Sensitivity of Fusion Neutronics to the Pre-Processing of Nuclear Data

by

Tanya Hutton

A thesis submitted to
The University of Birmingham
for the degree of
Doctor of Engineering (EngD)

School of Physics and Astronomy
The University of Birmingham
April 2016

Abstract

Nuclear data are the foundation of simulation and design in the nuclear industry. The success of commercialising thermonuclear fusion will be based on a set of highly accurate simulations used in design, optimisation and safety analyses.

This work focuses on the often overlooked, pre-processing stage of nuclear data. The effect of legacy methods in a fusion context is a concern within the community, but has never been quantified. The sensitivity of fusion neutronics to pre-processing was determined using a set of codes and methods developed as part of this thesis.

Legacy pre-processing methods demonstrated a difference between the processed and unprocessed distributions of up to 20%. Simple Monte-Carlo radiation transport simulations exhibited sensitivity within energy distributions for small models (< 5 mfp). Alternative data formats did not improve simulation results sufficiently to justify their implementation. Complex, fusion specific models showed a general insensitivity to the pre-processing when run to the current levels of statistical precision.

Future recommendations are to process all future data libraries into the cumulative tabulated probability format. Improved methods are not required at this stage as the core data libraries are incomplete and sometimes inaccurate. Only after the libraries have improved will pre-processing become significant.

Acknowledgements

I thank my supervisors, official and unofficial, academic and industrial. This project started in a very different form, and thanks go to Dr. Paul Norman and Dr. Andrew Davis for providing me with this opportunity. It has evolved under the guidance and support from Dr. Thomas Leadbeater and Prof. Peter Jones at the University of Birmingham, and Dr. Jean-Christophe Sublet, Dr. Lee Morgan and Lee Packer at Culham Centre for Fusion Energy (CCFE).

I acknowledge the support received from CCFE, sponsoring this project and providing direction. Thanks also go to the Nuclear EngD Centre at the University of Manchester and the EPSRC.

The University of Birmingham has become a home to me in the last 8 years, and along the way I have met some fantastic people. Whether it's through physics, climbing or drinking, they have all contributed to who I am today, and by proxy, this thesis.

This thesis would not have been possible without the continued love and support from my family. They have provided me with every opportunity and I cannot thank them enough - this is for you. You can finally see, and maybe even understand a little of what I've been doing for the last four years of my life.

Finally, thanks to my cat. Loki has kept me somewhat sane throughout this process and has suffered through many hours of me reading this thesis at her. She might just be the only cat in the world with an understanding of nuclear data.

Either the well was very deep, or she fell very slowly, for she had plenty of time as she went down to look about her, and to wonder what was going to happen next.

– Lewis Carroll, *Alice's Adventures in Wonderland*

Contents

List of Figures

List of Publications

Published

Nuclear data for fusion: Validation of typical pre-processing methods for radiation transport calculations. With J-Ch. Sublet, L. Morgan and T.W. Leadbeater. Published in *Fusion Engineering and Design*, 100:81-86, 2015.

In press

Pre-processing of nuclear data: optimising the relationship between accurate data representation and computational efficiency. With J-Ch. Sublet, L. Morgan and T.W. Leadbeater. Accepted for publication in *Annals of Nuclear Energy*, 2016.

In preparation

Sensitivity of fusion neutronics to the pre-processing of nuclear data. With L. Morgan, J. Shimwell and J-Ch. Sublet. In preparation for ND2016.

Turnip: A tool for investigating the validity of nuclear data in new regimes. With T.W. Leadbeater. In preparation.

List of Common Acronyms

MCNP	Monte-Carlo N-Particle
ENDF	Evaluated Nuclear Data File
ACE	A Compact ENDF
ITER	International Thermonuclear Experimental Reactor
DEMO	DEMOnstration reactor
JET	Joint European Torus
TBR	Tritium Breeding Ratio
TPR	Tritium Production Rate
EXFOR	EXchange FORmat
MAST	MegaAmp Spherical Tokamak
CCFE	Culham Centre for Fusion Energy
SINBAD	Sheilding and Integral Benchmark Archive and Database
PENDF	Point-wise ENDF

Chapter 1

Introduction

Nuclear fusion is a solution to the current energy crisis. Not only does it promise clean and safe energy, but it has the potential to produce sufficient energy to fuel civilisation for many thousands of years to come [1]. The culmination of over 50 years of international collaboration has brought fusion research to the point of constructing ITER, the International Thermonuclear Experimental Reactor. It is expected to be the first fusion device to achieve a net power output and will form the basis of future commercial power stations.

The ability to harness fusion energy, in a manner which is commercially viable, is an ongoing challenge. The nuclear physics behind the fusion process is well understood, and the knowledge base of controlling the plasma is growing, but the largest contributors delaying the demonstration of a commercially viable fusion reactor are the engineering elements. The conditions within a fusion device of reactor scale are expected to be exceptionally harsh. High temperatures and extreme radiation damage make the behaviour of components unpredictable as a device of this scale has never been realised before. ITER is one of many fusion experiments around the world designed to validate and test models of the plasma, interactions with surrounding components, candidate materials and power extraction. Results from these experiments will inform the designs of the next-generation

fusion devices. High fidelity computational modelling of nuclear processes is at the core of all fusion experiment and reactor design. This in turn requires high quality nuclear data and the proper treatment thereof. Many years have been spent collating and refining data for fission applications, yet important fusion relevant data are still left deficient. This is of great concern to the fusion and nuclear data communities, and as such there is an ongoing effort to rectify the situation. References [2–8] show a small insight into the continued pursuit of complete and accurate data libraries for fusion neutronics.

This work focuses on the treatment of nuclear data prior to their use in radiation transport simulations for fusion relevant scenarios. The current state of nuclear data and the methods used to manipulate them are highly biased towards fission energies and materials. Fusion regimes are centred around higher energies and non-standard materials, so in many cases the data and their treatment are still lacking. The findings are not limited to fusion, and will have consequences throughout the nuclear and high-energy particle physics fields.

1.1 Nuclear fusion

Nuclear fusion is the process of joining light nuclei with the resulting release of thermal energy from the reaction. The energy released corresponds to the nuclear mass difference between the fusing, and fused nuclei, i.e. the difference between initial and final mass states. The fusing nuclei must have sufficient kinetic energy to allow them to overcome the repulsive, long-range Coulomb force, and allow the attractive, short-range strong nuclear force to take over allowing the nuclear reaction.

1.1.1 History of nuclear fusion

Nuclear fusion experiments began in the 1930s in an effort to understand the physical process. The first tokamak was built in 1950 by the Soviet Union, which opened up the

potential engineering solutions to realise fusion as a power source [9]. Culham Centre for Fusion Energy (CCFE) is the UK's national laboratory for fusion research and has been at the forefront of the field since its commissioning in 1965. CCFE has developed and built the MAST (MegaAmp Spherical Tokamak) experiment and hosts JET (Joint European Torus) on behalf of the European collaboration. JET has been in operation since 1983 and is currently the world's largest magnetic confinement fusion experiment. It has undergone many upgrades and is the testbed for the ITER device, which is currently under construction. Over 50 years have been spent attempting to achieve a net energy output from the tokamak in order to create a commercially viable nuclear fusion reactor. While break-even has not yet been achieved, ITER is expected to demonstrate a Q-value [1] of greater than 10. DEMO, the next step in the fusion power roadmap is expected to demonstrate a Q-value greater than 25. A higher Q-value is needed to improve the commercial aspect of fusion power. The larger the Q-value, the more cost-effective each unit of energy put into the device will be. Figure 1.1 shows how fusion capabilities have improved over the last 50 years, and some of the key experiments in the journey to commercialising nuclear fusion are shown in Table 1.1.

[1]The Q-value here is the factor gain in energy, i.e. the energy production minus any losses with respect to the total energy consumption. Break-even occurs when Q=1 and Q>1 results in net power output.

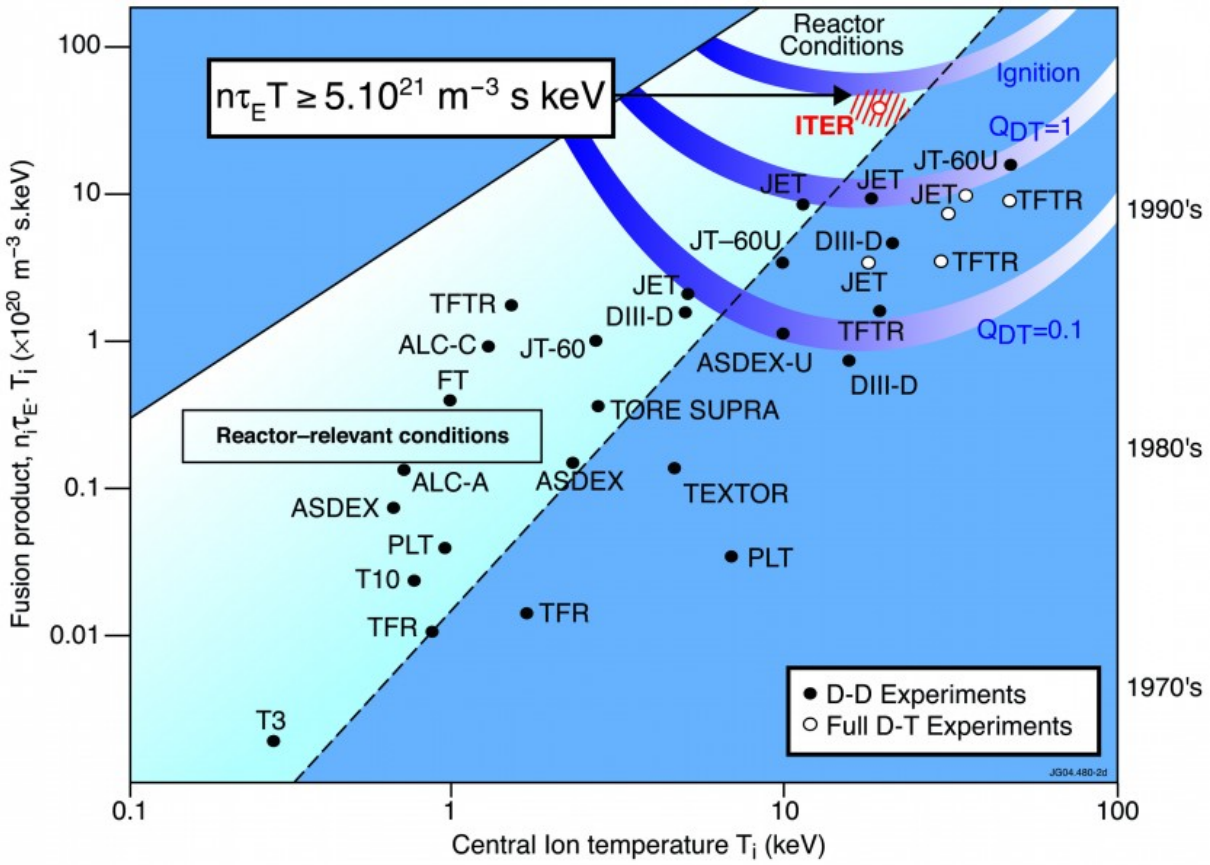

Figure 1.1: Progress of magnetic confinement fusion experiments as measured with the fusion triple product, which is a measure of the efficiency of confinement. Over time, fusion facilities have been able to burn plasma at increasingly high temperatures (T) and densities (n) for longer periods of time τ in an effort to achieve the reactor level conditions for commercial viability [10, p179].

Table 1.1: Details of key fusion experiments based on the tokamak design over the last 50 years [11, 12]

Name	**Date commissioned**	**Details**
T1	1950	First tokamak device, Russia
T3	1968	Highest temperatures and confinement times reached, Russia
TFR	1973	Plasma temperatures of 2 keV, France
JET	1983	Joint European Torus, produced 1.7 MW of power in 1991, 16 MW in 1997
TFTR	1983	Tokamak Fusion Test Reactor produced 10 MW of power in 1993, USA
Tore Supre	1988	Longest plasma duration time of over 6 minutes in 1996, France
JT-60	1985	Highest value of fusion triple product, Japan
ITER	$\sim$2025	Net power output, Q = 10
DEMO	$\sim$2050	First fusion device to put power on to grid, Q > 25

1.1.2 Confinement

There are three broad types of confinement for nuclear fusion, gravitational, inertial and magnetic [13, p49]. Nuclear fusion is the process which powers the stars, the gravitational pressure associated with the high mass produces the necessary conditions for fusion to occur. Inertial confinement involves the compression of a solid, supercooled fuel pellet, typically with laser beams. The evaporation of fuel from the surface of the pellet increases the density and temperature within the core sufficiently to induce the fusion reaction. Magnetic confinement is the basis of ITER, JET and many other fusion experiments, and is the most developed technique with respect to commercially viable fusion reactors at this point in time [14]. Specifically these experiments are based on the tokamak design; where a highly charged plasma is contained within a toroidal magnetic field to prevent the direct heat loading to the vessel walls and loss of plasma temperature. Figure 1.2 shows the schematic of ITER. A set of D-shaped magnets provide the toroidal field lines; the poloidal field is generated via the plasma current and works in conjunction with the outer magnets to maintain the shape and stability of the plasma. The tokamak is relatively simple to build given its symmetry, but it is difficult to control the plasma behaviour for long periods of time. An alternative to the tokamak is the stellarator. It is a toroidal device, but with helical magnetic fields. These fields are more stable for confining a plasma, but the largest operating example of the stellarator is the Wendelstein 7-X [15]. This design is much more complex to construct and is still in the early stages of implementation, so it is not a commercially viable option at present.

Magnetic fields are able to control the behaviour of the highly charged plasma, but the uncharged photons and neutrons produced are free to escape. After leaving the plasma the photons and neutrons interact with the surrounding materials and deposit their energy and/or escape. The energy with which these leave the plasma is dependent on the choice of fuel.

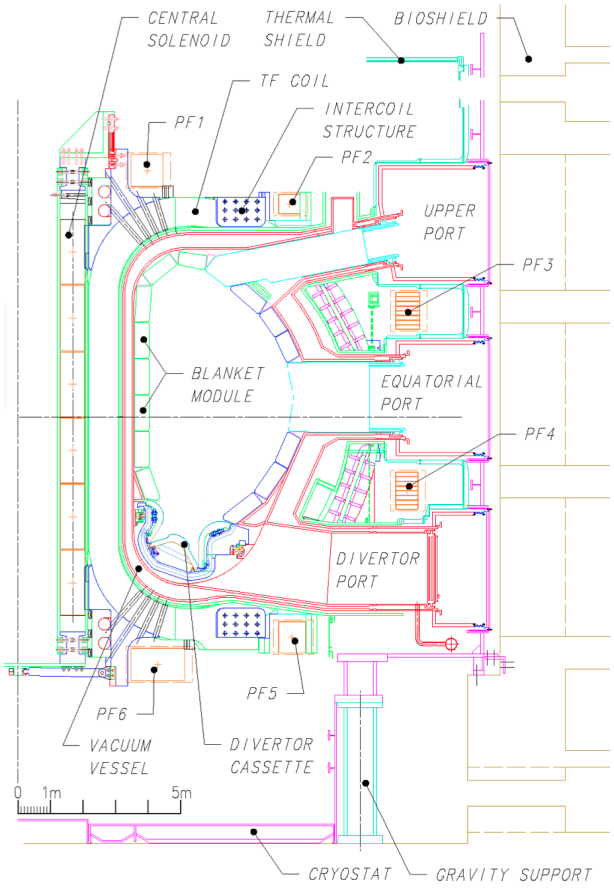

Figure 1.2: Cross-section through the ITER tokamak design. Plasma is contained within the D-shaped vessel with a combination of poloidal field (central solenoid and PF1-6) and toroidal field (TF) coils [16, p10].

1.1.3 Fuels

The fuel selection is based on a balance between interaction Q-values, fusion temperatures, abundances, and confinement parameters required to achieve net power out. The interaction Q-value is determined by the nuclear mass difference between the fusing and fused nuclei. Figure 1.3 shows the binding energy per nucleon with respect to atomic mass. In general, isotopes with an atomic mass below ^{56}Fe will produce energy when fused due to the mass difference.

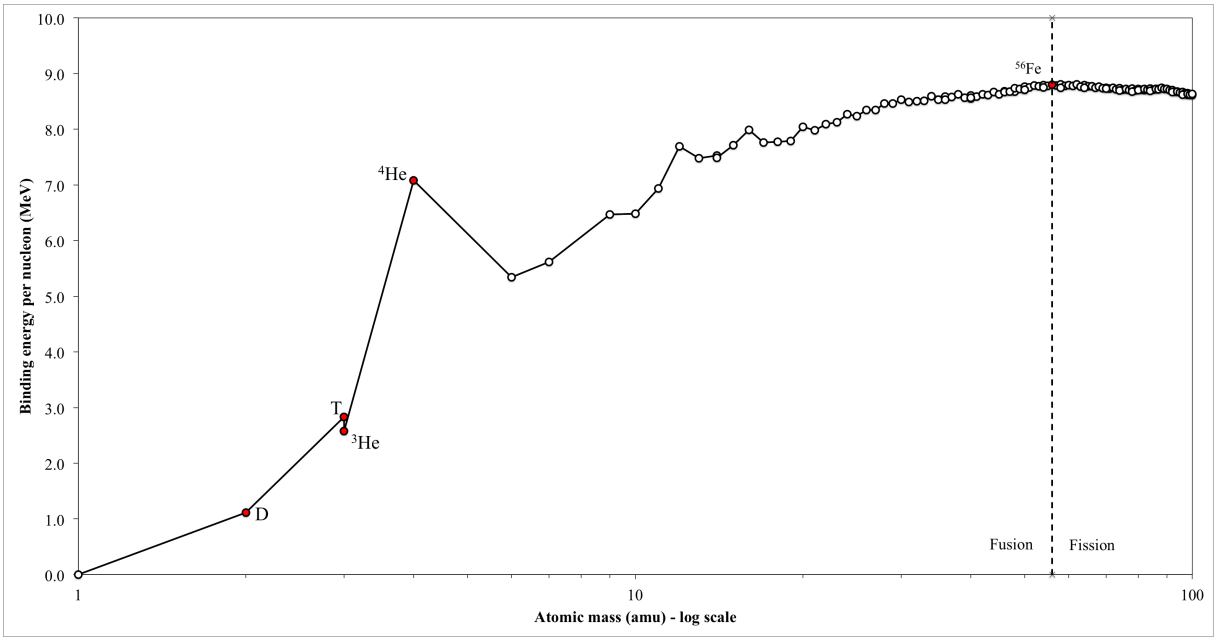

Figure 1.3: Average binding energy per nucleon with respect to atomic mass. Low mass nuclei of interest to fusion are highlighted in red. Up to ^{56}Fe, energy can be produced from fusing nuclei. Past this point, energy can be produced from fission.

Higher atomic masses require a higher temperature due to the larger number of protons within the nucleus, hence increasing the Coulomb repulsion factor. Low charge, and typically low mass, nuclei require a lower temperature to achieve fusion. The isotopes of hydrogen are the most practical options in terms of producing a fusion reactor. The two main candidates are deuterium-deuterium (D-D) or deuterium-tritium (D-T). Their reactions are given below:

$$^{2}H +^{2} H \rightarrow^{3} He + n \quad (Q = 3.3\ MeV)$$

$$^{2}H +^{2} H \rightarrow^{3} H + p \quad (Q = 4.0\ MeV)$$

$$^{2}H +^{3} H \rightarrow^{4} He + n \quad (Q = 17.6\ MeV)$$

D-D fusion has two pathways which are equally probable, whereas D-T fusion has only one primary pathway, though in both cases, there will be contributions from all reaction

channels. The high Q-value from the D-T reaction is due to the formation of the highly stable ^{4}He nucleus. The binding energy per nucleon is 7.1 MeV, compared to the 2.6 MeV in ^{3}He or 2.8 MeV in ^{3}H from the D-D reactions. The neutron producing reaction channels result in neutron energies of 2.5 MeV and 14.1 MeV respectively.

To determine which fuel is the best choice within a commercially viable fusion reactor, the Lawson criterion, as given in Equation 1.1.1 is used.

$$n\tau > \frac{12kT}{\langle\sigma v\rangle Q} \tag{1.1.1}$$

This equation balances plasma density, temperature and the confinement time required for break-even. The plasma density (n) multiplied by the confinement time (τ) must be greater than the given function of temperature (T) and average reaction rate $\langle\sigma v\rangle$. Where σ is the microscopic fusion cross-section and v is the average particle velocity [13, p131]. The temperature must be high enough to overcome the Coulomb barrier, minus any effects from quantum tunnelling, the ion density must be high enough that sufficient fusion reactions occur and the confinement time is a key factor in producing net power out.

Figure 1.4 shows the fusion cross-sections of different fuels as a function of the centre-of-mass energy. The cross-section for D-T fusion is consistently higher than that of D-D fusion below 1 MeV, resulting in a higher probability of fusion occurring at a lower energy, and temperature, with D-T fuel when compared to D-D fuel.

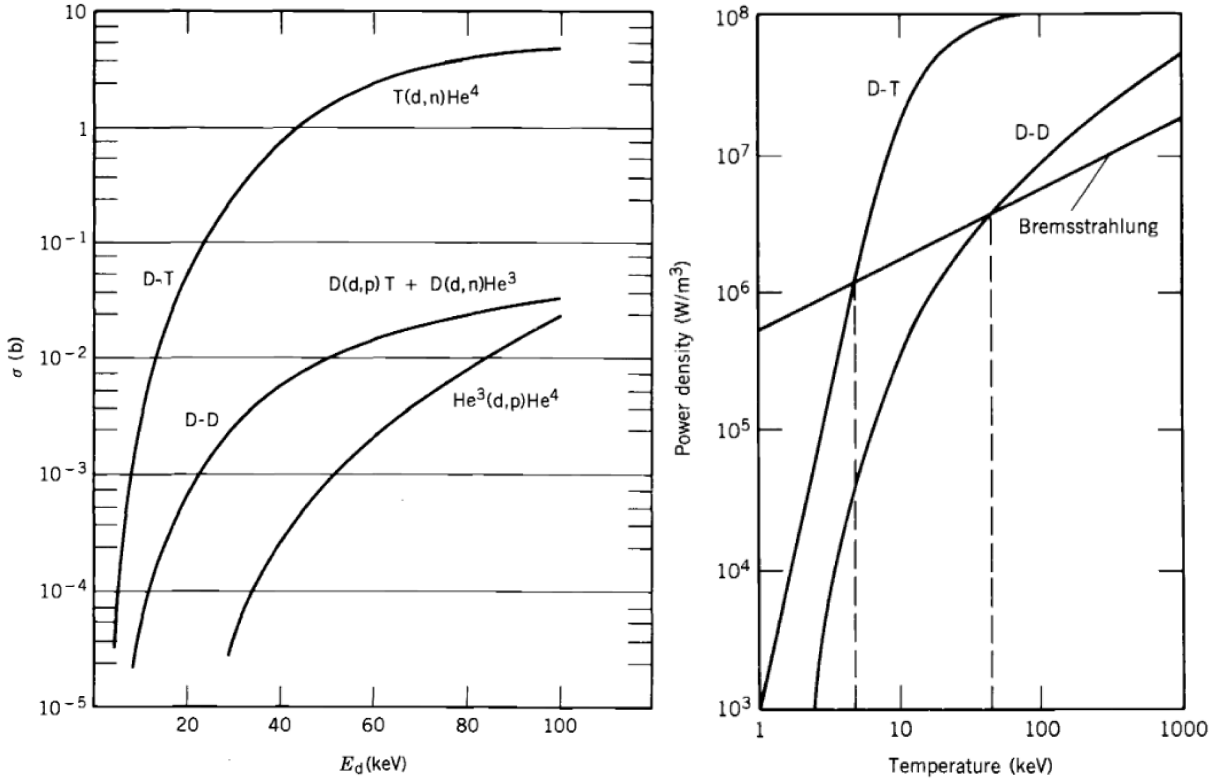

Figure 1.4: Left figure shows the interaction cross-sections for D-T and D-D fuels as a function of deuteron energy. Right figure shows the Bremsstrahlung losses in comparison to the power output of D-T and D-D fusion reactions with plasma temperature. Dashed lines show the temperature at which the fusion power output surpasses the Bremsstrahlung losses. [17, p533, p541] These plots do not show the losses associated with the neutrons escaping the plasma.

Equation 1.1.1 does not include radiative or conduction losses, i.e. Bremsstrahlung or neutrons leaving the plasma. It assumes that all energy produced from the fusion reaction is deposited within the plasma. This is not the case, so for D-T plasma a minimum temperature of 10 keV is required to overcome these losses. For D-D plasma this minimum temperature is around 100 keV. At these temperatures, the $n\tau$ Lawson criterion are $10^{20}\ sm^{-3}$ (D-T) and $10^{22}\ sm^{-3}$ (D-D), i.e. the density and confinement time requirements are two orders of magnitude lower for D-T fusion, in addition to requiring a lower overall plasma temperature. For these reasons, D-T fuel has been selected for use in ITER.

1.1.4 Tritium breeding

The main downfall of including tritium is the availability. It has a relatively short half-life of 12.3 years and the latest available inventory, as of 2011, was around 20 kg. ITER is expected to consume 0.8-1.2 kg per year of operation [18]. The predicted tritium inventory is not sufficient to fuel ITER and future fusion experiments so it must be bred within the device. Using the neutrons from the fusion reaction, tritium can be bred from lithium via the two reaction channels given below:

$$^{6}Li + n \rightarrow ^{3}H + ^{4}He \qquad (Q = 4.8\ MeV)$$
$$^{7}Li + n \rightarrow ^{3}H + ^{4}He + n^{*} \qquad (Q = -2.5\ MeV)$$

The ^{6}Li reaction is energetically favoured, with the capture of a thermal neutron, whereas the ^{7}Li reaction channel is a threshold reaction and the incoming neutron must have an energy above 2.5 MeV to initiate the reaction. The ^{6}Li tritium breeding yield is far higher than that from ^{7}Li due to the differences within the cross-sections, as seen in Figure 1.5. For self-sufficiency, ITER and future devices must breed enough tritium to maintain stable levels of fuel including any losses due to extraction efficiency and radioactive decay. The ability of a fusion reactor to be self sustaining is defined by the tritium breeding ratio (TBR) and is the rate of production over the rate of burning in the plasma [21]. The lithium is built into breeder modules which surround the plasma, as can be seen for ITER in Figure 1.2. ITER is an experimental reactor so the predicted TBR is low, whereas DEMO must satisfy a net TBR greater than 1. This is not possible with the one-to-one conversion of neutrons to tritium by the ^{6}Li reaction channel. Additional neutrons must be created via neutron multiplication reactions in ^{9}Be or ^{208}Pb depending on the blanket design. These types of engineering details will be tested within ITER.

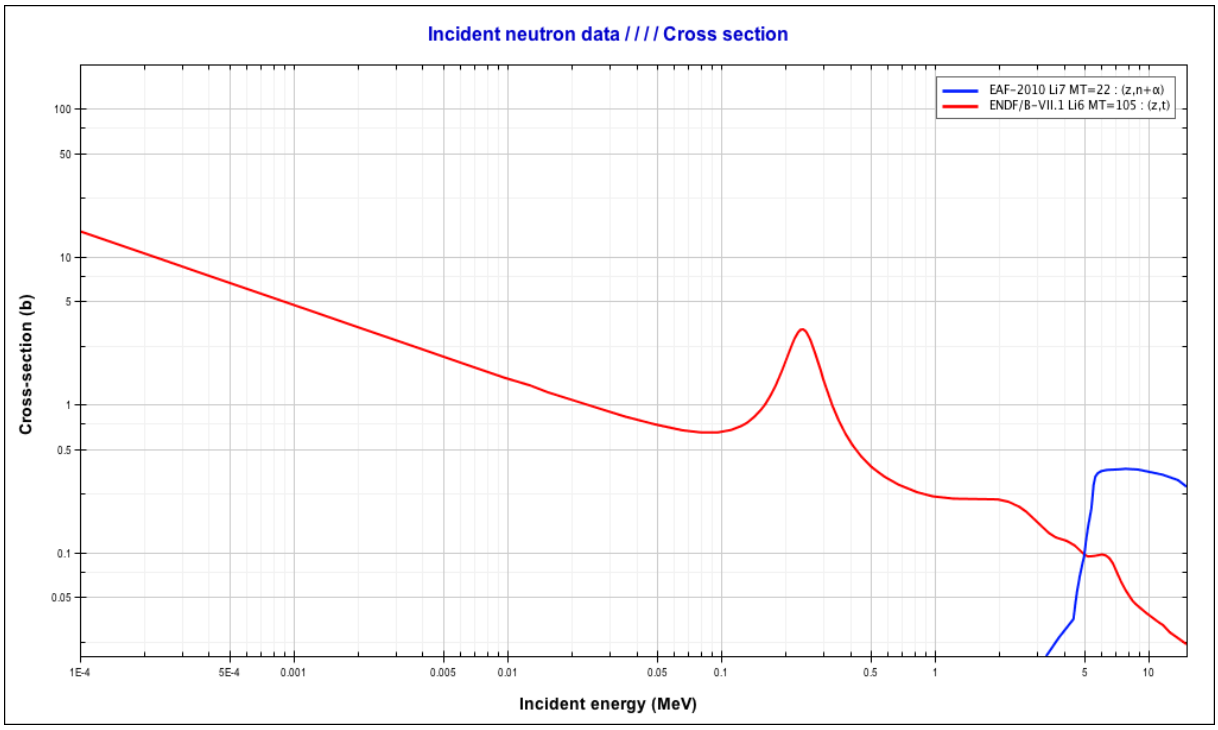

Figure 1.5: Cross-section data for tritium breeding candidates ^{6}Li and ^{7}Li with incident neutron energy. The probability of the ^{6}Li(n,t) reaction (red) behaves approximately as $1/\sqrt{E}$ or $1/v$, so a lower energy of incident neutron is preferred to maximise the tritium production. The ^{7}Li(n,nα) reaction (blue) is only energetically available when the neutron is above the threshold. Above 5.0 MeV this reaction is more productive than the ^{6}Li channel. Data taken from the ENDF/B-VII.1 [19] and EAF-2010 [20] libraries.

Neutrons are critical to a commercially viable fusion reactor. Not only are they crucial to tritium self-sufficiency, but neutrons are the main mechanism for heat extraction as they are able to escape the plasma. These neutrons determine the shielding requirements and the degree of radiation damage to vacuum vessel components. The neutron yields expected from the D-T plasma in ITER range from $10^{14} ns^{-1}$ up to $10^{21} ns^{-1}$ for the various modes of operation [22]. These neutron yields are far higher than have been seen before in an engineering environment, so the true extent of radiation damage and activation of components is unknown. This is another area in which ITER will be used to investigate the lifetime of components within the reactor, and the resulting activation.

1.2 Radiation transport for fusion

Radiation transport simulations are regularly performed to aid in the design of nuclear facilities. Within the realms of fusion reactors, its application is multifaceted. Simulations are used to determine the heat loading on the various components and their respective lifetimes. It is used in shielding design, particularly around the diagnostic components, and streaming. Other considerations are shutdown doses from photons due to neutron activation of the vessel, tritium breeding ratios and waste management. The radiation transport code of specific interest is MCNP (Monte-Carlo N-Particle) [23]. This is the industry standard Monte-Carlo radiation transport code, and has a long history within fission applications. It has more recently been applied to general-purpose scenarios, including fusion. Figure 1.6 shows the MCNP geometry for a 40° segment of the ITER device, which is used by CCFE and the wider fusion community for the various analyses mentioned above.

These radiation transport codes utilise nuclear data to determine the behaviour of radiation quanta within the system. These data undergo a series of processes in order for them to be applied within the radiation transport codes. For the simulation results to be an accurate depiction of the physical system, the data must replicate the physical interactions of the radiation with matter. One particular stage within the data cycle is the focus of this work; pre-processing. This is the preparation of general-purpose data files for use in specific applications such as MCNP. This stage was designed with fission in mind, so applying the same methods and techniques to fusion is not well validated and is often not considered.

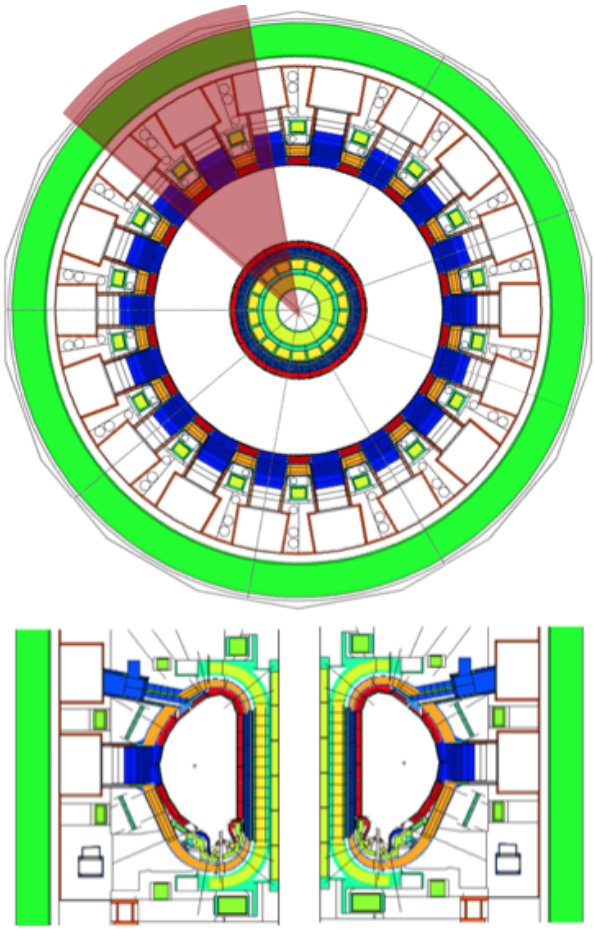

Figure 1.6: MCNP model of the ITER tokamak, the symmetry of the design allows for a 40° segment (highlighted in red) to be representative of the whole tokamak, reducing computation. Model dimensions are 20 m radially and 25 m axially, consisting of 8500 separate cells, 212 materials and 63 isotopes.[24]

1.3 Key objectives and commercial context

The priority was to rigorously test the current processed data formats and their applicability to fusion systems. This was done by first highlighting that a problem exists with applying legacy methods to non-legacy applications. From here several computational tools were developed to determine the effect of pre-processing on the simplest of systems up to engineering relevant control cases. In addition to the currently available pre-processed data formats, these tools allowed the investigation into alternative data formats, which would not otherwise be accepted by standard radiation transport codes. The

final stage was to determine the sensitivity of real-world simulations, which are currently used within the fusion community, to the pre-processing of nuclear data.

Commercial viability of a thermonuclear fusion reactor is dependent on the degree to which the uncertainty in the models and codes can be predicted, as these inform upon the final engineering tolerances. Benchmarking, validation and uncertainty quantification are key factors in determining the accuracy of codes and data. References [25–34] represent a small fraction of the work carried out in these areas (in fusion neutronics and elsewhere). This thesis has provided a framework for trialling and validating many aspects of nuclear data. While in this instance it has specifically been applied to the pre-processing of angular distributions for fusion neutronics, it has the potential to be utilised within many other areas of nuclear data.

Chapter 2

Theory

This Chapter covers each aspect involved in the transport of neutrons within a Monte-Carlo radiation transport code, along with the methods of quantifying differences within nuclear data. The way in which neutrons interact with matter is parameterised by nuclear data. These data form the basis of any radiation transport code, and the manner in which it is treated can impact upon the final simulation result.

2.1 Neutron interactions

The interactions of uncharged particles with matter are not affected by the Coulomb barrier and they are able to directly interact with a nucleus of any mass and at any energy [35, p. 116]. The change in energy and direction of the interacting particle are almost entirely stochastic in nature; this means that the path of any individual neutron (or photon) is unique and cannot be predicted exactly. In principle it is known how likely a neutron is to interact with a certain isotope, and how likely a certain interaction type is to occur.

The interactions of neutrons within a fusion scenario have far reaching consequences, including first wall neutron economy, heat extraction, tritium breeding and fuel efficiency [36, p. 19]. Table 2.1 shows the different interactions that are of importance to fusion neu-

tronics and the candidate materials associated with each component. This work focuses specifically on the transport of neutrons in fusion regimes, hence the following discussion is limited to the interaction of fast neutrons with materials of importance to fusion systems.

Table 2.1: Important neutron induced reactions within the area of fusion neutronics. For each area of interest, the relevant reaction channels and candidate materials for that purpose are given. The reaction channel shown as (n,Xn) refers to all neutron multiplication reactions and (n,abs) refers to the total of all absorption reactions. Adapted from [36, pp. 18-19]

Purpose	**Reactions**	**Candidate materials**
First wall neutron economy	All	Fe, Cr, Ni, Ti, V, Mo, Nb, W
Neutron multiplication	(n,Xn)	Be, Pb
Neutron moderation	(n,el), (n,n')	Be, B, C, O
Tritium breeding	(n,Xt), (n,n'α)	Li-compounds
Shielding	(n,el), (n,n'), (n,Xn), (n,abs)	B, C, O, Si, Ca, Fe, Ba
Transmutations	(n,p), (n,α), (n,d), (n,t), (n,^{3}He)	All
Neutron dosimetry	(n,γ), (n,p), (n,α), (n,n'), (n,Xn)	All
Radioactivity estimates	(n,p), (n,pn), (n,np), (n,α), (n,nα), (n,αn), (n,d), (n,t), (n,^{3}He)	All

There are two broad categories for the interaction of neutrons with matter, scattering events and absorption events. Scattering events produce an exit neutron (or retain the original with updated properties), which then continues to be transported within the system. Absorption events remove the original neutron from the system, but have a tendency to produce additional particles in the process, such as low mass nuclei, neutrons and photons. The primary neutron energy from the D-T plasma is 14.1 MeV, many

more interaction pathways are accessible at this energy when compared to fission type energies around 2.0 MeV. At 14.1 MeV the interacting neutrons are above the threshold for certain reaction types, such as inelastic scattering, charged particle production and neutron multiplication.

2.1.1 Scattering reactions

The elastic and inelastic scattering of neutrons are of particular interest to fusion reactors and this thesis. The high-energy (14.1 MeV) neutrons emitted from the D-T plasma are the primary method of extracting energy from the fusion reaction via energy deposition. The emitted neutrons interact with the surrounding vessel and, at the energy of interest, scattering is the preferred method of interaction despite the competing reaction channels. They can undergo many sequential scattering events and are moderated within the vessel and surroundings, depositing energy as heat. This heat must be deposited within appropriate regions for it to be extracted. Certain components, such as the supercooled magnets and diagnostics must be well shielded against the heating effects of the plasma and associated neutrons. Neutron scattering is important to all shielding requirements. The uncharged nature of the neutrons mean that they have long mean free paths and the level of shielding required is largely dependent on scattering and multiplication reactions. Large numbers of moderated neutrons are needed to induce the necessary amount of tritium breeding reactions for fuel self-sufficiency. Elastic and inelastic reactions are the primary energy loss mechanisms for neutrons to reach the required energy. The tritium-breeding capabilities of a fusion reactor are of vital importance to its commercial viability. If the reactor over-produces tritium, storage and proliferation become an issue. If the reactor under produces tritium, it is not possible to operate long term. The expected tritium requirement for ITER is below 0.1 kg per day and DEMO is around 0.5 kg per day operating at full power [37]. This far outweighs the current tritium inventory, so it

will be bred within the reactor itself. The understanding and control of neutron energy deposition within a reactor is fundamental to fusion engineering.

Elastic scattering, as defined by pure nuclear physics, refers to the quantum mechanical definition of a neutron scattering from a fixed nucleus and is governed by the wave-like properties of the neutron [38, p. 2]. In applied nuclear physics, elastic scattering is defined in a different way. The total kinetic energy of the system is conserved, but the nucleus is treated as a point that is able to recoil (and can exhibit thermal vibrations) post collision. Inelastic scattering in radiation transport theory is the absorption of a neutron by an interacting nucleus to form a compound nucleus. The nucleus then ejects a neutron minus the interaction threshold energy in the centre-of-mass frame. At high energies, multiple neutrons may be emitted after the interaction. In the case of radiation transport these are also considered to be inelastic scattering interactions [39, p. 7]. Each discrete level of excitation within the compound nucleus is considered as a separate interaction type.

In applied nuclear physics, the emission angle (and hence neutron energy after elastic scatter) is determined from probability distributions. These distributions correspond to the polar scattering angle, θ, but as these are three-dimensional problems the azimuthal angle, ω, needs to be included. This is treated as a canonical distribution between 0 and 2π radians, i.e. isotropic in ω. The three-dimensional scattering transform is shown in Equations 2.1.1 to 2.1.3 [39, p. 341]. Translating the incoming neutron vector $\vec{\Omega}$ into $\vec{\Omega}'$ through polar angle θ and azimuthal angle ω can be seen schematically in Figure 2.1. Typically the polar scattering component is represented by the scattering cosine, μ, where $\mu = \cos\theta$.

$$\Omega'_x = \frac{\sin\theta}{\sqrt{1-\Omega_z^2}}[\Omega_y \sin\omega - \Omega_y\Omega_x \cos\omega] + \Omega_x \cos\theta \tag{2.1.1}$$

$$\Omega'_y = \frac{\sin\theta}{\sqrt{1-\Omega_z^2}}[-\Omega_x \sin\omega - \Omega_z\Omega_y \cos\omega] + \Omega_y \cos\theta \tag{2.1.2}$$

$$\Omega'_z = \sin\theta\sqrt{1-\Omega_z^2}\cos\omega + \Omega_z \cos\theta \tag{2.1.3}$$

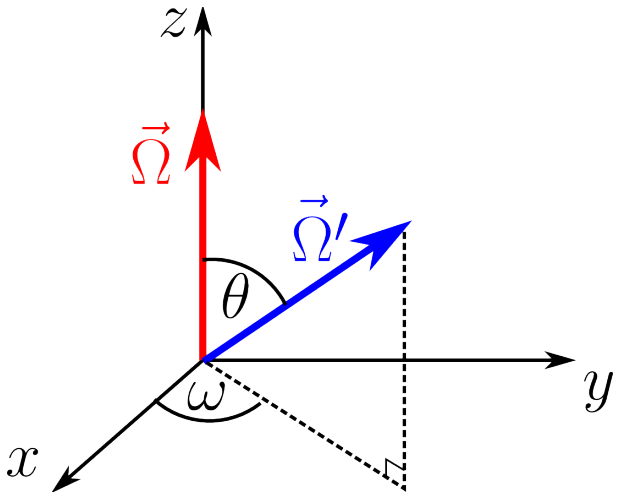

Figure 2.1: A particle with initial direction $\vec{\Omega}$ parallel to the z-axis, is transformed into the new direction $\vec{\Omega}'$ after a scattering event has occurred at the origin. It is transformed through polar angle θ and azimuthal angle ω. Note that this is the only case where Equations 2.1.1 to 2.1.3 do not hold. In this instance the transformation is performed with respect to the y-axis within radiation transport codes.

Elastic scatter

When a neutron interacts with matter via elastic scattering, the neutron collides with a nucleus and subsequently undergoes a change in energy and momentum; the lost energy is transferred to the nucleus as it recoils. The total energy of the two-body system is conserved, if the nucleus were fixed there would be no change in the neutron energy and an isotropic scattering distribution would result [38, p. 4].

The kinematics of the elastic scatter of a neutron from a nucleus varies based on the energy of the interacting neutron. At low energies (4 eV and below [23, p. 2-54]) the thermal

motion of the nucleus is significant with respect to the incoming neutron energy and must be accounted for [39, p. 342]. At higher energies the thermal motion is considered negligible and the interacting nucleus is considered at-rest in the lab frame. The energy loss of the neutron is related to the exit angle in the centre-of-mass frame. Low energy and low mass scattering tends to be relatively isotropic, and becomes more anisotropic with the increase of energy and/or mass. Figure 2.2 shows the schematics of the reaction in both lab and centre-of-mass frames.

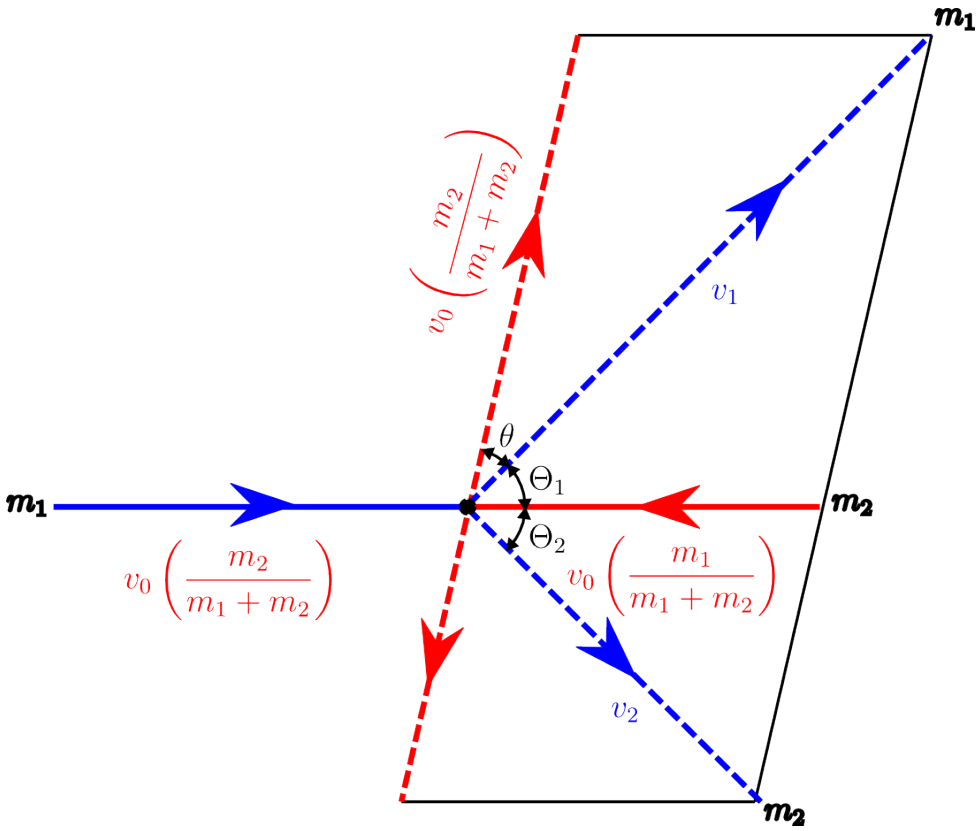

Figure 2.2: Schematic diagram of the elastic scattering kinematics (adapted from [35, p. 427]) in the lab (blue) and centre-of-mass (red) frame. The solid lines represent the pre-collision particle paths, and the dashed lines represent the post-collision paths. Note that the pre-collision path of the incident particle in the centre-of-mass frame is obscured by the pre-collision path in the lab frame. The masses of the incident (scattered) and target (recoil) particles are given by m_1 and m_2. The lab frame velocities of the incident particle are v_0 and v_1 pre- and post-collision, and v_2 is the recoil velocity post-collision. Scatter and recoil angles are Θ_1 and Θ_2 in the lab frame, and the scattered angle in the centre-of-mass frame is given by $\theta = (\frac{\pi}{2} - \Theta)$.

Based on energy and momentum conservation laws, the exit energy of the neutron (E_{out}) is given by Equation 2.1.4, where A is the mass (in amu) of the target nucleus, E_{in} is the incident neutron energy and θ_{cm} is the scattering angle in the centre-of-mass frame [40, p. 66].

$$E_{out} = E_{in}\left[\frac{1 + A^2 + 2A\cos\theta_{cm}}{(1+A)^2}\right] \tag{2.1.4}$$

The outgoing angle of the neutron in the centre-of-mass frame is governed by a set of energy dependent probability distributions. The maximum energy loss occurs with a π radian scattering angle (complete backscatter), and a minimum when the angle is zero (forward scatter). The equation for maximum energy loss simplifies to Equation 2.1.5.

$$E_{out} = E_{in}\frac{(A-1)^2}{(A+1)^2} = \alpha E_{in} \tag{2.1.5}$$

Lower mass nuclei are better at moderating neutrons as a higher proportion of energy is deposited per interaction when compared to higher mass nuclei. For example, a 1.0 MeV neutron elastically scattering from a ^{12}C nucleus has a maximum energy loss of 0.3 MeV at $\theta_{cm} = \pi$ radians, resulting in a 0.7 MeV neutron; whereas the same original neutron scattering from ^{56}Fe has a maximum energy loss of 0.1 MeV. Figure 2.3 shows the fractional energy loss with mass of the interacting nucleus.

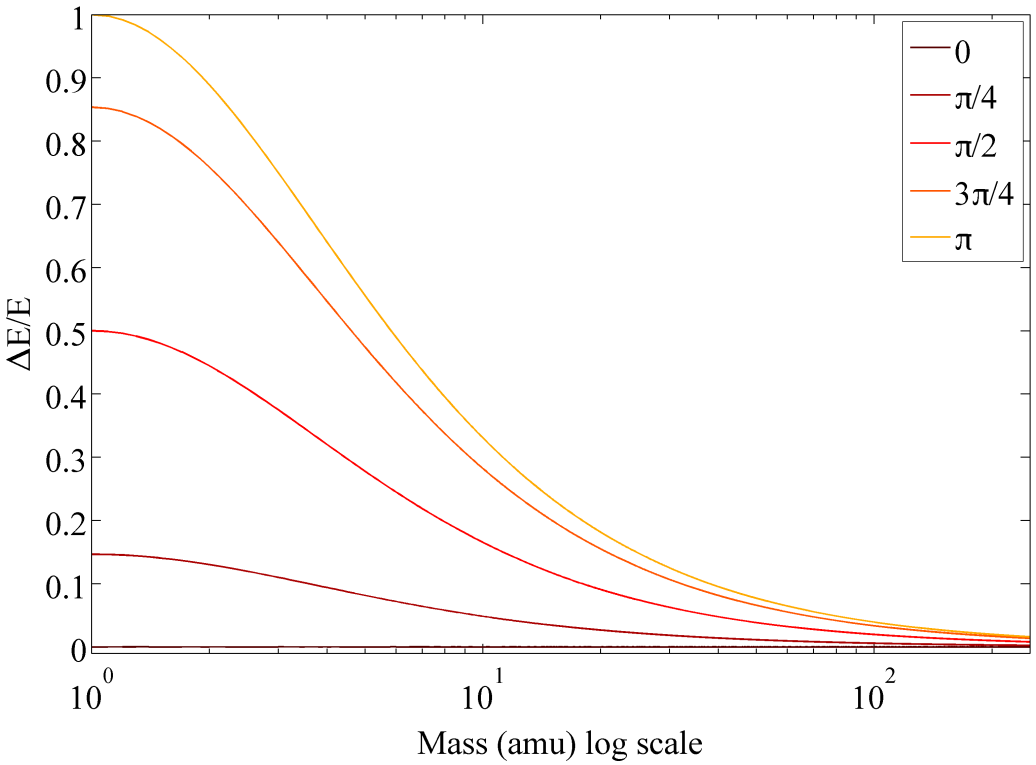

Figure 2.3: Fractional energy loss $\frac{\Delta E}{E}$ with respect to atomic mass on a log scale. Each line corresponds to the energy loss at scattering angles $0 \leqslant \theta \leqslant \pi$ radians. For ^{1}H, complete energy loss of the neutron occurs at $\theta = \pi$ radians. The neutron will retain some energy with all other nuclides at complete back scatter. For all masses, a zero scattering angle corresponds to zero energy loss.

Inelastic scatter

Inelastic scatter is a threshold reaction, where the incident neutron is captured by the target nucleus. This forms a compound nucleus in an unstable, excited state, which subsequently decays via the emission of at least one neutron and high-energy photon [41, p. 19]. These interactions only occur when the neutron is above the reaction Q-value and the energy loss of the neutron is directly related. For the case of the first inelastic level, the Q-value is equal to the first excited state within the nucleus. In general the neutron imparts sufficient energy to the nucleus that the constituent nucleons are raised in energy to the n^{th} excitation and de-excitation follows via particle and photon emission. Where more than one neutron is emitted from the interaction, i.e. neutron multiplication, the

Q-value is much higher, and hence are only available to high-energy neutrons.

The emission energy is independent of the angle, but the exit angle for each emitted neutron is governed by probability distributions, as with elastic scatters. For the case of non-multiplying inelastic scatters the lab-frame energy, post collision E_{out} is calculated with Equation 2.1.6. Where A is the atomic mass of the target nucleus, E_{in} is the incident neutron energy and Q is the energy threshold for the reaction [40, p. 67].

$$E_{out} = \left(\frac{A}{A+1}\right)^2 \left(E_{in} - \frac{A+1}{A}Q\right) \tag{2.1.6}$$

For nuclides with a high atomic mass, inelastic scattering is the most efficient reaction for energy loss. For example, a 14.1 MeV neutron incident on ^{56}Fe, inelastically scattering to the first level can reduce the neutron energy to 12.8 MeV (91% of the original), whereas a maximum energy loss via elastic scatter results in neutron energy of 13.1 MeV (93%). Figure 2.4 shows the fractional energy loss for multiple, identical and consecutive scatters in ^{56}Fe.

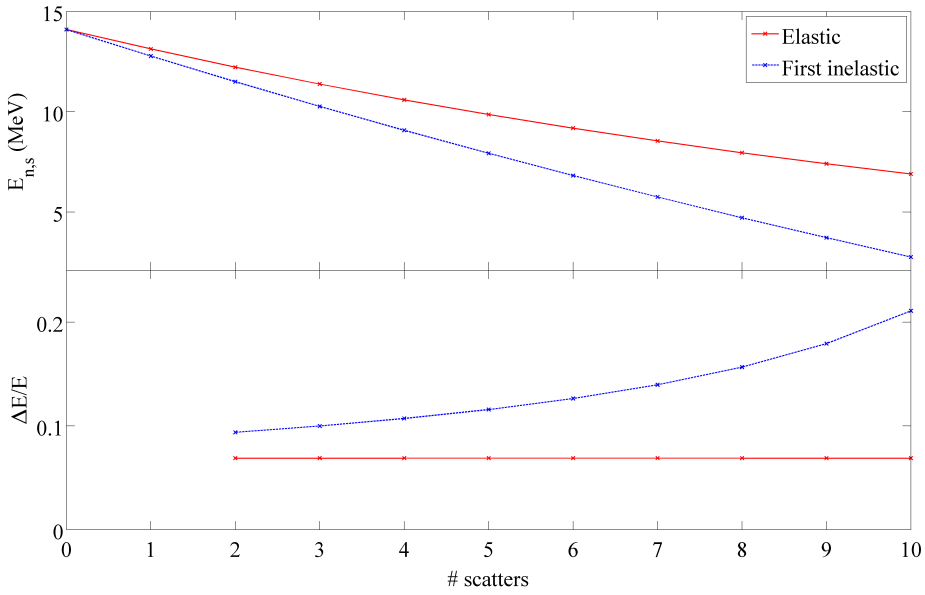

Figure 2.4: Comparison of energy loss via π radian elastic (red) and first level inelastic (blue) scattering for consecutive events in ^{56}Fe. Initial neutron energy is 14.1 MeV. Top figure shows the resulting neutron energy and the bottom shows the fractional energy loss per scatter for 0 to 10 scatters. Assuming complete back scatter for elastic events (maximum energy loss) a fixed 7% is lost per scatter. First level inelastic scatter results in a fixed energy loss of 847 keV per event. In this case, the inelastic scattering is more efficient at reducing the neutron energy.

2.1.2 Absorption reactions

An absorption reaction involves the loss of the original neutron to the interacting nucleus which is removed from the system. The two main mechanisms of neutron absorption are radiative capture and charged particle emission. For radiative capture, the incoming neutron is absorbed by the target nucleus. The now excited nucleus (plus one neutron) decays to the ground state via gamma emission (and possibly undergoes β^- emission). Charged particle emission reactions occur when the compound nucleus is sufficiently excited that it ejects a secondary particle. These secondary particles are typically low mass, such as hydrogen and helium nuclei. There are three main considerations for absorption or capture reactions within fusion scenarios; tritium breeding, transmutation and activation, and radiation damage.

The candidate breeder material for current reactor designs is lithium. Both ^{6}Li and ^{7}Li are capable of producing a triton after the capture of a neutron. The ^{6}Li reaction has an affinity for lower energy neutrons, and is more productive; whereas ^{7}Li produces the triton as part of a threshold reaction and removes energy from the system. (The reaction channels and cross-sections are given in Section 1.1.4, Figure 1.5). The ^{6}Li reaction channel absorbs the incoming neutron, removing it from the system. This one-for-one conversion would result in a tritium-breeding ratio (TBR) below 1 due to losses in the system [1]. DEMO is required to have a minimum tritium-breeding ratio of 1.1 [37]. Neutron multiplication (inelastic scatter) is necessary to achieve a satisfactory TBR. The accuracy with which tritium production is achieved and controlled can make the difference between a self-sustaining reactor and one that cannot maintain a stable level of fuel. Tritium self-sufficiency is an absolute requirement for a commercially viable fusion reactor.

When a neutron is captured to form a compound nucleus, the instability causes the nucleus to decay in an effort to form a stable state. For radiative capture de-excitation occurs via the emission of a high-energy photon, and secondary particle emission reactions can produce many different low-mass nuclei. Secondary particle emission changes the atomic number of the nuclide and hence it is transmuted. This new nuclide will have different properties from the original with respect to interaction and mechanical behaviours. This process of neutron absorption and consecutive decay contributes to the level of radiation within the system, particularly after shutdown. In a reactor scenario, the degree to which activation has occurred will determine the time required, after shutdown, before any repairs or maintenance can take place. Activation affects the shielding design of a reactor; a large volume of shielding will maintain safe radiation levels whilst the reactor is on-load, but will produce more secondary radiation as a result of activation than a smaller

[1] A tritium-breeding ratio below 1 means that the device is consuming more tritium than is being produced, a value greater than 1 means that more tritium is being produced than being consumed

volume. The design is carefully balanced between on-load and shutdown conditions. Figure 2.5 shows the main contributions to dose and activity from structural steel after 2.3 years of simulated irradiation in the DEMO fusion reactor [42].

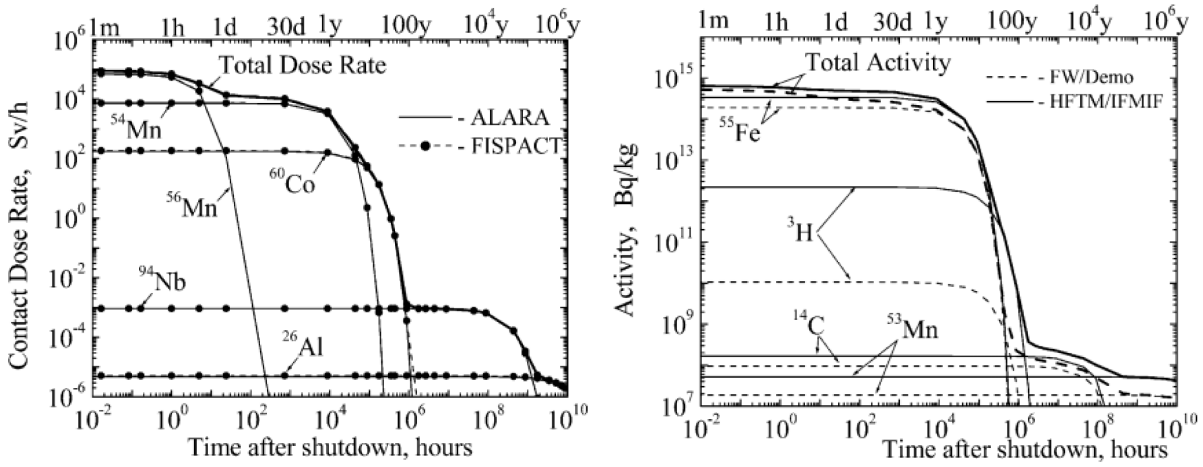

Figure 2.5: Simulated dose (left) and activity (right) profiles of the Eurofer steel after shutdown. Irradiation profiles are equivalent to those expected in DEMO over 2.3 years of operation. Immediately after shutdown the largest contributor to dose is ^{56}Mn, the largest contributor to activity is ^{55}Fe [42].

In the same way that neutron capture and subsequent decay of the compound nucleus causes activation, it also produces irreversible radiation damage. The absorption and decay results in a change in proton number, and hence is changed to a different element than the original. This transmutation affects the mechanical and chemical properties of the material, and when considering structural components, can be a big concern. One example of this is tungsten. It is transmuted into osmium and rhenium at high temperatures, changing a ductile material into a brittle one [43]. In addition to the transmutation of isotopes, secondary particle emission can cause gas production. When low mass nuclei, such as hydrogen and helium, are produced from a reaction, they tend to form neutral gas atoms and collect together along grain boundaries to form pockets of gas [44]. These can cause physical distortions to the components, such as swelling and cracking, again compromising the mechanical properties of the material.

All structural materials within ITER are expected to last the lifetime of the reactor, but

plasma facing components will have much shorter in-situ lifetimes due to the high neutron fluxes and will require periodic replacement. The useful lifetime of a component takes into account how many damaging radiation events are allowed before the structural properties of the material are compromised. Lifetime estimates are based on computational simulations, but the behaviour of materials at the expected temperatures and neutron fluxes is relatively unknown. Estimates will improve with the construction of IFMIF (International Fusion Materials Testing Facility) [45]. This facility is designed to deliver a lifetime equivalent neutron flux to test components, within much shorter time periods. A better understanding of the materials behaviour will improve the overall design and estimates of component lifetimes within a fusion reactor.

2.1.3 Interaction cross-sections

Interaction cross-sections are the key quantities used within radiation transport to determine how particles behave. The microscopic cross-section, σ, can be considered as the effective cross-sectional area of the target nucleus as seen by the incident particle and has units of *barns*; where 1 $barn = 10^{-24} cm^2$. This is the probability with which the particle will interact with the target nuclei and is shown schematically in Figure 2.6.

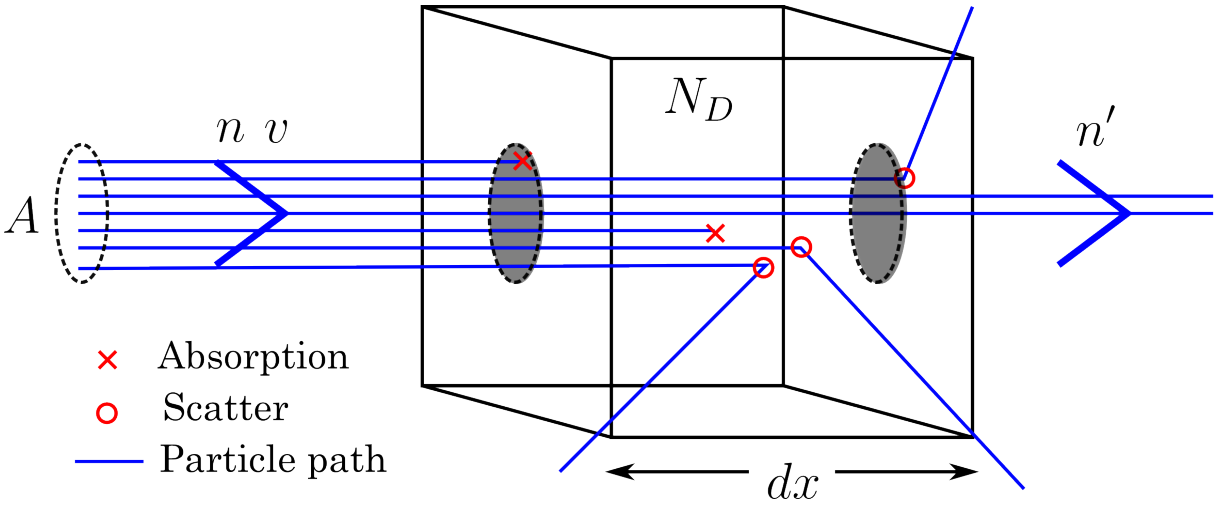

Figure 2.6: Schematic representation of neutron interaction cross-section, σ. A beam of neutrons, of density n ($neutrons\ cm^{-3}$), velocity v ($cm\ s^{-1}$) and area A (cm^2), is incident on a sample of thickness dx (cm) and number density N_D ($nuclei\ cm^{-3}$). The incident neutron flux ϕ is given by nv ($neutrons\ cm^{-2}\ s^{-1}$). As the neutrons interact within the target, they are removed from the beam, resulting in an unreacted beam density n'. The unreacted beam is related to the incoming beam density by $\frac{n'}{n} = \exp(-\sigma N_D x)$ (Equation 2.1.8), where σ is the interaction probability, or microscopic cross-section in units of *barns*.

For an incoming particle flux ϕ which is reduced to $\phi - d\phi$ through thickness dx of x, the fractional change in flux is given by Equation 2.1.7 [38, p. 25]. Where N_D is the number density of the target[2].

$$\frac{d\phi}{\phi} = -\sigma N_D dx \tag{2.1.7}$$

Integrating Equation 2.1.7 gives Equation 2.1.8, i.e. the unreacted particle flux reduces exponentially with target thickness, where ϕ_0 and ϕ_x are the initial and final fluxes, and x is the target thickness.

$$\frac{\phi_x}{\phi_0} = \exp\left(-\sigma N_D x\right) = \exp\left(-\Sigma x\right) \tag{2.1.8}$$

[2]The number density of a target is calculated by $\frac{\rho N_A}{m_A}$, where ρ ($g\ cm^{-3}$) is the target mass density, N_A (mol^{-1}) is Avagadro's constant and m_A ($g\ mol^{-1}$) is the atomic mass of the target nuclei.

The microscopic cross-section is dependent only on the incident particle energy and nuclide; it is independent of the global target properties. The macroscopic cross-section Σ (cm^{-1}) is the interaction probability per unit track length and is dependent on the target properties. This is calculated by multiplying the number density, N_D, by the microscopic cross-section σ. The reciprocal of the macroscopic cross-section provides the mean free path of neutrons within the target; with reference to Equation 2.1.8, this is the distance over which the intensity or particle flux is reduced by a factor of e. The reaction rate, R (s^{-1}), can be calculated by multiplying the incoming particle flux by the macroscopic cross-section and the volume of interaction, i.e. $R = \phi\Sigma V$.

The preceding discussion relates to the total interaction cross-section. This includes all individual scattering and absorption events which remove the particles from the phase space. Each of these interactions have a distinct microscopic cross-section, and the total is determined from the sum of each individual macroscopic cross-section. For a sense of the possible variation within the cross-sections, Figure 2.7 shows the energy dependence of the total, elastic and inelastic microscopic cross-sections with energy in ^{56}Fe. The total microscopic interaction cross-section varies by approximately 5 orders of magnitude within the energy range of 1 eV to 20 MeV (7 orders of magnitude).

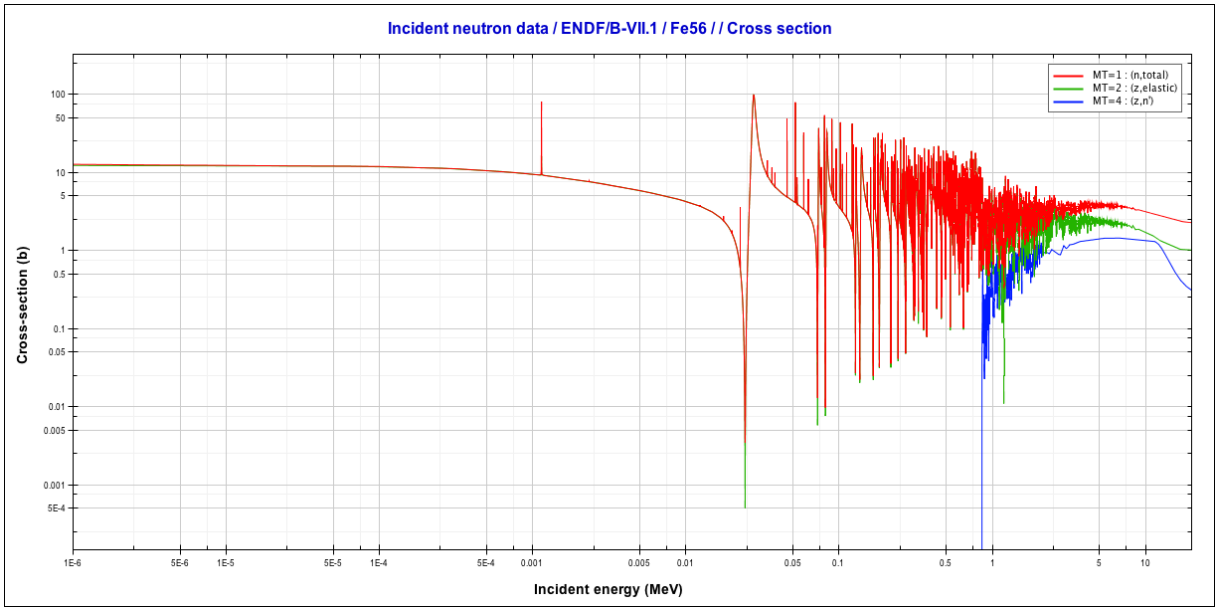

Figure 2.7: Interaction cross-sections as a function of incident neutron energy, for the total, elastic and inelastic reaction channels in ^{56}Fe. Three main regimes exist in most interaction cross-sections; at low energy the cross-section varies approximately as $1/\sqrt{E}$ or $1/v$. With increasing energy, more reaction channels become energetically available causing a deviation from this trend. The resonance region is characterised by the sharp peaks and troughs, in this example within the region of 1 keV to 2 MeV. Their locations correspond to the energy levels within the compound nucleus. At higher energies, the resonances occur more frequently and can no longer be resolved; hence this is the unresolved resonance region.

Differential and double-differential cross-sections

Where an interaction results in the emission of secondary particles, the properties with which they are produced are determined by the differential, and double-differential cross-sections. The differential cross-section can be with respect to energy or angle, it is the probability with which the exiting particle will have an energy E' or solid angle Ω in the centre-of-mass frame, as denoted by $\frac{d\sigma}{dE'}$ and $\frac{d\sigma}{d\Omega}$. The units of differential cross-section are $b\ sr^{-1}$, or $b\ eV^{-1}$. The double-differential cross-section is the probability that the exiting particle will have a specific energy and angle, and is written $\frac{d^2\sigma}{d\Omega dE'}$ or $\frac{d^2\sigma}{dE'd\Omega}$, in units of $b\ sr^{-1}\ eV^{-1}$. The solid angle, $d\Omega$, is subtended by $\sin\theta d\theta d\omega$, where θ is the polar angle and ω is the azimuthal angle. For scattering interactions, the system is symmetric in ω. Figure 2.8 shows the variation in polar angle distributions for different mass targets, energies and reaction channels.

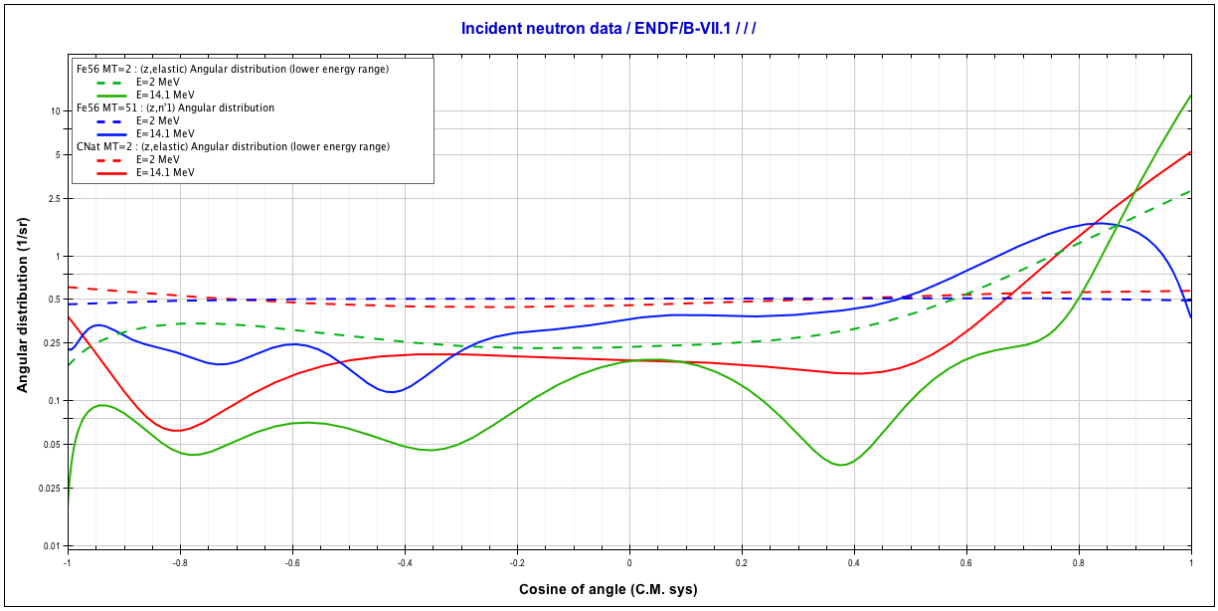

Figure 2.8: Differential probability distributions with respect to the scattering cosine, μ, in the centre-of-mass frame. Data are shown for the ^{nat}C(n,el), ^{56}Fe(n,el) and ^{56}Fe(n,n^{1*}) reaction channels. Dashed lines represent the lower energy, 2.0 MeV, distributions, and the solid lines represent the higher energy, 14.1 MeV, distributions. For the scatter of neutrons, low energy and low mass nuclei tend to have closely isotropic systems with respect to scattering cosine, when compared to high energies and masses the level of anisotropy increases.

2.1.4 Neutron transport equation

The transport of radiation within a certain region of phase space can be described by the neutron transport equation. This is a derivative of the Boltzmann transport equation [46], which is a general form for all radiation, both charged and uncharged. Setting the electromagnetic terms to zero results in the uncharged transport equation. All versions of the Boltzmann transport equation are conservation statements; they describe the net flow of radiation quanta through an element of phase space by quantifying the production and loss within that region. The phase space is defined within 7 dimensions, three in space, two in direction and one each in energy and time. This equates to the rate of change of neutrons within a small volume element dV, with energy dE about E and direction $d\vec{\Omega}$ about $\vec{\Omega}$. Figure 2.9 shows this differential phase space element schematically. The time dependent neutron density within this element is represented by $n(\vec{r}, E, \vec{\Omega}, t)dV\,dE d\vec{\Omega}$, where $n(\vec{r}, E, \vec{\Omega}, t)$ is the number of neutrons per unit time, t, at position $\vec{r}$.

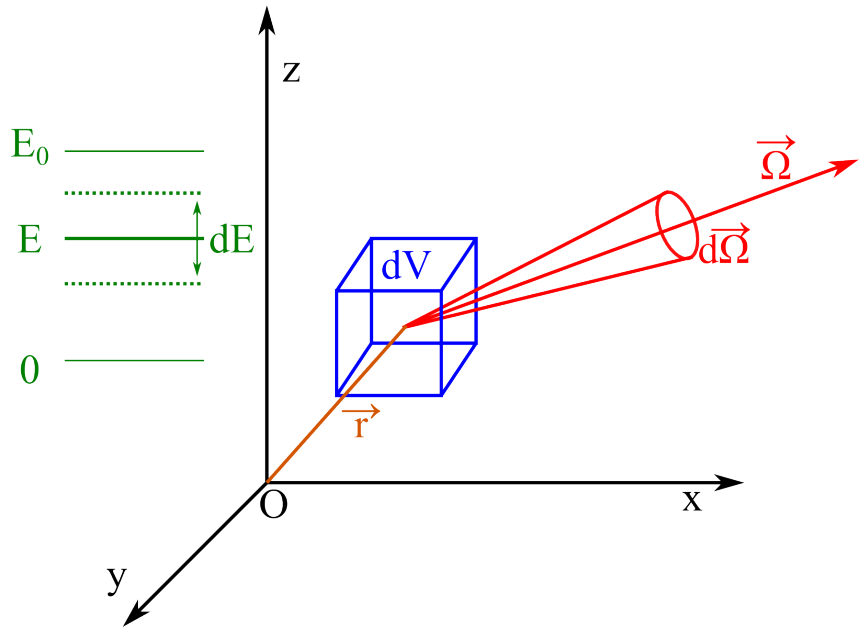

Figure 2.9: Schematic for the differential element of phase space $dV\,dE\,d\vec{\Omega}$ used within the neutron transport equation. The volume element, dV is three-dimensional, the directional element, $d\vec{\Omega}$ is two-dimensional with respect to polar and azimuthal angles, and the energy element dE. Adapted from [47, p. 440]

Neutrons within this differential element can be introduced via multiple mechanisms; an internal source within dV producing neutrons with energy dE about E and direction $d\vec{\Omega}$ about $\vec{\Omega}$, neutrons can flow into element dV from the surroundings with energy dE about dE and direction $d\vec{\Omega}$ about $\vec{\Omega}$, and finally neutrons can scatter within dV into the energy and directional ranges.

Neutrons within this differential element can be lost by absorption, i.e. the neutron undergoes a capture reaction that removes it from the system, flowing out of element dV and scattering out of the volume, energy or direction ranges. By balancing these production and loss components, the rate of change of neutrons is defined as $\frac{\partial}{\partial t} n(\vec{r}, E, \vec{\Omega}, t) dV dE d\vec{\Omega}$. All neutrons must be accounted for within the differential phase space element to provide continuity. The time-dependent, integro-differential form of the neutron transport equation is given in Equation 2.1.9.

$$\begin{aligned} &\frac{1}{v}\frac{\partial}{\partial t}\phi(\vec{r}, E, \vec{\Omega}, t) + \nabla.\vec{\Omega}\phi(\vec{r}, E, \vec{\Omega}, t) + \Sigma_{total}(E)\phi(\vec{r}, E, \vec{\Omega}, t) \\ &= \int_0^{4\pi} d\vec{\Omega}' \int_0^{\infty} \Sigma_{scatter}(E' \to E, \vec{\Omega}' \to \vec{\Omega})\phi(\vec{r}, E', \vec{\Omega}', t) dE' + S(\vec{r}, E, \vec{\Omega}, t) \end{aligned} \tag{2.1.9}$$

The neutron flux within element $dV dE d\vec{\Omega}$, is represented by $\phi(\vec{r}, E, \vec{\Omega}, t)$. The first term is the balancing component for the rate of change of neutron flux within the differential element, where v is the neutron speed and is proportional to $\sqrt{E}$. The second term is defined as the streaming term, and is the net flow of neutrons through the phase space. The third term is the loss of neutrons via interaction, this includes both scattering and absorption reactions, where Σ_{total} is the total macroscopic interaction cross-section. The fourth term corresponds to the in-scattering of neutrons from the surrounding regions, where $\Sigma_{scatter}(E' \to E, \vec{\Omega}' \to \vec{\Omega})$ is the macroscopic, double-differential scattering cross section from E' and $\vec{\Omega}'$ into the energy and direction of interest. This is integrated over

all possible incoming directions and energies. The final term is the source term for any neutron production localised within the region. All terms are given in units of particles per unit time.

2.2 Nuclear data

The neutron transport equation uses the macroscopic cross-sections and differential cross-sections to determine the probabilities with which each event will occur. These cross-sections are represented by nuclear data. These are the parameters used to describe the underlying physics of an engineering relevant problem. Within the scope of radiation transport simulations, the interaction cross-section data (with differential and double differential components) are included for each nuclide within a defined problem geometry. These are the basis of calculations to determine global parameters; for fusion these could be shut-down dose rates, tritium breeding ratios or shielding requirements. The quality and reliability of these global parameters are intimately linked to the nature of the input data; i.e. if these data do not properly represent the underlying physics, it is not possible to reliably predict the required engineering constraints. The discretisation introduced by pre-processing, or errors within the original evaluation contribute to this reduction of physical likeness. To design a commercially viable fusion reactor, the ability to quantify and minimise the errors within the nuclear data is vital.

Monte-Carlo based radiation transport results are typically quoted with a statistical error only; this makes the assumption that the underlying data and models are exact replications of the physical interactions. If this is not the case, the Monte-Carlo calculations are able to converge to a very high precision on a solution given sufficient histories, but this does not equate to accuracy. The data are subject to measurement and evaluation errors, and pre-processing discretisation, so all results should be quoted with an error that is sympathetic to these other sources in addition to the statistical component. Given the

multiple sources of error, this can be difficult to quantify. The method most widely used within the nuclear industry is the benchmarking of simulation results against physical experiments. There are a set of fusion specific, neutronics benchmarks within SINBAD (Shielding and Integral Benchmark Archive and Database) [48]. These benchmarks are quality checked, but many of these experiments are outdated and are missing information concerning experimental conditions and errors, so are limited in their applicability.

2.2.1 Nuclear data cycle

The data that are input to the radiation transport codes differ significantly from individual experimental measurements. For example, this could be the measurement of the total interaction cross-section at a single energy. Radiation transport codes require a full complement of data, i.e. they must cover the entire phase space of the application in materials, reaction channels, energy and angle. Evaluation is used to ensure the entire phase-space is covered, and when applied within a radiation transport system, informs the next iteration of experimental measurements in the areas most lacking. The cyclical nature of nuclear data can be seen in Figure 2.10. The little acknowledged pre-processing step is the focus of this work.

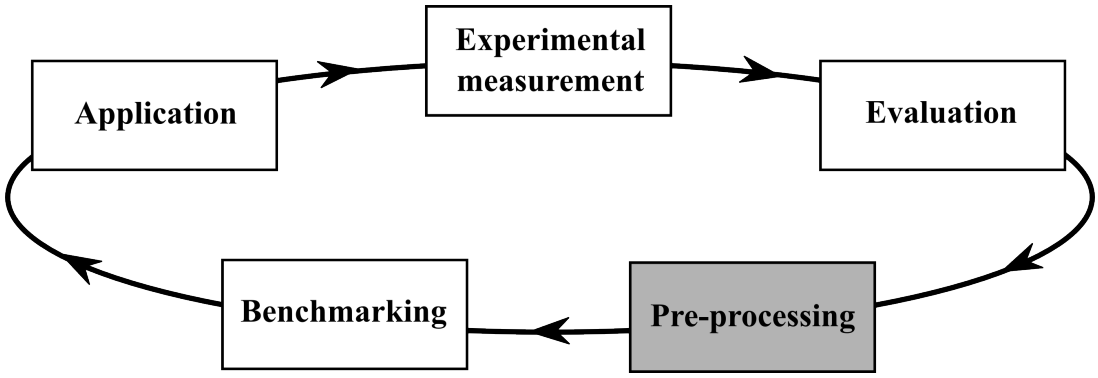

Figure 2.10: The nuclear data cycle, from experimental measurement through to application. For each iteration, the needs of the application determines the next set of experimental measurements. Each of the steps are described in the main text, Section 2.2.1.

Experimental measurement

The measurement technique depends on many different factors, the energy of the neutrons, quantity of interest and the availability of experimental facilities. Certain properties such as the individual reaction channel cross-sections, and differential and double differential components are more difficult to measure than integral quantities. When applied to a radiation transport problem, high quality experimental data are the grounding for a trustworthy simulation result. For each published measurement, the data are added to the EXFOR (EXchange FORmat) database with bibliographic and experimental details [49]. This database is managed by the NRDC (international network of Nuclear Reaction Data Centres) and is accessible internationally to encourage the sharing of data. The quality of the measurements are variable and they are sometimes incomplete; for example there could be uncertainty information missing, as is common with legacy data. There are also cases where conflicting information exists as seen in Figure 2.11. The measured values of the inelastic cross-section between 10 MeV and 20 MeV in ^{56}Fe are shown with the evaluated data.

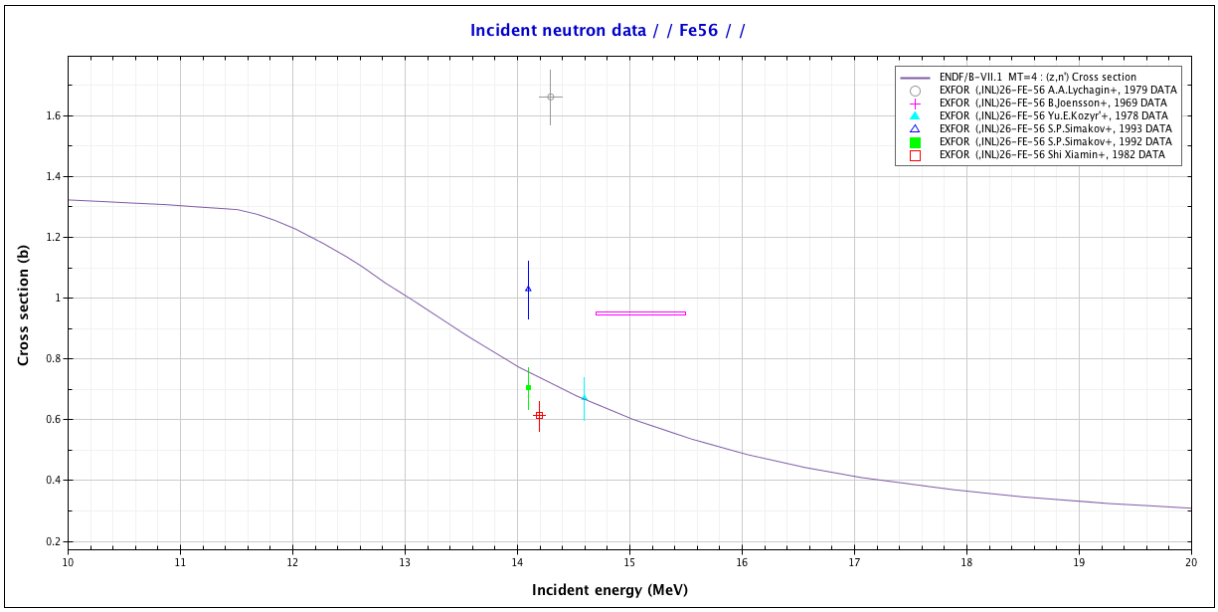

Figure 2.11: For the inelastic scatter of neutrons from ^{56}Fe cross-section data are shown within the range of 10-20 MeV. The individual points are taken from the EXFOR database and are from separate experiments. Where the points are equal in energy, none match within experimental errors and very little of the energy range is populated with experimental data. The line shows the evaluated cross-section data, taken from the ENDF/B.VII.1 library [19], this covers the entire energy range, though only two of the data points agree with the evaluation.

Evaluation

The details of this process are dependent on the evaluating authority, but the overall flow is generalised as a multi-step process for each nuclide. The evaluators determine which experimental data to include within the evaluation from EXFOR. Usually each authority has preferred sources, particularly when conflicting or incomplete information exists, as seen in Figure 2.11. These measured physical parameters are combined with well-known nuclear models in order to extract the fitted parameters, which allow for the extrapolation and interpolation over the entire range.

The precision and accuracy with which a measurement is made are crucial to the calculation of the fitted parameters. The degree to which the experimental data fit with the nuclear theory is quantified and recorded within covariance matrices. These matrices are now being included within the evaluation wherever possible, but there is no standard

format so the usefulness varies. The parameterisation of the experimental measurements, followed by the reduction to a tabulated format, is characterised as an evaluation [50, p. 2]. Each evaluation needs to contain data covering the entire range of the intended application, this usually covers energy from 10^{-5} eV to 20 MeV, and sometimes higher, in fusion applications. Any angular data must cover the full π radian range of exit angles. If the necessary experimental data are unavailable, the evaluator must rely on the systematics and nuclear models to provide the necessary information. The discrete experimental data points are used in conjunction with a nuclear model code, such as TALYS [51] to populate this phase space with tabulated data points and appropriate interpolation laws. The resulting parameters are output to the generalised ENDF-6 format (Evaluated Nuclear Data File)[50]. This is the standard method of nuclear data presentation for compatibility between authorities (Section 3.1 describes the ENDF format in detail). Each evaluation group packages the individual nuclides together to form libraries; the libraries typically used for fusion neutronics are listed in Table 2.2.

Table 2.2: List of fusion relevant, evaluated nuclear data libraries and their respective authorities [50, p. 5]

Library	**Authority/author**	**Newest release**
ENDF/B	United States Evaluated Nuclear Data File	ENDF/B-VII.1 (2011)
JEFF	NEA Joint Evaluated Fission and Fusion File (formerly JEF)	JEFF-3.2 (2014)
TENDL	Joint evaluation (IAEA, NRG, PSI), produced from TALYS nuclear model code	TENDL-2015 (2016)
FENDL	IAEA Fusion Evaluated Nuclear Data Library (ENDF/B + TENDL)	FENDL-3.1b (2015)

Pre-processing

As this is the main focus of this thesis, the pre-processing mechanisms and codes are described in detail in Section 2.2.2 and here the discussion is limited to the place of pre-processing within the nuclear-data cycle. The general purpose ENDF format is not

directly compatible with the majority of radiation transport codes; thus the data file must be processed into an application specific format. MCNP [23], Serpent [52] and FLUKA [53] require ACE format data (A Compact ENDF). These contain the interaction cross-section data, with resonances Doppler broadened to the set temperature, and any differential/double-differential components associated with the different reaction channels. The ENDF files are processed using a code such as NJOY [54] or PREPRO [55]; these codes use a series of modules to perform the necessary processing to create MCNP/Serpent/FLUKA compatible files.

Benchmarking

Benchmarking simulation data against physical experiments is used to verify the validity of the data libraries and measure the extent to which they can be applied. Experimental set-ups are converted into equivalent computational models, which replicate the physical geometry and source. Data libraries are then passed to this model and the simulation output data are compared to the original experimental measurements. If the two results differ significantly, the data are not considered suitable for that application. Databases of these experiments and their respective models have been compiled over many years, in addition to their relative qualities as a benchmark. The experiments are limited in scope, i.e. there are not sufficient benchmark experiments to cover the full range of applications so it is not possible to fully predict the effect of the data in every scenario. For fusion, the benchmarks are predominantly shielding experiments as given in SINBAD [48]. As the field continues to progress it is hoped that a broader range of fusion benchmarks will be produced.

Application

Once data files and libraries have successfully passed through the previous stages of the nuclear data cycle, they are considered suitable for use in applications. Typical fusion simulations include, but are not limited to, shielding, shut down doses, tritium-breeding and activation analyses. The use of the libraries in these applications informs the needs for the next iteration of the nuclear data cycle and future experimental campaigns.

2.2.2 Pre-processing

Within the context of the nuclear data cycle, radiation transport codes, such as MCNP, cannot accept the general purpose ENDF file without the appropriate conversion to ACE format. The ENDF files are created with all available information associated with the evaluation; not all of this is used directly for transport, while other parts are not in the correct format to be compatible with the code. The pre-processing step removes all extraneous data not directly required for the simulation and converts the remaining data into the accepted format. NJOY is the industry standard code for pre-processing data for MCNP via a series of individual modules. The first step is to produce a PENDF (Point-wise ENDF) containing the cross-section data for all reaction channels, whilst leaving the remainder of the file untouched. Some radiation transport codes, such as Tripoli [56], accept data in the PENDF format, subject to its own internal pre-processing. Figure 2.12 shows a typical module flow to produce the PENDF and ACE files. The PENDF production is the first stage of the two-step process to create the ACE; the second stage combines the PENDF with the processed differential and double differential data.

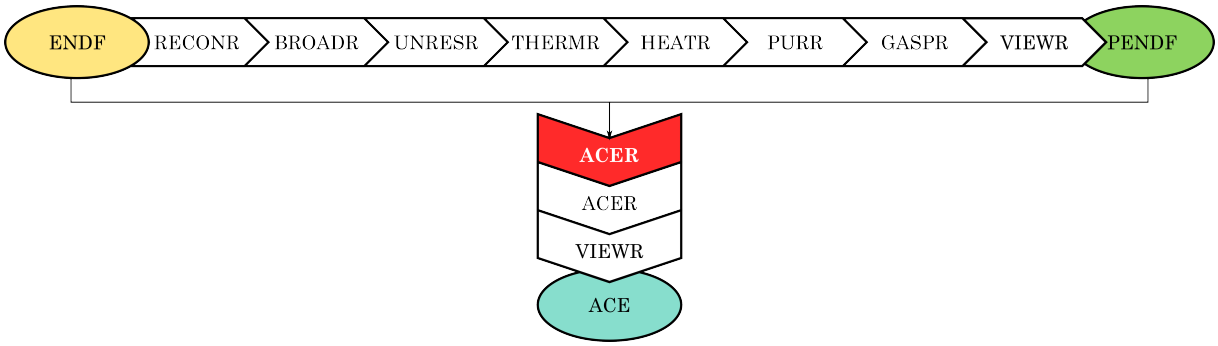

Figure 2.12: Module flow for NJOY to produce the ACE files for use in MCNP. The top chain of modules produces the PENDF file from one NJOY run, this is then passed, along with the original ENDF, to the ACER module as part of the second NJOY run. The ACER module is used twice; once to process the file and again to check for consistency using VIEWR. The descriptions of each of the modules can be found in Table 2.3. The ACER module highlighted in red is the main consideration for this work.

With reference to Figure 2.12, the first NJOY run reads in the ASCII ENDF file and converts it to binary with the MODER module (not shown). The description of each module used in the production of the PENDF is given in Table 2.3.

Table 2.3: Description of NJOY modules used within ENDF to PENDF conversion [54, 57]

Module	**Description**
RECONR	Creates a common energy grid, superimposes resonances and sums individual cross-sections to ensure that the totals are the exactly the sum of its parts
BROADR	Doppler broadening of the resonances at a set of user input temperatures
UNRESR	Calculates the effective cross-sections in the unresolved resonance region
THERMR	Thermal range cross-sections are calculated for coherent and incoherent, and free and bound scatters
HEATR	Produces the cross-section data associated with nuclear heating and radiation damage
PURR	Produces probability tables in the unresolved resonance region
GASPR	Produces point-wise cross-section data for the gas production reaction channels
VIEWR	Produce a set of plots to visually check the data

The cross-sections within the ENDF file are evaluated on a fine energy grid with resonances superimposed and Doppler broadened at the required temperature. Cross-sections in the unresolved region and thermal transport parameters are calculated. Nuclear heating, radiation damage and gas production are evaluated on a point-wise energy grid. Final visual checks can be made by inspecting the outputs from the VIEWR module.

While the resulting PENDF from these modules can be accepted by some codes at this point, NJOY must be run a second time to produce MCNP compatible files. The PENDF and the original ENDF are converted to binary with the MODER module, before being passed to the ACER module. For the sake of computational efficiency, the neutron and photon data within the PENDF and ENDF need to be further processed. For cross-section data, the linearly interpolated, point-wise data produced by RECONR and BROADR are used, thermal data are taken from the THERMR output and the unresolved probability tables are added from the PURR module.

Differential cross-section distributions, such as for two-body interactions, are represented by one of two formats within the ENDF: as a set of Legendre coefficients or as tabulated data points, with a given interpolation law. For the case of elastic and discrete level inelastic scattering, the distributions are always given in the centre-of-mass frame and tend to be represented by Legendre coefficients. This provides a continuous probability distribution over the entire angle and/or energy phase space. However, this format is unacceptable for MCNP and must be discretised to a form that is more computationally efficient. The two options are the legacy, 32 equal-probability histogram and the more recent, tabulated cumulative probability distribution with linear interpolation. Figure 2.13 shows an example of how the original form is represented with the two formats. The practicalities and methods of producing these data formats are seen in Sections 3.2 and 3.3.

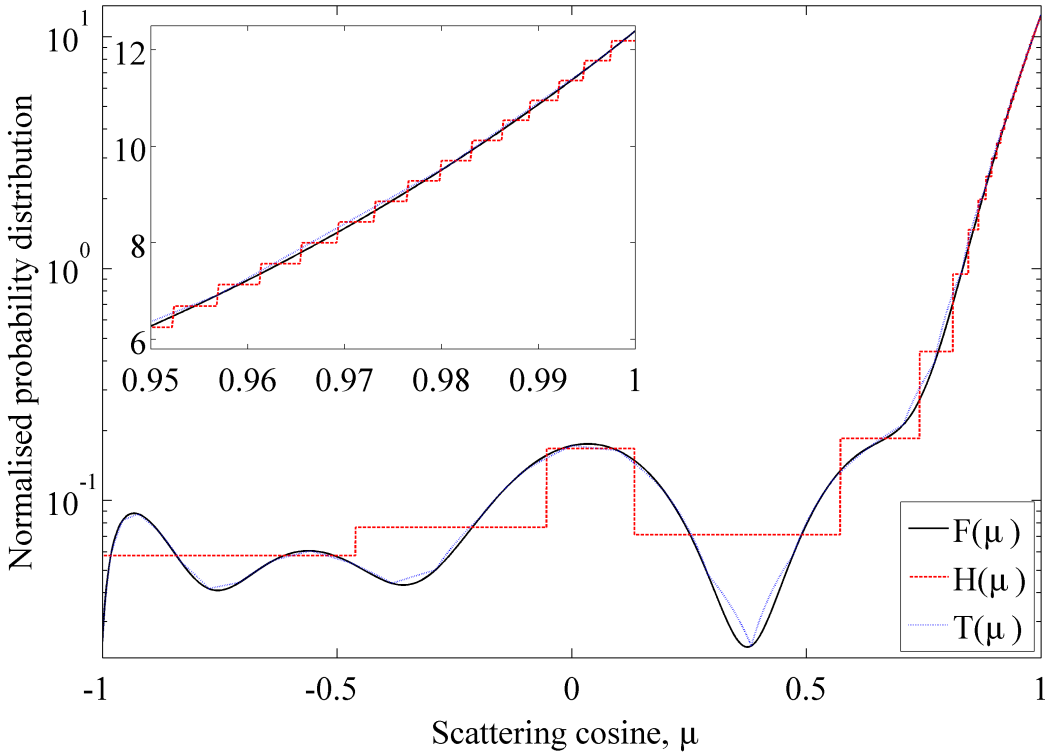

Figure 2.13: Plot shows the elastic scattering probability distributions for ^{56}Fe at 14.1 MeV, as a function of scattering cosine μ. The original, functional data ($F(\mu)$) is relatively well represented by the tabulated data ($T(\mu)$), but is crudely discretised by the 32 equal-probability histogram ($H(\mu)$). The inset shows the forward biased section of the distribution on a linear scale, a step-function from the histogram data overlays the functional data.

The legacy 32 equal-probability histogram format was created at a time when computational resources were scarce and efficiency was the key consideration for data formatting. Fission differential distributions are not badly affected by the application of the legacy method, as this is the regime for which it was designed. When this method is applied to higher energy regimes, such as for fusion, the distributions can be severely degraded. Tabulated cumulative probability distributions are slower to sample and tend to produce larger data files. Computational resources are more freely available and the additional runtime remains a consideration, but is not a limiting factor. After being checked for consistency with a second pass through ACER, each individual ACE file is combined within a library for use in MCNP, or similar.

The impact of pre-processing the angular component of differential and double-differential distributions has never truly been quantified. The following work demonstrates the problems associated with using legacy formats with non-legacy or non-standard applications, in addition to determining the sensitivity of full-scale Monte-Carlo radiation transport problems to these formats.

2.3 Accuracy quantification

To discuss the source and scale of errors within a radiation transport simulation, it is first necessary to consider the difference between accuracy and precision, specifically with respect to the Monte-Carlo approach. When a Monte-Carlo simulation is performed, the problem converges to the expectation value of the system with respect to the number of source histories. The precision of this result refers to the statistical spread of sampled values which have contributed to the final answer. With the Monte-Carlo method it is possible to achieve an arbitrary level of precision; the variance around the expectation value can be reduced to negligible levels by increasing the particle population within the model. The accuracy of this result is defined as the closeness of the expectation value to the true value. The precision of a simulation result can be improved via the methods described in Section 2.4.2, but to properly quantify the accuracy of that result is not as straightforward.

2.3.1 The nature of errors

Factors which affect the accuracy of a simulation result can be separated into three broad categories: code, model and human. All three of these categories contribute to the total systematic error, and their effect on the simulation result is not always obvious.

Code based factors include how the physics in the system is modelled, the uncertainties and quality of the input nuclear data and any bugs within the source code. The model

based factors affecting accuracy include the level of detail included in the geometry, the adequacy of the source definition and the definition of the materials. These types of error can be reduced by better replicating the physical set-up with a fine level of detail. Human based factors are largely due to user error in creating the input file, selecting the appropriate data libraries and the interpretation of the output data. All Monte-Carlo based codes perform basic statistical analyses on the output data, but it is ultimately the user who must determine if the results are correct. This requires a good knowledge of the model, expected values and how the Monte-Carlo method works, particularly if non-analogue techniques are employed. Of the three contributors to the systematic, the human element is the most difficult to predict and eradicate.

Errors or uncertainties within nuclear data can be further categorised into experimental, evaluation and pre-processing areas, as mentioned in Section 2.2. All three of these contribute to the overall systematic error of the simulation result, by quantifying and reducing them, the accuracy of the simulation can be improved. Improved confidence in the result could ultimately reduce the engineering tolerances built into fusion devices and hence reduce construction costs.

It is not possible to experimentally measure any nuclear data quantity, such as interaction cross-sections and differential distributions, without an associated measurement uncertainty. These values are associated with their own systematic and random errors as part of the experimental method. Data are made accessible to the evaluators via the EXFOR sharing platform with as much uncertainty information as possible; though many of the older experimental data are presented with little or no error information and it is impossible to retrieve.

The uncertainty introduced from the evaluation stage includes approximations or errors within the physics models, and the chosen experimental data used to refine the model fitting. Where many experimental data points are available within well understood en-

ergy regimes, the evaluated data can be highly reliable. When less standard materials, energies and reaction channels are considered, the current experimental data and models are insufficient in many cases. Particularly for the case of fusion, the level of nuclear data is severely lacking compared to fission relevant data. Recent libraries have been released with covariance matrices. These quantify how well the models fit the experimental data. In some cases data uncertainties have been propagated to simulation results with a total Monte-Carlo approach [58, 59]. This involves producing a set of randomised cross-sections within the uncertainty limits of the data using the TENDL-6 framework. The same model is used for all data sets and the subsequent results are compared. This is impossible when uncertainty information is missing from evaluations.

Pre-processing affects the nuclear data in various ways depending on the required output format and nature of the information being processed. Section 2.2.2 described the method of processing evaluated data into the MCNP compatible format. Uncertainties are introduced during the reconstruction of the cross-sections and resonances, and when the differential data are converted to the discretised format. The uncertainty introduced due to the pre-processing of the evaluated data files is often considered to be negligible. In the past, this may have been true, but for applications such as fusion neutronics this assumption no longer holds.

To quantify the effect of pre-processing nuclear data, specifically with respect to differential distributions, it is not useful to consider parameters such as the average difference between the processed and unprocessed distributions. The nature of the processing causes all differences to sum to zero over the range of the distribution. With reference to Figure 2.13, the ^{56}Fe(n,el) reaction channel at 14.1 MeV shows a highly featured back scatter region with a low probability, and an extreme forward bias. The forward peaked elastic scattering data, with a large rate of change, is poorly replicated with certain formats.

Within Figure 2.13, the two currently available processed formats are overlaid upon the

original distribution. The 32 equal-probability histogram is a crude representation of the original over the entire distribution. The tabulated data are capable of replicating the original distribution to a reasonable degree, with a sufficient number of points. In both cases, the sum of the differences over the μ range equal zero, and any standard calculation of mean difference are irrelevant. This is an artefact of the normalisation required to ensure safe sampling of the distributions; i.e. they cannot exceed the bounds of the Monte-Carlo sampling schemes. Applying the pre-processing methods described in Section 2.2.2 removes the smooth, functional description and replaces it with a coarse, discretised data set.

Two methods have been used to quantify how well the processed data represents the original: the maximum difference and the coefficient of variation. The first is a legacy method, and the second is one that is non-traditional for these applications. Each provides unique information concerning the distributions, and the shortfalls of each are covered by the other.

2.3.2 Maximum difference

The maximum difference is calculated from the evaluation of processed and original distributions on a fine μ grid. One is subtracted from the other and the largest residual, positive or negative, is taken as the maximum difference. Based on the data in Figure 2.13, the corresponding plot of differences is shown in Figure 2.14.

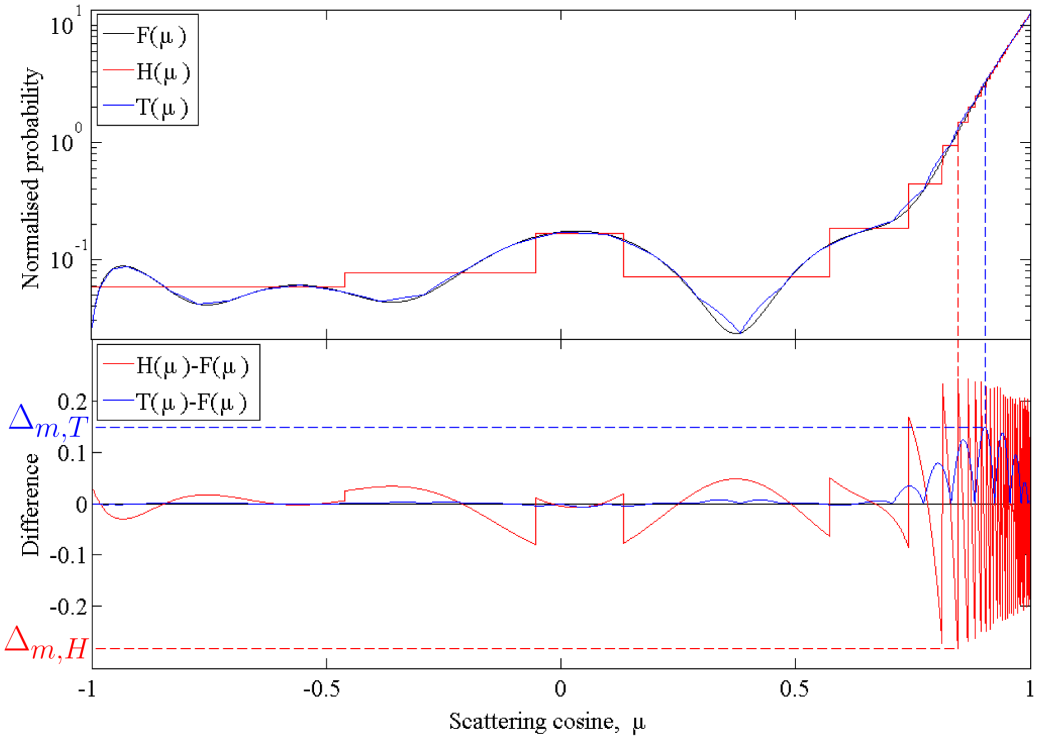

Figure 2.14: The top plot echoes the distributions in Figure 2.13 and the bottom plot shows the difference between the processed formats and the original as a function of μ. The maximum difference, Δ_m is marked for both histogram and tabulated formats. For the maximum difference, all other information is discarded.

The maximum difference, Δ_m, given by Equation 2.3.1, provides a measure of the largest mis-representation of the original data $F(\mu, E)$ by the processed form $G(\mu, E)$.

$$\Delta_m = \max_{-1 \leqslant \mu \leqslant 1} \left(F(\mu, E) - G(\mu, E) \right) \tag{2.3.1}$$

This is capable of quantifying both the nature and magnitude of any narrow or peaked features that are lost through pre-processing. The maximum difference must be normalised if it is to be comparable to the results from other isotopes. Historically this has been the standard technique for comparing two distributions within the field of nuclear data processing [47, p. 300], but it is not capable of providing a measure of how well

(or how badly) the processed format replicates the original over the entire range of the distribution, only how well it represents a single point.

2.3.3 Coefficient of variation

The coefficient of variation, C_v, provides a measure of how far, on average, the residual deviates from zero; this is used to determine the accuracy of the processed representation over the entire μ range. A C_v of zero corresponds to an exact replication and a large C_v corresponds to a poorly replicated distribution. This method is not typically used in nuclear data analyses, but works well to quantify the difference between two continuous distributions. The C_v is calculated using Equation 2.3.2, where the two distributions $F(\mu)$ (original) and $G(\mu)$ (processed) are evaluated on a fine grid of N points.

$$C_v = \frac{1}{\overline{F(\mu, E)}} \sqrt{\frac{\sum_{i=0}^{N} (F(\mu_i, E) - G(\mu_i, E))^2}{N}} \tag{2.3.2}$$

For clarity, this process is shown in Figure 2.15. By using the root mean square of the differences the magnitude of the differences is retained, regardless of their sense. The normalisation by the original distribution average allows for the comparison between methods, materials and reaction channels. The C_v is often expressed as a percentage and provides a measure of how closely the processed data follows the original over the entire range, with respect to the mean of the original distribution [60]. The main down-side of using the C_v, is that the nature of the differences are lost, i.e. it is impossible to determine from the C_v alone whether a region will be under or over sampled when compared to the original. The application of the maximum difference and coefficient of variation can be seen in Chapters 3 and 4.

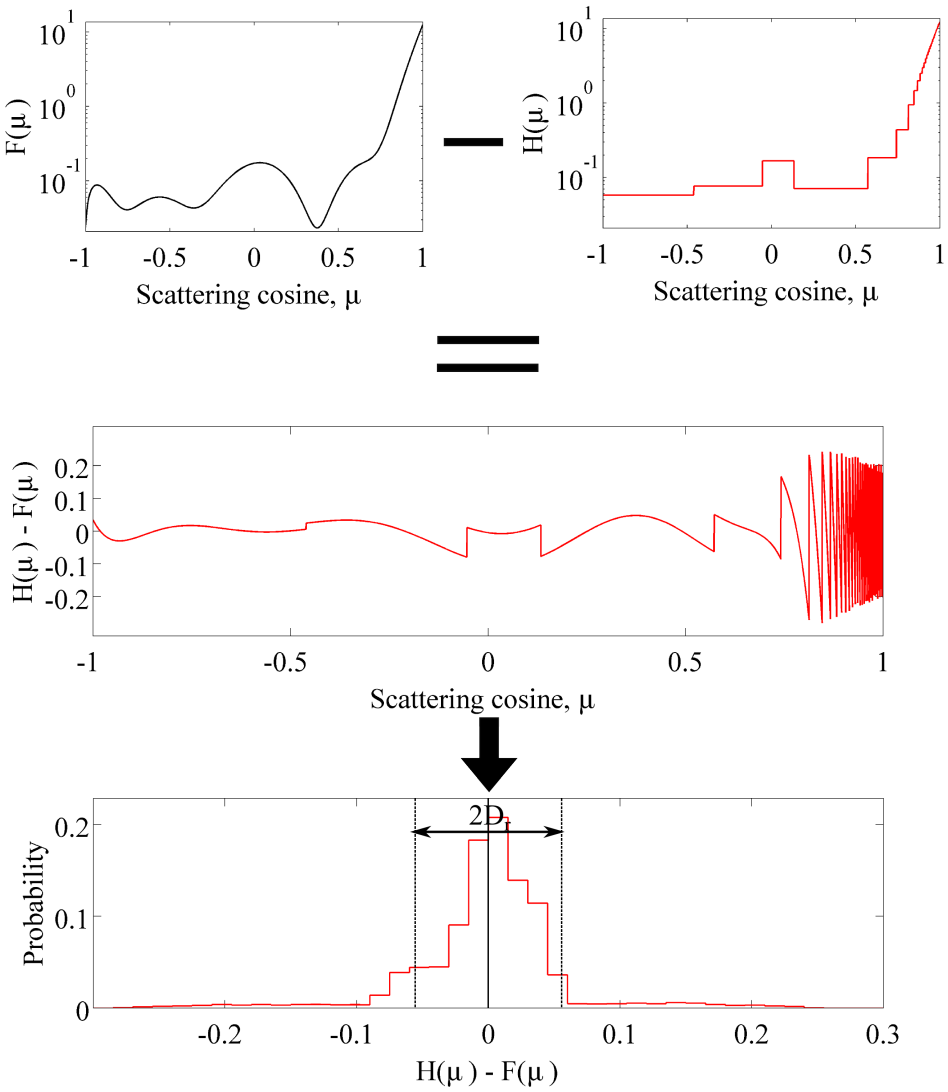

Figure 2.15: The process of calculating the root-mean-square deviation (D_r), which is subsequently used to calculate the coefficient of variation. Equation 2.3.2 can be written in terms of D_r; $C_v = D_r/\overline{F(\mu, E)}$. In this instance, the 32 equal-probability histogram, $H(\mu)$ is subtracted from the original, functional form $F(\mu)$ for 10^5 equally spaced samples in μ. The difference can be seen in the middle figure with respect to the scattering cosine. The bottom figure shows the histogrammed differences, the area marked by the dashed lines is equal to D_r for this data set. The solid black line marks the mean difference, this is always equal to zero.

2.4 Computational methods

Section 2.1 covered the nature of neutron interactions, and how their behaviour can be described by the neutron transport equation (Equation 2.1.9). The neutron behaviour within a system cannot be accurately solved with the full neutron transport equation. The phase space covered by this equation is seven dimensional; computational capabilities are insufficient to solve this equation for the entire phase space for anything but the simplest of problems [61].

In practice there are two methods for determining neutron behaviour within a system, deterministic or stochastic. Deterministic methods reduce the phase space of the neutron transport equation to allow for a direct solution. The phase space reduction is in the form of simplifying assumptions and discretisation. It provides a fast and repeatable, solution to an approximated model. These methods solve for the particle behaviour in the Eulerian reference frame, i.e. the overall observable neutron behaviour.

Stochastic methods directly replicate the behaviour of neutrons within the system based on empirical neutron interaction data. The neutrons can be transported in both space and time, without any simplifications. This is not a direct solution to the neutron transport equation, but provides statistical answers to integral quantities; i.e. they can provide an approximate answer to an exact model. This is a stochastic solution in the Lagrangian frame of reference, where the mean behaviour over many samples converges on the Eulerian solution.

2.4.1 Deterministic methods

There are several different approaches to solving Equation 2.1.9 deterministically in the steady state. These methods apply restrictions on the phase space but the approximations remove a level of physical representation and cannot reliably be applied to non-trivial problems. Deterministic solutions are fast to compute, as only a single calculation is re-

quired once the necessary simplifications have been made. For the same reason the results are always repeatable. The three main simplifications used in deterministic calculations are energy, angular and spatial.

The most common method of discretising the energy region of the transport equation is via the multi-group method. The continuous, energy dependent sections of Equation 2.1.9 are split into groups, most notably the interaction cross-section data. Complex, continuous cross-sections with respect to energy, such as those shown in Figure 2.7, are condensed to a set of tabulated data points. Each of these points are calculated from the flux weighted average of the continuous data within the chosen group limits. To retain as much physical information as possible, the groups must be of a high enough resolution to ensure a reasonable representation of the original data. It follows that a complex distribution with many resonances requires more groups than one with few resonances. The neutron transport equation reduces to Equation 2.4.1 with the multi-group approximation, where g is the energy associated with the group and all other parameters have their usual meaning. Equations 2.4.2 to 2.4.4 show the relationship of the group equation to the original neutron transport equation [47, p. 346].

$$\begin{aligned} &\frac{1}{v}\frac{\partial}{\partial t}\phi_g(\vec{r},\vec{\Omega},t) + \nabla.\vec{\Omega}\phi_g(\vec{r},\vec{\Omega},t) + \Sigma_{total,g}\phi_g(\vec{r},\vec{\Omega},t) \\ &= \frac{1}{4\pi}\sum_{g'}\int_0^{4\pi} d\vec{\Omega}' T(g',g'\rightarrow g,\vec{\Omega}'\rightarrow\vec{\Omega})\phi_{g'}(\vec{r},\vec{\Omega}',t) + S_g(\vec{r},\vec{\Omega},t) \end{aligned} \tag{2.4.1}$$

$$\phi_g(\vec{r},\vec{\Omega},t) = \int_{E'_g}^{E'_{g+1}} \phi(\vec{r},E,\vec{\Omega},t)dE' \tag{2.4.2}$$

$$\Sigma_{total,g}\phi_g(\vec{r},\vec{\Omega},t) = \int_{E'_g}^{E'_{g+1}} \Sigma_{total}(E')\phi(\vec{r},E',\vec{\Omega},t)dE' \tag{2.4.3}$$

$$T(g',g' \rightarrow g,\vec{\Omega}' \rightarrow \vec{\Omega})\phi_{g'}(\vec{r},\vec{\Omega}',t) =$$

$$\int_{E_g}^{E_{g+1}} dE \int_{E'_g}^{E'_{g+1}} \Sigma_{scatter}(E' \rightarrow E,\vec{\Omega}' \rightarrow \vec{\Omega})\phi(\vec{r},E',\vec{\Omega}',t)dE' \tag{2.4.4}$$

Angular discretisation takes two main forms, discrete ordinates (S_N) or spherical harmonics (P_N). The discrete ordinates method describes the angular variation of the neutron flux as a set of N discrete points with respect to μ. For the simple case of a one-dimensional, plane geometry, the neutron transport equation can be re-written as Equation 2.4.5 [62, p. 239].

$$\mu\frac{\partial}{\partial x}\phi_g(\mu,x) + \sigma_g^{S_N}(x)\phi_g(\mu,x) = \sum_{l=0}^{N} P_l(\mu)\sum_{g'}\sigma_{l,g\leftarrow g'}^{S_N}(x)\phi_{l,g'} + S_g(\mu,x) \tag{2.4.5}$$

The integral components are replaced with weighted Legendre polynomials. These weights are chosen to be as close to the expected flux as possible; for thermal systems the flux is well known. Within highly anisotropic systems it is far more difficult to estimate these weighting factors appropriately, hence this method is not typically used for fusion systems. The P_N method uses spherical harmonics to describe the angular variation of flux. The number of terms required to describe the flux is dependent on the complexity of the problem. This results in a semi-discrete version of the neutron transport equation that

can then be solved numerically. Equation 2.4.6 shows the P_N form of the Boltzmann transport equation, where $\sigma_{l,t,g}^{P_N}$ and $\sigma_{l,g\leftarrow g'}^{S_N}$ are the group averaged cross sections [62, p. 181].

$$\mu\frac{\partial}{\partial x}\psi_g(\mu,x)+\sum_{l=0}^{N}P_l(\mu)\sigma_{l,t,g}^{P_N}\psi_{l,g}=\sum_{l=0}^{N}P_l(\mu)\sum_{g'}\sigma_{l,g\leftarrow g'}^{S_N}(x)\psi_{l,g'}+S_g(\mu,x) \tag{2.4.6}$$

The final method covered here is the diffusion approximation. This is based on the assumption that neutrons behave similarly to a gas, in the sense that areas of high concentration will diffuse to areas of low concentration. The equivalent form of the neutron transport equation is shown in Equation 2.4.7.

$$\begin{aligned}\frac{1}{v}\frac{\partial}{\partial t}\phi(\vec{r},E,t)+\nabla.D(\vec{r},E)\nabla.\phi(\vec{r},E,t)+\Sigma_{total}(E)\phi(\vec{r},E,t)\\ =\int_0^\infty\Sigma_{scatter}(E'\rightarrow E)\phi(\vec{r},E',t)dE'+S(\vec{r},E,t)\end{aligned} \tag{2.4.7}$$

This equation only holds in the case of flux having a weak spatial and angular dependence and a slowly varying current density, i.e. no sources or sinks for neutrons. For situations such as homogenous fission reactor cores, this approximation is simple and accurate. It cannot however by used when the angular dependence of the flux is anisotropic, or if there are any materials boundaries within the problem.

Deterministic methods have their advantages, but all methods involve approximations and restrictions. This lends itself to some areas of fusion neutronics, but in general deterministic methods are not appropriate for this application.

2.4.2 Stochastic methods

The alternative to deterministic solutions to the neutron transport equation is to use stochastic methods. The most well known stochastic approach is the Monte-Carlo method. Neutron transport is largely stochastic so the Monte-Carlo method of simulating particle behaviour is a natural choice. This method is particularly advantageous as it is capable of simulating neutron behaviour in complex, three-dimensional systems with virtually no approximations. This method uses empirical nuclear data in conjunction with random sampling techniques to directly simulate the trajectory of neutrons within the system. By simulating many neutrons, their average behaviour within the system can be inferred.
The Monte-Carlo approach is primarily an integration technique and does not directly solve the neutron transport equation. However it is capable of producing engineering relevant, measurable quantities via the indirect solution of the integral form of the neutron transport equation. Directly simulating the Lagrangian particle behaviour allows the inference of the Eulerian behaviour over many samples. Monte-Carlo radiation transport was designed to study complex systems and covers the entire phase space of the transport equation. It is capable of handling many interacting components, including mixed particle fields. By combining known interaction probability laws, such as cross-section data into the Monte-Carlo process, the resulting particle tracks are abstract analogues of real world processes. The main downside to this method is the time of computation required. Deterministic methods require only one calculation once the problem has been properly defined. Monte-Carlo methods require large numbers of simulated particles to determine the average behaviour within a reasonable level of precision. However, it is possible to reduce the statistical uncertainties to negligible levels by increasing the particle population so the problem is, theoretically, only limited in accuracy by the input parameters. Large scale simulations, such as the DEMO or ITER models require a minimum of 10^9 source particles to fully populate the phase space. This equates to many days worth of

computational time.

Within Monte-Carlo radiation transport, the neutron transport equation can be reduced to a set of integral operators as shown in Equation 2.4.8.

$$\chi(\vec{r}, E, \vec{\Omega}, t) = S(\vec{r}, E, \vec{\Omega}, t) + C(\vec{r}, E' \rightarrow E, \vec{\Omega}' \rightarrow \vec{\Omega}) T(\vec{r}' \rightarrow \vec{r}, E, \vec{\Omega}).\chi'(\vec{r}', E', \vec{\Omega}', t) \tag{2.4.8}$$

The particle density is represented by χ at $\vec{r}$ with direction $\vec{\Omega}$ and energy E. The source operator, $S(\vec{r}, E, \vec{\Omega}, t)$, describes the external generation of particles. The collision operator, $C(\vec{r}, E' \rightarrow E, \vec{\Omega}' \rightarrow \vec{\Omega})$, describes all interactions including neutron capture. This determines the energy and direction components entering into E and $\vec{\Omega}$ at $\vec{r}$. The transport operator, $T(\vec{r}' \rightarrow \vec{r}, E, \vec{\Omega})$, moves the neutron between interaction sites and $\chi'(\vec{r}', E', \vec{\Omega}', t)$ is the initial particle density.

The analogue event chain of Monte-Carlo radiation transport can be described in terms of these operators. The transport is described for the simple example of a high-energy neutron incident on a block of tritium breeding material, as shown in Figure 2.16. It undergoes many different reactions which are stochastic in nature, producing additional particles throughout the history. Similar histories will be statistically equivalent, but will demonstrate different physical processes based on their relative probabilities.

A source neutron is generated with energy E and direction $\vec{\Omega}$. The path length to the next interaction, R is sampled from the total macroscopic interaction cross-section of the material. The transport operator then moves the neutron distance R along the direction vector $\vec{\Omega}$ to the interaction site. The collision operator determines the interacting nuclide, reaction type and emission properties based on the random sampling of the relative probabilities at energy E. The energy and direction of the neutron are updated if it has not been removed from the system via capture. This sequence of transport and collision is

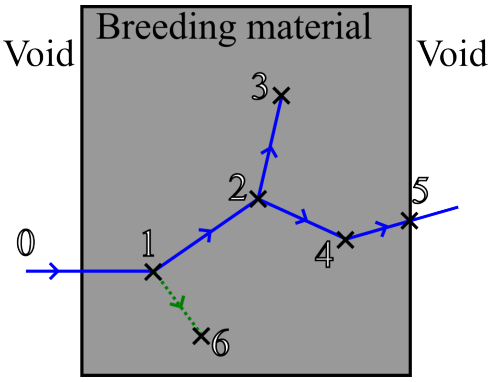

Figure 2.16: Example history of a neutron incident on a block of tritium breeding material. In this example, a source neutron generated at 0 undergoes an inelastic scatter at 1, with the production of a secondary photon. This is banked, and will be followed after the main neutron history has terminated. The neutron properties are updated, and it continues to 2, where it undergoes a neutron multiplication reaction. One of the neutrons is banked and the other continues to position 3, where it is captured in a tritium breeding reaction and terminated. The banked neutron is transported to position 4 where it undergoes an elastic scatter before leaking from the cell at position 5. The photon produced at 1 is followed to position 6 where it is captured and hence terminated.

repeated until the neutron is terminated, either through escaping the geometry, capture or a computational cut-off. Random sampling is at the heart of the Monte-Carlo method. Each time a random number is required, it is sampled uniformly between 0 and 1 and then mapped on to the relevant probability distribution. Figure 2.17 demonstrates this with the angular distribution for the elastic scatter of a 14.1 MeV neutron from ^{56}Fe. It is given with the integral form of the distribution which increases from 0 to 1. A random number, ξ, is sampled and the distribution solved to give the value of μ.

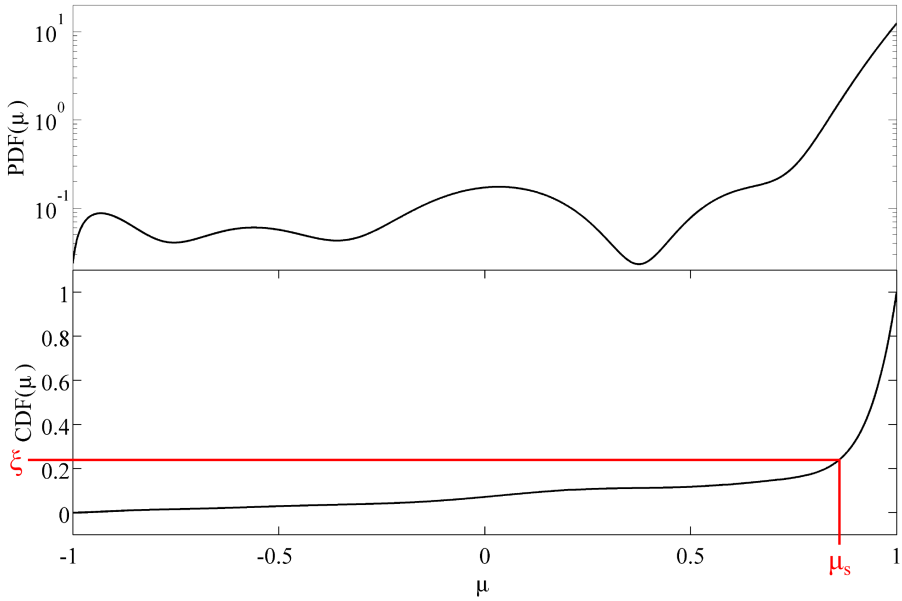

Figure 2.17: Top figure shows the normalised probability distribution function for the elastic scatter of 14.1 MeV neutrons from ^{56}Fe. This is integrated to give the cumulative probability distribution between 0 and 1. For a random number ξ sampled between 0 and 1, it is mapped onto the cumulative probability distribution to give the sampled value of μ_s.

Each history within the Monte-Carlo simulation will follow the same computational event chain. This is shown in Figure 2.18. Each area associated with the operators in Equation 2.4.8 are highlighted. The area not shown on this flow diagram is the geometry tracking. At each point, whether it is collision or transport, the neutron must know where it is within the geometry to determine which are the relevant interaction cross-sections.

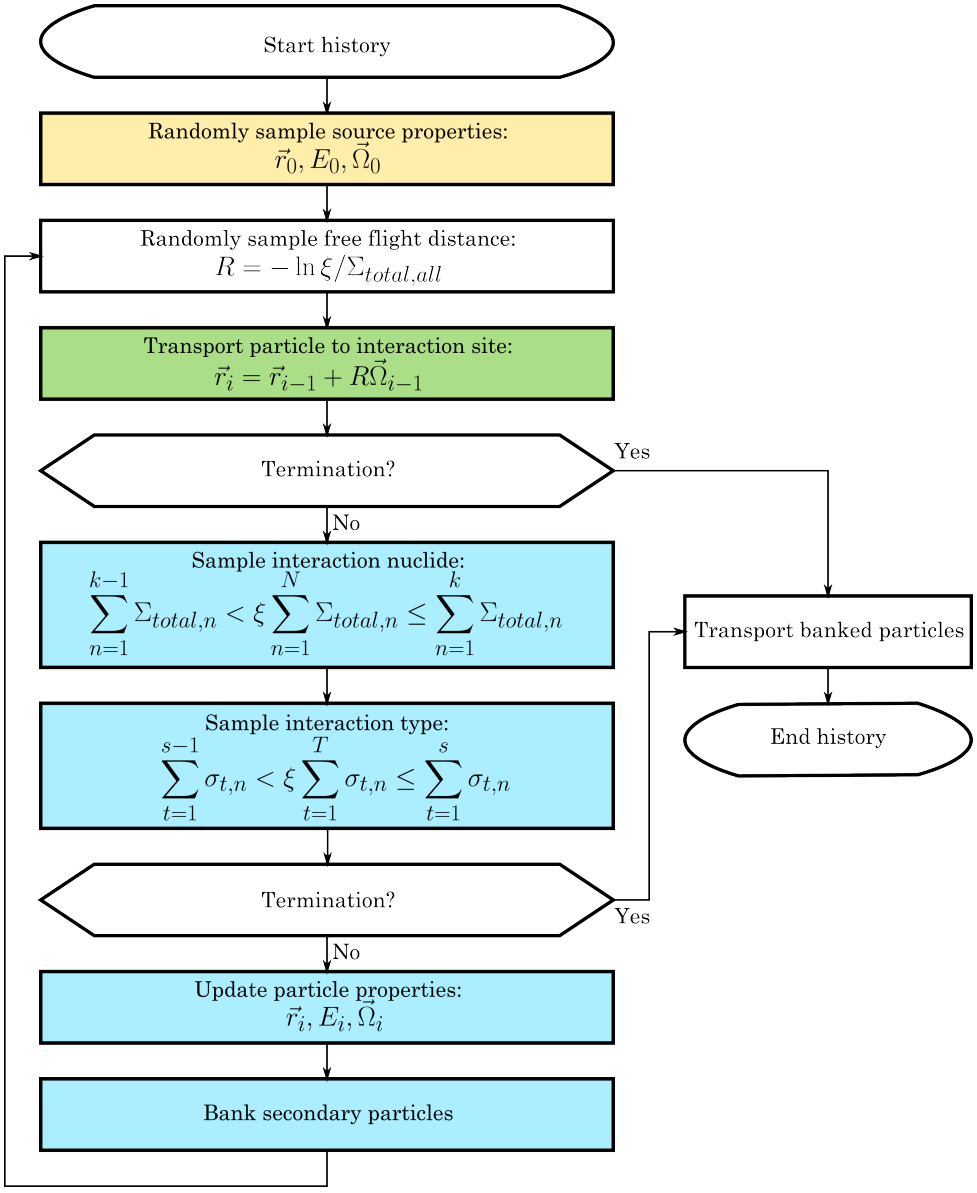

Figure 2.18: Monte-Carlo event chain for a particle history within analogue Monte-Carlo. With reference to Equation 2.4.8, the components associated with the source operator are highlighted in yellow, transport operator in green and the collision operator in blue. All random numbers (ξ) are sampled between 0 and 1. The total interaction cross-section for the material $\Sigma_{total,all}$ is the sum of all nuclides within the material. The total cross-sections for each nuclide $\Sigma_{total,n}$ are scaled to sum to 1. A random number is sampled and used to choose nuclide k of N. Once the nuclide has been selected, the same process occurs to select the interaction type, s of T from the microscopic cross-sections σ. If the particle has not been terminated, the properties are updated and the history continues. Any secondary particles are banked and transported once the main particle has terminated.

Simulating the neutron history is only part of the process. Tallying is used to acquire a specific quantity of interest, for example the particle flux through a surface, reaction rate or dose deposited within a volume. There are several different types of information that can be extracted, the simplest being the particle current, J, through a surface. This is represented by Equation 2.4.9, where, for analogue transport, the scored quantity is the number of particles crossing the surface.

$$J = \int dE \int dt \int dA \int d\vec{\Omega} |\vec{\Omega}.\vec{n}| \psi(\vec{r}, E, \vec{\Omega}, t) \tag{2.4.9}$$

The average surface flux, $\overline{\phi_s}$ is described by Equation 2.4.10, and is given in units of particles per unit area, where A is the area of the surface in question. The scored quantity is $1/(|\mu|A)$, where $|\mu| = |\vec{\Omega}.\vec{n}|$ and $\vec{n}$ is the vector for the surface normal.

$$\overline{\phi_s} = \frac{1}{A} \int dE \int dt \int dA \int d\vec{\Omega} \psi(\vec{r}, E, \vec{\Omega}, t) \tag{2.4.10}$$

The volume flux, $\overline{\phi_V}$, is calculated with Equation 2.4.11 and is given in units of particles per unit volume, where V is the volume of the cell. The scored quantity is T_l/V, where T_l is the track length of particles within the cell. The volume flux can be multiplied by interaction cross-sections to calculate a reaction rate within a volume.

$$\overline{\phi_v} = \frac{1}{V} \int dE \int dt \int dA \int d\vec{\Omega} \psi(\vec{r}, E, \vec{\Omega}, t) \tag{2.4.11}$$

These are the three most commonly used tallies in MCNP, though several others exist, such as point detectors, energy deposition and pulse height tallies. These are covered in

detail in reference [23, p. 2-80]. Each tally can be subdivided into energy, cosine and time bins. This converts the continuous integrals into sums over a discrete set of limits.

All tally results are subject to an associated statistical uncertainty. Central to the Monte-Carlo method is the law of large numbers; this describes the long-term stability of the mean of a variable. The calculated mean tends to the expectation value over a large number of sampled events. In the case of radiation transport, sufficient particle histories must be simulated for the expectation value to be reached within the desired level of statistical uncertainty. The central limit theorem is used to calculate a mean, standard deviation, variance and standard error. This states that the observations of a quantity, which is influenced by many independent and identically distributed random variables, tends to a normal distribution. These two principles form the basis of the statistical tests that are used to determine whether the problem has converged on the expectation value within the required limits of uncertainty.

For any tally, the sample mean, $\overline{x}$, is calculated as the sum of all contributions, x_i, over the number of source particles, N, as given in Equation 2.4.12.

$$\overline{x} = \frac{1}{N} \sum_{i=1}^{N} x_i \tag{2.4.12}$$

The standard deviation, S, is calculated from Equation 2.4.13, and can be approximated to $\sqrt{\overline{x^2} - \overline{x}^2}$ for a large number of histories. This demonstrates a $1/\sqrt{N}$ relationship, as with Poisson statistics, which should be replicated within the simulation results.

$$S = \sqrt{\frac{\sum_{i=1}^{N} (x_i - \overline{x})^2}{N - 1}} \tag{2.4.13}$$

The variance about the sample mean $S_{\overline{x}}^2$ is given by Equation 2.4.14 and is used to calculate the relative error, $R_{\overline{x}}$ (Equation 2.4.15) [63].

$$S_{\overline{x}}^2 = \frac{S^2}{N} \tag{2.4.14}$$

$$R_{\overline{x}} = \frac{S_{\overline{x}}}{\overline{x}} \tag{2.4.15}$$

This informs upon the confidence limits about the mean value; for an infinite number of particles, there is a 68% probability that the true value will lie within the range of $\overline{x}(1 \pm R_{\overline{x}})$, and a 98% probability that it will be within the range of $\overline{x}(1 \pm 2R_{\overline{x}})$ [23, p. 2-114]. This only holds if the Monte-Carlo technique has sufficiently sampled the source and populated the geometry, otherwise a statistically rare event can greatly affect the mean and standard deviation and the problem is not truly converged. If some regions of the phase space are left un-sampled, the mean and standard deviation are unlikely to converge to the expectation value of the problem. Additional statistical analyses are used in combination to mitigate the problem, see reference [64, pp. 35-39] for details on the ten standard statistical tests within MCNP.

The figure of merit (FoM) is used to study how well behaved a tally result is. This is calculated from Equation 2.4.16, where R is the relative error on the tally and T is the computational time.

$$FoM = \frac{1}{\sqrt{R_{\overline{x}}^2.T}} \tag{2.4.16}$$

This should tend to a constant value with increasing particle histories as $R_{\overline{x}} \propto 1/\sqrt{N}$ and $T \propto N$. The figure of merit measures the cost of error reduction in terms of computational

time and can be used to consider the efficiency of the computation. If the relative error is not significantly reduced with increasing particle histories, then the problem has converged (subject to agreement from the other statistical parameters) and no further histories are required.

To improve the figure of merit, and the overall convergence of the simulation, the relative error on tallies must be reduced without significantly increasing the computational time. Running more histories is computationally inefficient for an already expensive method, so a non-analogue approach to particle transport can be employed to reduce the standard deviation about the mean, and hence the relative error. Variance reduction methods are introduced to boost particle populations within the regions that are important to the tally by limiting the phase space of the problem. To increase particle numbers in the region of interest is not physical, as in the real world particle numbers are always conserved. To avoid this lack of physicality, all particles are assigned a statistical weight; this is the conserved property in non-analogue Monte-Carlo, instead of the absolute particle numbers. Full descriptions of the available variance reduction methods can be found in references [65, pp. 83-135], [39, pp. 327-339] and many other Monte-Carlo texts. The simplest variance reduction method involves energy or weight cut-offs. These terminate particles that are below the energy or weight, which would not significantly contribute to the tally. To preserve the weight, this is combined with Russian rouletting. Each particle below the weight or energy limit will be terminated based on a statistical game. The survival weight is calculated, and a random number sampled, if the survival weight is lower than the sampled number, the particle is killed. Otherwise the particle is allowed to continue until the next event, with a proportionally larger weight.

Another example of variance reduction is implicit capture. This adjusts the statistical weight of the interacting particle, rather than losing it to a capture reaction. In analogue radiation transport a neutron captured by the interacting nuclide is removed from the

system. The particle energy and weight is deposited at the point of capture and it can no longer contribute to the remainder of the problem. Alternatively, implicit capture reduces the particle weight based on the probability of capture with respect to the total microscopic interaction cross-section. The statistical weight is adjusted by $1 - \sigma_{capture}/\sigma_{total}$ for each collision, but the particle is allowed to continue transport within the system.

For models with dimensions that span over many mean free paths, importance mapping is used to encourage higher particle populations in the regions of interest. If a particle moves from a region of lower importance to a region of higher importance, the particle is split into a number of particles, corresponding to the difference in importance between regions. Each split particle carries a fraction of the original particle weight. For example, if a neutron, with weight of 1.0, moves from a region of importance 1 to a region of importance 2, it will split into two identical neutrons, each carrying a weight of 0.5. This is useful for regions that have a high interaction cross-section, such as shielding materials, as it forces an increased particle population throughout the volume. If particles are moving from a region of high importance to low importance, Russian roulette is played on the particle to determine whether it is allowed to continue or if it is terminated.

This section has presented the theoretical background to neutron interactions and their relation to nuclear data, and the computational methods which implement these data to determine the average particle behaviour within complex systems. This knowledge forms the basis of the tools developed to analyse the effect of the pre-processing of nuclear data on simulation results.

Chapter 3

Method

This Chapter covers the various methods and tools implemented within this work. It covers the ENDF format, and the embedded structures for the sections of interest. This is important to understand how to extract and parse the required data prior to processing. The various processing techniques are defined for current systems and additional schema developed as part of this project. The consequences of these various methods will be quantified with several tiers of analysis. At the lowest level, the original ENDF data are evaluated alongside the processed data and the goodness-of-fit measured. The various data are then implemented within a simple Monte-Carlo code, developed in-house to study the different sampling schema. A more complex (yet still simpler than full radiation transport codes) version of this code allows restricted, analogue transport of neutrons in standardised geometries. These simulations are used to benchmark the code against the industry standard MCNP and quantify the computational burden associated with the different data formats. The real-world effects of pre-processing angular distributions are investigated with MCNP fusion models, where only specific data formats are allowed. This information allows for the quantification of the sensitivity of large-scale neutronics simulations to the pre-processing of nuclear data, and the consequences of format choice on other computational parameters.

3.1 ENDF format

The ENDF files are used to store documented data evaluations in a computer-readable format that can then be processed as necessary for a particular application in nuclear technology (described in Section 2.2). The ENDF format was originally developed in 1966 due to a need for a simple, consistent format for the nuclear data files produced by different authorities [66]. Previously, each operating group had developed data sets that were specific to their own requirements. Unfortunately this limited the extent to which the data could be tested, as alternatives would be incompatible with the system. This new format was rigorous in the mathematical techniques and the interpolation methods used to ensure well-defined and repeatable results [50, p. 365]. The ENDF format was adopted within the nuclear technology community to use the same original data files, processed into application specific format for cross-comparison. There has been little change to these files and their structure since the adoption of the ENDF format internationally.

The radiation transport code MCNP requires ACE format data, the ENDF is converted to the accepted format with the pre-processing code NJOY. The general flow of NJOY processing of ENDF data files is well known and is documented in Section 2.2. It is common for NJOY users to black-box each module, resulting in potentially incorrect or unknown behaviour within the processed files. If the user is unaware of the different processes occurring within the NJOY modules, the consequences can be serious. Understanding the original data, before it is degraded with various processing mechanisms is vital. In the following sections, the behaviour of the ACER module of the NJOY code is replicated as far as possible. This allows the quantification of the effect of ACER processing on the differential and double differential cross-section distributions compared to the originals. By understanding the effect of current pre-processing methods, it is possible to then test alternative structures to improve the physical representation of the data and minimise

the computational burden. To replicate these ACER components, the data must first be extracted from the ENDF files, which in turn requires a detailed knowledge of the structures within these files. This section is kept as general purpose as possible, but the ^{56}Fe(n,el) reaction channel at 14.1 MeV is used to demonstrate the various techniques as appropriate.

3.1.1 Structure of ENDF tape

The ENDF files are strictly formatted with the ENDF-6 format, which has been adopted internationally to allow the ease of data transfer and use between agencies. The file, or tape as it is called historically, is a general-purpose ASCII file to allow machine independent data to be transferred. The interaction data are stored in a hierarchical structure, starting with the tape. The tape contains the materials (MAT), though typically these are individual to each nuclide. Each material contains files (MF), which determine the type of data being presented. Each file contains sections (MT), which usually determine the reaction type. Finally each file contains records (MR) that contain the information associated with the MAT, MF and MT. This can be seen visually in Figure 3.1. This nested structure is strictly conformed to by all ENDF data files to aid in the further, consistent processing of data.

The tape is delivered as a series of 80 character long lines in ASCII format. The first 66 characters of each line are dedicated to storing the record information, the next 4 characters contain the MAT number, the next 2 contain the MF number, the next 3 contain the MT number and the final 5 contain the record number, MR. While this format is easy to parse and process, there are limitations to the format that are now becoming an issue. Largely this is due to the restrictive width of the files which is a relic of the punchcard history. Each record can contain 6x11 character wide entries, which limits the precision of the data. Standard scientific notation is inefficient, and as such

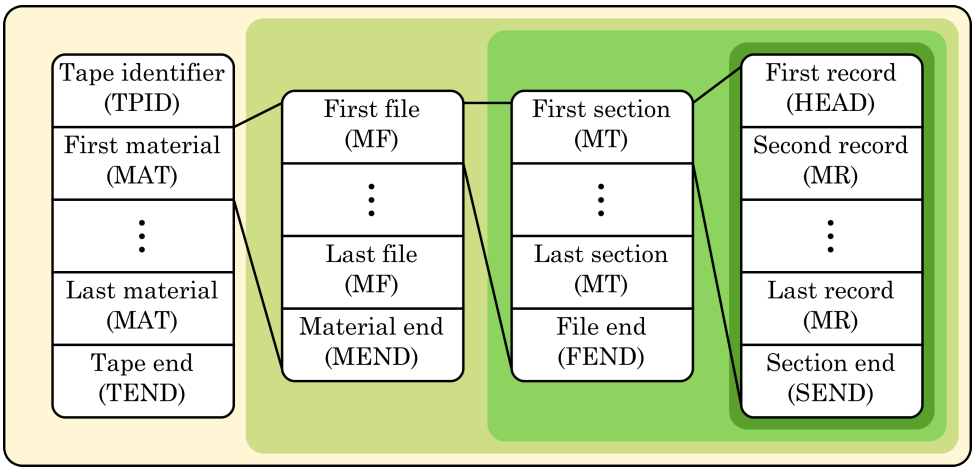

Figure 3.1: Schematic of the nested ENDF tape structure (adapted from [50, p. 28]). Each tape starts with an initial record for the identifier (TPID). The remainder of the tape is divided into materials (MAT). These materials can correspond to different nuclear isomers of one nuclide, or a range of nuclides depending on the evaluation. Each material contains multiple files (MF) for the different classes of information, e.g. cross-section, covariance, angular distributions etc. Each file is separated into sections (MT) corresponding to the various reaction types. The files are further divided into individual records (MR) which contain the evaluated parameters. The first record in a section (HEAD) provides information on how the following data are presented, i.e. number of entries, interpolation laws etc. At the end of each item (TAPE, MAT, MF or MT) there is an associated termination card (TEND, MEND, FEND or SEND).

the data are presented as a stripped down notation. This allows a maximum of 6 or 7 digit precision, depending on the value of the exponent; for example $\pm 1.23456 \pm nn$ or $\pm 1.234567 \pm n$. Full scientific notation can provide a maximum of 5 digit precision ($\pm 1.2345E \pm nn$), given the fusion energies cover up to 7 orders of magnitude, and the cross-sections over 5, this is ultimately a limiting factor in precision [4].

The MAT number identifies the target nuclide and its various isomer states. It is calculated based on $100Z + I$, where I is the isomer state and is equal to 25 for the lightest isotope in the ground state and increases by 3 for each consecutive isotope. The intermediary numbers are reserved for various excited states. For ^{56}Fe in the ground state the MAT number is 2631.

The MF number identifies the file, of which there are approximately 26 different types, a full list can be seen in the ENDF6 formats manual [50, p. 12]. The MF numbers of interest are shown in Table 3.1, specifically the differential data with respect to angle contained within MF 4 is the primary consideration.

Table 3.1: Partial list of file numbers (MF) and their descriptions.

MF	Description
1	Information and tape description
3	Cross-section data ($\sigma(E)$)
4	Angular distributions ($F(\mu, E)$)
6	Energy-angle distributions ($\sigma(\mu, E, E)$ or $\sigma(E', \mu, E)$)

The correlated energy-angle distributions within MF 6 are also affected by the pre-processing in similar ways but are not considered directly. MF 4 contains the probability distributions associated with the exit angle of the primary or secondary particles. This is referred to as $F(\mu, E)$, where μ is the scattering cosine of angle θ in the centre-of-mass frame and E is the energy of the incident particle. This is normalised as shown in Equation 3.1.1, but must be multiplied by the associated scattering cross-section ($\sigma_s(E)$) given in MF 3 to give the absolute differential cross-section $\sigma(\mu, E)$ in barns per steradian, as given by Equation 3.1.2 [50, p. 103].

$$\int_{-1}^{+1} F(\mu, E) d\mu = 1 \tag{3.1.1}$$

$$\sigma(\mu, E) = \frac{\sigma_s(E)}{2\pi} F(\mu, E) \tag{3.1.2}$$

The MT number identifies the section; this usually determines the interaction type and has approximately 800 options [50, p. 13], though in practice not all exist for each MAT. Here the focus is on scattering data, and the key MT values are given in Table 3.2.

Table 3.2: Partial list of section numbers (MT) and the associated reaction channels.

MT	Description
1	Total
2	Elastic scattering
50-90	Inelastic scattering, discrete levels 1 to 40
91	Inelastic continuum

Chapter 4 will demonstrate that these data are most affected by the processing methods at fusion energies, but it is not limited to these cases.

The MR number identifies the record index within each section. MR = 1 corresponds to the start of a section, and MR = 99999 corresponds to the end of a section. For the ^{56}Fe file from the ENDF/B-VII.1 library there are 168 combinations of MT-MF numbers, summing to 3×10^4 records. This is an average file size. From the same library, ^{3}He has the minimum file size with 8 combinations and 5×10^2 records in total, and ^{235}U has the maximum, with 132 combinations and 7×10^5 records in total.

3.1.2 Cross-section data MF 3

The internal structure of the MF3 file is demonstrated in the file excerpt in Figure 3.2. The records are split into the content and identifiers (MAT, MF, MT and MR), where the content is further divided into 6x11 character data entries. The main body of data is presented as energy (E), cross-section ($\sigma(E)$) pairs, where E is measured in eV and $\sigma(E)$ in barns. The HEAD of this file contains the necessary information for parsing the remaining data, and the FEND card flags the end of the file. The HEAD is always the first record in a section containing the material identifier, ZA, in mass terms $(1000.0 \times Z) + A$ and the atomic weight ratio (AWR) of the material, defined as the mass of the material with respect to that of a neutron. In the example shown in Figure 3.2, these values are 26056 and 55.454 respectively.

Content						MAT	MF	MT	MR
2.605600+4	5.545400+1	0	1	0	0	2631	3	2	1
0.000000+0	0.000000+0	0	0	1	2709	2631	3	2	2
2709	2					2631	3	2	3
1.000000-5	0.000000+0	2.530000-2	0.000000+0	4.000000+5	0.000000+0	2631	3	2	4
4.010000+5	-1.660000-3	4.500000+5	-1.660000-3	4.510000+5	-1.720000-3	2631	3	2	5
			⋮						
1.400000+8	6.242866-1	1.420000+8	6.125033-1	1.440000+8	6.039351-1	2631	3	2	905
1.460000+8	5.977444-1	1.480000+8	5.931886-1	1.500000+8	5.896000-1	2631	3	2	906
0.000000+0	0.000000+0	0	0	0	0	2631	3	0	99999
			⋮						
0.000000+0	0.000000+0	0	0	0	0	2631	0	0	0

Figure 3.2: Excerpt from ^{56}Fe, ENDF/B-VII.1 [67] for the cross-section data (MF3) for elastic scattering (MT2). In reference to Figure 3.1, the first record corresponds to the HEAD of the section (MR1) within the file with an additional two lines of variables which define the format of the data to follow (highlighted in blue). The main body (MR4 to MR906) contains the cross-section data, highlighted in green. The terminator for the section within the file, SEND, is highlighted in orange. The data continues to other MT values, before being terminated with the FEND card, highlighted in red.

The next two records contain the information concerning the reaction kinematics and the format of the subsequent data set. The first value of MR 2 is the mass-difference Q-value (QM) given in units of eV. This is defined by Equation 3.1.3 for a reaction of type $a + A \rightarrow b + c + ... + B$, where the m values are the masses associated with the reaction components in amu.

$$QM = [(m_a + m_A) - (m_b + m_c + ... + m_B)] \times (eV/amu) \tag{3.1.3}$$

The second entry is the reaction Q-value (QI) in eV. For ground state reactions with no complex breakup $QI = QM$. As Figure 3.2 shows the data for the elastic scatter reaction channel, these are both equal to zero. This is not the case for any of the inelastic or absorption reaction channels. The next entries of note are the final two on MR 2, and the first two on MR 3. These describe the format of the following data. Entry 5 on MR 2 is the number of different interpolation laws used within the data set, followed by the total

number of pairs. The entries on MR 3 are the number of pairs within the first and only (in this case) interpolation region, and the interpolation law associated with these pairs. Law 2 is a linear-linear interpolation. The available interpolation schemes are shown in Table 3.3 and are presented schematically in Figure 3.3.

Table 3.3: Available interpolation laws in ENDF format. Variables A and B are used as interpolation constants within the equations. These are calculated based on the (x, y) points being interpolated and are shown schematically in Figure 3.3.

Identifier	Law	Description	Equation
1	Histogram	y constant with x	$y = A$
2	Linear-linear	y linear with x	$y = Ax + B$
3	Linear-log	y linear with $\ln(x)$	$y = A\ln(x) + B$
4	Log-linear	$\ln(y)$ linear with x	$y = \exp(Ax)\exp(B)$
5	Log-log	$\ln(y)$ linear with $\ln(x)$	$y = \exp(B)x^A$

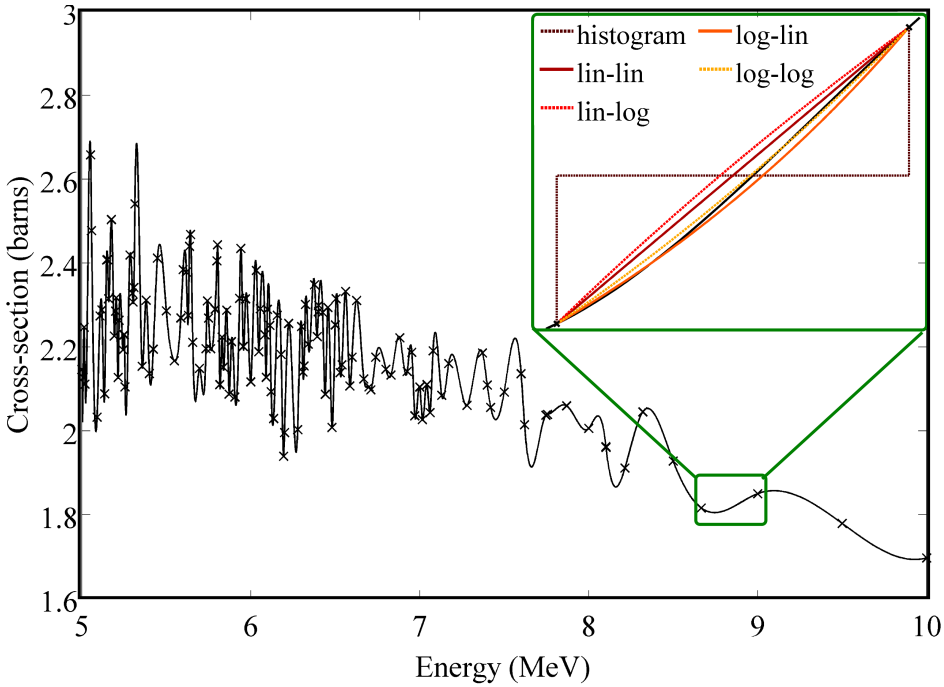

Figure 3.3: Elastic scattering cross-section data taken from ^{56}Fe, ENDF/B-VII.1 [67]. Figure shows the cross-section, $\sigma(E)$ measured in *barns* at the tabulated energy, E points in MeV. The data points are interpolated with one of the five possible laws (Table 3.3). The inset figure schematically shows the different interpolation laws. The choice of law is important to ensure the physical representation of the data is not lost [68].

3.1.3 Angular distributions MF 4

The angular distributions for incident neutron reactions are stored within MF 4. These distributions describe the probability with which emitted particles exit the interaction at a certain angle. Predominantly this is used for the elastic and the discrete level inelastic scattering reactions, where the kinematics of the reaction are easily derived from the exit angle or reaction Q-value. Where the exit angle and energy are strongly correlated, MF 6 is used in place of MF 4, as it provides a two-dimensional data set of angular distributions at each tabulated exit energy (or vice-versa). Figure 3.4 shows schematically how the data are presented in ENDF for the example of ^{56}Fe elastic scattering (MT 2).

Content						MAT	MF	MT	MR
2.605600+4	5.545400+1	0	3	0	0	2631	4	2	1
0.000000+0	5.545400+1	0	2	0	0	2631	4	2	2
0.000000+0	0.000000+0	0	0	1	375	2631	4	2	3
375	2					2631	4	2	4
0.000000+0	1.000000-5	0	0	2	0	2631	4	2	5
0.000000+0	0.000000+0					2631	4	2	6
0.000000+0	4.581000+4	0	0	4	0	2631	4	2	7
4.000000-2	1.200000-2	2.400000-3	3.400000-3			2631	4	2	8
		•							
		•							
		•							
0.000000+0	1.500000+8	0	0	1	73	2631	4	2	1433
73	2					2631	4	2	1434
-1.000000+0	1.10020-10	-9.990482-1	1.10020-10	-9.961947-1	7.855339-6	2631	4	2	1435
-9.914449-1	1.10020-10	-9.848078-1	7.999510-7	-9.762960-1	4.033438-6	2631	4	2	1436
		•							
		•							
		•							
9.914449-1	4.201373+1	9.961947-1	7.217055+1	9.990482-1	9.888429+1	2631	4	2	1458
1.000000+0	1.096701+2					2631	4	2	1459
0.000000+0	0.000000+0	0	0	0	0	2631	4	0	99999
		•							
		•							
		•							
0.000000+0	0.000000+0	0	0	0	0	2631	0	0	0

Figure 3.4: Excerpt from ^{56}Fe, ENDF/B-VII.1 for the angular distributions (MF 4) for elastic scattering (MT 2). MR 1-3 contain the HEAD of the section and an additional three lines of variables which define the format of the data to follow (highlighted in blue). The main body (green) is split into two sections, the angular distributions in terms of Legendre coefficients or tabulated (μ, $F(\mu, E)$) pairs. MR 5-6 shows the first entry at 10^{-5} eV described by two Legendre coefficients. MR 1433-1459 contains the tabulated data associated with an energy of 150 MeV. The section is terminated with the SEND card (orange). The data continues to other MT values, before the file is terminated with the FEND card (red).

This follows a similar format to the MF3 data in the sense of the first few records contain the information on how the data are to be interpreted, the actual data forms the central block and the section of the file is terminated with a SEND card. As above the first two entries on the HEAD record are the ZA and AWR numbers. The fourth entry in this record is the LTT value; this flag specifies the representation of the data. A value of 0 corresponds to all distributions being isotropic, a value of 1 corresponds to all distributions being described by Legendre coefficients and a value of 2 corresponds to all distributions being described by tabulated probability distributions of μ and $F(\mu, E)$. An LTT of 3, as shown in Figure 3.4, means that a combination of representations are used; at low energy, the distributions are described by Legendre coefficients and at high energy they are described with tabulated data. For elastic and discrete level inelastic scattering data, Legendre coefficients are preferred [50, p. 108]. Tabulated probabilities are only used when the distributions cannot be sufficiently described with a polynomial of the maximum allowed order (65) or below with non-negative probabilities.

Within the second record (MR 2) the fourth entry (LCT) describes the reference frame. A value of 1 is for lab frame data and a value of 2, as seen in the example, is for centre-of-mass frame data. The following two records mirror the format of MR2 and MR3 in the cross-section data. These indicate that there is one energy range, of which there are a total of 375 energy values. All of these 375 values are interpolated in energy with law 2, linear. The first distribution entry starts with MR5, this gives the temperature of the evaluation, T, in K and the energy, E, of the neutron in eV. In the example shown, the data are evaluated at $0K$ for energy 10^{-5}eV. As the data must cover the full phase space of particle transport, almost all data begins at 10^{-5}eV. As the lower energy entries are given in terms of Legendre coefficients, entry 5 on this record gives the highest order coefficient (NL) as well as the number of entries to read. It is always assumed that the zeroth component $a_0 = 1.0$ so is not included in the file, so the two values for a_1 and a_2 are given. Due

to the low energy, both coefficients are zero; this corresponds to isotropic scattering in the centre-of-mass frame. The remaining data in this regime follows the same pattern. At MR 1433 in this example, we see the transition to tabulated probability distributions. The temperature and energy are given as the first two entries. The remaining entries correspond to the number of regions within the tabulated data (1) and the total number of points (73). The subsequent record describes the number of points (73) governed by the interpolation law (2). The next record begins the tabulated (μ, $F(\mu, E)$) pairs. Finally the section of the file is terminated with the SEND card. To fully appreciate the quantity and complexity of the information stored within this small section of ENDF, Figure 3.5 shows the two-dimensional phase space covered by the Legendre coefficients.

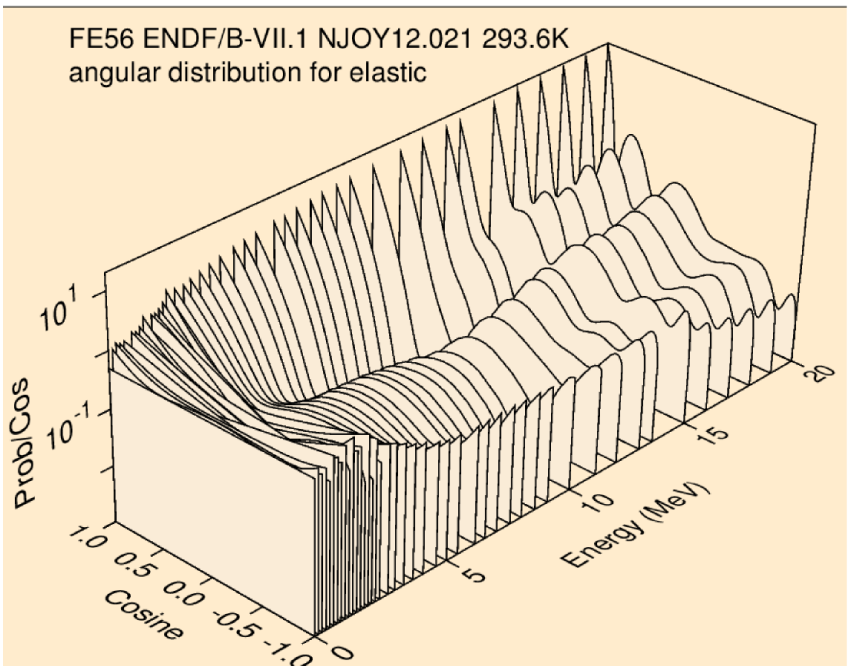

Figure 3.5: Elastic scattering data for ^{56}Fe, ENDF/B-VII.1 as plotted with NJOY2012. These data correspond to the Legendre coefficient portion of MF 4 MT 2 only, with energies between 10^{-5}eV up to 20MeV. The image gives the scattering probability ($F(\mu, E)$ or $Prob/Cos$) in terms of the scattering cosine (μ or $Cosine$) for each energy entry. Low energy corresponds to a relatively isotropic scattering probability in the centre-of-mass frame. At higher energies more features appear and a more anisotropic distribution results, i.e. forward scatter ($\mu = 1.0$) is more favoured than back scatter.

As mentioned in Section 3.1.1, these angular probability distributions, $F(\mu, E)$ are normalised within the μ range of $-1.0 \leqslant \mu \leqslant 1.0$, and are given in units of $(unit - cosine)^{-1}$. When described by Legendre coefficients, the functional form of the distribution is constructed by Equation 3.1.4, where μ is the scattered cosine, E is the incident energy, l is the order of the Legendre polynomial $P_l(\mu)$ up to NL and a_l is the l^{th} coefficient.

$$F(\mu, E) = \sum_{l=0}^{NL} \frac{2l+1}{2} a_l(E) P_l(\mu) \tag{3.1.4}$$

3.2 File handling

The previous Section, 3.1, described some of the intricacies of the ENDF format. This knowledge is used to parse and manipulate the data as required by the various pre-processing codes, such as NJOY. In order to understand the effect of NJOY on the data and further reduce it, parallel methods have been developed to study and quantify the effect of various techniques. These quantification methods will be described in Section 3.3, and this section is dedicated to the file handling and how the data are further manipulated. NJOY processes the ENDF into MCNP compatible ACE format with one of two data representations for the angular distributions. The equal-probability histogram, with a fixed 32 bins, or the tabulated cumulative probability distributions with linear interpolation. The former is a legacy format designed for times when computational resources were at a premium. It provides a compact data set that is very fast to sample at the cost of accurate data representation, particularly at fusion energies and higher. This is demonstrated in Chapter 4. The latter format can provide a better representation at the cost of computational efficiency. This is now the default method for any new data libraries, but many older libraries have been processed into the histogram format and are still in use.

Due to the strict formats required by MCNP it is impossible to trial variations or alternative formats other than the two specified; hence the need for a simple Monte-Carlo code. Turnip [1] is a Monte-Carlo code developed as part of this work for nuclear data sampling analyses. It is capable of replicating MCNP sampling schemes and implementing alternatives. Here it is used to evaluate different formats for angular distributions, but its potential application extends far beyond this. The development and processes of Turnip will be detailed in Section 3.3.1. The ENDF files are parsed and processed into Turnip compatible format using a code called Swede-Mash[2]. This has been designed to replicate NJOY processes for the two known MCNP formats and produces additional formats as required. Swede-Mash is constructed of three core modules, one to parse, one to process and one to analyse. Figure 3.6 shows the generalised flow diagram for Swede-Mash, and Sections 3.2.1, 3.2.2 and 3.2.3 will describe the three stages in detail.

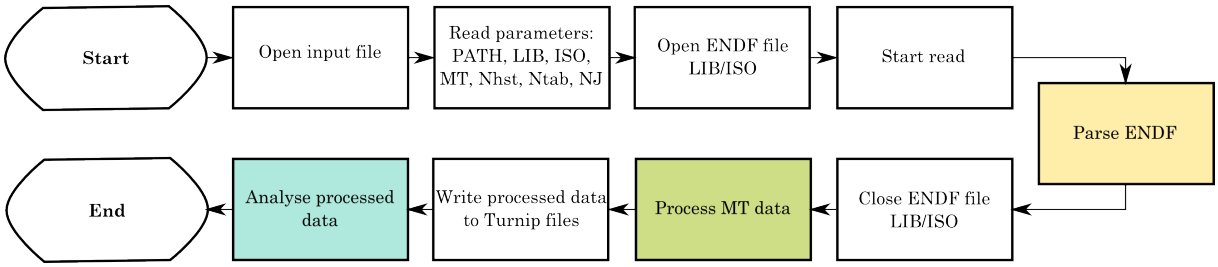

Figure 3.6: Global flow diagram for the Swede-Mash code. The main modules for processing, parsing and analysing are highlighted in yellow, green and blue. These three sections are detailed in sections 3.2.1, 3.2.2 and 3.2.3. The variables $PATH$, LIB, ISO, MT, $Nhst$, $Ntab$, and NJ correspond, respectively, to the home path of the source and input/output files, the library directory, the ENDF isotope to be parsed/processed, the reaction channel of interest, the number of histogram bins, the number of tabulated points and finally whether an NJOY input file is required.

[1]Note to the reader, Turnip is not an acronym and this code is not currently available to the public. Turnip was developed in collaboration with (T.W. Leadbeater, private communication) and is intended for future release.

[2]Note to the reader, Swede-Mash is not an acronym either. This was developed by the author as part of this thesis and is not publicly available at present, though future release is expected.

The required parameters for the parse/process are input via a basic text file. The structure is simple, each consecutive line is a different parameter, and any line starting with a # is considered as a comment. A basic input file is shown in Listing 3.1.

Listing 3.1: Example input file for Swede-Mash

```
# Path to home (PATH)
/Users/work/Documents/swede_mash
# library directory (LIB)
endfb_71
# Isotope to parse (ISO)
Fe56
# Reaction channel to parse (MT)
2
# Formats to process
# Number of EP bins (Nhst)
32
# Number of tabulated points (Ntab)
# min 2, or 0 for optimised
0
# NJOY (NJ)
# Produce NJOY input files for processing 0 = no, 1 = yes
0
```

The first parameter tells Swede-Mash where to find the source and is the home directory for any input or output files. The second parameter is the directory within the home path that contains the ENDF library. The next two entries correspond to the isotope and reaction channel of interest, in this example these are ^{56}Fe and the elastic scattering data. All associated MF data are extracted (cross-sections, differential data etc.) The following two entries determine the number of equal-probability histogram bins (the MCNP requirement is 32) and the number of tabulated points. If either of these values are set to 0, they are optimised based on a hard-coded level of accuracy. This is set to be a maximum of 1.0% variation between the original and processed formats. The final parameter is a flag to determine whether an NJOY input file is required for the given isotope.

3.2.1 Parse ENDF

The parsing module shown in Figure 3.6 has been designed to parse only the files of interest to this work as detailed in Table 3.1. Figure 3.7 shows a more detailed program flow within this module for MF 3 and 4. Swede-Mash is capable of parsing MF 5 and 6, but the methods are very similar to those used for MF 4 and are not shown. The complexity in parsing the ENDF is predominantly due to the ASCII format; while this is well suited to FORTRAN based interpreters it requires a different approach when using any other programming language, such as C. The file is parsed character by character for each line, with the relevant content being stored in character arrays before further parsing. As the format of each line within each section of file is dependent on the previous, a series of flags and switches were implemented to retain information on the section header parameters and what the next process should be. For data (versus information) containing entries the split content is then converted into floats and integers as required rather than the compact ENDF scientific notation. Each unique MT-MF combination associated with the target is parsed and output to temporary binary files to be used in the processing stages of Swede-Mash. So for the example input file shown in Listing 3.1, the data would be fully parsed for MF 3, MF 4, MF 5 and MF 6 if they exist in the file. Other associated files with the MT value such as information (MF 1) and covariances (MF 33) are extracted, but no further parsing is implemented for these files at this point.

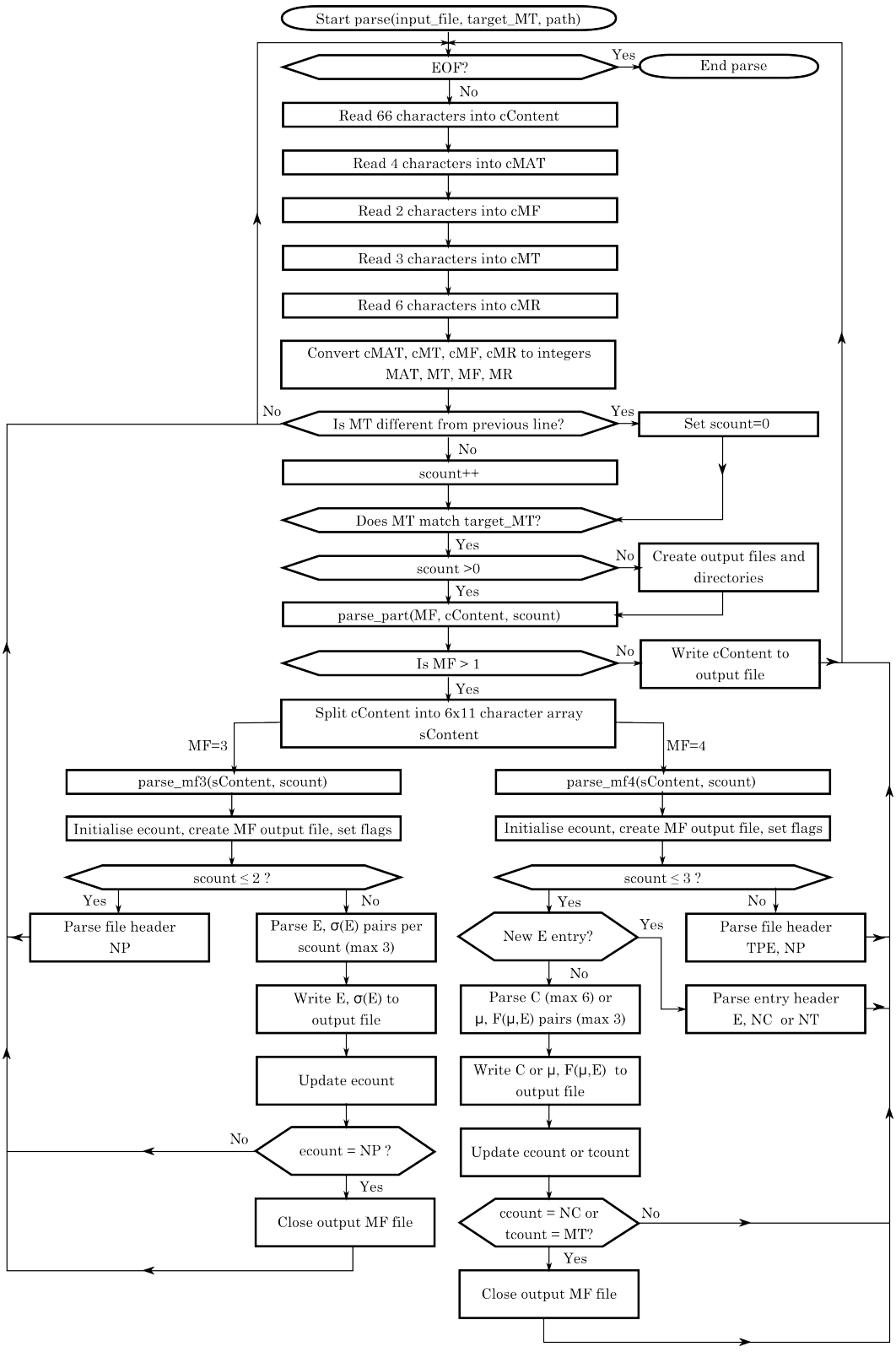

Figure 3.7: Simplified flow diagram for the parse module of Swede-Mash. ENDF input file is passed to this module with the target MT and home path. The ASCII file is read character by character and is assigned to the character arrays $cContent$, $cMAT$, cMF, cMT, cMR. When the MT value changes from the previous line it is checked against the target. If this is the required section, the content is parsed according to its file MF number. The three shown are for the information (1), cross-section (3) and angular data (4) files. Counters *scount*, *ecount*, *ccount*, *tcount* correspond to the line in section, energy, coefficient and tabulated points, and are used in combination with other flags to switch between processes.

3.2.2 Process ENDF

After the entire ENDF has been parsed and the relevant MT-MF combinations converted into a more accessible format, the various processing methods can be applied to the differential data. The flow diagram for this module can be seen in Figure 3.8. The case shown here applies to the processing of the angular distributions (MT 4). It is dependent on the initial input deck, namely the requested number of histogram bins and tabulated points. The processing treatment is applied to the differential data for all energy entries described by Legendre polynomials. The three formats presented in the following section are described by the computational methods used to produce the required data; these will be shown graphically in Section 4.1 for the example case of ^{56}Fe for clarity.

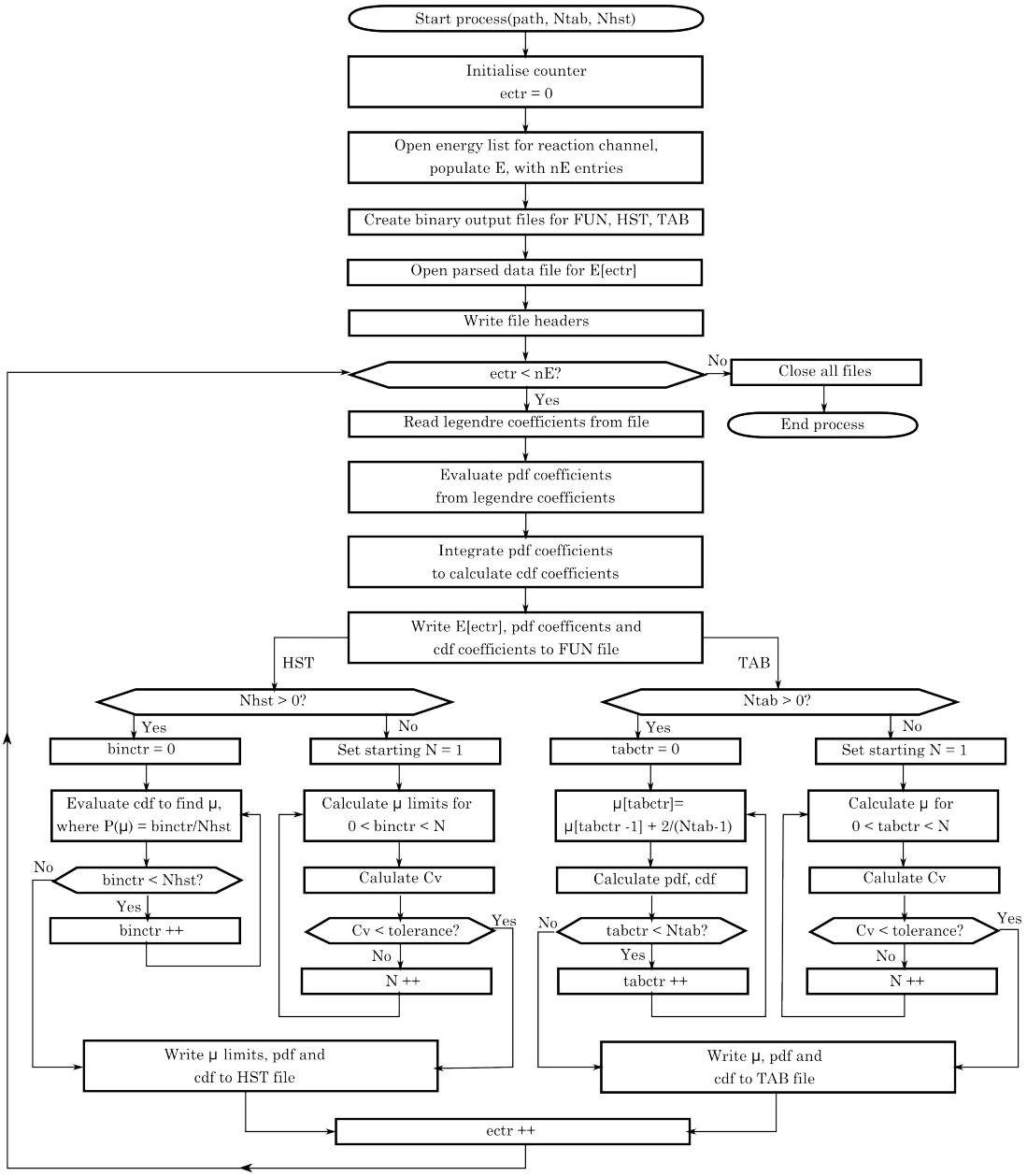

Figure 3.8: Flow diagram for the process module of Swede-Mash. This takes the parsed ENDF files and for each energy associated with the angular distributions, they are processed into the three formats; functional (FUN), histogram (HST) and tabulated (TAB) and written to Turnip binary files. These formats are described in detail within the text. Initially the Legendre coefficients are converted to plain polynomial coefficients, for both the normalised probability (pdf) and cumulative probability (cdf) distributions. These distributions are then used to calculate the histogram and tabulated formats, dependent on the number of bins/channels requested. For these two formats, an optimised option is available if $Nhst$ and/or $Ntab$ are equal to zero. This determines the optimum number of bins/points to achieve an accuracy below the set tolerance and is described within the text.

Functional

The first stage produces functional data that replicates the original Legendre polynomials and is referred to as the exact or original data. While this format is not strictly exact, it is the best possible case when processing the evaluated data, and is the standard to which all processed formats are compared. This format is not currently employed within any of the standard radiation transport codes. The coefficients are converted into standard polynomial coefficients, i.e. for each order the coefficients are summed. This representation is identical (subject to rounding errors and polynomial evaluation [69, p. 253]) to the original, but simplifies the numerical integration required throughout the processing module and is compatible with Turnip. Equation 3.1.4 shows how the ENDF probability distributions are constructed, whereas Equation 3.2.1 shows how the converted coefficients are used. Where p is the order of the polynomial from 0 to P and a_p is the p^{th} coefficient.

$$F(\mu, E) = \sum_{p=0}^{P} a_p.\mu^p \tag{3.2.1}$$

These data are written to the Turnip binary files with the general format as shown in Listing 3.2, both the normalised probability C_f and cumulative probability C_p coefficients are given.

Listing 3.2: Representation of functional data output format for Turnip

```
************************************************
Comment + Header [500 characters]
example file ZZAAA_MF_MT_FUN.bin
NE = number of energies
E(0),E(1)...E(NE-1) = energy values in eV
NP = number of coefficients
Cf = polynomial coefficient for orders 0 up to NP-1
Cp = polynomial coefficient for orders 1 up to NP
************************************************
NE
E(0) NP
Cf(0) Cp(1)
Cf(1) Cp(2)
...   ...
Cf(NP-1) Cp(NP)
...
E(NE-1) NP
Cf(0) Cp(1)
Cf(1) Cp(2)
...   ...
Cf(NP-1) Cp(NP)
```

Histogram

This simplified functional form, $F(\mu, E)$, is used to calculate the remaining formatted data, either in histogram or tabulated form. The equal probability histogram is calculated using Equation 3.2.2, where k is the bin number and K is the total number of bins required.

$$\int_{-1.0}^{\mu_k} F(\mu, E)d\mu = \frac{k}{K} \tag{3.2.2}$$

This histogram format can be represented as a piece-wise polynomial as given by Equation 3.2.3, where the c values correspond to the height of each bin to give an area of $1/K$. In this way, an arbitrary number of histogram bins can be systematically investigated in comparison to the MCNP specific 32 bin histogram.

$$H(\mu, E) = \begin{cases} c_0 & : -1.0 \leqslant \mu \leqslant \mu_1 \\ \ldots & \\ c_{K-1} & : \mu_{K-1} \leqslant \mu \leqslant 1.0 \end{cases} \tag{3.2.3}$$

The limits of each bin in μ are calculated based on an iterative solution to the integral of $F(\mu, E)$. This method only works due to the normalisation of the distributions and measures have been taken to ensure that this is preserved. Listing 3.3 depicts the format of these data as output to Turnip. For each energy entry, the full μ range must be accounted for; the number of bins as determined by NB are each associated with an upper and lower limit, so $MUL(0)$ will always equal -1.0 and $MUU(NB)$ will always be 1.0. Each bin is presented along with their normalised and cumulative probabilities. The histogram format is widely accepted by radiation transport codes, though the number of bins and the strictness varies; as previously mentioned MCNP will only accept 32 bins.

Listing 3.3: Representation of histogram data format for Turnip

```
************************************************
Comment + Header [500 characters]
example file ZZAAA_MF_MT_HST.bin
NE = number of energies
E(0),E(1)...E(NE-1) = energy values in eV
NB = number of histogram bins associated with E
MUL, MUU = lower/upper mu limits for binned data
PDF = histogram value of PDF
CDF = histogram value of CDF
************************************************
NE
E(0) NB
MUL(0)  MUU(0)  PDF(0)  CDF(0)
MUL(1)  MUU(1)  PDF(1)  CDF(1)
...     ...     ...     ...
MUL(NB) MUU(NB) PDF(NB) CDF(NB)
...
E(NE-1) NB
MUL(0)  MUU(0)  PDF(0)  CDF(0)
MUL(1)  MUU(1)  PDF(1)  CDF(1)
...     ...     ...     ...
MUL(NB) MUU(NB) PDF(NB) CDF(NB)
```

Tabulated

The final format considered is the tabulated cumulative distribution function, with linear interpolation between points. The user will input the required number of data points, t, these are then equally spaced in angle, θ, before being converted into cosines, μ. At each value of μ, the functional form is evaluated. These points are then linearly interpolated to produce a continuous distribution. Equation 3.2.4 shows the basic construction of this distribution.

$$T(\mu, E) = \begin{cases} T(-1.0) + (\mu + 1.0)\dfrac{(T(\mu_1) - T(-1.0))}{(\mu_1 + 1.0)} & : -1.0 \leqslant \mu \leqslant \mu_1 \\ \ldots & \\ T(\mu_{t-1}) + (\mu - \mu_{t-1})\dfrac{(T(1.0) - T(\mu_{t-1}))}{(1.0 - \mu_{t-1})} & : \mu_{t-1} \leqslant \mu \leqslant 1.0 \end{cases} \tag{3.2.4}$$

One additional step must be undertaken to preserve the normalisation required for the sampling. The total area is calculated simply with the trapezium rule, if this differs from the required value of 1.0, the height of all points are divided through by the area. The data are then output to the Turnip files, with the representation shown in Listing 3.4.

Listing 3.4: Representation of tabulated data output format for Turnip

```
************************************************
Comment + Header [500 characters]
example file ZZAAA_MF_MT_TAB.bin
NE = number of energies
E(0),E(1)...E(NE-1) = energy values in eV
NP = number of data points associated with E
MU = mu value
PDF = value of PDF evaluated at mu
CDF = value of CDF evaluated at mu
************************************************
NE
E(0) NP
MU(0) PDF(0) CDF(0)
MU(1) PDF(1) CDF(1)
...    ...     ...
MU(NP) PDF(NP) CDF(NP)
...
E(1) NP
MU(0) PDF(0) CDF(0)
MU(1) PDF(1) CDF(1)
...    ...     ...
MU(NP) PDF(NP) CDF(NP)
```

Optimisation

Swede-Mash has the additional capability to produce histogram and tabulated data that have been optimised with respect to file size, sampling time and accuracy. In a perfect world, it would be possible to have exact representations of differential data, with no loss of computational efficiency. Unfortunately this is not the case, to optimise just one of these parameters is typically at the cost of the others. To optimise based on file size, it is trivial to set the tabulated and histogram formats to have the minimum number of bins/points. In some instances this is not an issue, but with mid to high mass or high energy distributions this degrades the data to the point that they are non-physical. To optimise for sampling time, histogram format should always be used, but this again risks information loss. To optimise for accuracy, the C_v metric is used to quantify accuracy as introduced in Section 2.3. The implementation of the C_v will be described in Section 3.2.3 with respect to the analysis module of Swede-Mash. An arbitrary level of accuracy is hard-coded into Swede-Mash, currently this set to 1.0% to be commensurate with the

expected statistical precision of the Monte-Carlo simulations, but can easily be changed dependent on the level of optimisation required. To optimise the histogram data, a single histogram bin is used initially; the original and processed forms are evaluated on a fine μ grid and the differences used to calculate the C_v. If this is higher than the set level of accuracy, two bins are used and Swede-Mash will recalculate the μ limits and C_v. The process repeats iteratively with an increasing number of bins until the required accuracy is achieved. The data are then output in the same way to the Turnip files. For tabulated data, the method is identical, i.e. an initial number of points is trialled, if the resulting C_v is higher than the set-level, more points are added. This iterative process is much slower to produce the processed files than having a fixed number of bins/points, but the file handling and processing methods remain completely transparent and controllable.

3.2.3 Analyse ENDF

There is one analysis method built in to Swede-Mash that is used throughout the code and further data analyses; the general form of the C_v as given by Equation 2.3.2. In the practical sense, the original, functional form, $F(\mu, E)$ is evaluated for each energy at 10^5 equally spaced points in μ. The processed distributions are evaluated on the same μ grid; the values of $F(\mu, E)$, $H(\mu, E)$ and $T(\mu, E)$ are calculated according to Equations 3.2.1 to 3.2.4 in Section 3.2.2. The number of points was determined by a simple convergence study. Figure 3.9 shows the calculated C_v for the histogram format with respect to the unprocessed format for the elastic scattering reactions in ^{56}Fe and ^{nat}C. The C_v converged within negligible errors by 10^5 points and the spacing in μ is smaller than the calculated bin spacing ensuring all regions are sampled. In ^{56}Fe, the scattering distribution is highly anisotropic, hence C_v is much higher overall. The extreme forward bias means that the bins are very narrow in the forward scattering region and a high sampling rate is required to populate these bins for the C_v calculation and hence reduce the errors. Whereas ^{nat}C

is closely isotropic, so the bins are relatively evenly spaced in μ and a lower sampling rate is required to populate the entire distribution for the C_v calculation. This results in a much lower overall C_v and convergence at just 50 sampled points.

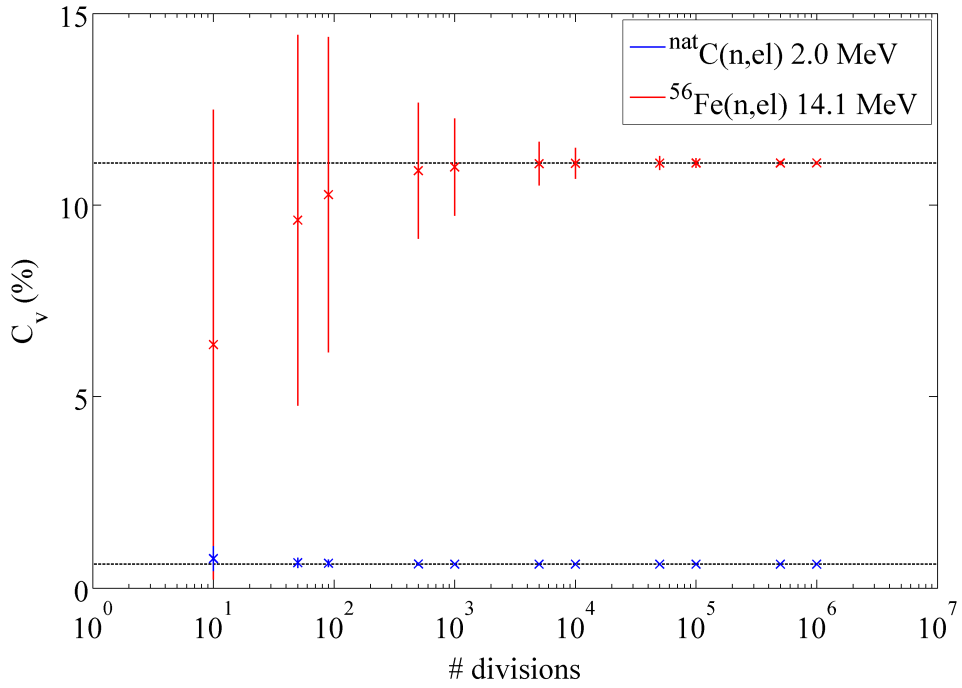

Figure 3.9: Convergence of the coefficient of variation as a function of sampled points. The red points correspond to the ^{56}Fe(n,el) fusion relevant reaction channel at 14.1 MeV, and the blue shows the ^{nat}C(n,el) fission relevant reaction channel at 2.0 MeV. These data are based on the processing of functional data into the 32 bin, equal-probability histogram shown in Figure 2.13.

The error on the coefficient of variation was calculated from the variance of the variance, in this case the variance of the root mean square deviation D_r (or standard deviation, s) with the number of samples. From [70, p. 199], the variance of the variance s^2, is defined by Equation 3.2.5, where N is the number of samples, μ_4 is the fourth central moment of the distribution and μ_2 is the second central moment, or variance.

$$\sigma^2(s^2) = \frac{1}{N}\left(\mu_4 - \frac{N-3}{N-1}\mu_2^2\right) \tag{3.2.5}$$

The kurtosis can be substituted into this equation as the normalised fourth moment as shown in Equation 3.2.6 [71, p. 27]. The kurtosis is a measure of how well the distribution represents the normal; in the case of a perfect normal distribution Equation 3.2.6 is equal to zero.

$$k = \frac{\mu_4}{\mu_2^2} - 3 \tag{3.2.6}$$

This is substituted into Equation 3.2.5, and given that the standard deviation is the square root of the variance the error on the standard deviation can be calculated from this relationship. For large values of N, the error on the standard deviation, $\sigma(s)$ is calculated from Equation 3.2.7.

$$\sigma(s) = s\sqrt{\frac{k+2}{N}} \tag{3.2.7}$$

Sampling at 10^5 intervals results in convergence to the expectation value of the C_v within negligible errors for ^{56}Fe(n,el). Fewer samples are required to achieve the same error for more isotropic systems, as shown for ^{nat}C(n,el), so the number of intervals is deliberately set to be conservative. The C_v will also be used to quantify the differences between Monte-Carlo simulation results. In these cases, assuming a high enough level of sampling, the error on the C_v will be calculated through the propagation of statistical errors as a conservative estimate. Figure 3.10 shows how this analysis module has been implemented within Swede-Mash.

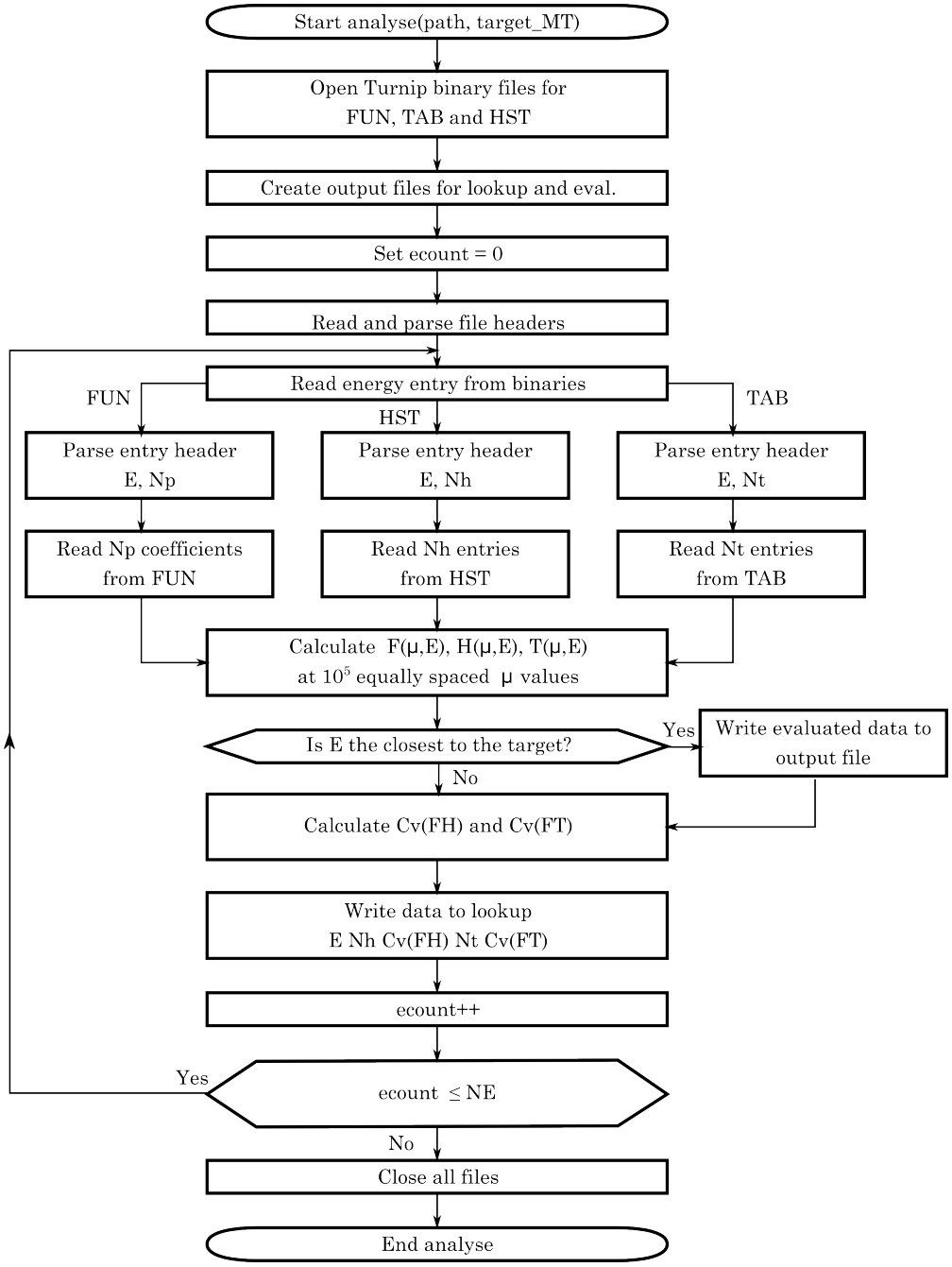

Figure 3.10: Flow diagram for the analyse module of Swede-Mash. Data are read from the processed Turnip binaries one energy entry at a time. The three formats are evaluated on a μ grid, of 10^5 equally spaced points. The C_v is calculated as a comparison between histogram-functional and tabulated-functional. Lookup tables of E, C_v pairs are produced, and full data are output for the entry closest to the target energy.

3.3 Analysis techniques

Section 3.2.3 covered the point-wise analysis for comparing the original ENDF format to the processed variations with respect to accuracy; this is measured with the coefficient of variation, C_v and the maximum difference, Δ_m. When using nuclear data within the Monte-Carlo radiation transport codes there are other considerations such as the computational burden associated with sampling from the different regimes. Simple Monte-Carlo geometries are used to study the propagation of pre-processing differences to macroscopic quantities, such as surface current. The model complexity is increased throughout the analyses to determine where and how, within a full working model, the data format effects global parameters.

Figure 1.6 within Chapter 1 showed the MCNP model of ITER that has been widely used within the fusion community to investigate the safety and efficiency of the proposed design. Due to the symmetry of the tokamak, a 40° segment is modelled to reduce the computational burden of running the full geometry. To populate this model with neutrons, 10^9 source particles are required. This equates to 9.4 days of total computational runtime [72], and 866 MB of nuclear data loaded into memory. To quantify the effect of the data processing, specifically with respect to fusion simulations, isotopes within close proximity to the plasma of the ITER model were analysed. The selected isotopes are listed below, and their purpose within a fusion reactor are given in Table 2.1:

- ^{1}H
- ^{16}O
- ^{52}Cr
- ^{56}Fe
- ^{63}Cu
- ^{90}Zr
- ^{184}W

At fusion energies of 14.1 MeV, scattering reactions are most probable. Figure 3.11 shows the proportion of elastic and inelastic scattering cross sections with respect to the other possible channels.

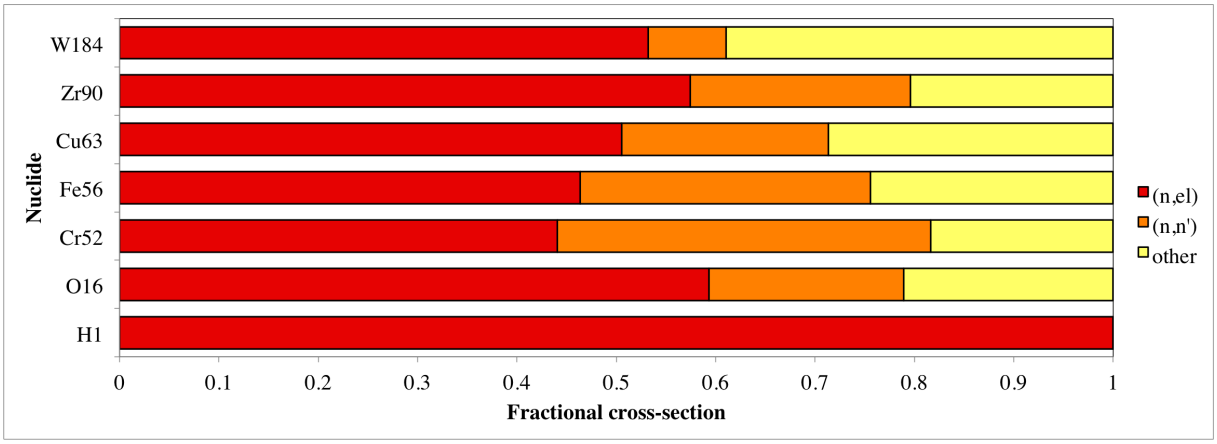

Figure 3.11: Interaction cross-sections shown as a fraction of the total for each fusion relevant isotope at 14.1 MeV. In all cases elastic scattering (red) is the largest contributor to the total. The inelastic scattering data (orange) are the sum of all inelastic levels and the remaining data are the sum of all other contributors.

For the relevant isotopes, Figure 3.12 shows the elastic scattering angular distributions on a polar plot. The elongated distributions correspond to a forward biased, and highly anisotropic distribution typically associated with high mass and high energy. These reactions are the most affected by the processing methods described in Section 3.2.2.

Though the scattering reactions for seven different fusion relevant isotopes are considered as part of this work, the detailed analysis of ^{56}Fe will be used for demonstration purposes in Chapter 4. This is a very well known and well studied isotope; it is a major component of steels in nuclear facilities and has been the subject of many years of experimentation and validation. The nuclear data files produced for ^{56}Fe are considered the best case in terms of original data, i.e. non-traditional materials such as ^{63}Cu are less well determined.

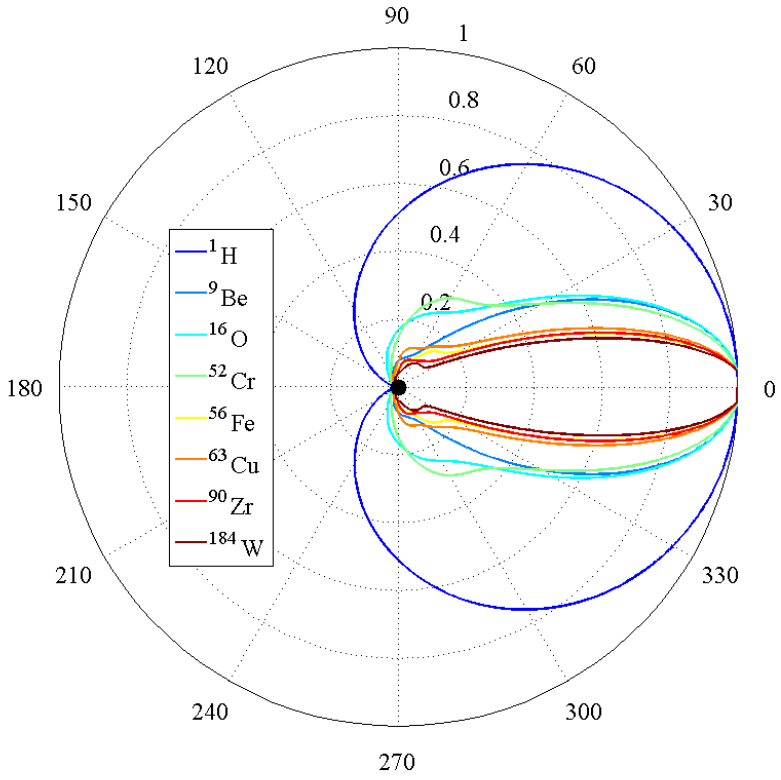

Figure 3.12: Polar plot of cumulative elastic scattering distributions for the fusion relevant isotopes at 14.1 MeV. An elongated distribution, such as ^{184}W implies highly anisotropic scattering; whereas the wider distributions, such as ^{1}H, implies closely isotropic scattering.

3.3.1 Simple Monte-Carlo

Traditional Monte-Carlo radiation transport codes such as MCNP are very strict in their input requirements, and limited in the potential output. It is non-trivial to encourage general purpose codes to provide the necessary information to properly investigate the effects of using different sampling schema for angular distributions. The Turnip Monte-Carlo code was developed as part of this work to study the sampling regimes without the complexities of geometry, variance reduction and transport. Swede-Mash produces processed binary files that are Turnip compatible, from the same ENDF libraries that are used (after processing with NJOY) in MCNP. The sampling mechanisms employed by

MCNP for current formats are replicated within Turnip, whilst allowing for higher levels of flexibility and additional sampling methods such as functional data. Figure 3.13 shows how the functional, histogram and tabulated formats are sampled.

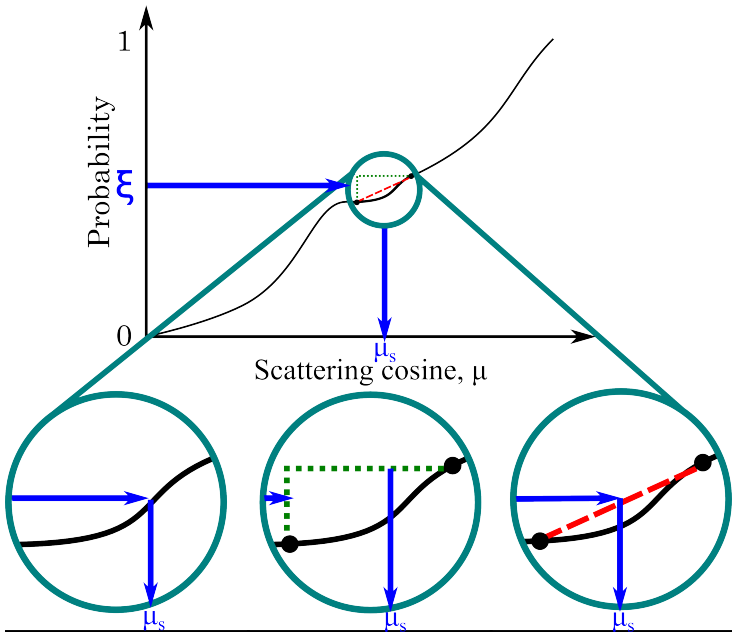

Figure 3.13: Schematic of the three sampling regimes implemented in Turnip. A cumulative probability distribution is shown with respect to scattering cosine, μ. For a random number ξ, sampled uniformly between 0 and 1, the sampled scattered cosine, μ_s is calculated from the cumulative distribution. The three formats, from left to right, are functional (black), histogram (green) and tabulated (red). Further details of the sampling schemes are given in Table 3.4. Figure is adapted from [73].

For the original, or functional data, the solution for μ is found iteratively, within a tolerance of 10^{-5}. The average number of calculations required is equal to $L \log_2 (10^5)$, where L is the order of the polynomial to be calculated. For example, a 4^{th} order polynomial, 67 calculations are required on average and for a 12^{th} order, 200 are required. For histogram data, only two calculations are ever required; one to locate the bin and the other to distribute the value of μ between the bin limits. Tabulated data requires a binary search of the points followed by an additional calculation to distribute μ between the linearly interpolated points. This equates to an average of $\log_2 (M) + 1$ calculations, where M is the number of tabulated points. For 100 tabulated points, this requires an average of 8 calculations. Table 3.4 summarises the three different sampling regimes.

Table 3.4: Summary of sampling schemes of angular data for the equal-probability histogram, tabulated and functional formats with respect to ξ sampled between 0 and 1.

Functional	**Histogram**	**Tabulated**
Binary search of polynomial of order L for solution $P(\mu) = \xi$ within a tolerance of 10^{-5}, where $P(\mu) = \sum_{l=0}^{L} a_l \mu^l$	Bin k of K selected based on $k = \lfloor K\xi \rfloor$. μ calculated by: $\mu = \mu_{k,L} + (K\xi - k)(\mu_{k,U} - \mu_{k,L})$	Binary search for channel m of M such that $P(\mu_m, L) \leqslant \xi < P(\mu_{m,U})$. Linearly interpolate between μ_m, L and μ_m, U to solve for μ.
Average number of calculations to solve: $L \log_2(10^5)$	Absolute number of calculations to solve: 2	Average number of calculations to solve: $\log_2(M) + 1$

The simplified flow diagram for Turnip can be seen in Figure 3.14, the user defines the isotope, reaction channel, sampling type, number of forced events, number of histories and initial neutron energy.

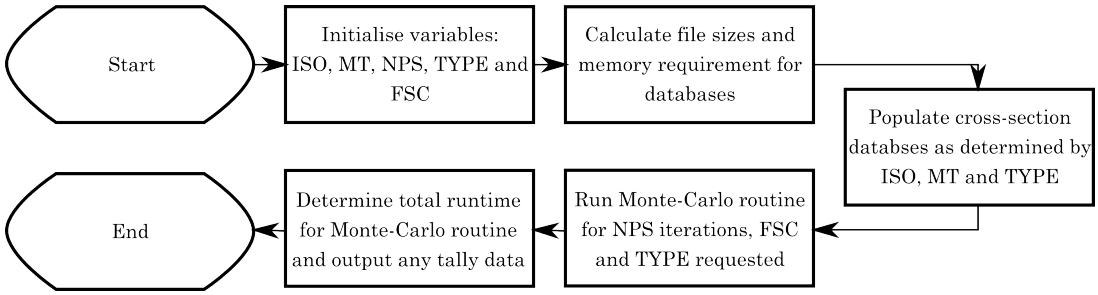

Figure 3.14: Generalised flow diagram for Turnip. The isotope, ISO, reaction channel, MT, sampling type, $TYPE$, and number of forced scatters, FSC are initialised. The databases are built and the Monte-Carlo routine is run for NPS source particles. The resulting tallies for angular and energy distributions are output at the end of the run, along with any computational run data. For Turnip, the Monte-Carlo routine follows the process as shown in Figure 3.15.

For each history, the neutron will undergo the specified number of forced events. As an example, two elastic scatters were forced with ^{56}Fe(n,el) at 14.1 MeV. All neutrons are born with the same energy and are mono-directional, it will sample the scattered cosine from the nearest data set to 14.1 MeV. The energy loss is calculated using Equation 2.1.4 and the direction vector calculated from Equation 2.1.1-2.1.3. The database will be

searched for the next closest energy entry to the updated energy and the scattering cosine sampled again. The final energy and direction are histogrammed for the output. This Monte-Carlo routine, without transport can be seen in Figure 3.15.

Turnip not only outputs the event data, but it also outputs the memory requirements for the files and built database sizes, and the time requirements for the different sections of the code, i.e the database set-up times, and Monte-Carlo loop. With the combination of computational and accuracy metrics, the overall value of each format can be quantified. It should be noted that the computational efficiency, whilst important, is considered to be of lower overall value than the accuracy of data representation.

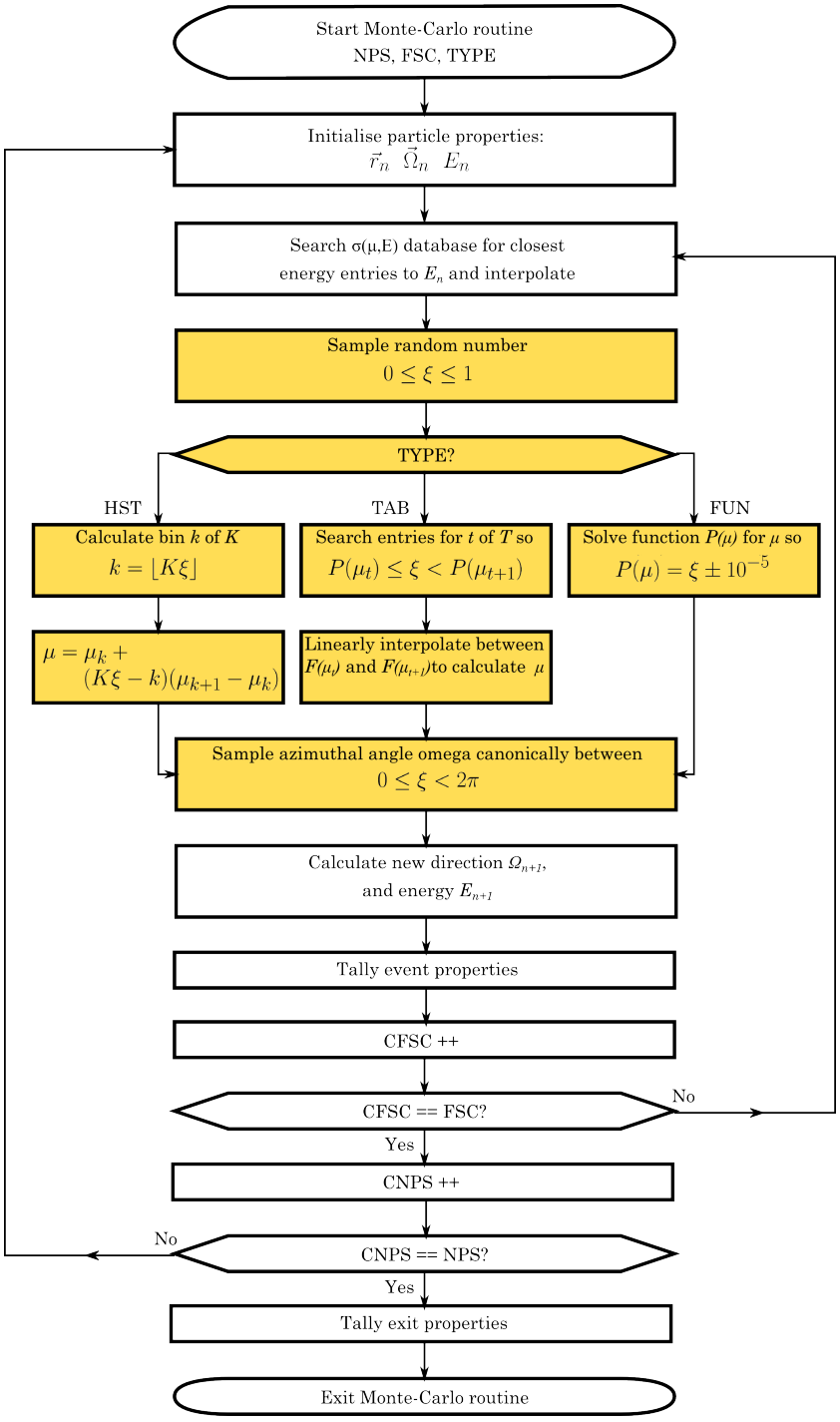

Figure 3.15: Monte-Carlo routine within Turnip, expansion from Figure 3.14. For *NPS* source particles, the angular distributions are sampled from the databases and the particle properties are updated. *FSC* determines the number of forced scatters requested, if this is greater than one, the particle is not transported in space but undergoes subsequent direction and energy changes based on the sampling of angular distributions. The event and exit particle properties are tallied. The stages highlighted in yellow correspond to the sampling of the different angular formats.

3.3.2 Full Monte-Carlo, basic geometries

To properly determine the implications of using the different data processing formats on a larger scale simulation, an increase in complexity is required. It is still necessary to retain complete control over the input data, so Turnip was altered to include transport. By using the original code as a base, transport through simple, known geometries was implemented. Further additions include the ability to sample multiple reaction channels and materials and several different source options. By increasing the complexity of the Monte-Carlo simulation to accommodate a more physical scenario, the clarity of the output data will be reduced. The implementation of geometry tracking, multiple reaction channels and materials will reduce the sampling of each distribution, potentially blurring the effects from different data formats. However, the level of control and flexibility required with respect to input data and sampling is retained. The flow diagram for the transport based Monte-Carlo routine can be seen in Figure 3.16, and replicates many of the techniques implemented in MCNP.

To validate the behaviour of Turnip, and to extrapolate the results to a more general Monte-Carlo code, equivalent geometries were constructed in MCNP. There are three main differences between the codes: data files, non-analogue transport and reaction channels. MCNP will accept the NJOY processed ACE files, whereas Turnip requires data files produced with Swede-Mash. MCNP will implement variance reduction techniques by default, but Turnip only transports neutrons through analogue methods. MCNP will sample from all available reaction channels, whereas Turnip will sample from elastic and discrete level inelastic only.

For a fair comparison, all possible non-analogue components were turned off within MCNP and both codes sampled from the 32 equal-probability histogram format. The MCNP data files were processed with the most recent version of NJOY, NJOY2012 for use in MCNP5. MCNP will tally neutron currents through surfaces, but this will include contributions

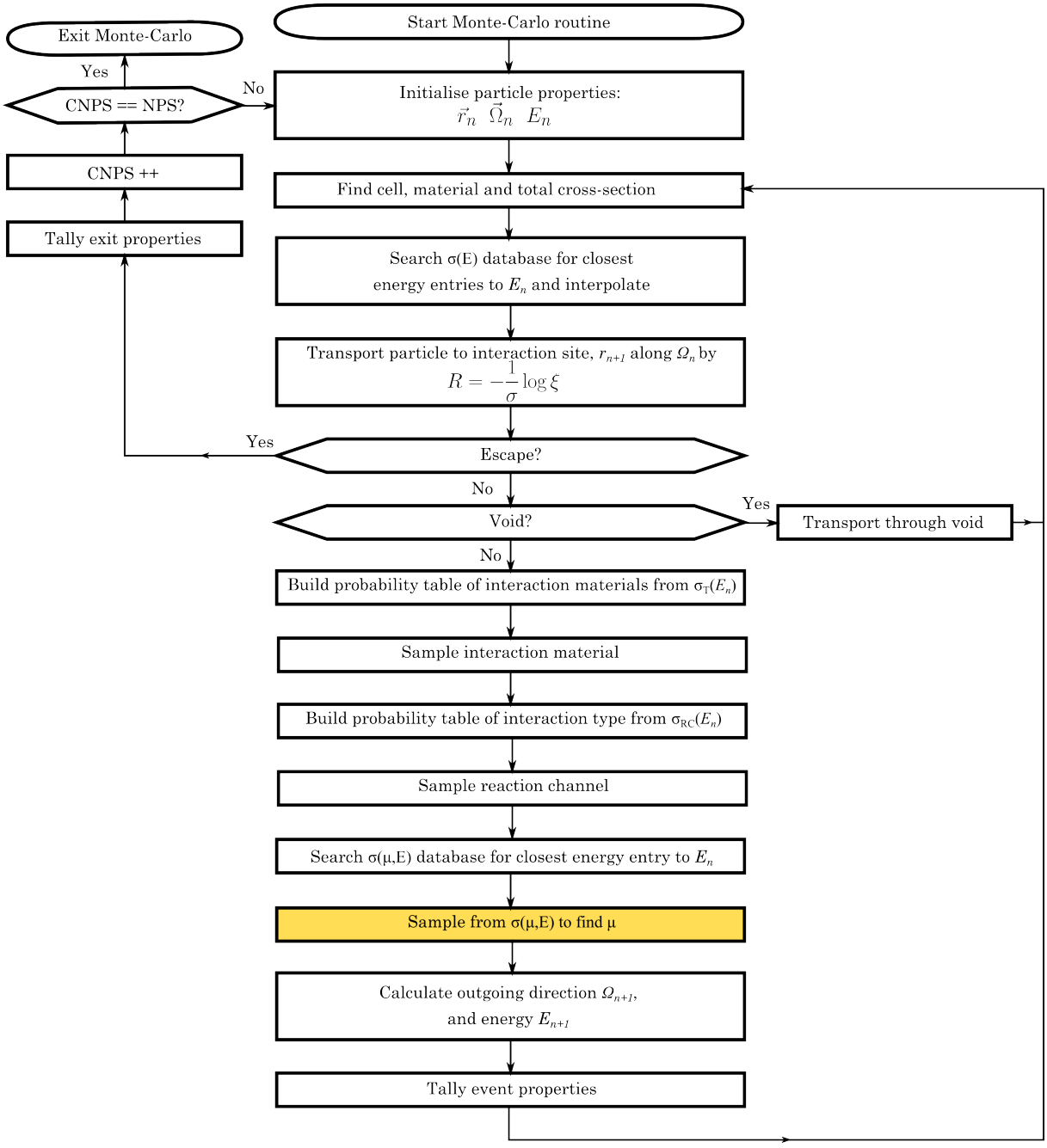

Figure 3.16: Flow diagram for the updated Turnip Monte-Carlo routine, this is closely related to the flow diagram in Figure 3.15, but with the inclusion of spatial transport throughout simple geometries, multiple reaction channel sampling and multiple material sampling. The stage highlighted in yellow corresponds to the sampling of the different angular formats. Figure 3.15 shows this stage in detail and is not repeated here.

from reaction channels that are not currently available in Turnip. To avoid this problem, all MCNP and Turnip simulations were required to produce an additional output file containing all particle histories and events (PTRAC files). From here, the MCNP output was filtered to exclude all reaction channels not implemented within Turnip. Both codes are then tallied in the same way, by counting the number of crossings at each surface and energy. The output files produced from MCNP required parsing into a similar format as Turnip for tallying purposes. This was achieved using an additional code known as Parsnip[3]. This was developed as part of this work, and has since been released as an open source tool within the wider MCNP community [74].

Benchmarking with simple geometries, prior to increasing the complexity, is to determine how comparable the two codes are, and validate the processing capabilities of SwedeMash in comparison to NJOY. By showing that MCNP and Turnip are comparable when presented with the same format of differential data, the potential implications of using non-standard formats can be inferred from their behaviour in Turnip.

Slab

A slab of isotopic, unit density material was created within Turnip and MCNP for each isotope. The source was defined as a mono-directional beam of 14.1 MeV neutrons, incident on the front face of the slab at (0,0,0). Additional surfaces were added to the volume to measure the energy dependent neutron current with respect to depth in the slab. To ensure that the results were comparable between codes and isotopes, the spacing of these surfaces was set to be in units of mean free path (MFP). This is calculated as $1/(\Sigma_{(n,el)}(14.1\text{MeV}) + \Sigma(n,n^{1*})(14.1\text{MeV}))$ for Turnip, where $\Sigma(14.1\text{MeV}) = N_D\sigma(14.1\text{MeV})$ is the macroscopic cross-section for each interaction type. The implemented sections were equivalent to a thin film (0.125 mean free paths), up to a thick slab (5 mean free paths). The example geometry can be seen in Figure 3.17.

[3]Parsnip is not an acronym, though it is a parser so the name is somewhat related.

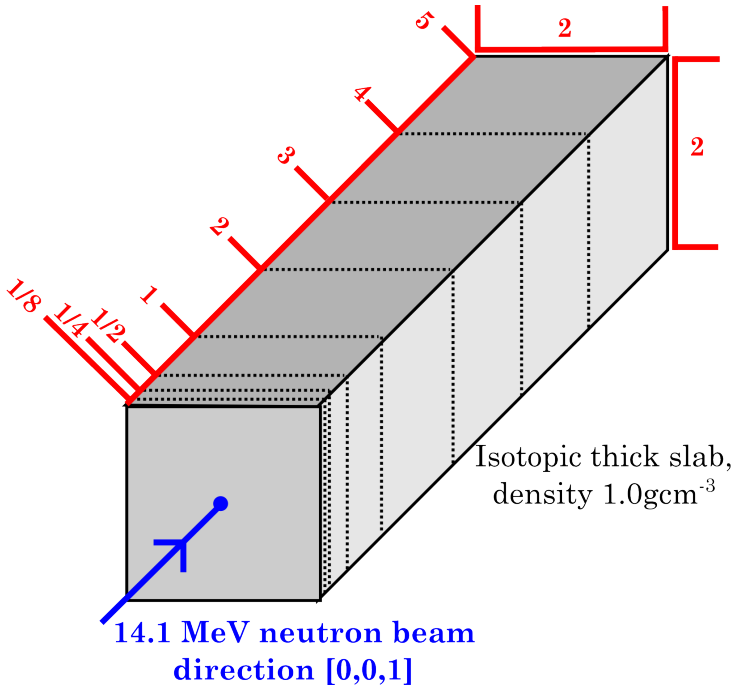

Figure 3.17: Schematic for the slab geometry run with Turnip and MCNP, dimensions shown in red are measured in terms of mean free paths. For each isotope at unit mass density, the mean free path of 14.1 MeV neutrons in the slab was calculated and used to set the slab thickness. The source used here is a mono-directional, mono-energetic beam of neutrons, incident on the front face of the slab at (0,0,0).

With reference to Equation 2.1.7, the spacing of the tally surfaces gives a range of expected interaction percentages between 12 and 99%, i.e. after 5 mean free paths only 1% of the beam should remain unreacted and the rest will have sampled from the processed differential distributions at least once. Given the width of the slab is 2 mean free paths, there is a high probability that neutrons will scatter out of the volume after a single event, and hence the neutron currents will reflect the input format. Turnip simulations were run with the standard MCNP data formats and additional alternatives. MCNP simulations were run with the 32 equal-probability histogram format, with the mean free path calculated as $1/\Sigma_{total}(14.1\text{MeV})$ for each isotope.

Sphere

A simple sphere, of isotopic, unit density material was simulated for each of the individual isotopes. An isotropic neutron source with an energy of 14.1 MeV, was placed at the centre of the sphere. Several concentric spherical surfaces were inserted for tallying, from which the neutron current throughout the volume was measured. Figure 3.18 shows the simple geometry used.

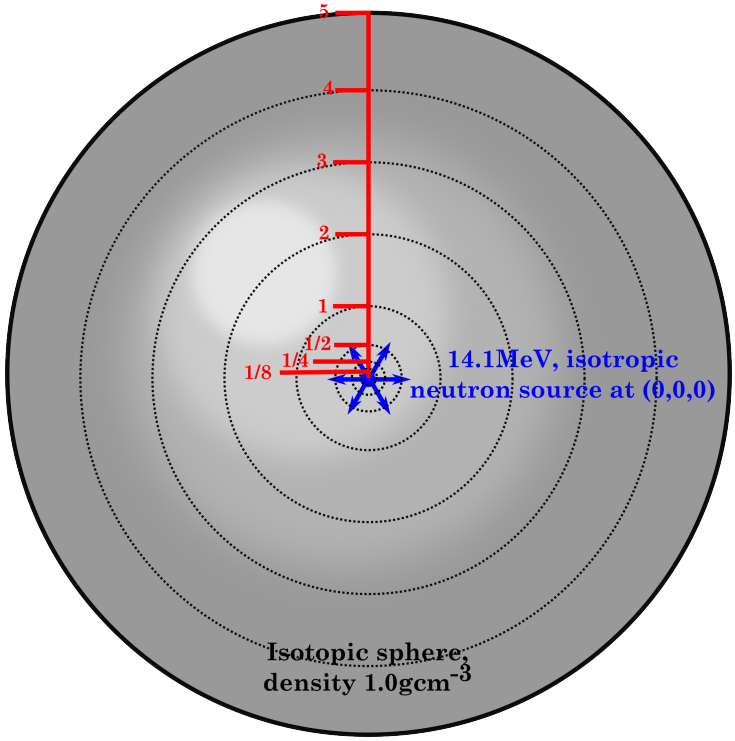

Figure 3.18: Schematic for the spherical simulation run with Turnip and MCNP, radial distances shown in red are measured in mean free paths. For each isotope at unit mass density, the mean free path of 14.1 MeV neutrons was calculated and used to set the shell radii. An isotropic, 14.1 MeV neutron source was placed at the origin. The region beyond the external bounding surface is set to be void, where no neutrons will be tracked.

Radii range from 0.125 mean free paths to 5 mean free paths. This was implemented in both Turnip and MCNP, and run with the same data formats as the slab geometry. The spherical geometry will demonstrate the differences due to data format over many scatters; the neutrons are born at the centre of the sphere and must pass through a minimum of 5

mean free paths of material before it can escape. There is a 1% chance that neutrons will have not interacted at the point at which they terminate. The remaining 99% will have scattered at least once, but the low escape probability of the geometry means that many consecutive scatters are likely.

3.3.3 Real-world MCNP simulations

To determine how real-world simulations are affected by the data format of differential data, two models, both widely used within the fusion community, have been chosen. The first model is the port-plug mock-up as seen in Figure 3.19 and is one of the standard ITER benchmark studies [75].

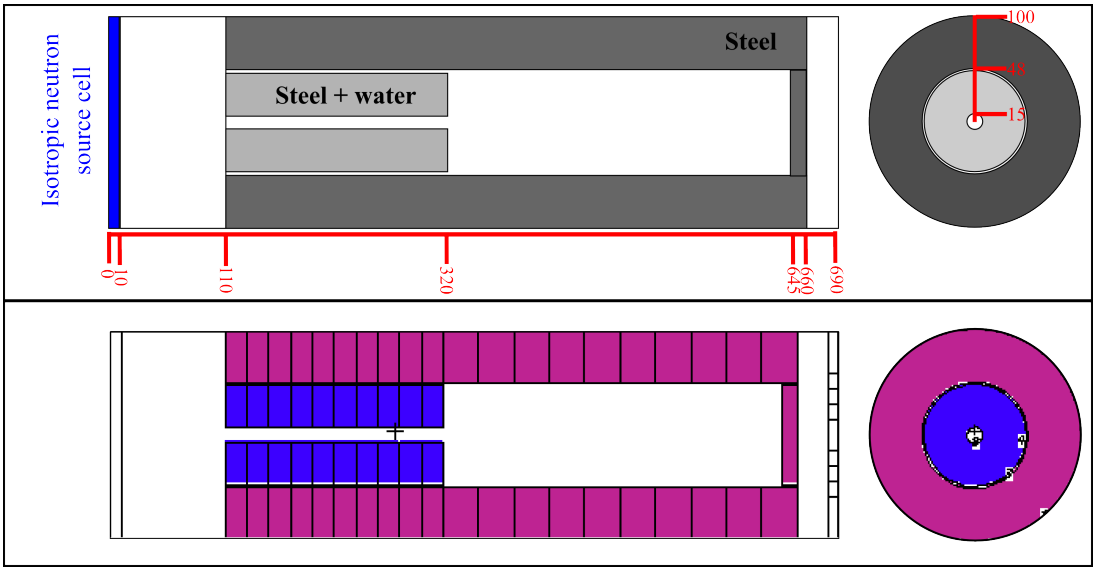

Figure 3.19: Top image shows the schematic of the port-plug y-z and x-y planes, with dimensions (red) given in cm. The bottom image shows the MCNP geometry, with additional surfaces introduced at intervals for tallying. The physical representation remains identical [75]. Note that additional features on the lower image are artefacts of the MCNP plotting utility and not the geometry.

This is typically used to calculate the shut-down dose after different simulated operational modes, and for variable lengths of irradiation time. Examples of its use can be found in [76] and [77]. The model is run with MCNP and a distributed neutron source. The resulting energy dependent neutron flux throughout the geometry is calculated. For shut-

down dose calculations these data are passed to an inventory code, such as FISPACT-II [78], where the activation due to the neutron flux is determined. The resulting gamma flux profile is inserted back into the original MCNP geometry and is run to determine the secondary particle behaviour. Between the two codes, a shut-down dose can be calculated at various cooling times. The port-plug model is a natural progression from the spheres and slabs, as the geometry is relatively simple and only contains two different materials. The second model is a segment of the DEMO tokamak, which is used for the testing of different tritium breeding scenarios and is shown in Figure 3.20.

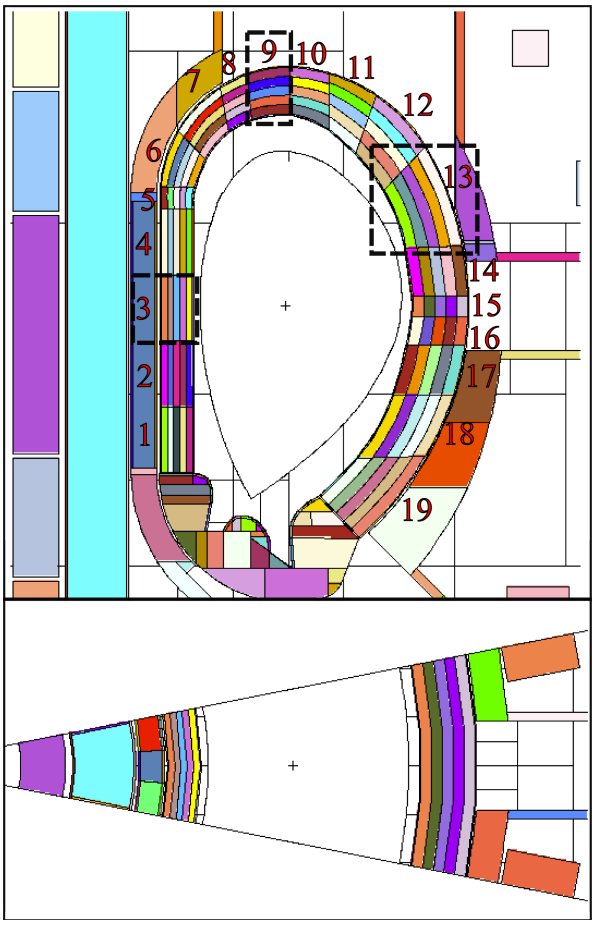

Figure 3.20: Segment (22.5°) of the proposed DEMO design [79]. The tritium breeding blanket modules surround the D-shaped plasma, and are split into 5 individual segments. This provides a spatial distribution of tritium production within the model. The sum of all blanket reaction rates is used to calculate the total tritium production. Each module has been numbered for clarity and the three modules highlighted (black dashed boxes) are used additionally to measure the net neutron current and other reaction rates.

The neutrons within the model are generated by a parametric plasma source, resulting in a range of energies, directions and spatial distribution. The neutron flux is determined within the different blanket modules and is subsequently used to calculate reaction rates for tritium production in each module. For a known level of tritium consumption and neutron flux, the breeding ratio can be calculated globally [80]. Table 3.5 shows the source parameters associated with the toroidal plasma neutron source. Each neutron is born within the plasma at 14.1 MeV, but will interact with the surrounding plasma, hence the energy will be distributed when it reaches the plasma facing components.

Table 3.5: Parameters for the plasma based neutron source in the DEMO model. The location of each source neutron was determined by the plasma density and the direction sampled as an isotropic emission.

Parameter	**Value**
Major radius (m)	9.00
Minor radius (m)	2.25
Triangularity	0.33
Peaking factor	1.30
Elongation	1.66
Energy (MeV)	14.10

These models are too complex in geometry, source definition and output requirements to be easily replicated in Turnip. The simulations were run with MCNP only, with the two allowed data formats. To obtain additional information from the different sampling mechanisms, the energy dependent neutron current was tallied at stages throughout both models and several reaction rates calculated. Based on these results, it is possible to determine the sensitivity of the models to the currently available formats, and infer the effect of alternative data formats based on the simple benchmarking in Turnip.

Chapter 4

Results

To demonstrate the effect of pre-processing angular distributions within the context of fusion neutronics, ^{56}Fe will be used as the exemplar case. Each stage of the analysis will be shown in detail, with comparisons where necessary. The effect of pre-processing on the remaining fusion relevant isotopes is presented as a summary of the same analyses. These cases are used to study the sampling process in detail, but are not physical. The results from the port-plug benchmark and DEMO models are presented to demonstrate the effect of data format on real-world simulations currently used within the fusion community.

4.1 ^{56}Fe example

The ENDF data file for ^{56}Fe was taken from the ENDF/B-VII.1 data library. This library is considered the industry standard, general-purpose library. For fusion the most common neutron energy within close proximity to the plasma is 14.1 MeV, the following analyses are performed on the ^{56}Fe(n,el) data set closest to this value. Even though this is one of the best evaluated materials, the closest angular distribution is at 13.57 MeV.

4.1.1 Point-wise

The ENDF file was passed to Swede-Mash to parse and extract the cross-sections and differential cross-sections associated with the elastic scattering reaction. The cross-sections

were directly output to Turnip compatible files. The differential data with respect to angle were primarily processed into the MCNP standard 32 equal-probability histogram format. Figure 4.1 shows the original and processed distributions for the fusion relevant ^{56}Fe(n,el) reaction channel, with an incident neutron of 14.1 MeV alongside the fission relevant ^{nat}C(n,el) reaction channel at 2.0 MeV.

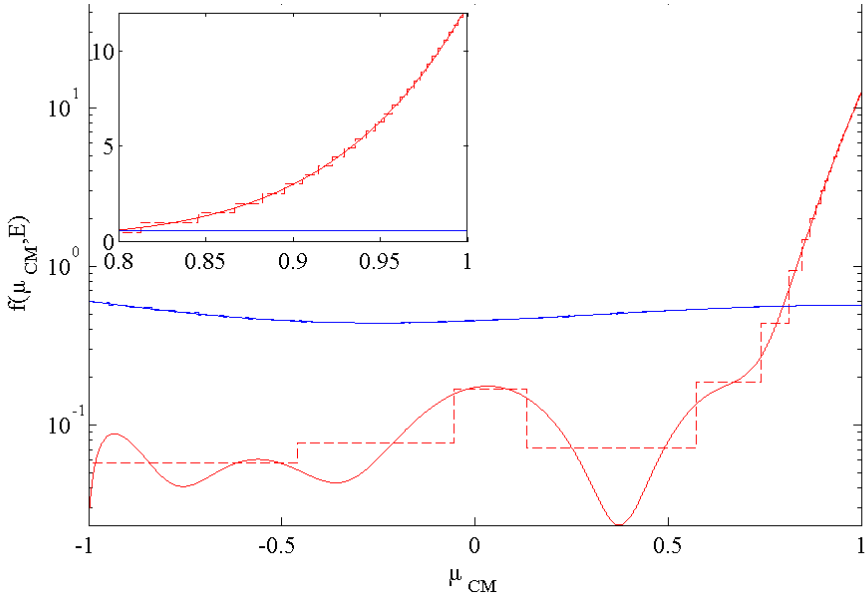

Figure 4.1: Elastic scattering distributions for the fission reaction, ^{nat}C(n,el) at 2.0 MeV (blue) and the fusion reaction ^{56}Fe(n,el) at 14.1 MeV (red). The normalised probability, $F(\mu)$ is given on a log scale as a function of the scattering cosine, μ, in the centre-of-mass frame. Processed distributions are shown with the dashed lines. The feature at $\mu = 0.4$ in the ^{56}Fe distribution is significantly overestimated by the processed format, though over the bin the average difference is zero. Inset shows the forward scattering region on a linear scale. Data for ^{nat}C and ^{56}Fe are taken from the ENDF-B/VII.1 library [81]

The distribution for ^{56}Fe at fusion energies is highly featured, with a distinct forward bias. The ^{nat}C reaction channel is closely isotropic by comparison. The overlaid processed distributions at these energies show the locations at which the original distributions are mis-represented. In ^{56}Fe the low-probability back-scattering region is crudely discretised by the processed format, most notably the region around $\mu = 0.4$. The original forward-scattering region is smooth, with a high rate of change and is converted into a step function that continually over and under samples the original (with an average difference of zero). For ^{nat}C, the isotropy and lack of features results in a close representation to the original.

The coefficient of variation for these angular distributions were calculated to be 11.0% and 0.6% respectively. Table 4.1 presents the accuracy metrics used within this analysis.

Table 4.1: Point-wise results of the fission and fusion reaction channels presented in Figure 4.1.

Reaction channel	Δ	C_v(%)
^{nat}C(n,el) 2.0 MeV	-0.012	0.6
^{56}Fe(n,el) 14.1 MeV	-0.265	11.0

The C_v represents the variation of differences between the original and processed forms. For ^{56}Fe the maximum difference is reach -0.265, or 53.0% if normalised. For ^{nat}C the maximum difference is -0.012, or 2.4%. It is worth noting, that within MCNP simulations, a statistical error of 10.0% or higher is considered to be questionable, below 10.0% the simulation result is generally considered reliable depending on the application [23, p1-7]. Ideally simulations will achieve a much lower level of statistical uncertainty, and higher level of convergence as a result, where differences in the pre-processing may become apparent.

The C_v analysis was performed on each of the elastic scattering distributions within the ^{56}Fe ENDF file. Figure 4.2 demonstrates how the C_v varies over the entire energy range covered by the distributions. At low energy ($10^{-5}-2.5$ MeV), the data density is far higher than the remainder of the file; of the 375 energy entries, 348 are below 2.5 MeV (93%) and only 27 exist between 2.5 and 20.0 MeV (7%). The low energy distributions show some variation between isotropic and slightly featured resulting in fluctuations within the C_v, i.e. featured distributions result in a higher C_v than closely isotropic ones. Between 2.5 MeV and 15.0 MeV the C_v is well behaved, tending to a constant value (11% for histogram format, 1% for tabulated by design). Above 15.0 MeV, there is a step change in C_v due to the increase in forward bias and additional features appearing in the distribution.

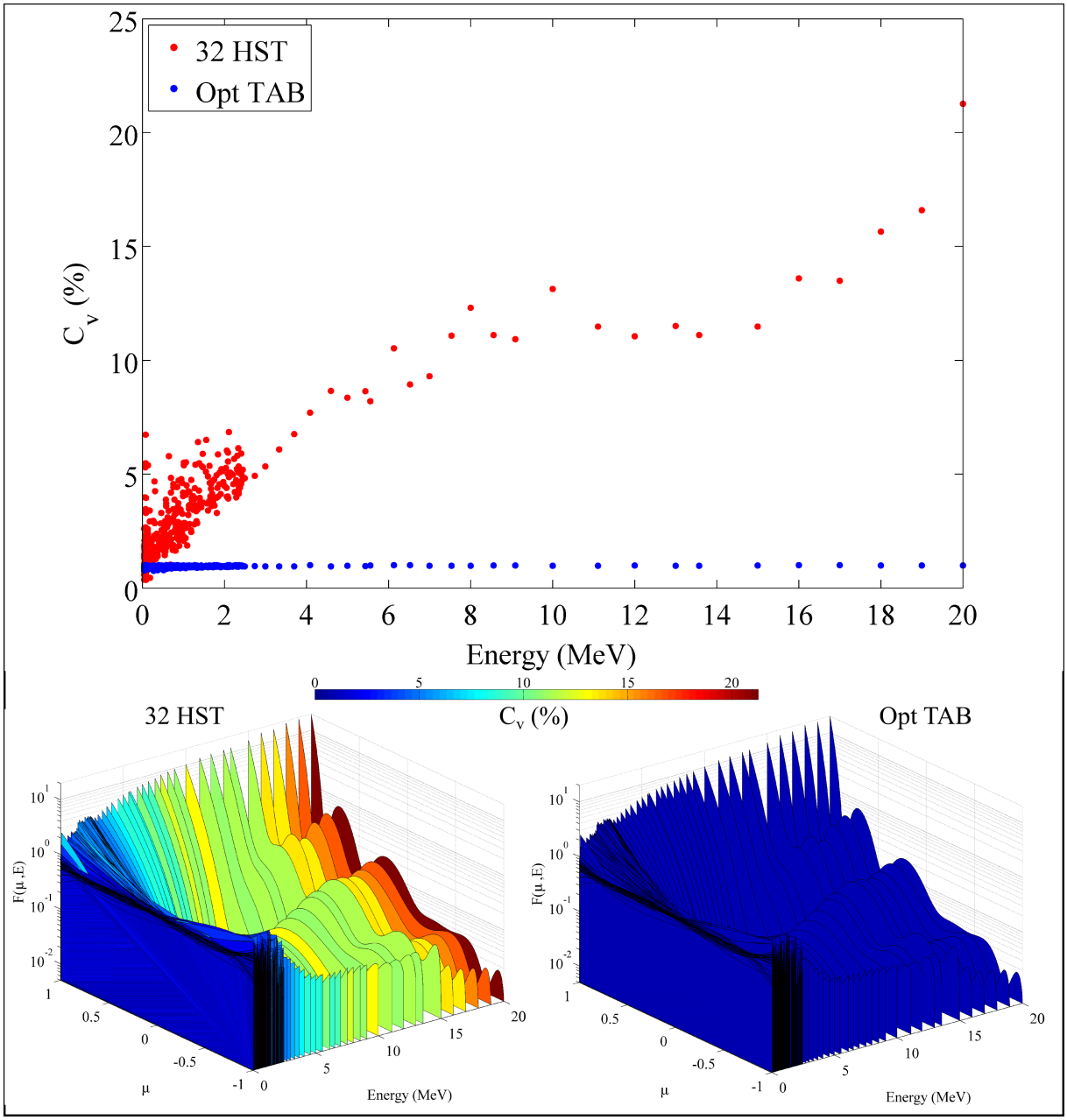

Figure 4.2: Top plot shows C_v as a function of energy for ^{56}Fe elastic scattering reaction for the accepted MCNP formats, 32 equal-probability histogram (red) and optimised tabulated data (blue). Bottom plots show original (un-processed) distributions with respect to energy and scattering cosine, coloured by the C_v associated with the processed format.

Up to this point, only single elastic scatters have been considered. Figure 4.3 shows the potential scattering distributions for two consecutive elastic events in ^{56}Fe, with an initial neutron energy of 14.1 MeV when sampled from the unprocessed and 32 equal-probability histogram format. The maximum energy loss from one scatter is less than the difference in energy to the next distribution, so there is a high probability that within a radiation transport simulation the same distribution will be sampled for further scatters. Comparing the histogram format to the functional format shows the point of largest difference at $\mu_1 = \mu_2 = 0.4$. This is consistent with the distributions in Figure 4.1, where the region of low probability at $\mu = 0.4$ is over estimated by the 32 equal-probability histogram format. At its worst point, the histogram format is 9 times that of the functional, though the absolute value is small. It is not possible to visualise further consequential scattering events with this method so a simple Monte-Carlo model is used to simulate further events.

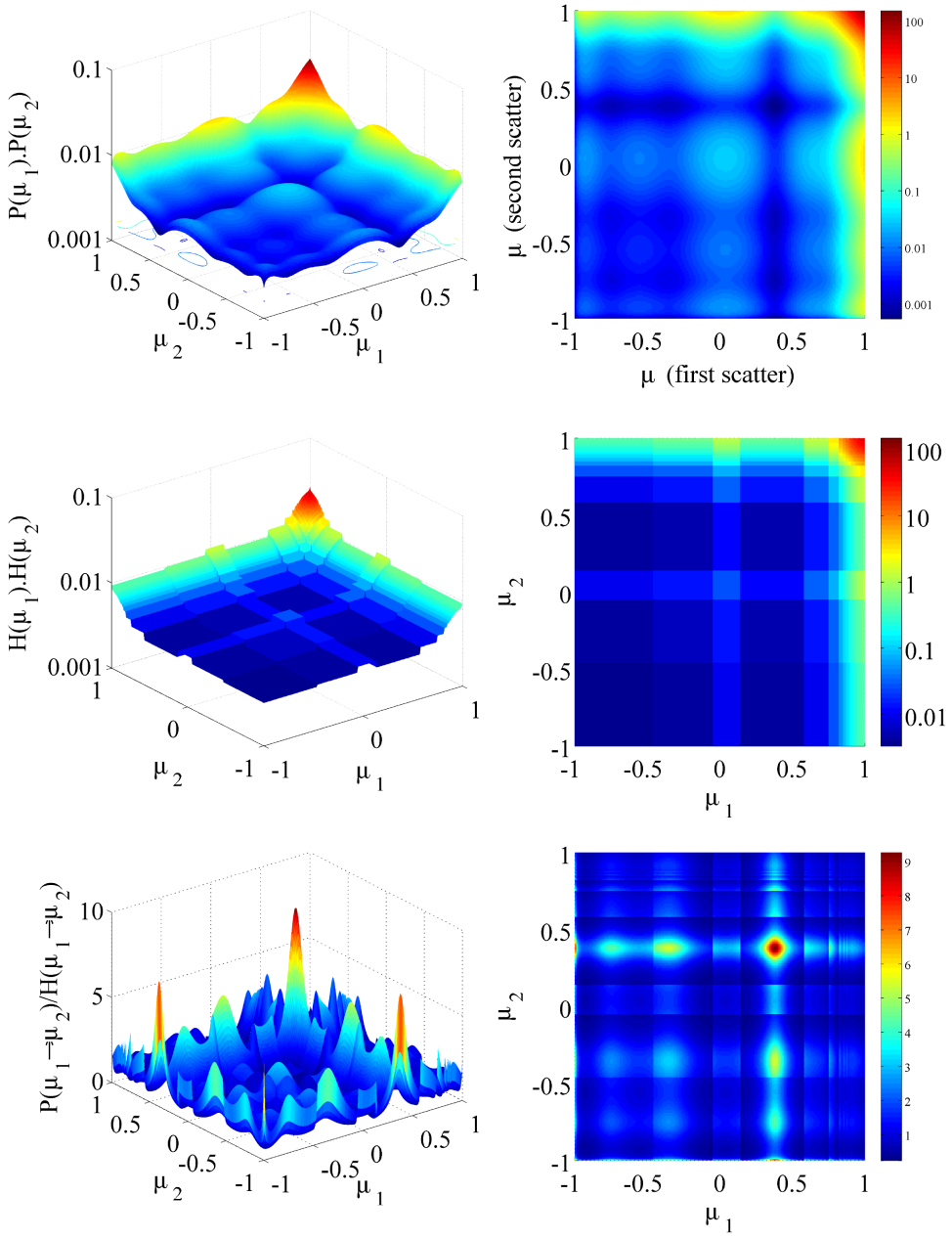

Figure 4.3: Probability of scattering into μ_1 followed by μ_2 for the original format (top) and the processed 32 bin histogram (middle). The difference between the two surfaces are shown in the bottom plots. The area of greatest difference corresponds to $\mu_1 = \mu_2 = 0.4$.

4.1.2 Turnip

The processed data from Swede-Mash was input to Turnip. Monte-Carlo simulations were performed to first validate the data files produced by Swede-Mash and the ability of Turnip to sample the various angular distributions.

Forced scatters, current formats

Within Turnip, a single scatter in ^{56}Fe at 14.1 MeV was forced for an increasing number of source neutrons, sampling from the functional, 32 bin histogram and optimised tabulated formats. The exit cosines were scored after a single event to reconstruct the input distribution in 100 equally spaced μ bins. The scored behaviour is shown in Figure 4.4 for $10^2 - 10^5$ source neutrons. At low numbers of source neutrons (below 10^4), the different formats are indistinct, and converge on to the expected distributions by 10^5 histories.

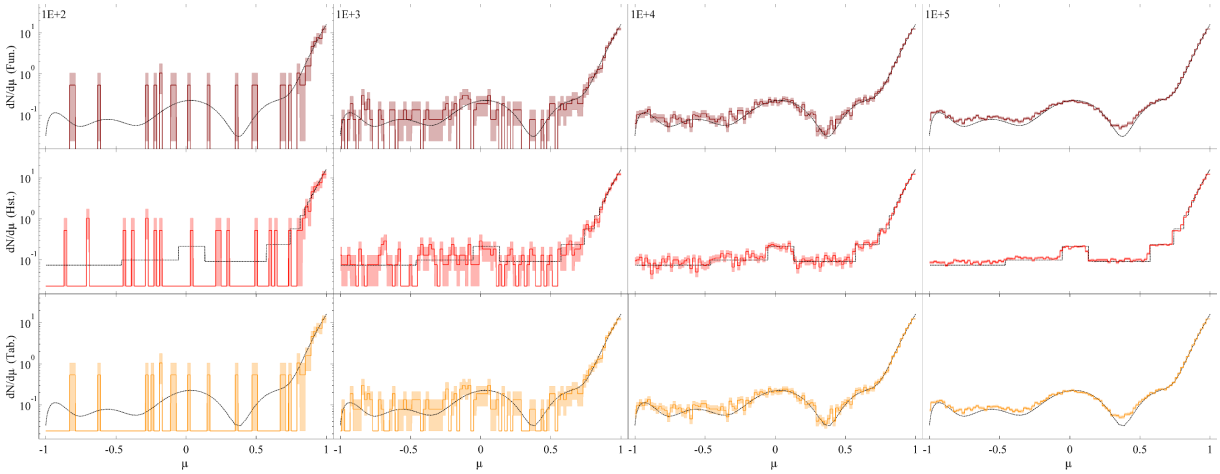

Figure 4.4: Plots show the scored exit angles with μ for an increasing number of source neutrons, after one forced elastic scatter in ^{56}Fe at 14.1 MeV. The top row shows the unprocessed, functional data, the middle shows the 32 equal-probability histogram and the bottom row shows the optimised tabulated data. The input distributions are shown on each plot with the black dashed line.

Figure 4.5 shows the calculated C_v for the histogram and tabulated formats as compared to the functional output distribution.

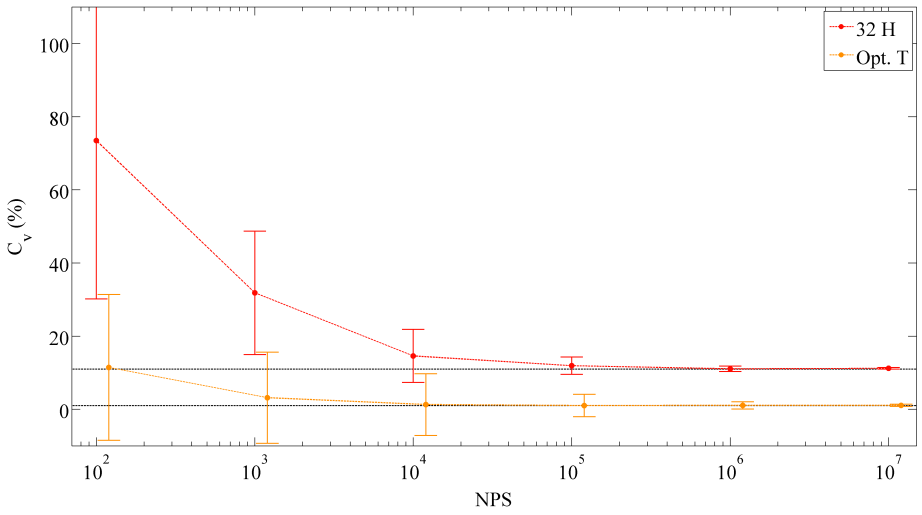

Figure 4.5: Calculated C_v with respect to the number of source neutrons (NPS log scale) for the distributions shown in Figure 4.4. These are calculated between the 32 bin histogram or optimised tabulated formats, compared to the equivalent functional distributions. Lines between points are shown to guide the eye only. Black dashed lines correspond to the expected C_v as calculated from the point-wise analysis.

Within errors, the histogram and tabulated formats are indistinct until a minimum of 10^5 source neutrons. After this point, the histogram and tabulated distributions converge to the expected values of 11% and 1% respectively. This suggests that for these distributions, they must be sampled a minimum of 10^5 times to be able to see the effect of implementing alternative data formats. If the distributions are not fully populated within a radiation transport simulation, the statistical uncertainty will reduce the possibility of seeing any variations.

Turnip has several modes of operation, it is possible to force a fixed number of scatters sampled from the angular distributions in three dimensions. In this case a 14.1 MeV source neutron is forced to interact at the origin, with a positive direction along the z-axis. It undergoes the angular change in the lab frame due to the sampled μ in the centre-of-mass frame and the energy is altered according to Equation 2.1.4. The particle is not

transported between collisions to generate statistics on the sampling process without the added complexities of a physical system. If one forced scatter is requested, the particle is then transported to the surface of an arbitrary sphere where its energy and exit angle are scored in the lab frame. Figure 4.6 presents the data for ^{56}Fe, with varying numbers of forced scatters for the processed formats.

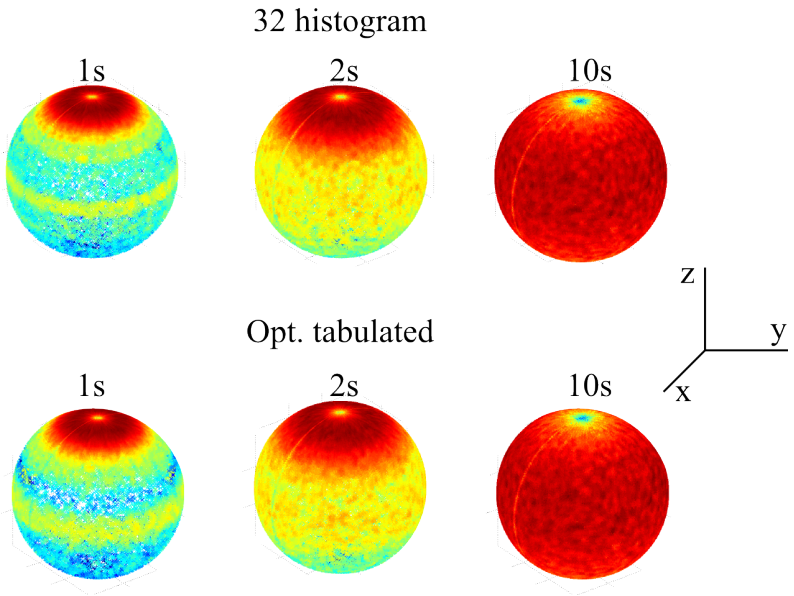

Figure 4.6: For a number of forced scatters, sampled from the 32 bin histogram (top row) and optimised tabulated (bottom row), the exit position is plotted on an arbitrary sphere. Colours denote the point density, where there is a large variation in colour, there is a large variation in the scattered distributions. For the cases where there is little colour variation, such as after 10 scatters, there is a relatively even distribution of points over the surface. Areas of low density as seen at the poles is an artefact of the conversion from μ to θ. Figure 4.7 plots these data with respect to μ.

The positions are plotted on the surface of the problem sphere and are coloured by the density of points. For one scatter it is possible to see the slight difference between the 32 histogram and optimised tabulated data. The histogram data has clear bands corresponding to the histogram bins, while tabulated data has a much smoother transition between regions of high and low probability. There is a visual difference between the two corresponding to the $\mu = 0.4$ region. After two scatters the difference between the formats has visibly reduced. After 10 scatters the resulting distributions for both formats result in isotropic distributions in μ.

For multiple scatters sampling from the 32 equal-probability histogram and optimised tabulated data, the scattering cosines in the lab frame were tallied. The resulting probability distributions can be seen in Figure 4.7. In conjunction with Figure 4.6 there are no obvious differences in the lab frame angular distributions after multiple scatters.

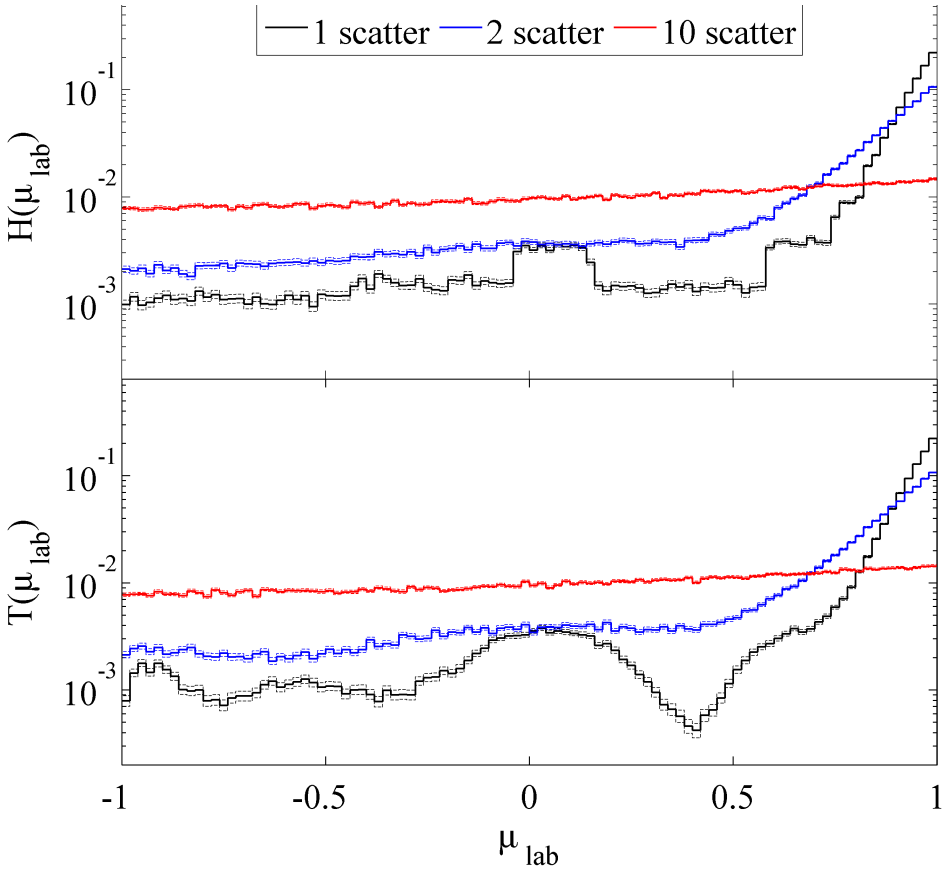

Figure 4.7: Sampled data from Turnip for 1, 2 and 10 forced scatters with respect to the lab frame cosine, μ_{lab} for 10^5 source neutrons. The top figure shows the resultant distributions for the 32 equal-probability histogram (H), the bottom figure shows the distributions for the optimised tabulated data (T).

Figure 4.8 shows the variation of C_v with the number of forced scatters. For each scatter, the resulting lab frame distribution of the functional data was considered to be the ideal case and compared to the equivalent distribution from the processed data.

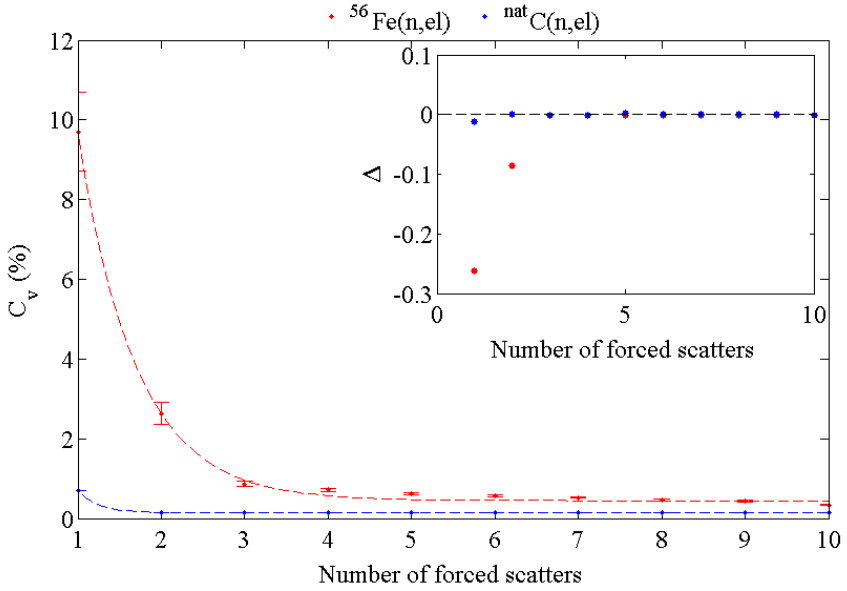

Figure 4.8: Calculated C_v as a function of forced scatters for ^{56}Fe(n,el) at 14.1 MeV (red) and ^{nat}C(n,el) at 2.0 MeV (blue), as sampled from the 32 equal-probability histogram format. Inset shows the behaviour of the maximum difference as a function of forced scatters. The dashed lines are fitted to the form $a\exp(-bN_s)+c$ [82].

The 32 equal-probability histogram has an initial C_v of 11.0%, which reduces to approximately 1.0% after 4 scatters. Beyond 4 scatters the C_v drops below the statistical limit on each tallied bin (1.0% Poisson error). The distributions in lab frame angle show no significant differences after multiple, statistically identical events. Each distribution in μ corresponds to a similar distribution in energy. To determine the effect of these different formats on the energy spectra, multiple scatters were forced and the resulting exit energy distributions calculated. Figure 4.9 shows the energy distributions for tabulated and histogram data for up to 9 scatters. The features associated with the different formats remain distinct for the first 5 scatters. After 5 scatters, the features begin to merge and the differences between the histogram and tabulated formats are reduced.

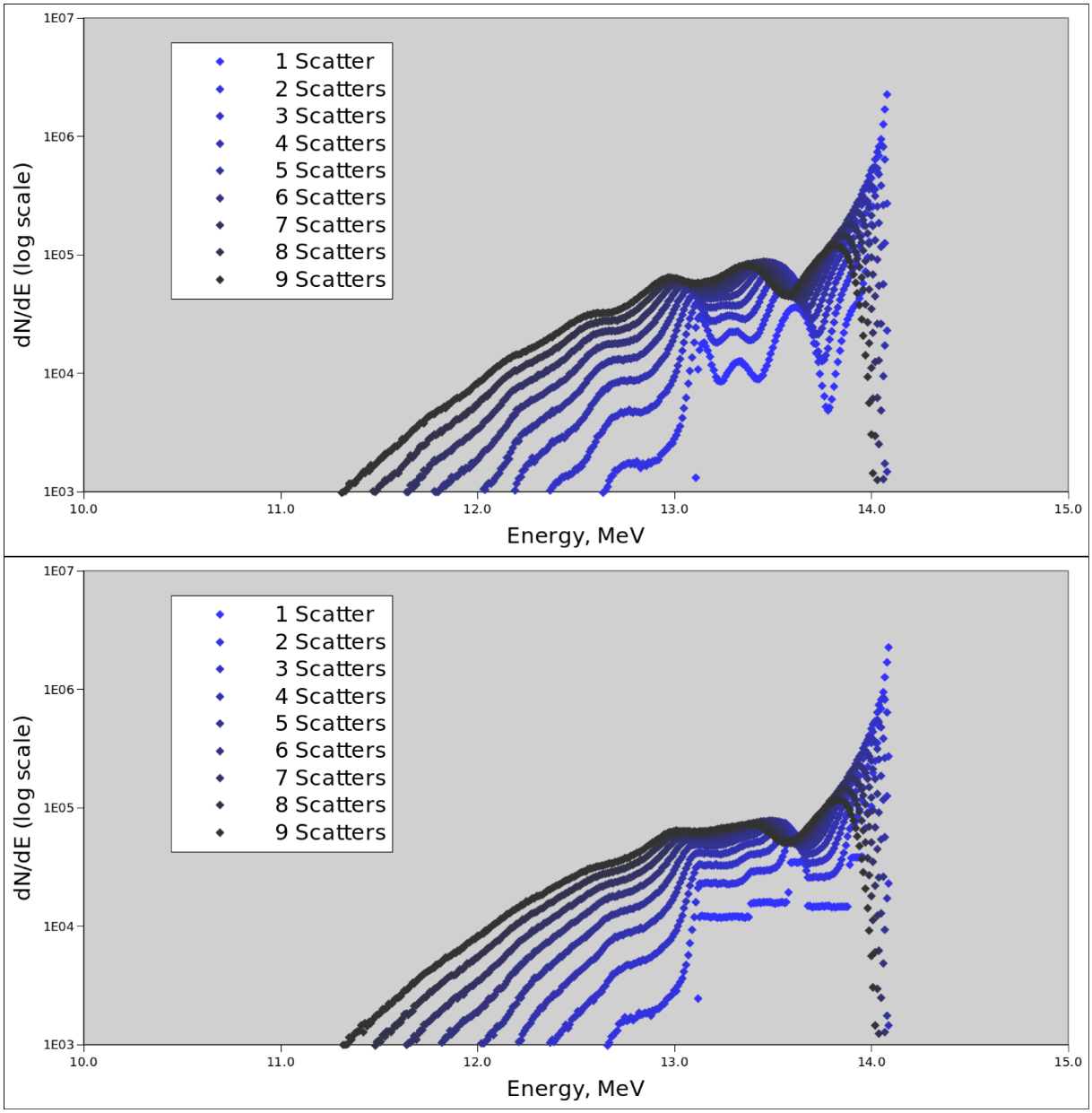

Figure 4.9: Resulting energy distributions for 1-9 forced scatters in Turnip. Top figure shows the data for the optimised tabulated format, and the bottom figure shows the equivalent for the 32 equal-probability bin histogram [T.W. Leadbeater, private communication].

Figure 4.10 shows the comparison of the 32 equal-probability histogram and optimised tabulated formats for multiple scatters with respect to the functional format. The optimised tabulated results are identical to the functional data within errors. Histogram data shows clear differences in comparison to the functional data. For one scatter, the maximum difference is a 203% over-estimate by the histogram format, for two scatters this reduces to 103%. By 9 scatters the maximum over-estimate is 13%.

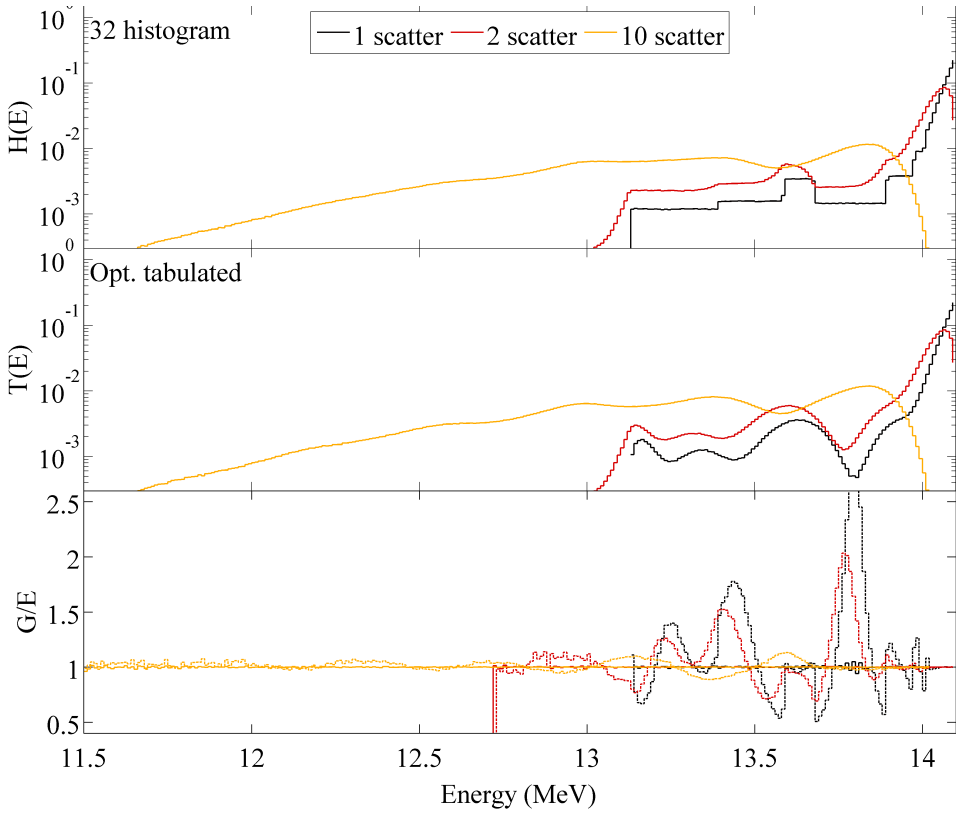

Figure 4.10: Comparison of energy distributions for 1, 2 and 10 scatters. The top plot shows the probability distribution for the 32 equal-probability histogram data and the middle shows the optimised tabulated data with respect to exit energy. The bottom plot shows the comparison of these scattered distributions with the equivalent for functional data. Dashed lines correspond to the histogram to functional comparison. Solid lines correspond to the tabulated to functional comparison, and are close to 1, so cannot be seen individually.

Figure 4.11 shows the calculated C_v with respect to the number of scatters for the energy spectra in each case. The resulting processed distributions are compared to the unprocessed functional energy distributions. A single scatter has a C_v of $10.4 \pm 1.3\%$, this remains constant within errors until after 5 scatters where the C_v is reduced. By 10 scatters, the C_v reduced to $6.7 \pm 0.6\%$. The initial scatter has the sharp forward scattering peak at 14.1 MeV, which shifts to lower energies with consecutive scatters and the associated energy loss. The other features, such as the $\mu = 0.4$ region of over representation by the histogram format, remain visible even after 10 scatters.

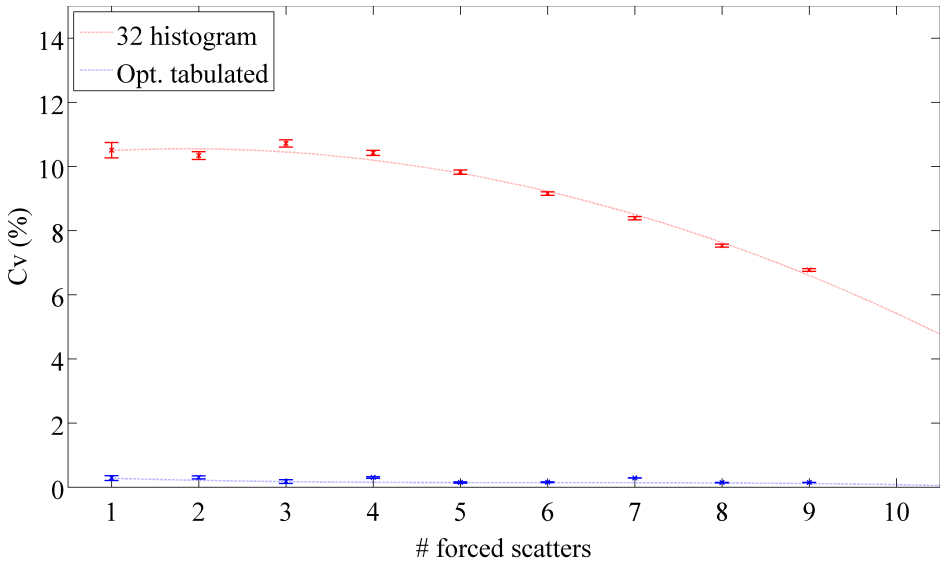

Figure 4.11: Calculated C_v with respect to the number of forced scatters. The values for 32 equal-probability histogram (red) and optimised tabulated formats (blue) are calculated with respect to the functional equivalent.

Figure 4.12 shows the variation of the energy distribution associated with the tabulated data for 1-90 forced scatters. The initial scatter always occurs at 14.1 MeV, and subsequent scatters at correspondingly lower energies. The original, asymmetric distribution tends to a symmetric distribution for a large number of forced events.

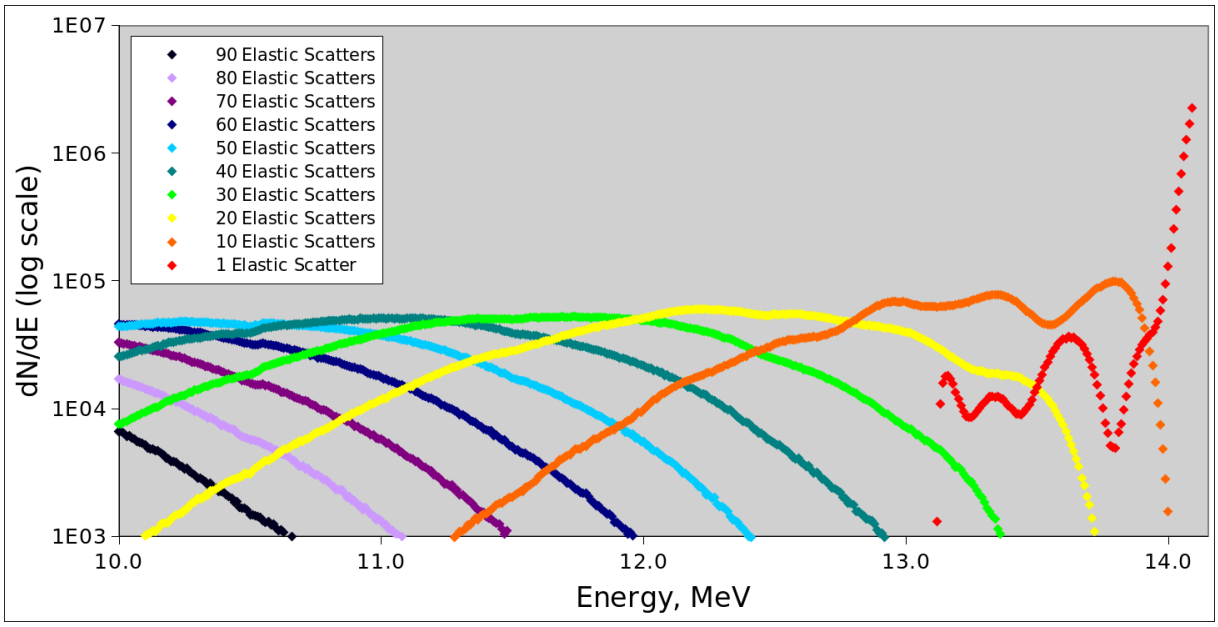

Figure 4.12: Resulting energy distribution from Turnip simulations with 1-90 forced scatters, sampled from the optimised tabulated data set with 10^9 source neutrons. The distributions lose features with increasing numbers of scatters, tending to the closely symmetric distribution at 40 scatters. For further scatters, the shape of the distribution remains the same, but the peak shifts to correspondingly lower energies. This behaviour is replicated for all other formats for a large number of forced scatters [T.W. Leadbeater, private communication].

The resulting energy spectra are more sensitive to the pre-processing format after multiple events when compared to the lab frame angular distributions. With reference to the angular distributions in Figures 4.7, the features associated with each format are indiscernible after two forced scatters. When the equivalent energy spectra are considered in 4.10, there is still a noticeable difference due to the input formats after 10 scatters. The rate of blurring, or loss of features, with respect to forced scatters is greater in the output angular distributions than the associated energy spectra, suggesting that energy spectra are more sensitive.

Forced scatters, alternative formats

The previous section showed the propagation of the two known pre-processing formats for a forced number of scatters. This section looks at the effect of using variations of current formats on the computational metrics associated with a simulation. The ^{56}Fe

elastic scattering data were processed with Swede-Mash to produce Turnip data files as described in Section 3.2.2. Formats range from a single histogram bin (isotropic in μ), up to 512 histogram bins. An equivalent number of tabulated formats were also produced. For each data set the C_v of the point-wise distribution was calculated within Swede-Mash. Turnip was then used to determine the computational runtime and memory requirements associated with each format. Figure 4.13 shows the variation of C_v with respect to energy for the histogram and tabulated data, for the 1-512 bins or channels.

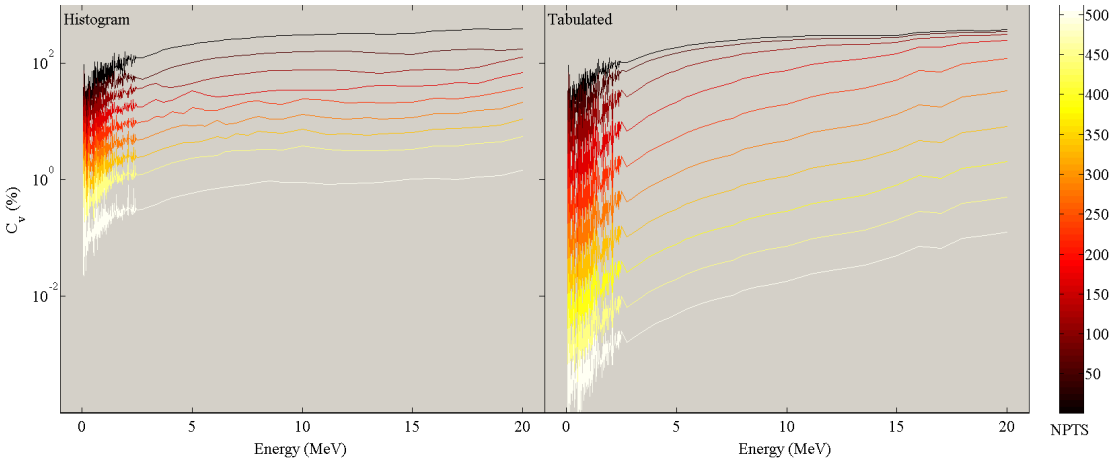

Figure 4.13: Calculated C_v values for the alternative histogram and tabulated formats. Plots show the C_v as a function of energy for all of the elastic scattering distributions within the ^{56}Fe ENDF file. The different colours represent the number of histogram bins (left) or tabulated channels (right).

The C_v increases with incident energy, which is consistent to the forward bias and highly featured back-scattering region associated with these distributions. Increasing the number of histogram bins or tabulated channels reduces the C_v.

For the first four distributions shown (1-8 bins/channels), the spacing between consecutive C_v distributions is larger for histogram data than tabulated, i.e. initially increasing the number of bins in the histogram format results in a larger reduction in C_v than increasing the number of tabulated channels. This effect is most obvious above 3.0 MeV. Equal-probability histogram μ limits are set based on the requirement that each bin must have an equal area and are calculated from the functional format, i.e. the location of the μ

value is closely linked to the original distribution and no normalisation is required. The μ values for the tabulated data are spaced evenly in θ, where $\theta = \arccos\mu$, i.e. they are arbitrarily set. The probability value associated with this μ calculated from the original, functional distribution and normalised to give an integral value of 1.0 over the μ range. This results in certain features, particularly in the extreme forward and back-scattering regions, being underestimated. Figure 4.14 shows the unprocessed and processed distributions for ^{56}Fe(n,el) reaction channel at 14.1 MeV for 1-4 bins or channels. After 4 channels, the tabulated μ values are more likely to coincide with a feature within the distribution and hence provide a better overall representation.

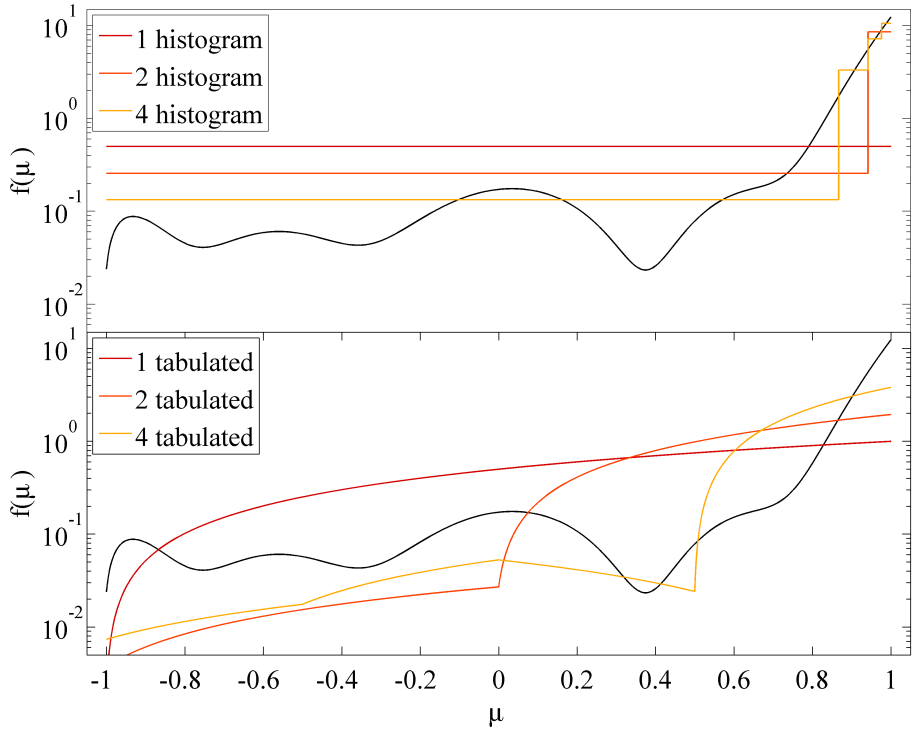

Figure 4.14: Angular distributions of ^{56}Fe(n,el) reaction channel at 14.1 MeV for variations on the histogram format (top) and tabulated format (bottom). In both cases, the unprocessed, functional form is represented with the black line. The μ limits for histogram data are determined from the integral of the functional form. The μ values associated with the tabulated data are equally spaced in θ, and probability values are normalised to give a total area of 1.0 for sampling. This results in an underestimate of the forward scattering region and a poor representation of the back-scatter features.

The C_v values for the tabulated format with 8-512 channels show a larger gradient than the histogram for an increasing number of points, i.e. the reduction in C_v with respect to an increased number of channels/bins is greater for tabulated format than the histogram format. This effect is largely due to the different interpolation methods, where the linear interpolation between points can better represent regions with a high rate of change compared to the histogram. The energy averaged C_v values were calculated for each format, and can be seen in Figure 4.15. If a limit of 1.0% average C_v were to be placed on the file, histogram format would require an average of 128 equal probability bins and tabulated data with linear interpolation would require an average of 32 channels. Low-energy distributions require fewer bins/channels to achieve the set precision and high-energy, anisotropic distributions require more.

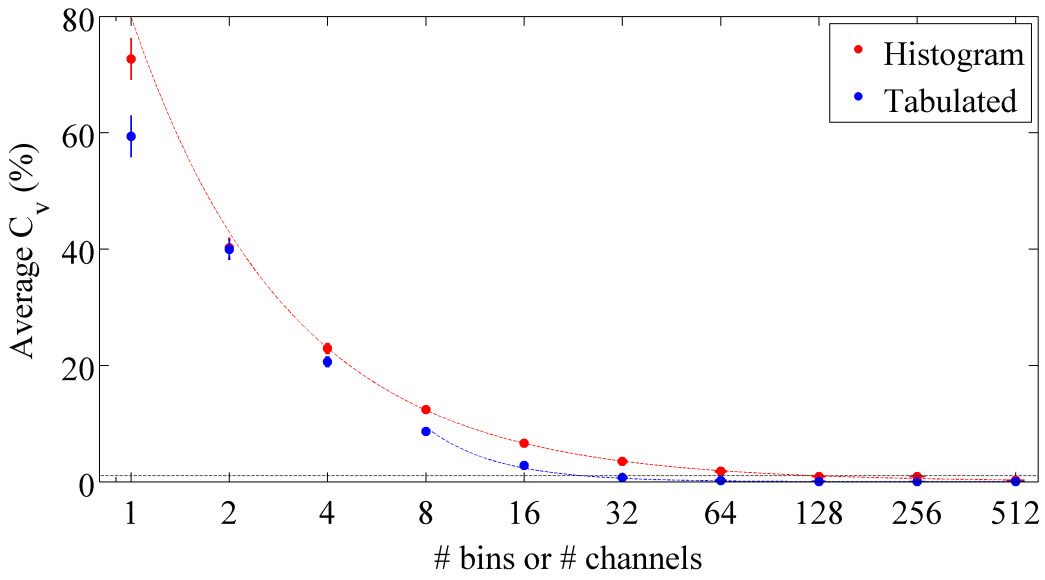

Figure 4.15: Figure shows the file averaged C_v as a function of bins/channels for the two data types. Histogram data (red) are consistently higher in C_v than tabulated data (blue) with linear interpolation. Dashed lines correspond to a fit of the form aN^b, this represents the histogram behaviour well for all calculated values and is proportional to N^{-1}. Tabulated data does not follow this form until after 4 channels, the reasons are described within the text. From 8-512 channels the tabulated C_v is proportional to N^{-2}

Figure 4.16 shows the C_v, time and database size as a function of the number of bins/channels for the distribution associated with an incident neutron of 14.1 MeV. All computational metrics are normalised by the result for the 32 equal-probability histogram format. This removes the variation between files where the number of energy entries vary and any computational differences between systems and methods. The normalised runtime and database size show the potential gains or losses with respect to the industry standard format. The runtime for tabulated and histogram formats behave as expected from Figure 3.13 and the discussion in Section 3.3.1. The histogram runtime is independent of the number of bins, whereas sampling from tabulated data increases linearly with the number of channels. The database sizes increase linearly with the number of points for both formats. For a 1.0% C_v on the 14.1 MeV elastic scattering distribution in ^{56}Fe, the requirements for each format are marked on the figure and the corresponding computational metrics calculated from this value. To achieve this level of accuracy, histogram format requires 428 bins and tabulated format requires 100 channels. There is no change in sampling time for the histogram format, but optimised tabulated data results in a 1.14 times increase. The memory requirement would increase in both cases, 13 times larger for the optimised histogram data and 3 times larger for the optimised tabulated data.

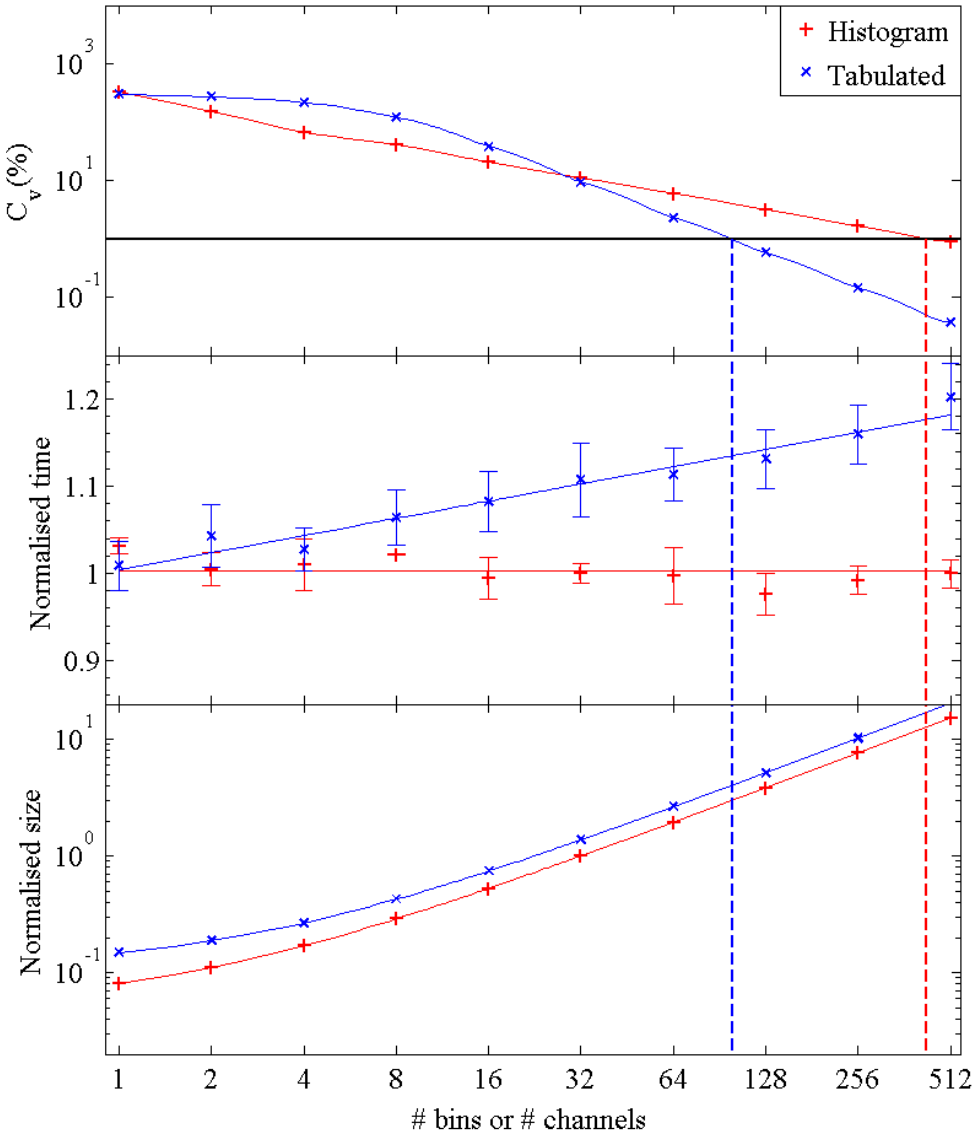

Figure 4.16: Combined data for point-wise analysis and Turnip runtime data. Analytical C_v values are calculated for 1-512 equal-probability histogram bins, and 1-512 tabulated channels. Runtime and database size are normalised by the current 32 equal-probability histogram format. Given a 1.0% C_v tolerance, the expected gains/losses in runtime and database size are shown with the dashed lines [73].

Slab geometry

The previous Monte-Carlo simulations demonstrate the use of identical event sampling to study how data formats affect the simulation in the Lagrangian system. In a practical radiation transport scenario, various Eulerian quantities are of interest such as particle currents, energy spectra and doses. The implementation of transport, geometry and multiple reaction types within Turnip enabled the study of the effects of data pre-processing on these quantities. The currently available reaction channels are elastic (MT2) and inelastic (MT51) collisions of neutrons in simple geometries.

A solid, unit-density block of ^{56}Fe, with dimensions of 2 MFP square and 5 MFP deep was created with intermediate tallying surfaces as shown in Figure 3.17 (Section 3.3.2). These simulations were run with 10^7 mono-energetic and mono-directional source neutrons. Each 14.1 MeV neutron was incident upon the front face of the block at (0,0,0), with direction vector [0,0,1].

Several different data formats were trialled as shown in Table 4.2. These distributions were chosen to demonstrate the best, worst, intermediate and current scenarios.

Table 4.2: Tested formats for the slab and sphere geometries. The C_v values shown correspond to the comparison of the various processed formats to the functional form for the ^{56}Fe(n,el) reaction channel at 14.1 MeV.

Format	C_v (%)	**Comments**
32 histogram	11.1	MCNP legacy method
Opt. tabulated	1.0	MCNP new method
Functional	0.0	Exact representation of input, best case
1 histogram	318.1	Isotropic in μ, worst case
8 histogram	41.2	Intermediate case
1 tabulated	297.7	Linear in μ, worst case
8 tabulated	119.2	Intermediate case

Section 4.1.2 showed that for a large number of scatters, regardless of the input format, the final energy and angular distributions in the lab frame tend to the equivalent of an isotropic distribution in μ. The severely degraded distributions, of one bin/channel, were

chosen to demonstrate the worst possible case in terms of accuracy, but the best case in terms of memory requirements.

Figure 4.17 shows the integrated surface currents mapped onto the slab geometry for the three tested histogram formats. The 32-equal probability histogram data are representative of the functional and optimised tabulated formats. The 8 equal-probability histogram is representative of the 8 channel tabulated format and the one bin histogram results are closely replicated with the one channel tabulated.

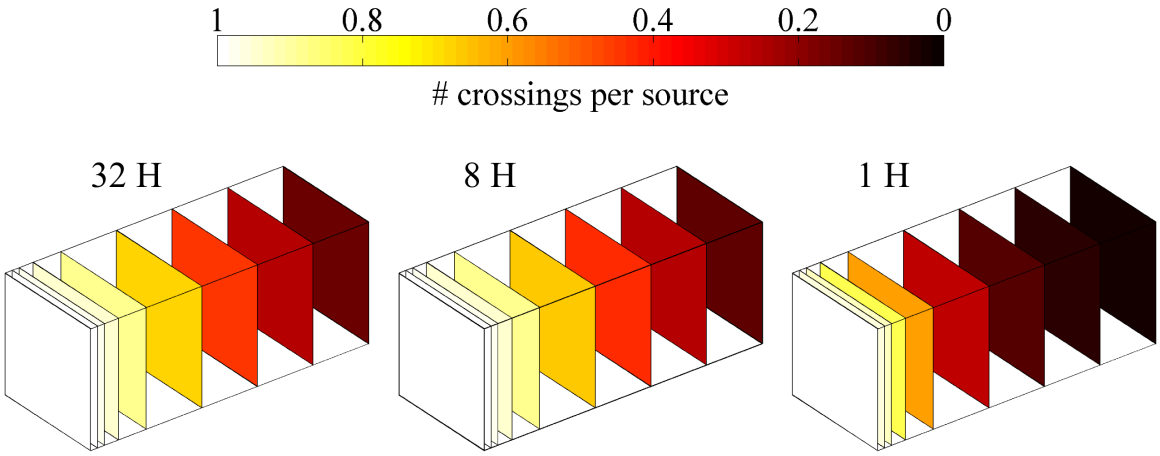

Figure 4.17: Total number of surface crossings per source neutron mapped as colour on to the slab geometry. The first plot shows the integrated values for the 32 equal-probability histogram format, the second for the 8 bin histogram and the third for the isotropic, one bin histogram.

Figure 4.18 shows the total number of surface crossings in the forward direction with respect to depth in the slab. Both of the MCNP formats are identical to the functional values within statistical errors. The 8 bin histogram and 8 channel tabulated formats agree with the functional values, within 1.0%, up to one mean free path. Past this point, the histogram underestimates the functional values by a maximum of 4.4% at 5 mean free paths and the tabulated data underestimates by a maximum of 30.4% at 5 mean free paths. The equal-probability format better represents the forward bias for this number of bins/channels (as also demonstrated within the calculated C_v values in Table 4.2), resulting in a smaller deviation from the functional values. The one bin histogram and

one channel tabulated formats under predict the functional values by 88.2% and 83.1% respectively at 5 mean free paths.

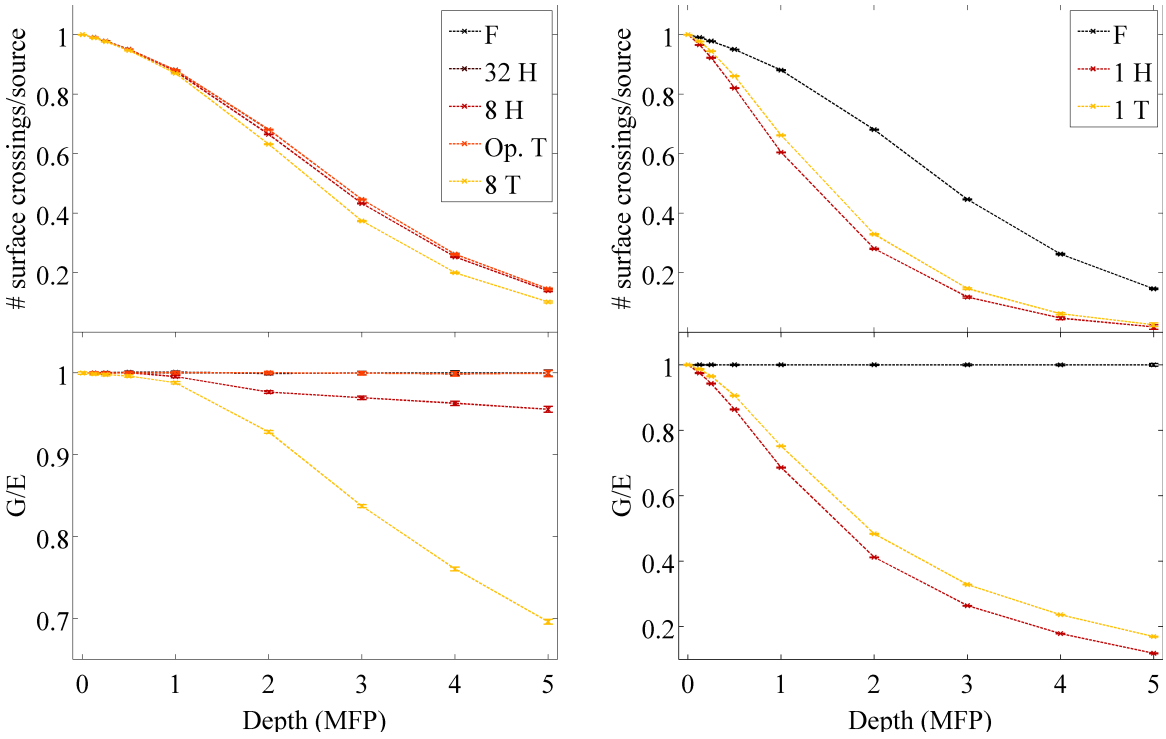

Figure 4.18: Total number of surface crossing events in the forward direction, per source neutron, with respect to depth in the slab geometry. The lefthand figures show the values for functional (F), 32 equal-probability histogram (32 H), 8 equal-probability histogram, optimised tabulated (Op. T) and 8 channel tabulated (8 T). The top plot shows the number of crossings in the forward direction per source and the bottom plot shows the comparison of processed (G) results with the exact, functional form (E) for each surface. The righthand figures show the same quantities for the functional (F), one bin histogram (1 H) and one channel tabulated (1 T) formats.

The two MCNP specific formats, and even the 8 bin histogram give a close representation of the functional data with respect to integral values. The one bin/channel formats and 8 channel tabulated representation significantly underestimate the number of surface crossings when compared to the functional data.

The cause of these differences are better understood with reference to Figure 4.19. This figure shows the tracks for the first 100 source neutrons throughout the geometry. For the 32 and 8 equal-probability histogram formats, the tracks diverge with depth but maintain an overall forward direction, with few back scattering events. The one bin histogram format shows a more isotropic distribution, with many surface crossing events between 0 and 2 mean free paths, but few beyond this point. All sampled μ values are equally probable, so there is a lower probability that neutrons will scatter in the forward direction and a higher probability that it will scatter out of the sides when compared to the other formats. Any events exiting through the bounding planes of the slab are terminated, if they exit through the x- and y-planes they will not be tallied and hence reduce the statistics.

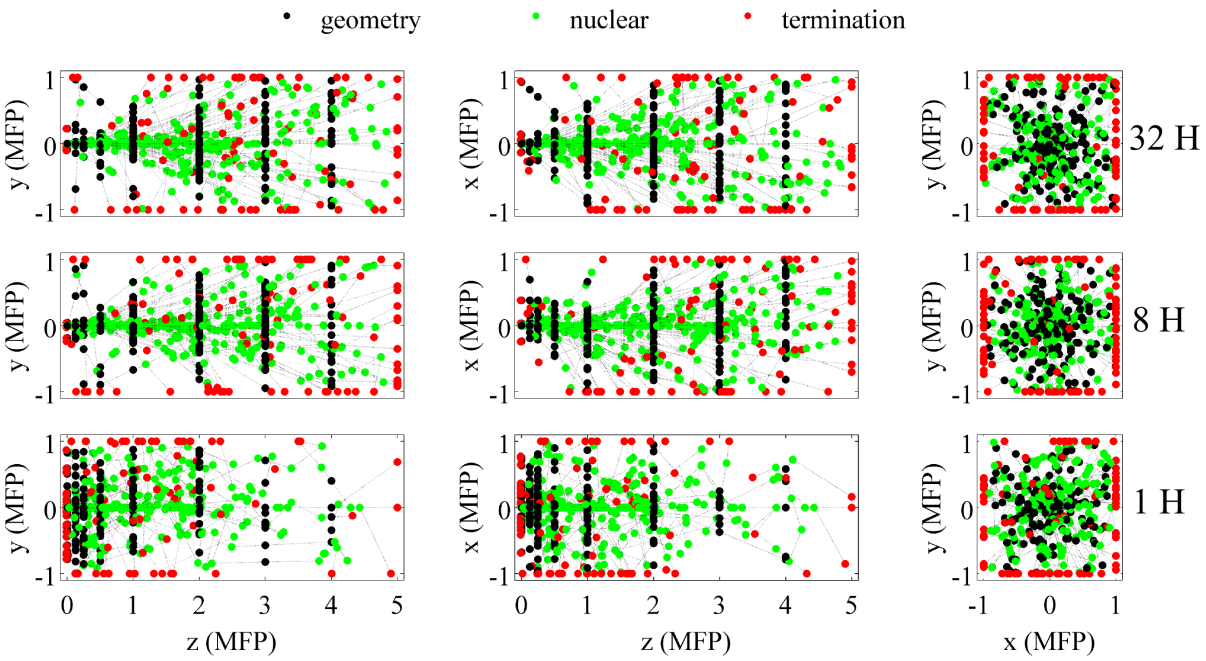

Figure 4.19: First 100 neutron tracks in the slab geometry. Points are coloured according to event type, surface crossing (black), nuclear collision (green) and termination (red). Top row shows the first 100 tracks sampled from the 32 equal-probability histogram, middle row is sampled from the 8 bin histogram and the bottom row shows the 1 bin histogram.

The differences between the two MCNP formats do not become overly apparent until the surface current is considered as a function of energy. The energy dependent, forward currents for 100 energy bins between 10.0 MeV and 14.1 MeV can be seen in Figure 4.20 for one mean free path. These are shown alongside the input distributions and the comparison with respect to the functional results.

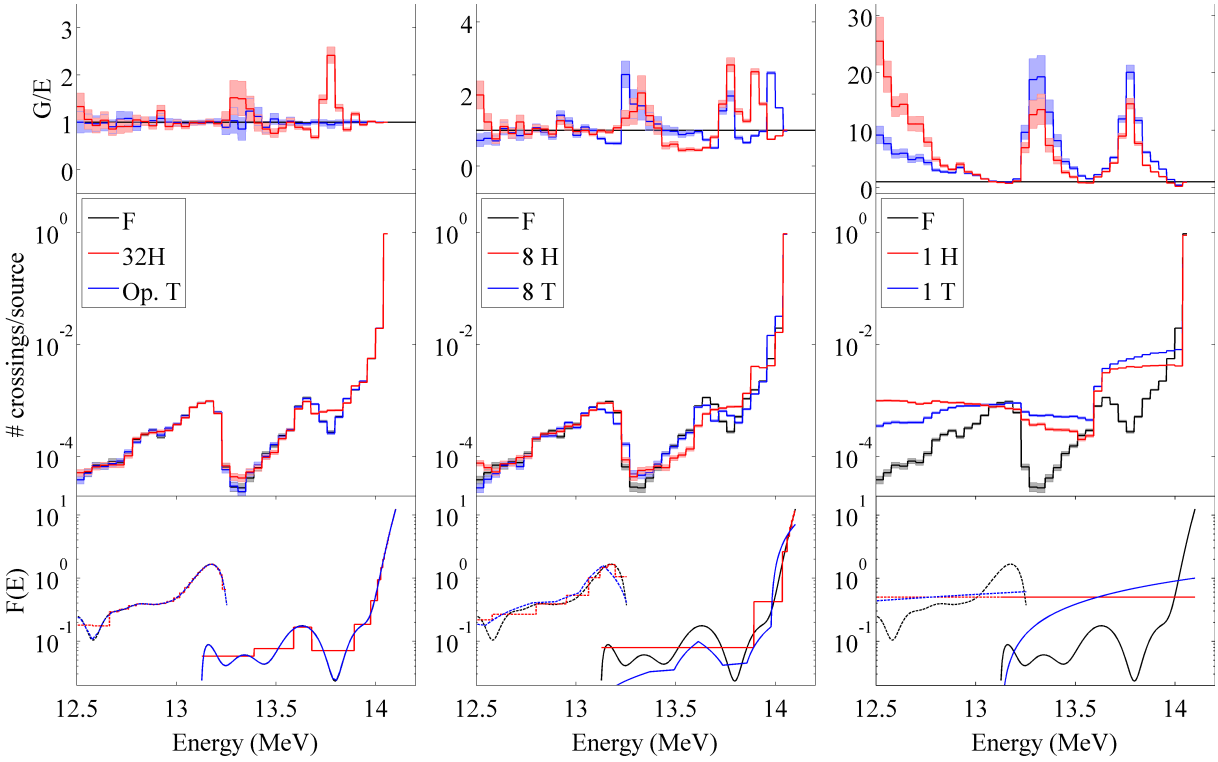

Figure 4.20: Three figures show the comparison of energy dependent surface currents in the slab at one mean free path. The first column corresponds to the current MCNP formats, the second shows the intermediate 8 bin/channel format data and the third column shows the severely degraded data formats. In all three figures the top plots compare the processed formats (G) to the unprocessed, functional format (E), and the middle plots show the number of surface crossings as a function of energy for the various formats. The bottom plots show the expected energy distributions from a 14.1 MeV source neutron, as calculated from the input data, for elastic (13.1-14.1 MeV, solid lines) and first level inelastic (12.7-13.2 MeV, dashed lines).

The 32 equal-probability histogram over-estimates the 13.6 MeV region ($\mu = 0.4$) by $240 \pm 21\%$. This is close to the maximum difference in the input distributions (280%). For the given energy bins, the difference between optimised tabulated data and the original functional form is indiscernible. The 8 bin and channel formats result in a maximum

overestimate of $280 \pm 18\%$ and $260 \pm 11\%$ respectively. The severely degraded one bin and channel formats overestimate the functional by $2550 \pm 420\%$ and $2010 \pm 120\%$ at the worst points.

Within the energy spectra for all formats, the number of surface crossings between 13.2 and 13.6 MeV is much lower than expected from the input distributions. This is an artefact from the geometry, where a scatter of $-1 < \mu < 0$ will remove a source neutron from the beam, where it may escape through the x- and y-planes without being scored further. The behaviour is most pronounced within the severely degraded one bin/channel formats.

All tested formats show the feature in the energy distribution between 12.7 and 13.2 MeV to varying degrees. This corresponds to the range of exit energies for a neutron after a first-level, inelastic scatter in ^{56}Fe at 14.1 MeV.

For the 100 energy bins, the C_v was calculated for all formats as a function of depth, as compared to the functional equivalent and is shown in Figure 4.21. Up to 1 MFP, the C_v increases with depth. This is due to the increasing number of scatters and reduction in unreacted beam. The C_v tends to a constant level, within errors, past two mean free paths and the values are given in Table 4.3.

This suggests that, should a similar simulation be run to the same level of convergence, using processed data could result in an additional uncertainty of up to 1.9% in the MCNP formats, or up to 247.5% for the alternative formats. The observed differences in the energy spectra could have far reaching consequences in terms of moderation, tritium breeding and activation analyses.

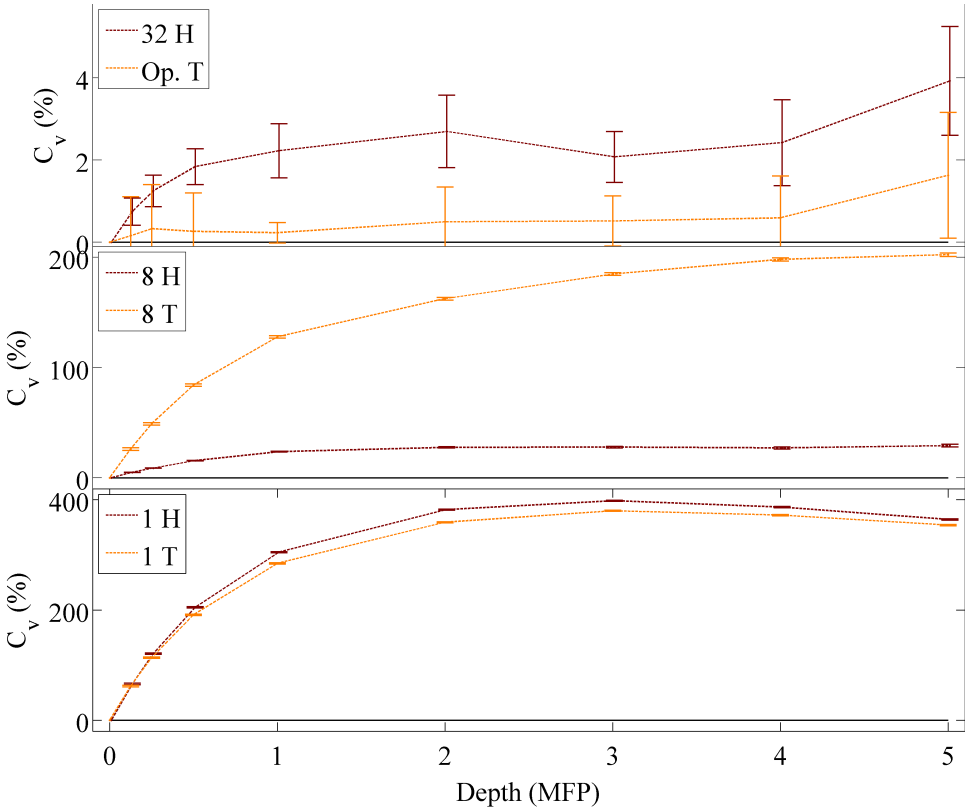

Figure 4.21: Calculated C_v as a function of depth for the slab geometry calculated from the energy spectra. The top plot shows the C_v for the 32 equal-probability histogram and optimised tabulated formats compared to the functional equivalent at each surface. The middle plot shows the C_v for the 8 bin histogram and 8 channel tabulated formats. The bottom plot shows the C_v for the one bin histogram and one channel tabulated. Dashed lines are to guide the eye only. From 0.125 to 1 MFP, the number of scatters increases and as a result the C_v increases. From 1 to 5 MFP the C_v tends to a constant level. The calculated errors on the C_v are determined from the propagation of the statistical uncertainty from each bin.

Table 4.3: Calculated C_v for each format in the slab geometry when compared to the functional energy spectra at each surface. Values are given for the region between 2 and 5 mean free paths, where the C_v reaches an approximate constant value

Format	Final C_v(%)
32 histogram	1.9 ± 0.4
8 histogram	18.3 ± 3.5
1 histogram	247.5 ± 48.4
Opt. tabulated	0.5 ± 0.1
8 tabulated	115.0 ± 24.4
1 tabulated	235.2 ± 46.3

Validation and comparison

To justify the use of the tools developed specifically for nuclear data sampling analyses, they must be validated against a well trusted, industry standard. In this case MCNP is considered to be the trusted case based on NJOY processed data libraries and is compared to the Swede-Mash and Turnip combination. There are some differences between the two that need to be addressed before the comparison. Turnip was run with the elastic and first level inelastic reaction channels only. MCNP will sample all available cross-sections within the ENDF file and so neutrons will have a correspondingly shorter mean free path in the same density material. This means that the mean free path of neutrons is not the same in Turnip and MCNP, so the geometry dimensions were altered accordingly. The additional reaction channels include neutron multiplication and the remaining inelastic levels in the case of ^{56}Fe, so the energy dependent surface current is expected to vary below 12.7 MeV (this is the lower energy limit from a single, first-level inelastic scatter). Turnip operates strictly as an analogue Monte-Carlo, whereas MCNP has certain variance reduction techniques implemented by default. Implicit capture was turned off in MCNP to allow the comparison to Turnip. To ensure consistency of tallied results, PTRAC files associated with each MCNP simulation were produced. These files contain the interaction data for every event and an equivalent was produced from the Turnip simulations. The MCNP produced PTRAC files were parsed using Parsnip [74] and then tallied in the same manner as Turnip. Figure 4.22 shows the scaled, energy dependent Turnip and MCNP surface current tallies. Both tallies are given in terms of number per source neutron, but an additional scaling factor was applied to the Turnip data to remove systematic differences between the two codes. This scaling factor was calculated as the integral of the MCNP tally over the Turnip tally for each surface.

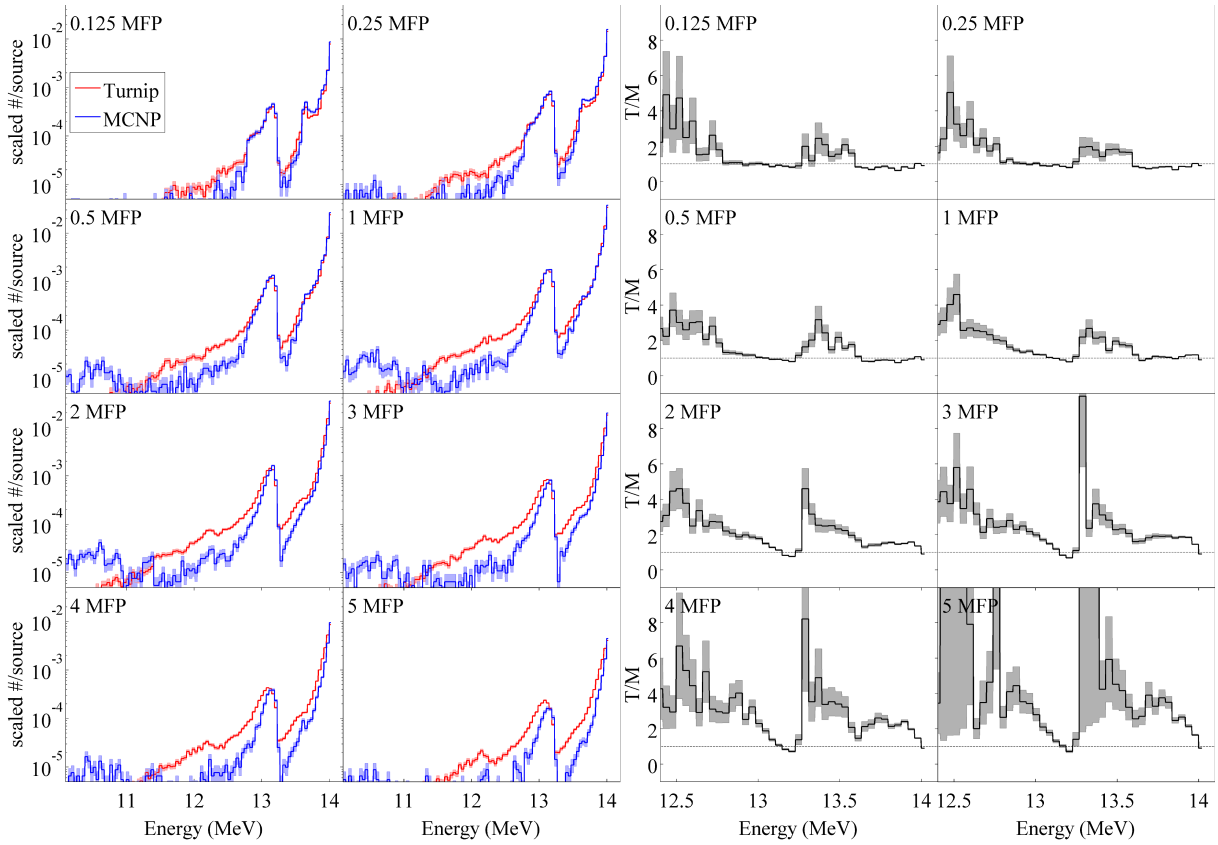

Figure 4.22: The left figure shows the scaled, energy dependent surface currents in the forward direction throughout the slab geometry. Both simulations were run with the 32 equal-probability histogram format data. Turnip values are shown in red, and the MCNP in blue. Both exhibit the forward bias and features expected from the first elastic scatter at 14.1 MeV, as well as the feature from the first inelastic level (12.7-13.2 MeV). Right figure shows the comparison of the Turnip values (T), as normalised by the MCNP values (M) for energies between 12.5 MeV and 14.1 MeV. A baseline is set to one to show the under or over estimate of Turnip with respect to MCNP. Shaded bars represent the statistical uncertainty associated with the distributions.

Within the range of one elastic scatter and the first inelastic scatter, MCNP and Turnip surface currents contain the same features at the same energies. Regions of low probability, and large statistical uncertainty within MCNP are overestimated by Turnip, most noticeably at lower energies (<12.5 MeV). In all cases Turnip tallies a larger integral value than MCNP for the same number of source neutrons. This is the result of the limited number of reaction channels currently implemented within Turnip. MCNP will terminate neutrons if they undergo capture reactions, and any captured neutron is unable to

contribute further to any tallies. MCNP will also sample from the additional reaction channels. MCNP tallies were derived from the PTRAC file rather than the inbuilt tallies. This allowed contributions from elastic and first level inelastic events to be tallied only, at the cost of statistics. This region is of low probability, so the fractional difference between the two is large, but the absolute contribution is small. Figure 4.23 shows the number of events of each type, per source neutron with respect to the exit energy from the reactions within MCNP.

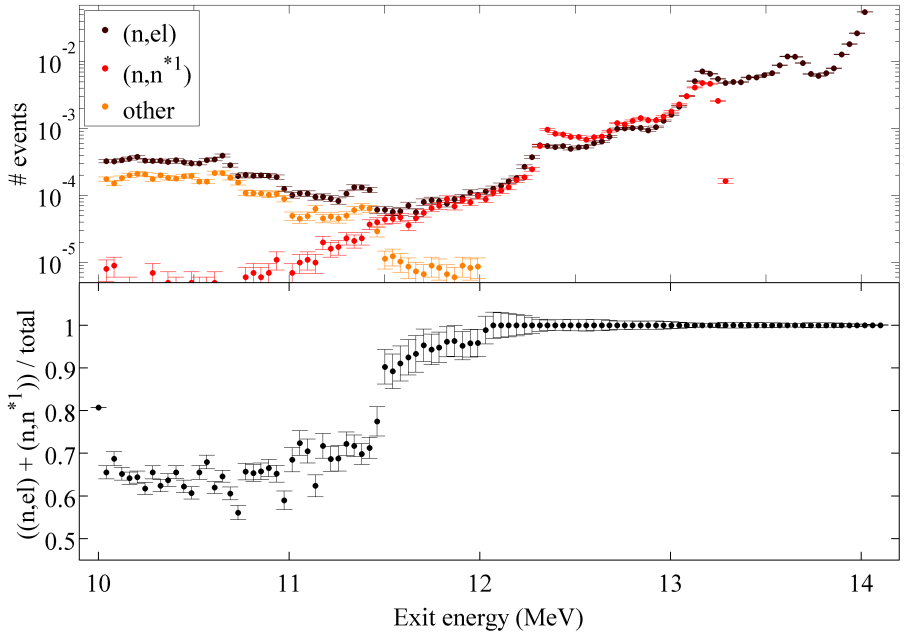

Figure 4.23: Top plot shows the number of events of each type per source over the entire slab geometry based on 14.1 MeV neutrons incident on ^{56}Fe. The abscissa shows the exit energy of the neutron, i.e. after the collision kinematics have occurred. The bottom plot shows the fraction of Turnip reaction channels (elastic and first inelastic) with respect to the total within MCNP. Exit energies above 12.0 MeV are due to elastic and first inelastic levels only in ^{56}Fe. Below 12.0 MeV the number of neutrons resulting from other reaction types increases. These are predominantly threshold reactions (further inelastic and multiplication), so for a 14.1 MeV neutron the maximum exit energy is dependent on the Q-value of the reaction.

The number of elastic scatters is continuous in the energy range shown, where all initial collisions occur at 14.1 MeV. The exit energy of neutrons after a single, first-level inelastic collision have a maximum value of 13.2 MeV. The other reaction channels include the remaining discrete level inelastic scatters, continuum inelastic scatters and neutron mul-

tiplication. These are all threshold reactions, hence the exit energy from these combined events has a maximum of 12.0 MeV. It is worth noting that within Figure 4.23, there is an increased contribution to elastic events below 11.4 MeV compared to the surface current tallies in Figure 4.22. The resulting neutrons produced from the other reaction channels continue to interact within the slab. These neutrons have a maximum energy of 12.0 MeV causing a low energy increase in event rate. These scattering events resulting from anything other than elastic and first level inelastic are discarded in the comparison of Turnip to MCNP.

The event rate of each interaction type with respect to depth in the slab was measured and is shown in Figure 4.24. These event rates are integrated over all energy values and given in units of interactions per source neutron. Throughout the volume, elastic scatter is consistently the most common reaction channel. Fewer events are recorded within the first few segments of the slab due to the reduced volumes of the segments, i.e. 0.125 MFP deep in the first segments and 1.0 MFP deep in the final segment. This agrees with the surface currents and worsening statistics with depth in the slab (Figure 4.22).

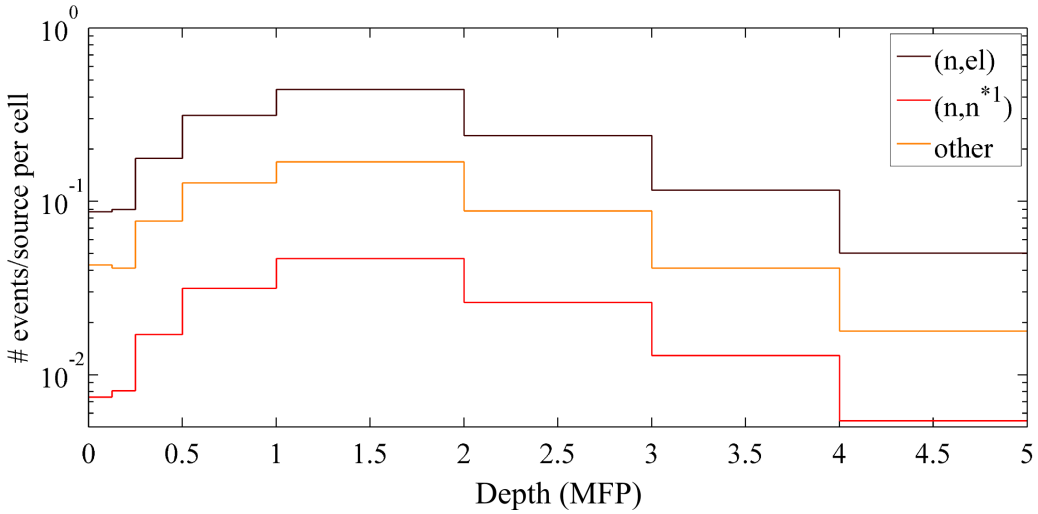

Figure 4.24: Event rate of reaction channels in ^{56}Fe with respect to depth in slab geometry. Integral values are calculated for each volume and normalised per source. Throughout the geometry, elastic and first-level inelastic reactions contribute between 68 and 76% of the total number of events.

Spherical geometry

The same analyses were performed with the spherical geometry as with the slab geometry. All 10^7 source neutrons are born at the origin of the sphere with an energy of 14.1 MeV and an isotropic angular distribution. The main difference between the geometries is that all source neutrons are contained. The only way in which they can escape and terminate is through a minimum of 5 mean free paths of material and all neutrons will score on each surface at least once. Figure 4.25 shows the total surface currents through the tallying spheres for the three variations of the histogram formats. These are representative of the remaining formats, where 32 bin, optimised tabulated and functional data all exhibit similar integral behaviour, as with the 8 bin/channel formats and one bin/channel formats. There are no obvious differences between the 32 bin and 8 bin histogram formats, but the one bin histogram exhibits very different behaviour.

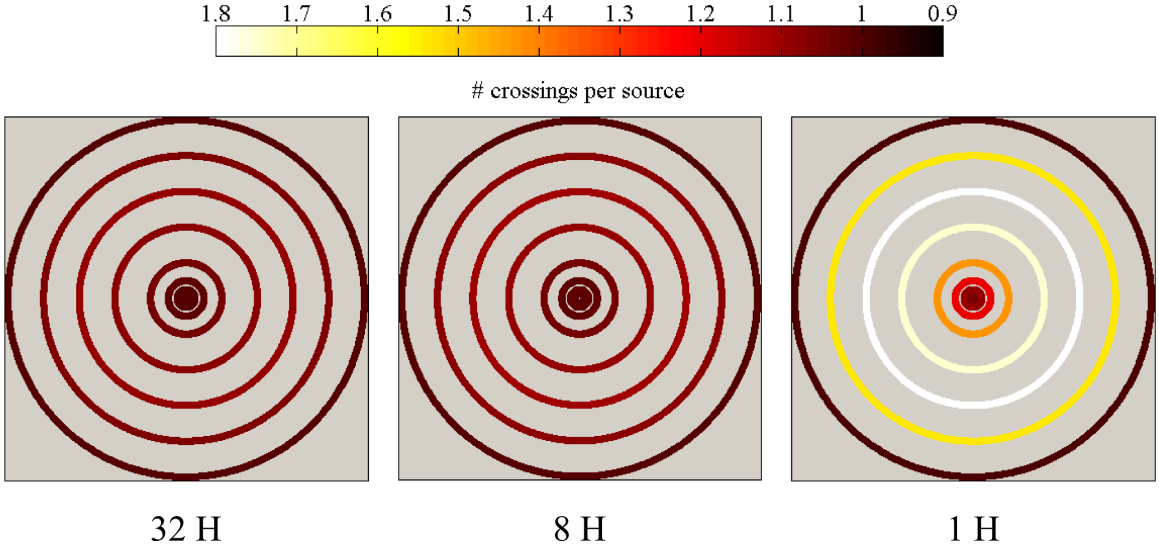

Figure 4.25: The total surface current over each spherical surface is coloured according to the number of surface crossings per source. Results shown for the 32 bin, 8 bin and 1 bin histogram distributions. All surfaces within the sphere have more than one crossing per source, with the exception of the external bounding surface. All neutrons leaving this surface terminated and are unable to return. One bin histogram sampling exhibits the greatest number of surface crossings, whereas the 32 and 8 bin are similarly behaved.

Figure 4.26 shows the comparison of integral values for all formats with respect to radius in the sphere geometry. There are no differences in the total surface crossings for functional, 32 equal-probability histogram and optimised tabulated formats within errors.

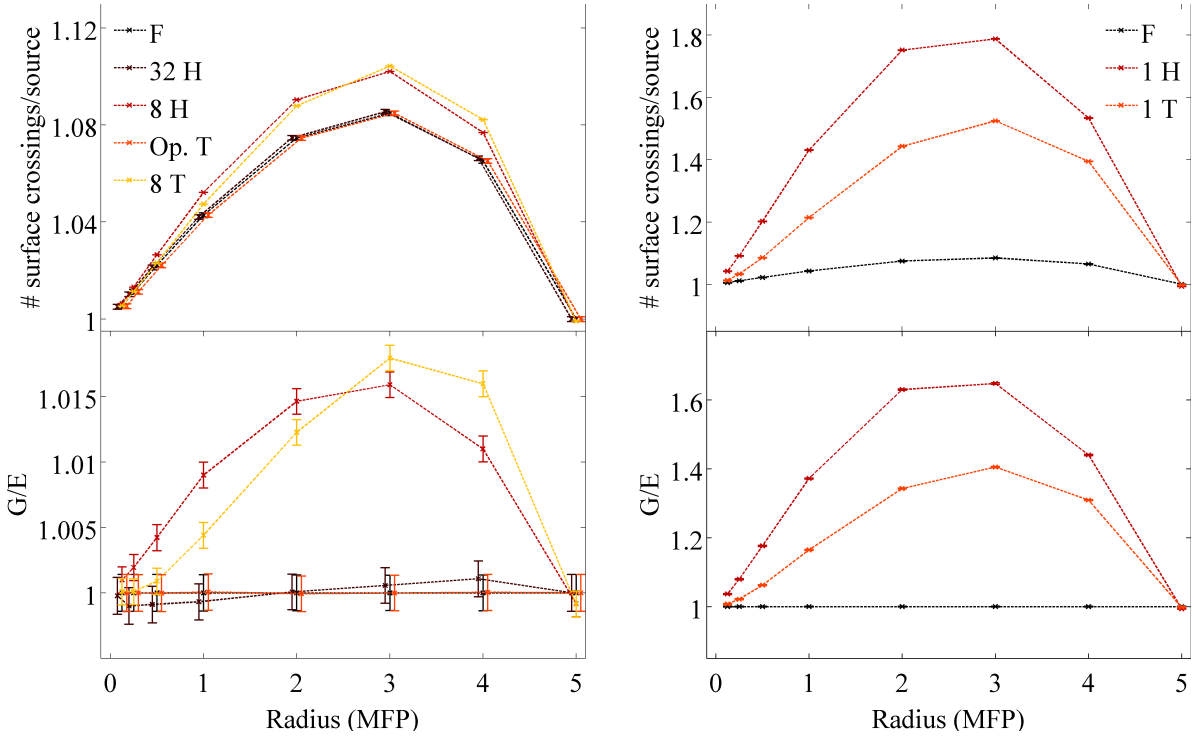

Figure 4.26: The total number of surface crossings for each format as a function of radius are shown in the top figures. The bottom figures show the processed result (G) as a fraction of the functional result (E). Left figure shows the distributions most closely replicating the functional result, whereas the right figure shows the three worst formats. Surfaces show an increase in crossings with radius up to 3 mean free paths due to the increase in surface area. Past this point neutrons are able to escape the geometry.

The 8 bin histogram format has a maximum difference of 1.02 ± 0.01 at 3 mean free paths, i.e. it overestimates the integral value by 2%. The 8 channel tabulated format has a maximum difference identical to the histogram equivalent within the given number of significant figures. The overall trend of the integral surface current for 8 channel tabulated data differs to the other distributions. The maximum difference still occurs at 3 mean free paths, but the distribution is biased towards the higher radii. The severely degraded one bin histogram and one channel tabulated formats overestimate the surface current within the main body of the sphere, with maximum differences of $65 \pm 1\%$ and

$41 \pm 1\%$ respectively. At a radius of 5 mean free paths, all processed formats agree with the functional data within errors. All neutrons born within the sphere eventually exit through this bounding surface, beyond which they are terminated.

In all cases, the maximum number of surface crossings occurs at 3 mean free paths. At this point, 71% of all source neutrons will have interacted. The maximum number of surface crossings here is a balance of scattered neutrons as there is no direction information tallied. Any neutrons which backscatter will likely pass through the previous surfaces at least one more time. If the neutron then undergoes another collision, the change in direction could cause additional surface crossings. Up to and including 3 mean free paths, there is a low chance for the neutron to escape (5%). Between 4 and 5 mean free paths, the neutrons have a much higher probability of leaving volume without further interactions reducing the number of surface crossings.

The main sources of these differences are demonstrated within the transport for the first 100 neutron tracks in Figure 4.27. The isotropic, or closely isotropic distributions have a far larger number of nuclear events within the main body of the sphere, resulting in a higher overall surface current. Distributions with a distinct forward bias have a higher probability that a neutron will interact and continue in the forward direction where it may escape.

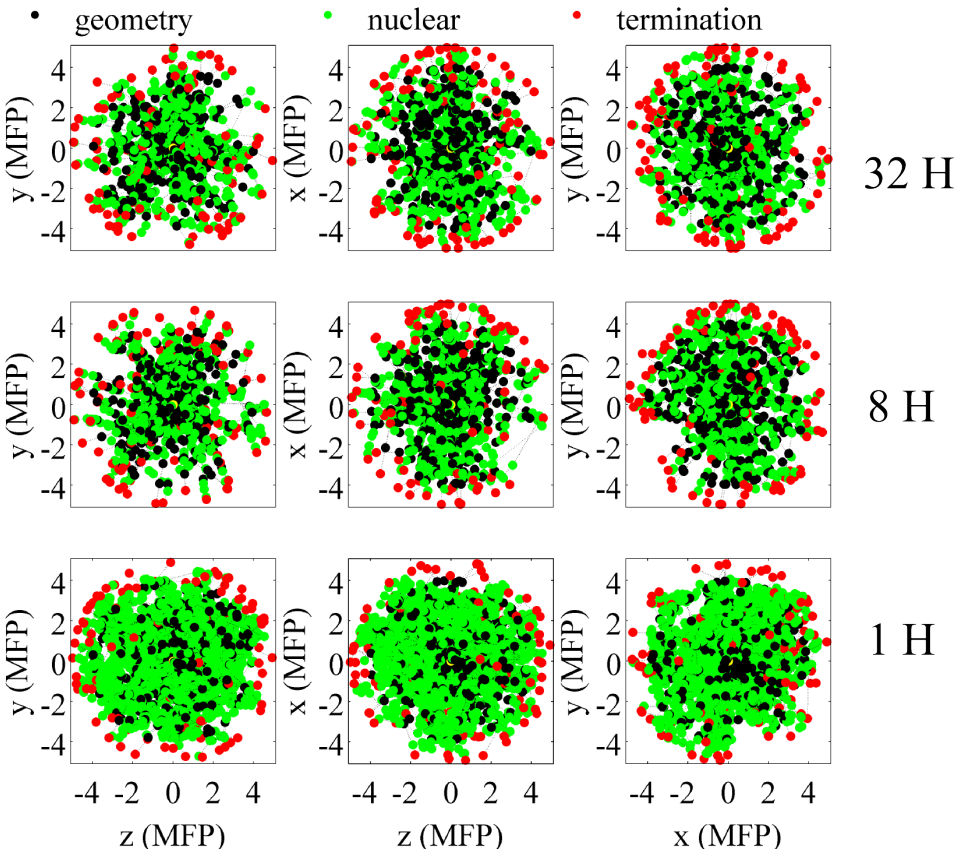

Figure 4.27: First 100 neutron tracks in the sphere geometry. Points are coloured according to event type, surface crossing (black), nuclear collision (green) and termination (red). Top row shows the first 100 tracks sampled from the 32 equal-probability histogram, middle row is sampled from the 8 bin histogram and the bottom row shows the 1 bin histogram.

For a radius of one mean free path, the energy dependent surface currents are shown in Figure 4.28. The closed surfaces of the spherical geometry result in a lower escape probability, so the neutrons are transported to lower energies than are seen in the slab geometry (Figure 4.20). The statistics are improved at lower energies, but the features associated with the first elastic and first-level inelastic scatters are less prominent. The integral behaviour of the 8 channel tabulated distributions (Figure 4.26) is explained with reference to its energy dependent behaviour. The region between 13.5 MeV and 13.9 MeV overestimates the functional distribution by 285%. This suggests that a larger proportion of neutrons are scattering into this region and hence the probability of multiple forward

scatters, and escape, is higher. The back scattering probability is correspondingly lower, so fewer neutrons will return through the preceding surfaces.

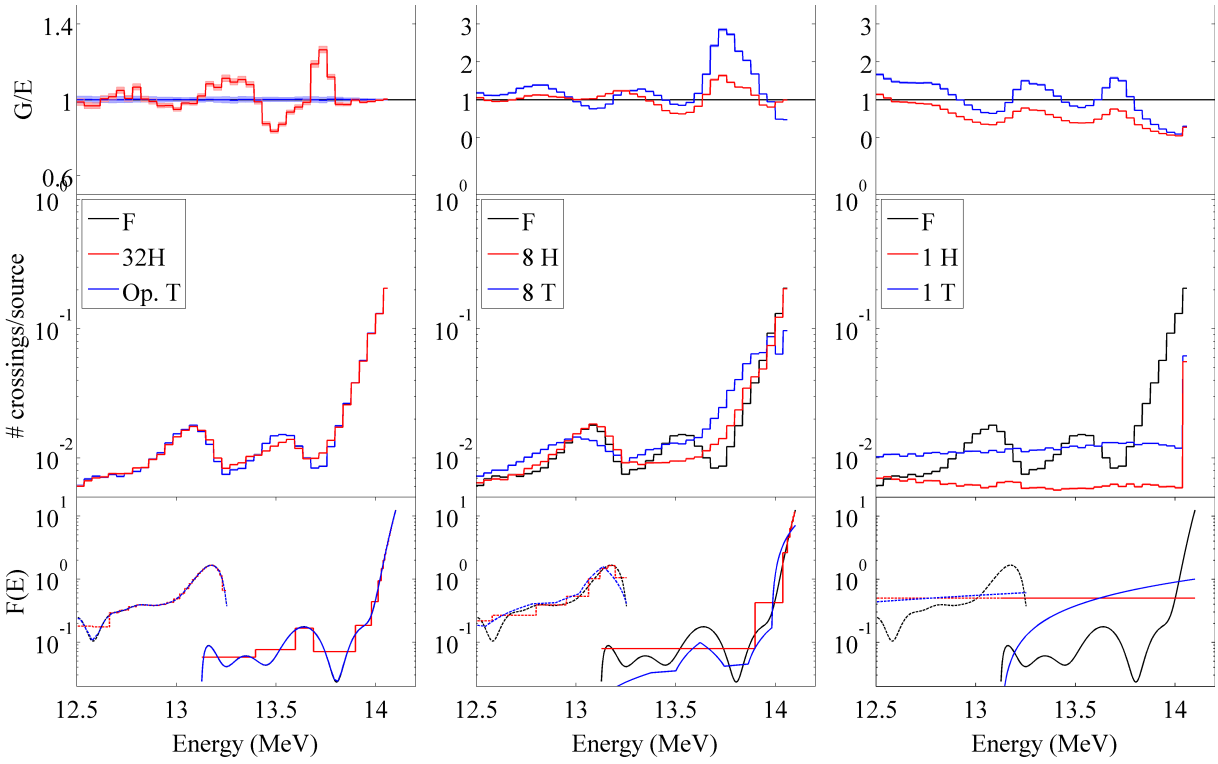

Figure 4.28: Three figures show the comparison of energy dependent surface currents in the sphere at a radius of one mean free path. The first column corresponds to the current MCNP formats, the second shows the intermediate 8 bin/channel format data and the third column shows the severely degraded data formats. In all three figures the top plots compare the processed formats (G) to the unprocessed, functional format (E), and the middle plots show the number of surface crossings as a function of energy for the various formats. The bottom plots show the expected energy distributions from a 14.1 MeV source neutron, as calculated from the input data, for elastic (13.1-14.1 MeV, solid lines) and first level inelastic (12.7-13.2 MeV, dashed lines).

The C_v of the resulting energy spectra for the different formats, with respect to the functional tallies, can be seen in Figure 4.29. Optimised tabulated data has an overall C_v of zero, i.e. the distributions from functional and optimised tabulated formats are identical within errors.

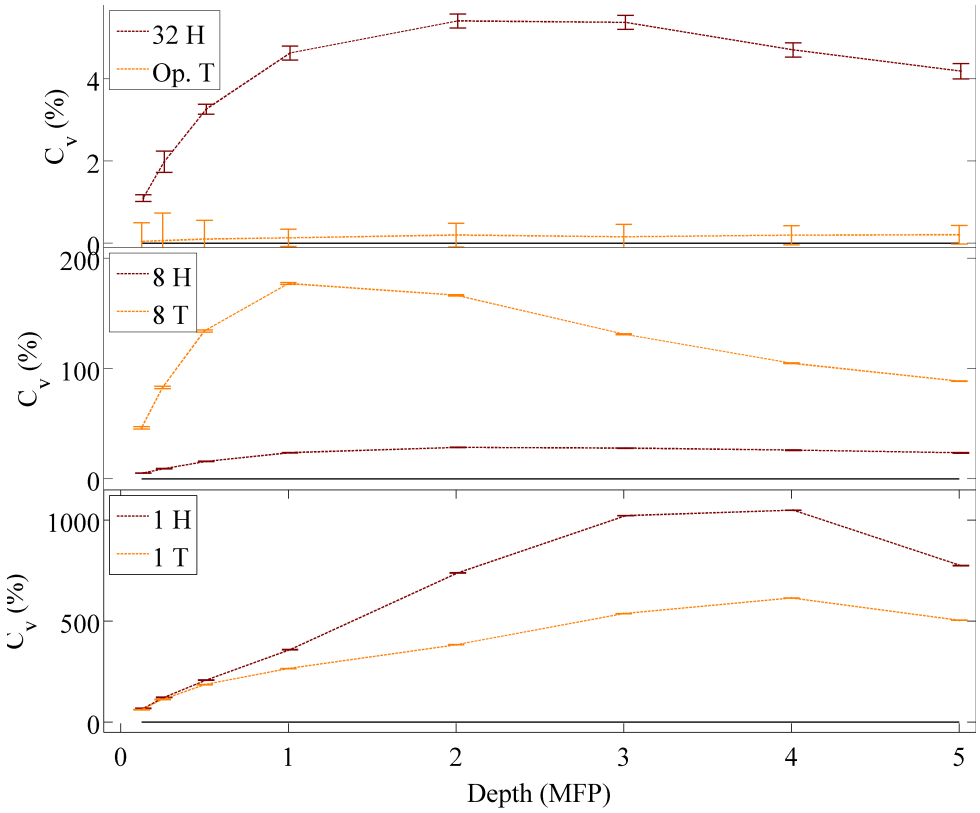

Figure 4.29: Calculated C_v as a function of radius for spherical geometry, as determined from the energy dependent neutron currents. The surface with the maximum C_v for each format varies, but all show an initial increase with radius up to the maximum value followed by a decrease in C_v by 5 mean free paths.

The maximum and final C_v values are given in Table 4.4. The calculated C_v values are reduced at 5 mean free paths due to each neutron crossing the surface once only at termination. The optimised tabulated format best replicates the functional distributions, followed by the 32-bin histogram format. There is no common scaling factor between these C_v values for the energy spectra and the C_v calculated from the input distributions (given in Table 4.2). However they do follow the general trend of a low C_v from input data corresponds to a low C_v in energy spectra and vice versa.

Table 4.4: Calculated C_v for each format in the spherical geometry when compared to the functional energy spectra at each surface. Values are given for the maximum, along with location, and the final C_v at the bounding sphere (radius 5 mean free paths)

Format	Maximum C_v(%)	Radius (MFP)	Final C_v (%)
32 histogram	5.4 ± 0.2	2	4.2 ± 0.2
8 histogram	28.5 ± 0.3	2	23.5 ± 0.3
1 histogram	1048.9 ± 1.0	4	774.5 ± 0.9
Opt. tabulated	0.2 ± 0.3	2	0.2 ± 0.2
8 tabulated	177.2 ± 0.9	1	88.6 ± 0.3
1 tabulated	614.4 ± 0.8	4	504.3 ± 0.7

Validation and Comparison

As with the slab geometry, the Turnip and MCNP energy distributions demonstrate the same features at the expected energies. Figure 4.30 shows the scaled surface currents for each of the surfaces within the sphere geometry. At small radii ($<$ 1 MFP), the two codes agree. At larger radii, Turnip overestimates the lower energy neutron current in comparison to MCNP. This same effect was seen in the slab data and is due to the lack of capture and other reaction channels within Turnip. The reaction channels within the MCNP simulation are the same as shown in Figure 4.23. The reduced statistics at low energy in MCNP limits the degree to which the two codes can be compared as it alters the scaling factor between the two.

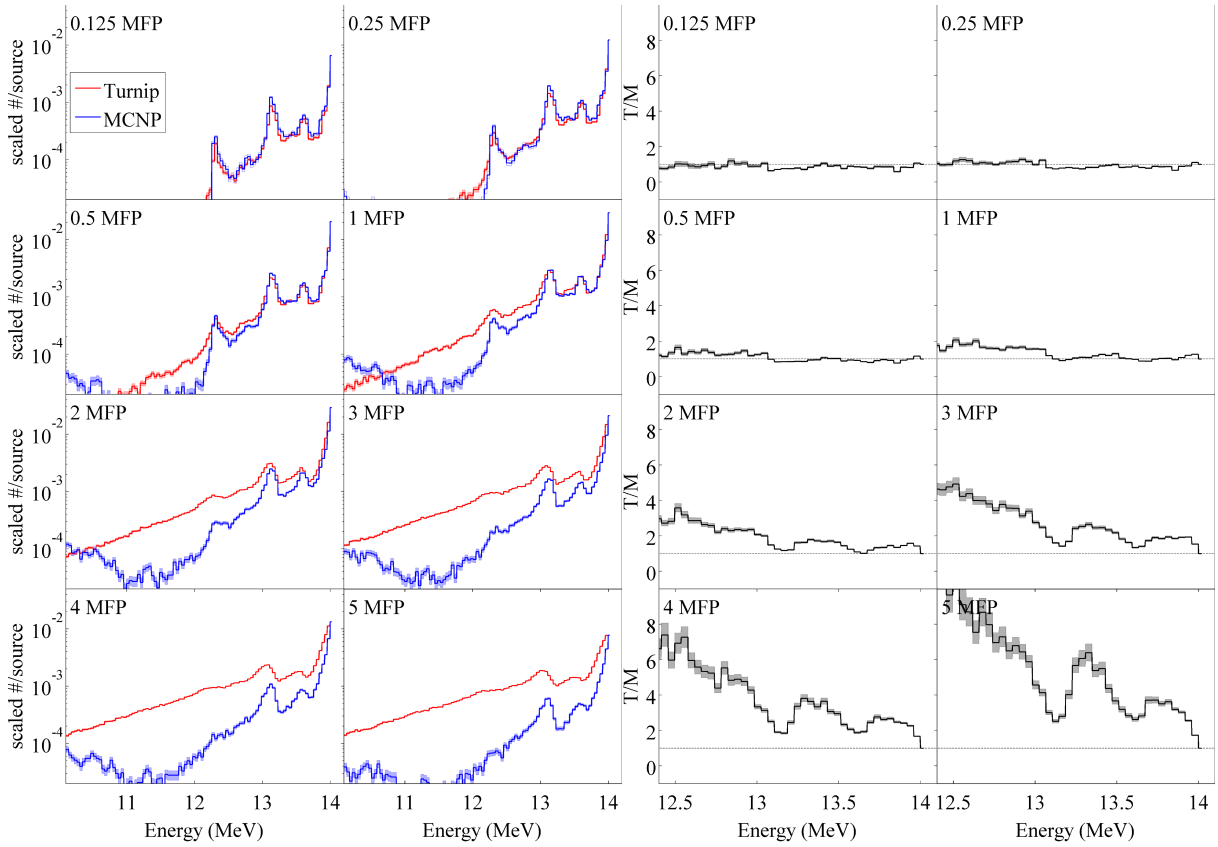

Figure 4.30: Normalised surface currents for Turnip (red) and MCNP (blue) are shown on the left for increasing radii. At small radii, with few interactions the two closely match. The difference increases with depth. This is due to the Turnip results being skewed by an increased number of lower energy events. MCNP results are lower in this region, with poor statistics due to the removal of capture reactions from the tally. The figure on the right shows the comparison of the Turnip (T) energy spectra with respect to the MCNP (M) energy spectra. A baseline is set at one and the shaded bars represent the statistical uncertainty.

4.1.3 Summary

The presented results for the example material of ^{56}Fe have demonstrated that the pre-processing of angular distributions does impact upon the Monte-Carlo simulation results at fusion energies. For the elastic scatter of neutrons at 14.1 MeV, the two MCNP accepted formats, of 32 equal-probability histogram and optimised tabulated, differ from the unprocessed distributions by 11.1% and 1.0% respectively. In general, this decreases with energy. When propagated within a Monte-Carlo simulation with a forced number

of scatters, the difference in the lab frame angular distributions become negligible past 4 events. However, the differences between the various formats remain visible within the resulting energy spectra for multiple events. After 9 scatters, the legacy 32 equal-probability histogram format is associated with a C_v of 6.7%. In all cases, the optimised tabulated format replicates the unprocessed distributions within statistical limits.

Variations upon the histogram and tabulated formats were tested to determine the effect upon simulation parameters. To achieve an accuracy optimised data set to be within 1.0% of the original distribution, the ^{56}Fe(n,el) reaction channel at 14.1 MeV would require a 13.4 times increase in the number of bins when compared to the current 32. For the equivalent of optimised tabulated data, the memory requirement increases 3.1 times. In both cases, an increased accuracy is at the cost of memory requirements, though the computational cost in sampling time is negligible for the histogram format, but is 14% larger for the tabulated format.

Several format variations were used to determine the sensitivity of full radiation transport simulations to the pre-processing. Degraded histogram and tabulated formats were used with the MCNP formats for comparison. The two MCNP formats showed little variation within the integral quantities, but the differences remained apparent when the energy spectra were considered. At the worst case, the 32 equal-probability histogram format differed from the unprocessed equivalent by 5.4% within the spherical geometry. Of the degraded formats, the 8 bin histogram was consistently better than the alternatives. The maximum difference between the integral values for the functional and 8 bin histogram were 4.4% and 2.0% within the slab and sphere geometries. Within the energy spectra, the differences were more apparent with a C_v of 28.5% as the worst case.

From these results it is clear that for ^{56}Fe(n,el) at 14.1 MeV the integral quantities considered are insensitive to the two currently used MCNP formats. The discrepancies only become apparent if the energy spectra are considered within the regions of greatest dif-

ference. Degraded data formats affect the integral and energy dependent quantities to varying degrees, though there are some minimal savings in memory requirement and sampling time.

4.2 Overview of multiple isotopes

The previous section demonstrated the effect of pre-processing on ^{56}Fe, predominantly with respect to the elastic scattering reaction channel at 14.1 MeV. This section presents an overview of the additional isotopes that are considered to be important for fusion neutronics within the context of this work.

4.2.1 Point-wise

For the isotopes shown in Table 4.5, the unprocessed data was extracted with Swede-Mash from the ENDF-B/VII.1 library, and the C_v calculated between the unprocessed functional distributions and the various processed formats. For histogram and tabulated data the C_v was calculated with the number of bins/channels as shown in Figure 4.31.

Table 4.5: Isotopes considered as part of this work. For each isotope the order of the polynomial for the unprocessed distribution is given as a guide to distribution complexity. These values are given for the elastic scatter of neutrons distribution energies closest to 14.1 MeV, note that none contain 14.1 MeV. The C_v is given for the legacy 32 equal-probability histogram format in comparison to the unprocessed functional form.

Isotope	Closest energy (MeV)	Order of polynomial	C_v (%)
^{1}H	14.0	8	< 1.0
^{16}O	14.0	10	< 1.0
^{52}Cr	14.2	18	12.8
^{56}Fe	13.6	12	11.1
^{63}Cu	14.5	14	12.7
^{90}Zr	14.0	28	14.5
^{184}W	14.0	34	20.0

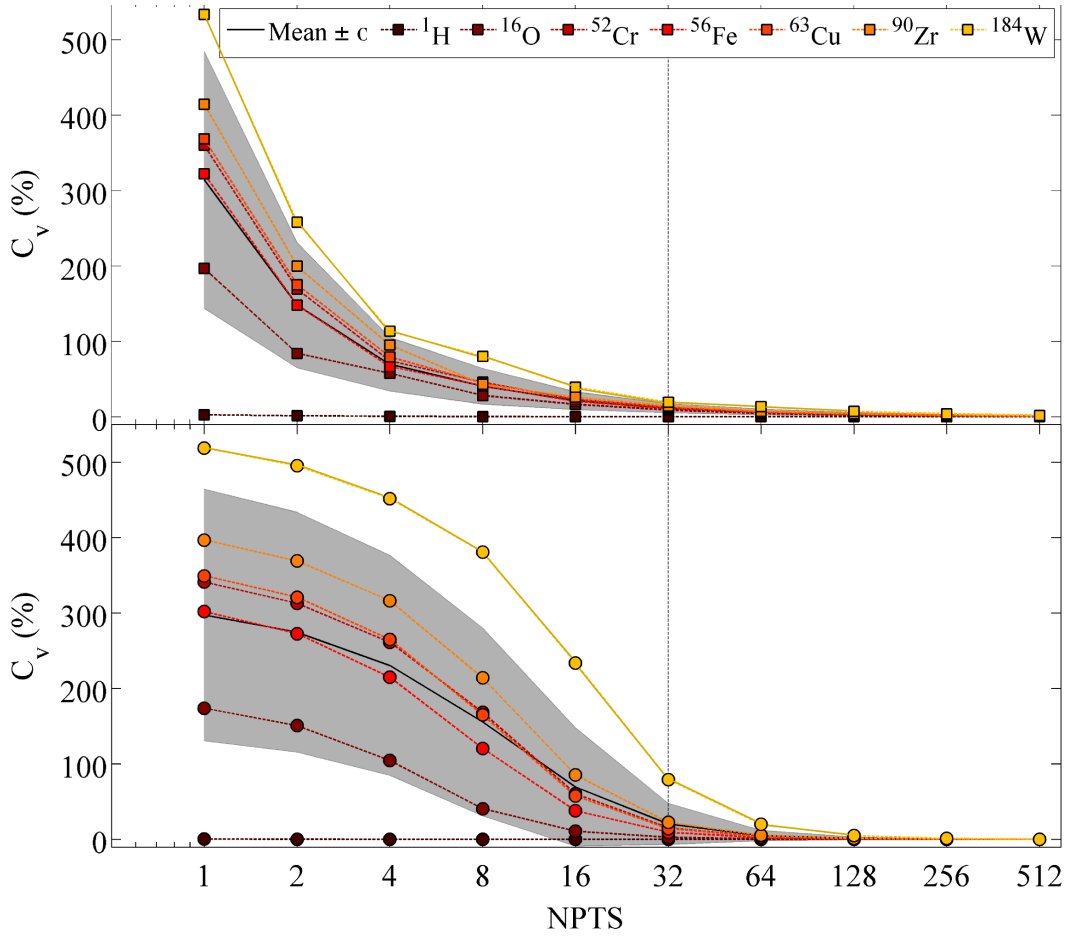

Figure 4.31: Calculated C_v for the seven considered isotopes as a function of the number of bins or channels. The top plot shows the values for the equal-probability histogram format and the bottom plot shows the values for the tabulated format with linear interpolation. The grey bar shows the mean value for all isotopes within one standard deviation. The vertical dashed line indicates the location of the currently implemented 32 equal-probability histogram.

Higher mass isotopes are consistent with a higher C_v due to the additional complexities and features associated with the distributions. Very low mass isotopes, such as ^{1}H are naturally tending towards the isotropic, so do not often require more than one bin/channel to produce a C_v of zero.

Between 1 and 64 bins the histogram format has a larger gradient in C_v than the tabulated format with linear interpolation. After 64 channels the tabulated data results in a lower C_v with additional channels than the histogram format. The tabulated points are equally

spaced in θ, and their location is not optimised. Equal-probability histograms with few bins are better at representing these distributions as the location of the bins is determined based on the initial distribution, i.e. they are not arbitrarily set.

4.2.2 Turnip

The processed data for each of the isotopes given in Table 4.5 were used within Turnip to determine the computational burden associated with the different data formats. These same data sets were used within the radiation transport simulations of the slab and sphere geometries.

Forced scatters

For each isotope, large numbers of identical, forced events were run to determine the runtime and the built database sizes. Figure 4.32 shows the normalised values for C_v, runtime and database size with respect to the current 32 equal-probability histogram format. The original functional data are an exact representation of the ENDF data, so accuracy is gained. For the distributions associated with ^{1}H, the behaviour of the C_v is different from the others. The accuracy optimised formats (histogram and tabulated) result in an increased C_v. The C_v associated with the 32 bin representation of this distribution is 0.1%, so 32 bins are unnecessary to replicate this distribution within a 1.0% C_v limit. This closely isotropic distribution only requires 4 equal-probability histogram bins or 1 tabulated channel to achieve this level of accuracy.

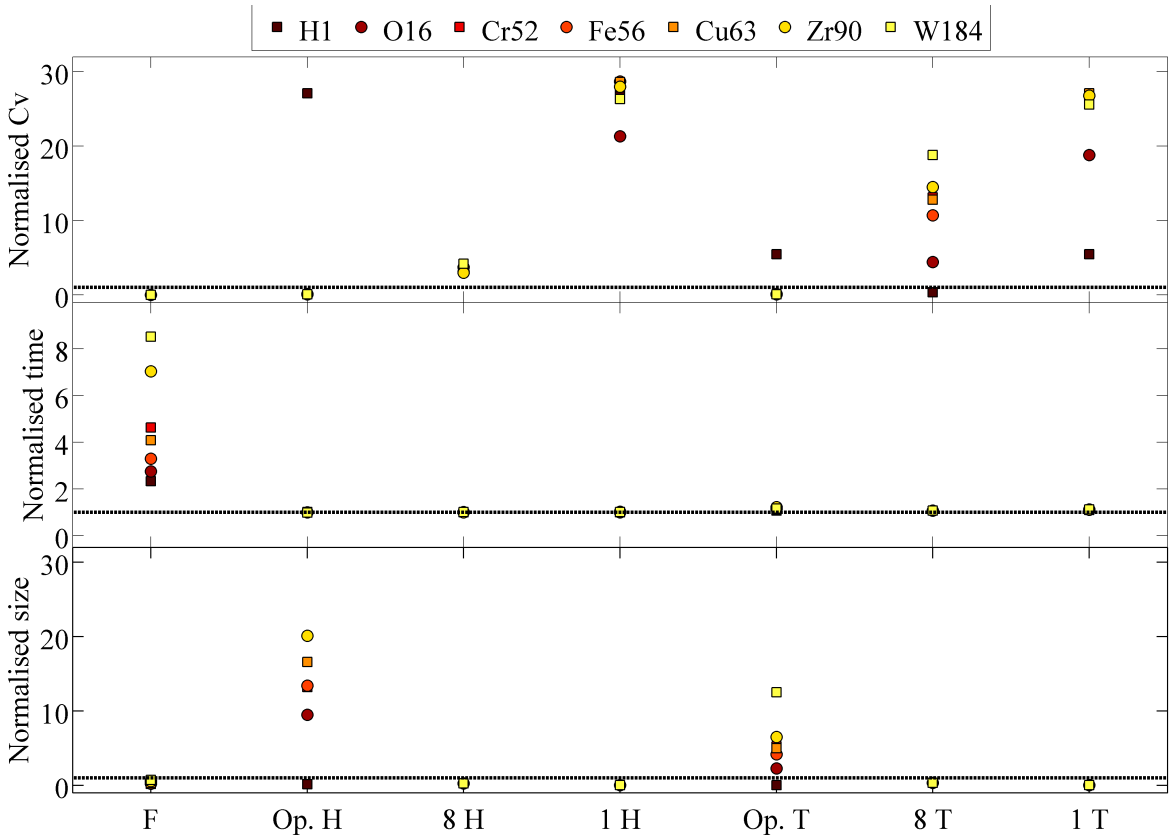

Figure 4.32: Potential gains and losses in terms of accuracy (top), sampling time (middle) and database size (bottom) for all tested isotopes and formats. The C_v, runtime and database sizes are normalised by the equivalent values for the 32 equal-probability histogram format. Any points below the black dashed line are gains in accuracy/time/size. Any points above the line are losses in accuracy/time/size. Where only one point is visible, the different isotopes have approximately the same value. Values correspond to the sampling of the elastic scattering distributions closest in energy to 14.1 MeV. Figure adapted from [73]

As expected from the ^{56}Fe results, the accuracy of representation is compromised by reducing the number of bins or channels for all isotopes. The 8 bin histogram has the least impact on the C_v, with an average of 3.6 ± 0.2 times the 32 bin histogram values. The worst case is for the one histogram bin, with an average of 26.8 ± 1.0 times the 32 bin histogram C_v.

With respect to sampling time, the functional format shows an increase for all isotopes, with an average of 4.7 ± 0.9 times the 32 equal-probability histogram. Tabulated formats show a slight increase in sampling time. The worst case of optimised tabulated data results

in a 1.2 ± 0.1 times average increase in sampling time. No formats reduce the sampling time, as equal probability histogram is the fastest format considered and is independent of the number of bins.

Some saving can be made in database sizes for the functional and degraded formats. One bin histogram and one channel tabulated result in an average database size 32 times smaller than the current format. The implementation of 8 bin histogram and 8 channel tabulated result in a database 4 times smaller. Functional data require a variable number of entries depending on the distribution complexity, but still results in an average reduction of 2.7 ± 0.1.

Figure 4.33 shows the product of the normalised C_v, time and size as a figure of merit for the various formats, and Table 4.6 describes the meaning of the possible values.

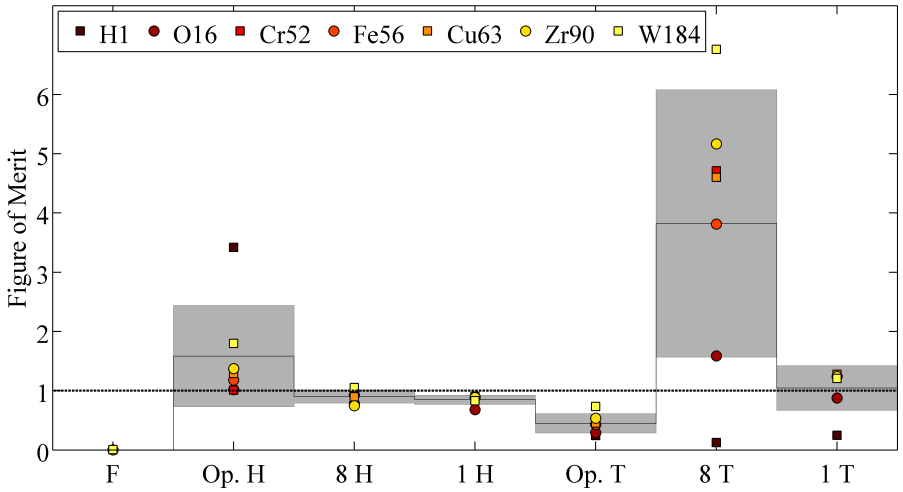

Figure 4.33: The figure of merit is defined as the product of the normalised C_v, time and size as shown in Figure 4.32. Grey bars show the average for each format, within one standard deviation. A figure of merit equal to one is associated with the current 32 equal-probability histogram.

Table 4.6: General meanings of the different values for the figure of merit.

Figure of Merit	**Meaning**
< 1.0	Net gain of accuracy, time and/or precision
1.0	No overall change in accuracy, time and/or precision
> 1.0	Net loss of accuracy, time and/or precision

These values should be used with caution, by combining all metrics within one single number the results could be misleading, particularly with the functional case. Naturally it is an identical representation of itself so by definition the C_v is zero. This results in a figure of merit equal to zero, but it is known that the computational time required to sample from these distributions is far longer than with any of the other formats. These values are dominated by the normalised C_v and should be considered alongside Figure 4.32 as a guide to compare the various formats. Excluding the functional format, the optimised tabulated data results in a net gain over the considered parameters for all isotopes. The increased memory requirements and sampling time are compensated by the gain in accuracy. One and 8 bin histograms show little variation from the current format. The one tabulated channel results in a gain for ^{1}H and ^{16}O, but a loss for the remaining isotopes. Optimised histogram increases the file sizes significantly, so the figure of merit results in no change or a net loss depending on the isotope. Finally, the worst case is the 8 channel tabulated format. For all isotopes except ^{1}H the increased C_v and sampling time are the largest contributors to the figure of merit.

For all previous stages of analysis, the optimised data formats were based on a 1.0% C_v limit. Figure 4.34 shows the computational consequences of optimising the angular distributions to 0.1, 1.0 and 10.0%. Given the previous work on ^{56}Fe, a C_v around 10.0% produces noticeably different energy spectra when sampled, 1.0% results in slight differences and 0.1% is considered to have negligible effects on simulation results. In an ideal world, the pre-processing would be an invisible stage within the nuclear data cycle, in the sense that the processed data would be indiscernible from the original. This is only possible by using the original format, or pre-processing with a sufficiently low tolerance, of 0.1% or below. This comes at a computational cost, so in reality the accuracy of pre-processing must be balanced in accordance with the implementation of the formats. Optimising the ^{1}H data formats reduces the number data points required, which is prop-

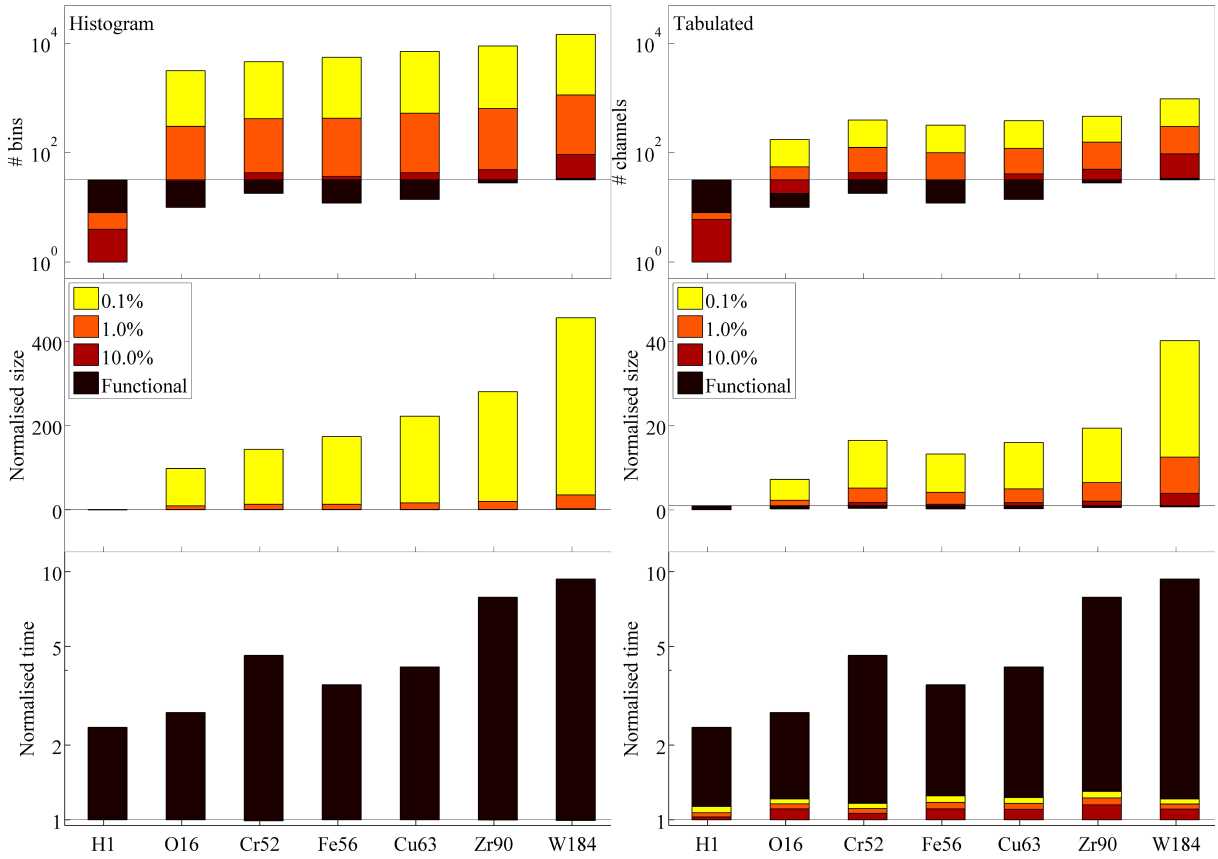

Figure 4.34: Effect of optimising histogram (left) and tabulated (right) angular distributions to within a tolerance of 0.1%, 1.0% and 10.0% with each isotope. The top plots show the number of bins or channels required to achieve the set tolerance (log scale). Baseline is set at 32 to show the variation with respect to the 32 equal-probability histogram format. Functional data (black) are shown on all figures for comparison. The number of coefficients normalised by 32 bins are shown in the top plots for comparison with the number of bins or channels. Middle plots show the database size, normalised by the value for 32 equal-probability bins and the the bottom plots show the normalised sampling time on a log scale. The baseline on both of these plots is set equal to 1. Adapted from [73].

agated through to savings in memory. In all other cases, the optimised data formats increase the memory requirement for the angular distributions. The number of bins/channels required is directly related to how complex the functional form is. The worst case of ^{184}W is described by a 34^{th} order polynomial, to optimise this requires 966, 301 or 95 tabulated channels for a C_v of 0.1%, 1.0% and 10.0% respectively. Histogram format requires 14,621, 1,143 or 93 bins for the same C_v values. Optimisation of the equal-probability histogram requires up to an order of magnitude more bins than channels in optimised

tabulated data. Functional data show a general reduction in file size, though higher mass nuclei with highly featured angular distributions may require a polynomial of order 32 or more.

Sampling time of the optimised histogram format is identical to the 32 bin case. Optimised tabulated data results in a small increase in sampling time that is proportional $\log_2 N$, where N is the number of tabulated channels. Functional data results in the largest sampling time up to 9 times that of the equal probability histogram. This increase is linear with respect to the order of the polynomial. The sampling time requirement for the different formats is consistent with the discussion in Section 3.3.1.

Simple Monte-Carlo with transport

Slab and sphere radiation transport simulations were run with all of the considered isotopes, with both elastic and first level inelastic reaction channels. The resulting C_v from the energy dependent surface currents can be seen in Figure 4.35 for the slab and 4.36 for the sphere.

The same general trend can be seen from both sets of simulations. The C_v from optimised tabulated data is mostly limited by statistics. The C_v for the 32 bin histogram is initially high, particularly for high mass isotopes and is consistent with previous analyses, before converging to an average of around 6% for both geometries. For the two MCNP specific formats, the C_v is larger in the slab geometry than the sphere. This variation is due to the high escape probability of neutrons after a single collision in the slab. The distributions for the initial scatter are echoed within the energy spectra, whereas in the spherical geometry the multiple surface crossings reduce the overall difference.

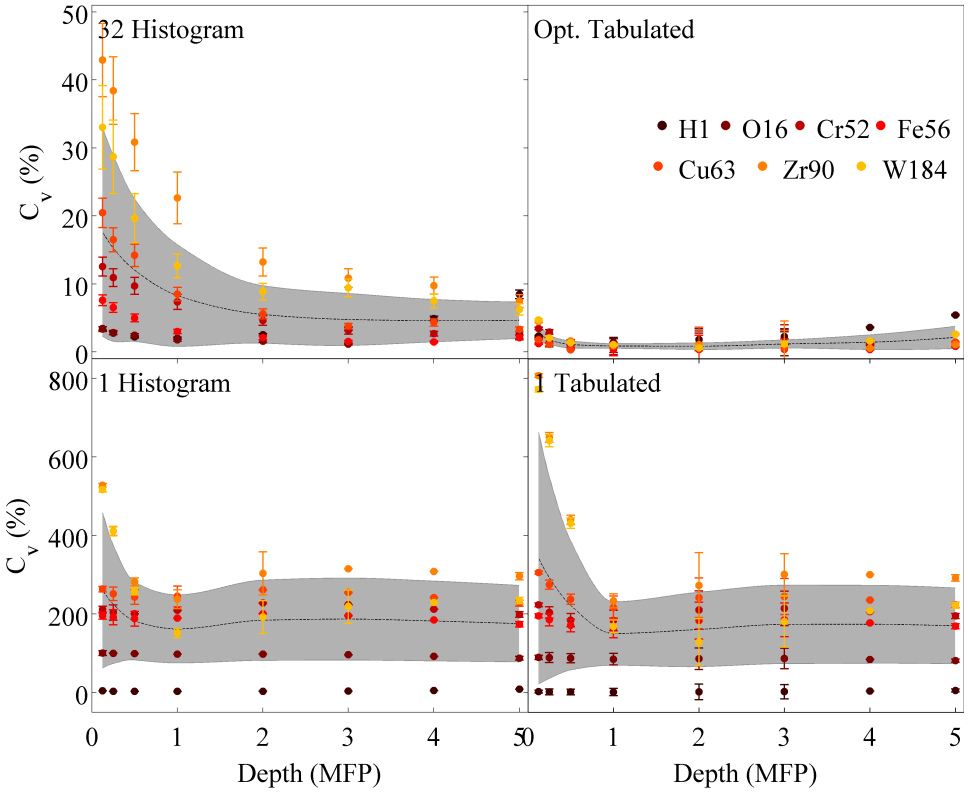

Figure 4.35: Calculated C_v as a function of depth for the fusion relevant isotopes in the slab geometry. The four formats 32 bin histogram, accuracy optimised tabulated, one bin histogram and one channel tabulated are shown. Grey bars show the mean value plus-minus one standard deviation.

The one bin histogram and one channel tabulated formats behave in the opposite sense. The transport is greatly affected by these degraded formats, the forward bias and featured back scatter region are lost. Within the slab, neutrons interact multiple times within the range of 0 to 1 mean free paths and escape, a reduced number of neutrons transport deeper into the slab. Within the sphere, the reduced escape probability results in a much higher density of surface crossings per source. Termination only occurs if the neutron reaches the bounding sphere (radius 5 mean free paths). The initial difference of the distributions is large, so the transport is correspondingly different between the MCNP and degraded data formats. This difference is compounded with multiple events in the case of comparing an isotropic to highly anisotropic system.

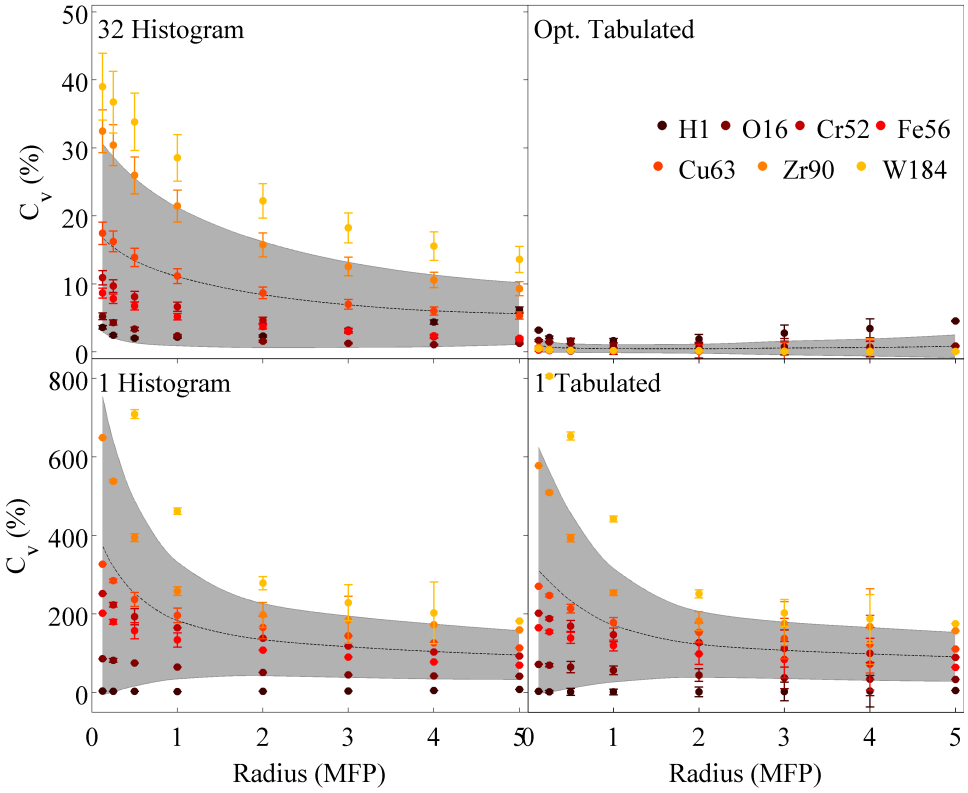

Figure 4.36: C_v as a function of radius for the seven different isotopes of interest. The four formats 32 bin histogram, accuracy optimised tabulated, one bin histogram and one channel tabulated are shown. Grey bars show the mean value plus-minus one standard deviation.

Runtime parameters were extracted from the simulations to determine the effect of the angular distribution format on a less trivial situation than forced scatters without transport. For each event Turnip must check the geometry for the location of the neutron, determine the interaction material and reaction channel, sample the angular distributions, calculate the outgoing neutron properties and transport it to the next interaction site. This continues until it is terminated through escape or energy cut-off. Figures 4.37 and 4.38 show the runtime for the different isotopes and formats, in the slab and sphere geometries. The data are given as the time per source, i.e. the total runtime, including database load times and transport divided by the number of source neutrons. This measures an average time per history.

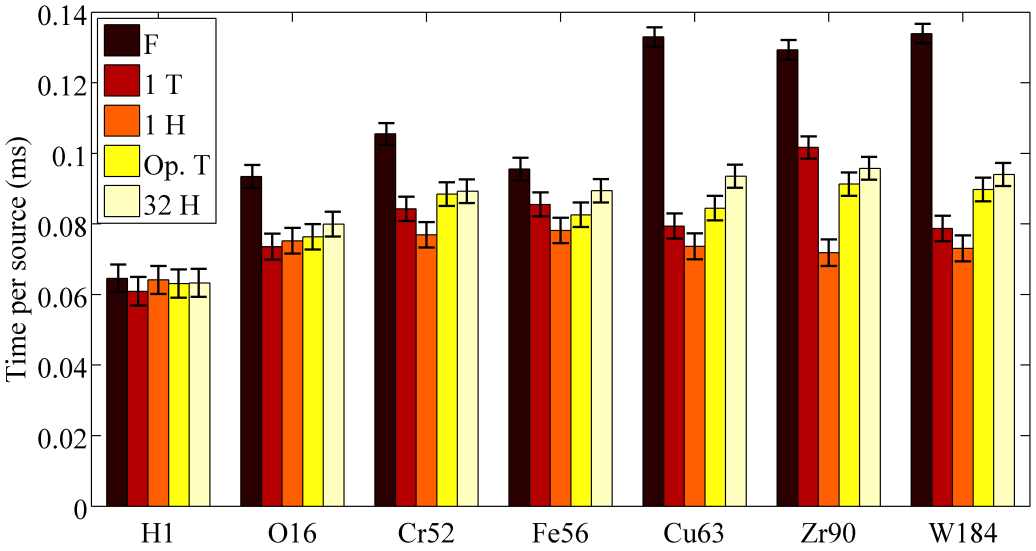

Figure 4.37: Runtime results for the slab geometry. For each isotope and data format, the runtime per source was calculated. Simulations were run with 10^7 source neutrons with initial energy of 14.1 MeV.

When ignoring transport, histogram formats were the fastest to sample from regardless of the number of bins. Within the slab, the forward biased formats (functional and MCNP formats) are associated with the neutrons undergoing very few events before they escape the geometry. The low number of events per history results in the expected runtime behaviour of the functional data, i.e. it is the slowest format to sample, though the effect is reduced when considering transport. Sampling from these distributions takes the longest time per source except for the case of ^{1}H, where all formats result in the same runtime within errors. All other formats are very similar in the average time per source.

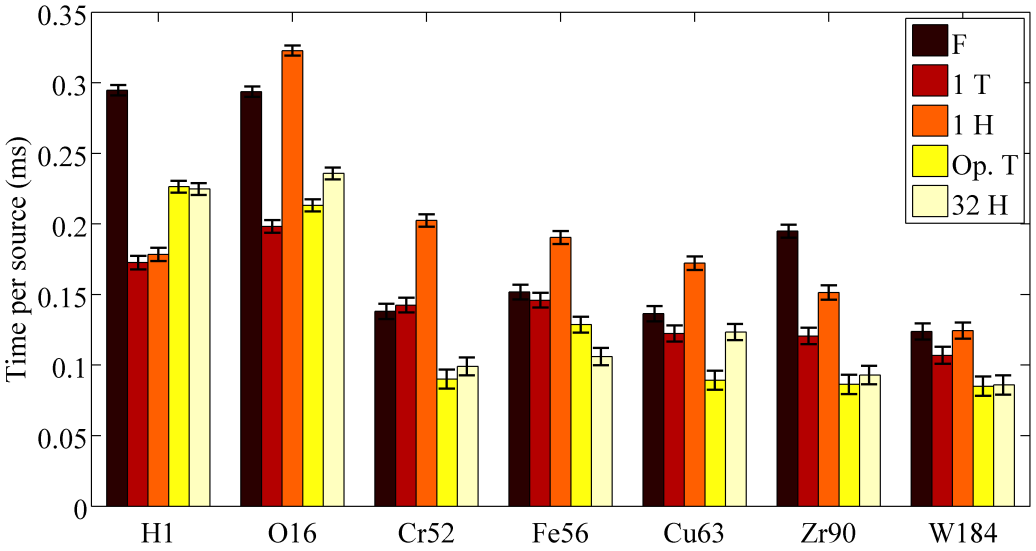

Figure 4.38: Runtime results for the sphere geometry. For each isotope and data format, the runtime per source was calculated. Simulations were run with 10^7 source neutrons with initial energy of 14.1 MeV.

In the spherical geometry the effects of the different formats become more distinct. For 5 of the 7 tested isotopes, the time per source for one histogram bin was equal to or higher than the functional data. When sampled from the one bin histogram data, interacting neutrons will scatter isotropically. When the neutron is enclosed in a large body of material it will scatter frequently within the volume increasing the overall runtime. Sampling from the functional format resulted in the longest average time per source in ^{1}H. Processed distributions with a distinct forward bias (32 histogram, optimised tabulated) resulted in very similar runtimes per source. In many cases, these formats resulted in lower runtimes as the anisotropy increased the probability that neutrons would forward scatter and exit the volume.

4.2.3 Summary

The isotope range considered within this section provides a representative sample of fusion relevant materials and reaction channels in close proximity to the plasma. The number of points required to optimise the processing of angular distributions is closely related to

the neutron energy, mass of the interacting nuclei and the reaction channel considered. Of the examples shown, the simplest case is ^{1}H(n,el) at 14.1 MeV. This low mass is associated with a closely isotropic angular distribution, even at fusion energies. Only one tabulated channel with linear interpolation (or 4 equal-probability histogram bins) are required to achieve a C_v below 1.0%. The current 32 equal-probability histogram format is excessive to achieve the required accuracy. The worst case is associated with the heaviest nuclide considered within this work, ^{184}W(n,el) at 14.1 MeV. The 32 equal-probability bin histogram results in a C_v of 20.0%. To optimise this within a 1.0% tolerance 301 tabulated channels or 1,143 histogram bins are required. The resulting increase upon memory requirements is significant for this case.

Combining the accuracy, runtime and database metrics, the largest overall improvement on the 32 equal-probability histogram is due to the accuracy optimised tabulated data. This balances the improved accuracy, with a small increase in sampling time and database size. If selecting a format based solely on accuracy, the unprocessed functional formats would be best. This format reduces memory requirements, but the cost in computational time is significant when considering forced scatters. The worst case of ^{184}W was 9 times slower than using the equal-probability histogram method.

When implemented within a radiation transport simulation, the geometry and isotope determines the overall impact of each different format. Geometries such as the slab have a high escape probability, and few interactions per source, so the differences in the resulting energy spectra closely replicate the differences between the input formats. The difference in time for each format per source neutron is noticeably different. The worst case shown was for ^{63}Cu, where functional data resulted in a 1.4 times increase in runtime in comparison to the 32 equal-probability bin format.

Within the spherical geometry, the escape probability is low, resulting in multiple events per source prior to escape. For the ^{56}Fe, 32 equal-probability case, an average of 4 collisions

per source occur within the slab but 8 collisions per source occur in the sphere. As more events are occurring within the sphere, a larger proportion of the total runtime is spent on transport and geometry than the sampling of angular distributions. Severely degrading the input data to one histogram bin makes the largest difference to sampling time (and accuracy). For the example of ^{56}Fe, each source undergoes an average of 4 collisions in the slab, but an average of 25 in the sphere. This not only reduces the overall computational efficiency, but reduces the degree to which the simulation is considered physical.

4.3 Real-world examples

Sections 4.1 and 4.2 demonstrated a small but noticeable effect on the energy spectra for varying levels of simulation complexity. Initial analyses considered the pure sampling of the angular distributions without transport. The next stage considered the effect of the sampling with transport in simple spherical and slab geometries. For these cases the effect of using different angular formats was most clearly demonstrated in regions were few scatters take place. Second order events, such as multiple scatters, have a tendency to reduce the sensitivity of the simulation to data format. In real-world fusion simulations, the models are far more complex than the control cases considered previously. They consist of many cells, materials and energies so the phase space of the problem is significantly larger. The port-plug benchmark is the next level of complexity, the geometry is relatively basic and only two materials are used. Table 4.7 describes some of the model parameters for each stage. The DEMO model described in Section 3.3.3 is one of the most complex geometries currently used for fusion neutronics analyses, yet this only covers a 22.5° segment of the entire device. This section presents the results from running these models with the two currently available MCNP formats; the legacy 32 equal-probability histogram and the optimised, cumulative tabulated data.

Table 4.7: MCNP model parameters for the simple and real-world models. Dimensions are given in units of meters or mean free path depending on the model. Materials correspond to the various combination of isotopes, for example steel or water. Isotopes are the individual components (and typically require separate data files) used within the materials.

	Slab	Sphere	Plug	DEMO
Cells	8	8	38	694
Materials	1	1	2	410
Isotopes	1	1	56	121
Dimensions	2mfp(h)	5mfp(h)	1m(h)	23m(h)
	2mfp(w)	5mfp(w)	1m(w)	7m(w)
	5mfp(d)	5mfp(d)	5m(d)	19m(d)
Source	Mono-energetic, mono-directional	Mono-energetic, isotropic	Mono-energetic, distributed	External parametric plasma source

4.3.1 Port-plug benchmark

This model was run with 10^9 source neutrons, and all possible variance reduction techniques were removed from the model. Surface currents were measured throughout the first half of the model, in both the steel and combined steel and water cells. The surface currents were measured on a coarse energy grid from 1.0 MeV to 20.0 MeV in 1.0 MeV intervals. Figure 4.39 shows the energy dependent surface current in the positive direction for the two data formats throughout the model. The data shown are for the inner volume of steel and water up to 105 cm deep into the model. The overall behaviour of the surface currents is similar to the steel sleeve. The statistics of the surface currents worsen with depth in the slab. On the source facing surface (depth 0 cm), there is a difference between the two energy distributions that is larger than the statistical errors.

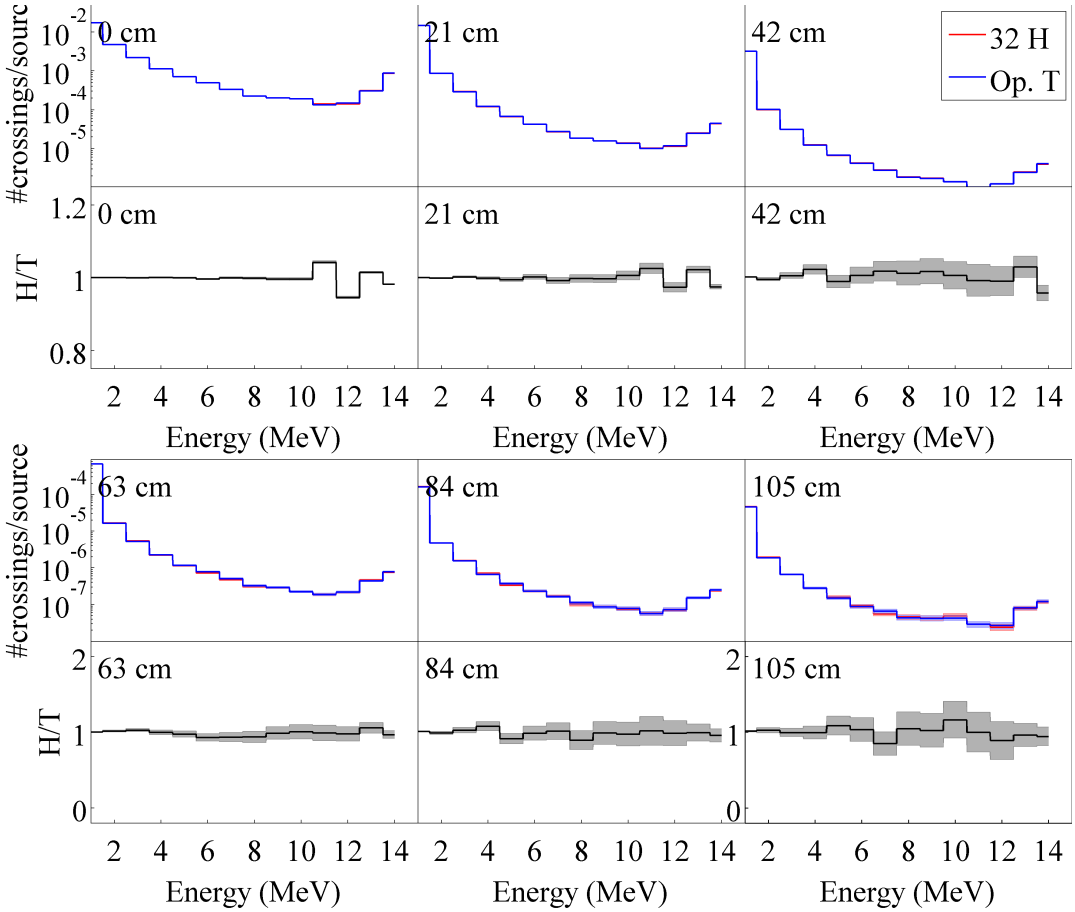

Figure 4.39: Energy dependent surface currents within port-plug model for the 32-equal probability histogram (red) and optimised tabulated formats (blue). Plots are grouped by the depth within the model for the steel and water volume. The top figure of each group shows the number of surface crossings per source within each energy bin, the bottom plots show the histogram data (H) as compared to the tabulated format (T). Shading corresponds to the statistical uncertainty associated with the solid line.

The histogram underestimates the surface current in the region between 11.0 MeV and 12.0 MeV by 5.5%. This is mostly compensated by an overestimate in the 10.0 MeV to 11.0 MeV region of 4.0%. The format of the angular distributions produces a small shift in energy, but given the finite number of neutrons the overall variation is zero. As the depth increases, the statistical uncertainty increases as more neutrons are escaping the volume. In most cases, the statistical error is larger than the differences between the formats and so any deviations are masked within the errors. In some cases, the energy bins are populated through a single event with a 100% statistical error.

Figure 4.40 shows the calculated C_v values from the energy dependent surface currents at each surface. In this case, it is not possible to compare the results from the processed formats with the unprocessed equivalent. From the previous analyses, the optimised tabulated format is a close alternative to the unprocessed data, so these results are considered to be the ideal case. The C_v is calculated as a comparison between the 32 equal-probability histogram with respect to the optimised tabulated data. The C_v shows a general increase with depth that was not seen in the previous studies.

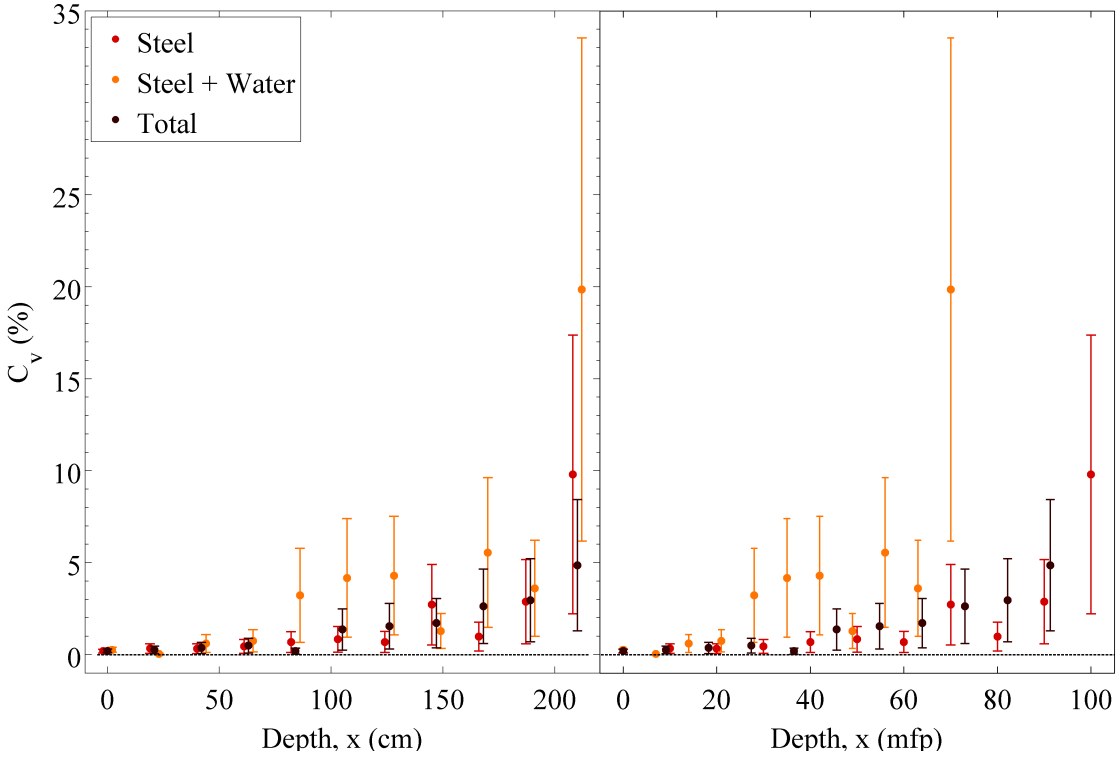

Figure 4.40: Calculated C_v between the energy dependent surface currents from the 32 equal-probability histogram and optimised tabulated formats. The C_v is calculated as a function of depth for the steel sleeve, steel and water inner volume and the total across both cells. The left plot shows the C_v with respect to the depth in cm, but the plot on the right shows the C_v with respect to depth in average mean free paths.

The average mean free path of neutrons, as calculated within MCNP, is 2.1 ± 0.1 cm within the steel sleeve and 3.0 ± 0.1 cm in the steel-water mixture. The results in Sections 4.1.2 and 4.1.2 demonstrated that thicknesses greater than one mean free path show a

reduced sensitivity to the input data format. The first 210 cm in this model equates to 100 mean free paths in steel, 70 mean free paths in the steel and water. Calculating a weighted value for the combined volumes, the average mean free path of neutrons within the geometry is 91 mean free paths. This model is up to 20 times larger than the slab geometries previously considered. The first two surfaces behave as expected, with a low C_v around 0.2%. The third surface and beyond shows an overall increase in C_v and associated errors. At 210 cm, the calculated values for the steel, steel and water, and total are $9.8 \pm 7.6\%$, $19.9 \pm 13.6\%$ and $4.9 \pm 3.6\%$ respectively. These are greater than zero, even when the large statistical errors are considered.

The steel and water C_v values are consistently higher than the others, and when considered as part of the total the C_v is reduced. This is largely due to the geometry, any neutron leaving the inner volume will have a high probability of entering the outer steel sleeve. The curved surface is entirely enclosed by the steel sleeve, so it will also be visited by neutrons leaving the inner surface of the sleeve. Any neutrons exiting the outer surface of the steel sleeve will be terminated, so the effect is less pronounced in this volume. The cross-material transport has altered the energy spectra of the tallies and it is not possible to quantify the contributions from each volume.

The original model, prior to the additions of surface current tallies, included a mesh based tally covering the entire volume. Within each voxel, the track length of the neutrons within the element is tallied, i.e. the contribution to the tally is proportional to how far the neutron travels within the mesh component. Figure 4.41 shows the neutron flux profiles based on the integral flux in each voxel alongside the relative error for the MCNP histogram and tabulated formats.

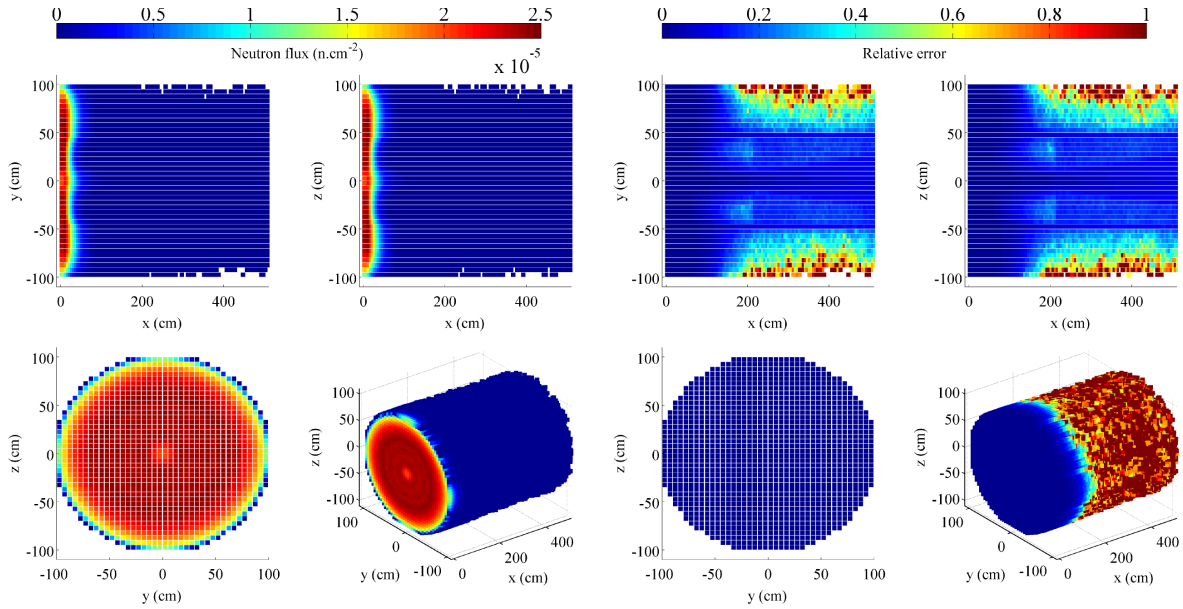

Figure 4.41: Integral tally results mapped over the port-plug model with the MCNP tabulated data. The left figure shows the integral neutron flux throughout the volume. The right figure shows the relative error for the same geometry. The 2-D plots show the values associated with the plane through the origin (x=0, y=0 or z=0).

The neutron flux is highest at the surface closest to the source, and reduces to half by 22 cm. By 500 cm, very few neutrons are transported to this depth. Past a depth of 100 cm, the relative error increases. The area worst affected is close to the outer surface of the steel sleeve. Any neutrons transported to that region have a high probability of escape, hence reducing the available statistics.

For each element in depth, the C_v was calculated based on the difference between the tabulated and histogram results for the energy dependent mesh tally and is shown in Figure 4.42. The C_v increases with depth up to 210 cm, where the volume of steel and water ends. After 210 cm each neutron is scored in void or steel only and so tends to a constant value. The statistics are poor past 210 cm, hence the large error on the C_v, but the lower limit varies between zero and 5.0%.

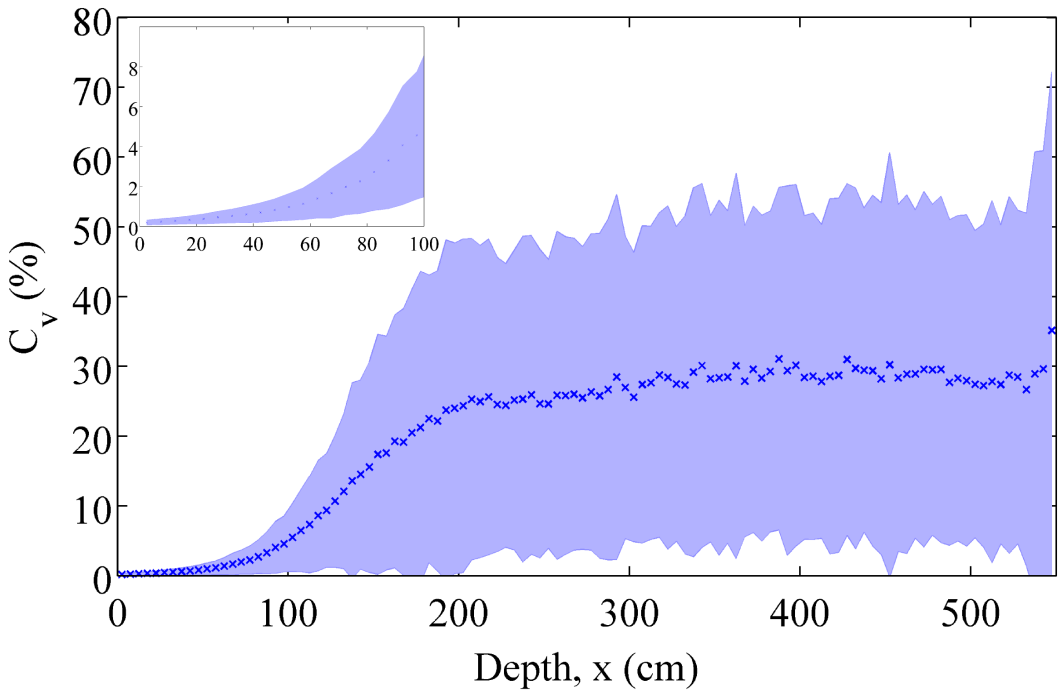

Figure 4.42: Calculated C_v for the comparison of the energy dependent histogram and tabulated mesh tallies in the port-plug model. The change in behaviour at 210 cm coincides with the end of the steel and water inner volume. Inset shows the first 100 cm in the slab, within errors the C_v is always non-zero.

The accuracy of tally results is important, but in a commercial context the efficiency of the simulation is also a contributing factor. This simulation was parallelised, splitting the source neutrons across seven separate threads to reduce the real-time of the simulation. Table 4.8 shows the runtime parameters for the port-plug simulation.

Table 4.8: Computational metrics of the port-plug model for the two MCNP data formats. The library size and the number of lines of data in the library, the number of collisions and total runtime are given. Final column shows compares the histogram and tabulated formats.

	32 histogram	Opt. tabulated	H/T
Size (MB)	1.342E+02	1.817E+02	0.739
Lines of data	1.184E+07	1.463E+07	0.809
Collisions	7.880E+09	7.874E+09	1.001
Runtime (min)	1.458E+04	1.487E+04	0.980

Using the optimised tabulated data results in a loss of efficiency in the computational sense. The library size as stored in memory is 26% larger than with the histogram format and the runtime is 2% longer. For this model, the tabulated format adds 5 hours on to the total runtime, but with parallelisation on seven cores the real-time increase is 43 minutes.

4.3.2 DEMO

This model was run with 10^8 histories, with each source neutron sampled from the standard fusion parametric plasma source with the parameters shown in Table 3.5 (Section 3.3.3). This number of source neutrons is an order of magnitude higher than the typical number used for this model to compensate the removal of variance reduction. The scale and complexity of the geometry and source definition results in high computational runtimes severely limiting the number of source particles allowed and hence limiting the statistics.

Section 3.3.3 described the model and highlighted the three modules used for in-depth analysis of the neutron behaviour. The neutron current was tallied for each surface within modules 3, 9 and 13 in 1.0 MeV bins between 1.0 MeV and 20.0 MeV. Figure 4.43 shows the energy dependent neutron current for module 3 over each surface.

Each surface is described by the depth within the module, this is not the same as distance from the source. Module 3 receives a high neutron flux from the plasma source due to its location, modules 9 and 13 receive a lower neutron flux reducing the available statistics in these modules. As with the port-plug benchmark, the area of largest difference is in the 10.0 MeV to 12.0 MeV region for the surface in closest proximity to the plasma. An overestimate of the tabulated format by the histogram is 4% at the worst point, but compensated for an underestimate on either side. At the next surface the difference has reduced to 2%.

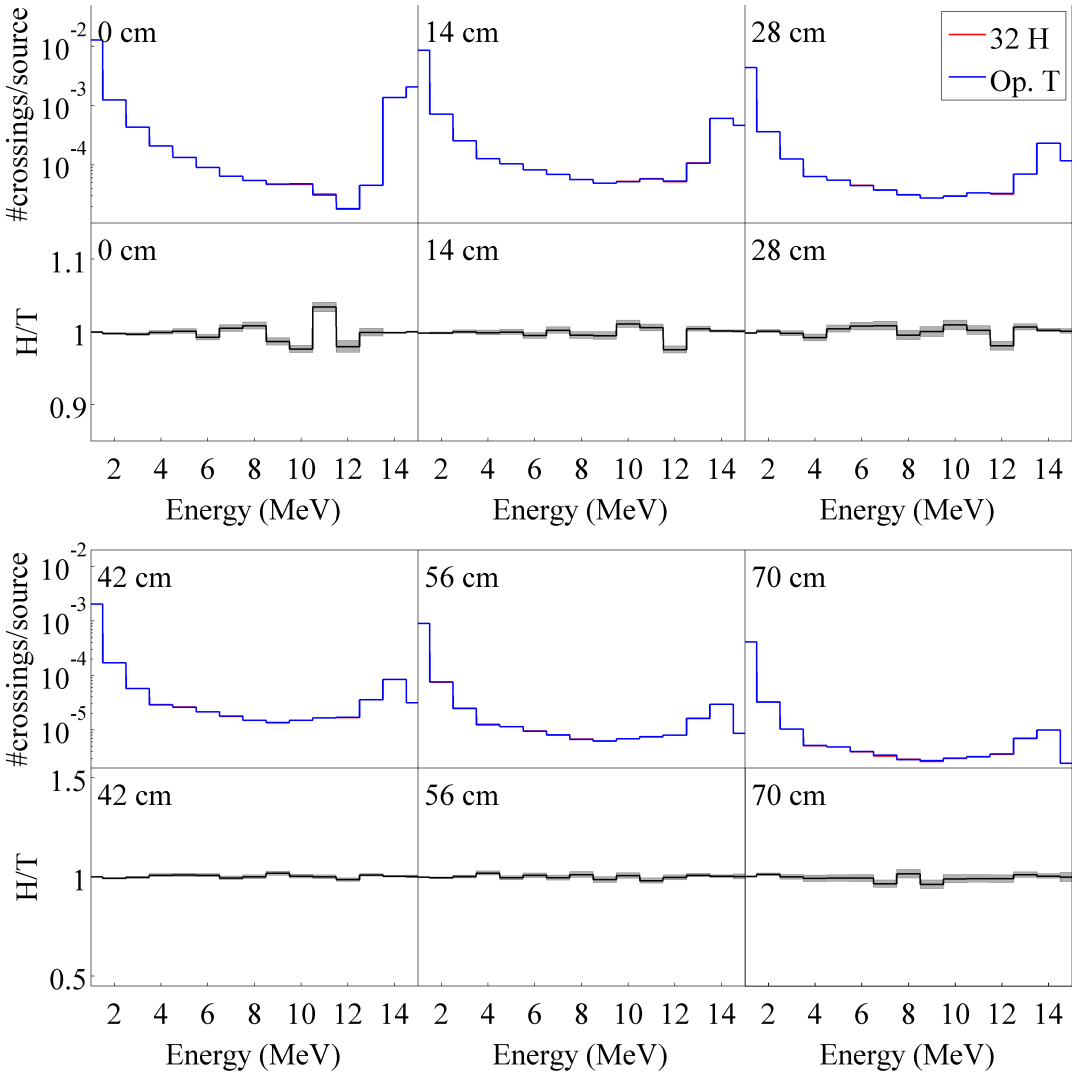

Figure 4.43: Energy dependent surface currents are given for breeder module 3 (see Figure 3.20) for each segment. Each set of plots corresponds to one of the surfaces within the module, the depth in the module is given in the top left corner. For each surface, the neutron current is given for the histogram (red) and tabulated (blue) formats. The bottom plots compare the two.

For each surface within these three modules, the C_v was calculated based on the comparison of energy dependent surface currents for the histogram and tabulated formats, and is shown in Figure 4.44. As with the port-plug model, there is an overall increase in C_v with distance from the source. The added scale and complexity of the DEMO model means the statistics are reduced and hence have a bigger impact on the tally result.

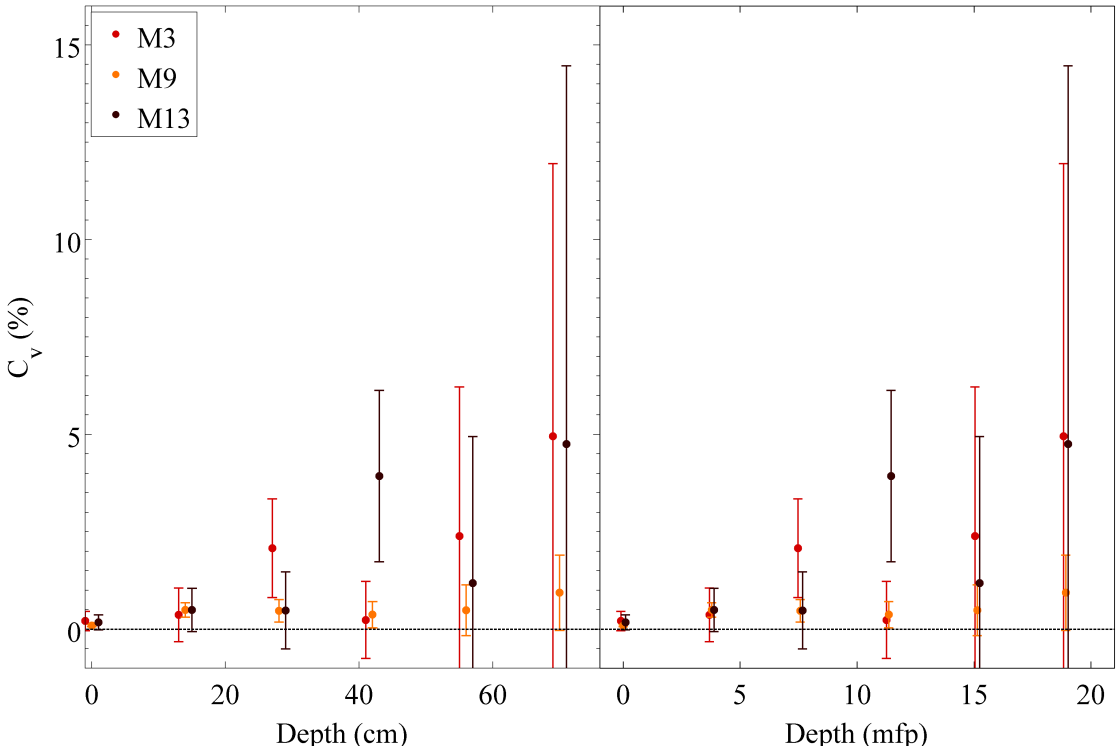

Figure 4.44: Calculated C_v comparing the energy dependent surface currents from the 32 equal-probability histogram and optimised tabulated formats in the DEMO model. The C_v is calculated as a function of depth for breeder modules 3, 9 and 13. The left plot shows the C_v with respect to the depth in cm, and the plot on the right shows the C_v with respect to depth in average mean free paths.

In all but a few cases, the two MCNP formats have a C_v of zero within errors. The case of highest C_v and smallest error is within module 13 at a depth of 42 cm. The value associated with this point is $3.9 \pm 2.2\%$. The statistics are the limiting factor on this analysis, and the spacing of the tally surfaces is equivalent to 3.8 times the average mean free path. This means that many events will occur within each module and the surface current on either side will be subject to contributions from second order effects. To reduce the spacing of these surfaces to one mean free path would be computationally expensive and as such has not been altered.

For each data format and module, the reaction rates were calculated for the tritium breeding reaction in lithium. These integral values were insensitive to the pre-processing of nuclear data within statistical limits. Figure 4.45 shows the tritium production rate per source neutron within the breeder modules.

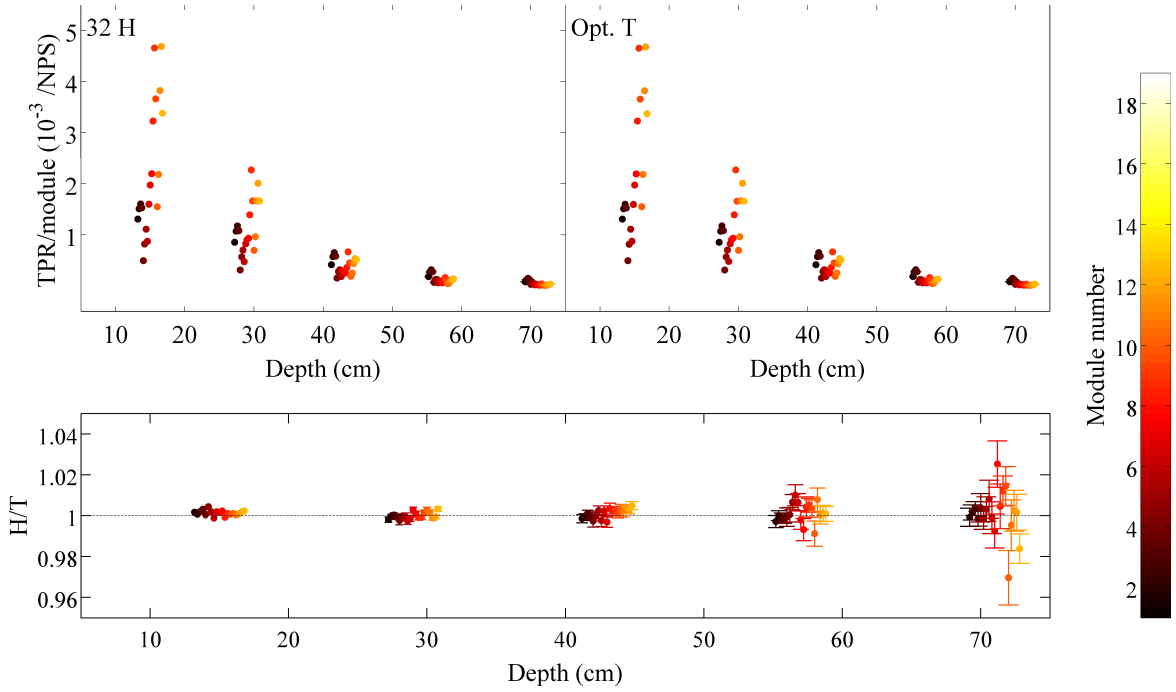

Figure 4.45: Tritium production rate (TPR) as calculated from the MCNP reaction rate tallies for each breeder module and segment. The top two plots show the TPR for the 32 bin histogram and optimised tabulated formats. The bottom plot shows the comparison of the two. Points are coloured by the module number.

For each segment in the module, the tritium production rate was compared for the histogram and tabulated formats. The segments close to the plasma show no variation between the formats, but there is an increase in variation with depth. This is largely dominated by statistical errors, though some modules show an over/underestimate of 2%. These values are summed to give the total tritium production rate per source neutron, and is proportional to the tritium breeding ratio if the time dependent neutron flux and tritium consumption are considered to be constant for both simulations. The differences in the total production rate, and hence tritium breeding ratio is $0.1 \pm 0.2\%$. Within errors

the difference is zero, so no clear differences are apparent at the current level of statistical precision.

Within this simulation, track length estimator tallies were set up to determine the neutron flux passing through all of the breeder modules. The range of this tally covers the entire energy space of the problem, from 10^{-10} MeV up to 19.5 MeV. Based on these energy dependent flux tallies, the C_v was calculated in each module for each segment. Again these C_v values were calculated comparing the 32 equal-probability histogram format to the standard of optimised tabulated data and can be seen in Figure 4.46. There is an obvious positive trend in C_v with distance from the plasma, though the difference is small. At 70 cm, the C_v is calculated to be $4.9 \pm 3.0\%$ in the worst case. The best case at 14 cm has a C_v of $0.6 \pm 0.1\%$. In all modules and surfaces, the difference between the two MCNP format is small, but not insignificant.

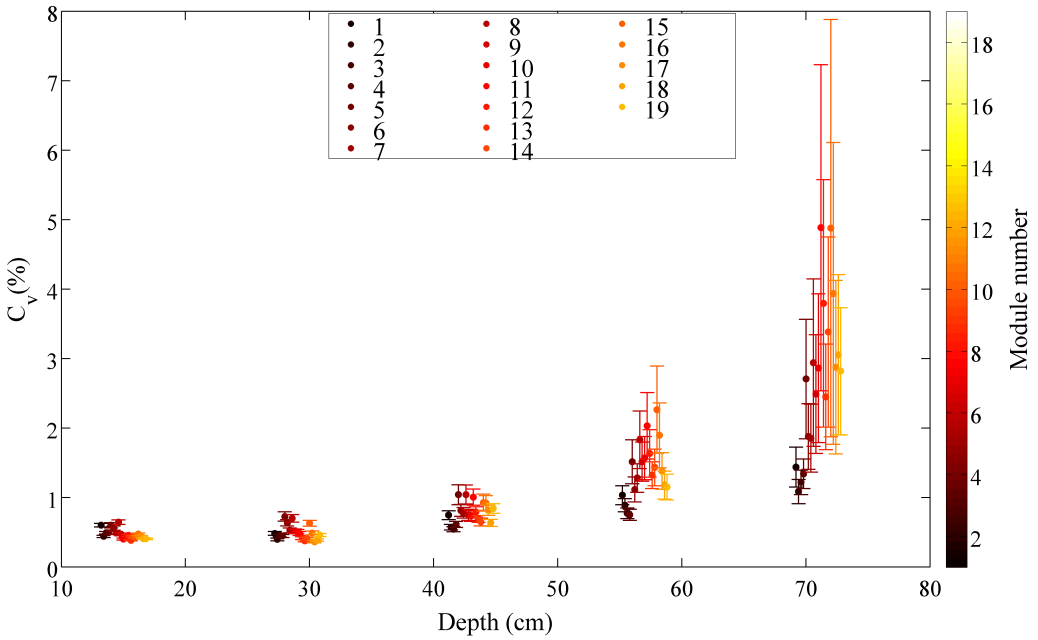

Figure 4.46: Calculated C_v with depth in breeder modules comparing the 32 equal-probability histogram to the optimised tabulated format for the energy dependent flux tallies. For each module, the values are given for each segment.

While the overall contribution from the different formats is small, the effect on the computational parameters are larger. Table 4.9 compares the memory and time requirements for the DEMO model. Optimised tabulated data increases the memory requirement by 10% compared to the histogram format. The total runtime is increased by 5%, this results in a total runtime increase of 19 hours. When run in parallel on seven cores, the real-time difference is closer to 3 hours.

Table 4.9: Computational parameters associated with the DEMO simulation. The library sizes are given in terms of memory requirement and number of lines in the data files alongside the total computational runtime. The final column compares the histogram (H) and tabulated (T) parameters.

	32 histogram	Op. tabulated	H/T
Size (MB)	9.024E+02	1.008E+03	0.895
Lines of data	2.491E+07	2.668E+07	0.933
Collisions	1.036E+10	1.036E+10	1.000
Runtime (min)	2.173E+04	2.286E+04	0.951

4.3.3 Summary

Investigating the effect of the pre-processing of angular data on real-world models is non-trivial given the limitations of MCNP and its accepted formats. There is a general correlation that the 32 equal-probability histogram and optimised tabulated formats result in a small, but quantifiable differences in energy spectra. The scale of the models involved severely limits the precision with which the consequences can be studied. The models were far larger in size than the control cases, for example the steel-water component in the port-plug model was 70 average mean free paths whereas the slab geometry had a maximum depth of 5 mean free paths. There were additional interactions between volumes of different materials, constructed of multiple isotopes and all available reaction paths contributed to the tallies. These may be the factors which contribute to the differences seen deep within the simulation results, but without improving the statistics it is not possible to determine the true source of these differences. As expected from the control

cases of spheres and slabs, the difference on integral values, such as the reaction rates for tritium breeding were insensitive to the format of the angular distributions.

For the port-plug benchmark and DEMO models, there was a computational cost to using the more accurate, optimised tabulated format. The memory requirements increased by 4% and 10% for the port-plug and DEMO model. This memory will be unavailable for use within the transport, and may be a contributor to the slower runtimes. The percentage difference in runtime was small (2% port-plug or 5% DEMO), but for models of this size it can result in a 5 or 19 hour increase in total computational time. If these models were to be run with additional source neutrons, the difference would increase accordingly.

Chapter 5

Discussion and Conclusions

The work contributing to this thesis has highlighted problem areas for the pre-processing stage within the nuclear data cycle. Prior to this the, effect of discretising the angular distributions had never been quantified for fusion neutronics. Legacy methods are appropriate for legacy applications, so this stage has been overlooked in many error propagation analyses. These methods must be used with caution when applied to fusion. The analyses were performed on the combination of NJOY processed data for use in the radiation transport code MCNP.

On a distribution-by-distribution basis, the legacy, 32 equal-probability histogram representation can cause significant differences between processed and unprocessed data for the fusion relevant reaction channels considered. The largest observed difference was 20% for the ^{184}W(n,el) reaction channel at 14.1 MeV. The equivalent fission reaction channels are better replicated with this format, with a difference of less than 1%. The lower average energy of the fission scenario typically relates to a closely isotropic distribution. However, at fusion energies the distributions tend to be extremely forward biased, with a low-probability but highly featured back scatter region. The 32 equal-probability bin format is a crude representation for these angular distributions.

The recently introduced alternative format for MCNP calculations is to use tabulated

cumulative distributions. These provide a far better representation of the unprocessed data, in comparison to the 32 equal-probability histogram format, as demonstrated within this work.

When propagated through multiple events, via Monte-Carlo sampling, the observed variations in the lab frame angular distributions tend to a constant level, i.e. isotropic in μ. The number of scatters required to reach this point is dependent upon how poor the processed distribution is in comparison to the original, as quantified with the C_v. A single scatter event retains the features associated with the tabulated and histogram format which far exceeds the statistical errors of the simulation. For the reaction channels considered the lab frame angular distributions were largely insensitive to input format after 4 consecutive scatters and tend to the isotropic case for both formats. This suggests that the format of the differential cross-sections contributes little to the simulation error, in this sense, for large problems with dimensions greater than 4 mean free paths. Caution should be used when using formatted data for problems with dimensions smaller than 4 mean free paths, provided the simulation is run with sufficient source neutrons to ensure convergence. In these cases, the simulation error should include the additional contribution from the pre-processing of nuclear data if the legacy method is used. The tools and methods developed as part of this work provide a rigorous basis for quantifying and propagating these errors.

The lab frame angular distributions quickly converge to an isotropic system, but the associated energy distributions continue to demonstrate differences after multiple events. These distributions retain some the differences associated with the first scatter that are not obscured by multiple events. By 10 forced elastic scatters in ^{56}Fe, the 32 equal-probability histogram resulted in an overall difference of 7% compared to sampling from the original function. For large numbers of forced events (> 40 scatters), the resulting energies tend to a Gaussian distribution that is replicated with both input formats. This

suggests that differences will be observed in the energy spectra if the geometry is less than 40 mean free paths in any one dimension. Furthermore this suggests that for extremely large geometries, or transport after many scatters, the angular distributions could be simplified to a uniform distribution that is memory efficient and fast to sample.

Differences could only be observed in these Monte-Carlo simulations if the problem had been run to convergence, with an associated statistical error below 1.0%. In this case the entire phase space of the input distributions are sampled. If the problem is not run with a sufficient number of samples, the differences are obscured by the large statistical uncertainties.

The development of a dedicated set of pre-processing and Monte-Carlo codes for nuclear data sampling analyses allowed the investigation into alternative data formats. These included the unprocessed, functional form, and variations upon the current histogram and tabulated formats. Functional data did not produce results that were significantly different in accuracy to the optimised, cumulative tabulated format. The differences were more apparent in terms of the computational parameters of the simulation, with functional sampling impractically slow to implement for high order Legendre polynomials. Functional format reduces the overall memory requirement compared to both MCNP formats. In comparison to the current histogram case, the sampling time per event increased by approximately one order of magnitude. Severely degraded data formats, such as the case of isotropic scatter and one channel tabulated data, were very poor representations of the original distributions. However, after many forced events the lab-frame angular distributions were identical to the current MCNP formats. The energy spectra also exhibited the same tendency to form a Gaussian distribution after many events.

A figure of merit was constructed to assess the balance between accuracy, sampling time and memory requirements. In comparison to the legacy, 32 equal-probability histogram case the optimised tabulated format results in an overall improvement. The sampling

time is fractionally longer, and the file sizes larger, but the gain in accuracy of data representation dominates the figure of merit.

Basic radiation transport simulations were performed with the various formats, the MCNP specific and alternatives. The transport of neutrons within slab and sphere geometries was less sensitive to the data format than the forced events. The addition of geometry components and multiple reaction channels reduces the probability that any particular distribution will be sampled. This reduction in statistics essentially blurs the differences which would otherwise have been observed. The two MCNP specific formats demonstrated a difference between the energy spectra, though the differences were less apparent over large distances (or equivalently, many scatters). The severely degraded formats altered the neutron transport sufficiently that the tallied quantities and energy spectra were up to an order of magnitude different when compared to the other formats. The savings in sampling time and memory requirement are not enough to justify the use of these formats in fusion simulations.

The model complexity was increased again to fusion specific models that are used regularly within the community. The phase space covered by these models in terms of the size, number of geometry elements, number of materials and energy is huge in comparison to the control cases. Due to this significant increase in complexity, the variations seen between the cumulative tabulated data and the legacy 32 equal-probability histogram are far smaller than seen in the control cases. The reduced statistics are the limiting factor in quantifying the differences due to the input format. At the current level of statistical precision, some small variations are seen in the energy spectra, but overall these simulations appear to be insensitive to the data format of the angular distributions. Integral values such as total neutron flux and reaction rates are insensitive to the MCNP specific formats of the differential cross-sections. In the considered cases, the observed differences were of the same scale as the statistical uncertainty. To confirm this, an

increased number of histories, and hence runtime, would be required. This suggests that past simulations run with the legacy data format are sufficiently accurate within statistical errors. In current and future simulations it would be recommended to implement the optimised tabulated format as standard, provided the increase in computational time is acceptable.

5.1 Further work

To build upon the foundations of this thesis, there are natural progressions to further evaluate the consequences of nuclear data pre-processing. The first would be to further develop Swede-Mash to parse the remaining reaction channels within the ENDF files, and to then implement these within Turnip. This would include capture and multiplication reactions. This would provide a better basis for comparing the Turnip results with MCNP, and improve the ability to extrapolate the Turnip results to a full radiation transport Monte-Carlo simulation.

The second would be to run the real-world models to a higher level of convergence than the standard case. This will be computationally expensive but will confirm how each of the data formats affect the transport, without statistical uncertainties obscuring the potential differences. In the future computation is likely to become more efficient and less expensive, suggesting that models may be run to the level of precision where these differences will become apparent.

It is uncommon for large models to be run without variance reduction. It would be interesting to investigate whether the the effects become more or less pronounced with increasing levels of variance reduction. This will help to improve the statistics in the regions of interest, and potentially clarify the differences in data format in these regions. Alternatively it may increase the contribution of second order events, reducing the overall ability to identify clear differences due to the pre-processing methods.

The full scale models considered within this work are widely used within the fusion community, but may not be the ideal case to investigate the true effect of pre-processing. The cases where these differences may become more apparent is in the field of nuclear metrology. The simulation of experimental set-ups for the measurement of integral and differential cross-sections have a much higher probability of highlighting the differences. Some suitable benchmarks are available within the SINBAD database, though the quality of these vary and can only be applied in certain cases. To really test the effect of angular distributions on simulation results, an experimental benchmark must be created with this purpose in mind.

Throughout this thesis, the C_v has been used as the metric to measure how well one distribution represents another. To aid in the propagation of pre-processing errors for the wider community, it would be useful to build a database containing the C_v for all isotopes, reaction channels and energies based on the different pre-processing formalisms. This lookup table would provide the typical radiation transport user with an understanding of how physical or unphysical the simulation results may be based on the choice of data pre-processing. By making this information accessible, it will encourage users to become more aware of the consequences of their choices in data library and to ask questions about how it is processed. A natural progression would then be to include dynamic measurement of the sampling statistics of each distribution at runtime. Used in combination with the C_v, this would allow for a quantitative measurement of data quality similar to the statistical tests built in to current Monte-Carlo radiation transport codes.

5.2 Key points and recommendations

- On a distribution-by-distribution basis, the legacy pre-processing formalisms result in differences up to 20% for fusion relevant reaction channels.

- Monte-Carlo sampling of the distributions for a forced number of events showed

that the resulting lab frame angular distributions were insensitive to the input data format after 4 scatters.

- After multiple forced events, the resulting energy distributions retain some of the differences associated with the input format. The energy distributions are sensitive to the data format until approximately 40 scatters, where the distributions tend to a Gaussian shape.

- The integral values, i.e. total neutron flux, reaction rates, are insensitive to the two MCNP specific data formats.

- The differences are only apparent if the sampled distributions are fully populated and problem is run to convergence with statistical errors below 1.0%.

- Testing alternative data formats showed no significant improvements to the overall simulation. Degraded data formats altered the physics sufficiently that the savings in runtime and memory were irrelevant. The exact functional format showed no significant difference when compared to the cumulative tabulated format, other than an unacceptable increase in runtime.

- Increased model complexity tends to suggest a lower sensitivity to the format of the angular distributions. A large phase space results in fewer samples of the same distribution. If the distributions are not fully populated, the differences are not apparent within the simulation output.

- Real-world models such as the port-plug benchmark and DEMO segment are mostly insensitive to the data format when they are run to the current levels of statistical precision.

The field of nuclear data is severely underrepresented within the nuclear community, it forms the basis of all nuclear simulations and is taken for granted by many end users.

The importance of high quality nuclear data cannot be stressed enough as the accuracy of the results are dependent on the accuracy of the data, i.e. garbage in, garbage out. Too many radiation transport users are unaware of where their data libraries come from, or the processing methods applied. Awareness within the community must be increased, so users can make an informed choice of data library and processing method that best suits the end application.

Within the field of fusion neutronics many of the currently used data libraries are not of a high enough standard. The materials are non-standard and little experimental data exist, particularly at fusion energies. This work has highlighted the areas in which the pre-processing stage of the nuclear data cycle can impact upon simulation results. This stage degrades the data as given in the evaluated files to varying degrees; but if the original data are not truly representative of the physics involved then pre-processing only succeeds at worsening the situation. The current MCNP formats are acceptable representations of the angular distributions given in the ENDF files, but the cumulative tabulated distributions provide a closer representation of the unprocessed data at the cost of computational efficiency. If computational efficiency is the priority, and the model is sufficiently large, the equal-probability histogram format is close enough to the original distribution to not significantly alter the transport. There is no need to implement complex alternatives to improve the accuracy of representation, when the unprocessed distributions are lacking in many cases.

The priority in nuclear data for fusion neutronics needs to be the furthering of experimental campaigns to bridge the gaps in current knowledge. This will naturally result in more complete data libraries over time. Only when these libraries are complete, and are an accurate depiction of the physics, will the effects of pre-processing become important.

List of References

[1] L.R. Grisham. Chapter 10 - nuclear fusion. In *Future Energy (Second Edition)*. Elsevier, 2014.

[2] U. Fischer, P. Batistoni, E. Cheng, and R. A. Forrest et al. Nuclear data for fusion energy technologies: Requests, status and development needs. In *International Conference on Nuclear Data for Science and Technology*, 2005.

[3] U. Fischer, P. Batistoni, R. Forrest, and C. Konno et al. Development needs of nuclear data for fusion technology. In *International Conference on Nuclear Data for Science and Technology*, 2007.

[4] D. E. Cullen. How Accurate Are Our Processed ENDF Cross Sections? Technical report, National Nuclear Data Centre, 2014.

[5] R.A. Forrest and J. Kopecky. Statistical analysis of cross sections—a new tool for data validation. *Fusion Engineering and Design*, 82:73–90, 2007.

[6] U. Fischer, P. Batistoni, Y. Ikeda, and M.Z Youssef. Neutronics and nuclear data: achievements in computational simulations and experiments in support of fusion reactor design. *Fusion Engineering and Design*, 51-52:556–680, 2000.

[7] R. A. Forrest. The role of nuclear data for fusion technology studies. *Nuclear Engineering and Design*, 241:4326–4330, 2011.

[8] A. Plompen. Summary report of the technical meeting on long-term needs for nuclear data development. Technical report, IAEA Nuclear Data Section, 2012.

[9] L.A. Artsimovich. Tokamak devices. *Nuclear Fusion*, 12(2):215, 1972.

[10] J. Mlyná. Focus On: JET. Technical Report EFD-R(07)01, European Centre of Fusion Research, May 2007.

[11] D. Meade. 50 years of fusion research. *Nuclear Fusion*, 50, 2010.

[12] V.D. Shafranov, B.D. Bondarenko, and C.A. Goncharov. On the history of the research into controlled thermonuclear fusion. Technical report, Russian Academy of Sciences, 2001.

[13] A. A. Harms, K. F. Schoepf, G. H. Miley, and D. R. Kingdon. *Principles of fusion energy*. World Scientific, 2000.

[14] T. Hamacher and A.M. Bradshaw. Fusion as a future power source: recent achievements and prospects. In *Proceedings of the 18th World Energy Congress*, 2001.

[15] H-S. Bosch, V. Bykov, R. Brakel, and P. van Eeten et al. Experience with the commissioning of the superconducting stellarator Wendelstein 7-X. *Fusion Engineering and Design*, 96–97:22 – 27, 2015. Proceedings of the 28th Symposium On Fusion Technology (SOFT-28).

[16] ITER Council. Final design report of the ITER engineering design activities. Technical Report G A0 FDR 4 01-06-28 R 0.2 , International Atomic Energy Agency, 2001.

[17] K Krane. *Introductory nuclear physics*. J Wiley Sons, 1988.

[18] M. Ni, Y. Wang, B. Yuan, and J. Jiang et al. Tritium supply assessment for ITER and DEMOnstration power plant. *Fusion Engineering and Design*, 88:2422 – 2426, 2013.

[19] M.B. Chadwick, M. Herman, P. Obložinský, and M.E. Dunn et al. ENDF/B-VII.1 Nuclear Data for Science and Technology: Cross Sections, Covariances, Fission Product Yields and Decay Data. *Nuclear Data Sheets*, 112:2887 – 2996, 2011. Special Issue on ENDF/B-VII.1 Library.

[20] L.W. Packer and J-Ch. Sublet. The European activation file: EAF-2010 decay data library. Technical Report CCFE-R (10) 02, EURATOM/CCFE fusion association, 2010.

[21] M. Abdou, E.L. Vold, C. Gung, and M. Youssef et al. Deuterium-tritium fuel self-sufficiency in fusion reactors. *Fusion Technology*, 9:250–285, 1986.

[22] P. Batistoni. On the absolute calibration of neutron measurements in fusion reactors. *Fusion Engineering and Design*, 105:58–69, 2016.

[23] X-5 Monte Carlo Team. *MCNP - Version 5, Vol. I: Overview and Theory*. Los Alamos National Laboratory, 2003.

[24] M.J. Loughlin, P. Batistoni, C. Konno, and U. Fischer et al. ITER nuclear analysis strategy and requirements. *Fusion Science and Technology*, 56:566–572, 2009.

[25] A. Hogenbirk. Extensive neutronic sensitivity - uncertainty analysis of a fusion reactor sheilding blanket. *Fusion Engineering and Design*, 24:275–286, 1994.

[26] G. Chiba, Y. Kawamoto, M. Tsuji, and T. Narabayashi. Estimation of neutronics parameter sensitivity to nuclear data in random sampling-based uncertainty quantification calculations. *Annals of Nuclear Energy*, 75:395–403, 2015.

[27] L. Petrizzi, M. Angelone, P. Batistoni, and U. Fischer et al. Benchmarking of Monte Carlo based shutdown dose rate calculations applied in fusion technology: From the past experience a future proposal for JET 2005 operation. *Fusion Engineering and Design*, 81:1417–1423, 2006.

[28] T. Noda, M. Fujita, H. Araki, and Akira Kohyama. Effect of nuclear data and impurities on the evaluation of induced activity of CVI SiCf/SiC composites. *Fusion Engineering and Design*, 61-62:711–716, 2002.

[29] A. A. Andrianov, Yu. A. Korovin, and I. S. Kuptsov. Interactive information system for preparation and verification of nuclear data in the high-energy range. *Journal of the Korean Physical Society*, 59:1096–1099, 2011.

[30] I. Kodeli. Use of Nuclear Data Sensitivity and Uncertainty Analysis for the Design Preparation of the HCLL Breeder Blanket Mockup Experiment for ITER. In *Proceedings of the International Conference Nuclear Energy for New Europe 2007*, 2008.

[31] K. Kondo, I. Murata, K. Ochiai, and N. Kubota et al. Verification of nuclear data for DT neutron induced charged-particle emission reaction of light nuclei. *Fusion Engineering and Design*, 82:2786–2793, 2007.

[32] R. Villaria, M. Angelonea, P. Batistonia, and U. Fischer et al. Validation of shutdown dose rate Monte Carlo calculations through a benchmark experiment at JET. *Fusion Engineering and Design*, 83:1782–1787, 2008.

[33] A. Davis. Validation report of a subset of available Monte Carlo codes using ENDF/B-VIIR.0 data. Technical report, CCFE, 2012.

[34] N. García-Herranz, O. Cabellos, J. Sanz, and J. Juan et al. Propagation of statistical and nuclear data uncertainties in Monte Carlo burn-up calculations. *Annals of Nuclear Energy*, 35:714–730, 2008.

[35] K.B. Mather and P. Swan. *Nuclear Scattering*. Cambridge University Press, 1958.

[36] D.E. Cullen, R. Muranaka, and J. Schmidt. *Applications in nuclear data and reactor physics*. World Scientific, 1986.

[37] B. Bornschein, C. Day, D. Demange, and T. Pinna. Tritium management and safety issues in ITER and DEMO breeding blankets. *Fusion Engineering and Design*, 88:466 – 471, 2013.

[38] R. Pynn. *Neutron applications in earth, energy and environmental sciences: neutron scattering applications and techniques*. Springer, 2009.

[39] E.E. Lewis and W.F. Miller. *Computational methods of neutron transport*. American Nuclear Society, 1993.

[40] L.L. Carter and E.D. Cashwell. *Particle transport simulation with the Monte-Carlo method*. USERD technical information centre, Oak Ridge, 1975.

[41] W.M. Stacey. *Nuclear reactor physics*. Wiley, 2007.

[42] U. Fischer, S. Simakov, U.V. Möllendorff, and P. Pereslavtsev et al. Validation of activation calculations using the Intermediate Energy Activation File IEAF-2001. *Fusion Engineering and Design*, 69(1–4):485 – 489, 2003.

[43] G.A Cottrell, R. Pampin, and N.P. Taylor. Transmutation and phase stability of tungsten armour in fusion power plants. *Fusion Science and Technology*, 50:89–98, 2006.

[44] M.R. Gilbert, S.L. Dudarev, D. Nguyen-Manh, and S. Zheng et al. Neutron-induced dpa, transmutations, gas production, and helium embrittlement of fusion materials. *Journal of Nuclear Materials*, 442:S755S760, 2013.

[45] Mario Pérez. The engineering design evolution of IFMIF: From CDR to EDA phase. *Fusion Engineering and Design*, 96–97:325 – 328, 2015.

[46] J. Singh. *Modern physics for engineers*. Wiley, 2004.

[47] D.G. Cacuci, editor. *Handbook of Nuclear Engineering: Nuclear Engineering Fundamentals*, volume 1. Springer, 2010.

[48] H.T. Hunter, D.T. Ingersoll, R.W. Roussin, and C.O. Slater. SINBAD: Shielding Integral Benchmark Archive and Database. Technical report, Oak Ridge National Laboratory, 1996.

[49] N. Otuka, E. Dupont, V. Semkova, and B. Pritychenko et al. Towards a More Complete and Accurate Experimental Nuclear Reaction Data Library (EXFOR): International Collaboration Between Nuclear Reaction Data Centres (NRDC) . *Nuclear Data Sheets*, 120:272 – 276, 2014.

[50] M. Herman and A. Trkov. ENDF-6 Formats manual. Technical Report BNL-90365-2009, Rev.2, National Nuclear Data Centre, 2011.

[51] A.J. Koning, S. Hilaire, and M.C. Duijvestijn. TALYS-1.0. In *Proceedings of the International Conference on Nuclear Data for Science and Technology*, 2008.

[52] J. Leppänen. *Serpent - a continuous-energy Monte Carlo reactor physics burnup calculation code*. VTT Technical Research Centre of Finland, 2015.

[53] A. Ferrari, P.R. Sala, A. Fasso, and J. Ranft. *FLUKA: a multi-particle transport code*. Stanford Linear Accelerator Center, 2005.

[54] R. E. MacFarlane, D. W. Muir, R. M Boicourt, and A. C. Kahler. *The NJOY Nuclear Data Processing System, Version 2012*. Los Alamos National Laboratory, 2013.

[55] D.E. Cullen. *PREPRO 2015: 2015 ENDF/B pre-processing codes*. International Atomic Energy Agency, 2015.

[56] E. Brun, F. Damian, C.M. Diop, and E. Dumonteil et al. TRIPOLI-4®, CEA, EDF and AREVA reference Monte Carlo code. *Annals of Nuclear Energy*, 82:151 – 160, 2015.

[57] R.E. MacFarlane and A.C. Kahler. Methods for Processing ENDF/B-VII with NJOY. *Nuclear Data Sheets*, 111:2739 – 2890, 2010. Nuclear Reaction Data.

[58] D. Rochman, W. Zwermann, S.C. van der Marck, and A.J. Koning et al. Efficicent use of monte-carlo: Uncertainty propagation. *Nuclear Science and Engineering*, 177(3):337–349, 2014.

[59] D. Rochman, A.J. Koning, S.C. van der Marck, and A. Hogenbirk et al. Nuclear data uncertainty propagation: Total monte carlo vs. covariances. *Journal of the Korean Physical Society*, 59:1236–1241, 2011.

[60] B.S. Everitt and A. Skrondal. *The Cambridge dictionary of statistics*. Cambridge University Press, 2010.

[61] N. Garis. Numerical solutions for the one-speed neutron transport equation in two-medium slabs and spheres. Technical report, Department of Reactor Physics, Chalmers University of Technology, 1991.

[62] S. Glasstone and G.I. Bell. *Nuclear Reactor Theory*. Van Nostrand Reinhold Company, 1970.

[63] P. K. Romano, N.E. Horelik, B.R. Herman, A.G. Nelson, and et al. 6. tallies — openmc documentation.

[64] J.K. Shultis and R.E. Faw. *An MCNP Primer*. Dept. of Mechanical and Nuclear Engineering, Kansas State University, 2011.

[65] J. Spanier and E.M. Gelbard. *Monte-Carlo principles and neutron transport problems*. Addison-Wesley Publishing Company, 1969.

[66] H.C. Honeck. ENDF/B - Specifications for an Evaluated Nuclear Data File for Reactor Applications. Technical report, Brookhaven National Laboratory, 1996.

[67] M.B. Chadwick, P.G. Young, and C.Y Fu. *ENDF/B-VII.1 MAT 2631*, September 1996.

[68] D. E. Cullen. ENDF Cross Sections are not Uniquely Defined. Technical report, Lawrence Livermore National Laboratory, 2010.

[69] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. *Numerical Recipes in C*. Cambridge University Press, 1992.

[70] S.S. Wilks. *Mathematical statistics*. Wiley, 1962.

[71] J.F. Kenney and E.S. Keeping. *Mathematics of statistics, Pt. 2*. Van Nostrand Reinhold Company, 1951.

[72] A. Davis and A. Turner. Comparison of global variance reduction techniques for Monte Carlo radiation transport simulations of ITER. *Fusion Engineering and Design*, 86:2698 – 2700, 2011. Proceedings of the 26th Symposium of Fusion Technology (SOFT-26).

[73] T. Hutton, J-Ch., L. Morgan, and T.W. Leadbeater. Pre-processing of nuclear data: optimising the relationship between accurate data representation and computational efficiency. (Accepted for publication in Annals of Nuclear Energy), 2016.

[74] T. Hutton. Parsnip (beta). https://sourceforge.net/p/ptracparser/wiki/ParsNIP%20home/, 2014.

[75] M. Loughlin. Conclusions of shutdown dose rate benchmark study. In *6th ITER neutronics meeting*, 2011.

[76] T. Eade, D. Stonell, and A. Turner. MCR2S unstructured mesh capabilities for use in shutdown dose rate analysis. *Fusion Engineering and Design*, 100:321 – 333, 2015.

[77] A.M. Ibrahim, D.E. Peplow, J.L. Peterson, and R.E. Grove. Analysis of shutdown dose rate in fusion energy systems using hybrid monte carlo/deterministic techniques. In *18th Topical Meeting of the Radiation Protection Shielding Division of ANS*, 2014.

[78] J-Ch. Sublet, J.W. Eastwood, and J.G. Morgan. *The FISPACT-II user manual*. CCFE, 2014.

[79] P. Pereslavtsev. Generation of the MCNP model that serves as a common basis for the integration of the different blanket concepts. EFDA D 2M7GA5 V.1.0 (2013).

[80] J. Shimwell, L. Morgan, S. Lilley, T. Eade, and at al. Spatially and temporally varying tritium generation in solid-type breeder blankets. *Fusion Engineering and Design*, 98–99:1868 – 1871, 2015.

[81] M.B. Chadwick, P.G. Young, and C.Y Fu. *ENDF/B-VII.1 MAT 600*, June 1996.

[82] T. Hutton, J-Ch. Sublet, L. Morgan, and T.W. Leadbeater. Nuclear data for fusion: Validation of typical pre-processing methods for radiation transport calculations. *Fusion Engineering and Design*, 100:81–86, 2015.