



Baseline Selection for CLIC Data

Ben & Matt



Overview



- **Goal: implement simple particle ID selectors** to use as baseline for deep learning algorithm performance studies
 - Use only simple quantities (ECAL and HCAL energies, moments, etc)
 - Implement both cut-based selectors and simple MVAs (will use TMVA)
- As a starting point, I wrote a python script that reads in the particle gun root files and dumps a small flat ntuple
 - Currently storing only a few quantities (total ECAL and HCAL energies, numbers of hits, particle gun info)
 - Additional variables will be stored as branches
 - These root files can then be passed to TMVA



Data Formats



root file

(output of DDSim collision+CLIC response)
custom data formats, e.g. HCalBarrelCollection



[Convert_to_txt.py](#)

text file

stores particle gun info, positions and
energies of ECAL/HCAL hits



[Convert_to_h5.py](#)

hdf5 file

stores particle gun info, positions and
energies of ECAL/HCAL hits

[See code from Matt here:](#)

[/afs/cern.ch/user/m/mazhang/Projects/JetCalo/ConversionScripts\)](#)



Data Formats



root file

(output of DDSim collision+CLIC response)
custom data formats, e.g. HCalBarrelCollection



[Convert_to_txt.py](#)

text file

stores particle gun info, positions and
energies of ECAL/HCAL hits



[Convert_to_h5.py](#)

hdf5 file

stores particle gun info, positions and
energies of ECAL/HCAL hits



root file

(flat ntuple)

stores particle gun info, total ECAL/HCAL
energies and numbers of hits
will add moments and other variables

[Convert_to_FlatNtuple.py](#)

See my code here:

[/afs/cern.ch/user/b/benhoob/public/MLProject](#)

See code from Matt here:

[/afs/cern.ch/user/m/mazhang/Projects/JetCalo/ConversionScripts\)](#)



Example Results

- Ran code on single-electron gun file with 100 events:
 - /eos/project/d/dshep/LCD/EleEscan/EleEscan_8_95.root
 - See file at: /afs/cern.ch/user/b/benhoob/public/MLProject/EleEscan_8_95_flatNtuple.root

```

root [1] flatNtuple->Print()
*****
*Tree   :flatNtuple: flatNtuple
*Entries :    100 : Total =          9967 bytes File Size =          4011 *
*       :       : Tree compression factor = 1.56
*****
*Br    0 :gunpx   : gunpx/F
*Entries :    100 : Total Size=          952 bytes File Size =          478 *
*Baskets :     1 : Basket Size= 32000 bytes Compression= 1.00
*.....*
*Br    1 :gunpy   : gunpy/F
*Entries :    100 : Total Size=          952 bytes File Size =          101 *
*Baskets :     1 : Basket Size= 32000 bytes Compression= 4.73
*.....*
*Br    2 :gunpz   : gunpz/F
*Entries :    100 : Total Size=          952 bytes File Size =          101 *
*Baskets :     1 : Basket Size= 32000 bytes Compression= 4.73
*.....*
*Br    3 :m       : m/F
*Entries :    100 : Total Size=          932 bytes File Size =          101 *
*Baskets :     1 : Basket Size= 32000 bytes Compression= 4.69
*.....*
*Br    4 :gunE    : gunE/F
*Entries :    100 : Total Size=          947 bytes File Size =          477 *
*Baskets :     1 : Basket Size= 32000 bytes Compression= 1.00
*.....*
*Br    5 :pdgID   : pdgID/I
*Entries :    100 : Total Size=          952 bytes File Size =          161 *
*Baskets :     1 : Basket Size= 32000 bytes Compression= 2.97
*.....*
*Br    6 :nhcalhits : nhcalhits/I
*Entries :    100 : Total Size=          972 bytes File Size =          332 *
*Baskets :     1 : Basket Size= 32000 bytes Compression= 1.45
*.....*
*Br    7 :necalhits : necalhits/I
*Entries :    100 : Total Size=          972 bytes File Size =          366 *
*Baskets :     1 : Basket Size= 32000 bytes Compression= 1.32
*.....*
*Br    8 :hcaltote : hcaltote/F
*Entries :    100 : Total Size=          967 bytes File Size =          481 *
*Baskets :     1 : Basket Size= 32000 bytes Compression= 1.00
*.....*
*Br    9 :ecaltote : ecaltote/F
*Entries :    100 : Total Size=          967 bytes File Size =          481 *
*Baskets :     1 : Basket Size= 32000 bytes Compression= 1.00
*.....*

```

particle gun info

total number of hits

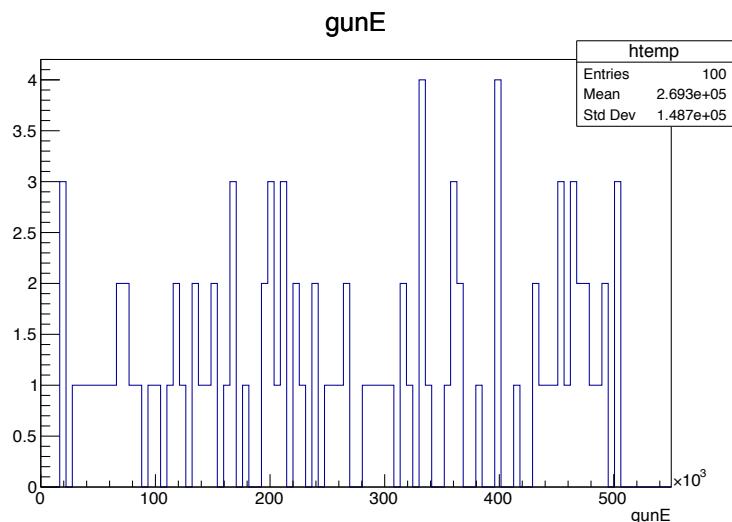
total energy (sum of hits)



Total Energies

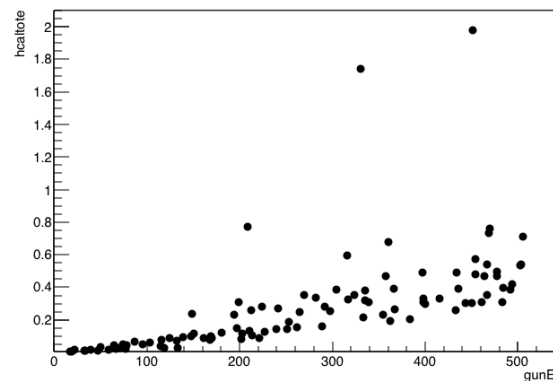
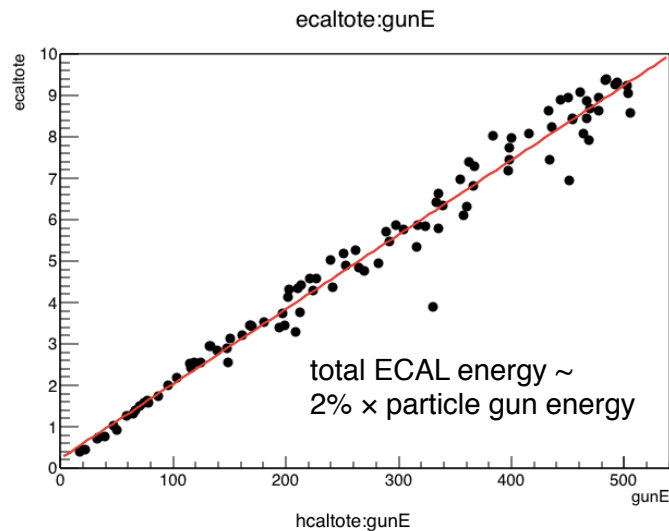


looks like a flat energy spectrum up to 500 GeV (expected)



Total HCAL energy is very small and correlated with with particle gun energy (this seems reasonable)

total ECAL energy is strongly-correlated with particle gun energy but off by a factor of ~ 50
→ **this seems strange**

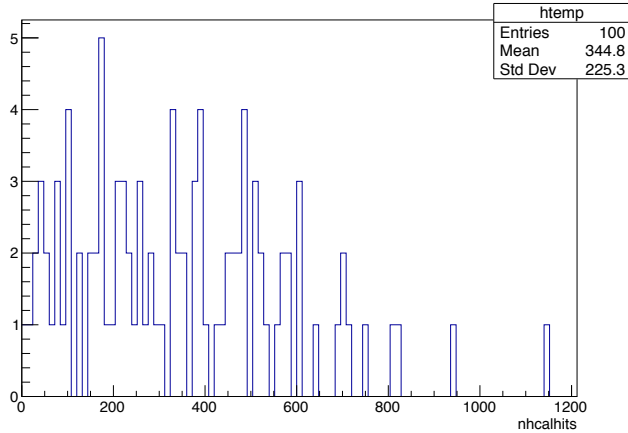




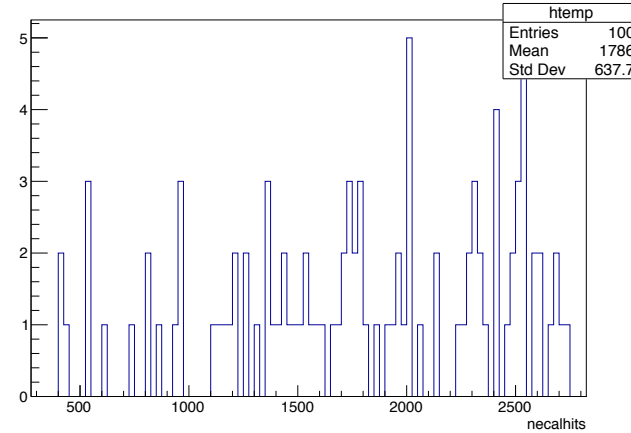
Number of Hits



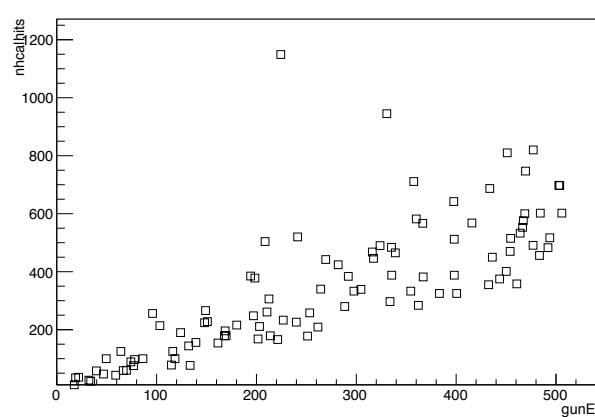
nhcalhits



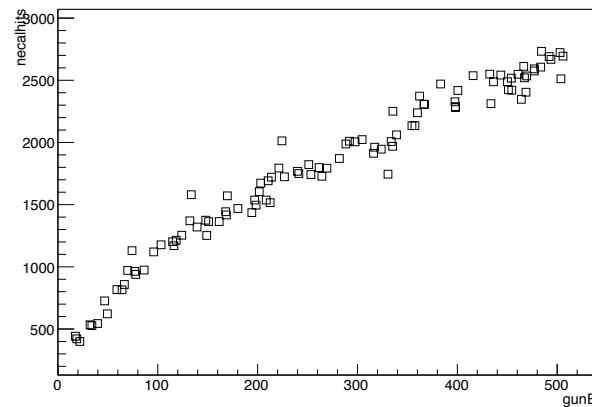
necalhits



nhcalhits:gunE



necalhits:gunE



- The total number of hits is correlated with the particle gun energy for both HCAL and ECAL



Summary



- I have written a script to convert the DDSim root files into flat ntuples, and stored a few basic branches
- I found that total ECAL energy $\sim 2\%$ of the particle gun energy, which is strange
- Next step is to compute simple variables and store them as branches, and then pass them to TMVA