Gamma/Proton discrimination for LHAASO-WCDA

WANG XIAOJIE

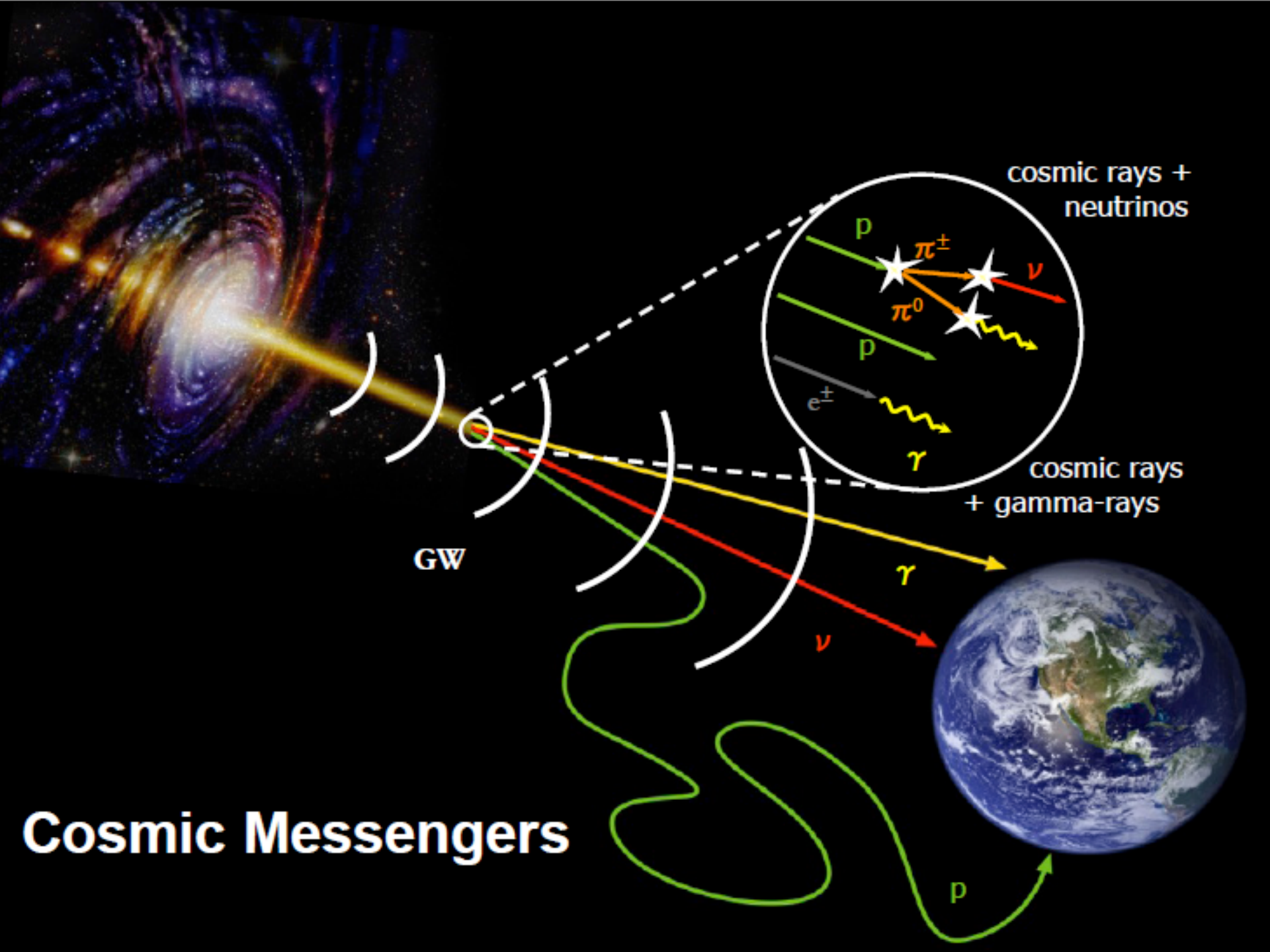
For the LHAASO collaboration

wangxiaojie@ihep.ac.cn

Madison WI , 26/07/2019

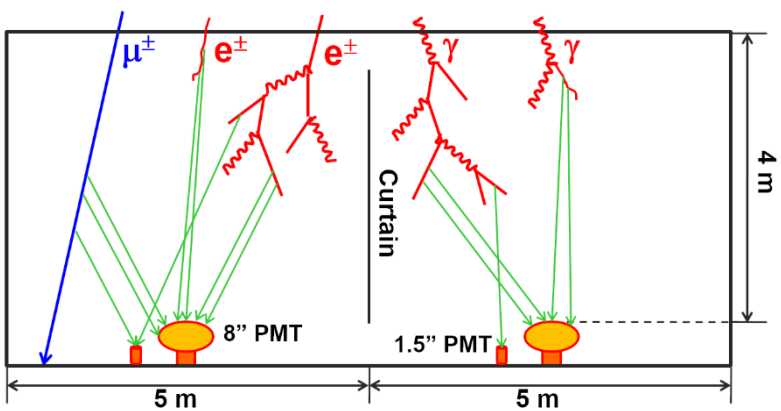
Outline

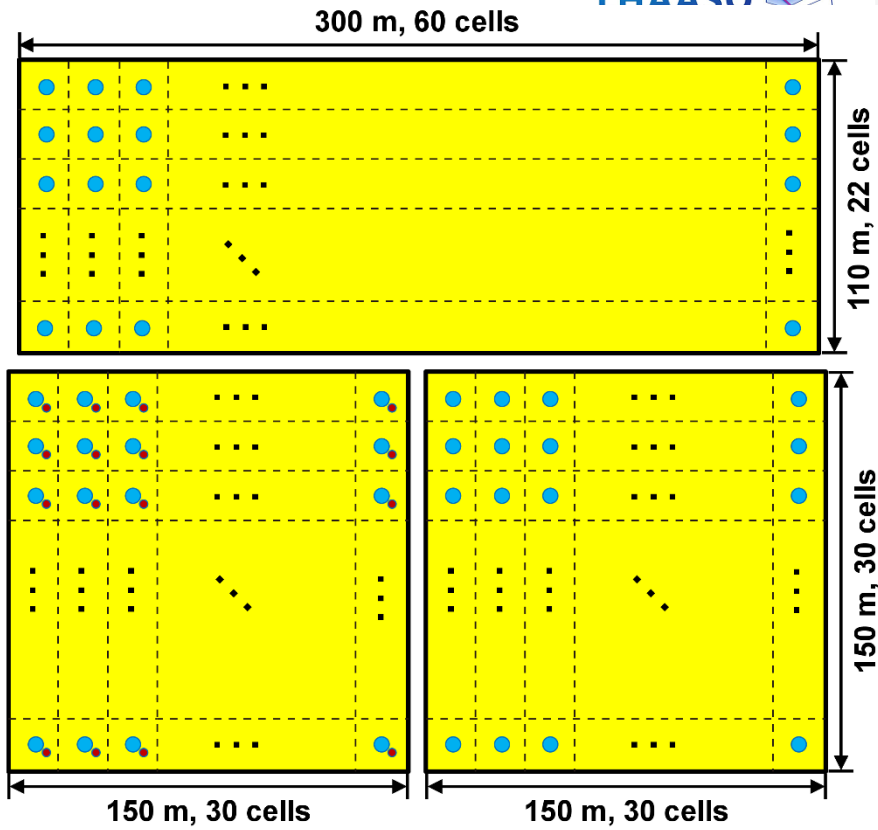
- Background introduction
- Methods
- Parameters & results & discussions
- Summary & outlook


cosmic rays +
neutrinos
p
$\pi^{\pm}$
ν
π^0
p
$e^{\pm}$
γ
cosmic rays
+ gamma-rays
GW
γ
ν
p
Cosmic Messengers

LHAASO-WCDA

- 3 pools
- 3120 cells
  - ✓ 2 PMTs/cell
- 1st pool
  - ✓ 900 cells





Background rejection

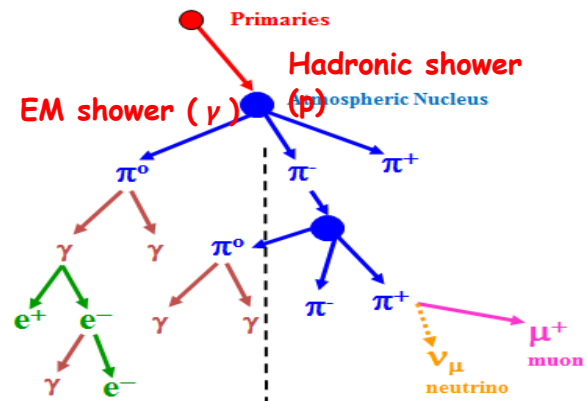
(*gamma/proton discrimination*)

is very important

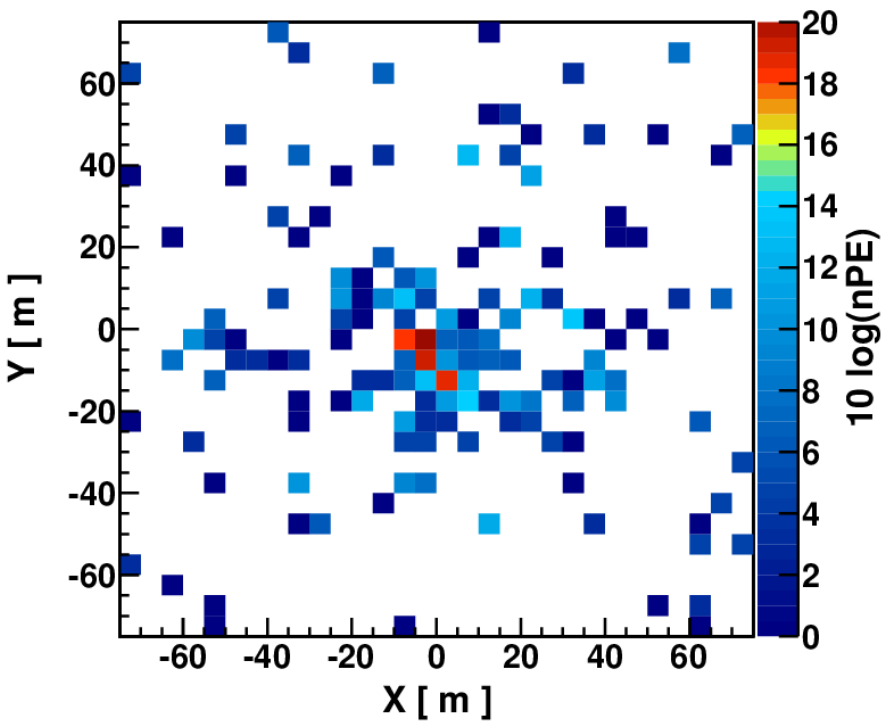
Mechanism & methods

Gamma

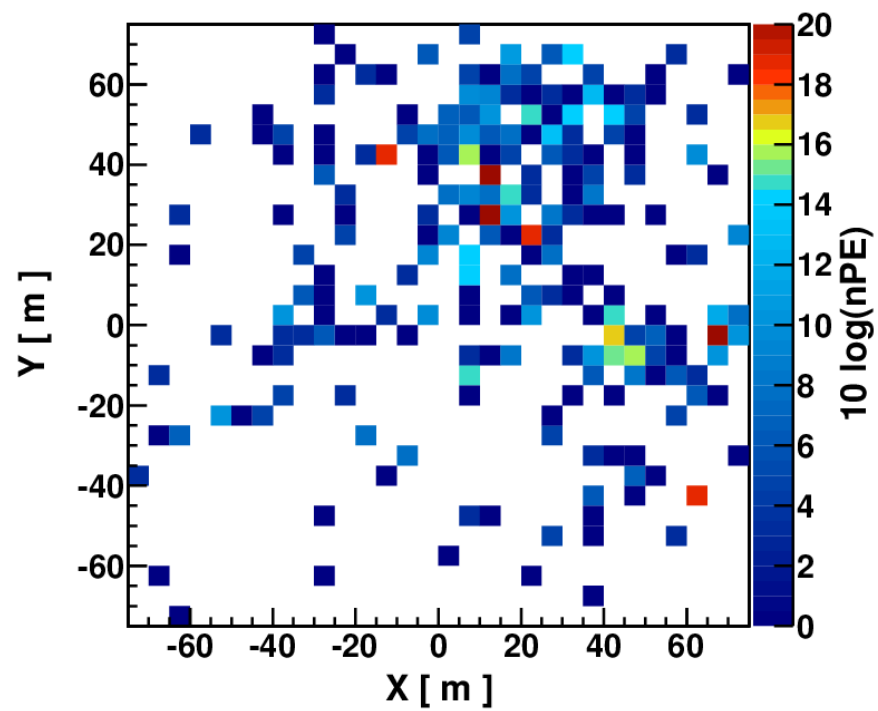
Proton



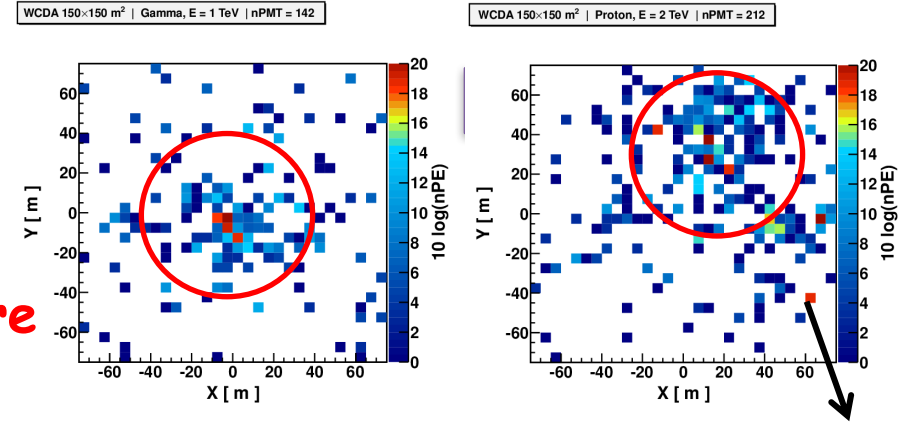
WCDA 150×150 m^2 | Gamma, E = 1 TeV | nPMT = 142



oton, E = 2 TeV | nPMT = 212



Discrimination variables



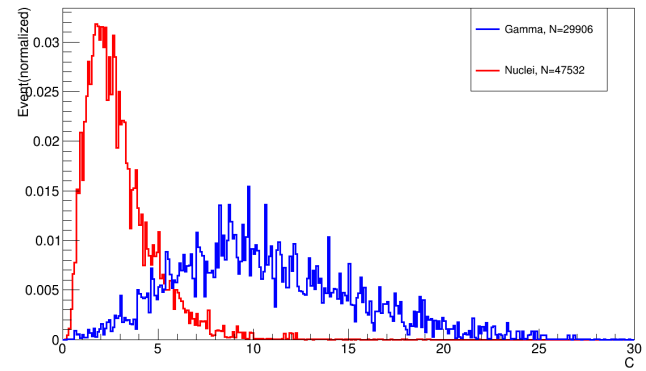
$cxPE_{45}$

Take advantages of muons & sub-core

1、 ***Compactness(C)*** $= \frac{nFit}{cxPE_{45}}$

$nFit$: ***number of PMTs involved in reconstruction***

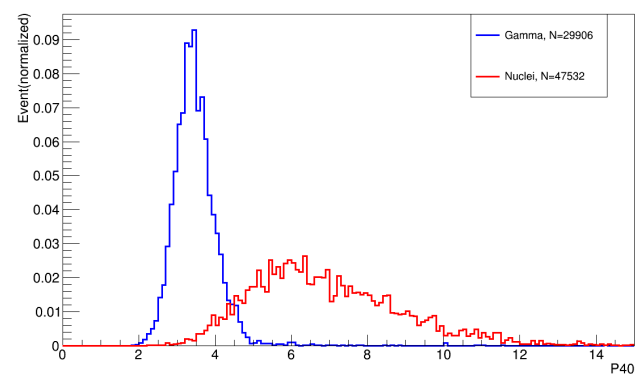
$cxPE_{45}$: maximum PE count outside the reconstruction core with a distance of 45m ;



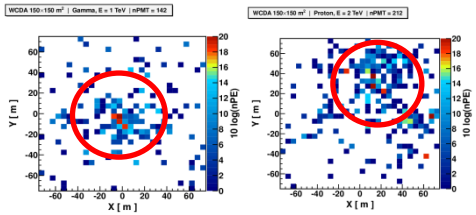
2、 **Density out(ρ_{40})** $= \frac{\sum PE_{40}}{\sum PMT_{40}}$

$\sum PE_{40}$: sum number of PEs outside core within 40m

$\sum PMT_{40}$: sum number of fired PMTs outside shower core with distance of 40m

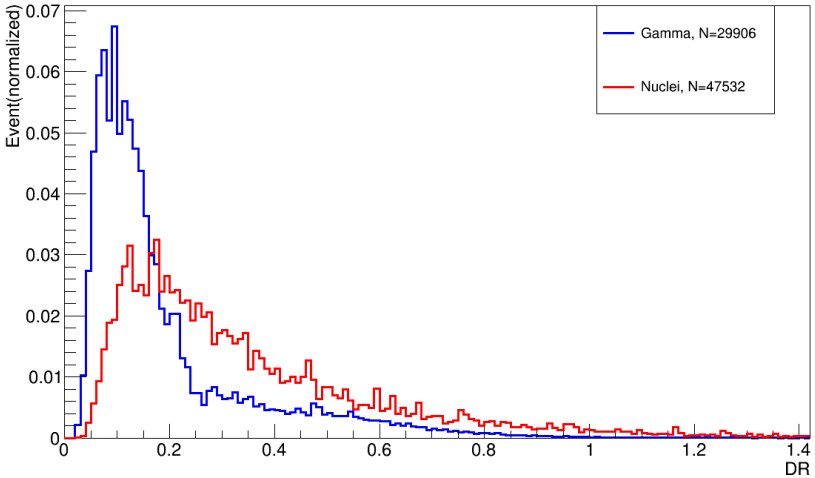


Discrimination variables



3、 Density raito(DR)

$$\mathrm{DR} = \frac{\sum PE_{50} / \sum PMT_{50}}{\sum PE_{10} / \sum PMT_{10}}$$

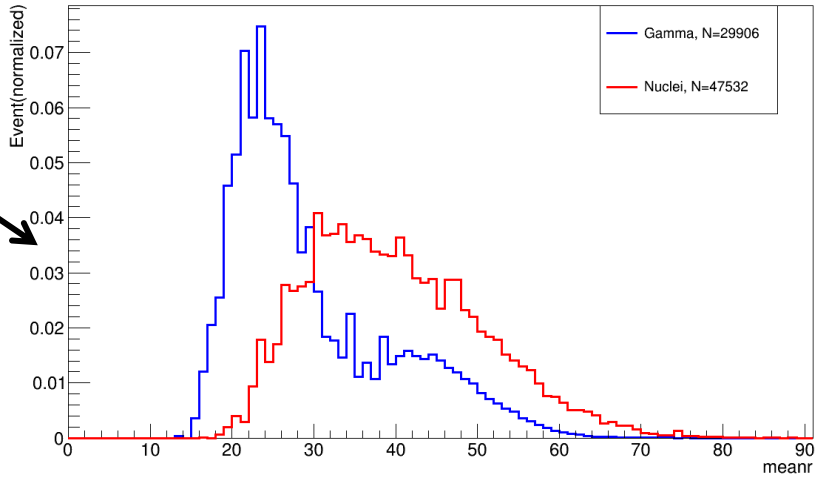


Use the lateral distribution of shower

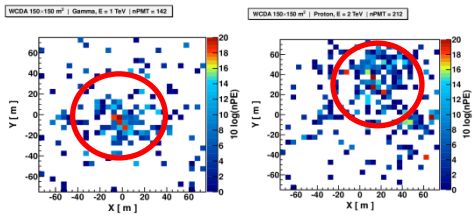
4、 <R> = $\frac{\sum PE_i R_i}{\sum PE_i}$

PE_i: number of nPE of the ith fired PMT ;
R_i : distance between the ith fired PMT to shower core

V. Grabski et al 32nd ICRC

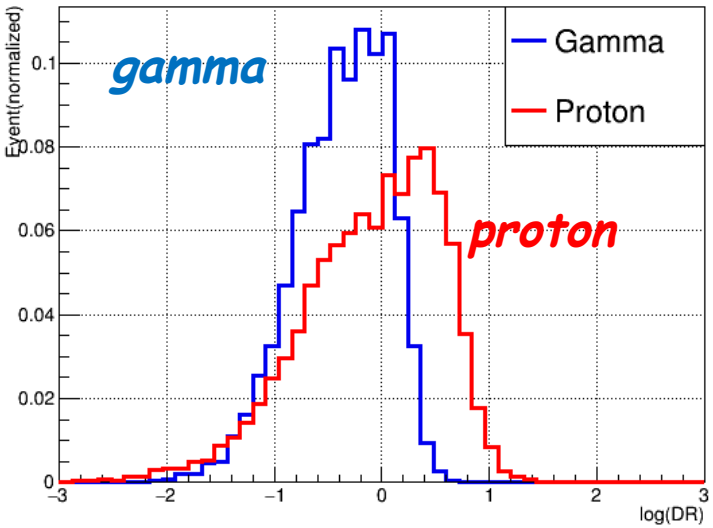


Discrimination variables



5、 r70

Minimum radius which contained 70% PE charge of the shower



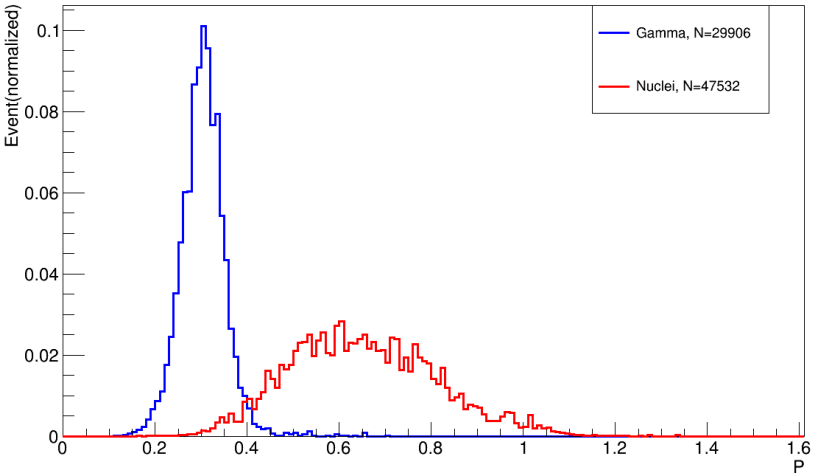
6、 P = $\frac{1}{N}\sum_{i=0}^{N}\frac{(\zeta_i-\langle\zeta_i\rangle)^2}{\sigma_{\zeta_i}{}^2}$

$\zeta_i = \log\mathbf{10}(Q_i)$

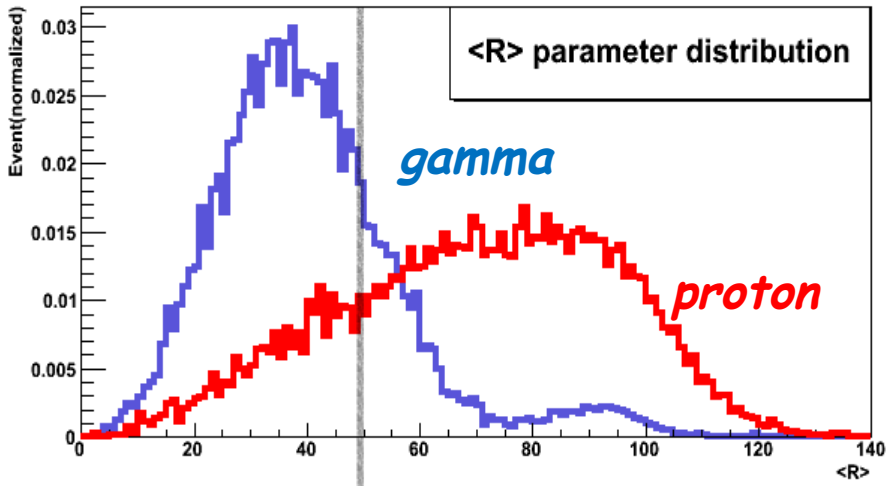
$\langle\zeta_i\rangle$: average ζ_i in an annulus with width

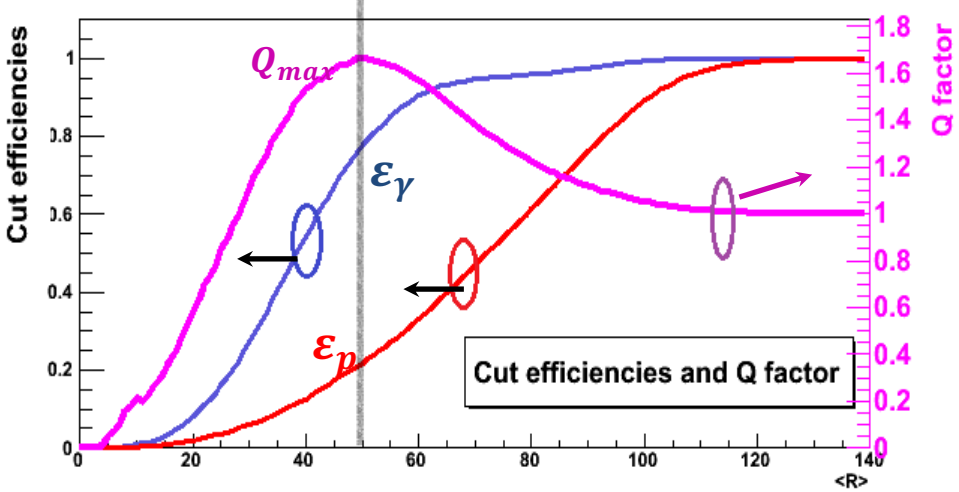
Of 5m

A. U. Abeysekara et al APJ 2017



Judgment parameter——Q factor





$$Q = \frac{\varepsilon_\gamma}{\sqrt{\varepsilon_p}}$$

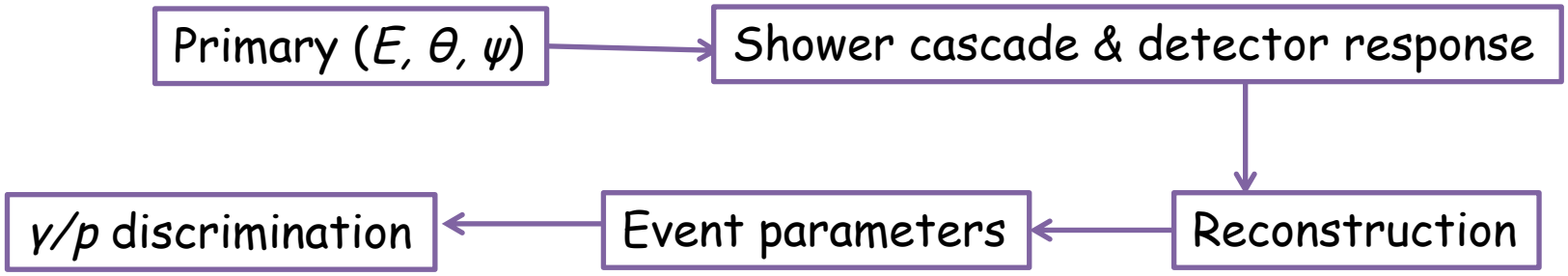
ε_γ: efficiency
ε_p: misidentification rate

For example,

Q_{max}=1.68
ε_γ=78%
ε_p=20%

Background rejection ratio is 80% , keeping 78% of signal

G/P separation procedure



MC simulation

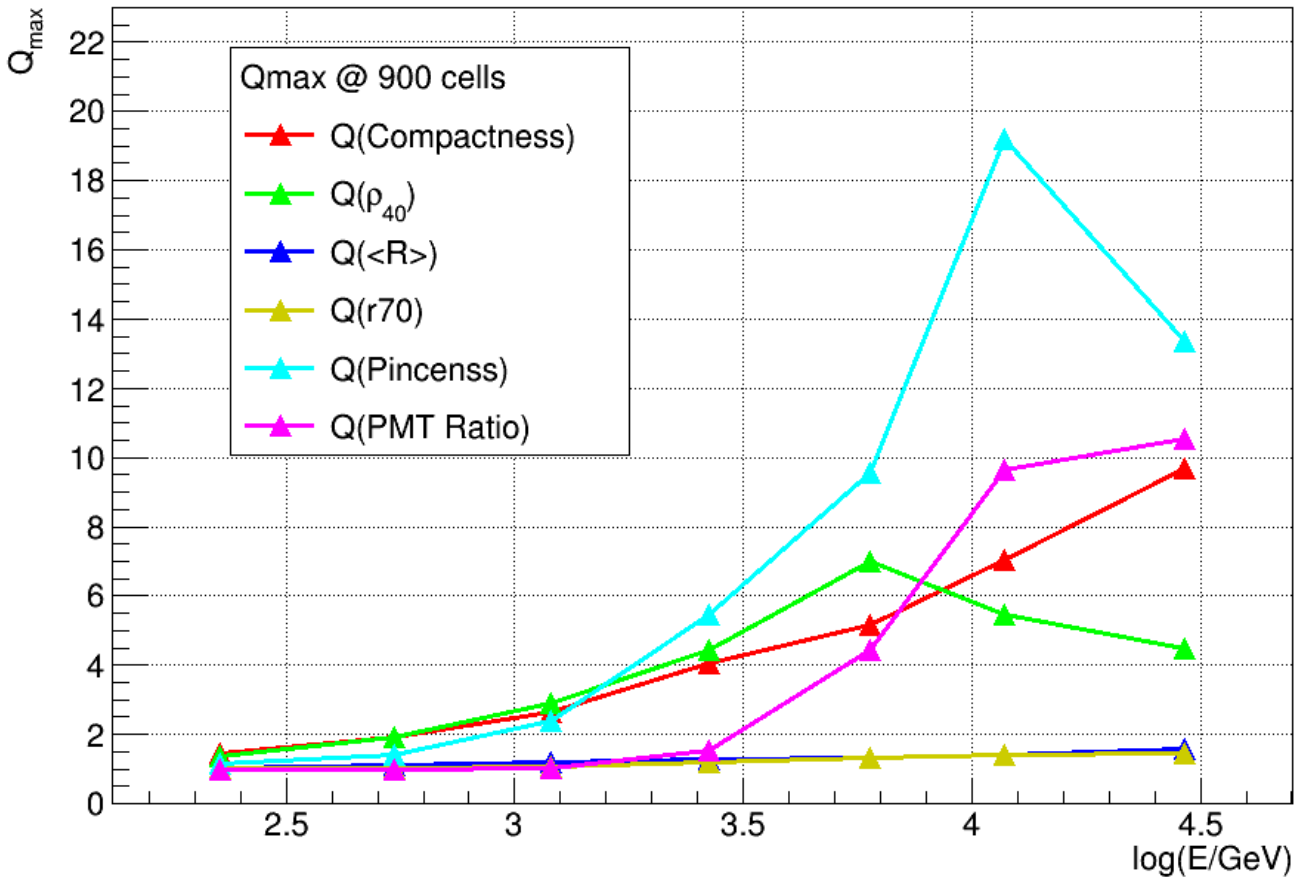
- Site: Daocheng @ 4410m a.s.l.
- Code: Corsika 6720 + QGSJET-II (*was EPOS*) + GHEISHA
- Primary: point source (γ)
  Spectrum & Flux: Crab measured by *HEGRA (astro-ph/0407118)* ⓟ2.05⊠10^{-6} $(E/GeV)^{-2.62}$ $cm^{-2}s^{-1}GeV^{-1}$.
  Energy:
  4 segments: 100 GeV-1-10-100 TeV
- Primary: background(*p*)
  Spectrum & flux: *J.R. Hoerandel, Astroparticle Physics 19 (2003) 193-220*
  Energy: same, but min energy = min(10,1.1⊠A).
- Energy cuts:
  50 (hadron), 50 (muon), 0.3 (electron), 0.3 (photon / pion) MeV

Events selection

cuts：

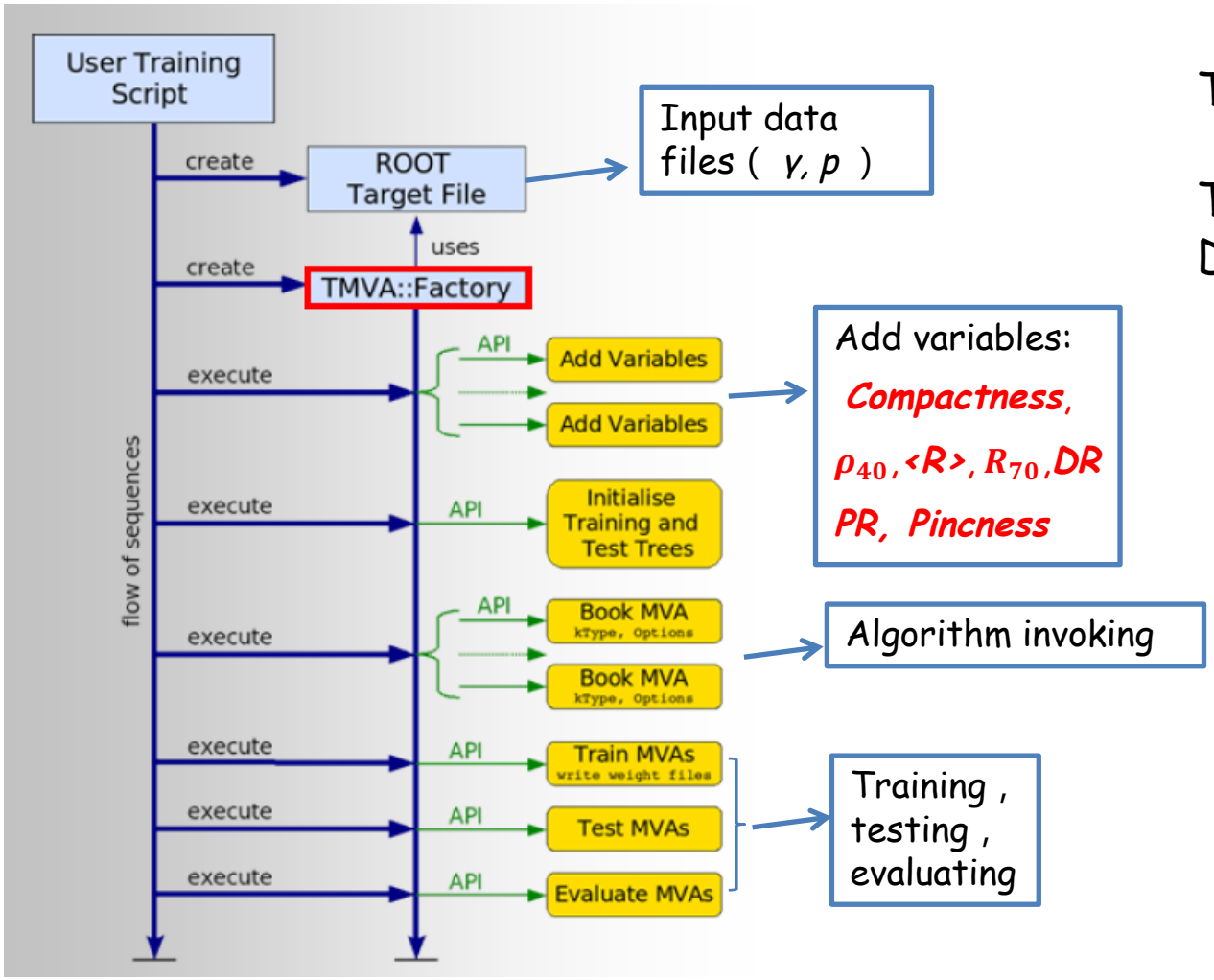
- reconstruction core fall in 150x150 m^2
- *zenith* : 0~45 °
- Gamma efficiency >50%

Single parameter result



TMVA

Analysis steps :



TMVA:

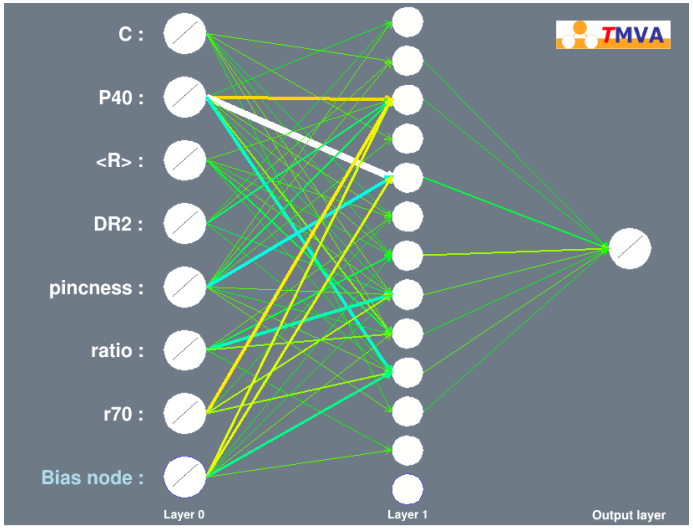
Toolkit for Multivariate Data Analysis with ROOT

Algorithms :

- Artificial Neural Network (ANN)
- Boosting decision trees (BDTG)

Artificial Neural Network (ANN)

1、 Input variable & neural network layout



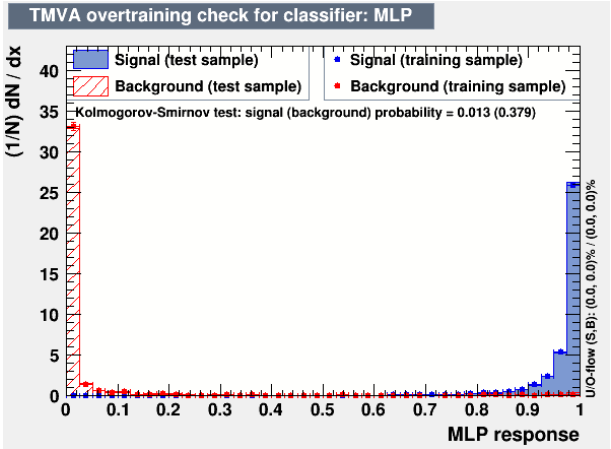
2、 Input data

Signal: 28856

Background: 27040

Half of them are used in training, the other half for testing

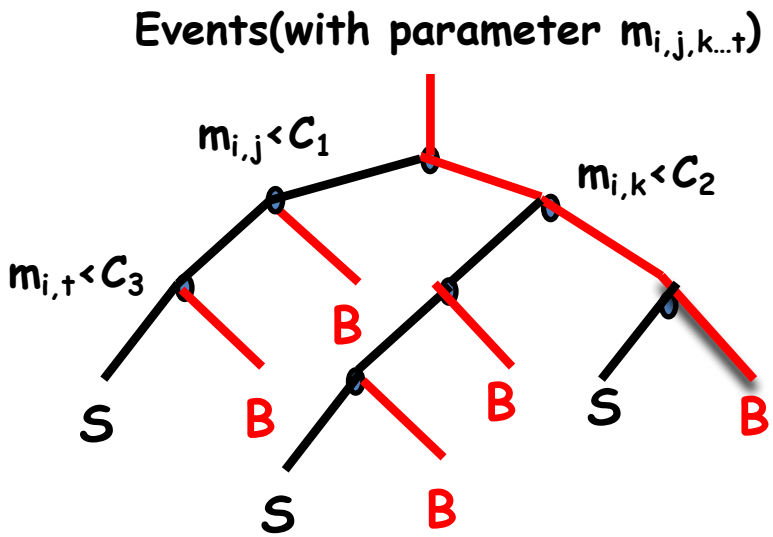
3、 ANN result (200 < *nfit* < 300)



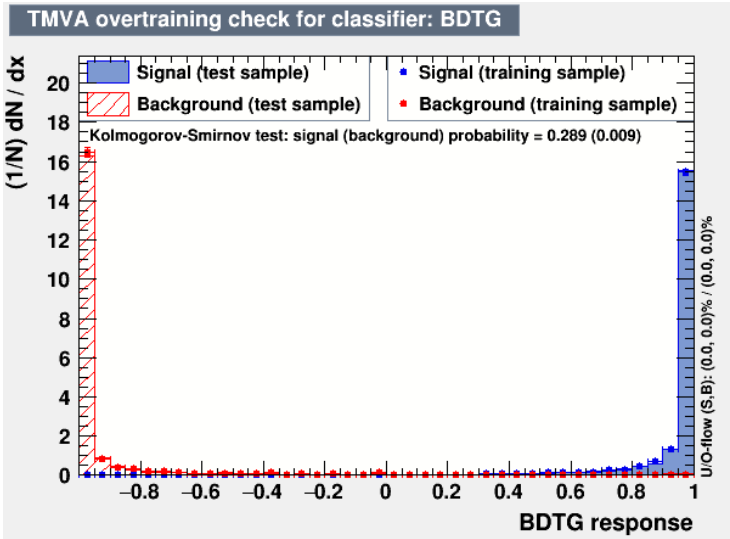
	P	C	MLP
Q max	9.55	5.18	7.78
ε_γ	62.9%	57.2%	54.7%
ε_p	0.4%	1.2%	0.5%

1.5 times of C !

Gradient Boost Dicision Tree(BDTG)



BDTG result (200 < nFit < 300)



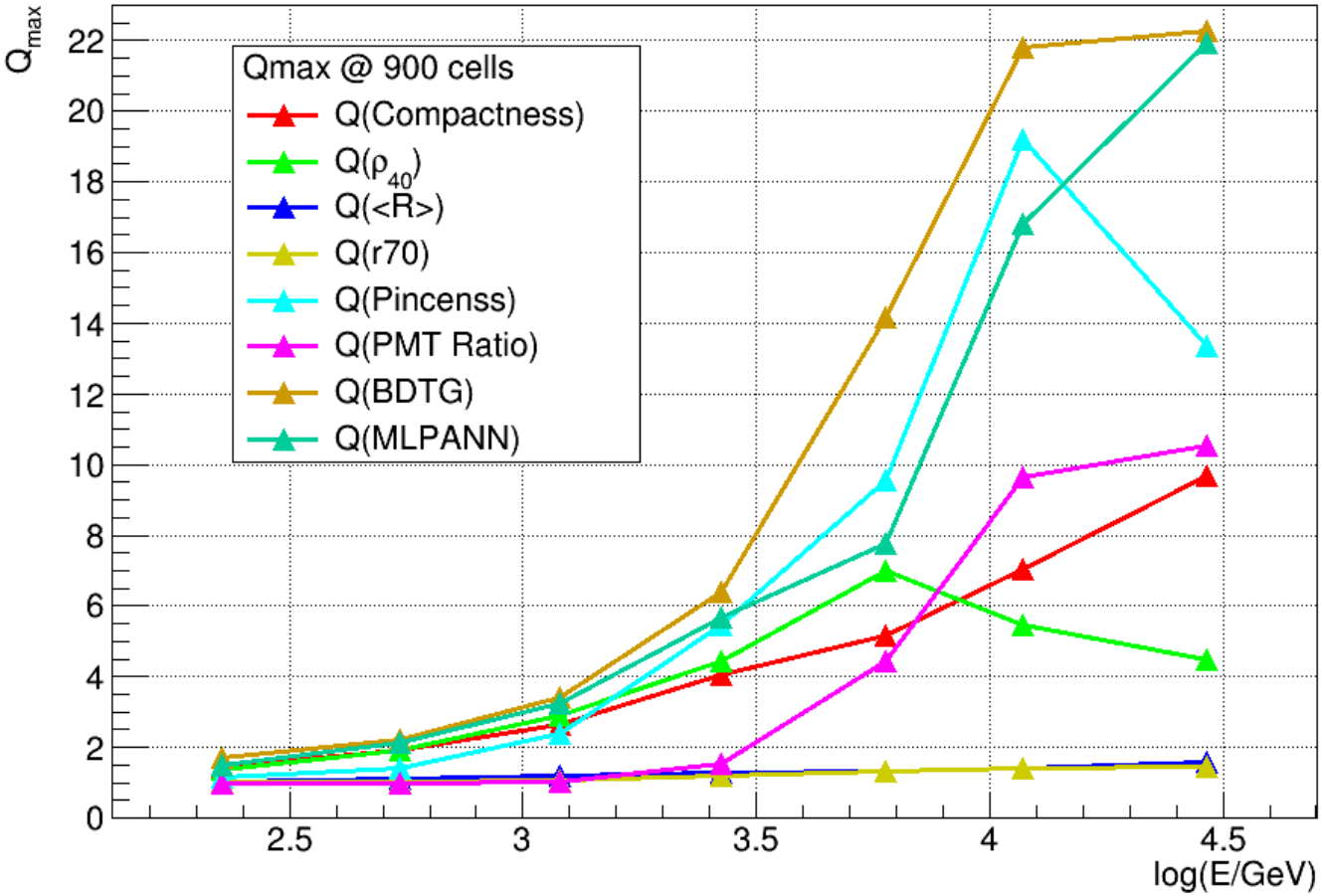
Input data

- ✓ Signal: 28856
- ✓ Background: 27040
- ✓ Half of them are used in training,
- ✓ the other half for testing

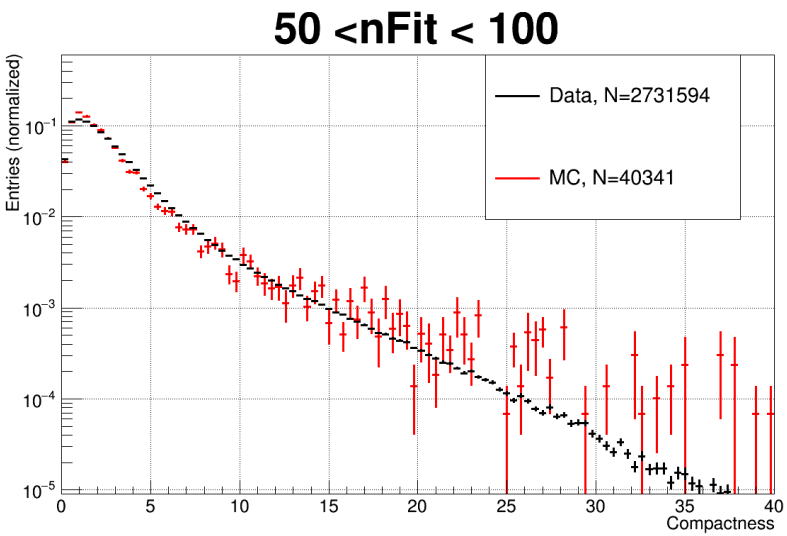
	P	C	BDTG
Q max	9.55	5.18	21.82
ε_γ	62.9%	57.2%	81.6%
ε_p	0.4%	1.2%	0.1%

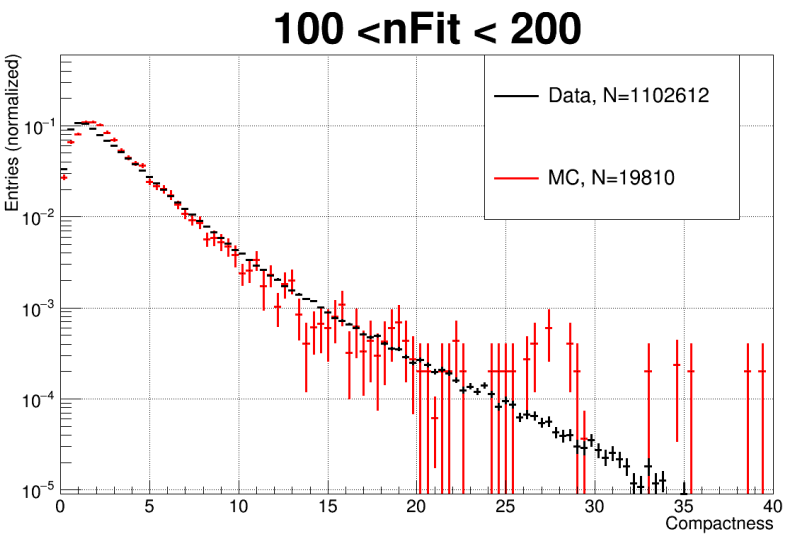
4.2 times of C !

Qmax vs. Energy

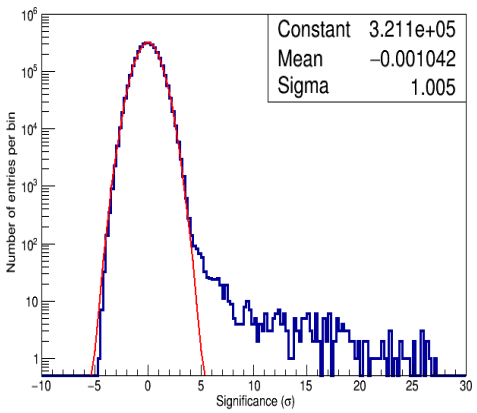


Simulation vs Data

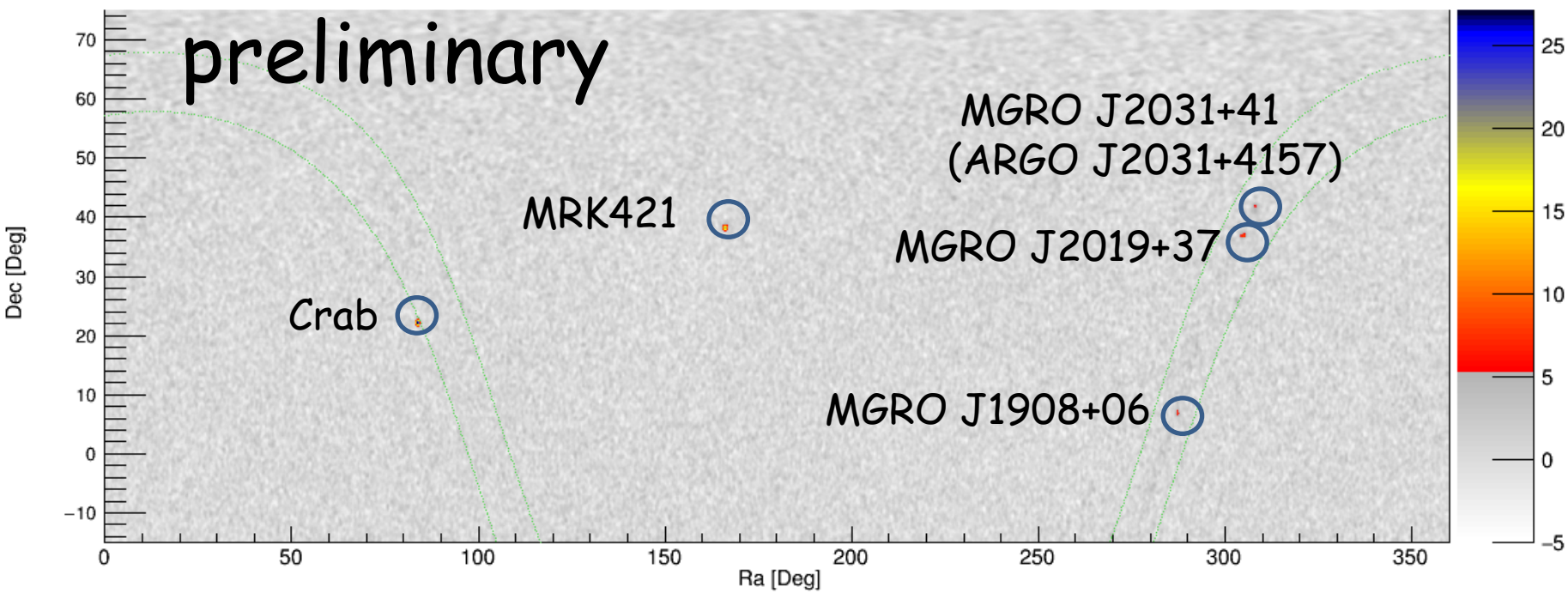




WCDA#1 Sky Map



- Data up to 2019-06-30
- Live time: 57 days
- Crab: 27σ



Summary & Outlook

- LHAASO-WCDA can separate gamma/proton well
- Energy >3TeV reject 99% of the background events
- Multivariate analysis can improve Q factor value obviously.
- Compactness distribution of data is basically agree with MC's
- There are still some problems need to be resolved , with more data, we hope better results

THANK YOU !

Back up

Training & testing sample

nfit bin	Number of signal	Number of background
10-20	6172	52322
20-50	69964	140948
50-100	141316	128340
100-200	113468	78656
200-300	28856	27040
300-500	12520	14468
500-900	1424	1612

input variable correlations