-optimized Au₁₀₂MBA₄₄ cluster [?].

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

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

356 The thresholded spring potentials are based on those previously proposed on by Pe-
357 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
358 bounds of the spring threshold r_t . The resulting restoring forces on the out-of-bound
359 atoms bring the system back within the bounds of the PDF, R_{\min} and R_{\max} , and
360 therefore preventing the system from exploding or collapsing. Otherwise, incorrect

refinements may result by having atomic pair distances out of the PDF bounds. κ is the spring constant in eV/Å and the Rw potential is converted from unitless to eV via multiplication by a conversion factor λ .

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

Model Parameters

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

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

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

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

388 **Au55: surface relaxed**

389 We first test our algorithm by solving the crystalline Au₅₅ (*c*-Au₅₅) cluster structure
390 from its PDF. The starting structure is taken as the bulk-cut cuboctahedron of Au₅₅
391 with a uniform bond length of 2.89 Å. Due to finite-size and surface effects, the DFT-
392 relaxed cluster structure shows a distinctively different bond length distribution as a
393 function of the bond's distance to the cluster center of mass, and therefore is difficult
394 to model with a small box approach which assumes an identical unit cell throughout
395 the whole system.

396 **Run Parameters**

397 R_{\min} and R_{\max} for this simulation were 2.45 Å and 11.4 Å, respectively, with $\delta Q =$
398 0.24 Å⁻¹. The simulation ran for approximately 34 minutes, over a total of ~40
399 thousand configurations. The results are shown in Fig. 3.1.

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

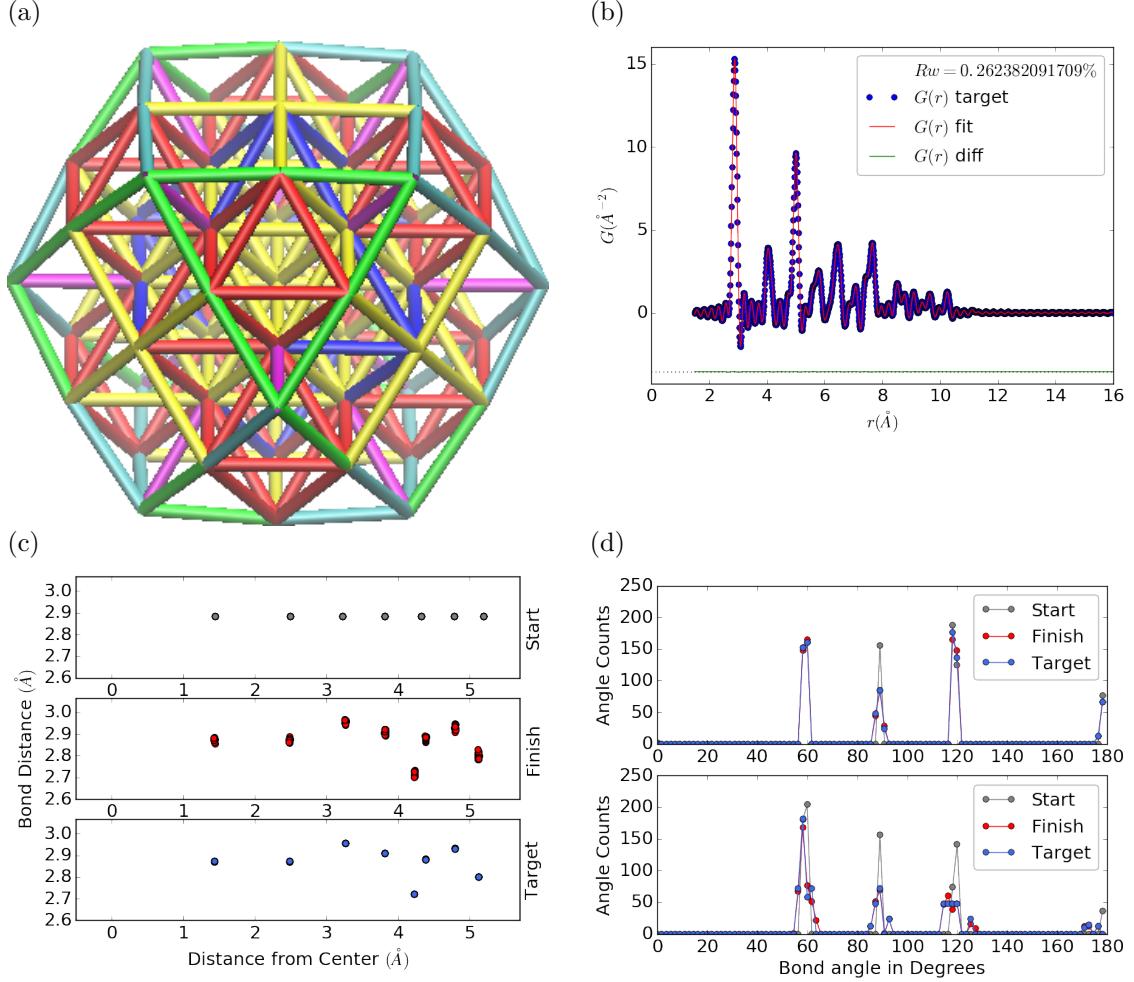


Figure 3.1: Au_{55} PDF fitting of DFT-optimized $c\text{-Au}_{55}$. (a) the final structural solution ($Rw=0.3\%$) with bond lengths color-coded by step of 0.05\AA , (b) the target PDF(blue dots) overlaid with the PDF of the final structure (solid red lines) with the difference in green lines offset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

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

Put this somewhere

.

413 **Au55: surface disordered**

414 In addition to surface relaxation, the structure of a cluster or nanoparticle is often
415 disrupted by the presence of defects and/or ligand bound to the surface. To mimic
416 such surface disorders, we took the DFT-optimized *c*-Au₅₅ structure from case I as
417 the starting structure and randomly displaced the surface atoms with a normal distri-
418 bution of $\sigma = 0.2 \text{ \AA}$. All atoms are allowed to move in the HMC simulation, including
419 the originally undisturbed core, which is a Au₁₃ cluster with O_h symmetry.

420 R_{\min} and R_{\max} for this simulation were 1.95 Å and 12.18 Å, respectively, with
421 $\delta Q = 0.23 \text{ \AA}^{-1}$. The simulation ran for approximately 3.6 hours, over a total of ~ 270
422 thousand configurations. The results of the simulation are shown in Fig. ??.

423 Overall, good agreement is found between PDFs of the target structure and the
424 final structural solution, even out to larger r , with an $Rw = 0.6\%$ starting from an
425 $Rw = 50.4\%$ (see Fig. S2). The radial bond distribution and angle distribution
426 show reasonably good agreement, but with lower degree of crystallinity in the final
427 structure compared to the target structure. The discrepancy is most obvious in
428 the core: despite the identical core structure in the starting and target structures,
429 the core atoms were displaced in the HMC simulations in order to achieve a “best”
430 solution. This is because PDF measures the global average of interatomic distances
431 between each atomic pair and does not contain direct information about the locality
432 of such pairs, e.g. on the surface or cores. If such information is obtained a priori, for
433 example, from theoretical prediction or other experimental measurements, the core
434 structure can then be fixed and excluded from HMC dynamics.

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

440 and high CN numbers. This is partly caused by the displacement of the core atoms
441 mentioned above, and partly by the lack of CN constraints for the PDF fitting, which
442 has been previously demonstrated in the case of α -Si [?]. Additional experimental
443 data, e.g. from EXAFS or NMR, may help to steer the simulations towards better
444 agreement in both PDF and CN distribution.

445 Au55: amorphous

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

451 R_{\min} and R_{\max} for this simulation were 2.6 Å and 11.26 Å, respectively, with
452 $\delta Q = 0.25 \text{ \AA}^{-1}$. The simulation ran for approximately an hour, over a total of ~ 87
453 thousand configurations. The results of the simulation are shown in Fig. ??.

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

466 **Au102: triple phase**

467 Our final benchmark is Au₁₀₂, whose structure was initially solved by Jadzinsky and
468 co workers using x-ray crystallography [?] and further confirmed by DFT studies [?].
469 The Au₁₀₂ structure consists of three main parts, a 49-atom Marks decahedron core,
470 two C_5 caps consisting of 20 atoms each, and 13 equatorial atoms. Unlike previous
471 cases, the multi-symmetry nature of the structure, i.e. each part has its own distinct
472 symmetry, poses a challenge for PDF-based solution of the structure. This is because
473 of the atomically centralized nature of the PDF, in which each atom “sees” a density
474 of other atoms surrounding it and has a strong tendency towards becoming the center
475 of the main symmetry group. Such tendency may lead to a solution where some of
476 the correct atomic symmetries are discarded in favor of the core symmetry.

477 **Starting from fcc structure**

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

480 R_{\min} and R_{\max} for this simulation were 2.7 Å and 16. Å, respectively, with $\delta Q =$
481 0.18 Å⁻¹. The simulation ran for approximately two hours, over a total of ~82
482 thousand configurations. The results of the simulation are shown in Fig. 3.6.

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

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

502 **Marks decahedron**

503 The starting structure, a Marks Decahedron, was generated by the ASE Cluster tool
504 with 2 atoms on the [100] face normal to the 5-fold axis, 3 atoms on the [100] plane
505 parallel to the 5-fold axis and, 1 atom deep Marks reentrance. This produced a
506 structure with 101 atoms which was extended by one more Au atom to fill out the
507 Au_{102} structure.

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

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

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

521 **Au147**

522 **3.2 PDF WITH ADPs**

523 **ADP 50**

524 1. Basic 50% larger magnitude

525 2. Random addition to APDs

526 3. Janus ADPs

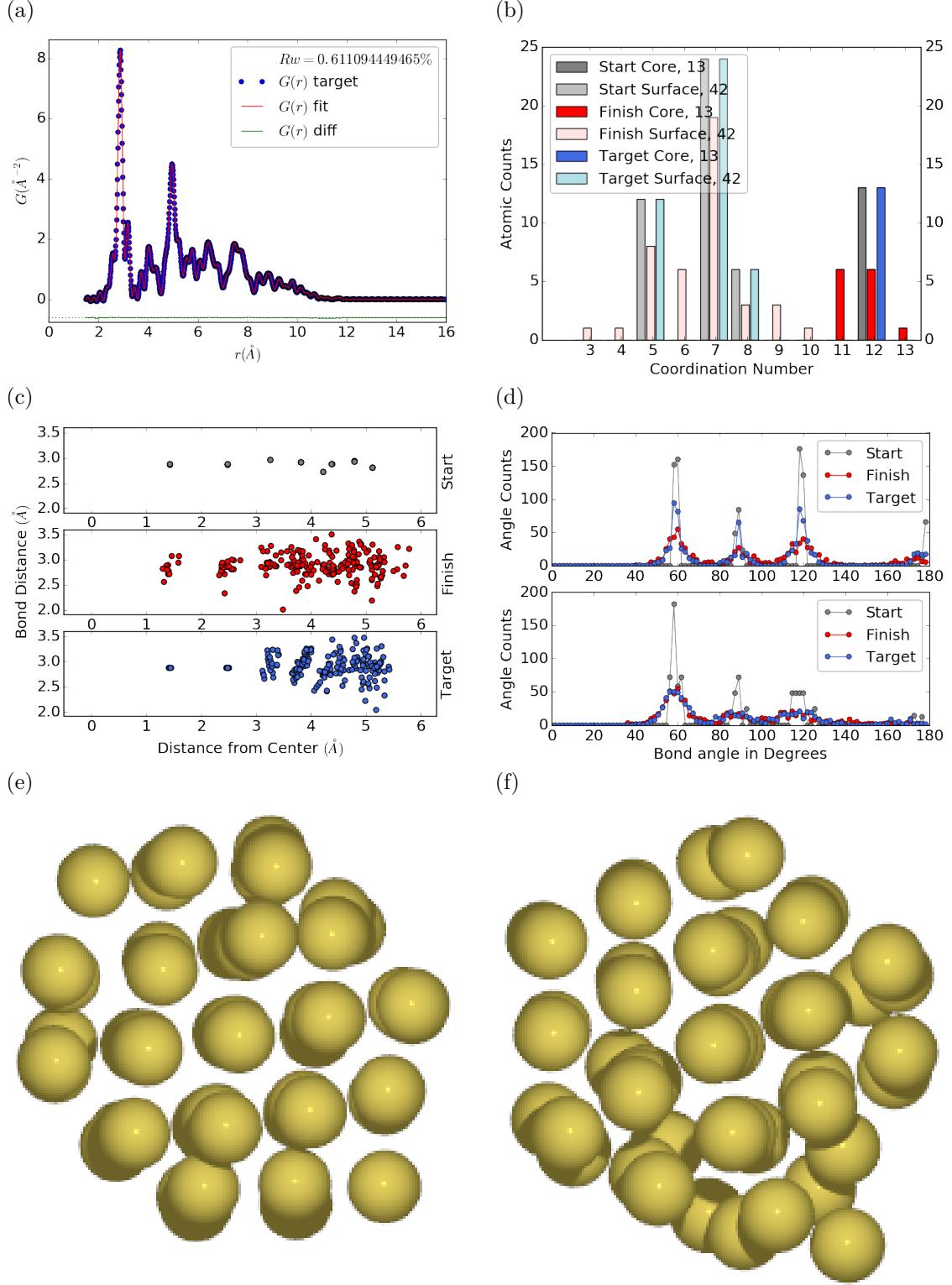


Figure 3.2: Au_{55} PDF fitting of surface-disordered Au_{55} . a) the target structure, b) the final structural solution ($Rw=0.6\%$), c) the comparison of PDFs, d) the CN distribution with the number of atoms in either the core or the surface, e) the radial bond distribution, and f) the bond angle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

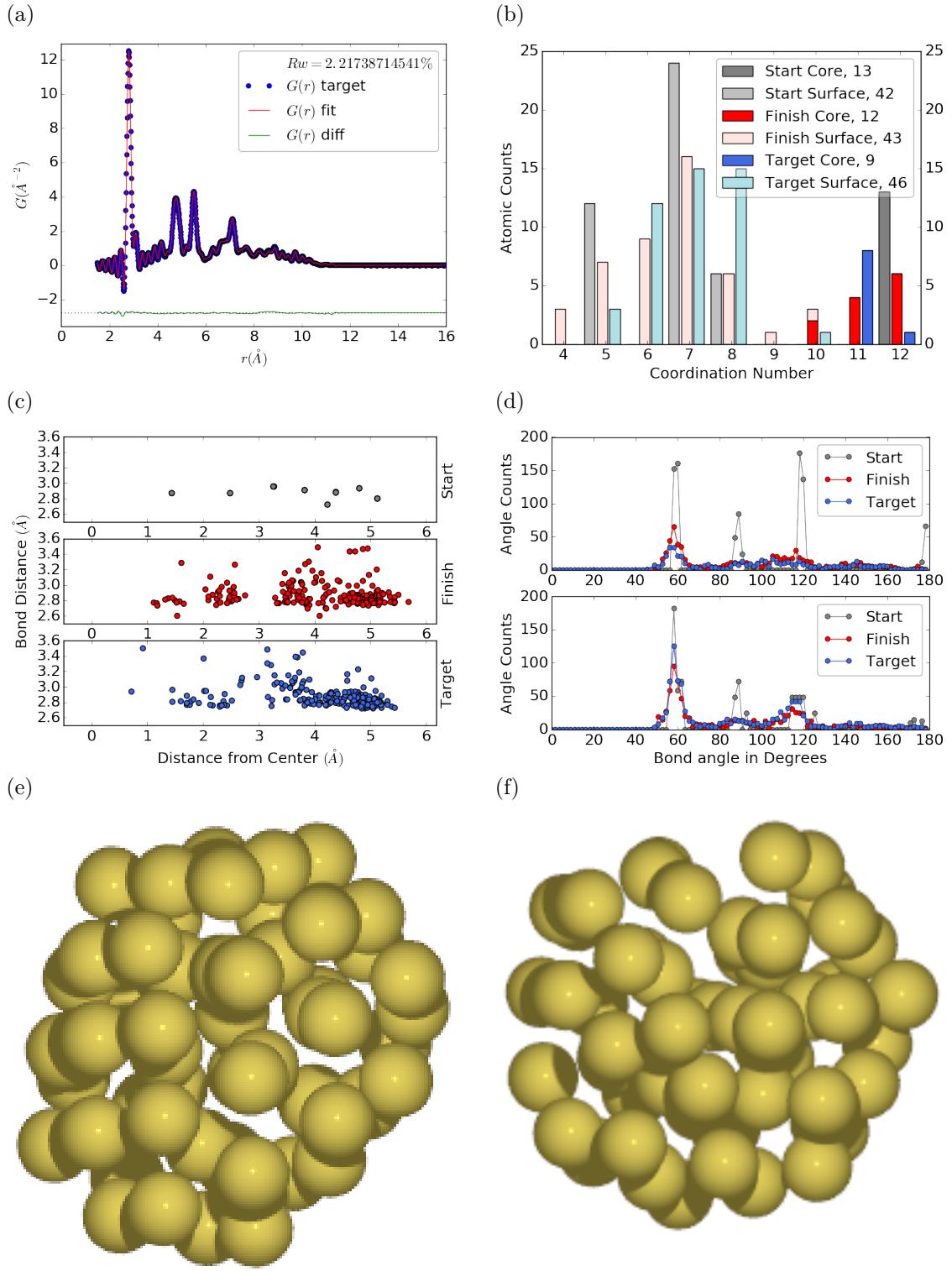


Figure 3.3: Similar to figure 3.2 for DFT-optimized amorphous Au_{55} .

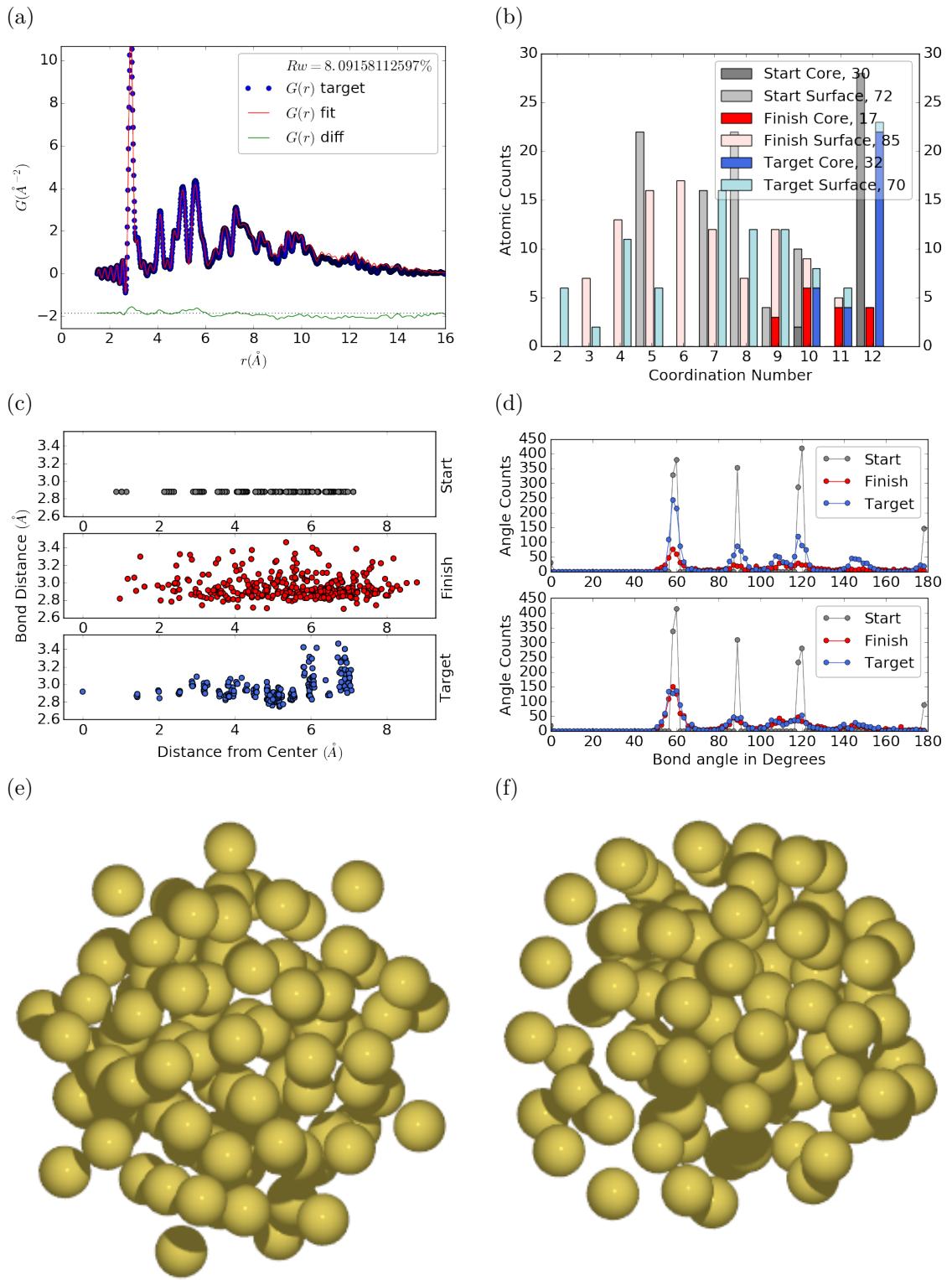


Figure 3.4: Similar to figure 3.2 for Au_{102} as in DFT-optimized $\text{Au}_{102}\text{MBA}_{44}$ cluster.

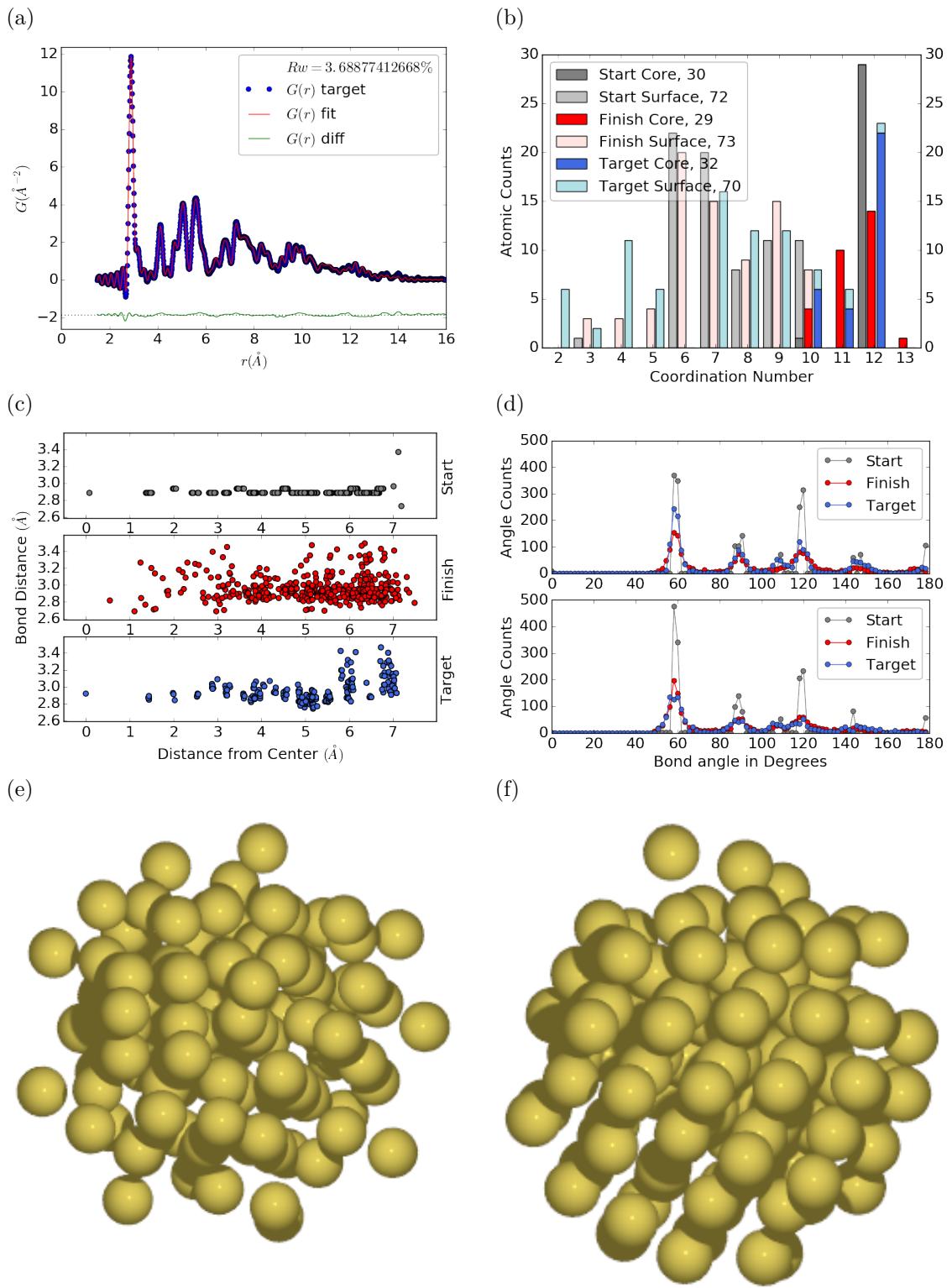


Figure 3.5: Similar to Fig. 3.6 with Marks decahedron as the starting structure.

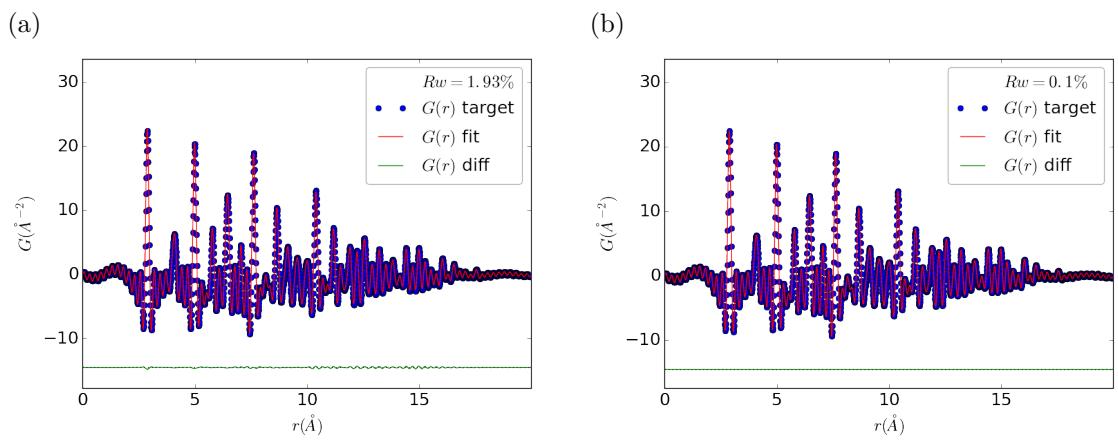


Figure 3.6

527

CHAPTER 4

528

X-RAY TOTAL SCATTERING DATA AQUESITION AND

529

PROCESSING

530

4.1 INTRODUCTION

531 X-ray total scattering experiments are generally performed at synchrotron light sources,
532 as only these sources can provide the needed flux, energy, and high momentum trans-
533 fer vectors needed to obtain reliable PDFs. [3, 4] Despite the need for a dedicated
534 facility to perform the total scattering experiments, the experiments themselves are
535 fairly forgiving, allowing for reactive gasious environments, experiment temperatures
536 ranging from 2 K to 1200 K, and even electrochemical cycling. [2, 8, 9] The rapid
537 PDF data aquesition associated with 2D area detectors creates a data management
538 problem, as 96 hours of beamtime could result in almost 10,000 images which need
539 to be associated with the experimental conditions and detector metadata. [3] Finaly,
540 all this data needs to be processed by masking bad pixels and regions, integrating
541 azimuthally, and converting the scattering data to the PDF. [6, 5, 10, 7, 1]

542

4.2 DETECTOR Q RESOLUTION

543 To properly azimuthaly integrate the images taken from the detector the Q resolution
544 of the pixels must be calculated. Integrating using even bins will cause pixels which
545 are not on the same ring to be binned together, causing the incorrect value of $I(Q)$
546 to be obtained and a larger standard deviation in the integrated data. To properly
547 calculate the Q resolution the resolution of each of the pixels in 2θ must be calculated.

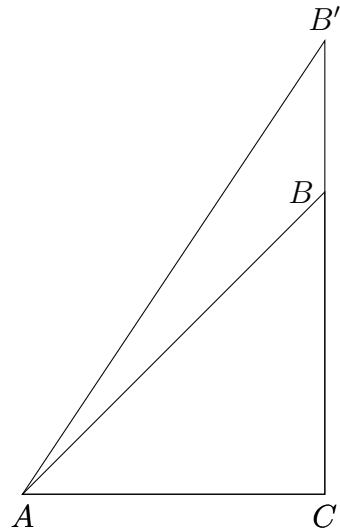


Figure 4.1: Scattering onto a flat detector

548 Figure 4.1 shows the scattering of x-rays onto a flat image plate detector. In this
 549 diagram the bottom of the n th pixel is B while the top is B' . The resolution of this
 550 pixel in 2θ is $\angle BAC - \angle B'AC$. Thus the resolution, calculated from the distances is

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

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

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

554 where λ is the x-ray wavelength.

555 For a Perkin Elmer image plate, like the one used at the NSLS-II's XPD and the
 556 APS's 11-ID-B, the resolution function is shown in 4.2. For the same detector the
 557 number of pixels per Q is shown in 4.3

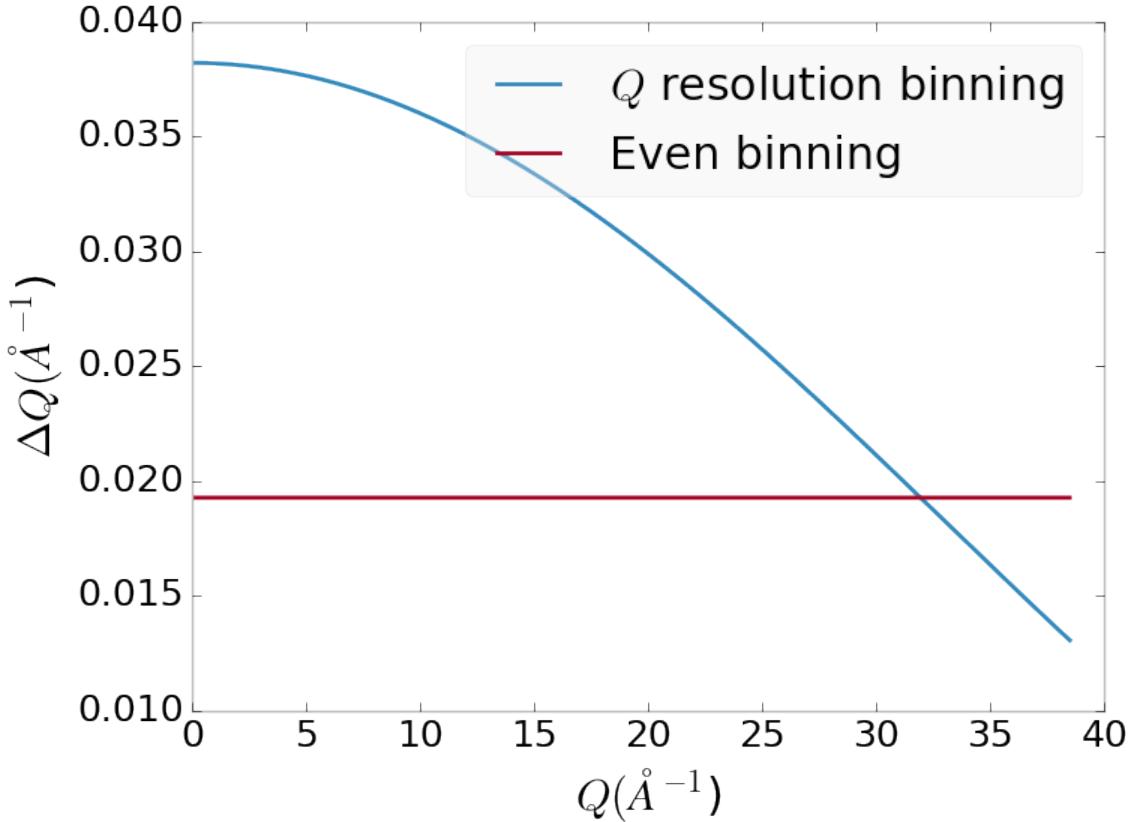


Figure 4.2: Q resolution as a function of Q .

558 4.3 AUTOMATED MASK GENERATION

559 **Introduction**

560 Detector masking is an important part of any x-ray scattering workflow as dead/hot
 561 pixels, streak errors, and beamstop associated features can be averaged into the data
 562 changing the signal and its statistical significance. While some features, like the
 563 beamstop holder, can be easily observed and masked by hand other are much more
 564 difficult to observe even on large computer monitors. Additionally, while dead/hot
 565 pixels and streaks are usually static the hot pixels associated with textured or sin-
 566 gle crystal scattering or cosmic rays are not. Thus, coming up with an automated
 567 method for finding such erroneous pixels is important, especially as high flux diffrac-
 568 tion beamlines can generate data very quickly.

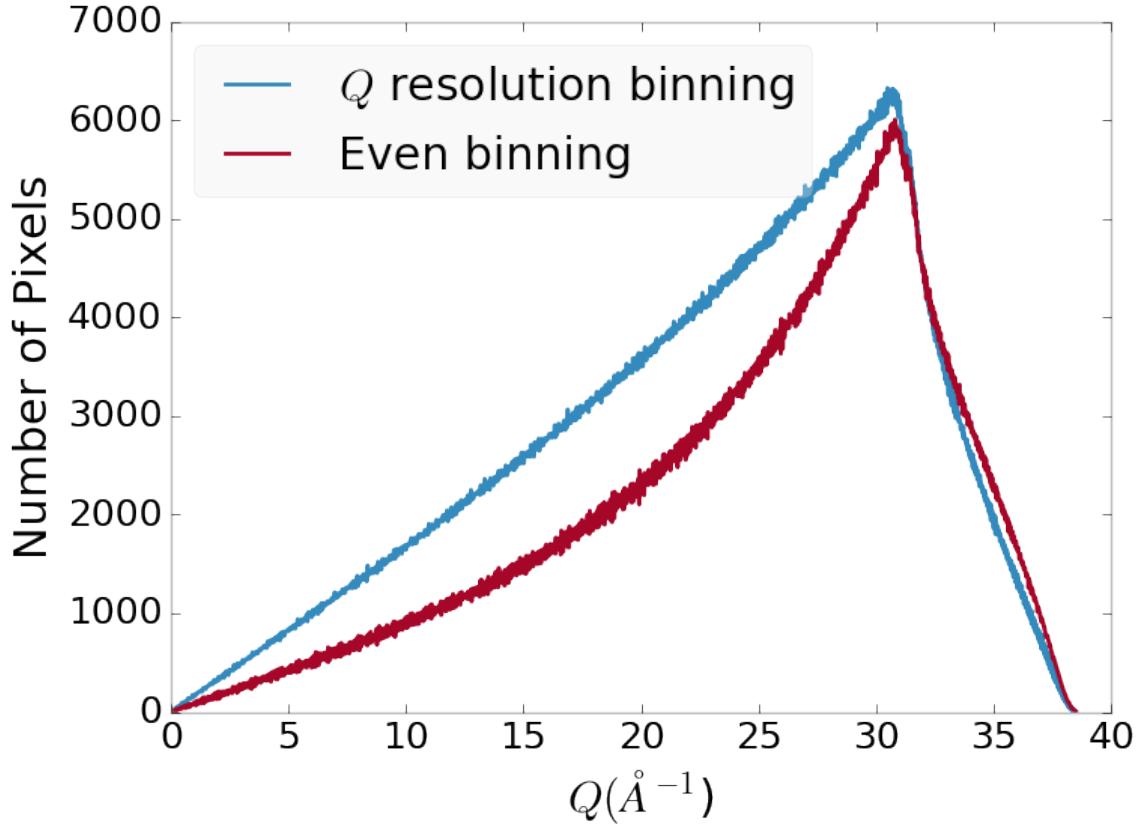


Figure 4.3: Number of pixels as a function of Q , binned at the Q resolution of the detector.

569 While this problem can be quite complex in the most general case, we can use the
 570 annular symmetry of the powder scattering pattern to our advantage, by comparing
 571 a pixel against pixels in the same ring. Since non-textured powder scattering should
 572 produce the same pixel intensity for a given ring we can mask any pixels which are α
 573 standard deviations away from the mean. This method relies on the aforementioned
 574 pixel binning algorithm, as using miss sized bins will cause some pixels which should
 575 be in separate rings to be put together, and others which should be in the same ring
 576 to be separated. In that case the masking algorithm will overestimate the number of
 577 pixels to be masked due to the additional statistical variation in the sample.

578 **Algorithm Design**

579 The masking algorithm procedure takes in the image and a description of the pixel
580 positions in either distance from the point of incidence or in Q . The image is then
581 integrated twice, producing both the mean $I(Q)$ and the standard deviation of each
582 $I(Q)$ ring. The mask is created by comparing the pixel values against each ring's
583 standard deviation and threshold α . Note that the threshold can be a function of
584 distance from the point of incidence or Q .

585 **Test Cases**

586 To study the effectiveness of the masking we ran the algorithm against both simulated
587 experimental data. In the case of the simulated data four systems were created: 1)
588 dead/hot pixels with varying numbers of defective pixels, 2) beamstop holder with
589 varying beamstop holder transmittance, 3) rotated beamstop holder with varying
590 beamstop holder transmittance, and 4) beamstop holder with dead/hot pixels. The
591 base scattering was produced by

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

592 where r is a pixel's distance from the beam point of incidence. The positions of
593 the dead/hot pixels were chosen at random as was the dead or hot nature of the
594 defect. Dead pixels had values from 0 to 10, while hot pixels had values from 200
595 to 255. The beamstop was positioned at the vertical center of the detector with an
596 initial width of 60 pixels and final width of 120 pixels. The height of the beamstop
597 was 1024 pixels. The beamstop was calculated to attenuate the x-ray scattering
598 signal at various transmittance, as various beamstop holder materials have different
599 transmittance. Two versions of the masking algorithm were run for each test case, one
600 using the standard even bin sizes for the integration step, and one where the bin sizes
601 are tuned to the pixel Q resolution as discussed in 4.2.

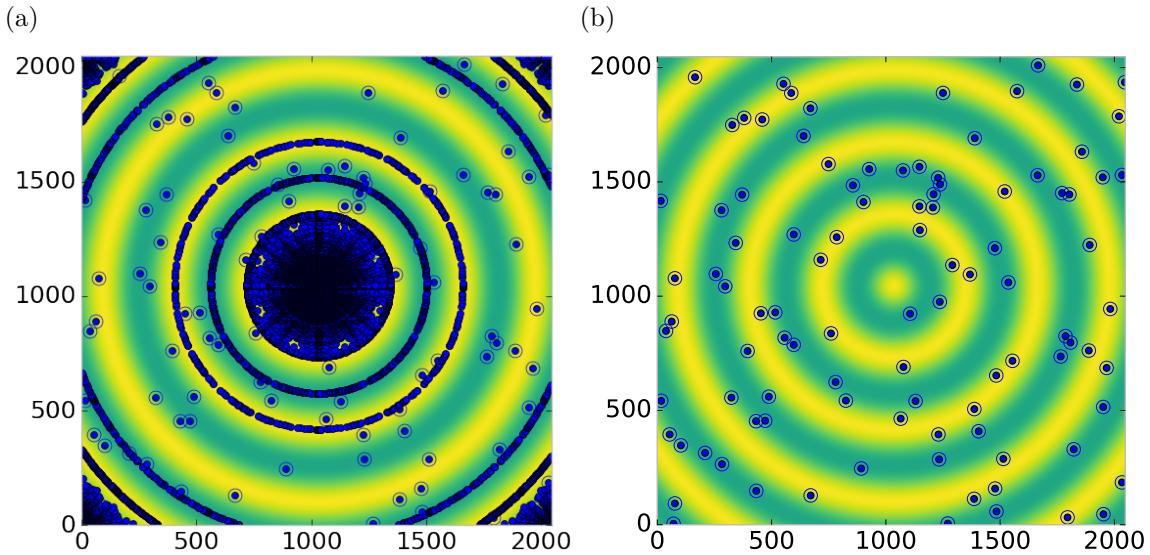


Figure 4.4: Generated dead/hot pixel masks for a detector with 100 bad pixels. a) the standard even bin mask and b) the Q resolution binned mask. The bad pixels are noted with open circles, masked pixels are noted with closed circles.

602 Results and Discussion

603 Figures 4.4-4.11 show the results of the masking algorithm on simulated images. The
 604 dead/hot pixel masking shows the importance of using the Q resolution based bin sizes
 605 as the even bin based mask have a tendency to over mask the image, removing pixels
 606 which contain valuable signal. This overmasking is caused by pixels being improperly
 607 associated with one another by the even bins. Figure 4.4 indicates that the masking
 608 algorithm, with the proper binning, masks the image perfectly, with no missed bad
 609 pixels or good pixels masked. This is not the case in figures 4.5 - 4.7 as we can see
 610 pixels which should have been masked but were not. Despite these missed pixels no
 611 pixels were improperly masked in any of the well binned images. These test cases
 612 are actually more difficult than experimental data, as the dynamic range of most
 613 detector causes the dead/hot pixels and single crystal/textured peaks to be orders of
 614 magnitude away from the desired signal.

615 The beamstop holder masks shown in figures 4.8 - 4.11, which were all run with

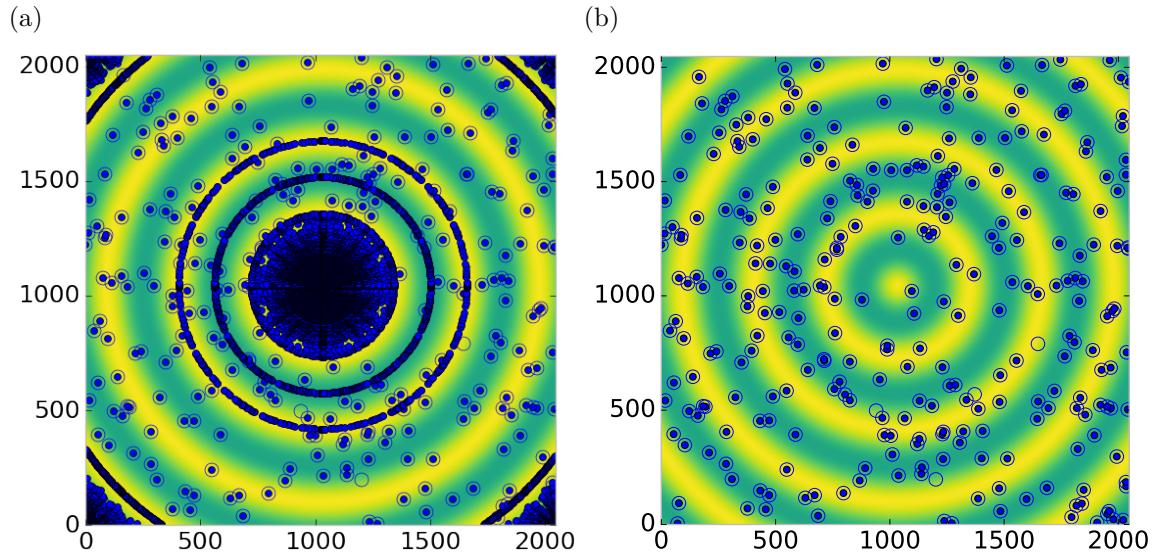


Figure 4.5: Generated dead/hot pixel masks for a detector with 300 bad pixels. a) the standard even bin mask and b) the Q resolution binned mask. The bad pixels are noted with open circles, masked pixels are noted with closed circles.

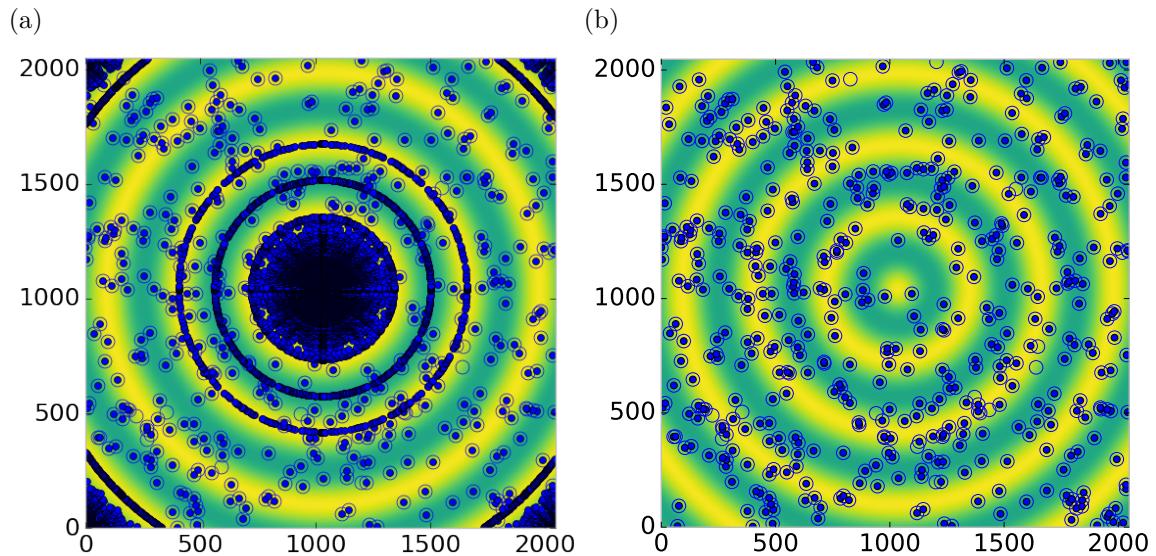


Figure 4.6: Generated dead/hot pixel masks for a detector with 500 bad pixels. a) the standard even bin mask and b) the Q resolution binned mask. The bad pixels are noted with open circles, masked pixels are noted with closed circles.

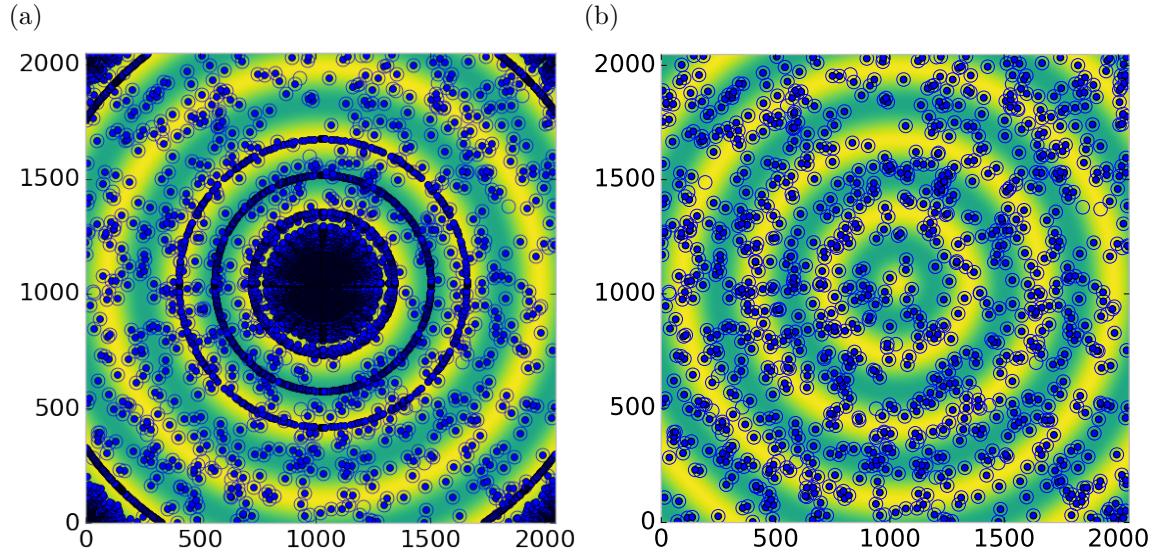


Figure 4.7: Generated dead/hot pixel masks for a detector with 1000 bad pixels. a) the standard even bin mask and b) the Q resolution binned mask. The bad pixels are noted with open circles, masked pixels are noted with closed circles.

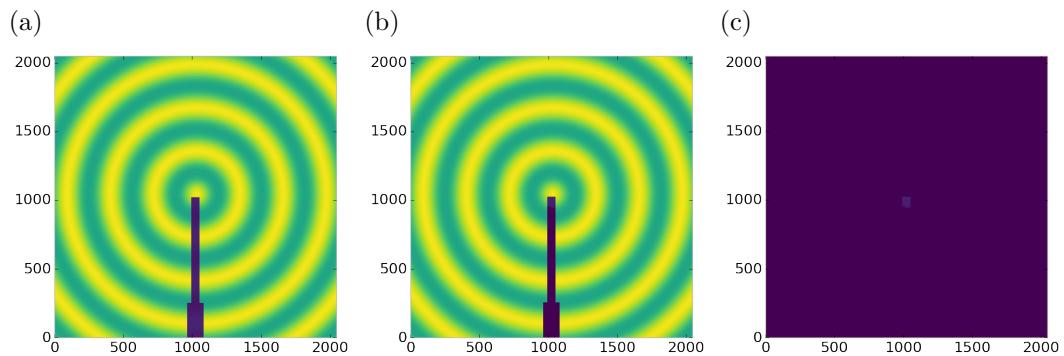


Figure 4.8: Generated beamstop holder masks for a beamstop holder with 10% transmittance. a) the raw image, b) the masked image, c) and the missed pixels

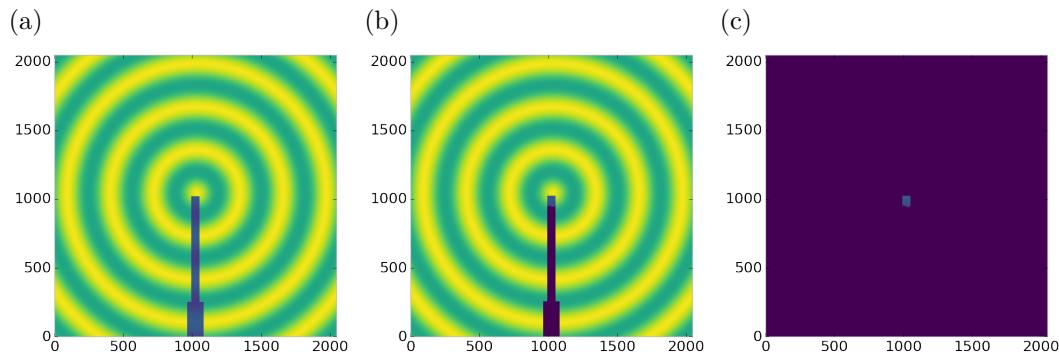


Figure 4.9: Generated beamstop holder masks for a beamstop holder with 30% transmittance. a) the raw image, b) the masked image, c) and the missed pixels

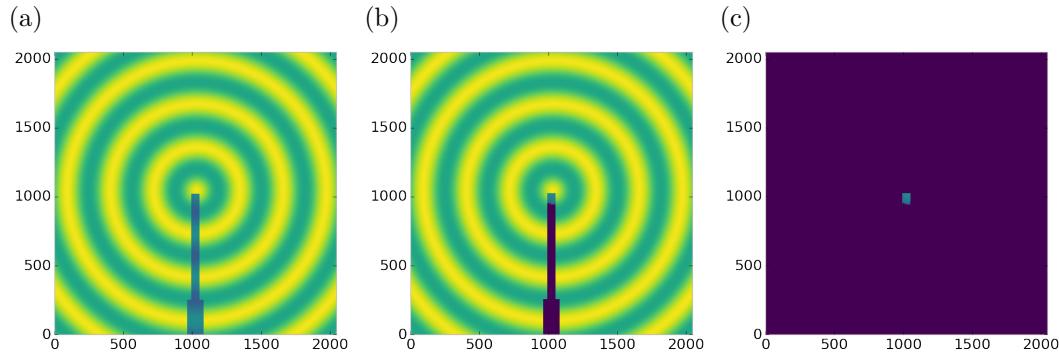


Figure 4.10: Generated beamstop holder masks for a beamstop holder with 50% transmittance. a) the raw image, b) the masked image, c) and the missed pixels

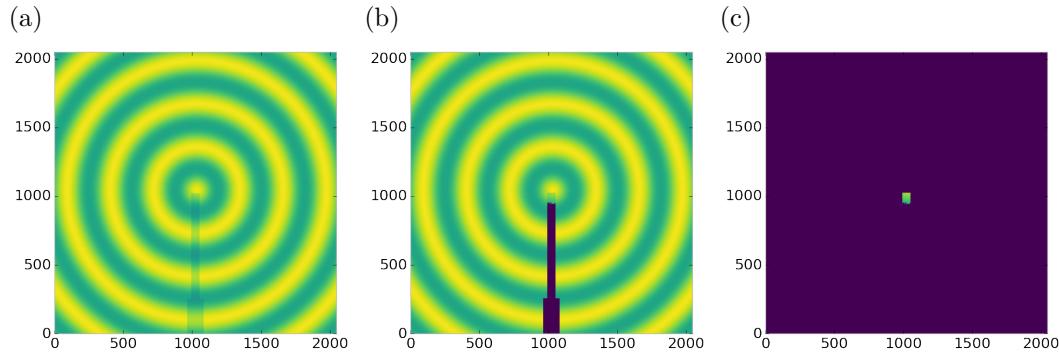


Figure 4.11: Generated beamstop holder masks for a beamstop holder with 90% transmittance. a) the raw image, b) the masked image, c) and the missed pixels

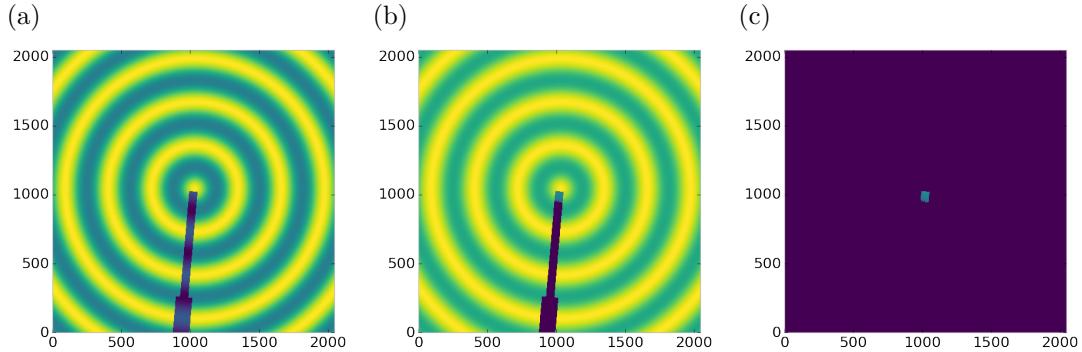


Figure 4.12: Generated beamstop holder masks which is rotated away from verticle

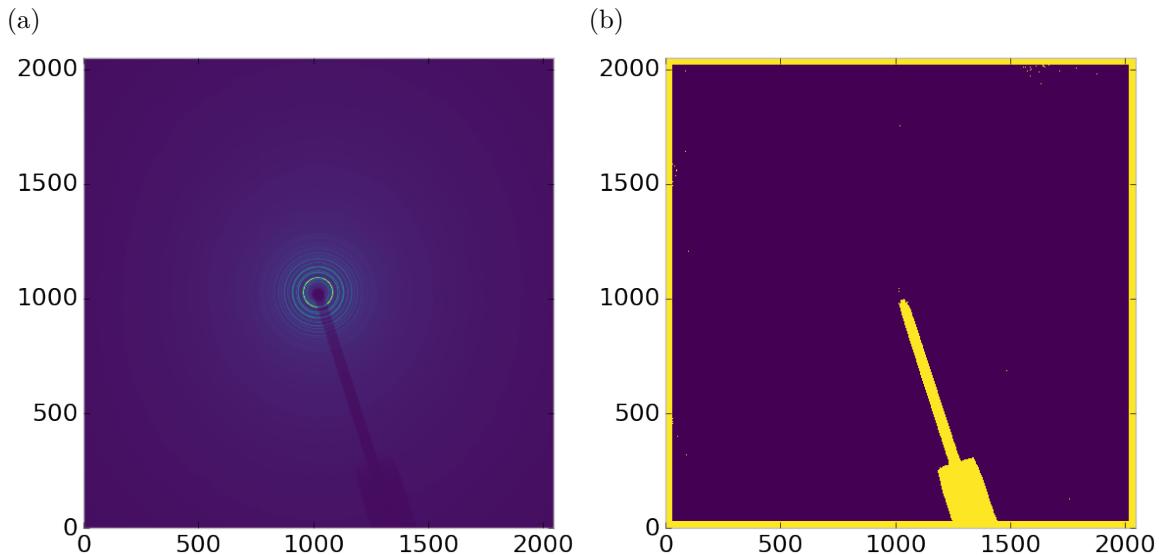


Figure 4.13: Masked experimental data. a) the raw image, b) the mask

the Q resolution binning show similar results across the transmittance range, missing only a small part of the beamstop holder near the point of incidence. Near this point the beamstop holder becomes a statistically significant part of the total number of pixels in a given ring, thus it can not be masked out using a statistical search of the rings. For most PDF and XRD studies this small area can be masked automatically by masking all the pixels who's distance from the point of incidence is smaller than a given radius r , or can be negelected outright as the area is not used in the analysis or refinement. Similar results were produced for beamstop holders which were rotated away from the vericle position, as shown in figure 4.12

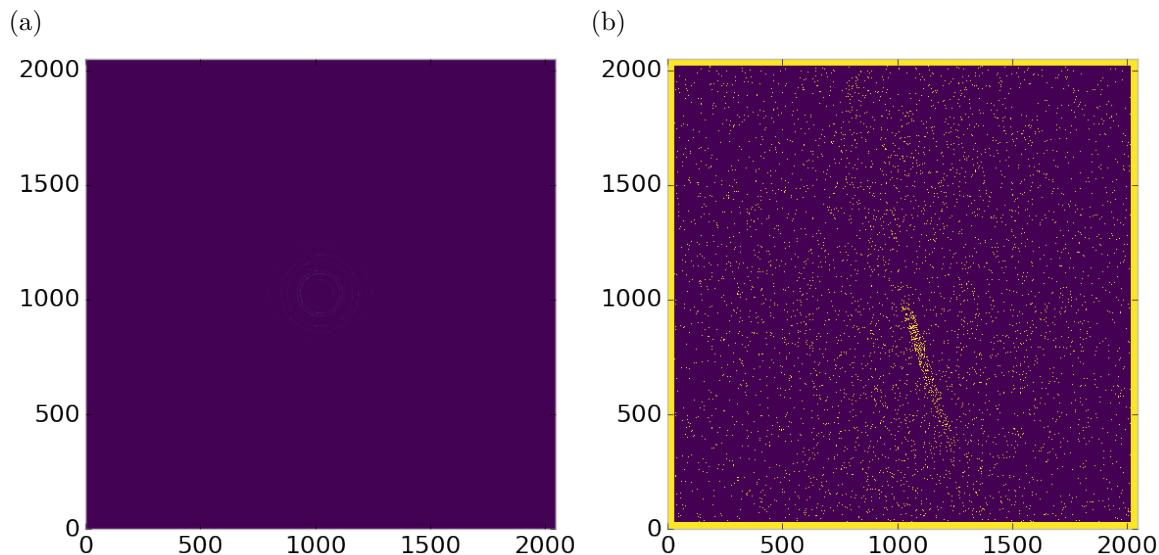


Figure 4.14: Masked experimental data with Pt single crystal signal. a) the raw image, b) the mask

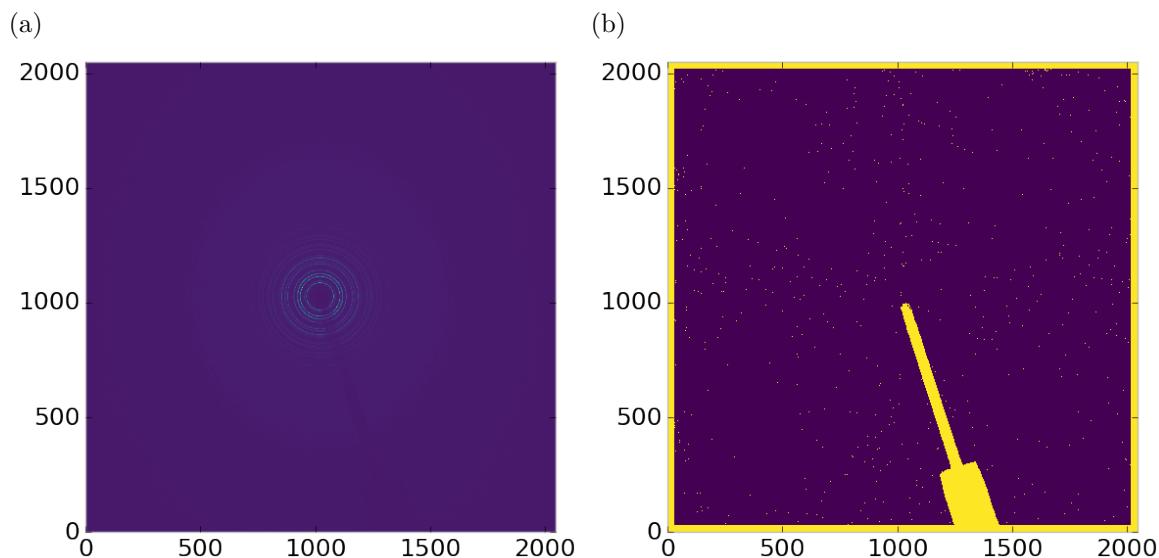


Figure 4.15: Masked experimental data with Pt single crystal signal using figure's 4.13 as a starting mask. a) the raw image, b) the mask

625 Working with actual experimental data, obtained at the Advanced Photon Source
626 beamline 11-ID-B, shows the difficulty of masking images which have low phonon
627 counts. While the masking of experimental data taken with longer exposures, con-
628 sisting of 250 .2 second shots, shown in figure 4.13 provides very sharp edges to the
629 beamstop holder, and very little extra masking beyond the occasional dead pixel, this
630 is not the case for the single crystal data. The single crystal data is more problem-
631 atic because of its short exposure time and low flux, with 500 frame at a .1 second
632 exposure and having shrunk the beam size. The low flux is to prevent the very strong
633 single crystal peaks from damaging the detector. However, this causes the image to
634 be less statistically viable than ideal, causing problems with the mask as seen in figure
635 4.14. This can be alleviated to some degree by using the previously generated mask
636 as a starting mask for the single crystal image, as shown in 4.15. While the masking
637 algorithm still produces many diffuse masked pixels, they are far fewer, this may
638 be due to the removal of the beamstop which could have contributed to the large
639 standard deviation in figure 4.14.

640 Conclusions

641 In this section the masking algorithm, which relies on both Q resolution based binning
642 and a statistical approach to azimuthal symmetry, was developed. The focus of
643 this algorithm was to remove many unwanted detector features associated with pixel
644 defect, beamstop holder associated scattering attenuation, and single crystal/texture
645 peaks. Simulated data was used to evaluate the beamstop holder and dead/hot pixel
646 masking capacity, while experimental data was used to check for single crystal and
647 texture based masking. Q resolution based binning was shown to be very important to
648 avoid overmasking. The ability of the mask writer to mask images is somewhat limited
649 by the overall statistical image quality, although some deficiencies can be obtained by
650 using previously generated masks as starting points. This masking algorithm is now

651 in use in the data processing workflow and will be available in scikit-beam soon.

652 4.4 AUTOMATED IMAGE AZIMUTHAL INTEGRATION

653 Using the Q resolution binning and masking developed in sections 4.2 and 4.3 the
654 images can be properly integrated. Generally, images are integrated by taking the
655 mean value of the pixels in a ring. However, other statistical measures of the average
656 value can be used, like the median.

657 Figures 4.16-4.18 show the importance of masking and the choice of average function.
658 All the figures were produced using the same dataset, 50 °C Pr_2NiO_4 taken at
659 the APS's 11-ID-B on a Perkin Elmer area detector. The automatic masking alpha
660 was 3 standard deviations from the mean. While it is difficult to observe the changes
661 the mask causes in the full $I(Q)$ plot (subfigures a) and b)), the standard deviation
662 plots show the effect of bad pixels on the data (subfigure c)). Subfigure c) for figures
663 4.16-4.18 shows that removal of the beamstop holder lowers the low Q standard de-
664 viation from around .1 to almost .01 out to 15 \AA^{-1} . The high Q subfigures d) and f)
665 in figures 4.16-4.18 show the “kink” effect of the detector edge and beamstop holder,
666 where there is a dip in the $I(Q)$ scattering when the rings include the edge of the
667 detector. This effect seems to be due to both errors in the edge pixel intensity and the
668 beamstop holder as masking of the edges only seems to provide only partial removal
669 of the issue. It is important to note that while integration using the mean of the
670 ring has issues with only the edge mask, as evidenced by the change in slope in 4.17
671 d) around 29.5 \AA^{-1} , the median integration does not include this error. Ideally the
672 detector would have a normal distribution of pixel intensity for a given ring, which
673 would imply an equivalency between the mean and median $I(Q)$ values. Despite the
674 closeness of the mean and median once the final mask has been created, it seems that
675 the median is more reliable, as it was less effected by the beamstop holder in figure
676 4.17. Thus, for subsequent integrations discussed in this work the median is used to

677 avoid any defective features that the masking algorithm may have missed.

678 4.5 CONCLUSIONS

679 This chapter developed and analyzed the proper data processing and reduction method-
680 ology for producing reliable $F(Q)$ data from x-ray total scattering measurements.
681 Binning at the Q resolution of the detector was found to be key to the data process-
682 ing. The primary outcome of using the Q resolution binning was an enhancement in
683 effectiveness for the masking algorithm, producing much fewer false positives for dead
684 pixels. This masking approach was then applied to the integration of experimental
685 data taken at the APD’s 11-ID-B beamline. The automatically generated masks, when
686 combined with edge masks, were found to greatly reduce the overall standard devi-
687 ation of the pixel intensity and produce a smoother $F(Q)$ at high Q , enabling the
688 use of much higher Q data in the PDF. Different statistical measures used in the
689 azimuthal integration was also compared. This comparison showed that the median
690 was a more reliable statistic for integration with data which had more detector de-
691 fects. However, upon properly masking it was shown that these metrics were almost
692 identical. The masking induced similarity between the mean and median shows that
693 the rings, when integrated, may form a Gaussian distribution. The distribution of
694 the pixel intensities for strongly and weakly scattering samples may be investigated
695 in future work.

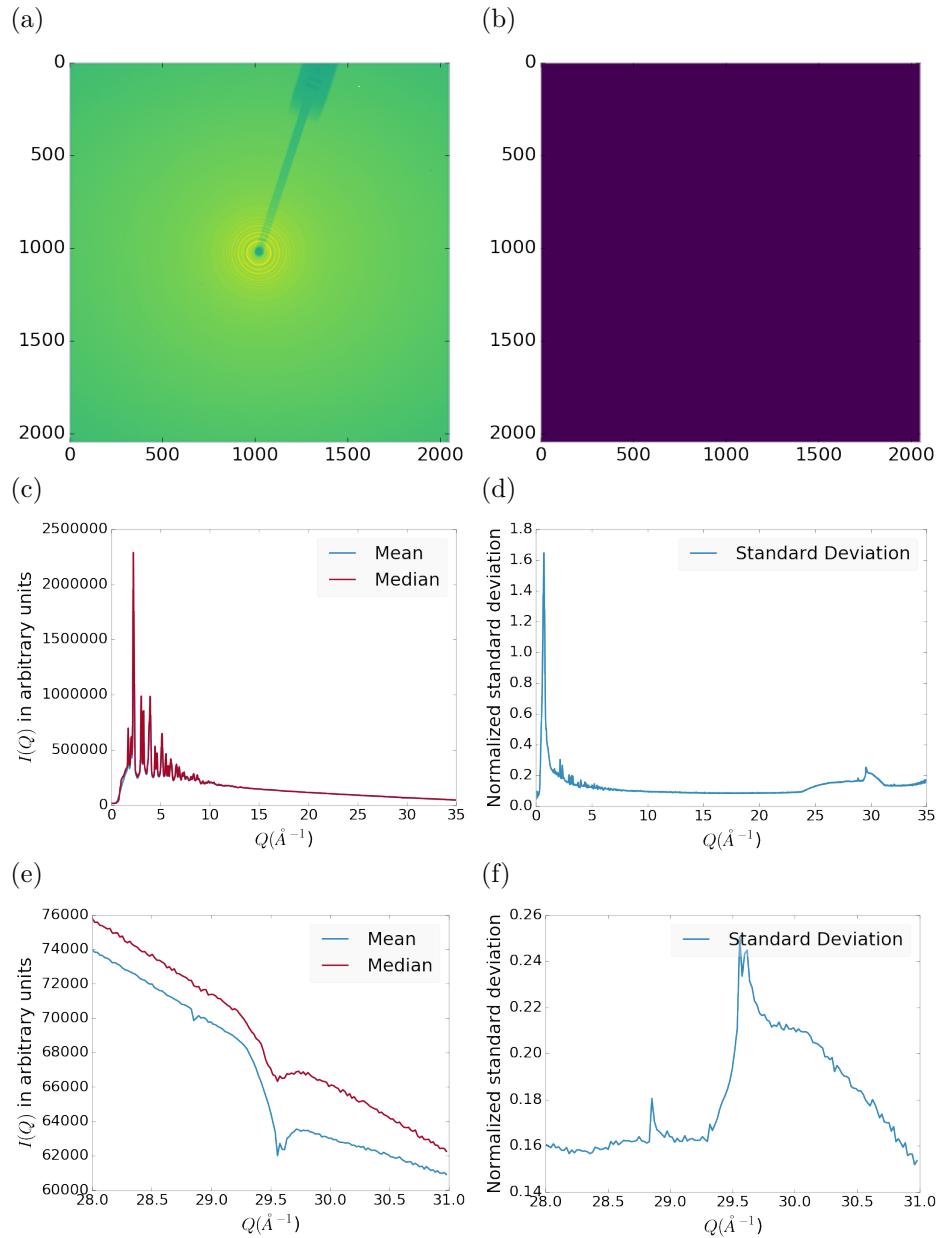


Figure 4.16: Masking, average, and standard deviation of an example x-ray total scattering measurement. This image was produced with no mask. a) the image, b) the mask, c) the mean and median values, d) the standard deviation (normalized to the median), e) a closeup of the 28 \AA^{-1} to 31 \AA^{-1} Q range for the mean and median, f) 28 \AA^{-1} to 31 \AA^{-1} Q range for the standard deviation

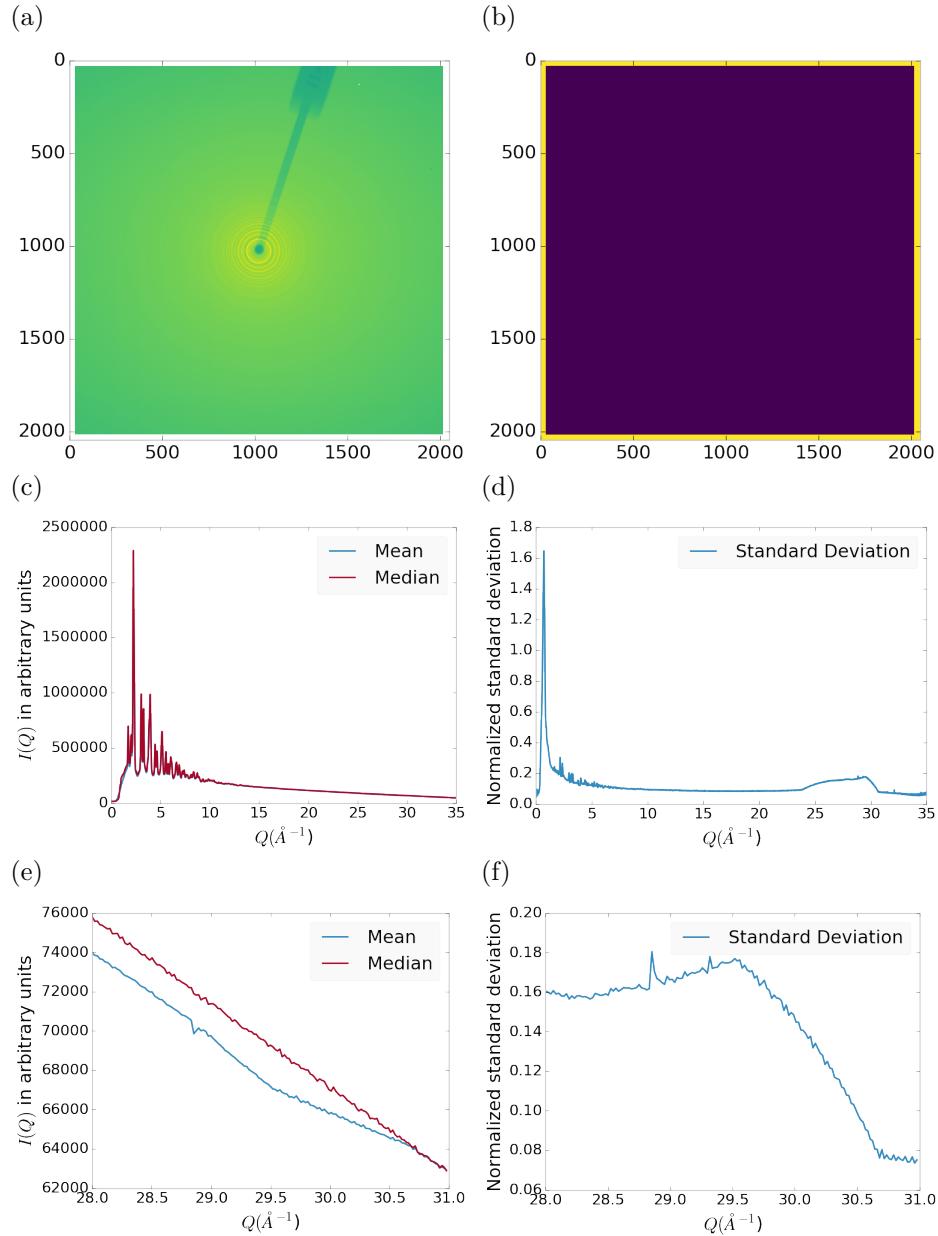


Figure 4.17: Masking, average, and standard deviation of an example x-ray total scattering measurement. This image was produced with only an edge mask. a) the image, b) the mask, c) the mean and median values, d) the standard deviation (normalized to the median), e) a closeup of the 28\AA^{-1} to 31\AA^{-1} Q range for the mean and median, f) 28\AA^{-1} to 31\AA^{-1} Q range for the standard deviation

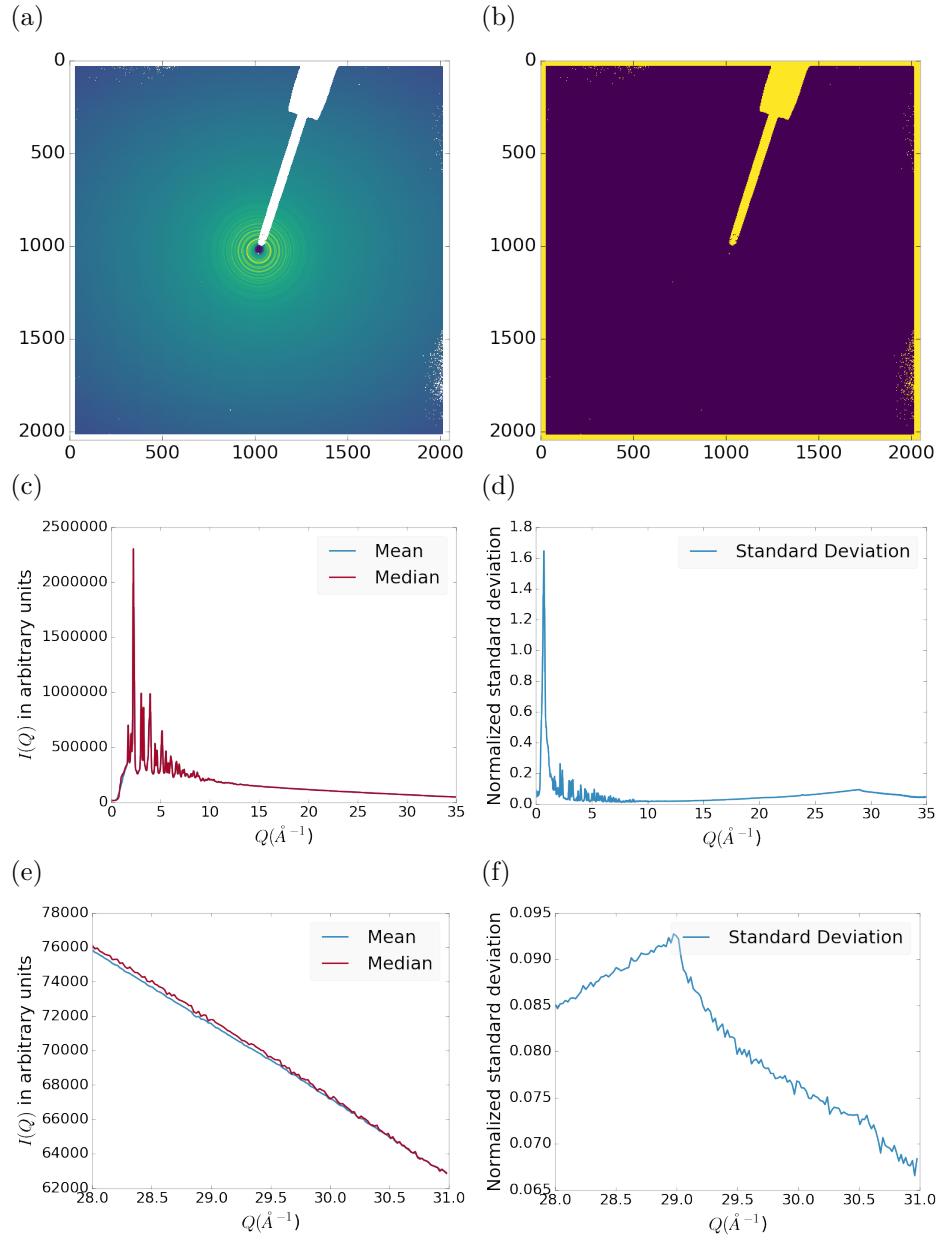


Figure 4.18: Masking, average, and standard deviation of an example x-ray total scattering measurement. This image was produced combining an edge mask and the automatically generated mask. a) the image, b) the mask, c) the mean and median values, d) the standard deviation (normalized to the median),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

696

CHAPTER 5

697

ANNEALING AND AGGREGATION OF 2NM

698

AU NANOPARTICLES

699 5.1 EXPERIMENTS

700 NP Synthesis

701 X-ray Total Scattering Measurements

702 5.2 DATA PROCESSING

703 5.3 DATA ANALYSIS

704 5.4 SIMULATION

705 5.5 STRUCTURAL ANALYSIS

706 5.6 CONCLUSIONS

707

CHAPTER 6

708

PHASE CHANGES AND ANNEALING DYNAMICS OF

709

Pr_2NiO_4 AND ITS DERIVATIVES

710 6.1 EXPERIMENTS

711 **Pr_2NiO_4 Synthesis**

712 **X-ray Measurements**

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

718 6.2 DATA PROCESSING

719

masking parameters

720

integration parameters

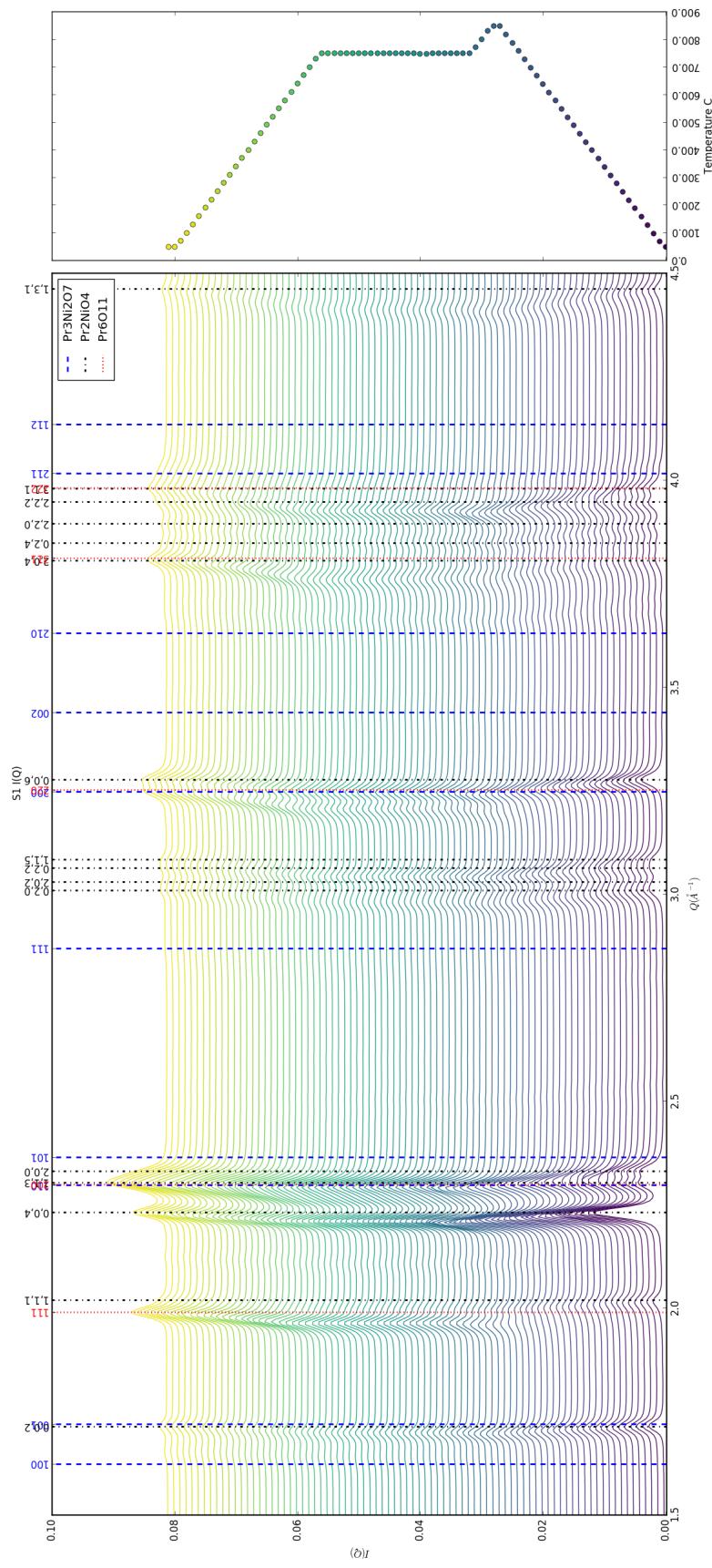
721

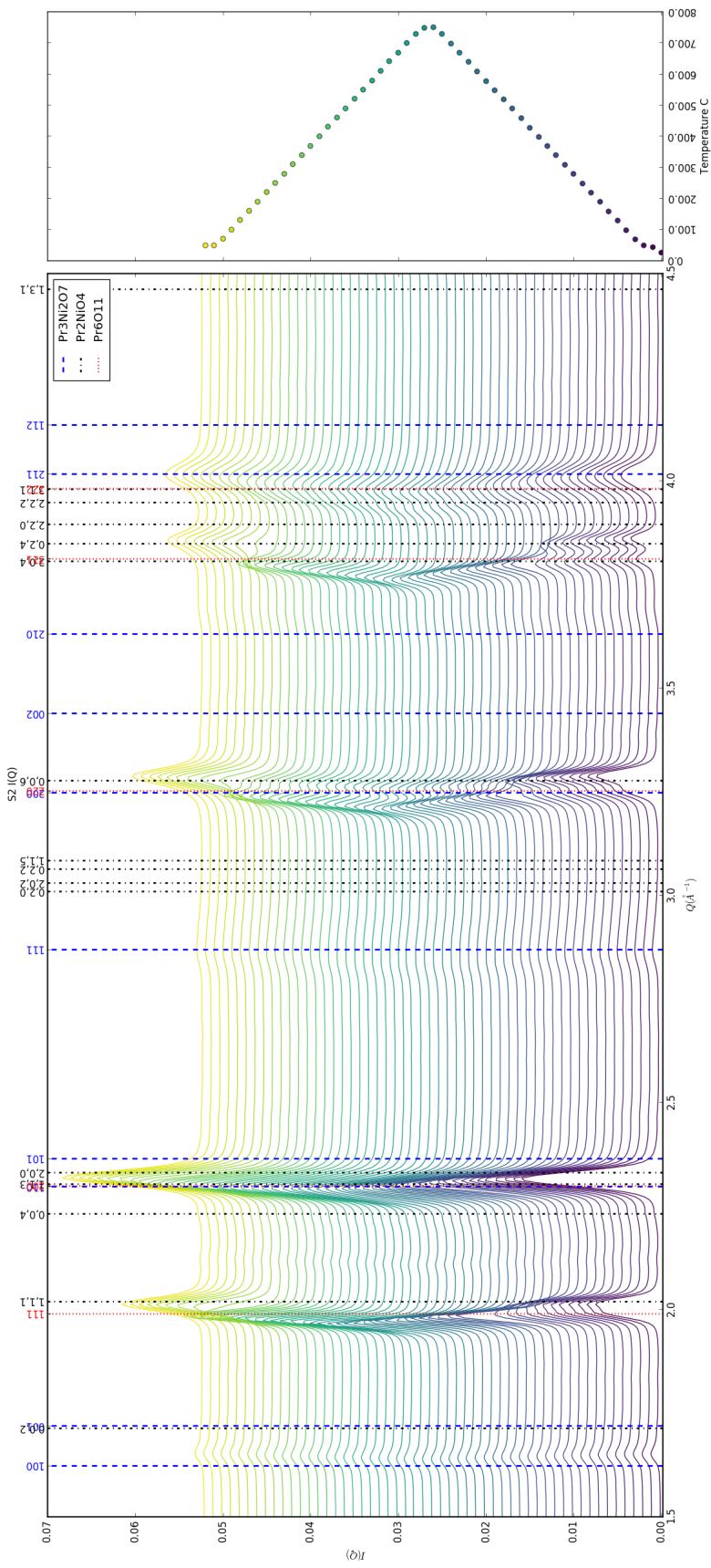
PDF parameters

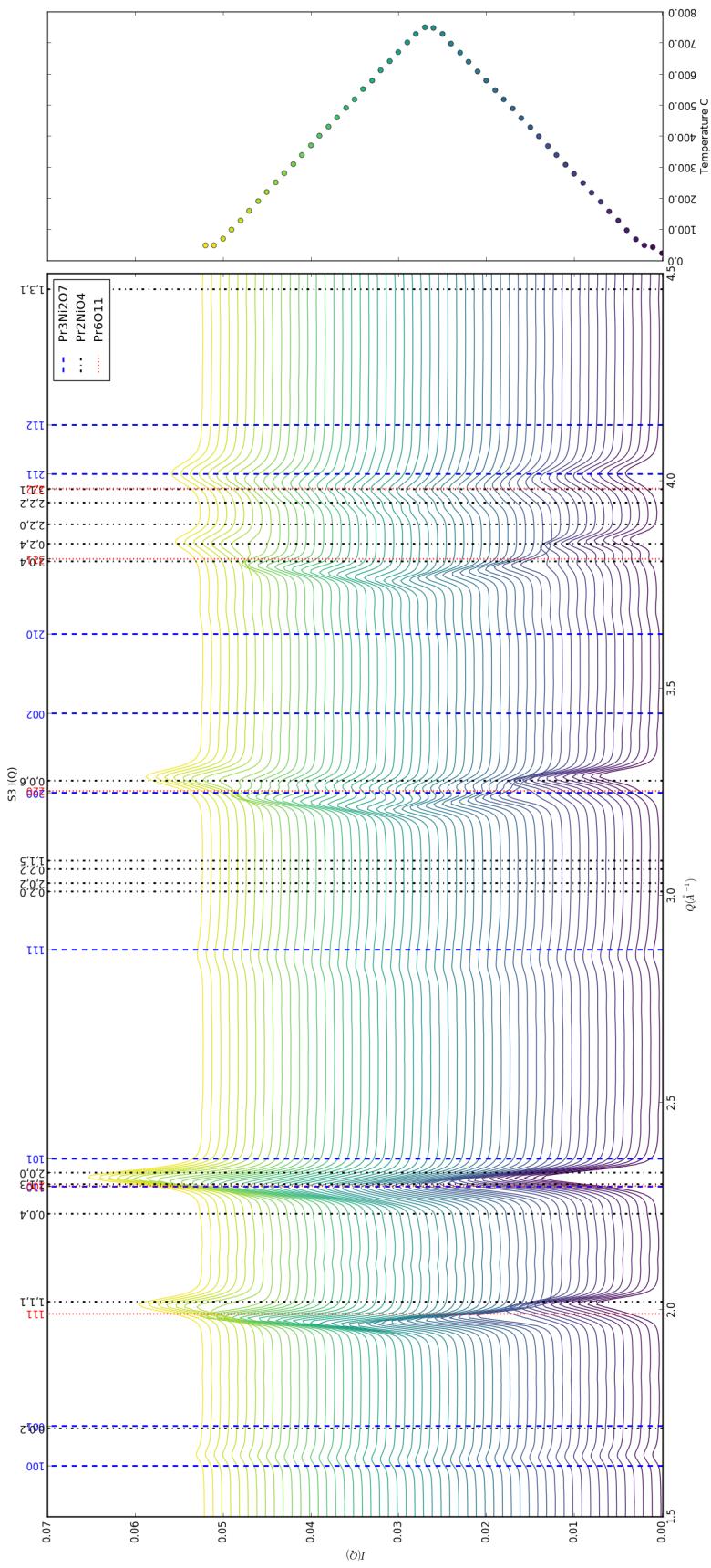
722 6.3 DATA ANALYSIS

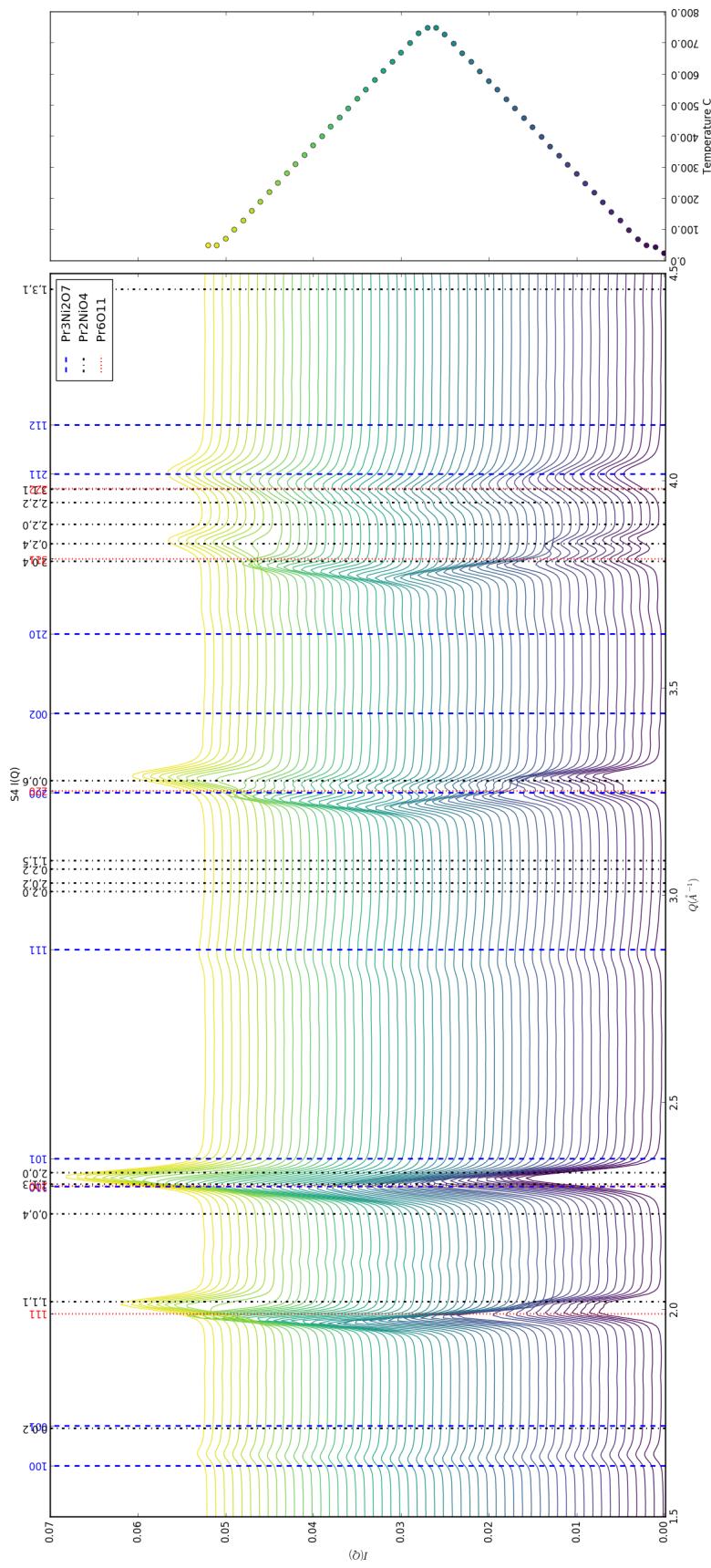
723 **Intra Sample Comparison**

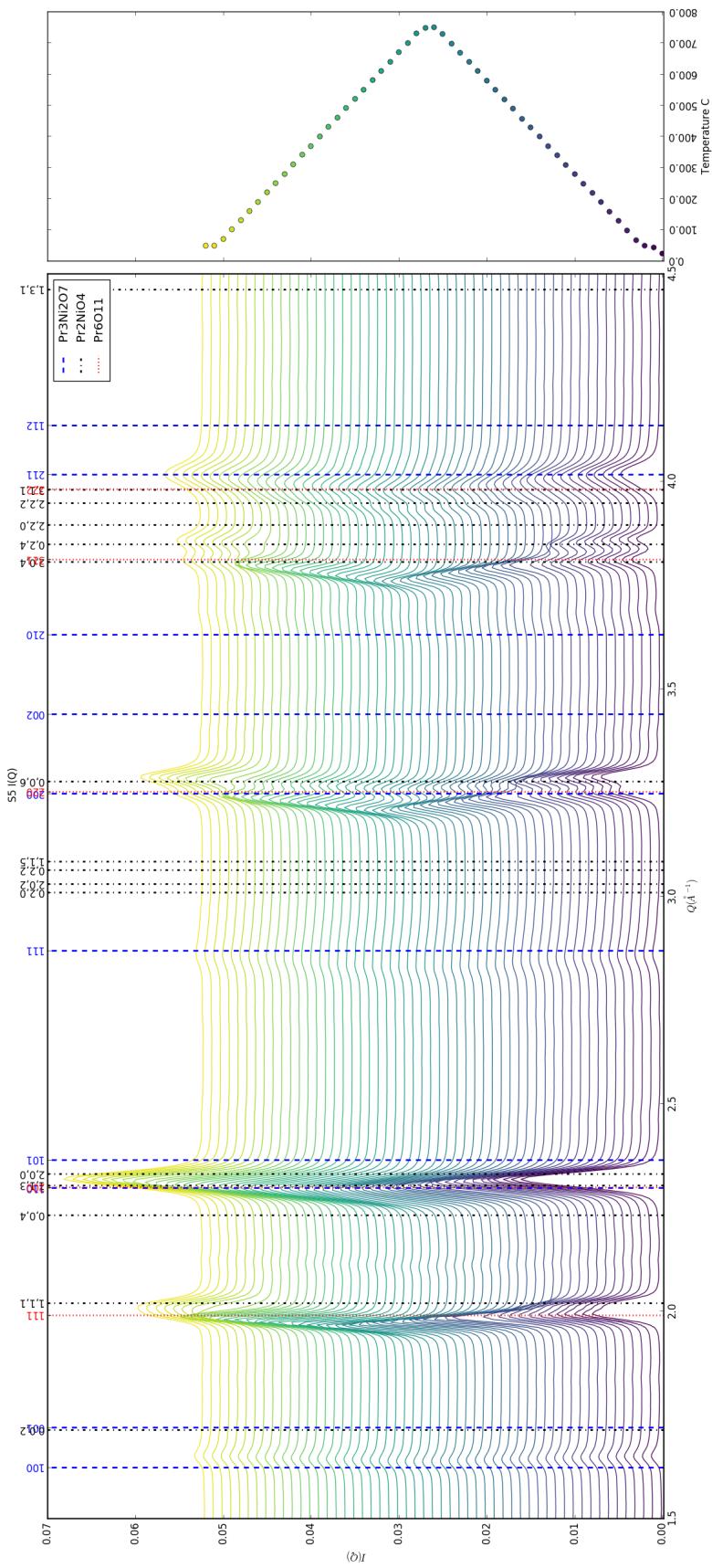
724 Changes in S1 but very little in S2-5.



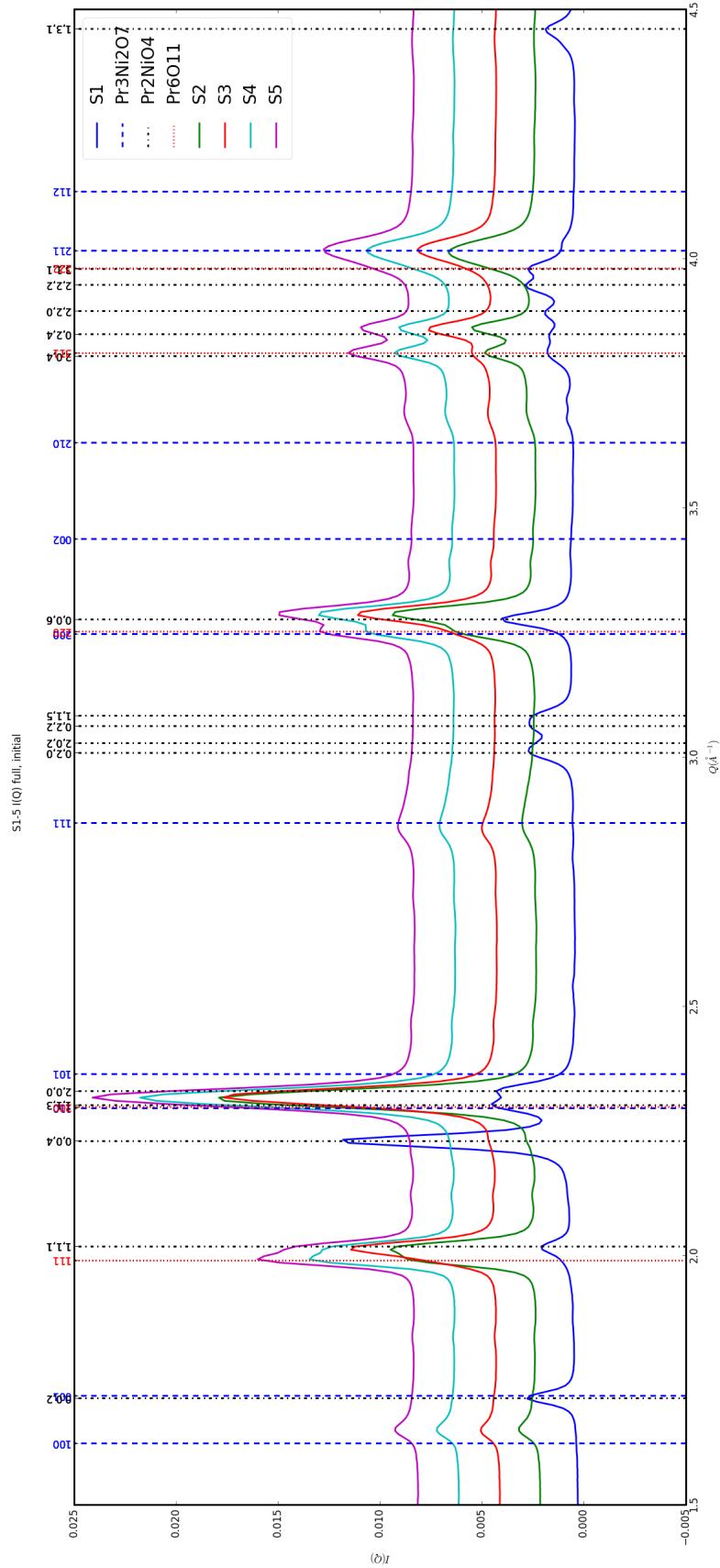


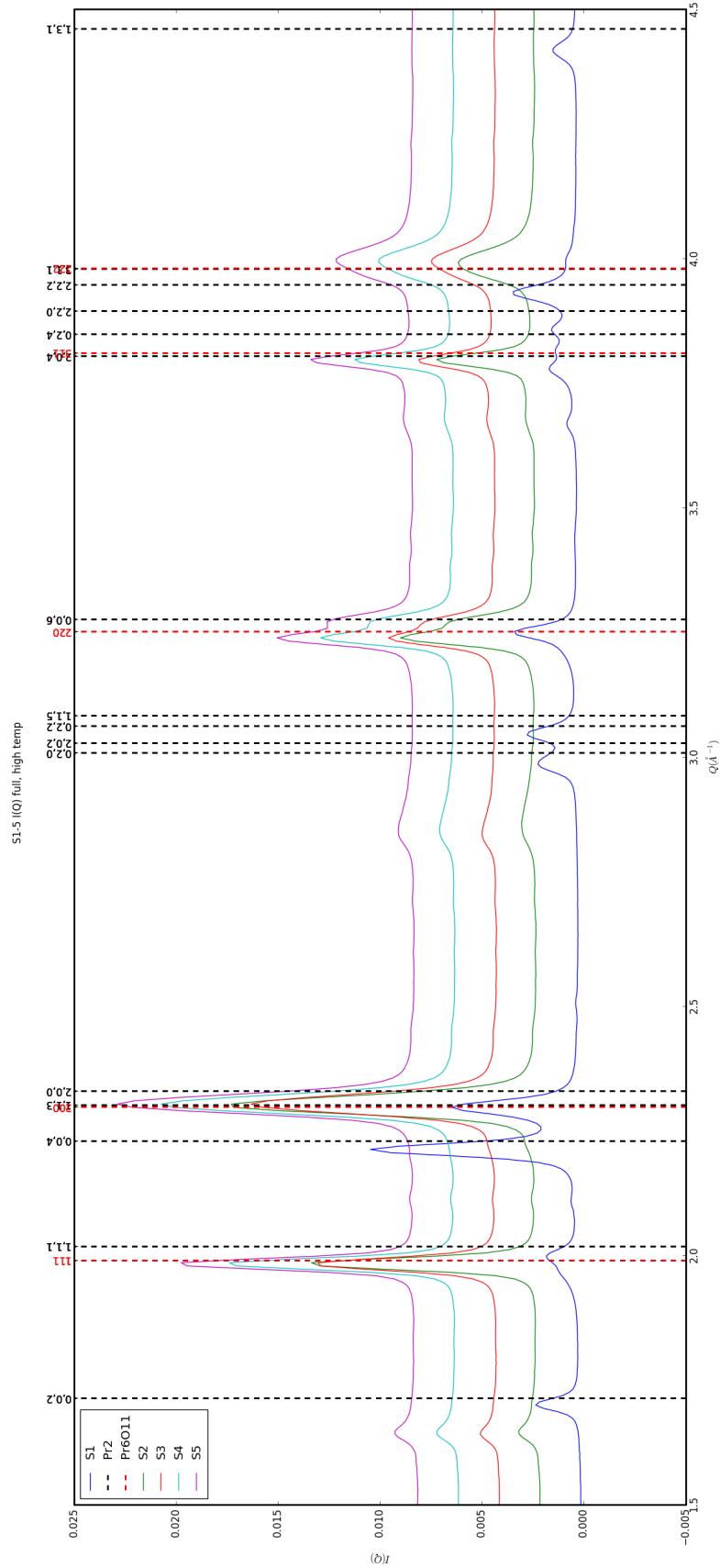


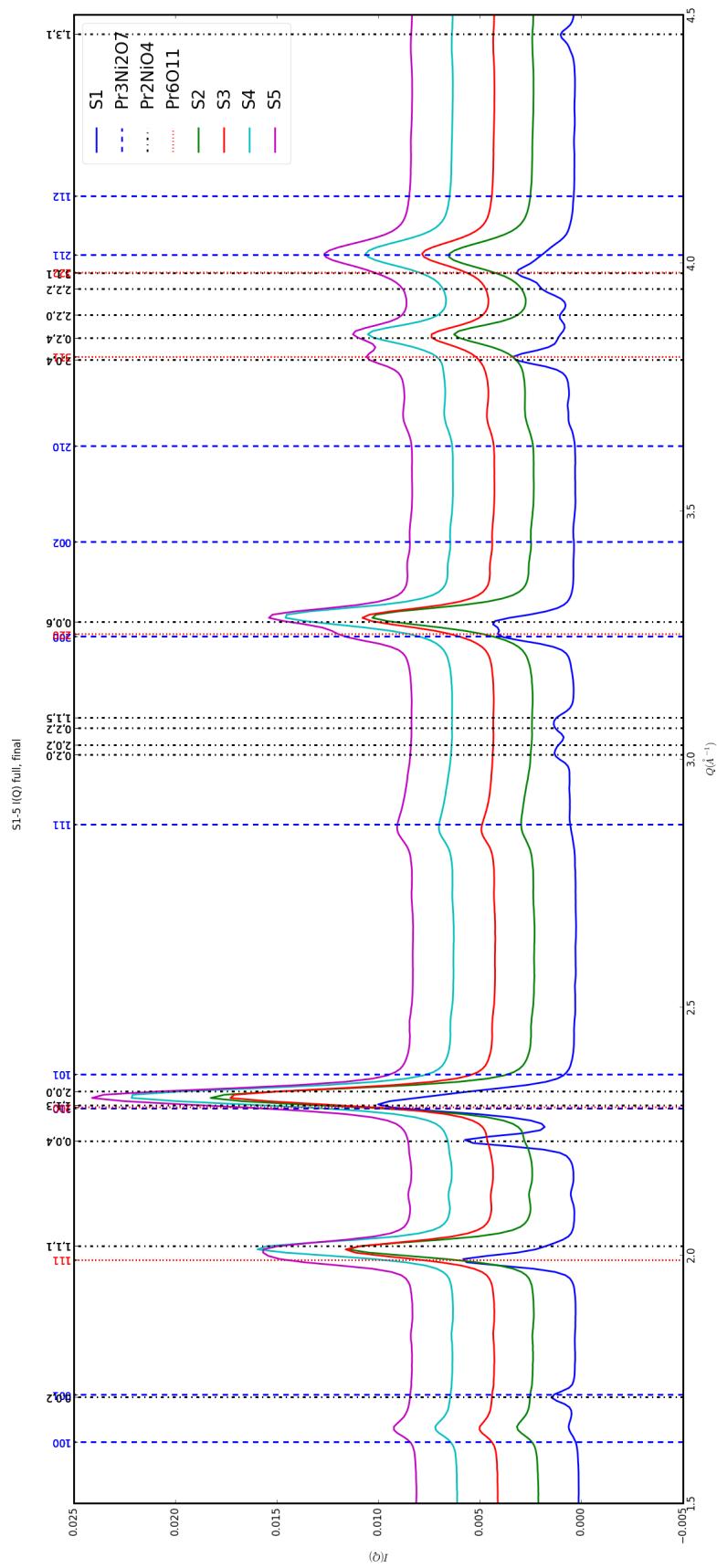


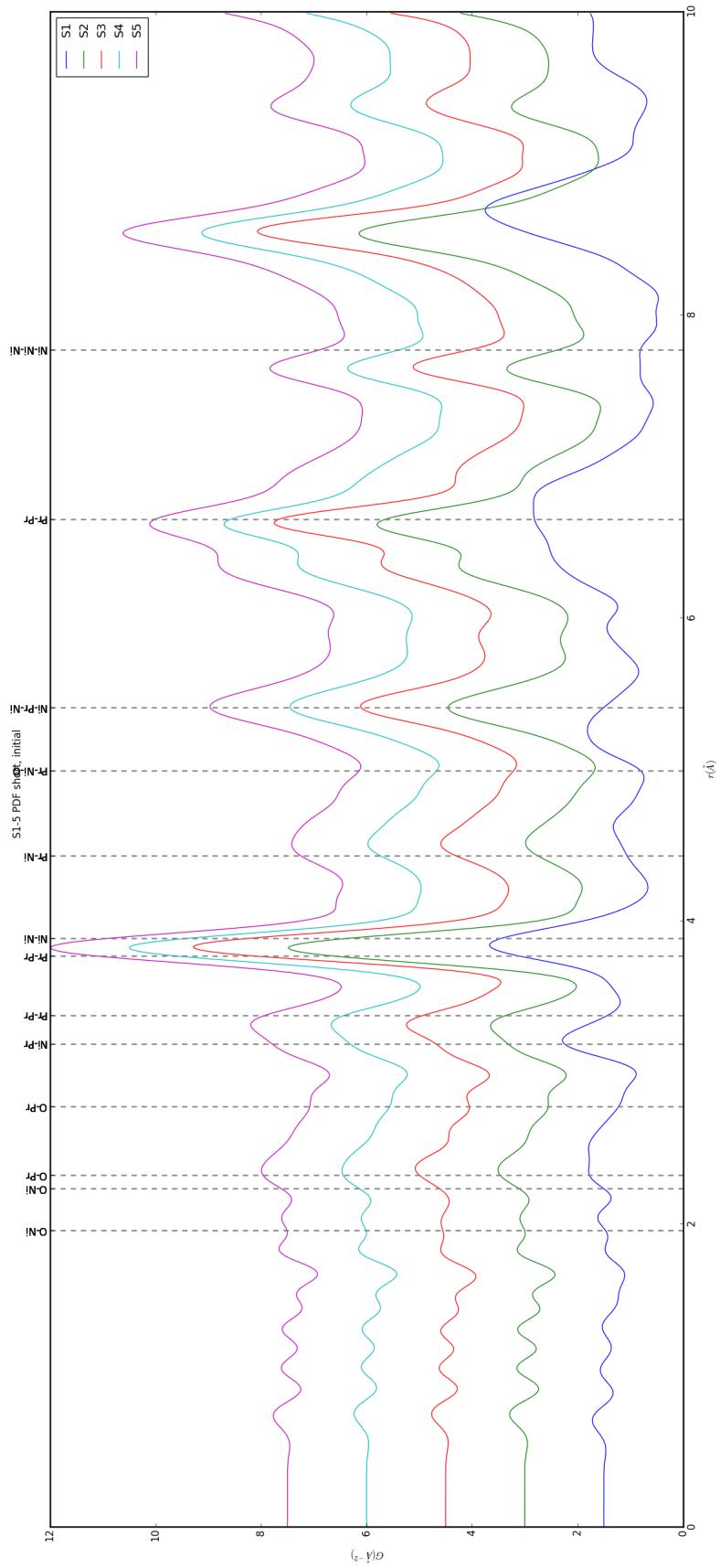


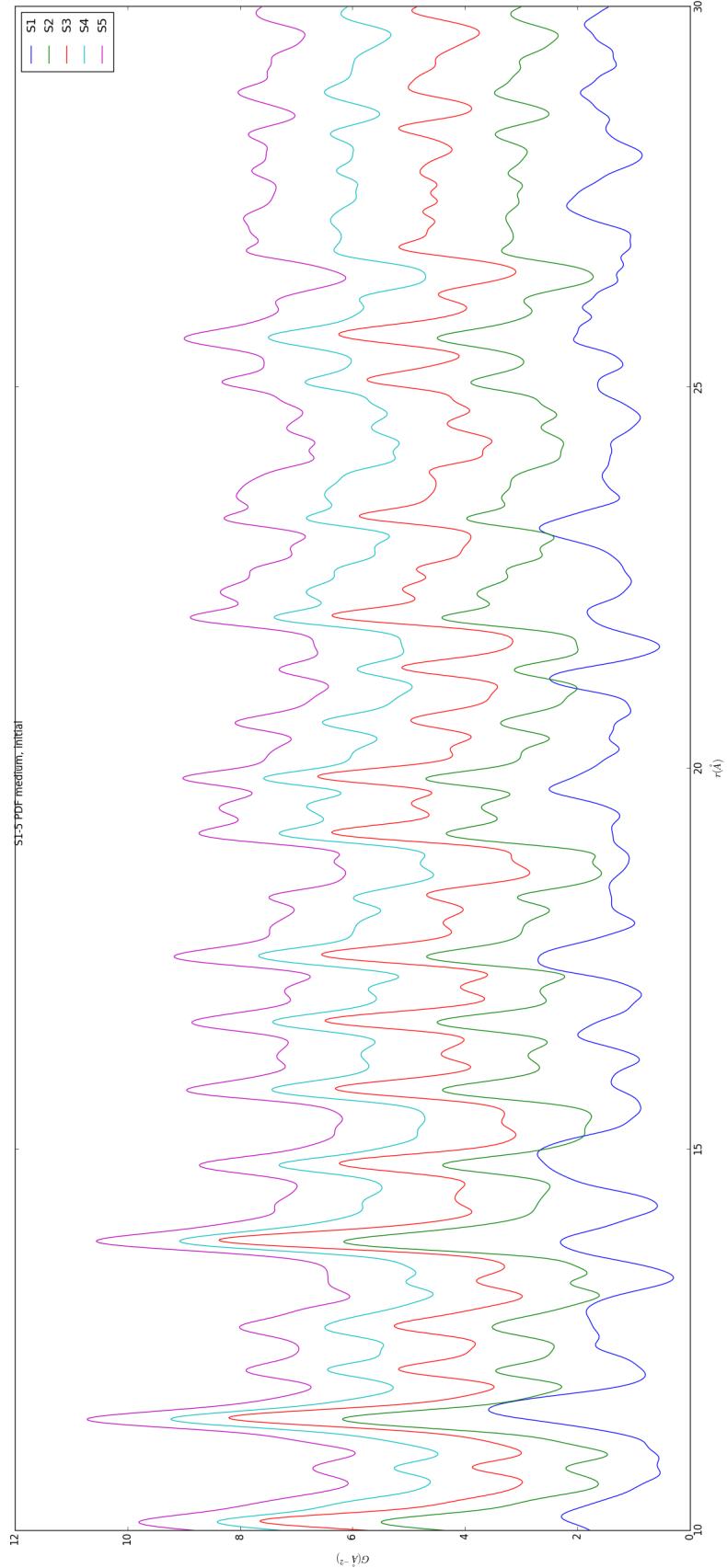
725 **Inter Sample Comparison**

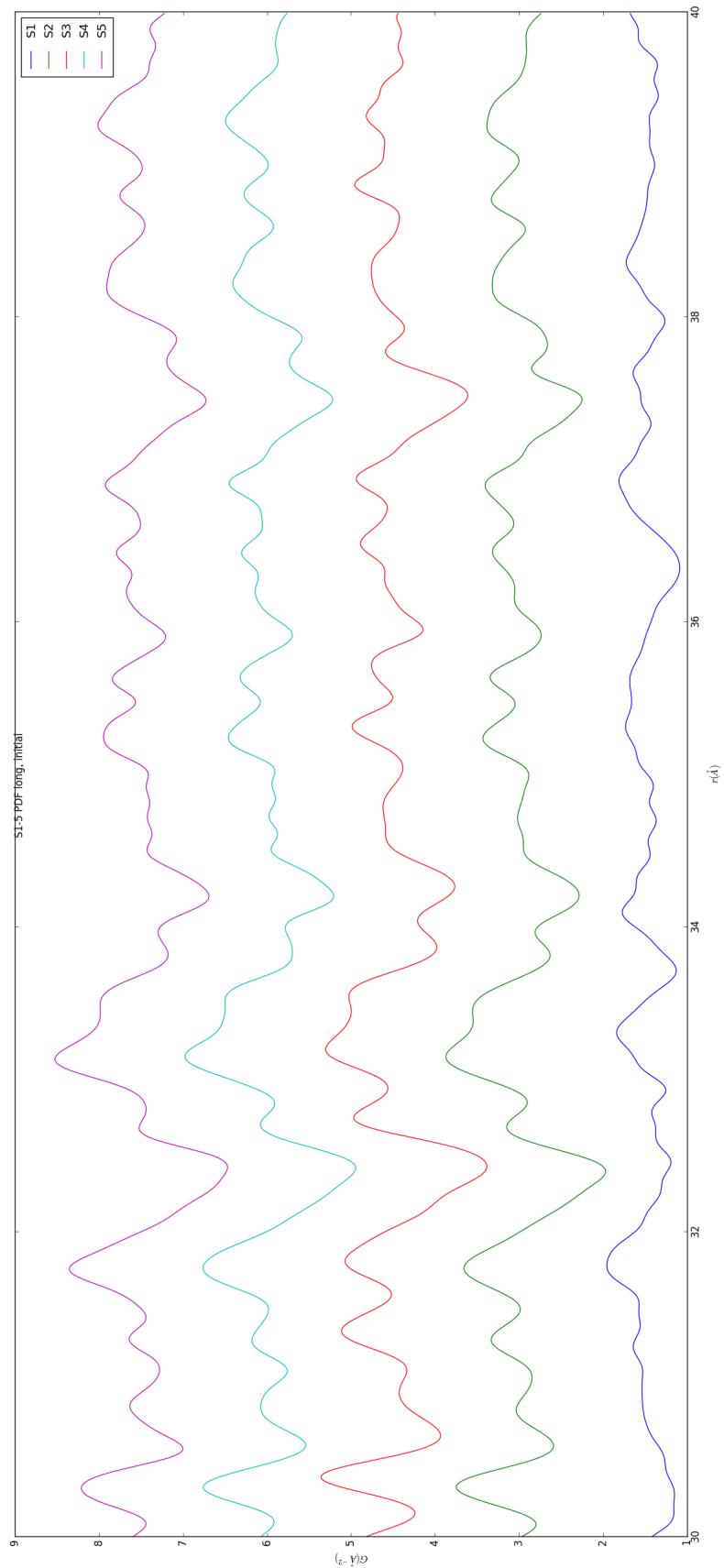


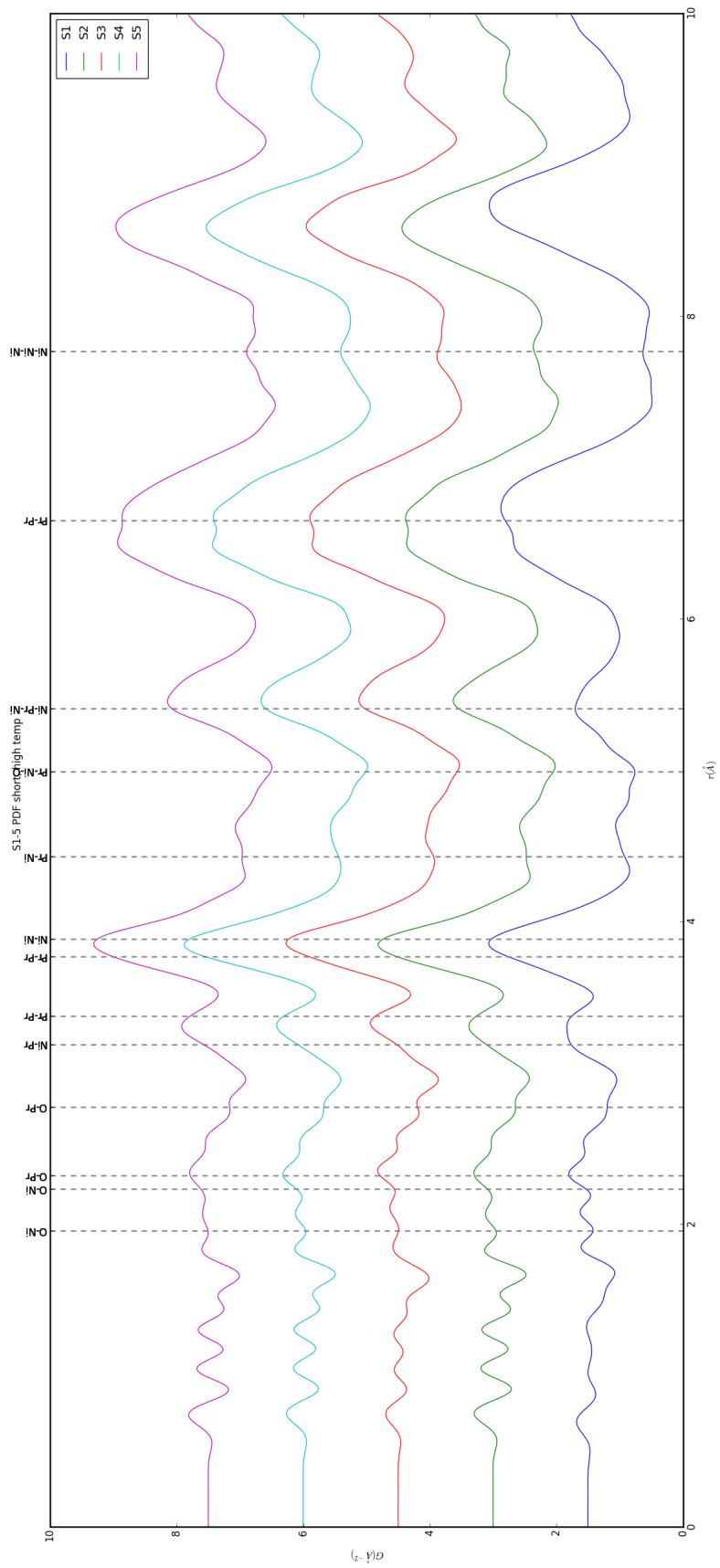


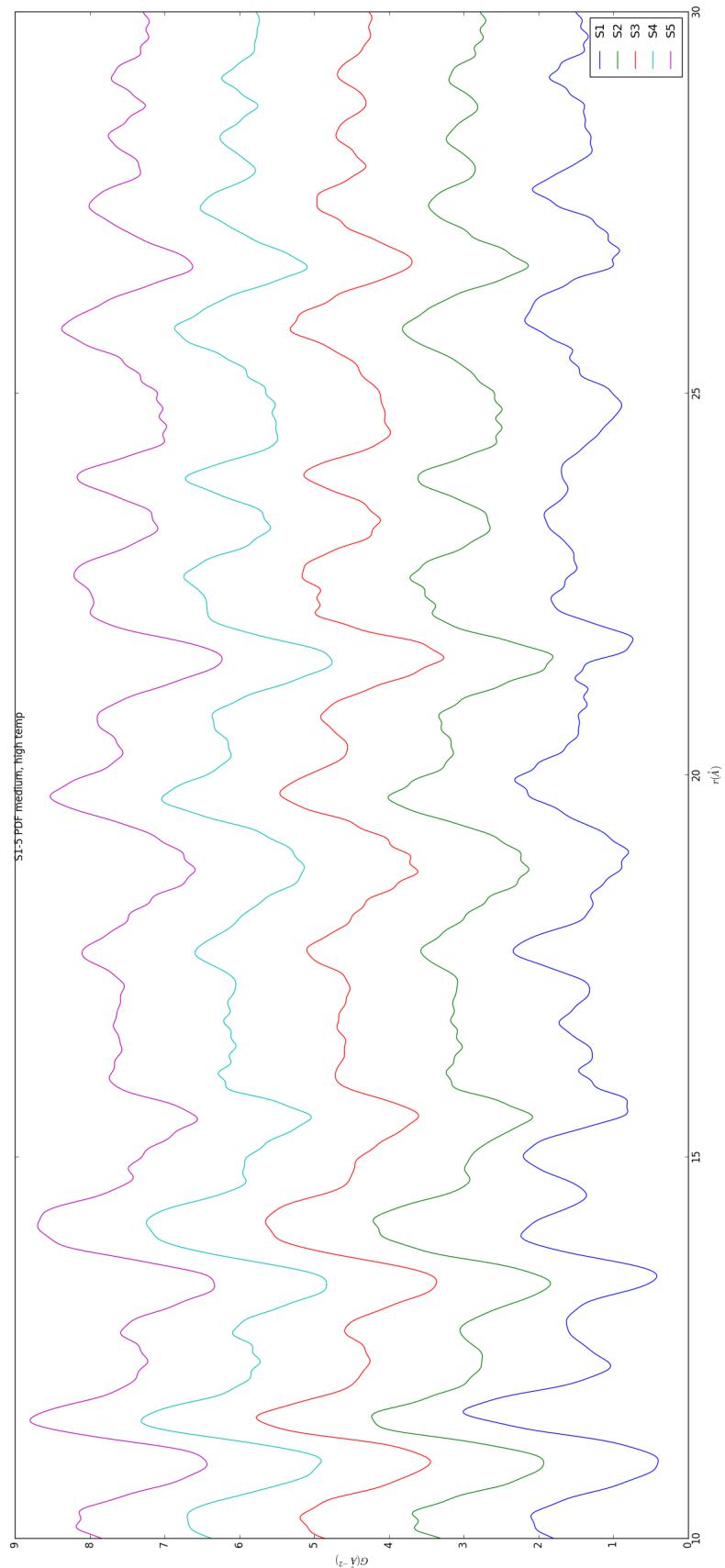


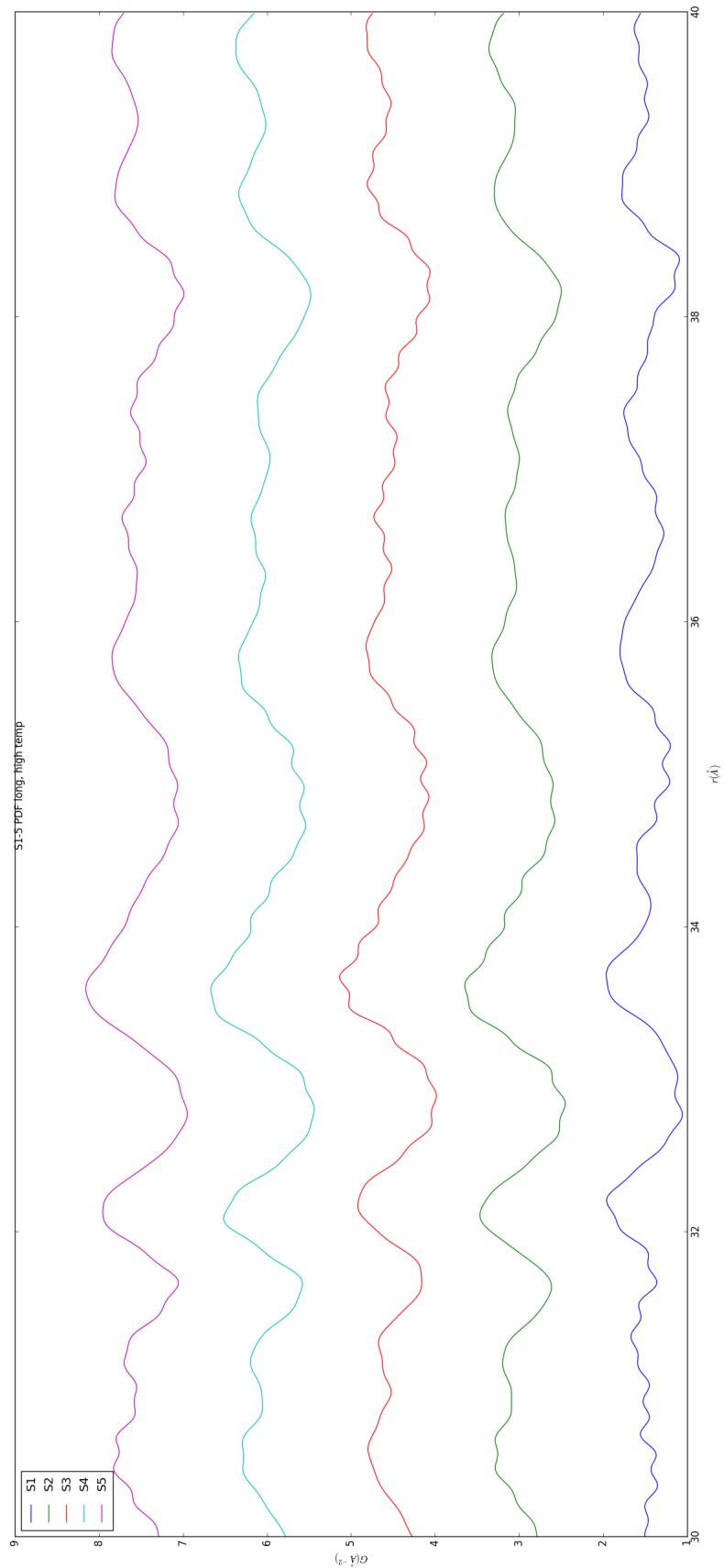


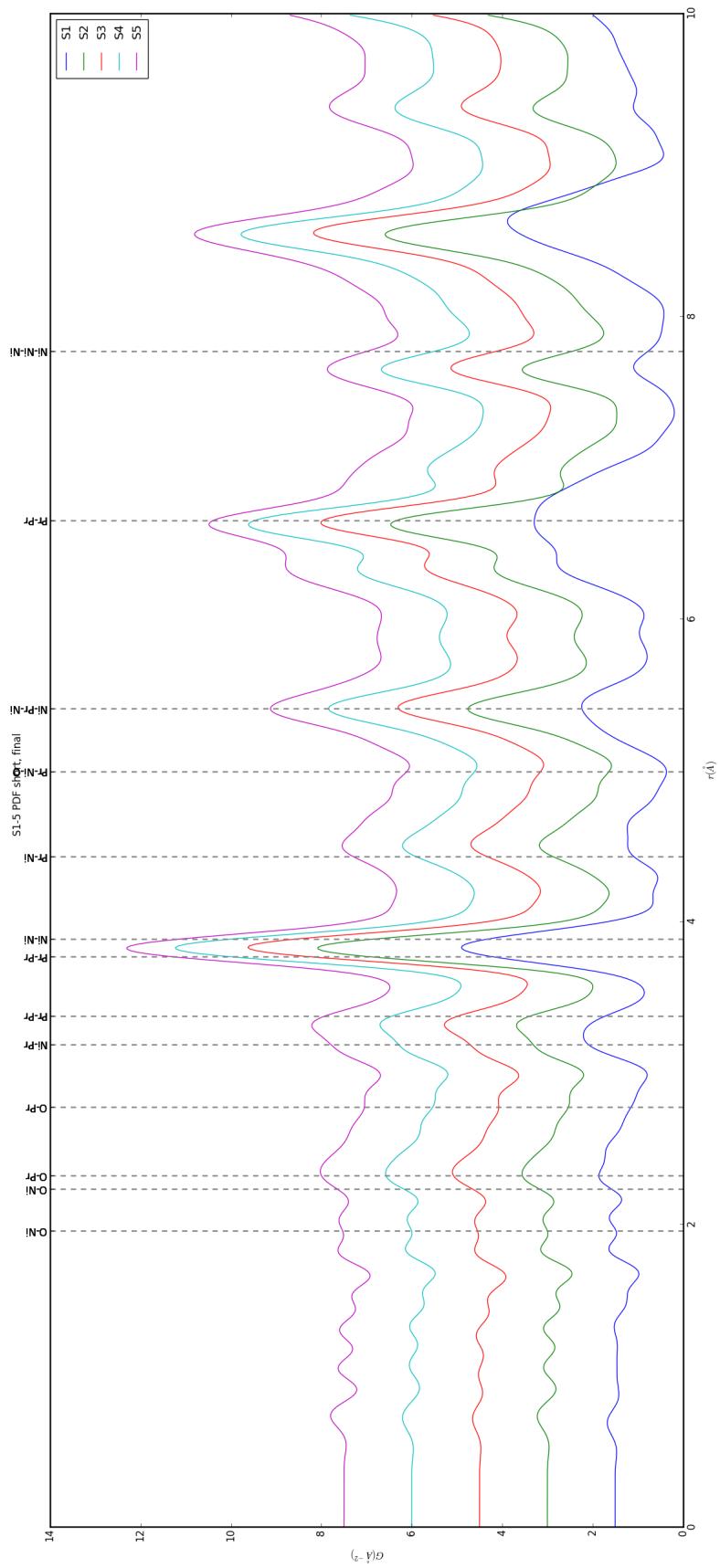


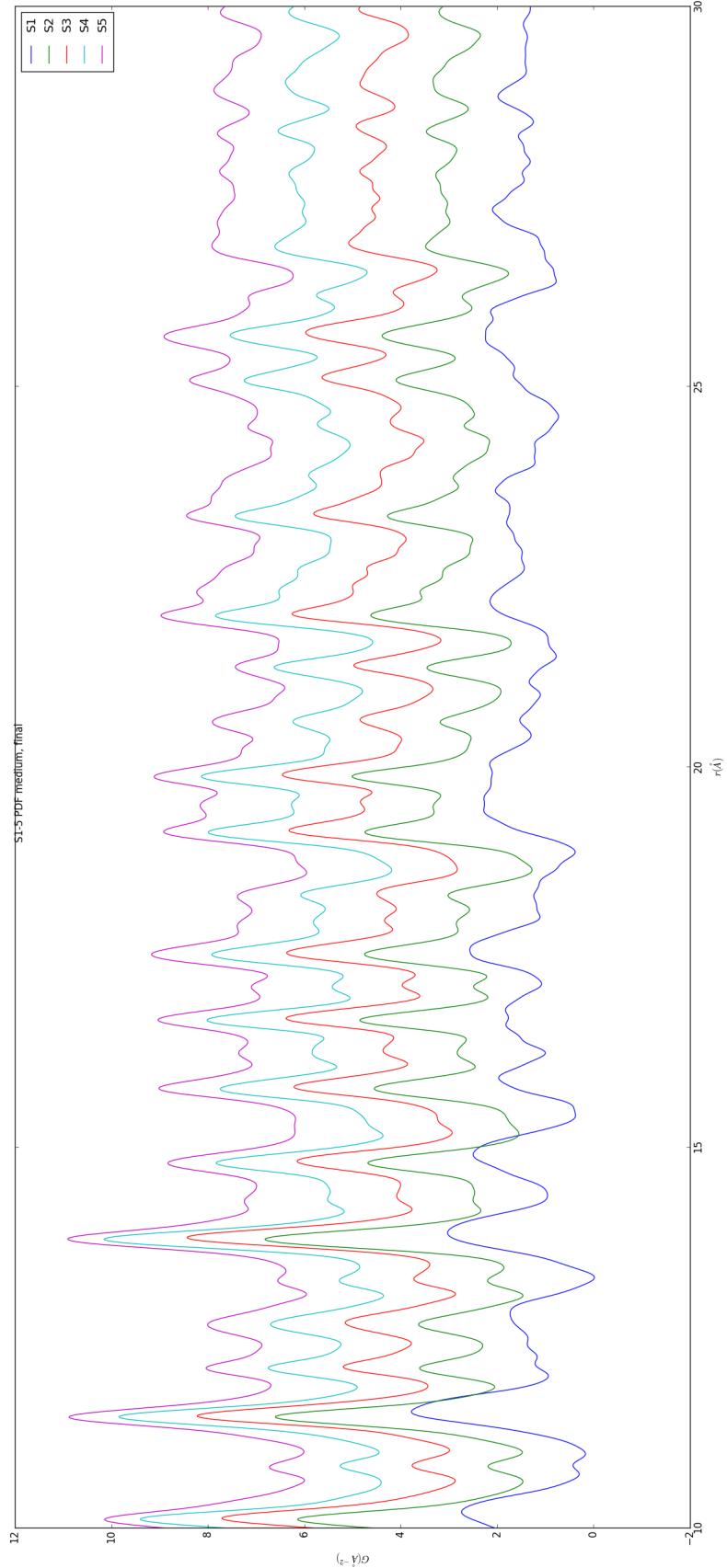


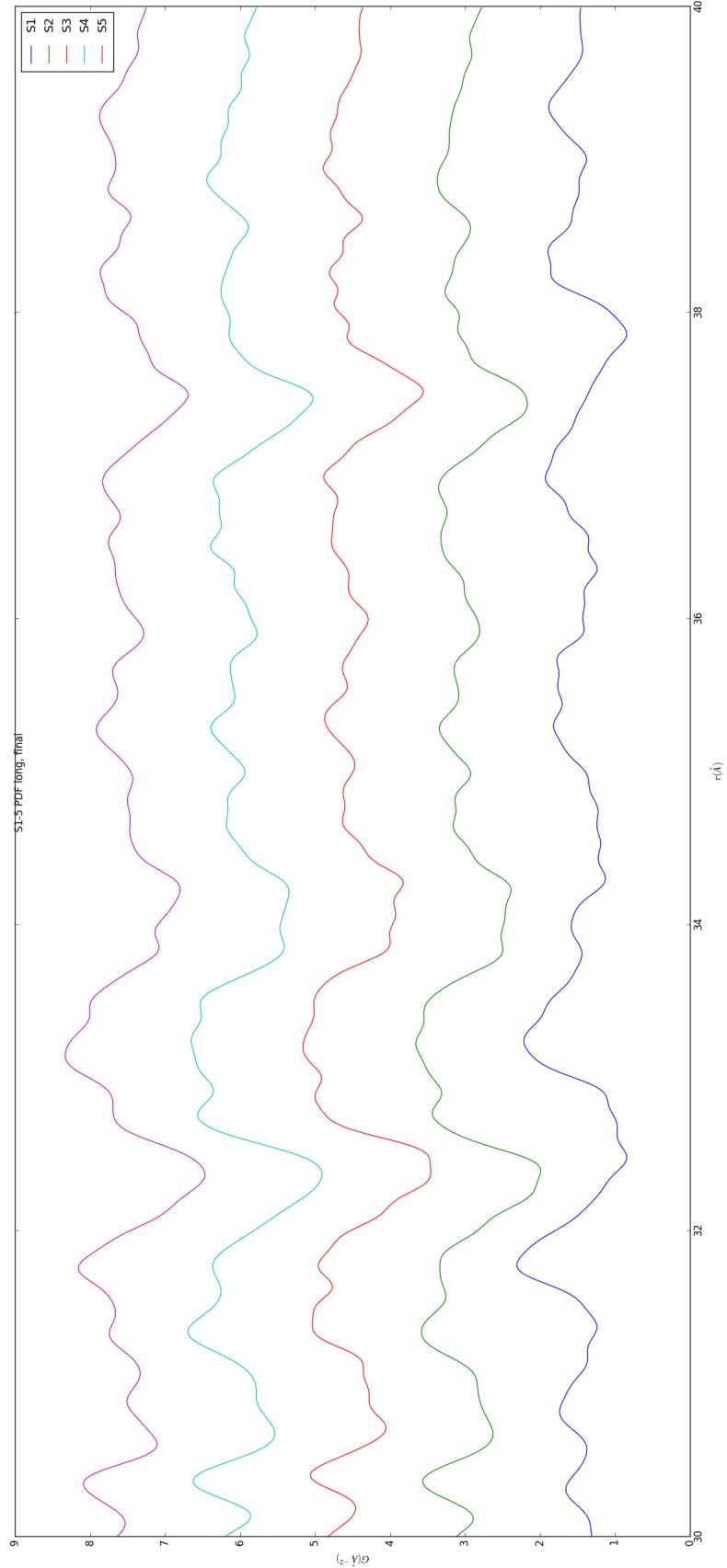












726 The PDF for S1 at operating temperature looks like S2-5 at the same temperature.

727 6.4 SIMULATION

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

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

740 6.5 CONCLUSIONS

741 X-ray total scattering and x-ray powder diffraction data was obtained on Pr_2NiO_4
742 powder samples annealed for various lengths of time. In-situ studies on the beamline
743 were performed to understand how the structure of each of these powders changes
744 at operating temperatures. The data was processed with the previously discussed Q
745 binning, masking, and integration methodology. The PDF results show very little
746 change in the structure for the as synthesized sample. However, the PDFs show a
747 large change in the previously annealed samples. These changes seem to reporduce
748 PDFs similar to the as-synthesized PNO at operating temperatures. This would seem
749 to imply that the source of the anamolus PNO phase/power density relationship may

750 be due to the adoption of an active structure upon heating which is universal despite
751 the amount of thermal degradation observed at room temperature. In contrast to the
752 PDF results, the XRD results seem to show signifigant changes in the PNO structure,
753 both with ex-situ and in-situ annealing. The XRDs show the degradation of the PNO
754 into various phases, potentially including Pr_2O_{11} , and higher ordered Pr based phases.
755 The discrepancy between these two results is quite interesting as it seems that the
756 XRD and PDF results are contradictory. Turbostratic diplacements between the
757 layers may be a cause of the PDF/XRD disagreement, as these changes would cause
758 very little change in the local structure observed in the PDF, while causing large
759 changes in the XRD.

760

CHAPTER 7

761

CONCLUSION

762

BIBLIOGRAPHY

- 763 [1] Simon J L Billinge and Takeshi Egami, *Underneath the Bragg Peaks: Structural
764 Analysis of complex Materials*, vol. Volume 16, Pergamon, 2012.
- 765 [2] Peter J. Chupas, Karena W. Chapman, Charles Kurtz, Jonathan C. Hanson,
766 Peter L. Lee, and Clare P. Grey, *A versatile sample-environment cell for non-
767 ambient X-ray scattering experiments*, Journal of Applied Crystallography **41**
768 (2008), no. 4, 822–824.
- 769 [3] PJ Chupas, X Qiu, JC Hanson, PL Lee, CP Grey, and SJL Billinge, *Rapid-
770 acquisition pair distribution function (RA-PDF) analysis*, Journal of Applied
771 Crystallography **36** (2003), 1342–1347.
- 772 [4] Timur Dykhne, Ryan Taylor, Alastair Florence, and Simon J L Billinge, *Data
773 requirements for the reliable use of atomic pair distribution functions in amor-
774 phous pharmaceutical fingerprinting.*, Pharmaceutical research **28** (2011), no. 5,
775 1041–8.
- 776 [5] P. Juhás, T. Davis, C. Farrow, and S. Billinge, *PDFgetX3 : a
777 rapid and highly automatable program for processing powder diffraction data into
778 total scattering pair distribution functions*, Journal of Applied Crystallography
779 **46** (2013), no. 2, 560–566 (en).
- 780 [6] Jérôme Kieffer and Dimitrios Karkoulis, *PyFAI, a versatile library for azimuthal
781 regrouping*, Journal of Physics: Conference Series **425** (2013), 202012.
- 782 [7] B R Pauw, *Corrigendum: Everything SAXS: small-angle scattering pattern col-
783 lection and correction (2013 J. Phys.: Condens. Matter 25 383201)*, Journal of
784 Physics: Condensed Matter **26** (2014), no. 23, 239501.
- 785 [8] Valeri Petkov, Shiyao Shan, Peter Chupas, Jun Yin, Lefu Yang, Jin Luo, and
786 Chuan-Jian Zhong, *Noble-transition metal nanoparticle breathing in a reactive
787 gas atmosphere.*, Nanoscale **5** (2013), no. 16, 7379–87.
- 788 [9] Erin L. Redmond, Brian P. Setzler, Pavol Juhas, Simon J. L. Billinge, and
789 Thomas F. Fuller, *In-Situ Monitoring of Particle Growth at PEMFC Cathode*

790 under Accelerated Cycling Conditions, Electrochemical and Solid-State Letters
791 **15** (2012), no. 5, B72 (en).

792 [10] X. Yang, P. Juhás, and S. J L Billinge, *On the estimation of statistical uncertain-
793 ties on powder diffraction and small-angle scattering data from two-dimensional
794 X-ray detectors*, Journal of Applied Crystallography **47** (2014), no. 4, 1273–1283.