APPLICATION OF A CONSTRAINED OPTIMIZATION TECHNIQUE TO THE IMAGING OF HETEROGENEOUS OBJECTS USING DIFFUSION THEORY

A Thesis

by

MATTHEW RYAN STERNAT

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

December 2009

Major Subject: Nuclear Engineering

APPLICATION OF A CONSTRAINED OPTIMIZATION TECHNIQUE TO THE IMAGING OF HETEROGENEOUS OBJECTS USING DIFFUSION THEORY

A Thesis

by

MATTHEW RYAN STERNAT

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Approved by:

Chair of Committee,	Jean C. Ragusa
Committee Members,	Wolfgang Bangerth
	William S. Charlton
Head of Department,	Raymond Juzaitis

December 2009

Major Subject: Nuclear Engineering

ABSTRACT

Application of a Constrained Optimization Technique to the Imaging of Heterogeneous Objects Using Diffusion Theory. (December 2009)

Matthew Ryan Sternat, B.S., Texas A&M University

Chair of Advisory Committee: Dr. Jean C. Ragusa

The problem of inferring or reconstructing the material properties (cross sections) of a domain through noninvasive techniques, methods using only input and output at the domain boundary, is attempted using the governing laws of neutron diffusion theory as an optimization constraint. A standard Lagrangian was formed consisting of the objective function and the constraints to satisfy, which was minimized through optimization using a line search method. The chosen line search method was Newton's method with the Armijo algorithm applied for step length control. A Gaussian elimination procedure was applied to form the Schur complement of the system, which resulted in greater computational efficiency.

In the one energy group and multi-group models, the limits of parameter reconstruction with respect to maximum reconstruction depth, resolution, and number of experiments were established. The maximum reconstruction depth for one-group absorption cross section or multi-group removal cross section were only approximately 6-7 characteristic lengths deep. After this reconstruction depth limit, features in the center of a domain begin to diminish independent of the number of experiments. When a small domain was considered and size held constant, the maximum reconstruction resolution for one group absorption or multi-group removal cross section is approximately one fourth of a characteristic length. When finer resolution then this is considered, there is simply not enough information to recover that many region's cross sections independent of number of experiments or flux to cross-section mesh

refinement.

When reconstructing fission cross sections, the one group case is identical to absorption so only the multi-group is considered, then the problem at hand becomes more ill-posed. A corresponding change in fission cross section from a change in boundary flux is much greater then change in removal cross section pushing convergence criteria to its limits. Due to a more ill-posed problem, the maximum reconstruction depth for multi-group fission cross sections is 5 characteristic lengths, which is significantly shorter than the removal limit.

To better simulate actual detector readings, random signal noise and biased noise were added to the synthetic measured solutions produced by the forward models. The magnitude of this noise and biased noise is modified and a dependency of the maximum magnitude of this noise versus the size of a domain was established. As expected, the results showed that as a domain becomes larger its reconstruction ability is lowered which worsens upon the addition of noise and biased noise.

To my father and mother, Louis Sternat Jr. and Patricia Sternat

ACKNOWLEDGMENTS

I would like to acknowledge my father and mother, Louis Sternat Jr. and Patricia Sternat, for their guidance, love and support through this journey. They always allowed me to pursue my passion for science. I would also like to acknowledge the rest of my family and friends for allowing me to become the person I am today.

I further acknowledge Dr. Jean C. Ragusa for his advice, advisory, and allowing me to work on this project and Dr. Wolfgang Bangerth for help with optimization methods and techniques.

TABLE OF CONTENTS

CHAPTER Page

CHAPTER Page

LIST OF TABLES

LIST OF FIGURES

FIGURE Page

FIGURE Page

FIGURE Page

CHAPTER I

INTRODUCTION

A. Objective

In the field of nuclear and global security, smuggling of special nuclear materials by transportation in containers on boats poses strong threat. To prevent this possible smuggling pathway, a detection system must be implemented that will have the ability to detect high enriched uranium (HEU) where current detection systems cannot. Due to self-shielding and long half-lives, uranium can be hard to detect through conventional methods, especially in large scale systems such as cargo containers. There are approximately 30,000 ships docking at the United States per year currently and efficient detection methods must be implemented. As of the 9/11 Commission Act of 2007, foreign seaports must scan 100 percent of the cargo entering the United States by 2012.[1]

A possible method of detection would be an active neutron imaging technique which would involve incident beams of neutrons upon the cargo container and neutron detectors surrounding the container. Using these detector readings and a constrained optimization technique, reconstructions of the material properties inside a container could be performed to determine the contents. We propose to address this parameter identification by posing it as an optimal control problem where a cost function is to be minimized. This cost function is defined as the difference between the boundary detector measurements and the boundary neutron fluxes computed from the inferred material properties inside the cargo. While many sets of material parameters may

---

The journal model is *Nuclear Science and Engineering*.

have the ability to reconstruct the outer detector readings, constraints upon these must be applied to limit the number of solutions. The valid constraint used in this work will involve the governing conservation law of neutron physics in the container, thereby limiting the solution of material parameters to a realistic or physical case.

This is an optimal control problem because the difference between the computed iterative solution at the boundary and the neutron detector readings must be minimized while satisfying the neutron transport equation or an approximation of it. The equations derived from the optimization process are nonlinear, naturally requiring a descent method to solve them. This problem is ill-posed because small changes in the material properties can often lead to large changes in the neutron fluxes at the boundaries. Application of iterative methods cannot guarantee convergence for any realistic initial guess due to the ill-posedness of this nonlinear problem.

B. Imaging

Active neutral particle imaging techniques involve illuminating a domain with beams of particles of known intensity and taking measurements around the domain of the boundary outflow in an orientation shown in Figure I-1. Active neutral particle imagining is performed to reconstruct information of the inside of the domain. The location, energy, and angle of incidence of the incoming particles can be varied and more information can be gathered. With multiple experiments of incoming beams around the domain, the material properties reconstruction satisfying all experiments at once can yield improved reconstruction.

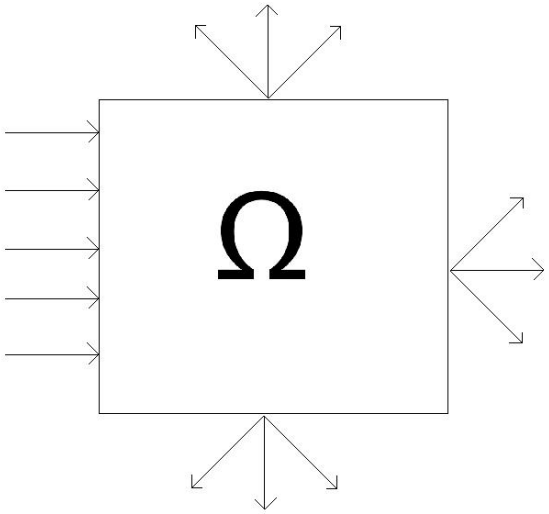


Fig. I-1. Example of incoming and outgoing particle currents

An example of this is optical tomography,[2] where a nonlinear system containing nonlinear combinations of the parameters intended to be reconstructed and the state variables is formed by the equations that define how light is transmitted and scattered through an object and often have no analytical solution.[3] By observing the light exiting the tissues, a reconstruction of the absorption and scattering coefficients inside the sample is performed.[3] These problems are solved iteratively using forward models to solve for the outgoing currents based on an initial guess on the interaction coefficients directly, and nonlinear optimization techniques to update the interaction coefficients[4].[5] This algorithm process is repeated until the iterative solution converges with the observed light exiting the tissues. This is very similar to the problem of special nuclear material (SNM) smuggling, but instead of biological matter, containers that can be up to many optical thicknesses deep are to be imaged using neutrons.

Another example of neutral particle imaging is in large ports for object detection. There are systems that use photons that operate in the 6-9 MeV range to image large

cargo containers. Most of currently implemented cargo imaging uses either x-rays or gamma rays. The x-ray systems are commonly used to ensure containers are empty without opening them or to determine contents of smaller containers where gamma rays are not needed. These types of systems are capable of producing images of large containers and trucks with spatial resolution of 9mm for the gamma systems and 2mm for x-ray systems.[6] While these types of systems can produce an image of the internal contents of a container, they cannot by themselves determine if fissile material is present. This is where a multigroup neutron imaging system would have the greatest impact. If a system were able to reconstruct fission cross sections to determine whether fissile material were present accurately, greater detection probability of smuggled HEU could be achieved.

Neutron imaging varies from gamma or x-ray imaging in the way they interact with matter quite differently then x-rays do, having a high interaction probability with hydrogen and much lower attenuation in heavier elements such as lead. While x-ray interaction probability is directly proportional to the atomic number of the material, neutron interaction is isotope-dependent causing both radiography mechanisms to excel in different media types.[7] Common examples of neutron radiography include nuclear fuel surveys, multi-phase flow imaging, and explosive device imaging. In the case at hand, HEU could easily be shielded from x-rays causing methods involving x-rays or radiation emitted from the material itself to be ineffective. Neutron interaction probabilities are energy-dependent, where neutrons of typical source energy have high scattering interaction probability in many materials, limiting the ability of larger scale imaging.

C. Optimization and Inverse Problem Solving

The majority of inverse problems or imaging techniques involve an optimization process in which a function is minimized or maximized by iterating the functions variables, often subject to constraints. The most commonly used methods to solve problems of any type involve iterative algorithms. In the optimization process, the optimum of a given function is obtained by solving the optimality conditions using an optimization algorithm. There is no universal optimization algorithm but instead a collection of algorithms in which each is valid for specific problem types.[8]

An example of application of inverse transport is the determination of interface locations in a multilayer domain of unknown dimensions. In this specific example, source gamma-rays were passed through a domain and observed at boundaries, then the location of the interfaces is solved for using optimization methods.[9] This is similar to the problem at hand except that instead of the material properties being known and the interface locations reconstructed, the material properties are unknown but reconstructed and assumed piecewise constant over a mesh.

D. Thesis Overview

The next chapter provides an in depth look at optimization methods from a mathematical standpoint. This chapter provides a complete step by step approach to optimization problems including specific methods.

Chapter III contains the development and implementation of the presented optimization methods to the inverse problem using diffusion theory.

Chapter IV presents the results of reconstructions of various domains. Many tests were performed in order to have an understanding of the workable space with respect to domain size, mesh size, number of experiments, and measurement location.

CHAPTER II

OPTIMIZATION METHODS

The goal of an optimization problem is to find the combination of parameters that optimize a given quantity subject to some restrictions or constraints.[10] The parameters that may be changed in the process of optimization are called control or decision variables while the restrictions on parameters are known as constraints.[10] Mathematically speaking, optimization is the minimization or maximization of an objective function defined by a problem statement and is subject to constraints on its variables. Often a vector x is formed that consists of the unknowns or parameters, f if the objective function, a scalar function of x, that we want to maximize or minimize, and a series of constraint functions, c_i, which are scalar functions of x that define constraints the unknown vector x must satisfy. Using this notation, the optimization problem can be written as shown in Equation 2.1.

$$\min_{x \in R^n} f(x) \text{ subject to } \begin{array}{l} c_i(x) = 0, \ i \in \xi \\ c_i(x) \geq 0, \ i \in \mathcal{I} \end{array} \tag{2.1}$$

where c_i can be an equality or inequality constraint and ξ and $\mathcal{I}$ are the sets of equality and inequality constraints.

This chapter provides an overview of optimization methods in general, starting with section A on optimization problem classifications.Section B provides the optimality conditions. The section of these optimality conditions is then detailed in section C using steepest descent and Newton's method. This chapter will then cover convergence criteria D, step-length control E and conclude with the Schur complement method employed to reduce the system's dimensions F.

A. Optimization Classifications

In deterministic optimization methods, first and second derivatives of the objective function, $f(x)$, need to be computed. These problems are classified by the type of their control variables and nature of their objective functions which are usually linear, quadratic, or fully nonlinear.[10] In certain cases, this function can be discontinuous and may contain integers and binary variables; these problems can only be optimized using discrete optimization methods for which derivatives are not defined. Other classes of problems, where the components of the given function are allowed to be real numbers can be optimized continuously. These continuous functions are normally easier to solve because they are often smooth and twice differentiable.[8]

When a problem is considered, it is classified by the nature of its objective function where some problems have constraints upon their variables and some do not. Problems that involve constrained variables are optimized using constrained optimization. Sometimes these constraints play a important role in determining the solution and an example of a constrained objective function can be seen in Figure II-1.

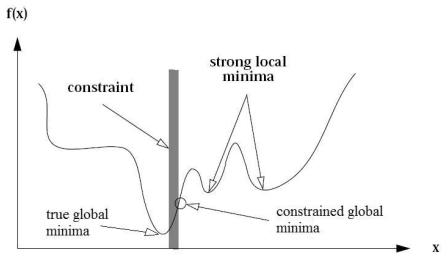


Fig. II-1. Example of a constrained objective function

For instance, in a budgetary problem, if the global solution lies outside the limits due to budgetary constraints, a local solution that lies within these constraints will be the best solution.[8] Whereas in fully unconstrained optimization, there are no limits

on any of the variables and the global minimum is the true function minimum as shown in Figure II-2.

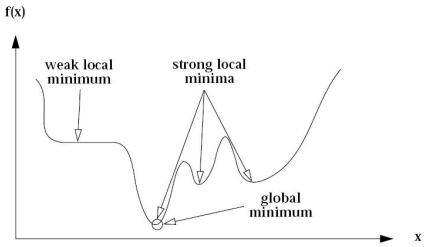


Fig. II-2. Example of an unconstrained objective function

Sometimes, even if there are minor constraints on a problems variables, if they do not interfere with the optimization algorithms, unconstrained optimization can be applied.[8]

B. Optimality Conditions

To find the minimum of $f(x)$, conditions are applied to find where $\nabla f = 0$. When constraints are upon $f(x)$, for example $c(x) = 0$, then a Lagrangian functional is introduced such as in Equation 2.2.

$$\min_{x \in R^n} f(x) \text{ subject to } c(x) = 0 \Leftrightarrow \mathcal{L}(x, \lambda) = \nabla f + \lambda \nabla c(x) \tag{2.2}$$

where λ is a Lagrange multiplier. A saddle point in $\mathcal{L}$ is found where $\nabla \mathcal{L} = 0$, which is

$$\frac{\partial \mathcal{L}}{\partial x} = 0 = \nabla f + \lambda \nabla c \tag{2.3}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = 0 = c(x) \tag{2.4}$$

where the first equation implies that $\nabla f \propto \nabla c$ and in the second equation the con-

straint arises $c(x) = 0$. These derivatives of the Lagrangian form a set of Karush-Kuhn-Tucker conditions or optimality conditions to be satisfied. Figure II-3 shows a simple iso-contour plot of $f(x)$ with a $c(x) = 0$ line solution.

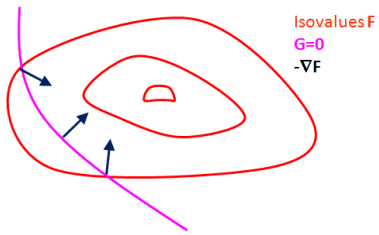


Fig. II-3. Iso-contour plot showing an objective function and a constraint

C. Line Search Methods

In a line search method, algorithms choose a direction p_k and search along this direction from the current iterate for a new iterate that is closer to the optimality conditions.[8] There are various methods that can be used to determine a line search direction along with many algorithms to determine how far in that direction to go.[8] The goal of this optimization problem is the minimization of $f(x)$ while satisfying any given conditions. At this minimum $\partial_j \mathcal{L}(x) = 0$ where j is any field variable in $\mathcal{L}(x)$. Just as in any iterative method, an initial guess is made and at this iterate $\nabla \mathcal{L}(x^k) \neq 0$. We will now describe two such techniques: the steepest descent and Newton's descent and then provide an example of a step-length control algorithm.

1. Steepest Descent

An obvious direction is the steepest descent. The steepest descent direction follows the opposite direction of the gradient, or the direction perpendicular to the iso-contours. For example in a simple two dimensional optimization scheme, this would be very

similar to a ball in a valley rolling to the bottom. This can be seen in Figure II-4, where X^* is the global minimum. The gradient of f is perpendicular to the iso-contour of $\mathcal{L}$.

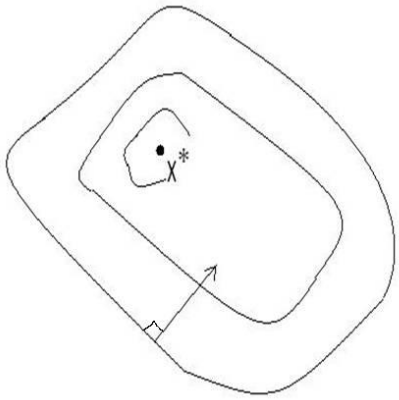


Fig. II-4. Steepest descent direction

In this steepest descent method, the descent direction is $p_k = -\nabla\mathcal{L}_k$ as shown in Equation 2.5.

$$x_{k+1} = x_k + \alpha p_k = x_k - \alpha\nabla\mathcal{L}_k \tag{2.5}$$

The steepest descent algorithm consists of the following:

1. Initialization: set k=0, set convergence criteria ϵ, choose x_k.
2. If $\|\nabla\mathcal{L}_k\| < \epsilon$ then exit, otherwise continue.
3. Compute $p_k = -\nabla\mathcal{L}_k$
4. Determine step length α_k (see Section 2.4).
5. Compute new update according to Eq. 2.5.
6. $k \leftarrow k + 1$ and go to 2.

This method requires only first derivatives, tends to get stuck in local minima, and is slowly converging while it iteratively takes steps in the gradient direction to a new solution with lower optimality condition. The steepest descent direction is updated at every step indexed by k and its progress is slow as some regions of indefinite curvature are encountered especially near a solution.[10] The convergence rate of this method is much slower than other higher order methods.

2. Newton's Method

A significantly more efficient higher order method can be derived from where the steepest descent method left off. Newton's method is comprised of second derivatives and is a curve of best fit method. This uses a line search direction other then the steepest descent, and is derived from the second order Taylor series approximation of $f(x_k + p)$ and is shown in Equation 2.6.

$$\mathcal{L}(x_k + p) \approx \mathcal{L}_k + p^T \nabla \mathcal{L}_k + \frac{1}{2} p^T \nabla^2 \mathcal{L}_k p = m_k(p) \tag{2.6}$$

An example of such direction can be shown in Figure II-5.

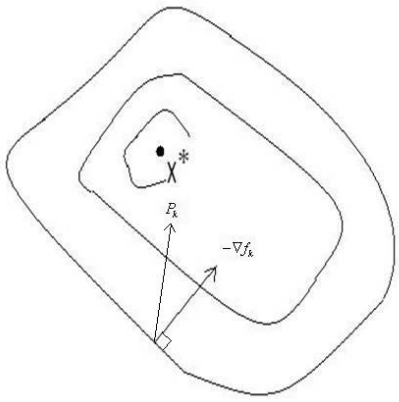


Fig. II-5. Newton's method vs. steepest descent direction

Using this second-order Taylor series approximation, the vector p that minimizes $m_k(p)$ is obtained by setting the derivative of $m_k(p)$ to zero leaving Equation 2.7.

$$\nabla \mathcal{L}_k + \nabla^2 \mathcal{L}_k p_k = 0 \tag{2.7}$$

where $\nabla^2 f_k = H$, the Hessian matrix. Then solving for p_k yields

$$p_k = -(\nabla^2 \mathcal{L}_k)^{-1} \nabla \mathcal{L}_k = -H^{-1} \nabla \mathcal{L}_k \tag{2.8}$$

$$x_{k+1} = x_k + \alpha p_k = x_k - \alpha H^{-1} \nabla \mathcal{L}_k \tag{2.9}$$

This Newtonian search direction tries to quadratically approximate a curve at iterate x_k and goes to the minimum of the quadratic fit. For a simple quadratic system, the minimum of $f(x)$ could be met after one step. Due to the nonlinearity and complexity of most systems, Newton's method often is applied where steepest descent methods will not converge.

The steepest descent and Newton's method are both of the form:

$$x_{k+1} = x_k + \alpha p_k = x_k - \alpha B^{-1} \nabla f_k, \tag{2.10}$$

with $B = I$, for steepest descent and $B = H^{-1}$, for Newton's method.

D. Convergence Criteria

The nonlinear system in Equation 2.4 will converge when the optimality conditions are satisfactory close to their solution. Some very nonlinear systems with random noise and bias will be very difficult to drive the optimality conditions close to the true solution. At the optimum, the optimality conditions will be met but when constraints are present the solution that is closest to the optimality conditions while

satisfying the constraints will be the global solution. For example, if the true global solution was unreachable due to constraints then the convergence criteria of the optimality conditions will not be achievable and the closest solution to these optimality conditions will be the solution.

E. Step-Length Selection Control and Algorithms

Now that a line search direction has been determined, how far to travel in that direction is established next. When the objective function is not smooth, a full Newtonian direction step may not lead to a reduction in optimality condition. Simple algorithms can be used to attempt to ensure the optimality conditions are lowered. Starting with the general *sufficient decrease* condition:

$$\mathcal{L}(x_k + \alpha p_k) - \mathcal{L}(x_k) < \alpha\lambda\nabla\mathcal{L}(x_k)^T p_k \tag{2.11}$$

where the descent direction derived from Newton's Method:

$$p_k = -H^{-1}\nabla\mathcal{L}(x) \tag{2.12}$$

where $\lambda \in (0, 1)$ is an algorithmic parameter typically around 10^{-4}. Beginning with $\alpha = 1$ repeatedly reduce α using any strategy that satisfies the general *sufficient decrease* condition.

$$\alpha_+ \in [\beta_{low}\alpha_c, \beta_{high}\alpha_c] \tag{2.13}$$

where $0 < \beta_{low} < \beta_{high} < 1$. The choice of $\beta = \beta_{low} = \beta_{high}$ is a simple rule in the Armijo algorithm shown below.[8]

1. Initialization: set α=1 and $\lambda \in (0, 1)$, set convergence criteria ϵ, choose x_k.

2. If $\mathcal{L}(x_k + \alpha p_k) - \mathcal{L}(x_k) < \alpha\lambda\nabla\mathcal{L}(x_k)^T p_k$, $x_{k+1} = x_k + \alpha p_k$. If not, continue.

3. Reduce α, return to step 2.

In an exact line search, the special case where λ leads to the exact minimum of $\mathcal{L}(x_c + \alpha p_k)$, is not only more expensive computationally but can often degrade the performance of the algorithm in general.

F. Schur Complement Method

The Schur Complement Method is a process of system simplification for a system involving a Karush-Kuhn-Tucker (KKT) type matrix.[8] It is a method that involves eliminating variables in a system to simplify and often obtain a linear system of one variable. One example of such matrix system can be shown in Equation 2.14.

$$\begin{bmatrix} G & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} p \\ \lambda \end{bmatrix} = \begin{bmatrix} g \\ h \end{bmatrix} \text{ where: } \dim(p) >> \dim(\lambda) \tag{2.14}$$

This KKT matrix has blocks of entries equal to zero and can easily simplified by simple algebra. A typical KKT matrix may have more rows of blocks, as many multivariate problems have multiple optimality conditions, but can be simplified in the same manner. After assuming G is positive definite, p is solved for in the first equation in 2.14 in terms of λ then substituted in the second equation leaving a system of λ alone as shown in Equation 2.15.

$$\lambda = A^{-T}(g - GA^{-1}h) \tag{2.15}$$

This smaller system is solved for λ and then the other vector variable p can be directly solved for as shown in Equation 2.16.

$$Gp = A^T\lambda - g \tag{2.16}$$

This method involves a matrix inversion of G and A^T which often results in signif-

icantly lower condition number than inverting the full original system matrix. This method of system simplification can be applied to reduce system runtime and improve computational efficiency.

CHAPTER III

INVERSE DIFFUSION MODELS

Neutron imaging is type of non-invasive inverse problem involving incoming and outgoing neutron beams where measurements are made only on the boundary of a domain. The neutron transport equation defines how neutrons behave in matter through various interaction types and can be used in inverse problems. An approximation to the transport model is diffusion theory, which introduces some simplicity for handling the angular dependence of the neutron population. In this thesis, we use the diffusion approximation to model the distribution of particles.

In inverse theory, many problems are ill-conditioned or are ill-posed, where a small variation in the input data causes a large change in the results.[11] In inverse diffusion methods, the flux solution to be solved for depends on unknown internal parameters of the domain. Generally, an initial guess is set for the domain parameters, then the flux is solved and the domain parameters are updated using optimization methods.

This chapter begins with an introduction to neutron diffusion theory and the application of a finite element method to solve neutron diffusion problems. The implementation of the previous chapters optimization methods applied to inverse diffusion models are described next, first deriving optimality conditions and then employing Newton's method to solve them. The optimum control problem is formulated for multiple experiments in the context of the multigroup diffusion approximation.

A. Neutron Diffusion Theory

The neutron diffusion equation is derived from the Boltzmann transport equation by integrating over all directions and using the diffusion theory expression for the neutron current derived from Fick's Law.[12] The one-group neutron diffusion equation, shown in Equation 3.1 and 3.2, is a phase-space dependent equation that relates the neutron scalar flux phase-space distribution across a domain to its nuclear properties.[12]

$$-\nabla \cdot D(\vec{r})\nabla\Phi + (\Sigma_a(\vec{r}) - \nu\Sigma_f(\vec{r}))\Phi = Q(\vec{r}) \text{ in } \Omega \tag{3.1}$$

$$\frac{\Phi}{4} + \frac{D(\vec{r})}{2}\partial_n\Phi = J^{inc}(\vec{r}) \text{ in } \partial\Omega \tag{3.2}$$

B. Finite Element Diffusion Solver

The forward diffusion models used in this problem are solved numerically using finite element methods. The finite element method is a numerical technique for finding approximate solutions of partial differential equations. This method differs from finite difference in such that finite difference methods approximate PDE equations while finite element methods approximate their solutions. Both of these methods discretize the domain into a mesh and the finite element method used in this work approximates the PDE's solution as a piecewise bi-linear function across each mesh cell.

In the finite element setting, the diffusion equation becomes:

$$\left[A_D + A_\Sigma + \frac{1}{2}M_{\partial\Omega}\right]\Phi = A\Phi = F \tag{3.3}$$

with:

1.

$$A_D(i,j) = \int_\Omega \nabla b_i \cdot \nabla b_j \tag{3.4}$$

If D is constant, it can be factored out and

$$A_D = D \times S \tag{3.5}$$

$$\text{with } S(i,j) = \int_\Omega D \nabla b_i \cdot \nabla b_j. \tag{3.6}$$

S is known as the stiffness matrix

2.

$$A_\Sigma(i,j) = \int_\Omega \Sigma b_i b_j \tag{3.7}$$

If Σ is constant, it can be factored out and

$$A_\Sigma = \Sigma \times M \tag{3.8}$$

$$\text{with } M(i,j) = \int_\Omega b_i b_j. \tag{3.9}$$

M is known as the mass matrix.

3.

$$M_{\partial\Omega}(i,j) = \int_{\partial\Omega} b_i b_j \tag{3.10}$$

4.

$$F(i) = \int_\Omega Q b_i + 2 \int_{\partial\Omega} J^{inc} b_i \tag{3.11}$$

5. Now, Φ is to be understood as a vector containing the flux values Φ_i at the nodes.

1. Finite Element Meshes

There are a variety of element meshes that can be implemented with finite element method. The most common of these are triangular and rectangular elements. After the domain has been broken elements, a set of piecewise polynomials are used for approximation. This must result in a function that is continuous with an integrable or continuous first or second derivative on the entire region. Polynomials of linear type in x and y in Equation 3.12 are often used with triangular elements and polynomials of bilinear type, shown in Equation 3.13 are used with rectangular elements.

$$\Phi(x, y) = a + bx + cy \tag{3.12}$$

$$\Phi(x, y) = a + bx + cy + dxy. \tag{3.13}$$

The two dimensional domain is broken up into finite element meshes. This consists of a fine mesh to be used in the flux solver and a coarse mesh that will be the regions where the cross section (taken to be piece-wise constant) are reconstructed. The difference between these two meshes is a refinement which is variable in each dimension of the domain. This refinement is necessary due to the ill-posed problem and lack of information required to solve this inverse diffusion problem. An example of these meshes and refinement is shown in Figure III-1.

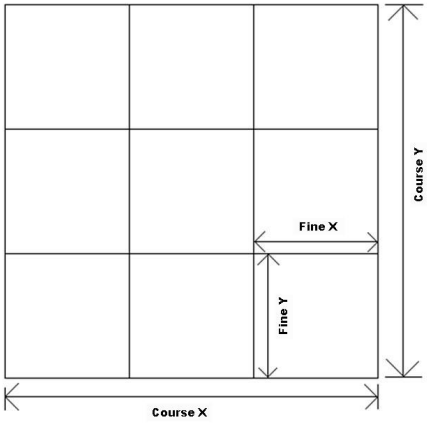


Fig. III-1. Example of the finite element meshes for the diffusion problem

2. Finite Element Methods

An attractive feature of the finite element method is its ability to handle complicated geometries with relative ease. This enables much more complicated domains and geometries to be solved where the finite difference method in its basic form is restricted to handle rectangular shapes and simple variations of. One reason for this is the finite element method's relative easy with which the boundary conditions are handled.[13] A lot of problems have boundary conditions involving derivatives and irregularly shaped boundaries which are difficult to handle using finite difference techniques.[13] The finite difference method handles these boundary conditions by approximating the derivative using a difference quotient at the grid points where irregular shaping of the boundary makes the grid point locations difficult.[13] The finite element method handles the boundary conditions in a functional's integral that is being minimized, which is independent of the particular boundary conditions of the problem itself.[13]

C. Optimization Functional

1. Misfit: To Minimize

When an iterative solution is considered, a function called the misfit is introduced which represents the iterative solutions's (Φ) distance from the measured solution (z) at the boundary. In the case at hand, a synthetic measured solution is used by evaluating the forward model with the true material parameters. While iterating to satisfy the optimality conditions, this misfit represents the distance of Φ from z at the boundary and should converge to zero as the solution is approached. In the case of the problem, the misfit will be defined by Equation 3.14.

$$\text{misfit} = \frac{1}{2}\int_{\partial\Omega'} \left[\Phi - z\right]^2, \tag{3.14}$$

where $\partial\Omega'$ represents the portion of $\partial\Omega$ where measurements are made. In the finite element setting:

$$\text{misfit} = \frac{1}{2}\left[\Phi - z\right]^T M_{meas}\left[\Phi - z\right], \tag{3.15}$$

where

$$M_{meas}(i,j) = \int_{\partial\Omega'} b_i b_j. \tag{3.16}$$

If the entire boundary is used to measure data, then $M_{meas} = M_{\partial\Omega}$. This misfit is directly used in the Lagrangian.

2. Lagrangian Functional

In constraint optimization problems, a $\mathcal{L}$ functional is formed and minimized consisting of two parts, one being the misfit representing distance from the true solution at the boundary and the other being constraints, here the diffusion equation acts as the governing equation, or in other words, $\mathcal{L} = \text{misfit} + \text{constraint}$. The optimality

conditions are derived from first derivatives of this Lagrangian with respect to each field variable and will be minimized in an iterative manner. The field variables are: Φ, λ, and Σ where Φ is the neutron flux, λ is the Lagrange multiplier or adjoint flux, and Σ is the set of piecewise continuous cross sections for the domain.

Constraints must be applied as several solution sets may satisfy the misfit conditions and application of constraints helps in selecting these solutions. The governing physics of the domain act as a constraint in the problem at hand and may help select a solution that physically realistic. Application of the neutron diffusion equation here will be the constraint of choice, but additional constraints may be implemented involving physical limits: $3\Sigma_a < \frac{1}{D}$, $\Sigma_a > 0$, $\Sigma_f < \Sigma_a$ and ensuring a domain remains subcritical ($k_{eff} < 1$). Most all problems objective functions are smooth enough to where the additional constraints are not needed. The goal of this optimization problem is to find the saddle point in a Lagrangian functional $\mathcal{L}$. If only Σ is to be determined, then

$$\mathcal{L}(\Phi, \lambda, \Sigma) = \frac{1}{2}\left[\Phi - z\right]^T M_{meas}\left[\Phi - z\right] + \lambda^T \left\{ \left[A_D + A_\Sigma + \frac{1}{2} M_{\partial\Omega} \right] \Phi - F \right\}. \quad (3.17)$$

The KKT optimality conditions arise as the derivatives of the Lagrangian with respect to each field variable and must be satisfied as the solution is approached. When $\mathcal{L}$ is the optimum, each of these optimality conditions will be satisfied.

$$\frac{\partial \mathcal{L}}{\partial \Phi} = M_{meas}\left[\Phi - z\right] + \left[A_D + A_\Sigma + \frac{1}{2} M_{\partial\Omega} \right]^T \lambda = 0, \quad (3.18)$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = \left[A_D + A_\Sigma + \frac{1}{2} M_{\partial\Omega} \right] \Phi - F = 0, \quad (3.19)$$

$$\frac{\partial \mathcal{L}}{\partial \Sigma} = \lambda^T \partial_\Sigma A \Phi = 0. \quad (3.20)$$

There are several features embedded in the presented optimality conditions, such as in Eq. 3.19 the constraint to the optimization problem arises as the diffusion residual must approach zero. In Eq. 3.18 the adjoint diffusion term arises with the misfit as a forcing term which too must approach zero as the method nears the solution. The Lagrange multiplier, λ, has a clear meaning as the adjoint flux. Equations 3.18-3.20 form a nonlinear system of equations to be satisfied. These KKT optimality conditions form a nonlinear system of equations, therefore Newton's method is employed.

D. Hessian System

Upon implementing Newton's method of optimization, the Hessian matrix must be formed. This Hessian matrix is the Jacobian matrix of the KKT optimality conditions which is composed of second derivatives of $\mathcal{L}$. The derivatives of the optimality condition are taken with respect to each field variable and put together to form a matrix.

$$\frac{\partial^2 \mathcal{L}}{\partial \Phi^2} = M_{meas} \tag{3.21}$$

$$\frac{\partial^2 \mathcal{L}}{\partial \lambda^2} = 0 \tag{3.22}$$

$$\frac{\partial^2 \mathcal{L}}{\partial \Sigma^2} = 0 \tag{3.23}$$

$$\frac{\partial^2 \mathcal{L}}{\partial \Sigma \partial \Phi} = \partial_\Sigma A^T \Phi \tag{3.24}$$

$$\frac{\partial^2 \mathcal{L}}{\partial \Sigma \partial \lambda} = \partial_\Sigma A \Phi \tag{3.25}$$

$$\frac{\partial^2 \mathcal{L}}{\partial \lambda \partial \Phi} = \left[A_D + A_\Sigma + \frac{1}{2} M_{\partial\Omega} \right] \tag{3.26}$$

Note that $M^T = M$ and $\left[A_D + A_\Sigma + \frac{1}{2} M_{\partial\Omega} \right]^T = \left[A_D + A_\Sigma + \frac{1}{2} M_{\partial\Omega} \right]$ in the case of one-group diffusion approximation. To simplify the system, the notation

$\left[A_D + A_\Sigma + \frac{1}{2}M_{\partial\Omega}\right] = A$ will be used, yielding the Hessian system below in Equation 3.27.

$$\begin{bmatrix} M_{meas} & A^T & \partial_\Sigma A\lambda \\ A & 0 & \partial_\Sigma A\Phi \\ \lambda^T\partial_\Sigma A & \Phi^T\partial_\Sigma A & 0 \end{bmatrix} \begin{bmatrix} \delta\Phi \\ \delta\lambda \\ \delta\Sigma \end{bmatrix} = - \begin{bmatrix} M_{meas}\left[\Phi - z\right] + A^T\lambda \\ A\Phi - F \\ \lambda^T\partial_\Sigma A\Phi \end{bmatrix} \tag{3.27}$$

where $\delta\Phi$, $\delta\lambda$, and $\delta\Sigma$ are updates and give the Newton iterate:

$$H\delta x^k = -F(x^k) \tag{3.28}$$

$$x^{k+1} = \delta x^k + x^k. \tag{3.29}$$

E. Implementation of Schur Complement

This Hessian system can be simplified to reduce run time by a Gauss elimination of $\delta\Phi$ and $\delta\lambda$ to arrive at a system with only $\delta\Sigma$. The main matrix that is inverted in the Schur complement solution has a lower condition number then the straight forward Hessian system.

The second row of the above Hessian system is solved first for $\delta\Phi$ in terms of $\delta\Sigma$ and constants.

$$A\delta\Phi + \partial_\Sigma A\Phi\delta\Sigma = -A\Phi + F \tag{3.30}$$

$$\delta\Phi = A^{-1}\left(-A\Phi + F - \partial_\Sigma A\Phi\delta\Sigma\right) \tag{3.31}$$

The first row of the above Hessian system is solved for $\delta\lambda$ in terms of $\delta\Phi$ and $\delta\Sigma$.

$$M_{meas}\delta\Phi + A^T\delta\Sigma + \partial_\Sigma A\lambda\delta\Sigma = -M_{meas}\left[\Phi - z\right] - A^T\lambda \tag{3.32}$$

$$\delta\lambda = A^{-T}\left(-M_{meas}\left[\Phi - z\right] - A^T\lambda - M_{meas}\delta\Phi - \partial_\Sigma A\lambda\delta\Sigma\right) \tag{3.33}$$

$$\delta\lambda = A^{-T}\begin{pmatrix} -M_{meas}\left[\Phi - z\right] - A^T\lambda - M_{meas}\cdot \\ \left[A^{-1}\left(-A\Phi + F - \partial_\Sigma A\Phi\delta\Sigma\right)\right] - \partial_\Sigma A\lambda\delta\Sigma \end{pmatrix} \tag{3.34}$$

These solutions for $\delta\Phi$ and $\delta\lambda$ can be plugged back into the third row of the above Hessian system to solve for $\delta\Sigma$. Starting with the third row from Equation 3.27:

$$\lambda^T M\delta\Phi + \Phi^T M\delta\lambda = -\lambda^T M\Phi \tag{3.35}$$

then filling $\partial\Phi$ and $\partial\lambda$ solutions:

$$\Phi^T\partial_\Sigma A\left(A^{-T}\begin{pmatrix} -M_{meas}\left[\Phi - z\right] - A^T\lambda - M_{meas}\cdot \\ \left[A^{-1}\left(-A\Phi + F - \partial_\Sigma A\Phi\delta\Sigma\right)\right] - \partial_\Sigma A\lambda\delta\Sigma \end{pmatrix}\right) + \lambda^T\partial_\Sigma A\left(A^{-1}\left(-A\Phi + F - \partial_\Sigma A\Phi\delta\Sigma\right)\right) = -\lambda^T\partial_\Sigma A\Phi \tag{3.36}$$

then grouping terms with and without $\delta\Sigma$:

$$\begin{aligned}&\left[-\lambda^T\partial_\Sigma AA^{-1}\left(\partial_\Sigma A\Phi\right) + \Phi^T\partial_\Sigma A\left[A^{-T}\left(-M_{meas}A^{-1}\left(-\partial_\Sigma A\Phi\right) + \partial_\Sigma A\lambda\right)\right]\right]\left[\delta\Sigma\right] \\ &= \begin{array}{c} -\lambda^T\partial_\Sigma A\Phi - \lambda^T\partial_\Sigma AA^{-1}\left(-A\Phi + F\right) - \Phi^T\partial_\Sigma AA^{-1}\cdot \\ \left(-M_{meas}\left[\Phi - z\right] - A^T\lambda - M_{meas}A^{-1}\left(-A\Phi + F\right)\right) \end{array}\end{aligned} \tag{3.37}$$

The operator created on the left hand side of Eq. 3.37 is the Schur complement for the system and will be called S. The right hand side will be called U for simple notation.

$$S = -\lambda^T\partial_\Sigma AA^{-1}\left(M\Phi\right) + \Phi^T\partial_\Sigma A\left[A^{-T}\left(-M_{meas}A^{-1}\left(-\partial_\Sigma A\Phi\right) + \partial_\Sigma A\lambda\right)\right] \tag{3.38}$$

$$\delta\Sigma = S^{-1}U \tag{3.39}$$

This method for the one group case creates a system for $\delta\Sigma$ only which is significantly smaller in size then the original system shortening iteration runtime.

F. Extension to Multiple Experiments

To also enable greater reconstruction abilities, multiple experiments can be performed over a domain, each experiment involving different source locations. Every experiment has a unique flux and adjoint solution to reconstruct the same cross sections to better the likelihood of success. The most logical choices are to break the boundary into halves, quarters, eighths, and sixteenths. The optimality conditions and misfit will be reduced for each of these experiment's flux solutions while optimizing the same set of parameters for the domain. This will provide much more data and enable greater reconstruction ability then a single experiment. Many runs will be done with this code to test the limits of reconstruction with respect to various elements of the domain such as mesh size, number of experiments, variable refinement, and domain size.

1. Optimality Conditions

The Lagrangian with multiple experiments will be a simple summation over the Lagrangian for each experiment. This Lagrangian for a total of I experiments is:

$$\mathcal{L}(\Phi, \lambda, \Sigma) = \sum_{i=1}^{I} \frac{1}{2} \left[\Phi_i - z_i\right]^T M_{meas} \left[\Phi_i - z_i\right] + \sum_{i=1}^{I} \lambda_i^T \left[A_i\Phi_i - F_i\right] \tag{3.40}$$

Similar to the single experiment case, the optimality conditions to be satisfied are derived from the first derivatives with respect to Φ_i and λ_i for each experiment

and the same Σ as before.

$$\frac{\partial \mathcal{L}}{\partial \Phi_i} = M_{meas}\left[\Phi_i - z_i\right] + A_i^T \lambda_i \ \forall i \tag{3.41}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_i} = A_i \Phi_i - F_i \ \forall i \tag{3.42}$$

$$\frac{\partial \mathcal{L}}{\partial \Sigma} = \sum_{i=1}^{I} \lambda_i^T \partial_\Sigma A \Phi_i \tag{3.43}$$

This creates more conditions for the system to be satisfied enabling greater reconstruction ability.

2. Hessian System

The Hessian system for multiple experiments is similar to the single experiment case. The corresponding second derivatives for the Hessian system were derived for each experiment forming a new Hessian matrix.

$$\frac{\partial^2 \mathcal{L}}{\partial \Phi_i^2} = M_{meas,i} \tag{3.44}$$

$$\frac{\partial^2 \mathcal{L}}{\partial \lambda_i^2} = 0 \tag{3.45}$$

$$\frac{\partial^2 \mathcal{L}}{\partial \Sigma^2} = 0 \tag{3.46}$$

$$\frac{\partial^2 \mathcal{L}}{\partial \Sigma \partial \Phi_i} = \partial_\Sigma A^T \lambda_i \tag{3.47}$$

$$\frac{\partial^2 \mathcal{L}}{\partial \Sigma \partial \lambda_i} = \partial_\Sigma A \Phi_i \tag{3.48}$$

$$\frac{\partial^2 \mathcal{L}}{\partial \lambda \partial \Phi} = \sum_{i=1}^{I} A_D + A_\Sigma + \frac{1}{2} M_{\partial\Omega} \tag{3.49}$$

where $M_{meas,i}$ is the mass matrix corresponding to experiment i's measurement location. This multiple experiment Hessian system has identical equations for Φ and λ but with an equation for each experiment. These experiments all operate over the same

set of cross sections, therefore the final equation has terms from every experiment.

$$\begin{bmatrix} M_{meas,1} & 0 & A_1^T & 0 & \cdots & \partial_\Sigma A\lambda_1 \\ 0 & M_{meas,2} & 0 & A_2^T & \cdots & \partial_\Sigma A\lambda_2 \\ A_1 & 0 & \ddots & 0 & \cdots & \partial_\Sigma A\Phi_1 \\ 0 & A_2 & 0 & 0 & \cdots & \partial_\Sigma A\Phi_2 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \lambda_1^T\partial_\Sigma A & \lambda_2^T\partial_\Sigma A & \Phi_1^T\partial_\Sigma A & \Phi_2^T\partial_\Sigma A & \cdots & 0 \end{bmatrix} \begin{bmatrix} \delta\Phi_1 \\ \delta\Phi_2 \\ \delta\lambda_1 \\ \delta\lambda_1 \\ \vdots \\ \delta\Sigma \end{bmatrix} = \begin{bmatrix} M_{meas,1}\left[\Phi_1 - z_1\right] + A_1^T\lambda_1 \\ M_{meas,2}\left[\Phi_2 - z_2\right] + A_2^T\lambda_2 \\ A_1\Phi_1 - F_1 \\ A_2\Phi_2 - F_2 \\ \vdots \\ \sum_{i=1}^{I} \lambda_i\partial_\Sigma A\Phi_i \end{bmatrix} \tag{3.50}$$

3. Schur Complement Modification

Every time another experiment is considered, another flux and adjoint correpsonding to that experiment will provide additional matrix equations in the Schur Complement. This final equation for $\delta\Sigma$ can be simplified and expressed as:

$$\delta\Sigma = \left[\sum_{i=1}^{I} S_i\right]^{-1} \cdot \left[\sum_{i=1}^{I} U_i\right] \tag{3.51}$$

where S_i is the Schur complement and U_i is the corresponding right hand side from Equations 3.37-3.39.

G. Multigroup Analysis

The final modifications to the code account for multiple energy groups of neutrons. This multi-group code will allow reconstruction of multigroup cross sections. This model includes reconstruction of fission cross sections, fission spectrum, group removal cross section, and intergroup scattering cross sections. The optimality conditions are again derived including first derivatives and the Hessian involving the second derivatives taken with respect to each of the new variables. It is supposed that this

model would enable greater acquisitions of realistic data that can be used to detect materials inside these multiple optical thickness thick objects.

Multigroup diffusion theory has more meaningful application to the problem at hand. The incident neutron beams can be classified due to their energy and with material constants for uranium, or other fissile material, such as χ_g, the fission cross sections for each group can be reconstructed to determine if fissile material is present in the cargo container. One example of a test case would be if the incident neutrons were only in the slow energy groups, but neutrons in fast energy groups were detected. Due to the nature that neutrons only have a reasonable probability to upscatter in the thermal Maxwellian range, those neutrons must have been born in the domain. That would be a greater chance of determination of SNM.

1. Multigroup Diffusion Theory

In neutron diffusion theory, neutrons can be classified by their energy and broken into groups. Due to scattering and fission, neutrons are able to be redistributed in energy based on the magnitude of their cross sections and fission spectrum, χ_g. In neutron diffusion theory equations of each group of neutrons can be formed with scattering terms that represent timerate densities of group to group scattering events.

$$-\nabla \cdot D_g \nabla \Phi_g + \Sigma_{r,g}\Phi = \chi_g \sum_{g'=1}^{G} \nu\Sigma_{f,g'}\Phi_{g'} + \sum_{g'=1, g\neq g'}^{G} \Sigma_{s,g'\to g}\Phi_{g'} \tag{3.52}$$

Where $\Sigma_{r,g} = \Sigma_{a,g} + \sum\limits_{g'=1,g\neq g'}^{G} \Sigma_{s,g\to g'}\Phi_g$ or removal from the group g due to absorption and outscatter. An example, the 2 group diffusion operator is given below:

$$A = \begin{bmatrix} -\nabla \cdot D_1 \nabla + \Sigma_{r,1} - \chi_1 \nu\Sigma_{f,1} & -\Sigma_{s,2\to 1} - \chi_1 \nu\Sigma_{f,2} \\ -\Sigma_{s,1\to 2} - \chi_2 \nu\Sigma_{f,1} & -\nabla \cdot D_2 \nabla + \Sigma_{r,2} - \chi_2 \nu\Sigma_{f,2} \end{bmatrix} \tag{3.53}$$

Everything remains unchanged for the optimization problem, except there are more parameters (Φ_g, λ_g, and Σ_g) for multiple energy groups creating the same $A\Phi = F$ system. The diffusion operator is no longer symmetric, $A^T \neq A$, because of scattering and fission.

2. Cross-Section Data for Various Materials

To gain a greater understanding of the reconstruction length-scale with respect to different materials, macroscopic cross sections for various materials are computed at fission spectrum average and 14 MeV energies. Using these cross sections, diffusion coefficients and diffusion lengths can be compared for different materials that may be present in a container. The macroscopic cross sections for various materials can be seen in Tables III-I - III-II[14].[15]

The steel composition used consisted of: 65.8% iron, 20.5% aluminum, 13.6% chromium, and 0.03% carbon. If fissile material were present, it may be shielded with a strong absorbing material such as borated polyethylene. Enriched boron is assumed in these computations at 90% B-10 and assuming uranium enriched to 80% U-235. The thermal, fission spectrum average, and 14.0 MeV macroscopic cross sections are computed and shown in Tables III-III - III-V[14].[15]

Table III-I. Cross-section data for various materials at 14.0 MeV

Material	Composition	Density [g/cc]	Σ_t [cm^{-1}]	Σ_s [cm^{-1}]	Σ_a [cm^{-1}]	D [cm]	L [cm]
Wood	CH_2O	0.5	0.0431	0.0403	0.0027	7.7299	52.6855
Fresh Wood	$(CH_2O)_{0.85} + (H_2O)_{0.15}$	0.575	0.0516	0.0482	0.0033	6.4505	43.5922
Plastic	C_2H_3Cl	1.39	0.0898	0.0766	0.0132	3.7094	16.7477
Steel	(see text)	7.85	0.2246	0.1618	0.0628	1.4834	4.8588
Aluminum	Al	2.7	0.1036	0.0742	0.0294	3.2159	10.4564

Table III-II. Cross-section data for various materials at fission spectrum average

Material	Composition	Density [g/cc]	Σ_t [cm^{-1}]	Σ_s [cm^{-1}]	Σ_a [cm^{-1}]	D [cm]	L [cm]
Wood	CH_2O	0.5	0.0821	0.0821	1.404E-05	4.0576	537.4704
Fresh Wood	$(CH_2O)_{0.85} + (H_2O)_{0.15}$	0.575	0.0990	0.0990	1.1939E-05	3.3659	530.9608
Plastic	C_2H_3Cl	1.39	0.1973	0.1973	3.7514E-05	1.6890	212.1866
Steel	(see below)	7.85	0.2643	0.2627	0.0015	1.2610	28.2883
Aluminum	Al	2.7	0.1963	0.19283	0.0035	1.6972	21.8490

Table III-III. Cross-section data for HEU in borated polyethylene at 14.0 MeV

Material	Composition	Density [g/cc]	Σ_t [cm^{-1}]	Σ_a [cm^{-1}]	Σ_f [cm^{-1}]	D [cm]	L [cm]
Polyethylene	CH_2	0.94	0.1085	0.0029	0.0	3.0715	32.3299
Borated Polyethylene	$(CH_2)_{0.95} + (B)_{0.05}$	0.978	0.1119	0.0042	0.0	2.9774	26.3211
HEU	U	19.0	0.3679	0.1318	0.0503	0.9058	2.6206
50/50 B.Poly./HEU	-	10.014	0.2399	0.0680	0.0251	1.3891	4.5165
90/10 B.Poly./HEU	-	2.785	0.1375	0.0170	0.0050	2.4232	11.9190

Table III-IV. Cross-section data for HEU in borated polyethylene at fission spectrum average

Material	Composition	Density [g/cc]	Σ_t [cm^{-1}]	Σ_a [cm^{-1}]	Σ_f [cm^{-1}]	D [cm]	L [cm]
Polyethylene	CH_2	0.94	0.2974	0.0060	0.0	1.1206	13.6610
Borated Polyethylene	$(CH_2)_{0.95} + (B)_{0.05}$	0.978	0.2984	0.0028	0.0	1.1169	19.8869
HEU	U	19.0	0.3843	0.1979	0.0448	0.8673	2.0933
50/50 B.Poly./HEU	-	10.014	0.3413	0.1003	0.0224	0.9764	3.1190
90/10 B.Poly./HEU	-	2.7852	0.3070	0.0223	0.0044	1.0856	6.9722

Table III-V. Cross-section data for HEU in borated polyethylene at thermal energy

Material	Composition	Density [g/cc]	Σ_t [cm^{-1}]	Σ_a [cm^{-1}]	Σ_f [cm^{-1}]	D [cm]	L [cm]
Polyethylene	CH_2	0.94	1.8719	0.0324	0.0	0.1780	2.3416
Borated Polyethylene	$(CH_2)_{0.95} + (B)_{0.05}$	0.978	22.7187	20.9564	0.0	0.0146	0.0264
HEU	U	19.0	24.1287	23.5404	20.0570	0.0138	0.0242
50/50 B.Poly./HEU	-	10.014	23.4237	22.2484	10.0285	0.01423	0.0252
90/10 B.Poly./HEU	-	2.7852	22.8597	21.2148	2.0057	0.0145	0.0262

CHAPTER IV

RESULTS

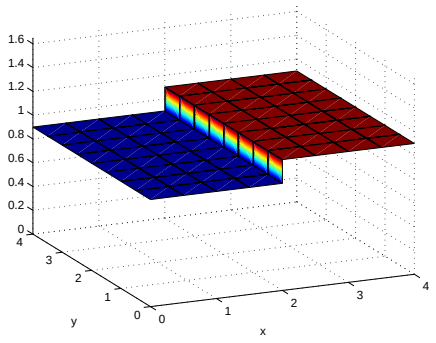
Several examples are presented to demonstrate the reconstruction limits with respect to different problem variables. The first example shows the importance of source and measurement location for imaging problems (Section A. The second example is a simple reconstruction of a homogeneous domain comparing convergence between the steepest descent and Newton's methods. Reconstructions with position dependence are presented next in (Section C), while the examples 4-5 (D) provides two cases displaying the effects on reconstruction resolution while increasing the number of experiments to demonstrate maximum reconstruction resolution. The example 6 shows the effects on reconstruction ability when the domain size is increased, which is tested using 8 and 32 experiments, to show the maximum reconstruction depth into a domain. Examples 7-11 include the addition of signal noise and signal bias to the synthetic measured solutions to better simulate actual detectors and the effect on reconstruction ability. Multigroup models are then shown, in examples 12-18, including various combinations of reconstructing multiple parameters (Σ_a and $\nu\Sigma_f$) in different energy groups.

There are three basic geometries that are used in the presented reconstruction examples. The first is a homogeneous domain which is used for basic reconstruction testing and misfit plotting in multigroup. The second consists of bars of various materials side by side. This simulates a piecewise constant one dimensional problem spread over a second dimension. The third is a centered strong absorber hidden in the center of a homogeneous domain, which is the typical material hidden inside another material example and is ideal for testing as the center of a domain typically has the

most error in reconstruction.

A. Example 1: Misfit Plots

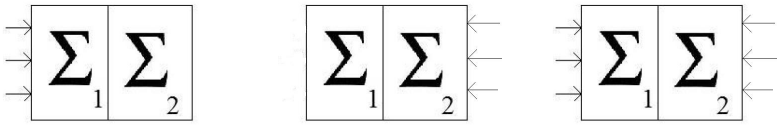
Consider a domain divided into two side-by-side homogeneous regions for which the cross section in each is to be reconstructed. An example of a such region is shown in Figure IV-1. This is a case of a two-parameter problem and a misfit surface plot can be graphed where the x-axis represents the inferred cross-section value in region 1, the y-axis represents the inferred cross-section value in region 2, and the z-axis represents the misfit as a function of the cross sections in the two regions. This graph will help understand the influence of the incoming source illumination on the ability to reconstruct one or both of the cross sections in this domain.



Region	D	Σ_a	$\nu\Sigma_f$
1	0.25	0.9	0.0
2	0.25	1.1	0.0

Fig. IV-1. Two-region domain of Example 1

The ability to reconstruct an internal parameter strongly depends on the illuminating source and measurement location. In this example measurements are taken on all sides and the source location is moved, similar results arise when the reverse case is considered. Two misfit surface plots were constructed using only one incident beam of neutrons on the side of one of the regions as shown in Figure IV-2 (cases a.) and b.)).



(a) Case a.) left incidence (b) Case b.) right incidence (c) Case c.) both beams present

Fig. IV-2. Ex. 1: Various cases of beam incidence

It can be noted in Figure IV-3, an elongated valley is produced for the misfit in the direction of the cross section whose side was not upon incident neutrons. This means that the cross section in this region can vary greatly while the misfit remains unchanged, illustrating the ill-posedness of the inverse problem.

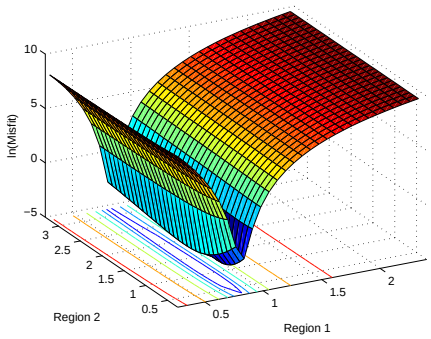


Fig. IV-3. Ex. 1: Misfit plot for case a.)

If the only incident neutron beam was moved to the side of the other region, an elongated valley is produced in the direction of the opposite cross section again as shown in Figure IV-4. Similarly in this case, the cross section in the rgion with no incoming neutron flux can vary greatly with respect to the other without significantly changing the true flux solution at the measurement points.

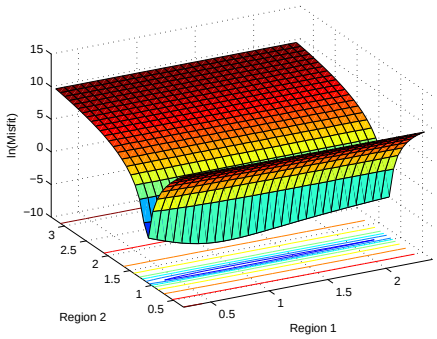


Fig. IV-4. Ex. 1: Misfit plot for case b.)

When both beams are on, one from each region side (case c.), a much smoother cone shape surface is produced as shown in Figure IV-5. It will be significantly easier to determine the cross section for both regions in this setting.

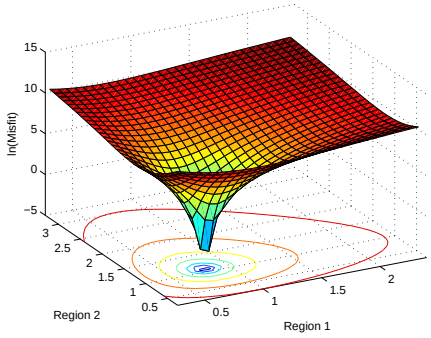
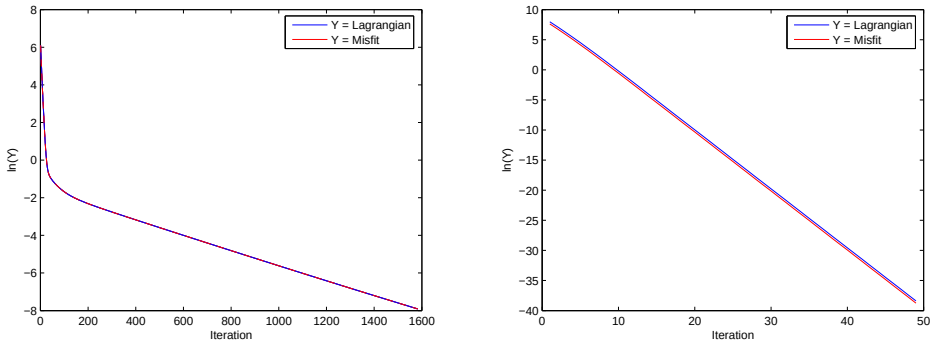


Fig. IV-5. Ex. 1: Misfit plot for case c.)

Having described and analyzed a 2-parameter case, the next models will include a fully position-dependent cross section. The information acquired here can be applied to the next model but cannot be shown visually. The model will have to find the minimum misfit of a multi-dimensional function so the incoming beams and measurements play a critical role in determining a solution and the ability to find this solution.

B. Example 2: Comparison of Convergence Between Steepest Descent Method and Newton's Method for a Homogeneous Problem

To show the decrease in misfit and Lagrangian as a function of iteration count, the optimization problem is run where these are stored for an example using a homogeneous domain consisting of: $D = 1$, $\Sigma_a = 0.3$ and $\nu\Sigma_f = 0$. This domain is of size 6 cm $\times$ 6 cm. These are simply stored then plotted after convergence, for this simple homogeneous problem starting relatively close ($\Sigma_a = 0.2$) to the true solution the misfit and Lagrangian are shown in Figure IV-6.



(a) Steepest Descent (b) Newton

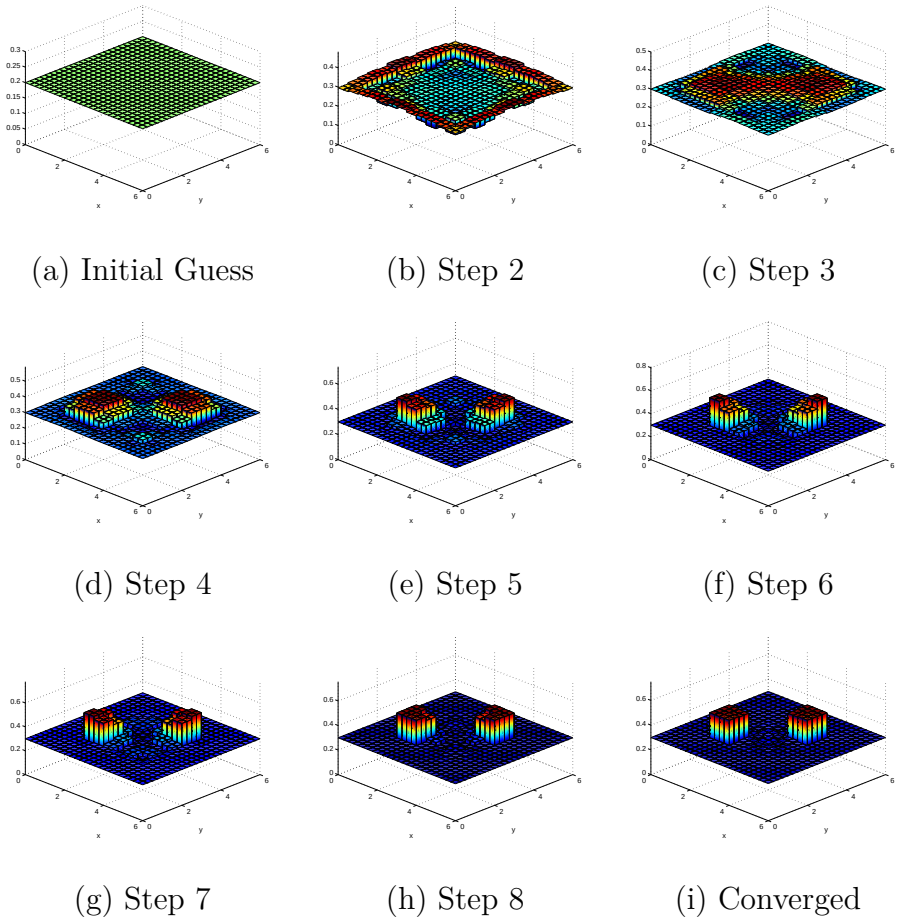
Fig. IV-6. Ex. 2: Convergence of misfit and Lagrangian for steepest descent and Newton's methods

As expected, these are both reduced and the reduction as a function of iteration count is exponential. Table IV-I shows a comparison of convergence statistics for this homogeneous domain. As expected, Newton's method outperforms the steepest descent method.

Table IV-I. Ex. 2: Convergence comparison between two line search methods

Method	Steepest Descent	Newton's
Time [s]	424.9	49.5
Iterations	1583	50
Order of $\mathcal{L}$	10^{-4}	10^{-17}

To better show convergence using Newton's method, a one-group problem consisting of a homogeneous domain is considered that contains two strong absorbers present inside and the cross section plotted at several steps during convergence. The initial guess for the domains cross section is a constant $\Sigma_a = 0.2$ where the true cross sections are $\Sigma_a = 0.3$ for the surrounding domain and $\Sigma_a = 0.5$ for the strong absorbers. An example of this problems convergence is shown in Figure IV-7 where each step represents approximately 10 iterations.



(a) Initial Guess (b) Step 2 (c) Step 3

(d) Step 4 (e) Step 5 (f) Step 6

(g) Step 7 (h) Step 8 (i) Converged

Fig. IV-7. Example 2: Convergence of dual strong absorbers in a homogenous domain

C. Example 3: Multiple Region Single Experiment Results

To reconstruct internal features of more complex domains, reconstructions are performed using using a mesh. The number of regions is determined using a coarse mesh based on the finer mesh used for the flux solver. This flux mesh (fine) consists of a variable refinement but a mesh of 2×2 finer than the material (coarse) mesh is often used. This enabled many various domains of varying cross section to be reconstructed at a user defined resolution.

The domain tested in this case was an 8×8 diffusion length domain consisting

of four vertical stripes of different materials which properties shown in Table IV-II. The true cross section of this domain is shown in Figure IV-8.

Table IV-II. Ex. 3: Cross-section data

Region	D	Σ_a	Σ_f
1	0.25	0.9	0.0
2	0.25	1.0	0.0
3	0.25	1.1	0.0
4	0.25	0.9	0.0

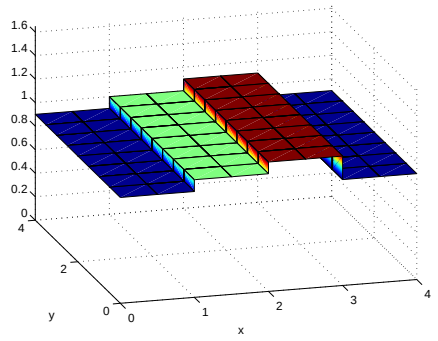


Fig. IV-8. Ex. 3: True cross section with position dependence

The coarse mesh here was setup as 8 by 8 in x and y directions respectively creating 64 regions with cross sections to be reconstructed. There was a refinement of 2 in both directions for the flux solver. This model converged after 8 Newton iterations with a misfit of 3.616e-15 and the optimality conditions were driven down to 1.06285e-11 in 1.47 seconds. The reconstructed flux and cross section can be shown in Figure IV-9.

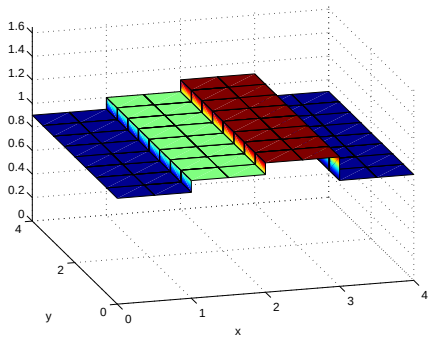


Fig. IV-9. Ex. 3: Reconstructed cross section with position dependence

To graphically show the precision of the reconstruction the difference between the true solution and reconstruction was taken and shown in Figure IV-10.

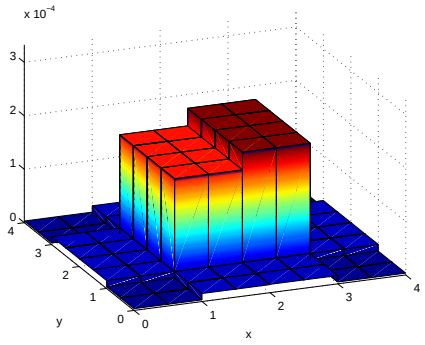


Fig. IV-10. Ex. 3: Error in cross section reconstruction with position dependence

This reconstructed solution agreed with the measured solution well as the error was on the order of 10^{-4}.

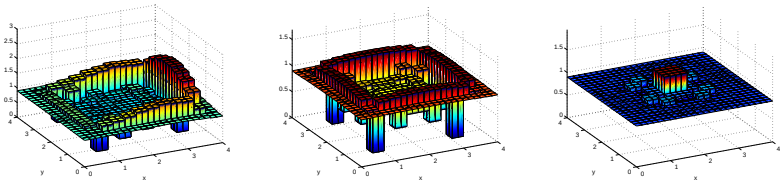
D. Multiple Experiment Results

When the incoming currents were split across the domain creating separate experiments, each with separate flux and adjoint solutions, the same set of cross sections for

the domain are reconstructed. This created more constraints on the data resulting in greater reconstruction ability.

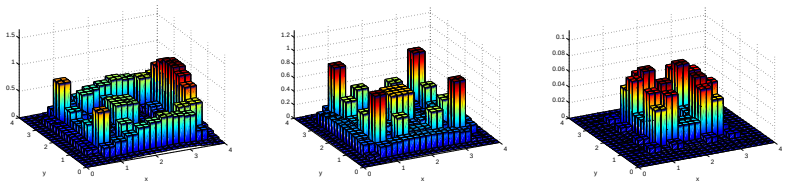
1. Example 4: Reconstruction Resolution Testing with Increasing Number of Experiments on a 4 cm × 4 cm Domain

The first example case shows the correlation between reconstruction resolution and the number of experiments on a domain consisting of a strong centered absorber. This domain consisted of $\Sigma_a = 1.3$, $\nu\Sigma_f = 0$ and $D = 0.3$ in the center region and $\Sigma_a = 0.9$, $\nu\Sigma_f = 0$ and $D = 0.3$ in the outer region. Three cases are considered for this example all of the same domain and mesh size, the first consisting of one experiment having a beam on a single side of the domain. The second case still consists of a single experiment but a beam on all four sides of the domain. The third case consists of eight experiments, each covering one half of each side. Measurements are made on all four sides every time. These cases reconstructions are shown in Figure IV-11 and corresponding errors in Figure IV-12.



(a) 1 Exp. 1 Beam (left) (b) 1 Exp. 4 Beams (1 beam per side) (c) 8 Experiments (each beam covering half a side

Fig. IV-11. Ex. 4: Effects on reconstruction resolution while increasing the number of experiments on a 4 cm × 4 cm domain



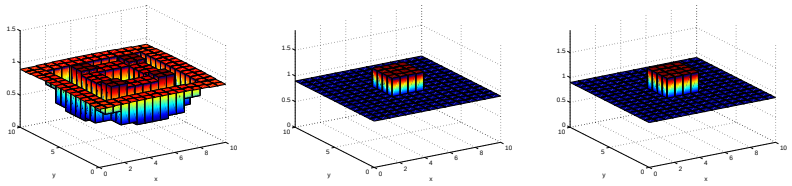
(a) 1 Exp. 1 Beam (left) (b) 1 Exp. 4 Beams (1 beam per side) (c) 8 Experiments (each beam covering half a side)

Fig. IV-12. Ex. 4: Error in reconstructions resolution testing while increasing the number of experiments on a 4 cm $\times$ 4 cm domain

This shows that increasing the number of experiments directly improves the maximum reconstruction resolution, enabling previously unsuccessful cases to have successful reconstructions. While this may be misleading, every problem has its resolution limits independent of the number of experiments. When reconstruction limits are approached, often additional experiments consist of too much noise to add any additional resolution refinement.

2. Example 5: Reconstruction Resolution Testing with Increasing Number of Experiments on a 10 cm $\times$ 10 cm Domain

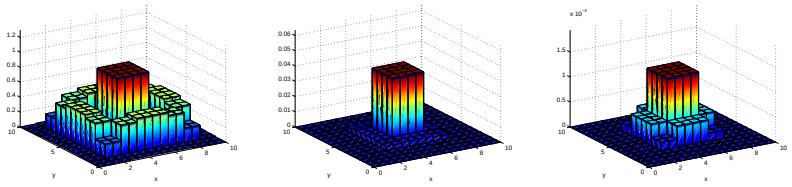
The next test example examines the effects of multiple experiments on reconstruction resolution for a larger domain. The same example problem of the centered strong absorber was considered here and reconstructions performed with 1, 8, and 32 experiments on a approximately 10 cm by 10 cm domain. With only 1 experiment consisting of beams on all four sides of the domain, Figure IV-13 shows the associated reconstruction with significant error and almost no recognizable features of the true solution.



(a) 1 Experiment (1 beam per side) (b) 8 Experiments (each beam covering half a side) (c) 32 Experiments (each beam covering $\frac{1}{8}$th of a side)

Fig. IV-13. Ex. 5: Reconstruction resolution testing of centered strong absorber in a 10 cm $\times$ 10 cm domain

When the number of experiments was set to eight, each covering one eighth of the domain or two per side this resulted in a better reconstruction where the error at the center is on the order of 10^{-2}. When the number of experiments was set to 32, each covering $\frac{1}{32}$ of the domain or eight per side this resulted in a much better reconstruction where the error at the center is on the order of 10^{-3} shown in Figure IV-14.

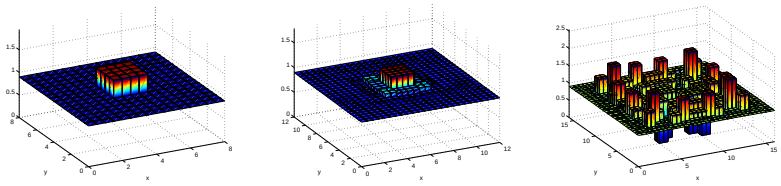


(a) 1 Exp. Error (1 beam per side) (b) 8 Exp. Error (each beam covering half a side) (c) 32 Exp. Error (each beam covering $\frac{1}{8}$th of a side)

Fig. IV-14. Ex. 5: Error in reconstructions for resolution testing of centered strong absorber in 10 cm $\times$ 10 cm domain

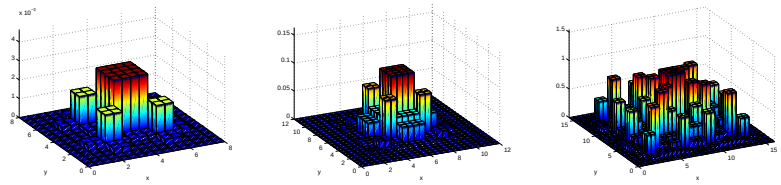
3. Example 6: Effects on Reconstruction When the Domain Size is Increased Using 8 and 32 Experiments

Another example problem, with the number of experiments held constant at eight, shows the effects of increased domain size on reconstruction ability. In this example, with the mesh size is held constant along with the number of experiments, the size of the domain is increased and the effects on reconstruction ability was observed. Three different sizes are used here approximately 8×8, 12×12, and 16×16 characteristic lengths. The reconstructions are shown in Figure IV-15 and the corresponding reconstruction error in Figure IV-16.



(a) 8 x 8 (b) 12 x 12 (c) 16 x 16

Fig. IV-15. Ex. 6: Effects on reconstructions when domain size is increased using eight experiments

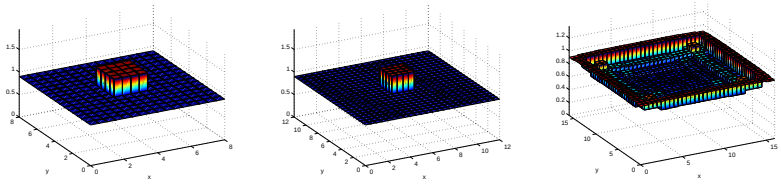


(a) 8 x 8 Error (b) 12 x 12 Error (c) 16 x 16 Error

Fig. IV-16. Ex. 6: Error in reconstructions when domain size is increased using eight experiments

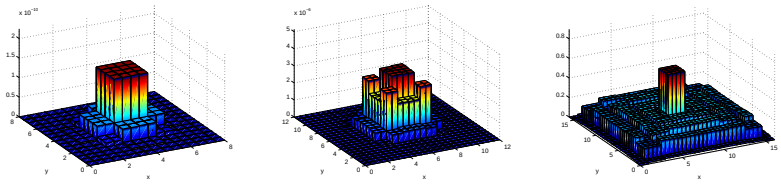
Using eight experiments, the it is only possible to reconstruct approximately

6-7 characteristic lengths deep into an object. As the domain size approached 16 characteristic lengths wide ($L = 1.054$cm in this example), there was too much signal degradation to reconstruct any domain characteristics. These same cases were run with 32 experiments, and the changes in reconstruction are shown below in Figure IV-17 and error Figure IV-18.



(a) 8 x 8 (b) 12 x 12 (c) 16 x 16

Fig. IV-17. Ex. 6: Effects on reconstructions when domain size is increased using 32 experiments



(a) 8 x 8 Error (b) 12 x 12 Error (c) 16 x 16 Error

Fig. IV-18. Ex. 6: Error in reconstructions when domain size is increased using 32 experiments

The implementation of 32 experiments shows improvement for the 12×12 case, while the 16×16 case showed no improvement. This proves that the addition of multiple experiments does provide additional reconstruction power, but reconstruction depth has severe limits independent of the number of experiments.

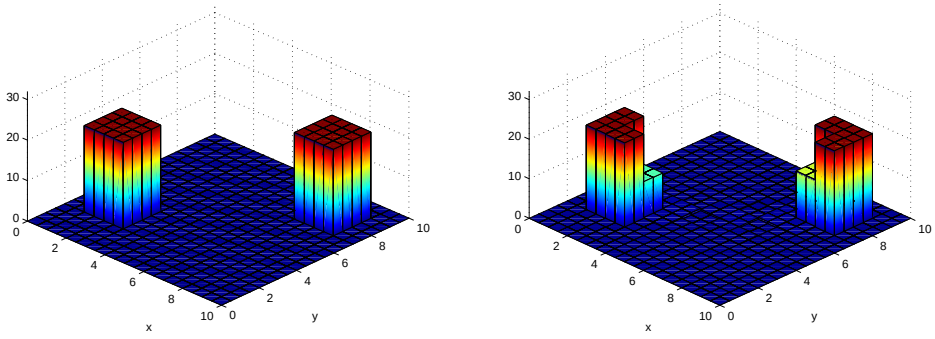
4. Example 7: Dual Strong Absorbers Embedded in a Large Highly Scattering Domain

The next domain reconstructed consisted of a large and highly scattering domain that contains two embedded identical strong absorbers. This domain size is over the previously mentioned maximum reconstruction size, but contains a very low absorption cross section. The domain parameters for the two regions are shown in Table IV-III.

Table IV-III. Ex. 7: Cross-section data for dual strong absorbers embedded in a large highly scattering domain

Region	D	Σ_a	Σ_s
Absorbers	0.016	21.37	0.012
Domain	0.467	0.0005	0.713

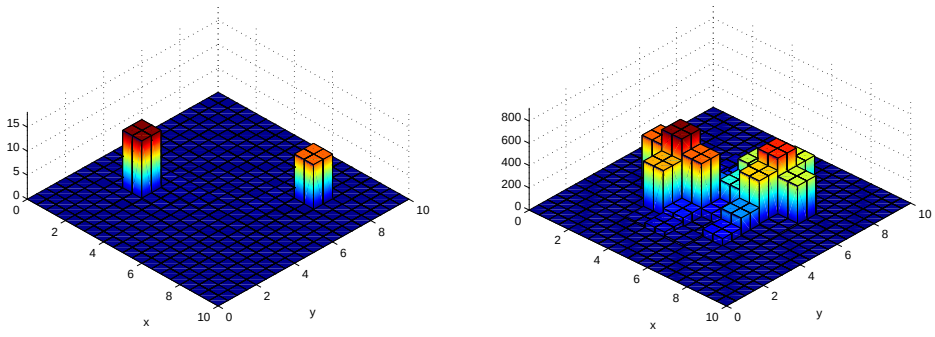
The domain is 10 cm $\times$ 10 cm where the outer region resembles carbon properties, while the absorbers embedded inside it resemble boron properties. The absorption cross section is reconstructed for this domain with an initial guess of a constant $0.008 cm^{-1}$. This reconstruction is performed using eight experiments on a coarse mesh of 10 $\times$ 10 with a refinement of 2 in both dimensions whose true and reconstructed cross sections are shown in Figure IV-19.



(a) True Cross Section (b) Reconstruction

Fig. IV-19. Ex. 7: Reconstruction of two strong absorbers in a large highly scattering domain

While this reconstruction may look very accurate, the magnitude of the parameters hides the reconstruction error. To better show this effect, the absolute error along with the relative error are shown in Figure IV-20.



(a) Absolute Error (b) Relative Error

Fig. IV-20. Ex. 7: Error in reconstruction of two strong absorbers in a large highly scattering domain

From the absolute error only, this reconstruction appears to be of excellent quality, but the relative error shows the cross sections in the center regions contain great

variance from the true solutions.

E. Addition of Signal Noise and Bias

To better simulate an actual boundary detector, signal noise and bias have been implemented into the synthetic measured solutions to simulate an actual radiation detector. The addition of noise results in a perturbation of the the synthetic measured solutions z, by a specified percentage using random numbers. The application of a bias would be similar but just averaged in a certain direction either positive or negative as opposed to zero. High accuracy boundary fluxes, or measured solutions, are crucial to achieve quality reconstructions, so as noise and bias are increased, it is expected that reconstruction ability will diminish. This amount of noise was varied and the effect on reconstruction observed.

While there are several ways to implement signal bias and noise, a constant noise distribution and a percentage based bias are used here shown in Equations 4.1-4.2.

$$z_{i,\text{noise}} = z_i(1 + \beta\epsilon_i) \tag{4.1}$$

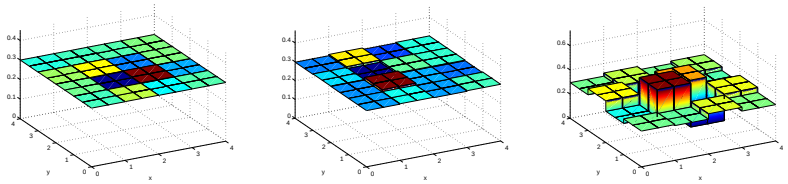
$$z_{i,\text{biased noise}} = z_i(1 + \delta + \beta\epsilon_i) \tag{4.2}$$

In Equation 4.1, z_i is the clean synthetic measured solution, β is the magnitude of the noise and $\epsilon_i \in (-1, 1)$ is the random number for measurement location i. In Equation 4.2, the same variable magnitude of noise is present but the addition of a flat percentage bias, $\delta \in (-1, 1)$, is added. The magnitude of δ is constant for all measurement points.

1. Addition of Random Signal Noise

a. Example 8: Signal noise added on a homogenous domain

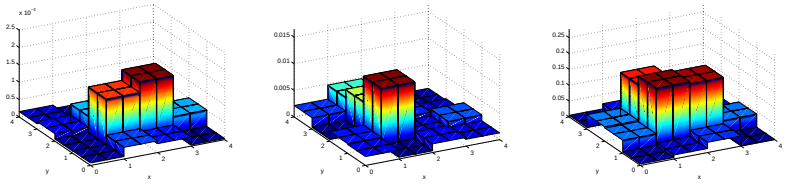
The optimal initial benchmark case would be one who's noiseless reconstruction ability is very high. The problem selected is a 4 cm × 4 cm homogeneous domain broken into 16 regions to reconstruct. This domain has the following properties: $\Sigma_a = 0.3$, $D = 1$, and $\nu\Sigma_f = 0$. All of the signal noise examples are reconstructed using eight experiments as this quantity of experiments is sufficient for noiseless reconstructions in all of the problems used. Three quantities of noise are applied to this domain of magnitudes of 0.01%, 0.1%, and 1.0% and reconstruction results shown in Figure IV-21.



(a) 0.01% Signal Noise (b) 0.1% Signal Noise (c) 1.0% Signal Noise

Fig. IV-21. Ex. 8: Reconstruction with signal noise of a homogeneous domain

The case of β =0.0% noise is not shown due to the high reconstruction ability of the β =0.01% case it is obvious what the true cross-section solution is. With β =0.01% and β =0.1% noise reconstruction ability is still high but as the noise magnitude approaches 1.0%, even this simple homogenous domain becomes almost indistinguishable. The resulting reconstruction errors are shown in Figure IV-22.



(a) 0.01% Signal Noise Error (b) 0.1% Signal Noise Error (c) 1.0% Signal Noise Error

Fig. IV-22. Ex. 8: Error in reconstruction with signal noise of a homogeneous domain

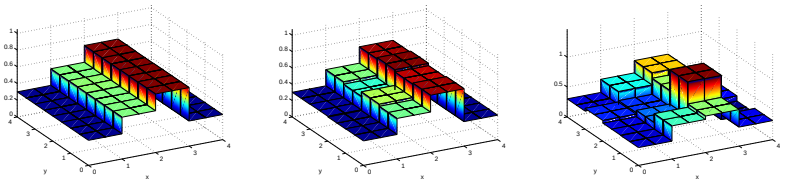
b. Example 9: Signal noise added to bars of various materials

The next problem implemented with signal noise is similar to Ex. 3 (Section C) which had high reconstruction ability and consisted of bars of various materials whose material properties are shown in Table IV-IV.

Table IV-IV. Ex. 9: Cross-section data for bars of various material

Region	D	Σ_a	Σ_f
1	1.0	0.3	0.0
2	1.0	0.5	0.0
3	1.0	0.7	0.0
4	1.0	0.3	0.0

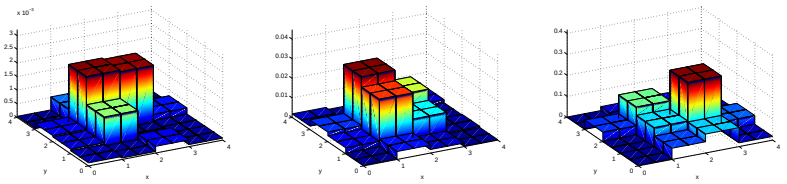
This domain size is approximately 4 × 4 characteristic lengths thick and is reconstructed using eight experiments. Due to the noiseless high reconstruction ability of this example, the same 0.01%, 0.1% and 1.0% quantities of signal noise are added and reconstructions shown in Figure IV-23.



(a) 0.01% Signal Noise (b) 0.1% Signal Noise (c) 1.0% Signal Noise

Fig. IV-23. Ex. 9: Reconstruction with signal noise of multiple materials

Just as in the case of the homogeneous domain, the 0.01% and 0.1% signal noise examples still provided quality reconstructions while features of the domain start to diminish in the center at 1.0% signal noise as shown in the resulting reconstruction error in Figure IV-24.



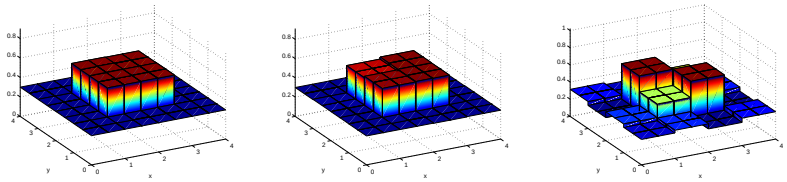
(a) 0.01% Signal Noise Error (b) 0.1% Signal Noise Error (c) 1.0% Signal Noise Error

Fig. IV-24. Ex. 9: Error in reconstruction with signal noise of multiple materials

c. Example 10: Reconstruction testing with signal noise on a centered strong absorber inside various size domains

The final example implemented with signal noise was the previously defined centered strong absorber. This was tested at multiple noise magnitudes along with multiple domain sizes in three cases approaching the maximum reconstruction size. This domain consists of a strong absorbing center region surrounded by a weaker absorbing domain. The center region has the following material properties $\Sigma_a = 0.6$, $\nu\Sigma_f = 0$,

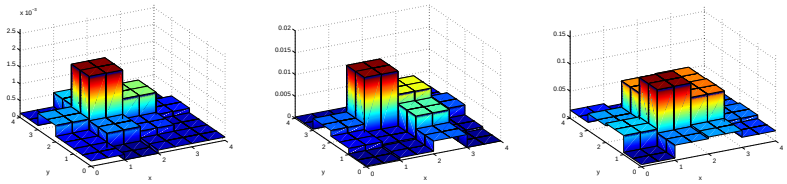
and $D = 1.0$, while the surrounding domain has $\Sigma_a = 0.3$, $\nu\Sigma_f = 0$, and $D = 1.0$. Three domain sizes are used here and reconstructions of the 4 cm $\times$ 4 cm ($L \approx 1$) domain are shown in Figure IV-25. In this example, due to the domains small size, reconstruction ability is high in the noiseless case so noise magnitudes of 0.01%, 0.1% and 1.0% are applied.



(a) 0.01% Signal Noise Error (b) 0.1% Signal Noise Error (c) 1.0% Signal Noise Error

Fig. IV-25. Ex. 10 Case 1: Reconstruction with signal noise of a centered strong absorber 4x4cm

As in the previous cases of similar domain size, the reconstruction ability diminished when the noise reaches 1.0% and reconstruction erros are shown in Figure IV-26.

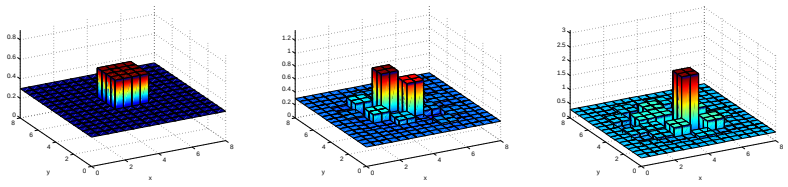


(a) 0.01% Signal Noise Error (b) 0.1% Signal Noise Error (c) 1.0% Signal Noise Error

Fig. IV-26. Ex. 10 Case 1: Error in reconstruction with signal noise of a centered strong absorber 4x4cm

The next case holds reconstruction resolution and material properties the same as the previous case of the centered absorber except the domain size is increased to 8

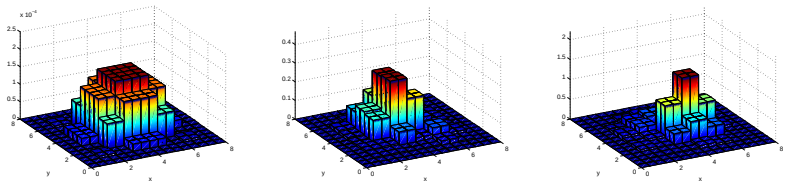
cm × 8 cm ($L \approx 1$). The material properties of both regions are the same as the 4 × 4 case. Two magnitudes of signal noise are applied in this case of 0.01% and 0.05% and reconstruction results are shown in Figure IV-27 along with the noiseless case.



(a) 0.0% Signal Noise Error (b) 0.01% Signal Noise Error (c) 0.05% Signal Noise Error

Fig. IV-27. Ex. 10 Case 2: Reconstruction with signal noise of a centered strong absorber 8x8cm

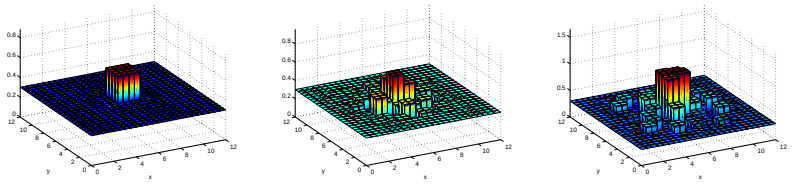
In the 0.01% signal noise case for this size domain, the model shows reasonably that there is a stronger absorber in the center. While the size of this absorber is incorrect and slightly smeared throught the center, the magnitude is a good approximation. In the 0.05% signal noise case, the entire centered absorber is smeared across the middle including one region whos magnitude is too high by a factor of 3.5. The error in this case is too high to consider this a successful reconstruction, as the features are hardly distinguishable. Thes reconstruction errors are shown in Figure IV-28.



(a) 0.0% Signal Noise Error (b) 0.01% Signal Noise Error (c) 0.05% Signal Noise Error

Fig. IV-28. Ex. 10 Case 2: Error in reconstruction with signal noise of a centered strong absorber 8x8cm

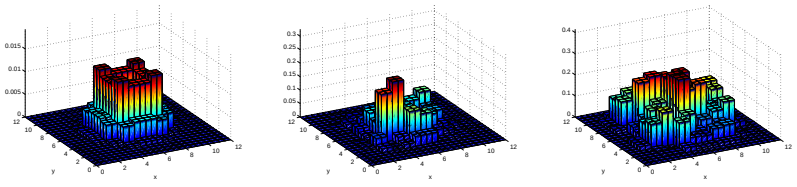
The third case was a domain of 12 cm× 12 cm in size ($L \approx 1$). This is close to the domain size reconstruction limit with no signal noise so reconstructions with any signal noise are expected to be poor. The same material properties of the 4 × 4 and 8 × 8 case are used. Two magnitudes of signal noise are applied in this case of 0.0005% and 0.005% and reconstruction results are shown in Figure IV-29 along with the noiseless case.



(a) 0.0% Noise (b) 0.0005% Noise (c) 0.005% Noise

Fig. IV-29. Ex. 10 Case 3: Reconstruction with signal noise of a centered strong absorber 12x12cm

This case already had limited reconstruction ability with 0.0% noise, therefore the addition of signal noise is expected to cause poor reconstructions. This case began to lose distinguishable features as error increased over 0.005% and resulting reconstruction error is shown in Figure IV-30.



(a) 0.0% Noise Error (b) 0.0005% Noise Error (c) 0.005% Noise Error

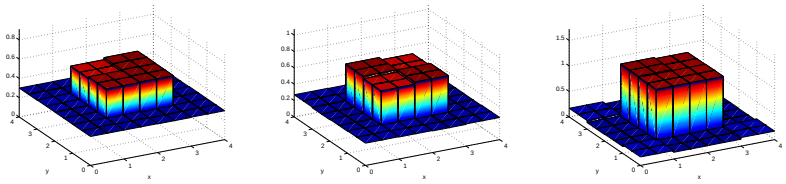
Fig. IV-30. Ex. 10 Case 3: Error in reconstruction with signal noise of a centered strong absorber 12x12cm

2. Addition of Signal Bias

The next example was implemented with signal bias was the previously defined centered strong absorber. This was tested at multiple noise strengths along with domain sizes up to approximately the maximum reconstruction size. One of the same domains tested with signal noise is chosen here and two problems tested with a bias with additional signal strength and attenuated signal strength. In these cases a signal bias is added of a specified magnitude, then random noise added of another magnitude to simulate an actual signal bias with variable variance and magnitude.

a. Example 11: Reconstructing with a positive signal bias of a centered strong absorber

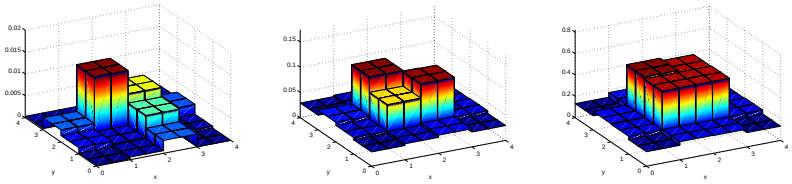
In the first case a positive signal bias of various magnitudes was added to the synthetic measured solutions then random signal noise of a 0.1% applied, and an attempt at reconstructions performed. The test domain was the 4×4 characteristic length centered strong absorber. This domain's properties consist of $\Sigma_a = 0.6$, $\Sigma_f = 0$, and $D = 1.0$ in the center and $\Sigma_a = 0.3$, $\Sigma_f = 0$, and $D = 1$ in the surrounding area. The positive bias is applied in two magnitudes of 1.0% and 5.0% and resulting reconstructions shown in Figure IV-31 including the 0.0% bias 0.1% noise case is shown for comparison.



(a) 0.0% Signal Bias (b) 1.0% Signal Bias (c) 5.0% Signal Bias

Fig. IV-31. Ex. 11: Reconstruction with a positive signal bias of a centered strong absorber

When a signal bias of 1.0% is added to the synthetic measured solution, the measured solution is higher then expected leading to a cross section that provides less attenuation. In the region surrounding the center absorber, the cross section is smaller then the true value as expected however, in the center region the magnitude of the cross section is higher then the true cross section. This is mainly due to the fact that the signal bias is percentage based for each element not a flat bias throughout the measurements. When the signal bias is increased to 5.0%, the cross section in the surrounding region is again lowered, while the center region is again increased. These reconstruction errors are shown in Figure IV-32.

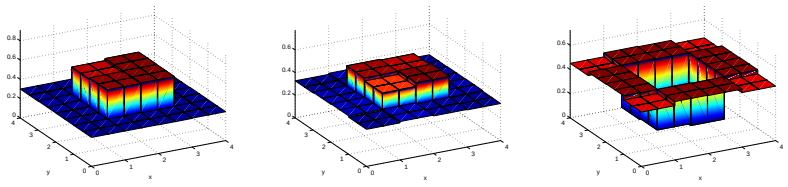


(a) 0.0% Bias Error (b) 1.0% Bias Error (c) 5.0% Bias Error

Fig. IV-32. Ex. 11: Error in reconstruction with a positive signal bias of a centered strong absorber

b. Example 12: Reconstructing with a negative signal bias on a centered strong absorber

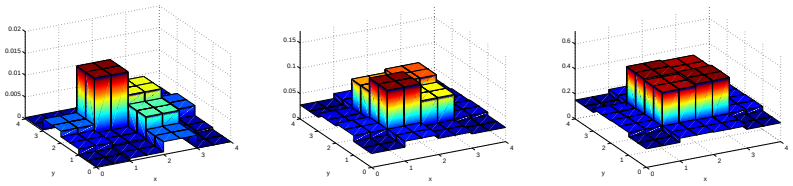
The same test as the positive signal bias, but with a negative signal bias instead. The same domain is tested here with -1.0% and -5.0% bias with 0.1% random signal noise. The same material properties as the positive bias are used here and reconstructions are shown in Figure IV-33.



(a) 0.0% Signal Bias (b) -1.0% Signal Bias (c) -5.0% Signal Bias

Fig. IV-33. Ex. 12: Reconstruction with a negative signal bias of a centered strong absorber

When a -1.0% bias is applied to the synthetic measured solutions, an abundance of attenuation is expected leading to an increase in reconstructed cross section. This is true for the surrounding region, this bias over accounts the attenuation in the surrounding region leaving the center region's cross section lower then expected. The same effects are shown in the -5.0% case to a greater effect and the reconstruction errors are shown in Figure IV-34.



(a) 0.0% Bias Error (b) -1.0% Bias Error (c) -5.0% Bias Error

Fig. IV-34. Ex. 12: Error in reconstruction with a negative signal bias of a centered strong absorber

F. Muli-group Results

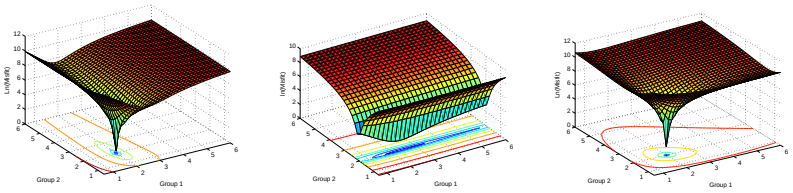
With the multigroup equations it is possible to reconstruct multiple parameters at once, for example the removal cross section of every group or a combination of removal cross section and fission cross section simultaneously. In the case of two parameters to be reconstructed, a homogeneous domain can be considered and a misfit surface plot can be constructed to observe the dimensional space of the misfit with respect to the two parameters.

1. Multi-group Misfit Plots

a. Example 13: Multigroup misfit plots of absorption cross sections only

The misfit surface code was modified to examine the variation of the misfit function in the case of a single region multigroup example with respect to various cross-section combinations. This enabled a visualization of reconstruction difficulty with respect to various parameter combinations. The first case that was tested consisted of a two group problem. $\chi_1 = 1$ and $\chi_2 = 0$, and the misfit was computed as a function of $\Sigma_{r,1}$ and $\Sigma_{r,2}$. This problem is very similar to the original one group problem in such there are two systems in energy coupled by parameters different than the ones being

reconstructed. Similarly to the one group models, $\Sigma_{r,i}$ can be reconstructed for any number of energy groups. With incident beams in both energy groups, this misfit surface plot is shown in Figure IV-35. As expected, with incoming beams in both groups, reconstruction of both parameters is fairly simple. If the incident neutrons were only in one group, for example the fast group, the ability to reconstruct was tested as well.



(a) Fast Group Incident Flux (b) Thermal Group Incident Flux (c) Both Group Incident Flux

Fig. IV-35. Ex. 13: Misfit surface plot of $\Sigma_{r,1}$ and $\Sigma_{r,2}$ for 2 group homogeneous region while varying source energy

In the case where incident neutrons are only in the fast group, the problem is more ill-posed as the misfit function has an elongated valley as opposed to a steep conic shape. While reconstruction ability is lower if incident neutrons are not in every group, the coupling of the groups provides enough data for reconstruction of parameters in the other groups. The misfit surface plot changed from a smooth conic shape to smeared cone in the $\Sigma_{r,2}$ direction. In the case where incident neutrons are only in the thermal group, an even more ill-posed problem is present. The reconstruction ability is strongly dependent on the coupling between the two group, down scattering or fission source, and when the source is weak in comparison to the incident neutrons, the problem may be very ill-posed. In multigroup models, incident beams in every group does add reconstruction ability.

b. Example 14: Multigroup misfit plots of fission cross sections only

The next problem space that was tested with the misfit surface plot consisted of $\Sigma_{f,1}$ and $\Sigma_{f,2}$ in a 2 group problem. When reconstructing fission cross sections, one must be careful due to physical limitations of the problem itself. When modifying $\nu\Sigma_f$, special care must be taken due to physical constraints such as criticality and $\Sigma_f \leq \Sigma_a$. In a realistic manner, if $\Sigma_f = \Sigma_a$ and $\nu\Sigma_f$ is still increased, due to Σ_f limitations would have to imply that ν was increasing. This can lead to criticality issues very quickly if conservative step length control is not implemented.

In this problem, when the fission cross section is modified by a large amount, and the multiplication factor of the domain becomes greater then one, the flux can result in a negative solution. The FEM diffusion solver implements the fission terms not as a fixed source but in the current solution, when the system approaches criticality, interesting results can occur such as negative fluxes in certain regions.

These physical limitations really limit the $\Sigma_{f,i}$ and $\Sigma_{f,j}$ workspace to a region around the true solution if they are the only parameters being modified. Now that physical limitations of simply modifying these parameters has been established, a visualization of the reconstruction space was performed. The misfit function was computed as a function of $\Sigma_{f,1}$ and $\Sigma_{f,2}$ of a homogeneous domain and results shown in Figure IV-36. This function space is much flatter than the Σ_r space where quadratic approximations in the Newtonian direction converged less slowly.

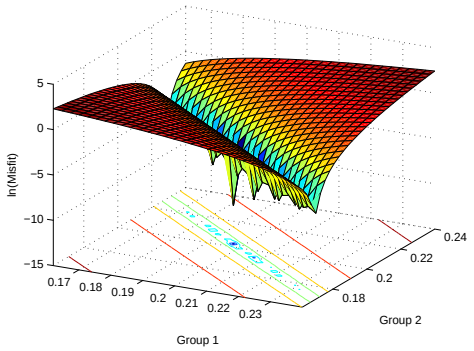


Fig. IV-36. Ex. 14: Misfit surface plot of $\Sigma_{f,1}$ and $\Sigma_{f,2}$ for 2 group homogeneous region

This $\Sigma_{f,i}$ vs. $\Sigma_{f,j}$ workspace has a misfit several orders of magnitude lower then the $\Sigma_{r,i}$ vs. $\Sigma_{r,j}$ workspace. This is partially due to the magnitude of the parameters themselves. Due to criticality concerns, any fissile material placed in a large container, would have to either be spread out or only be of a small quantity. From a reconstruction standpoint, this would lead to a fission cross section much lower then the total absorption cross section on the parameter grid due to the resolution in the reconstruction homogenization. In the misfit surface plot of the $\Sigma_{f,i}$ vs. $\Sigma_{f,j}$ workspace, the parameters can change significantly compared to the $\Sigma_{r,i}$ vs. $\Sigma_{r,j}$ workspace while maintaining a low misfit. This data can be extrapolated to the case of the centered fissile material expecting to limit the maximum reconstruction depth to be shorter than the non-fissile case. One extreme example of the ill-posedness of this type of problem is consisted of a very strong absorber with a fissile region with significantly lower fission cross section than the absorption cross section in the center of a domain, the fission rate, no matter what the fission cross section is, will have little effect on the boundary fluxes.

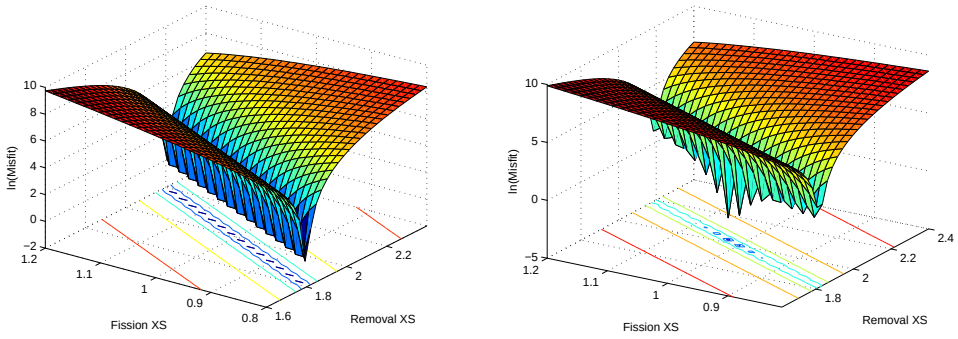
When χ is distributed across more then one group such as the previously mentioned 4 group example, similar improvements occur, such as the $\Sigma_{r,1}$ and $\Sigma_{f,1}$, but

the workspace for reconstruction is considerably smaller due to physical constraints.

c. Example 15: Multigroup misfit plots of mixed parameters

The next examples consisted of mixed parameters of type and group, for example $\Sigma_{r,i}$ and $\Sigma_{f,j}$. Reconstructions of this type are simpler then fission in multiple groups due to the parameters are coupled weaker with respect to their energy groups. For example, reconstruction of fission in both groups of two total groups, both parameters produce neutrons in group one and are strongly coupled by both the groups fluxes. While in a four group example, if $\Sigma_{f,4}$ and $\Sigma_{r,2}$ are reconstructed they are coupled weaker and are simpler to reconstruct such as reconstructions of single parameters are. Although mixed parameter reconstructions may have a limited workspace in the Σ_f dimension, reconstruction ability is significantly greater due to Σ_r flexibility.

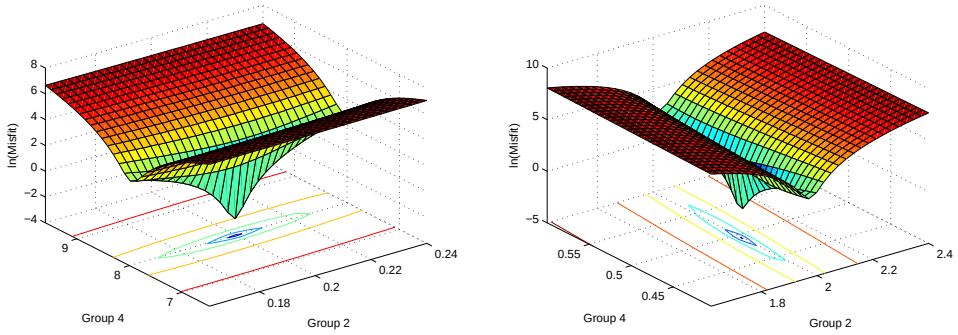
Next, the misfit is computed as a function of $\Sigma_{r,1}$ and $\Sigma_{f,1}$. This problem is very similar to the original one group problem in such that the addition of the second group provided no additional information to these two parameters as long as $\chi_2 = 0$. This function shows a line solution as the minimum as expected and is shown in Figure IV-37. If a 4 group problem is considered, and χ = [0.9 0.1 0.0 0.0] for the four corresponding groups and same misfit was computed, since $\chi_2 = 0.1$ additional information is provided to $\Sigma_{f,1}$. The resulting misfit surface plot is still an elongated valley, but instead of a line solution it has a minimum enabling reconstruction for this case.



(a) 2 Group (b) 4 Group

Fig. IV-37. Ex. 15: Misfit surface plot of $\Sigma_{r,1}$ and $\Sigma_{f,1}$ for 2 and 4 group homogeneous regions

One example of mixed parameter reconstruction is tested for a four group case. This example consists of $\chi = [0.9\ 0.1\ 0.0\ 0.0]$ where the misfit is computed as a function $\Sigma_{r,2}$ and $\Sigma_{f,4}$ then the same parameters with reverse groups and results shown in Figure IV-38.



(a) $\Sigma_{r,2}$ and $\Sigma_{f,4}$ (b) $\Sigma_{f,2}$ and $\Sigma_{r,4}$

Fig. IV-38. Ex. 15: Mixed parameter misfit surface plots for 4 energy groups

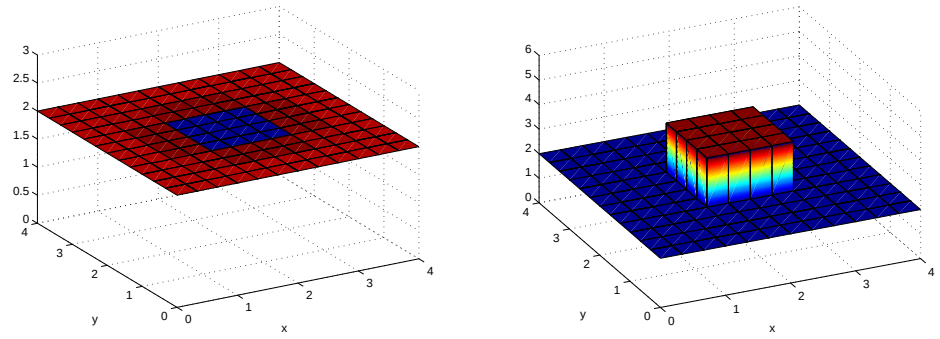
Instead of the typical Σ_r vs. Σ_r type conic surface, as shown before the misfit function is much flatter in the Σ_f space, so the misfit is expected to change slower

in the Σ_f direction of the Σ_r vs. Σ_f misfit plot. The same effect is shown when the parameters groups are switched.

2. Multigroup Reconstruction Results

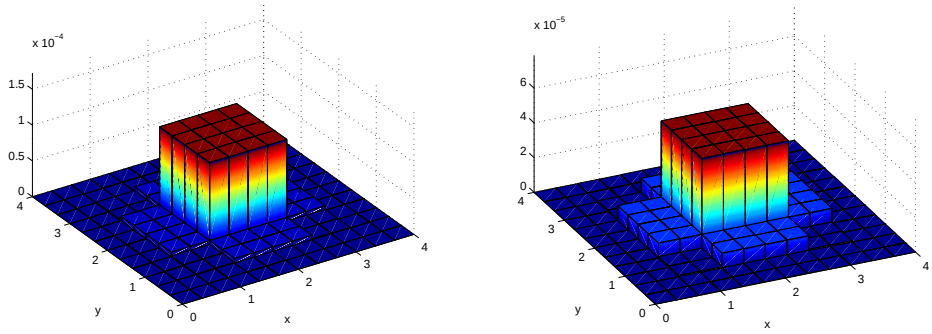
a. Example 16: Reconstruction of a thermal strong absorber

The first case consists of reconstruction of Σ_r for multiple groups. As mentioned previously, this example resembles weakly coupled one group cases solved simultaneously so high reconstruction ability is expected. This is essentially the base fundamental reconstruction and similar tests as the single group can be compared to test the reconstruction limits. The test case considered consisted of a homogeneous domain with a centered strong thermal absorber. This is similar to the previous centered strong absorber except that in this test case it only exists in group 2 of 2 total groups. The reconstruction for this problem is shown in Figure IV-39 and corresponding error in Figure IV-40.



(a) Group 1 Recon. (b) Group 2 Recon.

Fig. IV-39. Ex. 16: Multigroup reconstructions of a thermal centered absorber



(a) Group 1 Error (b) Group 2 Error

Fig. IV-40. Ex. 16: Error in multigroup recon. of a thermal centered absorber

Essentially the same reconstruction limits for one group problems are similar here for Σ_r reconstructions for multigroup problems. They can be viewed as separate systems only weakly coupled by fission and scattering terms and modification of the Σ_r does not have any limits with respect to the other cross sections making a large workspace in the reconstruction space.

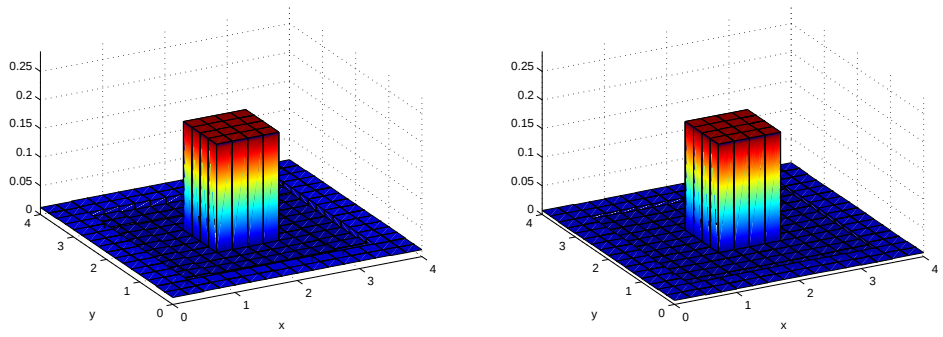
b. Example 17: Reconstruction of centered fissile material

The next example problem attempts to reconstruct the fission cross section of a centered strong fissile material. As mentioned before, the misfit space for the fission cross section is strongly dependent on the magnitude of the fission cross section and this example consists of a centered strong fissile region inside a domain. This problem consists of a coarse mesh of 8×8 or 64 regions to reconstruct and the domain properties shown in Table IV-V.

Table IV-V. Ex. 17: Domain parameters of centered fissile material

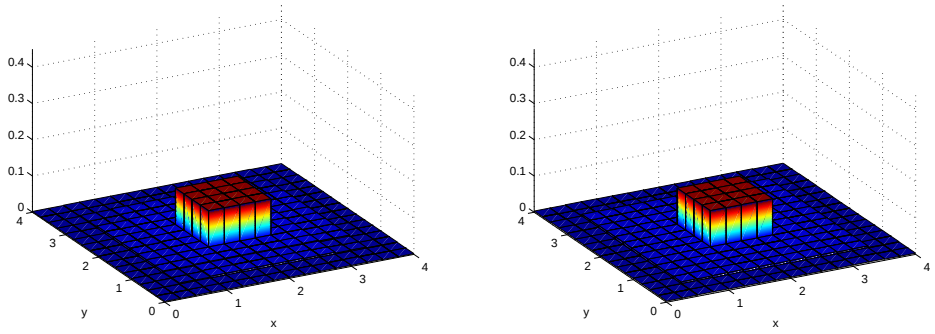
Region	D_1	D_2	$\Sigma_{a,1}$	$\Sigma_{a,2}$	$\Sigma_{f,1}$	$\Sigma_{f,2}$
Outer	0.5	0.5	2.0	3.0	0.01	0.01
Center	0.5	0.5	2.0	3.0	0.30	0.30

With an initial guess of $\Sigma_{f,1} = 0$ and $\Sigma_{f,2} = 0$ for the entire domain and, using eight experiments, the optimality conditions were iterated down to the order of 10^{-7}. This reconstruction is shown in Figure IV-41 and reconstruction error in Figure IV-42.



(a) Group 1 Recon. (b) Group 2 Recon.

Fig. IV-41. Ex. 17: $\Sigma_{f,1}$ and $\Sigma_{f,2}$ reconstructions of centered fissile material



(a) Group 1 Error (b) Group 2 Error

Fig. IV-42. Ex. 17: Error in $\Sigma_{f,1}$ and $\Sigma_{f,2}$ reconstructions of centered fissile material

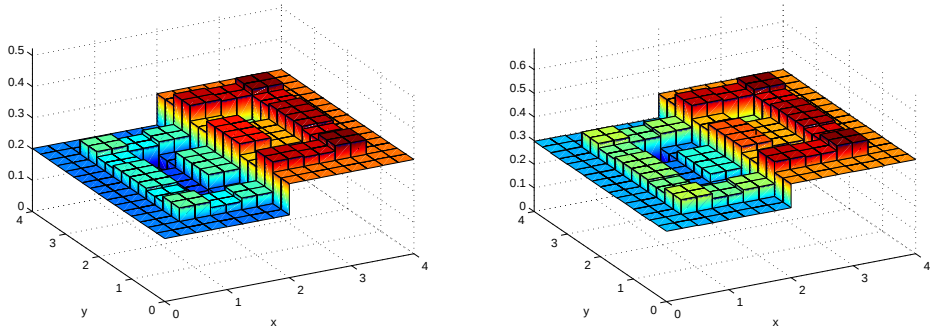
The reconstruction error is significantly greater then the previous removal or absorption cross-section reconstructions while the optimality conditions are driven down to the same magnitude. This is a result of the ill-posedness of the problem as shown by the flatness of the misfit plot. While the center region has the highest reconstruction error, it is still noticeable that there is stronger fissile material there.

c. Example 18: Maximum reconstruction depth testing for $\nu\Sigma_f$

Another fission reconstruction example demonstrates the maximum reconstruction depth. A simple domain of 8×8 consisting of side by side homogeneous pieces with a total of 64 regions of reconstruction is employed here. This domain has dimensions of approx 12 cm $\times$ 12 cm ($L \approx$ 1cm) and 16 experiments are used in this reconstruction each consisting of a beam covering one fourth of a side. This simple case is used as a benchmark to test fission cross-section reconstruction depth. This examples material properties are given in Table IV-VI.

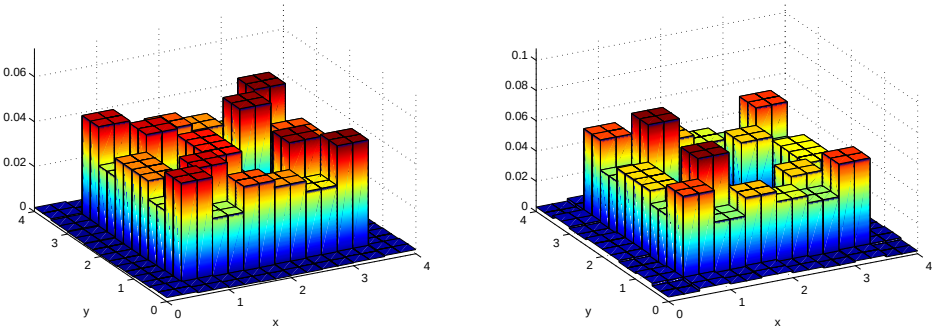
Table IV-VI. Ex. 18: Domain parameters of two region fissile material

Region	D_1	D_2	$\Sigma_{a,1}$	$\Sigma_{a,2}$	$\Sigma_{f,1}$	$\Sigma_{f,2}$
Left	0.33	0.33	2.0	3.0	0.2	0.3
Right	0.33	0.33	3.0	4.0	0.3	0.4



(a) Group 1 Recon. (b) Group 2 Recon.

Fig. IV-43. Ex. 18: $\Sigma_{f,1}$ and $\Sigma_{f,2}$ reconstructions of a two zone fissile step



(a) Group 1 Error (b) Group 2 Error

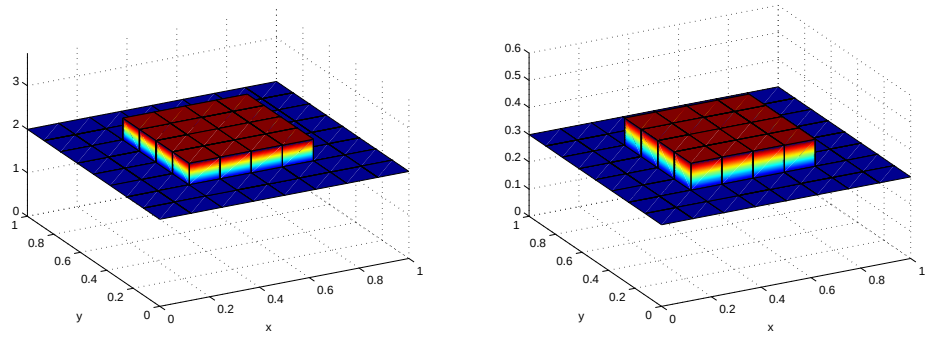
Fig. IV-44. Ex. 18: Error in $\Sigma_{f,1}$ and $\Sigma_{f,2}$ reconstructions of two zone fissile step

This reconstruction, in Figure IV-43 and error in Figure IV-44, was approxi-

mately the maximum depth any noticeable features can be distinguished for fission cross-section reconstructions. This reconstruction was approximately 6 characteristic lengths deep, or half the domain size, where noticeable internal features become indistinguishable. The maximum depth for fission parameter reconstruction is approximately 5-6 characteristic lengths which is 1-2 characteristic lengths shorter then removal cross-section reconstructions.

d. Example 19: Mixed parameter reconstructions

The final multigroup reconstruction problem consists of mixed parameter reconstructions. Due to the physical limitations on Σ_f with respect to Σ_a there is less freedom for parameter modification in these cases, causing the reconstruction space to be smaller. The first case was a 2-group problem where $\Sigma_{r,1}$ and $\Sigma_{f,2}$ are reconstructed. This example has the same paramters as shown in Table IV-VI. The domain has size of 1 cm $\times$ 1 cm ($L \approx 0.33$cm) and strong reconstruction ability is expected. Results from the reconstruction of $\Sigma_{r,1}$ and $\Sigma_{f,2}$ are shown in Figure IV-45.

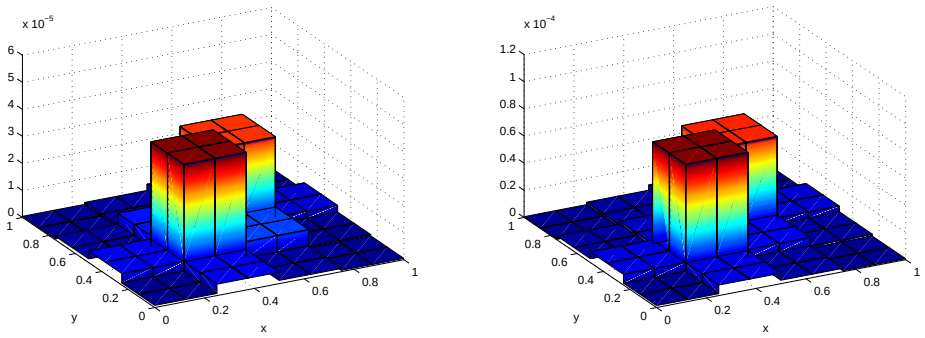


(a) $\Sigma_{r,1}$ (b) $\Sigma_{f,2}$

Fig. IV-45. Ex. 19: Mixed parameter reconstructions for a 2 group problem

As expected, the mixed parameter reconstruction is excellent and the error is

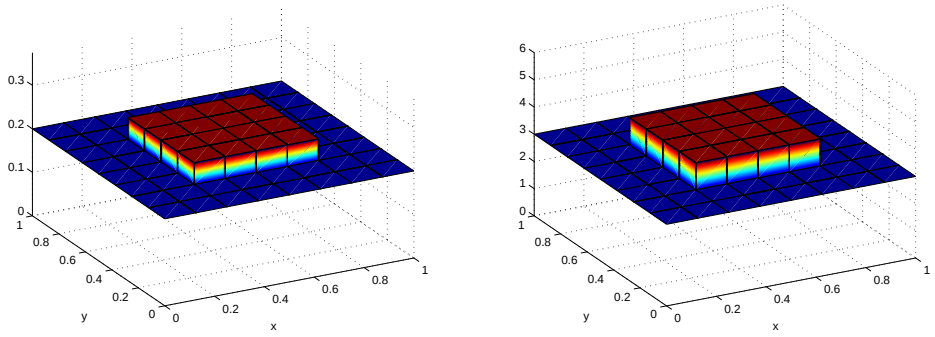
shown in Figure IV-46.



(a) $\Sigma_{r,1}$ (b) $\Sigma_{f,2}$

Fig. IV-46. Ex. 19: Error for mixed parameter reconstructions for a 2 group problem

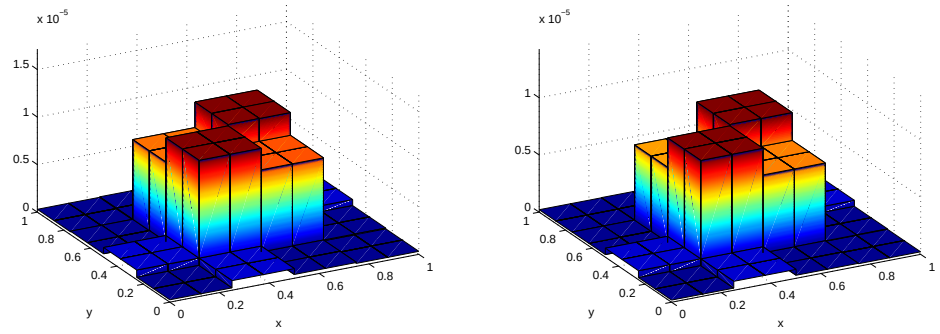
The reconstruction parameter groups are then and $\Sigma_{f,1}$ and $\Sigma_{r,2}$ are reconstructed and results shown in Figure IV-47.



(a) $\Sigma_{f,1}$ (b) $\Sigma_{r,2}$

Fig. IV-47. Ex. 19: Mixed parameter reconstructions for a 2 group problem

As in the previous example, an excellent reconstruction is obtained with resulting error in Figure IV-48.



(a) $\Sigma_{f,1}$ (b) $\Sigma_{r,2}$

Fig. IV-48. Ex. 19: Error for mixed parameter reconstructions for a 2 group problem

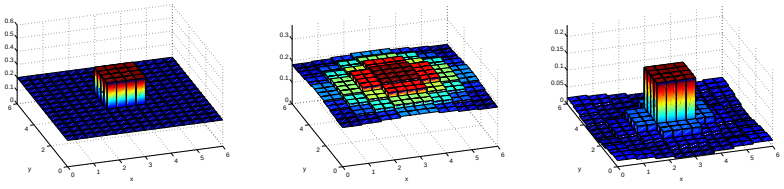
e. Example 20: Variation of incident neutron and measurement energy

A more ill-posed problem is at hand, as seen by the misfit surface plots in Figure IV-35, when incident neutrons are not present in all groups or measurements are not made in all groups. To demonstrate this, a two-group problem is considered where incoming neutrons are only present in group one and measurements are only made in group 2 and both groups removal cross section is attempted to be reconstructed. In a problem where measurements are made in a group that does not contain any incident neutrons, reconstruction ability is strongly dependent on the source of neutrons to this group by scatter or fission. For this example, the domain properties are shown in Table IV-VII.

Table IV-VII. Ex. 20: Domain parameters for a multigroup centered strong absorber

Region	D_1	D_2	$\Sigma_{a,1}$	$\Sigma_{a,2}$	$\Sigma_{f,1}$	$\Sigma_{f,2}$	$\Sigma_{s,1\to2}$	$\Sigma_{s,2\to2}$
Outer	1	1	0.2	0.2	0.05	0.1	0.14	0.0
Center	1	1	0.4	0.4	0.05	0.1	0.14	0.0

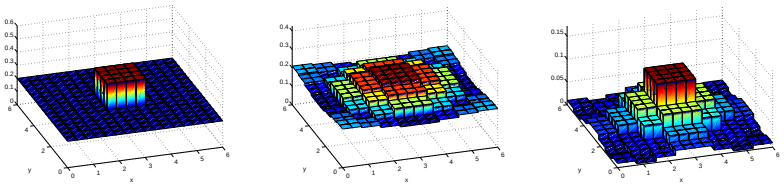
The results for the reconstruction in group one are shown in Figure IV-49.



(a) True $\Sigma_{r,1}$ (b) Recon. $\Sigma_{r,1}$ (c) Error

Fig. IV-49. Ex. 20: Reconstruction of $\Sigma_{r,1}$ with incident neutrons only in group one and measuring only in group 2

The results for the reconstruction in group two are shown in Figure IV-50.



(a) True $\Sigma_{r,1}$ (b) Recon. $\Sigma_{r,1}$ (c) Error

Fig. IV-50. Ex. 20: Reconstruction of $\Sigma_{r,2}$ with incident neutrons only in group one and measuring only in group 2

These reconstructions are not of high quality as the strong absorbers appear to be smeared throughout the domain. This due to the lack of information produced by the presence of incoming neutrons and measuring in single different groups. Even though this domain may have high reconstruction ability with incident neutrons and measurements in both groups, reducing these to one each greatly affects the reconstructions.

CHAPTER V

CONCLUSIONS

The inference or reconstruction of material parameters inside a domain is presented as a PDE-constrained optimization problem. In such problems, non-invasive inverse problem techniques are employed using only boundary input and output as information. An initial guess is performed, then an attempt to iterate to the solution is made where Newton's method implemented.

When a reconstruction is attempted on a domain there are several factors that play a key role in the ability of reconstruction no matter the complexity of the domain. Such factors are the location of measurements, location of incoming beams, and which parameter is being reconstructed. Through a misfit analysis, the location of incident and measurement location has been shown to severely affect the ability of reconstruction. For example in the two parameter case, or side by side homogeneous regions, if you only measure or have incoming particles on one side of the domain, the magnitude of the properties on the other side will have very little effect on the flux at the measurement points. Problems such as this are very ill-posed where the material parameters can change significantly whereas the resulting fluxes at the measurement points have little or no change. In such problems, the optimality conditions must be driven as low as possible to achieve meaningful results. This knowledge can be extracted to more complicated cases and problems can be approached with greater knowledge to provide the best possibility of reconstruction for any given domain.

The material parameter that is being reconstructed plays a role in the ability to reconstruct it based on its magnitude with respect to the other parameters. For example in a moderately thick domain of very strong absorption cross section, if the

center region had a very weak fission cross section, such few of these neutrons born from fission in the center actually make it to the measurement points therefor having negligible change on the boundary fluxes even with significant change in the center region's fission cross section therefore having negligible change on the misfit.

In the one group case with position dependent parameter reconstruction, an attempt at reconstructing the neutron absorption cross section was performed with incoming neutron beams on all four sides of the two dimensional domain. This model proves reconstruction of material parameters is possible with PDE-constrained optimization and more complicated applications should be implemented.

The concept of multiple experiments has been introduced next, where the neutron source is moved creating multiple separate flux and adjoint solutions while the problem optimizes the same set of cross sections. When multiple experiments were applied to this one group problem, a large increase in reconstruction ability was shown. Domains of certain complexity or size that no internal features distinguishable using only one experiment could now be reconstructed successfully. While the addition of multiple experiments proves greater reconstruction ability, no matter how many experiments are used, there still are reconstruction limits with respect to domain size and resolution. In the benchmark case of the centered strong absorber, the maximum reconstruction depth was limited to approximately 6-7 diffusion lengths deep from the boundary. When a domain of larger size is considered, even with a large number of experiments, the internal features of the domain cannot be reconstructed. Reconstruction resolution has limits as well, such that if the size of a domain is held constant and a finer and finer mesh considered, no matter how many experiments used, reconstruction resolution is limited to a fraction of an optical thickness.

The addition of signal noise or bias on the synthetic measured solutions severely affects the reconstruction ability of a domain. In small sized domains, such as 4

$\times$ 4, the addition of signal noise up to about 1.0% still allows reconstruction of distinguishable features of a domain while for larger domains, small quantities of signal noise around 0.005% to 0.01% may still be enough to destroy reconstruction ability. When a signal bias is applied to the synthetic measured solution, or signal noise averaged higher or lower then zero, it is expected that if the measured solutions are stronger, the cross sections be lower causing less attenuation and vice versa. The results from cases where reconstruction ability is high, this shows to be true although areas where reconstruction ability is low error increases significantly with the addition of signal noise or bias.

In the approximation, the same reconstruction limits do not always apply to the inference of various parameters from different energy groups. In the case of removal cross sections only to be reconstruction, these parameters do not couple the groups together in any way and is similar to reconstruction of two one-group problems simultaneously. In this reconstruction parameter configuration, the same reconstruction limits from the one group absorption problems arise here as well. The same maximum reconstruction depth of 6-7 diffusion lengths deep from the boundary is seen here.

If fission cross sections are to be reconstructed, a more complicated problem arises. In simple homogeneous domains, the fission cross section can be reconstructed easily just as the removal cross section. In more complicated problems, the ratio of fission cross section to absorption cross section may play a role in the ability to reconstruct. As mentioned before in the case of a strong absorber around a centered region of weak fissile material, due to the magnitude of the fission cross section with respect to the strength of the absorber, the misfit space in this problem is significantly flatter with respect to change in the cross section. The fission cross section in the center can change significantly without noticeable change in the boundary fluxes and

misfit. This problem is more ill-posed than the case of absorption or removal cross-section reconstruction as shown by the magnitude of the misfit surface plots. Through similar test problems, the maximum reconstruction depth of approximately 5 diffusion lengths is present for multigroup fission cross-section reconstructions. This is smaller then the removal or absorption reconstruction limits by a significant quantity of 1-2 characteristic lengths.

Subsequent work may include a transport model for the governing physics constraint. The same optimization methods would apply, except a transport operator would replace the diffusion operator. This will provide more accurate flux solutions especially in cases where the diffusion approximation is weak. Additional subsequent work may include the implementation of more constraints, for example using barrier methods, for complex problems where the optimization functional is not smooth.

REFERENCES

[1] J. Baker McNeill, "100 percent cargo container scanning: A global disaster," Research report, Kathryn and Shelby Cullom Davis Institute for International Studies, 2008.

[2] K. Ren, G. S. Abdoulaev, and A. H. Hielscher, "Optical tomography as a PDE-constrained optimization problem," *Inverse Problems*, vol. 21, pp. 1507–1530, 2005.

[3] Wolfgang Bangerth and Amit Joshi, "Adaptive finite element methods for the solution of inverse problems in optical tomography," *Inverse Problems*, vol. 24, pp. 034011/1–22, 2008.

[4] A. D. Klose and A. H. Hielscher, "The inverse source problem based on the radiative transfer equation in optical molecular imaging," *Journal of Computational Physics*, vol. 202, pp. 323–345, 2005.

[5] A. Gibson, M. Schweiger, and S. R. Arridge, "Computational aspects of diffuse optical tomography," Research report, Computing in Science & Engineering, Los Alamitos, CA, 2003.

[6] W. Hardin, "Cargo inspection: Imaging solutions wait for government's call," Research report, Machine Vision Online.

[7] V. Scipolo, "Scattered neutron tomography based on a neutron transport problem," M.S. thesis, Texas A&M University, College Station, Texas, May 2004.

[8] Jorge Nocedal and Stephen J. Wright, *Numcerical Optimization, 2nd edition*, Springer Science+Business Media, Madison, WI, 2006.

[9] J. A. Favorite and R. Sanchez, "An inverse method for radiation transport," *Radiation Protection Dosimetry*, vol. 116, pp. 482–4585, 2005.

[10] Computational Science Education Project, *Mathematical Optimization*, United States DOE, e-book, http://www.phy.ornl.gov/csep/, 1996.

[11] Murat Aydin and M. Akif Atalay, "Inverse neutron diffusion problems in reactor design," *Journal of Nuclear Science and Technology*, vol. 44, pp. 1142–1148, 2007.

[12] Weston M. Stacey, *Nuclear Reactor Physics*, Wiley-VCH GmbH & Co. KGaA, Weinheim, 2004.

[13] Richard L. Burden and J. Douglas Faires, *Numerical Analysis, 8th edition*, Thomson Brooks/Cole, Belmont, CA, 2005.

[14] Nuclear Data Evaluation Lab., "Table of nuclides," Tech. Rep., Korea Atomic Energy Research Institute, 2000.

[15] National Nuclear Data Center, "Evaluated nuclear data file (endf)," Tech. Rep., Brookhaven National Laboratory, 2009.

VITA

Matthew Ryan Sternat was born in 1982 to Louis J. Sternat Jr. and Patricia Sternat in Friendswood, TX. He graduated from Friendswood High School in 2001, received his Bachelor of Science in nuclear engineering from Texas A&M University in 2007 and his Master of Science in nuclear engineering from Texas A&M University in 2009. He is a member of Alpha Nu Sigma Honor Society and the National Scholars Honor Society. His primary interests are numerical reactor physics and application towards nuclear safeguards.

Department of Nuclear Engineering
Texas A&M University
3133 TAMU
College Station, TX 77843-3133

m.sternat@tamu.edu
matt.sternat@gmail.com