(Temporary Title)

Investigation of Bond Strain Effects on XANES Spectra by Supervised Machine Learning

by
Jeremy K. Thaller

Professor Anatoly Frenkel, Advisor
Brookhaven National Laboratory
Chemistry Division
New York, USA

Professor Wolfgang Schmahl, Advisor
Ludwig-Maximilians-Universität
Fakultät Geowissenschaft
München, Germany

A Master thesis submitted to the Faculty of Earth- and Environmental Sciences
of Ludwig-Maximilians-Universität München in the framework of MaMaSELF

April 27, 2021

Abstract

Your abstract will summarize your thesis in one or two paragraphs. This brief summary should emphasize methods and results, not introductory material.

Executive Summary

Your executive summary will give a detailed summary of your thesis, hitting the high points and perhaps including a figure or two. This should have all of the important take-home messages; though details will of course be left for the thesis itself, here you should give enough detail for a reader to have a good idea of the content of the full document. Importantly, this summary should be able to stand alone, separate from the rest of the document, so although you will be emphasizing the key results of your work, you will probably also want to include a sentence or two of introduction and context for the work you have done.

Acknowledgments

The acknowledgment section is optional, but most theses will include one. Feel free to thank anyone who contributed to your effort if the mood strikes you. Inside jokes and small pieces of humor are fairly common here . . .

Contents

List of Figures

Chapter 1

Introduction

The introduction is one of the most important pieces of your thesis. Here is a place for you to introduce the problem(s) on which you have worked and place them in the larger context of your field. You should aim to ensure that this section is completely understandable to virtually anyone - and certainly anyone with a sophomore-level grasp of physics. Presumably, this will include references to the literature.

In addition to setting your work into context, a second good idea for your introduction is to give a short outline for what the rest of your thesis will discuss. This is often done in the closing paragraph(s) of the introduction with sentences like "In the following chapters ..." and "Chapter 2 discusses ..." Tremendous detail is not required in this outline, but rather just a brief road map for the rest of the document.

1.1   X-ray Absorption Spectroscopy

I want to describe the problem we're trying to solve in this section. So I want to motivate the problem by describing XAFS a little so I can describe the limitations of XANES and why this project is useful. More in-depth explanations can be placed in chapter 2

X-ray absorption spectroscopy measures the absorption of high-energy photons by a sample as a function of energy [1]. The attenuation, or change in transmitted light intensity as a result of inelastic processes, is characterized by the Beer-Lambert Law (1.1). For an incident beam of intensity I_0, the transmitted intensity after interacting with an attenua- pick one tion/absorption coefficient of μ and a sample of thickness x is:

$$I = I_0 e^{-\mu x} \tag{1.1}$$

Above the absorption edge, the condensced state has characteristic absorption jumps where the incident photon's energy matches the binding energy of a core electron. At this energy, nearly all the photon's energy is absorbed by the core electron, resulting in the characteristic absorption-edges first observed in 1920 [2, 3].

1.1.1 Synchrotron Radiation

I'm not sure where this will end up going, but somewhere I'll have to write about the origin of synchrotron radiation

Synchrotron radiation was first observed by General Electric in Schenectady, New York [4]. Initially just a side effect of particle accelerator experiments, it has since grown to be an important and powerful source of high-energy electromagnetic radiation for structural determination. Compared to lab-scale x-ray generation for diffraction experiments, arguably the most important benefit of synchrotron radiation is its high brilliance. Synchrotron radiation creates a highly collimated beam of photons characterized by small divergence and spatial coherence. Additionally, synchrotron radiation is tunable across a wide spectrum (microwaves to hard X-rays) and capable of high flux, useful for short time-scale-dependent experiments or weak scatterers. Synchrotron radiation can be produced in a pulsed structure, and the incoming photons are highly polarized, either linearly or circularly depending on where the measurement system lies with respect to the plane of the synchrotron.

frequency range?

wordy

1.1.2 XAFS

X-ray Absorption Fine-Structure (XAFS) spectroscopy refers to the study of absorption spectra created from high-intensity x-ray interactions. As the energy of the incident radiation is increased, the photon's energy will eventually match the binding energy of a core-level electron. As a result, an "edge" in the spectrum will be observed. The location of these edges depends on the chemical and physical structure, as well as the electronic and vibrational state of the material, creating characteristic absorption oscillations.

(Here's a good source to cite a bunch: the 2000 review of XAFS [5]) [6]

absorption edges like fingerprints to identify elements. in 1920 Frische and Hertz observed peak shape, and 40 years later it was learned that this shape can be used to probe the short-range order.1971 e.a. Sterne expalin this effect. Fermi's golden rule. The inelastic interactions of the photon are related to characteristic energies. When the photon energies match the energy difference from the core-electron state to the first unoccupied level, the photon can be fully absorbed (before that only partially absorbed.)

$$\mu(E) = aE^{-3} + bE^{-4} \tag{1.2}$$

$$\mu(E) \propto \sum_{f}^{E_f > E_F} |\langle f|H_{int}|i\rangle|^2 \, \delta(E - E_F - Ef) \tag{1.3}$$

1.1.3 EXAFS

Extended X-ray Absorption Fine Structure (EXAFS)

1.1.4 XANES

1.2 Disorder in XAFS

1.3 Nanoparticles?

1.4 Machine Learning?

Probably want to talk about these papers in this section [7] [8].

1.5 Outline of the Thesis

Building off the historical context of measuring disordered nanoparticles in XANES spectra, this thesis begins with an in-depth description of

Chapter 2

Simulating Disorder

Neural networks rely on a large quantity of training data in order to be able to able to make any predictions. Specifically, to teach our neural network to predict the mean squared displacement (MSD), we must first generate a large quantity of training data comprised of XANES spectra, each labeled with a known MSD. Rather than spend countless hours generating this training data experimentally or simulating each possible disordered structure individually, we have generated this training data via clever statistical averaging of simulation data from simple, non-disordered structures. In this chapter, we explain the aforemention statistical weighting process in-depth –from creating the dataset for initial simulations to finally creating the neural network training data.

> Of just Au NPs? Can we generalize this to the sigma of the PRDF?

> Maybe give a brief outline of the process so you can understand the point of the sections as a walk-through before the very end

2.1 Generating Distortion Not Disorder

Instead of creating structures with a range of *disorder*, we begin instead by generating structures with a range of *distortion*. Here, the distinction is that distortion refers to isotropic expansion, or, equivalently, a radial shift in all atomic positions away from (or towards) the center atomic absorber.

Figure 2.1 only shows the xy-plane projection of the first 12 nearest neighbors, but the actual original structure consists of the first four shells (55 atoms) with a lattice constant of 4.0782 Åto match that of bulk Au. In reality, the nearest-neighbor distances for Au nanoparticles are likely smaller, but the important part is that the crystal structure is correct, especially since the original coordinates will only be one structure out of many.

> Citation? Wolfram Element Data?

A total of 91 FEFF input files are generated, each with the same center absorber located at $(0, 0, 0)$ but with the other atomic coordinates in a shifted location. They are generated with shifts on the range of -0.45 Å to $+0.45$ Å in increments of 0.01 Å. For example, the FEFF input file with the greatest inward shift has coordinates shifted 0.45 Å radially inwards towards the center absorber, and the FEFF input file with the largest outwards shift has
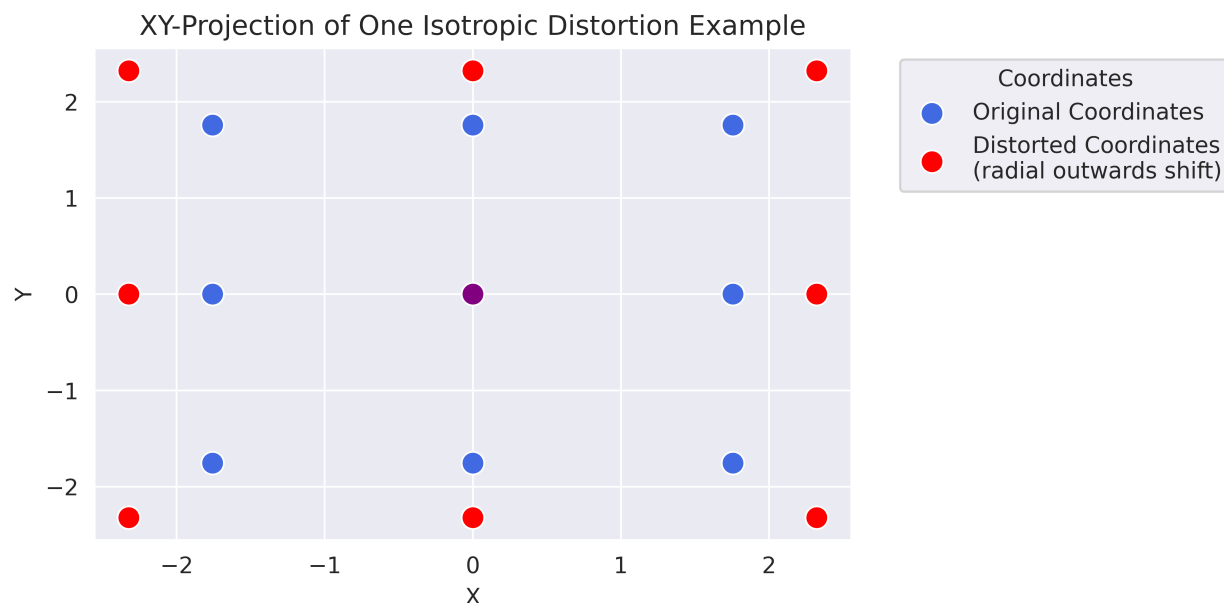
Figure 2.1: Each point represents an atom of an FCC unit cell projected onto the xy-plane. The blue atoms represent the original coordinates, and the red atoms represent the radially shifted coordinates. The center absorber atom is purple since its original position is the same as its distorted position

> Is it fair to say this? Au is FCC, but FCC unit cells do not have a center. I am choosing a different point such that it actually looks more like BCC

coordinates shifted 0.45 Å radially outwards away from the center absorber.
    Each FEFF input file is run with the following parameters:

```
1    SCF 4.6 0 30 .5 1
2    EDGE    L3
3    EXCHANGE     5    0.2 0.5
4    S02 1.
5    XANES    3.7 0.05    0.1
6    FMS 7
7
8    POTENTIALS
9    0         79         Au         -1         -1         0.
10   1         79         Au         -1         -1         0.
```

    Running the 91 simulations (one for each of the distorted structures) takes half an hour

> I mention it because I want to make a point of how much computationally faster doing this XANES averaging process is

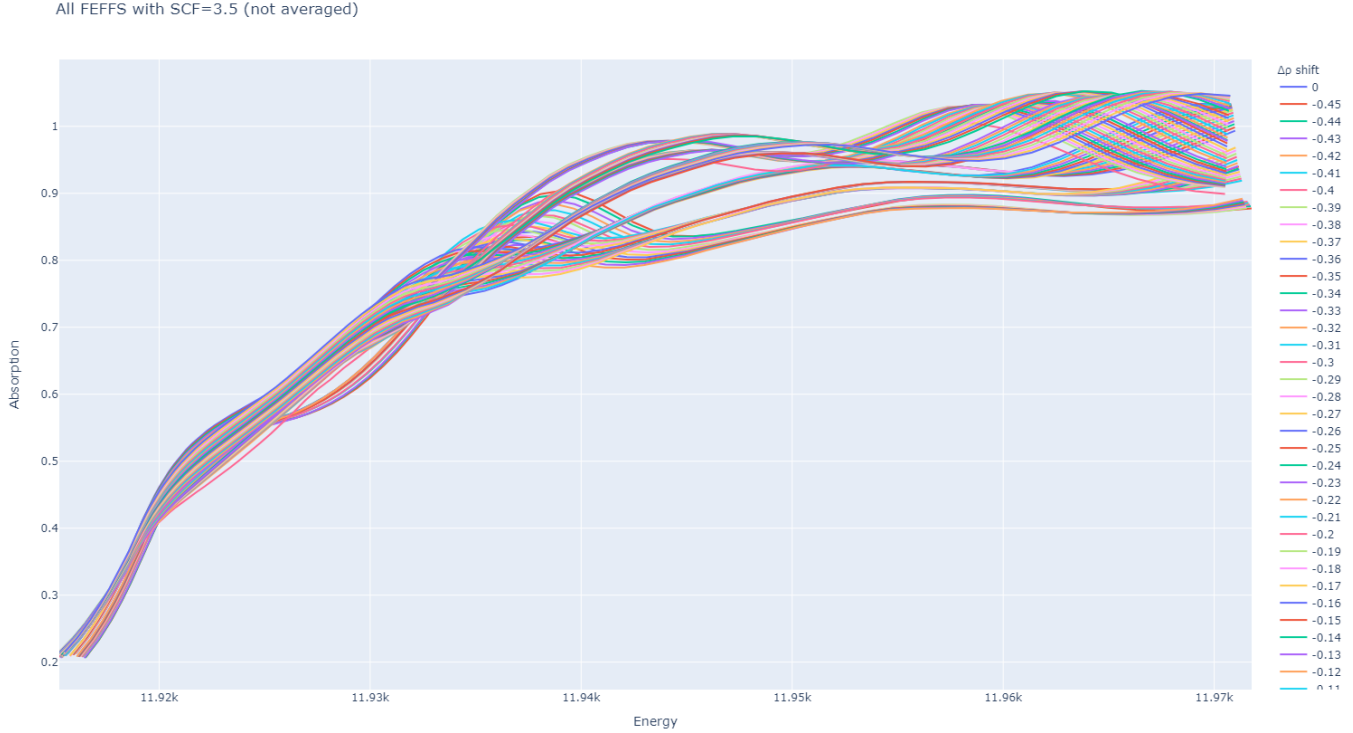. The resulting XANES spectra are plotted in figure 2.2.

Figure 2.2: TEMPORARY. Each spectrum represents the FEFF simulations result for a different distorted structure. That is to say, the crystal structure and center absorber are the same for each, the only parameter that varies is the euclidean distance from the center to the other coordinates.

2.2 Generating Disorder via Probability Distribution Averaging

Disorder in a system can be characterized by the Gaussian width, sigma, of the partial radial distribution function. The idea of our statistical averaging method emulates/simulates this width by weighting the simulated XANES spectra accordingly. For example, figure 2.3 depicts a histogram with $\sigma = 0.1$ Å. Each histogram bin represents a simulated XANES spectrum with a different isotropic displacement. For example, the bin at $\Delta\rho = 0.0$ Å represents the simulated XANES spectrum with no distortion, and the bin at $\Delta\rho = -0.2$ Å represents the simulated XANES spectrum with all the atomic coordinates shifted isotropically inwards towards the center absorber by 0.2 Å. The height of each bin, $f(\Delta\rho)$, represents the relative contribution of each simulated XANES spectrum towards the resulting weighted

spectrum. For visual clarity, figure 2.3 depicts only 40 bins; the actual weighting includes 91 bins ranging from -0.45 Å to $+0.45$ Å in increments of 0.01 Å.

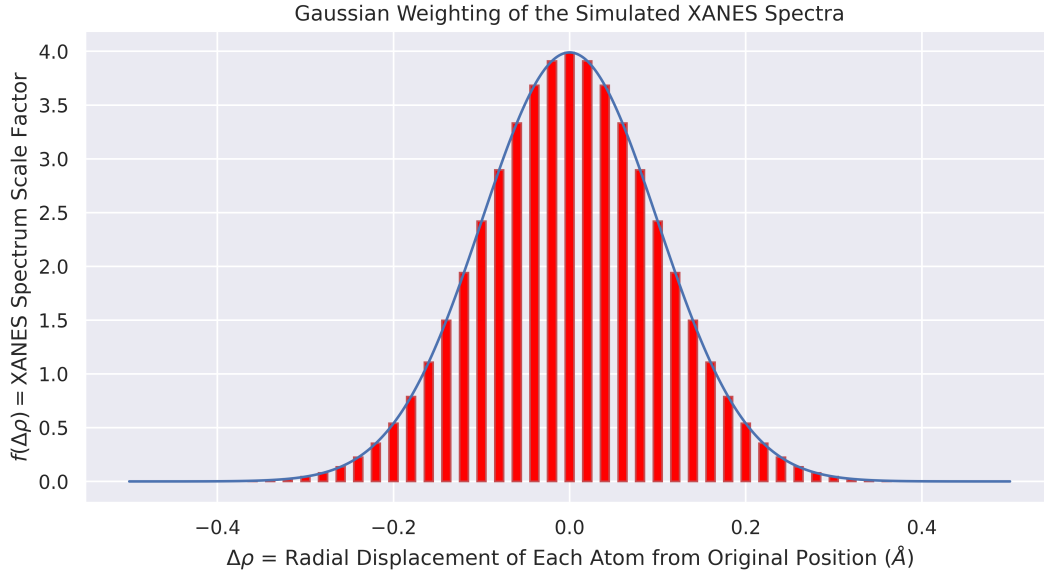Already mentioned earlier. Choose one.

Figure 2.3: A Gaussian distribution probability density function can be used to calculate the relative weight of each FEFF generated XANES spectrum towards one simulated, disordered spectrum. Each bin (red bar) represents a FEFF generated spectrum; the x-axis is the isotropic shift of the atomic positions, and the y-axis being the relative weight factor.

To calculate averaged XANES spectrum, $\langle \mu(E) \rangle$, using the histogram weighting of the gaussian in figure 2.3:

$$\langle \mu(E) \rangle = \frac{1}{S} \sum_{\Delta\rho=-.45}^{+.45} g\left(\Delta\rho \mid \mu = 0, \sigma = 0.1\right) \mu(E \mid \Delta\rho) \tag{2.1}$$

In the above equation $\Delta\rho$ is the isotropic, radial displacement of each atom from their original positions and $\mu(E \mid \Delta\rho)$ is the simulated FEFF spectrum for the given $\Delta\rho$ configuration. Furthermore, (2.1), S represents a standardization factor to negate the effect of the changing height of the Gaussian as a function of the variance, σ^2. In this way, only the relative heights of each bin matter for producing the averaged XANES spectrum. The standardization factor, S, can then be defined as:

$$S = \sum_{\Delta\rho=-.45}^{+.45} g\left(\Delta\rho \mid \mu = 0, \sigma = 0.01\right) \tag{2.2}$$

In both equations 2.1 and 2.2, the function g is just the typical Gaussian distribution probability density function:

$$g(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-(x-\mu)^2/2\sigma^2} \tag{2.3}$$

The above example only generates one (simulated) disordered XANES spectrum and does so via weighting of a Gaussian distribution with mean and variance equal to 0 and 0.01, respectively. To simulate systems with different degrees disorder, we can vary the shape of the probability density function. With a Gaussian distribution, we can only vary the mean and variance, but to simulate even more conditions, we can instead use the multivariate skew-normal distribution (eq 2.4) [9, 10]

$$f(x) = 2\phi(x)\Phi(\alpha x) \tag{2.4}$$

where $\phi(x)$ is the Gaussian PDF:

$$\phi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}} \tag{2.5}$$

and $\Phi(x)$ is the Gaussian CDF:

$$\Phi(x) = \int_{-\infty}^{x} \phi(t) \ dt \tag{2.6}$$

Equation 2.4 includes the shape parameter, α, which has the nice properties of producing a right-skewed distribution when positive and a left-skewed distibution when negative. When $\alpha = 0$ the distribution simply produces a the typical Gaussian distribution (eq. 2.3). By using equation 2.4, we can vary μ, σ, and α to vary the first four moments of the function: mean, standard deviation, skew, and kurtosis. Eighteen possible skew-norm weighting functions are plotted in Figure 2.4. To train the neural network, 1000 such weightings are used.

The disorder of the skew-norm generated, disordered spectrum is characterized by the mean-squared displacement of each atom from its original position ($\Delta\rho$), weighted in the same manner as the spectra.

Instead of characterizing the disordered spectrum by the standard deviation of the gaussian used to create it

Mathemtically, the equation to calculate the weighted mean squared displacement, MSD, is:

$$MSD = \frac{1}{S} \sum_{\Delta\rho=-.45}^{+.45} f\left(\Delta\rho \mid \mu, \sigma^2, \alpha\right) \tag{2.7}$$

where $f(x)$ is the skew-norm function in equation 2.4, and the MSD of each individual FEFF spectrum is equal to the isotropic distortion, $\Delta\rho$.
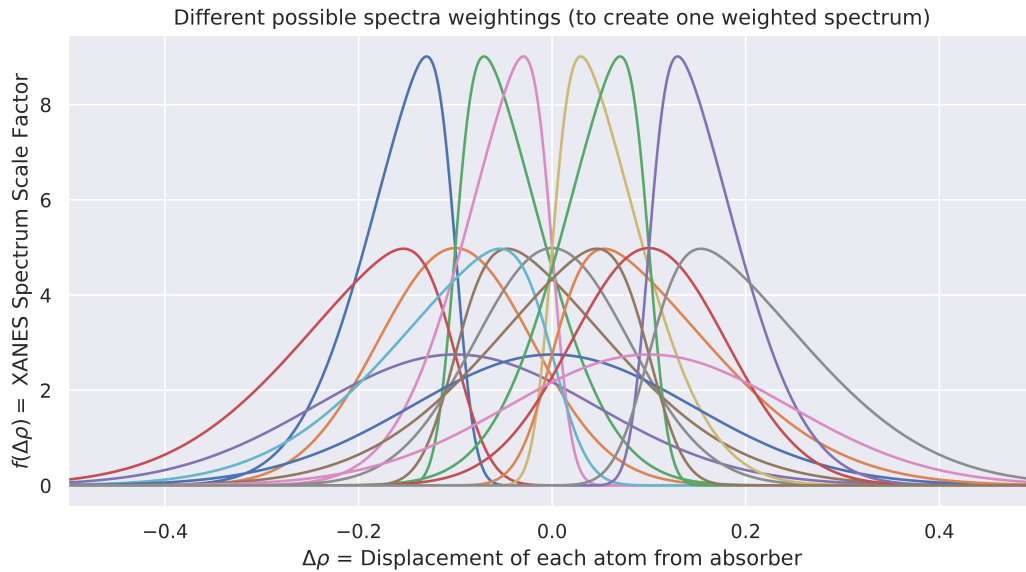
Figure 2.4: Eighteen skew-norm distributions plotted with all possible permutations of $\sigma = \{.08, .145\}$, $\mu = \{-.1, 0, .1\}$, and $\alpha = \{-5, 0, 5\}$. Each represents a possible way to produce a simulated, disordered spectrum from many FEFF-simulated, distorted spectra

2.3 Simulation vs. Experimental Data

To check our FEFF simulation parameters, as well as the validity of the gaussian-weighted disorder technique, we compare the simulation data to experimental data. In figure 2.5, both experimental and simulation spectra for bulk-like and nanoparticle scenarios are plotted. EXAFS fitting was used to characterize the disorder in the experimental measurements. For the bulk foil, this parameter was found to be $\sigma^2 = 0.0081(5)$ Å^2 , and for the 8 nm disordered particle, $\sigma^2 = 0.0102(8)$ Å^2. One simulated, disordered spectrum was weighted according to the gaussian $N(0, 0.09)$ to represent the disordered nanoparticle, and the other was weighted according to the gaussian $N(0, 0.038)$ to represent the Bulk. These weightings correspond to a MSD value that matches the measured σ^2 values for the experimental data.

Can some- one give me citations for these?

In Figure 2.5, the bulk Au foil spectrum is above the 8 nm nanoparticle spectrum (more absorption) until the peak around 11937 eV, where the NP absorption becomes higher. The two criss-cross again over the next two peaks, changing which material has the higher absorbance in an energy range. This change is more easily seen in figure 2.6, which plots the difference between the the bulk material and the nanoparticle absorption for both the experimental measurements and the simulations. The experimental and simulation difference-spectra follow the same trend with the exception of the peak around 11947 eV.

Figures 2.5 and 2.6 aren't meant to be perfect comparisons of simulations vs. experimental data. For one, the the experimental data compars a bulk spectrum to a nanoparticle. By constrast, both the simulation spectra are of the same size 55 atom cluster. Still, much
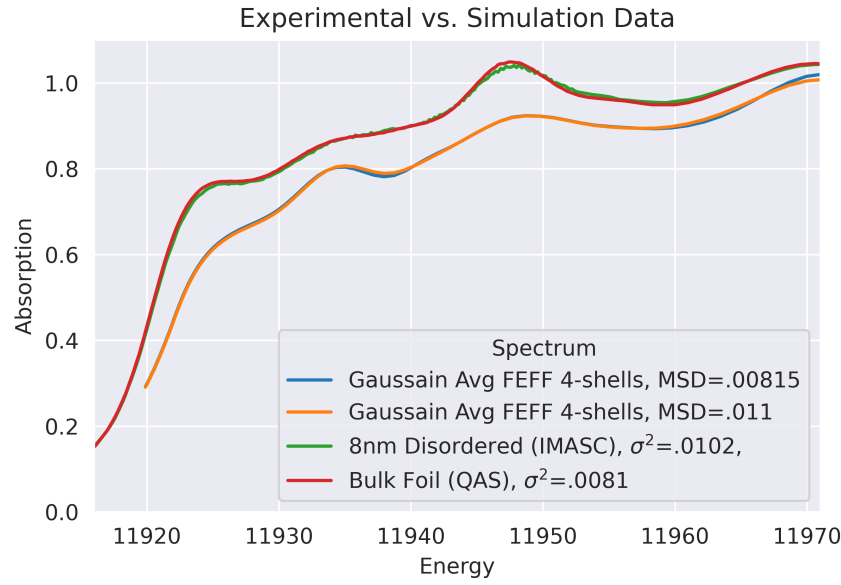
Experimental vs. Simulation Data

Figure 2.5: Comparing the bulk foil (red) measurement to the 8 nm disordered nanoparticle (green) measurement is an analog to comparing the simulated, non disordered FEFF spectrum (orange) to the simulated disordered spectrum (blue).
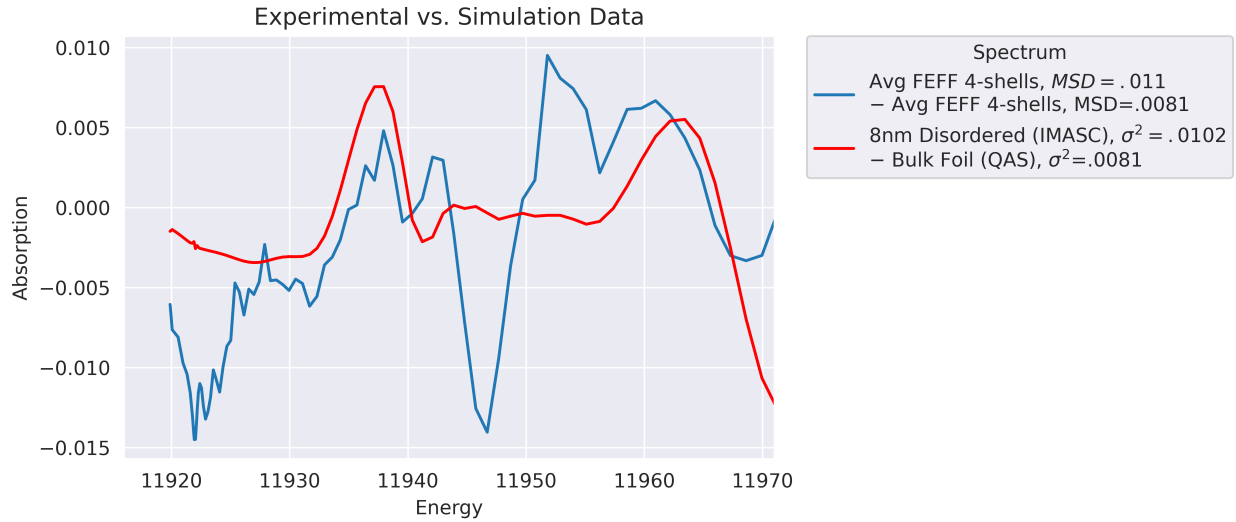

Figure 2.6: The difference between the nanoparticle spectrum and the bulk spectrum are plotted for the same data as in Figure 2.5. It is easier to see where the bulk and the nanoparticle absorption crisscross by plotting the difference.

of the disorder trends are coded in the simulation approach.

To test if the size information is also coded in our simulations we compare different size

simulations to experimental data in figure 2.7. As expected, more atoms in the simulation produces more bulk-like spectrum characterstics such as larger amplitude peaks.
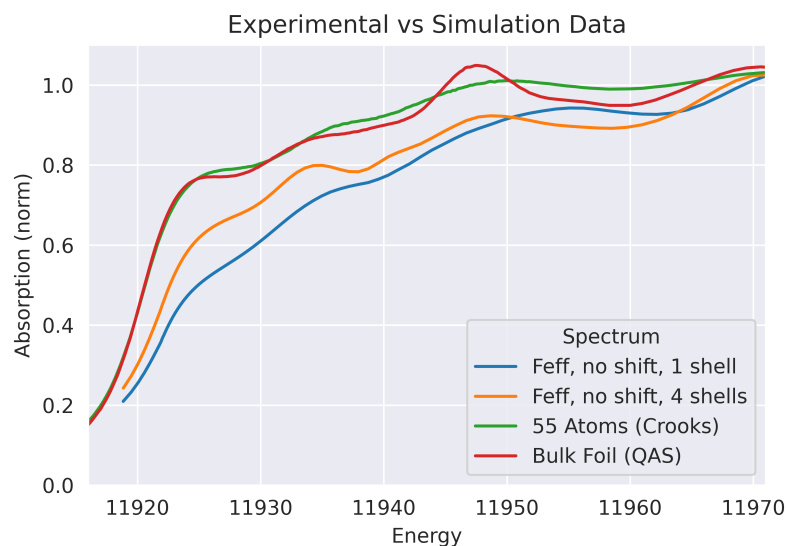

Figure 2.7: Comparing the bulk foil (red) measurement to the 55 atom nanoparticle (green) measurement is an analog to comparing the 13 atom simulated spectrum (blue) to the 55 atom simulated spectrum (orange).

Chapter 3

Machine Learning

An explanation of Machine Learning and Neural Networks in general

3.1   Autoencoders if they become useful

Talk about how autoencoders work. Give a nice broad explanation and really go into the math. Include some nice diagrams

Here's [11] a good source to read and model off of. Here [12] is another paper that might be interesting to read. It's about getting noise-free data from the original data using an autoencoder. Neat idea, and could actually be very relevant because they're using geophysical data.

Chapter 4

Results

Here I expect to showcase lots of nice figures and data to show how well the neural network works

4.1   temp

Appendix A

An appendix

Appendices are a good idea for almost any thesis. Your main thesis body will likely contain perhaps 40-60 pages of text and figures. You may well write a larger document than this, but chances are that some of the information contained therein, while important, does *not* merit a place in the main body of the document. This sort of content - peripheral clarifying details, computer code, information of use to future students but not critical to understanding your work . . . - should be allocated to one or several appendices.

Bibliography

[1] D. J. Gardenghi et al., Ph.D. thesis, Montana State University-Bozeman, College of Letters & Science (2012).

[2] H. Fricke, Physical Review **16**, 202 (1920).

[3] G. Hertz, Zeitschrift fuer Physik **3**, 19 (1920).

[4] F. R. Elder, A. M. Gurewitsch, R. V. Langmuir, and H. C. Pollock, Phys. Rev. **71**, 829 (1947), URL https://link.aps.org/doi/10.1103/PhysRev.71.829.5.

[5] J. J. Rehr and R. C. Albers, Reviews of modern physics **72**, 621 (2000).

[6] M. Newville, Reviews in Mineralogy and Geochemistry **78**, 33 (2014).

[7] J. Timoshenko, A. Anspoks, A. Cintins, A. Kuzmin, J. Purans, and A. I. Frenkel, Physical review letters **120**, 225502 (2018).

[8] J. Timoshenko, D. Lu, Y. Lin, and A. I. Frenkel, The Journal of Physical Chemistry Letters **8**, 5091 (2017).

[9] A. Azzalini and A. Capitanio, Journal of the Royal Statistical Society: Series B (Statistical Methodology) **61**, 579–602 (1999), ISSN 1467-9868, URL http://dx.doi.org/10.1111/1467-9868.00194.

[10] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, et al., Nature Methods **17**, 261 (2020).

[11] A. Ng et al., CS294A Lecture notes **72**, 1 (2011).

[12] D. Bhowick, D. K. Gupta, S. Maiti, and U. Shankar (2019), 1907.03278.