Journal of Physics: Conference Series **119** (2008) 032007 doi:10.1088/1742-6596/119/3/032007

Validating Geant4 versions 7.1 and 8.3 against 6.1 for *BABAR*

Swagato Banerjee,[a1] David N Brown,[b2] Chunhui Chen,[c3] David Côté,[d4] Gregory P Dubois-Felsmann,[e5] Igor Gaponenko,[b6] Peter C Kim,[e7] William S Lockman,[f8] Homer A Neal,[e9] Gabriele Simi,[c10] Alexandre V Telnov[g11] and Dennis H Wright[e12]

[a] University of Victoria, British Columbia, V8W 3P6, Canada
[b] Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA
[c] University of Maryland, College Park, MD 20742, USA
[d] Université de Montréal, Physique des Particules, Montréal, Québec, H3C 3J7, Canada
[e] Stanford Linear Accelerator Center, Stanford, CA 94309, USA
[f] University of California at Santa Cruz, Institute for Particle Physics, Santa Cruz, California 95064, USA
[g] Princeton University, Princeton, New Jersey 08544, USA

E-mail: [1] Swagato.Banerjee@slac.stanford.edu

E-mail: [2] Dave_Brown@lbl.gov

E-mail: [3] cchen23@slac.stanford.edu

E-mail: [4] cote@slac.stanford.edu

E-mail: [5] gpdf@slac.stanford.edu

E-mail: [6] IAGaponenko@lbl.gov

E-mail: [7] Peter.Kim@slac.stanford.edu

E-mail: [8] William.Lockman@slac.stanford.edu

E-mail: [9] Homer.Neal@slac.stanford.edu

E-mail: [10] Gabriele.Simi@slac.stanford.edu

E-mail: [11] avtelnov@princeton.edu

E-mail: [12] Dennis.Wright@slac.stanford.edu

Abstract. Since 2005 and 2006, respectively, Geant4 versions 7.1 and 8.3 have been available, providing: improvements in modeling of multiple scattering; corrections to muon ionization and improved MIP signature; widening of the core of electromagnetic shower shape profiles; newer implementation of elastic scattering for hadronic processes; detailed implementation of Bertini cascade model for kaons and lambdas, and updated hadronic cross-sections from calorimeter beam tests. The effects of these changes in simulation are studied in terms of closer agreement of simulation using Geant4 versions 7.1 and 8.3 as compared to Geant4 version 6.1 with respect to data distributions of: the hit residuals of tracks in *BABAR* silicon vertex tracker; the photon and K^0_L shower shapes in the electromagnetic calorimeter; the ratio of energy deposited in the electromagnetic calorimeter and the flux return of the magnet instrumented with a muon detection system composed of resistive plate chambers and limited-streamer tubes; and the muon identification efficiency in the muon detector system of the *BABAR* detector.

1. Introduction

The *BABAR* detector at the PEP-II e^+e^- storage ring at the Stanford Linear Accelerator Center collects data resulting from e^+e^- collisions at center-of-mass energies near 10.58 GeV, the mass of the $\Upsilon(4S)$ resonance. The response of the detector to the passage of particles through its material is simulated by code which is built on top of the Geant4 simulation tookit [1, 2]. Geant4 version 6.1 was used to generate simulated events and backgrounds corresponding to the data-taking periods 1999–2006. For the 2007 data-taking period, Geant4 version 7.1 is being used. For data to be taken during 2008, it is planned to use Geant4 version 8.3, and validation of this version is now being performed. There are also plans for a complete re-simulation of the entire 1999–2008 *BABAR* data set with Geant4 version 8.3.

In this note, we report on a subset of the validation results obtained by comparing data distributions with simulations performed using Geant4 versions 6.1, 7.1 and 8.3.

2. Detector and simulation Code

The *BABAR* detector is described in detail in Ref. [3]. Charged particles are reconstructed as tracks with a five-layer silicon vertex tracker (SVT) and a 40-layer drift chamber (DCH) inside a 1.5-T solenoidal magnetic field. Immediately outside the drift chamber is a ring-imaging Cherenkov detector (DIRC) used to identify charged pions, kaons and protons and provide additional electron-identification information. Outside of the DIRC is an electromagnetic calorimeter (EMC) consisting of 6580 CsI(Tl) crystals, used to identify photons and K_L^0's. The flux return of the solenoid is instrumented (IFR) with resistive-plate chambers and limited-streamer tubes, and is used to identify muons and K_L^0's.

In *BABAR*, simulation is carried out in four stages: event generation, propagation and decay of generated particles through the detector, digitization of the detector response and event reconstruction. The second stage uses the Geant4 toolkit to simulate the geometry and material composition of the PEP-II beamline and the detector components mentioned above, and specify all the physical processes required for modelling of particle propagation through the detector.

3. Geant4 updates

Though the specification of physics processes used in the *BABAR* simulation has not changed significantly since 2004, the Geant4 implementations of these processes and their underlying theoretical models have. These changes include the addition of previously neglected processes, tuning of existing code, newly developed code and bug fixes. The most important updates to the Geant4 version 6.1 toolkit are listed below.

3.1. Version 7.1 highlights

- **Pion cross-section improvements:** Compared to recent cross-section measurements, some of the total pion inelastic cross sections in earlier Geant4 versions were found to be about 25% too high for elements such as Cs and I between the momenta of 180 and 800 MeV/c. These were accordingly lowered. Various calorimeter results also indicated that π^+ cross sections for elements between Al and Fe were about 8% too high between 1 and 3 GeV/c, while for π^-, the Fe cross sections were 10% too low. These were corrected as well.

- **Ionization step-size correction:** In the *BABAR* simulation, low-momentum (< 120 MeV/c) pions were seen to have dramatically large energy losses which affected the drift-chamber pseudo-efficiencies, among other things. Within Geant4 version 6.1, the abnormally large energy loss was traced to large step sizes in the ionization energy loss code. In *BABAR*, this was first fixed at the simulation level by limiting the energy loss to reasonable values, and then by upgrading to Geant4 version 7.1.

- **Use of Bertini cascade model for kaon- and hyperon-induced reactions:** The Bertini cascade code was extended to provide detailed treatment of strange-particle interactions in the few-GeV/c region and below. The biggest improvement is expected in the simulation and analysis of K_L^0-initiated reactions in the EMC and IFR.

- **Improvements in multiple scattering:** The Highland formula, which describes the width of the central part of the multiple-scattering distribution, was modified, resulting in slightly broader electromagnetic showers. Other modifications in the multiple-scattering code have significantly decreased its dependence on the step size.

- **Improved stopping-power corrections:** Improved stopping-power corrections for hadrons and ions now agree with NIST data to within 2% in all cases. A slight improvement in simulation of the hadron energy deposition can be expected as a result.

3.2. Version 8.3 highlights

- **Improvments in multiple Coulomb scattering:** The multiple-scattering model now generates a transverse displacement as well as an angle deviation at the end of each step. The angular deviation and displacement are also correlated. This will result in a slightly broader shower shape, in agreement with data. An option is also provided that allows users to override the default step limitation invoked at the entrance to a new volume. This option is exercised in the latest version of the *BABAR* simulation.

- **Improved bremsstrahlung:** For hard bremsstrahlung emission, electron recoil is now included in the final state. This is expected to make the electromagnetic shower shape wider and to improve e/π discrimination and π^0 resolution.

- **Improved minimum-ionizing signature:** For muons and hadrons, the maximum energy transfer to delta-electrons now takes into account hadron size effects. This is important for ionization by very energetic hadrons.

- Z **and** A **dependence of pion cross sections:** The scaling of interpolated pion cross sections was changed from $Z^{2/3}$ to $A^{3/4}$ to better reflect the stronger pion absorption on nuclei. This results in slightly increased interpolated cross-section values.

- **Updated CLHEP version:** Geant4 versions 8.3 and later require the use of CLHEP 1.9 or 2.0. The *BABAR* version of CLHEP, which used to mostly conform with version 1.8, has been completely replaced by CLHEP 1.9 to allow integration of toolkit from Geant4 version 8.3.

4. Validation results

4.1. DCH pseudoefficiency of tracks

Pseudoefficiency is defined as the ratio of the number of tracks reconstructed in DCH normalized to those reconstructed in SVT, assuming that the SVT is 100% efficient in reconstructing well-measured tracks within its acceptance. As can be seen from Fig. 1, small improvements are noticed around 100 MeV/c, thanks to the ionization bug-fix in Geant4 version 7.1.

4.2. SVT hit residuals of tracks

Distributions for the SVT hit residuals are shown in Fig. 2 and Fig. 3 for tracks with momenta less than and more than 1 GeV/c respectively. Small improvements in Geant4 version 7.1 as

compared to data with Geant4 version 6.1 are clearly visible in these plots, which are due to improvements in the simulation of multiple scattering of the tracks.

4.3. Muon reconstruction and identification efficiency

Fig. 4 shows that with Geant4 version 7.1 the data/simulation ratio of the muon identification efficiency is closer to unity over a wider range of μ momenta than in Geant4 version 6.1, thanks to better simulation of multiple scattering of muons in IFR's iron yoke and brass absorber.

4.4. Shower shapes in EMC

Shower-shape profiles in the EMC are studied in terms of their lateral and second moments for control samples containing photons and K_L^0's. The lateral moment of the shower energy deposition [4] is defined as $\sum_{i=3}^{n} E_i r_i^2/(E_1 r_0^2 + E_2 r_0^2 + \sum_{i=3}^{n} E_i r_i^2)$, where the n crystals in the EMC cluster are ranked in descending order of deposited energy (E_i) , $r_0 = 5$ cm is the average distance between crystal centers, and r_i is the radial distance of crystal i from the cluster center. The second radial moment of the shower energy deposition is defined as $\sum_i E_i r_i^2$, where r_i is the radial distance of crystal i from the cluster center.

The photon control samples in the $1-5\,\mathrm{GeV}$ range are obtained from $e^+e^- \to \mu^+\mu^-\gamma$ events.

The K_L^0 control samples are obtained from $e^+e^- \to c\bar{c}$ events containing $D^{\star+} \to D^0(\to K_L^0\pi^+\pi^-)\pi^+$ decays, where charge conjugate modes are also implied. Events for the K_L^0 control sample are obtained by fitting the mass difference (Δm) between the $D^{\star+}$ and D^0 candidates, as shown in Fig. 5. The signal is parameterized with a double Gaussian function, and the background is parameterized with a polynomial function with an exponential damping term to account for kinematic constraints. The shower shapes of the K_L^0 candidates obtained from this fit to the data are compared with truth-matched distributions of K_L^0's obtained in simulation.

Figs. 6 and 7 show Geant4 versions 7.1 and 8.3 have better agreement with data than version 6.1, except for K_L^0's with lateral moment < 0.4 where none of the versions describe the data. The $\chi^2/d.o.f$ are 6.6, 2.7 and 2.2 between data and versions 6.1, 7.1 and 8.3 for photons and 35.5, 7.8 and 11.2 for K_L^0's for lateral moments, and 5.6, 3.0 and 2.4 between data and versions 6.1, 7.1 and 8.3 for photons and 45.6, 3.1 and 7.5 for K_L^0's for second moments, respectively.

4.5. K_L^0 yield in IFR vs EMC

The ratios of yields for K_L^0's reconstructed with IFR information only normalized to those reconstructed with the addition of EMC information in a sample of $B \to J/\psi K_L^0$ events are tabulated in Table 1. Although the ratios with Geant4 versions 7.1 and 8.3 are still smaller than those obtained from data, the improvment from inclusion of the Bertini cascade model in Geant4 version 7.1 is clearly visible in comparison with Geant4 version 6.1.

Table 1. Ratio of K_L^0 yields reconstructed only in IFR normalized to those reconstructed with additional EMC information for data and Geant4 simulation versions 6.1, 7.1 and 8.3.

Geant4 version 6.1 :	(0.51 ± 0.01)
Geant4 version 7.1 :	(0.64 ± 0.01)
Geant4 version 8.3 :	(0.64 ± 0.01)
BABAR data :	(0.74 ± 0.02)

5. Summary

The validation of recent Geant4 releases for possible use in the *BABAR* simulation has been more useful than originally expected. Significant progress in several areas has been made.

- **Electromagnetic shower shapes:** It has been a long-standing problem in *BABAR* that observed shower shapes in the EMC were somewhat broader that those predicted by Geant4. A strong feedback loop between *BABAR* and the Geant4 developers has resulted in significant improvement in this area, as seen in in Geant4 version 7.1 for the photon-induced shower shapes. More improvement was seen in version 8.3 in the simulation of multiple scattering, which is expected to further improve the simulation of the π^0 reconstruction efficiency.
- **Muon identification:** Improvements have also been seen in the agreement between muon efficiency data and simulation, as a result of improvements due to improved modelling of multiple scattering.
- **Bertini model for strange particles:** A clear improvement was seen in the simulation of K_L^0 interactions between Geant4 versions 6.1 and 7.1, as a result of inclusion of Bertini cascade models for kaons and lambdas.

These improvements, a successful migration from CLHEP 1.8 to CLHEP 1.9 and additional validation tests make it more likely that Geant4 version 8.3 will be used in the complete re-simulation of the entire *BABAR* dataset.

References

[1] S. Agostinelli *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **506**, 250 (2003).
[2] J. Allison *et al.*, IEEE Trans. Nucl. Sci. **53**, 270 (2006).
[3] B. Aubert *et al.* (*BABAR* Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **479**, 1 (2002).
[4] A. Drescher *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **237**, 464 (1985).

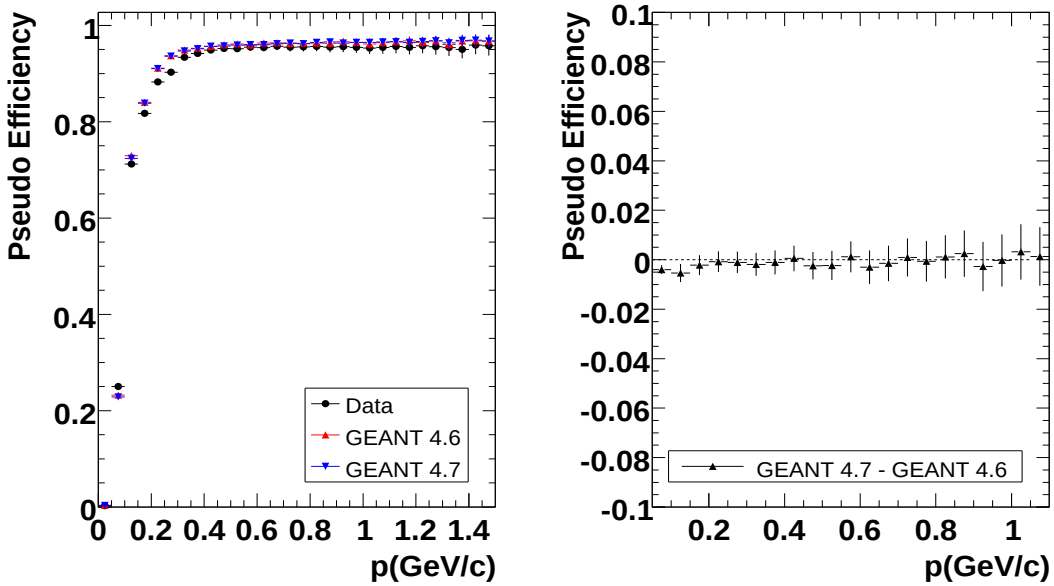

Figure 1. The left plot shows the pseudoefficiencies of tracks reconstructed in data (black dots) overlayed with Geant4 simulations with versions 6.1 (red upward triangles) and 7.1 (blue downward triangles). The right plot shows the difference of pseudoeffiencies around 100 MeV/c between Geant4 versions 6.1 and 7.1 as a result of the ionization bug-fix.

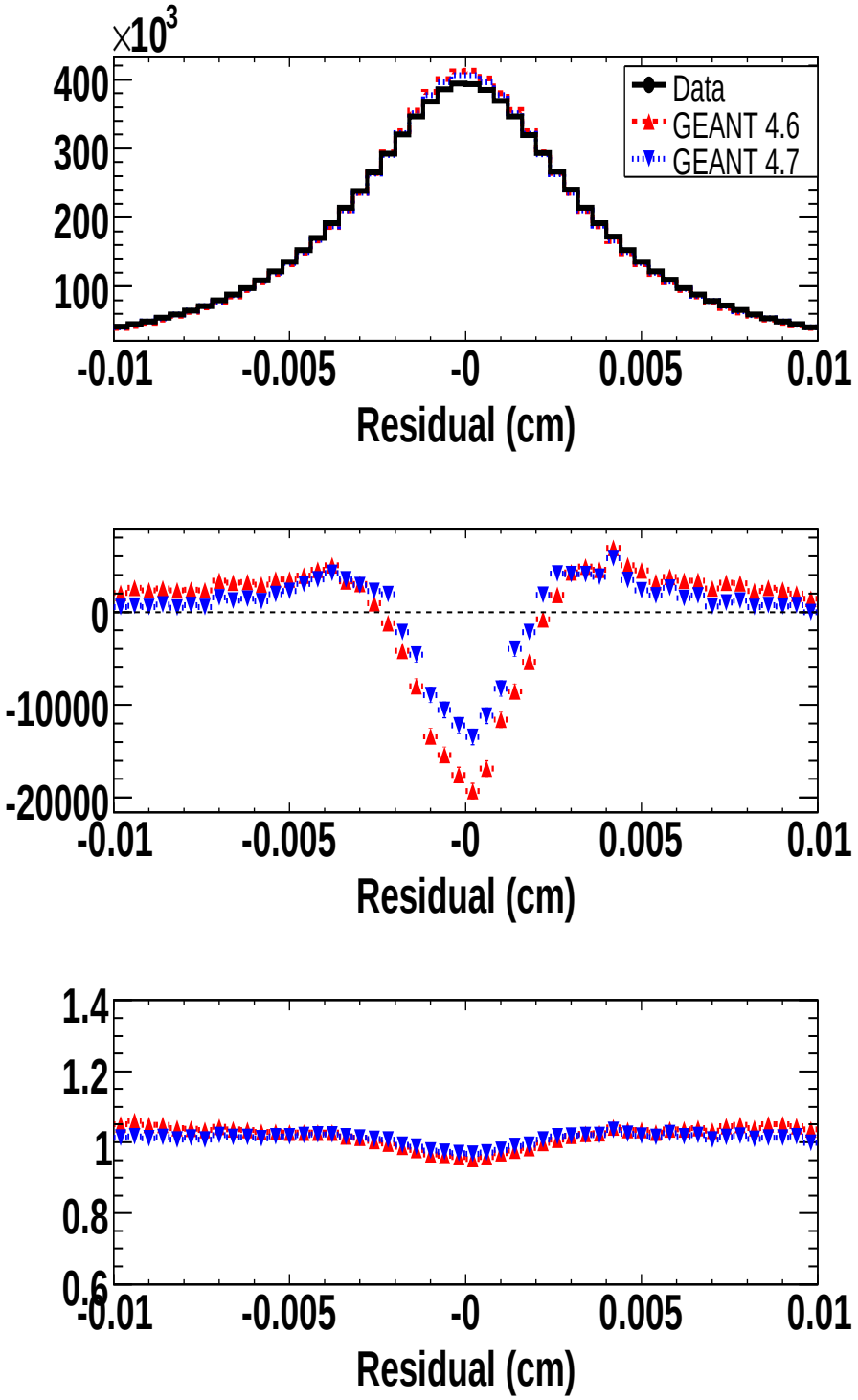

Figure 2. The top plot shows the distributions of SVT hit residuals in centimeters for tracks with momenta less than 1 GeV/c for data (black dots) and Geant4 simulation with versions 6.1 (red upward triangles) and 7.1 (blue downward triangles). The middle plot shows the differences between data and Geant4 simulation with version 6.1 (red upward triangles) and version 7.1 (blue downward triangles), and the bottom plot shows the ratio of data and different versions of Geant4 simulation.

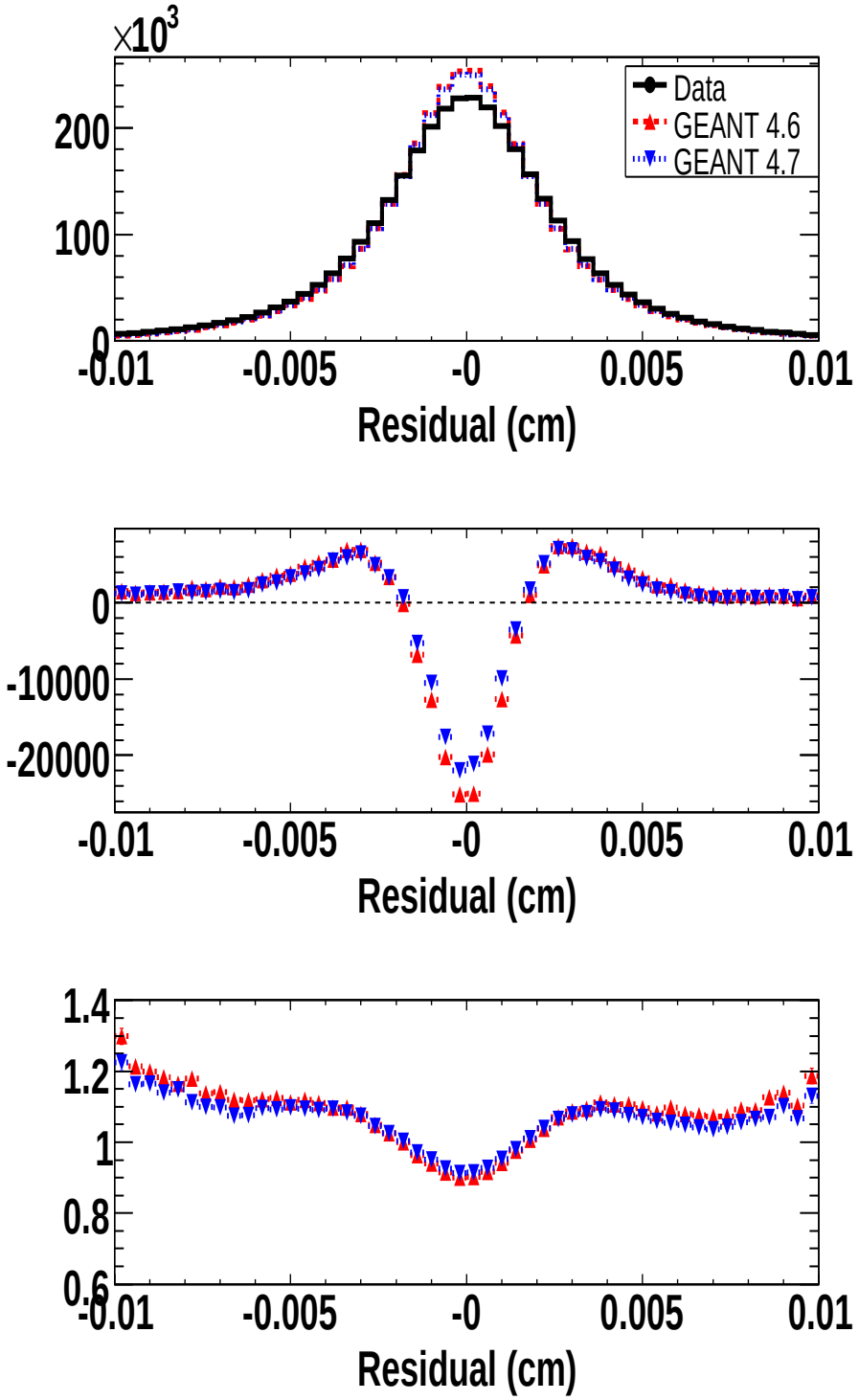

Figure 3. The top plot shows the distributions of SVT hit residuals in centimeters for tracks with momenta more than 1 GeV/c for data (black dots) and Geant4 simulation with versions 6.1 (red upward triangles) and 7.1 (blue downward triangles). The middle plot shows the differences between data and Geant4 simulation with version 6.1 (red upward triangles) and version 7.1 (blue downward triangles), and the bottom plot shows the ratio of data and different versions of Geant4 simulation.

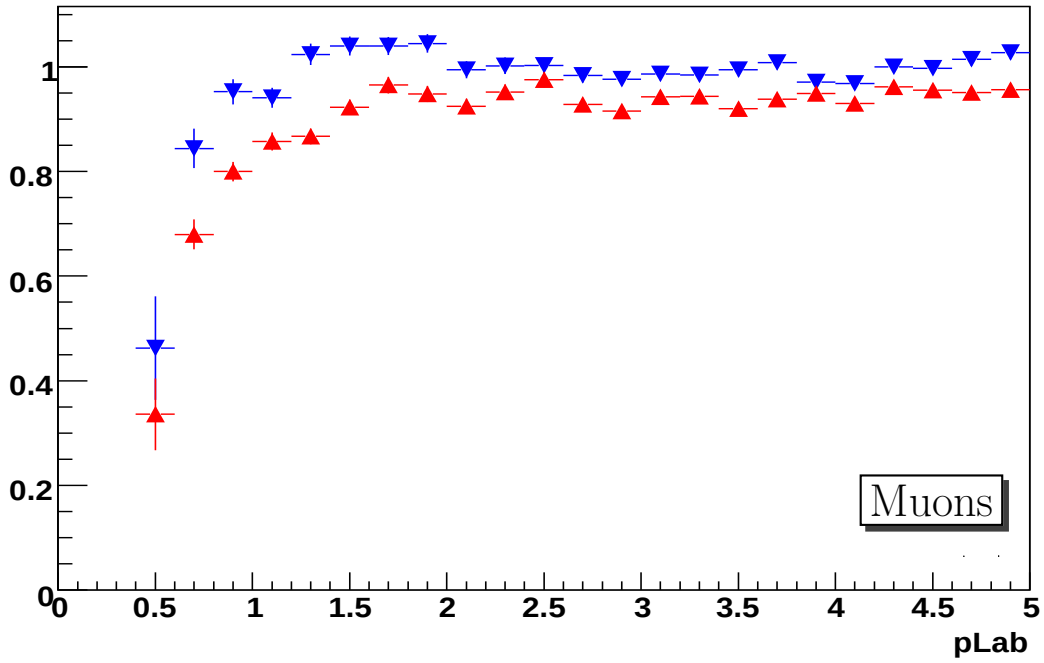

Figure 4. The data/simulation ratios of muon identification efficiencies with Geant4 version 6.1 (red filled upward triangles) and version 7.1 (blue filled downward triangles) as functions of the muon momentum in GeV/c.

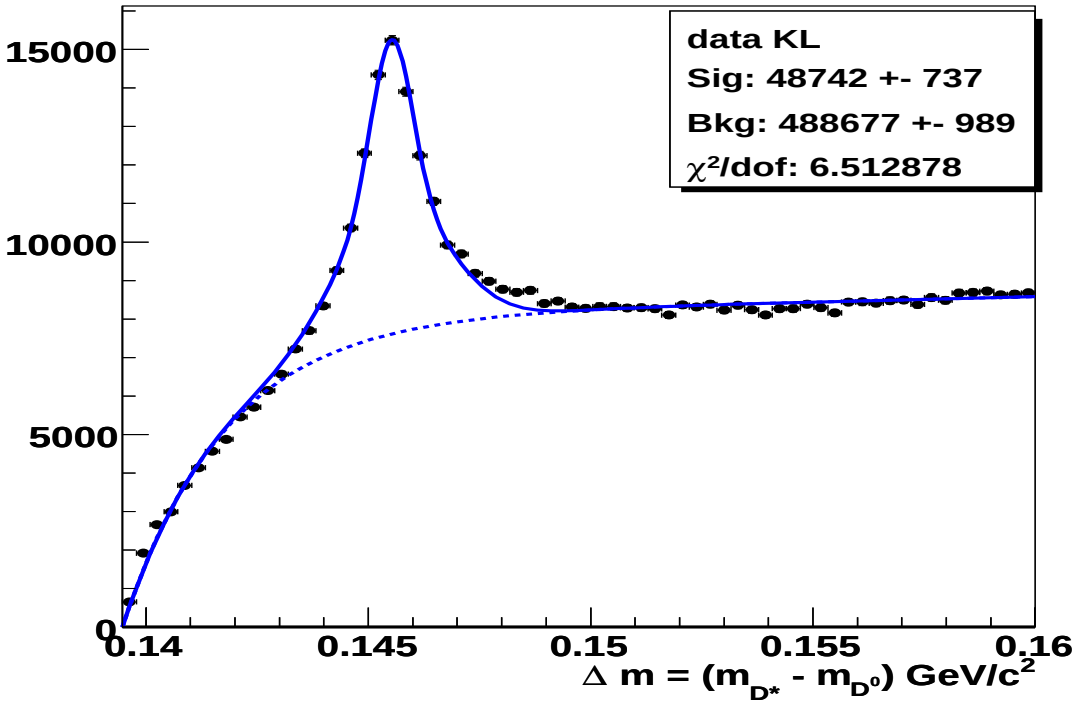

Figure 5. The data distribution of mass difference between $D^{\star+}$ and D^0 in $D^{\star+} \to D^0(\to K^0_L\pi^+\pi^-)\pi^+$ decays used as K^0_L control sample, shown along with the total fitted PDF (by solid line) and background component of the fitted PDF (by dotted line), as described in the text.

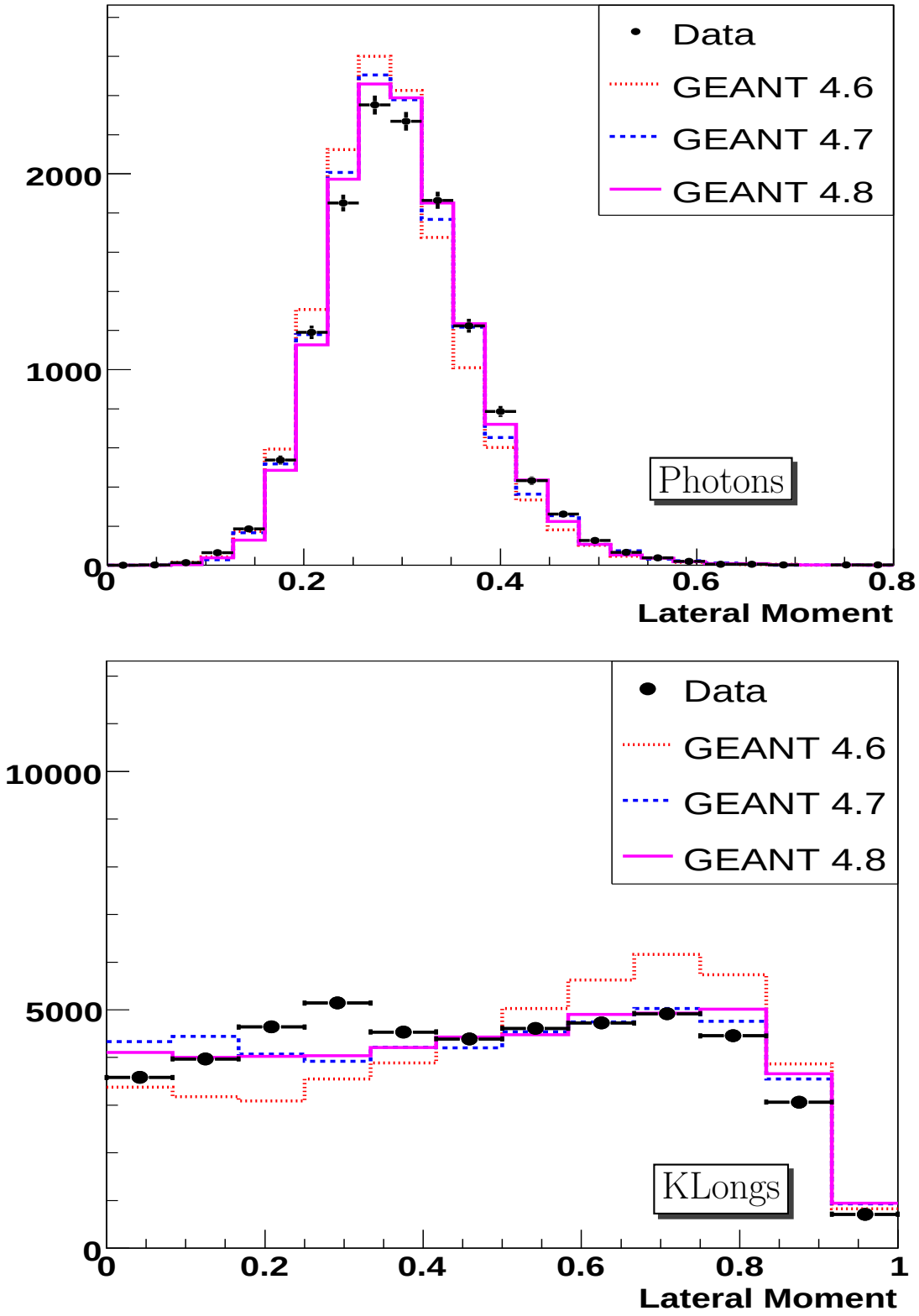

Figure 6. EMC lateral moment for photons (top) and K_L^0's (bottom) for Data (black dots), Geant 4.6 (red dotted line), Geant 4.7 (blue dashed line) and Geant 4.8 (magenta solid line).

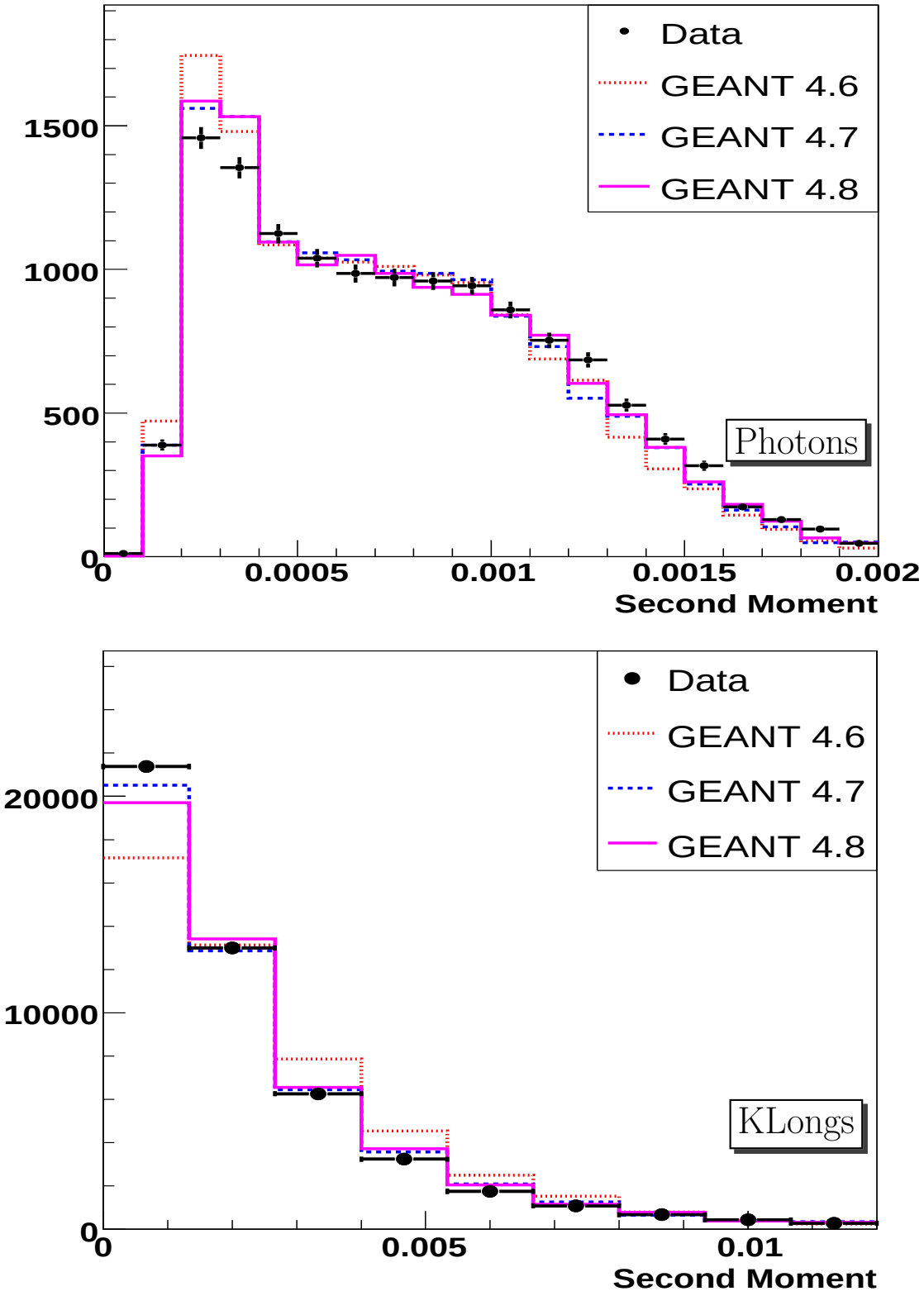

Figure 7. EMC second moment for photons (top) and K_L^0's (bottom) for Data (black dots), Geant 4.6 (red dotted line), Geant 4.7 (blue dashed line) and Geant 4.8 (magenta solid line).