**Energy sharing between close-by photons in the CMS HGCAL detector**

# **EM Showers Physics**

<https://pdg.lbl.gov/2011/reviews/rpp2011-rev-passage-particles-matter.pdf>

**EM Showers Introduction**

High energy

* photons generate e-p pairs, as long as they have sufficient energy (few )
* electrons/positrons emit photons via Bremsstrahlung (radiation produced decelerating as deflected by the field of another charged particle). Occurs above critical energy (wiki) or (Rossi, papers).

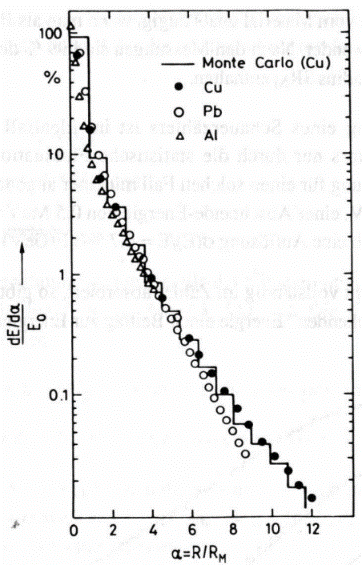
After they lost enough energy, they dissipate by ionization and excitation rather than by the generation of more shower particles, hence transversal size starts reducing.

**Shower Depth**

The longitudinal development is governed by the high-energy part of the cascade, hence the shower depth scales with the radiation length

where the radiation length (is the characteristic length of matter traversed by the phenomenon, defined as (note: )

* mean over which the electron loses all but 1/e (37%) of energy by bremsstrahlung
* 7/9 of the mean free path for pair production by a high energy photon

**Lateral Spread**

The lateral spread is mainly due to the multiple scattering of the electrons. Up to the shower maximum the shower is contained in a cylinder with radius < .

Beyond that point electrons are increasingly affected by multiple scattering, and the lateral size scales with the Molière radius of the material. The propagation of the photons causes deviations, however, roughly is contained in a cylinder of radius , within , within .

Note, *approximately*, where *.*

**Longitudinal Energy Deposition**

Can be useful to define, quantities

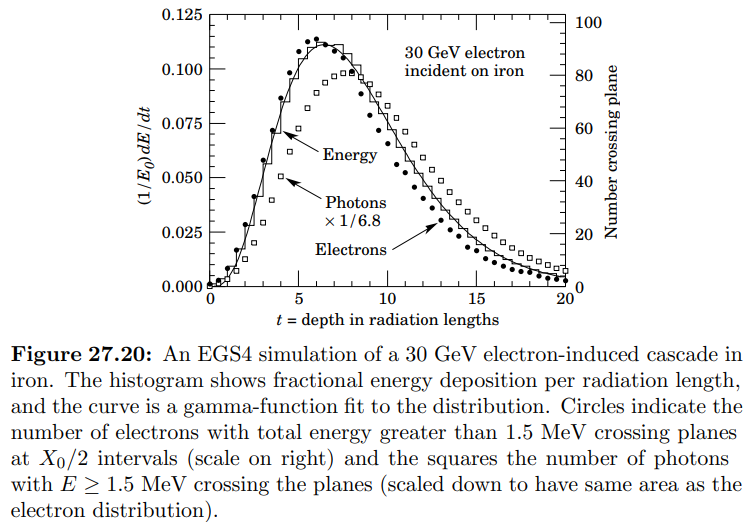
Note: with increasing depth, a larger fraction of the energy is carried by photons, which are also the main cause of possible small energy leakages.

The mean longitudinal profile of the energy deposition is reasonably well described (APART FOR FIRST 2 RADIATION LENGTHS) by a gamma distribution

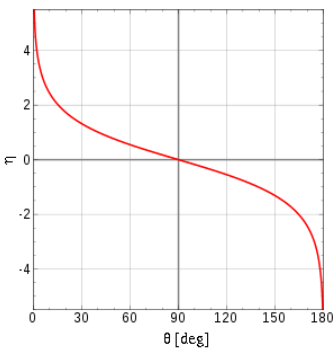
With shower maximum at as

respectively for <https://pdg.lbl.gov/2011/reviews/rpp2011-rev-passage-particles-matter.pdf> or <https://indico.cern.ch/event/318531/attachments/612850/843143/daniela_l5.pdf>. Note: .

Longitudinally, 95% energy deposited is contained in .



Lateral distributions are often represented as sum of 2 Gaussians ( more in G. Grindhammer et al., in Proceedings of the Workshop on Calorimetry for the Supercollider, Tuscaloosa, AL, March 13–17, 1989, edited by R. Donaldson and M.G.D. Gilchriese (World Scientific, Teaneck, NJ, 1989), p. 151; W.R. Nelson et al., Phys. Rev. 149, 201 (1966); G. Bathow et al., Nucl. Phys. B20, 592 (1970)).

**Pseudorapidity**

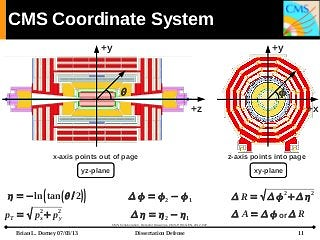
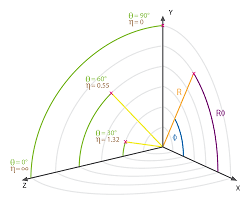
describes the angle of a particle relative to the beam axis. It is defined as

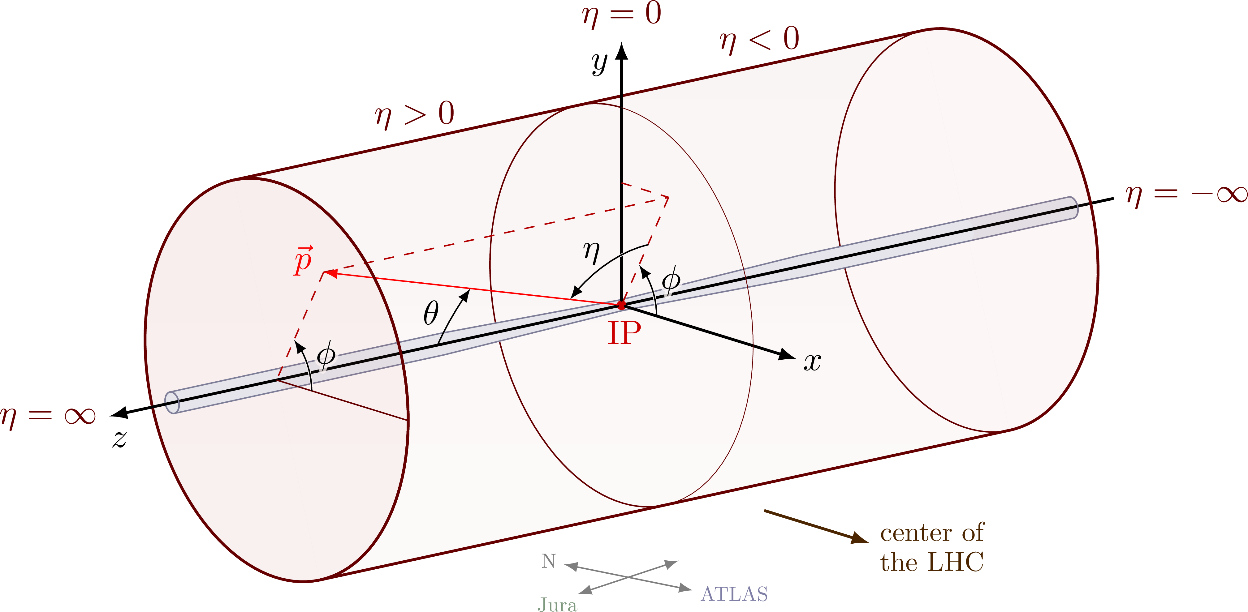
being the angle between and beam’s direction. Also:

being the longitudinal component of momentum, along the beam’s direction, whereas is the transversal component.

In massless/-velocity limits, coincides with rapidity, and between particles is independent of the longitudinal boost (invariant).

**CMS Coordinates ( Calorimeter Coordinates)**



Each layer is to the beam line, hence lies on xy plane. Layers are stacked one after the other longitudinally, i.e. layer # increases along z.

# **The High-Granularity Calorimeter**

**Introduction**

Assume to work with 140 – 200 collision events per bunch crossing.

**Granularity & Resolution**

* Lateral Granularity: good signal-to-noise ratio, allowing separation of close shower events and observation of narrow jets.
* Longitudinal Granularity: good EM energy resolution, patter recognition and discrimination against pileup.
* High Time Resolution: identification of different vertices and interactions, expecting a 25 ps resolution.

**Level-1 Trigger**

Level-1 Trigger separates interesting events from “bad” ones, based on multiple factors. The pointing capability of HGCAL opens the possibility of a dedicated trigger at Level-1 for displaced objects with a decay length, cτ, larger than a few centimetres.

Immagine che contiene testo

Descrizione generata automaticamente**Compartments**

**Longitudinally** (not transversally) it has 3 compartments for a total of 210.4 cm length:

1. Neutron Moderator: cm long.
2. Electromagnetic Compartment (CE-E): inner one, 28 layers for a total 34 cm width. Alternate layers contribute to L1 trigger primitives (bluish).
3. Hadronic Compartment (CE-H): outer one, all layers contributing to L1 trigger.
   1. FH: 12 forward layers for 60 cm (dark grey)
   2. BH: 12 backing layers for 100 cm (light grey).

Each layer is to the beam line, hence lies on the x-y plane. Layers are stacked one after the other longitudinally, i.e. layer # increases along z, as defined by the beam line (orangish).

**Reconstruction & Detector Performance**

**Aims**

5 dimensional set of row information: in severe 200 pileup env.

* Hadronic and EM showers are individually reconstructed and identified
* Charged hadrons matched to tracks reconstructed in the tracker.
* Jets (and missing energy) measured using best available estimations, are obtained mainly from tracker for charged hadrons and from calorimeter for neutral hadrons and photons.
* Track-cluster matching refines electron & converted photon reconstruction
* Fine granularity allows identification for electrons and photons.
* Calorimeter is sufficiently deep to provide excellent muon identification.

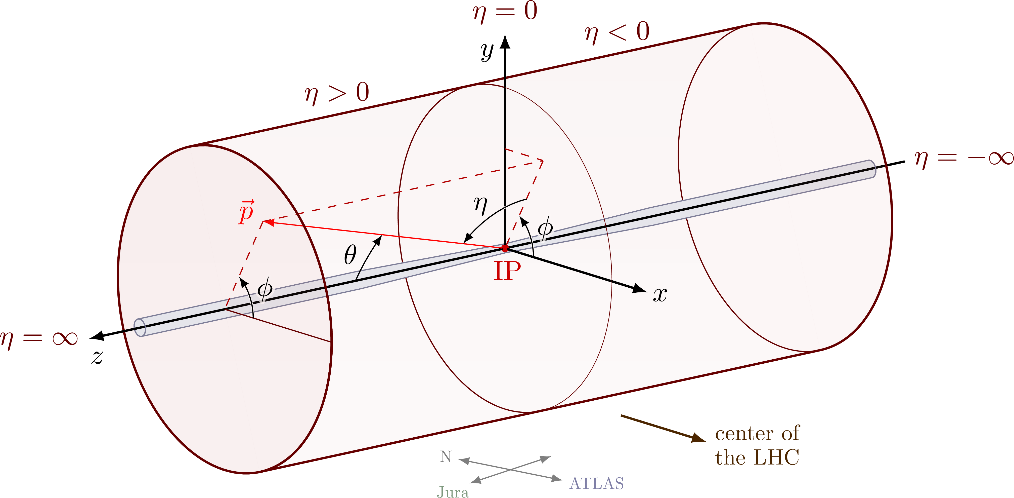
**EM Showers (5.1.3)**

Energy is estimated by summing cells in each layer in restricted area of 2.6 cm (20-40 cells) 🡪 misses a bit: the stochastic term due to transverse motion and phton counting, with fitted energy resolution .

Position resolution is better than 1mm in each layer effortlessly. Can be made much better, with sophisticated multivariate corrections for numerous subtle biases. Angular resolution is 7mrad (for ) and again it is expected to significantly improve.

EM showers have small containment radii and fast development, and usually give many cells with significant energy deposits: the shower of a pT = 2 GeV photon at η = 1.7 has an average of more than 20 cells above a 12 fC threshold, and the corresponding number for pT = 60 GeV is more than 100.

(Hadronic showers usually have a prompt core and later developing components that propagate laterally with respect to the shower axis, and the number of cells above a 12 fC threshold has large event-to-event fluctuations.)



**Clustering (5.1.4, 8.3, 10.1)**



First step in reconstruction: collect signals of energy deposited into clusters corresponding to showers, calibrated to reconstruct energy lost in absorber layers:

1. Construct 2D clusters in each of 52 layers (Section 10.1.2)
2. Collect them in 3D multi-cluster

It is suitable for implementation in GPUs, and may be useful as a future fast high level trigger (HLT) algorithm. However, it is expected an algorithm performing 3D clustering directly may achieve better performance.

Super-clustering: except photons not converting in tracker material (before 1st layer), electrons and photons generally start a shower before the 1st layer, giving rise to distinct clusters of energy spread in by -field, but with little spread.

Hadron showers also tend to result in many multi-clusters, though dominated by irregularity of shower itself and transverse development, rather than . Requires harder reconstructing algorithm, done with some assistance from Monte Carlo.

Immagine che contiene mappa

Descrizione generata automaticamente

Energy deposited (colour coded in log scale) by individual reconstructed hits of a VBF (*vector boson fusion*) jet and high photon in an environment of 200 interactions per bunch-crossing. The energy is projected onto one plane and longitudinally developed in 6 layers of CE-E.

**Stage1 – 2D Custering**: finds clusters by finding seed trigger cells (TC) with and secondary TCs with (actually energy density defined as a function of energy in chosen cell and in surrounding cells within some distance) and putting neighbours together, with a 5x5 maximum TCs along coordinate system directions, with seed in centre. Note: a cluster requires *at least* one seed TC. Each cluster is made of a memory that stores the TCs energy and an occupancy word, which maps all the TCs found within the cluster. Also defines energy (binned) map in space.

The algorithm already implemented should be the major part of the firmware design. However, it has some limitations: operations are limited to data within a single link so clusters are not formed from TCs from different links, TCs can be double counted (i.e. same TC assigned to more than one cluster), the clusters are limited in size because of the 5 × 5 arrays used, there is a limit of six seed TCs per link, and the energy map bin data mapping is not yet included. Allowing for these additional computations, the maximum total latency is estimated to be up to 1 µs.

Immagine che contiene testo

Descrizione generata automaticamente

Immagine che contiene tavolo

Descrizione generata automaticamente

Simulations show 2D cluster have only one seed. Average cluster has 70 bits.

**Stage2 – 3D clustering**: each 2D cluster can seed a 3D cluster and all 2D clusters lying within a distance in the projected plane smaller than some value (typically ) are added to the 3D cluster. Likelihood algorithm considers spread to decide whether to put them together or not. 3D clusters’ energy and position are found by a weighted sum of associated 2D, with each layer differently weighted (based on 4 possibilities: EM or hadron, with or without pileup removal, hadron used in general).

Also, energy map is built as a weighted sum of stage1 results.

Immagine che contiene testo

Descrizione generata automaticamente

Immagine che contiene tavolo

Descrizione generata automaticamente

**Machine Learning (Deep Neural Networks, 10.1.4**):Can find more differences in shower shapes of EM, muons and hadrons.

**SimClusters (10.1.5)**: Optimistic about ability to distinguish and separately cluster overlapping showers.

**Trigger Simulation & Performance (5.3)**

The output of TPG, i.e. trigger primitives, is a list of 3D clusters and an transverse energy map for each endcap.

Other calculated characteristics: fractions of energy in layers, shower widths and depths, internal structure such as local maxima (indicating overlapping particles), etc.

Primitives are used as input to central L1 correlator, which creates actual triggers using much more information (it accesses data from tracker, calorimeters, muon systems).

**Time Measurements (5.5)**

To make a time measurement we require at least three cells within a selection radius, ρ, of the shower axis, each with an energy deposit of >12 fC needed to fire the ToA. Showers that fail this requirement obtain no time measurement, resulting in a measured inefficiency, generally small, that is quantified below.

Great efficiency and resolution for photons. Trade off for hadrons.

**Electron/Photon Identification & Reconstruction**

Design and constraints of the TPG primitives have a large impact on electromagnetic object energy resolution and identification. They can be distinguished in central L1T correlator by matching L1 tracks, but for HGCAL they are treated very similarly.

An electron or photon is reconstructed from a single TPG 3D cluster.

A TPG-specific energy correction factor is applied to correct for the energy lost by the 2 MIPT cut and the clustering algorithms. The cluster is required to pass a selection based on its transverse and longitudinal shape:

* longitudinal profile selection uses the length of the cluster, the position of the start and the position in the layer with the maximum energy. These are used in a boosted decision tree to select the clusters that are compatible with an EM shower.
* transverse shape is measured by the cluster’s width in radial direction, computed in each layer included in the cluster, weighted according to the energy in the layer and combined. It results in an average radial width, used to identify EM objects.

To reduce pileup sensitivity of this transverse width, only trigger cells in the cluster’s core are used (cluster’s core is radius cone. Selection is chosen such that the total efficiency loss resulting from cuts is for . Tighter cuts with 10% efficiency loss are used above that threshold, reducing higher background.

**Physics Performance (11)**

**Photons (11.1.2)**

Photons are identified mainly through three components:

* Shower shape in calorimeters (distinguish EM vs Hadronic), using variables from multi-clustering.
* Track isolation and matching (rejects electrons & charged hadrons, leaves photons)
* Calorimeter-based isolation (additional rejection of neutral hadrons and photons in jets). Can find new set of observables which might be useful.

All variables are fed to a Boosted Decision Tree (BDT).

Note: efficiency is defined as the fraction of true reconstructed photons that pass a given threshold in BDT score output, and the photon misidentification probability is defined as the number of reconstructed background photons passing the same BDT score threshold divided by the total number of events processed.

example in section 11.2.

# **Energy Sharing & Separation**

**RESOLVING HOTONS FROM BACKGROUND AND MISEDINTIFICATION**

(the general problem, greatly already solved)

<https://journals.aps.org/prd/pdf/10.1103/PhysRevD.83.052005>

**Section V**

Immagine che contiene testo, quotidiano, documento, screenshot

Descrizione generata automaticamente

*NOTE: low energy showers deposit most of their energy in 2nd and 1st compartments, that’s why they are used in the following, but could find those quantities for all layers, or first, say, 10.*

*Here there are 3 long layers.*

*Strips are long constant cells covering a large phi range.*

**Appendix A**

Immagine che contiene testo

Descrizione generata automaticamente

Immagine che contiene testo

Descrizione generata automaticamente

Immagine che contiene testo

Descrizione generata automaticamente

Immagine che contiene testo

Descrizione generata automaticamente

[**https://arxiv.org/ftp/arxiv/papers/0901/0901.0512.pdf**](https://arxiv.org/ftp/arxiv/papers/0901/0901.0512.pdf) , from page 94

3 photon identification methods have been developed at present in ATLAS:

* a simple cut-based identification method
* a Log-likelihood-ratio-based identificatiomethod (LLR)
* covariance-matrix-based identification method (H-matrix).

**Quantities** (same as above, but better explanation)

Immagine che contiene testo

Descrizione generata automaticamente

Immagine che contiene testo

Descrizione generata automaticamente

**Covariance Method** (page 100)

10 photon shower-shape variables used:

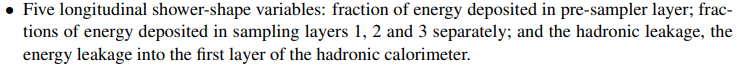


Immagine che contiene testo

Descrizione generata automaticamente

**RESOLVING TWO CLOSE PHOTONS**

[**https://arxiv.org/ftp/arxiv/papers/0901/0901.0512.pdf**](https://arxiv.org/ftp/arxiv/papers/0901/0901.0512.pdf)

Immagine che contiene testo

Descrizione generata automaticamente

Immagine che contiene testo

Descrizione generata automaticamente

[**https://www.osti.gov/etdeweb/servlets/purl/50345**](https://www.osti.gov/etdeweb/servlets/purl/50345)

Proposes a method in which associated to presence of 1 and 2 gammas in peak cell are compared, and based on their difference, one of the two hypotheses is chosen.

The calculation involves computations with detected and calculated energies and coordinates.

[**https://arxiv.org/pdf/1802.00672.pdf**](https://arxiv.org/pdf/1802.00672.pdf)

Doesn’t seem to show a method, just results. Oppositely the following is a comprehensive guide to the problem and results, though not to the method as such: [**https://twiki.cern.ch/twiki/pub/CALICE/CaliceAnalysisNotes/CAN-057.pdf**](https://twiki.cern.ch/twiki/pub/CALICE/CaliceAnalysisNotes/CAN-057.pdf)**.** For them, “The mixed event is considered to be successfully reconstructed if

• it contains exactly two reconstructed EM showers;

• energies and X, Z barycenter coordinates agree within ±20% and ±5 mm, respectively, of the energies and the coordinates reconstructed in the single shower events”.

**OTHER POSSIBLY USEFUL QUANTITIES**

[**https://link.springer.com/content/pdf/10.1140/epjc/s10052-012-1909-1.pdf**](https://link.springer.com/content/pdf/10.1140/epjc/s10052-012-1909-1.pdf)

**Immagine che contiene testo

Descrizione generata automaticamente**

Immagine che contiene testo

Descrizione generata automaticamente

**Machine Learning for HGCAL**

<https://www.hep.ph.ic.ac.uk/machine-learning/seminars.html>

* <http://www.hep.ph.ic.ac.uk/machine-learning/seminars/videos/200522_kieseler.mp4> also suggests deep convolutional networks. He suggests GravNet/GarNet as lighter faster yet flexible solutions, although still at initial stages. More info here: <https://www.researchgate.net/figure/Architecture-diagram-of-the-GravNet-graph-neural-network_fig1_354073314>.

**Deep Convolutional Network**

[**https://arxiv.org/pdf/2111.14939.pdf**](https://arxiv.org/pdf/2111.14939.pdf)

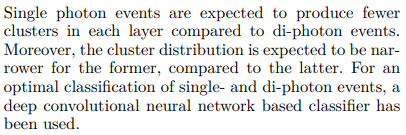


Immagine che contiene testo, quotidiano, documento

Descrizione generata automaticamente



Immagine che contiene testo

Descrizione generata automaticamente

Immagine che contiene tavolo

Descrizione generata automaticamente