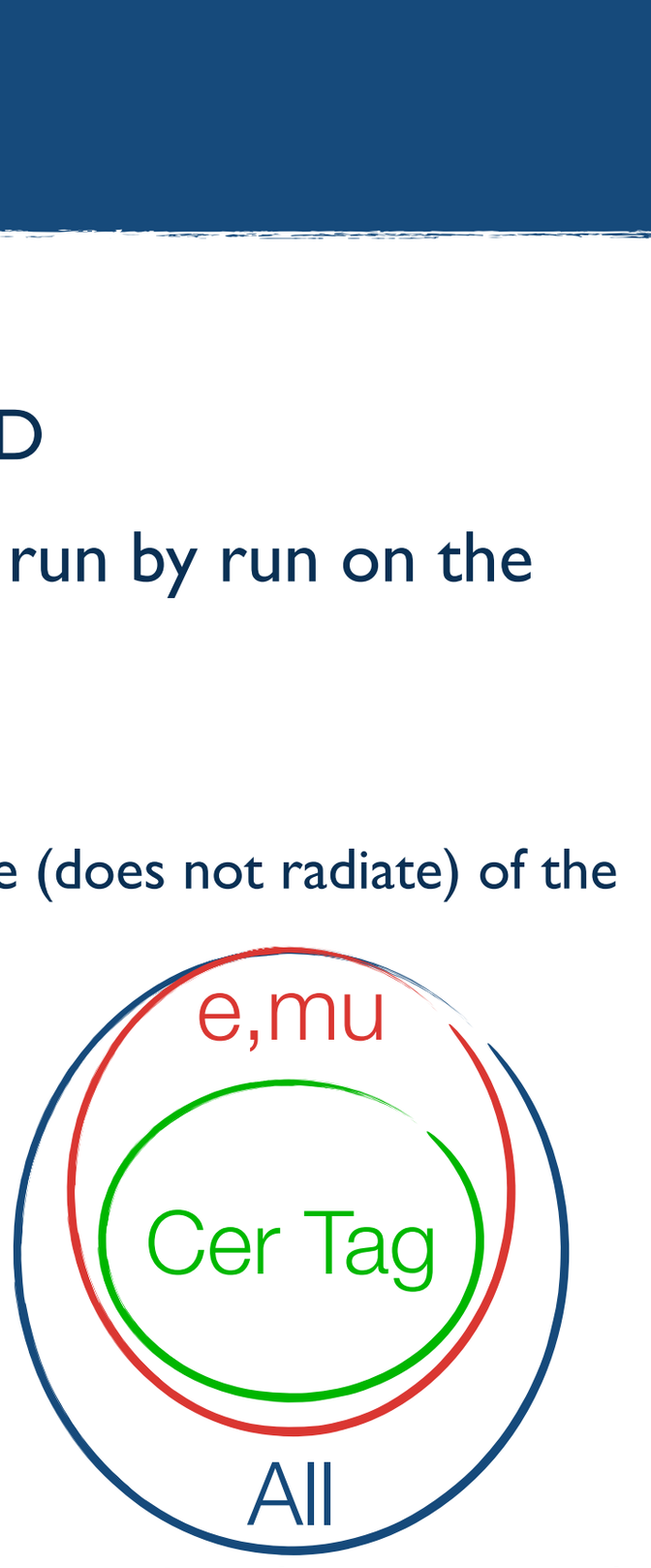
# Cherenkov data for SDHCAL - MST

Antoine Pingault
22/02/2017

# Cerenkov Data

- **Since December 2014**
  - Lots of run have been using cerenkov data for PID
  - Not always tagging same particles, have to check run by run on the elog to be sure.
  - Simple threshold detector were used:
    - Mass-dependent : Discriminate lighter and heavier particle (does not radiate) of the same energy/momentum.
    - => If tagging Pions, will also tag muons and electron
- Take it as inclusive tag:
  - All Particles should be correctly tagged
  - But not all particles are tagged
    - e.g case of tagging muons



# Cerenkov Data

- **In June 2015 (PS data), and both TB in 2016 we used 2 Cerenkovs**
  - Can select a given particle:
    - If Cer1 is set to tag Pions (and mu,e)
    - Cer 2 is set to tag muons (and e)
    - Data from (Cer2 - Cer1) will yield a 'pure' pion sample

- **From the data:**
- Cerenkov tag efficiency is ~ 60-70% for e, muons tagging:
  - Roughly was we see in TB (when setting a 'pure' muons beams)
  - I used "TrackCosTheta>0.999 && TrackLength>1150 && MaxRadius<20 && Nhit/Nlayer <3" for muon selection - Loosening/tightening the cut keep the ratio
- Expect lower efficiency for pi/k tagging
  - Full Cer efficiency not reached for pi/k while setting the threshold below the next particle type tagging

# Cerenkov Data

- **Bug in BIF Firmware for 2015/2016 Data**
  - Only happens when 2 Cerenkovs signals are plugged in
  - Data from one Cer are discarded from streamout
    - -> "Exception: DIF 3 T ? 0 129 Header problem" from streamout
  - Need to adapt DIFUnpacker.cc (done)
  - **Can adapt streamout to re-encode properly**
    - -> Consistent data in the RawCaloHit data across all TB. (few lines of code)

- Will discuss again with Laurent/Christophe for this consistency implementation

# Cerenkov Data

- **For now, encoding is TestBeam specific.**
- E.g. : December 2014
  - BIF_Id =1
  - cerenkov signal span over 7 clock and is delayed by 12 clock
- 2015 :
  - Bif_id =3
    - Signal span over 1-2clock and is in advance by 5-10 clocks depending on the run

- I already have implementation for different marlin processor: Streamout/ Trivent/sdhcal_analysis (Arnaud's algorithms)

# Cerenkov – How to extract the data?

- **From Streamouted data:**
  - Dif_id = 1(2014) or 3
  - Asic_id
    - Cer 1 = 1
    - Cer 2 = 129
  - I = 0…#CerHit in Trigger (first hit has I=0, second I=1 etc.)
  - J = 15
  - Time = 2-15 clock in advance
- If using Trivent, Need to add info in geometry file
  - 2014 -> (m3_bonneteau.xml)  1,57,0,0,1.,1.
  - 2015 ->(m3_bonneteau_avril2015.xml ) 3,57,0,0,1.,1.

# Cerenkov - TB specific

- **2014 Only 1 cerenkov:**
  - DifId = 1
  - Asic Id = 1
  - Threshold = 1
  - I = 0->NCer in trigger
  - J = 15
  - Time = 12 +- 2 clock in advance span over 7 clock

- **May 2015 Only 1 cerenkov:**
  - DifId = 3
  - Asic Id = 1
  - Threshold = 3
  - I = 0->NCer in trigger
  - J = 15
  - Time = 6 +- 2 clock in advance span over 1 clock

# Cerenkov - TB specific

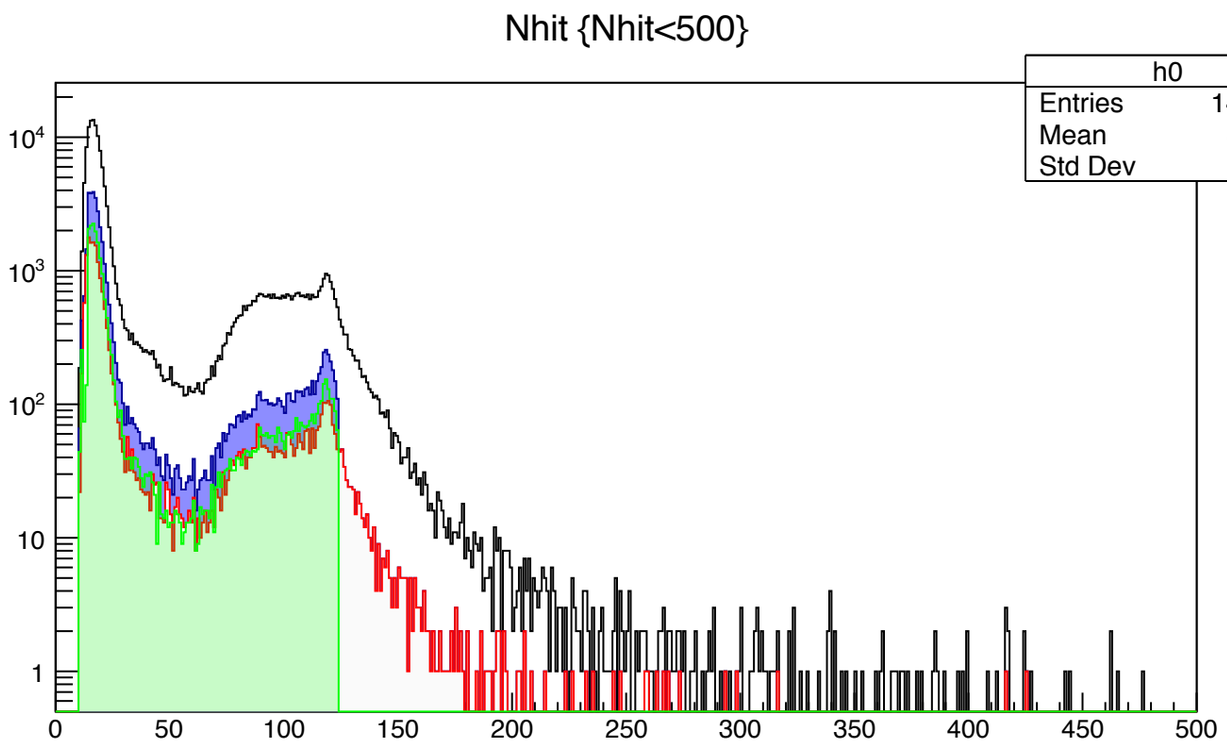
- **June 2015 (PS TestBeam) 2 cerenkovs:**
  - DifId = 3
  - Asic Id = 1(cer1) - 129 (cer2)
  - Threshold = 1(cer1) - 2 (cer2)
  - I = 0->NTotCer in trigger
  - J = 15
  - Time = 6 +- 2 clocks in advance span over 1 clock
- **October 2015 Only 1 cerenkov but 2 signals for some run (check ELOG):**
  - First is Cherenkov signal delayed by 10clocks and span 1 clock
  - Second is Cherenkov signal in coincidence with a scintillator on the beam line. 5 Clocks delay , span 1-2clock(200ns signal)
  - DifId = 3
  - Asic Id = 1 (Cer1) - 129(Cer2)
  - Threshold = 1 and 2 (both Cer)
  - I = 0->NCer in trigger
  - J = 15

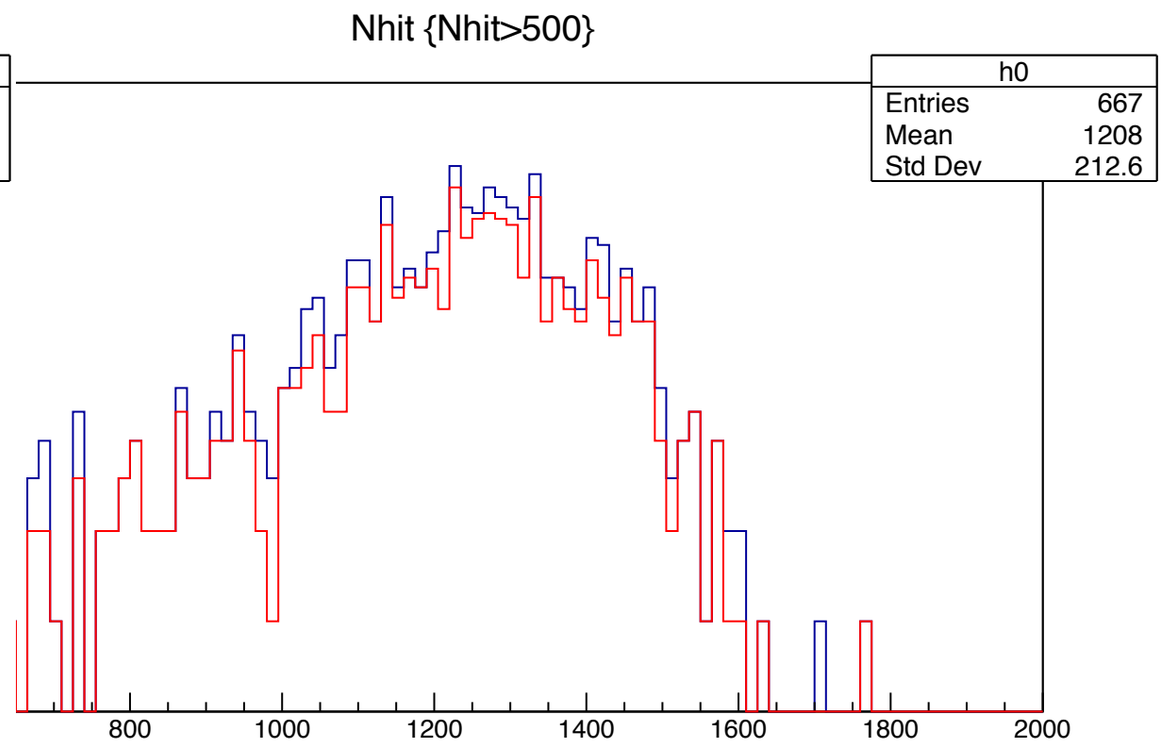
# Cerenkov - TB specific

- **2016 (both TB ) 2 cerenkovs:**
  - DifId = 3
  - Asic Id = 1(cer1) - 129 (cer2)
  - Threshold = 1-2 (both)
  - I = 0->NTotCer in trigger
  - J = 15
  - Time = 6 +- 2 clocks in advance span over 1 clock

# Cherenkov Data

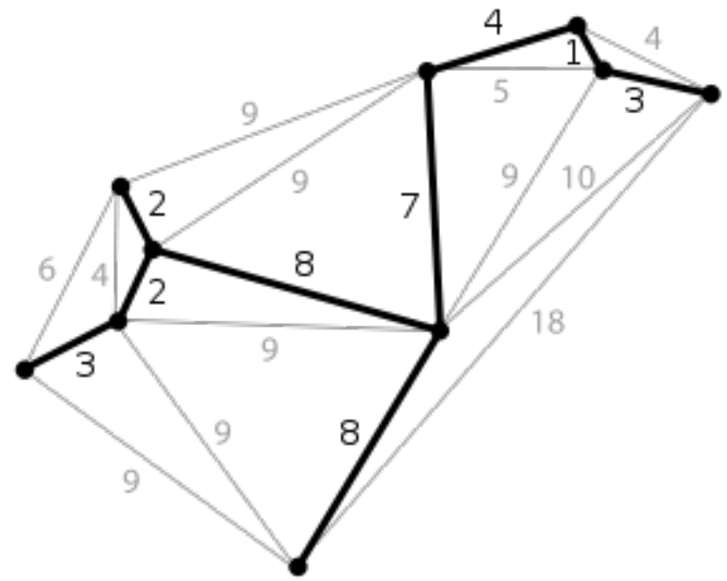
- **In June 2016 we had a CEDAR**
  - Ability to tag with good accuracy a given particle
  - Confused about the results though:
    - E.g. : On run 732817 - 70GeV Pi+
    - Cerenkov tagging all except protons | CEDAR tagging pions
      - left: green is CEDAR, red is Cerenkov, blue is both
    - Right, blue is no selection, red is CEDAR ( Cerenkov is the 'same')



# Minimum Spanning Tree ?

- Describe a set of point : **Nodes**
- Links between nodes : **Edges**
- Each edges is weighted
  - Basic weight can be the euclidian distance between the nodes



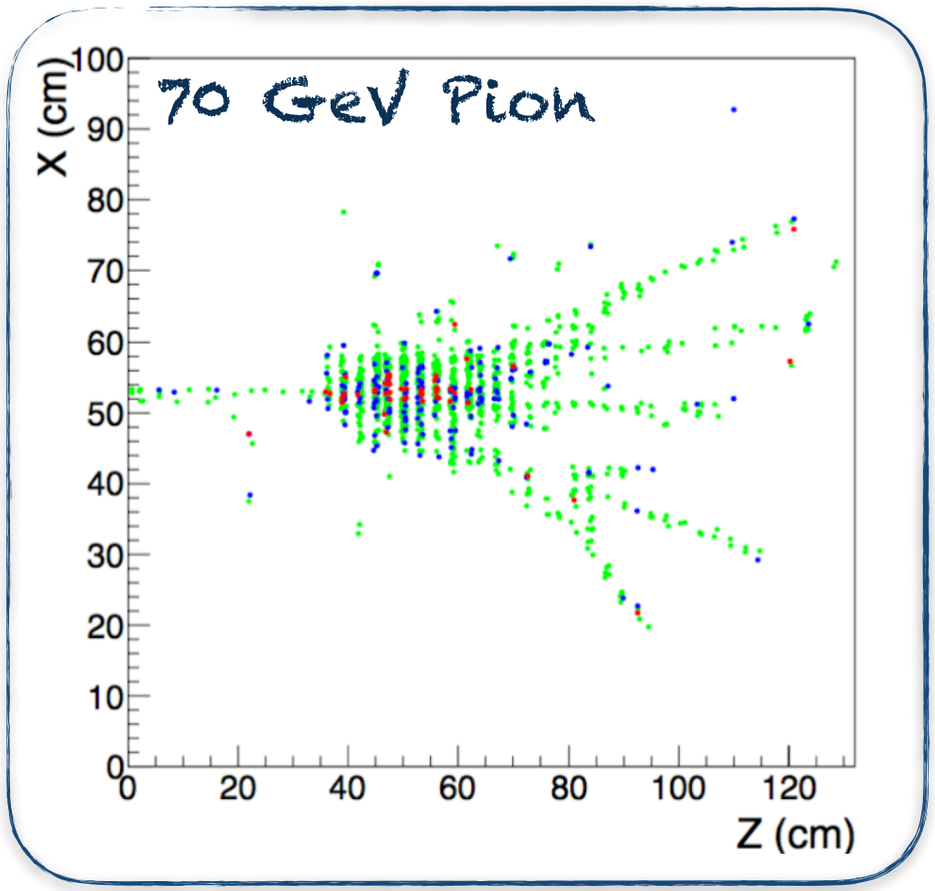

- All nodes are connected : **Spanning Tree**
- **Minimise** the sum of the connected Edge
- No cycle
- Undirected construction

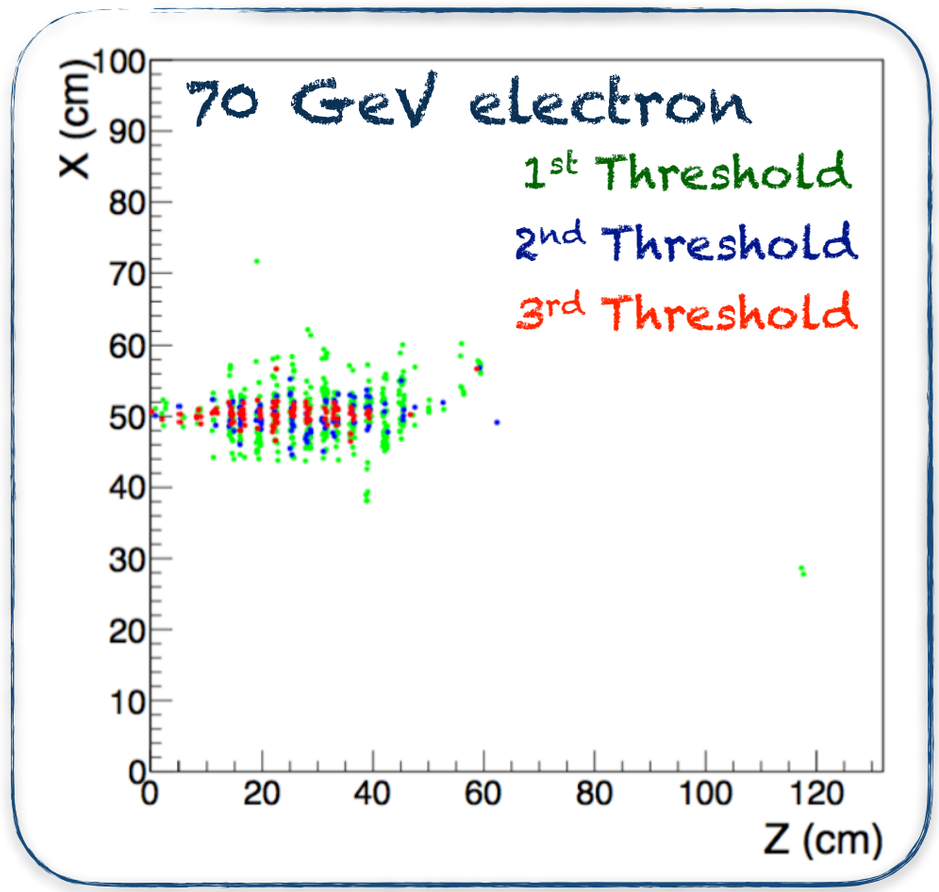
# Minimum Spanning Tree ?

- **Usually characterised by**
  - Length (weight) of the edges
    - Pure euclidian distance isn't particularly useful in our case
      - Need to define 'clever' metrics with weighting
    - Shape of the distribution should reflect density and uniformity of the tree's vertices
  - Numbers of edges / vertices
    - Dense tree expect to have more (electron like)
    - Tail of pion shower will tend to have branch structure : long chain of 2 edges vertices
  - Total weight of the tree
    - Expect heavier for long extended tree ( hadron like)

# Minimum Spanning Tree - SDHCAL

- Take advantage of shower topology difference



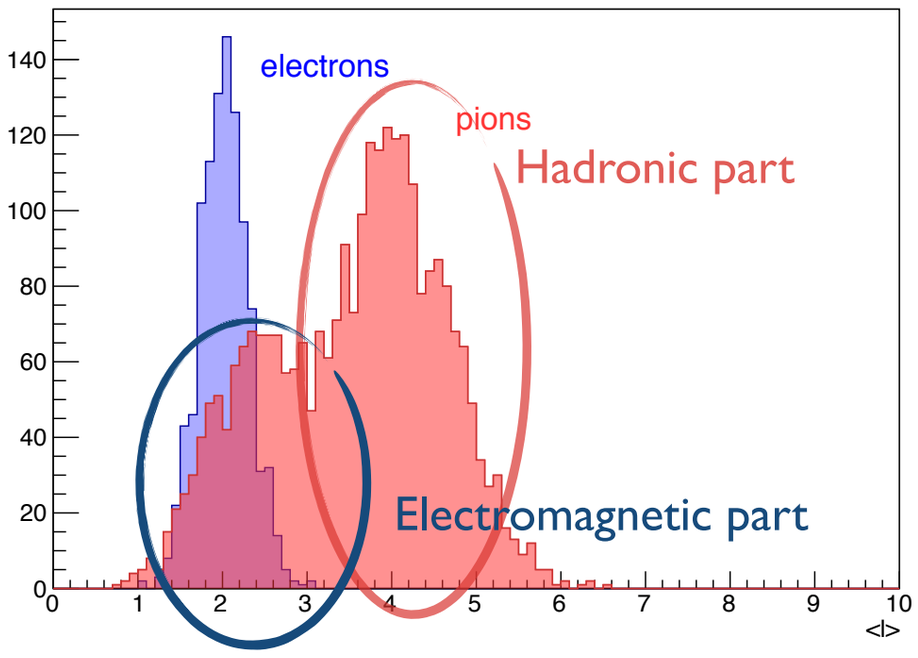
# Minimum Spanning Tree - Metric

- Use hits as nodes
- Metrics can be defined as : $l = \sqrt{\alpha l^2_{Euclid} - \beta hit^2_{treshold}}$

$$l^2_{Euclid} = \sqrt{a(\Delta r)^2 + b(\Delta z)^2}$$

- To help differentiate e from pi we can :
  - weight more along the transverse axis
  - weight more the 3rd threshold

- $< l_e > \quad < \quad < l_\pi >$
- Will accentuate the tree structure of Pions:
  - 1st threshold hit
  - longitudinal development



# Minimum Spanning Tree

- Other variables might also be interesting :
  - Normalised length $\bar{l} = l / < l >$
    - Should be effective in distinguishing trees with different structure (shape)
  - Average edge distance in Z or R
  - Their variance, covariance
    - Would add more info on the density of the shower

# MST Quick Summary

- Start the procedure at the CoG of the shower
  - Different starting point should not have an impact
- Find the nearest hit according to the metric
- repeat until all hits from the shower are linked together

- Few simple variable should be enough for e/pi separation
  - average edge length, edges/vertices, branch length, Total weight, etc.
- Might be trickier for other separation (p/pi)…not much thought for now.
- Definition of a good metric will probably be the most important parameter

- Will test over MC sample

# MST in other experiments

- Charm II implementation results with simple metrics
  - Can separate with good accuracy e/pi even with different energy