Optimization of Reconstruction Algorithm for BeamCal (ILC)

Lucia Bortko, DESY
on behalf of FCAL-collaboration

Group Meeting | DESY - Zeuthen | 12 May 2014

Search parameters for reconstruction Algorithm

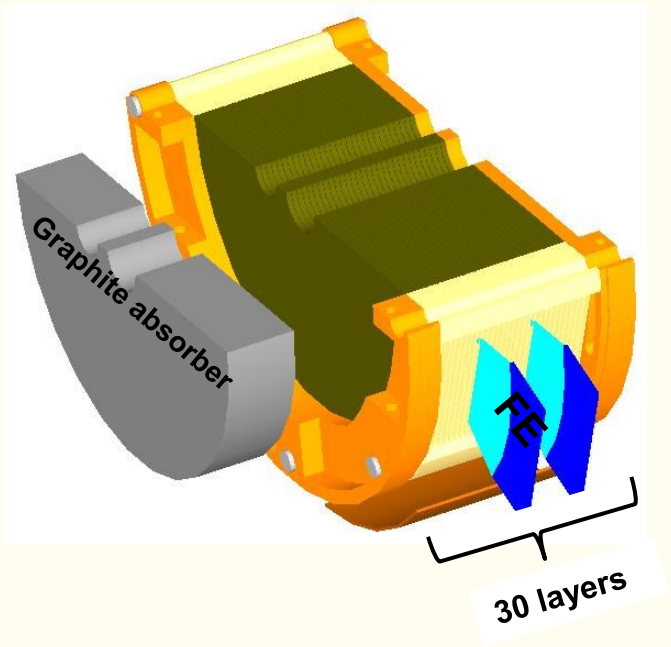
The goal: find optimal parameters of reconstruction algorithm

In my hands: deposited energy in each cell of calorimeter from shower and RMS of background(BG)

Parameters to apply:
- how many sigma(RMS) to apply
- which layers should be considered
- how many cells in a row

Requirements:
- fake rate < 2% (strictly!)
- increase: --efficiency of reconstruction
  --energy resolution
  --spatial resolution

Beam Calorimeter for ILC



- Tungsten absorber
- Diamond sensor
- Readout plane/air gap

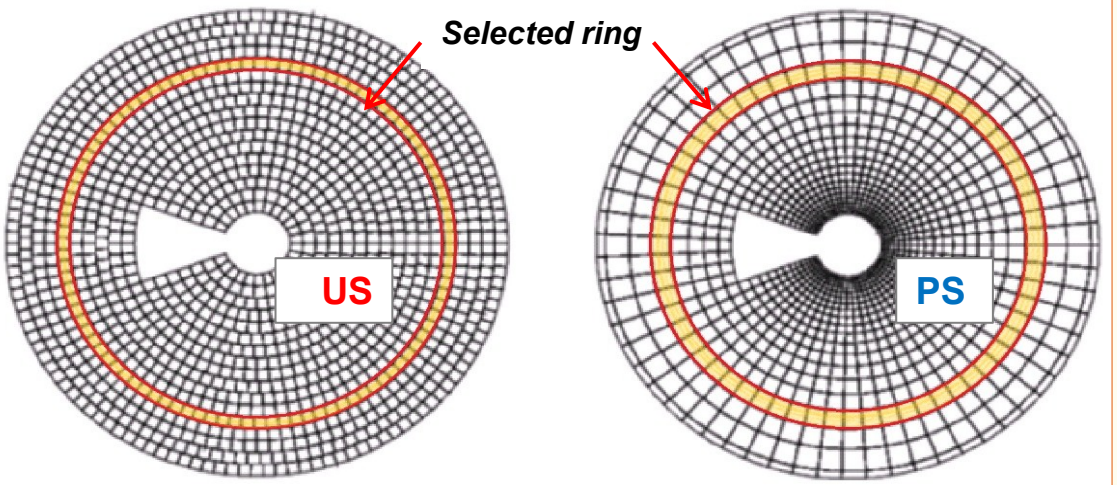
} $1\ X_0$

Size of pads:

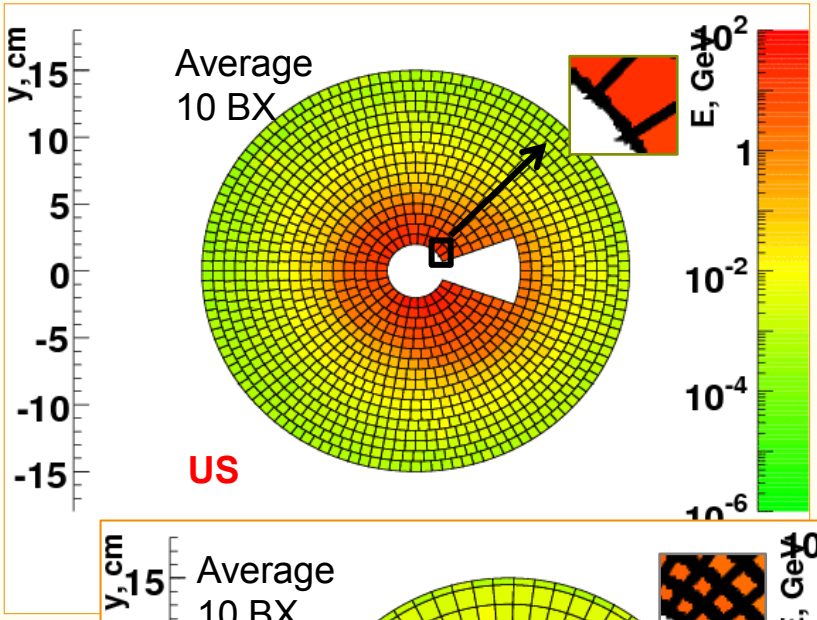
- US - similar, ≈7.65x7.65 mm
- PS - ~R, min 2.2x2.2, max 14.4x14.4mm

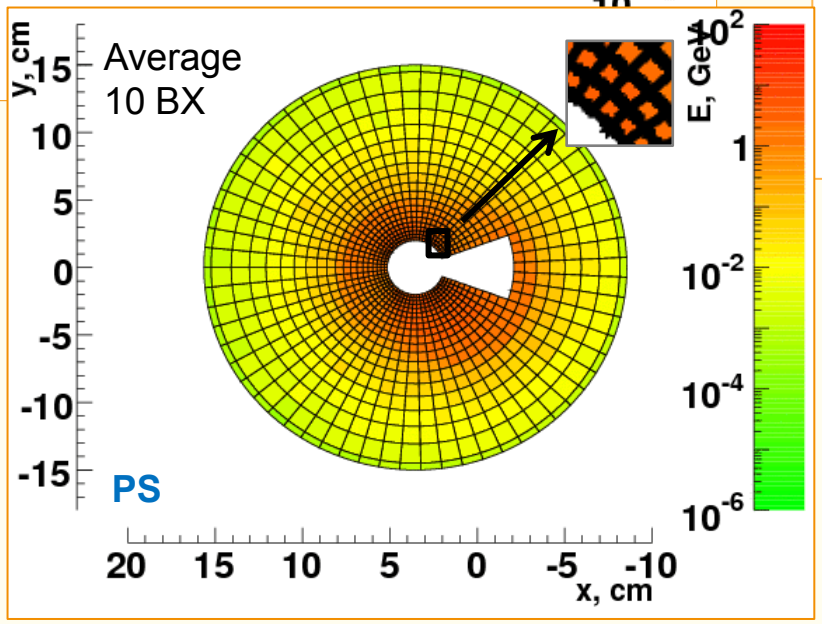
Number of pads:

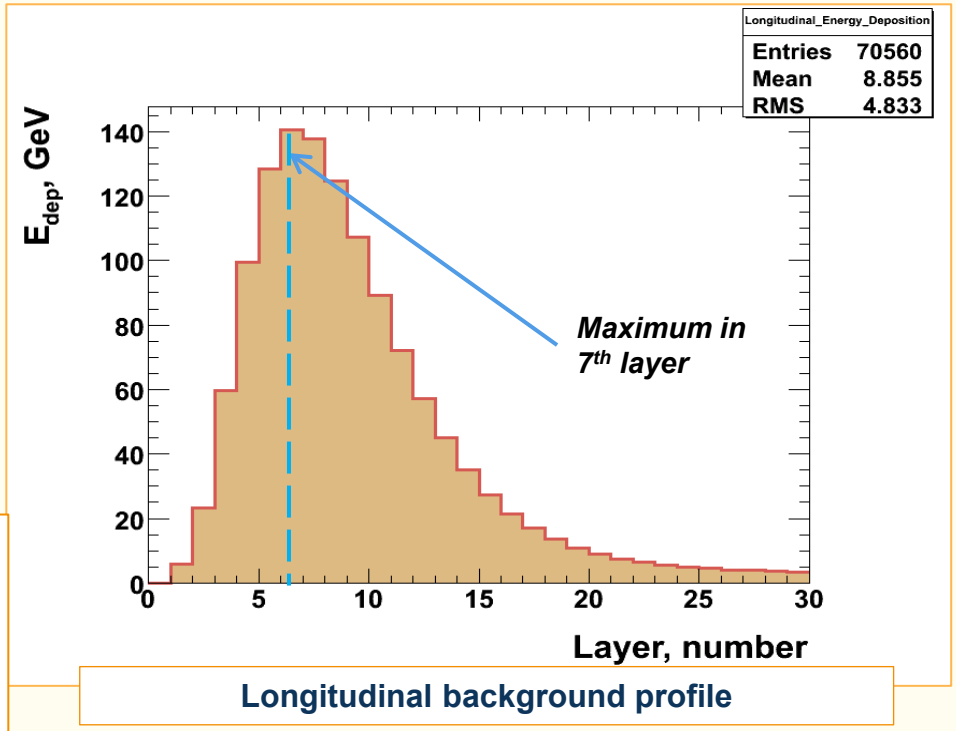
- For US and PS similar



Energy Deposition due to Beamstrahlung





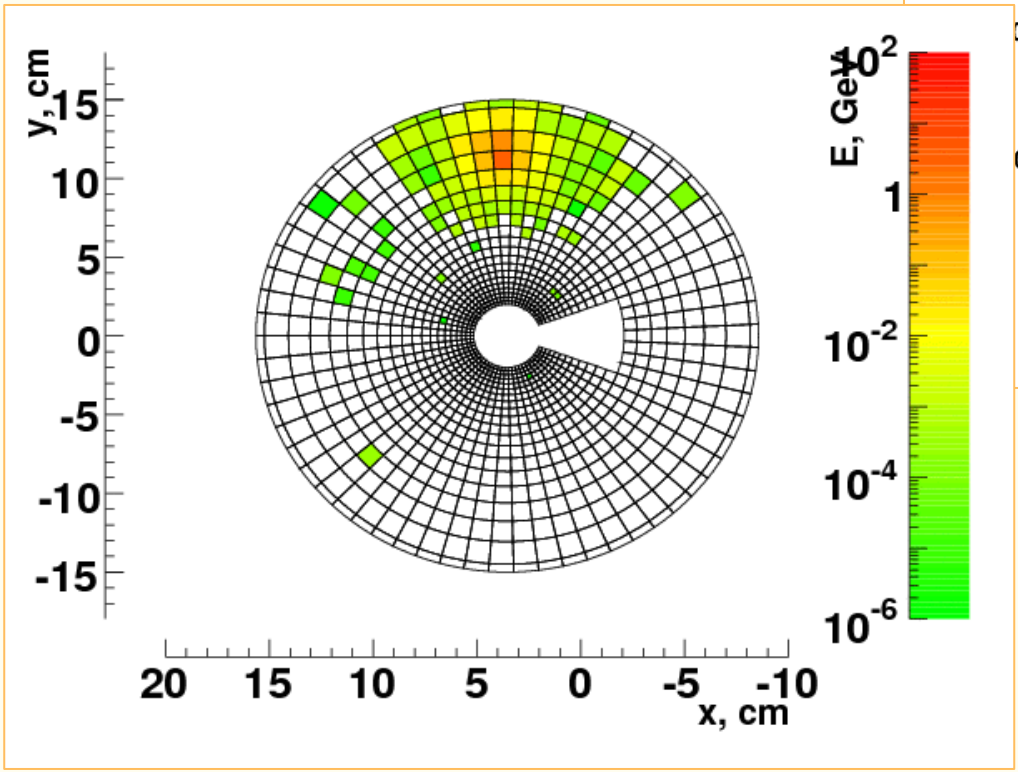


Longitudinal background profile

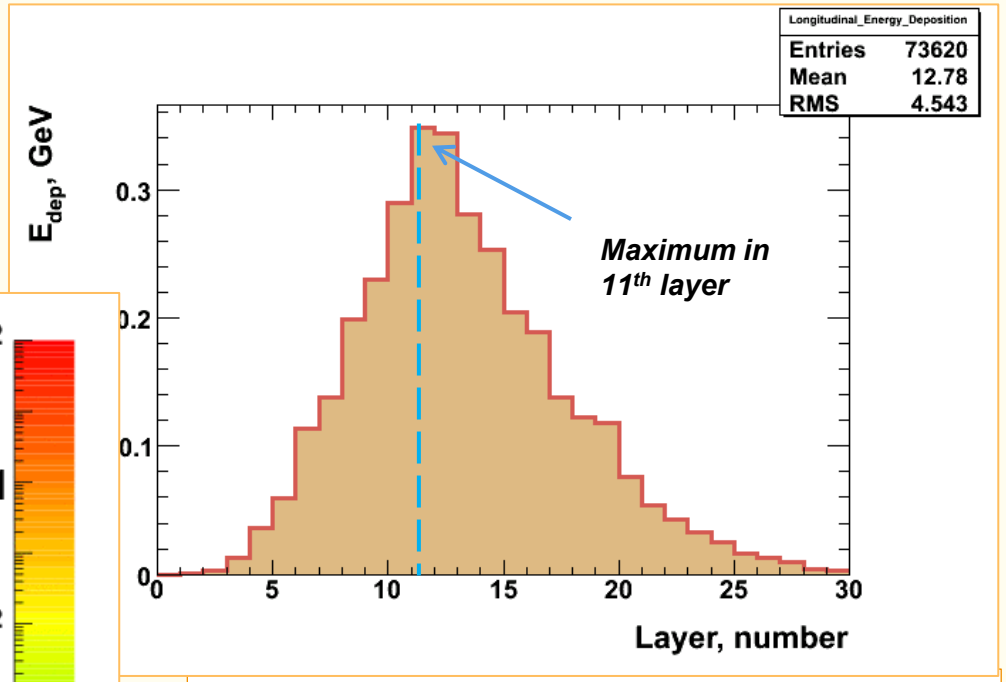
- Beamstrahlung (BS) pairs generated with Guinea Pig
- Energy deposition in sensors from BS simulated with BeCaS (Geant4) ⟶ *considered as background (BG)*

Shower from Single High Energy Electron

Example for 200 GeV electron:



Shower from 200- GeV electron



Longitudinal shower profile

Electron energy	Shower maximum
10 GeV	7 layer
20 GeV	8 layer
50 GeV	9 layer
100 GeV	10 layer

Processing tasks

Access to NAF2 LC group - to get more BG files

- On Friday I finally got access not only to NAF2-ILC, but also to directory of Aura Rosca
- On the path she gave me there were not that files, which I was searching for.

Batch system

- Recently the situation changed. Now it is allowed to send jobs only to SL6.
  There is no ROOT, no GEANT,..
- Therefore I will use batch system of ILC group in Hamburg. (Getting access to this in progress)
  Meanwhile I was using only my computer last weeks => took some time for simulations..

Does BeCaS work properly?

Why the files of showers are not creating (not simulating) sometimes?

- The problem was that the jobs sent to the farm sometimes were not finished.
  Now, working on my computer, I got all the showers all the time. => not an error of BeCaS

Why sometimes in simulated file there is no shower (only few cells of calorimeter have energy deposition)?

- It was events, that have coordinates of outer radii
- Did not find yet the size of LumiCal in BeCaS
- I checked by hands, that the “border” is on 13,64 cm (the whole calorimeter - 15 cm)

=> There are no problems with BeCaS

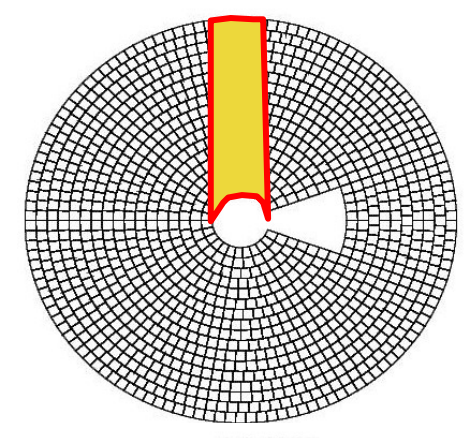
Simulation Showers

WAS:

- Rectangle area (therefore diff amount of pads on diff radii)
- Distribution: for US distribution of hittings was uniform

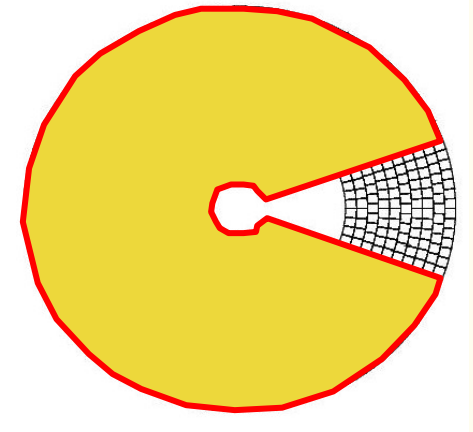
  for PS – density of hittings ~ $\frac{1}{r}$ inversly proportional to the r

  reason: to have similar amount events in each r (for SNR,..)



CHANGED:

- Sector area
- Distribution: for both segmentation made both distributions



Which distribution where to apply?

Proportional distribution (PD), close to reality

- efficiency calculating - need more events on small radii
- spatial resolution for PS(?) – need more statistics on small radii

Uniform distribution (UD)

- energy resolution (?) – need resolution on particular coordinate, shouldn't depend on distribution
- spatial resolution for US(?)

Algorithm

1. SH + BG – average_BG
2. Layers from 5 to 20
3. Energy threshold 5 RMS
4. Combine to towers
5. Search Max energetic tower
   * if there ≥ 13 cells (not necessarily sequent), search for neighbor towers
     * if in neighbor ≥ 9 cells & at least 1 neighbor

   => shower defined
   * Consider candidate towers to shower within Rm=1.2 cm or at least 8 pads around max energetic tower

   => shower created
6. Next shower: repeat step 5
7. For each shower calculate
   - R COG and Phi COG
   - Energy

Efficiency

1. Reconstruction showers on top of BG -> Number of ring rReco and phiReco
2. Reconstruct showers, no threshold applied (0*RMS,
   cause not all SH on small radii reconstructing) -> rTrue, phiTrue
3. If | rTrue - rReco| < Rm and |phiTrue – phiReco| < Rm,
   then shower reconstructed correctly
   and ratio rReco/rTrue = efficiency
4. Else (| Rtrue- Rreco| > Rm) - fake shower

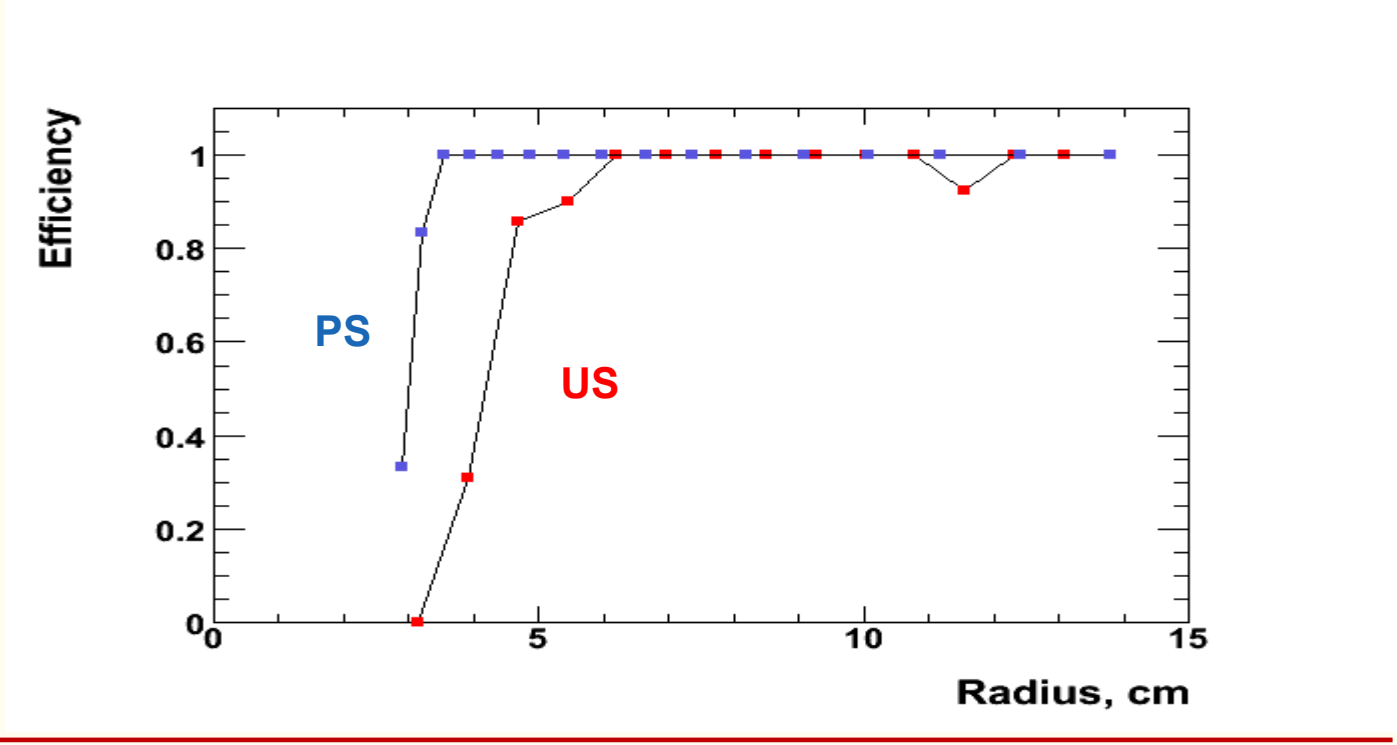
Efficiency 500 GeV

PS

TOTAL EVENTS=200
NUMBER NOT CREARTED SHOWERS=12
EQUAL_EVENTS=183
EQUAL WITHIN R_ MOLIERE EVENTS=0
NUMBER FAKE SHOWERS=2
NUMBER NOT RECONSTRUCTED EVENTS=3

US

TOTAL EVENTS=194
NUMBER NOT CREARTED SHOWERS=11
EQUAL_EVENTS=155
EQUAL WITHIN R_ MOLIERE EVENTS=0
NUMBER FAKE SHOWERS=2
NUMBER NOT RECONSTRUCTED EVENTS=26



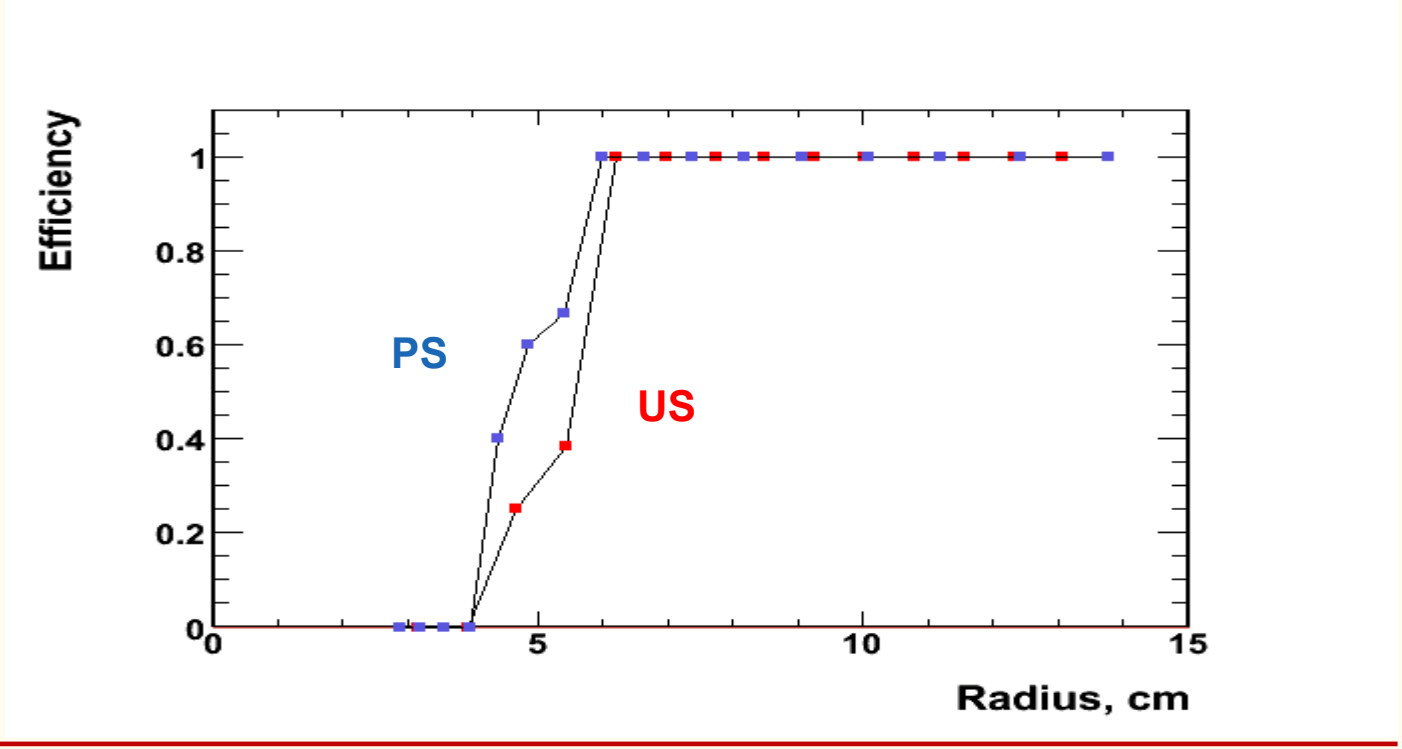
Efficiency 200 GeV

PS

TOTAL EVENTS=199
NUMBER NOT CREARTED SHOWERS=10
EQUAL_EVENTS=152
EQUAL WITHIN R_ MOLIERE EVENTS=0
NUMBER FAKE SHOWERS=2
NUMBER NOT RECONSTRUCTED EVENTS=35

US

TOTAL EVENTS=199
NUMBER NOT CREARTED SHOWERS=9
EQUAL_EVENTS=144
EQUAL WITHIN R_ MOLIERE EVENTS=1
NUMBER FAKE SHOWERS=1
NUMBER NOT RECONSTRUCTED EVENTS=44



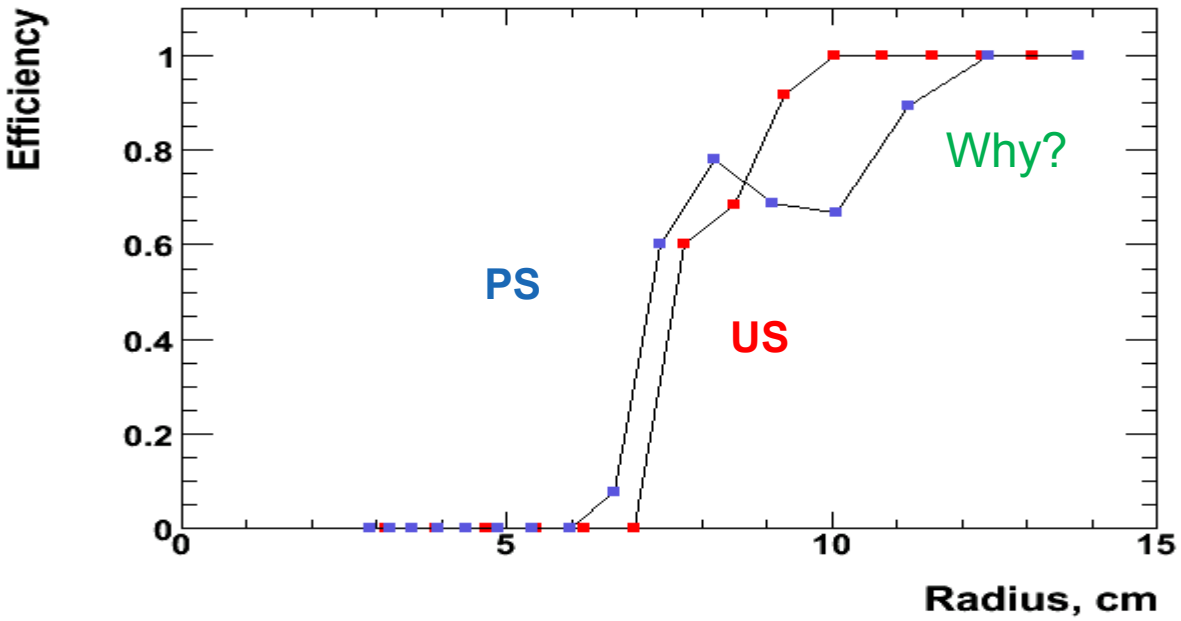
Efficiency 50 GeV

PS

TOTAL EVENTS=200
NUMBER NOT CREARTED SHOWERS=15
EQUAL_EVENTS=85
EQUAL WITHIN R_ MOLIERE EVENTS=0
NUMBER FAKE SHOWERS=0
NUMBER NOT RECONSTRUCTED EVENTS=100

US

TOTAL EVENTS=196
NUMBER NOT CREARTED SHOWERS=14
EQUAL_EVENTS=88
EQUAL WITHIN R_ MOLIERE EVENTS=0
NUMBER FAKE SHOWERS=3
NUMBER NOT RECONSTRUCTED EVENTS=91



Efficiency 500 GeV NEW

PS

TOTAL EVENTS=200

NUMBER RECONSTRUCTED WITHIN R_MOLIERE SHOWERS = 89.3%

NUMBER FAKE SHOWERS = 0.5%

NUMBER NOT RECONSTRUCTED SHOWERS = 10.2%

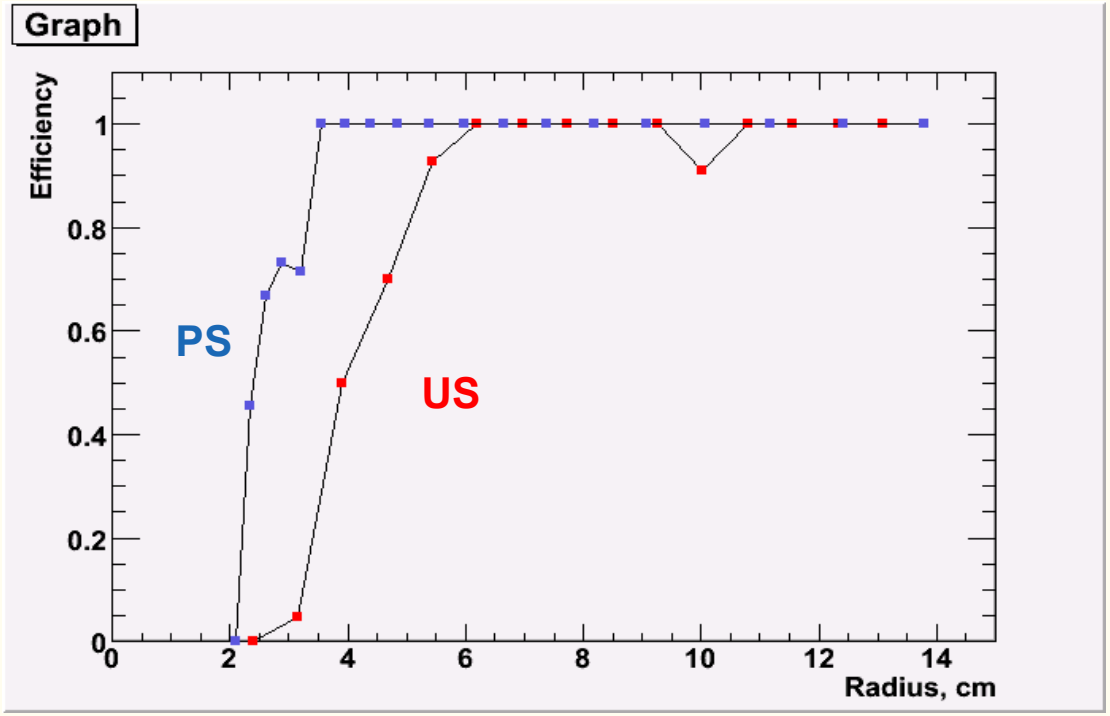
US

TOTAL EVENTS=200

NUMBER RECONSTRUCTED WITHIN R_MOLIERE SHOWERS = 61.7%

NUMBER FAKE SHOWERS = 0.0%

NUMBER NOT RECONSTRUCTED SHOWERS = 38.3%



Efficiency 200 GeV NEW

PS

TOTAL EVENTS=200

NUMBER RECONSTRUCTED WITHIN R_MOLIERE SHOWERS = 55.6%

NUMBER FAKE SHOWERS = 0.0%

NUMBER NOT RECONSTRUCTED SHOWERS = 44.4%

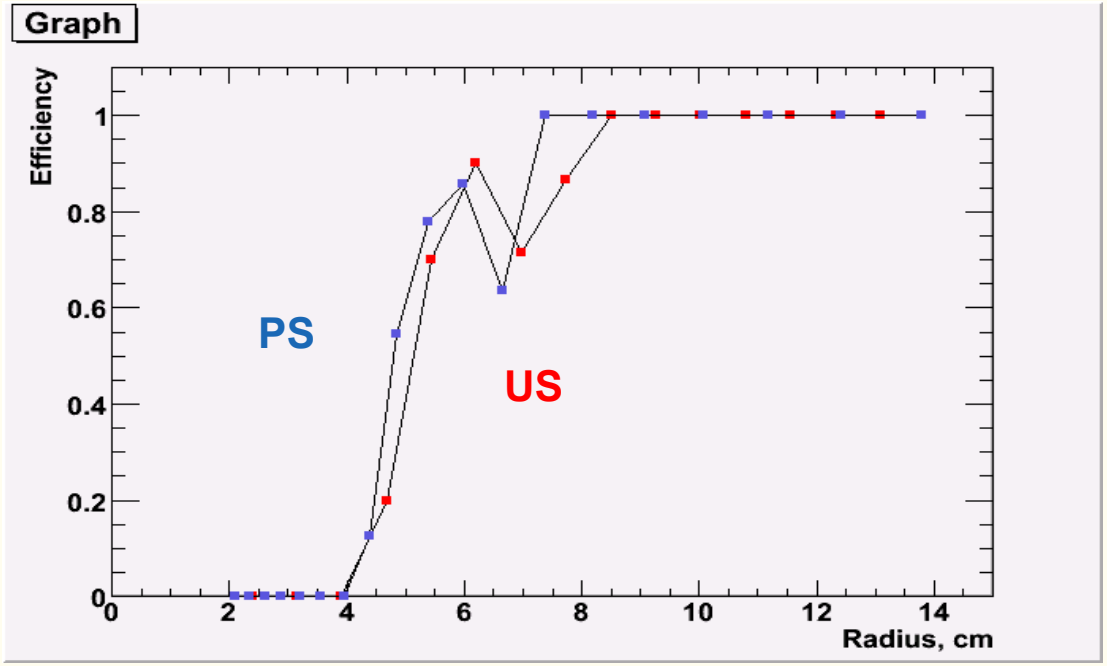
US

TOTAL EVENTS=200

NUMBER RECONSTRUCTED WITHIN R_MOLIERE SHOWERS = 52.0%

NUMBER FAKE SHOWERS = 2.5%

NUMBER NOT RECONSTRUCTED SHOWERS = 45.5%



Efficiency 50 GeV NEW

PS

TOTAL EVENTS=200

NUMBER RECONSTRUCTED WITHIN R_MOLIERE SHOWERS = 25.6%

NUMBER FAKE SHOWERS = 1.0%

NUMBER NOT RECONSTRUCTED SHOWERS = 73.3%

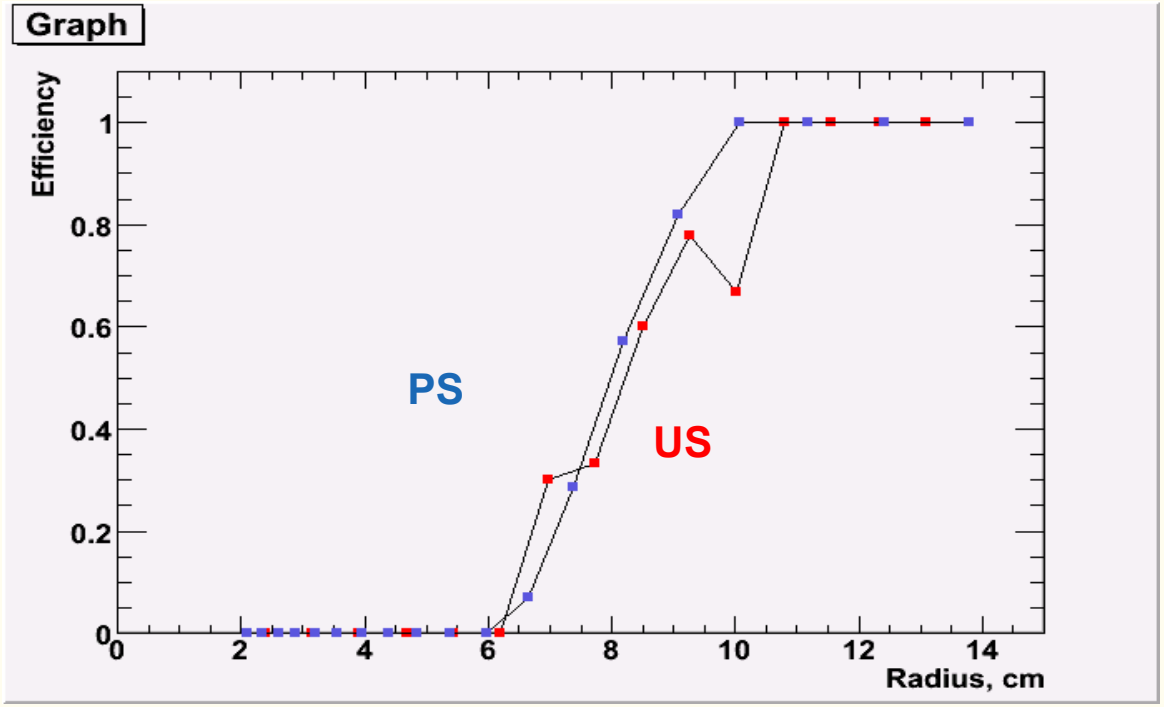
US

TOTAL EVENTS=200

NUMBER RECONSTRUCTED WITHIN R_MOLIERE SHOWERS = 23.0%

NUMBER FAKE SHOWERS = 2.6%

NUMBER NOT RECONSTRUCTED SHOWERS = 74.5%



Next things to prepare to FCAL talk

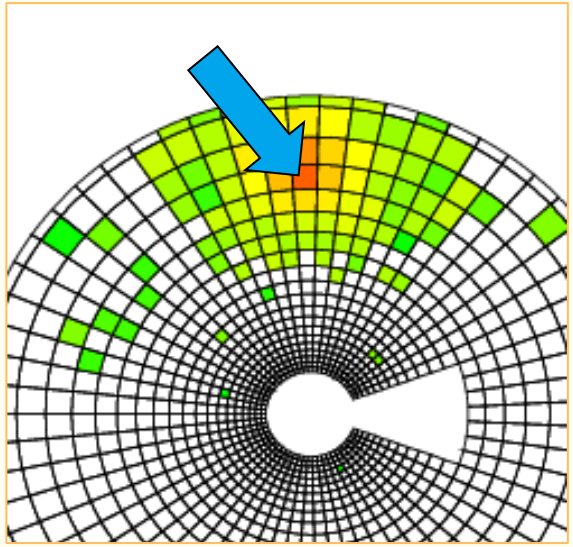
- Get more statistics for efficiency (500 events) and make error bars
- Get more BG files, simulate, get fake rate. If with current algorithm it is less then 2%, calculate energy resolution and spatial resolution (the programs for this more or less prepared, need just make some modifications)

Back up

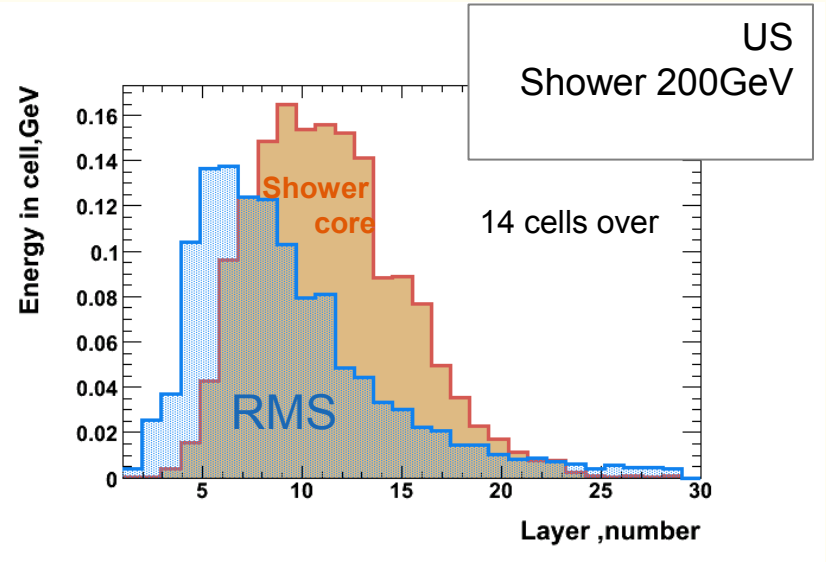
Idea

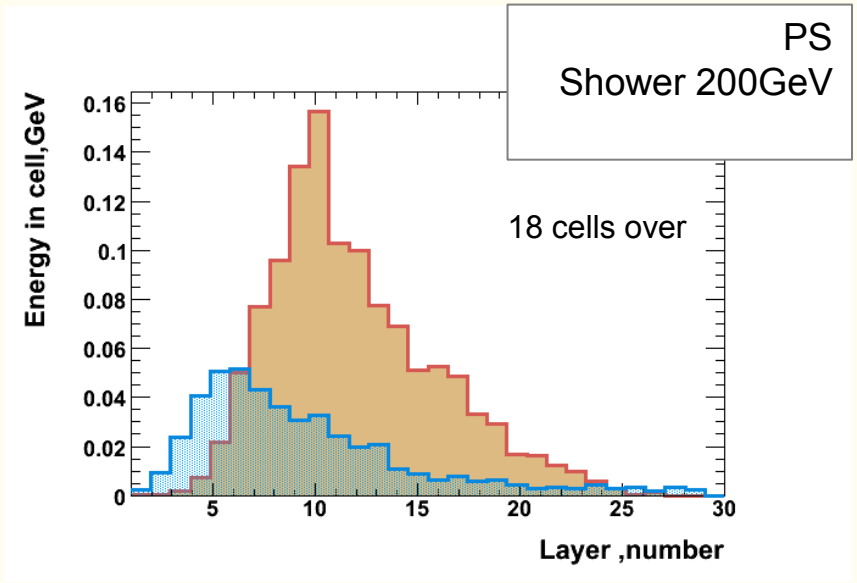
Compare energy deposition on small radii (most problematic area for reconstruction) along Z-axis for:

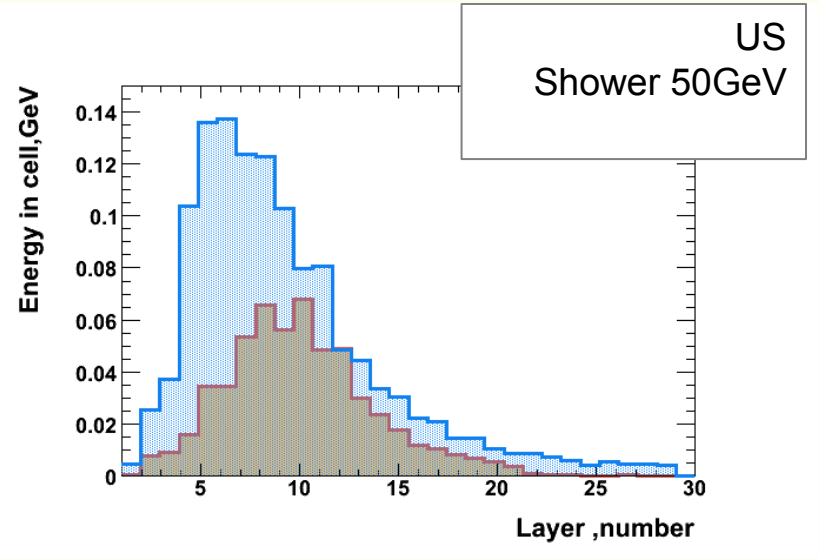
- tower of the shower core and tower of the RMS on small radii
- max energetic tower of (BG – average_BG) and
  tower of the RMS

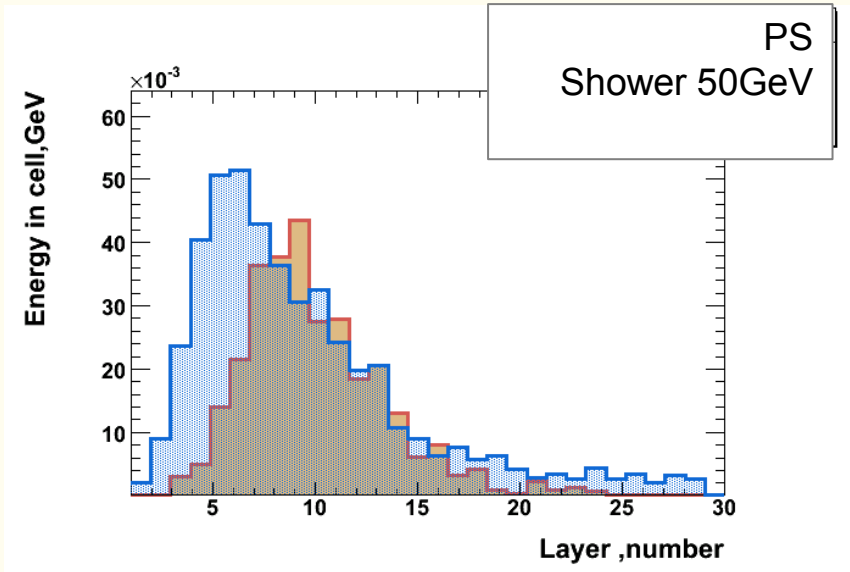


Tower profiles from Shower core and RMS on small R

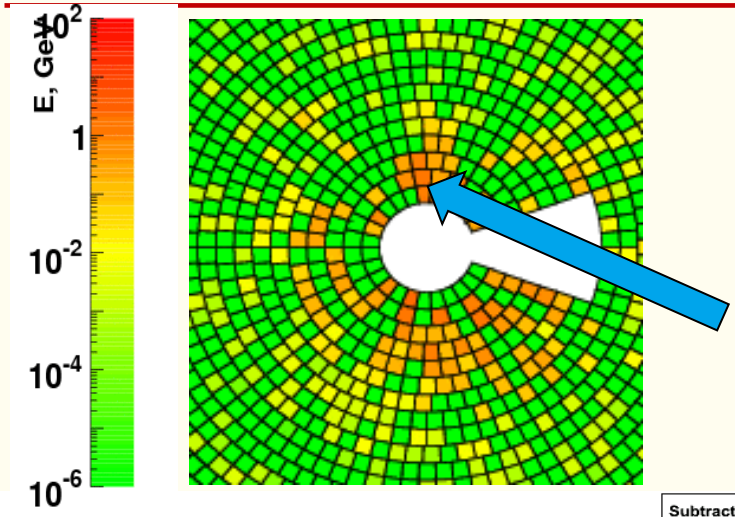




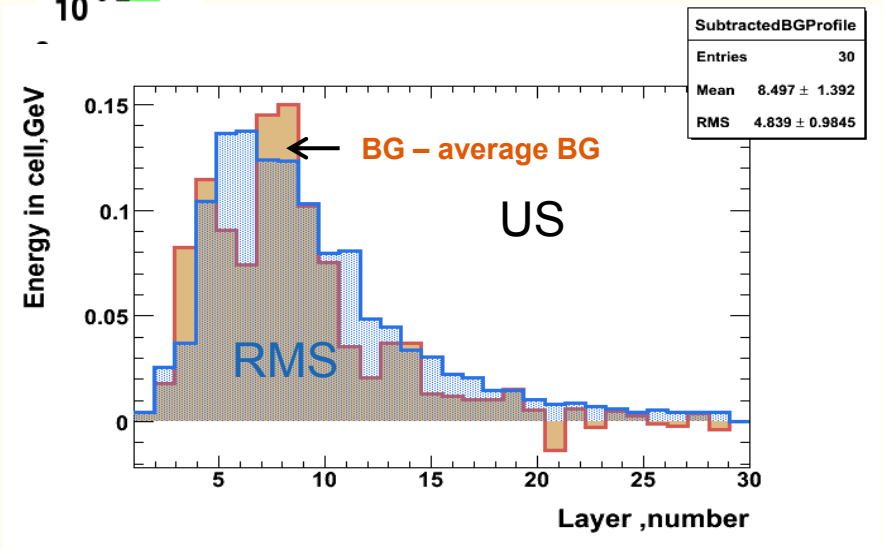


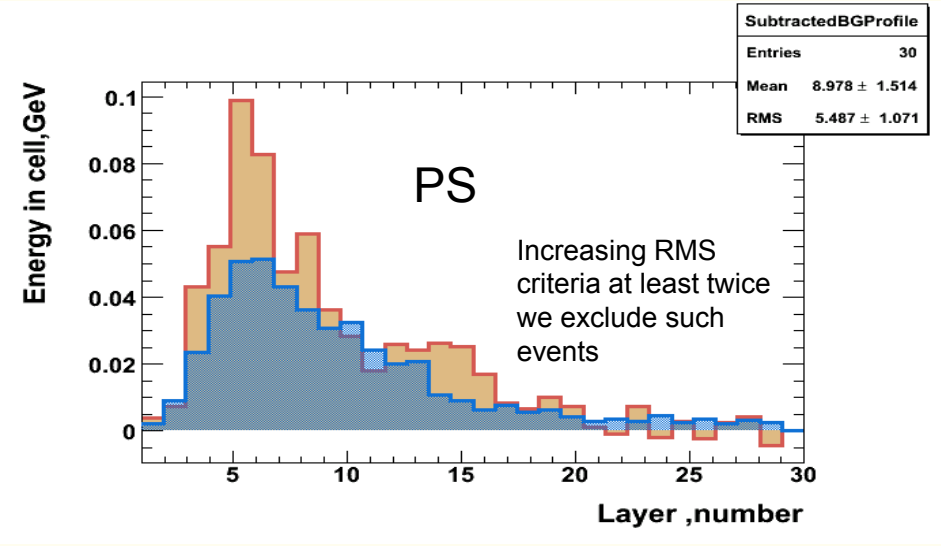


Tower profiles from Subtracted BG and RMS on small R



The tower with maximal energy deposition was chosen(the worst case)





But for showers(previous slide) we still have possibility to reconstruct, especially going further with radius

Choosing parameters. Fake Rate.

Source	Difference in conditions	Layers to be considered	RMS applyed	Min number of cells in a row	
				In SH max	In neighbor
Max SH Tower and RMS along Z comparison (previous slides)	**1 Tev**	**5-20** (25?)	**>2 RMS** (chosen 5 RMS)	**13**	**9**
Thesis of Katharina Kuznetsova, 2006	**500GeV ,** diff size of pads, type of segmentation - US	**4-17**	**3 RMS**	**10**	**6**
FCAL Paper, 2004	**500 GeV**	**2-20**	**5 RMS**	**9**	**6**

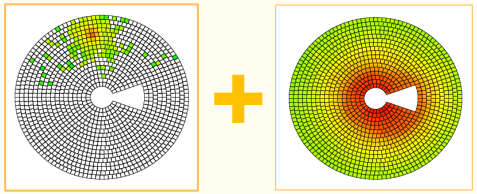
Checking fake rate (100 files were used)

	Layers to be considered	RMS applyed	Min number of cells in a row		Fake rate	
			SH max	Neighbor	US	PS
Case 1 (suitable)	**5-20**	**5 RMS**	**13**	**9**	**2 %**	**0 %**
Case 2 (relaxed)	**5-20**	**5 RMS**	**10**	**6**	**3%**	**3%**

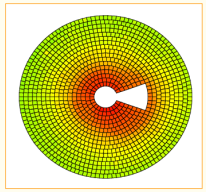
Old Algorithm

with BG

1. SH + BG

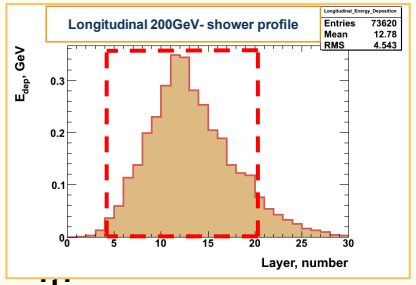


2. - average BG by 10th previous BX

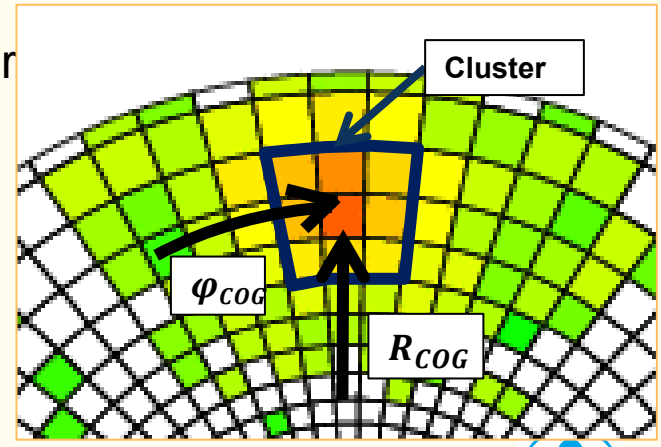


wo BG

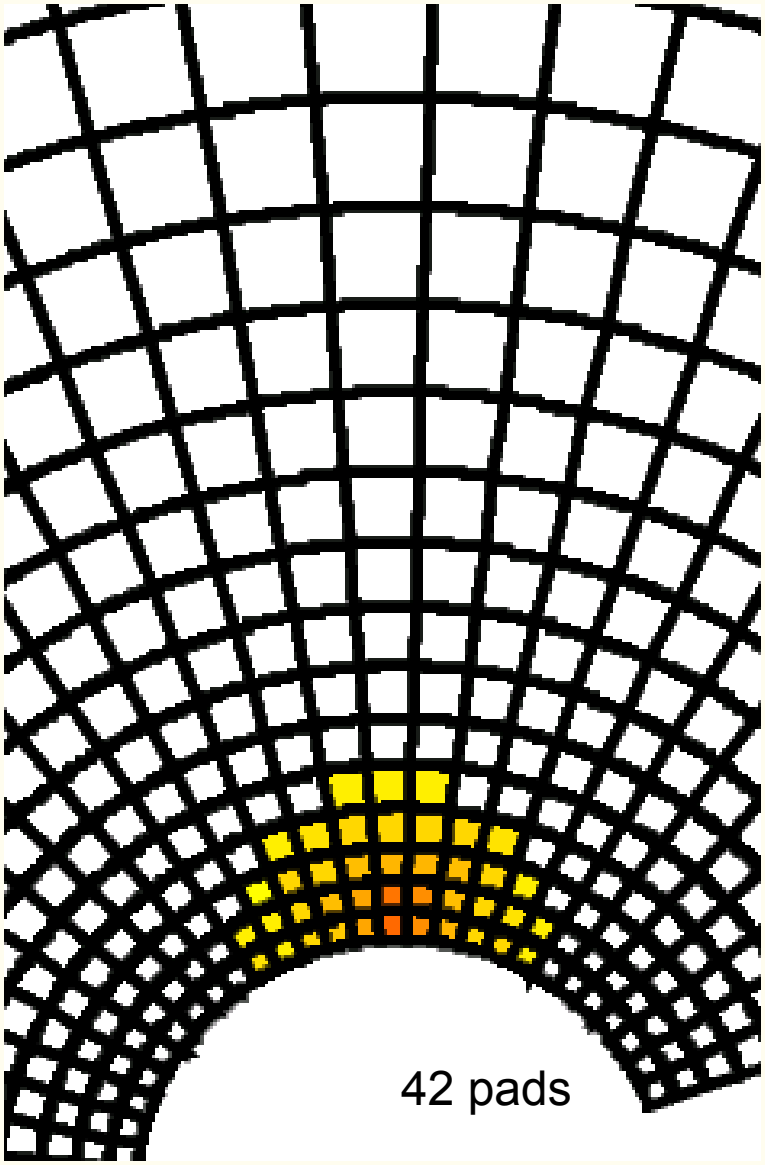
3. Select layers from 5th to 30th. Search for towers contains at least 10 sequent pads with Edep>0 along Z axis.



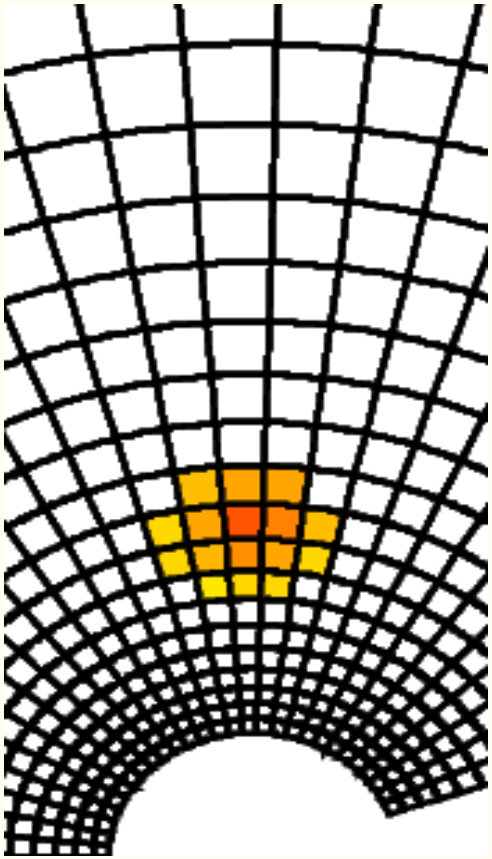
4. Searching in that towers tower with maximum energy deposition
5. Look on to 8 neighbor towers around that tower
6. Get output: $\mathbf{R_{COG}}$, $\boldsymbol{\varphi_{COG}}$, $\mathbf{E_{clu}}$



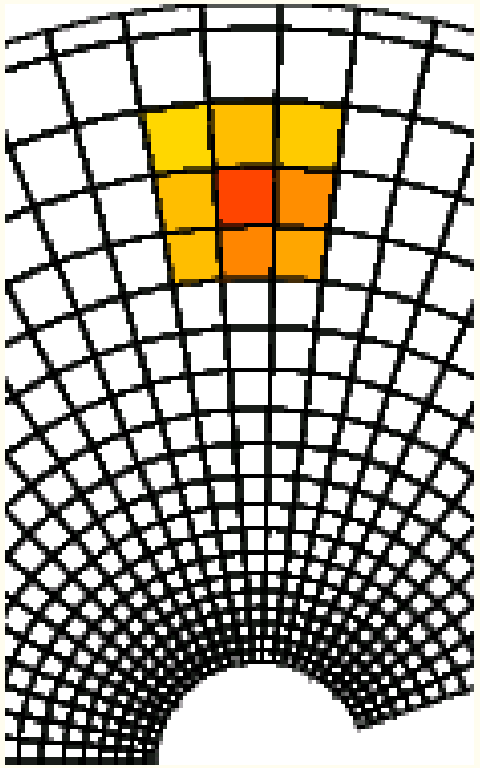
Moliere Radius



42 pads

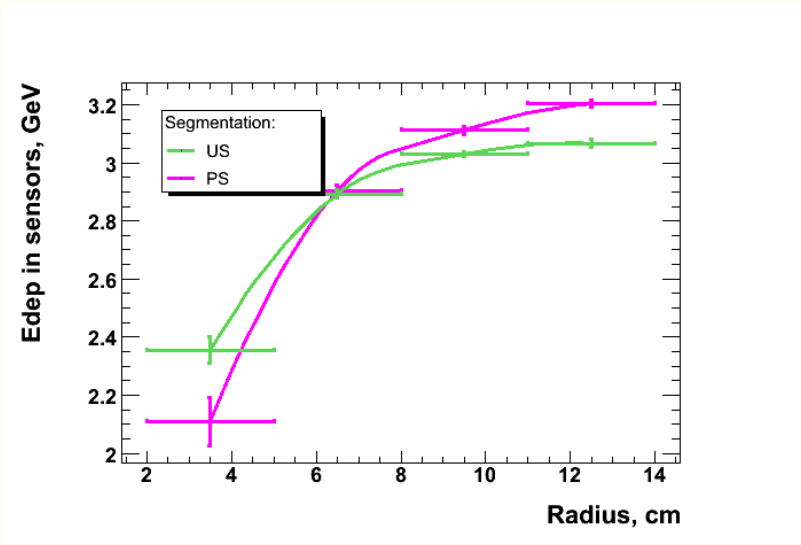


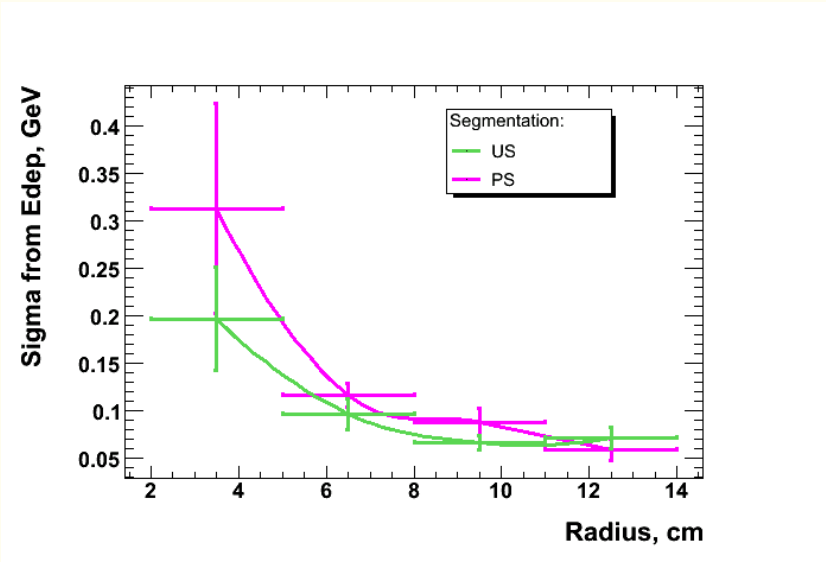
16 pads



8 pads

Energy resolution vs Radius





For 200Gev