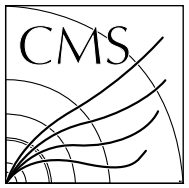
Available on CMS information server

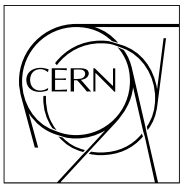
CMS NOTE-2008/025



The Compact Muon Solenoid Experiment

CMS Note

Mailing address: CMS CERN, CH-1211 GENEVA 23, Switzerland



05 September 2008

Calorimetry Task Force Report

S. Abdullin, F. Ambroglini, S. Banerjee, S. Beauceron, F. Beaudette, A. Bhatti, S. Bose, F. Chlebana, F. Cossutti, D. Jang, S. Y. Jun, S. Kunori, R. S. Kroeger, M. Paulini, M. Pierini, S. Piperov, S. Popescu, C. Rogan, F. Ronga, C. Rovelli, G. Safronov, S. Sharma, M. Spiropulu

Abstract

In this note we summarize the studies and recommendations of the calorimeter simulation task force (CaloTF). The CaloTF was established in February 2008 in order to understand and reconcile the discrepancies observed between the CMS calorimetry simulation and the test beam data recorded during 2004 and 2006. As the result of studies by the CaloTF a new version of Geant4 was developed and introduced in the CMS detector simulation leading to significanly better agreement with test beam data. Fast and flexible parameterizations describing showering in the calorimeter are introduced both in the Full Simulation (with a Gflash-like approach) and in the Fast Simulation. The CaloTF has developed a strategy to rapidly tune the CMS calorimeter simulation using the first LHC collision data when it becomes available. The improvements delivered by the CaloTF have been implemented in the software release CMSSW 2.1.0.

Executive Summary

The purpose of this document is:

1. To document the level of agreement between the test beam data and the Monte Carlo Simulation of calorimeter showers using standard CMS simulation software.
2. To study in detail and analyze the discrepancies so far observed between the test beam data and the available Monte Carlo simulations.
3. To investigate possible improvements in the simulations (e.g. physics models at the particle interaction level with the material of the CMS calorimeters) and provide methods for tuning simultaneously the electromagnetic and hadronic part of the simulated showers directly to data (test beam and collision data).
4. To provide a strategy for the CMS collaboration to iteratively implement improvements in the simulation both at the calorimeter level (collection of towers, jets) and the individual particle level.

Principal Findings

Progress has been made in understanding the disagreement between the test beam data and the available Monte Carlo calorimeter simulations at CMS. The implementation of the GFLASH-like parameterization allows for a faster simulation and provides greater flexibility to tune the Full GEANT4 simulation to the data. The development of methods to tune the hadronic showers and the direct comparisons with the Full Simulation is in progress. Finally the strategy for directly tuning the parameterizations of the CMS Fast Simulation both for the electromagnetic and hadronic showers is outlined and partially implemented using the test beam data as a reference. The work-flow and first estimates for the calibration needs of the calorimeter with collision data (starting already from dedicated trigger paths) is discussed.

Primary Recommendations

- The use of a new version of GEANT4 together with a new physics list provides a better agreement between the simulation and the test beam data. A better description of the energy response saturation of scintillation light emission is also included.
- The implementation and tuning of a GFLASH-like parameterization for the electromagnetic and hadronic showers in the CMS Full Simulation.
- The development and tuning of a Fast Simulation parameterization for hadronic showers.

Structure of this document

Section 1 describes the mandate of the CALOTF. It is followed by Section 2 where the current understanding of the GEANT4-based CMS calorimetry simulation is described together with the physics requirements and computing constraints. A new physics list together with an improved version of GEANT4 has been studied and shown to provide better agreement with the test beam data over a large range of measured energies and has been made the default starting with CMSSW_2_1_0. The fast GFLASH-based parameterization and tuning method for both the electromagnetic and hadronic shower is described in section 3. The parameter space and the available flexibility to tune the simulation using the test beam data is explored and comparisons with the full simulation are presented. The following section 4 describes the status of agreement with the test beam data and the full simulation as well as the methodology for tuning the CMS Fast Simulation electromagnetic and hadronic shower to the test beam data. The following three sections summarize the strategy for the use of LHC collision data (using dedicated trigger paths for iterative calibration) and the outlook given the work of this task force. Finally, an appendix is included providing additional material and references describing the test beam configuration and analysis.

Contents

1 Introduction and Charge

There are two detector simulation packages available for CMS. "FullSim" refers to a detailed detector simulation using the GEANT4 toolkit [1]. The detector geometry is described in detail and the particle interactions with the detector material is calculated from first physics principles. FullSim can take several minutes of CPU time per event. "FastSim" refers to a detector simulation using a simplified parameterization of the particle response and typically takes less than a second per event. In order to analyze the early LHC collision data and have reliable physics results the detector simulation should agree with the data to within 5% for kinematic distributions obtained from Standard Model enriched control samples.

Although FullSim provides a more accurate and precise simulation of the detector we expect to be able to tune FastSim to agree with FullSim to within 1% for the bulk of global kinematic and topology distributions important for physics analysis. FullSim will always be necessary as a validation reference for FastSim and very importantly as the appropriate detector simulation tool for the understanding of tails of distributions and small samples of "unusual" events where new physics may be contributing.

The measured CMS calorimeter response was determined using electron (H4) and hadron (H2) test beam data. A detailed description of the test beam configuration can be found in Sect. 9. Some discrepancies were observed when comparing the measured response to the full GEANT4 detector simulation. Namely:

- The Monte Carlo simulation overestimates the energy deposited in the lead tungstate crystals ($\mathrm{PbWO_4}$), particularly for low hadron beam momenta (below 10 GeV/c);
- The shower depth is too short in the simulation;
- The energy sharing between the hadron and electromagnetic calorimeters is inadequately simulated.

A calorimeter simulation task force (CALOTF) was established by the CMS management to investigate how the simulation can be improved to better describe the test beams results. Members of the task force were selected from the relevant DPG's, POG's, and the simulation groups. The charge of the task force is summarized below.

- For the full simulation of the ECAL and HCAL system:
  - Evaluate and "fix" or tune the shower models used by GEANT4 to improve the agreement between the simulation and the test beam data for the linearity of the response, the resolution, and the shower shape;
  - Implement the saturation effects (Birks' law [2]) in the scintillators;
  - Implement the contribution of Cerenkov light in the ECAL response;
  - Develop a GFLASH [3] based parameterization for the electromagnetic and hadron shower shapes tuned to the test beam data.
- For the fast simulation of the calorimeter system:
  - Tune the parameterization of the electromagnetic and hadron showers to the full simulation targeting better than 1% agreement;
  - Tune the shower parameterization to the available test beam data.
- Usage of collider data:
  - Provide a concise strategy to tune both the full and fast simulation to collider data. The strategy includes the specification of a trigger paths to record the necessary data as well as the tools for analysis and tuning of the simulation.

This report sumarizes the work of the CALOTF and outlines an agile program of tuning, validation, and cross-validation of the CMS calorimeter simulations. The retuning of the simulation, using the proceedures and tools developed by the CaloTF, will need to be performed with the first LHC collision data in order to respond to the needs of the physics research program of CMS. The first set of improvements have been included in CMSSW version 2.1.0.

2 Full GEANT4 Simulation

The full simulation of the CMS detector is implemented within the CMSSW framework [4] using the GEANT4 toolkit [1], the state-of-art simulation of the passage of particles through matter. All detector components, active as well as passive, are described in detail using a XML-based Detector Description (DD) formalism configurable at run time via a hierarchy of XML files. When configuring GEANT4 the user must specify what physics processes are to be used in the simulation. GEANT4 provides several physics models which can be selected by the CMS software at run time. The choice is motivated by the specific application, precision desired, and computing time available. Some of the models implemented for hadronic interactions include the Low Energy Parameterized model (LEP) [5]; the High Energy Parameterized model (HEP) [5]; the Binary cascade (BIC) [6]; the Bertini cascade (BERT) [7, 8]; the PreCompound model [8]; the Quark-Gluon String model (QGS) [9]; the CHiral Invariant PhaseSpace model (CHIPS) [10] and the Fritiof fragmentation model (FTF) [9]. The models used in GEANT4 to describe the interactions of hadrons are valid within limited energy regions. Several models are combined together in order to cover the full energy range of interest. Models for the electromagnetic as well as hadronic processes to be used in the simulation are grouped together in a collection referred to as a "physics list" [7]. Some of the available physics lists are the Low and High Energy Parametrized model (LHEP); the Quark-Gluon String model with the PreCompound model back-end (QGSP); the Quark-Gluon String model with the CHIPS generator back-end (QGSC) and the Fritiof fragmentation model with the PreCompound model back-end (FTFP). The cross sections used by the models have been carefully compiled from existing data. The formation of exclusive final states in the simulation process and the energy sharing of the final state particles are validated with data from various thin target experiments.

The default physics list used in software releases prior to CMSSW 2.1.0 has been QGSP_EMV. This physics list uses a string model to describe interactions of high energy hadrons while interactions at lower energies are described by the low energy parameterized (LEP) model. In the QGSP_BERT_EMV physics list, the LEP model is replaced with the Bertini cascade model for primary protons, neutrons, pions and kaons below $\sim$10GeV.

Test beam data for electromagnetic showers collected at the H4 beam line have been compared [11] with the predictions from GEANT4. The geometry description of the barrel electromagnetic calorimeter (EB) super-module used in the test beam utilizes the same description files as is used in the full CMS detector simulation. In addition, the geometry of the beam hodoscope is described in great detail for this comparison. The default physics list QGSP_EMV is used and particles are traced with the same production cuts as in the full CMS detector simulation (range of 1 mm for electrons, positrons, and photons in the region of the electromagnetic calorimeter). The linearity of the energy response for electron beams, measured using the energy sum of 5×5 crystals (E25) surrounding the crystal with the highest measured energy (E1), is well reproduced by the Monte Carlo simulation for energies between 20 GeV and 150 GeV. The transverse energy profile measured as the ratios E1/E9 or E1/E25 is reproduced by GEANT4 within an uncertainty of 0.4%.

The GEANT4 predictions for hadron showers are compared with data obtained from the H2 beam line (see Sect. 9.2) during 2004 and 2006. The test beam simulation shares many of the common elements with the simulation of the entire detector including the geometry description of calorimeter modules, production cuts, and handling of hits. The full simulation results did not adequately describe the data. The observed discrepancies are:

- The simulated energy deposited was higher than what is observed, particularly at low energies as shown in Figure 1;
- The mean energy response as a function of the beam energy is not adequately simulated for both the overall sample and for hadrons that give minimum ionizing signal (MIP) in the crystals as shown in Figure 2;
- The simulated showers are narrow and shorter as shown in Figure 3.

2.1 Tuning the Full Simulation

Figure 4 shows the measured energy distribution in the electromagnetic calorimeter for 5 GeV/c π^-s compared with the predictions from GEANT4 using the QGSP physics list. While the peak at around 400 MeV (due to MIP events in ECAL) is well reproduced by the Monte Carlo simulation, the peak due to showers in the ECAL occurs at higher energies than in data. A significant fraction of this energy deposit is associated with low energy ions and nucleons produced in the shower. The lower response in the data may be explained by the effects of saturation in scintillation light emission as observed in organic scintillator by Birks [2]. Saturation effects have been observed in inorganic crystals like BGO and BaF_2, however, it has not been experimentally studied in $PbWO_4$ crystals.

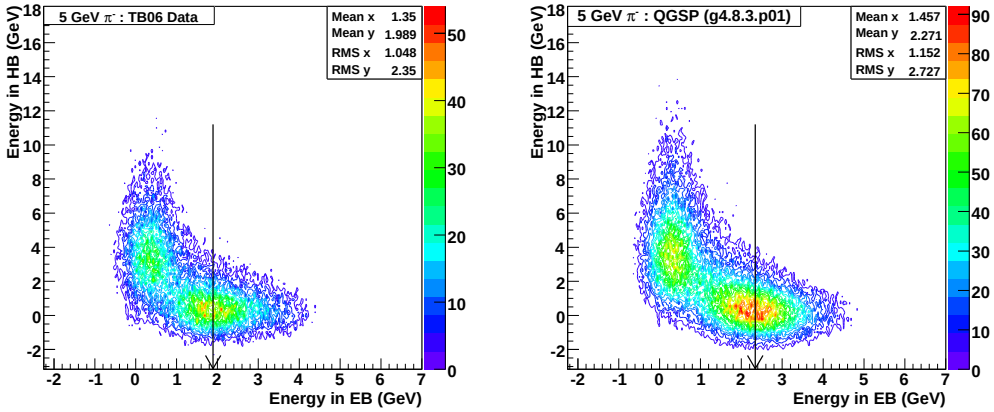


Figure 1: The measured energy in HCAL vs ECAL for 5 GeV π^- in the (a) data, (b) GEANT4 prediction using the QGSP physics list. The simulated energy in ECAL is higher than what is observed for the test beam.

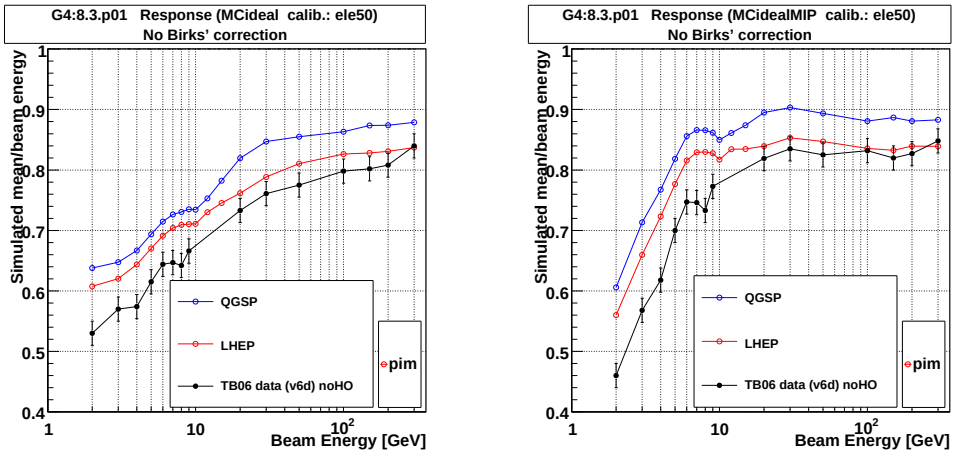


Figure 2: The mean response for a π^- beam as a function of beam momentum compared with the predictions of GEANT4 using the LHEP and QGSP physics lists for (a) all interactions, (b) MIP signal in the ECAL. The simulation is not adequately describing the data.

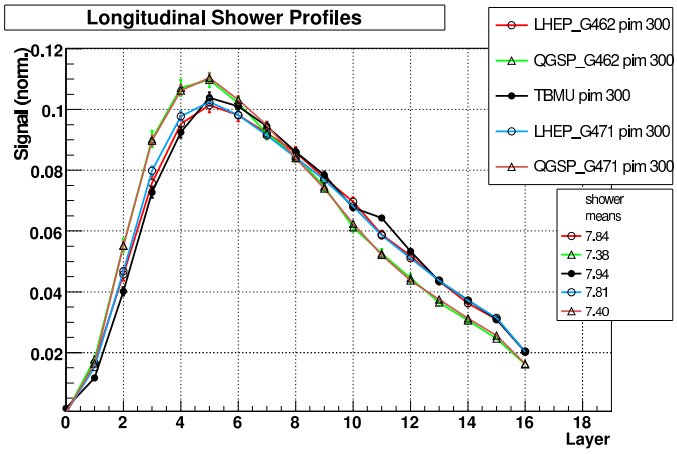


Figure 3: The longitudinal shower profile for 300 GeV/C π^- measured in the 2004 H2 test beam compared with the GEANT4 simulation for different physics lists. The simulated showers using the QGSP physics list are narrower and shorter compared to the data.

Enabling Birks' law in the simulation using the measured coefficient for BGO gives a better description of the measured energy in the ECAL as shown in Figure 4 (distribution in red). The description of energy deposits in the HCAL alone (for MIPS in ECAL) can be improved by the introduction of Birks' law in the plastic scintillator of HCAL.

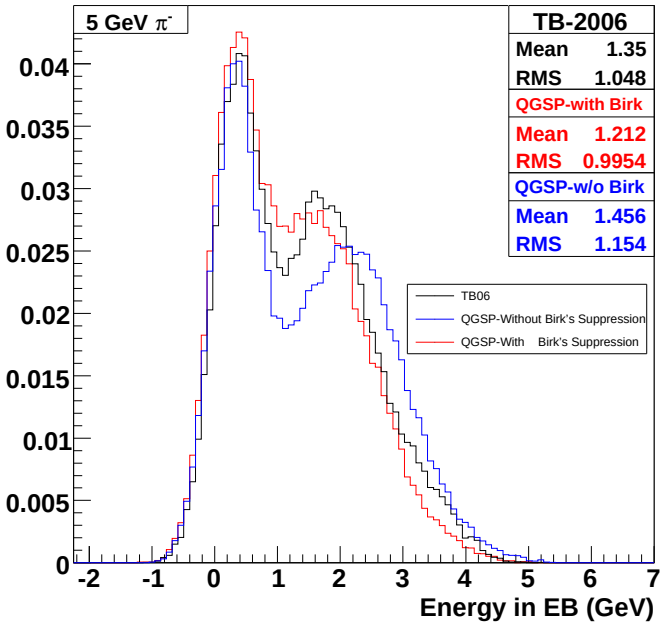


Figure 4: Test beam results for the energy measured (black) in the electromagnetic calorimeter compared with the GEANT4 simulation using the QGSP physics list without (blue) and with (red) Birks' law enabled.

Enabling Birks' law and using the QGSP (or QGSP_EMV) physics list results in a significant underestimation of the mean energy measurement. We recover the agreement by introducing a physics list that results in larger energy deposits at lower energies such as QGSP_BERT (or QGSP_BERT_EMV). The combination of the new physics list and Birks' law leads to a better agreement between data and simulation as shown in Figure 5 for the mean energy response.

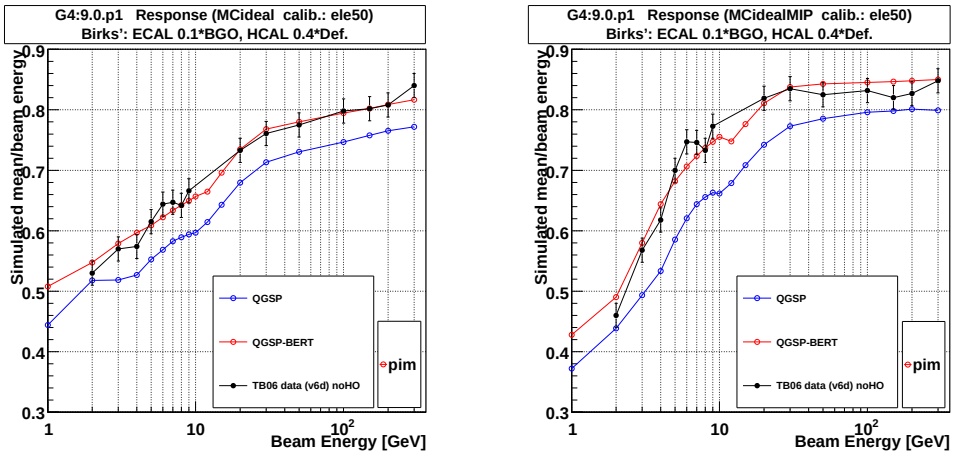


Figure 5: Mean energy response for the π^- beam in the calorimeter compared with the predictions of GEANT4 using the QGSP and QGSP_BERT physics lists for (a) all showers, (b) showers with a MIP signal in the ECAL.

An important consideration is to ensure that the use of Birks' law preserves the already good description of the electromagnetic showers. The effect on the absolute energy scale and on the transverse shape of the shower are studied on electrons samples using CMSSW 2.0.0pre6. Figure 6 shows the absolute energies deposited in the

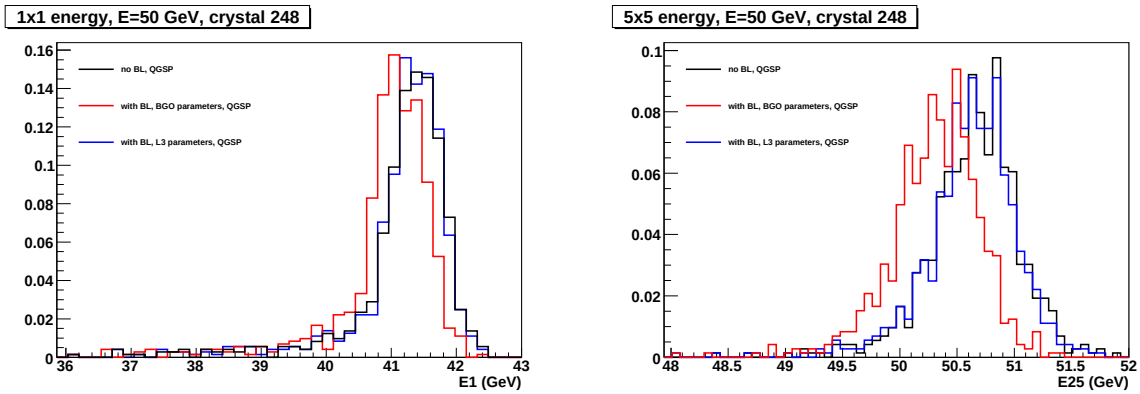


Figure 6: The effect of the Birks' law on the absolute energy scale, for the central crystal E1 (on the left) and a 5×5 crystal matrix E25 (on the right) for a sample of 50 GeV electrons simulated using FullSim with the QGSP physics list. Results for both the BGO and the L3 parameterizations are shown.

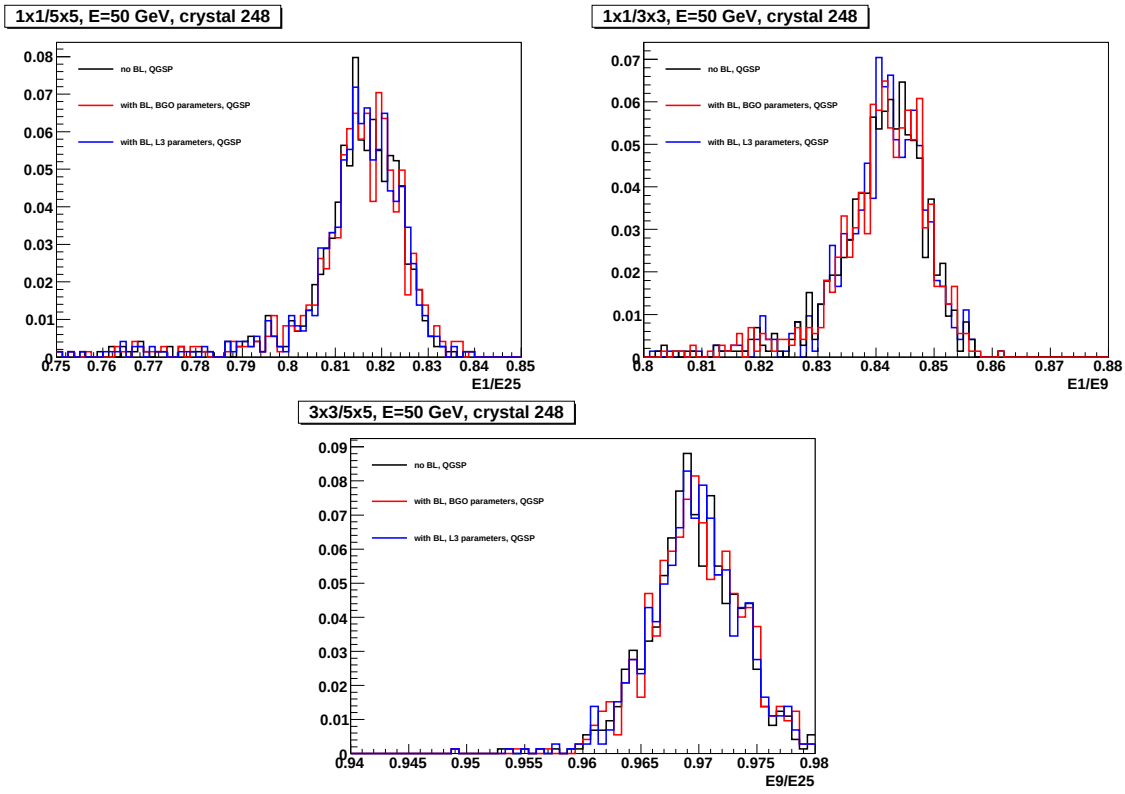


Figure 7: The effect of Birks' law on the transverse energy shape for a sample of 50 GeV electrons simulated using FullSim with the QGSP physics list. The ratios of different crystal matrices E1, E9 and E25 are shown. Both the BGO and the L3 parameterizations are shown.

central crystal (E1) and in a 5×5 crystal matrix (E25) while figure 7 shows the ratios of the energy deposits for E1/E25, E1/E9, and E9/E25. The crystal response depends strongly on the electron impact point on the crystal face, therefore only the electrons impacting close to the maximum containment point, defined as the impact point which yields the highest measured energy, of the crystal were kept for the analysis. The two figures refer to 50 GeV electrons impacting on crystal 248, which is located in the middle of module one far from module boundaries, but the results were checked to be qualitatively independent of the beam energy and the crystal position in the supermodule. There is almost no effect on the transverse shape while the absolute energy depends on the parameterization used. The L3 parameterization is in good agreement with the distributions obtained without Birks' law.

The response in absolute energy has been studied using the MC simulation as a function of the beam energy beyond the energies available at the H4 testbeam. Figure 8 shows the ratio of the E25 energy measured with and without Birks' law enabled, for both parameterizations. No slope is seen, and similar results are true for all modules, e.g. in the whole η acceptance of the barrel supermodule.

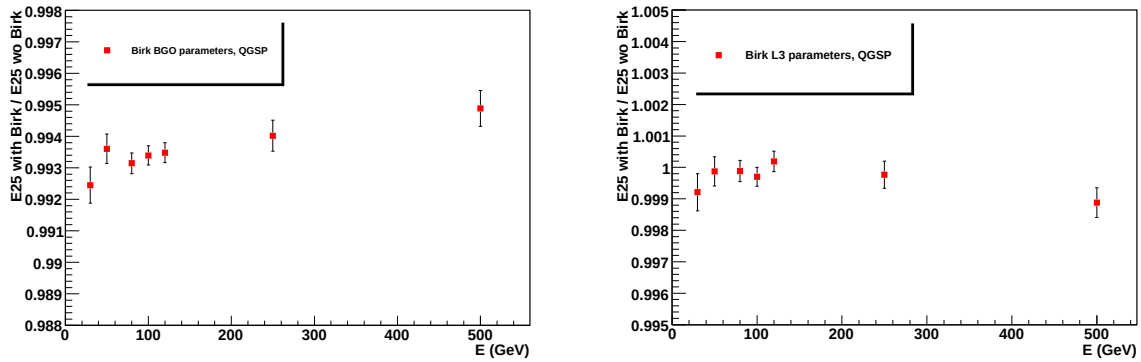


Figure 8: The ratio of the E25 energy measured with and without enabling Birks' law shown as a function of the beam energy. Results for both the BGO (left) and L3 (right) parameterizations are shown.

The distributions shown in the previous plots are obtained using the QGSP physics list, similar results are observed using the default QGSP_EMV or QGSP_BERT physics lists. We conclude that the use of Birks' law with the L3 parameterization preserves the description of the electromagnetic showers while the choice of the physics model used for hadronic interactions (basic QGSP or Bertini cascade) does not affect the electromagnetic showers.

2.2 Choice of GEANT4 Model

The Calorimeter Task Force worked together with the GEANT4 team to address some of the discrepancies observed between the simulation and the test beam results. As a result of this effort, a new version of GEANT4 (9.1.p02) was released and we recommend to use this version for the CMS detector simulation. This new version has improvements in the Bertini cascade code where Coulomb barrier effects are introduced for the de-excitation process and also in the treatment of quasi-elastic scattering for high energy region as described by the QGS model. The new GEANT4 version also has changes in the LHEP model resulting in an improved p_{T} distribution of hadrons from strange baryon reactions. We recommended to use the QGSP_BERT_EMV physics list rather than the default (QGSP_EMV). The electromagnetic physics list has been modified to include electromagnetic interactions for long-lived charged hadrons with c and b quarks. In addition, Birks' law correction is enabled by default in crystals (using the L3 parametrization for BGO) as well as in the plastic scintillator (with 40% of the suppression used in GEANT3).

The modifications give a better agreement with the mean energy response for all showers as well as showers starting in the hadron calorimeter (Figure 9). Also the transverse shower profile, measured in the ECAL (Figure 10), and the longitudinal shower profile, measured with the specially designed HB module (Figure 10) show the new parametrization describes the data better than the previously used physics lists QGSP and QGSP_EMV.

A critical measurement for the calibration of the HCAL is the fraction of MIP events as a function of beam energy. The data and the Monte Carlo simulation disagree (as seen in Figure 11) in the energy dependence of MIP fraction. The larger fraction in the simulation at low energy is attributed to an inadequate description of the quasi-elastic

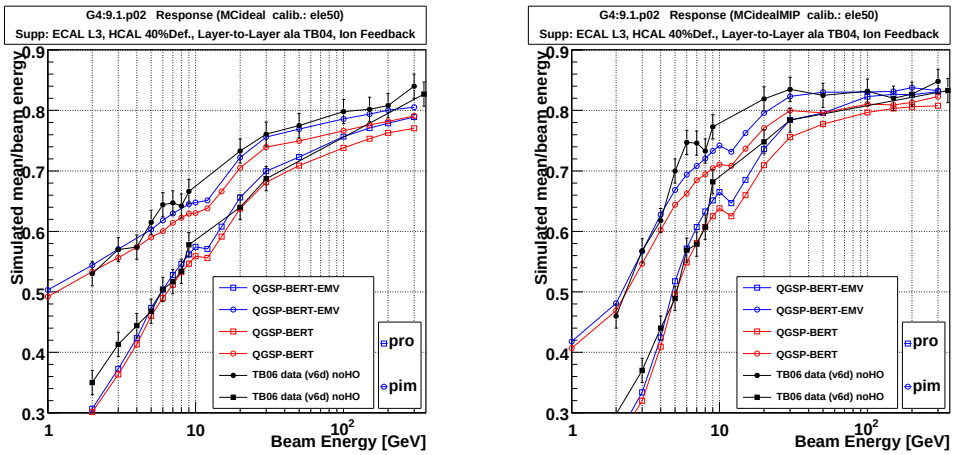


Figure 9: Mean energy response for π^- and proton beam in the calorimeter compared with the predictions of GEANT4 using the QGSP_BERT and QGSP_BERT_EMV physics lists for (left) all showers, (right) showers with a MIP signal in ECAL.

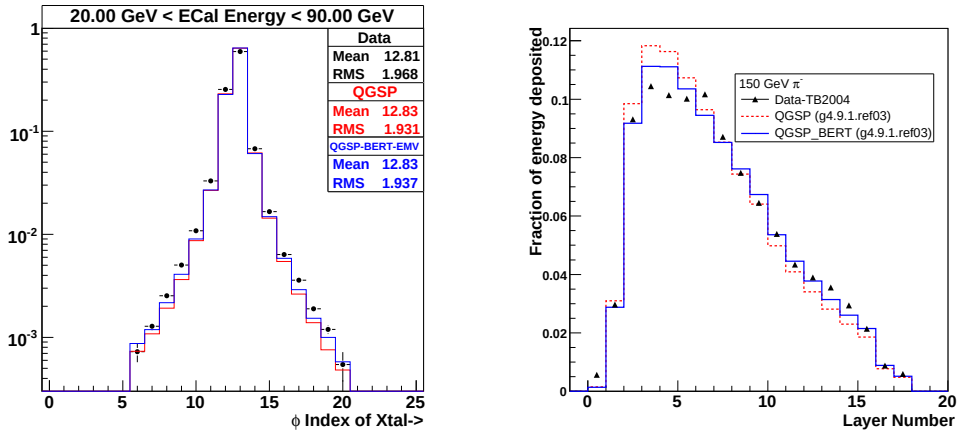


Figure 10: (left) Transverse shower profile of 100 GeV/c π^- and (right) longitudinal shower profile of 150 GeV/c π^- compared with the predictions of GEANT4 using the QGSP and QGSP_BERT_EMV/QGSP_BERT physics lists.

process in the Bertini model. This is presently under investigation by the GEANT4 team.

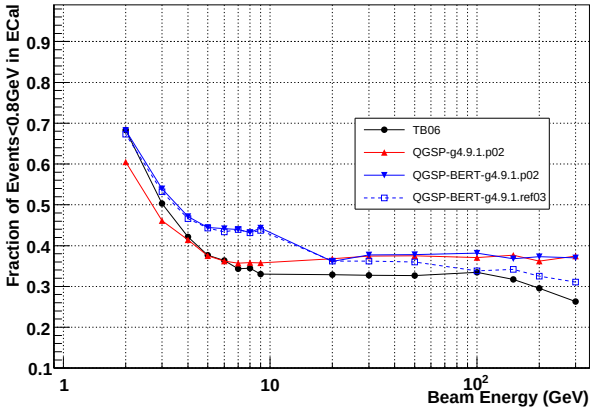


Figure 11: The fraction of π^-s with a MIP like signal in ECAL as a function of beam energy. Test beam results are compared with predictions of GEANT4 using the QGSP and QGSP_BERT physics lists. The disagreement below 10 GeV is attributed to an inadequate description of the quasi-elastic process in the Bertini model and is being investigated by the GEANT4 team.

2.2.1 Impact on Simulation Timing

The new physics list results in an improved agreement between the simulation and test beam results at the expense of an increased simulation time and larger event sizes. Samples of events of a few physics processes were produced using different physics lists: 100 minimum bias events, 20 $t\bar{t}$ and 20 $Z \rightarrow e^+e^-$. The CPU time and event size for the different samples are summarized in Table 1. As can be seen from the table, there is $\sim$40% increase in CPU time and $\sim$(50-80)% increase in event size. There is also $\sim$(5-10)% increase in memory usage.

A series of methods to reduce the simulation time and event size are currently being investigated. A significant fraction of additional hits occur at a delayed time and suppressing these hits would reduce the event size. The CPU time could be reduced by not tracking neutrons in GEANT4 that are outside of the time window of the detector readout and to eliminate the neutrinos from the treatment altogether.

	Minimum Bias	$t\bar{t}$	$Z \rightarrow e^+e^-$
CPU Time			
QGSP_EMV	1.00 (23.0 s)	1.00 (149.5 s)	1.00 (98.2 s)
QGSP	1.16	1.20	1.17
QGSP_BERT_EMV	1.41	1.46	1.36
QGSP_BERT	1.58	1.69	1.53
Event Size			
QGSP_EMV	1.00 (0.16 Mb)	1.00 (1.01 Mb)	1.00 (0.54 Mb)
QGSP	1.01	1.03	1.00
QGSP_BERT_EMV	1.52	1.77	1.60
QGSP_BERT	1.52	1.72	1.58

Table 1: CPU times and event sizes obtained using different physics lists for samples of minimum bias, $t\bar{t}$ and $Z \rightarrow e^+e^-$ events, relative to the QGSP_EMV physics list.

2.2.2 Impact on Jets

Enabling Birks' law and changing the default physics list affects the response of single particles. The impact on more complicated objects was studied by reconstructing jets. Separate QCD dijet samples with a $\hat{p_T}$ range of 80 - infinity GeV were produced using CMSSW 1.6.8 using the following 4 settings; i) Default physics list (QGSP_EMV) and Birks' law disabled, ii) Birks' law enabled, iii) QGSP_BERT, and iv) Birks' Law together with QGSP_BERT. Jets were reconstructed using the SISCone jet algorithm with a cone radius = 0.5.

Birks' law was enabled using the following settings:

```
replace g4SimHits.HCalSD.UseBirkLaw = true
replace g4SimHits.HCalSD.BirkC1 = 0.013
replace g4SimHits.HCalSD.BirkC2 = 9.6e-6
replace g4SimHits.ECalSD.UseBirkLaw = true
replace g4SimHits.ECalSD.BirkC1 = 0.046
replace g4SimHits.ECalSD.BirkC2 = 0.0
```

The default physics list was changed from QGSP_EMV to QGSP_BERT using the following setting:

```
replace g4SimHits.Physics.type = "SimG4Core/Physics/QGSP_BERT"
```

Figure 12 compares the inclusive jet p_T distribution for the four samples. The uncorrected calorimeter jet $p_T$is plotted with a $p_T >$ 30 GeV requirement. The jet p_T resolution is determined by exclusive [2)] matching the

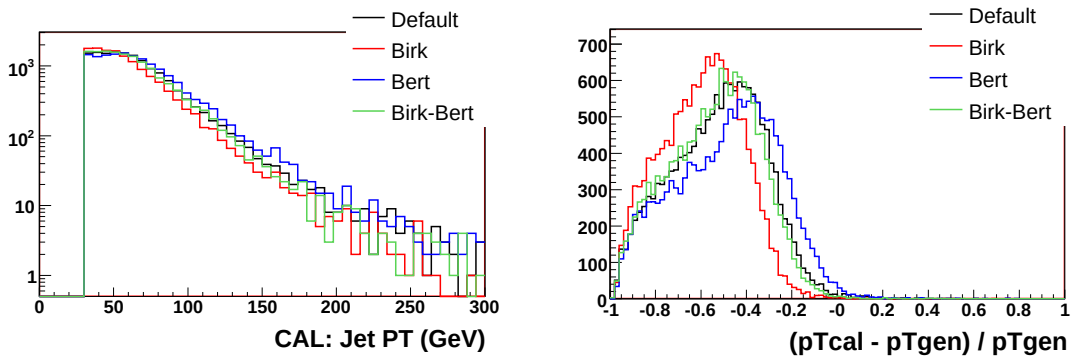


Figure 12: (left)The uncorrected inclusive jet p_T for jets reconstructed in the calorimeter (right) The jet p_T resolution.

reconstructed CaloJet to the GenJet (using $\Delta R = \sqrt{\delta\eta^2 + \delta\phi^2}$ matching). The p_T resolution distribution is also shown in Figure 12. The jet's position resolution is shown in Figure 13. Figure 13 also shows the jet p_T resolution as a function of the jet p_T . $\Delta p_T/p_T$ is determined for bins in p_T . The distribution is then fit to a Gaussian and the sigma is plotted.

Enabling Birks' law results in lowering the jet response, while using QGSP_BERT results in increasing the jet response. The effect on the jet position and p_T resolution of changing both the physics list from QGSP_EMV to QGSP_BERT and enabling Birks' law cancel out resulting in jets with similar resolutions as with the default physics list. For all results presented here the jet energy scale is determined using QGSP_EMV. In addition note that there is limited statistical sampling of the tails.

3 GFLASH in the Full Simulation

GFLASH is a parameterized simulation of electromagnetic (EM) and hadronic showers. It uses parameterizations to describe the longitudinal and lateral profiles in homogeneous and sampling calorimeters taking into account individual shower fluctuations and correlations. GFLASH provides a significantly faster calorimeter simulation than detailed GEANT4 based simulations, with greater flexibility of tuning the response to data. Details of the original GFLASH program can be found in Ref. [3] [12].

[2)] A GenJet is matched to one only CaloJet.

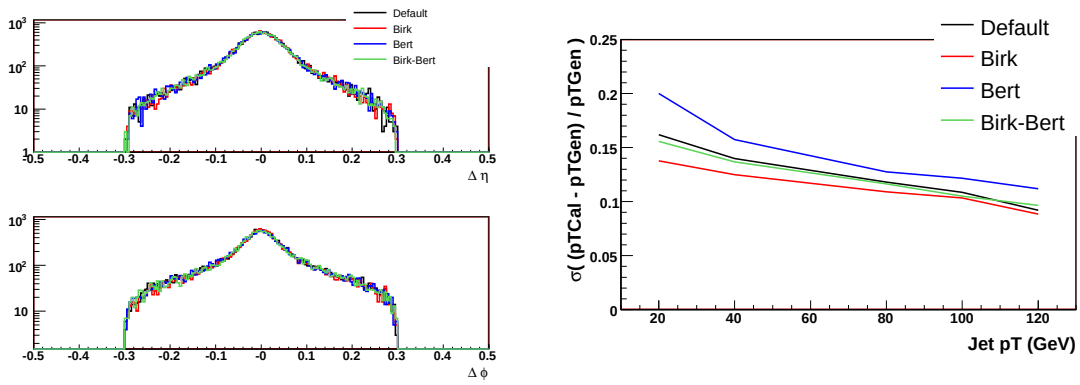


Figure 13: (left) Jet Position resolution (right) Jet p_{T} resolution as a function of jet p_{T} .

The simulation of showers in GFLASH is divided into the spatial distribution of deposited energy (E_{dp}) and the energy fraction of the deposited energy which is visible in the active medium. The spatial distribution of E_{dp} within the azimuthal symmetric calorimeter volume containing showers can be expressed as

$$dE_{dp}(\vec{r}) = \frac{E_{dp}}{2\pi} L(E,z) dz R(r,E,z) dr, \tag{1}$$

where L describes the longitudinal energy sub-profile along the shower depth z and R represents the lateral sub-profile with respect to the shower axis, and r is the radial distance perpendicular to the shower direction. The energy fraction of the deposited energy which is visible in the active medium is determined by the sampling fraction (or scaling factors determined by calibration) and their fluctuations.

In general, the EM shower parameterization can be applied to any homogeneous detector provided one can use the energy in units of critical energy (E_c) and the length in units of the radiation length (X_o). Since the CMS electromagnetic calorimeter (ECAL) consists of a homogeneous material (namely $\mathrm{PbWO_4}$) and is relatively long ($\sim 25 X_o$), the original GFLASH approach should work for the ECAL shower model without major modifications. We first tune both the longitudinal and lateral profile of GFLASH to those of GEANT4 to verify that the original parameterization is valid for the CMS detector and to study any CMS specific variations. Then, a minimal set of GFLASH parameters are changed to match the EM energy response in the $N \times N$ crystal tower between test beam data and simulated data of GFLASH.

The hadron shower parameterization is far more complicated due to π^0 contributions from both the primary and secondary inelastic hadronic interactions. The original GFLASH parameterization was developed for the H1 calorimeter assuming that the detector is described as a single effective medium within a repetitive sampling structure. The H1 parameterization, especially the longitudinal parameterization, is not suitable for the CMS calorimeter which has a hybrid structure consisting of crystal followed by the hadron calorimeter with some intermediate passive material. A new parameterization for the longitudinal hadronic shower is developed to take into account the specific structure of the CMS detector. The efffect of the strong magnetic field (4 T) is taken into account in the GFLASH simulation of the shower trajectory.

3.1 GFLASH Simulation of Electromagnetic Showers

The mean longitudinal energy profile of electromagnetic showers in GFLASH is described by

$$L = \frac{x^{\alpha-1} e^{-x}}{\Gamma(\alpha)}, \tag{2}$$

where $x = \beta z$ and z is the shower depth expressed in units of radiation length. The shower shape is characterized by a correlated pair of parameters, α and β, which correspond to the first and second moments of the Γ–distribution. The distribution has a maximum at $z_{max} = (\alpha - 1)/\beta$ along the longitudinal axis z. For each shower–i, a normally distributed correlated pair (α_i, β_i) is generated with the means (μ_α, μ_β), their fluctuations ($\sigma_\alpha, \sigma_\beta$) and the correlation ρ between α_i and β_i described by

$$\begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} = \begin{pmatrix} \mu_\alpha \\ \mu_\beta \end{pmatrix} + \begin{pmatrix} \sigma_\alpha & 0 \\ 0 & \sigma_\beta \end{pmatrix} \begin{pmatrix} \rho_+ & \rho_- \\ \rho_+ & -\rho_- \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} \tag{3}$$

where $\rho_{\pm} = \sqrt{(1 \pm \rho)/2}$ and ϕ_1 and ϕ_2 are normally distributed random numbers.

The lateral profile of energy distribution for both electromagnetic and hadronic showers is described by a simple ansatz,

$$R = \frac{2rR_0^2}{(r^2 + R_0^2)^2}, \tag{4}$$

where r, is in units of Moliére radius to the shower direction and with origin in the center of the shower, and $R_0(E, z)$ is a function of shower energy (E) and shower depth (z). The mean and the variance of $R_0(E, z)$ are parameterized as

$$\langle R_0 \rangle (E, z) = [R_1 + (R_2 - R_3 \log(E))z]^n, \tag{5}$$

$$V_{R_0}(E, z)/\langle R_0 \rangle^2 = [S_1 - (S_2 \log(E))(S_3 + S_4 z)]^2, \tag{6}$$

where n=2(1) for electromagnetic (hadronic) showers. R_1 represents the core component of the shower that is independent of the energy or depth. The term $(R_2 - R_3 \log(E))$ represents the spread of the energy deposition along the shower depth z. The relative fluctuation of a shower, $\sqrt{V_{R_0}(E, z)}/\langle R_0 \rangle$ decreases with energy while increasing with shower depth.

3.1.1 Tuning to the H4 Test Beam Data

In order to tune the EM shower shape and response of GFLASH, we use the data files from the H4 test beam taken in 2006. For each electron beam energy ($E_{\rm beam}$ = 20, 30, 50, 80, 120, 150 GeV), the nominal energy response in $N \times N$ crystals is obtained from the energy scan data on super module 16. Optimized weights are used for the amplitude reconstruction and the crystals are intercalibrated. The reference response of the detector is obtained by selecting only the electrons impacting close to the point of maximum containment which is determined for each crystal in order to disentangle the effects due to not fully containing the shower [11].

Simulated data using the full GEANT4 and GFLASH simulations are generated using the same set of configuration files for simulation and reconstruction;

```
Configuration/EcalTB/data/reco_application_tbsim_DetSim-Digi.cfg
Configuration/EcalTB/data/reco_application_tbsim_localreco.cfg
```

and they are then analyzed using the same analysis software as for the test beam data. Figure 14 and Figure 15 show the energy response in 5×5 crystals and the ratio of energy in 1×1 crystal to 5×5 for 50 GeV electrons. The overall energy spectrum as well as the lateral shower profiles are described by the tuned parameters of GFLASH for the full range of energies studied. These parameters are included in CMSSW_2_1.

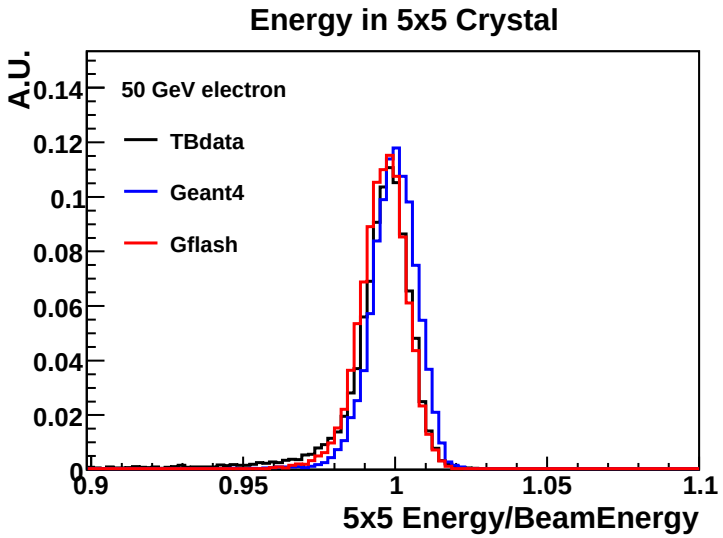


Figure 14: The energy response in 5×5 crystals between simulation (GFLASH and GEANT4) and test beam data for 50 GeV electrons.

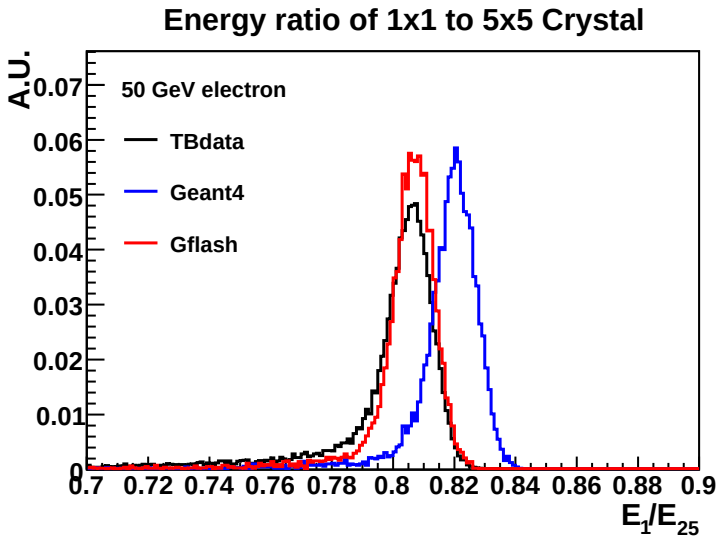


Figure 15: The ratio of energy in 1×1 crystal to 5×5 between simulation (GFLASH and GEANT4) and test beam data for 50 GeV electrons.

3.2 GFLASH Simulation of Hadronic Showers

In what follows we assume that the intrinsic representation of a shower in (r, z) for a given E in each detector region (ECAL, HCAL) is not changed by the starting point of the shower (i.e. albedo effects are ignored). The longitudinal profile of hadron showers can be then written as a combination of sub-profiles in two detector regions,

$$F = f_{\rm ECAL} F_{\rm ECAL} + f_{\rm hcal} F_{\rm hcal}, \tag{7}$$

where $F_{\rm ECAL}$ and $F_{\rm hcal}$ represent the hadronic shower profile in the ECAL and HCAL respectively. The coefficient f_i ($i = {\rm ECAL},\ {\rm hcal}$) is related to the fraction of the deposited energy in each detector,

$$E_i = E_{\rm inc} \int f_i F_i dV, \tag{8}$$

where dV only spans the region in each detector volume and $E_{\rm inc}$ and E_i are the energy of the incident particle and the deposited energy in each detector respectively. The $f_{\rm HCAL}$ parameter may depend on the starting point of the shower (z_s) while $f_{\rm ECAL}$ is assumed to be a constant..

The amount of energy deposited in the ECAL only depends on z_s and is determined by integrating $F_{\rm ECAL}(z - z_s)$ from z_s to the maximum z of the ECAL for an individual shower. The parameterization of $F_{\rm HCAL}$ depends on z_s when the shower starts within the HCAL volume. The amount of energy deposited in the HCAL is simulated by integrating $F_{\rm HCAL}$ from max(z_s,$z_{\rm min}$ of HCAL) to the maximum z of HCAL. In the current GFLASH model for hadronic showers, both $F_{\rm ECAL}$ and $F_{\rm HCAL}$ assume the same functional hypothesis consisting of the superposition of two Gamma distributions:

$$F_{\rm ECAL}({\rm or} F_{\rm HCAL}) = [cL(x_e; \alpha_e, \beta_e)dx_e + (1-c)L(x_h; \alpha_h, \beta_h)dx_h], \tag{9}$$

$$L = \frac{x_i^{\alpha_i - 1} e^{-x_i}}{\Gamma(\alpha_i)}, \quad x_i = \frac{\beta_i z}{d_i} \qquad (i = e, h), \tag{10}$$

where c is the fraction of the deposited energy by any π^0 produced from primary or secondary interactions, $L(x_e; \alpha_e, \beta_e)$ and $L(x_h; \alpha_h, \beta_h)$ are sub-profiles of the spatial energy distribution by π^0 and by any charged hadron tracks in the units of d_i ($d_e = X_o, d_h = \lambda_o$) respectively. In the original GFLASH parameterization of hadronic showers [3], the π^0 component is sub-divided into two parts. The contribution by π^0 produced from later interactions was separated from the contribution by the π^0 produced at the first inelastic interaction. Due to the passive material between the ECAL and the HCAL, the single set of longitudinal profiles (three Γ distributions) representing the entire calorimeter used in the original GFLASH approach may not be suitable to describe the CMS calorimeter. For this reason we introduce two sets of longitudinal profiles, $F_{\rm ECAL}$ and $F_{\rm HCAL}$ (four Γ distributions in total). While two Γ distributions for each calorimeter may be insufficient to provide a description of the longitudinal profile, their combination when the shower starts in the ECAL and extends to the HCAL should provide the

flexability to tune the GFLASH response to the data within the desired precision. Additional sub-profiles may still be introduced if needed.

Using the modified functions, we then parameterize the set of inputs describing the longitudinal profile. In particular, α and β of each Γ distribution can be semi–analytically estimated by using the mean ($m_1 = \alpha/\beta$) and variance ($V = \alpha/\beta^2$) of the distribution. To build an initial set of parameters, $\vec{x} = (c, \alpha_e, \beta_e, \alpha_h, \beta_h)$, we use the full CMS detector simulation with the default GEANT4 physics list.

3.2.1 Tuning to the H2 Test Beam Data

In order to tune the hadronic parameterization of GFLASH to 2006 test beam (H2) data, we follow the same procedure (framework, configuration files and analysis code) that was used for tuning the GEANT4 simulation of the test beam. We compare the result from GFLASH to the reference response of the 2006 pion test beam data. The file lists energy in EB (7x7 matrix), HB (4x3), HO (3x2) on one line per each event. Additionally we take into account the measured π/e ratio for the HCAL response. For each beam energy, we compare the total energy, the MIP energy, the EM energy and the HAD energy. The MIP energy is defined as the total (EM+HAD) energy for pions selected to be minimum ionizing in the EM section ($E_{EM} < 0.80$ GeV). Figure 16 shows comparisons of energy responses between simulation (GFLASH and GEANT4) and 20 GeV/c pion test beam data. The result

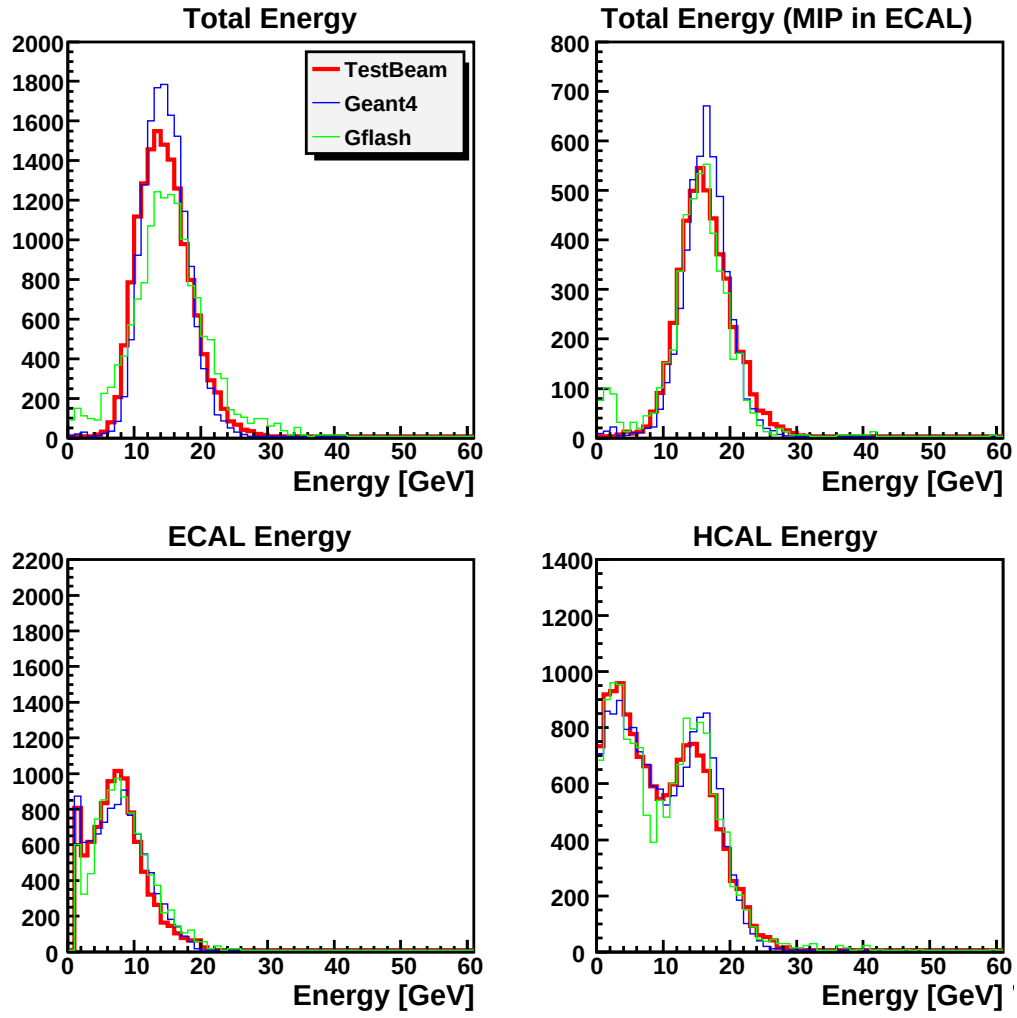


Figure 16: Energy responses between simulation (GFLASH and GEANT4) and 20 GeV pion test beam data: Total energy, MIP energy, EM energy, Had energy

of GFLASH was obtained with the earlier parameterization based on the average shower profile and its uncertainty before we implemented the procedure taking into account individual shower fluctuations and their corrections described in the previous section. The improved parameterization of hadron showers as a function of momentum and the shower starting point is nearly finalized and tuning GFLASH to H2 test beam data is in progress.

The steps to tune the longitudinal profiles of GFLASH to the test beam data is to tune 1) the ECAL response 2) the HCAL response and 3)the energy resolution and response for pions which are MIPs in the ECAL. We expect that the total energy response and the correlation between EM energy and HAD energy will be tuned simultaneously. The lateral profile of the hadron showers will be iteratively re-tuned.

4 Fast Simulation

FastSim provides a fast and fairly accurate simulation of the CMS detector. The same set of event generators available in FullSim are interfaced to FastSim. The event format after passing through the detector simulation is the same for both FullSim and FastSim and the same offline reconstruction code can be run on the data produced by the two different simulation packages. FastSim uses the same geometry description as FullSim. Whereas FullSim can take several minutes to simulate an event, FastSim can simulate the same event in seconds. To achieve this performance, a number of assumptions and simplified parameterizations describing showering are used. The details of the calorimeter shower development, both electromagnetic and hadronic, is parameterized in FastSim and is being tuned to the detailed GEANT4 CMS simulation. Experience gained from tuning FastSim to FullSim will be vital in quickly providing a more accurate detector simulation by tuning to the LHC collision data.

FastSim can also be tuned directly to the test beam data. To mimic the test beam configuration we deactivate the material effects of the tracker and switch off the magnetic field. In addition we adapt the particle gun input parameters to scan the region around the EM crystal. Different options have been used for the H2 (hadrons) and H4 (electrons) test-beam configurations. For timing performance and other technical reasons, no simulation of the hodoscopes that monitor the beam position and slope were implemented in FastSim and as a consequence the coordinates system used in FastSim and data studies are different: in FastSim the coordinate system is centered on the front face of a given crystal, while the hodoscope coordinates are used for the data. The axis of the two coordinate systems are identical and would be equivalent if the beam had no slope.

4.1 Electromagnetic Calorimeter Response

The electromagnetic shower simulation is performed in two steps. In the first step, the electron shower is developed following the GFLASH parameterization [12] where ECAL is treated as though it was a homogeneous medium. In the second step, the shower is translated onto the ECAL geometry. During this step, several detector effects, such as energy loss between the ECAL modules, shower spreading due to the magnetic field, longitudinal non-uniformity, and electronic noise are accounted for. With the detector effects properly included in the simulation, no additional tuning of the single shower parameterization was required to get an agreement with FullSim at the percent level; the Moliére radius of the lead tungstate had to be enlarged by about 10% to match the lateral profile predicted by the FullSim at the time when FastSim was first developed.

The data collected during the summer of 2006 at the H4 test beam have been used to intercalibrate the crystals within the super-modules. The absolute energy scale was not established and as a result, the total energy deposited in the ECAL cannot be directly compared with the simulation, however the lateral profile of the electron shower can be used to tune the simulation. The ratio of the energy (E1) deposited in the crystal the electron first enters over the energy in the 3x3 crystal window (E9) centered on this crystal is used to tune the shower lateral profile.

In the following, results for crystal number 248 are used for both the data and the simulation of a 120 GeV beam energy. This crystal is located in the middle of the first module of the calorimeter at a position away from the module boundaries. The measured energy in the crystal is strongly dependent on the impact point of the beam. For a given incident angle, the beam position resulting in the maximum energy in a given crystal defines the point of maximum containment $(\mathrm{X_{MC}}, \mathrm{Y_{MC}})$. The area of maximum containment is defined as the $(4\ \mathrm{mm})$x(4 mm) square centered on the point of maximum containment. The dependency of the energy contained in the crystal on the impact position is shown in Figure 17. The point of maximum containment corresponds to the maximum of the two curves.

The point of maximum containment has been determined in FastSim and the corresponding curves are presented in Figure 18. Although the bounds and the origin are different from those of Figure 17, the shapes are similar. The points spread below the plateau in the data are believed to be related to Bremsstrahlung photons contaminating the electron beam. The sensitivity to the lateral profile is related to the height of the plateau and it has not been attempted to reproduce this source of beam impurity.

Once the point of maximum containment has been determined, the ratio E1/E9 is plotted using events for which the electron impact point is within the area of maximum containment. In the case of data, about 83.5% of the energy contained in 3x3 crystal window is contained in the central crystal, as shown in Figure 19. In the case of FastSim,

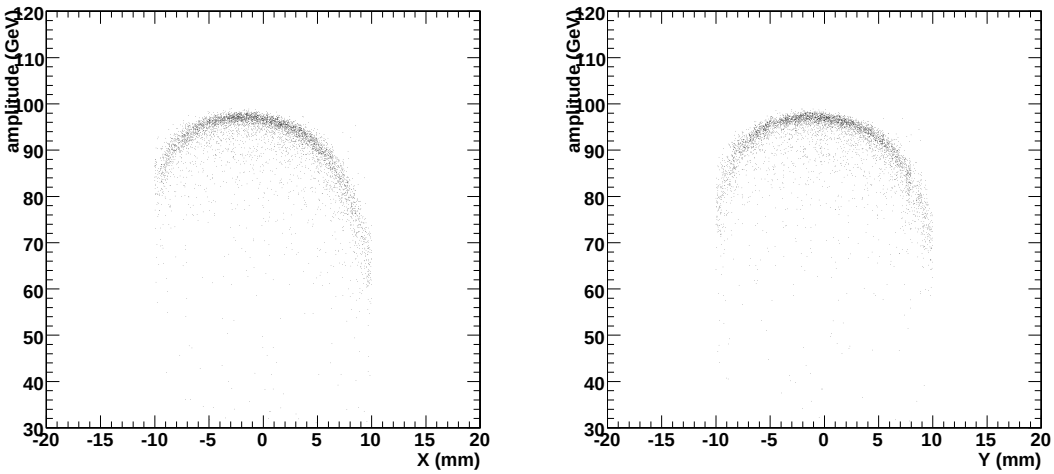


Figure 17: On the left (resp. right) plot, energy in the central crystal as a function of the X (Y) coordinate for $|\mathrm{Y} - \mathrm{Y_{MC}}| < 2\,\mathrm{mm}$ ($|\mathrm{X} - \mathrm{X_{MC}}| < 2\,\mathrm{mm}$) for 120 GeV electrons at the H4 test beam.

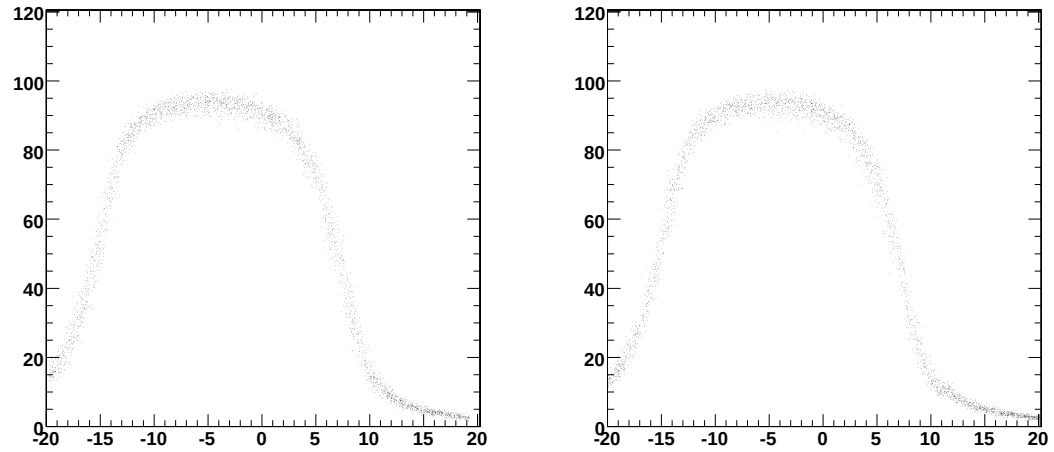


Figure 18: On the left (right) plot, energy in GeV in the central crystal is plotted as a function of the X (Y) coordinate in mm for $|\mathrm{Y} - \mathrm{Y_{MC}}| < 2\,\mathrm{mm}$ ($|\mathrm{X} - \mathrm{X_{MC}}| < 2\,\mathrm{mm}$) for 120 GeV electrons with FastSim.

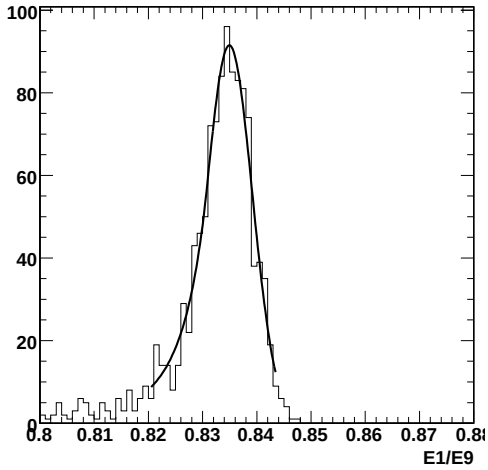


Figure 19: Distribution of the E1/E9 ratio on the test-beam data.

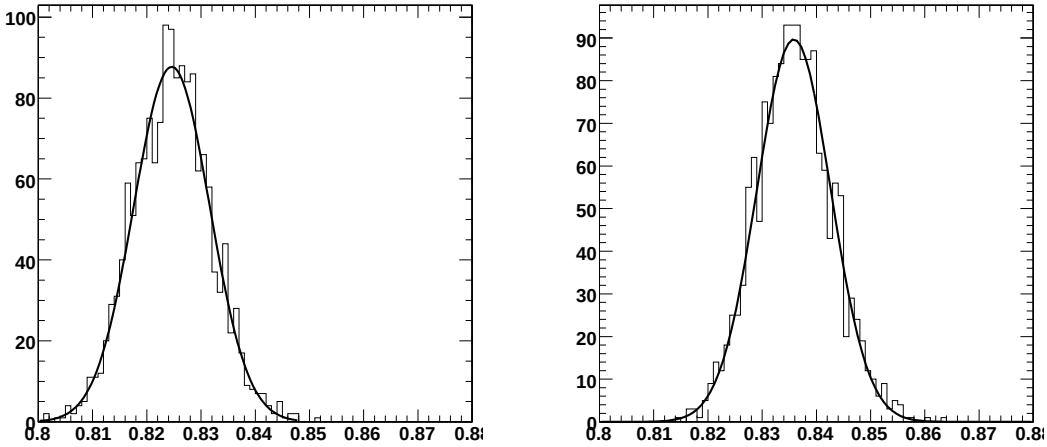


Figure 20: Distribution of the simulated E1/E9 ratio with FastSim tuned to the (old) FullSim data (left) and with the default setting of the Moliére radius (right).

using the default settings in CMSSW 1.8.4, the fitted mean value is 82.4% (see left plot of Figure 20) indicating that the simulated electromagnetic showers are slightly too wide. As mentioned earlier, the Moliére radius of the lead tungstate had been artificially enlarged in order to match the lateral profile obtained in the GEANT4-based simulation. When the Moliére radius is put back to the measured value, the E1/E9 ratio peaks at 83.4% (see right plot of Figure 20), in much better agreement with the test-beam data. A better understanding of the electron beam would be required to explain the differences in shapes between the data and the simulation.

As a result, the original Grindhammer parameterization used in FastSim provides a good description of the the lateral shower spread when its default parameters are used. The study should be extended, using data recorded at different beam energies, and the η dependency should also be studied. In addition the agreement of E1/E25 and E9/E25 should be checked.

4.1.1 Technical Implementation and Future Plans

Code to allow the Moliére radius to be modified is already in place in FastSim and the value giving the best agreement with the beam-data can be enabled by a simple parameter change. More options to control the lateral development of the core and the tail independently for electromagnetic showers have also been incorporated.

In addition to a more complete study of the lateral profile, for different values of η and energy, the data collected at

Table 2: Values of parameters for the "simple" simulation option

HCAL subdetector and particle type	Stochastic term	Constant term
HB hadrons	1.22	0.05
HE hadrons	1.30	0.05
HF hadrons	1.82	0.09
HF e-γ	1.38	0.05

the H4 test beam should allow for the so-called "umbrella effect" and the energy loss in the module boundaries to be tuned. The umbrella effect refers to the variation of the fraction of the incident particle energy reconstructed in a cluster as a function of η. This fraction decreases slightly with increasing η, leading to an umbrella shape when it is plotted as a function of η.

4.2 Hadronic Calorimeter Response

Currently there are three options to simulate the particle response in HCAL in FastSim:

- *simple*: Point-like energy deposited in HCAL only for hadrons (without ECAL), with no energy deposit from muons, and a simple analytical formula for the energy response using a Gaussian resolution;
- *medium*: idem, but the spacial energy distribution (both in HCAL and ECAL) is parameterized;
- *advanced*: idem, but with detailed tabulated single particle response (mean and Gaussian sigma) derived from FullSim.

The response in FullSim to pions, muons and e-γ (the latter - for HF only) has been determined using different versions of CMSSW. In the FullSim reconstruction of the single particle response the noise is turned off both in ECAL and HCAL, since in FastSim, the noise in both detectors is added on top of the simulated signal. In addition, only pions interacting in the calorimeters are selected, as particle interactions in the beam pipe and in the Tracker are simulated in FastSim. The reconstructed signal is collected in a 5×5 EcalPlusHcalTower matrix for all particle types. The Gaussian means and sigmas are determined for distributions sliced into η bins with a width of 0.1 (for hadrons) or divided into η regions (for muons) as a function of the particles' energy. The tabulated response is then used in FastSim.

4.2.1 Simple Response Option Parameters

The "simple" option uses the following formula for the energy response as a function of the initial particle energy:

$$E_{\rm resp} = E_{\rm in} \cdot p(1 + c \cdot \exp(n \cdot \log(s/E_{\rm in}))), \tag{11}$$

where $E_{\rm resp}$ denotes the energy response of the HCAL to particles with an incident energy of $E_{\rm in}$; p, c, n and s - parameters equal to 0.95, 1.0, 1.0 and 3 GeV respectively. Table 2 shows the values of the stochastic and constant terms in the formula for the energy resolution.

4.2.2 Advanced Response Simulation

The Gaussian mean and sigma of the hadron response are tabulated in a grid of particle energy and η. The energy values for HB/HE are: 1, 2, 3, 5, 9, 15, 20, 30, 50, 100, 150, 300, 1000 and 3000 GeV. All energies, except the two last ones match the available test beam data to facilitate tuning of the response and shower parameters to the data.

Linear interpolation (extrapolation) is used for the mean and sigma in (outside) the available energy range. Figure 21 shows a comparison of the transverse energy of single pions of $p_{\rm T}$=100 GeV/c as a function of pseudo-rapidity as reconstructed with FastSim (filled triangles) and FullSim (filled squares).

Once the deposited energy of the particle is simulated, it is distributed in the calorimeter using parameterized longitudinal and lateral shower profiles with variations of the shower shape and starting point based on the GFLASH approach [3, 12]. Figure 22 shows a simplified scheme of the hadron shower simulation. The shower starting point

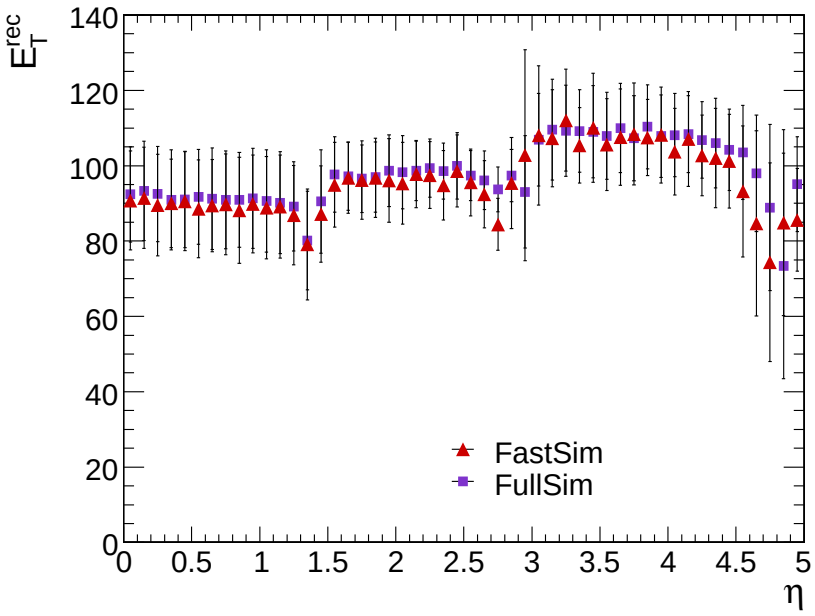


Figure 21: The reconstructed transverse energy of pions with p_T=100 GeV as a function of pseudo-rapidity. Triangles denote FastSim, while squares represent the FullSim results.

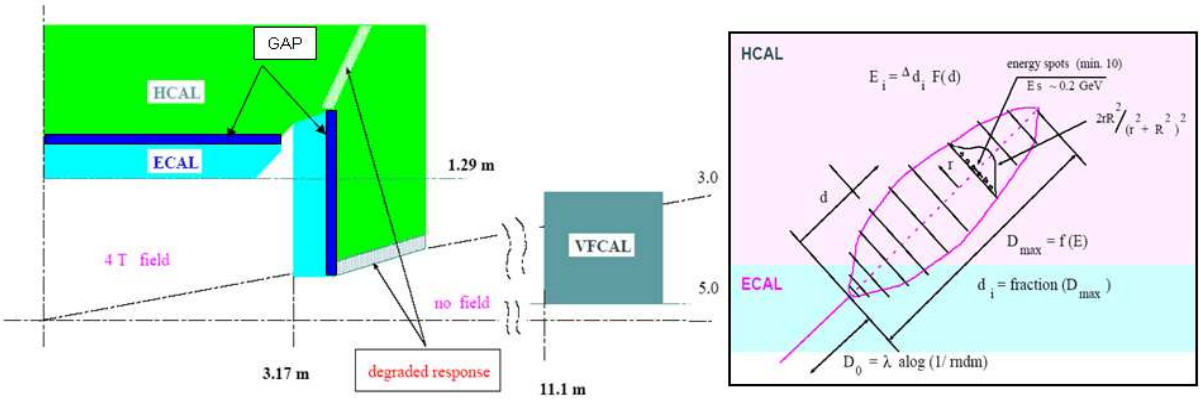


Figure 22: Hadron shower simulation sketch.

fluctuates according to the exponential law in terms of nuclear interaction lengths (λ). The energy spots are distributed according to parameterized longitudinal and transverse shapes.

There are some differences in the actual simulation in FastSim compared to the original GFLASH approach in order to speed up the execution time and to better tune the shower shape and the ECAL fraction of energy to results from FullSim. Figure 23 shows the comparison of the transverse shower shape between FastSim (dashed line) and FullSim (solid line) for pions of 100 GeV/c incident on the center of the HB tower. The upper plot shows the distribution of energy collected in this tower. The middle plot shows the distribution of energy collected in a 3×3 matrix of towers. The lower plot shows the ratio of the energy collected in the tower with respect to that collected in the 3×3 matrix.

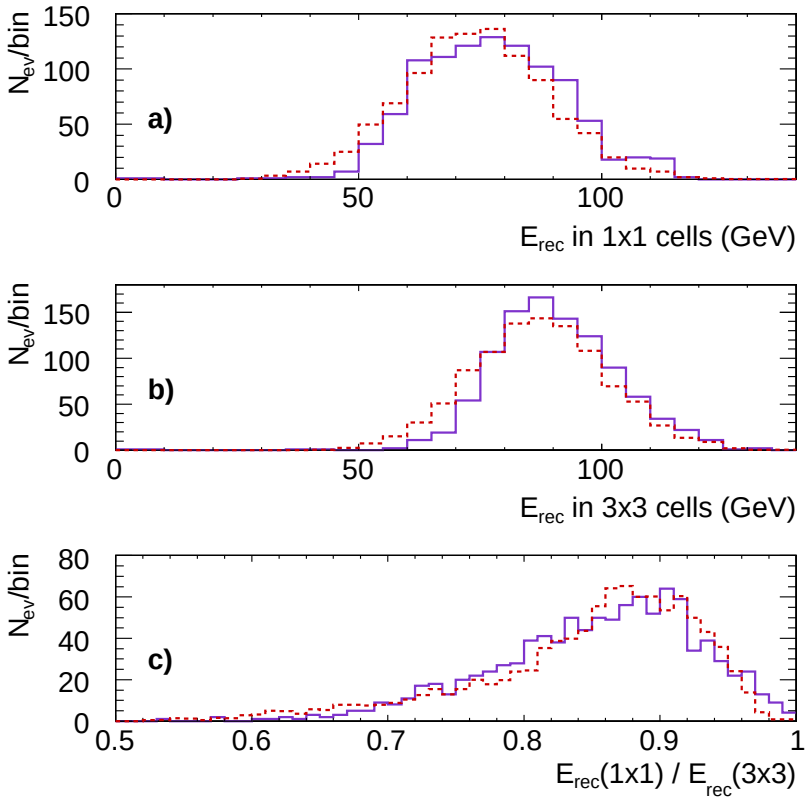


Figure 23: Transverse shower shape distributions for 100 GeV/c pions. More detail can be found in the text.

The muon response is simulated in FastSim in a similar way as for pions - on a grid of transverse momentum and pseudo-rapidity with interpolation/extrapolation. The p_T points are 10, 30, 100, 300 GeV. The response of HF to electromagnetic particles is evaluated with the full simulation for electrons with energies 30, 100, 300, 1000 and 3000 GeV as a function of pseudo-rapidity and is implemented in FastSim.

4.2.3 Jet reconstruction comparison between fast and full simulation

A comparison of reconstructed jet E_T between FastSim and FullSim using the Iterative Cone algorithm with a cone size $R = 0.5$ is shown in Figure 24. The ratio of the jet E_T energy to the Monte-Carlo E_T as a function of jet pseudo-rapidity for Monte-Carlo E_T=100-110 GeV is plotted. The agreement between FastSim and FullSim jets is satisfactory throughout the entire pseudorapidity and jet p_T range. The goal was to reproduce FullSim jets to within 5% over a broad range of jet p_T. The jet response in FastSim can be tuned by updating the single pion response tables used in FastSim to adjust for future changes in FullSim.

4.3 Tuning the Hadronic Response

4.3.1 Test Beam set-up tuning in FastSim and first comparisons

To fully benefit from the increased speed of simulating events using FastSim, especially in the context of producing large QCD multijet backgrounds, it is essential that the hadronic shower treatment encapsulates the relevant

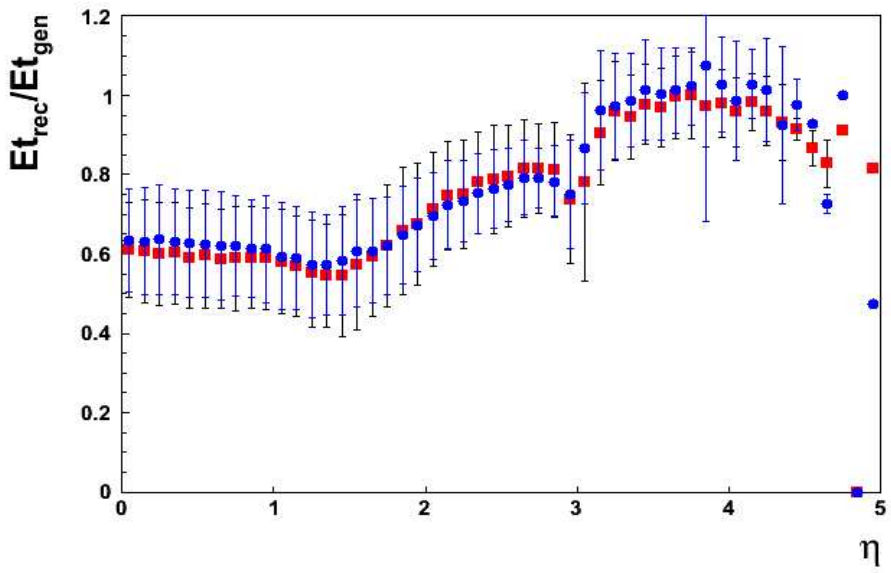


Figure 24: Ratio of the jet E_T energy to the Monte-Carlo E_T as a function of jet pseudo-rapidity for Monte-Carlo E_T=100-110 GeV. Filled squares denote the FastSim results, while filled circles represent the FullSim results.

physics. In the scope of the Calorimetry Simulation task force we are implementing shower parameterizations in FastSim and tuning them using the H2 Test Beam data.

For the transverse shower profile, the 2006 H2 test-beam data gives us a relatively good handle on the shower development. Here, an ECAL barrel supermodule (EB) was combined with a 4×3 matrix of HCAL barrel (HB) cells, along with a 3×2 matrix of the HCAL outer (HO) cells in the final CMS configuration with the DAQ and monitoring services. This set-up was placed in the H2 SPS beam-line, at the correct angle of incidence, to reproduce the calorimetric coverage of CMS at $\eta \approx 0.22$. Beam-data was acquired with a variety of beam configurations, including π^+, π^-, p and $\bar{p}$ mono-energetic beams over many different momenta ranging from a few GeV to a few hundred GeV.

In order to properly use the Test Beam data, we reproduce the test-beam set-up in FastSim, requiring the turning off of the tracker geometry which was not present in the test-beam. Additionally, special treatment of calorimetric noise, and beam conditions (spread and slope) is required in order to fully mimic the test-beam setup, where there was no zero suppression (included by default in FastSim due to CPU time constraints) and since the beam profile from the CMS interaction point does not naturally correspond to that in the H2 beam-line. In our study we assume the noise to be normally distributed, with the magnitude cross-checked with CRUZET data. Since the information from the wire chambers was not included in the data files we had access to, we implement a realistic beam profile by smearing the event primary vertex in the CMS z and in the CMS x/y plane perpendicular to the center of the test-beam setup, in order to recreate the beam spread in η and ϕ, respectively, since the beam profile from the CMS interaction point does not naturally correspond to that in the H2 beam-line.

Figure 25, 26, 27 and 28 illustrate the FastSim test beam setup tuning for the 300 GeV/c π^- beam. Similar tuning is performed for all of the different beam configurations studied. This is necessary, since systematic differences in the beam tuning will bias our understanding of the transverse beam profile.

Figures 29 and 30 show the electromagnetic fraction distribution using the current version (not tuned) of FastSim compared to the H2 TB data for 300 GeV/c π^- while figures 31 and 32 show the so-called "banana" plot of the HCAL vs ECAL energies for 5 GeV/c π^- and again for FastSim and the H2 TB data. These show a qualitative agreement that we should maintain after the tuning program is applied.

4.3.2 Tuning and Optimization

The transverse shower parameterization can be expressed as:

$$R(r) = \frac{2rR_{50}}{(r^2 + R_{50}^2)^2}, \tag{12}$$

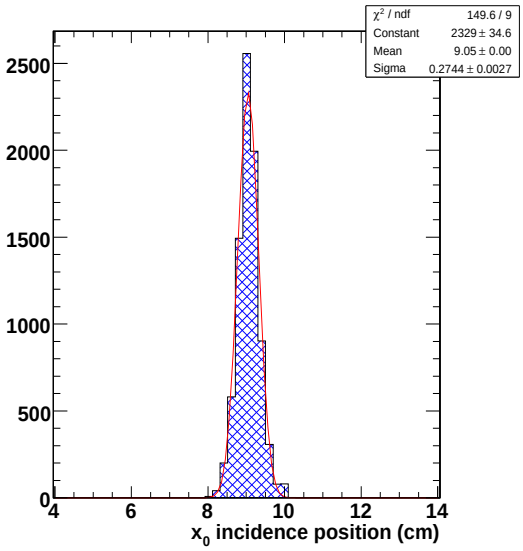


Figure 25: Beam profile in x (η) from the FastSim test beam simulation for 300 GeV/c π^- beam

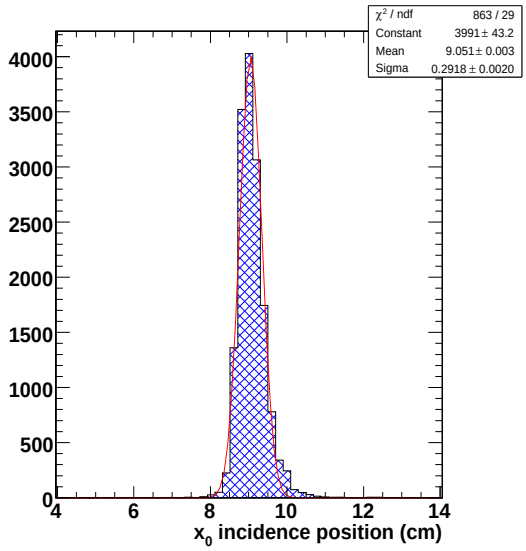


Figure 26: Beam profile in x (η) from H2 2006 TB data for 300 GeV/c π^- beam

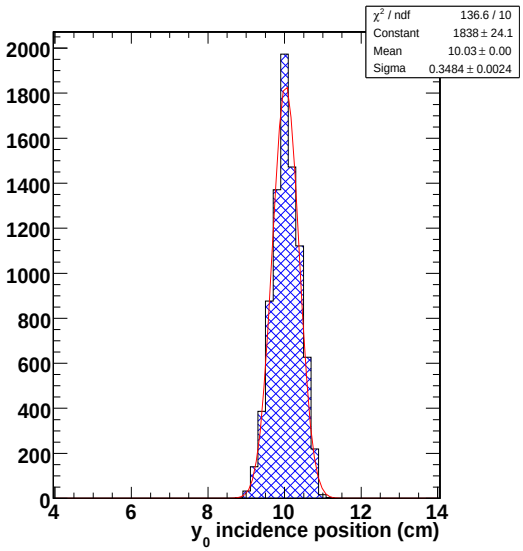


Figure 27: Beam profile in y (ϕ) from the FastSim test beam simulation for 300 GeV/c π^- beam

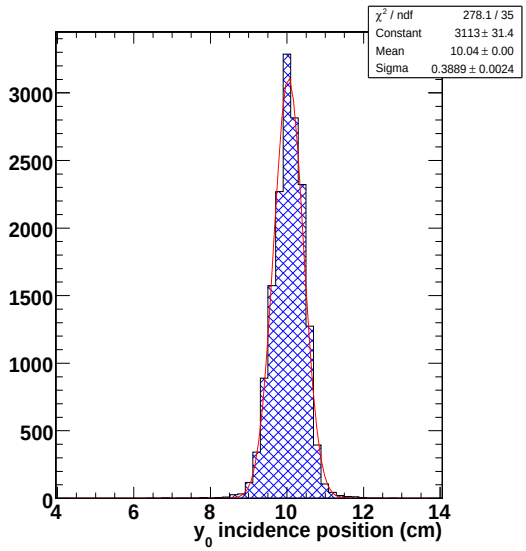


Figure 28: Beam profile in y (ϕ) from H2 2006 TB data for 300 GeV/c π^- beam

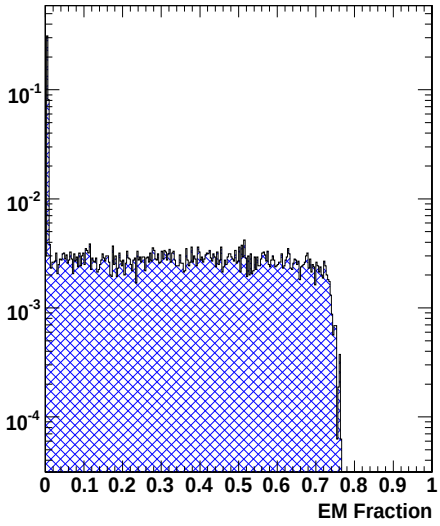


Figure 29: Electromagnetic fraction from the FastSim test beam simulation for 300 GeV/c π^- beam. Large EM fractions are not well represented currently in FastSim

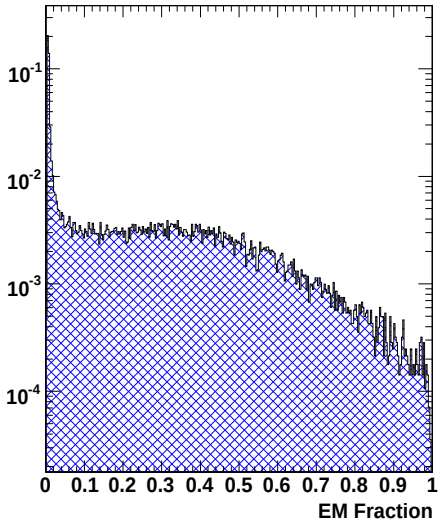


Figure 30: Electromagnetic fraction from H2 2006 TB data for 300 GeV/c π^- beam. Tuning of the transverse shower profile (and ECAL/HCAL energy sharing) should result in better agreement for this variable.

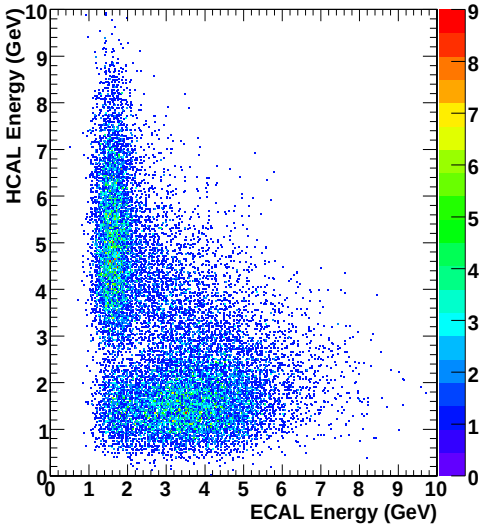


Figure 31: Reconstructed HCAL vs ECAL energies from the FastSim test beam simulation for 5 GeV/c π^- beam. The qualitative behavior is reasonably well captured in FastSim.

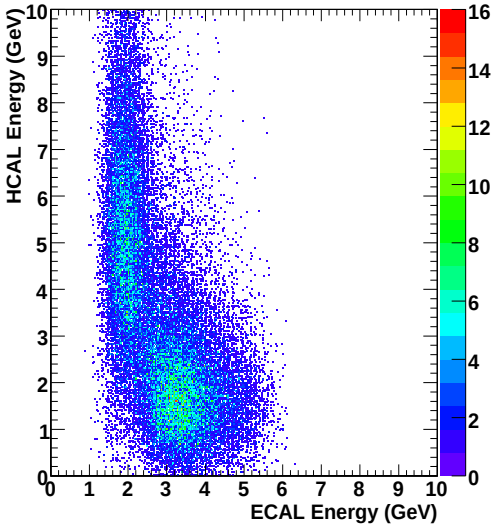


Figure 32: Reconstructed HCAL vs ECAL energies from H2 2006 TB data for 5 GeV/c π^- beam. It should be ensured that tuning of the transverse shower profile (and ECAL/HCAL energy sharing) should maintain this qualitative agreement for the low energy beam configurations.

where R_{50} and r are measured in units of Moliére radius. The energy and shower depth dependence of R_{50} are further represented by another 7 parameters, describing not only the mean but also the variance of the variable. In the Test Beam data there is no explicit information describing the transverse shower, but the granularity of the ECAL can be used in order to probe the shower shape. Specifically, the observable we construct is:

$$F(r) = \int_{z_{begin}}^{z_{end}} dz\, t(z) \int_0^r dr' s(r'), \quad f(r) = F(r)/F(\mathcal{R}), \tag{13}$$

where z_{begin} and z_{end} correspond to the depth of the hadronic shower starting point and back of the ECAL, respectively, and $t(z)$ describes the longitudinal shower profile. $\mathcal{R}$ corresponds to a large radius outside the spread of the shower, so that $f(r)$ is a normalized, cumulative radial energy distribution, as shown in Figure 33 for different π^- beam configurations.

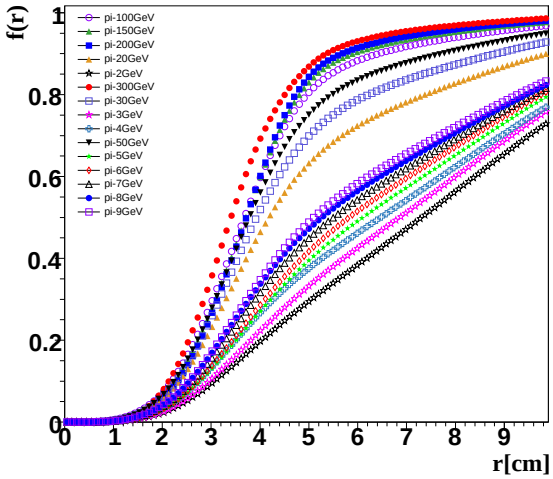


Figure 33: $f(r)$ normalized, cumulative radial energy distribution in the ECAL for a range of π^- beam configurations.

Due to the shower depth dependence of $f(r)$, and variations in the shower starting point between events, it is not viable to simply solve for the 7 parameters describing R_{50} analytically. Rather, we use the FastSim test beam setup as a toy Monte-Carlo, populating a space of the R_{50} 7 parameters through simulation of the different beam configurations and by varying the relevant parameters. Using the $f(r)$ observable to compare between FastSim and data via a χ^2 test then allows us to optimize the desired radial shower parameterization.

The values of R_{50} are distributed according to a log-normal probability density finction. Writing the mean and the variance parameters as,

$$\langle \mathcal{R} \rangle (E, z) = R_1 + (R_2 - R_3 \log(E))z, \tag{14}$$

$$Var[\mathcal{R}](E, z) = [(S_1 - (S_2 \log(E))(S_3 + S_4 z)\langle \mathcal{R} \rangle]^2, \tag{15}$$

we can fit for the parameters $S_1, ...S_n$ and $R_1, ..., R_n$ using the test beam data collected at different energies. Since the energy in each run is fixed, we can reproduce the dimensionality of the problem by writing the mean and sigma as

$$\langle \mathcal{R} \rangle = R_1 + \tilde{R}_2 z, \tag{16}$$

$$Var[\mathcal{R}] = (\tilde{S}_1 + \tilde{S}_2 z)^2 \langle \mathcal{R} \rangle^2. \tag{17}$$

In our optimization study we fit for $\tilde{S}_1$, $\tilde{S}_2$, R_1, and $\tilde{R}_2$. When a set of optimal values is available for each energy, we can obtain $S_1, ...S_n$ and $R_1, ..., R_n$ by fitting for the energy dependence.

The information available in the test beam data is insufficient to perform an analogous optimization of the longitudinal parameters. The varying π^0 content of the shower, variable starting depth and differences in non-compensation between the ECAL and HCAL material result in fluctuations of the ratio of the measured electromagnetic and hadronic portions of the shower event-by-event which is the primary handle for a longitudinal optimization program. To overcome this difficulty, a hybrid scheme using the full simulation is necessary in order to understand

how the hadronic shower content affects this ratio, and how to classify this content using the observable in the test beam in order to isolate the different effects. This is a useful exercise given that the available handles on the data are essentially identical to what will be present in *in situ* data from isolated tracks incident on the calorimeter.

We expect that the hadronic shower optimization that we are performing for FastSim is easily generalizable to be used in FullSim and will provide one of the first opportunities for tuning using the first LHC collision data. One should note that given the content of the available test beam data and the fact that we use FastSim for the simulation of the longituninal shape of the shower the approach we are following might not be adequate in that the reduced parameter space cannot be tuned to reproduce the development of the shower given that we do not know the beginning point of the shower and the relevance of the parametrization at a few or more interaction lenghts from its onset; in this case we would need to go back to the design of the parametrization and the understanding of the shower from first principles.

5 Using Collision Data to Tune the Simulation

The CMS detector simulation will be tuned with the LHC collision data. A large sample of isolated charged particles is needed to calibrate the HCAL detector. Multijet QCD production is the primary source of single charged particles and a dedicated trigger has been developed to collect large samples of them. These calibration events, referred to as "AlCaRaw", are collected by a High Level Trigger (HLT) path with a prerequisite on the Level-1 trigger. The Level-1 prerequisite consists of the logical `OR` of a Level-1 single jet trigger and a Level-1 single τ-jet trigger. Prescales will be used to keep the Level-1 trigger rate to $\sim$ 4.5 kHz for an instantaneous luminosity of 2×10^{30} - 2×10^{31} cm^{-2}s^{-1}.

The selection at the HLT requires a Level-1 τ-jet object not matched to the Level-1 tiggering jet object. Furthermore, an isolated pixel track with $p_{\rm T} > 20$ GeV/c is required to point to the Level-1 τ-jet object within a cone of radius R=0.5. The isolation requirement is that there are no pixel tracks with $p_{\rm T} > 2$ GeV/c within this cone. The tracking requirements limit the coverage to $|\eta| \leq 1.9$. When such a candidate is found, the event size is reduced by limiting the FED collections to be written out. Strip tracker and ECAL FEDs associated with the nominal tracks are saved together with all of the FEDs of the HCAL and Pixel tracker in order to allow the possibility of accessing all the timing information needed to perform HCAL calibration. The trigger will have to be highly prescaled in order to keep the rate within the bandwidth limitations. The total rate of such events is expected to be around 20 Hz or 2 MB/s. Figure 34 shows the expected number of isolated charged particles having energy greater than 15 GeV as a function of pseudo-rapidity in the 2×10^{31} cm^{-2}s^{-1} Level-1 trigger table. It is clear that the bulk of these tracks will be in the endcap region. There will be sufficient number of tracks at medium energies (below 50 GeV) for calibration purposes. However the number of high energy tracks (above 100 GeV) will not be adequate (Figure 35). This data will be used to intercalibrate the cells as well as to provide an estimate of the energy scale. The rich sample of charged pions collected by this trigger will be used to tune the simulation to the data.

6 Current and Future Work

Some discrepancies between the simulation and the test beam data still remain. Most notably i) the fraction of events with MIP in the ECAL is larger at lower beam energies and ii) the energy resolution in the simulation is better than what is observed in the data due to the narrower distribution of the total energy simulated in the HCAL (no ECAL in front). We note that the inadequate treatment of quasi-elastic interactions in the Bertini model is partially responsible for the first effect above.

The DREAM experiment [13] studied $\mathrm{PbWO_4}$ crystals and measured sizable contribution in light output from Cerenkov radiation. Attempts are now being made to model this in the GEANT4 simulation and studies are in progress to understand its implication, if any, in the context of CMS.

Finally, some rare events are observed in the test beam data that are not reproduced by the simulation. Specifically, events with large energy deposits are seen during the test beam runs in the HB and HF. Modeling in the simulation is not yet adequate to produce this type of events. Approximately 1.5% of hadron showers have shower particles passing through the ECAL's Avalanche Photo Diode (APD) resulting in $\sim$ 5% increase in the measured energy. Events with abnormally high energy in the EB+HB setup of the 2006 H2 test beam are characterized as a single crystal event with unusual time structure. These are compatible with the hypothesis of nuclear counter effect. Events with unusually large energy are seen in the test beam data with HF for muon and pion beams at certain η values. These are attributed to Cerenkov radiation in the window preceding the photo-cathode. To design and implement an efficient software filter for these abnormal events requires a proper simulation.

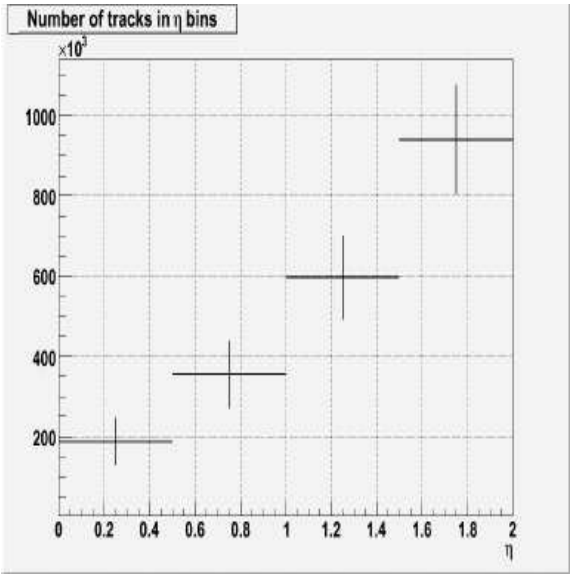


Figure 34: Expected number of isolated charged particles above 15 GeV as a function of η for one week running with the 2×10^{31} cm^{-2}s^{-1} Level-1 trigger table.

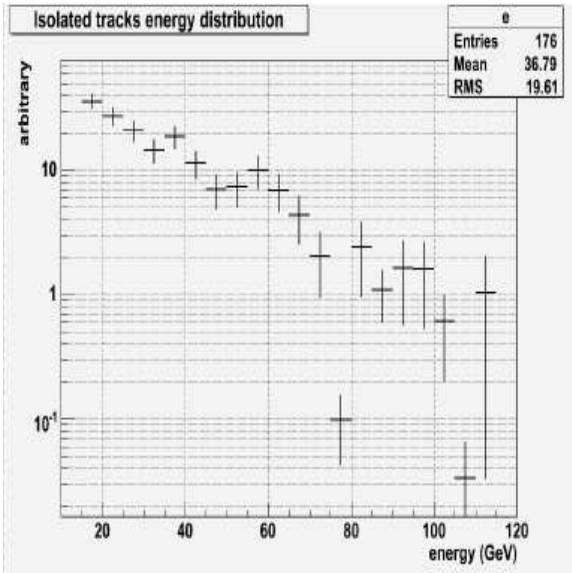


Figure 35: Expected energy spectrum of isolated charged particles to be obtained by the special calibration trigger path.

The issues above will be addressed in close collaboration with the GEANT4 team. The current work on the GFLASH to GEANT4 and FastSim to test beam data tuning and subsequent cross-validation is expected to be finalized before the first LHC collision data.

7 Summary and Outlook

The Calorimetry Simulation Task Force was charged with providing

- a more accurate calorimeter simulation,
- in-depth understanding of the handles within the GEANT4 toolkit used to simulate the passage of particles through matter,
- the development of a faster showering simulation using GFLASH-like parameterizations with the flexibility to rapidly tune FastSim to data.

The task force has delivered several improvements resulting in a significantly better agreement between the detector simulation and test beam data. The procedures, tools, and expertise developed by the CaloTF will assist the physics analysis of the experiment both at start-up and as we gain an increased understanding of the detector by providing accurate and fast calorimetry simulations that can be iteratively tuned to the data.

GEANT4 FullSim Report

We developed an improved version of GEANT4 for CMS in collaboration with the GEANT4 team. This new version has changes in the Bertini cascade model and in the LHEP model. In addition, we reccommend changing the default physics list from QGSP_EMV to QGSP_BERT_EMV and enabling Birks' law (describing the energy response saturation of scintillation light emission) with tuned parameters in the $\mathrm{PbWO_4}$ crystals and the plastic scintillator. These changes lead to a significantly better agreement between the test beam data and the simulation results. The mean energy response as a function of the beam energy for showers in both the ECAL and HCAL sections as well as for showers starting in the HCAL (when the particle behaves like a MIP in the electromagnetic calorimeter) is now well described. The lateral shower shape in the electromagnetic calorimeter and the longitudinal shower profile in the hadron calorimeter are in good agreement between the simulation and data. All improvements have

been committed to the CMS software code repository and are available in version CMSSW 2.1.0 of the detector simulation.

GFLASH-LIKE FullSim Report
A parameterization describing the electromagnetic and hadron showers is added to FullSim following a GFLASH-like approach. This results in a faster simulation and allows for greater flexibility when tuning the simulation to the data. The detailed tuning of the parameters is under way using test beam data. The optimization of the parameters describing EM showers is nearly final while those for hadron showers requires additional work. The functional form of the hadron shower parametrization has been extended in order to accommodate the inhomogeneous structure of the CMS calorimeter.

FASTSIM Report
The electromagnetic shower description used by FastSim has been tuned to the H4 test beam data. Tuning of the parameters in FastSim to also reproduce the FullSim results is underway. The hadronic shower description in FastSim started later in the course of this task force and work is ongoing. It is expected that the transverse hadron shower profile can be tuned directly to the test beam data using a hybrid approach, while the currently available test beam data are not adequate to address the tuning of longitudinal hadron show profile in FastSim. An approach that uses the results of the GFLASH longitudinal shower profile will be attempted.

LHC FIRST COLLISION DATA
Special calibration triggers have been developed to collect large samples of events with isolated tracks. The data collected with these triggers will be used for calibrating the detectors as well as for tuning the simulation of the electromagnetic and hadron showers for both FastSim and FullSim. The necessary software tools are available and being commissioned so that they can be used to quickly tune the simulation to the first collision data.

To summarize, the input to CMSSW 2.1.0 from the task force consists of

- A new patch for the GEANT4 package (4.9.1.p02);
- Choice of a new physics list (QGSP_BERT_EMV);
- Enabling Birks' law (using the L3 parameterization for BGO) with tuned parameters;
- The addition of a GFLASH-like parametrization for EM and hadron showers in FullSim;
- Tuned parameters for the EM shower parametrization for FastSim.

The items still under investigation are:

- Improvement of the Bertini cascade model for the description of quasi-elastics;
- Usage of bremsstrahlung of stable charged hadrons;
- Understanding Cerenkov effects in the crystal;
- Tuning parameters for the hadron showers in the GFLASH-like approach for full and fast simulation;
- Modeling abnormal events in the barrel as well as the forward calorimeter.

8 ACKNOWLEDGMENTS

We wish to thank the HCAL and ECAL DPG and JetMET POG CMS groups for feedback during the course of this work, Harvey Newman and Sezen Sekmen for carefully reading the manuscript of the report and their valuable comments.

9 Appendix:TB description

9.1 The H4 Test Beam

In summer 2006, nine ECAL super-modules were exposed to electrons at the H4 test beam facility at CERN [11]. The H4 beam line provided an electron beam with a narrow momentum spread, corresponding to an RMS of 0.09% between 15 GeV/c and 250 GeV/c. The beam contamination was negligible and the uncertainty on the absolute energy scale was $\sigma p/p = 25\%/p + 0.5\%$.

The super-modules were mounted on a rotating frame to allow the beam to be directed on each crystal of the super-module reproducing the same quasi-projective geometry of CMS. For each crystal the table position was chosen to maximize the energy fraction deposited in the crystal itself. Four planes of hodoscopes were used to reconstruct the trajectories of the electrons impinging on the crystals, with a resolution of about 150 μm in the two coordinates transverse to the beam. The hodoscopes were read out with each trigger. Events were triggered using the signals from plastic scintillation counters placed along the beam line which covered a transverse area of 20×20 mm^2. The final versions of the ECAL readout electronics, high and low voltage supply systems, cooling system, temperature and laser monitoring, data acquisition and data quality monitoring were all tested.

Inter-calibration data were collected on all the exposed super-modules at fixed beam momenta of 90 GeV/c and/or 120 GeV/c. Large statistics samples ($\sim$30k events) were collected on few super-modules at different pseudorapidities and energies for resolution and linearity studies. Finally crystals were exposed to a muon beam.

9.2 The H2 Test Beam

Two production wedges of barrel hadron calorimeter (HB) modules, one module of the hadron endcap (HE) prototype, and eighteen trays of outer hadron calorimeter (HO) were exposed to hadron beams in the H2 beam line of the Super Proton Synchrotron (SPS). Two sets of runs were used to tune hardonic showers. The runs during 2004 [14] used a prototype electromagnetic calorimeter by putting together 49 lead tungstate crystals in a matrix of 7×7. During 2006 [15], this matrix was replaced by a super-module of the barrel electromagnetic calorimeter (EB). The two dimensional movement of the platform in the ϕ and η directions allowed the beam to be directed onto any tower of the calorimeter mimicking a particle trajectory from the interaction point of the CMS experiment.

Four scintillation counters, located 3 meters upstream of the calorimeter, were used to trigger the system. The beam line was equipped with a set of Cerenkov detectors, time of flight counters, muon veto counters, beam halo vetoing system to select on identified single particle shower in the calorimeter.

The test beam was provided by a secondary beam obtained using 400 GeV/c protons from the SPS on a production target (T2) 590.9 m upstream of the calorimeter. Negative and positive beams with momenta between 15 GeV/c and 350 GeV/c were obtained from this beam line. Lower energy beams (VLE) were obtained using an additional target (T22) 97 m upstream to the calorimeter system with a 80 GeV/c secondary beam and a beam dump system. The tertiary beam used to provide beam momenta up to 9 GeV/c. Data was collected for negative and positive beams from 2 GeV/c to 350 GeV/c. The particle identification system worked well for low energy beams ($p < 10$ GeV/c) where showers where measured separately for $\pi^{\pm}$, $K^{\pm}$, protons and anti-protons. High energy measurements exist mainly for π^- and protons. Figure 36 summarizes the energy measurements from 2006 test beam setup.

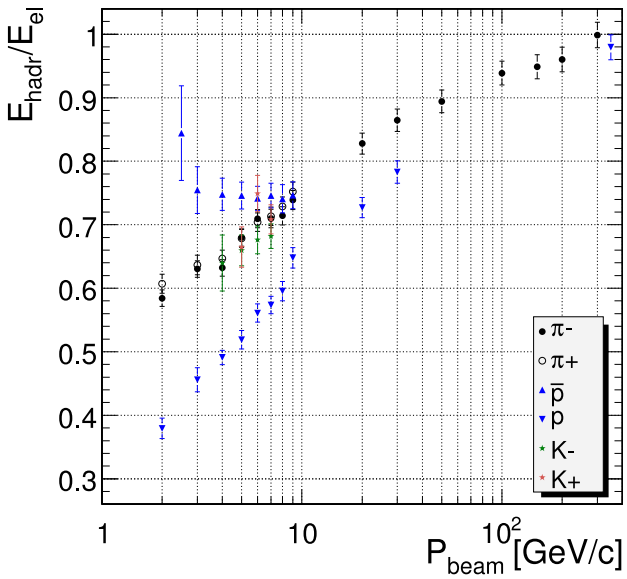


Figure 36: The mean energy response for different particle types as a function of beam momentum measured with the CMS calorimeter prototype in the H2 beam line during 2006.

References

[1] GEANT4: S. Agostinelli *et al.* , Nuclear Instruments and Methods **A506** (2003) 250;
J. Allison *et al.* , IEEE Transactions on Nuclear Science **53** (2006) 78.

[2] J. Birks, Theory and Practice of Scintillation Counting, Pergamon Press, 1964; J. B. Birks, Proc. Phys. Soc. **A64** (1951) 874.

[3] G. Grindhammer, M. Rudowicz, and S. Peters, *"The Fast Simulation of Electromagnetic and Hadronic Showers"*, Nucl. Instrum. and Methods, A290, 469 (1990).

[4] CMSSW: *http://twiki.cern.ch/twiki/bin/view/CMS/WorkBook*

[5] D. H. Wright *et al.*, AIP Conf. Proc. **867** (2006) 479; T. Koi *et al.*, AIP Conf. Proc. **896** (2007) 21.

[6] G. Folger *et al.*, Eur. Phys. J. A **21**407 (2004).

[7] D. H. Wright *et al.*, AIP Conf. Proc. **896** (2007) 11.

[8] A. Heikkinen, N. Stepanov and J. P. Wellisch, *In the Proceedings of 2003 Conference for Computing in High-Energy and Nuclear Physics (CHEP 03), La Jolla, California, 24-28 Mar 2003, pp MOMT008* [arXiv:nucl-th/0306008].

[9] G. Folger and J. P. Wellisch, *In the Proceedings of 2003 Conference for Computing in High-Energy and Nuclear Physics (CHEP 03), La Jolla, California, 24-28 Mar 2003, pp MOMT007* [arXiv:nucl-th/0306007].

[10] P. V. Degtyarenko, M. V. Kosov and H. P. Wellisch, Eur. Phys. J. A **8** (2000) 217.

[11] A. Zabi *et al.* , CMS DN-2007/005.

[12] S. Grindhammer and S. Peters, "The parameterized simulation of electromagnetic showers in homogeneous and sampling calorimeters,"*Int. Conf. of Monte-Carlo Simulation in High Energy and Nuclear Physcis, Tallahassee, Florida, USA(1993),* arXiv:hep-ex/0001020

[13] N. Akchurin *et al.* , Nuclear Instruments and Methods **A582** (2007) 474;
N. Akchurin *et al.* , Nuclear Instruments and Methods **A584** (2008) 273.

[14] G. Baiatian *et al.* , CMS Note 2006/143.

[15] S. Abdullin *et al.* , EPJ **C** submitted 2008.