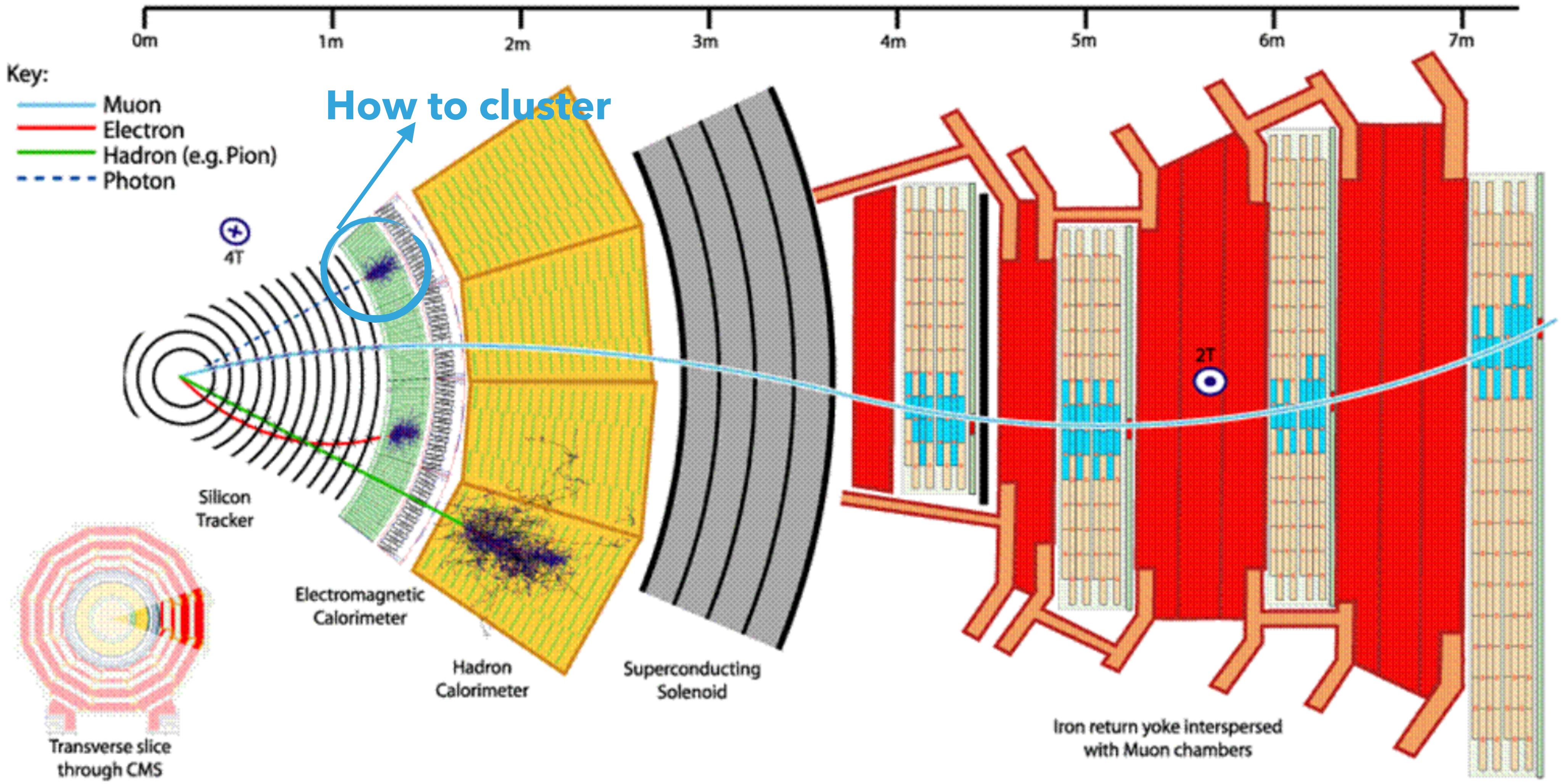
Clustering of neutral (γ and π^0) particles

SHILPI JAIN(TIFR)

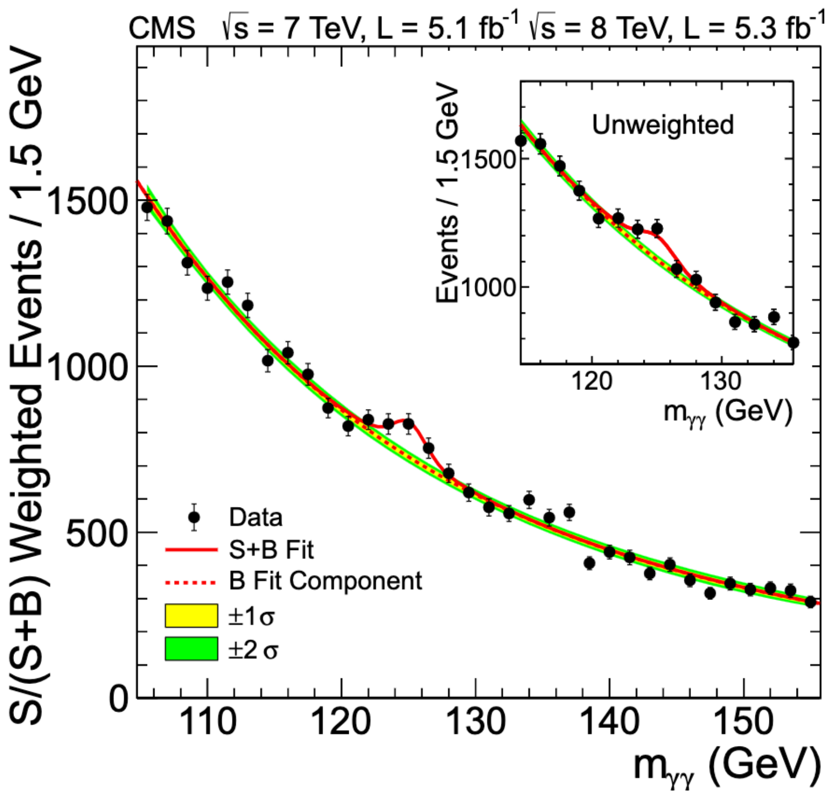
EHEP2024 School, TIFR Mumbai

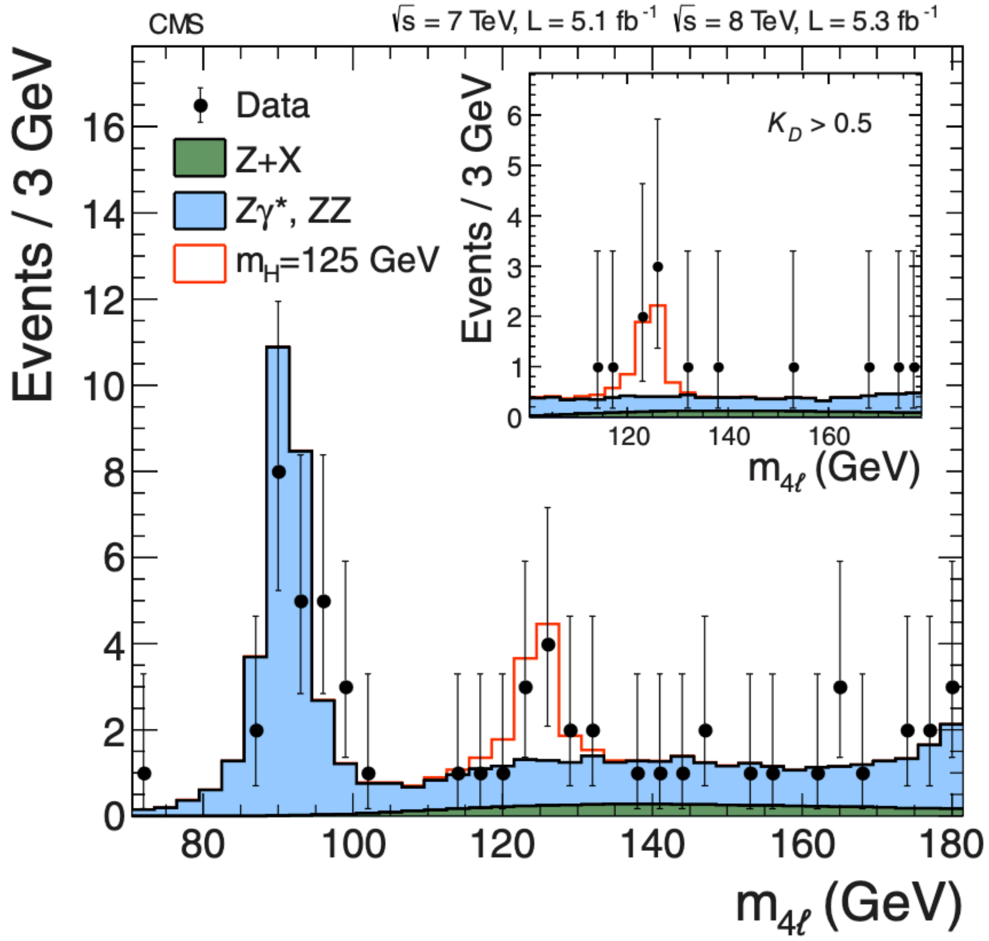
7th Feb, 2024

TODAY'S AIM


0m
1m
2m
3m
4m
5m
6m
7m
Key:
Muon
Electron
Hadron (e.g. Pion)
Photon
How to cluster
4T
2T
Silicon Tracker
Electromagnetic Calorimeter
Hadron Calorimeter
Superconducting Solenoid
Iron return yoke interspersed with Muon chambers
Transverse slice through CMS

ELECTRONS, PHOTONS AND PHYSICS SEARCHES

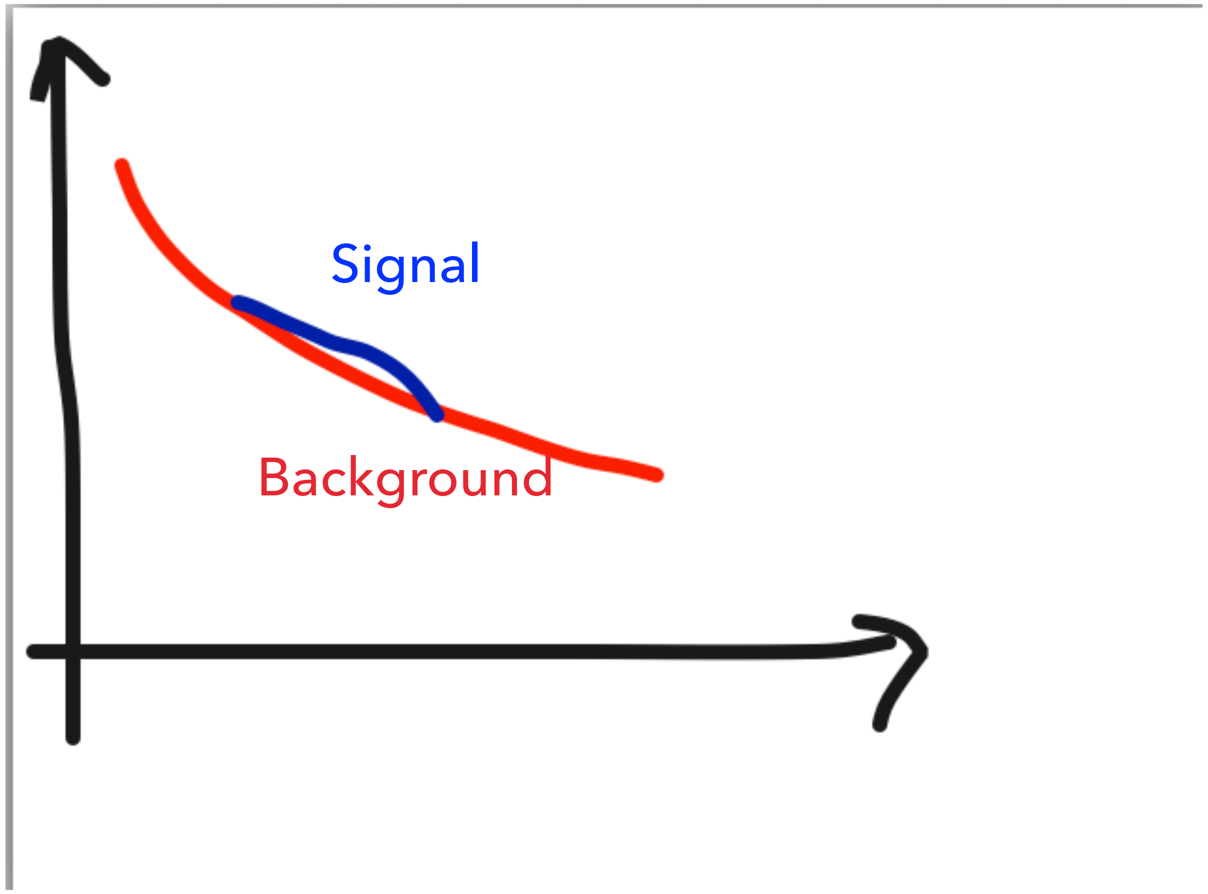




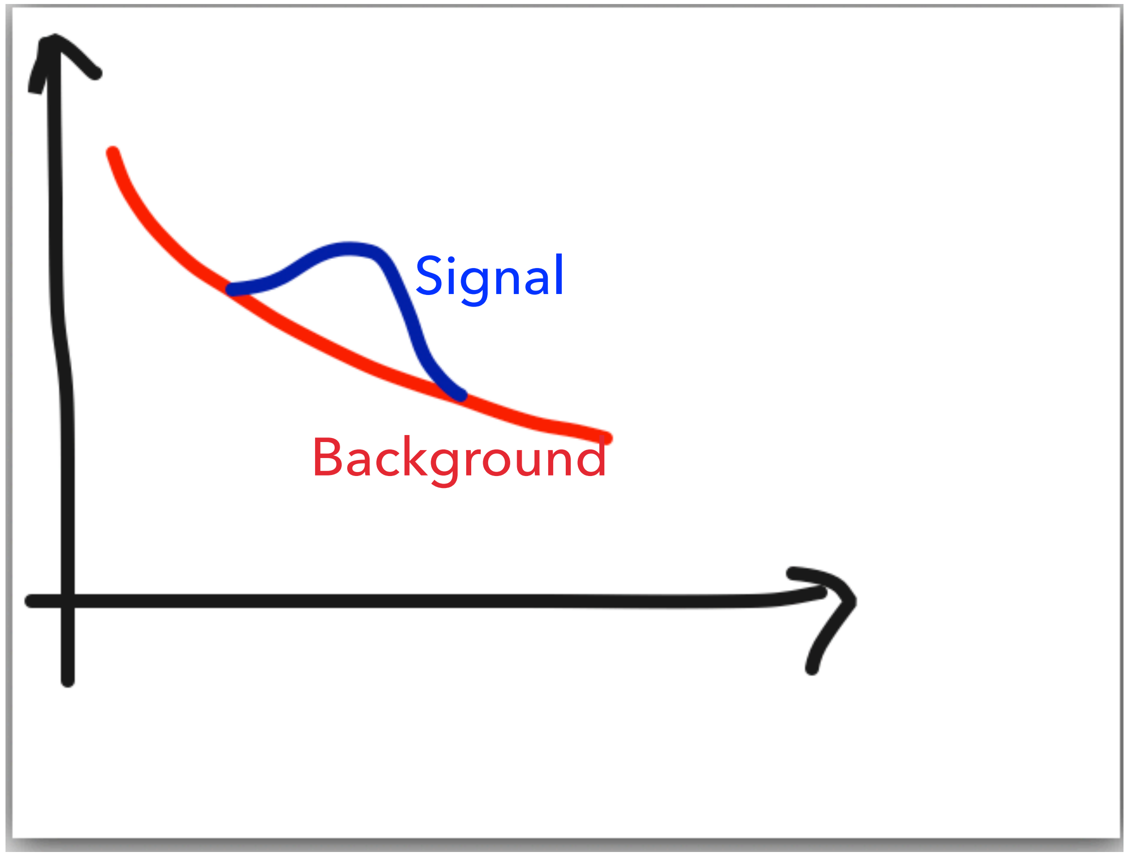
- Electron and photons are one of the most important signatures for the searches and measurements
- H–>γγ and H–>ZZ–>4leptons were two golden discovery channels
- For any search or measurement, it is important to make sure that all the final state objects (in today's case photons/ electrons) are calibrated and understood well
  - Clustering
  - Calibration
  - Mitigation of noise
  - Discrimination against pi0s

WHY IS RESOLUTION IMPORTANT

bad resolution

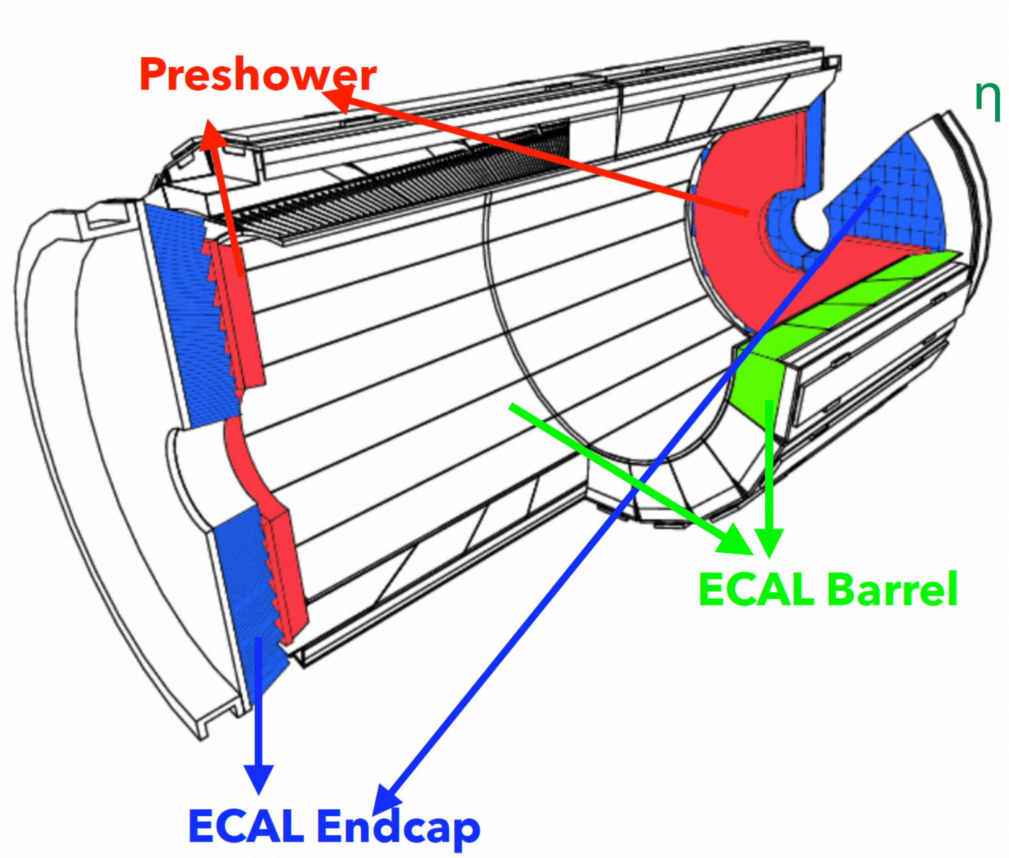


good resolution

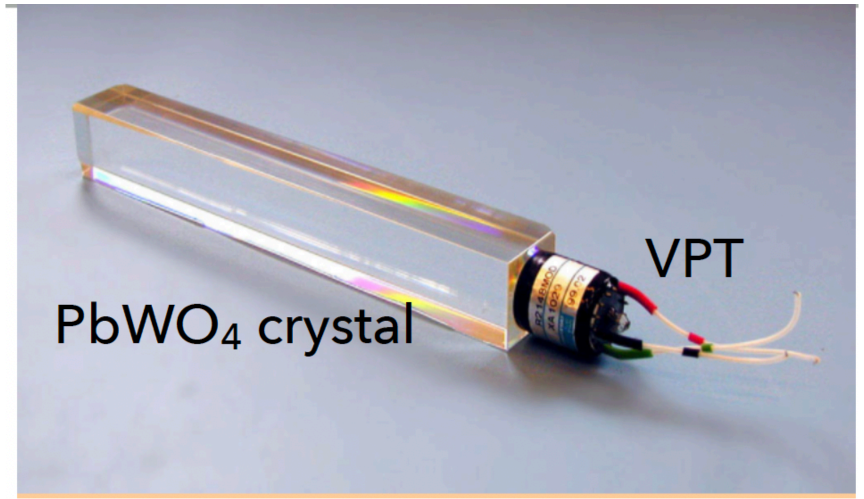


- In which scenario is it easy to discover signal?

CMS Electromagnetic Calorimeter (ECAL)



η is pseudorapidity



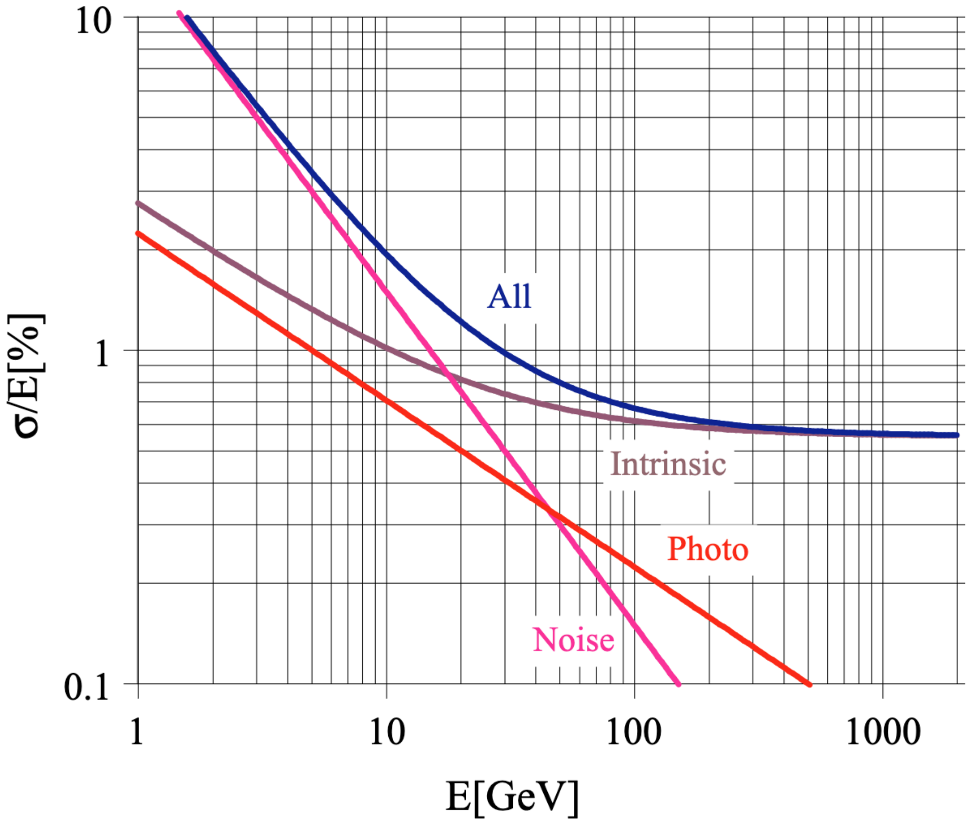
- Lead tungstate crystals ($PbWO_4$)
  - Preferred choice because of short radiation length (0.89 cm) and small moliere radius (2.2 cm)
- Barrel (|η| < 1.48): 61200 crystals read by Avalance Photo-Diodes (APDs)
- Endcaps (1.48 < |η| < 3): 14648 crystals read by Vaccum Photo-Triodes (VPTs)
- Dimensions of the crystal:
  - EB: 21.8 x 21.8 x 230 mm^3, Depth in X0 ~ 25.8
  - EE: 24.7 x 24.7 x 220 mm^3, Depth in X0 ~ 24.7

ECAL sits inside 3.8 T Magnetic field
This causes shower to spread in φ direction

- Short radiation length (X_0) = 0.89 cm. Reminder: In 1X0, the remaining energy of an electron is 1/e of the initial energy ($E = E0\ e^{-x/X0}$)
- Small Molière radius (R_M is the radius of a cylinder with 90% energy containment) = 2.19 cm
- Further effects broaden the shower in the ECAL:
  - Effect of magnetic field
  - material before the ECAL

RESOLUTION

Energy resolution of a homogeneous calorimeter



$$(\sigma/E)^2 = \underbrace{(a/\sqrt{E})^2}_{\text{Photo statistics}} + \underbrace{(\sigma_n/E)^2}_{\text{Noise}} + \underbrace{c^2}_{\text{Constant}} \quad \text{(E in GeV)}$$

$a = 2.8\%$

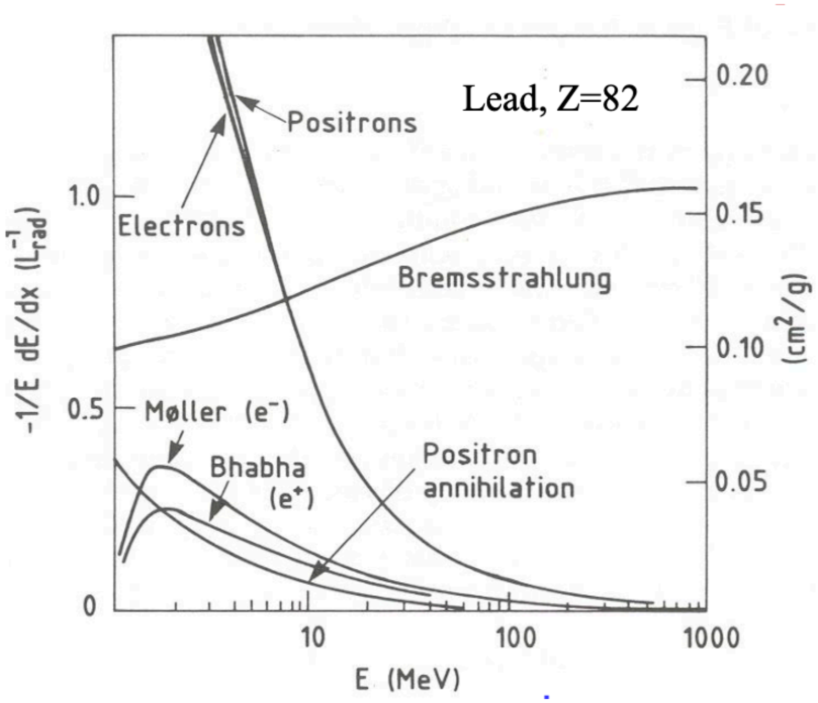
$\sigma_n = 12\%$

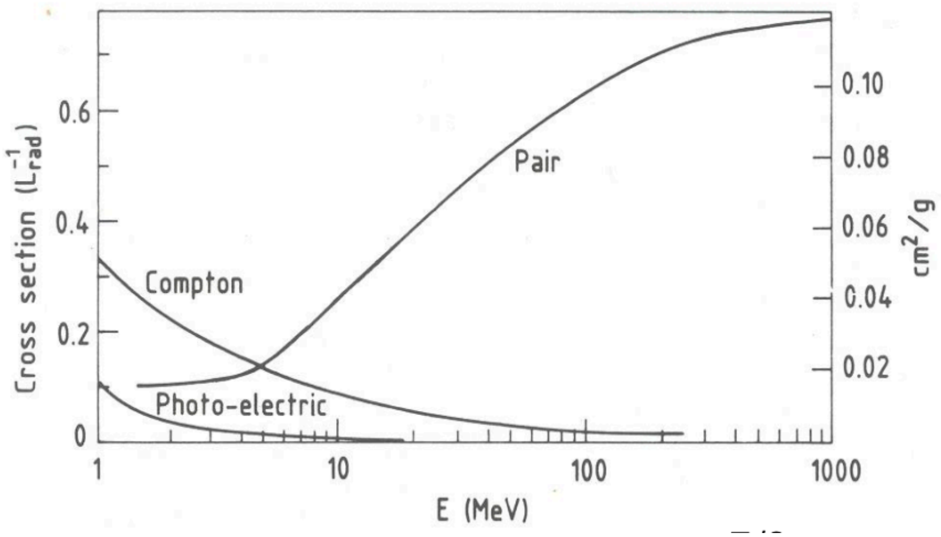
$c = 0.3\%$

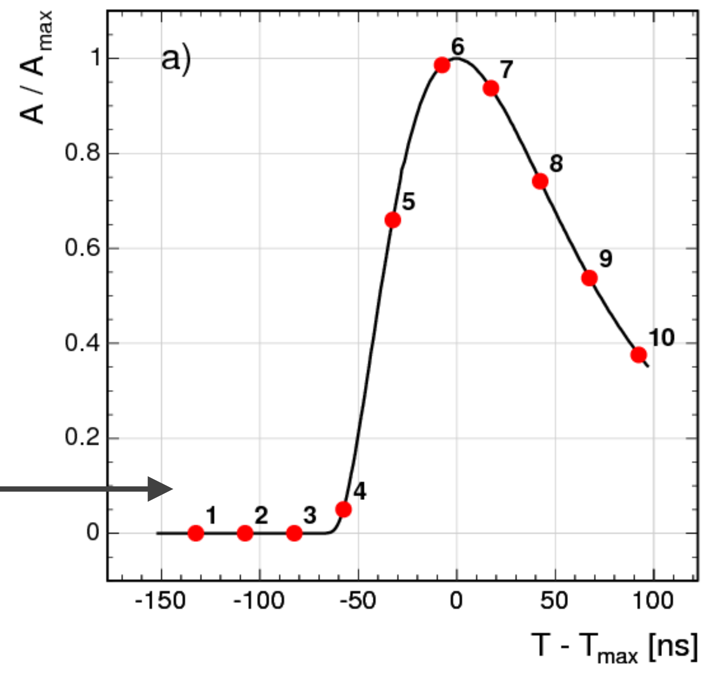
"+" means addition in quadrature

- Photo-statistics:
  - Due to number fluctuations in the number of shower particles
  - Noise: Electronic noise and the pileup energy
    - Corresponds to 40 MeV in the EB and 60 MeV in the EE
  - Constant term: Due to limited shower containment, longitudinal light collection, inter-calibration
    - Dominates for high energy electrons and photons

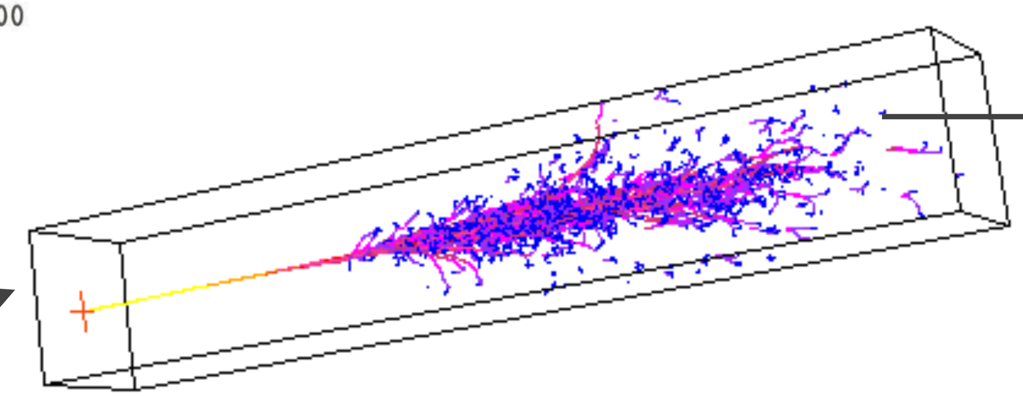
ENERGY DEPOSITION IN A CRYSTAL







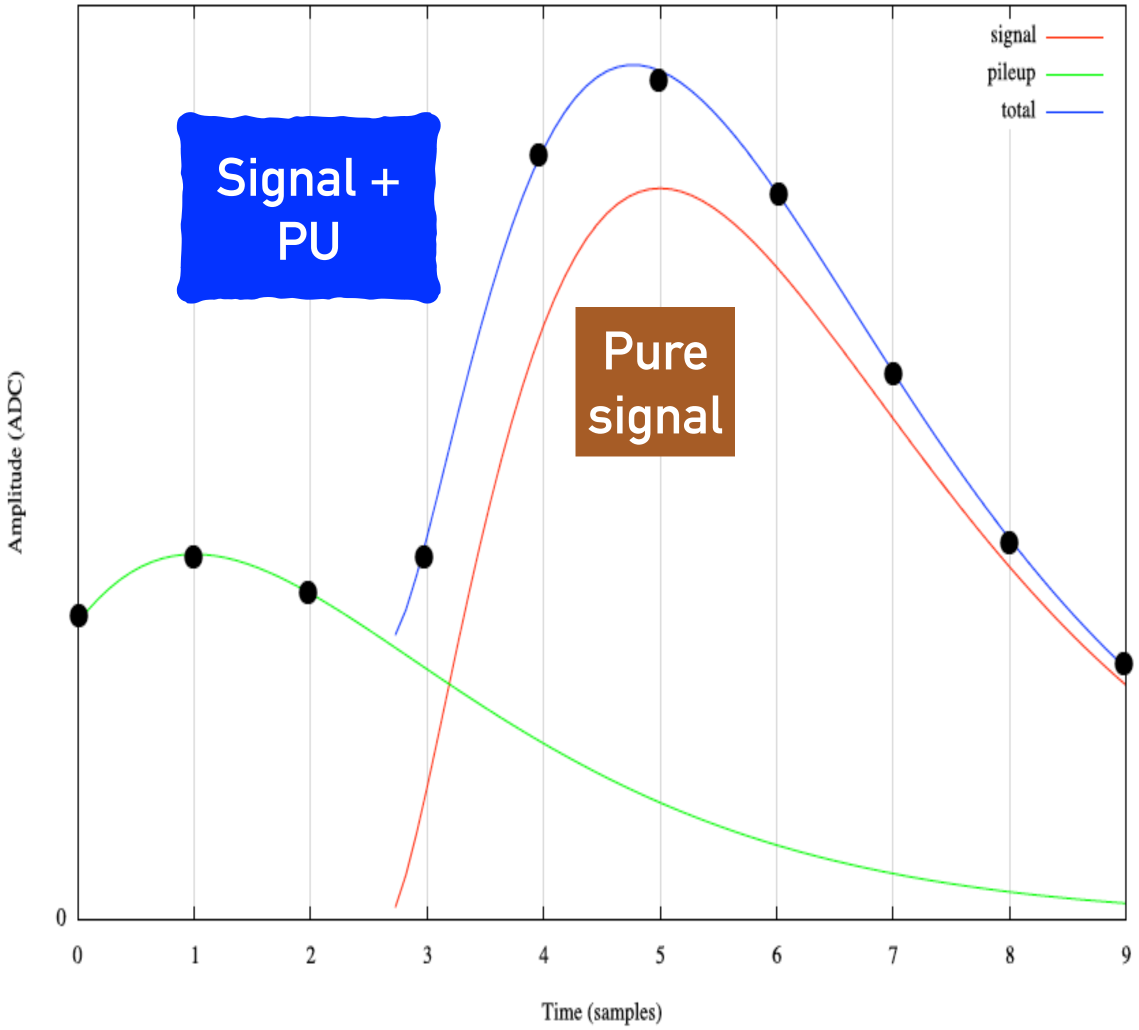
Relevance of the process is dependent
on the energy of the electron/photon



- Bremsstrahlung and pair product create shower.
- Ultimately the scintillation photons (detectable signal) are created due to ionization energy loss

AMPLITUDE RECONSTRUCTION

- Pulse shown in the previous slide is a pure signal pulse
- But that is not the complete scenario in pp collisions - there is PU!
- In the presence of out-of-time PU, the main signal gets energy from additional bunch crossing
  - Important to subtract this contribution
- To take care of this, in Run II, multi-fit method was developed



https://indico.cern.ch/event/292930/contributions/671061/attachments/547860/755142/edm-ecalreco-pu-21Aug2014.pdf

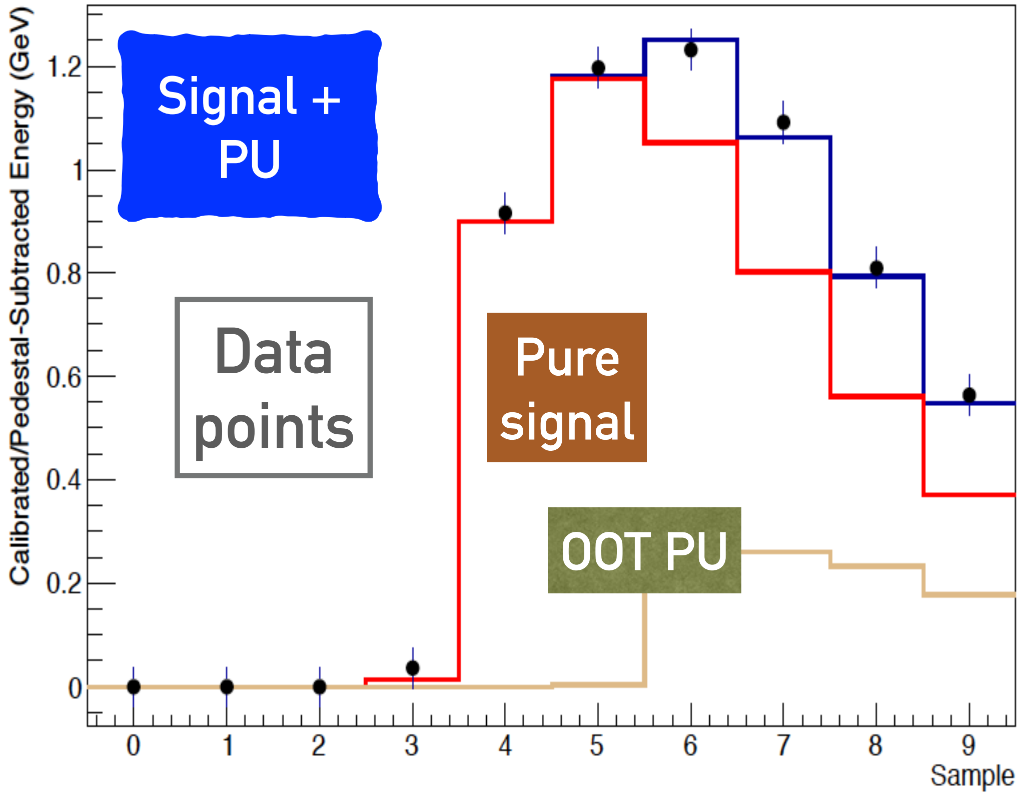
MULTI-FIT METHOD — PULSE SHAPE FITTING OF RUN II

- This method is based on template fitting
- Take the 9 pulse shape templates from Out Of Time (OOT) pileup (5 from previous and 4 from the next BXs)
  - These are essentially same as signal templates but shifted in multiples of 25 ns
- Minimize the X² given as:

$$\chi^2 = \sum_{i=1}^{N_{sample}} \frac{\left(s_i - \sum_{j=1}^{N_{pulse}} A_j p_{ji}\right)^2}{\sigma_i^2}$$

- Or in matrix formulation:

$$\chi^2 = (\mathbf{s} - \mathbf{pA})^T \boldsymbol{\Sigma}^{-1} (\mathbf{s} - \mathbf{pA})$$



- After fitting, we get fraction of each pulse shape contributing to signal
- Which means, we get the amplitude of the pure signal

AMPLITUDE TO GEV

Reference

$$E_{e,\gamma} = F_{e,\gamma} \cdot \left[G \cdot \sum_i \left(S_i(t) \cdot C_i \cdot A_i \right) + E_{ES} \right].$$

ADC2GeV

Time dependent correction

intercalibration constant

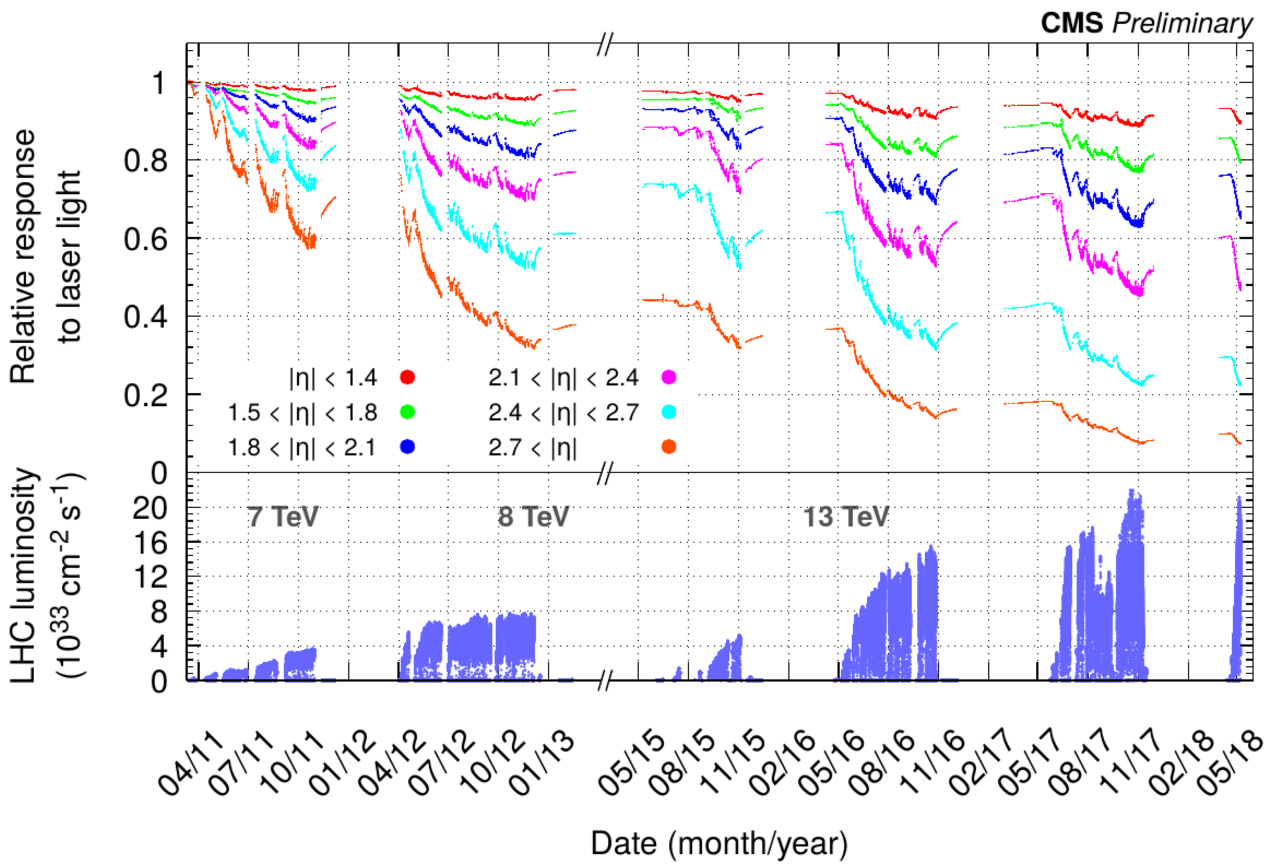
pulse amplitude

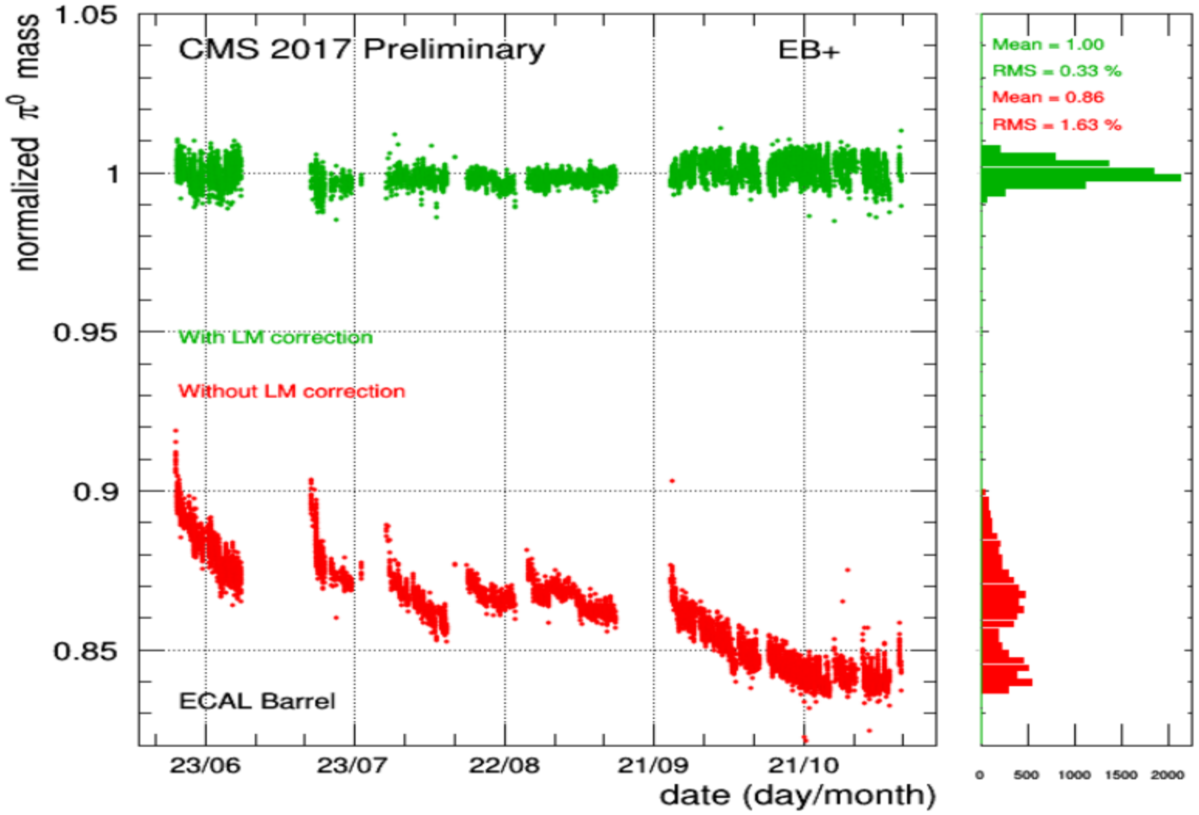
AMPLITUDE TO GEV

$$\Big| G \cdot \sum \Big(S_i(t) \cdot C_i \cdot$$

Time
dependent
correction

TIME DEPENDENT CORRECTIONS: S(T)





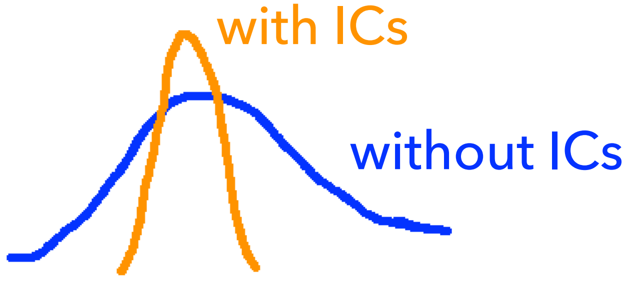
- ECAL response decreases with time
- Can degrade the performance
- Dedicated laser monitoring system corrects for the degraded response

AMPLITUDE TO GEV

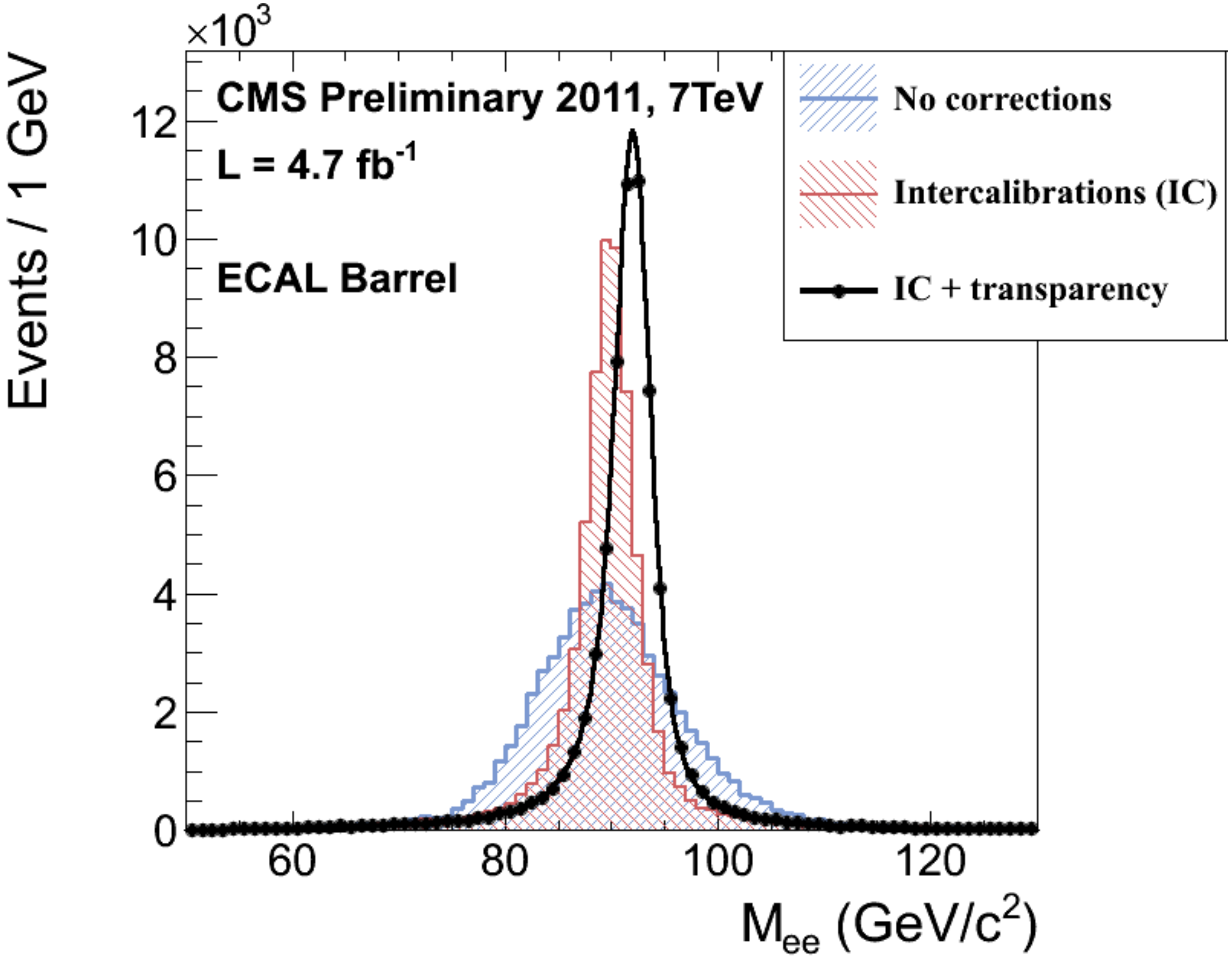
$$G \cdot \sum (S_i(t) \cdot \underset{\substack{\uparrow \\ \text{intercalibration} \\ \text{constant}}}{C_i} \cdot$$

INTERCALIBRATION CONSTANTS (IC) – THE C TERM

- All the crystals in ECAL do not have same response.
- What this means is that a 10 GeV electron can give a final reconstructed signal of 5 GeV in 1 crystal and full 10 GeV in another crystal.
- So essentially the intercalibration constant for the 1st crystal = 2.
- Resolution is written as: $\frac{\sigma(E)}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c$
  - where a: statistical term; b is the noise term and c is the constant term.
  - For high enough energies (typically > 100 GeV), c is the dominating term
- ICs have the dominant contribution in the constant term of the resolution
  - Important to have the precise estimation of ICs

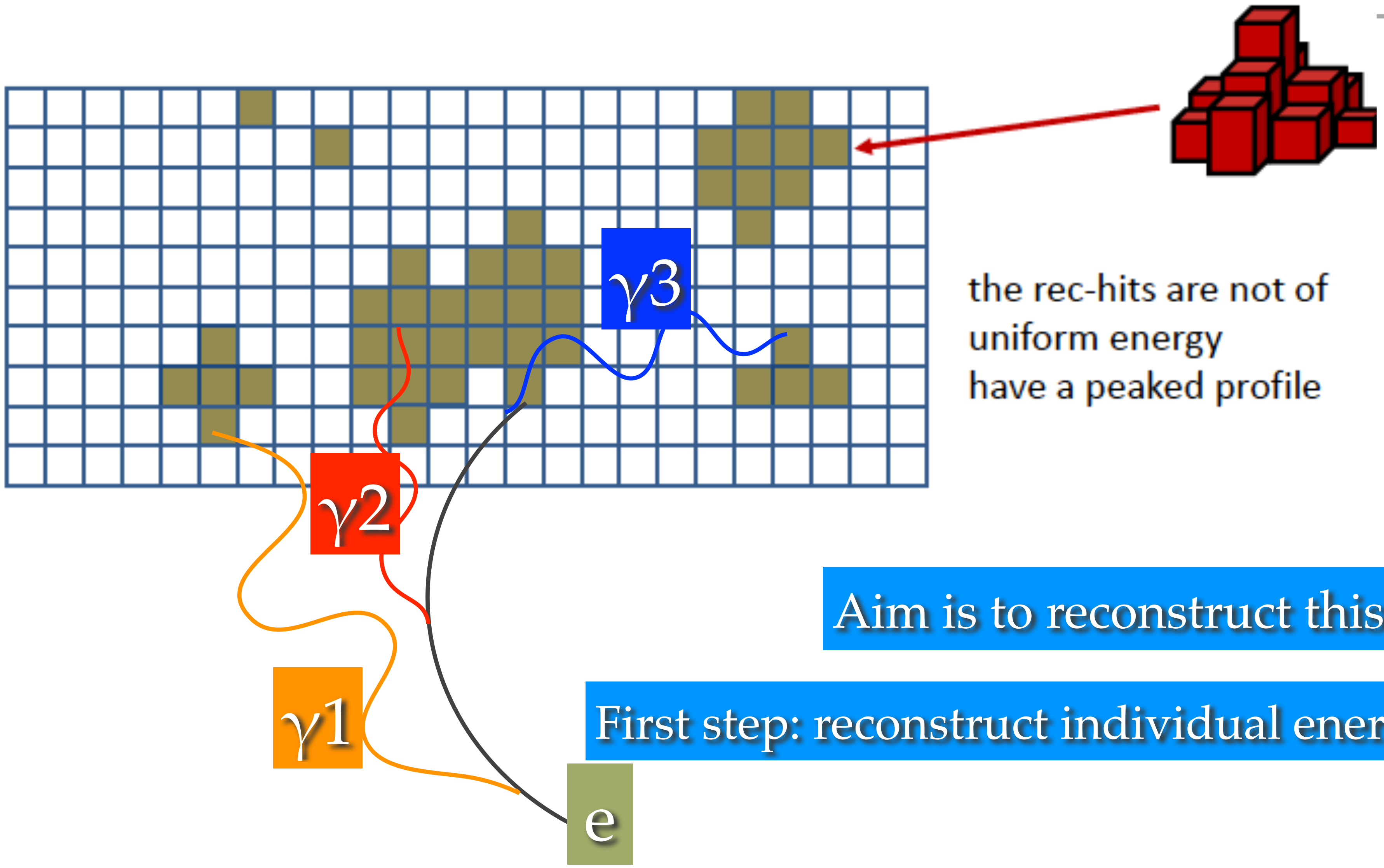


- To equilize the response, dedicated methods are performed to get the ICs:
  - phi symmetry method: Azimuthal symmetry of the energy distribution
  - pi0 and eta mesons: Exploits the peak position in the invariant mass distribution of pi0 or eta mesons
  - W/Z electrons: same as above but the mass is that of Z. In case of electrons coming from W, it is the E/p ratio which is used.
- Reference: https://arxiv.org/pdf/1306.2016.pdf



- Summary: transparency corrections essentially correct for the energy scale (and hence resolution). ICs correct for resolution mostly

ELECTRON AND PHOTON SHOWERS



the rec-hits are not of
uniform energy
have a peaked profile

Aim is to reconstruct this full energy

First step: reconstruct individual energy deposition

- An electron or a photon deposits energy in several crystals
- Aim: reconstruct the full energy of electron or a photon
- Once we have the crystal energies reconstructed, we need to reconstruct individual particles from deposit of a single photon/electron
  - i.e. in this picture reconstructing individual energy deposits from γ1, γ2, γ3 and the final electron
  - this is the PF (particle flow) clustering step - i.e. making small clusters
  - These small clusters are then combined to form super-clusters

TRADITIONAL CLUSTERING

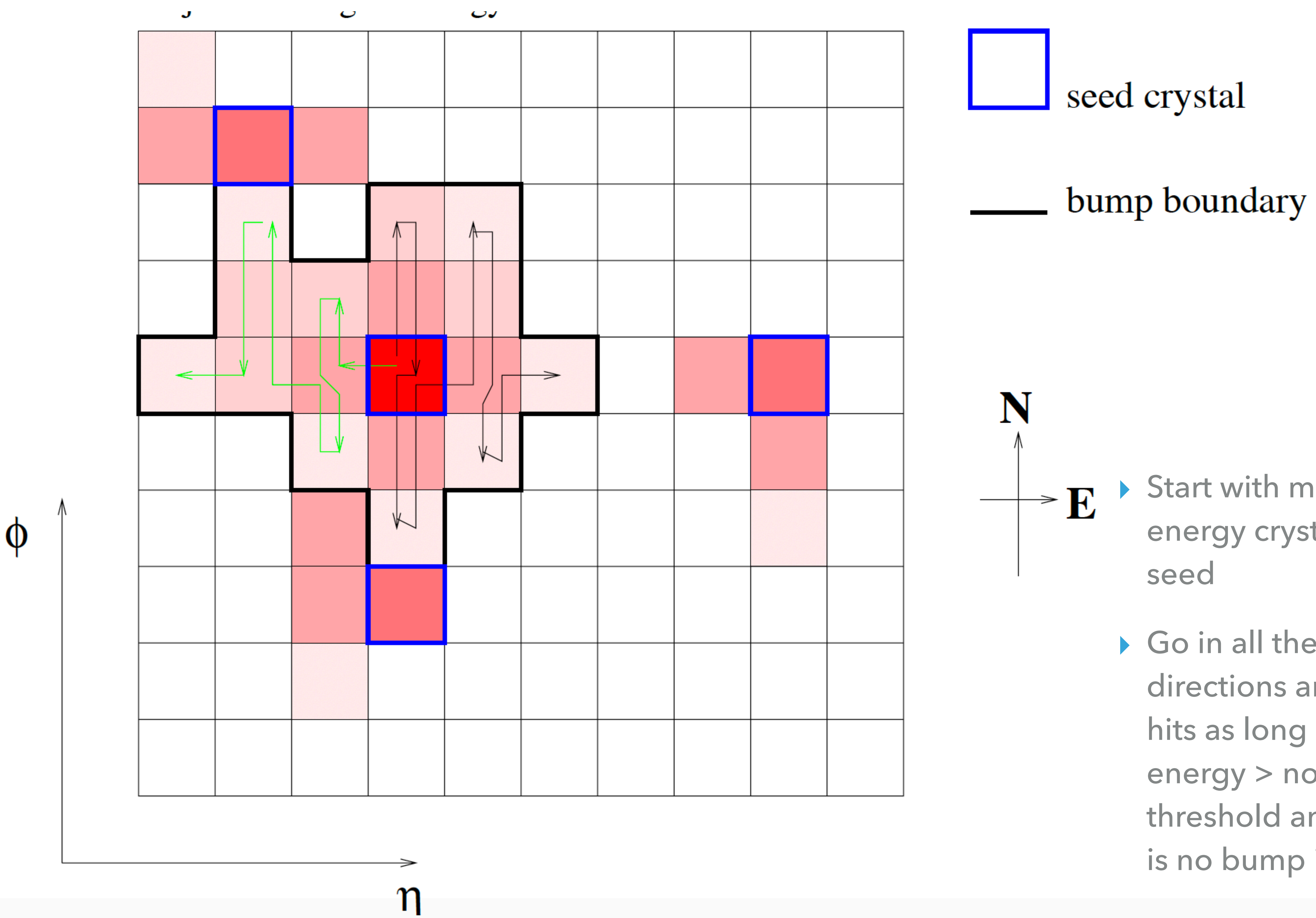
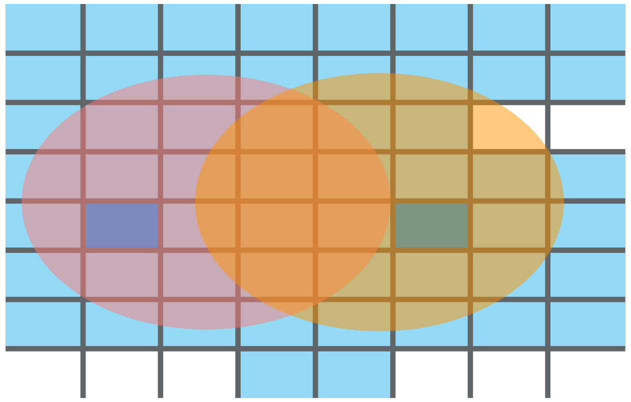


Fig. 1: Illustration of the Island clustering algorithm in the Barrel ECAL

- Start with maximum energy crystal –> seed
- Go in all the directions and collect hits as long as the hit energy > noise threshold and there is no bump in energy

2ND METHOD OF CLUSTERING: BUILD SMALL CLUSTERS

PF CLUSTERING I.E. RECHITS TO SMALL CLUSTERS



Aim: Reconstruct clusters of energy deposit from the individual particles

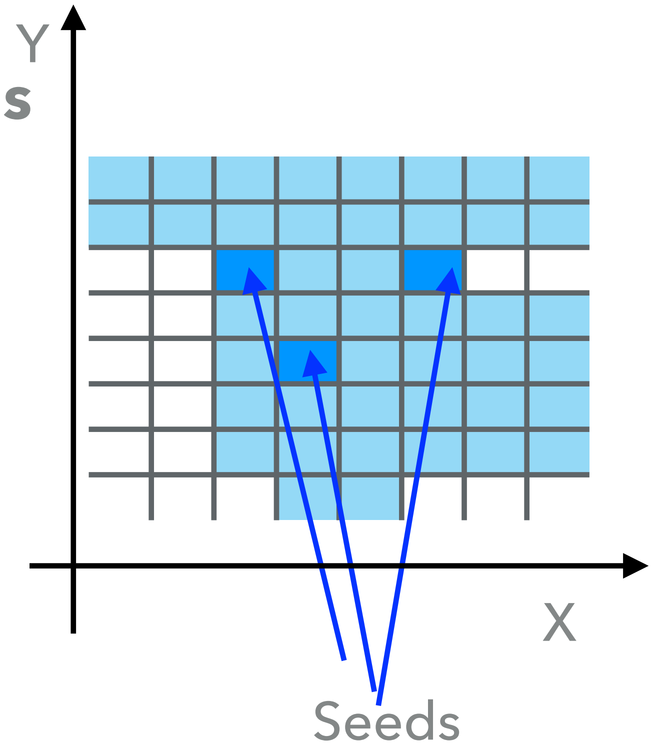
- A cluster can have energy contribution from N photons or an electron and a photon etc.

- It is important to separate the individual contributions.

- **Aim is to cluster the hits coming from the same particle and also determine the fraction of energy associated to a particle if a hit is shared**

- If there are M crystals in a cluster with N seeds, then these N seeds are coming N particles. The energy spread in other xtals from a particle is modeled using Gaussian, so N Gaussians
  - where seeds are crystals with local energy maxima above a threshold and are identified w.r.t the neighboring cells (8)

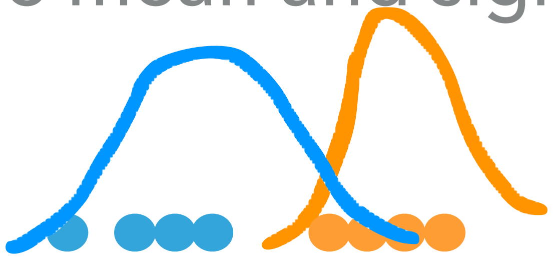
- For this purpose, an algorithm named **expectation maximization (EM) based on gaussian mixtures** is used



PROBLEM AT HAND

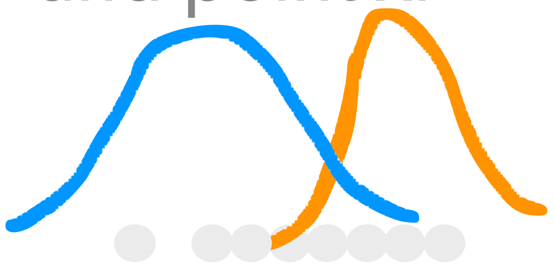
- Belongs to the class of clustering methods for ‘soft clusters’ in ML language (i.e. assigns probability)

- Problem 1: **Given a set of points from two gaussian**, we can evaluate the mean and sigma of the underlying Gaussians. μ = (x1 + …. + xn)/n



Straightforward

- Problem 2: **Given a set of points, and known parameters** of two gaussian, we can tell from which Gaussian those point most likely belong using Bayes’ theorem. Example if there are two gaussians ‘a’ and ‘b’ and point xi
  - P(a|xi) = P(xi|a) * P(a)/(P(xi|a) * P(a) + P(xi|b) * P(b))



PROBLEM 3

- Problem 3: Given a set of points, and no other information, we need to construct the gaussian distribution - **relevant to our case**
  - **In our case, we do not know the mean and sigma of the gaussian**
  - Use expectation-maximization algorithm
    - **In the first step, calculate posterior probability (P(a|xi)) of each point from each gaussian starting with some initial values of mean and sigma. In the next step, update the gaussian parameters using the knowledge from the first step**

Set of points/xtals

Determine:
What xtals belong together
and what the parameters
of the gaussians?

EXPECTATION-MAXIMIZATION ALGORITHM

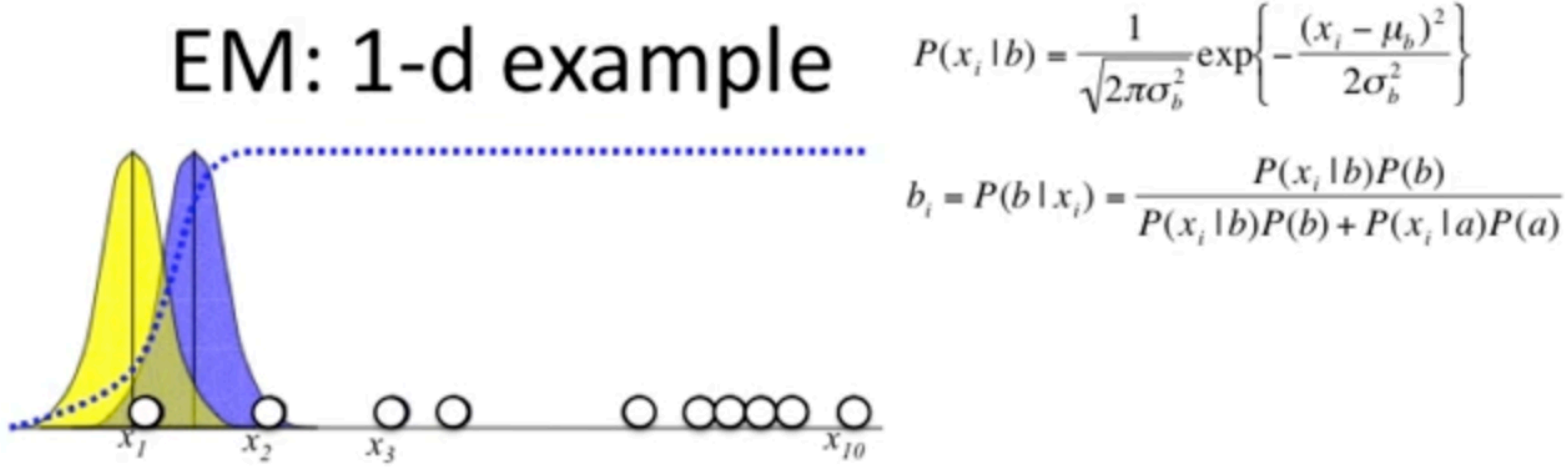
- Consider a total of M cells with N seeds
  - N particles –> N gaussians
- Aim essentially reduces to determining the parameters of the model are μ (space coordinate of the gaussian), and A (amplitude of that gaussian)
- **Expectation step**: Fraction of energy in a cell due to a gaussian is given by

$$\boxed{f_{ji} = \frac{A_i e^{-(\vec{c}_j - \vec{\mu}_i)^2/(2\sigma^2)}}{\sum_{k=1}^{N} A_k e^{-(\vec{c}_j - \vec{\mu}_k)^2/(2\sigma^2)}}.}$$

  - where j is the jth cell and i is the ith gaussian. μ is the space coordinate of the gaussian and c is the coordinate of the cell in question
- **Maximization step**: the parameters of the model (which are A and μ) are updated in the next step using the information from the first step

$$\boxed{\begin{aligned} A_i &= \sum_{j=1}^{M} f_{ji} E_j, \\ \vec{\mu}_i &= \sum_{j=1}^{M} f_{ji} E_j \vec{c}_j. \end{aligned}}$$

STEP I: EXPECTATION STEP



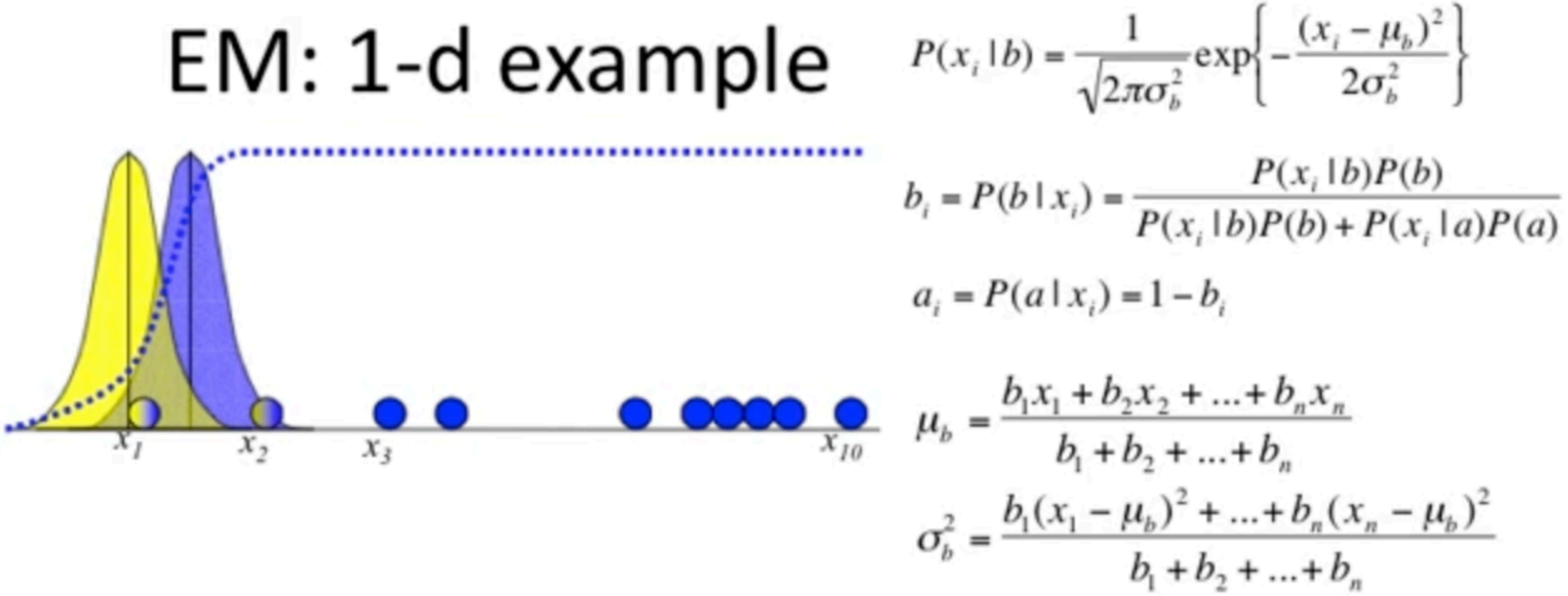
EM: 1-d example

$$P(x_i \mid b) = \frac{1}{\sqrt{2\pi\sigma_b^2}} \exp\left\{ -\frac{(x_i - \mu_b)^2}{2\sigma_b^2} \right\}$$

$$b_i = P(b \mid x_i) = \frac{P(x_i \mid b)P(b)}{P(x_i \mid b)P(b) + P(x_i \mid a)P(a)}$$

- Start with some initial values of means (μ) and sigma (σ) of two gaussians ‘a’ and ‘b’
    - Calculate the probability that given Gaussian ‘a’ with mean μ_a and σ_a, what is the probability of getting point x_i for Gaussian ‘a’
    - Same for Gaussian ‘b’

STEP 2: MAXIMIZATION STEP



EM: 1-d example

$$P(x_i \mid b) = \frac{1}{\sqrt{2\pi\sigma_b^2}} \exp\left\{-\frac{(x_i - \mu_b)^2}{2\sigma_b^2}\right\}$$

$$b_i = P(b \mid x_i) = \frac{P(x_i \mid b)P(b)}{P(x_i \mid b)P(b) + P(x_i \mid a)P(a)}$$

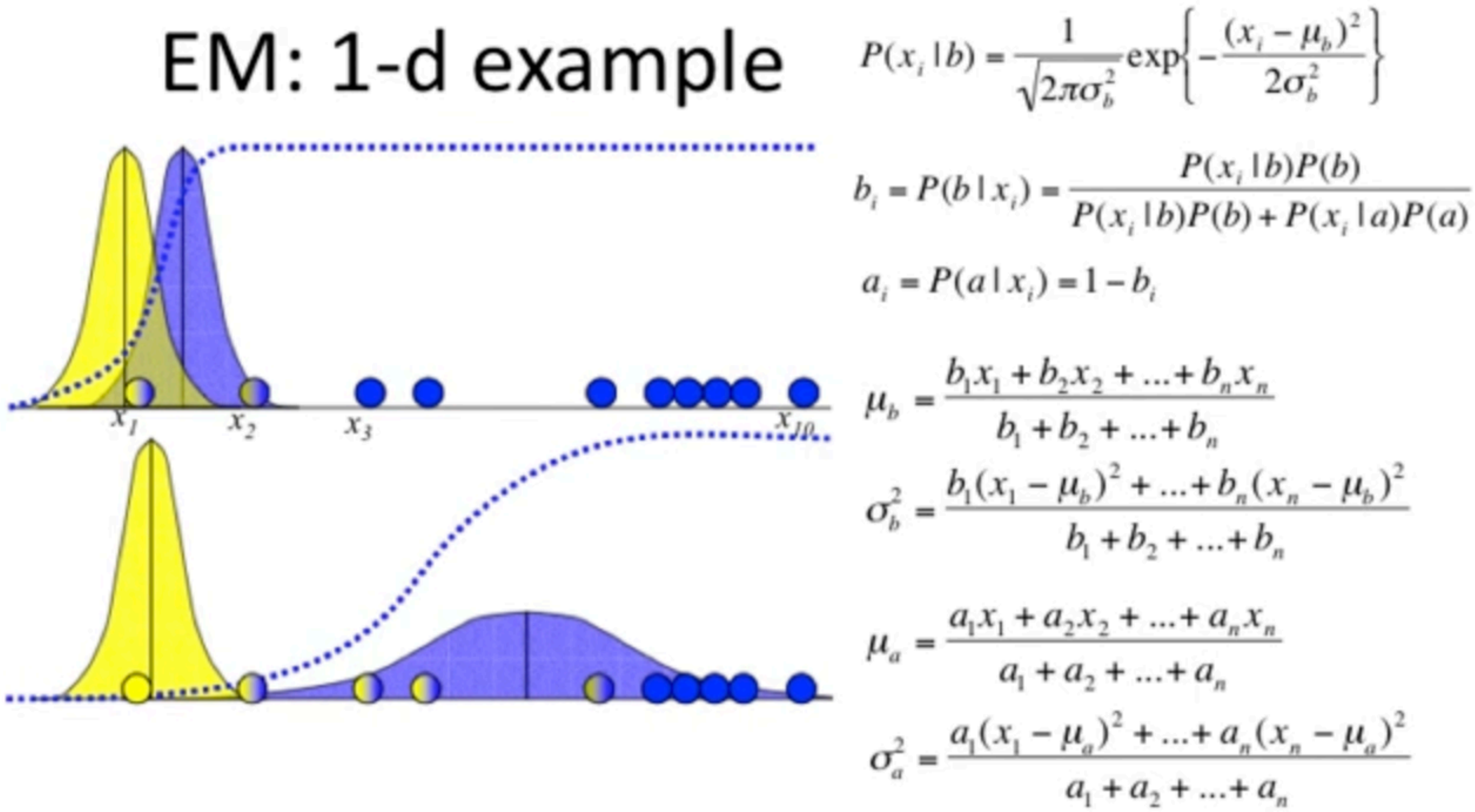
$$a_i = P(a \mid x_i) = 1 - b_i$$

$$\mu_b = \frac{b_1 x_1 + b_2 x_2 + \ldots + b_n x_n}{b_1 + b_2 + \ldots + b_n}$$

$$\sigma_b^2 = \frac{b_1(x_1 - \mu_b)^2 + \ldots + b_n(x_n - \mu_b)^2}{b_1 + b_2 + \ldots + b_n}$$

- For Gaussian ‘a’, calculate μ_a and σ_a again using these estimated probabilities by summing over all the hits
- Same for Gaussian ‘b’
- Now you have updated μ_a and σ_a for Gaussian ‘a’ and μ_b and σ_b for Gaussian ‘b’

STEP 3



EM: 1-d example

$$P(x_i \mid b) = \frac{1}{\sqrt{2\pi\sigma_b^2}} \exp\left\{ -\frac{(x_i - \mu_b)^2}{2\sigma_b^2} \right\}$$

$$b_i = P(b \mid x_i) = \frac{P(x_i \mid b)P(b)}{P(x_i \mid b)P(b) + P(x_i \mid a)P(a)}$$

$$a_i = P(a \mid x_i) = 1 - b_i$$

$$\mu_b = \frac{b_1 x_1 + b_2 x_2 + \ldots + b_n x_n}{b_1 + b_2 + \ldots + b_n}$$

$$\sigma_b^2 = \frac{b_1(x_1 - \mu_b)^2 + \ldots + b_n(x_n - \mu_b)^2}{b_1 + b_2 + \ldots + b_n}$$

$$\mu_a = \frac{a_1 x_1 + a_2 x_2 + \ldots + a_n x_n}{a_1 + a_2 + \ldots + a_n}$$

$$\sigma_a^2 = \frac{a_1(x_1 - \mu_a)^2 + \ldots + a_n(x_n - \mu_a)^2}{a_1 + a_2 + \ldots + a_n}$$

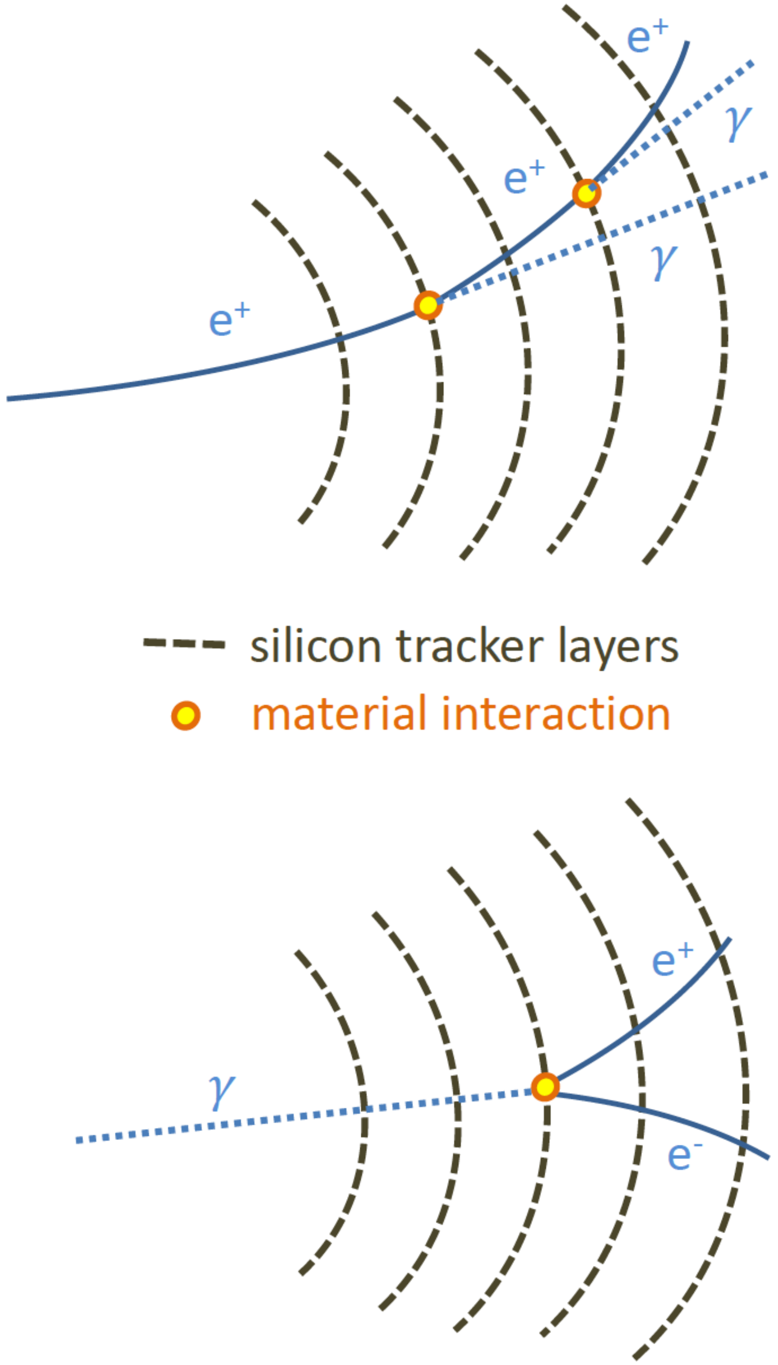
- With these updated μ_a and σ_a , go to step 1
- Repeat both the steps for let’s say 100 iterations

Illustration plotted using an adapted code from Satyaki Bhattacharya : here

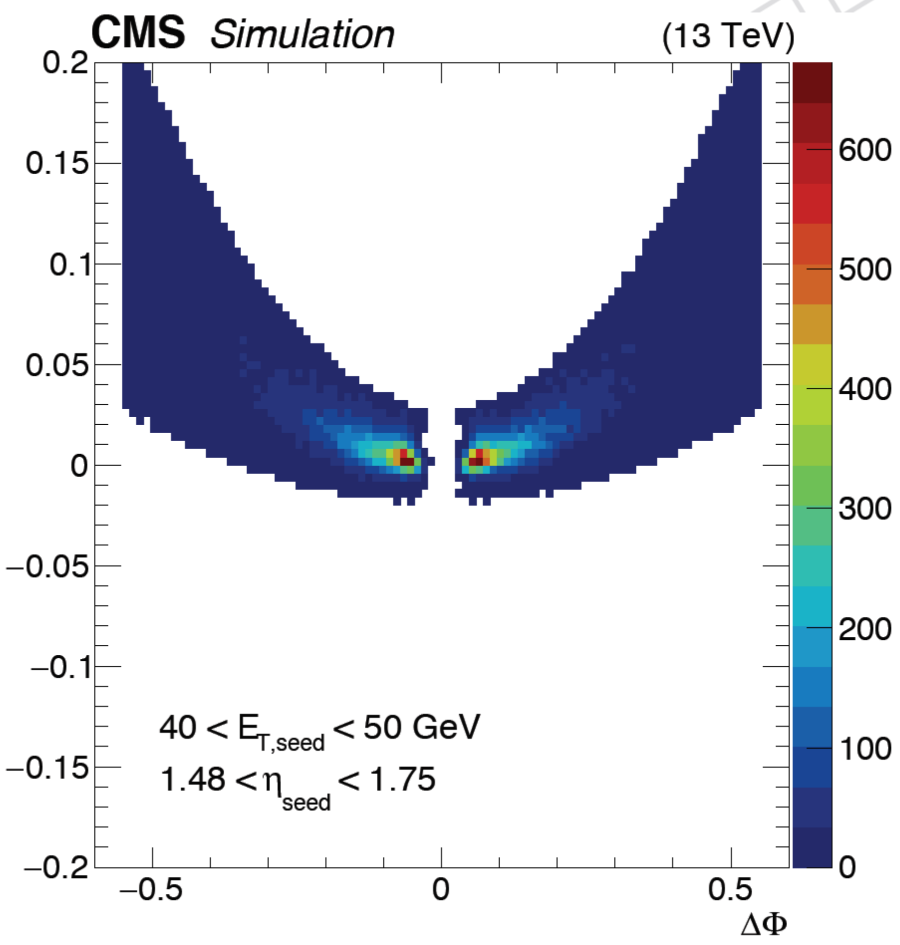
COLLECT SMALL CLUSTERS TO MAKE A BIG CLUSTER (‘SUPERCLUSTER’)

BASIC IDEA OF MUSTACHE SUPERCLUSTERING

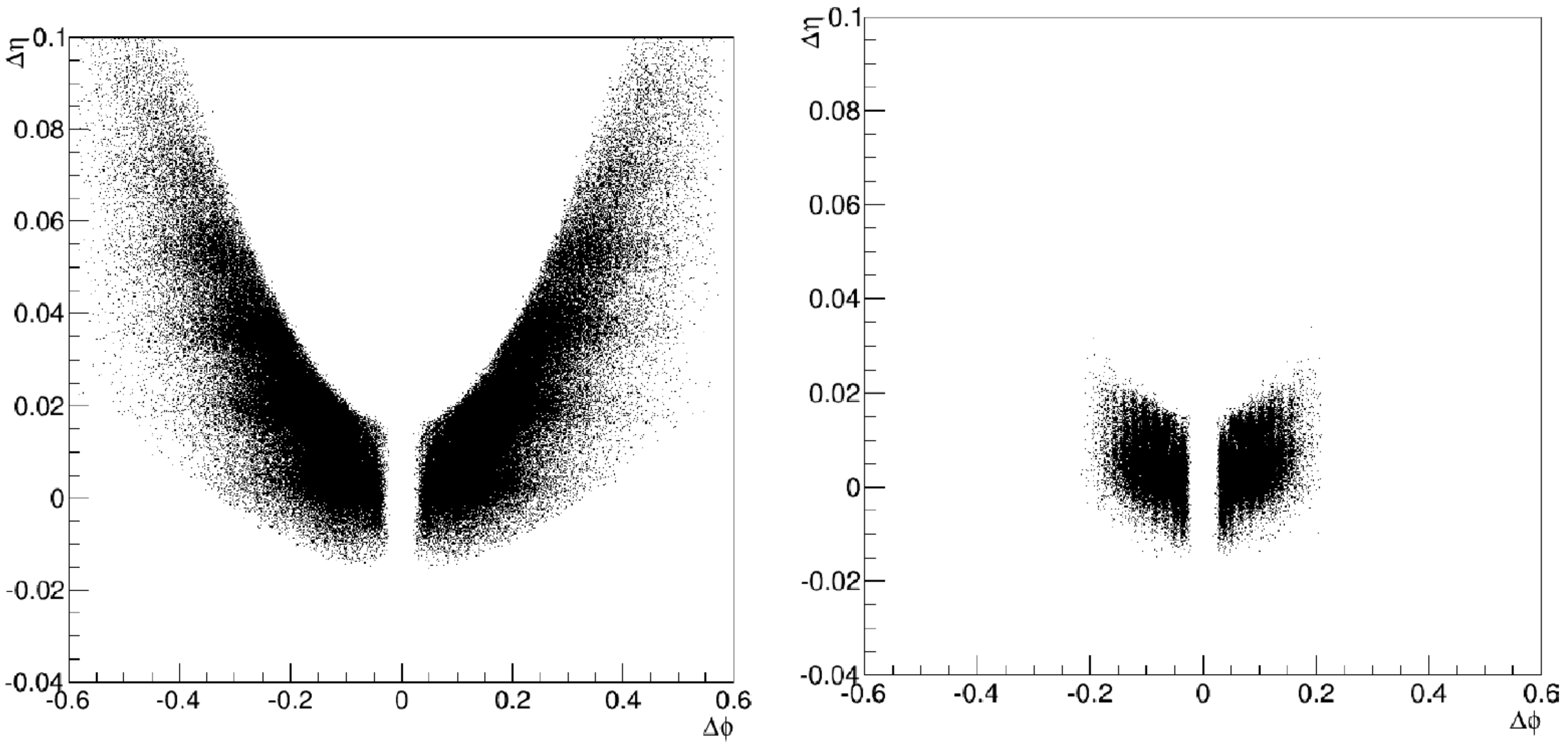
- EM showers usually have spread in η and (more in) φ because of the curved trajectory of electrons in the magnetic field
  - Bremsstrahlung from electron is tangential to the electron trajectory, whereas, due to the magnetic field in the Z direction, electron bends in phi
  - Similar happens for a photon which does a pair production



MUSTACHE SUPERCLUSTER



EGM-17-001



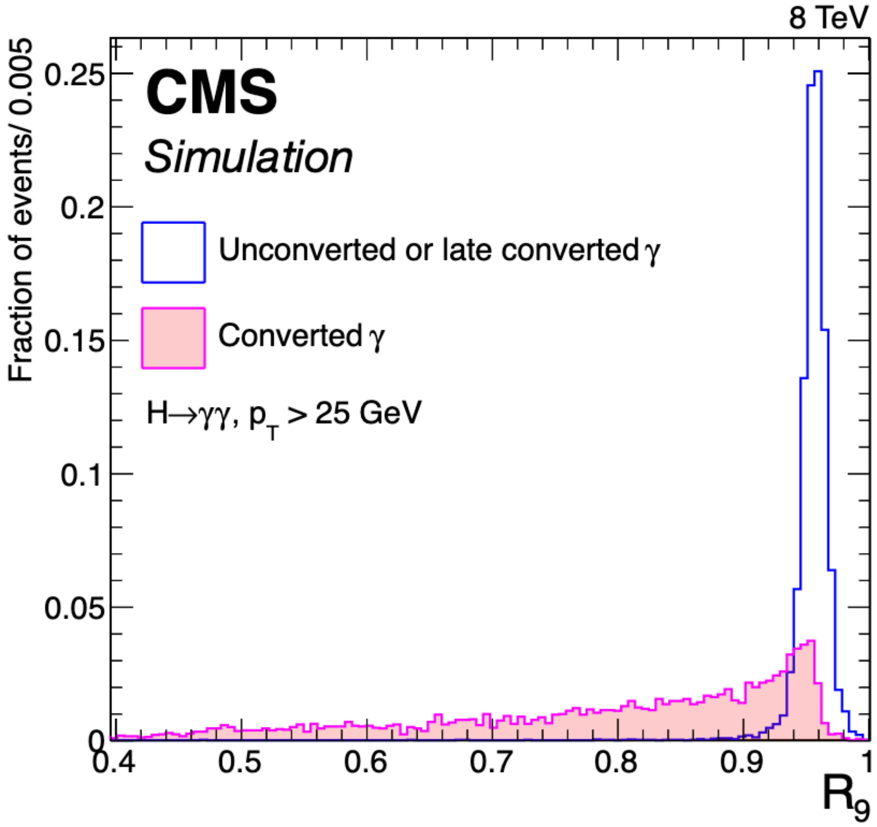
(a) $0.5 < E_T(\text{subcluster}) < 1.0$ GeV

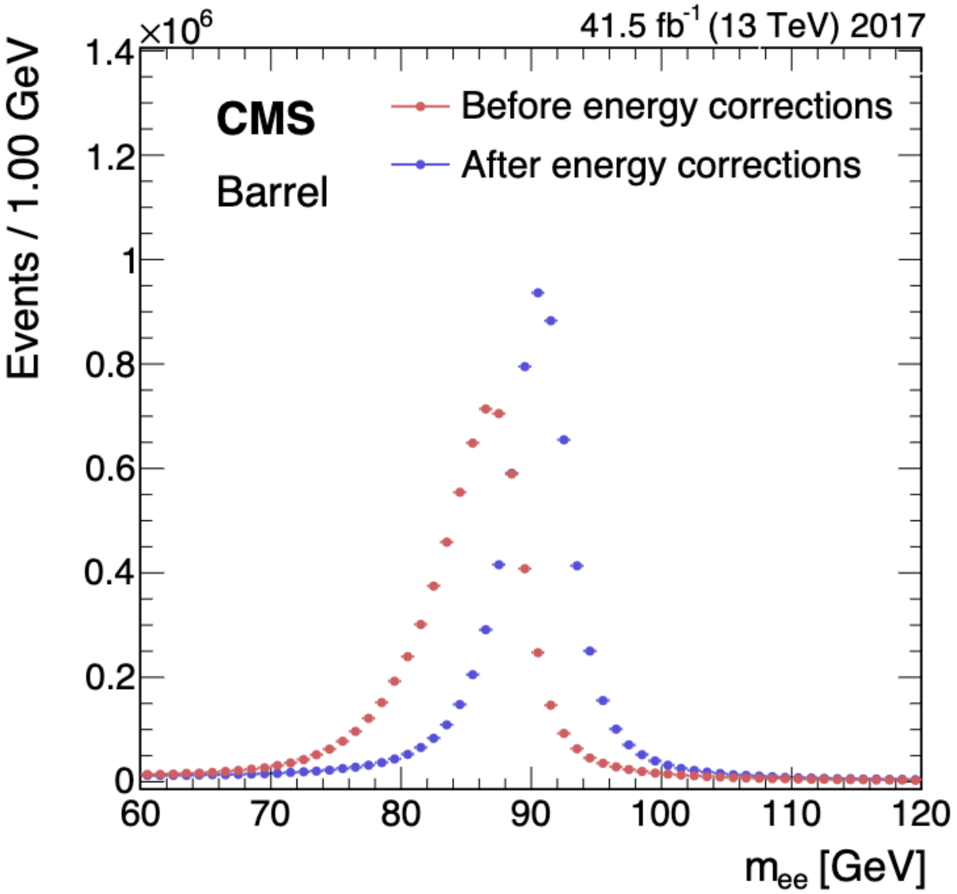
(b) $4 < E_T(\text{subcluster}) < 6$ GeV

ENERGY CALIBRATION

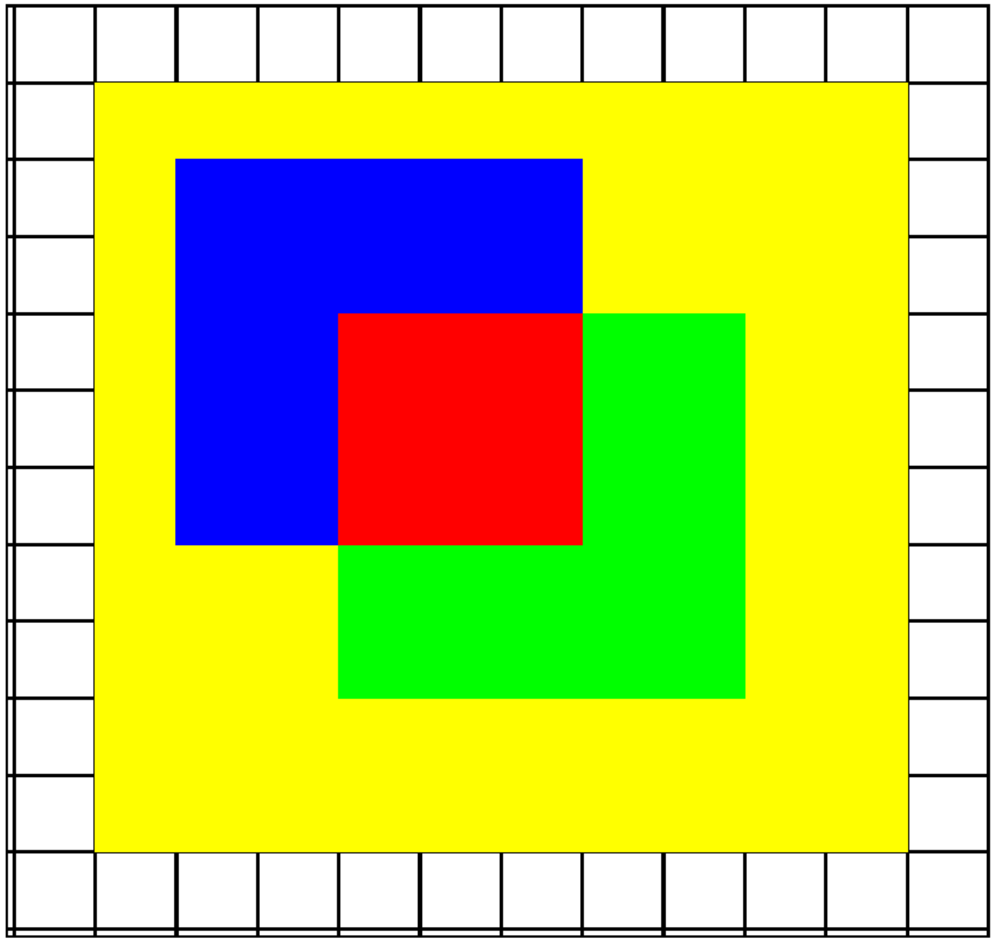
- Energy is lost in the tracker, support structure, gaps, can also remain unclustered
- Important to correct for it.
- Corrected by using photon variables e.g. shower shapes, information in HCAL, related information in the tracker to determine the energy loss and hence correct for it

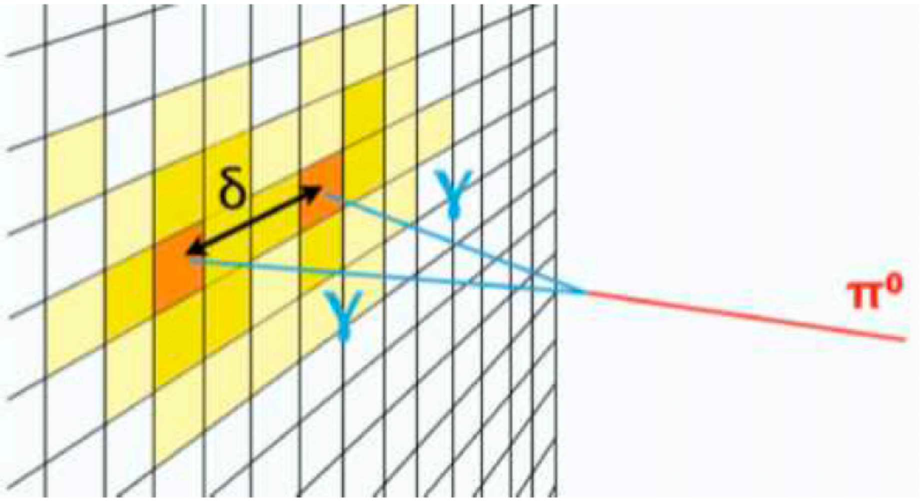
Variables such as: R9 = E3x3/Etotal



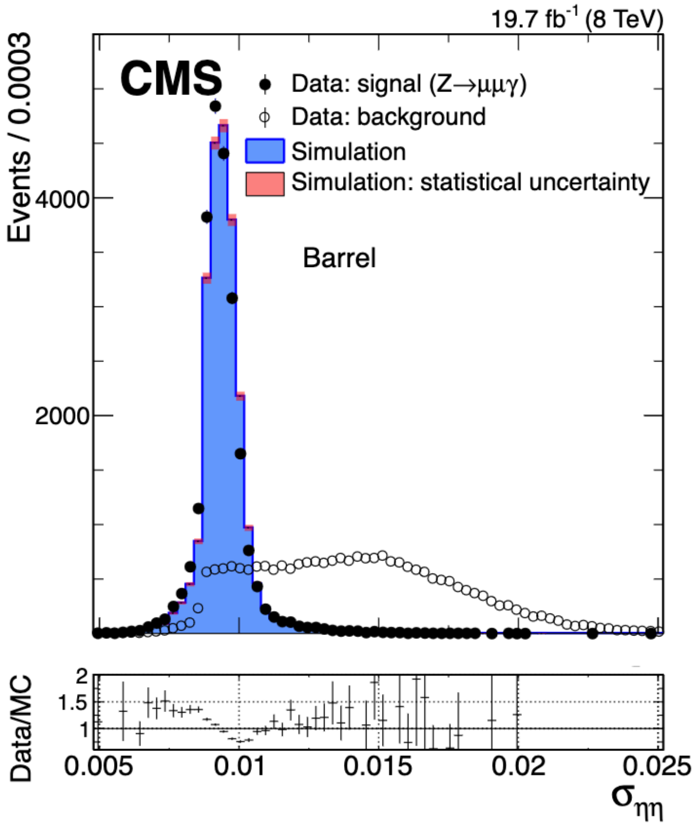


DISCRIMINATION BETWEEN PIONS AND SINGLE PHOTONS



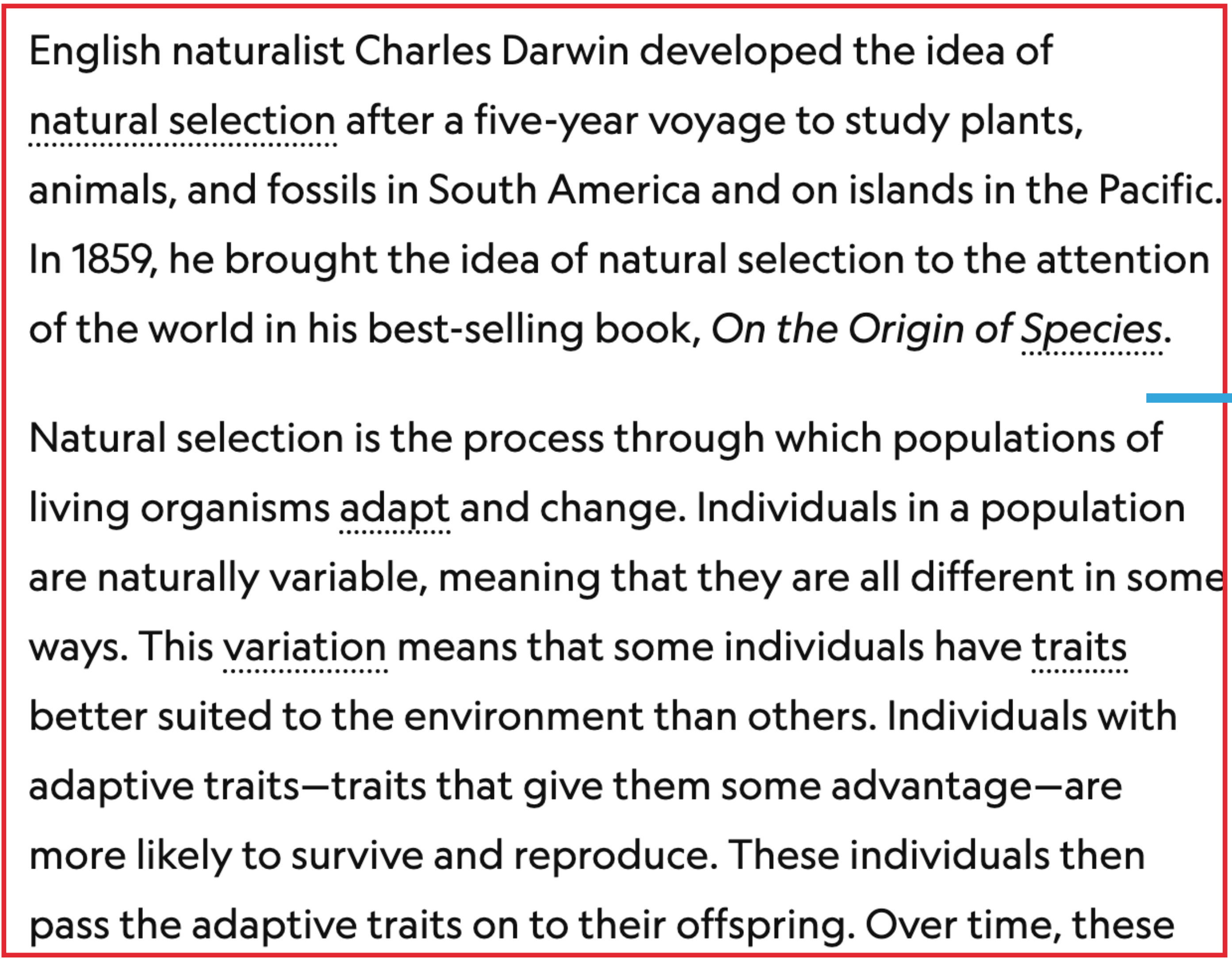


- Blue/Green : Individual photon
- Blue+Green : $\pi^0 \rightarrow \gamma\gamma$
- Yellow : Hadronic shower



CUT BASED USING GENETIC ALGORITHMS

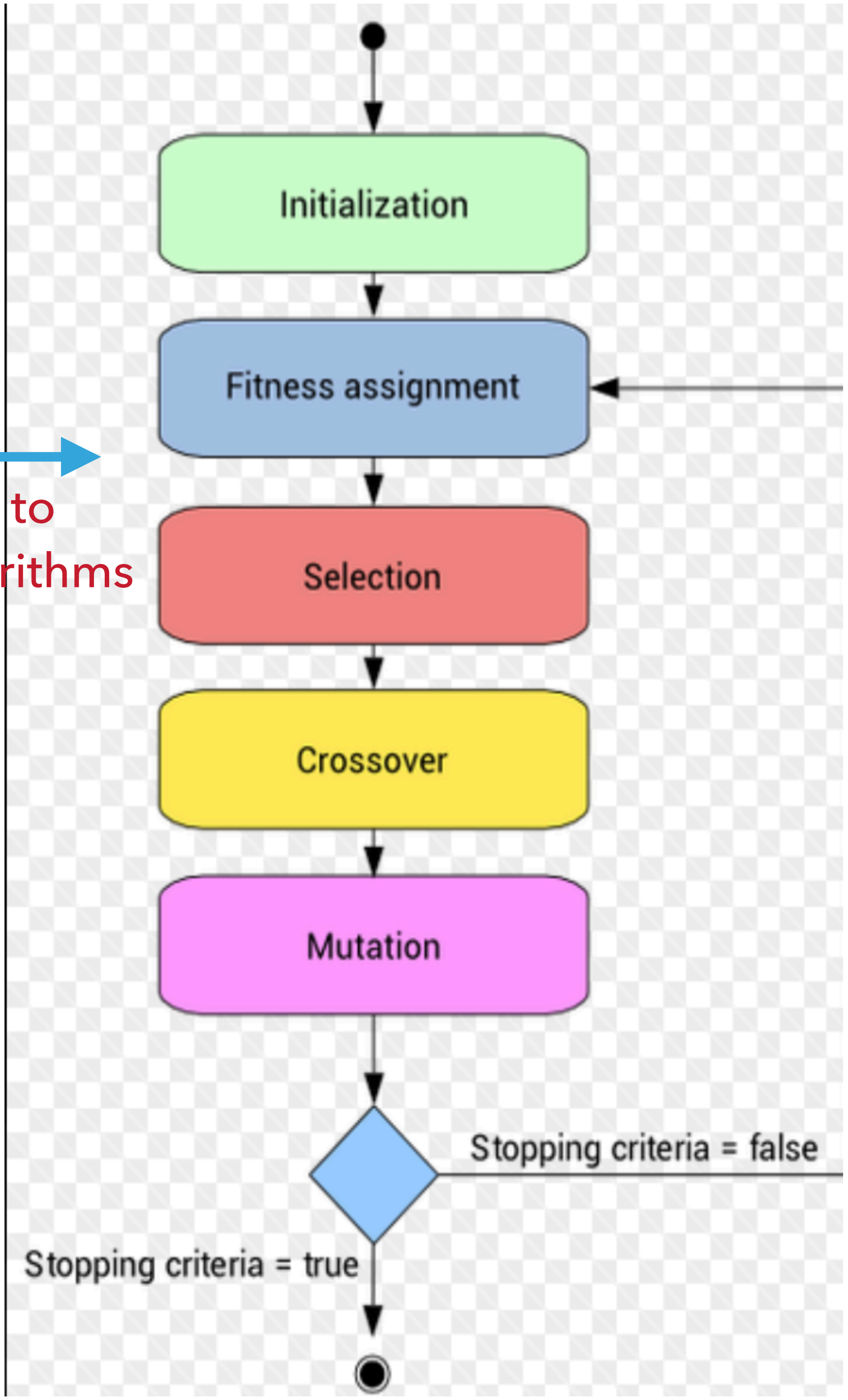
Explained in the context of photon ID



English naturalist Charles Darwin developed the idea of natural selection after a five-year voyage to study plants, animals, and fossils in South America and on islands in the Pacific. In 1859, he brought the idea of natural selection to the attention of the world in his best-selling book, *On the Origin of Species*.

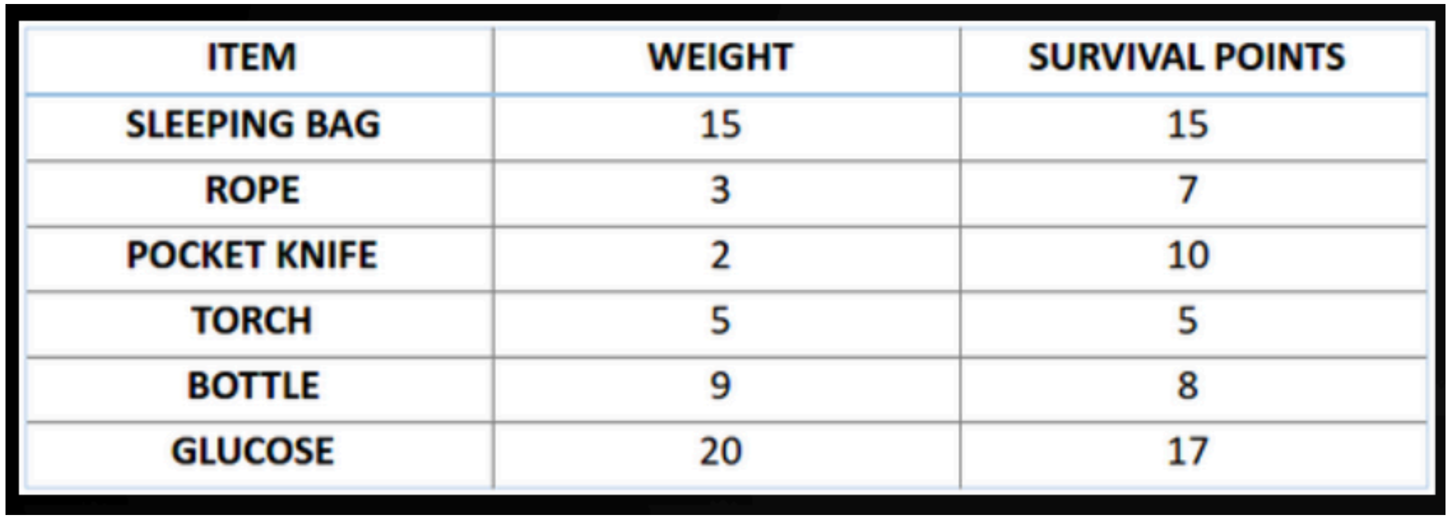
Natural selection is the process through which populations of living organisms adapt and change. Individuals in a population are naturally variable, meaning that they are all different in some ways. This variation means that some individuals have traits better suited to the environment than others. Individuals with adaptive traits—traits that give them some advantage—are more likely to survive and reproduce. These individuals then pass the adaptive traits on to their offspring. Over time, these

Application to search algorithms



- Based on the ideas of natural selection and genetics
- These algorithms simulate the process of natural selection
- It means that those species who can adapt to changes in their environment are able to survive and reproduce and go to the next generation

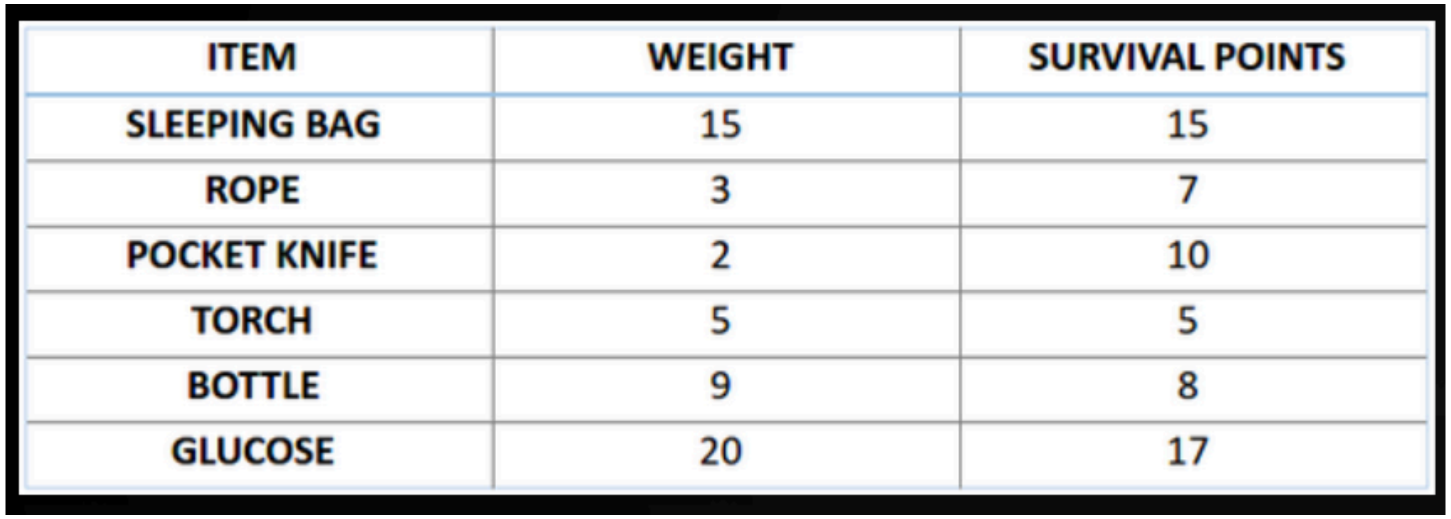
GENETIC ALGORITHMS: INITIALIZATION



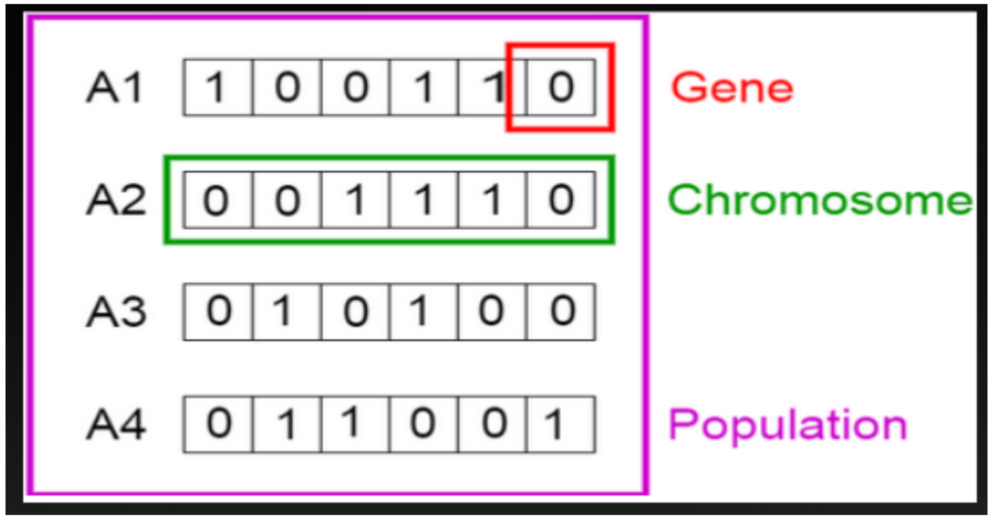
ITEM	WEIGHT	SURVIVAL POINTS
SLEEPING BAG	15	15
ROPE	3	7
POCKET KNIFE	2	10
TORCH	5	5
BOTTLE	9	8
GLUCOSE	20	17

- It can be understood by taking above example
- If you are allowed a maximum weight of 30 kg, how would you maximize your survival points

INITIALIZATION



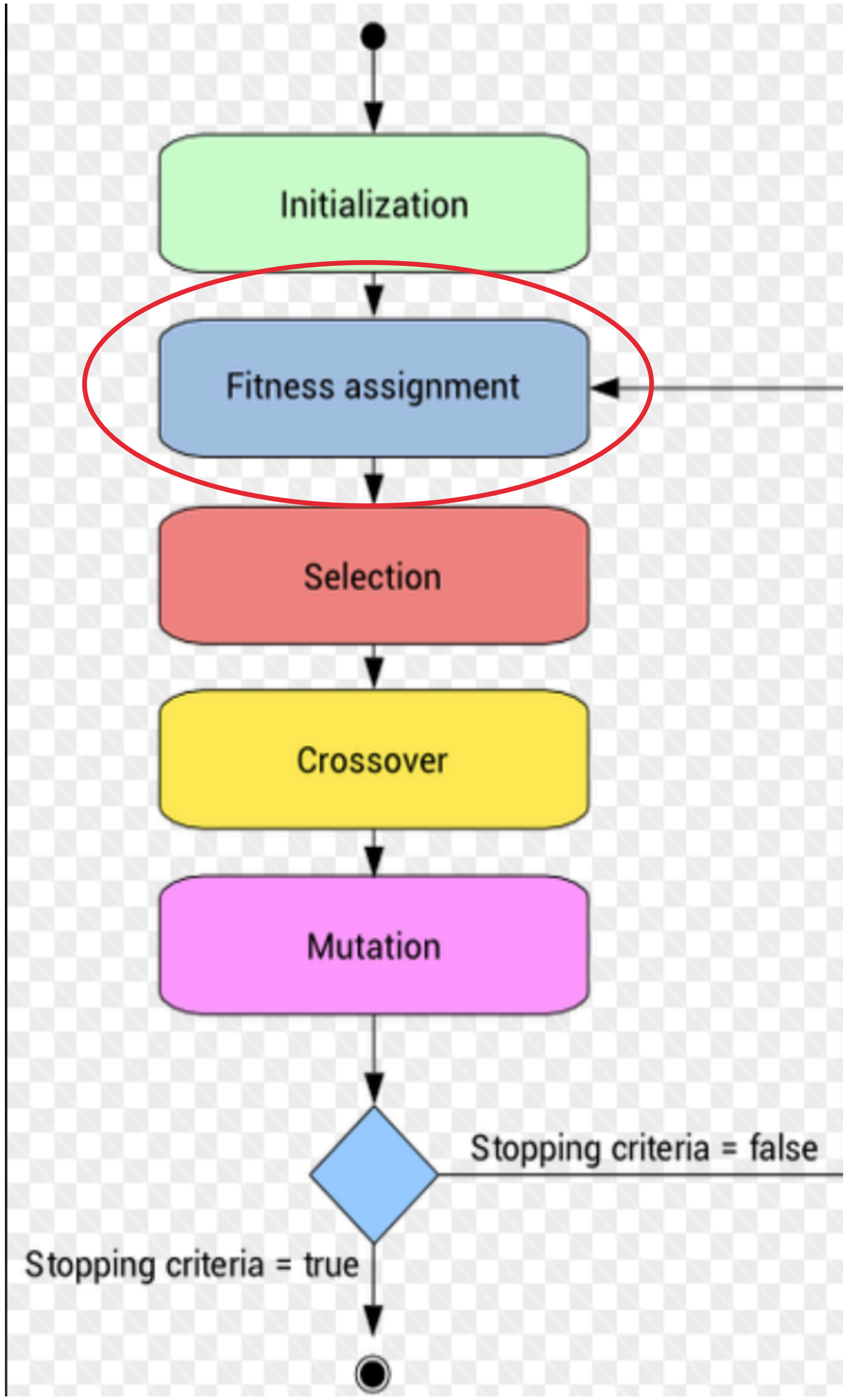
ITEM	WEIGHT	SURVIVAL POINTS
SLEEPING BAG	15	15
ROPE	3	7
POCKET KNIFE	2	10
TORCH	5	5
BOTTLE	9	8
GLUCOSE	20	17



- Number of ways in which selection can be made: 2^n with a constraint on the weight, where n = 5 items
- Relating it with the case of the cuts on variables in photon ID:
  - Gene <–> variable e.g. shower shape or isolation
  - Chromosome <–> photon object
  - Population <–> whole collection of photons
- Consider the sample as initial population

CUT BASED USING GENETIC ALGORITHMS

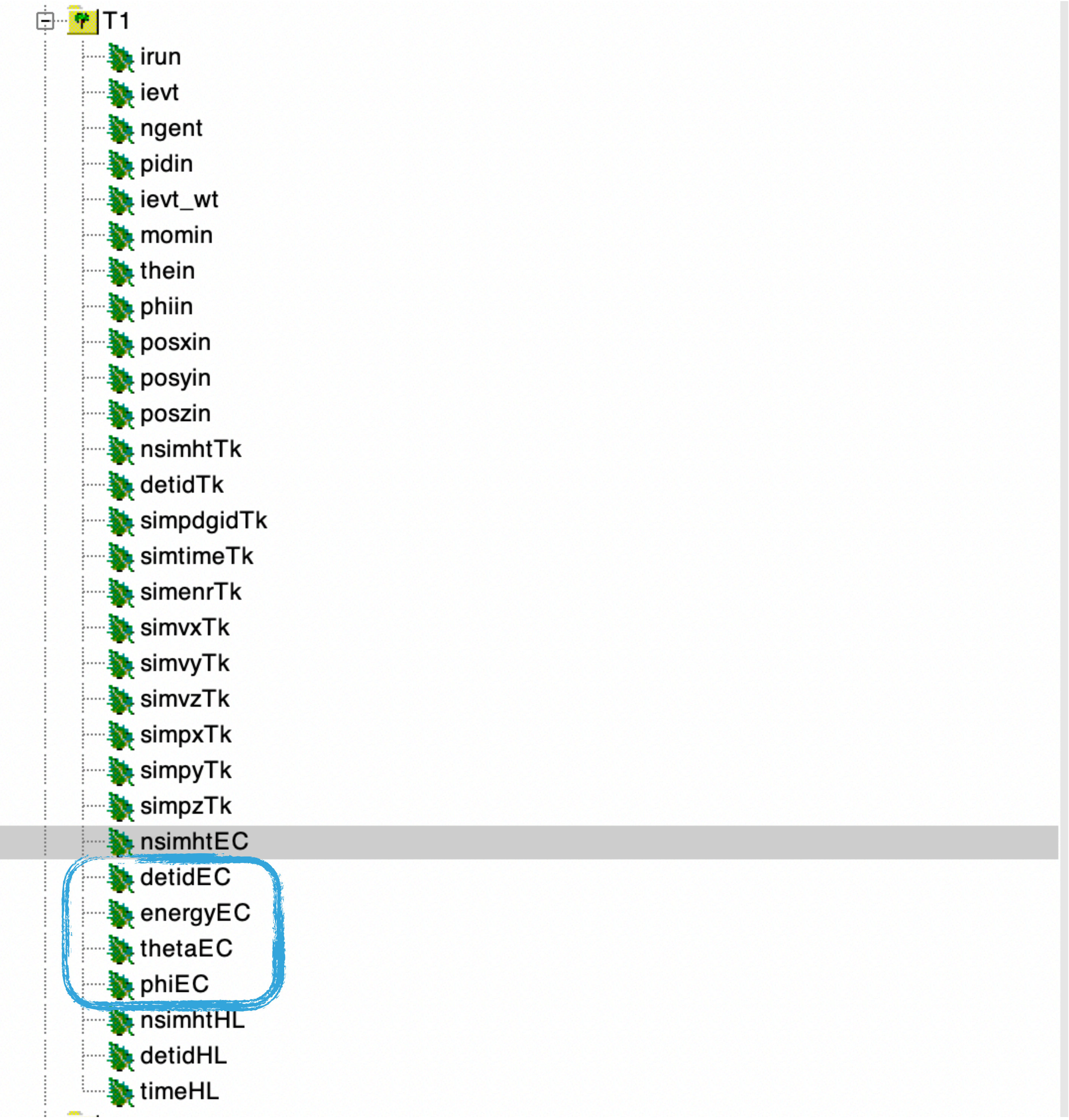
Explained in the context of photon ID



- In this case, you want the population which has the best background rejection for a given signal efficiency is the best
- Based on the fitness score, select the parents for cross-over

EXERCISES

ROOTFILE

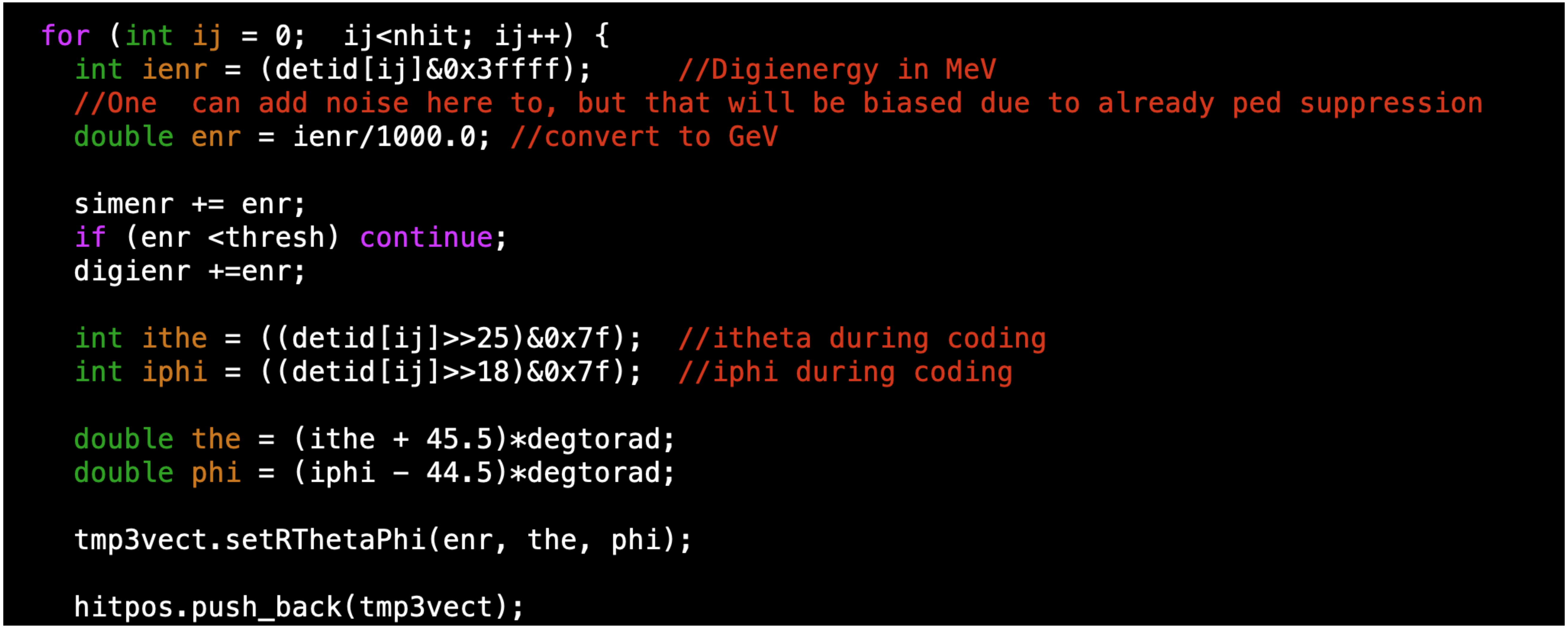


```
T1
  irun
  ievt
  ngent
  pidin
  ievt_wt
  momin
  thein
  phiin
  posxin
  posyin
  poszin
  nsimhtTk
  detidTk
  simpdgidTk
  simtimeTk
  simenrTk
  simvxTk
  simvyTk
  simvzTk
  simpxTk
  simpyTk
  simpzTk
  nsimhtEC
  detidEC
  energyEC
  thetaEC
  phiEC
  nsimhtHL
  detidHL
  timeHL
```

- Code repository:
  - git clone https://github.com/jainshilpi/EHEP2024.git
  - Ex 1 and 2: serc19_ecal_clustering.C
  - Ex 3(a): myEMalgo.C
  - Ex 3(b) and remaining (time permitted): serc19_ecal_clustering.C
  - Use Makefile for compilation

ROOTFILE

Already in the file



```
 for (int ij = 0;  ij<nhit; ij++) {
   int ienr = (detid[ij]&0x3ffff);     //Digienergy in MeV
   //One  can add noise here to, but that will be biased due to already ped suppression
   double enr = ienr/1000.0; //convert to GeV

   simenr += enr;
   if (enr <thresh) continue;
   digienr +=enr;

   int ithe = ((detid[ij]>>25)&0x7f);  //itheta during coding
   int iphi = ((detid[ij]>>18)&0x7f);  //iphi during coding

   double the = (ithe + 45.5)*degtorad;
   double phi = (iphi - 44.5)*degtorad;

   tmp3vect.setRThetaPhi(enr, the, phi);

   hitpos.push_back(tmp3vect);
```

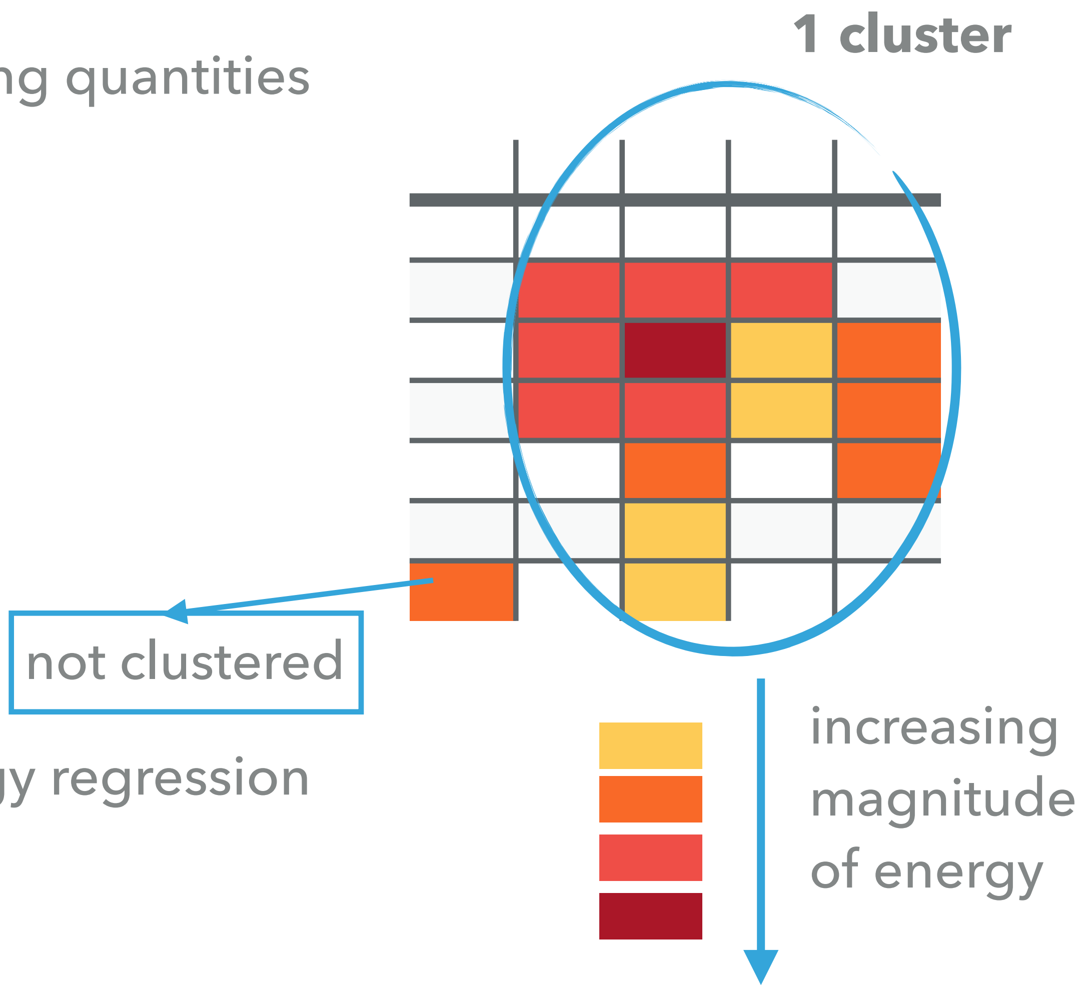
- How to use the above information to do the exercises?

```
for (int ij=0; ij<digienr.size(); ij++) {
double energyhit = digienr[ij].mag();
double cell_theta = digienr[ij].theta()*radtodeg;
double cell_phi    = digienr[ij].phi()*radtodeg;
}
```

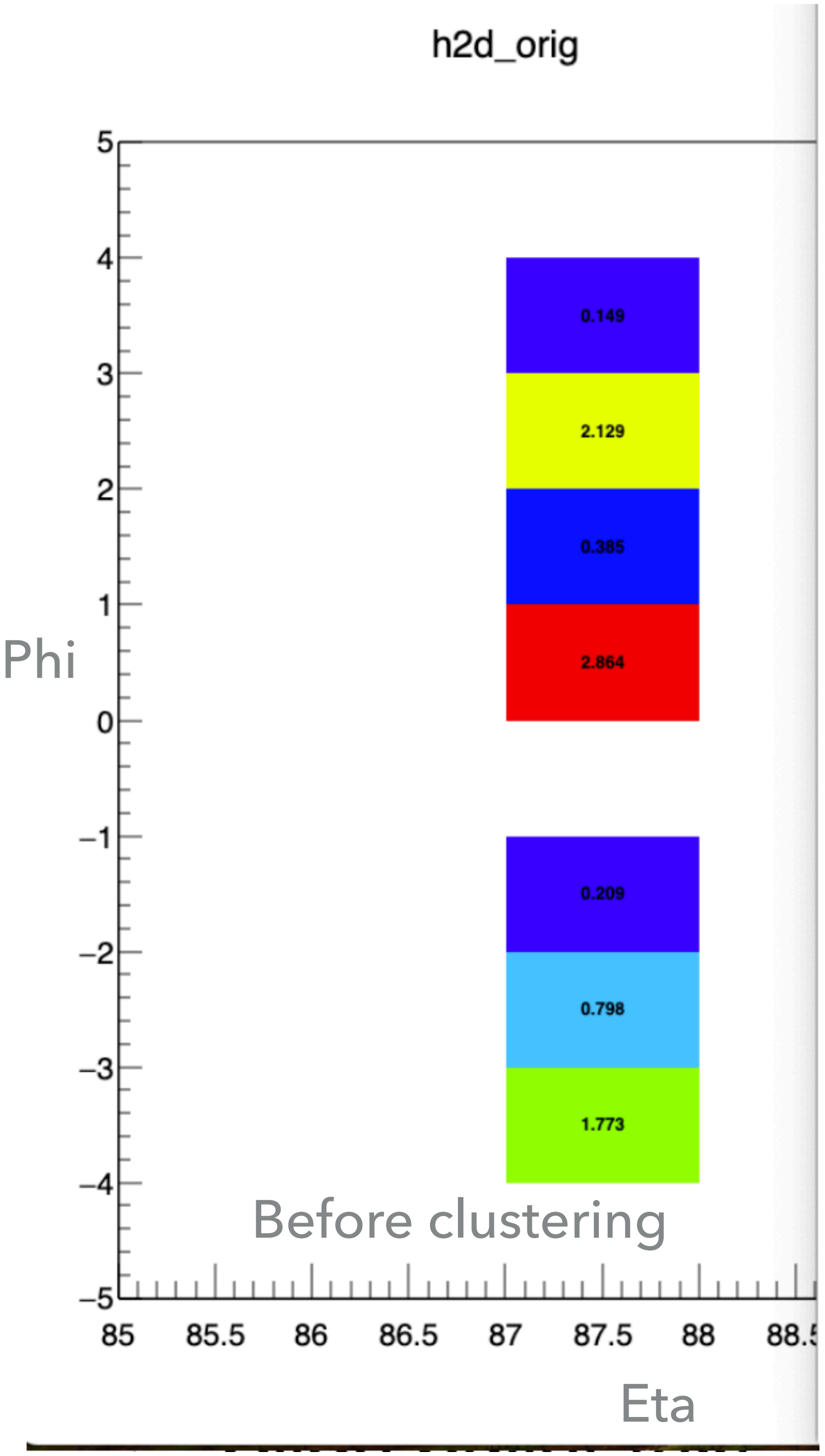
```
To calculate angle between two crystals:
double angle = cluster[jk].angle(digienr[kl].unit());
```

EXERCISE 1

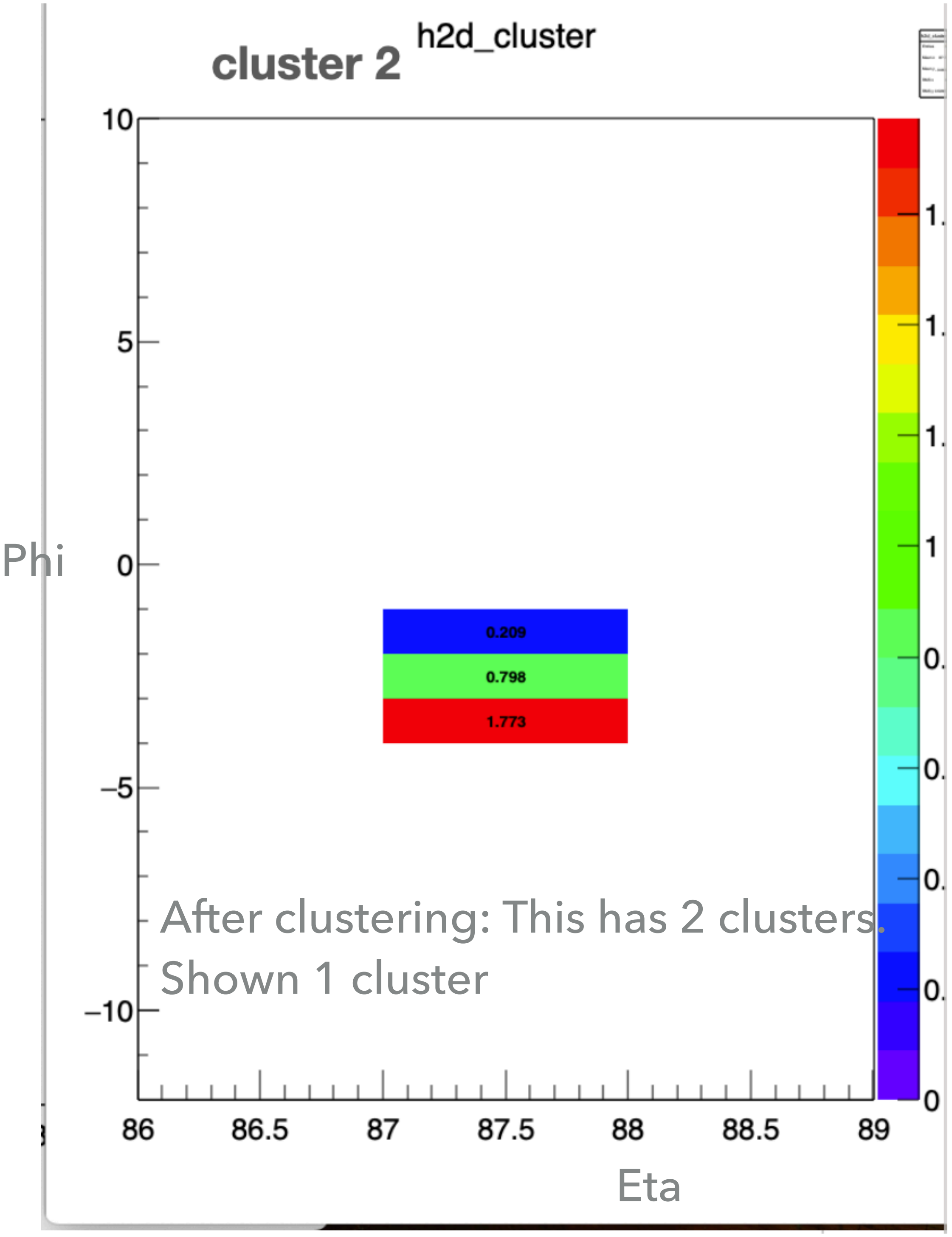
- Cluster the hits from a photon using basic clustering algorithm:
  - Make a collection of crystals passing seed threshold criteria. Take seed threshold energy = 0.5 GeV
  - Start with this collection, and collect crystals for each seed which are within 3x3 vicinity around it (passing a noise threshold of 0.04 GeV to start with) or in the vicinity of already clustered crystal.
    - Remove the hits which are being clustered from getting clustered again
- Associate each cluster with a photon and form the following quantities
  - E9, E25
  - Form moments:
    - dTheta = Σ(theta_hit - theta_seed)*energy_hit
    - Similar for dPhi and dPhidTheta
    - Diagonalize the matrix
  - Above quantities will be used for photon ID and energy regression
  - Save the photon



TESTS TO DO



Before clustering

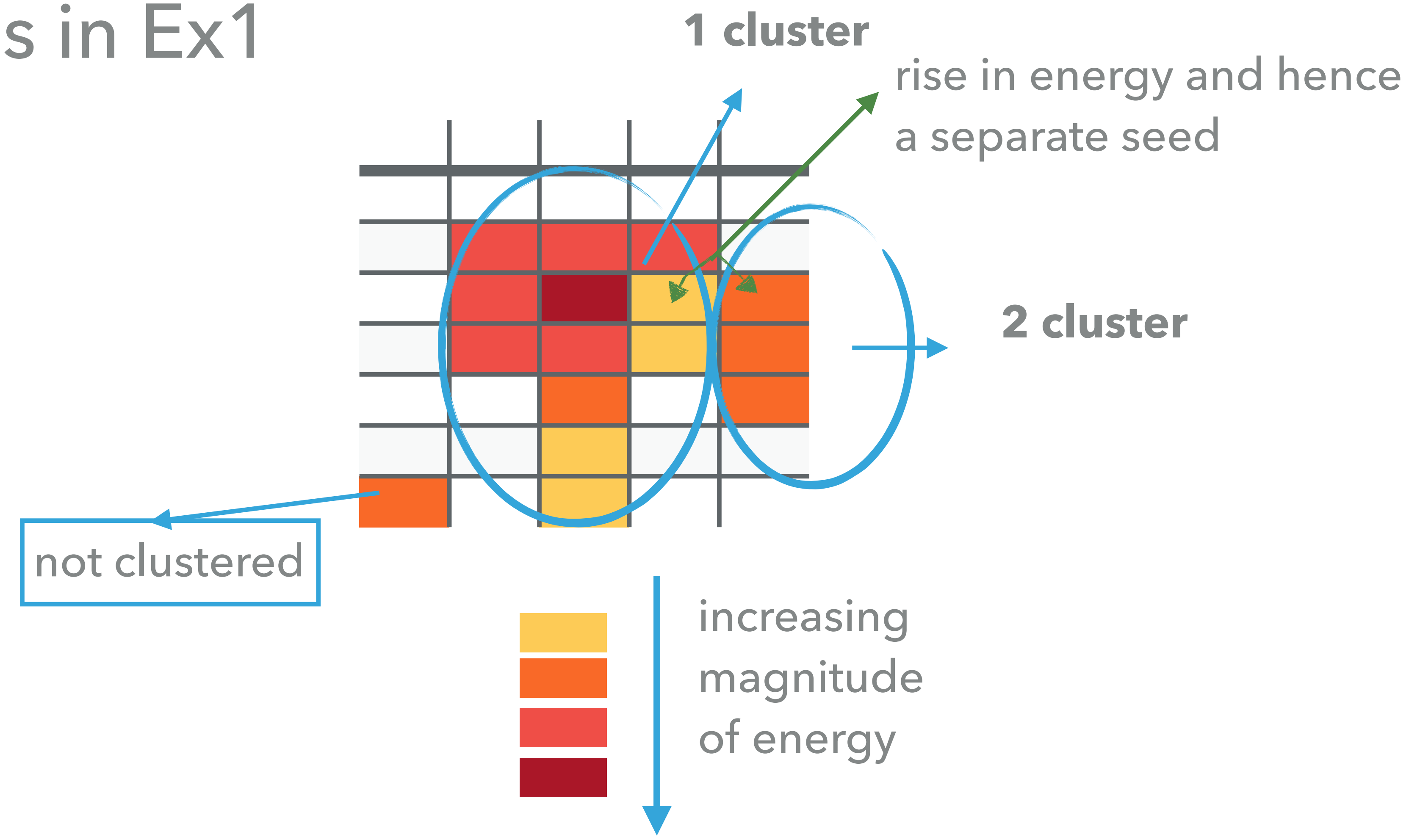


After clustering: This has 2 clusters.
Shown 1 cluster

- Plot the 2D histogram of theta VS phi of original hits
- Plot the 2D histogram of theta VS phi of the clustered hits

EXERCISE 2

- Add the condition in exercise 1 that while clustering there is no rise in energy as we go in one direction.
- If there is a rise, then take that hit as another seed
- Again compare as in Ex1


1 cluster
rise in energy and hence a separate seed
2 cluster
not clustered
increasing magnitude of energy

EXERCISE 3(A)

- Use EM algorithm
  - Start with 1-D pseudo data
  - You will see 3 gaussians with known mean and sigma are filled in a histogram.
  - Now consider each bin of the histogram as a crystal and apply EM algorithm
  - Check if you can get similar means and sigma of the gaussian as you started with

EXERCISE 3(B)

- Now that you have learnt how to use EM algorithm in1-D, use EM algorithm on the hits and extend to 2D EM algorithm
- Compare this with the algorithm in exercise 2 by forming π^0 mass

EXERCISE 2

- Use the quantities thus formed to determine the noise thresholds

EXERCISE 4

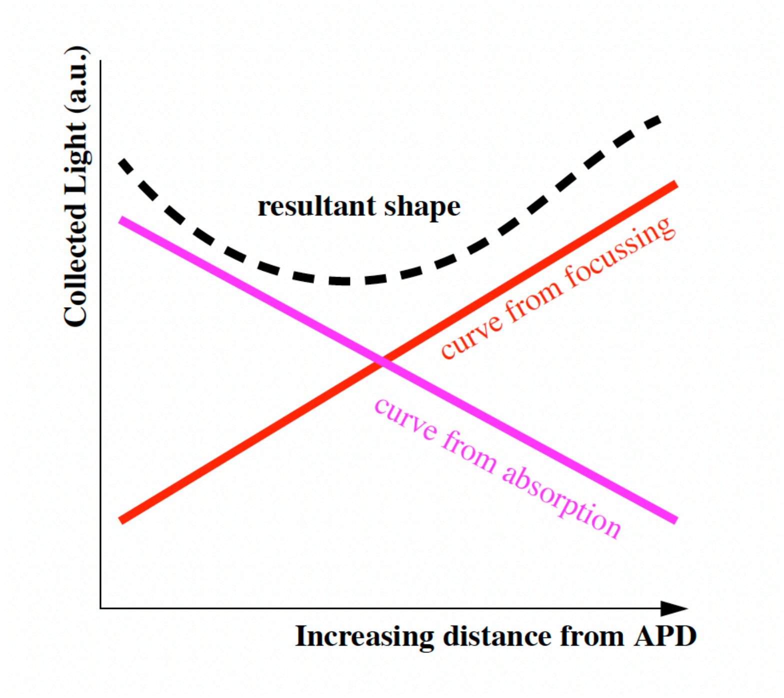
- Energy regression using the photon variables

EXERCISE 5

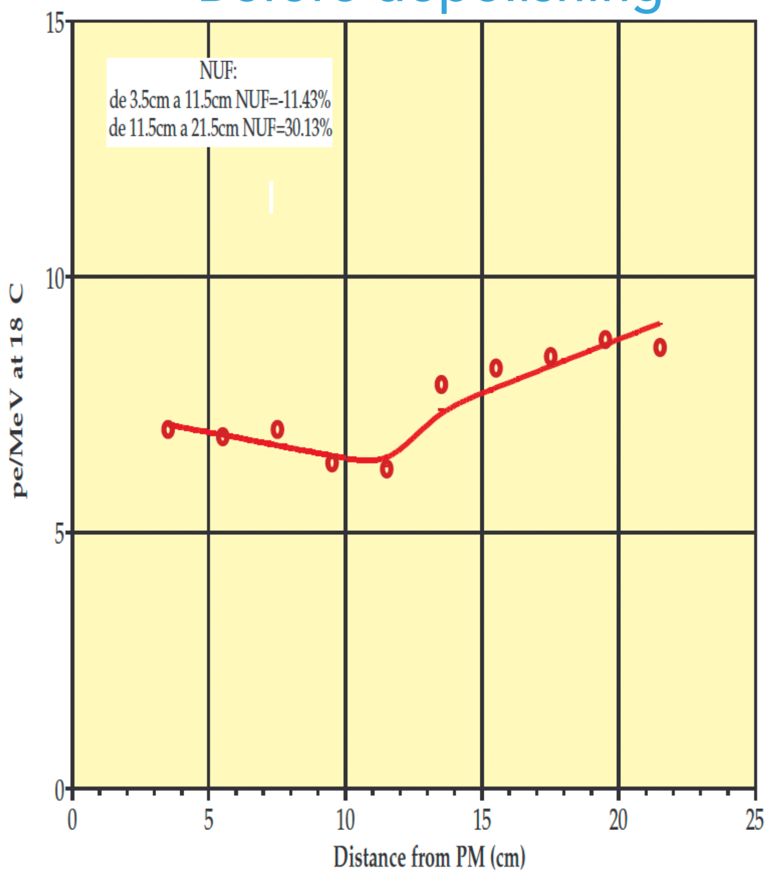
- Discrimination between pi0 and gamma using
  - Genetic algorithm

BACKUP

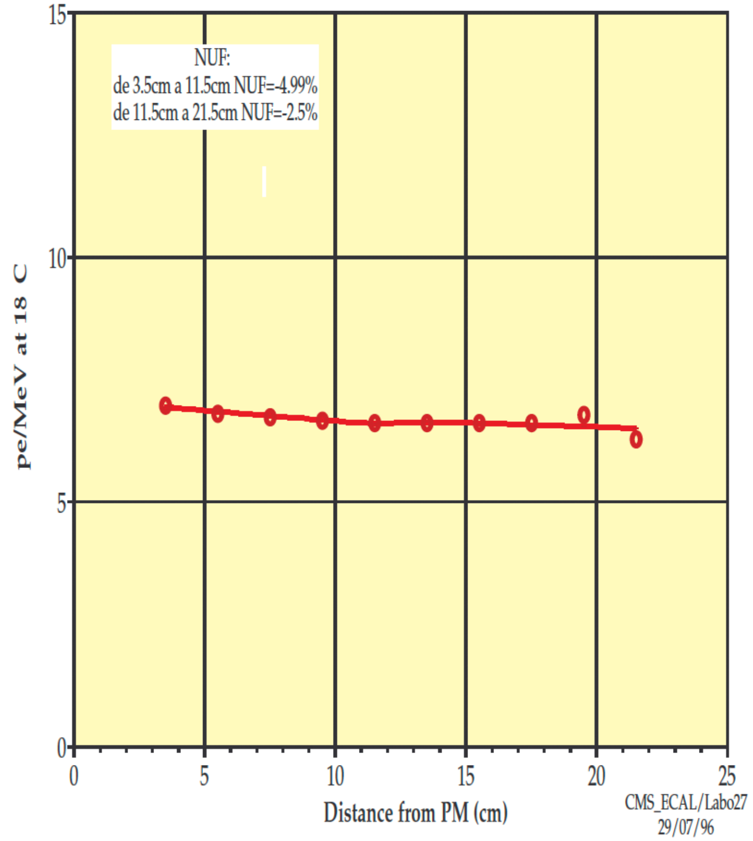
LIGHT COLLECTION EFFICIENCY

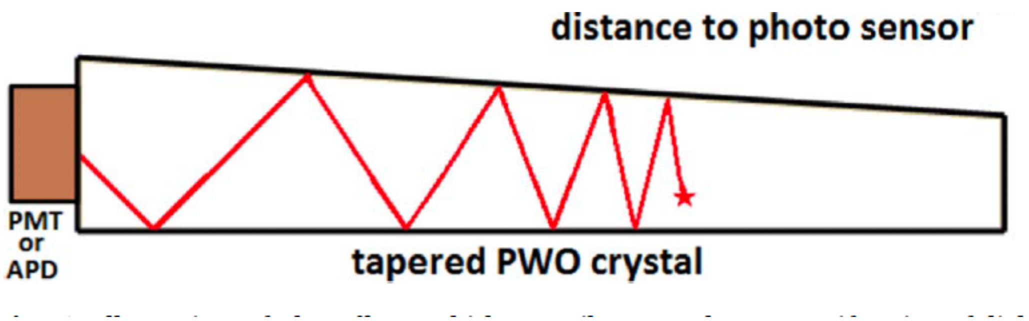


Before depolishing



After depolishing



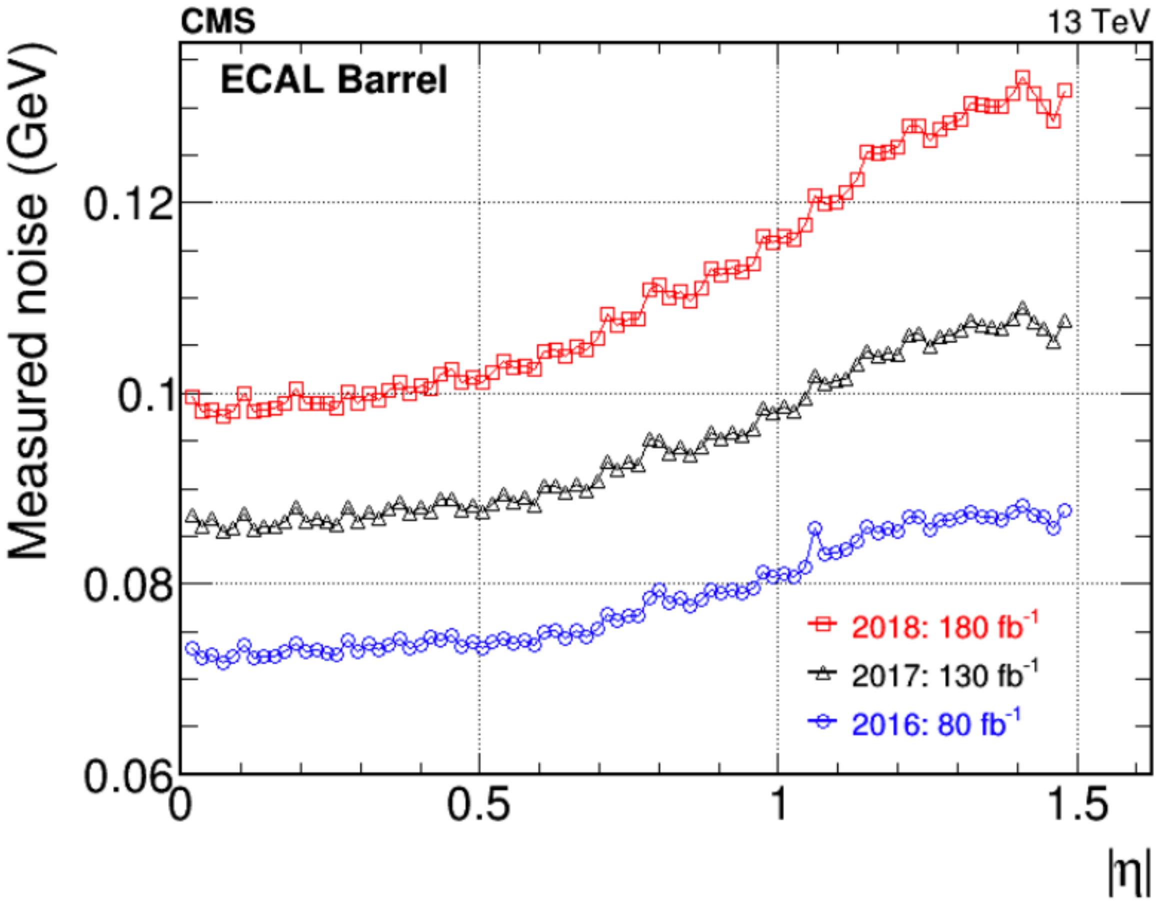


Depiction of focusing effect

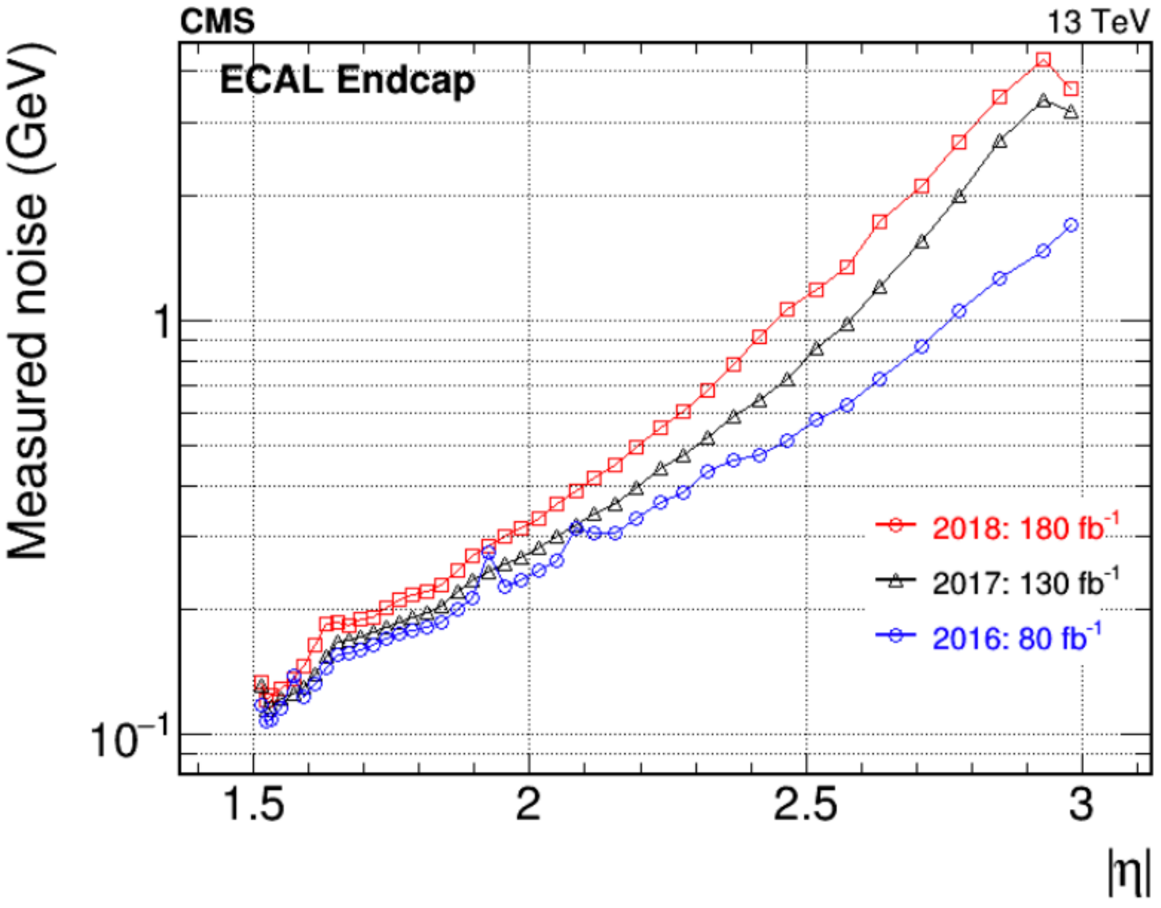
- The crystals that went inside the CMS detector had increased absorption lengths —> light does not get absorbed so fast as shown in the left plot above
  - Resulting light collection curves are dominated by focusing effects
  - This could introduce non-uniformity in the light collection and hence increase the constant term in the resolution
  - One side was depolished to reduce the focusing effect

References:
- Light collection uniformity of lead tungstate crystals for the CMS ECAL: NIM A 540(2005) 273-284
- NIM A 857 (2017) 1-6
- Internal CMS note: CMS CR 1998/004

Noise level in the ECAL Barrel and the Endcap



(c) EB, *E*



(d) EE, *E*

Noise increase in the EB due to increase in APD dark current

Noise increase in the EE due to decrease in transparency and

CMS-DP-2022/015

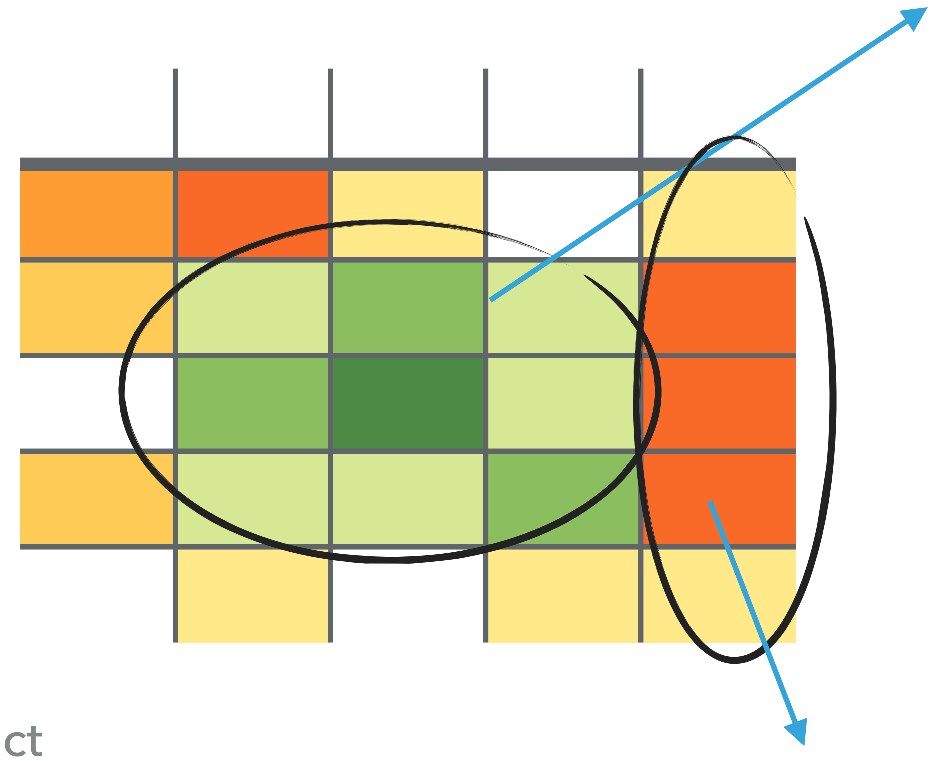
Effects of noise on physics

- Noise worsens the resolution of the detector

$$\frac{\sigma(E)}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c$$

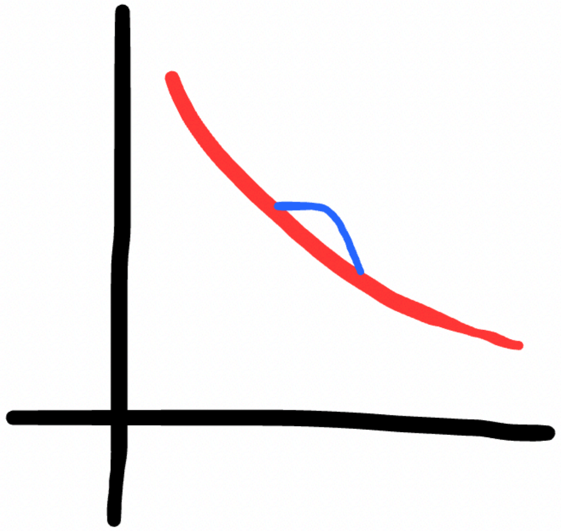
Noise term

Main shower

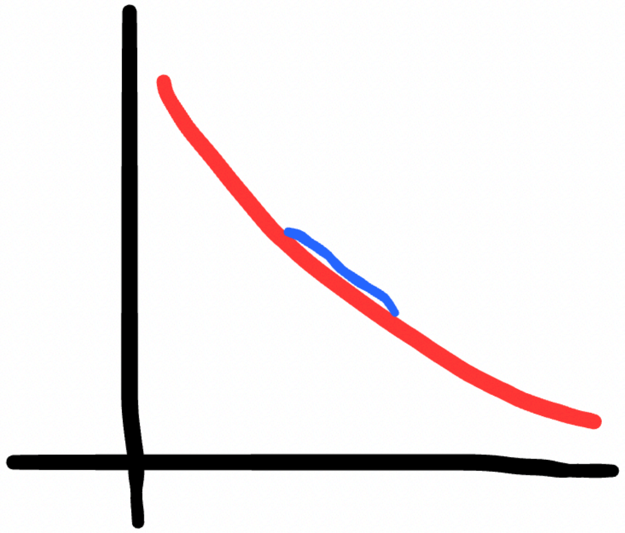


Noisy crystals
Color indicates
the amount of noise

  - Such noisy crystals are picked up in the clustering algorithm and worsen the total resolution of electron and photon object



Effect of low noise on the signal peak sitting on the background continuum

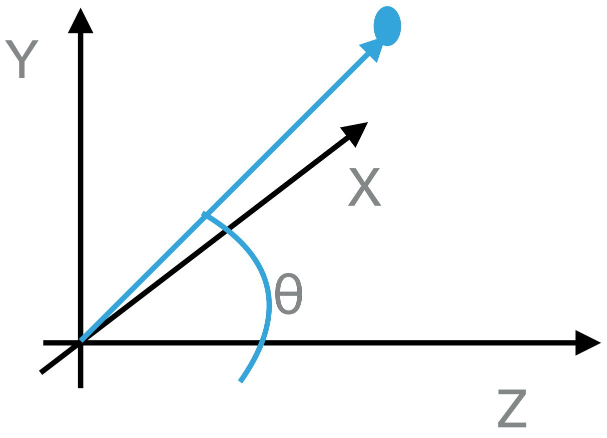


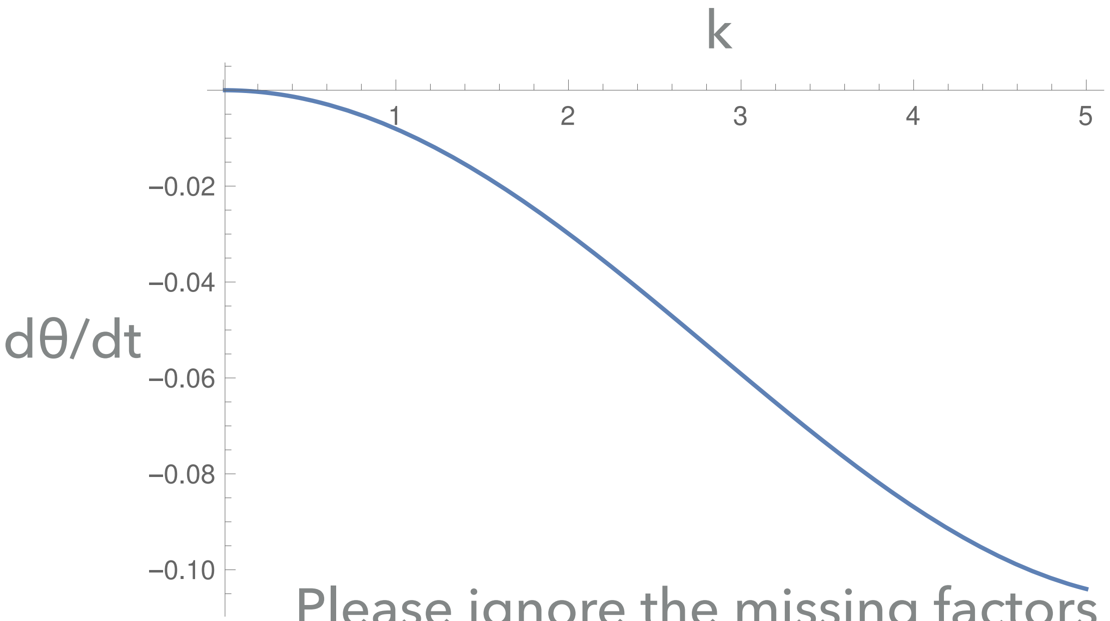
Effect of high noise on the signal peak sitting on the background continuum

- Remove noise by applying a threshold
- Threshold tuned in such a way that resolution is affected the least

BASIC IDEA OF MUSTACHE SUPERCLUSTERING

- Trajectory of a charged particle in magnetic field with mangetic field only in the Z direction:
  - X = sin(kt)/k ; Y = (1-cos(kt))/k at time t where k = 1./R where R is the radius of curvature
  - Z = Vz t; Vz = constant
  - R(t) = sqrt(x(t)^2 + y(t)^2 + z(t)^2)
  - tan(θ) = z(t)/R(t); Calculate d(θ)/dt



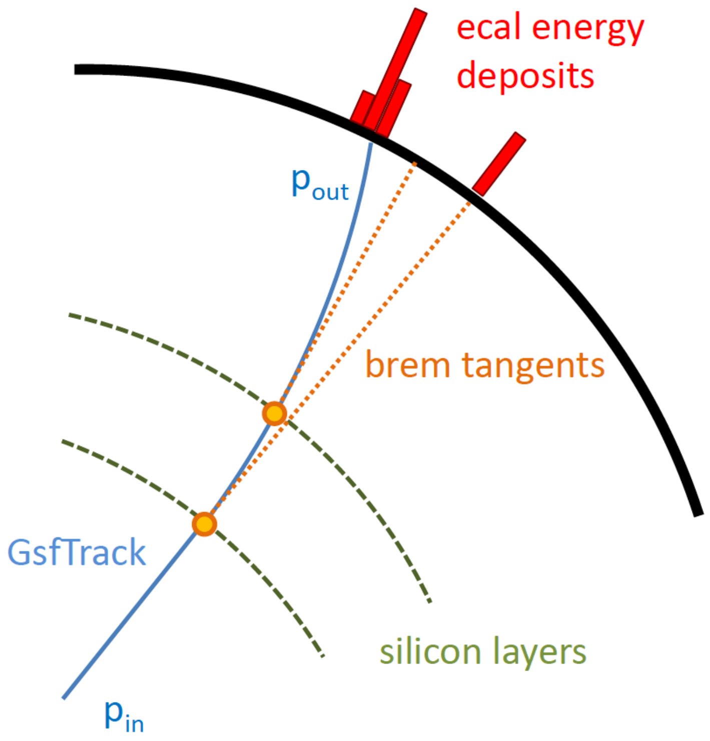


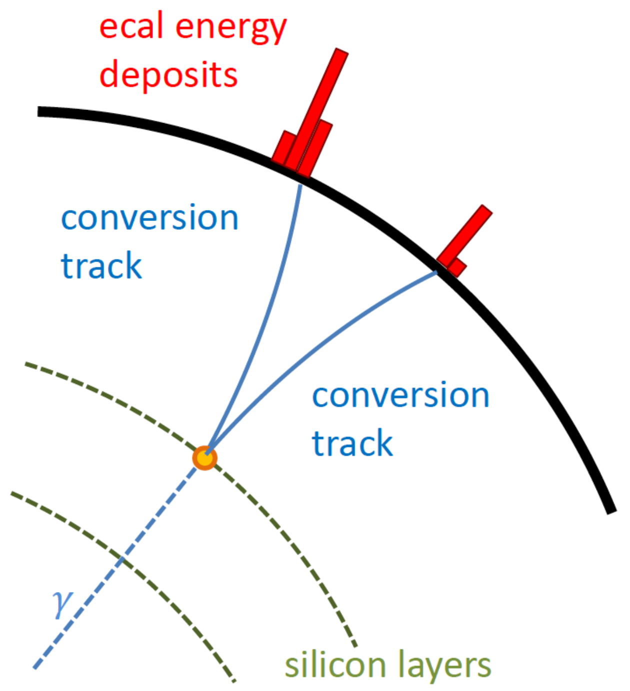
As R decreases, rate of change increases

Please ignore the missing factors to make arguments of sin and cosin dimensionless for now

REFINED SUPERCLUSTERS

Sam Harper





- The points at which track intersects a layer, a tangent is drawn and extended to the ECAL surface
- The cluster falling at tangent is included in the mustache SC to give the final refined SC
- All electrons and photons are made of these final refined SCs

EE TRANSVERSE FACE DIMENSIONS

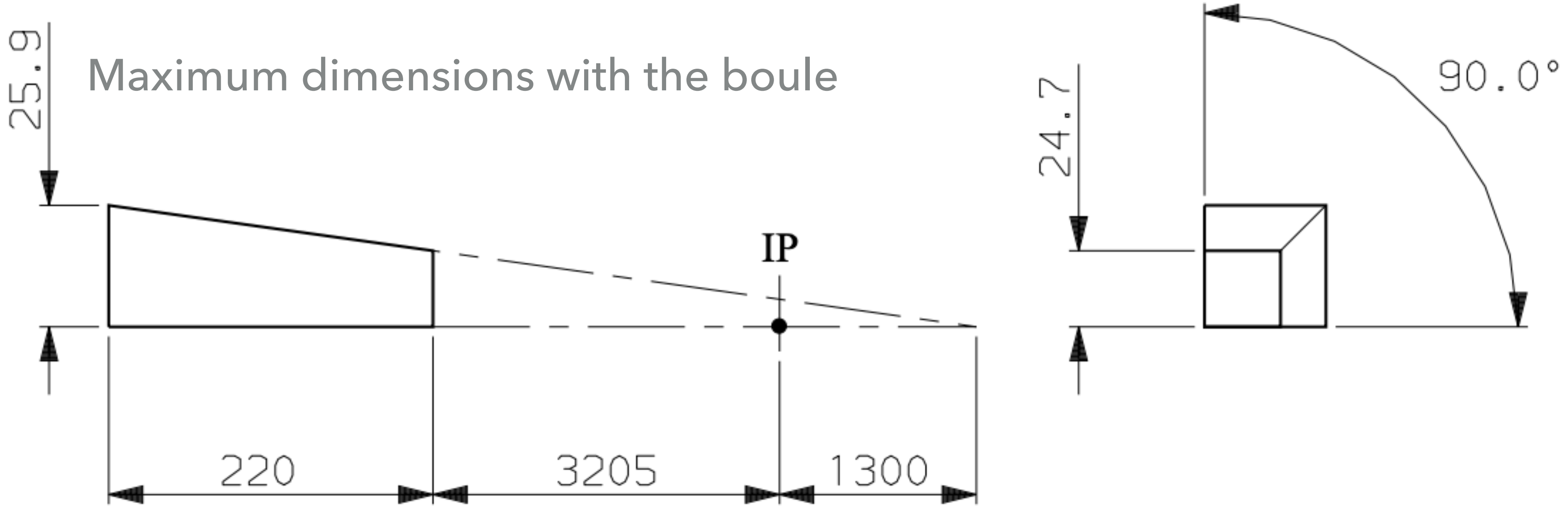
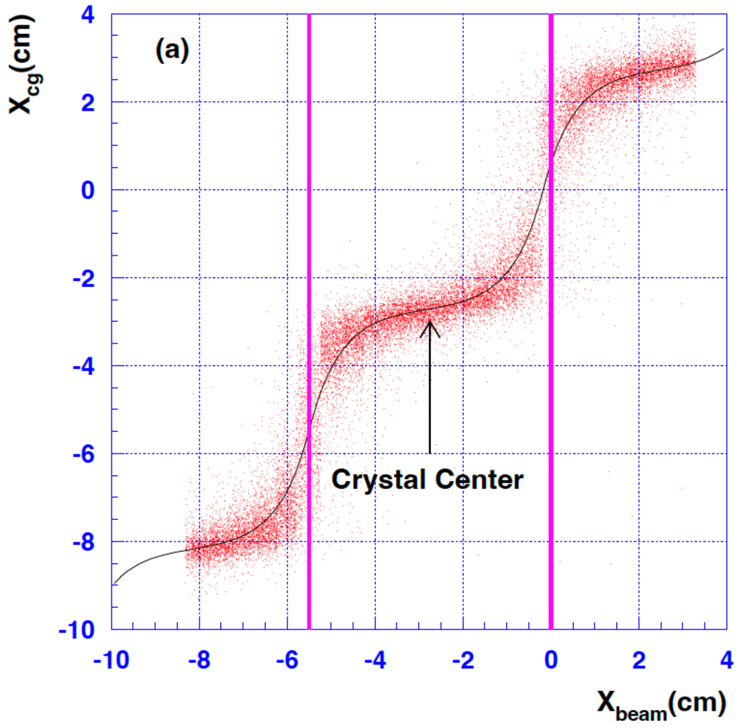
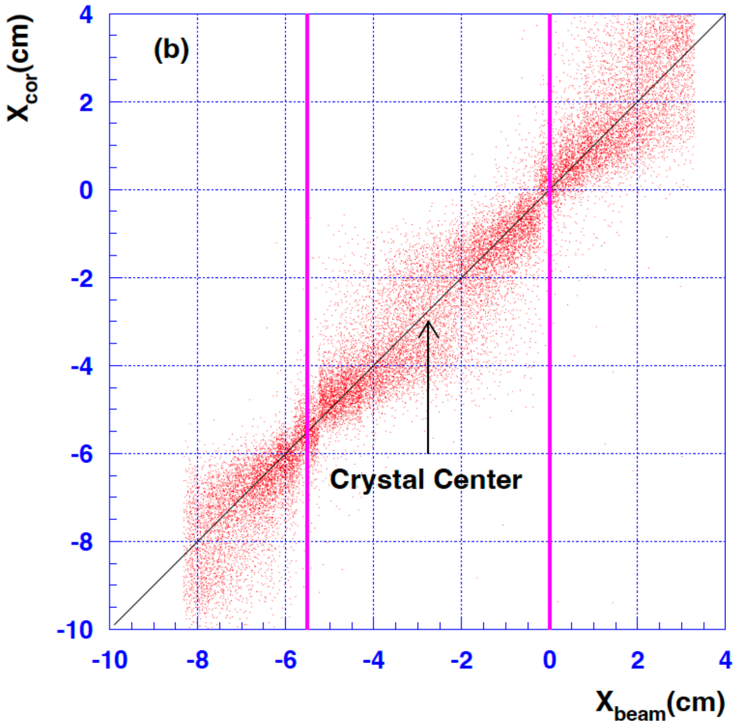
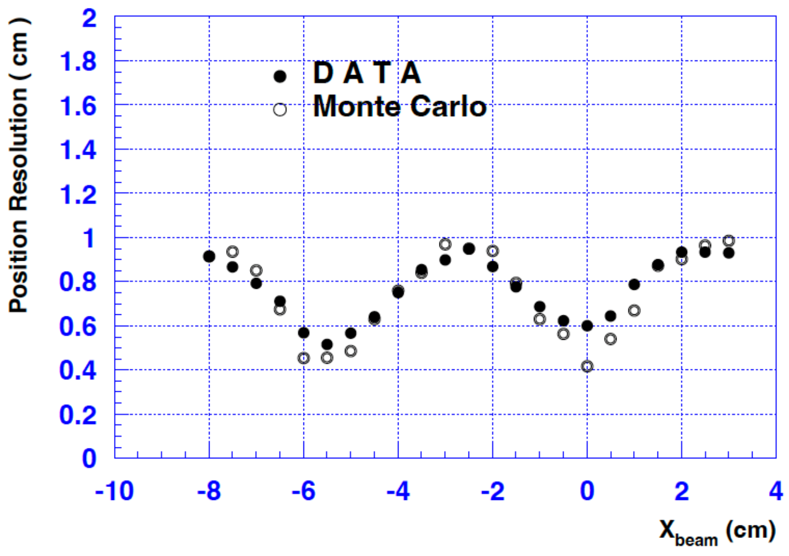


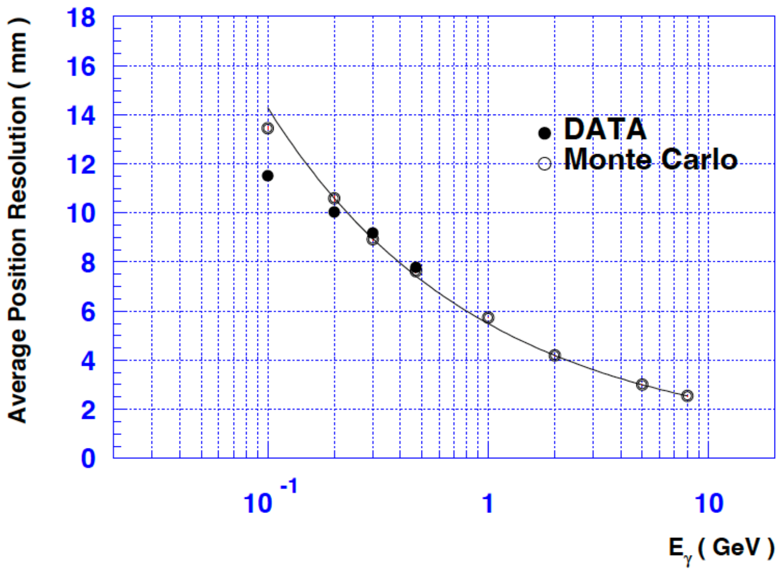
Fig. 3.28: EE crystal geometry (not to scale, dimensions in mm).

The geometric construction of the EE is based on a right-sided crystal with two tapering sides as shown in Fig. 3.28. The taper is defined by a line from a point 1300 mm from the far side of the intersection point, to the rear corner of the crystal. The taper defines the size of the front face of the crystal. The maximum crystal width, at the rear, that can be obtained with the current crystal boules is 25.9 mm. The corresponding front-face width is 24.7 mm. The taper on the crystal is small, only 1.2 mm over the full crystal length of 220 mm. Off-pointing to the far side of the intersection point is required in order to ensure maximum path length through the EE crystals.









- Position resolution is much lesser than the true detector size
-

ECAL READOUT CHAIN

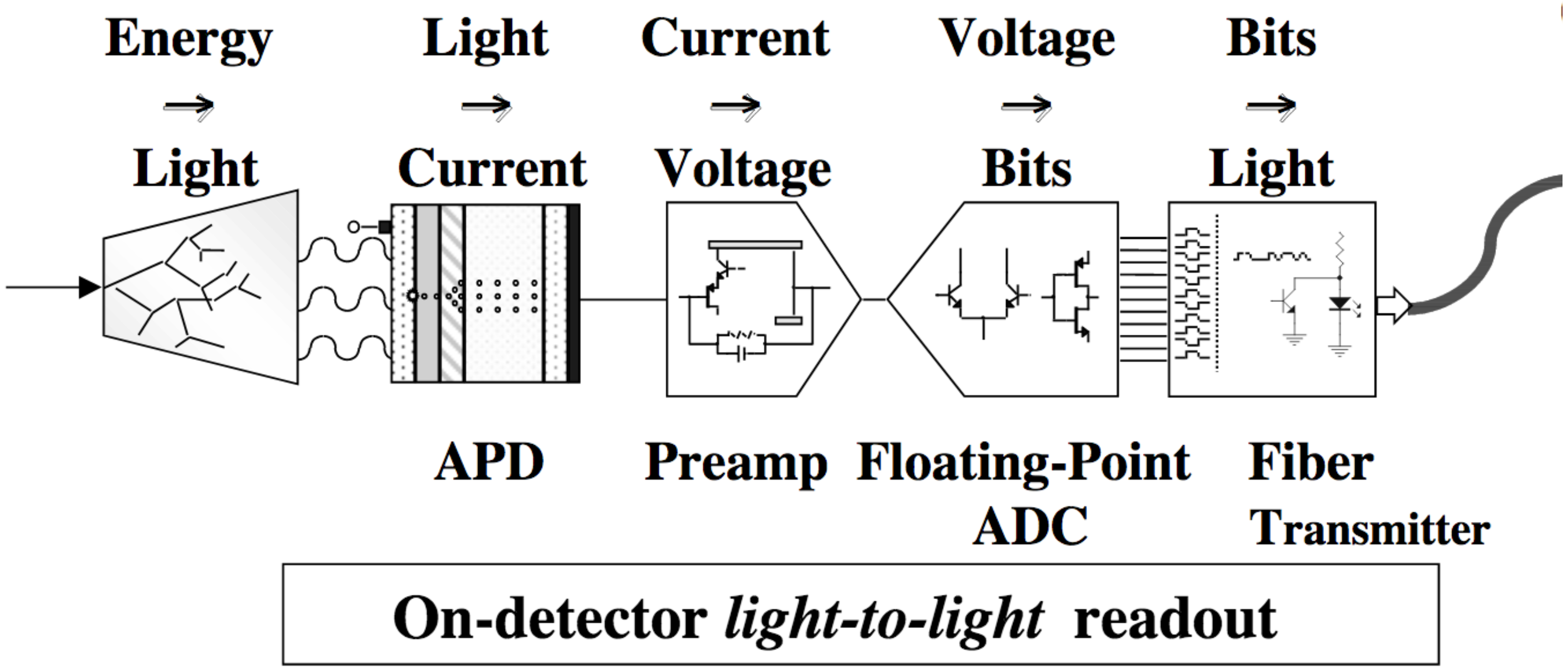
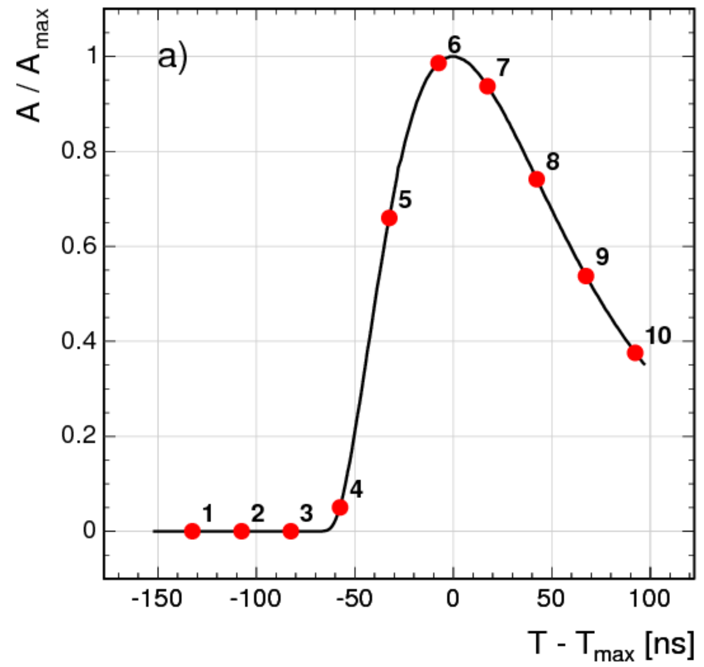
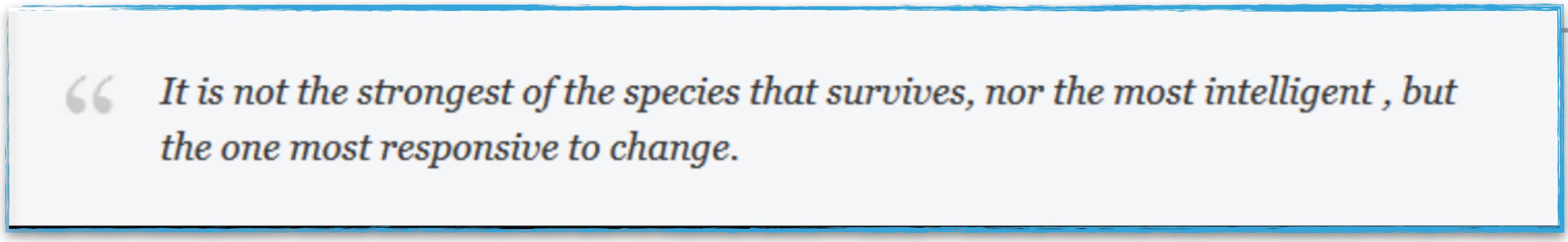
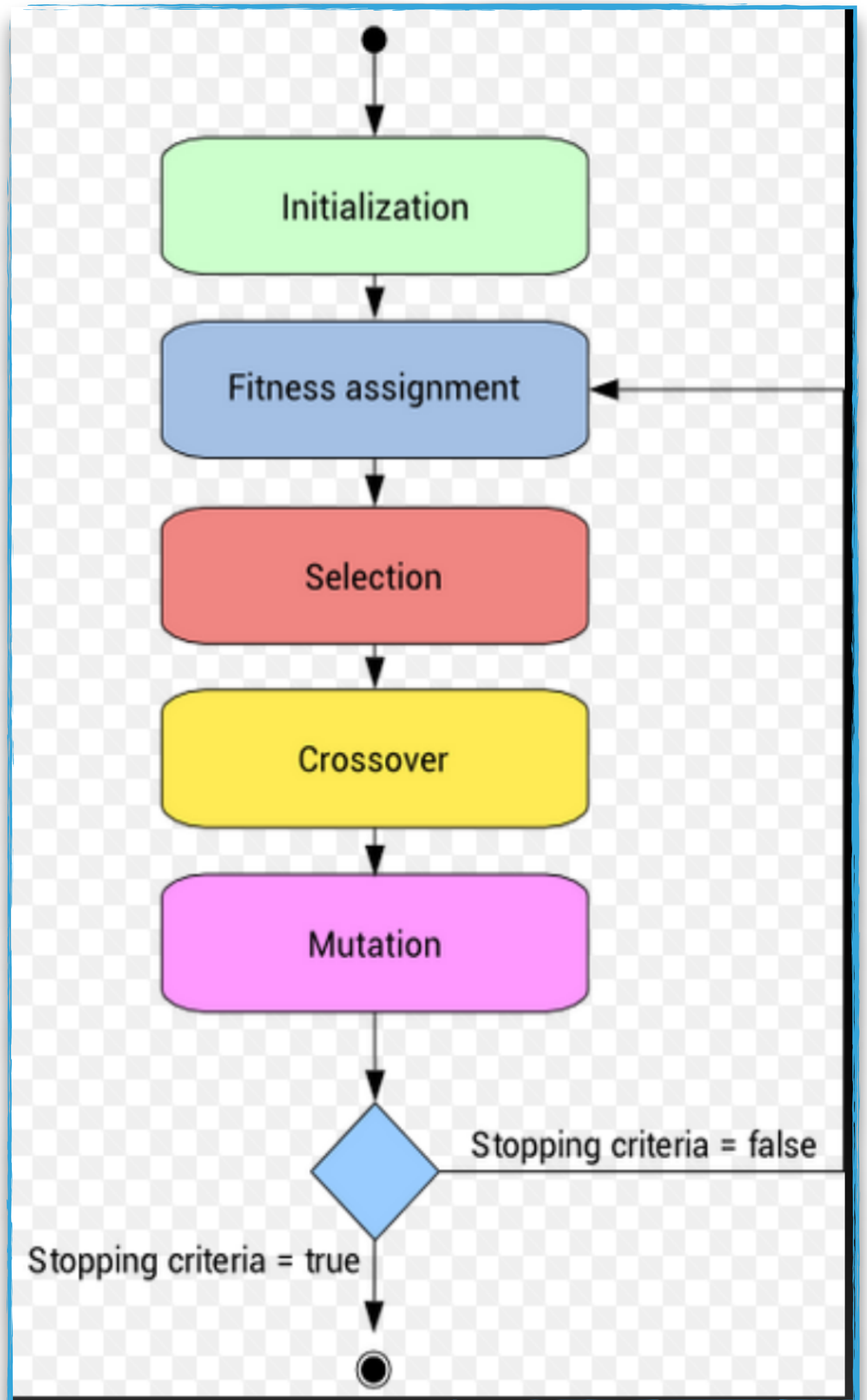


Fig. 5.4: ECAL readout chain.





> *It is not the strongest of the species that survives, nor the most intelligent , but the one most responsive to change.*



Borrowed from [Debabrata's](#) slides

- Variable: gene, event: chromosome, all events: population

- Select the most signal like events as parents

- Take the variables from each parent and again populate the next set of events to have the most signal like sample

MULTIVARIATE ANALYSIS USING BOOSTED DECISION TREES

- Introduced in HEP from the time of MiniBooNE (ν_μ–>ν_e) experiment where it was used for particle ID and it was demonstrated to be better than ANNs for their use case
- BDTs combine many "weak" classifiers to a single powerful classifier
- These are also used for regression that we learnt in the previous exercise

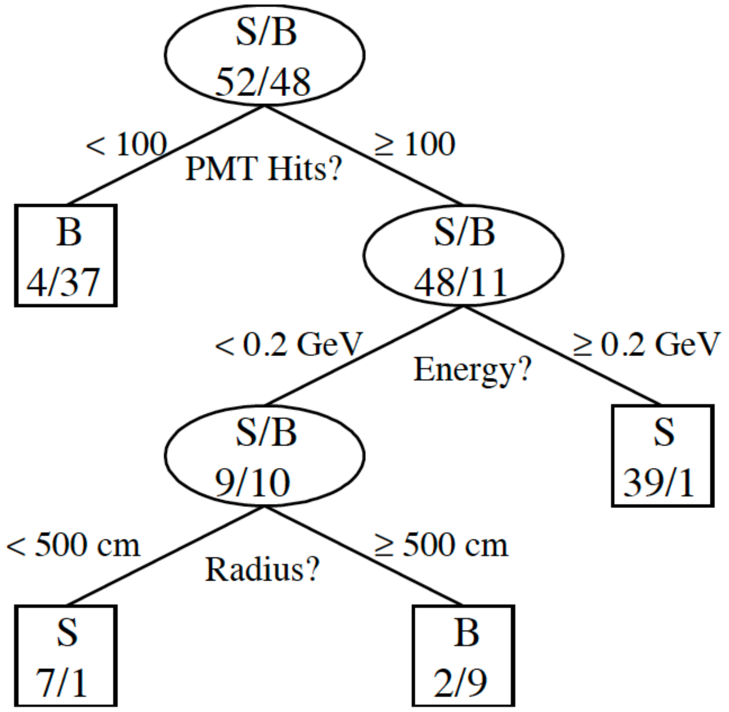
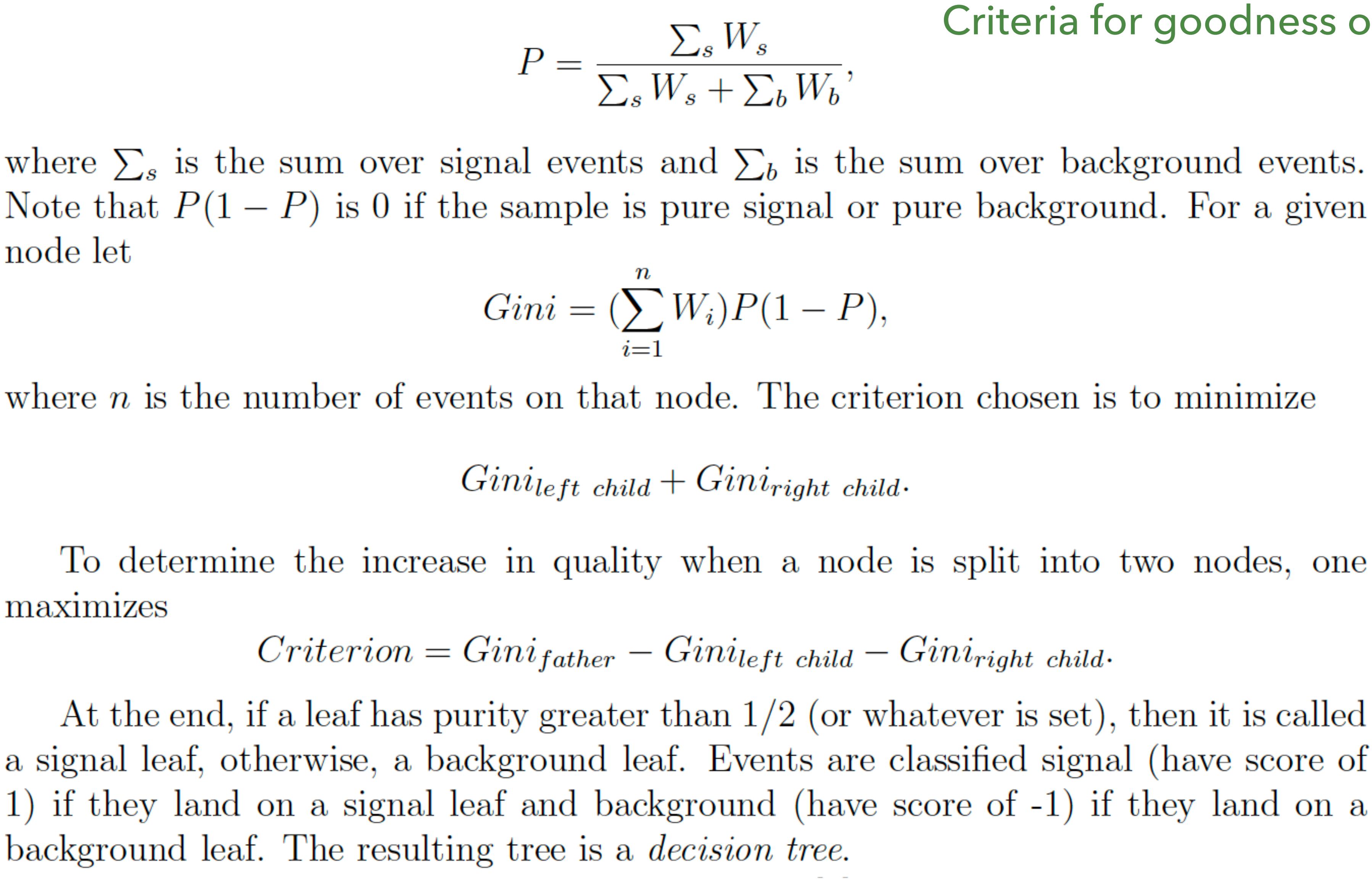


Figure 1: Schematic of a decision tree.

MULTIVARIATE ANALYSIS USING BOOSTED DECISION TREES

Criteria for goodness of separation



$$P = \frac{\sum_s W_s}{\sum_s W_s + \sum_b W_b},$$

where $\sum_s$ is the sum over signal events and $\sum_b$ is the sum over background events. Note that $P(1-P)$ is 0 if the sample is pure signal or pure background. For a given node let

$$Gini = (\sum_{i=1}^{n} W_i) P(1-P),$$

where n is the number of events on that node. The criterion chosen is to minimize

$$Gini_{left\ child} + Gini_{right\ child}.$$

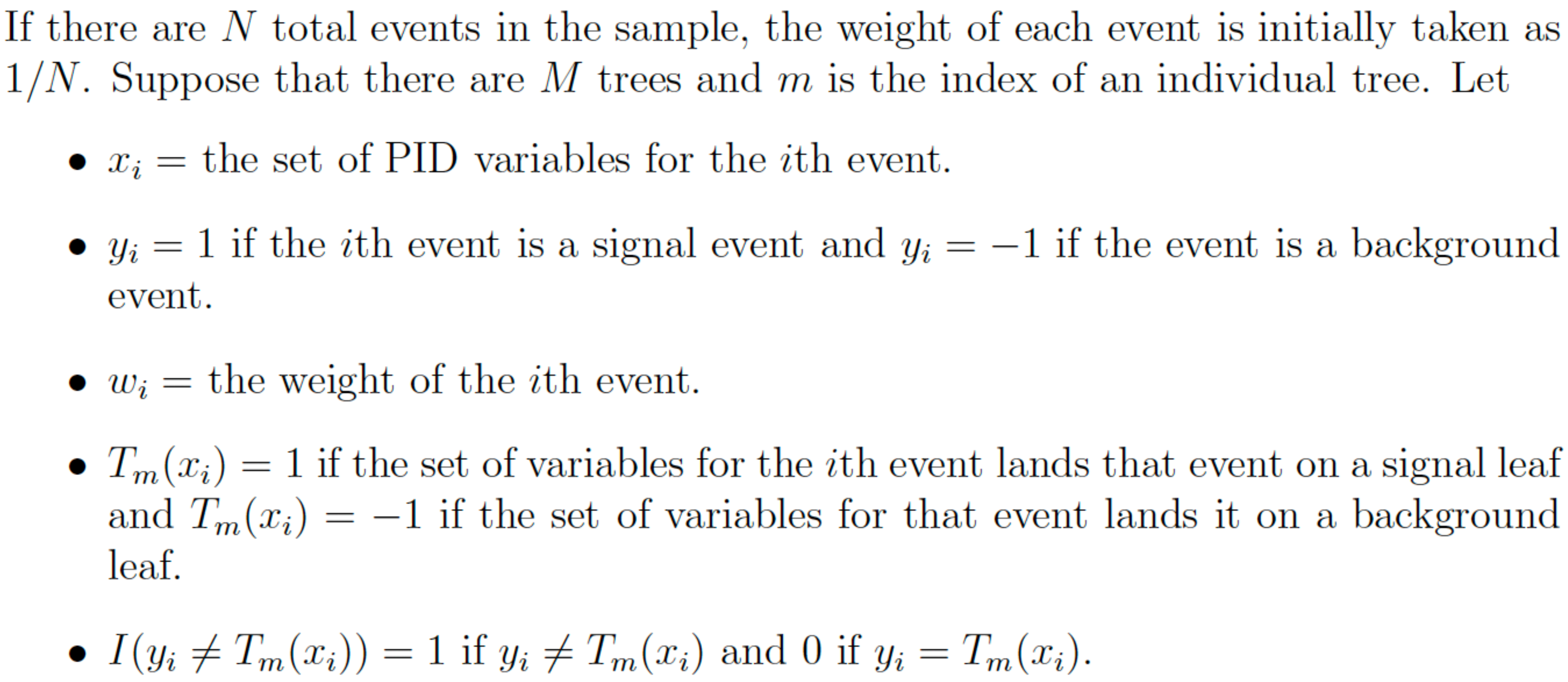
To determine the increase in quality when a node is split into two nodes, one maximizes

$$Criterion = Gini_{father} - Gini_{left\ child} - Gini_{right\ child}.$$

At the end, if a leaf has purity greater than 1/2 (or whatever is set), then it is called a signal leaf, otherwise, a background leaf. Events are classified signal (have score of 1) if they land on a signal leaf and background (have score of -1) if they land on a background leaf. The resulting tree is a *decision tree*.

TREATMENT OF WRONGLY ASSIGNED EVENTS

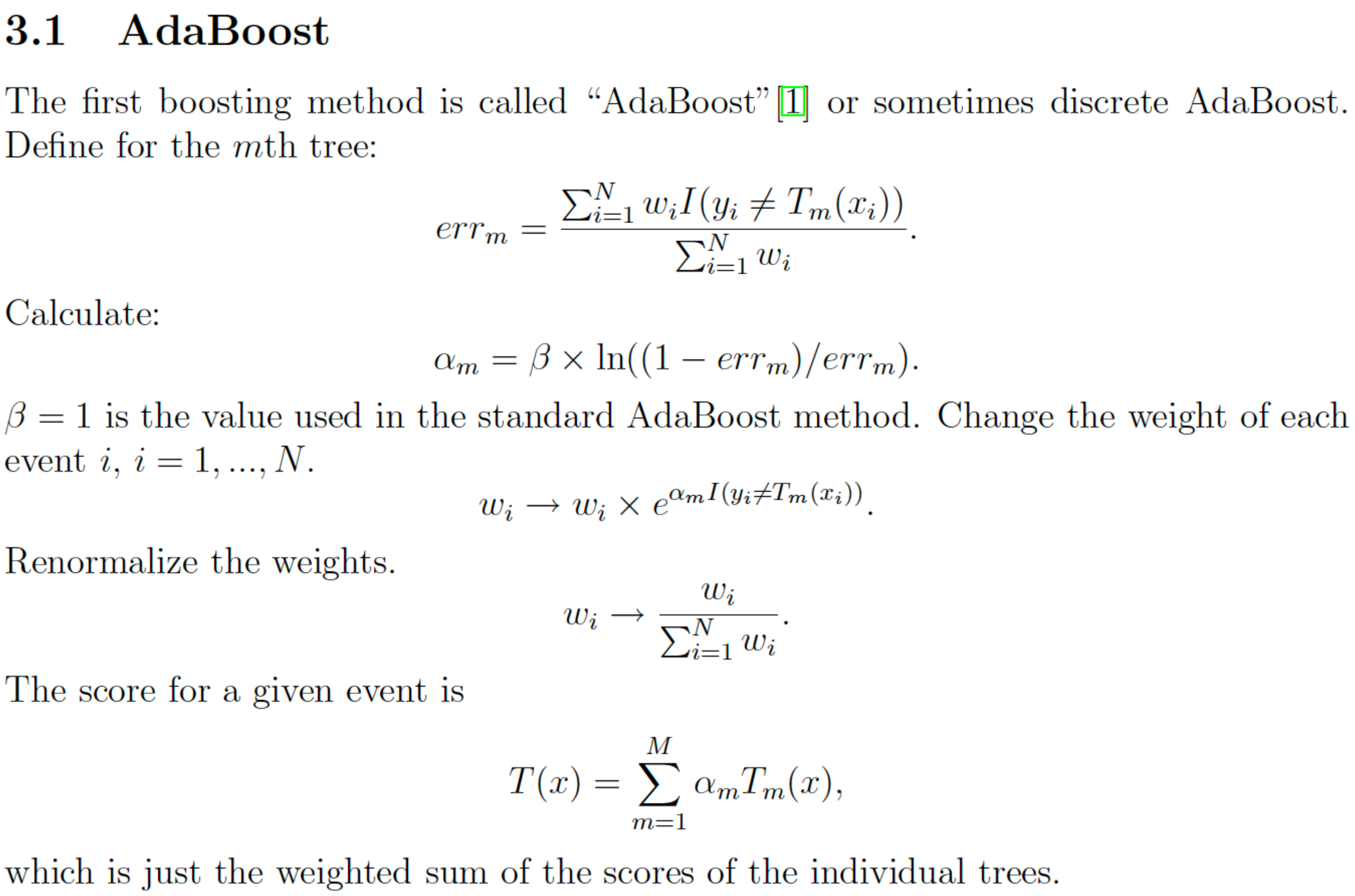
- Boost the events with wrong assignment with high weights



If there are N total events in the sample, the weight of each event is initially taken as $1/N$. Suppose that there are M trees and m is the index of an individual tree. Let

- x_i = the set of PID variables for the ith event.
- $y_i = 1$ if the ith event is a signal event and $y_i = -1$ if the event is a background event.
- w_i = the weight of the ith event.
- $T_m(x_i) = 1$ if the set of variables for the ith event lands that event on a signal leaf and $T_m(x_i) = -1$ if the set of variables for that event lands it on a background leaf.
- $I(y_i \neq T_m(x_i)) = 1$ if $y_i \neq T_m(x_i)$ and 0 if $y_i = T_m(x_i)$.

BOOSTING ALGORITHM



3.1 AdaBoost

The first boosting method is called "AdaBoost"[1] or sometimes discrete AdaBoost. Define for the mth tree:

$$err_m = \frac{\sum_{i=1}^{N} w_i I(y_i \neq T_m(x_i))}{\sum_{i=1}^{N} w_i}.$$

Calculate:

$$\alpha_m = \beta \times \ln((1 - err_m)/err_m).$$

$\beta = 1$ is the value used in the standard AdaBoost method. Change the weight of each event i, $i = 1, ..., N$.

$$w_i \rightarrow w_i \times e^{\alpha_m I(y_i \neq T_m(x_i))}.$$

Renormalize the weights.

$$w_i \rightarrow \frac{w_i}{\sum_{i=1}^{N} w_i}.$$

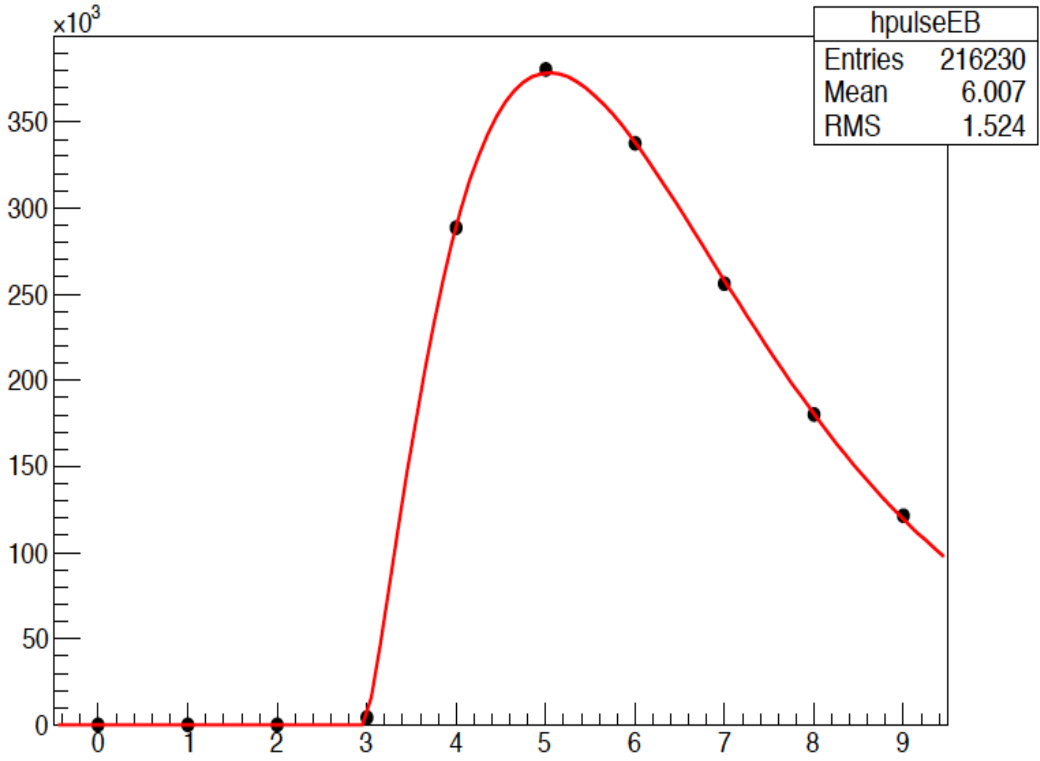
The score for a given event is

$$T(x) = \sum_{m=1}^{M} \alpha_m T_m(x),$$

which is just the weighted sum of the scores of the individual trees.

10 TIME SAMPLES —> ENERGY (IN RUN I, I.E. 2009 TO 2012)

- Ref: https://link.springer.com/article/10.1140/epjcd/s2006-02-002-x
- Based on digital filtering technique
- Amplitude is reconstructed by
  - $A = \Sigma w_i S_i$
  - S_i are the 10 time samples
  - w_i are the weights determined by reducing the χ^2
  - x^2 is calculated as follows:
    - Fit the pulse shape with a suitable function (F) (calle Function is given by:
    - Then form a x^2 between A and F and minimize



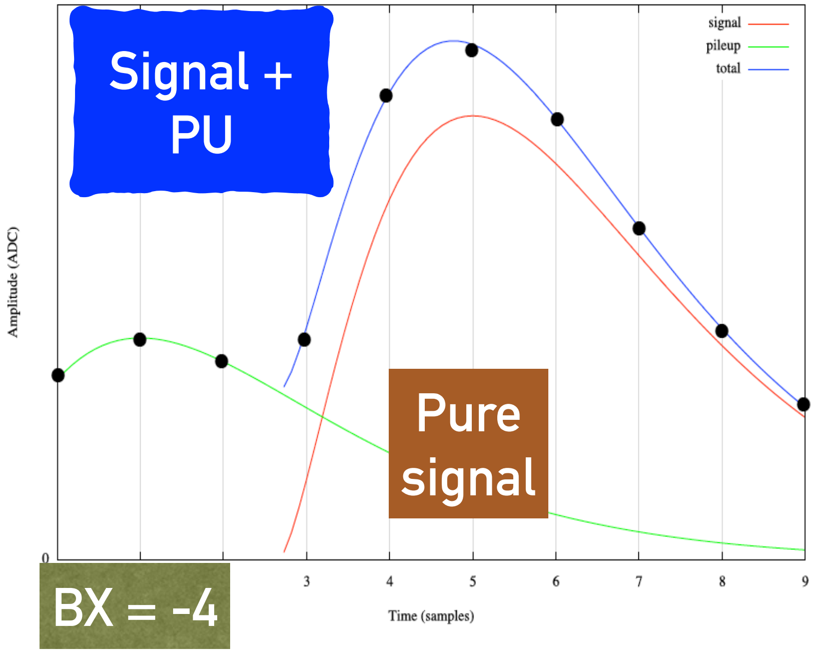
First 3/4 samples are pedestal

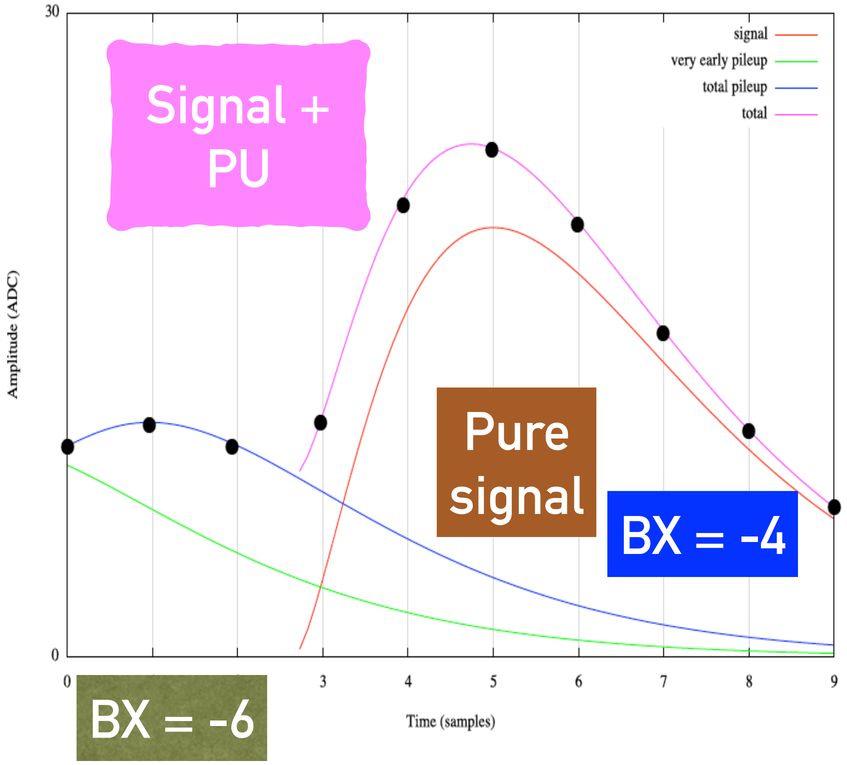
$$A(t) = Ped + A\left(1 + \frac{\Delta t}{\alpha\beta}\right)^{\alpha} \cdot e^{\frac{\Delta t}{\beta}}$$

https://indico.cern.ch/event/292930/contributions/671061/attachments/547860/755142/edm-ecalreco-pu-21Aug2014.pdf

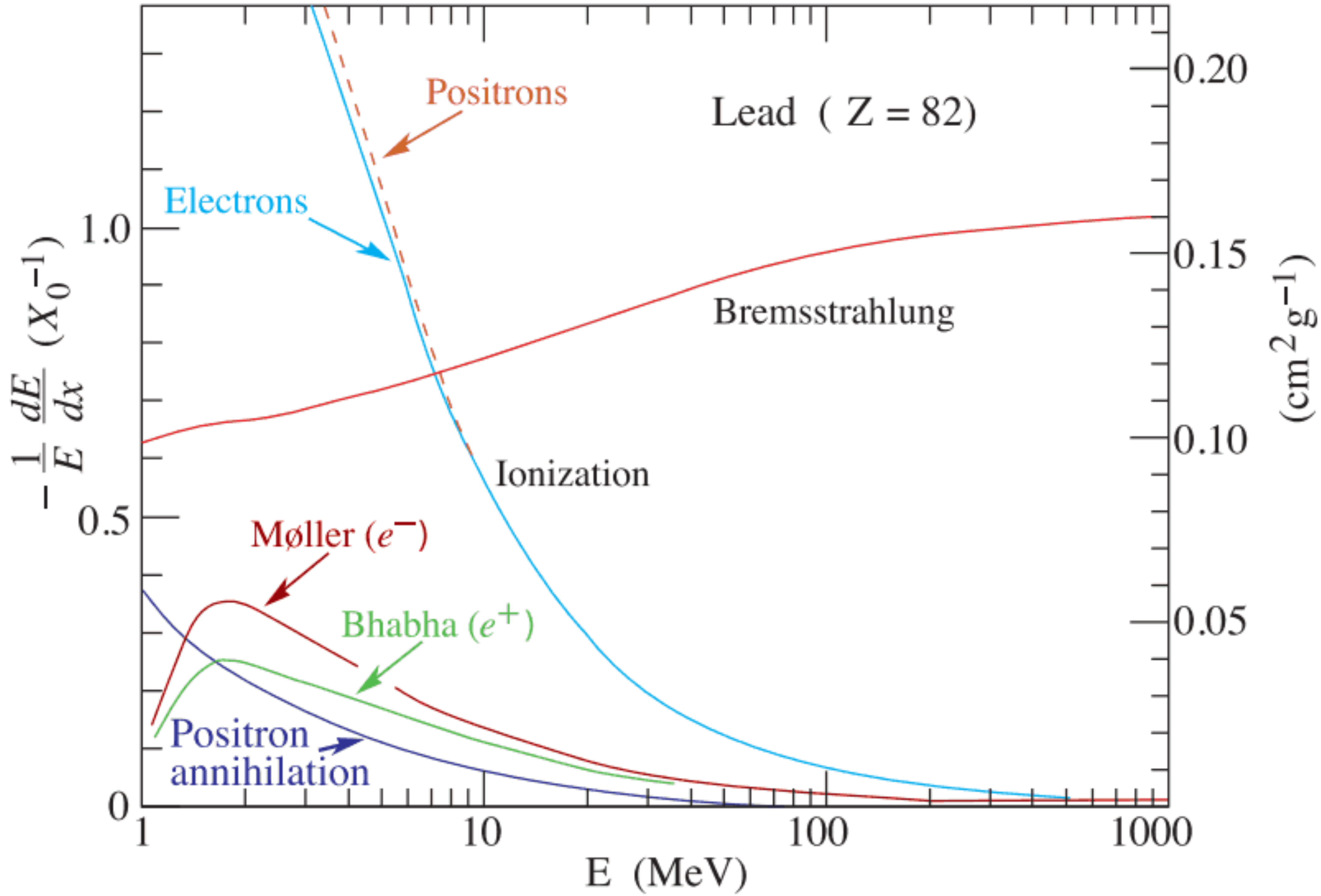
AMPLITUDE RECONSTRUCTION

- In going from Run I to Run II, the instantaneous luminosity has increased (by a factor of ~2)
  - # inelastic collisions per LHC bunch crossing has increased
  - Also, the bunch spacing has decreased from 50ns to 25ns –> **higher out-of-time PU**
- In the presence of out-of-time PU, the main signal gets energy from additional bunch crossing
  - Important to subtract this contribution
- To take care of this, in Run II, multi-fit method was developed

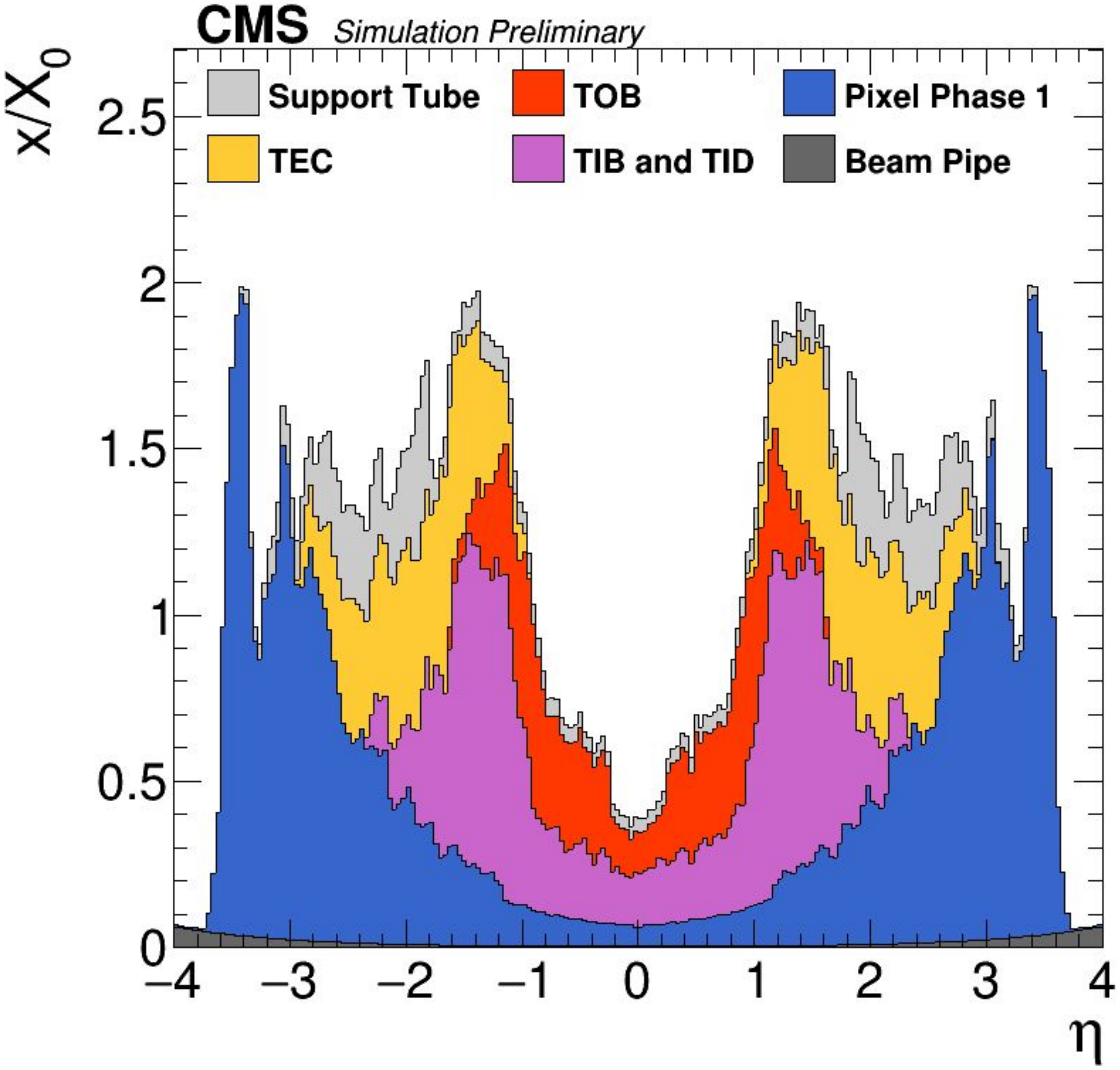




https://indico.cern.ch/event/292930/contributions/671061/attachments/547860/755142/edm-ecalreco-pu-21Aug2014.pdf


Lead (Z = 82)
Positrons
Electrons
Bremsstrahlung
Ionization
Møller (e^-)
Bhabha (e^+)
Positron annihilation
$-\frac{1}{E}\frac{dE}{dx}$ (X_0^{-1})
(cm^2 g^{-1})
E (MeV)
1.0
0.5
0
0.20
0.15
0.10
0.05
1
10
100
1000

CONVERSIONS IN CMS



- Significant amount of tracker material infront of the ECAL.
- Probability for a photon to convert to e+e- pair at eta ~1 is (1-exp(-7./9)) = 54%
    - Conversions are important to reconstruct!

CONVERSION RECONSTRUCTION - SINGLE AND DOUBLE LEG

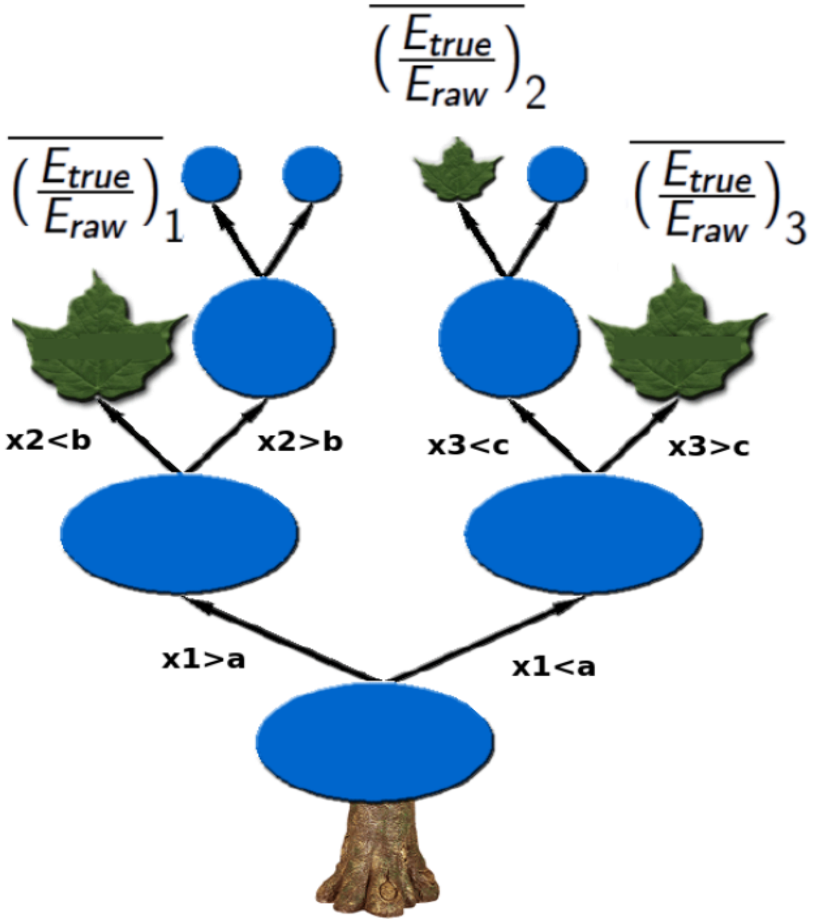
- Cross-section for conversion can be written as: $\frac{d\sigma}{dx} = \frac{A}{X_0 N_A}(1 - \frac{4}{3}x(1-x))$
    - where x is the momentum fraction of electron w.r.t the photon. Cross-section is symmetric in exchange of x and (1-x) and hence asymmetric conversions can happen
- Nearly Symmetric conversions (**double leg conversions**): Energy imparted to both the tracks is nearly the same. Two tracks can be reconstructed in the tracker
- Highly asymmetric conversions (**single leg conversion**) :
    - one of the tracks can be very low pT and hence may not be reconstructed within the tracker
- CMS uses dedicated algorithms to reconstruct the conversions in which 2 legs of the conversions are reconstructed and the case when only 1 leg is reconstructed

ENERGY CORRECTIONS

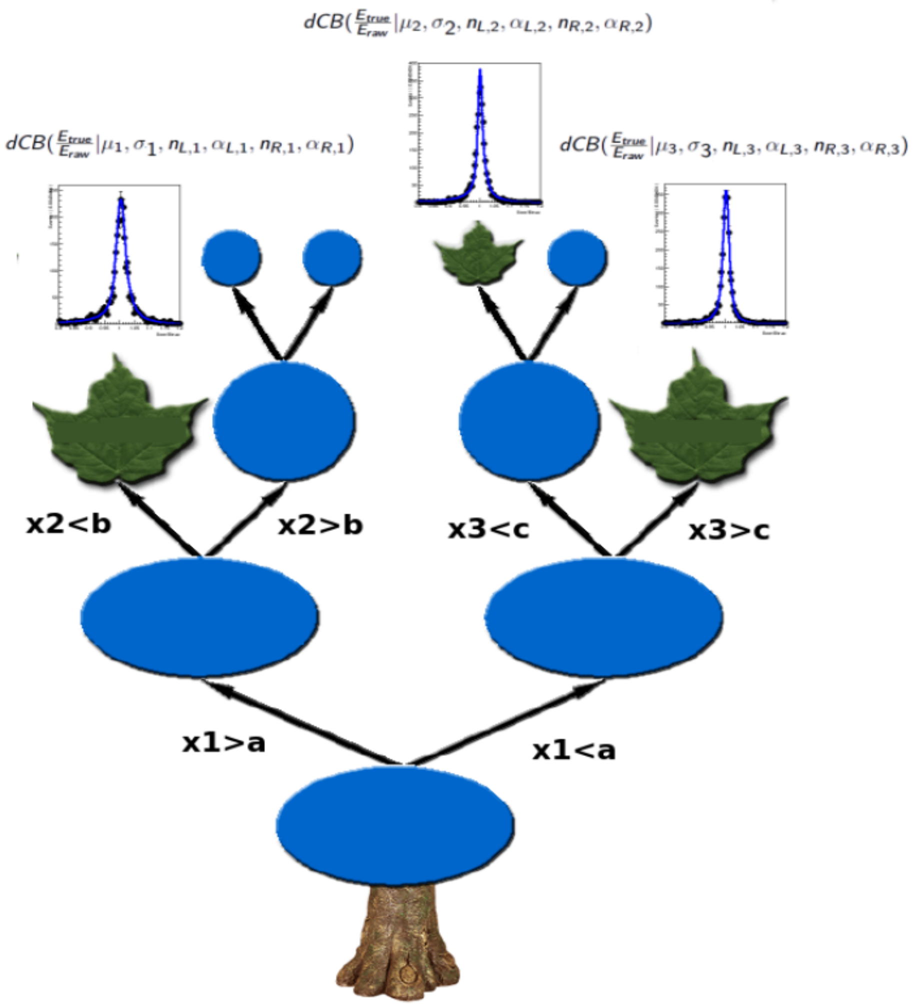
- The super-cluster algorithm still does not recover the full energy of electron/photon showers which is lost due to several reasons:
  - Longitudinal leakage, inter-modular gaps, energy lost in the tracker etc
  - To recover this lost energy and resolution, energy regression is performed
- A semi-parameteric regression based on BDTs is trained

A WORD ON BDT BASED SEMIPARAMETERIC REGRESSION

Traditionnal Regression


$\overline{\left(\frac{E_{true}}{E_{raw}}\right)_2}$
$\overline{\left(\frac{E_{true}}{E_{raw}}\right)_1}$
$\overline{\left(\frac{E_{true}}{E_{raw}}\right)_3}$
x2<b
x2>b
x3<c
x3>c
x1>a
x1<a

Semi-Parametric Regression


$dCB(\frac{E_{true}}{E_{raw}}|\mu_2, \sigma_2, n_{L,2}, \alpha_{L,2}, n_{R,2}, \alpha_{R,2})$
$dCB(\frac{E_{true}}{E_{raw}}|\mu_1, \sigma_1, n_{L,1}, \alpha_{L,1}, n_{R,1}, \alpha_{R,1})$
$dCB(\frac{E_{true}}{E_{raw}}|\mu_3, \sigma_3, n_{L,3}, \alpha_{L,3}, n_{R,3}, \alpha_{R,3})$
x2<b
x2>b
x3<c
x3>c
x1>a
x1<a

- Given the input variables X, we compute the prediction F(X) of the target variable y = Etrue/Eraw
- A regression tree divides training events into plenty of regions of the phase space (leaves), each leaf containing events corresponding to a similar value of the target
- At the end, the prediction F(X) corresponds to the weighted median of the individual tree prediction

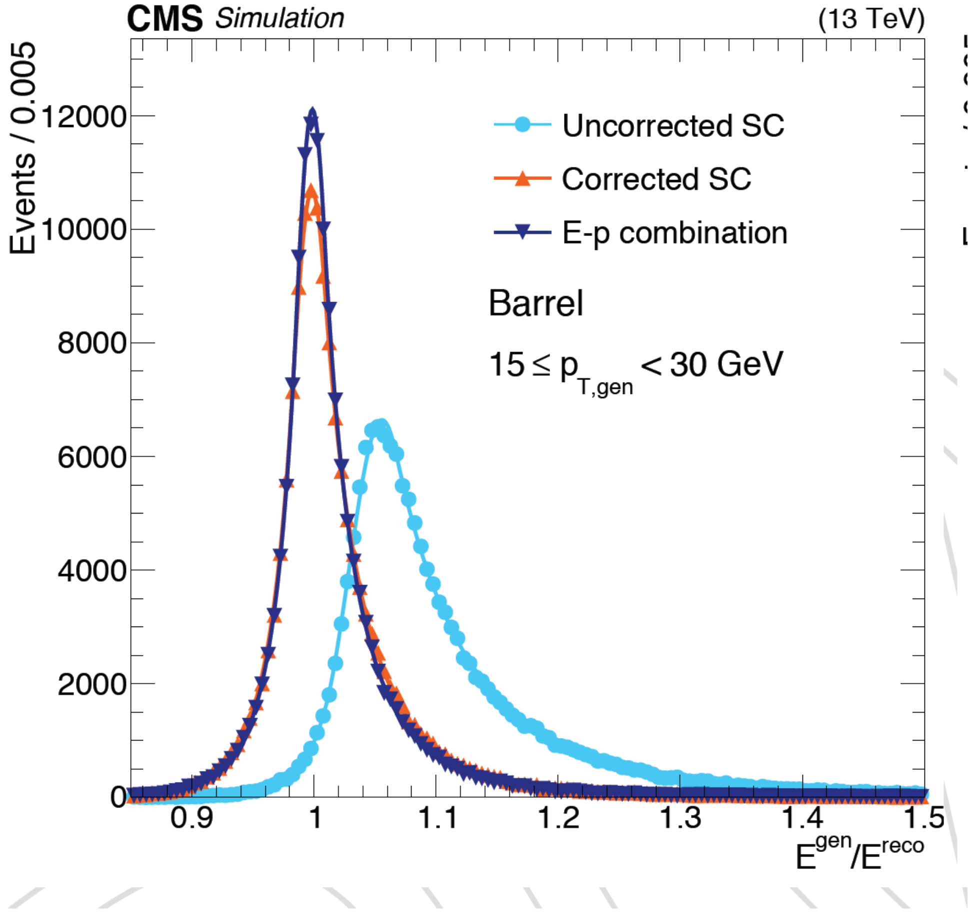
- Similar to traditional approach.
- Difference is that here instead of taking the weighted median of the individual tree prediction as the F(X), here we fit the full target distribution in each region of the phase space.
- So if the distribution is given by a double crystal ball (CB), this technique can be seen as a simultaneous regression of 6 targets which are the 6 parameters of the double CB

ELECTRON AND PHOTON REGRESSION

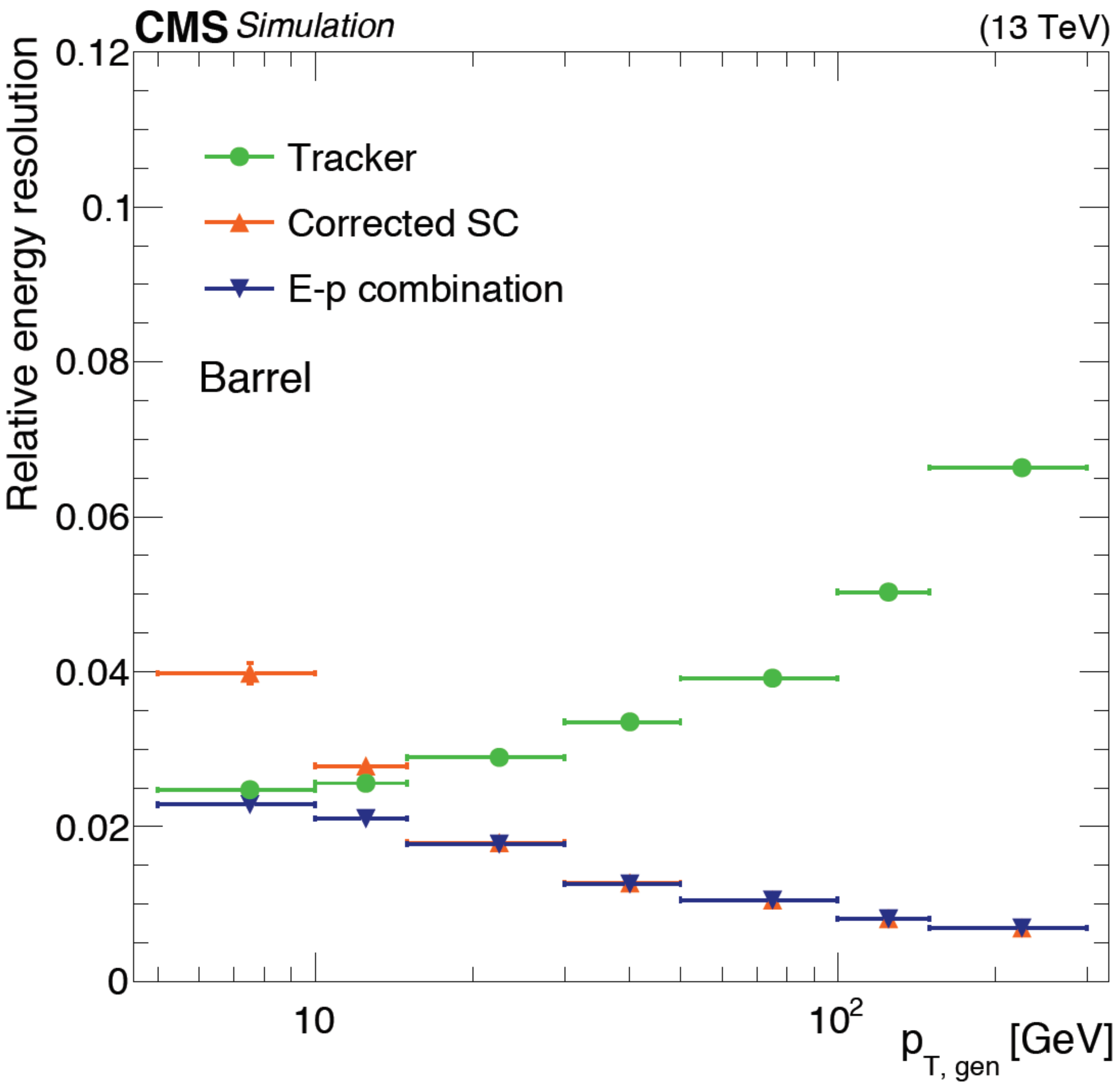
- Photons: Regressions is trained using all the ECAL related quantities (shower-shape variables which include R9, energy ratios etc)
    - Target is related to E_{true}/E_{raw}
- Electrons: In addition to above, here we have information from tracker as well. So another regression is trained which combines the ECAL corrected energy which is got in 1st regression and the tracker momentum

$$E^{\mathrm{reco}}_{\mathrm{combined}} = \frac{E_{\mathrm{ECAL}}/\sigma_E^2 + p_{\mathrm{tracker}}/\sigma_p^2}{1/\sigma_E^2 + 1/\sigma_p^2},$$

EGM-17-001



High tail is due to energy loss and low tail is due to energy over-estimation due to PU etc



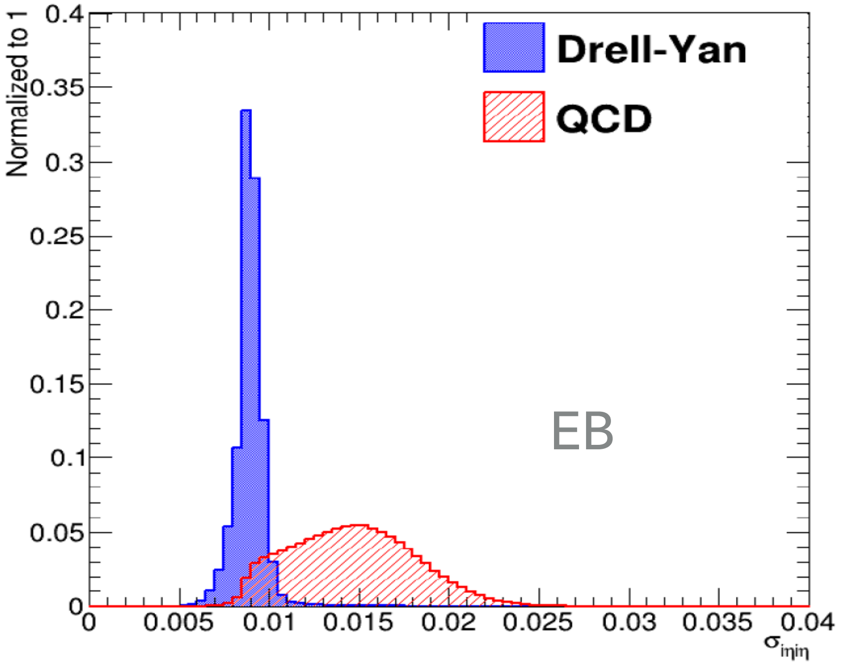
EGM-17-001

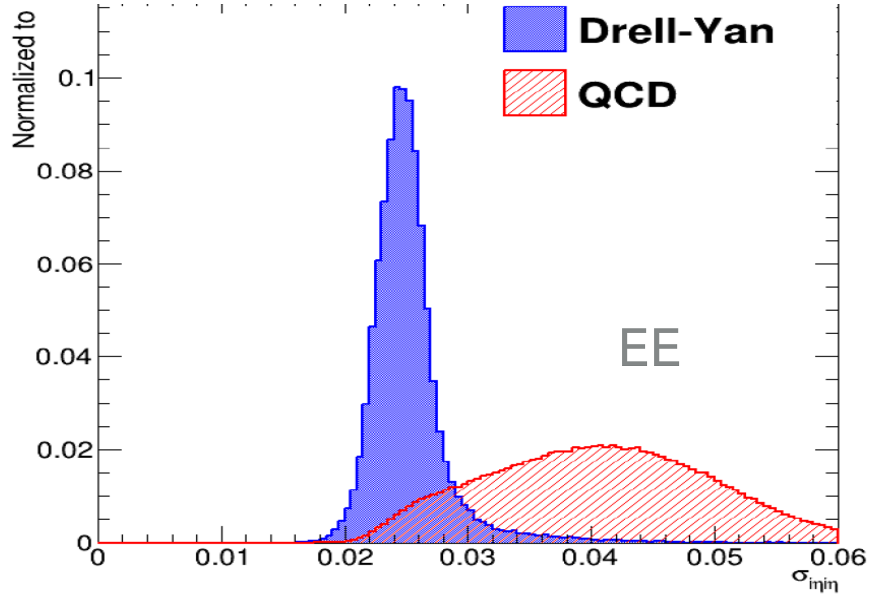
- At low pT, E-p combination brings huge improvement
- At high pT, tracker resolution becomes better and ECAL resolutions improves, the regression is dominated by the ECAL

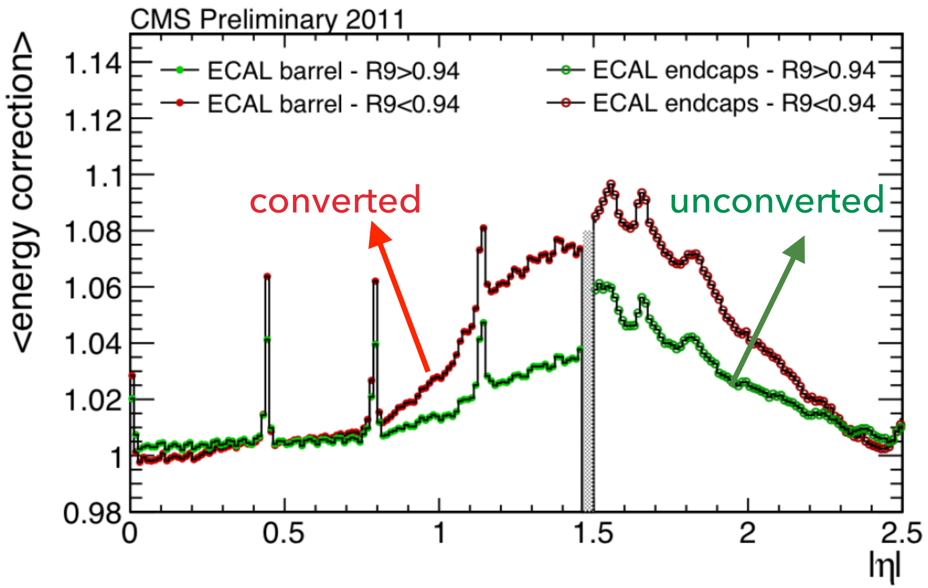
SHOWER SHAPE VARIABLE

$$\sigma_{i\eta i\eta} = \sqrt{\left(\frac{\Sigma_i^{5\times5} w_i\left(\eta_i - \bar{\eta}_{5\times5}\right)^2}{\Sigma_i^{5\times5} w_i}\right)}$$

- $\sigma_{\iota\eta\iota\eta}$:
  - wi = max(0,4.7 + ln(Ei/E5x5))
  - This is essentially the noise cut and tells that each crystal needs to have at least 0.9% of E5x5
  - Decreasing 4.7 to some other value will essentially tighten the noise cut
  - It is ιηιη and not ηη:
    - ιηιη essentially indicates that this distance in eta is estimated in crystal units
    - Using ηη makes it broader near the cracks (since the difference is eta is then larger)
  - **Quiz: why is mean of $\sigma_{\iota\eta\iota\eta}$ in EE is ~3 times higher than that in EB - the molier radius is the same?**
- R9 : E3x3/Esc (raw) - very important variable to tell which photons converted and which are unconverted
  - Used widely in Higgs analysis to categorize the higher sensitivity events from lower ones (S/sqrt(B))
  - An unconverted photon is expected to deposit ~94% of its energy in 3x3 array of crystals - high R9.
  - A converted photon would have lower energy in E3x3 and hence low R9

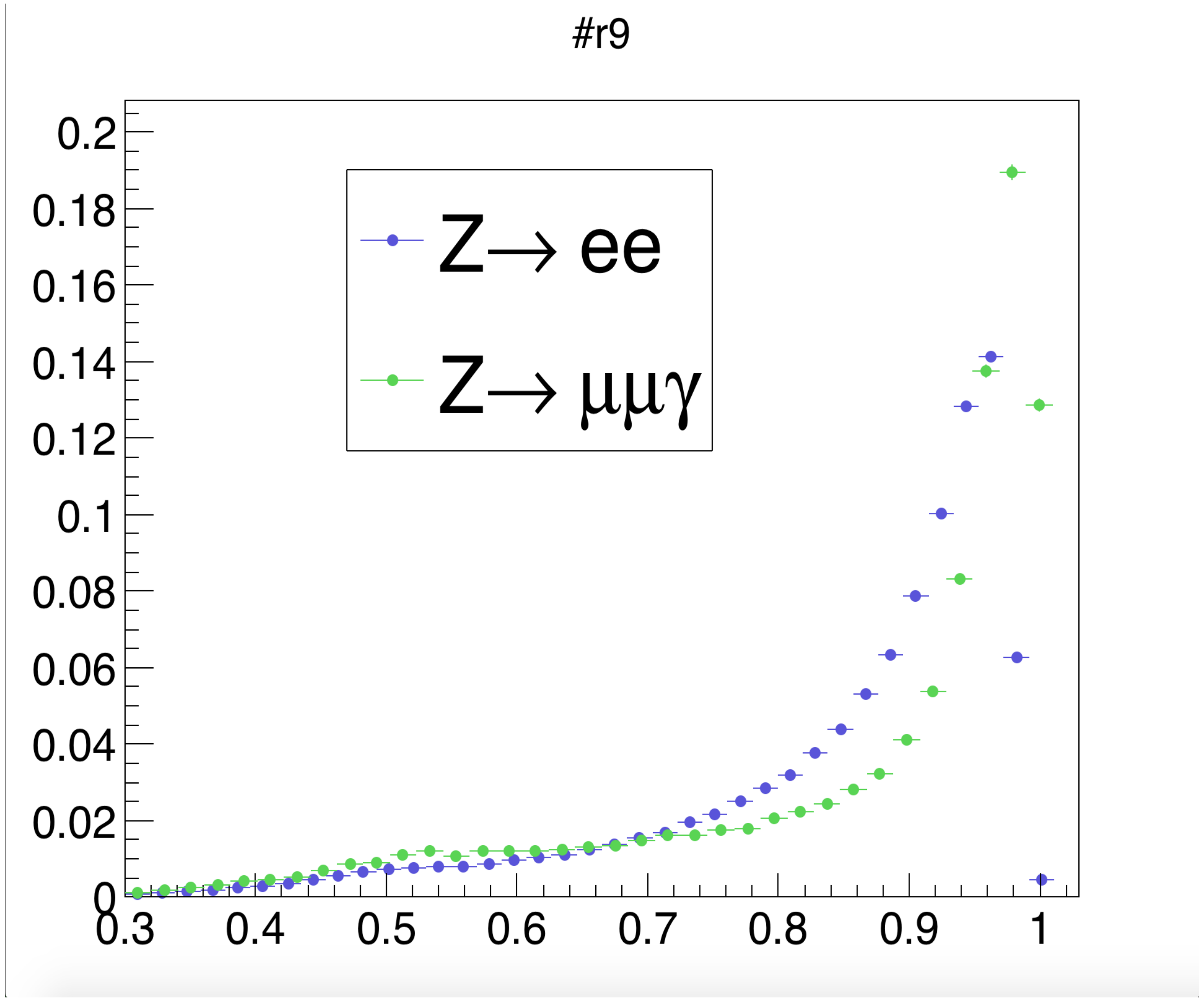


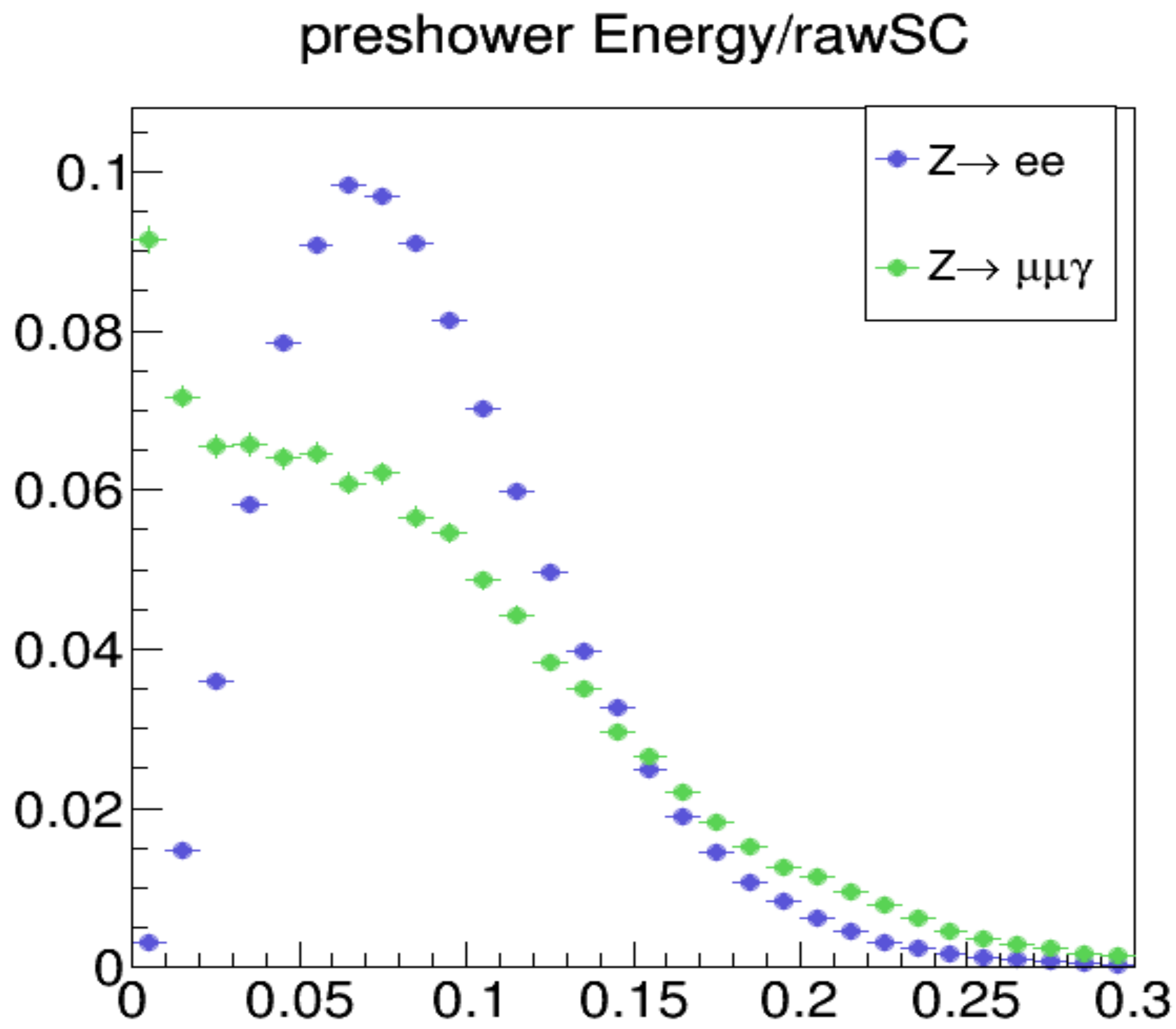




INTERESTING FACTS ABOUT ELECTRON AND PHOTON RECONSTRUCTION – II

- Difference between electrons and photons
- **Quiz: Why do we have such a difference in the shower shapes**


#r9
Z→ ee
Z→ μμγ


preshower Energy/rawSC
Z→ ee
Z→ μμγ