Nuclear Security Science
& Policy Institute

A Division of the
Texas Engineering Experiment Station

SCATTERED NEUTRON TOMOGRAPHY BASED ON A NEUTRON TRANSPORT INVERSE PROBLEM:

FINAL REPORT

Principle Investigator:

William S. Charlton, Ph.D.
Associate Professor, Nuclear Engineering Department
Director, Nuclear Security Science and Policy Institute
Texas A&M University
College station, TX 77843

This material is based upon work supported by the U S. Department of Energy under Award No. DE-FG07-04ID14612.

July 1, 2007

EXECUTIVE SUMMARY

Neutron radiography and computed tomography are commonly used techniques to non-destructively examine materials. Tomography refers to the cross-sectional imaging of an object from either transmission or reflection data collected by illuminating the object from many different directions. Typical neutron radiography and tomography analyze transmitted images in which a large fraction of the incident radiation is transmitted through the object without collisions. Many objects have small absorption rates but large scattering rates for neutrons. The scattering properties of the material increase the probability that multiple scattered neutrons could re-enter the beam and significantly decrease image resolution. Analysis of these structures using traditional techniques is a formidable task. Classical tomography fails to reconstruct the optical properties of thick scattering objects because it does not adequately account for the scattering component of the neutron beam intensity exiting the sample. In this work, a new method of computed tomography was developed which employs an inverse problem analysis of both the transmitted and scattered images generated from a beam passing through an optically thick object. All objectives of the project were successfully met.

This inverse problem makes use of a computationally efficient, two-dimensional forward problem based on neutron transport theory that effectively calculates the detector readings around the edges of an object. The forward problem solution uses a Step-Characteristic (SC) code with known uncollided source per cell, zero boundary flux condition and S_n discretization for the angular dependence. The calculation of the uncollided sources is performed by using an accurate discretization scheme given properties and position of the incoming beam and beam collimator. The detector predictions are obtained considering both the collided and uncollided components of the incoming radiation.

The inverse problem is referred as an optimization problem. The function to be minimized, called an objective function, is calculated as the normalized-squared error between predicted and measured data. The predicted data are calculated by assuming a uniform distribution for the optical properties of the object. The objective function depends directly on the optical properties of the object; therefore, by minimizing it, the correct property distribution can be found. The minimization of this multidimensional function is performed with the Polack Ribiere conjugate-gradient technique that makes use of the gradient of the function with respect to the cross sections of the internal cells of the domain. The code used to reconstruct the optical properties of the object from image data is called the Tomography Neutron Transport using Scattering code or TNTs.

The forward and inverse models have been successfully tested against numerical results obtained with MCNP (Monte Carlo Neutral Particles) and experimental images taken using the Texas A&M University (TAMU) Nuclear Science Center (NSC) imaging facility. These test showed excellent agreement between reconstructed and actual object properties. The reconstructions of several objects were successful. In fact, it was shown that TNTs can reconstruct simple objects using illumination of the object from only one side and measuring images from three sides of the object. The use of multiple beams around the object will significantly improve the capability of TNTs since it increases the number of constraints for the minimization problem.

The method developed here will allow for a more accurate determination of the location and size of any imperfections in objects and allow for more accurate reconstructions of the optical properties of objects. This research is the first of its kind using neutron transport theory as the

basis for image reconstruction and makes significant advances in methods used. These methods are computationally intensive but allow for much greater accuracy in image properties.

TABLE OF CONTENTS

1. INTRODUCTION

1.1. Overview

Tomography refers to the cross-sectional imaging of an object from either transmission or reflection data collected by illuminating the object from many different directions. The incident radiation must penetrate the object before being detected. Therefore the choice of radiation type is crucial. Neutron radiation is typically more penetrating than x-ray, γ-ray or charged particle radiation. Thus, neutron radiography can often be used to evaluate thick objects. Neutron radiation also typically serves as a complement to x-ray radiation.

Neutrons, compared to x-rays, have high interaction probability with hydrogen and a lower attenuation in several heavy elements which are "black" for x-rays (e.g. lead and bismuth).The attenuation of x-rays increases proportional to the atomic number (Z) of the material, whereas there is no direct relationship between neutron attenuation and Z. Thus neutron radiography is often complimentary to x-ray radiography[1].

The investigation of moisture and corrosion, the detection of explosives and adhesive connections, and the inspection of defects in nuclear fuel or in thick metallic samples[2] are examples where neutron radiography can be utilized favorably. Neutron radiography and tomography applications are, in fact, present in many fields: nuclear engineering, material characterization and the biomedical fields are the principal ones.

Other important examples of the versatility of the neutron radiography method are:

- surveys of nuclear fuel[2,3,4,5],
- imaging of two-phase flow[2,6],
- analysis of carbon-fiber composite airplane wings[2],
- imaging for explosive devices[7], and
- fast imaging of combustion[2].

However, neutrons, or near-infrared light are strongly scattered in many materials. Thus, standard back projection techniques, such as those applied in x-ray tomography, have been of limited success[8].

The objective of the work described here was to develop an algorithm and a computational code to solve the inverse tomography problem accounting for neutron scattering. The hypothesis is that it is possible to more accurately reconstruct the optical properties of an object using the information contained in the scattered radiation as well in as the transmitted radiation than using transmitted radiation alone. Sometimes very highly scattering problems are so difficult to solve that only the analysis of the scattered component leads to a solution[9].

1.2. The Early History of Neutron Radiography

The first radiographs were obtained in 1895, coinciding with the discovery of radiation. In 1985 Röntgen was experimenting with high voltage discharges in a vacuum tube. When photographic plates that had been stored nearby were developed, they were found to be blackened without any obvious cause. Reasoning that some unknown radiation from the high voltage discharge could be affecting the photographic emulsion, Röntgen carried out ad hoc experiments and quickly established that this was in fact the case. Realizing the importance of this effect, he rapidly

developed his discharge tube to produce more radiation and obtained a “negative” of welded plates of zinc[1].

The practical implications of this ability to produce shadow images of items which were opaque to light and thus reveal their internal structure were clear. Further development was rapid and widespread. Experimental x-ray radiographs were soon produced in laboratories in Europe and the USA[1]. The ferment of scientific and engineering endeavor that followed the demonstration of these completely new phenomena carried over into the new century with research into electronics and atomic physics.

The discovery of the neutron is credited to Chadwick. In 1932, he related and hypothesized on the work of Bothe, Becker, Curie and others and assumed that the penetrating radiation produced by bombarding beryllium with alpha particles was neither positively nor negatively charged. He called this radiation the neutron (from the Latin *neuter* meaning neither). He had identified a particle that, together with the proton, was one of the basic building blocks of matter.

The application of the neutron to produce radiographs quickly followed its discovery. Kallman and Kuhn in Germany in 1935 used neutrons produced by an accelerator to make radiographs but the quality was poor due to the weak and ill-defined beam. This coupled with the complexity of the apparatus and the fact that hours of exposure were required did not lead to early exploitation as had happened with x-rays. Nevertheless, the methods and potential of the technique were clearly demonstrated. The publication of their work was delayed by World War II. It was not until 1948 that they revealed the thoroughness of their investigation by describing most of the basic techniques in use today.

The development of nuclear reactors during and after the war increased the intensity of neutron fluxes available for experimental purposes by many orders of magnitude, but it was not until 1956 that further work on neutron radiography was reported. The first use of a beam of neutrons from a reactor to produce a radiograph was by Thewlis and Derbyshire. They used a beam from the 6MW graphite reactor BEPO at Harwell to produce good quality images having specific non-destructive testing applications such as voids in uranium and in “Boral”, a neutron shielding material fabricated from boron carbide and aluminum. Their radiograph showing the structure of a plant pointed to the usefulness of the technique in studying organic specimens. This is due to the high attenuation of neutrons by hydrogenous materials.

The technique developed slowly for several years until problems associated with the radiography of radioactive materials encouraged its more active revival. Several researchers reported their work in the early 1960’s. But it was principally the work of Berger[2] of Argonne Laboratories in USA, followed by Barton[8] at Birmingham University that led to its revival.

In 1968, there were 46 reactor facilities, three accelerators and five isotopic sources in use or being built. The history of neutron radiography stretches back 65 years; although, it is only in the last 30 that it has come to the forefront as an accepted method of non-destructive testing. In the last 10 years, it has reached the stature of World Conferences[1].

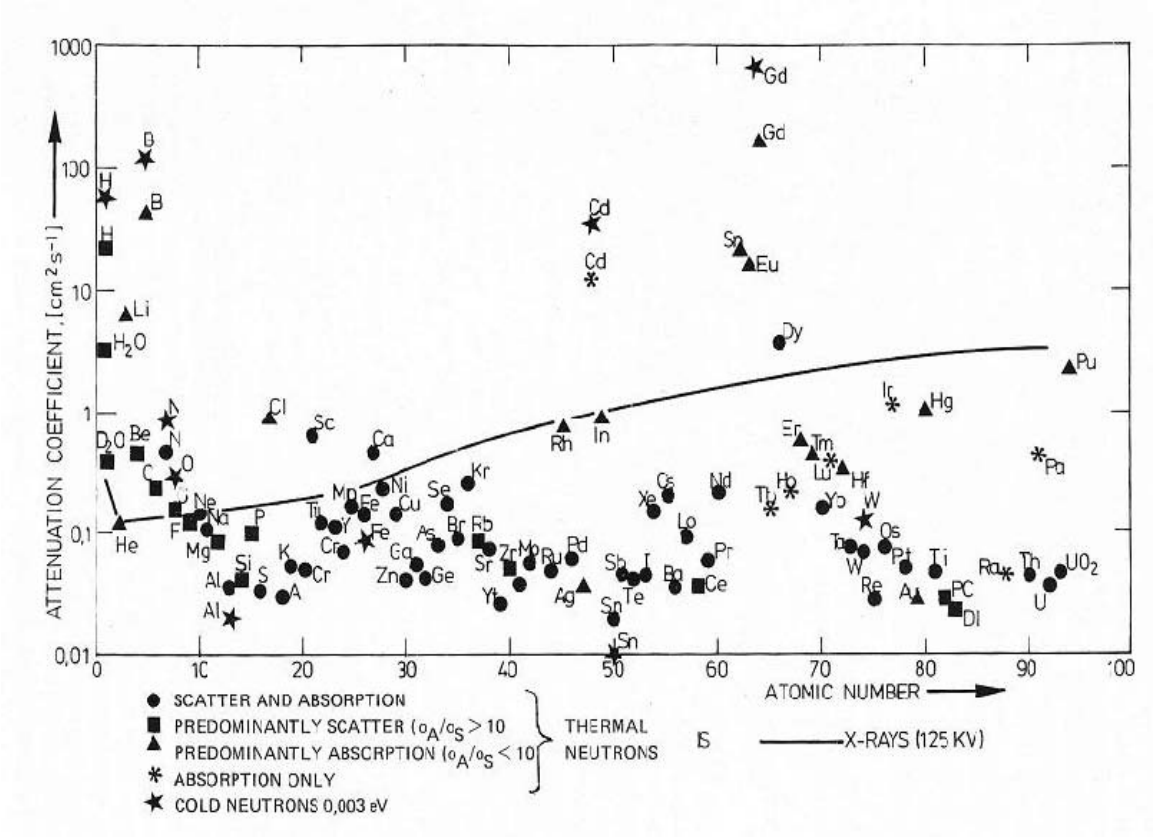


Fig. 1. Attenuation coefficient versus atomic number for all elements and for both thermal neutrons and x-rays.

1.3. Fundamentals of Neutron Radiography

Unlike the x-rays, which interact with the electron cloud, the neutron interaction is not characterized by a direct dependence on the atomic number of the object. There are practically no generalizations that can be made which relate neutron interaction characteristics to atomic mass or atomic number. Each interaction of a neutron with an isotope of a particular element is unique. Figure 1 depicts the relationship between the attenuation coefficient for neutrons and x-rays versus element atomic number. As can be seen, the x-ray line is a smooth function with atomic number but the neutron attenuation coefficient has no discernible pattern.

To produce a neutron radiograph a continuous supply of unbound neutrons is required and these must be directed onto the object to be radiographed. The object will scatter or absorb some of the radiation from the beam. The beam reaching the detector will have an intensity pattern representative of the structure of the object.

Neutron radiography involves three principal components:

- a suitable neutron beam
- an object of radiographic interest
- a device to record, either immediately or delayed by some time, the radiation intensity of the transmitted beam.

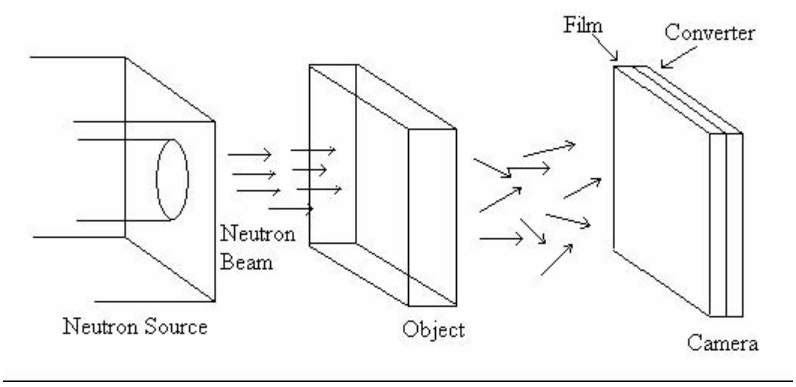


Fig. 2. Principal components of a schematic neutron radiography system.

Figure 2 provides a graphical depiction of these three system components for the case of imaging a non-radioactive object. After passing through the object, the beam that remains enters a detector that registers the fraction of initial radiation (I_o) intensity that has been transmitted by each point in the object. Any in-homogeneity in the object will show up as a change in radiation intensity reaching the detector.

The intensity of radiation (I) passing through an object of thickness t is given by:

$$I = I_0 e^{-\Sigma_t t} \tag{1}$$

where Σ_t is the total macroscopic cross section of the object. If a void defect, of width d, is present; then the intensity will be

$$I = I_0 e^{-\Sigma_t (t-d)} \tag{2}$$

If an inclusion is present as a defect, then the intensity will be

$$I = I_0 e^{-\Sigma_t (t-d) - \Sigma_a d} \tag{3}$$

where Σ_a is the macroscopic cross section of the inclusion.

This methodology [sometimes referred to as the simple exponential attenuation method (SEAM)] has been used successfully for numerous applications; however, when applied to highly scattering media (such as any low atomic number material) the scattering component of the neutron beam intensity exiting the sample is not adequately accounted for by SEAM. This tends to lead to decreased system resolution when these scattered neutrons are recorded at the image plane.

For example, Fig. 3 shows a neutron radiograph of two cadmium strips. The holes in the cadmium strips range from 50 μm in diameter to 1100 μm in diameter. For this object, which has a very high neutron absorption cross section and a very low neutron scattering cross section, all of the holes (and the gaps between the holes) are clearly visible.

Figure 4 shows a radiograph for a thick carbon composite structure. A 3175-μm diameter hole is present in the thick object. This hole is not visible in the radiograph even though it is much larger than the holes in the cadmium. The carbon fiber composite material has very low neutron absorption properties but reasonably high neutron scattering properties.

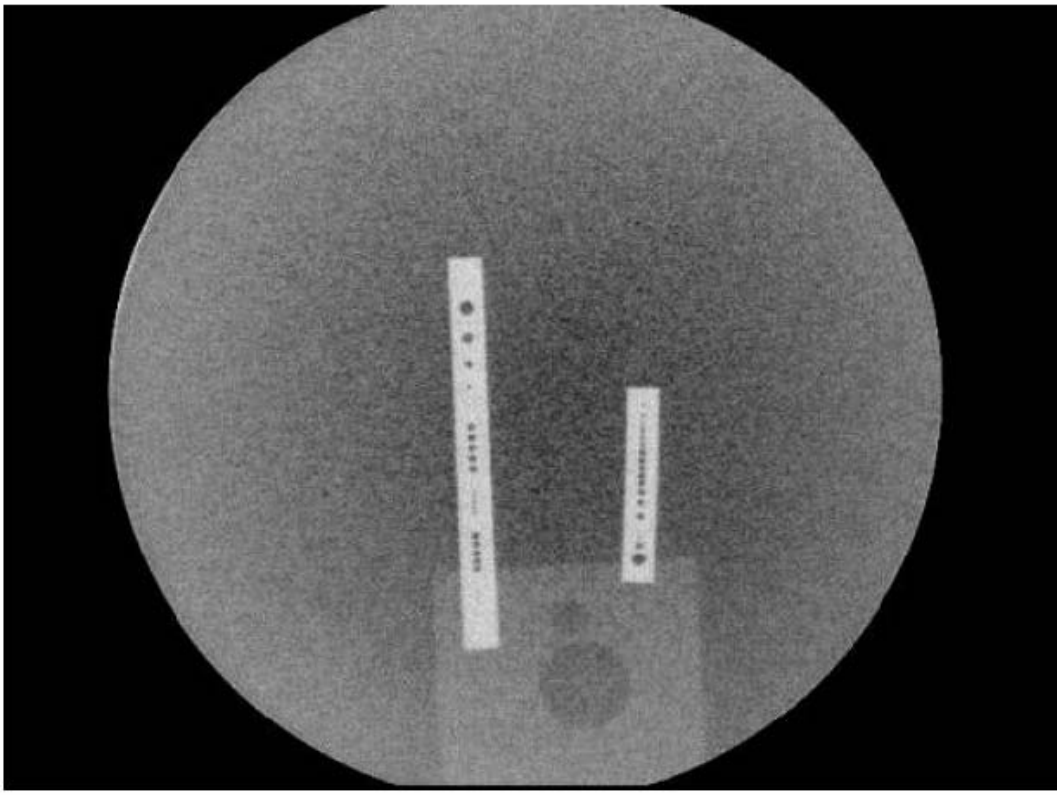
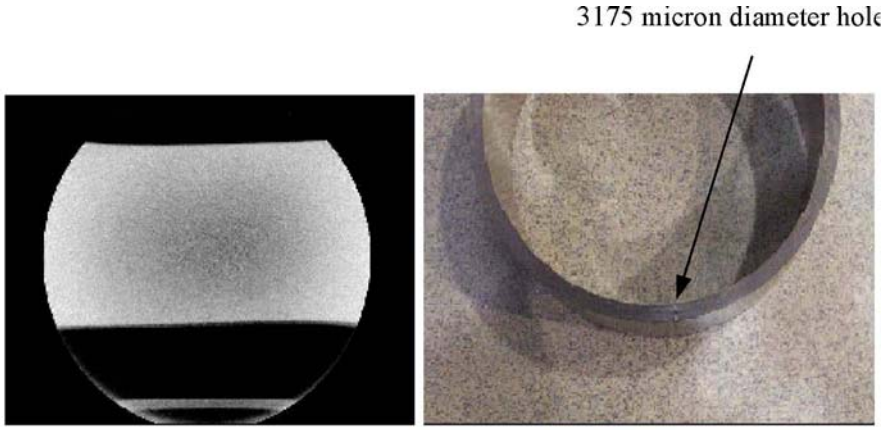


Fig. 3. Neutron radiograph of two cadmium strips displaying good image resolution for highly absorbing, low scattering materials. Each strip contains a number of different sized holes. The strips are supported on an aluminum plate.



(a) (b)

Fig. 4. Image on the left (a) is a neutron radiograph of a thick carbon fiber composite object with a 1/8th inch hole present. Image on the right (b) shows the object and the hole.

1.4. Neutron Transport Theory and Inverse Problem

A potential solution to the problem presented in the section above is proposed here. If the neutron-imaging camera was placed at some angle θ_o off the beam path as in Fig. 5, an image of the scattered neutron component from the surface of the sample in that direction could be acquired. It is proposed that if the scattered neutron component at a variety of angles around a sample can be measured, these scattered components could be used to aid in a better estimation of the source of neutron reactions in the sample. It is expected that this information would lead to a more accurate reconstruction of the surveyed object.

It is important to note that these scattered neutron images are fundamentally different from the transmission images. The source of neutrons in these images is (to a first-order approximation) the inverse of the source of images in the transmission experiment. That is in the transmission images the source of neutrons is the initial input beam and the image is created by removal of neutrons from that beam. In the scattered neutron images, the source of neutrons is the neutrons that are removed from the initial beam by scattering and the image created shows that source.

One of the more common tomographic techniques is the filtered back projection method[10,11]. In this technique, the projection data can be considered as line integrals along the neutron beam lines. For highly scattering objects these line integrals become complicated because of the source of scattered neutrons along the integrals which is difficult to characterize with transmission images.

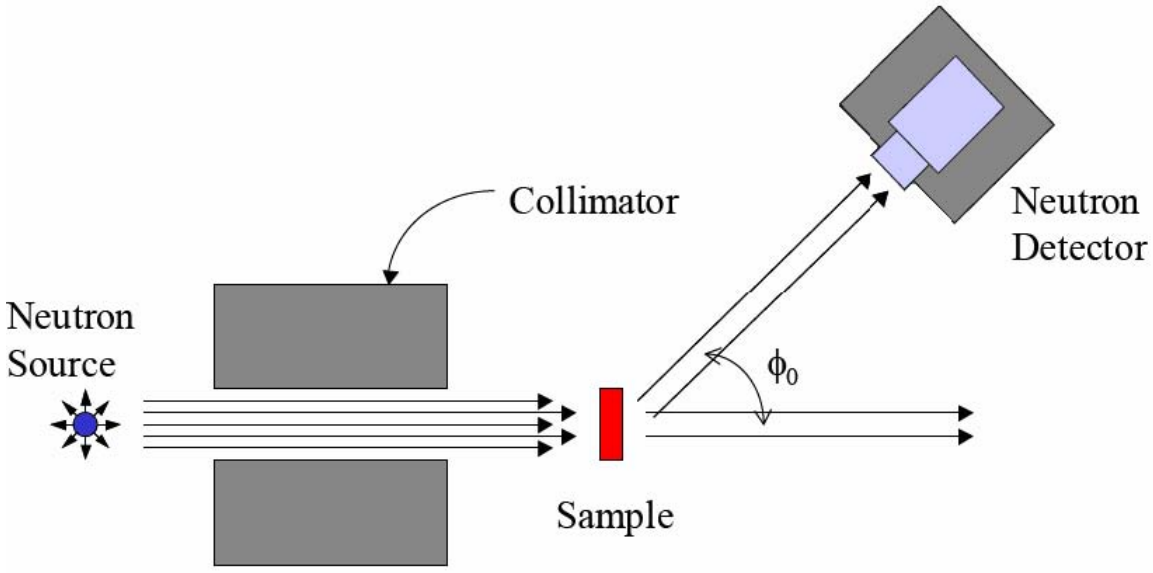


Fig. 5. Idealized setup for capturing scattered neutron images.

Suppose an object was imaged in which neutrons interacting in the sample could either scatter once in the sample, get absorbed in the sample, or travel uncollided through the

sample. Consider a sample such that the scattering rate in the object was higher than the absorption rate. In this instance, the transmission image would have a high degree of unsharpness due to the fact that the principle means of removal of neutrons from the beam is through scattering and depending on the angle of scatter, a high portion of these neutrons continue on through the object and interact in the detector in some location other than where they would interact if they continued uncollided. This leads to "fuzziness" of the resultant image. The scattered images measured, via a setup such as that shown in Fig.5, would record only the source of neutrons scattering along the line connecting the detector through the sample at an angle of θ_o. This allows for the determination of the scattering rates throughout the sample. These scattering rates could then be used to determine the contribution of scattered neutrons to the detector. With this contribution known, the image resulting only from the neutrons removed from the beam due to absorption can be determined. This would decrease the unsharpness in the resultant image.

In a realistic scenario, the transmitted image would consist of neutrons from the initial beam that passed through the object uncollided, had a single small angle scatter, or had multiple scatters. The scattered neutron image would consist of neutrons born from the scattering reactions in the sample from the initial beam and then passed uncollided through the sample or had additional scattering reactions in the sample. To properly simulate this effect the neutron radiation transport must be simulated through the sample accounting for all of the reaction rates in the sample.

Thus, a forward model capable of calculating the detector responses (both transmitted image and scattered image) for the beam passing through a known object is needed. In addition to this forward model, an inverse model is needed in order to pass information from the measured images back through the forward model to influence the "guess" of the initial object structure (essentially the cross section sets defining the object). The forward model can then be repeated using the more accurate guess. These iterations would continue until the calculated image matched the measured image to within some tolerance (i.e., an objective function is minimized). This is the fundamental concept behind the model-based iterative imaging reconstruction (MOBIIR) schemes[12]. MOBIIR schemes mainly differ in their choice of forward model and how the spatial distributions of the optical properties of the medium are updated.

Since a forward model is required, transport theory has to be introduced. The transport equation describes the movement of particles through a medium. It is an integro-differential equation with seven independent variables: space (3), direction (2), energy (1) and time (1). Therefore, only simple problems can be solved analytically and numerical methods must be applied to most problems of interest. Due to the large number of unknowns (e.g., a problem discretized with 10^N unknowns in each independent variable will have 10^{7N} unknowns) it is difficult to use direct numerical methods. Instead iterative schemes are employed to solve the transport equation.

Successive approximations of this complicated equation lead to the diffusion equation that has the great advantage of illustrating many of the important features of the particles behavior in the matter without the complexity of the transport equation. Diffusion theory provides a strictly valid mathematical description of the particle flux only if the assumption made in its derivation is satisfied. Specifically, this implies that Fick's law is valid. Fick's law is valid under the following conditions:

- Absorption is much less likely than scattering. This is satisfied for most of the moderating materials that are usually found in a nuclear reactor.
- There is a linear spatial variation of the particle distribution. This condition is satisfied a few mean free paths away from the boundary of large homogeneous media with relatively uniform source distributions.
- All scattering is isotropic in the center of mass reference frame. Typically this is a good assumption for most of the heavy atomic mass nuclei.

Diffusion theory has been widely used in tomography methods especially recently. A variety of optical methods based on MOBIIR schemes have been studied in the past[12-19]. While these studies have principally been in the area of low-energy x-ray medical imaging, they have led to a variety of creative methods and their general application can be extended to neutron imaging.

The forward model used in all of these methods is based on diffusion theory. Many authors solved the diffusion forward model by adopting a finite element method. Interesting methods used include up-wind or Gibbs schemes to reach a solution for the diffusion equation[6]. The efficiency of these methods is due to the easy analytical expression of the diffusion equation that leads to an easy analytical expression of the updating algorithm. For example perturbation theory method, widely used to update the guessed properties, is easily applicable to a problem in which the forward calculation is obtained with diffusion theory, due to the relatively simple analytical expression for the flux.

While accurate for some instances, for most problems of interest to neutron radiography, diffusion theory is insufficient for accurately describing the transport of neutrons through these objects[1]. Thus, this project will propose the use of a forward model based on neutron transport theory.

1.5. Report Overview

Section 2 introduces the development and implementation of the forward model based on neutron transport theory for predicting detector images from a thick highly scattering object. This methodology accurately and efficiently simulates the transport of neutrons through the object including at boundaries between highly scattering and highly absorbing regions. The accuracy of this method is crucial since uncertainties in the forward model propagate through the inverse model; however, computational speed is also crucial since in an iterative scheme, the forward model calculation may be used numerous times. Chapter II also contains the verification via MCNP of the validity of the code.

Section 3 contains the development and implementation of the inverse model. This model allows for the analysis of radiographic images to reconstruct the most likely details of the object that minimizes an objective function based on the measured images. It makes use of the forward model for predictions of the detector response to a specific object definition. The chapter will also involve the implementation of this methodology including integration of the forward and inverse models. It includes mathematical and numerical verification of the gradient calculated by an adjoint formulation.

Section 4 presents the result obtained from the reconstruction of different objects. These objects will range from highly scattering to moderately scattering and include defects of various sizes. The images used were obtained from both experimental data and from numerical simulations

using MCNP The reconstruction technique is therefore verified and some considerations are presented.

Section 5 presents conclusions and contains final considerations and suggestions for future work.

2. FORWARD MODEL

2.1. Background

The transport equation is a linear form of the Boltzmann equation, developed one century ago for the study of the kinetic theory of gases[20]. It describes the evolution of a particle distribution function in an infinitesimally small 7 phase-space (time, space, energy, direction) volume. The analytical solution of this equation is confined to very highly idealized cases often concerning astronomic problems with semi-infinite mediums[21]. The solution of the equation for more common but complicated problems is obtained through the use of numerical approximations and computational calculations. The application of transport theory is associated with a wide variety of fields and research topic such as:

- Nuclear Engineering
  - Reactor analysis[22]
  - Shielding and dose calculation[23]
- Rarefied gas dynamics
  - Sound propagation
  - Diffusion of molecules in gases
- Other
  - Traffic flow

The numerical solutions to the transport equation are divided into stochastic (Monte Carlo) and deterministic. The Monte Carlo method treats all the events that can occur to a particle in terms of probability functions. It tracks every particle from its "birth" until its "termination" (for many reasons such as absorption, leaking …) and makes the history of the particle[23]. By using a large number of histories it estimates the average particle behavior. This method is in general computationally more expensive than deterministic methods. The advantage is the possibility of simulating complex geometrical systems and physically complex histories.

Deterministic methods solve the transport equation by discretization of the phase-space volume in order to reduce the transport equation to a set of simpler algebraic equation. The discretization into energy groups lead to a multi-group transport equation. The transport equation can be expressed as an integro-differential equation or as an integral equation. The choice of spatial and angular discretization depends on the form of the equation. The form used for this project is the steady-state, one-group integrodifferential form that involves an angular integral and a first-order spatial derivative. Other forms are described elsewhere[21,24]. Different angular discretizations can be applied to simplify the angular integral into a set of differential equations. We choose to treat the angular dependence with a discrete ordinate (S_n) method in this work.

The S_n methods approximates the angular integral as a quadrature summation as will be presented in more detail later. The resulting set of equations is a system of partial differential equations in space that are spatially discretized to generate a set of algebraic equation.

The method chosen for the spatial discretization is the Step Characteristic (SC) method. It has been developed first by Lathrop[25]. Like for every other characteristic method the SC method transforms the S_n equation into a one-dimensional equation by rotating the axis of the coordinate system along the direction of motion (the characteristic line). Given the value

of the angular flux at a point along the characteristic line and known source term the characteristic equation can be analytically solved for the angular flux everywhere along the line. This will be presented in more detail later in the chapter.

Our choice for the forward model was the two dimensional SC method. In the decision many factors, other than accuracy and computational effort, were considered. The accuracy of this method for a typical neutron radiography problem is acceptable especially for an object with a scattering ratio less than unity. A small scattering ratio is also needed to have a relatively fast convergence. In addition, the SC is analytically simple. The resolving equations are normal exponential relations with the possibility to use them in the analytical process to update the cross sections. Accuracy and easy analytical expressions were the keys in the choice of the forward model.

In this work, only two-dimensional cases were considered, for simplicity. Thus, the model has been developed for a two dimensional case. This will allow for proper testing of the algorithm for the update of the cross sections and will limit the computational time required to acquire results. It is expected that the step from two to three dimensions will be straightforward once the inverse model has been tested and optimized.

2.2. Forward Model Description

A forward model was mathematically derived and computationally implemented to solve a set of problems with common characteristics. The purpose of this model was to predict detector readings around a rectangular object placed into a neutron beam. The beam is not necessarily mono-directional and may consider the physical divergence that a real beam experiences when collimated. Also, the detectors can be collimated such that they only record radiation coming from a particular direction. The beam-object-detector situation is schematically represented in Fig. 6.

In order to solve the problem described above, the computational and mathematical process was divided into three different parts. First, the beam and the spatial discretization of the object were defined. The object was divided into a set of spatial rectangular regions. Then the initial source of un-collided particles for every cell of the object was calculated.

Due to the operation of the first step, the problem has become a fixed-source, zero-boundary condition problem. In the second step, this is solved using SC with Sn angular discretization. The solution will be represented by the average angular flux for each cell and the x- and y-direction net-currents for every cell of the problem.

With the solution reached in the second step, it is possible to calculate the scattering source of every cell and calculate its contribute to the detector readings. By adding this component to the uncollided radiation entering every detector, the reading is complete.

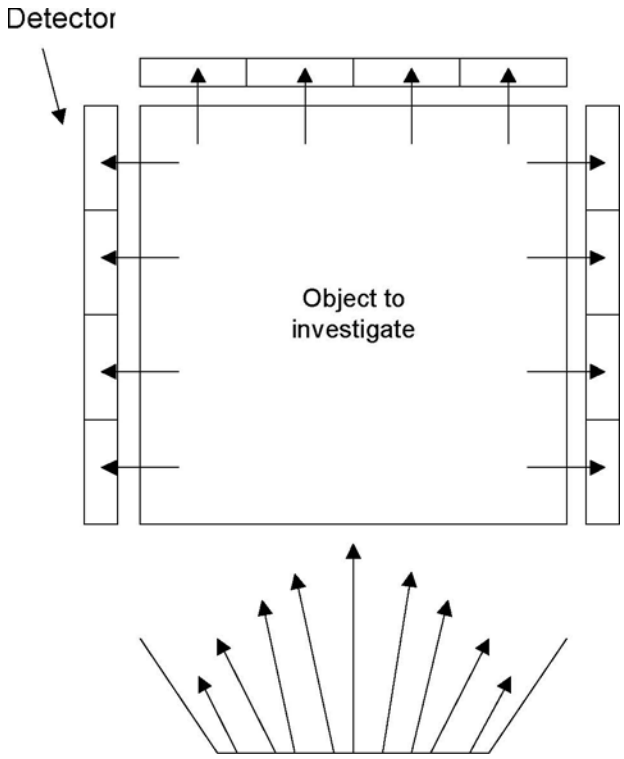


Fig. 6. Schematic of the problem.

2.2.1 Spatial Discretization

We begin with a graphical description of our approach to spatial discretization of an arbitrary rectangular Cartesian cell. The user inserts the dimension of the object in the x-y plane. The object is considered to be infinite and homogeneous in the third dimension. Then the number of regions in the x-and y-direction are inserted. The object is considered to be homogeneous over each region. For every region, the user specifies the number of cells in the x-direction and the number of cell in the y-direction. Cells that are into the same region are assumed to have the same initial optical properties. Figure 7 shows an object divided using the process described above. The object is divided into nine regions, three in the x-direction and three in the y-direction. Every region is then divided into cells. In particular, Region 2 in the x-axis is divided into five cells. Region 2 in the y-axis is divided into two cells. Thus Region (2, 2) has ten total cells in it. The cells are constructed by intersecting the lines of every sub-division. In particular, this problem consists of a total of ninety cells. Cells in the same regions have the same initial optical properties. The user introduces the optical properties per region. Every cell is denoted as unique with two indexes, i and j. In the example i goes from one to ten, j goes from one to nine and the cell is indicated as: cell (i,j).

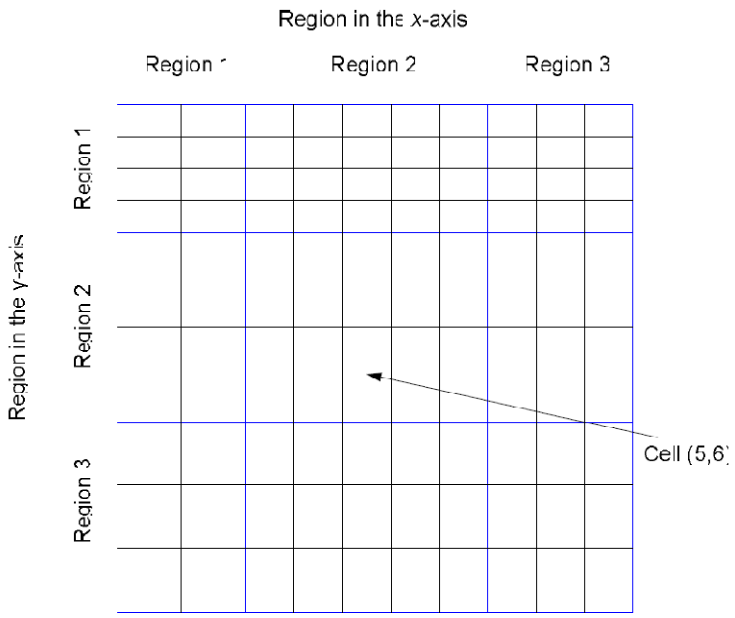


Fig. 7. Spatial discretization.

2.2.2 Uncollided Sources

The first step of the process to predict the detector readings is to calculate the fixed source of un-collided particles for every cell of the discretized object. First, the user inserts the properties of the beam window such as distances from the edges of the object and dimensions of the collimator.

In general, it is easy to calculate the component of radiation coming from a point source to a certain point into a medium:

$$\phi_{uncollided} = \left(\frac{source\ strength}{steradian}\right)\left(\begin{matrix}solid\ angle\ subtended\ by\ 1\ cm^2\\ area\ at\ the\ dose\ point\ perpendicular\\ at\ distance\ R\end{matrix}\right)\left(\begin{matrix}transmission\\ probability\ to\\ traverse\ R\\ without\ collision\end{matrix}\right)$$

$$\phi_{uncollided} = S(\Omega)\frac{1}{R^2}e^{-\sigma_t R} \quad (4)$$

where R is the total distance between the source point and the investigated one, S is the source strength, Ω defines a particular direction and σ_t is the total cross section of the medium. The above equation considers that the particle travels in a medium of constant total cross section. The formula can be easily generalized as follows:

$$\phi_{uncollided} = S(\Omega)\frac{1}{(r-r\prime)^2}e^{-\alpha(r-r\prime)} \quad (5)$$

where α(r-r') indicates the distance in mean free paths (MFP) between the point r and the point r'.

In case, the source is a planar source the un-collided radiation incident in a point is simply the integral of Eq. (5) along the plane, performed by using the angular notation introduced in Fig. 8.

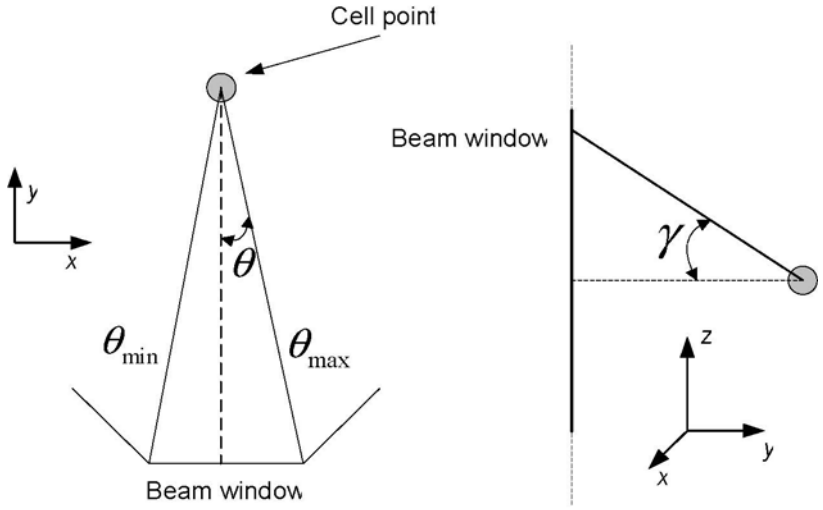


Fig. 8. Angular notation for the un-collided source calculation.

In Fig. 8, θ is the azimuthal angle and γ is the polar angle. Therefore, for a planar source finite in x and y and infinite in the z direction, the integral for the un-collided radiation in the point is:

$$\phi_{uncollided} = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} d\gamma \int_{\theta_{min}}^{\theta_{max}} d\theta \; S(\theta,\gamma) \frac{1}{(r-r')} e^{-\alpha(r-r')} \tag{6}$$

The angle θ, that together with r describes a point in the x and y plane, has a maximum and minimum value as boundaries of the integral, as shown in Fig. 6.

The code in order to perform the integral of Eq. (6) uses a quadrature set for the two angular dependencies. The interval between θ_{min} and θ_{max} is divided, following the trapezoidal rule, in as many angles as the user decides. The same is valid for the interval in the polar direction. The integral is therefore approached as a quadrature summation.

In order to calculate α, the distance in MFP between the source point and the investigated point, the code tracks back the particle along the line of conjunction recording the cell and the path that the particle goes trough. Every cell has its own total cross section and knowing the paths traveled in the cell the code can calculate the number of MFP traveled by the particle per cell that is simply the length of the path in cm multiplied by the total cross section.

In order to make a precise calculation of the un-collided flux per cell, the quadrature integral is performed at different points into the same cell and the result is then obtained averaging the results at the points into the cell. The number of locations per cell is user defined and the coordinates and weights of these locations are calculated by a Gauss –Legendre discretization technique. The choice of this tecnique is reliable since a Gauss-Legendre discretization of order N is the only quadrature set able to perform exact integrals of polynomial of degree up to 2N-1. Fig. 9 shows the process for a small number of cells. The number of locations into the cell to calculate the un-collided flux is user defined; in Fig. 9 it is four.

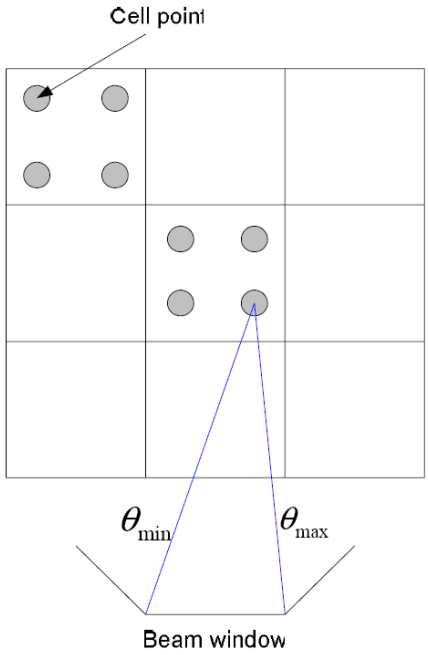


Fig. 9. Process to calculate the uncollided flux per cell.

By choosing an appropriate number of azimuthal and polar directions and a reasonable number of locations per cell, the code is able to perform an efficient and accurate calculation of the un-collided flux per cell. The un-collided flux per cell is then used as a fixed source in the source iteration technique to solve the transport equation in the cell. The accuracy of the forward method is greatly enhanced using the un-collided sources as driven force in the source iteration scheme. The accuracy of this technique lies in the treatment of the boundary conditions that are not discretized by following the same angular discretization of the iteration scheme but using the above described integral approach. The problem, therefore, has become a zero boundary, fixed source problem. To solve it we choose to use a SC scheme with angular Sn discretization and first grade polynomial approximation for the scattering source.

2.2.3 Step Characteristics Method

Using a general quadrature set of order K for the angular discretization, the general equation, or angular transport equation, in direction k is:

$$\mu_k \frac{\delta\psi(x,y)}{\delta x} + \eta_k \frac{\delta\psi(x,y)}{\delta y} + \sigma_{tot}\psi_k(x,y) = q_k(x,y) \tag{7}$$

where $\psi(x,y)$ is the angular flux at position (x,y), $\psi_k(x,y)$ is the angular flux in direction k at position (x,y), σ_{tot} is the macroscopic total interaction cross section in the cell $q_k(x,y)$ is the source term in direction k at location (x,y), and μ_k and η_k are the cosine and sine of direction k with respect to the x-axis. Note that the total cross section (σ_{tot}) is the sum of the scattering cross section (σ_{sca}) and the absorption cross section (σ_{abs}).

Considering the discretization presented in Fig. 7 if (x,y) are in cell (i,j) then:

$$q_{k,i,j} = S_{k,i,j} + \sigma_{sca} \sum_{k'=1}^{K} a_{k'} P_{kk'} \psi_{k',i,j} \tag{8}$$

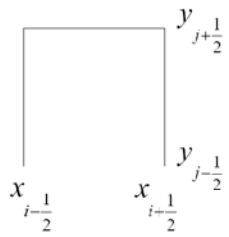


Fig. 10. Cell for the transport equation integration.

Equation (8) presents a piecewise constant source for the cell, where the probability to scatter in one direction is treated as generally dependent on the direction. This general dependence will be presented later in this section. To solve this equation a numerical method has to be applied and therefore a less complicate form has to be found. First, the equation is integrated over the cell shown in Fig. 10 as follows:

$$\frac{1}{\Delta x_i}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} dx \frac{1}{\Delta y_i}\int_{y_{i-\frac{1}{2}}}^{y_{i+\frac{1}{2}}} dy\left(\mu_k \frac{\delta\psi(x,y)}{\delta x}+\eta_k \frac{\delta\psi(x,y)}{\delta y}+\sigma_{tot}\psi_k(x,y)\right)$$
$$=\frac{1}{\Delta x_i}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} dx \frac{1}{\Delta y_i}\int_{y_{i-\frac{1}{2}}}^{y_{i+\frac{1}{2}}} dy q_k(x,y) \tag{9}$$

Equation (9), with some integral properties, simplifies as follows:

$$\frac{1}{\Delta x_i}\mu_k \frac{1}{\Delta y_i}\int_{y_{i-\frac{1}{2}}}^{y_{i+\frac{1}{2}}} dy\left(\psi\left(x_{i+\frac{1}{2}},y\right)-\psi\left(x_{i-\frac{1}{2}},y\right)\right)$$
$$+\frac{1}{\Delta y_i}\eta_k \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} dx\left(\psi\left(x,y_{j+\frac{1}{2}}\right)-\psi\left(x,y_{j-\frac{1}{2}}\right)\right)+\sigma_{tot}\psi_k(x,y)=$$
$$=q_k(x,y) \tag{10}$$

This is because the partial derivative in *x,* integrated between the *x*-cell range, gives the value of the flux (still a function of *y*) at $x_{i+1/2}$ minus the value of the flux (still a function of *y*) at $x_{i-1/2}$. The same is valid for the integral of the partial derivative in *y*. At this point, the integral in *y* and *x* are the average value of the flux along the right and left and the top and bottom edges of the cell:

$$\psi_{k,i\pm\frac{1}{2},j}=\frac{1}{\Delta y_j}\int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}}\psi_k\left(x_{i\pm\frac{1}{2}},y\right)dy \tag{11}$$

and

$$\psi_{k,i,j\pm\frac{1}{2}} = \frac{1}{\Delta x_i}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \psi_k\left(x, y_{j\pm\frac{1}{2}}\right)dx \tag{12}$$

For completeness, to derive Eq. (10) also the average cell-interior flux is presented:

$$\psi_{k,i,j} = \frac{1}{\Delta x_i}\frac{1}{\Delta y_j}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} dx \int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} \psi_k(i,j)\,dy \tag{13}$$

In the derivation of Eq. (10), the cross section has been averaged along the cell. The assumption in this case is that the properties are homogeneous in the cell. The cross sections are therefore kept constant in the derivation of the discretized transport equation per cell:

$$\sigma_{tot,i,j} = \frac{\int dx \int dy \psi(x,y)\sigma_{tot}}{\int dx \int dy \psi(x,y)} \tag{14}$$

With these definitions Eq. (10) assumes its final form for the cell *(i,j)* and direction *k*:

$$\mu_k \frac{\psi_{k,i+\frac{1}{2},j} - \psi_{k,i-\frac{1}{2},j}}{\Delta x_i} + \eta_k \frac{\psi_{k,i,j+\frac{1}{2}} - \psi_{k,i,j-\frac{1}{2}}}{\Delta y_j} + \sigma_{tot}\psi_{k,i,j} = q_{k,i,j} \tag{15}$$

Eq. (15) represents the equation that has to be solved in order to describe the behavior of the neutrons inside the medium. It is an equation in terms of four surface fluxes and one average flux defined as shown in Fig. 11.

Assuming that the intensity and the direction of the incoming flux are known, for example from the right bottom of the cell (mathematically $\mu_k > 0$, $\eta_k > 0$ and, $\psi_{i-1/2,j}$, $\psi_{i,j-1/2}$ are known), Eq. (15) contains three unknowns. In order to reach a solution of the system, two more independents equations must be found. It is important to note that at this point, the derivation contains only the assumptions inherent to the neutron transport equation [i.e. neutrons do not interact with other neutrons and no external forces (gravity, electromagnetic forces…) do not interact on the neutrons] and the assumption of constant cross sections over a cell. These assumptions are generally excellent for all problems of interest to this technique. To complete the derivation though, two additional equations need to be introduced to allow for a well-posed problem. The introduction of these two equations is the primary source of uncertainty in this derivation.

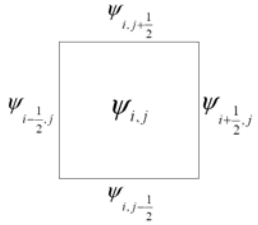


Fig. 11. Angular fluxes per cell representation.

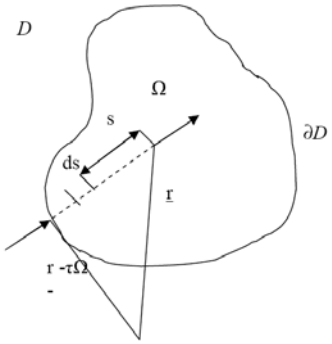


Fig. 12. Plot for the characteristic assumption.

The SC scheme has its own closure equations. An assumption made in the derivation of the Boltzmann equation is that the particles travel in straight lines between collisions. These straight lines are characteristics of Eq. (16) which can be written as:

$$\frac{d\psi}{ds} + \sigma_t \psi = S \tag{16}$$

where S is the source term in the considered direction, s is the distance shown in Fig. 12 and the derivative is along the direction of motion. Integrating Eq. (16) over a domain D with boundary ∂D, gives:

$$\psi = \psi_0 e^{-\sigma s} + \int_0^s S e^{-\sigma(s-s')} ds' \tag{17}$$

where ψ_o is the flux on the boundary, s is the distance from the boundary to the point in which ψ is evaluated and s' is the point in which the source S is evaluated. Physically, if there is a domain D with boundary ∂D (see Fig. 12), the angular flux, at any point $\underline{r}$ into the domain with a certain direction Ω is due to the summation of two contributions:

- the un-collided flux from some known point on the boundary with the same direction and
- the contribution from the source between $\underline{r}$ and the known point in the investigated direction.

For example, in the general case the angular flux at $\underline{r}$ with direction Ω will be:

$$\psi(\bar{r}, \Omega) = \psi\left(r - \tau \cdot \Omega, \Omega\right) e^{-\tau \Sigma_t} + \int_0^{\tau} ds' \cdot S_{tot}\left(r - s'\Omega, \Omega\right) e^{-s'\Sigma_t} \tag{18}$$

This was obtained from Fig. 12, omitting energy and time dependencies. This solution is correct in the case of a homogeneous domain but it can easily be derived also in the case of a non-homogeneous domain using the concept of optical distance. The above solution can be mathematically derived from the transport equation with known total source by using an integrating factor. Therefore this procedure can be used to differentiate the angular flux into the cell.

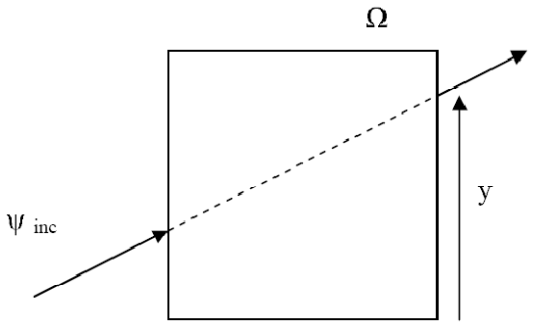


Fig. 13. Derivation of the edge fluxes.

In the case presented in Fig. 13:

$$\psi_k(x_{i+\frac{1}{2}}, y) = \psi_{inc} e^{-\tau\sigma_{tot,i,j}} + \int_0^{\tau} ds' \cdot \frac{q_{k,i,j}}{\sigma_{tot,i,j}} e^{-s'\sigma_{tot,i,j}} \tag{19}$$

Note that *k* indicates one particular angle discretization and in this case the one corresponding to the direction Ω.

Looking at all the possible directions of the angular discretization, it is possible to derive Eq.(19) for all the cases. The SC method derives the necessary closure equations in each quadrant by assuming that the source is constant over the cell and that the incident fluxes are constant over their respective edges.

2.2.4 Derivation of the Governing Equations

For clarity, the equations in a specific case are derived and then generalized. Assuming the direction *k* of the incident flux is well known, Fig. 14 will aid in the derivation of the governing equations. The cell has been divided in regions useful in the integration procedure.

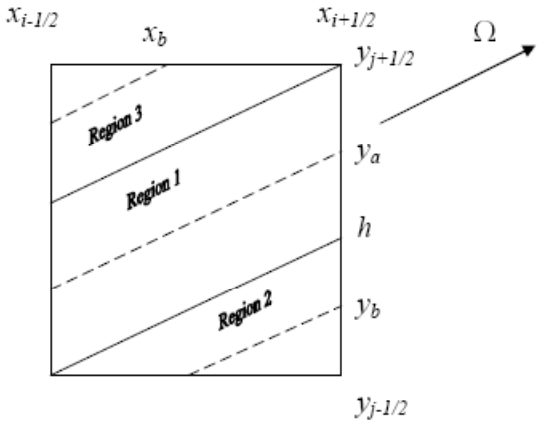


Fig. 14. Cell variables for governing equation derivation.

By definition, see Eq. (19), it can be derived:

$$\psi_k(x_{i+\frac{1}{2}}, y_a) = \psi_{k,i-\frac{1}{2},j} e^{-\sigma_t \frac{\Delta x_i}{\mu_k}} + \frac{q_{k,i,j}}{\sigma_t}\left(1 - e^{-\sigma_t \frac{\Delta x_i}{\mu_k}}\right) \tag{20}$$

$$\psi_k(x_{i+\frac{1}{2}}, y_b) = \psi_{k,i,j-\frac{1}{2}} e^{-\sigma_t \frac{\left(y_b - y_{j-\frac{1}{2}}\right)}{\eta_k}} + \frac{q_{k,i,j}}{\sigma_t}\left(1 - e^{-\sigma_t \frac{\left(y_b - y_{j-\frac{1}{2}}\right)}{\eta_k}}\right) \tag{21}$$

Eqs. (20) and (21) have been derived by using the closure assumption given by the SC scheme. The source and cross sections have been kept constant in the cell and the integrals have become functions only of geometric parameters.

Eqs. (11) and (12) define the average edge flux:

$$\psi_{k,i+\frac{1}{2},j} = \frac{1}{\Delta y_j} \int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} \psi_k(x_{i+\frac{1}{2}}, y) dy \tag{22}$$

that becomes in the case of Fig. 14:

$$\psi_{k,i+\frac{1}{2},j} = \frac{1}{\Delta y_j} \int_{h}^{y_{j+\frac{1}{2}}} \psi_k(x_{i+\frac{1}{2}}, y_a) dy_a + \frac{1}{\Delta y_j} \int_{y_{j-\frac{1}{2}}}^{h} \psi_k(x_{i+\frac{1}{2}}, y_b) dy_b \tag{23}$$

The first integral in Eq. (23) does not depend directly on y_a since Eq. (20) does not depend on y_a. Its solution will therefore be straightforward:

$$\left[\left(\psi_{k,i-\frac{1}{2},j} - \frac{q_{k,i,j}}{\Sigma_t}\right) e^{-\sigma_t \frac{\Delta x_i}{\mu_k}} + \frac{q_{k,i,j}}{\sigma_t}\right]\left(y_{j+\frac{1}{2}} - h\right)\frac{1}{\Delta y_j} \tag{24}$$

The second integral is instead more complicate and the result is:

$$\left[\left[\psi_{k,i,j-\frac{1}{2}} - \frac{q_{k,i,j}}{\sigma_t}\right]\left[1 - \frac{\eta_k}{\sigma_t} e^{-\Sigma_t \frac{\left(h - y_{j-\frac{1}{2}}\right)}{\eta_k}}\right] + \frac{q_{k,i,j}}{\sigma_t}\left(h - y_{j-\frac{1}{2}}\right)\right]\frac{1}{\Delta y_j} \tag{25}$$

Combining Eqs. (24) and (25), the solution for the case introduced in Fig. 14 is obtained:

$$\psi_{k,i-\frac{1}{2},j}=\left(\left(\psi_{k,i-\frac{1}{2},j}-\frac{q_{k,i,j}}{\Sigma_t}\right)\left(y_{j+\frac{1}{2}}-h\right)e^{-\Sigma_t\frac{\Delta x_i}{\mu_k}}+\frac{q_{k,i,j}}{\Sigma_t}\left(y_{j+\frac{1}{2}}-h\right)+\right.$$
$$\left.+\frac{\eta_k}{\Sigma_t}\left[\psi_{k,i,j-\frac{1}{2}}-\frac{q_{k,i,j}}{\Sigma_t}\right]\left[1-e^{-\Sigma_t\frac{\left(h-y_{j-\frac{1}{2}}\right)}{\eta_k}}\right]+\frac{q_{k,i,j}}{\Sigma_t}\left(h-y_{j-\frac{1}{2}}\right)\right)\frac{1}{\Delta y_j} \tag{26}$$

In Eq. (26) the parameter h can be substituted, by geometrical consideration, as follows:

$$h=\Delta x_i\frac{\eta_k}{\mu_k}+y_{j-\frac{1}{2}} \tag{27}$$

Rearranging the final form is obtained:

$$\psi_{k,i+\frac{1}{2},j}=\left(\psi_{k,i-\frac{1}{2},j}-\frac{q_{k,i,j}}{\sigma_t}\right)\left(1-\frac{\Delta x_i}{\mu_k}\frac{\eta_k}{\Delta y_j}\right)e^{-\sigma_t\frac{\Delta x_i}{\mu_k}}+$$
$$+\left(\psi_{k,i,j-\frac{1}{2}}-\frac{q_{k,i,j}}{\sigma_t}\right)\left(\frac{\Delta x_i}{\mu_k}\frac{\eta_k}{\Delta y_j}\right)\left[\frac{1-e^{-\sigma_t\frac{\Delta x_i}{\mu_k}}}{\sigma_t\frac{\Delta x_i}{\mu_k}}\right]+\frac{q_{k,i,j}}{\sigma_t} \tag{28}$$

This final form is the solution for the angular flux at the right edge of the cell given an incident beam coming from the left and bottom edge in direction k. The same relation can be derived for all the angular fluxes with all the different boundary condition applied (different direction k).

The generalized result is presented:

$$\psi_{k,i+\frac{m}{2},j}=\frac{q_{i,j}}{\sigma_{i,j}}+\nu_{k,i,j}\left(\psi_{k,i,j-\frac{n}{2}}-\frac{q_{i,j}}{\sigma_{i,j}}\right)\left(\frac{1-e^{\tau_{k,i,j}}}{\tau_{k,i,j}}\right)+\left(1-\nu_{k,i,j}\right)\left(\psi_{k,i-\frac{m}{2},j}-\frac{q_{i,j}}{\sigma_{i,j}}\right)e^{\tau_{k,i,j}} \tag{29}$$

and

$$\psi_{k,i,j+\frac{n}{2}}=\frac{q_{i,j}}{\sigma_{i,j}}+\alpha_{k,i,j}\left(\psi_{k,i-\frac{m}{2},j}-\frac{q_{i,j}}{\sigma_{i,j}}\right)\left(\frac{1-e^{\tau_{k,i,j}}}{\tau_{k,i,j}}\right)+\left(1-\alpha_{k,i,j}\right)\left(\psi_{k,i,j-\frac{n}{2}}-\frac{q_{i,j}}{\sigma_{i,j}}\right)e^{\tau_{k,i,j}} \tag{30}$$

The closure equations depend on the quadrant and direction of interest (Ω_k):

$$m=\begin{cases}+1 & \mu_k>0\\ -1 & \mu_k<0\end{cases} \tag{31}$$

and

$$m=\begin{cases}+1 & \eta_k>0\\ -1 & \eta_k<0\end{cases} \tag{32}$$

The constants used in the closure equations are defined as:

$$\tau_{m,i,j} = \sigma_{i,j} \min\left\{\frac{\Delta x_i}{|\mu_k|}, \frac{\Delta y_j}{|\eta_k|}\right\} \tag{33}$$

and

$$\alpha_{m,i,j} = \min\left\{1, \left|\frac{\mu_k \Delta y_j}{\eta_k \Delta x_i}\right|\right\} \tag{34}$$

and

$$\nu_{m,i,j} = \min\left\{1, \left|\frac{\eta_k \Delta x_i}{\mu_k \Delta y_j}\right|\right\} \tag{35}$$

From the balance equation, Eq. (15), the cell-interior angular flux is:

$$\psi_{k,i,j} = \frac{q_{k,i,j}}{\sigma_{tot}} - \mu_k \frac{\psi_{k,i+\frac{1}{2},j} - \psi_{k,i-\frac{1}{2},j}}{\Delta x_i \sigma_{tot}} - \eta_k \frac{\psi_{k,i,j+\frac{1}{2}} - \psi_{k,i,j-\frac{1}{2}}}{\Delta y_j \sigma_{tot}} \tag{36}$$

The discretized cell-interior angular flux is related to the scalar flux and currents as given by the discrete ordinates approximation:

$$\phi_{i,j} = \sum_{k=1}^{K} w_k \psi_{k,i,j} \tag{37}$$

$$J_{x_{i,j}} = \sum_{k=1}^{K} w_k \mu_k \psi_{k,i,j} \tag{38}$$

$$J_{y_{i,j}} = \sum_{k=1}^{K} w_k \eta_k \psi_{k,i,j} \tag{39}$$

The scattering cross section is considered linearly anisotropic. The scattering source, therefore, as defined in Eq. (8), is also linearly anisotropic and it can be calculated for every direction k from the scalar flux and the currents. Remembering the presence of the un-collided flux and currents per cell, these flux and currents will constitute fixed terms for the source:

$$q_{k,i,j} = \frac{\sigma_{sca}}{4\pi}\left(\phi_{i,j} + \phi_{i,j}^{uncollided}\right) + \frac{\sigma_{sca}}{4\pi} 3 g \mu_k \left(Jx_{i,j} + Jx_{i,j}^{uncollided}\right)$$
$$+ \frac{\sigma_{sca}}{4\pi} 3 g \eta_k \left(Jy_{i,j} + Jy_{i,j}^{uncollided}\right) + S_{k,i,j} \tag{40}$$

where g is called the anisotropic factor. g varies between 0 and 0.333 defining a scattering source more and more forward peaked. When g equals 0 the scattering source is isotropic.

Knowing the boundary conditions at the edges of the medium and using Eqs. (29) and (30), it is possible to find all the fluxes of the cells by iteratively sweeping through the medium. At the first iteration, the source will consist only of un-collided particles. The new sources created during

the sweeping process will be the sources of once-collided particle. Sweeping again through the object will solve for the once collided particles and create the source of the twice-collided particles. And so on until convergence, that is until the next source is within some ε of the previous one. The order of sweeping in the case of cell-centered spatial mesh and a know boundary conditions direction is presented in Fig. 15[24].

Lathrop[25] enumerated a list of desirable properties of a differencing scheme:

- it should be accurate in the sense that it has a small truncation error
- it should be simple (which we interpret to mean that it should involve a small number of numerical operations and should involve unknowns from within a single mesh cell)
- it should produce positive fluxes if the source and boundary fluxes are positive.

To this list, he added the requirements that

- the scheme has to be conservative, in the sense that a well defined relation exists among flows into and out of the cell and sources and absorptions in the cell and
- the scheme should be easily generalizable to all geometries[25].

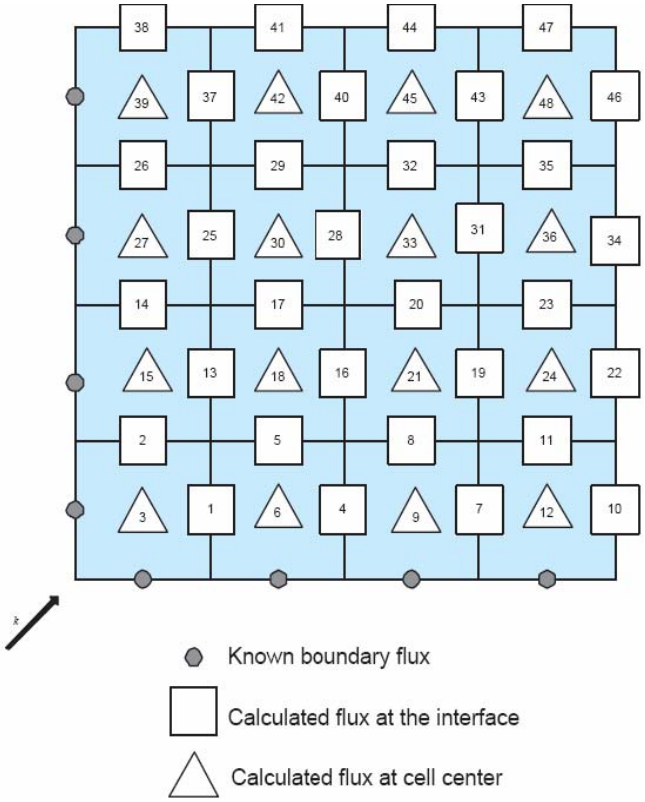


Fig. 15. Order of sweeping cell-centered spatial mesh for $\mu_k>0$, $\eta_k>0$.

Since the definition in Eqs. (11), (12) and (13) satisfy Eq. (15), the SC scheme is conservative. Given positive incoming fluxes and source, the out-going fluxes [due to Eqs. (29) and (30)] are

also positive. The truncation error, however, is something less than a second order error. It is possible to show that the relations of Eqs. (29) and (30) have first order truncation error which depends on $\sigma\Delta x/\mu$, $\sigma\Delta y/\eta$ and their ratio. The error is largest for large $\sigma\Delta x/\mu$ and $\sigma\Delta y/\eta$ or when their ratio is near zero. The relations approach second order accuracy for small $\sigma\Delta x/\mu$ and $\sigma\Delta y/\eta$ and when their ratio is near unity. The scheme is also difficult to generalize; there are not many geometries, beyond Cartesian, in which it is easy to integrate along the characteristics.

The main difficulty in using the SC scheme is its tendency to spread the beam along the object. A perfect forward beam will be transformed into a more spread beam due to the averaging of the flux along the edges of a cell.

The problem of interesting has the best geometry to develop the simple SC method since it has good boundary conditions treatment and highly angular-dependent fluxes. All the equations are easy exponential relations and are suitable for a further analytical analysis (as will be developed in Section 3).

At this point, the problem is solved in terms of average flux and net currents in x and y per every cell. With this information the second step is finished and the detector readings can be calculated. The detector reading will be constructed taking into account the un-collided radiation coming directly from the beam window and the collided radiation. The collided component of the detector will be simply the outgoing current at the edge correspondent to the detector. The details about the calculation of the detector readings are presented in the next sections after some notes about the angular quadrature set.

2.2.5 Level Symmetric Quadrature Set

In the derivation of the solution for the angular fluxes in the previous sections, one angle is needed to specify the angular distribution. The one used in this process is shown in Fig. 16. In multidimensional problems, two angular coordinates are required to specify the direction of neutron travel (Ω. If x_1, x_2 and x_3 are orthogonal spatial coordinates, μ, η, and ξ are the direction cosines of Ω with respect to these coordinates as indicated in Fig. 16.

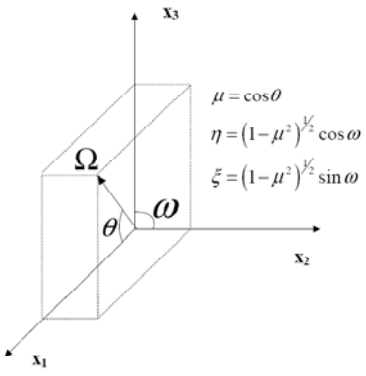


Fig. 16. Angular coordinate system.

Only two of the direction cosines may be specified independently, however, since Ω is a unit vector[23]. Hence they must satisfy:

$$\mu^2 + \eta^2 + \xi^2 = 1 \tag{41}$$

In three-dimensional problems $\psi(x,y,\Omega)$ must be determined over all eight octants of the unit sphere swept out by Ω. In two-dimensional geometries, the mirror symmetry of $\psi(x,y,\Omega)$ reduces the number of octants over which the angular dependence of $\psi(x,y,\Omega)$ must be determined to four.

The purpose of introducing a quadrature set for the direction Ω, is to calculate, in the transport equation, the following operator:

$$\int_{4\pi} d\Omega\, f(\Omega) \Rightarrow \sum_{m=1}^{M} p_m f(\Omega_m) \tag{42}$$

where p_m is the point weight and Ω_m are the "discrete ordinates." A "quadrature set" specifies the discrete ordinates at which the function is evaluated as well as the weight associated with each ordinate.

The quadrature points or "ordinates" in a level symmetric set are arranged on the principal octant in a triangular fashion on N/2 levels. Each level has $^{N}/_{2}$-n+1 points, where n=1…$^{N}/_{2}$. With this arrangement, there are $^{N(N+2)}/_{8}$ directions in the unit octant.

Each quadrature point represents a direction in the unit sphere specified by three direction cosines:

$$\Omega_m = \mu_i \hat{\mu} + \eta_j \hat{\eta} + \xi_k \hat{\xi} \tag{43}$$

Furthermore, the set of direction cosines obeys the relation[26-28]:

$$\mu_i^2 = \mu_1^2 + C(i-1) \tag{44}$$

with the constant C given by

$$C \equiv \frac{2(1-3\mu_1^2)}{N-2} \tag{45}$$

In addition to the ordinates, the weights in each octant must be determined. The normalization relation is:

$$\sum_{n=1}^{N(N+2)/8} w_n = 1 \tag{46}$$

More details about the fundamental equations and properties of the level symmetric quadrature set can be easily found in the literature[24].

2.2.6 Detector Reading

The last step of the forward method is the detector readings. The calculation of the detector responses is almost identical to the calculation of the initial un-collided sources except for the addition of the collided component. The detector readings consist of two components: the un-collided radiation and the collided radiation. The detector readings are given by:

$$P(\text{detector}) = P_{unc}(\text{detector}) + P_{coll}(\text{detector}) \tag{47}$$

where P indicates "prediction".

The un-collided particles are those that come directly from the beam window to the detector without undergoing a collision event, see Fig. 17. The calculation of this component is similar to that used for the un-collided sources:

- the detector is divided in points whose location into the detector and interval weights are calculated by using a Gauss-Legendre quadrature set
- for every point the un-collided current is calculated by performing an integral similar to the one of Eq. (6), again finding the interval for the θ angle, see Fig. (12)
- the reading of the detector is obtained averaging the readings of its points

The integral for the calculation of the un-collided radiation, performed as a quadrature summation is:

$$\begin{aligned} P_{un-collided}(\text{detector}) &= \sum_{r=1}^{\#\ \text{dec. points}} P_{inc}(\text{detector,point}) \\ &= \sum_{r=1}^{R} \sum_{poi=1}^{POI.} \sum_{az=1}^{AZ.} 2 w_r w_{poi} w_{az} \psi_{beam} e^{-\alpha ds} (\eta \text{ or } \mu) \end{aligned} \tag{48}$$

where:

- the first summation in r averages the detector point readings indicating with R the number of points per detector
- the second and third summations perform the integral along the two angles of the beam, coming from the window, decreased by the attenuation factor (in the exponent ds is the distance between beam window and detector point); POL and AZ are the maximum numbers of polar and azimuthal angles
- the angle to multiply with is μ or η. If the current along the x-axis is needed the multiplication will be with μ, otherwise with η.

The number of points per detector, and angles for the quadrature integral are user-defined. Again by using a reasonable number of these parameters leads to an efficient calculation of the uncollided radiation in the detectors.

The collided component is simply the current out of the cell-edge corresponding to the detector. An important requirement is that the detector is placed corresponding to the edge of a cell or multiple cells such that it is clear which cells edges constitutes the detector reading. The two components are summed together to have the total prediction. Fig. 17 presents this process graphically.

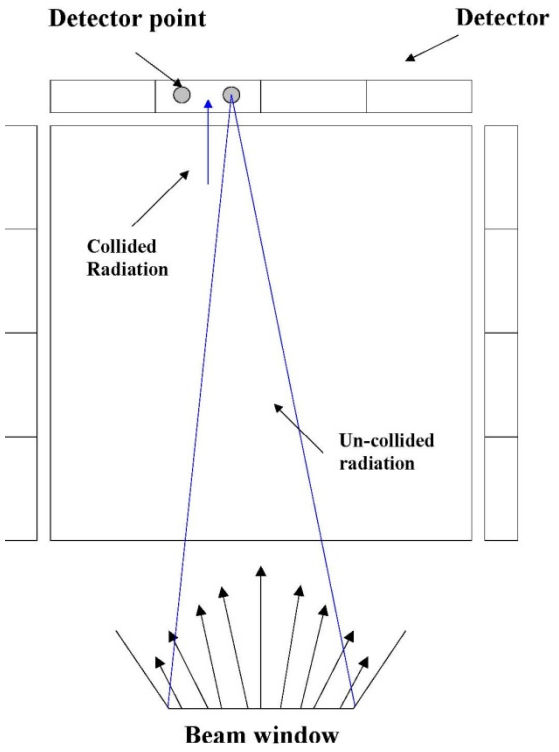


Fig. 17. Detector reading calculation.

2.3. Implementation and Computational Verification

The code was tested using different techniques and different testing problems. During the coding process every part of the code was tested against analytical simple calculations. At the end, when the coding was finished it was completely tested against MCNP results.

2.3.1 FORTRAN Coding

The code has been developed in Fortran 90 on a Microsoft Windows workstation with a 3.20 GHz Pentium 4 processor. Three input files are used for the forward model:

- input file for the geometry of the beam and collimator that indicates the order of the angular quadrature set and the number of each parameter used for the un-collided source calculation
- input file for the geometry of the problem. Dimension of the object, number of regions and cells per region
- input file for the cross sections per region.

The output of the code is contained on different files containing:

- summary of the problem
- scalar flux per cell
- current in x and y direction per cell flux

2.3.2 MCNP Decks

Two test objects were explicitly simulated using MCNP-5. One object was homogeneous. The other object was heterogeneous. The MCNP object has been discretized using a lattice subroutine such that the flux into the object can be compared with that calculated by the code. Fig. 18 presents a visual image of the heterogeneous test object showing with all the geometrical properties of the beam and the object.

2.3.3 Homogeneous Test Object

The object investigated is placed into a neutron beam collimated by using a slightly divergent collimator. In both MCNP and the code the fluxes inside the object (in the cell of each spatial discretization) were calculated. The object consisted of natural carbon with a 10 cm by 10 cm square area. The beam was symmetric about the centerline of the object and the beam window was 2 cm wide. The collimator was 3 cm long and with an exit window 4 cm wide. For this object no detectors were around the object. The testing and verification was performed to assure accuracy in the scalar flux in the object cells. The spatial discretization used in MCNP consisted of 20 by 20 uniformly distributed square cells. In the code to reach we used a grid of 200 by 200 cells. Post-calculation, groups of 10 by 10 cells were collapsed to calculate the averaged flux which could then be compared with MCNP. In MCNP, reflective boundary conditions in the third dimension were used to construct a real two dimensional problem. A summary of the properties and the result of the comparison is shown in Table I.

The fluxes were compared at three different positions inside the object: at the entry (first row of cells), at the centerline, and at the exit. The g factor, which measures the anisotropic nature of the scattering, was varied between 0 and 0.07 to determine its affect on the result. The results are presented in Figs. 18-20. The plots present in the abscissa the cell number and in ordinate the strength of the sources inside that cell, with respect to the g factor.

With a g factor equal to 0.07, the MCNP result and the code calculation agree within 2% everywhere. These results demonstrate that the forward model can accurately calculate flux values that agree with MCNP simulations. To ensure code verification and code reliability a heterogeneous object test was performed as well.

TABLE I Summary of the homogeneous case test problem

Object Dimensions		
x length	10	cm
y length	10	cm
Collimator Dimensions		
Beam window length	2	cm
Collimator window length	3	cm
Distance beam window object	4	cm
Distance collimator window object	1	cm
CODE DISCRETIZATION		
Discretization Forward Problem		
# x cell	200	
# y cell	200	
Order quadrature set (n of Sn)	8	
# polar direction	20	
Discretization for Unc. Sources		
# points per cell	1	
# azimuthal angles	20	
# polar angles	20	
Optical Properties		
Scattering cross section	0.538678	cm^{-1}
Absorption cross section	0.000382	cm^{-1}
Total cross section	0.53906	cm^{-1}
g factor	variable	
MCNP INPUT		
# lattice cells	20 by 20	
# particles	10^8	

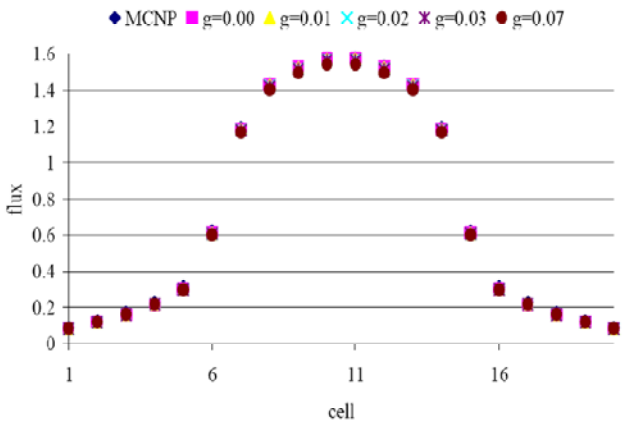


Fig. 18. Comparison of scalar fluxes for the bottom cells using different values for *g*.

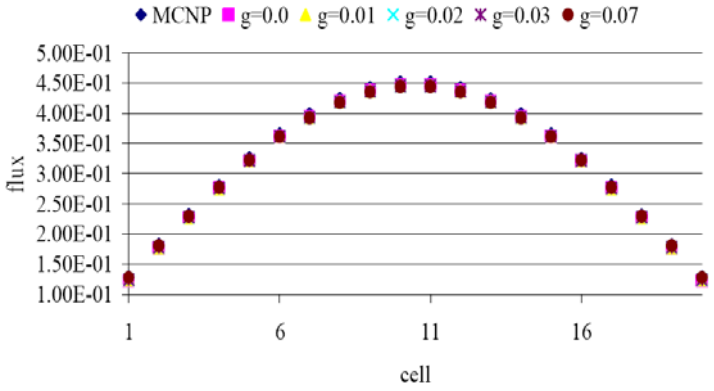


Fig. 19. Comparison of the scalar fluxes for the center cells using different values for *g*.

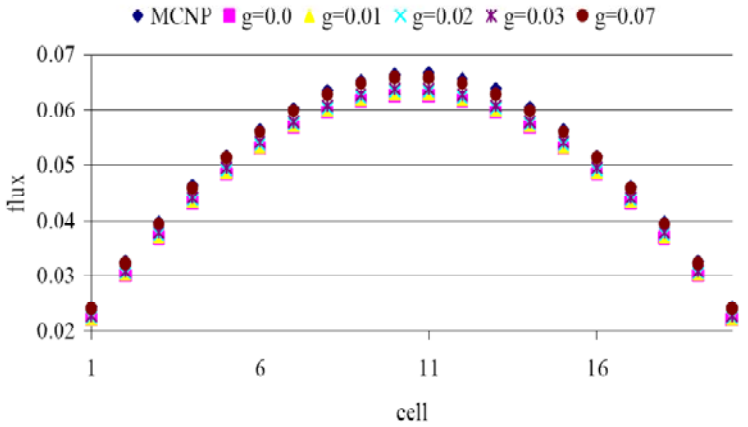


Fig. 20. Comparison of the scalar fluxes for the top cells using different values for g.

2.3.4 Heterogeneous Test Object

The heterogeneous object was similar to the previous object except for the inclusion of a 1 cm by 1.5 cm boron defect. This defect was placed 1 cm away from the top and right edges. All the variables related with the beam remained the same. Fig. 21 shows a graphical depiction of the object. A summary of the optical properties of the object is listed in Table II.

The results of the comparison are presented in Figs. 22-25. The plots present in abscissa the cell number and ordinate the strength of the source. As in the homogeneous case, the data is collected at different location throughout the object: entry (first row of cells), object centerline and exit (last row of cells). In addition, for the heterogeneous case, a plot is taken also at the centerline of the boron defect (Fig. 24).

The MCNP and code calculation agree within 5% along the top edge of the object. This level of accuracy is sufficient for the purpose of this study. Together with the homogeneous test, the heterogeneous test shows accurate calculations for the code. The forward code appears to be correctly implemented and verified.

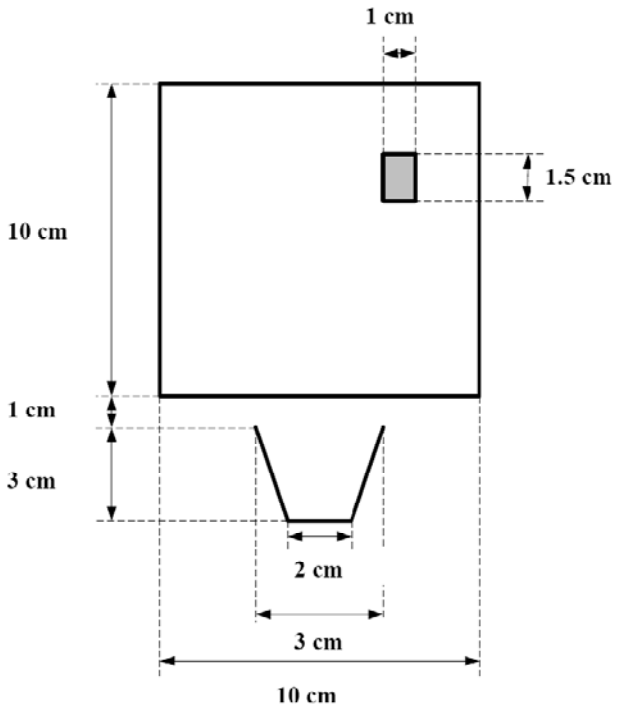


Fig. 21. Heterogeneous test object.

TABLE II Properties of the heterogeneous case

Object Dimensions		
x length	10	cm
y length	10	cm
Collimator Dimensions		
Beam window length	2	cm
Collimator window length	3	cm
Distance beam window object	4	cm
Distance collimator window object	1	cm
CODE DISCRETIZATION		
Discretization Forward Problem		
# x cell	200	
# y cell	200	
Order quadrature set (n of Sn)	8	
# polar direction	20	
Discretization for Unc. Sources		
# points per cell	1	
# azimuthal angles	20	
# polar angles	20	
Optical Properties Background		
Scattering cross section	0.538678	cm^{-1}
Absorption cross section	0.000382	cm^{-1}
Total cross section	0.53906	cm^{-1}
g factor	0.07	
Optical Properties Intrusion		
Scattering cross section	0.293788	cm^{-1}
Absorption cross section	525.7772	cm^{-1}
Total cross section	526.071	cm^{-1}
g factor	0.07	
MCNP INPUT		
# lattice cells	20 by 20	
# particles	10^8	

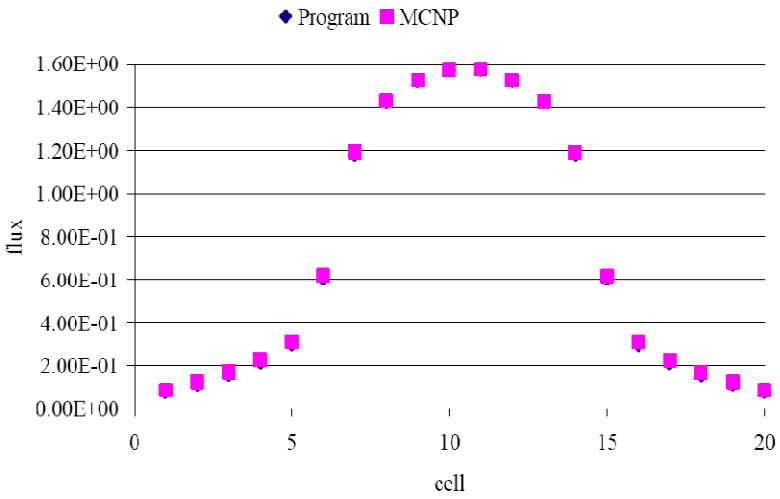


Fig. 22. Comparison between MCNP and code scalar fluxes for cell at the bottom part of the object.

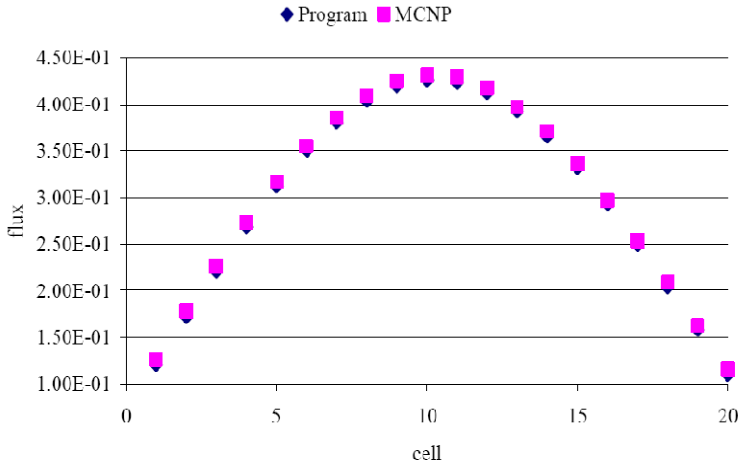


Fig. 23. Comparison between MCNP and code scalar fluxes for the cells at the center of the object.

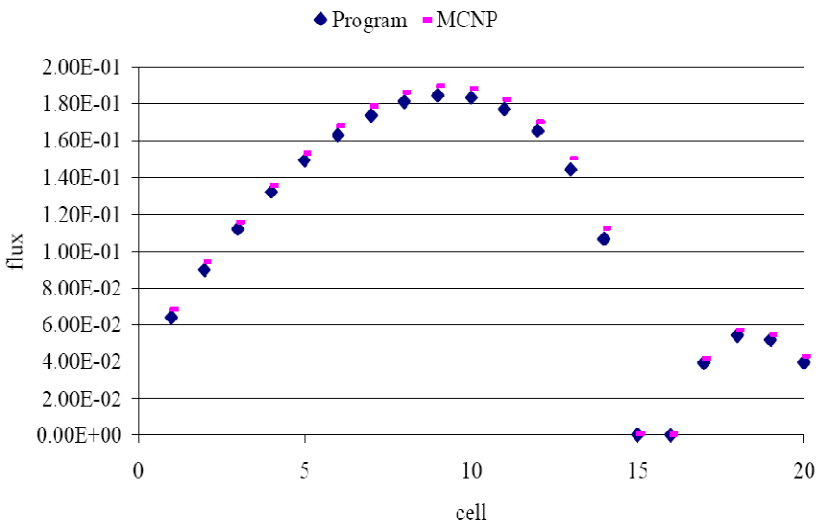


Fig. 24. Comparison between MCNP and code scalar fluxes for the cells at the boron defect centerline.

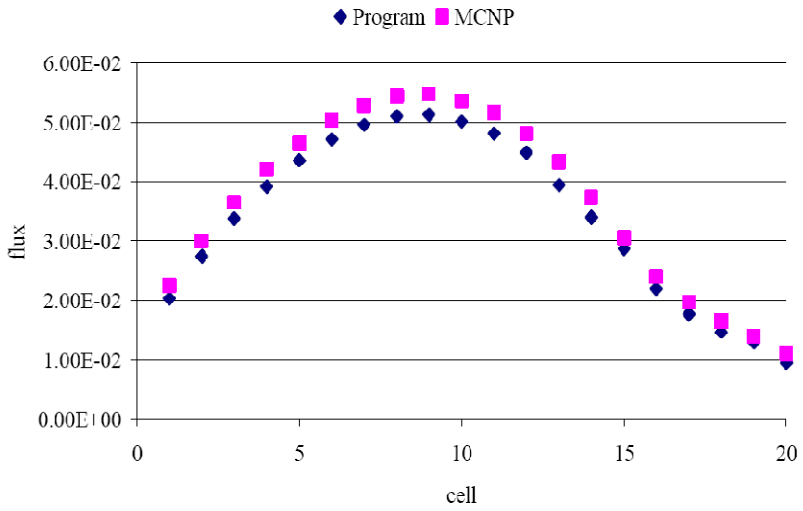


Fig. 25. Comparison between MCNP and code scalar fluxes for the cells at the top part of the object.

3. INVERSE MODEL

The focus of this section is to present the scheme to solve the inverse problem and therefore determine the optical properties distribution for the object given a measurement of transmitted and scattered image. Given the measurements performed on the surface of the object and the predicted measurements calculated by using the forward method presented in the previous chapter, the scheme has to be able to reconstruct the cross sectional image of the optical properties of the media investigated. The imaging problem is treated as an optimization problem, in which an objective function is minimized. The objective function has been calculated as a normalized-squared error between predicted and measured data.

The inverse model makes use of the conjugate gradient scheme for the minimization and it finds the gradient of the objective function with respect to all optical properties. This gradient is the major step of the updating scheme and its calculation is performed using an adjoint differentiation algorithm that allows for an efficient and accurate calculation.

The complete code, consisting of both forward and inverse model, will be referred as TNTs (Tomography Neutron Transport using Scattering).

3.1. Perturbation Methods

The majority of available reconstruction algorithms are based on perturbation methods[29-34]. The limited application of this scheme comes from its main assumption that the variations between the optical properties of the medium to reconstruct and those used as an initial guess in the forward model are small. These methods are also computationally expensive because they involve the inversion of full ill-conditioned Jacobian matrixes[7].

Perturbation theory is the study of the effects of small disturbances in the mathematical model of a physical system. Assuming that the optical properties of the unknown object are a small perturbation of an estimated distribution it is possible to reconstruct the object by using a perturbation model. Experimental measurement are taken along the boundary of the investigated medium and compared with the prediction given by a forward method that makes use of the estimated distribution. If ξ_e, the estimated distribution of optical properties, is close to ξ_r, the real distribution, a Taylor series for the measurements (M) at the boundary locations can be performed:

$$\underline{M} = f\left(\underline{\xi_e}\right) + J\left(\underline{\xi_e}\right)\left(\underline{\xi_r} - \underline{\xi_e}\right) + \left(\underline{\xi_r} - \underline{\xi_e}\right)^T H\left(\underline{\xi_e}\right)\left(\underline{\xi_r} - \underline{\xi_e}\right) \tag{49}$$

where f represents the forward model and J and H are the Jacobian and the Hessian of the forward scheme. All the underlined quantities are vectors. If $\underline{M}_{exp}$ is the vector of the measured data, then the difference between experimental values and predicted values $[f(\xi_e)]$ is defined as follows:

$$\underline{M}_{exp} - f\left(\underline{\xi_e}\right) \cong \underline{M} - f\left(\underline{\xi_e}\right) = J\left(\underline{\xi_e}\right)\underline{\Delta\xi} + \underline{\Delta\xi}^T H\left(\underline{\xi_e}\right)\underline{\Delta\xi} \tag{50}$$

where $\underline{\Delta\xi}=(\underline{\xi_r}-\underline{\xi_e})$.

If the second-order term on the right hand side of Eq. (50) is neglected, the dependence between the differences in the measurements is linearly related to the difference in the properties and $\underline{\Delta\xi}$ can easily be found:

$$\underline{M}_{\exp} - f\left(\underline{\xi}_e\right) = J\left(\underline{\xi}_e\right)\Delta\underline{\xi} \;\Rightarrow\; \Delta\underline{\xi} = J\left(\underline{\xi}_e\right)^{-1}\left(\underline{M}_{\exp} - f\left(\underline{\xi}_e\right)\right) \tag{51}$$

Knowing the starting estimated distribution $\underline{\xi}_e$ and $\underline{M}_{exp}$-$f(\underline{\xi}_e)$ leads to the calculation of $\Delta\underline{\xi}$ by the inversion of a full ill-conditioned Jacobian matrix. The approach is usually generalized to an iterative method that at every step evaluates a new Jacobian matrix and a new $\Delta\underline{\xi}$. A regularization term is often used to reduce to a more diagonally dominant matrix the ill-conditioned J matrix[7]. Although such matrixes can be efficiently constructed the method becomes intractable as the size of the problem domain increases.

3.2. Gradient-Based Iterative Scheme

Another approach is to regard the problem as the optimization of an objective function representing the sum-squared difference between measured and predicted detector readings. In Section II, a forward method for the calculation of detector readings was introduced. The predicted measurements are function of the optical properties of the entire object:

$$P = f(\sigma_{scattering}, \sigma_{total}, g_{anisotropicfactor}) = f(\sigma_{sca}, \sigma_{tot}, g) \tag{52}$$

Under the assumption of a maximum-likelihood approach to the solution for the inverse problem, the objective function is defined as:

$$E(\sigma_{sca}, \sigma_{tot}, g) = \frac{1}{2}\sum_{i=1}^{m}\left(M_{\exp} - f(\sigma_{sca}, \sigma_{tot}, g)\right)^2 \tag{53}$$

where m indicates the total number of measurements performed along the domain. From Eq. (53) it is clear that the objective function depends on every optical property of the object. In particular, in the case of a discretized domain such as the one used in the solution of the forward problem, E depends on the properties of every cell:

$$E(\sigma_{sca}(i,j), \sigma_{tot}(i,j), g(i,j)) \qquad i=1..I \text{ and } j=1..J \tag{54}$$

where I and J to indicate the total number of cells in the x and y direction. The problem therefore can be referred as an optimization of the multidimensional function called the objective function.

There are different methods to approach the solution of this problem. The classical unconstrained multivariable optimization is usually approached with Newton-type iterative algorithms that have better results in terms of accuracy and number of iteration than the steepest descent method[36]. However, it is generally very difficult to calculate the Hessian of the function.

In the method of steepest descent, see Figs. 26 and 27, $f(x)$ is evaluated at a certain point $\underline{x}(0)$. The method slides down to the bottom of the paraboloid as shown in Fig. 27. A series of $\underline{x}(1)$, $\underline{x}(2)$, $\underline{x}(3)$,... are taken until the problem is close enough to the solution $\underline{x}_f$. In taking the step, the steepest descent method chooses the direction in which f decreases most rapidly which is the direction opposite to the gradient $f'(\underline{x})$. Indicating with r_i the residual, which in terms of matrix multiplication is nothing but how far the problem is from the solution, it can be shown that r_i=-$f'(x_i)$ and

$$x_{i+1} = x_i + \alpha r_i \tag{55}$$

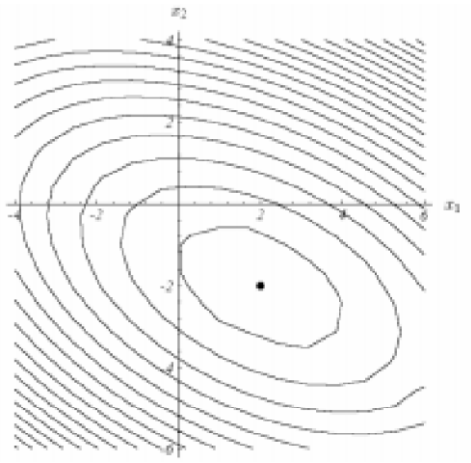


Fig. 26. Contours of a perfect quadratic form (with minimum in (2,-2)). Each ellipsoidal has constant f(x).

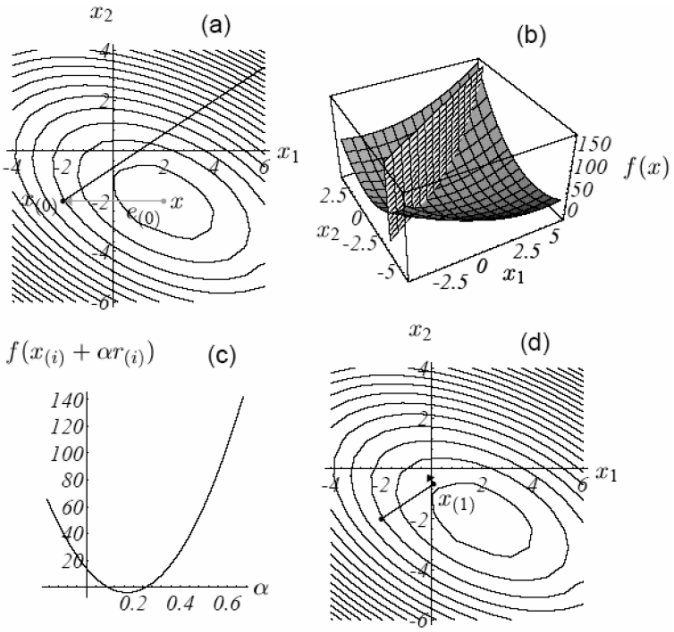


Fig. 27. Method of steepest descent: (a) starting at x(0)=(-2,2), the method steps in the direction of –f', (b) Intersection between these two surfaces, (c) The bottom point of the parabola is the target of the line search, and (d) The gradient at the bottom point is orthogonal to the gradient at the previous step (see Fig. 28).

A line search is the procedure that chooses α to minimize *f* along the line. On the line search *f* is minimized where the gradient is orthogonal to the search line (see Figs. 28 and 29). A variety of line search methods can be used, either utilizing gradient information, such as the secant method, or using only function evaluations such as the quadratic fit method. Often an exact line search is too computationally expensive due to the large number of function derivative

computations. Experience shows that exact line search, minimizing the objective function as accurate as possible, is not necessary for Newton's methods. Typically, a full step is taken (α=1) and if $f(\underline{x}_{i+1})>f(\underline{x}_i)$, then we back-track towards $\underline{x}_i$. Quadratic interpolations are available and inexact line search methods such as Armijo's Rule define the bounds for acceptable step lengths which guarantee convergence. For conjugate gradient or steepest descent methods the precision in the line search is instead vital.

In developing TNTs Brent's method was used[37]. An interval that brackets the minimum is given as input, then Brent's method approaches the minimum by choosing at each step between quadratic fit and golden search.

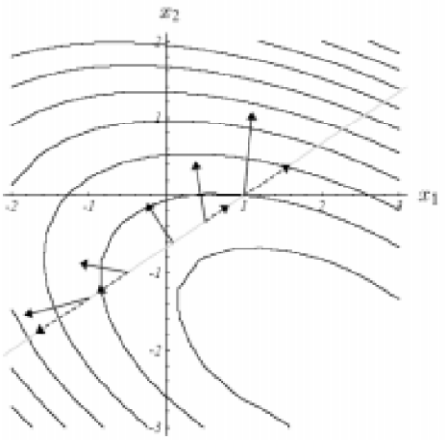


Fig. 28. The gradient f' is shown at different locations along the search line. On the search line f is minimized where the gradient is orthogonal to the search line

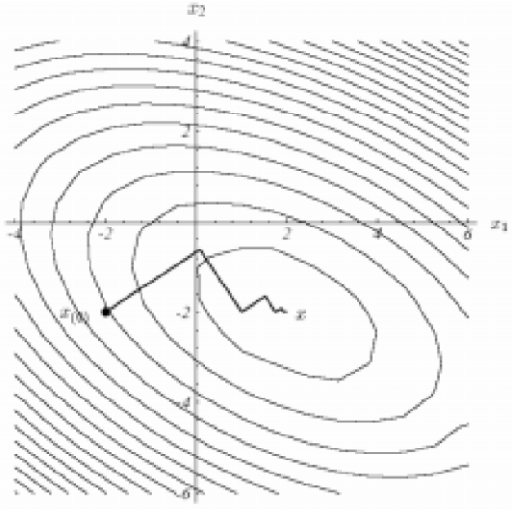


Fig. 29. Here the steepest descent starts at (-2,2) and converges at (2,-2).

The method of the conjugate gradient (CG) is slightly different from the steepest descent. At every step instead of moving along a direction orthogonal to the previous one the CG moves along an *A*-orthogonal direction (*A* is the matrix that defines the quadratic dependence of the function with respect to all the variables). From a more understandable point of view, the CG method tries to minimize the residual instead of the objective function itself. In order to do that, the new direction in every iteration is calculated with a linear interpolation between the old direction and the new gradient. The coefficient of this interpolation (β) varies with the method chosen to couple the CG scheme. The interesting property of CG is that it finds at every step the best solution within the bounds of where it is been allowed to explore. The best property, though, is that the CG can be used not only to find the minimum point of a quadratic form, but to minimize any continuous function $f(\underline{x})$ for which the gradient f' can be computed.

Under the assumption of an effective calculation of the gradient of the objective function $E(\sigma_x)$ (where x can indicate total, scattering or g factor), the choice for the minimization has been the non-linear conjugate-gradient method coupled with the Polack-Ribiere method or with the Fletcher-Reeves formula[38]. The algorithm for this is given by:

Algorithm 1: Non linear conjugate-gradient method

Choose an initial set of variables $\underline{\sigma_x}^1$

Set the initial search direction $\underline{h}^1=\underline{r}^1$. Define $\underline{r}^1=-\nabla E(\underline{\sigma_x}^1)$

Set iteration counter $j=0$

Repeat

Find α_j that minimizes $E(\underline{\sigma_x}^j+\alpha^j\underline{h}^j)$

$\underline{\sigma_x}^{j+1}=\underline{\sigma_x}^j+\alpha_{min}\underline{h}^j$

$\underline{r}^{j+1}=-\nabla E(\underline{\sigma_x}^{j+1})$

$\underline{h}^{j+1}=\underline{r}^{j+1}+\beta^j\underline{h}^j$

$j=j+1$

Until $||\nabla E(\underline{\sigma_x}^j)||<\varepsilon$

In nonlinear CG, the residual is always set to the negative of the gradient $\underline{r}^j=-\nabla E(\sigma_x^j)$. The search directions are computed by Gram-Schmidt[39] conjugation of the residuals as with linear CG. Performing a line search along this search direction is much more difficult than in the linear case.

In linear CG, there are several equivalent expressions for the value of β. In nonlinear CG, these different expressions are no longer equivalent; researchers are still investigating the best choice. Two choices are the Fletcher-Reeves formula, which is used in linear CG for its ease of computation, and the Polak-Ribiere formula:

$$\beta_{FR}^{\,j+1} = \frac{\left(r^{j+1}\right)^T r^{j+1}}{\left(r^j\right)^T r^j}, \quad \beta_{PR}^{\,j+1} = \frac{\left(r^{j+1}\right)^T \left(r^{j+1} - r^j\right)}{\left(r^j\right)^T r^j} \tag{56}$$

The Fletcher-Reeves method converges if the starting point is sufficiently close to the desired minimum. The Polak-Ribiere method can, in rare cases, cycle infinitely without converging. However, Polak-Ribiere often converges much more quickly. Fortunately, convergence of the Polak-Ribiere method can be guaranteed by choosing $\beta=\max(0,\beta_{PR})$. Using this value is equivalent to restarting CG if $\beta_{PR}<0$. To *restart* CG is to forget the past search directions, and start a new CG in the direction of steepest descent. Fig. 30 shows the path followed by the Polack-Ribiere CG to find the minimum of the functions. As can be seen, the directions are not orthogonal with the previous ones.

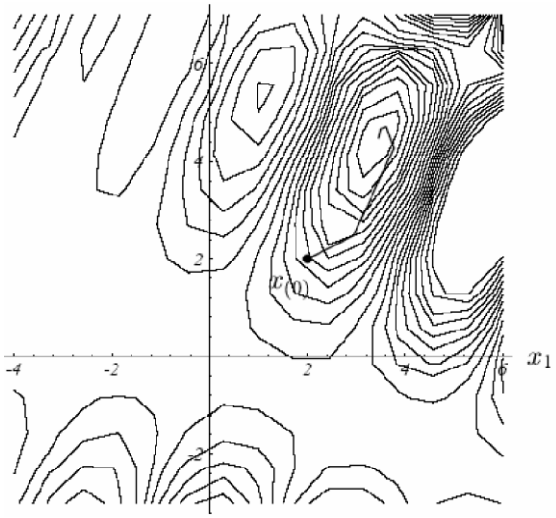


Fig. 30. Convergence path of the Polack Ribiere path for a function with many local minima and maxima.

Application of the Conjugate-gradient for the optimization of an objective function is used in a great variety of situations: engineering design, non-linear regression, and lately neural net training. For the purpose of optical tomography, the optimization of the objective function by CG is very reliable, under the assumption of a possible gradient calculation.

3.3. Gradient Calculation

The gradient of the objective function represents the most challenging part of the entire process. Eq. (53) expresses the objective function in terms of the optical properties of the medium but the dependences are more complicated and directly related to the forward process. In order to understand the effect of each cross section on the objective function, the forward problem has to be analyzed step by step.

In Section 2, it was shown how the detector readings consist of two contributions, one due to particles that pass uncollided through the medium and one due to the particles that undergo

scattering events in the medium before reaching the detectors. The objective function can therefore be expressed in a more precise way as:

$$E = \frac{1}{2}\sum_{i=1}^{m}\left(M_{exp} - P_{collided} - P_{uncollided}\right)^2 \tag{57}$$

This different way to express the objective function underlines another important concept. Both the uncollided and collided processes involve as fundamental a variable the total cross section. It is the total cross section only that determines the probability for a particle to pass through a medium without collisions. The streaming process itself used in the forward model to calculate the fluxes on the surfaces of the cell is affected only by the total cross section. It is therefore more convenient to think of the objective function as dependent on scattering cross section and total cross section, and anisotropy factor.

The best way to investigate the dependence of the function on the above optical properties it is by analyzing the two prediction components separately: uncollided gradient and collided gradient calculation.

3.3.1 Uncollided Gradient Calculation

Uncollided radiation passes through the medium without collision. The forward model calculates this component by performing a numerical integral along the beam window to calculate the radiation arriving at one detector point. Multiple contributes from multiple points into the same detector are summed together with respect to the corresponding weights (see Section 2). If a particular point of a detector is taken into consideration it is clearly understandable that the total cross section of all the cells between beam window and detector point affect the result. In Fig. 31 these cells are marked with “x”.

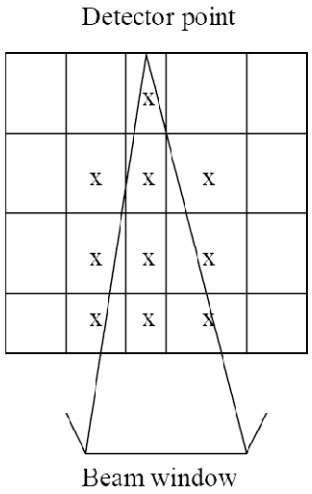


Fig. 31. Detector point and beam window connection. The cells that come into play in the detector reading calculation are indicated with an “x”.

TNTs is able to calculate the effect of each cross section on the uncollided component of the radiation for each point of each detector. From Section 2, the derivative of this component over the total cross section is easy to calculate and the result is:

$$\frac{\delta P_{\text{uncollided}}(\text{detector,point})}{\delta\sigma_{tot}(cell)} = 2w_{\gamma}w_{polar}e^{-\tau}ds(\text{cell,point})\cdot(|\mu| \text{ or } |\eta|) \tag{58}$$

where all the variables have been introduced previously.

The total derivative of the uncollided prediction in one detector as a function of a cell total cross section is then:

$$\frac{\delta P_{unc}(\text{detector})}{\delta\sigma_{tot}(cell)} = \sum_{n=1}^{N} w_{point}\frac{\delta P_{unc}(\text{detector,point})}{\delta\sigma_{tot}(cell)} \tag{59}$$

It is therefore straightforward to calculate the derivative of the objective function over the total cross section of a cell due to the uncollided component. From Eq. (57), applying the chain rule gives:

$$\begin{aligned}\frac{\delta E}{\delta\sigma_{tot}(cell)} &= \frac{\delta E}{\delta P_{unc}(\text{detector})}\frac{\delta P_{unc}(\text{detector})}{\delta\sigma_{tot}(cell)} = \\ &= \sum_{dec=1}^{D}\left\{\left(M_{\exp} - P_{col}(\text{detector}) - P_{unc}(\text{detector})\right)\right. \\ &\left.\sum_{n=1}^{N} w_{point}\left(2w_{\gamma}w_{polar}e^{-\tau}ds\cdot(|\mu| \text{ or } |\eta|)\right)\Big|_{n^{th}\text{ point}}\right\}_{dec^{th}\text{ detector}}\end{aligned} \tag{60}$$

TNTs applies Eq. (60) at all the cells and constructs a matrix of contributions of the single cell cross section to the objective function due to un-collided radiation. The uncollided part is easy to analyze and easy to compute. All the calculations were made along the normal forward process after the convergence of the solution without inserting further steps or iterations.

3.3.2 Collided Gradient Calculation

The analytical expression of the collided component of the gradient is also straightforward. The implementation into a computational code though is less immediate and represents an interesting application of an adjoint formulation. In the collided part the radiation passes through different steps before it reaches the detectors. In Section 2, the general solution concept was outlined: uncollided source for all cells are constructed, then these sources represent the driving force for the step-characteristic source iteration scheme that iteratively constructs the sources per cell of all collided particles. The total cross section and the scattering cross sections appear in all of these steps.

The scattering cross section comes into play only in the construction of the sources per cell since the strength of a scattering-source depends directly on the value of the scattering cross section (see Section 2). The total cross section, instead, affects all the streaming processes including the one used for the creation of the un-collided sources per cell. It is easier to understand this by looking at the grid in Fig. 32.

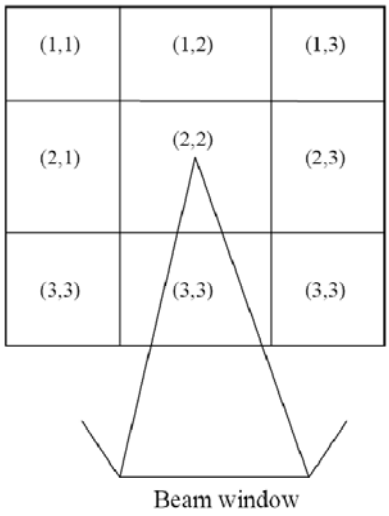


Fig. 32. Spatially discretized medium for forward and gradient calculations.

Assuming that the only changing properties is the scattering cross section in cell (2,2), let's follow the steps of the forward method to see how this change affects the objective function. An important note is that the change of the scattering cross section has to be done keeping the total cross section constant.

- Uncollided source calculation: the source in cell (2,2) changes its strength because the scattering cross section determines linearly the strength of the source. Note that the streaming part of this first process is not affected by the change of the scattering cross section since depends only on the total cross section of the cells that the particle go through. The sources in all the other cells remain unchanged.
- The change in the source at the first iteration changes the fluxes at the edges of the cell itself. It is intuitive to understand that each iteration will transfer the changes of these fluxes throughout the problem, affecting all the fluxes and sources in all the other cells. It is important to note that keeping constant all the total cross sections in the medium leads to un-changed streaming processes. The fluxes therefore will change throughout the iterations only because of the effect of the fluxes in cell (2,2). It is therefore necessary to understand how this "information" travels into the problem between fluxes and sources.
- The fluxes obtained at the end of the iterative scheme together with the uncollided radiation construct the objective function. Again the un-collided radiation is not affected by the change of the scattering cross section in cell (2,2) since it depends only on the total cross sections in the medium.

It is clear, from this first approximate analysis, that the relation between objective function and total cross section is more complicated. Again following the steps of the forward scheme it is possible to understand better how this happens [the changing cross section is the one of cell (2,2)]:

- The streaming process between the beam window and the cell itself is now affected. A change in the total cross section causes the un-collided flux and current in the cell to change. Therefore, the un-collided source is changed in its strength. But in this case not

only the source in cell (2,2) is changed. All the cell in which the un-collided source is constructed by particles streaming through cell (2,2) are affected. In Fig. 32 these cells are the ones of the first row. In general, all the uncollided sources of cells above the cell taken into consideration are affected by the total cross section of this cell.

- At the first iteration then the un-collided source in cell (2,2) creates the edge fluxes. The effect of the cross section is therefore passed to the edges of the cell itself. Different from the scattering case, at the first iteration all the sources affected by the total cross section of cell (2,2) pass the information to their edges. Each iteration will spread the effect of the total cross section from all the affected cells to the entire medium. The streaming process is un-changed in all the cell but the one taken into consideration, cell (2,2). In this cell at each iteration also the streaming term affects the fluxes and the source.
- At the convergence of the forward model both the component of the objective functions bring information about the change in the total cross section of cell (2,2).

The effect of a change in the total cross section is clearly more complicate than the effect due to a change in the scattering property. Let's try to write this in an approximate mathematical way; indicating with *(m,n)* the cells at the edge used as detector reading and *(i,j)* the generic problem cell:

$$\begin{aligned}\frac{\delta E}{\delta \sigma_{sc}(i,j)} &= \left.\frac{\delta E}{\delta \sigma_{sc}(i,j)}\right|_{\text{coll}} + \cancel{\left.\frac{\delta E}{\delta \sigma_{sc}(i,j)}\right|_{\text{un-coll}}} \\ &= \sum_{n=1}^{N}\sum_{m=1}^{M} \frac{\delta E}{\delta J_{\text{coll}}(m,n)} \frac{\delta J_{\text{coll}}(m,n)}{\delta \sigma_{sc}(i,j)} \\ &= \sum_{n=1}^{N}\sum_{m=1}^{M} \frac{\delta E}{\delta J_{\text{coll}}(m,n)} \frac{\delta J_{\text{coll}}(m,n)}{\delta \sigma_{sc}(i,j)}\end{aligned} \tag{61}$$

where J indicates the current calculated at the outer surface of the cell (that is the current that contributes to the detector reading calculation). The same can be written for the total cross section:

$$\begin{aligned}\frac{\delta E}{\delta \sigma_{tot}(i,j)} &= \left.\frac{\delta E}{\delta \sigma_{tot}(i,j)}\right|_{\text{coll}} + \left.\frac{\delta E}{\delta \sigma_{tot}(i,j)}\right|_{\text{un-coll}} \\ &= \sum_{n=1}^{N}\sum_{m=1}^{M} \frac{\delta E}{\delta J_{\text{coll}}(m,n)} \frac{\delta J_{\text{coll}}(m,n)}{\delta \sigma_{tot}(i,j)} + \left.\frac{\delta E}{\delta \sigma_{tot}(i,j)}\right|_{\text{un-coll}} \\ &= \sum_{n=1}^{N}\sum_{m=1}^{M} \frac{\delta E}{\delta J_{\text{coll}}(m,n)} \frac{\delta J_{\text{coll}}(m,n)}{\delta \sigma_{tot}(i,j)} + \left.\frac{\delta E}{\delta \sigma_{tot}(i,j)}\right|_{\text{un-coll}}\end{aligned} \tag{62}$$

The uncollided component was presented in the previous section. It is vital now to express in a better way the dependence of the current in one cell with respect to the cross sections of another cell. This is the core of the adjoint calculation. As seen before, the contributions to the cross sections are different, but in both cases, the information of every single cell is transmitted to the outer boundary of the problem through the edge fluxes. It is possible to conceive of a mammoth chain rule that brings information from a single cell to the outer boundary. It is easy to draw numerous different paths that a particle can follow from a cell to the boundaries traveling across consecutive surfaces and sources. In general these paths are infinite because mathematically there is always a residual radiation that scatters without being absorbed; but this residual will be negligible after a certain number of collisions.

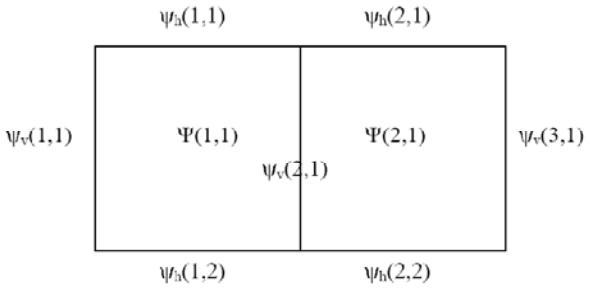


Fig. 33. Simple 2 by 1 cells problem to introduce the gradient adjoint calculation.

The best way to take into consideration the passage of information from a cell to an adjacent cell is by constructing a matrix with the derivative of the flux in one cell over the flux in another cell as its components. A two cell problem is taken into consideration to introduce the matrix and its properties. Fig. 33 illustrates this and in particular defines the symbol of the fluxes at the edges and the average flux per cell. With this configuration it is possible to derive the chain rule to understand how a cross section affects the fluxes, but it is also possible to derive the chain rule to relate a flux with all the adjacent ones. The matrix for the flux derivatives is shown in Fig. 34.

Only fluxes of the same cell are directly related, and the average flux per cell only depends on the edge fluxes of the cell and on itself. With these considerations some terms in the matrix are zero as shown in Fig. 35.

$\frac{\partial \overline{\psi}_{(1,1)}}{\partial \overline{\psi}_{(1,1)}}$	$\frac{\partial \overline{\psi}_{(1,1)}}{\partial \overline{\psi}_{(2,1)}}$	$\frac{\partial \overline{\psi}_{(1,1)}}{\partial \psi^{V}_{(1,1)}}$	$\frac{\partial \overline{\psi}_{(1,1)}}{\partial \psi^{V}_{(2,1)}}$	$\frac{\partial \overline{\psi}_{(1,1)}}{\partial \psi^{V}_{(3,1)}}$	$\frac{\partial \overline{\psi}_{(1,1)}}{\partial \psi^{H}_{(1,1)}}$	$\frac{\partial \overline{\psi}_{(1,1)}}{\partial \psi^{H}_{(1,2)}}$	$\frac{\partial \overline{\psi}_{(1,1)}}{\partial \psi^{H}_{(2,1)}}$	$\frac{\partial \overline{\psi}_{(1,1)}}{\partial \psi^{H}_{(2,2)}}$
$\frac{\partial \overline{\psi}_{(2,1)}}{\partial \overline{\psi}_{(1,1)}}$	$\frac{\partial \overline{\psi}_{(2,1)}}{\partial \overline{\psi}_{(2,1)}}$	$\frac{\partial \overline{\psi}_{(2,1)}}{\partial \psi^{V}_{(1,1)}}$	$\frac{\partial \overline{\psi}_{(2,1)}}{\partial \psi^{V}_{(2,1)}}$	$\frac{\partial \overline{\psi}_{(2,1)}}{\partial \psi^{V}_{(3,1)}}$	$\frac{\partial \overline{\psi}_{(2,1)}}{\partial \psi^{H}_{(1,1)}}$	$\frac{\partial \overline{\psi}_{(2,1)}}{\partial \psi^{H}_{(1,2)}}$	$\frac{\partial \overline{\psi}_{(2,1)}}{\partial \psi^{H}_{(2,1)}}$	$\frac{\partial \overline{\psi}_{(2,1)}}{\partial \psi^{H}_{(2,2)}}$
$\frac{\partial \psi^{V}_{(1,1)}}{\partial \overline{\psi}_{(1,1)}}$	$\frac{\partial \psi^{V}_{(1,1)}}{\partial \overline{\psi}_{(2,1)}}$	$\frac{\partial \psi^{V}_{(1,1)}}{\partial \psi^{V}_{(1,1)}}$	$\frac{\partial \psi^{V}_{(1,1)}}{\partial \psi^{V}_{(2,1)}}$	$\frac{\partial \psi^{V}_{(1,1)}}{\partial \psi^{V}_{(3,1)}}$	$\frac{\partial \psi^{V}_{(1,1)}}{\partial \psi^{H}_{(1,1)}}$	$\frac{\partial \psi^{V}_{(1,1)}}{\partial \psi^{H}_{(1,2)}}$	$\frac{\partial \psi^{V}_{(1,1)}}{\partial \psi^{H}_{(2,1)}}$	$\frac{\partial \psi^{V}_{(1,1)}}{\partial \psi^{H}_{(2,2)}}$
$\frac{\partial \psi^{V}_{(2,1)}}{\partial \overline{\psi}_{(1,1)}}$	$\frac{\partial \psi^{V}_{(2,1)}}{\partial \overline{\psi}_{(2,1)}}$	$\frac{\partial \psi^{V}_{(2,1)}}{\partial \psi^{V}_{(1,1)}}$	$\frac{\partial \psi^{V}_{(2,1)}}{\partial \psi^{V}_{(2,1)}}$	$\frac{\partial \psi^{V}_{(2,1)}}{\partial \psi^{V}_{(3,1)}}$	$\frac{\partial \psi^{V}_{(2,1)}}{\partial \psi^{H}_{(1,1)}}$	$\frac{\partial \psi^{V}_{(2,1)}}{\partial \psi^{H}_{(1,2)}}$	$\frac{\partial \psi^{V}_{(2,1)}}{\partial \psi^{H}_{(2,1)}}$	$\frac{\partial \psi^{V}_{(2,1)}}{\partial \psi^{H}_{(2,2)}}$
$\frac{\partial \psi^{V}_{(3,1)}}{\partial \overline{\psi}_{(1,1)}}$	$\frac{\partial \psi^{V}_{(3,1)}}{\partial \overline{\psi}_{(2,1)}}$	$\frac{\partial \psi^{V}_{(3,1)}}{\partial \psi^{V}_{(1,1)}}$	$\frac{\partial \psi^{V}_{(3,1)}}{\partial \psi^{V}_{(2,1)}}$	$\frac{\partial \psi^{V}_{(3,1)}}{\partial \psi^{V}_{(3,1)}}$	$\frac{\partial \psi^{V}_{(3,1)}}{\partial \psi^{H}_{(1,1)}}$	$\frac{\partial \psi^{V}_{(3,1)}}{\partial \psi^{H}_{(1,2)}}$	$\frac{\partial \psi^{V}_{(3,1)}}{\partial \psi^{H}_{(2,1)}}$	$\frac{\partial \psi^{V}_{(3,1)}}{\partial \psi^{H}_{(2,2)}}$
$\frac{\partial \psi^{H}_{(1,1)}}{\partial \overline{\psi}_{(1,1)}}$	$\frac{\partial \psi^{H}_{(1,1)}}{\partial \overline{\psi}_{(2,1)}}$	$\frac{\partial \psi^{H}_{(1,1)}}{\partial \psi^{V}_{(1,1)}}$	$\frac{\partial \psi^{H}_{(1,1)}}{\partial \psi^{V}_{(2,1)}}$	$\frac{\partial \psi^{H}_{(1,1)}}{\partial \psi^{V}_{(3,1)}}$	$\frac{\partial \psi^{H}_{(1,1)}}{\partial \psi^{H}_{(1,1)}}$	$\frac{\partial \psi^{H}_{(1,1)}}{\partial \psi^{H}_{(1,2)}}$	$\frac{\partial \psi^{H}_{(1,1)}}{\partial \psi^{H}_{(2,1)}}$	$\frac{\partial \psi^{H}_{(1,1)}}{\partial \psi^{H}_{(2,2)}}$
$\frac{\partial \psi^{H}_{(1,2)}}{\partial \overline{\psi}_{(1,1)}}$	$\frac{\partial \psi^{H}_{(1,2)}}{\partial \overline{\psi}_{(2,1)}}$	$\frac{\partial \psi^{H}_{(1,2)}}{\partial \psi^{V}_{(1,1)}}$	$\frac{\partial \psi^{H}_{(1,2)}}{\partial \psi^{V}_{(2,1)}}$	$\frac{\partial \psi^{H}_{(1,2)}}{\partial \psi^{V}_{(3,1)}}$	$\frac{\partial \psi^{H}_{(1,2)}}{\partial \psi^{H}_{(1,1)}}$	$\frac{\partial \psi^{H}_{(1,2)}}{\partial \psi^{H}_{(1,2)}}$	$\frac{\partial \psi^{H}_{(1,2)}}{\partial \psi^{H}_{(2,1)}}$	$\frac{\partial \psi^{H}_{(1,2)}}{\partial \psi^{H}_{(2,2)}}$
$\frac{\partial \psi^{H}_{(2,1)}}{\partial \overline{\psi}_{(1,1)}}$	$\frac{\partial \psi^{H}_{(2,1)}}{\partial \overline{\psi}_{(2,1)}}$	$\frac{\partial \psi^{H}_{(2,1)}}{\partial \psi^{V}_{(1,1)}}$	$\frac{\partial \psi^{H}_{(2,1)}}{\partial \psi^{V}_{(2,1)}}$	$\frac{\partial \psi^{H}_{(2,1)}}{\partial \psi^{V}_{(3,1)}}$	$\frac{\partial \psi^{H}_{(2,1)}}{\partial \psi^{H}_{(1,1)}}$	$\frac{\partial \psi^{H}_{(2,1)}}{\partial \psi^{H}_{(1,2)}}$	$\frac{\partial \psi^{H}_{(2,1)}}{\partial \psi^{H}_{(2,1)}}$	$\frac{\partial \psi^{H}_{(2,1)}}{\partial \psi^{H}_{(2,2)}}$
$\frac{\partial \psi^{H}_{(2,2)}}{\partial \overline{\psi}_{(1,1)}}$	$\frac{\partial \psi^{H}_{(2,2)}}{\partial \overline{\psi}_{(2,1)}}$	$\frac{\partial \psi^{H}_{(2,2)}}{\partial \psi^{V}_{(1,1)}}$	$\frac{\partial \psi^{H}_{(2,2)}}{\partial \psi^{V}_{(2,1)}}$	$\frac{\partial \psi^{H}_{(2,2)}}{\partial \psi^{V}_{(3,1)}}$	$\frac{\partial \psi^{H}_{(2,2)}}{\partial \psi^{H}_{(1,1)}}$	$\frac{\partial \psi^{H}_{(2,2)}}{\partial \psi^{H}_{(1,2)}}$	$\frac{\partial \psi^{H}_{(2,2)}}{\partial \psi^{H}_{(2,1)}}$	$\frac{\partial \psi^{H}_{(2,2)}}{\partial \psi^{H}_{(2,2)}}$

Fig. 34. Matrix of flux over flux derivatives.

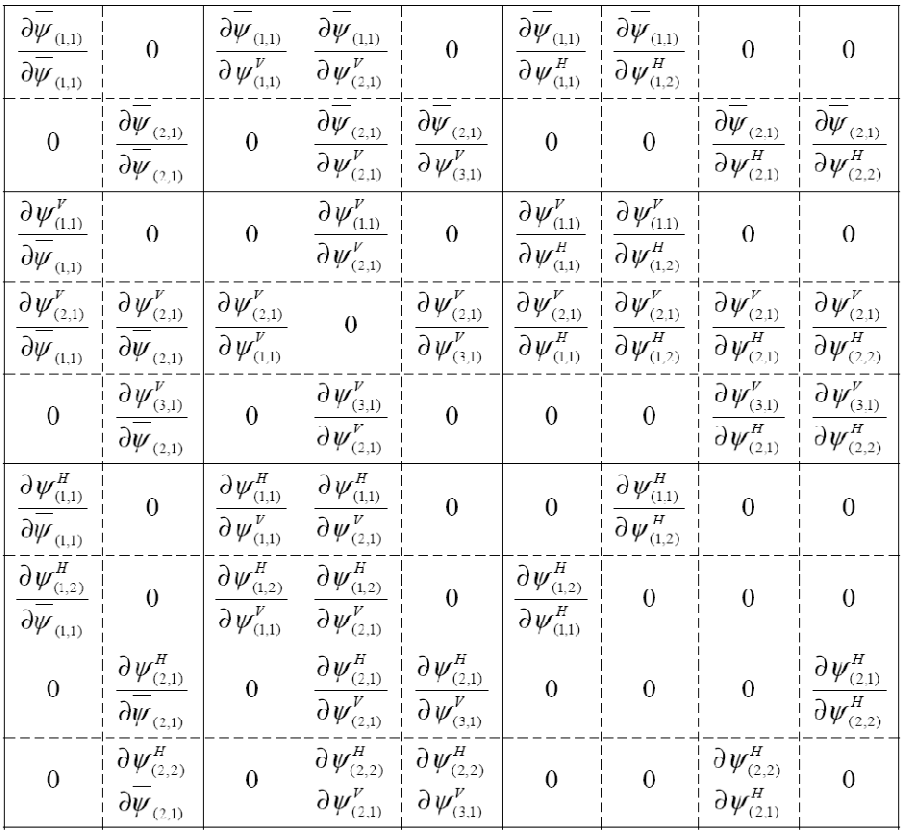


Fig. 35. Simplified matrix with dependencies between edge and average fluxes.

Another important consideration that has not been introduced previously is about the direction *k* in which all these relations are derived. Every element of the matrix in Fig. 35 is in reality a *K* by *K* matrix where *K* represents the number of directions that the S_n angular discretization scheme is using. Every derivative is therefore taken as follow:

$$\frac{\delta\psi^{k'}(i,j)}{\delta\psi^{k''}(m,n)} \tag{63}$$

where not only the cells of reference are changing but also the directions of each flux. This leads to a more complicated expression for the final form of the matrix but it is still straightforward to derive.

The matrix, that by itself simply shows the dependence between all the fluxes with respect to all the other fluxes of their cell, is a very powerful tool in the gradient calculation. It will be shown in detail later how the multiplication of the matrix by itself *n* times leads to an interesting result. It will represent how every flux changes as a function of all the fluxes that are *n*-steps away from it. A step is consider as a forward iteration or better it can be seen as a path that "connects" adjacent surface fluxes or surface fluxes with average fluxes. Every multiplication of the matrix by itself *n* times will gather all the *n*-step paths from every investigated flux and therefore, multiplying the matrix infinite times corresponds to finding all the infinite paths that a particle can undergo to reach a point.

The matrix introduced before will be referred from now on as matrix A with all the characteristics and all the hidden features explained previously. It is possible now to proceed with the analysis of the gradient expressed in Eqs. (61) and (62). We will introduce the vector v which contains the derivative of the objective function with respect to all the surface and average fluxes of the problem. It is easy to note that this derivative will be non-zero only when it is with respect to fluxes of surfaces that correspond to detector position and to directions that contribute to the detector responses:

$$\frac{\delta E}{\delta \psi^k(p,q)} = \begin{cases} 0 & \text{if surface (p,q) is not detector position} \\ & \text{or k is not outer direction} \\ \text{value} & \text{otherwise} \end{cases} \tag{64}$$

Multiplying v and A together n times gives a vector of derivatives of the objective function over fluxes n steps away from the detectors position. The tools presented so far are general and they don't depend on anything other than the geometry of the problem and the properties of the medium. The properties, in particular, have been considered constants in the derivation of the matrix and the vector. All the derivatives have been taken with respect to fluxes only and not cross sections. To close the chain rule, we need to insert the derivative of every flux with respect to the cross sections of each cell. As seen before this will be easier for the scattering dependence since the fluxes of a cell depend only on the scattering property of their cell; it will be more complicate for the total cross section since the average flux of every cell depends, in general, on the total cross section of all the cells in the problem. A new matrix to contain all these derivatives must therefore be constructed. It will be called matrix B and will have many zeros in the scattering case and few in the total cross section case. The multiplication of matrix B with A^n and vector v will represent therefore the derivative of the objective function over the cross section of a cell due to particles that from the cell reach the detectors position in n steps.

Finally, the gradient of the objective function over the scattering and total cross section can be represented as:

$$\frac{\delta E}{\delta \sigma_{sca}} = \sum_n B_{sca} A^n v \tag{65}$$

$$\frac{\delta E}{\delta \sigma_{tot}} = \sum_n B_{tot} A^n v \tag{66}$$

The maximum value of n is the one that makes the product converge. Thinking in term of steps, the maximum n will be exactly equal to the number of iterations of the forward problem.

Let's see in more detail for the case introduced in Fig. 33 what $A^n v$ will give [considering only one detector on the vertical surface (1,3)]:

$$A^0 v = \frac{\delta E}{\delta \psi_V(1,3)} \tag{67}$$

$$A^1 v = \frac{\delta E}{\delta \psi_V(1,3)} \frac{\delta \psi_V(1,3)}{\delta \psi_V(1,2)} + \frac{\delta E}{\delta \psi_V(1,3)} \frac{\delta \psi_V(1,3)}{\delta \psi_{II}(2,1)} + \\ + \frac{\delta E}{\delta \psi_V(1,3)} \frac{\delta \psi_V(1,3)}{\delta \psi_H(2,2)} + \frac{\delta E}{\delta \psi_V(1,3)} \frac{\delta \psi_V(1,3)}{\delta \psi(2,1)} \tag{68}$$

and the same is true for n>2.

Eq.(67), derived for n=0, is the derivative of the objective function due to fluxes *0 steps* away from the detector, Eq.(68) instead is the derivative with respect to fluxes that are *1 step* away from the detector. This has been derived by simply multiplying the matrix and the vector and eliminating all the terms that are zero because of angular considerations. For example:

$$\frac{\delta \psi_V(1,2)}{\delta \psi(2,1)} \frac{\delta \psi_V(1,3)}{\delta \psi_V(1,2)} = 0 \tag{69}$$

since the directions in which $\psi(2,1)$ affects $\psi(1,2)$ are different from the directions in which $\psi(1,2)$ affects $\psi(1,3)$. Clearly in Eq. (67) and (68) the chain rule it's in term of steps from the detector position (where the objective function is calculated) and it can be derived by inspection without too much effort, at least for small n where all the paths can be easily listed. The complicated calculation of the gradient is therefore reduced to a matrixes multiplication and summation. The main task is therefore the construction of these matrixes.

It's interesting to note that the size of the domain doesn't affect the construction of the matrixes. The chain rule, obviously longer for finer meshes problems, is carried out correctly by the multiplication of A with v as many times as necessary to consider all the particle paths that contribute to the objective function. A single cell problem will be presented to derive the components of the matrixes.

3.3.3 Matrixes Components

To simplify the derivation the presented one-cell problem will consider an incident flux in direction k incident on bottom and left surfaces, as in Fig. 36.

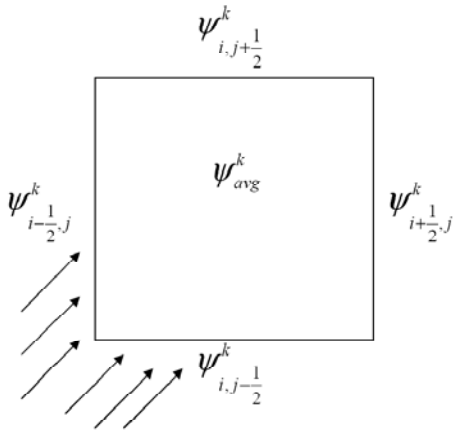


Fig. 36. One-cell problem for the derivation of the matrixes components.

To write the equation for a single cell problem we consider a cell-by-cell discretization of the discrete ordinate transport equation. This can be called *closed linear one-cell functional* method. Given such a method from Section 2, and the problem defined in Fig. 36, the following equation for all the exiting angular fluxes from one cell can be written:

$$\begin{bmatrix} \psi^K_{i+\frac{1}{2},j} \\ \psi^K_{i,j+\frac{1}{2}} \end{bmatrix} = \begin{bmatrix} \frac{v(1-e^{-\tau})}{\tau} & (1-v)e^{-\tau} \\ (1-a)e^{-\tau} & \frac{a(1-e^{-\tau})}{\tau} \end{bmatrix} \begin{bmatrix} \psi^k_{i,j-\frac{1}{2}} \\ \psi^k_{i-\frac{1}{2},j} \end{bmatrix} + \begin{bmatrix} 1-\frac{v(1-e^{-\tau})}{\tau}-(1-v)e^{-\tau} \\ 1-(1-a)e^{-\tau}-\frac{a(1-e^{-\tau})}{\tau} \end{bmatrix} \frac{1}{4\pi\sigma_{tot}} Q \tag{70}$$

or

$$\underline{\psi}^k_{s,out} = A^k_{s\leftarrow s}\underline{\psi}^k_{s,in} + A^k_{s\leftarrow v}\underline{Q} \tag{71}$$

where we have defined:

- $\underline{\psi}^k_{s,out}$ = vector of outgoing angular fluxes from the cell (all outgoing surfaces for direction k)
- $\underline{\psi}^k_{s,out}$ = vector of incoming angular fluxes to the cell (incoming surfaces for direction k)
- $\underline{Q}$ = vector of total source-rate densities in the cell (isotropic)
- $A^k_{s\leftarrow s}$ = transmission matrix for the cell for direction k (each element depends on cell size and total cross section)
- $A^k_{s\leftarrow v}$ = volume-to-surface matrix for the cell for direction k (each element depends on cell size and total cross section)

We have assumed isotropic scattering for simplicity, but this can be readily extended to anisotropic scattering. Eq. (70) will be used for the derivative calculation. Eq. (71) has been introduced to be used later.

A similar equation for the cell-averaged flux can be written:

$$\psi^k_{AVG} = \begin{bmatrix} \frac{\mu}{\Delta x\sigma_{tot}}\frac{v(1-e^{-\tau})}{\tau} - \frac{\eta}{\Delta y\sigma_{tot}}\left((1-a)e^{-\tau}-1\right) & \frac{\mu}{\Delta x\sigma_{tot}}\left((1-v)e^{-\tau}-1\right) - \frac{\eta}{\Delta y\sigma_{tot}}\frac{a(1-e^{-\tau})}{\tau} \end{bmatrix} \begin{bmatrix} \psi^k_{i,j-\frac{1}{2}} \\ \psi^k_{i-\frac{1}{2},j} \end{bmatrix} +$$
$$+ \left[1-\frac{\mu}{\Delta x\sigma_{tot}}\left(1-\frac{v(1-e^{-\tau})}{\tau}-(1-v)e^{-\tau}\right)-\frac{\eta}{\Delta y\sigma_{tot}}\left(1-(1-a)e^{-\tau}-\frac{a(1-e^{-\tau})}{\tau}\right)\right]\frac{1}{4\pi\sigma_{tot}}Q \tag{72}$$

or

$$\underline{\psi}^k_{AVG} = B^k_{v\leftarrow s}\underline{\psi}^k_{IN} + B^k_{v\leftarrow v}\underline{Q} \tag{73}$$

where:

- $\underline{\psi}^k_{s,out}$ = vector of cell-interior angular fluxes (direction k)
- $B^k_{s\leftarrow s}$ = surface-to-volume matrix for the cell for direction k
- $B^k_{s\leftarrow v}$ = volume-to-volume matrix for the cell for direction k

Integrating Eq.(73) gives an equation for the cell-interior scalar flux:

$$\begin{aligned}\phi &= \sum_k w_k \underline{\psi}^k_{avg} = \\ &= \sum_k w_k B^k_{v\leftarrow s} \underline{\psi}^k_{s,in} + \underline{Q}\sum_k w_k B^k_{v\leftarrow v} = \\ &= \sum_k w_k B^k_{v\leftarrow s} \underline{\psi}^k_{s,in} + D\underline{Q}\end{aligned} \tag{74}$$

where $\underline{\phi}$ is the vector of cell-interior scalar fluxes and D is volume-to-volume matrix for the cell (direction integral of $B^k_{s\leftarrow v}$). Note, from Section 2, that $\underline{Q}$ depends on $\underline{\phi}$:

$$\underline{Q} = \sigma_s \underline{\phi} + \underline{S} \tag{75}$$

From Eq. (22), it is possible to derive:

$$\begin{bmatrix} \dfrac{d\psi^k_{i-\frac{1}{2}}}{d\sigma_s} \\ \dfrac{d\psi^k_{i,j+\frac{1}{2}}}{d\sigma_s} \end{bmatrix} = \begin{bmatrix} 1-\dfrac{v(1-e^{-\tau})}{\tau}-(1-v)e^{-\tau} \\ 1-(1-a)e^{-\tau}-\dfrac{a(1-e^{-\tau})}{\tau} \end{bmatrix} \frac{1}{4\pi\sigma_t}\frac{\delta\underline{Q}}{\delta\sigma_s} \tag{76}$$

$$\frac{\delta\underline{\psi}^k_{s,out}}{\delta\sigma_s} = A^k_{s\leftarrow v}\frac{\delta\underline{Q}}{\delta\sigma_s} \tag{77}$$

$$\begin{aligned}\begin{bmatrix} \dfrac{d\psi^k_{i-\frac{1}{2}}}{d\sigma_t} \\ \dfrac{d\psi^k_{i,j+\frac{1}{2}}}{d\sigma_t} \end{bmatrix} &= \begin{bmatrix} \dfrac{ve^{\tau}}{\sigma_t}-\dfrac{v(1-e^{-\tau})}{\tau c} & -c(1-v)e^{-\tau} \\ -c(1-a)e^{-\tau} & \dfrac{ae^{\tau}}{\sigma_t}-\dfrac{a(1-e^{-\tau})}{\tau c} \end{bmatrix} \begin{bmatrix} \psi^k_{i,j-\frac{1}{2}} \\ \psi^k_{i-\frac{1}{2},j} \end{bmatrix} + \\ &+ \begin{bmatrix} \dfrac{-1}{\sigma_t}+\dfrac{2v(1-e^{-\tau})}{\sigma_t\tau}-\dfrac{ve^{-\tau}}{\sigma_t}+\dfrac{(1-v)e^{-\tau}}{\sigma_t}+(1-v)ce^{-\tau} \\ \dfrac{-1}{\sigma_t}+\dfrac{2a(1-e^{-\tau})}{\sigma_t\tau}-\dfrac{ae^{-\tau}}{\sigma_t}+\dfrac{(1-a)e^{-\tau}}{\sigma_t}+(1-a)ce^{-\tau} \end{bmatrix} \frac{1}{4\pi\sigma_t}Q + \\ &+ \begin{bmatrix} 1-\dfrac{v(1-e^{-\tau})}{\tau}-(1-v)e^{-\tau} \\ 1-(1-a)e^{-\tau}-\dfrac{a(1-e^{-\tau})}{\tau} \end{bmatrix} \frac{1}{4\pi\sigma_t}\frac{\delta\underline{Q}}{\delta\sigma_t}\end{aligned} \tag{78}$$

or

$$\frac{\delta\underline{\psi}^k_{s,out}}{\delta\sigma_t} = \frac{\delta A^k_{s\leftarrow s}}{\delta\sigma_t}\psi^k_{s,in} + \frac{\delta A^k_{s\leftarrow v}}{\delta\sigma_t}Q + A^k_{s\leftarrow v}\frac{\delta\underline{Q}}{\delta\sigma_t} \tag{79}$$

where all the quantities and variables are evaluated in cell *(i,j)*.

In the case of a multi-cell problem, the derivatives are calculated as shown in Eqs. (77) and (79) paying attention to the dependence of the uncollided sources of a cell on the total cross section of another cell:

$$\frac{\delta \underline{\psi}^{k}_{s,out}(i,j)}{\delta \sigma_s(m,n)} = A^{k}_{s\leftarrow v}(i,j)\frac{\delta Q(i,j)}{\delta \sigma_s(m,n)}\delta_{i,m}\delta_{j,n} \tag{80}$$

$$\frac{\delta \underline{\psi}^{k}_{s,out}(i,j)}{\delta \sigma_t(m,n)} = \left(\frac{\delta A^{k}_{s\leftarrow s}(i,j)}{\delta \sigma_t(m,n)}\underline{\psi}^{k}_{s,in}(i,j) + \frac{\delta A^{k}_{s\leftarrow v}(i,j)}{\delta \sigma_t(m,n)}\underline{Q}(i,j)\right)\delta_{i,m}\delta_{j,n} + A^{k}_{s\leftarrow v}(i,j)\frac{\delta \overline{Q}(i,j)}{\delta \sigma_t(m,n)} \tag{81}$$

Equations (80) and (81) consider all the dependencies of a flux on the cross sections of other cells. They constitute the building equations for the construction of matrixes *Bsca* and *Btot*.

Using the same constitutive equations [Eqs. (70) and (71)] it is possible to derive the components for the construction of matrix *A*. First, we consider the derivative of the surface fluxes over the cell-averaged flux:

$$\frac{\delta \psi^{k}_{i=\frac{1}{2},j\pm\frac{1}{2}}}{\delta \overline{\psi}^{k'}_{i,j}} = \frac{1}{4\pi\sigma_{tot}}\frac{\delta Q^{k}_{i,j}}{\delta \overline{\psi}^{k'}_{i,j}}\left(1 - v\left(\frac{1-e^{-\tau}}{\tau}\right) - (1-v)e^{-\tau}\right) \tag{82}$$

or

$$\frac{\delta \psi^{k}_{i\pm\frac{1}{2},j\pm\frac{1}{2}}}{\delta \overline{\psi}^{k'}_{i,j}} = \frac{\sigma_{sca}(i,j)}{\sigma_{tot}(i,j)}\frac{w(k')}{4\pi}A^{k}_{s\leftarrow v}(i,j) \tag{83}$$

Then we consider the derivative of the surface fluxes over the other surface fluxes of its cell:

$$\frac{\delta \psi_{i\pm\frac{1}{2},j,k}}{\delta \psi_{i\mp\frac{1}{2},j,k'}} = (1-v)e^{-\tau}\delta_{k,k'} \tag{84}$$

$$\frac{\delta \psi_{i=\frac{1}{2},j,k}}{\delta \psi_{i,j\mp\frac{1}{2},k'}} = v\left(\frac{1-e^{-\tau}}{\tau}\right)\delta_{k,k'} \tag{85}$$

and

$$\frac{\delta \psi_{k,i,j\pm\frac{1}{2}}}{\delta \psi_{k',i+\frac{1}{2},j}} = \alpha\left(\frac{1-e^{\tau}}{\tau}\right)\delta_{k,k'} \tag{86}$$

$$\frac{\delta \psi_{k,i,j\pm\frac{1}{2}}}{\delta \psi_{k',i,j\pm\frac{1}{2}}} = (1-\alpha)e^{\tau}\delta_{k,k'} \tag{87}$$

Now we consider the cell-averaged derivatives. These are acquired from Eq. (72). The derivative of the cell-averaged flux over the surface fluxes is:

$$\frac{\delta\psi^{k}_{avg}}{\delta\psi^{k'}_{i,j-\frac{1}{2}}}=\left(\frac{\mu}{\Delta x\sigma_{tot}}\frac{v(1-e^{-\tau})}{\tau}-\frac{\eta}{\Delta y\sigma_{tot}}\left((1-a)e^{-\tau}-1\right)\right)\Bigg|_{k'} \tag{88}$$

$$\frac{\delta\psi^{k}_{avg}}{\delta\psi^{k'}_{i-\frac{1}{2},j}}=\left(\frac{\mu}{\Delta x\sigma_{tot}}\left((1-v)e^{-\tau}-1\right)-\frac{\eta}{\Delta y\sigma_{tot}}\frac{a(1-e^{-\tau})}{\tau}\right)\Bigg|_{k'} \tag{89}$$

or

$$\frac{\delta\psi^{k}_{avg}}{\delta\psi^{k'}_{s,in}}=B^{k'}_{v\leftarrow s} \tag{90}$$

The derivative of the cell-averaged flux over itself is:

$$\frac{\delta\psi^{k}_{avg}}{\delta\psi^{k'}_{avg}}=\left(\begin{array}{l}1-\frac{\mu}{\Delta x\sigma_{tot}}\left(1-\frac{v(1-e^{-\tau})}{\tau}-(1-v)e^{-\tau}\right)\\-\frac{\eta}{\Delta y\sigma_{tot}}\left(1-(1-a)e^{-\tau}-\frac{a(1-e^{-\tau})}{\tau}\right)\end{array}\right)\Bigg|_{i,j,k'}\frac{\sigma_{sca}(i,j)}{\sigma_{tot}(i,j)}\frac{w(k')}{4\pi} \tag{91}$$

or

$$\frac{\delta\psi^{k}_{avg}}{\delta\psi^{k'}_{avg}}=B^{k'}_{v\leftarrow v}\frac{\sigma_{sca}(i,j)}{\sigma_{tot}(i,j)}\frac{w(k')}{4\pi} \tag{92}$$

Equations (88), (89) and (91) are the components of the *A* matrix. Angular consideration has to be done in the construction of every block in matrix *A*.

To complete the calculation of the tools for the adjoint calculation only the vector *v* need to be explicitly expressed:

$$\frac{\delta E}{\delta\psi^{k}_{i\pm\frac{1}{2},j}}=\begin{cases}\Delta y(i,j)\,w_k\,|\mu_k|\left(P(i,j)-M(i,j)\right) & \text{if (i,j) = detector position}\\ 0 & \text{if k is outer direction}\end{cases} \tag{93}$$

$$\frac{\delta F}{\delta\psi^{k}_{i,j\pm\frac{1}{2}}}=\begin{cases}\Delta x(i,j)\,w_k\,|\eta_k|\left(P(i,j)-M(i,j)\right) & \text{if (i,j) = detector position}\\ 0 & \text{if k is outer direction}\end{cases} \tag{94}$$

$$\frac{\delta E}{\delta\psi^{k}_{avg}}=0 \tag{95}$$

Again it should be noted that the derivative of the objective function over the angular flux of a cell is different from zero only if the surface is a detector surface and if the direction is the outgoing direction. Thus, all the components of the matrixes and the vector have been mathematically derived. It is necessary to verify the correctness of these relations.

3.3.4 Mathematical Verification

It is possible, for the simple case of the one-cell problem, to derive the gradient of the objective function both analytically and with the adjoint method. If the adjoint model is correct the result should be the same or at least approximately the same for a large n where n is the number of time the matrix A is multiplied by itself.

Consider a one-cell problem, the one of Fig. 36, with known incident fluxes and known fixed source $\underline{S}$, where $\underline{S}$ is a function of the total cross section of the cell. The goal is to know how the total and scattering cross sections in the cell affect the outgoing fluxes. That is:

$$\frac{\partial}{\partial \sigma_t}\underline{\psi}_{s,out} \quad \text{and} \quad \frac{\partial}{\partial \sigma_s}\underline{\psi}_{s,out} \tag{96}$$

From Eq. (71) we acquire:

$$\begin{aligned}
\frac{\partial}{\partial \sigma_s}\underline{\psi}^k_{s,out} &= \cancel{\left[\frac{\partial}{\partial \sigma_s}A^k_{s\leftarrow s}\right]}\underline{\psi}^k_{s,in} + \frac{\partial}{\partial \sigma_s}\left[A^k_{s\leftarrow v}\underline{Q}\right] \\
&= \cancel{\left[\frac{\partial}{\partial \sigma_s}A^k_{s\leftarrow v}\right]}\underline{Q} + A^k_{s\leftarrow v}\left[\frac{\partial}{\partial \sigma_s}\underline{Q}\right] \\
&= A^k_{s\leftarrow v}\frac{\partial}{\partial \sigma_s}\left[\sigma_s\underline{\phi} + \underline{S}\right] = A^k_{s\leftarrow v}\frac{\partial}{\partial \sigma_s}\left[\sigma_s\underline{\phi}\right] + \cancel{A^k_{s\leftarrow v}\frac{\partial}{\partial \sigma_s}\left[\underline{S}\right]} \\
&= A^k_{s\leftarrow v}\underline{\phi} + \sigma_s A^k_{s\leftarrow v}\frac{\partial}{\partial \sigma_s}\left[\underline{\phi}\right]
\end{aligned} \tag{97}$$

and

$$\begin{aligned}
\frac{\partial}{\partial \sigma_t}\psi^k_{s,out} &= \left[\frac{\partial}{\partial \sigma_t}A^k_{s\leftarrow s}\right]\psi^k_{s,in} + \frac{\partial}{\partial \sigma_t}\left[A^k_{s\leftarrow v}\underline{Q}\right] \\
&= \left[\frac{\partial}{\partial \sigma_t}A^k_{s\leftarrow s}\right]\underline{\psi}^k_{s,in} + \left[\frac{\partial}{\partial \sigma_t}A^k_{s\leftarrow v}\right]\underline{Q} + A^k_{s\leftarrow v}\left[\frac{\partial}{\partial \sigma_t}\underline{Q}\right] \\
&= \left[\frac{\partial}{\partial \sigma_t}A^k_{s\leftarrow s}\right]\underline{\psi}^k_{s,in} + \left[\frac{\partial}{\partial \sigma_t}A^k_{s\leftarrow v}\right]\underline{Q} + \sigma_s A^k_{s\leftarrow v}\left[\frac{\partial}{\partial \sigma_t}\underline{\phi}\right]
\end{aligned} \tag{98}$$

Since this is a one-cell problem, we have neglected the cell indexes. We will now solve for the scalar flux. Equations (74) and (75) give:

$$\left[I - D\sigma_s\right]\underline{\phi} = \sum_k w_k B^k_{v\leftarrow s}\underline{\psi}^k_{s,in} + D\underline{S} \tag{99}$$

which implies that

$$\underline{\phi} = \left[I - D\sigma_s\right]^{-1}\left[\sum_k w_k B^k_{v\leftarrow s}\underline{\psi}^k_{s,in} + D\underline{S}\right] \tag{100}$$

Note the following:

- D is the sum of the volume-to-volume B matrices, which contain the total cross section and cell dimensions but not the scattering cross section.
- The surface-to-volume B matrices also do not contain the scattering cross section, but they do contain the total cross section and cell dimensions.

An important consideration for the computational application can be shown, before to directly calculate the derivative of the scalar flux in Eq. (100) with respect to the cross sections of the problem. The source iteration applied to the one-cell problem would generate the following solution:

$$\underline{\phi}^{(1)} = \sum_k w_k B^k_{v\leftarrow s} \underline{\psi}^k_{s,in} + D\left[\sigma_s \underline{\phi}^{(0)} + \underline{S}\right] \tag{101}$$

$$\begin{aligned}\underline{\phi}^{(2)} &= \sum_k w_k B^k_{v\leftarrow s} \underline{\psi}^k_{s,in} + D\left[\sigma_s\left(\sum_k w_k B^k_{v\leftarrow s} \underline{\psi}^k_{s,in} + D\left[\sigma_s \underline{\phi}^{(0)} + \underline{S}\right]\right) + \underline{S}\right] \\ &= \left[I + \sigma_s D\right]\left[\sum_k w_k B^k_{v\leftarrow s} \underline{\psi}^k_{s,in} + D\underline{S}\right] + \left[\sigma_s D\right]^2 \underline{\phi}^{(0)}\end{aligned} \tag{102}$$

Continuing this, it can be found that

$$\phi^{(l)} = \left[I + \sigma_s D + (\sigma_s D)^2 + \ldots + (\sigma_s D)^{l-1}\right]\left\{\sum_k w_k B^k_{v\leftarrow s} \psi^k_{s,in} + D\underline{S}\right\} + \left[\sigma_s D\right]^l \phi^{(0)} \tag{103}$$

and

$$\phi^{(converged)} = \left[\sum_{l=0}^{\infty} (\sigma_s D)^l\right]\left\{\sum_k w_k B^k_{v\leftarrow s} \psi^k_{s,in} + D\underline{S}\right\} + \left[\lim_{l\to\infty}\left[\sigma_s D\right]^l\right]\phi^{(0)} \tag{104}$$

If $\sigma_s D$ has spectral radius less than unity, then:

$$\left[\sum_{l=0}^{\infty} (\sigma_s D)^l\right] = \left[I - \sigma_s D\right]^{-1} \tag{105}$$

and

$$\left[\lim_{l\to\infty}\left[\sigma_s D\right]^l\right]\underline{\phi}^{(0)} = \underline{0} \tag{106}$$

Note also that $\sigma_s D$ is the iteration matrix in this problem. As long as $\sigma_s < \sigma_t$, $\sigma_s D$ will have spectral radius less than unity. Therefore it can be concluded that Eqs. (105) and (106) hold for the problem, and that, as a result, Eqs. (100) and (104) are equivalent.

Plugging Eq. (104) in Eq. (71) to obtain an expression for the exiting angular flux leads to:

$$\psi^k_{s,out} = A^k_{s\leftarrow s}\psi^k_{s,in} + A^k_{s\leftarrow v}\left[\sigma_s\left|I - D\sigma_s\right|^{-1}\left(\sum_{k'} w_k B^{k'}_{v\leftarrow s}\psi^{k'}_{s,in} + D\underline{S}\right) + \underline{S}\right] \tag{107}$$

This important result is shown to note how the final angular flux is totally dependent on the converged solution for the scalar flux. This leads to an important computational application: the gradient of the objective function can be calculated at the end of the forward process making use of the converged solution. This fact, not at all intuitive, shows that the iteration scheme during its process of solving the problem generates un-converged fluxes and currents that don't add information to the gradient evaluation. From the computational point of view the gradient calculation can be performed at the last iteration of the forward model. The result of this investigation doesn't agree with the methodology used in other computed tomography papers[35], where the gradient is calculated in a different way along the forward iterative process.

Starting with Eq. (100):

$$\frac{\partial}{\partial \sigma_s}\underline{\phi} = \left(\frac{\partial}{\partial \sigma_s}\left[I - D\sigma_s\right]^{-1}\right)\left[\sum_k w_k B^k_{v\leftarrow s}\underline{\psi}^k_{s,in} + D\underline{S}\right] \tag{108}$$

It is necessary to take the derivative of the inverse matrix in this expression. If D were just a constant (which it is for methods like SC that use only one scalar-flux unknown per cell), then it would be

$$\frac{\partial}{\partial \sigma_s}\left(\left[I - D\sigma_s\right]^{-1}\right) = D\left[I - D\sigma_s\right]^{-2} \tag{109}$$

It can be shown that it is also correct when D is a matrix. Using the previous result Eq. (105):

$$\begin{aligned} D\left[I - D\sigma_s\right]^{-2} &= D\left[I + D\sigma_s + \left(D\sigma_s\right)^2 + \ldots\right]\left[I + D\sigma_s + \left(D\sigma_s\right)^2 + \ldots\right] \\ &= D\left[I + 2D\sigma_s + 3\left(D\sigma_s\right)^2 + 4\left(D\sigma_s\right)^3 + \ldots\right] \end{aligned} \tag{110}$$

This agrees with differentiating the other form for the inverse:

$$\begin{aligned} \frac{\partial}{\partial \sigma_s}\sum_{l=0}^{\infty}\left(D\sigma_s\right)^l &= \sum_{l=0}^{\infty} lD\left(D\sigma_s\right)^{l-1} = D\sum_{l=1}^{\infty} l\left(D\sigma_s\right)^{l-1} \\ &= D\left[I + 2D\sigma_s + 3\left(D\sigma_s\right)^2 + 4\left(D\sigma_s\right)^3 + \ldots\right] \end{aligned} \tag{111}$$

Therefore, plugging Eq. (111) into Eq. (108) gives the important result:

$$\begin{aligned} \frac{\partial}{\partial \sigma_s}\underline{\phi} &= \left(\left[I - D\sigma_s\right]^{-2} D\right)\left[I - D\sigma_s\right]\underline{\phi} = \left[I - D\sigma_s\right]^{-2}\left[I - D\sigma_s\right]D\underline{\phi} \\ &= \left[I - D\sigma_s\right]^{-1} D\underline{\phi} \end{aligned} \tag{112}$$

The same exists for the derivative of the scalar flux with respect to the total cross section:

$$\begin{aligned} \frac{\partial}{\partial \sigma_t}\underline{\phi} &= \sum_k w_k\left(\frac{\partial}{\partial \sigma_t}B^k_{v\leftarrow s}\right)\underline{\psi}^k_{s,in} + \left(\frac{\partial}{\partial \sigma_t}D\right)\underline{Q} + D\frac{\partial}{\partial \sigma_t}\left(\sigma_s\underline{\phi} + \underline{S}\right) \\ &= \sum_k w_k\left(\frac{\partial}{\partial \sigma_t}B^k_{v\leftarrow s}\right)\underline{\psi}^k_{s,in} + D'\underline{Q} + \sigma_s D\frac{\partial}{\partial \sigma_t}\underline{\phi} \end{aligned} \tag{113}$$

which implies that:

$$\begin{aligned}\frac{\partial}{\partial\sigma_t}\underline{\phi} &= \left[I-\sigma_s D\right]^{-1}\left\{\sum_k w_k\left(\frac{\partial}{\partial\sigma_t}B^k_{v\leftarrow s}\right)\underline{\psi}^k_{s,in}+D'\underline{Q}\right\}\\ &= \left[I+\sigma_s D+(\sigma_s D)^2+(\sigma_s D)^3+\ldots\right]\left\{\sum_k w_k\left(\frac{\partial}{\partial\sigma_t}B^k_{v\leftarrow s}\right)\underline{\psi}^k_{s,in}+D'\underline{Q}\right\}\end{aligned} \tag{114}$$

Substituting Eqs. (112) and (114) respectively into Eqs. (97) and (98):

$$\begin{aligned}\frac{\partial}{\partial\sigma_s}\underline{\psi}^k_{s,out} &= A^k_{s\leftarrow v}\underline{\phi}+\sigma_s A^k_{s\leftarrow v}\left[I-D\sigma_s\right]^{-1}D\underline{\phi}\\ &= A^k_{s\leftarrow v}\left\{I+\left[I-D\sigma_s\right]^{-1}\sigma_s D\right\}\underline{\phi}\\ &= A^k_{s\leftarrow v}\left\{I-D\sigma_s+\sigma_s D\right\}\left[I-D\sigma_s\right]^{-1}\underline{\phi}\\ &= A^k_{s\leftarrow v}\left[I-D\sigma_s\right]^{-1}\phi\\ &= A^k_{s\leftarrow v}\left[I+D\sigma_s+(D\sigma_s)^2+(D\sigma_s)^3+\ldots\right]\underline{\phi}\end{aligned} \tag{115}$$

and

$$\begin{aligned}\frac{\partial}{\partial\sigma_t}\psi^k_{s,out} &= \left[\frac{\partial}{\partial\sigma_t}A^k_{s\leftarrow s}\right]\psi^k_{s,in}+\left[\frac{\partial}{\partial\sigma_t}A^k_{s\leftarrow v}\right]\underline{Q}\\ &\quad+\sigma_s A^k_{s\leftarrow v}\left[I-\sigma_s D\right]^{-1}\left\{\sum_k w_k\left(\frac{\partial}{\partial\sigma_t}B^k_{v\leftarrow s}\right)\underline{\psi}^k_{s,in}+D'\underline{Q}\right\}\\ &= \left[\frac{\partial}{\partial\sigma_t}A^k_{s\leftarrow s}\right]\underline{\psi}^k_{s,in}+\left[\frac{\partial}{\partial\sigma_t}A^k_{s\leftarrow v}\right]\underline{Q}\\ &\quad+\sigma_s A^k_{s\leftarrow v}\left[I+\sigma_s D+(\sigma_s D)^2+(\sigma_s D)^3+\ldots\right]\\ &\quad\left\{\sum_k w_k\left(\frac{\partial}{\partial\sigma_t}B^k_{v\leftarrow s}\right)\underline{\psi}^k_{s,in}+D'\underline{Q}\right\}\end{aligned} \tag{116}$$

The last two equations are the analytical solution for the calculation of the derivative of the exiting fluxes with respect to the cross sections in the one-cell problem. The adjoint calculation introduced in the previous section needs to give the same result in order to be tested, at least for this simple case.

First is the calculation of the derivative of the fluxes over the scattering cross section. Recall that in the adjoint method:

$$\frac{\partial}{\partial\sigma_s}\underline{\psi}^k_{s,out} = \sum_n A^n B^T \tag{117}$$

where the matrix *A* in Eq.(117) is not to be confused with the transmission matrix introduced for the *closed linear one-cell functional* method.

Performing this multiplication for *n* less and equal to two leads to:

$$\begin{bmatrix} \frac{\delta\psi^k_{i+\frac{1}{2},j}}{\delta\sigma_s} \\ \frac{\delta\psi^k_{i,j+\frac{1}{2}}}{\delta\sigma_s} \end{bmatrix} = \begin{bmatrix} \frac{d\psi^k_{i+\frac{1}{2},j}}{d\sigma_s} \\ \frac{d\psi^k_{i,j+\frac{1}{2}}}{d\sigma_s} \end{bmatrix} + \sum_{k'} \begin{bmatrix} \frac{\delta\psi^k_{i+\frac{1}{2},j}}{\delta\psi^{k'}_{avg}} \\ \frac{\delta\psi^k_{i,j+\frac{1}{2}}}{\delta\psi^{k'}_{avg}} \end{bmatrix} \frac{\delta\psi^{k'}_{avg}}{\delta\sigma_s} + \sum_{k'} \begin{bmatrix} \frac{\delta\psi^k_{i+\frac{1}{2},j}}{\delta\psi^{k'}_{avg}} \\ \frac{\delta\psi^k_{i,j-\frac{1}{2}}}{\delta\psi^{k'}_{avg}} \end{bmatrix} \sum_{k''} \begin{bmatrix} \frac{\delta\psi^{k'}_{avg}}{\delta\psi^{k''}_{avg}} \\ \frac{\delta\psi^{k'}_{avg}}{\delta\psi^{k''}_{avg}} \end{bmatrix} \frac{\delta\psi^{k''}_{avg}}{\delta\sigma_s} + \dots \tag{118}$$

that is:

$$\frac{\partial}{\partial\sigma_s}\underline{\psi}^k_{s,out} = \frac{\partial}{\partial\sigma_s}\underline{\psi}^k_{s,out} + \sum_{k'} \frac{\delta\underline{\psi}^k_{s,out}}{\delta\psi^{k'}_{avg}} \frac{\delta\psi^{k'}_{avg}}{\delta\sigma_s} + \sum_{k'} \frac{\delta\underline{\psi}^k_{s,out}}{\delta\psi^{k'}_{avg}} \sum_{k''} \frac{\delta\psi^{k'}_{avg}}{\delta\psi^{k''}_{avg}} \frac{\delta\psi^{k''}_{avg}}{\delta\sigma_s} + \dots \tag{119}$$

and using our previous results:

$$\begin{aligned} \frac{\partial}{\partial\sigma_s}\underline{\psi}^k_{s,out} &= A^k_{s\leftarrow v}\phi + \sum_{k'} w_k A^k_{s\leftarrow v}\sigma_s B^{k'}_{v\leftarrow v}\phi + \sum_{k'} w_{k'} A^k_{s\leftarrow v}\sigma_s \sum_{k'} B^{k'}_{v\leftarrow v} w_{k''}\sigma_s B^{k''}_{v\leftarrow v}\phi + \dots = \\ &= A^k_{s\leftarrow v}\phi + A^k_{s\leftarrow v}\phi\sigma_s \sum_{k'} w_k B^{k'}_{v\leftarrow v} + A^k_{s\leftarrow v}\phi\sigma_s\sigma_s \sum_{k'} w_{k'} B^{k'}_{v\leftarrow v} \sum_{k'} w_{k'} B^{k''}_{v\leftarrow v} + \dots = \\ &= A^k_{s\leftarrow v}\phi\left(I + \sigma_s \sum_{k'} w_{k'} B^k_{v\leftarrow v} + \sigma_s{}^2 \sum_{k'} w_{k'} B^{k'}_{v\leftarrow v} \sum_{k'} w_{k''} B^{k''}_{v\leftarrow v} + \dots \right) = \\ &= A^k_{s\leftarrow v}\phi\left(I + D\sigma_s + \left(D\sigma_s\right)^2 + \dots \right) \end{aligned} \tag{120}$$

The result is in perfect accordance with the analytical result obtained in Eq. (115). We can do the same for the scattering derivative:

$$\begin{bmatrix} \frac{\delta\psi^k_{i+\frac{1}{2},j}}{\delta\sigma_{tot}} \\ \frac{\delta\psi^k_{i,j+\frac{1}{2}}}{\delta\sigma_{tot}} \end{bmatrix} = \begin{bmatrix} \frac{d\psi^k_{i+\frac{1}{2},j}}{d\sigma_{tot}} \\ \frac{d\psi^k_{i,j+\frac{1}{2}}}{d\sigma_{tot}} \end{bmatrix} + \sum_{k'} \begin{bmatrix} \frac{\delta\psi^k_{i+\frac{1}{2},j}}{\delta\psi^{k'}_{avg}} \\ \frac{\delta\psi^k_{i,j+\frac{1}{2}}}{\delta\psi^{k'}_{avg}} \end{bmatrix} \frac{\delta\psi^{k'}_{avg}}{\delta\sigma_{tot}} + \sum_{k'} \begin{bmatrix} \frac{\delta\psi^k_{i+\frac{1}{2},j}}{\delta\psi^{k'}_{avg}} \\ \frac{\delta\psi^k_{i,j-\frac{1}{2}}}{\delta\psi^{k'}_{avg}} \end{bmatrix} \sum_{k''} \begin{bmatrix} \frac{\delta\psi^{k'}_{avg}}{\delta\psi^{k''}_{avg}} \\ \frac{\delta\psi^{k'}_{avg}}{\delta\psi^{k''}_{avg}} \end{bmatrix} \frac{\delta\psi^{k''}_{avg}}{\delta\sigma_{tot}} + \dots \tag{121}$$

that is:

$$\frac{\partial}{\partial\sigma_{tot}}\underline{\psi}^k_{s,out} = \frac{d}{d\sigma_{tot}}\underline{\psi}^k_{s,out} + \sum_{k'} \frac{\delta\underline{\psi}^k_{s,out}}{\delta\psi^{k'}_{avg}} \frac{\delta\psi^{k'}_{avg}}{\partial\sigma_{tot}} + \sum_{k'} \frac{\delta\underline{\psi}^k_{s,out}}{\delta\psi^{k'}_{avg}} \sum_{k''} \frac{\delta\psi^{k'}_{avg}}{\delta\psi^{k''}_{avg}} \frac{\delta\psi^{k''}_{avg}}{\partial\sigma_{tot}} + \dots \tag{122}$$

And simplifying again:

$$\frac{\partial}{\partial\sigma_{tot}}\underline{\psi}^{k}_{s,out}=\frac{\delta A^{k}_{s\leftarrow s}}{\delta\sigma_t}\psi^{k}_{IN}+\frac{\delta A^{k}_{s\leftarrow v}}{\delta\sigma_t}Q+\sum_{k'}w_{k'}A^{k}_{s\leftarrow v}\phi\left(\frac{\delta B^{k'}_{v\leftarrow s}}{\delta\sigma_t}\psi^{k'}_{IN}+\frac{\delta B^{k'}_{v\leftarrow s}}{\delta\sigma_t}Q\right)+$$
$$+\sum_{k'}w_{k'}A^{k}_{s\leftarrow v}\phi\sum_{k''}B^{k'}_{v\leftarrow v}w_{k''}\sigma_s\left(\frac{\delta B^{k''}_{v\leftarrow s}}{\delta\sigma_t}\psi^{k'}_{IN}+\frac{\delta B^{k''}_{v\leftarrow s}}{\delta\sigma_t}Q\right)+\ldots$$
$$=\frac{\delta A^{k}_{s\leftarrow s}}{\delta\sigma_t}\psi^{k}_{IN}+\frac{\delta A^{k}_{s\leftarrow v}}{\delta\sigma_t}Q+A^{k}_{s\leftarrow v}\phi\sum_{k'}w_{k'}\left(\frac{\delta B^{k'}_{v\leftarrow s}}{\delta\sigma_t}\psi^{k}_{IN}+\frac{\delta B^{k}_{v\leftarrow s}}{\delta\sigma_t}Q\right)+$$
$$+A^{k}_{s\leftarrow v}\phi\sigma_s D\sum_{k''}w_{k''}\left(\frac{\delta B^{k''}_{v\leftarrow s}}{\delta\sigma_t}\psi^{k''}_{IN}+\frac{\delta B^{k''}_{v\leftarrow s}}{\delta\sigma_t}Q\right)$$
$$=\frac{\delta A^{k}_{s\leftarrow s}}{\delta\sigma_t}\psi^{k}_{IN}+\frac{\delta A^{k}_{s\leftarrow v}}{\delta\sigma_t}Q+A^{k}_{s\leftarrow v}\phi(I+\sigma_s D+\ldots)\left[\sum_{k}w_k\left(\frac{\delta B^{k}_{v\leftarrow s}}{\delta\sigma_t}\psi^{k}_{IN}\right)+D'Q\right] \tag{123}$$

Equation (123) is essentially the same as the analytical results of Eq. (116). To find the gradient, we simply multiply the results by the vector v. The adjoint model gives the correct interpretation of the gradient of the objective function with respect to the cross sections in the discretized domain according to the analytical result for the one-cell problem.

3.4. FORTRAN Coding

TNTs has been developed in Fortran 90 on a Microsoft Windows workstation with a 3.20 GHz Pentium 4 processor. It uses the results of the forward model and the experimental data, inserted by the user as input, to calculate the gradient. All the variables, such as cross sections and geometric measures, are passed from the forward model to the inverse model as common variables.

3.5. Numerical Verification

Another possibility to verify the gradient calculation is by calculating it numerically for a simple problem. By definition, the gradient of a one variable function can be approximated as follows:

$$\frac{\delta f(x)}{\delta x}\approx\frac{f(x+\varepsilon)-f(x)}{\varepsilon} \tag{124}$$

where ε is an infinitesimally small constant. This means that by changing slightly one cross section and leaving all the others un-changed and using Eq. (124) leads to the calculation of the gradient of the objective function with respect to the changed property:

$$\frac{\delta E(\sigma(i,j))}{\delta\sigma(i,j)}\approx\frac{E(\sigma(i,j)+\varepsilon)-E(\sigma(i,j))}{\varepsilon} \tag{125}$$

The gradient calculated numerically for all the cells of the domain can be therefore compared with the one obtained by using the adjoint formulation to give an estimation of the reliability of TNTs.

A test with a two by two cell homogeneous object is presented. The optical properties were 0.9 cm^{-1} and 1.2 cm^{-1} for scattering and total cross section, respectively. One by one every cross section has been changed adding 0.001 cm^{-1} to its value and the gradient has been calculated. It is important to note that when the scattering cross section has been changed also the absorption in the cell has been changed to keep the total cross section constant.

Tables III and IV contain the scattering gradients for the 2x2 problem from the numerical results and from TNTs result, respectively. The percent error between the numerical and TNTs results is shown in Table V. Tables VI and VII contain the total cross section gradient calculated numerically and by TNTs respectively. The percent error of the total cross section gradient is presented in Table VIII.

TABLE III Numerical result for the scattering gradient

0.11388464	0.11388464
0.84114470	0.84114470

TABLE IV TNTs (adjoint calculation) results for the scattering gradient

0.11387102	0.11387102
0.84102871	0.84102871

TABLE V Percent error in scattering gradients

0.012	0.012
0.014	0.014

TABLE VI Numerical results for the total cross section gradient

-0.11985335	-0.11985335
-1.18359350	-1.18359350

TABLE VII TNTs results for the total cross section gradient

-0.11986997	-0.11986997
-1.18379570	-1.18379570

TABLE VIII Percent error in total cross section gradient

0.014	0.014
0.017	0.017

The difference between the numerical and TNTs calculated values is negligible. The error assuming the numerical result is exact is less than 0.02%. This verification shows the effectiveness of the adjoint formulation in the gradient calculation.

3.6. Reconstruction

In this section, we will verify that the derived method can reconstruct an object. The geometric properties of the object and the collimator in question are shown in Fig. 37. A summary of these properties is given in Table IX.

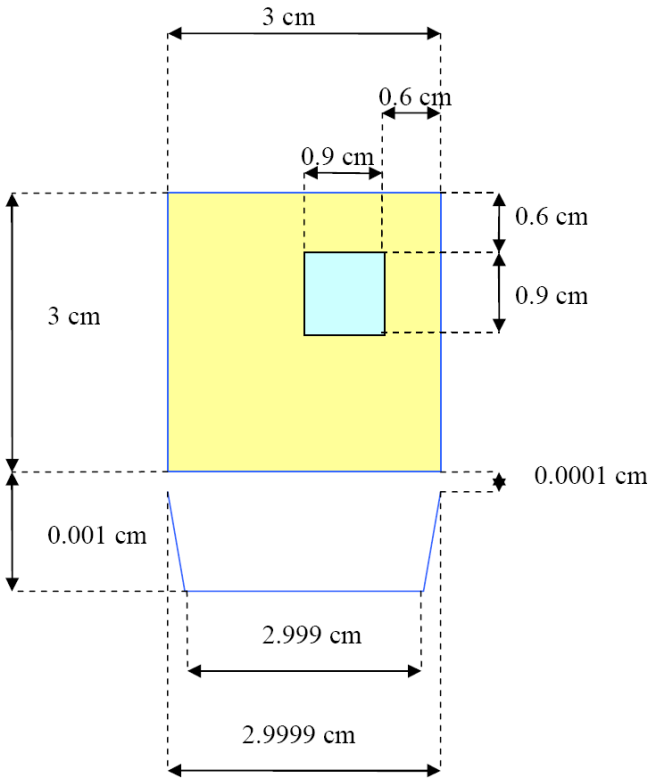


Fig. 37. View of the object and the collimator for the reconstruction verification of the process.

TABLE IX Geometric, discretization and optical parameters of the object

Object Dimensions		
x length	3	cm
y length	3	cm
Collimator Dimensions		
Beam window length	2.999	cm
Collimator window length	2.9999	cm
Distance beam window object	0.001	cm
Distance collimator window object	0.0001	cm
Discretization Forward Problem		
# x cell	20	
# y cell	20	
Order quadrature set (n of Sn)	4	
# polar direction	20	
Discretization for Unc. Sources		
# points per cell	1	
# azimuthal angles	20	
# polar angles	20	
Discretization Detectors		
# detector points	1	
# azimuthal angles	20	
# polar angles	20	
Optical Properties Background		
Scattering cross section	1.8	cm^{-1}
Absorption cross section	0.5	cm^{-1}
Total cross section	2.3	cm^{-1}
g factor	0.02	
Optical Properties Intrusion		
Scattering cross section	0.5	cm^{-1}
Absorption cross section	2	cm^{-1}
Total cross section	2.5	cm^{-1}
g factor	0.02	

The object designed for the reconstruction was optically thick. Due to its properties it was around 7 mean free paths long and 7 mean free paths wide. Less than 0.1% of the initial radiation was transmitted radiation in the detectors. A tomographic technique applied to this object to attempt its reconstruction would have found the task hard to accomplish. Most of the radiation reached the detectors after multiple scattering that would have decrease the resolution of a radiographic image. The problem is therefore interesting to solve with the gradient based computed tomography introduced. The optical distribution of the real object is shown in Figs. 38 and 39.

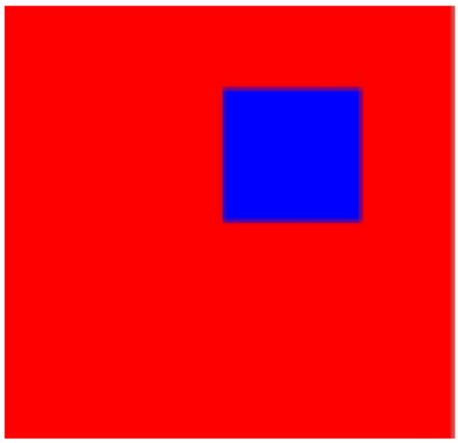


Fig. 38. Scattering cross section distribution of the real object.

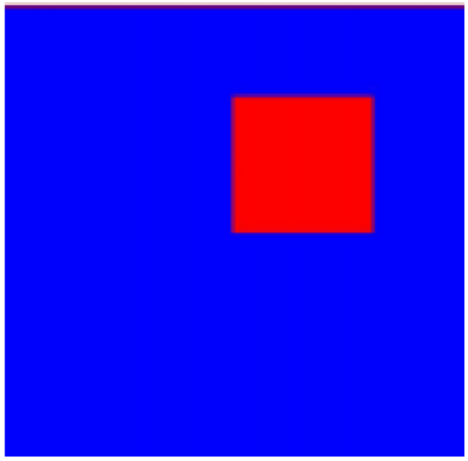


Fig. 39. Total cross section distribution of the real object.

The object has been discretized with a 20 by 20 grid. The number of unknown variables were therefore 800, since the reconstruction was made by using only scattering and total cross section. There were 60 detector measurements that were simply the result given by the forward model with as input the real object. To attempt the reconstruction the initial guess for the

optical properties was equal to the properties of the background object, the part without the intrusion. The results after 300 iterations (almost 7 hours) are shown in Figs. 40 and 41.

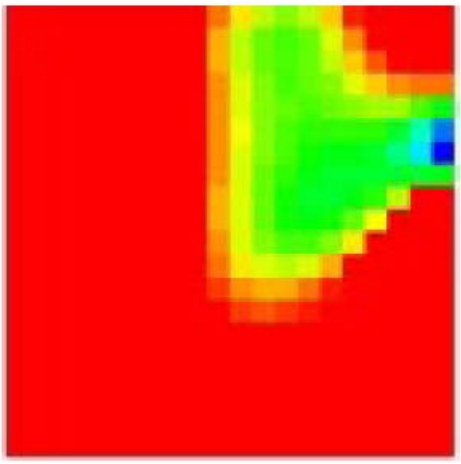


Fig. 40. Scattering distribution of the reconstructed object.

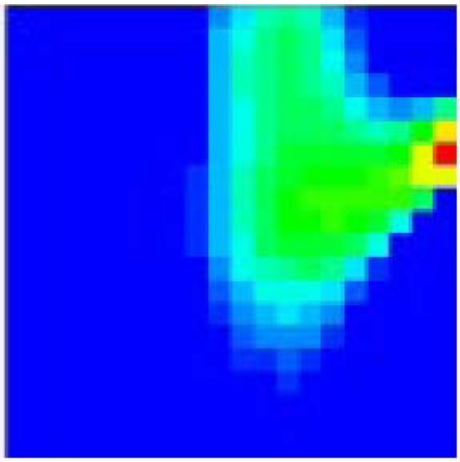


Fig. 41. Total cross section distribution of the reconstructed object.

Both the optical distribution for the reconstructed object show the presence of a different object inside the homogeneous one used as initial guess. The green part in Fig. 40 and Fig. 41 is more absorbing and less scattering than the background medium. In particular the scattering cross section at the center of the green spot of Fig. 40 is around 1.5 cm^{-1} instead of the real 0.5 cm^{-1} and the absorption cross section is around 0.8 cm^{-1} instead of 2.0 cm^{-1}.

The intrusion is well detected in its position, but the shape tends to be less representative of the real object. Due to the poor discretization (only 400 cells) the reconstructed area is “bounded” to the edge of the object were the detectors are placed. Since the process steps back from the boundaries of the object to reconstruct the optical distribution it is reasonable that the

reconstructed area, in case of few cell between the defect and the detector, is bounded to the boundaries. The process changes the cell properties in the right direction, and it presents an intrusion in the right position which is more absorbing and less scattering than the surrounding medium. The objective function decreased three order of magnitude from it's initial value showing a correct behavior of the inverse model.

The result presented corresponds to that using the conjugate gradient (CG) scheme minimization technique. The same result was calculated with steepest descent (SD) scheme for comparison. In particular, the final optical distributions were the same for both methods but, as it was expected, the CG took considerably less iteration to converge to the solution. A plot of the objective function with respect to number of iteration obtained with both methods is shown in Fig. 42.

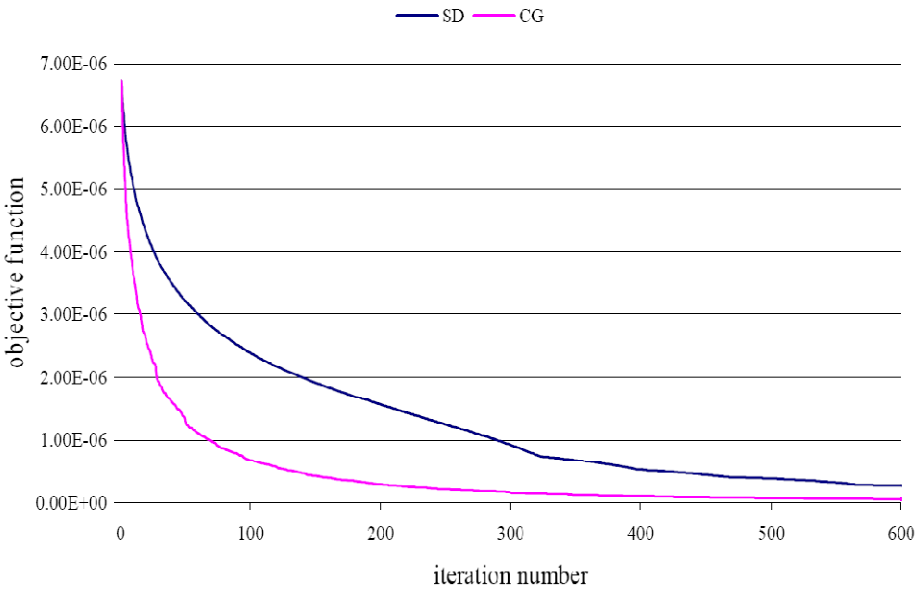


Fig. 42. Objective function with respect to iteration number for CG and S.

In the analysis of the above problem other important features were investigated. In particular nothing in the literature has been found regarding the shape of the line along which TNTs finds the minimum for the step α.

3.7. Line Search, Negativity and Gradient Perpendicularity

The method to use for the line search, such the step α to minimize the line could be found efficiently required knowledge of the shape of the line. In particular, since the objective function has in general more than one minimum it could also be that the line along the negative gradient contains more local minima. In order to verify that at the first iteration of the problem presented above, the line $f(\alpha)=E(\underline{\sigma}-\alpha\nabla E(\underline{\sigma}))$ has been calculated for different values of α. These results are shown in Fig. 43.

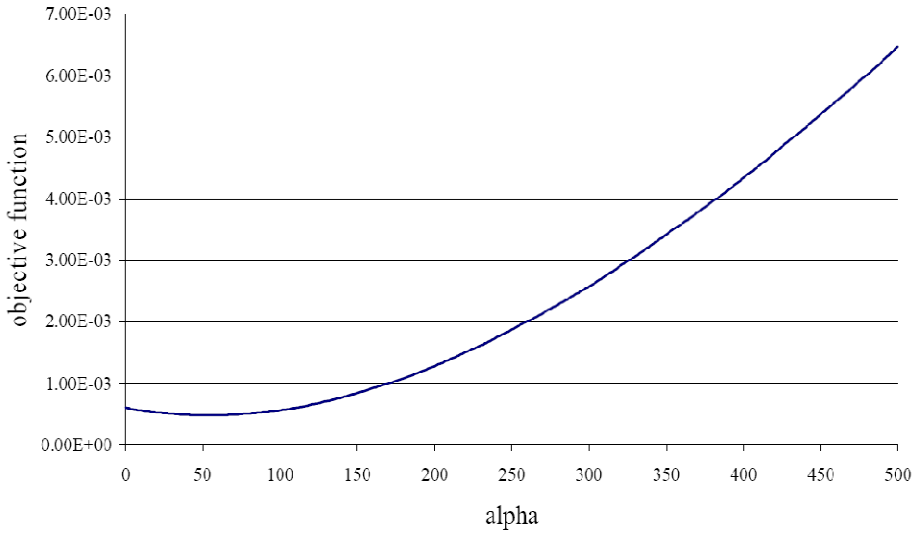


Fig. 43. Objective function with respect to alpha.

The behavior showed in Fig. 43 is interesting. Not only is there only one minimum but the shape of the function is essentially that of a second order polynomial. A polynomial fitted through the points of the function has an R^2 of 0.9992. A cubic has an R^2 equal to 1.

Since this result seems due to the particular homogeneous guess of the first iteration, a plot of the same function starting from a completely heterogeneous object (every cell different from the others) is presented in Fig. 44.

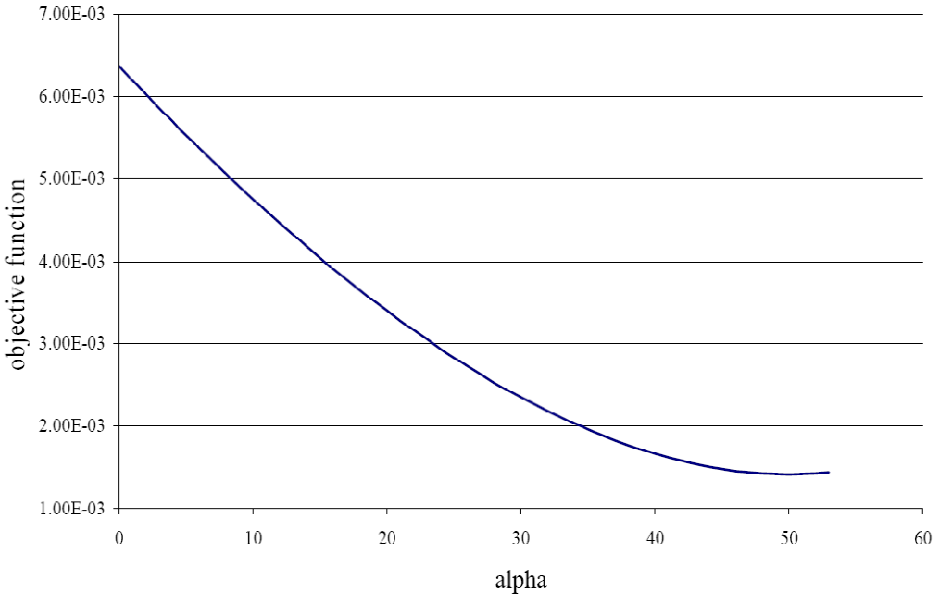


Fig. 44. Objective function with respect to alpha for an heterogeneous object.

At α>53, one cross section would turn negative therefore TNTs stops. If there is a minimum point of the line before TNTs encounters the alpha that makes negative a cross section, a new gradient is calculated at the minimum point, otherwise the gradient is calculated in the point with the maximum alpha that assures positivity of all the optical properties. In the case of Fig. 44, there is a minimum around α=50. Again there is the same behavior as before even with a more complex initial guess.

There are two important considerations that can be made around the example presented above:

- during the line search, cross sections can become negative
- the function of alpha seems to be a smooth function with only one minimum (it cannot be elevated as general rule without further investigation)

When a cross section turns negative before a minimum can be found, TNTs comes back to the maximum α to assure the positivity of all the optical properties and calculates a new gradient. It moves therefore along another direction until a global minimum for the objective function is found.

For the minimization of a function without any knowledge about its gradient many algorithms have been developed. A golden section search is designed to handle, in effect, the worst possible case of function minimization[37]. But, if the function is nicely parabolic near the minimum, then the parabola fitted through any three points ought to reach in a single leap the minimum or at least a point very close to it. But a method that uses only parabolic fittings is not likely to succeed in practice. The task is to use a scheme which relies on sure-but-slow technique, like golden section search, when the function is not cooperative and switches over to a parabolic fit when the function allows. A method with this capability is the Brent's method that is up to the task in all particulars. In the worst possible case, where the parabolic steps are useless, the method will approximately alternate between parabolic steps and golden sections, converging in due course by virtue of the latter[8].

When a minimum is found the gradient calculated at that location (that particular α) should be perpendicular to the previous gradient. For the gradients calculated by TNTs the dot product of the two gradients at the alpha for the minimum is around 0.01, proving a good perpendicularity of the directions.

4. RECONSTRUCTION RESULTS

This section presents the result of different object reconstructions. The reliability of TNTs and its algorithm was tested using data obtained with MCNP-5 and from experimental measurements of simple test objects. The MCNP results are considered equivalent to experimental results since many successful benchmarks have been done for the performance of MCNP in similar problems. All objects simulated and reconstructed were essentially two-dimensional objects (infinite in the z-direction). All reconstruction were done using one neutron group. The simulations used a monoenergetic thermal neutron beam and the experiments used the predominately thermal neutron beam available at the TAMU NSC.

4.1. MCNP Decks

Some objects reconstructed by TNTs were explicitly simulated using MCNP. The MCNP simulation contained:

- planar source finite in x and y directions and infinite along the z-axis to simulate the beam window introduced in TNTs (see Section 2)
- object, finite in x and y directions and infinite along the z-axis, with different material properties in it
- detectors around the edges of the object

In order to have an infinite homogeneous third direction, so that the problem can be assumed two-dimensional, the MCNP model contains reflecting planes perpendicular to the z-directions. The rectangular beam and the divergent collimator are shown in Fig. 45. A spherical source is placed behind the collimator. The source energy is 1E-6 MeV and the spatial distribution follows a power law of order 1 centered at the origin of the sphere (the power law is the right distribution to make the beam at the window isotropic as it is in TNTs). The radius of the source is larger enough to cover the entire beam window. The materials for the object that we chose to use are carbon, boron, hydrogen and nitrogen because of their constant cross sections around energies close to the one chosen for the source. This choice was made because TNTs is not a multi-group code and it uses constant cross sections without knowledge about the neutron energy.

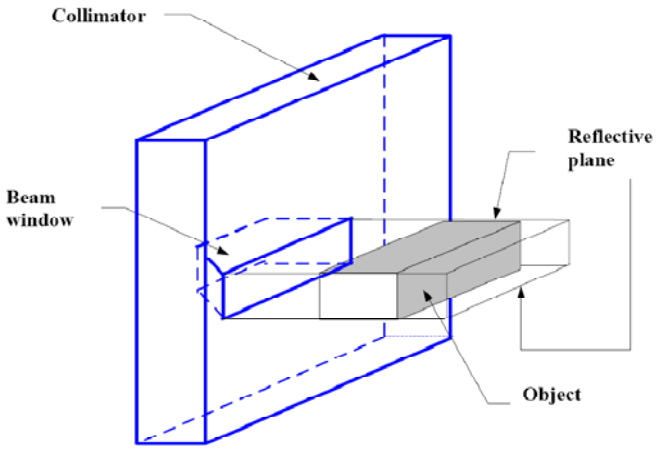


Fig. 45. MCNP design.

The particles incident to the collimator are stopped and only those that enter the beam window can reach the object. The collimator is divergent as is the one inserted in the input file for TNTs. The objective is to have a beam coming from the beam window with the same spatial and angular distribution as the beam in TNTs. The particles out of the collimator encounter the object or the reflective planes so that the third infinite and homogeneous direction is created.

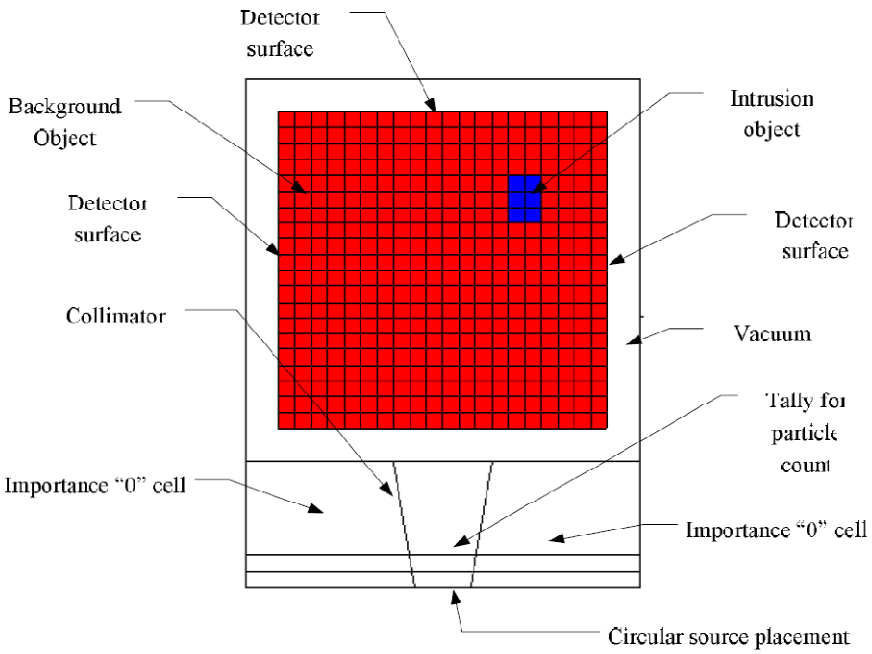


Fig 46. xy view of the MCNP layout.

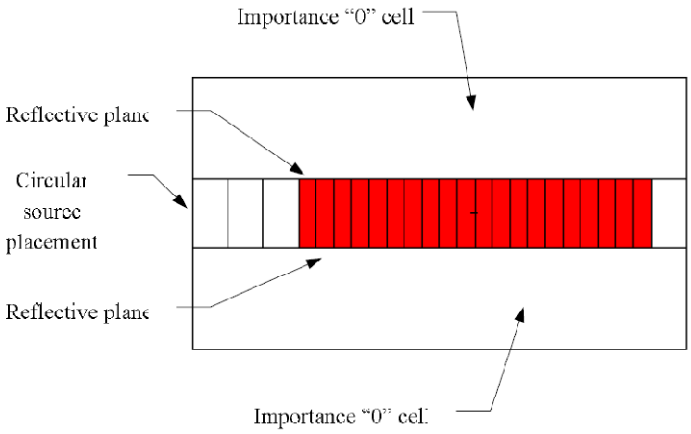


Fig. 47. xz view of the MCNP layout.

The three outer surfaces of the object as shown in Fig. 46 are used to tally the detector readings. The MCNP tally used is a surface current tally. There are 20 detectors, equally spaced, on each side for a total of sixty measurements. Figures 47 and Fig. 46 show the geometrical layout for the problem considered.

The importance "0" cells shown in Fig. 46 and Fig. 47 represent the collimator. In MCNP importance "0" defines a cell in which particles are not tracked. Therefore, the particles entering these cells are "terminated". Only the particle entering the beam window can reach the object as shown in all the pictures. These particles are counted, using a surface current tally, in order to normalize the results on the effective number of particle that form the beam.

4.2. MCNP Calculations

To run the cases a parallel version (PVM) of MCNP was used. This helped to reduce the computational time and to increase the particle number for better statistics. The cases were run on a LINUX cluster consisting of 21 nodes. Each node had a single 3.2 GHz INTEL Pentium 4 processor.

4.3. Results from MCNP Simulated Objects

Five different objects were simulated in MCNP. The results from these simulations were inputted as "measured" images to TNTs and a distribution of optical properties was reconstructed. A description of each of the objects and the results of the reconstructions are given below.

4.3.1 Reconstruction I

The first object reconstructed is shown in Fig. 48. The properties of the object are presented in Table X. They have been chosen so that the problem is approximately seven MFP thick.

The measurements along the three non-lightened edges were interpolated to calculate the values for the "experimental" TNTs input. TNTs, in fact, uses a 40 by 40 grid with a total of 120 measurements. A summary of the problem is presented in Table XI.

By performing the reconstruction of the object only changing the scattering and total cross sections, the task is represented by the minimization of a function of 3200 variables having 120 measurements. Even if TNTs changes scattering and total cross sections, the results are presented in terms of scattering and absorption distribution. This is because it is easier to understand the property of the object and of the extraneous body in terms of these cross sections.

The result of the reconstructed scattering distribution is shown in Fig. 49.

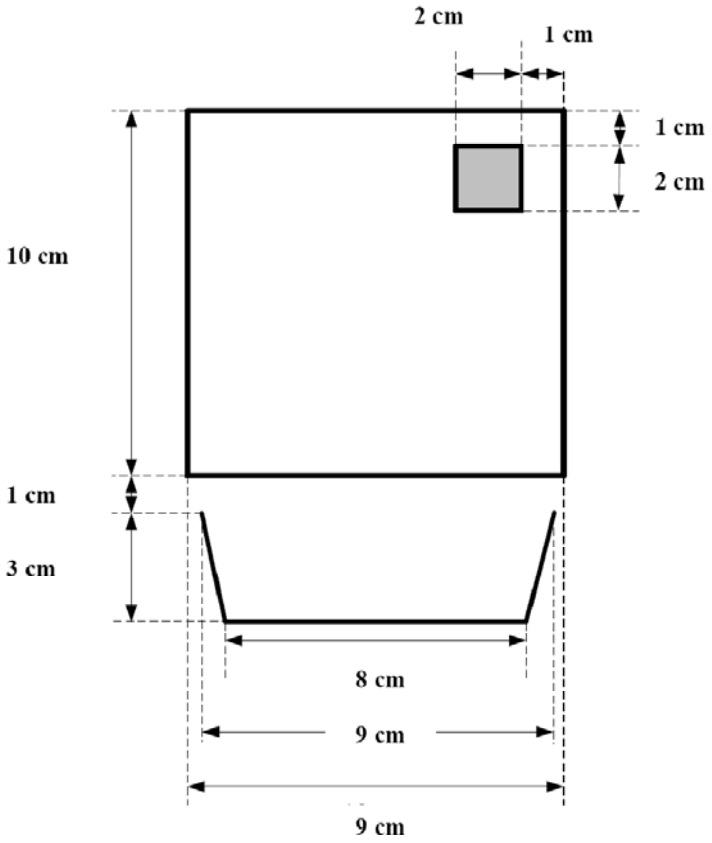


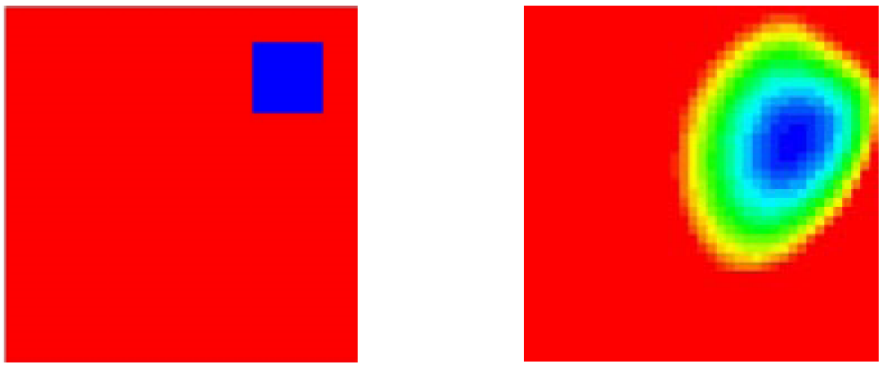
Fig. 48. Object used in first reconstruction process.

TABLE X Properties of the investigated object

Background object:		
Polyethylene (C2H4)		
density	0.3	g/cm^3
σ_{sca}	5.8834E-01	cm^{-1}
σ_{abs}	8.5955E-03	cm^{-1}
G	0.02	
Intrusion object		
Boron Nitride (BN)		
density	0.06	g/cm^3
σ_{sca}	2.1078E-02	cm^{-1}
σ_{abs}	1.1142E+00	cm^{-1}
g	0.02	

TABLE XI TNTs discretization and optical properties

Discretization Forward Problem		
# x cell	40	
# y cell	40	
Order quadrature set (n of Sn)	8	
# polar direction	20	
Discretization for Unc. Sources		
# points per cell	1	
# azimuthal angles	20	
# polar angles	20	
Discretization Detectors		
# detector points	1	
# azimuthal angles	20	
# polar angles	20	
Optical Properties Background		
Scattering cross section	5.8834E-01	cm^{-1}
Absorption cross section	8.5955E-03	cm^{-1}
Total cross section	5.9694E-01	cm^{-1}
g factor	0.02	



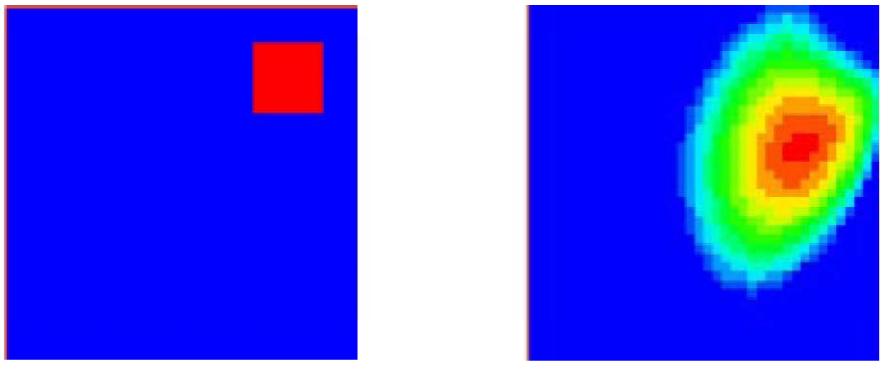
(a) (b)

Fig. 49. Scattering Cross Section from First Object Reconstruction. (a) Image of the scattering property of the real object. (b) Image of the scattering property of the reconstructed object.

The reconstructed scattering distribution presents interesting properties:

- the intrusion has been identified
- the location of the intrusion is slightly different from the real one, in particular is correct the distance from the right edge but is slightly off the distance from the top edge
- the scattering cross section in the center of the intrusion is around 0.4 cm^{-1}that is far from the real one (0.021 cm^{-1}) but represents a change in the cross section equal to 32% of the initial value in the right direction

For the absorption cross section reconstruction the result is presented in Fig. 50.



(a) (b)

Fig. 50. Absorption Cross Section from First Object Reconstruction. (a) Image of the absorption property of the real object. (b) Image of the absorption property of the reconstructed object.

The reconstructed absorption properties present interesting properties:

- the intrusion has been identified
- the location of the intrusion is slightly different from the real one, in particular is correct the distance from the right edge but is slightly off the distance from the top edge
- the absorption cross section in the center of the intrusion is around 0.997E-3 cm^{-1} that is far from the real one (1.11 cm^{-1}) but represents a change in the cross section equal to 16% of the initial value in the right direction

A more absorbing and less scattering intrusion has therefore been found with this method. It is important to notice that these results have been obtained lightening only one edge of the object.

In the minimization process, for the reconstruction of the optical distribution of the object, the objective function dropped by almost two orders of magnitude reaching a stable value. The norm of the gradient, initially of the order 10^{-7}, determined the criterion to stop the process when it reached order 10^{-12}. The final distribution was almost completely reached within the first 10 iterations. After that the process changed slightly the properties until

convergence of the gradient. The time required to reach a reasonable distribution like the one presented above is between 1 and 3 hours.

4.3.2 Reconstruction II

The second object tested is shown in Fig. 51. Its optical properties are summarized in Table XII. The object has an extraneous body in it that is a tenth of the width of the object itself. The input files for TNTs are summarized in Table XIII. The reconstructions for both the properties are presented in Figs. 52 and 53.

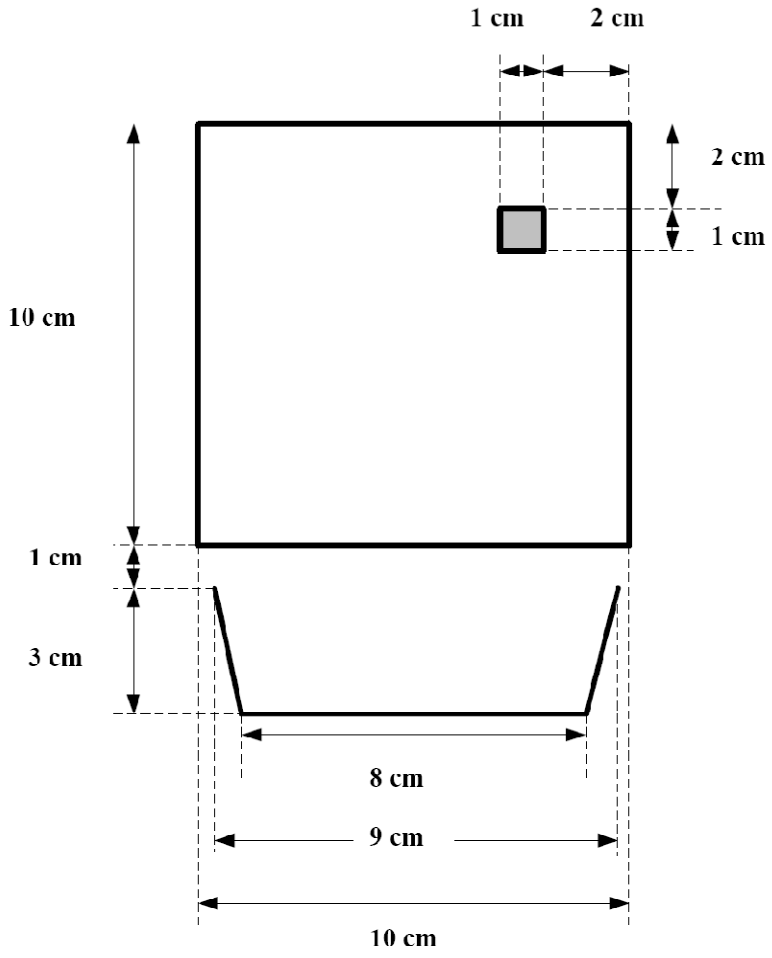


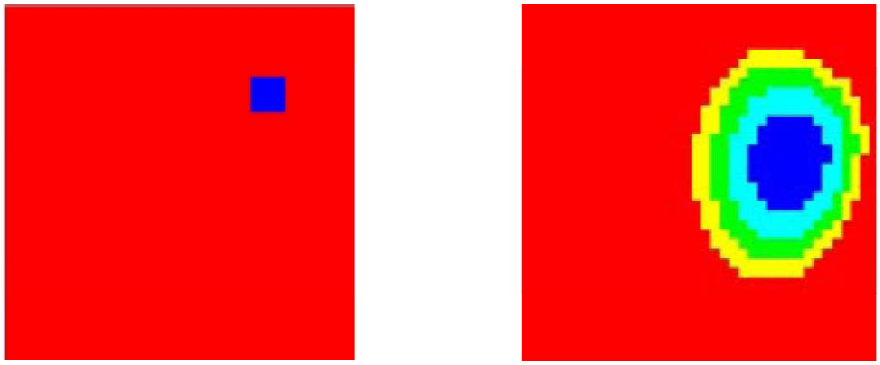
Fig. 51. Second reconstruction object and beam properties.

TABLE XII Optical properties of second reconstruction object

Background object:		
Carbon (C)		
density	3	g/cm^3
σ_{sca}	0.712850935	cm^{-1}
σ_{abs}	0.000505397	cm^{-1}
g	0.055506063	
Intrusion object		
Boron (N)		
density	0.1	g/cm^3
σ_{sca}	0.011942621	cm^{-1}
σ_{abs}	21.37305892	cm^{-1}
g	0.061665588	

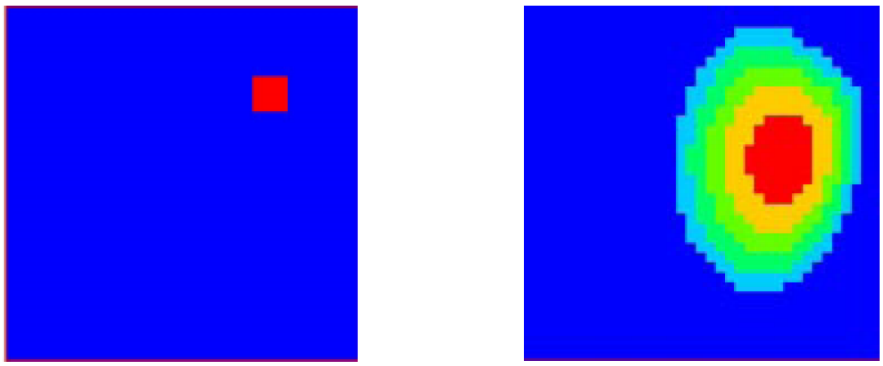
TABLE XIII TNTs discretization parameters and optical properties

Discretization Forward Problem		
# x cell	40	
# y cell	40	
Order quadrature set (n of Sn)	8	
# polar direction	20	
Discretization for Unc. Sources		
# points per cell	1	
# azimuthal angles	20	
# polar angles	20	
Discretization Detectors		
# detector points	1	
# azimuthal angles	20	
# polar angles	20	
Optical Properties Background		
Scattering cross section	0.712850935	cm^{-1}
Absorption cross section	0.000505397	cm^{-1}
Total cross section	0.713355632	cm^{-1}
g factor	0.055	



(a) (b)

Fig. 52. Scattering Cross Section from Second Object Reconstruction. (a) Image of the scattering property of the real object. (b) Image of the scattering property of the reconstructed object.



(a) (b)

Fig. 53. Absorption Cross Section from Second Object Reconstruction. (a) Image of the absorption property of the real object. (b) Image of the absorption property of the reconstructed object.

There are common considerations that can be made on both the reconstructions:

- the extraneous object is detected
- again, as in the previous section, the distance from the top is slightly off but the one from the right edge is correct
- the properties of the reconstructed object and its dimension are not close to those of the real object but still the process moved the cross section in the right direction, to create an area less scattering and more absorbing

The process and algorithms were able to detect the extraneous body and to locate it almost precisely into the object. Even if the properties were not the ones of the real object, still TNTs was able to find a less scattering and more absorbing area that corresponds to the intrusion material.

4.3.3 Reconstruction III

A similar object to the previous one was reconstructed in which the intrusion material has a scattering cross section comparable to the scattering cross section of the background medium. The properties of this material are shown in Table XIV.

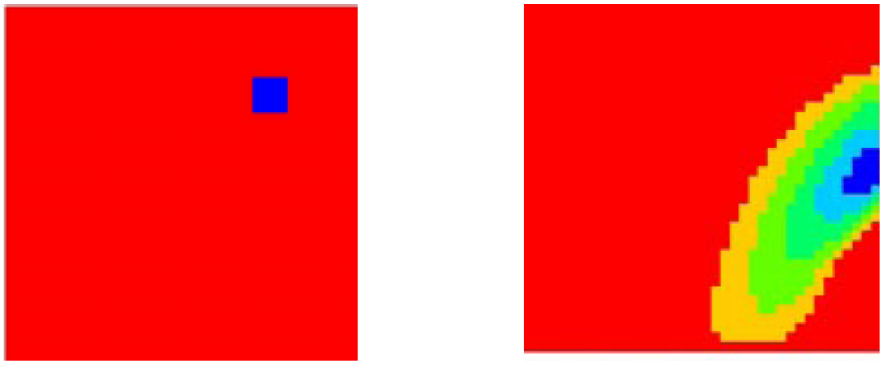
TABLE XIV Property distribution

Background object:		
σ_{sca}	0.9	cm^{-1}
σ_{abs}	0.3	cm^{-1}
g	0.02	
Intrusion object		
σ_{sca}	0.3	cm^{-1}
σ_{abs}	12.4	cm^{-1}
g	0.02	

TABLE XV Discretization and optical properties for TNTs input file

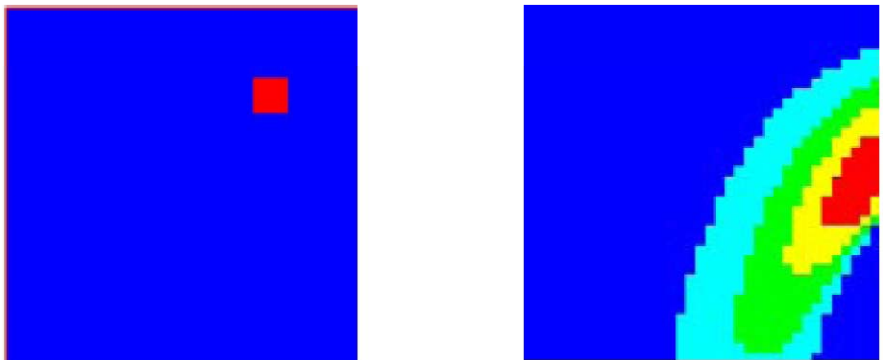
Discretization Forward Problem		
# x cell	40	
# y cell	40	
Order quadrature set (n of Sn)	8	
# polar direction	20	
Discretization for Unc. Sources		
# points per cell # azimuthal angles # polar angles	1 20 20	
Discretization Detectors		
# detector points	1	
# azimuthal angles	20	
# polar angles	20	
Optical Properties Background		
Scattering cross section	0.9	cm^{-1}
Absorption cross section	0.3	cm^{-1}
Total cross section	1.2	cm^{-1}
g factor	0.02	

In particular the scattering cross section is comparable numerically in both the materials. In the previous section there was at least an order of magnitude difference between the scattering cross sections of the two materials. In this simulation they are within a factor of 3. The experimental data for this reconstruction have been calculated directly from TNTs and not from MCNP as in the other reconstructions. The process used the input properties shown in Table XV. The results are shown in Figs. 54 and 55.



(a) (b)

Fig. 54. Scattering Cross Section from Third Object Reconstruction. (a) Image of the scattering property of the real object. (b) Image of the scattering property of the reconstructed object.



(a) (b)

Fig. 55. Absorption Cross Section from Third Object Reconstruction. (a) Image of the absorption property of the real object. (b) Image of the absorption property of the reconstructed object.

It is very interesting to notice that even if the geometric measures of the extraneous body into the object are the same of the previous reconstruction the results are different. The comparison of Fig. 54b and Fig. 55b with Fig. 52b and Fig. 53b underlines the difference in the reconstruction due to the change in the optical properties. In particular, when the scattering cross section of the extraneous object is comparable in order with the scattering cross section of the background object the extraneous body is not well defined in the reconstruction. In Fig. 54b and Fig. 55b, the area where the process changed the properties is bounded to the boundary

edge like if the scattering events throughout the entire object covered the effect of the absorption of the extraneous body. The scattering still present in the intrusion (only a factor of 3 less than in the background) increases the "noise" in the detectors. In other words, especially for the detection of little objects, it seems that the process to reconstruct properly relies more on the "lack" of scattering sources than on the importance of the absorption in the extraneous body cells. In particular in the II reconstruction the low value of the scattering cross sections in the extraneous body led to "zero source" in its cells. The effect on the detection on this edge was more important than in this case where the body is still as scattering as the background and strongly absorbing. Also in this case the reconstruction started a little lower on the y-axis in comparison with the real position of the extraneous object.

4.3.4 Reconstruction IV

The next reconstruction involves the presence of two extraneous bodies into the background object as shown in Fig. 56. The properties of the object and the defect are listed in Table XVI. Again the experimental results were obtained by using MCNP. TNTs had the initial guess for the material and the discretization parameters presented in Table XVII. The process reached a converged solution in 5 iterations (almost an hour of computational time) dropping the objective function of one order of magnitude from 10^{-6} to 10^{-7}. The results obtained are presented in Fig. 57 and Fig. 58.

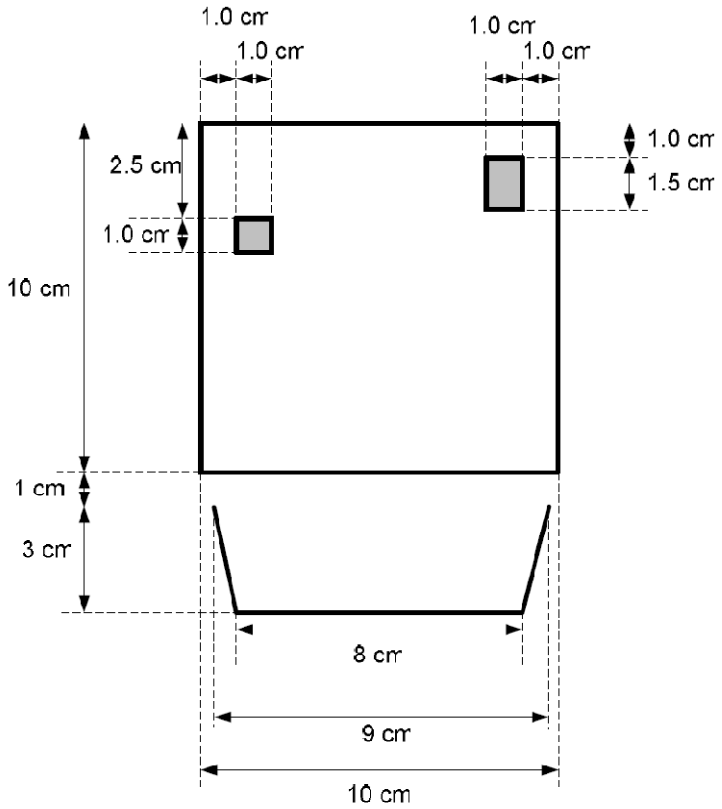


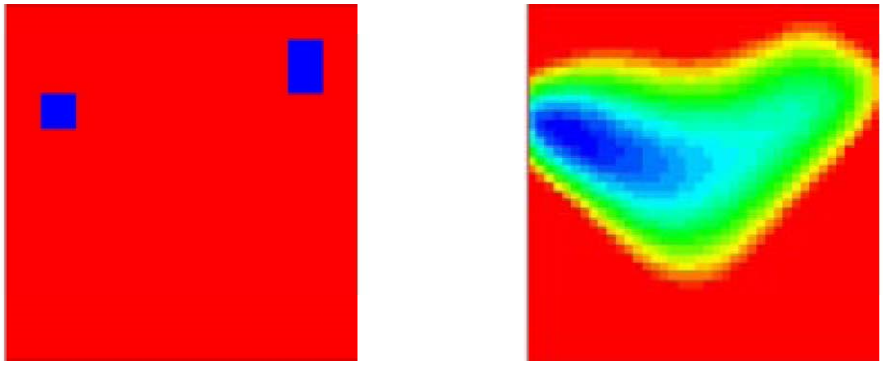
Fig. 56. Third reconstruction object and beam properties.

TABLE XVI Optical properties of materials for reconstruction

Background object:		
Polyethylene (C2H4)		
density	0.3	g/cm^3
σ_{sca}	5.8834E-01	cm^{-1}
σ_{abs}	8.5955E-03	cm^{-1}
G	0.02	
Intrusion object		
Boron Nitride (BN)		
density	0.06	g/cm^3
σ_{sca}	2.1078E-02	cm^{-1}
σ_{abs}	1.1142E+00	cm^{-1}
g	0.02	

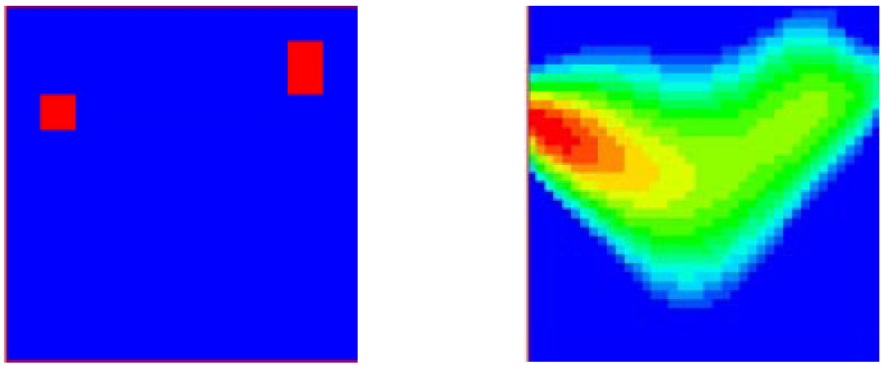
TABLE XVII Discretization parameters and optical properties for the initial guess

Discretization Forward Problem		
# x cell	40	
# y cell	40	
Order quadrature set (n of Sn)	8	
# polar direction	20	
Discretization for Unc. Sources		
# points per cell	1	
# azimuthal angles	20	
# polar angles	20	
Discretization Detectors		
# detector points	1	
# azimuthal angles	20	
# polar angles	20	
Optical Properties Background		
Scattering cross section	5.8834E-01	cm^{-1}
Absorption cross section	8.5955E-03	cm^{-1}
Total cross section	5.9694E-01	cm^{-1}
g factor	0.02	



(a) (b)

Fig. 57. Scattering Cross Section from Fourth Object Reconstruction. (a) Image of the scattering property of the real object. (b) Image of the scattering property of the reconstructed object.



(a) (b)

Fig. 58. Absorption Cross Section from Fourth Object Reconstruction. (a) Image of the absorption property of the real object. (b) Image of the absorption property of the reconstructed object.

It is clear from intuition and from the images that the reconstruction of an object with two extraneous bodies is more difficult to obtain. In both the reconstructions, the body closer to the left edge has been detected, but the one closer to the right edge is not well defined. Even if not well defined it is visible the presence of "something" else that affect the shape of the reconstructed area. At first sight it seems incorrect that the body closer to two edges (the top and right edge) is the one that is less defined in the converged image for the optical distribution. The other body, the one closer to the left edge, is though also closer to the lightened edge. The calculation of the gradient as introduced in Chapter III showed how the gradient is in part function of the scalar flux of the cell. Due to its position the extraneous body on the left is in an area with higher scalar flux. Its gradient is therefore larger than the gradient of the body on the right. It seems that the process tends to minimize the objective function by looking at every moment at the spot with higher gradient because it is the one that affects the objective function the most; then after that spot has been moved enough toward the direction of the minimum for the objective function the process looks for other possible spots to adjust. This is the reason why

there is an effect of the presence of the second body, the one closer to the right edge; but this body is not well defined because by the time that the process tries to change that area the property already changed have brought the objective function to a minimum (local minima).

4.3.5 Reconstruction V

To prove this supposition a new object is considered with a different geometric disposition for the two bodies in order to make one predominant, in terms of the gradient, over the other. The size of the bodies is also changed but it is the same for both of them. The body closer to the left edge has been moved closer to the bottom edge, so that it will be into a higher flux region. The other body has been increased in size so that its effect on the detection measurements will be more consistent (Fig. 59). The properties are shown in Table XVIII. The discretization parameters for the forward and inverse model are the same used in all the previous examples and again the initial guess for the properties are the ones of the background object as shown in Table XIX. The real optical property distribution and the reconstruction results are shown in Fig. 60 and Fig. 61.

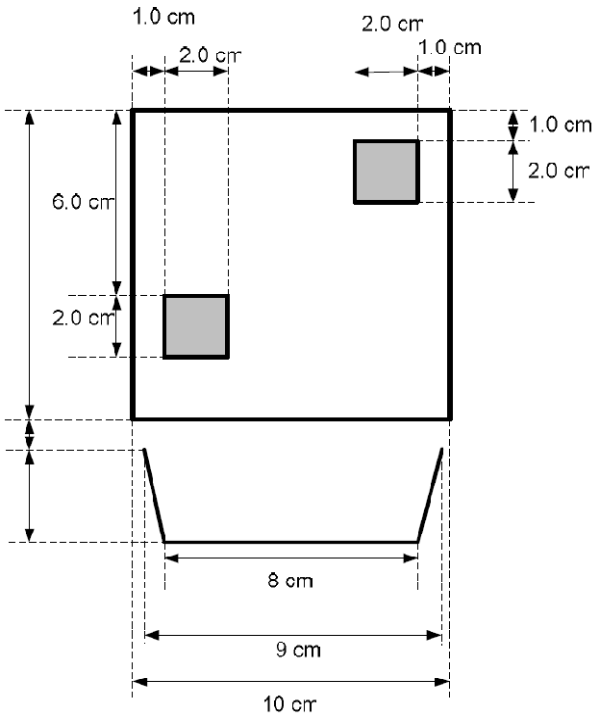


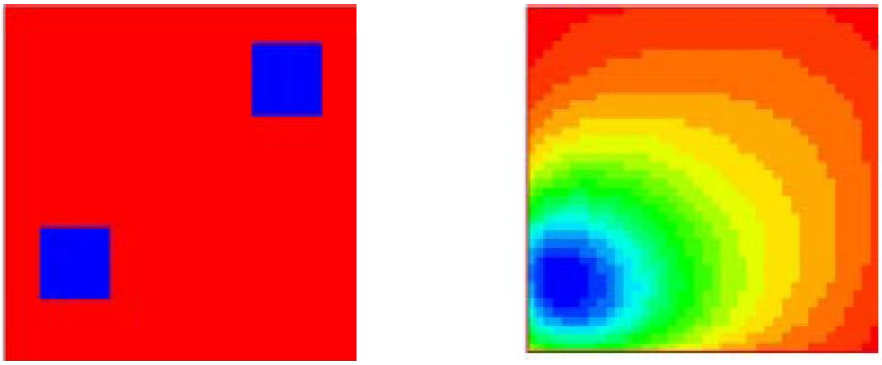
Fig. 59. Beam and object geometric properties.

TABLE XVIII Background and extraneous object optical properties

Background object:		
Carbon (C)		
density	3	g/cm^3
σ_{sca}	0.712850935	cm^{-1}
σ_{abs}	0.000505397	cm^{-1}
g	0.055506063	
Intrusion object		
Boron (N)		
density	0.1	g/cm^3
σ_{sca}	0.011942621	cm^{-1}
σ_{abs}	21.37305892	cm^{-1}
g	0.061665588	

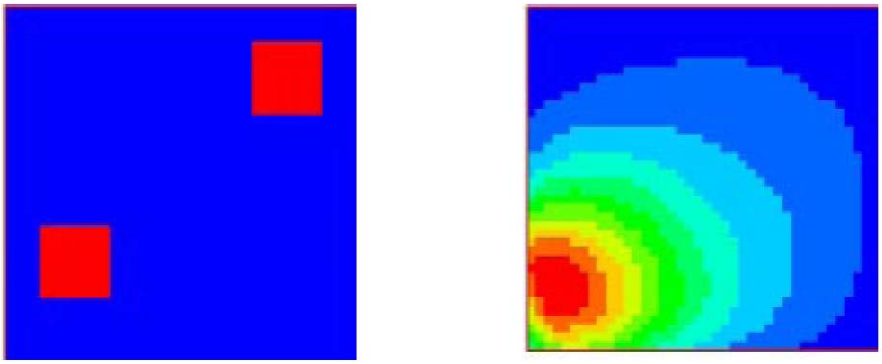
TABLE XIX TNTs discretization parameters and initial guess

Discretization Forward Problem		
# x cell	40	
# y cell	40	
Order quadrature set (n of Sn)	8	
# polar direction	20	
Discretization for Unc. Sources		
# points per cell	1	
# azimuthal angles	20	
# polar angles	20	
Discretization Detectors		
# detector points	1	
# azimuthal angles	20	
# polar angles	20	
Optical Properties Background		
Scattering cross section	5.8834E-01	cm^{-1}
Absorption cross section	8.5955E-03	cm^{-1}
Total cross section	5.9694E-01	cm^{-1}
g factor	0.02	



(a) (b)

Fig. 60. Scattering Cross Section from Fifth Object Reconstruction. (a) Image of the scattering property of the real object. (b) Image of the scattering property of the reconstructed object.



(a) (b)

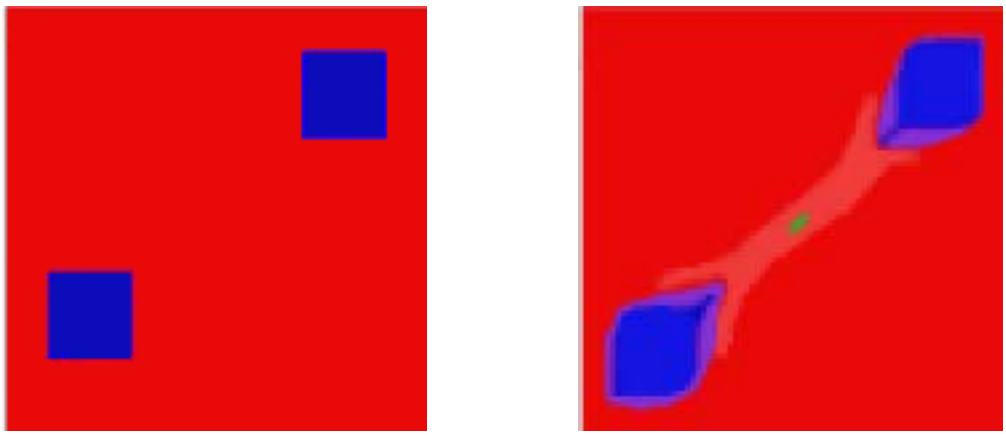
Fig. 61. Absorption Cross Section from Fifth Object Reconstruction. (a) Image of the absorption property of the real object. (b) Image of the absorption property of the reconstructed object.

As expected the body closer to the bottom edge, the lightened one, has a predominant gradient over the other extraneous body. It is so predominant that the algorithm fails in the reconstruction of the optical distribution of the entire object. This simulation seems to confirm the consideration that the process reconstructs before the spot with the higher gradient and then the other spots in order of gradient. The algorithm was successful in the reconstruction of one object and it fails in the reconstruction of the other.

4.4. Results from Measured Images

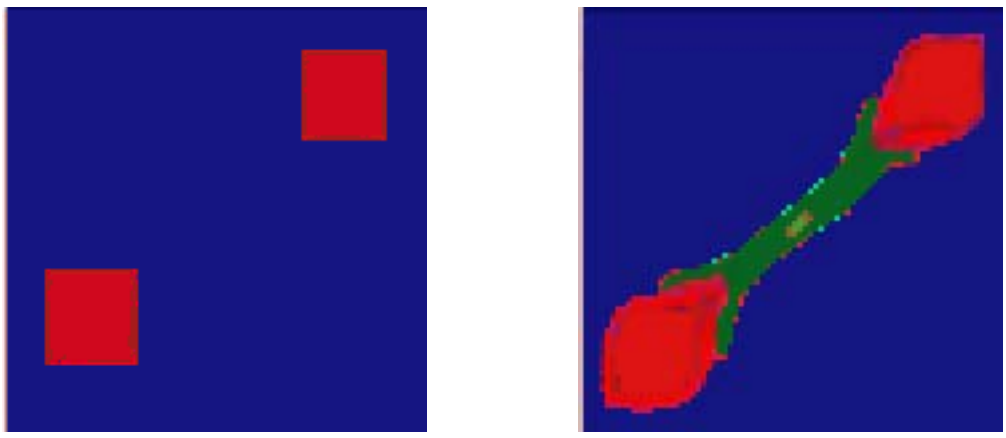
The results from the simulated MCNP data was very promising, especially considering the ability of TNTs to reconstruct the complete image without rotating the object. However, the results from the fifth object reconstructed showed that as the object becomes more heterogeneous, there is a definite need to illuminate the object from multiple sides. To test the ability of TNTs to reconstruct and actual object, an object with properties essentially identical to those used in Reconstruction V was created. This object consisted of a high density polyethylene slab with two boron nitride intrusions similar to those shown in Fig. 59. The object was

illuminated on the bottom (using the thermal neutron beam at the TAMU NSC and images collected on the other three sides. The object was then rotated 90° and illuminated again with images collected on the other three sides. The object was then rotated again by 90° and illuminated again with images collected on the other three sides. The real optical property distribution and the reconstruction results are shown in Figs. 62 and Fig. 63. The use of illumination from three sides instead of one has significantly increased the capability of the system in identifying the location of the intrusions as well as their properties. While the system still has some difficulties in properly reconstructing the interior of the object, it is expected that if more images were taken this limitation would diminish. Also, it should be noted that the optical thickness of this object would make any existing tomographic technique essentially unable to provide any useful reconstruction.



(a) (b)

Fig. 62. Scattering Cross Section from Object Reconstruction. (a) Image of the scattering property of the real object. (b) Image of the scattering property of the reconstructed object.



(a) (b)

Fig. 63. Absorption Cross Section from Object Reconstruction. (a) Image of the absorption property of the real object. (b) Image of the absorption property of the reconstructed object.

5. CONCLUSIONS

An image reconstruction algorithm and code for optical tomography were developed based on a neutron transport inverse problem. This algorithm does not require small perturbations or advanced knowledge of a reference medium for successful execution. TNTs consists of three components:

- A forward model used to predict the detector readings along the surfaces of the object assuming a certain distribution of optical properties. The forward model is based on a SC scheme for the solution of the neutron transport equation with level symmetric discretization for the angular dependence, zero boundary condition and uncollided sources per cell.
- An analysis scheme in which an objective function, defined as the sum of the squared difference between predicted responses and experimental measurements, is defined. The objective function is minimized in order to find the expected optical distribution of the real object.
- An updating scheme, which iteratively changes the optical properties of the medium based on the gradient of the objective function over every optical property. The gradient is calculated with an adjoint calculation.

This work constitutes the first effort to use a forward model based on transport theory rather than diffusion or radiative transfer applied to tomography reconstruction. The forward model has been successfully tested against numerical results obtained with MCNP and from experimental measurements showing excellent agreements. This forward model is more accurate than the models based on diffusion or on radiative transfer in most applications. The generality in the definition of the incoming radiation beam allows TNTs to be used in a great variety of physical situation.

The main accomplishment was the calculation of the gradient of the objective function with respect to the optical properties of the medium for the transport-based forward model. A mathematical proof, via sensitivity equation, was presented to show the effectiveness of the adjoint calculation. The adjoint calculation was performed in a different manner than was done previously by other authors and was tested numerically to prove its reliability. The adjoint calculation, fundamental for this application, should be considered for further future study. The adjoint calculation is capable of finding the gradient quickly in comparison to other techniques such as finite-difference methods or perturbation theory. It can be used in a variety of other application such as homogenization of the cross section in reactor analysis.

The reconstructions of several objects illuminated from only one side were successful. In the case of single intrusion TNTs was always able to detect the intrusion. However the predicted position was slightly different from the real position. It could locate the defect precisely along the x-axis, but less precisely along the y-axis. The reason of this behavior can be related to the fact that SC tends to "move" more forward the particle with respect to the real solution (MCNP in our case). This is a direct effect of the averaging of the angular flux along the edge of a single cell. In the case of the double body object TNTs was able to reconstruct partially the optical distribution. The most important defect, in terms of gradient, was correctly located and reconstructed. Difficulties were discovered in the location and reconstruction of the second defect. Nevertheless, the results are exceptional considering they were obtained by lightening the object from only one side.

TNTs was also used to perform reconstructions for objects illuminated from three sides. Significant improvements in the reconstruction were shown; however, TNTs still tends to locate the intrusion closer to the surface in the reconstruction than it is in the actual object.

There are many possibilities for improvement of the current technique and algorithm including:

- The construction of matrix A and vector v (see Section 3) is performed by storing all the blocks involved in the calculation of the gradient. This is a memory and time consuming process. Since the only matrix dependent on the forward solution through the angular fluxes is matrix B, the operation involving matrix A and vector v should be substitute with functions. Instead of storing all of the information, a function, calculating the same information, can be called whenever necessary. This saves memory and computational time and will allow the process to use a finer grid for the discretization of the domain
- TNTs should be modified to include a full 3-dimensional, time-dependent, and energy discretized forward model to assure generality and to increase the capability of the updating scheme

The method developed here has great potential for significantly advancing the state of tomographic neutron radiography. This method can greatly aid in the surveying of thick, highly-scattering objects. With the addition of the upgrades mentioned above it is expected that this methodology will find application in many fields

REFERENCES

1. J. C. DOMANUS, *Practical Neutron Radiography*, Kluwer Academic, Washington DC (1992).
2. H. BERGER, *Radiography, Methods, Capabilities and Applications*, Elsevier, New York (1965).
3. E. LEHMANN, P. VONTOBEL, "Utilisation of Epithermal Neutrons for Investigations with Strong Neutron Absorbers and Nuclear Fuel Elements," *Paul Scherrer Institute Annual Report, Annex GFA,* Villingen, Switzerland (1998).
4. F. GROSCHEL, "Neutron Radiography of Irradiated Fuel Rod Segments at the SINQ: Loading, Transfer and Irradiation Concept," *Proceedings of the 3rd International Topic Meeting on Neutron Radiography (ITMNR-**3**),* 16-19 March, Lucerne, Switzerland (1998).
5. H.K. JENSSEN, "Neutron Radiography of Irradiated Fuel Rods: An Approach to Improve Spatial Resolution in Neutron Radiographs," *Proceedings of the Plenary Meeting of European Working Group for Hot Laboratories and Remote Handling,(IFE/KR/F-98/159)*, 14 September, Windscale, Norway (1998).
6. R.J. ADRIAN, "Particle-Imaging Techniques for Experimental Fluid Mechanics," *Annual Rev.Fluid Mech,* **23**, 261 (1991).
7. C. SUBBIAH, K. VENKATA RAO, K. VISWANATHAN and S.K. ATHITHAN, "Evaluation of Pyrotechnic Devices by Neutron Radiography," *Trends in NDE Science & Technology; Proceedings of the 14th World Conference on Non-Destructive Testing,* **3**, 1461 (1996).
8. J.P. BARTON, "Contrast Sensitivity in Neutron Radiography" *Appl. Materials Res, **4**,* 90 (1965).
9. J.M. OLLINGER, "Model-based Scatter Correction for Fully 3D PET," *Phys Med Biol,* **41**, 153 (1996).
10. G.T. HERMAN, *Image Reconstruction from Projections,* Academic Press, New York (1980).
11. A.C. KAK, *Principles of Computerized Tomographic Imaging*, IEEE Press, New York (1987).
12. A.H. HIELSCHER, "Gradient Based Iterative Image Reconstruction Scheme for Time-Resolved Optical Tomography," *IEEE Trans. Medical Imaging*, **18**(3), 262 (1999).
13. H. JIANG, K.D. PAULSEN, L. OSTERBERG, "Optical Image Reconstruction Using DC Data: Simulations and Experiments," *J. Opt. Soc. Am. A*, **13**, 253 (1995).
14. A.J. DAVIES, B. CHRISTIANSON, L.C.W. DIXON,, R. ROY, P. VAN DER ZEE, "Reverse Differentiation and the Inverse Diffusion Problem," *Adv. Eng. Software*, **28**, 217 (1997).
15. Z. ZHOU, R.M. LEAHY, J. QI "Approximate maximum likelihood hyperparameter estimation for Gibbs priors," *IEEE Trans. on Image Processing*, **6**, 844 (1997).
16. S.S. SAQUIB, "ML Parameter Estimation for Markov Random Fields with Applications to Bayesian Tomography," *IEEE Trans. Image Processing*, **7**(7), 1029 (1998).
17. A.D. KLOSE, "Optical Tomography Using the Time-Independent Equation of Radiative Transfer – Part 1: Forward Model," *J. Quantitive Spectroscopy and Radiative Transfer*, **72**, 691 (2002).
18. A.D. KLOSE, "Optical Tomography Using the Time-Independent Equation of Radiative Transfer – Part 2: Inverse Model," *J. Quantitive Spectroscopy and Radiative Transfer*, **72**, 715 (2002).
19. N.J. McCORMICK, "Inverse Problems for Biomedical Characterization and Imaging," *3rd Int. Conf. on Inverse Problems in Engineering: Theory and Practice*, Port Ludlow, Washington, June 13-18 (1999).
20. J.J. DUDERSTADT and L.J. HAMILTON, *Nuclear Reactor Analysis*, John Wiley & Sons, New York (1976).

21. R. SANCHEZ and N.J. McCORMICK, “A Review of Neutron Transport Approximations,” *Nucl. Sci. Eng.*, **80**, 481 (1982).
22. W.M. STACEY, *Nuclear Reactor Physics*, John Wiley and Sons, New York (2001).
23. J.K. SHULTIS and R.E. FAW, *Radiation Shielding,* American Nuclear Society, Chicago, IL (2000).
24. E.E. LEWIS and W.F. MILLER JR., *Computational Methods of Neutron Transport*, John Wiley & Sons, New York (1984).
25. K.D. LATHROP, “Spatial Differencing of the Transport Equation: Positivity vs. Accuracy,” *J. Comput. Phys.*, **4**, 475 (1969).
26. Y. WANG and Y. BAYAZITOGLU, “Wavelets and the Discrete Ordinate Method for Radiative Heat Transfer in a Two-Dimensional Rectangular Enclosure with a Nongray Medium,” *Journal of Heat Transfer*, **124**(5), 971 (2002).
27. K.D. LATHROP and B.G. CARLSON, “Discrete Ordinates Angular Quadrature of the Neutron Transport Equation,” *LA-3186*, Los Alamos Scientific Laboratory (1965).
28. K.D. LATHROP and B.G. CARLSON, “Transport Theory -- The Method of Discrete Ordinates,” *Computing Methods in Reactor Physics*, H. Greenspan, C. N. Kelber, and D. Okrent (eds.), Gordon and Breach, New York (1968).
29. S.R. ARRIDGE, “The Forward and Inverse Problems in Time-Resolved Infrared Imaging,” *Proc. SPIE-Int. Soc. Opt. Eng.*, **11**, 35 (1993).
30. S.R. ARRIDGE and M. SCWEIGER, “Photon Measurement Density Function. Part I: Analytical Forms,” *Appl. Opt.,* **34**, 7395 (1995).
31. S.R. ARRIDGE and M. SCHWEIGER, “Photon Measurement Density Function. Part II: Finite Element results,” *Appl. Opt.,* **34**, 8026 (1995).
32. S.R. ARRIDGE and M. SCHWEIGER, “A System for Solving the Forward and Inverse Problem in Optical Spectroscopy and Imaging,” *Advances in Optical Imaging and Photon Migrations* (OSA Trends in Optics and Photonic Series) Optical Society of America, **2**, 263 (1996).
33. R.L. BARBOUR and H.L. GRABER, “A Perturbation Approach for Optical Diffusion Tomography Using Continuous-Wave and Time-Resolved Data,” *Proc. SPIE-Int. Soc.Opt. Eng.,* **11**, 87 (1993).
34. H.L. GRABER, J.H. CHANG, R. ARONSON and R.L. BARBOUR, “A Perturbation Model for Imaging in Dense Scattering Media: Derivation and Evaluation of Imaging Operators,” *Proc. SPIE-Int. Soc. Opt. Eng.,* **11**, 121 (1993).
35. A.H. HIELSCHER, A.D. KLOSE and K.M. HANSON, “Gradient-Baesd Iterative Image Reconstruction Scheme for Time-Resolved Optical Tomography,” *IEEE Transaction on Medical Imaging,* **18**, 262 (1999).
36. F.S. ACTON, *Numerical Methods That Work,* Mathematical Association of America, Washington DC (1990).
37. W.H. PRESS, B.P. FLANNERY, S.A. TEUKOLSKY, and W.T. VETTERLING, *Numerical Recipes (Fortran)*, Cambridge University Press, New York, 276 (1989).
38. E. POLACK *Optimization: Algorithms and Consistent Approximations*, Springer-Verlag, New York (1997).
39. G. ARFKEN, “Gram-Schmidt Orthogonalization," *Mathematical Methods for Physicists,* 3rd ed., Academic Press, Orlando, FL, 516 (1985).