Electron flux

Analysis of fiducial events

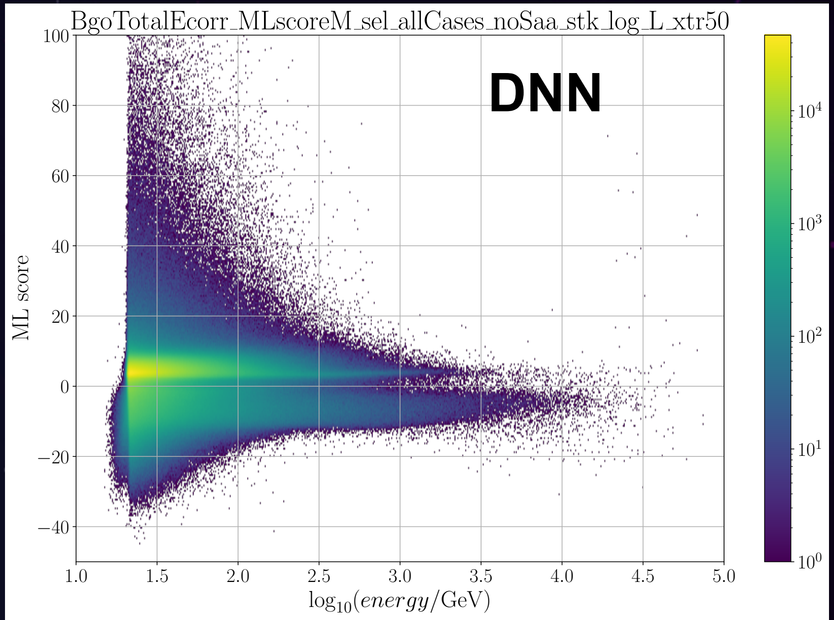
UNIVERSITÉ DE GENÈVE

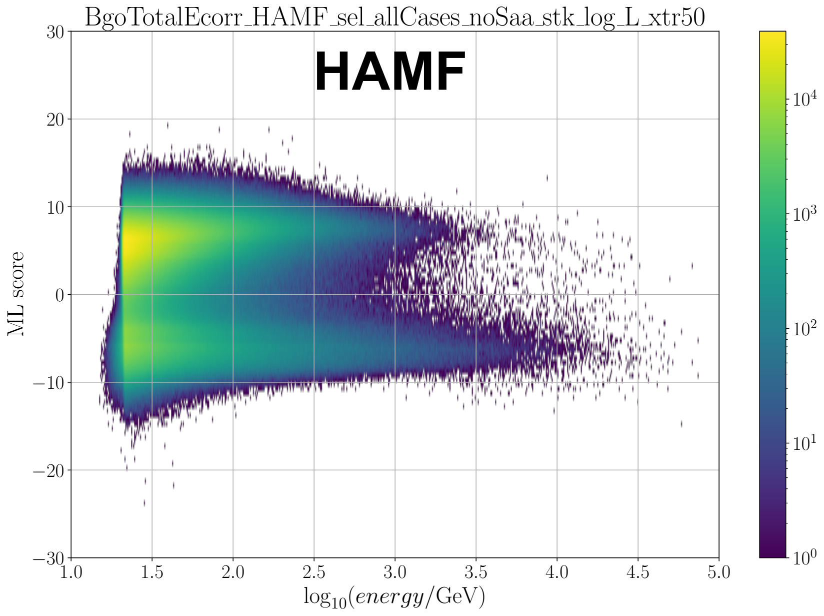
Swiss National Science Foundation

Paul Coppin

Overview

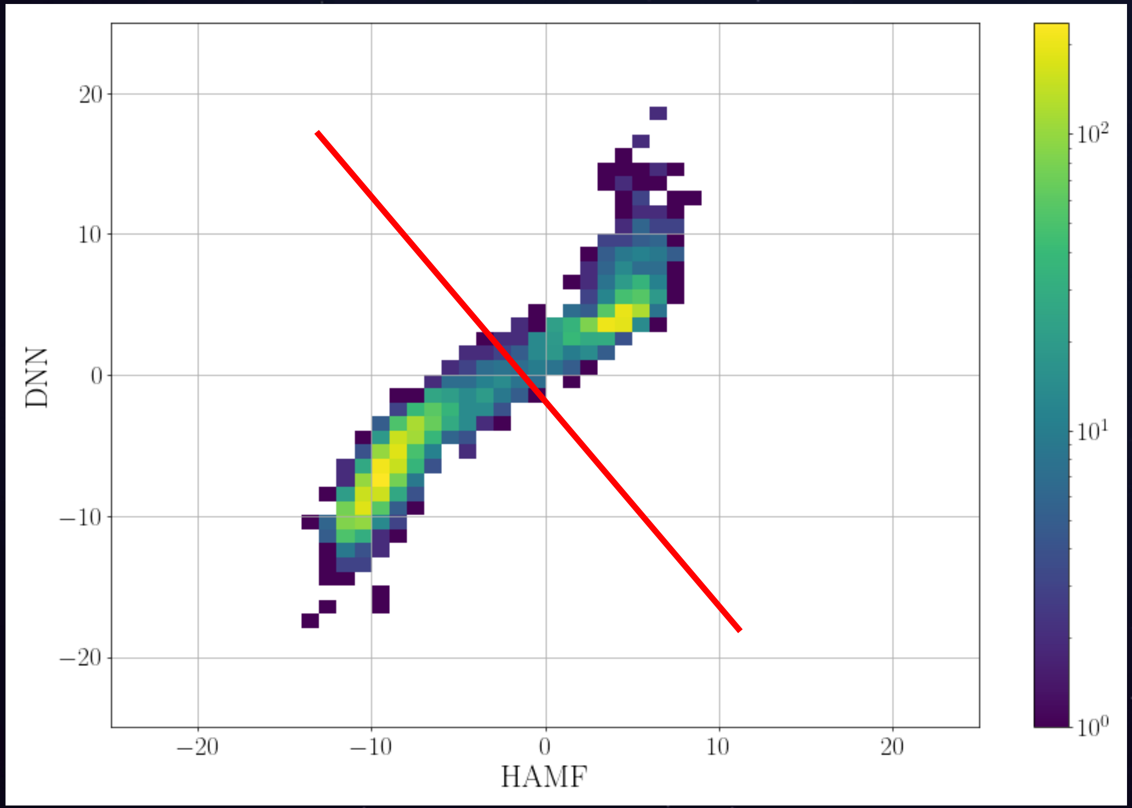
- Electron flux for contained events
- Pre-selection by Prof. Wu
- Machine learning classifier by:
  - Droz David (DNN)
  - Li Manbing (HAMF)
- Flux:
  - Extend to 20 TeV
  - Evaluate background uncertainty
  - Significance of break at 4 TeV





Combined score

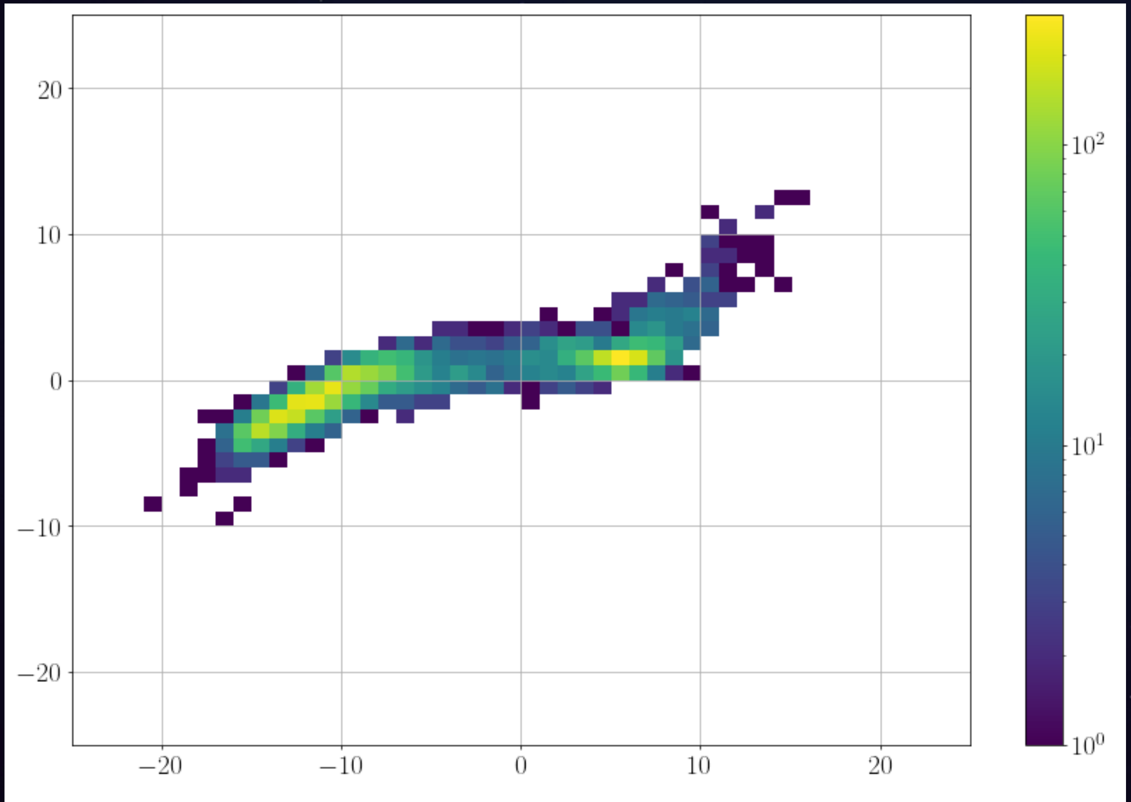
- Place a cut on both HAMF & DNN score



Proton: bottom left
Electron: top right
Showing one energy bin: 758.6 to 871.0 GeV

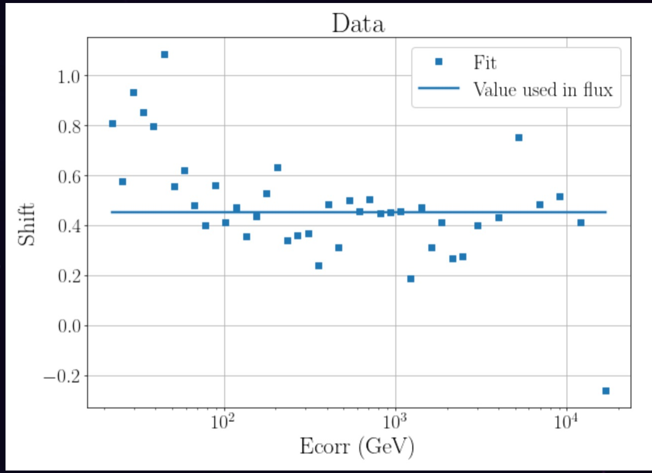
Combined score

- Place a cut on both HAMF & DNN score



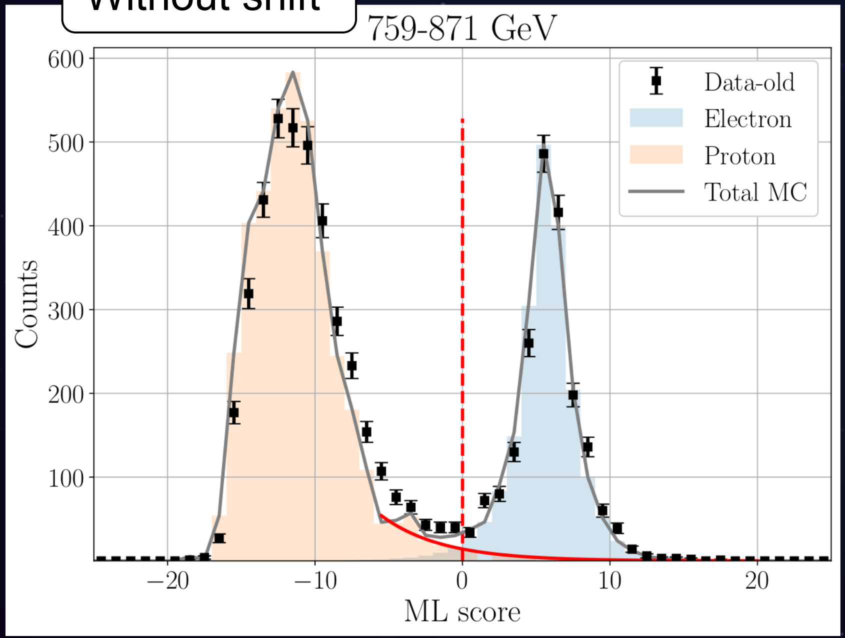
Rotated by 30 degrees
→ Give more weight to HAMF model

Charge distributions

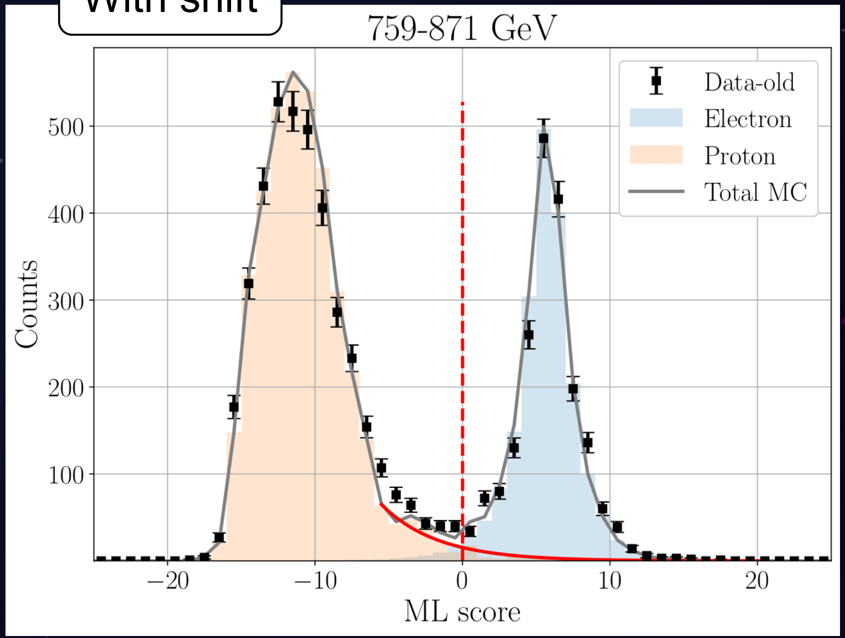


- Good match between data and MC after minor shift of $score \rightarrow score + 0.45$ for proton

Without shift

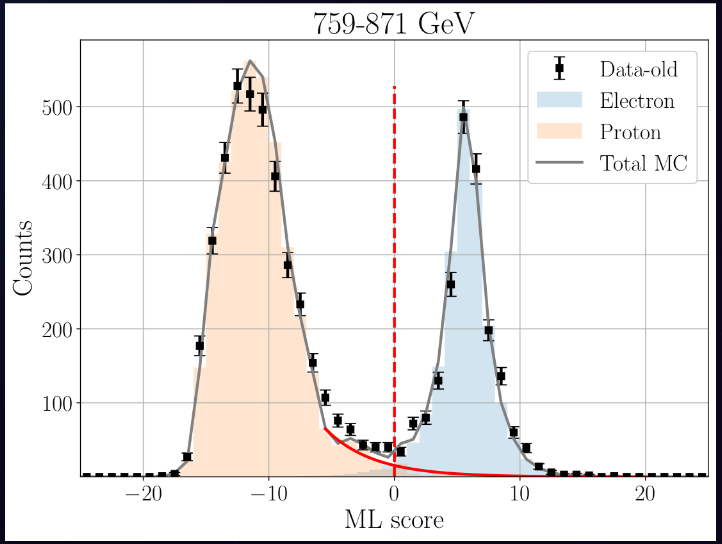


With shift

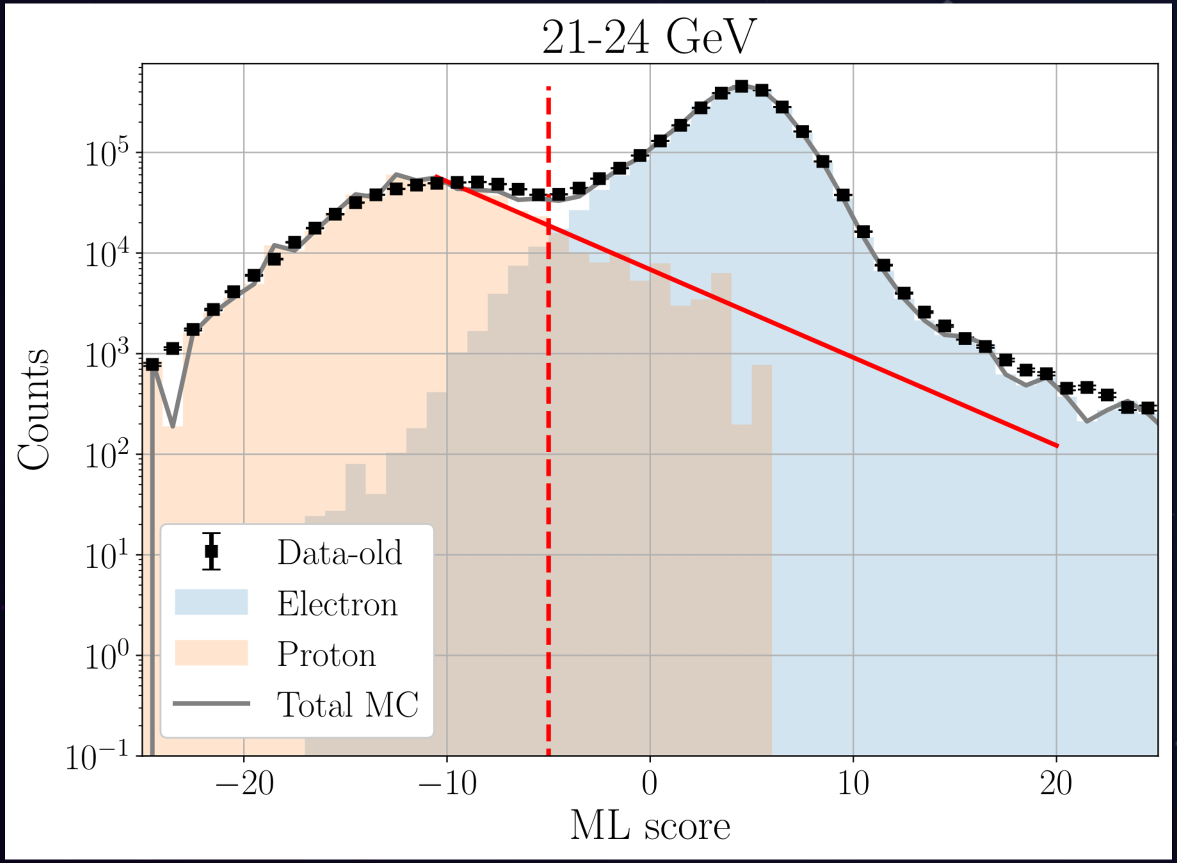


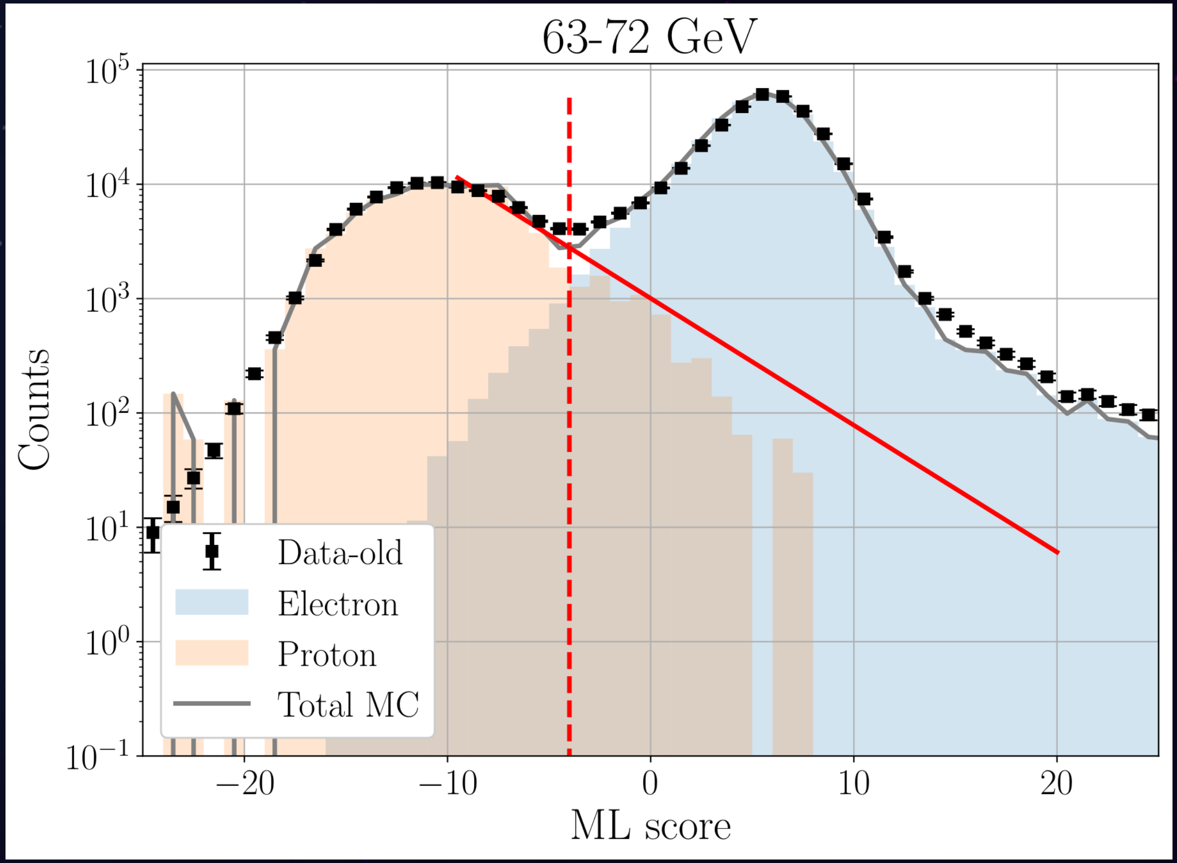
Cut on ML value

- Per energy bin:
  - Fit height of electron and proton MC to data
  - Determine ML-cut based on $\alpha = \frac{signal}{\sqrt{signal+background}}$
    - Shift cut to the right if it improves α
    - Else, use same cut as previous energy
    - Don't go beyond score of 4

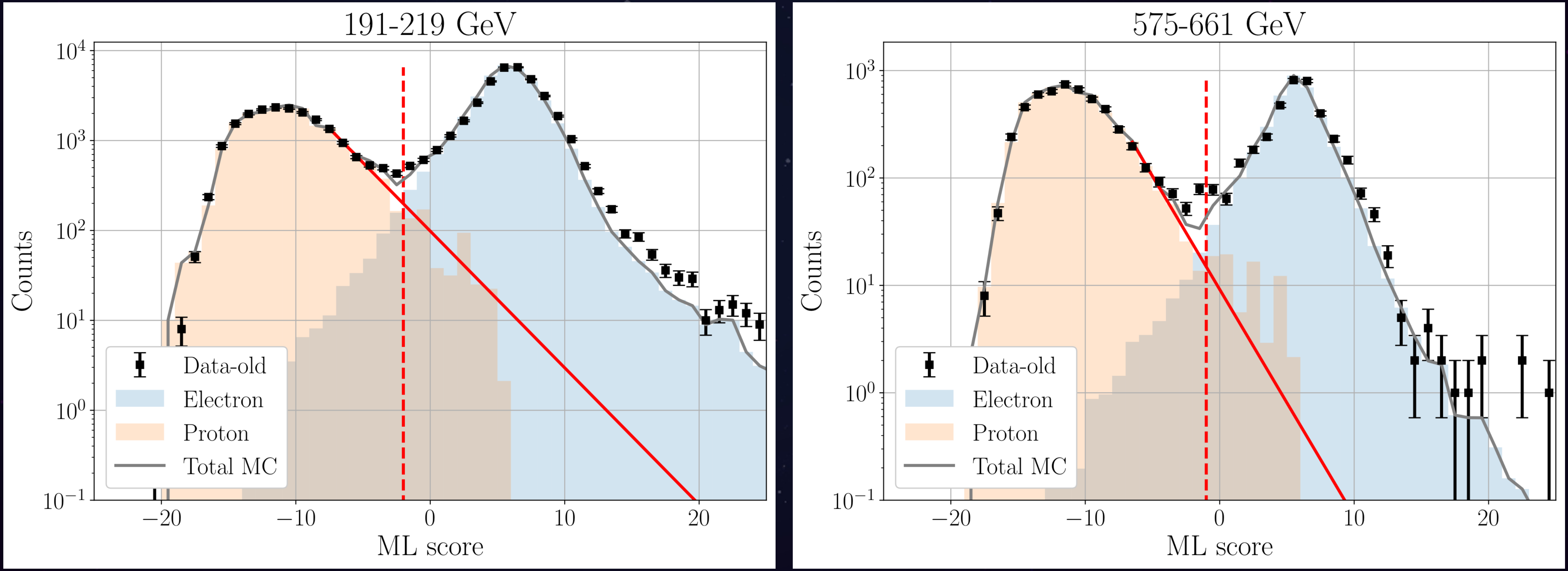


Score distributions

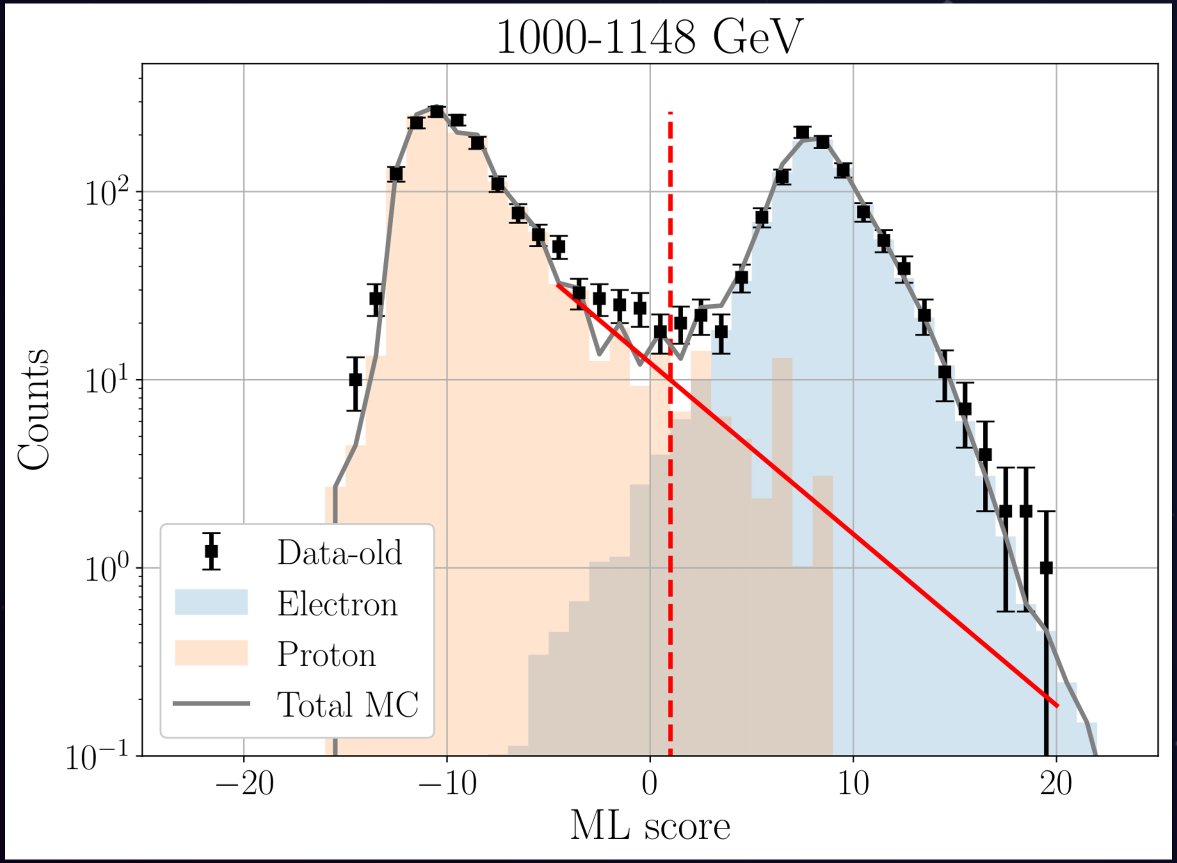


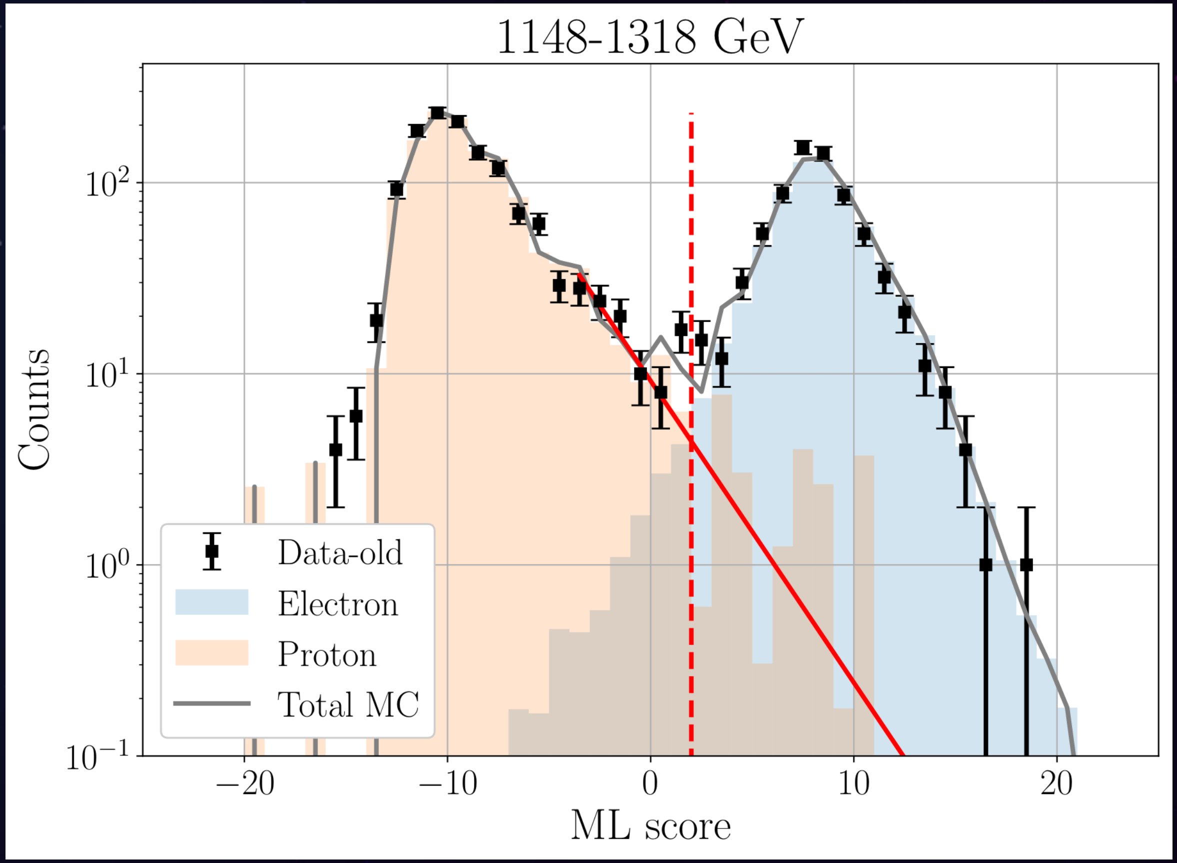


Score distributions

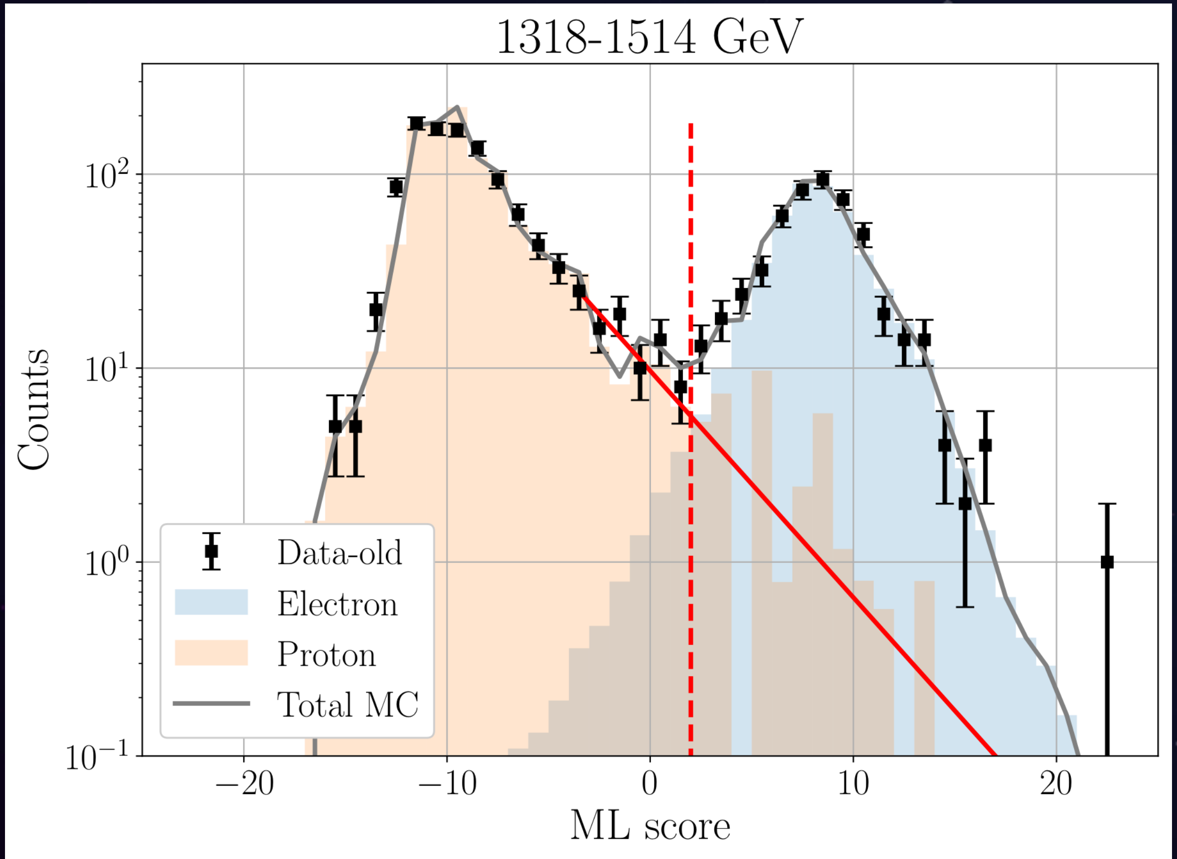


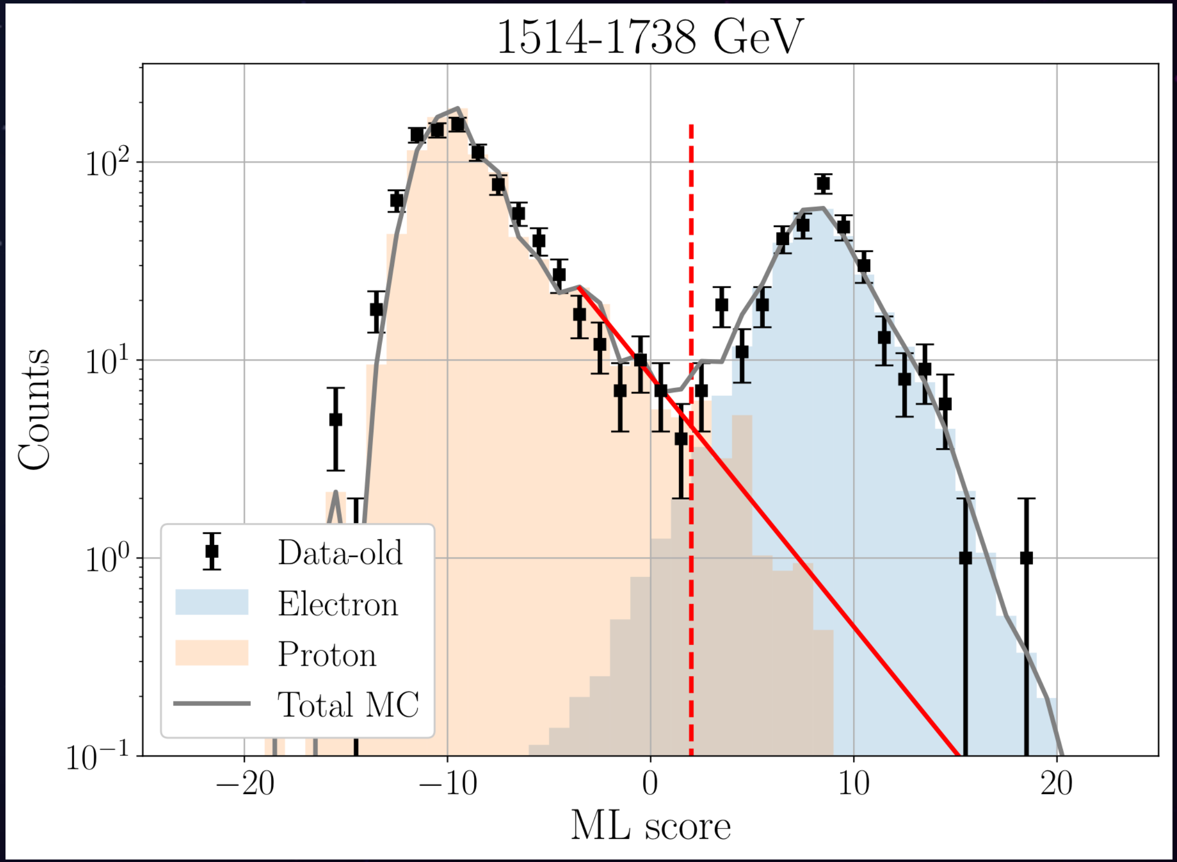
Score distributions



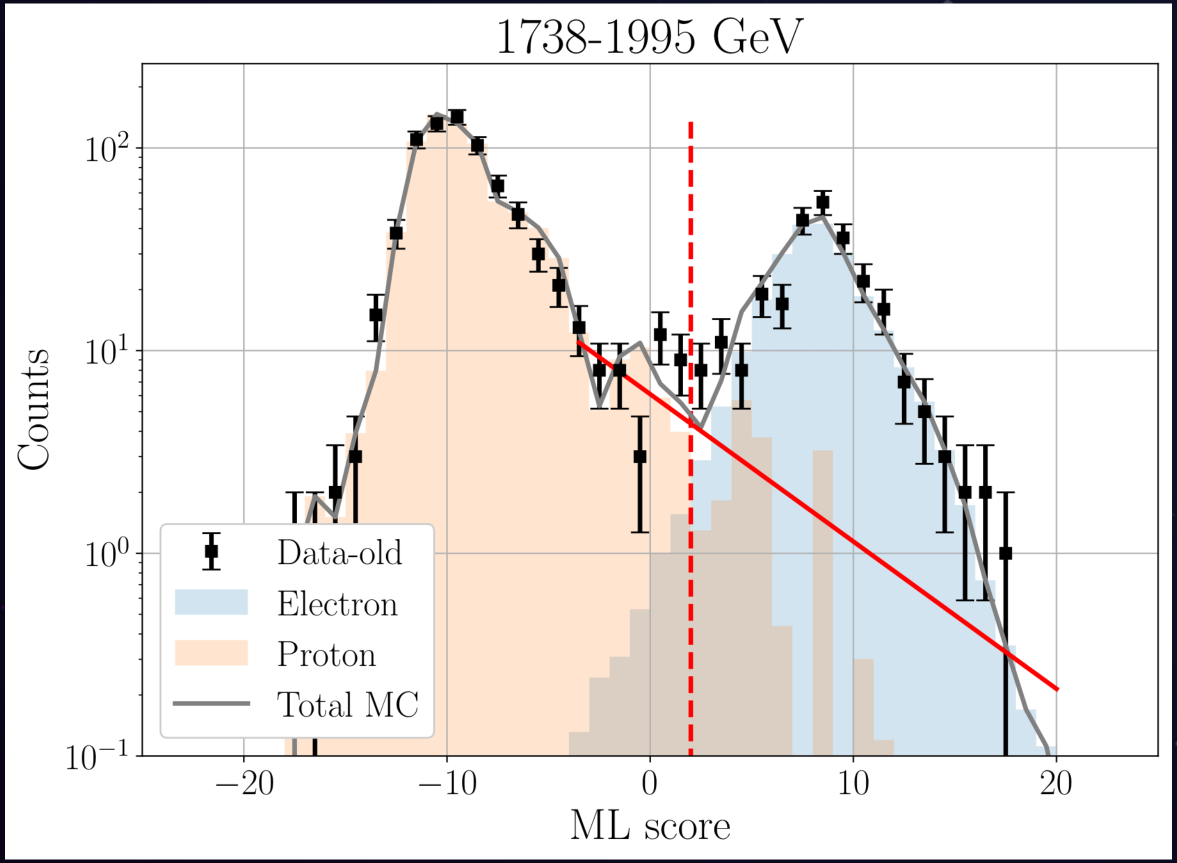


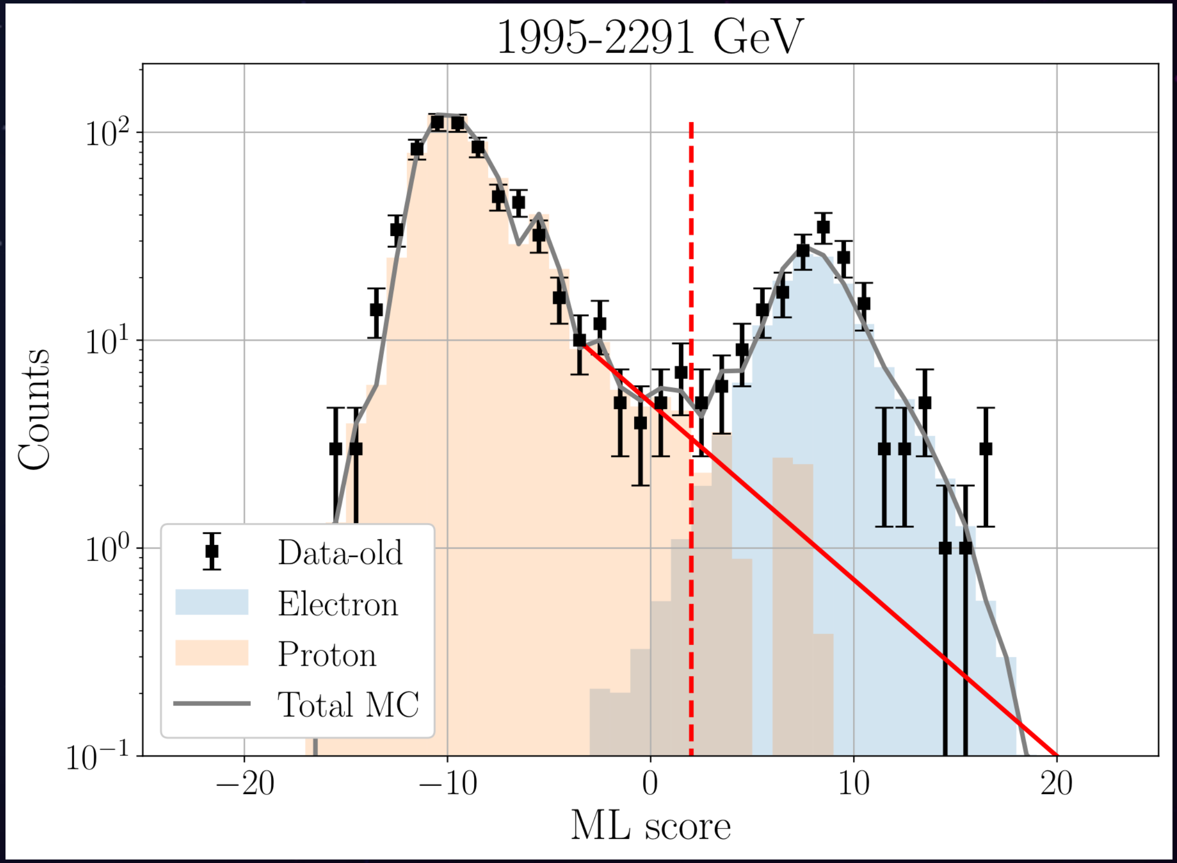
Score distributions



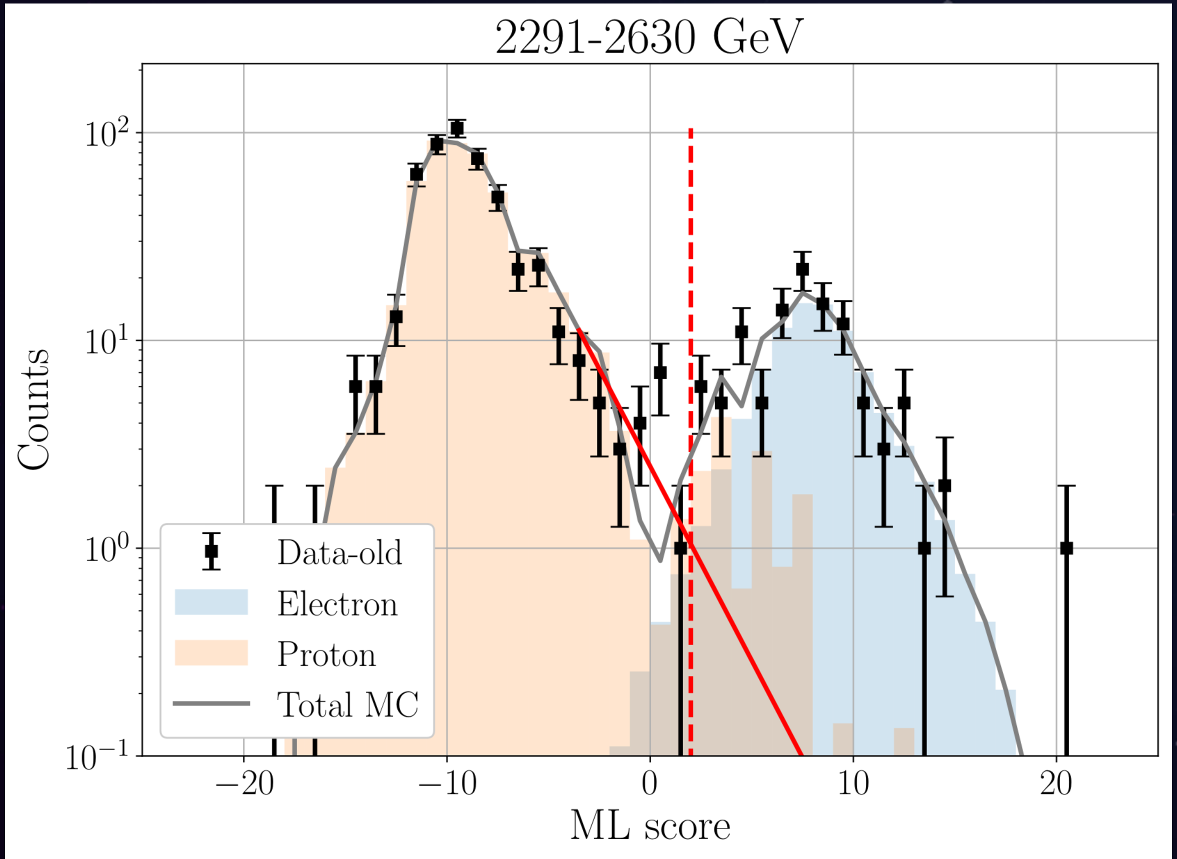


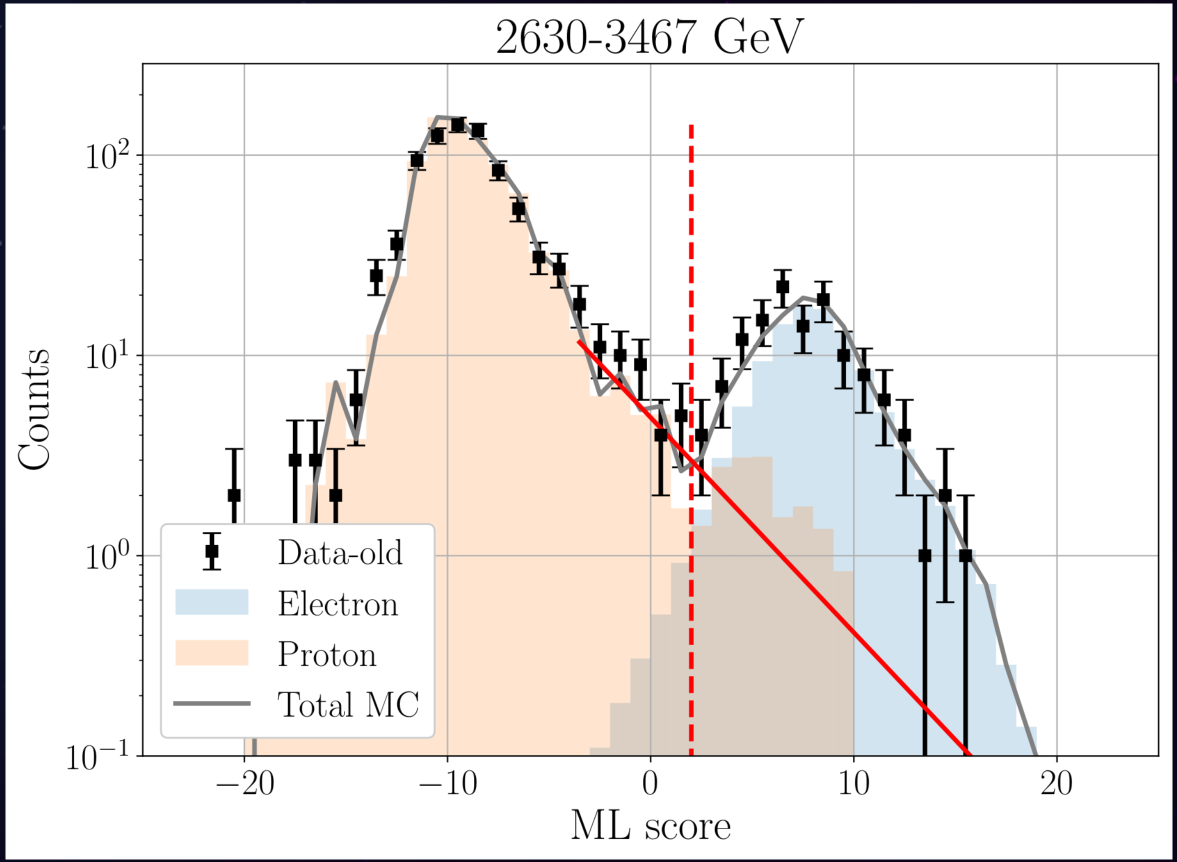
Score distributions



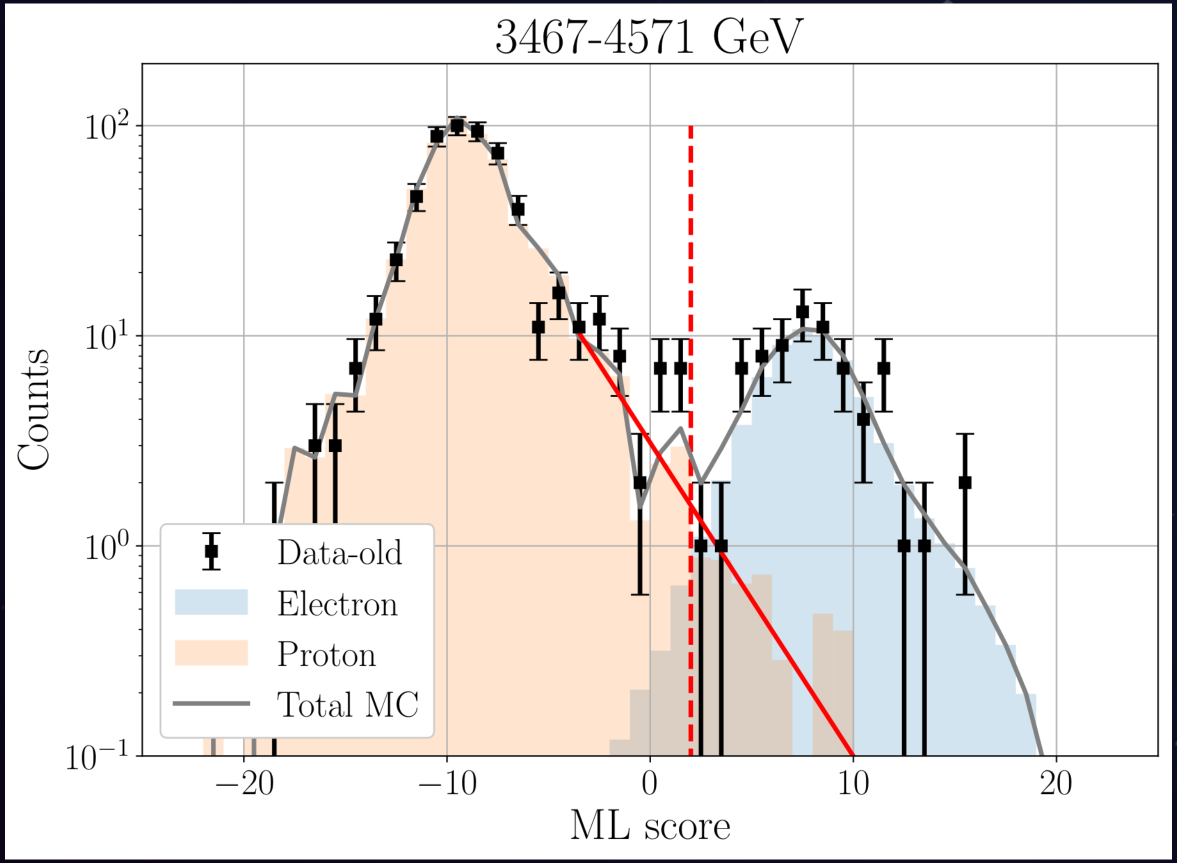


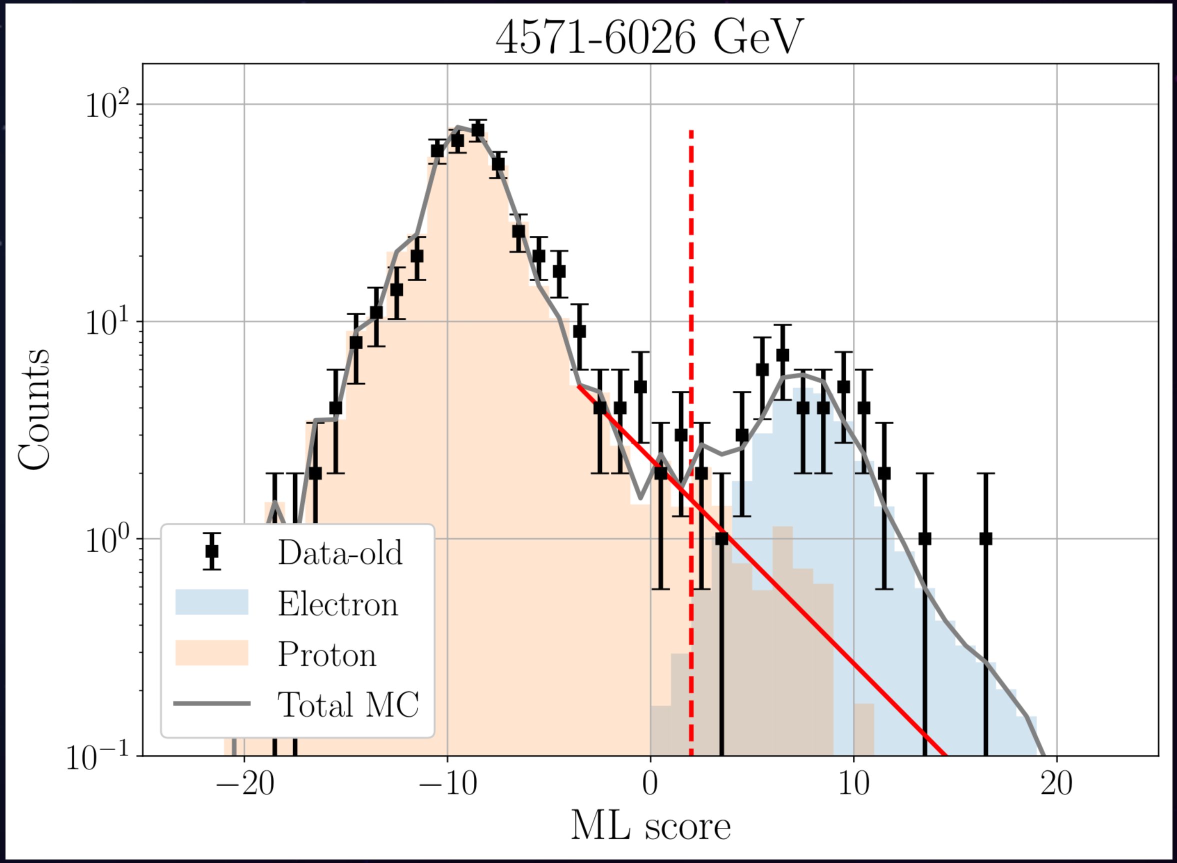
Score distributions



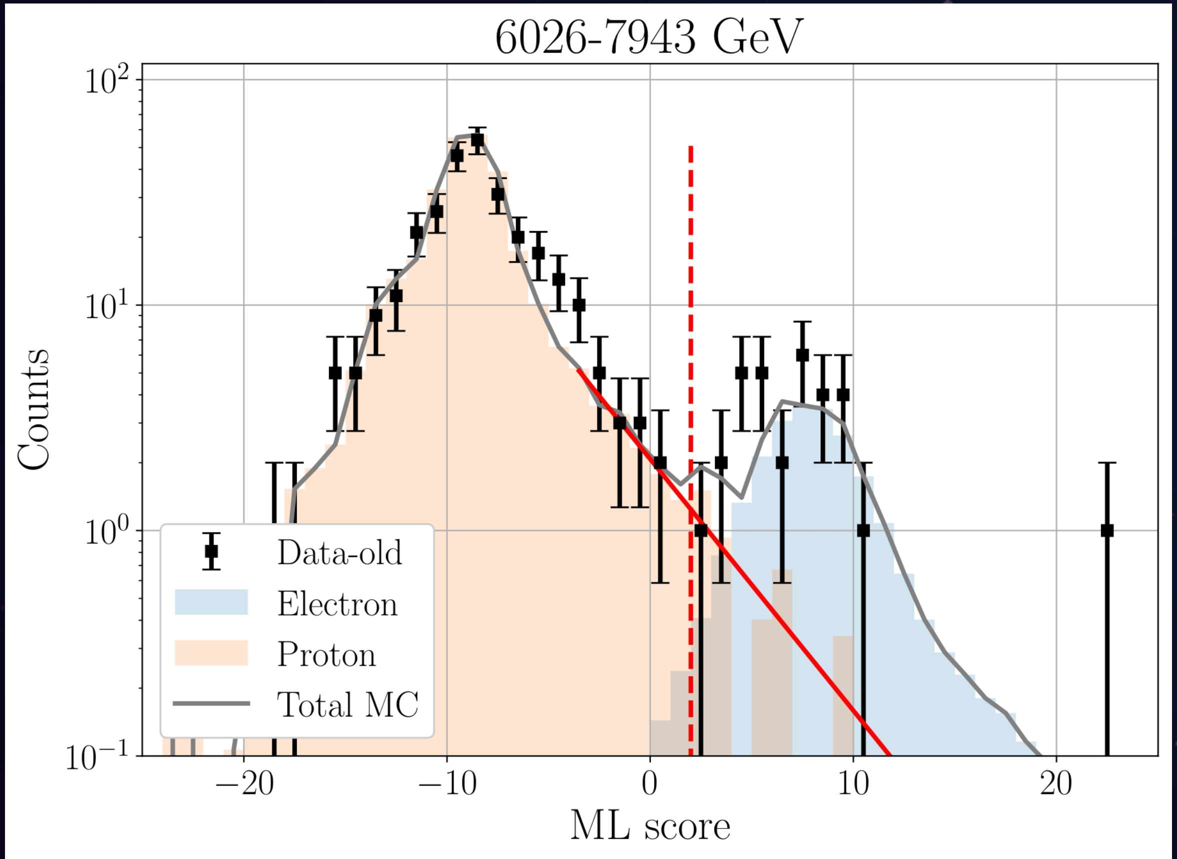


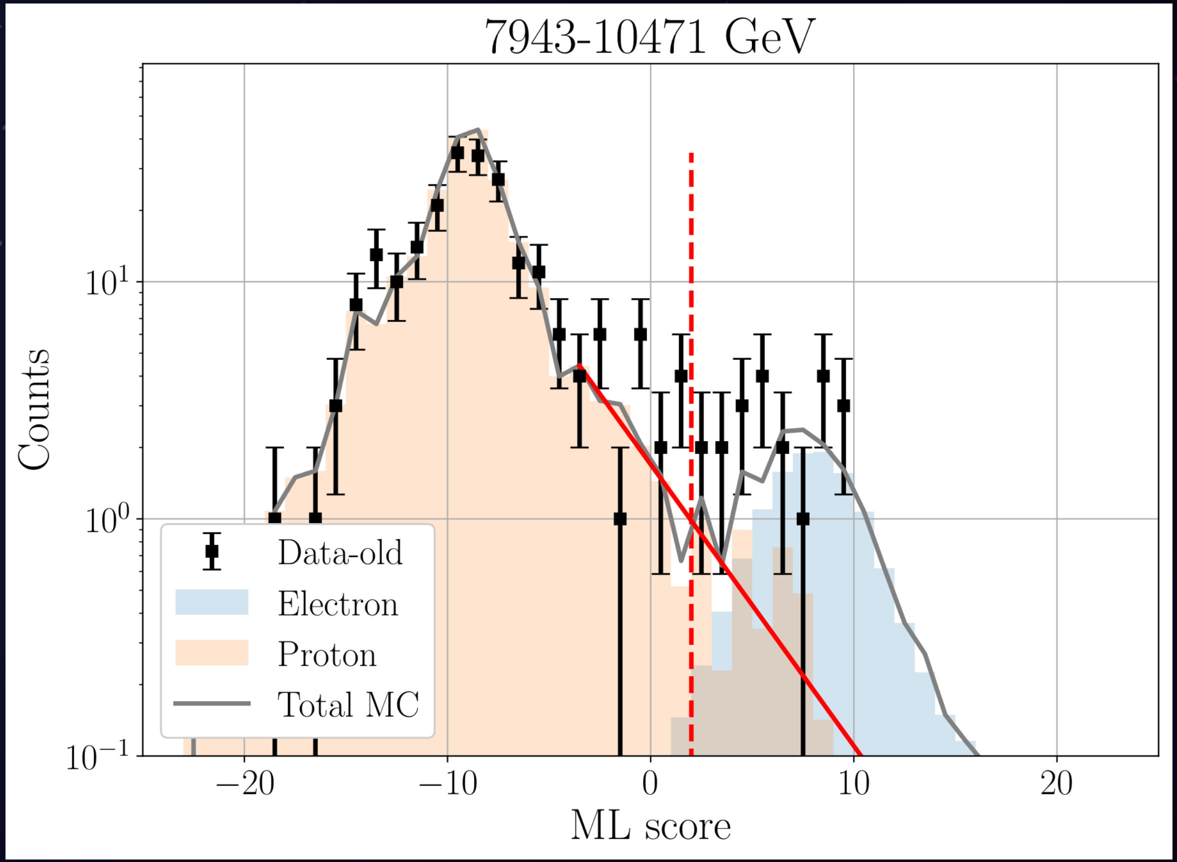
Score distributions



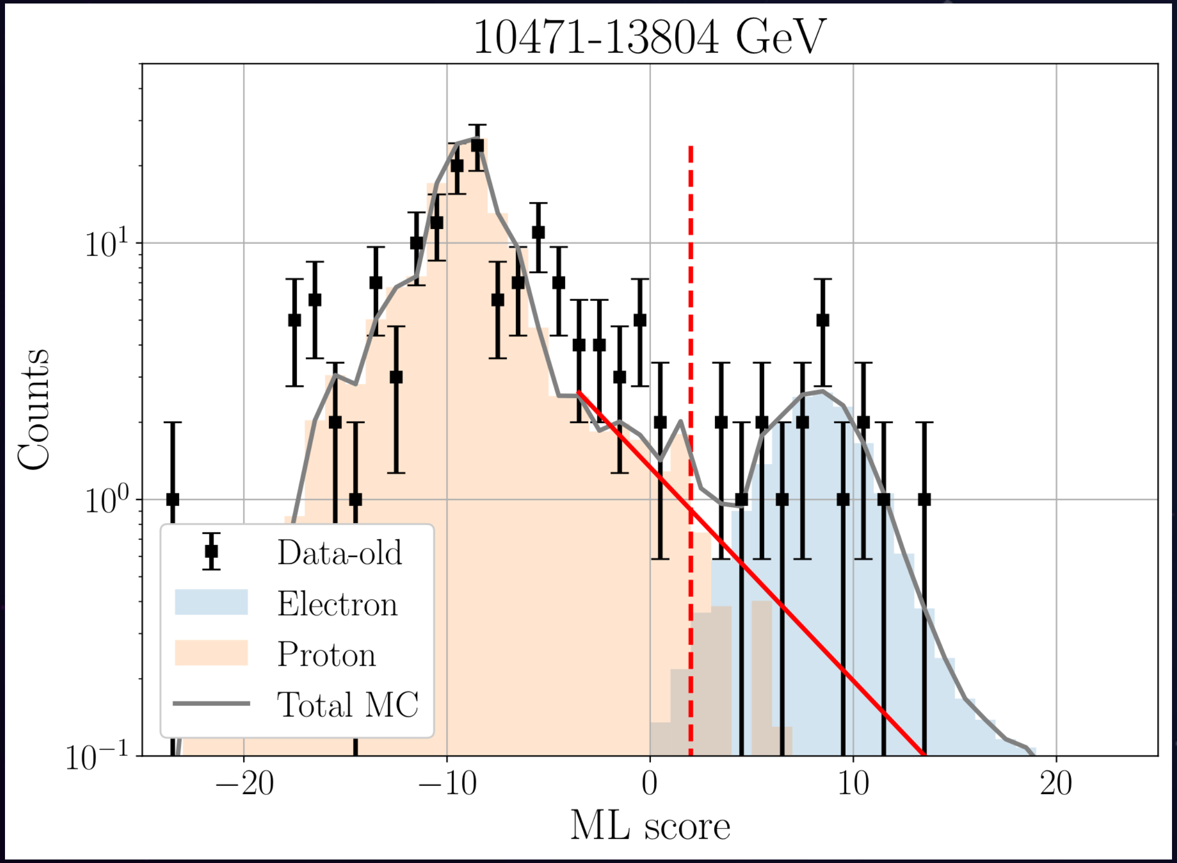


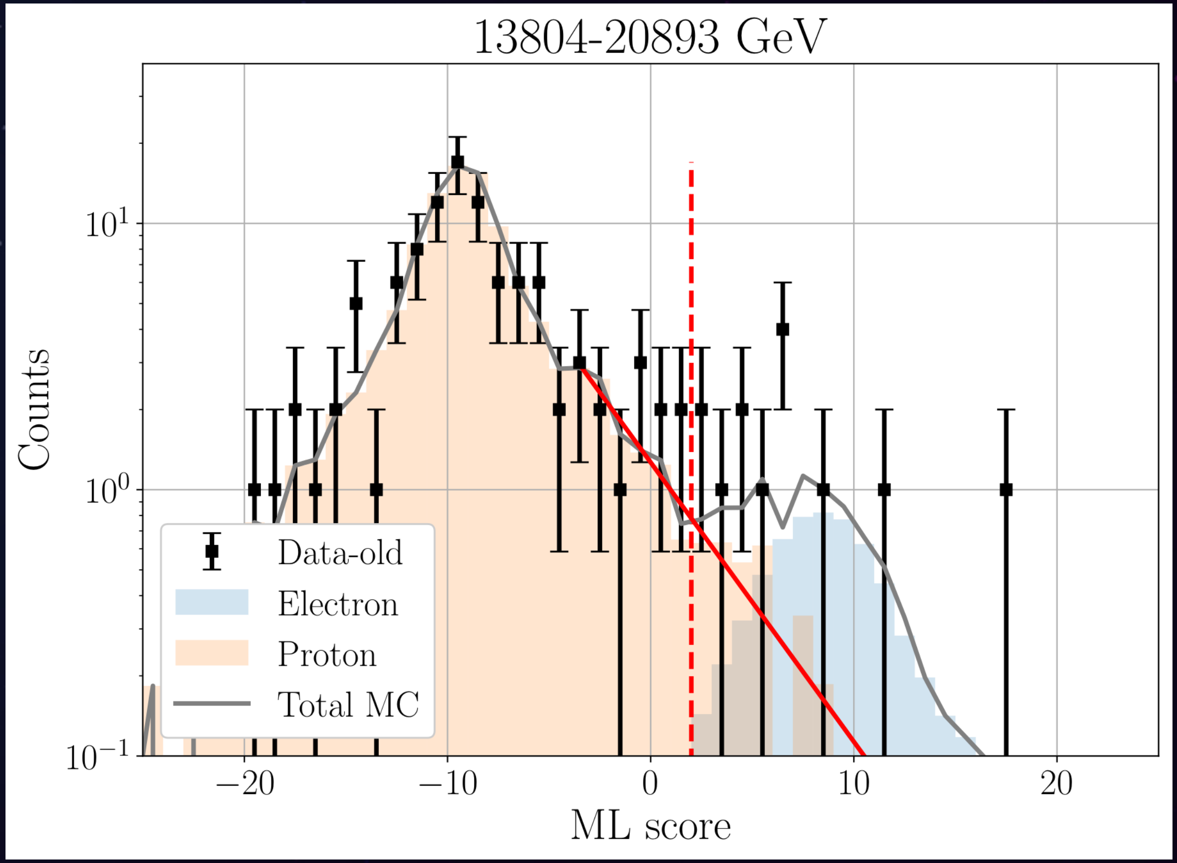
Score distributions



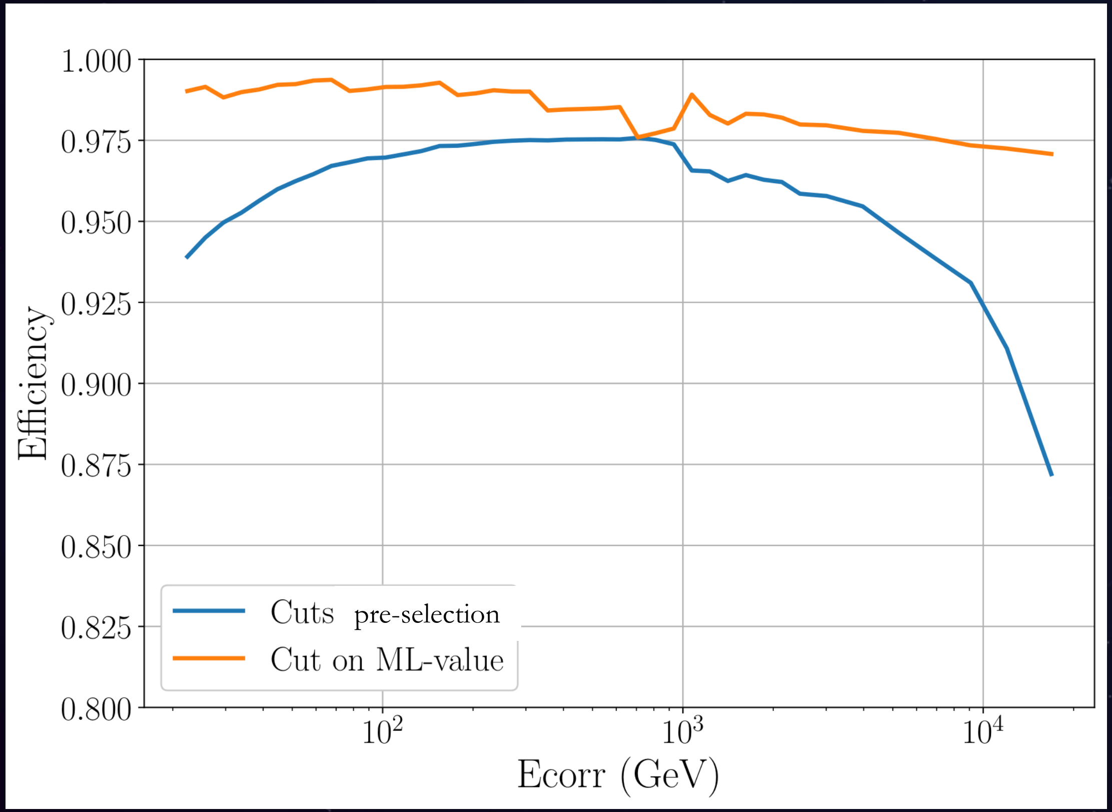


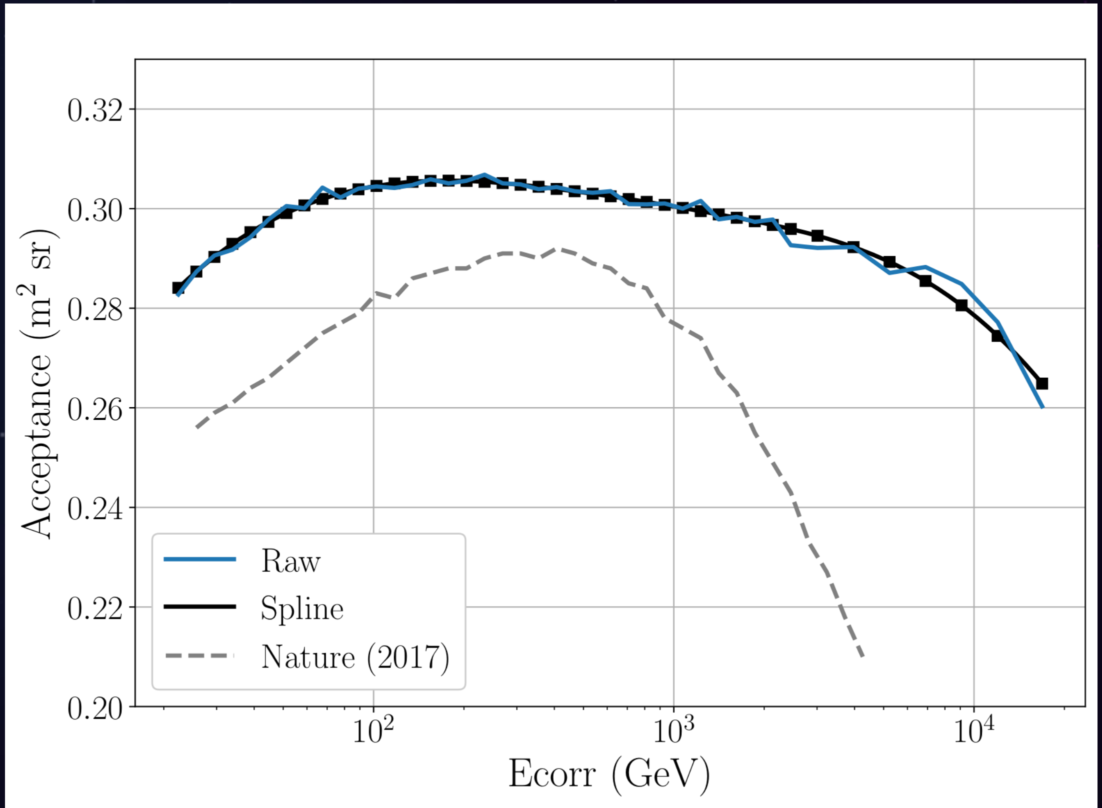
Score distributions





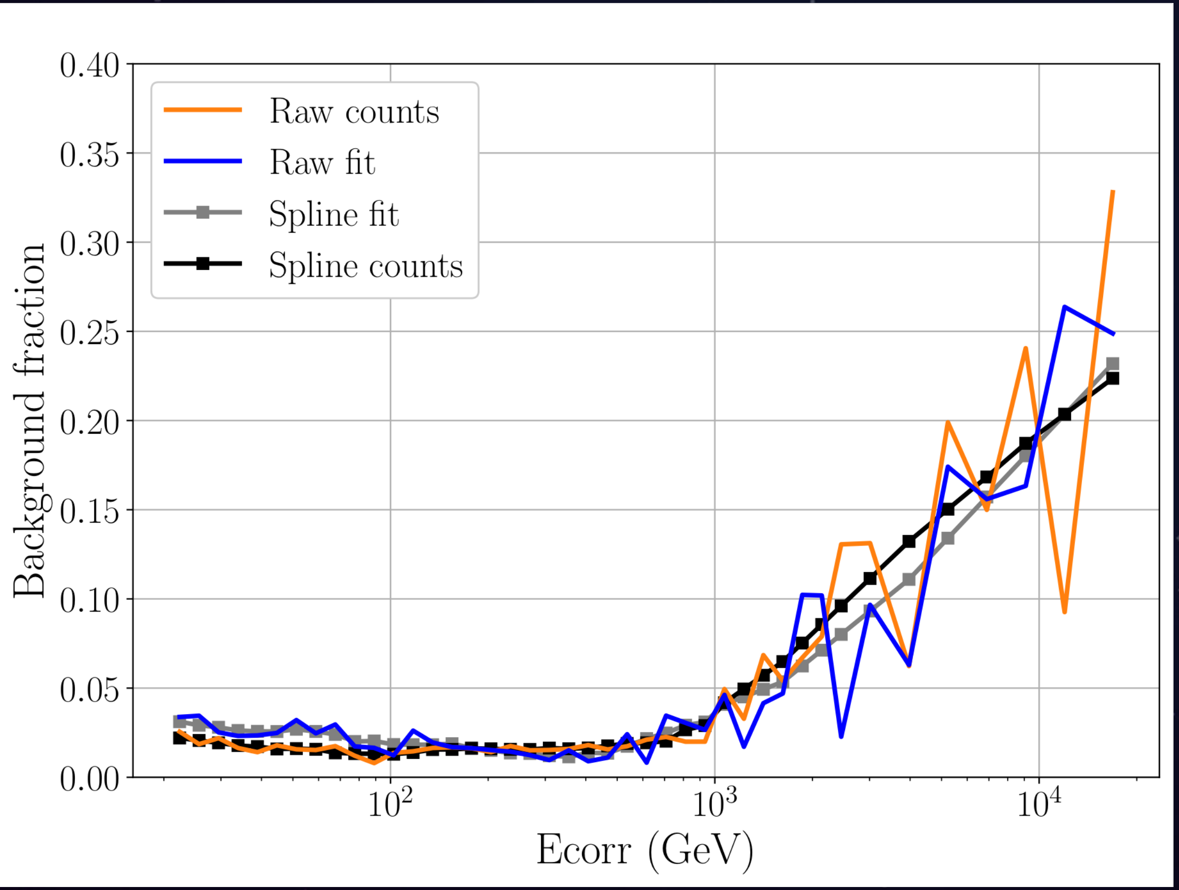
Efficiency and acceptance





Background fraction

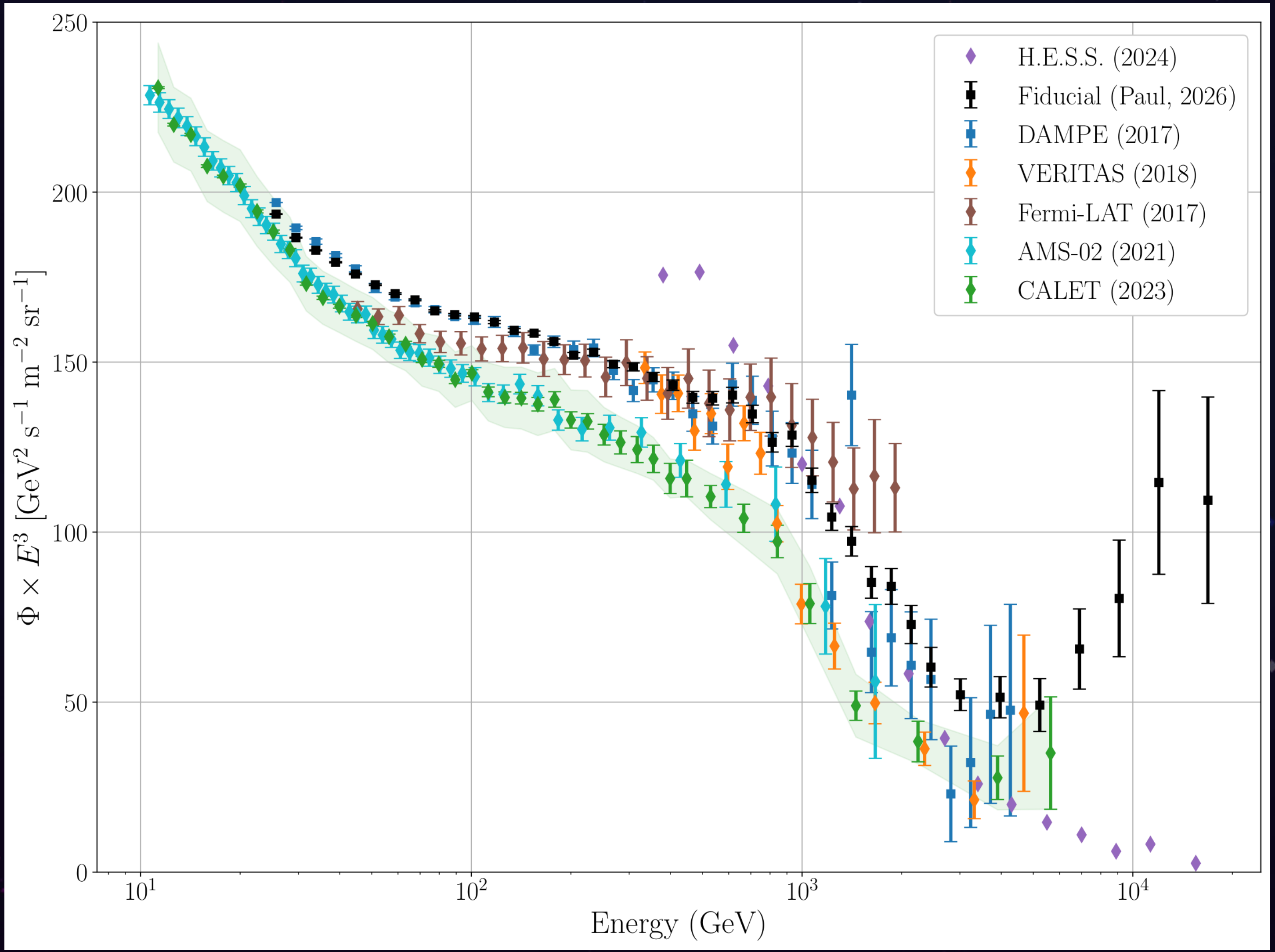
Counting: `MC-proton / (MC-electron + MC-proton)`



Livetime

- Extracted from database:
  - ROOT.DmpSvcLiveTime.GetInstance(). GetLiveTime(T_min, T_max)
  - 2015-12-30 to 2025-10-01
  - 236375620 seconds

Fiducial flux



Flux computation: 2 approaches

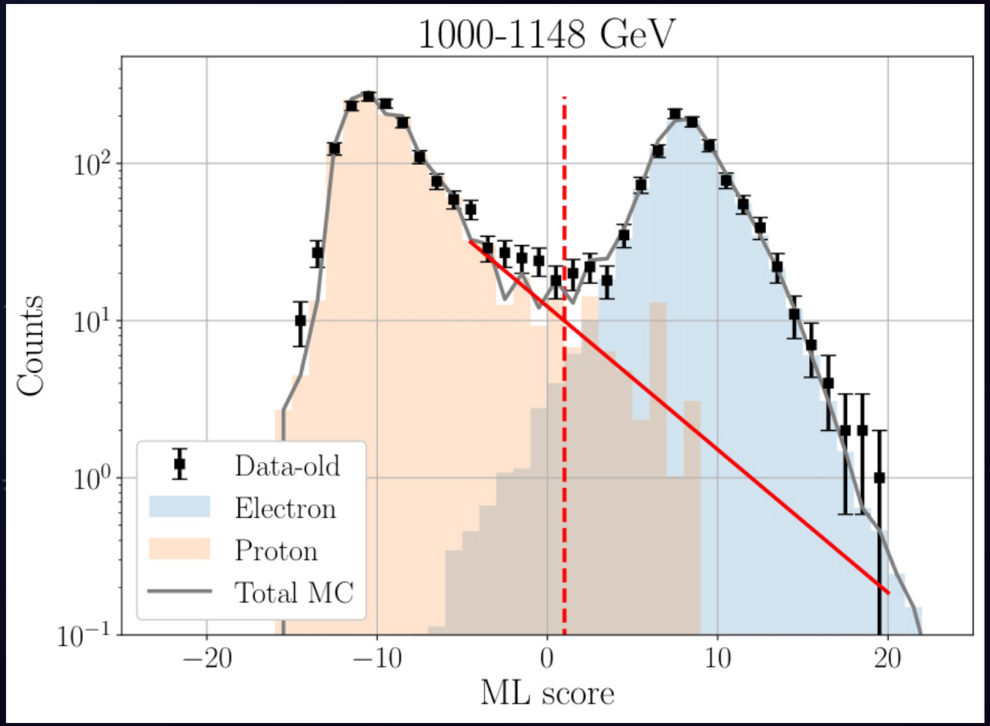
- Cut and count:
  - Impose a fixed cut → red line
  - Background fraction:

$$\text{bg} = \frac{N^{MC}_{proton}(\Omega > cut)}{N^{MC}_{proton}(\Omega > cut) + N^{MC}_{electron}(\Omega > cut)}$$

  - Efficiency:

$$\text{eff} = \frac{N^{MC}_{electron}(\Omega > cut)}{N^{MC}_{electron}}$$

  - $\text{counts} = \dfrac{(1-\text{bg}) \cdot N^{data}(\Omega > cut)}{\text{eff}}$

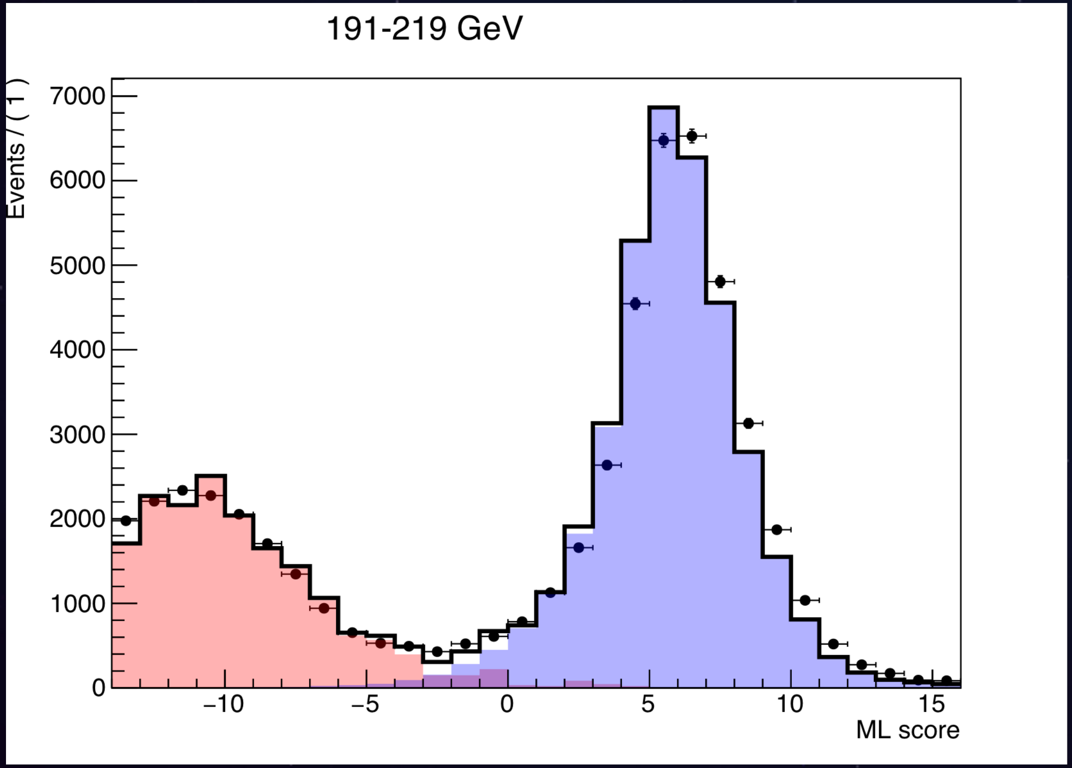


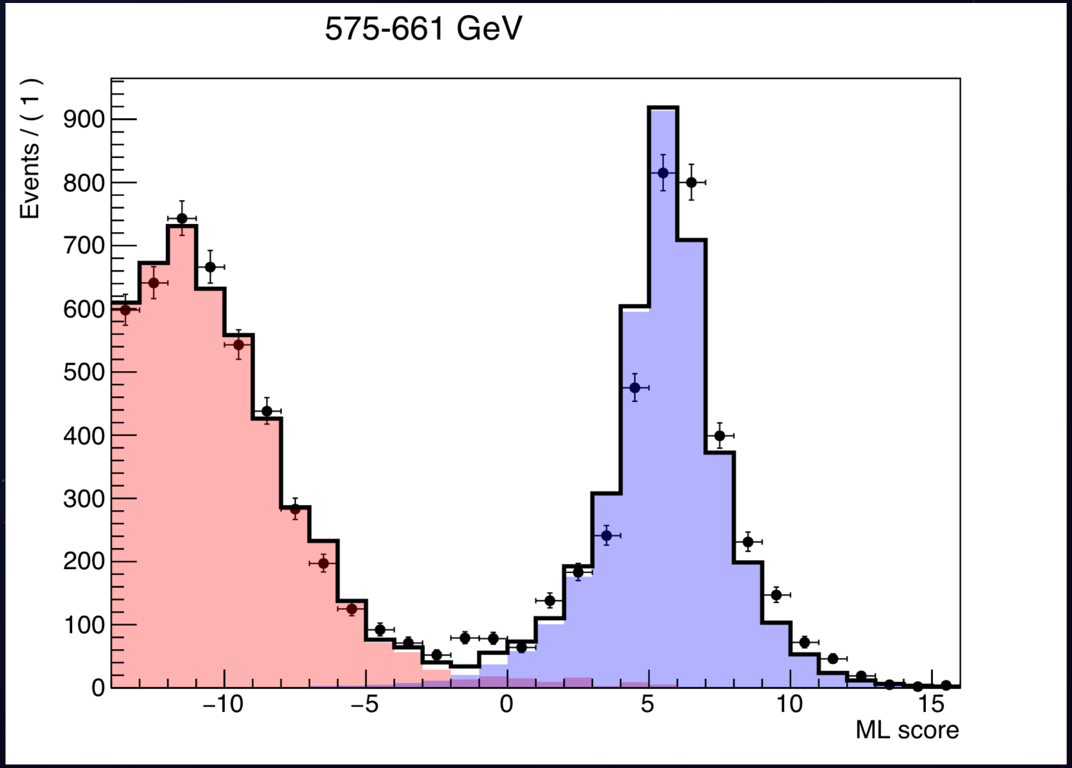
Flux computation: 2 approaches

- Barlow-Beeston:
  - Based on this paper: 10.1016/0010-4655(93)90005-W
  - Fit MC pdf templates to data, accounting for low MC statistics
  - Abstract:
    Analysis of results from HEP experiments often involves estimation of the composition of a sample of data, based on Monte Carlo simulations of the various sources. Data values (generally of more than one dimension) are binned, and because the numbers of data points in many bins are small, a χ^2 minimisation is inappropriate, so a maximum likelihood technique using Poisson statistics is often used. This note shows how to incorporate the fact that the Monte Carlo statistics used are finite and thus subject to statistical fluctuations.
  - Default method in ROOT when fitting pdf templates to data
  - Commonly used in big collaborations like Super-Kamiokande

Flux computation: 2 approaches

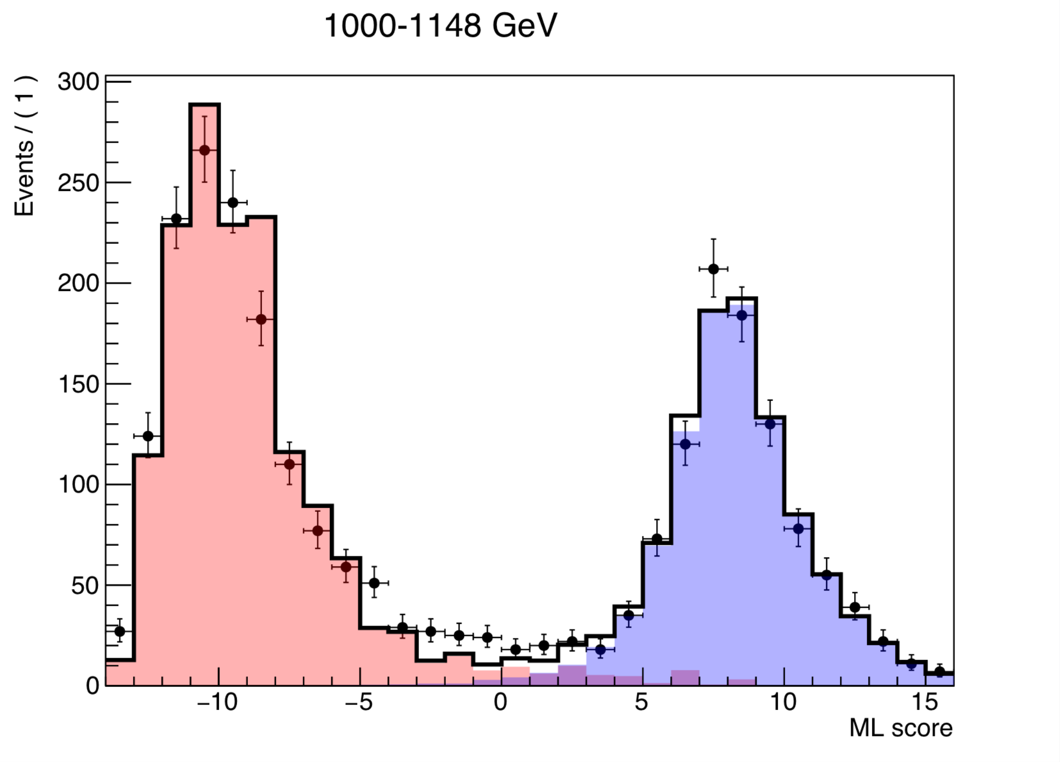
- Barlow-Beeston:

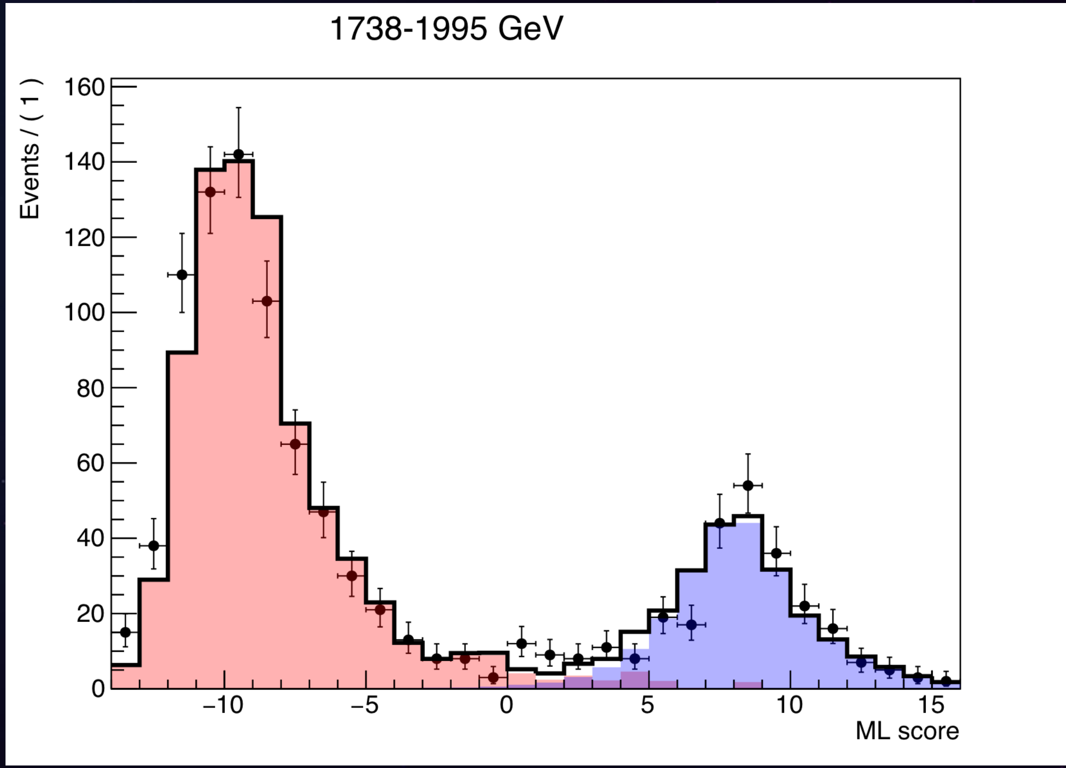




Flux computation: 2 approaches

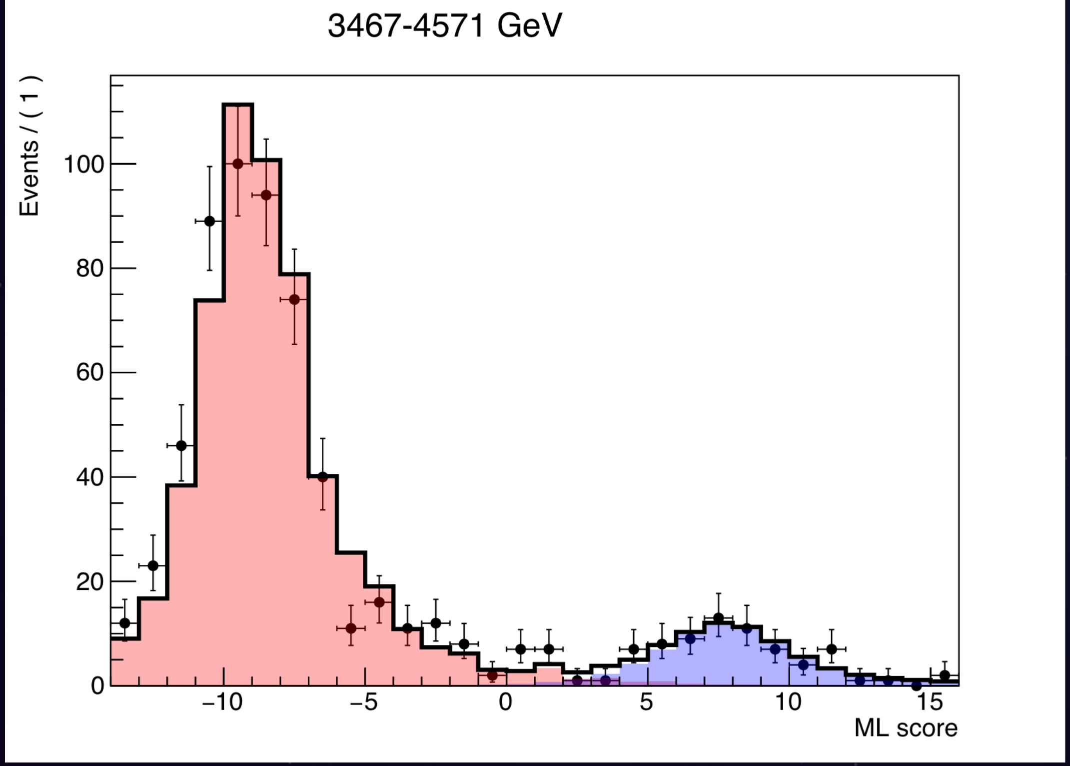
- Barlow-Beeston:

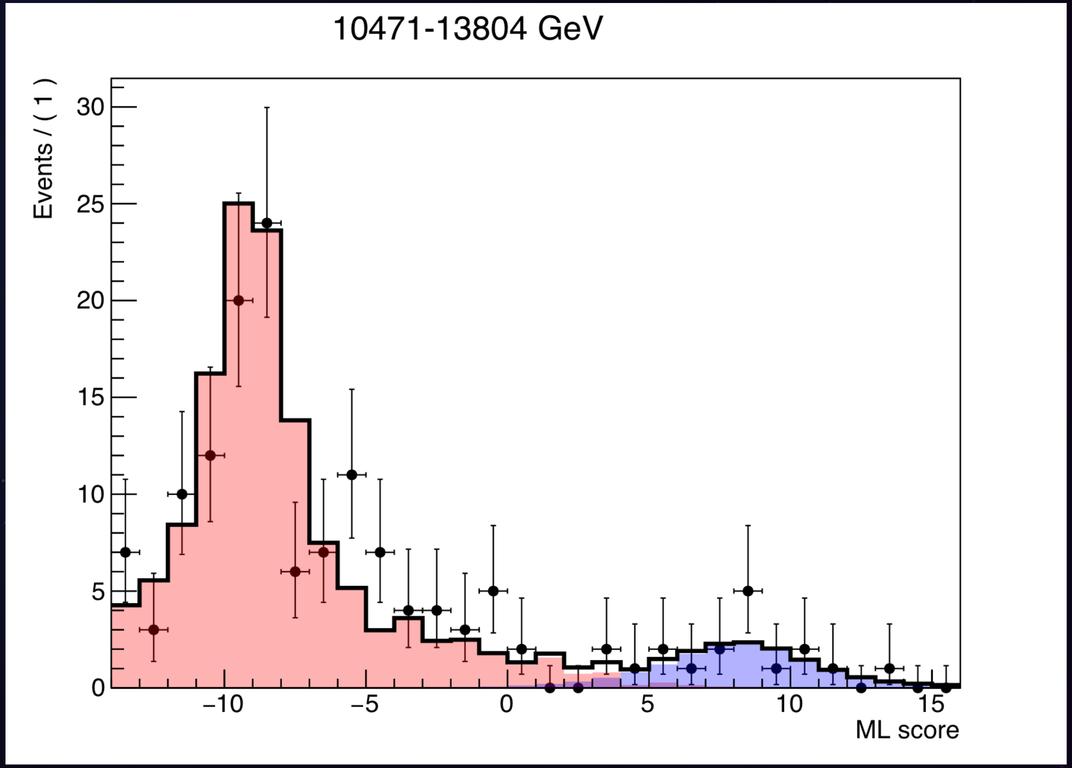




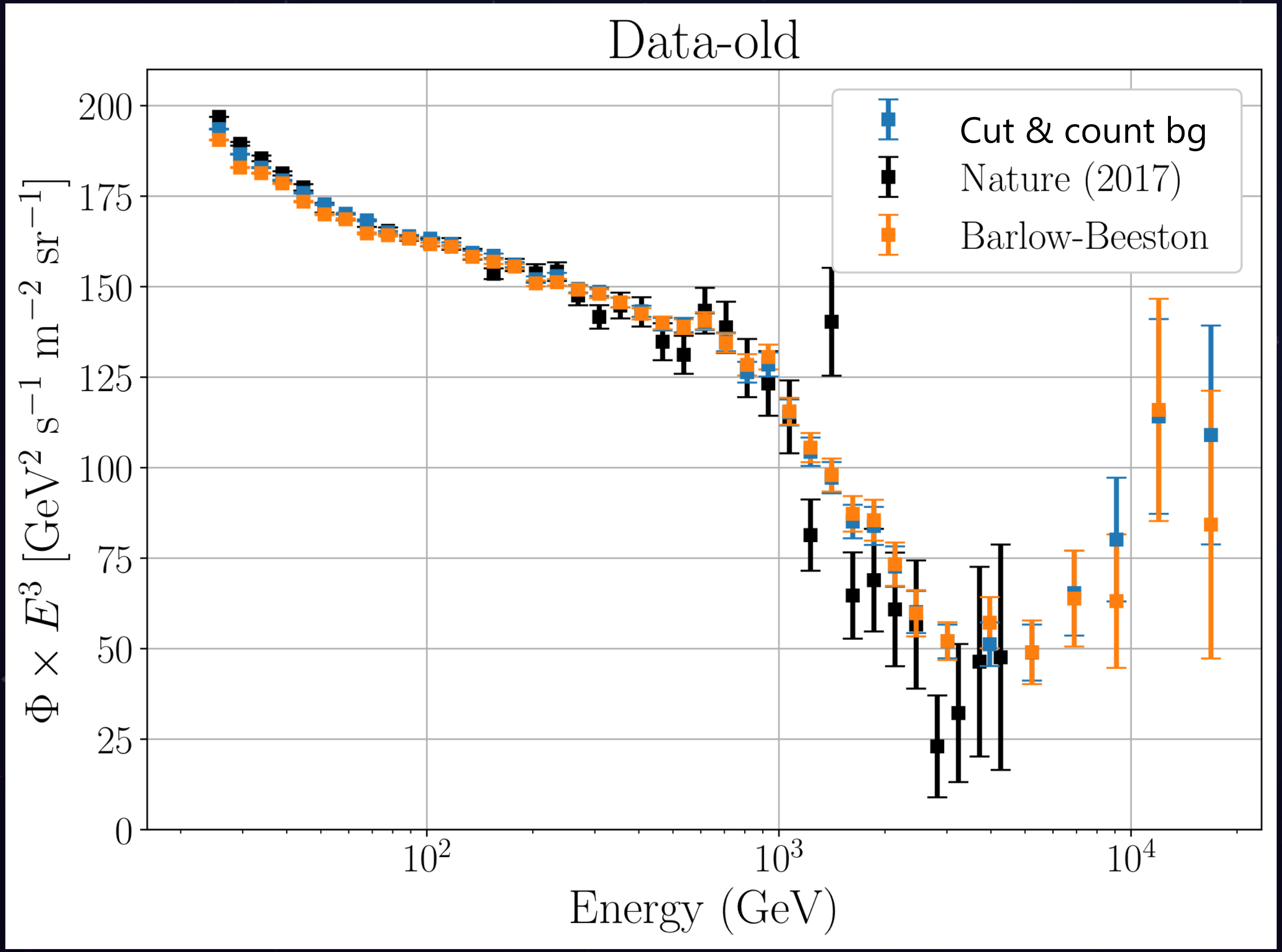
Flux computation: 2 approaches

- Barlow-Beeston:

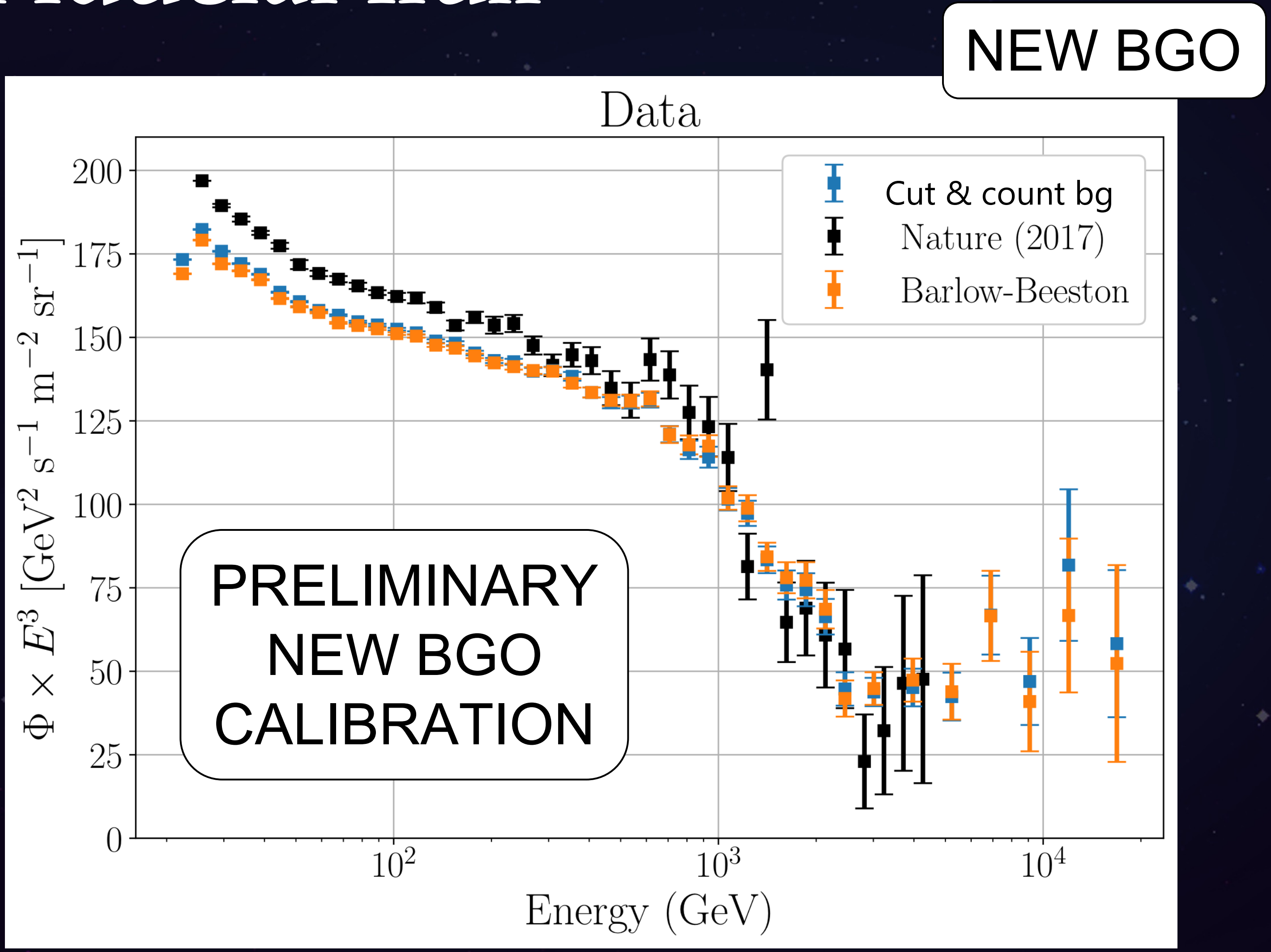




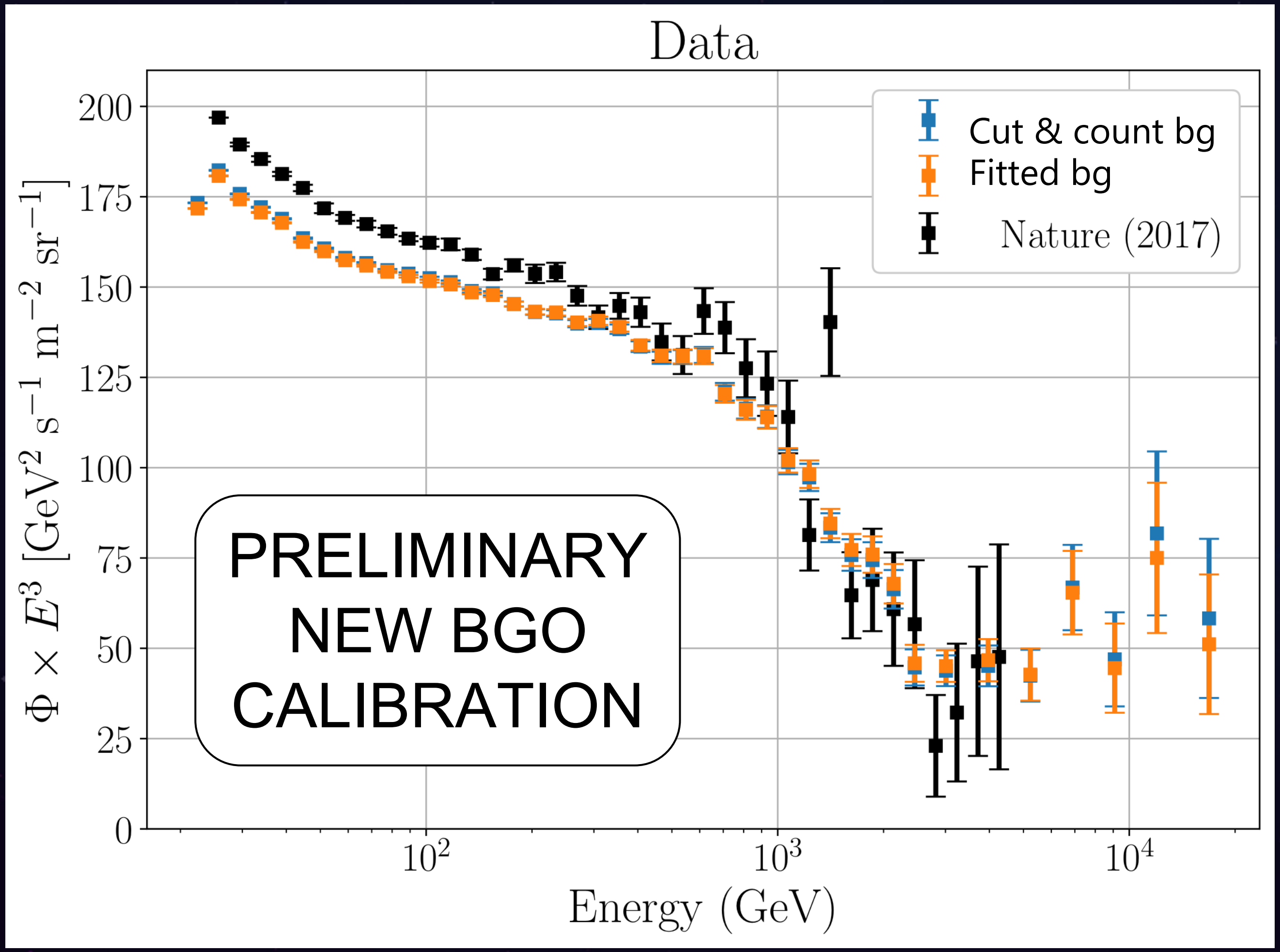
Fiducial flux



Fiducial flux



Fiducial flux

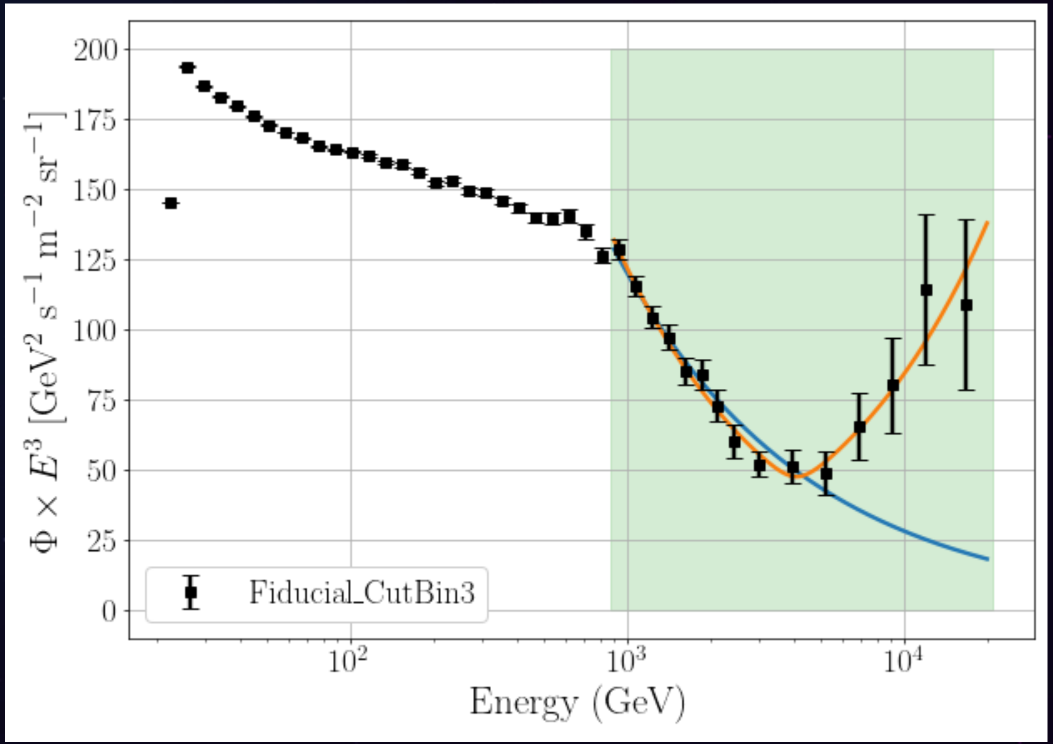


Break significance

- Energy > 0.9 TeV
- Statistical error only
- Two models:
  - Single powerlaw (2 param)
    → Bad fit ($\chi^2 \sim 42.1$)
    Goodness-of-fit: $p = 3.1 \cdot 10^{-5}$
  - Broken powerlaw (4 param)
    → Very good fit ($\chi^2 \sim 4.0$)
    Goodness-of-fit: $p = 0.947$
- Hypothesis test:
  $p = 5.3 \cdot 10^{-9}$ or 5.7 sigma

Best-fit parameters: $\Phi = \Phi_0 \cdot \left(\frac{E}{TeV}\right)^{-\gamma} \cdot \left[1 + \left(\frac{E}{E_{br}}\right)^{\frac{\Delta\gamma}{s}}\right]^{s}$

- $\Phi_0 = (1.22 \pm 0.01) \cdot 10^{-7}$
- $\gamma = 3.72 \pm 0.03$
- $(4.1 \pm 0.3) \cdot 10^3\ GeV$
- $\Delta\gamma = (1.43 \pm 0.12)$
- $s = 0.1$ (fixed)



Break significance

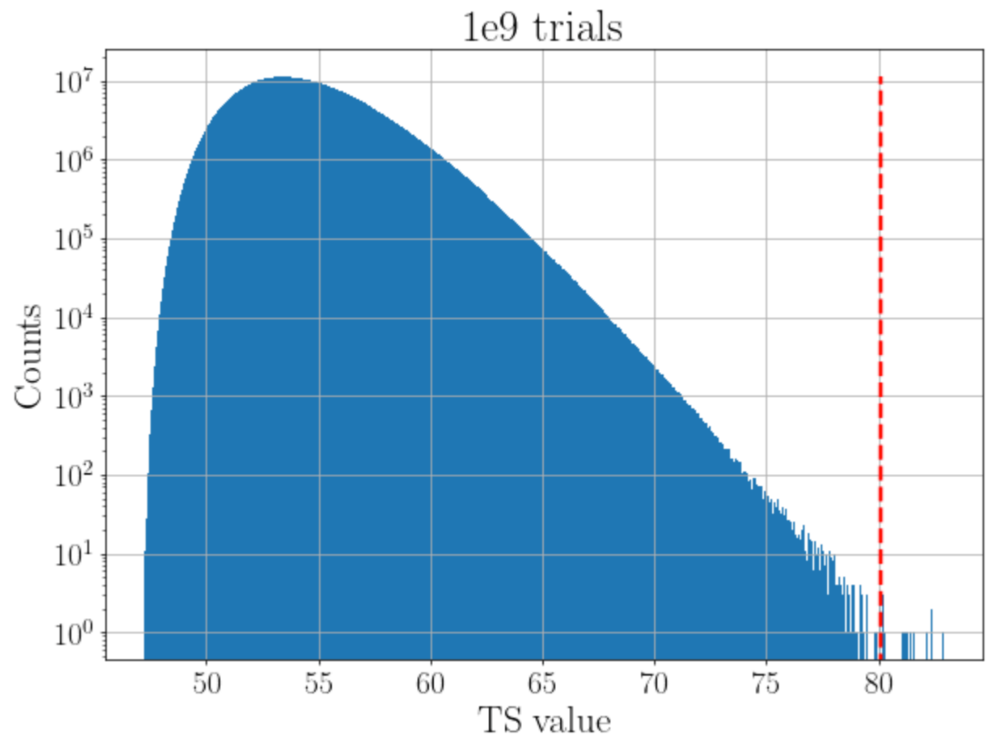
- Energy > 0.9 TeV
- Statistical + systematic error
- Hypothesis test:
  $p = 5.3 \cdot 10^{-9}$ or 5.7 sigma
  (no sys)

  $p = 1.6 \cdot 10^{-7}$ or 5.1 sigma
  (5% sys, 2 nuisance param)

  $p = 1.7 \cdot 10^{-5}$ or 4.1 sigma
  (10% sys, 2 nuisance param)

Break significance

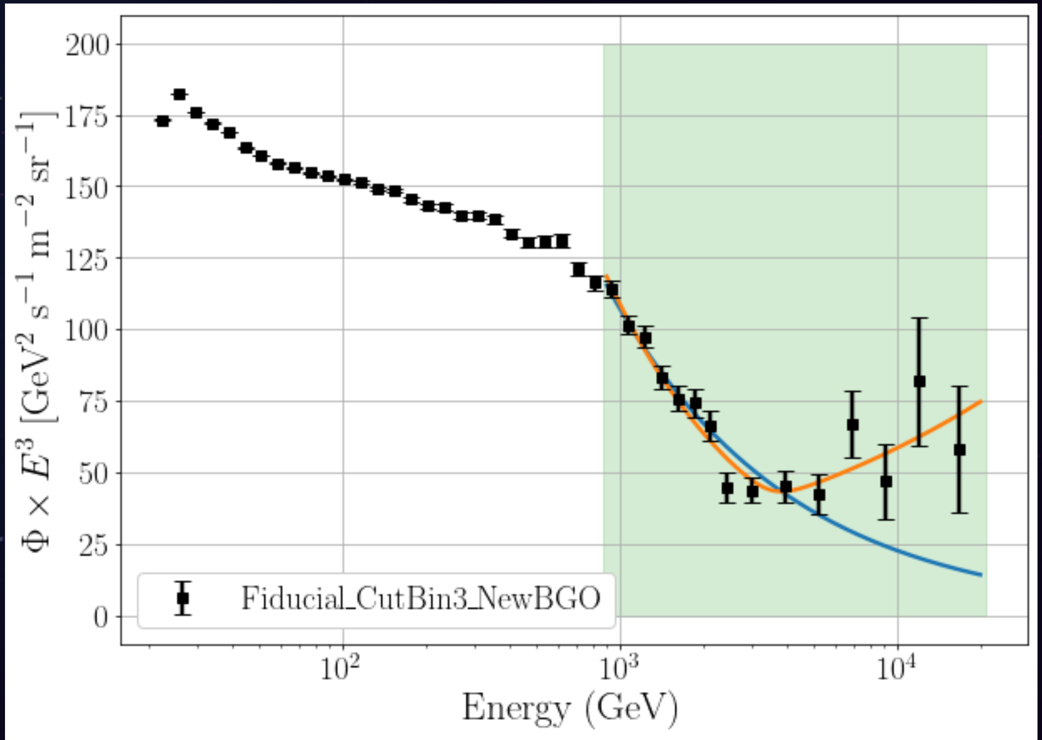
- Alternative approach:
  - Use fit to calculate number of expected events $\mu = \phi \cdot Acceptance$
  - Calculate Poisson likelihood of observed counts → TS
  - Pseudo-trials → TS-distribution
  - Result:
    - p-value: 1.4e-08
    - Significance: 5.5 sigma



Break significance (for new BGO)

- Energy > 0.9 TeV
- Statistical error only
- Two models:
  - Single powerlaw (2 param)
    → Bad fit ($\chi^2 \sim 38.7$)
    Goodness-of-fit: $p = 1.2 \cdot 10^{-4}$
  - Broken powerlaw (4 param)
    → Very good fit ($\chi^2 \sim 13.1$)
    Goodness-of-fit: $p = 0.22$
- Hypothesis test:
  $p = 2.8 \cdot 10^{-6}$ or 4.5 sigma

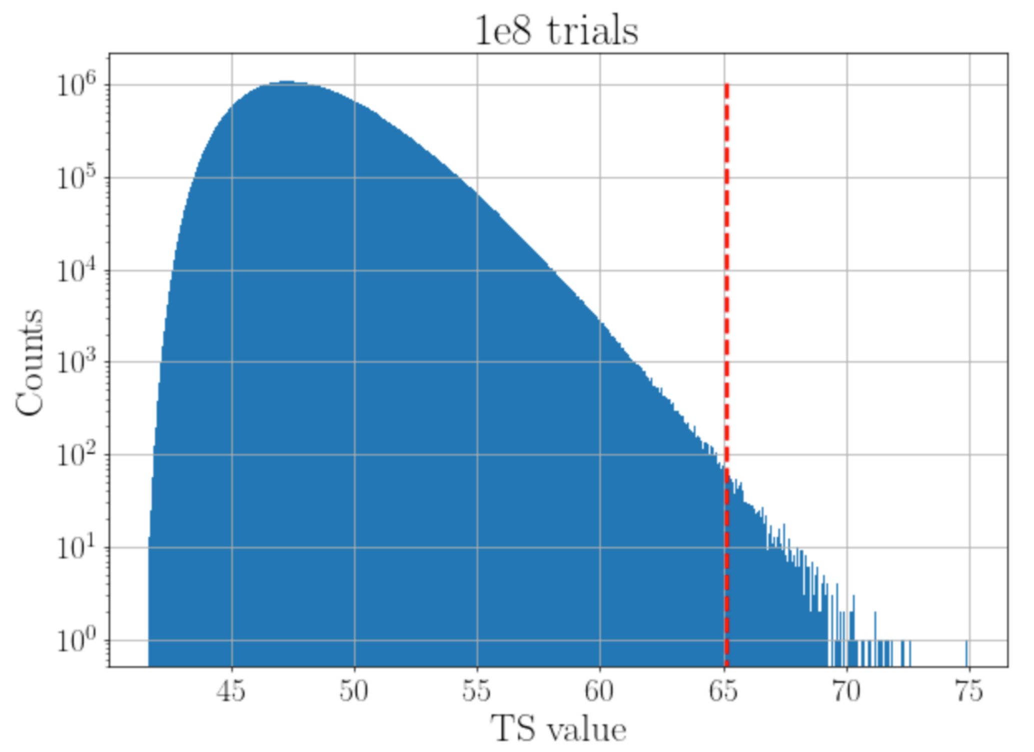
NEW BGO



Break significance (for new BGO)

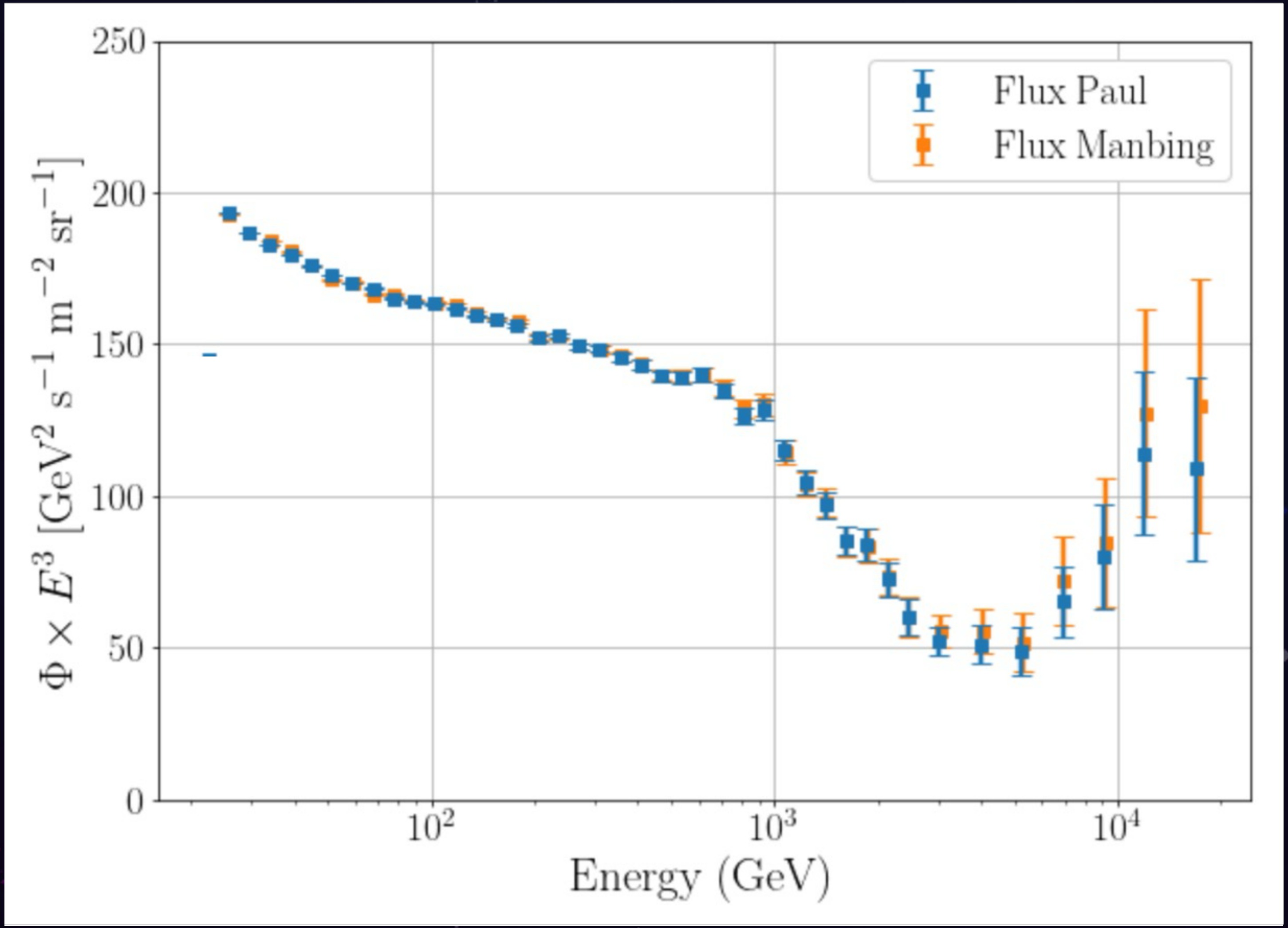
- Alternative approach:
  - Use fit to calculate number of expected events
    $\mu = \phi \cdot Acceptance$
  - Calculate Poisson likelihood of observed counts → TS
  - Pseudo-trials → TS-distribution
  - Result:
    - p-value: 1.2e-05
    - Significance: 4.2 sigma

NEW BGO



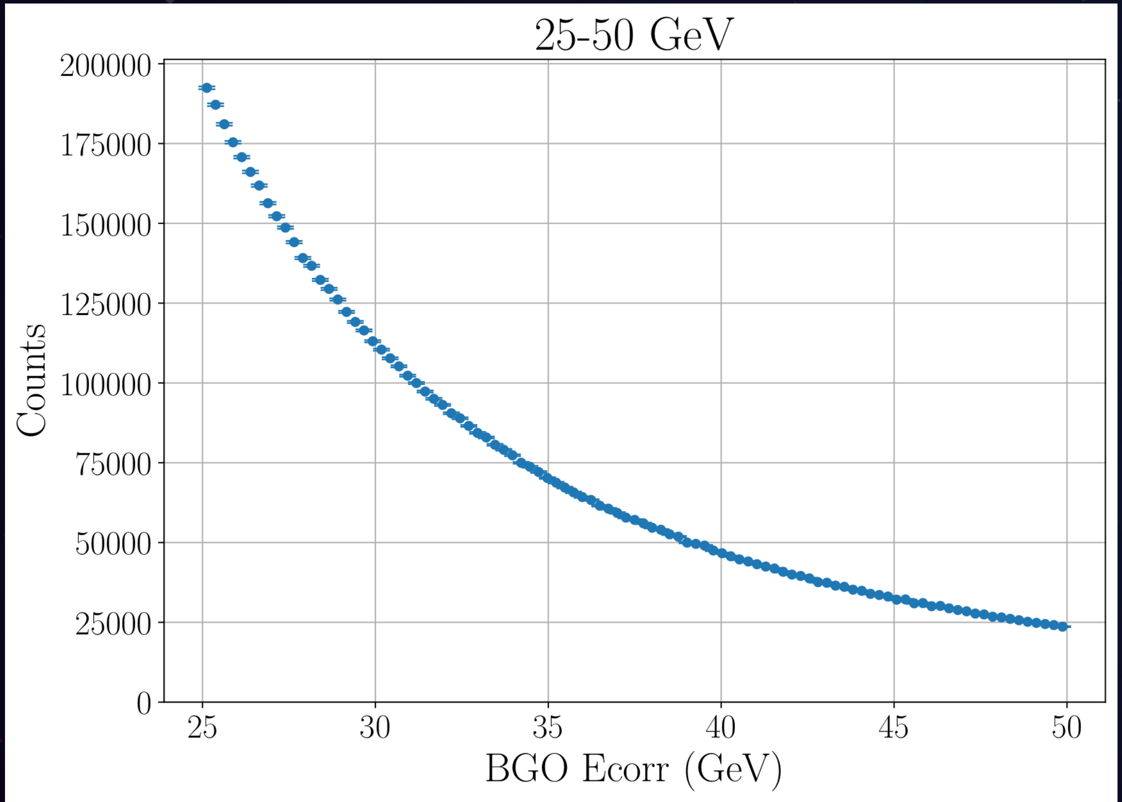
Extra plots

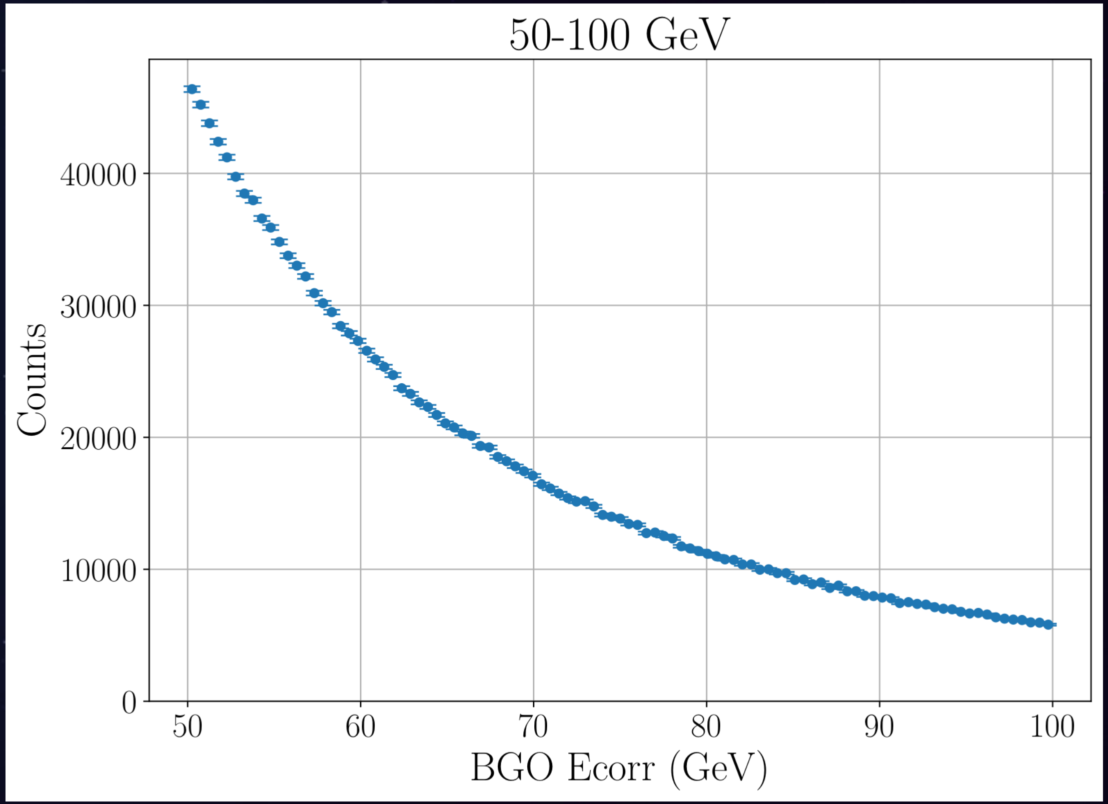
Fiducial flux



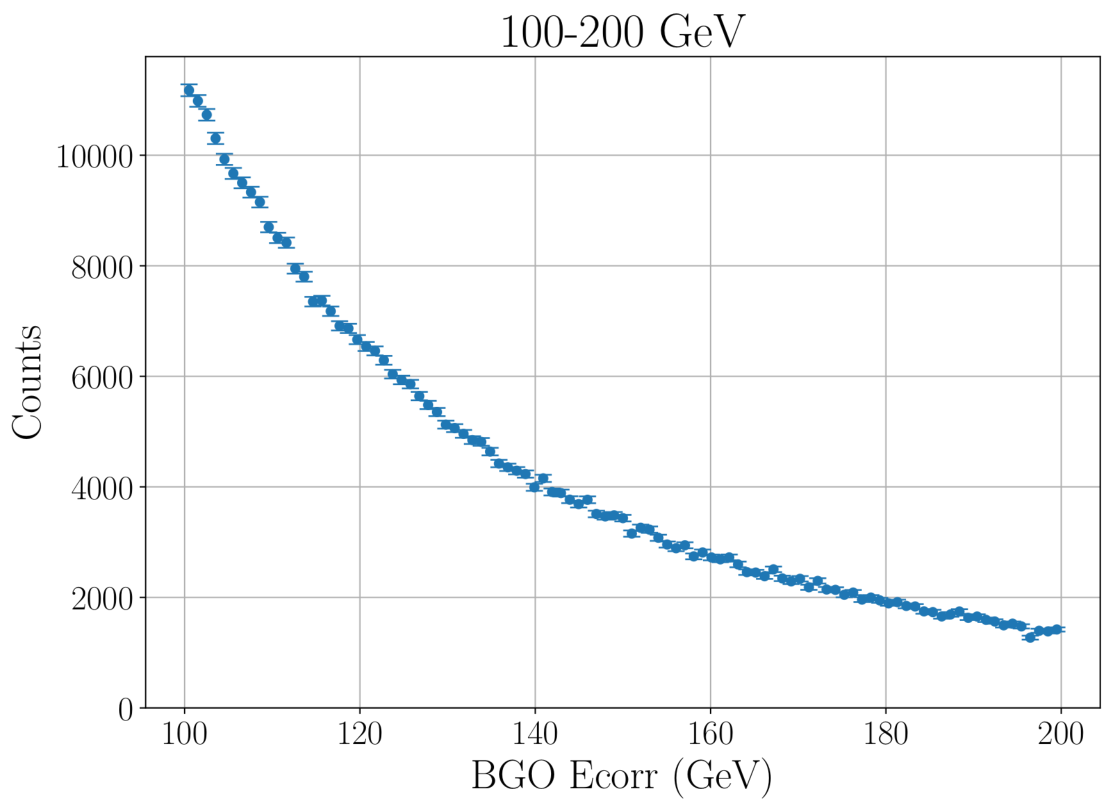
Fully independent analysis code lead to very consistent result. Minor difference at highest energy is due to different cut on ML-score, leading to different efficiencies and background rates.

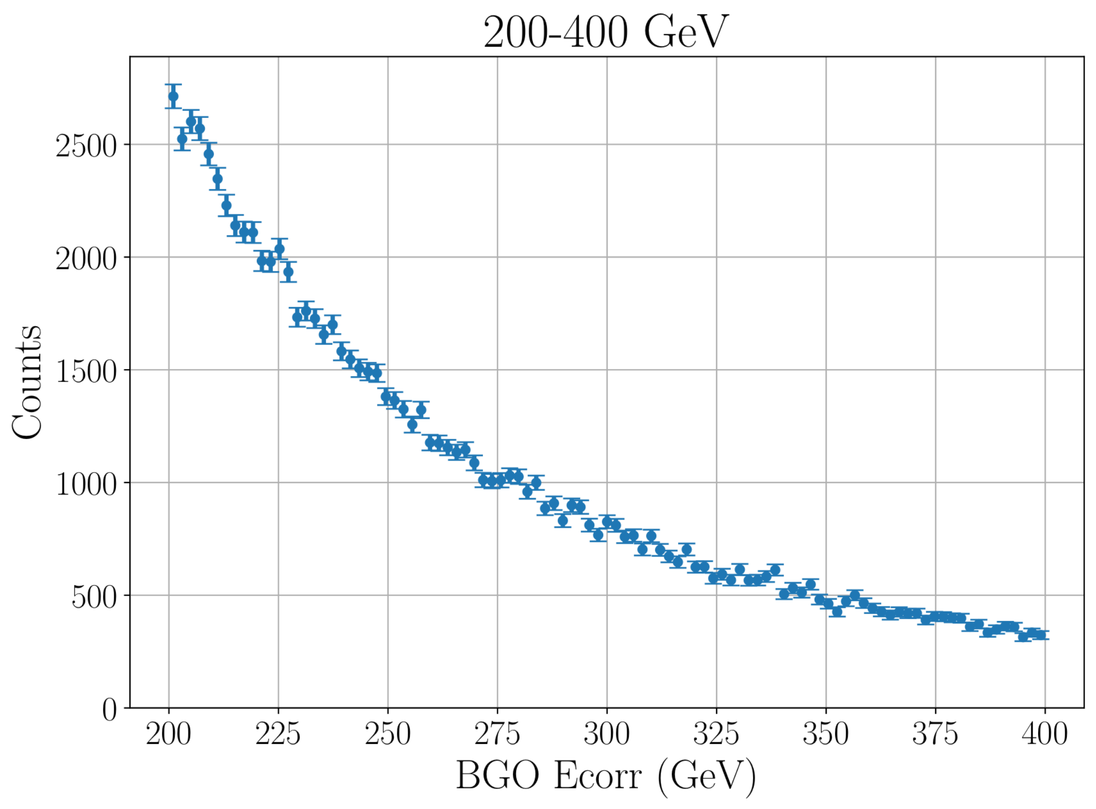
Raw counts for Ω>0 with fine binning



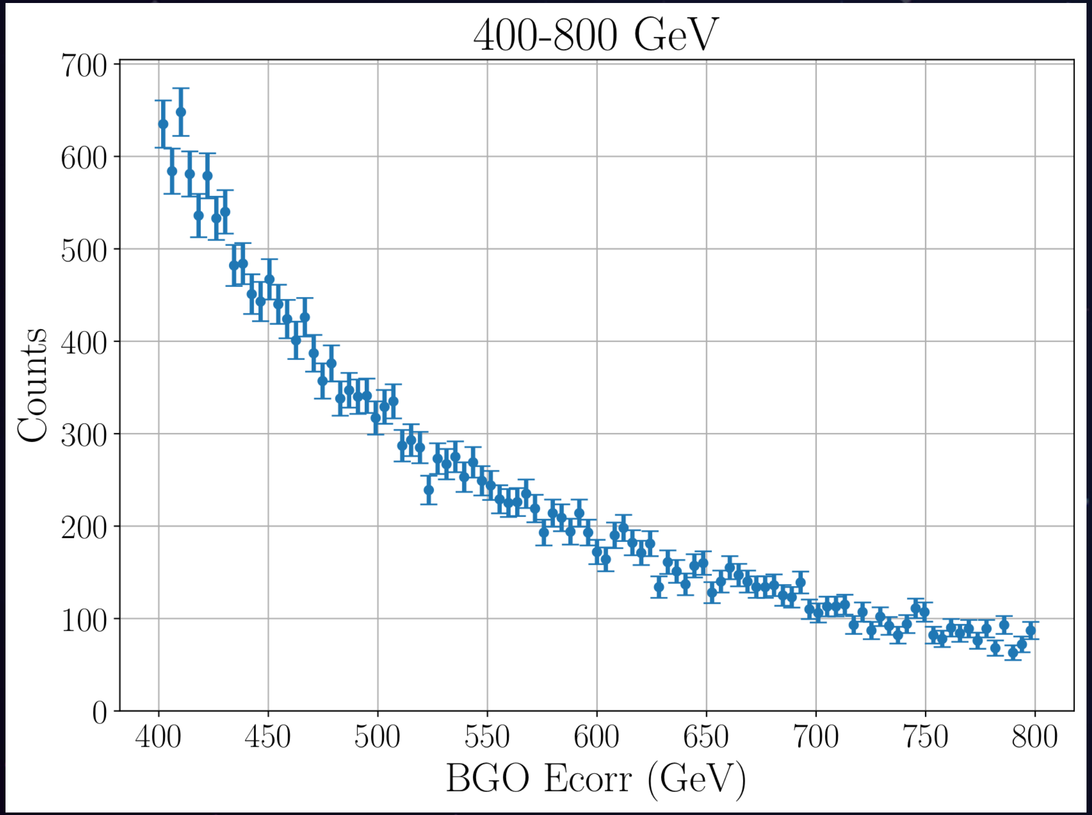


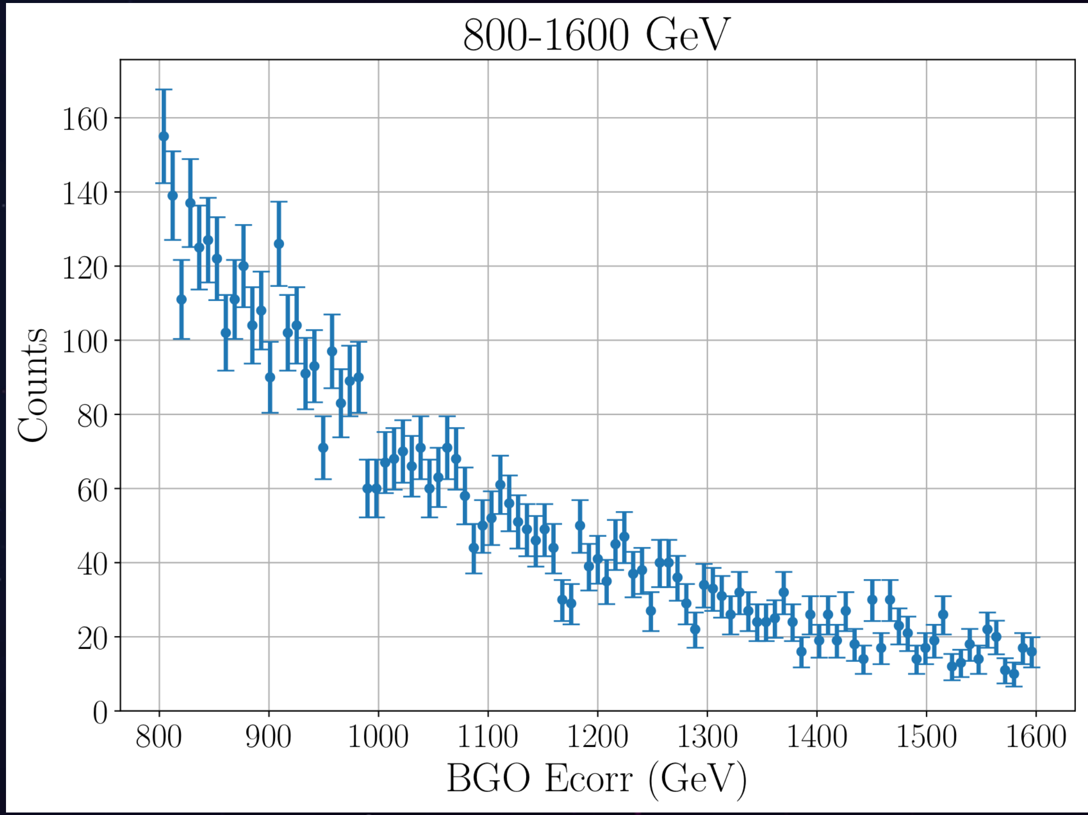
Raw counts for Ω>0 with fine binning



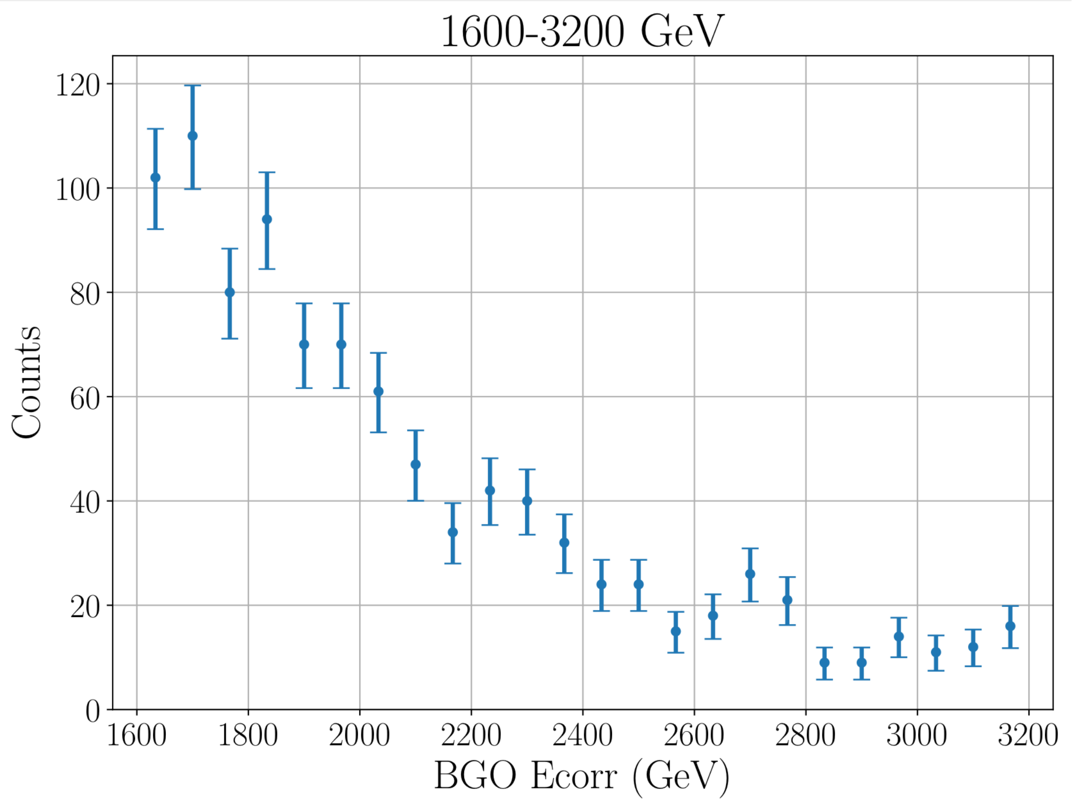


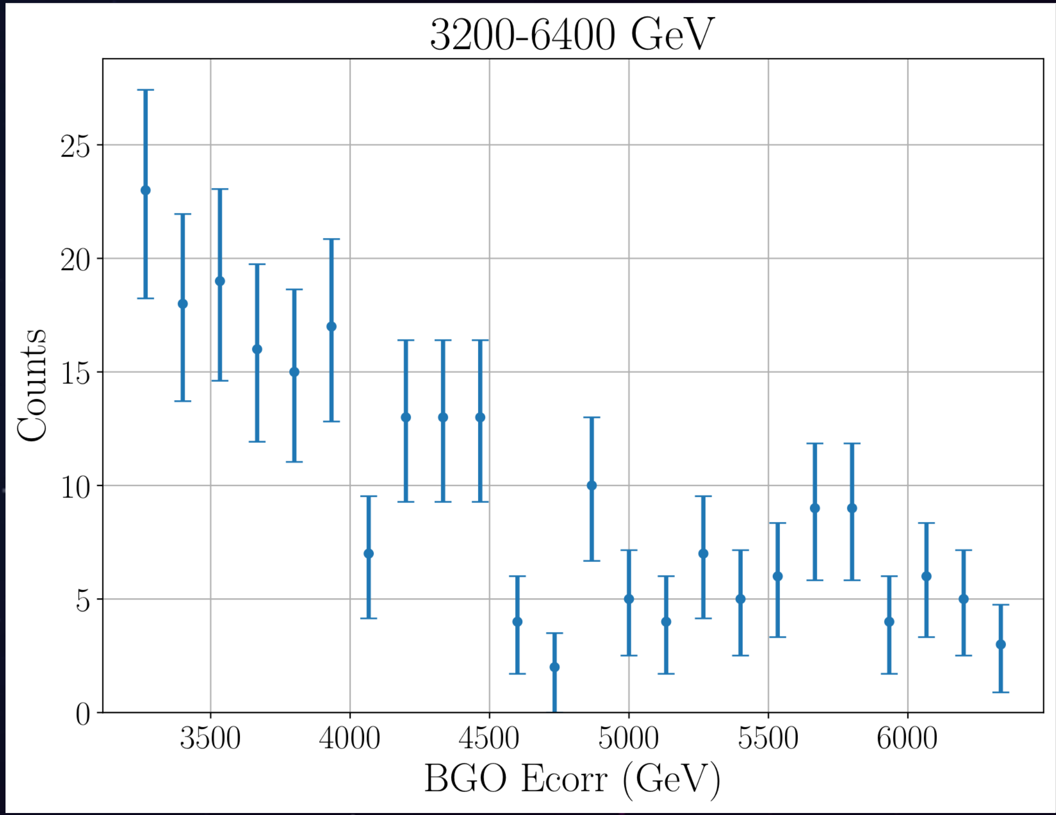
Raw counts for $\Omega>0$ with fine binning



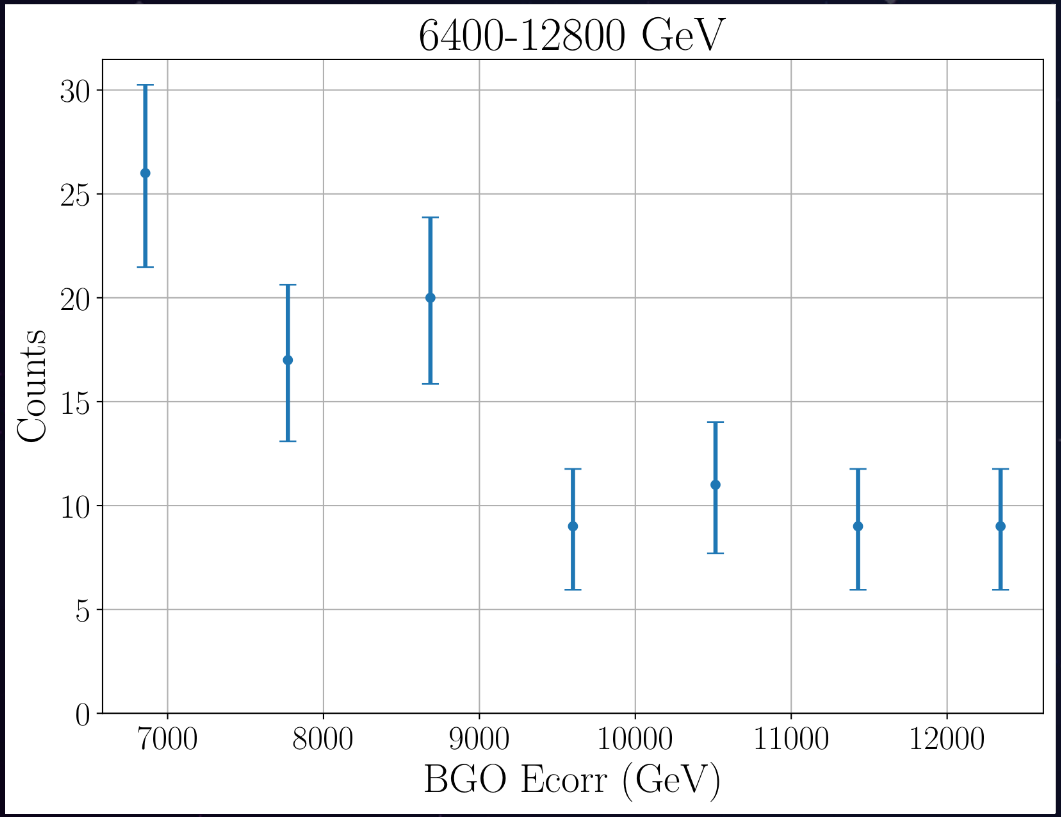


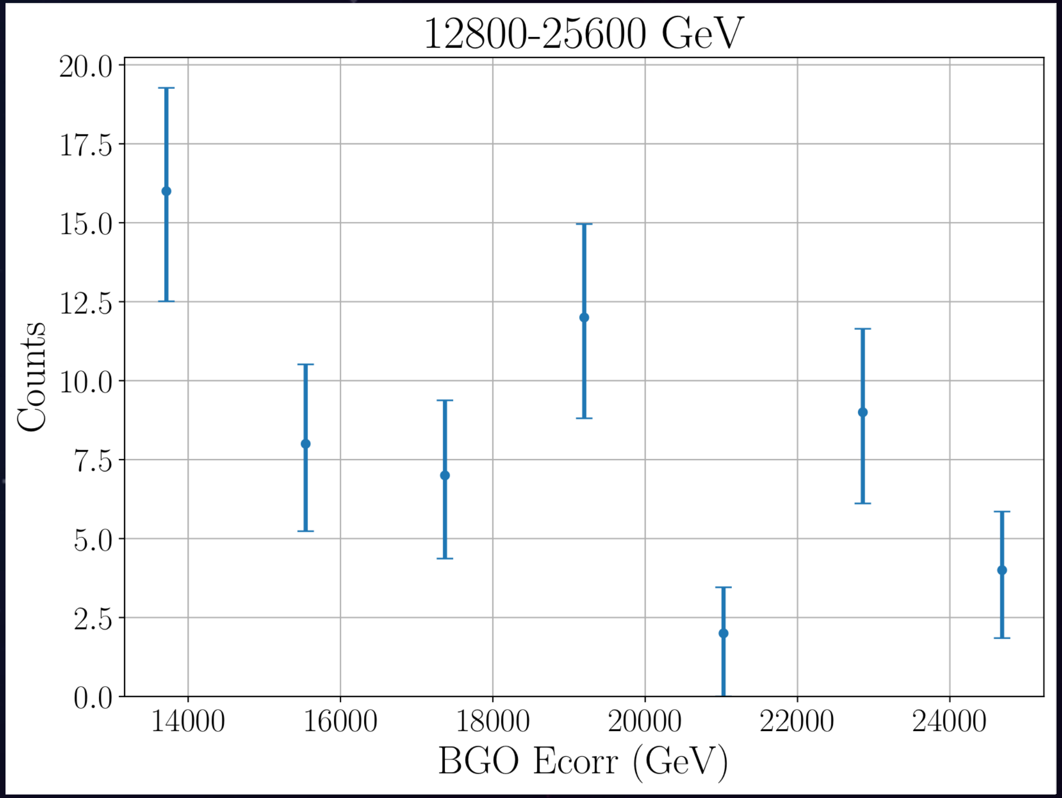
Raw counts for $\Omega > 0$ with fine binning





Raw counts for Ω>0 with fine binning





Electron flux

Analysis of top-fiducial events

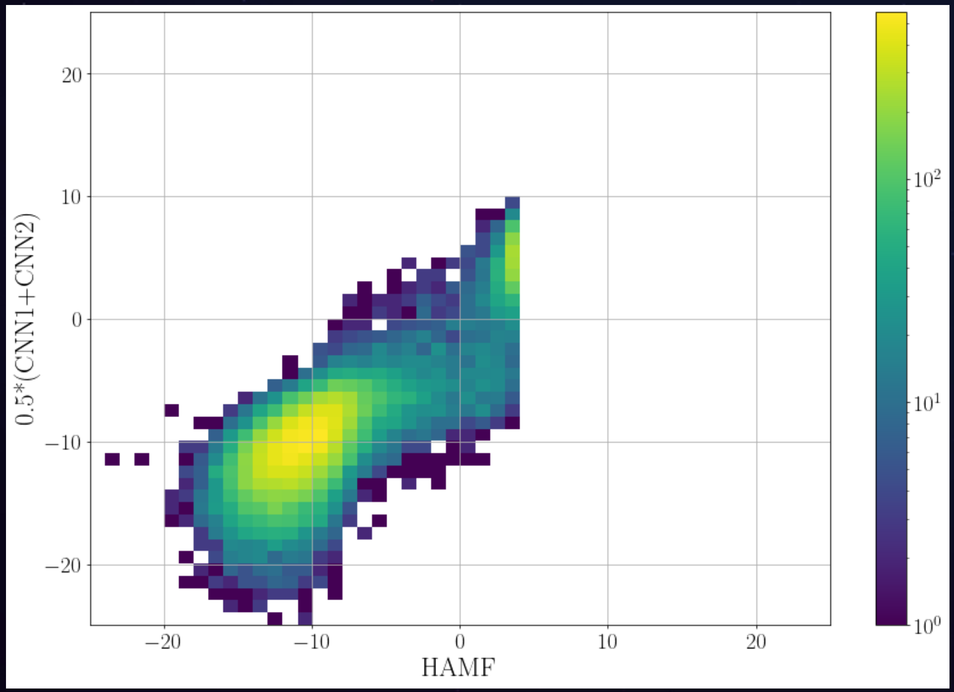
UNIVERSITÉ DE GENÈVE

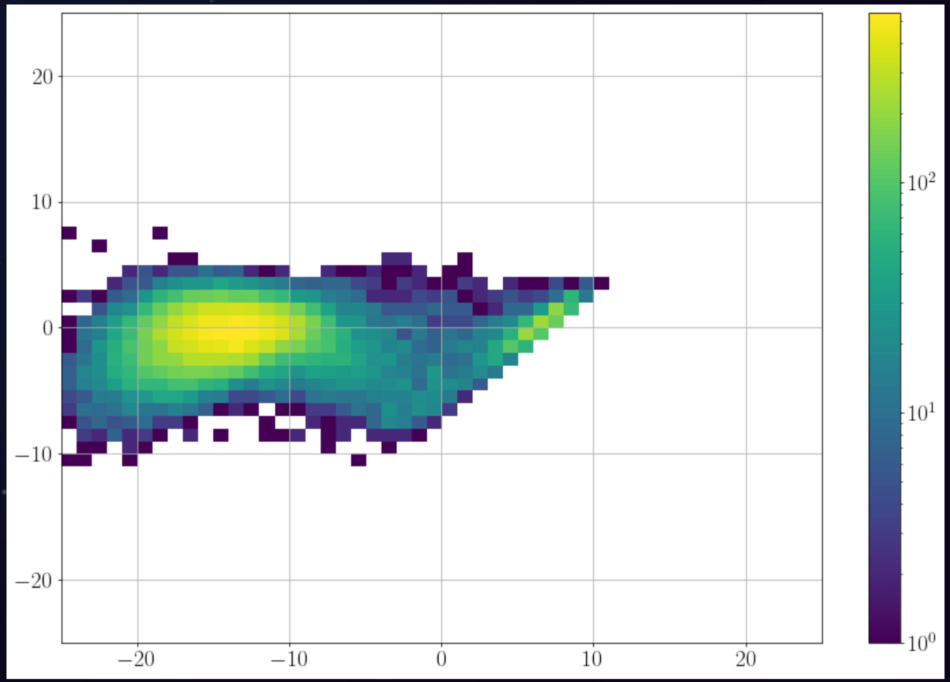
Swiss National Science Foundation

Paul Coppin

TopFiducial analysis

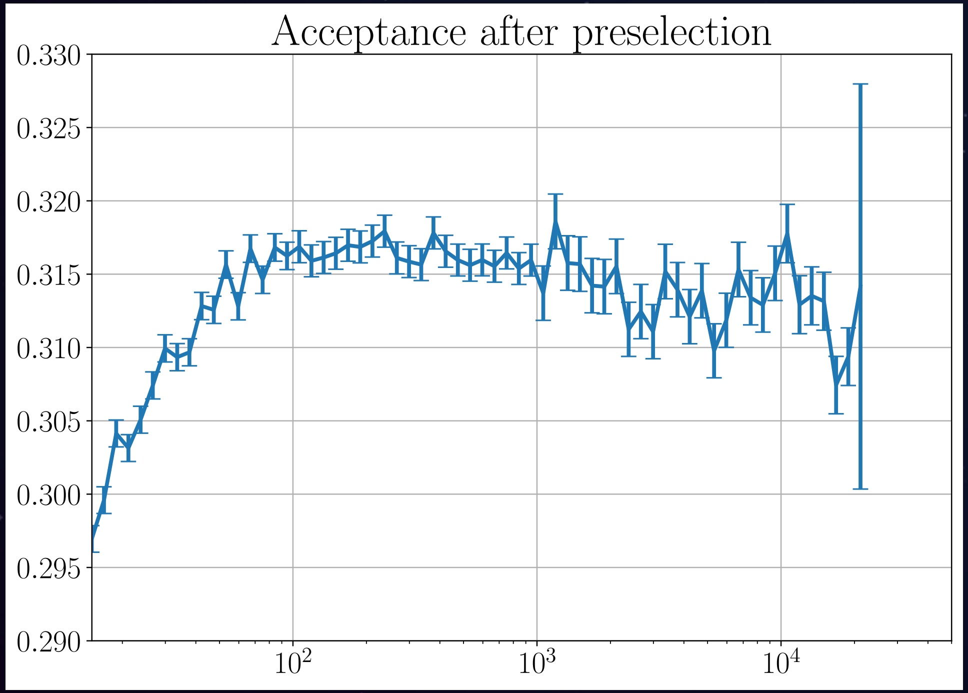
- Similar procedure to fiducial analysis:
  Ω obtained by rotating HAMF and (mean) CNN models

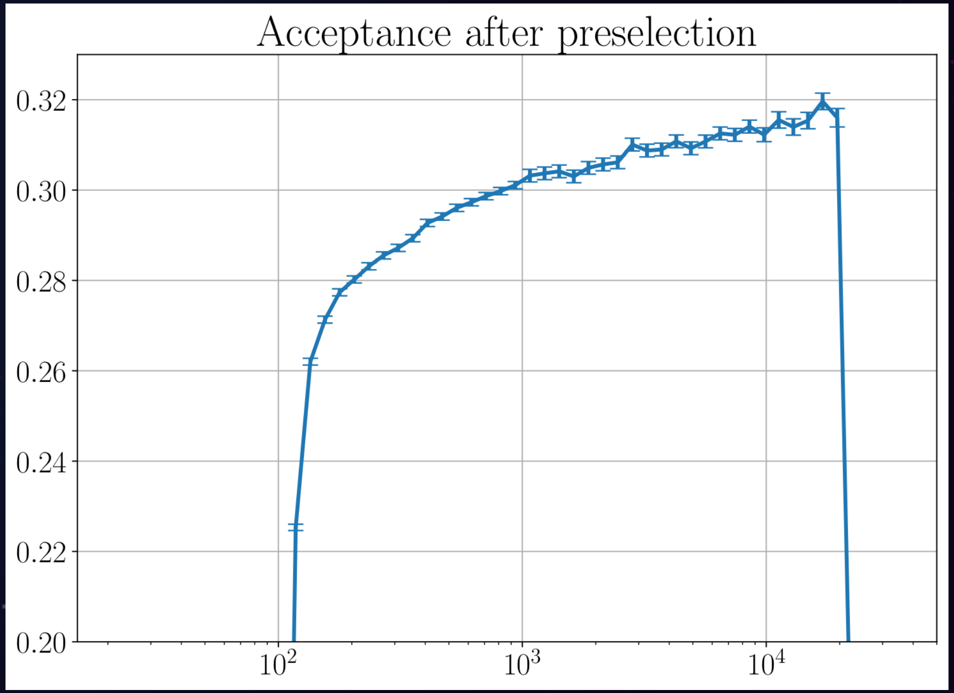




Same trick as before, rotate over 30 degrees.

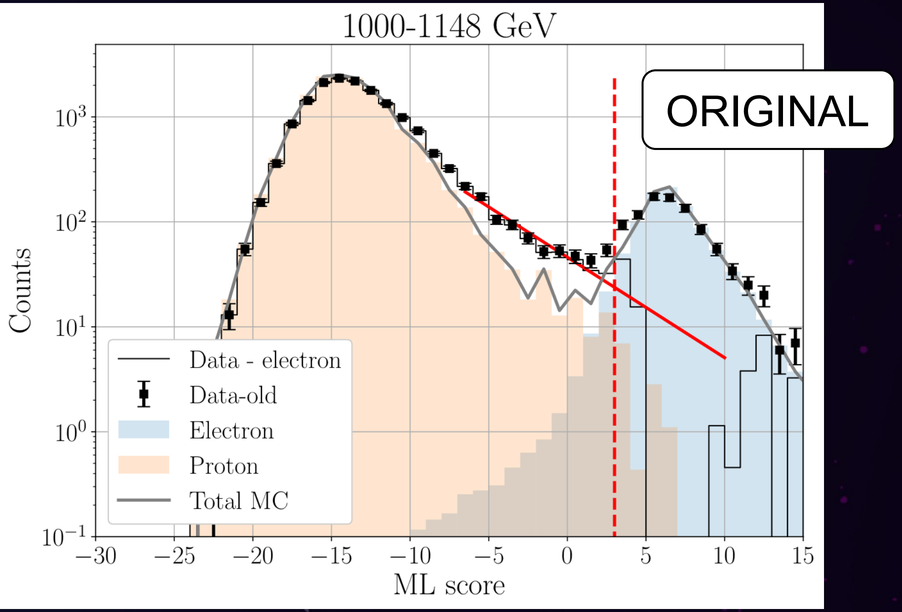
Acceptance after preselection

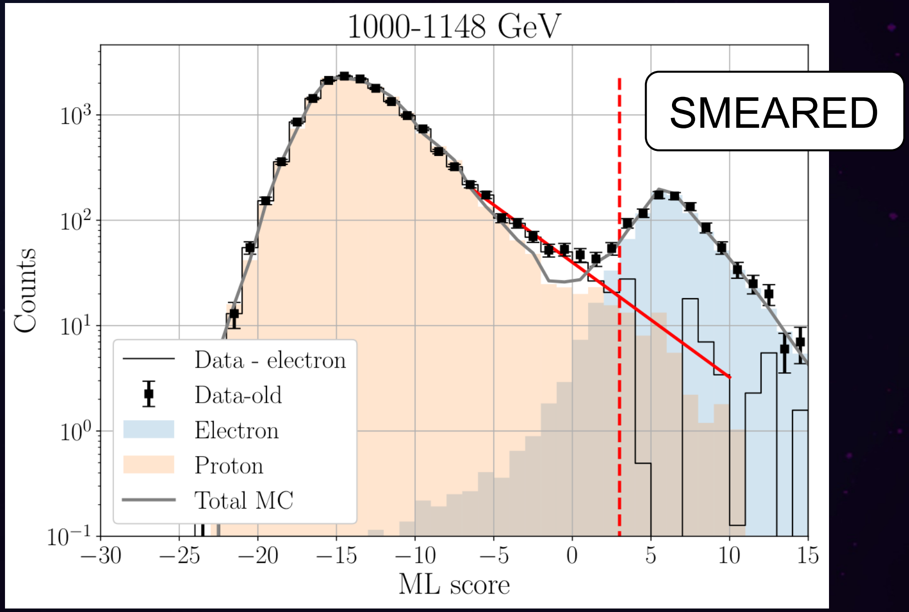




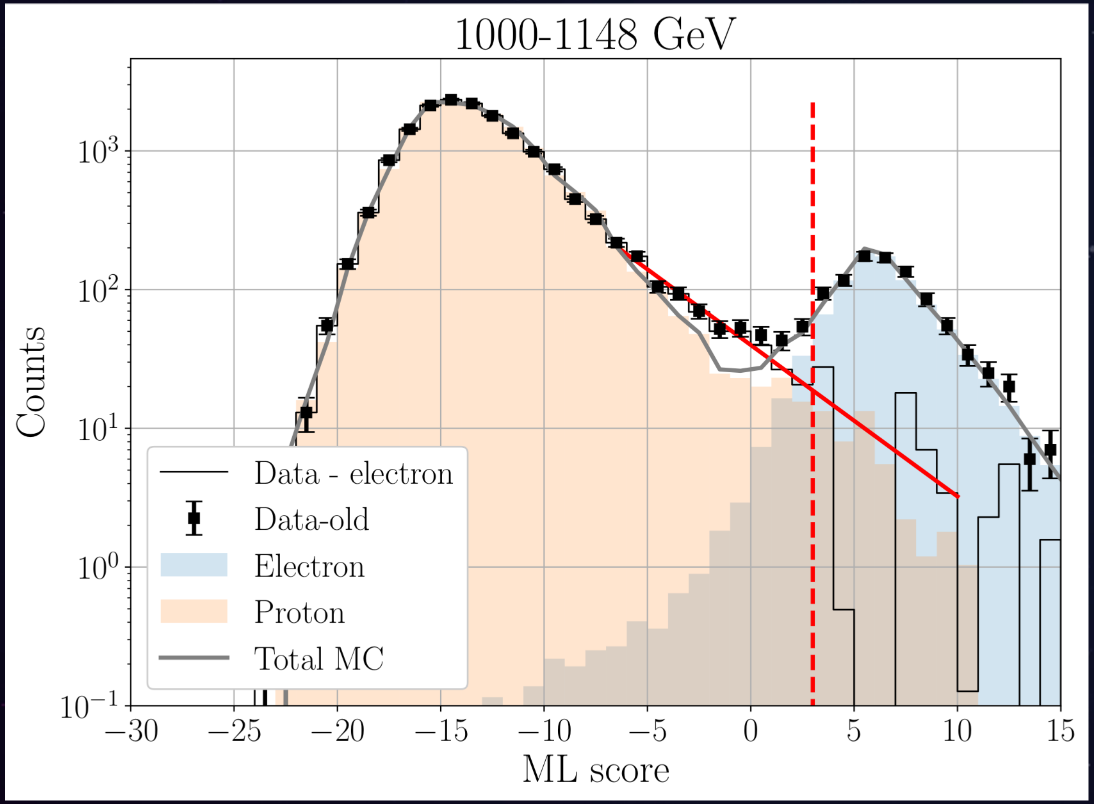
Data-MC agreement

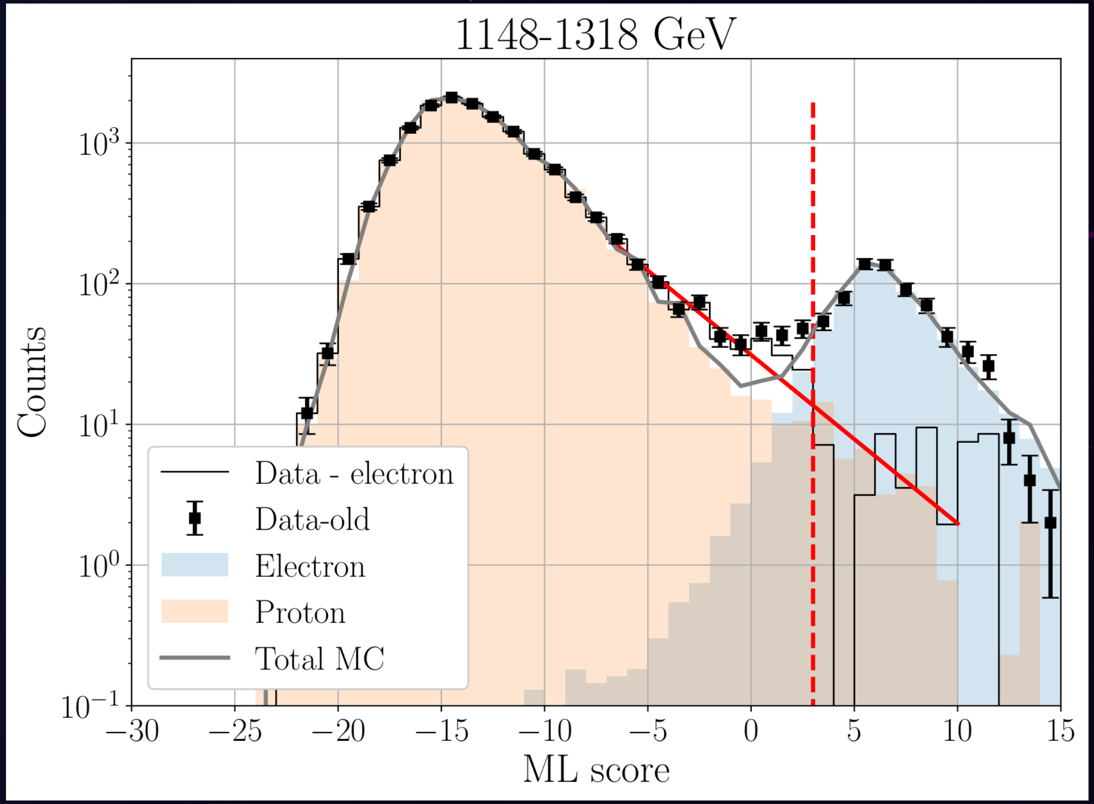
- Smear MC to data
  - Shift electrons left by 0.25
  - Shift proton right by 0.5
  - Increase width by 16% (electron & proton)
- Background estimation:
  - Count: based on proton MC
  - Fit data minus electron histogram → integrate fit (simple linear model, to be refined)



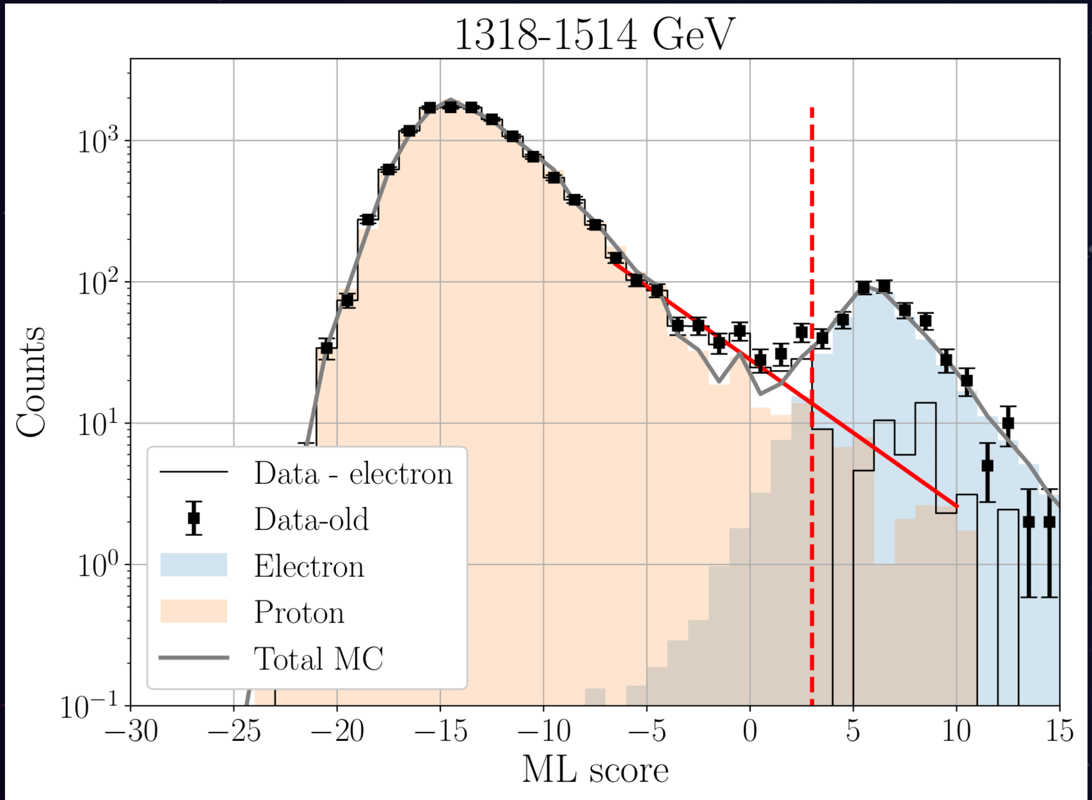


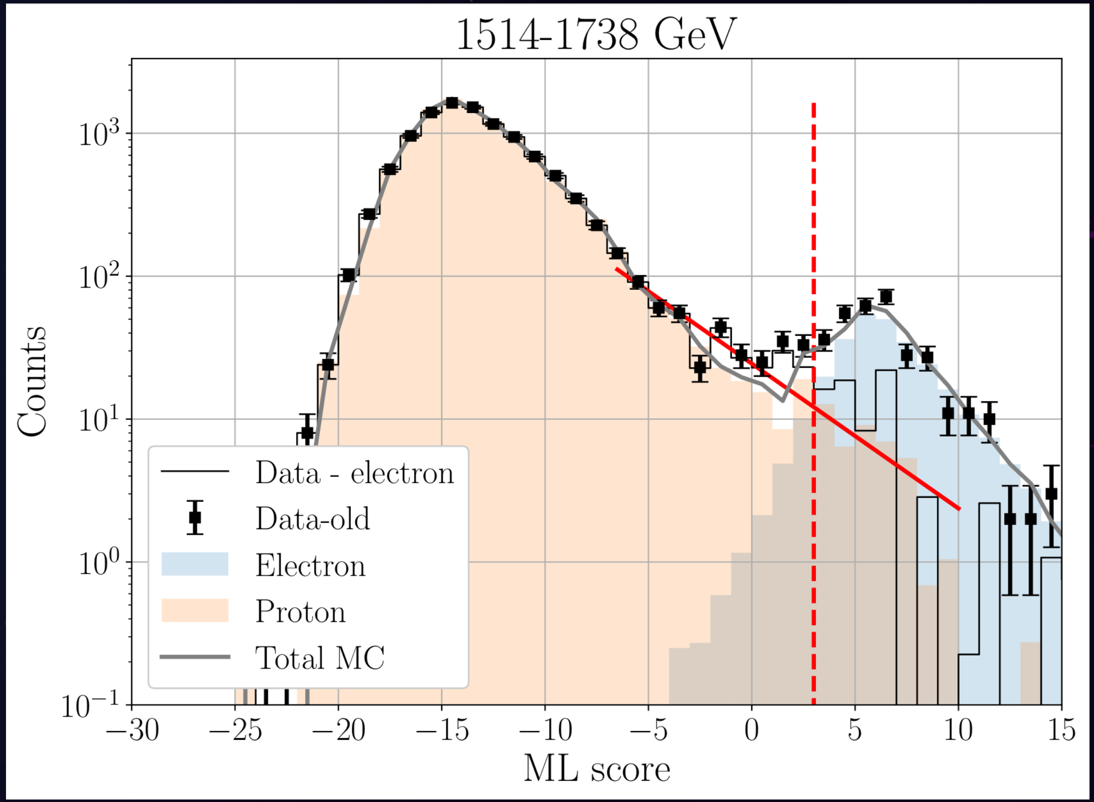
Score distributions



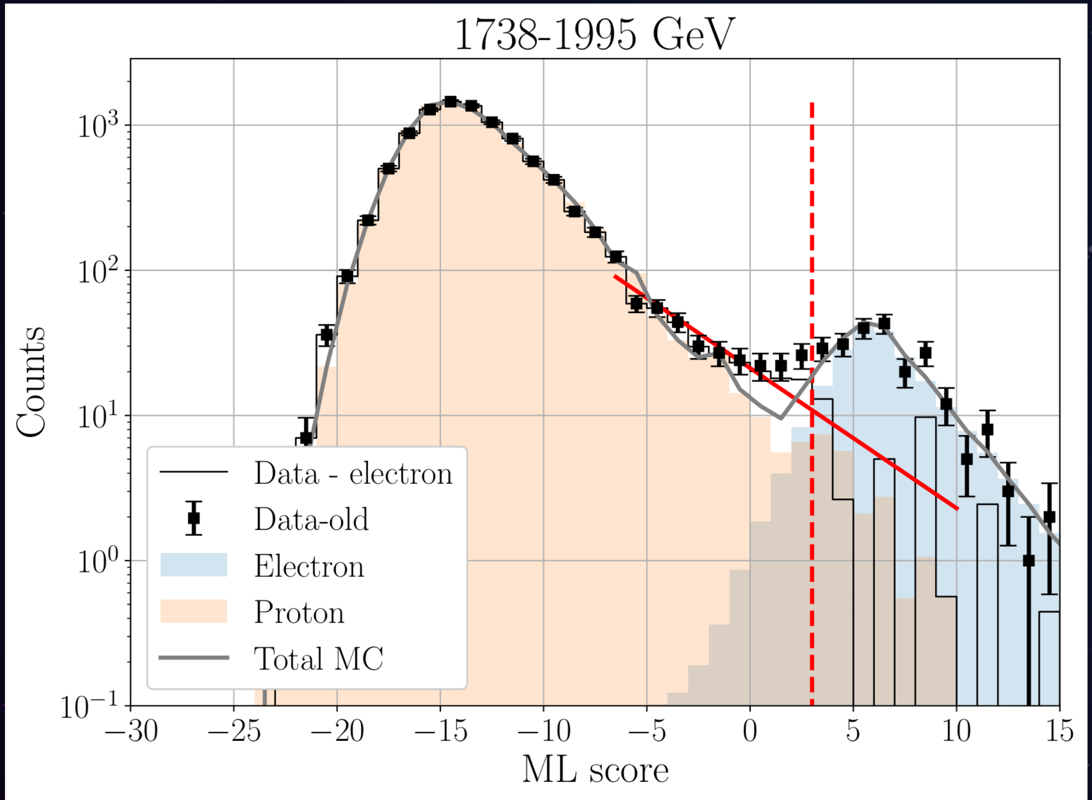


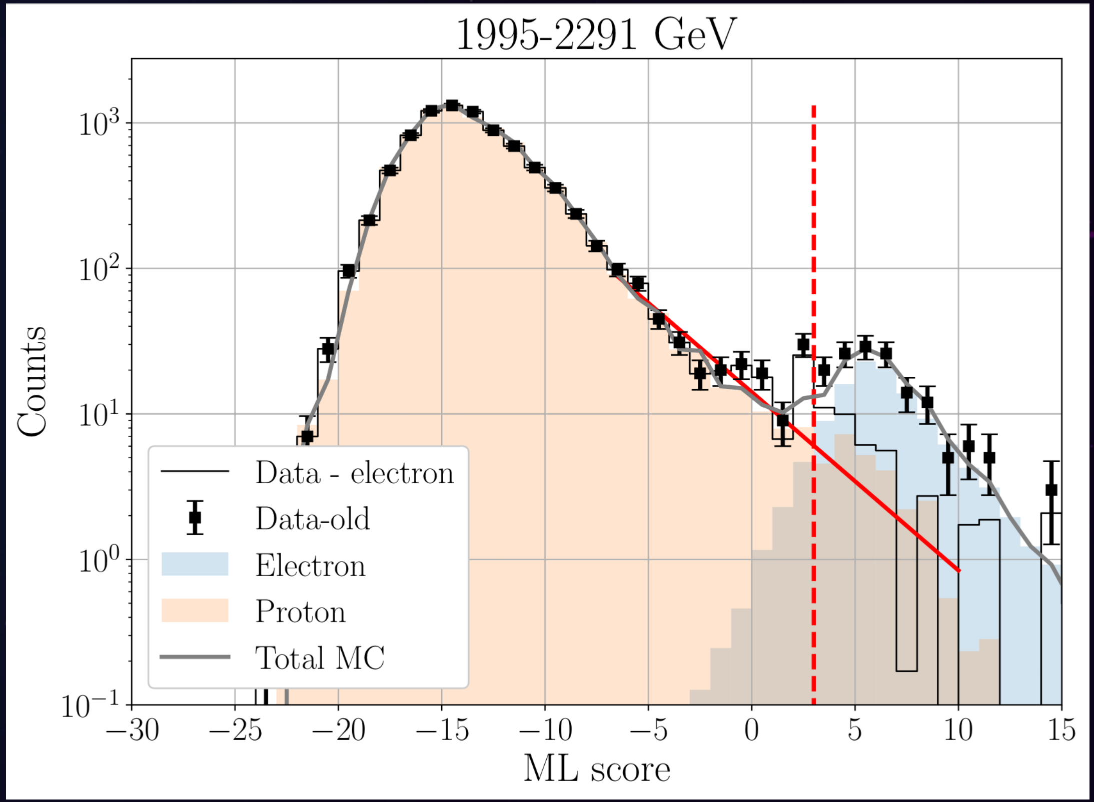
Score distributions



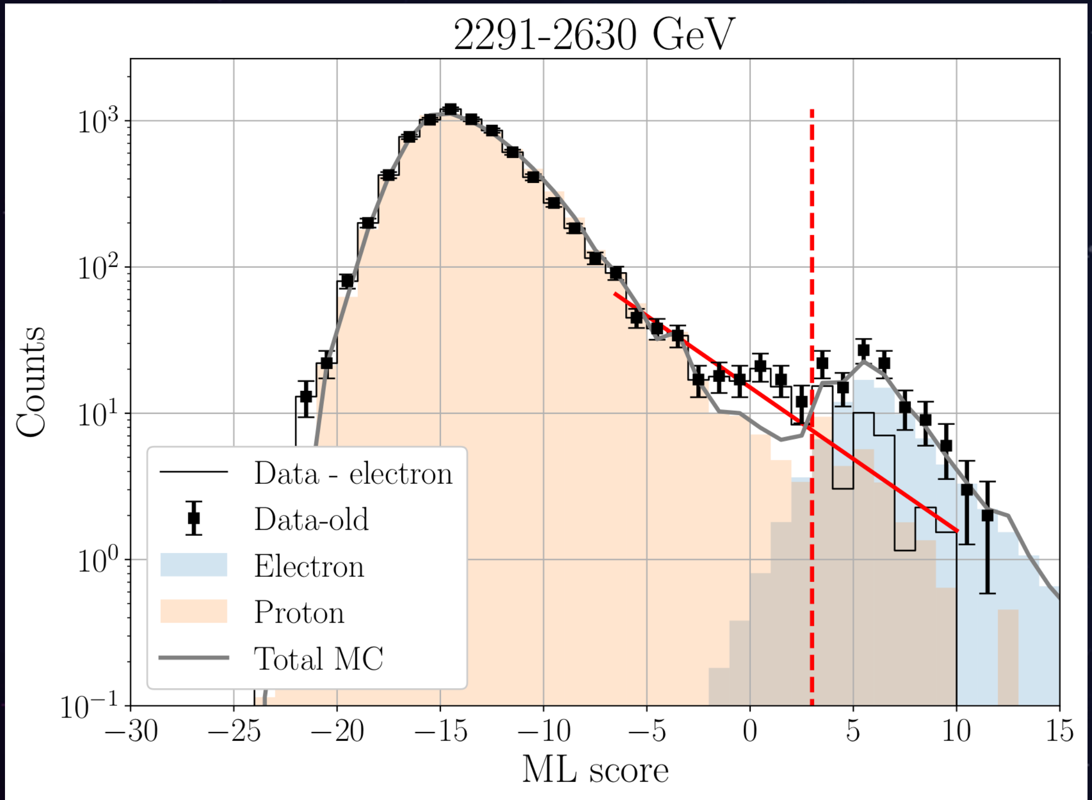


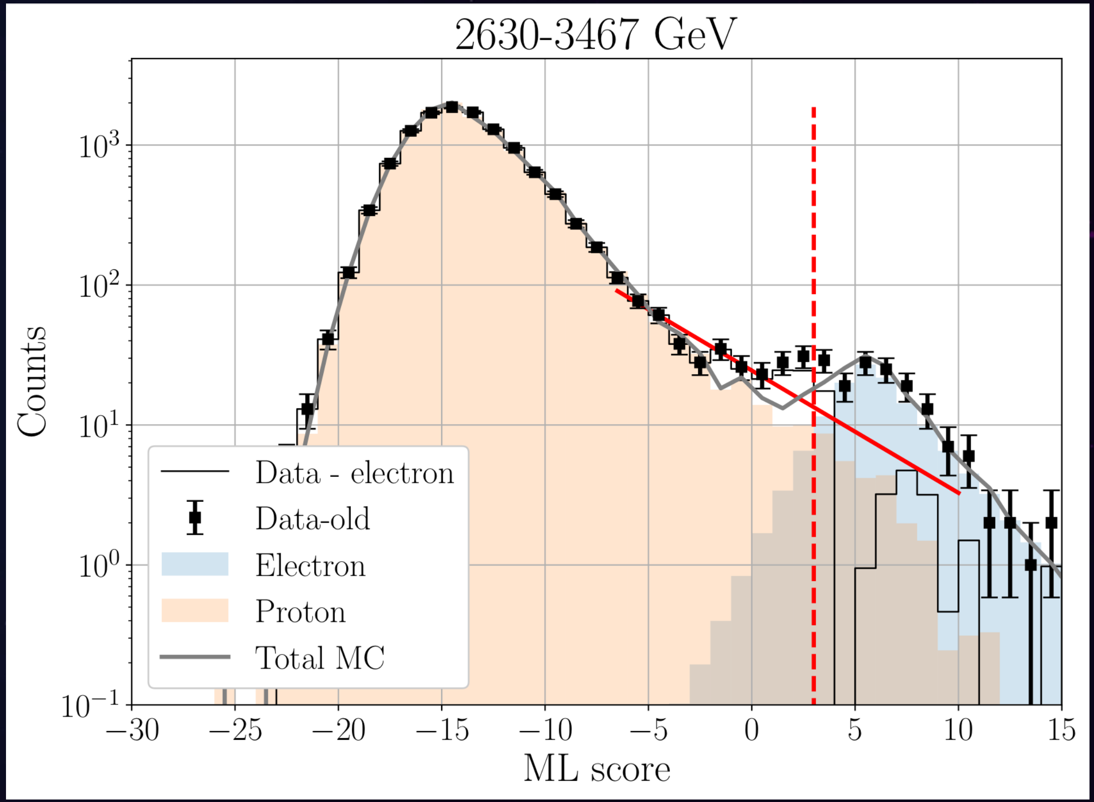
Score distributions



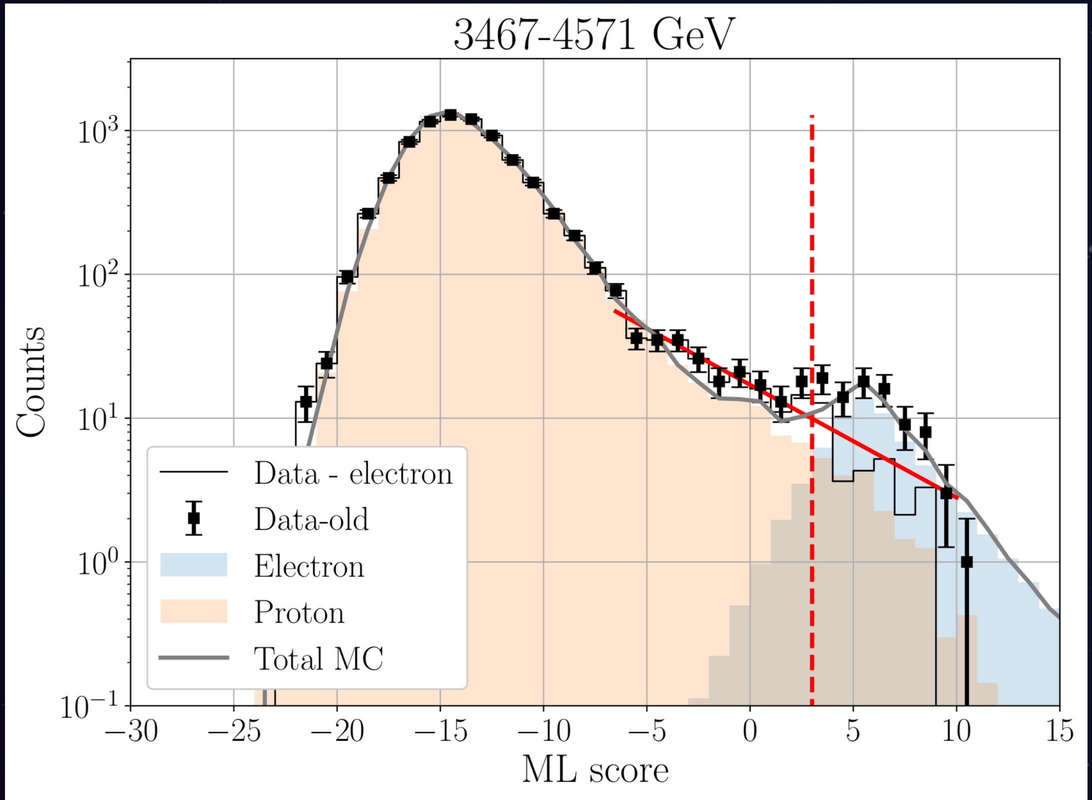


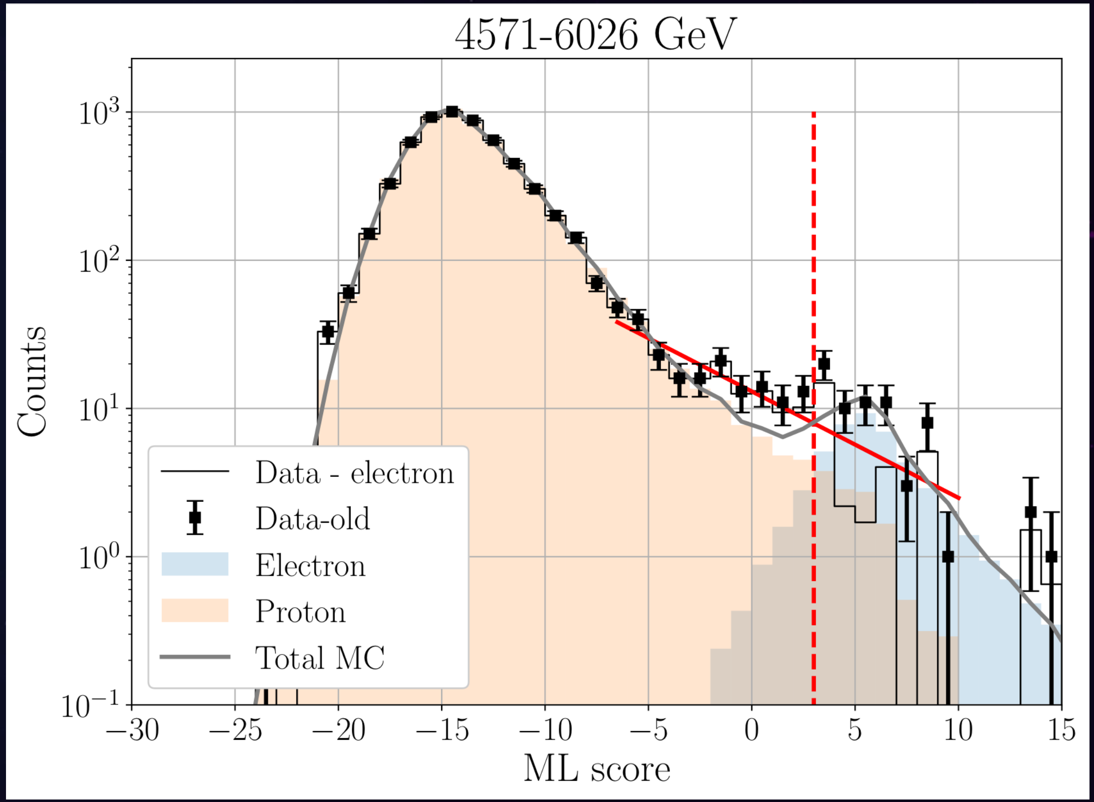
Score distributions



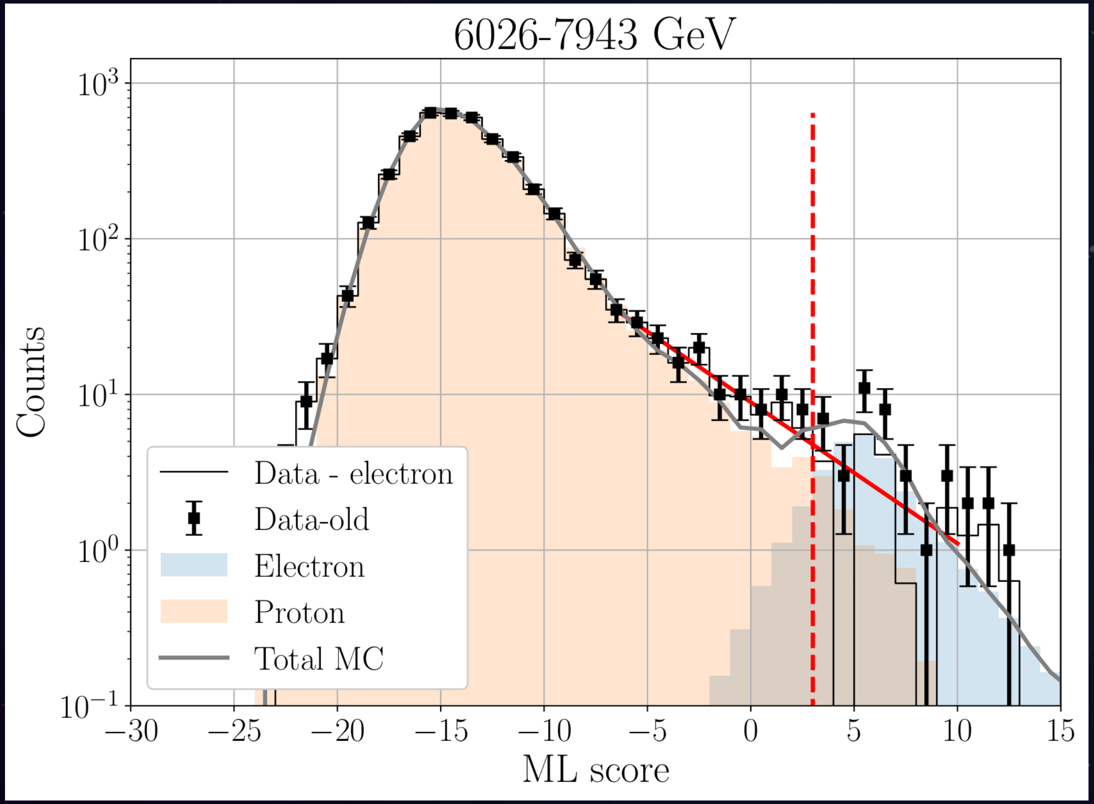


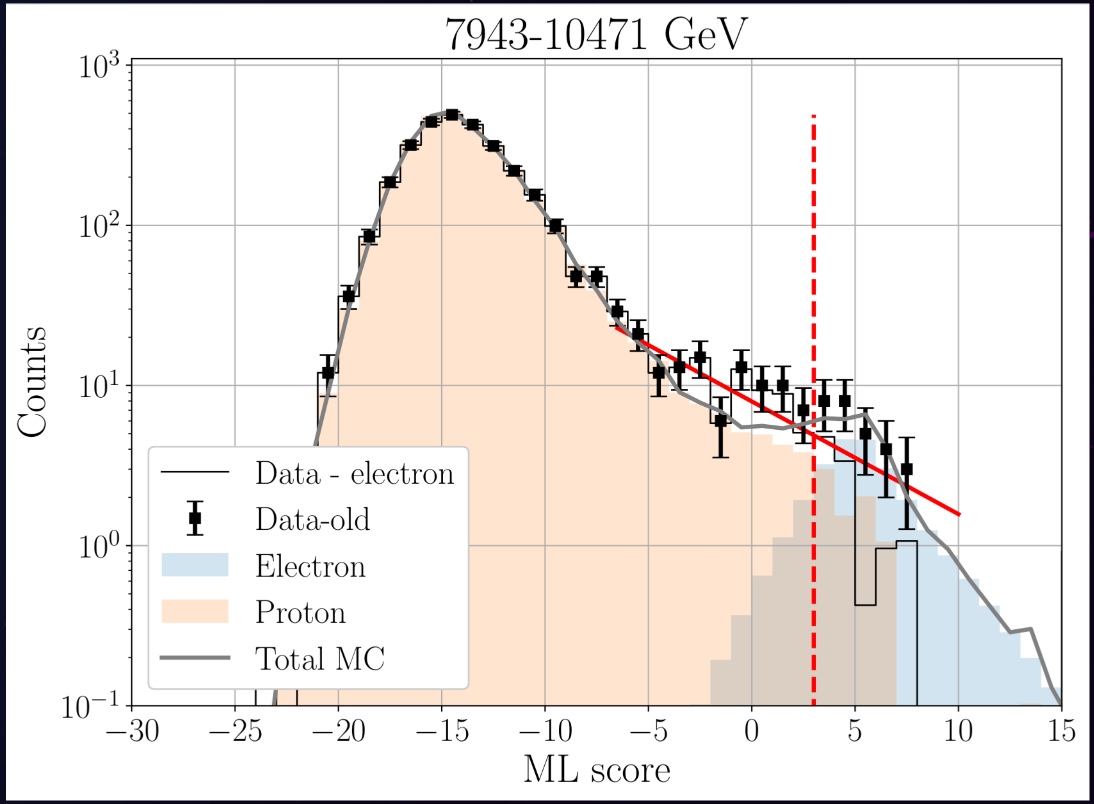
Score distributions



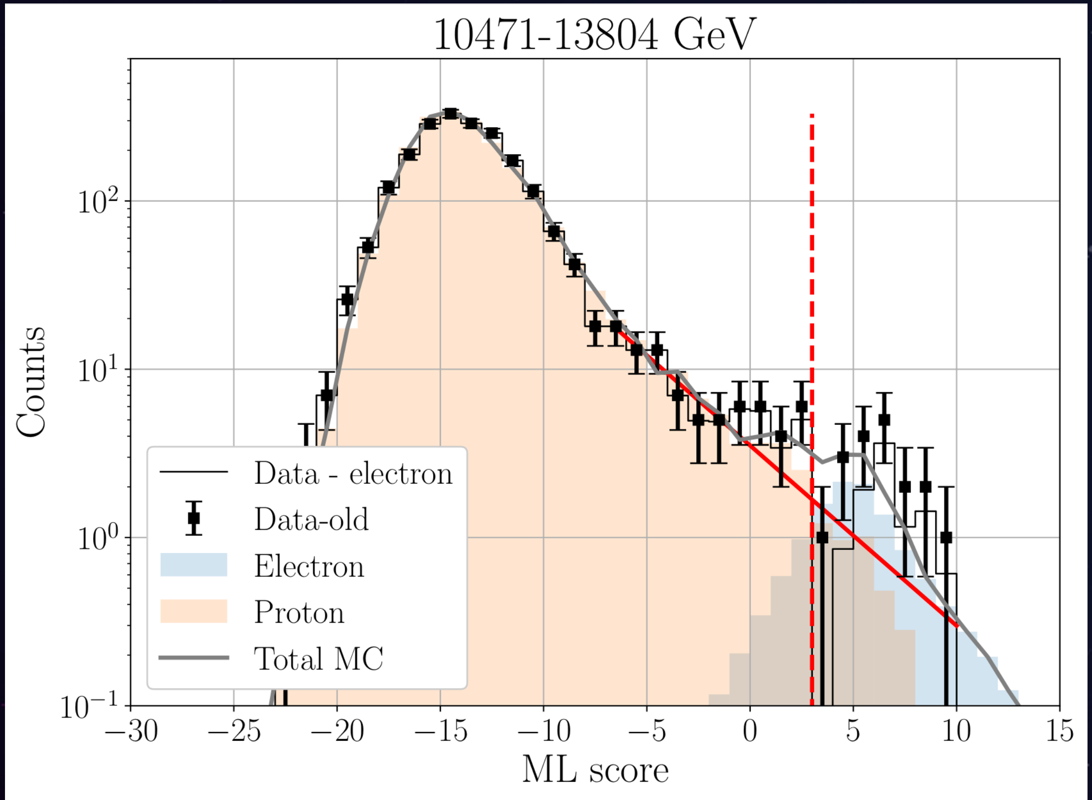


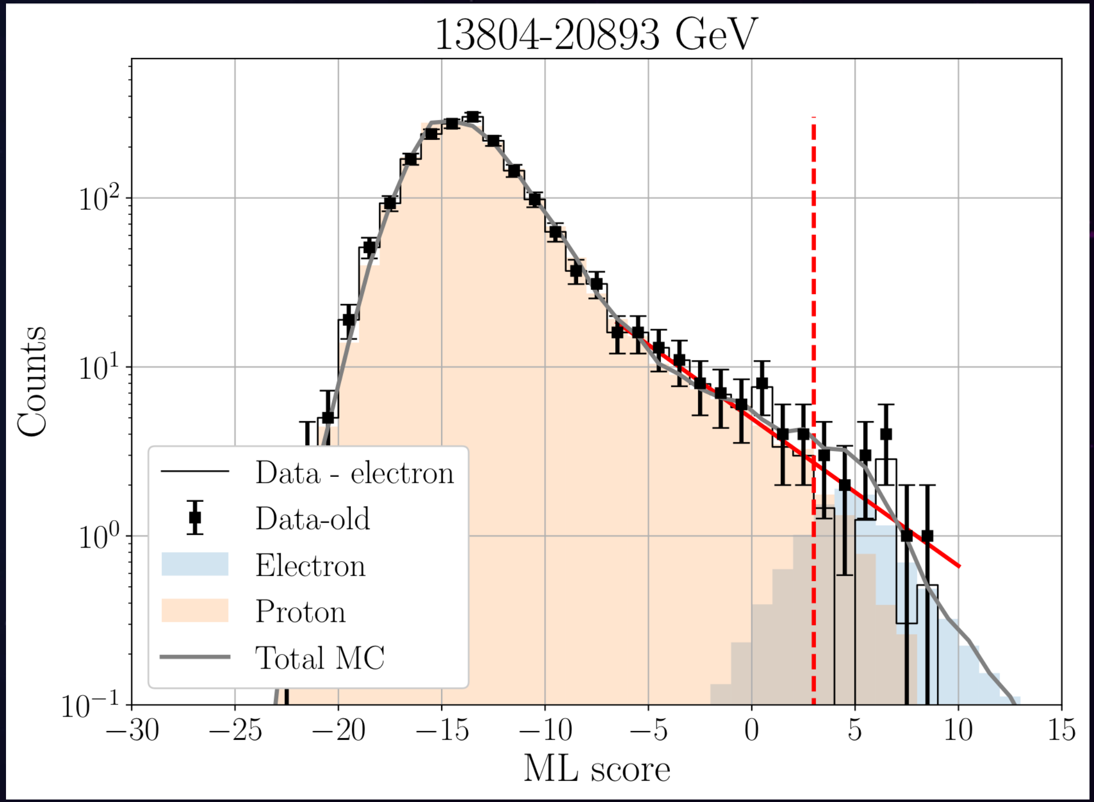
Score distributions



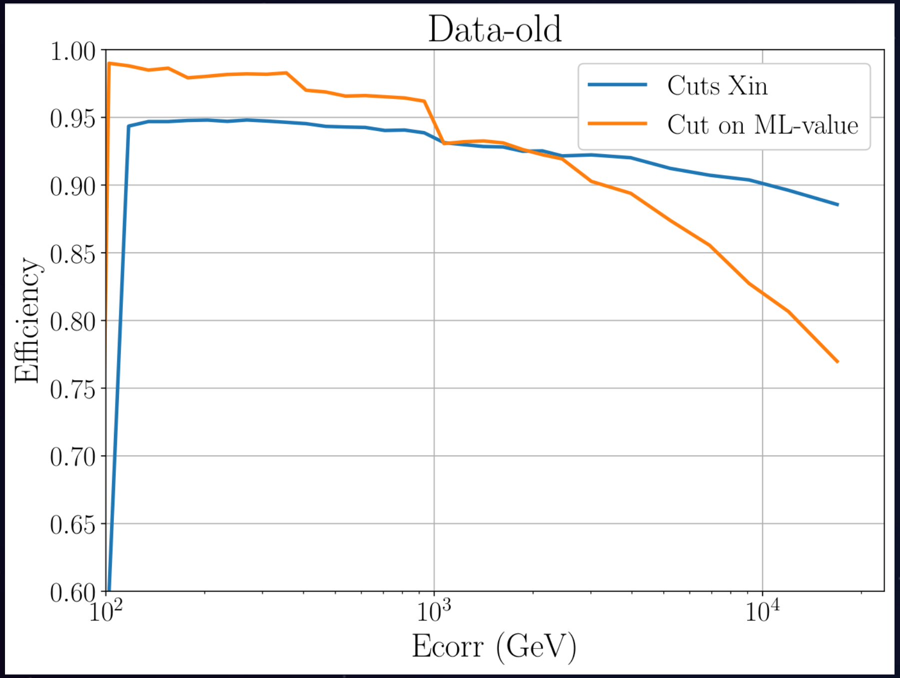


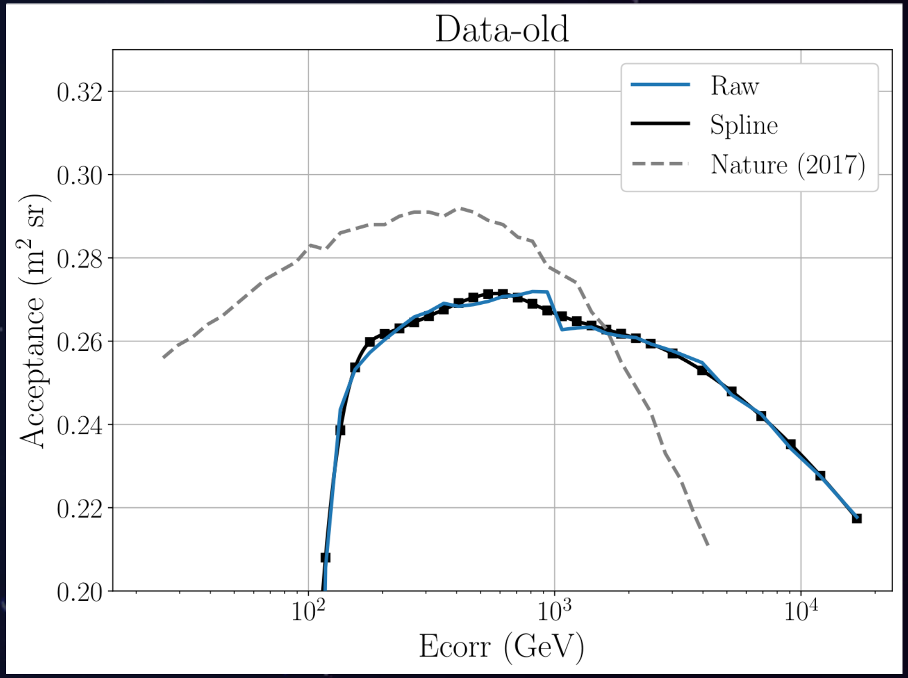
Score distributions



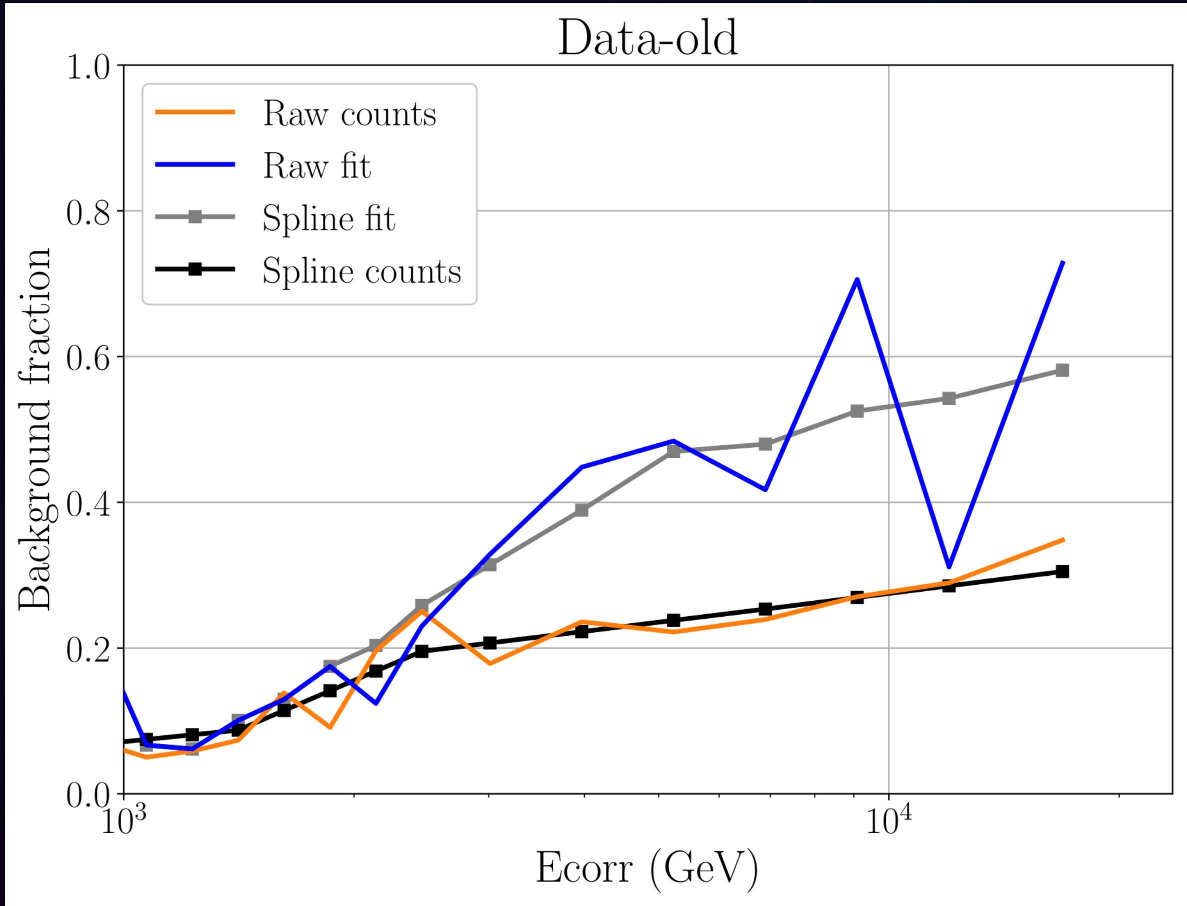


Selection efficiency and acceptance



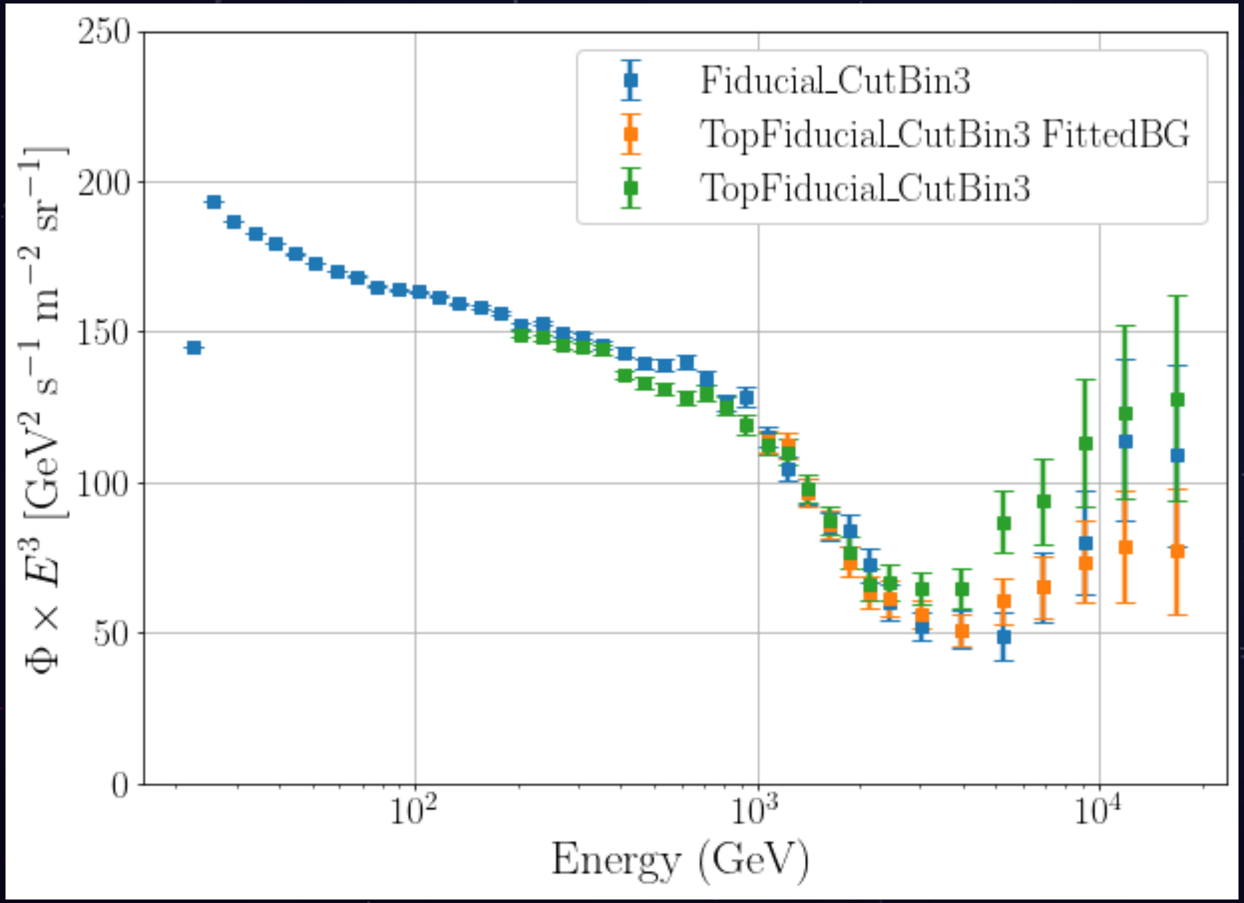


Background



Background can be reliably estimated up to ~2.5 TeV, after which the systematics significantly increase.

Flux (preliminary result)



Excellent agreement up to ~2.5 TeV!

At higher energies, depending on choice of background model:

- Counts → likely underestimates
- Fit → likely overestimates

result over- or undershoots fiducial.

Even with very conservative background, a break is this clearly visible in the spectrum.