Exploring a new g/h separation models on the HAWC Observatory

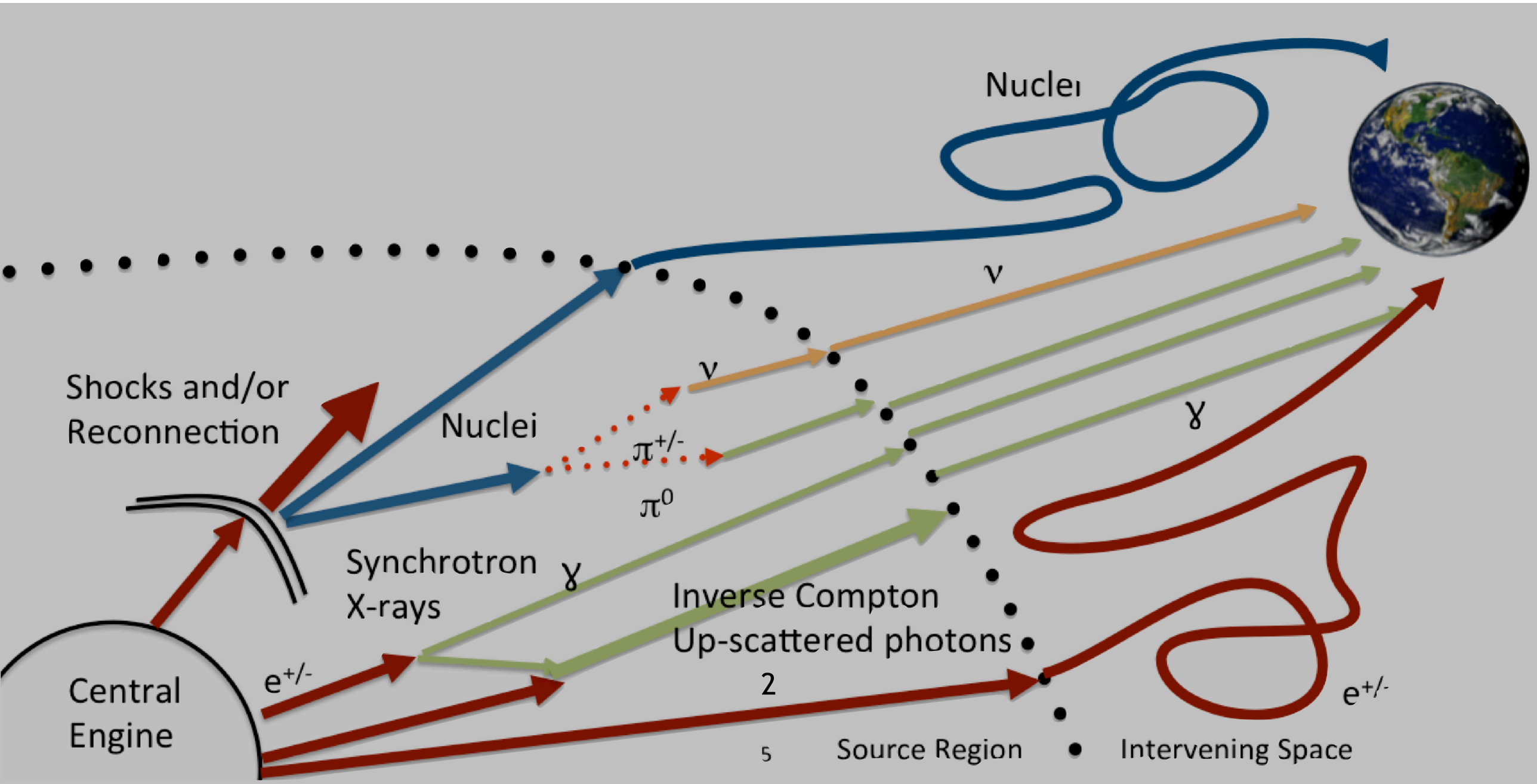
Tomás Capistrán Rojas
INAOE-MSU-HKU

Meeting of the Cosmic Rays Section of the Mexican Physical Society

November 26th, 2019

Why study Gamma-rays?

Gamma-rays is not deflected by magnetic fields.



Pretz, J. (2015), Highlights from the High Altitude Water Cherenkov Observatory.

Gamma-ray Observatories

Wide-field
Continuous Operation

TeV Sensitivity

Fermi
AGILE
EGRET

HAWC
ARGO
Milagro

VERITAS
HESS
MAGIC
FACT

Pretz, J. (2015), Highlights from the High Altitude Water Cherenkov Observatory.

Mapping the Northern Sky in High-Energy Gamma Rays

HAWC
High Altitude Water Cherenkov
Gamma-Ray Observatory

HAWC Observatory

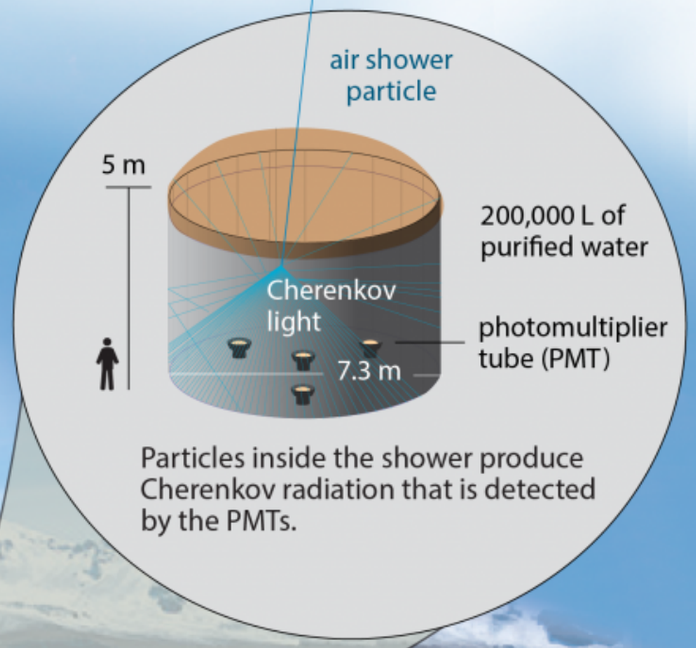
HAWC operates day and night, providing a large field of view for the observation of the highest energy gamma rays.

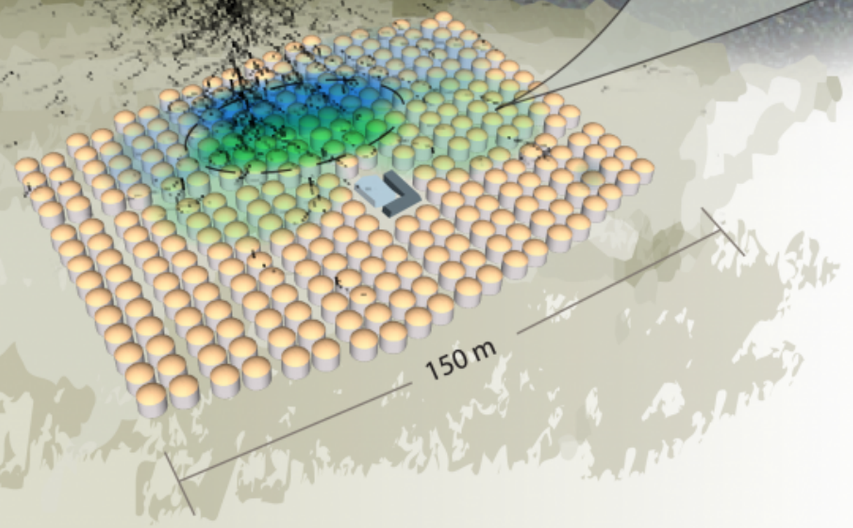
Pico de Orizaba
(5,626 m)

HAWC is located at 4,100 m above sea level, covering an area of 20,000 m^2.

Water Cherenkov tank

HAWC comprises an array of 300 tanks that record the particles created in gamma-ray and cosmic-ray showers.

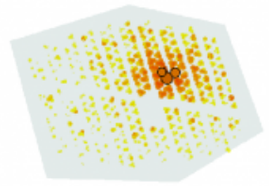




Gamma rays vs cosmic rays

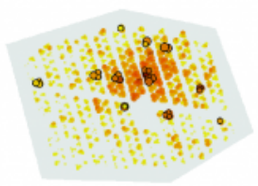
HAWC selects gamma rays from among a much more abundant background of cosmic rays.

gamma-ray shower



"hot" spots concentrate around the core

cosmic-ray shower



"hot" spots are more dispersed

HAWC
High Altitude Water Cherenkov
Gamma-Ray Observatory

Mapping the Northern Sky in High-Energy Gamma Rays

HAWC Observatory

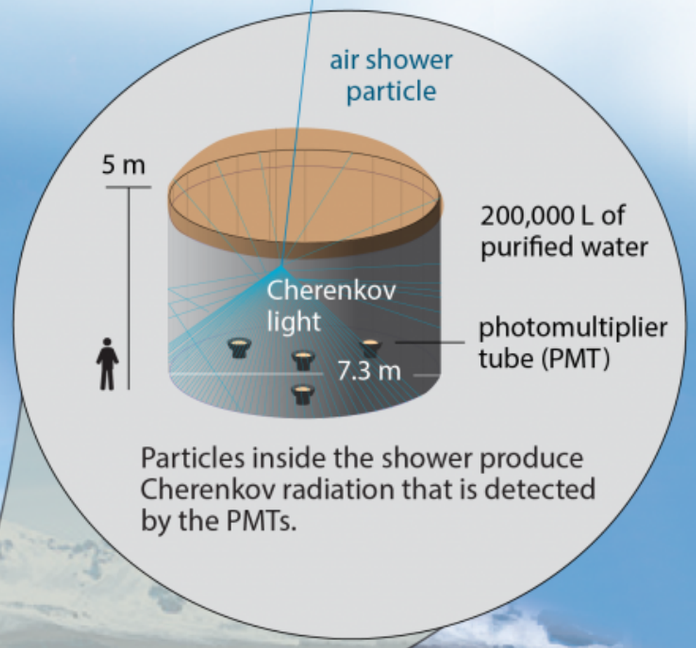
HAWC operates day and night, providing a large field of view for the observation of the highest energy gamma rays.

Pico de Orizaba
(5,626 m)

HAWC is located at 4,100 m above sea level, covering an area of 20,000 m^2.

Water Cherenkov tank

HAWC comprises an array of 300 tanks that record the particles created in gamma-ray and cosmic-ray showers.



Particles inside the shower produce Cherenkov radiation that is detected by the PMTs.

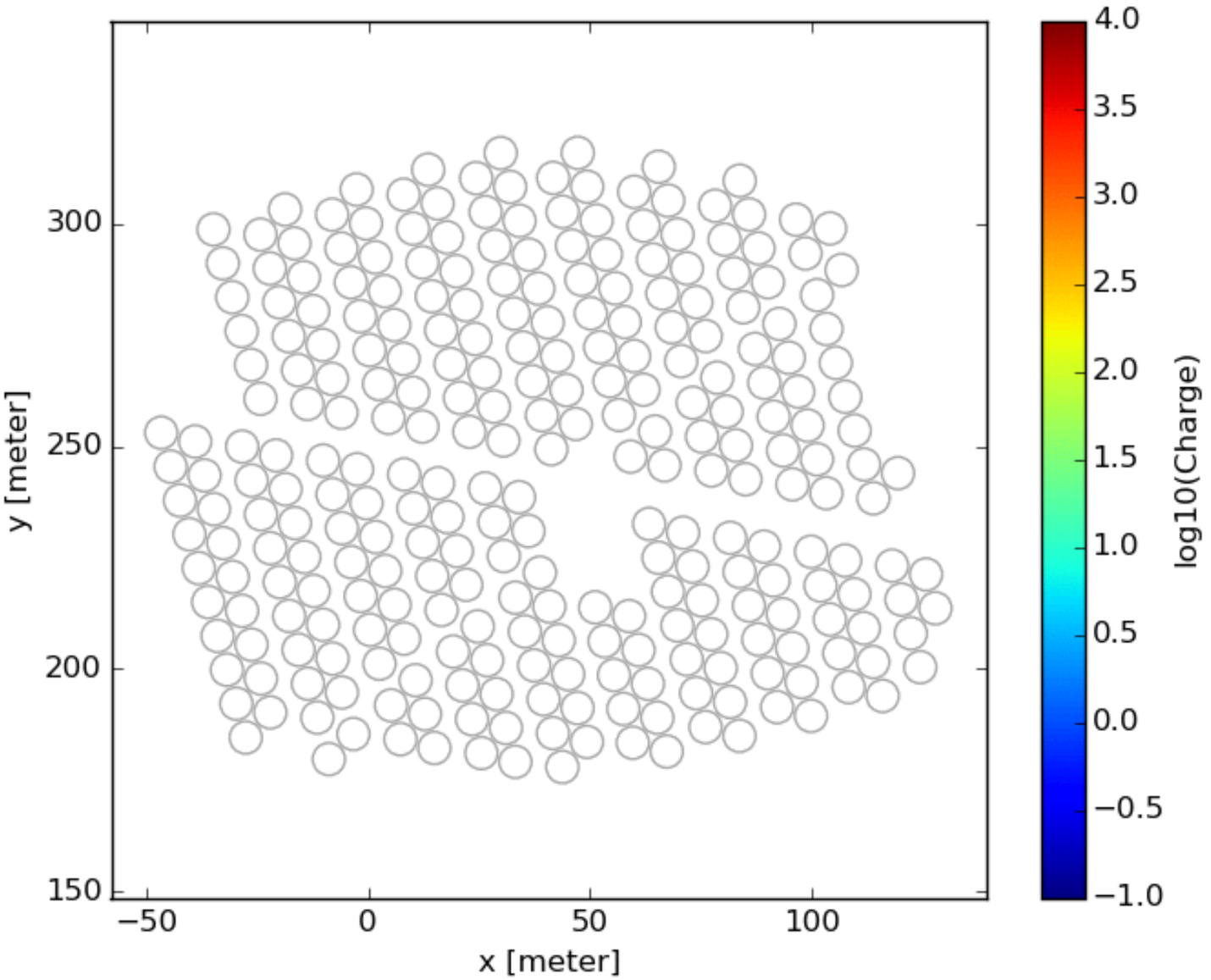
Gamma rays vs cosmic rays

HAWC selects gamma rays from among a much more abundant background of cosmic rays.

Main background: Hadronic cosmic ray

- **Crab nebula: 400 photons/day**
- **Background: 15,000 cosmic ray/second**

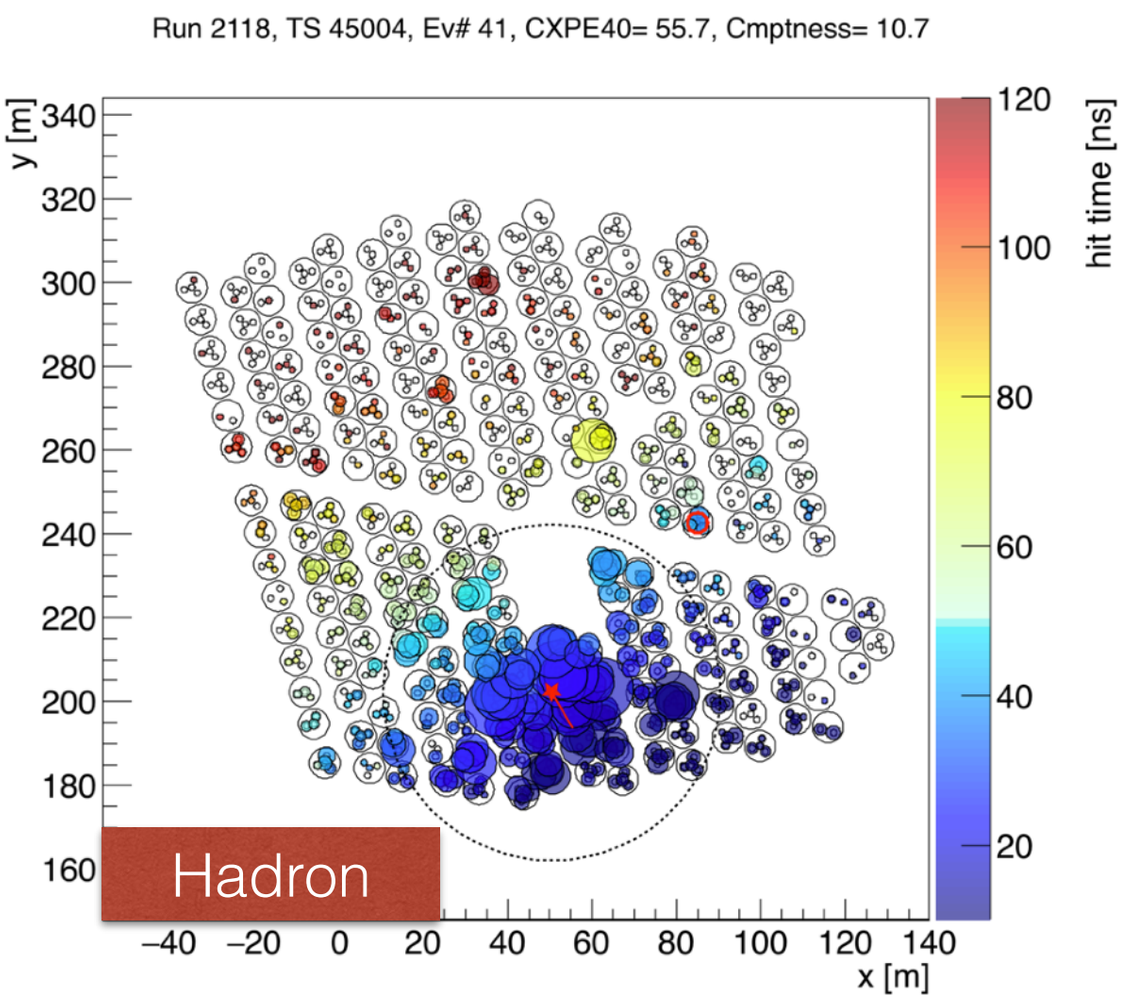
Event simulation detected by HAWC

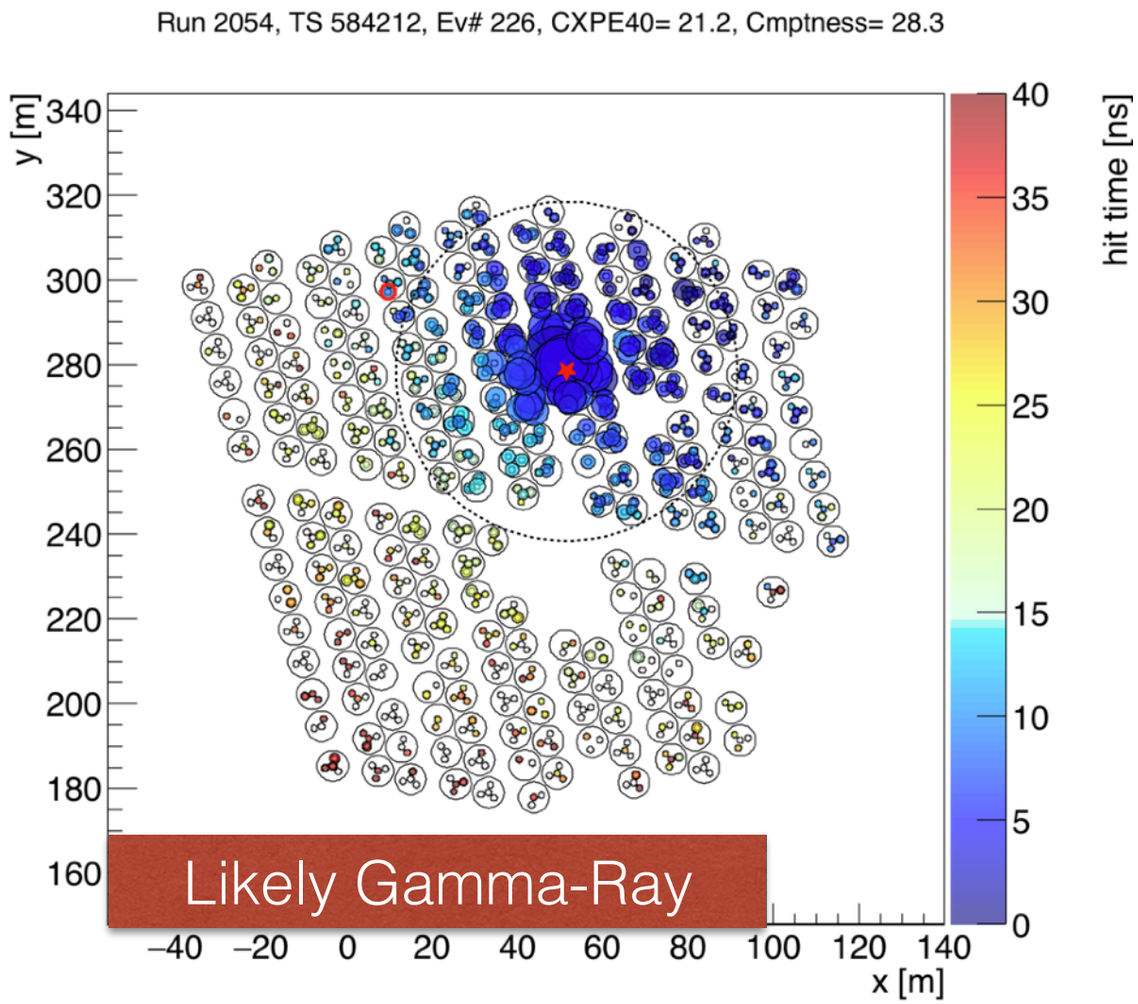


A. Timing information allows us determining where the particle comes.

B. Energy deposition in each PMT:
   - Primary particle energy.
   - The shower core.
   - Gamma or Hadron?

Gamma Vs Hadron

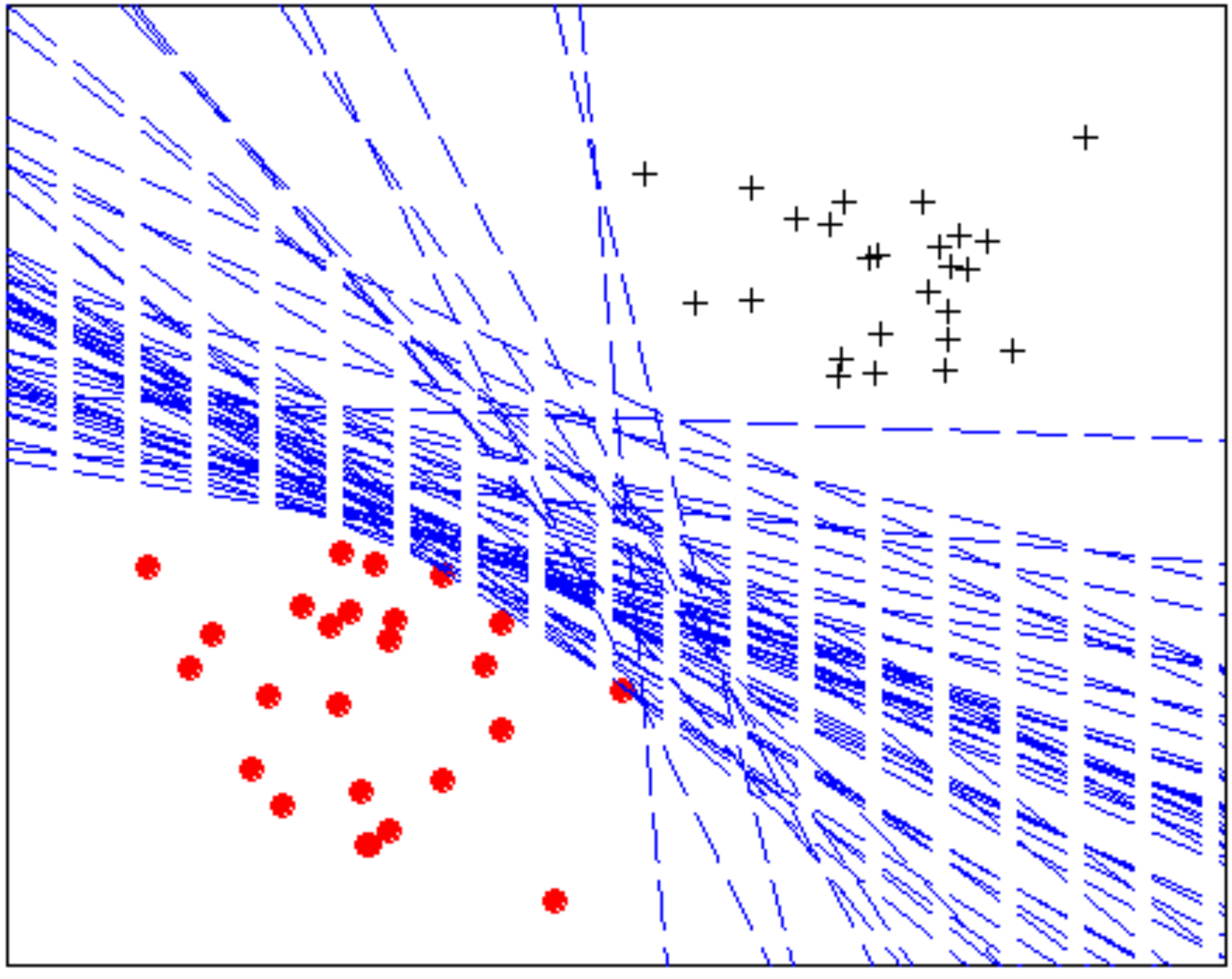




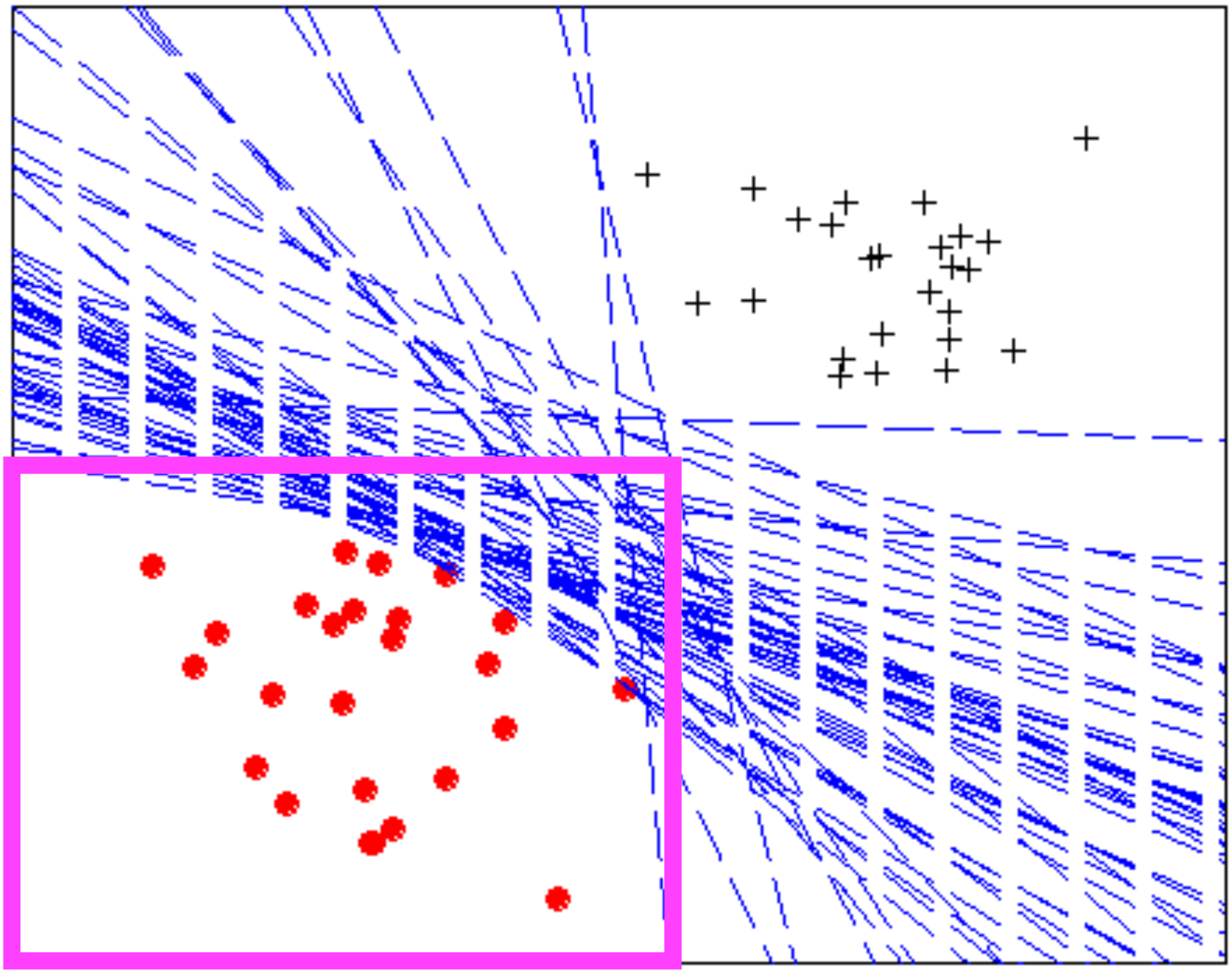
Task: Distinguishing between gammas and hadrons

http://www.hawc-observatory.org/observatory/ghsep.php

How recognize the particle?



Rectangle cut



Standard cuts

It is the official method in HAWC Observatory. It can describe as rectangle cut.

$$PINC <= Cut_{PINC} \;\&\&\; LiC <= Cut_{LiC}$$

Where

$$LiC = Log_{10}(\frac{CxPE40}{nHitSP20})$$

$$PINCness = \frac{\sum_{i=0}^{N} \frac{(\log q_i - \bar{\log} q_i)^2}{\sigma^2_{\log q_i}}}{N}$$

Bins:

1. **fhit:** nHitSP20/nChAvail

The fraction of the PMTs hit

2. **ebin:** logNNenergyV2

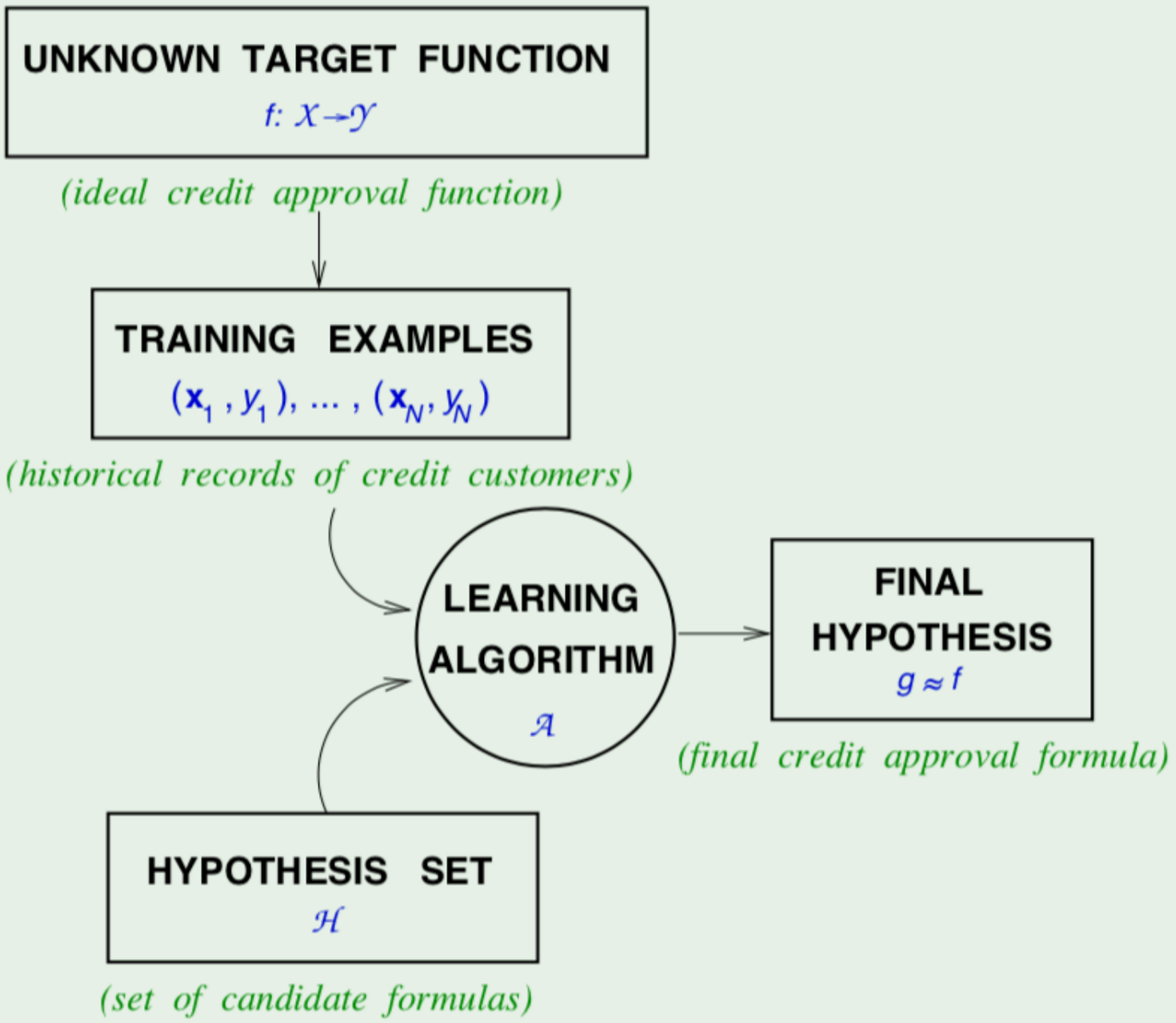
Energy estimator using a Neural Network

fhin	min fhin	max fhin
0	4.4%	6.7%
1	6.7%	10.5%
2	10.5%	16.2%
3	16.2%	24.7%
4	24.7%	35.6%
5	35.6%	48.5%
6	48.5%	61.8%
7	61.8%	74.0%
8	74.0%	84.0%
9	84.0%	100.0%

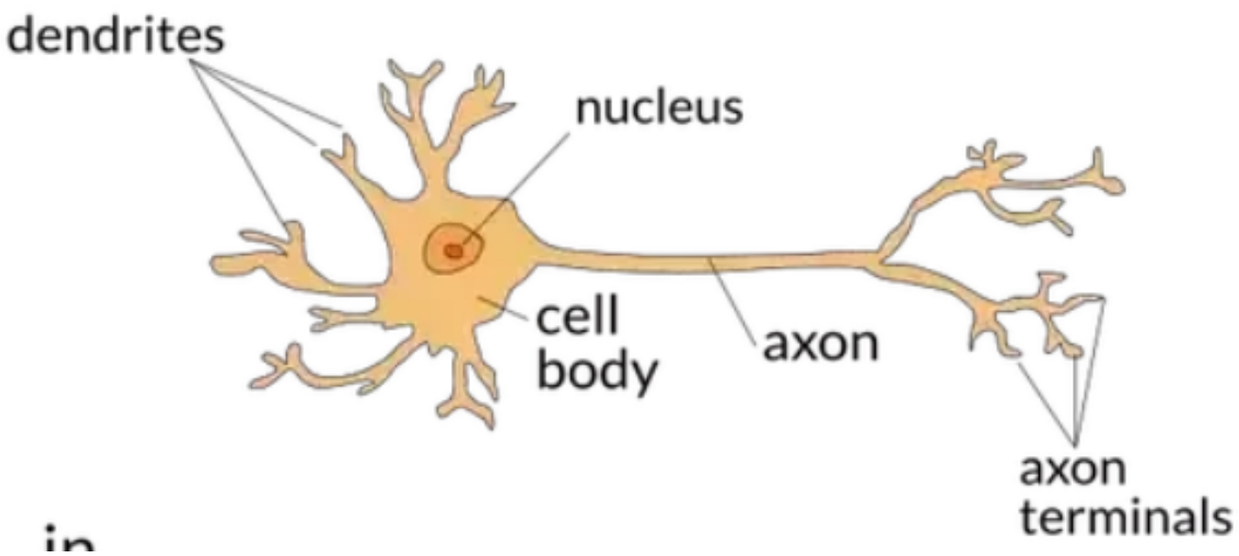
ebin	min ebin	max ebin	min ebin (GeV)	max bin (Gev)
0	2.50	2.75	316.23	562.34
1	2.75	3.00	562.34	1000.00
2	3.00	3.25	1000.00	1778.28
3	3.25	3.50	1778.28	3162.28
4	3.50	3.75	3162.28	5623.41
5	3.75	4.00	5623.41	10000.00
6	4.00	4.25	10000.00	17782.79
7	4.25	4.50	17782.79	31622.78
8	4.50	4.75	31622.78	56234.13
9	4.75	5.00	56234.13	100000.00
10	5.00	5.25	100000.00	177827.94
11	5.25	5.50	177827.94	316227.77

$\hat{E} = \log_{10}(E / 1\ \text{GeV})$

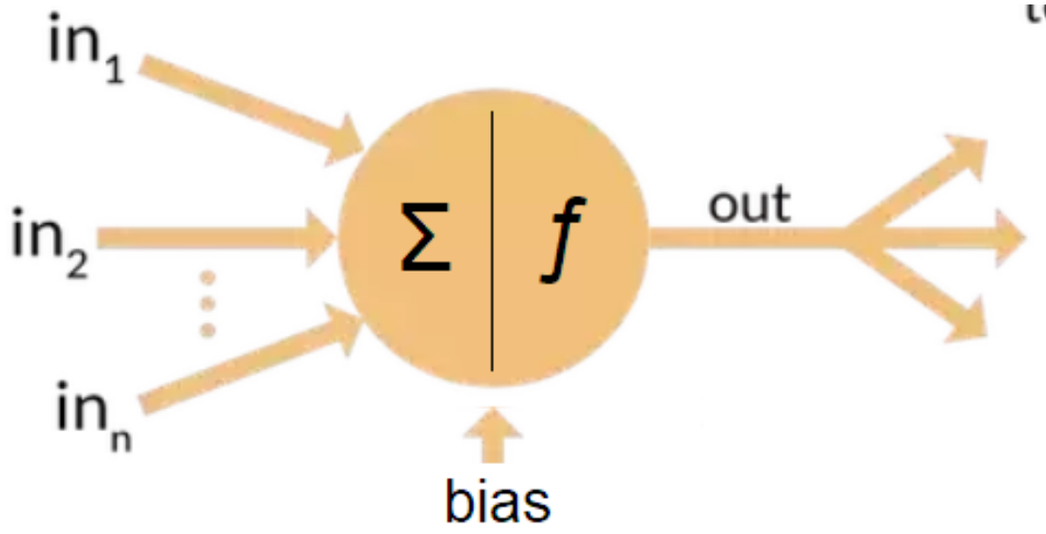
Learning from data

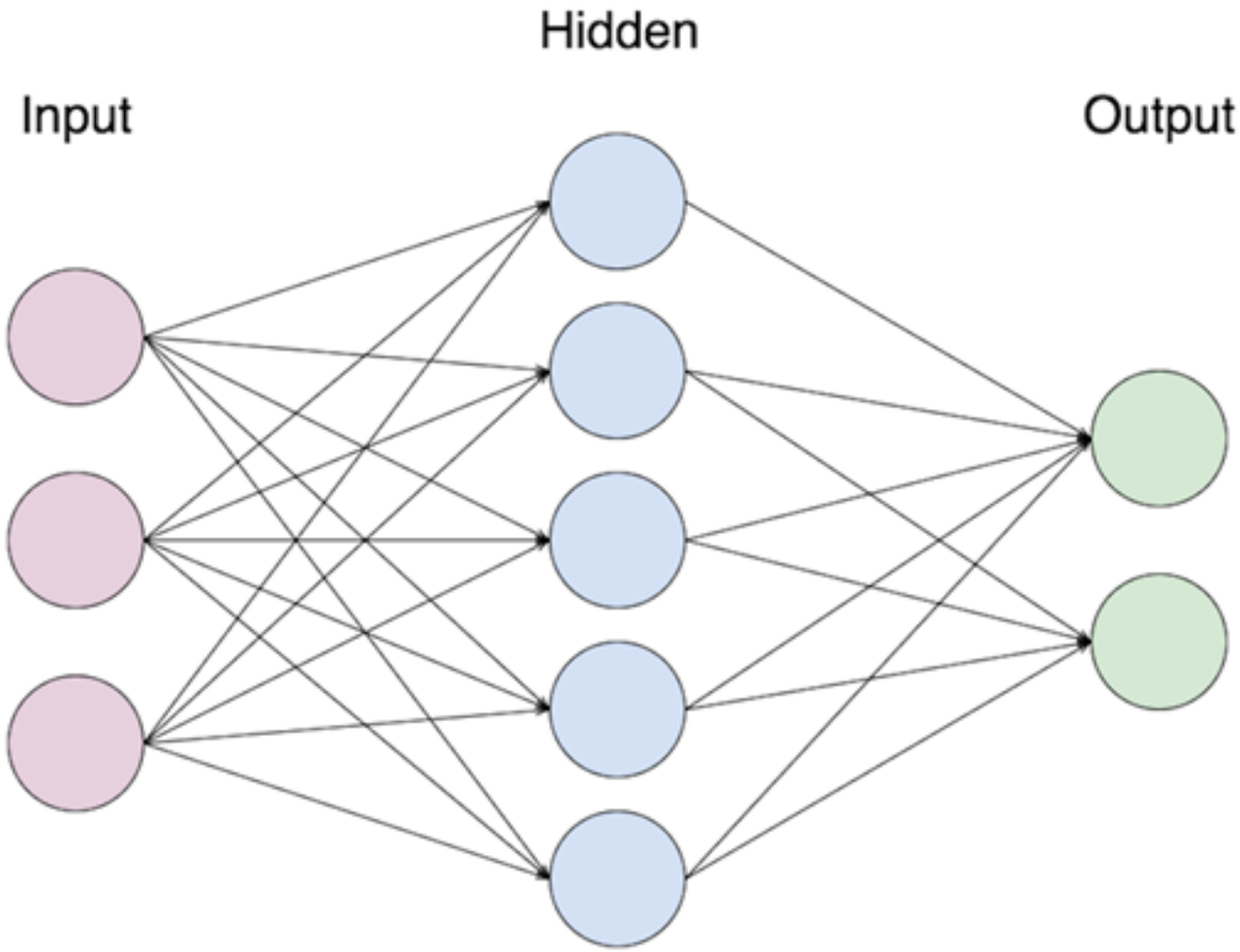


Neural Network

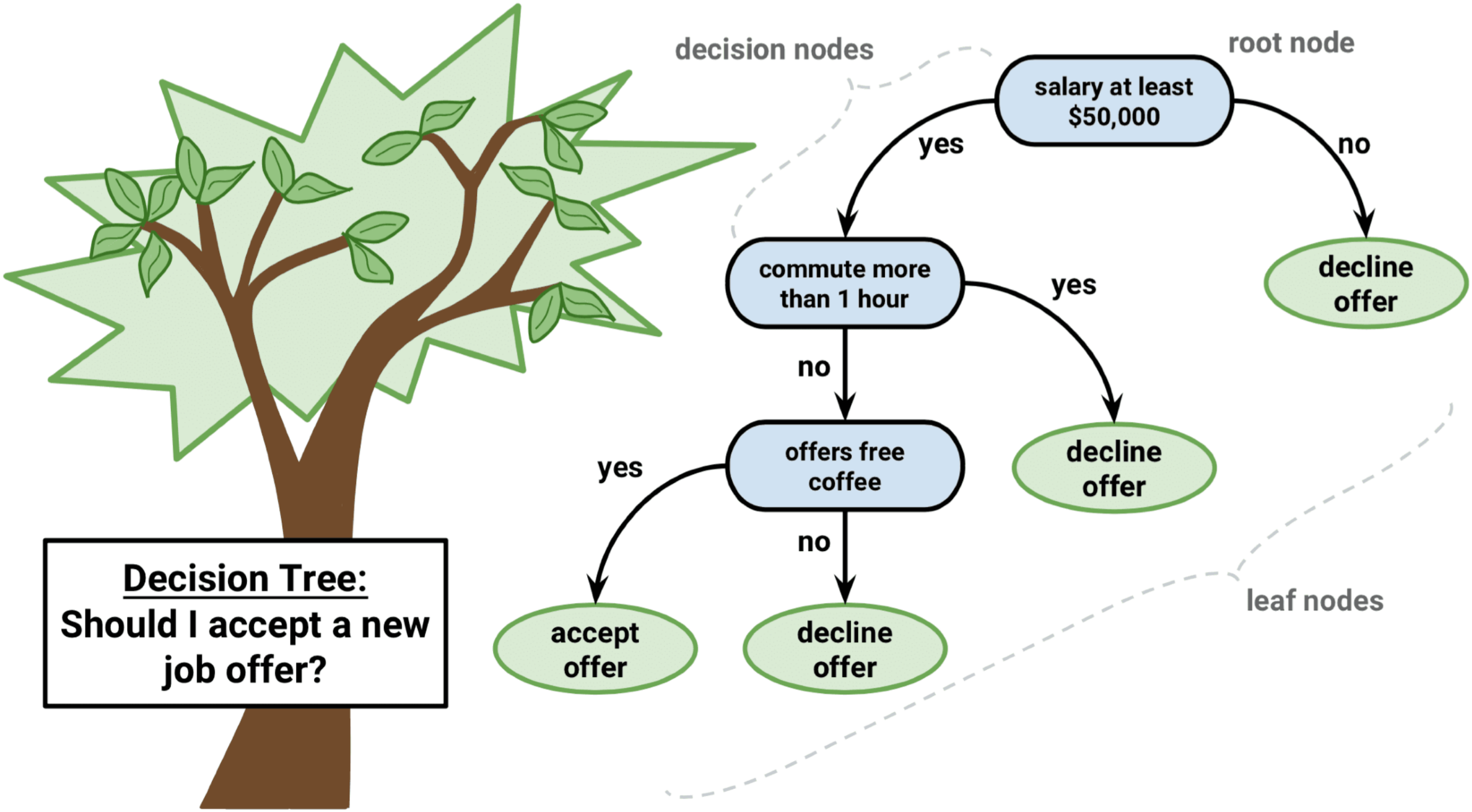


~

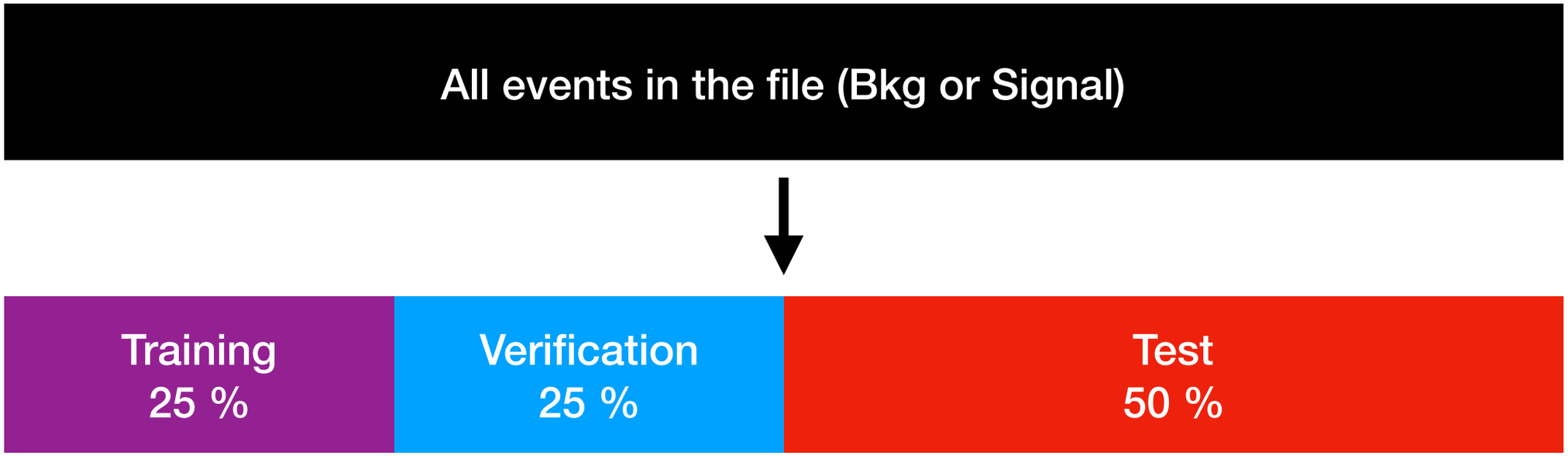




Boosted Decision Tree



Train a NN and BDT



A. Input parameters:

- LIC = log10(CxPE40 / nHitSP20)
- PINC
- logNNEnergyV2
- disMax
- LDFAmp
- LDFChi2
- fbin = nHitSP20 / nChAvail

MLT configuration:

Neural Network (NN):

B. **Architecture**: 7 : 10 : 10 : 1

C. Models trained with **TMVA**

Boosted Decision Tree (BDT):

B. **Model** with 500 tree

C. Models trained with python (**Xgboost** package)

Both Models

D. **Don't use Physical weight**

E. **Models trained**

- **`Low fbin:`** `0.044 to 0.162`
- **`Medium fbin:`** `0.162 to 0.485`
- **`High fbin:`** `0.485 to 1.000`

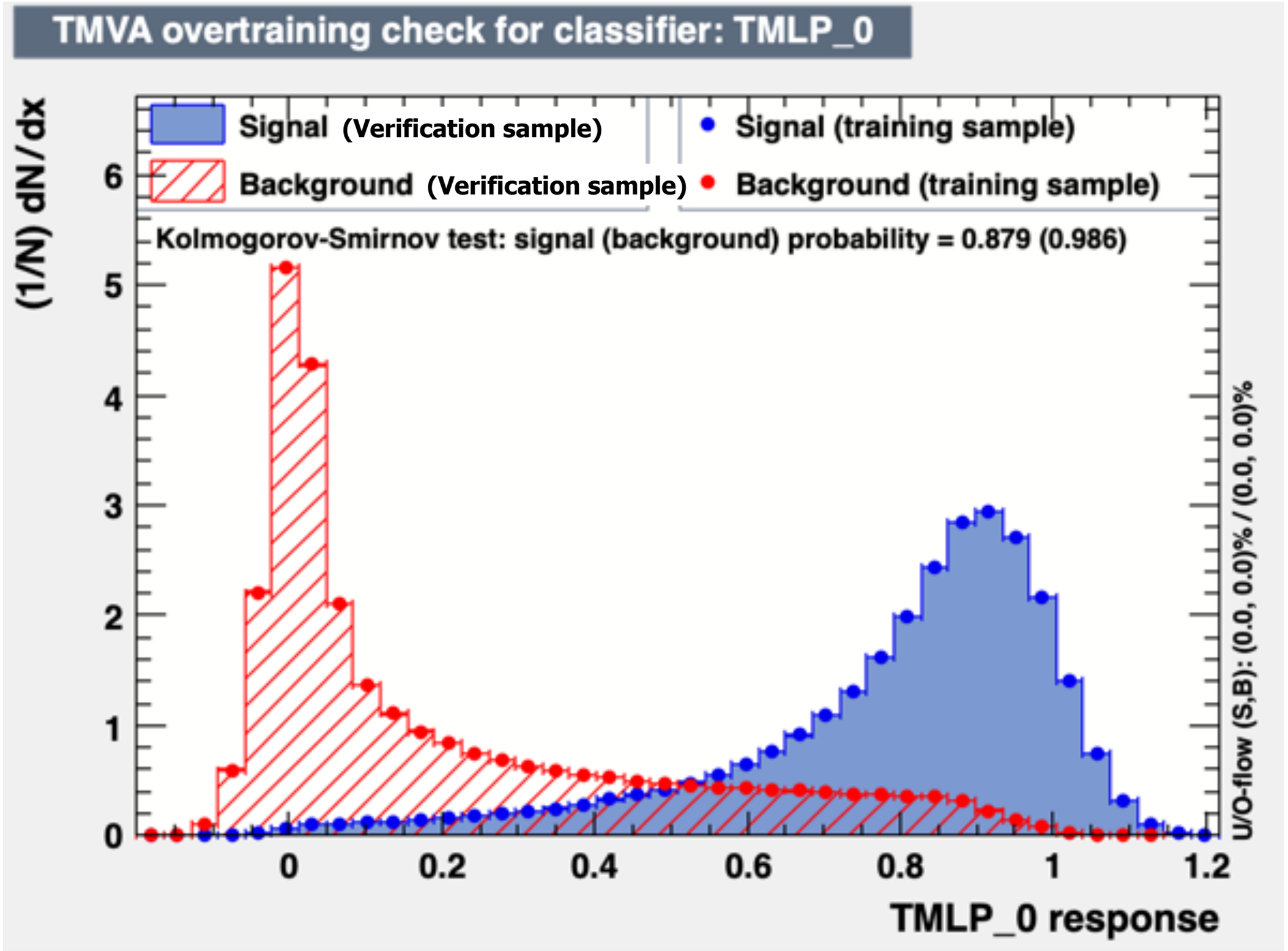
F. **Apply Quality cuts**

- rec.angleFitStatus==0
- rec.coreFitStatus==0
- rec.nChTot>=800
- rec.nChAvail>0.9*rec.nChTot

G. **Target**

- **`Signal`** `= 1`
- **`Background`** `= 0`

After Training: NN model for low fbin

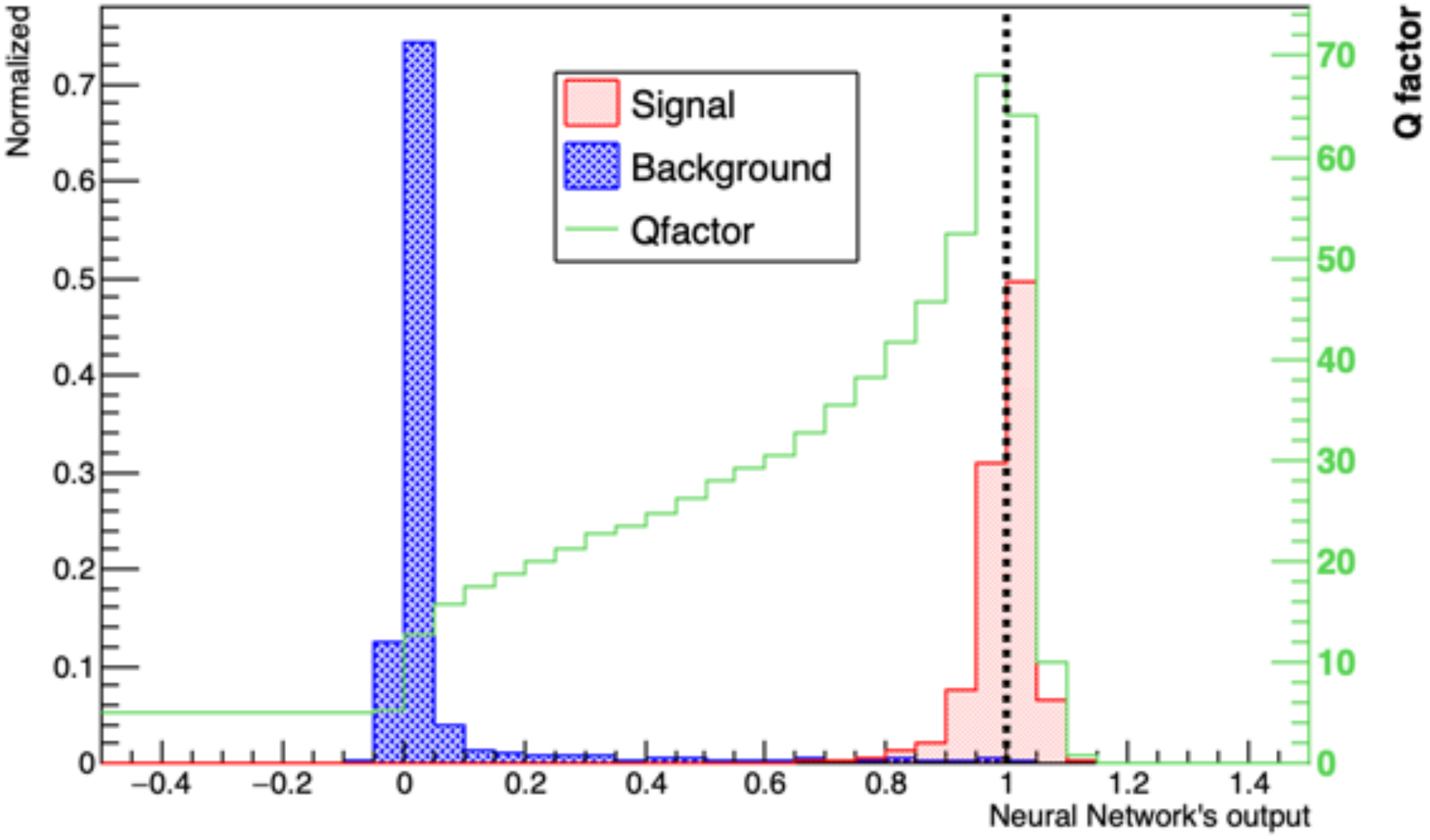


Find the cuts

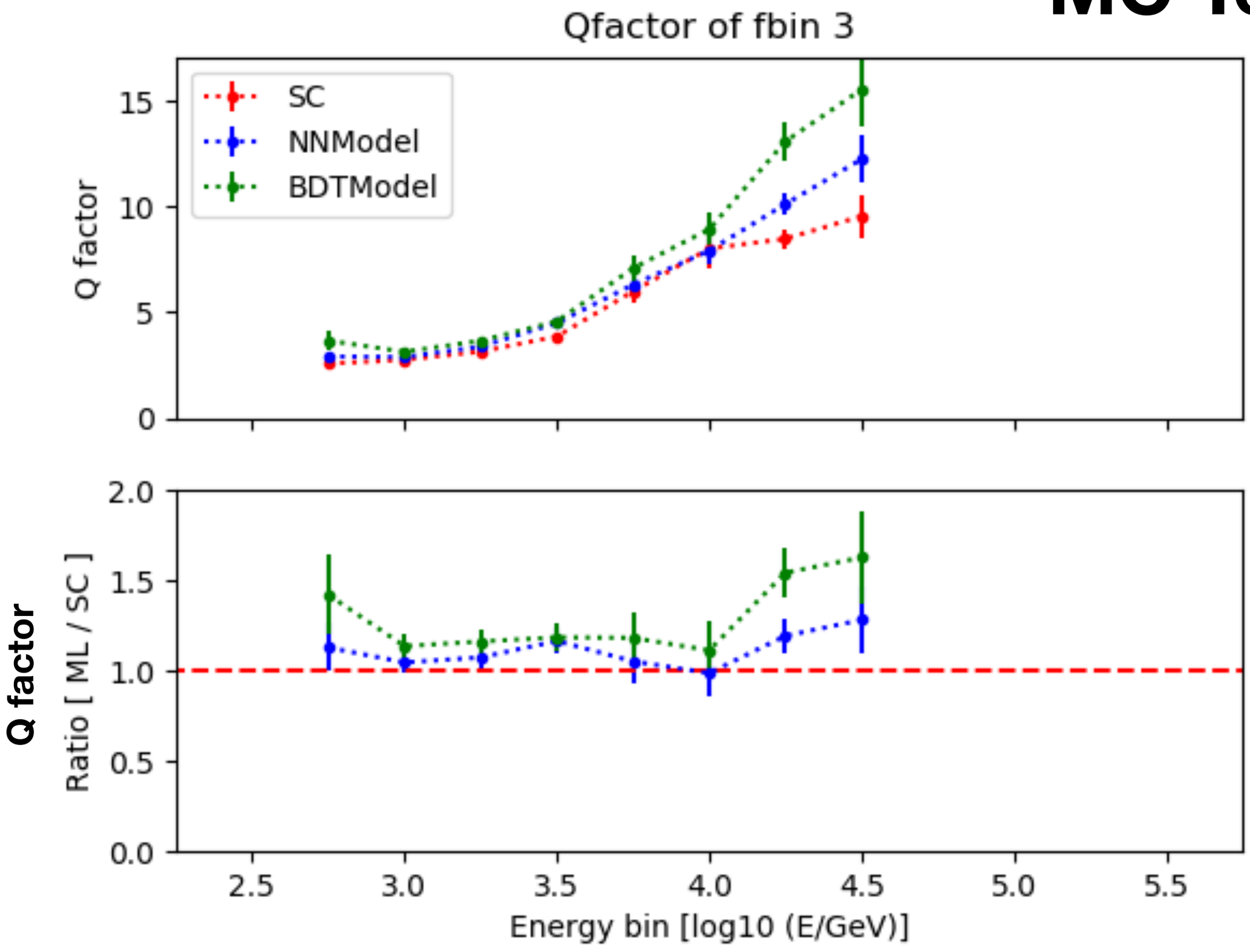
Conditions:

- **Gamma efficiency > 50%**
- **Hadron efficiency > 0.1%**

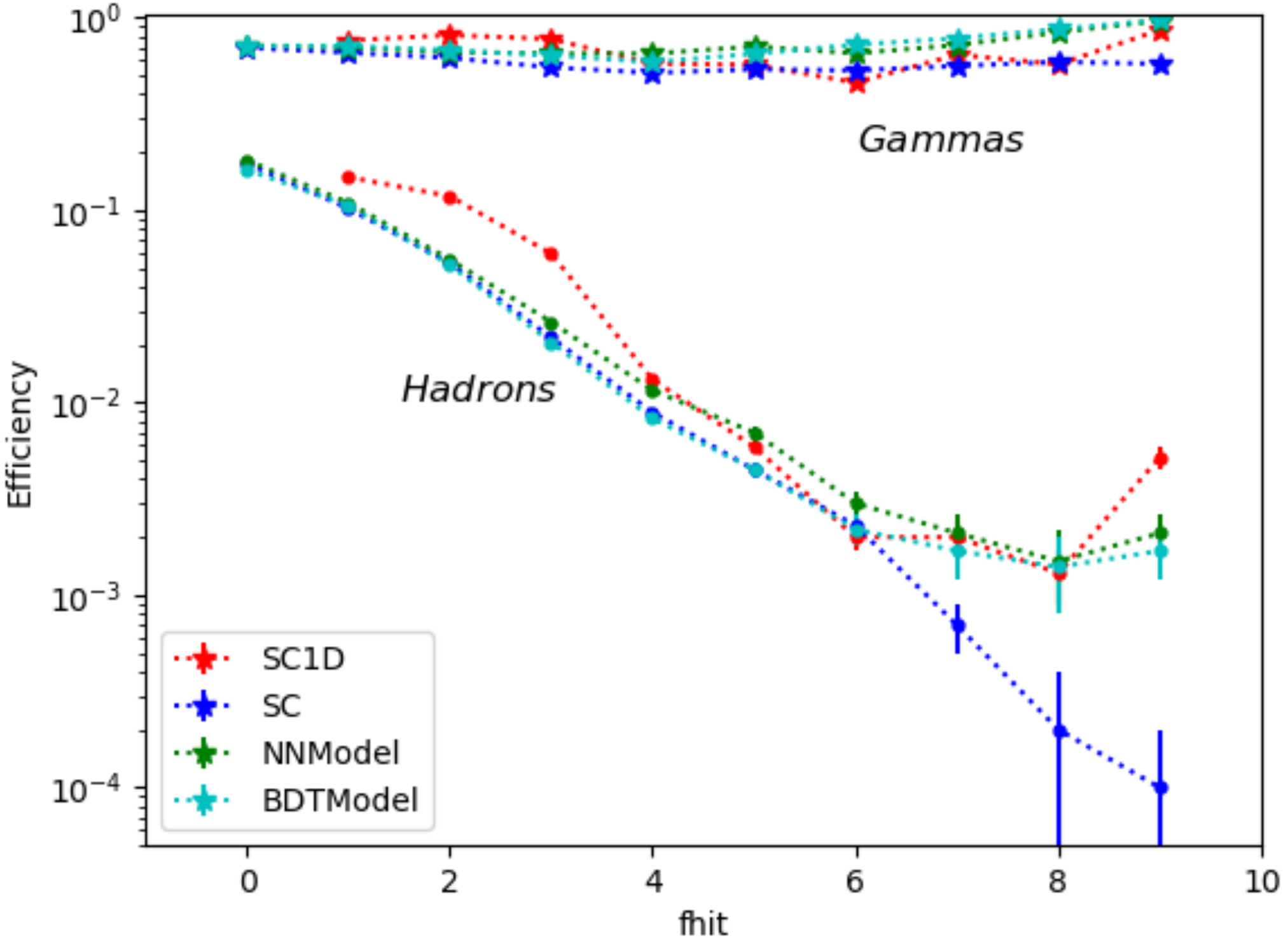
Example of fhit 7, ebin 3.75



MC Test


Qfactor of fbin 3
SC
NNModel
BDTModel
Q factor
0
5
10
15
Q factor
Ratio [ML / SC]
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
Energy bin [log10 (E/GeV)]

MC Test



SC1D - https://iopscience.iop.org/article/10.3847/1538-4357/aa7555

Maps

Data used:

A. **Period :** from 2015/11/06 to 2017/12/20

B. **Duration:** ~837 days

1. Crab Nebula:

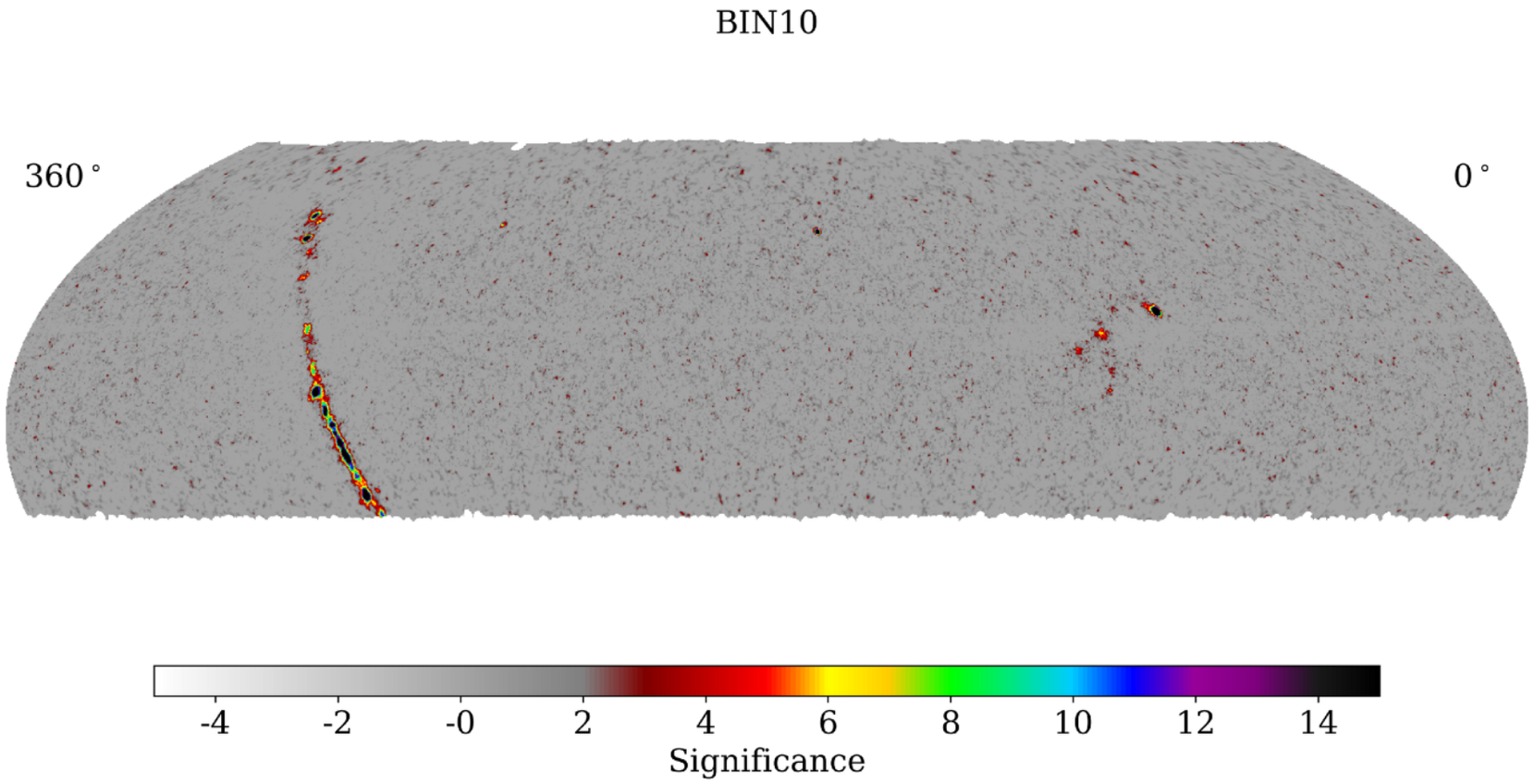
- RA : 83.6332
- DEC: 22.0145

2. Markarian 421:

- RA : 166.1138
- DEC: 38.2088

3. Markarian 501:

- RA : 253.4675
- DEC: 39.7604



4. List of 2nd HAWC Catalog (Use combine maps)

Crab Significance using BDT

Energy Bin lower bound (log (E/GeV))

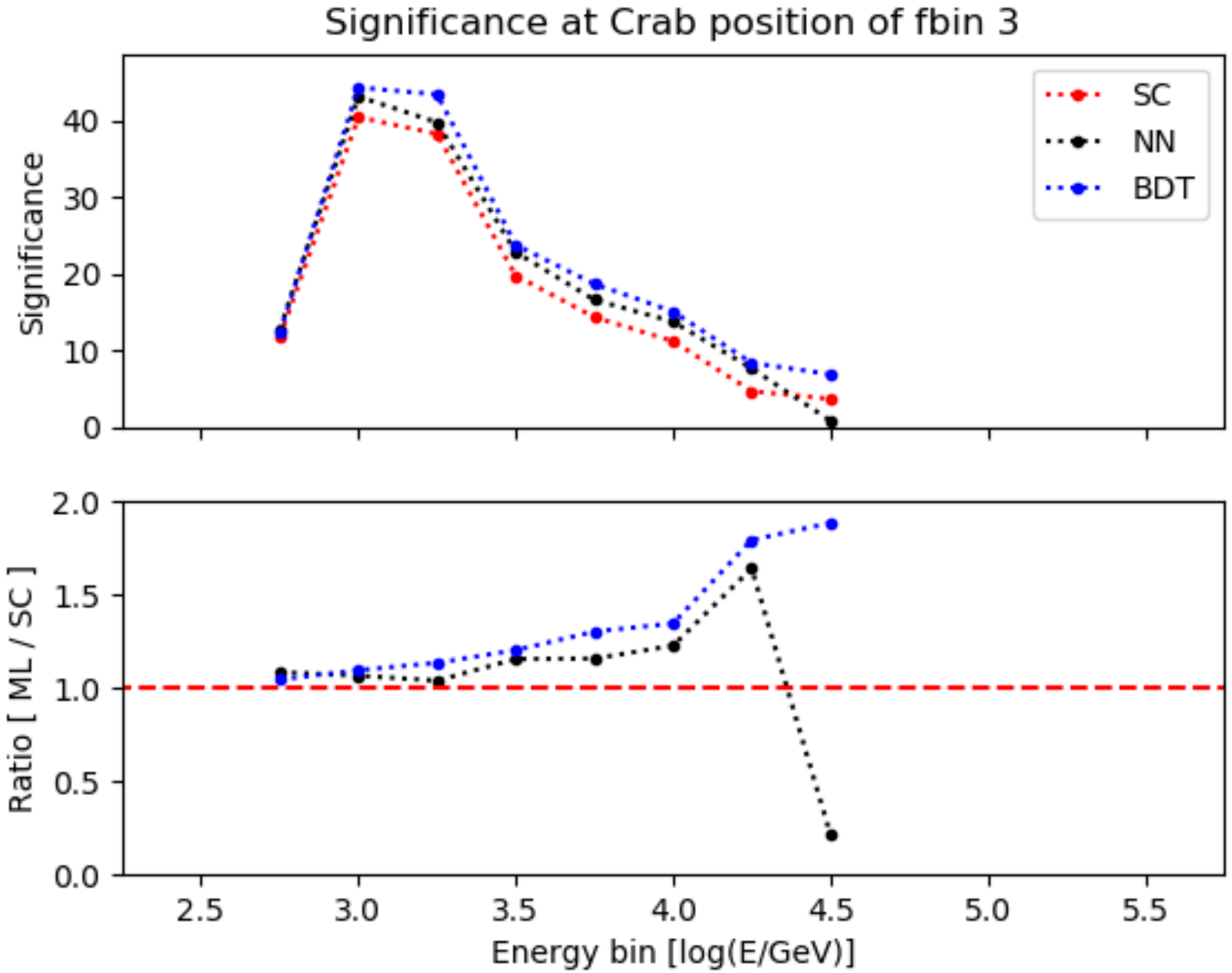
2.50 2.75 3.00 3.25 3.50 3.75 4.00 4.25 4.50 4.75 5.00 5.25

Fractional Bin

0 1 2 3 4 5 6 7 8 9

Significance

0 2 4 6 8 10 12 14


Significance at Crab position of fbin 3
Significance
40
30
20
10
0
SC
NN
BDT
Ratio [ML / SC]
2.0
1.5
1.0
0.5
0.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
Energy bin [log(E/GeV)]

Significance at the source position

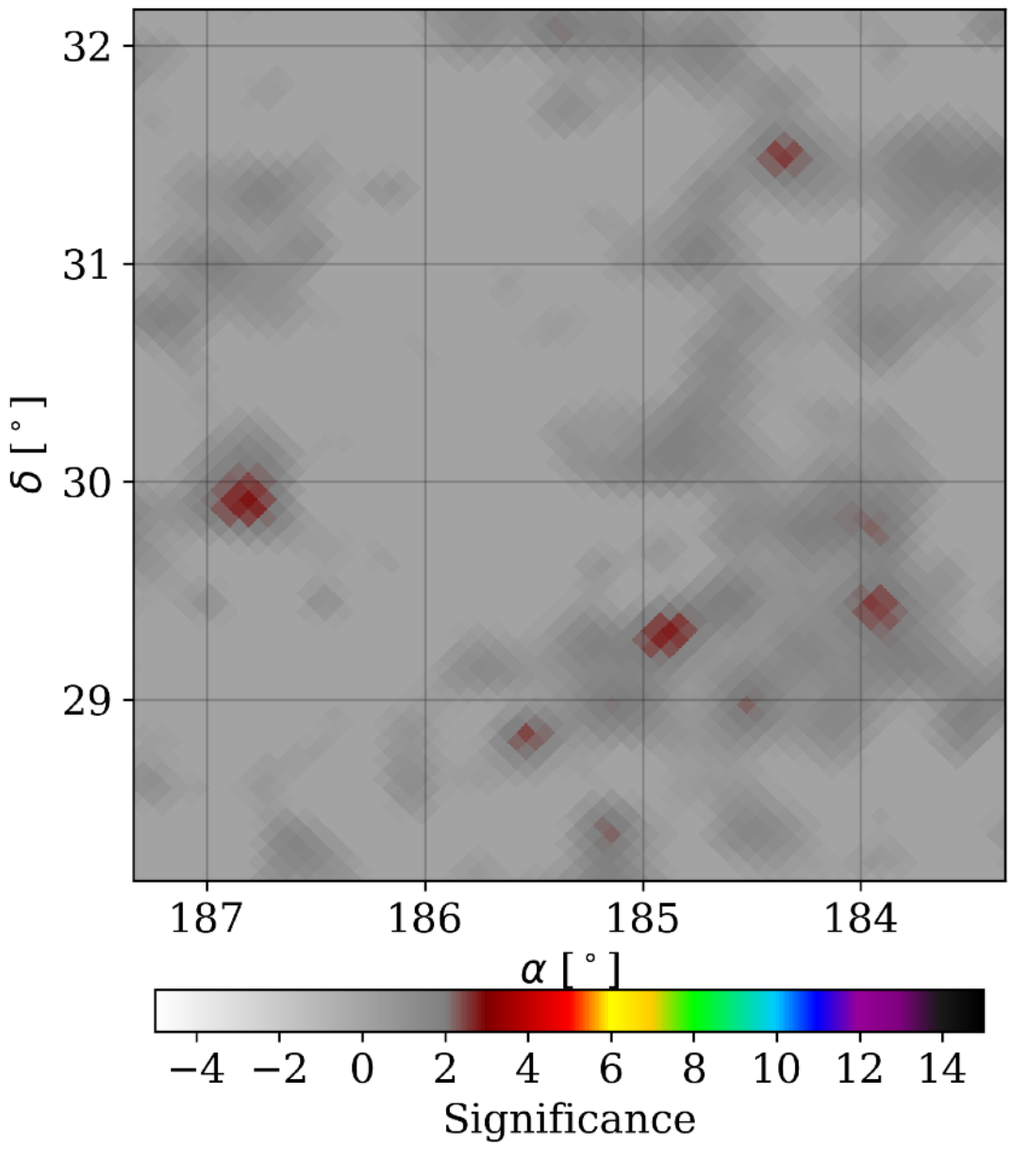
fbin	Crab		Mrk 421		Mkr 501	
	NN / SC	BDT /SC	NN / SC	BDT /SC	NN / SC	BDT /SC
0	-3.1%	5.5%	2.0%	1.7%	-	-
1	-0.4%	2.4%	-6.2%	-2.2%	11.3%	21.1%
2	1.1%	5.1%	-4.2%	2.4%	6.2%	29.4%
3	6.0%	15.4%	3.8%	11.6%	-16.5%	-20.6%
4	9.5%	9.2%	11.6%	5.4%	19.9%	-14.0%
5	-2.2%	12.2%	1.9%	16.7%	14.4%	50.8%
6	-21.5%	7.3%	-3.6%	22.2%	-59.4%	14.1%
7	3.1%	5.5%	22.6%	15.9%	12.1%	29.5%
8	7.2%	6.4%	-10.1%	-51.3%	-13.5%	8.2%
9	9.2%	9.0%	26.1%	-60.9%	117.9%	81.3%
1-9	0.7%	9.6%	1.4%	9.0%	-4.0%	12.4%
0-9	0.7%	9.6%	1.2%	8.6%	-4.9%	11.9%

All source using official fhit

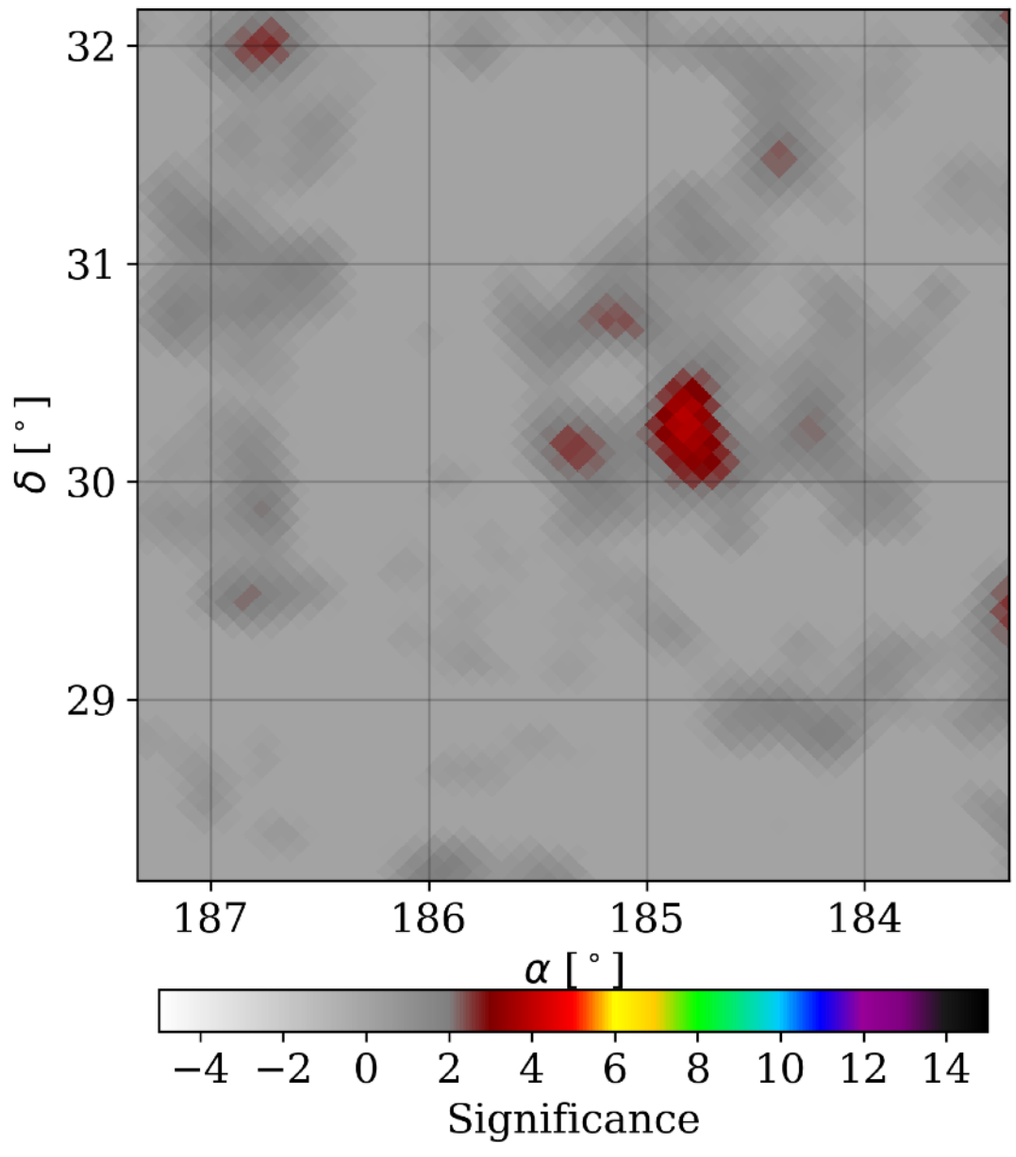
SOURCE	SC	NN	BDT	NN/SC	BDT/SC
J0534+220	155.74	156.87	170.69	0.73%	9.60%
J1104+381	35.26	35.96	38.63	1.99%	9.56%
J1825-134	31.32	34.06	35.76	8.75%	14.18%
⋮					
J0630+186	4.94	3.43	3.60	-30.57%	-27.13%
J2003+348	4.70	4.91	5.17	4.47%	10.00%
J1922+169	4.46	5.09	4.59	14.13%	2.91%
J1918+158	4.12	3.68	3.14	-10.68%	-23.79%
⋮					
1ES_1215+303	2.20	3.89	4.02	76.82%	82.73%
J0709+108	1.96	3.16	2.90	61.22%	47.96%
PG_1218+304	1.95	3.89	4.02	99.49%	106.15%
1ES_2344+514	1.83	1.54	3.76	-15.85%	105.46%

PG 1218+304
using bin 1-9

SC



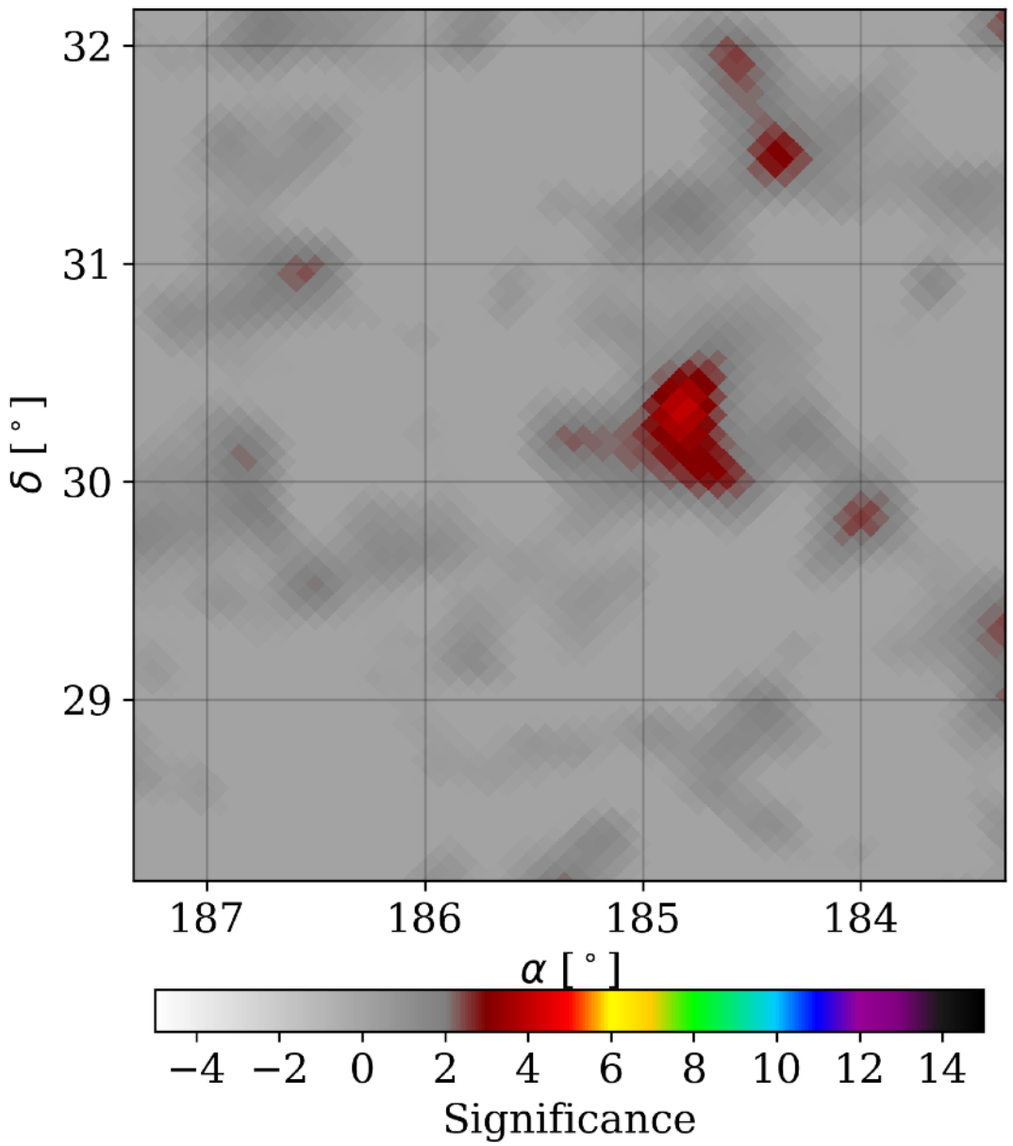
NN



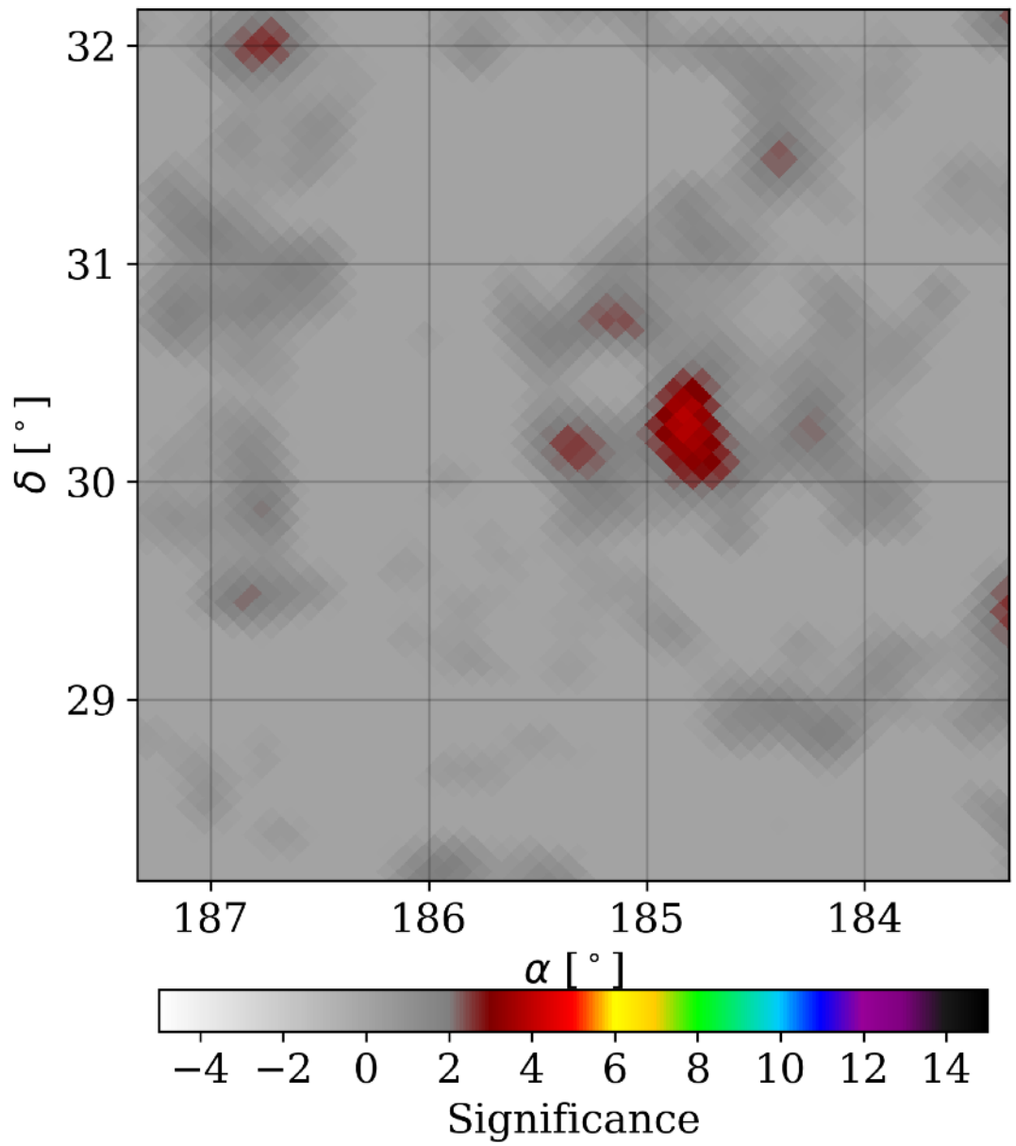
DEC: 30.167 , RA: 185.337

PG 1218+304
using bin 1-9

BDT



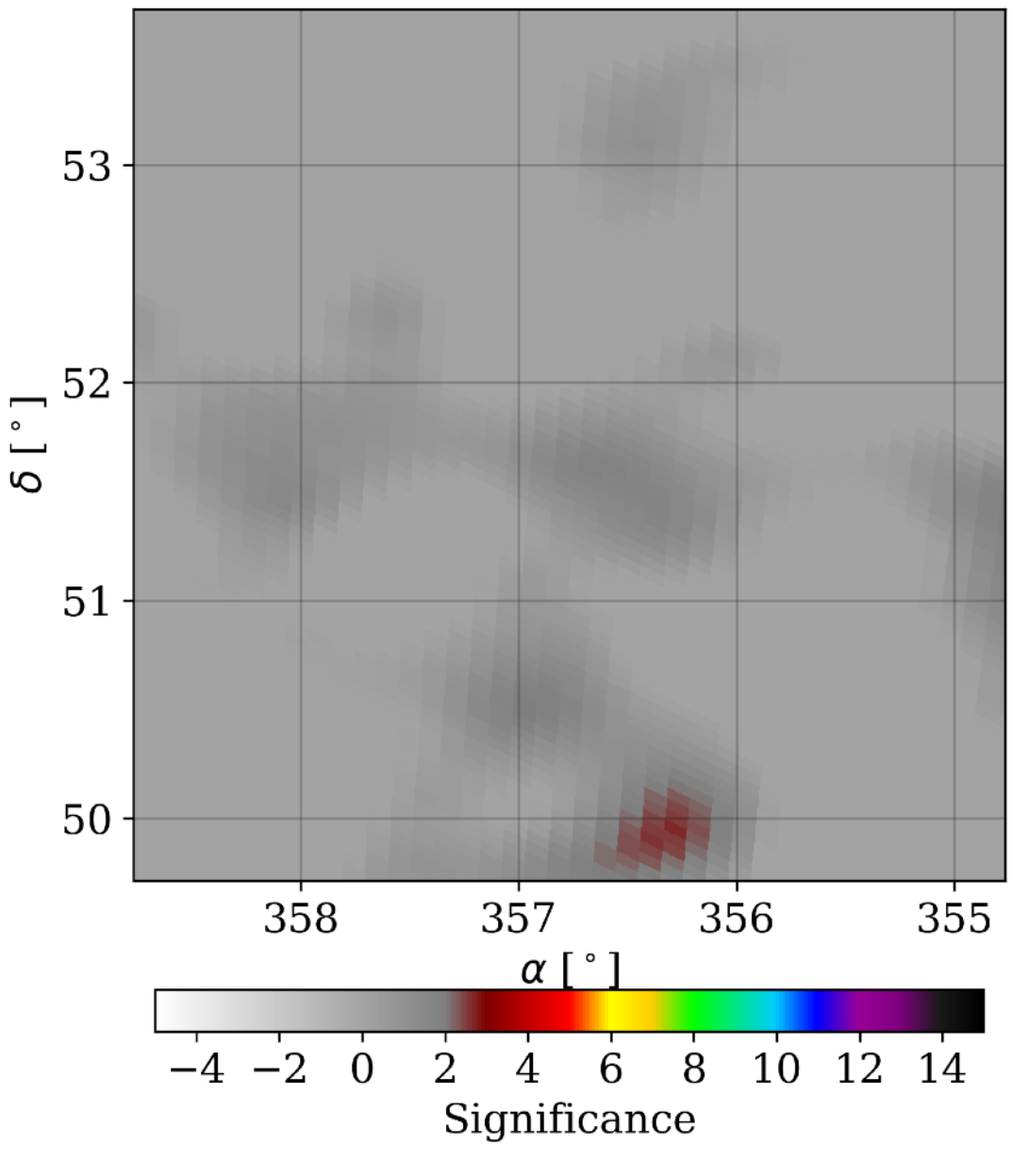
NN



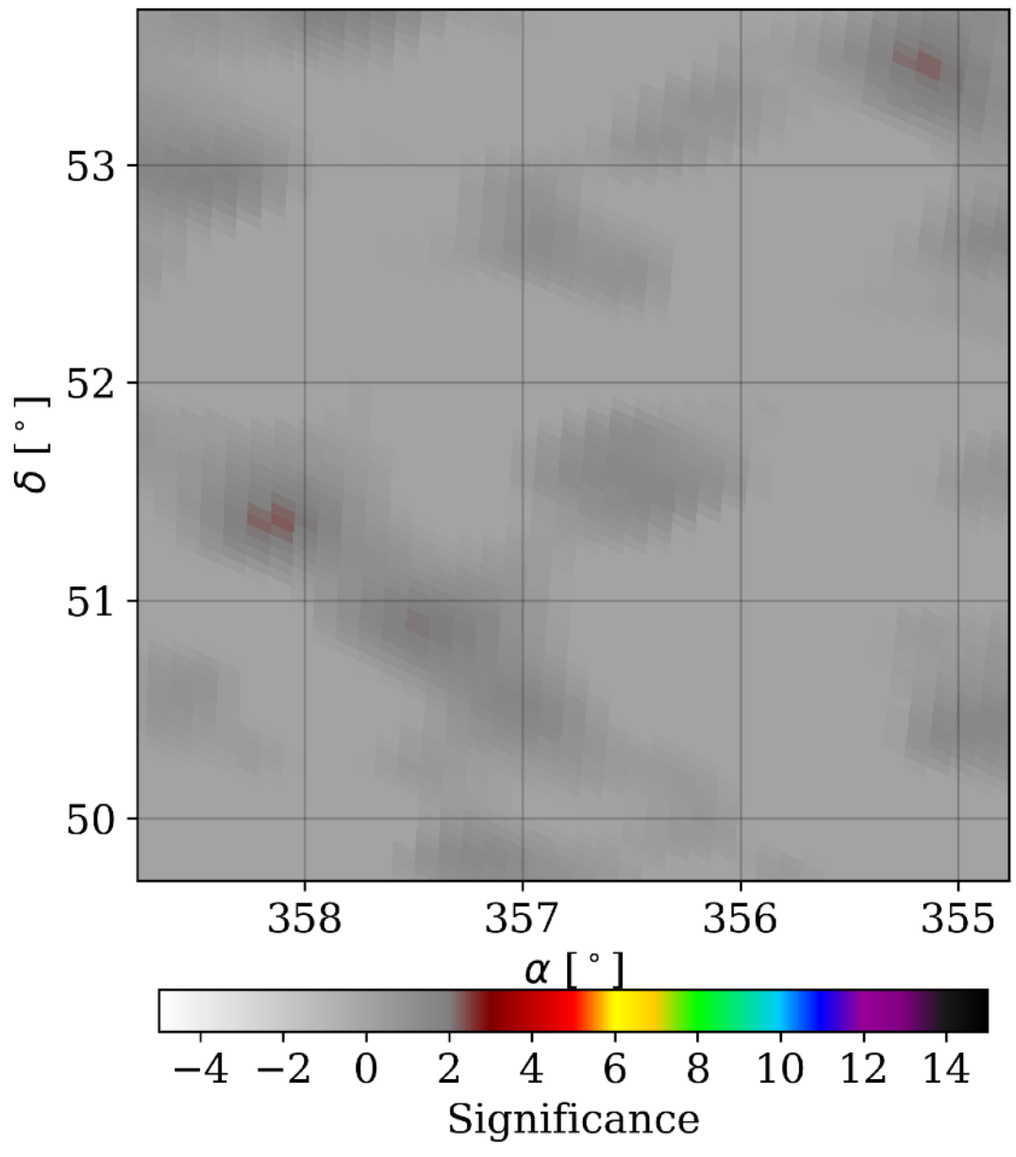
DEC: 30.167 , RA: 185.337

1ES_2344+514
using bin 1-9

SC



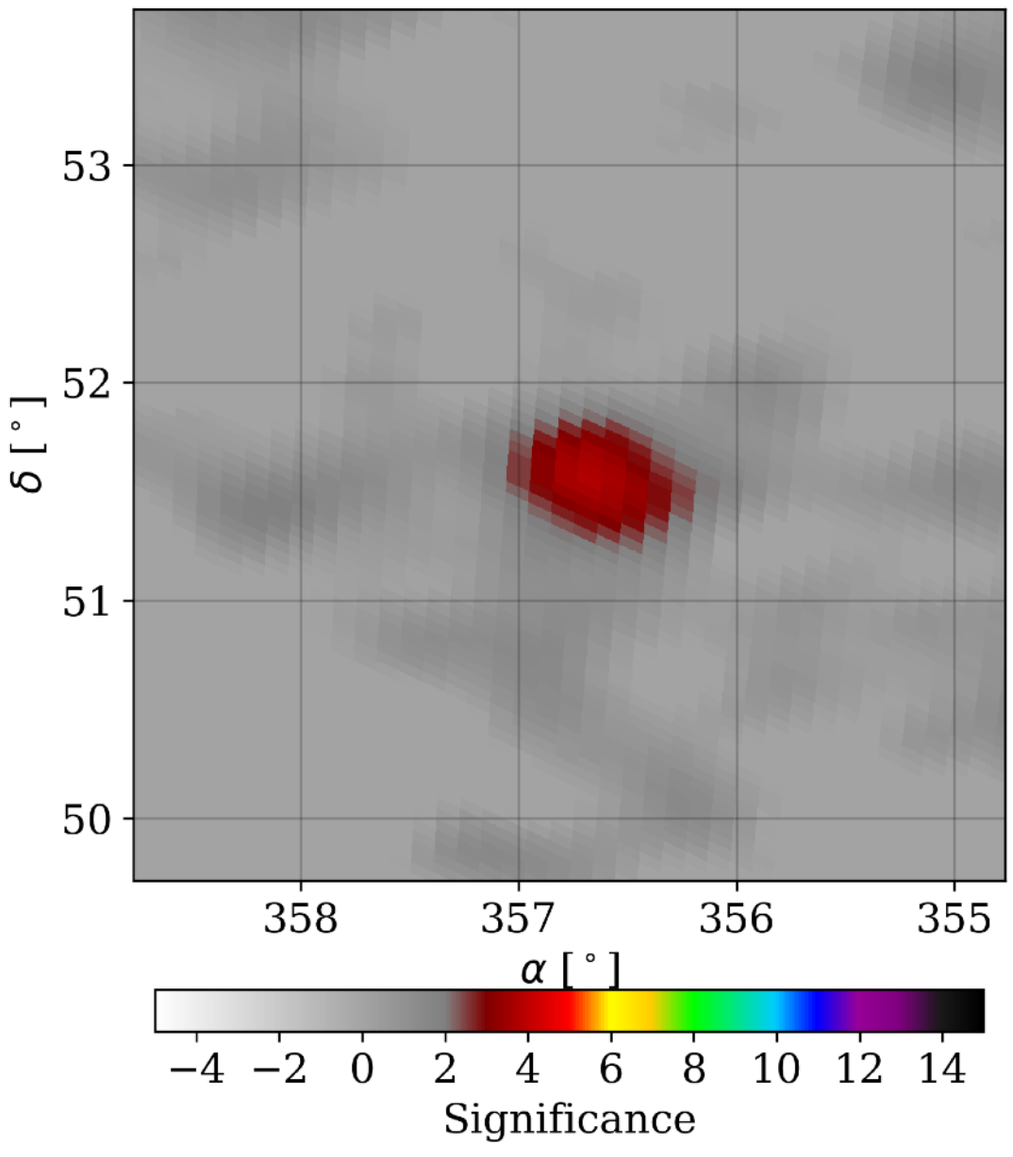
NN



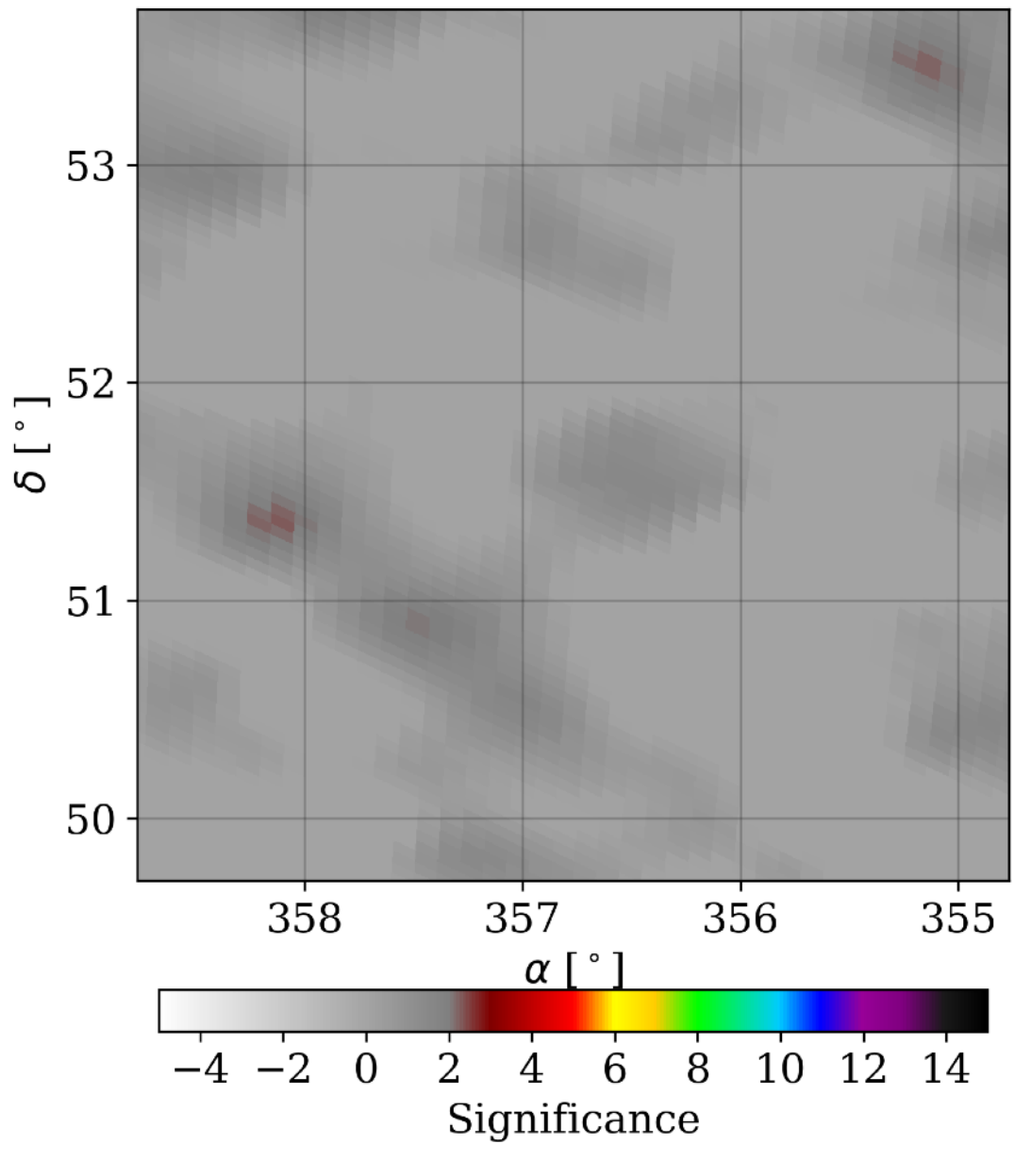
 DEC: 51.7136 , RA: 356.7667

1ES_2344+514
using bin 1-9

BDT



NN



 DEC: 51.7136 , RA: 356.7667

Summary

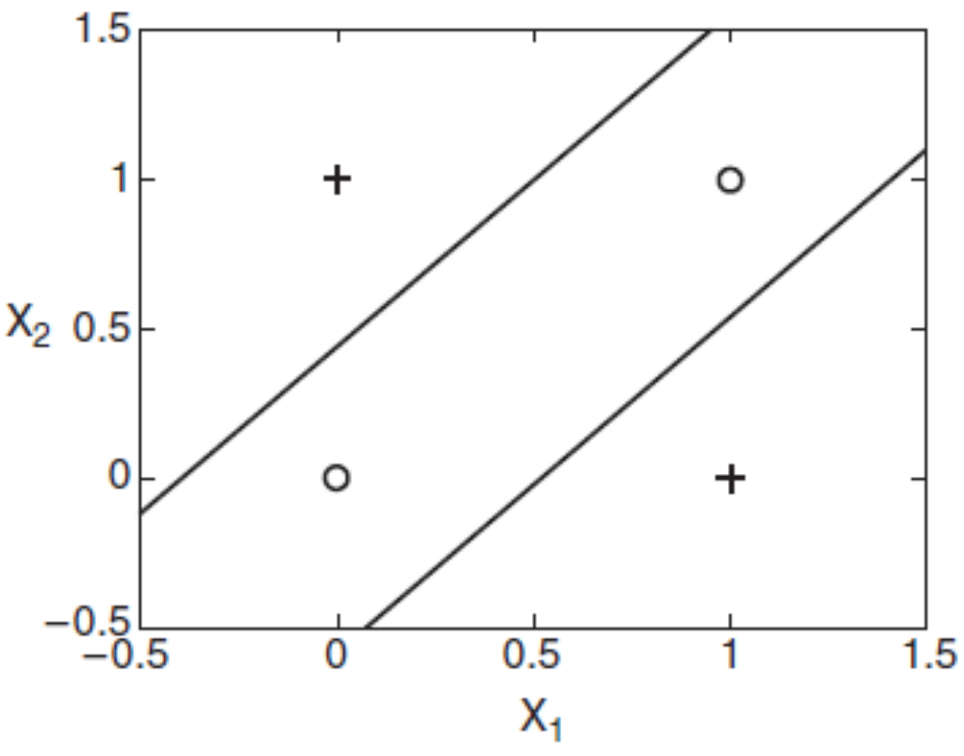
- A g/h separation model were built using the MLT.
- These MLT models where compare with the SC, and get successful results using MC data.
- The MLT has a good results using the Crab Nebula and Mrk 421.

Thanks

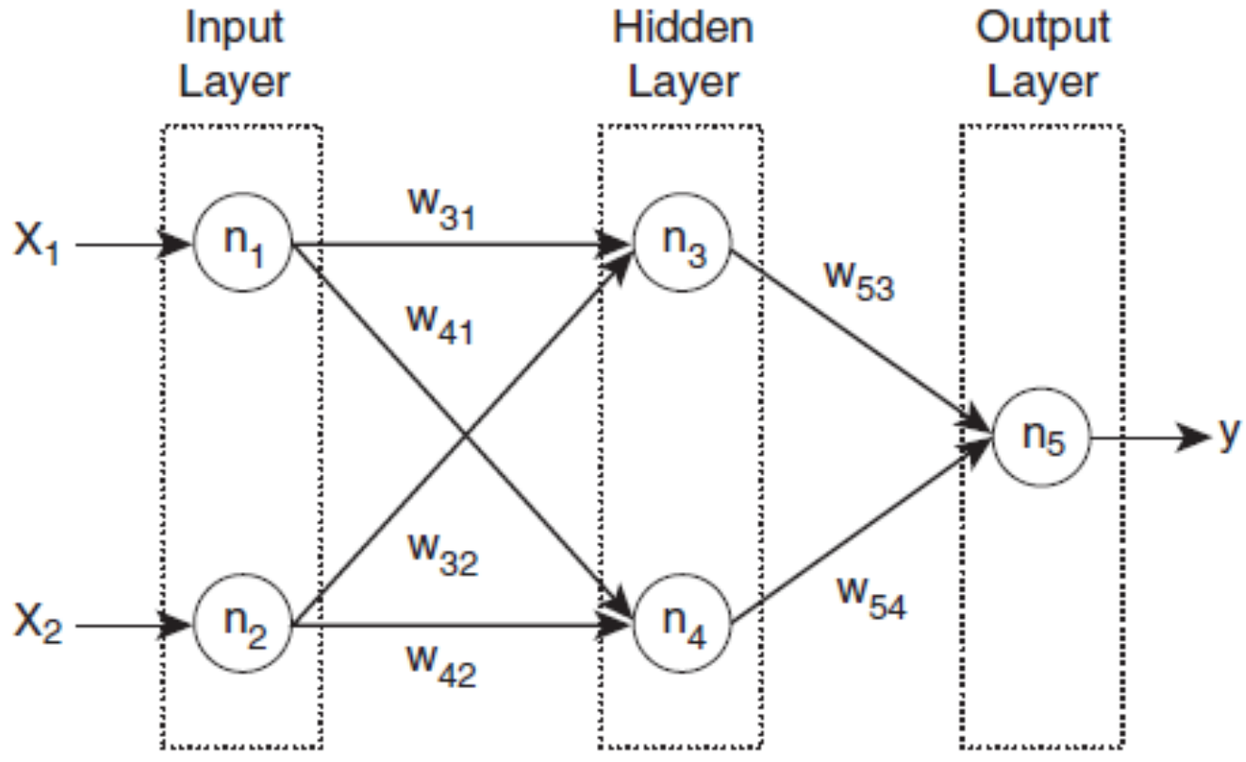
Backslide

Multilayer Neural Network

- It is a nonlinear classifier



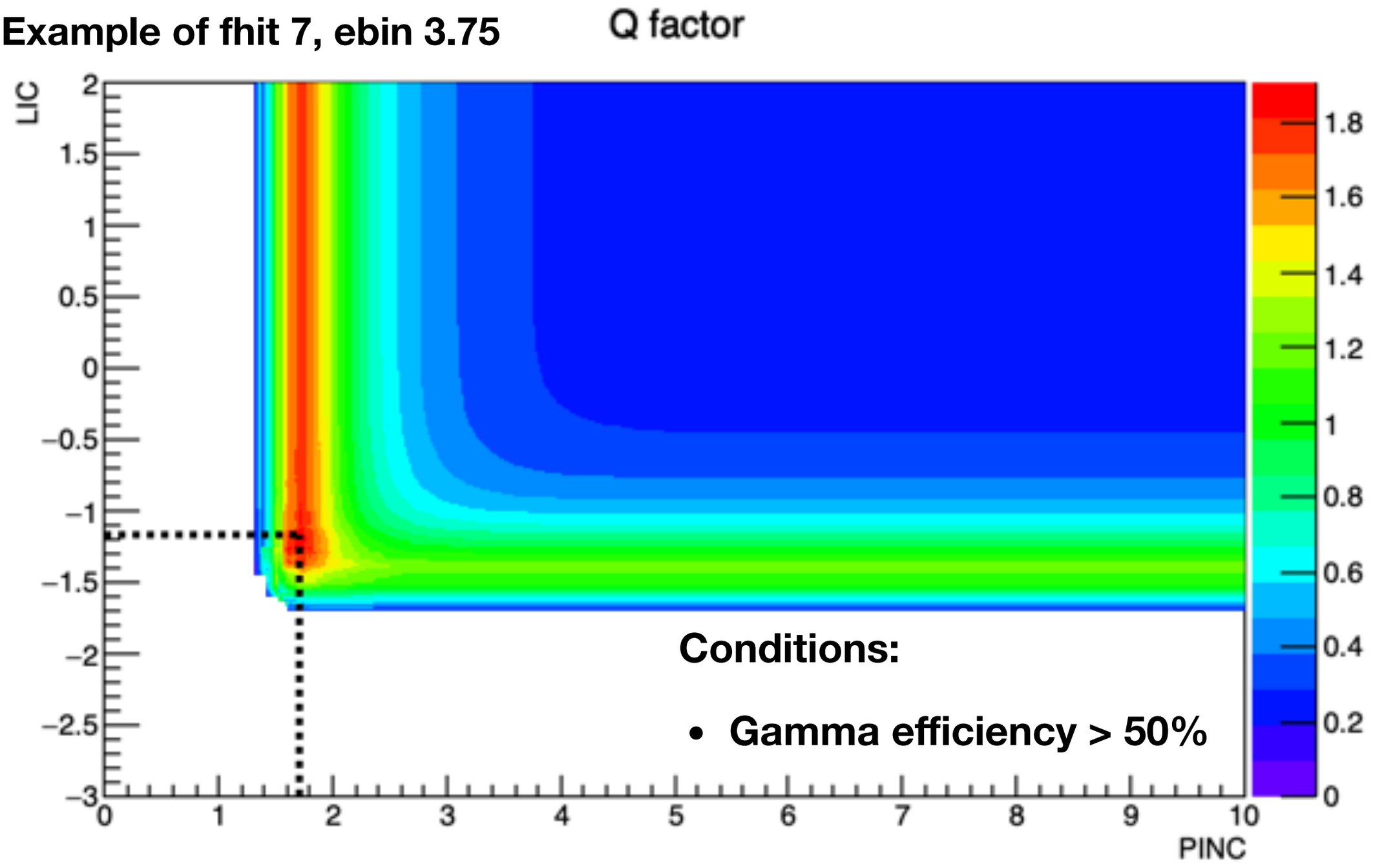
(a) Decision boundary.



(b) Neural network topology.

Standard Cut

Example of fhit 7, ebin 3.75



Crab Significance using SC

Energy Bin lower bound (log (E/GeV))

2.50 2.75 3.00 3.25 3.50 3.75 4.00 4.25 4.50 4.75 5.00 5.25

Fractional Bin

0 1 2 3 4 5 6 7 8 9

Significance

0 2 4 6 8 10 12 14

Crab Significance using NN

Energy Bin lower bound (log (E/GeV))

2.50 2.75 3.00 3.25 3.50 3.75 4.00 4.25 4.50 4.75 5.00 5.25

Fractional Bin

0 1 2 3 4 5 6 7 8 9

Significance

0 2 4 6 8 10 12 14

Crab Significance using BDT

Energy Bin lower bound (log (E/GeV))

2.50 2.75 3.00 3.25 3.50 3.75 4.00 4.25 4.50 4.75 5.00 5.25

Fractional Bin

0 1 2 3 4 5 6 7 8 9

Significance

0 2 4 6 8 10 12 14

Combine ebin to get a fbin map

fbin	Crab			Mrk 421			Mkr 501		
	SC	NN	BDT	SC	NN	BDT	SC	NN	BDT
0	15.16	14.69	15.99	8.10	8.26	8.24	-	-	-
1	27.57	27.47	28.22	13.11	12.30	12.82	3.79	4.22	4.59
2	44.13	44.60	46.36	16.25	15.56	16.64	2.89	3.07	3.74
3	62.39	66.14	71.97	19.10	19.82	21.32	5.34	4.46	4.24
4	69.71	76.34	76.15	19.66	21.95	20.72	5.13	6.15	4.41
5	71.33	69.74	80.05	14.99	15.28	17.49	3.76	4.30	5.67
6	61.52	48.32	65.99	9.13	8.80	11.16	4.95	2.01	5.65
7	47.70	49.18	50.32	5.40	6.62	6.26	2.24	2.51	2.90
8	32.75	35.10	34.84	1.19	1.07	0.58	2.67	2.31	2.89
9	28.70	31.34	31.29	0.23	0.29	0.09	1.12	2.44	2.03
1-9	155.74	156.87	170.69	35.26	35.74	38.43	10.62	10.20	11.94
0-9	156.33	157.45	171.31	35.99	36.42	39.10	10.63	10.11	11.90