Chapter 7
of
Radiotherapy Radiation Physics: A Monte Carlo Approach

David W. O. Rogers

latexed October 10, 2024

Dave Rogers homepage and contact information

Chapter 7

Monte Carlo

Learning Objectives

- to know enough about Monte Carlo to properly use it in treatment planning systems.
- to understand the basic Monte Carlo techniques of photon and electron transport.
- to understand the role of photon and especially electron low-energy transport cutoffs.
- to understand the difference between Class I and Class II electron transport algorithms.
- to be aware of the major widely available general purpose Monte Carlo codes.
- to be aware of some terminology specific to the EGSnrc Monte Carlo code since it will be used elsewhere in this book.
- to understand the meaning and importance of the Fano test of a Monte Carlo code.
- to understand the difference between variance reduction techniques (VRTs) and approximate efficiency improving techniques (AEITs) and how a few examples work.
- to learn how statistical uncertainties are calculated using the batch method or the history by history method
- to understand how to improve statistics in a calculation without using VRTs or AEITs
- to appreciate where systematic uncertainties in Monte Carlo calculations may come from and how to estimate them

Prerequisites/Necessity of this chapter

- Ch 3
- Essential background for effective use of Monte Carlo in the clinic but details of section 7.V. are only important for people doing their own Monte Carlo calculations.

Contents

7.I. What is Monte Carlo

The AAPM's TG105 report on Monte Carlo treatment planning states: "The Monte Carlo technique for the simulation of the transport of electrons and photons though bulk media consists of using knowledge of the probability distributions governing the individual interaction probability distributions governing the individual interactions of electrons and photons in materials to simulate the random trajectories of individual particles. One keeps track of physical quantities of interest for a large number of histories to provide the required information about the average quantities" [1,2]. As Monte Carlo techniques now play a fundamental role in much of radiotherapy physics, this chapter outlines some of the basics of the technique. Given my history, the EGSnrc code is used as an example, but the general methods are the basis of all the major codes which are discussed in section 7.VII..

7.I.i. A simple photon transport example

Consider a simple example of photon transport where the only relevant interactions are Compton scattering and pair production where the total cross section is:

$$\Sigma_{\text{tot}} = \Sigma_{\text{comp}} + \Sigma_{\text{pair}}. \qquad [\text{cm}^{-1}] \tag{7.1}$$

The first questions are, how far does the photon go before interacting, and then what interaction will occur? The distance x to the interaction is given by (see section 7.I.ii.c below):

$$x = -\frac{1}{\Sigma_{\text{tot}}} \ln \xi_1, \qquad [\text{cm}] \tag{7.2}$$

where ξ_1 is a random number uniformly distributed on $(0, 1]$. The value of x is exponentially distributed on $[0, \infty)$. [n1] Once the pathlength has been determined, sample a second random number ξ_2 and determine whether a Compton or pair interaction has occurred via:

$$\xi_2 \leq \frac{\Sigma_{\text{comp}}}{\Sigma_{\text{tot}}} \Rightarrow \text{Compton interaction} \tag{7.3}$$

$$\xi_2 > \frac{\Sigma_{\text{comp}}}{\Sigma_{\text{tot}}} \Rightarrow \text{pair interaction} \tag{7.4}$$

This simple algorithm demonstrates the major techniques of Monte Carlo simulation. The simple log of the random number is used in many situations. Sampling from more complex distributions (e.g., the angle of the electrons and photons after the interactions in the simple example) can be much more complicated and require much more cross section data. Kawrakow gives a detailed description of all the algorithms used in EGSnrc in the code's manual[3]. But the process is fundamentally the same. A knowledge of the probability distribution (in this case, the exponential distribution with the total cross section) is used in combination with random numbers to determine the physical quantities (in this case, the distance to the interaction and which type).

Fig. 7.1 shows a more complete algorithm for the transport of photons in a real geometry where all interactions are considered and the geometry plays a role. The concept of a 'STACK' is important since during a simulation of just one initial particle there can be

many secondary and/or scattered particles and these are placed on a 'STACK' which is basically a list of all the particles that need to be followed. All codes need to systematically track them until they fall below some energy cutoff or they escape the geometry of interest.

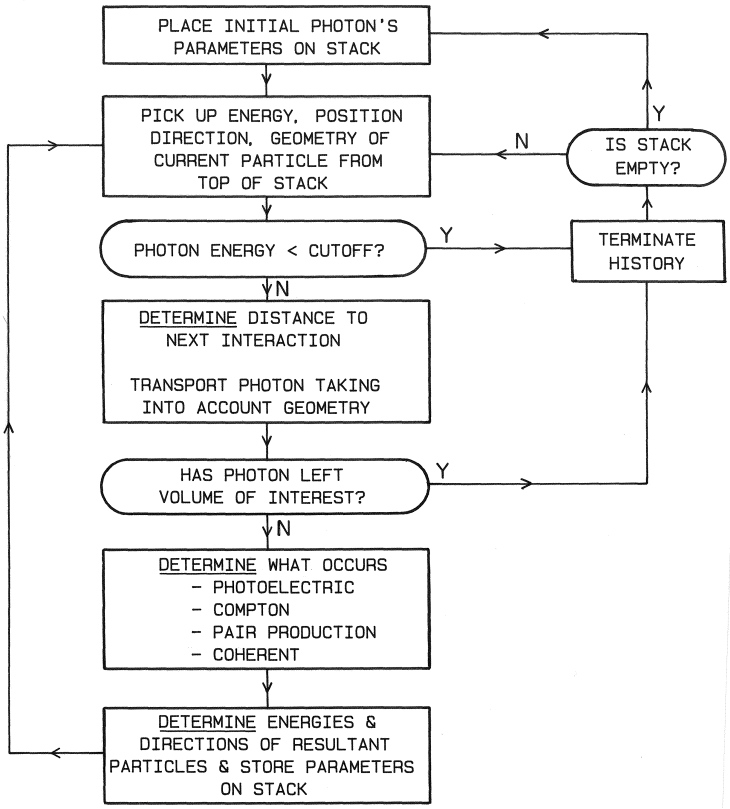

Figure 7.1: Logic of a Monte Carlo algorithm for photon transport. 'DETERMINE' means to get the relevant physical quantity using knowledge of the physics probability distributions and some random numbers. From ref[2].

7.I.ii. Sampling

7.I.ii.a Random number generation

An essential component of all Monte Carlo codes is the generation of random numbers as seen even in the simple example in the previous section. Researchers have spent their lives developing good random number generators (RNGs) but great care is required to make sure the sequences are actually random and that they do not repeat. L'Ecuyer has given a detailed history of random number generation[4] and has also developed a battery of tests for the testing of RNGs[5]. It is best to use well tested and packaged RNGs since good quality numbers are essential. Two such packages are the ranmar generator proposed by Marsaglia and Zaman in 1991[6] and the ranlux generator written by James[7] based on the algorithm of Lüsher[8].

Pseudo-random numbers are those generated by software algorithms as opposed to those based on some random physical process (e.g., rolling dice). An important feature is that given

the same initialization, pseudo-random generators produce the same sequence of random numbers. This can be essential when debugging Monte Carlo programs. So-called quasi-random numbers or Sobol random numbers are also generated by software algorithms but are more uniformly sampled[9,10]. These can improve the efficiency of some calculations despite taking longer to calculate.

An important feature of RNGs is their period, i.e., how many numbers can it generate before repeating the sequence. The period of the ranlux generator is said to be about 10^{171} and the ranmar generator can have periods of 10^{171} up to 10^{527} so there is little chance of running out of random numbers[6,7]. Other useful features of RNGs are the ability to store their current state and be restarted at the same point and/or to produce completely independent sequences when initiated with different initializations. This latter feature is mandatory for calculations which are to be run in parallel on separate machines since one needs to know the runs are independent.

7.I.ii.b Probability distribution functions: PDFs

A probability distribution function (PDF), $p(x)$, gives the probability of obtaining the value x in the interval $[x - \mathrm{d}x/2, x + \mathrm{d}x/2]$. It must be non-negative and can be defined on any region $[a, b]$ where $\pm\infty$ are allowed. It should be normalized so that

$$\int_a^b p(x)\mathrm{d}x = 1. \tag{7.5}$$

The associated cumulative distribution function(CDF) is defined as

$$c(x) = \int_a^x p(t)\mathrm{d}t, \tag{7.6}$$

and obviously $c(b) = 1$ which follows from the normalization in Eq. 7.5.

7.I.ii.c Direct or Inverse Sampling

Given the cumulative distribution function of a probability distribution function, one can sample a value x by taking a random number, ξ, which, like the CDF, lies between 0 and 1, and solve for x using

$$\xi = c(x) \Rightarrow x = c^{-1}(\xi). \tag{7.7}$$

In principle, $c(x)$ can be determined by numerical integration as in Eq. 7.6 but ideally can be done analytically. As an example of the latter, consider the PDF for a particles distance to its next interaction, *viz.*[n2]

$$p(x) = \Sigma_{\mathrm{tot}} e^{-\Sigma_{\mathrm{tot}} x}, \tag{7.8}$$

from which the CDF is

$$c(x) = \Sigma_{\mathrm{tot}} \int_0^x e^{-\Sigma_{\mathrm{tot}} t}\mathrm{d}t \qquad = \Sigma_{\mathrm{tot}} \left[\frac{e^{-\Sigma_{\mathrm{tot}} t}}{\Sigma_{\mathrm{tot}}}\right]_0^x \qquad = 1 - e^{-\Sigma_{\mathrm{tot}} x}. \tag{7.9}$$

Setting $\xi = c(x)$ leads to

$$\begin{aligned} \xi &= 1 - e^{-\Sigma_{\mathrm{tot}} x} \Rightarrow 1 - \xi = e^{-\Sigma_{\mathrm{tot}} x} \Rightarrow \ln(1-\xi) = -\Sigma_{\mathrm{tot}} x \\ \Rightarrow x &= -\frac{1}{\Sigma_{\mathrm{tot}}} \ln(1-\xi) \qquad \equiv \qquad x = -\frac{1}{\Sigma_{\mathrm{tot}}} \ln(\xi), \end{aligned} \tag{7.10}$$

which is just the result used in section 7.I.i. for sampling the distance travelled by a photon before interacting. The final equivalence follows since, if ξ is uniformly distributed between 0 and 1, so is $1-\xi$. Note that this latter, oft-used expression requires that the RNG cannot return 0 although 1 is allowed.

7.I.ii.d Rejection sampling

If the inverse of the cumulative distribution function, $c(x)$, is difficult to invert, the rejection sampling technique may be appropriate, especially if the probability distribution is not too peaked. The rejection sampling algorithm to sample an arbitrary $p(x)$ distribution is:

- scale the PDF to its maximum value: $g(x) = p(x)/p(x_{\max})$ so $g(x_{\max}) = 1$. If finding $x_{\max}$ is complex, one can use any value for $p(x_{\max})$ which is an upper bound on $p(x)$.
- select a random number, ξ_1, uniformly in $(0, 1]$ and convert this to an x value in the range of the probability distribution, i.e., $[a, b]$ using a simple transformation $x = a + (b-a)\cdot\xi_1$. This assumes a and b are both finite.
- select a second random number, ξ_2 and consider x as accepted if $\xi_2 \le g(x)$, otherwise reject x and start again with another ξ_1.

The efficiency of this method, i.e., the fraction of the time that a particular x value is accepted is given by the ratio of the area under the PDF between a and b and the area where the random numbers are placing points, i.e., $\epsilon = \int_a^b p(x)\mathrm{d}x \Big/ (p(x_{\max})(b-a).)$ The flatter the PDF, the better the efficiency although compared to the inverse sampling technique, 2 random numbers are needed for every accepted value rather than 1. This can nonetheless be more efficient if calculating and inverting the CDF is computationally expensive.

In practice, rejection sampling may not follow the steps above in an obvious manner. For example, consider sampling points in a circle of radius R. Start by sampling points randomly in a box defined by $-1 \le x \le 1$ and $-1 \le y \le 1$. This is done using $x = 2\cdot\xi_1 - 1$ and $y = 2\cdot\xi_2 - 1$ so that x and y are between -1 and 1. Choose to accept these values if $x^2 + y^2 \le 1$, i.e., the point is in a circle of radius 1. Then return $x = R\cdot x$ and $y = R\cdot y$, otherwise use 2 new random numbers and start again until the values are accepted.

7.I.ii.e Realistic sampling

While the previous 2 sections introduce some basic methods, in practice, sampling the realistic probability distributions common in Monte Carlo transport can require very sophisticated sampling routines. These combine methods and use many ingenious algorithms for sampling related to processes such as bound Compton scattering or bremsstrahlung emission. It is best left to the experts!

7.I.iii. Electron transport basics: discrete vs continuous energy loss

As a charged particle slows down it undergoes thousands of elastic collisions with atoms and nuclei. These change the direction of the particle but are too numerous to model individually and hence most Monte Carlo codes use a condensed history (CH) technique to model charged particle transport. In CH transport the particle's path is broken into a

series of multiple-scattering steps in which the effect of many interactions are grouped. A multiple-scattering theory (see section **??**) accounts for the angular deflections during the step while at the same time the electron loses energy, either in a continuous mode via creation of low-energy secondaries and excitations or in discrete interactions creating higher-energy secondaries, either electrons or brems photons. Fig. 7.2 shows this schematically where the energy lost in the continuous mode is shown in the hatched areas immediately around the electron path. Each multiple-scattering step is thought of as a straight path with an angular deflection at the end of the step. When a secondary electron (or brems) is created, this both decrements the initial electron's energy and deflects the initial electron (for class II CH transport algorithms: see below).

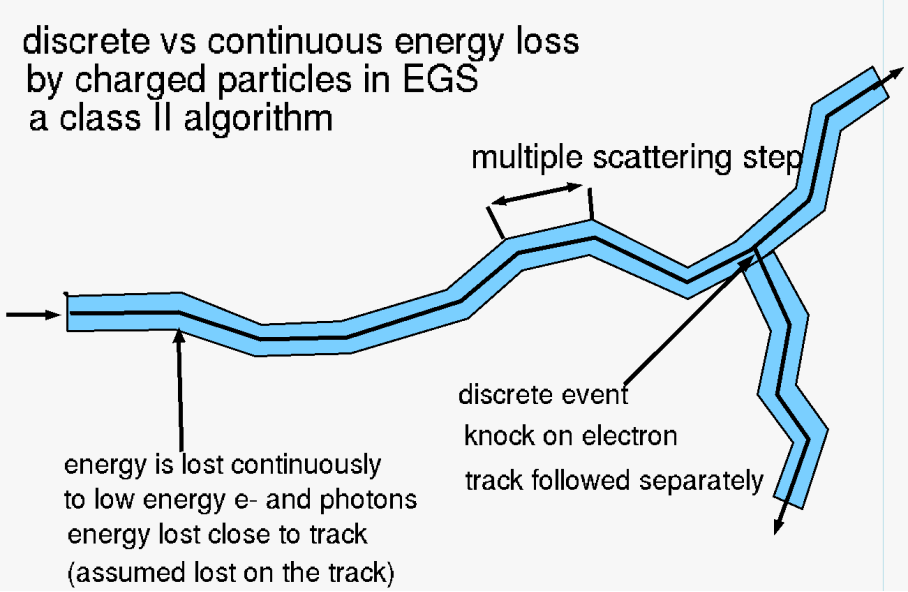

Figure 7.2: Schematic of an electron step for a Class II electron transport algorithm. Based on ref[2].

7.I.iii.a **Definition of production thresholds (AE and AP)**

It is useful to have notation for the thresholds for the creation of secondary electrons or brems. The terminology for EGSnrc is used here for convenience, but all condensed history electron-photon Monte Carlo codes have equivalent (but not always identical) quantities.

AE is the threshold for the production of secondary electrons and plays the same role as Δ does for restricted stopping powers. For historical reasons, the value of AE includes the rest mass of the charged particle (511 keV).

$E_{\rm sec} > AE$: treat the interactions discretely and follow the secondaries.

$E_{\rm sec} \leq AE$: treat the interactions as part of the continuous energy loss mechanism represented by the restricted electronic stopping power, $(L(\Delta)/\rho)_{\rm el}$.

AP is the threshold for the production of brems photons:

$E_{\rm brems} > AP$: treat the interactions discretely and follow the brems photons.

$E_{\rm brems} \leq AP$: treat the interactions as part of the continuous energy loss mechanism represented by the restricted radiative stopping power, $(L(\Delta)/\rho)_{\rm rad}$.

7.I.iii.b **Definition of transport cutoffs: (ECUT, PCUT)**

In electron and photon transport, the particles are tracked until they exit the geometry of interest or have an energy below some cutoff. When the particle is below the cutoff, its energy is deposited locally (except for positrons which annihilate and the radiative photons escape). These cutoffs are called ECUT and PCUT for charged particles and photons respectively. Logic requires ECUT $\geq$ AE and PCUT $\geq$ AP since it doesn't make sense to track particles to an energy below the lowest energy at which they can be created. In electron-photon Monte Carlo simulations, the electron transport takes the vast majority of the computing time. Furthermore, since, loosely speaking, step sizes are a constant fraction of the current energy and computing time is proportional to the number of steps taken, this means it takes almost as long to track an electron from 20 MeV to 10 MeV as from 20 keV to 10 keV. Hence the overall computing time is very sensitive to the value of ECUT used. Since the range of low-energy electrons is very short (see section **??**), in many situations it is not necessary to track them. For example, the CSDA range of 100 keV electrons is 0.14 mm but as seen in Fig. 7.3(left), few of them travel that far in their initial direction. So unless the specific situation has a geometric scale of the order of 0.1 mm in water, there is little need to transport electrons below that energy. Fig. 7.3(right) shows remarkably little variation even for an ECUT value of nearly 4 MeV kinetic energy where the CSDA range of electrons is 2 cm. In this case, because the electrons' directions are nearly isotropic at depth, the effects of terminating the history is much less than expected based on the residual range. Perhaps more importantly, the dose vs depth curves with ECUT values of 700 and 1000 keV lie on top of each other. There is no need to track electrons below 489 keV kinetic energy in this case.

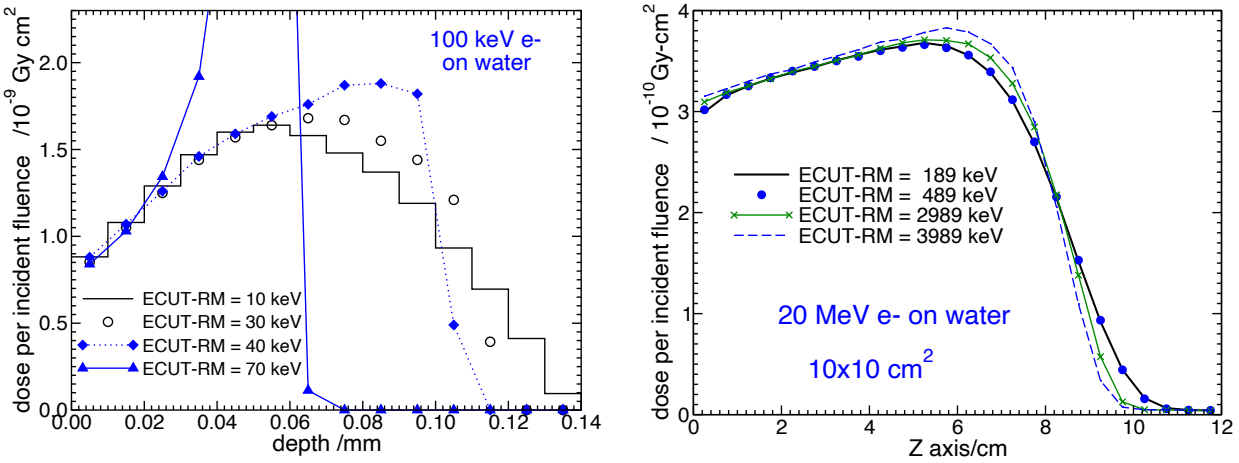

Figure 7.3: Variation in 100 keV (left) and 20 MeV (right) electron beam dose vs depth curves as a function of ECUT, the lowest energy to which electrons are transported. The rest mass, RM, is subtracted from ECUT to show the kinetic energies. Scoring regions are 5 mm thick in the right panel.

7.I.iv. Class I vs Class II electron transport algorithms

In his seminal paper on condensed history (CH) Monte Carlo transport algorithms, Martin Berger[11] differentiated between 2 classes of CH algorithms (see Fig. 7.4).

In a Class I algorithm on each multiple-scattering step of length t, the energy loss is determined by sampling from an energy-loss straggling distribution, $\Delta E(t)$ (e.g., a Landau distribution[12]). The probability of creating a secondary is sampled independently and if created, it has an energy of E_δ. The energy deposited locally is given by $E_{\rm dep} = \Delta E(t) - E_\delta$. Note that for any given step in the Class I algorithm, energy is not conserved since E, the energy at the end of the step, does not take into account the specific value of E_δ since it is already accounted for in the energy-loss straggling distribution. There is energy conservation in a statistical sense since the distribution $\Delta E(t)$ takes into account these energy losses as well as the energy loss to sub-threshold events.

In a Class II CH algorithm, the step length can be determined by the distance to the creation of a secondary particle (secondary electron or brems photon). The energy at the end of the step accounts for the continuous energy loss to sub-threshold events and the explicit energy loss from creating a secondary electron or brems photon. In Class II algorithms there is energy conservation on every step and an explicit correlation between the final energy and the energy of the secondary.

In contrast to the Class II algorithms where the step size can be limited by the creation of a secondary above AE or AP, in a Class I algorithm, the step size is defined through a predetermined set of energies. Of course, geometrical boundaries between materials can also affect the step size in both classes.

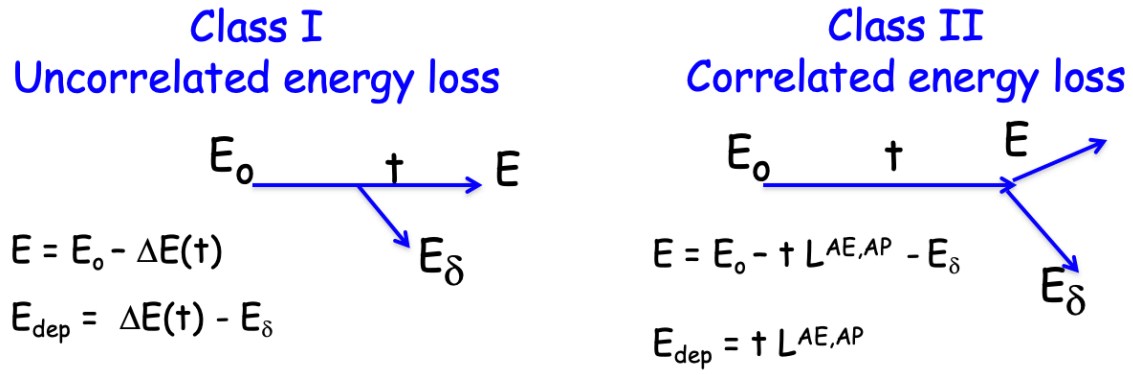

Figure 7.4: Schematic of Class I vs Class II condensed history algorithms. $\Delta E(t)$ represents an energy-loss straggling distribution which is sampled for the step length t and $L^{\rm AE,AP}$ is the restricted total stopping power for the AE,AP cutoffs defined in sec 7.I.iii.a. The step length t is different in each case.

7.I.iv.a 20 MeV electrons through a thin slab

To further understand how electron transport algorithms work, it is useful to consider a simple example: the spectra of electrons emerging from a 2.5 mm slab of water irradiated by 20 MeV electrons. Fig. 7.5 shows the angular distribution of the emerging electrons as calculated in the CSDA model or in a simulation allowing the creation of electrons or

photons above 10 keV kinetic energy. Although the creation of secondaries can deflect the primary electrons, the close agreement between the CSDA and primaries from the full simulation shows that these deflections are not critical in this case. There are also relatively few secondary electrons which escape from the slab. Fig. 7.6 shows the spectrum of emerging electrons when the CSDA is used. Virtually all electrons have the same energy which is given by the pathlength times the unrestricted total stopping power for 19.7 MeV (on average) electrons. There is a very small lower-energy tail (only visible on this 5-cycle log scale) caused by electrons taking a slightly longer path due to multiple scattering deflections.

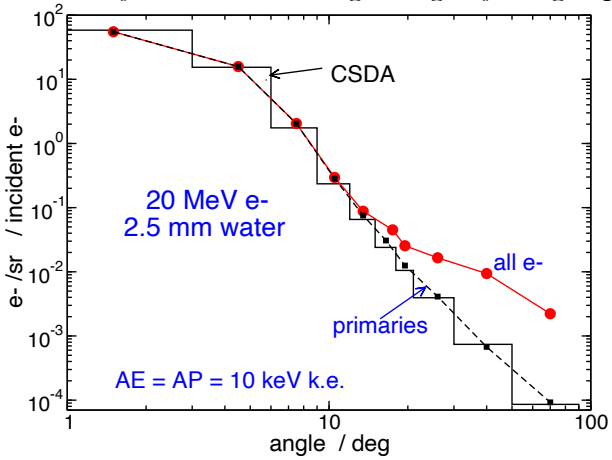

Figure 7.5: Angular distribution of 20 MeV electrons through 2.5 mm of water as calculated using a CSDA model (histogram) or a simulation creating electrons and photons with kinetic energies above 10 keV where either all electrons (circles) or only the primary electrons (dash line, squares) are scored.

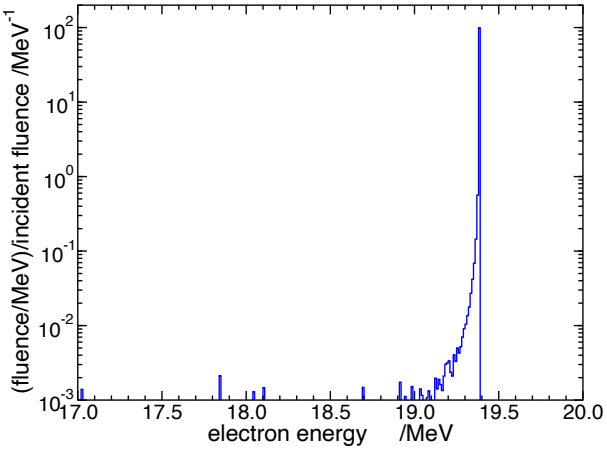

Figure 7.6: Spectrum of electrons after passing through a 2.5 mm slab of water. Simulation was done in CSDA.

Fig. 7.7 presents the emerging spectra of electrons where the Class II calculations are done with various AE, AP values of 1 MeV, 100 keV, 10 keV or 1 keV kinetic energies. For this thin 2.5 mm slab of water, the vast majority of the electrons pass through the slab with the same pathlength, just slightly greater than 2.5 mm. The average energy loss in all cases is 618 keV. The peaks are from electrons which have not created any secondaries and since the restricted stopping power that governs their energy loss is lower, the energy losses are less and the peak's energies higher as the values of AE and AP decrease. If an electron creates a secondary electron or brems, the energy lost to the secondary is subtracted from the value at the peak. Since all secondaries in the calculations shown in the upper-left panel are at least 100 keV in energy, there is a 100 keV 'valley' below the peak. In this upper-left panel there is distinct increase in the number of electrons with energies at least 1 MeV below the peak. These electrons have produced secondary electrons above the AE threshold of 1 MeV kinetic energy and thus have lost at least 1 MeV. This increase in the number of electrons per energy bin at least 1 MeV below the peak implies that the cross section for creating 1 MeV brems photons is considerably smaller than that for creating secondary electrons of similar energy. This spectrum of energies is clearly not physical but represents an artifact of the Class II algorithm and the use of relatively high values of AE and AP. The upper-right

panel of Fig. 7.7 uses much lower thresholds and here the spectrum is much more realistic although a 10 keV gap is still visible below the high-energy peak.

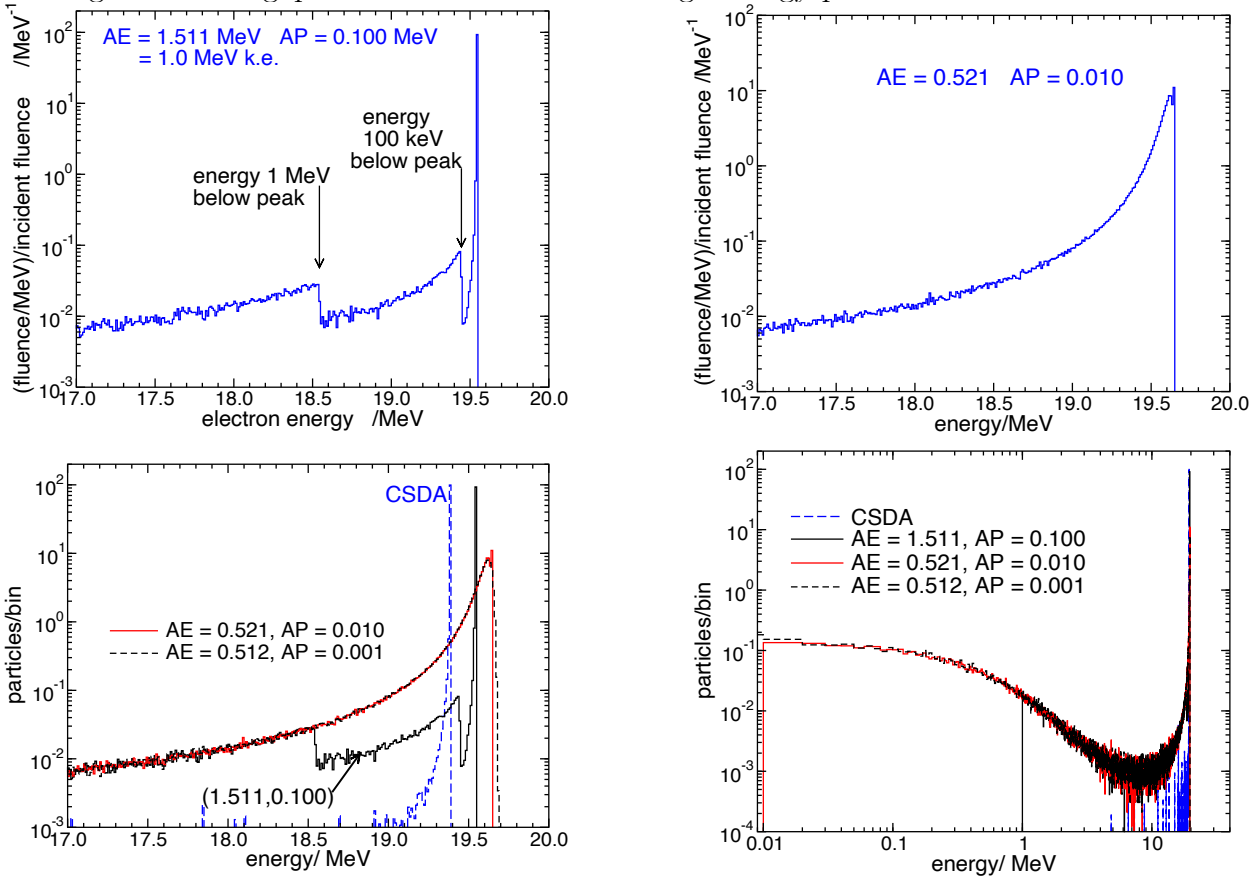

Figure 7.7: Energy spectra of 20 MeV electrons after passing through a 2.5 mm slab of water. Upper-left panel shows a calculation creating secondary electrons with kinetic energy $>$ 1 MeV (AE=1.511 MeV) and brems photons with energies >100 keV (AP=100 keV). The upper-right panel uses thresholds of 10 keV kinetic energy for secondary electrons and brems photons. The lower panels compare the spectra for several different values of AE and AP. The calculation times were 77, 84, 144 and 1130 s for decreasing AE,AP values respectively. The mean energy loss of primaries escaping the slab is 618 keV in all cases.

The incident electrons cannot lose more than 1/2 of their energy while creating a secondary electron although they can lose virtually all of their energy while creating a brems photon. In this thin slab of water, 17% of the energy loss is via brems with 99% of brems photons above 200 keV and 80% above 2 MeV. The low-energy electrons in the lower-right panel of Fig. 7.7 will be a mixture of secondary electrons and primary electrons emitting high-energy brems.

The lower panels of Fig. 7.7 show the same spectra but compare calculations with various AE, AP settings. As the AE, AP values get lower, the spectra become more realistic in nature, the peak energy continues to increase as the restricted stopping powers decrease and the computing time takes much longer. The mean energy loss of the primary electrons escaping the slab is the same 618 keV in all cases. The lower-right panel shows the entire spectrum down to 10 keV but this includes all the secondary electrons. The vertical line at

1 MeV shows that for the AE=1.511 MeV (i.e., 1 MeV kinetic energy) case, there are no electrons below 1 MeV since they are not created below 1 MeV and it is a characteristic of EGSnrc that there can be no electron transport below AE.

When a similar calculation is done using a Class I algorithm, the spectrum looks like the calculation here with AE, AP values of 1 keV kinetic energy because the energy-loss straggling distribution correctly takes into account the straggling. [n3].

While the spectra in Fig. 7.7 look quite different, if they are used as initial spectra to calculate a dose vs depth curve in a water phantom they all produce virtually identical results.

7.I.v. Determining electron step size

Step size effects play an important role in determining the accuracy of an electron transport algorithm (see next two sections) and this accuracy is related to how well multiple scattering of the charged particles (see section **??**) is handled. Various aspects of electron transport, such as the correction for the true curved pathlength vs the straight path assumed in each multiple-scattering step(see Fig. 7.4) become more accurate as step sizes decrease. However, the shorter the step, the longer the calculation time and if the steps are too short, the multiple-scattering algorithms can break down since there is no longer multiple scattering. Hence, to ensure as long a step as possible, an important part of any transport algorithm is the pathlength correction (PLC) which relates the actual pathlength traversed by the charged particle to the straight-line step taken in the simulation. There can be restrictions on the maximum amount of multiple scattering allowed in a step to ensure that the pathlength correction is accurate and this leads to a restriction on the pathlength of each step.

But other things also affect individual step sizes. Multiple-scattering steps cannot cross boundaries so the distance to a boundary in the current direction can limit the step size. Multiple-scattering theories assume scattering in an infinite medium, so if a particle step is close to a boundary with another medium, even if it is not directed at the boundary, the step size may need to be restricted (see further discussion below). The step size is also shortened in a Class II algorithm (see section 7.I.iv.) when a discrete event is about to occur, either creating a brems photon or a secondary electron. This distance is strongly affected by the energy thresholds for these events (see Figures **??** and **??**). A convenient way to to ensure that the various electron transport algorithms don't break down is to limit the maximum fractional energy loss per step (e.g., the `ESTEPE` parameter in EGSnrc).

There is a complex interplay of all these components that determine the length of a given step, and these play an important role in determining the computing time required for a calculation.

7.I.vi. Electron transport algorithms

Even ignoring the issue of boundary crossing (see section 7.I.vii.), the algorithms for transport of charged particles is much more complex than that for photons. Fig. 7.8 shows various algorithms of increasing complexity. Different Monte Carlo codes have used a variety of these options with various approaches to determining the pathlength correction (PLC) or the lateral correlation algorithm (LCA). The algorithm of Kawrakow(Fig. 7.8d) implemented

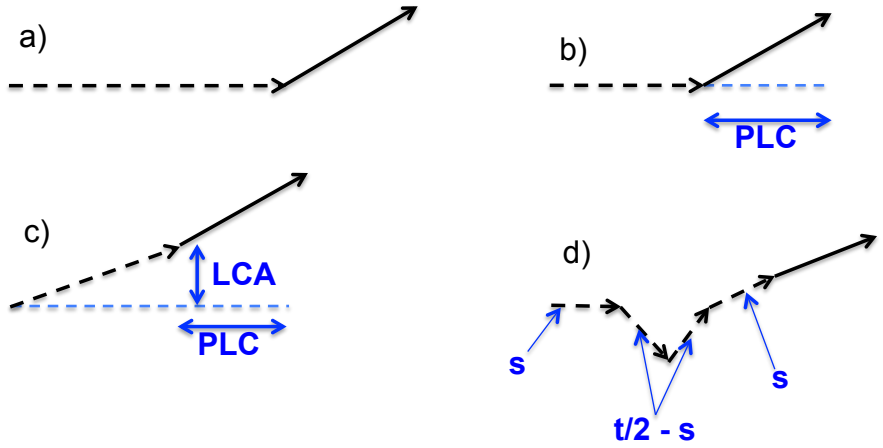

Figure 7.8: Electron transport algorithms of increasing sophistication. In panel a) the electron travels the full step size in the original direction and then samples a multiple-scattering angle. In panel b), the step taken is reduced by a pathlength correction (PLC) to account for the actually curved nature of the path. In panel c) a lateral deflection based on a lateral correlation algorithm(LCA) is added to account for the change in direction during the step. Panel d) represents the electron transport algorithm of Kawrakow discussed in the text and implemented in EGSnrc[13].

in EGSnrc is one of the most robust and efficient since it allows for relatively large steps while maintaining accuracy[13,14]. The algorithm consists of the following.

- break each step of curved pathlength t (see section 7.I.v. about determining t), into two substeps
- select polar and azimuthal multiple-scattering angles independently for each sub-step
- final direction is from combining both angles
- transport electron a distance s in original direction (s selected randomly as a fraction of $t/2$)
- apply first multiple-scattering angle to the direction
- transport electron a distance $t/2 - s$ in that direction
- assign the electron a new direction based on second multiple-scattering angle with respect to the initial direction and transport a distance $t/2 - s$
- transport the electron a distance s in the final direction.

This algorithm, when implemented with an accurate multiple-scattering theory produces remarkably accurate results for relatively large step sizes. Fig. 7.9 demonstrates that, with the exception of the very low-probability case of electrons scattered directly backwards, the algorithm matches the results of a single-scattering calculation for a difficult case with a large amount of multiple scattering.[n4]

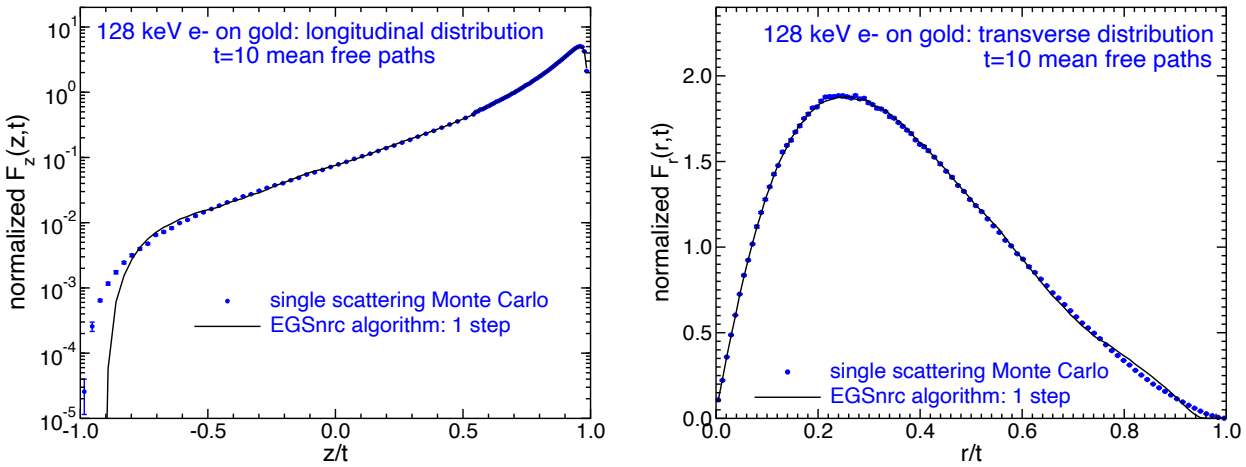

Figure 7.9: For the case of 128 keV electrons travelling about 10 mean free paths in a large gold phantom, a comparison of normalized F_z(z,t) and F_r(r,t) values as a function of z/t or r/t for either an analogue single-scattering calculation or a single step (2 substeps) application of the EGSnrc algorithm described above. F_z(z,t) is the probability of finding the electron at z after travelling path-length t normalized to unit total probability and F_r(r,t) is the corresponding value for finding the electron at a lateral distance r after path-length t. Based on data in figures from ref[14].

7.I.vii. Boundary crossing issues

In section 7.I.v. it was pointed out that boundaries with different geometric regions can cause the overall step size to be reduced. The most obvious reason is that as one crosses into the new region, various cross sections may change and hence one needs to stop or otherwise deal with the boundary. However, even if not crossing the boundary, since multiple-scattering theories assume transport in an infinite medium, assumptions break down if the particle is too close to a boundary. Dealing with boundary crossing is perhaps the most difficult part of modelling electron transport.

In addition to the step size restrictions discussed in section 7.I.v., one approach[15] to boundary crossing is to limit the curved pathlength for any given step to the minimum of $t_{\rm max}$ or $t_{\rm perp}$ where $t_{\rm max}$ is the maximum step for which the multiple-scattering PLC holds and $t_{\rm perp}$ is the perpendicular distance to the closest boundary. This means the charged particle can be considered to be moving entirely in the current medium. However it also means the charged particle will never reach the boundary since its straight-line step size is less than $t_{\rm max}$ or $t_{\rm perp}$ due to the pathlength correction. The solution is to ignore multiple scattering once the particle is sufficiently close to the boundary and transport the particle to the boundary. Since this is only a very short final step, the effect of ignoring the PLC is negligible.

However, as Smyth and Foote pointed out[16,17], having a charged particle stop at a boundary leads to an artifact in the calculation of ion chamber response if it causes a new multiple-scattering angle to be selected at the boundary. To overcome this problem, Kawrakow's algorithm[13], rather than ignoring multiple scattering when very close to the boundary, switches to single-scattering mode. The electron steps are now straight between interactions. The electron still stops on the boundary to sample from new cross sections if the material changes,

but continues in a straight line until the next single scattering. As the particle leaves the boundary it continues in single-scattering mode until it is considered far enough from the boundary to recommence the full multiple-scattering transport algorithm.

Using these complex algorithms along with accurate multiple-scattering algorithms and cross sections, the accuracy of transport algorithms can be assessed using various forms of what is known as a Fano test described below in section 7.VI..

7.II. Structure of electron beam dose vs depth curves

Although general aspects of electron beam dose vs depth curves are dealt with in section **??**, the present section highlights how different aspects of electron transport algorithms affect Monte Carlo calculations of electron beam dose vs depth curves. This is done by turning on and off various physical processes to show their effects. Fig. 7.10 presents CSDA calculations for broad parallel beams (BPBs) of 20 MeV electrons where no secondary electrons or photons are created and hence no energy-loss straggling is present. The histogram does not include multiple scattering whereas the curve with symbols includes multiple scattering. The peak at the end of the range of the particles is because all particles reach the same depth before stopping and depositing their remaining kinetic energy (189 keV). The size of this peak is a calculational artifact which depends on the width of the bin the histories terminate in and on the cutoff energy, ECUT. With no multiple scatter the dose falls slightly with distance from the surface as the total stopping power decreases as the energy falls (see Fig. **??**) until the electrons reach below 1 MeV. At this energy the stopping power increases and leads to an increased dose just before the peak at the end of their range. The other curve includes the effect of multiple scattering which leads to a lateral spreading of the electrons which thus shortens the depth of penetration of most electrons and an increase in the dose before dose maximum because the fluence has increased. In this case, the depth-straggling is entirely caused by the lateral multiple scattering since every electron has travelled the same distance along its individual path (review Fig. **??**).

Fig. 7.11 presents the case of no multiple scatter for the same 20 MeV BPB of electrons. The curve with no energy-loss straggling (see section **??**) is in principle an identical simulation to the same curve in Fig. 7.10 but the peak's height is very different due to the different size of depth-bin. For the other two cases there is energy-loss straggling due either to the creation of brems photons (one case) or secondary electrons (a second case). Those electrons which do not undergo any discrete interactions have a pathlength greater than the CSDA range (Fig. **??**) and hence the dose vs depth curves stretch past R_{csda}. There is a buildup region near the surface for the case where secondary electrons are created until the secondaries reach a type of equilibrium.

Fig. 7.12 shows dose vs depth curves for the same 20 MeV BPB of electrons with multiple scatter turned on and 4 different levels of straggling. The no-straggling case leads to the highest maximum dose due to the multiple scattering.

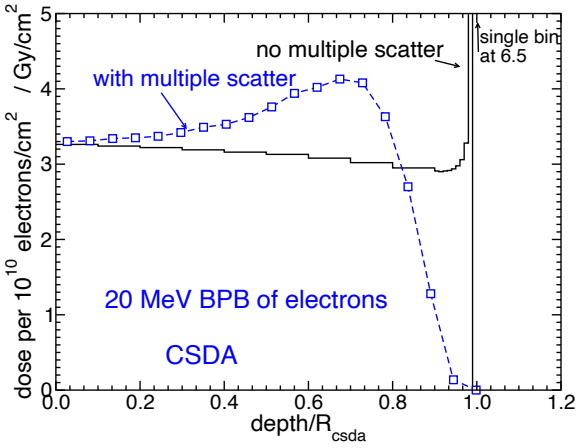

Figure 7.10: Broad parallel beam of 20 MeV electrons on a water phantom in the CSDA. The histogram has no multiple scatter whereas the smooth curve does. From data in ref[2].

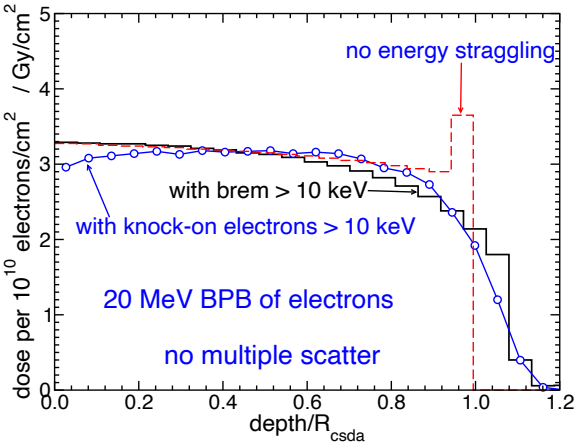

Figure 7.11: As in Fig. 7.10 but with no multiple scatter and either no energy-loss straggling (i.e., CSDA) or with straggle from creation of secondary e- in one case or brems photons in the other.

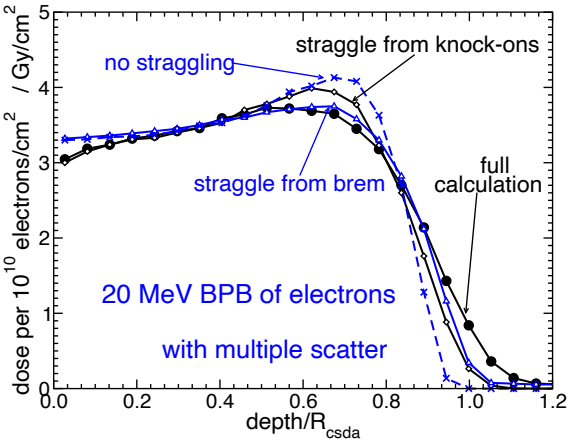

Figure 7.12: As in Fig. 7.11 but with multiple scatter and also a full calculation with both types of energy straggling.

7.III. Statistical uncertainties

The results of Monte Carlo calculations are inherently statistical and thus any result must include an estimate of the statistical uncertainty on the value. These estimates can be determined in a variety of ways. The two most common methods are the 'batch method' and the 'history-by-history method'.

7.III.i. Batch method of estimating statistical uncertainties

The simplest method for calculating statistical uncertainties consists of breaking the simulation into N batches, calculating the result independently for each batch and then determine the final result as the average of the batch results with an uncertainty related to the root-mean-square deviations of the batch results. Consider a calculation done in N separate batches. For any scored quantity X, the value X_i is determined separately for each batch i. The estimate of the uncertainty in the mean value of X is[n5]

$$s_{\overline{X}} = \sqrt{\frac{\sum_{i=1}^{N}(X_i - \overline{X})^2}{N(N-1)}}, \tag{7.11}$$

where N is the number of batches, X_i is the mean value of X in batch i and $\overline{X}$ is the mean value of X evaluated over all batches, i.e.:

$$\overline{X} = \frac{1}{N}\sum_{1}^{N} X_i. \tag{7.12}$$

Note that $s_{\overline{X}}$ is an estimate of the uncertainty on the mean value of the quantity of interest It is not an estimate of the width of the distribution of values. There can be situations in which it is important to be aware of the underlying width of the distribution of values, but the batch technique will not provide that since it is determining the distribution of the average values of each batch, and this will be narrower as the number of histories in each batch increases. See a related discussion in section **??** about stochastic and non-stochastic quantities.

The quantity $s_{\overline{X}}$ is an estimate of the 68% confidence limit on the value of $\overline{X}$, i.e., there is a 68% probability that the true mean value of the parameter X is in the interval $(\overline{X} - s_{\overline{X}}, \overline{X} + s_{\overline{X}})$. This is the so-called $k = 1$ uncertainty estimate and the $k = 2$ estimate doubles the uncertainty estimate and provides a 95% confidence interval. These confidence limits assume that X is normally distributed.

A problem with the batch method is that the uncertainty on the uncertainty is quite large, especially for N=10 batches. Also, it requires N extra memory locations for every parameter being estimated. If calculating the dose to each voxel in a phantom with $128 \times 128 \times 128$ voxels, 20,971,520 extra locations are required if using 10 batches.

7.III.ii. History by history method

Rewrite Eq. 7.11 using Eq. 7.12 so that for any scored quantity X,

$$s_{\overline{X}} = \sqrt{\frac{1}{N-1}\left(\frac{\sum_{i=1}^{N} X_i^2}{N} - \left(\frac{\sum_{i=1}^{N} X_i}{N}\right)^2\right)}, \tag{7.13}$$

where

- N is the number of statistically independent (primary) histories
- X_i is all the contributions to X during primary history i
- $s_{\overline{X}}$ is the uncertainty on $\overline{X}$, the mean value of X evaluated over all primary histories.

There are several things to note.

- large sample size, $N \Rightarrow$ small uncertainty on uncertainty
- grouping by primary histories ensures that any correlations between particles (e.g., from a phase space source) are accounted for
- scoring $\sum_{i=1}^{N} X_i$ and $\sum_{i=1}^{N} X_i^2$ on the fly $\Rightarrow$ 2 rather than N_{batch} locations per X
- Perceived problem: Updating $\sum_{i=1}^{N} X_i$ and $\sum_{i=1}^{N} X_i^2$ for each voxel at the end of each primary history can be computationally intensive.

This last point meant that for many years this technique was not used since in the example above of $128 \times 128 \times 128$ voxels this technique would require looping over 2.1 million voxels at the end of each history to update the usually small subset of voxels that had been affected in that history. However, in 2000 Salvat proposed a simple algorithm which overcame this problem[n6] whereby the individual X_i and X_i^2 values were only updated each time they were affected on a history that differed from the primary history number the previous time the quantity X_i was affected. In the meantime, a temporary storage location keeps track of the contributions to X_i and X_i^2 during the current primary history. At the end of the run, all of the temporary storage locations must be used to update the final values for each X_i.

To score correlated quantities, e.g., $\overline{E_\gamma} = E_\gamma/N_\gamma$, one needs, for the case of C= X/Y:

$$\frac{s_{\overline{C}}}{\overline{C}} = \sqrt{\left(\frac{s_{\overline{X}}}{\overline{X}}\right)^2 + \left(\frac{s_{\overline{Y}}}{\overline{Y}}\right)^2 - \frac{2\text{cov}(X,Y)}{(N-1)\left(\overline{X}\,\overline{Y}\right)}}. \tag{7.14}$$

where the covariance between X and Y is

$$\text{cov}(X,Y) = \frac{\sum_{i=1}^{N} X_i Y_i}{N} - \frac{\sum_{i=1}^{N} X_i \sum_{i=1}^{N} Y_i}{N^2}. \tag{7.15}$$

This means that $\sum_{i=1}^{N} X_i Y_i$ must also be scored throughout the run, but for situations in which X and Y are correlated, the covariance is greater than zero and hence the third term in Eq. 7.14 can substantially reduce the uncertainty compared to just using the first 2 terms which ignores any correlation (see further discussion in section 7.V.v.).

In Fig. 7.13 the smoothness of the history by history estimate of the uncertainty on the doses on the central axis of an 18 MeV $e-$ beam shows the significant improvement in the

uncertainty on the uncertainty by using this technique. This figure also shows that the uncertainty estimates are improved by using 40 batches rather than 10. However note that the average uncertainty on all the voxels in the beam for which the doses were greater than 50% is the same within 0.1%, i.e., all 3 methods were estimating the same uncertainties.

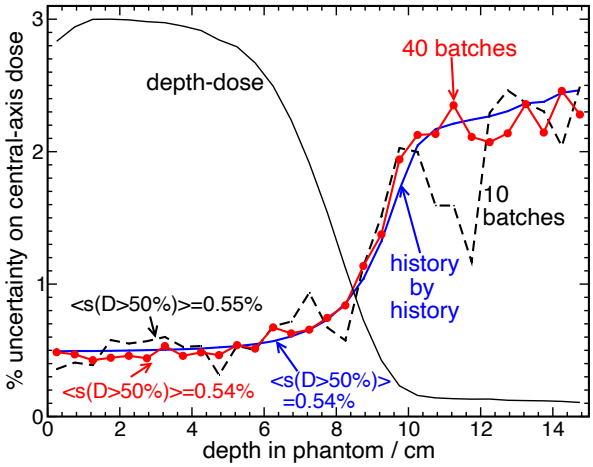

Figure 7.13: Three estimates of the uncertainties on the dose vs depth curve on the central axis of an 18 MeV 20× 20 cm^2 $e-$ beam simulated using BEAMnrc. The same phase-space file and number of histories (50x10^6) were used in all cases. Average uncertainty on doses >50% of D_{max} are shown. Based on data in ref[18].

7.III.iii. Variation of statistical uncertainty with number of histories

Generally the quantities being scored in Monte Carlo radiation transport problems are part of a Poisson process and the statistical uncertainty is directly related to the number of independent events that have been scored. For a sufficient number of histories, the distribution becomes normal and the estimated variance on the parameter of interest (i.e., the estimated square of the uncertainty) is inversely proportional to N, the number of independent histories, i.e.,

$$s^2_{\overline{X}} \propto 1/N \qquad \text{or} \qquad s_{\overline{X}} \propto 1/\sqrt{N}. \tag{7.16}$$

Fig. 7.14 shows that for an example case of dose in a phantom irradiated by a ^{60}Co beam, the relationships in Eq. 7.16 are verified. These relationships are very useful in practice since it means to reduce the uncertainty by a factor of 2, the number of histories must be increased by a factor of $2^2 = 4$. There are other ways to increase the number of events that are scored related to a parameter and these are discussed below in section 7.V.ii.a.

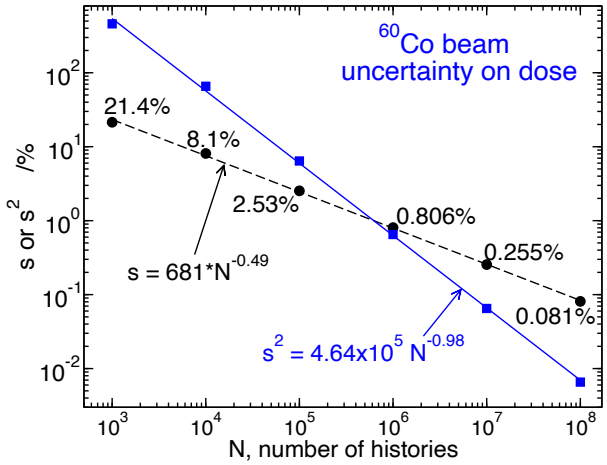

Figure 7.14: Estimates of the uncertainty, $s_{\overline{X}}$ or the variance, $s^2_{\overline{X}}$ on the central-axis dose between 1 to 2 cm depth in a 5 cm thick phantom irradiated by a 100 cm^2 ^{60}Co beam as a function of the number of histories, i.e., incident photons. The straight lines show least squares fits.

7.IV. Estimating systematic uncertainties

While the statistical uncertainty can often be the dominant uncertainty in a calculation, with enough computing power or clever variance reduction techniques it can be made arbitrarily small. But the systematic uncertainty can also play an important role in any calculation and must also be addressed.

7.IV.i. Uncertainty from code's algorithms

Systematic uncertainties can occur from a variety of causes. The first is related to the accuracy of the actual code itself. Have the various sampling algorithms and transport mechanics been implemented properly? This can be hard to establish ab initio but comparisons between different codes under well specified conditions can be useful. In some cases, situations can be set up so that there is a known result based on an independent method. The Fano test discussed later (see section 7.VI.) is an example where the dose in a cavity filled with a gas of the same material as the wall can be specified in terms of fundamental cross section data or energy conservation. This is a very difficult test to pass and has been used to improve the algorithms in a variety of general purpose codes.

7.IV.ii. Uncertainty from imperfect knowledge of geometry

A second source of uncertainty is the accuracy of the geometric model being used. What are the materials, their densities, their dimensions? These factors will have very different effects depending on the calculation at hand. For example when calculating something like an ion chamber's k_Q factor (see section **??**), the result is not very sensitive to the radius of the ion chamber being modelled because what is being calculated is the ratio of the chamber's reading in different beam qualities, and the mass of the gas actually cancels out. In contrast, if the absolute response of the ion chamber is being calculated, there would be a very strong

dependence on the radius. A 0.5 mm error in a 3 mm radius chamber would lead to a 30% error. Similarly, specifying the exact location of a TLD close to a brachytherapy source can lead to significant differences in the absolute dose to the TLD but only have a minor effect on the ratio of the doses to the TLD and the water at the same point, despite the strong $1/r^2$ variation in the doses. This class of uncertainties due to errors in the geometry can be a major issue when modelling accelerator heads or x-ray units where many of the details are proprietary. But reasonable estimates of the uncertainty on the final calculated values can be made by redoing the calculation varying one uncertain parameter at a time.

7.IV.iii. Uncertainty from cross sections

Another major source of Monte Carlo uncertainties is the uncertainties in the cross sections being used. This can vary considerably depending on the energies involved and the specific quantities being scored. For example, if 1% is taken as a conservative estimate on the uncertainty on the water photon cross sections in the radiotherapy energy range[19,20] then the uncertainty on the calculated dose at 5 cm depth in a water phantom in a ^{60}Co beam is about 0.85% but only 0.29% at 10 cm depth in an 18 MeV beam since the cross section and hence the attenuation is lower. But if what is calculated is the ratio of the dose to the water and the dose to the cavity of an ion chamber, the uncertainty in the water cross section has virtually no effect since both quantities are affected in a similar way.

At the other extreme, if the reduction in air kerma rate by a thick attenuator is being calculated, a small error in the cross section can lead to massive changes in the result. For example, a 3% change in the lead cross section for ^{137}Cs photons (this is the range of literature values) causes a 30% change in the air kerma rate behind a lead 1000 times attenuator.

7.IV.iv. Estimating uncertainties using error propagation

A general method for estimating the uncertainty due to the uncertainties in cross sections and geometric or other parameters is to use the general equation for error propagation, i.e.:

$$u_{f(\vec{X})} = \left[\sum_{i=1}^{n} \left(\frac{\partial f(\vec{X})}{\partial x_i} \right)^2 u^2(x_i) \right]^{\frac{1}{2}}, \tag{7.17}$$

where $u_{f(\vec{X})}$ is the uncertainty on $f(\vec{X})$, $\vec{X}$ represents the many variables which f depends on, the x_i are the individual variables/quantities that $f(\vec{X})$ depends on, e.g., each of the electron or photon cross sections or some physical dimension in the model, and $u^2(x_i)$ is the variance (uncertainty squared) on x_i e.g., 1% on the photon cross section. Make the approximation that:

$$\left(\frac{\partial f(\vec{X})}{\partial x_i} \right) = \frac{\Delta f(\vec{X})}{\Delta x_i}, \tag{7.18}$$

where $\Delta f(\vec{X})$ is the change in $f(\vec{X})$ when quantity x_i is changed by Δx_i. Taking Δx_i to be $u(x_i)$ gives:

$$u_{f(\vec{X})} = \left[\sum_{i=1}^{n} (\Delta f(\vec{X}))_i^2 \right]^{\frac{1}{2}}. \tag{7.19}$$

This can be a tedious process when there are many materials and dimensions involved, but provides a reasonable estimate of the overall uncertainty in the quantity being scored, *viz.*, $f(\vec{X})$. As will is discussed in section **??** this technique has been applied in detail to calculating the uncertainty on calculated k_Q values[21–23] taking into account uncertainties in electron and photon cross sections as well as chamber dimensions and uncertainty in the variation of $(W/e)_{air}$. However even these studies suffer from the inherent weakness of knowing the uncertainties on the cross sections and various assumptions must be made, e.g., that the photon cross sections are uniformly either high or low by some uncertainty, or that the theory behind calculating electron stopping powers is accurate and the only uncertainty is due to uncertainty in the mean ionization, I value or that uncertainties in photon cross sections are either correlated because they are based on the same theory or un-correlated. In these cases, comparisons to experimental data become important to establish the impact of any inaccuracies in the various cross section theories involved.

7.IV.v. Comparison to measured values

The ultimate test of the accuracy of Monte Carlo calculations is comparison the measured values, but this is not as simple as one would like. It is often very difficult to get sufficient information about the experimental setup to ensure an accurate Monte Carlo model can be made. Nonetheless, if there is agreement with experiment at the 10% *vs* 0.1% level, this gives an indication of the uncertainty of the calculations and the experiment at the same time. The ideal scenario is to have a high-quality measurement from a well specified experiment and a detailed uncertainty estimate. If this is then compared with a Monte Carlo calculated result which has a complete uncertainty analysis as well, then comparing the two sets of results and seeing if they agree within uncertainties can give confidence in the uncertainty estimate on the Monte Carlo results. Unfortunately, this is rarely the case, and/or the argument is circular since many experimental results depend on corrections calculated with Monte Carlo. Nonetheless, an example of such a comparison in the case of measured and calculated k_Q factors for many ion chambers is given later in section **??** and this confirms the methods described above provide reasonable uncertainty estimates for that application.

7.V. Improving the efficiency of a calculation

Monte Carlo calculations potentially take a large amount of computing time and hence one is interested in minimizing the time taken to achieve a given statistical precision. For a calculation which takes a computing time of T and yields s^2, an estimate of the variance on the quantity of interest, the efficiency, ϵ, is defined as

$$\epsilon = \frac{1}{s^2 T}. \tag{7.20}$$

This is a useful metric since, as shown in Eq. 7.16, $s^2 \propto 1/N$ where N is the number of particle histories and generally $T \propto N$ so that $s^2 T$ and ϵ are independent of N. The goal being to have ϵ as high as possible for the quantity of interest. Note that the absolute value of ϵ is meaningless but the relative values for 2 or more calculations can tell which is more efficient, i.e., will produce a given uncertainty in the shorter computing time. Note also that

the efficiency depends very much on the specific quantity of interest, so, e.g., a technique might greatly increase the efficiency of the dose calculated in the buildup region of a photon dose vs depth curve while at the same time making the efficiency of the calculated dose at 20 cm depth much worse.

Broadly speaking, there can be 2 classes of algorithms for improving the efficiency of calculations[1,24]. In one class, the approximate efficiency improving techniques (AEITs), one usually speeds up the calculation by making some approximations in the transport which both speed it up and, ideally, have no effect on the quantity of interest. The second class are true variance reduction techniques (VRTs) in which mathematical "tricks" are played which improve the statistical uncertainty on the quantity of interest (usually at the expense of some other quantity) but which still lead to an unbiased estimate of the quantity of interest. In practice, many techniques are a combination of both classes of algorithms. In the following a sampling of various techniques are described to give an indication of the possibilities, but the subject is deserving of an entire book.

7.V.i. Weighting

An important concept used in many variance reduction algorithms is that of a particle's weight. Usually a particle's weight starts as 1.0 but various VRTs adjust the weight in order to remain unbiased. In a simple example discussed below, when using particle splitting, a single particle may be replaced by N_{split} particles. To ensure an unbiased estimate of physical quantities, the weight is reduced by a factor of $1/N_{\text{split}}$. In all scoring routines, the quantity of interest must be "weighted" by the particle's weight, w. For example, if the energy deposited in a given region is being scored, rather than scoring the usual $E_{\text{dep}} = t(\text{d}E/\text{d}\rho x)_{\text{el}}$, on each step, $E_{\text{dep}}.w = E_{\text{dep}}/N_{\text{split}}$ is scored.

7.V.ii. Approximate energy improving techniques (AEITs)

The efficiency of many Monte Carlo calculations can be improved by using "approximate energy improving techniques" or AEITs after it is established that the approximations involved do not affect the accuracy of the simulation.

7.V.ii.a Optimized geometry

In section 7.III.iii. it was shown that the estimate of the variance, s^2, is proportional to 1 over the number of independent histories, $1/N$. Increasing the number of events being scored can be done by means other than just increasing the total number of histories. For example, when scoring a dose vs depth curve, if the area of the scoring region perpendicular to the beam's direction is increased, the number of events scored will increase at the same rate as the area increases as long as the dose distribution is flat radially. Hence $s^2_{\overline{X}} \propto 1/\text{area}$ which means $s_{\overline{X}} \times r$ should be constant as a function of scoring region radius r. Fig. 7.15 demonstrates this is the case when scoring either kerma or dose in a 25 MeV uniform photon beam with a radius of 5.6 cm. This is an AEIT in the sense that the radius used must be small enough that the quantity being scored is uniform over the entire region otherwise it will not represent the quantity throughout the region, and in particular, not the central-axis values.

In Fig. 7.15 the uncertainty when scoring the kerma is much higher because the number of events scored depends only on the photon interactions within the scoring region whereas for the dose, there are events from the electrons generated by the many more photon interactions taking place throughout a distance corresponding to the buildup depth in this very high-energy beam.

Fig. 7.16 shows the product of $s_{\overline{X}} \times \sqrt{\text{thickness}}$ as a function of $\sqrt{\text{thickness}}$. For the kerma calculations, the number of events scored is proportional to the thickness of the scoring region and hence $s_{\overline{X}} \times \sqrt{\text{thickness}}$ is constant, i.e., the uncertainty goes down as the thickness increases. But in the case of the dose, the uncertainty depends very weakly on the thickness (it decreases but only slightly) and thus the product increases quite dramatically.

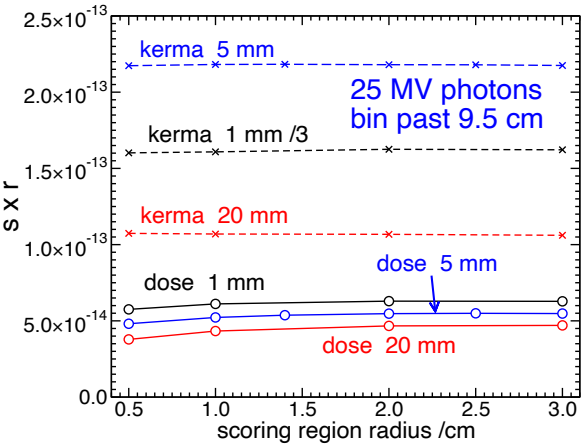

Figure 7.15: Product of $s_{\overline{X}}$ and scoring region radius for a parallel 25 MeV photon beam of radius 5.6 cm. Dose and kerma are scored on the central axis at depths starting at 9.5 cm in cylinders of varying thicknesses and radii. Data for kerma in 1 mm thick regions are divided by 3 to fit on same scale.

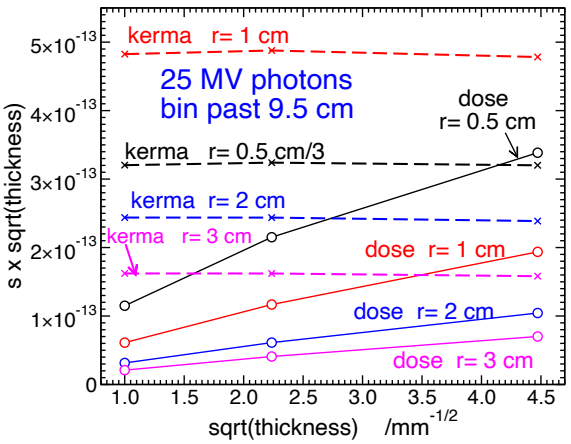

Figure 7.16: Product of $s_{\overline{X}}$ and the square root of the thickness of the scoring region for both the kerma and the dose in the same beam as in Fig. 7.15.

7.V.ii.b **Reciprocity theorem**

A useful application of symmetry makes use of the reciprocity theorem[25]. If the dose or other quantity is scored in a small region around the central axis of a normally incident, uniform, parallel beam with a much larger radius, then a lot of computing time is spent tracking particles which never reach the region near the central axis. The reciprocity theorem relates the dose delivered by a beam with a large radius to a narrow central-axis region to the dose delivered to a large radial region by a beam with the small radius. The advantage is that all histories contribute to the quantity being scored and in practice the central-axis dose can be scored for many different beam radii in the same simulation[n7] using[2]

$$D(r_b, 0) = \frac{E_{\text{dep}}(0, r_b)}{Nt}, \tag{7.21}$$

where $D(r_b, 0)$ is the central-axis dose per incident fluence from a parallel uniform beam of radius r_b incident on a slab phantom, $E_{\text{dep}}(0, r_b)$ is the energy deposited in a cylinder of radius r_b when irradiated by a normally incident pencil beam of N particles, and t is the thickness (in g/cm^2) of the slab of interest.

While the above is expressed in terms of dose, the theorem can be applied to other quantities such as fluence or stopping-power ratios. But note that this technique only applies to normally incident parallel beams on slab phantoms which are effectively semi-infinite laterally. Also, it is in a section about AEITs whereas it could be considered a true VRT in the sense that it provides an unbiased result, but the approximation is the restrictions on the geometry.

As an example of the power of this technique, consider 6 MeV electrons incident on a water phantom where the quantity of interest is the dose in a 1 mm thick region of radius 5 mm at 1.45 cm depth when beams of radii corresponding to 2x2 (1.128), 10x10(5.64) and 20x20 (11.28) cm^{2}are incident. The calculations using the reciprocity theorem with a beam of 5 mm radius and scoring in regions to 1.128, 5.64 or 11.28 cm are 10.1, 293 and 1,154 times more efficient than doing the calculations with incident beams of the different radii and scoring in the central 5 mm region. In addition, the reciprocity based calculation obtained the results for all radii in one run as opposed to needing 3 separate runs in the straight forward approach.

7.V.ii.c Selection of appropriate energy cutoffs

Considerable computing time can be saved by the appropriate choice of the electron threshold energy for creating electrons(AE in EGSnrc), and the cutoff for transporting electrons(ECUT). In photon beams past $D_{\max}$, the kerma and absorbed dose are very similar to each other away from interfaces (see Fig. **??**). This means completely ignoring electron transport does not change the dose a great deal so there can be considerable latitude in selecting ECUT and AE in photon beams, depending on the situation. As seen in Fig. 7.3 for electrons on a water phantom, there is virtually no change in a 20 MeV electron dose vs depth curve for energy cutoffs of 500 keV kinetic energy or below whereas for 100 keV electrons there is considerable variation.

Table 7.1 demonstrates the changes in efficiency associated with the selection of energy thresholds (AE and ECUT values) in a 6 MV photon beam. The first observation is that the dose distributions determined with any of the parameters in Table 7.1 (and below in Table 7.2) are identical within statistics of 0.5%. Increasing the AE/ECUT values from 0.512 to 0.700 MeV increases the efficiency by a factor of 9 for (2 mm)^3 voxels and a factor of 18 for (5 mm)^3. Going from the smaller to larger voxels increase the efficiency by a factor of more than 10 in most cases, of which a factor of about 2 comes from increased speed per history while most of it comes from the lower uncertainty because of the larger voxels (see section 7.V.ii.a). In varying the AE/ECUT values, most of the gains come from the ECUT variation although the value of AE can make a factor of 2 difference.

In summary, selection of the lowest energies to create and/or transport electrons can make a large difference in calculation times and in some situations, as in this example, have no effect on the results. However there are situations in which thresholds have a marked effect on the calculated result and it is important to establish what the appropriate thresholds are.

Table 7.1: Variation in relative efficiency as a function of the threshold energy for creating secondary charged particles (AE, total energy) and as a function of the cutoff energy below which electrons are not transported (ECUT, total energy) when calculating dose delivered by a 100 cm^2 6 MV photon beam incident on a water phantom with either $(2\ mm)^3$ or $(5\ mm)^3$ voxels. Efficiency is relative to the average statistical uncertainty on all voxels with $D > 0.5D_{max}$ and overall is normalized to the (512,512) $(2\ mm)^3$ case and separately, in brackets, to the (512,512) $(5\ mm)^3$ case. No range rejection is used (see Table 7.2 for range rejection effects). Calculations with DOSXYZnrc.

		ϵ	
AE/MeV	ECUT/MeV	$(5\ mm)^3$	$(2\ mm)^3$
0.512	0.512	8.04 (1.00)	1.00
0.512	0.521	38.6 (4.80)	3.54
0.512	0.700	50.5 (6.28)	4.78
0.521	0.521	47.8 (5.95)	4.28
0.521	0.700	127 (15.8)	7.88
0.700	0.700	148 (18.4)	9.21

7.V.ii.d **Range rejection**

Range rejection is a technique which can save a very substantial amount of computing time. Basically range rejection terminates a charged particle's history and its energy is deposited locally if the particle cannot escape from the current region, i.e., if its range is less than the distance to the nearest boundary. A conservative estimate of the rejection range is

$$R_{rej} = \int_{E_{min}}^{E} \frac{dE'}{(L/\rho)_{tot}(E', \Delta)}, \qquad [g/cm^2] \tag{7.22}$$

where E is the particle's current kinetic energy (in MeV) and $(L/\rho)_{tot}(E', \Delta)$ is the total mass restricted stopping power (in MeV/(g/cm^2)) with a threshold of Δ (AE). This range is considerably larger than the CSDA range as seen in Fig. 7.17 where, for 1 (10) MeV electrons in water the rejection range for a 1 keV production threshold is 42% (63%) longer than R_{csda} and for a 10 keV threshold is 21% (40%) longer. The corresponding ratios at 1 (10) MeV for Ta are 80%(154%) and 39%(106%) for 1 and 10 keV thresholds respectively. These big differences are a result of pathlength energy straggling as shown in Fig. **??** where electron pathlengths are seen to often go much further than the CSDA ranges but the actual pathlengths experienced do not reach the theoretical limits given by the rejection ranges in Fig. 7.17. Also, the pathlength is much longer than the average net physical distance travelled and hence the rejection ranges of Eq. 7.22 are conservative estimates. For energies near thresholds the CSDA range is actually larger than the rejection ranges because the CSDA range takes into account the range to a complete stop whereas the rejection ranges are only to the threshold energies, which is what is needed in the calculations although may cause errors if the very low-energy electrons play a role in a very small geometry.

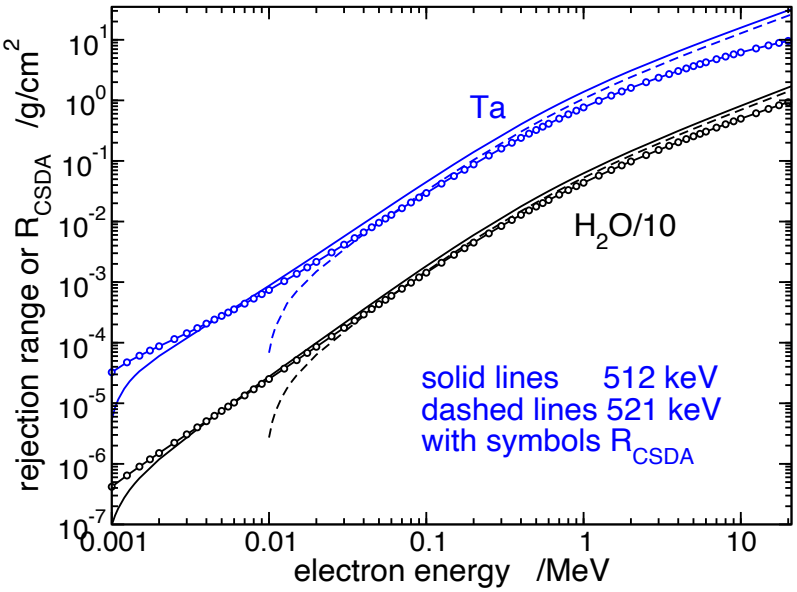

Figure 7.17: Values of the rejection ranges for water and tantalum as given by Eq. 7.22 compared to the ICRU Report 37 values of the CSDA range, $R_{\rm csda}$, given by Eq. ??. Values are given for 2 thresholds for secondary particle creation, 1 or 10 keV kinetic energy. Values for water are divided by 10 for clarity.

Table 7.2 shows that range rejection can provide substantial improvements in a calculation's efficiency, up to a factor of 4 when using AE=ECUT=0.512 MeV and progressively less for higher values of AE and ECUT. There is slightly more gain for the larger voxels for a given value of AE/ECUT. Range rejection is more effective the lower the values of AE/ECUT since, as seen from Table 7.1, it takes about 4 times as long to track electrons down to 0.512 MeV rather than to 0.521 MeV and these low-energy electrons have a short range so that range rejection terminates transport of most low-energy electrons. There is not much variation in efficiency as the upper limit for applying range rejection is increased past 1 MeV (0.489 MeV kinetic energy) since not a great amount of time is spent tracking these higher-energy electrons.

Table 7.2: Variation in relative efficiency for the same calculation as in Table 7.1 but as a function of the threshold energy for creating secondary charged particles (AE, total energy) and as a function of the upper limit total energy below which range rejection is applied. An upper limit of 0 means no range rejection is applied. Values with no bracket normalized to no range rejection (2 mm)^3 case and values in brackets to no range rejection value for that AE value.

AE=ECUT/MeV	Upper limit for range rejection/MeV	ϵ	
		(5 mm)^3	(2 mm)^3
0.512	0	8.04 (1.00)	1.00
0.512	1.00	32.7 (4.07)	2.93
0.512	5.00	31.2 (3.88)	3.00
0.521	0	47.8 (1.00)	4.28(1.00)
0.521	1.00	86.0 (1.80)	5.84(1.36)
0.521	5.00	87.8 (1.84)	5.77(1.35)
0.700	0	148 (1.00)	9.21(1.00)
0.700	1.00	148 (1.00)	9.21(1.00)
0.700	5.00	158 (1.07)	9.20(1.00)

Approximation involved in range rejection Range rejection as described above is not a true variance reduction technique because terminating an electron's history eliminates the possibility of that electron producing a brems photon which escapes the current region and thus can deposit energy elsewhere. The probability of this happening can be controlled by restricting range rejection to occurr only below an energy threshold so that the higher-energy electrons, which will produce the majority of the brems are not subject to range rejection.

Another approach, which makes range rejection into a true variance reduction technique is to play Russian Roulette about whether to do the range rejection when the electron is initially eligible (see section 7.V.iii.a) and when it is not done, increase the weight of any brems which may be given off, possibly splitting them.

Generally speaking range rejection is appropriate when dose is the quantity of interest. However, if one is scoring the energy spectrum of electrons in various regions, range rejection is not appropriate since the lower-energy electrons would be missing close to any boundaries.

7.V.ii.e Region of interest (ROI) range rejection

In the previous section, range rejection terminated any history that could not escape the current region. With the exception of the potential for brems issue, application of that technique leads to appropriate energy deposition in all regions in a geometry. But there are situations where the only quantity of interest is in one region of the geometry, e.g., determining the dose to an ion chamber's cavity in a large water tank. In such cases it is useful employ a range rejection based on the distance to the ROI. The rejection range is either established based on the actual materials between the current location and the ROI or established based on the material in the geometry with the longest rejection ranges. This rejection range is used to decide whether to terminate the history. In many situations this can be very efficient (e.g., an ion chamber in a water tank) but in geometries containing relevant air or vacuum regions, the latter technique is useless.

7.V.iii. Variance reduction techniques (VRTs)

7.V.iii.a Russian Roulette and splitting

Russian Roulette Russian Roulette is a true variance reduction technique whereby one randomly terminates a specific particle with a given probability, say $1/N_{RR}$, and if the particle survives, its weight is increased by a factor of N_{RR}. This may be done, for example, in cases where the particle is heading away from the region of interest but by maintaining the high-weight survivors the simulation allows for the unexpected events. This can often require fine tuning since if high-weight particles actually do contribute to the quantity of interest, they bring a large statistical uncertainty. As an example, consider a simulation of an accelerator beam which includes many particles far from the centre of the beam where the dose is being calculated. One can play Russian Roulette with all particles outside a given distance from the central axis but then must study the overall efficiency of the calculation as a function of the radius selected since too small a radius will cause the calculation to be less efficient than not using Russian Roulette.

Splitting The opposite VRT from Russian Roulette is particle splitting. When particles are heading towards a region of interest, the number of particles is increased by a factor $N_{\rm split}$ and the weight decreased by a factor of $1/N_{\rm split}$. This can be effective when the quantity of interest is away from the main radiation field, or the region of interest is small, or when it is expensive to get the initial particle. An example of the latter situation is when an accelerator is being modelled and the quantity of interest is the dose distribution in a phantom. Splitting all photons as they enter the phantom can be very effective since it is computationally expensive to obtain the photon from the accelerator.

Systematic photon splitting Photon splitting can be made more efficient by causing the $N_{\rm split}$ photons to interact systematically throughout the phantom. To do this, using a single random number, ξ, the number of mean free paths to the interaction of the i-th photon is given by

$$\lambda_i = -\ln\left[1-\frac{\xi+i-1}{N_{\rm split}}\right] = -\ln\left[1-\frac{\xi}{N_{\rm split}}-\frac{i-1}{N_{\rm split}}\right], \tag{7.23}$$

rather than the usual $\lambda_i = -\ln(1-\xi)$ from Eqs. 7.10 and 7.41. Since these mean free paths are spread out equally within the log, it means they are spread out farther at depth in the phantom as illustrated in Fig. 7.18. When the photons interact, Russian Roulette is played with the secondary photons with a probability of $1/N_{\rm split}$ so that surviving photons have the initial photon's weight. All charged particles are tracked and have a weight of $1/N_{\rm split}$ times the original photon's weight. If these charged particles produce any photons (e.g., brems or annihilation), Russian Roulette is played with these photons, again with probability $1/N_{\rm split}$ so the surviving photons again have the initial particle's weight. In all cases when a photon survives, it undergoes splitting again.

Fig. 7.19 demonstrates that the efficiency gains can be significant. The gains vary considerably with the value of $N_{\rm split}$ and clearly too large a value is counter productive. The gains are also very dependent on the geometry being modelled and although the no splitting ($N_{\rm split}=1$) efficiency is much higher for the larger 5 mm voxels than for the 2 mm voxels,

the corresponding gain in efficiency with splitting is much less. The overall efficiency gain would be considerably more in both cases if the incident photons were from an accelerator simulation and thus more costly than this example using a pre-calculated spectrum.

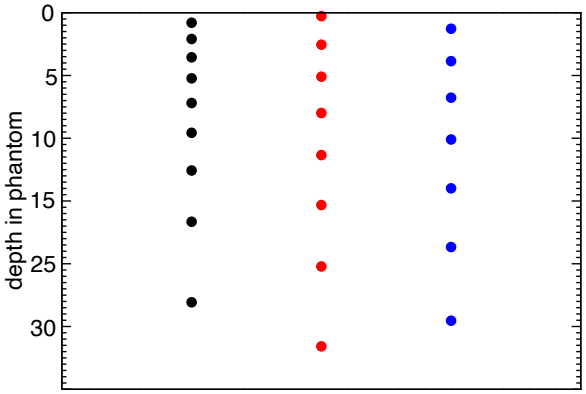

Figure 7.18: Depths of photon interactions for three 6 MV photons incident on a 30 cm thick water phantom with $N_{\rm split} = 10$. Locations depend both on the random number, ξ, sampled for each track, but also on the photon mean free path which depends on photon energy.

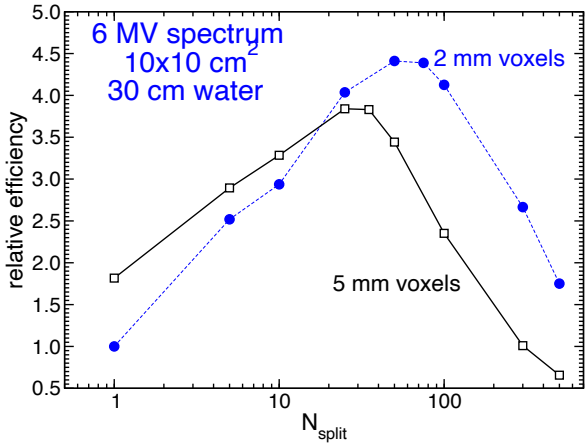

Figure 7.19: Efficiency as a function of $N_{\rm split}$ with systematic splitting for a 10×10 cm^2 6 MV photon beam at an SSD of 100 cm incident on a (30 cm)^3 water phantom with 2 or 5 mm voxels. Efficiencies are all relative to that for $N_{\rm split} = 1$ and 2 mm voxels. The value of s^2 used was the average in all voxels in the phantom with $D > 0.5 D_{\rm max}$.

Brems Splitting: Uniform, Selective or Directional Another example is the splitting of brems in a linac's target. The transport of the initial electron requires considerable computing time so when it is established that a brems event is to occur, efficiency can be increased by sampling the brems photon $N_{\rm split}$ times and reducing the weight of each brems photon by $1/N_{\rm split}$. This has the added advantage of actually sampling more of the brems spectrum for each event. This is referred to as Uniform Brems Splitting (UBS) to contrast it with Selective Brems Splitting (SBS)[26] in which the splitting number varies depending on an estimate, given the current direction of the electron, of how often the brems photons are likely to be going towards the phantom. SBS and UBS are often used in conjunction with Russian Roulette because a large number of the brems photons inevitably end up in the linac's primary collimator and rarely escape. An even more sophisticated technique is Directional Brem Splitting (DBS)[27] which is the standard in the BEAMnrc code and available in TOPAS[28]. This algorithm uses Kawrakow's clever techniques to only generate brems photons aimed at the field of interest and a wide variety of other "tricks", including Russian Roulette, to ensure an efficient simulation. Fig. 7.20 demonstrates the very significant efficiency gains to be made using these various techniques and the very large splitting numbers which can be useful. Kawrakow[29] developed an analytic theory which explains the efficiency dependence on scoring zone size and splitting number when using DBS or UBS.

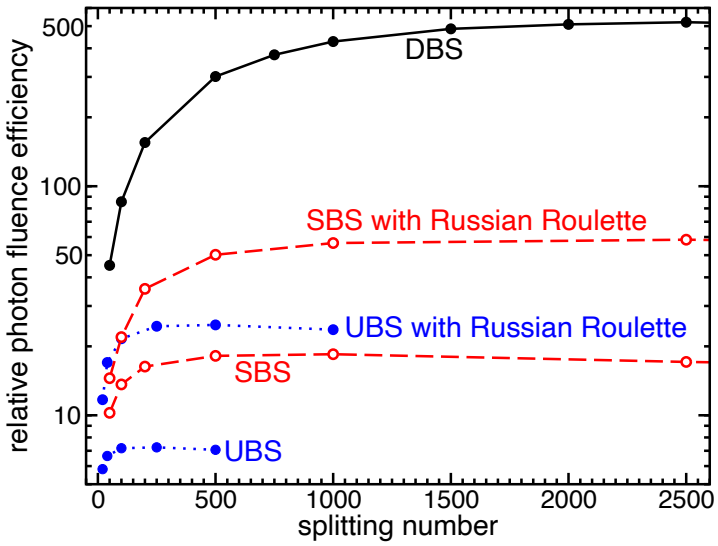

Figure 7.20: Relative efficiency for calculating photon fluence within a 10×10 cm^2 field of a 6 MV photon beam as a function of brems splitting number. Fluence uncertainty is summed for the 81 1×1 cm^2 regions centred within the 10×10 cm^2 field. Efficiencies shown are relative to total photon fluence efficiency with no splitting. For UBS and SBS, efficiencies are shown with Russian Roulette on and off. The field size parameter used with SBS was 30 cm[26]. DBS results have electron splitting off. Minimum splitting numbers are 20 (UBS) and 50 (SBS,DBS). Based on data from ref[27].

7.V.iii.b **Photon forcing and pathlength biasing**

Consider a simulation to calculate the dose in a thin slab of material irradiated by photons. The vast majority of the photons will pass through the slab without interacting. Photon forcing overcomes this problem. Assume the slab, or more generally, any geometry, is X mean free paths thick in the photon's direction. The number of photons escaping is $N_0 e^{-X}$ so the number of interactions is $N_0(1-e^{-X}) \approx N_0(1-(1-X)) = N_0 X$ for small X. Now, instead of selecting the number of mean free paths for the photon to travel as $\gamma_{\text{mfp}} = -\ln(1-\xi)$ where ξ is a random number between 0 and 1(as in Eq. 7.10), use

$$\gamma_{\text{mfp}} = -\ln(1-\xi(1-e^{-X})) \overset{X<<1}{\approx} -\ln(1-\xi X) \overset{X<<1}{\approx} \xi X, \tag{7.24}$$

which limits the number of mean free paths to between 0 and X and hence inside the geometry. But since the actual number of interactions would have been $1-e^{-X}$, the particle's weight is reduced by the factor $1-e^{-X} \overset{X<<1}{\approx} X$.

As an example, consider a slab of material that is 0.01 mean free paths thick (e.g., 0.062 cm of water irradiated by a 1.25 MeV photon beam). Only $1-e^{-0.01} = 1\%$ of the photons will normally interact in the slab of water. By using Eq. 7.24 every photon will be forced to interact with the resulting particles having their weight reduced by a factor of $1-e^{-0.01} = 0.0100$.

Another very useful technique with photon beams is pathlength biasing which can be used to preferentially cause the photon to interact earlier than would naturally occur, e.g., to study details in the buildup region of a dose vs depth curve. It can also be used to cause

the photon to penetrate further in the material which can be useful in shielding studies. In this technique, the photon's mean free path is selected as

$$\gamma_{\mathsf{mfp}} = -\frac{1}{1 - C\cos\theta}\ln\xi, \tag{7.25}$$

where C is a constant and θ is the angle between the photon's current direction and the direction in which bias is to be applied. For a normally incident photon, when the intention is to change the efficiency with respect to depth in the phantom, $\cos\theta = 1.0$. Hence

$$C < 0 \Rightarrow \gamma_{\mathsf{mfp}} \text{ decreases} \tag{7.26}$$
$$0 < C < 1 \Rightarrow \gamma_{\mathsf{mfp}} \text{ increases} \tag{7.27}$$

In these cases, to ensure an unbiased result the weight must be multiplied by

$$\frac{W'}{W} = \frac{1}{1 - C\cos\theta} e^{-\gamma_{\mathsf{mfp}}\ C\cos\theta}, \tag{7.28}$$

where the weight factor depends on the value of γ_{mfp} selected. The weight is reduced in regions where the biasing increases the number of photon interactions.

Table 7.3 shows how pathlength biasing can selectively increase the efficiency in the region of interest (near the surface for $C < 0$ or at great depth for $0 < C < 1$) while at the same time reducing the efficiency in the opposite region. As seen by the case of $C = 0.9$ in the 30 cm phantom, there can be excessive biasing which in this case causes the efficiency to decrease in both regions. The example of a 200 cm water phantom is, of course, unrealistic, but pathlength stretching ($0 < C < 1$) can be very useful in shielding calculations.

Table 7.3: Variation, when using pathlength biasing as a function of the biasing parameter C, of the relative efficiency in the calculated on-axis doses either near the surface or at depth in both a 30 cm or 200 cm thick water phantom irradiated by a 100 cm^2 beam of 6 MV photons.

	$\epsilon_{\text{relative}}$	
	depths in 30 cm phantom	
C	4 to 6 mm	28 to 28.5 cm
0	1.000	1.000
-3	2.65	0.044
-6	2.73	0.002
0.5	0.69	1.17
0.9	0.29	0.724
	depths in 200 cm phantom	
C	1 to 1.5 cm	194-196 cm
0	1.000	1.00
0.5	0.51	6.74
0.9	0.10	17.1

7.V.iv. Cross section enhancement

7.V.iv.a Brems cross section enhancement(BCSE)

In section ?? it was shown that a 100 keV electron slowing in a tungsten target would only give off a brems photon about 5% of the time. This makes simulating x-ray targets very inefficient as most of the time is spent tracking electrons which produce no photons. The brems cross section enhancement technique (BCSE) was introduced to overcome this problem[30]. The brems cross section is artificially increased by a factor, f_{enh}, thereby giving off f_{enh} times as many brems photons. Each photon must have its weight reduced by a factor of $1/f_{\text{enh}}$ to ensure an unbiased result. Furthermore, the energy of the electron cannot be decreased each time is creates a brems photon, so to make the simulation unbiased the energy is only decremented if a random number is sampled to be less than $1/f_{\text{enh}}$.[n8]

Using BCSE alone for a 130 kV x-ray tube, it was found that f_{enh} values of between 10^4 and 10^5 led to efficiency improvements by a factor of up to 1600[30]. When simulating the 3-D geometry around a 50 keV electronic brachytherapy source, the efficiency improvements were up to 10,000 for voxels up to (1 mm)^3 for f_{enh} values close to 5×10^4. If used in conjunction with uniform brems splitting, an additional 13% to 44% improvement in the efficiency is obtained although the f_{enh} and UBS's N_{split} values are much lower (e.g., 500 and 100 respectively). It has also been shown that combining BCSE with DBS improves the efficiency in linac simulations by 20% to 40% over using DBS alone[30].

7.V.iv.b Photon cross section enhancement (XCSE)

Similar to BCSE, considerable gains in efficiency can be made by artificially increasing the photon cross section by a factor f_{enh}. As implemented in the egs_chamber application[31] this enhancement can also be arranged to only occur in specific regions of interest, e.g., around an ion chamber in a phantom. When the photon interacts, it creates secondary particles, each with a weight of $1/f_{\text{enh}}$. By playing Russian Roulette on the scattered photons with a probability of $1/f_{\text{enh}}$, any surviving scattered photon will have the original photon's weight. As in the BCSE case, the initial photon is not affected by the interaction unless a selected random number is $< 1/f_{\text{enh}}$, in which case the original photon is not kept. In this algorithm the number of photons remains the same on average, all with their original weight and there are f_{enh} times as many primary electrons with weights $1/f_{\text{enh}}$.

The efficiency gains with this technique are a complicated function of the size of the region in which XCSE is applied and the value of the f_{enh} factor. Efficiency increases up to a factor of more than 100 (as high as 600 for ideal situations) are possible with f_{enh} values in the range of 30 to 500[31]. As a starting point, the XCSE region should surround the scoring region of interest by a distance equal to the depth of dose maximum for the photon component of the beam.

7.V.v. Correlated sampling

Correlated sampling can be a powerful method for increasing calculational efficiency when the quantity of interest is the ratio of, or difference between quantities due to differences in the geometries. The technique has been widely used for calculating detector correction

factors[31–33], e.g., when asking, what is the difference in an ion chamber's response in a phantom when its wall is made of graphite or water i.e., to calculate the P_{wall} correction factor(see section **??**).

The technique consists of defining 2 or more geometries where the only differences are in the materials in a finite part of the geometry, e.g., such as the solid region in Fig. 7.21. As a particle enters the region with differences, as in path I in the figure, the state of the particle (energy, direction, charge, weight *etc*) and the status of the random number generator are saved. The particle is transported through the remainder of its history using the geometry option 1, even after it leaves the region with differences. The history is then restarted at the boundary of the region with geometry differences using the saved particle and RNG status and transported in the second geometry. If a particle does not enter the region with differences, as in path II in the figure, the quantities of interest are scored for both geometry options.

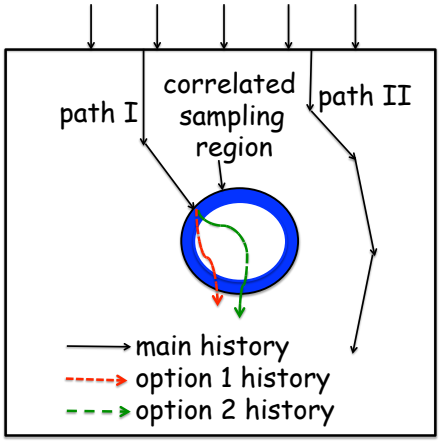

Figure 7.21: Schematic highlighting the correlated sampling method described in text.

Compared with running the simulation twice with the two overall geometries, this method saves time since the path II histories and the path I histories up to the point of entering the region with geometry differences, are only simulated once. There is some overhead with saving and restarting the histories as they enter the regions with differences.

But another major saving can come from the correlations between the two or more geometries. Recall Eq. 7.14 for the case of scoring a ratio, C= X/Y. The fractional uncertainty on the ratio of two quantities is

$$\frac{s_{\overline{C}}}{\overline{C}} = \sqrt{\left(\frac{s_{\overline{X}}}{\overline{X}}\right)^2 + \left(\frac{s_{\overline{Y}}}{\overline{Y}}\right)^2 - \frac{2\text{cov}(X,Y)}{(N-1)\left(\overline{X}\ \overline{Y}\right)}}. \tag{7.29}$$

where $\text{cov}(X,Y)$(Eq. 7.15) is the covariance between the scored quantities in geometry I or geometry II. The degree of correlation can be quantified by the correlation coefficient ρ_{xy} given by[n9]

$$\rho_{xy} = \frac{\text{cov}(X,Y)}{s_X s_Y}. \tag{7.30}$$

Then, using $s_{\overline{X}} = s_X/\sqrt{N}$,

$$\frac{s_{\overline{C}}}{\overline{C}} = \sqrt{\left(\frac{s_{\overline{X}}}{\overline{X}}\right)^2 + \left(\frac{s_{\overline{Y}}}{\overline{Y}}\right)^2 - 2\rho_{xy}\frac{N}{N-1}\left(\frac{s_{\overline{X}}}{\overline{X}}\right)\left(\frac{s_{\overline{Y}}}{\overline{Y}}\right)}. \tag{7.31}$$

Generally when using correlated sampling, $s_{\overline{X}}/\overline{X} \approx s_{\overline{Y}}/\overline{Y}$ so their product is comparable to the first 2 squared terms. For correlated quantities, $0 \leq \rho_{xy} \leq 1.0$ where $\rho_{xy} = 1$ means perfect correlation. For situations where the materials in the two geometries are reasonably similar (e.g., water and graphite), and since they start with the same random number seeds, the histories in the two geometries can often be very similar, i.e., correlated and hence the correlation term can greatly reduce the uncertainty in the ratio.

The overall gain in efficiency using correlated sampling can be thought of as[33]

$$G = \frac{k}{1+\alpha(k-1)}\frac{s^2_{\text{uncorr}}}{s^2_{\text{corr}}}, \tag{7.32}$$

where α is the fractional increase in calculation time for each geometry option beyond the first one when using correlated sampling and k is the number of geometry options being considered. The improvement in the ratio of uncertainties is dependent on how correlated the various geometries are. For example, with a 1.8 cm long, 5 mm diameter alanine pellet at 5 cm depth in water irradiated by a ^{60}Co beam, because alanine and water are similar in properties, the correlation coefficient is 0.97 and there is only a 1.8% increase in the computing time so the gain in efficiency is over a factor of 60 compared to doing 2 separate calculations with and without the pellet in place[33]. For a 2 mm long, 0.564 mm radius Al_2O_3 pellet at 30 cm depth in water, the correlation coefficient is only 0.75 but there is only a 0.4% increase in computing time. This leads to an efficiency gain of 8[33].

The correlated sampling technique can be very powerful under the right circumstances and has played a major role in calculating correction factors for ion chambers and other detectors. It is a core component of the widely used egs_chamber application[31].

7.V.vi. Tracklength scoring

For low-energy photon calculations such as those for brachytherapy, the quantity of interest can be the collision kerma since this is essentially equal to the dose when there is CPE(see section **??**). When the quantity of interest in a region is the collision kerma, it can be estimated using its relationship to the energy fluence, *viz.*, $D \approx K_c = \Psi\,(\mu_{\text{en}}/\rho) = E_\gamma \phi\,(\mu_{\text{en}}/\rho)$. Eq. **??** gives $\phi = \frac{1}{V}\sum_j t_j$ where the t_j are the photon tracklengths in region i. The tracklength estimator of the collision kerma in region i is

$$D_i \approx K_{c,i} = \frac{1}{V_i}\sum_j E_j t_j \left(\frac{\mu_{\text{en}}}{\rho}\right)_{\text{j}}, \tag{7.33}$$

where $(\mu_{\text{en}}/\rho)_{\text{j}}$ is the mass energy absorption coefficient at energy E_j for the medium in region i. This is in contrast to the analogue estimator of the dose or kerma given by

$$D_i = \frac{1}{m_i}\sum_j \Delta E_j, \tag{7.34}$$

where ΔE_j is the energy deposited in region i by electrons generated by photon interaction in the region or possibly nearby if electron transport is being simulated.

The advantage of tracklength scoring is that there is a contribution to the dose by every photon passing through the region of interest rather than waiting for a photon interaction in or near the region. In a series of simulations of clinic-like scenarios using the egs_brachy application (e.g., a prostate treatment with 100 ^{125}I seeds in voxels of $(1\text{ mm})^3$ or $(2\text{ mm})^3$), the efficiency improvement from using tracklength vs analogue (no electron transport) estimators varied by factors of 10 to 300[34].

This technique is based on the assumption that $D = K_c$. Fig. 7.22 shows the values of D/K as a function of depth in water for a 25 keV photon beam (relevant for LDR brachytherapy). At this energy $\overline{g}$ is 2×10^{-4} (see Fig. **??**) so D/K is the same as D/K_c. The figure demonstrates that once past a buildup region, $D = K$. But for regions very close to the surface there can be significant electron transport effects, as much as 40% in the first 0.01 mm, 4% averaged over the first 0.1 mm and only 0.4% averaged over a 1 mm bin, but in practice all of the effect is from the first 0.01 mm. The range of 25 keV electrons in water is 0.015 mm as seen in Fig. 7.17.

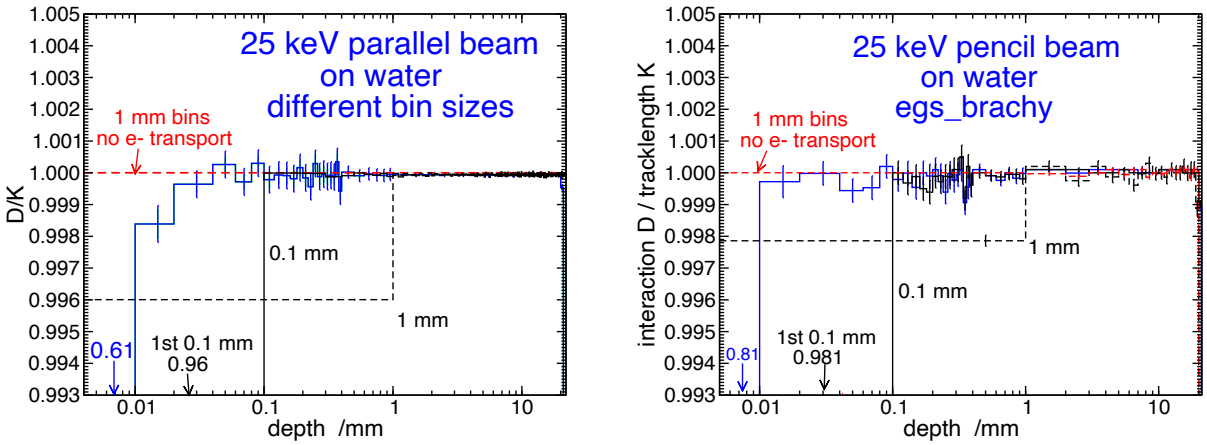

Figure 7.22: Monte Carlo calculated ratios of dose over kerma as a function of depth for 25 keV photons incident on water. Values are shown for differing depth-bin sizes near the surface. DOSRZnrc calculated the left panel and scored kerma directly whereas egs_brachy calculated the right panel and compares the dose from interaction scoring to the dose determined by tracklength scoring of the collision kerma. The surface bins are off-scale as shown. Differences in the surface bins are not understood.

7.V.vii. Exponential tracklength and next flight estimators

Williamson has introduced a variety of other techniques which can be useful for photon transport calculations when kerma is being sought[35,36]. These are mostly for calculating kerma at a series of points. The exponential tracklength estimator not only scores the tracklength when a photon passes through the volume of interest but takes into account the potential contribution of any photon directed at the volume. Kerma from photon i is

$$K_i = \frac{1}{m_i}\sum_j E_j \left(\frac{\mu_{en}}{\mu}\right) e^{-\mu s_j}\left(1 - e^{-\mu t_j}\right), \tag{7.35}$$

where, $E_j\left(\frac{\mu_{en}}{\mu}\right) = E_{tr}$(from Eq. **??**), $m_i = \rho V_i$, s_j is the distance from where the photon of energy E_j started heading towards the voxel of interest and t_j is the tracklength in the voxel. If the next interaction is in or past the voxel of interest, $s_j = 0 \Rightarrow e^{-\mu s_j} = 1$. For thin voxels, $\mu t_j << 1 \Rightarrow (1 - e^{-\mu t_j}) \approx \mu t_j$ so that for interactions in or past the voxel of interest, Eq. 7.35 reduces to the standard tracklength estimator, Eq. 7.33. The advantage is that even if the photon interacts before reaching the voxel, it contributes to the kerma, albeit weighted appropriately by $e^{-\mu s_j}$.

This approach is taken further using a next flight estimator[35]. Every time a photon interacts in the geometry, the probability is calculated of the photon scattering towards the point of interest(as opposed to voxel of interest) and the kerma calculated as

$$K = \frac{\sum_i E(\theta_i) P(\theta_i|E_{i-1}) e^{-\mu d_i} \left(\frac{\mu_{en}}{\rho}\right)}{d_i^2}, \tag{7.36}$$

where E_{i-1} is the photon's energy prior to scattering, θ_i is the angle of scattering required to go towards the point of interest, $E(\theta_i)$ is the energy of the photon scattered at that angle (determined from the kinematics), d_i is the distance to the point of interest, $(\mu_{\rm en}/\rho)$ the mass energy coefficient at energy E_i and $P(\theta_i|E_{i-1})$ is the probability per solid angle of a photon of energy E_{i-1} scattering to that direction given by

$$P(\theta_i|E_{i-1}) = \frac{\frac{d\sigma}{d\Omega}(\theta_i, E_{i-1})}{\sigma}. \tag{7.37}$$

A problem occurs when the interactions are close to the point of interest and the small d_i^2 term causes large fluctuations in the estimates. This can be overcome using a bounded next flight estimator[35] which, for any photon interaction inside a sphere of radius R about the point of interest, the $e^{\mu d_i}/d_i^2$ term in Eq. 7.36 is replaced with its mean value, *viz.*, $3(1 - e^{-\mu R})/(\mu R^3)$. This basically assumes the interactions within the sphere are uniform.

7.V.viii. Overview on improving efficiency

As mentioned at the start of this section (7.V.), an entire book could be devoted to describing efficiency improving techniques, and especially variance reduction techniques related to radiation transport Monte Carlo. The various techniques introduced in this section are meant to give an idea of the types of techniques that are available. For details about the many other techniques used, there are a wide variety of papers and reviews available[24,31,37–41]

7.VI. The Fano test

The Fano test, based on Fano's theorem (see section **??**), plays an important role in verifying the accuracy of a Monte Carlo code's modelling of electron transport. This is a very stringent test and any Monte Carlo code being used to calculate doses in geometries with significant density variations, especially ion chamber response, should be shown to pass this test. It is a test of the code's transport algorithms but not its cross sections.

There are several forms of the test and initially the commonly used version was applied to modelling ion chambers (some history in refs[16,42]). That version of the Fano test is

discussed later after a few more cavity-theory tools are developed (see section **??**). The test described here originated from a version[43] which starts directly with electron sources thereby avoiding issues about unattenuated and scatter-free photon beams which are needed if photons are the initial particles. The geometry considered must be effectively infinite and consist of the same medium everywhere with density which can vary. Despite the densities being different, the density effects used to calculate the stopping powers must be artificially the same to fulfill the Fano theorem's requirement that all media be the same. To establish charged particle equilibrium, electron sources emit I, a fixed number of electrons per mass throughout the geometry. For a simulation with N histories, $I = N/(\text{mass in simulation})$. Fano's theorem means that the electron fluence is identical everywhere and because it is all the same medium, this means the absorbed dose is the same everywhere. For monoenergetic electrons of energy E_0, conservation of energy means the dose everywhere in a simulation with N histories is

$$D = IE_0, \tag{7.38}$$

since the rhs is just the total energy per mass for a given region. This result does not require the electrons to be isotropic nor monoenergetic (although E_0 would be replaced by E_{ave} for a spectrum). However, if there is a magnetic field present, the Fano theorem no longer holds in general[44] and either the electrons must be emitted isotropically in a general magnetic field or the electrons are emitted in a spatially uniform manner and the magnetic field is in a constant direction with a strength proportional the local density[44].

The term "effectively infinite" is required to ensure CPE and in practice means that for regions where scoring is done, the geometry's boundaries must be farther than the maximum distance that a particle can travel. To minimize this distance, any photons created are not tracked and their energy deposited locally. The absolutely maximum distance a charged particle can travel is given by the rejection range, R_{rej}, defined above by Eq. 7.22. As seen in Fig. **??** , the actual pathlength while slowing down can be significantly greater than the CSDA range, R_{csda}, and calculated R_{rej} values (see Fig. 7.17) represent a conservative upper limit.

Given the above conditions, the Fano test can be applied with any geometry as long as the media in all regions are the same, albeit with differing densities. In fact, by filling any geometry of interest with a single medium of varying densities and assuming the accuracy of the code's transport algorithm, the Fano test essentially becomes a test on the geometry package, including the masses assigned to each region. However, the more common use has been to verify the accuracy of a given code and typically this has been done using a geometry with a low-density region between two high-density "wall" regions as seen in Fig. 7.23. [n10] is used This provides a severe test because of the big change in densities and because backscatter, which is very hard to model accurately, plays a significant role.

The most efficient way to run the test is by using the reciprocity theorem (section 7.V.ii.b) and consider a line source of initial electrons on the central axis with initial numbers being constant per mass and scoring the energy deposited in the gas to a radius that includes all particles. The reciprocity theorem means that this is equivalent to the dose on the central axis for sources everywhere in the cylinder out to radius R.

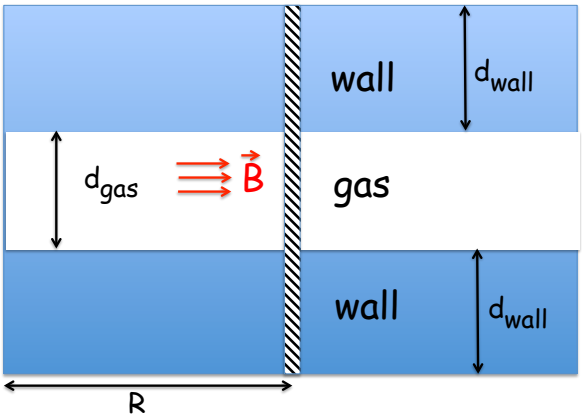

Figure 7.23: A typical cylindrical geometry for the Fano test in a magnetic field where the materials in the wall and gas regions only differ in their densities. By making use of the reciprocity theorem(section 7.V.ii.b) the electron sources can be constrained to an isotropic line source on the central axis with a constant number of initial electrons per mass throughout the cylinder. The radius of the cylinder, R, must be large enough that no particles initially on the central axis can reach the outer boundary, even in the gas. The wall thickness, d_{wall}, must be large enough that no particle from the gas can get through the wall. If a mag field is present, it can be in any direction and should be tested in several. If there is no mag field present, the initial electrons need not be isotropic.

7.VII. General purpose Monte Carlo codes

There are many general purpose Monte Carlo codes used in medical radiation physics. To properly describe them in detail would require another book, so the following will give a very brief introduction to 4 of the most widely used and cite references to their respective literatures (with apologies for my emphasis on the package I have been involved with for about 40 years).

One general observation is needed about low-energy simulations where quantum mechanical effects start to play a major role. Thomson and Kawrakow state[45]: "Classical trajectory MC simulations of low-energy electron transport do not model quantum effects which become non-negligible as electron energy decreases (wavelength increases)." and in their earlier paper[46]: "The classical Monte Carlo transport treatment is questionable for sub-1 keV electrons in condensed water as uncertainties on position and momentum must be large (relative to electron momentum and mean free path) to satisfy the quantum uncertainty principle." This does not necessarily invalidate the results of such low-energy simulations, but their accuracy can depend very much on what is being modelled.

7.VII.i. EGSnrc

The EGSnrc package[13,47–49] consists of a series of routines for simulating electron and photon transport in arbitrary geometries. It includes an efficient algorithm for simulating electron transport in magnetic and electric fields[50]. The user accesses these routines through a C++ or fortran application. As well as those mentioned in separate paragraphs below,

many complete applications are distributed as part of the package. For example, DOSRZnrc (dose in cylindrical geometry), SPRRZnrc (stopping-power ratios in cylindrical geometry), FLURZnrc (fluence in cylindrical geometry), EDKnrc (energy deposition kernels), g (mass energy absorption coefficients and radiative yield) and egs_fac (correction factors for free-air chambers). Simulations are controlled by user created ASCII input files (which can be created using GUIs for many applications). One of the main strengths of the EGSnrc system is that by default it runs very accurately and no adjustments are needed to, e.g., pass the Fano test, although by appropriate selection of some parameters the simulations can be made more efficient, e.g., as discussed in section 7.V.ii.c. The low-energy limit of transport is 1 keV.

The package started from the EGS4 Code System[51]. The core routines are written in a fortran preprocessor language called mortran3[52] as are some of the older applications, but most recent applications are C++ based and the interface can be entirely in C++. The package and documentation are available at https://github.com/nrc-cnrc/EGSnrc and a users community is on-line at https://www.reddit.com/r/EGSnrc/ or https://github.com/nrc-cnrc/EGSnrc/discussions.

BEAMnrc and DOSXYZnrc The BEAMnrc[53,54] application simulates external beam radiotherapy electron and photon sources with detailed models of the accelerator heads and applicators or ^{60}Co units involved. It includes a large number of very powerful variance reduction techniques which are optimized for this particular application. The DOSXYZnrc application[55] simulates dose distributions in a rectilinear phantom which can be generated from a CT dataset. It can be run with a full accelerator simulation as an input. Both codes and their auxiliary analysis applications come with GUI's for preparing the input files and packages to display the geometry or dose distributions. Both codes are mortran3/fortran based. The issue of selecting parameters for the incident electron beam can be a problem. Various solutions have been described. One early paper[56] matched central axis dose vs depth curves and off-axis ratios to determine parameters such as the beam size, mean energy and divergence. A more recent paper[57] cites the extensive literature on this important topic.

egs_chamber The egs_chamber[31] C++ application was designed to very efficiently simulate detailed models of ion chambers in phantoms irradiated by incident beams which could be from known spectra or using fully modelled BEAMnrc sources. It has been widely used for calculating the correction factors required for clinical dosimetry protocols(e.g., ref[58]).

egs_brachy The egs_brachy C++ application[34] is optimized for calculating dose distributions in patients being treated by brachytherapy. It can fully model electron transport or turn it off and apply tracklength scoring (section 7.V.vi.). Simulating typical brachy treatments with up to 100 seeds and an average 2% precision in (2 mm)^3 voxels can be done in 10s of seconds on a single cpu[34]. The distribution includes many vetted seed and phantom models[59,60]. Unlike the codes mentioned above, which are all part of the EGSnrc distribution, egs_brachy is available at https://github.com/clrp-code/egs_brachy and requires a separate installation of EGSnrc.

7.VII.ii. PENELOPE

The PENELOPE package[61–65] is a general purpose code for electron and photon transport. It has extensive cross section data especially at lower energies. It has been used for many

dosimetry related studies and with care can accurately simulate ion chamber response[43]. There are associated packages designed to simplify radiotherapy applications. The PRIMO package for Windows machines[66] pulls together various packages such as PENEASY and PENEASYLINAC[67] to facilitate the Monte Carlo simulation of linacs together with their electron applicators and multileaf collimators.

The PENELOPE package is available at https://www.oecd-nea.org/tools/abstract/detail/nea-1525/.

7.VII.iii. Geant4

The Geant4 package[68–70] is a general purpose system for transport of "all" particles which was originally developed for high-energy physics applications but has been extended to very low energies and attempts to simulate DNA damage[71]. The package is basically a set of routines which the user pulls together to handle their specific interest. There are several powerful packages for handling specific applications. For example the TOPAS package can model a passive scattering or scanning proton beam treatment head, calculate patient dose distributions based on CT images and work with phase space files[72]. The GATE package[73] was original released in 2004 for simulations involving PET and SPECT imaging[74] and in a major upgrade in 2006 was extended to cover CT and radiotherapy applications[75,76]. Similarly, Gamos provided a framework for using Geant4 with a user-friendly interface[77]. The ALGEBRA platform[78] is designed to handle brachytherapy dose calculations. It reports achieving 2% uncertainty in $(2\text{ mm})^3$ voxels in 13 min for ^{125}I prostate treatments. A recent review references many Geant4 applications in medicine[70].

G4-Med, a benchmarking and regression testing system for Geant4 applications in medical physics has been set up[79]. There are currently 18 tests described at https://twiki.cern.ch/twiki/bin/view/Geant4/G4MSBG.

Geant4 is available at https://geant4.web.cern.ch/. Geant4 has many different options related to a simulation and it requires careful tuning to ensure accuracy for difficult calculations such as ion chambers[80–82] or electron backscattering where often only the single-scattering mode appears to work although the Goudsmit-Saunderson multiple scattering model is often quite accurate[83]. There have been recommendations about which options and parameters to use for radiotherapy applications[84].

7.VII.iv. MCNP

The MCNP code system developed at LANL goes back a very long way. MCNP used to stand for Monte Carlo Neutron Photon but now stands for Monte Carlo N-Particle. The currently maintained version is MCNP6[85,86]. Its electron transport is a Class I algorithm (section 7.I.iv.) which can be traced back to the ETRAN system at NIST[87,88] via the ITS package at Sandia[89]. Earlier MCNP versions have been used extensively for TG-43 calculations for brachytherapy sources (see many references in e.g., ref[90]). For calculations of electron dose point kernels in water there are issues about selecting the correct transport parameters[91] and similar issues related to calculating electron backscattering coefficients[92].

7.VII.v. FLUKA

The FLUKA package was originally developed for particle physics applications[93] but is used for medical physics applications, especially related to ion beam therapy[94,95].

It is available via http://www.fluka.org/fluka.php.

7.VII.vi. Comparisons of Monte Carlo codes

The ultimate test of Monte Carlo codes is comparison to experiments but a lot can be learned from comparing various codes against each other since in that case the geometries and initial particles are well defined.

Lee et al.[96] used the FANO test with or without magnetic fields present to compare the electron transport in EGSnrc, MCNP6, PENELOPE and Geant4. Using the geometry in Fig. 7.23 they investigated initial electron energies of 10 keV, 100 keV, 1 MeV and 3 MeV and field strengths from 0 T to 3 T.

Using Malkov's magnetic field algorithm[50] and EGSnrc's default parameters, the test was passed at the 0.1% level for all cases.

After setting the 8 relevant step parameters(presumably in light of the recommendations in ref[43]), even with B=0, PENELOPE showed Fano discrepancies up to 0.9%, especially at 1 and 3 MeV. Running PENELOPE in single scattering mode there were still discrepancies greater than 0.2% and 0.5% at 1 and 3 MeV despite taking 4 to 19 times as long to compute. For B>0 T the discrepancies were between 1.2 and 2.1% at 3 T for energies above 100 keV or between 0.3 and 0.9% for mag fields up to 1.5 T. Using much shorter steps the discrepancies were reduced to 0.5% at 3 T.

Calculations with MCNP6.1 show discrepancies with the expected doses of up to 40% although, at the cost of computing time, using very short step sizes reduced the discrepancies to within 0.4%.

Using reasonable default parameters with Geant4 Lee *et al.* saw discrepancies from +8% to -30% but mostly in the range of 1-2% for energies other than 100 keV. Using the strict parameters recommended by O'Brien *et al.*.,[81] discrepancies were restricted to <2% up to 1.5 T but were -15% and -6% for 1 and 3 MeV electrons at 3 T.

Archambault and Mainegra-Hing[97] compared EGSnrc, MCNP5, PENELOPE and Geant4 using a very simple geometry. They scored the energy deposition in spheres of water with radii ranging from 2.5 mm to 4.5 cm irradiated by a point source of mono-energetic electrons 5 cm from the front of the spheres with either air or vacuum between the source and the sphere. Electron energies of 0.5, 1.0 and 5 MeV were studied. For calculations done in vacuum in single-scattering mode, at 500 keV all codes agreed within 0.1%. At 5 MeV in single scattering mode, Geant4 showed differences relative to EGSnrc and PENELOPE of up to 0.8% for the smallest spheres but agreed within 0.2% for radii >1.5 cm. When air was present, at 500 keV PENELOPE was 5% low compared to EGSnrc and Geant4 for the smallest radii but agreed within a fraction of a percent above a radius of 1.5 cm.

The more interesting comparisons are those using the normal condensed history (CH) algorithms. The results were compared to the same code's single-scattering results. In vac-

uum, EGSnrc and PENELOPE agreed with their own single-scattering results within 0.1%. Geant4 and MCNP5 showed discrepancies of 3% and 5% respectively for the smallest radius sphere. For other radii Geant4 was within 1% of its single-scattering results but MCNP5 varied by 2%, especially for the 5 MeV electrons. When air is added to the simulation EGSnrc's CH result differed from its single-scattering result for the smallest radius by 0.7% but agreed within 0.1% for R≥1 cm. Using standard parameter settings, Geant4, PENELOPE and MCNP5 were 200%, >100% and 600% high respectively for the smallest radius and the 500 keV electrons. PENELOPE was within 3% for radii ≥1 cm. Using optimized settings in-air, Geant4 and MCNP5 were within 6% and PENELOPE within 2% and better than 0.1% for R≥1 cm. EGSnrc did not require optimized settings.

This study additionally compared the efficiencies of the calculations. Using the default settings, Geant4 was as much as 3 times more efficient than EGSnrc and 6 times more efficient than PENELOPE. However the more relevant figure is for the optimized settings which gave the more accurate results. Here EGSnrc was between 100 (for the smallest R) and 10 times more efficient.

Archambault[98] showed similar results comparing EGSnrc, Geant4 and MCNP5 but for a high Z medium and looking at brems production as well.

7.VIII. Bin-size effects on calculated dose distributions

When calculating dose vs depth curves or any dose distribution with a Monte Carlo code, what is actually calculated is the dose averaged over a bin of finite depth in the dose vs depth curve case, Δz, i.e.:

$$D_{MC}(z_0) = \frac{1}{\Delta z}\int_{z_0-\frac{\Delta z}{2}}^{z_0+\frac{\Delta z}{2}} D(z)dz. \tag{7.39}$$

If $D(z)$ is a constant, this reduces to $D_{MC}(z_0) = D(z)$, the average is just equal to the constant value. However, if $D(z)$ is varying, Kawrakow[99] points out that:

$$D_{MC}(z_0) = D(z_0)\left[1. + \frac{D''(z_0)(\Delta z)^2}{24D(z_0)} + ... + O((\Delta z)^4)\right], \tag{7.40}$$

where $D_{MC}(z_0)$ is the average dose in the bin of width (Δz) centred a z_0, $D(z_0)$ is the dose at the mid-point of the bin, and $D''(z_0)$ is the second derivative at z_0 of the dose with respect to depth. This equation is derived using a Taylor expansion of the dose $(D(z) = D(z_0) + D'(z_0)\Delta z + \frac{1}{2!}D''(z_0)(\Delta z)^2 + ...)$ in eqn 7.39.

Eq. 7.40 means that for any dose vs depth curve where the rate of change is constant (i.e., $D''(z_0) = 0$), the dose at the middle of the bin is almost exactly equal to the average dose over the bin. However, in the build-up region the second derivative is not zero and so small bins must be used to ensure that this is the case.

Hunter[100] looked at pure photon beam dose vs depth curves and found that bins as small as 0.2 mm are sometimes required to get 0.1% accuracy in the buildup region of calculated dose vs depth curves for pure photon beams although the constraint is not so tight if electron contamination is included in the calculation because the buildup region is less dramatic (i.e.,

$D''(z_0)$ is smaller). A more critical situation is the region with a steep fall off around a brachytherapy source. Fig. 7.24 shows the error in assuming the dose averaged over a shell of a given thickness is equal to the dose at the mid-point of the shell when there is a perfect $1/r^2$ dose distribution.

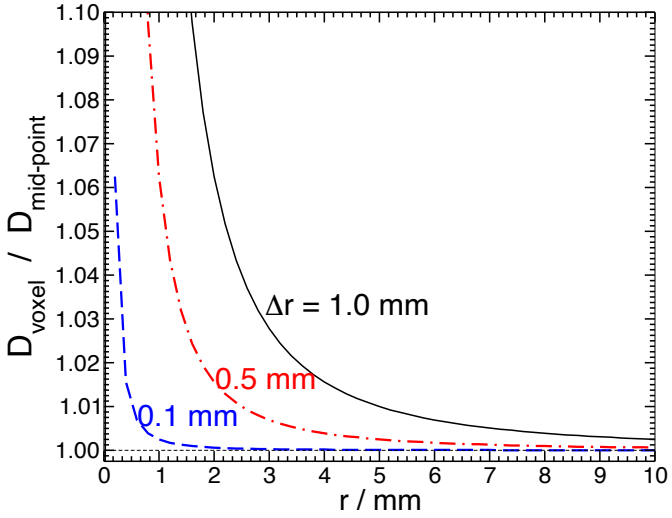

Figure 7.24: Ratio of the average dose in spherical shells of thickness Δr to the dose at the midpoint of the shell for a point source with a $1/r^2$ distribution for shell thicknesses of 0.1, 0.5 and 1 mm. Data from ref[101].

Notes

Chapter 7: Monte Carlo

n1 Recall that $\Sigma_{\text{tot}} = \mu$ (section **??**) and that the photon's mean free path, γ_{mfp}, equals $1/\mu = 1/\Sigma_{\text{tot}}$(Eq. **??**). This means that if the number of mean free paths to be travelled, N_{mfp}, is taken as

$$N_{\text{mfp}} = -\ln \xi_1 \tag{7.41}$$

then

$$N_{\text{mfp}} \times \gamma_{\text{mfp}} = -\ln \xi_1 \times \frac{1}{\mu} = -\frac{1}{\Sigma_{\text{tot}}} \ln \xi_1 = x \tag{7.42}$$

n2 The multiplying constant, Σ_{tot}, in the expression $p(x) = \Sigma_{\text{tot}} e^{-\Sigma_{\text{tot}} x}$ is required to normalize the PDF as required by Eq. 7.5:

$$\int_0^\infty f e^{-\Sigma_{\text{tot}} x} \mathrm{d}x = f \times \frac{1}{\Sigma_{\text{tot}}} = 1 \Rightarrow f = \Sigma_{\text{tot}}. \tag{7.43}$$

n3 The reported computing times of 77, 84, 144 and 1130 seconds in Fig. 7.7 were for calculations on a 1986 era computer and would be much less today but the relative time would be about the same.

n4 In the originally published version[14] of Fig. 7.9, there are comparisons to the generally less-accurate results for algorithms used in general purpose codes available in 1998.

n5 Note that the proper notation for an estimate of the square root of the variance of a parameter is $s_{\overline{X}}$ as opposed to $\sigma_{\overline{X}}$ which is the square root of the true value of variance of the parameter (which is an unknown).

n6 This algorithm is described in the appendix to a paper by Sempau *et al.*[102] which was co-authored by Salvat.

n7 The ability to score results for many radii at once is the reason that so many of this book's examples as a function of beam radius are calculated for parallel beams (e.g., Fig. **??** to Fig. **??**) .

n8 When using BCSE, there can be a large number of electrons generated and if the interest is in the photons, e.g., from an x-ray target, the it is useful to play Russian Roulette with any of the generated electrons.

n9 In ref[33] we mistakenly defined ρ_{xy} in terms of $s_{\overline{X}}$ and $s_{\overline{Y}}$ rather than s_X and s_Y. However, the CSnrc code actually calculated the covariance correctly. Thanks to Alexandra Kravchuk for helping me find this error.

n10 The geometry in Fig. 7.23 appears to be the most severe form of the test in the sense that if the geometry of a cylindrical ion chamber is used, the test can be passed for situations which fail the test using the slab geometry of Fig. 7.23[50].

References

[1] I. J. Chetty *et al.*, Report of the AAPM Task Group No. 105: Issues associated with clinical implementation of Monte Carlo-based photon and electron external beam treatment planning, Med. Phys. **34**, 4818 –4853 (2007). (pp 9 and 29)

[2] D. W. O. Rogers and A. F. Bielajew, Monte Carlo techniques of electron and photon transport for radiation dosimetry, in *The Dosimetry of Ionizing Radiation,Vol III*, edited by K. R. Kase, B. E. Bjärngard, and F. H. Attix, pages 427 – 539, Academic Press, 1990. (pp 9, 10, 13, 22, and 30)

[3] I. Kawrakow, E. Mainegra-Hing, D. W. O. Rogers, F. Tessier, and B. R. B. Walters, The EGSnrc Code System: Monte Carlo simulation of electron and photon transport, NRC Technical Report PIRS-701, National Research Council Canada, Ottawa, Canada. https://nrc-cnrc.github.io/EGSnrc/doc/pirs701-egsnrc.pdf, 2019. (p 9)

[4] P. L'Ecuyer, History of Uniform Random Number Generation, in *Proceedings of the 2017 Winter Simulation Conference*, edited by W.K.V. Chan *et al.*, pages 202 – 230, , 2017. (p 10)

[5] P. L'Ecuyer and R. Simard, TestU01: A C Library for Empirical Testing of Random Number Generators, ACM Trans. Math. Softw. **33**, 22:1 – 22:40 (2007). (p 10)

[6] G. Marsaglia and A. Zaman, A new class of random number generators, Annals of Applied Probability **1**, 462 – 480 (1991). (pp 10 and 11)

[7] F. James, RANLUX: A Fortran implementation of the high-quality pseudorandom number generator of Lüscher, Computer Phys. Commun. **79**, 111 – 114 (1994). (pp 10 and 11)

[8] M. Lüscher, A portable high-quality random number generator for lattice field theory simulations, Comp. Phys. Commun. **79**, 100 – 110 (1994). (p 10)

[9] I. M. Sobol', , USSR Computational Mathematics and Mathematical Physics **vol. 7, no. 4**, 86–112 (1967). (p 11)

[10] I. A. Antonov and V. M. Saleev, , USSR Computational Mathematics and Mathematical Physics **vol. 19, no. 1**, 252–256 (1979). (p 11)

[11] M. J. Berger, Monte Carlo Calculation of the penetration and diffusion of fast charged particles, in *Methods in Comput. Phys.*, edited by B. Alder, S. Fernbach, and M. Rotenberg, volume 1, pages 135 – 215, Academic, New York, 1963. (p 15)

[12] L. Landau, On the energy loss of fast particles by ionization, J. Phys. USSR **8**, 201 (1944). (p 15)

[13] I. Kawrakow, Accurate condensed history Monte Carlo simulation of electron transport. I. EGSnrc, the new EGS4 version, Med. Phys. **27**, 485 – 498 (2000). (pp 19, 20, and 45)

[14] I. Kawrakow and A. F. Bielajew, On the condensed history technique for electron transport, Nucl. Inst. Meth. B **142B**, 253 – 280 (1998). (pp 19, 20, and 51)

[15] A. F. Bielajew and D. W. O. Rogers, PRESTA: The Parameter Reduced Electron-Step Transport Algorithm for electron Monte Carlo transport, Nucl. Inst. Meth. B **18**, 165 – 181 (1987). (p 20)

[16] V. G. Smyth, Interface effects in the Monte Carlo simulation of electron tracks, Med. Phys. **13**, 196 – 200 (1986). (pp 20 and 43)

[17] B. J. Foote and V. G. Smyth, The modelling of electron multiple-scattering in EGS4/PRESTA and its effect on ionization-chamber response, Nucl. Inst. Meth. **B100**, 22 – 30 (1995). (p 20)

[18] B. R. B. Walters, I. Kawrakow, and D. W. O. Rogers, History by history statistical estimators in the BEAM code system, Med. Phys. **29**, 2745 – 2752 (2002). (p 25)

[19] E. S. M. Ali, B. Spencer, M. R. McEwen, and D. W. O. Rogers, Towards a quantitative, measurement-based estimate of the uncertainty in photon mass attenuation coefficients at radiation therapy energies, Phys. Med. Biol. **60**, 1641 – 1654 (2015). (p 27)

[20] J. H. Hubbell, Review of photon interaction cross section data in the medical and biological context, Phys. Med. Biol. **44**, R1 – R22 (1999). (p 27)

[21] J. Wulff, J. T. Heverhagen, K. Zink, and I. Kawrakow, Investigation of systematic uncertainties in Monte Carlo-calculated beam quality correction factors, Phys. Med. Biol. **55**, 4481 – 4493 (2010). (p 28)

[22] B. R. Muir and D. W. O. Rogers, Monte Carlo calculations of k_Q, the beam quality conversion factor, Med. Phys. **37**, 5939 – 5950 (2010).

[23] B. R. Muir, M. R. McEwen, and D. W. O. Rogers, Beam quality conversion factors for parallel-plate ionization chambers in MV photon beams, Med. Phys. **39**, 1618 – 1631 (2012). (p 28)

[24] D. Sheikh-Bagheri, I. Kawrakow, B. Walters, and D. W. O. Rogers, Monte Carlo simulations: efficiency improvement techniques and statistical considerations, in *Proc of AAPM 2006 Summer School: Integrating New Technologies into the Clinic: Monte Carlo and Image-Guided Radiation Therapy*, edited by B. H. Curran, J. E. Balter, and I. J. Chetty, pages 71 – 91, Med. Phys. Pub., Madison, WI, 2006. (pp 29 and 43)

[25] ICRU, Radiation Dosimetry: Electron beams with energies between 1 and 50 MeV, ICRU Report 35, ICRU, Washington D.C., 1984. (p 30)

[26] D. Sheikh-Bagheri, Monte Carlo study of photon beams from medical linear accelerators; optimization, benchmark and spectra., PhD. thesis, Carleton University, Ottawa (1999). (pp 36 and 37)

[27] I. Kawrakow, D. W. O. Rogers, and B. Walters, Large efficiency improvements in BEAMnrc using directional bremsstrahlung splitting, Med. Phys. **31**, 2883 – 2898 (2004). (pp 36 and 37)

[28] J. Ramos-Mendez, C. R. Ortiz, J. Schuemann, H. Paganetti, and B. Faddegon, TOPAS simulation of photoneutrons in radiotherapy: accuracy and speed with variance reduction, Phys. Med. Biol. **69**, 11NT01(9pp) (2024). (p 36)

[29] I. Kawrakow, On the efficiency of photon beam treatment head calculations, Med. Phys. **32**, 2320 – 2326 (2005). (p 36)

[30] E. S. M. Ali and D. W. O. Rogers, Efficiency improvements of x-ray simulations in EGSnrc user-codes using Bremsstrahlung Cross Section Enhancement (BCSE), Med. Phys. **34**, 2143 – 2154 (2007). (p 39)

[31] J. Wulff, K. Zink, and I. Kawrakow, Efficiency improvements for ion chamber calculations in high energy photon beams, Med. Phys. **35**, 1328– 1336 (2008). (pp 39, 40, 41, 43, and 46)

[32] C.-M. Ma and A. E. Nahum, Calculation of absorbed dose ratios using correlated Monte Carlo sampling, Med. Phys. **20**, 1189 – 1199 (1993).

[33] L. A. Buckley, I. Kawrakow, and D. W. O. Rogers, CSnrc: correlated sampling Monte Carlo calculations using EGSnrc, Med. Phys. **31**, 3425 – 3435 (2004). (pp 40, 41, and 51)

[34] M. Chamberland, R. E. P. Taylor, D. W. O. Rogers, and R. M. Thomson, egs_brachy: a versatile and fast Monte Carlo code for brachytherapy, Phys. Med. Biol. **61**, 8214 – 8231 (2016). (pp 42 and 46)

[35] J. F. Williamson, Monte Carlo evaluation of kerma at a point for photon transport problems, Med. Phys. **14**, 567 – 576 (1987). (pp 42 and 43)

[36] Z. Li, J. F. Williamson, and H. Perera, Monte Carlo calculation of kerma to a point in the vicinity of media interfaces, Phys. Med. Biol. **38**, 1825 – 1840 (1993). (p 42)

[37] I. Kawrakow and M. Fippel, Investigation of variance reduction techniques for Monte Carlo photon dose calculation using XVMC, Phys. Med. Biol. **45**, 2163 – 2184 (2000). (p 43)

[38] I. Kawrakow and B. R. B. Walters, Efficient photon beam dose calculations using DOSXYZnrc with BEAMnrc, Med. Phys. **33**, 3046 – 3056 (2006). (p 43)

[39] M. Fippel, Variance Reduction Techniques, in *Monte Carlo Techniques in Radiation Therapy*, edited by J. Seco and F. Verhaegen, pages 29 – 42, Taylor and Francis, Boca Raton, FL, 2013.

[40] M. Rodriguez, J. Sempau, and L. Brualla, A combined approach of variance-reduction techniques for the efficient Monte Carlo simulation of linacs, Phys. Med. Biol. **57**, 3013 – 3024 (2012). (p 43)

[41] B. R. B. Walters and I. Kawrakow, A "HOWFARLESS" option to increase efficiency of homogeneous phantom calculations with DOSXYZnrc, Med. Phys. **34**, 3794 – 3807 (2007). (p 43)

[42] D. W. O. Rogers, Fifty years of Monte Carlo simulations for medical physics, Phys. Med. Biol. **51**, R287 – R301 (2006). (p 43)

[43] J. Sempau and P. Andreo, Configuration of the electron transport algorithm of PENELOPE to simulate ion chambers, Phys. Med. Biol. **51**, 3533 – 3548 (2006). (pp 44, 47, and 48)

[44] H. Bouchard, J. de Pooter, A. Bielajew, and S. Duane, Reference dosimetry in the presence of magnetic fields: conditions to validate Monte Carlo simulations, Phys. Med. Biol. **60**, 6639 – 6654 (2015). (p 44)

[45] R. M. Thomson and I. Kawrakow, Quantum versus classical Monte Carlo simulation of low-energy electron transport in condensed amorphous media, Physica Medica **54**, 179 – 188 (2018). (p 45)

[46] R. M. Thomson and I. Kawrakow, On the Monte Carlo simulation of electron transport in the sub-1 keV energy range, Med. Phys. **38**, 4531 – 4534 (2011). (p 45)

[47] I. Kawrakow, E. Mainegra-Hing, D. W. O. Rogers, F. Tessier, and B. R. B. Walters, The EGSnrc code system: Monte Carlo simulation of electron and photon transport, Technical Report PIRS-701, National Research Council Canada, Ottawa, Canada. http://nrc-cnrc.github.io/EGSnrc/doc/pirs701-egsnrc.pdf, 2020. (p 45)

[48] I. Kawrakow, egspp: the EGSnrc C++ class library, Technical Report PIRS–899, National Research Council Canada, Ottawa, Canada, 2005.

[49] I. Kawrakow, E. Mainegra-Hing, F. Tessier, R. Townson, and B. R. B. Walters, The EGSnrc C++ class library, Technical Report PIRS–898(2020), National Research Council Canada, Ottawa, Canada. http://nrc-cnrc.github.io/EGSnrc/doc/pirs898/index.html, 2020. (p 45)

[50] V. N. Malkov and D. W. O. Rogers, Charged particle transport in magnetic fields in EGSnrc, Med. Phys. **43**, 4447 – 4458 (2016). (pp 45, 48, and 51)

[51] W. R. Nelson, H. Hirayama, and D. W. O. Rogers, The EGS4 code system, Report SLAC–265, Stanford Linear Accelerator Center, Stanford, California, 1985. (p 46)

[52] A. J. Cook, Mortran3 user's guide, SLAC Computa. Res. Group Tech. Memo. No. CGTM 209 (1983). (p 46)

[53] D. W. O. Rogers, B. A. Faddegon, G. X. Ding, C.-M. Ma, J. Wei, and T. R. Mackie, BEAM: A Monte Carlo code to simulate radiotherapy treatment units, Med. Phys. **22**, 503 – 524 (1995). (p 46)

[54] D. W. O. Rogers, B. Walters, and I. Kawrakow, BEAMnrc Users Manual, Technical Report PIRS-794, National Research Council Canada, Ottawa. http://nrc-cnrc.github.io/EGSnrc/doc/pirs794-dosxyznrc.pdf, 2020. (p 46)

[55] B. R. B. Walters, I. Kawrakow, and D. W. O. Rogers, DOSXYZnrc Users Manual, NRC Report PIRS 794revB , https://nrc–cnrc.github.io/EGSnrc/doc/pirs794–dosxyznrc.pdf (2017). (p 46)

[56] D. Sheikh-Bagheri and D. W. O. Rogers, Sensitivity of megavoltage photon beam Monte Carlo simulations to electron beam and other parameters, Med. Phys. **29**, 379 – 390 (2002). (p 46)

[57] P. Papaconstadopoulos, I. R. Levesque, S. Aldelaijan, K. O'Grady, S. Devic, and J. Seuntjens, Modeling the primary source intensity distribution: reconstruction and intercomparison of six Varian TrueBeam sources, Phys. Med. Biol. **64**, 135005 (12pp) (2019). (p 46)

[58] M. R. McEwen, L. A. DeWerd, G. S. Ibbott, D. S. Followill, D. W. O. Rogers, S. M. Seltzer, and J. P. Seuntjens, Addendum to the AAPM's TG–51 protocol for clinical reference dosimetry of high-energy photon beams, Med. Phys. **41**, 041501(20pp) (2014). (p 46)

[59] H. Safigholi, M. J. P. Chamberland, R. E. P. Taylor, C. H. Allen, M. P. Martinov, D. W. O. Rogers, and R. M. Thomson, Update of the CLRP TG-43 parameter database for low-energy brachytherapy sources, Med. Phys. **47**, 4656 – 4669 (2020). (p 46)

[60] H. Safigholi, M. J. P. Chamberland, R. E. P. Taylor, M. P. Martinov, D. W. O. Rogers, and R. M. Thomson, Update of the CLRP TG-43 parameter database for high-energy brachytherapy sources, Med. Phys. **50**, 1928 – 1941 (2023). (p 46)

[61] F. Salvat, *PENELOPE-2018: A Code System for Monte Carlo Simulation of Electron and Photon Transport*, OECD Nuclear Energy Agency, https://www.oecd-nea.org/tools/abstract/detail/nea-1525/, 2019. (p 46)

[62] J. Baro, J. Sempau, J. M. Fernandez-Varea, and F. Salvat, PENELOPE: an algorithm for Monte Carlo simulation of the penetration and energy loss of electrons and positrons in matter, Nucl. Inst. Meth. B **100**, 31 – 46 (1995).

[63] J. Sempau, E. Acosta, J. Baro, J. M. Fernandez-Varea, and F. Salvat, An algorithm for Monte Carlo simulation of coupled electron-photon transport, Nucl. Inst. Meth. B **132** (1997).

[64] J. Sempau, J. M. Fernandez-Varea, E. Acosta, and F. Salvat, Experimental benchmarks of the Monte Carlo code PENELOPE, Nucl. Inst. Meth. B **207**, 107 – 123 (2003).

[65] F. Salvat and J. M. Fernandez-Varea, Overview of physical interaction models for photon and electron transport used in Monte Carlo codes, Metrologia **46**, S112 – S138 (2009). (p 46)

[66] M. Rodriguez, J. Sempau, and L. Brualla, PRIMO: A graphical environment for the Monte Carlo simulation of Varian and Elekta linacs, Strahlentherapie und Onkologie **189**, 881 – 886 (2013). (p 47)

[67] J. Sempau, A. Badal, and L. Brualla, A PENELOPE-based system for the automated Monte Carlo simulation of clinacs and voxelized geometries—application to far-from-axis fields, Med. Phys. **38**, 5887 – 5895 (2011). (p 47)

[68] S. Agostinelli *et al.*, GEANT4 – a simulation toolkit, Nucl. Inst. Meth. A **506**, 250 – 303 (2003). (p 47)

[69] J. Allison *et al.*, Recent developments in Geant4, Nucl. Inst. Meth. A **835**, 186 – 225 (2016).

[70] S. Incerti, J. M. Brown, and S. Guatelli, Advances in Geant4 applications in medicine, Physica Medica **70**, 224 – 227 (2020). (p 47)

[71] S. Incerti *et al.*, Geant4-DNA example applications for track structure simulations in liquid water: A report from the Geant4-DNA Project, Med. Phys. **45**, e722 – e739 (2018). (p 47)

[72] J. Perl, J. Shin, J. Schümann, B. Faddegon, and H. Paganetti, TOPAS: An innovative proton Monte Carlo platform for research and clinical applications, Med. Phys. **39**, 6818 – 6837 (2012). (p 47)

[73] D. Sarrut *et al.*, The OpenGATE ecosystem for Monte Carlo simulation in medical physics, Phys. Med. Biol. **67**, 184001 (18pp) (2022). (p 47)

[74] S. Jan *et al.*, GATE: a simulation toolkit for PET and SPECT, Phys. Med. Biol. **49**, 4543–4561 (2004). (p 47)

[75] S Jan *et al.*, GATE V6: a major enhancement of the GATE simulation platform enabling modelling of CT and radiotherapy, Phys. Med. Biol. **56**, 881 – 901 (2011). (p 47)

[76] D. Sarrut *et al.*, A review of the use and potential of the GATE Monte Carlo simulation code for radiation therapy and dosimetry applications, Med. Phys. **41**, 064301 (14pp) (2014). (p 47)

[77] P. Arca *et al.*, Gamos: A framework to do Geant4 simulations in different physics fields with an user-friendly interface, Nucl. Inst. Meth. A **735**, 304 – 313 (2014). (p 47)

[78] H. Afsharpour, G. Landry, M. D'Amours, S. Enger, B. Reniers, E. Poon, J.-F. Carrier, F. Verhaegen, and L. Beaulieu, ALGEBRA: ALgorithm for the heterogeneous dosimetry based on GEANT4 for BRAchytherapy, Phys. Med. Biol. **57**, 3273 – 3280 (2012). (p 47)

[79] P. Arce et al., Report on G4-Med, a Geant4 benchmarking system for medical physics applications developed by the Geant4 Medical Simulation Benchmarking Group, Med. Phys. **48**, 19–56 (2021). (p 47)

[80] E. Simiele and L. DeWerd, On the accuracy and efficiency of condensed history transport in magnetic fields in GEANT4, Phys. Med. Biol. **63**, 235012(13pp) (2018). (p 47)

[81] D. J. O'Brien, D. A. Roberts, G. S. Ibbott, and G. O. Sawakuchi, Reference dosimetry in magnetic fields: formalism and ionization chamber correction factors, Med. Phys. **43**, 4915 – 4927 (2016). (p 48)

[82] K.-S. Baumann, L. Derksen, M. Witt, S. Adeberg, and K. Zink, The influence of different versions of FLUKA and GEANT4 on the calculation of response functions of ionization chambers in clinical proton beams, Phys. Med. Biol. **68**, 24NT01(11pp) (2023). (p 47)

[83] P. Dondero, A. Mantero, V. Ivanchencko, S. Lotti, T. Mineo, and V. Fioretti, Electron backscattering simulation in Geant4, Nucl. Inst. Meth. B **425**, 18 – 25 (2018). (p 47)

[84] P. Arce, J. I. Lagares, J. D. Azcona, and P.-B. Aguilar-Redondo, A proposal for a Geant4 physics list for radiotherapy optimized in physics performance and CPU time, Nucl. Inst. Meth. A **964**, 163755(9pp) (2020). (p 47)

[85] T. Goorley *et al.*, Initial MCNP6 Release Overview, Nucl. Technol. **180**, 298 – 315 (2012). (p 47)

[86] G. Hughes, Recent developments in low-energy electron/photon transport for MCNP6, Prog. Nucl. Sci. Technol. **4**, 454 – 458 (2014). (p 47)

[87] M. J. Berger and S. M. Seltzer, Calculation of energy and charge deposition and of the electron flux in a water medium bombarded with 20-MeV electrons, Ann. N. Y. Acad. of Sci. **161**, 8 – 23 (1969). (p 47)

[88] S. M. Seltzer, Electron-photon Monte Carlo calculations: the ETRAN code, Int'l J. Appl. Radiation and Isotopes **42**, 917 – 941 (1991). (p 47)

[89] J. A. Halbleib, R. P. Kensek, T. A. Mehlhorn, G. D. Valdez, S. M. Seltzer, and M. J. Berger, ITS Version 3.0: The Integrated TIGER Series of Coupled Electron/Photon Monte Carlo Transport Codes, Sandia report SAND91-1634 (1992). (p 47)

[90] M. J. Rivard *et al.*, Supplement 2 for the 2004 update of the AAPM Task Group No. 43 Report: Joint recommendations by the AAPM and GEC-ESTRO, Med. Phys. **44**, e297 – e338 (2017). (p 47)

[91] R. Antoni and L. Bourgois, Evaluation of the new electron-transport algorithm in MCNP6.1 for the simulation of dose point kernel in water, Nucl. Inst. Meth. B **412**, 102 – 108 (2017). (p 47)

[92] A. Poškus, Evaluation of computational models and cross sections used by MCNP6 for simulation of electron backscattering, Nucl. Inst. Meth. B **368**, 15 – 27 (2016). (p 47)

[93] A. Fasso, A. Ferrari, J. Ranft, and P. R. Sala, FLUKA: Present status and future developments, Proceedings of the 4th Conference on calorimetry in High Energy Physics **World Sci., Singapore**, 493 – 502 (1993). (p 48)

[94] T.T. Böhlen *et al.*, The FLUKA Code: Developments and Challenges for High Energy and Medical Applications, Nuclear Data Sheets **120**, 211 – 214 (2014). (p 48)

[95] G. Battistoni *et al.*, The FLUKA Code: An Accurate Simulation Tool for Particle Therapy, Front. in Oncol. **6**, 116(24pp) (2016). (p 48)

[96] J. Lee, J. Lee, D. Ryu, H. Lee, and S.-J. Ye, Fano cavity test for electron Monte Carlo transport algorithms in magnetic fields: comparison between EGSnrc, PENELOPE, MCNP6 and Geant4, Phys. Med. Biol. **63**, 195013(11pp) (2018). (p 48)

[97] J. P. Archambault and E. Mainegra-Hing, Comparison between EGSnrc, Geant4, MCNP5 and Penelope for mono-energetic electron beams, Phys. Med. Biol. **60**, 4951 – 4962 (2015). (p 48)

[98] J. P. Archambault, Monte Carlo calculations of electrons impinging on a copper target: A comparison of EGSnrc, Geant4 and MCNP5, Appl. Radiat. Isot. **132**, 129 – 134 (2018). (p 49)

[99] I. Kawrakow, On the effective point of measurement in megavoltage photon beams, Med. Phys. **33**, 1829 – 1839 (2006). (p 49)

[100] C. R. Hunter, Monte Carlo Simulation Of Depth Dose Curves And Humidity Correction Factors For The $^{90}Sr^{90}Y$ Beta Radiation Source, Carleton University Honours Project Report, Physics Dept , 1–59 (2007). (p 49)

[101] R. E. P. Taylor, G. Yegin, and D. W. O. Rogers, Benchmarking BrachyDose: voxel-based EGSnrc Monte Carlo calculations of TG–43 dosimetry parameters, Med. Phys. **34**, 445 – 457 (2007). (p 50)

[102] J. Sempau, A. Sánchez-Reyes, F. Salvat, H. O. ben Tahar, S. B. Jiang, and J. M. Fernández-Varea, Monte Carlo simulation of electron beams from an accelerator head using PENELOPE, Phys. Med. Biol. **46**, 1163 – 1186 (2001). (p 51)

Index