Influence of Geant4 Physics List on Simulation Accuracy and Performance

Róbert Breier[1], Alexander Fuss[2, 3], Holger Kluck[2], Valentyna Mokina[2], Veronika Palušová[4*], Pavel Povinec[1],

1 Department of Nuclear Physics and Biophysics, Faculty of Mathematics, Physics and Informatics, Comenius University, 84248 Bratislava, Slovakia
2 Institut für Hochenergiephysik der Österreichischen Akademie der Wissenschaften, Wien, 1050, Austria
3 Atominstitut, Technische Universität Wien, Wien, 1020, Austria
4 Institute of Experimental and Applied Physics, Czech Technical University in Prague, Prague, CZ-11000, Czech Republic
* veronika.palusova@cvut.cz

September 23, 2022

14th International Conference on Identification of Dark Matter
Vienna, Austria, 18-22 July 2022
doi:10.21468/SciPostPhysProc.?

Abstract

A main goal of current low background physics is the search for rare and novel phenomena beyond the Standard Model of particle physics, e.g. the scattering off of a potential Dark Matter particle inside a $CaWO_4$ crystal or the neutrinoless double beta decay of Ge nucleus. The success of such searches depends on a reliable background prediction via Monte Carlo simulations. A widely used toolkit to construct these simulations is GEANT4, which offers a wide choice of physics models, so-called physics lists. To facilitate the selection of physics lists for simulations of $CaWO_4$ and Ge targets, we quantify their impact on the total energy deposition for several test cases.

1 Introduction

Rare event search experiments, such as the search for dark matter or neutrinoless double beta decay, depend crucially on a reliable and verified background prediction. Commonly, these predictions are based on Monte Carlo simulations of the relevant background sources. A widely used toolkit to create these simulations is GEANT4 [1–3].

In literature, the predefined physics lists[1] of GEANT4 and their settings are validated for a wide range of physics processes and observables, e.g. electron back scattering [4]. However, to our knowledge, no such study exist about the impact of the used physics list on the total energy deposited by radioactive decays in $CaWO_4$ or Ge target. As pointed out in [4], an observable like this is the result of several physics processes. Hence, it can only be assessed

[1]For the sake of simplicity, we include also the *physics constructors* in the term of *physics lists*.

for a specific use case, and it is usually precarious to extrapolate it from studies based on a different use case.

With this work we aim to provide a dedicated study for this missing use case: examine the impact of different physics lists on the total energy deposition for several configurations of our test case, i.e. combinations of radioactive contaminants, target material ($CaWO_4$ and Ge) and target thickness. The goal is to give an assessment to what extent the selected physics list affect the simulated observable of total energy deposition, i.e. how compatible different physics lists are. To assess the compatibility in a qualitative and objective way, we adopt the methodology established in [4–7]. To give some guidance which physics list to choose in case of compatibility, we consider also the computing performance.

2 Methodology

The strategy of this study is focused on the simulation of total energy deposited in a bulk crystal of two target materials, i.e. $CaWO_4$ and Ge. Simulations were produced with GEANT4 version 10.6.3 released in 2020. *Test cases*, characterized by target material, target thickness and radioactive contaminant, are simulated for each GEANT4 model, characterized by physics list. The impact of different physics lists on the total energy deposition for several *configurations* of our test case is assessed by the means of statistical methods.

2.1 Model Configurations

GEANT4 provides several predefined configurations of electromagnetic (EM) physics lists called constructors, which offer users an easy way to implement the "best set" of physics models for electromagnetic physics. Twelve EM physics constructors were are used during this work: G4EmStandardPhysics_option1, 2, 3 and 4, G4EmLivermorePhysics, G4EmLivermorePolarized-Physics, G4EmPenelopePhysics, G4EmStandardPhysics, G4EmStandardPhysicsGS, G4EmLow-EPPhysics, G4EmStandardPhysicsWVI, and G4EmStandardPhysicsSS.

We treat the physics constructors as one parametric-models with the range cut for secondary particle production as free parameter. In a GEANT4 simulation, secondary particles that are unable to travel at least the production cut value are not produced, but their energy is deposited locally. In order to study the impact of such parameter and to possibly improve the performance of simulations, each simulation was performed for five production cut values: 100 nm, 1 μm, 1 mm, 1 cm, 10 cm. In this work, one GEANT4 model configuration is characterized by physics constructor and cut value, yielding together 60 model configurations.

2.2 Test Cases

This study covers two target materials - $CaWO_4$ or Ge of two different thicknesses - thin (50 μm) and bulky (32 mm). Six common radioactive contaminants in chosen targets are considered - low Q-value β emitters (^{228}Ra, ^{210}Pb), high Q-value β emitters (^{208}Tl, ^{210}Tl) and α emitters (^{211}Bi, ^{234}U).

Each one of the 24 test cases is reproduced in the simulation for each model configuration, yielding together 1440 datasets. Based on the GEANT4 manual, we chose the most accurate model, G4EmStandardPhysics_option4 and 1 mm cut, as reference dataset.

2.3 Statistical Analysis

Quantitative comparisons between simulated test case and reference datasets is determined by statistical tests and is performed in two steps. In the first step, we address the question of how well a simulated dataset distribution is described by the distribution of the reference dataset by appropriate test statistics (Goodness-of-fit tests). The second step is based on categorical tests, which determine if the difference in compatibility observed across our GEANT4 model configurations is statistically significant. This statistical significance is evaluated in two cases: for simulations using *different* physics constructors, which produce unpaired samples, and for simulations using the *same* physics constructor but differ in a secondary feature, i.e. in production cut, which are related and produce paired samples.

As the statistical tests are not uniformly sensitive to differences in distributions at all values, a variety of statistical tests is applied in each step of analysis to minimize the possibility of introducing systematic effects in validation of results. Three independent goodness-of-fit tests were performed to test compatibility: Anderson-Darling (AD) [8], Kolmogorov-Smirnov (KS) [9, 10] and χ^2 test [11, 12]. The tests used to compare groups of categorical data for a significant difference are χ^2 test of independence and Fisher's exact test [13] for unpaired data, and McNemar's test [14] for paired data. In each test, the significance level was chosen at $\alpha = 5\%$.

3 Results

We show first results of the first step, i.e. comparing the relative compatibility of energy deposition of several test cases with the reference model configuration. Figure 1 shows the example of energy deposition per single event of ^{210}Tl decay in $CaWO_4$ for G4EmStandardPhysics_option1 and two cut values compared with the reference model. The outcome of χ^2 goodness-of-fit tests is reported in the form of efficiency, which is defined as the fraction of test cases where the p-value resulting from the χ^2 tests is larger than $\alpha = 5\%$. This quantifies the capability of model configuration to produce statistically consistent results with the reference simulation. Results of the efficiency for GEANT4 models is shown in Figure 2.

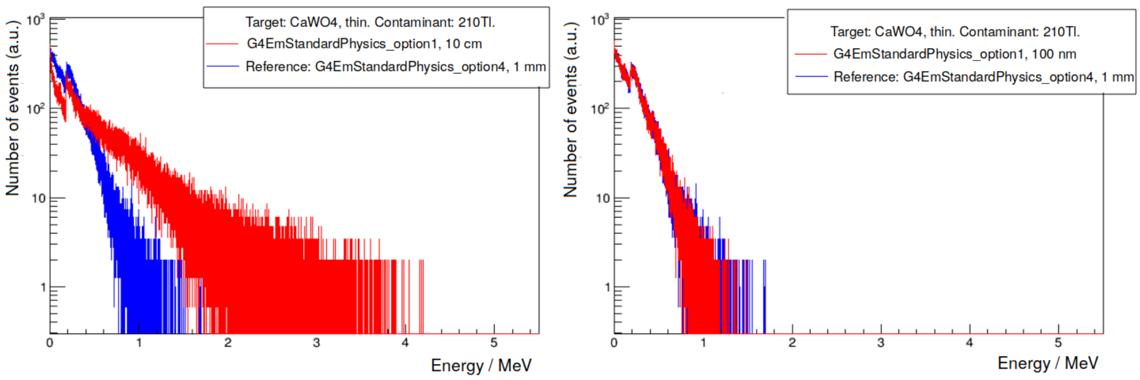

Figure 1: Energy deposition per single event in $CaWO_4$ target for 2 model configurations. Left - example of failed test of compatibility, right - example of passed test of compatibility.

The analysis identifies G4EmLivermore as a model that is robust against changes of the production cut. Hence, tuning of the production cut may improve the performance of simulations without affecting the outcome of the simulation. Analysis with AD and KS goodness-of-fit tests, and categorical analysis based on the results of compatibility, are currently conducted.

Besides maximizing the accuracy of simulations, one also wants to optimize the computational performance in terms of CPU time. Figure 3 shows the fraction of average run time of GEANT4 model configurations relative to the reference model datasets. The decreased performance of single scattering models (G4EmStandardWVI and G4EmStandardSS) is expected.

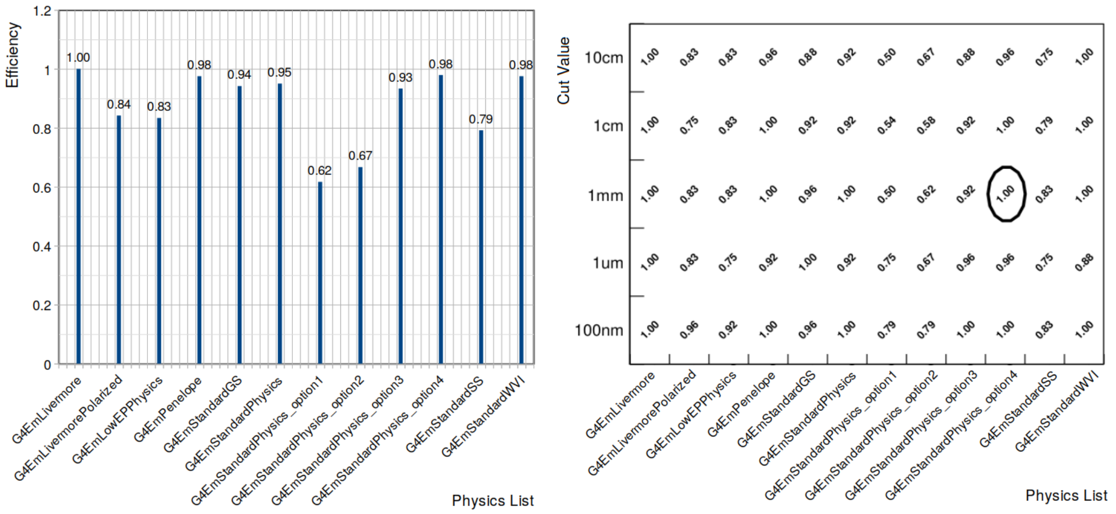

Figure 2: Efficiency of GEANT4 physics lists (left) and model configurations (right) based on the results of χ^2 goodness-of-fit test. The black ellipse marks the reference configuration.

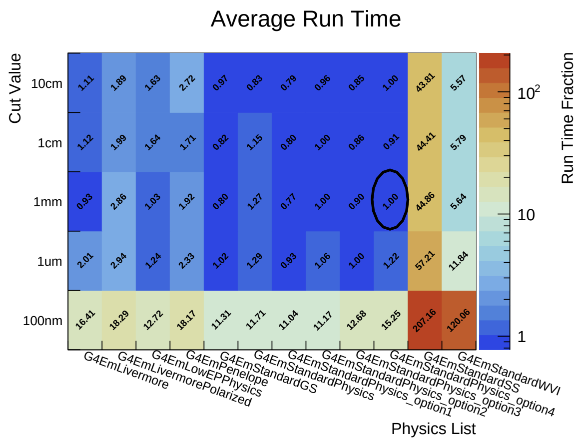

Figure 3: Average run time of GEANT4 model configurations relative to the reference model. The black ellipse marks the reference configuration.

Except for the 100 nm case, most values for the production cut affect the relative performance by $\mathcal{O}(10\,\%)$.

4 Conclusion

With this study, we aim to determine the compatibility between total energy deposition of different GEANT4 EM physics constructors compared to the most accurate G4EmStandardPhysics_option4. Besides the effects due to the physics list, this study also evaluates the sensitivity to features such as the production range cut value. First results of this analysis could already provide guidance to users in optimizing the model configuration of GEANT4 based background prediction studies.

Acknowledgements

This work has been funded through the Sonderforschungsbereich (Collaborative Research Center) SFB1258 'Neutrinos and Dark Matter in Astro- and Particle Physics', by the Austrian science fund (FWF): I5420-N, W1252-N27, P34778-N and by Austria's Agency for Education and Internationalisation (project SK 06/2018). The Bratislava group acknowledges a partial support provided by the Slovak Research and Development Agency (projects SK-AT-2017-001, APVV-15-0576 and APVV-21-0377).

References

[1] S. Agostinelli *et al.*, *GEANT4–a simulation toolkit*, Nuclear Instruments and Methods A **506**, 250 (2003), doi:10.1016/S0168-9002(03)01368-8.

[2] J. Allison *et al.*, *Geant4 developments and applications*, IEEE Transactions on Nuclear Science **53**, 270 (2006), doi:10.1109/TNS.2006.869826.

[3] J. Allison *et al.*, *Recent developments in Geant4*, Nuclear Instruments and Methods A **835**, 186 (2016), doi:10.1016/j.nima.2016.06.125.

[4] T. Basaglia, M. C. Han, G. Hoff, C. H. Kim, S. H. Kim, M. G. Pia and P. Saracco, *Quantitative test of the evolution of geant4 electron backscattering simulation*, IEEE Transactions on Nuclear Science **63**(6), 2849 (2016), doi:10.1109/tns.2016.2617834.

[5] S. Guatelli, A. Mantero, B. Mascialino, M. G. Pia and V. Zampichelli, *Validation of geant4 atomic relaxation against the nist physical reference data*, IEEE Transactions on Nuclear Science **54**(3), 594 (2007), doi:10.1109/TNS.2007.894814.

[6] A. Lechner, M. G. Pia and M. Sudhakar, *Validation of geant4 low energy electromagnetic processes against precision measurements of electron energy deposition*, IEEE Transactions on Nuclear Science **56**(2), 398 (2009), doi:10.1109/tns.2009.2013858.

[7] M. Batic, G. Hoff, M. G. Pia, P. Saracco and G. Weidenspointner, *Validation of geant4 simulation of electron energy deposition*, IEEE Transactions on Nuclear Science **60**(4), 2934 (2013), doi:10.1109/tns.2013.2272404.

[8] T. W. Anderson and D. A. Darling, *A test of goodness of fit*, Journal of the American Statistical Association **49**(268), 765 (1954), doi:10.1080/01621459.1954.10501232.

[9] A. L. Kolmogorov, *Sulla determinazione empirica di una legge di distribuzione*, Giornale dell'Istituto Italiano Degli Attuari **4**, 83 (1933).

[10] N. V. Smirnov, *On the estimation of discrepancy between empirical curves of distribution for two independent samples*, Moscow University Mathematics Bulletin **2**, 3 (1939).

[11] K. Pearson, *X. on the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling*, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **50**(302), 157 (1900), doi:10.1080/14786440009463897.

[12] W. G. Cochran, *The χ^2 Test of Goodness of Fit*, The Annals of Mathematical Statistics **23**(3), 315 (1952), doi:10.1214/aoms/1177729380.

[13] R. A. Fisher, *On the Interpretation of χ^2 from Contingency Tables, and the Calculation of P*, doi:10.2307/2340521 (1922).

[14] Q. McNemar, *Note on the sampling error of the difference between correlated proportions or percentages*, doi:10 (1947).