A proposal to validate the SDHCAL simulation in iLCSoft

SDHCAL collaboration meeting

Camilo Carrillo, Mary-Cruz Fouz and Héctor García

Ciemat
Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas

26/02/2020

Summary feedback from last week:

- When running the reconstruction, the SDHCAL digitizer (the Marlin processor SimDigital) produces a number of AIDA/ROOT trees.
- Among them, there are SimDigitalGeom and SimDigitalStep containing detailed information on the hit and step contents before the digitisation.
- This was mainly done to check the correctness of the cellid versus digitisation geometry.
- They run by default with the SDHCAL digitizer.
- They are stored in an output file produced by the AIDAprocessor with all other processors dropping their histo/trees there.
- When AIDA is outputed as ROOT, the typical file has a name XXXX AIDA.root containing one ROOT TDirectory per processor so with the digitiser trees in the MySimDigital TDirectory.
- There is also another interesting tuple named -CollectionStat- that is filled after the digitisation
  int NsimHit, NrecoHit, N1, N2, N3

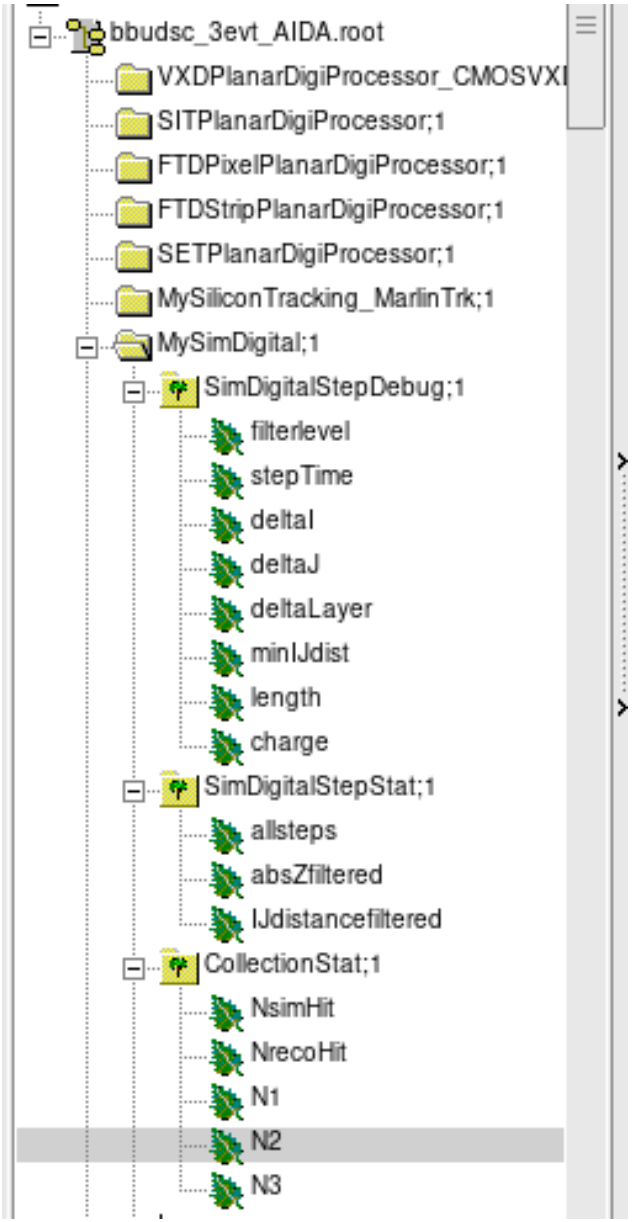
Running the full iLCSoft simulation

- We started with a π^+ gun with the following energies 40 GeV, 60 GeV and 80 GeV. To cover the same energy range as the test beam 2018.
- Some technial details:
  - iLCSoft v02-00-01
  - The default sequences used for simulation where used `StandardConfig/production/ddsim_steer.py` Tunned to fit a random-gun as input.
  - The reconstruction step was done with Marlin by default but changing the default AHCAL for our SDHCAL specified with the –compactFile flag `ILD/compact/ILD_l5_v02/ILD_l5_v02.xml`

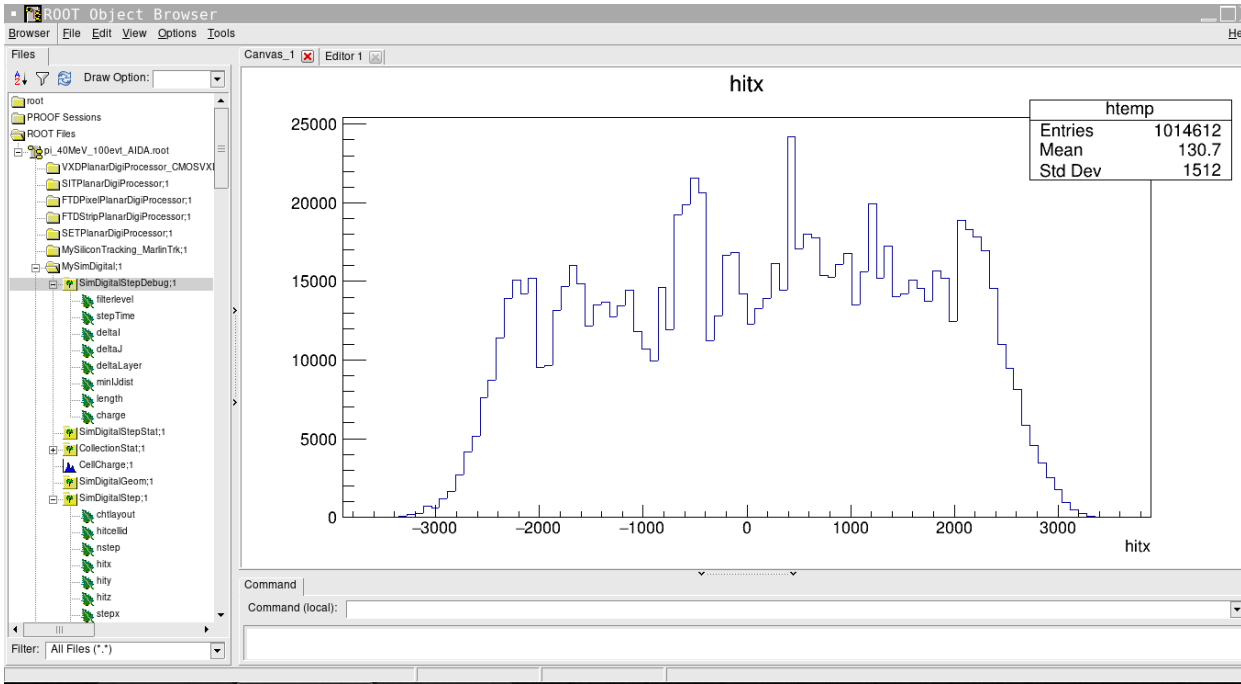
```
carrillo@gaeuicali1:~/ilcsoft/full_chain/ILDConfig-02-00-02/StandardConfig/production> $cat ddsim_40.txt
ddsim --outputFile=./pi_40MeV_100evt_SIM.slcio --compactFile $lcgeo_DIR/ILD/compact/ILD_l5_v02/ILD_l5_v02.xml --steeri
gFile=ddsim_steer_40.py --enableGun --gun.particle pi+ --gun.energy 40*GeV  --gun.distribution uniform --numberOfEvent
 1000

carrillo@gaeuicali1:~/ilcsoft/full_chain/ILDConfig-02-00-02/StandardConfig/production> $cat marlin_40.txt
Marlin MarlinStdReco.xml --constant.lcgeo_DIR=$lcgeo_DIR  --constant.DetectorModel=ILD_l5_o2_v02 --constant.OutputBase
ame=pi_40MeV_100evt --global.LCIOInputFiles=pi_40MeV_100evt_SIM.slcio
carrillo@gaeuicali1:~/ilcsoft/full_chain/ILDConfig-02-00-02/StandardConfig/production> $
```

  - The considered energy values were 40 GeV, 60 GeV and 80 GeV with 10kEvent each.
  - The datasets are available here: `http://wwwae.ciemat.es/~carrillo/calice/datasets/`.

The AIDA tree

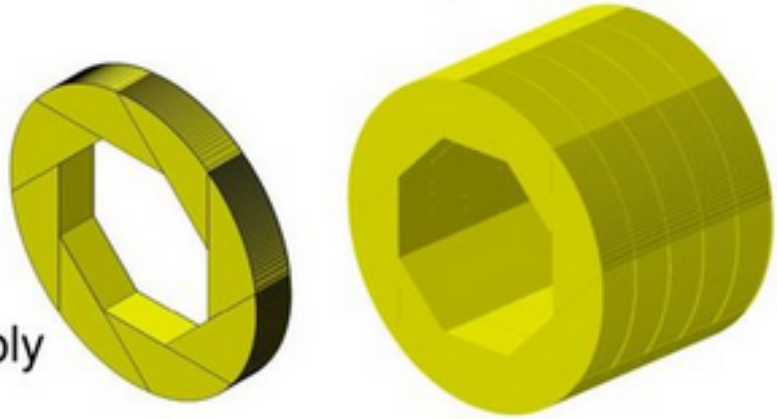


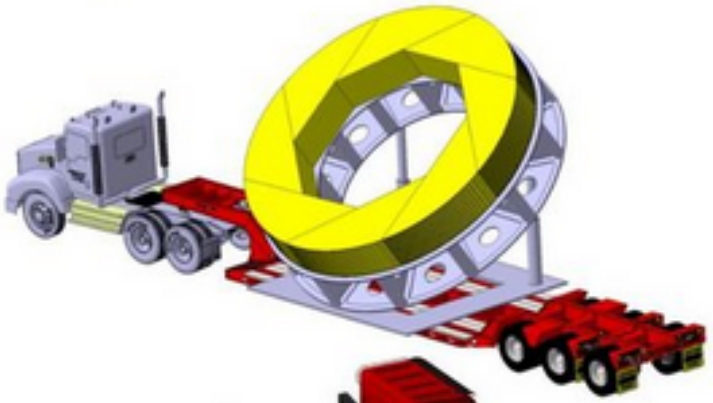
- ddsim step is done (simulation)
- marling step still running (digitization/reconstruction) with SDHCAL (O2)
- distributions a priory look healthy.

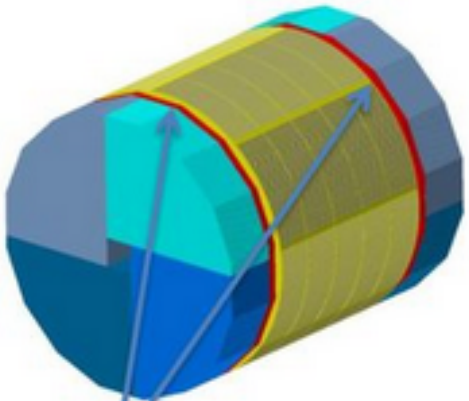


Geometry information, geometry

- SDHCAL barrel consists of 5 wheels
- Each wheel consists of 8 modules
- Weight of a wheel is 88 (+37) ton
- Two scenarios of the wheel structure assembly
  - Assembly in industry / campus
  - Assembly in assembly hall
- Each SDHCAL endcap consists of 4 modules
- Each module weighs 50 (+22.5) ton

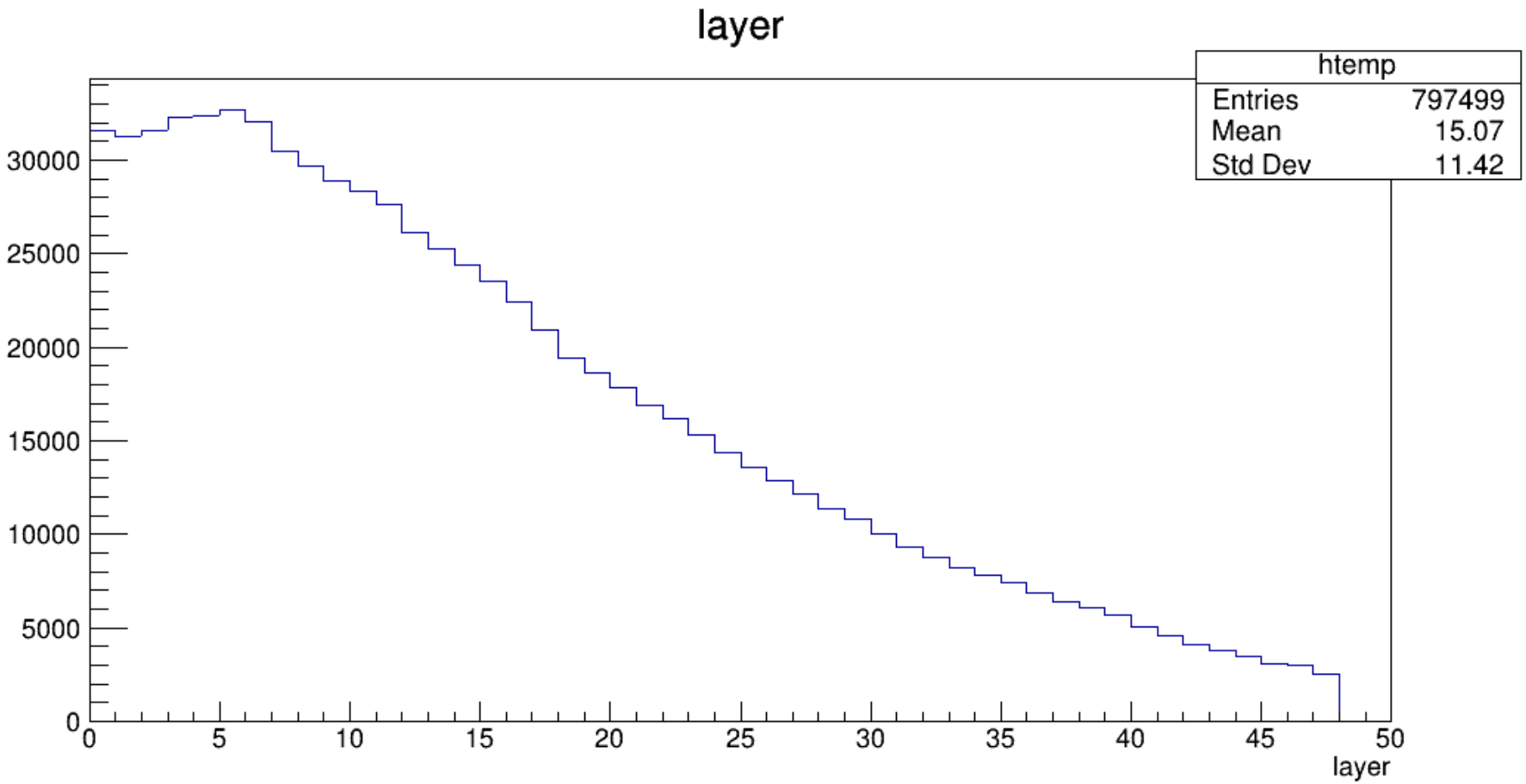




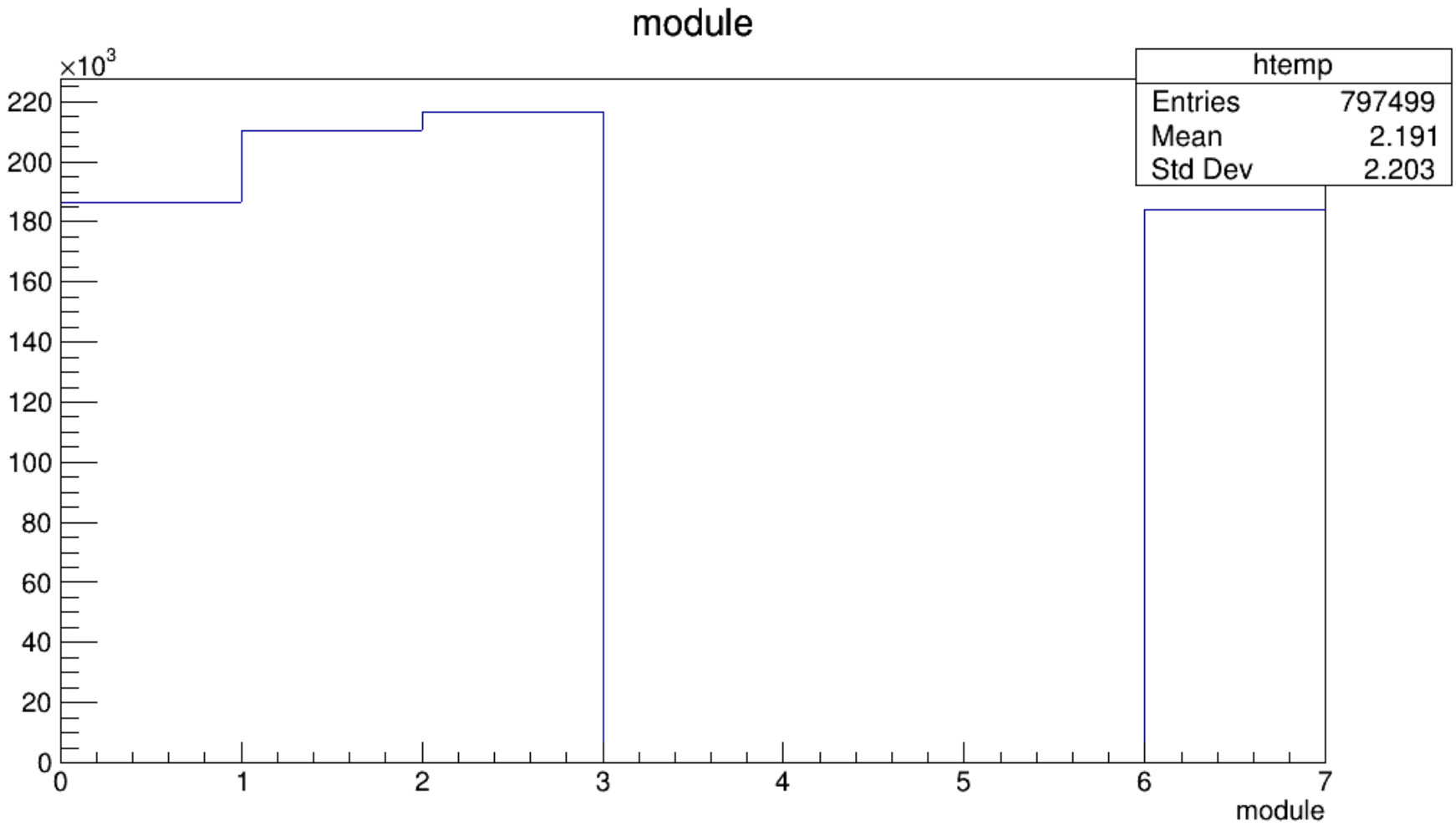


21

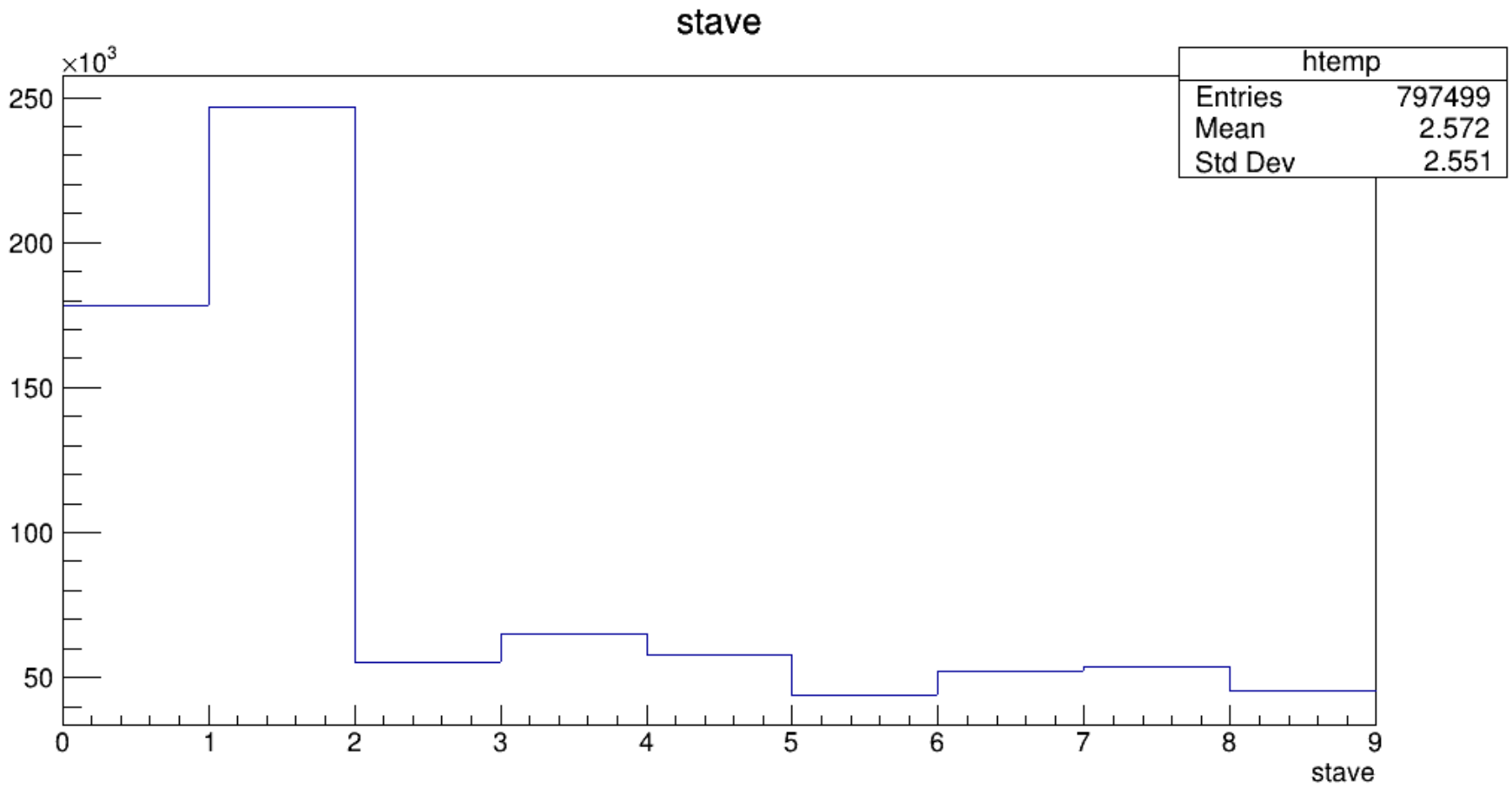
Geometry information, layer



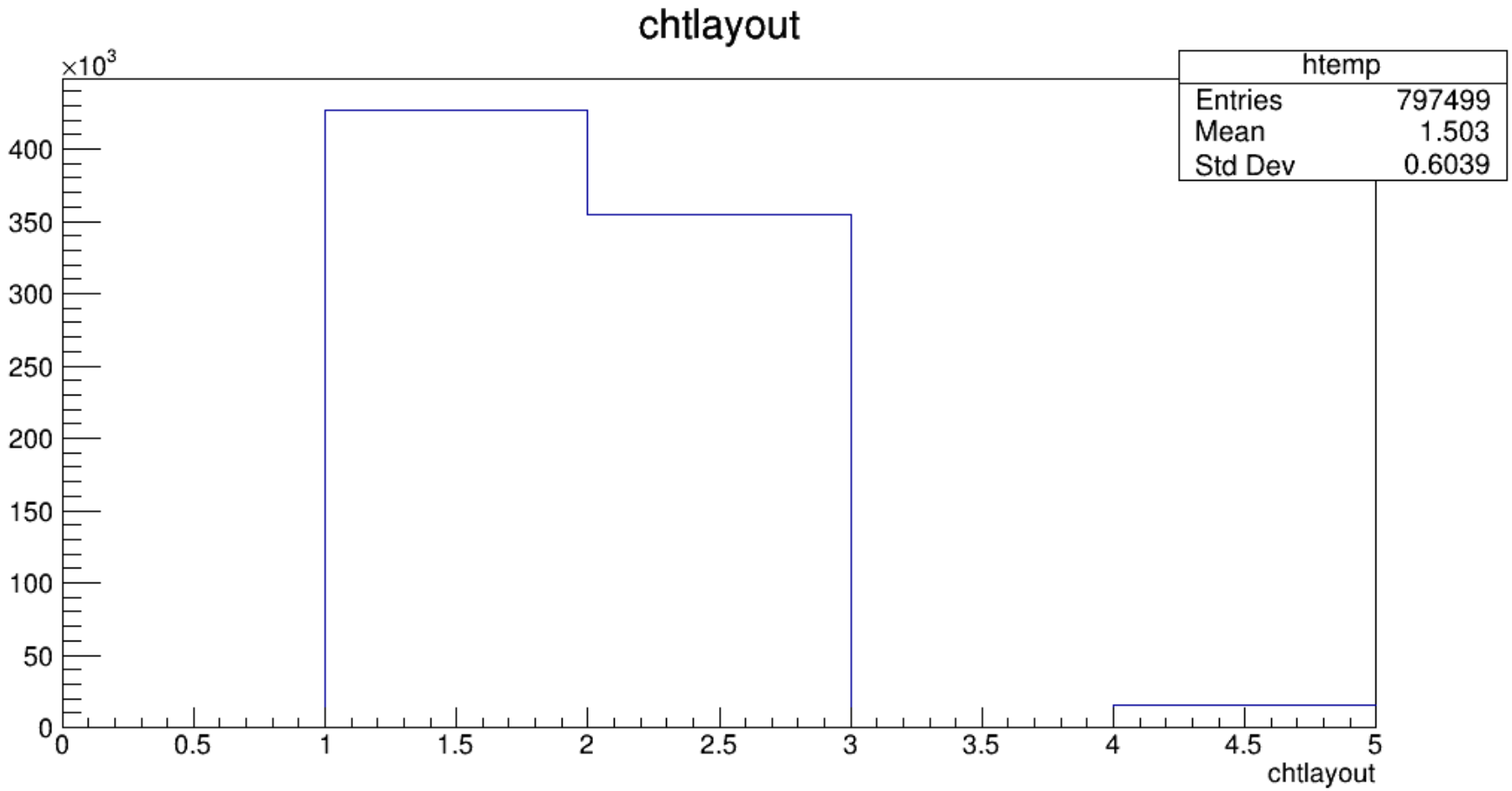
Geometry information, module



Geometry information, stave

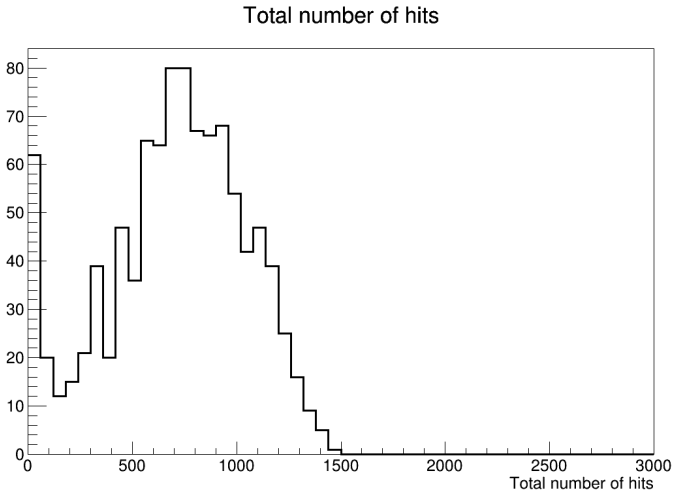


Geometry information, chtlayout

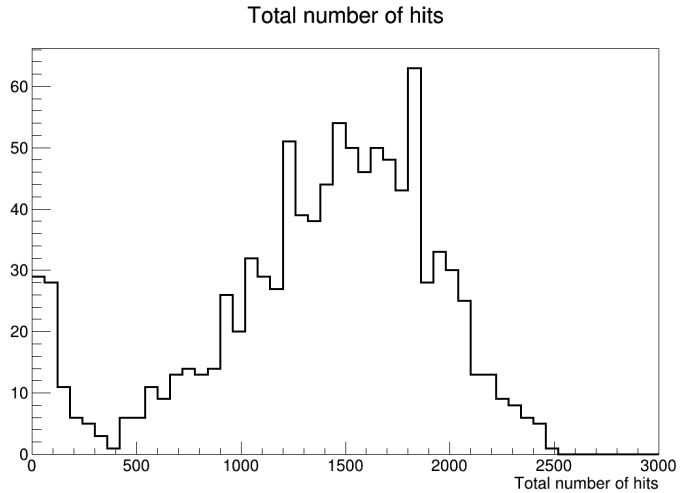


random flat gun, fixed energy Number of hits π^+

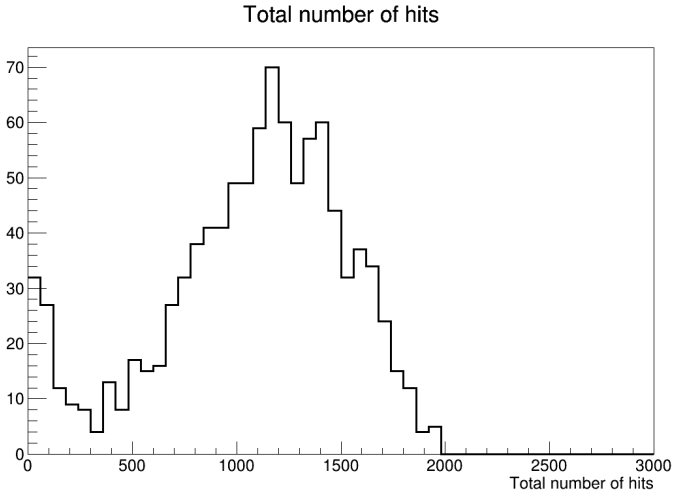
$\pi^+ 40\ GeV$



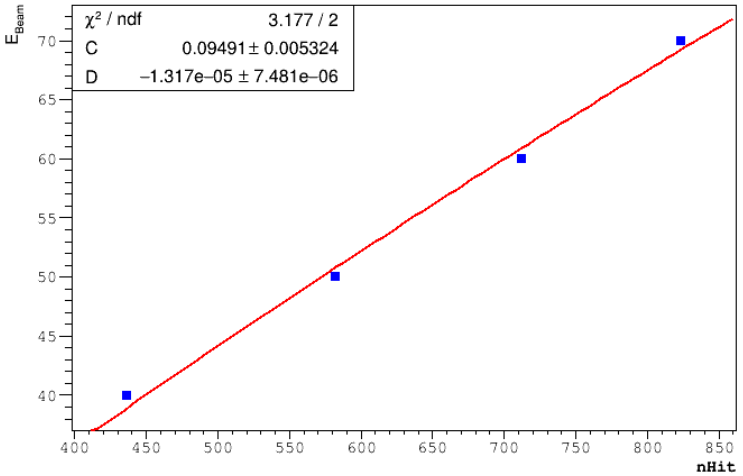
$\pi^+ 80\ GeV$



$\pi^+ 60\ GeV$

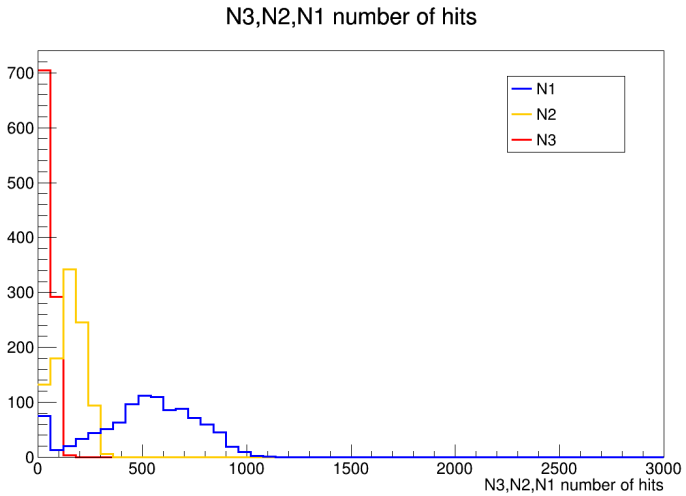


$\pi^+ TB2018$ (/2)?

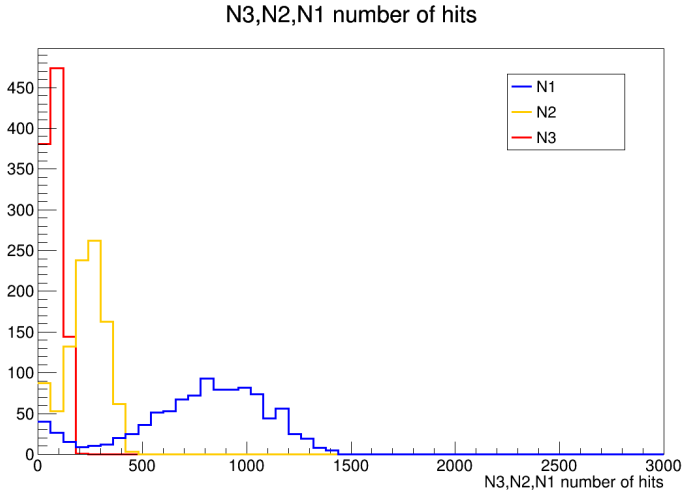


random flat gun, fixed energy Number of hits by thershold

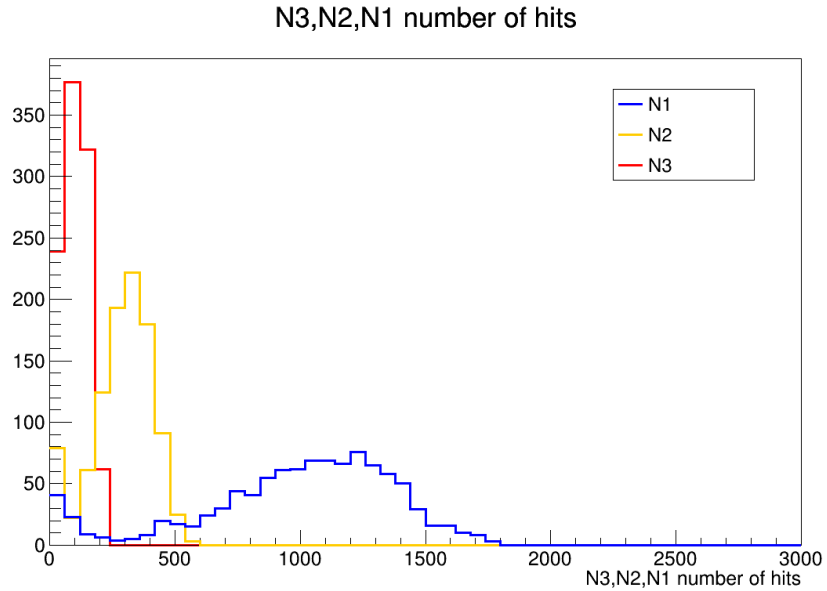
$\pi^+ 40\ GeV$



$\pi^+ 60\ GeV$



$\pi^+ 80\ GeV$



- Distributions behaving as expectedd

Conclusions

- Finally access to all information from the SDHCAL simulation and reconstructions.
- To get deeper in the understanding of the detector would be good to try our own energy regression (also to hunt possible bugs).
- Hector sharing his code in our gitlab respository (new calibration may be needed).
  `https://gitlab.cern.ch/hegarcia/mc-higgs-channel/tree/master/EReco`
- Still pending to understand/fix the factor 2 observed in the number of hits!
- Planing to come back to discuss these results with Gerald on Friday / Next week.

Backup
Slides from last week

Slides for release validation of SDHCAL in iLCSoft (proposal)

The following slides are the proposal/comparions we suggest to validate a given release or MC production for the SDHCAL in the framework of iLCSoft.

We focus on 3 studies that will give an overview of the detector simulation:

- Hit multiplicity.
- Longitudinal and radial profiles.
- Linearity of the reconstructed energy w.r.t MC thruth.

Each item is studied and compared in 3 scnearios (test beam 2018, uds sample, 250 GeV 2020 simulation [when done]).

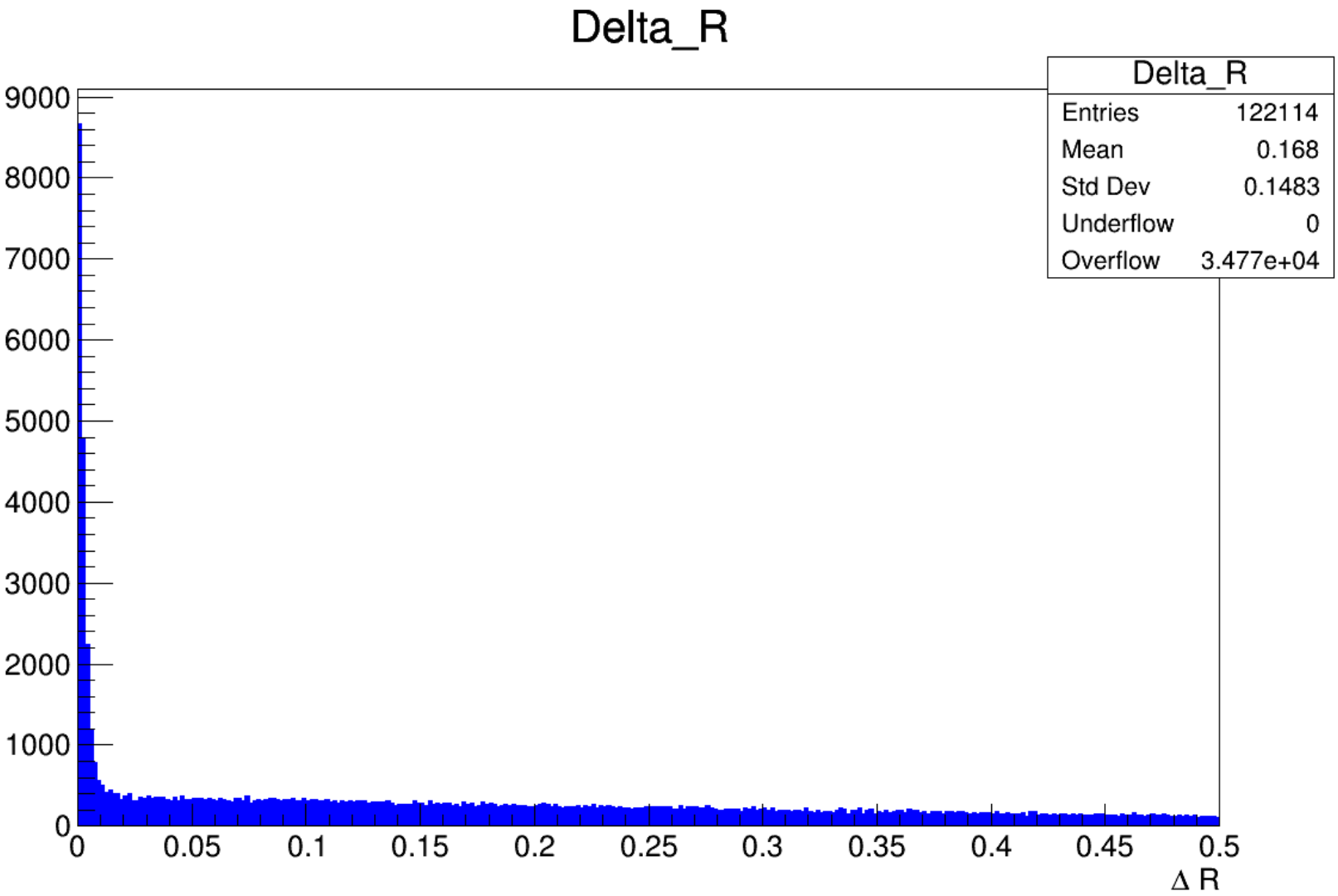
The workflow of the study for a MC (uds) sample:

The slcio file obtained from Dirac was “LCTupliced” :

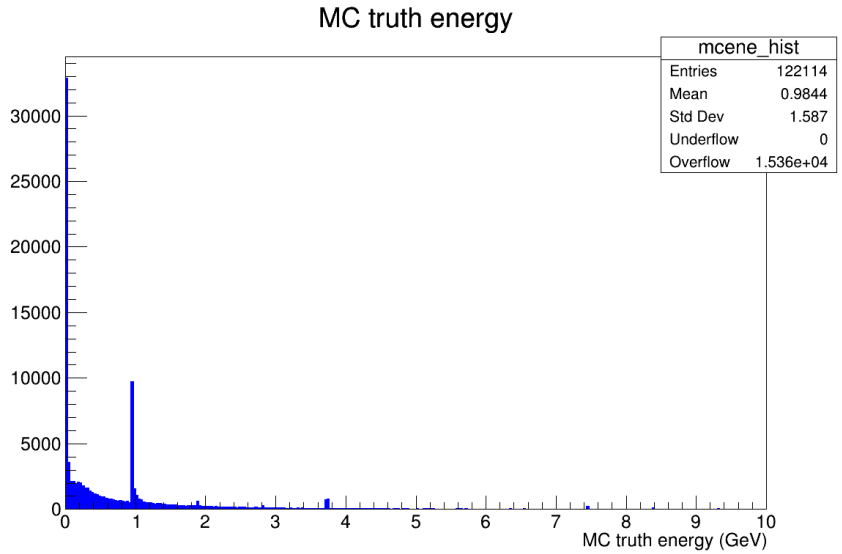
```
Marlin MarlinStdRecoLCTuple.xml
       --global.GearXMLFile=gear_ILD_l5_o2_v02.xml
       --global.LCIOInputFiles=/pool/calice/hectorgc/MC_samples/DIRAC/rec/1-calib/uds/ILD_l5_o2_v02_nobg/v02-00-01/rv0\
2-00-01.sv02-00-01.mILD_l5_o2_v02_nobg.E1-calib.I110069.Puds30.e0.p0.n001_001.d_rec_00010760_452.slcio
       -MyAIDAProcessor.FileName=uds_LCTuple
```

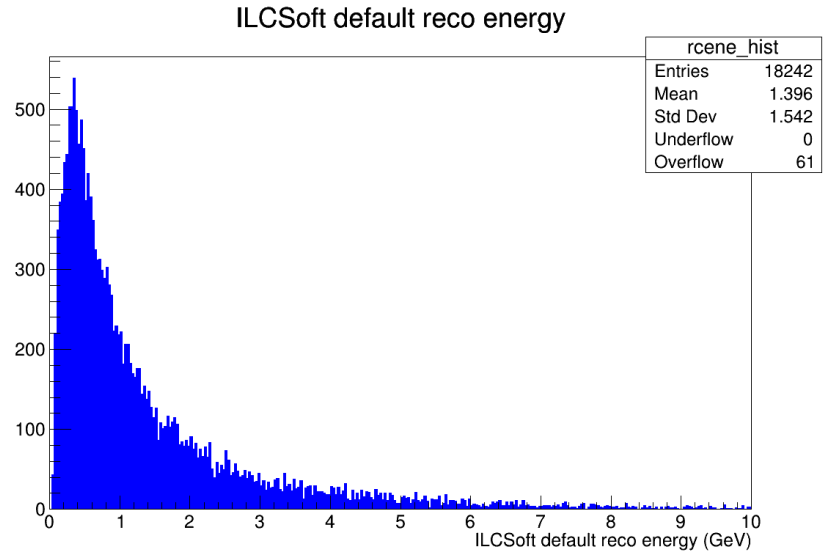
- Providing the right xml files, and correct ilcSoft enviroment.
- Everything runs in the cluster dedicated to CALICE in CIEMAT.
- Once the root-tree is obtained the iterations are done as with any other root tree.

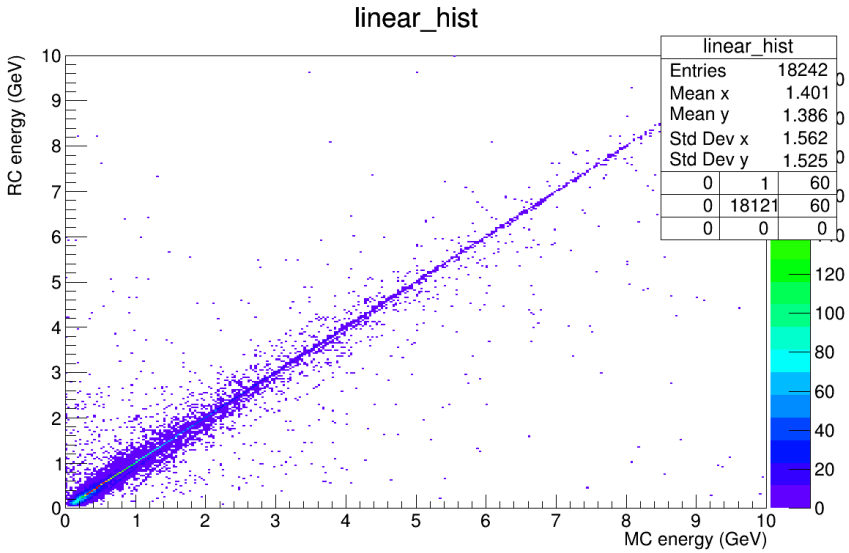
MC/reco matching: ΔR matching, all particles

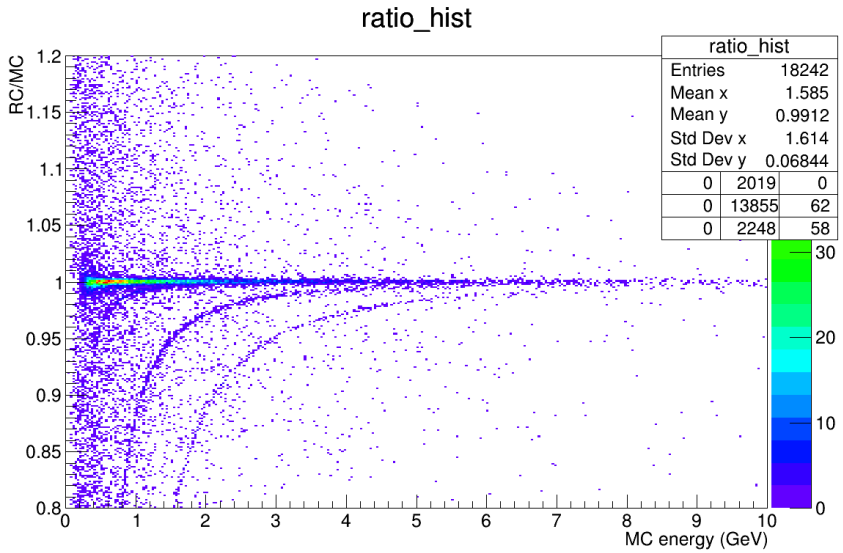


Energy Distributions, all particles



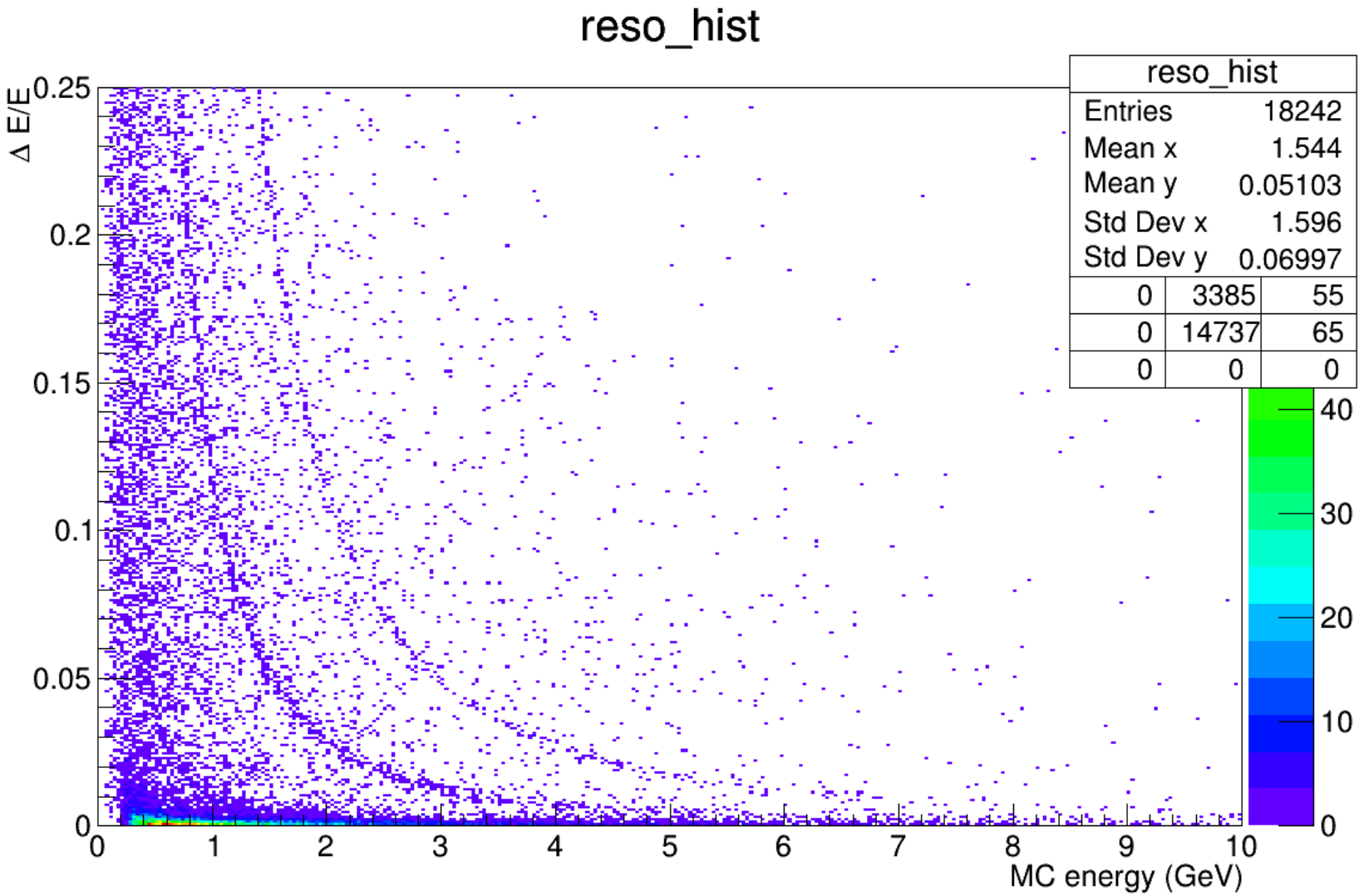




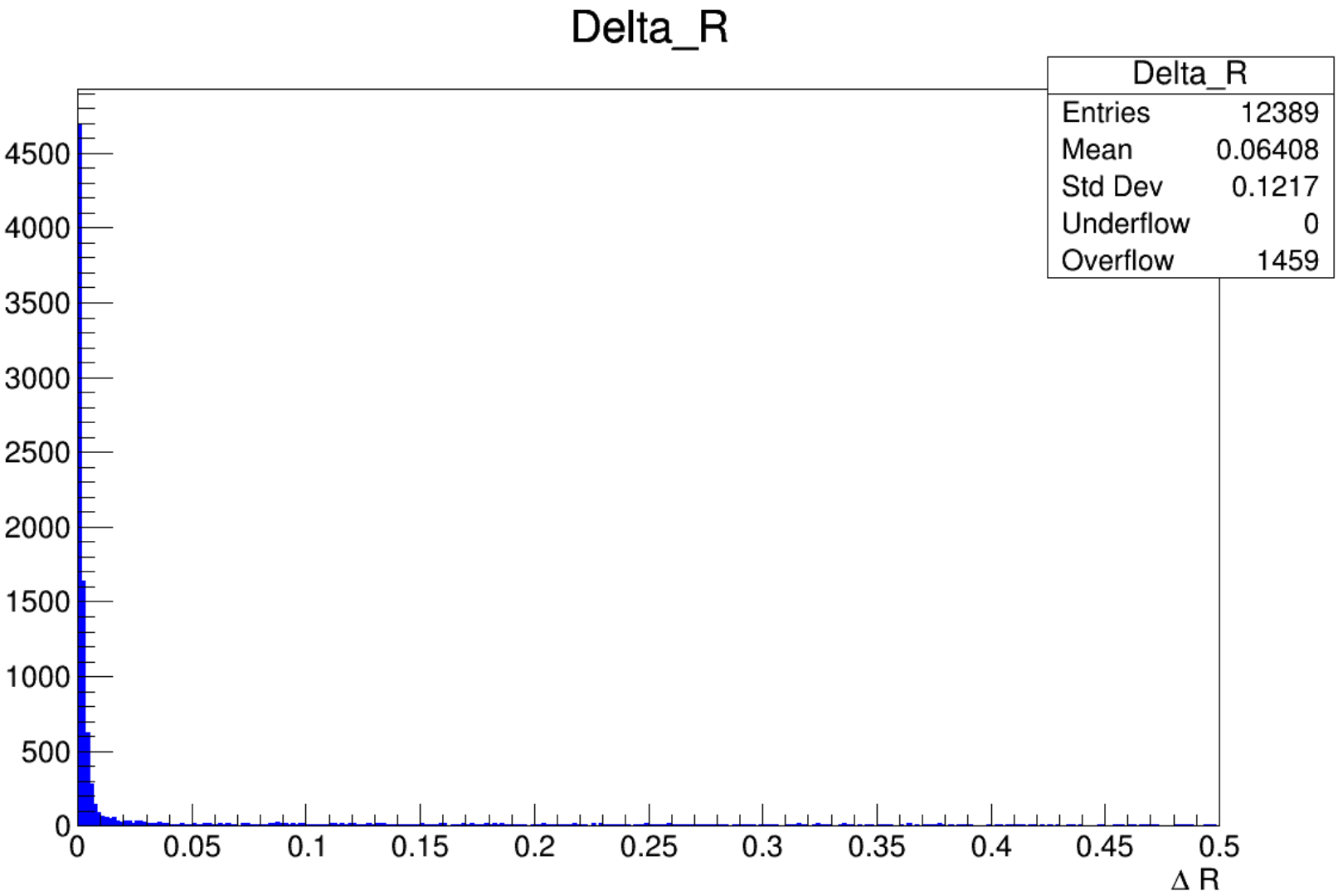


$min\Delta R < 0.01$, all particles included.

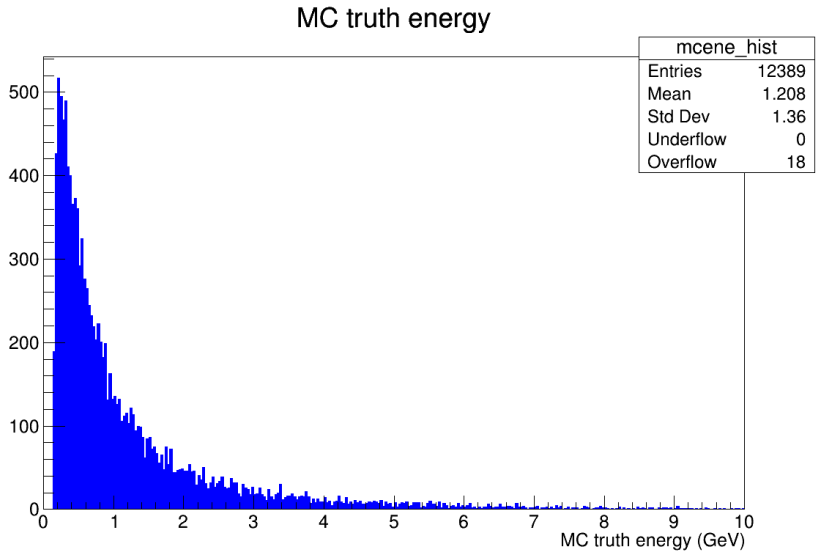
$\Delta E/E$, all particles.

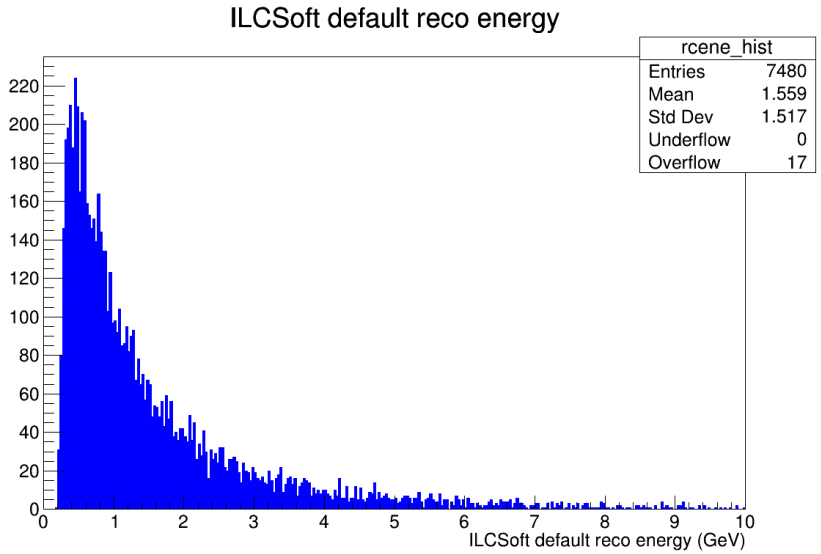


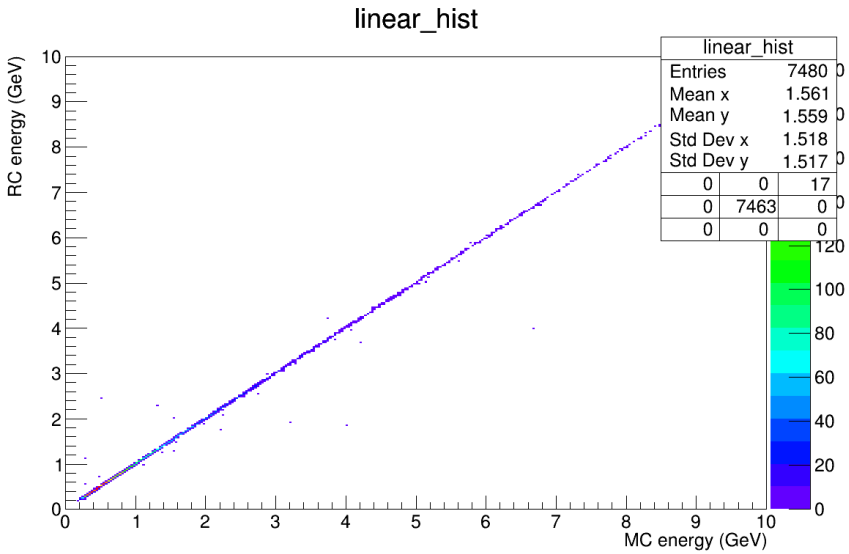
Now we focus in pions ΔR matching, $\pi^{\pm}$

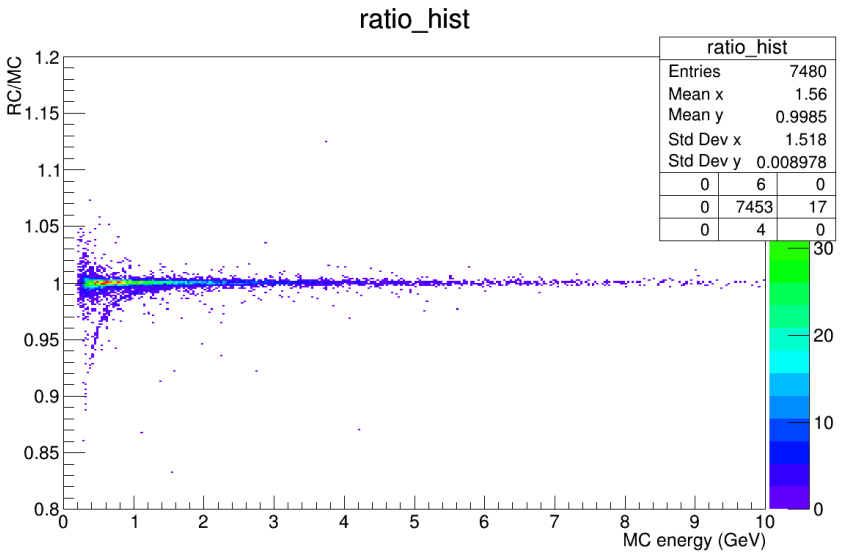


Energy Distributions, $\pi^{\pm}$



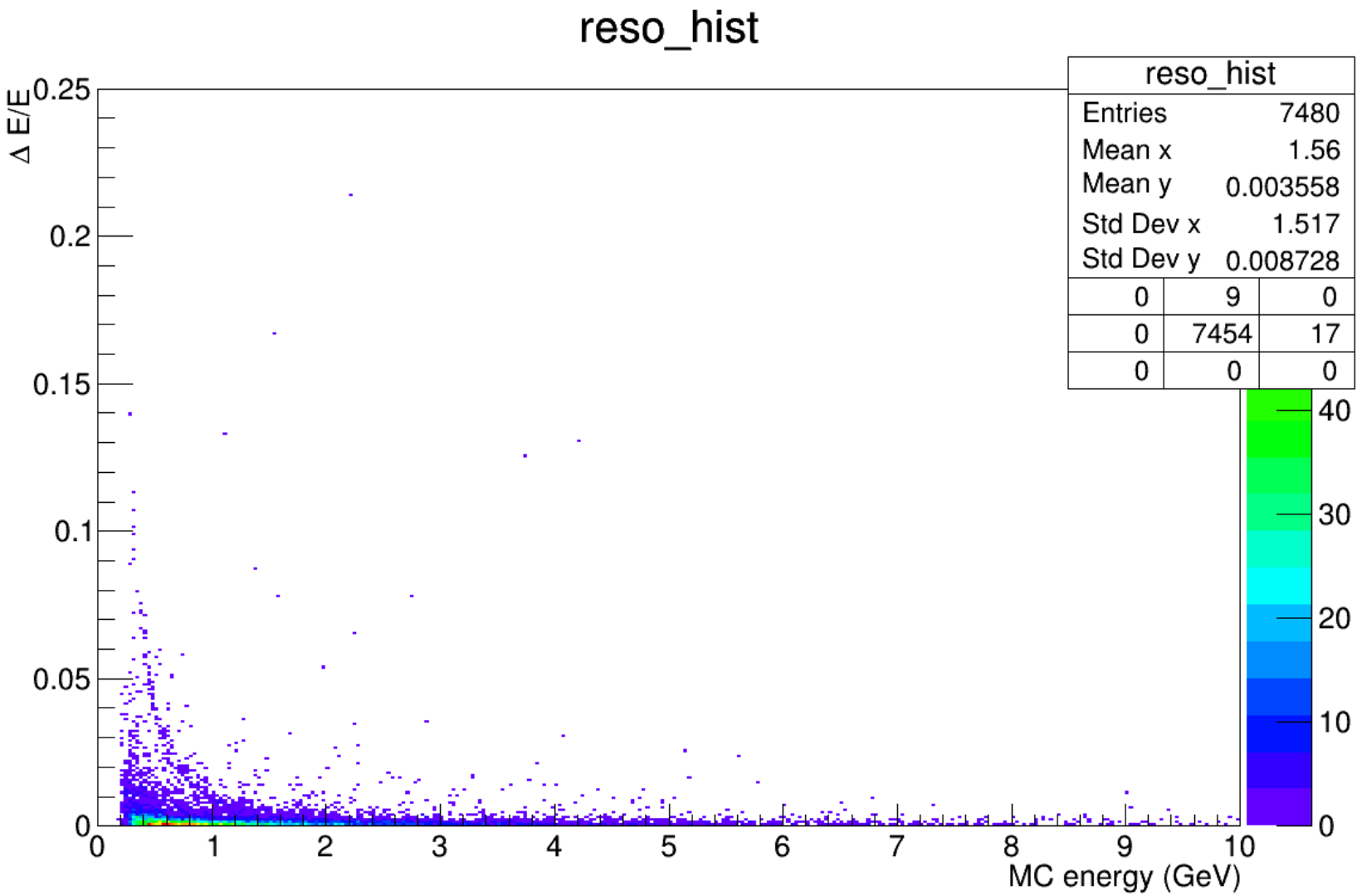






$min\Delta R < 0.01$, only $\pi^{\pm}$

$\Delta E/E, \pi^{\pm}$

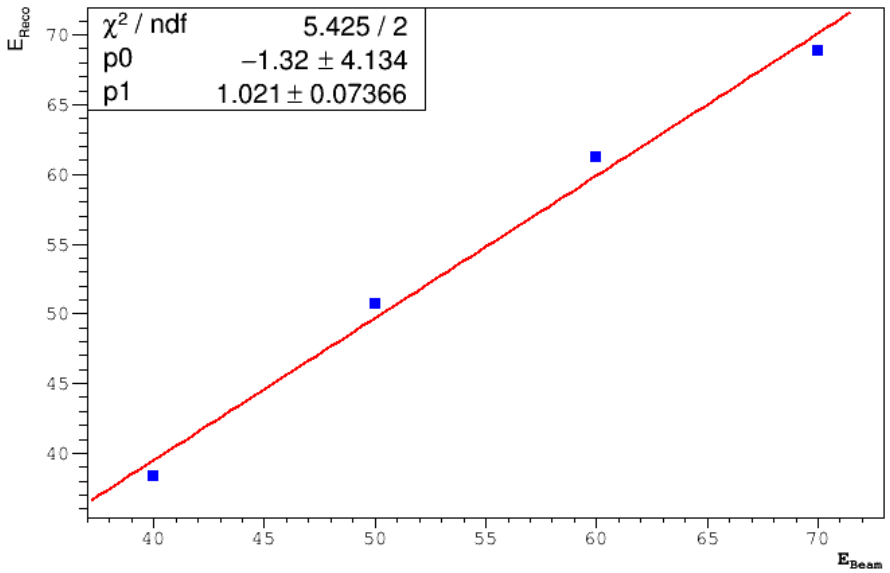


All distributions, all particles

- Set of plots, including resolution plots:
  `http://wwwae.ciemat.es/~carrillo/calice/`
- gitlab repository with all code (nice-cern log-in required):
  `https://gitlab.cern.ch/carrillo/mc-higgs-channel/tree/master/Analysis`

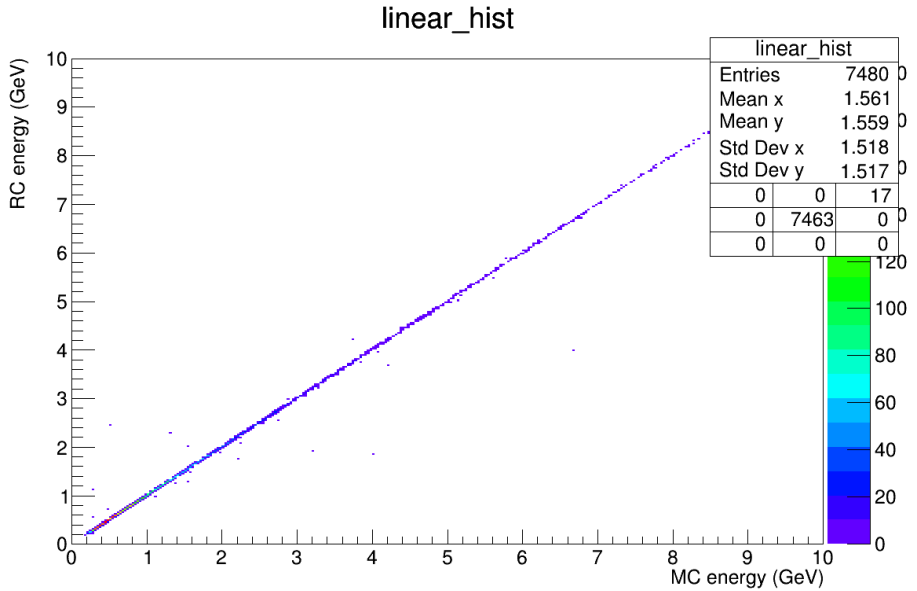
Energy Linearity ($\pi^{\pm}$) proposed for validation

test beam 2018



- Plot from test-beam 2018 analysis (check Hectors slides).
- Discrete energy spectrum.
- No MC truth but well defined energy for the pion beams.

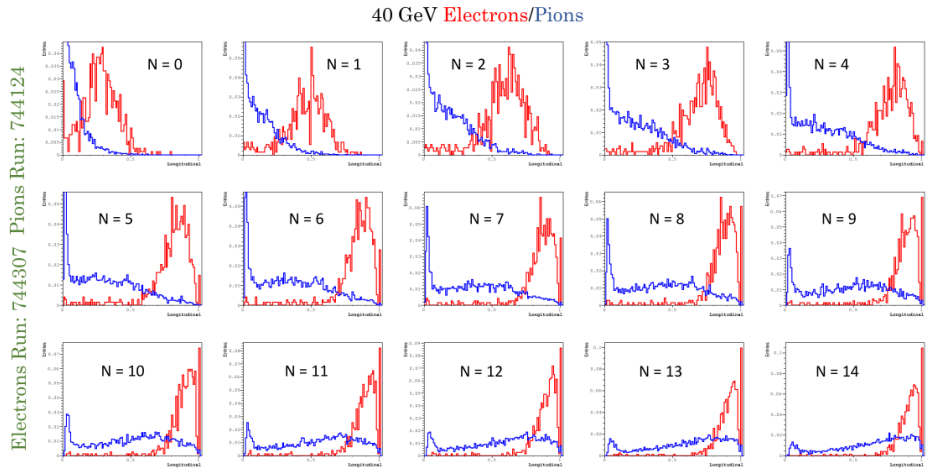
uds sample



- Plot with LCTuple workflow.
- Continous energy spectrum but needs to be enlarged (may be with the 250GeV)
- Still pending the desentanglement of SDHCAL stand alone energy.

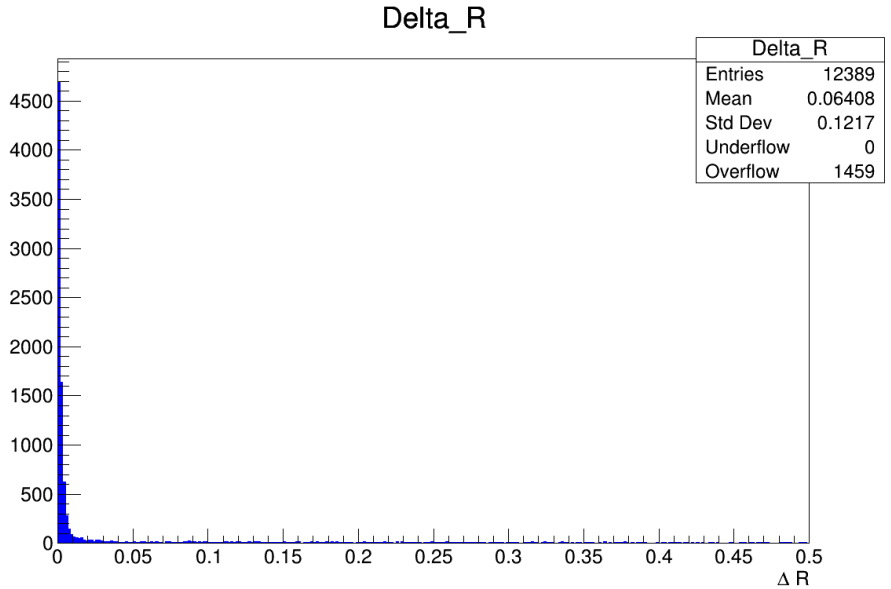
Longitudinal and radial profile ($\pi^{\pm}$) proposed for validation

test beam 2018 (ΔR)



- How wide and deep the shower gets (depends on the material in front and the thersholds).
- Could be either measured with ΔR or as average number of hits per plane / layers.
- A measurement related to the width (size of the smallest cone containing the shower) still to be done with the TB data analysis.

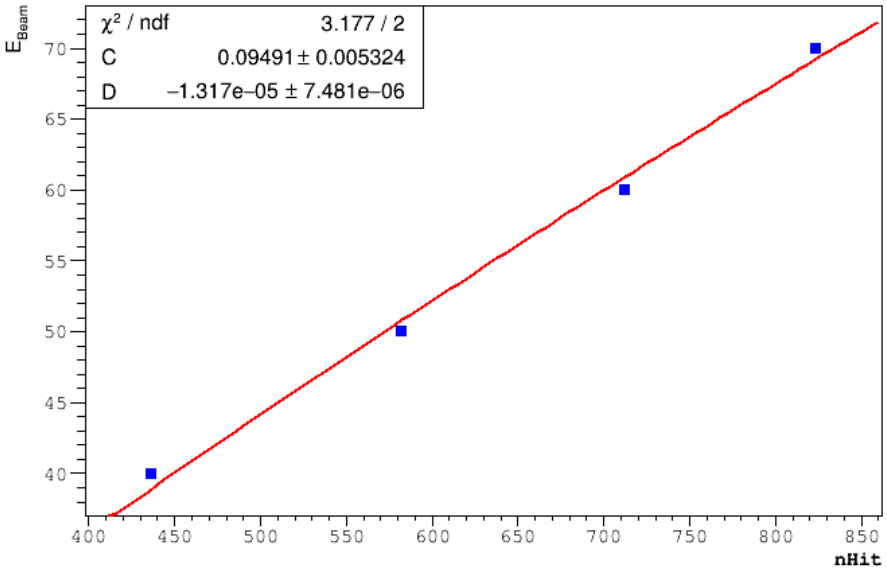
uds sample



- **The two quantities are not equivalent** but are related to the size of the shower.
- To produce the number of hits associated to a given shower we are still facing an issue accessing the SDHCAL hits belonging to a given shower in the LCTuple.
- The ΔR plots already shown in previous slides will give us a measurement of the size of the showers observed for the pions (convoluted with the error in the measurement of the size of the cone and the error in the direction of the shower.

Hit Multiplicity ($\pi^{\pm}$) proposed for validation

test beam 2018 (Hector)



uds sample

to be done

- The number of hits associated to a particle passing any thershold (and another one passing each specific thershold).
- Still issues accessing the information at single hit level in the ntuples.

Where such validation can centrally run?

iLCSoft use Travis-CI for the iLCSoft Continous Integration. The following things are currently setup[1]:

- Run every night nightly builds. These are managed by our CLIC friends. The software is deployed on CVMFS.
- PR for all packages: the code is compiled against the nightly build and unit tests are run. But they don't include any check on the data itself, just code.
- In the ILDConfig package: there is nothing to compile in this package it is only configuration scripts, a check that the standard software chain for ILD is running.
- In general we use CVMFS for deploying the software, either in the CLIC CVMFS or the ILC one.

For continous integration checks are done only at Software level. May be here is not a good place to add our SDHCAL goodness checks. For the moment it will be a stand alone script running on top of LCTuples.

[1] Remi private exchange

Conclusions, Short term plans

Validation of the reconstruction chain, from simulated hits to standalone energy measurement with the SDHCAL

- Solve issue accessing single hits at the LCTuples level.
- Stablish a simple goodness check for the 3 studies that are proposed. Or, what do we expect from the SDHCAL in a given simulation.
- Study The next large-scale simulation of samples at 250 GeV (which should be done by now?).
- Once we are happy, we present the proposal and results in the ild-software-conveners and ild-physics-conveners.

All software will be public and stored in gitlab@cern already pointed out. Suggestions are very welcome.

Learning iLCSoft

In order to build up the validation tool and extend the timming studies. We need to add the SimHit and Digi information to the LCTuple:

- The steering xml file for the LCTuples already contain the option:

Branch: master ▾ LCTuple / example / lctuple.xml — Find file | Copy path

gaede - added support for TrackerHits — 308ef53 on May 15, 2013

2 contributors

152 lines (127 sloc) | 7.28 KB — Raw | Blame | History

```
<?xml version="1.0" encoding="us-ascii"?>
<!-- ?xml-stylesheet type="text/xsl" href="http://ilcsoft.desy.de/marlin/marlin.xsl"? -->
<!-- ?xml-stylesheet type="text/xsl" href="marlin.xsl"? -->

<!-- Loading shared library : /Users/fgaede/marlin/mymarlin/lib/libmymarlin.0.1.0.dylib (libmymarlin.dylib)-->

<!--##########################################
    #                                        #
    #     Example steering file for marlin   #
    #                                        #
    ##########################################-->


<marlin xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:noNamespaceSchemaLocation="http://ilcsoft.desy.de/marlin/marlin.xsd
```

Learning iLCSoft

- and also our steering file (the one used for the uds sample):

```
<processor name="MergeSimCaloHits" type="MergeCollections">
  <!--MergeCollections creates a transient subset collection that merges all input collections -->
  <!--Names of all input collections-->
  <parameter name="InputCollections" type="StringVec">
    BeamCalCollection
    EcalBarrelCollection
    EcalEndcapRingCollection
    EcalEndcapsCollection
    HcalBarrelRegCollection
    HcalEndcapRingCollection
    HcalEndcapsCollection
    LumiCalCollection
    YokeBarrelCollection
    YokeEndcapsCollection
    LHCalCollection
  </parameter>
 <!--Optional IDs for input collections - if given id will be added to all objects in merged collections as ext<CollID>()-->
  <parameter name="InputCollectionIDs" type="IntVec">
    25 20 21 29 22 23 30 24 27 31
  </parameter>
  <parameter name="OutputCollection" type="string">SimCalorimeterHits </parameter>
  <!--verbosity level of this processor ("DEBUG0-4,MESSAGE0-4,WARNING0-4,ERROR0-4,SILENT")-->
  <parameter name="Verbosity" type="string">DEBUG </parameter>
</processor>
```

- It seems there is a merged collection of all our SimCaloHits. Including the Ecal.

Learning iLCSoft

- However I can not find the branches/leafs in the obtained TTree.

```
carrillo@gaeuicali1:~/ilcsoft/ntuplizer/mc-higgs-channel/Analysis> $root -b uds_LCTuple.root
root [0]
Attaching file uds_LCTuple.root as _file0...
(TFile *) 0x30f5290
root [1] MyLCTuple->Show()
======> EVENT:-1
 evevt           = 0
 evrun           = 0
 evwgt           = 0
 evtim           = 0
 evsig           = 0
 evene           = 0
 evpoe           = 0
 evpop           = 0
 evnch           = 0
 nmcp            = 0
 nrec            = 0
 npid            = 0
 ntrk            = 0
 ntrst           = 0
 nsth            = 0
 nsch            = 0
 r2mnrel         = 0
root [2]
```

- Is a recompilation needed?, how?
- Reco Hits are not even listed in the steering file. Would need to dig deeper.
- lets say work in progress...