Decomposition Approach to Parametric Nonconvex Regression; Nuclear Resonance Analysis

Jordan L. Armstrong
University of Tennessee, Knoxville, jarmst37@vols.utk.edu

To the Graduate Council:

I am submitting herewith a thesis written by Jordan L. Armstrong entitled "Decomposition Approach to Parametric Nonconvex Regression; Nuclear Resonance Analysis." I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Industrial Engineering.

Hugh Medal, Major Professor

We have read this thesis and recommend its acceptance:

Jim Ostrowski, Vladimir Sobes

Accepted for the Council:
Dixie L. Thompson

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Decomposition Approach to Parametric Nonconvex Regression; Nuclear Resonance Analysis

A Thesis Presented for the

Master of Science

Degree

The University of Tennessee, Knoxville

Jordan L. Armstrong

December 2021

Acknowledgements

Abstract

Parameterized nonconvex regression is a difficult problem for any optimization solver packages, often resulting in approximations and linearizations of the problem in order to be able to arrive a solution, if the problem is even solvable at all. These changes to the initial problem are largely dependent upon having appropriate domain knowledge and still often times result in a sizable gap between the achieved solution and the best true solution. We propose a novel method of decomposing the global problem into small, overlapping windows. Thus, the independent windows are now solvable. Subsequently, we offer a novel, sequential method of parameter cardinality and parameter value agreement in order to stitch the windows back together to arrive at the solution to the initial global problem. While this method is problem agnostic, we demonstrate the successful results of its application to the nuclear data analysis problem of properly characterizing the resonances of the capture cross section for Copper-63. By being able to solve the 100 resonance problem, this method demonstrates it can solve up to the thousands of possible resonances an isotope can have within a single spin group.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction and Background

This chapter will serve as a general outline of how global and nonconvex optimizations problems are solved, while nuancing the advantages and disadvantages of these different approaches. By giving this background, it helps demonstrate a partial explanation for why this algorithms cannot handle such large, nonconvex optimization formulations such as the one posed by the nuclear resonance problem. Additionally, the motivation behind why the nuclear resonance characterization problem is an important is highlighted, as solving this problem has great implications throughout its field. Lastly, the chapter concludes with a brief high level overview of how BARON operates, which was the selected mathematical optimization software for this project.

1.1 Optimization Background

The general nonlinear programming problem currently has no effective methods for being solved [3]. Problems spanning hundreds of variables can be intractable, while even simplistic problems with tens of variables can be difficult. Thus, methods for solving the general nonlinear programming problem take on many varying approaches, each involving their own compromises. However, the difficulty not only steeply increases for nonlinear problems, but is even more apparent for nonconvex

problems, as "in fact the great watershed in optimization isn't between linearity and nonlinearity, but convexity and nonconvexity" [4].

What makes nonconvex optimization challenging is potentially having many local minima, the potential existence of saddle points and very flat regions, and having widely varying curvature. Thus, this class of problems is NP-hard [5]. When having close initialization and favorable geometry, local convergence to a local minima via iterative methods, such as gradient descent, can be proven, but often times these stronger properties (such as the strict saddle property) for local convergence cannot be shown [6, 7]. Multistart implementations of iterative methods can help get around the close initialization issues, but at the expensive of a much greater runtime and not much to be gained about the true global solution. Methods for solving nonconvex problems like a stochastic gradient descent can easily end up finding a saddle point, local maximum, or region of general flatness. Therefore, these techniques require more advanced analysis and methods to even demonstrate that a solution is a truly a local minima and that is when these methods even converge to a solution, which is rarely guaranteed for a nonconvex function [8].

There are two more general methods for solving nonconvex problems. The first being convex relaxation, where in simplest form the objective function is convexififed by its convex envelope, and the constraints are relaxed with their convex hull [9]. However, even these problems can be computationally intractable, hence requiring even further hierarchical relaxations. Once a solvable convex problem is reached, one must use convex analytic mathematics to demonstrate the solution is global solution to the original problem, which is not a simple task. This methodology is the typical go-to for sparse vector and low-rank matrices problems of structure signal recovery [10]. One of the major shortcomings of convex relaxation is its computational burden, as the resulting formulation is often times a generic semi-definite program. Furthermore, while convex relaxation is convenient due to it building upon the nice mechanics of convex analysis, it also do not ensure correctness to the original nonconvex problem, whether that be the tractability of the relaxed problem or in how

close the relaxed solution is to the true global solution. For instance, in nonneagtive low-rank approximation, convex relaxations are not efficient and in sparse PCA, these relaxations are tractable but are shown to be suboptimal in their solutions [11, 12]. The second approach is the transforming the optimization space, such as with the always concave Lagrangian dual problem [7]. Other transformations include solving the trust-region subproblem or approximating the MAX-CUT problem [13, 14]. The major issue with these types of approaches is that they are highly problem-specific and involves subsequent local refinements.

Being able to solve nonconvex problems is important and interesting, and is not just a small niche of optimization problems. However, due to the details already examined, these problems greatly range in difficulty. For example, Principal Component Analysis (PCA), finding the dominant eigenvalue-eigenvector pair of a positive semidefinite symmetric matrix, is nonconvex. However, the convex methods of gradient descent (standard, stochastic, momentum, variance reduction, etc) all work in solving PCA and explicit convergence rates can be demonstrated. But nonconvex problems can get much harder, for instance with the mathematics behind deep neural networks, where proving convergence is almost always impossible. Therefore, this class of problems can be hard to generalize methods for, but is an important class to be able to solve even with its list of possible things that can go wrong of converging to a bad local minimum, saddle point, or region of low gradient magnitude, or large curvature causes diverging steps.

Often times, methods for solving these types of problems result in a search for a solution that is only locally optimal, which is usually faster and more widely applicable, however are sensitive to the initial guess of the solution, sensitive to algorithm parameters, and little information is known about the gap between the local and global solutions [3]. Being able to solve global optimization problems is an important group of problems that range from safety verification of nuclear systems to protein structure predictions to radio signal propagation [15, 16, 17]. Various methods such as Broyden's method, inexact Newton methods, and tensor

3

methods can be implemented to solve nonlinear problems, but must be done locally and are not guaranteed to converge [18, 19]. On the other hand, attempting to perform global optimization is at the cost of efficiency, with worst-case complexity growing exponentially with problem size [3]. Typically global optimization methods necessitate computing lower bounds on the optimal value, either via replacing nonconvex constraints with convex, looser constraints or by solving the Largrangian dual problem [20].

The main classes of global optimization methods are deterministic methods and stochastic methods. Two of the most successful deterministic methods are cutting plane algorithms and branch & bound methods. Cutting plane algorithms focus upon approximating the feasible region by a finite set of closed half spaces and solve that sequence using linear programs [21]. The downsides of these methods are numerical instability, requiring a high number of cuts to make solid progress towards the solution, as well as a lot of work needing to be performed to determine optimal cuts. Branch & bound attempts to avoid complete enumeration of the solution space by solving a relaxed version of the problem, branching on a variable creating two nodes, and obtaining relative upper and lower bounds of the nodes. Many questions arise in the best implementation of branch & bound such as which node to process first, which variable to choose for branching, and how often to run heuristics to help find a good solution [22]. The cutting plane algorithms can be combined with branch & bound to be an improved method called branch and cut. While stochastic methods include direct Monte-Carlo sampling and parallel tempering and rely upon random simulations in order to find the solution [23]. These deterministic and stochastic global optimization methods also do not include various types of heuristics such as ant colony optimization (a probabilistic technique for finding good paths through graphs), simulated annealing (a metaheuristic for approximating global optimization in a large search space), or graduated optimization (solving a greatly simplified problem, and progressively transforming that problem until it is equivalent to the original problem) [24, 25, 26].

It is important to note that while this method is for curve fitting, a method choosing to denoise the data was not desirable. The example of the nuclear data resonance analysis problem exemplifies this concern as many of the resonances widths are very small and can be overlapping with one another, often times a small resonance being located on the tails of a large resonance. Thus, most denoising methods would lose the details of the little resonance and would show it being absorbed by the larger resonance. Typical regularization methods such as Tikhonov regularization or lasso regularization were avoided, as the parameters in these problems have physical meaning to them and the goal is not to necessarily keep them small, but keep them close to their true underlying values and distributions [27, 28]. On that note, future work may look at adding maximum likelihood penalties to the objective function, in particularly for problems where all of the parameters have their well-known distributions, but that currently has not been needed to obtain quality results, therefore is not worth the extra computational burden as the objective function will already be cumbersome due to its nonconvex nature [29].

We are proposing a "fast" method that can achieve a global solution to a nonconvex problem without providing an initial guess or rely on proper local initializations, without having to approximate or relax the objective function or constraints, and letting the model determine the appropriate number of parameters. While still having a few hyperparameters to select, this method should be applicable to a whole class of nonconvex problems. The implementation will be demonstrated upon the results of its application to nuclear data analysis.

1.2 Nuclear Resonance Background

The current practice for resonance evaluation is laborious and time consuming as it involves substantial man-hours to evaluate experimental data and is reliant on specific expert judgement. Additionally, these processes are susceptible to errors and have limited reliable reproducibility. Modern AI/ML capabilities hold the potential

for faster, more reproducible, and more reliable evaluation of nuclear data. These methods would reduce the workload of the human workforce allowing focus on high expertise-high impact tasks, improve the information archiving and preservation, and enable more detailed analyses of cross section uncertainty. Given the number of prospective benefits, now is the time to leverage AI/ML for nuclear data evaluation, and this project is a significant development toward that objective.

The United States Department of Energy (DOE) Office of Nuclear Physics understands the opportunities associated with the application of AI/ML methods to the nuclear data pipeline. The Nuclear Data Interagency Working Group Funding Opportunity Announcement states, "Of particular interest are applications of artificial intelligence and machine learning to the nuclear data pipeline challenges" [30]. This view is strongly supported by the community of scientists working in nuclear data research. In the Proceedings of the Workshop for Applied Nuclear Data (WANDA) it is explicitly stated that, "The fast spread and impact of machine learning and artificial intelligence models to diverse areas of physical sciences indicate their tremendous potential to address critical issues and potential bottlenecks in the nuclear data pipeline" [31]. Introduction of AI/ML methods into the nuclear data pipeline will accelerate the access of the end users to the new nuclear data evaluations and the impact that new evaluations will have on predictive modeling and simulations related to the DOE mission. The current timeline is between two and five years from initial funding of the project to final evaluation of the experimental data.

Accurate and reliable nuclear data is fundamental to nuclear physics and nuclear astrophysics research. At low incident neutron energies, roughly below one MeV for most nuclides, the nuclear level density above the neutron separation energy is small and the interaction proceeds through individual levels that can be separated experimentally. It is vitally important to carefully determine the characteristics of these individual resonances for successful calculation of the neutron reaction cross sections. In turn, the cross sections are the absolute-fundamental input data for predictive modeling and simulation in nuclear power systems with a thermal neutron

spectrum, where neutrons slow down through the resonance region and individual resonances can be identified in the flux depressions at those energies. Furthermore, detailed knowledge of the resonance region is important for associated fuel cycle operations, national security and non-proliferation applications, shielding studies, materials analysis, medical radioisotope production, diagnosis and radiotherapy.

In the period of 1950s to 1980s, nuclear data experienced a rapid expansion and many experiments were performed to measure precisely these resonance properties. Since then, computational modeling and simulation of nuclear systems has had its own period of rapid expansion. The resolution of the detailed behavior of the systems, which can currently be modeled on a computer, has grown several orders of magnitude since the 1980s. Nuclear data evaluators now find themselves back in a scenario where the predictive power of radiation transport codes is limited by the input nuclear data. In turn, this has economic, safety, and security limits on the applications relying on modeling and simulation. For example, safeguards and homeland security applications rely on hybrid methods of radiation detection and computational solutions of the inverse radiation transport problem [32]. Computational modeling limited by nuclear data can limit the ability to detect special nuclear materials of interest [33].

There are three shortcomings with the nuclear data evaluations carried out during that initial, fruitful period. First, the nuclear data was measured and evaluated to the precision compatible with the computational modeling capabilities of the day which has now become a limiting factor. Second, the nuclear data evaluations relied heavily on expert judgement and were scarcely documented which makes it difficult, often impossible, to go back and make targeted adjustments or corrections to the evaluations without completely restarting the evaluation process. Lastly, uncertainty estimates on the evaluated cross section were typically not done due to the lack of computational capability at the time to propagate such uncertainty through the calculations. Some crude estimates of this uncertainty data have been put in after the fact. The availability of reliable uncertainty information for evaluated nuclear

data is crucial to understanding the predictive power of computational modeling and simulation.

Since the 1980 nuclear data evaluation has remained an active field with major new nuclear data library releases in 1990 (ENDF/B-VI), 2006 (ENDF/B-VII), 2011 (ENDF/B-VII.1), and 2018 (ENDF/B-VIII.0). For the major nuclides of interest, almost no nuclear data evaluations remain from before 1980 in the US Evaluated Nuclear Data File [34]. A major, international collaborative effort was recently completed to update the nuclear data evaluations of the six major isotopes, ^{1}H, ^{16}O, ^{56}Fe, ^{235}U, ^{238}U, ^{239}Pu [35]. There are also currently three international collaborations lead by the International Atomic Energy Agency (IAEA) taking place to update the evaluations for the light-, medium-, and heavy-weight isotopes [36]. The INDEN and CIELO are highlighted here as international collaboration efforts which show a continued interest in the development of nuclear data evaluation. Many national-level and individual nuclear data evaluation projects are also on-going. Recent years, have seen a rapid expansion in thermal neutron scattering law evaluations, fission product yield, and decay data. The development of fission theory and microscopic models in an ever-on-going effort of interest to both the applied nuclear science community as well as fundamental physics. On the other hand, the physical model in the resonance region has not changed basically since it was described in 1958 in the seminal paper by Lane and Thomas [37]. Some interest remains in the description of the physics governing the resonance region, but is mainly limited to the mathematics of R-Matrix theory [38, 39, 40]. With respect to this project, this is good, because unlike the high-energy region, in the resonance region, the R-Matrix theory of nuclear reactions is not considered to be susceptible to model defects.

Even though there is not much progress in the physics governing the resonance region, resonance evaluations continue to be updated in-line with other nuclear data regions. Resonance evaluations were part of the new evaluation the CIELO project and are currently part of all three groups of the INDEN project. Resonance evaluations take a long time. There are three main reasons. First, compiling

the historical and new experimental data is time consuming. In this process, the resonance evaluator must make sure to provide compile as complete of a representation of the experiment as possible; carefully understanding the measurement details and associated uncertainties. Recent efforts on trying to understand and quantify uncertainty in experimental measurements have revealed this to be a monumental task [41]. Collecting this information is valuable not only for accurate evaluation of the resonance parameters, but it is also important to properly estimate the uncertainty in the evaluation. The process is further muddled by the incompleteness of information about the experiment [42].

Secondly, R-Matrix theory can only predict the cross sections if all of the resonances have been correctly identified. Currently, resonance identification is done manually by expert nuclear data evaluators. This is a laborious, time-consuming process that is far from perfect. The fact that resonance evaluators cannot correctly identify all of the resonances is well documented. See for example, [43, 44], that show plots of missing resonances from the evaluation and a imbalance between the number of resonances and the distribution of their properties among the different spin groups. Currently, the feedback on missing resonances is manually introduced by the evaluator through an iterative process of; identify and classify resonances, evaluate statistical distribution of resonance properties, adjust the spin group assignments, and repeat. On the other hand, getting the spin group assignment correct is important for the evaluation of the average resonance properties and the strength function which informs the unresolved resonance region and the high energy region evaluations. Furthermore, missing small resonances leads to a drop in background of the capture cross section which can only be compensated by including an artificial, pointwise background cross section, which violates the unitarity of the R-Matrix physics.

Lastly, the covariance evaluation. In the resonance region, it is well-documented that a systematic evaluation of the uncertainty on the evaluated cross section is unreliably low. This is explicitly addressed explicitly in Section IV.E.6. in the User's Manual of the world's most popular resonance evaluation code, SAMMY [45]

"uncertainties for evaluated cross sections reproduced by propagating the resonance parameter covariance matrix have historically been regarded as 'too small.' In fact, cross section uncertainties based solely on the resonance parameter covariance matrix are indeed too small."

The fact that a systematic propagation of uncertainty through the SAMMY evaluation code leads to cross section uncertainties which are "too small," ends up as additional work for the resonance evaluator to search for a non-systematic, artisanal, way to make the estimated uncertainty not "too small." Traditionally, resonance region evaluators have had to artificially inflate the uncertainty estimates to manually account for this.

Uncertainty in nuclear data has large, though yet unquantified impact on nuclear engineering applications using evaluated nuclear data. Nuclear data uncertainty can be systematically propagated as uncertainty on the prediction of computational simulations. For some applications, nuclear data can be the dominating source of uncertainty. In the particular example of the design of next generation nuclear reactors, nuclear data uncertainty has been demonstrated to result in over 100% uncertainty on the magnitude of reactivity feedback coefficients [46]. This means that there is so much uncertainty in the nuclear data that for a particular reactor design, it is not possible to predict whether a certain feedback coefficient will be positive or negative. Reactor companies are then forced to design-in additional engineering safety margins to account for this uncertainty. While, uncertainty in nuclear data will not prevent new reactors from being constructed, the uncertainty will impact the economic efficiency of those reactors through the design choices that can be made based on predictive modeling.

With the rapid raise of AI/ML research in the last decade, AI and ML algorithms have begun making their first appearances in the nuclear data pipeline. Starting with the experimental data side of the pipeline, AI/ML algorithms are being proposed as natural language processing of archived publications of experimental measurements.

Furthermore, there is currently an international effort to make the EXFOR and other online experimental nuclear databases machine-readable [47]. In the evaluation-step of the nuclear data pipeline, AI/ML applications are currently being explored to aid with computationally intensive nuclear physics calculations, such as microscopic fission theory [48]. In validation, classification and outlier detection algorithms are being proposed to identify nuclear data which needs improvement [49]. AI algorithms have been proposed to help in the design of optimized critical experiments for targeted nuclear data validation. To the best knowledge of the author, there have not been any efforts to automate the labor-intensive, manual process of resonance evaluation. Initial efforts of application of AI/ML techniques in the resonance region by Brown are noted [50], however, these efforts are not aimed at automating resonance evaluation based on experimental data.

1.3 Optimization Solver BARON

The optimization problems presented in this work were implemented via the Branch and Reduce Optimization Navigator (BARON), a software produced by The Optimization Firm [51, 52]. BARON is a branch & bound nonlinear, mixed integer global optimization solver. First, a short overview of some of the preprocessing methods will be given. The first part of the preprocessing is a feasibility based range reduction, also known as a "Poor Man's Linear Program". The following explanation goes with Figure 1.1. The solid outer box represents the initial domain of the node, where the solid inner box represents the reduced domain by considering each of the constraints individually. The dotted boundary represents solving for the limits on each coordinate with all constraints active. Lastly, the shaded region is the true feasible region. Additionally, contained in this preprocessing is lower bounding. The lower bounding algorithms are fairly complex but involve recursive arithmetic relaxation, product disaggregation relaxation, using convex/concave envelopes, and
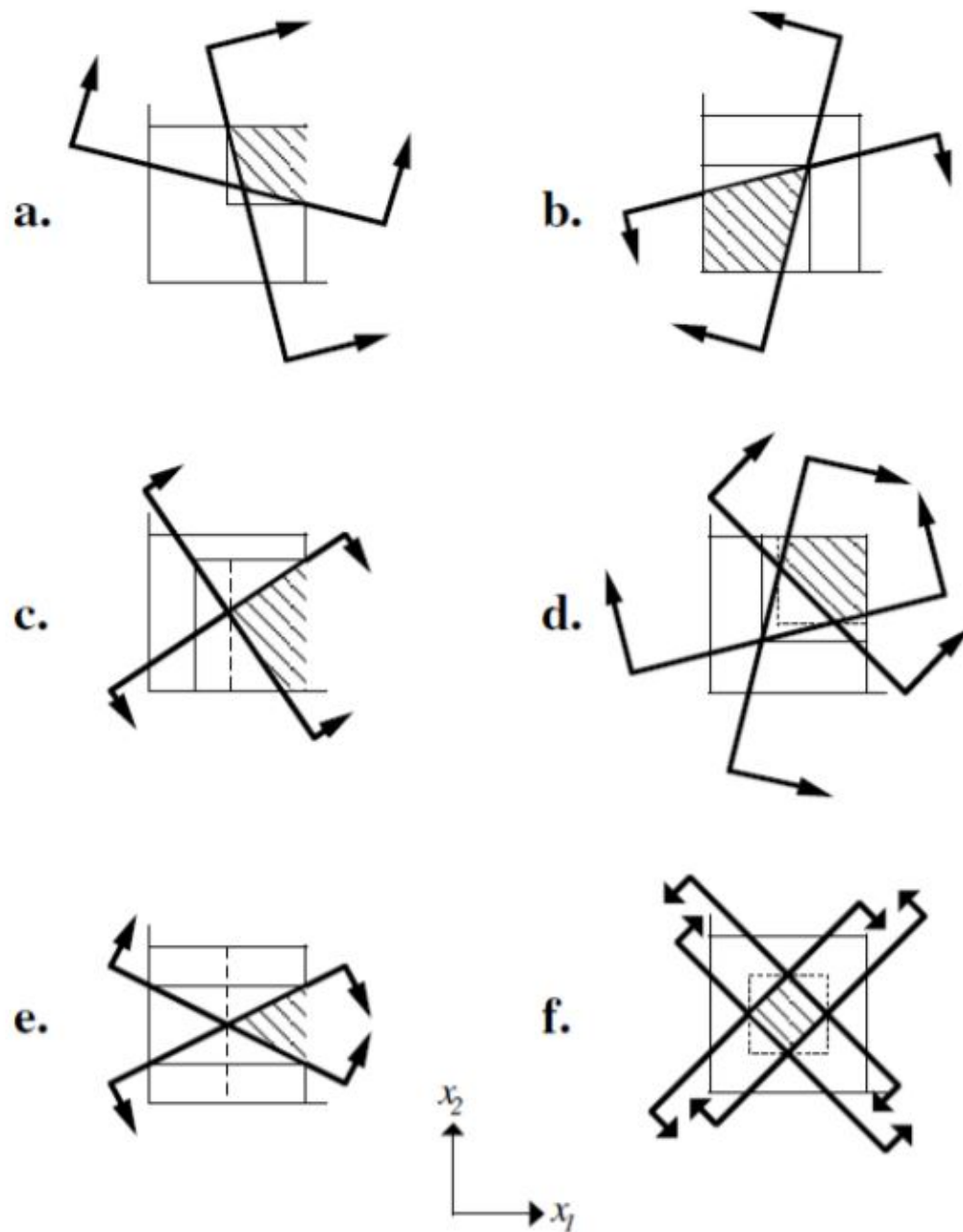
Figure 1.1: A diagram helping to explain the Poor Man's LP [2].

recursive sums and products relaxations [53]. After the lower bound is created, a linear approximation is formed in order to increase computational speed.

When allowing BARON to perform branching, BARON uses a rectangular subdivision scheme, resulting in a single variable being chosen for each branching. The variable that contributes the most to the relaxation gap is the variable that is chosen to branch upon. In the lower bounding step, copious amounts of additional variables can be introduced, so the variable selection for the branching is a non trivial process. For BARON's point selection in branching, on occasion the branching happens at the midpoint of the variable range, but otherwise is at the solution of the lower bounding problem. Lastly, for node selection, default BARON implementation is a composite value based on lower bound, violation (sum of violations of all variables), and order of creation [2].

BARON is a branch-and-reduce algorithm due to the range reduction techniques it employs, such as the feasibility-based range reduction of the "Poor Man's Linear Program" and various complex optimality-based range reduction methods [54]. In implementing branch & bound routines, enhancing the performance of the bounding procedure at each node of the search tree is achieved via stressing range reduction. During both the pre-processing and post-processing steps, the search space and reduce the relaxation gap is contracted by applying these tests to each subproblem of the search tree [54]. A majority of the reduction tests are duality based and are utilized when the relaxation is convex and solved by an algorithm that provides both the dual and primal solution of the relaxed problem [54].

1.4   Contributions

This section highlights the contributions of this work, both to industrial engineering (IE) and nuclear engineering (NE). The goal on the IE side is not to propose a new mechanic of how to solve an optimization formulation, such as improvements upon branch and bound or a new, improved iteration method for nonconvex problems.

However, the result is a way to change the problem formulation to take an intractable problem and now make it tractable. More specifically, break these intractable problems down in small, overlapping windows, where the independent windows are tractable and then stitching these solutions together to arrive at a global solution. Thus, the key novelties are the window decomposition into small, overlapping windows as well as the window stitching enforced through the cardinality agreement and parameter agreement routines. This procedure is problem agnostic, but the class of problem it is focused upon solving are nonconvex regression problems with large number of parameters (and ultimately choosing how many parameters) and problems that require a dense amount of the nonconvex function evaluations across the problem space. Furthermore, the results of this work are exemplified in its additional contributions to the NE field by offering a fast, reproducible, and automated process for nuclear data evaluation, where previously the problems in that field had to be solved by meticulous work by hand by expert evaluators. By being able to solve the 100 resonance problem for a single spin group, this model can solve for any number of resonances in a spin group.

Chapter 2

Methods

2.1 Data Generation

One of the nice advantages of the nuclear data problem is that the training set of data is "infinite" as we are able to generate as many synthetic experimental sets given the specific experimental conditions as are needed and compare our results to the true parameters and cross section. This large training size allows the building up of the fitting statistics and bias of the model's fit before having to evaluate the real experimental data. Synthetic data is required for the training as the real experimental data is limited to around three to ten data sets depending on the experiment.

For a given nuclear reaction, we have the target nucleus with mass A and angular momentum I, while the incoming particle has angular momentum i. Lastly, the spin statistical factor g_{J_α} for total angular momentum J is

$$g_{J_\alpha} = \frac{2J + 1}{(2i + 1)(2I + 1)}.\tag{2.1}$$

The value of J affects the size of the resonance while the value of ℓ affects the shape of the resonance, where the resonance are the "jumps" in the cross section data. The cross section values represent the probability of a given nuclear reaction taking place as a function of the kinetic energy of the incident neutron. The kinetic energy

dependent functions of the wave number $k_\alpha(E)$ and a relation to the center-of-mass momentum $\rho(E)$ are

$$k_\alpha(E) = \sqrt{\frac{2M_n}{\hbar}} \frac{A}{A+1} \sqrt{E}$$

$$\rho(E) = k_\alpha(E) a_c,$$

(2.2)

where M_n is the mass of a neutron, $\hbar$ is Planck's constant, and a_c is the scattering radius. The hard-sphere penetrability (penetration factor) $P(E)$ and the level shift factor $S(E)$ are also dependent upon the orbital angular momentum as given in Table 2.1. The average reduced neutron amplitude $< \gamma_{n,\ell=0}^2 >$ is derived from that orbital angular momentum's strength function S_0 to be

$$< \gamma_{n,\ell=0}^2 > = \sqrt{\frac{4\hbar < S_{\ell=0} >< D_{\ell=0} > (A+1)}{A a_c (2J+1) \sqrt{2m_n}}},$$

(2.3)

where $\hbar$ is Planck's constant, $< S_{\ell=0} >$ and $< D_{\ell=0} >$ are the average strength function and average level spacing for the given value of $\ell = 0$, m_n is the mass of a neutron, A is the atomic mass of the target nucleus, a_c is the scattering radius, and J is the total angular momentum.

A given cross section can be parameterized using the Single-Level Breit Wigner (SLBW) formalism, which is an approximation to the Reich Moore (RM) formalism. The SLBW approximation removes the complex variable dependency that is found in the intermediate variables of the RM formalism, which is required as most solvers cannot handle optimization over the not ordered field of the complex plane. In using this approximation, the cross section is examined as a summation of independent resonances, thus ignoring channel interference. Other possible approximations include the multilevel Breit-Wigner or the Adler-Adler approximations, but first the SLBW formalism is used as it is the simplest approximation to see if a more intricate approximation is required [55, 56]. Thus, the cross section can be expressed as

16

Table 2.1: The kinetic energy dependent functions of the hard-sphere penetrability $P(E)$ and the level shift factor $S(E)$ given as a function of the orbital angular momentum ℓ.

ℓ	$P_\ell(E)$	$S_\ell(E)$
0	$\rho(E)$	0
1	$\dfrac{\rho(E)^3}{\rho(E)^2+1}$	$\dfrac{-1}{1+\rho(E)^2}$
ℓ	$\dfrac{\rho(E)^2 P_{\ell-1}(E)}{(\ell-S(E)_{\ell-1})^2+P_{\ell-1}^2(E)}$	$\dfrac{\rho^2(E)(\ell-S_{\ell-1})}{(\ell-S_{\ell-1})^2+P(E)_{\ell-1}^2} - \ell$

$$\sigma(E, E_\lambda, \Gamma_\gamma, \gamma_n^2) = \frac{\pi g_{J_\alpha}}{k_\alpha(E)^2} \sum_{j=1}^{\text{\# of Resonances}} \frac{2P(E)\gamma_{n_j}^2 \Gamma_{\gamma_j}}{(E - E_{\lambda_j})^2 + \left(\frac{2P(E)\gamma_{n_j}^2 + \Gamma_{\gamma_j}}{2}\right)^2}, \qquad (2.4)$$

where $k_\alpha(E)$ is defined in (2.2), $P(E)$ is defined in Table 2.1, and g_{J_α} is defined in (2.1). The three parameters for each resonance are the resonance energy E_λ that represents the kinetic energy location of that resonance, as well as the two partial widths of the capture width Γ_γ and the squared reduced neutron amplitude γ_n^2. The total width at half maximum of the resonance is $\Gamma_\gamma + 2P(E)\gamma_n^2$. For a given isotope and spin group of that isotope, there is a well-known average neutron width, average capture width, and average kinetic energy difference between consecutive resonances (level spacing). The average level spacing $< D >$ follows the Wigner-Distribution,

$$p(D) = \frac{\pi}{2} \left(\frac{D}{< D >}\right)^2 \exp\left(-\pi \left(\frac{D}{2 < D >}\right)^2\right), \qquad (2.5)$$

where D is the distance between two given resonances [57]. The two partial widths $\Gamma_n(E)$ (where $\Gamma_n(E) = 2P(E)\gamma_n^2$) and Γ_γ follow the Porter-Thomas distribution for $x = \frac{\Gamma}{<\Gamma>}$,

$$p(x) = \frac{\nu}{2G(\frac{\nu}{2})} \left(\frac{\nu x}{2}\right)^{\frac{\nu}{2}-1} \exp\left(\frac{-\nu x}{2}\right), \qquad (2.6)$$

where G is the Gamma function and ν is the degrees of freedom, which for the neutron width $\nu = 1$ and for the capture width $\nu = \infty$ [58]. An example of a true cross section curve generated using the SLBW formalism and randomly sampled RM parameters from their distributions is shown in Figure 2.1.

Lastly, we need to know the kinetic energy density of the incident neutrons, or how many electron volts separate each of the experimental data points. This value was arbitrarily chosen to be one data point every five electron volts. With all of this experimental parameters in hand and the distributions in (2.5) and (2.6), a synthetic
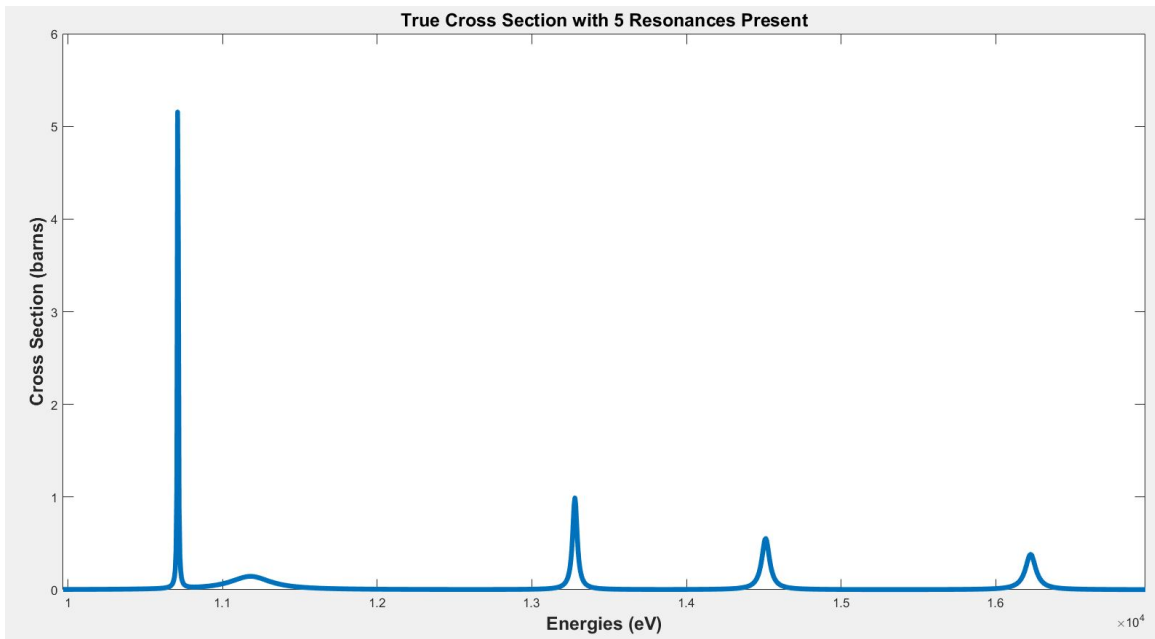
Figure 2.1: An example figure of a true cross section plot with five resonances being randomly sampled from their parameter's distributions.

true cross section curve can be generated. Again, this synthetic true cross section will closely resemble that of a true cross section of a reaction.

Due to the variety of experimental considerations such as Doppler broadening, resolution broadening, detector efficiency, and uniformity of the sample, the resultant cross section contains a varying level of noise [59]. The data generation simulates this noise using the following transformation

$$\sigma_{\text{experiment}} = N(a\sigma_{\text{true}} + b, \sqrt{a\sigma_{\text{true}} + b})$$
$$\text{std}(\sigma_{\text{experiment}}) = \frac{\sqrt{a\sigma_{\text{true}} + b}}{a}, \tag{2.7}$$

where a and b are chosen parameters to control the magnitude of the noise, $N(x, y)$ represents the normal distribution with mean x and standard deviation y, and σ is the cross section value. Figure 2.2 demonstrates the same underlying true cross section, with the difference in the plots being the amount of noise for a different set of values for a and b. Modeling the noise as normal is an approximation for the Poisson noise that is present in the experiment, however the normal distribution is a good approximation of the Poisson distribution with parameter λ for high values of λ and later parts of the methods will require an assumption of the noise being normally distributed. However, one of the disadvantages in this choice of noise model is that it will allow for negative cross section data points to be generated. Again, the cross section is the probability of a specific nuclear reaction taking place, so physically the cross section cannot be zero. However, the presented methodology is invariant to an additive translation of the background, which would take the data away from zero.

2.2   Independent Window Decomposition

Instead of attempting to solve for all of the resonances at once, the problem is decomposed into small, independent, and overlapping windows. Attempting to achieve the global optima by solving the problem as a whole is too difficult of a
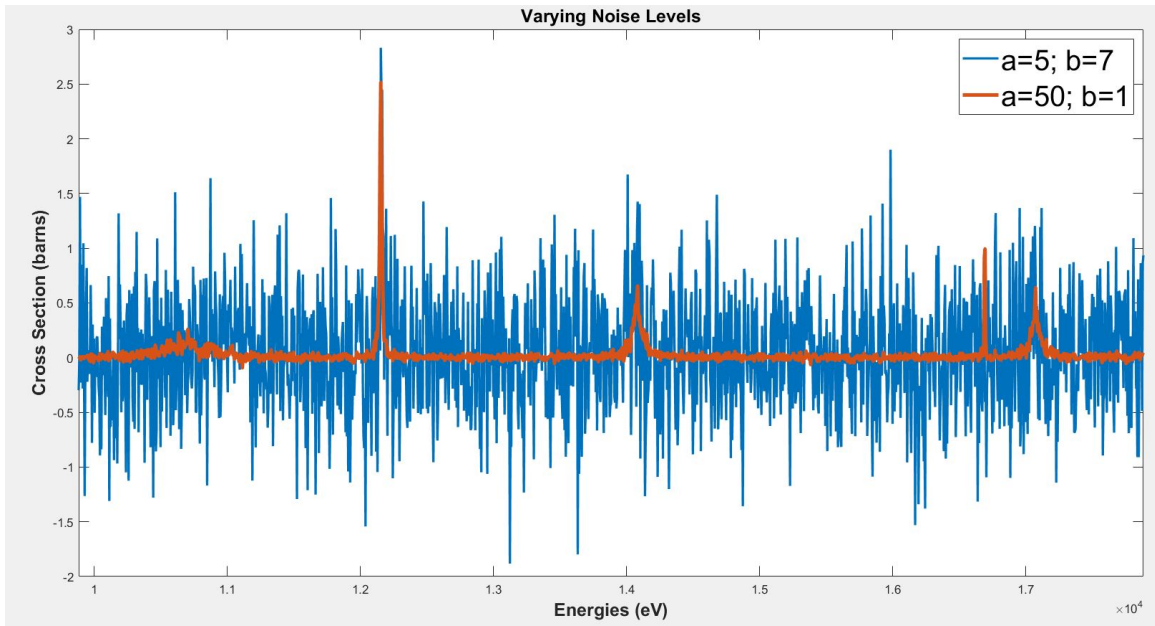
Figure 2.2: An example demonstrating how the values of *a* and *b* affect the noise present in the data. The true cross section value is the same for these two noise realizations with five resonance being present. Note this is a unique random sampling of the resonances and not the same resonances as Figure 2.1.

problem to solve, thus why such an approach is needed. This decomposition takes what would be a difficult, if not impossible, optimization formulation to solve and now proposes many quick problems to solve. Thus, the neutron's kinetic energy region is divided into M equal sized overlapping segments: $[0, m_1], [m_1 - \epsilon, m_2], \ldots, [m_{M-1} - \epsilon, m_M]$, where ϵ is the amount of overlap. By having the overlapping region between consecutive windows, it helps control for overfitting in a singular window, as the windows will later be forced to agree upon their solutions, as well as helping account for the local interactions present on either side of the overlapping region. The single window optimization formulation is

$$\min \quad \sum_{j=1}^{m} \left(\sigma_{\exp_j} - \frac{\pi g_{J_\alpha}}{k_\alpha(E_j)^2} \sum_{k=1}^{K} \frac{2P(E_j)\gamma_{n_k}^2 \Gamma_{\gamma_k}}{(E_j - E_{\lambda_k})^2 + \left(\frac{2P(E_j)\gamma_{n_k}^2 + \Gamma_{\gamma_k}}{2} \right)^2} \right)^2 \qquad \text{(2.8a)}$$

$$\text{s.t.} \quad L_1 z_k \leq \gamma_{n_k}^2 \leq U_1 z_k, \qquad\qquad\qquad\qquad\qquad k = 1, \ldots, K$$
$$\text{(2.8b)}$$

$$L_2 z_k \leq \Gamma_{\gamma_k} \leq U_2 z_k, \qquad\qquad\qquad\qquad\qquad k = 1, \ldots, K$$
$$\text{(2.8c)}$$

$$L_3 z_k \leq E_{\lambda_k} \leq U_3 z_k, \qquad\qquad\qquad\qquad\qquad k = 1, \ldots, K$$
$$\text{(2.8d)}$$

$$E_{\lambda_k} \leq E_{\lambda_{k+1}}, \qquad\qquad\qquad\qquad\qquad k = 1, \ldots, K - 1$$
$$\text{(2.8e)}$$

for a window of m kinetic energy points measuring the experimental data points of $(E_j, \sigma_{\exp_j})$. The values of L and U represent the minimum and maximum allowable values for each of the parameters. For the two partial widths, these values are found via the Porter-Thomas Distribution in (2.6) [58]. For the resonance energy, the minimum and maximum kinetic energy values of the window are used. The first three constraints of (2.8b),(2.8c), and (2.8d) are semi-continuous variable constraints implemented via the binary variable z_k which is 1 if that resonance is present or 0 if it is absent. Semi-continuous variables disallows the model from overfitting the windows by attempting to fit to noise with width values that are either too small

or too large compared to the distributions. The last constraint (2.8e) simply orders the resonance energies of the resonances found in the window, thus preventing the model from finding multiple solutions with the same objective function value due to permutations on the order of the resonances. Lastly, the value of K represents the number of resonances allowed to be in the window. Since this number is unknown, we provide the upper bound based upon the Wigner Distribution in (2.5) and allow the model to "zero out" resonances if there are too many [57]. Currently the window size is chosen to allow for the maximum window size will allowing for the problem to still be solvable and be solvable in a reasonable amount of time.

2.3 Cardinality Agreement

After each of the individual windows are solved, the results must be forced to agree in order to stitch the windows back together to achieve the best global solution to the overall problem. First, the windows must agree on the number of parameters used to fit each of the overlapping regions, which for the nuclear resonance problem, means the windows must agree on the number of resonances found in the overlap. Therefore, the individual window problem is resolved but now with a new objective function that is penalizing the squared error of the fit based upon the disagreement in the cardinality of the parameters in the overlap. The constraints remain the same as found in (2.8). The objective function for left window becomes

$$
\min \sum_{j=1}^{m} \left(\sigma_{\exp_j} - \frac{\pi g_{J_\alpha}}{k_\alpha(E_j)^2} \sum_{k=1}^{K_{\mathrm{L}}} \frac{2P(E_j)\gamma_{n_k}^2 \Gamma_{\gamma_k}}{(E_j - E_{\lambda_k})^2 + \left(\frac{2P(E_j)\gamma_{n_k}^2 + \Gamma_{\gamma_k}}{2}\right)^2} \right)^2 \tag{2.9}
$$
$$
+ \alpha(|z_{\mathrm{Overlap_{Left}}}| - \max(z_{\mathrm{Overlap_{Left}}}, z_{\mathrm{Overlap_{Right}}}))^2,
$$

while the objective function for the right window becomes

$$\min \sum_{j=1}^{m} \left(\sigma_{\exp_j} - \frac{\pi g_{J_\alpha}}{k_\alpha(E_j)^2} \sum_{k=1}^{K_{\mathrm{R}}} \frac{2P(E_j)\gamma_{n_k}^2 \Gamma_{\gamma_k}}{(E_j - E_{\lambda_k})^2 + \left(\frac{2P(E_j)\gamma_{n_k}^2 + \Gamma_{\gamma_k}}{2} \right)^2} \right)^2 \tag{2.10}$$
$$+ \alpha(|z_{\mathrm{Overlap}_{\mathrm{Right}}}| - \max(z_{\mathrm{Overlap}_{\mathrm{Right}}}, z_{\mathrm{Overlap}_{\mathrm{Right}}}))^2,$$

where

$$K_{\mathrm{L}} = \begin{cases} |z_{\mathrm{Left}}|, & \text{if } |z_{\mathrm{Overlap}_{\mathrm{Right}}}| < |z_{\mathrm{Overlap}_{\mathrm{Left}}}| \\ \min(|z_{\mathrm{Left}}| + (|z_{\mathrm{Overlap}_{\mathrm{Right}}}| - |z_{\mathrm{Overlap}_{\mathrm{Left}}}|, K), & \text{otherwise} \end{cases}$$

and

$$K_{\mathrm{R}} = \begin{cases} |z_{\mathrm{Right}}|, & \text{if } |z_{\mathrm{Overlap}_{\mathrm{Right}}}| > |z_{\mathrm{Overlap}_{\mathrm{Left}}}| \\ \min(|z_{\mathrm{Right}}| + (|z_{\mathrm{Overlap}_{\mathrm{Left}}}| - |z_{\mathrm{Overlap}_{\mathrm{Right}}}|, K), & \text{otherwise,} \end{cases}$$

and K is the maximum allowable number of resonances in a window from (2.8). The calculation of K_{R} and K_{L} is allowing for the maximum number of parameters in the overlap from each of the two windows to be the same, and BARON to determine if any of those then need to be "zeroed out". For the nuclear resonance problem, it is preferred that the data be overfit with too many parameters rather than underfit, thus the choice of the max function in (2.9) and (2.10), however this choice can be changed to the min function if it is better for a problem to be underfit. The penalty term α starts at a small value, but is dependent upon the amount of noise in the experiment and the window size, as a larger window requires a larger initial α and doubles on each iteration until the two windows agree on the number of parameters or a maximum number of iterations is reached. Additionally, the model is only allowed to update the RM values found for the resonances inside of the overlapping region (or zero out extra resonances in the overlap). The values of the RM parameters for

24

resonances found in the window from the independent window solve, but outside of that given overlapping region are not allowed to be changed at this step. Currently, a simple α updating routine is implemented, where α starts off at a very small value and doubles on every iteration until convergence of the cardinality is reached.

2.4   Parameter Value Agreement

Now that the windows agree on the number of parameters present to fit the data, they must agree on the values of those parameters. The objective function of the squared error of the predicted fit is now penalized based upon the disagreement in the variable values. Again, the constraints remain the same as the independent window formulation. Thus, the new objective function for the left window will be

$$
\min \sum_{j=1}^{m} \left(\sigma_{\exp_j} - \frac{\pi g_{J_\alpha}}{k_\alpha(E_j)^2} \sum_{k=1}^{K_{\mathrm{L}}} \frac{2P(E_j)\gamma_{n_k}^2 \Gamma_{\gamma_k}}{(E_j - E_{\lambda_k})^2 + \left(\frac{2P(E_j)\gamma_{n_k}^2 + \Gamma_{\gamma_k}}{2} \right)^2} \right)^2
$$
$$
+ \beta \left\| \frac{\vec{\theta}_{\mathrm{Overlap_{Left}}} - \vec{\bar{\theta}}_{\mathrm{Overlap}}}{\vec{\bar{\theta}}_{\mathrm{Overlap}}} \right\|_2^2 ,
\tag{2.11}
$$

and the new objective for the right window is

$$
\min \sum_{j=1}^{m} \left(\sigma_{\exp_j} - \frac{\pi g_{J_\alpha}}{k_\alpha(E_j)^2} \sum_{k=1}^{K_{\mathrm{R}}} \frac{2P(E_j)\gamma_{n_k}^2 \Gamma_{\gamma_k}}{(E_j - E_{\lambda_k})^2 + \left(\frac{2P(E_j)\gamma_{n_k}^2 + \Gamma_{\gamma_k}}{2} \right)^2} \right)^2
$$
$$
+ \beta \left\| \frac{\vec{\theta}_{\mathrm{Overlap_{Right}}} - \vec{\bar{\theta}}_{\mathrm{Overlap}}}{\vec{\bar{\theta}}_{\mathrm{Overlap}}} \right\|_2^2 ,
\tag{2.12}
$$

where K_{L} and K_{R} are the number of resonances found in the windows from the output of the cardinality agreement routine, β is the disagreement penalty term, $\vec{\theta}_{\mathrm{Overlap_{Left}}}$ is a vector of the parameter values found in the overlap from the left window, $\vec{\theta}_{\mathrm{Overlap_{Right}}}$ is a vector of the parameter values found in the overlap from the

right window, $\bar{\vec{\theta}}_{\text{Overlap}} = 1/2(\vec{\theta}_{\text{Overlap}_{\text{Left}}} + \vec{\theta}_{\text{Overlap}_{\text{Right}}})$, and $\|\cdot\|_2$ represents the Euclidean norm. The norm is taken of the relative error of each of the parameters to help account for the disparity in the magnitudes of the variables. For instance, in the nuclear data problem, the parameter's magnitudes can range from millielectron volts up to megaelectron volts. Currently, a simple β updating routine is implemented, where β starts off at a very small value and doubles on every iteration until convergence of the parameter's value is reached with maximum disagreement of a single parameter being one percent.

2.5   Uncertainty Estimation

Now that each window has agreed on the number and average values for each of the parameters, an uncertainty estimate needs to be provided. The result of the squared error regression will provide an estimate on the average RM parameters. However, since this is an experiments measurement with the cross sections data points having their respective uncertainties, these values cannot be known exactly. Thus, the model not only will provide an estimate for the average RM parameters, but provide this uncertainty on the parameters the reflect the experimental uncertainty. Additionally, an uncertainty on the predicted RM parameters and thus the predicted cross section needs to be provided for future calculations that utilized the cross section such as with reactor safety and criticality. This estimation is performed via quantile regression, which will minimize the median absolute deviation between our predicted cross section and the provided experimental data. The quantile regression formulation is

$$\min \quad \sum_{\ell=1}^{\beta} \rho_\tau \left(\frac{g_{J_\alpha}\pi}{k_\alpha(E_\ell)^2 \mathrm{std}(\sigma_{\exp_\ell})} \left(\sum_{j=1}^{\Omega} \frac{2P(E_\ell)\gamma_{n_j}^2 \Gamma_{\gamma_j}}{(E_\ell - E_{\lambda_j})^2 + \left(\frac{2P(E)\gamma_{n_j}^2 + \Gamma_{\gamma_j}}{2}\right)^2} \right) - \sigma_{\exp_\ell} \right) \right)$$

$$\text{s.t.} \qquad \gamma_{n_j}^2 \in \{N_1, M_1\}$$

$$\Gamma_{\gamma_j} \in \{N_2, M_2\}$$

$$E_{\lambda_j} \in \{N_3, M_3\},$$

$$E_{\lambda_j} \leq E_{\lambda_{j+1}} \qquad j \in \{1, \dots, \Omega - 1\}$$

$$(2.13)$$

where $k_\alpha(E)$ is defined in (2.2), $P(E)$ is defined in Table 2.1, g_{J_α} is defined in (2.1), Ω is the number of resonances in the window determined by the fit in as the output from the agreement routines, and $(E_\ell, \sigma_{\exp_\ell})$ are the β experimental data points. The check function $\rho_\tau(u)$ is defined as

$$\rho_\tau(u) = \tau \max(u, 0) + (1 - \tau) \max(-u, 0) \qquad (2.14)$$

for the τth quantile. For instance, if 10% of the errors are positive and 90% of the errors are negative, then the output would be the median of the 90th percentile. Quantile regression allows various quantile's medians to be provided for each of the RM parameters for each resonance, which then can be converted to the variances for each of those RM parameters and results in the desired covariance matrix. Adding the weighted term of the standard deviation of the experimental data helps deal with the nonlinearities present. Note that this model has to be run for each required τ value for each of the windows.

2.6 Bayesian Update

Upon completion of solving the optimization formulations for the average RM parameter values and uncertainties, Bayes' method, which is sometimes called "generalized least squares" is implemented. This method is an example of how domain

knowledge for a problem can be added as an additional part of the process in order to improve the accuracy of the overall fitting routine. There are three basic assumptions made, the prior joint pdf is a joint normal, the likelihood function is a joint normal, and the true value is a linear function of the parameters. Therefore, the posterior joint pdf is also a joint normal. Let $P = \{P_k\}$ for $k = 1$ to K to be the set of all parameters of the theoretical model to be considered with the respective covariance matrix being M. This version of Bayes' Equations is

$$M' = (M^{-1} + G^t V^{-1} G)^{-1}$$
$$P' = P + M' G^t V^{-1} (D - T),$$
(2.15)

where D is the vector of experimental data points (length L) with V being the covariance matrix of the experimental data [45]. The "sensitivity matrix" G is composed of the partial derivatives of T_n with respect to the parameters P_k, evaluated at $P = \bar{P}$:

$$G_{n,k} = \left.\frac{\partial T_n}{\partial P_k}\right|_{P=\bar{P}} \quad \text{for} \quad \begin{cases} n = 1 \text{ to } L \\ k = 1 \text{ to } K \end{cases}.$$
(2.16)

Thus, G has a dimension of $L \times K$, as T is a vector of the L data points and P is a vector of the K Reich-Moore parameters. Note that in the case that the prior covariance matrix M is infinite on the diagonal, then this version of the Bayes' Equations become the least-squares equations. The Bayesian Update results in the updated predictions for the mean values of the RM parameters P' with their corresponding updated covariance matrix M'. Performing the Bayesian update is unique enough from the above optimization formulations, as it is performed on all the resonances at once, and not window by window, hence focusing on the global fit of the problem and not numerous local fits.

Chapter 3

Results and Discussion

3.1 Implementation of the Model

The solver for this work was BARON, which a short background on how BARON works is provided in Chapter 1. BARON was implemented in MATLAB via the MATLAB-BARON interface found at https//minlp.com/matlab-baron-interface. The software was run on a Yoga 260 Laptop (ThinkPad)-Type 20FE, which has an Intel(R) Core(TM) i7-6500U CPU @2.50 GHz 2.69 GHz processor and 15.4 GB of usable RAM. The processor is an ultra low voltage dual-core system on a chip (SoC) based on the Skylake architecture that implements hyper-threading of the two CPU cores at 2.5-3.1 GHz.

3.2 Specific Isotope and Angular Momentums

While this method can be used for characterizing the nuclear resonances for any isotope, the numerical experiments presented focus upon ^{63}Cu. This isotope was chosen as it is more simple to model in terms of the types of reactions that are present, as the total cross section here is only comprised of elastic and capture reactions. In this work, the cross section that is being represented is actually only the capture cross section. However, ^{63}Cu also presents a unique challenge as it has two $\ell = 0$

29

spin groups, which is not common. Lastly, the research group has readily available access to the true experimental measurements for ^{63}Cu, as this work will compare its performance to real experimental data in future developments.

Table 3.1 gives the nuclear parameters to match an instance of a ^{63}Cu experimental that are used for the presented numerical results that would be used in Equation 2.2. Additionally, Table 3.1 shows these average parameter values for ^{63}Cu, which has an average level spacing of 722 eV, an average capture width of 0.500 eV, and a derived average squared reduced neutron amplitude of 152.8079 eV for the s-wave spin group (equation for derivation in Equation 2.3) [1]. For now, we choose $J = 1$ and the current work only focuses upon the the orbital angular momentum $\ell = 0$, which is the s-wave spin group. This selection is a subset of the spins for ^{63}Cu, which is identified to have $\ell = 0$ with $J \in \{1, 2\}$ as well as $\ell = 1$ with $J \in \{0, 1, 2, 3\}$. Hence, by replacing the values in Table 3.1 with the values for any isotope, given reaction, and correct value of ℓ (as well as choosing the right functional forms of penetration factor and level shift factor in Table 2.1 based upon ℓ), a complete set of synthetic training data can be generated. For the given parameters in Table 3.1 for this experiment of ^{63}Cu, this window size was allowing for a maximum of five resonances per window, which was choosing a value of twice the average level spacing in terms of the neutron's kinetic energy per window. The window overlap was selected to be one-half the window size, to allow for each of the experimental data points to be covered by two consecutive windows.

3.3 Case Study Demonstrations

This section walks through the methodology of the project by demonstrating the data generation, the average value and quantile regression working on a single window, and the window stitching working for two windows. The goal with showcasing these small examples is that it is much easier to visually see what is happening when only having a few resonances present.

Table 3.1: Nuclear and experimental parameters for a specific ^{63}Cu target. The constants of a_c of 6.7 fermi expressed as 10^{-12} cm and $\sqrt{\frac{2M_n}{\hbar}}$ in units of 10^{-12} cm $\sqrt{eV}^{-1}$ are given for convenience. The average values are from Mughabghab [1].

A	I	i	l	a_c	$\sqrt{\frac{2M_n}{\hbar}}$	$< D_{\ell=0} >$	$< \gamma^2_{n,\ell=0} >$	$< \Gamma_{\gamma,\ell=0} >$
62.929599	1.5	0.5	0	0.67	0.002197	722 eV	152.8079 eV	0.5 eV

3.3.1 Single Window Solutions

The merit of the single window problem in both achieving the proper number of resonances, proper predicted average parameter values, and that the quantile regression for uncertainty quantification works will be demonstrated. There are two presented example of solving the single window problem, a one resonance example and a three resonance example. These examples also give an idea of the variance of each RM parameter in their samplings from their distributions in Equations 2.5 and 2.6, The single resonance problem has parameters of $E_\lambda = 12.626$ keV, $\gamma_n^2 = 307.101$ eV, and $\Gamma_\gamma = 0.547$ eV. While the three resonance problem has parameters of $E_\lambda \in \{11.664, 12.182, 12.544\}$ keV, $\gamma_n^2 \in \{399.839, 443.780, 101.438\}$ eV, and $\Gamma_\gamma \in \{0.550, 0.513, 0.479\}$ eV. After generating the true cross sections that are seen in red in Figure 3.1, the noise model generates the blue experimental data in Figure 3.1 by allowing $a = 50$ and $b = 1$ for Equation 2.7. The errors bar shown are ± 1 standard deviations.

First, the single window optimization formulation in Equation 2.8 calculates the predicted average value for the RM parameterization that will generate the predicted average values for the cross sections. Remember, the formulation allows for up to 5 resonances to be fit in the data, but in both these cases, the model correctly identifies only 1 and 3 resonances, respectfully, and "zeros" out the RM parameters for those extra allowed resonances. Figure 3.2 demonstrates the fit from the squared error fit. In both cases the squared error of the predicted cross section fit is better than that of the true cross section fit with 0.357 versus 0.376 for the single resonance example and 0.809 versus 0.885 for the three resonance example. The slight decrease in the squared error fit is characteristic of the proposed method when the true cross section in the window is fit well.

The relative ℓ_1 norm weighted by the number of kinetic energy points is given as

$$\frac{\left\| \frac{\hat{\sigma} - \sigma_{\text{true}}}{\sigma_{\text{true}}} \right\|_1}{|\sigma_{\text{true}}|}, \tag{3.1}$$
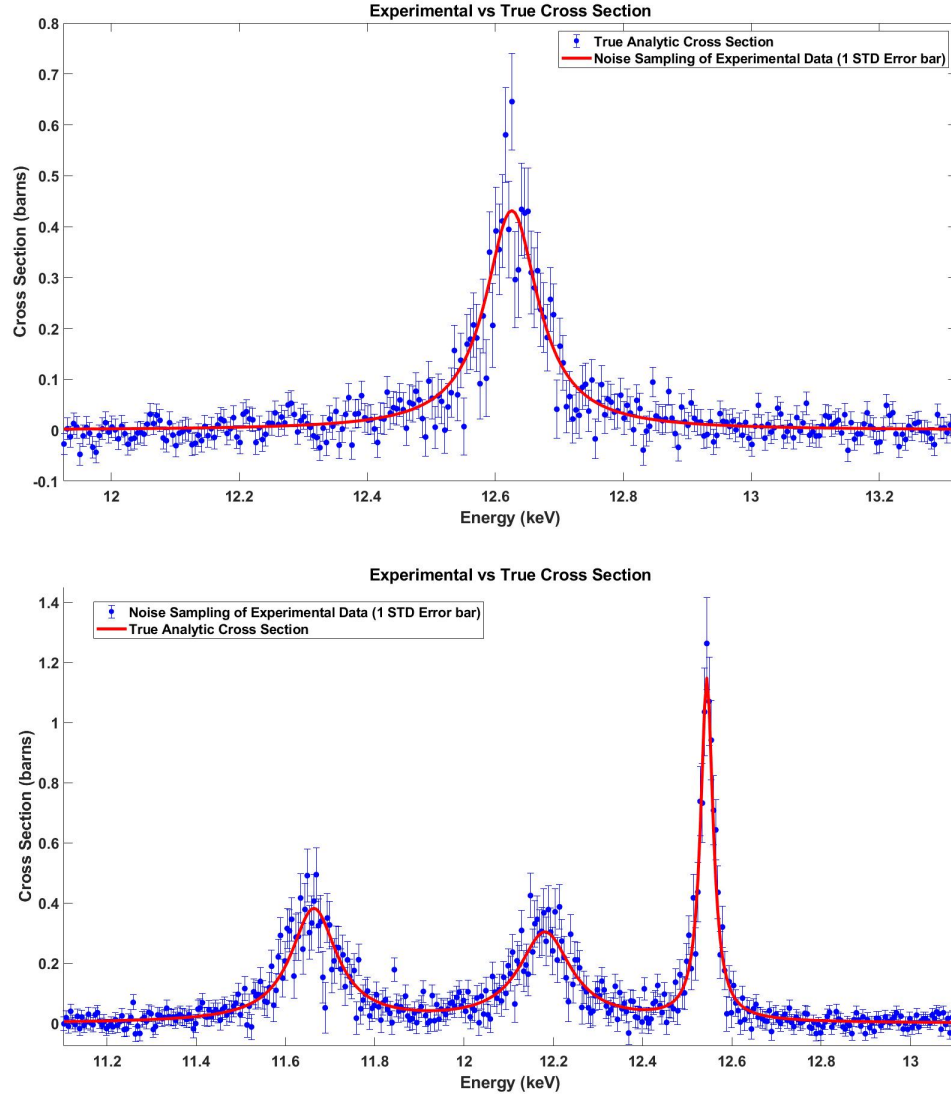
Figure 3.1: The top figure shows the true versus experimental cross section for RM parameters of $E_\lambda = 12.626$ keV, $\gamma_n^2 = 307.101$ eV, and $\Gamma_\gamma = 0.547$ eV. The bottom figure shows the same for three resonances with $E_\lambda \in \{11.664, 12.182, 12.544\}$ keV, $\gamma_n^2 \in \{399.839, 443.780, 101.438\}$ eV, and $\Gamma_\gamma \in \{0.550, 0.513, 0.479\}$ eV, The experimental data was generated with noise model parameters of $a = 50$ and $b = 1$ for both data sets.
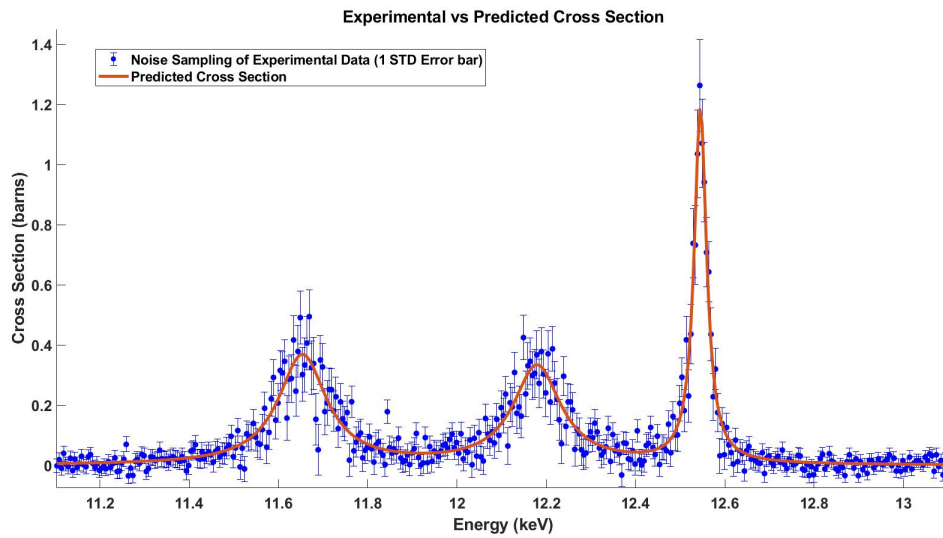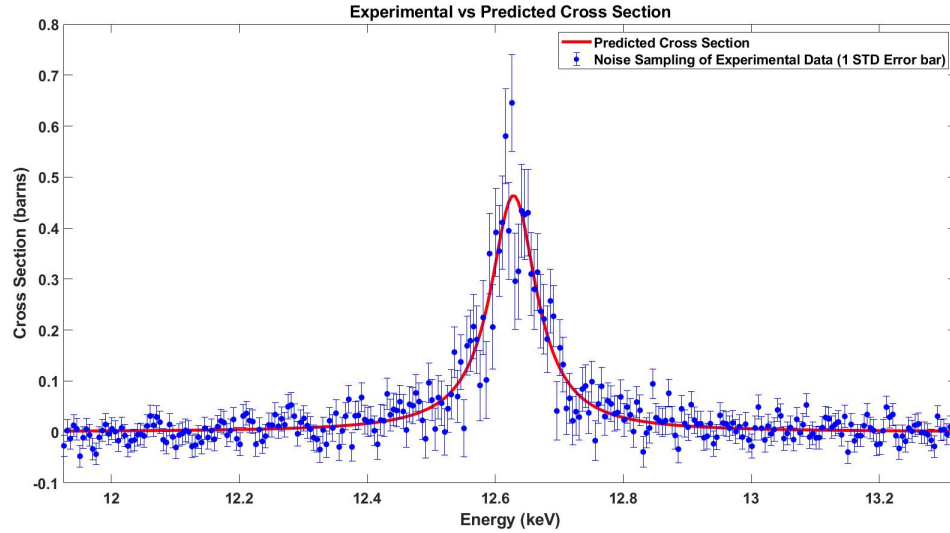
Figure 3.2: In both plots in this figure, the predicted cross section as a result of the single window optimization is in red, while the experimental data is in blue. Note that in both cases the squared error of the predicted cross section fit is better than that of the true cross section fit with 0.357 versus 0.376 for the single resonance example and 0.809 versus 0.885 for the three resonance example.

where $\hat{\sigma}$ is the predicted cross section from the model, σ_{true} is the true cross section, $|\sigma_{\text{true}}|$ is the number of data points, and $\|\cdot\|_1$ is the L_1 norm. This metric evaluated to 0.157 for the single resonance example and 0.058 for the three resonance example (the three resonance example has a larger kinetic energy span as well). Now that the average predicted values for the RM parameters (and thus the cross section) have been determined, as the model is measuring the experimental values that contain their own uncertainty, the model must be able to provide an appropriate uncertainty on its predicted values as well. Thus, the quantile regression and Bayesian update routines are run to generate this uncertainty on the predicted cross section. Figure 3.3 demonstrates how by varying the τ value, the degree to which the cross section is being under predicted versus overpredicted is changing. For instance, how a higher value of τ results in more overpredicted cross section data points. The key takeaway from the quantile regression is that the minimum squared error of the various τ's fit of the data occurs at $\tau = 0.5$, which the solution for $\tau = 0.5$ should be similar to the solution from squared error fit. Then as the τ value is either increased or decreased from $\tau = 0.5$, the squared error is a non-decreasing function. Additionally, it can be seen visually in Figure 3.3 how varying the τ results in the desired uncertainty envelope especially when examining the base of the resonance and the peak of the resonance. Further implications of the quantile regression results will be examined below in "Uncertainty Quantification" section.

3.3.2   Window Stitching Solutions

The previous subsection gave two illustrative examples demonstrating the data generation process and the single window optimization problem working for predicting the correct number of parameters for the window, a quality estimate of the parameter's average values, and a proper uncertainty estimate on each parameter. This subsection gives an example of demonstrating the window cardinality agreement routine in action. An example of the parameter value agreement is not given as most
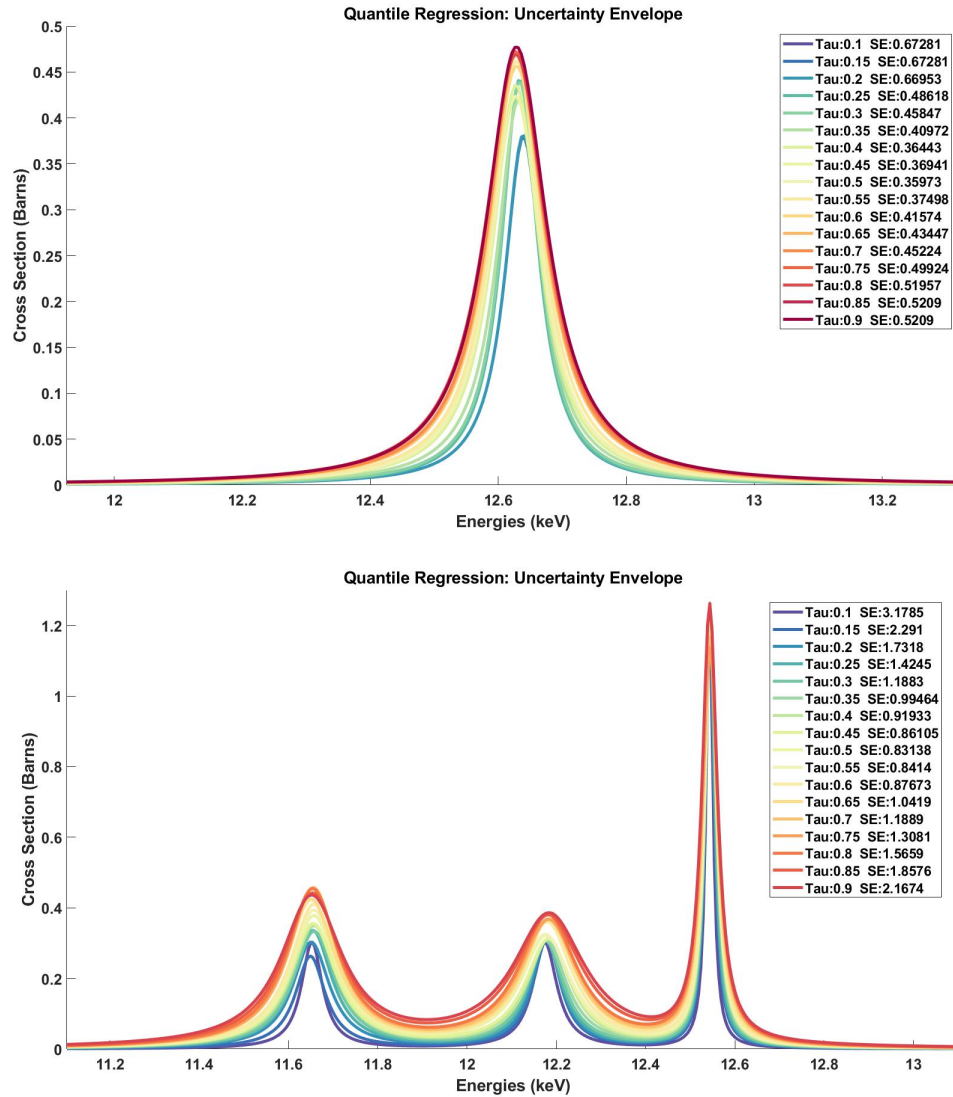
Figure 3.3: The output of quantile regression for both the single and triple resonance example. The key things to note is not to try to make out the colors of the various τ's on the graph, but notice how they form an uncertainty envelope on at the base of the resonance and at the peak. The other key aspect is the squared error values take a minimum at the $\tau = 0.5$ as expected.

of these value agreements resulted in very minimal changes to the predicted values, and so these examples are not as visual and noteworthy as the cardinality agreement.

For these two windows, on the initial solve, the first window found three resonances at $2.3790, 3.1423,$ and 3.1573 keV while the second window identified zero resonances in its window. The overlapping region here would be from 2.7398 keV to 3.4598 keV, so the first window identified two resonances in the overlap while the second window identified none. The interesting part about this example is that both windows are wrong about the overlap. The true resonances covering the span of these two windows have energies of 2.3790 keV and 3.3157 keV. Therefore, there is only one true resonance in the overlap, so the first window overpredicted by saying there were two resonances and the second underpredicted by saying there were none. After one iteration of the cardinality agreement routine, the two windows agreed on there only being a single resonance in the overlapping region. Now the first window correctly only had resonances at 2.3790 keV and 3.1573 keV while the second window had its resonance at 3.1573 keV. Note that how the two windows exactly matched on the values for that resonance simply by making them agree on the number of resonances. Another observation about this example is that the cardinality agreement routine resulted in the squared error fit of both of the windows decreasing. The squared error of the first window decreased from 6.943 to 5.813 (while now having two instead of three resonances), and the second window decreased from 240.474 to 0.495. Figure 3.4 and 3.5 illustrates visually how the results for the two windows change as a before and after of the one iteration of the cardinality agreement for this overlapping region for the left and right windows, respectfully.

3.4 Resonance Ladder Performance

Across ten unique samplings of the 50 resonance problem, attempting to solve for all 50 resonances in a single optimization formulation resulted in BARON erroring out due to "insufficient memory for data structures" after approximately 20 minutes

Figure 3.4: The example of showing how the left window's number of resonances in the overlapping regions changes from overpredicting with two resonances to being correct with one resonance. This window initially tried to fit two resonances to the rightmost resonance in this window. As a result of the cardinality agreement, the squared error for this window went down from 6.943 to 5.813. The black vertical lines represent the resonance energies of the predicted resonances.
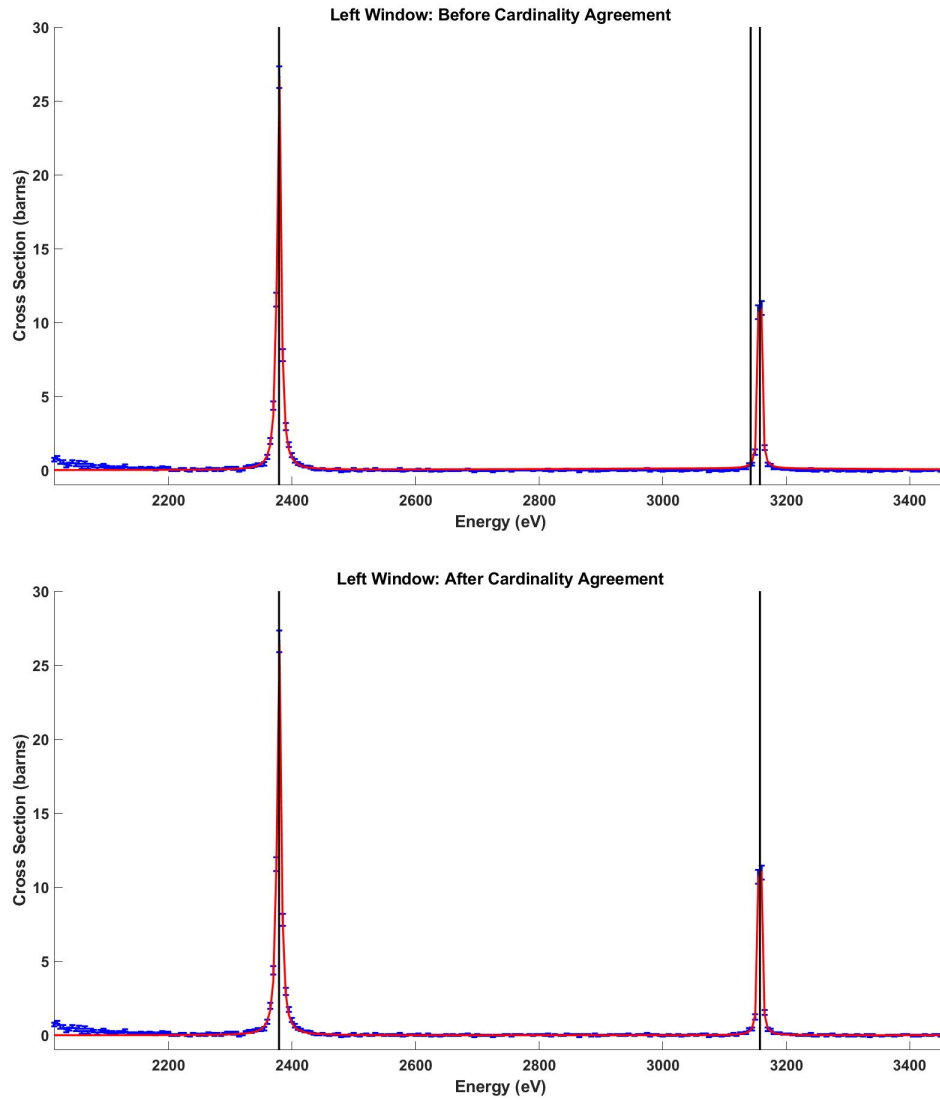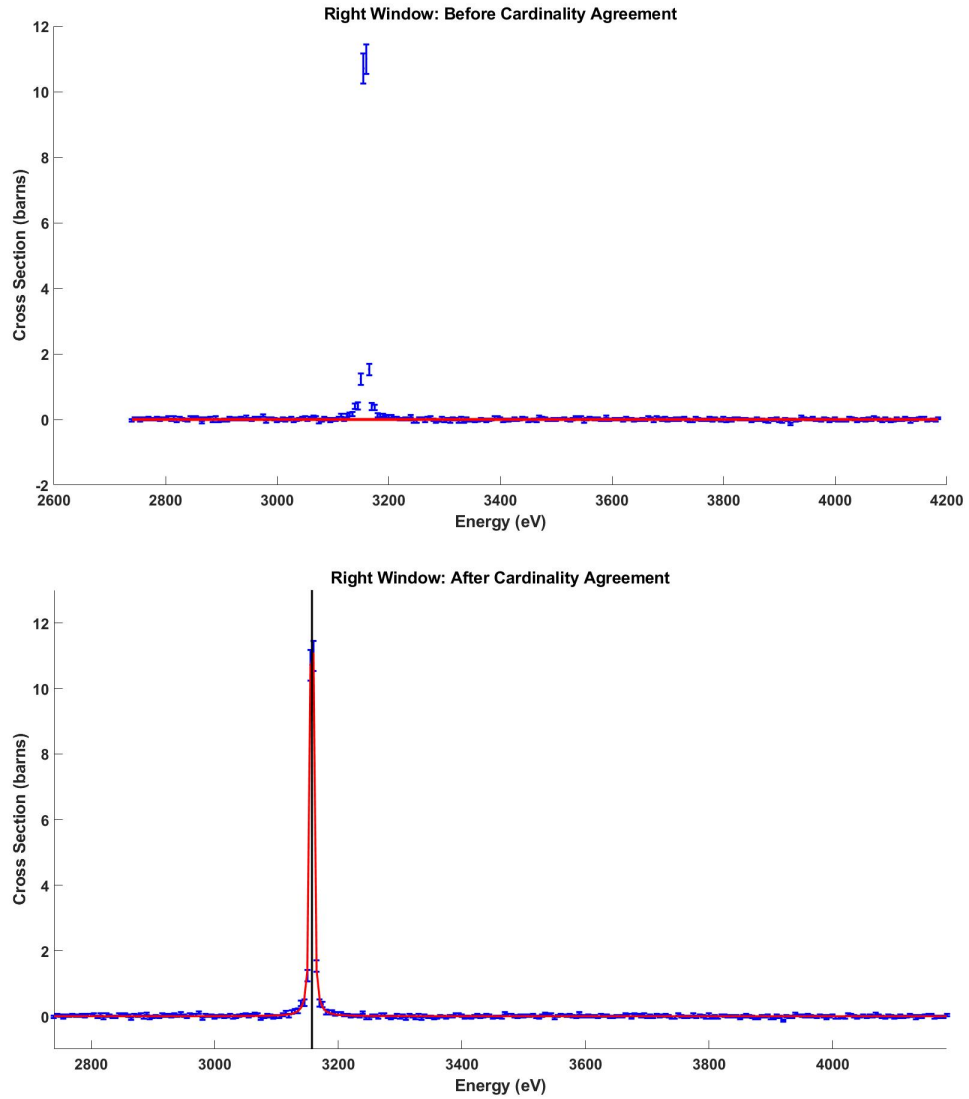
38

Figure 3.5: The example of showing how the right window's number of resonances in the overlapping regions changes from underpredicting with zero resonances to being correct with one resonance. As a result of the cardinality agreement, the squared error for this window went down from 240.474 to 0.495. The black vertical lines represent the resonance energies of the predicted resonances.

in all ten cases. Thus, producing no solution or even estimate of a solution to this problem. However, below it will be demonstrated how the windowed approach can solve the 100 resonance problem on the order of hours and arrive at a good solution. First, note the unique structure of the independent small windows. The 100 resonance problem is composed of approximately 100 windows (due to the choice in window size and overlap), therefore the 100 resonance problem is not the ability to solve for that group of 100 resonances, but the ability of the algorithm to solve 100 unique windows and combine those solutions together. Thus, the 100 window problem is the same as 5 realizations of the 20 window problem, or 10 realizations of the 10 window problem, and so on in terms of the RM parameter sampling. Remember, the RM parameters are sampled independently from their distributions in Equations 2.5 and 2.6, so these 100 windows are 100 unique windows. However, solving the 100 resonance problem is actually a more illustrative demonstration of the results of this method compared to say 10 realizations of the 10 window problem or 5 realizations of the 20 window problem. It is more illustrative because the 100 resonance ladder covers a much larger kinetic energy range, which is important as shape of the resonance also has an dependence on the kinetic energy of the incident neutron. Therefore, a larger kinetic energy range gives the model a larger range of resonance sizes it has to be able to fit over.

Solving for a ladder of any number of resonances is solely dependent on how well the single window problem can be solved. Note that if the single window problem can be solved really well, then there will no longer be a need to have overlapping windows as all of the overlapping regions will already agree. However, an iterative approach will be still be needed to handle the possible effects of one window on the next (as the windows are solved initially as though they were independent even though they are not), such as having the tails of resonance in one window run over into the next window.

This argument is further exemplified in Figure 3.6, where the squared errors for the true versus experimental cross section (in blue) and the predicted versus
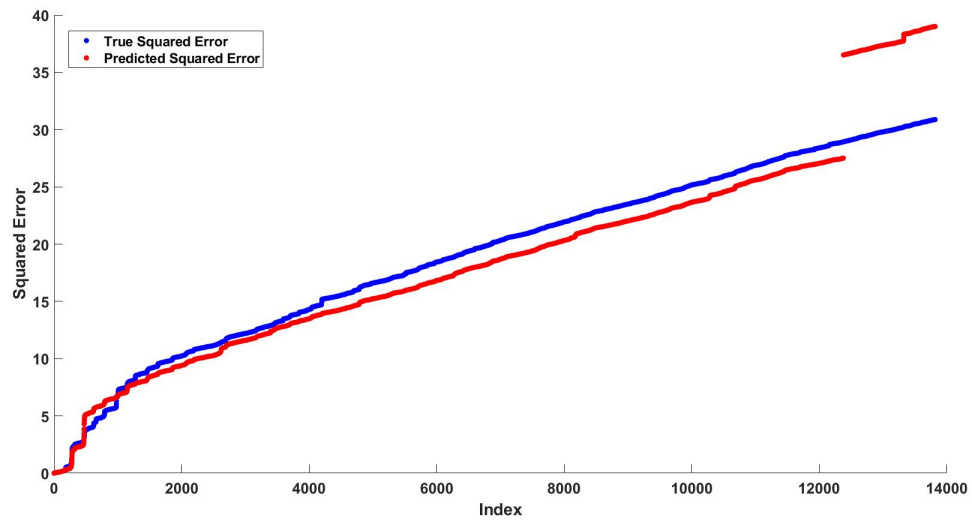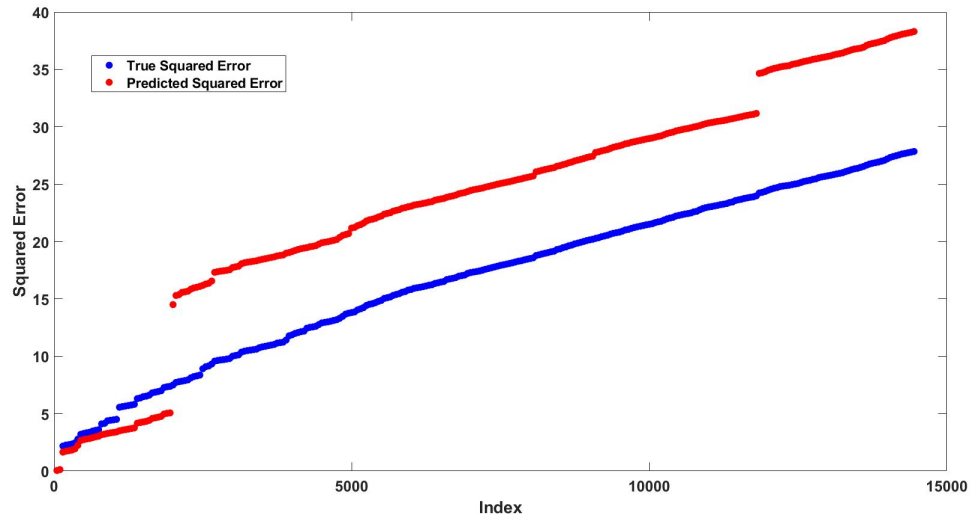
Figure 3.6: The squared error expressed as a function of neutron's kinetic energy in the resonance ladder. Each of these errors were calculated pointwise, so every point represents five electron volts. Each of the two plots are across unique 100 resonance ladders.

experimental cross section (in red) are shown. As the experimental data is composed of incident neutron kinetic energies every five electron volts, these metrics are calculated pointwise with each index representing that five electron volt spacing. For the first 100 resonance ladder, the predicted squared error starts off much lower than the true squared error, then has two distinct points where the predicted squared error has significant increases in value. But other than those two jumps, the predicted squared error appears to climb at the same rate as the true squared error, thus demonstrating the ability of the model to correctly identify the true cross section. Similarly, for the second 100 resonance ladder, there are two distinct jumps in the predicted squared error, otherwise the predicted and true squared error again appear to increase at the same rate. Those squared error jumps occur as the model did not identify a handful of very small resonances-resonances that are composed of only a singular data point. Due to the nature of the resonance formalism, these very narrow widths also are very tall, thus having a large contribution to the squared error if missed. However, it is not ideal for the model to necessarily fit to these one point resonances as that one point could easily be an experimental error or anomaly in the data collection when examining real data, so fitting these singular points needs to be performed by the expert evaluator who can better determine the holistic picture of the data set.

Now, Figure 3.7 would not be available for real experimental data, but demonstrates a very similar behavior where the ability of the model to characterize the true cross section is constant across a majority of the resonance ladder except for a few key places where there is a big jump in this error. This behavior demonstrates that only a few places in the predicted cross section curve need the attention of the nuclear evaluator in order to drastically decrease the fit on the true cross section. As a specific example of BARON handling these windowed problems well, Figure 3.8 demonstrates the ability to correctly identify both true resonances present in the experimental data set, even though to the human evaluator this small window would be very difficult to properly characterize. Thus, this is a prime example of BARON

42
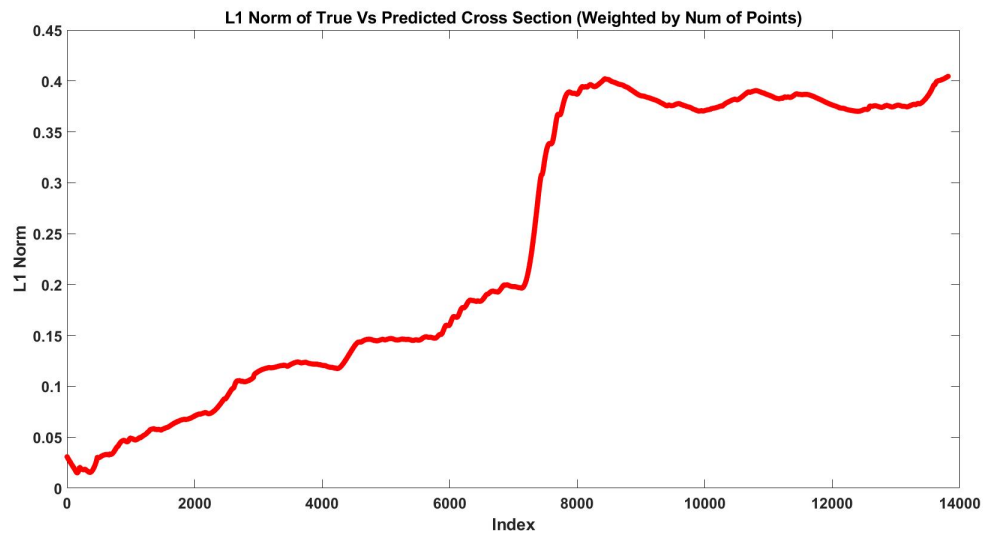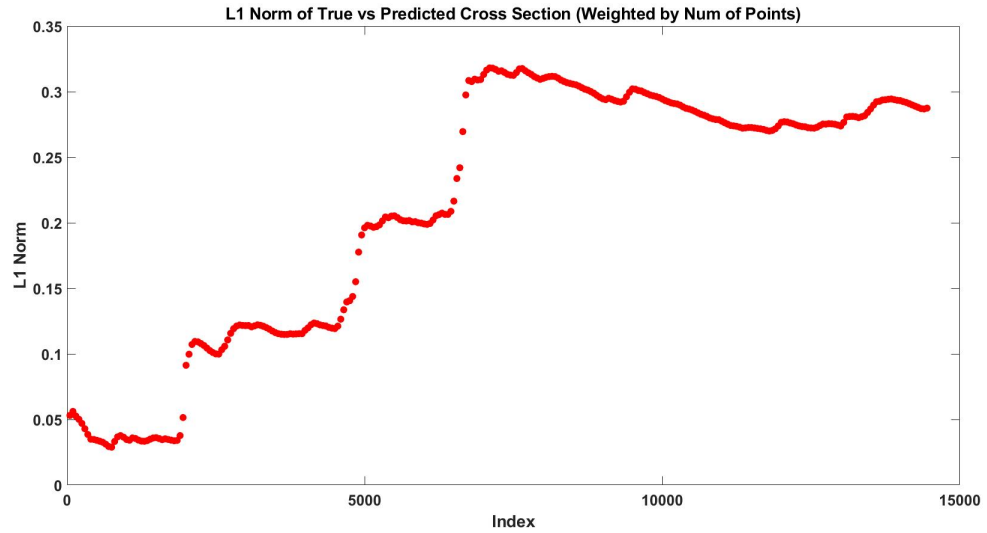
Figure 3.7: The weighted L_1 norm expressed as a function of neutron's kinetic energy in the resonance ladder. Each of these errors were calculated pointwise, so every point represents five electron volts. Each of the two plots are across unique 100 resonance ladders.
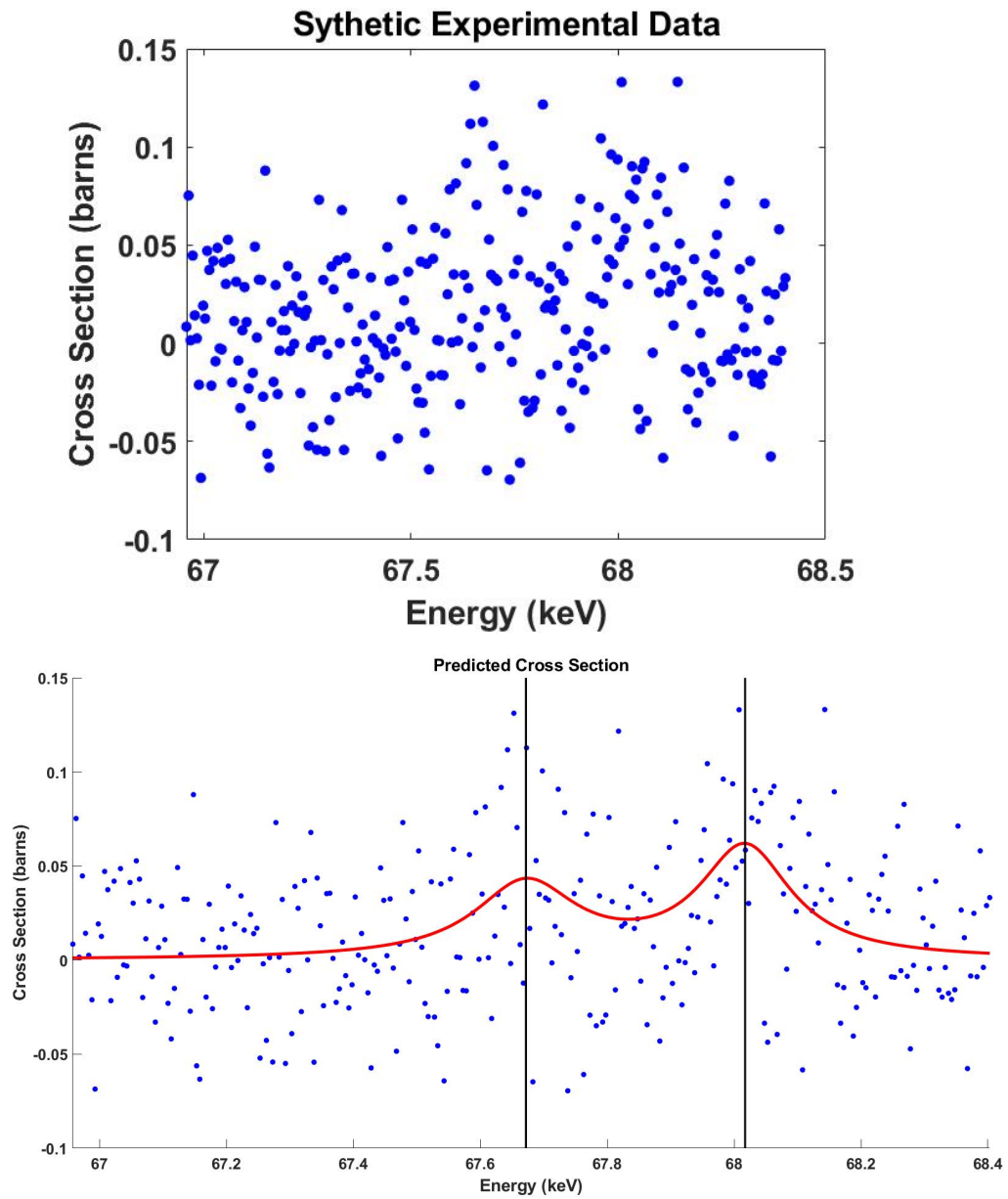
Figure 3.8: An example of resonance cluster where the experimental data has little apparent structure, yet the formulation was correctly able to identify both resonances present. The black vertical lines represent the resonance energy of the identified resonances.

being able to handle the one extreme case of resonance clusters close together when the resonances are small (the other extreme being properly characterizing big, isolated resonances, which would be easy to do by an evaluator and BARON does well with too). This figure also gives a visual example of the difficulty of resonance clusters as the resonances bleed into one another, thus making it harder to characterize their widths.

3.5 Uncertainty Quantification

As the model is predicting the RM parameterization of the cross section via experimental measurements, there are subsequent uncertainties with the model's predicted parameter values. Thus, in predicting these values, an appropriate uncertainty needs to be calculated for them. The squared error fit of the independent windows solution gives the prediction of the average value of the cross section, and goal of the quantile regression calculations is to provide an experimentally appropriate error envelope around this average predicted value. Figure 3.9 shows the relationship between the chosen value of τ and the amount of cross section data points that are overpredicted for the experimental cross section (in red) and the true cross section (in blue).

Upon visual inspection of the figure, both sets of data points have an apparent functional form to them. Upon further investigation, the true cross section overprediction curve was determined to follow the logistic function, which is the cumulative distribution function (CDF) of the logistic distribution with the form

$$f(x; \mu, s) = \frac{1}{1 + \exp\left(\frac{-(x-\mu)}{s}\right)} = \frac{1}{2} + \frac{1}{2}\tanh\left(\frac{x-\mu}{2s}\right), \tag{3.2}$$

where x is the random variable, μ is the mean, and s is a scale parameter proportional to the standard deviation. This fit for the single resonance example has a mean of $\mu = 0.5272$ and scaling parameter of $s = 0.01018$, with the corresponding fitting
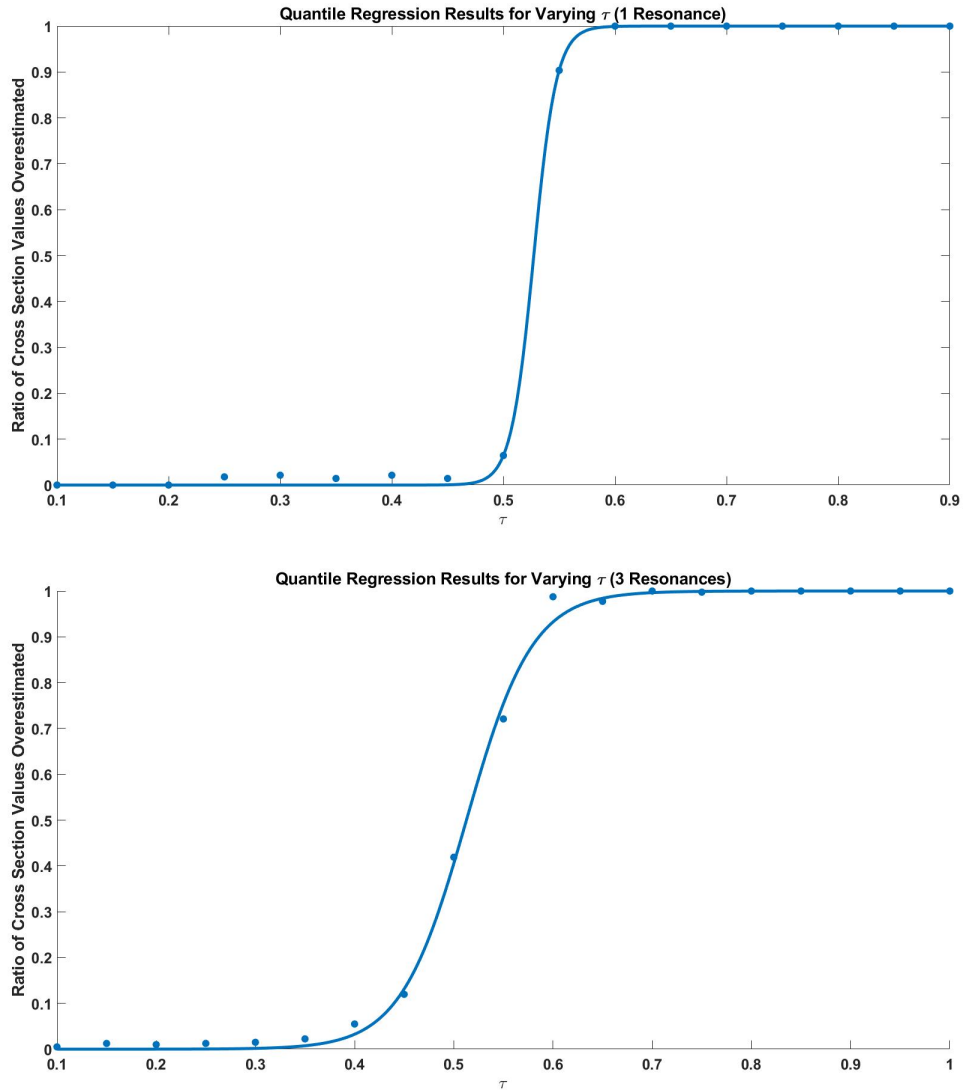
Figure 3.9: The effect of varying τ with the amount of true cross section data points overpredicted. The logistic CDF was fit to the single resonance curve with $\mu = 0.5272$ and $s = 0.01018$ and to the three resonance curve with $\mu = 0.5128$ and $s = 0.03325$. The adjusted R^2 was above 0.99 for both fits.

statistics of $R^2 = 0.9996$ and root mean squared error of 0.0104. While the three resonance example has a mean of $\mu = 0.5128$ and scaling parameter of $s = 0.03325$, with the corresponding fitting statistics of $R^2 = 0.9985$ and root mean squared error of 0.0186. Therefore, it is evident that the quantile regression results on providing the uncertainty on the true cross section values follow a well known and well characterized probability distribution. The probability density function for the logistic distribution is

$$f(x; \mu, s) = \frac{\exp\left(\frac{-(x-\mu)}{s}\right)}{s\left(1 + \exp\left(\frac{-(x-\mu)}{s}\right)\right)^2} = \frac{1}{4s}\operatorname{sech}^2\left(\frac{x-\mu}{2s}\right), \tag{3.3}$$

again where x is the random variable, μ is the mean, and s is a scale parameter proportional to the standard deviation. The scale factor s relates to the variance as $\mathrm{Var} = \frac{s^2\pi^3}{3}$. The connection of this result to quantile regression can be made as the inverse cumulative distribution function of the logistic distribution is the quantile function, which is a generalization of the logit function. The logistic distribution closely resembles the normal distribution in the shape of the distribution, but has higher kurtosis (heavier tails) and is a special case of the Tukey lambda distribution. The logistic distribution appears in many applications ranging from the direct machine learning application of logistic regression (modeling categorical dependent variables) to in physics where the logistic probability density function has the same functional form as the derivative of the Fermi function [60].

Ultimately determining the functional form of the ratio of overpredictions of the experimental cross section is not important, as the model is concerned with determining an estimate and uncertainty on the true underlying cross section. It is worthwhile to note that the true cross section uncertainty can be characterized well, even though the optimization formulation is fitting to the experimental data points. Having the logistic curve fit that data as it does, enable the nuclear data evaluator to not need to calculate the quantile regression procedure for all of the presented τ values (which all have to be run independently, being computational expensive), but allows him/her to achieve the entire functional fit by simply running a two values of

τ (need two values to find the two unknowns of the CDF: the mean μ and scaling parameter s) and then selecting what size uncertainty envelope he/she ultimately wants. Again, the key takeaway from the quantile regression results is that a well known probability distribution is established to characterize the uncertainty on the true cross section, which this characterization happens by only being able to see the experimental data.

Chapter 4

Conclusions and Recommendations

4.1 Future Model Improvements

While the model has appeared to work well, there are a few key directions for future work to improve the methodology to obtain even better results.

4.1.1 Penalty Function

The model's solution to optimization problem is greatly affected by the shape of the penalty function. Transformations such as a positive number scaling of the penalty function does not affect the solution, since this simply scales the objective function [3]. Currently, the quadratic penalty function $\phi(u) = u^2$ is implemented, which yields the Euclidean norm approximation. One issue with the quadratic penalty function is that it gives a heavy weighting to large residuals, making it sensitive to outliers in the data. For example, when comparing the L_1-norm and L_2-norm, for small residuals u, $\phi_1(u) \gg \phi_2(u)$ and for large residuals $\phi_1(u) \ll \phi_2(u)$. Therefore, the residual distribution for the L_1 approximation will have more very small and zero residuals, while the L_2 approximation will have less large residuals. This residual distribution difference is demonstrated in Figure 4.1 for a penalty function of $\phi(u) = \|u\|^p$. There is little incentive for the L_2 penalty to drive small residuals even smaller.

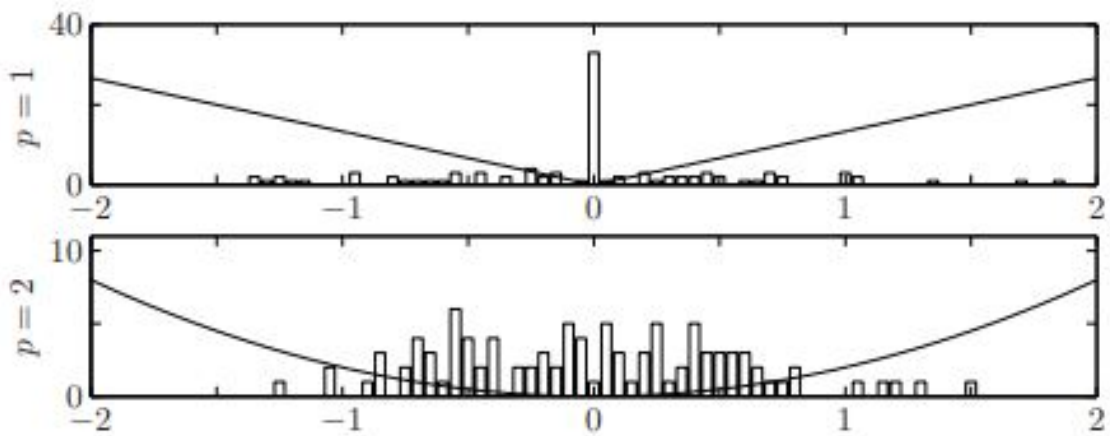**Figure 4.1:** The residual distribution difference when comparing the L_1 approximation to the L_2 approximation for the penalty function of $\phi(u) = \|u\|^p$. The $p = 1$ residual distribution has more zero and very small residuals while having many more relatively large residuals, while the $p = 2$ has less large valued residuals and many modest valued residuals. Figure is from Boyd [3].

This behavior of the penalty function's effect on the solution to the optimization problem is especially important for the nuclear resonance evaluation, as the model is having to fit noisy, experimental data with the goal of extracting the underlying true cross section. Thus, in evaluating the fits of presented by BARON, there is this similar trend of having few large residuals, but then having a large number of "modest" residuals. It is evident that the L_2 penalty function is doing a good job of getting a close solution to the true cross section, but there is still room to improve upon that solution, and that improvement being limited not by the noise in the data or by the formulation, but by the drawbacks in the L_2 objective. For instance, take a resonance that is composed of ten data points. If one of those ten points is an outlier compared to the other nine in its representation of the true cross section, the current formulation is going to give a high weighting to not allowing that outlier's residual to be large and subsequently have a slightly worse fit on the other nine points, and thus the performance of measuring the true underlying cross section is slightly decreased. There are two proposed procedures for improving these solutions, one being take the L_2 solution as the starting point for a new optimization problem, and only allow the RM parameters to be slightly changed from that solution, but now use a different penalty function that penalizes the small residuals more in hopes that the new penalty shifts more of the small residuals towards zero and by bounding the RM parameters around the L_2 solution, that will disallow any large residuals from forming. The downside to this method is now a new optimization problem has to be solved for each of the windows, which is computationally expensive. However, this problem may be able to be accounted for in the initial solve by slightly modifying the penalty function by using penalty function approximation, such as

$$\phi(u) = \begin{cases} u^2 & |u| \leq M \\ M^2 & |u| > M. \end{cases} \tag{4.1}$$

Penalty function approximation preserves the Euclidean norm for residual values that are less than M, however it places a set weight upon residuals larger than M, thus

ignoring these large residuals. Note that this penalty function is not convex, however the formulation is already dealing with a nonconvex optimization problem. But there are other choices for a penalty function to help with outliers, while keeping the penalty function convex would be robust penalty functions. The least sensitive ones are those for where $\phi(u)$ grows linearly such as the L$_1$ approximation or using the robust least squares (Huber penalty function) of

$$\phi_{\text{hub}}(u) = \begin{cases} u^2 & |u| \leq M \\ M(2|u| - M) & |u| > M, \end{cases} \tag{4.2}$$

which acts like least-squares for small residuals and then like an L$_1$ for large residuals [3]. The convex function closest to the outlier penalty function is the Huber penalty function [3]. Note that the simple case of only using the L$_2$ penalty function gave a good solution to the resonance fitting problem, and future work can refine this penalty function to improve even more upon the solution.

4.1.2 Window Size

Currently, the model has a static window size, which for this experimental data was twice the average level spacing. This value was determined heuristically to be the best maximum window size as it would allow for up to five resonances to be present in a window, and any more resonances being allowed saw a drastic decrease in the tractability of the problem. However, not all windows are created the same in the resonance ladder. Some windows had a single, large resonance isolated by itself while other windows had "resonance clusters" being present where there were multiple resonances located within a small kinetic energy range and thus causing direct interference with one another. These examples are to show that the complexity of each window is far from being constant. Figure 4.2 demonstrates that although the level spacing stays constant throughout the resonance ladder, the total widths of the resonances increase as the kinetic energy of the incident neutron increases (a direct
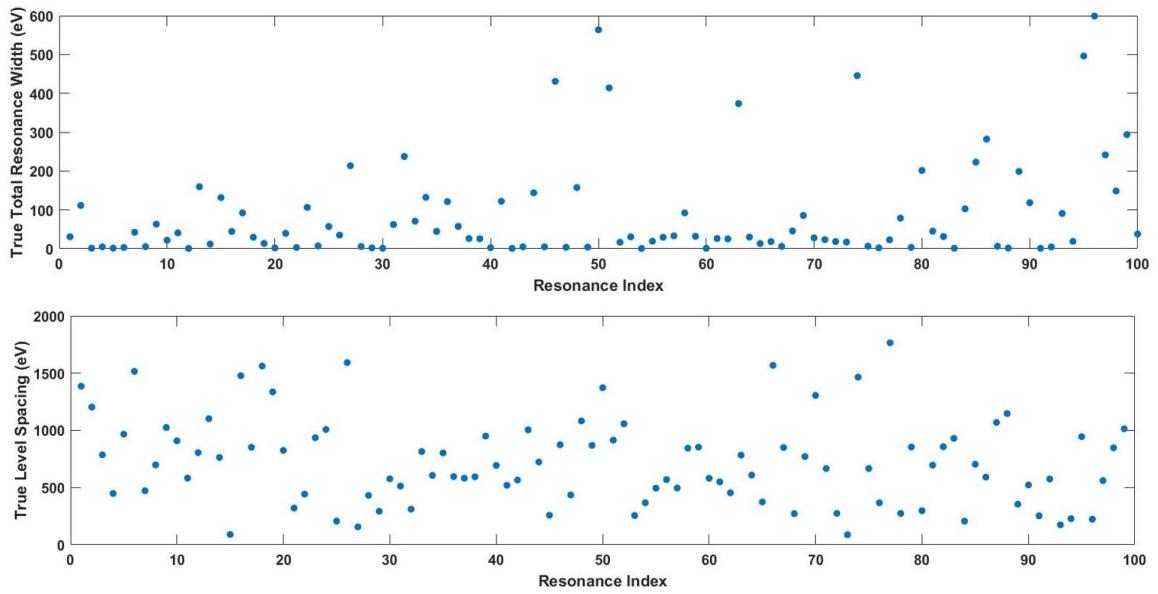
Figure 4.2: The effects of increasing incident neutron's kinetic energy of the resonance ladder has on the true total widths and the true level spacing of the resonances in that ladder. The level spacing stays constant while the true total widths increase, as the kinetic energy of the resonance increases.

effect of the penetrability factor), thus the resonance clusters become more prevalent as the neutron's kinetic energy increases.

Due to the differing complexity of these windows, BARON is able to obtain a similar squared error fit in both windows (comparing the predicted to true cross section), but that result does not have a similar error when comparing the fit to the true underlying cross section. That disparity is due to a good fit (having a lot of small residuals, but not a lot of very small or zero residuals) gets the solution close in the complex window, but to fully characterize that underlying true cross section, the residuals need to be very small. On the other hand, having very small residuals is not required in the single isolated resonance window, as those small residuals are good enough to characterize that easily identifiable resonance. Thus, the requirements for obtaining a high quality characterization of the true cross section can vary greatly from window to window. So one solution is modifying the penalty functions based upon this complexity as previously mentioned or have a dynamic window size selection routine. The dynamic window size would allow the model to have a very large window when they are few resonances and/or clearly isolated resonances, and then have a much smaller window when dealing with the complex resonance clusters. This dynamic window size should decrease the overall runtime of the method as it should create less windows as a whole as a large number of windows right now have one or two resonances so a majority of those could be collapsed into a single window, and having less windows and windows with better solutions will require less iterations of the stitching routines on the back end of the algorithm.

On a similar note to dealing with resonance clusters, this is a problem with manual nuclear data evaluation as well. Eventually the resolved resonance region will reach an incident neutron kinetic energy where the cross section fluctuations and resonances still exist but are no longer able to be resolved due to experimental resolution. Thus, that is an important future question to answer in how this model's resolved resonance region compares to that of manual evaluation. As if the neutron's kinetic energy where

the unresolved region starts for this model is less than that of an evaluator, this method still has utility as the evaluator can use it for the model's resolved region, get a good starting guess for those resonances, and finish out the rest of the calculations by hand. Or the model's unresolved region could start at a higher kinetic energy of the incident neutron, and thus be able to give a more accurate characterization of the cross section as a whole and and thus improve the nuclear understanding of these cross sections. Currently, estimates are formed in the unresolved resonance region based upon average RM values found in the resolved region and propagating those into the unresolved region by effective cross sections from probability tables as a consideration of self-shielding effects or via Monte Carlo methods [61, 62].

4.2 Nuclear Resonance Limitations

While the focus of this work is upon the methodology of being able to transform large, nonconvex, intractable problems, into small, tractable ones, there are a few points of the specific points of its application to the nuclear resonance problem that should be delineated.

4.2.1 Accuracy of Synthetic Data

The first is that this application hinges upon being able to generate a large training set of synthetic experimental data to build the fitting statistics of the model before its application to real experimental data (again a benefit as real experimental data sets for a given reaction are very limited in size). The key concern here is what if the real experimental data is not accurately modeled by the synthetic data. Then all of these fitting statistics that have been built up for a given experiment would be useless. This question is being answered by a fellow student in the NE research group, where he is demonstrating that the entire set of real experimental data is a subset of the synthetic experimental data. If his work is successful in not being able

to demonstrate that a real experimental data set is distinguishable from a group of synthetic data sets, then this problem is no longer a concern. However, if he can demonstrate there is a distinguishable difference that ultimately cannot be accounted for in fixing the synthetic data generation, then this project's methodology will have to be slightly shifted in only being able to train upon real experimental data sets, and will lose its ability to generate the fitting statistics, but will still (in theory) be able to fit the real data sets well.

4.2.2 Missing and Extra Resonances

The other key limitation is the problem of dealing with extra or missing resonances generated from BARON's fitting of the experimental data. While the goal of the project is to best extract the true underlying cross section, and not necessarily perfectly find every resonance, by being able to remove extra resonances or go back and find missing ones, this could potentially increase how well we are able to calculate the true cross section. However, the benefit of having the synthetic training data is that there is the ability to characterize what types of resonances fall into the two categories of being missed or extra. For instance, nearly all of the missed resonances fall below a "minimum detectable limit". This minimum detectable limit was taken to be half the distance between the kinetic energy points of the incident neutron, so a resonance of this width would only be accounted for by a single data point. By demonstrating the majority of missed resonances fall into this category, the model can be trusted to a higher degree by the nuclear evaluator as an evaluator is unlikely to identify a single data point spike as a resonance either, and would likely call it an anomaly in the experiment or a mistake in the data record. On a similar note, almost all of the extra resonances are in the resonance cluster areas, where the model could fit the cluster decently well, but adding an extra resonance made that cluster's fit even better. Again, the nuclear evaluator can then see this behavior, update the

values for the RM parameters for the resonances that should be in that cluster and delete any erroneous resonances that were added.

It is worth noting the current optimization formulation does not incorporate the probability distributions, which easily could have been introduced to the objective function. However, this incorporation was not performed in order to try to fit with the simplest objective function first to see if a more intricate one was required and because the work for the missing and extra resonances is a direct implementation of the Wigner and Porter-Thomas distributions. The work by this student is using the distributions to calculate the probability of each found resonance actually being an extra resonance and calculating what are the most likely regions where a resonance was missed. His algorithm will then be added to the end of the routines presented in this work, and will inform BARON of where to attempt to add or remove a resonance and see how that ultimately affects the fit. These two algorithms will be back and forth in their iterations until an improved fit is no longer achieved.

4.3   Final Remarks

An additional benefit of having the synthetic training data is the ability to compare the predicted RM parameters and cross section to the true values. However, with real data, this is obviously not possible. One of the key remaining questions is how good actually are this method's good fits. All of these metrics are comparing the model's output to the true values, such as finding 95 resonances when they are 100 resonances present, or comparing the model's squared error to the true squared error. However, what truly matters in terms of performance, is how these metrics compare to how well a nuclear evaluator would characterize their parameterization of the same data set. With that, currently the model is being demonstrated with 100 resonances and a single spin group. Within the next year, the model will move to handling multiple spin groups, those the functional forms of the equations change (see Table 2.1 as well as the resonance cluster issue being more prevalent, as well as incorporating

the missing/extra resonance identification algorithm. Once multiple spin groups are able to be handled, then the model can evaluate real experimental data and have its results compared to that of ENDF (human evaluators), which being able to have that comparison is approximately one to two years away.

Lastly, as a reminder the method itself is problem agnostic. It is intended to be applied to a whole class of parameterized curve fitting problems that deal with objective functions that are nonconvex and require a high number of function evaluations, and even allows the model to choose how many parameters to include in the parameterization. A fairly good global solution is able to be achieved by only supplying a minimum and maximum allowable values on those fitting parameters, and using some iteration to determine the optimal window size and window overlap for a particular problem. There was no initial solutions given to the solution, the method was shown to be fast, and the final global solution was close to the true underlying value. This method is not perfect in its current form, with several possible methods of improvement noted, but it is able to take a hard, intractable optimization problem and give its user a high quality approximation of the best global solution.

Bibliography

[1] S. F. Mughabghab. Atlas of neutron resonances : resonance parameters and thermal cross sections z=1-100, 2006. vii, 30, 31

[2] Thomas Lipp. Baron: Branch and reduce optimization navigator a high level overview for cme334, Dec 2012. viii, 12, 13

[3] Stephen P. Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge Univ. Pr., 2011. x, 1, 3, 4, 49, 50, 52

[4] R Tyrrell Rockafellar. Lagrange multipliers and optimality. *SIAM review*, 35(2):183–238, 1993. 2

[5] Katta G. Murty and Santosh N. Kabadi. Some np-complete problems in quadratic and nonlinear programming. *Mathematical Programming*, 39(2):117–129, 1987. 2

[6] Rong Ge, Furong Huang, Chi Jin, and Yang Yuan. Escaping from saddle points—online stochastic gradient for tensor decomposition. In *Conference on learning theory*, pages 797–842. PMLR, 2015. 2

[7] Dimitri P. Bertsekas. *Nonlinear programming*. Athena Scientific, 1999. 2, 3

[8] Nicolas Boumal, P-A Absil, and Coralia Cartis. Global rates of convergence for nonconvex optimization on manifolds. *IMA Journal of Numerical Analysis*, 39(1):1–33, 2018. 2

[9] Jean-Baptiste Hiriart-Urruty and Lemarechal Claude. *Convex analysis and minimization algorithms II: Advanced theory and bundle methods.* Springer-verlag, 1996. 2

[10] D. Amelunxen, M. Lotz, M. B. McCoy, and J. A. Tropp. Living on the edge: Phase transitions in convex programs with random data. *Information and Inference*, 3(3):224–294, 2014. 2

[11] Stephen A. Vavasis. On the complexity of nonnegative matrix factorization. *SIAM Journal on Optimization*, 20(3):1364–1377, 2010. 3

[12] K.W. Regan. On superlinear lower bounds in complexity theory. *Proceedings of Structure in Complexity Theory. Tenth Annual IEEE Conference.* 3

[13] Franz Rendl and Henry Wolkowicz. A semidefinite framework for trust region subproblems with applications to large scale minimization. *Mathematical Programming*, 77(1):273–299, 1997. 3

[14] Michel X. Goemans and David P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM*, 42(6):1115–1145, 1995. 3

[15] Yann Richet, Grégory Caplin, Jérôme Crevel, David Ginsbourger, and Victor Picheny. Using the efficient global optimization algorithm to assist nuclear criticality safety assessment. *Nuclear Science and Engineering*, 175(1):1–18, 2013. 3

[16] CA Floudas, JL Klepeis, and PM Pardalos. Global optimization approaches in protein folding and peptide docking. *DIMACS series in discrete mathematics and theoretical computer science*, 47:141–171, 1999. 3

[17] Jill K Nelson, Megan U Hazen, and Maya R Gupta. Global optimization for multiple transmitter localization. In *MILCOM 2006-2006 IEEE Military Communications conference*, pages 1–7. IEEE, 2006. 3

[18] Jorge Nocedal and Stephen Wright. *Numerical Optimization.* Springer New York, 2006. 4

[19] Philip E. Gill, Walter Murray, and Margaret H. Wright. *Practical optimization.* Society for Industrial and Applied Mathematics, 2020. 4

[20] David G. Luenberger. *Optimization by vector space methods.* J. Wiley and Sons, 1969. 4

[21] Paul C Gilmore and Ralph E Gomory. A linear programming approach to the cutting-stock problem. *Operations research*, 9(6):849–859, 1961. 4

[22] Michele Conforti, Cornuejols Gerard, and Giacomo Zambelli. *Integer Programming.* Springer International Publishing, 2014. 4

[23] Robert H Swendsen and Jian-Sheng Wang. Replica monte carlo simulation of spin-glasses. *Physical review letters*, 57(21):2607, 1986. 4

[24] Monmarche Nicolas. *Artificial ants: From collective intelligence to real-life optimization and beyond.* ISTE, 2010. 4

[25] Scott Kirkpatrick, C Daniel Gelatt, and Mario P Vecchi. Optimization by simulated annealing. *science*, 220(4598):671–680, 1983. 4

[26] Hossein Mobahi and John W Fisher. On the link between gaussian homotopy continuation and convex envelopes. In *International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition*, pages 43–56. Springer, 2015. 4

[27] A. N. Tihonov and Arsenin Vasilij Jakovlevic. *Solutions of ill-posed problems.* V. H. Winston, 1977. 5

[28] Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288, 1996. 5

[29] Richard O. Duda, Peter E. Hart, and David G. Stork. *Pattern classification.* 5

[30] Office Of Science Nuclear Physics and Keith Janowski. volume DE-FOA-0002440. 2020. 6

[31] Catherine E Romano, Lee Bernstein, Teresa Bailey, Rike Bostelmann, David A Brown, Yaron Danon, RJ Casperson, M Devlin, Bethany Goldblum, Jeremy Conlin, et al. Proceedings of the workshop for applied nuclear data: Wanda 2020. Technical report, Oak Ridge National Lab.(ORNL), Oak Ridge, TN (United States), 2020. 6

[32] Karen A Miller and William S Charlton. An inverse transport approach to radiation source location for border security. In *Annual Meeting on the European Safeguards Research and Development Association. Aix-en-Provence France*, 2007. 7

[33] JM Hall, JA Pruet, DA Brown, M Descalle, GW Hedstrom, and SG Prussin. Modeling the production of beta-delayed gamma rays for the detection of special nuclear materials. Technical report, Lawrence Livermore National Lab., Livermore, CA (US), 2005. 7

[34] David A Brown, MB Chadwick, R Capote, AC Kahler, A Trkov, MW Herman, AA Sonzogni, Y Danon, AD Carlson, M Dunn, et al. Endf/b-viii. 0: the 8th major release of the nuclear reaction data library with cielo-project cross sections, new standards and thermal scattering data. *Nuclear Data Sheets*, 148:1–142, 2018. 8

[35] MB Chadwick, Erwan Dupont, E Bauge, A Blokhin, O Bouland, DA Brown, R Capote, A Carlson, Y Danon, C De Saint Jean, et al. The cielo collaboration: neutron reactions on 1h, 16o, 56fe, 235,238 u, and 239pu. *Nuclear Data Sheets*, 118:1–25, 2014. 8

[36] R Capote, P Dimitriou, and G Schnabel. Inden - international nuclear data evaluation network, Jun 2021. 8

[37] A. M. Lane and R. G. Thomas. R-matrix theory of nuclear reactions. *Rev. Mod. Phys.*, 30:257–353, Apr 1958. 8

[38] GM Hale, Ronald E Brown, and Nelson Jarmie. Pole structure of the j $\pi=3/2+$ resonance in he 5. *Physical review letters*, 59(7):763, 1987. 8

[39] CR Brune. An alternative parameterization of r-matrix theory. *Nuclear Physics A*, 718:472–474, 2003. 8

[40] Goran Arbanas, Vladimir Sobes, Andrew Holcomb, Pablo Ducru, Marco Pigni, and Dorothea Wiarda. Generalized reich-moore r-matrix approximation. In *EPJ Web of Conferences*, volume 146, page 12006. EDP Sciences, 2017. 8

[41] Denise Neudecker, Brooke Ellen Hejnal, Fredrik Tovesson, Morgan Curtis White, Donald L Smith, Diane Elizabeth Vaughan, and R Capote. Template for estimating uncertainties of measured neutron-induced fission cross-sections. *EPJ Nuclear Sciences and Technologies*, 4(LA-UR-17-29963), 2018. 9

[42] Roberto Capote. *Lessons learned from recent evaluations and experimental needs.* IAEA, Dec 2018. 9

[43] MT Pigni, M Herman, and P Obložinskỳ. Extensive set of cross-section covariance estimates in the fast neutron region. *Nuclear science and engineering*, 162(1):25–40, 2009. 9

[44] Doug Soltesz, MAA Mamun, AV Voinov, Z Meisel, BA Brown, CR Brune, SM Grimes, H Hadizadeh, M Hornish, TN Massey, et al. Determination of the zn 60 level density from neutron evaporation spectra. *Physical Review C*, 103(1):015802, 2021. 9

[45] Office of Scientific, Technical Information, and Nancy M Larson. UT-BATTELLE, LLC, Aug 1984. 9, 28

[46] Nicholas W Touran and Jinan Yang. Sensitivities and uncertainties due to nuclear data in a traveling wave reactor. In *Proc. Int. Conf. PHYSOR*, pages 1–5, 2016. 10

[47] A Lewis, D Neudecker, and A Koning. Developing an automatically readable, comprehensive and curated experimental reaction database, Mar 2021. 11

[48] Paulo Bedaque, Amber Boehnlein, Mario Cromaz, Markus Diefenthaler, Latifa Elouadrhiri, Tanja Horn, Michelle Kuchera, David Lawrence, Dean Lee, Steven Lidia, et al. Report from the ai for nuclear physics workshop. *arXiv preprint arXiv:2006.05422*, 2020. 11

[49] Denise Neudecker, M Grosskopf, M Herman, Wim Haeck, P Grechanuk, S Vander Wiel, Michael Evan Rising, AC Kahler, N Sly, and Patrick Talou. Enhancing nuclear data validation analysis by using machine learning. *Nuclear Data Sheets*, 167:36–60, 2020. 11

[50] David A Brown. A tale of two tools: mcres. py, a stochastic resonance generator, and grokres. py, a resonance quality assurance tool. Technical report, Brookhaven National Lab.(BNL), Upton, NY (United States), 2018. 11

[51] Aida Khajavirad and Nikolaos V Sahinidis. A hybrid lp/nlp paradigm for global optimization relaxations. *Mathematical Programming Computation*, 10(3):383–421, 2018. 11

[52] Hong S Ryoo and Nikolaos V Sahinidis. Global optimization of nonconvex nlps and minlps with applications in process design. *Computers & Chemical Engineering*, 19(5):551–566, 1995. 11

[53] Mohit Tawarmalani and Nikolaos V Sahinidis. Global optimization of mixed-integer nonlinear programs: A theoretical and computational study. *Mathematical programming*, 99(3):563–591, 2004. 13

[54] Nikolaos V Sahinidis. Baron branch and reduce optimization navigator user's manual version 4.0, Jun 2000. 13

[55] Hideki Takano and Yukio Ishiguro. Multi-level correction to breit-wigner single-level formula. *Journal of Nuclear Science and Technology*, 14(9):627–639, 1977. 16

[56] Fritz H Fröhner. *New techniques for multi-level cross section calculation and fitting*. Kernforschungszentrum Karlsruhe, 1980. 16

[57] Eugene P Wigner. On the quantum correction for thermodynamic equilibrium. In *Part I: Physical Chemistry. Part II: Solid State Physics*, pages 110–120. Springer, 1997. 18, 23

[58] Charles E Porter and Robert G Thomas. Fluctuations of nuclear reaction widths. *Physical Review*, 104(2):483, 1956. 18, 22

[59] FH Fröhner. Evaluation and analysis of nuclear resonance data. 2000. 20

[60] N. Balakrishnan. *Handbook of the logistic distribution*. Dekker, 1992. 47

[61] ME Dunn and LC Leal. Calculating probability tables for the unresolved-resonance region using monte carlo methods. *Nuclear science and engineering*, 148(1):30–42, 2004. 55

[62] JM Otter, RC Lewis, and LB Levitt. U3r: A code to calculate unresolved resonance cross section probability tables. Technical report, Atomics International, Canoga Park, Calif., 1972. 55

Vita

Originally from Florida, Jordan Armstrong spent most of his childhood in Longwood, Florida. After graduating as valedictorian from the Institute of Engineering at Lyman High School, he attended the United States Air Force Academy in Colorado Springs, Colorado. Upon his graduation from the Academy, he earned a Bachelors of Science in Applied Mathematics and Physics with his minor in Nuclear Weapons and Strategy. Additionally, with his graduation, he commissioned as Second Lieutenant in the United States Air Force. His first assignment as an Operations Research Analyst was continuing his education at the University of Tennessee, Knoxville to pursue his Masters of Science degree in Industrial and Systems Engineering. After graduation, he will continue to serve out his commitment in the Air Force as an Operations Research analyst with his next assignment being at Randolph Air Force Base in San Antonio, Texas where he will be performing data analysis and modeling for the Air Force Personnel Center.