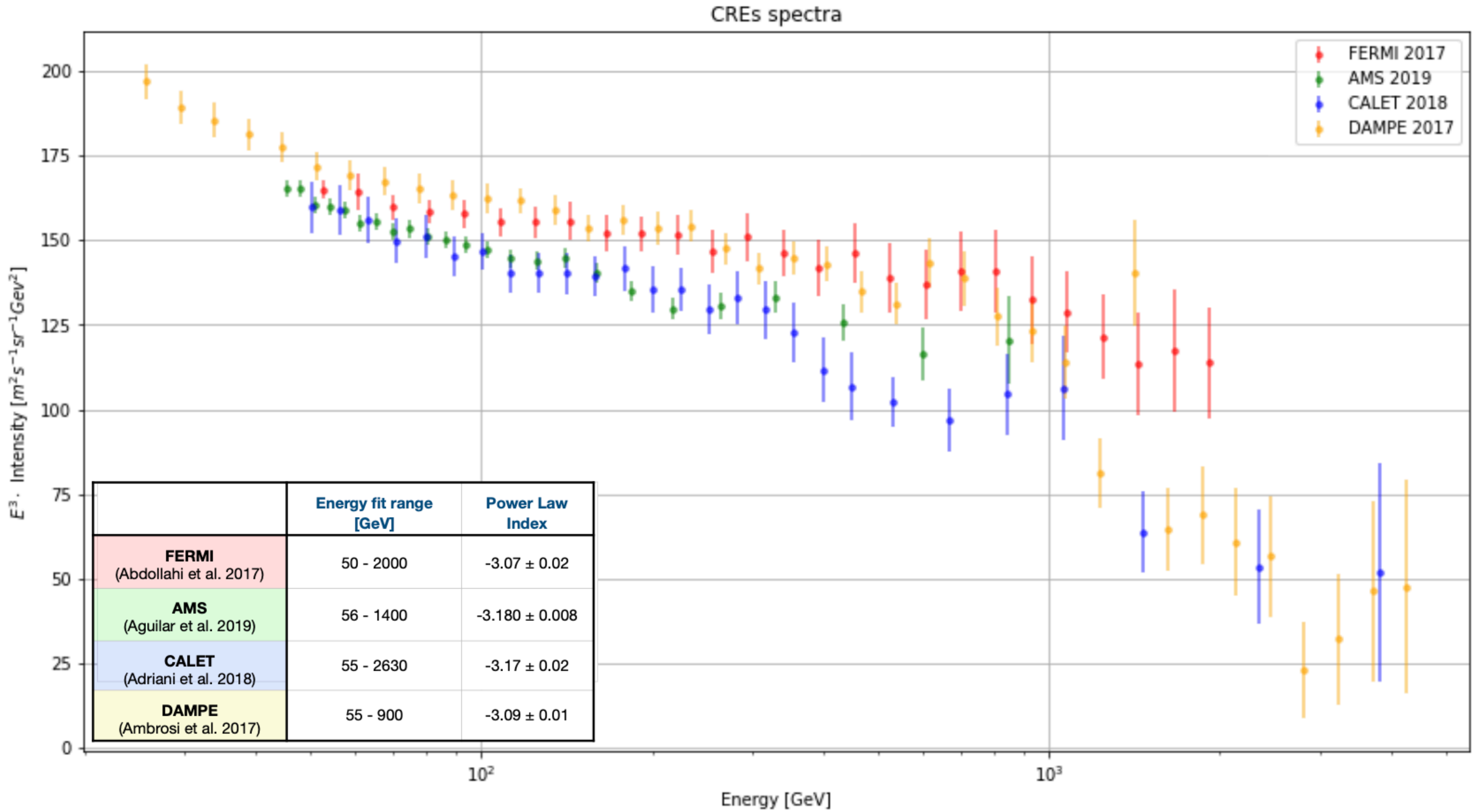
Machine Learning for the measurement of the Cosmic-Ray Inclusive Electron Spectrum with Fermi LAT

Nicolò Cibrario
On behalf of Eleonora Barbano and Raffaella Bonino

University of Torino, Italy
nicolo.cibrario@unito.it

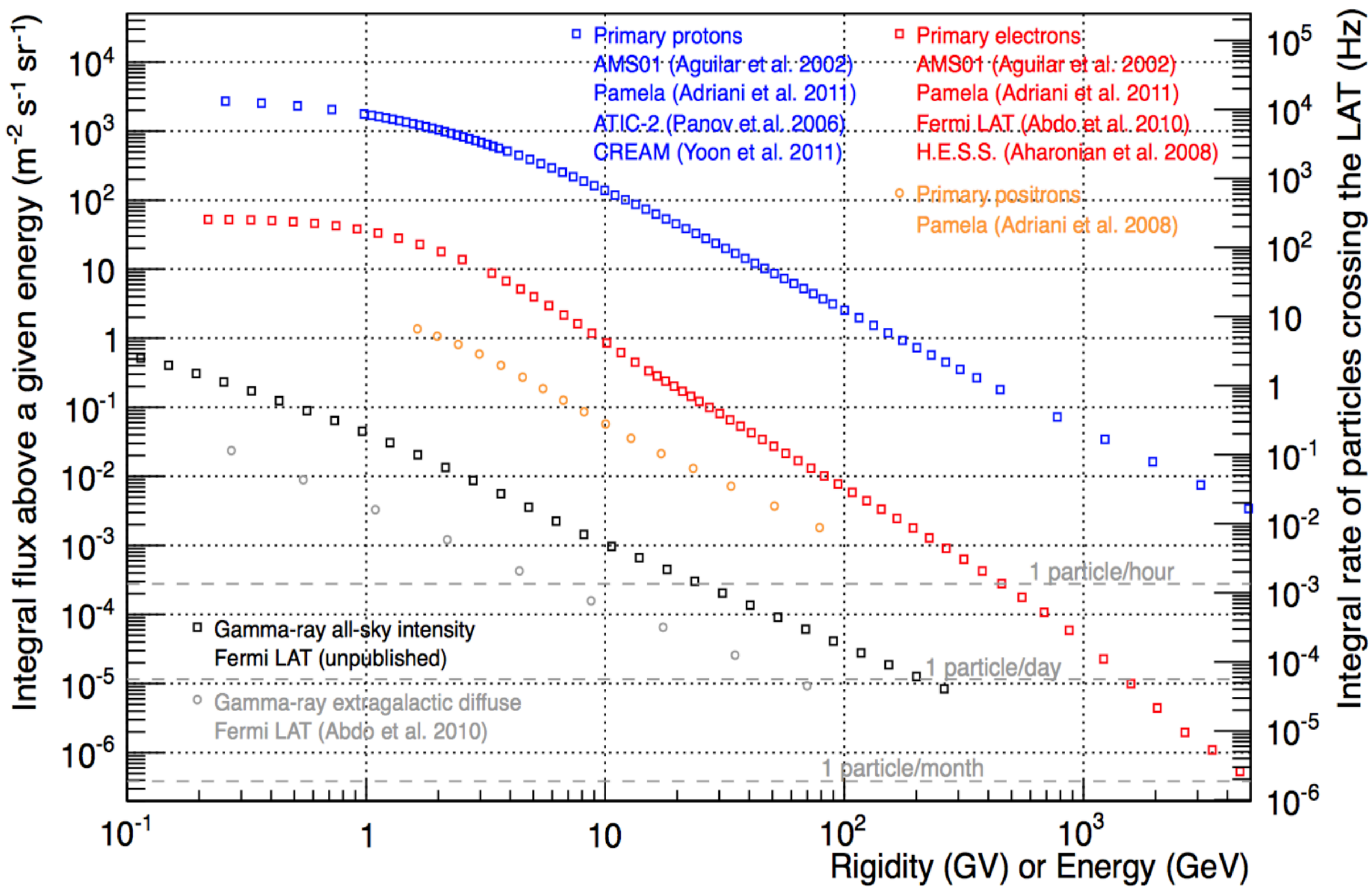
State of the art



	Energy fit range [GeV]	Power Law Index
FERMI (Abdollahi et al. 2017)	50 - 2000	-3.07 ± 0.02
AMS (Aguilar et al. 2019)	56 - 1400	-3.180 ± 0.008
CALET (Adriani et al. 2018)	55 - 2630	-3.17 ± 0.02
DAMPE (Ambrosi et al. 2017)	55 - 900	-3.09 ± 0.01

Analysis outline

Objective: Discriminate between electrons and background (mainly protons)



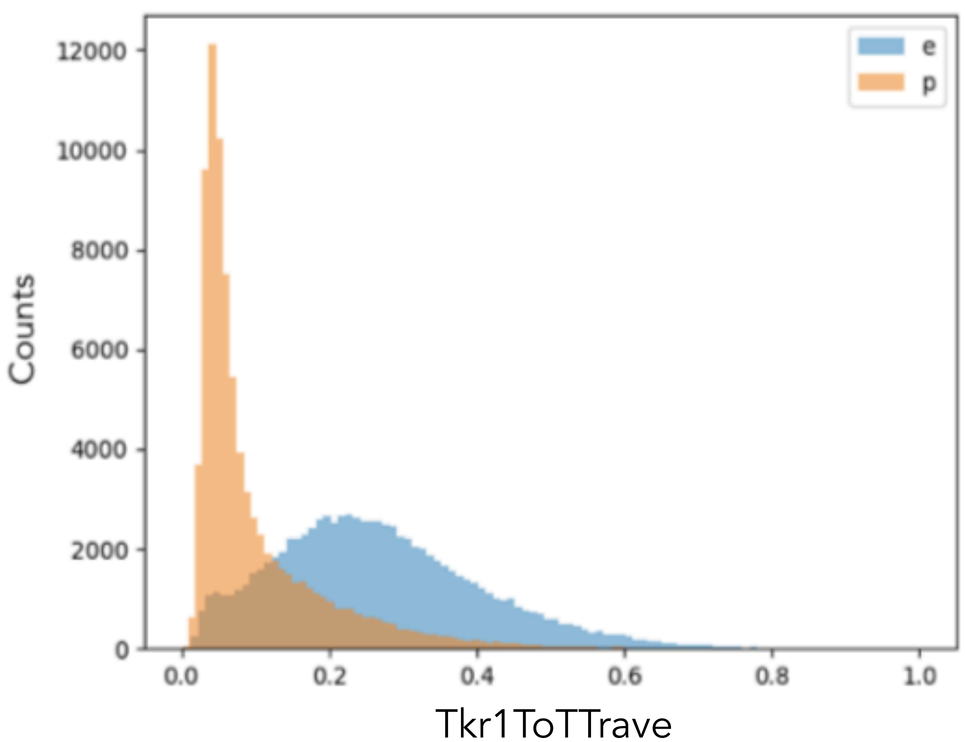
2017 Analysis

Selection of variables → Boosted Decision Trees (BDT) → Template Fitting → CRE Spectrum

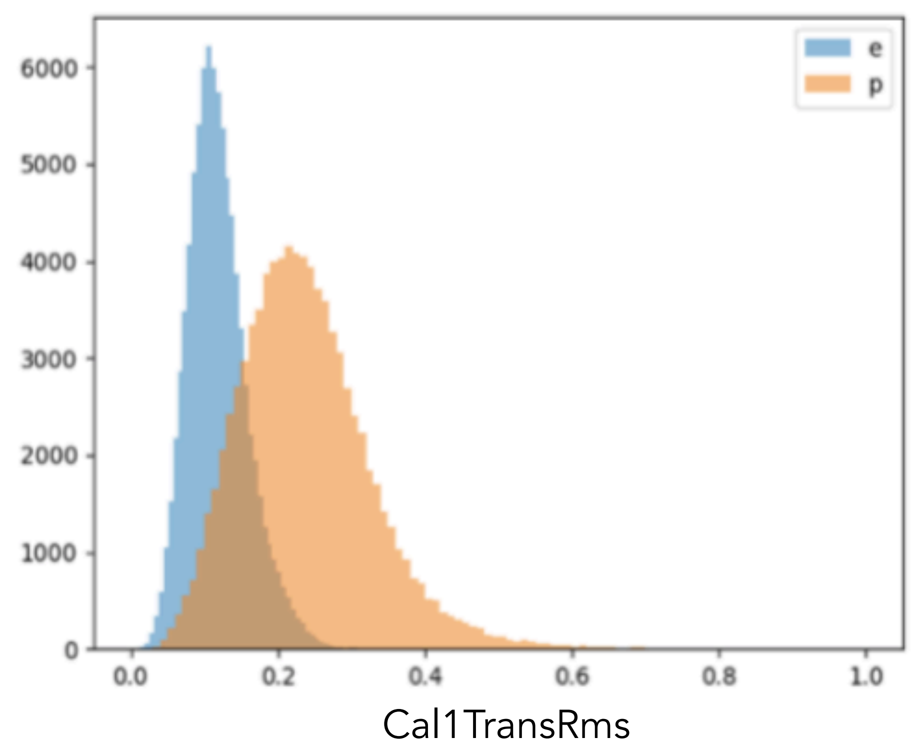
A subset of variables is needed:

highlight the differences between electrons and protons events

Average Time over Threshold for the hits



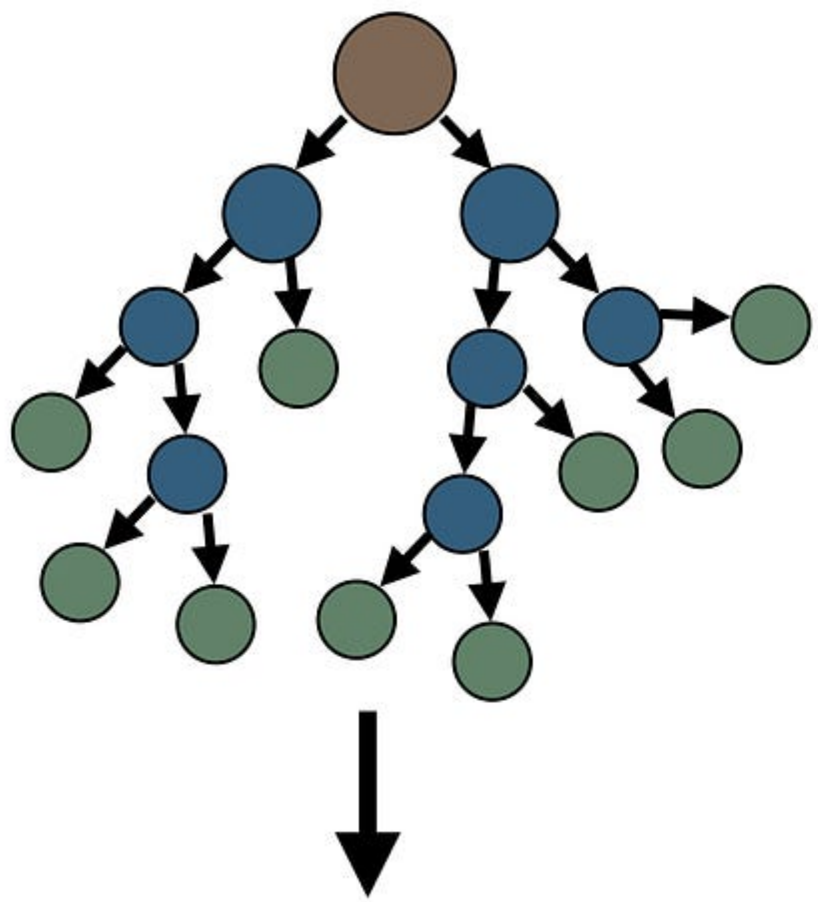
RMS of transverse position measurements in the cluster



2017 Analysis

Selection of variables → Boosted Decision Trees (BDT) → Template Fitting → CRE Spectrum

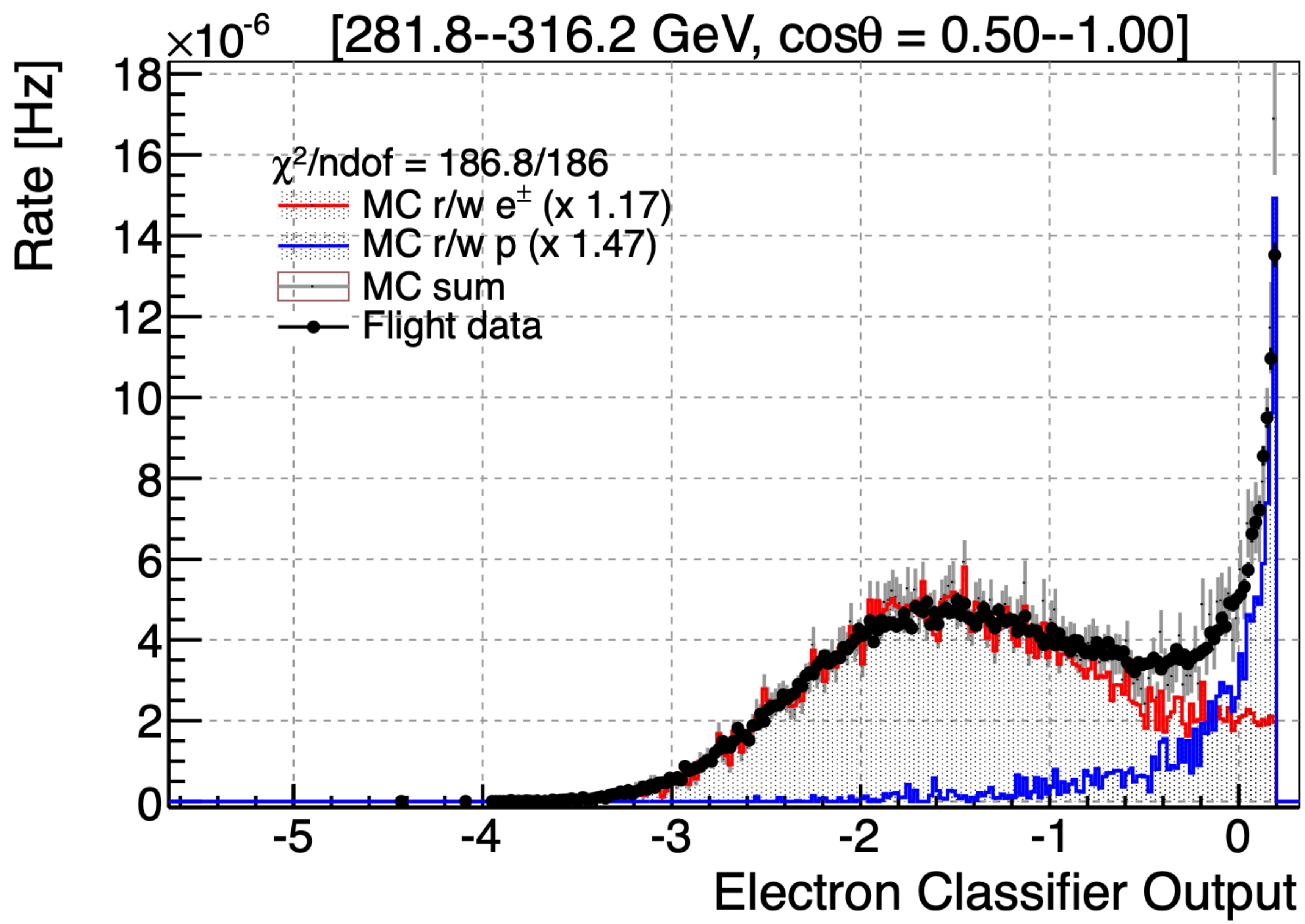
They are trained to assign to particles a variable **p: ranging from -1 (indicating protons-like events) to 1 (indicating electrons-like events)**. Log(1-p) is calculated, to highlight the region where electrons and protons overlap.



$p \in [-1,1]$

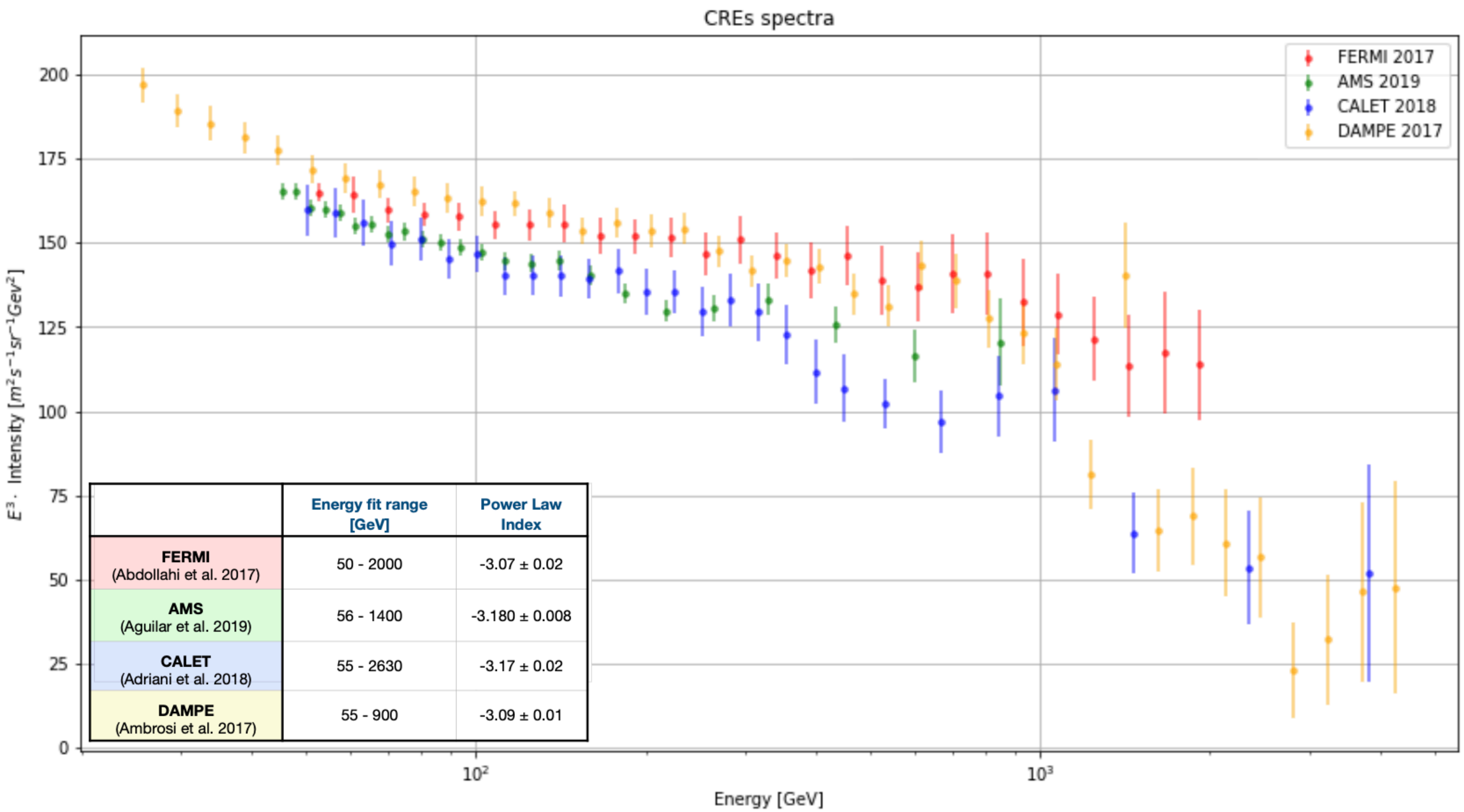
2017 Analysis

Selection of variables → Boosted Decision Trees (BDT) → **Template Fitting** → CRE Spectrum



2017 Analysis

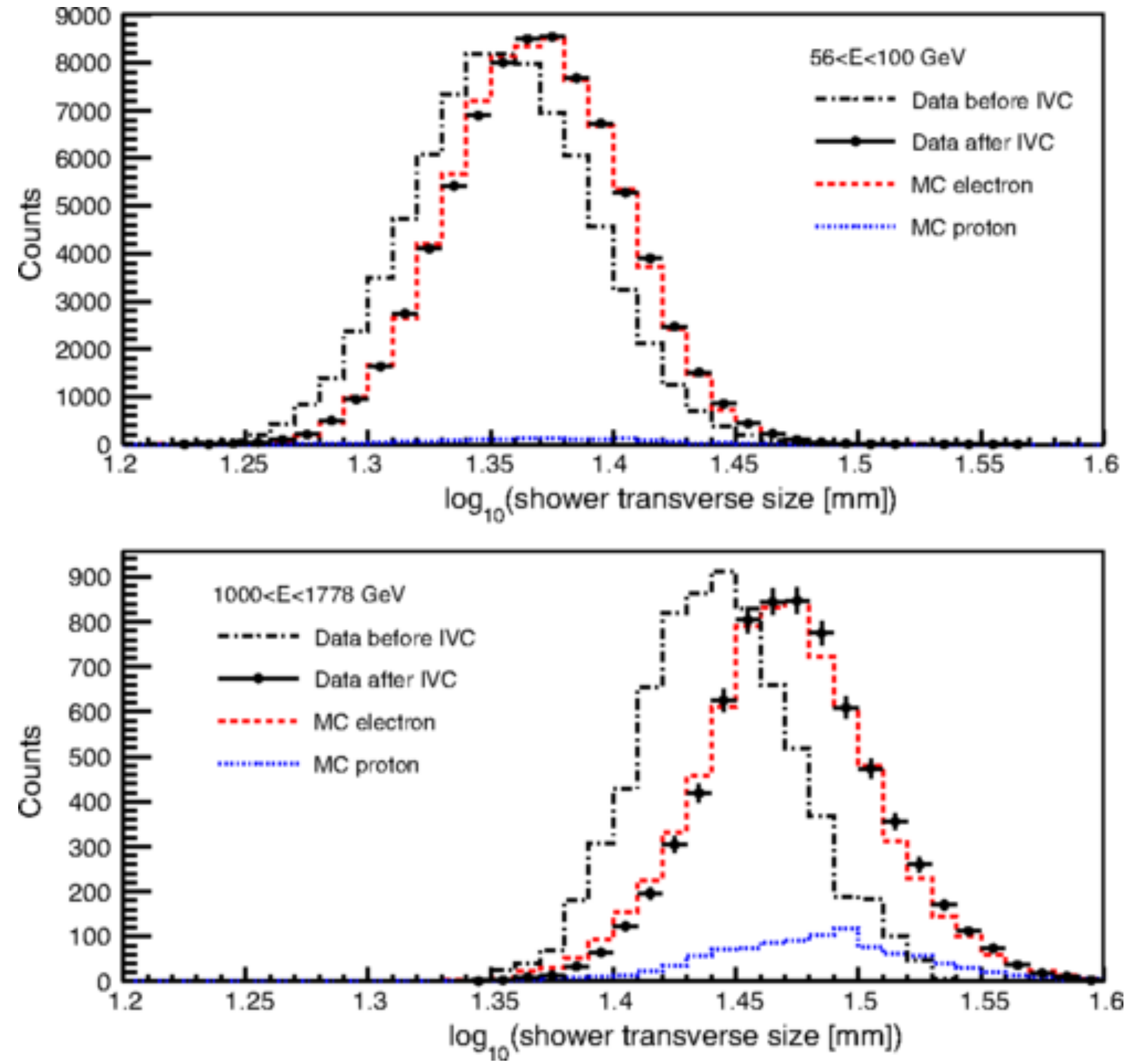
Selection of variables → Boosted Decision Trees (BDT) → Template Fitting → CRE Spectrum



	Energy fit range [GeV]	Power Law Index
FERMI (Abdollahi et al. 2017)	50 - 2000	-3.07 ± 0.02
AMS (Aguilar et al. 2019)	56 - 1400	-3.180 ± 0.008
CALET (Adriani et al. 2018)	55 - 2630	-3.17 ± 0.02
DAMPE (Ambrosi et al. 2017)	55 - 900	-3.09 ± 0.01

IVC Corrections

Need to correct the MC-data agreement with the IVC corrections: introduction of a systematic which is difficult to quantify



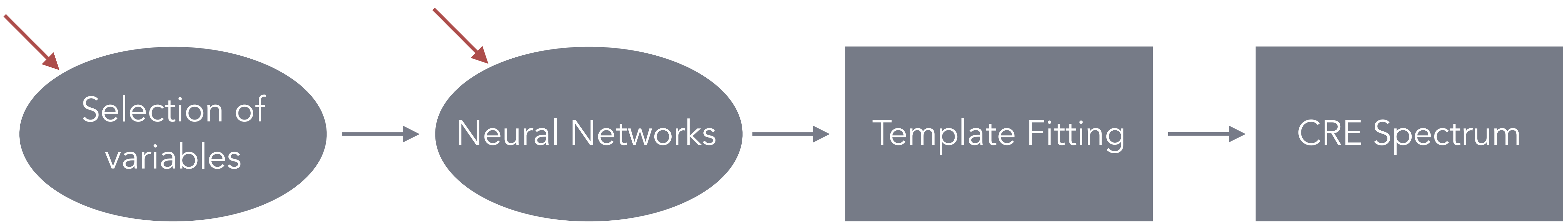
Neural Networks

Possible improvements:

Detect and handle **non-linear relations** among variables

Overcome overfitting by some regularizing steps

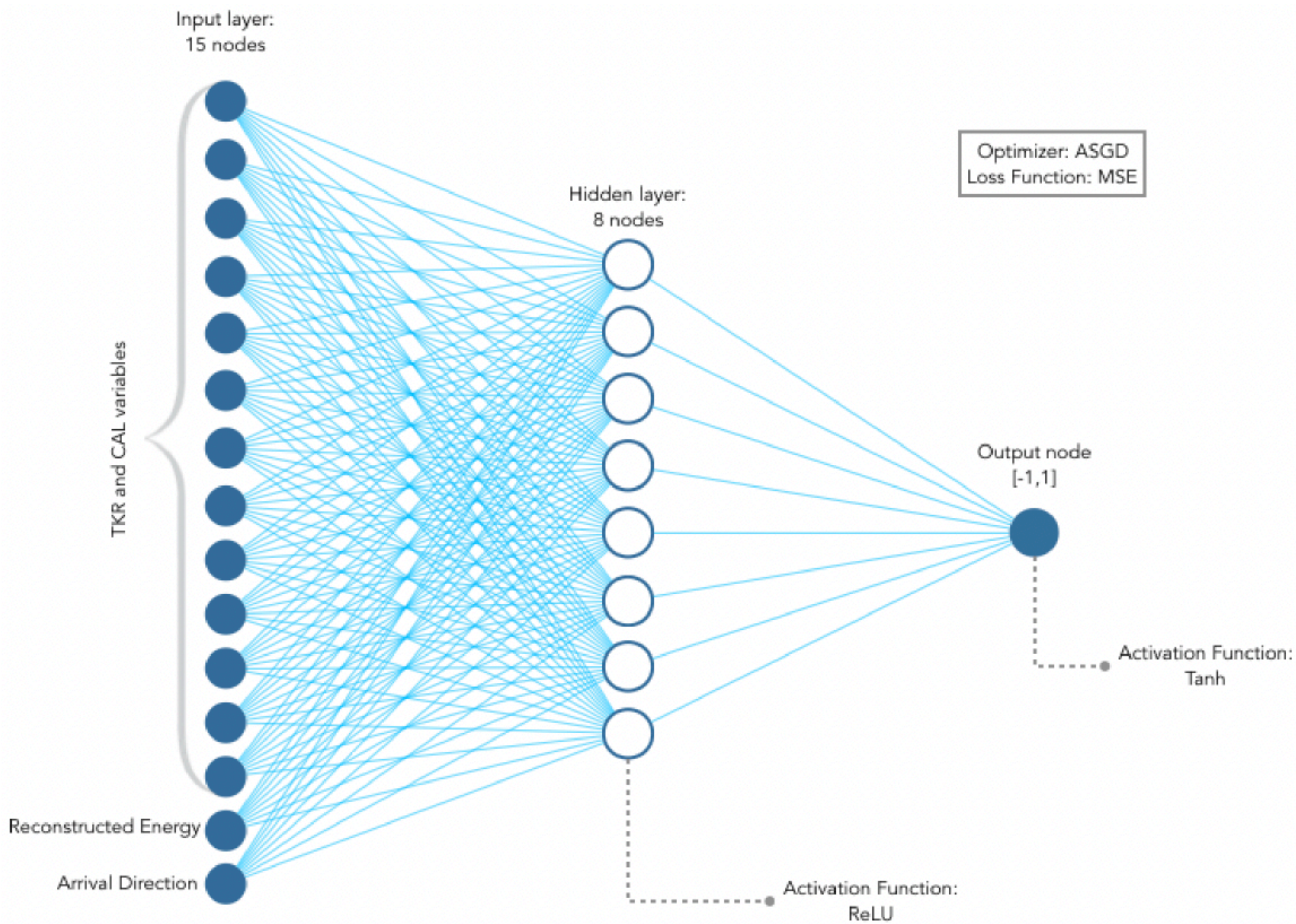
Less dependency on MC-data agreement, as no cut is applied to variables



Neural Networks

Selection of variables → **Neural Networks** → Template Fitting → CRE Spectrum

Ensamble of shallow Neural Networks used for assigning the variable p



Unsupervised Learning

Possible improvement:
Full independence from MC, uncertainties could be reduced

Complication:
Difficulties in dealing with very different sizes of populations: proton background is dominant wrt electron signal

Selection of variables → Dimensionality Reduction → Clustering

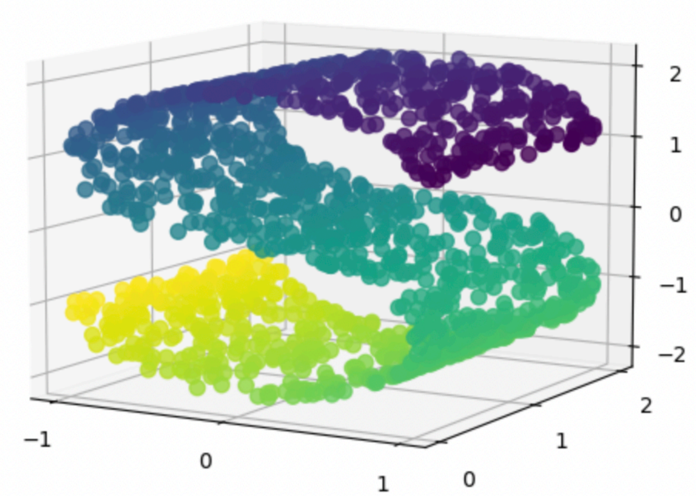
Unsupervised Learning

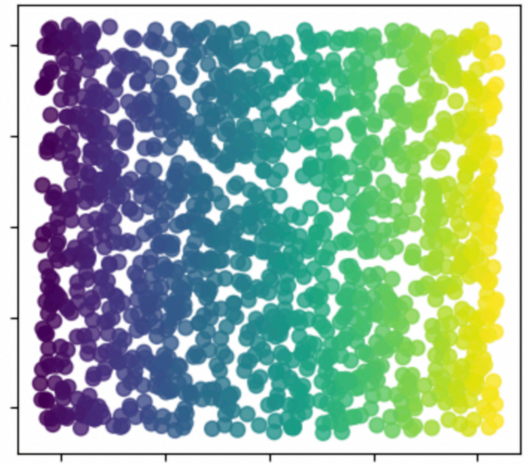
Selection of variables → **Dimensionality Reduction** → Clustering

Selected algorithm:
Locally Linear Embedding (LLE)

Why dimensionality reduction?

- Clustering algorithms usually perform better with low dimensions
- Possibility of introducing an evaluation of uncertainty

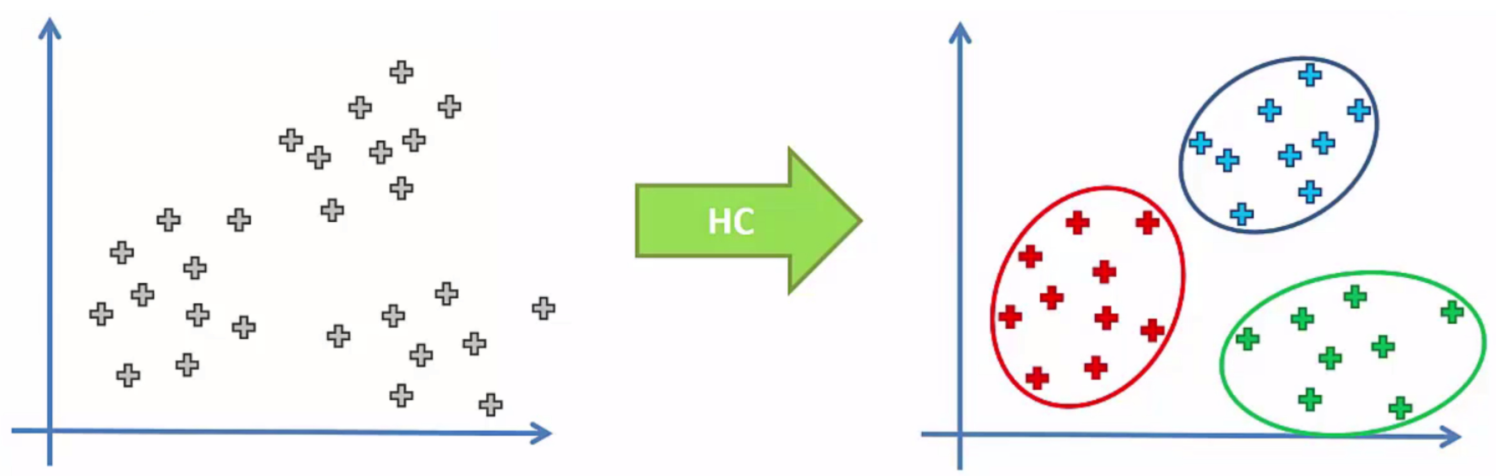




Unsupervised Learning

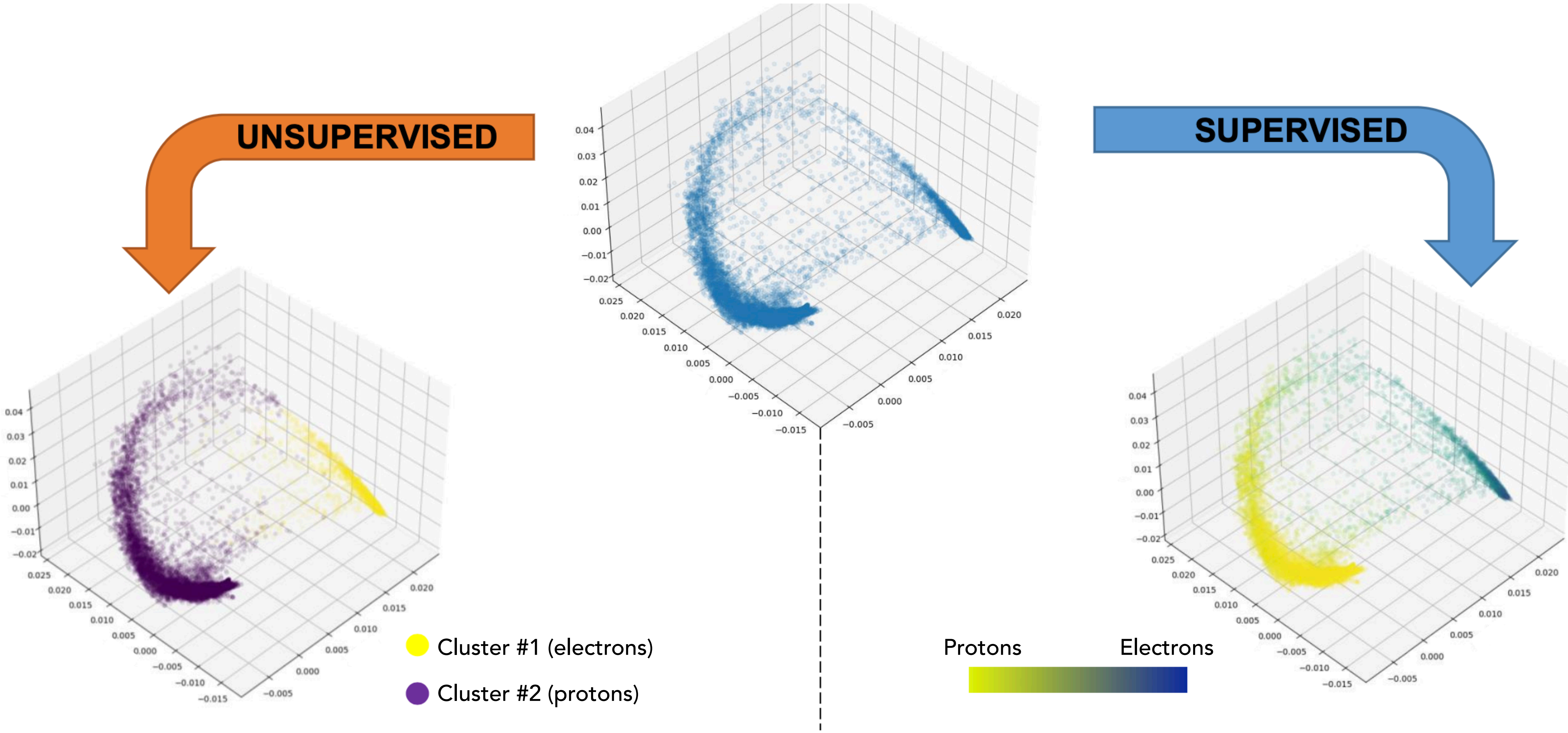
Selection of variables → Dimensionality Reduction → Clustering

Selected algorithm:
Hierarchical Clustering



Unsupervised Learning

Example of the unsupervised learning results on LAT data



Conclusions

- We developed two new **Machine Learning analyses for computing the CRE spectrum with Fermi LAT data.**
- We conducted a **Supervised Learning analysis using Neural Networks**. We are running tests to confirm the definitive spectrum.
- We found some promising results using **Unsupervised Learning techniques,** which were never applied before to Fermi LAT data.

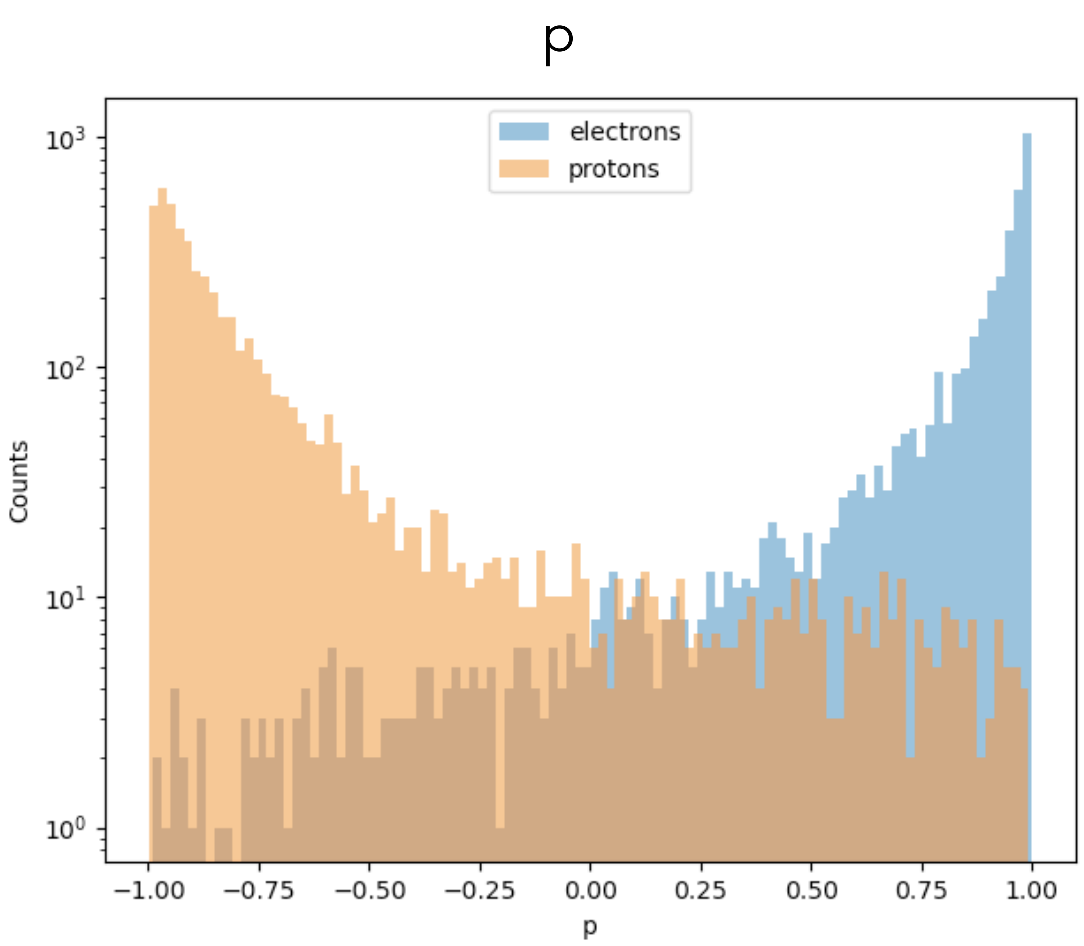
Thank you for the attention

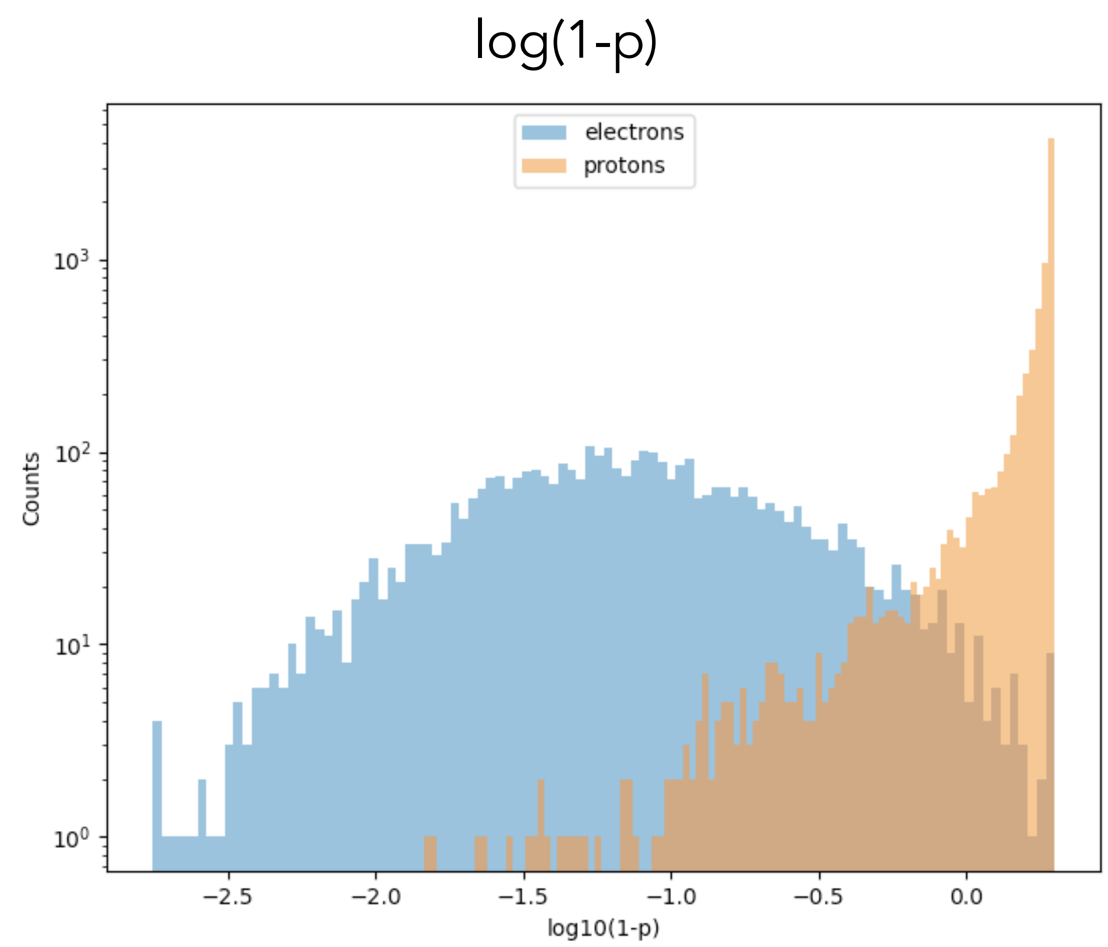
&

Stay tuned!

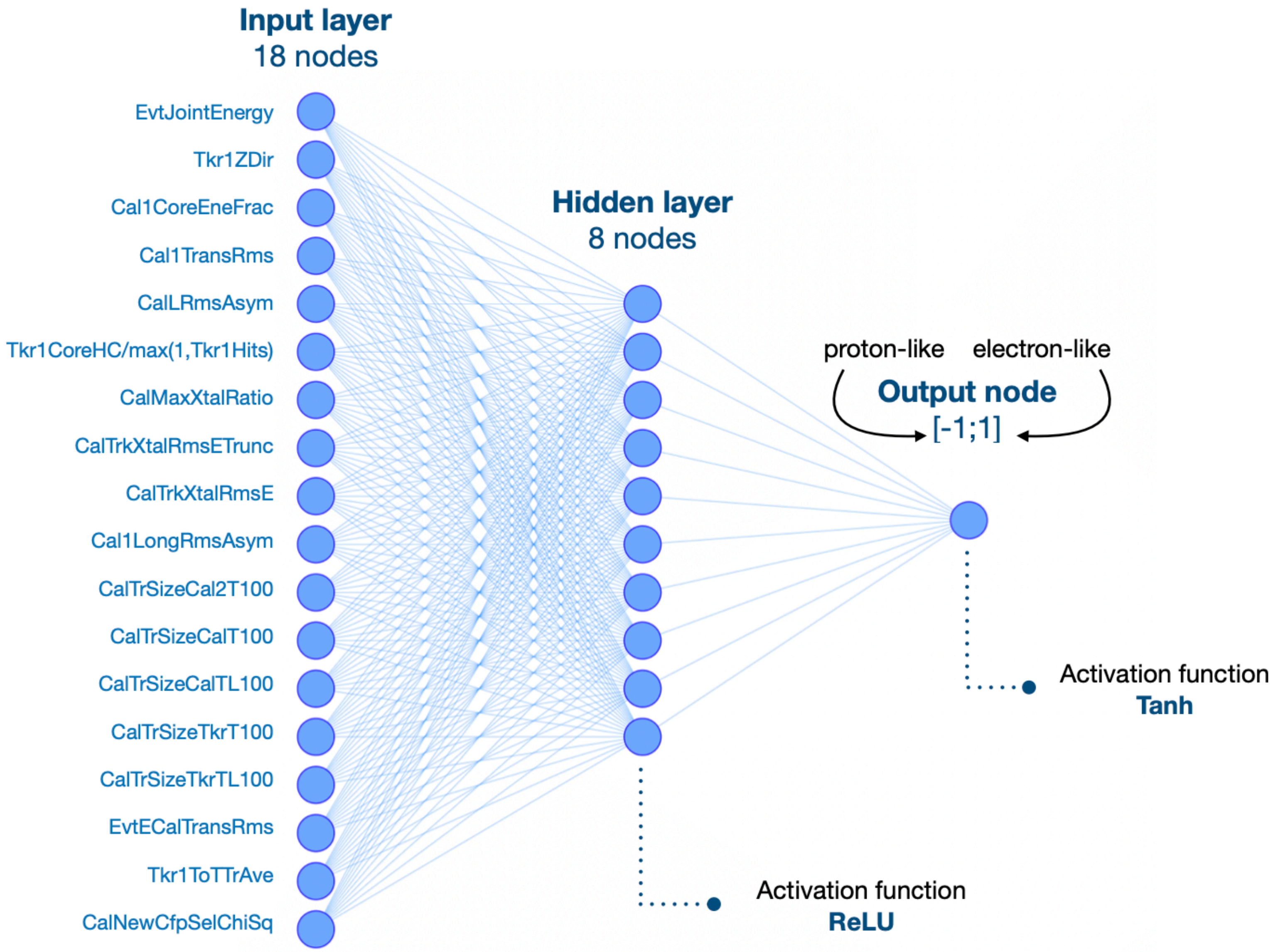
Backup

log(1-p)


p
electrons
protons
Counts
10^3
10^2
10^1
10^0
−1.00
−0.75
−0.50
−0.25
0.00
0.25
0.50
0.75
1.00
p

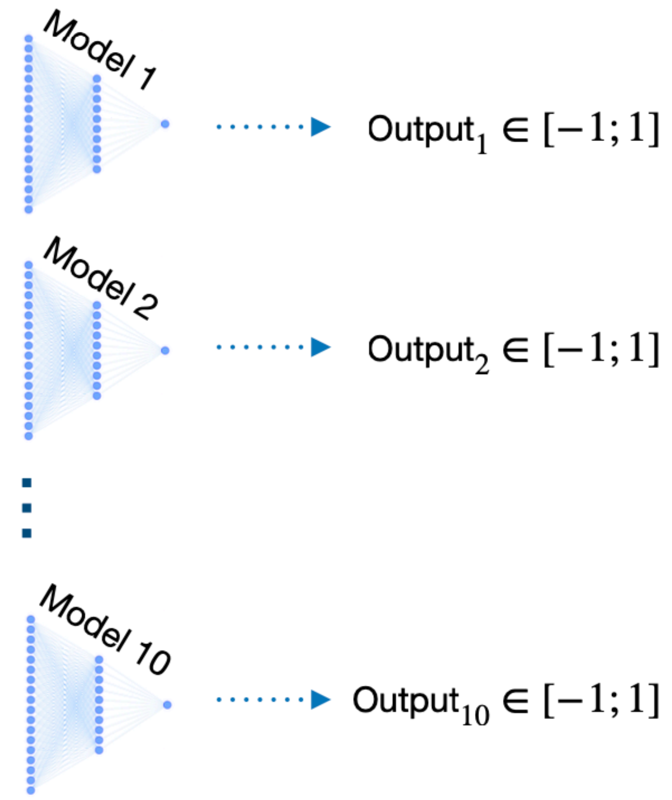

log(1-p)
electrons
protons
Counts
10^3
10^2
10^1
10^0
−2.5
−2.0
−1.5
−1.0
−0.5
0.0
log10(1-p)

NN details

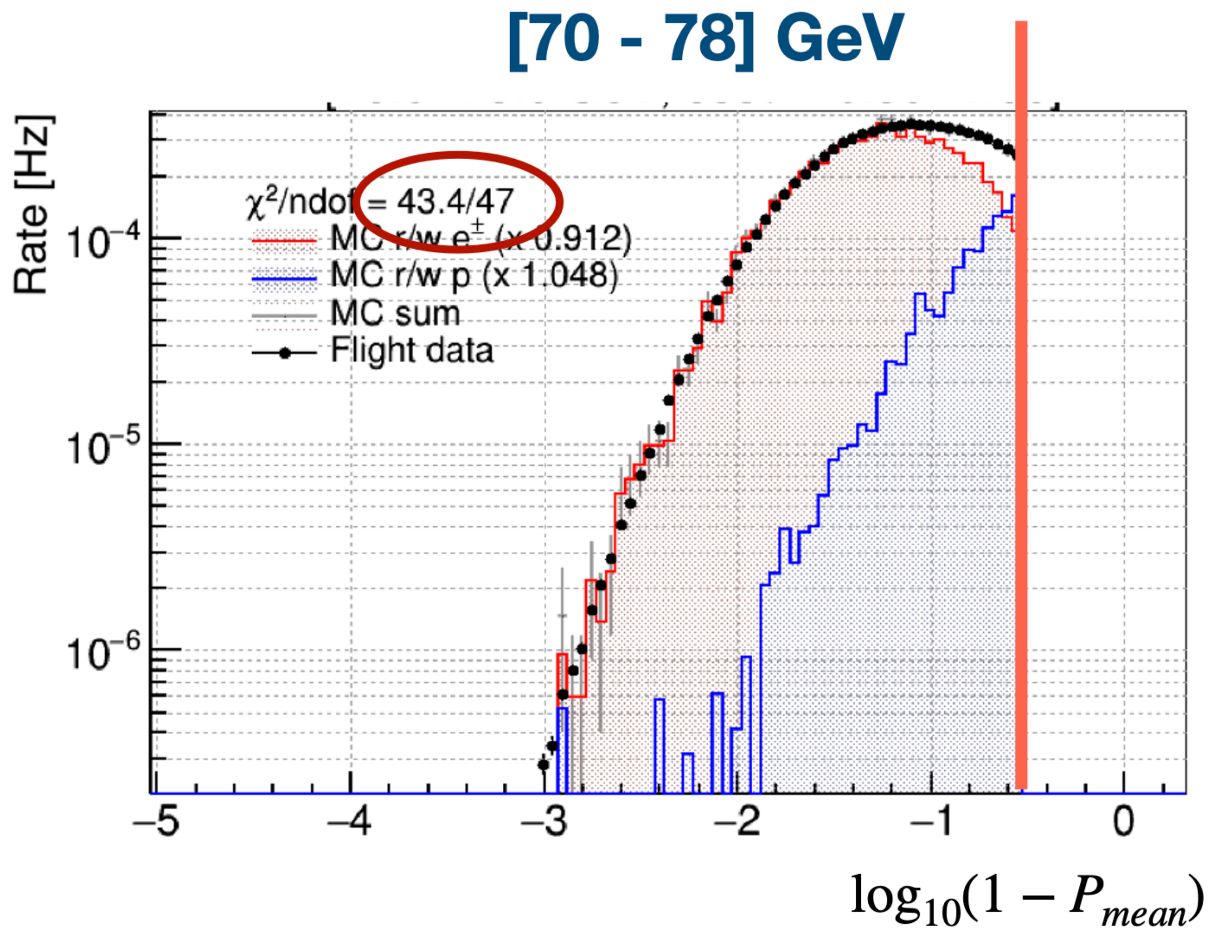


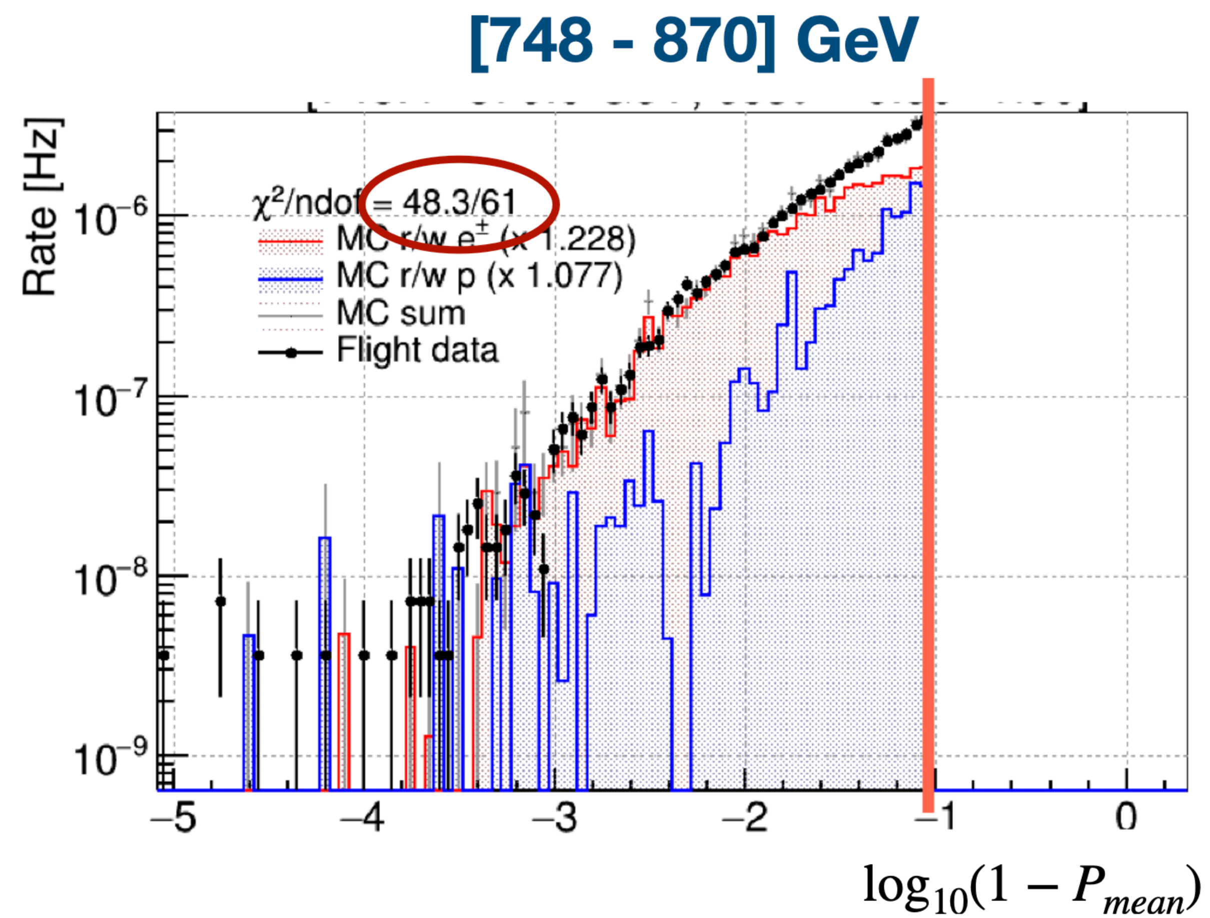
"Ensemble learning" method

Training of multiple models to get a solid prediction



Selection cut





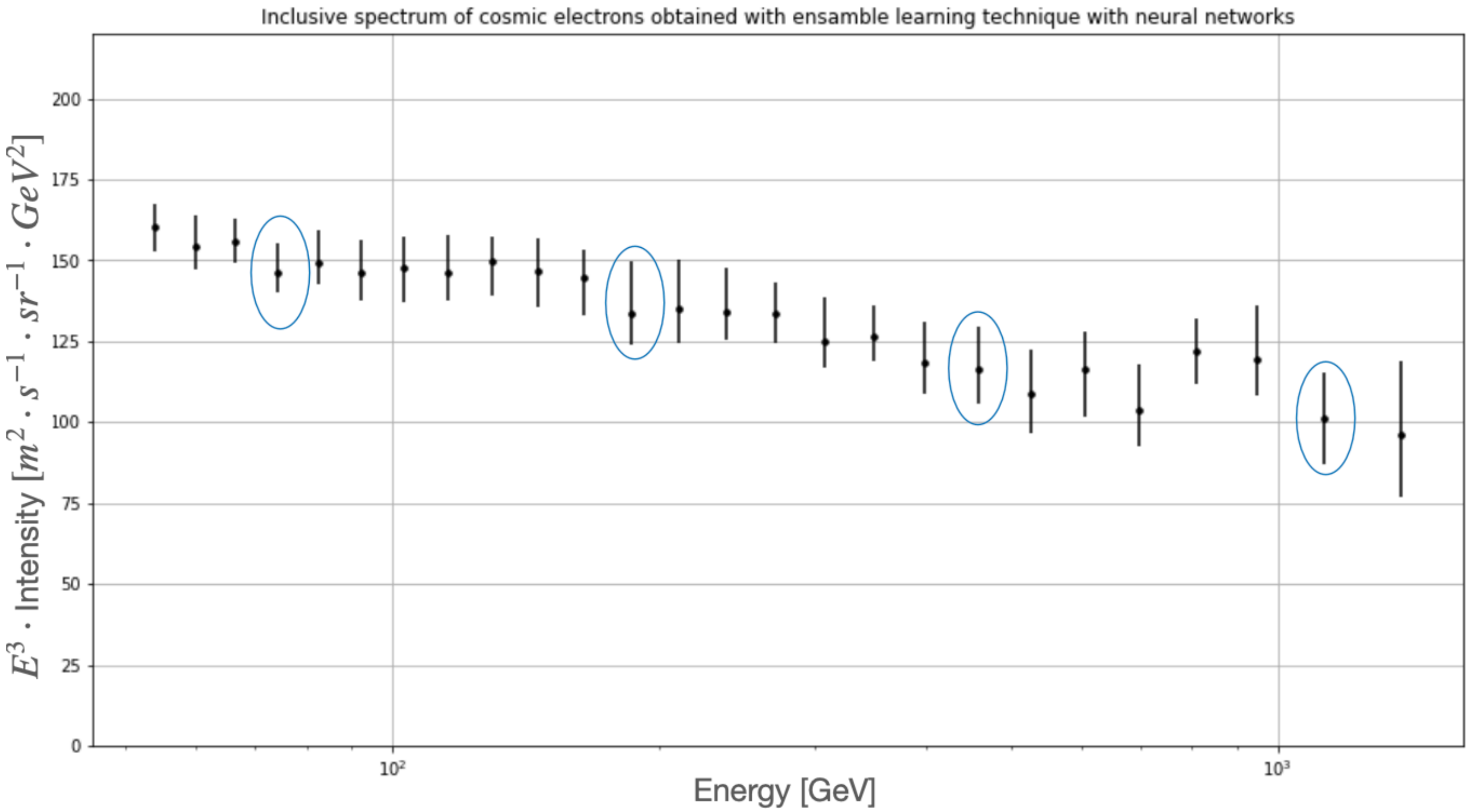
Uncertainties

Statistical uncertainty σ_{stat}

Uncertainty on the acceptance σ_{eff}

Uncertainty on the contamination σ_{cont}

$$\sigma_{tot} = \sqrt{\sigma_{stat}^2 + \sigma_{eff}^2 + \sigma_{cont}^2}$$



Energia [GeV]	σ_{stat}	σ_{eff}	σ_{cont}
74	0.1%	5.5%	3.2%
165	0.2%	7.0%	4.8%
400	0.5%	9.35%	8.2%
1127	1.7%	13.95%	13.7%

Dimensionality reduction

Is there an optimal dimension?