

SOLVING ATOMIC STRUCTURES 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

Mark Uline, Committee Member

Jochen Lauterbach, Committee Member

Thomas Vogt, Committee Member

Lacy Ford, Vice Provost and Dean of Graduate Studies

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

DEDICATION

For Diane & Donald Wright

My first scientific advisers

ACKNOWLEDGMENTS

ABSTRACT

Engineering the next generation of materials, especially nanomaterials, requires a detailed understanding of the material's underlying atomic structure. Understanding these structures we can obtain a better understanding of the structure-property relationship, allowing for a property driven rational design of the material structure on the atomic level. Even more importantly, understanding the structure in-situ will allow the reversal of the flow of information, translating stimuli and responses on the macroscopic system level to changes in the atomic structure. Despite the importance of atomic structures for designing materials, solving the atomic structure of materials is difficult both experimentally and computationally. Atomic pair distribution functions (PDFs) have been shown to provide information on atomic structure, although extracting the PDF from x-ray total scattering measurements can be difficult. Computationally, translating the PDF to an atomic structure requires the search of a very high dimensional space and can be computationally expensive.

This work aims to address these issues by developing 1) novel statistical mechanical approaches to solving material structures, 2) fast simulation of x-ray total scattering and atomic pair distribution functions (PDFs), and 3) data processing procedures for experimental x-ray total scattering measurements. First, experimentally derived potential energy surfaces (PES) and the statistical mechanical ensembles used to search them are developed.

Then the mathematical and computational framework for the PDF and its gradients will be discussed. The combined PDF-PES-ensemble system will be benchmarked against a series of nanoparticle structures to ascertain the efficiency and ef-

fectiveness of the system. Experimental data processing procedures, which maximize the usable data, will be presented. Finally, preliminary results from experimental x-ray total scattering measurements will be discussed. This work presents one of the most complete end-to-end systems for processing and modeling x-ray total scattering PDF data, potentially allowing for high-throughput structural solution.

TABLE OF CONTENTS

DEDICATION	iii
ACKNOWLEDGMENTS	iv
ABSTRACT	v
LIST OF TABLES	xii
LIST OF FIGURES	xii
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	11
Ensemble description	11
Grand Canonical Monte Carlo	11
GCMC biasing	12

CHAPTER 2 ATOMIC PAIR DISTRIBUTION FUNCTION: THEORY AND COMPUTATION	14
2.1 Theory	14
Derivation	14
Analytically 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	20
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	51
 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	26
Figure 3.2 Au ₅₅ PDF fitting of surface-disordered Au ₅₅	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 <i>Q</i> resolution as a function of <i>Q</i>	39
Figure 4.3 Number of pixels as a function of <i>Q</i> , binned at the <i>Q</i> resolution of the detector.	40
Figure 4.4 Generated dead/hot pixel masks for a detector with 100 bad pixels.	42
Figure 4.5 Generated dead/hot pixel masks for a detector with 300 bad pixels.	43
Figure 4.6 Generated dead/hot pixel masks for a detector with 500 bad pixels.	43
Figure 4.7 Generated dead/hot pixel masks for a detector with 1000 bad pixels.	44
Figure 4.8 Generated beamstop holder masks for a beamstop holder with 10% transmittance.	44
Figure 4.9 Generated beamstop holder masks for a beamstop holder with 30% transmittance.	45

Figure 4.10 Generated beamstop holder masks for a beamstop holder with 50% transmittance.	45
Figure 4.11 Generated beamstop holder masks for a beamstop holder with 90% transmittance.	45
Figure 4.12 Generated beamstop holder masks which is rotated away from vertical	46
Figure 4.13 Masked experimental data.	46
Figure 4.14 Masked experimental data with Pt single crystal signal.	47
Figure 4.15 Masked experimental data with Pt single crystal signal using figure's 4.13 mask as a starting mask.	47
Figure 4.16 Masking, average, and standard deviation of an example x-ray total scattering measurement with with no mask.	50
Figure 4.17 Masking, average, and standard deviation of an example x-ray total scattering measurement with with only an edge mask.	51
Figure 4.18 Masking, average, and standard deviation of an example x-ray total scattering measurement with combining an edge mask and the automatically generated mask.	52

TODO LIST

2	Need more references	2
3	Why is atomistic engineering important	3
4	Barriers to atomistic engineering	3
5	How are we going to attack this problem	3
6	Explain the origin of the potentials, especially χ^2 and Rw	6
7	Need to discuss Monte Carlo more, including Reverse Monte Carlo	9
8	Need to put HMC into the Stat Mech Formalism, it is kinda canonical, except that it has a random momenta which is specified and not temperature.	9
10	Need the official criteria for HMC	10
11	Talk about the advances NUTS gives us	10
12	What limitations in RMC/HMC/NUTS/Cannonical ensembles in general force us to move to GCMC	11
14	Include figure which shows the configurational biasing map	13
15	cite something with debye waller most likely simon's paper	14
16	Need to include PBC ADP math	16
17	Also should include PBC gradients, although they are trivial, maybe?	17
18	How does this compare against the Ewald simulation technique for ionic solutions	17
19	use the figure from the presentation, it is better	17
20	I wish we had a picture of the GPU profiling	17
21	Include Speed Benchmarks Here	21
22	This entire section needs some rewritting to distinguish this from the paper .	22
23	Put this somewhere	26

24	masking parameters	55
25	integration parameters	55
26	PDF parameters	55
27	Need more references	

28

INTRODUCTION

29

Why is atomistic engineering important

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

34

Barriers to atomistic engineering

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

38

How are we going to attack this problem

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

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

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

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

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

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

58

CHAPTER 1

59

STATISTICAL MECHANICAL ENSEMBLES AND POTENTIAL ENERGY SURFACES

61 1.1 INTRODUCTION

62 The approach taken in this work for solving the atomic structures of materials is one
63 of optimization. The plan is to develop a potential energy surface (PES) which has
64 minima associated with atomic structures who's properties match the experimentally
65 observed properties. Thus, the various positional variables of the structure can be
66 solved by optimizing the structure against the PES. This approach is popular in the
67 PDF community for solving the structure of materials using both extensive large box
68 models and simpler small box models.

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

74 1.2 POTENTIAL ENERGY SURFACES

75 A PES simply describes the potential energy of the system as a function of all its
76 relevant coordinates in phase space, essentially providing a mapping $\mathbb{R}^n \rightarrow \mathbb{R}$, where \mathbb{R}
77 is the set of real numbers and n is the number of positional parameters in the system.
78 Usually these coordinates are the positions of the atoms q and their conjugate the

79 momenta p . Note that there could be more variables associated with the system,
80 for instance the magnetic moments of the atoms could play a role in describing the
81 system. In this magnetic system there would be positional variables for the atomwise
82 spin vectors and their "momenta". Application of the term "momenta" might seem
83 odd here, as the magnetic spin does not have a mass or a velocity. However, since the
84 magnetic "position" is defined on the PES we need to describe its conjugate variable
85 to properly formulate Hamiltonian dynamics and the kinetic portion of the PES.

86 Experimentally Derived Potential Energy Surfaces

87 Explain the origin of the potentials, especially χ^2 and Rw

88 Generally PESs are obtained from purely computational experiments including:
89 ab-initio DFT, classical approximations via the embedded atom method, or even
90 parameter driven models with experimentally fitted parameters. However, one can
91 derive a PES from an experiment which describes how well the model reproduces the
92 experimental data. In this case one needs a theoretical and computational framework
93 mapping the atomistic variables of the simulation to the same space of the data
94 obtained from the experiment. This allows the experiment to be compared directly
95 against the predicted data via an experimentally derived PES.

96 Potentials

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

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

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

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

101

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

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

105 The Rw and χ^2 potentials have been reported numerous times. [?] However, the
 106 INVERT potential is fairly new and aims to incorporate descriptions of the structural
 107 symmetry into the PES. [2, 3] In the case of the INVERT potential NMR or other
 108 symmetry sensitive data is used to describe the number of unique atomic coordina-
 109 tions. This is then used to describe the number of unique atomwise pair distribution
 110 functions, thus causing systems with more or less unique coordination environments
 111 to be higher in energy. This approach has been shown to be useful for C_{60} and other
 112 systems which are highly symmetric, creating a PES with an easier to find minima.
 113 [2, 3] However, many times this kind of data is unavailable when refining the struc-
 114 ture causing the potential to be less useful. Additionally, this potential introduces
 115 an element of user bias as the refiner must decide, based on some spectroscopic data,
 116 how many unique environments are in the material. This bias could be removed by
 117 using one of the other potentials with a method for simulating the observed spectra,
 118 allowing the computational system decide what structures properly reproduce all the
 119 observed data.

120 **Forces**

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

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

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

123

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

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

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

130 1.3 ENSEMBLES

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

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

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

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

142 Analyticly one generally defines a partition function, which describes the sum of
 143 probabilities over all potential atomic states.

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

144 where P_i is the probability of the i th state and is a function of the total energy of
145 that state. This partition function can then be used to obtain the probability of any
146 specific state. The relationship of the probability of a state to the state's energy
147 and other properties depends on the ensemble being used. For the microcanonical
148 ensemble the probability of a state is:

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

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

151 However, for the canonical ensemble the probability is:

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

152 Monte Carlo Modeling

153 Monte Carlo can be used to simulate a statistical mechanical ensemble which can
154 not be solved analytically. In most Monte Carlo systems the ensemble is simulated by
155 randomly changing one of the system parameters and comparing the energy of the
156 new system against the energy of the old system. If the energy of the new system is
157 lower than the current energy then the new configuration is accepted. Otherwise the
158 new system is rejected unless

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

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

161 Need to discuss Monte Carlo more, including Reverse Monte Carlo

162 Hamiltonian Monte Carlo

163 Need to put HMC into the Stat Mech Formalism, it is kinda canonical, except
that it has a random momenta which is specified and not temperature.

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:

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

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

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

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

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

166 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.

169 Need the official criteria for HMC

170 **No-U-Turn-Sampling**

171 Talk about the advances NUTS gives us

172 **Grand Canonical Ensemble**

What limitations in RMC/HMC/NUTS/Cannonical ensembles in general force us
173 to move to GCMC

174 **Ensemble description**

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

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

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

181 **Grand Canonical Monte Carlo**

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

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

185 Similarly the removal of an atom has the probability:

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

186 However, both of these equations depend of the overall simulation volume and the
187 thermal wavelength, which is undesirable as these are not really properties that we
188 are of interest to these simulations. Thus, we roll them into the definition of the
189 chemical potential, essentially setting the base chemical potential to counteract these
190 effects. This makes certain that our simulation does not change if we change the

191 overall cell volume. A GCMC move consists of creating a new atomic configuration,
192 where an atom has been added or removed, and checking the above criteria. However,
193 previous results have shown that this method is computationally expensive in dense
194 liquids, and exceedingly expensive in solid materials. The long simulation times
195 are due to the random nature of the atomic additions or removals which produce:
196 over-tightly packed atoms, atoms in the middle of nowhere, or unphysical vacancies.
197 These configurations are rejected by the GCMC criteria but their probability of being
198 sampled is much higher than configurations which are lower in energy, since the
199 number of incorrect ways to add/remove atoms is much larger than the correct ways.
200 Thus we have implemented methods for biasing the atomic addition positions and
201 the atomic removals toward configurations which are more likely to be accepted.

202 GCMC biasing

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

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

213 where x, y, z and a, b, c are the resolutions and cell side lengths in the cardinal di-
214 rections, respectively. While this does help to limit the total probability space it
215 does not tell us which voxels are likely to lead to better configurations, leading to

216 many rejected atomic additions. To combat this issue we can weigh the individual
217 voxels, giving more probability to voxels which show promise and less to those with
218 less likelihood to be accepted.

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

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

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

227 Include figure which shows the configurational biasing map

228

CHAPTER 2

229

ATOMIC PAIR DISTRIBUTION FUNCTION: THEORY AND COMPUTATION

231 2.1 THEORY

232 To properly understand the PDF and its limitations we need to derive its mathemat-
233 ics. The PDF has been previously derived many times so it is not re-derived here.
234 This discussion of the PDF and its gradients use the notation of Farrow and Billinge.
235 [6]

236 **Derivation**

237 **Analytically Gradients**

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

247 cite something with debye waller most likely simon's paper

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

248 where

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

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

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

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

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

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

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

252 where

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

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

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

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

253 The derivatives are as follows:

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

254 where

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

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

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

256 Thus

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

257 **Without ADPs**

258 Without ADPs the equations simplify down to

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

259 and

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

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

262 **Periodic Boundary Conditions**

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

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

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

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

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

268 boxes. Neglecting the thermal motion portion:

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

269 where

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

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

270

271 Need to include PBC ADP math

272 Also should include PBC gradients, although they are trivial, maybe?

273 How does this compare against the Ewald simulation technique for ionic solutions

274 **2.2 COMPUTATION**

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

279 **HPC and GPUs**

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

284 **GPUs and Parallelization**

285 use the figure from the presentation, it is better

286 I wish we had a picture of the GPU profiling

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

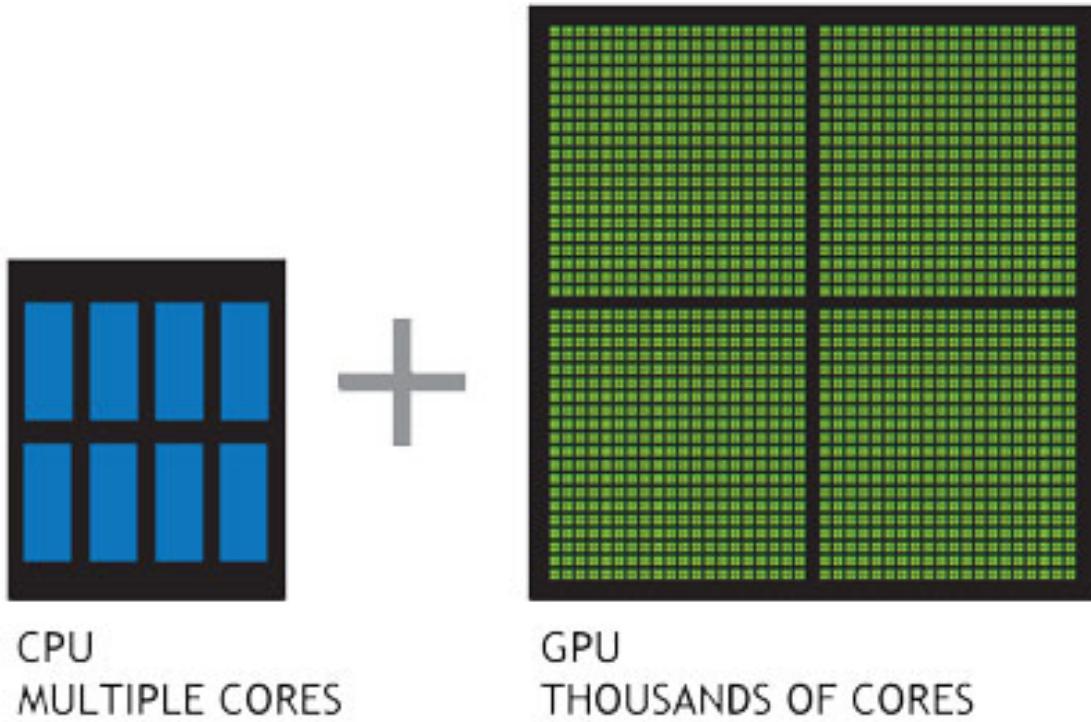


Figure 2.1: Comparison of the CPU and GPU chip architectures

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

297 **Map from ij space to k space**

298 The above equations, although formally correct, are very inefficient. $F(Q)$ and its
 299 gradient are indexed over all the atoms twice, however there are symmetries that
 300 allow us to only compute over the atom pairs essentially mapping from an $n \times n$ space,
 301 ij space, to a $\frac{n(n-1)}{2}$ space, k space. For $F(Q)$ we apply the following mapping where
 302 E denotes the atomic coordinates in ij space, E' denotes $F(Q)$ before the summation
 303 in ij space, B denotes the atomic pairs in k space, B' denotes $F(Q)$ in k space, and
 304 Z denotes the final summed $F(Q)$. For the operators, ϕ denotes the mapping from
 305 ij space to k space $k = j + i * \frac{i-1}{2}$, ψ and ψ' denote the $F(Q)$ operation in ij and k

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

306 space, respectivly. Σ denotes the sum over all the atoms.

307 To properly define Σ' we must establish whether $F(Q)$ is an even function. We
 308 can accomplish this by examining each of the portions of $F(Q)$, $\alpha, \beta, \tau, \Omega$. Ω is even,
 309 since r_{ij} is the interatomic distance, which is the same despite a flip of indicies, Q
 310 does not depend on the atomic indicies, and since Qr_{ij} is even so is $\sin Qr_{ij}$. Thus,
 311 Ω is even. Providing similar analysis to τ we can see that while \vec{u}_{ij} is odd, so is
 312 the unit displacement vector between the two atoms, thus the two odds cancel out.
 313 Intuitivly this makes sense, since the $F(Q)$ equation is fundamentally interested in the
 314 interatomic distances which is even. Thus, switching atom indicies does not change
 315 $F(Q)$. Due to the even nature of the $F(Q)$ operator the Σ' operator sums over all the
 316 atom pairs, and multiplies by two to reflect the double counting of the Σ operator.

For the gradient a similar mapping is used:

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

317

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

323 GPU Memory Allocation

324 While GPUs are very fast computational engines they tend to be memory bound.
325 While a gradient array for a 10nm Au nanoparticle, consisting of 31,000 atoms and
326 half a billion unique distances, occupies 1.5 TB of memory a single GPU's RAM
327 allotment varies from 4GB on a NVIDIA GTX970 to 24 GB on a NVIDIA Tesla K80.
328 Thus, it is important to determine exactly how many atoms can fit on a GPU of
329 arbitrary size as a function of the number of atoms and the Q range. The memory
330 required per array is:

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

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

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

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

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

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

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

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

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

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

331 where n is the number of atoms, k is the number of unique distances, Q is the scatter
332 vector, and the $[=]$ operator denote the number of single precision floating point
333 values in memory. Each of the above arrays are used in the computation and thus
334 must be able to be held in memory. Thus the number of atom pairs that can fit on
335 a GPU with am bytes of available memory is:

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

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

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

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

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

338 Thus the pair allotment is:

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

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

$$\vec{\nabla}\Omega = 3kQ \quad (2.36)$$

$$\vec{\nabla}F_k(Q) = 3kQ \quad (2.37)$$

$$\vec{\nabla}F_n(Q) = 3nQ \quad (2.38)$$

341 Thus the gradient allotment is:

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

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

$$\vec{\nabla}\tau = 3kQ \quad (2.40)$$

343 Thus the allocation is:

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

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

Include Speed Benchmarks Here

347

348

CHAPTER 3

349

BENCHMARKING

350

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

351

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

352

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.

355

3.1 PDF

356

The formation of NPs with both crystallographic and non-crystallographic structures [12] and with different chemical patterns [7] 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 [8, 11]. We used optimized structures from the Density Functional Theory (DFT) as target structures and generated the corresponding PDF, G_{obs} , according to

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

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

372 Spin-polarized DFT calculations were carried out using the Vienna ab initio sim-
373 ulation package (VASP) [9, 10] within the Perdew-Burke-Ernzerhof (PBE) exchange-
374 correlation functional [13]. The projected augmented wave method [1] and a kinetic
375 energy cutoff of 400 eV were used. Structural optimization was performed until the
376 total energy and ionic forces were converged to 10^{-6} eV and 10 meV/Å, respectively.
377 The amorphous Au₅₅ structures were generated by simulated annealing using the
378 classical embedded atom method potential [16]. Different annealing temperatures
379 between 1200 K and 1670 K (bulk melting temperature of Au) were used and the
380 thermally equilibrated structures were cooled down to 300 K before minimization at
381 0 K. Further optimization using DFT leads to total energies that vary within 1-2
382 eV among different amorphous structures and the lowest energy one was used as the
383 target structure. The target structure of Au₁₀₂ was taken as the Au₁₀₂ core of the
384 DFT-optimized Au₁₀₂MBA₄₄ cluster [11].

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

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

388 The thresholded spring potentials are based on those previously proposed on by Pe-
389 terson [14], 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
390 bounds of the spring threshold r_t . The resulting restoring forces on the out-of-bound
391 atoms bring the system back within the bounds of the PDF, R_{\min} and R_{\max} , and
392 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

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

420 **Au55: surface relaxed**

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

428 **Run Parameters**

429 R_{\min} and R_{\max} for this simulation were 2.45 Å and 11.4 Å, respectively, with $\delta Q =$
430 0.24 Å⁻¹. The simulation ran for approximately 34 minutes, over a total of ∼40
431 thousand configurations. The results are shown in Fig. 3.1.

432 The PDF, radial bond distribution, and bond angle distribution show good agree-
433 ment between the target and final fitted structures, with a Rw of 0.3% whereas Rw
434 of the starting structure is as high as 44.8%. DFT calculations yield a total energy of
435 the final structure very close to that of the target structure (within a few meV). The
436 success in the fitting is largely attributed to the factor that the target structure is
437 only locally (and mildly) disturbed from its bulk-like counterpart and therefore there
438 is no need to overcome any high PES barriers to reach the correct solution. As shown
439 below, the situation is rather different for much more disordered target structures.
440 Interestingly, the small-box solution using PDFgui[5] yields a rather large Rw of 43%,

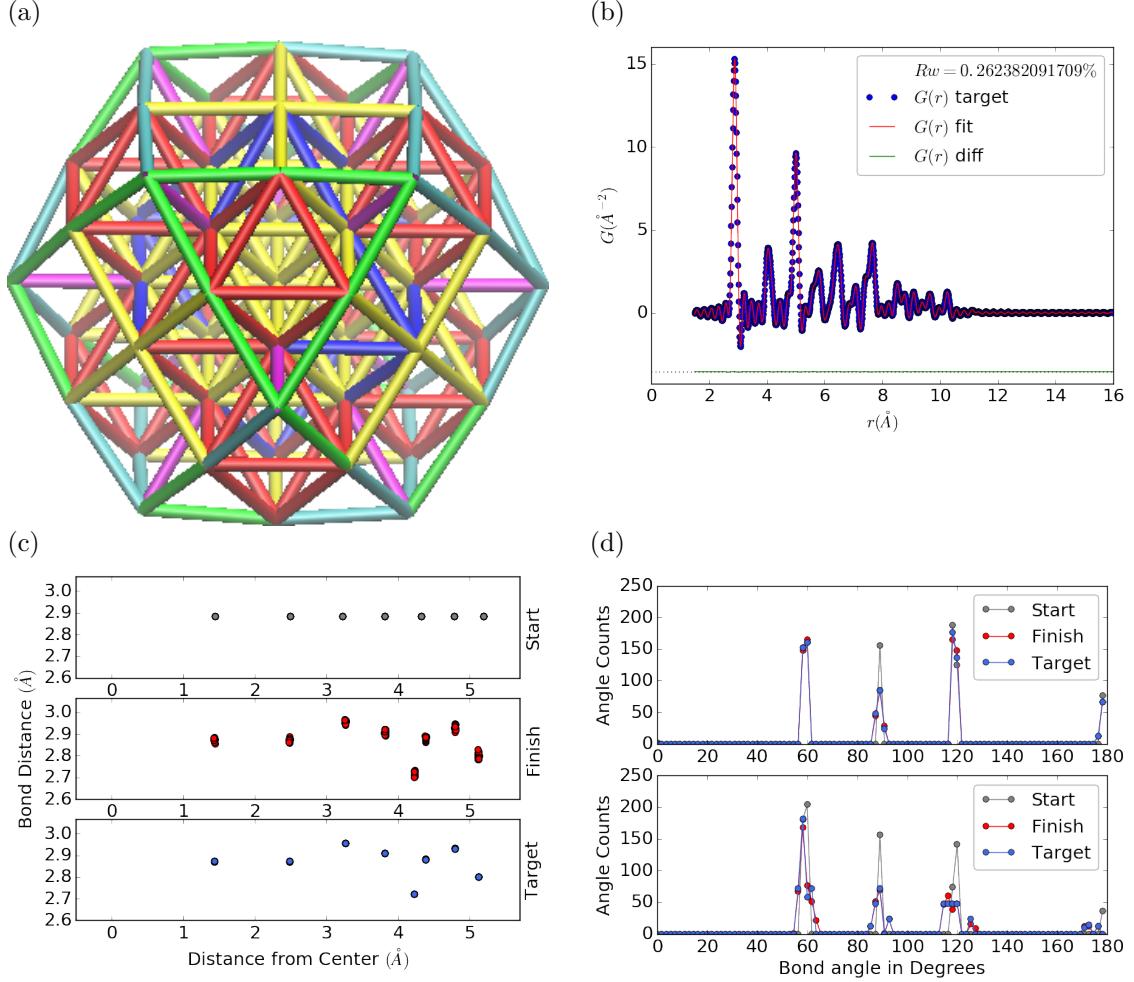


Figure 3.1: Au_{55} PDF fitting of DFT-optimized $c\text{-}\text{Au}_{55}$. (a) the final structural solution ($Rw=0.3\%$) with bond lengths color-coded by step of 0.05\AA , (b) the target PDF(blue dots) overlaid with the PDF of the final structure (solid red lines) with the difference in green lines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

443 Put this somewhere

444 .

445 **Au55: surface disordered**

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

452 R_{\min} and R_{\max} for this simulation were 1.95 Å and 12.18 Å, respectively, with
453 $\delta Q = 0.23 \text{ \AA}^{-1}$. The simulation ran for approximately 3.6 hours, over a total of ~ 270
454 thousand configurations. The results of the simulation are shown in Fig. ??.

455 Overall, good agreement is found between PDFs of the target structure and the
456 final structural solution, even out to larger r , with an $Rw = 0.6\%$ starting from an
457 $Rw = 50.4\%$ (see Fig. S2). The radial bond distribution and angle distribution
458 show reasonably good agreement, but with lower degree of crystallinity in the final
459 structure compared to the target structure. The discrepancy is most obvious in
460 the core: despite the identical core structure in the starting and target structures,
461 the core atoms were displaced in the HMC simulations in order to achieve a “best”
462 solution. This is because PDF measures the global average of interatomic distances
463 between each atomic pair and does not contain direct information about the locality
464 of such pairs, e.g. on the surface or cores. If such information is obtained a priori, for
465 example, from theoretical prediction or other experimental measurements, the core
466 structure can then be fixed and excluded from HMC dynamics.

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

472 and high CN numbers. This is partly caused by the displacement of the core atoms
473 mentioned above, and partly by the lack of CN constraints for the PDF fitting, which
474 has been previously demonstrated in the case of α -Si [2]. Additional experimental
475 data, e.g. from EXAFS or NMR, may help to steer the simulations towards better
476 agreement in both PDF and CN distribution.

477 Au55: amorphous

478 Next, we turn to the case in which the entire cluster structure is disordered. We used
479 a DFT-optimized amorphous Au_{55} ($a\text{-Au}_{55}$) as the target structure, and the DFT-
480 relaxed $c\text{-Au}_{55}$ cluster from Case I as the starting structure. The total energy of
481 $a\text{-Au}_{55}$ was computed to be *lower* than that of $c\text{-Au}_{55}$ by as large as 2.9 eV, consistent
482 with the 3.0 eV found in previous DFT work [4].

483 R_{\min} and R_{\max} for this simulation were 2.6 Å and 11.26 Å, respectively, with
484 $\delta Q = 0.25 \text{ \AA}^{-1}$. The simulation ran for approximately an hour, over a total of ~ 87
485 thousand configurations. The results of the simulation are shown in Fig. ??.

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

498 **Au102: triple phase**

499 Our final benchmark is Au₁₀₂, whose structure was initially solved by Jadzinsky and
500 co workers using x-ray crystallography [8] and further confirmed by DFT studies [11].
501 The Au₁₀₂ structure consists of three main parts, a 49-atom Marks decahedron core,
502 two C_5 caps consisting of 20 atoms each, and 13 equatorial atoms. Unlike previous
503 cases, the multi-symmetry nature of the structure, i.e. each part has its own distinct
504 symmetry, poses a challenge for PDF-based solution of the structure. This is because
505 of the atomically centralized nature of the PDF, in which each atom “sees” a density
506 of other atoms surrounding it and has a strong tendency towards becoming the center
507 of the main symmetry group. Such tendency may lead to a solution where some of
508 the correct atomic symmetries are discarded in favor of the core symmetry.

509 **Starting from fcc structure**

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

512 R_{\min} and R_{\max} for this simulation were 2.7 Å and 16. Å, respectively, with $\delta Q =$
513 0.18 Å⁻¹. The simulation ran for approximately two hours, over a total of ~82
514 thousand configurations. The results of the simulation are shown in Fig. 3.6.

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

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

534 Marks decahedron

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

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

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

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

553 **Au147**

554 **3.2 PDF WITH ADPs**

555 **ADP 50**

556 1. Basic 50% larger magnitude

557 2. Random addition to APDs

558 3. Janus ADPs

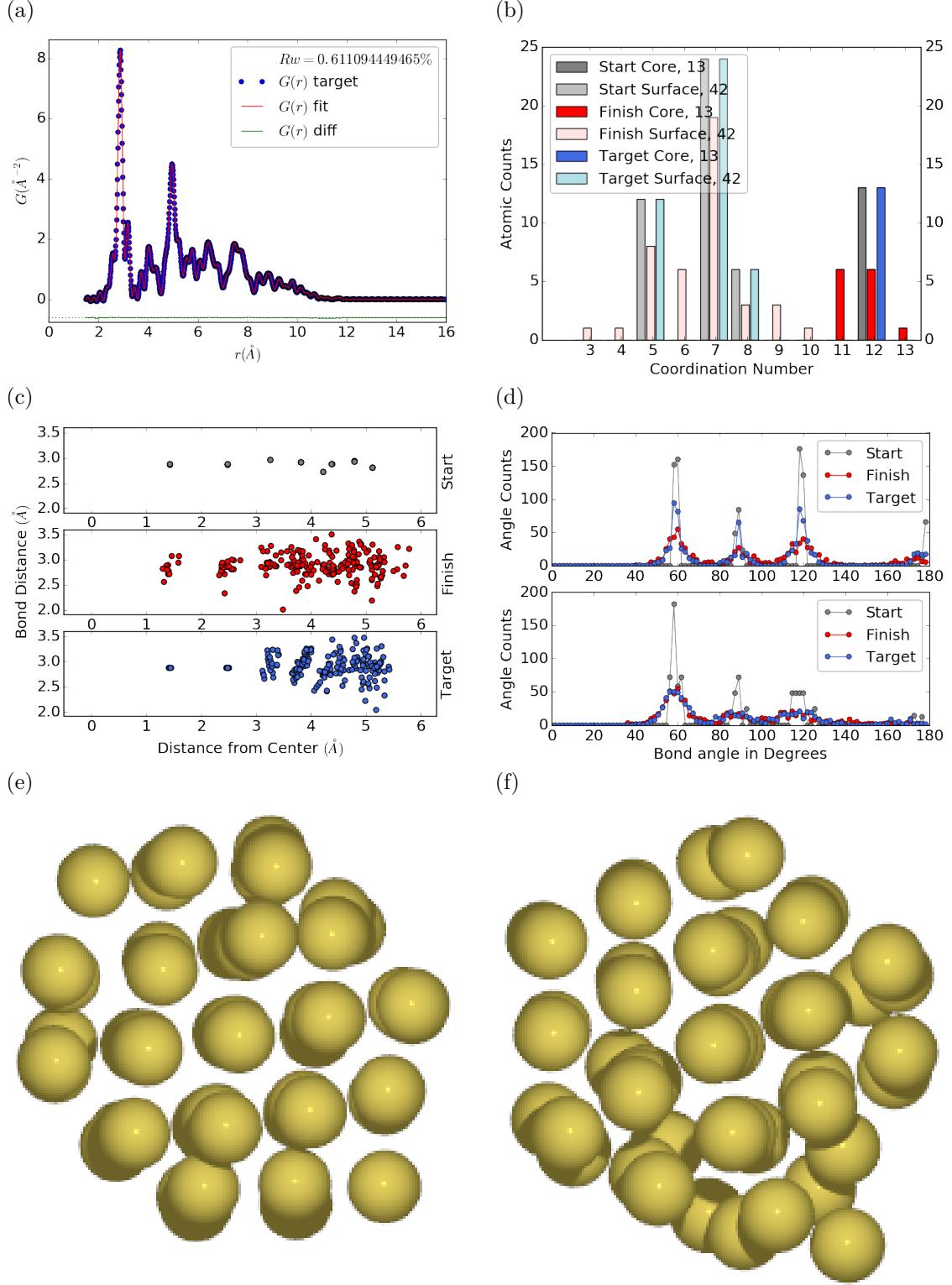


Figure 3.2: Au_{55} PDF fitting of surface-disordered Au_{55} . a) the target structure, b) the final structural solution ($Rw=0.6\%$), c) the comparison of PDFs, d) the CN distribution with the number of atoms in either the core or the surface, e) the radial bond distribution, and f) the bond angle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

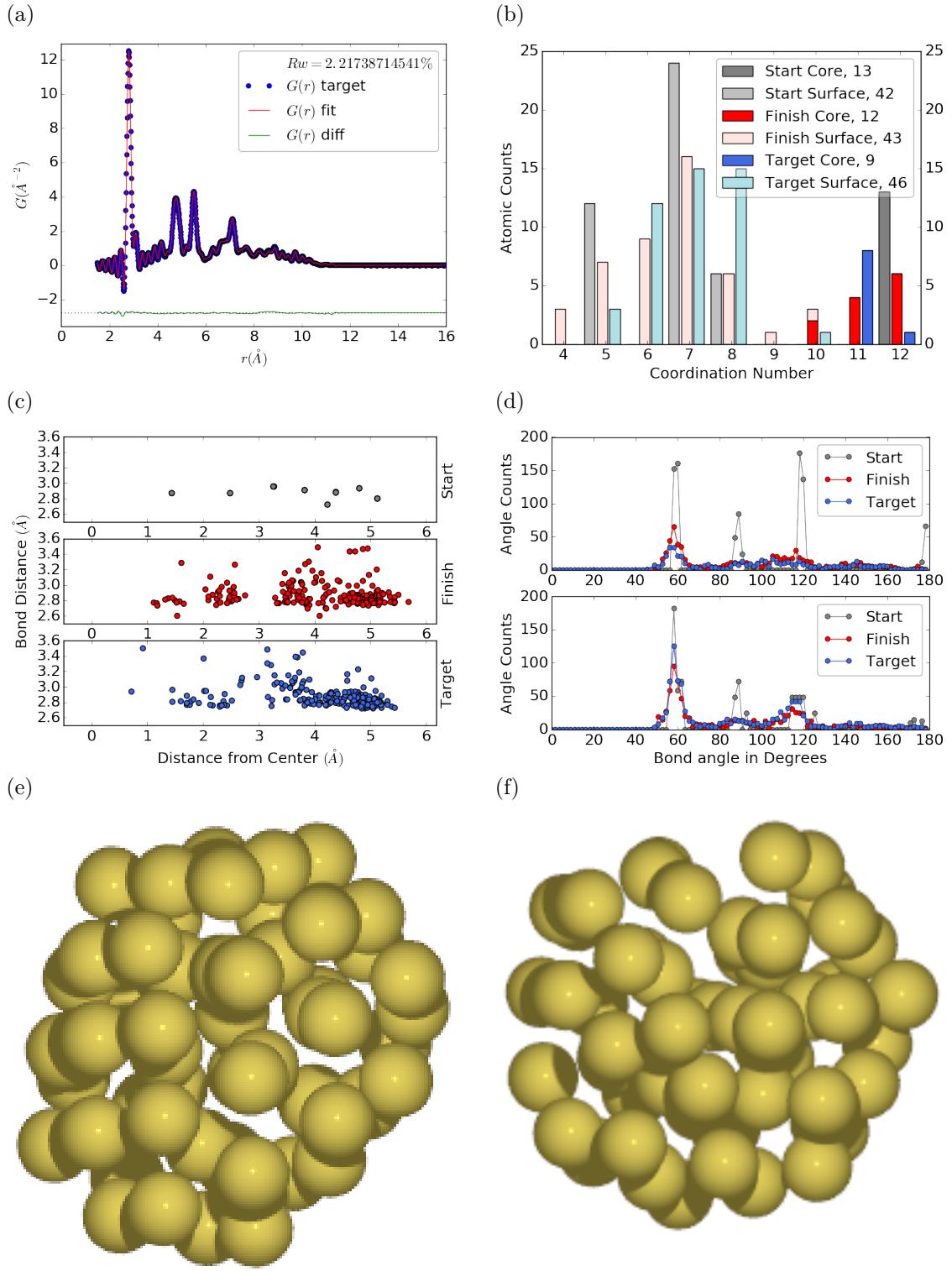


Figure 3.3: Similar to figure 3.2 for DFT-optimized amorphous Au₅₅.

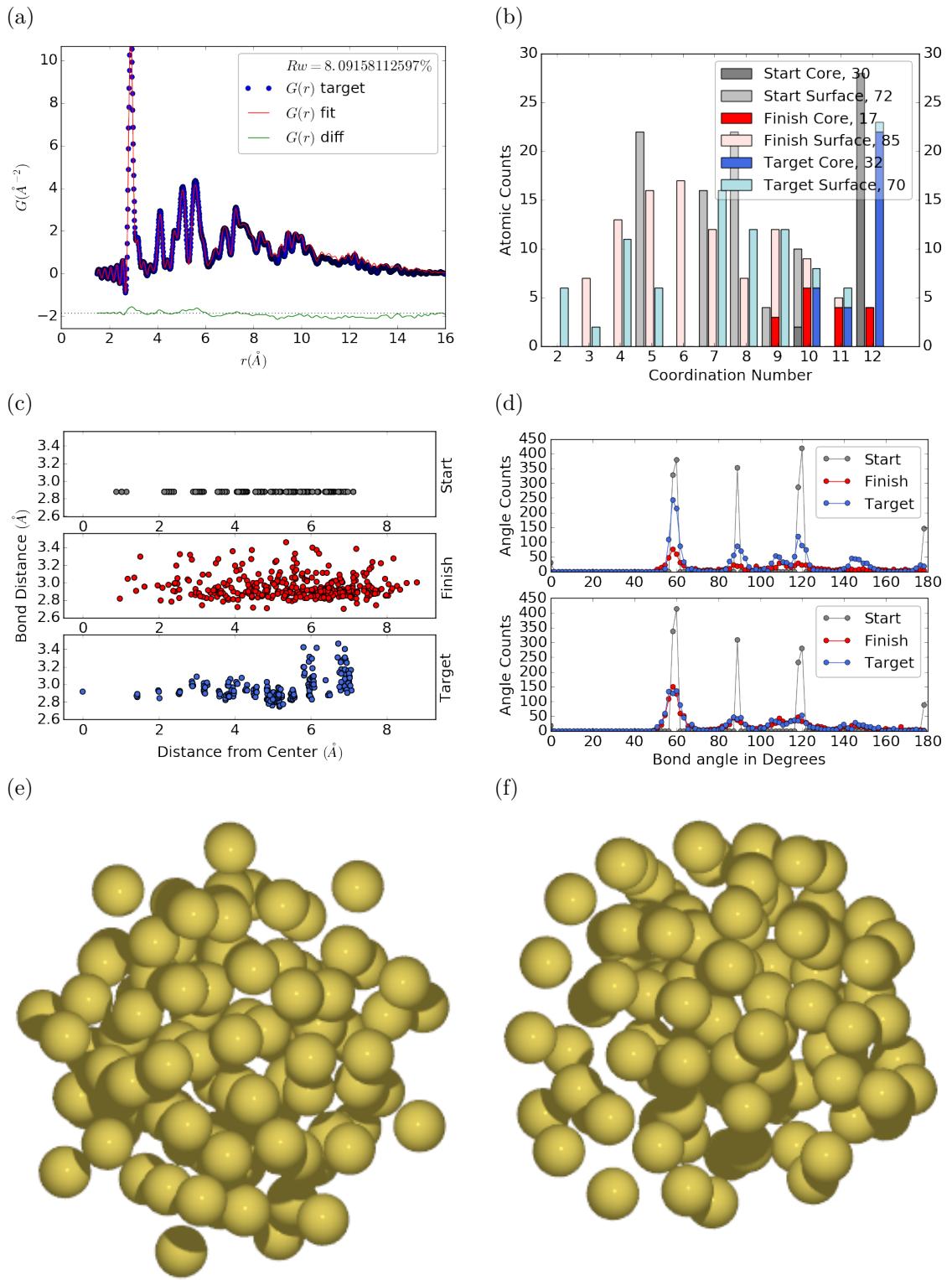


Figure 3.4: Similar to figure 3.2 for Au_{102} as in DFT-optimized $\text{Au}_{102}\text{MBA}_{44}$ cluster.

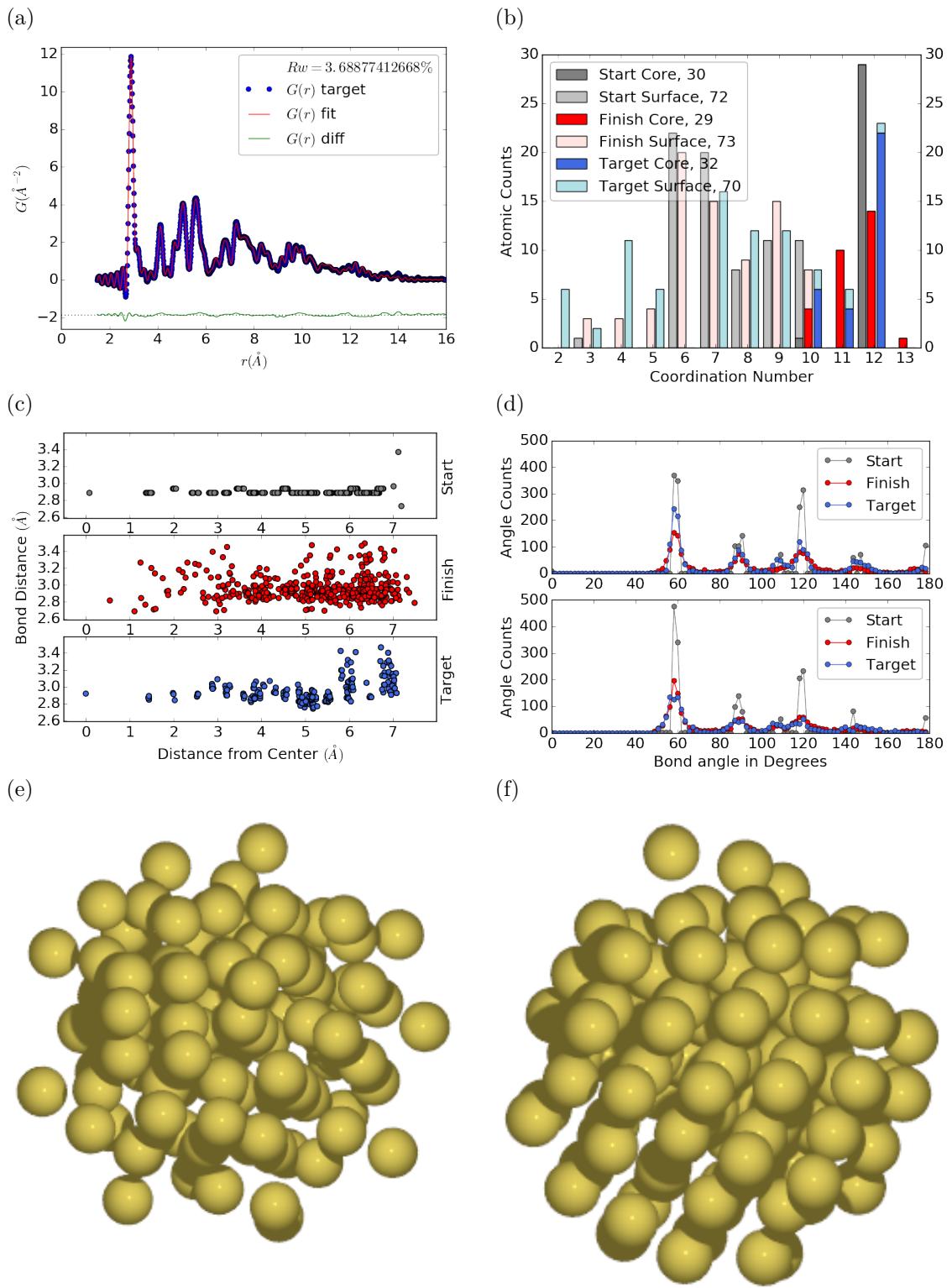


Figure 3.5: Similar to Fig. 3.6 with Marks decahedron as the starting structure.

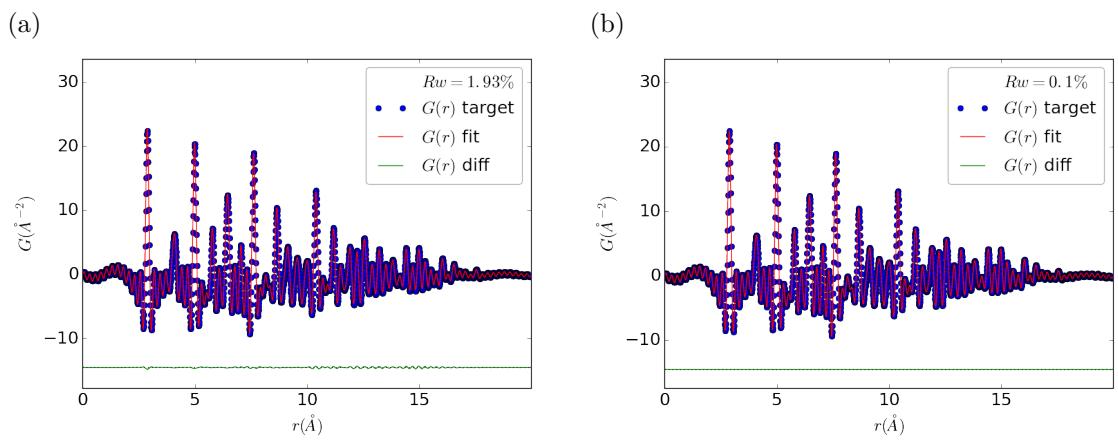


Figure 3.6

559

CHAPTER 4

560

X-RAY TOTAL SCATTERING DATA ACQUISITION AND

561

PROCESSING

562

4.1 INTRODUCTION

563 X-ray total scattering experiments are generally performed at synchrotron light sources,
564 as only these sources can provide the needed flux, energy, and high momentum trans-
565 fer vectors needed to obtain reliable PDFs. [?, ?] Despite the need for a dedicated
566 facility to perform the total scattering experiments, the experiments themselves are
567 fairly forgiving, allowing for reactive gasious environments, experiment temperatures
568 ranging from 2 K to 1200 K, and even electrochemical cycling. [?, 15, ?] The rapid
569 PDF data aquesition associated with 2D area detectors creates a data management
570 problem, as 96 hours of beamtime could result in almost 10,000 images which need
571 to be associated with the experimental conditions and detector metadata. [?] Finaly,
572 all this data needs to be processed by masking bad pixels and regions, integrating
573 azimuthally, and converting the scattering data to the PDF. [?, ?, ?, ?, ?]

574

4.2 DETECTOR Q RESOLUTION

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

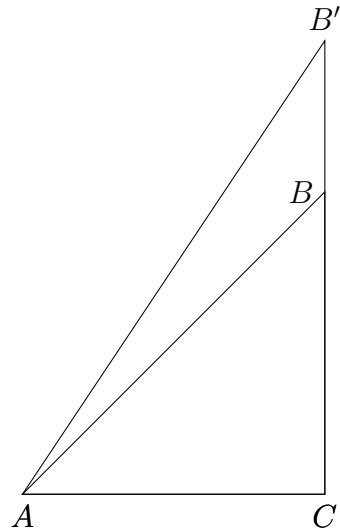


Figure 4.1: Scattering onto a flat detector

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

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

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

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

586 where λ is the x-ray wavelength.

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

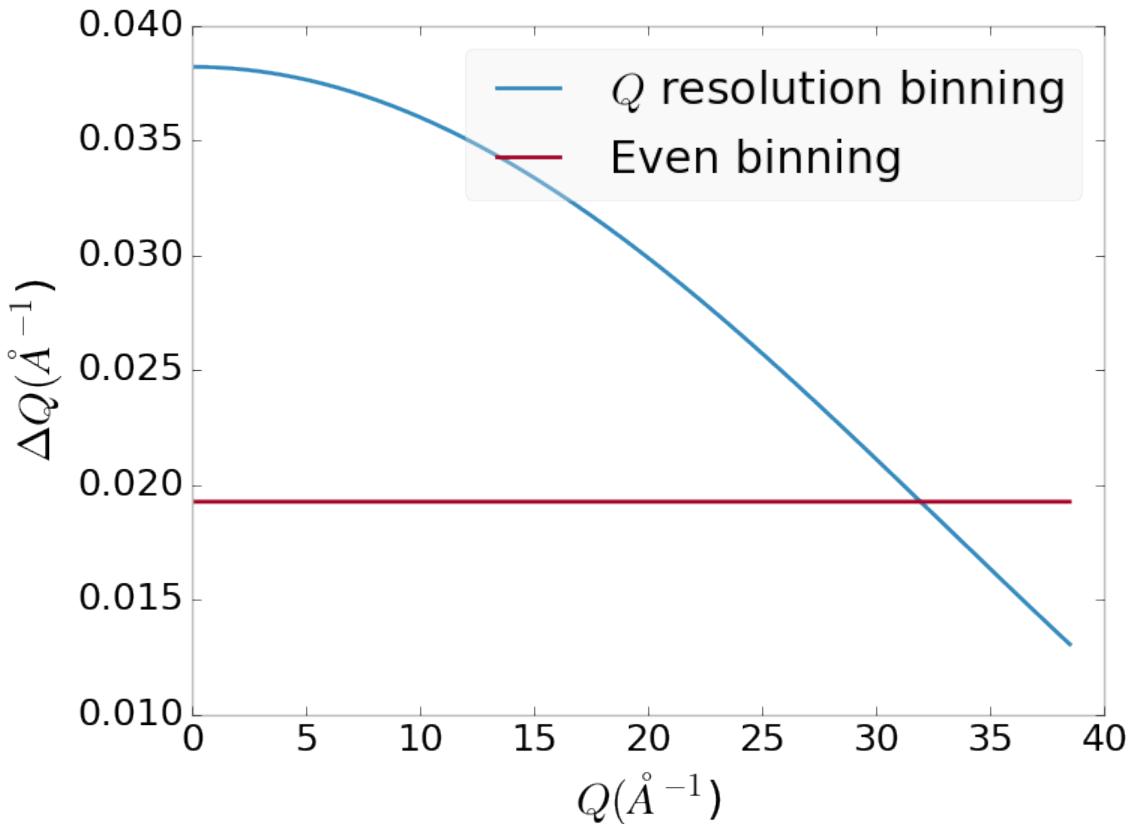


Figure 4.2: Q resolution as a function of Q .

590 4.3 AUTOMATED MASK GENERATION

591 **Introduction**

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

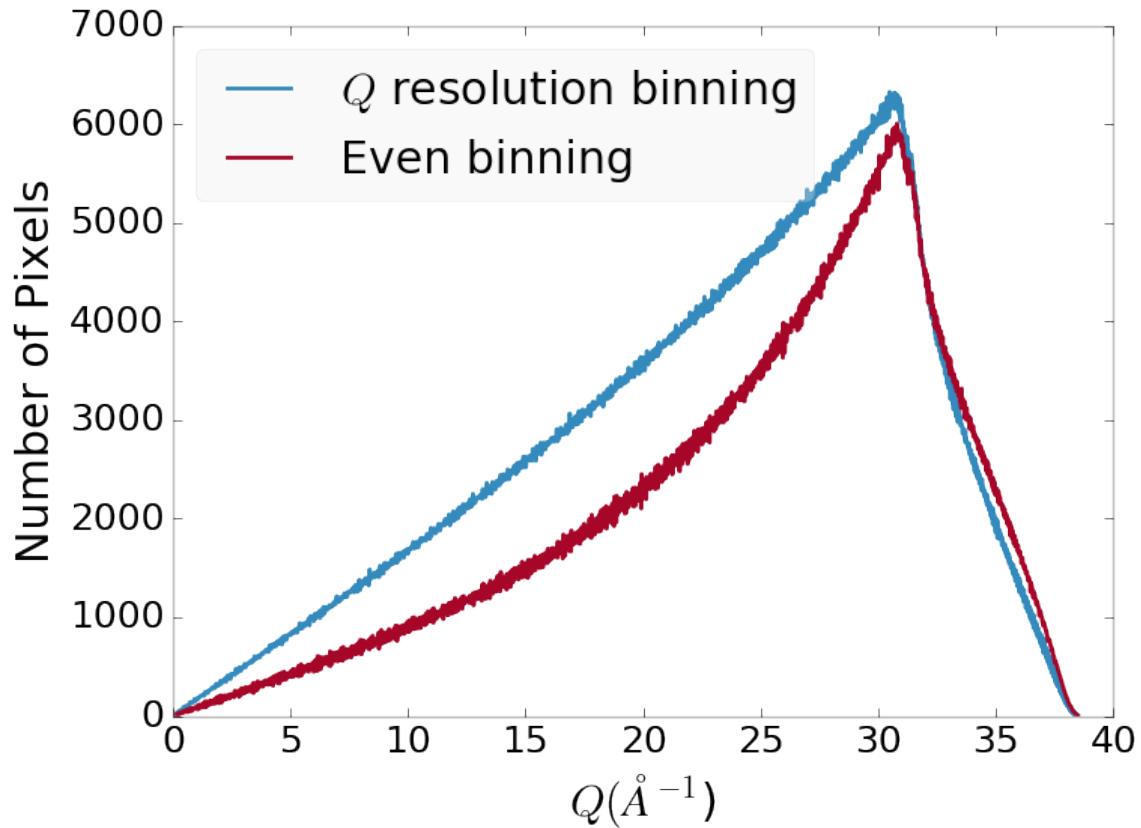


Figure 4.3: Number of pixels as a function of Q , binned at the Q resolution of the detector.

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

610 **Algorithm Design**

611 The masking algorithm procedure takes in the image and a description of the pixel
612 positions in either distance from the point of incidence or in Q . The image is then
613 integrated twice, producing both the mean $I(Q)$ and the standard deviation of each
614 $I(Q)$ ring. The mask is created by comparing the pixel values against each ring's
615 standard deviation and threshold α . Note that the threshold can be a function of
616 distance from the point of incidence or Q .

617 **Test Cases**

618 To study the effectiveness of the masking we ran the algorithm against both simulated
619 experimental data. In the case of the simulated data four systems were created: 1)
620 dead/hot pixels with varying numbers of defective pixels, 2) beamstop holder with
621 varying beamstop holder transmittance, 3) rotated beamstop holder with varying
622 beamstop holder transmittance, and 4) beamstop holder with dead/hot pixels. The
623 base scattering was produced by

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

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

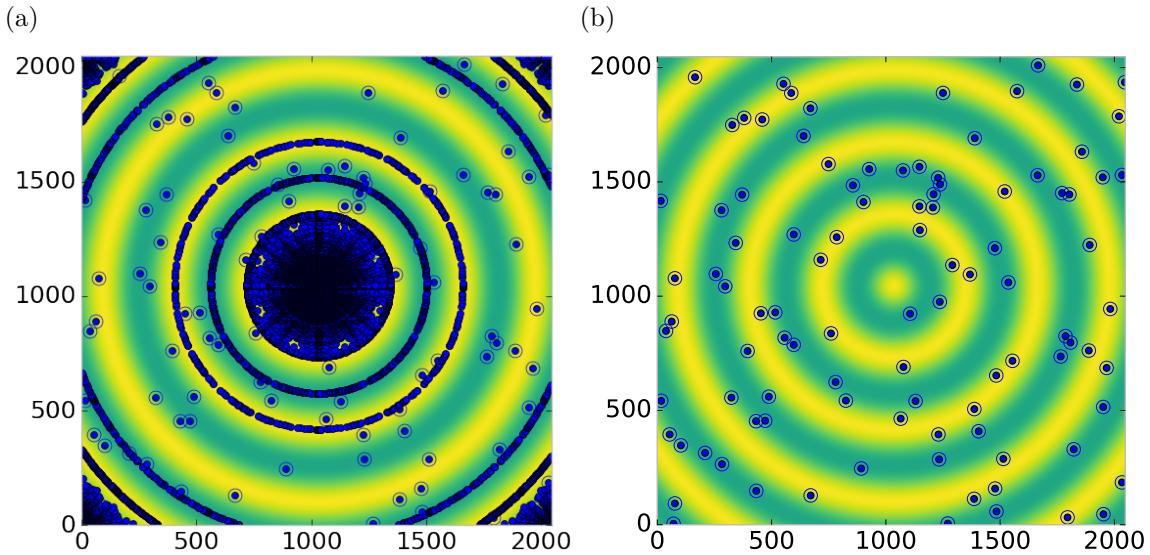


Figure 4.4: Generated dead/hot pixel masks for a detector with 100 bad pixels. a) the standard even bin mask and b) the Q resolution binned mask. The bad pixels are noted with open circles, masked pixels are noted with closed circles.

634 Results and Discussion

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

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

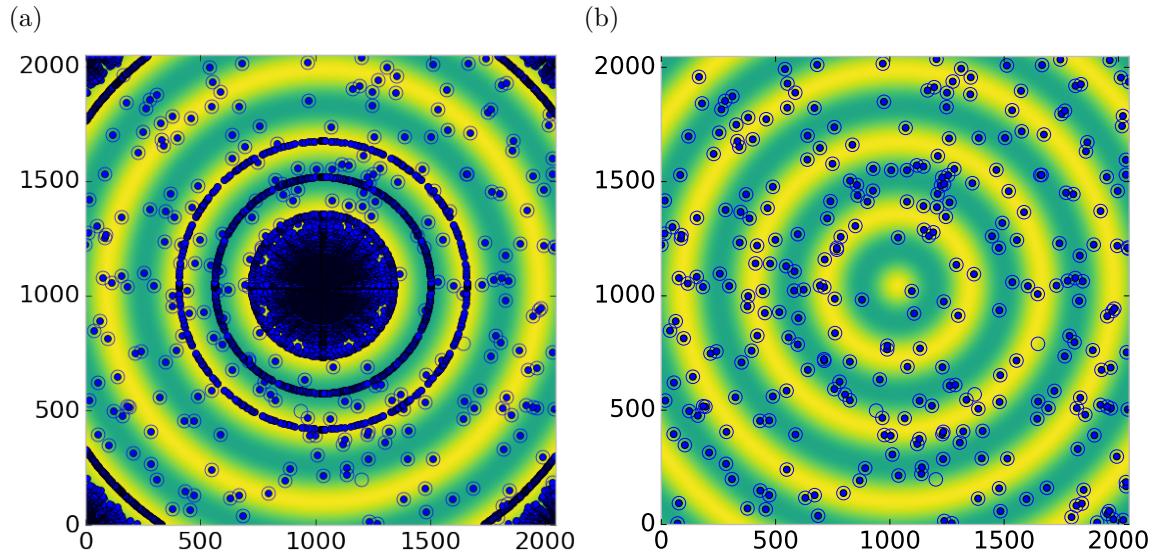


Figure 4.5: Generated dead/hot pixel masks for a detector with 300 bad pixels. a) the standard even bin mask and b) the Q resolution binned mask. The bad pixels are noted with open circles, masked pixels are noted with closed circles.

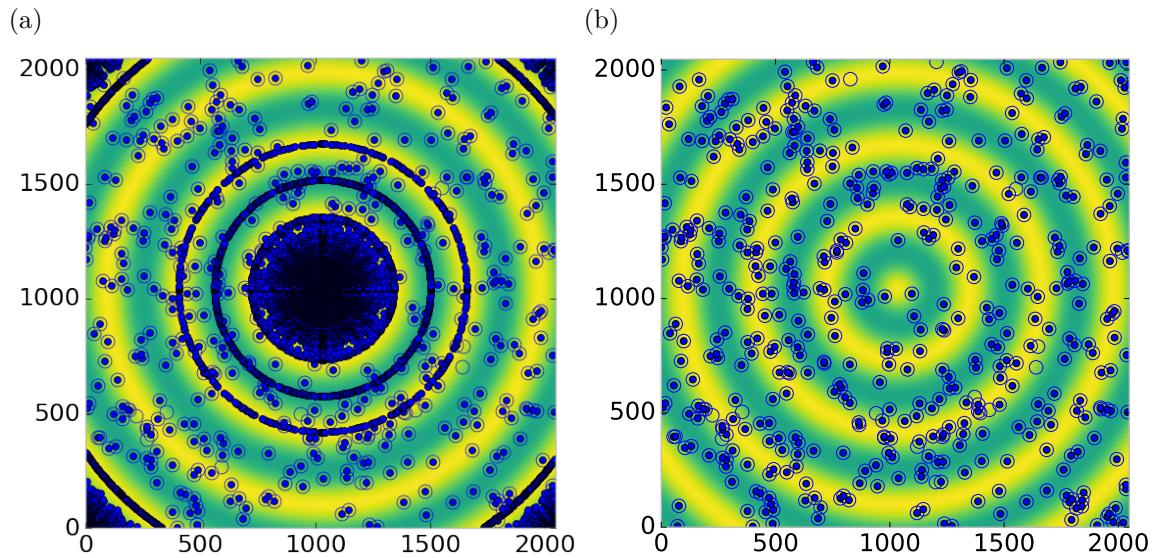


Figure 4.6: Generated dead/hot pixel masks for a detector with 500 bad pixels. a) the standard even bin mask and b) the Q resolution binned mask. The bad pixels are noted with open circles, masked pixels are noted with closed circles.

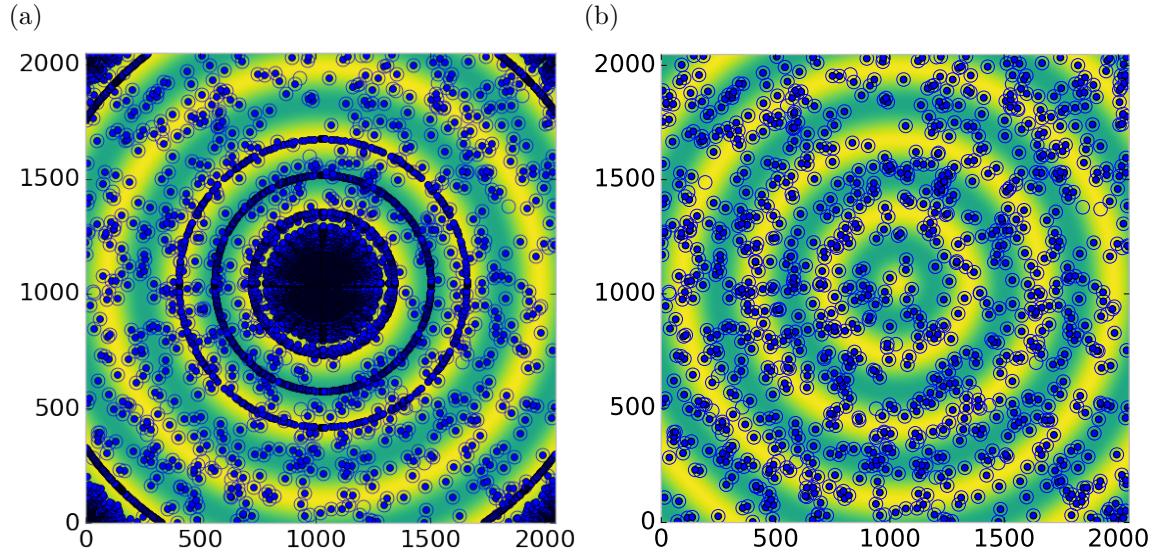


Figure 4.7: Generated dead/hot pixel masks for a detector with 1000 bad pixels. a) the standard even bin mask and b) the Q resolution binned mask. The bad pixels are noted with open circles, masked pixels are noted with closed circles.

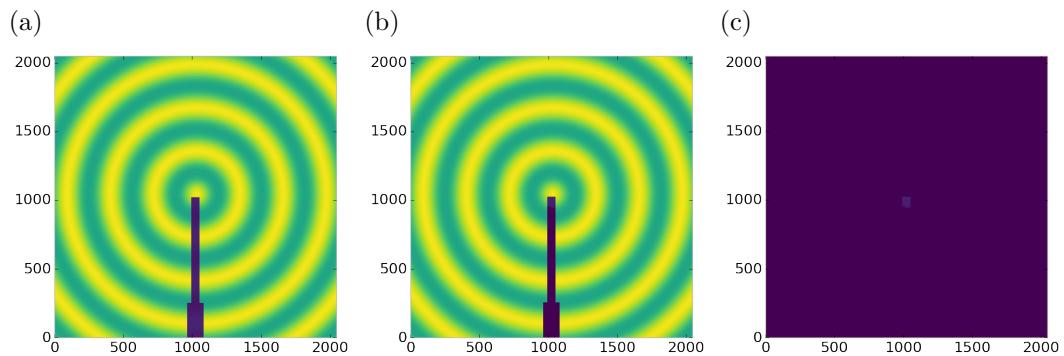


Figure 4.8: Generated beamstop holder masks for a beamstop holder with 10% transmittance. a) the raw image, b) the masked image, c) and the missed pixels

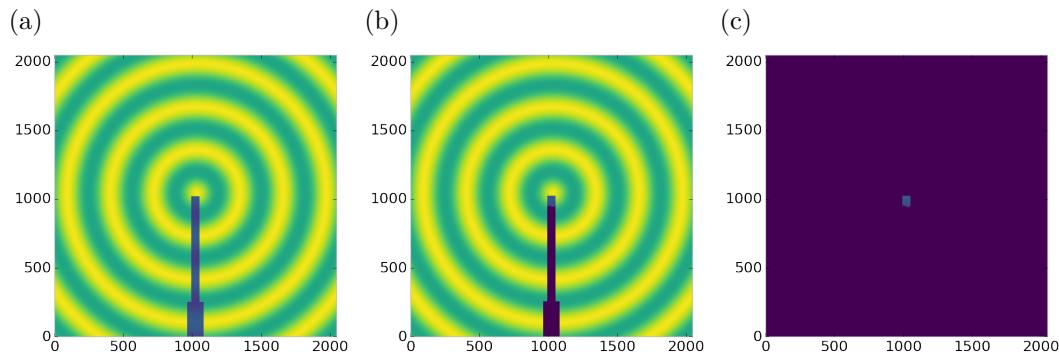


Figure 4.9: Generated beamstop holder masks for a beamstop holder with 30% transmittance. a) the raw image, b) the masked image, c) and the missed pixels

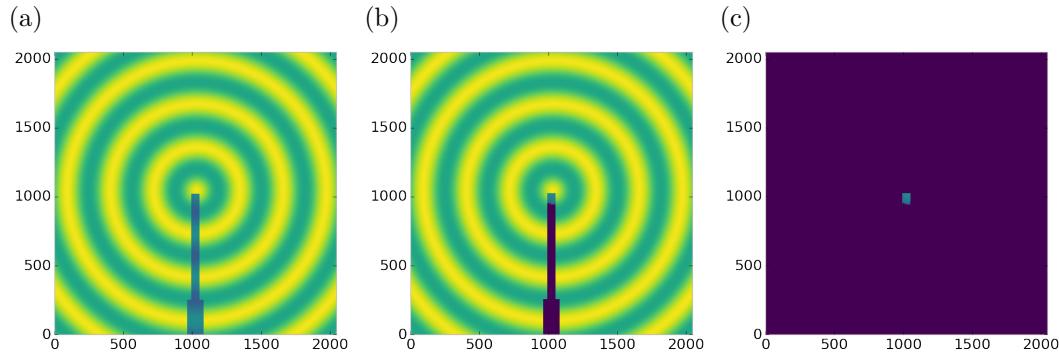


Figure 4.10: Generated beamstop holder masks for a beamstop holder with 50% transmittance. a) the raw image, b) the masked image, c) and the missed pixels

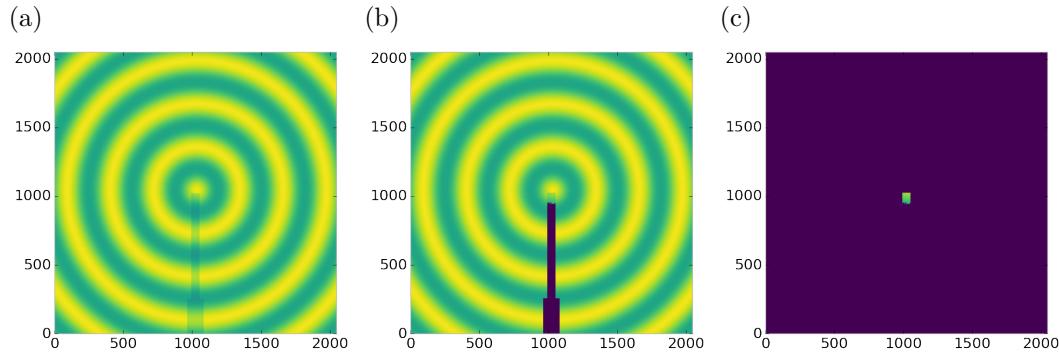


Figure 4.11: Generated beamstop holder masks for a beamstop holder with 90% transmittance. a) the raw image, b) the masked image, c) and the missed pixels

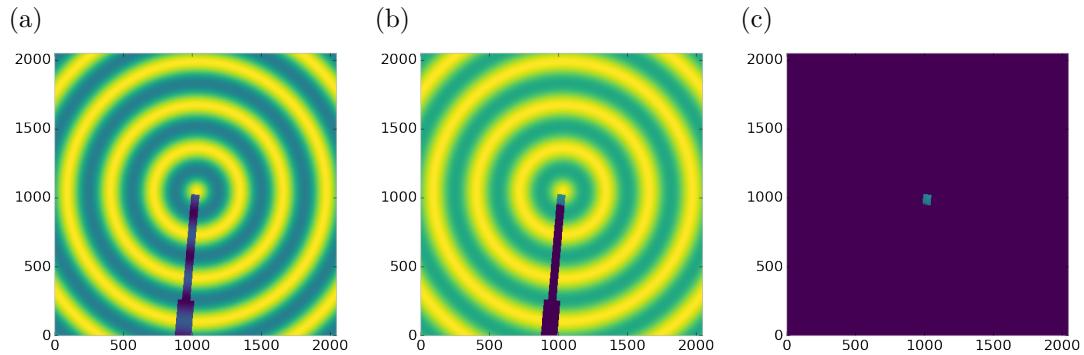


Figure 4.12: Generated beamstop holder masks which is rotated away from vertical

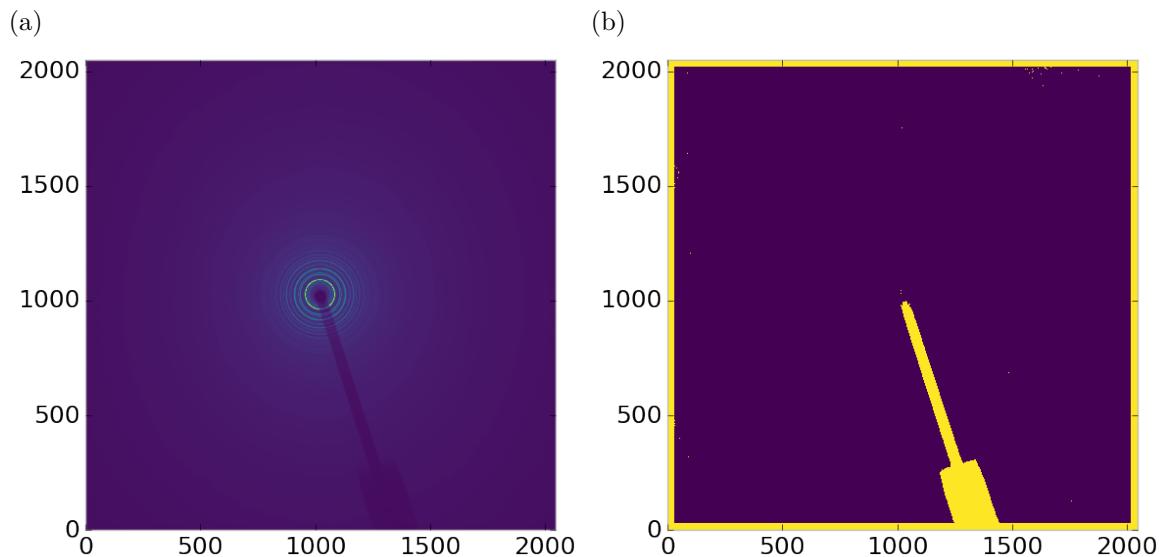


Figure 4.13: Masked experimental data. a) the raw image, b) the mask

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 neglected 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 vertical position, as shown in figure 4.12

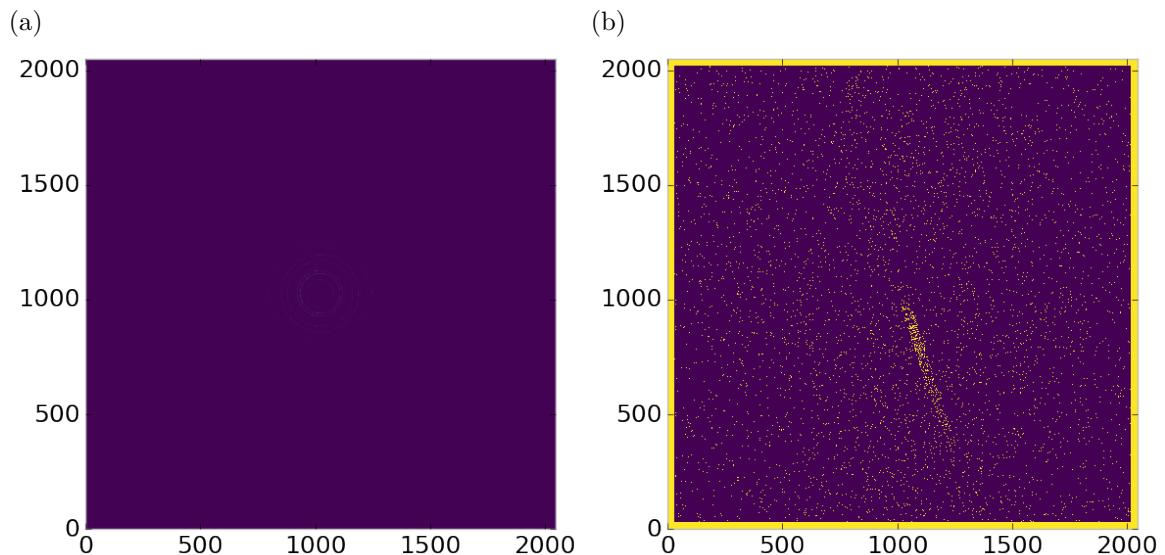


Figure 4.14: Masked experimental data with Pt single crystal signal. a) the raw image, b) the mask

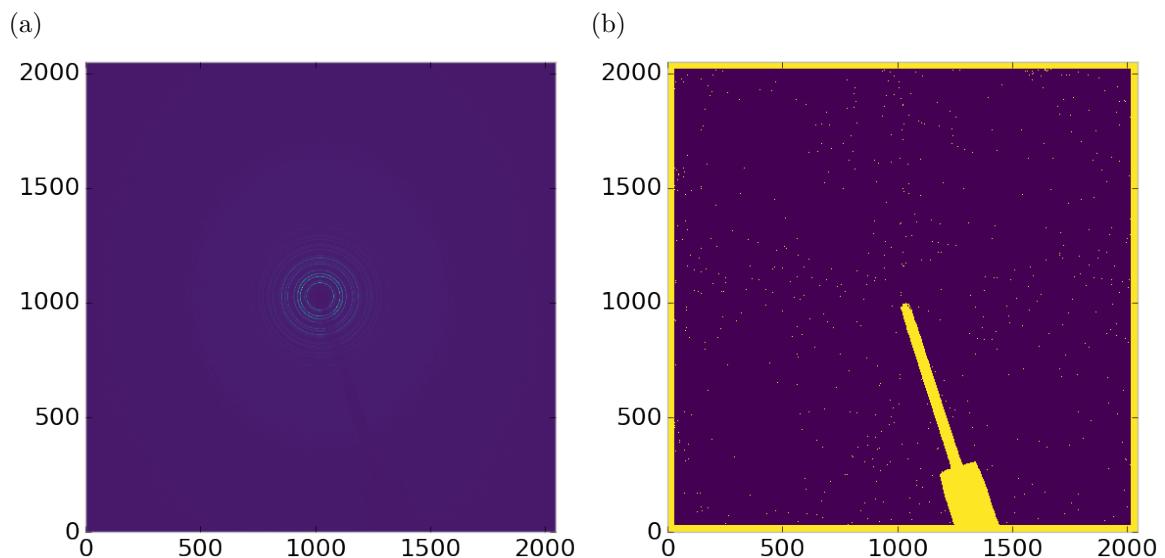


Figure 4.15: Masked experimental data with Pt single crystal signal using figure's 4.13 mask as a starting mask. a) the raw image, b) the mask

657 Working with actual experimental data, obtained at the Advanced Photon Source
658 beamline 11-ID-B, shows the difficulty of masking images which have low photon
659 counts. While the masking of experimental data taken with longer exposures, con-
660 sisting of 250 .2 second shots, shown in figure 4.13 provides very sharp edges to the
661 beamstop holder, and very little extra masking beyond the occasional dead pixel, this
662 is not the case for the single crystal data. The single crystal data is more problem-
663 atic because of its short exposure time and low flux, with 500 frame at a .1 second
664 exposure and having shrunk the beam size. The low flux is to prevent the very strong
665 single crystal peaks from damaging the detector. However, this causes the image
666 to be less statistically viable then ideal, causing problems with the mask as seen in
667 figure 4.14. This can be alleviated to some degree by using the previously generated
668 mask as a starting mask for the single crystal image, as shown in 4.15. While the
669 masking algorithm still produces many diffuse masked pixels, they are far fewer, this
670 may be due to the removal of the beamstop which could have contributed to the large
671 standard deviation in figure 4.14.

672 **Conclusions**

673 In this section the masking algorithm, which relies on both Q resolution based bin-
674 ning and a statistical approach to azimuthal symmetry, was developed. The focus of
675 this algorithm was to remove many unwanted detector features associated with pixel
676 defect, beamstop holder associated scattering attenuation, and single crystal/texture
677 peaks. Simulated data was used to evaluate the beamstop holder and dead/hot pixel
678 masking capacity, while experimental data was used to check for single crystal and
679 texture based masking. Q resolution based binning was shown to be very important
680 to avoid over-masking. The ability of the mask writer to mask images is somewhat
681 limited by the overall statistical image quality, although some deficiencies can be
682 obtained by using previously generated masks as starting points. This masking algo-

683 rithm is now in use in the data processing workflow and will be available in scikit-beam
684 soon.

685 4.4 AUTOMATED IMAGE AZIMUTHAL INTEGRATION

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

690 Figures 4.16-4.18 show the importance of masking and the choice of average func-
691 tion. All the figures were produced using the same dataset, 50 °C Pr₂NiO₄ taken at
692 the APS’s 11-ID-B on a Perkin Elmer area detector. The automatic masking alpha
693 was 3 standard deviations from the mean. While it is difficult to observe the changes
694 the mask causes in the full $I(Q)$ plot (subfigures a) and b)), the standard deviation
695 plots show the effect of bad pixels on the data (subfigure c)). Subfigure c) for figures
696 4.16-4.18 shows that removal of the beamstop holder lowers the low Q standard de-
697 viation from around .1 to almost .01 out to 15 Å⁻¹. The high Q subfigures d) and f)
698 in figures 4.16-4.18 show the “kink” effect of the detector edge and beamstop holder,
699 where there is a dip in the $I(Q)$ scattering when the rings include the edge of the
700 detector. This effect seems to be due to both errors in the edge pixel intensity and the
701 beamstop holder as masking of the edges only seems to provide only partial removal
702 of the issue. It is important to note that while integration using the mean of the
703 ring has issues with only the edge mask, as evidenced by the change in slope in 4.17
704 d) around 29.5 Å⁻¹, the median integration does not include this error. Ideally the
705 detector would have a normal distribution of pixel intensity for a given ring, which
706 would imply an equivalency between the mean and median $I(Q)$ values. Despite the
707 closeness of the mean and median once the final mask has been created, it seems that
708 the median is more reliable, as it was less effected by the beamstop holder in figure

709 4.17. Thus, for subsequent integrations discussed in this work the median is used to
 710 avoid any defective features that the masking algorithm may have missed.

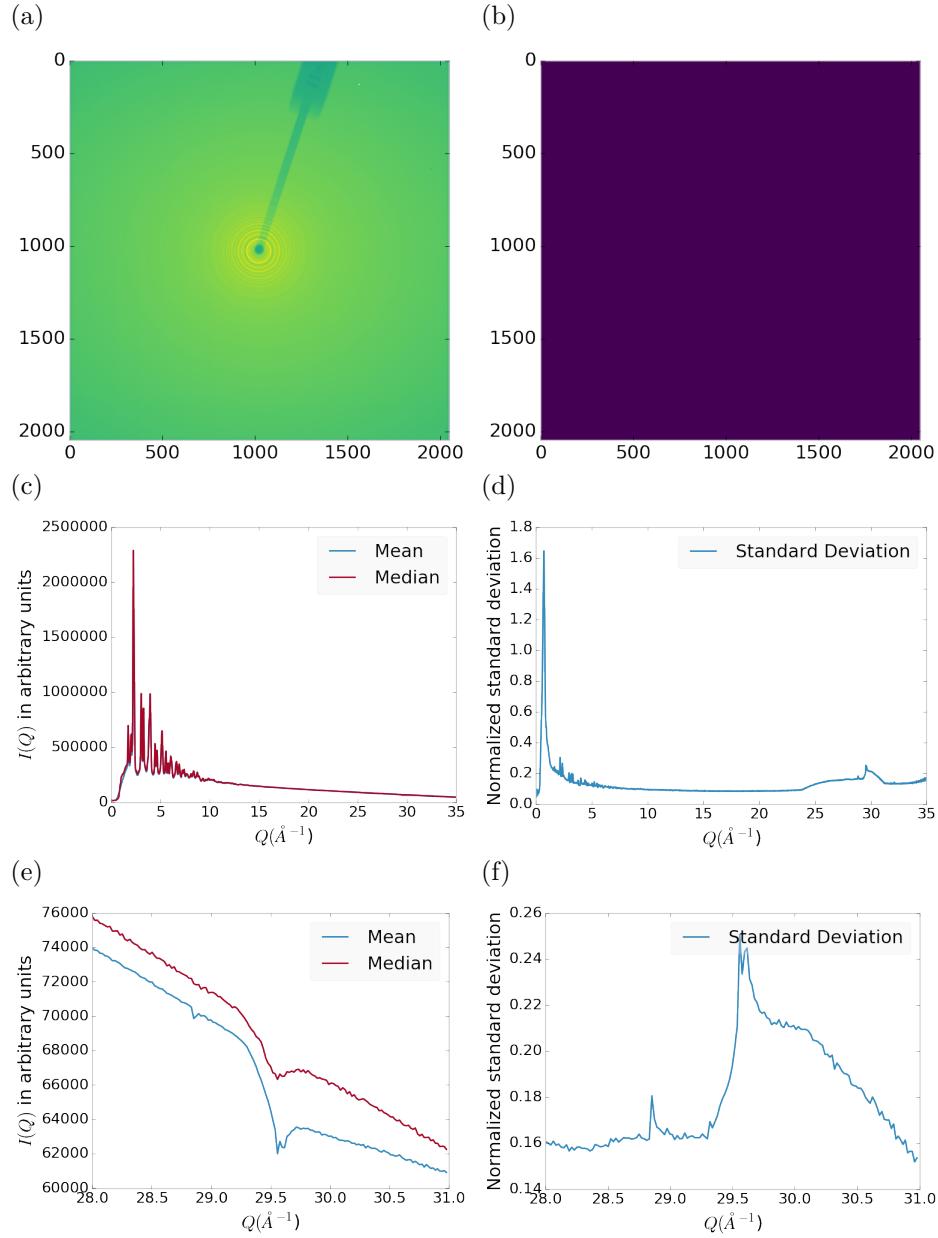


Figure 4.16: Masking, average, and standard deviation of an example x-ray total scattering measurement. This image was produced with no mask. a) the image, b) the mask, c) the mean and median values, d) the standard deviation (normalized to the median), e) a closeup of the 28 \AA^{-1} to 31 \AA^{-1}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

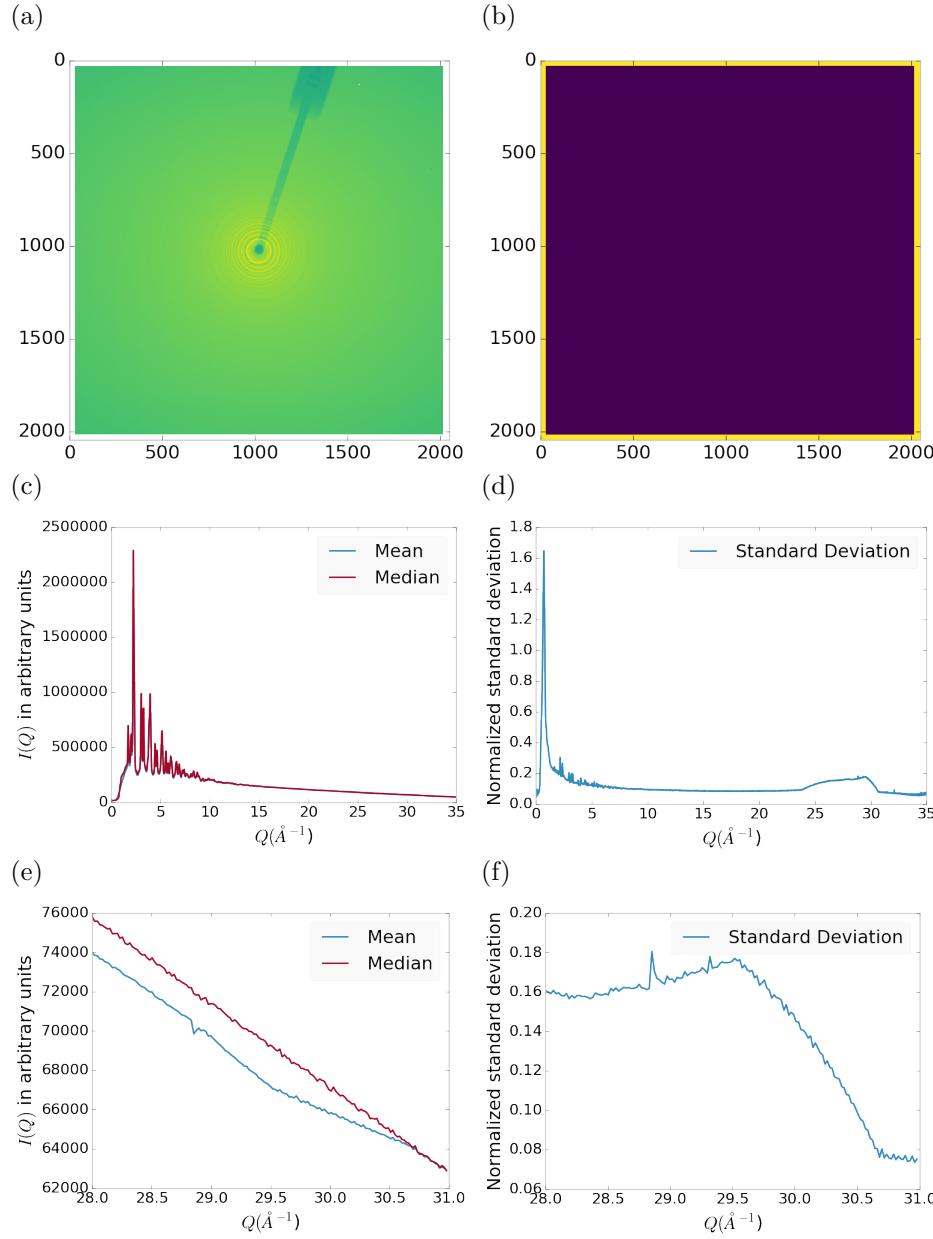


Figure 4.17: Masking, average, and standard deviation of an example x-ray total scattering measurement. This image was produced with only an edge mask. a) the image, b) the mask, c) the mean and median values, d) the standard deviation (normalized to the median), e) a closeup of the 28 \AA^{-1} to 31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

711 4.5 CONCLUSIONS

712 This chapter developed and analyzed the proper data processing and reduction method-
 713 ology for producing reliable $F(Q)$ data from x-ray total scattering measurements.

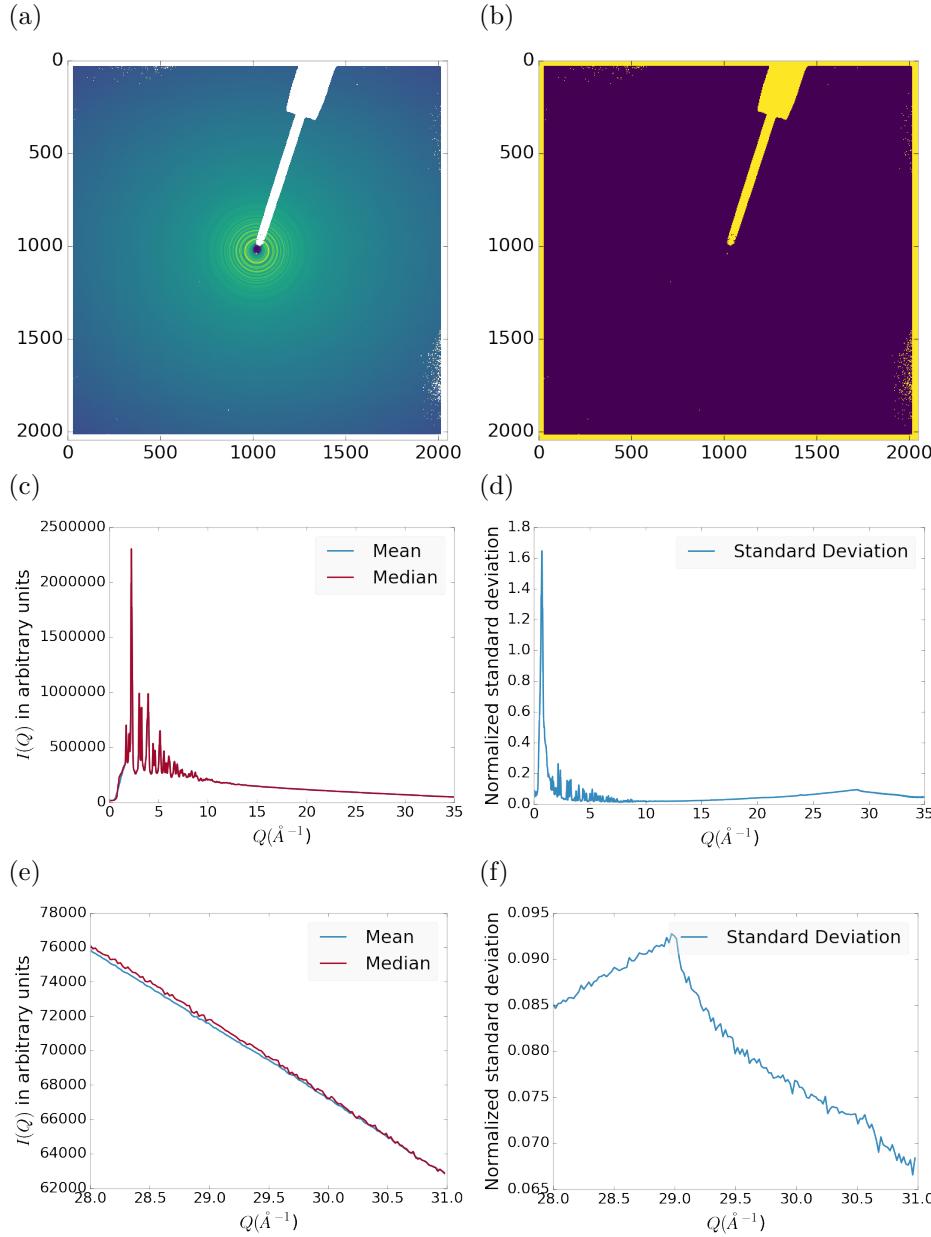


Figure 4.18: Masking, average, and standard deviation of an example x-ray total scattering measurement. This image was produced combining an edge mask and the automatically generated mask. a) the image, b) the mask, c) the mean and median values, d) the standard deviation (normalized to the median), 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

714 Binning at the Q resolution of the detector was found to be key to the data process-
 715 ing. The primary outcome of using the Q resolution binning was an enhancement in
 716 effectiveness for the masking algorithm, producing much fewer false positives for dead

717 pixels. This masking approach was then applied to the integration of experimental
718 data taken at the APD’s 11-ID-B beamline. The automatically generated masks, when
719 combined with edge masks, were found to greatly reduce the overall standard devi-
720 ation of the pixel intensity and produce a smoother $F(Q)$ at high Q , enabling the
721 use of much higher Q data in the PDF. Different statistical measures used in the
722 azimuthal integration was also compared. This comparison showed that the median
723 was a more reliable statistic for integration with data which had more detector de-
724 fects. However, upon properly masking it was shown that these metrics were almost
725 identical. The masking induced similarity between the mean and median shows that
726 the rings, when integrated, may form a Gaussian distribution. The distribution of
727 the pixel intensities for strongly and weakly scattering samples may be investigated
728 in future work.

729

CHAPTER 5

730

ANNEALING AND AGGREGATION OF 2NM

731

AU NANOPARTICLES

732 5.1 EXPERIMENTS

733 NP Synthesis

734 X-ray Total Scattering Measurements

735 5.2 DATA PROCESSING

736 5.3 DATA ANALYSIS

737 5.4 SIMULATION

738 5.5 STRUCTURAL ANALYSIS

739 5.6 CONCLUSIONS

740

CHAPTER 6

741

PHASE CHANGES AND ANNEALING DYNAMICS OF

742

Pr_2NiO_4 AND ITS DERIVATIVES

743 6.1 EXPERIMENTS

744 **Pr_2NiO_4 Synthesis**

745 **X-ray Measurements**

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

751 6.2 DATA PROCESSING

752

masking parameters

753

integration parameters

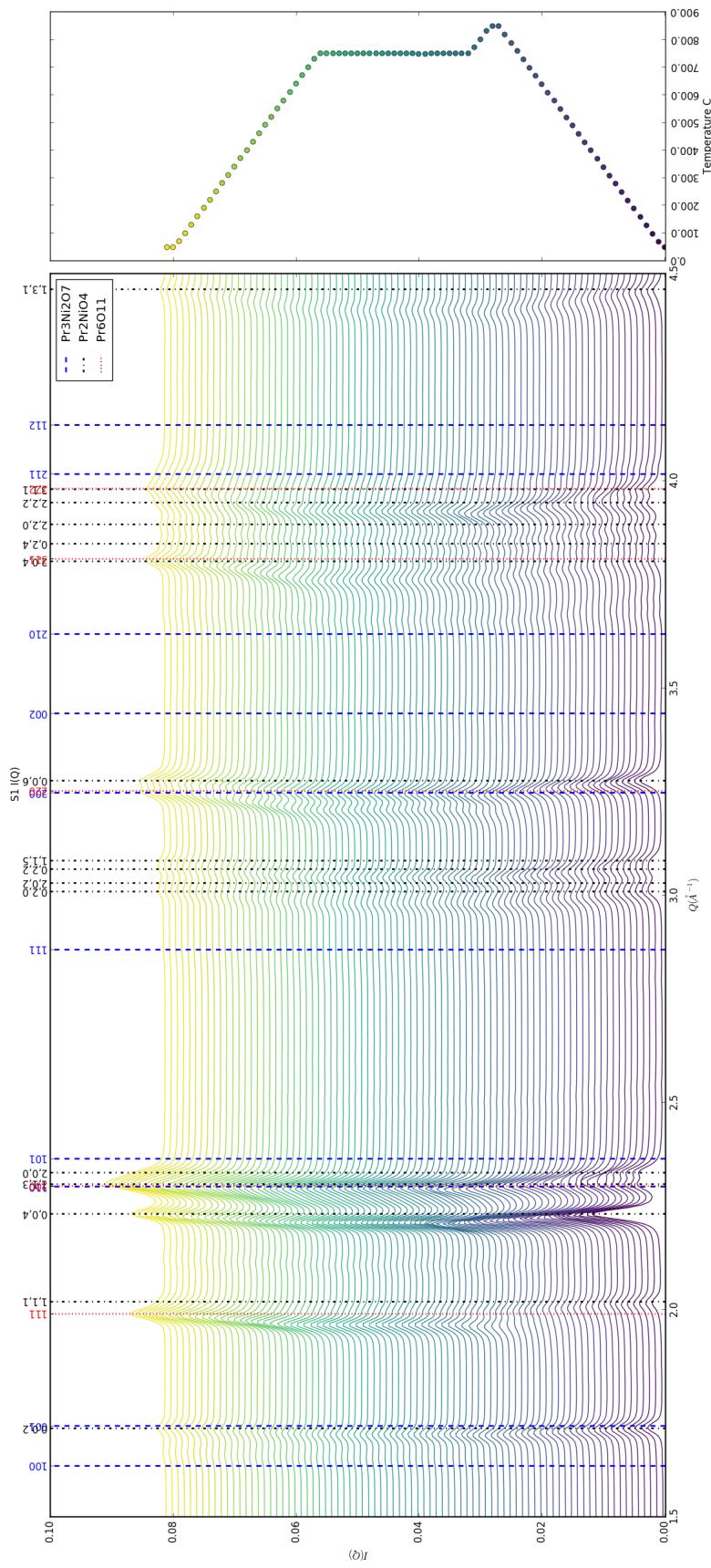
754

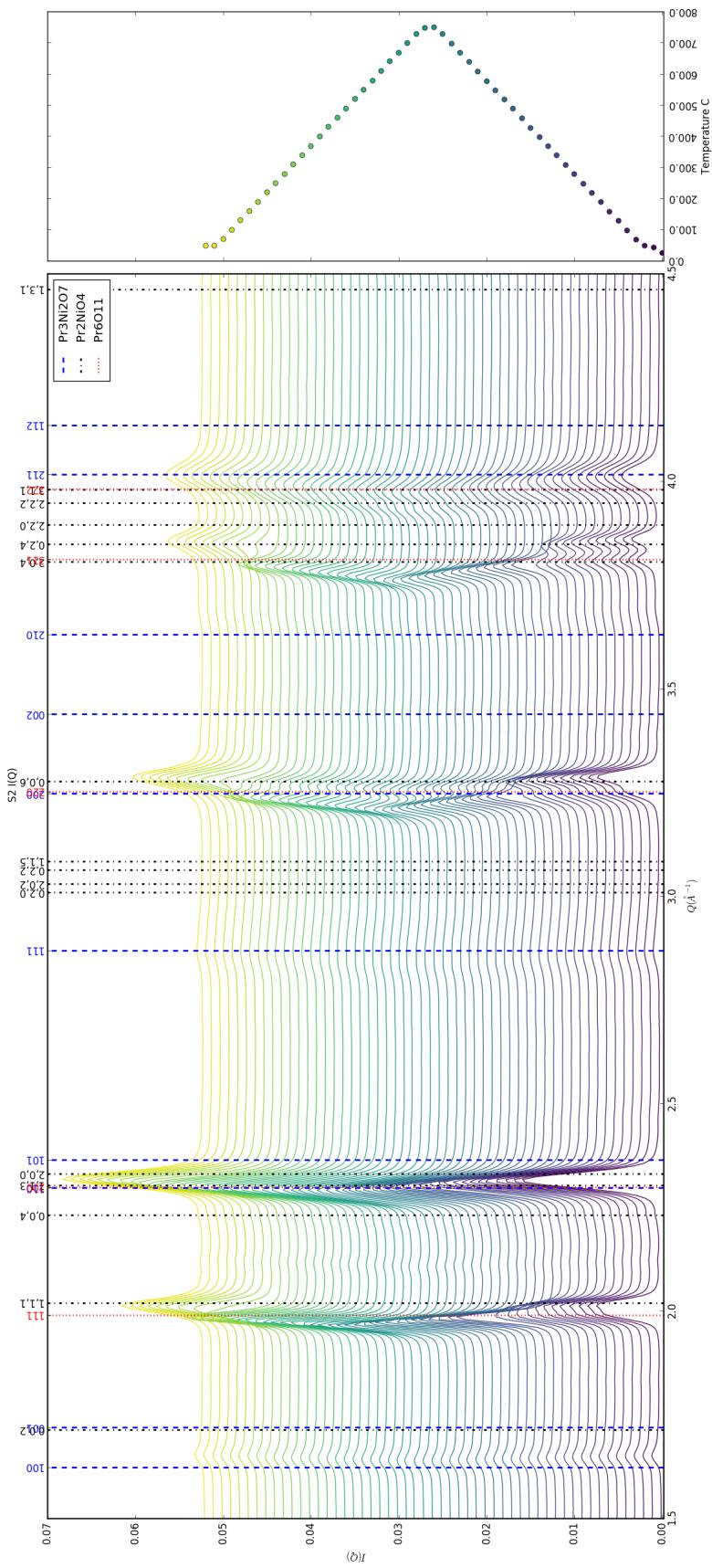
PDF parameters

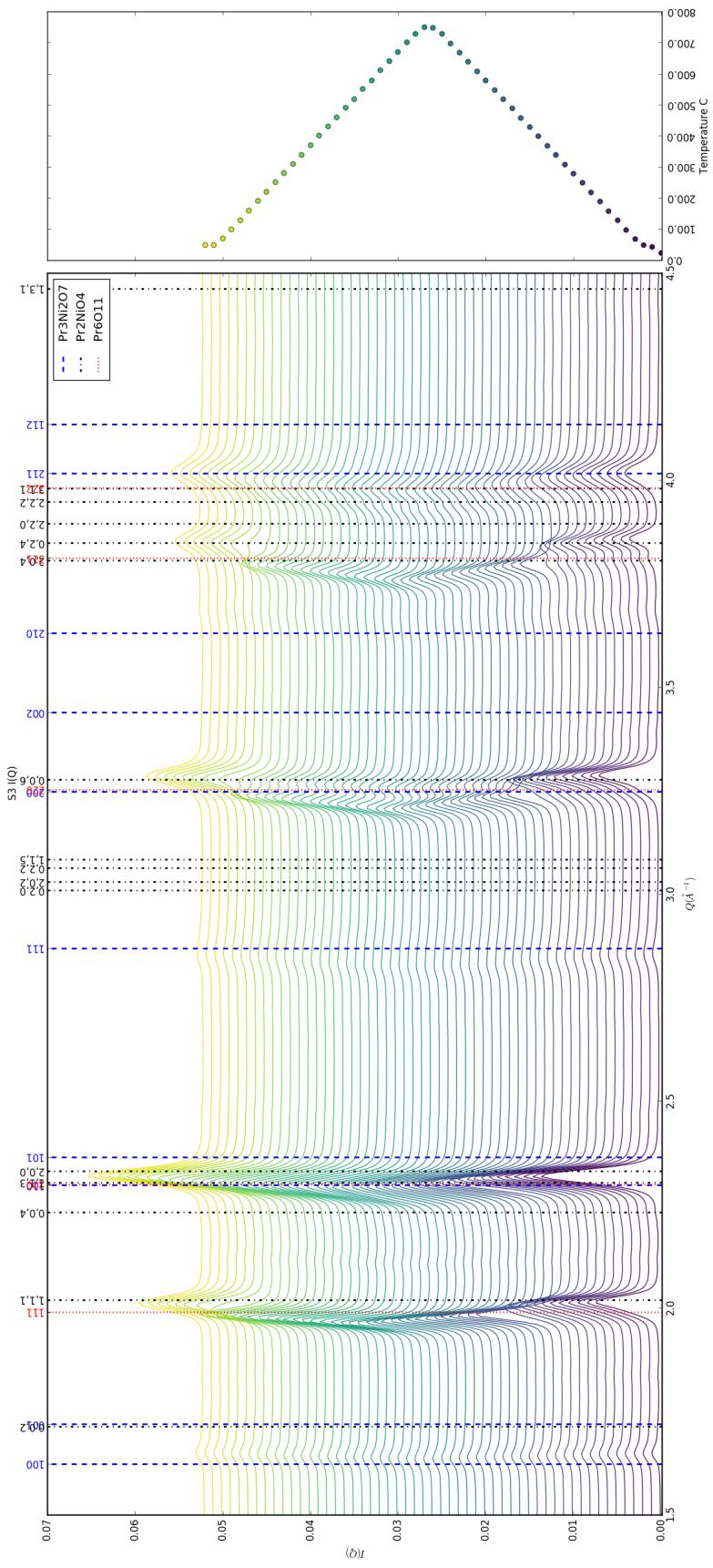
755 6.3 DATA ANALYSIS

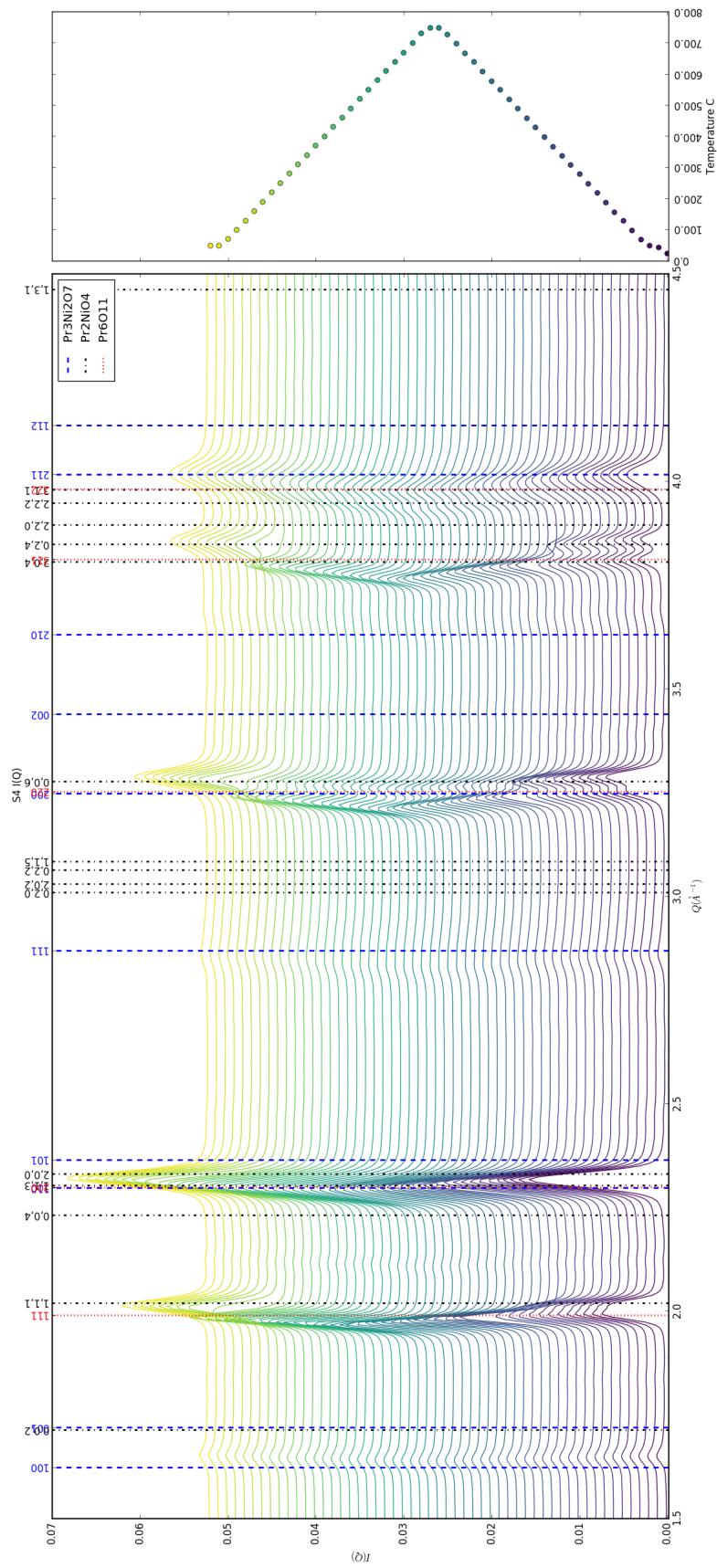
756 **Intra Sample Comparison**

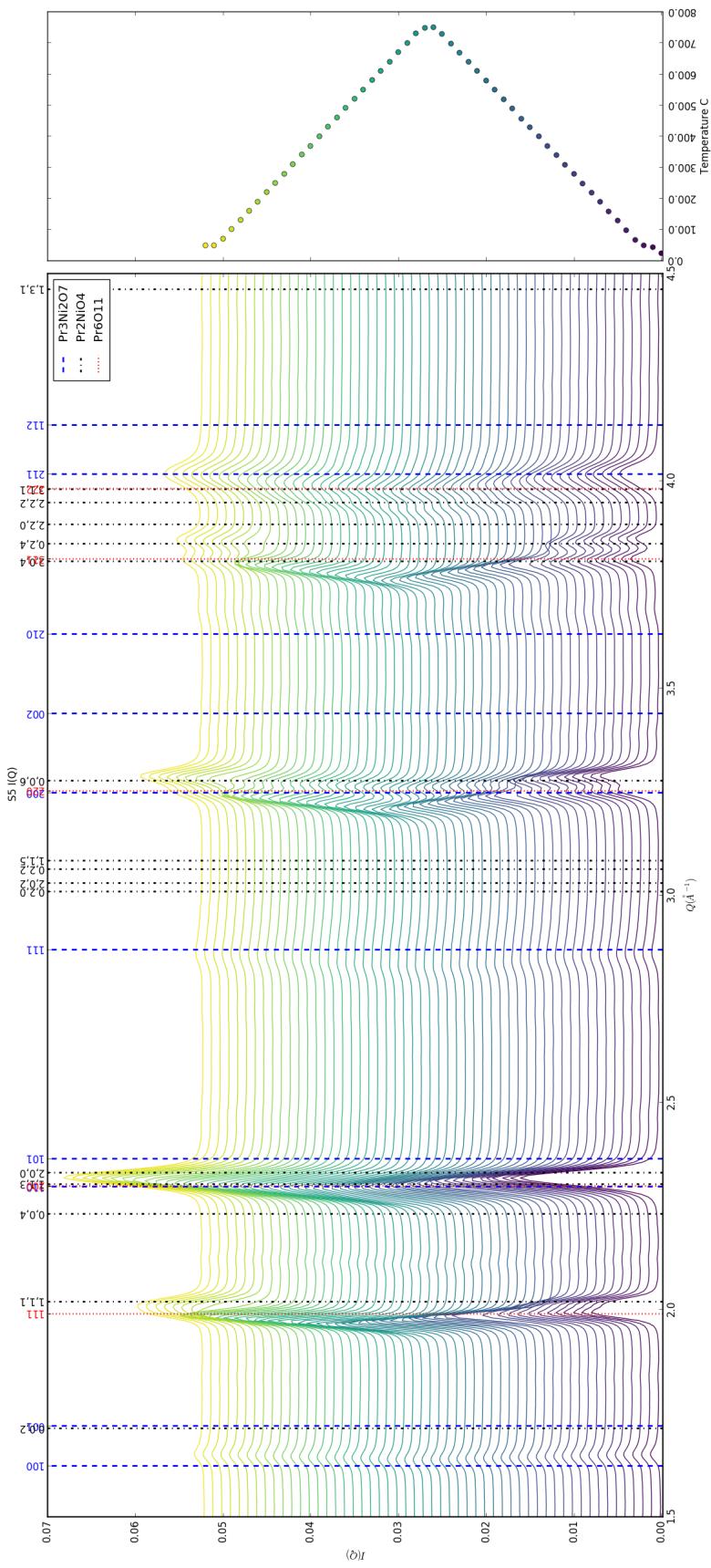
757 Changes in S1 but very little in S2-5.



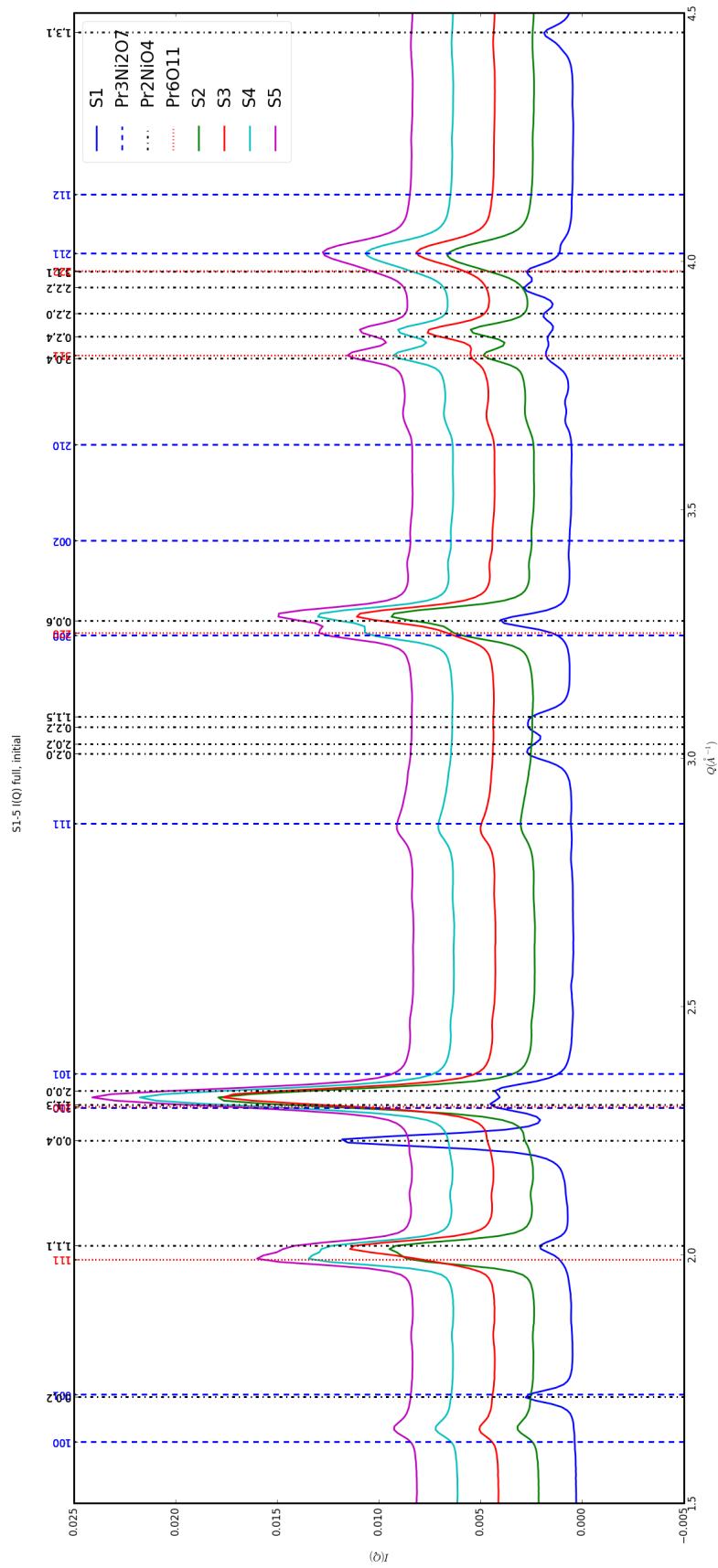


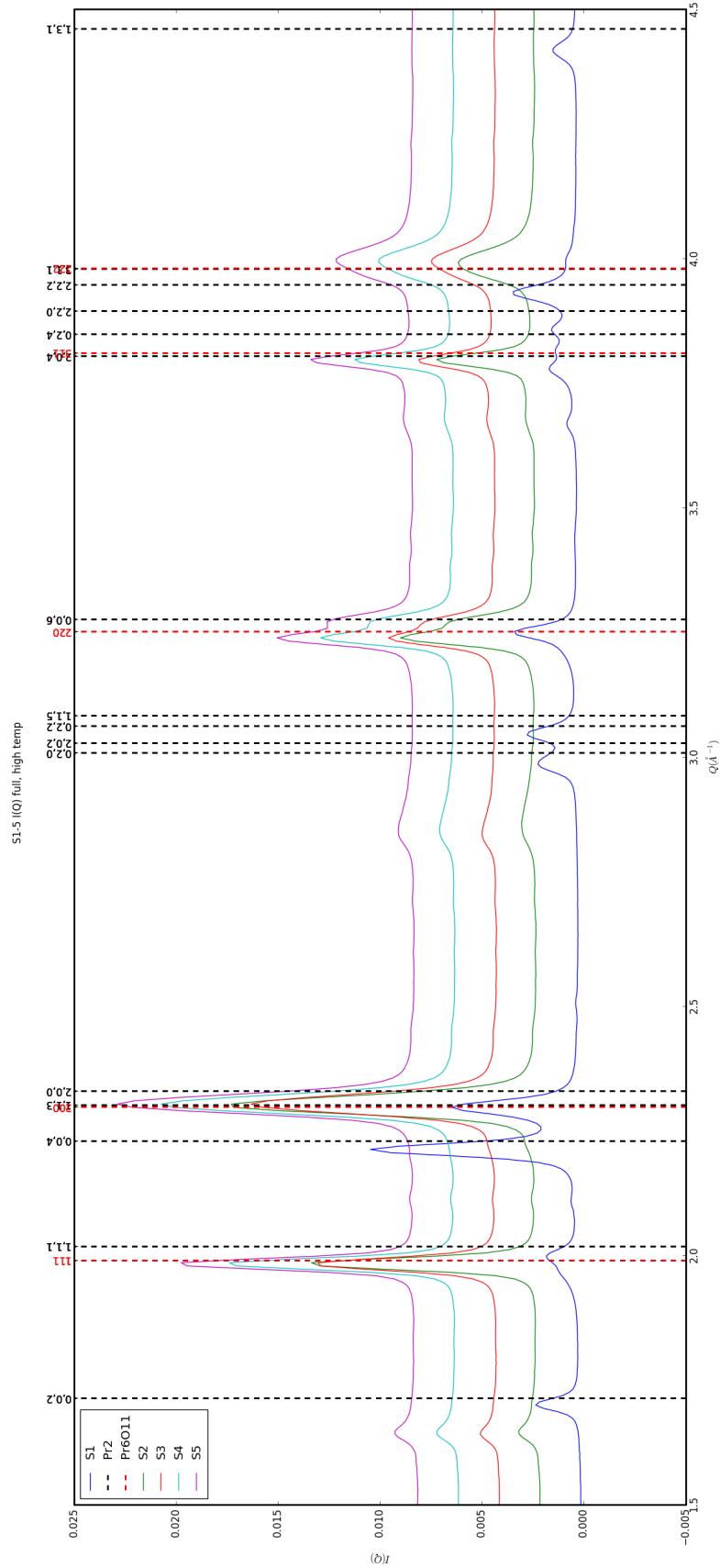


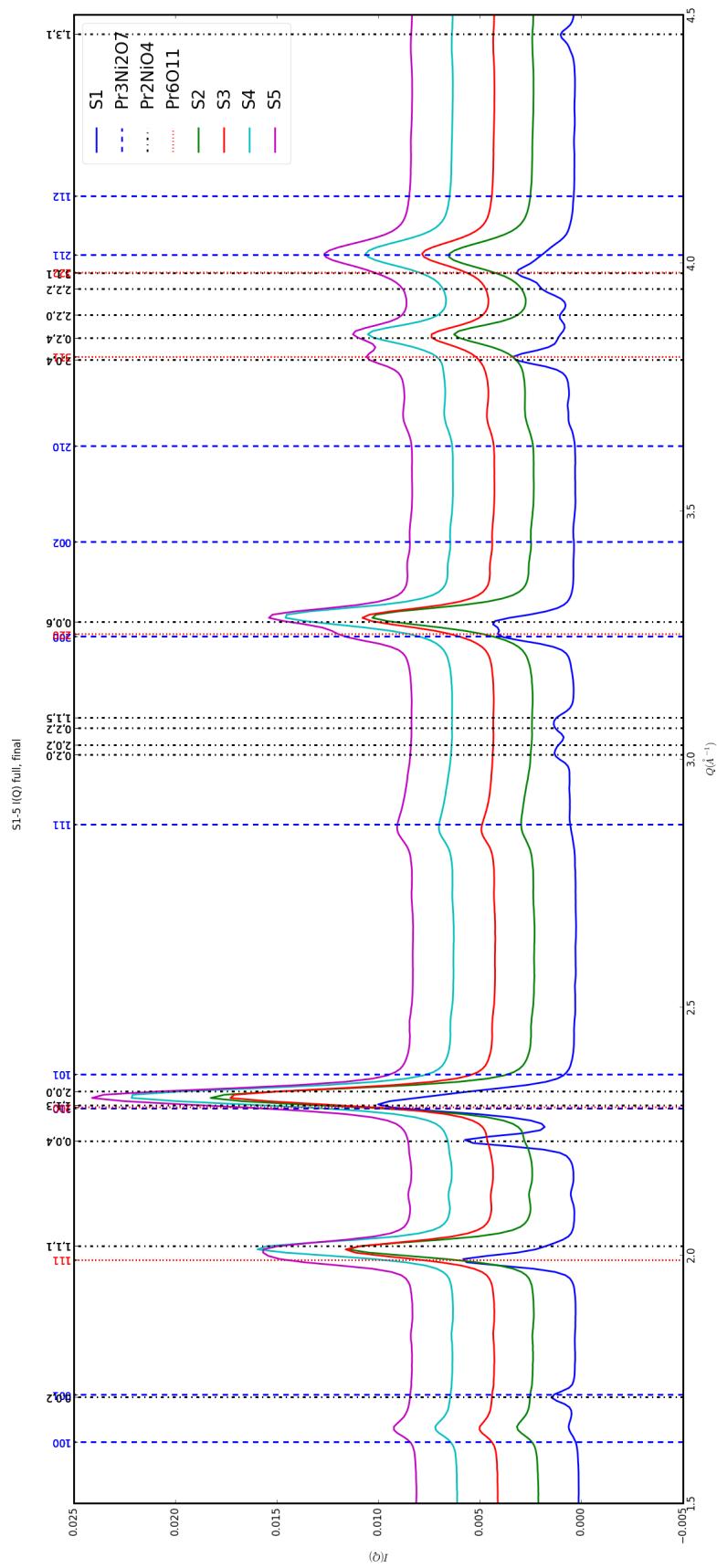


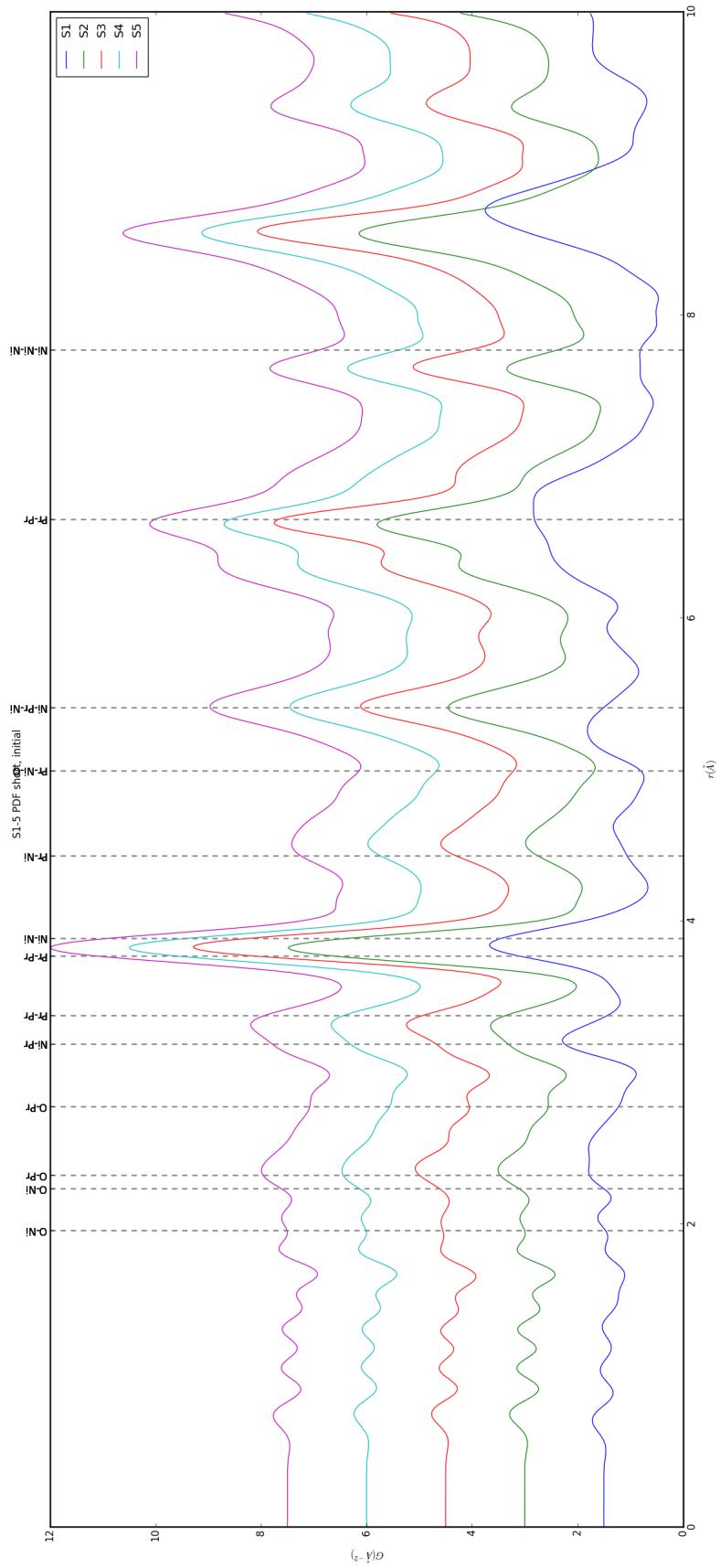


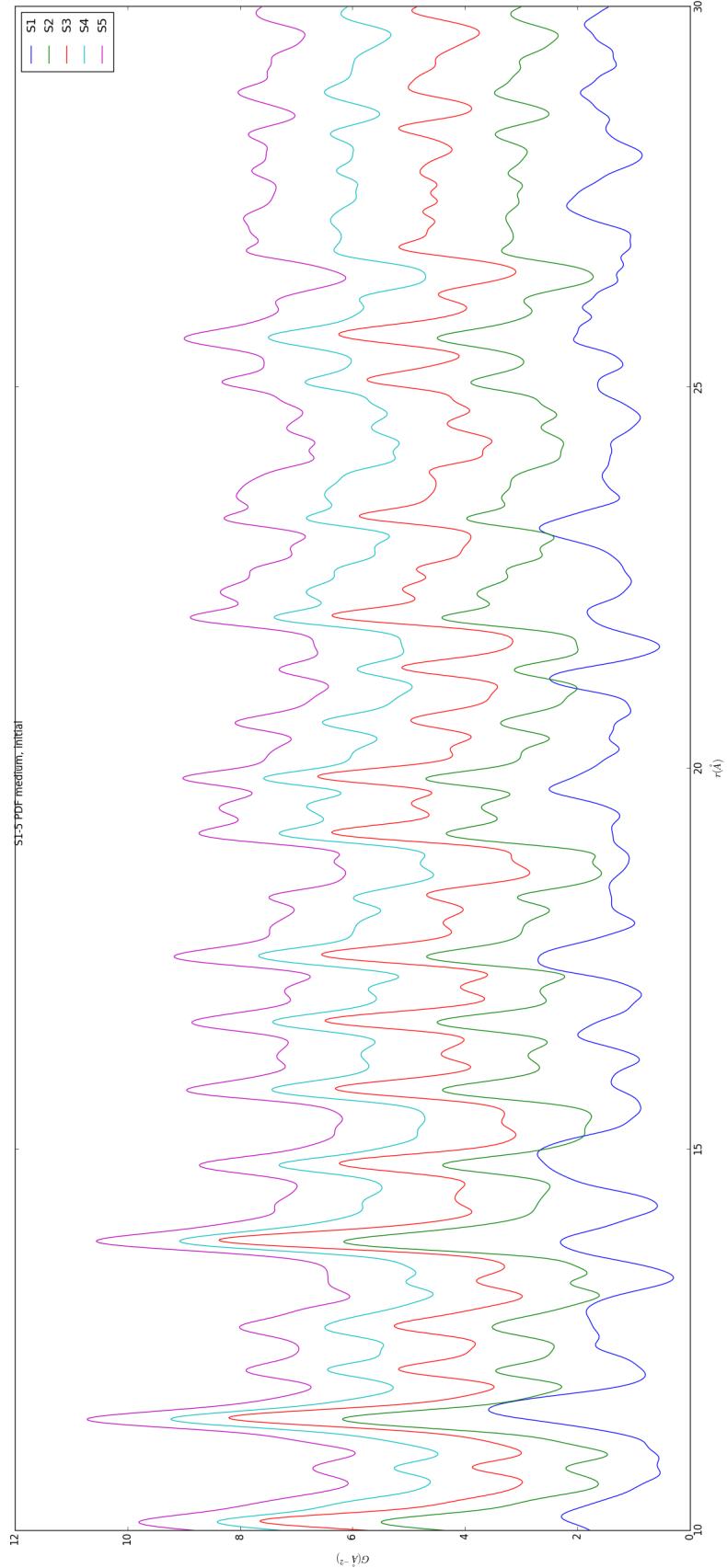
758 Inter Sample Comparison

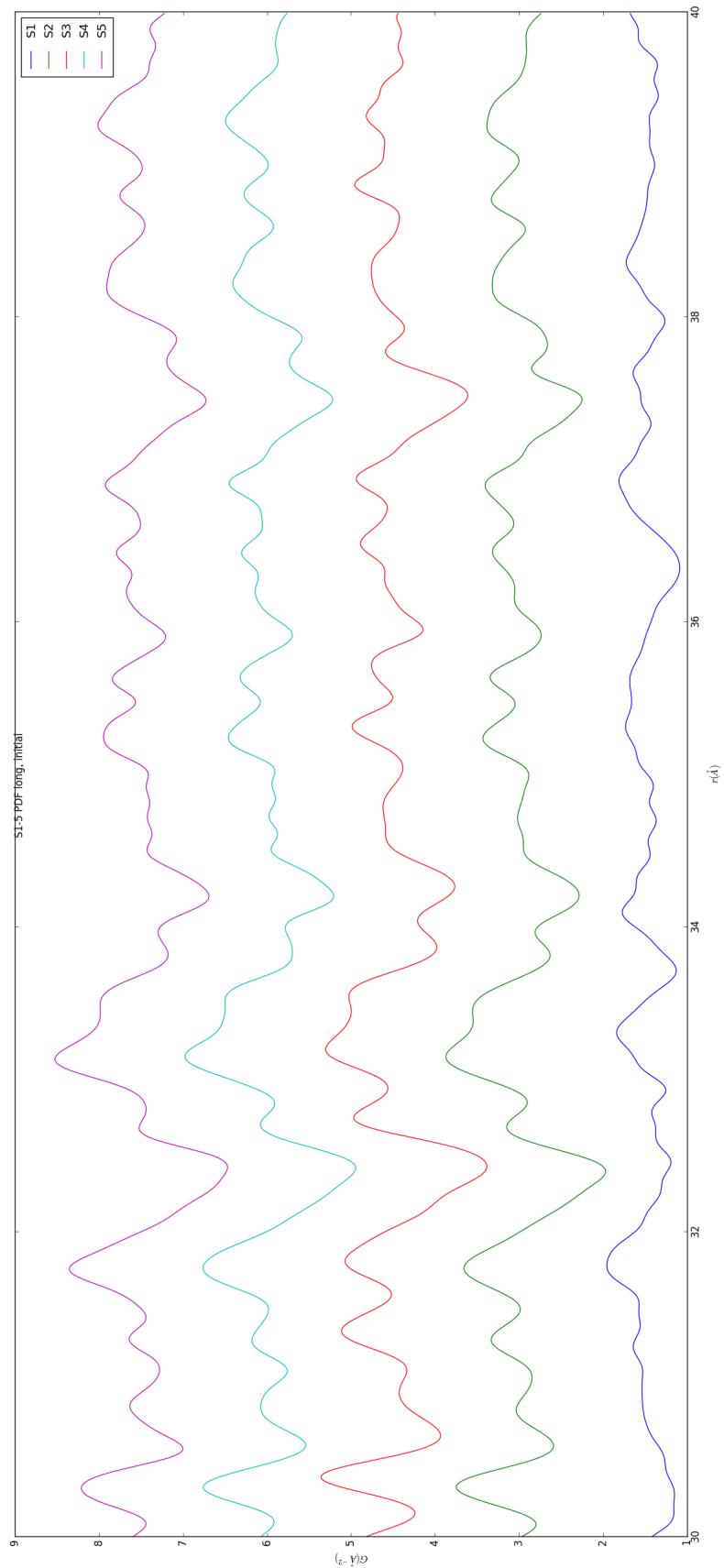


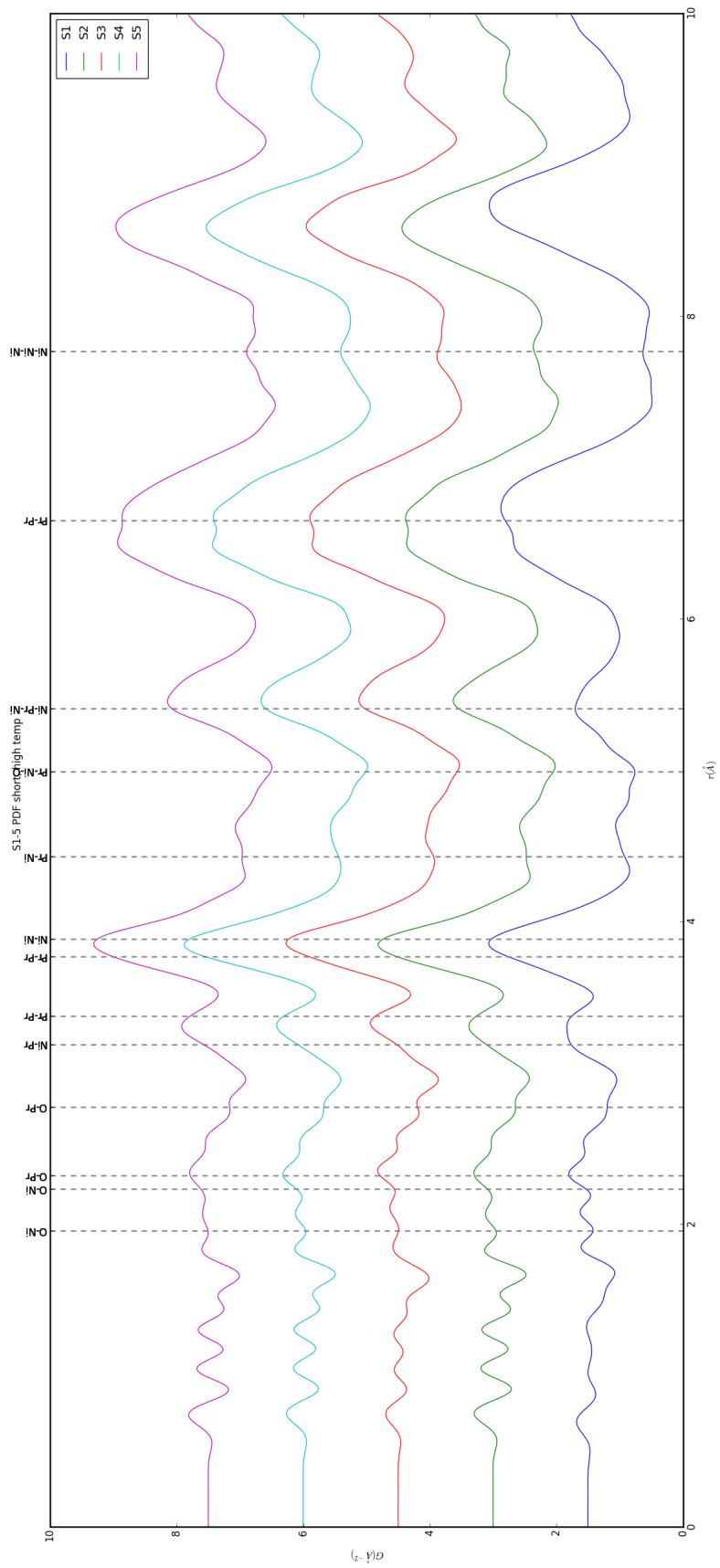


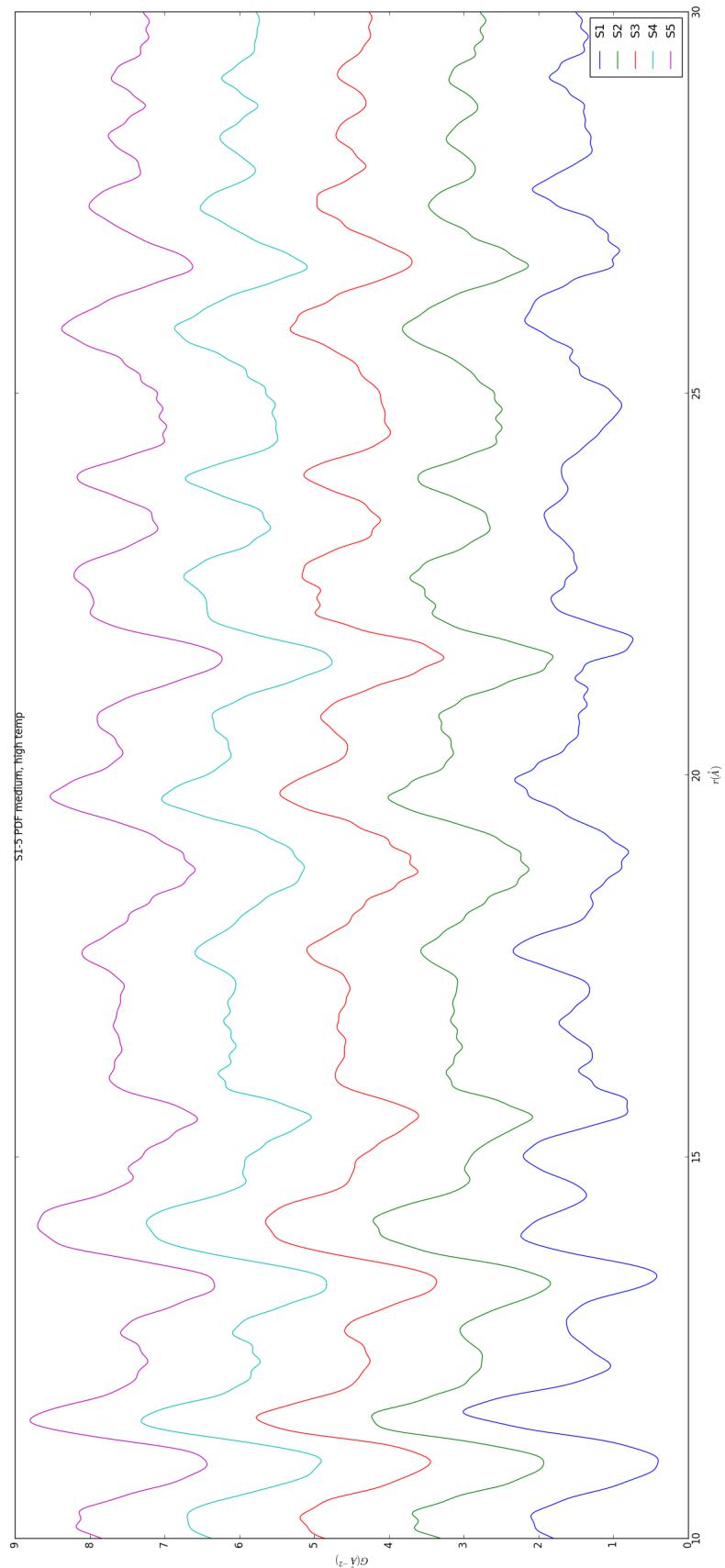


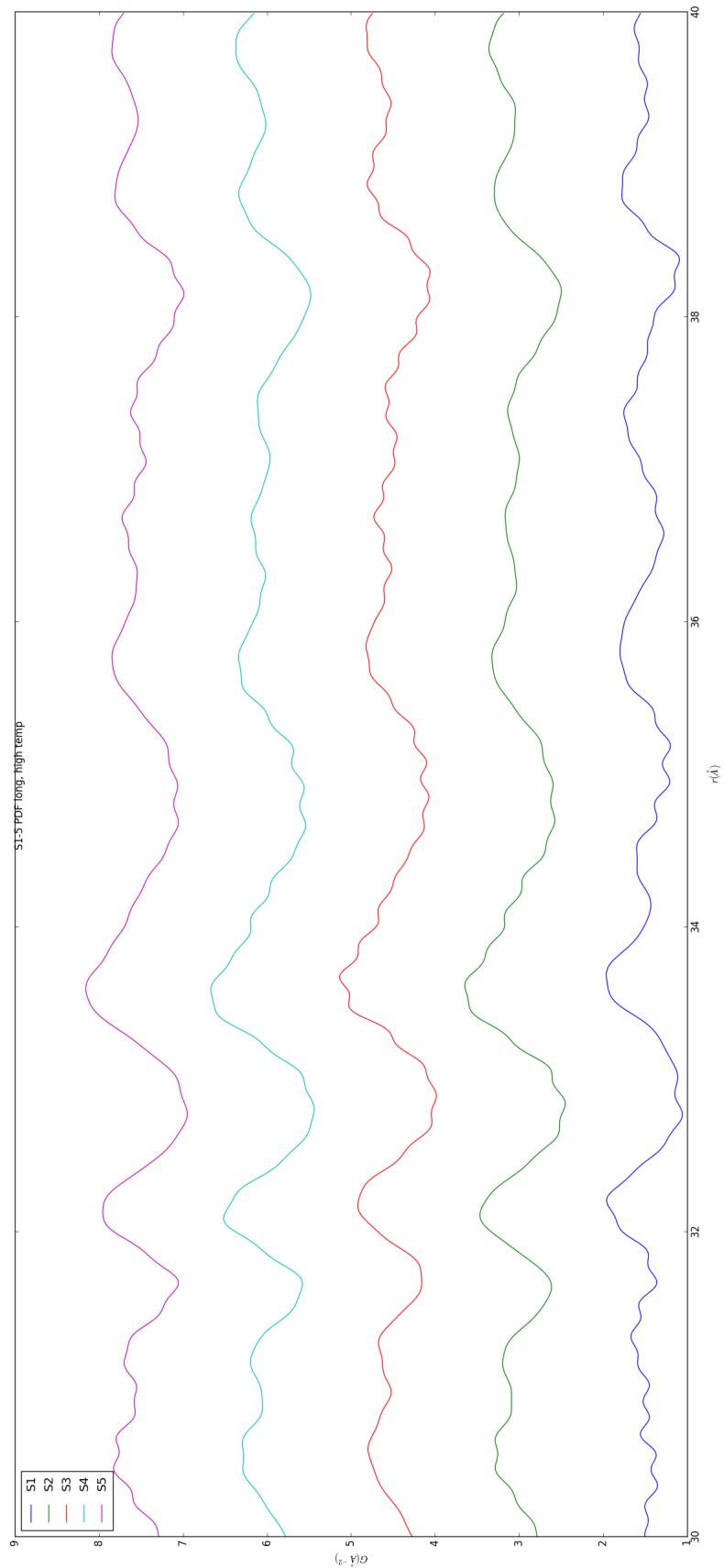


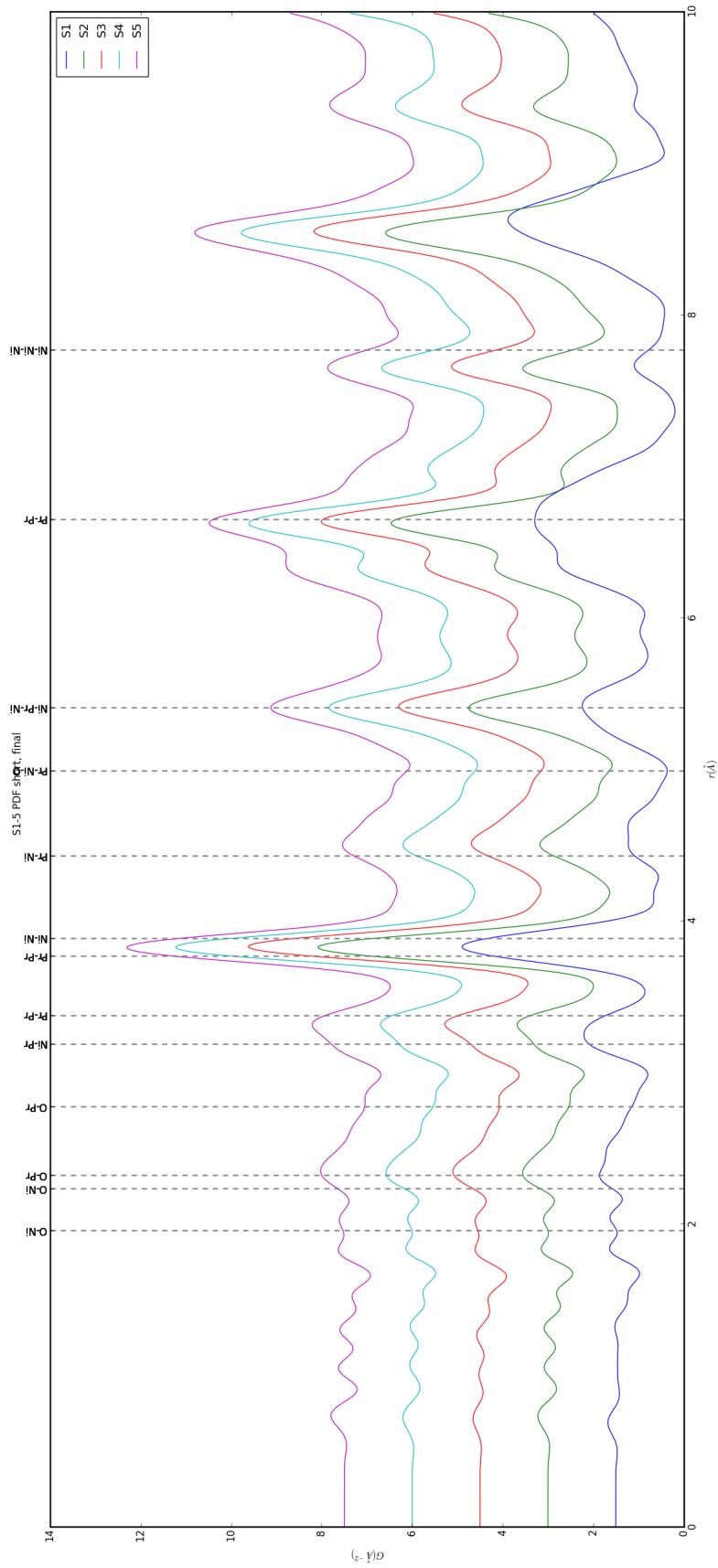


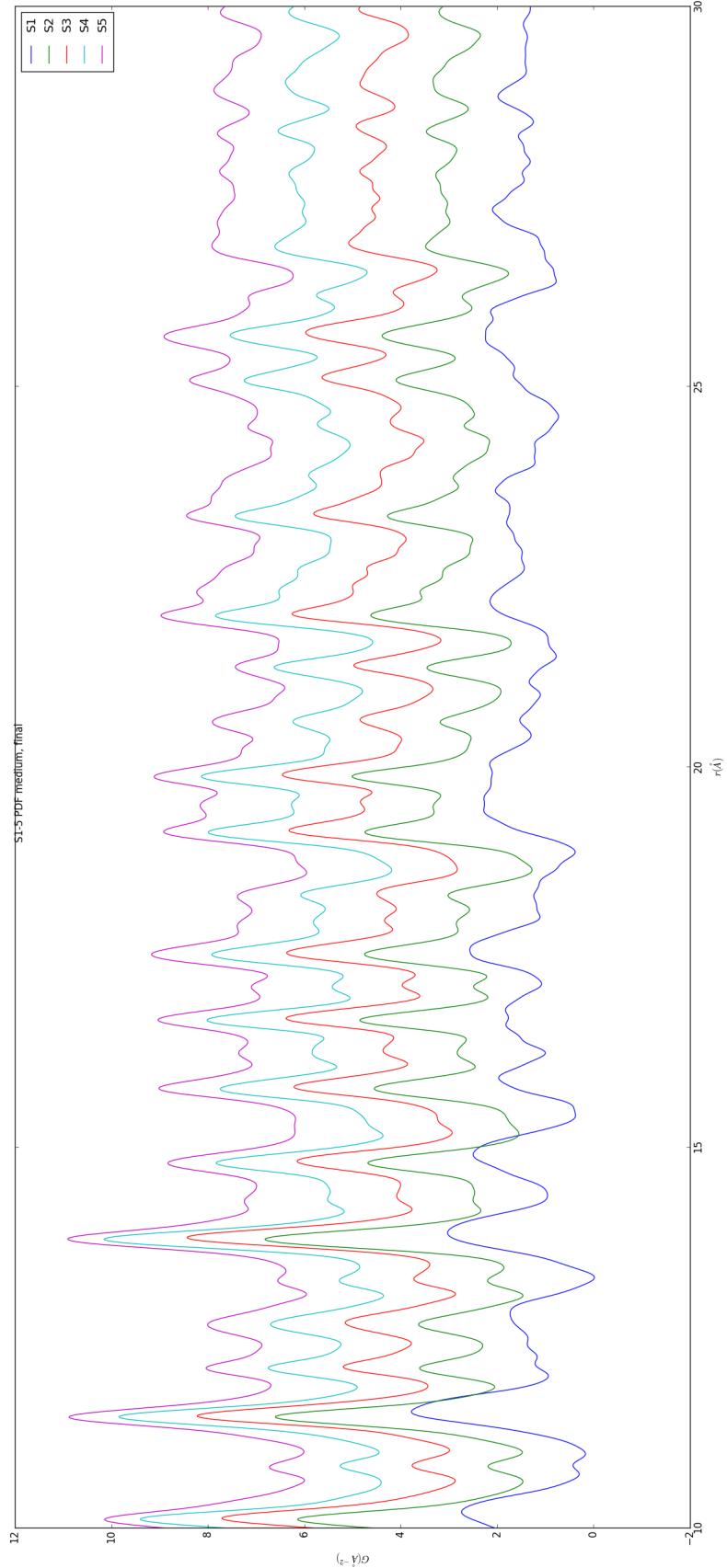


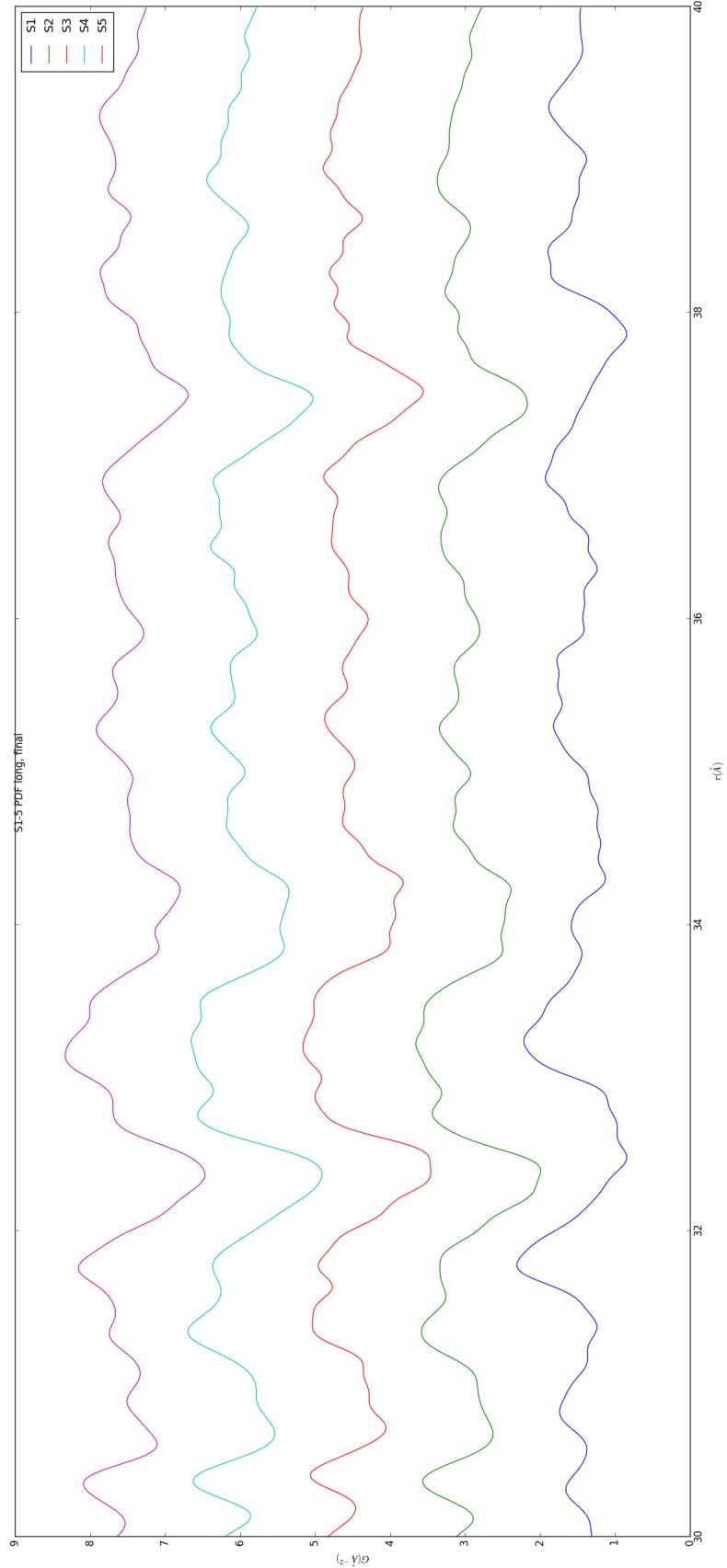












759 The PDF for S1 at operating temperature looks like S2-5 at the same temperature.

760 6.4 SIMULATION

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

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

773 6.5 CONCLUSIONS

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

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

793

CHAPTER 7

794

CONCLUSION

BIBLIOGRAPHY

- 796 [1] P. E. Blöchl, *Projector augmented-wave method*, Physical Review B **50** (1994),
797 no. 24, 17953–17979.
- 798 [2] Matthew J. Cliffe, Martin T. Dove, D. a. Drabold, and Andrew L. Goodwin,
799 *Structure determination of disordered materials from diffraction data*, Physical
800 Review Letters **104** (2010), no. 12, 1–4.
- 801 [3] Matthew J Cliffe and Andrew L Goodwin, *Nanostructure determination from the
802 pair distribution function: a parametric study of the INVERT approach.*, Journal
803 of physics. Condensed matter : an Institute of Physics journal **25** (2013), no. 45,
804 454218.
- 805 [4] Juarez L F Da Silva, Hyoung Gyu Kim, Maurício J. Piotrowski, Maurício J. Pri-
806 eto, and Germano Tremiliosi-Filho, *Reconstruction of core and surface nanopar-
807 ticles: The example of Pt 55 and Au55*, Physical Review B - Condensed Matter
808 and Materials Physics **82** (2010), no. 20, 1–6.
- 809 [5] C L Farrow, P Juhas, J W Liu, D Bryndin, E S Božin, J Bloch, Th Proffen, and
810 S J L Billinge, *PDFfit2 and PDFgui: computer programs for studying nanostruc-
811 ture in crystals.*, Journal of Physics. Condensed Matter : an Institute of Physics
812 journal **19** (2007), no. 33, 335219.
- 813 [6] Christopher L Farrow and Simon J L Billinge, *Relationship between the atomic
814 pair distribution function and small-angle scattering: implications for modeling
815 of nanoparticles.*, Acta Crystallographica Section A Foundations of Crystallog-
816 graphy **65** (2009), no. Pt 3, 232–9 (en).
- 817 [7] Riccardo Ferrando, Julius Jellinek, and Roy L Johnston, *Nanoalloys: From The-
818 ory to Applications of Alloy Clusters and Nanoparticles*, Chemical Reviews **108**
819 (2008), no. 3, 846–904.
- 820 [8] Pablo D Jadzinsky, Guillermo Calero, Christopher J Ackerson, David A Bushnell,
821 and Roger D Kornberg, *Structure of a thiol monolayer-protected gold nanoparti-
822 cle at 1.1 Å resolution.*, Science (New York, N.Y.) **318** (2007), no. 5849, 430–433.

- 823 [9] G. Kresse and J. Hafner, *Ab initio molecular dynamics for liquid metals*, Physical
824 Review B **47** (1993), no. 1, 558–561.
- 825 [10] ———, *Ab initio molecular-dynamics simulation of the liquid-
826 metalâŠamorphous-semiconductor transition in germanium*, Physical Review
827 B **49** (1994), no. 20, 14251–14269.
- 828 [11] Yan Li, Giulia Galli, and Fran ois Gygi, *Electronic structure of thiolate-covered
829 gold nanoparticles: Au102(MBA)44*, ACS Nano **2** (2008), no. 9, 1896–1902.
- 830 [12] L D Marks, *Experimental studies of small particle structures*, Reports on
831 Progress in Physics **57** (1994), no. 6, 603–649 (en).
- 832 [13] John P. Perdew, Kieron Burke, and Matthias Ernzerhof, *Generalized Gradient
833 Approximation Made Simple*, Physical Review Letters **77** (1996), no. 18, 3865–
834 3868.
- 835 [14] Andrew A. Peterson, *Global optimization of adsorbate-surface structures while
836 preserving molecular identity*, Topics in Catalysis **57** (2014), no. 1-4, 40–53.
- 837 [15] Valeri Petkov, Shiyao Shan, Peter Chupas, Jun Yin, Lefu Yang, Jin Luo, and
838 Chuan-Jian Zhong, *Noble-transition metal nanoparticle breathing in a reactive
839 gas atmosphere.*, Nanoscale **5** (2013), no. 16, 7379–87.
- 840 [16] H. W. Sheng, M. J. Kramer, A. Cadien, T. Fujita, and M. W. Chen, *Highly
841 optimized embedded-atom-method potentials for fourteen FCC metals*, Physical
842 Review B - Condensed Matter and Materials Physics **83** (2011), no. 13, 134118.