Physics in Medicine & Biology

IPEM Institute of Physics and Engineering in Medicine

PAPER

Comparison of Geant4 multiple Coulomb scattering models with theory for radiotherapy protons

To cite this article: Anastasia Makarova *et al* 2017 *Phys. Med. Biol.* **62** 5959

IOP Publishing | Institute of Physics and Engineering in Medicine Physics in Medicine & Biology

Phys. Med. Biol. **62** (2017) 5959–5974 https://doi.org/10.1088/1361-6560/aa6ce3

Comparison of Geant4 multiple Coulomb scattering models with theory for radiotherapy protons

Anastasia Makarova[1,3], **Bernard Gottschalk**[2] **and Wolfgang Sauerwein**[1]

[1] NCTeam, Department of Radiation Oncology, Medical Faculty, University of Duisburg-Essen, Essen 45122, Germany
[2] Harvard University Laboratory for Particle Physics and Cosmology, 18 Hammond St., Cambridge, MA 02138, United States of America
[3] Global Station for Quantum Medical Science and Engineering, Global Institution for Collaborative Research and Education (GI-CoRE), Hokkaido University, Sapporo 060-8648, Japan

E-mail: anastasia.makarova@uni-due.de

Received 26 October 2016, revised 14 March 2017
Accepted for publication 12 April 2017
Published 6 July 2017

Abstract

Usually, Monte Carlo models are validated against experimental data. However, models of multiple Coulomb scattering (MCS) in the Gaussian approximation are exceptional in that we have theories which are probably more accurate than the experiments which have, so far, been done to test them. In problems directly sensitive to the distribution of angles leaving the target, the relevant theory is the Molière/Fano/Hanson variant of Molière theory (Gottschalk *et al* 1993 *Nucl. Instrum. Methods Phys. Res.* B **74** 467–90). For transverse spreading of the beam in the target itself, the theory of Preston and Koehler (Gottschalk (2012 arXiv:1204.4470)) holds.

Therefore, in this paper we compare Geant4 simulations, using the Urban and Wentzel models of MCS, with theory rather than experiment, revealing trends which would otherwise be obscured by experimental scatter. For medium-energy (radiotherapy) protons, and low-Z (water-like) target materials, Wentzel appears to be better than Urban in simulating the distribution of outgoing angles. For beam spreading in the target itself, the two models are essentially equal.

Keywords: Geant4, proton, MCS, pencil beam

(Some figures may appear in colour only in the online journal)

1. Introduction

Monte Carlo (MC) simulations are the gold standard for dose calculations in proton radiotherapy. Geant4 is perhaps the most popular MC, particularly if we take into account packages based on it such as GATE (Jan *et al* 2004) and TOPAS (Perl *et al* 2012). It is therefore important that the physics models in Geant4 be validated. The present paper focuses on multiple Coulomb scattering (MCS) in the Gaussian approximation.

A 1993 paper Gottschalk *et al* (1993)[4] (hereinafter Go93) contains measurements of the rms projected Gaussian angle $(\theta_x)_{\text{rms}}$ at 158.6 MeV incident energy for 14 materials, with 115 material/thickness combinations in total. Go93 also summarizes the variations of Molière theory, covering low-Z targets, thick targets, compounds and mixtures, and the Gaussian approximation.

The CERN URL http://vnivanch.web.cern.ch/vnivanch/verification/verification/electromagnetic/MSCP/geant4-10-02-patch-01/ links to a site comprising 14 graphs showing, for six Geant4 models, the difference in $(\theta_x)_{\text{rms}}$ between Geant4 simulations and the Go93 experimental data. A summary graph shows χ^2/N for each material and Geant4 setup (our figure 7 is similar). The machinery to produce these graphs is described in a 2013 note by Schwarz (2013); the most recent graphs were evidently generated around 15FEB2016. We discuss this work at greater length below. Suffice it to say here that the Wentzel model agrees best with experiment, as we will also find. However, because of experimental scatter, the CERN graphs give little insight into the dependence of either the Wentzel or Urban models on target material and/or thickness.

Fortunately, in the special case of MCS we have the luxury of an accurate theory free of adjustable parameters. Comparison with measurements in Go93 on many different target materials and thicknesses shows the Moliére/Fano theory to be accurate to better than 1% on the average. It appears to break down only when the target is thicker by ≈ 97% of the mean proton range. (MCS theories for thick targets depend upon the relation between proton energy and depth. That relation breaks down for near-stopping targets because of range straggling.) In this paper we take advantage of that by provisionally assuming that the Molière/Fano/Hanson variant (Moliere (1948), Bethe (1953)[5], Gottschalk *et al* (1993)) appropriate to the Gaussian approximation (hereinafter 'Hanson') is ground truth insofar as MCS is concerned. We justify that assumption by comparing the Go93 experimental data to Hanson. We then compare, also with Hanson, Geant4 computations of MCS using the Wentzel and Urban models. Further analysis reveals the dependence of the Wentzel and Urban models on target material and thickness, trends which would otherwise be obscured by experimental error. As to speed, the Wentzel model takes some 15% more CPU time than the Urban model.

The Goudsmit–Saunderson (GS) model (G4GoudsmitSaundersonModel) is available in Geant4 only for positrons and electrons (Geant4 Collaboration 2016). If nonetheless we alter the Geant4 code and apply GS to protons, it exceeds Hanson theory by more than 60%. The Urban model is, in essence, GS altered to fit experimental data, and as such is even used for electrons and positrons in the default EM physics lists.

Go93 is a 'target/drift' experiment: the spread in projected angle $(\theta_x)_{\text{rms}}$, introduced by MCS in a target, is converted into a transverse spread x_{rms} by a drift region, approximated by

[4] We have discovered the following errors in Gottschalk *et al* (1993).: equation (2) should read

$$\Xi(\chi) = \frac{1}{\pi}\frac{\chi_c^2}{(\chi^2+\chi_a^2)^2}$$

and in table 1 the heading α should read α^2 and $\times 10^9$ under χ_c^2 should read $\times 10^6$

[5] Four entries in the second column (the Gaussian) of table II are slightly incorrect (A. Cormack, priv. comm.) but the error (corrected in our programs is at worst 1%)

a large air gap, in which the additional scattering is small. Given the effective thickness of the air gap, $(\theta_x)_{\mathrm{rms}}$ may be inferred from measured x_{rms}.

Another class of experiments might be termed 'beam spreading'. The transverse spreading x_{rms} of an incident pencil beam *in the target itself* is measured as a function of depth. Such experiments are also regarded as tests of MCS models (Grevillot *et al* 2010, Matysiak *et al* 2013) and also obey an accurate, experimentally tested theory (Preston and Koehler 1968, Gottschalk 2012) with no adjustable parameters. We include them for completeness, though our primary emphasis is on $(\theta_x)_{\mathrm{rms}}$.

A side issue that will arise is the dependence of computed $(\theta_x)_{\mathrm{rms}}$ (for thick targets) on the range-energy relation of protons in the target material. That relation (judging by differences between standard tables (Janni 1982b, Berger *et al* 1993)) is uncertain to 1–2%, affecting independently both the Hanson computation and the Geant4 simulations. We will show that the numerical effect on either is small.

2. Methods

2.1. Experiment

For completeness, we summarize the Go93 experiment. A well collimated 158.6 MeV proton beam was directed onto the target and transverse scans were taken with a small Si diode 100 cm distant from the upstream target face. (A microwave diode was the smallest dosimeter available at the time. Its over-response at low energies (Koehler 1967) did not matter because the protons leaving the target all have the same energy.) Each scan was fit with a Gaussian on a constant background to find x_{rms}. Target-out 'air' scans were taken and analyzed similarly and their x_{rms} was subtracted in quadrature to correct for beam size, scattering in air, and detector size. Finally, corrected x_{rms} was converted to $(\theta_x)_{\mathrm{rms}}$ using an effective drift length that took into account the effective scattering point in the target.

Go93 compared $(\theta_x)_{\mathrm{rms}}$ with Highland's formula, a parameterization of Molière/Bethe/Hanson theory (Highland 1975, 1979). We will, instead, use Molière/Fano/Hanson theory directly, which should be slightly better. In all, Go93 studied 14 target materials of potential interest in proton radiotherapy, spanning the periodic table. Target thicknesses ranged from very thin to somewhat greater than the mean proton range.

2.2. Theory

2.2.1. Target/drift experiments. In the Gaussian approximation the 2D distribution $f(\theta)$ of polar angle θ is given by

$$f(\theta)\,\theta\,\mathrm{d}\theta\,\mathrm{d}\phi = \frac{1}{2\pi\,\theta_0^2}\mathrm{e}^{-\frac{1}{2}\left(\frac{\theta}{\theta_0}\right)^2}\,\theta\,\mathrm{d}\theta\,\mathrm{d}\phi \tag{1}$$

where

$$\theta_0 = (\theta_x)_{\mathrm{rms}} = \theta_{\mathrm{rms}}/\sqrt{2} \tag{2}$$

Equation (1) is valid to $\theta \approx 2.5\,\theta_0$, where the Molière single scattering tail becomes appreciable (Gottschalk 2012). This Gaussian region contains about 96% of the protons and therefore dominates proton radiotherapy dose calculations.

As Hanson *et al* (1951) first observed, the best Gaussian fit to Molière theory is obtained, not by merely using the first (Gaussian) term in Molière's expansion of $f(\theta)$, but by letting

$$\theta_0 = \chi_c \sqrt{B - 1.2}/\sqrt{2} \tag{3}$$

where χ_c is Molière's characteristic single scattering angle and B is his reduced target thickness. For the Molière/Fano/Hanson computation of θ_0 we find these quantities using Molière's rather than Bethe's form of the theory (Z^2 rather than $Z(Z+1)$) and using the Fano correction for low-Z targets (see Go93). The appropriate formulas are embodied in Fortran program LOOKUP (Gottschalk 0000). (In a minor improvement, LOOKUP uses cubic spline interpolation, rather than a polynomial fit, to interpolate range-energy tables.) We used the default MIXED range-energy table, namely ICRU49 (Berger *et al* 1993) except for Nylon, Zn and brass which are Janni (1982b).

For thin targets (negligible energy loss) θ_0 depends only on the initial value of

$$pv = \frac{(T/mc^2)+2}{(T/mc^2)+1} T \tag{4}$$

where p, v, T and mc^2 are proton momentum, speed, kinetic energy and rest energy (Gottschalk 2012). (In the clinical regime $3 \leqslant T \leqslant 300$ MeV the fraction multiplying T ranges from 2 to 1.76 so pv is roughly twice the kinetic energy.)

For thick targets, the integrals in Molière theory depend on the relation of pv to depth in the target, and the range-energy relation comes into play. To estimate this effect we replaced MIXED, the range-energy table one would use nowadays, by Janni (1982a), the tables (now outdated) used by Go93. The largest change in θ_0, for near-stopping Pb, was 3.1%, and for most material/thickness combinations it was far smaller. We will perform an analogous test in the Geant4 simulations. Table 1 lists typical values of θ_0 (Hanson) for reference.

2.2.2. Beam spreading experiments. Unlike Molière theory, which is complicated, beam spreading in a homogeneous slab follows just two rules first derived and tested experimentally by Preston and Koehler (1968). They hold for protons or heavier ions stopping in any material at any incident energy. A modern derivation is given in Gottschalk (2012).

The first rule is that the rms transverse spread $\sigma_x(R)$ at end-of-range R is proportional to range. The constant of proportionality is

$$\frac{\sigma_x(R)}{R} = \frac{E_s\, z}{2\,(pv)_{\mathrm{R}/2}} \sqrt{\frac{R}{X_S}} \tag{5}$$

which despite appearances is very nearly independent of R. $E_s = 15.0$ MeV, z is the particle charge number, pv is evaluated at the T value corresponding to $R/2$, and X_S is the scattering length (Gottschalk 2010) of the material. In Lexan, for instance, $\sigma_x(R) = 0.0210\,R$. Values of $\sigma_x(R)/R$ and ρX_S for many other materials are given in Gottschalk (2012).

The second rule is that, at any lesser depth $z < R$,

$$\frac{\sigma_x(z)}{\sigma_x(R)} = \left[2\,(1-t)^2 \ln\left(\frac{1}{1-t}\right) + 3\,t^2 - 2\,t\right]^{1/2} \quad , \quad t \equiv z/R \tag{6}$$

which, with equation (5), completely describes beam spreading in a homogeneous slab. Equations (5) and (6) assume an ideal incident beam, so measurements of $\sigma_x(z)$ must be corrected for initial beam size, divergence and emittance. These can be significant in beams designed for pencil beam scanning, but will not concern us. We will simply assume an ideal beam.

Table 1. Selected results: experiment, theory and Geant4 simulations. Material, thickness and θ_0 (exptl) are from Go93. θ_0 (Hanson) is from Molière/Fano/Hanson theory (section 2.2). θ_0 (Urban) and θ_0 (Wentzel) are from Gaussian fits to Geant4 runs (sections 2.3.1–2.3.3). First and last fitted points and some intermediate ones are given.

Material	Thickness (g cm^{-2})	θ_0 (exptl) (mrad)	θ_0 (Hanson) (mrad)	θ_0 (Urban) (mrad)	θ_0 (Wentzel) (mrad)
Beryllium	0.0572	0.993	0.980	0.876	0.967
	1.820	6.394	6.596	6.096	6.331
	20.313	43.848	44.601	41.798	42.589
Polystyrene	0.347	3.346	3.289	2.980	3.237
	15.751	42.031	41.973	39.039	40.304
Carbon	0.316	3.084	3.172	2.911	3.139
	1.616	7.728	7.846	7.268	7.630
Lexan	0.094	1.762	1.651	1.480	1.643
	1.455	7.436	7.523	6.834	7.254
Nylon	0.093	1.727	1.653	1.479	1.659
	3.010	10.656	11.529	10.499	11.035
Lucite	0.366	3.558	3.544	3.194	3.498
	1.449	7.579	7.610	6.931	7.370
Teflon	0.055	1.626	1.353	1.244	1.378
	1.072	6.918	7.037	6.484	6.928
	19.908	64.003	64.274	61.234	62.358
Aluminum	0.216	3.534	3.587	3.314	3.613
	2.173	13.104	12.995	12.038	12.733
	21.245	87.103	81.996	78.369	80.453
Copper	0.045	2.204	2.102	1.995	2.193
	1.450	14.327	14.671	13.875	14.799
	24.250	118.561	117.658	114.698	115.435
Zinc	0.190	4.884	4.825	4.518	4.968
	0.379	7.131	7.096	6.667	7.240
Brass	1.342	14.120	14.394	13.655	14.451
	24.398	115.851	120.982	118.213	118.697
Tin	0.0875	4.113	3.730	3.586	3.945
	0.345	8.106	8.074	7.756	8.411
Lead	0.029	2.304	2.320	2.302	2.540
	0.907	16.093	16.585	16.309	17.179
	31.566	175.421	186.292	190.846	178.819
Uranium	3.630	36.942	37.688	37.961	37.905
	17.430	95.288	101.524	104.097	99.147

2.3. Geant4 setup

2.3.1. Physics. We used Geant4 10.02, the latest release at the time of writing. Electromagnetic physics was based on the `G4EmStandardPhysics_option4` physics constructor class providing the most accurate models available in standard and low energy categories. The physics constructor was modified to allow different models and different parameters of the MCS process to be activated for protons. Stopping power tables were limited to the range 0–200 MeV and the number of bins was increased to 50 per decade according to the recommendations of Grevillot *et al* (2010). Other parameters, such as the step limitation function for the stopping process and other physical processes, were left at their default values.

Two MSC models, `G4UrbanMscModel` and `G4WentzelVIModel` were tested with the default step limitation and lateral displacement parameters. Then, the impact of those parameters on the results was investigated. Special attention should be paid to the way the step limitation, lateral displacement options and other parameters of MCS and other electromagnetic processes are defined. Static class `G4EmParameters` was added recently to the Geant4 library specifically for this purpose. Its methods `SetMuHadLateralDisplacement()` and `SetMscMuHadStepLimitType()` control switching the lateral displacement and type of step limitation algorithm for the hadronic MCS process on or off. By default, these parameters are set to `false` and `fMinimal` respectively.

A stricter step limitation slightly underestimates, in general, the MCS angle for small thicknesses. That is beneficial only for the Wentzel model applied to high-Z materials. Otherwise, this parameter only increases the deviation by 2–3%. The increase in execution time is roughly 2%.

Using the lateral displacement option for hadrons seems to affect the result randomly by 1–3% for both models. It is preferable to use the default, with lateral displacement switched off for hadrons and on for lighter particles. This parameter seems to have no effect on execution time.

The `G4CoulombScattering` process was not tested. Though it provides accuracy comparable with solving the diffusion equation, it is far too slow and thus inapplicable in proton therapy calculations.

Readers interested in how these models are implemented in the code may consult the Geant4 physics manual (Geant4 Collaboration 2016).

2.3.2. Material properties. Table 2 lists properties of the fourteen materials in this study. Names beginning with G4 indicate Geant4 default compositions and I_{G4} (mean excitation energy) values. For the others, we used densities and compositions from Go93 and computed I_{G4} by the internal Geant4 procedure (Geant4 Collaboration 2016). The mean projected range R_{G4} corresponds to the depth at which most protons stop as found in an auxiliary simulation.

As noted earlier, for thick targets θ_0 depends on the range-energy relation. Generally, when reconciling an MC with (say) an experimental Bragg peak, one can fine-tune either the incident energy or I. Here, however, we *must* use I since θ_0 depends directly on the incident energy via equation (4) even for thin targets. To change I by a reasonable amount we can adjust it to reproduce, via Geant4, some well-known range-energy table other than the one that follows from Geant4 defaults. Somewhat arbitrarily, we choose Janni (1982a), the first comprehensive tables used in proton radiotherapy and the ones used in Go93.

Accordingly, for each simulation with I_{G4}, we performed a second with I_{adj} adjusted to yield a range R_{adj} closely matching the range R_{Janni} from Janni (1982a) as given in Go93. These quantities are also given in table 2 as is the percent difference between R_{adj} and R_{G4}. That reflects the difference between two plausible range-energy relations for an arbitrary assortment of fourteen materials.

2.3.3. Scoring and analysis. Unlike Schwarz (0000) we did *not* (even approximately) simulate the Go93 experiment. Instead, a point mono-energetic mono-directional 158.6 MeV proton source was placed in front of the material slab and 1 M protons were traced from the source to the last step in the slab. For maximum efficiency, the polar angle θ of particles emerging from the last step, weighted by $1/\theta$, was scored in an annulus of radius θ and bin width $d\theta$ using the `G4CsvAnalysisManager` class. That histogram was then fitted with a Gaussian (see equation (1)) to find θ_0.

Table 2. Material properties used in Geant4 runs. I_{G4}, R_{G4}: default Geant4 mean ionization potential and mass range; I_{adj}, R_{adj}: the same with I adjusted to match R_{Janni}, the mass range stated in Go93 from a polynomial fit to Janni (1982a); final column: deviation of R_{adj} from R_{G4}.

Material	Geant4 material or $g\ cm^{-3}$, frac. wt.			I_{G4} (eV)	R_{G4} ($g\ cm^{-2}$)	I_{adj} (eV)	R_{adj} ($g\ cm^{-2}$)	R_{Janni} ($g\ cm^{-2}$)	$R_{adj}/R_{G4} - 1$ (%)
Beryllium	G4_Be			63.7	21.333	60.4	21.099	21.108	−1.10
Polystyrene	G4_POLYSTYRENE			68.7	17.682	62.3	17.494	17.504	−1.06
Carbon	G4_C			81.0	19.513	74.3	19.278	19.270	−1.20
		C	0.741						
Lexan	1.20	O	0.185	68.4	17.790	65.5	17.666	17.667	−0.70
		H	0.074						
		C	0.549						
Nylon	1.13	O	0.244	64.8	17.250	62.1	17.190	17.195	−0.35
		N	0.107						
		H	0.100						
		C	0.600						
Lucite	1.20	O	0.320	68.5	17.614	67.1	17.594	17.584	−0.11
		H	0.081						
Teflon	G4_TEFLON			99.1	20.883	106.5	21.003	21.008	0.57
Aluminum	G4_Al			166.0	22.401	155.6	22.155	22.158	−1.10
Copper	G4_Cu			322.0	26.265	289.9	25.947	25.923	−1.21
Zinc	G4_Zn			330.0	26.235	312.8	26.007	25.985	−0.87
		Cu	0.615						
Brass	8.489	Zn	0.352	333.7	26.439	319.8	26.325	26.345	−0.43
		Pb	0.033						
Tin	G4_Sn			488.0	30.678	437.2	30.188	30.159	−1.60
Lead	G4_Pb			823.0	35.844	754.4	35.196	35.209	−1.81
Uranium	G4_U			890.0	37.148	834.4	36.788	36.776	−0.97

2.4. Graphs and trend analysis

Let the *percent deviation* of e.g. experiment from Hanson theory be defined as

$$D_{EH} \equiv 100 \times \left(\frac{\theta_0(E)}{\theta_0(H)} - 1 \right) \tag{7}$$

where, if $D > 0$, the quantity under test (E) is greater than ground truth (H). We plot D_{EH}, D_{UH} and D_{WH} (for experiment, Urban and Wenzel, respectively) to exactly the same scales to facilitate comparison. The abscissa (target mass thickness in g/cm^2) is logarithmic because of the known behavior of MCS with target thickness (see Go93).

At fixed energy, only the dependence of θ_0 on target material and target thickness remain to be explored. To summarize the compliance of experiment, Urban and Wentzel to Hanson theory, we fit the data in figures 1–3 with straight lines. We exclude near stopping targets ($>0.9\times$ mass range) where Molière theory fails because range straggling destroys the relation between pv and depth. MCS in this region is of minor importance in proton radiotherapy, the residual range being so small that the proton direction hardly matters.

Finally we plot the slope D' (thickness dependence, %/decade) and mean value $\langle D \rangle$ (material dependence, %) of the fitted lines for E, U and W (figures 4–6). Again, we use exactly the same scales to facilitate comparison.

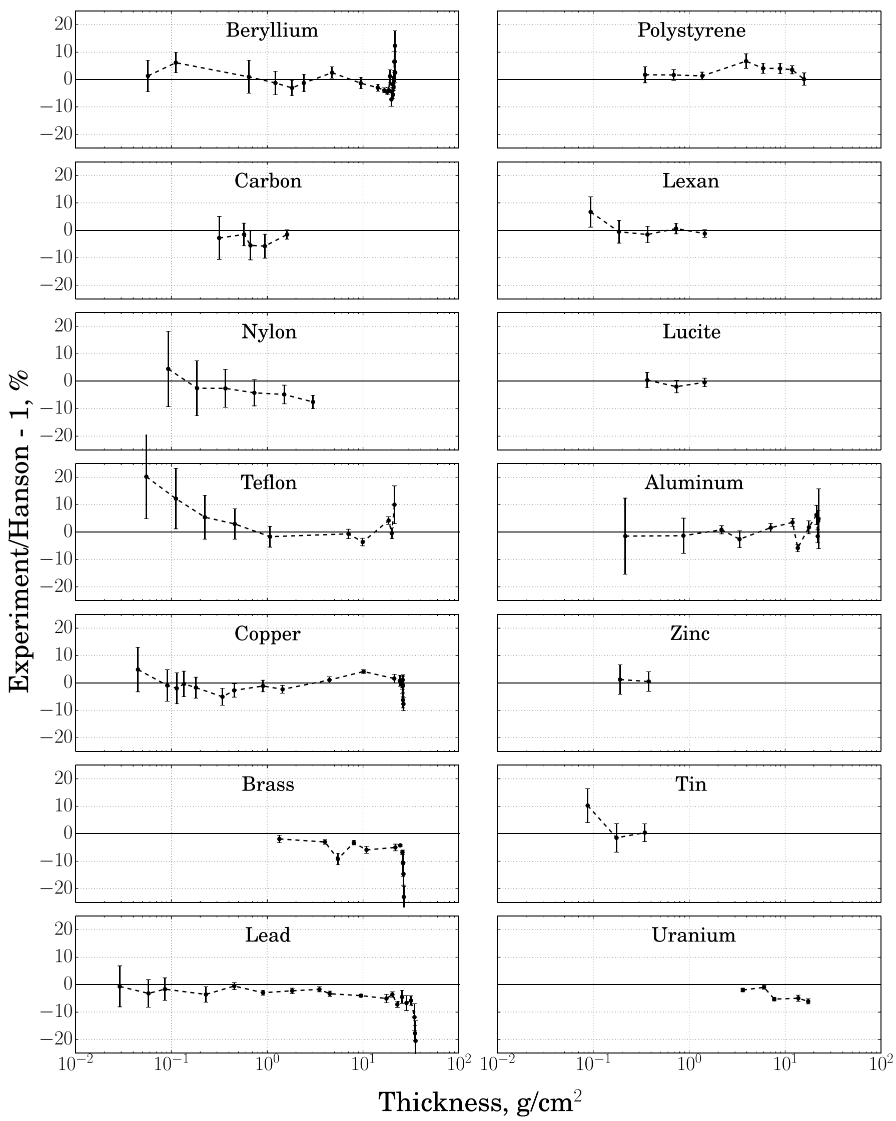

Figure 1. Comparison of experimental data from Go93 with 'Hanson', the Molière/Fano/Hanson theory computed using LOOKUP and the MIXED range-energy table.

3. Results

Table 1 gives, for selected points, the target material, g cm^{-2}, and measured θ_0 from Go93, the computed θ_0 (Hanson) from LOOKUP using the MIXED range-energy table, and finally θ_0 (Urban) and θ_0 (Wentzel) from the Geant4 simulations using the Geant4 default I values.

Percent deviations of θ_0 (exptl), θ_0 (Urban), and θ_0 (Wentzel) from θ_0 (Hanson) are shown in figures 1–3 respectively. 1σ experimental errors in figure 1 were taken from Go93.

Figure 1, taking into account the experimental error, shows that Hanson theory indeed describes the measurements with the possible exception of the thickest Brass, Pb and U points where theory may be some 4% high (or experiment 4% low) as already noted in Go93 for highest-Z materials.

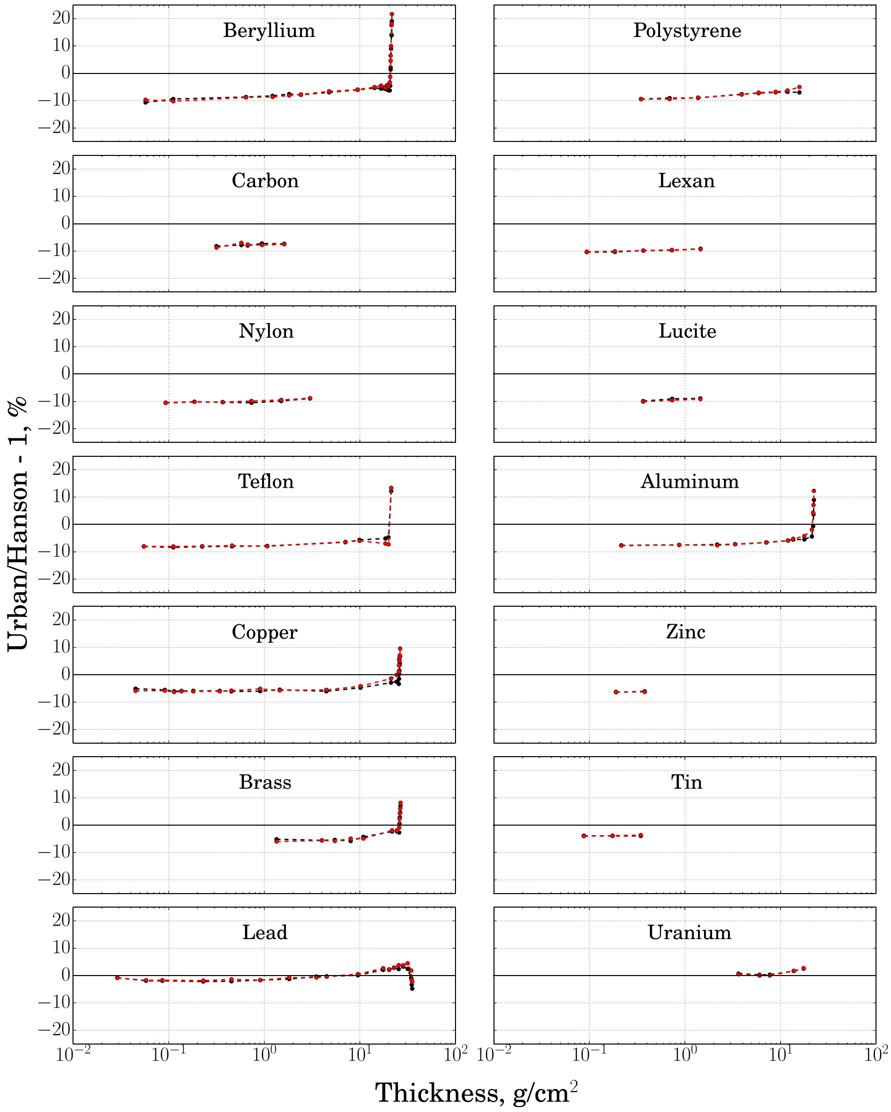

Figure 2. Comparison of the Geant4 Urban simulation with Hanson. Black points: simulation with Geant4 default values of I. Red (grey) points: simulation with I tuned to fit the Janni66 range-energy tables (Janni 1982a).

Figures 2 and 3 show the comparison of Geant4 simulation with Urban and Wentzel models with Hanson theory. Typical behavior for both models is a deviation from theory which is nearly linear in the logarithm of target thickness, has a small positive or negative slope, and some average offset from 0. MC statistical errors are small and are already implied by the non-smoothness of the lines with dots bigger than error bars.

Figure 4 quantifies the trends seen in figure 1. Teflon and Sn are obvious outliers, almost certainly due to experimental error given the much better agreement of neighboring materials. In particular Be, Al and Cu, for each of which a full range of thicknesses was measured, agree with theory very well on average.

Figures 5 and 6 quantify and summarize the trends seen in figures 2 and 3. Averaged over target thickness, the Urban model is ≈8% low for low-Z targets, smoothly approaching ≈0% for Pb and U. The variation with $\log_{10}$(target thickness) is roughly linear, with a slope of 1% − 2%/decade.

By contrast, the Wentzel model is ≈4% low for low-Z targets, agrees with theory at mid-range, and is ≈4% high for high-Z targets. Material dependence is noticeably less smooth than the Urban model. Thickness dependence is slightly greater than the Urban model and opposite in sign, say −2% to −4%/decade.

4. Discussion

Four other studies known to us test the Geant4 MCS model. Only one (Grevillot *et al* 2010) is published.

4.1. CERN web site

This site, already mentioned, explores six Geant4 configurations. The only documentation appears to be the note by Schwarz (0000) based on which there are two major differences with the present work.

First and foremost, Geant4 is compared with experimental measurements rather than theory.

Second, Schwarz (0000) describes a partial simulation of the Go93 experiment, unlike our method which merely scored θ of protons emerging from the target. In Schwarz (0000) an ideal beam enters the target and proton hits are scored on a finely divided measuring plane (to avoid detector size effects) 100 cm downstream of the target entrance face. The intervening gap is void (to avoid scattering in air). A Gaussian is fitted to what is effectively the transverse fluence (rather than dose). To convert its $x_{\rm rms}$ to θ_0 the effective scattering point is calculated according to Go93.

This procedure seems somewhat roundabout compared to simply scoring angles emerging from the target, but it seems to account for everything except incident beam size. Beam size may explain why Geant4 is consistently low for the thinnest Be targets on this Web site.

Figure 7, somewhat similar to a figure on the CERN site, summarizes the goodness-of-fit χ^2/N for the six configurations. Two Geant4 configurations are significantly worse than the four others, which are indistinguishable. Of those, opt4 + elastic corresponds most closely to our Urban and WVI + elastic to our Wentzel. The main point of figure 7 is that direct comparison with experiment, using χ^2/N as a figure of merit, is a poor way to evaluate the different models. It is too sensitive to how individual points and errors happen to fall out, and suggests material dependencies in Geant4 that cannot possibly be real.

4.2. Fuchs et al

This poster presentation (Fuchs *et al* 2015) tests numerous Geant4 releases. Simulated θ_0 is obtained directly, by scoring angles emerging from the target, or indirectly by back projecting dose profiles. The two methods agree. Those values of θ_0 are then compared with the Go93 measurement for every material/thickness combination. The '10.1 mod EM Wentzel VI' release is found to be best, with an average error of only $-1.2 \pm 3.3\%$, in substantial agreement with our figure 6. The range, −17.9%–11.2%, is of course much larger owing to comparison with experiment rather than theory.

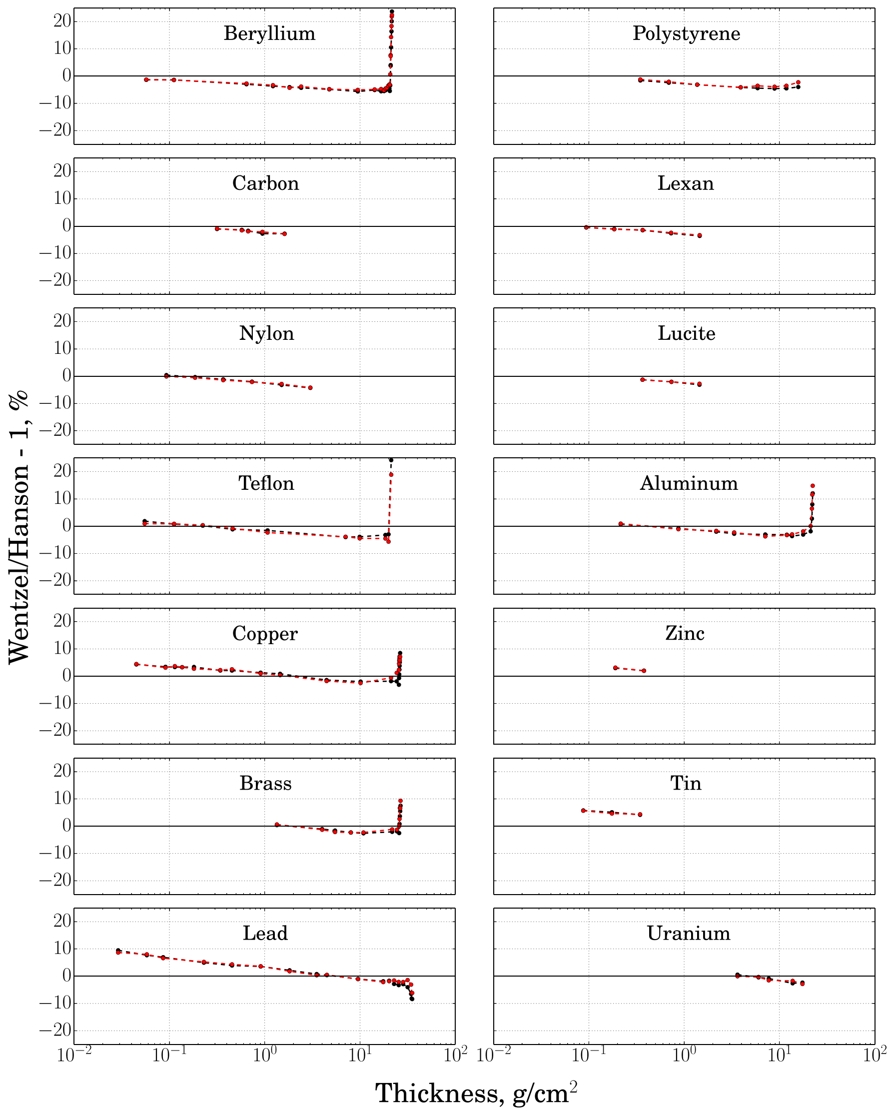

Figure 3. Comparison of the Geant4 Wentzel simulation with Hanson. Black points: simulation with Geant4 default values of I. Red (grey) points: simulation with I tuned to fit the Janni66 range-energy tables (Janni 1982a).

4.3. Matysiak et al

In this poster presentation (Matysiak *et al* 2013) Matysiak *et al* develop an MC tool to assess the accuracy of the Eclipse pencil beam model.

First, to select the best Geant4 model, they consider spreading of an ideal beam in a 7.5 cm water equivalent Lexan range shifter (RS) at six incident energies 120 – 226.7 MeV, each at four step sizes. (The RS thickness is 6.507 cm.) Urban (option 3) and Wentzel (option 4) simulations are compared with σ_x values from 'analytical calculations using Molière scattering'. We computed our own σ_x values, finding $\sigma_x(R)/R = 0.0213$ from equation (5) and $\sigma_x(6.507\ \mathrm{cm})$ values from equation (6) consistently 8% higher than Matysiak's. Therefore,

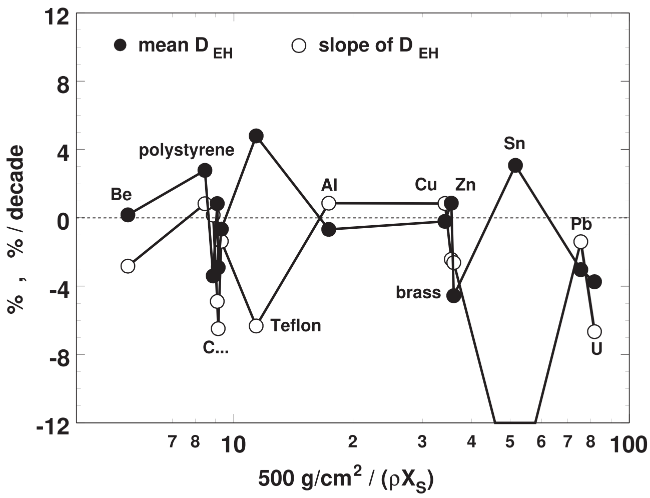

Figure 4. Fitted experimental results. The horizontal scale, (500 g cm^{-2})/(mass scattering length), is arbitrary. Filled circles are mean D_{EH} (%); open circles are slope D'_{EH} (%/decade). 'C…' stands for C, Lexan, Nylon and Lucite in that order.

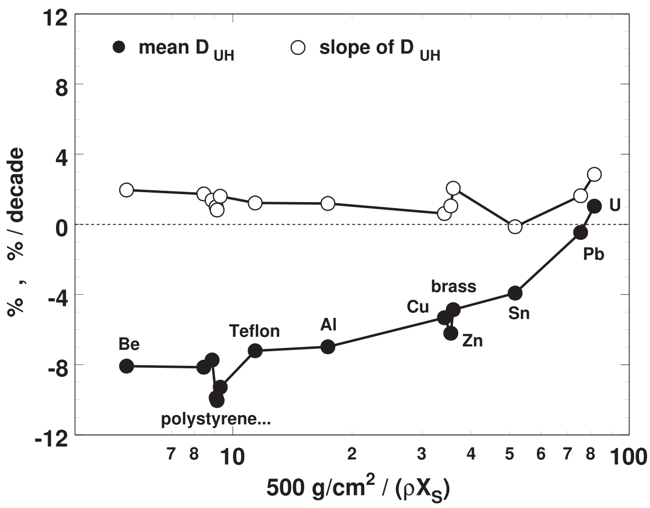

Figure 5. Fitted Urban results. The horizontal scale, (500 g cm^{-2})/(mass scattering length), is arbitrary. Filled circles are mean D_{UH} (%); open circles are slope D'_{UH} (%/decade). 'polystyrene…' stands for polystyrene, C, Lexan, Nylon and Lucite in that order.

over the energy range, Wentzel beam spreading in Lexan is either found to be ≈4 to 0% high (Matysiak theory) or ≈4 to 8% low (our theory). The Urban model is found to be step-size dependent and therefore not used further by Matysiak.

Having compared simulated beam spreading in a homogeneous slab with theory, Matysiak *et al* proceed to compare a target/drift experiment with measurement, using the RS as the target. First, they determine the incident beam size, divergence and emittance by fitting measurements of the open beam in air. Next, inserting the RS and using those open beam parameters, they simulate and measure transverse fluence distributions at five locations along the beam

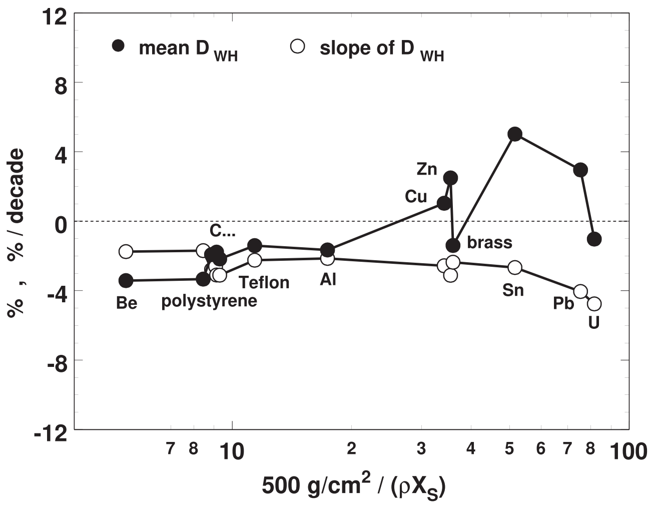

Figure 6. Fitted Wentzel results. The horizontal scale, (500 g cm^{-2})/(mass scattering length), is arbitrary. Filled circles are mean D_{WH} (%); open circles are slope D'_{WH} (%/decade). 'C…' stands for C, Lexan, Nylon and Lucite in that order.

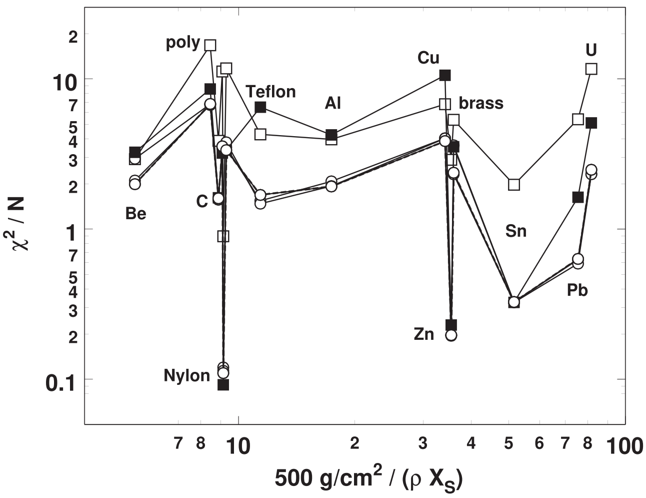

Figure 7. Summary of χ^2/N for CERN web site runs dated 15FEB2016. The abscissa is arbitrary, with some materials labeled for reference. All Geant4 configurations are prefaced emstandard. Empty squares = opt3 + elastic, full squares = opt0 + none. The remaining four, better and essentially indistinguishable, are opt4 + elastic, opt0 + elastic, WVI + elastic, WVInoDisp + elastic, all shown as empty circles.

axis covering a range of 35 cm, fitting with Gaussians to find simulated and measured σ_x. Geant4 Wentzel is high by $\approx 7\%$ at 120 MeV improving to $\approx 0\%$ at 226.7 MeV. It is better in the *y* direction than in the *x* direction (presumably the bend plane).

Unlike beam spreading, the target/drift experiment comes close to a direct test of the Geant4 MCS model. Assuming the energy dependence to be largely due to sensitivity to beam parameters, and allowing for some experimental error, the Matysiak study is not inconsistent with our finding that Wentzel is $\approx 2\%$ low in the neighborhood of Lexan see figure 6.

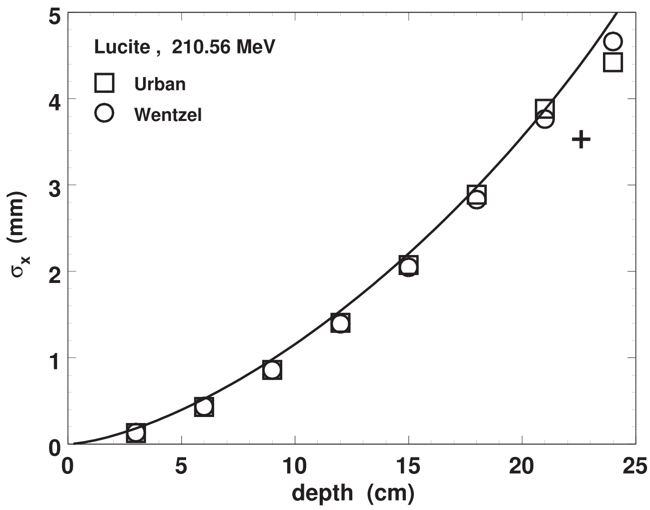

Figure 8. Beam spreading in Lucite (PMMA) at 210.56 MeV incident. Line: theory of Preston and Koehler (1968). Points: simulation with Geant4 Urban and Wentzel models as described in text. For both, $\langle \text{error} \rangle = -0.14$ mm and $\langle \text{error} \rangle / \langle \sigma_x \rangle = -6.4\%$. Cross: point at 22.6 cm depth per Grevillot *et al* (2010).

4.4. Grevillot et al

Grevillot *et al* (2010) optimized GEANT4 settings for proton pencil beam scanning simulations using GATE. In the section relevant here, they found (their figures 10 and 11) that GATE-simulated beam spreading in PMMA (Lucite) underestimated experiment by an energy dependent amount reaching 20% at 210.56 MeV incident. They measured σ_x using EBT radiochromic film in a PMMA phantom. Quantitive dosimetry with radiochromic film is an exacting technique subject to nonlinear dose response, LET dependence and sensitivity to scanning technique (Niroomand-Rad *et al* 1998). Indeed, they describe their own results as ‘preliminary’ and ‘qualitative’.

Even so, this paper invites the question whether Geant4 simulations of beam spreading in PMMA (and presumably, other water-like materials) could possibly be low by as much as 20%. In figure 8 we compare a Geant4 simulation, with the settings described above, with the theory of Preston and Koehler (1968) as summarized by equations (5) and (6). There is little difference between the Urban and Wentzel models. Both are poor near end-of-range and do well elsewhere. For both, average difference between MC and the theory is $\langle \text{error} \rangle = -0.14$ mm and $\langle \text{error} \rangle / \langle \sigma_x \rangle, = -6.4\%$ including the last point. The possibility of a 20% shortfall in Geant4 at 22.6 cm (cross) is ruled out.

5. Summary

Measurements of the distribution of outgoing angles from a target test the MCS model of a Monte Carlo program and normally, models are validated directly against such experimental data. We have argued that, in the special case of MCS, theory may be taken as ground truth because it is free of adjustable parameters and agrees, on average, with a very large body of data. That reveals trends in the models that would otherwise be obscured by experimental scatter.

We have concentrated on the target/drift configuration, which measures outgoing angles. First, we justified the ground truth assumption by comparing experiment with the Molière/Fano/Hanson theory. We then compared Geant4 simulations, using the Urban and Wentzel models, with the same theory.

Our figures 5 and 6 give Wentzel a slight advantage in proton radiotherapy where the materials of greatest interest are water-like. For the highest-Z materials Urban is at least as good.

For completeness, we discussed beam-spreading experiments, also considered to be tests of the MCS model. Here the relevant theory is that of Preston and Koehler (1968) as re-derived in Gottschalk (2012) and summarized by our equations (5) and (6). Geant4 simulation of beam spreading in PMMA agrees with theory to a fraction of a millimeter or 6.4% for both models, contradicting the finding of Grevillot *et al* (2010).

Acknowledgments

BG is indebted to the Physics Department of Harvard University for continuing support.

References

Berger M J, Inokuti M, Andersen H H, Bichsel H, Powers D, Seltzer S M, Thwaites D, Watt D E, Paul H and Sternheimer R M 1993 Stopping powers and ranges for protons and alpha particles *ICRU Rep.* (https://doi.org/10.1093/jicru/os25.2.Report49)

Bethe H 1953 Moliere's theory of multiple scattering *Phys. Rev.* **89** 1256–66

Fuchs H, Vatnitsky S, Stock M, Georg G and Grevillot L 2015 Evaluation of GATE/Geant4 multiple Coulomb Scattering precision for a 160 MeV proton beam *Proc. Annual Congress of the Austrian Society of Medical Physics, Wiener Neustadt*

Gottschalk B 2014 Software (BGware.zip), including Windows executables, data files and Fortran source code, free download at http://users.physics.harvard.edu/~gottschalk/ or the Particle Therapy Co-Operative Group (PTCOG) website http://ptcog.web.psi.ch/

Geant4 Collaboration 2016 *Physics Reference Manual* Version: geant4 10.3 (9 December 2016)

Gottschalk B 2010 On the scattering power of radiotherapy protons *Med. Phys.* **37** 352–67

Gottschalk B 2012 Techniques of proton radiotherapy: transport theory (arXiv:1204.4470)

Gottschalk B, Koehler A, Schneider R, Sisterson J and Wagner M 1993 Multiple Coulomb scattering of 160 MeV protons *Nucl. Instrum. Methods Phys. Res.* B **74** 467–90

Grevillot L, Frisson T, Zahra N, Bertrand D, Stichelbaut F, Freud N and Sarrut D 2010 Optimization of Geant4 settings for proton pencil beam scanning simulations using GATE *Nucl. Instrum. Methods Phys. Res.* B **268** 3295–305

Hanson A O, Lanzl L H, Lyman E M and Scott M B 1951 Measurement of multiple scattering of 15.7 MeV electrons *Phys. Rev.* **84** 634–7

Highland V L 1975 Some practical remarks on multiple scattering *Nucl. Instrum. Methods* **129** 497–9

Highland V L 1979 Some practical remarks on multiple scattering *Nucl. Instrum. Methods* **161** 171

Jan S *et al* 2004 GATE: a simulation toolkit for PET and SPECT *Phys. Med. Biol.* **49** 4543

Janni J F 1982a Energy loss, range, path length, time-of-flight, straggling, multiple scattering, and nuclear interaction probability *At. Data Nucl. Data Tables* **27** 147–339

Janni J F 1982b Proton range-energy tables, 1 keV–10 GeV *At. Data Nucl. Data Tables* **27** 147–339

Koehler A M 1967 Dosimetry of proton beams using small silicon diodes *Radiat. Res. Suppl.* **7** 53–63

Matysiak W, Yeung D, Slopsema R and Li Z 2013 Evaluation of Geant4 for experimental data quality assessment in commissioning of treatment planning system for proton pencil beam scanning mode *Int. J. Radiat. Oncol. Biol. Phys.* **87** S739 (And accompanying poster at http://astro2013.abstractsnet.com)

Moliere G 1948 Theorie der Streuung schneller geladener Teilchen II Mehrfach-und Vielfachstreuung *Z. Nat.forsch.* A **3** 78–97

Niroomand-Rad A *et al* 1998 Radiochromic film dosimetry: recommendations of AAPM radiation therapy committee task group 55 *Med. Phys.* **25** 2093–115

Perl J, Shin J, Schümann J, Faddegon B and Paganetti H 2012 TOPAS: an innovative proton Monte Carlo platform for research and clinical applications *Med. Phys.* **39** 6818–37

Preston W and Koehler A 1968 The effects of scattering on small proton beams Harvard Cyclotron Laboratory. Unpublished manuscript. A facsimile is available in `BGdocs.zip` at http://users.physics.harvard.edu/~gottschalk

Schwarz S 2013 Validation of Geant4 proton scattering a summer student project http://cds.cern.ch/record/1603093/files/Students-Report_CERNSummerSchool2013_SachaSchwarz.pdf