Towards an interval particle transport monte carlo method

Ander Gray

Institute for Risk and Uncertainty, University of Liverpool, UK. E-mail: ander.gray@liverpool.ac.uk

Andrew Davis

Culham Centre for Fusion Energy, UKAEA, UK. E-mail: andrew.davis@ukaea.uk

Edoardo Patelli

Institute for Risk and Uncertainty, University of Liverpool, UK. E-mail: epatelli@liverpool.ac.uk

Monte carlo has for a long time been the high fidelity model of choice in particle transport. The current state of the art in uncertainty propagation in particle transport is the so called Total Monte Carlo (TMC) method, a method which relies on the repeated execution of the same transport simulation. This however is often computationally intractable even with modern high performance computing standards. In this paper we review the TMC method, with a slight modification, and propose an alternative based on interval analysis; which provides a robust bound on the uncertainty in a single model evaluation.

Keywords: Particle Transport, Nuclear Fusion, Monte Carlo, Uncertainty Propagation, Nuclear Data, Total Monte Carlo, probability bound analysis.

1. Introduction

Particle transport, or radiation transport, is the study of how radiation propagates through matter: how it interacts, what with, how it effects the matter, and its general manipulation. The application area for particle transport methods is very wide. A partial list of current physics and engineering research areas interested in methods for radiation transport is:

- Solar radiation shielding of spacecraft
- Study of the radiative layers of stellar interiors
- Dynamics of charged particles produced from collisions in particle accelerators
- Proton ray therapy in cancer research
- Neutron population dynamics in reactors
- Ray tracing methods in optics

Monte Carlo (MC) has for a long time been the high fidelity method of choice for particle transport[a], due to its simplicity and ability to handle complicated geometries. As apposed to solving a mass balance equation, such as the Boltzmann equation, an agent based modelling approach is taken. Here the paths of individual particles are tracked from their birth to death, with their state at each point in their history being stochastically updated following physical interaction laws. Although the paths of individual particles may not be useful, the global behaviour of a large number of simulated particles is physical.

There are three inputs that are fundamental to a monte carlo particle transport simulation: the particle source, the geometry of the problem in question, and a definition of particle-matter interaction rules. The source defines the initial states of the particles: their birth location, momenta and their type (neutron, electron, proton or other). It is usually defined as a probability distribution, which is sampled at the beginning of a particles history defining that histories initial state. The geometry defines the universe of the problem in question. For example, it could be a model of a human body in radiation therapy applications, or a fuel pin cell model of fission reactor core. Along with the geometric boundaries of the problem, the material within the geometry is also defined: the density of the material and its nuclide composition[b]. Finally the interaction rules of the particles with the material (and its constituent nuclides) needs to be defined. In nuclear applications, this is known as nuclear data. Nuclear data usually provides three things for every nuclide-interaction pair:

[a] It was proposed around the time nuclear technology was first being developed

[b] It is important to make the distinction between different isotopes. Sometimes the lattice structure of the material needs to also be defined, but for most applications the particle energies are large enough for the material properties to be treated as isotropic

Proceedings of the 29th European Safety and Reliability Conference.
Edited by Michael Beer and Enrico Zio

ISBN: 981-973-0000-00-0 :: doi: 10.3850/981-973-0000-00-0_output

(i) **Interaction cross sections**: characterises the probability of a particular reaction
(ii) **Exit energy/angle distributions**: if the reaction has an exit particle, its next energy/angle state is sampled
(iii) **Covariances**: the uncertainty, auto-correlation and cross-correlation information of the interaction cross sections

The distance that the particle travels in each material is also sampled from a distribution derived from the nuclear data. Figure 1 shows a simplified flowchart for constructing particle histories. Since the transport model itself is stochastic, there is an inaccuracy associated with using a limited particle budget. Ideally an uncertainty quantification method should handle and propagate uncertainties from all three of these inputs; and should also distinguish this uncertainty from the inaccuracy of the monte carlo estimator.

Recently there has been a growing interest in methods for quantifying and propagating uncertainty in particle transport MC, particularly for nuclear technologies. Most of these methods are for nuclear data uncertainty propagation, and are mostly either perturbation theory or sensitivity based methods (Aufiero et al. (2016)). These methods have been effective for fission, but are yet to be successful for fusion applications. This is because a key assumption in perturbation methods is that the response of model is linear in the range of the uncertain input. This is an assumption that is too strong for fusion applications due to the severity of the uncertainty in the nuclear data inputs required in fusion[c].

The current most rigorous method for propagation in fusion is the so called Total Monte Carlo method, a second order monte carlo uncertainty propagation method. However this is usually intractable even with modern high performance computing standards. In this paper we propose an alternative to monte carlo for nuclear data uncertainty propagation based on interval analysis, which not only provides rigorous bounds on the uncertainty in a single simulation, but we believe it can also be extended to propagate source and geometric uncertainty. In the following section we discuss Total Monte Carlo in greater detail.

2. Total Monte Carlo: second order monte carlo

The creation of the nuclear data inputs for nuclear applications is a research field by its own right, known as nuclear data evaluation. Here it is the nuclear data evaluators task to statistically mix

[c]There has been a considerably greater academic and industrial inertial in fission research. As a result the nuclear data inputs for fission have been well studied, with relatively low uncertainties

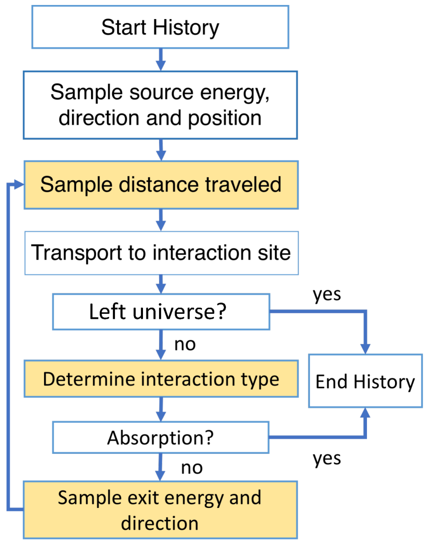

Fig. 1.: A flowchart showing a simplified history of a particle. The shaded boxes show where the sampling is derived from nuclear data

experimental reaction data with reaction models to produce his/her best estimate of the nuclear data quantities plus uncertainty. Due to the difficulty and cost of conducting nuclear reaction experiments, experimental data is sparse or often not present for the vast majority of nuclides and reactions (excluding fission related nuclides, which have been extensively studied).

The Total Monte Carlo (TMC) method was originally proposed by Koning and Rochman (2008) for general nuclear data evaluation and uncertainty propagation. The method relies on a nuclear reaction model code, such as TALYS (Koning and Rochman (2012)), that can produce a complete evaluated data set. They use Bayesian updating, a now popular uncertainty characterisation method, to construct distributions for the input parameters of TALYS with the available experimental data. This input distribution can then be propagated through TALYS to produce a set, or distribution, of evaluated nuclear data quantities. They then suggest that either this set is condensed into a mean evaluation plus a covariance matrix, which can then be propagated using traditional perturbation methods, or that the particle transport simulation is repeated for the random instances of the evaluation: ie. by monte carlo. Only this second method has so far been effective for fusion (Rochman et al. (2010)). Figure 2. shows the TAYLS evaluated Fe56 elastic cross section viewed at various dimensions.

Since the cross section is a particle energy dependant function, an uncertain (in a probabilis-

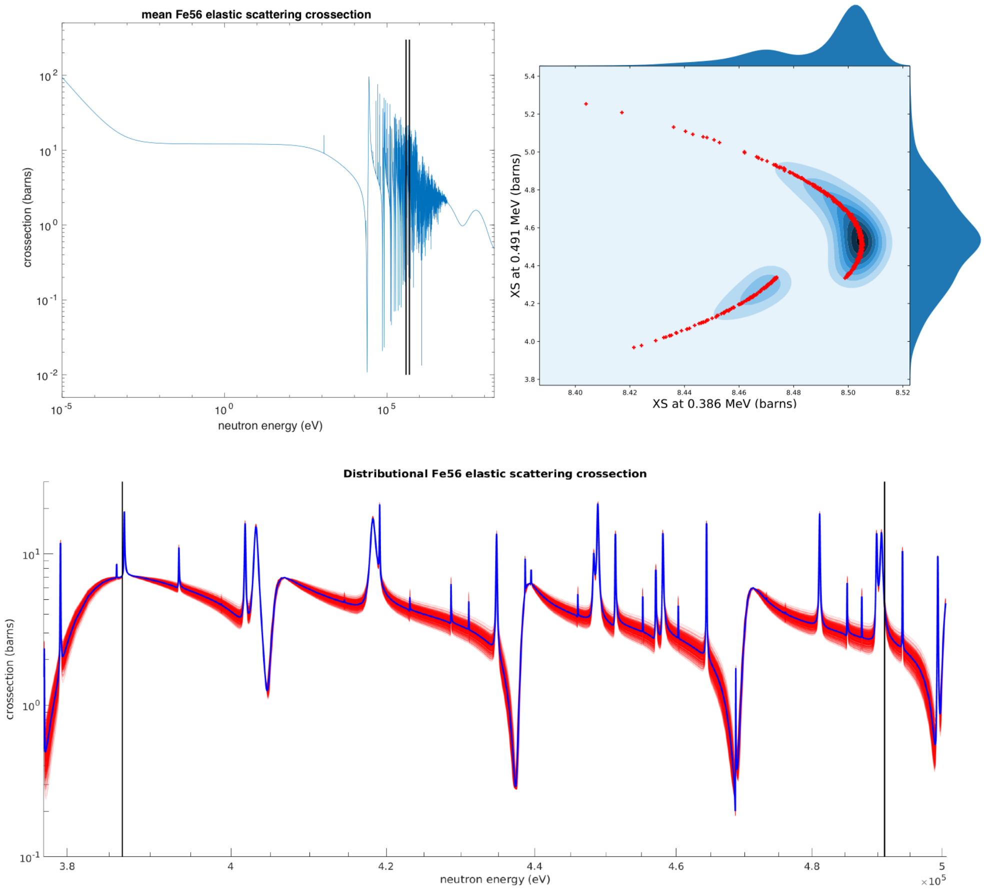

Fig. 2.: Random TALYS evaluated Fe56 elastic scattering cross section viewed at various dimensions. The top left shows the mean cross section, with two marked black lines. The bottom shows an expanded view about these lines, with random instances of the cross section in red. The top right shows a two dimensional slice of the distribution at the marked lines, with the samples in red and a kernel density estimate of the marginals and the multivariate in blue.

tic sense) cross section can be thought of as a stochastic process; where every point in energy is a random variable with a precise distribution and with a precise statistical dependence defined between each random variable. The Bayesian updating in the TMC method provides samples from this stochastic process. The top right image in figure 2 shows a two dimensional slice of this process, with the samples in red and a kernel density estimation of the marginals and the multivariate in blue. What should be noted is that not only are the marginals non-Gaussian, but the statistical dependence is also non-linear. If one condenses all the random instances of the cross section into a mean evaluation and covariance matrix, then this distributional and non-linear dependence is lost.

The great success of TMC is that all of this complicated multivariate statistical information is captured in the set of random evaluations; and that this information can be propagated through any nuclear application in a non-intrusive way by monte carlo. For particle transport applications, this uncertainty propagation scheme can be considered as second order monete carlo: that is, since each execution of the MC particle transport simulation produces a distribution (a mean value and deviation corresponding to the monte carlo error), then repeated execution will produce a density of distributions. Figure 3. is a illustration of this. The overall dispersion of the distributions is due to the uncertain nuclear data input, with the variance of each individual distribution being due to the monte carlo error of that particular simulation, which in turn is dependant on the uncertain input also. In general, this density of distributions is the output to second order monte carlo, and is

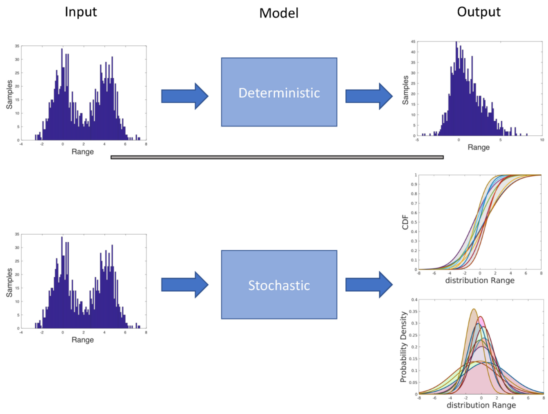

Fig. 3.: An illustration of monte carlo uncertainty propagation in deterministic and stochastic models.

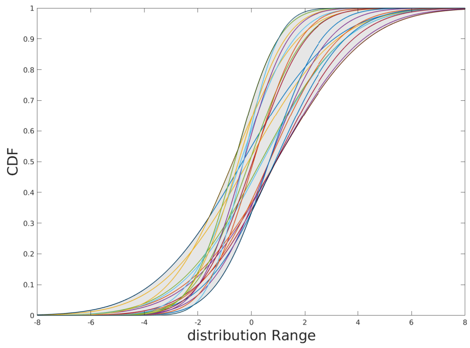

Fig. 4.: P-box representation of a second order distribution

what should be used in further uncertainty and risk assessment. The statistical properties of second order distributions are themselves distributional, i.e. the mean and variance are distributional with a statistical dependency between them. This can make second order distributions difficult to interpret and manipulate, and so often approximations are made. For example, the analyst might only use the distribution of means if they are confident that the monte carlo error is negligible. Rochman et al. (2010) in their Total Monte Carlo method suggest that this second order distribution should be condensed into a singular distribution with the following statistical properties:

$$\begin{cases} \mu_{Total} = \frac{1}{n}\sum_{i=1}^{n} \mu_i \\ \sigma_{Total}^2 \approx \overline{\sigma}_{MC}^2 + \sigma_{ND}^2 \end{cases} \quad (1)$$

where μ_{Total} and σ_{Total} are the mean and standard deviation of this condensed distribution, with μ_i being the mean of each individual simulation. $\overline{\sigma}_{MC}$ is the mean of all the individual monte carlo errors and σ_{ND} being the deviation of the nuclear data, which they give to be:

$$\begin{cases} \sigma_{ND}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(\mu_i - \mu_{Total})^2 \\ \overline{\sigma}_{MC}^2 = \frac{1}{n}\sum_{i=1}^{n} \sigma_{MC,i}^2 \end{cases} \quad (2)$$

There are some arguments that can be made against representing the uncertainty this way. Firstly, any distinction between the monte carlo error and the uncertainty from the nuclear data is lost. Equivalently, this distribution mixes the frequentest and Bayesian interpretations of probability: which represent aleatory and epistemic uncertainty respectively. Theses are generally conflicting and should not be aggregated. This is because the cross section a single but unknown function, and does not have a frequency. The monte carlo error alternatively is rooted in the variability of transport model itself. Secondly, this representation assumes a gaussian distribution. If one wishes to perform further analysis or propagate this uncertainty, we believe that propagating this distribution would lead to an underestimation of the uncertainty.

If one needs to reduce the second order distribution, our preferred representation is an imprecise distribution or a probability box (Ferson et al. (2015)): a distribution with interval moments. This can be created from taking the outter two envelopes of the all the CDFs of the second order distribution. This is shown in figure 4 as the grey region between all of the CDFs in the second order distribution. This gives two bounding distributions representing the two extrema of the probability mass. With this representation the distinction between the nuclear data uncertainty and the monte carlo error is not lost. The nuclear data uncertainty is represented by the difference between these bounding CDFs. With a decreasing nuclear data uncertainty, theses bounds eventually meet creating a precise distribution. The monte carlo error is captured in the interval variance of the p-box. With a decreasing monte carlo error, the bounds steepen eventually creating an interval. The exact density within these bounds is lost in this representation. Instead the interpretation is that any distribution function that lies within these bounds is a candidate for the precise simulation, making no statement on how the probability mass is distributed within. This precise distribution function is retrieved if the inputs are known exactly. We believe that this is a more honest characterisation of the TMC output; which will yield the correct bounds in further analysis and propagation.

For fusion applications, Total Monte Carlo is usually intractable even with modern high performance computing standards. A full fusion reactor simulation would require of the order of $\sim 10^{12}$ particle histories for an acceptable monte carlo error. Assuming that modern cpus can run around 1000 histories per second, a particle transport MC calculation in parallel on 1000 cpu cores would take $\sim$1.5 weeks to complete. Repeating this calculation is completely impractical. We therefore investigate alternatives to monte carlo for uncertainty propagation, which can be performed in a single model query and yields robust bounds on the uncertainty: the above mentioned p-box output. This we base on interval analysis, and discuss in the following section.

3. Interval particle transport

Ferson and Sentz (2016) describe a general framework for extending agent based models to incorporate and handle uncertainty within the simulation. Incorporating uncertainty into the calculation itself would allow for a near automatic propagation of the uncertainty, removing the need for the above described monte carlo. Interestingly they also compare this intrusive propagation method to monte carlo, and they find that monte carlo generally provides an underestimation of the uncertainty in complex systems such as this. They describe a number of features that must be added to an agent based model for it to be generalised to an uncertain model. Following their description, the following must be incorporated into particle transport MC:

(i) Characterise stochastic drivers imprecisely
(ii) Specify particle attributes, tallies, geometries, material properties and particle sources as uncertain numbers
(iii) Execute rules in a way that respects uncertainty in their conditional clauses

The above framework can work for all uncertain numbers (scalars, distributions, intervals, p-boxes, fuzzy numbers or other), including the interactions between uncertain numbers of different classes. We have selected to use intervals as our model of uncertainty, since interval algebra has been extensively studied (Moore (1979)). In this section we describe how a number of these features may be incorporated into the particle transport MC algorithm for propagating nuclear data uncertainty.

3.1. *Particle interactions: imprecise events*

The stochastic drivers in particle transport MC are the distributions which are sampled and change the state, or attributes, of the particle. These are the yellow shaded boxes in figure 1. Since they are directly derived from the nuclear data input, the first step is to convert the stochastic process cross section (figure 2) provided by TMC into something akin to an interval process. This can be quite simply done by taking the maximum and minimum values over all samples at every point in energy. This can be thought of constructing a convex hull or a 2 dimensional interval around all of the samples of the multivariate distribution in the top right of figure 2. Following the interpretation that the interval represents the set of all probability distribution functions defined on the real numbers within its range (Ferson et al. (2007)), then this multivariate, including its non-linear dependence, is contained in this set. Although the exact distribution is lost, it will still be represented and propagated by the interval.

The attributes of the particle are: energy (scalar), direction (3 vector), position (3 vector) and alive state (boolean). A weight, or a probability mass, (scalar) is also often included for advanced monte carlo simulation, such as importance sampling, but it will be omitted from the current description; although we believe that it can also be included in this framework. The above particle characteristics will be intervals, which are changed or updated from sampling of the stochastic drivers. We will only consider two possible events in the particles history: elastic scattering, which changes the particles energy and angle, and absorption, which will kill the particle and end its history.

The probability of an absorption event is given by:

$$P(abs) = \frac{\sigma_{abs}(E)}{\sigma_{abs}(E) + \sigma_{scat}(E)} \tag{3}$$

where σ_{abs} is the absorption cross section dependant on the incident particles energy, and where σ_{scat} is the elastic scattering cross section. Whenever an event is to be sampled, a U(0,1) random number is selected and compared to this value. If it falls bellow, an absorption event occurs and the particles alive state is changed to a zero, ending the particles history. Since $P(abs) = 1 - P(scat)$, above $P(abs)$ a scattering event occurs, where the next energy/angle state of the particle is sampled from a distribution derived from the physical laws of elastic scattering[d]. Since σ_{abs} and σ_{scat} are intervals, we have an interval probability for the event. By rearranging the above relation to:

$$P(abs) = \frac{1}{1 + \frac{\sigma_{scat}}{\sigma_{abs}}} \tag{4}$$

[d] The derivation is out of the scope of this paper, but this function can be extended to include the interval particle attributes as inputs

and by inserting the cross sections as intervals and following the interval algebra, the interval probability for this event can be found to be:

$$[\underline{P}(abs), \overline{P}(abs)] = \left[\frac{1}{1+\frac{\overline{\sigma_{scat}}}{\underline{\sigma_{abs}}}}, \frac{1}{1+\frac{\underline{\sigma_{scat}}}{\overline{\sigma_{abs}}}}\right] \quad (5)$$

where $[\underline{\sigma_{abs}}, \overline{\sigma_{abs}}]$ and $[\underline{\sigma_{scat}}, \overline{\sigma_{scat}}]$ are the bounds on the interaction cross sections. By rearranging the equation like this the double counting problem often encountered in interval methods is avoided.

Similar to the precise case, if the uniform random number falls bellow $\underline{P}(abs)$ or above $\overline{P}(abs)$ then either absorption or scattering occur. If it falls between these bounds, then we are presented with a situation where we cannot say for certain which of these events has occurred. In this situation, both options must be concurrently considered and propagated in the simulation. This is the main technical challenge in this scheme. For this simple case where we are only considering absorption and scattering, if the event is imprecise we can change the particles alive state to $[0, 1]$, and now consider it to be both alive and dead. Since an absorption event ends the particles history, only the scattering event needs to be further propagated, and so the particle continues its history as if scattering has occurred, but with an updated $[0, 1]$ alive state. The effect of this uncertainty in the simulation is that whenever this particle contributes to a tally, the lower bound of this contribution will be a zero for every contribution in the particles history after this imprecise event.

In general there are multiple events that can occur in a particles history: fission and inelastic scatter are examples. In such cases, where the particle continues to live but has different energy/angle outputs, we believe that the appropriate action is to split the particles attributes and propagate both options. We believe that this action is also appropriate when the particles interval position intersects a geometric barrier, which itself can be uncertain. This however has remains untested. In the following section we describe how these rules can be used to construct particle histories.

3.2. *Particle histories: imprecise exponential branching walks*

A particles history begins by sampling the particle source, the stochastic driver which defines the particles initial attributes. Uncertainty in the source can be included by defining this stochastic driver as a p-box, which may be sampled to give the particle random interval states. The distance that particles travel between events follows an exponential distribution, the parameter of which depends on the properties of the material the particle resides in. While the particle lives, this exponential distribution is sampled defining the particles next position. From its birth to death, the particle follows an exponential random walk. The distance to next interaction is defined by:

$$d = -\frac{ln(\xi)}{\Sigma_t(E)} \quad (6)$$

where ξ is a random number between 0 and 1, and Σ_t is known as the total macroscopic cross section, a cross section which is constructed from the sum of all interaction cross sections the particle can have with the particular material it resides in. This quantity depends on material properties such as density, whose uncertainty may also be represented as an interval. If Σ_t is an interval, the above exponential distribution is a p-box. An interval sample can be generated from this p-box by:

$$[\underline{d}, \overline{d}] = \left[-\frac{ln(\xi)}{\overline{\Sigma}_t}, -\frac{ln(\xi)}{\underline{\Sigma}_t}\right] \quad (7)$$

The interval $[\underline{d}, \overline{d}]$ is added to the particles interval position in the direction of its travel, defining the particles next uncertain interaction location. Following the above description, imprecise particle histories may be simulated. Due to the imprecise events, the same particle history will branch into different reaction channels, splitting the convex hull defined in the particles phase or attribute space. Interestingly if one reduces all uncertainty to zero, the precise transport algorithm is retrieved.

Figure 5 shows various instances of a particles history in its phase space. For simplicity only the energy and distance from source is shown. The top left shows a precise history. The top right shows a particle history with an initial uncertainty in energy. The black box is the convex hull simulated in the interval MC simulation, with individual monte carlo samples shown as points. All the histories simulated by monte carlo fall within this interval. The bottom left figure shows an expanded view of the final interval state of the history of the top right. What can be observed is that the complicated correlated structure produced by monte carlo falls within the interval bounds. Finally the bottom right show a history where initial energy and position are uncertain, with also uncertainty in the nuclear data. Again, all histories fall within the interval bounds. Interestingly, the uncertainty in the particle position increases the more imprecise events it undergoes, suggesting that the further from the source you would like to tally the greater the uncertainty will be.

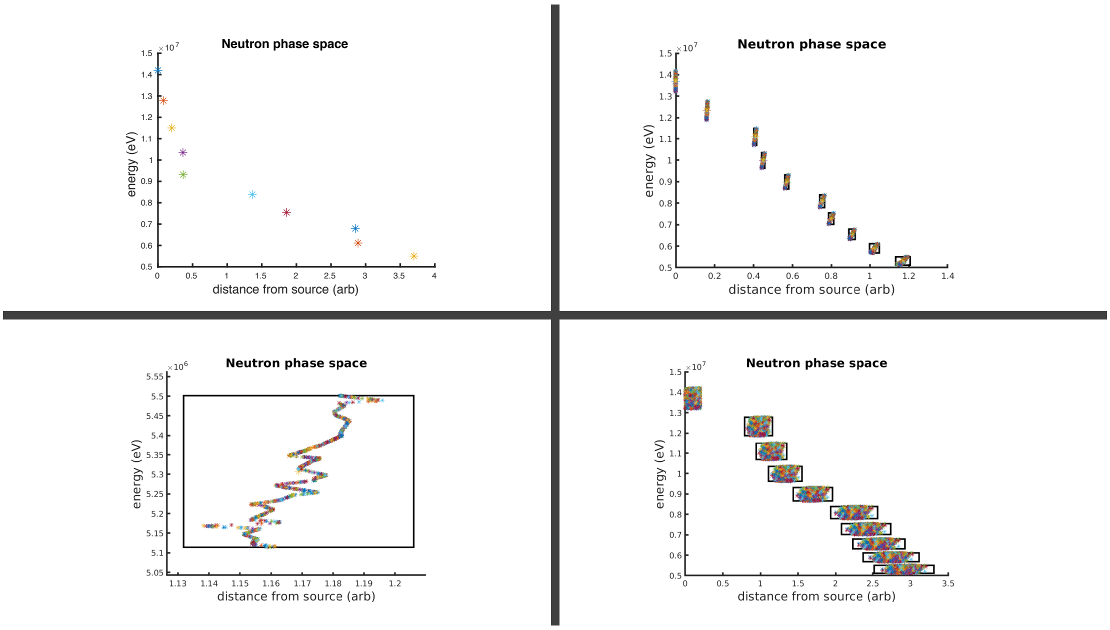

Fig. 5.: Various imprecise particle histories compared to monte carlo. Top left shows the precise case. The top right shows a history with an initial uncertainty in energy. Bottom left shows an expanded view of the convex hull of the particles final state of the top right history. Bottom right shows a history with an initial position and energy uncertainty and imprecise events

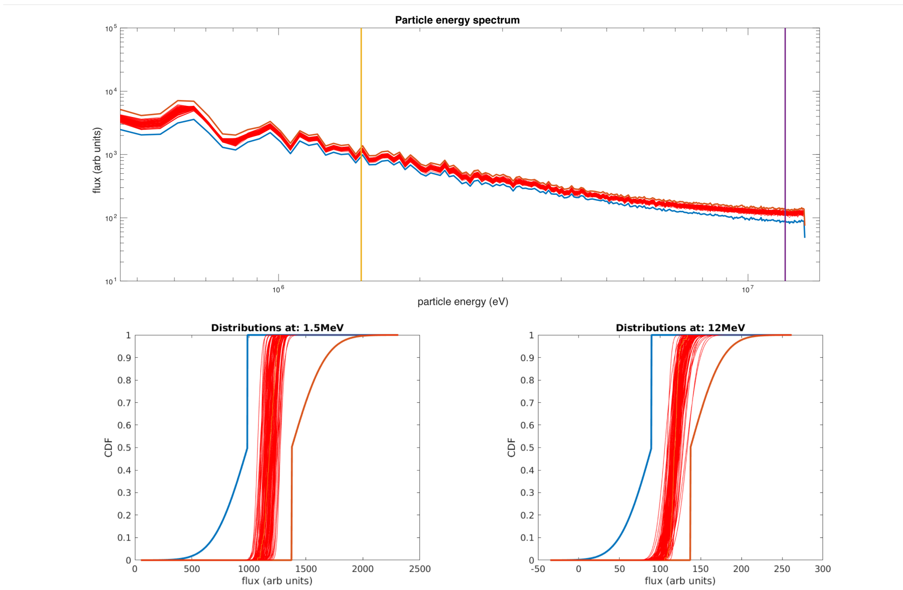

Fig. 6.: The energy spectrum output of an interval MC simulation, compared to monte carlo. Top image shows this spectrum, with the means of the individual monte carlo simulations in red and the bounds on the mean from the interval calculation in blue and brown. The bottom two images shows the p-boxes at the two marked lines in the spectrum, with the individual monte carlo simulation in red

3.3. *Descriptive statistics for tallies: p-boxes*

An entire interval transport simulation may now be run. The main quantity of interest is the particle energy spectrum, the number of particles in a particular volume at a particular energy; since most other transport properties can be derived from it. The monte carlo estimator for flux spectrum is:

$$\phi(E) = \frac{1}{N}\sum_{n} d_i \quad (8)$$

a weighted sum of the distances particles have travelled at a particular energy within that volume. Identical transport simulations may be run, with different random number seeds, producing variations in the monte carlo estimator. The mean and variance over all of the repeated simulations is taken, giving an estimate of the monte carlo error. The interval extension to this estimator is straight forward, where a the lower bound of the estimator is found using the lower bound of the distance, and the upper bound using the upper distance. If a particle undergoes an imprecise absorption event, the contribution to the lower bound is zero from that event forward. By repeating the interval simulation with different random number seeds, a distribution of intervals is created. Ferson et al. (2007) describe an algorithm for performing descriptive statistics on interval data. In this case an interval value for the mean and the variance is produce, which may be used to construct a p-box over this interval data. Figure 6 shows the output of the interval transport simulation with comparison to monte carlo (the source energy has been excluded). The above images shows the bounds on the mean from the interval transport, with the means of the individual monte carlo simulations in red. All means fall within these bounds. The bottom two images show the distributions expanded at the marked lines in the spectrum. In red is the second order distribution produced by Total Monte Carlo, with the p-box produced from the interval particle transport in blue and brown. All Total Monte Carlo samples fall within the distribution bounds of the p-box.

4. Conclusions and open questions

In this paper we present an alternative to monte carlo for uncertainty propagation in particle transport, based on interval analysis. For a very simple simulation setup, without any geometric barriers and only considering elastic scattering and absorption events, the interval transport can propagate uncertainty in a single transport simulation, bounding the monte carlo result. For this method to be generally applicable there are a number of open questions:

- The action to be taken when the particle intersects a geometric barrier
- The simulation of a material with multiple nuclides
- The concurrent propagation of interval hulls which have been split from imprecise events

The last of these is the most challenging. We believe that if these questions are answered, uncertainties from all inputs to the transport model can be propagated in a single simulation. All of the above depends on the quality of the nuclear data evaluation.

Acknowledgement

We thank Scott Ferson and Marco De Anglis (both University of Liverpool). Without their input, this uncertain output would not have been possible. This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom research and training programme 2014-2018 and 2019-2020 under grant agreement No 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

References

Aufiero, M., M. Martin, and M. Fratoni (2016). Xgpt: extending monte carlo generalized perturbation theory capabilities to continuous-energy sensitivity functions. *Annals of Nuclear Energy 96*, 295–306.

Ferson, S., V. Kreinovich, L. Grinzburg, D. Myers, and K. Sentz (2015). Constructing probability boxes and dempster-shafer structures. Technical report, Sandia National Lab.(SNL-NM), Albuquerque, NM (United States).

Ferson, S., V. Kreinovich, J. Hajagos, W. Oberkampf, and L. Ginzburg (2007). Experimental uncertainty estimation and statistics for data having interval uncertainty. *Sandia National Laboratories, Report SAND2007-0939 162*.

Ferson, S. and K. Sentz (2016). Epistemic uncertainty in agent-based modeling. In *7th international workshop on reliable engineering computing*.

Koning, A. J. and D. Rochman (2008). Towards sustainable nuclear energy: Putting nuclear physics to work. *Annals of Nuclear Energy 35*(11), 2024–2030.

Koning, A. J. and D. Rochman (2012). Modern nuclear data evaluation with the talys code system. *Nuclear data sheets 113*(12), 2841–2934.

Moore, R. E. (1979). *Methods and applications of interval analysis*, Volume 2. Siam.

Rochman, D., A. Koning, and S. Van Der Marck (2010). Exact nuclear data uncertainty propagation for fusion neutronics calculations. *Fusion Engineering and Design 85*(5), 669–682.