Karlsruher Institut für Technologie

Numerical Methods and Machine Learning for Prompt-Gamma Neutron Activation Analysis

Zur Erlangung des akademischen Grades eines

DOKTORS DER NATURWISSENSCHAFTEN

von der KIT-Fakultät für Mathematik des
Karlsruher Instituts für Technologie (KIT)
genehmigte

DISSERTATION

von

Alexander Jesser, M.Sc.

Tag der mündlichen Prüfung:	23.04.2024
Erstgutachter:	Prof. Dr. Martin Frank
Zweitgutachter:	Prof. Dr. Christian Wieners
Drittgutachter:	Prof. Ryan G. McClarren, PhD

Ich versichere wahrheitsgemäß, die Arbeit selbstständig angefertigt, alle benutzten Hilfsmittel vollständig und genau angegeben und alles kenntlich gemacht zu haben, was aus Arbeiten anderer unverändert oder mit Abänderungen entnommen wurde.

I declare that I have developed and written the enclosed thesis completely by myself, and have not used sources or means without declaration in the text.

. . . .
Alexander Jesser
Karlsruhe, 26.02.2025

Acknowledgements

First and foremost, I would like to thank my advisor Martin Frank for the opportunity and freedom to work on this interesting research topic. I also thank him for his good and patient supervision.

I would also like to thank Christian Wieners for being my second referee. His valuable advice helped me to improve this thesis.

Further, I am thankful to Ryan McClarren for being my third referee. I would also like to thank him for the good cooperation over the last two years.

I want to express my sincere gratitude to all colleagues at KIT and AiNT for creating a pleasent working atmosphere. In particular, I would like to thank Christopher Helmes for years of fruitful collaboration and the late Andreas Havenith for his mentoring.

My special thanks goes to Kai Krycki. His guidance and supervision was essential for completing this thesis.

Finally, I would like to thank my family for their unwavering support. I am especially grateful to my mother for her continuous support during all phases of my studies and enabling my studies in the first place.

Preface

Prompt-gamma neutron activation analysis (PGNAA) is a radio-analytical measurement technique that allows for a non-destructive elemental analysis of a sample. For this purpose, the sample is irradiated with neutrons which are absorbed by atomic nuclei within the sample. The atomic nuclei are in an excited state after the absorption. During the transition into a stable ground state, gamma rays with nuclide specific energies are emitted, which can be used for the analysis of the sample. PGNAA has applications in many fields, including chemistry, the material sciences, geology and archeology [4]. While PGNAA is usually used to analyze small samples at research reactors [73], the interest to also apply this method to larger samples has grown for years. A promising application is the final deposition of radioactive waste. In order to ensure a safe and sustainable storage of the waste, it is demanded in Germany that waste acceptance criteria regarding the material (elemental) composition of the waste need to be met. For this purpose, various hazardous substances need to be quantified beforehand. So far, also no metrological solution is applied as a standard, requiring a recourse visual inspection by opening waste drums. This procedure leads to the exposure of personal to radiation and potentially hazardous substances while also contaminating working material, thereby increasing the total amount of waste that needs to be deposited. To overcome this situation, a measurement device called QUANTOM [22, 23] has been developed that is based on large sample PGNAA. It is capable of performing a non-destructive analysis of elemental masses of radioactive wastes in 200-l-drums.

In order to evaluate large sample PGNAA measurements, it is necessary to improve the mathematical methods currently used in the field. Especially solving the neutron transport equation within an acceptable time frame is of major importance. The governing equation for neutrons is the linear Boltzmann equation:

$$\begin{aligned}&\Omega \cdot \nabla \Psi(x,\Omega,E) + \Sigma_t(x,E)\Psi(x,\Omega,E)\\ &= \int_{\mathbb{S}^2}\int_{E_{min}}^{E_{max}} \Sigma_s(x,\Omega',\Omega,E',E)\Psi(x,\Omega',E')dE'd\Omega' + Q(x,\Omega,E),\end{aligned} \tag{0.1}$$

where the neutron flux $\Psi(x,\Omega,E)$ is usually interpreted as the number of particles at position x with energy E traveling in direction Ω. Here, $Q(x,\Omega,E)$ is the external source while the so-called cross sections Σ_t and Σ_s are material-dependent.

In the following, the novelty of this research and its importance for PGNAA applications is outlined. First, we will use so-called deterministic neutron transport methods and describe the development of SPARC, a solver based on this methods which is specifically developed for the eveluation of PGNAA measurements. Within the large sample PGNAA evaluation process, deterministic calculations offer immense improvements in computation time. Then, based on the deterministic model it is possible to formulate the complete evaluation procedure as a pde-constrained optimization problem. This formulation allows for an investigation of the uniqueness of the solution, which is, to the authors knowledge, the first one for PGNAA problems. Finally, we will investigate Physics Informed Neural Networks (PINNs) as another potential improvement for the neutron transport computations. The neutron transport in large parts of PGNAA measurement facilities takes place in the so-called diffusive regime, so we investigate PINNs in this specific case. We analyze the shortcomings of the methods and introduce a diffusive scaling to overcome these problems.

This thesis is structured as follows. In Chapter 1, we discuss the basic principles of PGNAA and its application on large samples. In Chapter 2 we discuss neutron transport, especially the linear Boltzmann equation (Eq. 0.1) and its boundary conditions. So-called deterministic methods to solve the neutron transport equation are the topic of Chapter 3. We introduce the most commonly used discretization methods. This includes the multigroup approximation and the $\mathrm{SP_N}$ approximation of the neutron transport equation (Eq. 0.1), which are the main methods we use throughout this thesis. Chapter 4 describes the development of SPARC, a neutron transport solver based on the $\mathrm{SP_N}$ approximation. The implementation and verification of SPARC as well as the methods used for multigroup neutron transport are discussed in this Chapter. As described in Chapter 5, the evaluation of the QUANTOM measurements relies on the calculation of so-called partial (n, γ)-cross sections. In order to compute these, neutron transport simulations of the entire measurement facility are necessary. However, the Monte Carlo codes previously used to compute the neutron flux offer insufficient performance for this application. This Chapter shows that the deterministic neutron transport solver SPARC can be used to compute the neutron fluxes with sufficient accuracy while offering a performance that is significantly better than that of the Monte Carlo Code OpenMC [85]. For this purpose, numerical results for four different reference materials are presented. The extension of SPARC to also enable adjoint neutron transport simulations is the topic of Chapter 6. In Chapter 7, we consider the evaluation of QUANTOM measurements as a parameter estimation problem. Unlike most PGNAA measurement facilities, QUANTOM not only contains gamma detectors, but incorporates a large number of neutron detectors, offering additional information for elemental composition reconstruction. We introduce a formulation of the problem that incorporates the different detectors and employ an adjoint-based method to determine the elemental composition. Uniqueness of solutions in this context is studied, and the results offer novel insights for PGNAA applications. While the performance of neutron transport simulations for large sample PGNAA can be significantly improved by the use of deterministic solvers like SPARC, neutron transport simulations remain a bottleneck of the entire evaluation workflow. Therefore, in Chapter 8, we investigate an alternative approach for neutron transport simulations, PINNs [79] with ReLU activation functions. PINNs provide a method to solve partial differential equations (PDE) using neural networks. We study the analogy between PINNs and Least-Squares Finite Elements (LSFE) which lies in the shared approach to reformulate the PDE solution as a minimization of a quadratic functional. We prove that in the diffusive regime, the correct limit is not reached, and introduce a diffusive scaling that can be applied to overcome this issue. These results are in agreement with known results for first-order LSFE. Numerical results are provided that support these theoretical findings. This thesis is concluded in Chapter 9 with a discussion, conclusion and outlook.

Statement of Authorship

This thesis partially relies on previously published work. In this paragraph, the contributions of the author will be clarified. The author of this thesis was the first and corresponding author of the publications mentioned in this paragraph. The main parts of Chapter 5 have been published in [48] together with Kai Krycki and Martin Frank. The author carried out the numerical simulations. Section 6.4 shows a previously unpublished benchmark of the adjoint solver integrated in SPARC and the adjoint solver integrated in ARTEMIS. The ARTEMIS computations were carried out by René van Geemert, while the author carried out the computations with SPARC. The author also implemented the adjoint solver in SPARC. Chapter 7 is based on [49], a paper published together with Kai Krycki and Martin Frank. The approach for the investigation of uniqueness originates in a discussion of Martin Frank with the author. The author was also responsible for the implementation of the necessary computer codes and the numerical simulations. Chapter 8 is based on [50], published together with Kai Krycki, Ryan G. McClarren and Martin Frank. The main idea, to study the analogy between PINNs and LSFE, is due to Martin Frank. The author was responsible for the implementation of all necessary computer codes as well as the numerical simulations. The analysis in Sec. 8.4 was also carried out by the author.

Contents

Appendices

List of Figures

List of Tables

1. Prompt-Gamma Neutron Activation Analysis

While many radio-analytical techniques have lost importance, prompt-gamma neutron activation analysis (PGNAA) has been rapidly developing in the past decades. It has many applications, e.g. in chemistry, the material sciences, geology and archeology [4]. This methods allows for the non-destructive elemental analysis of a sample based a radiative neutron capture process. PGNAA allows for both a qualitative and quantitative analysis of the elemental composition of a sample. The most comprehensive overview of Prompt-Gamma Neutron Activation Analysis can be found in [73], while [82] provides a more concise review of the PGNAA methodology. Due to the reduced availability of many large-scale neutron sources required for PGNAA, an important topic of recent research is the improvement of the analysis workflow [94].
This chapter is structured as follows. In Sec. 1.1, the basic principles of PGNAA are introduced. In Sec. 1.2 photopeak efficiencies are discussed. The chapter is concluded by an introduction into large sample PGNAA in Sec. 1.3.

1.1. Basic Principles

The nuclear reaction underlying PGNAA can be described as follows. A sample is continously irradiated with neutrons. The atomic nuclei within the sample absorb neutrons and form a compound nuclei. The compound nuclei are then in an excited state. During the transition into the stable ground state these nuclei emmit gamma rays with energies that are characteristic for the absorbing isotope. This process is depicted in Fig. 1.1. While in PGNAA the transition into the ground state typically occurs within a picosecond or even less time, in delayed-gamma neutron activation analysis (DGNAA) far more time is required. The compound nuclei are radioactive and decay with a half life that can be even several minutes. In the following, we will only investigate PGNAA cases since prompt gamma rays have a far higher intensity than delayed gamma rays which makes the former ones far more suitable for a quantitative analysis. While PGNAA in principle allows for a reconstruction of the quantities of specific nuclides, usually only an elemental analysis is desired. In the following, we will thus assume that all nuclides are present with their natural abundance and discuss PGNAA only as a method for elemental analysis.

Fig. 1.2 depicts an abstract sketch of a PGNAA measurement. A neutron source, e.g. a neutron generator, emits neutrons that are absorbed within the sample. The excited nuclei in the sample then emit gamma rays with energies characteristic for each element. The gamma detector then measures the gamma spectrum.

A cutout of an exemplary gamma spectrum measured with the MEDINA measurement facility [69] at the research center Jülich (FZJ) is depicted in Fig. 1.3. While there are peaks that contribute only to the background, e.g. escape-peaks, most peaks can be matched with the peaks

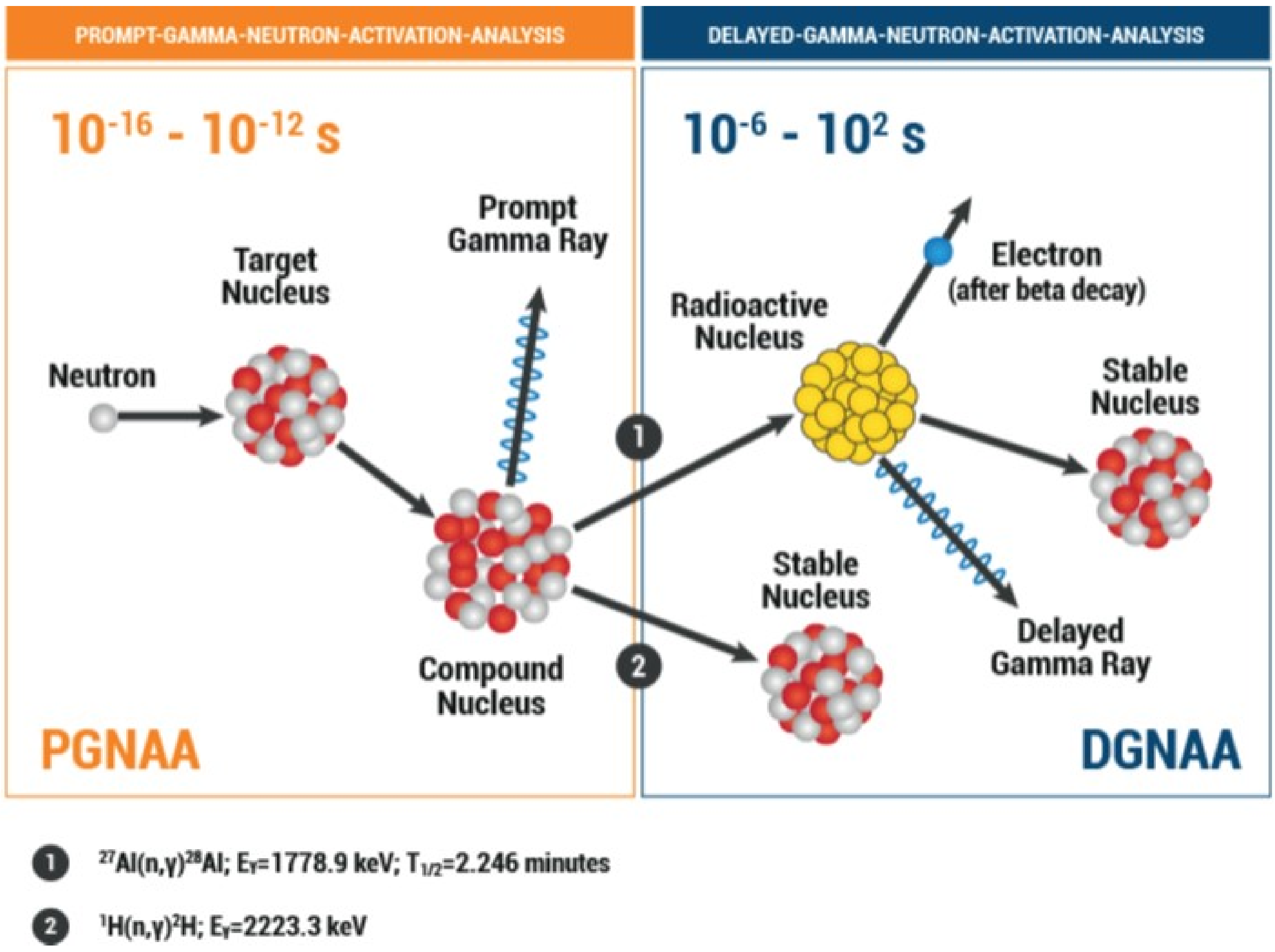


Figure 1.1.: Prompt and Delayed Gamma Neutron Activation Analysis. Neutrons are absorbed by target nuclei and induce the emission of gamma rays with energies characteristic for the target nucleus. While the compound nuclei emits prompt gamma rays within typically less then one picosecond, in the delayed cased the radioactive nucleus might have a life in the order of minutes.

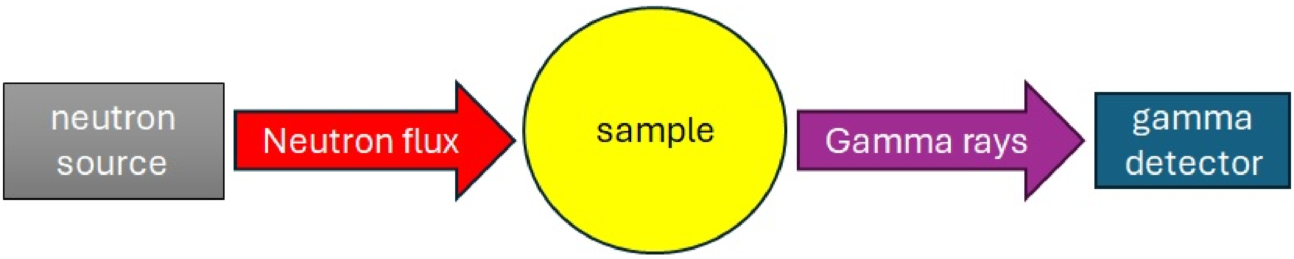


Figure 1.2.: Abstract sketch of a PGNAA measurement. Neutrons emitted by a source are absorbed within the sample. The sample then emits gamma rays with energies characteristic for each element. The gamma detector then measures the energy and quantity of the incoming gamma rays.

characteristic for specific elements. Said characteristic peaks can be obtained from a database [21]. This allows for a qualitative analysis of the composition of the sample. For a quantitative analysis more information is needed. First, it is necessary to estimate the background and subtract it form the spectrum. Then, for each characteristic peak the remaining area under the peaks needs to be determined, since these areas are identical to the peak count rates for said peak.

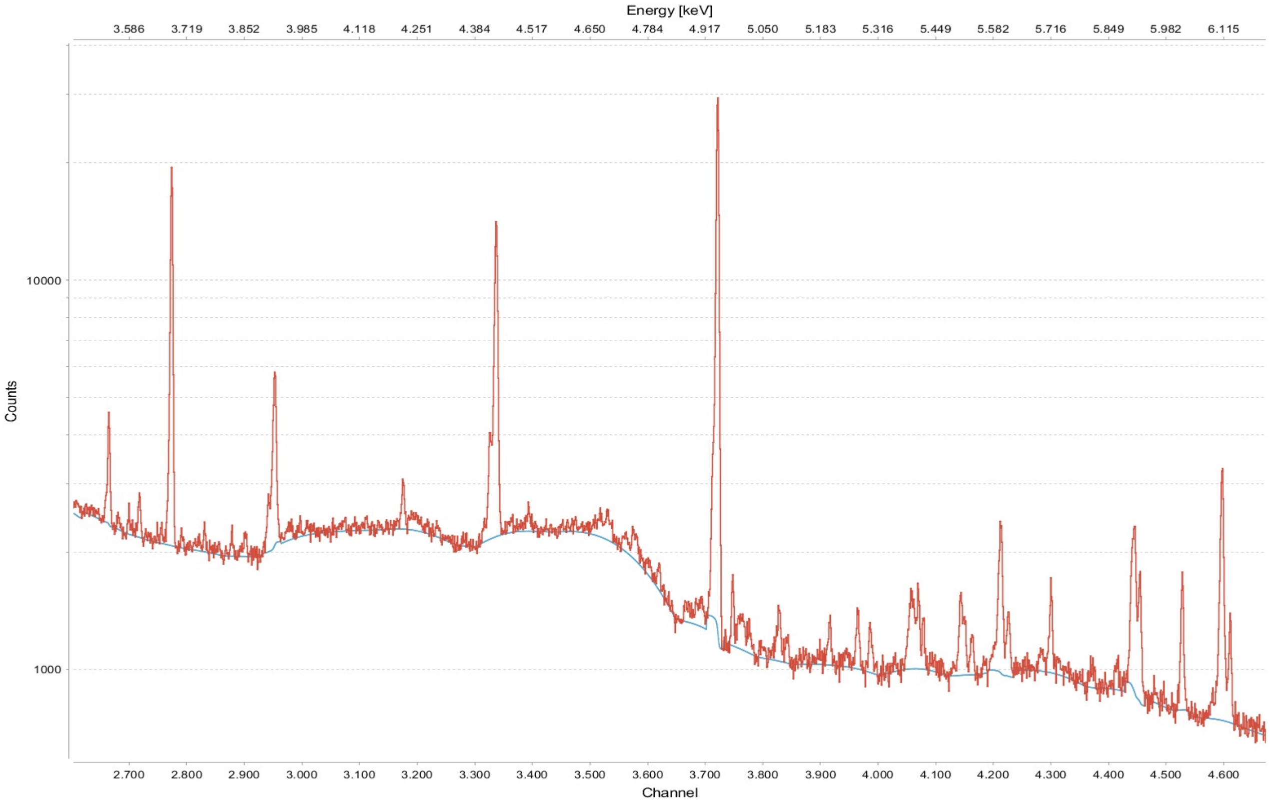


Figure 1.3.: Cutout of a gamma spectrum measured with the MEDINA measurement facility at the research center Jülich (FZJ). The estimation of the background is depicted in blue.

The material reconstruction procedure then aims at finding an elemental composition of the sample that reproduces the correct peak count rates with as little deviation as possible. Details of this procedure as well as an alternative approach will be discussed in chapter 7. For now, we will provide the underlying physical model for the peak counting rates. The most general expression for the photopeak count rate is

$$P_{E_{\gamma,l}} = \frac{N_A}{M_l} \int_0^{\infty} \int_{V_{sam}} \rho_l(x) \epsilon_{E_{\gamma,l}}(x) \sigma_{E_{\gamma,l}}(E) \phi(E,x) dx dE, \tag{1.1}$$

where $\rho_l(x)$ is the mass density of the examined element, $\epsilon_{E_{\gamma,l}}$ is the so-called photopeak efficiency, $\sigma_{E_{\gamma,l}}$ is the so-called partial (n,γ)-cross section, $\phi(E,x)$ is the neutron flux within the sample, with the Avogadro constant N_A and the molar mass M_l. The sample domain is denoted as $V_{sam} \subset \mathbb{R}^3$. In this thesis, we denote the neutron energy as E and the gamma ray energy as $E_{\gamma,l}$, where the index l denotes the element l to which the peak can be assigned. While the energy-dependent partial (n,γ)-cross sections are tabulated, both the photopeak efficiency and the neutron flux need to be computed. We introduce a nodel for the photopeak efficiency in the next section, while the neutron flux $\phi(E,x)$ and its computation are discussed in detail in the following chapters. It is important to note that both the photopeak efficiency and the neutron flux depend on the material composition in the facility and the sample. Therefore, both quantities depend on $\rho_l(x)$, necessitating an iterative evaluation process (see Fig. 1.5).

1.2. Photopeak Efficiency

The photopeak efficiency $\epsilon_{E_{\gamma,l}}$ is defined as the fraction of gammas which are detected at full energy. It is important to note that only unscattered photons that create a signal in the detector allow for the identification of elements. Since photons lose energy during the scattering, scattered photons do not have energies characteristic for specific nuclides anymore and contribute only to the background. The photopeak efficiency is given by

$$\epsilon_{E_{\gamma,l}}(x) = \epsilon_{geom,E_{\gamma,l}}(x) \cdot \epsilon_{att,E_{\gamma,l}}(x) \cdot \epsilon_{det,E_{\gamma,l}}, \tag{1.2}$$

where $\epsilon_{att,E_{\gamma,l}}(x)$ is the photon attenuation, $\epsilon_{geom,E_{\gamma,l}}(x)$ is the geometric efficiency and $\epsilon_{det,E_{\gamma,l}}$ is the detector efficiency. The photopeak efficiency is dependent of the location x where the photon is emitted. The three constituents of the photopeak efficiency can be interpreted as follows. The geometric efficiency $\epsilon_{geom,E_{\gamma,l}}(x)$ is the probability that a photon generated at a point x within the sample is emitted into a direction that allows for reaching the detector without scattering. The photon attenuation $\epsilon_{att,E_{\gamma,l}}(x)$ is dependent on the materials the photons traverse and their path length within these materials and is a measure for the decrease in gamma ray intensity along this path. The detector efficiency $\epsilon_{det,E_{\gamma,l}}$ is the probability for a photon that reaches the detector to actually induce a signal. It is mainly determined by detector physical effects. Since there exists no closed analytical expression for the full photopeak efficiency, we will introduce a simplistic model for the photopeak efficiencies, which is sufficient for the purpose of this thesis. It should be noted that all results remain valid when using experimentally obtained calibration curves, more sophisticated analytical models, or calibration curves based on extensive Monte Carlo simulations. The geometric efficiency in this setup is given by

$$\epsilon_{geom,E_{\gamma,l}}(x) = \frac{1}{\pi} \arctan\left(\frac{r_{det}(x)}{d_{det}(x)}\right), \tag{1.3}$$

where $r_{det}(x)$ is the radius of the detector and $d_{det}(x)$ the distance to the center of the detector. The photon attenuation is given by

$$\epsilon_{att,E_{\gamma,l}}(x) = e^{-\mu \cdot d(x)}, \tag{1.4}$$

where μ is the attenuation coefficient and d the traversed distance. Hereby the attenuation coefficient is mainly determined by the average density of the material, which can be easily determined in practice. The detector efficiency $\epsilon_{det,E_{\gamma,l}}$ can be interpreted as the probability for a photon entering the detector to contribute to the measured signal and has to be determined by photon transport simulations. For a given energy and detector, it is a constant. For simplicity it is set to one in the following.

1.3. Large Sample PGNAA

While PGNAA is usually used to analyze small samples [73], the interest to also apply this method to larger samples has grown for years. A promising application is the final deposition of radioactive waste. For the disposal of low level (LLW) and intermediate level (ILW) radioactive waste in Germany a final deep geological repository called Konrad is currently expected to go into operation around the year 2030. In the process of disposal of nuclear wastes in Konrad one of the major challenges is the compliance with waste acceptance criteria in order to assure a safe and sustainable storage of the waste, and especially the proof of compliance. Radiological waste acceptance criteria can be verified using passive measurements of the gamma emitting radionuclide inventory of waste drums [29, 30] and correlation of the remaining radionuclides. However, the waste acceptance criteria regarding the material (elemental) composition of the waste are far more difficult to meet. These requirements were defined on the basis of the results of a site-specific safety assessment and necessitate the quantification of various hazardous substances (e.g. mercury, cadmium, copper or arsenic) [12], leading to a comprehensive documentation of the material composition of the waste [45]. The existing documentation is usually not sufficient, especially for so-called historic waste [42]. So far, also no metrological solution is applied as a standard, requiring a recourse visual inspection by opening waste drums. This procedure leads to the exposure of personal to radiation and potentially hazardous substances. It also contaminates working material, thereby increasing the total amount of waste that needs to be deposited.

To overcome this situation, a measurement device called QUANTOM [22, 23] has been developed that is capable of performing a non-destructive analysis of elemental masses of radioactive wastes in 200-l-drums. The determination of the elemental composition of radioactive waste drums with QUANTOM is based on large sample PGNAA. Free neutrons are produced by a built-in neutron generator, which is in fact a small linear accelerator where Deuterium is accelerated to a Deuterium target. The resulting neutrons neutrons have an energy of 2.45 MeV. Since the (n, γ) reaction rates are far higher for lower energies, the neutrons need to be moderated, that means their speed and thereby energy needs to be reduced. This is achieved by their interaction with the surrounding material, a massive graphite layer.

Because of the large volume samples under consideration in QUANTOM, a spatially resolved measurement is carried out corresponding to a complete surface scan. This is achieved by lifting the drum into 4 different positions and rotating the drum so that it can be measured from 6 different angles for each position. With two gamma detectors, a total of 48 gamma spectra is recorded. The gamma spectra are evaluated jointly based on a model of evaluation which accounts for the different measurement positions and angles [23]. In order to quantitatively evaluate the neutron flux in the measurement chamber, additional neutron detectors are included in QUANTOM. Gamma spectra are recorded using high purity Germanium detectors (HPGe), while the total neutron flux is measured using 3He counter tubes, as shown in Fig. 1.4.

It is important to note that the standard use of PGNAA comprises the analysis of small samples at research reactors [62]. In the presence of a strong, mono-energetic neutron flux and a small sample size, a constant neutron spectrum can be assumed in the sample. Then the model of evaluation reduces to a simple proportionality of element mass m_l (index l indicating a specific element) and net photopeak count rate $P_{E_{\gamma,l}}$ which is obtained by evaluating the gamma spectrum. Eq. (1.1) then effectively reduces to

$$P_{E_{\gamma,l}} = \frac{N_A}{M_l} \cdot m_l \cdot \epsilon_{E_{\gamma,l}} \cdot \sigma_{E_{\gamma,l}} \cdot \Phi_{tot}. \tag{1.5}$$

with the total neutron flux

$$\Phi_{tot} = \int_0^\infty \int_{V_{sam}} \phi(E, x) dx dE. \tag{1.6}$$

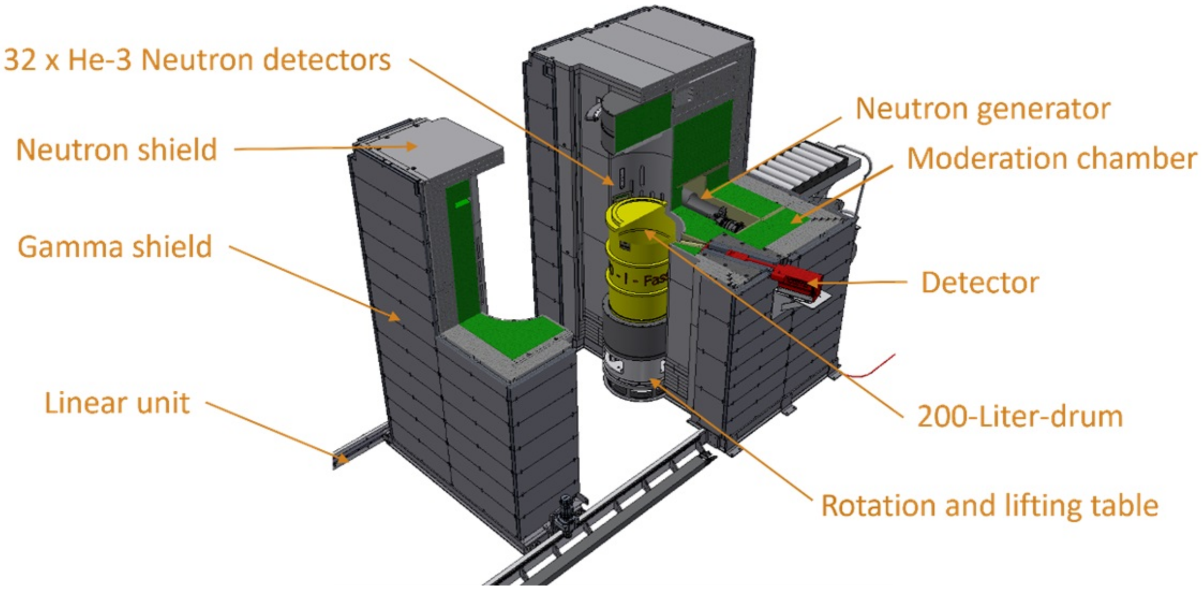


Figure 1.4.: Cut-view of the QUANTOM measurement device (taken from [23]).

Hence, element masses can be obtained by inverting the proportionality.
For large samples, aforementioned assumptions are not met, leading to a much more complicated evaluation process. In these complex cases, where the dependence of the neutron flux and the photopeak efficiency on the elemental composition can not be neglected, PGNAA evaluations are solved using an iterative formulation. Hereby, the neutron flux and photopeak efficiency for a given elemental composition are computed and, using Eq. (1.5), the updated elemental composition is determined. The element masses in the $(n+1)-$th iteration are thus computed as follows:

$$(m_l)_{n+1} = \frac{P_{E_{\gamma,l}} M_l}{N_A \cdot \epsilon_{E_{\gamma,l}}[(m_1)_n, ..., (m_L)_n)] \cdot \sigma_{E_{\gamma,l}} \cdot \Phi_{tot}[(m_1)_n, ..., (m_L)_n)]}, \tag{1.7}$$

where L is the total number of elements. The iterative process terminates when the relative change of compositions between two iterations becomes negligible. The above described workflow for the evaluation of large sample PGNAA measurements is depicted in Fig. 1.5.

In [46, 2] this approach has been studied for the evaluation of one single PGNAA measurement using a HPGe detector. Under mild assumptions and based on a relative formulation, the authors were able to show the existence of a solution for the PGNAA setting described in [46, 2]. Nevertheless, the uniqueness of a solution cannot be guaranteed in this setting, although, in accompanying experimental studies ambiguous solutions did not occur. The setup in QUANTOM differs in several aspects from the one considered in [46, 2]. The sample volume in QUANTOM is larger than considered in standard literature before, hence spatial distribution of neutron flux and photopeak efficiencies need to be carefully represented in the simulation. Measurements for the neutron flux in QUANTOM present additional information regarding the reconstruction of elemental compositions. Furthermore, the problem formulation of [2], where relative compositions are determined, cannot be applied to the more complex measurement task of QUANTOM. The model of evaluation in this work therefore aims at the determination of absolute quantities.

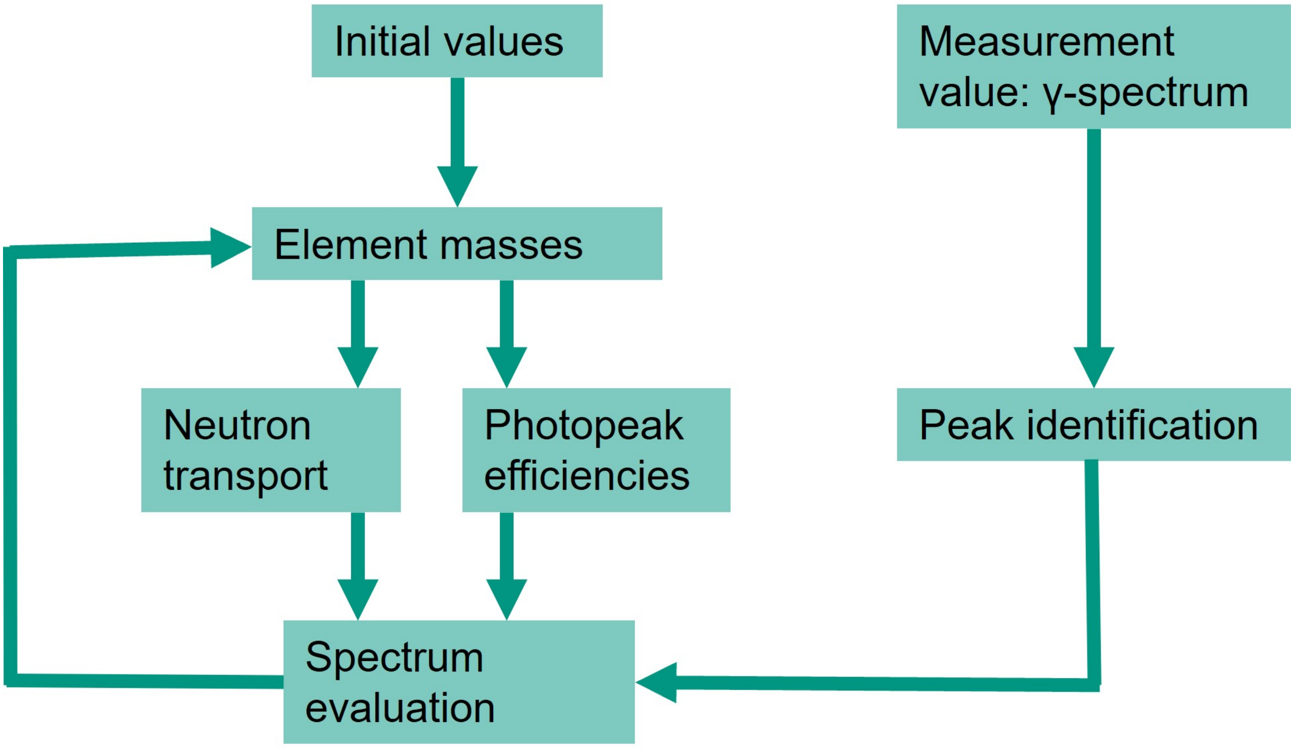


Figure 1.5.: Abstract sketch of a large sample PGNAA evaluation workflow. Starting form initial values for the element masses, the neutron flux and photopeak efficiences are computed. The evaluation of the measured γ-spectrum consists of the identification of the peaks and the determination of the associated peak counting rates. The counting rates are then used together with the neutron flux and photopeak efficiences to update the element masses according to Eq. (1.7).

2. Neutron Transport

In this chapter, the basics of neutron transport will be discussed. In Sec. 2.1, the linear Boltzmann equation, the governing equation for neutron transport problems, is introduced. In Sec. 2.2 we will discuss the most common boundary conditions. In Sec. 2.3, we describe criticality computations, since they belong to the most important neutron transport applications. We conclude this chapter with a brief introduction into the most important neutron particle interactions in Sec. 2.4.

2.1. Linear Boltzmann Equation

The equation for the transport of neutrons through a (in general) inhomogeneous medium is the linear Boltzmann equation, a partial differential equation for the neutron flux Ψ. The steady-state neutron transport equation without fission and an with external source term $Q(x, \Omega, E)$ is given by

$$\begin{aligned}&\Omega \cdot \nabla \Psi(x, \Omega, E) + \Sigma_t(x, E)\Psi(x, \Omega, E)\\&= \int_{\mathbb{S}^2} \int_{E_{min}}^{E_{max}} \Sigma_s(x, \Omega', \Omega, E', E)\Psi(x, \Omega', E')dE'd\Omega' + Q(x, \Omega, E).\end{aligned} \tag{2.1}$$

The neutron flux $\Psi(x, \Omega, E)$ is usually interpreted as the number of particles at position x with energy E traveling in direction Ω. The phase space of the neutron flux Ψ consists of the space $x \in D$ where $D \subset \mathbb{R}^3$ is a bounded domain, the energy $E \in [E_{min}, E_{max}] \subset \mathbb{R}_+$ and the direction of flight $\Omega \in \mathbb{S}^2$. We use the convention

$$\Omega = \begin{pmatrix} \sqrt{1-\mu^2}\cos(\varphi) \\ \sqrt{1-\mu^2}\sin(\varphi) \\ \mu \end{pmatrix} \tag{2.2}$$

with the azimuthal angle $\varphi \in [0, 2\pi]$ and the polar angle $\theta \in [0, \pi]$. We define $\mu = \cos(\theta)$ and therefore $\mu \in [-1, 1]$. Here Σ_t is the total macroscopic cross-section, which can be determined by

$$\Sigma_t(x, E) = \int_{E_{min}}^{E_{max}} \int_{\mathbb{S}^2} \int_{\mathbb{S}^2} \Sigma_s(x, \Omega', \Omega, E', E)d\Omega d\Omega' dE' + \Sigma_a(x, E) \tag{2.3}$$

where Σ_a is the absorption cross section. Σ_s is the double-differential scattering-cross section. It is a measure for the probability of a particle scattering from direction Ω' and energy E' into direction Ω and energy E after the collision. The cross sections are material parameters and will be discussed in in more detail in Subsec. 3.1.2.

Eq. (2.1) is a balance equation. Its left hand side describes the loss of particles: the first term is the advection term and describes the free flight of particles through the domain, leading to them streaming out of the state (x, Ω, E). The second term describes the loss of particles through absorption and out-scattering, indicated by Eq. (2.3). The right hand side of Eq. (2.1) describes the gain of particles. The first term describes particles entering the state (x, Ω, E) through in-scattering, while the second term is an external source term.

2.2. Boundary Conditions and Uniqueness

While partial differential equations equations describe how a system involves due to physical laws, boundary conditions describe which system we are considering exactly. Therefore, we need to equip the linear Boltzmann equation (Eq. 2.1) with suitable boundary conditions. For the linear Boltzmann equation a variety of different boundary conditions is used, depending on the application. In the following, we introduce the boundary conditions most often used for neutron transport. Let $D \subset \mathbb{R}^3$ be a bounded domain with outer boundary ∂D and $n(x)$ the outer normal vector. Then we have:

- *Explicit boundary conditions:* A known function describes the incoming flux. We denote the incoming neutron flux as Ψ^-. The boundary conditions for the linear Boltzmann equation then read:

$$\Psi(x, \Omega, E) = \Psi^-(x, \Omega, E) \text{ for } \Omega \cdot n(x) < 0,\ x \in \partial D. \tag{2.4}$$

- *Vacuum boundary conditions:* A special case are vacuum boundary conditions, where incoming flux is zero:

$$\Psi(x, \Omega, E) = 0 \text{ for } \Omega \cdot n(x) < 0,\ x \in \partial D. \tag{2.5}$$

- *Albedo boundary conditions:* Let Ω' be the direction of the outgoing flux and $\alpha \in \mathbb{R}^{\geq 0}$. Then the incoming neutron flux is equal to the outgoing neutron flux times an known albedo coefficient α. The so-called albedo boundary conditions are given by

$$\Psi(x, \Omega, E) = \alpha \Psi(x, \Omega', E) \text{ for } \Omega \cdot n(x) < 0,\ x \in \partial D,\ \Omega \cdot n(x) = -\Omega' \cdot n(x). \tag{2.6}$$

- *Specular reflective boundary conditions:* Here all outgoing particles are reflected back like in a perfect mirror. Specular reflective boundary conditions are a special case of albedo boundary conditions and read

$$\Psi(x, \Omega, E) = \Psi(x, \Omega', E) \text{ for } \Omega \cdot n(x) < 0,\ x \in \partial D,\ \Omega \cdot n(x) = -\Omega' \cdot n(x). \tag{2.7}$$

- *White boundary conditions:* Here all neutrons leaving the system through the boundary are emitted back into the domain with an isotropic distribution in angle. We symbolically integrate over all $\Omega' \in \mathbb{S}^2$ with $\Omega' \cdot n(x) > 0$, so that the white boundary conditions read:

$$\Psi(x, \Omega, E) = \frac{\int_{\Omega' \cdot n(x) > 0} |\Omega' \cdot n(x)| \Psi(x, \Omega', E) d\Omega'}{\int_{\Omega' \cdot n(x) > 0} |\Omega' \cdot n(x)| d\Omega'} \text{ for } \Omega \cdot n(x) < 0,\ x \in \partial D \tag{2.8}$$

In this thesis, mostly vacuum boundary conditions will be used. The boundary conditions above specify the incoming flux and we have the following theorem that ensures the uniqueness of the solution of the linear Boltzmann equation.

Theorem 2.2.1. *Let $D \subset \mathbb{R}^3$ be a bounded domain with boundary ∂D and $\Gamma^- = \{(x,\Omega) \in \partial D \times \mathbb{S}^2 : \Omega \cdot n(x) < 0\}$. Let the following assumptions be met:*

- *Let $\Sigma_t \in L^\infty(D \times [E_{min}, E_{max}]])$ be non-negative.*
- *Let Σ_s be non-negative. There exist constants $M_a, M_b > 0$ such that:*

$$\int_{\mathbb{S}^2}\int_{E_{min}}^{E_{max}} \Sigma_s(x,\Omega',\Omega,E',E)dE'd\Omega' < M_a$$
$$\int_{\mathbb{S}^2}\int_{E_{min}}^{E_{max}} \Sigma_s(x,\Omega',\Omega,E',E)dEd\Omega < M_b.$$

- *There exists a $\Sigma_a \in L^\infty(D \times [E_{min}, E_{max}]])$ with $\Sigma_a > 0$ so that Σ_t and Σ_s fulfill Eq. (2.3).*
- *Let $Q \in L^\infty(D \times \mathbb{S}^2 \times [E_{min}, E_{max}]])$ be non-negative.*
- *Let $\Psi(x,\Omega,E)$ be non-negative $\forall (x,\Omega) \in \Gamma^-$ and $\Psi \in L^\infty(\Gamma^- \times [E_{min}, E_{max}])$.*

Then the linear Boltzmann equation (Eq. 2.1) equipped with any of the boundary conditions (Eqs. 2.4-2.8) has a unique weak solution $\Psi \in L^\infty(D \times \mathbb{S}^2 \times [E_{min}, E_{max}]])$.

For the proof, see [7].

2.3. Criticality Computations

Since they belong to the most important applications of the neutron transport equation, we will introduce criticality computations in this section. In order to do this, we need to derive the linear Boltzmann equation with fission. Fission is a reaction where a neutron is absorbed by a nucleus and induces the splitting of the nucleus into two daughter nuclei and additional neutrons. While most neutrons are prompt neutrons, i.e. generated instantaneously, generation of delayed neutrons is also possible. Since the fraction of delayed neutrons is usually small, we will assume here that all neutrons are prompt neutrons for simplicity. The gain of neutrons via fission can be generated by an additional source term

$$Q_f(x,E) = \chi_f(E)\int_{\mathbb{S}^2}\int_{E_{min}}^{E_{max}} \nu_f(E')\Sigma_f(E')\Psi(x,\Omega',E')dE'd\Omega', \tag{2.9}$$

where χ_f is prompt fission spectrum, the energy distribution of the fission neutrons, Σ_f is the fission cross section, ν_f is the neutron multiplicity, which is the average number of neutrons produced by fission of a particular nuclide. It should be noted that ν_f is an average over all possible fission reactions and thus not necessarily a natural number. By adding Eq. (2.9) to the right hand side of Eq. (2.1), we obtain the Linear Boltzmann equation with a fission source:

$$\begin{aligned}&\Omega \cdot \nabla\Psi(x,\Omega,E) + \Sigma_t(x,E)\Psi(x,\Omega,E)\\ &= \int_{\mathbb{S}^2}\int_{E_{min}}^{E_{max}} \Sigma_s(x,\Omega',\Omega,E',E)\Psi(x,\Omega',E')dE'd\Omega' + Q(x,\Omega,E)\\ &\quad + \chi_f(E)\int_{\mathbb{S}^2}\int_{E_{min}}^{E_{max}} \nu_f(E')\Sigma_f(E')\Psi(x,\Omega',E')dE'd\Omega'.\end{aligned} \tag{2.10}$$

In steady state reactor calculations, so called k-eigenvalue problems are often considered [78]. In this case external sources are absent, i.e. $Q(x, \Omega, E) = 0$. It is assumed that it is possible to control the expected number of fission neutrons. To this end, the fission source in Eq. (2.10) is modified by a constant factor $1/k$ with $k > 0$. The neutron transport equation then reads

$$\begin{aligned} &\Omega \cdot \nabla\Psi(x, \Omega, E) + \Sigma_t(x, E)\Psi(x, \Omega, E) \\ &= \int_{\mathbb{S}^2} \int_{E_{min}}^{E_{max}} \Sigma_s(x, \Omega', \Omega, E', E)\Psi(x, \Omega', E')dE'd\Omega' \\ &\quad + \frac{1}{k}\chi_f(E) \int_{\mathbb{S}^2} \int_{E_{min}}^{E_{max}} \nu_f(E')\Sigma_f(E')\Psi(x, \Omega', E')dE'd\Omega'. \end{aligned} \tag{2.11}$$

Vacuum boundary conditions (Eq. 2.5) are imposed. $\Psi = 0$ is always a trivial solution of Eq. (2.11). The goal is to find the largest k for which a non-trivial solution exists. This k is referred to as criticality is the system. The system can be in one of three possible states depending on the value of k:

- $k < 1$: The system is subcritical. Not enough neutrons are produced to replace the ones lost by absorption and leakage, i.e. leave the domain. The fission source must be increased for a steady-state solution to exist.
- $k = 1$: The system is critical. Fission produces as many neutrons as are absorbed or leak out of the system. A nontrivial solution of Eq. (2.11) exists.
- $k > 1$: The system is supercritical. More neutrons are produced than lost. The fission source must be decreased for a steady state solution to exist.

The calculation of the largest k for which Eq. 2.11 has a non-trivial solution is one of the most important calculations for the design and operation of nuclear power reactors. More details can be found in [78].

2.4. Neutron Particle Interactions

In the following, we will describe the physical reactions contributing to the absorption and scattering neutron cross sections Σ_a and Σ_s that we introduced above as well as the concept of neutron thermalization.

For our purposes, the only relevant absorption reaction is the radiative capture of a neutron, the (n, γ) reaction. That means that a neutron is absorbed by an atomic nuclei and its kinetic energy is then released via emission of gamma rays. As described above, these gamma rays can then be used for the detection of elements.

For scattering, two processes occur: elastic and inelastic neutron scattering. In elastic scattering interactions, the momentum and kinetic energy is conserved. Inelastic scattering results in a loss of kinetic energy to the elevation of the energy state of the atomic nucleus. This energy is emitted in the form of a gamma ray as the nucleus reverts to the ground state. Inelastic scattering typically occurs for high neutron energies, say above 10 keV, since the neutron kinetic energy must exceed a certain threshold energy in order to excite the first excited state of the nucleus [26]. It should be noted that it is possible for a neutron to loose a large amount of energy in an inelastic scattering reaction since much of its kinetic energy may be converted into the energy of excitation of the target nucleus.

For higher neutron energies, it is a good approximation to assume that the neutron scatters with a free atomic nucleus which is at rest. This is erroneous for lower energies, say below 10 eV, however. In this energy range, the thermal energy of the nucleus cannot be neglected anymore, so the thermal motion of the nucleus has to be taken into account. In addition to that, the chemical binding and crystalline effects of the target molecule need to be accounted for. This can be done with the use of so-called $S(\alpha, \beta)$-tables, where α and β stand for the unit-less momentum and energy transfer, respectively. A detailed introduction into this topic can be found in [93].

3. Deterministic Neutron Transport Methods

In the following, we discuss the discretization schemes for the neutron transport equation used in deterministic solvers. Because the neutron transport equation depends on each of its variables in significantly different ways, the discretization methods for each variable also vary vastly. In Sec. 3.1, the energy discretization with the multigroup approximation is introduced. In Sec. 3.2, we discuss the different methods for the angular discretization. The spatial discretization of the neutron transport equation can be carried out in many ways. While originally relatively simple finite difference schemes where favored, later on more sophisticated methods like finite element, nodal, characteristics or corner balanced methods were introduced [58]. In this work, we rely on the finite element method [3, 63].

3.1. Energy Discretization with the Multigroup Approximation

The multigroup approximation of the neutron transport equation is almost universally used to discretize the energy variable [78]. The resulting multigroup transport equations have a structure that closely resembles the original transport equation.
The multigroup approximation requires a finite number G of so-called energy groups to be chosen:

$$E_G < E_{G-1} < \dots < E_1 < E_0 \tag{3.1}$$

The g-th groups contains all particles with energies between E_g and E_{g-1}. Energy groups are per convention ordered so that the energy decreases as the group index increases. We assume that E_0 and E_G have been chosen in a way that ensures that the number of particles with energies higher than E_0 or smaller than E_G is negligible.
Our objective is to obtain an equation for the angular group flux

$$\Psi_g(x,\Omega) = \int_g \Psi(x,\Omega,E)dE. \tag{3.2}$$

For brevity we employ the shorthand notation

$$\int_g dE = \int_{E_g}^{E_{g-1}} dE \tag{3.3}$$

during this chapter.
By integrating Eq. (2.1) between E_g and E_{g-1} we obtain

$$\begin{aligned}\Omega \cdot \nabla \int_g \Psi(x,\Omega,E)dE + \int_g \Sigma_t(x,E)\Psi(x,\Omega,E)dE \\ = \sum_{g'=1}^{G} \int_{\mathbb{S}^2} \int_g \int_{g'} \Sigma_s(x,\Omega',\Omega,E',E)\Psi(x,\Omega',E')dE'dEd\Omega' + \int_g Q(x,\Omega,E)dE.\end{aligned} \tag{3.4}$$

By defining the total group cross section

$$\Sigma_{t,g}(x,\Omega) = \frac{\int_g \Sigma_t(x,E)\Psi(x,\Omega,E)dE}{\int_g \Psi(x,\Omega,E)dE}, \tag{3.5}$$

the group-to-group scattering cross-section

$$\Sigma_{s,gg'}(x,\Omega',\Omega) = \frac{\int_g \int_{g'} \Sigma_s(x,\Omega',\Omega,E',E)\Psi(x,\Omega',E')dE'dE}{\int_g \Psi(x,\Omega',E')dE'} \tag{3.6}$$

and the external group source

$$Q_g(x,\Omega) = \int_g Q(x,\Omega,E)dE, \tag{3.7}$$

we obtain the multigroup form of the neutron transport equation:

$$(\Omega\cdot\nabla + \Sigma_{t,g}(x,\Omega))\Psi_g(x,\Omega) = \sum_{g'=1}^{G}\int_{\mathbb{S}^2} \Sigma_{s,gg'}(x,\Omega',\Omega)\Psi_{g'}(x,\Omega')d\Omega' + Q_g(x,\Omega). \tag{3.8}$$

3.1.1. Energy Separability

While Eq. (3.8) is exact, it is not very useful in practice. The total group cross section and the group-to-group-scattering cross section depend on the solution of the continuous energy problem, which is not known. In the multigroup approximation, an approximation for Ψ is specified and introduced into the right sides of Eq. (3.5) and Eq. (3.6).

To proceed, we will use the assumption that the group flux is separable in energy. Suppose that the angular flux within each energy group can be approximated by a product of a function of the angle $h(x,\Omega)$ and a function of the energy $f(E,x)$:

$$\Psi(x,\Omega,E) = f(E,x)h(x,\Omega) \text{ for } E_g < E \leq E_{g-1} \tag{3.9}$$

We introduce Eq. (3.9) into the right sides of Eq. (3.5) and Eq. (3.6). Then angular dependent function $h(x,\Omega)$ cancels out and we obtain for the total group cross section

$$\Sigma_{t,g}(x) = \frac{\int_g \Sigma_t(x,E)f(E,x)dE}{\int_g f(E,x)dE}, \tag{3.10}$$

and the group-to-group scattering cross-section:

$$\Sigma_{s,gg'}(x,\Omega',\Omega) = \frac{\int_g \int_{g'} \Sigma_s(x,\Omega',\Omega,E',E)f(E',x)dE'dE}{\int_g f(E',x)dE'}. \tag{3.11}$$

With Σ_t now independent of the angle, we obtain the the following multigroup form of the neutron transport equation:

$$(\Omega\cdot\nabla + \Sigma_{t,g}(x))\Psi_g(x,\Omega) = \sum_{g'=1}^{G}\int_{\mathbb{S}^2} \Sigma_{s,gg'}(x,\Omega',\Omega)\Psi_{g'}(x,\Omega')d\Omega' + Q_g(x,\Omega). \tag{3.12}$$

The boundary conditions are obtained by integrating Eq. (2.4) over the energy groups so that we obtain

$$\Psi_g(x,\Omega) = \Psi_g^-(x,\Omega) \text{ for } \Omega\cdot n < 0,\ x \in \partial D, \tag{3.13}$$

where $\Psi_g^-(x,\Omega)$ is the incident group flux.

3.1.2. Computation of Multigroup Cross Sections

In order to solve Eq. (3.12), we need to compute multigroup cross sections defined in Eqs. (3.10) and (3.11) first. We assume that we can write the macroscopic total cross section as

$$\Sigma_t(x, E) = \sum_i \rho_i(x)\sigma_{t,i}(E), \tag{3.14}$$

where ρ_i is the number density for the ith isotope. The microscopic cross sections $\sigma_{t,i}$ are nuclear-physical properties of the isotopes and determined through measurements. Evaluated nuclear data is available through databases like the ENDF [20]. In comparison, $f(E)$ is far more difficult to obtain since it depends a great deal on the characteristics of the system under analysis, e.g. the geometry and exact material composition. This necessitates computation of multigroup cross-sections for each distinct application case.
Codes like NJOY [65] that are mainly based on approximate models where for a long time the primary mean to compute multigroup cross-sections. However, computational resources have become more easily available, so Monte Carlo codes have emerged as an alternative over the last years. These codes offer improved accuracy and ease of use. At the moment, Monte Carlo codes like Serpent [59] and OpenMC [85, 10] are in wide-spread use. The multigroup cross-sections used for this thesis are computed using OpenMC.

3.1.3. Solution of the Multigroup Equations for Fixed Source problems

The multigroup form of the neutron transport equation is a system of G coupled partial differential equations. Even on modern computing systems, directly solving such a system sets very high demands on the RAM for large values of G, so it is usual to solve this system iteratively. In the following, the two most common iterative methods for multigroup neutron transport are introduced. Both methods are closely related to matrix splitting methods used for solving linear systems. By formally defining the multigroup transport operator H by

$$H_{gg'}(x,\Omega)\Psi_{g'}(x,\Omega) = \delta_{gg'}(\Omega\cdot\nabla + \Sigma_{t,g}(x))\Psi_{g'}(x,\Omega) - \int_{\mathbb{S}^2} \Sigma_{s,gg'}(x,\Omega',\Omega)\Psi_{g'}(x,\Omega')d\Omega' \tag{3.15}$$

we can rewrite Eq. (3.8) as an operator acting on an unknown vector valued quantity:

$$H \cdot \Psi = Q, \tag{3.16}$$

We can now split H, that means separating an invertible submatrix B:

$$H = B + (H - B) \tag{3.17}$$

By defining

$$M = B^{-1}(B - H) \tag{3.18}$$

where B^{-1} is the inverse of B we obtain the fixed point equation

$$\Psi = M \cdot \Psi + B^{-1}Q \tag{3.19}$$

that can be solved with an iteration of the form:

$$\Psi^{(n+1)} = M \cdot \Psi^{(n)} + B^{-1}Q. \tag{3.20}$$

That means $M(x,\Omega)$ is the iteration matrix. To avoid the inversion of differential operators, we rewrite Eq. (3.20) as

$$B\Psi^{(n+1)} = (B-H)\cdot\Psi^{(n)} + Q. \tag{3.21}$$

The first method introduced here closely resembles the Gauss-Seidel method used to solve linear equations. We split the left lower triangular matrix from H. The iterative procedure is then given by

$$H_{gg}\Psi_g^{(n+1)} = Q_g - \sum_{g'<g} H_{gg'}\Psi_{g'}^{(n+1)} - \sum_{g'>g} H_{gg'}\Psi_{g'}^{(n)} \tag{3.22}$$

The splitting of the left lower triangular matrix can be physically motivated: in most applications so-called downscattering dominates, that means neutrons loose energy when scattering. This process is described by the left lower triangular components of H, so it is close at hand to split this part of the matrix. Upscattering, that means neutrons gain energy when the scattering, is then effectively a small correction.
The second method is the iteration on the scattering source. This methods splits the identity from H. For linear systems, this method is known as Richardson iteration. The iterative procedure is given by

$$(\Omega\cdot\nabla + \Sigma_{t,g})\Psi_g^{(n+1)}(x,\Omega) = Q_{s,g}^{(n)}(x,\Omega) \tag{3.23}$$

with the scattering source

$$Q_{s,g}^{(n)}(x,\Omega) = \begin{cases} Q_g(x,\Omega) & n=0 \\ \sum_{g'=1}^{G}\int_{\mathbb{S}^2}\Sigma_{s,gg'}(x,\Omega',\Omega)\Psi_{g'}^{(n)}(x,\Omega')d\Omega' & n>0 \end{cases} \tag{3.24}$$

and the starting value

$$\Psi_g^0(x,\Omega) = 0. \tag{3.25}$$

The neutron flux is then given by

$$\Psi_g(x,\Omega) = \sum_{n=0}^{\infty}\Psi_g^{(n)}(x,\Omega). \tag{3.26}$$

We can give a physical interpretation of this iterative method. $\Psi_g^{(0)}$ includes only uncollided neutrons, while $\Psi_g^{(n)}$ includes all neutrons that collided exactly n times. This directly connects the physical number of neutron scattering processes and the convergence behaviour of the method, which is one of the major reasons for the popularity of this method.
Splitting methods are constructed in a way that ensures that they are contractions that can be characterized by the contraction number

$$c = \lim_{n\to\infty}\frac{||\Psi^{(n+1)}-\Psi^{(n)}||_{L^1(D)}}{||\Psi^{(n)}-\Psi^{(n-1)}||_{L^1(D)}} \tag{3.27}$$

where the L^1 norm is defined as

$$||\Psi(x,\Omega)||_{L^1(D)} = \sum_{g=1}^{G}\int_{\mathbb{S}^2}\int_D |\Psi_g(x,\Omega)|dxd\Omega \tag{3.28}$$

The limit of the iteration method is the fixed point Ψ_*. The distance to Ψ_* and therefore the absolute error is a posteriori given by

$$||\Psi^{(n+1)} - \Psi^{(n)}||_{L^1(D)} \leq \frac{c}{c-1}||\Psi^{(n+1)} - \Psi^{(n)}||_{L^1(D)} \tag{3.29}$$

The relative error is thus given by

$$E_{c,rel} = \frac{c}{c-1}\frac{||\Psi^{(n+1)} - \Psi^{(n)}||_{L^1(D)}}{||\Psi^{(n+1)}||_{L^1(D)}}. \tag{3.30}$$

3.2. Angular Discretization

With the definition

$$S_g(x,\Omega) = \sum_{g' \neq g}^{G} \int_{\mathbb{S}^2} \Sigma_{s,gg'}(x,\Omega',\Omega)\Psi_{g'}(x,\Omega')d\Omega' + Q_g(x,\Omega) \tag{3.31}$$

we can write Eq. (3.12) for each g as

$$(\Omega \cdot \nabla + \Sigma_{t,g}(x))\Psi_g(x,\Omega) = \int_{\mathbb{S}^2} \Sigma_{s,gg}(x,\Omega',\Omega)\Psi_g(x,\Omega')d\Omega' + S_g(x,\Omega). \tag{3.32}$$

This is the within-group form of the multigroup equation, with a group-to-group source $S_g(x,\Omega)$ representing both the external source and the scattering from groups $g' \neq g$ into g. In this form, the equation for each group g has the form of a mono-energetic transport equation. For the rest of this subsection, we concentrate on the angular discretization, so the group index g will be omitted for brevity. We can now introduce the standard methods for the angular discretization of the neutron transport equation [78]: the discrete ordinates method ($\mathrm{S_N}$ method) and a moment method, the method of spherical harmonics ($\mathrm{P_N}$ method). We will introduce the later in a simplified setting, the one-dimensional slab geometry. After that, we introduce the simplified $\mathrm{P_N}$ equations ($\mathrm{SP_N}$ equations), which we are a used throughout this thesis, in greater detail. It should be noted the the $\mathrm{SP_1}$ equation is the neutron diffusion equation which is also in widespread use for neutron transport computations.

3.2.1. The Discrete Ordinates Method

The discrete ordinates method demands Eq. (3.32) to be satisfied only at a set of discrete ordinates. This ordinates can then be physically interpreted as the directions in which the particles travel. As a starting point, a set of quadrature points $\{\Omega_1, ..., \Omega_N\} \in \mathcal{S}^2$ with weights $\{w_1, ..., w_N\}$ is defined. A common choice are Gauss points in μ with the corresponding weights and equidistant points in ϕ. The angular-integrated neutron flux can then be approximated using this quadrature:

$$\Psi(x) = \int_{\mathbb{S}^2} \Psi(x,\Omega)d\Omega \approx \sum_{n=1}^{N} w_n \Psi(x,\Omega_n). \tag{3.33}$$

The within-group neutron transport equation discretized with the $\mathrm{S_N}$ method yields the system of equations

$$(\Omega_n \cdot \nabla + \Sigma_t(x))\Psi_n(x) = \sum_{n'=1}^{N} w_{n'}\Sigma_s(\Omega'_{n'}, \Omega_n)\Psi(x, \Omega'_{n'}) + S(x, \Omega_n) \tag{3.34}$$

where

$$\Psi_n(x) = \Psi(x, \Omega_n). \tag{3.35}$$

The $\mathrm{S_N}$ equations are a set of linear advection equations that is coupled by the scattering term on the right-hand side.

3.2.2. Spherical Harmonics

An alternative approach for the angular discretization, which we will persue within this thesis, are moment models. We will start by introducing the $\mathrm{P_N}$ equations for a slab geometry, where the neutron flux is expanded in Legendre polynomials.
In slab geometry, the medium is assumed to be infinite in two dimensions (here the y- and z-directions) and finite in one dimension (here the x-direction). The medium is therefore bounded by two parallel planes which form a slab through which the neutrons travel. The now one-dimensional spatial coordinate x is used to describe the position of a neutron within the slab, where $x \in D$ and $D = [x_l, r_r] \subset \mathbb{R}^1$. The direction of neutron movement is described by the polar angle $\mu = \cos(\theta)$, which is here defined as the angle between the neutron's direction of travel and the x-axis. Due to its symmetry, the problem is independent of the azimuthal angle φ. Furthermore, we assume an isotropic source, that is Q independent of the angle.
The neutron transport equation with isotropic scattering in slab geometry [18] reads

$$\left(\mu\frac{d}{dx} + \Sigma_t(x)\right)\Psi(x, \mu) = \frac{1}{2}\int_{-1}^{1} \Sigma_s(x)\Psi(x, \mu')d\mu' + Q(x). \tag{3.36}$$

We expand the flux in Legendre polynomials

$$\Psi(x, \mu) = \sum_{n=0}^{N} \frac{2n+1}{4\pi}\phi_n(x)P_n(\mu) \tag{3.37}$$

with the moments of the flux defined by

$$\phi_n(x) = \int_{-1}^{1} P_n(\mu)\psi(x, \mu)d\mu \tag{3.38}$$

and the Legendre polynomials $P_n(\mu)$. We also define:

$$\Sigma_n(x) = \begin{cases} \Sigma_t(x) - \Sigma_s(x) & n = 0 \\ \Sigma_t(x) & n \geq 1 \end{cases} \tag{3.39}$$

Then the slab geometry $\mathrm{P_N}$ equations read:

$$\begin{aligned} \frac{d\phi_1}{dx}(x) + \Sigma_0\phi_0(x) &= Q(x) \\ \frac{n}{2n+1}\frac{d\phi_{n-1}}{dx}(x) + \frac{n+1}{2n+1}\frac{d\phi_{n+1}}{dx}(x) + \Sigma_n(x)\phi_n(x) &= 0 \text{ for } n > 0 \end{aligned} \tag{3.40}$$

For the derivation, see e.g. [70]. Eqs. 3.40 form a system of $N+1$ moment equations with $N+2$ unknowns, so a closure is required to achieve a well-posed system. For our purposes, it is sufficient to close the equations by setting $\phi_{N+1}=0$. It should be mentioned however that other closures are also possible [39]. To obtain the Marshak boundary conditions an Nth order expansion where N is odd, we take the Legendre polynomial moments of 2.4 and obtain

$$\begin{aligned}\int_0^1 P_n(\mu)\Psi(x_l,\mu)d\mu &= \int_0^1 P_n(\mu)\Psi^-(x_l,\mu)d\mu \\ \int_0^{-1} P_n(\mu)\Psi(x_r,\mu)d\mu &= \int_0^{-1} P_n(\mu)\Psi^-(x_r,\mu)d\mu \end{aligned} \tag{3.41}$$

for $n=1,3,5,...,N$. Recall that $D=[x_l,x_r]$, so x_l is the left boundary of the system and x_r is the right boundary of the system. If N is even, the subject of boundary conditions is more subtle [60]. Expanding Ψ using Eq. 3.37 then gives

$$\begin{aligned}\sum_{l=0}^{N}\frac{2l+1}{4\pi}\phi_n(x_l)\int_0^1 P_l(\mu)P_n(\mu)d\mu &= \int_0^1 P_n(\mu)\Psi^-(x_l,\mu)d\mu \\ \sum_{l=0}^{N}\frac{2l+1}{4\pi}\phi_n(x_r)\int_0^{-1} P_l(\mu)P_n(\mu)d\mu &= \int_0^{-1} P_n(\mu)\Psi^-(x_r,\mu)d\mu. \end{aligned} \tag{3.42}$$

The numerical solution of the slab geometry $\mathrm{P_N}$ equations is discussed in Chapter 8.

3.2.3. Simplified Spherical Harmonics

While it is possible to generalize the $\mathrm{P_N}$ to three dimensions by using full spherical harmonics instead of just Legendre polynomials, this has serious drawbacks. For instance, instead of the set of N equations we obtain for the slab geometry, we obtain a set of $(N+1)^2$ equations. In order to avoid this, we use the $\mathrm{SP_N}$ approximation.

The $\mathrm{SP_N}$ approximation was originally introduced by Gelbard [33] in an ad-hoc way. Over the years, a theoretical foundation has been established. There are several ways to use asymptotic analysis to derive the $\mathrm{SP_N}$-equations [77],[27],[32]. A derivation of the $\mathrm{SP_3}$ equations based on variational analysis can be found in [11]. This derivation also yields the appropriate boundary and material-interface conditions. A detailed review of the theory underpinning the $\mathrm{SP_N}$ equations can be found in [70]. In this section, we have $x\in D$ where $D\subset\mathbb{R}^3$ is bounded. The so-called second-order form of the $\mathrm{SP_N}$ equations read

$$\begin{aligned} -\nabla\cdot\frac{1}{3\Sigma_1}\nabla\phi_0-\nabla\cdot\frac{2}{3\Sigma_1}\nabla\phi_2+\Sigma_0 &= Q \\ -\nabla\cdot\frac{n(n-1)}{(2n+1)(2n-1)\Sigma_{n-1}}\nabla\phi_{n-2}-\nabla\cdot\frac{(n+1)(n+2)}{(2n+1)(2n+3)\Sigma_{n+1}}\nabla\phi_{n+2} & \\ -\nabla\cdot\left(\frac{n^2}{(2n+1)(2n-1)\Sigma_{n-1}}+\frac{(n+1)^2}{(2n+1)(2n+3)\Sigma_{n+1}}\right)\nabla\phi_n+\Sigma_n\phi_n &= 0 \\ \text{for } n=2,4,...,N-1. & \end{aligned} \tag{3.43}$$

The second-order form of the $\mathrm{SP_N}$ is useful in practice since the equations are a system of coupled diffusion equations.

In the following, the second order forms of $\mathrm{SP_1}$ and $\mathrm{SP_3}$ will be written out explicitly, since those will be used in the later chapters. We assume isotropic scattering, that is $\Sigma_{sn} = 0$ for $n > 0$, and consequently $\Sigma_n = \Sigma_t$ for $n > 0$. Furthermore, we will use vacuum boundary conditions, which are Marshak boundary conditions with the incoming flux set to zero. The SP_1 equation then reads

$$-\nabla \cdot \frac{1}{3\Sigma_t}\nabla\phi_0 + \Sigma_0\phi_0 = Q. \tag{3.44}$$

with the boundary condition

$$\frac{1}{2}\phi_0(x) + \frac{1}{3\Sigma_t(x)} n \cdot \nabla\phi_0(x) = 0,\ x \in \partial D \tag{3.45}$$

where n is the outer normal vector. The SP_3 equations are given by

$$\begin{aligned} -\nabla \cdot \frac{1}{3\Sigma_t}\nabla\phi_0 - \nabla \cdot \frac{2}{3\Sigma_t}\nabla\phi_2 + \Sigma_0\phi_0 &= Q \\ -\nabla \cdot \frac{2}{15\Sigma_t}\nabla\phi_0 - \nabla \cdot \frac{11}{21\Sigma_t}\nabla\phi_2 + \Sigma_t\phi_2 &= 0. \end{aligned} \tag{3.46}$$

With some algebra, we obtain the equivalent system

$$\begin{aligned} -\nabla \cdot \frac{1}{3\Sigma_t}\nabla\phi_0 - \nabla \cdot \frac{2}{3\Sigma_t}\nabla\phi_2 + \Sigma_0\phi_0 &= Q \\ \nabla \cdot \frac{9}{35\Sigma_t}\nabla\phi_2 + \Sigma_t\phi_2 &= \frac{2}{5}(\Sigma_0\phi_0 - Q), \end{aligned} \tag{3.47}$$

which will be used for this thesis. The vacuum boundary conditions for this system read:

$$\begin{aligned} \frac{1}{2}\phi_0 + \frac{1}{3\Sigma_t}(n \cdot \nabla)(\phi_0 + 2\phi_2) &= -\frac{5}{8}\phi_2 \\ \frac{3}{8}\phi_2 + \frac{9}{35\Sigma_t}(n \cdot \nabla)\phi_2 &= \frac{3}{40}\phi_0,\ x \in \partial D. \end{aligned} \tag{3.48}$$

The numerical solution SP_1 and SP_3 equations is discussed in the following chapter.

4. Development of SPARC - a Neutron Transport Solver based on the $\mathrm{SP_N}$ Approximation

In this chapter, we will describe the implementation of neutron transport solver SPARC. The development of said solver was necessary since the available neutron transport solvers are unsuitable for our application, large sample PGNAA analysis.
While Monte Carlo codes like MCNP [83] or OpenMC [85] offer a good accuracy, they are computationally too expansive for our application, especially when not only a good energy resolution but also a high spatial resolution is demanded. However, thanks to their high accuracy, they are well suited for benchmarks.
Most deterministic codes like Attila [72] and SCALE [99] use the $\mathrm{S_N}$-approximation. These codes are also not suitable for our application. On one hand, these codes are usually developed with other applications in mind, e.g. SCALE is optimized for nuclear reactor simulations and analysis. On the other hand, we expect that solver based on the $\mathrm{SP_N}$-approximation has a better performance while still offering sufficient accuracy.
All this factors combined lead to the development of neutron transport solver based on the $\mathrm{SP_N}$ approximation. This chapter is organized as follows: in Sec. 4.1, the implementation of the solver is discussed. In Sec. 4.2, we introduce the method of manufactured solutions and use it to verify our solver. Finally, we discuss the expansion of the solver on multigroup neutron transport in Sec. 4.3.

4.1. Implementation

In order to solve the $\mathrm{SP_1}$ equation (Eq. 3.44) and the $\mathrm{SP_3}$ equations (Eqs. 3.47) for this specific application, a special neutron transport code named SPARC was developed. It is completely based on open source libraries. For the spatial discretization of the $\mathrm{SP_3}$ equations, the finite element method was used since it features great geometrical flexibility. An additional benefit is that the finite element method works very well with Poisson-like equations, in this case the odd order $\mathrm{SP_N}$-equations. A flowchart of the implementation is depicted in Fig. 4.1. The FEniCS library [3, 63] with PETSc [6] as linear algebra backend is used for the implementation. It offers a large amount of linear solvers and is complemented by hypre [28], a library of high performance preconditioners. For this work a GMRES solver [88] is used in conjunction with an algebraic multigrid preconditioner [86]. As a mesh generator, Gmsh [34] is used. The multigroup cross sections for SPARC are computed using the OpenMC Monte Carlo code [85, 10]. While Monte Carlo methods are computationally too expensive for our application, this does not apply to the computation of multigroup cross sections. The multigroup cross sections are precomputed and stored in a database, so the Monte Carlo computation needs to be carried out only once and not for each sample as it is the case for the neutron flux.

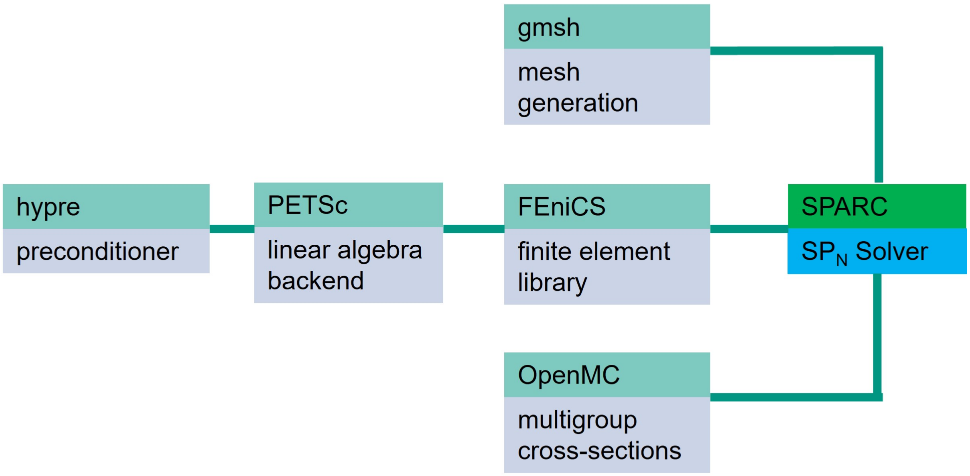


Figure 4.1.: Flowchart of the SPARC implementation.

4.2. Verification

In order to verify the implementation, we use the method of manufactured solutions [84]. This method is used to test whether a code solves the underlying physical equations correctly. We will first introduce the method using a two dimensional poisson equation with an additional linear term as an example. After that, two three dimensional test cases with the SP_3-equations will be presented, one of them featuring two energy groups.

4.2.1. Method of Manufactured Solutions

The method of manufactured solution works as follows. We start with a partial differential equation, here a Poisson equation with an additional linear term:

$$-\frac{1}{3}\Delta f + \frac{1}{10}f = Q \text{ in } D \tag{4.1}$$

Here, D is a square stretching form -1 to 1 in both x and y directions. Now we choose an analytical solution

$$f_{ana} = x^2 - y^2 \tag{4.2}$$

We insert this solution into the left hand side of the differential equation (Eq. 4.1) and obtain the source term:

$$Q = \frac{1}{10}(x^2 - y^2). \tag{4.3}$$

We use Dirichlet boundary conditions:

$$f_{ana} = x^2 - y^2 \text{ on } \partial D. \tag{4.4}$$

We expect that if the implementation is correct, then the relative error of the solution decreases as the mesh is refined. The relative errors in this section are to be understood as relative L^2 errors and computed as follows:

$$\xi_{rel} = \sqrt{\frac{\int_D |f_{FE} - f_{ana}|^2 dx}{\int_D |f_{ana}|^2 dx}}, \tag{4.5}$$

where f_{FE} is the finite element solution and f_{ana} is the analytical solution obtained with the method of manufactured solution.

Fig. 4.2 depicts the solution of Eq. (4.1) with aforementioned source and boundary conditions.

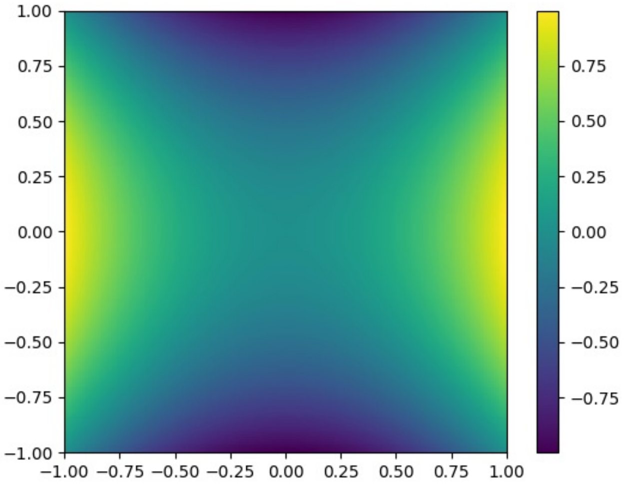


Figure 4.2.: Finite Element solution of Eq. (4.1). The analytical solution is given by $f = x^2 - y^2$.

Tab. 4.1 shows the relative error of the Finite Element Solution obtained with FEniCS. A mesh with triangular Finite Elements is used. As can be seen, the already small error of 0.6 % obtained for a mesh with a finite element size of 0.2 can be further reduced by refining the mesh. For a finite element size of 0.067, the error is only 0.07 %. The values in Tab. 4.1 are depicted in Fig. 4.3. As can be seen, the linear Finite Element method converges quadratically in L^2 norm.

Table 4.1.: Error of the Finite Element Solution of Eq. (4.1).

Finite Element size	0.2	0.1	0.067
relative error	0.60 %	0.15 %	0.07 %

4.2.2. Verification of the SP$_3$ Solver

Equipped with the method of manufactured solutions, we can now verify the implementation of the SP$_3$ equations. The emphasis in this subsection is the verification of the results within the domain and not on the boundary, so we will use Dirichlet boundary conditions as a simplification. A manufactured solution with Marshak boundary conditions is introduced for the verification of the adjoint solver in Chapter 6.
In the test case, we solve the SP$_3$-Equations (Eq. 3.46) for the cross sections $\Sigma_t = 1.0$ and $\Sigma_0 = 0.1$. The domain D is a cube stretching from -1 to 1 in each direction. With the method of method of manufactured solutions, we can derive an analytical solution for this equation. We define the auxiliary function:

$$f(x, y, z) = \cos(x)\cos(y)\cos(z). \tag{4.6}$$

The moments of the neutron flux are $\phi_0 = f$ and $\phi_2 = -\frac{7}{45}f$ with the source term $Q = \frac{71}{90}f$. The Dirichlet boundary conditions are $\phi_0 = f$ on ∂D and $\phi_2 = -\frac{7}{45}f$ on ∂D. The Finite Element

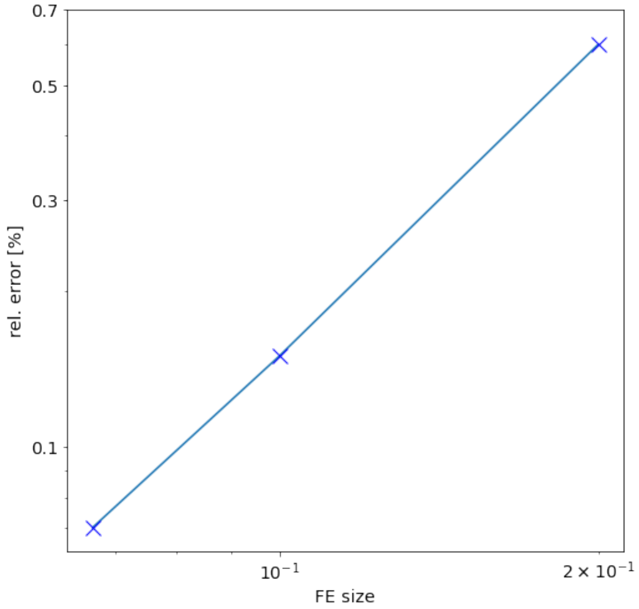
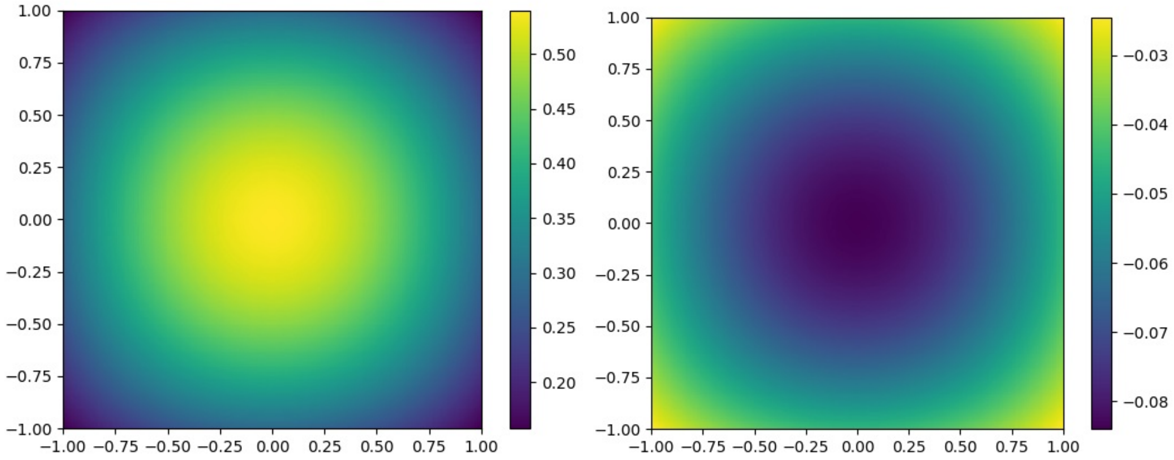


Figure 4.3.: Relative error of the Finite Element solution of Eq. (4.1) depending on the mesh size.

solution for a mesh with 162000 tetrahedral linear finite elements is depicted in Fig. 4.4. The figure shows the font side of the cube. The error of the finite element solution is shown in Tab. 4.2. While the error is over 2 % for the coarse mesh with a finite element size of 0.2, it reduces as the mesh is refined. For a mesh with a finite element size of 0.067, the error is only about 0.3 %. Fig. 4.5 depicts the results for ϕ_0. As can be seen, the linear Finite Element method converges quadratically in L^2 norm.



(a) Finite element solution of ϕ_0.

(b) Finite element solution of ϕ_2.

Figure 4.4.: Finite element solution of a mono-energetic SP$_3$ test case generated with the method of manufactured solutions in a cubic geometry. Depicted is the front side of the cube ($z = -1$).

The mono-energetic SP$_3$ test case can be expanded two include a second energy group. We assume that only downscattering from group one to group two occurs, but no upscattering from group two to one. Such cases are not uncommon in reactor physics, where computations are often carried out for one fast and one thermal energy group with no significant upscattering.

Table 4.2.: Error of the finite element solution for a mono-energetic SP_3 test case generated with the method of manufactured solutions in a cubic geometry.

Finite Element size	0.2	0.1	0.067
rel. error ϕ_0	2.95 %	0.75 %	0.33 %
rel. error ϕ_2	2.30 %	0.57 %	0.25 %

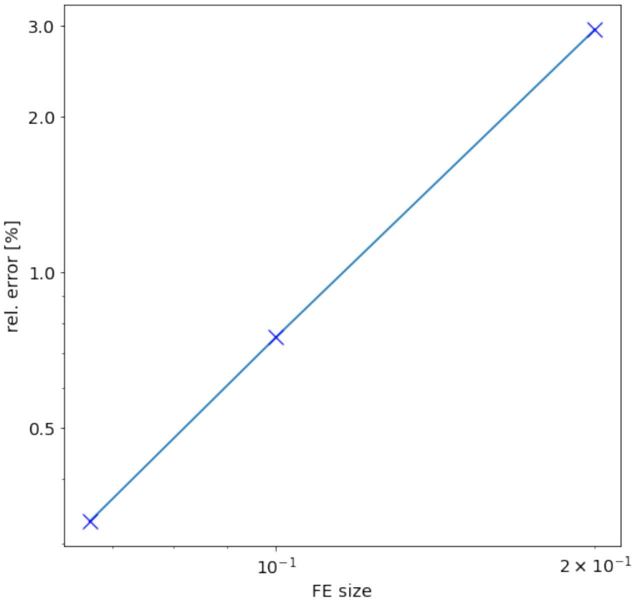


Figure 4.5.: Relative error of the Finite Element solution depending on the mesh size for a mono-energetic SP_3 test case generated with the method of manufactured solutions in a cubic geometry.

Therefore, the moments of the flux from group one, which we denote now $\phi_{0,1}$ and $\phi_{2,1}$, where the second index indicates the group, are identical to the last test case. For simplicity, we set the cross section for the scattering from group one to group two as $\Sigma_{s,21} = 1.0$. There is no scattering form group two to one, so $\Sigma_{s,12} = 0.0$. The analytical solution for group two is then $\phi_{0,2} = \frac{90}{71} f$ and $\phi_{2,2} = -\frac{14}{71} f$. The results for the SP_3 test case with two energy groups are shown in Tab. 4.3. The results for group one are identical to the former test case, which is expected. For group two, the error is also decreasing as the mesh is refined, form about 4 % for the coarsest mesh to below 0.5 % for the finest mesh. The errors for group two are also larger than for group one. This is not unexpected. The source term for group two is proportional to $\phi_{0,1}$, which is only determined as accurately as shown in Tab. 4.3. The discretization error in the computation of $\phi_{0,1}$ and the discretization error of group two add up, leading to a larger error. Fig. 4.6 depicts the error depending on the finite element size for $\phi_{0,2}$. As can be seen, the convergence of the linear Finite Element method in L^2 is quadratic as all is smooth.

Table 4.3.: Error of the finite element solution for a SP_3 test case with two energy groups generated with the method of manufactured solutions in a cubic geometry.

Finite Element size	0.2	0.1	0.067
rel. error $\phi_{0,1}$	2.95 %	0.75 %	0.33 %
rel. error $\phi_{2,1}$	2.30 %	0.57 %	0.25 %
rel. error $\phi_{0,2}$	4.25 %	1.09 %	0.49 %
rel. error $\phi_{2,2}$	3.74 %	0.96 %	0.43 %

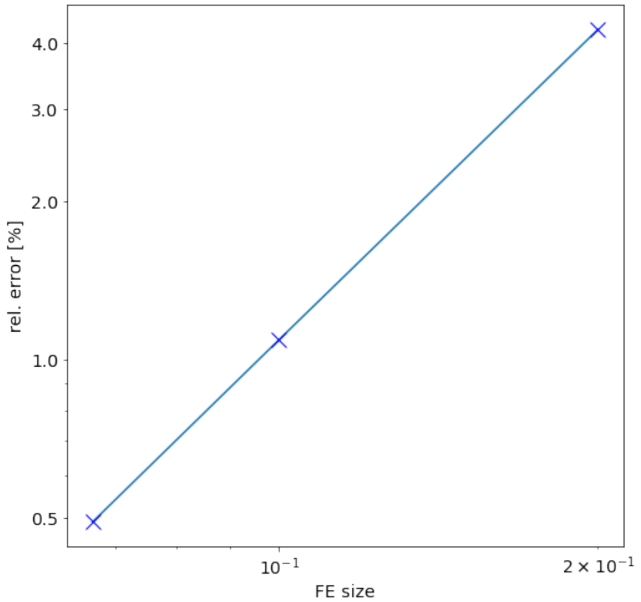


Figure 4.6.: Depiction of relative error of the Finite Element solution of $\phi_{0,2}$ depending on the mesh size for a SP_3 test case with two energy groups generated with the method of manufactured solutions in a cubic geometry.

4.3. Numerical Methods for Multigroup Neutron Transport

In this section, we will discuss the implementation of the multigroup neutron transport in SPARC in detail. We investigate a simplified setting, a pure graphite cube, in this chapter.

Graphite is the main constituent of the measurement facilities and used as the moderator, so it is close at hand to investigate this setting before investigating more complex geometries introduced in the following chapters. For the graphite cube, we examine the convergence behavior of the Gauss Seidel and Source Iteration methods introduced in Subsec. 3.1.3. We also show how important upscattering effects are for this application and determine up to which energy these effects need to be included. Upscattering means that neutrons gain energy when scattering. This effect is described by the components of the group-to-group scattering matrix $\Sigma_{s,gg'}$ which are above the diagonal.
Recall that the multigroup form of the neutron transport equation is given by

$$(\Omega \cdot \nabla + \Sigma_{t,g}(x))\Psi_g(x,\Omega) = \sum_{g'=1}^{G} \int_{\mathbb{S}^2} \Sigma_{s,gg'}(x,\Omega',\Omega)\Psi_{g'}(x,\Omega')d\Omega' + Q_g(x,\Omega). \tag{4.7}$$

For the angular discretization, we use the $\mathrm{SP_N}$ approximation introduced in Subsec. 3.2.3. The multigroup form of the $\mathrm{SP_3}$-equations (see [11]) in the case of an isotropic source and isotropic scattering read

$$\begin{aligned} &-\nabla \cdot \frac{1}{3\Sigma_{t,g}}\nabla\left(\phi_{0,g}(x) + 2\phi_{2,g}(x)\right) + (\Sigma_{t,g} - \Sigma_{s,gg})\phi_{0,g}(x) = S_g(x) \\ &-\nabla \cdot \frac{9}{35\Sigma_{t,g}}\nabla\phi_{2,g}(x) + \Sigma_{t,g}\phi_{2,g}(x) = \frac{2}{5}\left((\Sigma_{t,g} - \Sigma_{s,gg})\phi_{0,g}(x) - S_g(x)\right), \end{aligned} \tag{4.8}$$

where $\phi_{0,g}$ and $\phi_{2,g}$ are the zeroth respectively the second Legendre moment of the neutron flux in the gth energy group. The source term of the gth group is given by

$$S_g(x) = \sum_{g \neq g'} \Sigma_{s,gg'}\phi_{0,g'} + Q_g. \tag{4.9}$$

and the boundary conditions read

$$\begin{aligned} \frac{1}{2}\phi_{0,g} + \frac{1}{3\Sigma_{t,g}}(n\cdot\nabla)(\phi_{0,g} + 2\phi_{2,g}) &= -\frac{5}{8}\phi_{2,g} \\ \frac{3}{8}\phi_{2,g} + \frac{9}{35\Sigma_{t,g}}(n\cdot\nabla)\phi_{2,g} &= \frac{3}{40}\phi_{0,g}, \ x \in \partial D. \end{aligned} \tag{4.10}$$

We determine the weak solution of the system with the Finite Element method (see [3, 63]), so we have $\phi_{0,g} \in H^1(D)$ and $\phi_{2,g} \in H^1(D)$. The $\mathrm{SP_1}$-equation, also widely known as neutron diffusion equation, is obtained from Eqs. (4.8) by setting the higher order moments to zero. In this section, we use the group structure with 44 energy groups from [99], so $G = 44$. We have the cubic domain $D = (-50, 50)^3$. The mesh contains 48000 Finite Elements with a maximum edge length of 5 and and 9261 nodes. We use the SPARC code implemented as described in Sec. 4.1 as the Finite Element solver.
In the following, we investigate the solution of the multigroup equations for pure graphite. Graphite is the moderator used in all measurement facilities discussed within this thesis and also main constituent of said facilities. Therefore, we study the solution of the multigroup equations in a simplified setting containing only graphite before examining more complex cases. The results shown in this Subsection are computed for a homogeneous graphite cube with an edge length of 100 cm and a density 1.825 g/cm^3. A mono-energetic point source with an energy of 2.45 MeV is placed at the center of the cube. For the computation of the multigroup cross sections $\Sigma_{t,g}$ $(1 \leq g \leq 44)$ and $\Sigma_{s,gg'}$ $(1 \leq g \leq 44,\ 1 \leq g' \leq 44)$ we use the OpenMC Monte Carlo code [85, 10] (see Subsec. 3.1.2).

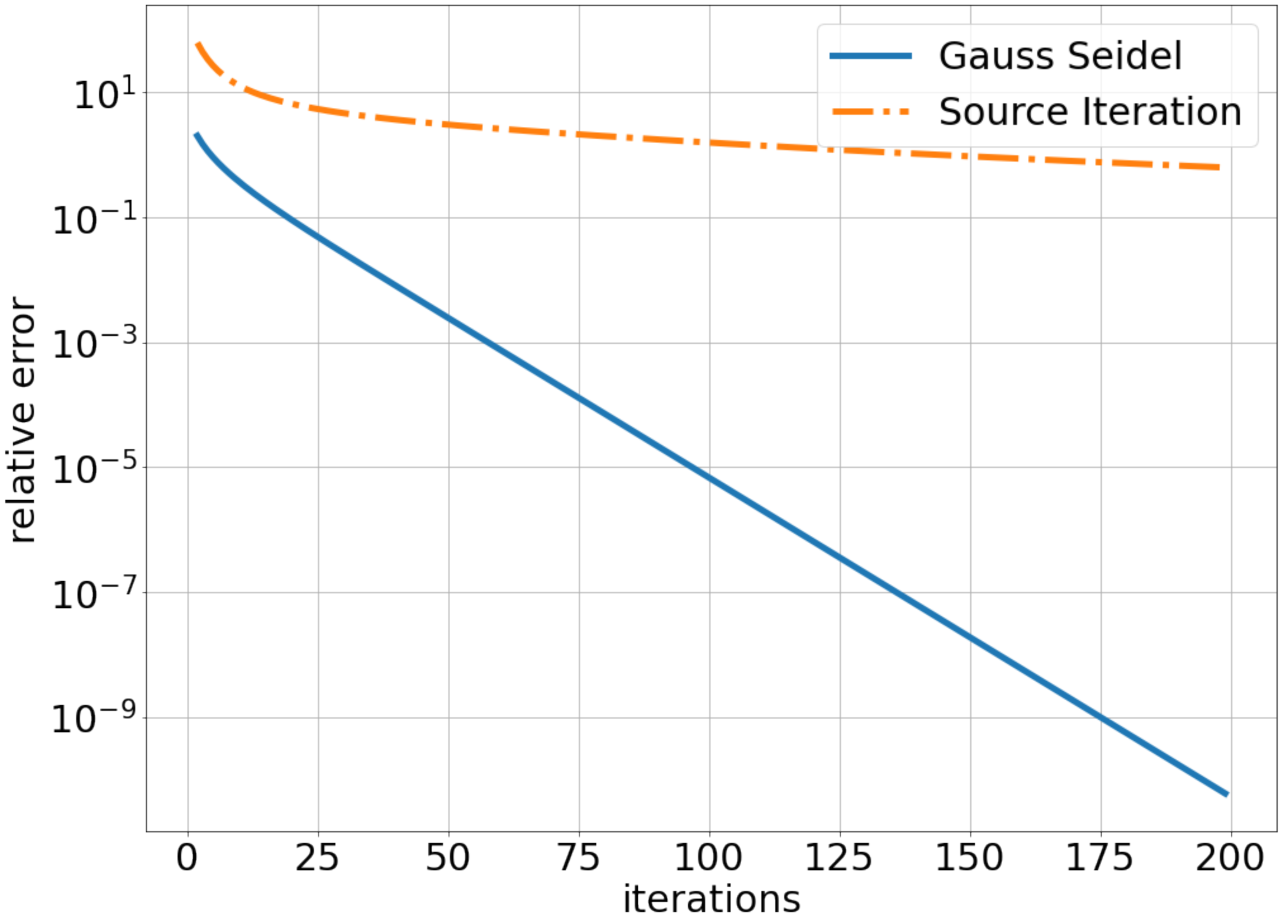


Figure 4.7.: Convergence behaviour for the Gauss Seidel (Eq. 3.22) and Source Iteration methods (Eq. 3.23) described in Subsec. 3.1.3 for a homogeneous graphite cube with an edge length of 100 cm and a mono-energetic point source with an energy of 2.45 MeV in the center. The relative error is the distance to the fix point of the iteration.

Fig. 4.7 shows a comparison of the convergence behaviour for the Gauss Seidel and the Source Iteration methods introduced in Subsec. 3.1.3. As can be seen, the Gauss Seidel method performs much better. The contraction factors are 0.995 for the Source Iteration and 0.889 for the Gauss Seidel method. Therefore, only the Gauss Seidel method is used in the following.

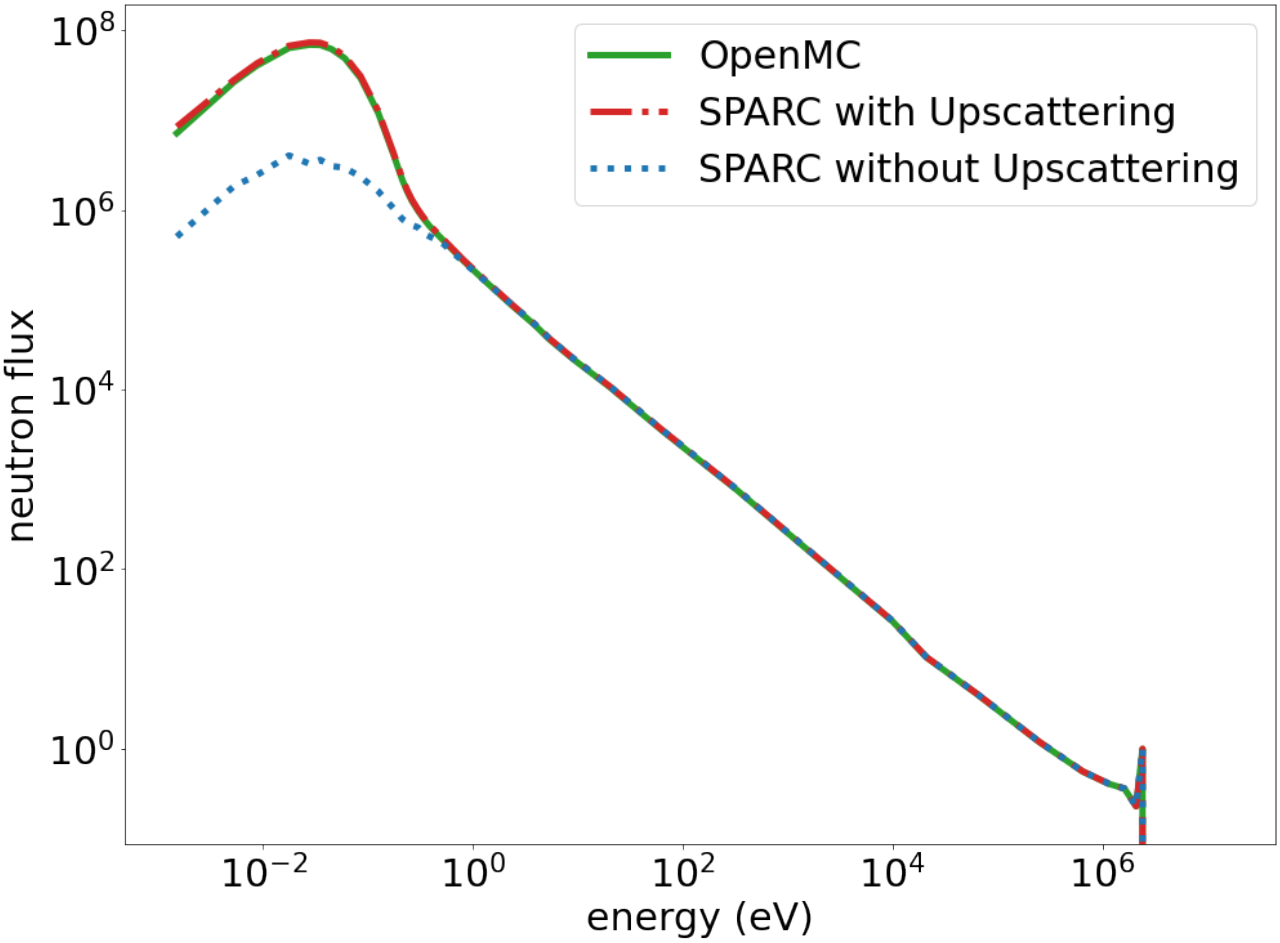


Figure 4.8.: Neutron fluxes computed with OpenMC and SPARC for a homogeneous graphite cube with an edge length of 100 cm and a mono-energetic point source with an energy of 2.45 MeV in the center. The SPARC results with upscattering agree well with OpenMC. Without upscattering, there are large deviations between the SPARC results and OpenMC for energies below 1 eV.

In many applications upscattering can be neglected. Therefore, it is sufficient to solve the multigroup equations successively from higher to lower energies without using any iterative method. Fig. 4.8 shows that this is not the case here. While SPARC with upscattering agrees well with OpenMC, SPARC without upscattering deviates by over one order of magnitude in the low energy region.

While upscattering is very important at low energies, its effect is negligible at higher energies. Thus, upscattering effects are usually cut off at application-specific energies so that the Gauss Seidel iterations only need to include the lower part of the energy groups, which improves the performance significantly. Fig. 4.9 depicts the ratio of SPARC without upscattering and SPARC with upscattering. While it is 1 for higher energies and thus the difference is negligible, below 0.1 eV the neutron flux is underestimated by over an order of magnitude if upscattering is not included. The energy of 3 eV was chosen as a cutoff, indicated by the red line.

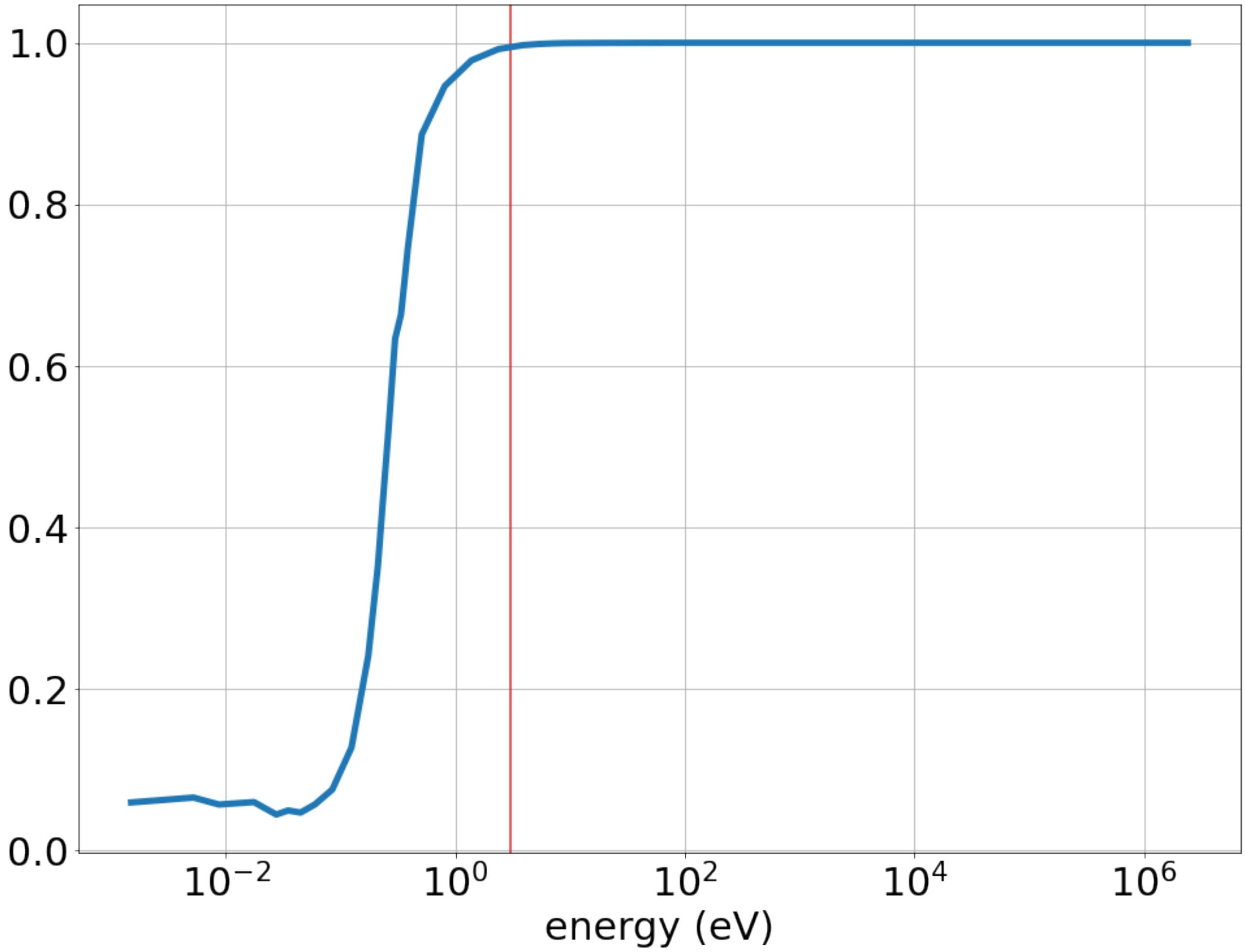


Figure 4.9.: Ratio of SPARC without upscattering and SPARC with upscattering from Fig. 4.8. The deviation between SPARC with and without upscattering is negligible for energies over 3 eV. However, without upscattering, the neutron flux for energies below 0.1 eV is underestimated by over one order of magnitude.

The neutron fluxes were then computed with upscattering included only below 3 eV and upscattering included for all energies. The relative deviation between the results was only 0.13 %, while the performance improved by a factor of 1.8, which is a favorable trade-off.

5. Computation of Partial Cross Sections for PGNAA

In this chapter, we will demonstrate the applicability of the neutron transport code SPARC, which was introduced in the previous chapter, for the computation of partial (n,γ)-cross sections. The partial (n,γ)-cross sections are an important quantity for the evaluation of large sample PGNAA measurements, and the neutron flux is a necessary parameter for their computation. The setup we investigate is the measurement facility QUANTOM, which is described in detail in Sec. 1.3. The measurement facility QUANTOM is being developed for the material analysis of radioactive waste packed up in 200-l-drums. It enables a spatially resolved elemental analysis based on prompt gamma neutron activation analysis (PGNAA). While the complex geometry of the measurement facility suggests the use of Monte Carlo methods, for the practical application their performance is insufficient. The inversion method iterates over the material composition and a full neutron transport simulation is necessary for each measurement position in every iteration, accounting for hours of simulation time per iteration. To ensure that the simulations can be carried out within an acceptable time frame for a large scale industrial application, we use the deterministic neutron transport code SPARC which is specially developed for this purpose based on the $\mathrm{SP_N}$ approximation of the linear Boltzmann equation.

Since the neutron spectra in the measurement facility are not strictly thermal, the used diffusive $\mathrm{SP_N}$ approximations are not perfectly suited to determine neutron transport at the highest order of accuracy. On the other hand, computational efficiency, user-friendly implementation and the possibility to handle complex 3D geometries are important factors when choosing the deterministic transport model. This can be achieved using SPARC for the simulation of neutron transport. The question arises, if the approximation in the neutron transport model still allows a calculation of the partial cross sections at a sufficient level of accuracy. Therefore, in this chapter, we study the applicability of diffusive $\mathrm{SP_N}$ approximations in partial cross sections for PGNAA in the geometrical setting of the facility. In a simulation study we consider four typical matrix materials and compare cross sections for all elements of the periodic table to reference results obtained by Monte-Carlo simulations.

To the authors knowledge, the only other approach to use a deterministic multi-group code for PGNAA was described in [14], where the neutron transport simulations were carried out with Attila [72]. However, the main focus there lies on a new method for generating multi-group neutron-photon cross-sections which is verified for a series of simple benchmark problems. The applicability of an approximate neutron transport model in the calculation of partial (n,γ)-cross sections was not a subject of the study.

The structure of this chapter is as follows. In Section 5.1 we will introduce partial (n,γ)-cross sections. Section 5.2 is dedicated to the numerical method, geometrical setup and implementation details. Computational results regarding the applicability of the approximate neutron transport model in computing the quantity of interest are presented in Section 5.3, for which simulation studies with four typical matrix material where carried out.

5.1. Partial Cross Sections

In order to define the partial (n,γ)-cross sections, it is necessary to introduce two auxiliary quantities first, the neutron energy distribution and the isotopic capture cross section. For a given neutron flux $\phi(E,x)$ and a domain $D \subset \mathbb{R}^3$ the neutron energy distribution can be defined as

$$\upsilon(E) = \frac{\int_D \phi(E,x)dx}{\int_{E_{min}}^{E_{max}} \int_D \phi(E,x)dxdE}. \tag{5.1}$$

For an incident neutron flux that is not mono-energetic the isotopic capture cross section is defined as

$$\sigma_{iso} = \int_{E_{min}}^{E_{max}} \sigma_{a,iso}(E)\upsilon(E)dE, \tag{5.2}$$

where $\sigma_{a,iso}(E)$ is the energy-dependent microscopic absorption cross section for the given isotope. The partial (n,γ)-cross sections are then defined as

$$\sigma_{E_{\gamma,l}} = I_{E_{\gamma,l}} \cdot \theta_{iso} \cdot \sigma_{iso}, \tag{5.3}$$

where σ_{iso} is the isotopic capture cross section, θ_{iso} is the natural abundance of the given isotope in the element and $I_{E_{\gamma,l}}$ is the so called intensity of the gamma line. It can be interpreted as the number of photons with energy $E_{\gamma,l}$ emitted per capture. According to [21], the energy dependence of the intensity in the thermal energy range is negligible. The argumentation found there utilizes previous work on neutron capture models, based either on statistical theory [8] or direct capture [54, 74, 75]. The partial (n,γ)-cross sections can be characterized as the probability of producing a gamma ray with energy $E_{\gamma,l}$ per atom of the given element per neutron capture. Since the gamma lines are characteristic for each isotope, PGNAA can in principle identify single isotopes. In most applications including ours this is not needed, a reconstruction of the element masses is sufficient. The natural isotope ratio applies also in our case so the natural abundance θ is incorporated in Eq. (5.3) to make sure that the partial (n,γ)-cross sections directly relate the measured values and the masses of the elements. A more detailed introduction into this topic can be found in [81]. The values for the isotopic capture cross sections σ_{iso}, the natural abundances θ_{iso} and the intensities $I_{E_{\gamma,l}}$ that are used in this Chapter are extracted from the ENDF/B-VIII.0 database [13].

5.2. Methodology and Model Description

In order to compute the partial (n,γ)-cross sections, it is necessary to compute the neutron flux first. In this chapter, we solve the SP$_1$ and SP$_3$ approximation of the neutron transport equation (see Subsec. 3.2.3) with SPARC. Details of the implementation are described in Subsec. 4.1. The multigroup form of the SP$_3$-equations in the case of an isotropic source and isotropic scattering read

$$\begin{aligned} &-\nabla \cdot \frac{1}{3\Sigma_{t,g}} \nabla \left(\phi_{0,g}(x) + 2\phi_{2,g}(x)\right) + (\Sigma_{t,g} - \Sigma_{s,gg})\phi_{0,g}(x) = S_g(x) \\ &-\nabla \cdot \frac{9}{35\Sigma_{t,g}} \nabla \phi_{2,g} + \Sigma_{t,g}\phi_{2,g} = \frac{2}{5}\left((\Sigma_{t,g} - \Sigma_{s,gg})\phi_{0,g} - S_g(x)\right), \end{aligned} \tag{5.4}$$

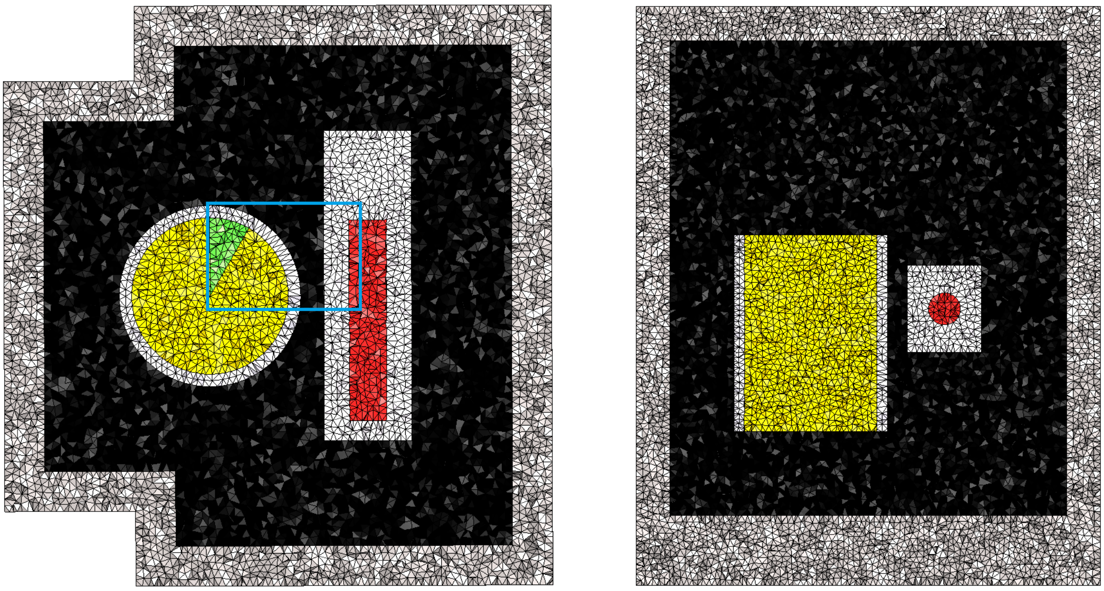
where $\phi_{0,g}(x)$ and $\phi_{2,g}(x)$ are the zeroth respectively the second Legendre moment of the neutron flux in the gth energy group. The source term of the gth group is given by

$$S_g(x) = \sum_{g \neq g'} \Sigma_{s,gg'} \phi_{0,g'}(x) + Q_g(x). \tag{5.5}$$

and the boundary conditions read

$$\begin{aligned} \frac{1}{2}\phi_{0,g} + \frac{1}{3\Sigma_{t,g}}(n \cdot \nabla)(\phi_{0,g} + 2\phi_{2,g}) &= -\frac{5}{8}\phi_{2,g} \\ \frac{3}{8}\phi_{2,g} + \frac{9}{35\Sigma_{t,g}}(n \cdot \nabla)\phi_{2,g} &= \frac{3}{40}\phi_{0,g}, \; x \in \partial D. \end{aligned} \tag{5.6}$$

We determine the weak solution of the system with the Finite Element method (see [3, 63]), so we have $\phi_{0,g} \in H^1(D)$ and $\phi_{2,g} \in H^1(D)$. The SP$_1$-equation, also widely known as neutron diffusion equation, is obtained from Eqs. (5.4) by setting the higher order moments to zero. In this section, we again use the group structure with 44 energy groups from [99], so $G = 44$. In order to solve the SP$_3$ equations for this specific application, the neutron transport code SPARC was developed, which is discussed in detail in Sec. 4. The multigroup cross sections (see Subsec. 3.1.2) $\Sigma_{t,g}$ $(1 \leq g \leq 44)$ and $\Sigma_{s,gg'}$ $(1 \leq g \leq 44,\ 1 \leq g' \leq 44)$ are computed using the OpenMC Monte Carlo code [85, 10]. OpenMC also yields the neutron flux which is used for benchmarking purposes. The computation of the neutron flux using Monte Carlo is described in [55]. It should be noted that OpenMC solves the full linear Boltzmann equation (Eq. 2.1) without using the multigroup and SP$_\text{N}$ approximations like implemented in SPARC.



(a) Measurement facility: horizontal cross section **(b)** Measurement facility: vertical cross section

Figure 5.1.: Slice through the Finite Element mesh used by SPARC. Depicted are the drum content (yellow), a sector of the drum (green), the neutron generator (red), the moderating graphite (black), the outer shielding (grey) and air gaps within the facility (white). Fig. 5.2 shows the picture section within the blue frame in greater detail.

The geometric model used for both SPARC and OpenMC is a simplification of the CAD representation of the measurement facility. Figure 5.1 depicts a horizontal and a vertical cross

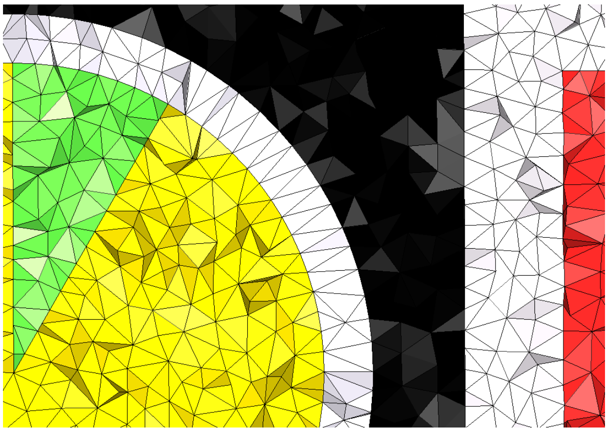


Figure 5.2.: Picture section of Fig. 5.1a in greater detail. The variation of the finite elements in size and form ensures that the mesh preserves the geometric properties of the model. The slice does not divide single finite elements, and since the mesh is unstructured, not all finite elements lie on the same height. The varying brightness of the finite element facets indicates the different angles between the image plane and the facets.

section through the Finite Element mesh used by SPARC. It shows the most important components of the facility, namely the neutron generator (red), the graphite moderator (black), the drum (yellow) and the outer radiation shielding (gray) as well as some air gaps within the facility. Figure 5.1a also shows a drum sector (green). As a future improvement it is planned to divide the whole drum into 48 sectors with possibly different filling materials in order to enable inhomogeneous drum fillings. The model has a length of 219 cm, a width of 232 cm and a height of 273 cm. In reality, the facility is larger, however the parts outside the radiation shielding have a negligible influence on the neutron flux in the drum. Monte Carlo Simulations show that only 0.55 % - 0.56 % of the neutrons leave the modeled parts of the facility. The shielding consists of highly absorbing borated polyethylene, so possible room shine effects with neutrons that leave the facility and then return are insignificant. The outer parts of the facility can therefore be neglected for this study. The characteristic length of each finite element is not larger than 2.5 cm, resulting in a mesh of the measurement facility that consists of 3.88 mio. finite elements. The exact size of the finite elements varies to adapt better to the geometry, resulting in an unstructured mesh. This ensures that the discontinuous transitions in the cross sections on material interfaces are represented correctly. Fig. 5.2 depicts the picture section in the blue frame in Fig. 5.1a in greater detail, showing how the finite elements vary in size and form to adapt to the material boundaries.

Since there is no significant amount of transuranic elements present in the measurement facility, we can assume that there is no fission, but a mono-energetic point source with an energy of 2.45 MeV, the energy of the deuterium-deuterium fusion in the neutron generator. The neutrons are emitted isotropically. It is also remarkable that the huge amount of graphite used as a moderator causes strong upscattering even at room temperature, the neutron flux with and without upscattering differ by an order of magnitude for thermal neutron energies.

In order to represent a wide variety of possible drum contents, the simulations are carried out with four different filling materials. The chosen materials were concrete, polyethylene, cast iron and lead. The exact material composition can be found in [71] under the material numbers 97, 248, 162 and 271. The first two materials have a relatively low density. They were chosen since concrete is a filling that is used very often, while polyethylene represents light mixed wastes quite well. Cast iron and lead represent regular metallic and heavy metallic materials, respectively. In the following the drum consists mainly of one of these materials. They serve as matrix materials, and small amounts of other materials that are embedded in this matrix. It is assumed that the small amounts of other materials have a negligible influence on the neutron flux, so the transport calculations are carried out taking only the matrix material into account.

After the neutron transport simulations are carried out, the isotopic and the partial (n,γ)-cross sections are calculated for the each chemical element using the computed neutron energy distribution densities. In the following, the results for the γ-lines with the highest σ_{E_γ} for the respective element are shown. It is reasonable to concentrate on these, since these are typically the best detectable and therefore most relevant ones for the practical application.

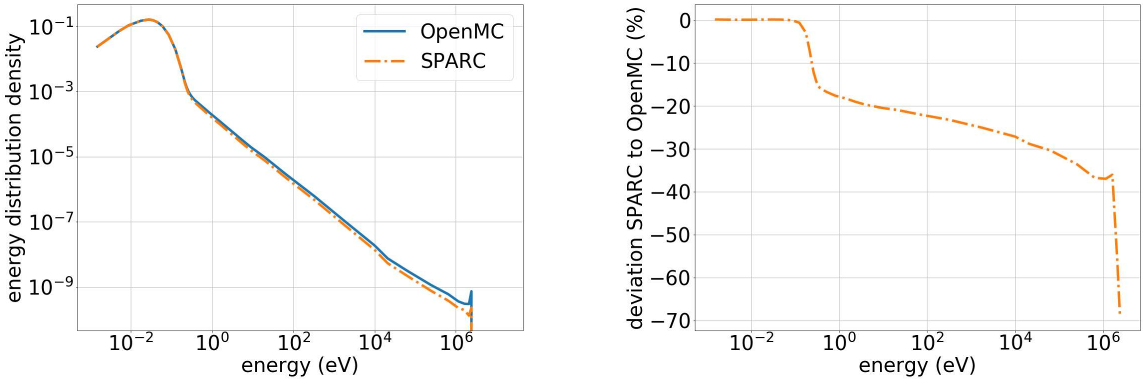
In this analysis, the isotopic and therefore partial (n,γ)-cross sections are not computed for the full energy range, but for three distinct energy ranges. The first energy range includes neutrons with energies smaller than 0.25 eV and will be named thermal region in the following discussion. It is the most important energy range, since the high microscopic cross sections in this region in combination with a neutron flux that is nearly completely thermalized ensure that these energies have a dominant influence on the final result. The second energy range will be named resonant energy range and include neutrons with energies between 0.25 eV and 1 MeV. The inclusion of the resonant energy range is necessary in this case since the neutron flux is not fully thermalized. For experiments at research reactors, where the neutron flux is essentially a thermal one [61], resonant neutrons can usually be neglected. Neutrons with energies greater than 1 MeV, here named fast neutrons, have an insignificant influence on the final result. The absorption cross sections in this region are by orders of magnitude smaller than the cross sections in the thermal energy range, while the fast neutron flux in the facility is also significantly smaller than the thermal neutron flux. In the following, the integration boundaries in Eqs. (5.1) and (5.2) are adjusted in order to compute the cross sections for the thermal and resonant energy range, while the fast energy range is neglected for the reasons mentioned above. Consequently, to compute the thermal partial (n,γ)-cross sections $\sigma_{E_\gamma,therm}$, we use $[E_{min}, E_{max}] = [0, 0.25]$ within Eqs. (5.1)-(5.3). For the resonant partial (n,γ)-cross sections $\sigma_{E_\gamma,res}$, we use $[E_{min}, E_{max}] = [0.25, 10^6]$ within Eqs. (5.1)-(5.3).

5.3. Results and Discussion

In this section, the results for the partial (n,γ)-cross sections (see Eq. 5.3) will be discussed for the application cases introduced in Section 5.2. The necessary neutron energy distributions (see Eq. 5.1) were computed with the SP$_\text{N}$ neutron transport code SPARC and with the OpenMC Monte Carlo code. The purpose of this analysis is to show that the SP$_\text{N}$ approximation, although having a higher model error than Monte Carlo, is nevertheless appropriate for computing σ_{E_γ} for the evaluation of PGNAA spectra. In the following, the OpenMC solutions serve as a reference solutions for SPARC. All so-called errors are to be understood as deviations of SPARC from OpenMC. The results have been divided in two subgroups, Subsec. 5.3.1 concentrates on materials with lower density, while dense metallic materials will be discussed in Subsec. 5.3.2. In Subsec. 5.3.3 the performance of SPARC and OpenMC is compared.

5.3.1. Light materials

Figure 5.3a shows the neutron energy density for a drum filled with concrete, computed with both SPARC and OpenMC. In Figure 5.3b the relative deviation of the SPARC results in comparison to OpenMC is depicted. For thermal energies, the results for both codes are almost identical. The deviation gets larger for higher energies, it reaches its maximum for fast neutrons where SPARC underestimates the OpenMC flux by over 60 percent. For the application at hand the quantity of interest is not the neutron flux, but the partial (n,γ)-cross sections. However, these results show that the SP_N approximation suffers from relatively large errors if used for a full neutron transport simulation in this application case.



(a) absolute values **(b)** relative deviation of SPARC to OpenMC

Figure 5.3.: Neutron energy density $\upsilon(E)$ within a drum filled with concrete calculated with the neutron flux computed with OpenMC and SPARC.

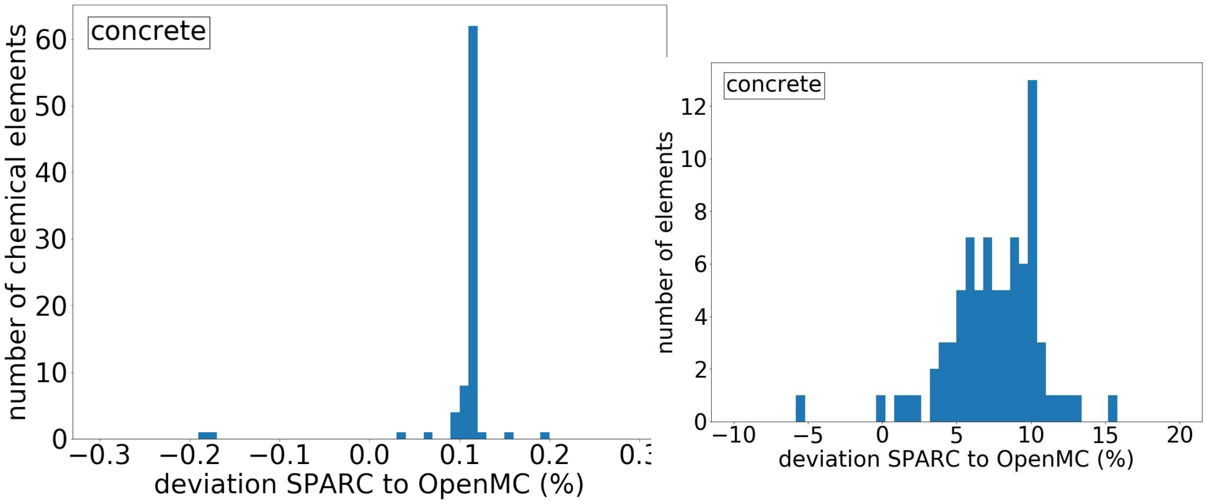
Figure 5.4a depicts the relative deviation of the thermal partial (n,γ)-cross sections computed with SPARC in comparison to OpenMC as a histogram. The average deviation of the values is only 0.10 %. It jumps into the eye that nearly all elements lie in the same bin. This is plausible since the microscopic capture cross sections for nearly all elements show the same functional behaviour in this energy range, namely one divided by the square root of the energy. In the thermal energy range, the partial (n,γ)-cross sections computed with SPARC and OpenMC are therefore almost indistinguishable. The results for the resonant partial (n,γ)-cross sections are shown in Figure 5.4b. Due to the higher model error on the neutron flux computed by SPARC, the average deviation for the resonant partial (n,γ)-cross sections also higher, namely 7.55 %. Since every element has a unique resonance structure and therefore unique cross sections in this energy range, the deviations for the different elements are spread out over a broad energy range, a strong contrast when compared with the results for thermal energies.

When the integration in Eqs. 5.1 and 5.2 are, as mentioned before, adjusted to compute the cross sections for the thermal and resonant energy range separately, one sees that the resonant energy density $\upsilon_{res}(E)$ computed with SPARC overestimates OpenMC for smaller energies and underestimates OpenMC for higher energies. Since smaller energies have a higher impact on the resulting partial cross sections due to higher cross sections and and a higher absolute flux, an overestimation for smaller energies leads directly to an overestimation of the quantity of interest and therefore the assymetry as seen in figure 5.4b.

In order to judge the suitability of SPARC to compute partial (n,γ)-cross sections for the application at hand, it is necessary to introduce the reaction rate R, given by

$$R = \sigma_{E_\gamma,therm}\Phi_{therm} + \sigma_{E_\gamma,res}\Phi_{res}. \tag{5.7}$$

In the equation above $\sigma_{E_\gamma,therm}$ and $\sigma_{E_\gamma,res}$ are the thermal and resonant partial (n,γ)-cross sections, while Φ_{therm} and Φ_{res} are the thermal respectively resonant neutron flux. The fluxes



(a) Thermal partial (n,γ)-cross sections **(b)** Resonant partial (n,γ)-cross sections

Figure 5.4.: Relative deviation of partial (n,γ)-cross sections for a drum filled with concrete computed with SPARC from OpenMC.

for this study are simulated, while they will be determined by measurement devices in the final evaluation process.

The relative deviation of the reaction rates computed with the partial (n,γ)-cross sections from SPARC in comparison to the values computed with OpenMC for a drum filled with concrete is depicted in Figure 5.5. The average deviation is only 0.62 %, with all outliers limited to the low single-digit percentage range. This result shows that despite the model error of SPARC when describing non-thermal neutrons, the (n,γ)-reaction rates can be computed with a comparatively small error. Since the deviation between SPARC and OpenMC is much smaller than the expected measurement error of about 5-10 %, the computation of the partial (n,γ)-cross sections can be carried out with the SP_1 method as implemented in SPARC as an alternative to the costly Monte Carlo computations.

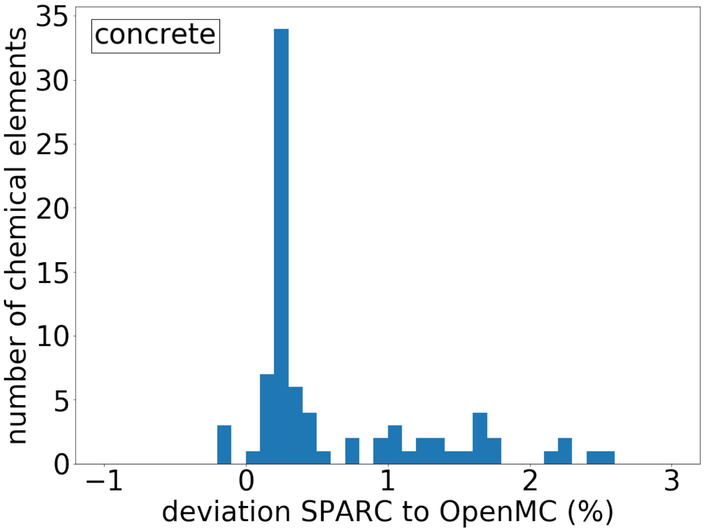
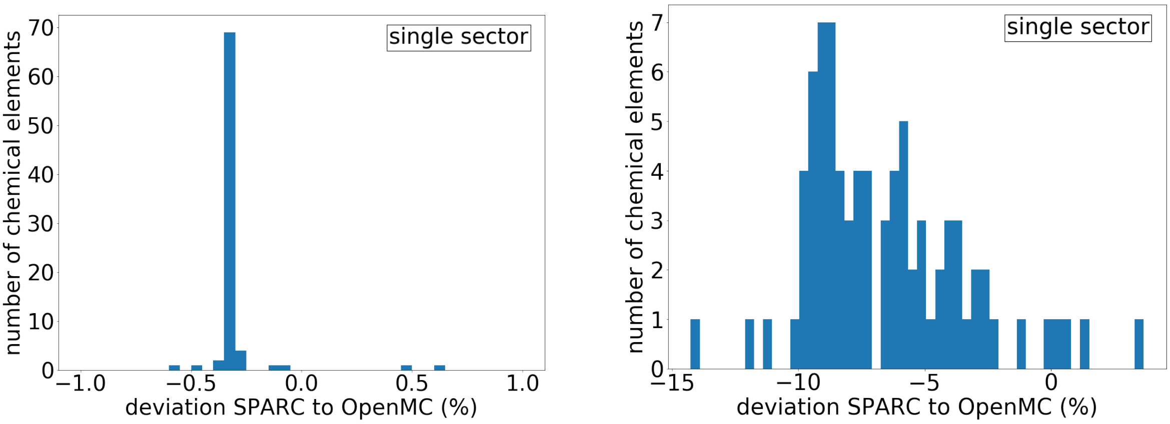


Figure 5.5.: Relative deviation of reaction rates computed with the partial (n,γ)-cross sections from SPARC in comparison to the values computed with OpenMC for a drum filled with concrete.

In order to ensure that σ_{E_γ} can be computed locally and not only averaged over the complete drum, the same computations have been carried out for a single sector, as depicted in Figure 5.1. The partial (n,γ)-cross sections for the single sector filled are depicted in Figure 5.6. The average error on the (n,γ)-cross sections in the thermal energy region is -0.31 %. Similarly to the complete drum, almost all elements lie in the same bin, although the relative deviation to OpenMC is slightly larger. In the resonant energy region, the average error is -6.57 %. While for the whole drum the resonant partial (n,γ)-cross sections are overestimated for almost all elements, the resonant partial (n,γ)-cross sections are underestimated for the single sector. The resulting reaction rates for the single sector in depicted in Fig. 5.7. The average error on the reaction rates was -0.56 %. Thus, one sees a slight underestimation of the reaction rate, whereas for the whole drum one sees a slight overestimation.



(a) Thermal partial (n,γ)-cross sections **(b)** Resonant partial (n,γ)-cross sections

Figure 5.6.: Relative deviation of partial (n,γ)-cross sections for a single sector of a drum filled with concrete computed with SPARC from OpenMC.

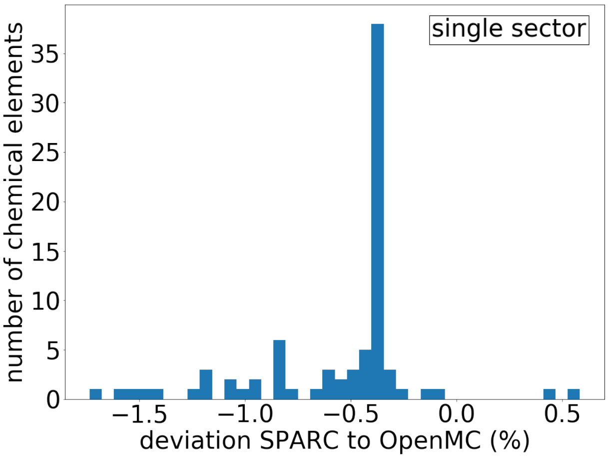
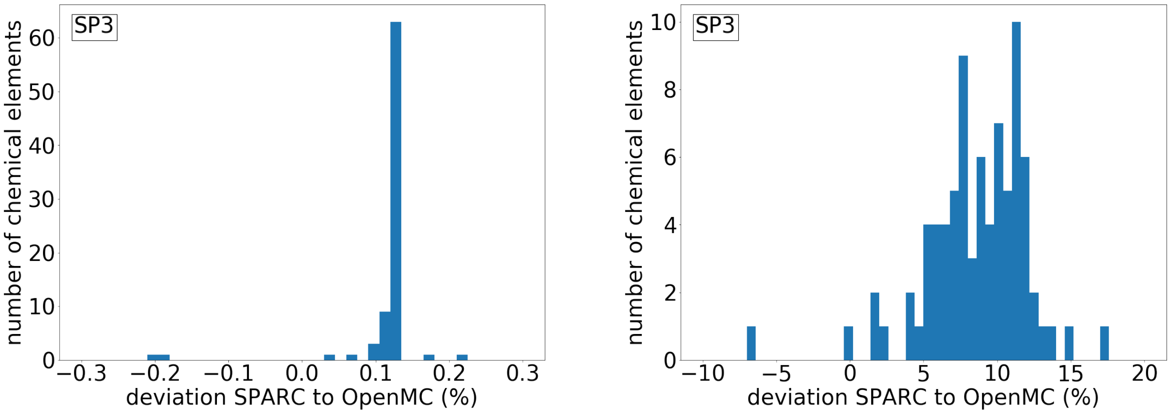


Figure 5.7.: Relative deviation of reaction rates computed with the partial (n,γ)-cross sections from SPARC in comparison to the values computed with OpenMC for a single sector of a drum filled with concrete.

Since SPARC also enables SP_3 computations, it was investigated whether a further improvement of the results using the higher order approximation is possible. The results are depicted in Fig. 5.8. With the SP_3 approximation, the error on the thermal partial (n,γ)-cross sections is 0.11 %, while it is 8.59 % for resonant energies. The results for the reaction rates are shown in Fig.

5.9. The error on the reaction rates is 0.71 %. This error is slightly higher than the error obtained using the SP_1 approximation. However, since the deviations between both results are much smaller than the model error, one should see both methods as yielding indifferent results. Hereinafter, only the SP_1 approximation will be used due to the better performance.



(a) Thermal partial (n,γ)-cross sections **(b)** Resonant partial (n,γ)-cross sections

Figure 5.8.: Relative deviation of partial (n,γ)-cross sections for a drum filled with concrete computed with SPARC using the SP_3 approximation from OpenMC.

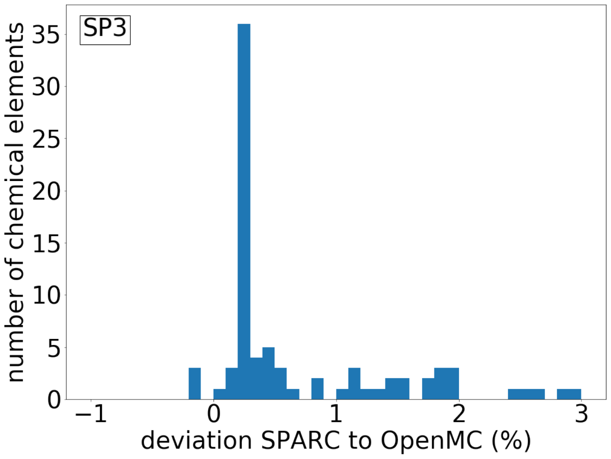


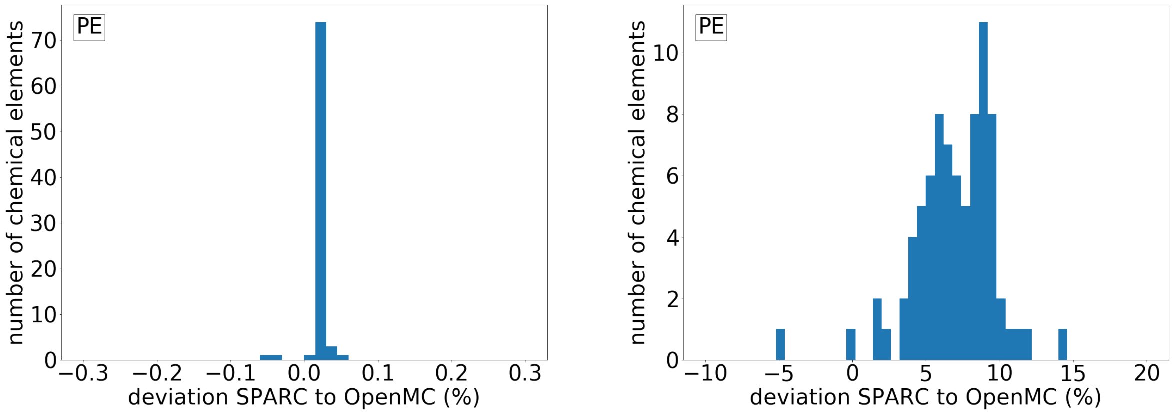
Figure 5.9.: Relative deviation of reaction rates computed with the partial (n,γ)-cross sections from SPARC using the SP_3 approximation in comparison to the values computed with OpenMC for a drum filled with concrete.

In this study, the group structure with 44 energy groups from SCALE [99] is used for the energy discretization. In order to test its suitability, a convergence study has been performed where the SPARC results are compared to the OpenMC results for the respective group structure. The results are shown in Table 5.1. The energy groups structures with 22 groups and 88 groups are generated via uniting two adjacent energy groups and dividing one group in two in a way that the logarithmized width of both groups is the same, respectively. The error on the thermal and resonant cross partial (n,γ)-cross sections, $\Delta\sigma_{E_\gamma,therm}$ and $\Delta\sigma_{E_\gamma,res}$ decreases with a finer discretization, as well as the error on the reaction rates ΔR. However, while the error is significantly reduced when using 44 groups instead of 22, the improvement by using 88 groups instead of 44 is minor. This shows that 44 groups are a good choice, further refinement does not reduce the error enough to justify the additional computational effort.

Table 5.1.: Influence of the energy discretization for a fixed mesh

Energy Groups	$\Delta\sigma_{E_\gamma,therm}$	$\Delta\sigma_{E_\gamma,res}$	ΔR
22	3.40 %	19.28 %	4.51 %
44	0.10 %	7.56 %	0.62 %
88	0.10 %	7.33 %	0.61 %

The results for the other light material, polyethylene, are very similar to the results for concrete. As can be seen Fig. 5.10, the results are nearly the same. For PE, there are less outliers in the thermal energy region, resulting in a smaller deviation of the SPARC results in comparison to OpenMC, with an average deviation of 0.03 % for $\sigma_{E_\gamma,therm}$. For the resonant energy region, the results for PE and concrete are almost identical, the deviation of the SPARC results in comparison to OpenMC is with 6.90 % for $\sigma_{E_\gamma,res}$ slightly smaller than for concrete. Fig. 5.11 depicts the relative difference in the reaction rates computed with SPARC in comparison to OpenMC. The results for PE are again similar to those for concrete, but with a smaller error of just 0.39 %.



(a) Thermal partial (n,γ)-cross sections

(b) Resonant partial (n,γ)-cross sections

Figure 5.10.: Relative deviation of partial (n,γ)-cross sections for a drum filled with PE computed with SPARC from OpenMC.

5.3.2. Metallic Materials

The partial (n,γ)-cross sections for lead are depicted in Figure 5.12. For thermal neutron energies, the average error on the cross sections is 0.87 %. Nearly all elements are concentrated in a small area around the average error, yielding a picture similar to concrete, although with a larger error. For resonant neutron energies the error is 14.87 %, with the results for the different elements distributed over a broad area. Figure 5.13 depicts the error on the reaction rates. The average error is 3.74 %, with many elements concentrated in the same bin, similar to concrete. In general, the error is larger than for concrete, but this is expected. Lead has a much higher absorption cross section, leading to a higher error of the neutron fluxes computed with the SP_1 approximation and therefore the partial (n,γ)-cross sections and reaction rates.

For cast iron, the partial (n,γ)-cross sections are depicted in Fig. 5.15. In the thermal energy range, the results are qualitatively similar to the results for lead. However, the errors are larger. The error on the partial (n,γ)-cross sections for the thermal energy range is 1.70 %. For the

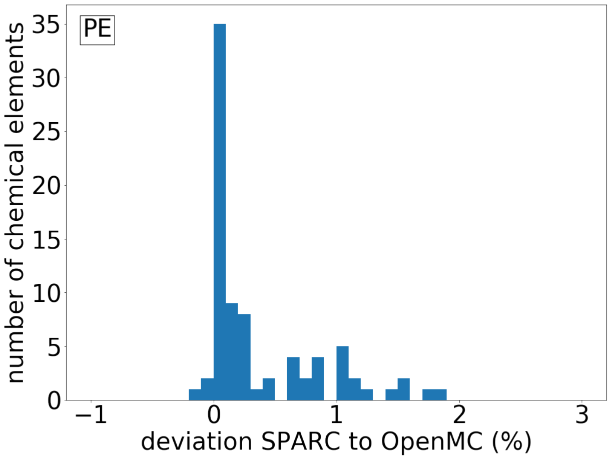
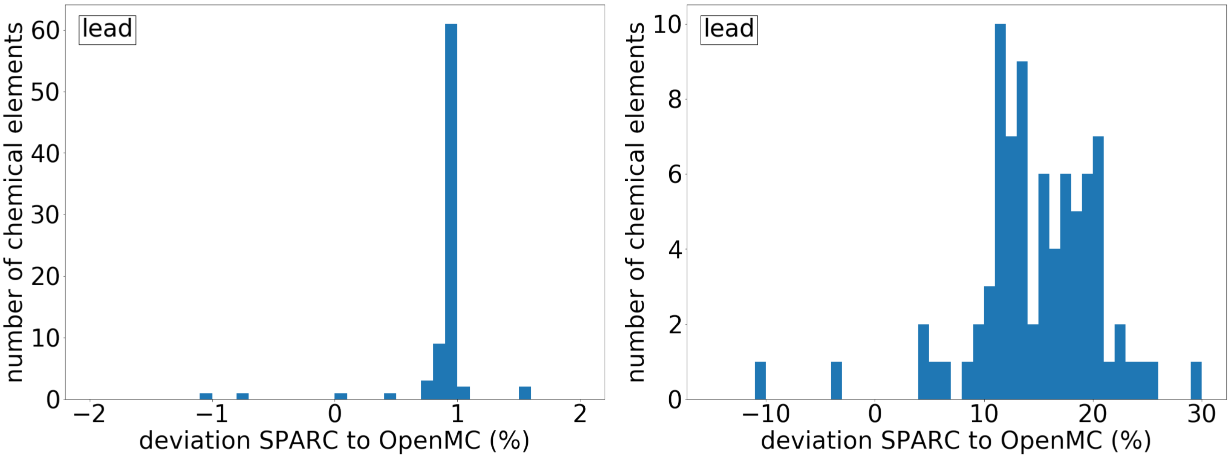


Figure 5.11.: Relative deviation of reaction rates computed with the partial (n,γ)-cross sections from SPARC in comparison to the values computed with OpenMC for a drum filled with PE.



(a) Thermal partial (n,γ)-cross sections

(b) Resonant partial (n,γ)-cross sections

Figure 5.12.: Relative deviation of partial (n,γ)-cross sections for a drum filled with lead computed with SPARC in comparison to OpenMC.

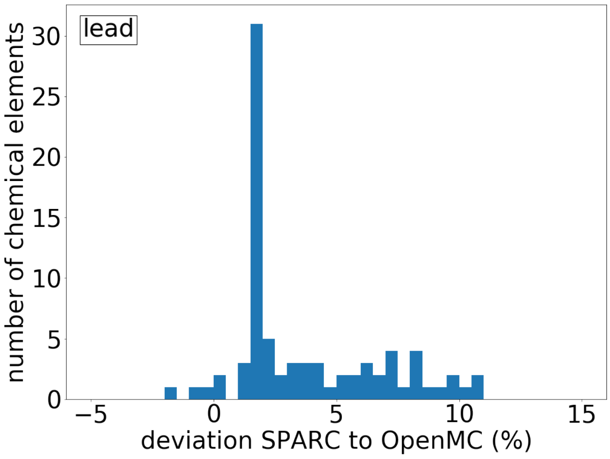
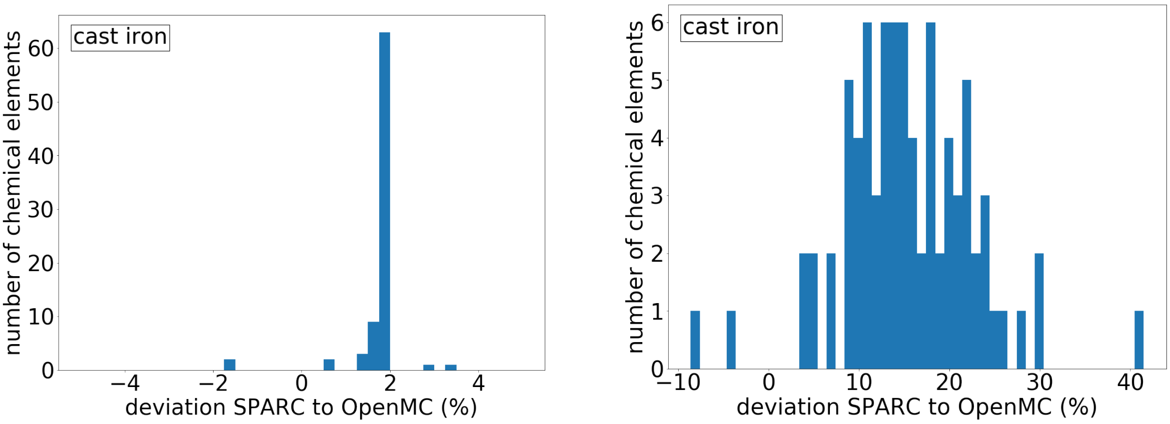


Figure 5.13.: Relative deviation of reaction rates computed with the partial (n,γ)-cross sections from SPARC in comparison to the values computed with OpenMC for a drum filled with lead.



(a) Thermal partial (n,γ)-cross sections

(b) Resonant partial (n,γ)-cross sections

Figure 5.14.: Relative deviation of partial (n,γ)-cross sections for a drum filled with cast iron computed with SPARC in comparison to OpenMC.

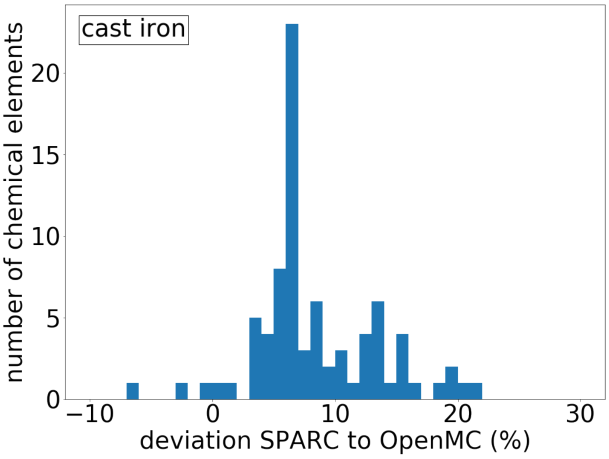


Figure 5.15.: Relative deviation of reaction rates computed with the partial (n,γ)-cross sections from SPARC in comparison to the values computed with OpenMC for a drum filled with cast iron.

resonant energy range, the error on the partial (n,γ)-cross sections is 15.40 %. The reaction rates are depicted in Fig. 5.15. The error is 8.35 % for the reaction rates and therefore much higher than for the other tested materials.

5.3.3. Performance Comparison

The computations were executed on a cluster consisting of AMD Ryzen Threadrippers 2950X, each with 32 cores. Since the OpenMC computations were executed on nine nodes to reduce the total runtime while the SPARC computations were executed on one node, the runtimes were converted to core hours. OpenMC as a Monte Carlo code scales nearly linear, hence the inter-node communication has only a minor influence on the runtime and the core hours are well comparable. The runtimes are shown Table 5.2. The minimum speed up factor of SPARC in comparison to OpenMC is 20 for polyethylene, getting as large as 32 for lead. This shows that SPARC features a significant performance improvement over OpenMC as it is needed for this application.

Table 5.2.: Runtime comparison for the neutron transport computations in core hours.

material	OpenMC runtime	SPARC runtime	speed up factor
concrete	824.32 h	34.97 h	23.57
polyethylene	803.68 h	40.57 h	19.81
cast iron	787.20 h	26.50 h	29.71
lead	918.64 h	28.64 h	32.08

6. Adjoint Neutron Transport with SPARC

In this chapter, the implementation and verification of the adjoint solver included in SPARC is described. The adjoint solver will be needed to solve the parameter estimation problem described in the next chapter. Since it was shown in chapter 5 that the SP_3 approximation does not improve the results enough to justify the additional computational effort, we will only introduce the adjoint SP_1 equation.

This chapter is organized as follows: in Sec. 6.1 we will derive the one-group and the multigroup adjoint SP_1 equation. In Sec. 6.2, the implementation of the adjoint SP_1 solver in SPARC will be discussed. The verification of the adjoint solver with manufactured solutions is described in Sec. 6.3. We conclude the chapter with a benchmark with the ARTEMIS code in Sec. 6.4.

6.1. Derivation of Adjoint Equations

For the sake of simplicity, we rely in this chapter on formal computations in L^2 spaces, since an exact analysis based on operators in suitable Banach spaces would be much more complex. Nevertheless, this is sufficient to derive the correct adjoint equations (see Sec. 2.10 in [95]).

For (potentially different) Hilbert spaces U and V and a bounded, linear operator $A : U \to V$ the adjoint operator $A^* : V \to U$ can be characterized by the property:

$$\langle Au, v\rangle = \langle u, A^*v\rangle \quad \forall u \in U, v \in V. \tag{6.1}$$

Here $\langle , \rangle$ denotes a scalar product. In the following, we will use the scalar product

$$\langle u, v\rangle = \int_D u(x)v(x)dx. \tag{6.2}$$

This is the standard inner product in the Hilbert space $L^2(D)$, where $D \subset \mathbb{R}^3$ is bounded.

6.1.1. Adjoint One-Group SP_1 Equation

In the following, we will derive the adjoint one-group SP_1 equation and the boundary conditions that will be needed in the next chapter. We formally define the operator $T : H^1(D) \to H^1(D)$ as

$$T\phi(x) = -\nabla \cdot \left(\frac{1}{3\Sigma_t(x)}\nabla\phi(x)\right) + \Sigma_0(x)\phi(x), \tag{6.3}$$

so that we can write the SP_1 equation (Eq. 3.44) as

$$T\phi = Q. \tag{6.4}$$

Then we have:

$$\langle T\phi, \phi^* \rangle = \int_D (T(x)\phi(x))\phi^\star(x)dx. \tag{6.5}$$

By formal integration by parts and using Marshak boundary conditions (Eq. 3.45) as well as assuming that ϕ and ϕ^* are smooth, we obtain

$$\langle T\phi, \phi^* \rangle = \int_D \phi(x)(T(x)\phi^\star(x))dx - \int_{\partial D} \frac{1}{2}\phi(x)\phi^\star(x) - \frac{1}{3\Sigma_t(x)}\phi(x)\frac{\partial}{\partial n}\phi^\star(x)ds. \tag{6.6}$$

The boundary terms in Eq. (6.6) need to vanish. This is only the case if the relation

$$\frac{1}{2}\phi^\star(x) + \frac{1}{3\Sigma_t(x)}\frac{\partial}{\partial n}\phi^\star(x) = 0 \tag{6.7}$$

is fulfilled on the boundary. We can now write the right hand-side Eq. (6.6) as a scalar product:

$$\langle T\phi, \phi^* \rangle = \int_D \phi(x)(T(x)\phi^\star(x))dx = \langle \phi, T\phi^* \rangle \tag{6.8}$$

Consequently, the operator $T : H^1(D) \to H^1(D)$ is self-adjoint.

6.1.2. Adjoint Multigroup SP$_1$ Equations

To derive the adjoint multigroup SP$_1$ equations, we use the scalar product

$$\langle u, v \rangle = \sum_{g=1}^{G} \int_D u_g(x)v_g(x)dx. \tag{6.9}$$

We define the multigroup operator

$$N_{gg'}(x)\phi_{0,g'}(x) = \delta_{gg'}T_{g'}(x)\phi_{0,g'}(x) - \left(\sum_{g'=0}^{G} \Sigma_{s,gg'}(x)\phi_{0,g'}(x) \right) \tag{6.10}$$

and the diffusion-advection operator

$$T_{g'}\phi_{0,g'}(x) = \left(-\nabla \cdot \frac{1}{3\Sigma_{t,g'}}\nabla + \Sigma_{t,g'} \right) \phi_{0,g'}(x). \tag{6.11}$$

Then the multigroup form of the SP$_1$ equation reads:

$$\sum_{g'=1}^{G} N_{gg'}(x)\phi_{0,g'}(x) = Q_g(x), \quad 1 \leq g \leq G. \tag{6.12}$$

To derive the adjoint equations, we have

$$\langle N\phi_0, \phi_0^* \rangle = \sum_{g=1}^{G} \int_D \left(\sum_{g'=1}^{G} \delta_{gg'}(T_{g'}(x)\phi_{0,g'}(x))\phi_{0,g}^*(x) - \sum_{g'=0}^{G} \Sigma_{s,gg'}\phi_{0,g'}(x)\phi_{0,g}^*(x) \right) dx \tag{6.13}$$

We partially integrate Eq. (6.13) and contract indices to obtain:

$$\begin{aligned}\langle N\phi_0, \phi_0^*\rangle = &\int_D \left(\sum_{g=1}^{G} \phi_{0,g}(x)(T_g(x)\phi_{0,g}^*(x)) - \sum_{g,g'=1}^{G} \Sigma_{s,gg'}\phi_{0,g'}(x)\phi_{0,g}^*(x) \right) dx \\ &- \sum_{g=1}^{G} \int_{\partial D} \frac{1}{2}\phi_{0,g}(x)\phi_{0,g}^{\star}(x) - \frac{1}{3\Sigma_{t,g}(x)}\phi_{0,g}(x)\frac{\partial}{\partial n}\phi_{0,g}^{\star}(x)ds \qquad (6.14)\end{aligned}$$

boundary terms in Eq. (6.14) need to vanish:

$$\frac{1}{2}\phi_{0,g}^{\star}(x) - \frac{1}{3\Sigma_{t,g}(x)}\frac{\partial}{\partial n}\phi_{0,g}^{\star}(x) = 0, \quad 1 \leq g \leq G \qquad (6.15)$$

We swap the indices of the components of the group-to-group scattering matrix on the right-hand side of Eq. (6.14) so that we can write:

$$\begin{aligned}\langle N\phi_0, \phi_0^*\rangle &= \int_D \left(\sum_{g}^{G} \phi_{0,g}(x)(T_g(x)\phi_{0,g}^*(x)) - \sum_{g,g'=1}^{G} \phi_{0,g'}\Sigma_{s,gg'}(x)\phi_{0,g}^*(x) \right) dx \\ &= \int_D \left(\sum_{g}^{G} \phi_{0,g}(x)(T_g(x)\phi_{0,g}^*(x)) - \sum_{g,g'=1}^{G} \phi_{0,g'}\Sigma_{s,g'g}^{T}(x)\phi_{0,g}^*(x) \right) dx \\ &= \langle \phi_0, N^*\phi_0^*\rangle \qquad (6.16)\end{aligned}$$

Therefore we have the adjoint multigroup operator

$$N_{gg'}^*(x)\phi_{0,g'}^*(x) = \delta_{gg'}T_{g'}(x)\phi_{0,g'}^*(x) - \left(\sum_{g'=0}^{G} \Sigma_{s,gg'}^{T}(x)\phi_{0,g'}^*(x) \right) \qquad (6.17)$$

Note that the transport operator is still self-adjoint, i.e. $T_g = T_g^*$, but the group-to-group scattering matrix is transposed.

6.2. Implementation in SPARC

Since the transport operator is self-adjoint, we can use the SP_1 solver described in chapter 4 as a basis for the implementation of the adjoint SP_1 solver. Most parts can adopted from the forward solver without change, with the notable exception of the coupling between the energy groups. As described in 3.1, the lower triangular components of the group-to-group scattering matrix are usually by far larger than upper triangular components. Therefore, it is usual for neutron transport problems to iterate from lower to higher group indices in the Gauss-Seidel method (see 3.1.3). Since the transpose of the group-to-group scattering matrix enters the adjoint multigroup SP_1 equations, it is reasonable to iterate from larger to lower indices instead since we have larger upper triangular components now. We expect faster convergence with this approach.

6.3. Verification with Manufactured Solutions

In this section, we derive a manufactured solution for the adjoint SP_1 equations with Marshak boundary conditions. We investigate a two-dimensional test case with the domain $D = (-1, 1)^2$.

In order to derive the manufactured solution, we need to alter the methodology described in Subsec. 4.2.1. In the following, we first determine a function that fulfills the Marshak boundary conditions and then insert it into the adjoint SP_1 equations to derive the source term.
We use the approach

$$\phi_0(x,y) = f_x(x)f_y(y) \tag{6.18}$$

and aim to determine f_x and f_y. Inserting Eq. (6.18) into the Marshak boundary conditions (Eq. 3.45) yields for the left and right boundary

$$\begin{aligned} \frac{1}{2}f_x(1) + \frac{1}{3\Sigma_t(1)}\frac{\partial}{\partial x}f_x(1) &= 0 \\ \frac{1}{2}f_x(-1) - \frac{1}{3\Sigma_t(-1)}\frac{\partial}{\partial x}f_x(-1) &= 0. \end{aligned} \tag{6.19}$$

With the polynomial approach

$$f_x(x) = a_0 + a_1 x + a_2 x^2 \tag{6.20}$$

we obtain a linear system

$$\begin{aligned} \frac{1}{2}a_0 + \frac{1}{2}a_1 + \frac{1}{2}a_2 + \frac{1}{3\Sigma_t}(a_1 + 2a_2) &= 0 \\ \frac{1}{2}a_0 - \frac{1}{2}a_1 + \frac{1}{2}a_2 + \frac{1}{3\Sigma_t}(a_1 - 2a_2) &= 0 \end{aligned} \tag{6.21}$$

with the solution

$$\begin{aligned} a_0 &= -\left(\frac{4}{3\Sigma_t} + 1\right)a_2 \\ a_1 &= 0 \end{aligned} \tag{6.22}$$

Therefore,

$$f_x(x) = \frac{4}{3\Sigma_t} + 1 - x^2 \tag{6.23}$$

is a suitable choice for our purpose. Similar computations for the upper and lower boundary yield

$$f_y(y) = \frac{4}{3\Sigma_t} + 1 - y^2 \tag{6.24}$$

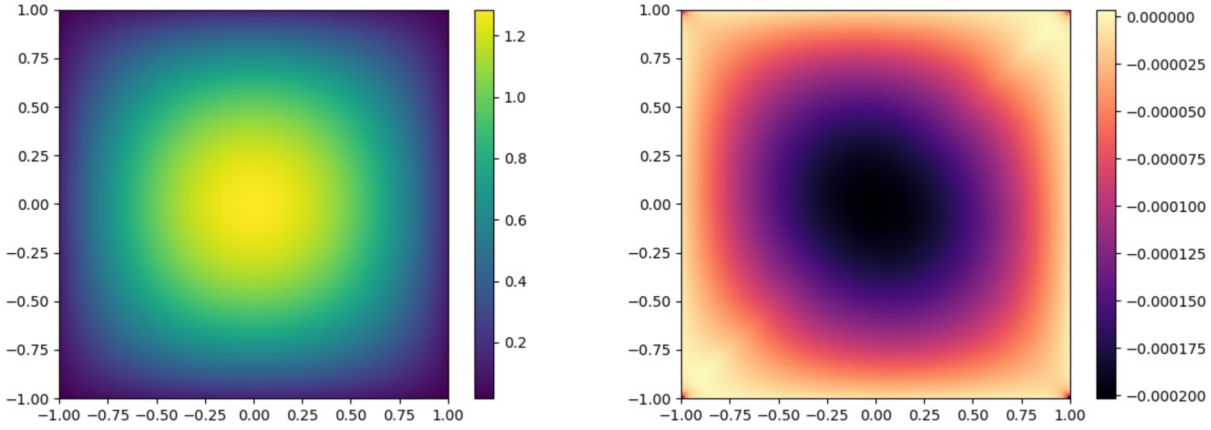
Inserting our approach $\phi_0(x,y) = f_x(x)f_y(y)$ into the SP_1 equations (Eq. 3.44) yields the source term:

$$Q(x,y) = \frac{2}{3\Sigma_t}(f_x(x) + f_y(y)) + (\Sigma_t - \Sigma_0)f_x(x)f_y(y) \tag{6.25}$$

This test case can be transformed to include two energy groups. We choose the source term $Q_g(x,y) = \frac{2}{3\Sigma_t}(f_x(x) + f_y(y))$ for $g = 1,2$ and the group-to-group scattering matrix

$$\begin{pmatrix} \Sigma_t & \Sigma_t - \Sigma_0 \\ \Sigma_t - \Sigma_0 & \Sigma_t \end{pmatrix}. \tag{6.26}$$

It can easily be checked that $\phi_{0,g}(x,y) = f_x(x)f_y(y)$ for $g = 1, 2$.
Since it does not only test the spatial convergence, but also the correct implementation of the modified Gauss Seidel algorithm, we will investigate the latter test case. We choose the values $\Sigma_t = 10$ and $\Sigma_0 = 9$. Fig. 6.1a depicts the Finite Element solution of $\phi^*_{0,1}$ for a mesh with 80000 finite elements, while the absolute error can be seen in Fig. 6.1b. The relative errors of $\phi^*_{0,1}$ and $\phi^*_{0,2}$ for different spatial resolutions are shown in Tab. 6.1. While the error for a finite element size of 0.04 is about one per mille, it can be reduced by refining the mesh. For a finite element size of 0.01, the relative error is only 0.014 % for the first group and 0.015 % for the second group.



(a) Finite Element Solution of $\phi^*_{0,1}$ **(b)** Absolute error of $\phi^*_{0,1}$

Figure 6.1.: Finite element solution of an adjoint SP1 test case generated with the method of manufactured solutions in a cubic geometry. Depicted are the results for the first energy group computed with 80000 Finite Elements, corresponding to a finite element size of 0.01.

Table 6.1.: Error of the finite element solution for an adjoint SP_1 test case with two energy groups generated with the method of manufactured solutions in a square geometry.

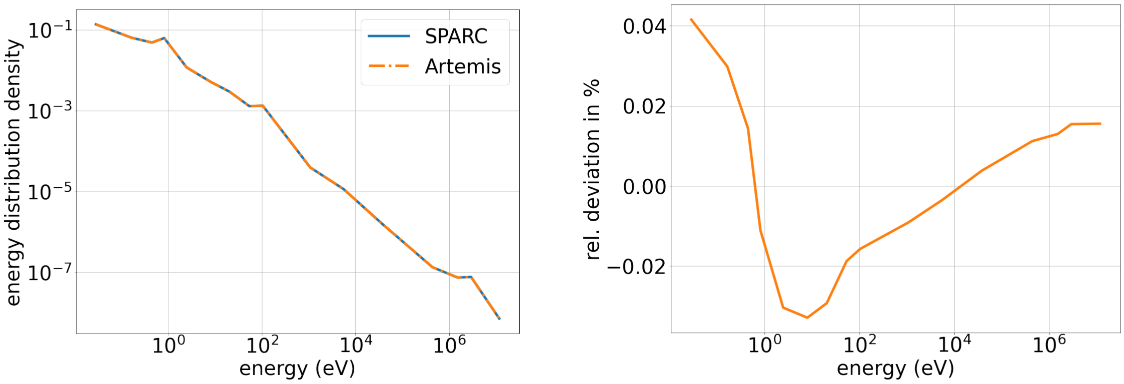
Finite Element size	0.04	0.02	0.01
rel. error $\phi^*_{0,1}$	0.115 %	0.034 %	0.014 %
rel. error $\phi^*_{0,2}$	0.116 %	0.035 %	0.015 %

6.4. Benchmark with ARTEMIS[1]

As a final step in the verification of the adjoint solver integrated in SPARC, a benchmark with ARTEMIS [44] was carried out. ARTEMIS enables the numerical solution of adjoint neutron transport equations based on an adjoint multi-group nodal diffusion formulation derived from the simplified lower-order NEM-M0 interface current approach [31]. More details and a demonstration of ARTEMIS capability to solve adjoint equations can be found in [96].

---

[1] ARTEMIS is a trademark or a registered trademark of Framatome or its affiliates, in the U.S.A. or other countries



(a) Comparison of the energy distribution densities

(b) Relative deviation of SPARC to ARTEMIS

Figure 6.2.: Benchmark of SPARC and ARTEMIS for a graphite cube with 40 cm edge length. In the computations 16 energy groups were used. A point source was placed at the center of the cube in group 16.

In the following, we compare the adjoint energy distribution densities, which we define analogously to the neutron energy distribution density as

$$\upsilon^*(E) = \frac{\phi_0^*(E)}{\int_0^\infty \phi_0^*(E)dE}. \tag{6.27}$$

The setup of the benchmark is as follows: the domain is a graphite cube with 40 cm edge length. A group structure with 16 energy groups is used. The generation of this group structure structure is described in [90]. A point source is placed in the center of the cube for the lowest energy group (group 16). There is no external source for the other groups. The results are depicted in Fig. 6.2. As can be seen, the SPARC results agree very well with the ARTEMIS results. The relative deviation of SPARC to ARTEMIS is smaller than 0.05 % for all energies. Considering that ARTEMIS and SPARC use different numerical methods, this is an excellent result.

7. Numerical Solution of a Parameter Estimation Problem arising in PGNAA

In large sample PGNAA, the dependence of the neutron flux on the elemental composition can not be neglected. Thus PGNAA evaluations are solved using an iterative formulation. In [46, 2] this approach has been studied for the evaluation of one single PGNAA measurement using a HPGe detector. Based on a relative formulation the authors were able to show the existence of a solution under mild assumptions. Nevertheless, the uniqueness of a solution cannot be guaranteed in this setting, although, in accompanying experimental studies ambiguous solutions did not occur. Furthermore, the problem formulation of [46, 2], where relative compositions are determined, cannot be applied to the more complex measurement task of QUANTOM. The model of evaluation in this work therefore aims at the determination of absolute quantities. Further on, the setup we investigate in this chapter differs in several aspects from the one considered in [46, 2]. The sample in QUANTOM (see Sec. 1.3) is a 200l drum with a weight of several hundred kilograms. Therefore, it is much larger than the samples considered in standard literature before, which typically have a weight of a few grams. Hence, the spatial distribution of neutron flux and photopeak efficiencies need to be carefully represented in the simulation. Measurements for the neutron flux also present additional information regarding the reconstruction of elemental compositions. This additional information cannot be included in the standard iterative evaluation procedure [46, 2], so an alternative approach is needed.

Hence, we present an alternative to the standard iterative approach which is motivated by the spatially resolved measurement and additional measurement equipment for the neutron flux. We consider a mathematical formulation of the evaluation of QUANTOM measurements as a parameter estimation problem. Based on this formulation, we make use of gradient based methods to solve for the elemental composition. We only consider the parameter estimation problem with respect to neutron transport model, the dependence of photon transport to element composition is much weaker and can therefore be neglected.

This problem is studied in a simplified setting which makes it possible to also address the question of uniqueness of a solution. The material is assumed to be homogeneous, so only one measurement position of the HPGe detector is considered. Both types of detectors, for gamma spectra and total neutron flux, are considered in this setting. Elemental compositions are represented by the vector ρ of all corresponding elemental densities ρ_l, which are parameters of the neutron transport equation and therefore a natural candidate for the reconstruction. In this setting we investigate the uniqueness of solutions to the parameter estimation problem. Throughout this work we restrict the numerical setting to two space dimensions. This is sufficient to study and present all major effects arising in the problem setup. The problem setup is in parts comparable to Diffuse Optical Tomography (DOT), as a similar governing equation is used as an approximate model for particle transport. The connection between the two problems are discussed in the conclusive part of this chapter.

This chapter is structured as follows. In Section 7.1 we recall the SP_1-approximation of the neutron transport equation as well as the photopeak efficiencies and introduce model functions for the detector response which, combined, represent the forward model. Section 7.2 comprises a formulation of the parameter estimation problem for QUANTOM in the form of a PDE-constrained optimization problem. The key results of this work are contained in Section 7.3, where we investigate uniqueness of solutions to the parameter estimation problem. Section 7.4 comprises a solution method of the PDE-constrained optimization problem via a formal Lagrange method. In the following Section 7.5, we describe the computational implementation of the parameter estimation problem, based on an already existing neutron transport solver. We conclude the chapter by presenting numerical results which support our findings are presented in Section 7.6.

7.1. Forward Model

As mentioned before, two different kinds of detectors are used in QUANTOM to detect the total neutron flux and the energy resolved emitted gammas, respectively. For both detector types a model for the counting rates is required. Figure 7.1 shows the abstract measurement setup. Neutrons are constantly emitted by a neutron source Q. As a result, a steady state neutron flux field is formed in the graphite and the sample, while the graphite moderator provides a thermal neutron spectrum both in the graphite and the sample. It is important to note here, that the neutron flux field depends on the elemental composition of the sample, since the neutrons directly interact with its atoms. Both the sample and graphite are activated by the free neutrons, resulting in element-specific gamma emissions. The detectors are placed in proximity to the sample in order to collect the related measurement signal. Activation of the graphite can be neglected since scattering is dominant in the moderator. Furthermore, detectors can be collimated to the sample to further decrease this effect.

In order to evaluate the counting rates of the detectors, it is necessary to introduce models for the gamma and photon transport first. The transport of gammas and neutrons can be described by the linear Boltzmann equation [19]. For both types of transport we make use of approximate models. For the neutron transport equation, the SP_1-equation, also known as neutron diffusion equation, is used as an approximation. The applicability of the SP_1-equation for this setup was shown in a previous publication [48]. Note that in this work we apply the one-group SP_1-equation, since we can assume a thermal neutron flux. We denote the computational domain with D and its boundary with ∂D. The equation then reads

$$-\nabla \cdot \left(\frac{1}{3\Sigma_t(x)} \nabla \phi(x) \right) + \Sigma_a(x)\phi(x) = Q(x) \quad \forall x \in D \tag{7.1}$$

with the Marshak boundary conditions

$$\frac{1}{2}\phi(x) + \frac{1}{3\Sigma_t(x)} \frac{\partial}{\partial n}\phi(x) = 0 \quad \forall x \in \partial D \tag{7.2}$$

where $\phi(x)$ is the neutron flux and $Q(x)$ the source term while $\Sigma_t(x)$ and $\Sigma_a(x)$ are the total and absorption cross sections of the neutrons. It is important to note that the total and absorption cross sections within the sample can be expressed using the elemental densities ρ_l as well as the

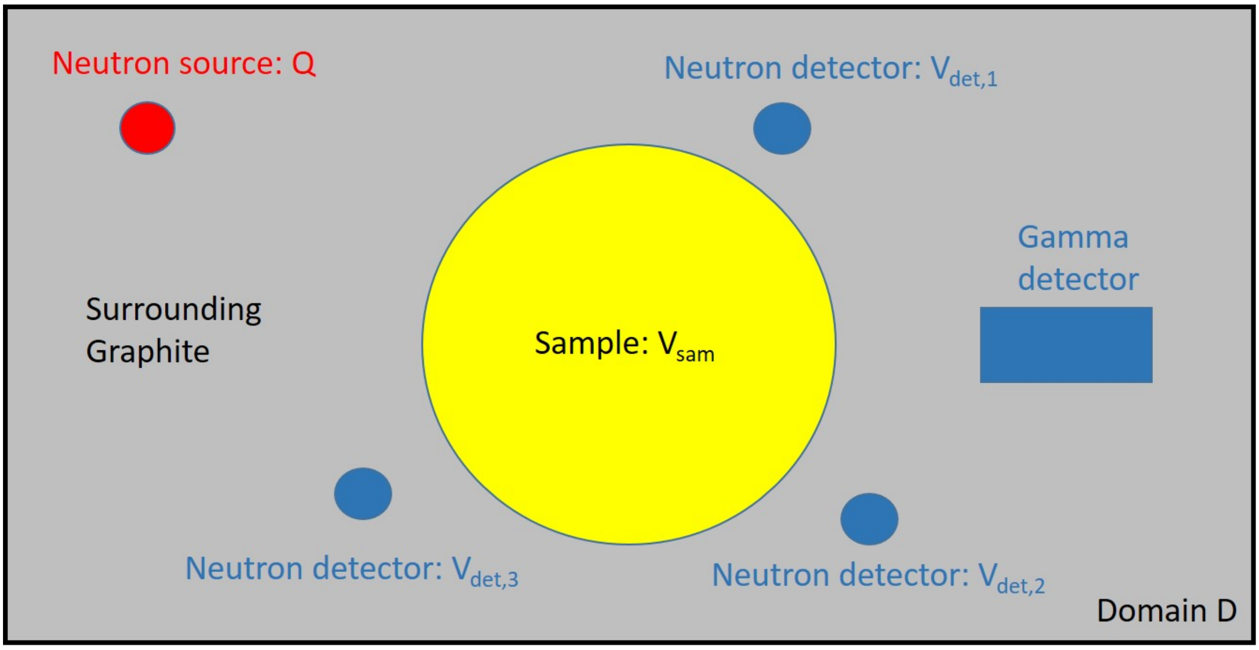


Figure 7.1.: Abstract sketch of the measurement setup, including the neutron source, sample and both types of detectors in a surrounding graphite moderator. In this chapter, the computational domain will be denoted as D, the sample volume as V_{sam} and the volumes of the neutron detectors as $V_{det,n}$ with $n = 1, 2, 3$.

total and absorption cross sections of the elements, σ_{tl} and σ_{al}, while these cross sections are independent of ρ anywhere else, leading to the representation:

$$\Sigma_t(x) = \begin{cases} \sum_l \sigma_{tl}\rho_l & x \in V_{sam} \\ M_t(x) & \text{else} \end{cases}$$
$$\Sigma_a(x) = \begin{cases} \sum_l \sigma_{al}\rho_l & x \in V_{sam} \\ M_a(x) & \text{else} \end{cases} \tag{7.3}$$

We denote the sample volume with $V_{sam} \subset D$. Since the sample is homogeneous, Σ_t and Σ_a are constant within the sample. $M_t(x)$ and $M_a(x)$ are piecewise linear functions that represent the cross sections in the actual facility.
According to Eq. (1.5), photon transport is relevant for the evaluation of a single peak as it determines the so-called photo-peak efficiency $\epsilon_{E_{\gamma,l}}$. This quantity represents the fraction of gammas which are detected at full energy. Consequently, scattered photons do not contribute to $\epsilon_{E_{\gamma,l}}$, allowing relatively simple models to calculate photo-peak efficiencies, rather than employing full photon transport simulations. The main effects which have to be considered are absorption, the geometry of the measurement setup and detector physics. Photopeak efficiency is a parameter dependent on gamma energy. Therefore, evaluating a measurement requires either point-wise values for all gamma energies considered or determining a (gamma-)energy-dependent calibration curve for a given measurement facility or setup, accounting for all these effects. This calibration curve provides photopeak efficiencies for any given gamma-energy and can be obtained experimentally, by measuring at given energies using calibration emitters and then interpolating between measured values. Alternatively, calibration curves can be determined using either analytical models or Monte Carlo simulations. As prompt gamma energies range between 50 keV up to 12 MeV, experimental calibration is of limited use, making modeling or simulation-based calibrations preferable. Optimal performance is typically achieved

using Monte Carlo simulations with MCNP [83] to account for for the intricate physics within a HPGe detector. Further details on the calibration of gamma detectors can be found in [36]. In principle, the photopeak efficiency also depends on the composition of the sample, which changes in every iteration of the evaluation (Eq. (1.7)) (as the current result of the measurement, not the physical composition!). However, the tabulated values in [89] suggest that the mass attenuation coefficients for most of the materials relevant for the application are similar. Therefore, the sample's impact on the photo-peak efficiency is primarily governed by its average density, a known parameter that remains constant during the evaluation. Thus, we assume photo-peak efficiencies to be constant throughout the iterative evaluation.

We can now define the counting rates of detectors. The neutron detector counting rates are given by

$$R_n[\phi] = \int_{V_{det,n}} \sigma_d \cdot \phi(x) dx, \tag{7.4}$$

where σ_d is the so-called detector cross section. The integral is carried out over the volume of the n-th detector $V_{det,n} \subset D$. See Fig. 7.1 for a sketch of the setup where the detectors are depicted in blue and Fig. 7.3 for the Finite Element mesh, where the neutron detector volumes are depicted in red.

In the following, we use the model for the photopeak efficiency that was introduced in Sec. 1.1. The photopeak efficiency then reads

$$\epsilon_{E_{\gamma,l}}(x) = \frac{1}{\pi} \arctan\left(\frac{r_{det}(x)}{d_{det}(x)}\right) e^{-\mu \cdot d(x)}, \tag{7.5}$$

where r_{det} and d_{det} are given functions. In contrast to the neutron detectors, the gamma detectors do resolve the energy of the measured photons. The resulting gamma spectrum can then be analyzed for element-specific photopeaks. Assuming a homogeneous sample, the counting rate for photons forming a photopeak at energy E_γ specific for element l is given by

$$P_{E_{\gamma,l}}[\rho, \phi] = \int_{V_{sam}} \epsilon_{E_{\gamma,l}} \rho_l \sigma_{E_{\gamma,l}} \phi(x) dx, \tag{7.6}$$

where $\sigma_{E_{\gamma,l}}$ is the so-called partial (n,γ)-cross-section and ρ_l is the density of the element l within the sample. Eq. (7.6) is a generalization of Eq. (1.5), accounting for the additional spatial variation of the parameters $\epsilon_{E_{\gamma,l}}$ and ϕ. The integral in Eq. (7.6) is carried out over the volume where the gamma photons are produced, the sample volume $V_{sam} \subset D$. A sketch of the setup is depicted with the sample volume in yellow is depicted in Fig. 7.1. Finite Element mesh, where the sample volume is also depicted in yellow, is shown in Fig. 7.3.

7.2. Formulation as PDE-constrained optimization problem

The parameter estimation problem to be considered in this work is the following: given measurements from the neutron and gamma detectors, determine the elemental densities ρ_l. These are themselves parameters of the neutron transport equation. The parameter estimation problem can be formulated as a PDE-constrained optimization problem with SP$_1$-equation as a constraint. Therefore, it can be written as an optimal control problem [43] with the vector of the material densities $\rho = (\rho_1, ..., \rho_L)$ as the control variable and the neutron flux ϕ as the state

variable.
The forward model provides a map from parameter to data space. Given the measurements for the peak counting rates $\overline{P_{E_{\gamma,l}}}$ and the neutron counting rates $\overline{R_n}$, an objective function including peak counting rates for all elements and the counting rates for all neutron detectors can be defined :

$$J_{det}[\rho,\phi]=\sum_{l=1}^{L}\frac{w_\gamma}{\overline{P_{E_{\gamma,l}}}^2}(P_{E_{\gamma,l}}[\rho,\phi]-\overline{P_{E_{\gamma,l}}})^2+\sum_{n=1}^{N}\frac{w_d}{\overline{R_n}^2}(R_n[\phi]-\overline{R_n})^2. \tag{7.7}$$

Here, w_γ and w_d are weights that determine the contribution of the different detectors to the objective function, while L is the number of elements within the sample and N the number of neutron detectors. In the following, we assume that we want to reconstruct the densities of two elements and have three neutron detectors available, so we have $L = 2$ and $N = 3$. The weight of the neutron detectors is set to $w_d = \frac{1}{2}$, while the weight of the gamma detectors is set to $w_\gamma = \frac{3}{4}$ to ensure an equal weight of both terms despite the higher number of neutron detectors. In addition, we include a term punishing deviations from the known total mass $\overline{M}$:

$$J_m[\rho]=\frac{w_m}{\overline{M}^2}(c_1\rho_1+c_2\rho_2-\overline{M})^2 \tag{7.8}$$

Here c_1 and c_2 are material-specific constants and $w_m = \frac{1}{2}$. Since the total mass $\overline{M}$ can be measured with high accuracy, we assume in the following that the error on the total mass is negligible. This term effectively acts as a regularization. Its effect is investigated in Sec. 7.6.2. The complete objective function is then given by

$$J[\rho,\phi]=J_{det}[\rho,\phi]+J_m[\rho]. \tag{7.9}$$

The parameter estimation problem can then be formulated as minimizing the objective function with the SP$_1$-equation as a constraint, so we want to solve the problem

$$\begin{aligned}\min_\rho J[\rho,\phi] \text{ subject to } & T(x;\rho)\phi(x)=Q(x) \text{ in } D\\ \text{and } & \frac{1}{2}\phi(x)+\frac{1}{3\Sigma_t(x;\rho)}\frac{\partial}{\partial n}\phi(x)=0 \text{ on } \partial D.\end{aligned} \tag{7.10}$$

We have two elements, so $\rho\in\mathbb{R}^2_+$. The differential operator in Eq. (7.10) is defined by

$$T(x;\rho)\phi(x)=-\nabla\cdot\left(\frac{1}{3\Sigma_t(x;\rho)}\nabla\phi(x)\right)+\Sigma_a(x;\rho)\phi(x), \tag{7.11}$$

while D is the domain with corresponding boundary ∂D. Outside the sample, Σ_t and Σ_a are independent of ρ. Thus, the differential operator T depends on ρ on a part of the domain.

7.3. Uniqueness

The objective function defined in Eq. (7.7) is convex. However, two fluxes ϕ that are identical within V_{sam} and $V_{det,n}$ $(n = 1, 2, 3)$ but deviate in the regions between would lead to the same value of J, so the objective function is not strictly convex. Therefore, standard theorems on uniqueness cannot be applied, necessitating the following investigation of uniqueness.

In this chapter, we investigate the uniqueness of the solution of the unperturbed parameter estimation problem, i.e. without noise in the measurement data. First, we introduce a change of variables to simplify the discussion. Thereafter, we investigate the objective function.

7.3.1. Change of variables

Within this section, we introduce a change of variables [5] to simplify the following discussion:

$$\psi = \gamma\phi \quad \text{with} \quad \gamma = \frac{1}{\sqrt{3\Sigma_t}}, \tag{7.12}$$

with $\Sigma_t \in C^2(D)$. By substituting Eq. (7.12) into Eq. (7.1) we obtain the Helmholtz-like equation:

$$\nabla^2\psi + \eta\psi = \frac{Q}{\gamma} \quad \text{with} \quad \eta = \frac{\nabla^2\gamma}{\gamma} + \frac{\Sigma_a}{\gamma^2}. \tag{7.13}$$

By substituting Eq. (7.12) into Eq. (7.2) we then obtain the boundary conditions

$$\frac{1}{2\gamma}\psi(x) + \gamma^2\frac{\partial}{\partial n}\left(\frac{1}{\gamma}\psi(x)\right) = 0 \quad \forall x \in \partial D \tag{7.14}$$

with the outer normal vector n. Since the cross sections as defined in Eq. (7.3) are dependent on ρ only within the sample volume V_{sam}, both the source term and the boundary conditions are independent of ρ. The parameter η in Eq. (7.13) is only varied within V_{sam} and effectively reduces there to

$$\eta = \frac{\Sigma_a}{\gamma^2} = 3\Sigma_a\Sigma_t \tag{7.15}$$

due to the homogeneity of the sample. For the investigation of uniqueness, the term punishing deviations from the total mass M is not needed, so we set w_m in Eq. (7.8) to zero. Using the definition of ψ in Eq. (7.12), we then obtain a transformed cost functional

$$J[\rho,\psi] = \sum_{l=1}^{L}\frac{w_\gamma}{\overline{P_{E_{\gamma,l}}}^2}(P_{E_{\gamma,l}}[\rho,\psi] - \overline{P_{E_{\gamma,l}}})^2 + \sum_{n=1}^{N}\frac{w_d}{\overline{R_n}^2}(R_n[\psi] - \overline{R_n})^2. \tag{7.16}$$

with neutron counting rates given by

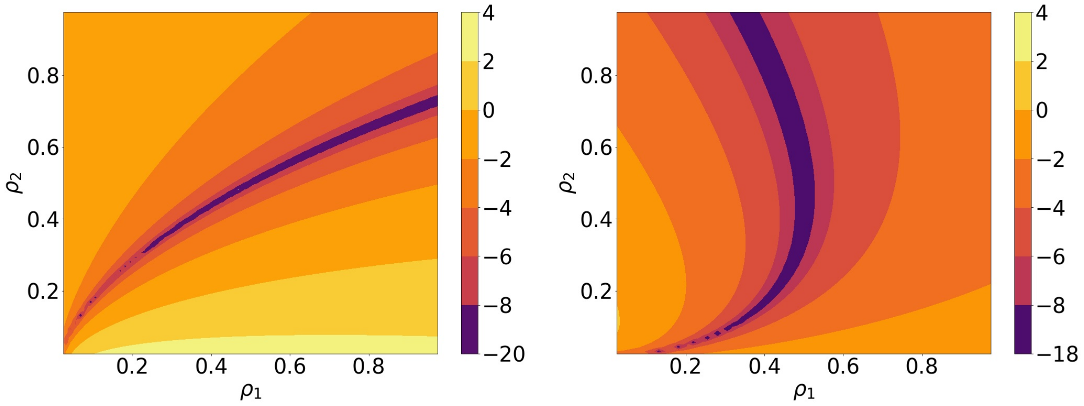
$$R_n[\psi] = \int_{V_{det,n}} \sigma_d\frac{1}{\gamma}\psi(x)dx, \tag{7.17}$$

and gamma counting rates given by

$$P_{E_{\gamma,l}}[\rho,\psi] = \int_{V_{sam}} \epsilon_{E_{\gamma,l}}(x)\rho_l\sigma_{E_{\gamma,l}}\frac{1}{\gamma}\psi(x)dx. \tag{7.18}$$

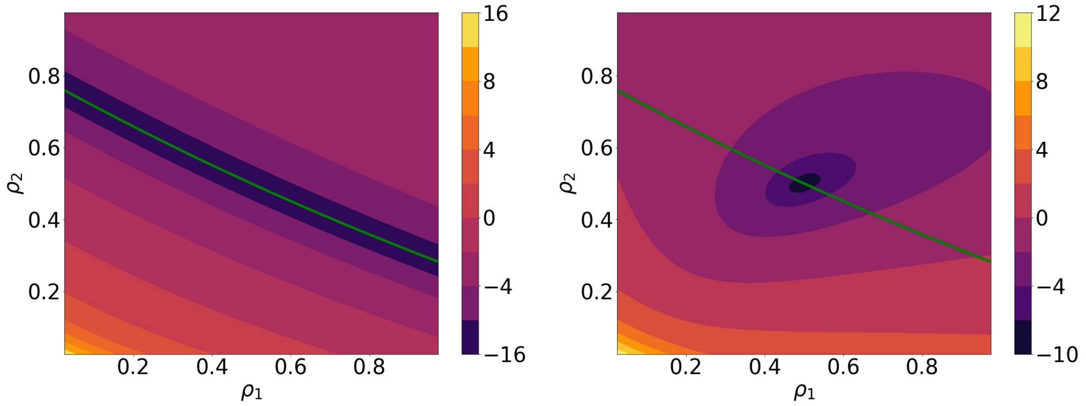
7.3.2. Objective Function

Figure 7.2 shows an exemplary objective function where the cross sections of the elements are set to $\sigma_{t1} = \sigma_{t2} = 1$, $\sigma_{a1} = 0.1$ and $\sigma_{a2} = 0.5$. We have the peak counting rates $\overline{P_{E_{\gamma,1}}} = 494589.3$ and $\overline{P_{E_{\gamma,2}}} = 945893.0$. The neutron counting rates read $\overline{R_1} = 182705.9$, $\overline{R_2} = 3958782.8$ and $\overline{R_3} = 7150.9$. The different terms in the objective function defined in Eq. (7.16) are displayed separately. Figure 7.2 (a) and (b) show the objective function including only the first and the second photopeak, respectively. It can be seen that there exists no unique minimum in both



(a) Objective function including the first Photopeak

(b) Objective function including the second Photopeak



(c) Objective function including all neutron detectors

(d) Full objective Function

Figure 7.2.: Subfigures (a)-(c) depict the different terms of the objective function after decomposition, subfigure (d) depicts the full objective function. The analytically computed minimum of the objective function of the neutron detectors is green. A logarithmic scale is used for all plots.

figures, but instead a large 'valley' and thus many admissible values for ρ.

Since the neutron counting rates, as can be seen from Eqs. 7.13 and 7.17, only depend on the parameter η, they can only be used to reconstruct this single parameter, independent of the number of detectors. Thus additional detectors can possibly reduce the measurement error, but cannot reconstruct additional parameters. Therefore, in Fig. 7.2(c) the sum over the partial objective functions for all detectors is shown. It should be noted that all combinations of Σ_t and Σ_a that lead to the same parameter η give rise to the same solution ψ, since both the boundary conditions and the source term are independent of ρ. Thus the neutron detectors alone are not sufficient to uniquely reconstruct the material composition of the sample, although they might reconstruct the parameter η. In fact, for this specific problem one can derive an analytic expression that allows to compute the density of one material for a given density of the other material that leads to a desired parameter η. By inserting Eqs. (7.3) into Eq. (7.15) for the case $x \in V_{sam}$, we obtain

$$\eta = 3(\sigma_{a1}\rho_1 + \sigma_{a2}\rho_2)(\sigma_{t1}\rho_1 + \sigma_{t2}\rho_2) \tag{7.19}$$

By solving this quadratic equation for ρ_2 and using that $\rho_2 > 0$, we obtain a relation between ρ_2 and ρ_1:

$$\rho_2 = \frac{-(\sigma_{a1}\sigma_{t2} + \sigma_{a2}\sigma_{t1})\rho_1 + \sqrt{(\sigma_{a1}\sigma_{t2} + \sigma_{a2}\sigma_{t1})^2\rho_1^2 - 4(\sigma_{a1}\sigma_{t1}\rho_1^2 - \eta/3)\sigma_{a2}\sigma_{t2}}}{2\sigma_{a2}\sigma_{t2}}. \tag{7.20}$$

If the value of η belonging to the minimum is used, Eq. (7.20) can be used to compute the expected position of the minimal values. It is depicted in Fig. 7.2(c) as the green curve. It can be seen that the green curve lies within the valley of the numerical minimal values.
From Figure 7.2(a)-(c) it is apparent that neither the measurements of the different gamma lines nor the neutron flux measurements alone have a unique minimum. In combination however, we obtain an objective function with a minimum contained within a small region, well suitable for a reconstruction, as depicted in Fig. 7.2(d).

7.4. Solution via formal Lagrange Method

In order to solve the problem posed in Eq. (7.10), we use a formal Lagrange method, from which we derive an expression for the gradient of the objective function, which in turn can be used for a gradient based optimization scheme. In the following, D is the computational domain with the corresponding boundary ∂D.
The Lagrangian for this problem is given by the objective function with a Lagrange multiplier $\phi^\star$ and the SP$_1$-equation as a constraint:

$$\mathcal{L}[\rho, \phi, \phi^\star] = J[\rho, \phi] - \int_D \phi^\star(x)(T(x;\rho)\phi(x) - Q(x))dx. \tag{7.21}$$

By formal integration by parts and using Eq. (7.2), we obtain

$$\begin{aligned}\mathcal{L}[\rho, \phi, \phi^\star] = J[\rho, \phi] &- \int_D \phi(x)(T(x;\rho)\phi^\star(x))dx - \int_D Q(x)\phi^\star(x)dx \\ &+ \int_{\partial D} \frac{1}{2}\phi(x)\phi^\star(x) + \frac{1}{3\Sigma_t(x)}\phi(x)\frac{\partial}{\partial n}\phi^\star(x)ds.\end{aligned} \tag{7.22}$$

Since the operator $T(x;\rho)$ is self-adjoint (see Subsec. 6.1.1), i.e. $T(x;\rho) = T^\star(x;\rho)$, we can identify $\phi^\star$ with the adjoint flux. The boundary terms in Eq. (7.22) need to vanish. This is only the case if the relation

$$\frac{1}{2}\phi^\star(x) + \frac{1}{3\Sigma_t(x)}\frac{\partial}{\partial n}\phi^\star(x) = 0 \tag{7.23}$$

is fulfilled on the boundary. The derivatives of the Lagrangian can be used within an iterative scheme to reconstruct the elemental densities ρ. Differentiation of Eq. (7.21) with respect to the adjoint flux $\phi^\star$ yields the weak form of the forward SP$_1$-equation, while the differentiation with respect to the flux ϕ yields the weak form of the adjoint SP$_1$-equation

$$T^\star(x;\rho)\phi^\star(x) = Q^\star(x) \tag{7.24}$$

with the adjoint source

$$
\begin{aligned}
Q^\star &= \frac{\delta J}{\delta \phi}[\rho, \phi] \\
&= \sum_{l=1}^{L} \frac{2w_\gamma}{\overline{P_{E_{\gamma,l}}}^2}\left(P_{E_{\gamma,l}}[\rho,\phi] - \overline{P_{E_{\gamma,l}}}\right) \int_{V_{sam}} \epsilon_{E_{\gamma,l}}(x)\rho_l \sigma_{E_{\gamma,l}} dx \\
&\quad + \sum_{n=1}^{N} \frac{2w_d}{\overline{R_n}^2}\left(R_n[\phi] - \overline{R_n}\right) \int_{V_{det,n}} \sigma_d dx.
\end{aligned} \tag{7.25}
$$

and the boundary condition defined in Eq. (7.23). The derivative of Eq. (7.21) with respect to the density of the l-th element is given by

$$
\begin{aligned}
\frac{\delta \mathcal{L}}{\delta \rho_l}[\rho, \phi, \phi^\star] = &\frac{\sigma_{t,l}}{\sum_j \sigma_{t,j}\rho_j} \int_{V_{sam}} \nabla\phi \cdot \nabla\phi^\star dx - \sigma_{a,l} \int_{V_{sam}} \phi\phi^\star dx \\
&+ \frac{2w_\gamma}{\overline{P_{E_{\gamma,l}}}^2}\left(\int_{V_{sam}} \epsilon_{E_{\gamma,l}}\sigma_{E_{\gamma,l}}\rho_l\phi dx - \overline{P_{E_{\gamma,l}}}\right) \int_{V_{sam}} \epsilon_{E_{\gamma,l}}\sigma_{E_{\gamma,l}}\phi dx \\
&+ \frac{2w_m c_l}{\overline{M}^2}(c_1\rho_1 + c_2\rho_2 - \overline{M}).
\end{aligned} \tag{7.26}
$$

Using the results for ϕ and $\phi^\star$, the gradient of the Lagrangian with respect to the different elemental density ρ_l is computed and can be used in a gradient-based optimization scheme.

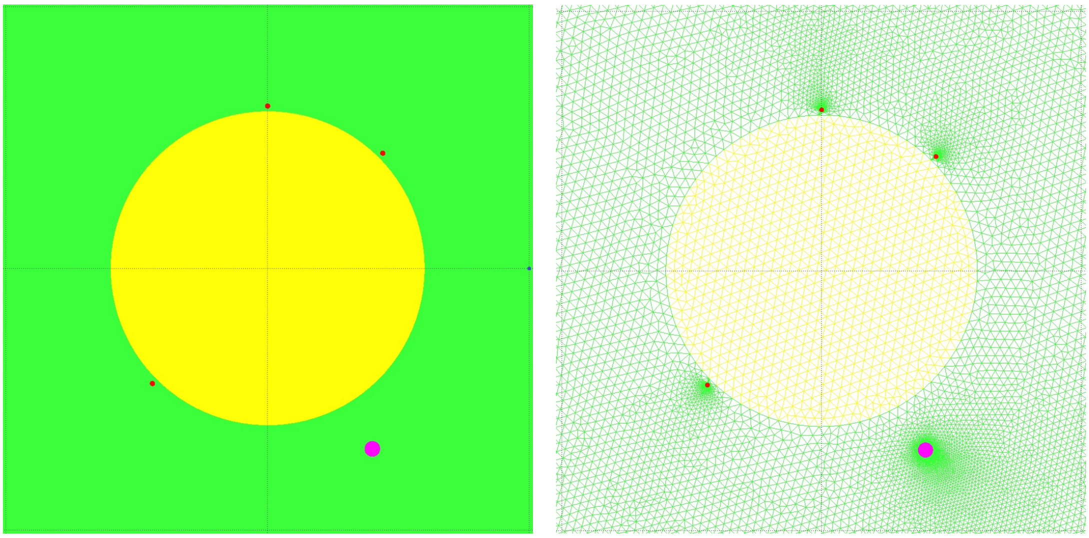
7.5. Implementation

For the neutron transport computations, the neutron transport code SPARC introduced in chapter 4 is used. As described in chapter 6, SPARC also enables adjoint computations, using the fact that the SP$_1$-equation is self-adjoint. However, a modification of the source term is necessary so that it equals the one introduced in Eq. (7.25). The gradient of the Lagrangian can than be computed as detailed out in Sec. 7.4.

The model used for the following simulation studies is shown in Fig. 7.3. It depicts a zoom into the relevant part of the computational domain. The remaining parts are filled solely with more moderating material. It depicts the moderator (green), the sample (yellow), the gamma detector (purple), the neutron detectors (red) and on the left side the source (in blue). Fig. 7.3(a) shows the computational domain, while Fig. 7.3(b) shows the Finite Element Mesh used for the reconstruction including approx. 84.000 Finite Elements.

7.6. Numerical Results

In this section, we present numerical results which show the applicability of the herein developed theory. In subsection 7.6.1, we first present a comparison of the gradient computed with the Optimal Control formalism as introduced in Section 7.4 and a Finite Difference approach in order to verify our implementation. Thereafter, we show exemplary minimizations for which the computed gradient is used within an L-BFGS-B scheme. In subsection 7.6.2 we study both



(a) Computational Domain

(b) Computational Mesh used for Reconstruction.

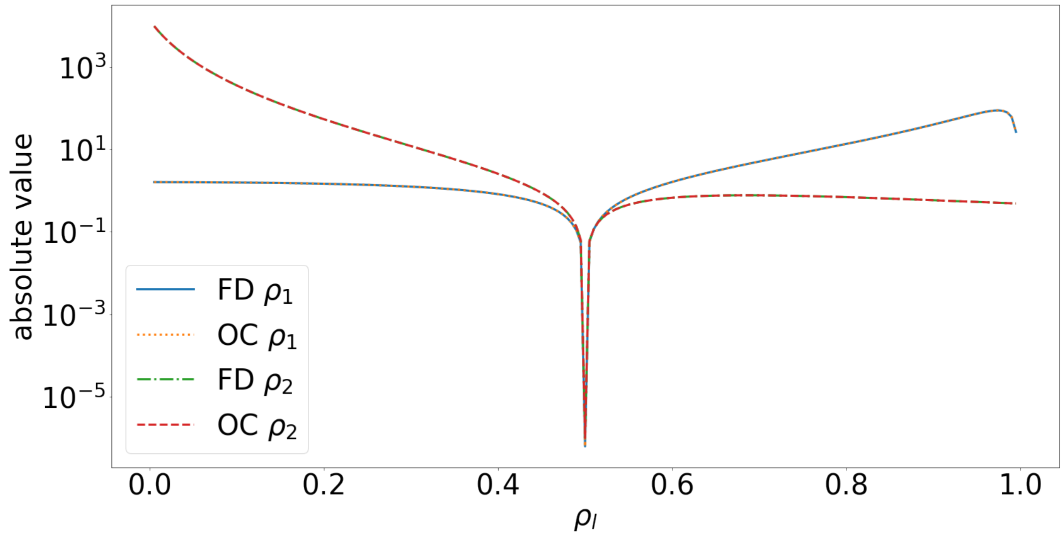
Figure 7.3.: Computational Model used in SPARC including the different regions of the computational domain D (left side) and the Finite Element mesh used for the reconstruction (right side). Depicted are the moderator (green), the sample volume V_{sam} (yellow), the gamma detector (purple), the neutron detector volumes $V_{det,n}$ $(n = 1, 2, 3)$ (red) and on the left side the source (in blue).

the stability of the reconstruction under given measurement errors as well as the applicability of our method for different materials. Throughout our simulation studies we have used fixed numerical grids, but the investigation of the effect of measurement noise of varying intensities in subsection 7.6.2 shows that we are not running into problems related to the so-called inverse crime in accordance with [100].

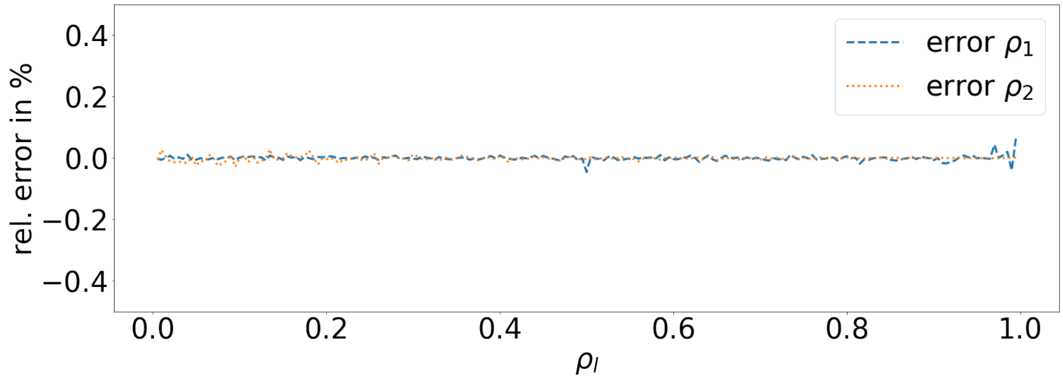
7.6.1. Verification of the Implementation

In order to verify the implementation, the gradient of the objective function is computed with both Finite Differences and the Optimal Control formalism. Fig. 7.4 depicts an exemplary case which compares the gradient components along the line segment defined by $\rho_2 = 1 - \rho_1,\ \rho_1 \in [0, 1]$. In the upper subfigure the absolute value of the gradient components is taken so that a logarithmic scale can be applied for better visualization. The plot shows a very good agreement between the results for both ρ_1 and ρ_2. In the lower subfigure, the difference between the gradients computed with both methods is depicted. The relative error is smaller than 0.1% for both ρ_1 and ρ_2.

While Fig. 7.4 depicts the verification of the gradient itself, Fig. 7.5 depicts the gradient-based minimization process. The gradients computed with the Optimal Control formalism are used to carry out the minimization using the L-BFGS-B [102] algorithm implemented in SciPy [98]. Fig. 7.5 depicts the objective function as well as three exemplary minimizations. The minimizations are represented by the blue, cyan and green lines, while the objective function is represented by a contour plot. It should be noted that each color in the contour plot represents an order of magnitude, so a major challenge here is to minimize a function which values vary



(a) Comparison of the gradient components computed with Finite Differences and the Optimal Control Formalism.



(b) Relative deviation of the gradient components computed with the Optimal Control formalism in comparison to the results obtained with Finite Differences.

Figure 7.4.: Comparison of the partial derivatives of the objective function computed with Finite Differences and the Optimal Control Formalism along the line segment defined by $\rho_2 = 1 - \rho_1, \;\; \rho_1 \in [0, 1]$.

over many orders of magnitude. For each case, the minimum (purple) is found within only a few iterations.

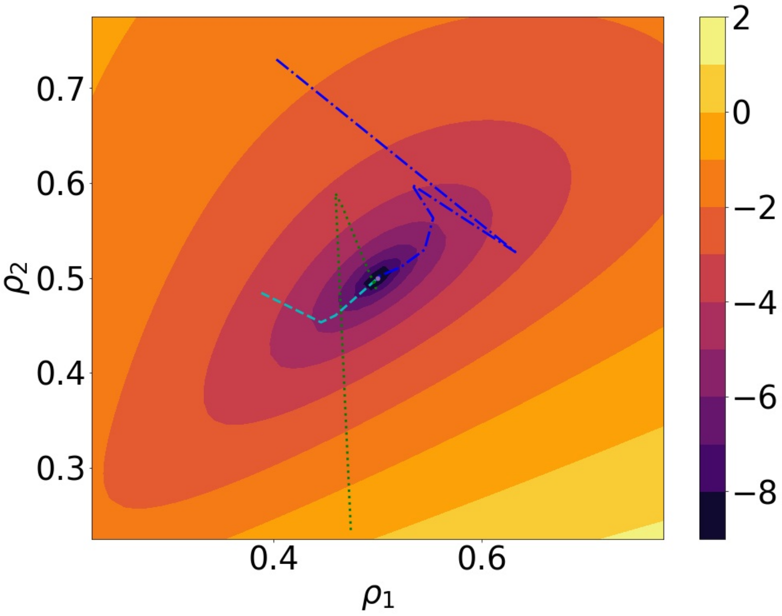


Figure 7.5.: The logarithmised objective function (see Eq. 7.9) with blue, cyan and green lines showing three exemplary minimizations using the L-BFGS-B algorithm.

7.6.2. Simulation Studies

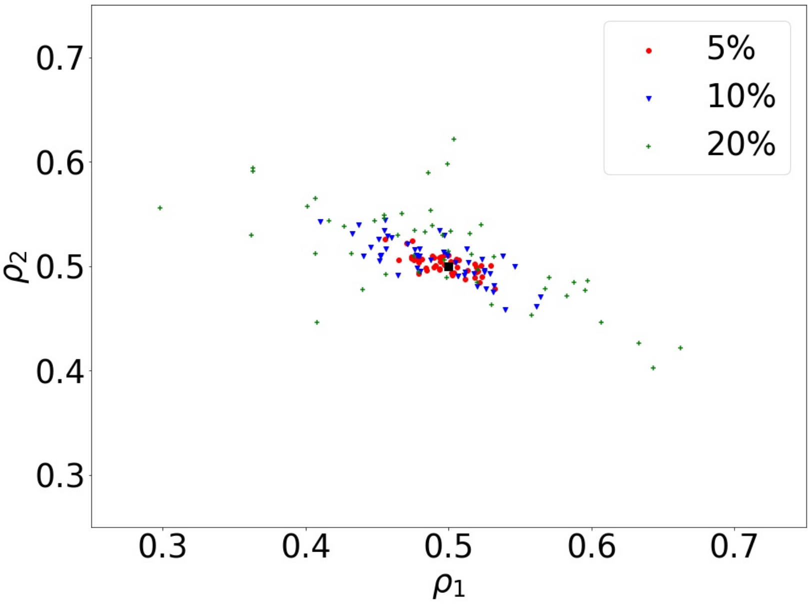


Figure 7.6.: Reconstructed values of ρ for normally distributed measurements errors of the counting rates $\overline{P_{E_{\gamma,l}}}$ and $\overline{R_n}$. 50 Simulations were carried out for each error level. The average errors on the reconstruction and the number of needed L-BFGS-B iterations can be found in tab. 7.1.

In order to investigate the stability of the reconstruction under measurement errors, a simulation study has been carried out with normally distributed errors. The standard deviations were as

Table 7.1.: Average absolute errors on the reconstruction (see Eq. 7.27) of ρ_1/ρ_2 for three different error levels as as well as the average number of needed L-BFGS-B iterations. For each case, 50 simulations have been carried out. The reconstructed values are depicted in Fig. 7.6. The used material properties are the same as defined in Tab. A.1 for the case Mixed Scattering/Absorption 1.

std. dev./total value measurement [%]	rec. error ρ_1 [%]	rec. error ρ_2 [%]	#iterations
5	3.1	1.4	6.3
10	6.1	3.4	6.3
20	11.7	8.0	6.6

Table 7.2.: Average absolute errors on the reconstruction of ρ_1/ρ_2 (see Eq. 7.27) for four different material combinations as well as the average number of needed L-BFGS-B iterations. The measurement values have a normally distributed error of 10%. For each material combination, 50 simulations have been carried out. The material properties can be found in Tab. A.1. The upper half of the table shows the reconstruction with the term $J_m[\rho]$ (see Eq. 7.8) punishing deviations from the total Mass $\overline{M}$ which acts as a regularization term. The lower half of the table shows the reconstruction without said term. For both cases, the material properties and the measurement values are identical.

reconstruction with regularization	rec. error ρ_1 [%]	rec. error ρ_2 [%]	#iterations
Mixed Scattering/Absorption 1	5.4	2.6	6.3
Mixed Scattering/Absorption 2	5.2	5.5	5.5
Scattering Dominated	6.1	5.1	5.8
Absorption Dominated	5.9	6.1	4.9
reconstruction without regularization			
Mixed Scattering/Absorption 1	50.0	51.8	11.0
Mixed Scattering/Absorption 2	23.7	23.3	5.8
Scattering Dominated	63.1	66.1	9.0
Absorption Dominated	46.5	48.9	9.9

large as 5%, 10%, and 20% of the simulated counting rates. For each error level 50 Simulations have been carried out. The reconstructed values are depicted in Fig. 7.6. The error-free reconstruction is depicted as the black point in the center, while the reconstructed values for ρ are depicted in red, blue, and green, depending on the error level. It can be seen that the reconstructed values for ρ are scattered in a wider area for larger error levels. Tab. 7.1 shows the reconstruction errors for the different error levels. The errors are to be understood as mean absolute errors, that is

$$\xi_{rel} = \sum_{i=0}^{50} \frac{|\rho_{rec} - \rho_{tr}|}{\rho_{tr}}, \tag{7.27}$$

where ρ_{rec} is the reconstructed value for ρ and ρ_{tr} is the true value of ρ. For all three cases the reconstruction errors are much smaller than the errors on the counting rates. At the same time the number of needed L-BFGS-B iterations is nearly independent from the error on the simulated counting rates.

In order to test the applicability of our method, four cases with different material properties have been investigated. The materials themselves are artificial, but using parameters that are realistic for the application. The first two cases are combinations of a highly scattering and a highly absorbing material in each case. The third and fourth cases investigate combinations of materials with similar neutronics properties, in one case two highly scattering materials and in the other case two highly absorbing materials. The material properties used for the test cases can be found in Tab. A.1, while the counting rates are given in Tab. A.2. The upper half of Tab. 7.2 shows the errors of the reconstruction for the four different cases. It can be seen that the reconstruction of ρ works well for all investigated cases. The error on the counting rates is 10%, while the error on the reconstructed densities is typically 5% to 6%. In one case the error on ρ_2 is as low as 2.6%. This seems to be an exception, however.

The regularizing effect of the term punishing deviations from the total mass introduced in Eq. (7.8) can be seen by comparing the upper and the lower half of Tab. 7.2. The lower half shows the the reconstruction results with the weight of the term w_m set to zero. The same initial and measurement values as before have been used. Without the regularization, the error of the reconstruction massively increases, from values between 2.6% and 6.1% to values between 23.3% and 66.1%. The average number of iterations also increased considerably, from approximately five to six to values between nine and eleven. The second test case is the sole exception. These results underline the importance of the regularization term in order to increase the accuracy of the proposed method.

8. Numerical Robustness of PINNs for Multiscale Transport Equations

Even though the performance of neutron transport simulations for large sample PGNAA can be significantly improved by the use of deterministic solvers (see chapter 5) instead of more costly Monte Carlo computations, neutron transport simulations remain a bottleneck of the entire evaluation workflow (see Fig. 1.5). Therefore, it is of great interest to investigate alternative approaches for neutron transport simulations. One such approach are so-called Physics Informed Neural Networks (PINNs).

PINNs have recently attracted a great deal of attention. First proposed in [79] in 2019, PINNs provide a method to solve partial differential equations (PDEs) using (deep) neural networks (NN) by incorporating the corresponding equation in a loss function in a training process. Successively, points in the computational domain of the PDE are chosen and the weights of the NN are adapted to solve the PDE pointwise, resulting in an approximation of the solution on the complete computational domain. In that way, the numerical solution of the PDE corresponds to training of a NN. First computational packages are available which therefore provide easy to use black-box solver for PDEs, what makes the method attractive for users.

Although already heavily used in various applications [52, 16, 17, 24, 87], only a limited number of results regarding the numerical analysis of PINNs exist so far [25]. For the classes of linear second order elliptic and parabolic PDEs convergence of the discretized solution to the PDE solution can be shown in a strong sense [91]. A more detailed numerical analysis of PINNs (convergence, error bounds, stability etc.) needs to be developed based on these first results in order to gain a deeper understanding of this undoubtedly powerful and useful method.

With this chapter, we contribute to the numerical analysis of PINNs by studying an analogy between PINNs and the Least Squares Finite Element method (LSFE). For a specific class of PDEs (neutron transport equations) we study similarities based on an existing theory for LSFE for numerical solution in the so-called diffusive regime (cf. [66], [67], [68]). The LSFE is a numerical approach used to solve partial differential equations and involves reformulating the problem into a minimization of a least-squares functional. A numerical solution is obtained by solving for a finite element representation of the corresponding variational formulation.

The analogy between PINNs and the LSFE lies in their shared approach of reformulating the solution of differential equations as the minimization of a (typically quadratic) functional. In LSFE, this involves the construction of a least-squares functional based on the residual of the differential equation, which, when minimized, yields the solution to the problem in a variational form. Similarly, PINNs define a loss function based on the residuals of the differential equation and corresponding boundary conditions. Minimizing this loss function drives the neural network to approximate the solution. Hence, although not explicitly stated in variational terms, the minimization of the loss function in PINNs can be viewed as finding an approximate solution that satisfies the differential equation in an integral sense across the domain. Using an activation function such as ReLU for PINNs, the corresponding neural network can be interpreted as a piecewise linear approximation of the solution comparable to the finite element representation.

For this reason, we expect both methods to show comparable behavior in situations where numerically resolving multiscale effects in a PDE becomes difficult. A prominent and widely studied example is neutron transport in the so-called diffusive regime, which is briefly introduced in the next section. This example is also of great relevance for large-sample PGNAA measurement facilities, since the neutron transport in moderator materials like graphite takes place in the diffusive regime.
This chapter is structured as follows. We start by introducing neutron transport in the diffusive regime in section 8.1. Sections 8.2 and 8.3 introduce PINNs and LSFE respectively as numerical methods for the solution of PDEs. Here, we also introduce the notations of the methods that we use in the following. The main result is stated in Section 8.4. We prove that numerical solutions of PINNs in the diffusive regime do not converge to the correct solution, and show that a diffusive scaling can be applied to overcome this, in full agreement with the theoretical results for LSFE. We conclude this chapter with numerical results that are presented in Section 8.5.

8.1. Neutron Transport in the Diffusive Regime

To simplify the following analysis, we will not investigate the full linear Boltzmann equation (Eq. 2.1), but the so-called *slab geometry* case [60] (one space and directional dimension), where all relevant effects in the diffusive regime can be shown. Here, the medium is assumed to be infinite in two dimensions (e.g., the y- and z-directions) and finite in the x-direction. The medium is therefore bounded by two parallel planes which form a slab through which the neutrons travel. In the following, we denote the x-coordinates of the planes that limit the slab by x_l for the left plane and x_r for the right plane. The (one-dimensional) spatial coordinate x is used to describe the position of a neutron within the slab, the direction of neutron movement becomes a one-dimensional angle $\mu = \cos(\theta)$, where θ is the angle between the neutron's direction of travel and the x-axis. Fig. 8.1 depicts the *slab geometry* case.

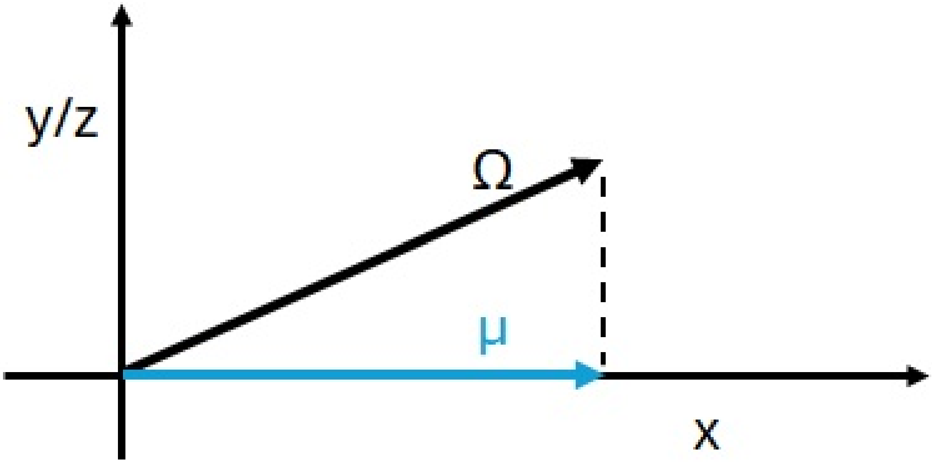


Figure 8.1.: In *slab geometry*, transport is projected onto the x-axis. The problem is symmetric in the y/z-plane.

In the following, we assume an isotropic source, i.e. Q independent of μ. We have $x \in D$ where $D = [x_l, x_r] \subset \mathbb{R}$. The steady state, single energy group neutron transport equation in slab geometry is then formulated as

$$L\Psi(x,\mu) = Q(x) \tag{8.1}$$

with source term Q and operator

$$L\Psi(x,\mu) = \left(\mu\frac{\partial}{\partial x} + \sigma_t(I-P) + \sigma_a P\right)\Psi(x,\mu), \tag{8.2}$$

where I denotes the identity operator and P the projection onto the space of all L^2-functions independent of angle

$$(P\Psi)(x) = \frac{1}{2}\int_{-1}^{1}\Psi(x,\mu)d\mu. \tag{8.3}$$

Boundary conditions can be introduced by a boundary operator B, so that

$$B\Psi(x,\mu) = g(x,\mu)\ \forall(x,\mu)\in\Gamma^-, \tag{8.4}$$

with $\Gamma^- = \{(x,\mu)\in\partial D\times[-1,1] : n(x)\cdot\mu < 0\}$, where ∂D is the boundary of the computational domain D and n is the outer normal vector. The boundary conditions in this form specify the flux distribution entering D over ∂D. We introduce the boundary conditions in detail later in Section 8.5.

The multiscale effects in this setting originate from two different scales relevant to this problem: the mean-free path is small compared to the size of the physical domain. The numerical instability of PINNs that we discuss in this chapter is a consequence of the different scales and the stiffness they induce. In the diffusive regime scattering dominates the effects of free transport or absorption and the mean-free path of particles is small. The diffusive regime can typically be characterized by a small parameter $\varepsilon > 0$, which can be interpreted as the ratio of the mean-free path to the physical size of the domain of computation. Introducing such $\varepsilon > 0$ we scale the cross sections and source term accordingly [57] as

$$\sigma_t \to \frac{1}{\varepsilon},\ \sigma_a \to \alpha\varepsilon,\ Q \to \varepsilon Q, \tag{8.5}$$

where α is assumed to be $\mathcal{O}(1)$. By applying this scaling on Eq. (8.1) we obtain

$$L_\varepsilon\Psi^\varepsilon(x,\mu) = \varepsilon Q(x) \tag{8.6}$$

where the neutron flux in the diffusion limit is denoted as Ψ^ε and the scaled operator L_ε reads

$$L_\varepsilon\Psi^\varepsilon(x,\mu) = \left(\mu\frac{\partial}{\partial x} + \frac{1}{\varepsilon}(I-P) + \varepsilon\alpha P\right)\Psi^\varepsilon(x,\mu). \tag{8.7}$$

It can be shown [56] that in the limit of $\varepsilon \to 0$ the solution of the transport equation Ψ^ε converges to a function ϕ_0, which is independent of μ and the solution of an associated diffusion equation

$$-\frac{\partial}{\partial x}\frac{1}{3\sigma_t}\frac{\partial\phi_0}{\partial x}(x) + \sigma_a\phi_0(x) = Q(x). \tag{8.8}$$

For numerical methods it is therefore of key importance to reproduce this behavior on a discretized level. A method that ensures this is called asymptotic preserving [47]. For the LSFE method a corresponding analysis has been carried out (cf. [66], [67], [68]) and we build upon these results for the investigation of PINNs. A related analysis using PINNs for multiscale time dependent linear transport equations has been carried out in [51]. Using LSFE without additional scaling (see Section 8.4), the numerical solution produces an incorrect limit as $\varepsilon \to 0$. We show in this chapter that the same happens when PINNs are applied to solve this class of equations and that the same scaling as derived for LSFE can be applied to ensure a correct limiting behavior for PINNs, underlining the close connection between the two numerical methods.

8.2. Physics Informed Neural Networks

Physics Informed Neural Networks, PINNs for short, are a numerical method for solving general partial differential equations [79, 52]. A neural network serves as an interpolation function for the solution of the PDE. It is obtained by successively selecting points in the computational domain and adjusting the parameters of the neural network to obtain the desired output, which is commonly referred to as training the NN. The method is therefore mesh-free and can serve as a black-box solver for a PDE, since only the form of the equation needs to be known. The mathematical formulation of the method is briefly presented below.
In the following, we use a notation similar to [91]. Let $\Psi^M(x,\mu) : \mathbb{R}^2 \to \mathbb{R}^1$ be a feed-forward neural network (FNN) with M layers. Then its output is recursively defined by

$$\begin{aligned} \Psi^0(x,\mu) &= (x,\mu) \in \mathbb{R}^2 \\ \Psi^m(x,\mu) &= W^m \sigma(\Psi^{m-1}(x,\mu)) + b^m \quad \in \mathbb{R}^{M_l},\ 2 \le m \le M-1 \\ \Psi^M(x,\mu) &= W^M \sigma(\Psi^{M-1}(x,\mu)) + b^M \quad \in \mathbb{R}^1 \end{aligned} \tag{8.9}$$

with the weight matrices $W^m \in \mathbb{R}^{N_m, N_{m-1}}$, the biases $b^m \in \mathbb{R}^{N_m}$ and the activation function σ. Let $\vec{n} = \{N_0, ..., N_M\} \in \mathbb{N}^M$ with $\theta = \{W^m, b^m\}_{1 \le m \le M}$ be the parameter space of the network. Since a neural network $\Psi^M(x,\mu)$ depends on the architecture $\vec{n}$ and its parameters θ, we can denote it as $\Psi^M = \Psi^M(x,\mu,\vec{n},\theta)$. For a fixed $\vec{n}$, we write $\Psi^M(x,\mu,\theta)$.
A numerical solution of a PDE using such a FNN can be obtained by incorporating the PDE residual into the loss function during the training process, thus obtaining a 'physics informed' neural network. In the following, we assume that a unique classical solution of the PDE exists. The PDE and its boundary conditions are given by

$$L[\Psi](x,\mu) = Q(x)\ \forall x \in D,\ B[\Psi](x,\mu) = g(x,\mu)\ \forall (x,\mu) \in \Gamma^-. \tag{8.10}$$

Here L denotes a differential operator, Q is a source term and B is a boundary operator. Hence, the PINN maps the spatial variable x to an approximate solution of the PDE via $\Psi^M(x,\mu) \approx \Psi(x,\mu)$. The loss function of the neural network is given by

$$\text{Loss}[\Psi] = \frac{1}{|T_D|} \sum_{(x,\mu) \in T_D} (L[\Psi](x,\mu) - Q(x))^2, \tag{8.11}$$

where T_D denotes a set of training points within the domain D. This set is often generated using a quasi-random low-discrepancy sequence like the Sobol sequence [92]. The training process of the PINN is the equivalent to solving the PDE in classical methods. It is a minimization of the loss function in the set of functions V which is defined by

$$V = \{\Psi^M(\cdot,\cdot,\theta) : \mathbb{R}^2 \to \mathbb{R}^1 | \theta = \{W^m, b^m\}_{1 \le m \le M}\}. \tag{8.12}$$

In general, V is not a vector space. The boundary conditions are incorporated by a restriction of the set of functions so that

$$\begin{aligned} V_b = \{\Psi^M(\cdot,\cdot,\theta) : \mathbb{R}^2 \to \mathbb{R}^1 | \theta = \{W^m, b^m\}_{1\leq m \leq M} \\ \text{where } B\Psi^M(x,\mu) = g(x,\mu)\ \forall (x,\mu) \in \Gamma^-\}. \end{aligned} \tag{8.13}$$

Alternatively, the boundary condition can be included in the loss function. We will describe this approach in Subsec. 8.5.2. For now, we can write the training process of the PINN as a minimization of the loss function in the set of functions V_b:

$$\Psi_{min} = \min_{\Psi \in V_b} \text{Loss}[\Psi]. \tag{8.14}$$

Equivalently, we can interpret it as a minimization of the loss function in the parameter space with gradient-based optimizers such as Adams [53] or L-BFGS [15]:

$$\Psi_{min} = \min_{\theta} \text{Loss}[\Psi](\theta). \tag{8.15}$$

It is important to note that $\Psi^M(x,\mu)$ using the rectified linear unit (ReLU) as activation function is piecewise linear. Moreover, it can be shown that Neural Networks with ReLU activation functions and sufficiently many layers can reproduce all piecewise linear basis functions from the finite element method [40].

The problem considered in this work depends, in principle, on the architecture of the underlying NN, including the number of layers and neurons. However, our analysis remains valid regardless of the specific architecture, provided that a feed-forward NN is used, which is sufficiently large to accurately approximate the solution to the PDE. The theoretical analysis focuses on networks with ReLU activation functions, while the case of hyperbolic tangent activation functions is also explored in the numerical experiments. In general, using different activation functions (or a combination of them) or Non-Feedforward Neural Networks might lead to different results in the diffusion limit.

8.3. Least-Squares Finite Elements

The Least-Squares Finite Element (LSFE) method relies on a finite element discretization of a variational formulation of the underlying PDE. The least-squares formulation of Eq. (8.1) (neutron transport in slab geometry) is given by

$$\min_{\Psi \in U_b} F[\Psi] \quad \text{with} \quad F[\Psi] = \int_{x_l}^{x_r} \int_{-1}^{1} (L\Psi(x,\mu) - Q(x,\mu))^2 d\mu dx \tag{8.16}$$

where the boundary condition (Eq. 8.4) is incorporated by a restriction of the function space:

$$U_b = \{u(x,\mu) \in L^2(D) : \mu\frac{\partial u}{\partial x} \in L^2(D) \text{ and } u|_{\Gamma^-} \in L^2(\Gamma^-) \text{ where } Bu(x,\mu) = g(x,\mu)\ \forall (x,\mu) \in \Gamma^-\} \tag{8.17}$$

For Ψ to be a minimizer of F, it is a necessary condition that the first variation vanishes for all admissible $u \in U_b$, resulting in the problem: find $\Psi \in U_b$ such that

$$a(\Psi, u) = l(u) \tag{8.18}$$

for all admissible $u \in U_b$ with the bilinear form

$$a(\Psi, u) := \int_{x_l}^{x_r} \int_{-1}^{1} L\Psi(x,\mu) Lu(x,\mu) d\mu dx \tag{8.19}$$

and the functional

$$l(u) = \int_{x_l}^{x_r} \int_{-1}^{1} Q(x,\mu) Lu(x,\mu) d\mu dx. \tag{8.20}$$

The system is then discretized by replacing the function space U with a finite dimensional subspace $U_h \subset U_b$. Consequently, we obtain a finite dimensional linear system for the discretized neutron flux Ψ_h:

$$a(\Psi_h, u_h) = l(u_h) \ \forall u_h \in U_h. \tag{8.21}$$

If $\{\lambda_p\}$ is a basis of U_h, we can represent the solution as an expansion in said basis:

$$\Psi_h = \sum_p c_p \lambda_p, \tag{8.22}$$

where c_p are the expansion coefficients. This allows us to write Eq. (8.21) as a linear system

$$A\Psi_\lambda = b \tag{8.23}$$

where the components of A and l are given by $A = a(\lambda_p, \lambda_q)_{pq}$ and $b = l(\lambda_p)_p$, while Ψ_λ is the vector of expansion coefficients c_p.
For the spatial discretization in the variable x first-order Lagrange polynomials are used as basis functions. The discretization in the angular variable μ is with the method of moments is described in Subsec. 8.5.1.
A general introduction to the least-squares finite element methods can be found in [9]. Details about least-squares finite element solutions of the neutron transport equation in the diffusive regime are discussed in [67].
A key difference between LSFE and PINNs is that, after parametrization, LSFE requires the solution of a linear problem (Eq. 8.23), whereas PINNs requires the solution a nonlinear optimization problem (Eq. 8.15). This means that for the PINN we look for weights and biases, and for LSFE we look for basis expansion coefficients.

8.4. PINNs in the asymptotic limit

In the diffusive limit $\varepsilon \to 0$ as introduced in Section 8.1 the neutron transport equation becomes singular and numerical methods often fail to produce valid approximations in this regime. Fig. 8.2 shows the behavior of a numerical solution obtained with a PINN with ReLU activation functions and a first-order LSFE solution in the asymptotic limit ($\varepsilon = 10^{-4}$), compared to an analytical reference solution. This testcase and its setting will be discussed in detail in Subsec. 8.5.3.

It can be seen that the neutron fluxes resulting from both numerical methods are close to zero and far away from the reference solution. The numerical analysis of the LSFE method yields an explanation to why this behavior occurs, and the methodological similarities between LSFE

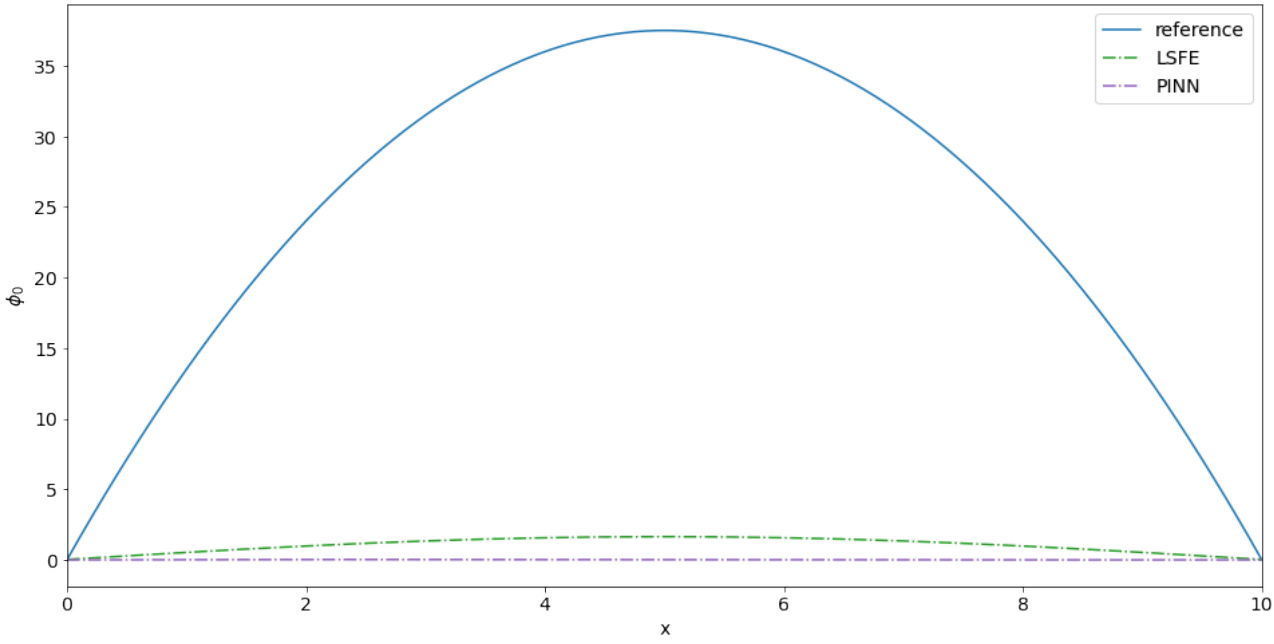


Figure 8.2.: Results for PINN with ReLU activation functions and first order LSFE in the asymptotic limit ($\varepsilon = 10^{-4}$) compared to an analytical reference solution. The setting is discussed in detail in Subsec. 8.5.3.

and PINNs suggested a corresponding behavior for PINNs.
In the following, we provide a numerical analysis that shows that a PINN with ReLU activation functions does in general not yield a correct solution in the diffusion limit.

We can express the solution of the transport equation by an expansion in the Legendre moments in angle. For $n \in \mathbb{N}$, let $P_n(\mu)$ be the normed n-th order Legendre polynomial on the interval $[-1, 1]$. In this section, we use the normed Legendre polynomials defined as $p_n(\mu) = \sqrt{2n+1}P_n(\mu)$ to simplify the coefficients. Let

$$\phi_n(x) = \int_{-1}^{1} p_n(\mu)\Psi(x,\mu)d\mu \tag{8.24}$$

be the n-th angular moment of Ψ. Then the expansion of the neutron flux in angle is given by

$$\Psi(x,\mu) = \sum_{i=0}^{\infty} p_n(\mu)\phi_n(x). \tag{8.25}$$

Furthermore, we note that in the diffusion limit, it is sufficient to take only the first two moments of the expansion into account [18].
In this situation we consider the set of functions

$$V_h = \{v_h \in V_b : v_h(x,\mu) = \phi_0(x) + \mu\phi_1(x), \text{where } \phi_0, \phi_1 \in \mathcal{P}(T_h)\}, \tag{8.26}$$

where $\mathcal{P}(T_h)$ denotes the space of piecewise linear functions on the partition T_h of the slab. Here the term partition corresponds to an interval within the slab, i.e. $[x_1, x_2] \subset [x_l, x_r]$. In the diffusive regime, the training of a PINN corresponds to the solution of the following minimization problem:

$$\Psi^{\varepsilon}_{min} = \min_{\Psi^{\varepsilon} \in V_h} \frac{1}{|T_D|} \sum_{i,j} (L_\varepsilon \Psi^{\varepsilon}(x_i,\mu_j) - \varepsilon Q(x_i,\mu_j))^2 \tag{8.27}$$

In the following, we use the fact that for all $\xi > 0$ it is possible to find a number of training points $|T_D(\xi)|$ so that the absolute difference between the discrete sum divided by the number of training points and the integral is smaller than the parameter ξ:

$$\left| \frac{1}{|T_D(\xi)|} \sum_{i,j} (L_\varepsilon \Psi^\varepsilon(x_i, \mu_j) - \varepsilon Q(x_i, \mu_j))^2 - \int_{x_l}^{x_r} \int_{-1}^{1} (L_\varepsilon \Psi^\varepsilon(x, \mu) - \varepsilon Q(x, \mu))^2 d\mu dx \right| < \xi. \quad (8.28)$$

This is ensured by the sequence we use to construct the training points. The training points are generated with a Sobol sequence which is constructed in a way that the sum divided by the number of (training) points converges to the integral [92]. Hence, if we replace in the following the term 'minimum' by 'ξ-suboptimal solution', e.g. the L^2-norms difference of the minimum and the ξ-suboptimal solution is smaller than ξ, the results remain valid. For the sake of readability, in the following, we neglect this sublety and we assume that a function Ψ^ε_{min} that minimizes Eq. (8.29) also minimizes Eq. (8.27) and use the term 'minimum'. Hence, instead of Eq. (8.27) we directly consider the following continuous minimization problem:

$$\min_{\Psi \in V_h} F[\Psi] \quad \text{with} \quad F[\Psi] = \int_{x_l}^{x_r} \int_{-1}^{1} (L_\varepsilon \Psi^\varepsilon(x, \mu) - \varepsilon Q(x, \mu))^2 d\mu dx. \quad (8.29)$$

To summarize, in the following analysis Ψ^ε_{min} denotes the neural network trained on a Sobol sequence T_D being a solution of Eq. (8.27). This neural network is an approximate solution of the neutron transport equation and is furthermore a piecewise linear, hence integrable function. By making use of the property of ξ-suboptimality we also consider Ψ^ε_{min} to minimize the continuous functional Eq. (8.29).

8.4.1. Characterization of the minimizer

In the following, we formally expand the solution Ψ in powers of ε and derive relations between the expansion coefficients. We then use these relations to prove that the PINN method with ReLU activation functions yields an incorrect solution in the diffusion limit. For our analysis we build on the theory developed in [80] for LSFE, which we adapt for PINNs.

Lemma 8.4.1. *Let the Functional F and the set of functions V_h be as given in Eqs. (8.29) and (8.26). Suppose Ψ^ε_{min} minimizes F restricted to V_h with $g = 0$. Suppose further that $\varepsilon \leq 1$ and that Ψ^ε_{min} has an expansion*

$$\Psi^\varepsilon_{min}(x, \mu) = \phi^\varepsilon_0(x) + \mu \phi^\varepsilon_1(x) \quad (8.30)$$

with

$$\phi^\varepsilon_0(x) = \sum_{\nu=0}^{\infty} \varepsilon^\nu \eta_\nu(x) \quad \text{and} \quad \phi^\varepsilon_1(x) = \sum_{\nu=0}^{\infty} \varepsilon^\nu \delta_\nu(x), \quad (8.31)$$

where η_ν and δ_ν are independent of ϵ. Then we have:

$$\delta_0(x) = 0 \quad \text{and} \quad \eta'_0(x) = -\delta_1(x). \quad (8.32)$$

Proof. By inserting Eq. (8.30) into Eq. (8.7) and using that P is a projection onto the zeroth moment ($P\Psi = \phi_0$), while $I - P$ projects onto moment one and all higher order moments, we obtain

$$L_\varepsilon \Psi^\varepsilon_{min} = \mu \frac{\partial \phi^\varepsilon_0}{\partial x} + \mu^2 \frac{\partial \phi^\varepsilon_1}{\partial x} + \frac{1}{\varepsilon}\mu\phi^\varepsilon_1 + \varepsilon\alpha\phi^\varepsilon_0 \tag{8.33}$$

Inserting Eq. (8.31) and sorting the terms in the order of ε yields

$$L_\varepsilon \Psi^\varepsilon_{min} = \frac{1}{\varepsilon}[\mu\delta_0] + [\mu\eta'_0 + \mu^2\delta'_0 + \mu\delta_1] + \mathcal{O}(\varepsilon). \tag{8.34}$$

By inserting Eq. (8.34) into Eq. (8.29), we obtain a representation of F as a power series

$$F(\Psi^\varepsilon_{min}) = \sum_{\nu=-2}^{\infty} \varepsilon^\nu F_\nu(\Psi_h) \tag{8.35}$$

with

$$\begin{aligned} F_{-2}(\Psi^\varepsilon_{min}) &= \int_{x_l}^{x_r}\int_{-1}^{1} \mu^2\delta_0^2(x)d\mu dx \\ &= \frac{2}{3}\int_{x_l}^{x_r} \delta_0^2(x)dx \end{aligned} \tag{8.36}$$

and

$$\begin{aligned} F_{-1}(\Psi^\varepsilon_{min}) &= 2\int_{x_l}^{x_r}\int_{-1}^{1} \mu^2\eta'_0(x)\delta_0(x) + \mu^3\delta_0\delta'_0 + \mu^2\delta_0(x)\delta_1(x)d\mu dx \\ &= \frac{4}{3}\int_{x_l}^{x_r} \eta'_0(x)\delta_0(x) + \delta_0(x)\delta_1(x)dx. \end{aligned} \tag{8.37}$$

Since $0 \in V_h$ and Ψ^ε_{min} minimizes F, we obtain for $|\varepsilon| \leq 1$

$$F(\Psi^\varepsilon_{min}) \leq F(0) = \varepsilon^2 \int_{x_l}^{x_r} q(x)^2 dx \leq \int_{x_l}^{x_r} q(x)^2 dx \tag{8.38}$$

Therefore, we must have $F_{-2}(\Psi^\varepsilon_{min}) = 0$ and $F_{-1}(\Psi^\varepsilon_{min}) = 0$, since otherwise $F(\Psi)$ diverges for $\varepsilon \to 0$, which would contradict Eq. (8.38). We conclude that

$$\delta_0(x) = 0 \tag{8.39}$$

Using Eq. (8.39), we can restrict the set of functions to

$$\begin{aligned} W_h = \{w_h \in V : w_h(z,\mu) = \sum_{\nu=0}^{\infty} \varepsilon^\nu \eta_\nu(x) + \mu \sum_{\nu=1}^{\infty} \varepsilon^\nu \delta_\nu(x), \\ w_h(x_l) = 0 \text{ for } \mu < 0,\ w_h(x_r) = 0 \text{ for } \mu > 0\}. \end{aligned} \tag{8.40}$$

A necessary condition for the minimum is that the derivative of F with respect to Ψ is zero:

$$\int_{x_l}^{x_r}\int_{-1}^{1} (L_\varepsilon \Psi^\varepsilon_{min}(x,\mu) - \varepsilon Q(x,\mu)) L_\varepsilon \Psi^\varepsilon_{min}(x,\mu) d\mu dx = 0. \tag{8.41}$$

We have

$$\begin{aligned} L_\varepsilon \Psi^\varepsilon_{min} &= \mu \frac{\partial \phi^\varepsilon_0}{\partial x} + \mu^2 \frac{\partial \phi^\varepsilon_1}{\partial x} + \frac{1}{\varepsilon}\mu\phi^\varepsilon_1 + \varepsilon\alpha\phi^\varepsilon_0 \\ &= [\mu\eta_0' + \mu\delta_1] + \varepsilon[\mu\eta_1' + \mu^2\delta_1' + \mu\delta_2 + \alpha\eta_0] + \mathcal{O}(\varepsilon^2) \end{aligned} \tag{8.42}$$

and by inserting Eq. (8.42) into Eq. (8.41) we obtain

$$\int_{x_l}^{x_r} \int_{-1}^{1} \mu^2(\eta_0' + \delta_1)^2 d\mu dx + \varepsilon I_1 + \mathcal{O}(\varepsilon^2) = 0 \tag{8.43}$$

with

$$\begin{aligned} I_1 &= 2\int_{x_l}^{x_r} \int_{-1}^{1} \mu[\alpha\eta_0\eta_0' + \alpha\delta_1\eta_0] + \mu^2[\eta_0'\eta_1' + \eta_0'\delta_2 + \delta_1\eta_1' + \delta_1\delta_2] + \mu^3[\delta_1'\eta_0' + \delta_1\delta_1'] d\mu dx \\ &= \int_{x_l}^{x_r} \int_{-1}^{1} 2\mu^2[\eta_0'\eta_1' + \eta_0'\delta_2 + \delta_1\eta_1' + \delta_1\delta_2] d\mu dx. \end{aligned} \tag{8.44}$$

In Eq. (8.43) $\mathcal{O}(1)$ and $\mathcal{O}(\varepsilon)$ terms show up only on the left-hand side. Therefore, both terms need to vanish, what is implied by the relation

$$\eta_0' = -\delta_1. \tag{8.45}$$

This shows that only terms of order ε^2 remain, what ends the proof. ☐

Using Lemma 8.4.1, we can prove the main theorem on PINN solutions in the diffusion limit.

Theorem 8.4.2. *Let the functional F and the set of functions V_h be as given in Eqs. (8.29) and (8.26). Suppose that Ψ^ε_{min} minimizes F restricted to V_h. Suppose further that $\varepsilon \leq 1$ and that Ψ^ε_{min} has an expansion as defined in Eqs. (8.30) and (8.31).*
Then we have:

$$\Psi^\varepsilon_{min} \to 0 \text{ as } \varepsilon \to 0 \text{ pointwise, for all } x \in D. \tag{8.46}$$

Proof. If both $\eta_\nu(x)$ and $\delta_\nu(x)$ are continuous piecewise linear functions, it follows from Eq. (8.45) that η_0 is a linear function. With the boundary conditions we obtain $\eta_0 = 0$. Therefore, we have $\Psi^\varepsilon_{min} \to 0$ as $\varepsilon \to 0$. Since all η, δ are continuous, the convergence holds pointwise, for all $x \in D$. ☐

As a consequence, PINNs using feed forward neural networks with ReLU activation functions do not give a correct approximation of the neutron flux ϕ in the diffusion limit, except for $Q = 0$.

8.4.2. Scaling for the numerical solution

As shown in the main theorem, directly solving the neutron transport equation using a PINN with ReLU activation functions leads to a method that does not preserve the diffusion limit. Solving the neutron transport equation using first order least-squares finite elements leads to similar results ([67], [97]).

In [67], a scaling of the equations that solves this issue is proposed. In the following, we also use this scaling for PINNs.
Recall that

$$(P\Psi)(x) = \frac{1}{2}\int_{-1}^{1} \Psi(x,\mu)d\mu. \tag{8.47}$$

Using this operator, which is a projection onto the first Legendre moment, we define the scaling operator

$$S = P + \tau(I - P). \tag{8.48}$$

with $\tau \in \mathbb{R}^{\geq 0}$. By applying S on both sides of Eq. (8.6), we obtain

$$S\mu\frac{\partial\Psi}{\partial x} + \frac{\tau}{\varepsilon}(I-P)\Psi + \varepsilon\alpha P\Psi = \varepsilon Q. \tag{8.49}$$

We use the same expansion for Ψ as in Eq. (8.30) to project on the first two Legendre moments. The systems of scaled equations then reads

$$\begin{aligned} \frac{1}{3}\frac{\partial\phi_1^\varepsilon}{\partial x} + \alpha\varepsilon\phi_0^\varepsilon &= \varepsilon Q \\ \tau\frac{\partial\phi_0^\varepsilon}{\partial x} + \tau\phi_1^\varepsilon \quad &= 0 \end{aligned} \tag{8.50}$$

In the following, we are only interested in the behavior of the leading orders, so we take to terms in Eqs. (8.31) only up to order $\mathcal{O}(\varepsilon)$ in account. In addition, we use that $\delta_0 = 0$ according to Eq. (8.32). Therefore, we obtain

$$\begin{aligned} \phi_0 &= \eta_0 + \varepsilon\eta_1 \\ \phi_1 &= \varepsilon\delta_1. \end{aligned} \tag{8.51}$$

By inserting Eqs. (8.51) into Eqs. (8.50) and neglecting higher order terms, we obtain

$$\begin{aligned} \varepsilon\frac{1}{3}\frac{\partial\delta_1}{\partial x} + \varepsilon\alpha\eta_0 &= \varepsilon Q \\ \tau\frac{\partial\eta_0}{\partial x} \quad + \ \tau\delta_1 &= 0. \end{aligned} \tag{8.52}$$

The case $\tau = 1$ corresponds to the unscaled equation, where the first equation is $\mathcal{O}(\varepsilon)$ and the second equation $\mathcal{O}(1)$, so that Eqs. (8.52) are unbalanced for $\epsilon \to 0$. Choosing $\tau = \mathcal{O}(\varepsilon)$ results in a balancing of the terms in orders of ε. This is a common strategy to improve the convergence behavior. The numerical results in the following section demonstrate that the scaling leads to the correct diffusion limit. Here we use $\tau = \sqrt{\sigma_a/\sigma_t} = \sqrt{\alpha}\epsilon$, since this scaling directly relates to the physical parameters.

8.5. Numerical Results

In this section we present numerical results which illustrate the theory developed in Sec. 8.4. It is structured as follows: in Subsec. 8.5.1 we introduce the $\mathrm{P_N}$ method which was used for the angular discretization. In Subsec. 8.5.2 the implementation is described. In Subsec. 8.5.3 and Subsec. 8.5.4 we present numerical results for PINNs with ReLU activation function and first order LSFE to demonstrate how the scaling leads to a huge improvement in accuracy for both methods. In Subsec. 8.5.5 we demonstrate that the scaling can also improve the solution for activation functions other than ReLU.

8.5.1. Angular Discretization

For the angular discretization of Eq. (8.7), we use the method of moments (see 3.2.2), which is a spectral Galerkin method in μ [41]. Let $\phi_n(x)$ be the n-th angular moment of Ψ as defined in Eq. (8.24). Recall that by taking the first N Legendre polynomial moments of Eq. (8.1) and using the recursion relations for Legendre polynomials, we obtain the slab geometry $\mathrm{P_N}$ equations:

$$\frac{\partial \phi_1}{\partial x} + \sigma_a \phi_0 = Q \tag{8.53}$$

$$\frac{n}{2n+1}\frac{\partial \phi_{n-1}}{\partial x} + \frac{n+1}{2n+1}\frac{\partial \phi_{n+1}}{\partial x} + \sigma_n = 0 \ (n > 0). \tag{8.54}$$

We close the equations by setting

$$\phi_n = 0, \qquad n > N. \tag{8.55}$$

Note while we used the normed Legendre polynomials in the previous section to simplify the equations, we do not norm the Legendre polynomials here since otherwise we would not obtain the standard formulation of the $\mathrm{P_N}$ equations in this case. In the following we use vacuum boundary conditions, which assume no incoming flux, and reflective boundary conditions, which assume that all outgoing particles are reflected back into the domain. For an Nth order expansion where N is odd, there are $(N+1)/2$ vacuum boundary conditions that read

$$\begin{aligned} 0 &= \sum_{n=0}^{N} \frac{2n+1}{2}\phi_n(x_l)\int_0^1 P_{2m-1}(\mu)P_n(\mu)d\mu \text{ for } m = 1,2,...,(N+1)/2 \\ 0 &= \sum_{n=0}^{N} \frac{2n+1}{2}\phi_n(x_r)\int_0^{-1} P_{2m-1}(\mu)P_n(\mu)d\mu \text{ for } m = 1,2,...,(N+1)/2 \end{aligned} \tag{8.56}$$

where $P_n(\mu)$ is the n-th order Legendre polynomial on the interval $[-1,1]$. Reflective boundary conditions are obtained by setting the odd moments to zero at the boundary. More details about the boundary conditions for slab geometry $\mathrm{P_N}$ equations can be found in [37]. It is known, that in the limit $\varepsilon \to 0$ a solution of Equation 8.7 converges to the solution of a corresponding diffusion equation, which is independent of the angular variable μ. Therefore, in practice a small N ($N = 1$ or $N = 3$) is sufficient as an approximate model.

8.5.2. Implementation

The PINN solver used for the numerical computations was implemented using the python library DeepXDE [64] with tensorflow [1] as a backend. The boundary conditions were included in the loss function as an implementation of the restriction of the underlying set of functions. In order to do this, the term

$$\mathrm{Loss}_{\partial D}[\Psi] = \frac{w_{\partial D}}{|T_{\partial D}|}\sum_{m=1}^{(N+1)/2}\sum_{x\in T_{\partial D}} (B_m[\Psi](x) - g(x))^2, \tag{8.57}$$

where $w_{\partial D}$ is a weight, $T_{\partial D}$ a set of training points on the boundary and B_m the operator representing the mth boundary condition introduced in Subsec. 8.5.1, was added to the loss function (Eq. 8.11). For vacuum boundary conditions, we have $g = 0$ and the operator B_m reads

$$B_m[\Psi](x_l) = \sum_{n=0}^{N} \frac{2n+1}{2}\phi_n(x_l) \int_0^1 P_{2m-1}(\mu)P_n(\mu)d\mu \text{ for } m = 1, 2, ..., (N+1)/2$$
$$B_m[\Psi](x_r) = \sum_{n=0}^{N} \frac{2n+1}{2}\phi_n(x_r) \int_0^{-1} P_{2m-1}(\mu)P_n(\mu)d\mu \text{ for } m = 1, 2, ..., (N+1)/2 \qquad (8.58)$$

For reflective boundary conditions, we have $g = 0$ and

$$B_m[\Psi](x) = \phi_{2m+1}(x) \text{ for for } m = 1, 2, ..., (N+1)/2,\ x \in T_{\partial D}. \qquad (8.59)$$

We implemented a LSFE solver using the FEniCS library [3, 63]. We use PETSc [6] as linear algebra backend in conjunction with hypre [28], a library of high performance preconditioners. For this work we use a combination of a GMRES solver [88] and an algebraic multigrid preconditioner [86]. The boundary condition was implemented by adding the term

$$a_b(\Psi, u) := \int_{x_l}^{x_r} \int_{\partial D} B\Psi(x,\mu)Bu(x,\mu)d\mu ds \qquad (8.60)$$

to the least-squares bilinear form (Eq. 8.19).
The errors in this section are computed as follows: we define G points that are equidistantly distributed over the computational domain and compute the squared difference between the PINN or LSFE neutron flux and the reference solution divided by the squared reference solution. The relative error is then given by

$$\xi_{rel} = \sum_{g=0}^{G} \frac{(\phi_{0,PINN/LSFE}(x_g) - \phi_{0,\mathrm{ref}}(x_g))^2}{(\phi_{0,\mathrm{ref}}(x_g))^2}. \qquad (8.61)$$

We use the zeroth flux moment since it is identical to the angular-integrated total flux. Since PINNs may converge to different solutions from different initial values [76], we train PINNs from random initialization three times. The PINN results shown in this section are to be understood as the average of the three networks.

8.5.3. Asymptotic Test of the Diffusion Limit

The neutron transport equation in the diffusion limit is given by

$$\left(\mu\frac{\partial}{\partial x} + \frac{1}{\varepsilon}(I-P) + \varepsilon\alpha P\right)\Psi(x,\mu) = \varepsilon Q(x,\mu). \qquad (8.62)$$

With the source term

$$Q(x,\mu) = 1 + \alpha\phi_0(x) \qquad (8.63)$$

and vacuum boundary conditions, the neutron flux $\Psi(x,\mu)$ tends asymptotically to the solution of the diffusion equation:

$$\phi_0(x) = -\frac{3}{2}x^2 + 15x. \qquad (8.64)$$

In the following, we investigate the asymptotic behavior of the PINN and LSFE solutions in the diffusion limit. We choose $\alpha = 10^{-2}$. It should be noted that for $\varepsilon = 10^{-2}$, this choice leads to a problem described in [97].

For the PINN computations, we use feed forward neural networks, each with 5 hidden layers and 50 nodes per layer. An Adams optimizer with learning rate $l_r = 2.5 \cdot 10^{-4}$ is used for the training process. A total of 300 training points is used. For the LSFE computations, we use a mesh with 20 first order Lagrange Finite elements.

The results for three different values of ε are depicted in Fig. 8.3, while the relative errors can be found in Tab. 8.1.

For $\varepsilon = 10^{-2}$, all errors are below 1%. The scaling reduces the already small error even further for both PINN and LSFE. For $\varepsilon = 10^{-3}$, the unscaled solutions deviate significantly from the reference. While the error for the PINNs is 36.8%, the LSFE error is 18.9%. The scaling however reduces the error for both methods massively, the scaled PINN error is only 0.3%, while the scaled LSFE error is 0.5%. For $\varepsilon = 10^{-4}$, the unscaled solutions are close to zero. This is expected, since the unscaled solution is supposed to converge to zero as $\varepsilon \to 0$ according to the analysis in Sec. 8.4. The scaled solutions however have still a small error, 0.6% for the PINN and 1.8% for the LSFE.

This results are in accordance with the theoretical predictions in Sec. 8.4. As expected, for smaller values of ε the unscaled versions of PINN and LSFE converge against zero. The scaled solution as also always better than the unscaled solution. The results for both methods are similar, however the exact numerical values deviate, as one would expect for two methods with vast differences in the implementation.

Table 8.1.: Relative errors in Fig. 8.3, computed with Eq. 8.61.

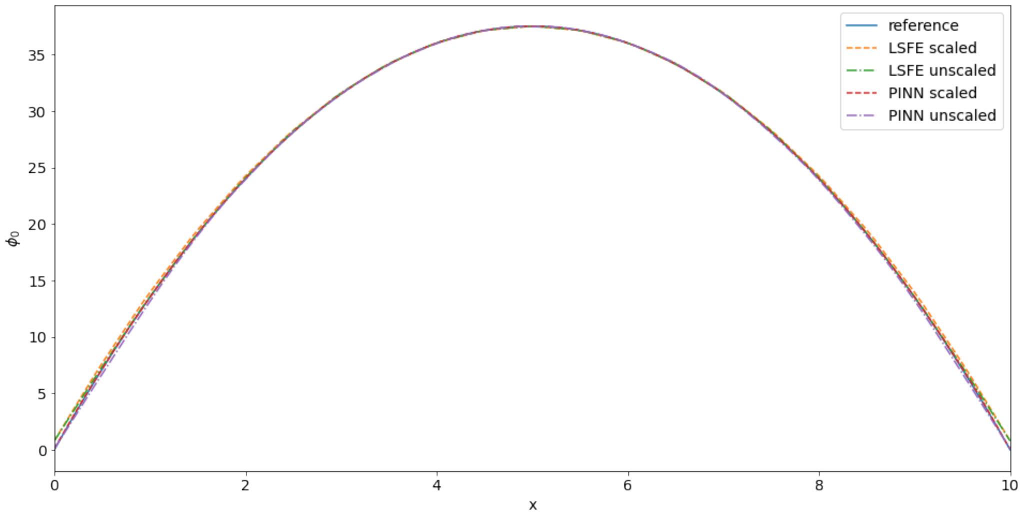
ε	10^{-2}	10^{-3}	10^{-4}
PINN unscaled	0.6%	36.8%	100.0%
PINN scaled	0.1%	0.3%	0.6%
LSFE unscaled	0.7%	18.9%	95.8%
LSFE scaled	0.3%	0.5%	1.8%

8.5.4. Diffusive Test with an Interface

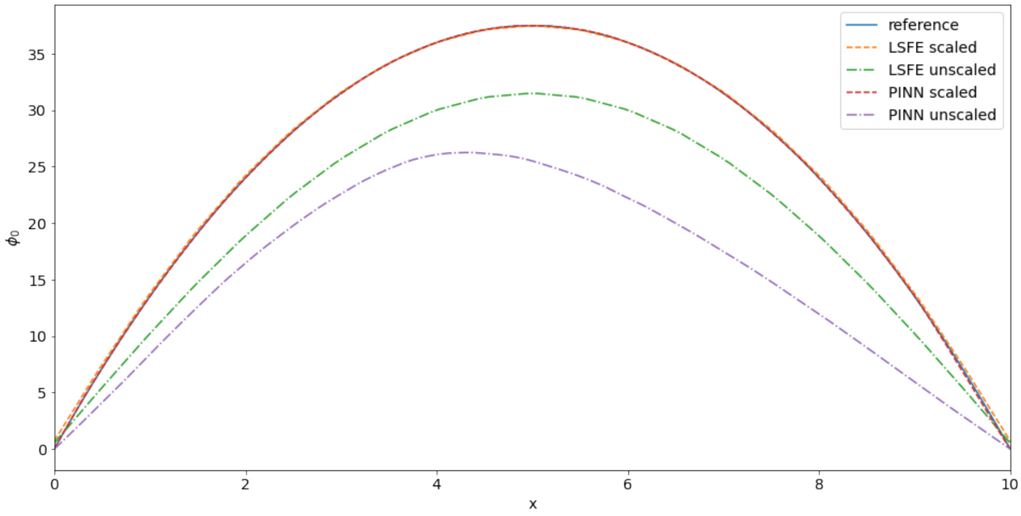
In this subsection we investigate a problem with an internal interface. The left side of the slab consists of a pure absorber ($\sigma_a = \sigma_t = 2,\ x \in (0,2)$), while the right side is a strong scatterer ($\sigma_t = 100,\ \sigma_a = 10^{-4},\ x \in (2,10)$) with only a very small absorption cross section. A constant source $Q = 1$ is applied on the left side of the slab, i.e. for $x \in (0,2)$. On the left boundary, reflective boundary conditions are imposed, while on the right boundary vacuum boundary conditions are used.

For the PINN computations, we use the same architecture and learning rate as in Subsec. 8.5.3. For this testcase, 450 training points within the domain are used. For the LSFE computations we use a mesh with 50 first order Lagrange Finite elements.

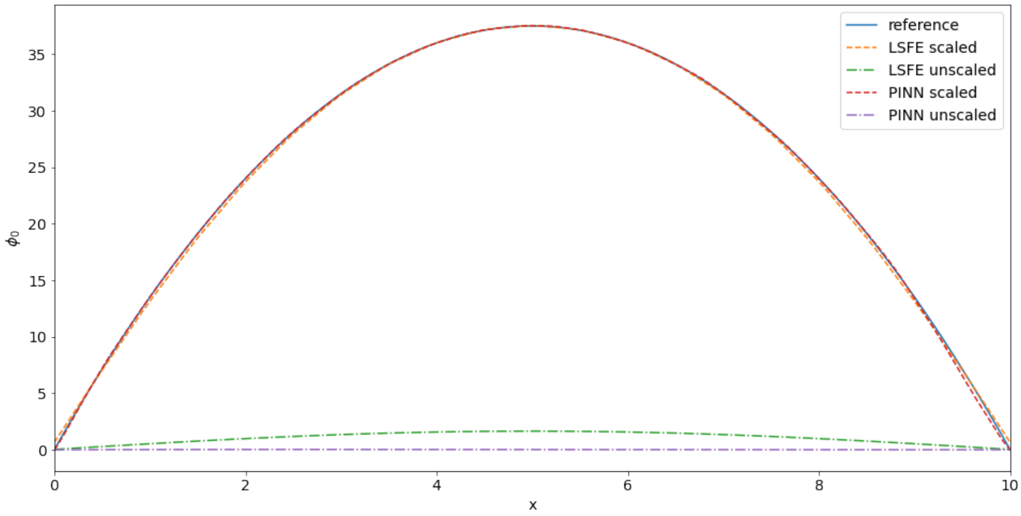
The reference solution was computed with the OpenMC Monte Carlo Code [85]. For details on solving the neutron transport equation with Monte Carlo methods, see [55]. OpenMC solves the full neutron transport equation (Eq. 2.1). That means it does not approximate the neutron angle, but includes the full angular dependence, e.g. all Legendre moments. Since the higher order moments are small in the diffusive regime, we do not expect large deviations, but we still use the P$_3$-approximation for the angular discretization to minimize potential errors.



(a) $\varepsilon = 10^{-2}$



(b) $\varepsilon = 10^{-3}$



(c) $\varepsilon = 10^{-4}$

Figure 8.3.: Results for PINN and LSFE in an asymptotic test of the diffusion limit for three different values of ε. Depicted are the results for the zeroth Legendre moment ϕ_0, which is identical to the angular-integrated total flux.

Fig. 8.4 depicts the results for the test with an internal interface. In both cases, the scaling leads to a huge improvement of the solution. While the unscaled PINN solution deviates from the reference by 10.6%, the scaled PINN solution deviates by only 1.4%. The LSFE are similar. While the error for the unscaled LSFE is 9.9%, the error of the scaled solution is only 0.6%. We again see the similarity of the PINN und LSFE solutions.

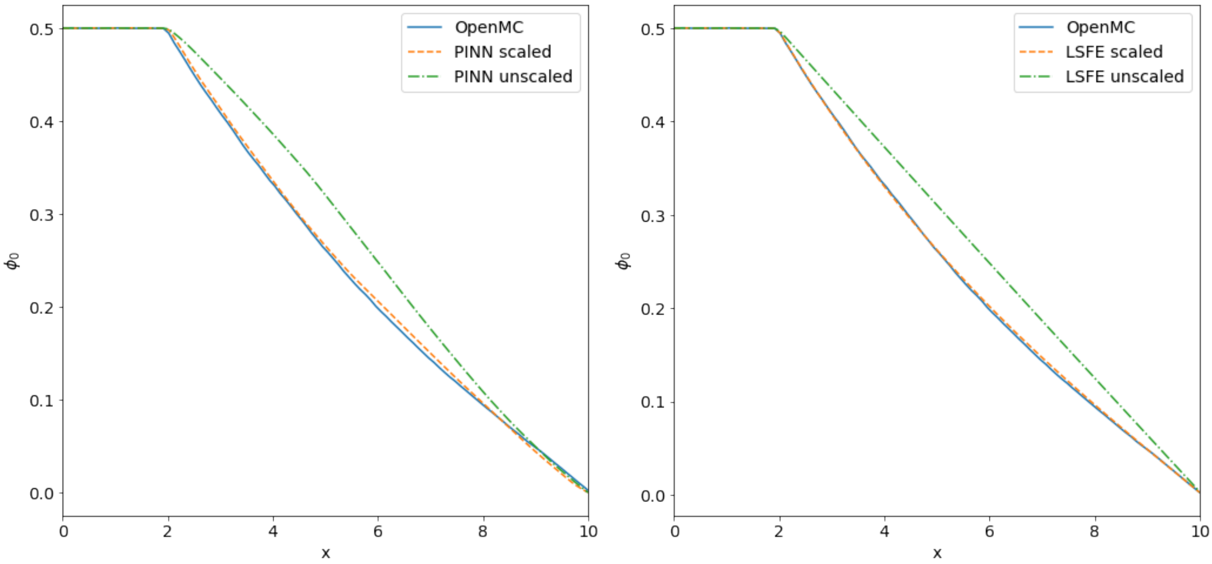


Figure 8.4.: Results for PINN and LSFE in a diffusive test with an internal interface. In both cases, the scaling leads to a huge improvement of the solution. Depicted are the results for the zeroth Legendre moment ϕ_0.

8.5.5. Diffusive Test with Hyperbolic Tangent Activation Functions

While the analysis in Subsec. 8.4.1 shows that PINNs with ReLU activation functions lead to a method that does not preserve the diffusion limit, the scaling introduced in Subsec. 8.4.2 to correct this does not assume the usage of a specific activation function. Therefore, it is close at hand to investigate the effect of the scaling on PINN with a different activation function. In this subsection, we investigate the problem with the internal interface introduced in Subsec. 8.5.4 for PINN with hyperbolic tangent activation functions. We use the same number of training points as before. As an optimizer L-BFGS is used since it showed better convergence properties for this testcase than the Adams optimizer. Fig. 8.5 depicts the results. The unscaled solution deviates form the reference by 23.4% and is therefore worse than the unscaled solution obtained with ReLU activation functions. The error of the scaled solution is only 0.5% and therefore smaller than in the case with ReLU activation functions. This shows that the proposed scaling leads to a vast improvement of the solution not only for ReLU activation functions, but also for hyperbolic tangent activation functions.

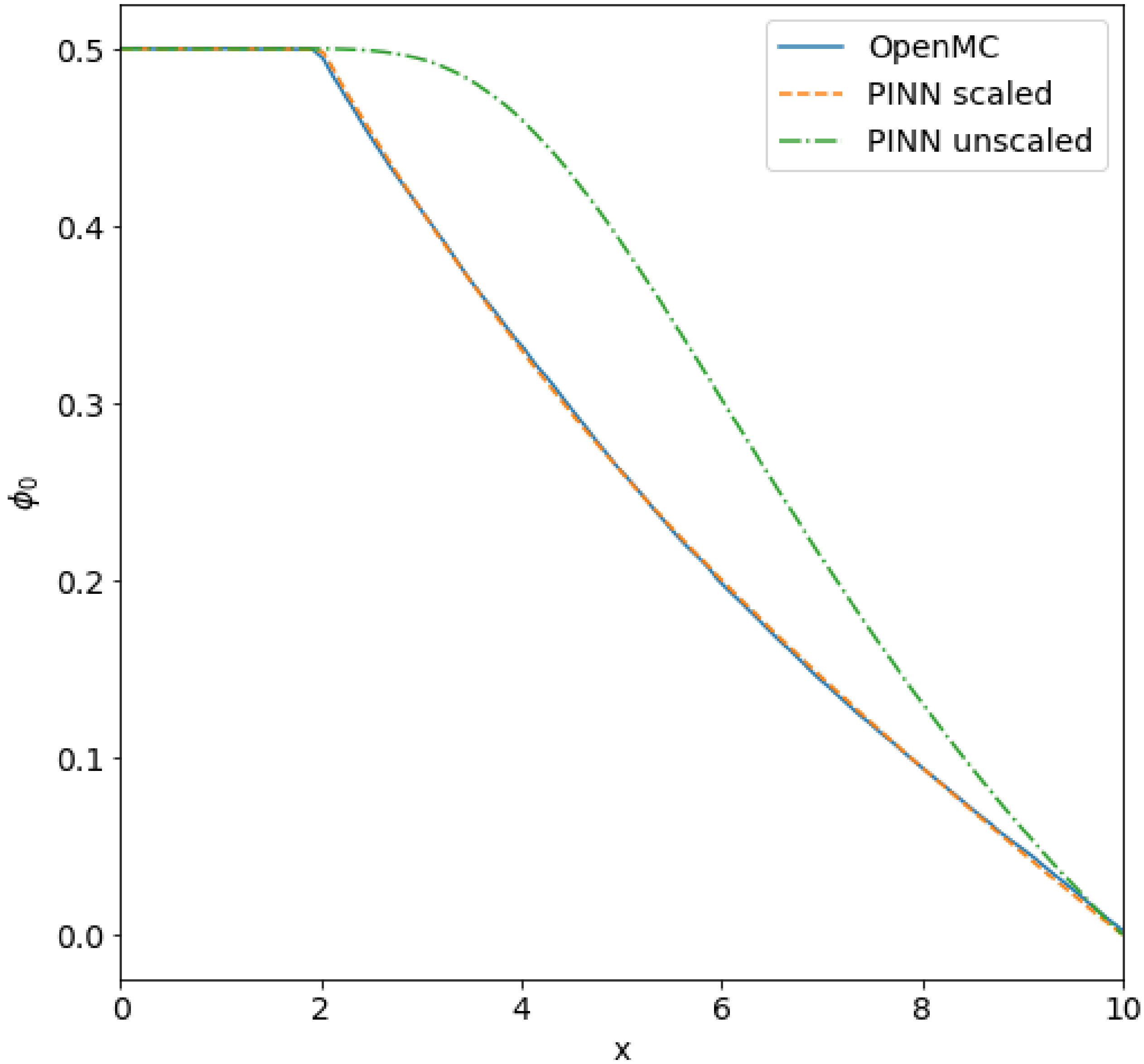


Figure 8.5.: Results for PINN with hyperbolic tangent activation functions for the diffusive test with an internal interface discussed in Subsec. 8.5.4. The scaling leads to a huge improvement of the solution. Depicted are the results for the zeroth Legendre moment ϕ_0.

9. Discussion, Conclusion and Outlook

In this work, we investigate numerical methods and machine learning for PGNAA. In the following, we summarize and discuss the key results and give an outlook on future research. This Chapter is divided into three parts. First, we will discuss the development of the deterministic neutron transport code SPARC and its application on the computation of partial cross section for PGNAA. Thereafter, we discuss the parameter estimation problem introduced in Chapter 7. We conclude with a discussion of the numerical robustness of PINNs and their connection to LSFE.

Computation of Partial Cross Section for PGNAA with SPARC

In Chapter 4, the development and verification of SPARC, an internal neutron transport code based on open source libraries, was described. Using SPARC, it was shown in Chapter 5 that partial (n,γ)-cross sections for PGNAA can be computed with neutron fluxes obtained using the SP_1- and SP_3-equations. Despite the relatively high model error for higher neutron energies, the partial (n,γ)-cross sections can be computed with sufficient accuracy. This is caused by the fact that the model error in the thermal energy range is minimal and the reaction rates are strongly dominated by the contribution of thermal neutrons. For lighter filling materials, the most relevant application case, the average error on the reaction rates is less than a single percent. For metallic materials the error is larger, e.g. about 4 % or 8 % in the test cases discussed above, however still acceptable. In practice, the drums will typically be filled with compacted scrap metal instead of massive metal like in these test cases. Due to technical limitations, for example of in-drum compactors, the resulting density is expected to be smaller than in the shown test cases. It is expected that the error will also be smaller. The SPARC runtime in the shown test cases was 20 to 30 times better than OpenMC using the same compute cluster. It is expected that the performance advantage of SPARC will be even larger when applied to the more detailed models of the measurement facility that will be used for future computations.

Parameter Estimation in PGNAA

In Chapter 7 we have considered a new kind of parameter estimation problem related to PGNAA where two different types of detectors are used, one to record a gamma spectrum and additional detectors for the total neutron flux in direct proximity to the sample. This parameter estimation problem arises in the measurement facility QUANTOM, where radioactive waste drums are analyzed for their elemental composition. We invert a combined neutron-photon transport model to obtain a vector of elemental densities. In a simplified setting considering a homogeneous sample and one-group neutron transport we investigated uniqueness of a solution. This was necessary since the objective function is not strictly convex, so the standard theory on uniqueness was not applicable.

To the current state of the art PGNAA spectra are evaluated in an iterative manner by solving a fixed point problem [2, 46]. This naturally reflects the dependence of the elemental composition and the neutron flux distribution in the sample. The method proposed in this work offers an alternative. Formulating the parameter estimation problem as an optimization problem we make use of a gradient descent method to solve for the elemental composition as a parameter of the neutron transport model. As a side effect, our method yields the elemental composition as absolute quantities in grams or kilograms. Furthermore, this approach is especially useful when several measurements are combined, since all measurements can be combined in the objective function. For the measurements in QUANTOM the relative formulation in [2, 46] does not apply. Hence, our method offers more flexibility to deal with a broader range of measurement tasks related to PGNAA. In situations where both methods may be applied (standard PGNAA with one measurement and one detector type) it is not obvious which methods yields the best performance in terms of convergence and computation times. This will be a topic of future research.

Furthermore, we see a close connection to the theory of Diffuse Optical Tomography (DOT). In both DOT and the PGNAA problems presented in Chapter 7, the governing equation is a diffusion equation. The measurement of the flux alone is not sufficient to ensure a unique reconstruction. In DOT, it is not possible to reconstruct both the scattering and absorption coefficients from the photon flux measurements alone [5]. Similarly, we have seen that the neutron flux is insufficient to reconstruct the elemental densities, as discussed in Sec. 7.3. Only the combination with the energy resolved measurements of the gamma detector, which yields measurement results for each element individually, ensures uniqueness. While in DOT no general uniqueness result exists due to the range of different measurables and unknowns [35], it is still possible to obtain uniqueness results for certain scenarios using additional assumptions on the scattering and absorption coefficients that are to be reconstructed [38]. For PGNAA comparable studies do not exist yet since this work represents the first result in this direction. Therefore, a complete study of possible necessary and sufficient conditions for existence and uniqueness of solutions should be a focus of future work.

Throughout Chapter 7, we assumed homogeneity of the sample. This case is of practical importance for PGNAA, since most measurements are carried out with homogeneous samples, especially at research reactors [73]. Monitoring of the total neutron flux is applied by monitor materials in this case. Hence, our results may be adapted also for this situation.

Our investigation of uniqueness covered the homogeneous case. However, the case of an inhomogeneous sample is also of great interest. In the QUANTOM facility, this situation is covered by performing numerous measurements at different measurement positions, corresponding to a complete surface scan of the drum by the HPGe detector. Measurements of the total neutron flux by the 3He detectors are also performed in every measurement position. A joint evaluation of all measurement positions shall account for inhomogeneities in the drum. Nevertheless, an investigation of the uniqueness of a solution in this situation needs to be performed in the future. It should be noted that the gradient based method in introduced in Sec. 7.4 holds for the inhomogeneous case. Since more measurements are necessary, the objective function will include more terms. The method relies then on the assumption of homogeneity in discretized parts of the drum.

Numerical Robustness of PINNs

As can be seen in Chapter 5, the performance of neutron transport simulations for large sample PGNAA can be significantly improved by the use of deterministic solvers instead of more costly Monte Carlo computations. However, the neutron transport simulations remain a bottleneck of the entire evaluation workflow. Therefore, we investigated PINNs [79], an alternative approach to solve PDEs. In Chapter 8, we investigated the numerical stability of PINNs for multiscale transport problems. As an exemplary problem we studied the example of the neutron transport in the so-called diffusive regime. This example is great relevance for large-sample PGNAA measurement facilities, since the neutron transport in moderator materials like graphite takes place in the diffusive regime. We used an analogy between PINNs and LSFE that lies in their shared approach of reformulating the solution of differential equations as the minimization of a (typically quadratic) functional. By making use of this analogy we were able to build on a theory for LFSE and adapt it for PINNs. It was shown that PINNs with ReLU activation functions yield incorrect approximate solutions in the diffusion limit, similar to first-order LSFE. It was also demonstrated that a scaling can be applied to PINNs that leads to the correct diffusion limit. These theoretical findings were underlined with numerical results.

A formal proof for the convergence of the scaled solution remains a goal for future research. For LSFE, a corresponding result exists. The proof uses the fact that the quadratic functional needs to be minimized for all admissible test functions. Then, by choosing specific test functions, in addition to the relations in lemma 8.4.1 it can be shown, that the leading order in the formal expansion of ϕ_0^ε satisfies a variational form of the diffusion equation. Details can be found in [80]. Since in the PINN context no direct analogon for the test functions exists, this cannot be mimicked with PINNs. Therefore, we cannot directly adopt the theory for LSFE and a different way to formally prove the convergence of the scaled PINN solution needs to be found.
We see the possibility to build on further results of the existing theory for LSFE and adapt it for PINNs, too. For instance, there are cases known for LSFE were scalings other than the one we applied to PINNs are used. Taking additional parameters, such as the medium optical thickness or the local cell size, into account, these scalings provide a further improved accurracy. In [101], as an example, such a scaling is introduced which is especially useful for optically thin materials.
We showed that unscaled PINNs with ReLU activation functions do not converge to the correct diffusion limit. First order LSFE exhibit the same behavior. For LSFE, the already existing theory shows that using higher order finite elements instead of a scaling can be sufficient to migitate this issue, even though applying the scaling still improves the results numerically [80]. It is not apparent whether a PINN analogy to higher order finite elements exists and how it would look like. However, it is clear that simply choosing other activation functions is not sufficient, as our testcase with the hyperbolic tangent (see Subsec. 8.5.5) shows.

Bibliography

[1] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dandelion Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. URL `https://www.tensorflow.org/`. Software available from tensorflow.org. 8.5.2

[2] H. Akkurt, J.P. Holloway, and L.E. Smith. The fixed point formulation for large sample PGNAA – Part 2: experimental demonstration. *Nuclear Instruments and Methods in Physics Research A*, 522:545–557, 2004. 1.3, 7, 9

[3] M. Alnæs, J. Blechta, J. Hake, A. Johansson, B. Kehlet, A. Logg, C. Richardson, J. Ring, M. Rognes, and G. Wells. The FEniCS Project Version 1.5. *Archive of Numerical Software*, 3(100):9–23, 2015. 3, 4.1, 4.3, 5.2, 8.5.2

[4] D.L. Anderson and Z. Kasztovszky. Applications of PGAA with Neutron Beams. In G. Mjolnár, editor, *Handbook of Prompt Gamma Activation Analysis*, pages 137—-172. Kluwer Academic Publishers, Dordrecht, The Netherlands, 2004. (document), 1

[5] S. Arridge and W. Lionheart. Nonuniqueness in diffusion-based optical tomography. *Opt. Lett.*, 23(11):882–884, 1998. doi: 10.1364/OL.23.000882. 7.3.1, 9

[6] S. Balay, S. Abhyankar, M. Adams, J. Brown, P. Brune, K. Buschelman, L. Dalcin, A. Dener, V. Eijkhout, W. Gropp, D. Karpeyev, D. Kaushik, M. Knepley, D. May, L. Curfman McInnes, R. Mills, T. Munson, K. Rupp, P. Sanan, B. Smith, S. Zampini, H. Zhang, and H. Zhang. PETSc users manual. Technical Report ANL-95/11 - Revision 3.14, Argonne National Laboratory, 2020. 4.1, 8.5.2

[7] C. Bardos, R. Dautray, A. Craig, I.N. Sneddon, J.L. Lions, M. Cessenat, A. Kavenoky, P. Lascaux, B. Mercier, O. Pironneau, et al. *Mathematical Analysis and Numerical Methods for Science and Technology: Volume 6 Evolution Problems II*. Mathematical Analysis and Numerical Methods for Science and Technology. Springer Berlin Heidelberg, 1999. ISBN 9783540661023. 2.2

[8] J. Blatt and V. Weisskopf. *Theoretical Nuclear Physics*. Wiley, New York, 1960. 5.1

[9] P.B. Bochev and M.D. Gunzburger. *Least-Squares Finite Element Methods*. Applied Mathematical Sciences. Springer New York, 2009. ISBN 9780387689227. 8.3

[10] William Boyd, Adam Nelson, Paul K. Romano, Samuel Shaner, Benoit Forget, and Kord Smith. Multigroup Cross-Section Generation with the OpenMC Monte Carlo Particle Transport Code. *Nuclear Technology*, 205(7):928–944, 2019. doi: 10.1080/00295450.2019.1571828. 3.1.2, 4.1, 4.3, 5.2

[11] P. Brantley and E. Larson. The Simplified P_3 Approximation. *Nuclear Science and Engineering*, 134:1–21, 2000. 3.2.3, 4.3

[12] P. Brennecke. Requirements on Radioactive Waste for Disposal (Waste Acceptance Requirements of December 2014) - Konrad Repository - . Technical Report SE-IB-29/08-REV-2, BGE, in German, 2014. URL `https://www.bge.de/fileadmin/user_upload/Konrad/Wesentliche_Unterlagen/Endlagerungsbedingungen_Konrad/Endlagerungsbedingungen_Konrad_Stand_12_2014.pdf`. 1.3

[13] D.A. Brown, M.B. Chadwick, R. Capote, A.C. Kahler, A. Trkov, M.W. Herman, A.A. Sonzogni, Y. Danon, A.D. Carlson, M. Dunn, D.L. Smith, G.M. Hale, G. Arbanas, R. Arcilla, C.R. Bates, B. Beck, B. Becker, F. Brown, R.J. Casperson, J. Conlin, D.E. Cullen, M.-A. Descalle, R. Firestone, T. Gaines, K.H. Guber, A.I. Hawari, J. Holmes, T.D. Johnson, T. Kawano, B.C. Kiedrowski, A.J. Koning, S. Kopecky, L. Leal, J.P. Lestone, C. Lubitz, J.I. Márquez Damián, C.M. Mattoon, E.A. McCutchan, S. Mughabghab, P. Navratil, D. Neudecker, G.P.A. Nobre, G. Noguere, M. Paris, M.T. Pigni, A.J. Plompen, B. Pritychenko, V.G. Pronyaev, D. Roubtsov, D. Rochman, P. Romano, P. Schillebeeckx, S. Simakov, M. Sin, I. Sirakov, B. Sleaford, V. Sobes, E.S. Soukhovitskii, I. Stetcu, P. Talou, I. Thompson, S. van der Marck, L. Welser-Sherrill, D. Wiarda, M. White, J.L. Wormald, R.Q. Wright, M. Zerkle, G. Žerovnik, and Y. Zhu. ENDF/B-VIII.0: The 8^{th} major release of the nuclear reaction data library with CIELO-project cross sections, new standards and thermal scattering data. *Nuclear Data Sheets*, 148:1 – 142, 2018. ISSN 0090-3752. doi: https://doi.org/10.1016/j.nds.2018.02.001. Special Issue on Nuclear Reaction Data. 5.1

[14] K. Burns, L. Smith, C. Gesh, and M. Shaver. Multi-group transport methods for high-resolution neutron activation analysis. In *Proc. Int. Conf. on Mathematics, Computational Methods and Reactor Physics (M&C2009)*, Saratoga Springs, New York, May 3–7, 2009. American Nuclear Society. 5

[15] Richard H. Byrd, Peihuang Lu, Jorge Nocedal, and Ciyou Zhu. A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing*, 16: 1190–1208, September 1995. ISSN 1064-8275. doi: 10.1137/0916069. 8.2

[16] Shengze Cai, Zhicheng Wang, Sifan Wang, Paris Perdikaris, and George Karniadakis. Physics-informed neural networks for heat transfer problems. *J. Heat Transfer*, 143(6), 2021. doi: https://doi.org/10.1115/1.4050542. 8

[17] Cai, Shengze and Mao, Zhiping and Wang, Zhicheng and Yin, Minglang and Karniadakis, George. Physics-informed neural networks (pinns) for fluid mechanics: a review. *Acta Mechanica Sinica*, 37(12), 2021. doi: 10.1007/s10409-021-01148-1. 8

[18] K.M. Case and P.F. Zweifel. *Linear Transport Theory*. Addison-Wesley Publishing Company, 1967. 3.2.2, 8.4

[19] C. Cercignani. *The Boltzmann Equation and Its Applications*. AMS, 1988. 7.1

[20] M.B. Chadwick, M. Herman, P. Obložinský, M.E. Dunn, Y. Danon, A.C. Kahler, D.L. Smith, B. Pritychenko, G. Arbanas, R. Arcilla, R. Brewer, D.A. Brown, R. Capote, A.D. Carlson, Y.S. Cho, H. Derrien, K. Guber, G.M. Hale, S. Hoblit, S. Holloway, T.D. Johnson, T. Kawano, B.C. Kiedrowski, H. Kim, S. Kunieda, N.M. Larson, L. Leal, J.P. Lestone, R.C. Little, E.A. McCutchan, R.E. MacFarlane, M. MacInnes, C.M. Mattoon, R.D. McKnight, S.F. Mughabghab, G.P.A. Nobre, G. Palmiotti, A. Palumbo, M.T. Pigni, V.G. Pronyaev, R.O. Sayer, A.A. Sonzogni, N.C. Summers, P. Talou, I.J. Thompson, A. Trkov, R.L. Vogt, S.C. van der Marck, A. Wallner, M.C. White, D. Wiarda, and P.G. Young. ENDF/B-VII.1 nuclear data for science and technology: Cross sections, covariances, fission product yields and decay data. *Nuclear Data Sheets*, 112(12):2887 – 2996, 2011. ISSN 0090-3752. doi: 10.1016/j.nds.2011.11.002. Special Issue on ENDF/B-VII.1 Library. 3.1.2

[21] H. Choi, R. Firestone, R. Lindstrom, G. Molnár, S. Mughabghab, R. Paviotti-Corcuera, Z. Revay, A. Trkov, V. Zerkin, and C. Zhou. Database of Prompt Gamma Rays from Slow Neutron Capture for Elemental Analysis. Technical Report STI/PUB/1263, International Atomic Agency, 2007. 1.1, 5.1

[22] L. Coquard, A. Havenith, J. Kettler, K. Krycki, B. Fu, G. Nordhardt, J. Hummel, T. Veltkamp, S. Wegener, O. Schumann, and T. Koeble. Non-destructive Material Characterization of Radioactive Waste Packages with QUANTOM. In *WM2020 Conference*, Phoenix, Arizona, 3 2020. (document), 1.3

[23] L. Coquard, J. Hummel, G. Nordhardt, M. Georgi, A. Havenith, K. Krycki, B. Fu, C. Helmes, M. Heidner, F. Simons, T. Köble, and O. Schumann. Non-destructive verification of materials in waste packages using QUANTOM. *EPJ Nuclear Sci. Technol.*, 9 (5), 2023. doi: https://doi.org/10.1051/epjn/2022043. (document), 1.3, 1.4

[24] Salvatore Cuomo, Vincenzo Schiano Di Cola, Fabio Giampaolo, Gianluigi Rozza, Maizar Raissi, and Francesco Piccialli. Scientific machine learning through physics–informed neural networks: Where we are and what's next. *Journal of Scientific Computing*, 88(3), 2022. doi: 10.1007/s10915-022-01939-z. 8

[25] Tim De Ryck and Siddhartha Mishra. Numerical analysis of physics-informed neural networks and related models in physics-informed machine learning. *Acta Numerica*, 33: 633–713, 2024. doi: 10.1017/S0962492923000089. 8

[26] J.J. Duderstadt and L.J. Hamilton. *Nuclear reactor analysis*. John Wiley and Sons, New York, 1976. 2.4

[27] J. E. Morel Edward W. Larsen and John M. McGhee. Asymptotic derivation of the multigroup p1 and simplified pn equations with anisotropic scattering. *Nuclear Science and Engineering*, 123(3):328–342, 1996. doi: 10.13182/NSE123-328. 3.2.3

[28] R. Falgout and U. Meier Yang. hypre: A library of high performance preconditioners. In *International Conference on Computational Science*, Amsterdam, The Netherlands, April 21—24, 2002. Springer. 4.1, 8.5.2

[29] P. Filß. Specific activity of large-volume sources determined by a collimated external detector. *Kerntechnik*, 54(3):198–201, 1989. doi: https://doi.org/10.1515/kern-1989-540326. 1.3

[30] P. Filß. Relation between the activity of a high-density waste drum and its gamma count rate measured with an unshielded Ge-detector. *Applied Radiation Isotopes*, 48(8):805–812, 1995. doi: 10.1364/OL.23.000882. 1.3

[31] H. Finnemann, F. Bennewitz, and M. Wagner. Interface current techniques for multidimensional reactor calculations. *Atomkernenergie,(Germany)*, 30(2), 1977. 6.4

[32] Martin Frank, Axel Klar, Edward W. Larsen, and Shugo Yasuda. Time-dependent simplified PN approximation to the equations of radiative transfer. *Journal of Computational Physics*, 226(2):2289–2305, 2007. ISSN 0021-9991. doi: https://doi.org/10.1016/j.jcp.2007.07.009. 3.2.3

[33] E. Gelbard. Application of Spherical Harmonics Methods to Reactor Problems. Technical Report WAPD-BT-20, Bettis Atomic Power Laboratory, 1960. 3.2.3

[34] C. Geuzaine and F. Remacle. Gmsh: a three-dimensional finite element mesh generator withbuilt-in pre- and post-processing facilities. *International Journal for Numerical Methods in Engineering*, 79(11):1309–1331, 2009. 4.1

[35] A P Gibson, J C Hebden, and S R Arridge. Recent advances in diffuse optical imaging. *Physics in Medicine and Biology*, 50(4):R1, feb 2005. doi: 10.1088/0031-9155/50/4/R01. URL `https://dx.doi.org/10.1088/0031-9155/50/4/R01`. 9

[36] G. Gilmore. *Practical Gamma-Ray Spectrometry.* John Wiley and Sons, 2008. 7.1

[37] Steven P. Hamilton and Thomas M. Evans. Efficient solution of the simplified pn equations. *Journal of Computational Physics*, 284:155–170, 2015. ISSN 0021-9991. doi: https://doi.org/10.1016/j.jcp.2014.12.014. 8.5.1

[38] Bastian Harrach. On uniqueness in diffuse optical tomography. *Inverse Problems*, 25(5): 055010, apr 2009. doi: 10.1088/0266-5611/25/5/055010. URL `https://dx.doi.org/10.1088/0266-5611/25/5/055010`. 9

[39] Cory Hauck and Ryan McClarren. Positive p_n closures. *SIAM Journal on Scientific Computing*, 32(5):2603–2626, 2010. doi: 10.1137/090764918. 3.2.2

[40] Juncai He, Lin Li, Jinchao Xu, and Chunyue Zheng. Relu deep neural networks and linear finite elements. *Journal of Computational Mathematics*, 38(3):502–527, 2020. doi: https://doi.org/10.4208/jcm.1901-m2018-0160. 8.2

[41] Jan S. Hesthaven, Sigal Gottlieb, and David Gottlieb. *Spectral Methods for Time-Dependent Problems.* Cambridge Monographs on Applied and Computational Mathematics. Cambridge University Press, 2007. 8.5.1

[42] F. Himmerkus, R. Steiner, and L. Dorr. Requalification of historical intermediate level waste (ILW) for the repository KONRAD on basis of non destructive assay (NDA) and re-evaluated nuclide vectors. In *WM2015 Conference*, Phoenix, Arizona, March 2015. URL `https://www.osti.gov/biblio/22822805`. 1.3

[43] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich. *Optimization with PDE Constraints.* Springer, 2009. doi: 10.1007/978-1-4020-8839-1. URL `https://doi.org/10.1007/978-1-4020-8839-1`. 7.2

[44] Hobson, G. and Bolloni, H.W. and Breith, K.A. and Curca-Tivig, F. and Dall'Osso, A. and van Geemert, R. and Heinecke, J. and Hartmann, B. and Merk, S. and Porsch, D. and Pothet, B., and Tiles, V. ARTEMIS: The core simulator of AREVA NP's next generation coupled neutronics/thermal-hydraulics code system ARCADIA. In *Proceedings PHYSOR 2008*, Interlaken, Switzerland, 2008. 6.4

[45] J. Hoelscher, I. Harms, V. Bieling, and M. Ast. Risks of Hazardous Substances in Radioactive Waste for the Konrad Repository in Germany. In *WM2017 Conference*, Phoenix, Arizona, March 2017. URL `https://www.osti.gov/biblio/22802493`. 1.3

[46] J.P. Holloway and H. Akkurt. The fixed point formulation for large sample PGNAA – Part 1: theory. *Nuclear Instruments and Methods in Physics Research A*, 522:529–544, 2004. 1.3, 7, 9

[47] J. Hu, S. Jin, and Q. Li. Chapter 5 - asymptotic-preserving schemes for multiscale hyperbolic and kinetic equations. In Rémi Abgrall and Chi-Wang Shu, editors, *Handbook of Numerical Methods for Hyperbolic Problems*, volume 18 of *Handbook of Numerical Analysis*, pages 103–129. Elsevier, 2017. doi: https://doi.org/10.1016/bs.hna.2016.09.001. 8.1

[48] A. Jesser, K. Krycki, and M. Frank. Partial Cross-Section Calculations for PGNAA Based on a Deterministic Neutron Transport Solver. *Nuclear Technology*, 208(7):1114–1123, 2022. doi: 10.1080/00295450.2021.2016018. (document), 7.1

[49] A. Jesser, K. Krycki, and M. Frank. Numerical solution of a parameter estimation problem arising in prompt-gamma neutron activation analysis. *Applied Mathematics in Science and Engineering*, 32(1), 2024. doi: 10.1080/27690911.2024.2336170. (document)

[50] A. Jesser, K. Krycki, R. McClarren, and M. Frank. Numerical robustness of pinns for multiscale transport equations, 2024. URL `https://arxiv.org/abs/2412.14683`. (document)

[51] Shi Jin, Zheng Ma, and Keke Wu. Asymptotic-preserving neural networks for multiscale time-dependent linear transport equations. *Journal of Scientific Computing*, 94, 01 2023. doi: 10.1007/s10915-023-02100-0. 8.1

[52] G.E. Karniadakis, I.G. Kevrekidis, and L. Lu et al. Physics-informed machine learning. *Nature Review Physics*, 3:422—440, 2021. doi: https://doi.org/10.1038/s42254-021-00314-5. 8, 8.2

[53] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. In *International Conference for Learning Representations*, 2015. 8.2

[54] A. Lane and J. Lynn. Theory of radiative capture in the resonance region. *Nuclear Physics*, 17:563–585, 1960. 5.1

[55] B. Lapeyre, E. Pardoux, R. Sentis, A. Craig, and F. Craig. *Introduction to Monte Carlo Methods for Transport and Diffusion Equations*. Introduction to Monte Carlo Methods for Transport and Diffusion Equations. Oxford University Press, 2003. ISBN 9780198525936. 5.2, 8.5.4

[56] E. W. Larsen. Diffusion theory as an asymptotic limit of transport theory for nearly critical systems with small mean free path. *Ann. Nuclear Energy*, 7:249–255, 1980. doi: https://doi.org/10.1016/0306-4549(80)90072-9. 8.1

[57] E. W. Larsen. The asymptotic diffusion limit of discretized transport problems. *Nuclear Science and Engineering*, 112(4):336—-346, 1992. doi: https://doi.org/10.13182/NSE92-A23982. 8.1

[58] Edward W. Larsen and Jim E. Morel. Advances in Discrete-Ordinates Methodology. In Yousry Azmy and Enrico Sartori, editors, *Nuclear Computational Science, A Century in Review*, pages 1–84. Springer, Berlin, 2010. doi: 10.1007/978-90-481-3411-3. 3

[59] Jaakko Leppänen, Maria Pusa, Tuomas Viitanen, Ville Valtavirta, and Toni Kaltiaisenaho. The serpent monte carlo code: Status, development and applications in 2013. *Annals of Nuclear Energy*, 82:142–150, 2015. ISSN 0306-4549. doi: 10.1016/j.anucene.2014.08.024. 3.1.2

[60] E.E. Lewis and W.F. Miller. *Computational Methods of Neutron Transport*. Wiley, 1984. ISBN 9780471092452. 3.2.2, 8.1

[61] R. Lindstrom and Z. Reváy. Beams and Facilities. In G. Mjolnár, editor, *Handbook of Prompt Gamma Activation Analysis*, pages 31–58. Kluwer Academic Publishers, Dordrecht, The Netherlands, 2004. 5.2

[62] R. Lindstrom and Ch. Yonezawa. Samples and Standards. In G. Mjolnár, editor, *Handbook of Prompt Gamma Activation Analysis*, pages 59–70. Kluwer Academic Publishers, Dordrecht, The Netherlands, 2004. 1.3

[63] A. Logg, K. Mardal, and G. Wells. *Automated Solution of Differential Equations by the Finite Element Method*. Springer, 2012. doi: 10.1007/978-3-642-23099-8. 3, 4.1, 4.3, 5.2, 8.5.2

[64] Lu Lu, Xuhui Meng, Zhiping Mao, and George Em Karniadakis. DeepXDE: A deep learning library for solving differential equations. *SIAM Review*, 63(1):208–228, 2021. doi: 10.1137/19M1274067. 8.5.2

[65] Robert Macfarlane, Douglas W. Muir, R. M. Boicourt, Albert Comstock Kahler, III, and Jeremy Lloyd Conlin. The NJOY Nuclear Data Processing System, Version 2016. Technical report, Los Alamos National Laboratory, 1 2017. 3.1.2

[66] T. Manteuffel and K. Ressel. Multilevel methods for transport equations in the diffusive regime. *NASA Langley Research Center, The Sixth Copper Mountain Conference on Multigrid Methods, Part 2*, 1993. 8, 8.1

[67] T. Manteuffel and K. Ressel. Least-squares finite element solution of the neutron transport equation in diffusive regimes. *SIAM Journal on Numerical Analysis*, 35 (2), 1998. doi: https://doi.org/10.1137/S0036142996299708. 8, 8.1, 8.3, 8.4.2

[68] T. Manteuffel, K. Ressel, and G. Starke. A Boundary Functional for the Least-Squares Finite- Element Solution of Neutron Transport Problems. *SIAM Journal on Numerical Analysis*, 37 (2):556–586, 2000. doi: https://doi.org/10.1137/S0036142998344706. 8, 8.1

[69] E. Mauerhofer and A. Havenith. The MEDINA facility for the assay of the chemotoxic inventory of radioactive waste packages. *Journal of radioanalytical and nuclear chemistry*, 1:1–15, 2014. ISSN 1588-2780. doi: 10.1007/s10967-014-3210-2. 1.1

[70] R. McClarren. Theoretical Aspects of the Simplified Pn Equations. *Transport Theory and Statistical Physics*, 15(2-4):73–109, 2010. 3.2.2, 3.2.3

[71] R. McConn, C. Gesh, R. Pagh, R. Rucker, and R. Williams. Compendium of Material Composition Data for Radiation Transport Modeling. Technical Report PNNL-15870 Rev. 1, Pacific Northwest National Laboratory, 2011. 5.2

[72] J. McGhee, T. Wareing, D. Barnett, K. Thompson, A. Samardzie, and I. Davis. Attila User's Manual. *Transpire Inc*, 15, 2007. 4, 5

[73] G. Molnar. *Handbook of Prompt Gamma activation Analysis with Neutron Beams*. Kluwer Academic Publishers, 2004. (document), 1, 1.3, 9

[74] S. Mughabghab. Verification of the Lane–Lynn theory of direct neutron capture. *Phys. Lett. B*, 81:93–97, 1979. 5.1

[75] S. Mughabghab, M. Lone, and B. Robertson. Quantitative test of the Lane–Lynn theory of direct radiative capture of thermal neutrons by ^{12}C and ^{13}C. *Phys. Rev. C*, 26:2698–2701, 1982. 5.1

[76] Guofei Pang, Lu Lu, and George Em Karniadakis. fPINNs: Fractional Physics-Informed Neural Networks. *SIAM Journal on Scientific Computing*, 41(4):A2603–A2626, 2019. doi: 10.1137/18M1229845. 8.5.2

[77] G.C. Pomraning. Asymptotic and variational derivations of the simplified PN equations. *Annals of Nuclear Energy*, 20(9):623–637, 1993. ISSN 0306-4549. doi: https://doi.org/10.1016/0306-4549(93)90030-S. 3.2.3

[78] Anil K. Prinja and Edward W. Larsen. General Principles of Neutron Transport. In Dan Gabriel Cacuci, editor, *Handbook of Nuclear Engineering*, pages 427–542. Springer US, Boston, MA, 2010. doi: 10.1007/978-0-387-98149-9_5. 2.3, 2.3, 3.1, 3.2

[79] M. Raissi, P. Perdikaris, and G.E. Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378, 1:686–707, 2019. doi: https://doi.org/10.1016/j.jcp.2018.10.045. (document), 8, 8.2, 9

[80] Klaus J. Ressel. *Least-squares finite-element solution of the neutron transport equation in diffusive regimes*. Phd thesis, University of Colorado, Denver, 1994. 8.4.1, 9

[81] Z. Revály and T. Belgya. Principles of the PGAA method. In G. Mjolnár, editor, *Handbook of Prompt Gamma Activation Analysis*, pages 1–30. Kluwer Academic Publishers, Dordrecht, The Netherlands, 2004. 5.1

[82] Zsolt Revay. Determining elemental composition using prompt gamma activation analysis. *Analytical chemistry*, 81:6851–6859, 08 2009. doi: 10.1021/ac9011705. 1

[83] Michael Evan Rising, Jerawan Chudoung Armstrong, Simon R. Bolding, Forrest Brooks Brown, Jeffrey S. Bull, Timothy Patrick Burke, Alexander Rich Clark, David A. Dixon, Robert Arthur Forster, III, Jesse Frank Giron, Tristan Sumner Grieve, Henry Grady Hughes, III, Colin James Josey, Joel Aaron Kulesza, Roger Lee Martz, Austin P. McCartney, Gregg Walter McKinney, Scott William Mosher, Eric John Pearson, Clell Jeffrey Solomon, Jr., Sriram Swaminarayan, Jeremy Ed Sweezy, Stephen Christian Wilson, and Anthony J. Zukaitis. MCNP® Code Version 6.3.0 Release Notes. Technical Report LA-UR-22-33103, Rev. 1, Los Alamos National Laboratory, Los Alamos, NM, USA, January 2023. URL `https://www.osti.gov/biblio/1909545`. 4, 7.1

[84] P.J. Roache. *Fundamentals of verification and validation.* Hermosa, Albuquerque, 2009. ISBN 9780913478127. 4.2

[85] P. Romano, N. Horelik, B. Herman, A. Nelson, B. Forget, and K. Smith. OpenMC: A state-of-the-art Monte Carlo code for research and development. *Annals of Nuclear Energy*, 82:90–97, 2015. (document), 3.1.2, 4, 4.1, 4.3, 5.2, 8.5.4

[86] J. Ruge and K. Stüben. Algebraic multigrid. In S. McCormick, editor, *Multigrid methods, volume 3 of Frontiers in Applied Mathematics*, pages 73–130. SIAM, 1987. 4.1, 8.5.2

[87] R. Molinaro S. Mishra. Physics informed neural networks for simulating radiative transfer. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 270:107705, 2021. doi: https://doi.org/10.1016/j.jqsrt.2021.107705. 8

[88] Y. Saad and M. Schultz. GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems. *Siam Journal on Scientific and Statistical Computing*, 7: 856–869, 1986. 4.1, 8.5.2

[89] Stephen Seltzer. Tables of X-Ray Mass Attenuation Coefficients and Mass Energy-Absorption Coefficients, NIST Standard Reference Database 126, 1995. URL `http://www.nist.gov/pml/data/xraycoef/index.cfm`. 7.1

[90] Seubert, A. and Langenbuch, S. and Zwermann, W. Solution of the Stationary State of the PWR MOX/UO-2 Core Transient Benchmark. In *Proceedings PHYSOR 2006*, Vancouver, Canada, 2006. 6.4

[91] Y. Shin, J. Darbon, and G.E. Karniadakis. On the convergence of physics informed neural networks for linear second-order elliptic and parabolic type pdes. *Communications in Computational Physics*, 28(5):2042–2074, 2020. doi: https://doi.org/10.4208/cicp.OA-2020-0193. 8, 8.2

[92] I.M. Sobol. The distribution of points in a cube and the accurate evaluation of integrals. *Zh. Vychisl. Mat. i Mat. Phys.*, 7:784–802, 1967. 8.2, 8.4

[93] G. L. Squires. *Introduction to the Theory of Thermal Neutron Scattering.* Cambridge University Press, Cambridge, 3 edition, 2012. 2.4

[94] László Szentmiklósi, Zsolt Revay, Jozsef Östör, and Boglarka Maróti. Improved analytical workflow for prompt-gamma activation analysis. *Journal of Radioanalytical and Nuclear Chemistry*, 333:3325–3333, 2024. doi: 10.1007/s10967-023-09071-4. 1

[95] F. Tröltzsch. *Optimal control of partial differential equations: Theory, Methods and Applications.* Vieweg+Teubner Verlag, in German, Wiesbaden (Germany), 2009. 6.1

[96] René van Geemert. Use of gpt for stabilization and acceleration of search mechanisms in industrial core computations. *Annals of Nuclear Energy*, 136, 2020. ISSN 0306-4549. doi: 10.1016/j.anucene.2019.107013. 6.4

[97] E. Varin and G. Samba. Spherical harmonics finite element transport equation solution using a least-squares approach. *Nuclear Science and Engineering*, 151(2):167–183, 2005. doi: 10.13182/NSE05-A2538. 8.4.2, 8.5.3

[98] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020. doi: 10.1038/s41592-019-0686-2. 7.6.1

[99] W. Wieselquist, R. Lefebvre, and M. Jessee. SCALE Code System. Technical Report ORNL/TM-2005/39, Version 6.2.4, Oak Ridge National Laboratory, 2020. 4, 4.3, 5.2, 5.3.1

[100] Armand Wirgin. The inverse crime, 2004. URL `doi.org/10.48550/arXiv.math-ph/0401050`. 7.6

[101] Weixiong Zheng and Ryan G. McClarren. Accurate least-squares pn scaling based on problem optical thickness for solving neutron transport problems. *Progress in Nuclear Energy*, 101:394–400, 2017. ISSN 0149-1970. doi: https://doi.org/10.1016/j.pnucene.2017.06.001. 9

[102] Ciyou Zhu, Richard H. Byrd, Peihuang Lu, and Jorge Nocedal. Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization. *ACM Trans. Math. Softw.*, 23:550–560, 1997. 7.6.1

A. Appendix

A.1. Material Properties used in Sec. 7.6

Table A.1.: Material Properties

	σ_{t1}	σ_{a1}	σ_{t2}	σ_{a2}	c_1	c_2
Mixed Scattering/Absorption 1	1.00	0.01	0.50	0.25	1	2
Mixed Scattering/Absorption 2	1.00	0.05	1.00	0.25	1	1
Scattering Dominated	1.00	0.01	1.00	0.05	1	1
Absorption Dominated	0.50	0.30	0.50	0.25	2	2

A.2. Counting Rates used in Sec. 7.6

Table A.2.: Counting Rates

	$\overline{P_{E_{\gamma,1}}}$	$\overline{P_{E_{\gamma,2}}}$	$\overline{R_1}$	$\overline{R_2}$	$\overline{R_3}$
Mixed Scattering/Absorption 1	1857521.7	18575217.0	501285.4	10453010.7	17870.5
Mixed Scattering/Absorption 2	7229203.8	72292038.0	1073168.0	18575479.3	41900.1
Scattering Dominated	903042.7	9030427.0	355318.6	7915378.5	12529.2
Absorption Dominated	1619960.2	16199602.0	515428.9	10706125.8	18424.6