

A Toolkit to Study Sensitivity of the Geant4 Predictions to the Variations of the Physics Model Parameters

Laura Fields¹, Krzysztof Genser¹, Robert Hatcher¹, Michael Kelsey², Gabriel Perdue¹, Hans Wenzel¹, Dennis H. Wright², Julia Yarba¹

¹ Fermi National Accelerator Laboratory[†], Batavia, IL, USA

² SLAC National Accelerator Laboratory, Stanford, CA, USA

E-mail: yarba_j@fnal.gov

Abstract. Geant4 is the leading detector simulation toolkit used in high energy physics to design detectors and to optimize calibration and reconstruction software. It employs a set of carefully validated physics models to simulate interactions of particles with matter across a wide range of interaction energies. These models, especially the hadronic ones, rely largely on directly measured cross-sections and phenomenological predictions with physically motivated parameters estimated by theoretical calculation or measurement. Because these models are tuned to cover a very wide range of possible simulation tasks, they may not always be optimized for a given process or a given material. This raises several critical questions, e.g. how sensitive Geant4 predictions are to the variations of the model parameters, or what uncertainties are associated with a particular tune of a Geant4 physics model, or a group of models, or how to consistently derive guidance for Geant4 model development and improvement from a wide range of available experimental data. We have designed and implemented a comprehensive, modular, user-friendly software toolkit to study and address such questions. It allows one to easily modify parameters of one or several Geant4 physics models involved in the simulation, and to perform collective analysis of multiple variants of the resulting physics observables of interest and comparison against a variety of corresponding experimental data. Based on modern event-processing infrastructure software, the toolkit offers a variety of attractive features, e.g. flexible run-time configurable workflow, comprehensive bookkeeping, easy to expand collection of analytical components. Design, implementation technology, and key functionalities of the toolkit are presented and illustrated with results obtained with Geant4 key hadronic models.

1. Introduction

Geant4 [1] is the leading detector simulation toolkit used to design detectors and to optimize calibration and reconstruction software in such domains as high energy and nuclear physics, space science, reactor safety, and medical applications. It offers a set of carefully validated physics models to simulate interactions of particles with matter across a wide range of interaction energies. Because Geant4 cannot offer a single modeling algorithm to cover the entire energy domain from zero to the TeV scale, for all known processes and particles, a combination of models is typically used to perform a simulation task. Each hadronic model relies on a substantial number of parameters involved in calculations associated with modeling the interactions. This naturally leads to the questions of how sensitive the Geant4 predictions are to the variations of

these parameters, and what uncertainties are associated with a Geant4 physics model, or a group of models, involved in simulation and optimization of a detector design. As of today, at least 3 Geant4 hadronic models, PreCompound [2], Bertini cascade [3] (Bertini) and Fritiof [4] (FTF), offer a configuration interface that allows to vary up to 39 parameters and 17 switches, in order to study the sensitivity of the simulated results. The Geant4 collaboration is currently working to expand the number of configurable hadronic models and to identify additional parameters and switches to include in the configuration interface. However, while studying the sensitivity of the Geant4 simulated results, it is important to prevent the use of physically incorrect settings of the model parameters or the neglect of correlations among parameters. Thus, an important part of advertising to the users community configurable parameters of hadronic models is documentation of the validity range of each parameters, and the correlations of parameters, where applicable. This challenge has motivated the Geant4 collaboration to start a new initiative and develop a toolkit to study the effects of varying model parameters on the simulated results, and to explore the associated uncertainties. The primary goal is to estimate the physically meaningful range of validity of the different model parameters and to determine their optimal values and uncertainties from a global fit to relevant data. Model parameters exposed to the users will be thoroughly documented. As an additional benefit, understanding variations in the parameters and how they can possibly impact systematics may provide valuable input towards tuning Geant4 models by the authors.

2. Toolkit Features

We have designed and implemented a comprehensive, modular, user-friendly software package that allows the user to randomly sample in the multiparameter space and to run simulation with multiple configurations of the Geant4 models. Another feature is the possibility to run a collective analysis of multiple variants of the resulting physics observables of interest and simultaneous benchmark versus relevant experimental datasets (one or many, as applicable). For benchmarking simulated results versus relevant experimental data or to perform statistical analysis, programmatic (C++) access to the DoSSiER repository [7] is also implemented. Geant4 version 10.3.patch01 or later is required to utilize this software. The toolkit is based on a modern event-processing framework Art [5] and offers a variety of features, such as a user-friendly model configuration API, flexible run-time configurable workflow, geometry and sensitive detector setup in GDML [6] if needed, comprehensive bookkeeping, analysis of multiple variants of the resulting physics observables of interest, and an extensible design. The toolkit components and the workflow are schematically presented in figure 1.

3. Preliminary Results

In our initial study we explored sensitivity of the Bertini and FTF simulated results to the variations of their parameters. An example of the effects of varying selected Bertini parameters on the simulated energy deposit in the LArIAT [9] Liquid Argon detector is shown in figure 2. The simulation uses the QGSP_FTFP_BERT physics list [1, 8] that includes the Bertini covering the 0-9.9 GeV energy range. Predictions using the default values of Bertini parameters were compared with results obtained with several alternative settings of these parameters within a physically meaningful range. In this study the varied parameters were the nuclear radius and the “internal” hadron-nucleon cross-section for an incident particle traveling through the target nucleus (“Radius Scale” and “XSec Scale” in figures 2 and subsequent figures). The toolkit has also been used to model hadron-nucleus interactions with various settings of the Geant4 Bertini or FTF parameters and to perform initial statistical analysis of the simulated results by benchmarking them versus relevant data from HARP [10, 11] and ITEP-771 [12] experiments. Results are given as χ^2 per number of degrees of freedom (χ^2/ndf). The χ^2 is calculated for each simulated spectrum and the corresponding data spectrum. Figure 3 shows preliminary results

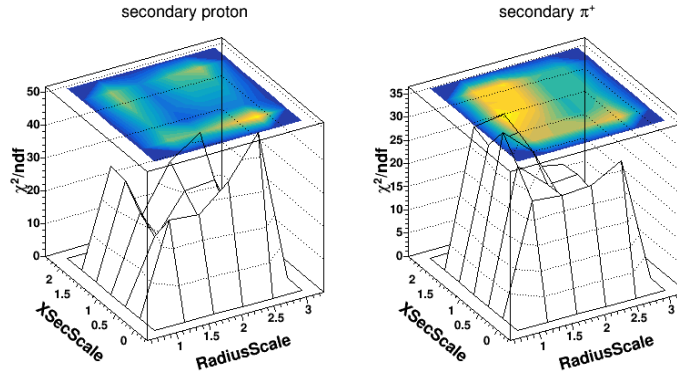


Figure 3. Statistical comparison of the simulated results for proton (left) and π^+ (right) production by a 5 GeV/c proton beam interacting with a Lead nucleus. Results are shown in a form of χ^2/ndf as a function of Bertini nuclear radius scale (Radius Scale) and the “internal” hadron-nucleon cross-section scale (XSec Scale) settings. The χ^2 is calculated for each simulated spectrum, as obtained for a given pair of radius scale and cross-section scale settings, and the relevant experimental data spectrum [10, 11, 12]. The ndf is the number of degrees of freedom.

on the χ^2/ndf as a function of the Bertini parameters values in the 2-parameters space.

We also varied the parameter that controls modeling of the trailing effect (“Trailing Radius”). The trailing effect takes into account the local depletion of the nuclear density following an intra-nuclear collision; it is roughly the radius (in fm) of the depleted region. In this study we randomly selected 100 settings of the parameter uniformly distributed within its validity range. Preliminary results are shown in figure 4 as χ^2/ndf versus the value of Bertini trailing radius. Results of modeling proton production in hadron-nucleus interactions with different variants of Bertini appear to favor the trailing radius setting of approximately 0.7 while by default this feature is disabled (i.e. trailing radius is set to zero). However, it is important to remember that there exist correlations among the Bertini parameters. Figure 5 illustrates the effect of varying nuclear radius on simulated hadron production in hadron-nucleus interactions as modeled with the trailing radius set to zero (default) or to 0.7. While both sets of χ^2/ndf as a function of nuclear radius are sensitive to the nuclear radius value, they exhibit substantially different patterns. In figures 4 and 5 the χ^2 is calculated for each simulated spectrum and the relevant experimental data spectrum [10, 11, 12].

We have performed a similar study using FTF to model hadron production in hadron-nucleus interaction and to explore the sensitivity of the simulated observables. Preliminary results are shown in figures 6 and 7. Figure 6 illustrates how the χ^2/ndf depends on the value of one of the FTF parameters that control modeling of the nuclear destruction of the target nucleus [13] (“FTFP NUCDESTR.P1.TGT” in figure 6). The parameter in question may be set to a fixed value (e.g. 1.0 in the Geant4 release 10.3.patch01) or it can be expressed as a function of the number of nucleons in the target nucleus (as projected in future Geant4 releases). In this study we varied this parameter assuming that its value is the same for all types of the target nuclei, and we simulated hadron production in proton-Carbon or proton-Lead interactions at 5 GeV/c. The χ^2 is calculated for each simulated spectrum and the relevant experimental data spectrum [10, 11, 12]. As presented in figure 6, the distributions of χ^2/ndf as a function of NUCDESTR.P1.TGT obtained for light (Carbon) or heavy (Lead) nuclei exhibit substantially different patterns, which is especially clear for the secondary protons. This is likely to confirm that the parameter should depend on the number of nucleons in the target nucleus. However,

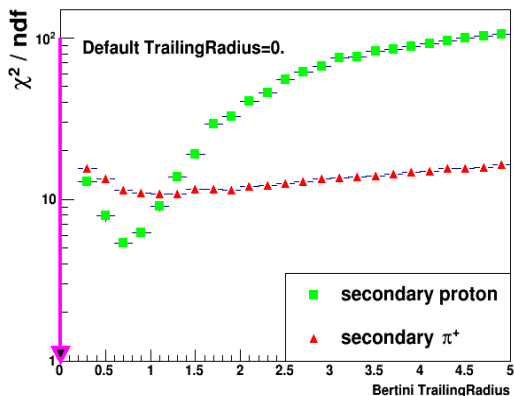


Figure 4. Statistical comparison of the simulated results for proton and π^+ production in proton-Lead interactions at 5 GeV/c. Results are shown in a form of χ^2/ndf as a function of Bertini trailing radius. Experimental data from HARP [10, 11] and ITEP-771 [12] are used to calculate the χ^2/ndf .

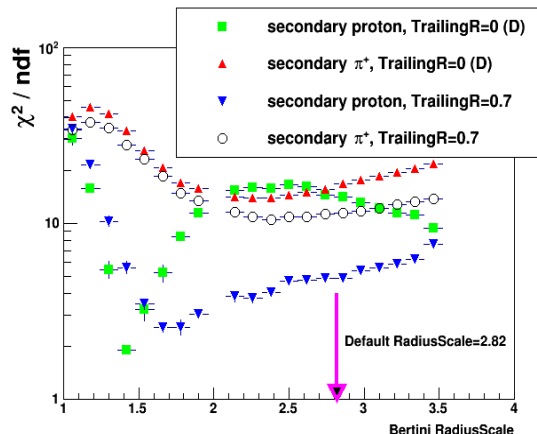


Figure 5. Statistical comparison of the simulated results on proton and π^+ production in proton-Lead interactions at 5 GeV/c. The χ^2/ndf versus Bertini nuclear radius is shown for different settings of the trailing radius. Experimental data are from HARP [10, 11] and ITEP-771 [12].

the correlation among FTF parameters should be thoroughly explored before a more definite statement is made. Figure 7 illustrates how the simulated results on hadron production in hadron-nucleus interactions depend on one of the parameters, average transverse momentum (“FTFP BARYON_AVRG_PT2” in figure 7) that is involved in modeling excitation of the participating hadron as a QCD string (the string is further decayed through the fragmentation mechanism). The χ^2 is calculated for each simulated spectrum and the relevant experimental data spectrum [14].

4. Future Plans

Our nearest-term task is to understand errors and correlations of the Geant4 hadronic models parameters via simulation and analysis in the multiparameter space, across a wide range of energies, beam particle types, and target nuclei. In addition to the possibility to run various types of Geant4-based simulation in the multiparameters space, we have also developed an interface to the Professor package [15] that has been used for tuning a number of Monte Carlo event generators. Professor relies on the bin-wise parameterisation of the results simulated in the multiparameters to a low order polynomial function; then it uses such function to construct the $\chi^2 = \sum_{bin} ((interpolation - data)^2 / error^2)$ and to further numerically minimize it. The obvious benefit of such approach is that it substantially reduces the number of simulated samples required for the study. We plan to obtain more results from the Professor-based statistical analysis for such key Geant4 hadronic models as PreCompound, Bertini, and FTF.

5. Summary

In response to requests from the user community, we are developing a software toolkit to explore the impact of varying Geant4 model parameters on the simulated physics results, to refine the validity ranges of the parameters, and to numerically determine correlations among them. Comparison of simulated results versus experimental data, including statistical analysis, can be done via programmatic access to the DoSSiER repository. The toolkit was used to study the

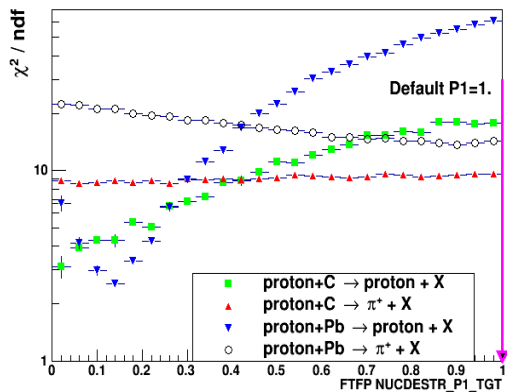


Figure 6. Statistical comparison of the simulated results for proton and π^+ production in proton-Carbon or proton-Lead interactions at 5 GeV/c. Results are shown as χ^2/ndf versus one of the FTF parameters involved in modeling nuclear destruction of the target nucleus. The χ^2/ndf is calculated versus data from HARP [10, 11] and ITEP-771 [12].

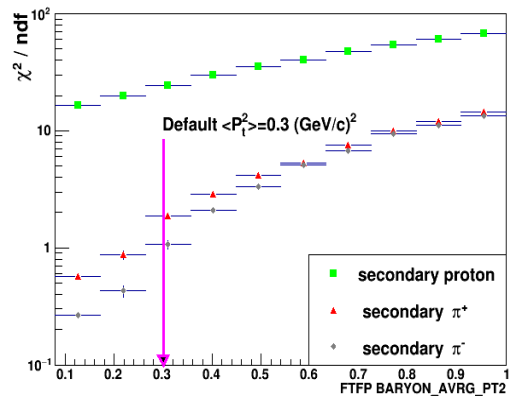


Figure 7. Statistical comparison of the simulated results for proton and pion production in proton-Carbon interactions at 158 GeV/c. Results are shown as χ^2/ndf versus one of the FTF parameters involved in modeling excitation of the participating hadron. The χ^2/ndf is calculated versus spectra measured by NA49 [14].

effects of varying Geant4 hadronic model parameters on simulated observables, with Geant4 Bertini and FTF models as the use-cases. Selected results are included in this presentation for illustration. Our near-term plans include further development of the toolkit and use of the Professor tuning toolkit for analysis in the multiparameter space.

Acknowledgments

† Operated by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the United States Department of Energy.

References

- [1] Allison J et al. 2016 *Nucl. Instrum. Meth. A* **835** 186
- [2] <http://cern.ch/geant4/UserDocumentation/UsersGuides/PhysicsReferenceManual/fo/PhysicsReferenceManual.pdf>
- [3] Wright D H and Kelsey M 2015 *Nucl.Instrum.Meth. A* **804** 175
- [4] Uzhinsky V 2013 *International Conference on Calorimetry for the High Energy Frontier* 260
- [5] <http://art.fnal.gov/>
- [6] Chytracek R, McCormick J, Pokorski W and Santin G 2006 *IEEETrans.Nucl.Sci.* **53** 2892
- [7] Wenzel H et al., 2017 DoSSiER: Database of Scientific Simulation and Experimental Results *Preprint* FERMILAB-CONF-16-508-CD
- [8] <http://cern.ch/geant4/UserDocumentation/UsersGuides/ForApplicationDeveloper/html>
- [9] <http://lariat.fnal.gov/>
- [10] Catanesi M G et al. 2008 *Phys.Rev. C* **77** 05520
- [11] Apollonio M et al. 2009 *Phys.Rev. C* **80** 035208
- [12] Bayukov Yu et al. 1983 Preprint ITEP-83-148; Bayukov Yu et al. 1985 *Sov.J.Nucl.Phys.* **42** 116
- [13] <http://geant4.web.cern.ch/geant4/UserDocumentation/UsersGuides/ForToolkitDeveloper/html>
- [14] <https://spshadrons.web.cern.ch/spshadrons>
- [15] <http://professor.hepforge.org>