

ABSTRACT

Title of dissertation: Search for Pair Production of
Third-Generation Scalar Leptoquarks
and R-Parity Violating Stops
in Proton-Proton Collisions at $\sqrt{s} = 8 \text{ TeV}$

Kevin Pedro, Doctor of Philosophy, 2014

Dissertation directed by: Professor Sarah C. Eno
Department of Physics

insert abstract here

Search for Pair Production of Third-Generation Scalar Leptoquarks
and R-Parity Violating Stops in Proton-Proton Collisions at
 $\sqrt{s} = 8 \text{ TeV}$

by

Kevin Pedro

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2014

Advisory Committee:
Professor Sarah C. Eno, Chair/Advisor

© Copyright by
Kevin Pedro
2014

Dedication

To my parents, Philip and Lisa

Acknowledgments

insert acknowledgments here

Table of Contents

List of Tables	vii
List of Figures	viii
List of Abbreviations	x
1 Introduction	1
2 Theoretical Motivations	2
2.0 The Standard Model	2
2.1 Leptoquarks	2
2.2 R-Parity Violating Supersymmetry	2
3 Compact Muon Solenoid Experiment	3
3.0 The Large Hadron Collider	3
3.1 Tracker	3
3.2 Electromagnetic Calorimeter	3
3.3 Hadronic Calorimeter	3
3.4 Solenoid	3
3.5 Muon System	3
3.6 Trigger	3
3.7 Luminosity Measurement	3
4 Event Reconstruction	4
4.1 Event Generation	5
4.2 Detector Simulation	5
4.2.1 Full Simulation	5
4.2.2 Fast Simulation	5
4.2.3 Digitization and Reconstruction	5
4.3 Particle Flow	5
4.4 Tracks and Vertices	5
4.5 Electrons	5
4.6 Muons	5

4.7	Taus	5
4.8	Jets	5
4.9	b-tagging	5
5	Data Analysis	6
5.1	Data Samples	7
5.1.1	Observed Data	7
5.1.2	Monte Carlo	7
5.2	Selection and Optimization	7
5.2.1	Object Identification	7
5.2.1.1	Muons	7
5.2.1.2	Electrons	7
5.2.1.3	Taus	7
5.2.1.4	Jets	7
5.2.2	Event Selection	7
5.2.2.1	Preselection	7
5.2.2.2	Main Selection	7
5.2.2.3	Final Selections	7
5.3	Background Estimations	7
5.3.1	Irreducible Background (ttbar)	7
5.3.2	Reducible Background (fake tau)	7
5.3.3	Reducible Background (QCD)	7
5.3.4	Other Backgrounds	7
5.4	Systematic Uncertainties	7
5.5	Results	7
6	Calorimeter Upgrades	8
6.1	Phase 1 Simulations	8
6.1.1	HE Radiation Damage Model	8
6.1.2	Jet Studies with Radiation Damage	16
6.2	Phase 2 Simulations	16
6.2.1	Validation of Upgrade Standalone Simulation	16
6.2.1.1	Sampling Factors	18
6.2.1.2	Birks' Law for the Electromagnetic Calorimeter	21
6.2.1.3	Dead Material	22
6.2.1.4	CMSSW Settings	22
6.2.1.5	Pion Results	23
6.2.2	Tests of Physics Effects on Pion Response and Resolution	23
6.2.3	HE Rebuild/Extension + Shashlik ECAL Physics Studies	27
6.3	Hadronic Fast Simulation	27
6.3.1	Retuning of Hadronic Response	28
6.3.2	MIP Fraction in Hadronic Showers	34
6.4	Dose Rate Effects	38
6.4.1	Dose Rate Effect Models	38
6.4.2	Scintillator Radiation Damage Studies	38

7	Conclusions	39
A	Full CLs Shape-Based Limits	40
B	Event Displays	43
C	Table of Monte Carlo Datasets	44
D	CMS Collaboration	45
	Bibliography	46

List of Tables

6.1	List of test settings	26
-----	---------------------------------	----

List of Figures

6.1	Exponential fits to 2012 HCAL laser calibration data, for layers 1 and 7 in select towers. Data for the positive (HEP) and negative (HEM) sides of the HE are shown separately. The average of the HEP and HEM values for each scaling constant D is used for the radiation damage model. [1]	9
6.2	Interpolation and extrapolation of the scaling constants D for each layer of each tower in the HE. [1]	10
6.3	Relative signal in each layer of each tower of the HE at (a) 500 fb^{-1} and (b) 3000 fb^{-1}	11
6.4	Depth segmentation schemes in the HCAL for (a) Phase 0 with HPDs as the photodetectors and (b) a proposal for the Phase 1 upgrade with SiPMs as the photodetectors. Locations of the front-end electronics (FEE) for the HB and the HE are also shown. [2]	12
6.5	Recalibration factors for the Phase 0 depth segmentation scheme. The default recalibration cutoff of 20 is shown as a dotted line on each plot.	14
6.6	Recalibration factors for the proposed Phase 1 depth segmentation scheme. The default recalibration cutoff of 100 is shown as a dotted line on each plot.	15
6.7	Plot of the SiPM dark current for different radiation doses. [2]	17
6.8	Geant4 visualization of the standalone simulation with particle showers from a 500 GeV pion (top) and 500 GeV electron (bottom). From left to right: preshower (dark blue), EE (gold), dead material (purple), HE (blue and white).	19
6.9	Energy distributions from 10 000 pions at 50 GeV, fit with Gaussians, for the standalone simulation (left) and the CMSSW full simulation (right).	24
6.10	The ECAL energy vs. the HCAL energy from 10 000 pions at 50 GeV, for the standalone simulation (left) and the CMSSW full simulation (right).	24
6.11	Comparison between the standalone simulation and the CMSSW full simulation for pion energy response (left) and resolution (right). . . .	25

6.12	Comparison of energy response for pions at 8 GeV (left) and 50 GeV (right), using test settings given in Table 6.1.	26
6.13	Comparison of energy resolution for pions at 8 GeV (left) and 50 GeV (right), using test settings given in Table 6.1.	27
6.14	Plots of MIP percentage vs. energy for $i\eta = 1$ (in the barrel) and $i\eta = 20$ (in the endcap).	36
6.15	Plots of MIP percentage vs. energy and η for the entire calorimeter system.	37
A.1	Comparison of the observed value (red line) to the probability densities for H_0 (background only, blue line) and H_1 (signal + background, brown line) as a function of the log likelihood ratio. Green area: CL_{s+b} , yellow area: $1 - CL_b$. From [11].	42

List of Abbreviations

ALICE	A Large Ion Collider Experiment
APD	Avalanche Photodiode
APV	Atomic Parity Violation
ATLAS	A Toroidal LHC ApparatuS
BPTX	Beam Pick-up Timing for the eXperiments
BRW	Buchmüller-Rückl-Wyler
BSM	Beyond Standard Model
CDF	Cumulative Distribution Function
CERN	European Organization for Nuclear Research
CL	Confidence Level
CMS	Compact Muon Solenoid
CMSSW	CMS Software
CP	Charge-Parity
CPU	Central Processing Unit
CSC	Cathode Strip Chamber
CTF	Combinatorial Track Finder
DAQ	Data Acquisition
DT	Drift Tube
EB	ECAL Barrel
ECAL	Electromagnetic Calorimeter
EE	ECAL Endcap
EM	Electromagnetic
FCNC	Flavor-Changing Neutral Current
FSR	Final-State Radiation
GSF	Gaussian Sum Filter
GUT	Grand Unified Theory
HB	HCAL Barrel
HCAL	Hadron Calorimeter
HE	HCAL Endcap
HEEP	High Energy Electron Pairs
HERA	Hadron-Electron Ring Accelerator
HF	HCAL Forward
HO	HCAL Outer
HPD	Hybrid Photodiode
HLT	High-Level Trigger
IP	Interaction Point
ISR	Initial-State Radiation
L1	Level 1
L1A	Level-1 Accept
LEP	Large Electron-Positron Collider
LHC	Large Hadron Collider
LHCb	Large Hadron Collider beauty
LQ	Leptoquark
LO	Leading Order

mBRW	minimal Buchmüller-Rückl-Wyler
MB	Muon Barrel
MC	Monte Carlo
ME	Muon Endcap
MET	Missing Transverse Energy
MIP	Minimum Ionizing Particle
NLO	Next-to-Leading Order
NNLO	Next-to-Next-to-Leading Order
PD	Primary Dataset
PF	Particle Flow
PDF	Parton Distribution Function
PDF	Probability Density Function
PMT	Photomultiplier Tube
PS	Proton Synchrotron, Preshower
PSB	Proton Synchrotron Booster
QED	Quantum Electrodynamics
QCD	Quantum Chromodynamics
RBX	Readout BoX
RF	Radio Frequency
RMS	Root Mean Square
RPC	Resistive Plate Chamber
RPC	R-Parity Conserving
RPV	R-Parity Violating
SiPM	Silicon Photomultiplier
SLHA	SUSY Les Houches Accord
SM	Standard Model
SPS	Super Proton Synchrotron
SUSY	Supersymmetry
TCS	Trigger Control System
TEC	Tracker End Cap
TIB	Tracker Inner Barrel
TID	Tracker Inner Disks
TOB	Tracker Outer Barrel
TPG	Trigger Primitive Generator
TTC	Timing, Trigger and Control
VPT	Vacuum Phototriode
WLS	Wavelength-Shifting

Chapter 1: Introduction

Chapter 2: Theoretical Motivations

2.0 The Standard Model

2.1 Leptoquarks

2.2 R-Parity Violating Supersymmetry

Chapter 3: Compact Muon Solenoid Experiment

3.0 The Large Hadron Collider

3.1 Tracker

3.2 Electromagnetic Calorimeter

3.3 Hadronic Calorimeter

3.4 Solenoid

3.5 Muon System

3.6 Trigger

3.7 Luminosity Measurement

Chapter 4: Event Reconstruction

4.1 Event Generation

4.2 Detector Simulation

4.2.1 Full Simulation

4.2.2 Fast Simulation

4.2.3 Digitization and Reconstruction

4.3 Particle Flow

4.4 Tracks and Vertices

4.5 Electrons

4.6 Muons

4.7 Taus

4.8 Jets

4.9 b-tagging

Chapter 5: Data Analysis

5.1 Data Samples

5.1.1 Observed Data

5.1.2 Monte Carlo

5.2 Selection and Optimization

5.2.1 Object Identification

5.2.1.1 Muons

5.2.1.2 Electrons

5.2.1.3 Taus

5.2.1.4 Jets

5.2.2 Event Selection

5.2.2.1 Preselection

5.2.2.2 Main Selection

5.2.2.3 Final Selections

Chapter 6: Calorimeter Upgrades

6.1 Phase 1 Simulations

6.1.1 HE Radiation Damage Model

Radiation damage to the HE plastic scintillator tiles and wavelength-shifting fibers causes a reduction in scintillation light output. This light loss or darkening is modeled by an exponential degradation function, with specific parameters for each HE tile. These parameters were derived from 2012 HCAL laser calibration data, which exists for layers 1 and 7. Figure 6.1 shows exponential fits of relative light yield vs. integrated luminosity in fb^{-1} from the laser data for selected towers. The parameter D is a scaling constant for the exponential degradation. A smaller value of D means that the tile darkens more rapidly. The value of D varies per tile based on the pseudorapidity and layer locations of the tile, which determine the dose received, and also the size of the tile, which determines the mean path length that light must travel to escape the tile.

The scaling constants from layers 1 and 7 are interpolated for layers 2-6 and extrapolated for layers 8-17, as shown in Fig. 6.2. The values from layer 1 are used for layers 0 and -1. This specifies a radiation damage model for the entire HE.

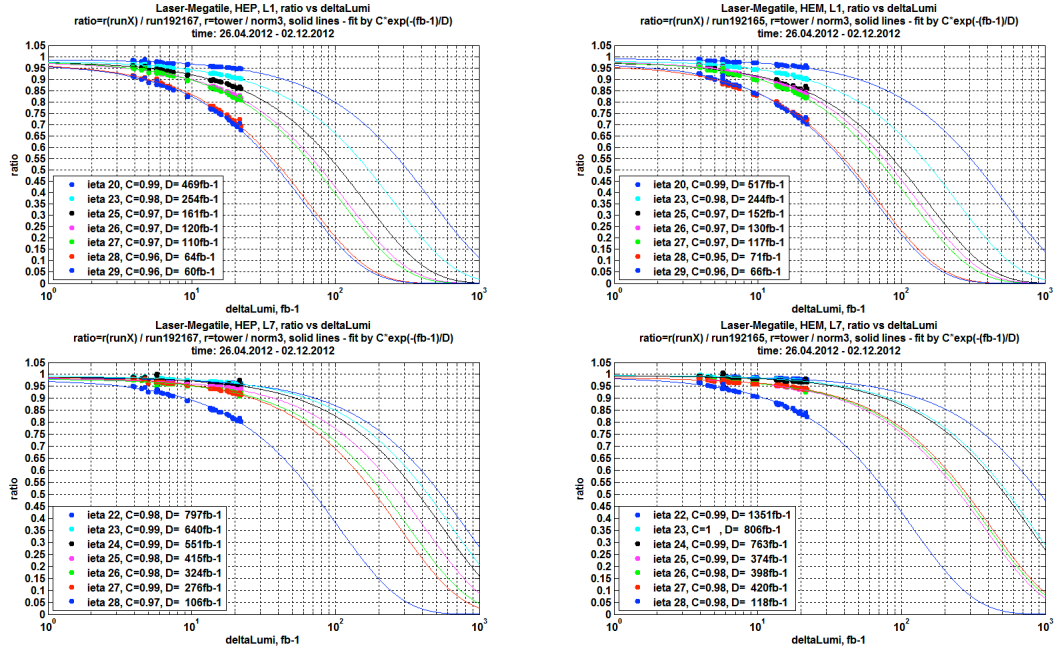


Figure 6.1: Exponential fits to 2012 HCAL laser calibration data, for layers 1 and 7 in select towers. Data for the positive (HEP) and negative (HEM) sides of the HE are shown separately. The average of the HEP and HEM values for each scaling constant D is used for the radiation damage model. [1]

Figure 6.3 shows the relative signal in each layer of each tower of the HE at different integrated luminosity values, based on this radiation damage model. When the LHC center-of-mass energy increases to 14 TeV, a given integrated luminosity value will correspond to a higher amount of dose than it would at 8 TeV when the laser calibration measurements were made. To account for this, the scaling constants are divided by a factor of 1.2, based on FLUKA calculations of the difference in particle flux for 8 TeV and 14 TeV.

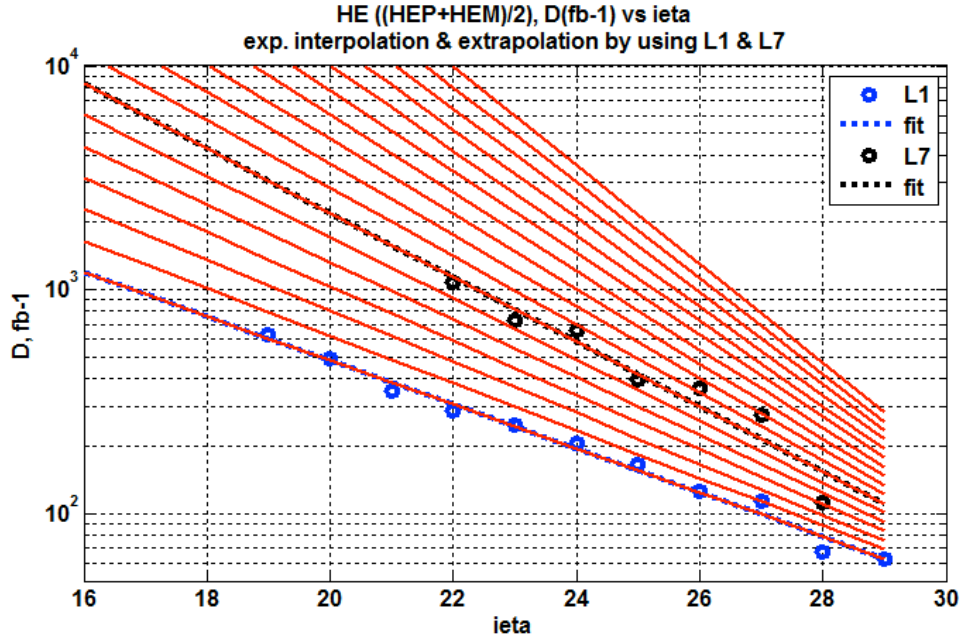


Figure 6.2: Interpolation and extrapolation of the scaling constants D for each layer of each tower in the HE. [1]

The HE photodetectors will be recalibrated to compensate for the darkening of the scintillators and fibers. In a given tower, the light output from several layers is sent to a single photodetector. Each group of layers assigned to a photodetector in this way is called a depth. The specific arrangement of layers into depths, called

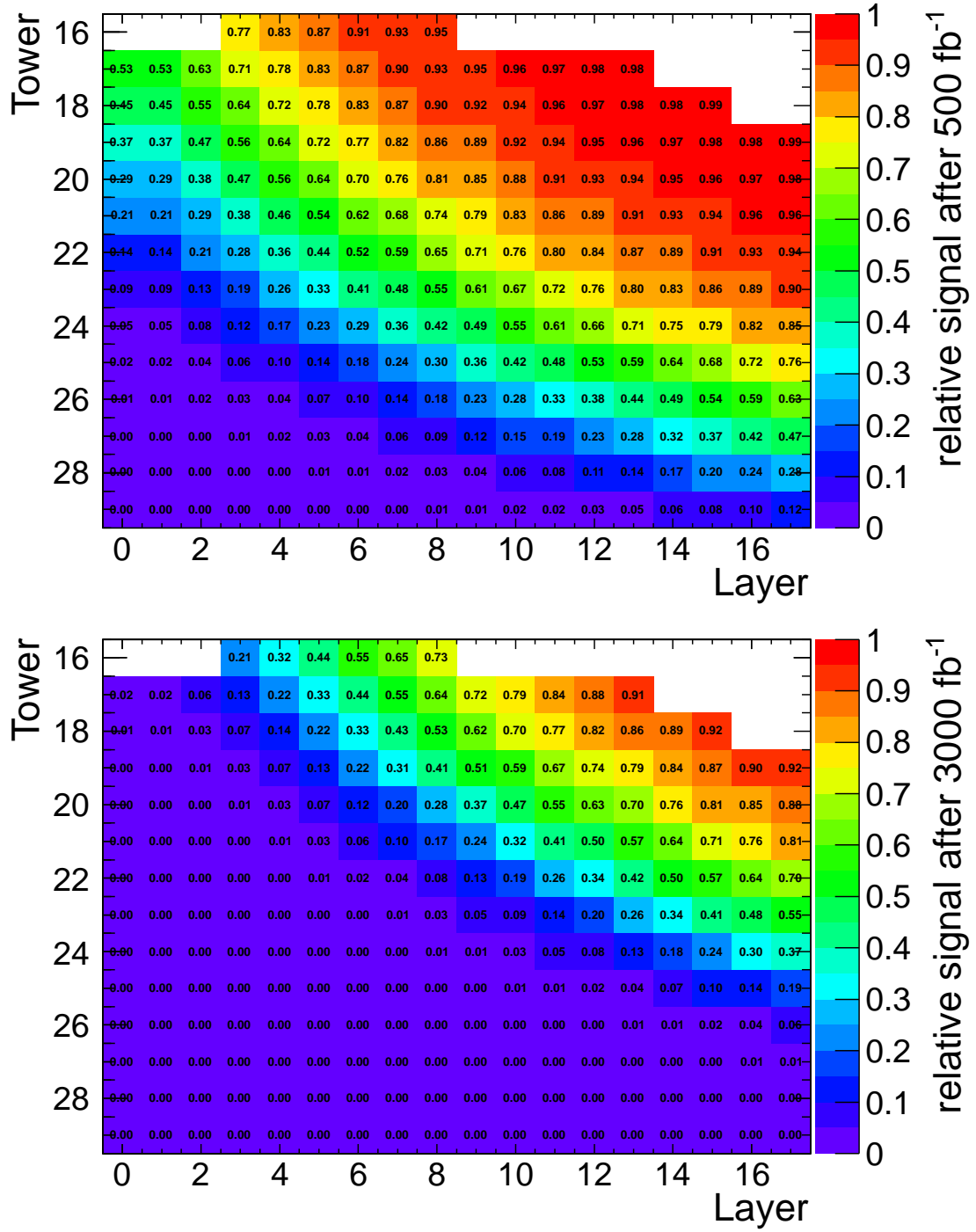


Figure 6.3: Relative signal in each layer of each tower of the HE at (a) 500 fb^{-1} and (b) 3000 fb^{-1} .

the depth segmentation scheme, will be changed during the Phase 1 upgrade of the HCAL. Figure 6.4 shows the current depth segmentation scheme and a potential upgrade scheme.

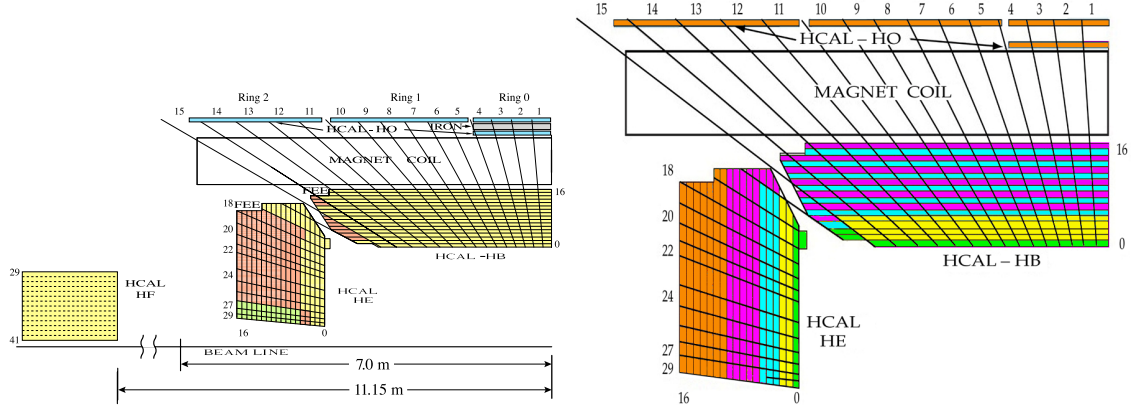


Figure 6.4: Depth segmentation schemes in the HCAL for (a) Phase 0 with HPDs as the photodetectors and (b) a proposal for the Phase 1 upgrade with SiPMs as the photodetectors. Locations of the front-end electronics (FEE) for the HB and the HE are also shown. [2]

The HE radiation damage model has a specific degradation curve for each layer of each tower. This means that recalibration factors can be calculated for any depth segmentation line scheme, as long as the initial relative weighting of layers within a depth is known. The weighting is determined by tabulating mean energy deposits $\langle E \rangle(\ell, i\eta, 0)$ from the full simulation of 100,000 single pion events, where each pion has an energy of 50 GeV and is generated uniformly in $1.6 < \eta < 3.0$ and in ϕ . The simulation is run at 0 fb^{-1} , so no darkening is simulated. With these mean energy weights, the radiation damage model, and a depth segmentation scheme,

recalibration factors can be calculated for any integrated luminosity:

$$\langle E \rangle(\ell, i\eta, L) = e^{-L/D(\ell, i\eta)} \langle E \rangle(\ell, i\eta, 0) \quad (6.1)$$

$$\langle E \rangle(d, i\eta, L) = \sum_{\ell \in d} \langle E \rangle(\ell, i\eta, L) \quad (6.2)$$

$$f(d, i\eta, L) = \frac{\langle E \rangle(d, i\eta, 0)}{\langle E \rangle(d, i\eta, L)} \quad (6.3)$$

Equation (6.1) shows how to use the radiation damage model to find the weights at a given integrated luminosity value, based on the weights at 0 fb^{-1} . Equation (6.2) is a sum of the weights for the layers in a given depth, determined by the depth segmentation scheme. Equation (6.3) gives the calculation for the recalibration factor at a given depth and integrated luminosity. In these equations, $\langle E \rangle$ is the average energy used as a weight (for layers or for depths), ℓ is the layer number, $i\eta$ is the tower number, L is the integrated luminosity in fb^{-1} , d is the depth number, and f is the recalibration factor.

The recalibration factors for the specified Phase 0 and Phase 1 depth segmentation schemes are shown in Figs. 6.5 and 6.6, respectively. In practice, a maximum recalibration cutoff will be assigned so that photodetector signals and noise are not multiplied by absurdly large values. The default cutoff values are 20 for HPDs and 100 for SiPMs, shown on the figures.

Radiation does not have a large effect on the pedestal widths of the HPDs. However, the SiPMs' dark current grows with increasing radiation dose, which leads photodetector noise to increase with integrated luminosity. The simulation of SiPMs

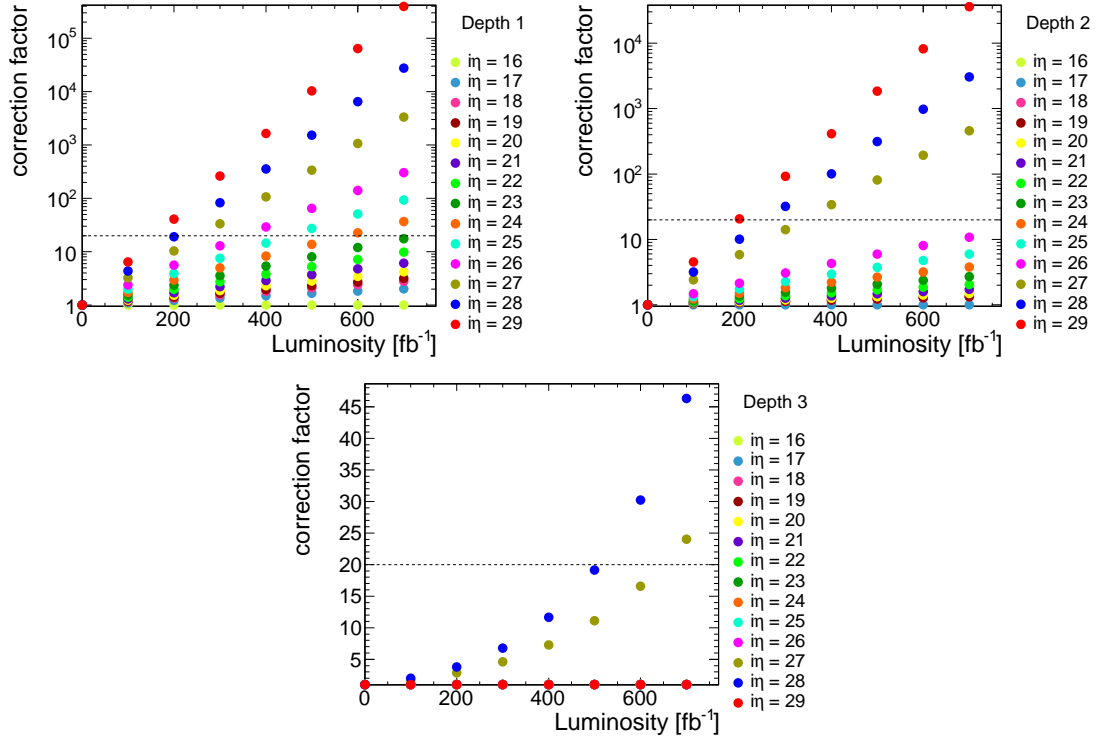


Figure 6.5: Recalibration factors for the Phase 0 depth segmentation scheme. The default recalibration cutoff of 20 is shown as a dotted line on each plot.

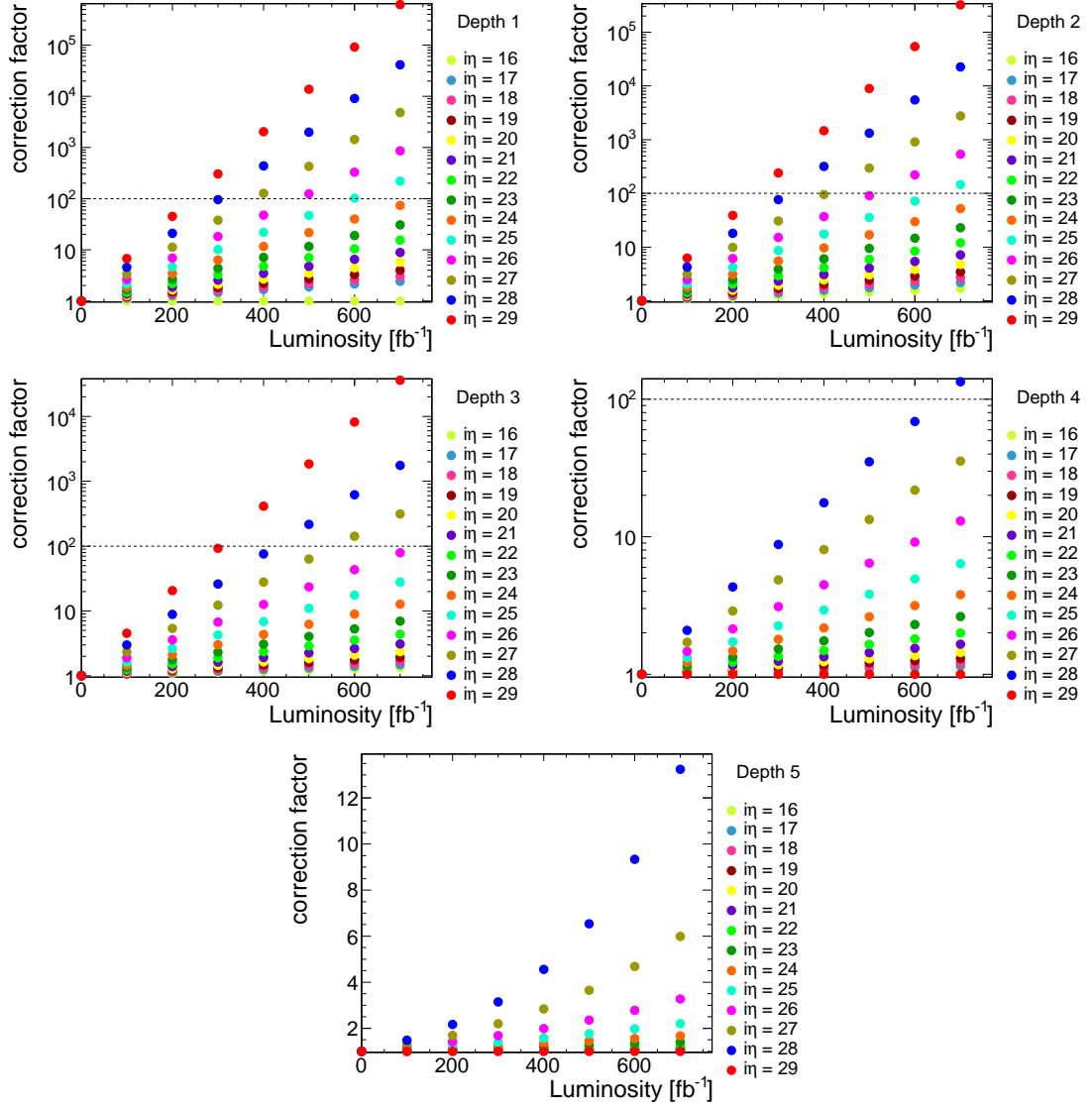


Figure 6.6: Recalibration factors for the proposed Phase 1 depth segmentation scheme. The default recalibration cutoff of 100 is shown as a dotted line on each plot.

accounts for these luminosity-dependent pedestal widths with the following equations, based on measurements shown in Fig. 6.7:

$$L_{\text{eff}} = \max(L - 200 \text{ fb}^{-1}, 0) \quad (6.4)$$

$$\sigma_{\text{SiPM}}^{(\text{HB})} = 5 + 1.7\sqrt{L_{\text{eff}}} \text{ [fC]} \quad (6.5)$$

$$\sigma_{\text{SiPM}}^{(\text{HE})} = 5 + 0.7\sqrt{L_{\text{eff}}} \text{ [fC]} \quad (6.6)$$

Equation (6.4) accounts for the planned installation of the SiPMs after 200 fb^{-1} have already been collected. Equations (6.5) and (6.6) show the scaling of the pedestal widths with the square root of integrated luminosity. The increase in dark current also changes the pedestal means and zero suppression thresholds, and the simulation accounts for these changes.

6.1.2 Jet Studies with Radiation Damage

6.2 Phase 2 Simulations

6.2.1 Validation of Upgrade Standalone Simulation

In order to simulate different options for new calorimeters, the FCAL Task Force created a standalone Geant4 simulation with a simplified geometry based on the CMS endcap detector specifications as given in Ref. [3]. The simulation geometry uses the same coordinate system as the real CMS geometry, with z as the beamline and x, y as the transverse directions. The detector components all have the shape

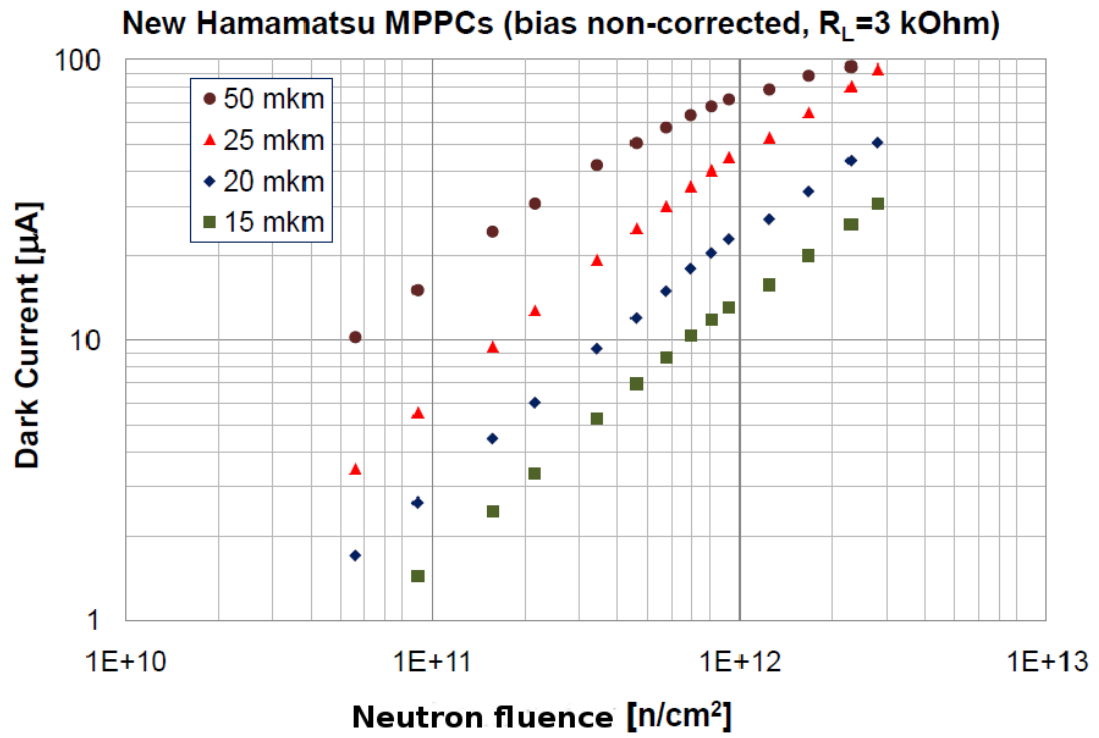


Figure 6.7: Plot of the SiPM dark current for different radiation doses. [2]

of rectangular prisms with transverse dimensions $540\text{ cm} \times 540\text{ cm}$. The preshower is simulated as 3 cm of aluminum, 1.5 cm of lead, and a G10 cryogenic plate. The EE is constructed as a 22 cm deep homogeneous medium of PbWO_4 . Between the EE and the HE, there is a section of dead material, which is modeled as 23.4 cm of a mixture of copper and polymer to simulate cables. The HE is 149.6 cm deep with 17 sampling layers, each consisting of 79 mm brass, 3.7 mm plastic scintillator, air gaps, and aluminum shielding. At the front of the HE there is a 15 mm zero layer, including 9 mm plastic scintillator, air gaps, and aluminum shielding. Energy deposits in the zero layer are weighted with a factor of 0.5 to account for its larger size. Figure 6.8 shows a side view of the simulated geometry with particle showers propagating through the calorimeters.

To ensure that the essential physics and detector behavior are present in the standalone simulation, it was validated against the CMSSW full simulation using pions and the settings given in Sec. 6.2.1.4. The validation compared the pion response and resolution from SimHits, since the standalone simulation does not include a simulation of electronics or event reconstruction. Several subtle effects and details had to be taken into account in order to reproduce the CMSSW full simulation results in the standalone simulation.

6.2.1.1 Sampling Factors

The HCAL is a sampling calorimeter which collects only a fraction of the total energy of incident particles. A conversion factor is needed to estimate how

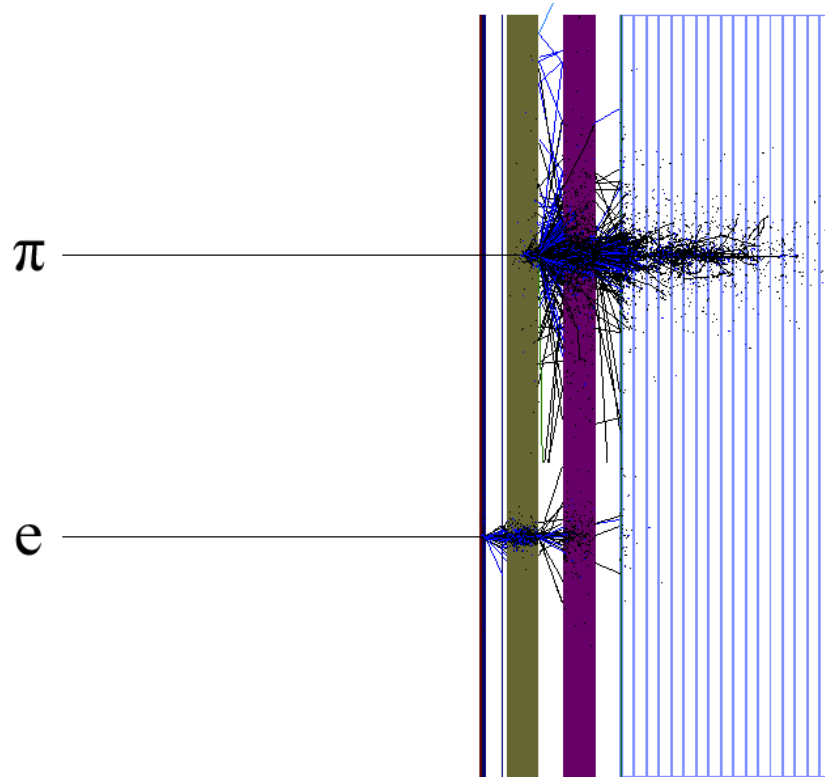


Figure 6.8: Geant4 visualization of the standalone simulation with particle showers from a 500 GeV pion (top) and 500 GeV electron (bottom). From left to right: preshower (dark blue), EE (gold), dead material (purple), HE (blue and white).

the collected energy corresponds to the original particle’s energy. There are numerous methods for calculating this sampling factor, and variations in it can lead to significant differences in energy response. To ensure similarity to the CMSSW full simulation, the sampling factor algorithm for the standalone simulation follows the CMSSW algorithm, described in Ref. [4], as closely as possible. The simplified geometry of the standalone simulation means that there is no structure of $i\eta$ towers and therefore no need to iterate the algorithm to account for the spread of energy over different towers.

Sampling factors are calculated by selecting from a sample of 50 GeV pions only the pions which act as minimum ionizing particles (mips) in the ECAL, so each shower is largely contained within the HCAL and a comparison to the incident energy of 50 GeV is appropriate. Based on the energy distribution of pions in the ECAL, a mip threshold of $E_{\text{ecal}} < 1 \text{ GeV}$ was chosen. The sampling factor is calculated from a sample of pions which pass this threshold:

$$f_{\text{hcal}} = \frac{E_{\text{gen}}}{E_{\text{hcal}}^*}, \quad (6.7)$$

where E_{hcal}^* is the peak, or most probable, energy deposited in the HCAL from that pion sample. Applying Eq. (6.7) to both the standalone simulation and the CMSSW full simulation, with settings which will be specified below, gives values $f_{\text{hcal}}^{\text{standalone}} = 178$ and $f_{\text{hcal}}^{\text{CMSSW}} = 183$, showing good agreement.

6.2.1.2 Birks' Law for the Electromagnetic Calorimeter

Organic scintillators have a nonlinear energy response due to quenching. Birks' Law [5] describes this nonlinearity with several parameterizations comparing the scintillator light output to the energy deposited by a particle. Inorganic scintillators, too, may have a nonlinear energy response, which requires a different parameterization. During the development of the CMSSW full simulation, it was found that the agreement of simulation with data for the ECAL improved when using the parameterization developed for the BGO detector at the L3 experiment [6]. The BGO values must be used directly, as the nonlinearity of PbWO_4 has not been measured. This parameterization has also been included in the standalone simulation:

$$\frac{dL}{dr} = w \frac{dE}{dr}, \quad (6.8)$$

$$w = 1 - s \cdot \log \left(k \frac{dE}{dr} \right), \quad (6.9)$$

where dL/dr is the light output, dE/dr is the energy deposited, and w is the weight factor which introduces the nonlinearity. The parameter s is an empirical slope factor, and k is related to Birks' constant kB . After w is calculated in Eq. (6.9), its bounds are checked: if it is lower than a parameter c , it is set equal to c ; if it is higher than 1, it is set equal to 1. The BGO values of these parameters, which are applied to PbWO_4 , are $k = 0.03333$, $s = 0.253694$, and $c = 0.1$.

6.2.1.3 Dead Material

Initial versions of the standalone geometry included only a small amount, 1.5 mm, of dead material between the ECAL and the HCAL. However, in the real detector there is a much larger amount of dead material, consisting of cables, support structures, and cooling systems. Using the material budget tool described in Ref. [7], the CMSSW full simulation geometry was scanned at $\eta = 2$, roughly the center of the endcap, for various ϕ values. The depths of all materials between the end of the ECAL crystals and the start of the HCAL structure are summed to find a range of dead material varying in ϕ between $0.57 \lambda_0$ and $0.63 \lambda_0$. Geant4 has an internal formula for the nuclear interaction length, which calculated $\lambda_0 = 389.98 \text{ mm}$ for the copper-polymer material used to simulate the dead material. The depth of this material was increased to $234 \text{ mm} = 0.6 \lambda_0$, the average of the scanned material budget range, and the volume was placed to be centered between the ECAL and the HCAL as depicted in Fig. 6.8.

6.2.1.4 CMSSW Settings

There are several settings in the CMSSW full simulation which can be modified to increase the similarity with the standalone simulation. A calorimeter-only geometry was selected in order to remove the tracker, which is not included in the standalone simulation. The magnetic field was turned off; though a magnetic field could instead be turned on in the standalone simulation, turning off the field prevents lower-energy particles from being deflected away from the center of the

endcap, $\eta = 2$, where the CMSSW calorimeter geometry most closely resembles the standalone geometry. Finally, by default CMSSW uses a customized physics list QGSP_FTFP_BERT_EML [8], which was changed to the standard QGSP_BERT_EMV to match the standalone simulation.

6.2.1.5 Pion Results

Using the sampling factors given in Sec. 6.2.1.1, reconstructed energies for pions are calculated as:

$$E_{\text{rec}} = E_{\text{ecal}} + f_{\text{hcal}} \cdot E_{\text{hcal}}. \quad (6.10)$$

Energy response and resolution are calculated from a Gaussian fit based on the mean m and standard deviation s of the energy distributions, over a range $[m - 2s, m + s]$ to account for the high tail. Figure 6.9 shows example energy distributions and fits for 50 GeV pions, and Fig. 6.10 shows two-dimensional plots of the ECAL energy vs. the HCAL energy. Both figures show good agreement between the standalone simulation and the CMSSW full simulation. This agreement extends over a large range of pion energies, 1 GeV to 3 TeV, as demonstrated in Fig. 6.11.

6.2.2 Tests of Physics Effects on Pion Response and Resolution

The pion response shown in Fig. 6.11 is highly nonlinear, especially at lower energies. The simplicity and flexibility of the standalone simulations makes it straightforward to test the various physics and geometrical effects identified in the preceding section. In order to test each effect individually, the simulation starts from the “old

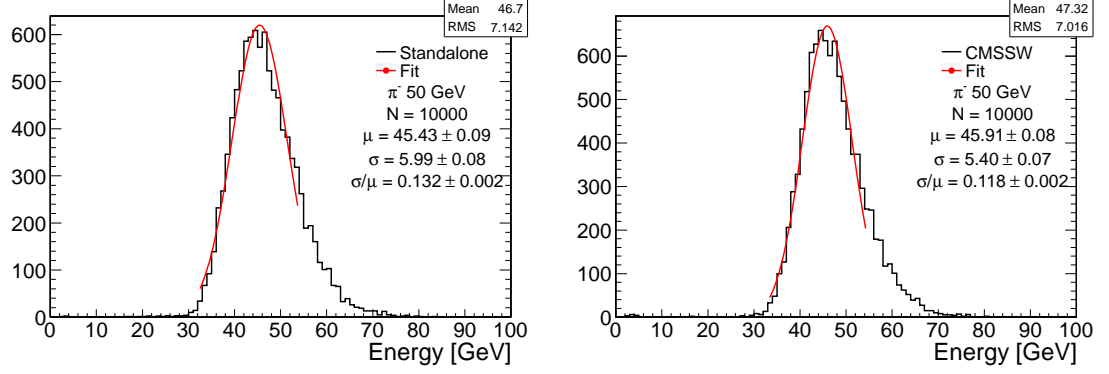


Figure 6.9: Energy distributions from 10 000 pions at 50 GeV, fit with Gaussians, for the standalone simulation (left) and the CMSSW full simulation (right).

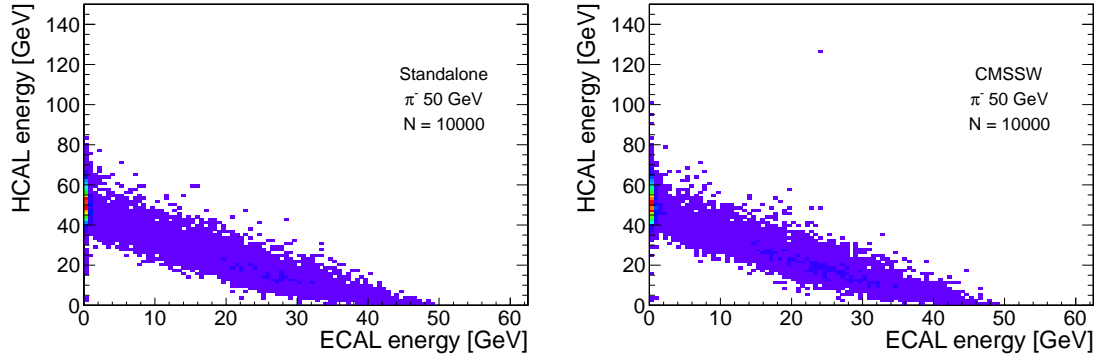


Figure 6.10: The ECAL energy vs. the HCAL energy from 10 000 pions at 50 GeV, for the standalone simulation (left) and the CMSSW full simulation (right).

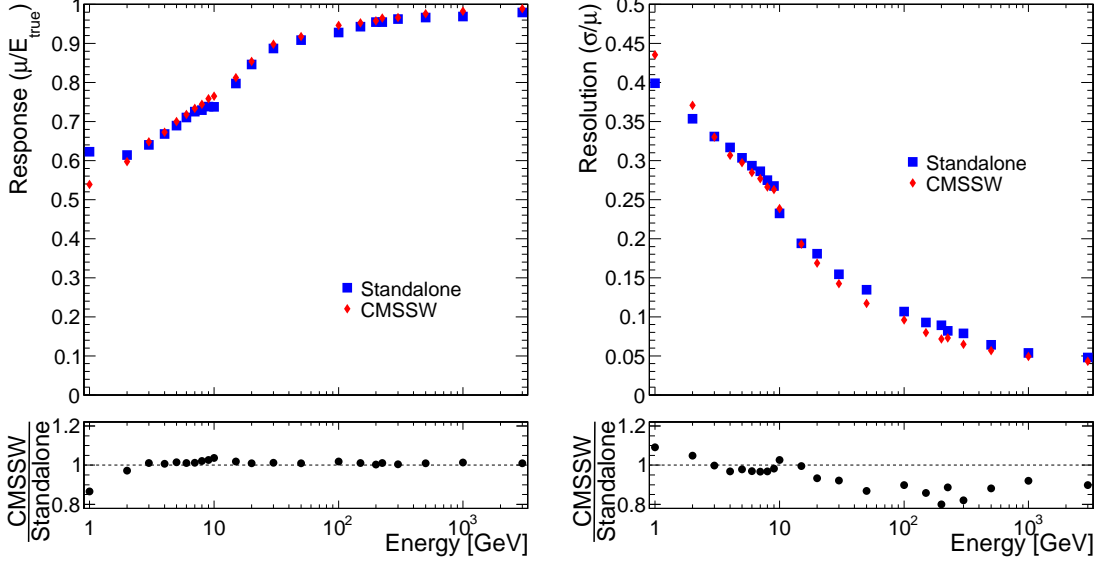


Figure 6.11: Comparison between the standalone simulation and the CMSSW full simulation for pion energy response (left) and resolution (right).

default” settings of the standalone simulation, before the validation against the CMSSW full simulation. The “old default” settings are: no preshower, minimal dead material (1.5 mm), and no Birks’ Law in the ECAL or the zero layer. From here, each test setting is turned on by itself, without any of the others. Setting 5 is an additional test, removing Birks’ Law from all of the HCAL scintillating layers. Changing the zero layer weight from 0.5 to 1.0 was also tested. The “new default” settings are those which were used to validate the standalone simulation; they consist of the “old default” settings with test settings 1, 2, 3, 4 applied in concert. For each test, the sampling factor f_{hcal} was recalculated from 50 GeV pions simulated with the appropriate settings.

Figure 6.12 shows the energy response from a Gaussian fit for 8 GeV and 50 GeV pions. The dead material between the ECAL and the HCAL has the largest

Table 6.1: List of test settings

0. Old default: no preshower, no Birks' law in the ECAL or zero layer, 1.5 mm of dead material
1. 0 + inclusion of preshower
2. 0 + Birks' Law in the ECAL
3. 0 + dead material increased to 234 mm
4. 0 + Birks' Law in the zero layer
5. 0 + no Birks' Law in the HCAL
6. New default: 0 + 1 + 2 + 3 + 4

effect on response, and Birks' Law in the HCAL scintillators also has a large effect. The preshower detector in front of the ECAL has a larger effect on the lower-energy pion response, as expected. Weighting the zero layer equal to the other layers increases the response and also improves the resolution for higher-energy pions when dead material is present, as shown in Fig. 6.13. The presence of the preshower detector is the largest contribution to the resolution for lower-energy pions.

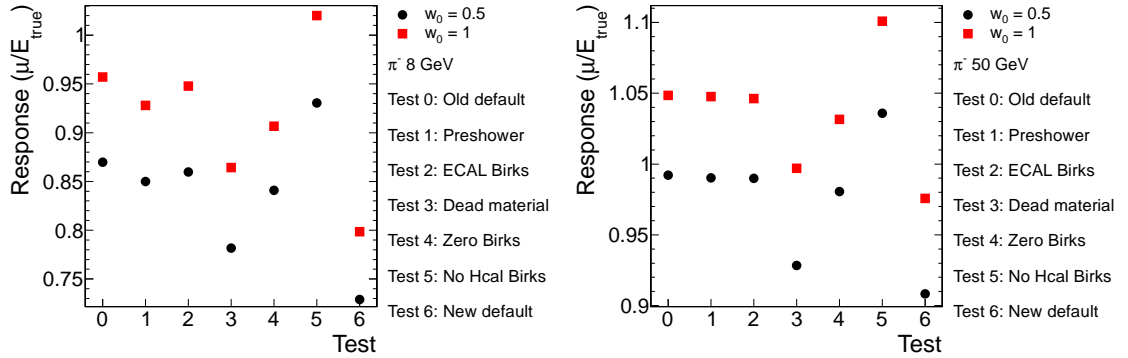


Figure 6.12: Comparison of energy response for pions at 8 GeV (left) and 50 GeV (right), using test settings given in Table 6.1.

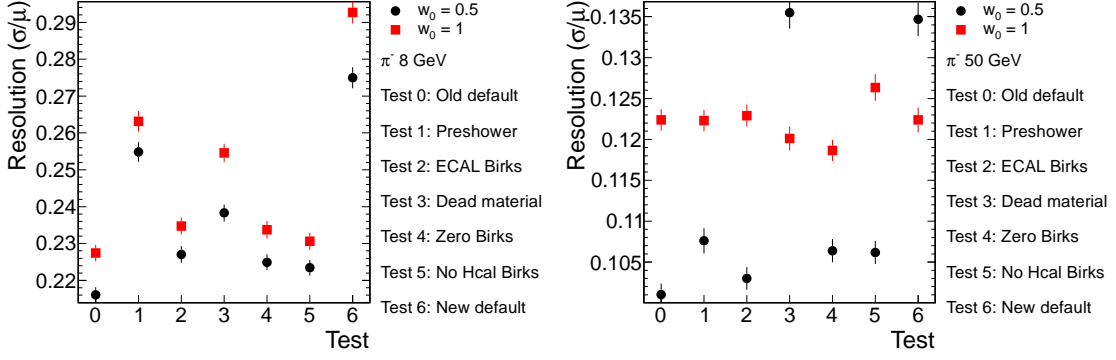


Figure 6.13: Comparison of energy resolution for pions at 8 GeV (left) and 50 GeV (right), using test settings given in Table 6.1.

6.2.3 HE Rebuild/Extension + Shashlik ECAL Physics Studies

6.3 Hadronic Fast Simulation

The CMS hadronic fast simulation is based on the mean shower parameterization from GFLASH [9], containing one term for the hadronic component and another term for the π^0 (electromagnetic) component. Various random fluctuations for the longitudinal and transverse development and the energy deposition of the shower are included in the simulation. However, the GFLASH parameterization for hadronic showers was only developed for single homogeneous calorimeters. The CMS HCAL is a sampling calorimeter with a non-compensating homogeneous crystal ECAL and dead material in front of it. The final energy measured by this combined calorimeter system for a given event is the product of not just shower fluctuations, but also sampling fluctuations, non-linearity, and geometric effects. To include these other effects, the energy deposits from the fast simulation must be multiplied by a random factor generated according to hadronic energy distributions from the full simulation

(“smeared”).

6.3.1 Retuning of Hadronic Response

At first, Gaussian distributions were used for the hadronic energy response smearing. The parameters for the Gaussian distributions were taken from the means and standard deviations of energy distributions for single pion samples from the full simulation. The energy distributions were constructed by collecting reconstructed calorimeter hits in a cone of $\Delta R < 0.5$, divided into bins of $\Delta\eta = 0.1$ based on the incident pion direction. These samples were simulated at different energies and the Gaussian parameters were linearly interpolated for intermediate energies. Additional η - and energy-dependent factors were derived to improve agreement with the full simulation, accounting for any remaining differences between the simulations in an ad-hoc, rather than physical, way.

The fast simulation of the calorimeters originally used a custom reconstruction process. However, to make the simulations easier to maintain and upgrade, the fast simulation was changed to use the full digitization and reconstruction, which are described in Ch. 4.2.3, without significant loss in speed. This change invalidated the ad-hoc correction factors described above, which were derived from the custom reconstruction process for the response smearing based on reconstructed hits. A new, more precise tuning, based on simulated hits, would ideally work well enough to omit any ad-hoc correction factors.

A simple Gaussian distribution does not capture potential large tails in the

pion energy distributions. To account for the behavior of both the low- and high-energy tails, a double-sided Crystal Ball function is used. This function has a Gaussian core with two power-law tails. The Crystal Ball function is continuous, and despite being defined in a piecewise manner, so is its first derivative. In order to generate random numbers distributed according to a double-sided Crystal Ball function, inversion sampling is used. The cumulative distribution function of the double-sided Crystal Ball function is invertible, enabling the random generation to be computed analytically, up to the calculation of the error function erf .

To describe this method, first the probability distribution function (PDF) $f(x)$ and the cumulative distribution function (CDF) $F(x)$ must be defined in general:

$$\int_{-\infty}^{\infty} f(x) dx = 1 \quad (6.11)$$

$$F(x) = \int_{-\infty}^x f(x') dx' \quad (6.12)$$

$$y \equiv F(x) \rightarrow F^{-1}(y) = x \quad (6.13)$$

Equation (6.13) shows the basis of inversion sampling: the inverted CDF $F^{-1}(y)$ produces a random number distributed according to the PDF $f(x)$.

The double-sided Crystal Ball function has six parameters as noted below,

with a few additional definitions and conditions:

$$\vec{p} = (\mu, \sigma, a_L, n_L, a_R, n_R) \quad (6.14)$$

$$d_L = n_L/a_L \quad (6.15)$$

$$d_R = n_R/a_R \quad (6.16)$$

$$n_L, n_R > 1 \quad (6.17)$$

$$a_L, a_R > 0 \quad (6.18)$$

Here, μ is the mean of the Gaussian core, σ is the width of the Gaussian core, a_L is the starting location of the left tail, n_L is the steepness of the left tail, and a_R and n_R act similarly for the right tail. The parameter combinations d_L and d_R appear frequently in the definition of the Crystal Ball function, so it is useful to define them. The conditions on n_L and n_R ensure that the Crystal Ball function has a finite integral and can be used as a PDF, while the conditions on a_L and a_R ensure that the function is continuous. The PDF of the Crystal Ball, using these parameters, is given in Eq. (6.19), with the normalization given in Eq. (6.20).

$$f(x; \vec{p}) = N \cdot \begin{cases} \exp\left(-\frac{a_L^2}{2}\right) \cdot \left[\frac{1}{d_L} \left(d_L - a_L - \frac{x-\mu}{\sigma}\right)\right]^{-n_L} & \text{for } \frac{x-\mu}{\sigma} \leq -a_L \\ \exp\left(-\frac{1}{2} \left(\frac{x-\mu}{\sigma}\right)^2\right) & \text{for } -a_L < \frac{x-\mu}{\sigma} < a_R \\ \exp\left(-\frac{a_R^2}{2}\right) \cdot \left[\frac{1}{d_R} \left(d_R - a_R + \frac{x-\mu}{\sigma}\right)\right]^{-n_R} & \text{for } \frac{x-\mu}{\sigma} \geq a_R \end{cases} \quad (6.19)$$

$$N = \frac{1}{\sigma \left[\frac{d_L}{n_L-1} \cdot \exp\left(-\frac{a_L^2}{2}\right) + \sqrt{\frac{\pi}{2}} \left(\operatorname{erf}\left(\frac{a_L}{\sqrt{2}}\right) + \operatorname{erf}\left(\frac{a_R}{\sqrt{2}}\right) \right) + \frac{d_R}{n_R-1} \cdot \exp\left(-\frac{a_R^2}{2}\right) \right]} \quad (6.20)$$

The PDF can be integrated according to Eq. (6.12) to produce the CDF in Eq. (6.21). The same CDF is shown in shortened form in Eq. (6.22), with some abbreviations defined to make the inverse CDF in Eq. (6.23) more readable.

$$F(x; \vec{p}) = \sigma N \cdot \left\{ \begin{array}{ll} \frac{d_L}{n_L-1} \exp\left(-\frac{a_L^2}{2}\right) \left[\frac{1}{d_L} \left(d_L - a_L - \frac{x-\mu}{\sigma}\right)\right]^{-n_L+1} & \text{for } \frac{x-\mu}{\sigma} \leq -a_L \\ \frac{d_L}{n_L-1} \exp\left(-\frac{a_L^2}{2}\right) + \sqrt{\frac{\pi}{2}} \operatorname{erf}\left(\frac{a_L}{\sqrt{2}}\right) + \sqrt{\frac{\pi}{2}} \operatorname{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right) & \text{for } -a_L < \frac{x-\mu}{\sigma} < a_R \\ \frac{d_L}{n_L-1} \exp\left(-\frac{a_L^2}{2}\right) + \sqrt{\frac{\pi}{2}} \operatorname{erf}\left(\frac{a_L}{\sqrt{2}}\right) \\ + \sqrt{\frac{\pi}{2}} \operatorname{erf}\left(\frac{a_R}{\sqrt{2}}\right) + \frac{d_R}{n_R-1} \exp\left(-\frac{a_R^2}{2}\right) \\ + \frac{d_R}{1-n_R} \exp\left(-\frac{a_R^2}{2}\right) \left[\frac{1}{d_R} \left(d_R - a_R + \frac{x-\mu}{\sigma}\right)\right]^{-n_R+1} & \text{for } \frac{x-\mu}{\sigma} \geq a_R \end{array} \right. \quad (6.21)$$

$$= \sigma N \cdot \left\{ \begin{array}{ll} B_L \left[\frac{1}{d_L} \left(d_L - a_L - \frac{x-\mu}{\sigma}\right)\right]^{-n_L+1} & \text{for } \frac{x-\mu}{\sigma} \leq -a_L \\ A_L + C_L + \sqrt{\frac{\pi}{2}} \left(1 - \operatorname{erfc}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right) & \text{for } -a_L < \frac{x-\mu}{\sigma} < a_R \\ A_L + C_L + C_R + A_R \\ + B_R \left[\frac{1}{d_R} \left(d_R - a_R + \frac{x-\mu}{\sigma}\right)\right]^{-n_R+1} & \text{for } \frac{x-\mu}{\sigma} \geq a_R \end{array} \right. \quad (6.22)$$

$$x = \begin{cases} \mu + \sigma \left(-d_L \left[\frac{y}{\sigma N} / B_L \right]^{\frac{1}{-n_L+1}} - a_L + d_L \right) & \text{for } y < \sigma N A_L \\ \mu + \sigma \sqrt{2} \operatorname{erfc}^{-1} \left[1 - \sqrt{\frac{2}{\pi}} \left(\frac{y}{\sigma N} - A_L - C_L \right) \right] & \text{for } \sigma N A_L \leq y \leq \sigma N (A_L + C_L + C_R) \\ \mu + \sigma \left(d_R \left[\frac{\frac{y}{\sigma N} - A_L - C_L - C_R - A_R}{B_R} \right]^{\frac{1}{-n_R+1}} + a_R - d_R \right) & \text{for } y > \sigma N (A_L + C_L + C_R) \end{cases} \quad (6.23)$$

This inverse CDF is used for the inversion sampling procedure. A uniform random number y is generated and used as the input to Eq. (6.23), producing x which is distributed according to the double-sided Crystal Ball function.

Charged pion events are used to determine the double-sided Crystal Ball parameter values for the hadronic response smearing. For each of several energy values (1, 2, 3, 5, 9, 11, 15, 20, 30, 50, 100, 150, 225, 300, 1000, 3000 GeV), a sample of 500,000 pion events was generated, uniformly distributed in $|\eta| < 5.0$, and simulated using the full simulation. Vertex smearing and the magnetic field were disabled, so the direction of the incident pion would directly predict the location of energy deposits in the calorimeter. Energy from ECAL and HCAL simulated hits was collected in a cone of $\Delta R < 0.5$. The HCAL energy was calibrated by multiplying by the η -dependent sampling factors determined in the full simulation.

For each energy, the sample was divided into bins of $\Delta\eta = 0.1$. Each resulting energy distribution was fit to a double-sided Crystal Ball function. For these fits,

the default energy range was $[0.1E, 2E]$, and the parameter limits were: $\mu, \sigma > 0$; $0 \leq a_L, a_R \leq 10$; $1.01 \leq n_L, n_R \leq 200$. Technically, the Crystal Ball function approaches a Gaussian function when $a \rightarrow \infty$. However, because the a parameters are in units of σ , $a = 10$ is effectively infinite in practice. The chance of a 10σ event occurring randomly, if the events are normally distributed, is approximately 1 in 66 sextillion. The energy range and parameter values were slightly altered in some cases when the χ^2 value of the fit indicated that the fitting algorithm was trapped in a bad local minimum.

For the response smearing, the parameter values resulting from the fits are linearly interpolated for intermediate energies that were not explicitly simulated. The parameter values, including this linear interpolation, are shown in Fig...

6.3.2 MIP Fraction in Hadronic Showers

In the CMS hadronic shower fast simulation, the shower starting depth s is simulated using an exponential distribution. Integrate to find the cumulative distribution for inversion sampling, where $x \in [0, 1]$ is a uniformly distributed random number:

$$f(s) = e^{-s} \tag{6.24}$$

$$F(s) = \int_0^s f(s') ds' = 1 - e^{-s} \tag{6.25}$$

$$x \equiv F(s) \rightarrow F^{-1}(x) = -\ln(1 - x) = \ln\left(\frac{1}{1-x}\right) = s \tag{6.26}$$

In the last step, the fact that x is a uniformly distributed random number in $[0, 1]$ is used to take $(1 - x) \rightarrow x$.

The condition which decides if the shower will start in ECAL is based on a comparison between the depth of ECAL d_{ecal} and the starting depth s . If the shower does not start in ECAL, the incident hadron is considered to be a MIP (minimum ionizing particle) in ECAL.

$$\frac{d_{\text{ecal}} - s}{d_{\text{ecal}}} > 0.1 \quad (6.27)$$

$$\rightarrow 0.9d_{\text{ecal}} > s \text{ (for pion showers starting in ECAL)} \quad (6.28)$$

$$\rightarrow 0.9d_{\text{ecal}} \leq s \text{ (for pions which are MIPs in ECAL)} \quad (6.29)$$

$$\rightarrow d \equiv 0.9d_{\text{ecal}} \text{ (the minimum starting distance for MIPs)} \quad (6.30)$$

Since $f(s)$ is a probability distribution, it has area 1 in $[0, \infty]$. The area for $s = d.. \infty$, i.e. when $s \geq d$, should be equal to the probability p that the particle is a MIP. In order to solve this problem, the distribution must be transformed to introduce a free parameter:

$$f(s, \lambda) = \lambda e^{-\lambda s} \quad (6.31)$$

$$s = \frac{1}{\lambda} \ln \left(\frac{1}{x} \right) \quad (6.32)$$

Now the integral can be solved to require the correct MIP fraction:

$$p = \int_d^\infty ds \lambda e^{-\lambda s} = e^{-\lambda d} \quad (6.33)$$

$$\rightarrow \lambda = \frac{1}{d} \ln \left(\frac{1}{p} \right) \quad (6.34)$$

Since d is determined by detector geometry, for any $p \in (0, 1)$, λ can be found to ensure the correct MIP fraction. The final result is just a scaling by $1/\lambda$ of the original equation for randomly generating s from the uniform random number x . In practice, p can be determined from full simulation as a function of incident particle energy and η . The easiest way to do this would be to store values for each η and the same energy points that are used in HCALResponse, and then interpolate for intermediate energies. (See Figures 6.14 and 6.15 for examples.) For energies outside that range, the first or last values should be used rather than extrapolating, since extrapolating could produce $p \leq 0$ or $p \geq 1$, which would create unphysical values of λ , i.e. $\lambda \notin (0, \infty)$.

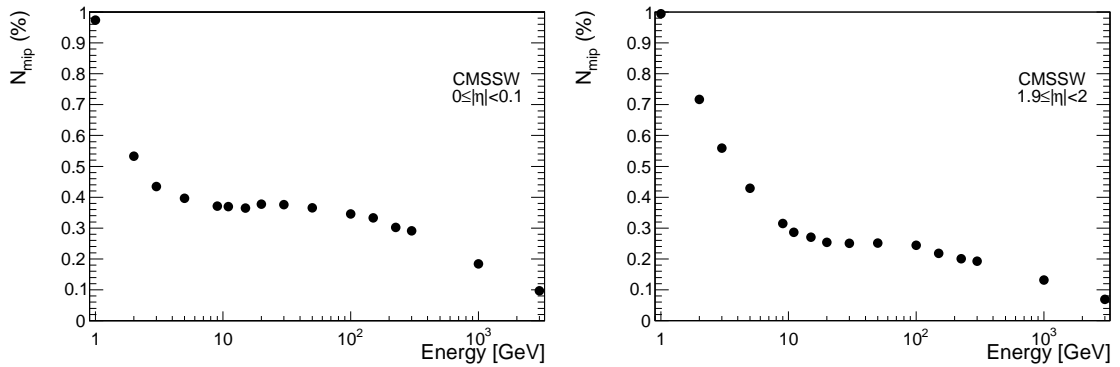


Figure 6.14: Plots of MIP percentage vs. energy for $i\eta = 1$ (in the barrel) and $i\eta = 20$ (in the endcap).

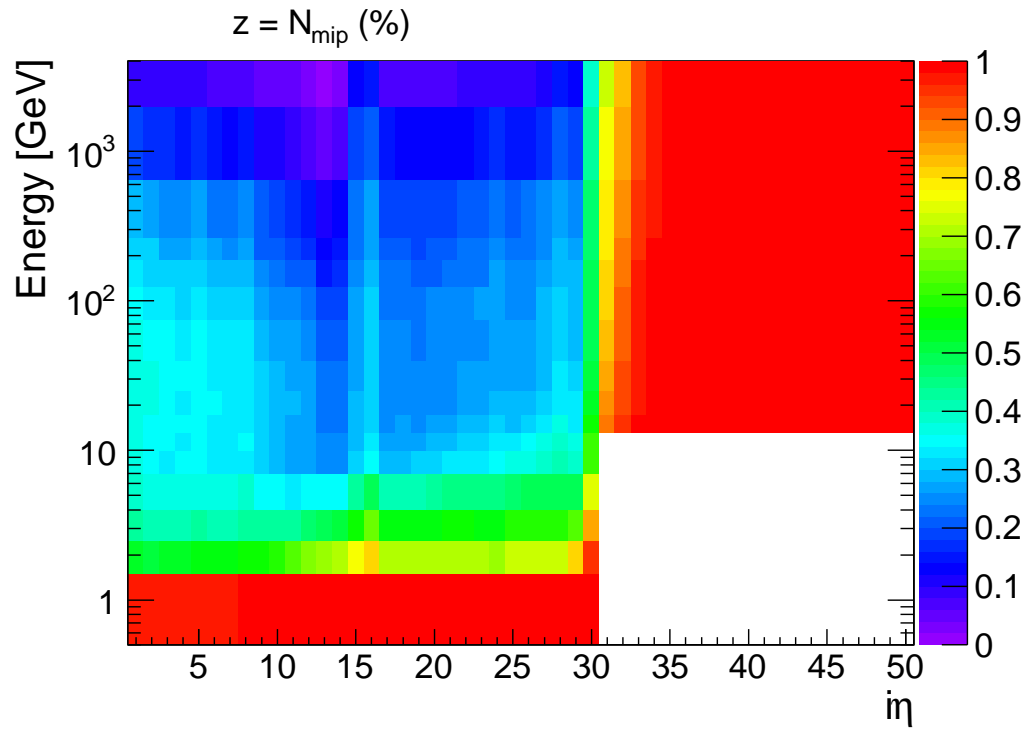


Figure 6.15: Plots of MIP percentage vs. energy and η for the entire calorimeter system.

6.4 Dose Rate Effects

6.4.1 Dose Rate Effect Models

6.4.2 Scintillator Radiation Damage Studies

Chapter 7: Conclusions

Appendix A: Full CLs Shape-Based Limits

To set limits using the modified frequentist CL_s procedure [10], two hypotheses are defined. The first is the null or background-only hypothesis H_0 or b , and the second is the alternate or signal plus background hypothesis H_1 or $s + b$.

$\mathcal{P}(\theta; N_{H_i})$ is defined as the Poisson probability to observe θ events in data given the hypothesis H_i which predicts N_{H_i} events. This probability can be defined generally for the whole sample, but also per bin for a histogram of some quantity, e.g. S_T , and/or per channel.

To obtain this probability, it is necessary to integrate over all of the nuisance parameters:

$$\mathcal{P}(\theta; N_{H_i}) = \int \text{Poisson}(\theta; N_{H_i}, \eta) f(\eta) d\eta \quad (\text{A.1})$$

where f is the probability density function (PDF) for the nuisance parameter η .

With those definitions, the test statistic \mathcal{Q} is written as a ratio of likelihoods for a basic counting experiment:

$$\mathcal{Q} = \frac{\mathcal{P}(\theta; N_{H_1})}{\mathcal{P}(\theta; N_{H_0})} \quad (\text{A.2})$$

Splitting into bins and channel gives:

$$\mathcal{Q} = \prod_{i=e,\mu} \prod_{j=0}^{n_{\text{bin}}} \frac{\mathcal{P}_{i,j}(\theta; N_{H_1})}{\mathcal{P}_{i,j}(\theta; N_{H_0})} \quad (\text{A.3})$$

For simplicity of computation, another form of the test statistic can be defined using the log likelihood ratio:

$$q = -2 \ln \mathcal{Q} \quad (\text{A.4})$$

To evaluate the test statistic as a function of the number of observed events θ , many simulated pseudo-experiments are performed. For each hypothesis, θ is varied according to the probability distribution of that hypothesis, and the value of \mathcal{Q} (or q) is kept for each θ value. To get \mathcal{Q} for the actual number of observed events, \mathcal{Q}_{obs} , the same procedure is followed using $\theta = N_{\text{obs}}$. The CL_{s+b} and CL_b variables correspond to the probability for \mathcal{Q}_{obs} to be greater than the \mathcal{Q} values obtained for the hypotheses H_1 and H_0 , respectively. When using q , the observed value should be smaller than the value for the hypothesis. A visual example of these variables is shown in Fig. A.1.

$$\text{CL}_{s+b} = \mathcal{P}(\mathcal{Q} \leq \mathcal{Q}_{\text{obs}} | H_1) = \mathcal{P}(q \geq q_{\text{obs}} | H_1) \quad (\text{A.5})$$

$$\text{CL}_b = \mathcal{P}(\mathcal{Q} \leq \mathcal{Q}_{\text{obs}} | H_0) = \mathcal{P}(q \geq q_{\text{obs}} | H_0) \quad (\text{A.6})$$

$$\text{CL}_s = \text{CL}_{s+b} / \text{CL}_b \quad (\text{A.7})$$

To set a mass limit on the signal hypothesis, the calculation of CL_s is repeated for different signal masses. Masses with $\text{CL}_s < 1 - \alpha$ are excluded at the α confidence

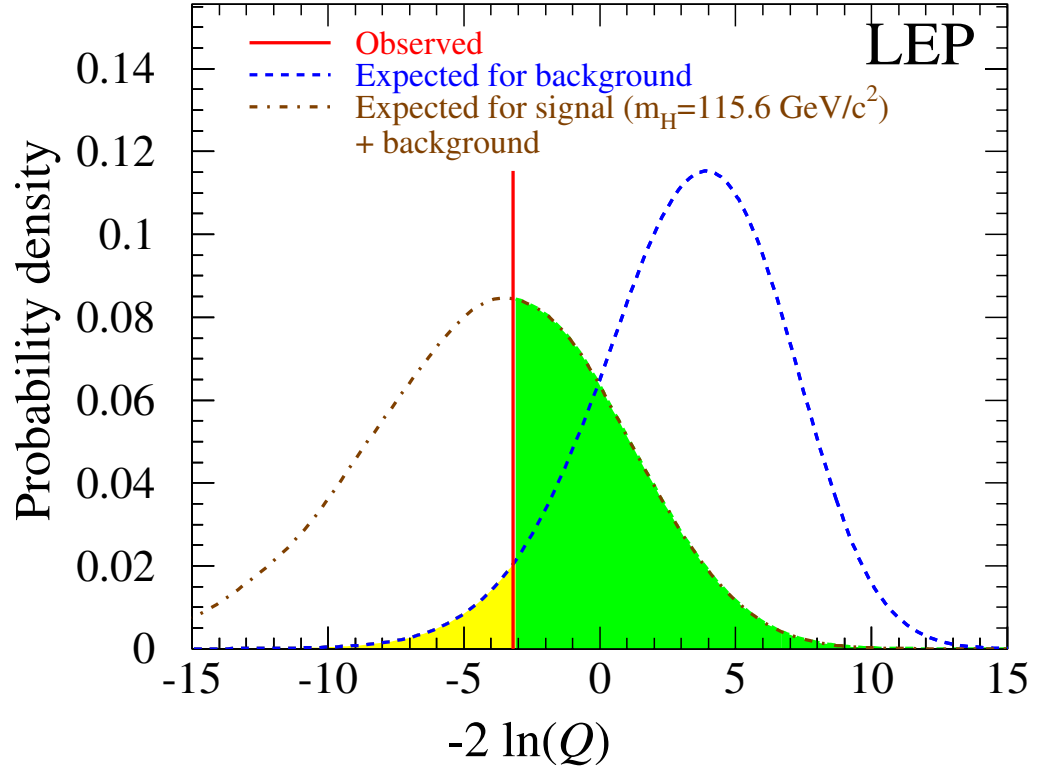


Figure A.1: Comparison of the observed value (red line) to the probability densities for H_0 (background only, blue line) and H_1 (signal + background, brown line) as a function of the log likelihood ratio. Green area: CL_{s+b} , yellow area: $1 - CL_b$. From [11].

level, typically 95%.

Appendix B: Event Displays

Appendix C: Table of Monte Carlo Datasets

Appendix D: CMS Collaboration

Bibliography

- [1] P. de Barbaro and V. Epshteyn. <https://indico.cern.ch/contributionDisplay.py?confId=254223&contribId=1>, May, 2013. viii, 9, 10
- [2] J. Mans, “CMS Technical Design Report for the Phase 1 Upgrade of the Hadron Calorimeter”, Technical Report CERN-LHCC-2012-015. CMS-TDR-10, CERN, Geneva, Sep, 2012. viii, 12, 17
- [3] CMS Collaboration, “The CMS experiment at the CERN LHC”, *JINST* **03** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004. 16
- [4] D. Gerbaudo, “HCAL Sampling Factors”. <https://twiki.cern.ch/twiki/bin/view/CMS/HcalSamplingFactors>. 20
- [5] J. Birks, “Theory and Practice of Scintillation Counting”. Pergamon Press, 1964. 21
- [6] S. Banerjee, “Data driven approach to calorimeter simulation in CMS”, *J. Phys. Conf.* **219** (2010), no. 3, 032030. 21
- [7] S. Banerjee and S. Sharma, “Material Budget in the Hadron Calorimeter”. CMS-IN-2009-004, Jan, 2009. 22
- [8] S. Abdullin et al., “Calorimetry Task Force Report”, Technical Report CMS-NOTE-2010-007, CERN, Geneva, Mar, 2010. 23
- [9] G. Grindhammer, M. Rudowicz, and S. Peters, “The fast simulation of electromagnetic and hadronic showers”, *Nucl. Instrum. Meth. A* **290** (1990) 469, doi:10.1016/0168-9002(90)90566-0. 27
- [10] A. L. Read, “Modified frequentist analysis of search results (the CLs method)”, CERN Report CERN-OPEN-2000-005, 2000. 40
- [11] A. L. Read, “Presentation of search results: the CL_s technique”, *J. Phys. G* **28** (2002), no. 10, 2693. ix, 42