CERN-LCGAPP-2004-10
July 26, 2004

Geant4 hadronic physics validation with LHC test-beam data: first conclusions

C. Alexa[a,c], J. Apostolakis[a], S. Banerjee[a,b], S. Constantinescu[c], A. De Roeck[a], S. Dita[c], A. Dotti[d], D. Elvira[e], F. Gianotti[a], A. Kiryunin[f,h], A. Lupi[d], C. Roda[d], D. Salihagic[f,i], P. Schacht[f], P. Strizenec[g], H.-P. Wellisch[a]

[a] *CERN, PH Division, Switzerland*
[b] *Tata Institute of Fundamental Research, Mumbai, India*
[c] *INFN, Particle Physics Department, Bucharest, Romania*
[d] *INFN, Sezione di Pisa, Italy*
[e] *Fermi National Accelerator Laboratory, USA*
[f] *Max-Planck-Institut für Physik, Werner-Heisenberg-Institut, Müenchen, Germany*
[g] *Institute of Experimental Physics, Kosice, Slovakia*
[h] *On leave of absence from Institute for High Energy Physics, Protvino, Russia*
[i] *On leave of absence from University of Montenegro, Podgorica, Serbia and Montenegro*

Abstract:
In the framework of the LCG Simulation Physics Validation Project, we present first conclusions about the validation of the Geant4 hadronic physics lists based on comparisons with test-beam data collected with three LHC calorimeters: the ATLAS Tilecal, the ATLAS liquid-argon hadronic end-cap, and the CMS HCAL. An effort has been made to ensure a consistent approach for the three detectors.
Outstanding issues which require additional work, in particular on the simulation side, are also discussed.

1. Introduction

An accurate description of hadronic interactions and hadronic showers over a very broad energy range is the most challenging part of any package which will be used to simulate the LHC experimental environment.

Validation of the Geant4 and Fluka hadronic physics is being performed, in the framework of the LCG Simulation Physics Validation Project [1], through two complementary approaches: comparison with test-beam data collected with LHC calorimeter modules or prototypes addresses global shower features, which are particularly relevant to calorimetric measurements; in addition, the so-called "simple benchmark" studies, where the simulation is compared to simple experiments performed at nuclear facilities, provide stringent tests of individual microscopic collisions, like those occurring in the particle interactions with the thin layers of the LHC tracking detectors, and are relevant for both calorimeters and tracking detectors.

In this note we report first conclusions about the validation of Geant4 hadronic physics, based on comparisons of the simulation with test-beam data collected with three LHC calorimeters: the ATLAS Fe-scintillator Tilecal, the ATLAS Cu-liquid argon hadronic end-cap, and the CMS brass-scintillator HCAL. The results shown here have been obtained from detailed simulations of the test-beam set-up of the three detectors, from the beam-line layout to the features of the readout chains. The simulated data were produced with two Geant4 (version 5.2p02) physics lists: QGSP 2.7 and LHEP 3.6. Simulation and data were compared mainly for three crucial observables: the pion energy resolution, the e/π ratio, and the pion longitudinal and lateral shower profiles.

An effort has been made to follow, as much as possible, a coherent approach for the three calorimeters. Indeed, a consistent study across detectors which are based on different technologies and/or materials should provide more complete and robust results than those obtained by the individual experiments.

Based on the outcome of these studies, we have tried to draw some preliminary LHC-wide conclusions about the adequacy of Geant4 hadronic physics. This has been done by comparing the simulation validation results against the quantitative physics requirements collected in Ref. [2]. Aspects which necessitate more studies, both on the data and simulation sides, are also mentioned, and the strategy for future validation work is outlined.

This note is organized as follows. Section 2 presents the main features of the two physics lists used in these studies. Section 3 describes the test-beam set-up for the three calorimeters and the analysis procedures. Comparison results between data and simulation are shown in Section 4. Finally Section 5 is devoted to some conclusions and to the discussion of outstanding issues that require additional work.

2. Description of the used physics lists

Geant4 [3] comes with a very rich collection of modeling possibilities for hadronic physics. It covers the full spectrum of interactions of hadrons and leptons with matter, from thermal neutron transport to multi-TeV events at LHC, and to higher energies for cosmic-ray applications. Geant4 hadronic physics (GHAD) is a tool-kit where each component represents a particular piece of physics. Hence a physics simulation engine needs to be assembled from the components in the tool-kit, in order to provide a complete and consistent description of the hadronic, nuclear, and neutron physics relevant to the application. This is a non-trivial task, and hence the Hadronic Physics Working Group (HPWG [4]) provides a set of educated guesses for frequently occurring use-cases involving hadronic interactions. At present time, 19 simulation engines are supported for 24 different use-cases. These are released with respect to particular versions of Geant4, and versioned, to allow inter-comparison of results stemming from different studies.

The physics simulation engines validated in the course of the present study are those recommended for calorimeter studies by the HPWG. Four physics lists are suggested for this use-case: LHEP, QGSP, QGSC and FTFP. These differ mainly in the physics of the strongest determining factor in calorimeter simulations: the initial reaction of the primary hadron inducing the shower.

In the case of LHEP, parameterization-driven modeling is used for all hadronic interactions. LHEP uses the most CPU-effective hadronic modeling available in GHAD, namely the LEP and HEP parameterized models for inelastic scattering. This type of modeling parameterizes features of the final states of individual inelastic reactions, like multiplicities and inelasticity. As a result it is very fast, but resonances are not produced, and the secondary angular distributions for O(100 MeV) reactions cannot be described in detail. The total cross-sections, used to calculate the location of individual reaction sites, are inherited from Geant3.21 [5].

The second physics list, QGSP, uses theory-driven modeling for the reactions of energetic pions, kaons, and nucleons. It employs the quark-gluon string model (QGS, [6]) for the soft and fast 'punch-through' interactions of the projectile with nucleons from the nuclear medium. The string excitation cross-sections are calculated in quasi-eikonal approximation. Concurrency of reactions is assumed, resulting in a delta-function in the correlated sampling of the parton momenta, to enforce energy conservation across all reactions in the case of multiple participants. Minimum and maximum string-mass constraints are introduced. This enables the use of the QGS model at relatively low center-of-mass energies (without Kopylov's algorithm). A pre-equilibrium decay model [7] with an extensive evaporation phase is employed to describe the behavior of the damaged nucleus. QGSP uses Barashenkov's pion cross-sections [8], and the Wellisch-Axen systematics for nucleon-induced reactions [9], currently also in use in the MCNP code system [10].

In the other two simulation engines, the theoretical parts of the modeling are varied with respect to QGSP. In QGSC, the intra-nuclear transport after the initial soft reactions is changed from a pre-compound model to chiral invariant phase-space decay [11] (multi-Quasmon fragmentation). In FTFP, the string excitation in the quark-gluon string model is replaced by a diffractive Ansatz, similar to what is used in FRITJOF [12], and the Lund fragmentation functions replace the softer fragmentation functions needed for the dressed quark-partons of the quark-gluon string model.

In all the above lists, at energies below a few GeV, and for inelastic interactions of strange baryons and anti-baryons, the most CPU-effective models available are used. It is worth noting that here low-energy final states are simulated only approximately, and some effects, like low-energy pion absorption, are entirely neglected.

For the purpose of this study, two physics lists out of the four described above were selected: LHEP and QGSP.

3. Test-beam set-up and analysis procedure

In this Section we describe briefly the detector and test-beam layouts for all three calorimeters. We also discuss the reconstruction procedures used to extract the incident particle energy from the raw data.

Details about the simulation parameters used in each experiment are also given. The data are compared to Geant4 and, in some cases, also to Geant3. However, since the different experiments do not use the same Geant3 hadronic packages, simulation results obtained with Geant3 do not necessarily compare to each other.

3.1 ATLAS Tilecal

The Tile Calorimeter [13] is a sampling device made of steel and scintillating tiles as absorber and active materials respectively. The detector consists of a cylindrical structure with an inner radius of 2280 mm and an outer radius of 4230 mm. It is subdivided into a 5640 mm long central barrel and two 2910 mm long extended barrels. Each cylinder is made of 64 ϕ-modules.

The scintillating tiles lie in the $r - \phi$ plane and span the width of the module in the ϕ direction. Wavelength shifting (WLS) fibers running radially collect the light from the tiles along their two open edges. Readout cells are then defined by grouping together a set of fibers into a photomultiplier (PMT), to obtain a three dimensional segmentation.

Radially, the calorimeter is segmented into three layers, approximately 1.4, 3.9 and 1.8 interaction lengths (λ) thick at $\eta = 0$. The $\Delta\eta \times \Delta\phi$ segmentation is 0.1×0.1 in the first two layers and 0.2×0.1 in the last radial layer.

Most of the Tile calorimeter measurements were performed in the CERN SPS H8 beam line. The latter was instrumented in a standard fashion, with one threshold Cherenkov counter used for beam particles separation, a pair of delay-line wire chambers each one providing the beam coordinates in the plane perpendicular to the beam, and scintillator counters to define a trigger. The scintillators selected a beam spot of size 2-3 cm, while the chambers allowed the reconstruction of the particle impact point on the calorimeter face to better than 1 mm.

The measurements reported in this note were obtained in 2002, when two extended barrel modules, one barrel module and the barrel module zero (see Figure 1) were exposed to pions and electrons between 20 and 180 GeV. Also shown are results based on low-energy pion data (1, 2, 3, 5 and 9 GeV) taken in 2003 with a set-up consisting of module zero and two barrel modules.

The analysis of the data started by selecting pions, electrons or protons using the Cherenkov counter, which provides electron separation from pions and protons up to $\sim$ 50 GeV, and pion-proton separation up to 200 GeV. For about half of the energy points

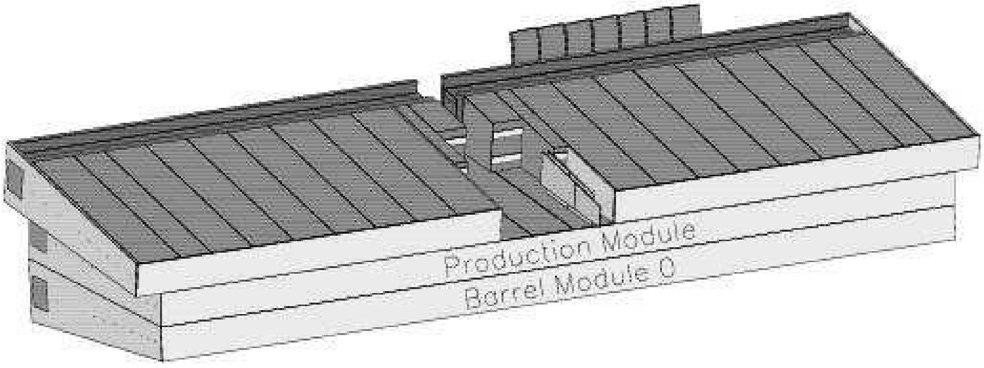

Figure 1: Schematic view of the 2002 ATLAS Tilecal test-beam set-up.

the Cherenkov information was insufficient or unavailable. In these cases particles were identified on the base of their longitudinal shower profile, which also removed the muon contamination in the beam.

The two PMT of each calorimeter cell were equalized with a cesium source. The equalization uniformity over a whole module was within 1%. In each event the PMT signals were sampled in nine time slices at 40 MHz. The first time slice defined, for each event and channel, the pedestal level, which was directly subtracted from the energy calculation. The total energy released by a particle was obtained by summing all the PMT signals. For each geometry and beam energy, the energy scale was set using electrons.

A Geant4-based simulation package which describes in full detail the detector modules and the test-beam layout has been developed [14]. In the simulation, the range cut was set to 1 mm. A study of the electron energy response as a function of this cut was made [15] and no dependence was found over the region 1-$10^4 \mu m$. Statistics of 5000 events per energy point and per pseudo-rapidity (η) value were produced. The starting point for energy reconstruction was the total energy released by a shower in the active material, which was calibrated to the incident energy using electrons as for the real data. The contributions of the electronic noise (pedestal smearing) and photostatistics were measured with data and included in the simulated samples at the analysis level. These effects are particularly relevant for particle energies below 20 GeV. The effect of the beam smearing was also included in the simulation.

Systematic uncertainties arised mainly from the knowledge of the beam contamination and of the calibration linearity. If the information of the Cherenkov counter was not available, as it was the case for data samples collected at high energies and a few samples at low energies (3, 5 GeV), electrons and pions were selected on the base of their longitudinal shower shape, as mentioned above. This procedure may introduce geometry- and/or energy-dependent biases. The corresponding systematic uncertainties were obtained by comparing results from different analyses. The systematic error related to the calibration non-linearity was deduced from dedicated calibration runs. The overall systematic uncertainty on the e/π ratio was preliminarily estimated to be about 8% in the worst case (more detailed studies are being carried out).

3.2 ATLAS HEC

The Hadronic End-Cap Calorimeter (HEC) [16] is a sampling device with flat copper absorber plates interleaved with liquid-argon gaps. The HEC is structured in two wheels (HEC1 and HEC2). Each wheel consists of 32 azimuthal modules, has an outer diameter of about 4 m, and a depth of 0.82 m (HEC1) or 0.96 m (HEC2). The thickness of the copper plates is 25 mm for HEC1 and 50 mm for HEC2. Annular spacers define 8.5 mm gaps between the absorber plates. The liquid-argon gaps are instrumented with a readout structure based on the principle of the electrostatic transformer (EST). The transverse granularity of the calorimeter is $\Delta\eta\times\Delta\varphi = 0.1\times2\pi/64$ $(0.2\times2\pi/32)$ for $|\eta| < 2.5$ $(|\eta| > 2.5)$. There are four longitudinal compartments, and the total detector thickness is $\sim$103 X_0 and > 10 λ.

During the years 2000-2001, the production calorimeter modules were exposed to test beams, as a part of the standard quality control procedure during HEC construction. The beam tests have been carried out in the H6 beam line of the CERN SPS, using secondary and tertiary beams. The beam line was instrumented with scintillator counters and scintillator walls for trigger purposes, a Cherenkov counter which provided electron/pion separation up to 80 GeV, and four multi-wire proportional chambers for beam position and angle determination. More than 2200 runs were recorded, each run typically consisting of 20000 events, over energy ranges 6-150 GeV for electrons, 10-200 GeV for charged pions and 120, 150, 180 GeV for muons. The collected data were processed using a digital filtering technique for the signal reconstruction. Electrons were discriminated from pions by using the Cherenkov information, where available, and/or on the base of the energy depositions in the rear HEC modules (third and fourth longitudinal compartments). All details, as well as data analysis results, can be found in Ref. [17].

Comparisons of experimental data with detailed Monte Carlo simulations have been performed. A dedicated program [18] based on the Geant3 package was developed and routinely used. New code [19] was written for the Geant4-based simulation of the HEC test-beam. The geometry of the test-beam set-up in the Geant4 and Geant3 simulations is identical, which allows for direct comparisons with least biases due to possible differences in the geometry. The geometry description is very detailed, and includes not only the copper plates and liquid-argon gaps, but also the polyimide electrodes and copper pads. The cryostat and all beam elements, such as multi-wire proportional chambers and scintillator counters (distributed over a distance of 30 m), are included as well. The simulation also contains virtual leakage detectors, placed laterally and longitudinally around the actual HEC modules, to estimate the energy leakage. The use of these leakage detectors also speeds up the simulation, as all particles entering them are immediately stopped, and their energy is deposited on the spot.

In the Geant4 simulations of the HEC test-beam the range cut was set to 20 μm. This value was obtained by comparing the largest signals in one HEC cell observed for 100 GeV electron data with the signals predicted by Geant4 for different range cuts [20]. For the Geant3 simulations, version 3.21 and the GCALOR hadronic shower package were used. Particle tracking thresholds were set to 100 keV and secondary production cuts for photons

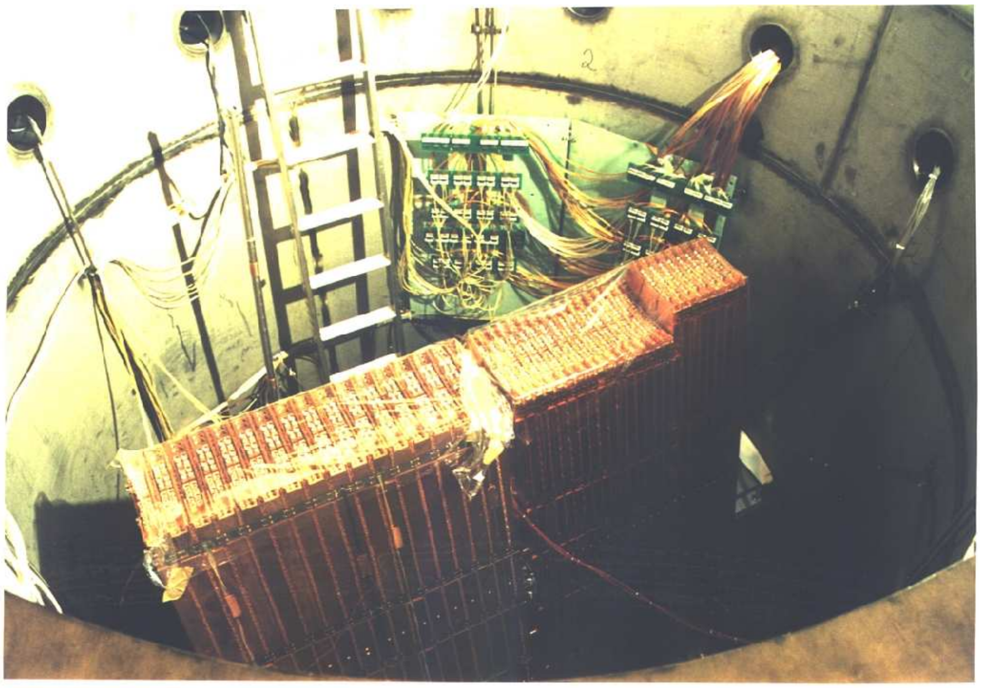

Figure 2: ATLAS HEC test-beam set-up.

and electrons were set to 1 MeV.

Simulations of the response of test-beam modules to negative charged pion beams of different energies (from 10 to 200 GeV) and to electron beams (6-147.8 GeV) for one selected impact point were carried out. The typical statistics was 5000 events per energy point. The analysis of the simulated samples followed rather closely the experimental procedure [17]. Prior to any reconstruction, the visible energy in the rear modules (HEC2 wheel) was multiplied by a factor of two to take into account the difference in sampling fractions between the two HEC wheels. For the particle energy reconstruction, clusters of fixed size were used, i.e. the same set of HEC cells was selected for energy measurements in all events and for all beam energies at a given impact point. Applying the experimental procedure of cluster-size selection to simulated pion data, one obtained the following results: 54, 46 and 62 cells for LHEP, QGSP and Geant3, respectively. The experimental value is 51 cells. The use of clusters of fixed size for the energy reconstruction allowed evaluation of the electronics noise (including possible correlations between HEC cells). To compare experimental and simulated data, the noise was subtracted from the experimental data.

Systematic uncertainties in the experimental results have been preliminarly evaluated. The systematic error on the HEC energy scale was estimated to be $\pm 2\%$.

3.3 CMS HCAL

The CMS central hadron calorimeter (HCAL) [21] is a sampling device divided into one barrel and two end-cap parts. It consists of 18 layers of plastic scintillators interleaved with brass absorbers. Scintillation light is collected using WLS fibers, read out with hybrid photodiodes. The layers are grouped in towers of granularity $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$.

The HCAL has been tested with two different prototypes and set-up in 1996 and 2002.

In 1996, a set-up consisting of a prototype electromagnetic calorimeter (ECAL) and a prototype hadron calorimeter was tested in the CERN H2 beam line [22]. Data were taken using beams of muons, electrons and hadrons, with energies from 10 to 300 GeV. The HCAL prototype was based on the hanging file structure, in which copper alloy absorber plates varying in thickness from 2 to 8 cm were interposed with scintillator tiles equipped with WLS fibers. Each scintillator tile was read out independently with a photomultiplier tube located at the end of the optical cable. This allowed detailed studies of the shower longitudinal profile (see Section 4.3). The total depth of the prototype corresponded to 10.1 λ, and the transverse size of the plastic scintillator tiles was 64 cm $\times$ 64 cm. A schematic view of the setup is shown in Figure 3.

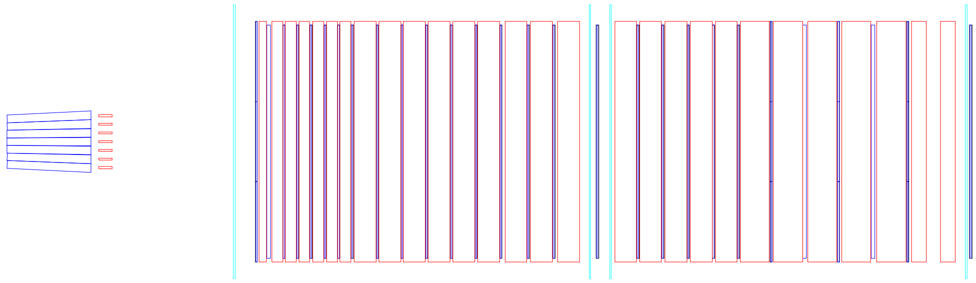

Figure 3: A schematic view of the 1996 CMS HCAL test-beam set-up.

The prototype ECAL detector consisted of a matrix of 7×7 $PbWO_4$ crystals, each 23 cm long (25.8 X_0, 1.1 λ) and 2 cm $\times$ 2 cm transverse size. The ECAL module was placed approximately 50 cm upstream of the HCAL prototype. The crystals were read out by Avalanche Photo Diodes with a gain of about 50.

The combined calorimeter system was put inside a magnetic field of maximum strength 3T. The magnet was oriented in such a way that the lines of the B-field were parallel to the scintillator planes, as it is in the CMS barrel configuration. In the first five interaction lengths of the HCAL, the magnetic field was uniform to within 10%. Data were taken with different settings of the field value (including no field) and with two detector configurations (with and without the ECAL in front). Here only results obtained from the HCAL standalone test are reported.

The HCAL was calibrated by equalizing the response of each layer to 225 GeV muons. This method provided an accuracy of $\approx 3\%$, and was independently checked with radioactive sources.

The data were compared with simulation results based on Geant3 (version 3.21, interfaced to the GHEISHA hadronic package) and Geant4. Details of the sensitive and absorber layers were described, but the beam-line detectors were not included in the simulation. In Geant4, a cutoff of 700 μm was used for all particles. In the Geant3 simulation, 100 keV thresholds were chosen for photons, electrons, muons and charged hadrons, whereas the cutoff for neutrons was set at 10 keV. Non-uniformities in light collection in the crystals were simulated using light collection curves obtained from measurements. The electronic noise was studied with random triggers in the data and was added to individual channels in the simulation.

The simulated and experimental data were reconstructed in the same way. Each HCAL channel was calibrated with muons. Therefore, for a configuration with HCAL alone, the energy was obtained by (i) converting energy depositions to number of MIPs; (ii) weighing the MIPs in each layer with the absorber thickness in front; (iii) normalizing the total number of weighted MIPs to the nominal beam energy using 100 GeV pions.

Systematic errors in the measurements arised from calibration uncertainties and beam contamination. The calibration uncertainty was studied using different calibration procedures (muons versus radioactive source). Beam contamination uncertainties were studied by changing the selection cuts in the data sample. The overall uncertainty on the pion energy scale is at the level of 5%.

In 2002, another beam test of the HCAL was performed (TB02). The experimental set-up (Figure 4) consisted of an electromagnetic calorimeter, two hadronic barrel wedges (HB), and two Outer Hadronic layers (HO). A moving table was used to change the relative direction of the beam with respect to the detector, thus allowing tests of the different HB towers as a function of ϕ and η. The beam line was equipped with two wire chambers and two scintillator counters to trigger the events and monitor the beam position. The detector included a mock-up of the ECAL, consisting of an aluminum box containing a 7×7 $PbWO_4$ crystal matrix. The ECAL box was followed by two HB wedges positioned together as in the real CMS set-up. A 310 mm thick aluminum slab accounted for the solenoid material, and preceded the outer hadronic calorimeter. The HO consisted of two (one) layers of scintillator in the central (outer) ring, with a 195 mm thick iron plate in-between acting as a tail catcher. For the HB and HO, the light from the scintillators was transported to Hybrid Photo Diodes through clear fibers. The signal was then processed by a Charge Integration Electronics which gave a short output pulse. The two HB wedges used for the

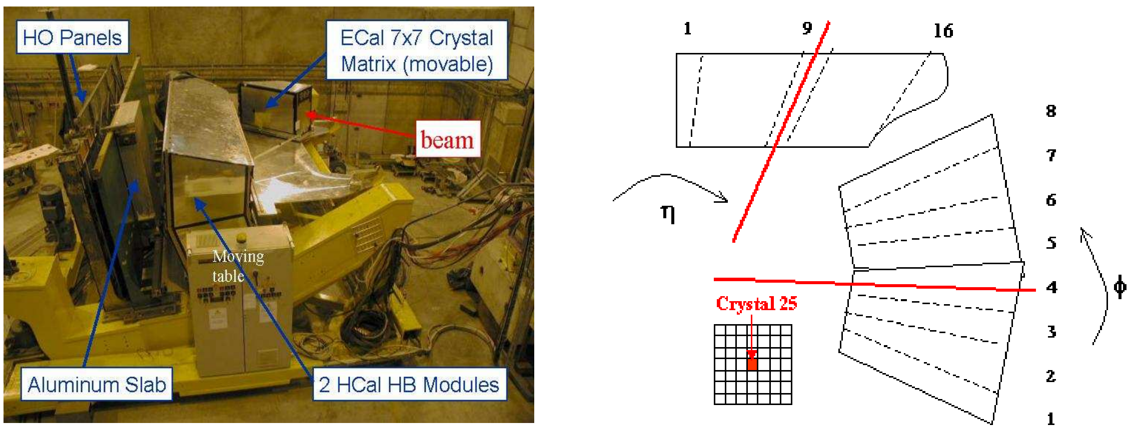

Figure 4: Left: set-up of the 2002 CMS HCAL test-beam. Right: position of tower $\eta = 9$, $\phi = 4$.

TB02 test were production modules for the actual CMS detector. The thickness of the brass layers 2-16 was 6 cm for layers 2-10 and 6.6 cm for layers 11-16. The steel layers, 1 and 17, were 7.45 cm and 8.9 cm thick respectively. The thickness of the scintillation tiles was 0.4 cm in layers 2-16, and 0.8 cm in layers 1 and 17. The HB modules were segmented into 0.087×0.087 towers in $\eta \times \phi$.

The energy scale of the ECAL matrix was derived using a 100 GeV electron beam, whereas the HB energy scale was obtained in several steps. First, a Co^{60} source was used to inter-calibrate the individual scintillator tiles. Second, a beam of 100 GeV electrons was directed onto a given HB tower to estimate the ADC counts-to-GeV conversion factor. In the third step, a beam of 50 GeV pions depositing little energy in the ECAL (less than 2 GeV) was used to derive the final energy scale.

The collected data were used to measure the pion energy resolution and response in the 20-300 GeV energy range. In order to validate the physics of the Geant4 simulation, samples of 2000-10000 events were generated in the same range. The simulation followed the detector configuration, geometry and calibration very closely, including the effect of electronic noise in the HB, added as a Gaussian term with a σ=520 GeV centered about the total energy deposited in a tower. As in the TB02 experiment, the simulated pion beam was generated in the direction of the $\eta = 9$, $\phi = 4$ tower, and passed through the central crystal of the ECAL matrix. The Geant4 production range cuts for secondary particles were set to 1 mm for electrons and positrons and 10 mm for photons, except in the ECAL where the later cut was set to 1 mm.

The pion energy was reconstructed, in the same way in the data and in the simulation, as the sum of the energy measured in the ECAL and HB detectors: $E_{\pi^-} = a \times E_{ECAL5\times5} + b \times E_{HB5\times5}$, where a, b are the calibration factors, $E_{ECAL5\times5}$ is the total energy deposited in a 5×5 crystal matrix around the central crystal, and $E_{HB5\times5}$ is the total energy deposited in a 5×5 HB tower matrix around the central tower.

Most of the results presented in this note were obtained from the 1996 test-beam data, with the exception of some results discussed in Section 4.1.3. Therefore, figures and numbers in the following Sections refer by default to these data, unless otherwise stated.

4. Results

Comparisons between test-beam data from three LHC calorimeters and simulation packages (two Geant4 physics lists and Geant3) are shown in the following three Sections for three important observables: the pion energy resolution, the e/π ratio and the pion shower profile. Some general remarks are presented in Section 4.4.

4.1 Pion energy resolution

4.1.1 ATLAS Tilecal

The energy resolution σ/E for pions was studied in the energy range from 3 to 180 GeV and at angles of incidence corresponding to η values from 0.25 to 0.65. The resolution σ and the peak energy E were obtained from Gaussian fits to the reconstructed energy spectra. The pion energy resolution was fitted with the expression:

$$\frac{\sigma}{E} = \frac{a}{\sqrt{E_{beam}}} + b \tag{4.1}$$

where E_{beam} is the beam energy, a is the sampling term and b is the constant term.

The results are presented in Figure 5 as a function of the beam energy for the data and for the two Geant4 physics lists, LHEP 3.6 and QGS P2.7. Figure 6 shows the ratio of energy resolutions between the simulation and the test-beam data. The agreement Geant4-data is at the level of $\sim \pm 12\%$ at $\eta = -0.35$.

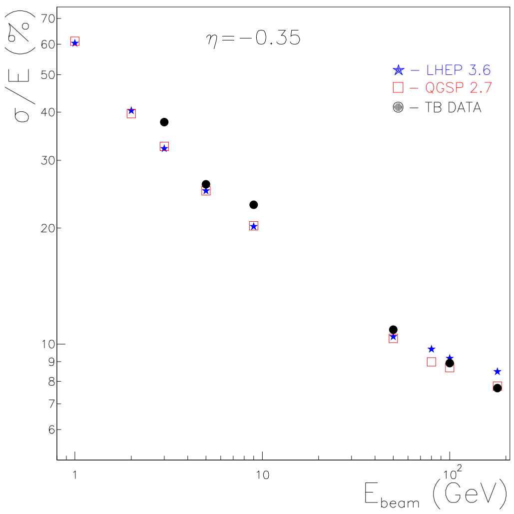

Figure 5: Pion energy resolution in the ATLAS Tilecal as a function of beam energy at $\eta = -0.35$.

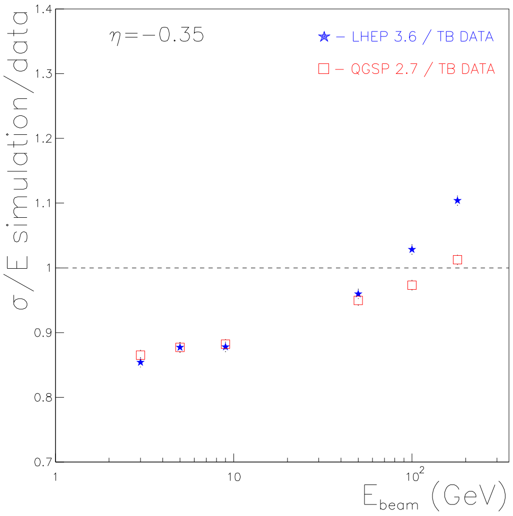

Figure 6: Ratio between simulation and data for the pion energy resolution in ATLAS Tilecal at $\eta = -0.35$.

Table 1 summarizes the results of the fits obtained at various pseudo-rapidity values for both test-beam data and simulations. The Geant4 predictions are in good agreement with the data, whereas Geant3 GCALOR gives a too small constant term.

	QGSP 2.7		**LHEP 3.6**		**GCALOR**		**Data**	
η	a	b	a	b	a	b	a	b
0.25	41.3	6.27	33.1	6.45	36.0	3.35	44.2	4.72
0.35	38.1	4.87	31.2	5.89	41.8	2.02	48.0	4.11
0.45	40.7	4.14	35.1	5.17	41.9	1.84	53.0	3.42
0.55	36.9	4.43	41.5	4.19	41.2	1.64	48.2	3.54
0.65	38.0	4.04	35.7	4.57	48.9	0.94	30.1	5.83

Table 1: Sampling ($a[\%\sqrt{\text{GeV}}]$) and constant ($b[\%\sqrt{\text{GeV}}]$) terms of the pion energy resolution in Tilecal as a function of rapidity η, as obtained from simulations and test-beam data. The errors of the fits to the simulated and experimental data are included in the last digit.

4.1.2 ATLAS HEC

To evaluate the energy resolution for pions, Gaussian curves were fitted to the reconstructed energy distributions. The peak value E and the standard deviation σ from these fits were used to calculate the resolution $R = \sigma/E$. The energy dependence of the resolution was then fitted with the formula:

$$\frac{\sigma}{E} = \frac{a}{\sqrt{E_{beam}}} \oplus b \tag{4.2}$$

The energy resolution as a function of the beam energy is presented in Figure 7 for the data and for the Geant4 and Geant3 simulations, together with results of the fit to Equation 4.2. The resulting sampling and constant terms are given in Table 2. Ratios of the simulation to data for the energy resolution are shown in Figure 8. Both Geant4 lists describe the pion energy resolution quite well (QGSP is slightly better). Geant3 predicts a systematically too good energy resolution.

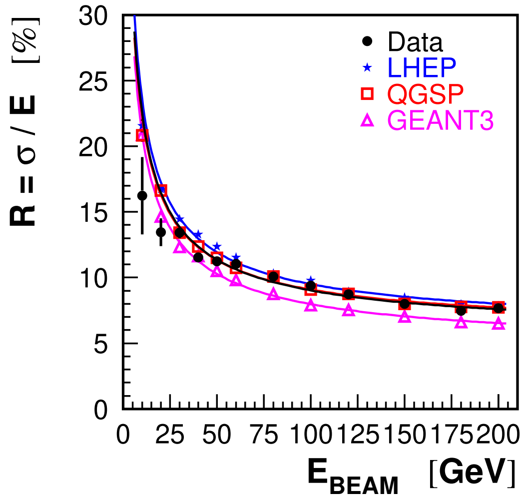

Figure 7: Energy resolution for pions in the ATLAS HEC as a function of the beam energy. The lines indicate the results of fits to Equation 4.2.

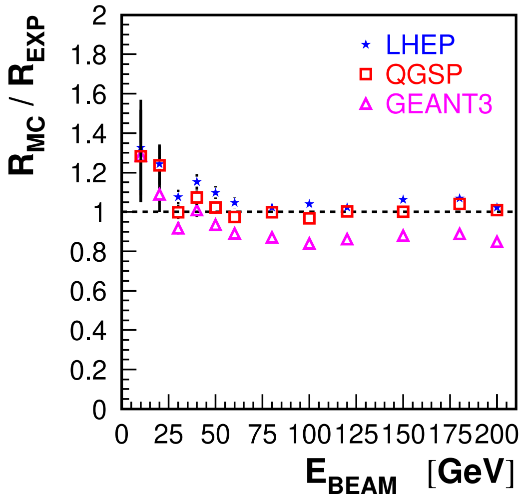

Figure 8: Ratio between simulation and data for the pion energy resolution in the ATLAS HEC.

	Data	LHEP 3.6	QGSP 2.7	GCALOR
Sampling term a [$\%\sqrt{\text{GeV}}$]	69±1	72.0±0.8	67.6±0.7	64.7±0.6
Constant term b [$\%\sqrt{\text{GeV}}$]	5.8±0.1	6.2±0.1	6.1±0.1	4.7±0.1

Table 2: Sampling and constant terms of the pion energy resolution in ATLAS HEC, as obtained from simulations and test-beam data.

4.1.3 CMS HCAL

Figure 9 shows a comparison of the energy spectra for 100 GeV pions in the 1996 HCAL standalone set-up between data and various physics lists. The two Geant4 physics lists

LHEP and QGSP are in good agreement with the data while Geant3 predicts a somewhat broader distribution. The RMS spread for data, LHEP, QGSP and Geant3 is 9.1 ± 0.1 GeV, 9.9 ± 0.1 GeV, 9.3 ± 0.1 GeV and 10.8 ± 0.1 GeV respectively, where the errors are statistical only.

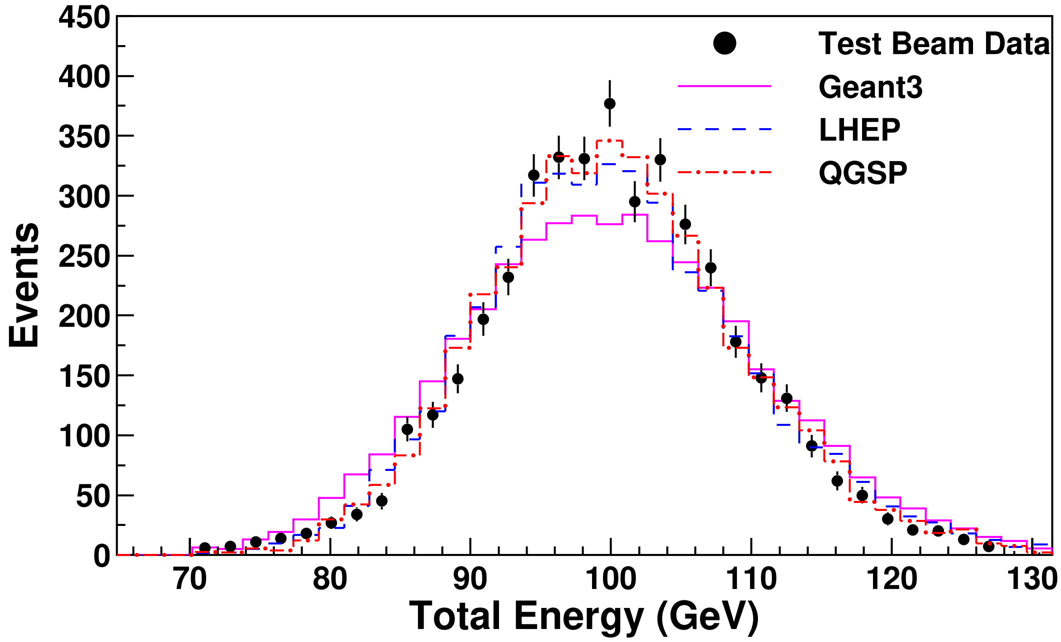

Figure 9: Energy measured in the CMS HCAL 1996 standalone test-beam set-up with 100 GeV pions. The experimental data are compared to various simulation packages.

Figure 10 shows the ratios of the mean measured energies to the nominal beam energies as a function of beam energy. A non-linear behaviour is observed, particularly at low energies. This effect, due to $e/h > 1$, is also visible in different simulation models. The ratio of the energy response between simulation and data is shown in Figure 11. LHEP and QGSP are consistent with the data within 2% over the entire energy range. The Geant3 prediction shows a larger disagreement but is still within the systematic uncertainties of these measurements.

The measured pion energy resolution as a function of the beam energy is presented in Figure 12. The data show a significantly better resolution than all simulations at low energies, which is due to a smaller high energy tail in the data than in the models. Ratios between simulation and data for the pion energy resolution are plotted in Figure 13 as a function of beam energy. Both Geant4 physics lists provide a good description at high energies, with QGSP doing a better job (agreement better than 5%). The agreement at energies below 30 GeV is only at the level of 12%. Finally, in Table 3 we present the sampling and the constant terms of the pion energy resolution obtained by a fit with Equation 4.1.

Results obtained with the TB02 set-up are shown in Figure 14, and compared to the two physics lists. The vertical bars represent the experimental statistical and systematic

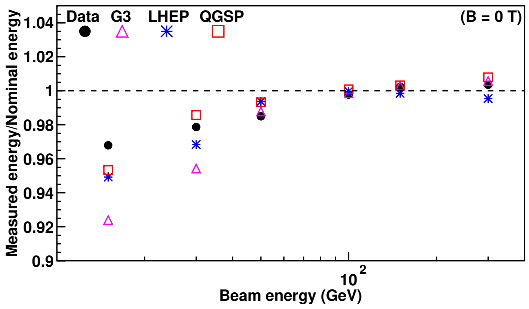

Figure 10: Energy response to pions, normalized to the beam energy, as a function of the beam energy in the CMS HCAL.

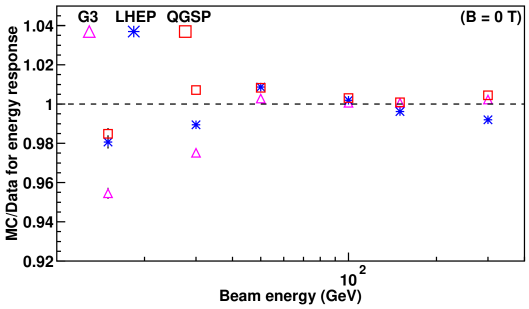

Figure 11: Ratios between simulation and data for the reconstructed pion energy in the CMS HCAL.

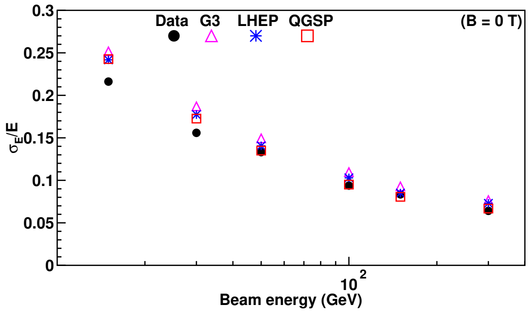

Figure 12: Pion energy resolution in the CMS HCAL as a function of beam energy.

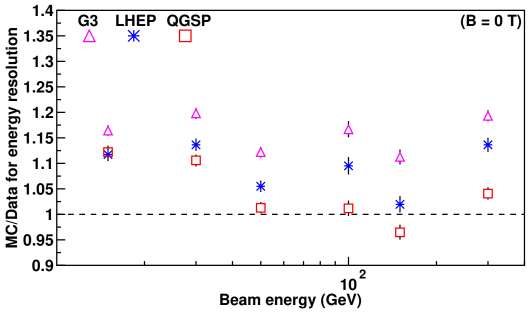

Figure 13: Ratio between simulation and data for the pion energy resolution in the CMS HCAL.

	Data	LHEP 3.6	QGSP 2.7	GHEISHA
Sampling term a [$\%\sqrt{\mathrm{GeV}}$]	71.6±1.0	71.5±1.0	74.1±0.9	82.9.7±1.2
Constant term b [$\%\sqrt{\mathrm{GeV}}$]	1.6±0.1	2.5±0.1	1.7±0.1	2.4±0.1

Table 3: Sampling and constant terms of the pion energy resolution in CMS HCAL, as obtained from simulations and test-beam data.

uncertainties added in quadrature. Systematic errors include contributions from the beam contamination (dominant at low energy) and calibration. Within the experimental uncertainties, there is good agreement between the data and the simulation, although the results are inconclusive below 50 GeV because of the large error bars. In general, the absolute energy response in Geant4 is 3-5% lower than in the data. Part of this difference may be attributed to the fact that for this preliminary study the energy scale calibration of the HCAL was not obtained in exactly the same way in the data and in the simulation. This issue requires more investigation. Concerning the energy resolution, which was calculated as σ/E as in the previous cases, in general the agreement with the data is good for both lists, with LHEP giving a few percent worse performance than QGSP.

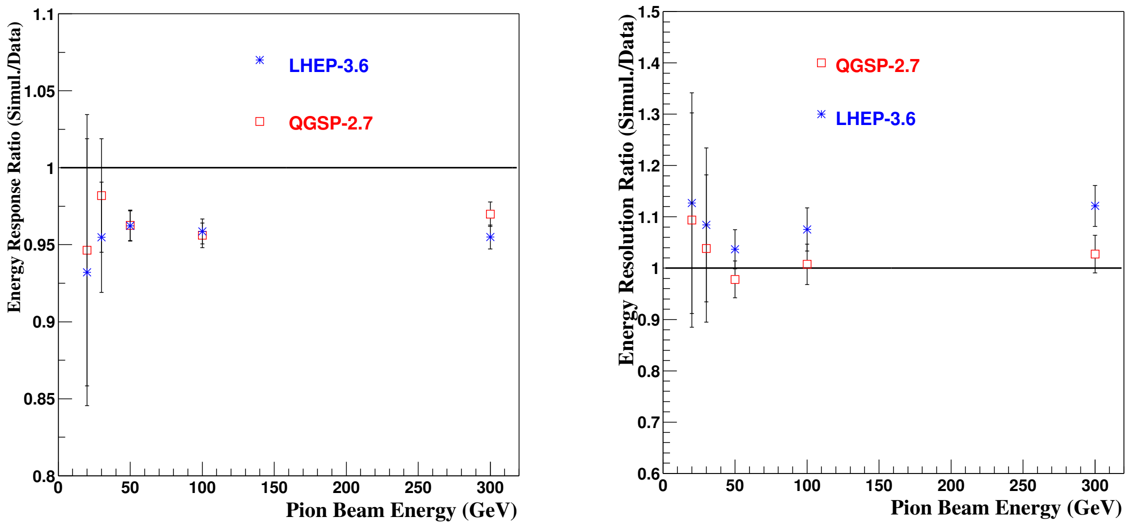

Figure 14: For the CMS TB02 set-up, the ratios between the simulated and measured energy response (left) and resolution (right) for negative pions, as a function of energy.

4.2 e/π ratio

The e/π variable is defined as the ratio between the calorimeter response to electrons and charged pions.

4.2.1 ATLAS Tilecal

The e/π ratio as a function of beam energy is shown in Figure 15 for data and simulation, and the ratios between the two are presented in Figure 16. LHEP 3.6 predicts higher values than QGSP 2.7, and the agreement with the data is $\sim 6\%$ for LHEP and $\sim 3\%$ for QGSP. For lower-energy particles at $\eta = -0.35$ (Figure 17), data and simulation agree to about 10%. Finally, Figure 18 shows the results obtained with a special geometry configuration, where the beam impinged on the lateral face of the calorimeter modules (90° incidence). The differences between the measured e/π ratios presented in Figures 15, 17 and 18 can be in large part attributed to uncorrected systematic uncertainties (the data samples differ in the energy range, geometry and analysis procedure).

4.2.2 ATLAS HEC

The e/π ratio in the HEC is presented in Figure 19 as a function of the beam energy for the data and for the Geant4 and Geant3 simulations. Ratios between simulations and data are presented in Figure 20. Geant4 describes the data with a few percent precision, whereas Geant3 predicts systematically smaller values of e/π.

4.2.3 CMS HCAL

The e/π ratio measured with the 1996 HCAL standalone set-up is presented in Figure 21 as a function of beam energy. The energy dependences for data and simulation are very

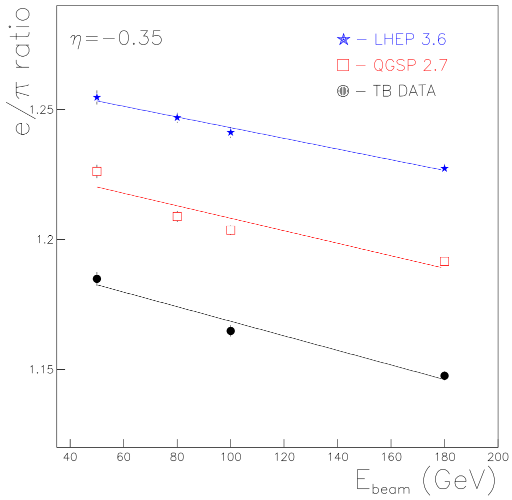

Figure 15: e/π ratio in ATLAS Tilecal at $\eta = -0.35$ as a function of energy.

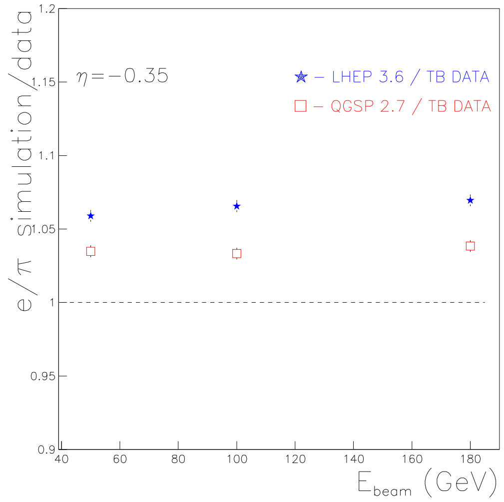

Figure 16: Ratio between simulation and data for the e/π ratio in ATLAS Tilecal at $\eta = -0.35$.

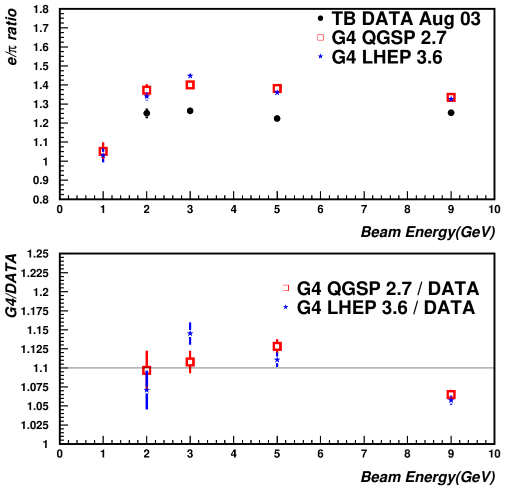

Figure 17: e/π ratio in ATLAS Tilecal at $\eta = -0.35$ as a function of energy (top), and ratio between simulation and data (bottom).

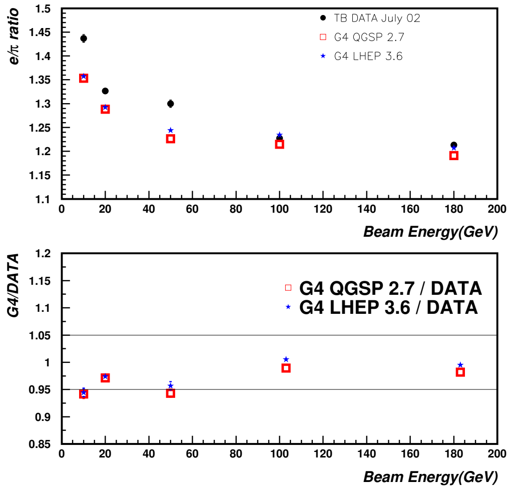

Figure 18: e/π ratio in ATLAS Tilecal at 90° incidence as a function of energy (top), and ratio between simulation and data (bottom).

similar. As can be seen also from Figure 22, LHEP and QGSP predict 3-5% smaller e/π values than measured in the data, while Geant3 predicts 3-5% higher values.

4.3 Shower profiles

4.3.1 ATLAS Tilecal

Longitudinal and transverse shower shapes were investigated by studying the energies de-

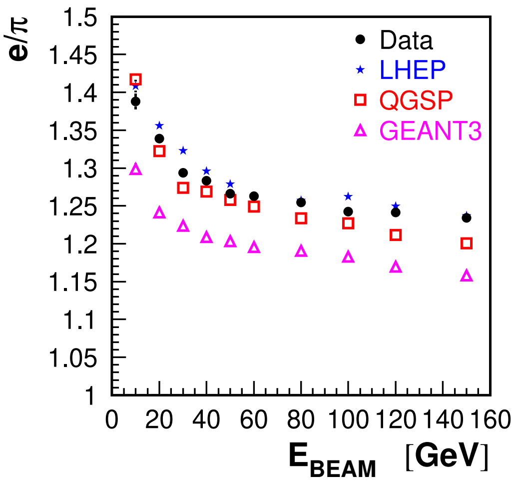

Figure 19: e/π ratio in ATLAS HEC as a function of energy.

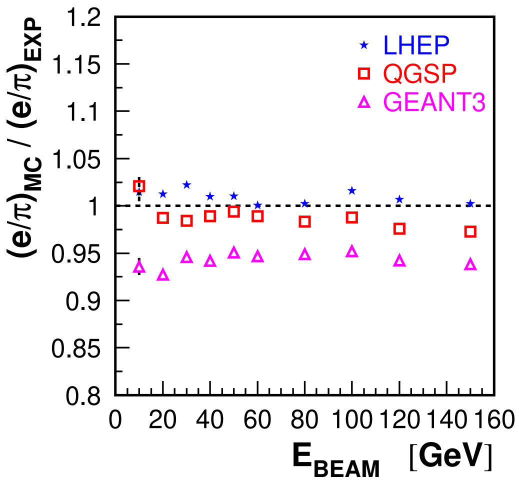

Figure 20: Ratio between simulation and data for the e/π ratio in ATLAS HEC.

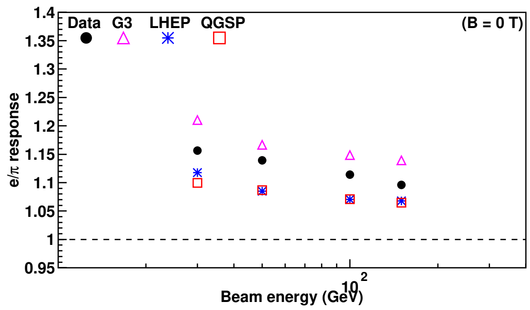

Figure 21: e/π ratio in CMS HCAL as a function of energy.

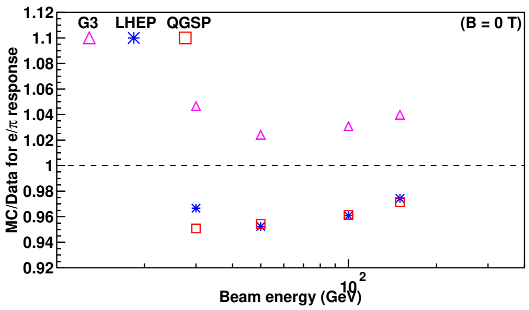

Figure 22: Ratio between simulation and data for the e/π ratio in CMS HCAL.

posited in the various cells and longitudinal compartments of the test modules. First, the energy released in the central module (Barrel) and the energy released in the modules situated above and below the central one (M0+EB), were considered. The ratio between these energies and the total deposited energy is presented in Figure 23. Both QGSP and LHEP give more energy in the central module than the data, and less energy in the lateral modules. Therefore the transverse shower profile is more compact in the simulation than in the data.

Figure 24 shows the energy deposited by pions in each longitudinal compartment of Tilecal for the data and the simulated events. The data have more energy in the first compartment than the simulation and less in the other two layers. The discrepancy increases with energy.

4.3.2 ATLAS HEC

Longitudinally the HEC calorimeter is segmented into four layers, which allows the longi-

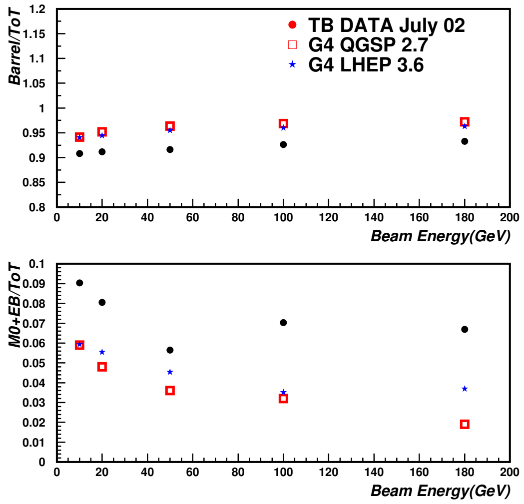

Figure 23: Ratios between the energy deposited in the central module (top) and in the lateral modules (bottom) and the total pion energy in ATLAS Tilecal.

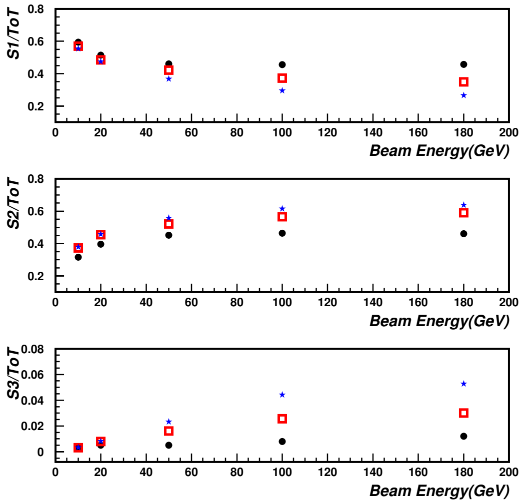

Figure 24: Fractions of the pion shower energy deposited in the various longitudinal compartments of ATLAS Tilecal as a function of the beam energy.

tudinal development of hadronic showers to be studied. As in the Tilecal case, a sensitive variable is the ratio between the energy reconstructed in a layer and the total energy reconstructed in the pion cluster. These ratios are presented in Figure 25 as a function of energy. Ratios between simulation and data are plotted in Figure 26.

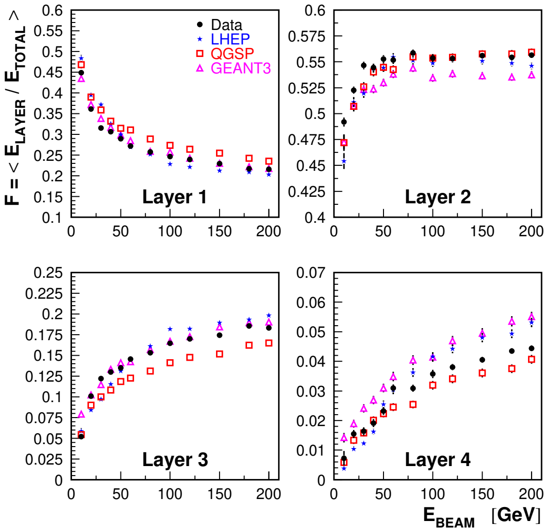

Figure 25: Fractions of the pion shower energy deposited in the various longitudinal layers of the HEC, as a function of the beam energy.

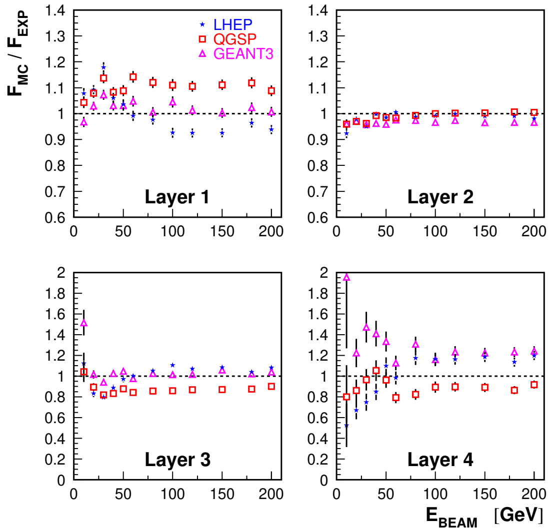

Figure 26: Ratios between simulation and data for the pion energy fractions in the ATLAS HEC layers.

Approximately half of the pion signal is concentrated in the second layer, which is rather well described by all three simulation codes. On the other hand, the agreement for the energy fractions in the first and last layers is quite modest. QGSP predicts more energy in the front part than the data and less energy in the rear for almost all beam energies. This means that hadronic showers in this model start earlier and are more compact. LHEP shows different trends at different beam energies: at small energies it is rather close to QGSP, whereas at high energies it predicts a later start of the shower. In general, Geant3 describes the longitudinal profiles of hadronic showers better than Geant4.

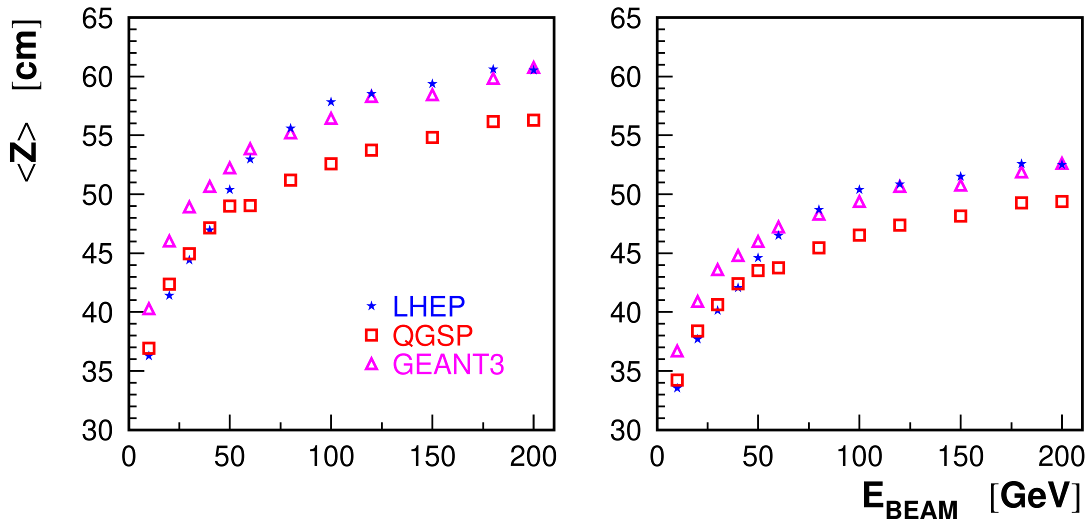

Figure 27: Average longitudinal position of π^- showers in the HEC copper plates (left) and in the liquid-argon gaps (right) for various simulation packages, as a function of the incident energy.

To better understand the differences between the various models, additional simulation studies have been performed. Negative charged pions of various initial energies were tracked through the calorimeter starting from the front face (in order to minimize the impact of the upstream material). The information of the energy depositions in the 42 copper plates and 40 liquid-argon gaps was recorded. The energy density (dE/dZ) deposited in copper and liquid-argon as a function of the calorimeter depth Z could therefore be obtained. The average Z-position of the shower is shown in Figure 27 for various beam energies. It can be seen that the predictions of the various packages differ rather significantly. In QGSP, the hadronic showers develop earlier, while in Geant3 they develop deeper in the calorimeter for all incident energies. The LHEP model is close to QGSP at low energies and to Geant3 at high energies. Similar conclusions can be drawn for the longitudinal spread of hadronic showers, shown in Figure 28.

4.3.3 CMS HCAL

The longitudinal shower profile for 100 GeV pions in the HCAL calorimeter is shown in Figure 29, where the fraction of deposited energy is plotted as a function of the layer

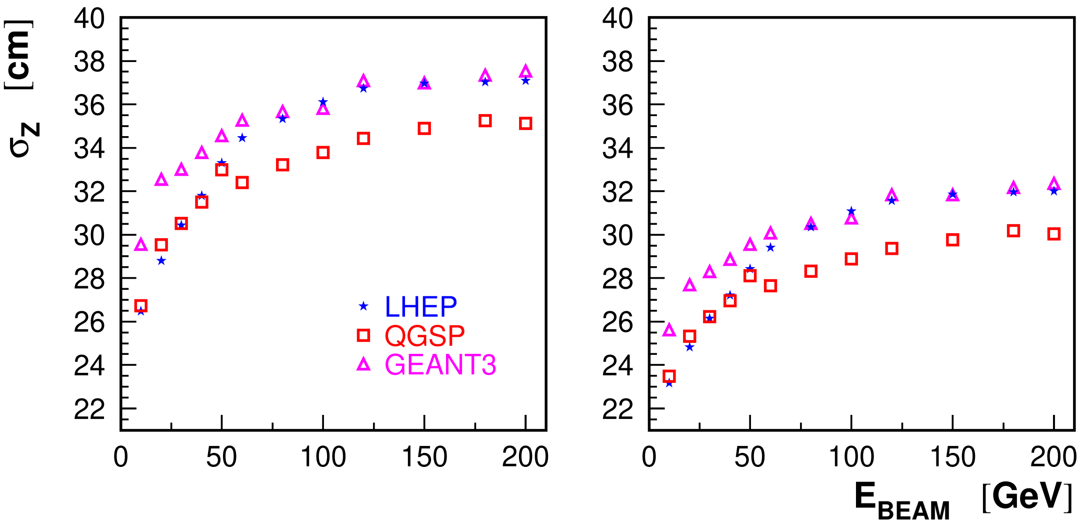

Figure 28: Spread of pion showers in the longitudinal direction in the HEC copper plates (left) and in the liquid-argon gaps (right) for various simulation packages, as a function of the beam energy.

number. As in the Tilecal and HEC cases, QGSP predicts more compact showers than the parametrised models (LHEP and Geant3 GEISHA). The data lie between the two models up to layer 23, and agree better with the parametrised models beyond. Figure 30 presents the ratios between the simulation and the data points shown in Figure 29. The agreement between data and models is good for layers between 3 and 16, whereas deeper in the calorimeter the data show significantly larger energy depositions.

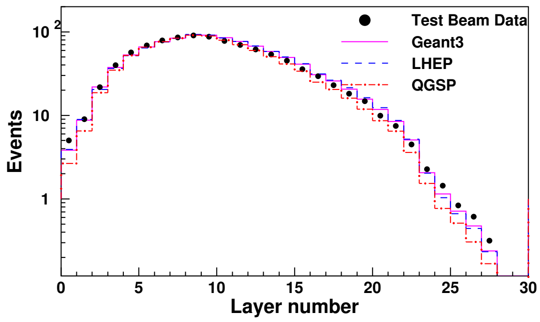

Figure 29: Longitudinal shower profile in the CMS HCAL for 100 GeV pions.

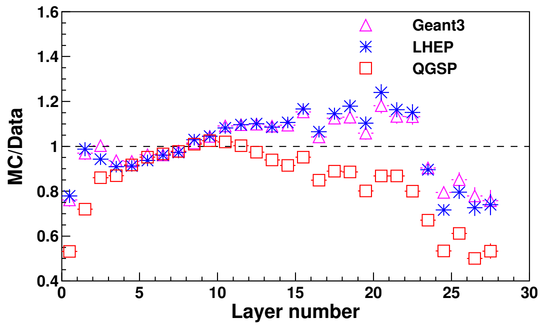

Figure 30: Ratio between simulation and data for the longitudinal shower profile of 100 GeV pions in the CMS HCAL.

The average longitudinal shower profile for pions of various energies is presented in Figure 31, where the ratios between simulation and data are given. The QGSP physics list provides the best description of the data (accuracy within 2%), while the two parametrised models show agreements at the level of 5% over the entire energy range. Similarly, the RMS spreads of the longitudinal shower profiles are shown in Figure 32. All models provide a fair description of the data, although in this case the agreement for the parametrised models is better than for QGSP.

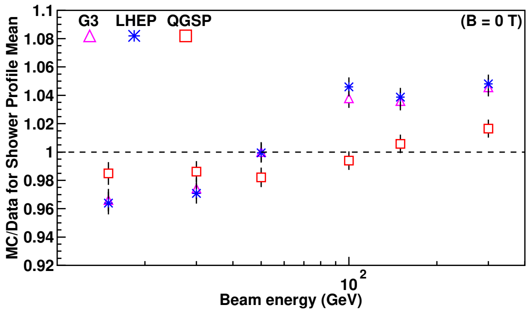

Figure 31: Ratio between simulation and data for the mean pion longitudinal shower profile in CMS HCAL, as a function of the beam energy.

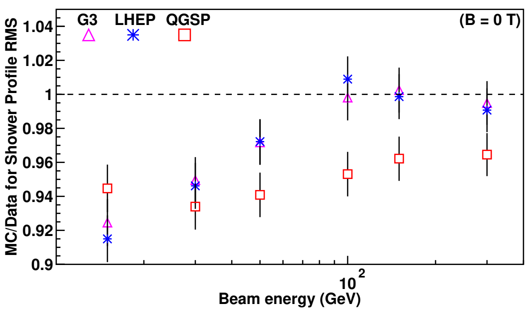

Figure 32: Ratio between simulation and data for the RMS spread of the pion longitudinal shower profile in CMS HCAL, as a function of the beam energy.

4.4 Summary and remarks

Table 4 summarizes the main features of the tested calorimeters and the beam-energy ranges considered in this study.

Detector	Type	Long. segm.	Total (compartment) thickness (λ)	Thickness in X_0	Beam energy
ATLAS-Tile	Fe/Sci	3	7.1(1.4/3.9/1.8)	$\sim$ 60	1 - 180 GeV
ATLAS-HEC	Cu/LAr	4	10.2 (1.5/2.9/3.0/2.8)	103	10 - 200 GeV
CMS-HCAL	Brass/Sci	28	10.1	$\sim$101	10 - 300 GeV

Table 4: Main features of the tested calorimeters and the test-beam energy ranges for this study.

The results for pion energy resolution presented in the previous sections are collected in Figures 33 and 34, where all three calorimeters are shown together. At energies above $\sim$30 GeV the agreement between the two Geant4 physics lists and the data is better than 10% for all detectors, whereas at lower energy significant discrepancies observed. For the e/π ratio (Figures 35 and 36), the agreement between simulation and data is within $\pm$5% for all calorimeters, and even slightly better in the case of the QGSP package.

It should be noted that the experimental points are affected by systematic uncertainties at the level of a few percent, as discussed in Section 3. These may explain, at least in part, the low-energy regions in Figures 33-34 and the systematic shifts in Figures 35-36. For instance, data collected with the HEC calorimeter during two different test-beam periods show a few percent discrepancy on the measured e/π ratios. Full understanding and possible corrections of these uncertainties are among the goals of ongoing studies.

On the simulation side, different versions of Geant4 (e.g. 5.2, 6.0, 6.2) give variations in the energy scale for electrons by up to 1.5-2%. This is also a point which requires more investigation.

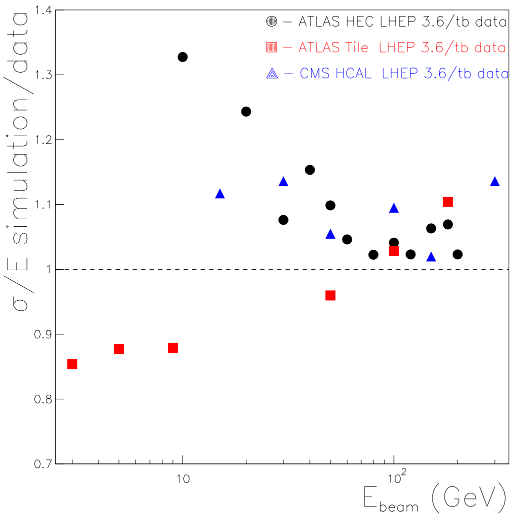

Figure 33: Ratio between the LHEP physics list and test-beam data of ATLAS Tilecal, ATLAS HEC and CMS HCAL, for the pion energy resolution as a function of energy.

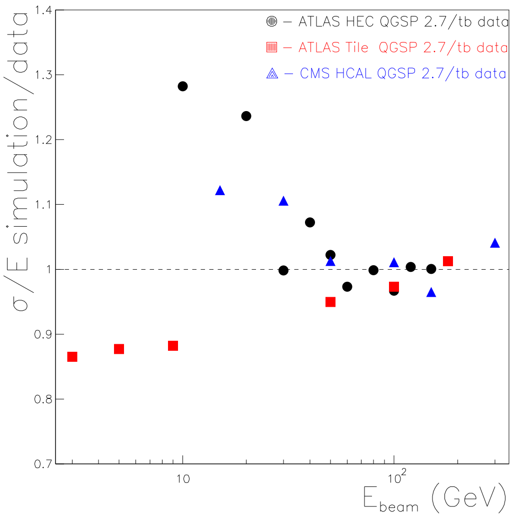

Figure 34: Ratio between the QGSP physics list and test-beam data of ATLAS Tilecal, ATLAS HEC and CMS HCAL, for the pion energy resolution as a function of energy.

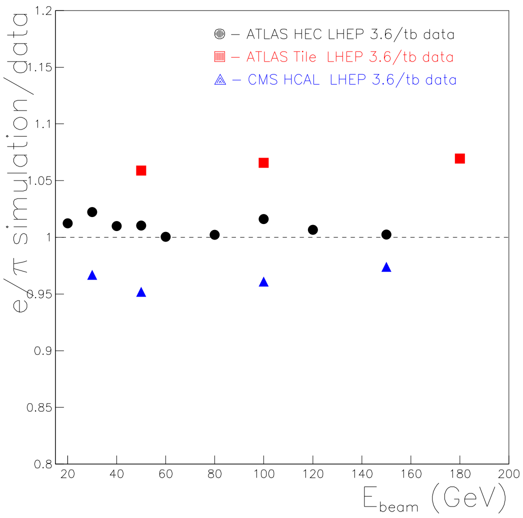

Figure 35: Ratio between the LHEP physics list and test-beam data of ATLAS Tilecal, ATLAS HEC and CMS HCAL for the e/π ratio as a function of energy.

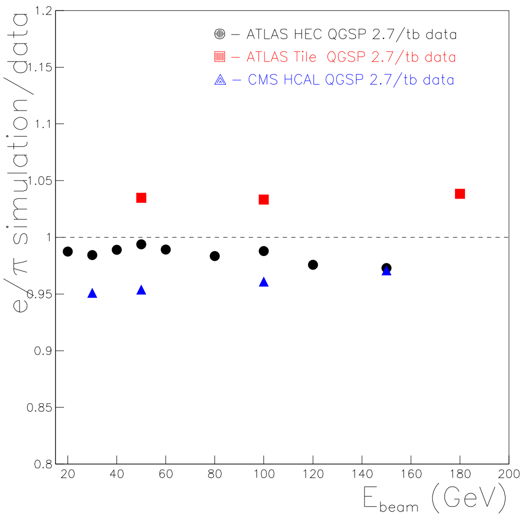

Figure 36: Ratio between the QGSP physics list and test-beam data of ATLAS Tilecal, ATLAS HEC and CMS HCAL for the e/π ratio as a function of energy.

5. Conclusions and future plans

First attempts to test two Geant4 hadronic physics lists (QGSP 2.7 and LHEP 3.6) in a coherent way using the test-beam data from three different LHC hadron calorimeters have been presented.

The main conclusions of these studies can be summarized as follows:

- Energy resolution. A good description of the calorimeter hadron energy resolution is important for several physics studies at the LHC, in particular for a precise understanding of the detector missing transverse energy resolution [2]. For beam energies above ~30 GeV, the contributions of electronic noise and photostatistics are negligible in the data. In this energy range, the agreement between simulation and data is within 10% (Figures 33 and 34), and is particularly good (at the level of 5%) for QGSP. The latter turns out to be the preferred physics list by all calorimeters for this performance issue. More work is needed to clarify the discrepancies simulation-data at energies below 30 GeV (up to 30%), which become more and more important with decreasing energy. On the data side, this is the region where instrumental effects coming for instance from noise are important (e.g. the shift in the Tilecal data visible in Figure 6 may be attributed to a change of electronics gain). A more detailed evaluation of these contributions by the experiments is therefore desirable. In addition, the Tilecal pion data collected at energies below 10 GeV may suffer from residual beam contaminations which need to be better understood. On the simulation side, the hadronic working group of Geant4 plans to achieve a better understanding of the influence of multiple scattering on the energy resolution.

  We note that comparisons between simulation and data for electrons and muons may also help to clarify these issues.

- e/π ratio. The required precision on the simulation of this quantity, coming from e.g. the sensitivity of quark Compositeness studies at the LHC, is at the level of a few percent [2]. The level of agreement simulation-data (Figures 35 and 36) is within ±5% above 20 GeV. The ATLAS HEC data (Figure 20) and the CMS HCAL (Figure 22) data are equally well described by both lists, whereas the ATLAS Tilecal data show a better agreement with QGSP. The energy-independent +5% (−5%) scale offset for Tilecal (CMS HCAL) may indicate a problem with the absolute energy scale calibration, which requires more investigations on the data side. At energies below 10 GeV only Tilecal data are available (Figure 17). Here the e/π ratio is systematically higher (by about 10%) in the Geant4 simulation (for both lists) than in the data, which may indicate an underestimation of the electron contamination in the low-energy pion beam. Detailed studies of the response linearity of all three calorimeters for electron and pion beams over the largest possible energy range would be useful to understand and hopefully improve these results.

  On the simulation side, the description provided by the hadronic packages has reached a level of accuracy such that one is sensitive also to the impact of the electron signal simulation on the e/π ratio. There is presently an uncertainty of 1.5-2% on this simulation, as noted above. In addition, the hadronic working group of Geant4 plans to study the influence of multiple scattering and gamma-nuclear reactions on this quantity. The goal of reaching a few percent predictive power is very challenging, and turns the physics into a totally coupled system.

- Shower profiles. The level of accuracy in the simulation of hadron shower profiles has an impact for instance on particle identification (e.g. e/π and e/jet separation), and on the estimation of energy losses in the dead regions of the detector. The latter affect the description of the hadron and jet energy resolutions, and of the tails in the event missing transverse energy distribution[2]. Shower profiles also define the energy sharing between different calorimeter technologies and among calorimeter cells, and hence directly influence weighting techniques used to reconstruct the pion and jet energies in non-compensating detectors.

  All studies have been made so far for energies above 10 GeV. In general, the QGSP list predicts showers which are systematically more compact than those predicted by LHEP. Concerning the transverse shower profiles, only Tilecal results are available for the time being (Figure 23). The agreement simulation-data is within a few percent for the shower core (Figure 23 top) and worse in the tails (Figure 23 bottom). The data showers are wider than the simulation, which may be due, at least in part, to electronic noise and upstream material effects. In the ATLAS HEC, the number of cells in a pion cluster, in combination with the longitudinal shower shape, also indicates that QGSP showers are more compact in the transverse plane than the data and LHEP showers are wider. Both QGSP and LHEP showers are significantly more compact than Geant3 showers.

  Concerning the longitudinal profiles, where results from all three calorimeters are available, the behavior is different for the different detectors. For ATLAS Tilecal, showers are more compact in the data than in both physics lists, whereas for CMS HCAL and ATLAS HEC the data are in between QGSP and LHEP. In general, the relative agreement simulation-data is at the level of a few percent in the shower core and 10-20% in both the start and tails of the showers. An energy-dependent disagreement emerges from a comparison between the two physics lists alone (Figures 27 and 28).

  On the simulation side, the Geant4 hadronic working group plans to study the link between shower shapes in calorimetry and elementary cross-sections, which will help to understand the details of this detector performance aspect.

In conclusion, the preliminary results presented in this note show that Geant4 reproduces the pion test-beam data of three different LHC calorimeters in a generally satisfactory way. In particular, the description of the e/π ratio and of the energy resolutions are close to requirements. No clear preference for one of the two investigated lists, QGSP and LHEP, emerges at this stage, although the data seem to indicate a slightly better overall agreement with QGSP.

To push further the sensitivity of these comparisons, some systematic effects which seem to plague the data (see e.g. Figures 35 and 36) need to be clarified, and experimental uncertainties, in particular affecting the (challenging) low-energy region below $\sim$30 GeV, need to be understood. It should be noted that only test-beam set-up and test-beam data of high quality and well under control can allow studies of the pion energy response (resolution, absolute energy scale) to the few percent level. Understanding the shower profiles requires much more work, in particular on the simulation side, and no firm conclusions can be drawn on this issue for the time being.

The availability of new test-beam data from all LHC experiments on a time scale of a few months (in particular data from the ATLAS combined test-beam runs, and data from low-energy CMS runs) will allow a new set of extensive studies to be undertaken, and will provide important benchmarks and constraints to test both Geant4 and Fluka.

References

[1] Physics Validation Project, http://lcgapp.cern.ch/project/simu/validation/

[2] A. De Roeck, F. Gianotti, A. Morsch, W. Pokorski, *Simulation physics requirements from the LHC experiments, CERN-LCGAPP-2004-02.*

[3] S. Agostinelli *et al.* [Geant4 Collaboration], Nucl. Instrum. Meth. A **506**, 250 (2003).

[4] J.P. Wellisch, http://cmsdoc.cern.ch/ hpw/GHAD/HomePage/

[5] R. Brun, F. Bruyant, A.C. McPherson, M. Maire, P. Zanarini, CERN Data Handling Division, DD/EE/84–1, 1987. Also, Geant - Detector Description and Simulation Tool, CERN Program Library Long Write-up W5013, IT Division, API group, CERN.

[6] A. B. Kaidalov, K. A.Ter-Martirosyan, Phys. Lett. **B117** 247 (1982); Sov. J. Nucl. Phys. **39**, 979 (1984).

[7] V. Lara and J. P. Wellisch, CHEP 2000: Computing in high energy and nuclear physics, 52-55.

[8] V.S. Barashenkov, Dubna Preprint P2-90-158 (1990).

[9] H. P. Wellisch and D. Axen, Phys. Rev. C **54**, 1329 (1996).

[10] MCNP: Judith F. Briesmeister (Los Alamos), LA-7396-M, 1986.

[11] P. V. Degtyarenko, M. V. Kossov and H. P. Wellisch, Eur. Phys. J. A **9**, 421 (2000); Eur. Phys. J. A **9**, 411 (2000); Eur. Phys. J. A **8**, 217 (2000).

[12] H. Pi, *An event generator for interactions between hadrons and nuclei: FRITIOF version 7.0*, Comput. Phys. Commun. 71 (1992).

[13] ATLAS Collaboration, *Tile Calorimeter Technical Design Report*, CERN/LHCC 96-42 (1996).

[14] A. Solodkov and V. Tsulaia, *Tilecal Beam Test Simulation Application in the FADS/Goofy Framework (Geant4)*, ATLAS Internal Note ATL-TILECAL-2003-002.

[15] ATLAS Tilecal Simulation, Setting G4 parameters, http://tilecal.web.cern.ch/tilecal/Simulation/G4settings/G4settings.html

[16] ATLAS Collaboration, *Liquid Argon Calorimeter Technical Design Report*, CERN/LHCC/96-41 (1996).

[17] B. Dowler *et al.*, Nucl. Instrum. Meth. A**482**, 94 (2002).

[18] A.E. Kiryunin and D. Salihagić, *Monte Carlo for the HEC Prototype: Software and Examples of Analysis*, ATLAS Internal Note ATL-HEC-063 (1998).

[19] A. Kiryunin, D. Salihagić and P. Strizenec, *Geant4 Simulation of the HEC Test Beam Data (Status Report)*, ATLAS Internal Note ATL-HEC-090 (1999).

[20] D. Benchekroun *et al.*, *Comparisons of the Electron and Muon Testbeam Response of the ATLAS Liquid Argon Calorimeters with Geant4 Simulations*, ATLAS Internal Note ATL-LARG-2004-003 (2004).

[21] CMS Collaboration, *Hadron Calorimeter Technical Design Report*, CERN/LHCC/97-31 (1997).

[22] V. V. Abramov *et.al.*, Nucl. Instrum. Meth. A**457**, 75 (2000).