

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

Lacy Ford, Vice Provost and Dean of Graduate Studies

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

DEDICATION

To Diane & Donald Wright

My first scientific advisers

To see a World in a Grain of Sand

ACKNOWLEDGMENTS

This work would not have been possible without my entire ensemble of colleagues, advisors, friends, and family. To Dr. Sanjit Ghose thank you very much for getting me started in the x-ray field and sticking with me, even when I went into the mathematical deep end. To Dr. Yan Li, thank you very much for always pushing my simulations to be the best they could be, and pusing me to search every possible nook and cranny of parameter space.

I'd also like to thank the NSLS-II XPD team for their experimental, intelecual, and moral support. To Dr. Thomas Caswell, Dr. Eric Dill, and Dr. Dan Allen, I would not be $1e^{-6}$ of the programmer I am today with out your guidence, conversation, and explinations.

To Prof. Scopatz, thank you very much for your insights, conversation, and being a spectacular developer role model.

To my many readers, I'd like to say thank you, if my thesis hasn't put you to sleep already.

To my parents, I don't know if any language could correctly render the magnitude of my gratitude for everything you have done for me.

Last, but certainly not least, Prof. Zhou. I am certain that I could not have gotten a better start in research science than working with you.

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 effectiveness of the system. Experimental data processing proce-

dures, 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 FIGURES	x
TODO LIST	1
CHAPTER 1 INTRODUCTION	2
CHAPTER 2 STATISTICAL MECHANICAL ENSEMBLES AND POTENTIAL ENERGY SURFACES	4
2.1 Introduction	4
2.2 Potential Energy Surfaces	4
Experimentally Derived Potential Energy Surfaces	5
Potentials	5
Forces	6
2.3 Ensembles	7
Monte Carlo Modeling	8
Hamiltonian Monte Carlo	9
No-U-Turn Sampling	11
Grand Canonical Ensemble	11
Ensemble description	12
Grand Canonical Monte Carlo	12
GCMC biasing	13
2.4 Conclusions	16

CHAPTER 3 ATOMIC PAIR DISTRIBUTION FUNCTION: THEORY AND COMPUTATION	17
3.1 Introduction	17
3.2 Theory	18
Derivation	18
Analytically Gradients	19
Without ADPs	20
Periodic Boundary Conditions	21
3.3 Computation	21
HPC and GPUs	22
GPUs and Parallelization	22
Map from ij space to k space	23
GPU Memory Allocation	25
Speed and Scaling of PDF Computation	27
3.4 Conclusions	28
CHAPTER 4 BENCHMARKS	29
4.1 Introduction	29
Target Setup	30
Model Parameters	31
4.2 Structural Solutions	32
Case I: crystalline Au ₅₅	32
Case II: Au ₅₅ with surface disorder	33
Case III: amorphous Au ₅₅	35
Case IV: ligand-protected Au ₁₀₂	37
Starting from fcc structure	37
Starting from Marks Decahedron	39
4.3 Discussion and Conclusion	40
CHAPTER 5 X-RAY TOTAL SCATTERING DATA ACQUISITION AND PRO- CESSING	43
5.1 Introduction	43
5.2 Detector Q resolution	43

5.3	Automated Mask Generation	46
	Introduction	46
	Algorithm Design	47
	Test Cases	47
	Results and Discussion	48
	Conclusions	54
5.4	Automated Image Azimuthal Integration	55
5.5	Conclusions	60
 CHAPTER 6 PHASE CHANGES AND ANNEALING DYNAMICS OF Pr_2NiO_4 AND ITS DERIVATIVES		61
6.1	Introduction	61
6.2	Experiments	62
	Pr_2NiO_4 Synthesis	62
	X-ray Measurements	62
6.3	Data Processing	62
6.4	Data Analysis	63
	Intra Sample Comparison	63
	PDF	63
	$I(Q)$	68
	Inter Sample Comparison	73
6.5	Simulation	76
6.6	Conclusions	76
 CHAPTER 7 CONCLUSION		78
 BIBLIOGRAPHY		81
 APPENDIX A SUPPLEMENTAL INFORMATION: PHASE CHANGES AND AN- NEALING DYNAMICS OF Pr_2NiO_4 AND ITS DERIVATIVES . . .		87
	Intra Sample Comparison	87
	Inter Sample Comparison	100

LIST OF FIGURES

Figure 2.1 Addition biasing with a Lennard Jones potential.	15
Figure 3.1 Comparison of the CPU and GPU chip architectures	22
Figure 3.2 Speed comparison of CPU and GPU implementations	27
Figure 4.1 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, c) the radial bond distribution, and d) bond angle distribution.	33
Figure 4.2 Au ₅₅ PDF fitting of surface-disordered Au ₅₅ . (a) the comparison of PDFs, (b) the CN distribution, (c) the target structure, (d) the radial bond distribution, (e) the bond angle distribution, and (f) the final structural solution ($Rw=0.6\%$)	34
Figure 4.3 Similar to figure 4.2 for DFT-optimized amorphous Au ₅₅	36
Figure 4.4 Similar to Fig. 4.2 for Au ₁₀₂ as in DFT-optimized Au ₁₀₂ MBA ₄₄ cluster.	38
Figure 4.5 Similar to Fig. 4.4 with Marks decahedron as the starting structure.	39
Figure 5.1 Scattering onto a flat detector	44
Figure 5.2 Q resolution as a function of Q	45
Figure 5.3 Number of pixels as a function of Q , binned at the Q resolution of the detector.	46
Figure 5.4 Generated dead/hot pixel masks for a detector with 100 bad pixels.	49
Figure 5.5 Generated dead/hot pixel masks for a detector with 300 bad pixels.	49

Figure 5.6 Generated dead/hot pixel masks for a detector with 500 bad pixels.	50
Figure 5.7 Generated dead/hot pixel masks for a detector with 1000 bad pixels.	50
Figure 5.8 Generated beamstop holder masks for a beamstop holder with 10% transmittance.	51
Figure 5.9 Generated beamstop holder masks for a beamstop holder with 30% transmittance.	51
Figure 5.10 Generated beamstop holder masks for a beamstop holder with 50% transmittance.	51
Figure 5.11 Generated beamstop holder masks for a beamstop holder with 90% transmittance.	52
Figure 5.12 Generated beamstop holder masks which is rotated away from vertical	52
Figure 5.13 Masked experimental data.	53
Figure 5.14 Masked experimental data with Pt single crystal signal.	53
Figure 5.15 Masked experimental data with Pt single crystal signal using figure's 5.13 mask as a starting mask.	54
Figure 5.16 Masking, average, and standard deviation of an example x-ray total scattering measurement with with no mask.	57
Figure 5.17 Masking, average, and standard deviation of an example x-ray total scattering measurement with with only an edge mask.	58
Figure 5.18 Masking, average, and standard deviation of an example x-ray total scattering measurement with combining an edge mask and the automatically generated mask.	59
Figure 6.1 PDF as a function of temperature for as synthesized PNO show- ing the full PDF	64
Figure 6.2 PDF as a function of temperature for as synthesized PNO show- ing a close up on the short range section	65
Figure 6.3 PDF as a function of temperature for PNO annealed at 750 °C for 25 hours showing the full PDF	66

Figure 6.4 PDF as a function of temperature for PNO annealed at 750 °C for 25 hours showing a close up on the short range section	67
Figure 6.5 $I(Q)$ as a function of temperature for as synthesized PNO showing the full XRD	69
Figure 6.6 $I(Q)$ as a function of temperature for as synthesized PNO showing a close up on the low Q section	70
Figure 6.7 $I(Q)$ as a function of temperature for PNO annealed at 750 °C for 25 hours showing the full XRD	71
Figure 6.8 $I(Q)$ as a function of temperature for PNO annealed at 750 °C for 25 hours showing a close up on the low Q section	72
Figure 6.9 Comparison of PNO sample PDFs as a function of annealing time high-temp	74
Figure 6.10 Comparison of PNO sample $I(Q)$ as a function of annealing time high-temp	75
Figure A.1 PDF as a function of temperature for PNO annealed at 750 °C for 50 hours showing the full PDF	88
Figure A.2 PDF as a function of temperature for PNO annealed at 750 °C for 50 hours showing a close up on the short range section	89
Figure A.3 PDF as a function of temperature for PNO annealed at 750 °C for 100 hours showing the full PDF	90
Figure A.4 PDF as a function of temperature for PNO annealed at 750 °C for 100 hours showing a close up on the short range section	91
Figure A.5 PDF as a function of temperature for PNO annealed at 750 °C for 200 hours showing the full PDF	92
Figure A.6 PDF as a function of temperature for PNO annealed at 750 °C for 200 hours showing a close up on the short range section	93
Figure A.7 $I(Q)$ as a function of temperature for PNO annealed at 750 °C for 50 hours showing the full XRD	94
Figure A.8 $I(Q)$ as a function of temperature for PNO annealed at 750 °C for 50 hours showing a close up on the low Q section	95

Figure A.9 $I(Q)$ as a function of temperature for PNO annealed at 750 °C for 100 hours showing the full XRD	96
Figure A.10 $I(Q)$ as a function of temperature for PNO annealed at 750 °C for 100 hours showing a close up on the low Q section	97
Figure A.11 $I(Q)$ as a function of temperature for PNO annealed at 750 °C for 200 hours showing the full XRD	98
Figure A.12 $I(Q)$ as a function of temperature for PNO annealed at 750 °C for 200 hours showing a close up on the low Q section	99
Figure A.13 Comparison of PNO sample PDFs as a function of annealing time at room temperature	101
Figure A.14 Comparison of PNO sample PDFs as a function of annealing time at room temperature	102
Figure A.15 Comparison of PNO sample PDFs as a function of annealing time at room temperature	103
Figure A.16 Comparison of PNO sample PDFs as a function of annealing time at room temperature	104
Figure A.17 Comparison of PNO sample PDFs as a function of annealing time at operating temperature	105
Figure A.18 Comparison of PNO sample PDFs as a function of annealing time at operating temperature	106
Figure A.19 Comparison of PNO sample PDFs as a function of annealing time at operating temperature	107
Figure A.20 Comparison of PNO sample PDFs as a function of annealing time at operating temperature	108
Figure A.21 Comparison of PNO sample PDFs as a function of annealing time cooled back to room temperature	109
Figure A.22 Comparison of PNO sample PDFs as a function of annealing time cooled back to room temperature	110
Figure A.23 Comparison of PNO sample PDFs as a function of annealing time cooled back to room temperature	111

Figure A.24 Comparison of PNO sample PDFs as a function of annealing time cooled back to room temperature	112
Figure A.25 Comparison of PNO sample $I(Q)$ as a function of annealing time at room temperature	113
Figure A.26 Comparison of PNO sample $I(Q)$ as a function of annealing time at operating temperature	114
Figure A.27 Comparison of PNO sample $I(Q)$ as a function of annealing time cooled back to room temperature	115

1

TODO LIST

2 CHAPTER 1

3 INTRODUCTION

4 Engineering materials and chemicals on the atomic scale has long been a goal for the
5 chemistry, physics, materials science, and chemical engineering fields. Realizing this
6 goal could lead to durable fuel cell catalysts, more bioavailable pharmaceuticals, and
7 radiation damage resistant spacecraft shielding. Before we can even think of making
8 atomistically exact structures, durable structures, or structures which change in re-
9 producible ways, we need to know the atomic structure exactly. This work addresses
10 these issues by developing a methodology for solving the structure of nanomaterials
11 by matching experimental x-ray scattering data with simulated atomic structures.

12 Chapter 2 develops the statistical mechanical system used to match the theoretical
13 structure. §2.2 focuses on the development of potential energy surfaces, including
14 potential energy and force equations, which have minima where experimental results
15 and simulated structures agree the most. §2.3 will discuss statistical mechanical
16 ensembles which are used to search for minima on the potential energy surface.

17 Chapter 3 will discuss the mathematical and computational development of the
18 atomic pair distribution function (PDF). §3.3 will focus on the rapid graphical pro-
19 cessing unit based calculation of the PDF and its gradients.

20 Chapter 4 will discuss the benchmarking of the the combined statistical mechan-
21 ical optimizer and PDF calculation systems against a series of theoretical nanoparti-
22 cles, focusing on understanding limitations of the method and structure reproduction.

23 Chapter 5 will focus on the acquisition of experimental data, its management, and
24 processing. §5.2, 5.3, and 5.4 will discuss the derivation of the Q resolution function,

25 the automated masking of 2D area detectors for x-ray total scattering measurements
26 using the previously derived Q resolution, and the impact of different averaging meth-
27 ods and masks on azimuthal integration, respectively.

28 Chapter 6 will discuss preliminary experimental results investigating the phase
29 changes and local structure of Pr_2NiO_4 , revealing the influence of thermal history on
30 the structure.

31

CHAPTER 2

32

STATISTICAL MECHANICAL ENSEMBLES AND POTENTIAL ENERGY SURFACES

34 2.1 INTRODUCTION

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

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

47 2.2 POTENTIAL ENERGY SURFACES

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

52 momenta p . Note that there could be more variables associated with the system,
53 for instance the magnetic moments of the atoms could play a role in describing the
54 system. In this magnetic system there would be positional variables for the atom-wise
55 spin vectors and their "momenta". Application of the term "momenta" might seem
56 odd here, as the magnetic spin does not have a mass or a velocity. However, since the
57 magnetic "position" is defined on the PES we need to describe its conjugate variable
58 to properly formulate Hamiltonian dynamics and the kinetic portion of the PES.

59 Experimentally Derived Potential Energy Surfaces

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

68 Potentials

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

$$\chi^2 = \sum_{a=a_{\min}}^{a_{\max}} (A_{\text{obs}} - \alpha A_{\text{calc}})^2 \quad (2.1)$$

$$Rw = \sqrt{\frac{\sum_{a=a_{\min}}^{a_{\max}} (A_{\text{obs}} - \alpha A_{\text{calc}})^2}{\sum_{a=a_{\min}}^{a_{\max}} A_{\text{obs}}^2}} \quad (2.2)$$

$$\chi^2_{\text{INVERT}} = \frac{1}{N} \sum_j \sum_r [A_{\text{obs}}(r) - \alpha A_{j\text{calc}}(r)]^2 \quad (2.3)$$

73

$$\alpha = \frac{\sum_{a=a_{\min}}^{a_{\max}} A_{\text{obs}} A_{\text{calc}}}{\sum_{a=a_{\min}}^{a_{\max}} A_{\text{calc}}^2} = \frac{\vec{A}_{\text{obs}} \cdot \vec{A}_{\text{calc}}}{|\vec{A}_{\text{calc}}|^2} \quad (2.4)$$

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

77 The Rw and χ^2 potentials have been reported numerous times. [46, 35, 7, 36, 48]
 78 Essentially these potentials measure the least squares distance between the observed
 79 scattering and the predicted scattering providing a way to quantify the agreement
 80 between the model and experiment. While RW and χ^2 are now standard in the PDF
 81 community, the INVERT potential is fairly new and aims to incorporate descriptions
 82 of the structural symmetry into the PES. [10, 11] In the case of the INVERT poten-
 83 tial NMR or other symmetry sensitive data is used to describe the number of unique
 84 atomic coordinations. This is then used to describe the number of unique atom-wise
 85 pair distribution functions, thus causing systems with more or less unique coordi-
 86 nation environments to be higher in energy. This approach has been shown to be
 87 useful for C_{60} and other systems which are highly symmetric, creating a PES with an
 88 easier to find minima. [10, 11] However, many times this kind of data is unavailable
 89 when refining the structure causing the potential to be less useful. Additionally, this
 90 potential introduces an element of user bias as the refiner must decide, based on some
 91 spectroscopic data, how many unique environments are in the material. This bias
 92 could be removed by using one of the other potentials with a method for simulat-
 93 ing the observed spectra, allowing the computational system decide what structures
 94 properly reproduce all the observed data.

95 **Forces**

$$\vec{\nabla} \chi^2 = -2 \sum_{a=a_{\min}}^{a_{\max}} \left(\alpha \frac{\partial A_{\text{calc}}}{\partial \gamma_{i,w}} + A_{\text{calc}} \frac{\partial \alpha}{\partial \gamma_{i,w}} \right) (A_{\text{obs}} - \alpha A_{\text{calc}}) \quad (2.5)$$

96

$$\vec{\nabla}Rw = \frac{Rw}{\chi^2} \sum_{a=a_{\min}}^{a_{\max}} (\alpha \frac{\partial A_{\text{calc}}}{\partial \gamma_{i,w}} + A_{\text{calc}} \frac{\partial \alpha}{\partial \gamma_{i,w}})(\alpha A_{\text{calc}} - (A_{\text{obs}})) \quad (2.6)$$

97

$$\vec{\nabla}\chi_{\text{INVERT}}^2 = \frac{-2}{N} \sum_{a=a_{\min}}^{a_{\max}} \sum_j (\alpha \frac{\partial A_{j,\text{calc}}}{\partial \gamma_{i,w}} + A_{j,\text{calc}} \frac{\partial \alpha}{\partial \gamma_{i,w}})(A_{\text{obs}} - \alpha A_{j,\text{calc}}) \quad (2.7)$$

98

$$\frac{\partial \alpha}{\partial \gamma_{i,w}} = \frac{(\sum_{a=a_{\min}}^{a_{\max}} A_{\text{obs}} \frac{\partial A_{\text{calc}}}{\partial \gamma_{i,w}} - 2\alpha \sum_{a=a_{\min}}^{a_{\max}} A_{\text{calc}} \frac{\partial A_{\text{calc}}}{\partial \gamma_{i,w}})}{\sum_{a=a_{\min}}^{a_{\max}} A_{\text{calc}}^2} \quad (2.8)$$

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

104 2.3 ENSEMBLES

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

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

$$\mathcal{H}(q, p) = U(q) + K(p) \quad (2.9)$$

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

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

$$\Xi = \sum_i P_i(q, p) \quad (2.10)$$

118 where P_i is the probability of the i th state and is a function of the total energy of
119 that state. This partition function can then be used to obtain the probability of any
120 specific state. The relationship of the probability of a state to the state's energy and
121 other properties depends on the ensemble being used.

122 For the canonical ensemble the partition function is probability is:

$$Q(N, V, T) = \exp\left(\frac{-\mathcal{H}(q, p)}{k_b T}\right) \quad (2.11)$$

123 where k_b is the Boltzmann constant and T is the temperature of the system. [37]

124 Monte Carlo Modeling

125 Monte Carlo can be used to simulate a statistical mechanical ensemble which can not
126 be solved analytically. In most Monte Carlo systems the ensemble is simulated by
127 randomly changing one of the system parameters and comparing the energy of the
128 new system against the energy of the old system. If the energy of the new system is
129 lower than the current energy then the new configuration is accepted. Otherwise the
130 new system is rejected unless

$$\exp\left(\frac{-\Delta E}{E_T}\right) < u \quad (2.12)$$

131 where u is a random number $[0, 1)$ and E_T is the thermal energy characteristic to the
132 system. The ability of Monte Carlo modeling to accept “bad” moves allows the system
133 to hop out of local energy minima during the search for the global minimum. Reverse
134 Monte Carlo (RMC) is similar to Monte Carlo except it uses χ^2 as the PES.[36]

135 Despite the utility of RMC, and its wide use in the x-ray scattering community, as
136 Hoffman and Gelman state “Not all MCMC [Markov Chain Monte Carlo] algorithms

137 are created equal".[24] RMC, similar to standard Monte Carlo simulations, samples
 138 from the PES at random, usually by translating atoms in the system randomly. This
 139 creates a less efficient, random walk based, exploration of the PES.[24, 38] Thus,
 140 methods for suppressing this random walk nature, while still searching the potential
 141 energy surface fully are needed.

142 Hamiltonian Monte Carlo

143 Hybrid or Hamiltonian Monte Carlo (HMC) can help to address some of these issues.
 144 HMC was developed originally in the lattice quantum chromodynamics community
 145 and provides a more efficient, more scalable approach to PES sampling for Monte
 146 Carlo.[15, 39] In HMC the PES is explored using Hamiltonian dynamics, essentially
 147 following the gradient of the PES to find more acceptable configurations.

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

$$\frac{dq_i}{dt} = \frac{\partial \mathcal{H}}{\partial p_i} = p_i \quad (2.13)$$

$$\frac{dp_i}{dt} = -\frac{\partial \mathcal{H}}{\partial q_i} = -\vec{\nabla}U \quad (2.14)$$

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

$$p_i(t + \delta t/2) = p_i(t) - \frac{\delta t}{2} \frac{\partial}{\partial q_i} U(q(t)) \quad (2.15)$$

$$q_i(t + \delta t) = q_i(t) + \delta t * p_i(t + \delta t/2) \quad (2.16)$$

$$p_i(t + \delta t) = p_i(t + \delta t/2) - \frac{\delta t}{2} \frac{\partial}{\partial q_i} U(q(t + \delta t)) \quad (2.17)$$

148 Note that $\frac{\partial}{\partial q_i}$ is the gradient with respect to q where i denotes the i th atom being

149 moved. Using this notation the gradient is

$$\vec{\nabla}U = \begin{bmatrix} \frac{\partial U}{\partial q_{0,x}} & \frac{\partial U}{\partial q_{0,y}} & \frac{\partial U}{\partial q_{0,z}} \\ \vdots & \frac{\partial U}{\partial q_{i,w}} & \vdots \\ \frac{\partial U}{\partial q_{n,x}} & \frac{\partial U}{\partial q_{n,y}} & \frac{\partial U}{\partial q_{n,z}} \end{bmatrix} = \begin{bmatrix} \vec{\mathcal{F}}_0 \\ \vdots \\ \vec{\mathcal{F}}_i \\ \vdots \\ \vec{\mathcal{F}}_n \end{bmatrix} \quad (2.18)$$

150 where $\frac{\partial}{\partial q_{i,w}}$ is the derivative with respect to q where w denotes direction of the deriva-
 151 tive (x , y , or z), n is the number of atoms and U is the potential which depends on q ,
 152 and $\vec{\mathcal{F}}_i$ is the "force" on the i th atom. Using these equations new potential configura-
 153 tions are proposed from the PES. These proposals are checked against the standard
 154 Metropolis criteria discussed above, except that the change in potential energy ΔE
 155 is replaced with the change in the Hamiltonian $\Delta\mathcal{H}$. Note that while this sampling
 156 closely simulates the canonical ensemble, it is not exactly the same. Usually the
 157 canonical ensemble is formulated as microcanonical ensembles in contact with an in-
 158 finite heat bath at a given temperature, or a set of microcanonical ensembles which
 159 exchange thermal energy. However, the HMC ensemble presented here has a momen-
 160 tum bath instead of a temperature bath. One could imagine the atoms sitting in a
 161 simulation box which has walls which can toggle their thermal exchange. Initially the
 162 box starts in the momentum bath, allowing the atoms to come to equilibrium with
 163 the bath momentum. The box is then removed from the bath causing it to become
 164 adiabatic. Hamiltonian dynamics are then propagated inside the box, essentially run-
 165 ning a microcanonical simulation. Once the dynamics are finished the energy of the
 166 system is checked with the Metropolis criteria and the box is reintroduced to the
 167 momentum bath and the process starts again.

168 **No-U-Turn Sampling**

169 Two parameters must be specified in HMC simulations, the step size δ and the number
170 of steps N . The step size is critical to the stability of the fitting procedure: with
171 a too small δ the simulation runs inefficiently producing structures too close to the
172 previous, whereas with a too big δ the linear approximation for the forces breaks
173 down and often the simulated NP explodes. The number of steps to take during the
174 dynamics is equally important and an inappropriate choice may result in backtracking
175 or random walk characteristics in the simulations. In this work, we employ the No-
176 U-Turn Sampling (NUTS) method recently proposed by Hoffman and Gelman to
177 address this issue [24]. In the NUTS method δ and N are dynamically computed
178 by examining the ratio of accepted to rejected configurations as well as whether
179 or not the simulation has started to take a U-turn. The U-Turn criteria makes
180 certain that the simulation stops when it begins to backtrack, preventing excess
181 computation on configurations that have very little new information to offer. The
182 use of NUTS leaves us with two simulation parameters: the simulation temperature
183 and the target acceptance. Hoffman and Gelman have empirically shown that the
184 ideal target acceptance, which governs the dynamics time steps, is .65, which we have
185 used for all of the simulations here. The simulation temperature sets the magnitude
186 of the random starting momenta for the atoms at the beginning of each dynamics
187 run [24].

188 **Grand Canonical Ensemble**

189 While NUTS-HMC simulations provide a system to find minima on PESs, the sim-
190 ulation is fundamentally run in the Canonical Ensemble thus the variables in the
191 simulation are limited to a fixed number of particles, simulation volume, and thermal
192 energy. Fixing the thermal energy and simulation volume is not a problem, as they
193 are not variables of interest in the final structure. However, specifying the number of

194 atoms in the system can be problematic, as the exact number of atoms in a sample
195 can be difficult to count or a sample could have a distribution of particle sizes. Thus,
196 a new ensemble needs to be used to allow the number of atoms to vary as a function
197 of the PES. This new ensemble is the Grand Canonical Ensemble.

198 **Ensemble description**

199 In the Grand Canonical Ensemble (GCE) two sets of variables are allowed to change,
200 the atomic positions, and the total number of atoms and their associated identi-
201 ties. These two variables are controlled by temperature, or average momentum, and
202 chemical potential. The partition function is

$$\Xi = \sum_{N=0}^{\infty} Q(N, V, T) \exp \frac{N\mu\beta}{T} \quad (2.19)$$

203 where $Q(N, V, T)$ is the Canonical partition function discussed above, μ is the chem-
204 ical potential. [37] This is translated into a Monte Carlo system, producing Grand
205 Canonical Monte Carlo (GCMC).

206 **Grand Canonical Monte Carlo**

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

$$\min[1, \frac{V}{(N+1)\Lambda(T)^3} e^{-\beta\Delta U + \beta\mu}] \quad (2.20)$$

210 Similarly the removal of an atom has the probability:

$$\min[1, \frac{(N)\Lambda(T)^3}{V} e^{-\beta\Delta U - \beta\mu}] \quad (2.21)$$

211 However, both of these equations depend of the overall simulation volume and the
212 thermal wavelength, which is undesirable as these are not really properties that we
213 are of interest to these simulations. Thus, we roll them into the definition of the

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

227 GCMC biasing

The basic idea of GCMC biasing is mapping, in 3 dimensions, where an atomic addition or removal is most likely to be accepted. Thus, the simulation volume is broken up into voxels, 3 dimensional volumes which are contained by the total simulation volume, with a pre-set size. Each voxel is given a probability of being chosen for a trial insertion where the probability is:

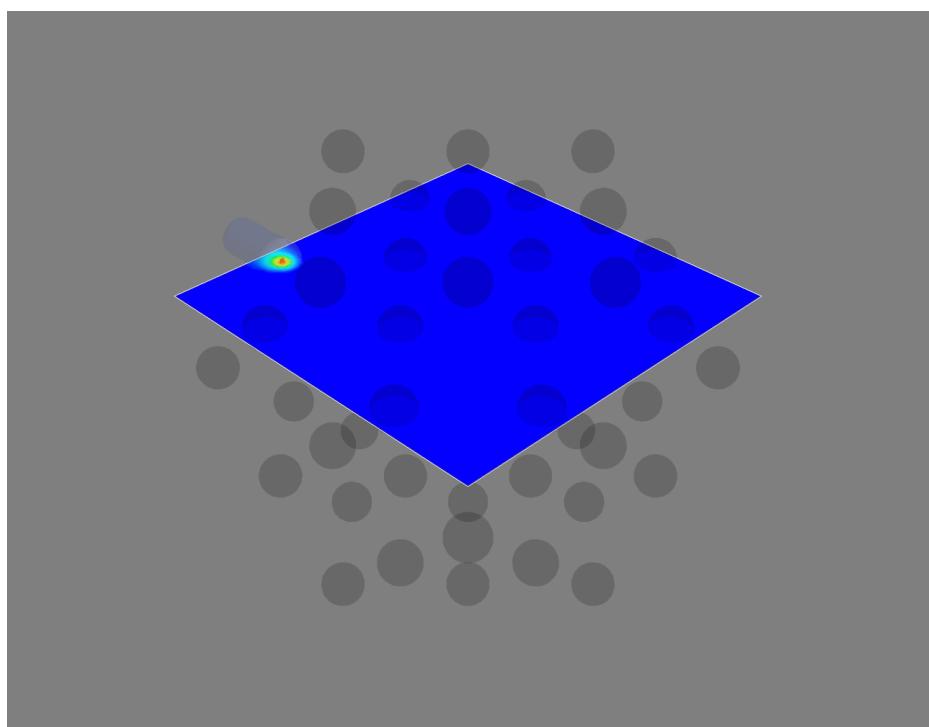
$$P_{i,j,k} = \frac{e^{\beta\Delta U_{i,j,k}}}{\sum_{i,j,k} e^{\beta\Delta U_{i,j,k}}} \quad (2.22)$$

where $\Delta U_{i,j,k}$ is the change in energy. However, calculating $\Delta U_{i,j,k}$ can be particularly expensive, especially when calculating scattering from atomic positions. The computational expense can be mitigated by using a cheaper potential, if only for the evaluation of the voxel energy, as previously shown. Similar to previous work we can use the Lennard Jones potential to approximate the addition potential, lowering the computational burden. [52]

239 Atomic deletion follows a similar biasing procedure, calculating the energy of each
240 atom and biasing the probability of each atom to be chosen for removal by its energy.
241 This way atoms which add the most energy to the system are more likely to be
242 removed.

243 Figure 2.1a shows an example map for atomic addition in a Au54 atom system,
244 with an Au55 atom target. Figure 2.1b shows the results of a few GCMC insertions
245 with biasing, showing the focusing of the simulation on the missing atom. The high
246 density of insertions around the missing atom would not have been possible without
247 the biasing.

(a)



(b)

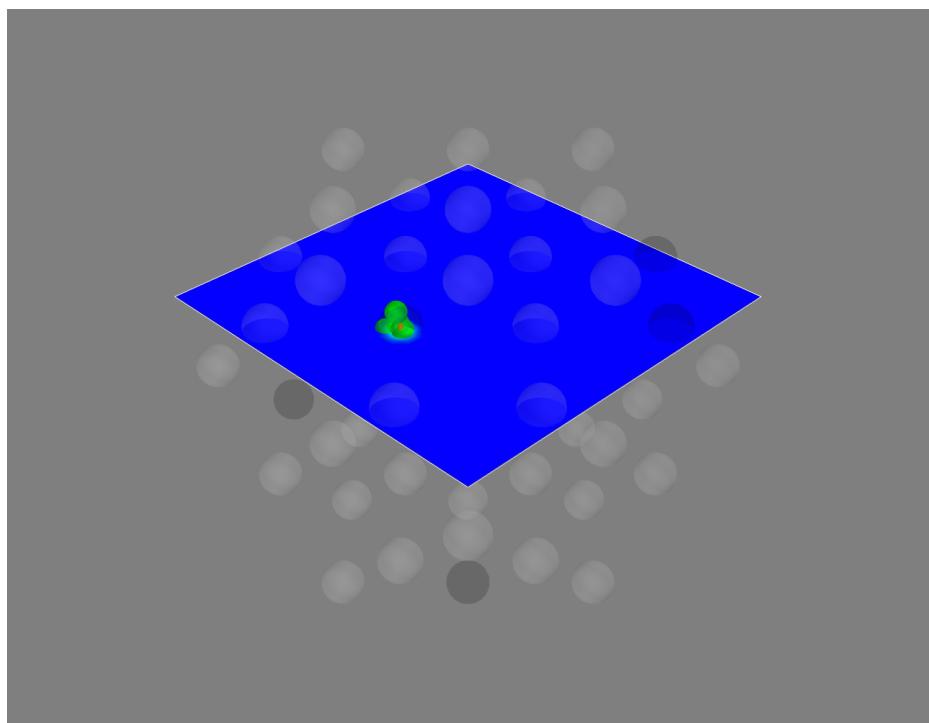


Figure 2.1: These figures show slices of the three dimensional addition potential energy surface. a) shows the addition probability, with the red area most probable. b) shows the results of trial insertions, where the green spheres denote where insertions were attempted.

248 2.4 CONCLUSIONS

249 In this chapter we have presented the development of both PES and the statistical
250 mechanical ensembles used to search them. We expanded the classical concept of
251 a PES to a more general mapping from positional variable space to energy space.
252 This expansion allowed for the implementation of experimentally derived PES, where
253 the disagreement between experimental and computed results can be included in the
254 PES. Common experimental PESs were discussed, and their forces derived. The
255 implementation of various statistical mechanical ensembles, used for searching the
256 PES for minima, was also discussed with a special focus on No-U-Turn-Sampling
257 Hamiltonian Monte Carlo. Grand Canonical Monte Carlo was also discussed, with
258 an emphasis on the us of biasing to increase the overall acceptance rate. Future
259 work in this area may include the development of PESs which leverage 2 dimensional
260 data, like STEM images, or ensembles which help to eliminate tuned parameters like
261 parallel tempering.

262

CHAPTER 3

263

ATOMIC PAIR DISTRIBUTION FUNCTION:

264

THEORY AND COMPUTATION

265

3.1 INTRODUCTION

Atomistic structural insight is essential for understanding and controlling a material's properties and functions, which has led to some of the most exciting advances in modern materials science and engineering. X-ray diffraction techniques are one of the most powerful tools for probing atomic structures with ultimate precision. Traditionally, thousands of diffraction peaks are analyzed using refinements of a structural model with few parameters to determine the 3D structure of bulk single crystals with high precision [21]. However, real engineered materials differ from ideal single crystals by showing a complexity in morphology, crystallite size, and atomic structure. The X-ray Powder Diffraction method (XPD) is among the most widely used methods for solving the structure of micro-crystals. The XPD technique utilizes hundreds of diffraction peaks and constrains the refinement of the structural model to few parameters in order to resolve the structure [42]. Solving the atomic structure becomes difficult using traditional x-ray diffraction techniques when the size of the material or its important features is reduced to the nanometer scale with non-periodic or short-periodic atomic arrangements. Materials consisting of particles with sizes less than a few tens of nanometers, often called nanoparticles (NPs), are structurally more complex than their bulk cousins. This structural complexity is often attributed to the large number of surface atoms which have incomplete coordination spheres [53],

284 surface relaxation [25], and surface environment effects [47, 23, 34]. These effects
285 make the precise determination of 3D atomic structure of NPs far more complicated
286 and problematic [4].

287 Over the years many advances have been made to address the famous “nanostruc-
288 ture problem”[4], for example, by using the atomic Pair Distribution Function (PDF)
289 analysis of x-ray and neutron total scattering data [17, 40, 48], PDF analysis combined
290 with molecular dynamics simulations [56, 22], bulk crystallography approaches [26],
291 and others. Attempts with non-diffraction based approaches have also been made,
292 including Transmission Electron Microscopy (TEM) [12], Raman spectroscopy[29],
293 Extended X-ray Absorption Fine-Structure Spectroscopy (EXAFS)[20] and Nuclear
294 Magnetic Resonance (NMR) [2].

295 In this chapter the PDF and its gradients will be derived. These expressions,
296 when combined with the PES and statistical mechanical treatment from chapter 2,
297 will allow for the solution of atomic structures. This chapter will also develop a
298 computational framework for evaluating the PDF and its gradients using Graphical
299 Processing Units (GPUs) to enable fast structural solution.

300 3.2 THEORY

301 To properly understand the PDF and its limitations we need to derive its mathemat-
302 ics. The PDF has been previously derived many times so it is not re-derived here.
303 This discussion of the PDF and its gradients use the notation of Farrow and Billinge.
304 [18]

305 **Derivation**

306 Many of the above techniques require the gradient of the PES. This in turn requires
307 the gradient of the PDF to be derived. Mathematically treating thermal vibrations
308 will also be discussed in this section. Systems which are truly extended materials, like

309 powders with particle sizes larger than 10nm, are best formulated as systems with
 310 periodic boundaries. Thus, the equations for a periodically bound PDF need to be
 311 developed as well, with their gradients.

312 Analytically Gradients

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

$$F(Q) = \frac{1}{N\langle f \rangle^2} \sum_{j \neq i} f_i^*(Q) f_j(Q) \exp(-\frac{1}{2}\sigma_{ij}^2 Q^2) \frac{\sin(Qr_{ij})}{r_{ij}} \quad (3.1)$$

322 where

$$\sigma_{ij}^2 = (\vec{u}_{ij} * \hat{d}_{ij})^2 \quad (3.2)$$

$$\vec{u}_{ij} = \vec{u}_i - \vec{u}_j \quad (3.3)$$

$$\hat{d}_{ij} = \frac{\vec{d}_{ij}}{r_{ij}} \quad (3.4)$$

$$r_{ij} = \|\vec{d}_{ij}\| \quad (3.5)$$

$$\vec{d}_{ij} = \begin{bmatrix} q_{ix} - q_{jx} \\ q_{iy} - q_{jy} \\ q_{iz} - q_{jz} \end{bmatrix} \quad (3.6)$$

323 where Q is the scatter vector, f_i is atomic scattering factor of the i th atom, and r_{ij} is
 324 the distance between atoms i and j and has q dependence. [27] For simplicity's sake

325 we will break up $F(Q)$ so that

$$F(Q) = \alpha \sum_{j \neq i} \beta_{ij} \tau_{ij} \Omega_{ij} \quad (3.7)$$

326 where

$$\alpha = \frac{1}{N \langle f \rangle^2} \quad (3.8)$$

$$\beta_{ij} = f_i^*(Q) f_j(Q) \quad (3.9)$$

$$\tau_{ij} = \exp\left(-\frac{1}{2} \sigma_{ij}^2 Q^2\right) \quad (3.10)$$

$$\Omega_{ij} = \frac{\sin(Qr_{ij})}{r_{ij}} \quad (3.11)$$

327 The derivatives are as follows:

$$\frac{\partial}{\partial q_{i,w}} F(Q) = \alpha \sum_j \beta_{ij} \left(\frac{\partial \tau_{ij}}{\partial q_{i,w}} \Omega_{ij} + \tau_{ij} \frac{\partial \Omega_{ij}}{\partial q_{i,w}} \right) \quad (3.12)$$

328 where

$$\frac{\partial \Omega_{ij}}{\partial q_{i,w}} = \frac{Q \cos(Qr_{ij}) - \Omega_{ij}}{r_{ij}^2} (q_{i,w} - q_{j,w}) \quad (3.13)$$

$$\frac{\partial \tau_{ij}}{\partial q_{i,w}} = \frac{\sigma_{ij} Q^2 \tau_{ij}}{r_{ij}^3} ((q_{i,w} - q_{j,w}) \sigma_{ij} - (u_{i,w} - u_{j,w}) r_{ij}^2) \quad (3.14)$$

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

330 Thus

$$\frac{\partial}{\partial u_{i,w}} F(Q) = \alpha \sum_j \beta_{ij} \frac{\partial \tau_{ij}}{\partial u_{i,w}} \Omega_{ij} \quad (3.15)$$

$$\frac{\partial \tau_{ij}}{\partial u_{i,w}} = -\frac{\sigma_{ij} Q^2 \tau_{ij}}{r_{ij}} (q_{i,w} - q_{j,w}) \quad (3.16)$$

331 **Without ADPs**

332 Without ADPs the equations simplify down to

$$F(Q) = \frac{1}{N \langle f \rangle^2} \sum_{j \neq i} f_i^*(Q) f_j(Q) \frac{\sin(Qr_{ij})}{r_{ij}} \quad (3.17)$$

333 and

$$\frac{\partial}{\partial q_{i,w}} F(Q) = \alpha \sum_j \beta_{ij} \frac{\partial \Omega_{ij}}{\partial q_{i,w}} \quad (3.18)$$

334 use of these equations, when ADPs are not appropriate (like at cryogenic tempera-
335 tures), greatly speeds up the computation.

336 Periodic Boundary Conditions

337 Periodic boundary conditions can be helpful when simulating extended solids or large
338 nanoparticles. In this case all the non-crystallinity is contained within the simulation
339 box and the box is repeated to create the longer distance peaks observed in the PDF.
340 To perform this we can break up the Debye equation into two main parts, the part
341 that describes the interatomic distances within the simulation box and those between
342 boxes. Neglecting the thermal motion portion:

$$F(Q) = \frac{1}{N \langle f \rangle^2} \left(\sum_{j \neq i} f_i^*(Q) f_j(Q) \frac{\sin(Qr_{ij})}{r_{ij}} + \sum_{i,j} f_i^*(Q) f_j(Q) \frac{\sin(QR_{ij})}{R_{ij}} \right) \quad (3.19)$$

343 where

$$R = |\vec{r} + \vec{\nu}| \quad (3.20)$$

$$\vec{\nu} = \gamma_1 * \vec{a} + \gamma_2 * \vec{b} + \gamma_3 * \vec{c} \quad (3.21)$$

344 where γ_i is the number of copies of the simulation box in the i th direction, and $\vec{a}, \vec{b}, \vec{c}$
345 are the lattice or superlattice directions.

346 3.3 COMPUTATION

347 Simply deriving the equations for the PDF is not enough. The many body nature of
348 the PDF equation make analytical solution of the structure from the PDF impossible.
349 Thus, the PDF must be computed from a structural candidates and compared against
350 experimental results to evaluate the reliability of the model. These computations were
351 implemented using high performance or high throughput computing methods (HPC
352 or HTC) and Graphical Processing Units to provide quicker solutions.

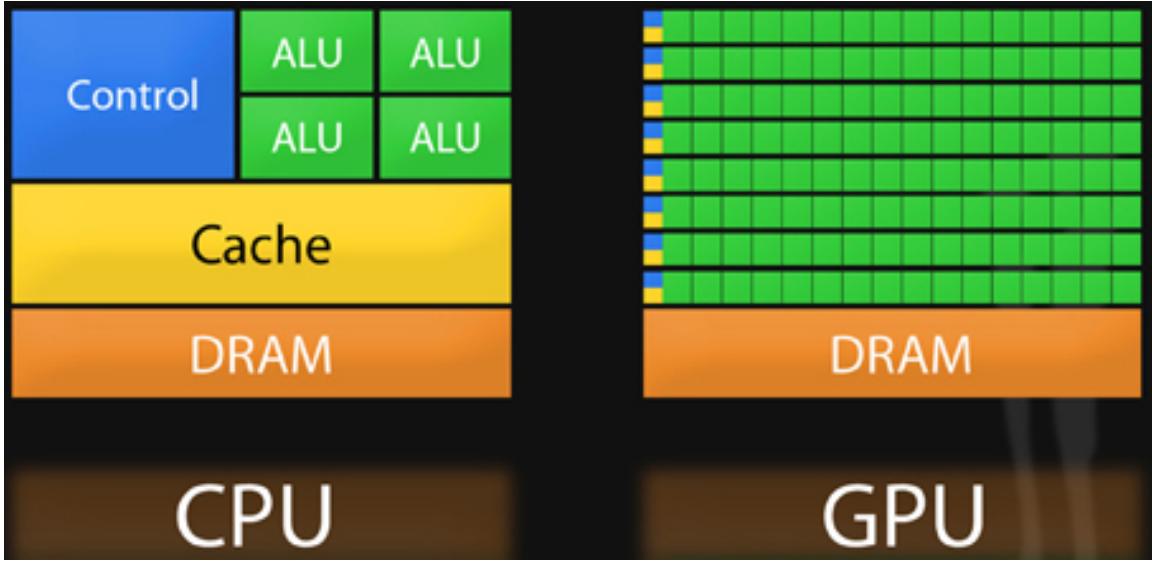


Figure 3.1: Comparison of the central processing unit (CPU) and GPU chip architectures from [6]. The arithmetic logic unit (ALU) are the arithmetic logic units which perform the mathematical operations, the dynamic random-access memory (DRAM) holds most of the data, although it is slower to access, the Cache holds rapidly accessed data, and the Control controls the execution of software. Note the greater number of ALUs on the GPU, the comparatively smaller cache, and the allocation of caches and controls to entire rows of processors.

353 HPC and GPUs

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

358 GPUs and Parallelization

359 Computing the PDF is an embarrassingly parallel problem. The basic procedure is
 360 to calculate the reduced structure factor $F(Q)$ for each atom pair and momentum
 361 transfer vector, sum over all the atom pairs, and Fourier transform the structure to
 362 the PDF. The first part of this procedure is perfectly parallelizable, as each atom pair
 363 is separate from the others. The summation over all the atomic reduced structure fac-

364 tors can be parallelized via distributed summing. Lastly the FFT can be parallelized
 365 using existing parallel Fast Fourier Transform (FFT) algorithms.

366 Graphical Processing Units (GPUs) are particularly well suited to the task of
 367 computing PDFs. GPU chip architecture is designed to perform many tasks simulta-
 368 neously by having potentially thousands of cores. Figure 3.1 shows the comparison
 369 of CPU and GPU architectures. As the figure shows the GPUs have a very different
 370 layout of computational processors (ALUs) and memory. While each ALU is simpler
 371 on the GPU, requiring the instructions to be less demanding in terms of memory,
 372 there are many more of them. The greater number of processors allows each atomic
 373 pairing to be placed on its own processor, so long as the math can be broken into sim-
 374 pler operations. The equations are broken up on the GPUs into various pieces which
 375 correspond to the α, β, τ and Ω as shown in equations 3.8-3.11 and sub-equations as
 376 needed. For example, while β is computed in one step, Ω requires the computation
 377 of the displacement array, then the distance array and finally the Ω array. The exact
 378 breakdown of processes, how the problems are broken down and spread across the
 379 processor has been optimized for speed and reliability.

380 **Map from ij space to k space**

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

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

385 E denotes the atomic coordinates in ij space, E' denotes $F(Q)$ before the summation
 386 in ij space, B denotes the atomic pairs in k space, B' denotes $F(Q)$ in k space, and
 387 Z denotes the final summed $F(Q)$. For the operators, ϕ denotes the mapping from
 388 ij space to k space $k = j + i * \frac{i-1}{2}$, ψ and ψ' denote the $F(Q)$ operation in ij and k
 389 space, respectively. Σ denotes the sum over all the atoms.

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

For the gradient a similar mapping is used:

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

401
 402 In this mapping, however, we use the $\tilde{\phi}\Sigma$ operator. This operator simultaneously
 403 performs a reverse mapping from k to ij space, and a summation with the correct
 404 symmetry. In this case the ψ and ψ' operators, which denote the $\vec{\nabla}F(Q)$ operator
 405 in ij and k space, are antisymmetric. Intuitively this makes sense as an extension
 406 of Newton's Second Law, since each particle's interaction is felt oppositely by its

407 partner.

408 **GPU Memory Allocation**

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

$$q[=]3n \quad (3.22)$$

$$d[=]3k \quad (3.23)$$

$$r[=]k \quad (3.24)$$

$$scatter[=]nQ \quad (3.25)$$

$$normalization[=]kQ \quad (3.26)$$

$$\Omega[=]kQ \quad (3.27)$$

$$F_k(Q)[=]kQ \quad (3.28)$$

$$Sum[=]kQ \quad (3.29)$$

$$Sum2[=]kQ \quad (3.30)$$

$$F(Q)[=]Q \quad (3.31)$$

416 where n is the number of atoms, k is the number of unique distances, Q is the scatter
417 vector, and the $[=]$ operator denote the number of single precision floating point
418 values in memory. Each of the above arrays are used in the computation and thus
419 must be able to be held in memory. Thus the number of atom pairs that can fit on

420 a GPU with am bytes of available memory is:

$$k_{perGPU} = \frac{1}{16Q + 16} (-4Qn - 4Q + am - 12n) \quad (3.32)$$

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

$$adps = 3n \quad (3.33)$$

$$\sigma = k \quad (3.34)$$

$$\tau = kQ \quad (3.35)$$

423 Thus the pair allotment is:

$$k_{perGPU} = \frac{-4Qn - 4Q + am - 24n}{20Q + 20} \quad (3.36)$$

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

$$\vec{\nabla}\Omega = 3kQ \quad (3.37)$$

$$\vec{\nabla}F_k(Q) = 3kQ \quad (3.38)$$

$$\vec{\nabla}F_n(Q) = 3nQ \quad (3.39)$$

426 Thus the gradient allotment is:

$$\vec{\nabla}k_{perGPU} = \frac{-16Qn + am - 12n}{32Q + 16} \quad (3.40)$$

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

$$\vec{\nabla}\tau = 3kQ \quad (3.41)$$

428 Thus the allocation is:

$$\vec{\nabla}k_{perGPU} = \frac{-16Qn + am - 24n}{48Q + 20} \quad (3.42)$$

429 These equations were solved by sympy as their validity is very important to the overall
430 reliability of the software. If the GPU is over allocated then the system may crash
431 or return meaningless results.

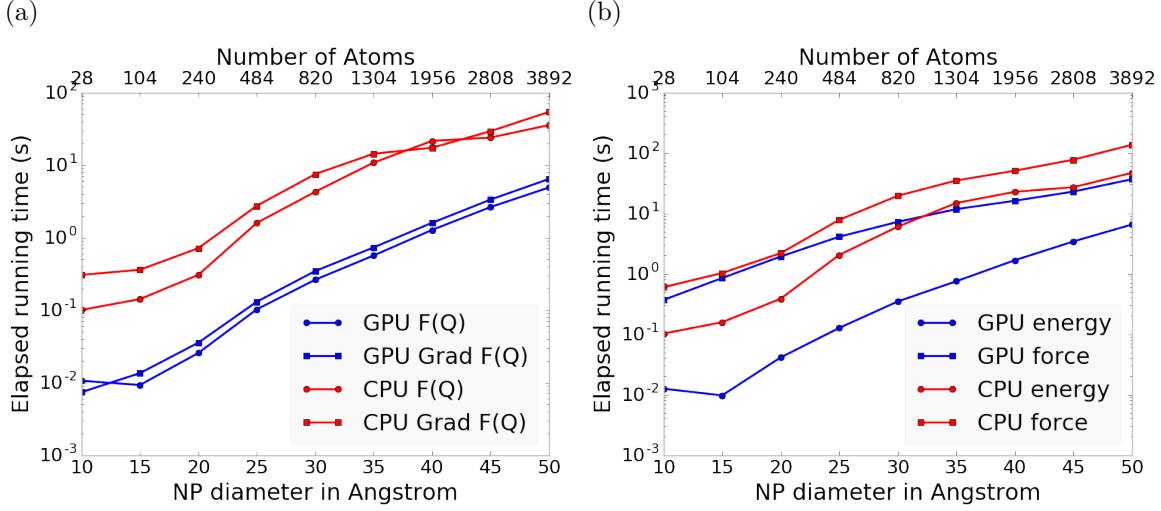


Figure 3.2: Speed comparison of CPU and GPU implementations. a) shows the time to compute the $F(Q)$ by itself. b) shows the time to compute the Rw based energy for Au NPs of various sizes, which includes computing $F(Q)$, its FFT, and the Rw .

432 Speed and Scaling of PDF Computation

433 To understand exactly how much the GPUs speed up the computation of $F(Q)$ and
 434 the PDF a series of time studies were run Au nanoparticles of varying size. Figure
 435 3.2 shows the results of these time studies. CPU and GPU calculations were carried
 436 out on an Intel i7-4820K @3.70GHz Quad-Core and one NVIDIA GTX970s, respec-
 437 tively. The $F(Q)$ computations show a 100x to 10x speedup using the GPUs over
 438 the CPUs. Additionally, the $\vec{\nabla}F(Q)$ and $F(Q)$ computations seem to have similar
 439 computation time and scaling relationships on the GPU. This implies that the two
 440 processes may have similar bottlenecks, most likely in the $F(Q)$ computation work-
 441 flow. This relationship is similarly preserved, although to a lesser extent, in the CPU
 442 scaling.

443 Interestingly, the tight run time relationship between $F(Q)$ and its gradient are
 444 not preserved in the Rw based force calculations. While the energy calculations are
 445 very similar to the $F(Q)$ calculations in terms of run time, the GPU and CPU force
 446 calculations are much closer, with the GPU calculations being much slower. This is

447 due to the force bottleneck being the $3n$ FFT operations which must be performed
448 on the $\vec{\nabla}F(Q)$ array to produce the $\vec{\nabla}\text{PDF}$ array. While the GPU is leveraged
449 to perform the FFT, the data must be loaded off the GPU and back on, causing a
450 potential slowdown. Larger systems of atoms were not tried as the CPU computation
451 quickly becomes very slow. Even higher GPU speedup is expected on more advanced
452 GPUs like the NVIDIA Tesla series.

453 3.4 CONCLUSIONS

454 In this chapter we developed the gradients of the PDF in the discrete and periodic
455 boundary condition case. We also developed the computational implementation of
456 the PDF equations. This implementation emphasized use of GPUs to compute the
457 PDF and its gradient. The GPU software was further sped up by mapping the com-
458 putation to atom pairs rather than atom by atom. Finally, the speed of the GPU
459 implementation was checked against the CPU implementation via speed benchmark-
460 ing.

461

CHAPTER 4

462

BENCHMARKS

463 4.1 INTRODUCTION

464 Having developed the ensemble, PES, and PDF mathematics in chapters 2 and 3
 465 this chapter will now benchmark the system to check its ability to produce resonable
 466 structures from PDFs. The benchmarks shown here are certainly not exahustive, nor
 467 are they the even all the benchmarks run against the NUTS-HMC system. However,
 468 they do provide a good window into the workings, limitations, and abilities of the
 469 system.

470 The formation of NPs with both crystallographic and non-crystallographic struc-
 471 tures [34] and with different chemical patterns [19] are well documented. For sim-
 472 plicity, we chose monometallic Au clusters as benchmarks and considered two groups
 473 of structures with different size and degrees of structural disorder in order to assess
 474 the reliability and efficiency of our HMC method for solving atomic structures from
 475 PDFs. The first group consists of Au_{55} clusters with different degrees of disorder,
 476 including a crystalline cluster structure in O_h (Octahedral) symmetry, a structure
 477 with a disordered surface, and an amorphous structure. The second group consists
 478 of the crystallographically solved Au_{102} structure as in the $\text{Au}_{102}\text{MBA}_{44}$ nanocrystals
 479 [26, 33]. We used optimized structures from the Density Functional Theory (DFT)
 480 as target structures and generated the corresponding PDF, G_{obs} , according to

$$G_{\text{obs}} = \frac{2}{\pi} \int_{Q_{\min}}^{Q_{\max}} Q[S_{\text{obs}}(Q) - 1] \sin(Qr) dQ \quad (4.1)$$

481 where S_{obs} is the target structure's structure factor. Since all the target structures

were optimized by DFT at zero Kelvin the target and model PDF profiles were calculated at zero temperature, with no atomic displacement parameters (ADPs). However, ADPs would have a considerable impact on the calculation of the PDF, especially for nanoparticles at non-zero temperatures.

Target Setup

Spin-polarized DFT calculations were carried out using the Vienna ab initio simulation package (VASP) [32, 31] within the Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional [43]. The projected augmented wave method [5] and a kinetic energy cutoff of 400 eV were used. Structural optimization was performed until the total energy and ionic forces were converged to 10^{-6} eV and 10 meV/Å, respectively. The amorphous Au_{55} structures were generated by simulated annealing using the classical embedded atom method potential [50]. Different annealing temperatures between 1200 K and 1670 K (bulk melting temperature of Au) were used and the thermally equilibrated structures were cooled down to 300 K before minimization at 0 K. Further optimization using DFT leads to total energies that vary within 1-2 eV among different amorphous structures and the lowest energy one was used as the target structure. The target structure of Au_{102} was taken as the Au_{102} core of the DFT-optimized $\text{Au}_{102}\text{MBA}_{44}$ cluster [33].

All systems were solved using a PES which consists of a linear combination of Rw , the repulsive and attractive thresholded spring potentials. The total potential energy in the Hamiltonian in Eq. (2.9) is expressed as:

$$U(q) = U_{Rw}(q) + U_{\text{spring}}(q, R_{\min}) + U_{\text{spring}}(q, R_{\max}) \quad (4.2)$$

The thresholded spring potentials are based on those previously proposed on by Petersen [44], 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 bounds of the spring threshold r_t . The resulting restoring forces on the out-of-bound

atoms bring the system back within the bounds of the PDF, R_{\min} and R_{\max} , and therefore preventing the system from exploding or collapsing. Otherwise, incorrect solutions 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. (4.2) 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. (4.1) with a step of $\delta R = .01 \text{ \AA}$, $Q_{\min} = 0.1 \text{ \AA}^{-1}$, $Q_{\max} = 25.0 \text{ \AA}^{-1}$. 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

$$\delta Q = \frac{\pi}{R_{\max} + \frac{12\pi}{Q_{\max}}} \quad (4.3)$$

The HMC simulation was run with $N = 300$ iterations, a target acceptance rate

530 of 0.65, and an average starting momentum for each NUTS iteration of 10 eVfs/Å.
531 Both repulsive and attractive spring potentials are used with $\kappa = 200$ eV/Å and
532 thresholds matching R_{\max} and R_{\min} of the PDF, respectively. $\lambda = 300$ eV was used
533 as conversion factor for Rw . Each simulation was run with a pair of Nvidia GTX970
534 graphics cards, with one card partially occupied with desktop visualization.

535 4.2 STRUCTURAL SOLUTIONS

536 **Case I: crystalline Au₅₅**

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

544 R_{\min} and R_{\max} for this simulation were 2.45 Å and 11.4 Å, respectively, with
545 $\delta Q = 0.24 \text{ \AA}^{-1}$. The simulation ran for approximately 34 minutes, over a total of
546 ~40 thousand configurations. The results are shown in Fig. 4.1.

547 The PDF, radial bond distribution, and bond angle distribution show good agree-
548 ment between the target and final fitted structures, with a Rw of 0.3% whereas Rw
549 of the starting structure is as high as 44.8%. DFT calculations yield a total energy of
550 the final structure very close to that of the target structure (within a few meV). The
551 success in the fitting is largely attributed to the factor that the target structure is
552 only locally (and mildly) disturbed from its bulk-like counterpart and therefore there
553 is no need to overcome any high PES barriers to reach the correct solution. As shown
554 below, the situation is rather different for much more disordered target structures.

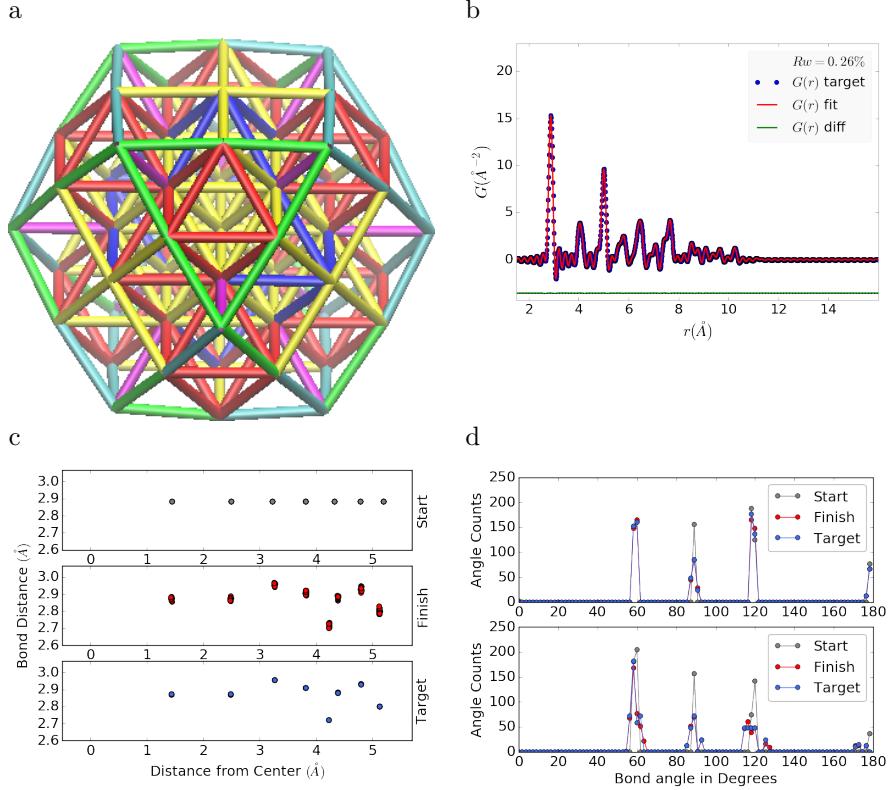


Figure 4.1: Au_{55} PDF fitting of DFT-optimized $c\text{-Au}_{55}$. a) the final structural solution ($Rw=0.3\%$) with bond lengths color-coded by step of 0.05\AA , b) the target PDF(blue dots) overlaid with the PDF of the final structure (solid red lines) with the difference in green lines offset below, c) the radial bond distribution, and d) bond angle distribution.

555 Interestingly, the small-box solution using PDFgui[17] yields a rather large Rw of
 556 43%, due to the failure to fit the surface contracted atoms with a unit cell. The
 557 PDF fits of the starting structure and small-box solution are shown in Fig. S1 in the
 558 Supporting Information.

559 Case II: Au_{55} with surface disorder

560 In addition to surface relaxation, the structure of a cluster or nanoparticle is often
 561 disrupted by the presence of defects and/or ligand bound to the surface. To mimic
 562 such surface disorders, we took the DFT-optimized $c\text{-Au}_{55}$ structure from case I as
 563 the starting structure and randomly displaced the surface atoms with a normal distri-

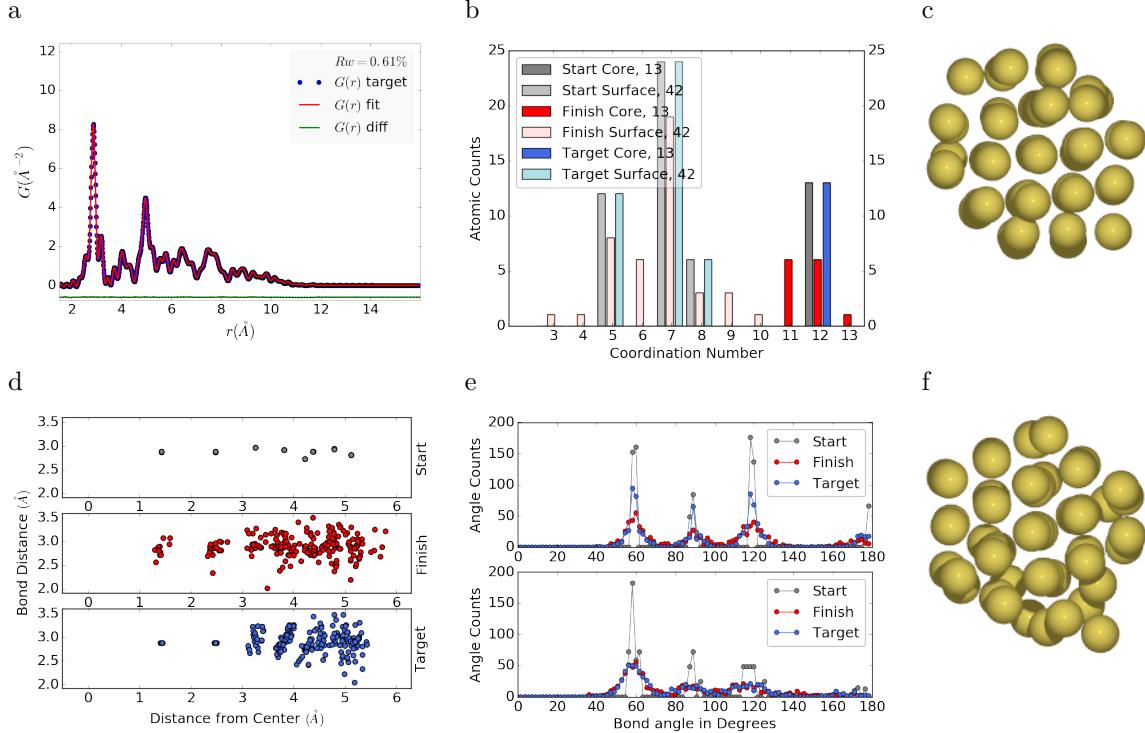


Figure 4.2: Au_{55} PDF fitting of surface-disordered Au_{55} . (a) the comparison of PDFs, (b) the CN distribution, (c) the target structure, (d) the radial bond distribution, (e) the bond angle distribution, and (f) the final structural solution ($Rw=0.6\%$)

564 bution of $\sigma = 0.2 \text{ \AA}$. All atoms are allowed to move in the HMC simulation, including
 565 the originally undisturbed core, which is a Au_{13} cluster with O_h symmetry.

566 R_{\min} and R_{\max} for this simulation were 1.95 \AA and 12.18 \AA , respectively, with
 567 $\delta Q = 0.23 \text{ \AA}^{-1}$. The simulation ran for approximately 3.6 hours, over a total of ~ 270
 568 thousand configurations. The results of the simulation are shown in Fig. 4.2.

569 Overall, good agreement is found between PDFs of the target structure and the
 570 final structural solution, even out to larger r , with an $Rw = 0.6\%$ starting from an
 571 $Rw = 50.4\%$ (see Fig. S2). The radial bond distribution and angle distribution
 572 show reasonably good agreement, but with lower degree of crystallinity in the final
 573 structure compared to the target structure. The discrepancy is most obvious in
 574 the core: despite the identical core structure in the starting and target structures,
 575 the core atoms were displaced in the HMC simulations in order to achieve a “best”

576 solution. This is because PDF measures the global average of interatomic distances
577 between each atomic pair and does not contain direct information about the locality
578 of such pairs, e.g. on the surface or cores. If such information is obtained a priori, for
579 example, from theoretical prediction or other experimental measurements, the core
580 structure can then be fixed and excluded from HMC dynamics.

581 Similar discrepancies are found in the CN distribution. Since the initial displace-
582 ments of the surface atoms are relatively mild, the interatomic connectivities remain
583 more or less the same and therefore the target structure has an identical CN distri-
584 bution to the starting (unperturbed) structure. This is, however, not the case for
585 the final fitted structure, which shows discernible differences, especially at the low
586 and high CN numbers. This is partly caused by the displacement of the core atoms
587 mentioned above, and partly by the lack of CN constraints for the PDF fitting, which
588 has been previously demonstrated in the case of α -Si [10]. Additional experimental
589 data, e.g. from EXAFS or NMR, may help to steer the simulations towards better
590 agreement in both PDF and CN distribution.

591 Case III: amorphous Au₅₅

592 Next, we turn to the case in which the entire cluster structure is disordered. We used
593 a DFT-optimized amorphous Au₅₅ (*a*-Au₅₅) as the target structure, and the DFT-
594 relaxed *c*-Au₅₅ cluster from Case I as the starting structure. The total energy of
595 *a*-Au₅₅ was computed to be *lower* than that of *c*-Au₅₅ by as large as 2.9 eV, consistent
596 with the 3.0 eV found in previous DFT work [13].

597 R_{\min} and R_{\max} for this simulation were 2.6 Å and 11.26 Å, respectively, with
598 $\delta Q = 0.25 \text{ \AA}^{-1}$. The simulation ran for approximately an hour, over a total of ~ 87
599 thousand configurations. The results of the simulation are shown in Fig. 4.3.

600 Our PDF fitting yielded a final structure of Rw of 1.7%, whereas that of the
601 initial structure is as high as 76.1% (see Fig. S3), due to the drastically different

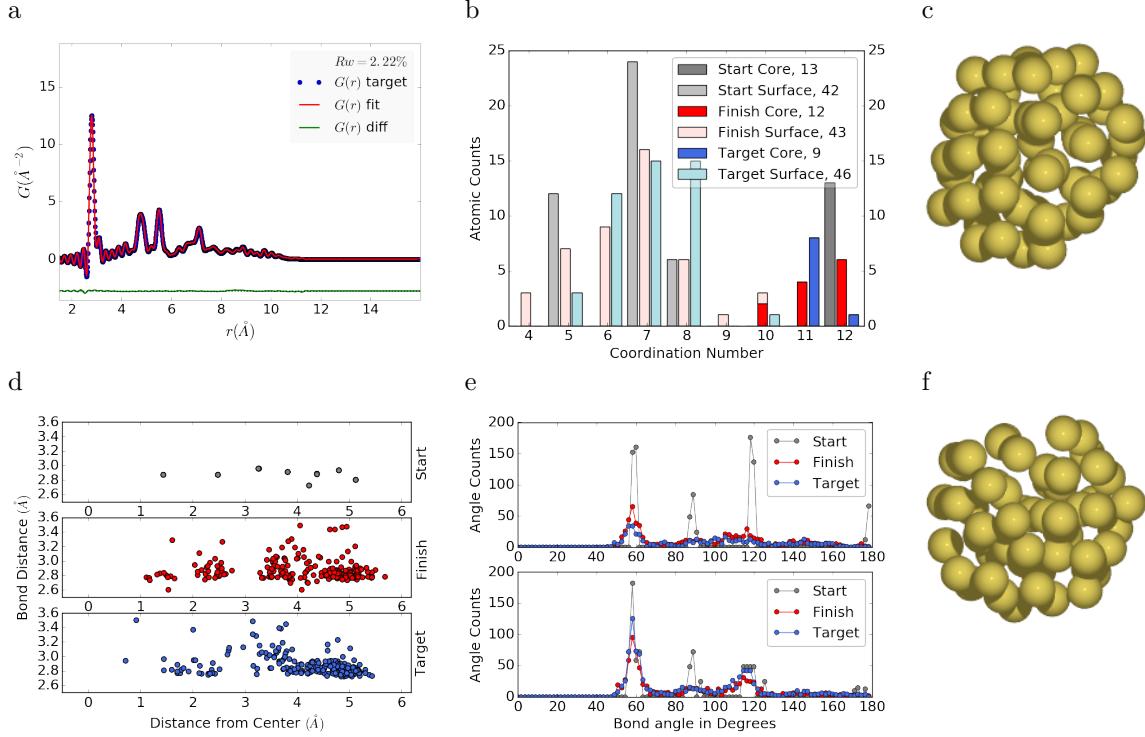


Figure 4.3: Similar to figure 4.2 for DFT-optimized amorphous Au_{55} .

602 atomic structure of the crystalline and amorphous Au_{55} clusters. Overall reasonable
 603 agreement in PDF, bond angle distribution, and radial bond distance distribution
 604 was found, and the wide spread of the bond lengths was qualitatively reproduced.
 605 However, the mismatch in CNs is problematic, partly due to the lack of information
 606 and/or constraints on the CNs. The total energy of the final structure is computed to
 607 be ~ 6 eV higher than that of the target structure and the difference is substantially
 608 larger than the variation among different amorphous structures computed by DFT
 609 ($\Delta E_{\text{tot}} \sim \pm 1\text{-}2$ eV). Such a fitting result, despite the rather small Rw , clearly
 610 indicates the importance of complementary informations and/or constraints necessary
 611 for reliably solving disordered NP structures from PDF.

612 **Case IV: ligand-protected Au₁₀₂**

613 Our final benchmark is Au₁₀₂, whose structure was initially solved by Jadzinsky and
614 co workers using x-ray crystallography [26] and further confirmed by DFT studies
615 [33]. The Au₁₀₂ structure consists of three main parts, a 49-atom Marks decahedron
616 core, two C₅ caps consisting of 20 atoms each, and 13 equatorial atoms. Unlike
617 previous cases, the multi-symmetry nature of the structure, i.e. each part has its own
618 distinct symmetry, poses a challenge for PDF-based solution of the structure. This is
619 because of the atomically centralized nature of the PDF, in which each atom “sees”
620 a density of other atoms surrounding it and has a strong tendency towards becoming
621 the center of the main symmetry group. Such tendency may lead to a solution where
622 some of the correct atomic symmetries are discarded in favor of the core symmetry.

623 **Starting from fcc structure**

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

626 R_{\min} and R_{\max} for this simulation were 2.7 Å and 16. Å, respectively, with $\delta Q =$
627 0.18 Å⁻¹. The simulation ran for approximately two hours, over a total of ~82
628 thousand configurations. The results of the simulation are shown in Fig. 4.4.

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

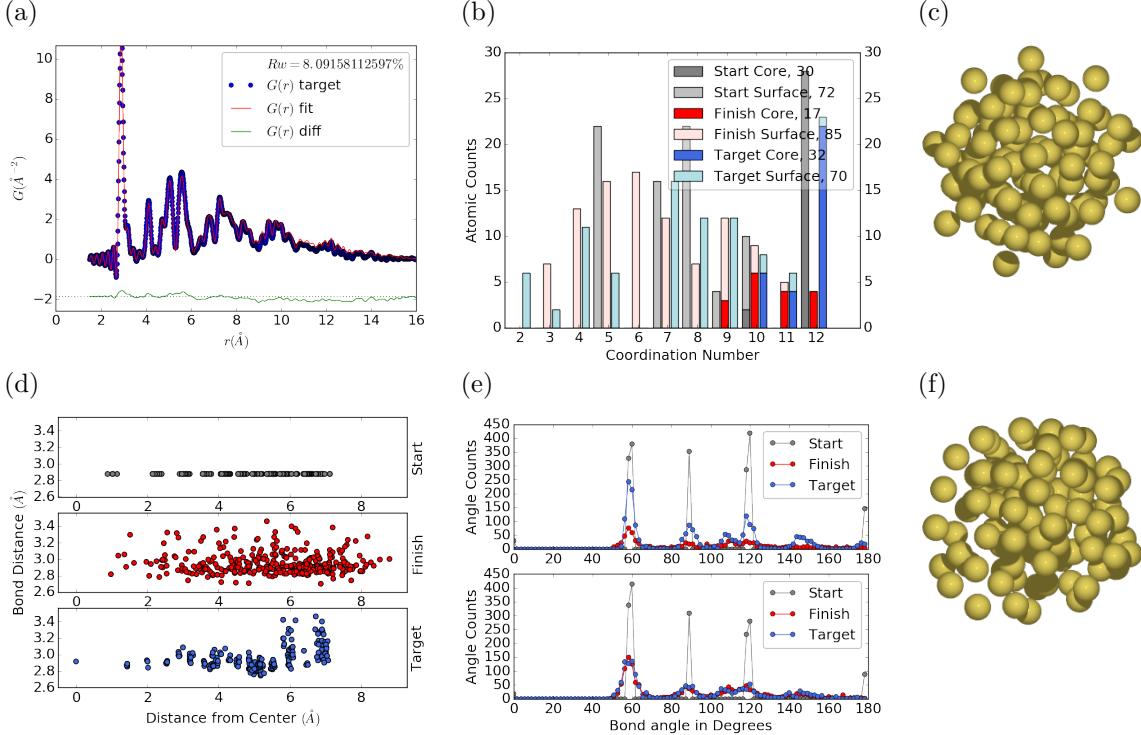


Figure 4.4: Similar to Fig. 4.2 for Au_{102} as in DFT-optimized $\text{Au}_{102}\text{MBA}_{44}$ cluster.

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

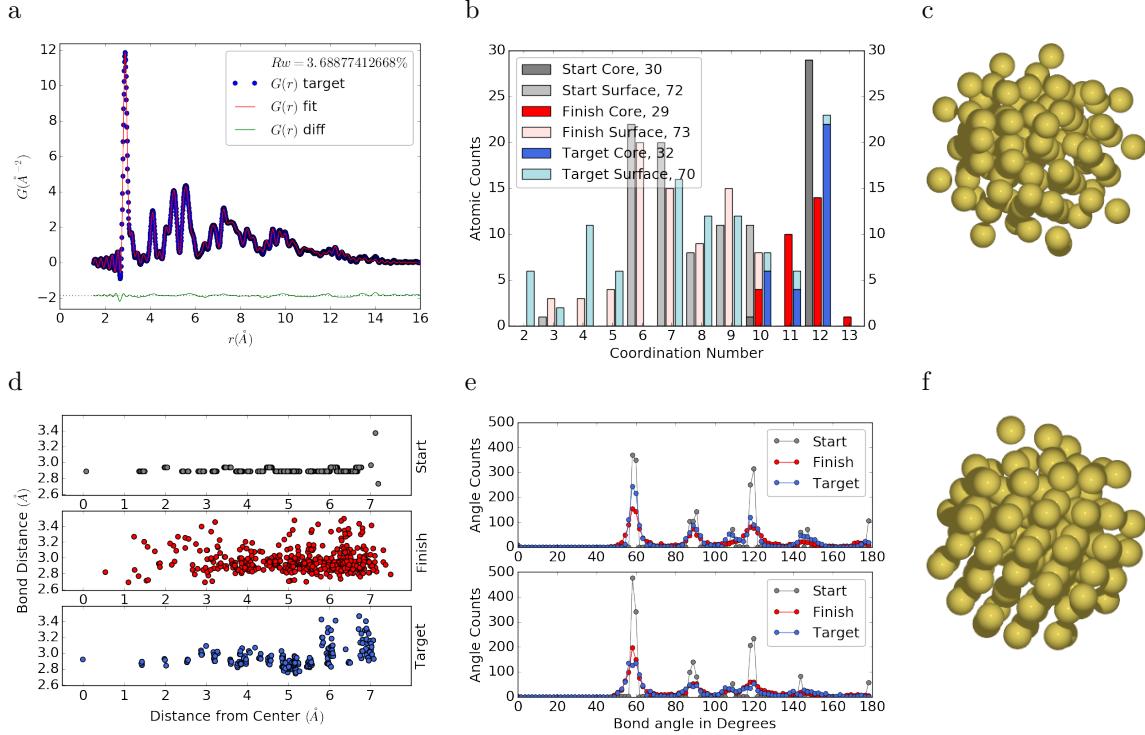


Figure 4.5: Similar to Fig. 4.4 with Marks decahedron as the starting structure.

648 Starting from Marks Decahedron

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

654 R bounds and Q resolution were the same as the previous case. The simulation
 655 ran for approximately 2.5 hours over a total of ~ 90 thousand configurations. The
 656 results of the simulation are shown in Fig. 4.5.

657 The starting structure of Marks decahedron ($Rw=56.6\%$, see Fig. S5) yielded
 658 a better structural solution, with a final Rw of 3.3%. However, the discrepancies at
 659 high R remains as in the previous case. By examining the final structure, we can see
 660 that these high R errors are due to a lack of the two 20-atom caps and 13 equatorial

661 atoms. Similarly, the radial bond distance distribution displays a diffusive behavior
662 unlike the bond length clustering in the target structure. Compared to the previous
663 case, the agreement in the CN and bond angle distributions are improved, with the
664 latter capturing nearly all peaks in the target structure with the exception of the 110
665 bond angle. Relatively large discrepancies are found in the CN distribution at the
666 low and high ends.

667 4.3 DISCUSSION AND CONCLUSION

668 The challenge of NP structure determination from X-ray Diffraction stems from the
669 absence of long range atomic ordering and the presence of non-trivial disordered
670 structures. PDF data analysis is one of the successful experimental methods for
671 solving NP structures, which has been tested on varieties of NP structures using
672 non-crystallographic large box hybrid methods [46]. In this work, we developed a
673 novel, HMC-based method to solve the structures of NPs from PDFs. We chose
674 monometallic Au clusters as benchmarks and generated synthetic experimental PDF
675 data from DFT-optimized cluster structures. Two cluster sizes (Au_{55} and Au_{102})
676 and different degree of disorders (crystalline, surface disordered, amorphous, ligand-
677 protected) were considered, which pose further challenges in addition to the lack of
678 long-range order in finite-size systems.

679 The quality of the PDF fit, as measured by the agreement factor Rw , is expected
680 to provide a basis for the comparison between the fit structure and the target struc-
681 ture. The HMC method is designed to find minima in the Rw PES and therefore the
682 ultimate criterion for its success is the magnitude of the Rw values. This criterion has
683 been met in the cases of surface relaxed $c\text{-}\text{Au}_{55}$, surface distorted-Au55, and $a\text{-}\text{Au}_{55}$
684 systems, where we obtained an excellent fits with $Rw < 2\%$. With most of the mod-
685 eling methods a PDF fit with an $Rw < 15\%$ would be considered a solved structure
686 [46, 1, 35, 51, 45]. However, a low Rw does not necessarily translate into the correct

687 structural solution and the discrepancies may be caused by several different factors.
688 For example, the disordered surfaces associated with NPs produces a fall off in the
689 PDF profile, resulting in underestimated NP sizes in the fitted structures [46]. In ad-
690 dition, as the PDF profile depends on the average core and surface pair distributions,
691 the structural solution may not be unique especially for disordered systems. This is
692 best manifested in the example of amorphous Au₅₅, in which the final structure pro-
693 duces a rather small R_w but has distinctively different structural features and a much
694 higher total energy compared to the target structure. In order to faithfully solve the
695 structures, other fitting metrics, such as the atomic coordination numbers, may need
696 to be considered. For ordered or partially ordered systems, we have demonstrated
697 a reasonable agreement between the final and target CN distributions, whereas the
698 discrepancy increases for disordered structures. A hybrid HMC method combining
699 with other experimental (EXAFS, NMR) or computational (DFT, force field) meth-
700 ods is expected to provide more physical constraints, e.g. CN distribution, in the
701 PDF fitting and result in better structural agreement.

702 The case of ligand-protected Au₁₀₂, a crystallographically well solved NP system
703 by X-ray diffraction, provides a unique test for the HMC algorithm due to its multi-
704 symmetry nature and very short range ordering. The structure of Au₁₀₂ was solved
705 using two different starting crystal structures, i.e. fcc and Mark decahedron. The R_w
706 values support the Marks decahedron structure of the cluster core over the fcc one,
707 but discrepancies are found at high R values. This could be explained as limitation
708 of the method while using PDF profile fit for high symmetry local structures such
709 as cap atoms in Au₁₀₂, which has a C_5 symmetry. This is an intrinsic problem with
710 the PDF, which measures the global inter-atomic distances, thus causing each atom
711 to see itself at the epicenter of the majority symmetry. Use of additional techniques,
712 like EXAFS, which has an emphasis on the local coordination spheres, may help
713 with finding better minima by breaking the dominant symmetry and providing a low

714 energy path to compartmentalized structures.

715 In summary, we present the development of a novel Hamiltonian Monte Carlo
716 based method for finding atomic structure solutions to PDFs. We derived the ana-
717 lytical derivatives of Rw , $G(r)$, and $F(Q)$, which were used to guide the Hamiltonian
718 dynamics to follow the gradient of the PES towards minimum energies. The ma-
719 trix element algebra inherent to the Debye sum and its derivatives lent itself to be
720 written as GPU kernels, allowing for the speedup of the HMC algorithm by massive
721 parallelization. The GPU implementation is 100 to 4x faster than the comparable
722 multi-core CPU enabled code, depending on the size of the NP. The robustness and
723 accuracy of the HMC method was benchmarked against a set of Au cluster struc-
724 tures of different size and degree of disorder. Overall, low Rw values were obtained
725 for all the benchmarks, with most of the structural features reproduced as shown by
726 various structural metrics. The Au_{102} cluster, which contains both high symmetry
727 local structures and disordered surface regions, posed the largest challenge. Although
728 the fitted structure shows some dependence on two starting configurations, neither
729 of them was able to fully solve the target structure. Future work which employs ex-
730 perimentally derived potentials, including EXAFS, or computational potentials that
731 contain chemical bonding information, may help to solve these particularly difficult
732 cases. Work is underway to extend this method to the grand canonical ensemble and
733 parallel tempering, which will allow for solving structures with unknown number of
734 atoms and compositions and provide better stochastic mixing of structures.

735

CHAPTER 5

736

X-RAY TOTAL SCATTERING DATA ACQUISITION AND PROCESSING

738

5.1 INTRODUCTION

739 X-ray total scattering experiments are generally performed at synchrotron light sources,
740 as only these sources can provide the needed flux, energy, and high momentum trans-
741 fer vectors needed to obtain reliable PDFs. [9, 16] Despite the need for a dedicated
742 facility to perform the total scattering experiments, the experiments themselves are
743 fairly forgiving, allowing for reactive gaseous environments, experiment temperatures
744 ranging from 2 K to 1800 K, and even electrochemical cycling. [8, 47, 49] The rapid
745 PDF data acquisition associated with 2D area detectors creates a data management
746 problem, as 96 hours of beamtime could result in almost 10,000 images which need
747 to be associated with the experimental conditions and detector metadata. [9] Finally,
748 all this data needs to be processed by masking bad pixels and regions, integrating
749 azimuthally, and converting the scattering data to the PDF. [30, 28, 54, 41, 3]

750

5.2 DETECTOR Q RESOLUTION

751 To properly azimuthaly integrate the images taken from the detector the Q resolution
752 of the pixels must be calculated. Azimuthal integration is the process of deviding the
753 pixels up into “bins”. Each bin has a set width, usually in Q , which describes which
754 pixels can go into the bin. During the integration process each pixel is placed into
755 its correspondng bin. Finally a statistical measurement of the average of all the

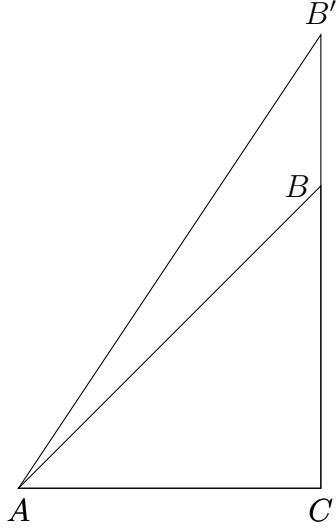


Figure 5.1: Scattering onto a flat detector

pixels in each bin is taken to produce the $I(Q)$ data. Although commonly performed, integrating using evenly spaced bins will cause pixels which are not on the same ring to be binned together, causing the incorrect value of $I(Q)$ to be obtained and a larger standard deviation in the integrated data. To properly calculate the Q resolution the resolution of each of the pixels in 2θ must be calculated. Figure 5.1 shows the scattering of x-rays onto a flat image plate detector. In this diagram the bottom of the n th pixel is B while the top is B' . The resolution of this pixel in 2θ is $\angle BAC - \angle B'AC$. Thus the resolution, calculated from the distances is

$$\Delta 2\theta = \arctan \frac{b}{d} - \arctan \frac{t}{d} \quad (5.1)$$

where d is the sample to detector distance (AC in figure 5.1), b is the distance to the bottom of a pixel (CB in figure 5.1)), and t is the distance to the top of that pixel (CB' in figure 5.1)). Note that these distances need to have been corrected for detector tilt and rotation. Thus the resolution of a pixel in Q is

$$\Delta Q = \frac{4\pi(\sin \arctan \frac{b}{d} - \sin \arctan \frac{t}{d})}{\lambda} \quad (5.2)$$

where λ is the x-ray wavelength.

769 This effect is analagous to looking at windows on a very tall building. The windows
770 are all the same size, but due to the nature of the perspecitve, the windows seem to
771 shrink as one looks higher up the building.

772 For a Perkin Elmer image plate, like the one used at the NSLS-II's XPD and the
773 APS's 11-ID-B, the resolution function is shown in 5.2. For the same detector the
number of pixels per Q is shown in 5.3

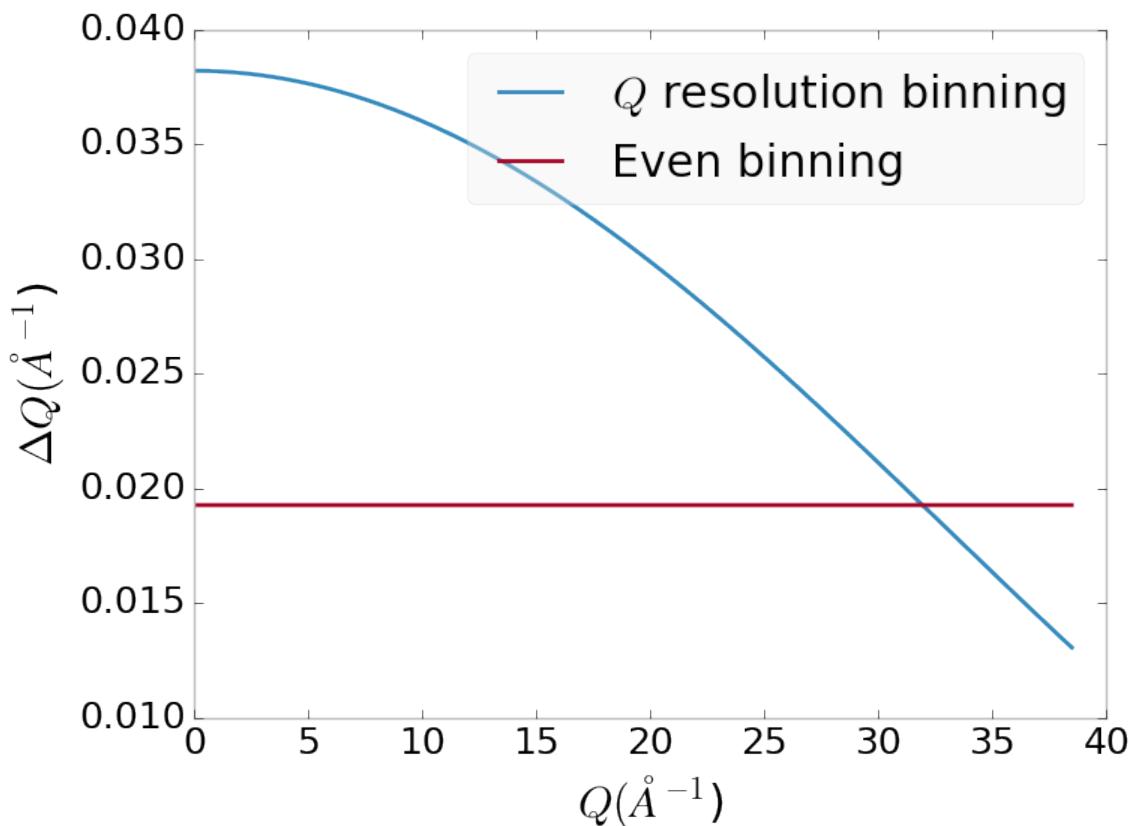


Figure 5.2: Q resolution as a function of Q .

774

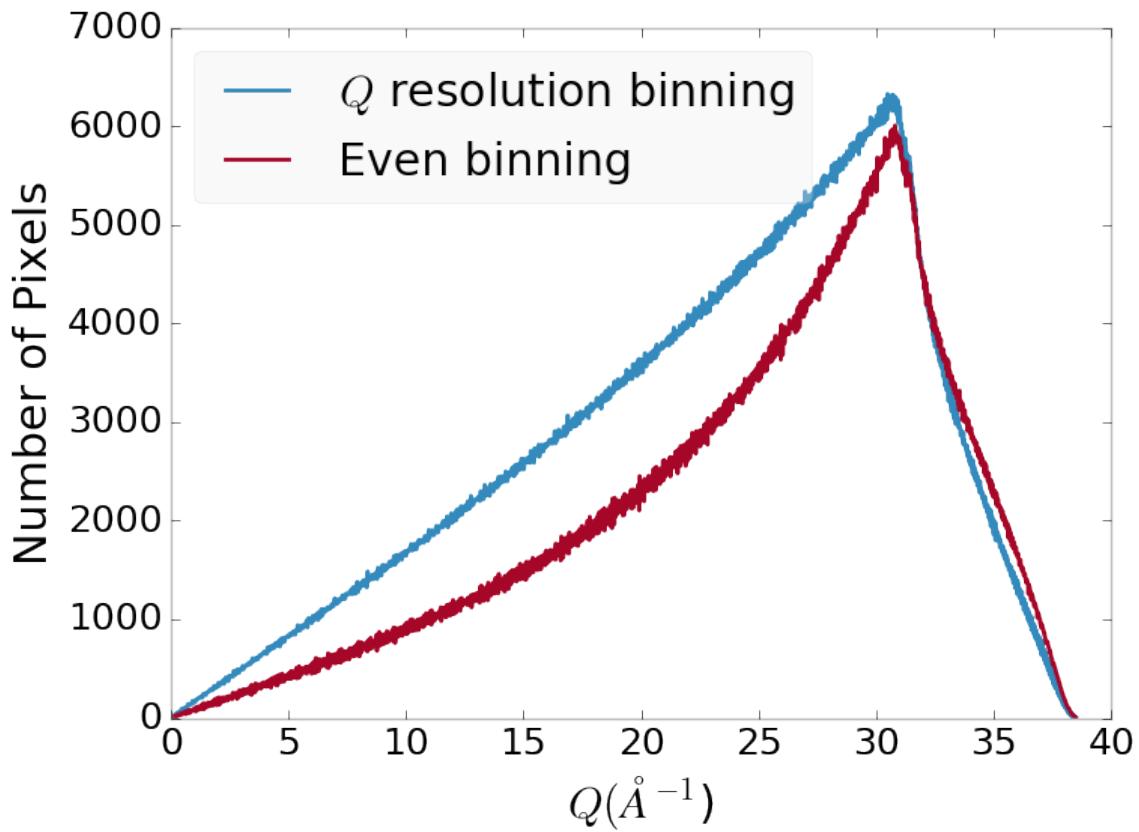


Figure 5.3: Number of pixels as a function of Q , binned at the Q resolution of the detector.

775 5.3 AUTOMATED MASK GENERATION

776 **Introduction**

777 Detector masking is an important part of any x-ray scattering workflow as dead/hot
 778 pixels, streak errors, and beamstop associated features can be averaged into the data
 779 changing the signal and its statistical significance. While some features, like the
 780 beamstop holder, can be easily observed and masked by hand other are much more
 781 difficult to observe even on large computer monitors. Additionally, while dead/hot
 782 pixels and streaks are usually static the hot pixels associated with textured or sin-
 783 gle crystal scattering or cosmic rays are not. Thus, coming up with an automated
 784 method for finding such erroneous pixels is important, especially as high flux diffrac-

785 tion beamlines can generate data very quickly.

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

795 **Algorithm Design**

796 The masking algorithm procedure takes in the image and a description of the pixel
797 positions in either distance from the point of incidence or in Q . The image is then
798 integrated twice, producing both the mean $I(Q)$ and the standard deviation of each
799 $I(Q)$ ring. The mask is created by comparing the pixel values against each ring's
800 standard deviation and threshold α . Note that the threshold can be a function of
801 distance from the point of incidence or Q .

802 **Test Cases**

803 To study the effectiveness of the masking we ran the algorithm against both simulated
804 and experimental data. In the case of the simulated data four systems were created:
805 1) dead/hot pixels with varying numbers of defective pixels, 2) beamstop holder with
806 varying beamstop holder transmittance, 3) rotated beamstop holder with varying
807 beamstop holder transmittance, and 4) beamstop holder with dead/hot pixels. The

808 base scattering was produced by

$$I = 100 \cos(50r)^2 + 150 \quad (5.3)$$

809 where r is a pixel's distance from the beam point of incidence. The positions of
810 the dead/hot pixels were chosen at random as was the dead or hot nature of the
811 defect. Dead pixels had values from 0 to 10, while hot pixels had values from 200
812 to 255. The beamstop was positioned at the vertical center of the detector with an
813 initial width of 60 pixels and final width of 120 pixels. The height of the beamstop
814 was 1024 pixels. The beamstop was calculated to attenuate the x-ray scattering
815 signal at various transmittance, as various beamstop holder materials have different
816 transmittance. Two version of the masking algorithm were run for each test case, one
817 using the standard even bin sizes for the integration step, and one where the bin sizes
818 are tuned to the pixel Q resolution as discussed in 5.2.

819 Results and Discussion

820 Three main studies were run each examining a different aspect of the simulated
821 or experimental studies. These included, masking bad pixels, masking a beamstop
822 holder, and masking experimental data. Figures 5.4-5.11 show the results of the
823 masking algorithm on simulated images. The dead/hot pixel masking shows the
824 importance of using the Q resolution based bin sizes as the even bin based mask have
825 a tendency to over mask the image, removing pixels which contain valuable signal.
826 This over-masking is caused by pixels being improperly associated with one another
827 by the even bins. Figure 5.4 indicates that the masking algorithm, with the proper
828 binning, masks the image perfectly, with no missed bad pixels or good pixels masked.
829 This is not the case in figures 5.5 - 5.7 as we can see pixels which should have been
830 masked but were not. Despite these missed pixels no pixels were improperly masked
831 in any of the well binned images. These test cases are actually more difficult than

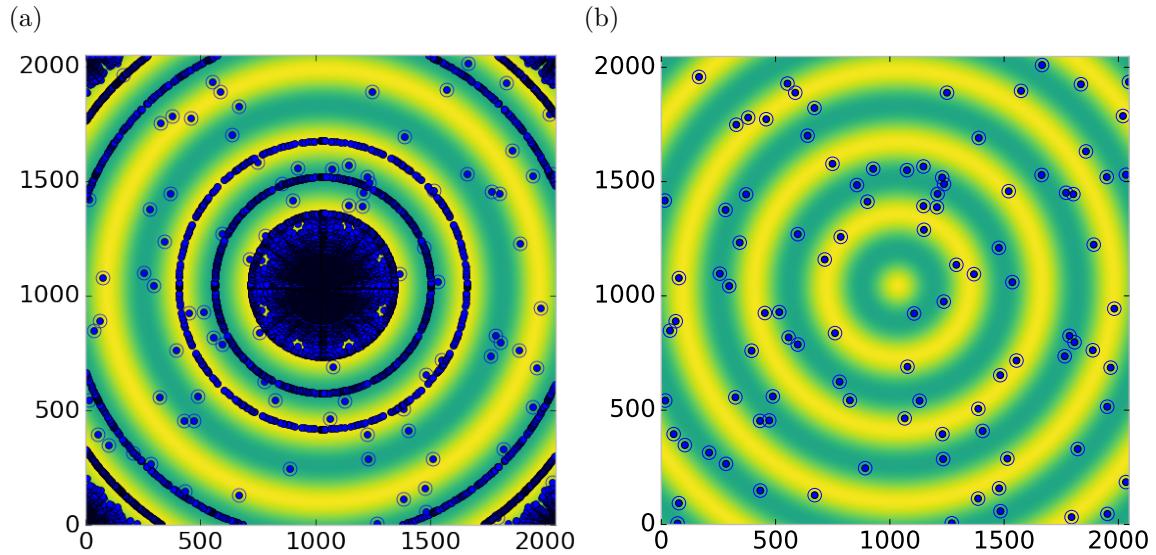


Figure 5.4: Generated dead/hot pixel masks for a detector with 100 bad pixels. a) the standard even bin mask and b) the Q resolution binned mask. The bad pixels are noted with open circles, masked pixels are noted with solid circles.

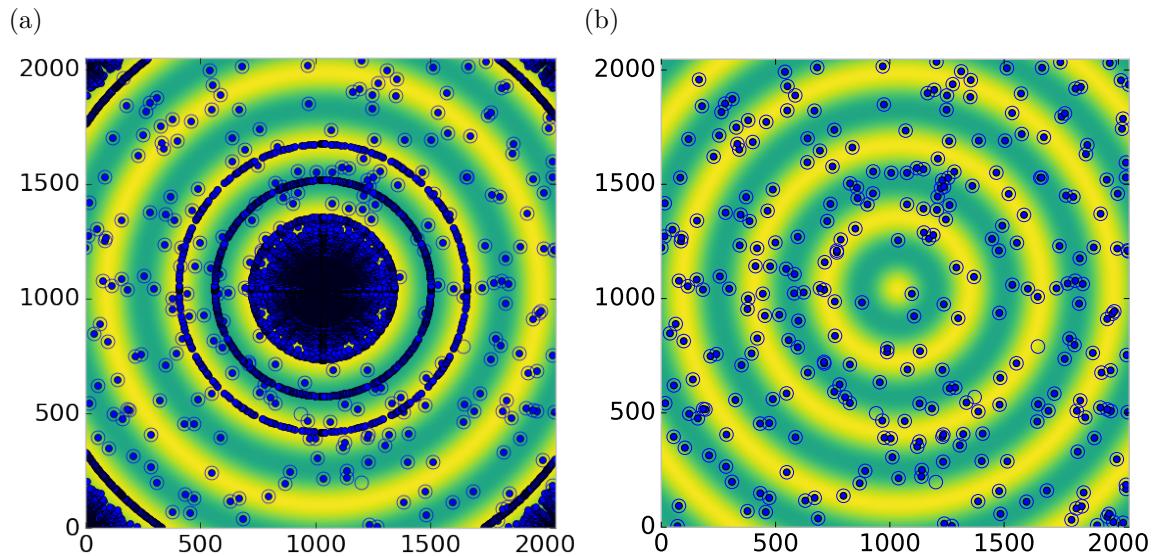


Figure 5.5: Generated dead/hot pixel masks for a detector with 300 bad pixels. a) the standard even bin mask and b) the Q resolution binned mask. The bad pixels are noted with open circles, masked pixels are noted with solid circles.

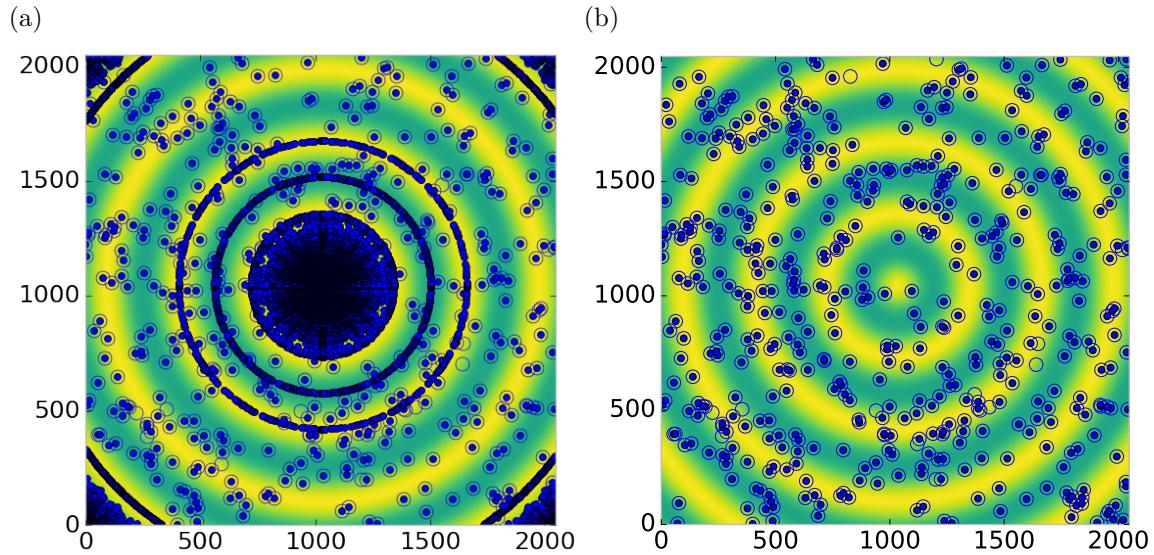


Figure 5.6: Generated dead/hot pixel masks for a detector with 500 bad pixels. a) the standard even bin mask and b) the Q resolution binned mask. The bad pixels are noted with open circles, masked pixels are noted with solid circles.

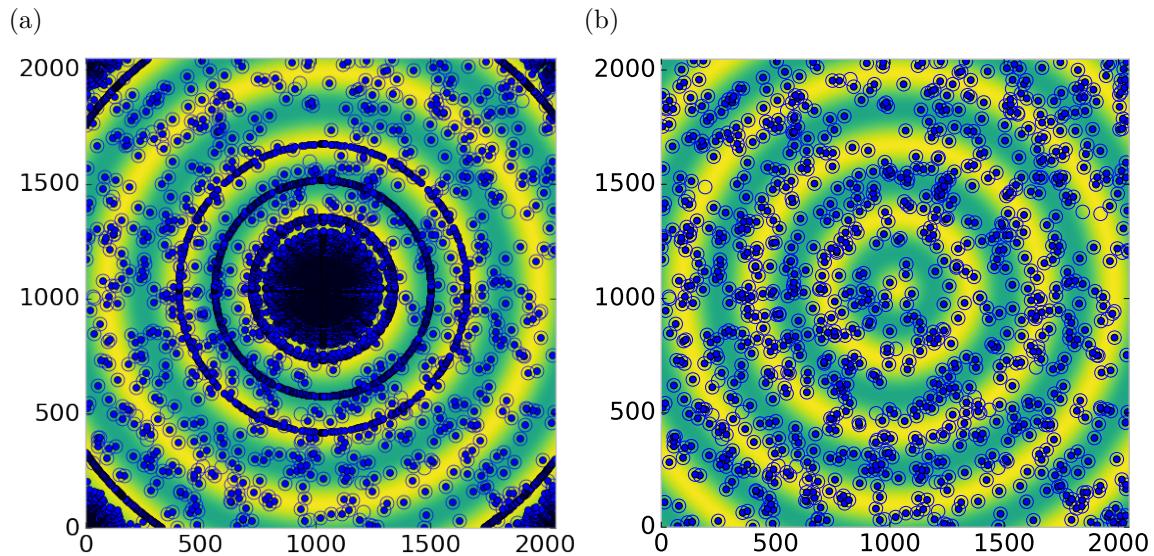


Figure 5.7: Generated dead/hot pixel masks for a detector with 1000 bad pixels. a) the standard even bin mask and b) the Q resolution binned mask. The bad pixels are noted with open circles, masked pixels are noted with solid circles.

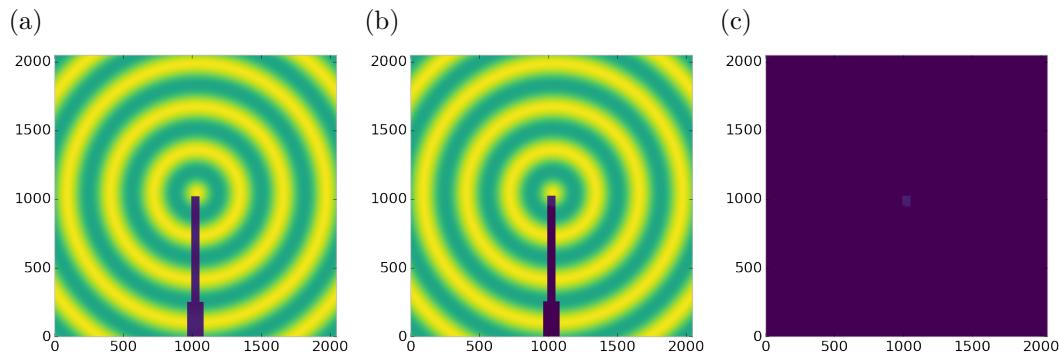


Figure 5.8: Generated beamstop holder masks for a beamstop holder with 10% transmittance. a) the raw image, b) the masked image, c) and the missed pixels

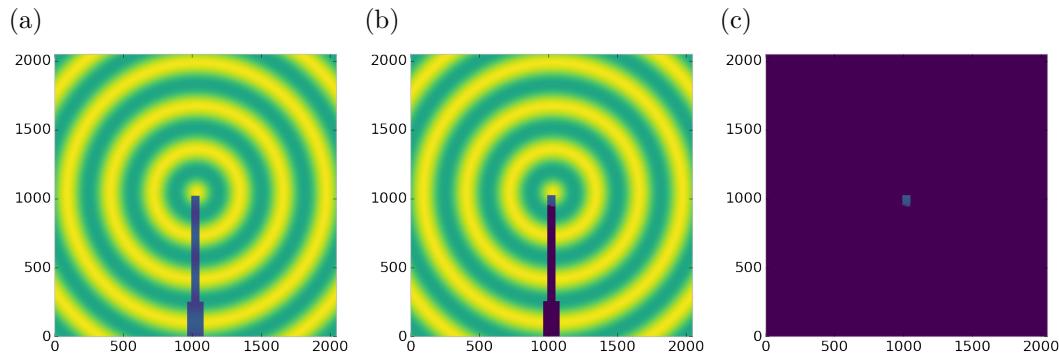


Figure 5.9: Generated beamstop holder masks for a beamstop holder with 30% transmittance. a) the raw image, b) the masked image, c) and the missed pixels

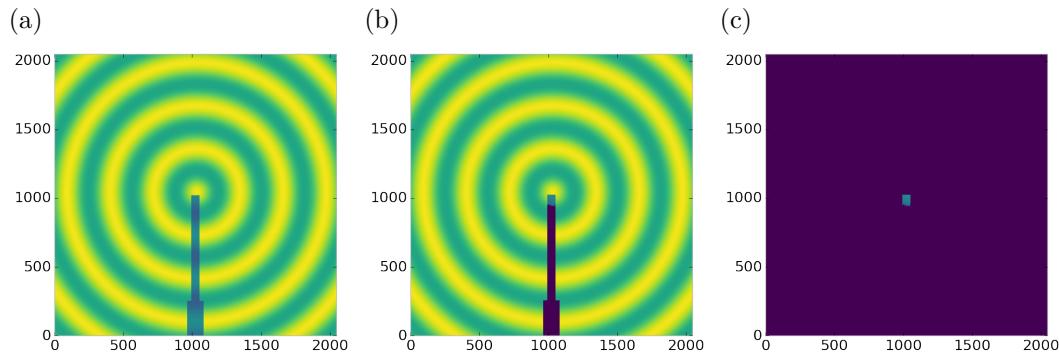


Figure 5.10: Generated beamstop holder masks for a beamstop holder with 50% transmittance. a) the raw image, b) the masked image, c) and the missed pixels

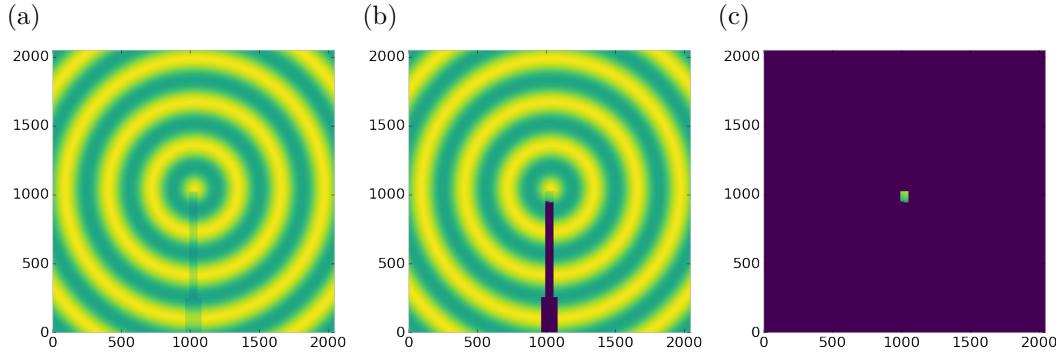


Figure 5.11: Generated beamstop holder masks for a beamstop holder with 90% transmittance. a) the raw image, b) the masked image, c) and the missed pixels

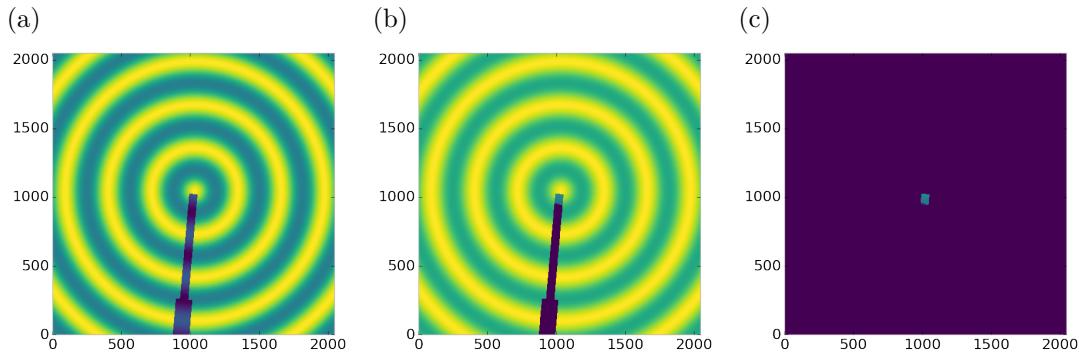


Figure 5.12: Generated beamstop holder masks which is rotated away from vertical

832 experimental data, as the dynamic range of most detector causes the dead/hot pixels
 833 and single crystal/texture peaks to be orders of magnitude away from the desired
 834 signal.

835 The beamstop holder masks shown in figures 5.8 - 5.11, which were all run with
 836 the Q resolution binning show similar results across the transmittance range, missing
 837 only a small part of the beamstop holder near the point of incidence. Near this point
 838 the beamstop holder becomes a statistically significant part of the total number of
 839 pixels in a given ring, thus it can not be masked out using a statistical search of the
 840 rings. For most PDF and XRD studies this small area can be masked automatically
 841 by masking all the pixels who's distance from the point of incidence is smaller than a
 842 given radius r , or can be neglected outright as the area is not used in the analysis or
 843 refinement. Similar results were produced for beamstop holders which were rotated

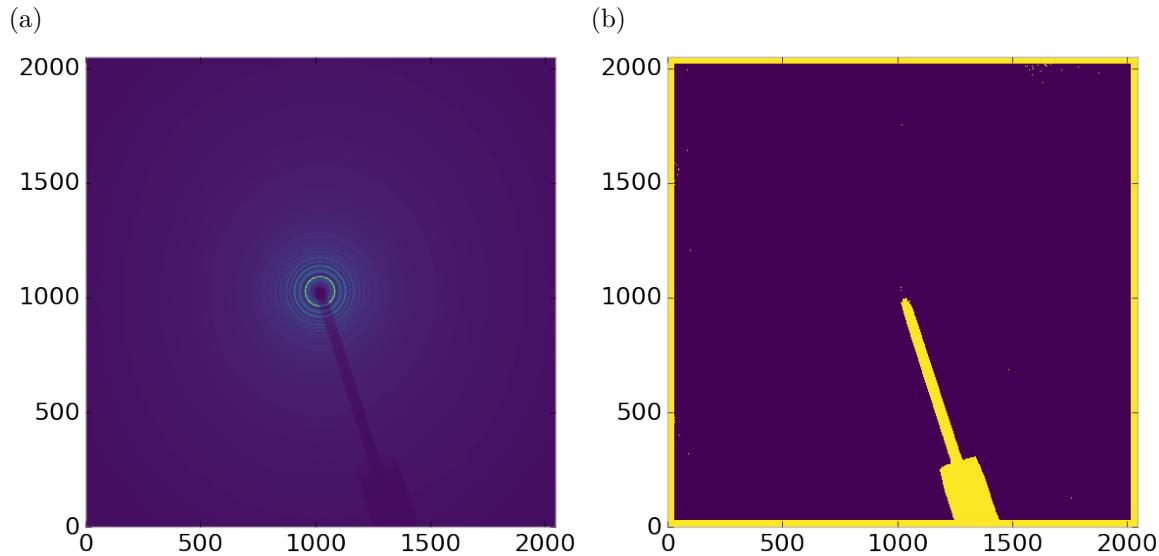


Figure 5.13: Masked experimental data. a) the raw image, b) the mask

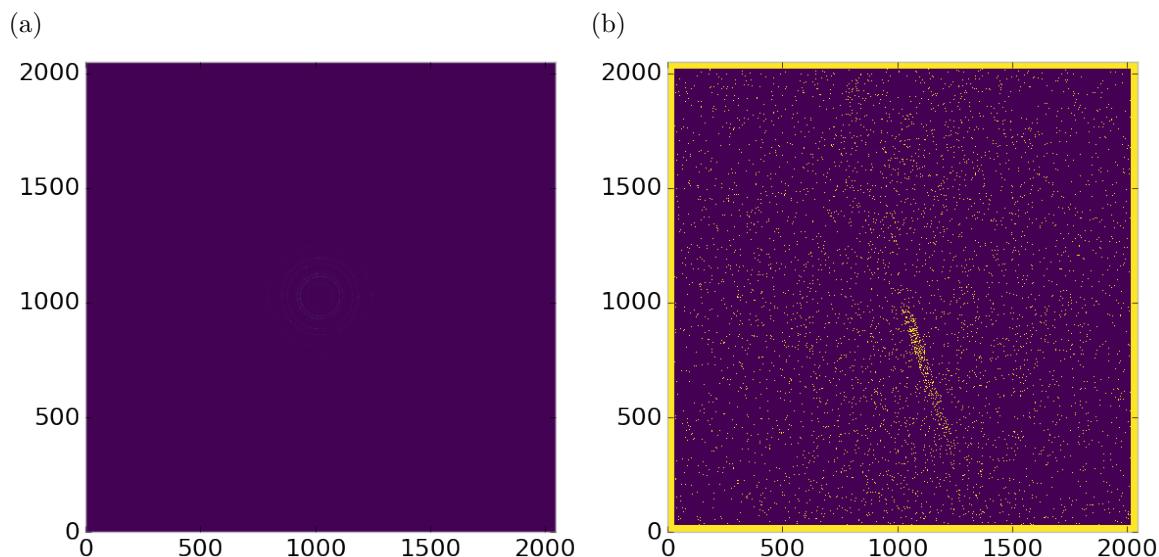


Figure 5.14: Masked experimental data with Pt single crystal signal. a) the raw image, b) the mask

844 away from the vertical position, as shown in figure 5.12

845 Working with actual experimental data, obtained at the Advanced Photon Source
 846 beamline 11-ID-B, shows the difficulty of masking images which have low photon
 847 counts. While the masking of experimental data taken with longer exposures, con-
 848 sisting of 250 .2 second shots, shown in figure 5.13 provides very sharp edges to the

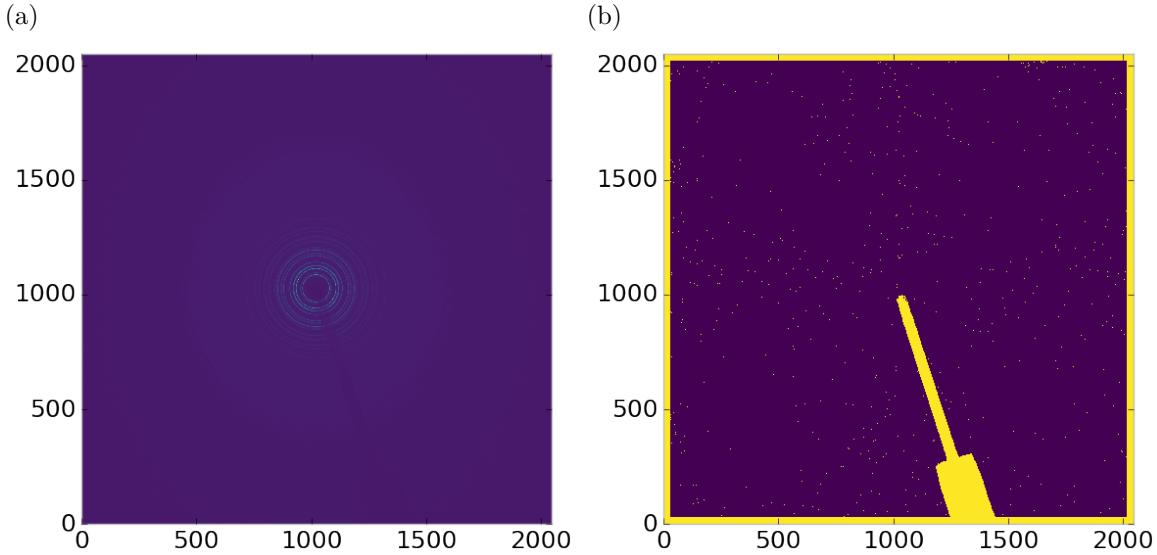


Figure 5.15: Masked experimental data with Pt single crystal signal using figure's 5.13 mask as a starting mask. a) the raw image, b) the mask

beamstop holder, and very little extra masking beyond the occasional dead pixel, this is not the case for the single crystal data. The single crystal data is more problematic because of its short exposure time and low flux, with 500 frame at a .1 second exposure and having shrunk the beam size. The low flux is to prevent the very strong single crystal peaks from damaging the detector. However, this causes the image to be less statistically viable than ideal, causing problems with the mask as seen in figure 5.14. This can be alleviated to some degree by using the previously generated mask as a starting mask for the single crystal image, as shown in 5.15. While the masking algorithm still produces many diffuse masked pixels, they are far fewer, this may be due to the removal of the beamstop which could have contributed to the large standard deviation in figure 5.14.

Conclusions

In this section the masking algorithm, which relies on both Q resolution based binning and a statistical approach to azimuthal symmetry, was developed. The focus of

863 this algorithm was to remove many unwanted detector features associated with pixel
864 defect, beamstop holder associated scattering attenuation, and single crystal/texture
865 peaks. Simulated data was used to evaluate the beamstop holder and dead/hot pixel
866 masking capacity, while experimental data was used to check for single crystal and
867 texture based masking. Q resolution based binning was shown to be very important
868 to avoid over-masking. The ability of the mask writer to mask images is somewhat
869 limited by the overall statistical image quality, although some deficiencies can be
870 obtained by using previously generated masks as starting points. This masking algo-
871 rithm is now in use in the data processing workflow and will be available in scikit-beam
872 soon.

873 5.4 AUTOMATED IMAGE AZIMUTHAL INTEGRATION

874 Using the Q resolution binning and masking developed in sections 5.2 and 5.3 the
875 images can be properly integrated. Generally, images are integrated by taking the
876 mean value of the pixels in a ring. However, other statistical measures of the average
877 value can be used, like the median. Note that all the integrations done here use the
878 pixels as they are, without pixel splitting, minimizing the covariance of the resulting
879 $I(Q)$.[54]

880 Figures 5.16-5.18 show the importance of masking and the choice of average func-
881 tion. All the figures were produced using the same dataset, 50 °C Pr_2NiO_4 taken at
882 the APS’s 11-ID-B on a Perkin Elmer area detector. The automatic masking alpha
883 was 3 standard deviations from the mean. While it is difficult to observe the changes
884 the mask causes in the full $I(Q)$ plot (subfigures a) and b)), the standard deviation
885 plots show the effect of bad pixels on the data (subfigure c)). Subfigure c) for figures
886 5.16-5.18 shows that removal of the beamstop holder lowers the low Q standard de-
887 viation from around .1 to almost .01 out to 15 \AA^{-1} . The high Q subfigures d) and f)
888 in figures 5.16-5.18 show the “kink” effect of the detector edge and beamstop holder,

889 where there is a dip in the $I(Q)$ scattering when the rings include the edge of the
890 detector. This effect seems to be due to both errors in the edge pixel intensity and the
891 beamstop holder as masking of the edges only seems to provide only partial removal
892 of the issue. It is important to note that while integration using the mean of the
893 ring has issues with only the edge mask, as evidenced by the change in slope in 5.17
894 d) around 29.5 \AA^{-1} , the median integration does not include this error. Ideally the
895 detector would have a normal distribution of pixel intensity for a given ring, which
896 would imply an equivalency between the mean and median $I(Q)$ values. Despite the
897 closeness of the mean and median once the final mask has been created, it seems that
898 the median is more reliable, as it was less effected by the beamstop holder in figure
899 5.17. Thus, for subsequent integrations discussed in this work the median is used to
900 avoid any defective features that the masking algorithm may have missed.

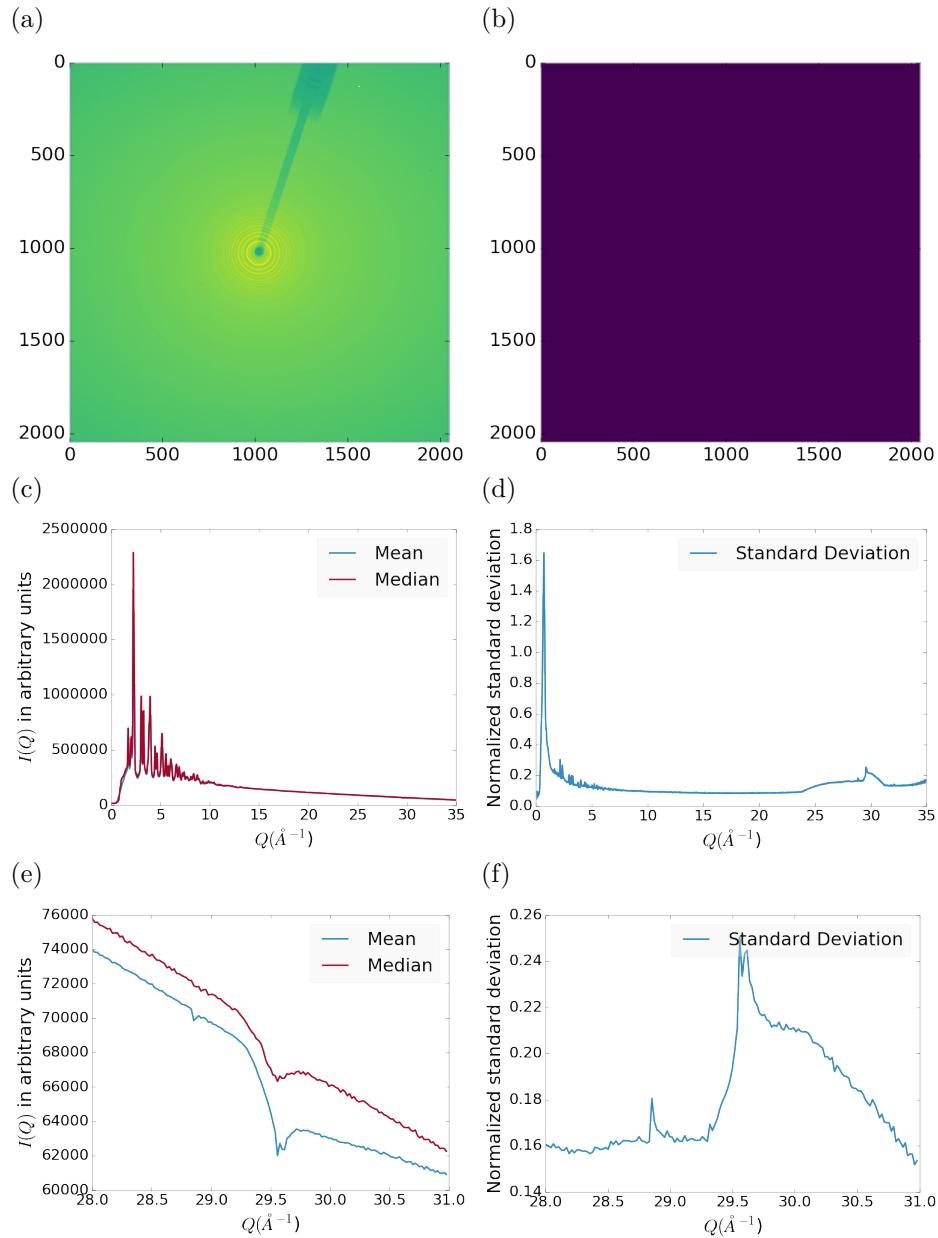


Figure 5.16: Masking, average, and standard deviation of an example x-ray total scattering measurement. This image was produced with no mask. a) the image, b) the mask, c) the mean and median values, d) the standard deviation (normalized to the median), e) a closeup of the 28 \AA^{-1} to 31 \AA^{-1} Q range for the mean and median, f) 28 \AA^{-1} to 31 \AA^{-1} Q range for the standard deviation

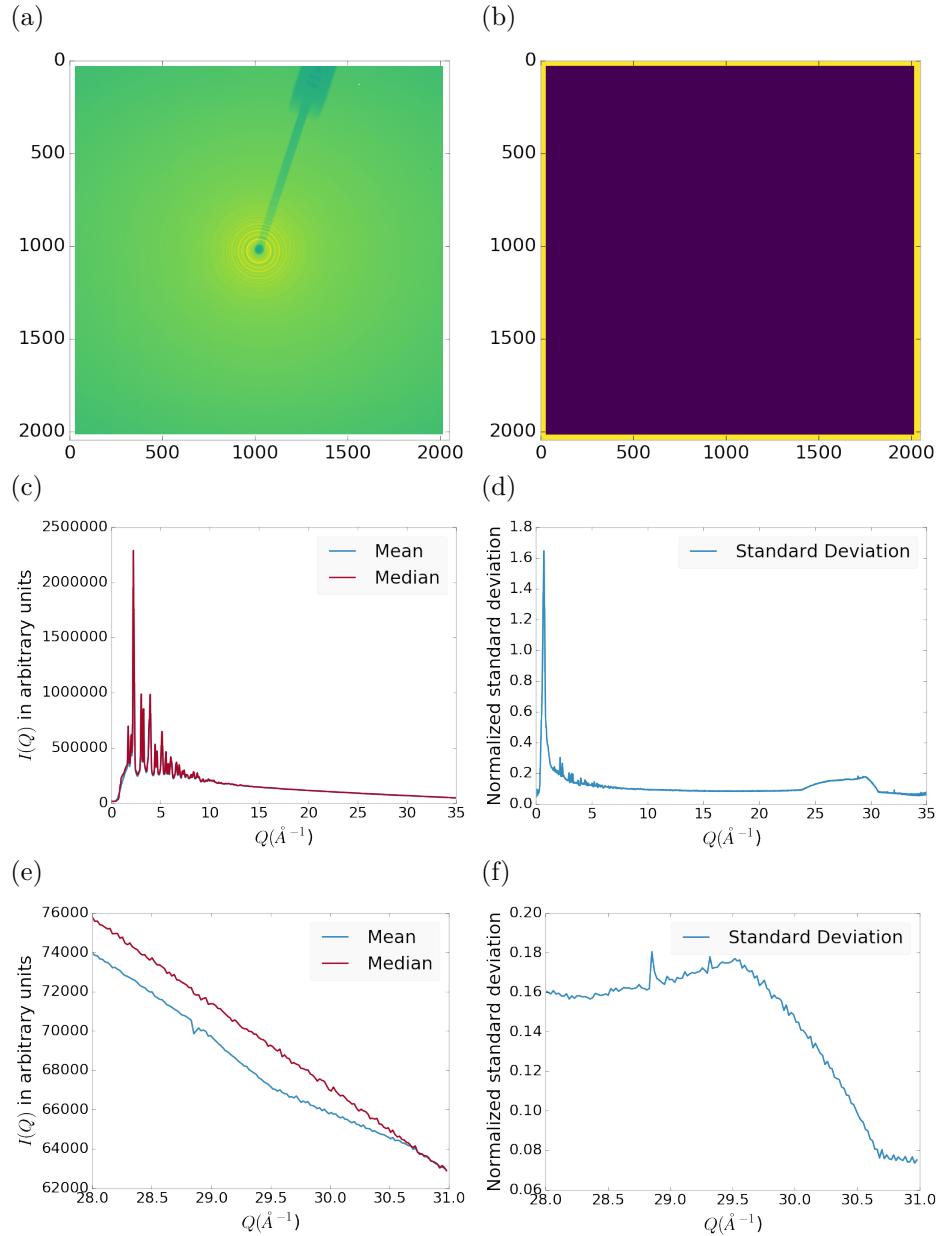


Figure 5.17: Masking, average, and standard deviation of an example x-ray total scattering measurement. This image was produced with only an edge mask. a) the image, b) the mask, c) the mean and median values, d) the standard deviation (normalized to the median), e) a closeup of the 28 \AA^{-1} to 31 \AA^{-1} Q range for the mean and median, f) 28 \AA^{-1} to 31 \AA^{-1} Q range for the standard deviation

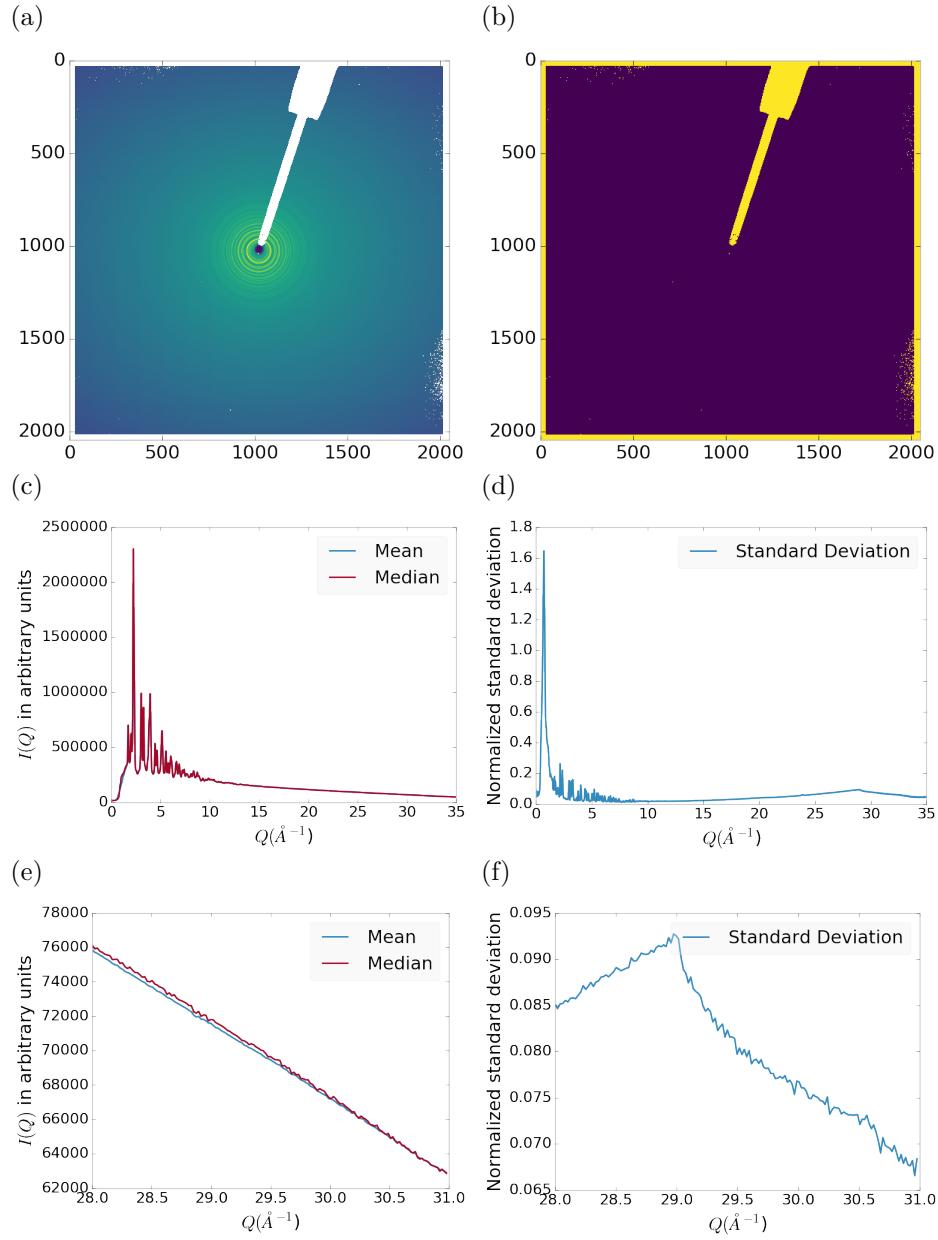


Figure 5.18: Masking, average, and standard deviation of an example x-ray total scattering measurement. This image was produced combining an edge mask and the automatically generated mask. a) the image, b) the mask, c) the mean and median values, d) the standard deviation (normalized to the median), 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

901 5.5 CONCLUSIONS

902 This chapter developed and analyzed the proper data processing and reduction method-
903 ology for producing reliable $F(Q)$ data from x-ray total scattering measurements.
904 Binning at the Q resolution of the detector was found to be key to the data process-
905 ing. The primary outcome of using the Q resolution binning was an enhancement
906 in effectiveness for the masking algorithm, producing much fewer false positives for
907 dead/hot pixels. This masking approach was then applied to the integration of ex-
908 perimental data taken at the APD's 11-ID-B beamline. The automatically generated
909 masks, when combined with edge masks, were found to greatly reduce the overall
910 standard deviation of the pixel intensity and produce a smoother $F(Q)$ at high Q ,
911 enabling the use of much higher Q data in the PDF. Different statistical measures
912 used in the azimuthal integration was also compared. This comparison showed that
913 the median was a more reliable statistic for integration with data which had more
914 detector defects. However, upon properly masking it was shown that these metrics
915 were almost identical. The masking induced similarity between the mean and me-
916 dian shows that the rings, when integrated, may form a Gaussian distribution. The
917 distribution of the pixel intensities for strongly and weakly scattering samples may
918 be investigated in future work.

919 CHAPTER 6

920 PHASE CHANGES AND ANNEALING DYNAMICS OF
921 Pr_2NiO_4 AND ITS DERIVATIVES

922 6.1 INTRODUCTION

923 Pr_2NiO_4 (PNO) electrodes provide higher power density than $\text{La}_{0.8}\text{Sr}_{0.2}\text{MnO}_3$ (LSM),
924 and is more stable than $(\text{La}_{0.60}\text{Sr}_{0.40})_{0.95}(\text{Co}_{0.20}\text{Fe}_{.80})\text{O}_{3-\text{x}}$ (LSCF), which is known to
925 rapidly degrade in performance. [57] PNO's high performance between 600-900 °C is
926 associated with its high activity towards the oxygen reduction reaction (ORR), which
927 stems from PNO's high oxygen diffusion and surface exchange coefficients, substantial
928 oxygen over-stoichiometry, and large oxygen ion conduction paths through the unit
929 cell. [55] Despite these advantages, PNO's tendency to partially decompose into
930 PrOx and other phases is particularly challenging. [14] Full cell operation after 500
931 hours at 750 °C and 0.8 V shows major decomposition of the parent PNO phase,
932 while the performance degrades by only 4%. Such significant changes in phase and
933 relatively small changes in performance further assure the necessity for understanding
934 the phase evolution in nickelate cathodes during operation. To address these disparity
935 in performance and phase degradation PDF and XRD analysis may be able to examine
936 these issues from both long and short range ordering perspectives.

937 6.2 EXPERIMENTS

938 **Pr₂NiO₄ Synthesis**

939 Pr₂NiO₄ was synthesized using the standard approach, as detailed in the work by
940 Dogdibegovic et. al. [14] The nickelate powder was initially aprepared via the
941 glycine-nitrate process. This was followd by thermal annealing at 1080 °C for 10
942 hours in air.

943 **X-ray Measurements**

944 X-ray total scattering and x-ray powder diffraction experiments were performed at
945 the APS's 11-ID-B beamline. An x-ray energy of 86.7 keV, .145 Åwas provided
946 by the beamline monochromator. The detector was moved between a 20cm and a
947 95 cm sample to detector distance to measure the x-ray total scattering and x-ray
948 diffraction patterns. Various PNO samples were annealed on the beamline during
949 x-ray measurement.

950 6.3 DATA PROCESSING

951 The data was calibrated at each of the detector positions using a CeO₂ standard
952 via pyFAI. [30] The images were corrected for a .95 x-ray polarizetion. Masks were
953 produced for both the foreground and background images. The forground masks were
954 produced using both a 30 pixel edge mask and a 2.5σ automatic mask as discussed
955 in chapter 5. The background masks were produced by using the forground mask as
956 a starting mask with a 2.5σ automatic mask.

957 The foreground and background images were then integrated using the Q resolu-
958 tion binning discussed in chapter 5. The resulting $I(Q)$ data were corrected for their
959 number of frames and I_{00} . Finally the corrected background $I(Q)$ was subtracted
960 from the foreground $I(Q)$.

961 Each PDF was generated with a Q_{min} of 1.5, Q_{max} of 29., R_{poly} of .9, R_{max} of 40.
962 descriptions of these parameters can be found in the work by Juhas et. al. [28]

963 6.4 DATA ANALYSIS

964 **Intra Sample Comparison**

965 **PDF**

966 As figures 6.1 and 6.2 show the as synthesized PNO undergoes very little change in
967 structure according to the PDF. The PDF does show some broadening at around 3.5
968 and 8.5 Å, but the peak shifts themselves are fairly limited. This implies that the as
969 synthesized PNO structure is stable at least for the 1 hour that the sample was held
970 at 750 °C.

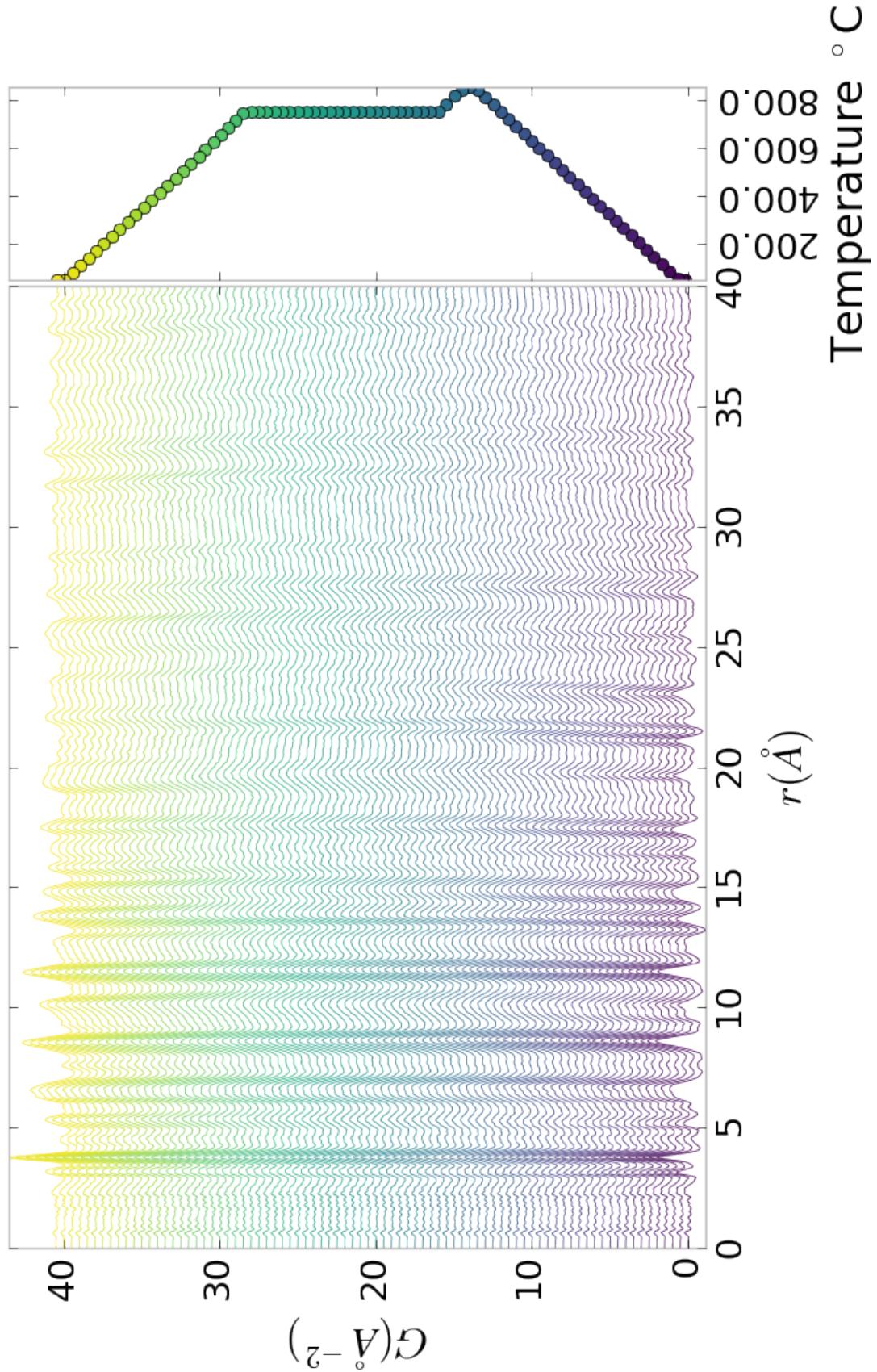


Figure 6.1: PDF as a function of temperature for as synthesized PNO showing the full PDF

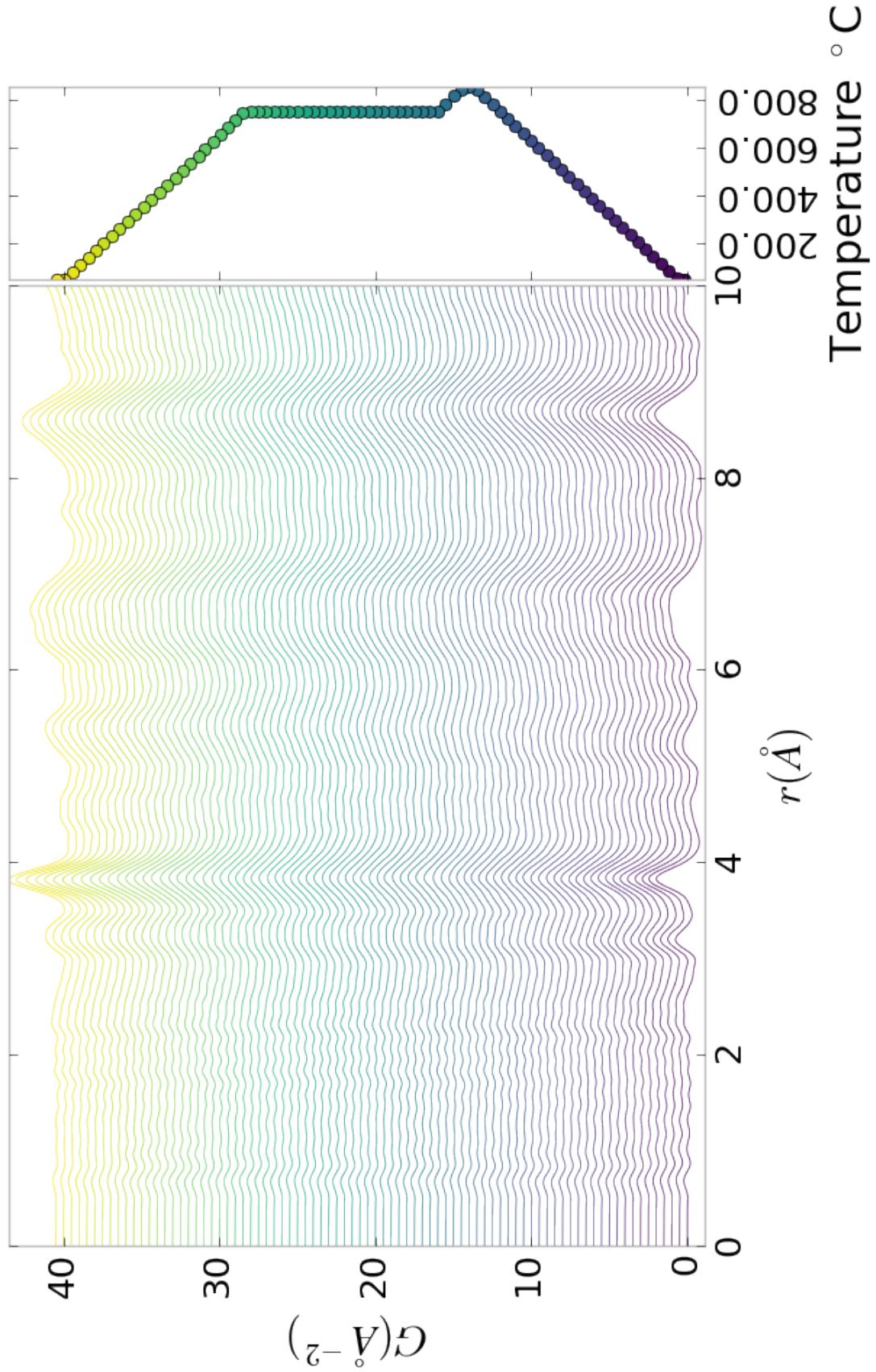


Figure 6.2: PDF as a function of temperature for as synthesized PNO showing a close up on the short range section

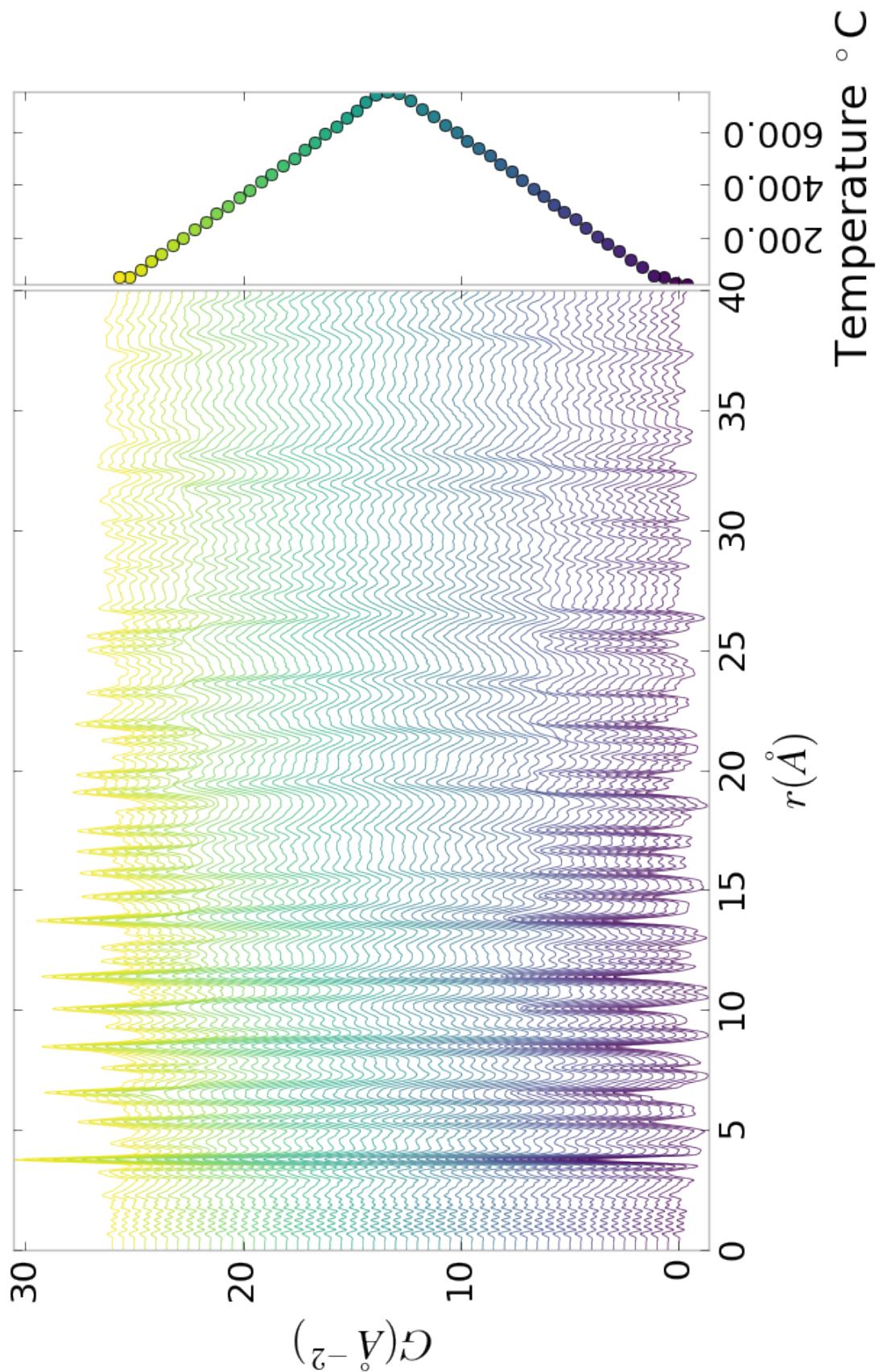


Figure 6.3: PDF as a function of temperature for PNO annealed at 750 °C for 25 hours showing the full PDF

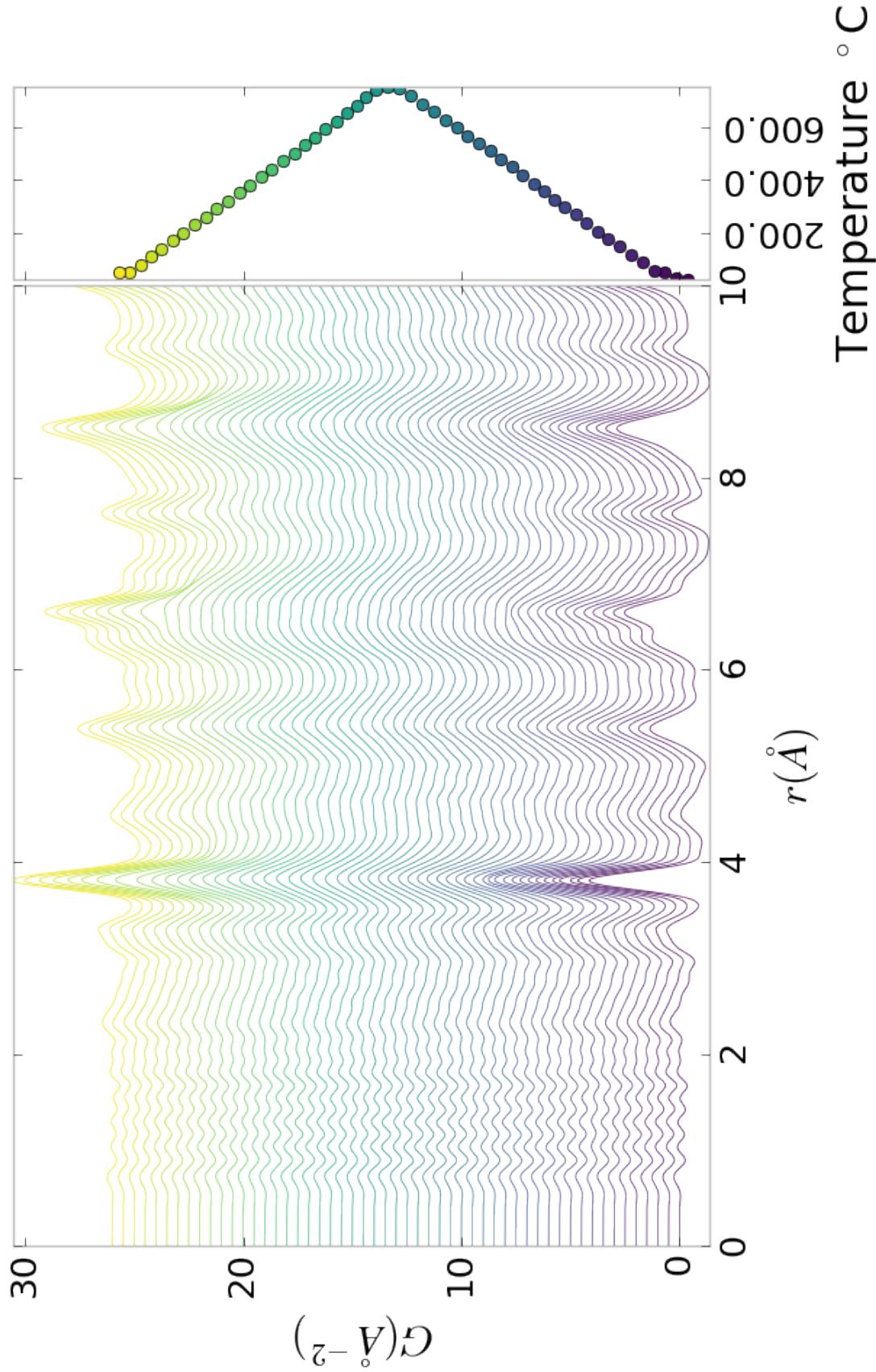


Figure 6.4: PDF as a function of temperature for PNO annealed at $750\text{ }^{\circ}\text{C}$ for 25 hours showing a close up on the short range section

971 $I(Q)$

972 The annealed samples figures, 6.3 and 6.4, tell a rather different story. In this case the
973 PDF shows significant peak shifts and broadening, especially at higher interatomic
974 distances. Some peaks completely disappear, like the peak at 12 Å. Similar results were
975 also observed for samples with longer annealing times, as shown in the appendix.

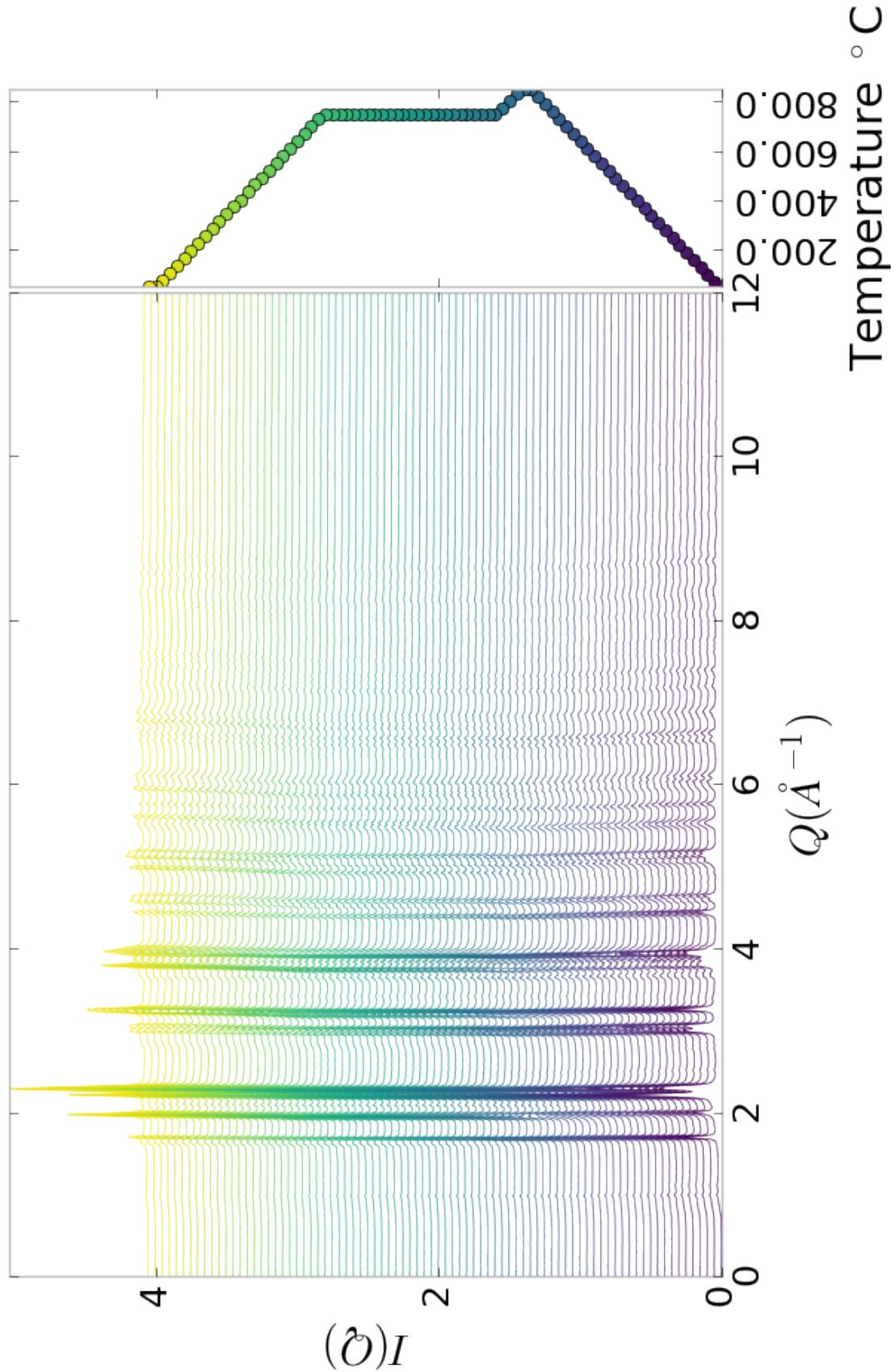


Figure 6.5: $I(Q)$ as a function of temperature for as synthesized PNO showing the full XRD

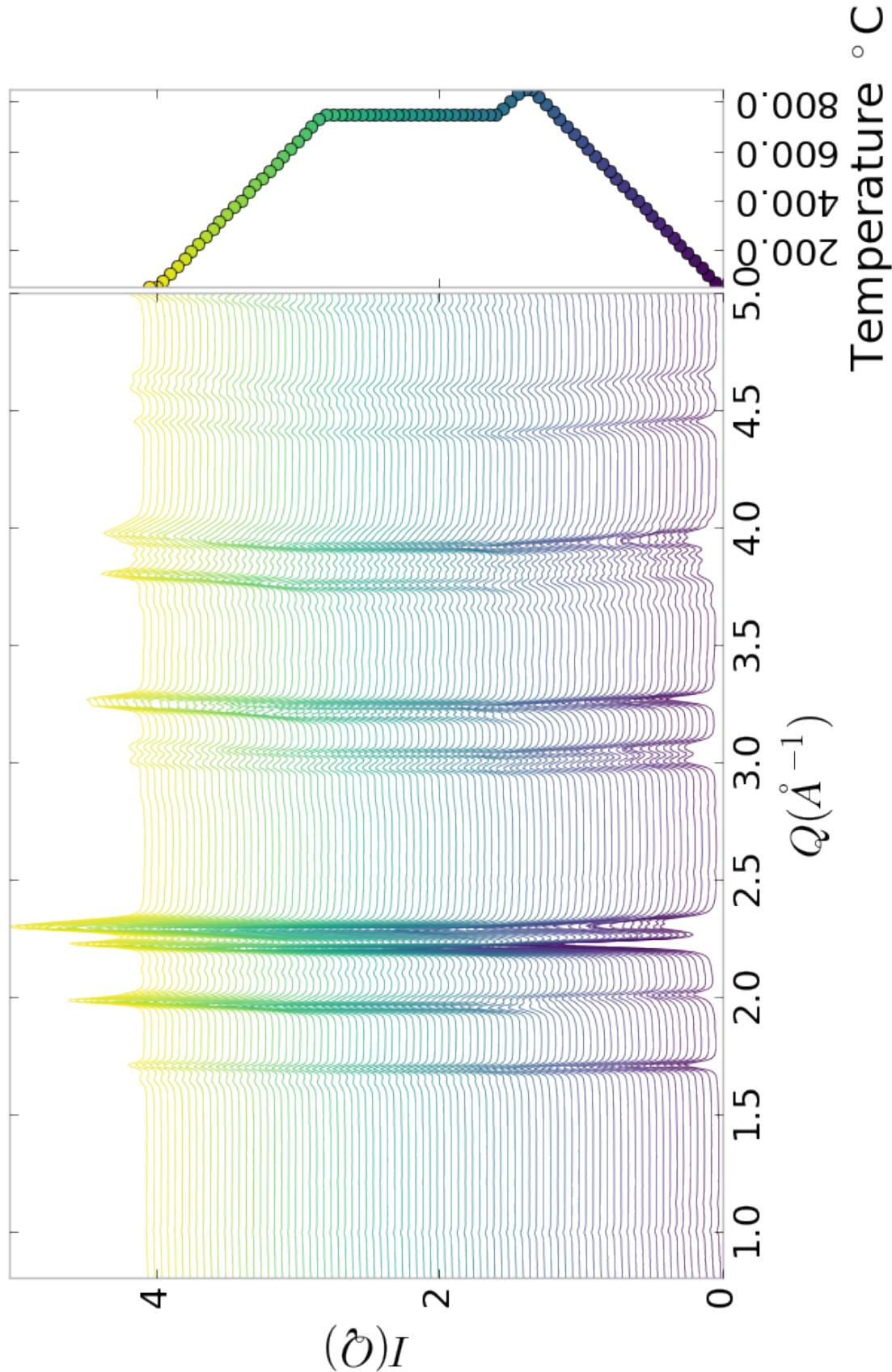


Figure 6.6: $I(Q)$ as a function of temperature for synthesized PNO showing a close up on the low Q section

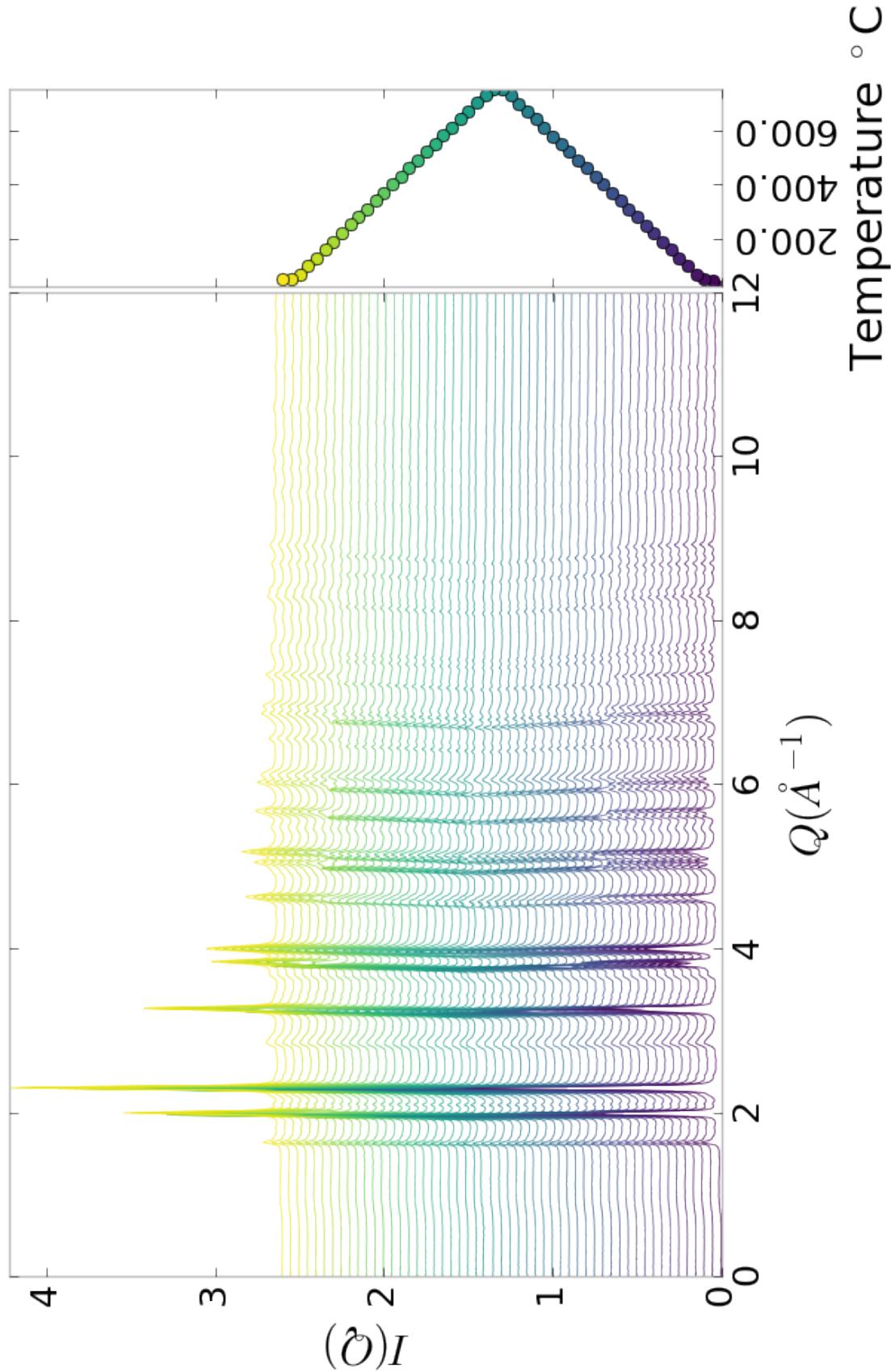


Figure 6.7: $I(Q)$ as a function of temperature for PNO annealed at 750 °C for 25 hours showing the full XRD

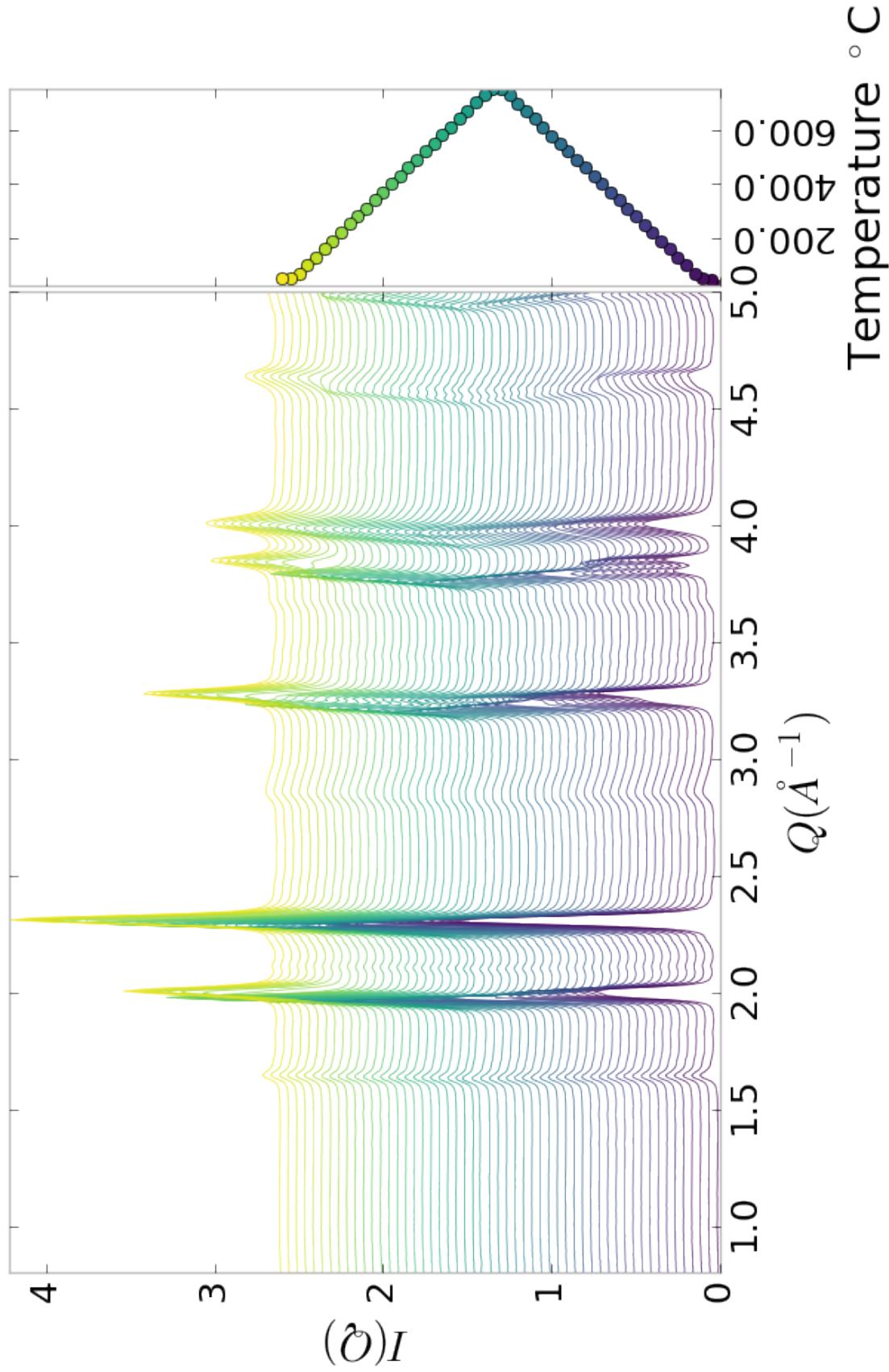


Figure 6.8: $I(Q)$ as a function of temperature for PNO annealed at 750 °C for 25 hours showing a close up on the low Q section

976 **Inter Sample Comparison**

977 Figures A.26 and A.20 show a very interesting contrast. Figure A.26 show significant
978 differences in the $I(Q)$ between the as-synthesized and annealed PNO, which could
979 be associated with the more degradation present in the annealed samples. However,
980 figure A.20 shows very little difference in the PDF between the various annealing
981 times. This discrepancy seems to point to some kind of disorder which changes the
982 interatomic distances very little but changes the symmetry enough to change the
983 Bragg reflections.

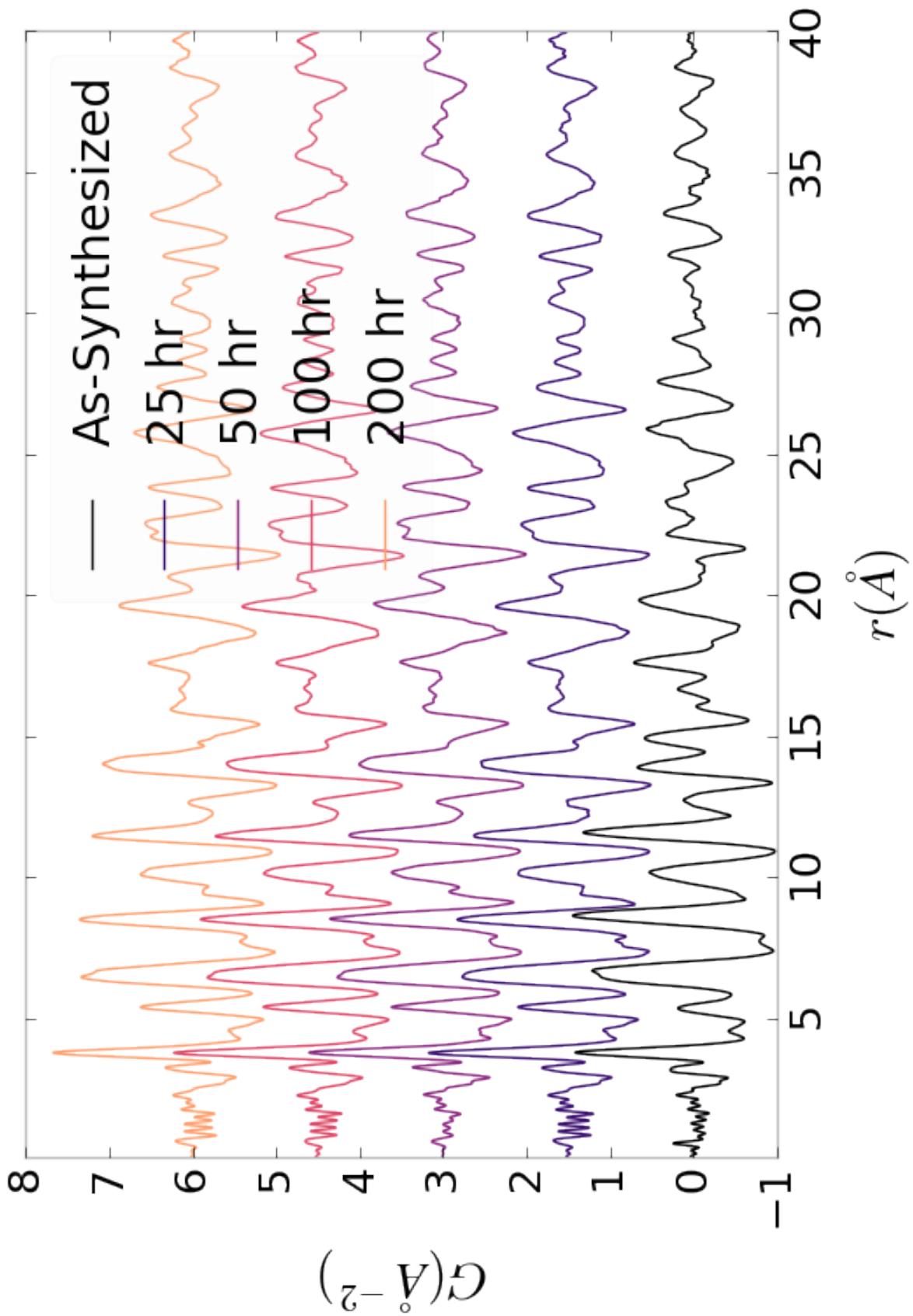


Figure 6.9: Comparison of PNO sample PDFs as a function of annealing time high-temp

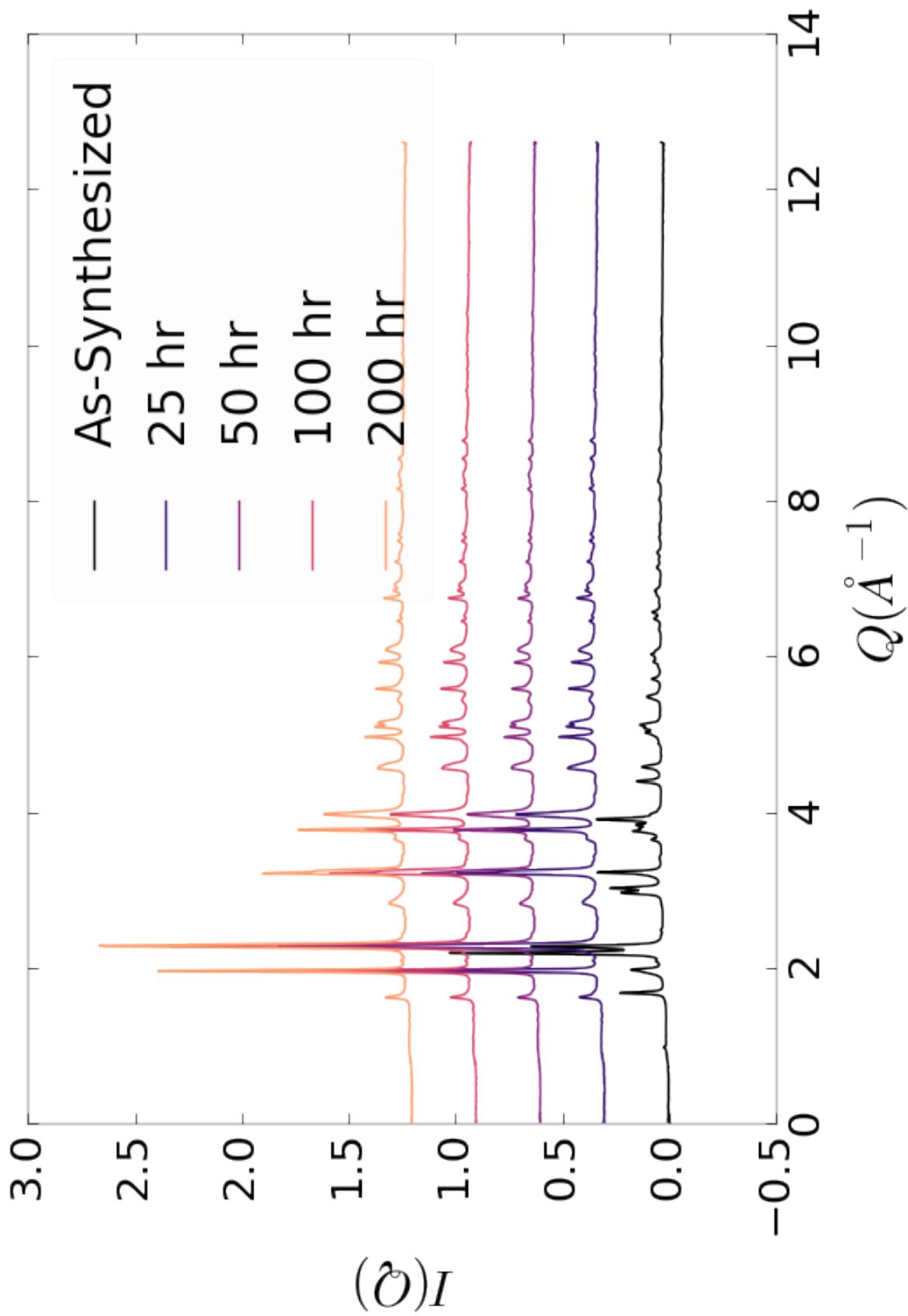


Figure 6.10: Comparison of PNO sample $I(Q)$ as a function of annealing time high-temp

984 6.5 SIMULATION

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

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

997 6.6 CONCLUSIONS

998 X-ray total scattering and x-ray powder diffraction data was obtained on Pr_2NiO_4
999 powder samples annealed for various lengths of time. In-situ studies on the beamline
1000 were performed to understand how the structure of each of these powders changes
1001 at operating temperatures. The data was processed with the previously discussed Q
1002 binning, masking, and integration methodology. The PDF results show very little
1003 change in the structure for the as synthesized sample. However, the PDFs show a
1004 large change in the previously annealed samples. These changes seem to reporduce
1005 PDFs similar to the as-synthesized PNO at operating temperatures. This would seem
1006 to imply that the source of the anamolus PNO phase/power density relationship may
1007 be due to the adoption of an active structure upon heating which is universal despite

1008 the amount of thermal degradation observed at room temperature. In contrast to the
1009 PDF results, the XRD results seem to show signifigant changes in the PNO structure,
1010 both with ex-situ and in-situ annealing. The XRDs show the degradation of the PNO
1011 into various phases, potentially including Pr_2O_{11} , and higher ordered Pr based phases.
1012 The discrepancy between these two results is quite interesting as it seems that the
1013 XRD and PDF results are contradictory. Turbostratic diplacements between the
1014 layers may be a cause of the PDF/XRD disagreement, as these changes would cause
1015 very little change in the local structure observed in the PDF, while causing large
1016 changes in the XRD.

1017 CHAPTER 7

1018 CONCLUSION

1019 The work here presents one of the most complete end to end approaches to pro-
1020 cessing, analyzing, and simulating atomic pair distribution function data. The goals
1021 of this work were to build a modular, quick, and robust method for handling both
1022 experimental PDF data and solving atomic structures from said data.

1023 The statistical mechanical PES solvers were designed to robustly find atomic
1024 solutions which are global minima of the PES. This was accomplished by using some
1025 of the most advanced Monte Carlo algorithms and samplers. The analytical equations
1026 for the PES and its gradients were derived to provide the quickest searches.

1027 The PDF gradients were derived and implemented as GPU kernels to furthur
1028 speed up the PES search. The inclusion of the GPUs, combined with the atom pair
1029 mapping, were found to provide a 10x to 100x speedup over a multiprocessed CPU
1030 methodology.

1031 The extensive benchmarking of the NUTS-HMC system presented in chapters 2
1032 and 3 showcased the system's robustness, speed, and effectiveness. Interestingly it
1033 seems the the simulations also helped to elucidate the relationship between Rw and
1034 the resulting fit of the secondary metrics, including radial bond distribution. This is
1035 particularly important as it begins to establish Rw goals and a relationship between
1036 Rw and the confidence that features from the underlying structure that the PDF
1037 represents are reproduced by the structural model which is produced my Monte Carlo
1038 modeling. It seems that the theshold for acceptable Rw in Monte Carlo modeling
1039 needs to be quite lower than the current literature standards to properly reproduce

1040 the structure.

1041 A novel data processing workflow was also developed which focused on using Q
1042 resolution binning to create masks automatically and azimuthally integrate. The Q
1043 resolution binning provided a significant improvement in the automated masking
1044 robustness, leading to much fewer false positives, as shown by a series of masks
1045 generated on simulated and experimental data. The effect of these masks on the
1046 median and mean azimuthal integration was also discussed, establishing masks as
1047 very important to the removal of the high Q “kink” seen in 2D area detector data.
1048 Furthermore, a comparison was drawn between the median and mean integration,
1049 showing the median to be more reliable than the mean when working with data that
1050 could have residual detector defects. Overall the masking scheme was shown to reduce
1051 the standard deviation of the data significantly.

1052 Finally, preliminary results of x-ray total scattering measurements on Pr_2NiO_4
1053 were presented. Interestingly, these results show a strong discrepancy between the
1054 PDF and $I(Q)$ data. Where the PDF shows a very static as synthesized structure,
1055 despite annealing, the associated $I(Q)$ data shows peak movement and formation. For
1056 the pre-annealed samples both the $I(Q)$ and PDF data show peak changes. Inter-
1057 estingly, the PDFs of the as synthesized and pre-annealed samples show very similar
1058 local structure at operating temperatures.

1059 Despite all the work presented here, there is of course much more to be done.
1060 Implementing new ensembles, like Parallel Tempering, and faster Grand Canonical
1061 Monte Carlo, may help to find solutions faster and with less user based parameter
1062 tuning. Building the mathematics and software to quickly compute the data from
1063 other atomistic experiments, including EXAFS, STEM, and neutron scattering, could
1064 help to produce structures which more fully describe all the available experimental
1065 data. Implementing the existing codebase in a more general High Performance Com-
1066 putting context would allow for the solution of much larger particles, and extended

1067 solids. Furthur benchmarking will help to probe the robustness of the algorithm
1068 with other systems, including systems with periodic boundary conditions. Faster
1069 scattering data processing will enable a quicker total turn around time from taking
1070 experimental images to producing atomic structures.

1071 Even without these enhancements it is expected that this work will become a stan-
1072 dard method for solving atomic structures from x-ray total scattering experiments,
1073 having presented one of the most complete end to end x-ray processing and analysis
1074 systems.

BIBLIOGRAPHY

- 1076 [1] Milinda Abeykoon, Christos D. Malliakas, Pavol Juhás, Emil S. Božin, Mer-
1077 couri G Kanatzidis, and Simon J L Billinge, *Quantitative nanostructure char-*
1078 *acterization using atomic pair distribution functions obtained from laboratory*
1079 *electron microscopes*, Zeitschrift fur Kristallographie **227** (2012), no. 5, 248–256.
- 1080 [2] Alexandre A. Arnold, Victor Terskikh, Qian Ying Li, Rafik Naccache, Isabelle
1081 Marcotte, and John A. Capobianco, *Structure of NaYF₄ upconverting nanopar-*
1082 *ticles: A multinuclear solid-state NMR and DFT computational study*, Journal
1083 of Physical Chemistry C **117** (2013), no. 48, 25733–25741.
- 1084 [3] Simon J L Billinge and Takeshi Egami, *Underneath the Bragg Peaks: Structural*
1085 *Analysis of complex Materials*, vol. Volume 16, Pergamon, 2012.
- 1086 [4] Simon J L Billinge and Igor Levin, *The problem with determining atomic struc-*
1087 *ture at the nanoscale.*, Science (New York, N.Y.) **316** (2007), no. 5824, 561–565.
- 1088 [5] P. E. Blöchl, *Projector augmented-wave method*, Physical Review B **50** (1994),
1089 no. 24, 17953–17979.
- 1090 [6] Gary K. Chen and Yunfei Guo, *Discovering epistasis in large scale genetic as-*
1091 *sociation studies by exploiting graphics cards*, Frontiers in Genetics **4** (2013),
1092 no. DEC, 1–12.
- 1093 [7] Joshua J Choi, Xiaohao Yang, Zachariah M Norman, Simon J L Billinge, and
1094 Jonathan S Owen, *Structure of methylammonium lead iodide within mesoporous*
1095 *titanium dioxide: Active material in high-performance perovskite solar cells*,
1096 Nano Letters **14** (2014), no. 1, 127–133.
- 1097 [8] Peter J. Chupas, Karena W. Chapman, Charles Kurtz, Jonathan C. Hanson,
1098 Peter L. Lee, and Clare P. Grey, *A versatile sample-environment cell for non-*
1099 *ambient X-ray scattering experiments*, Journal of Applied Crystallography **41**
1100 (2008), no. 4, 822–824.
- 1101 [9] PJ Chupas, X Qiu, JC Hanson, PL Lee, CP Grey, and SJL Billinge, *Rapid-*

- 1102 *acquisition pair distribution function (RA-PDF) analysis*, Journal of Applied
1103 Crystallography **36** (2003), 1342–1347.
- 1104 [10] Matthew J. Cliffe, Martin T. Dove, D. A. Drabold, and Andrew L. Goodwin,
1105 *Structure determination of disordered materials from diffraction data*, Physical
1106 Review Letters **104** (2010), no. 12, 1–4.
- 1107 [11] Matthew J Cliffe and Andrew L Goodwin, *Nanostructure determination from the*
1108 *pair distribution function: a parametric study of the INVERT approach.*, Journal
1109 of physics. Condensed matter : an Institute of Physics journal **25** (2013), no. 45,
1110 454218.
- 1111 [12] Chunhua Cui, Lin Gan, Marc Heggen, Stefan Rudi, and Peter Strasser, *Composi-*
1112 *tional segregation in shaped Pt alloy nanoparticles and their structural behaviour*
1113 *during electrocatalysis.*, Nature materials **12** (2013), no. 12, 765–771.
- 1114 [13] Juarez L F Da Silva, Hyoung Gyu Kim, Maurício J. Piotrowski, Maurício J. Pri-
1115 eto, and Germano Tremiliosi-Filho, *Reconstruction of core and surface nanopar-*
1116 *ticles: The example of Pt55 and Au55*, Physical Review B - Condensed Matter
1117 and Materials Physics **82** (2010), no. 20, 1–6.
- 1118 [14] Emir Dogdibegovic, Christopher J. Wright, and Xiao-Dong Zhou, *Stability and*
1119 *Activity of $(Pr_{1-x}Nd_x)_2NiO_4$ as Cathodes for Solid Oxide Fuel Cells: I. Quantifi-*
1120 *cation of Phase Evolution in Pr_2NiO_4* , Journal of the American Ceramic Society
1121 **5** (2016), 1–5.
- 1122 [15] Simon Duane, A. D. Kennedy, Brian J. Pendleton, and Duncan Roweth, *Hybrid*
1123 *Monte Carlo*, Physics Letters B **195** (1987), no. 2, 216–22.
- 1124 [16] Timur Dykhne, Ryan Taylor, Alastair Florence, and Simon J L Billinge, *Data*
1125 *requirements for the reliable use of atomic pair distribution functions in amor-*
1126 *phous pharmaceutical fingerprinting.*, Pharmaceutical research **28** (2011), no. 5,
1127 1041–8.
- 1128 [17] C L Farrow, P Juhas, J W Liu, D Bryndin, E S Božin, J Bloch, Th Proffen, and
1129 S J L Billinge, *PDFfit2 and PDFgui: computer programs for studying nanostruc-*
1130 *ture in crystals.*, Journal of Physics. Condensed Matter : an Institute of Physics
1131 journal **19** (2007), no. 33, 335219.
- 1132 [18] Christopher L Farrow and Simon J L Billinge, *Relationship between the atomic*
1133 *pair distribution function and small-angle scattering: implications for modeling*

- 1134 *of nanoparticles.*, Acta Crystallographica Section A Foundations of Crystallography **65** (2009), no. Pt 3, 232–9 (en).
- 1136 [19] Riccardo Ferrando, Julius Jellinek, and Roy L Johnston, *Nanoalloys: From Theory to Applications of Alloy Clusters and Nanoparticles*, Chemical Reviews **108** (2008), no. 3, 846–904.
- 1139 [20] Anatoly Frenkel, *Solving the 3D structure of metal nanoparticles*, Zeitschrift für Kristallographie **222** (2007), no. 11, 605–611.
- 1141 [21] Carmelo Giacovazzo, Hugo Luis Monaco, Gilberto Artioli, Davide Viterbo, Marco Milanesio, Gastone Gilli, Paola Gilli, Giuseppe Zanotti, Giovanni Ferraris, and Michele Catti, *Fundamentals of Crystallography*, Oxford University Press, 1992.
- 1145 [22] Benjamin Gilbert, Jasmine J. Erbs, R. Lee Penn, Valeri Petkov, Dino Spagnoli, and Glenn A. Waychunas, *A disordered nanoparticle model for 6-line ferrihydrite*, American Mineralogist **98** (2013), no. 8-9, 1465–1476.
- 1148 [23] Benjamin Gilbert, Feng Huang, Hengzhong Zhang, Glenn A. Waychunas, and Jillian F Banfield, *Nanoparticles: strained and stiff.*, Science (New York, N.Y.) **305** (2004), no. 5684, 651–654.
- 1151 [24] Md Matthew D. Hoffman and Andrew Gelman, *The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo*, The Journal of Machine Learning Research **15** (2014), no. 2008, 1593–1623.
- 1154 [25] W J Huang, R Sun, J Tao, L D Menard, R G Nuzzo, and J M Zuo, *Coordination-dependent surface atomic contraction in nanocrystals revealed by coherent diffraction.*, Nature Materials **7** (2008), no. 4, 308–313.
- 1157 [26] Pablo D Jadzinsky, Guillermo Calero, Christopher J Ackerson, David A Bushnell, and Roger D Kornberg, *Structure of a thiol monolayer-protected gold nanoparticle at 1.1 Å resolution.*, Science (New York, N.Y.) **318** (2007), no. 5849, 430–433.
- 1160 [27] I. K. Jeong, R. H. Heffner, M. J. Graf, and S. J. L. Billinge, *Lattice dynamics and correlated atomic motion from the atomic pair distribution function*, (2002), 9.
- 1163 [28] P. Juhás, T. Davis, C. L. Farrow, and S. J. L. Billinge, *PDFgetX3 : a rapid and highly automatable program for processing powder diffraction data into total*

- 1165 *scattering pair distribution functions*, Journal of Applied Crystallography **46**
1166 (2013), no. 2, 560–566 (en).
- 1167 [29] A. Kassiba, M. Makowska-Janusik, J. Bouclé, J. Bardeau, A. Bulou, and
1168 N. Herlin-Boime, *Photoluminescence features on the Raman spectra of quasistoi-*
1169 *chiometric SiC nanoparticles: Experimental and numerical simulations*, Physical
1170 Review B **66** (2002), no. 15, 1–7.
- 1171 [30] Jérôme Kieffer and Dimitrios Karkoulis, *PyFAI, a versatile library for azimuthal*
1172 *regrouping*, Journal of Physics: Conference Series **425** (2013), 202012.
- 1173 [31] G. Kresse and OJ. Hafner, *Ab initio molecular-dynamics simulation of the liquid-*
1174 *metal-amorphous-semiconductor transition in germanium*, Physical Review B **49**
1175 (1994), no. 20, 14251–14269.
- 1176 [32] G. Kresse and J. Hafner, *Ab initio molecular dynamics for liquid metals*, Physical
1177 Review B **47** (1993), no. 1, 558–561.
- 1178 [33] Yan Li, Giulia Galli, and François Gygi, *Electronic structure of thiolate-covered*
1179 *gold nanoparticles: Au102(MBA)44*, ACS Nano **2** (2008), no. 9, 1896–1902.
- 1180 [34] L D Marks, *Experimental studies of small particle structures*, Reports on
1181 Progress in Physics **57** (1994), no. 6, 603–649 (en).
- 1182 [35] A. S. Masadeh, E. S. Božin, C. L. Farrow, G. Paglia, P. Juhas, S. J. L. Billinge,
1183 A. Karkamkar, and M. G. Kanatzidis, *Quantitative size-dependent structure and*
1184 *strain determination of CdSe nanoparticles using atomic pair distribution func-*
1185 *tion analysis*, Physical Review B - Condensed Matter and Materials Physics **76**
1186 (2007), no. 11, 115413.
- 1187 [36] R L McGreevy and L Pusztai, *Reverse Monte Carlo Simulation: A New Tech-*
1188 *nique for the Determination of Disordered Structures*, Molecular Simulation **1**
1189 (1988), no. 6, 359–367.
- 1190 [37] Donald A. McQuarrie, *Statistical Mechanics*, University Science Books, Sausalito,
1191 CA, 2000.
- 1192 [38] Radford M Neal, *Probabilistic Inference Using Markov Chain Monte Carlo Meth-*
1193 *ods*, Intelligence **45** (1993), no. September, 144.

- 1194 [39] Radford M. Neal, *MCMC Using Hamiltonian Dynamics*, Handbook of Markov
1195 Chain Monte Carlo (Steve Brooks, Andrew Gelman, Galin L. Jones and Xiao-Li
1196 Meng, eds.), Chapman and Hall/CRC, 2011, pp. 113–162.
- 1197 [40] Katharine Page, Taylor C. Hood, Thomas Proffen, and Reinhard B. Neder,
1198 *Building and refining complete nanoparticle structures with total scattering data*,
1199 Journal of Applied Crystallography **44** (2011), no. 2, 327–336 (en).
- 1200 [41] B R Pauw, *Corrigendum: Everything SAXS: small-angle scattering pattern col-*
1201 *lection and correction (2013 J. Phys.: Condens. Matter 25 383201)*, Journal of
1202 Physics: Condensed Matter **26** (2014), no. 23, 239501.
- 1203 [42] Vitalij Pecharsky and Peter Zavalij, *Fundamentals of Powder Diffraction and*
1204 *Structural Characterization of Materials*, Springer Science, 2009.
- 1205 [43] John P. Perdew, Kieron Burke, and Matthias Ernzerhof, *Generalized Gradient*
1206 *Approximation Made Simple*, Physical Review Letters **77** (1996), no. 18, 3865–
1207 3868.
- 1208 [44] Andrew A. Peterson, *Global optimization of adsorbate-surface structures while*
1209 *preserving molecular identity*, Topics in Catalysis **57** (2014), no. 1-4, 40–53.
- 1210 [45] Valeri Petkov, Youngmin Lee, Shouheng Sun, and Yang Ren, *Noncrystallographic*
1211 *atomic arrangement driven enhancement of the catalytic activity of Au nanoparticles*, Journal of Physical Chemistry C **116** (2012), no. 50, 26668–26673.
- 1213 [46] Valeri Petkov, Binay Prasai, Yang Ren, Shiyao Shan, Jin Luo, Pharrah Joseph,
1214 and Chuan-Jian Zhong, *Solving the nanostructure problem: exemplified on metal-*
1215 *alloy nanoparticles*, Nanoscale **6** (2014), no. 17, 1–11.
- 1216 [47] Valeri Petkov, Shiyao Shan, Peter Chupas, Jun Yin, Lefu Yang, Jin Luo, and
1217 Chuan-Jian Zhong, *Noble-transition metal nanoparticle breathing in a reactive*
1218 *gas atmosphere.*, Nanoscale **5** (2013), no. 16, 7379–87.
- 1219 [48] Th. Proffen and R. B. Neder, *DISCUS: a Program for Diffuse Scattering and*
1220 *Defect-Structure Simulation*, Journal of Applied Crystallography **30** (1997), 171–
1221 175.
- 1222 [49] Erin L. Redmond, Brian P. Setzler, Pavol Juhas, Simon J. L. Billinge, and
1223 Thomas F. Fuller, *In-Situ Monitoring of Particle Growth at PEMFC Cathode*
1224 *under Accelerated Cycling Conditions*, Electrochemical and Solid-State Letters
1225 **15** (2012), no. 5, B72 (en).

- 1226 [50] H. W. Sheng, M. J. Kramer, A. Cadien, T. Fujita, and M. W. Chen, *Highly*
1227 *optimized embedded-atom-method potentials for fourteen FCC metals*, Physical
1228 Review B - Condensed Matter and Materials Physics **83** (2011), no. 13, 134118.
- 1229 [51] Chenyang Shi, Majid Beidaghi, Michael Naguib, Olha Mashtalir, Yury Gogotsi,
1230 and Simon J L Billinge, *Structure of Nanocrystalline Ti₃C₂ MXene Using Atomic*
1231 *Pair Distribution Function.*, Physical Review Letters **112** (2014), no. March,
1232 125501.
- 1233 [52] Randall Q. Snurr, Alexis T. Bell, and Doros N. Theodorou, *Prediction of ad-*
1234 *sorption of aromatic hydrocarbons in silicalite from grand canonical Monte Carlo*
1235 *simulations with biased insertions*, The Journal of Physical Chemistry **97** (1993),
1236 no. 51, 13742–13752 (EN).
- 1237 [53] Chang Q. Sun, *Size dependence of nanostructures: Impact of bond order defi-*
1238 *ciency*, Progress in Solid State Chemistry **35** (2007), no. 1, 1–159.
- 1239 [54] X. Yang, P. Juhás, and S. J L Billinge, *On the estimation of statistical uncertain-*
1240 *ties on powder diffraction and small-angle scattering data from two-dimensional*
1241 *X-ray detectors*, Journal of Applied Crystallography **47** (2014), no. 4, 1273–1283.
- 1242 [55] Masatomo Yashima, Makiko Enoki, Takahiro Wakita, Roushown Ali, Yoshitaka
1243 Matsushita, Fujio Izumi, and Tatsumi Ishihara, *Structural Disorder and Dif-*
1244 *fusional Pathway of Oxide Ions in a Doped Pr₂NiO₄ -Based Mixed Conductor*,
1245 Journal of the American Chemical Society **130** (2008), no. 9, 2762–2763.
- 1246 [56] Hengzhong Zhang, Bin Chen, Jillian F. Banfield, and Glenn A. Waychunas,
1247 *Atomic structure of nanometer-sized amorphous TiO₂*, Physical Review B **78**
1248 (2008), no. 21, 214106.
- 1249 [57] X.-D. Zhou, J.W. Templeton, Z. Nie, H. Chen, J.W. Stevenson, and L.R. Pedersen,
1250 *Electrochemical performance and stability of the cathode for solid oxide fuel*
1251 *cells: V. high performance and stable Pr₂NiO₄ as the cathode for solid oxide fuel*
1252 *cells*, Electrochimica Acta **71** (2012), 44–49.

1253

APPENDIX A

1254

SUPPLEMENTAL INFORMATION: PHASE CHANGES AND

1255

ANNEALING DYNAMICS OF Pr_2NiO_4 AND ITS

1256

DERIVATIVES

1257 **Intra Sample Comparison**

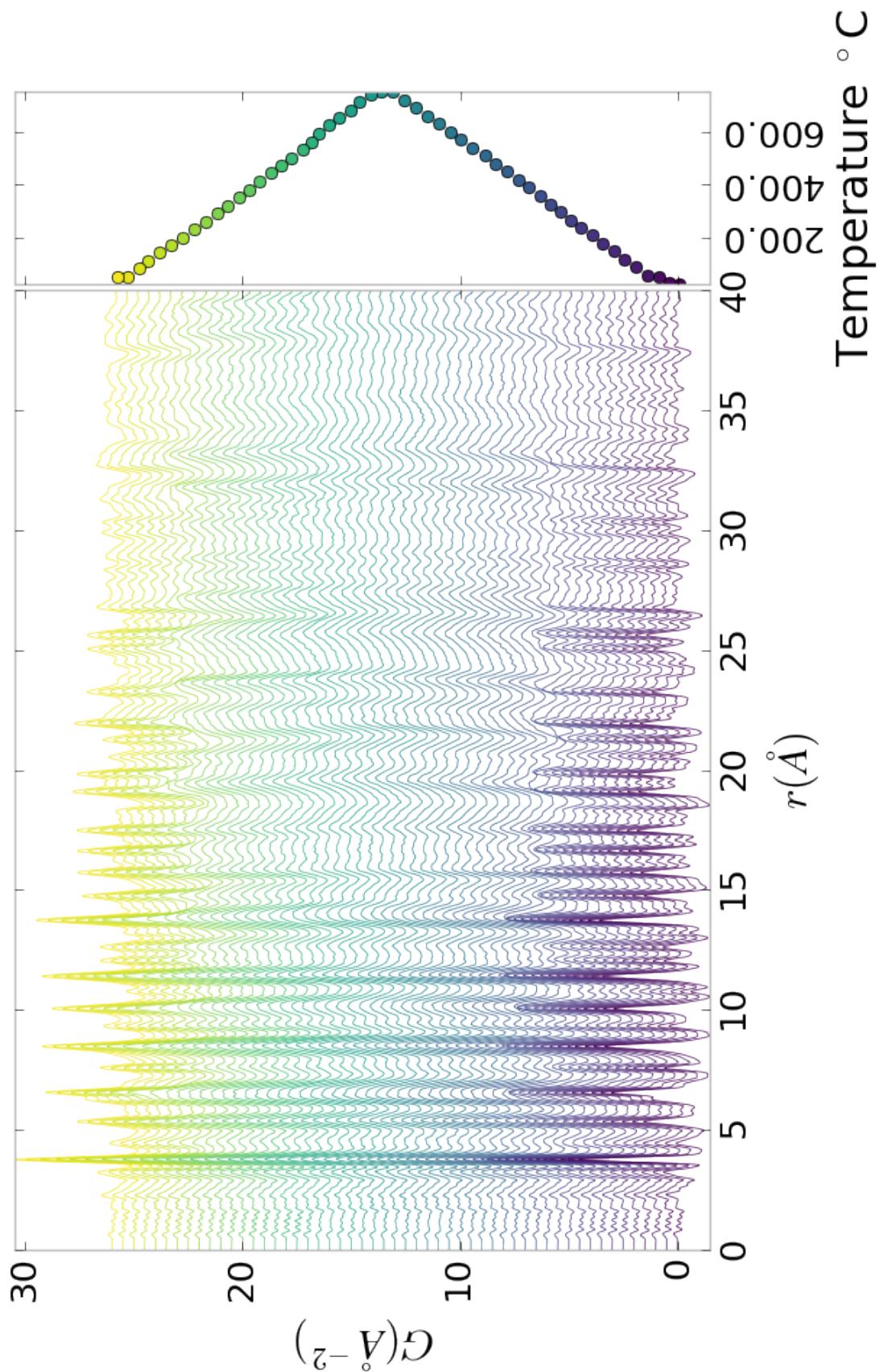


Figure A.1: PDF as a function of temperature for PNO annealed at 750 °C for 50 hours showing the full PDF

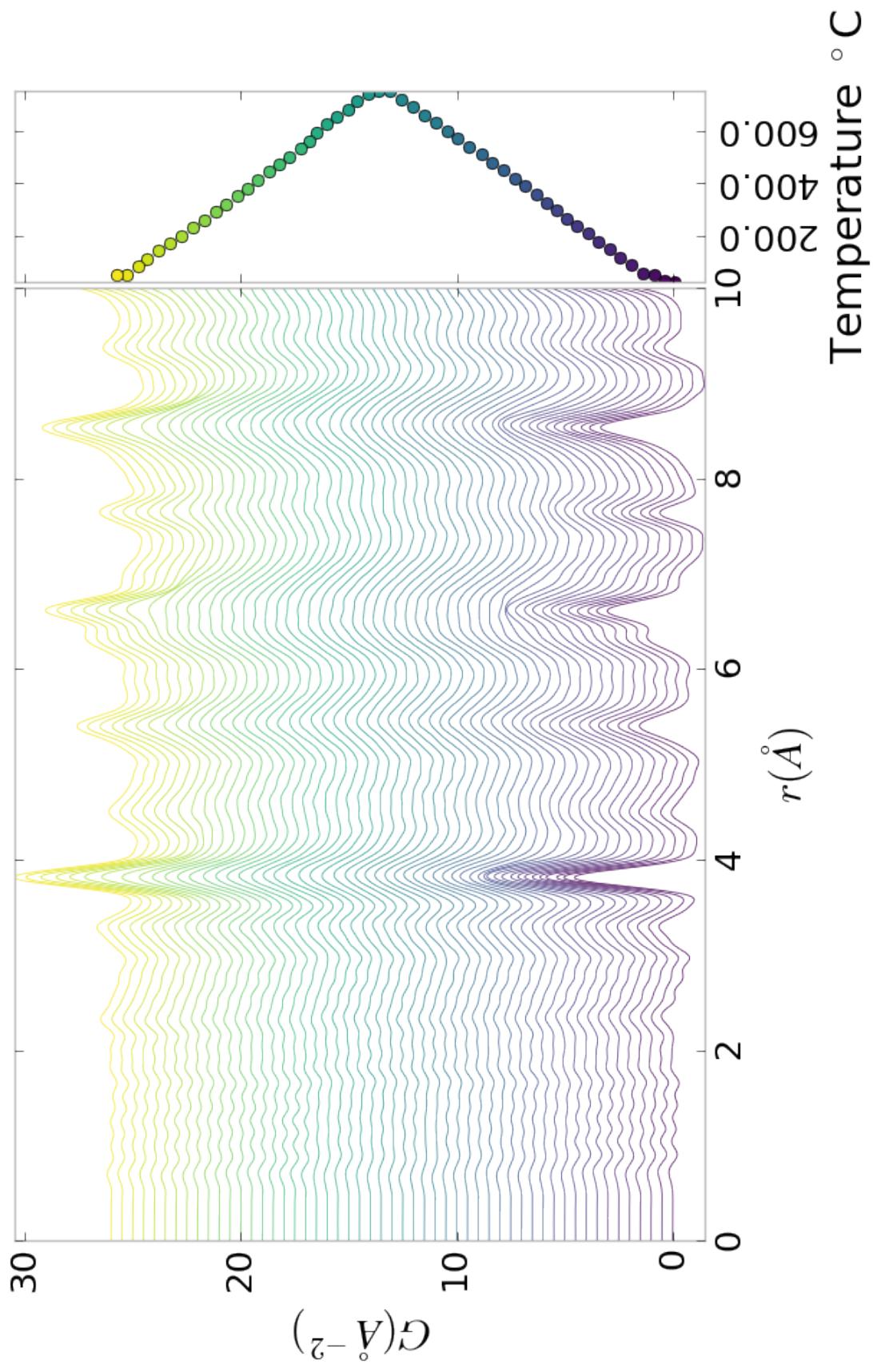


Figure A.2: PDF as a function of temperature for PNO annealed at 750 $^{\circ}\text{C}$ for 50 hours showing a close up on the short range section

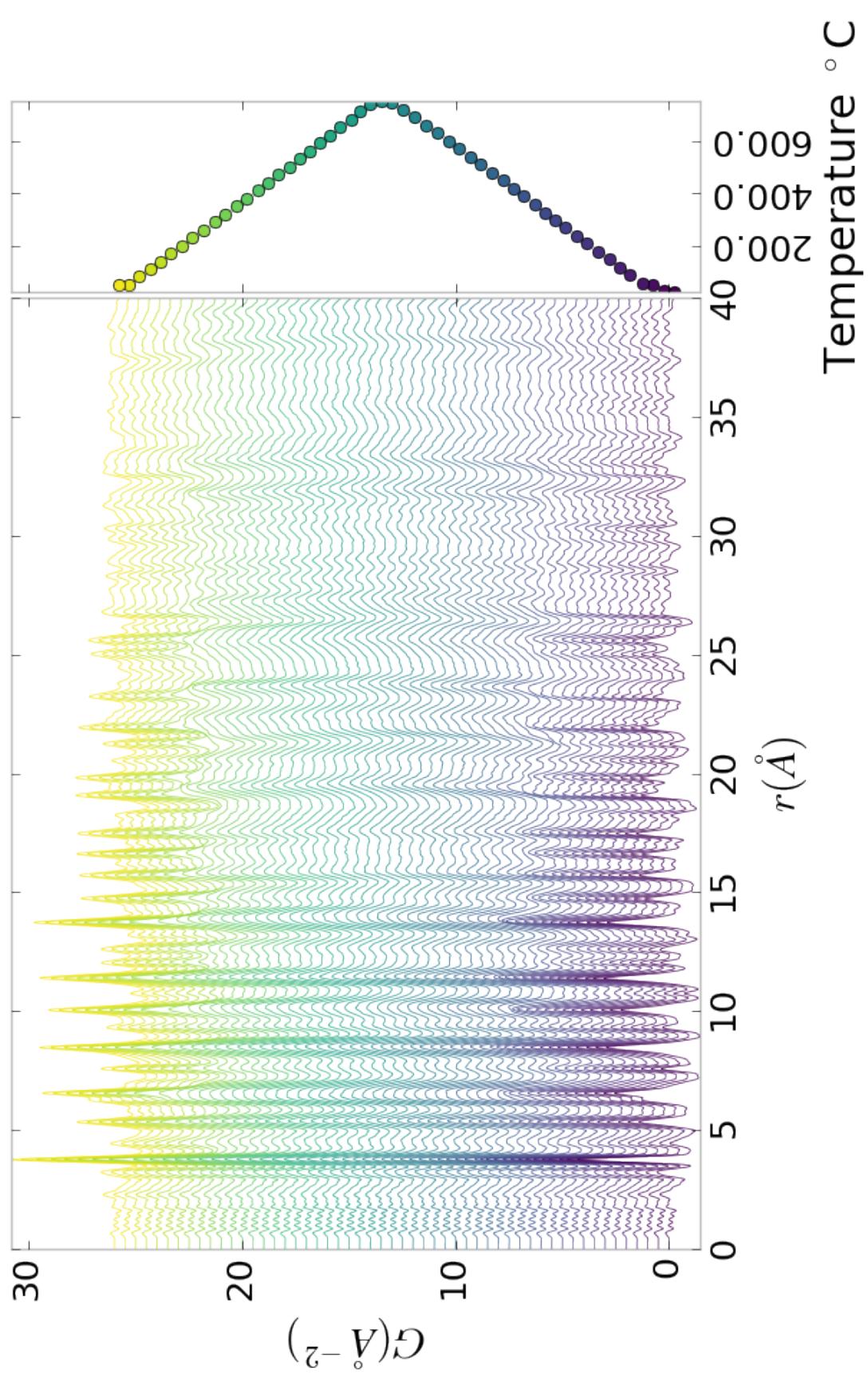


Figure A.3: PDF as a function of temperature for PNO annealed at 750 $^{\circ}\text{C}$ for 100 hours showing the full PDF

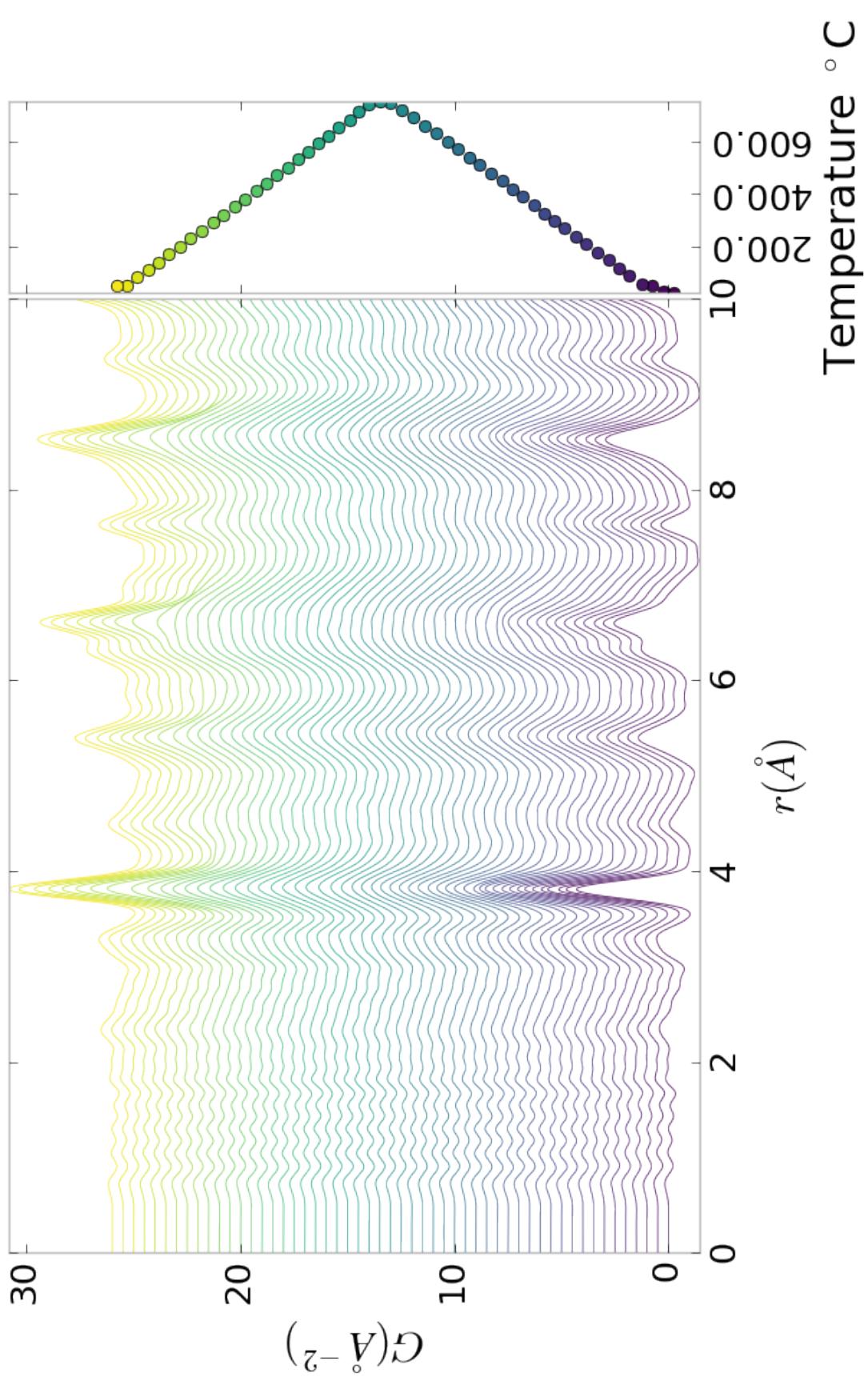


Figure A.4: PDF as a function of temperature for PNO annealed at 750 °C for 100 hours showing a close up on the short range section

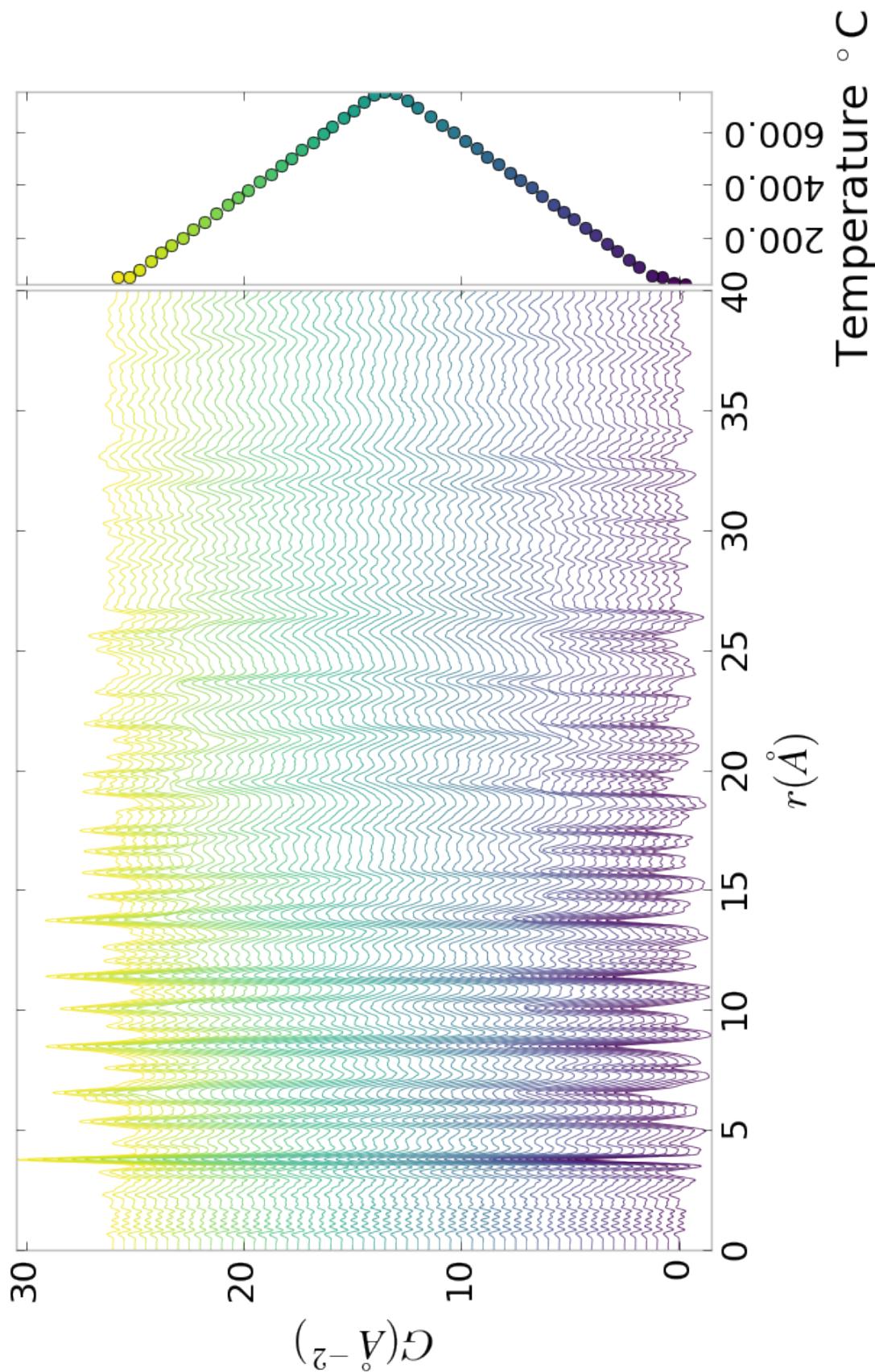


Figure A.5: PDF as a function of temperature for PNO annealed at 750 $^{\circ}\text{C}$ for 200 hours showing the full PDF

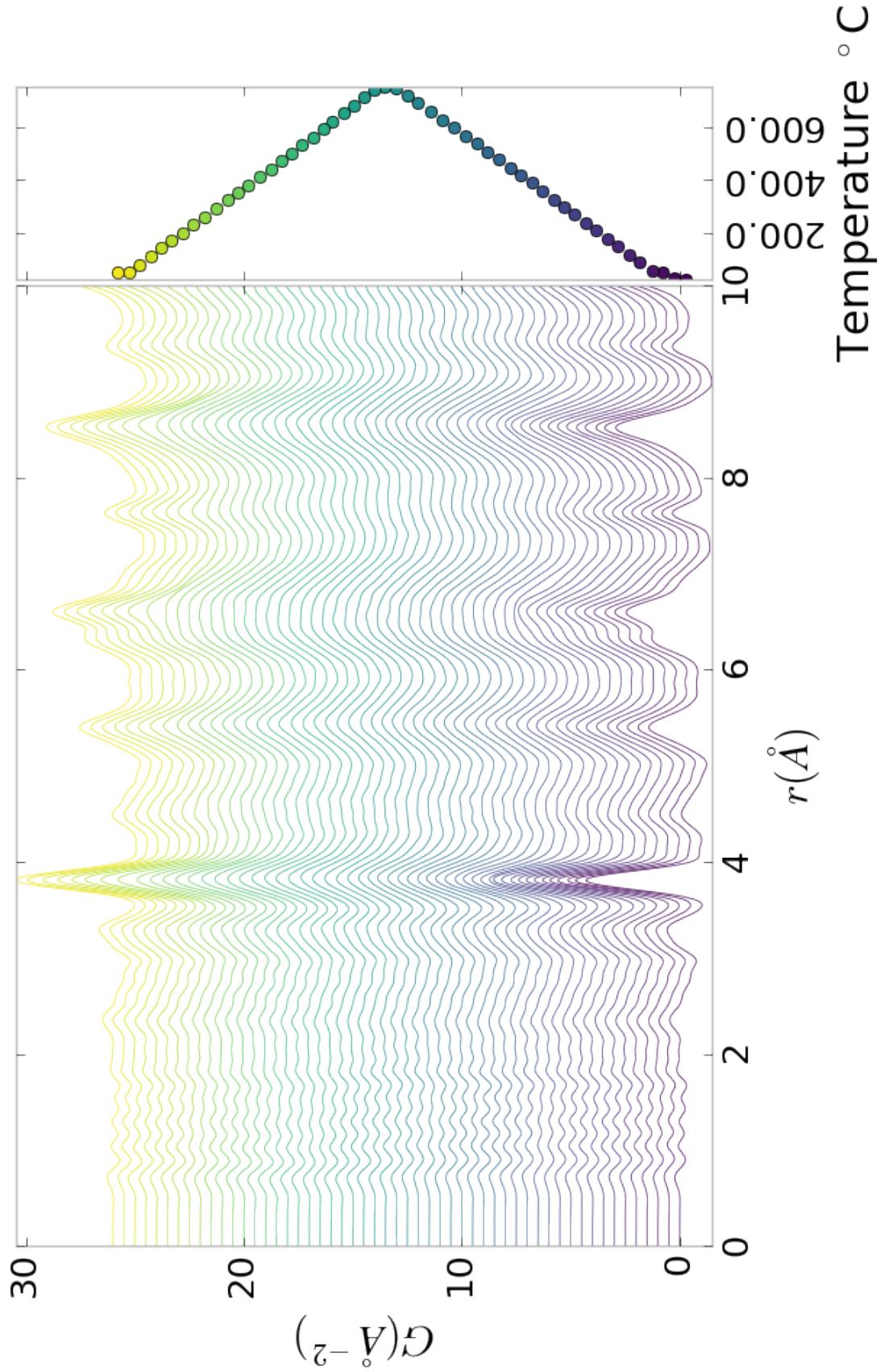


Figure A.6: PDF as a function of temperature for PNO annealed at 750 $^{\circ}\text{C}$ for 200 hours showing a close up on the short range section

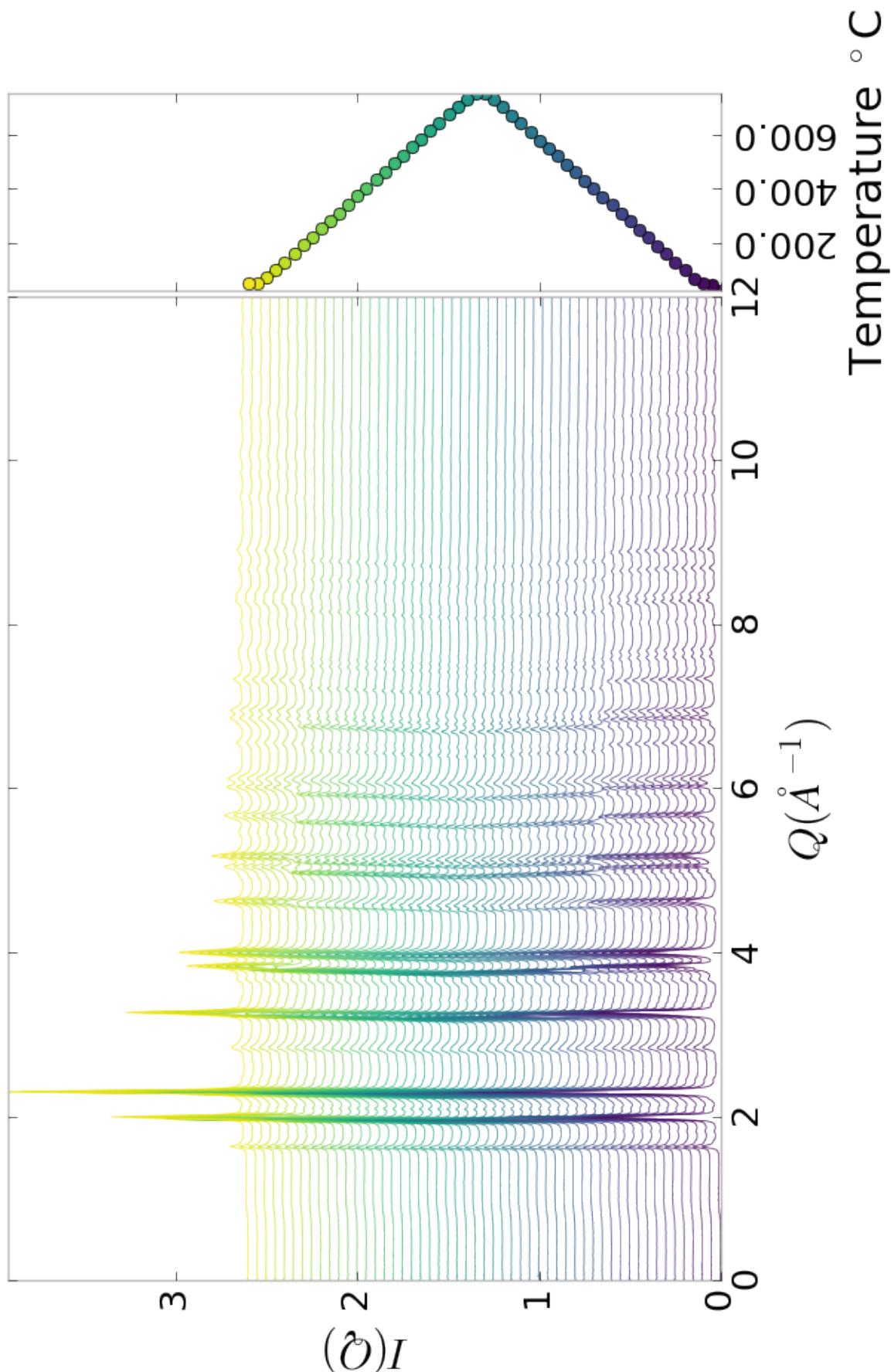


Figure A.7: $I(Q)$ as a function of temperature for PNO annealed at 750 °C for 50 hours showing the full XRD

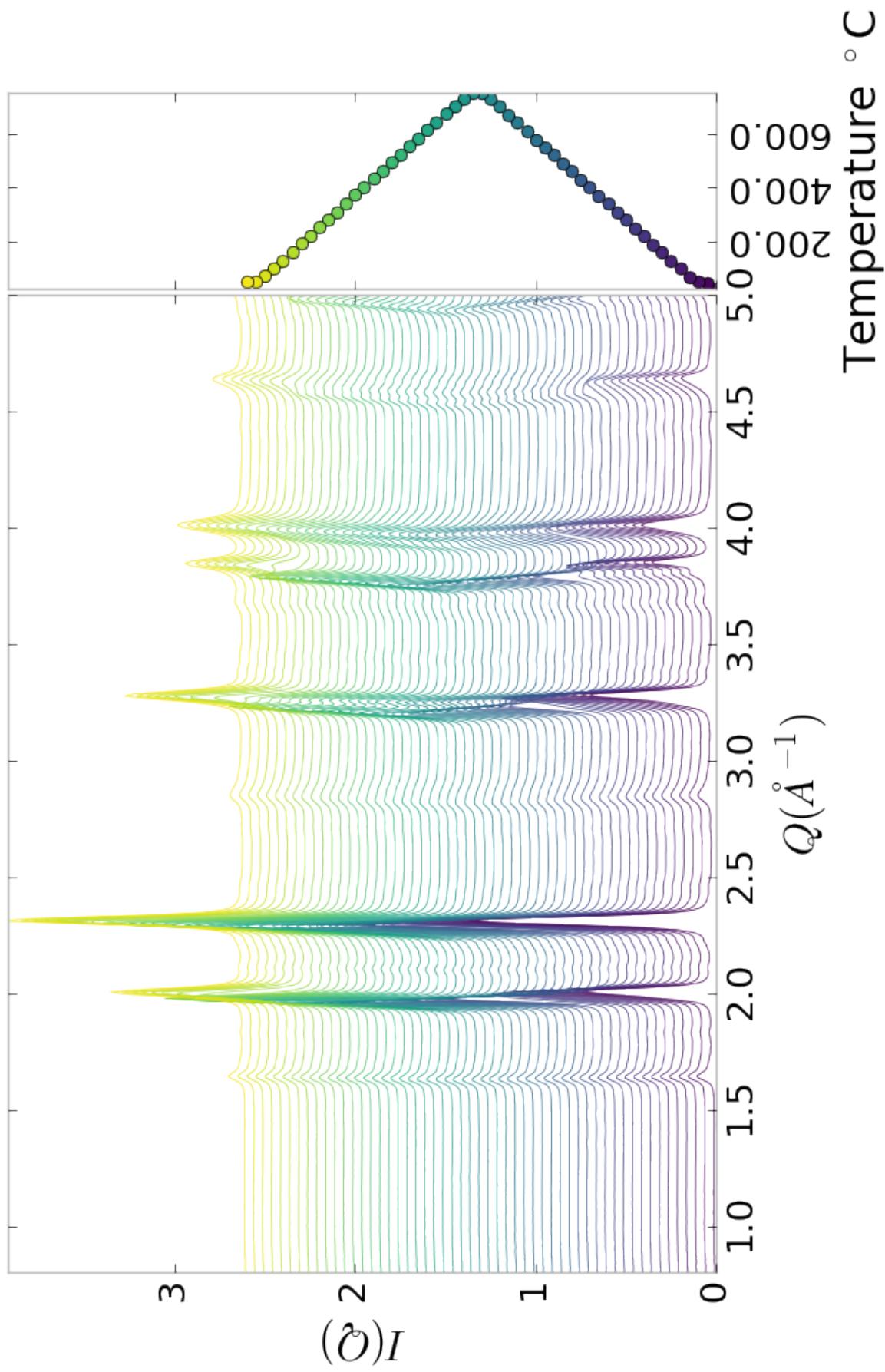


Figure A.8: $I(Q)$ as a function of temperature for PNO annealed at 750 °C for 50 hours showing a close up on the low Q section

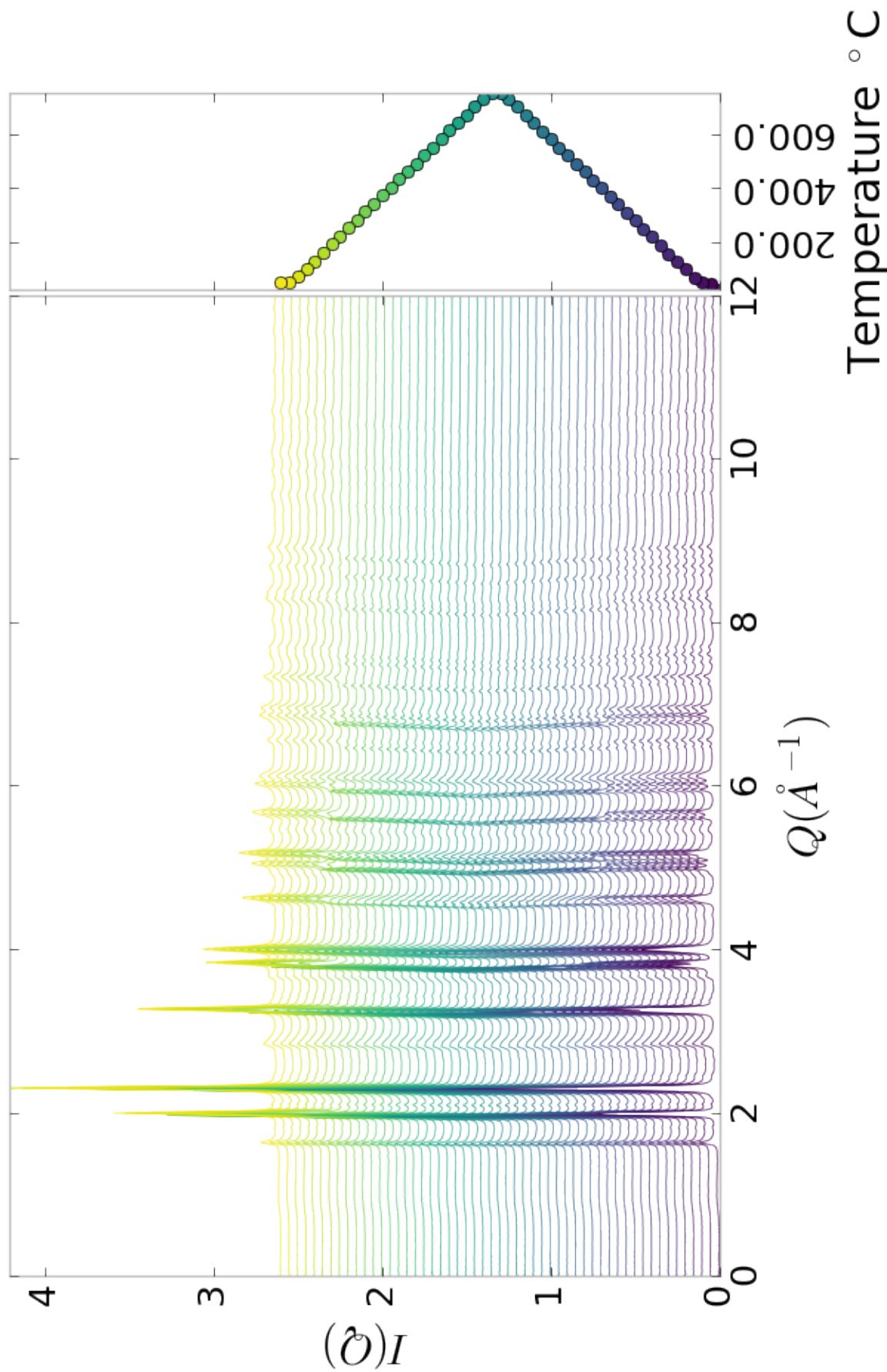


Figure A.9: $I(Q)$ as a function of temperature for PNO annealed at $750\text{ }^{\circ}\text{C}$ for 100 hours showing the full XRD

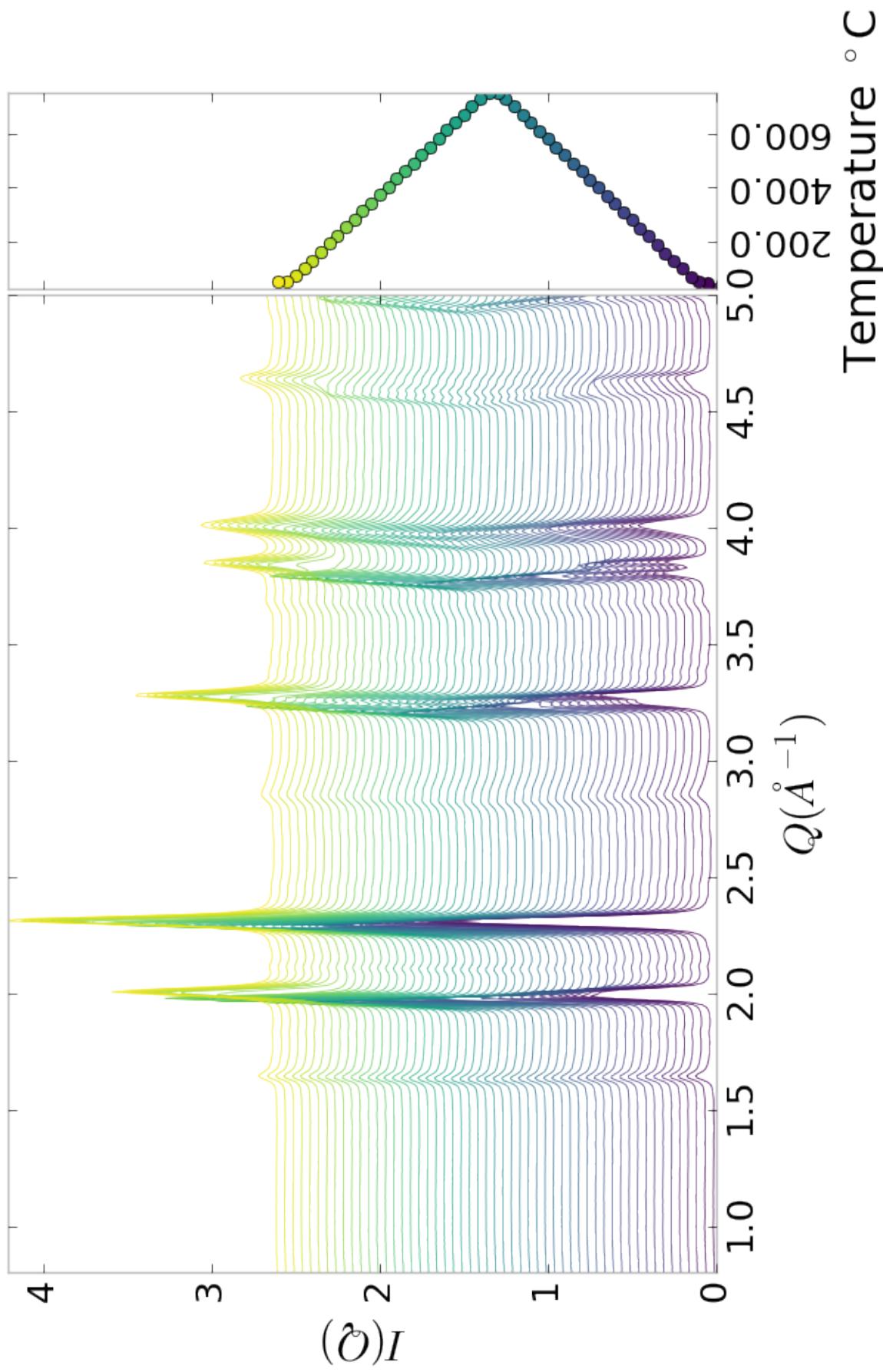


Figure A.10: $I(Q)$ as a function of temperature for PNO annealed at $750\text{ }^{\circ}\text{C}$ for 100 hours showing a close up on the low Q section

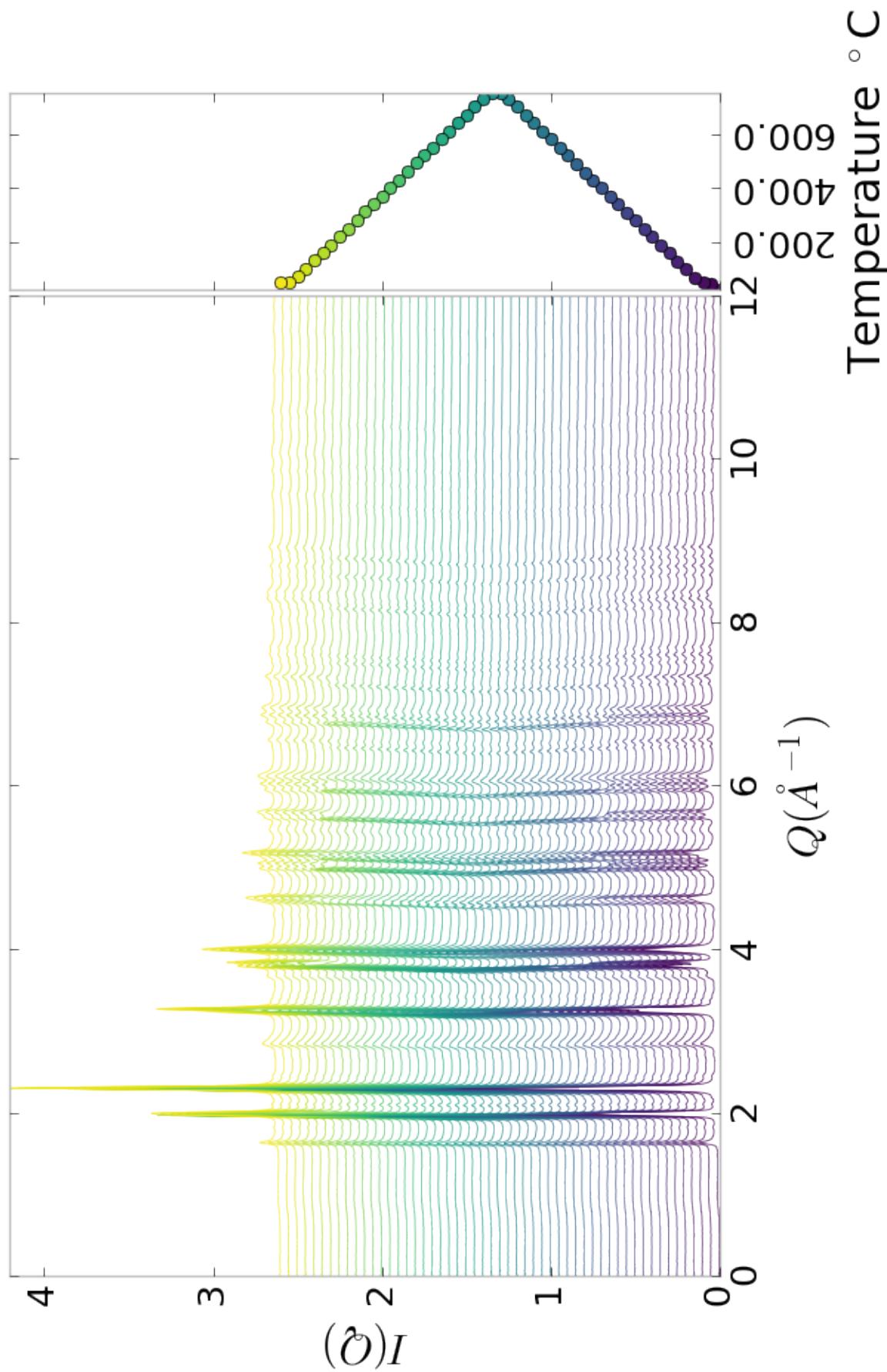


Figure A.11: $I(Q)$ as a function of temperature for PNO annealed at 750 $^{\circ}\text{C}$ for 200 hours showing the full XRD

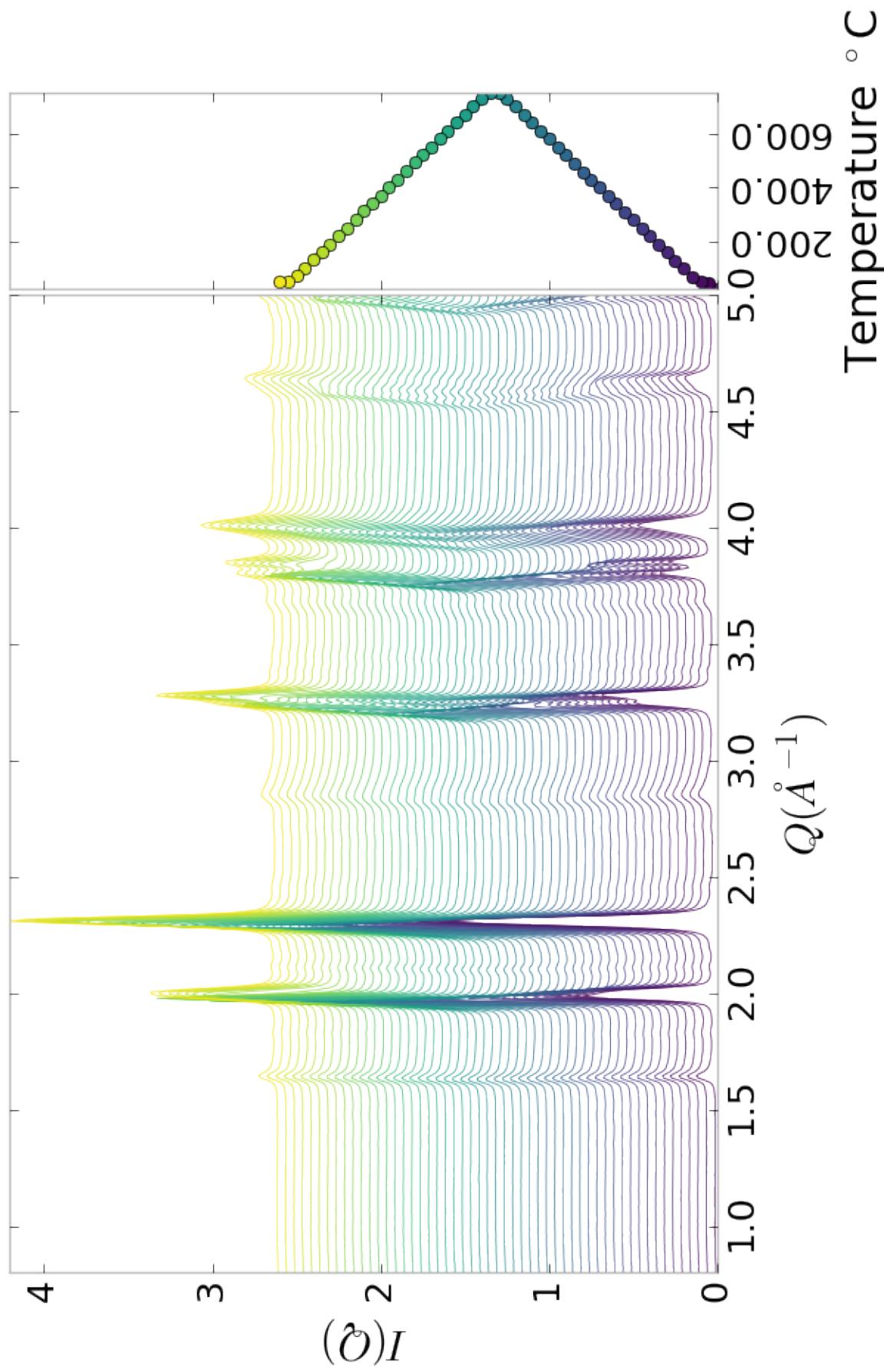


Figure A.12: $I(Q)$ as a function of temperature for PNO annealed at 750 °C for 200 hours showing a close up on the low Q section

1258 **Inter Sample Comparison**

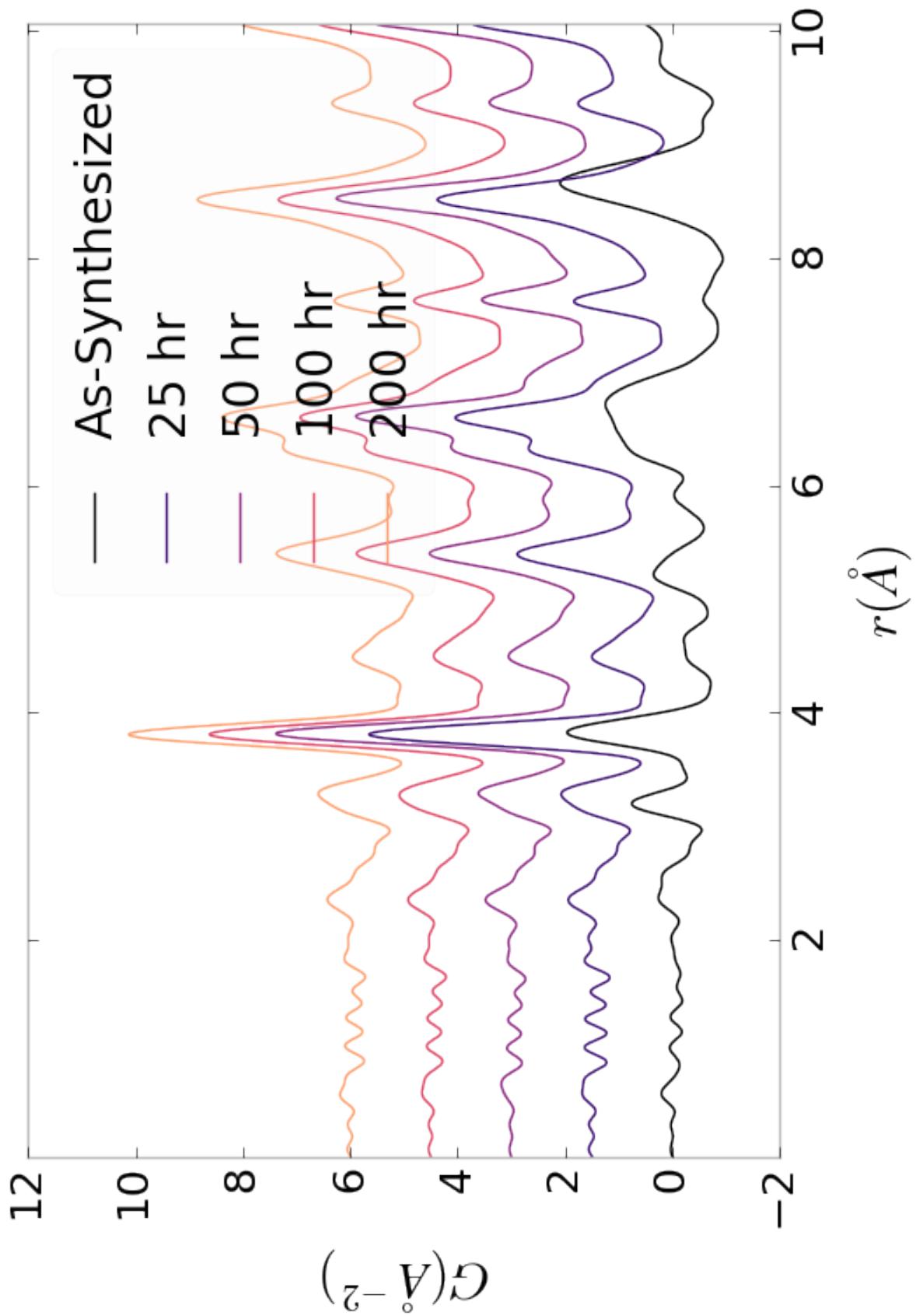


Figure A.13: Comparison of PNO sample PDFs as a function of annealing time at room temperature

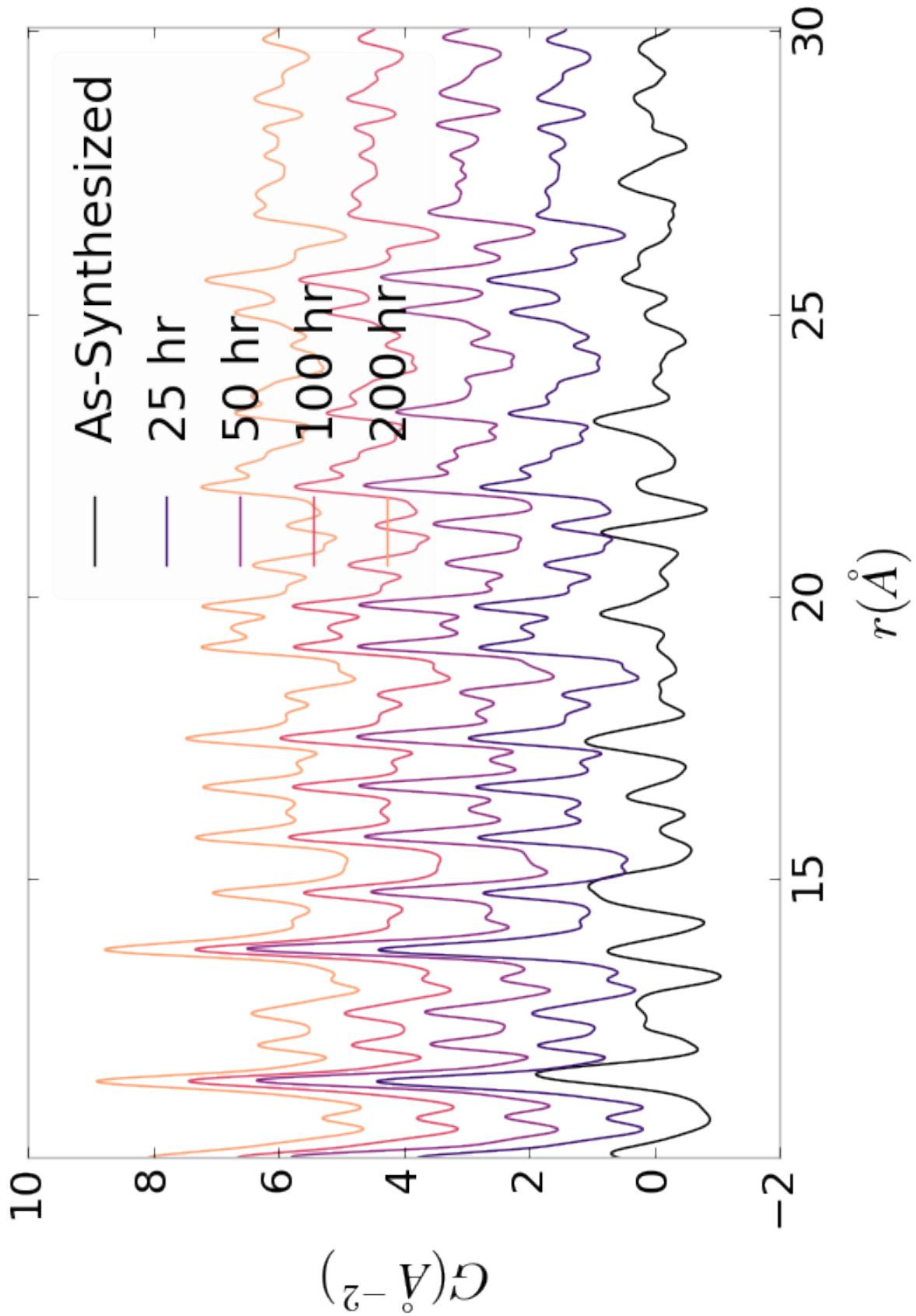


Figure A.14: Comparison of PNO sample PDFs as a function of annealing time at room temperature

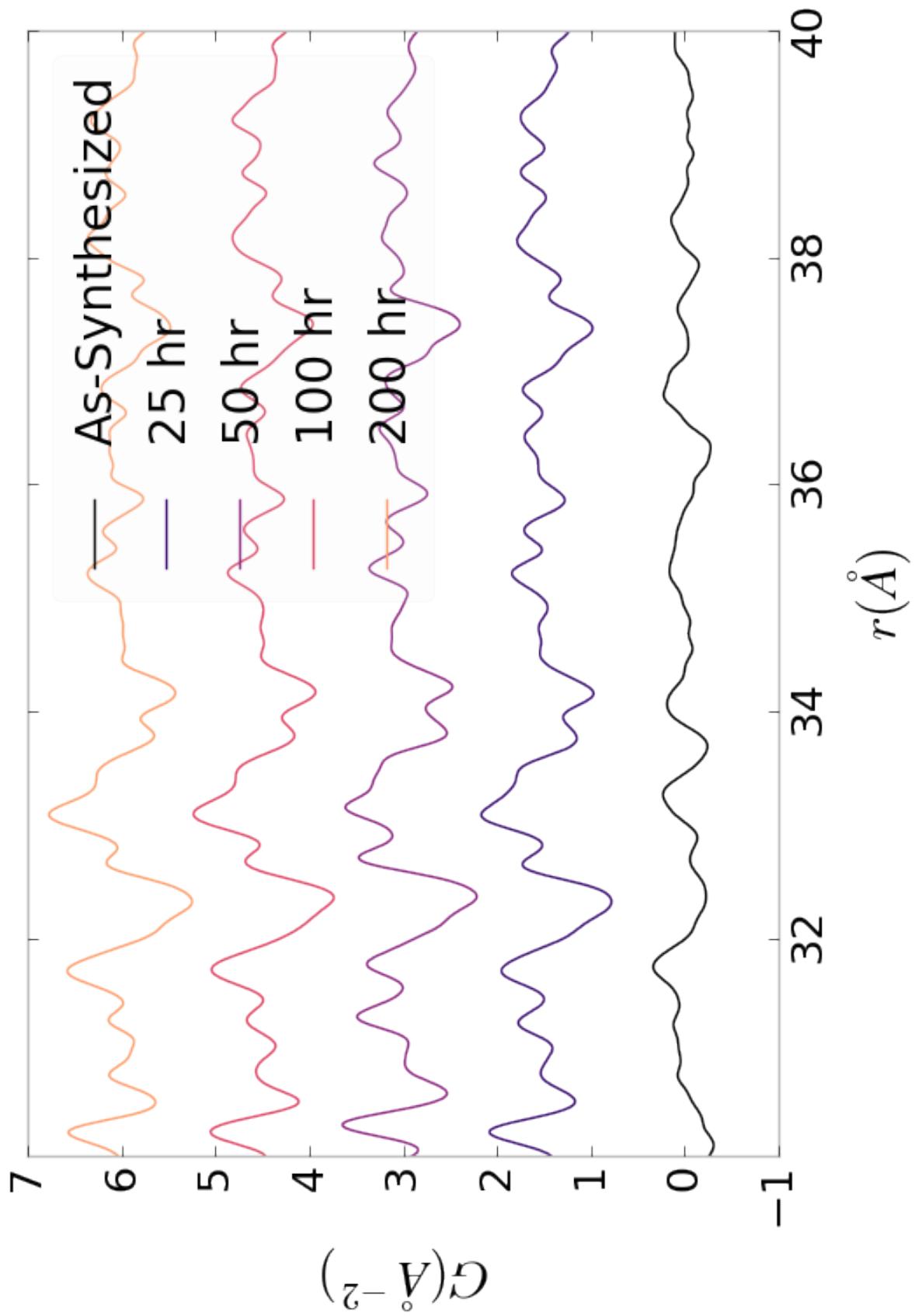


Figure A.15: Comparison of PNO sample PDFs as a function of annealing time at room temperature

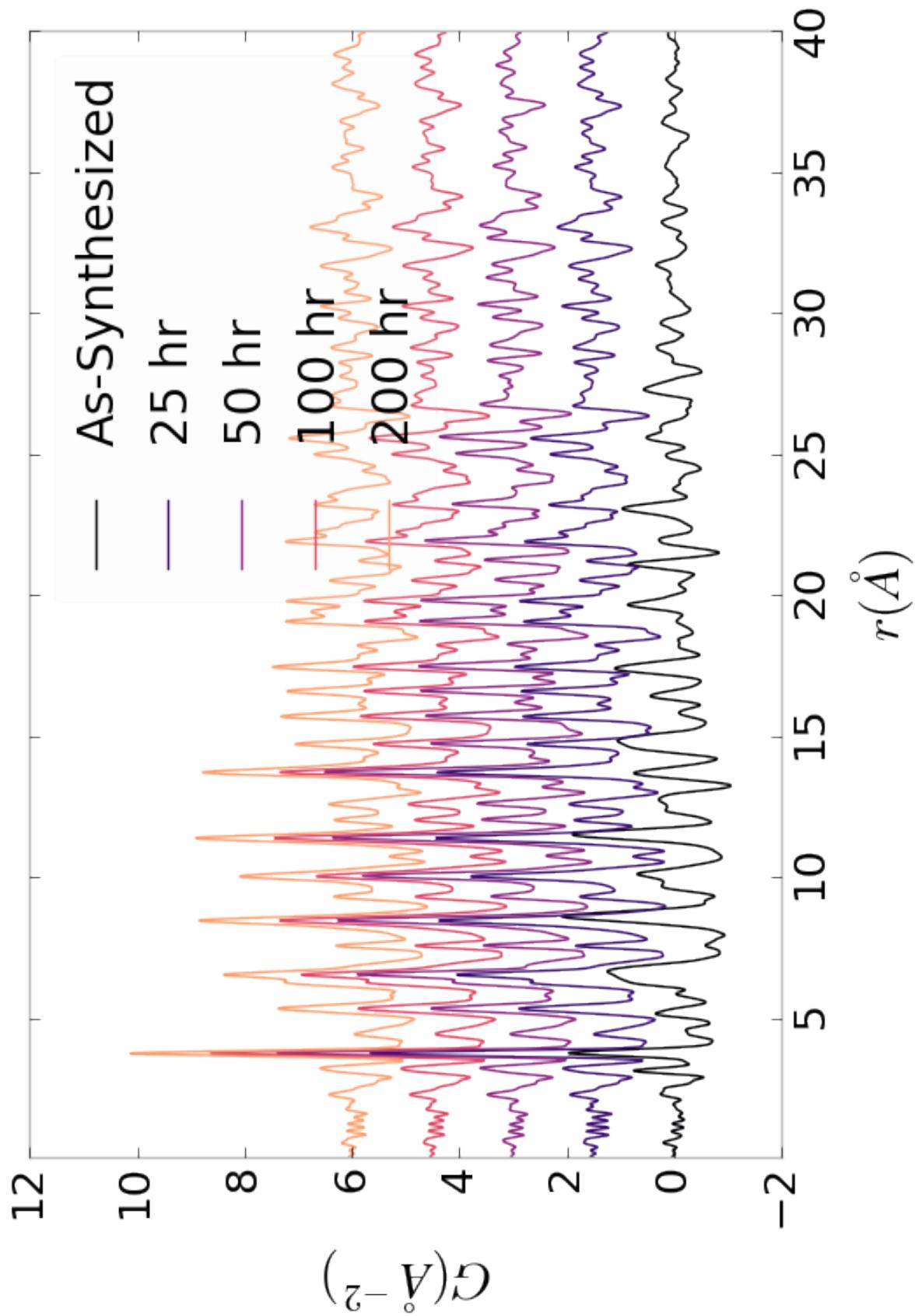


Figure A.16: Comparison of PNO sample PDFs as a function of annealing time at room temperature

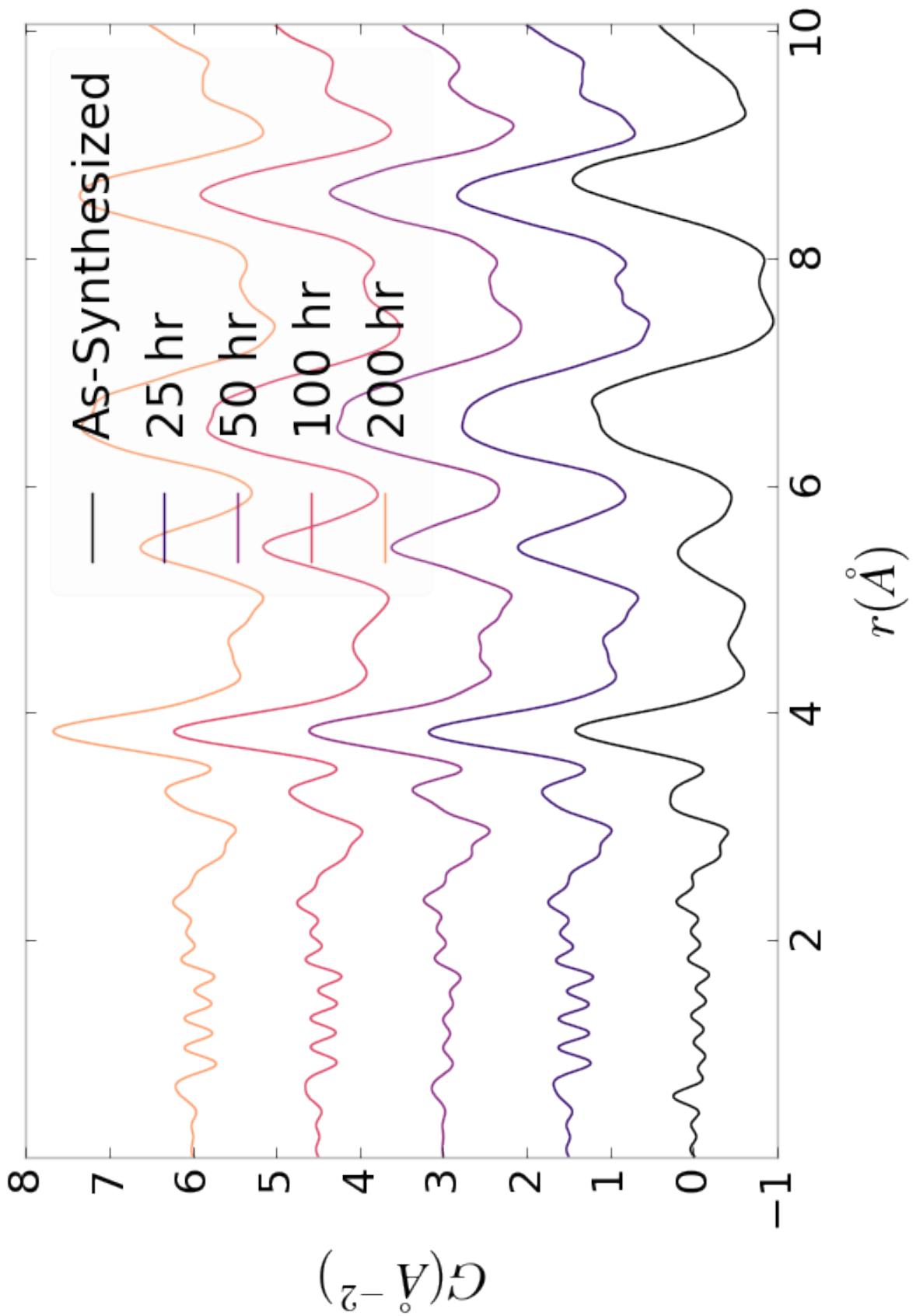


Figure A.17: Comparison of PNO sample PDFs as a function of annealing time at operating temperature

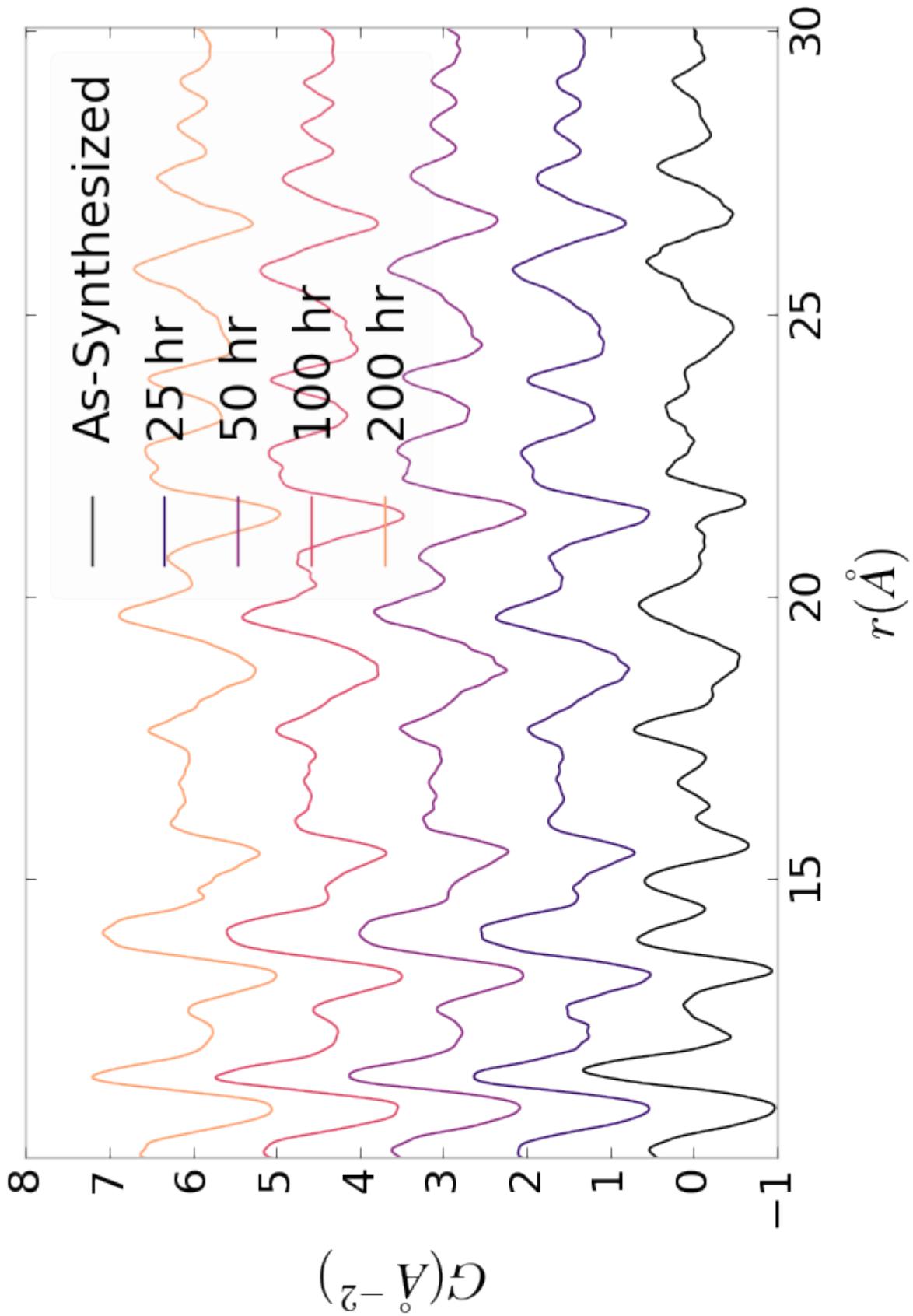


Figure A.18: Comparison of PNO sample PDFs as a function of annealing time at operating temperature

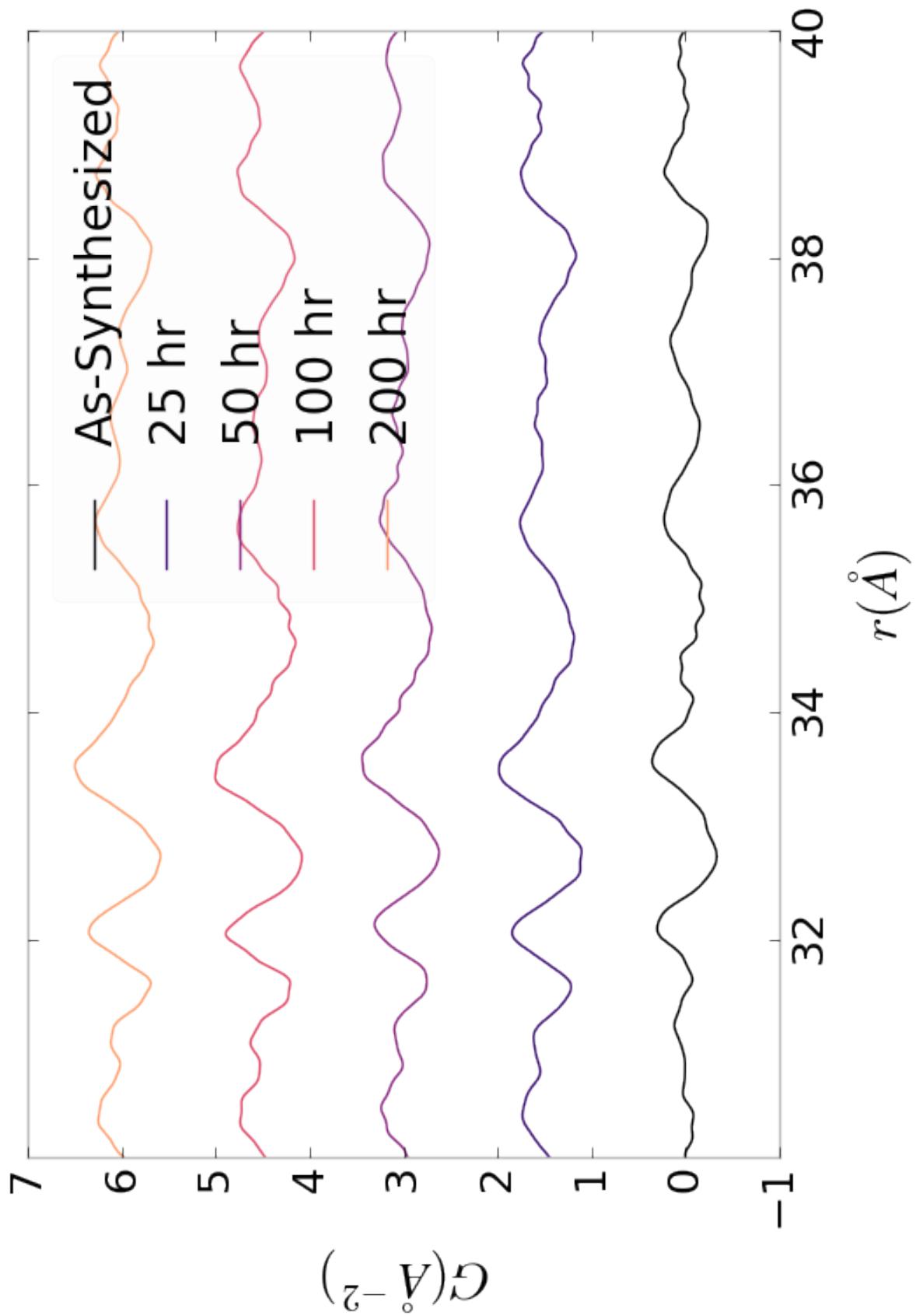


Figure A.19: Comparison of PNO sample PDFs as a function of annealing time at operating temperature

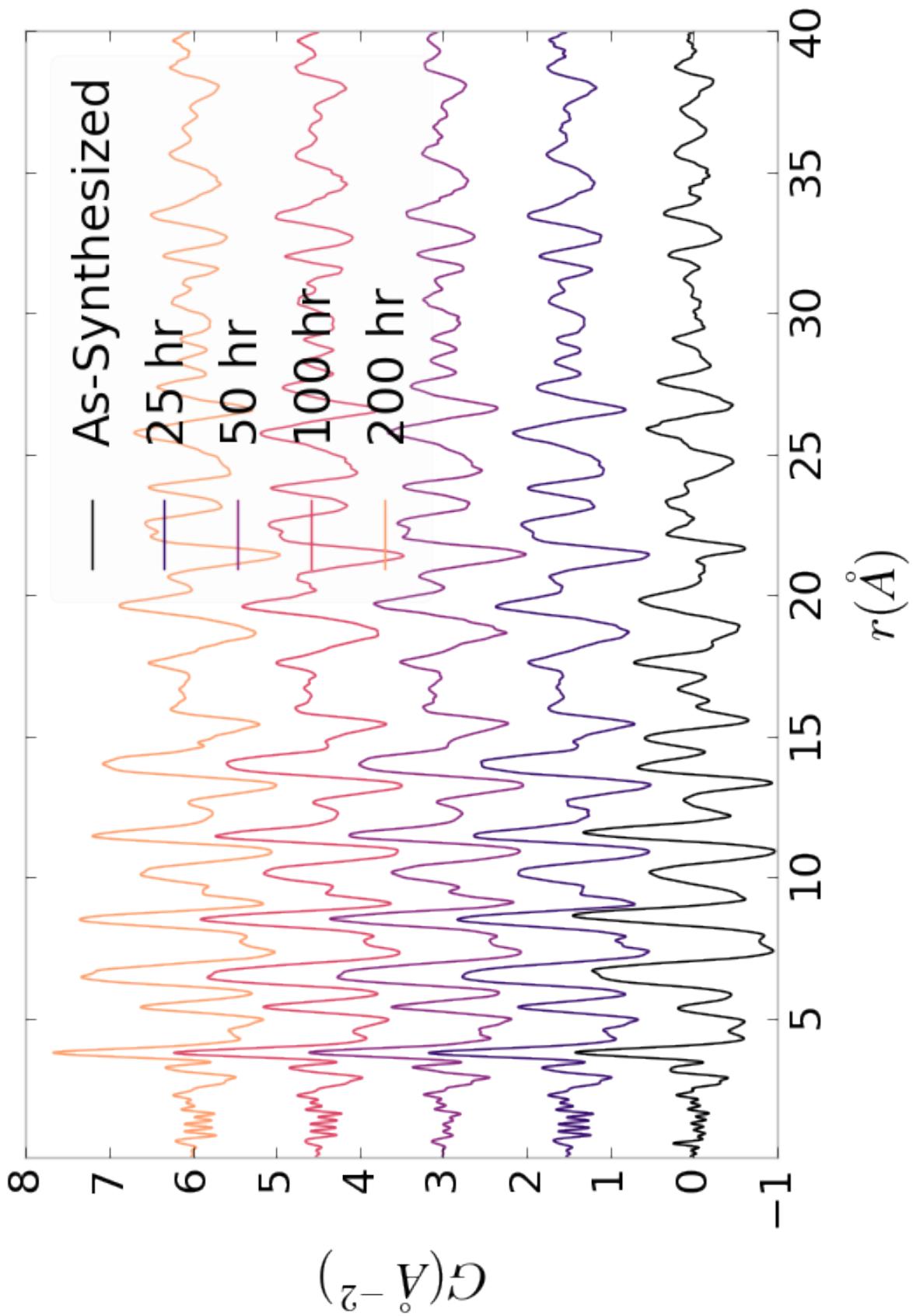


Figure A.20: Comparison of PNO sample PDFs as a function of annealing time at operating temperature

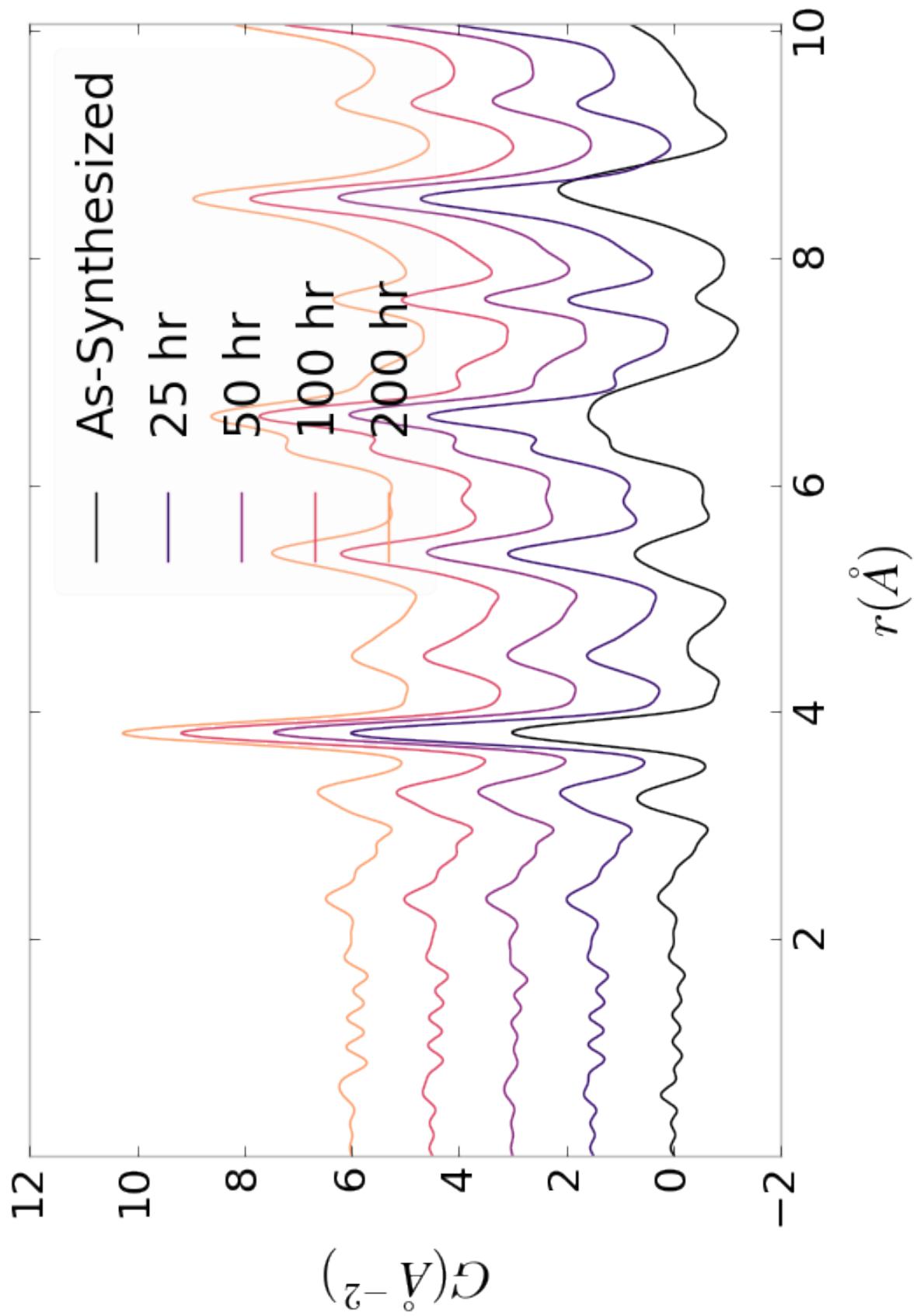


Figure A.21: Comparison of PNO sample PDFs as a function of annealing time cooled back to room temperature

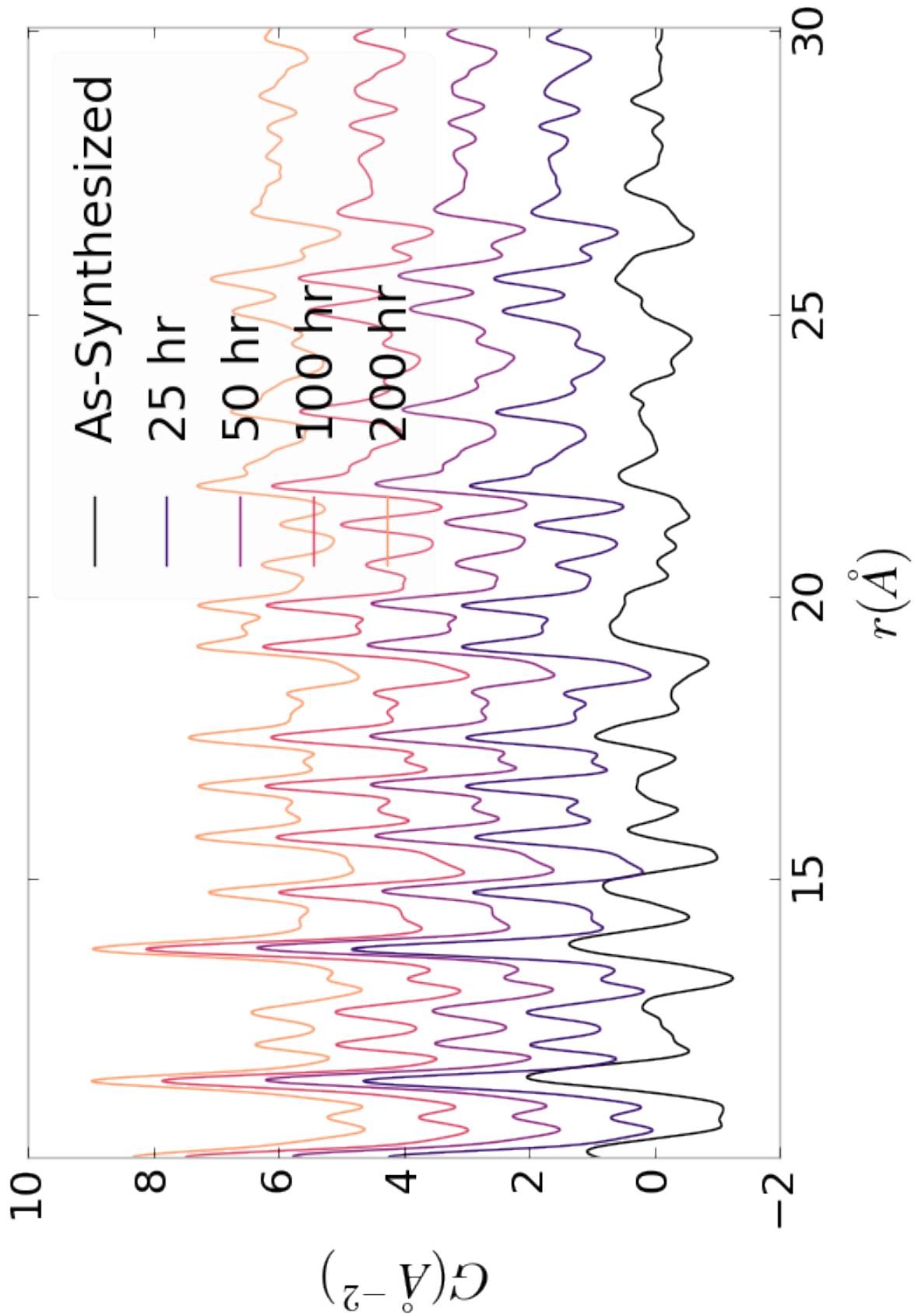


Figure A.22: Comparison of PNO sample PDFs as a function of annealing time cooled back to room temperature

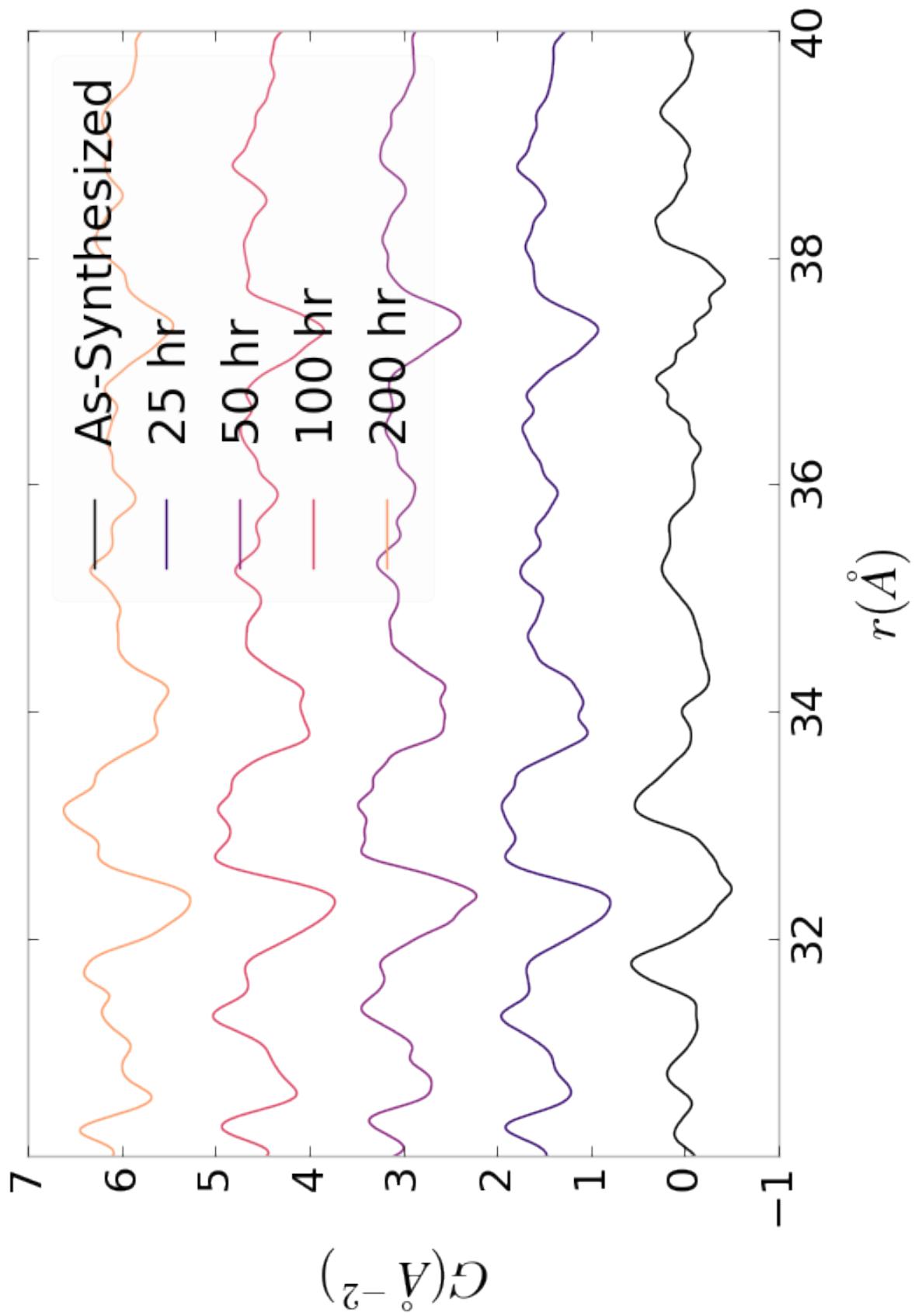


Figure A.23: Comparison of PNO sample PDFs as a function of annealing time cooled back to room temperature

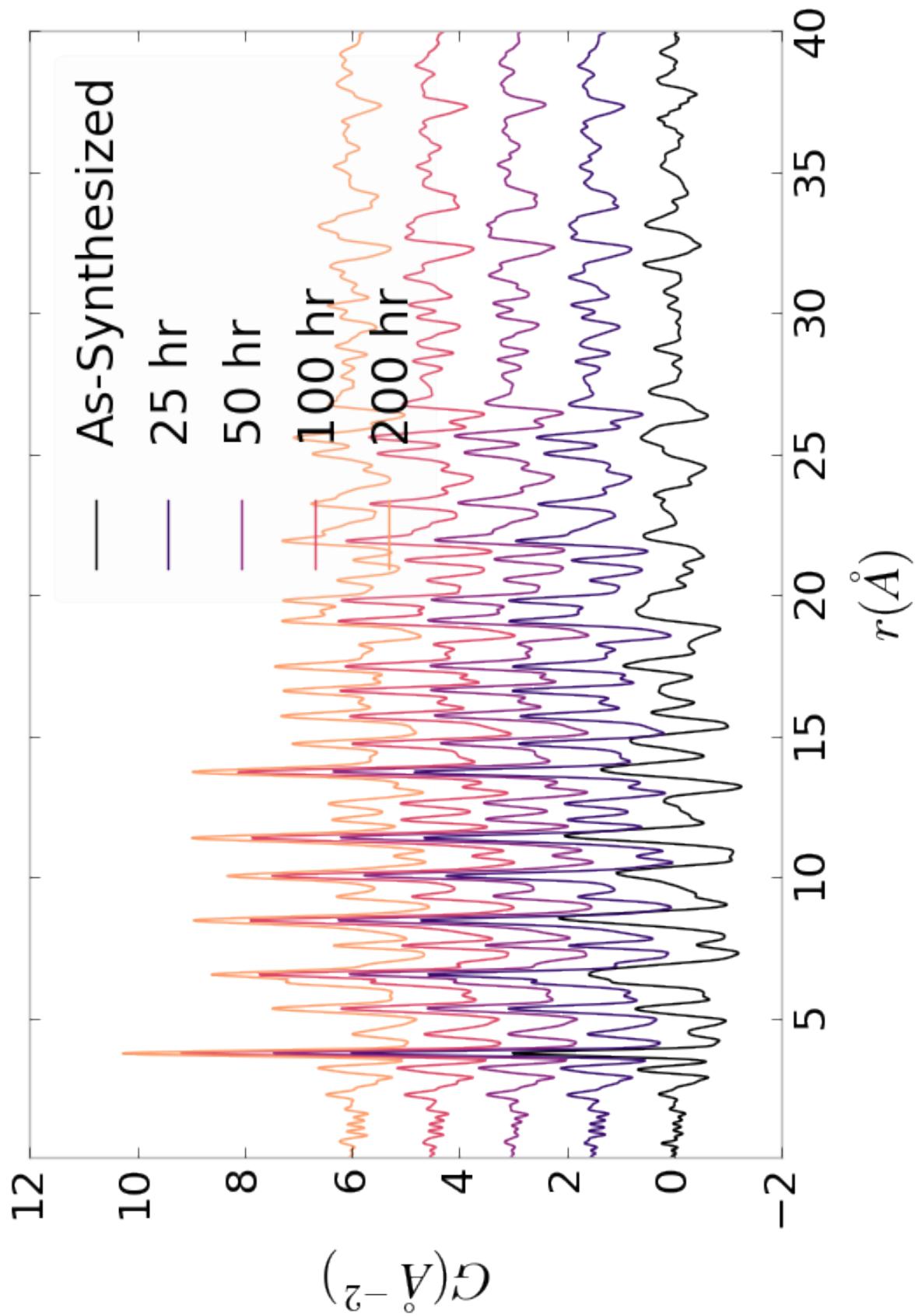


Figure A.24: Comparison of PNO sample PDFs as a function of annealing time cooled back to room temperature

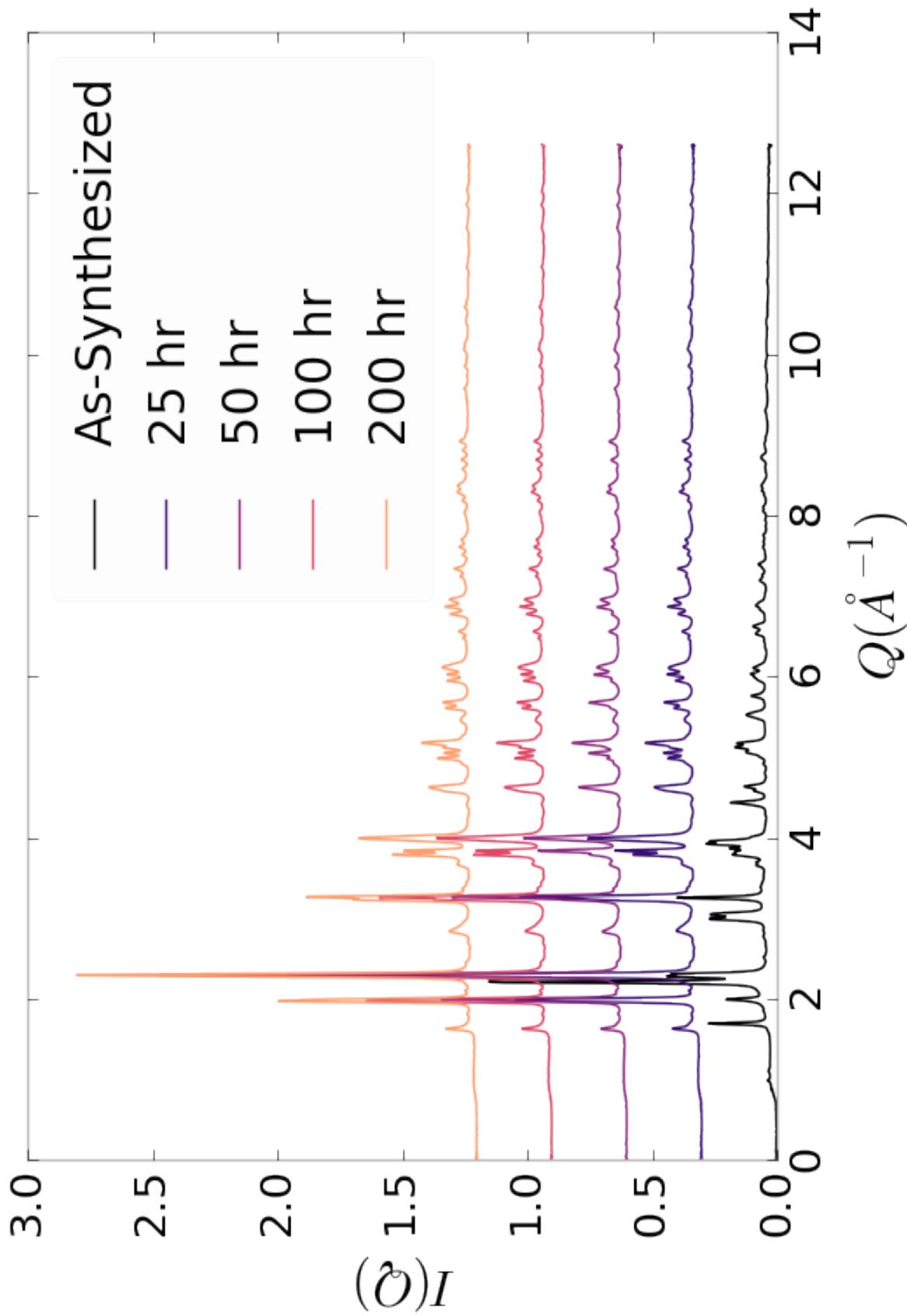


Figure A.25: Comparison of PNO sample $I(Q)$ as a function of annealing time at room temperature

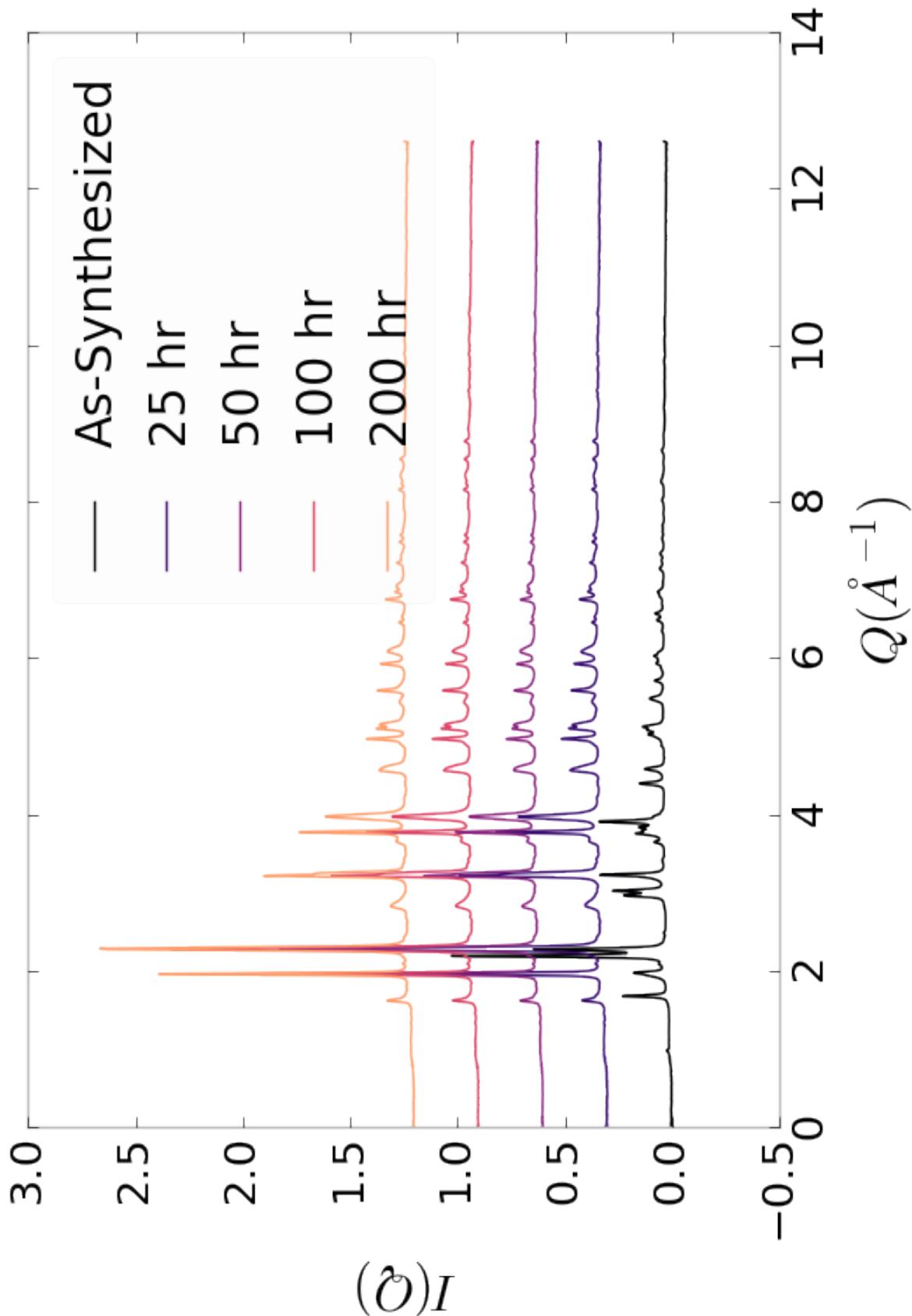


Figure A.26: Comparison of PNO sample $I(Q)$ as a function of annealing time at operating temperature

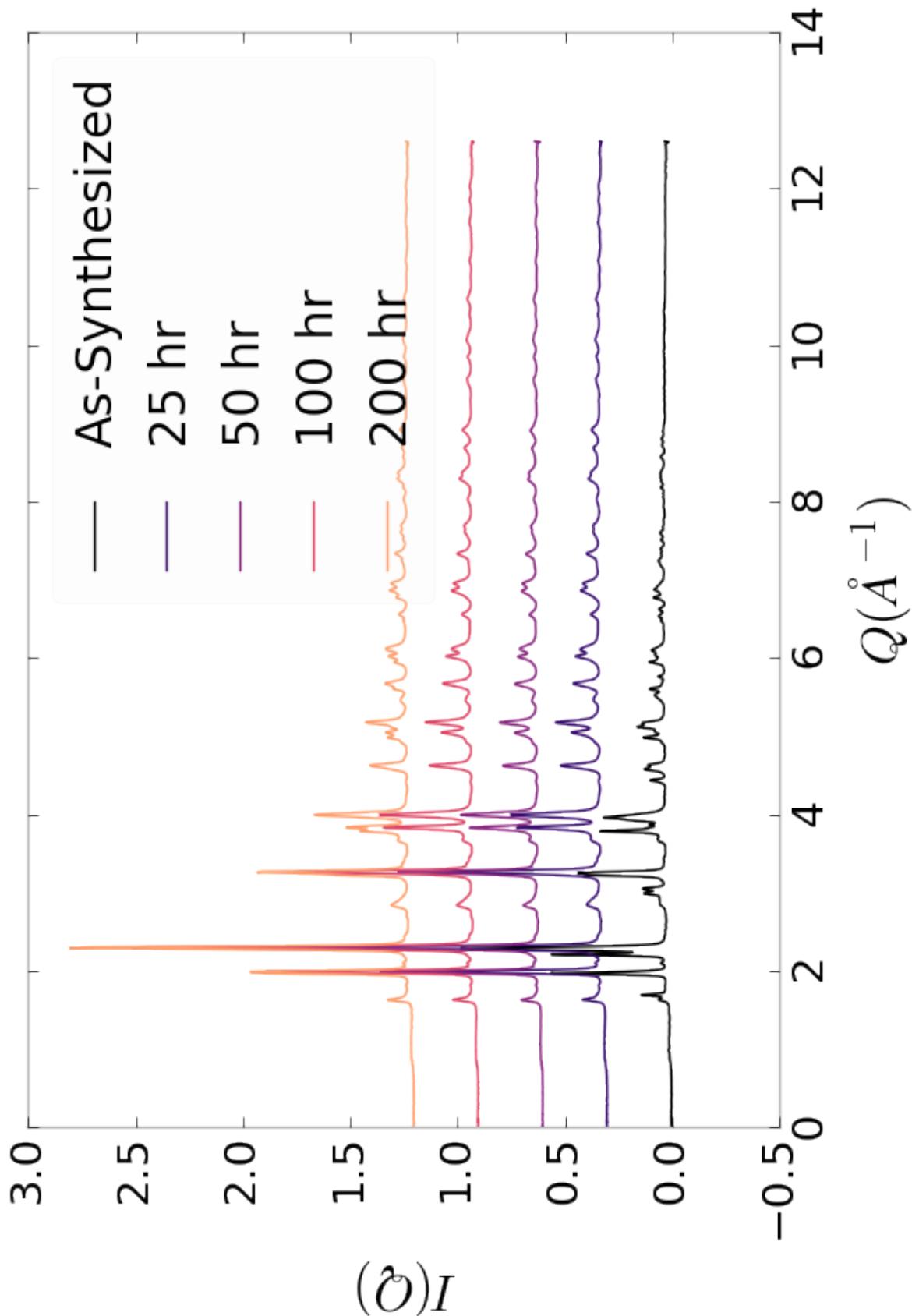


Figure A.27: Comparison of PNO sample $I(Q)$ as a function of annealing time cooled back to room temperature