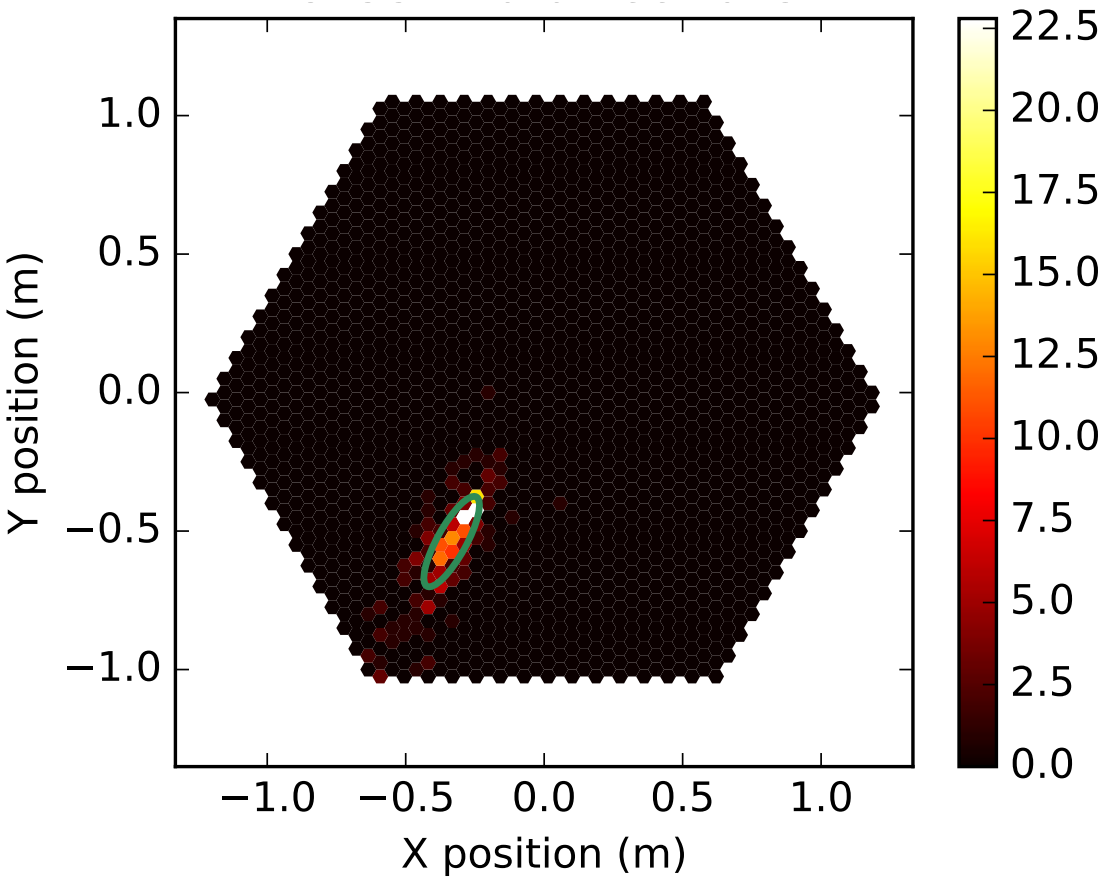
to recap my workflow

Tino Michael
CEA Saclay, Irfu/SAp

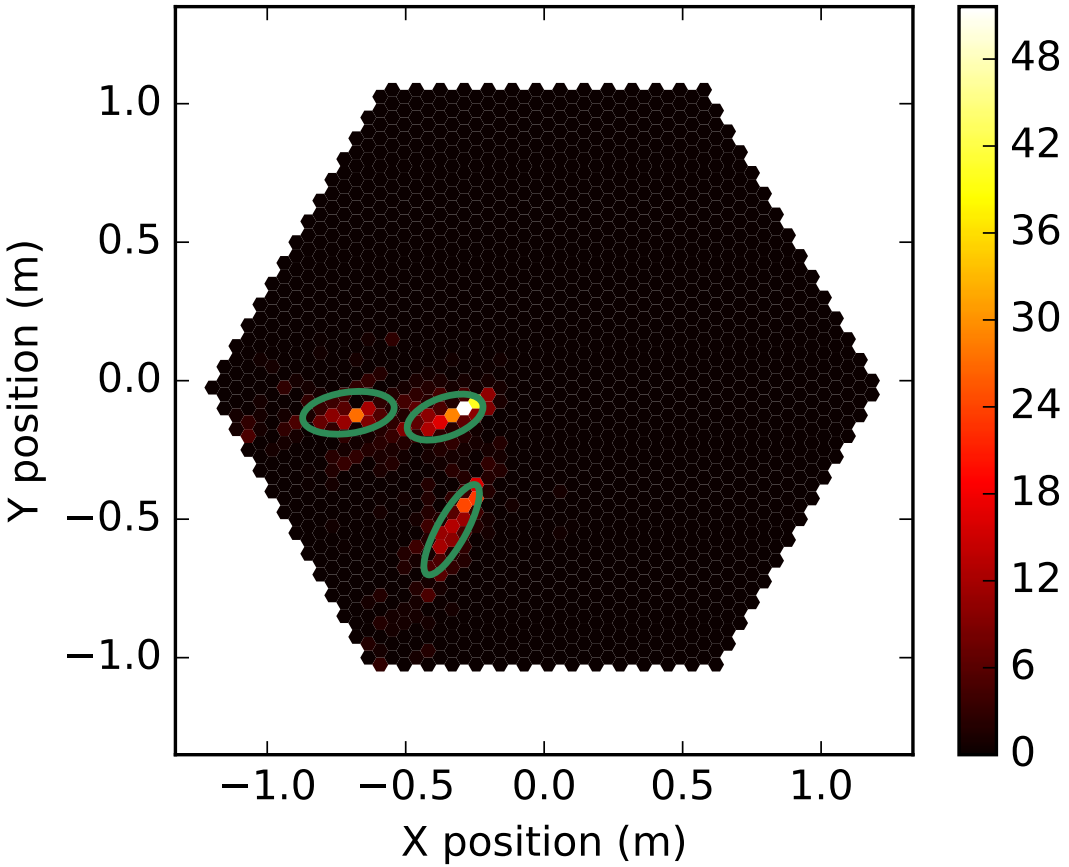
Group Meeting
2017-04-20

Shower Reconstruction



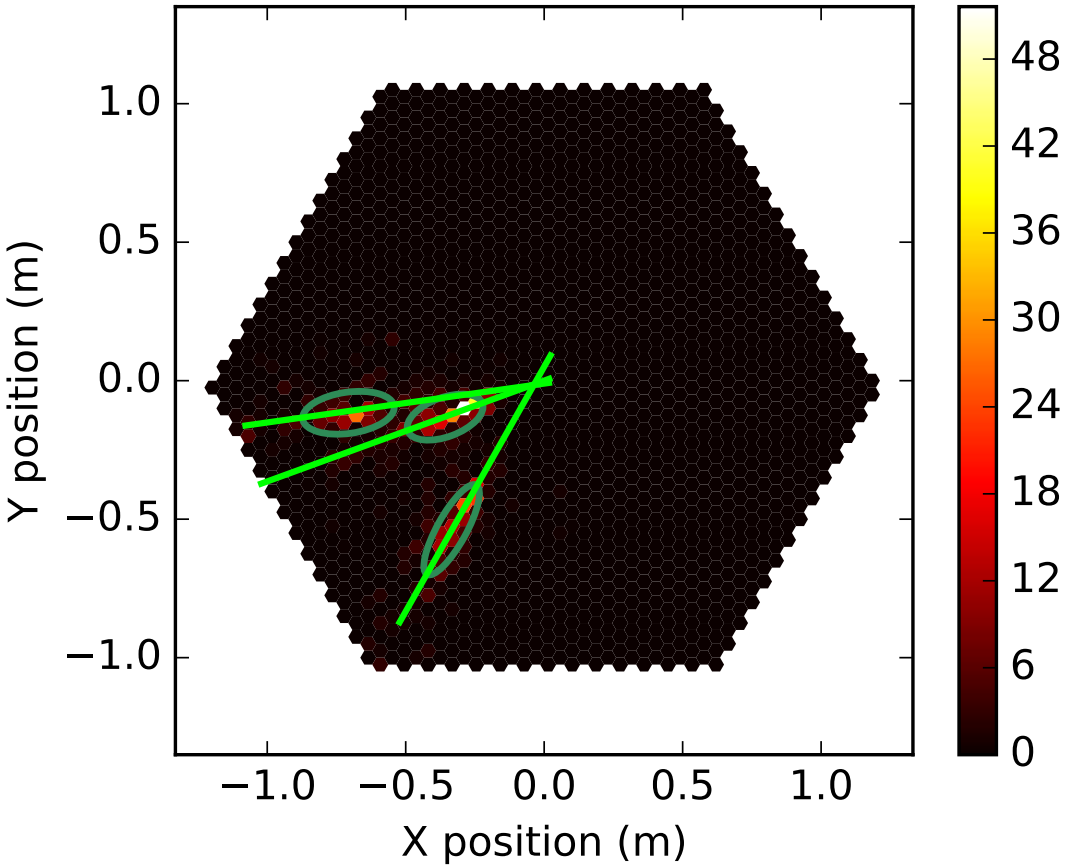
- construct an ellipsis with moments of the shower image:
  Hillas Parametrisation

Shower Reconstruction



- construct an ellipsis with moments of the shower image: *Hillas Parametrisation*
- combine images from different cameras

Shower Reconstruction



- construct an ellipsis with moments of the shower image: *Hillas Parametrisation*
- combine images from different cameras
- intersection of their ellipsis axes is the shower origin

Next Steps

Photon / Proton Discrimination

- Protons pose major background
- Event rate about 10^5 times above Photons

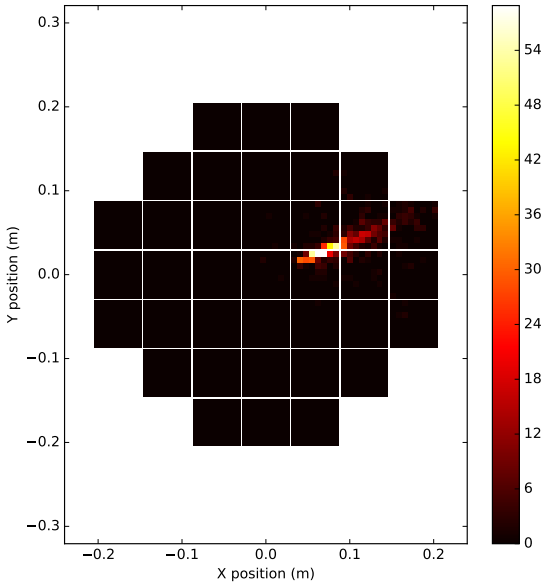
H.E.S.S. methode:

- reducing total signal on camera, length and width of ellipsis and their variances from all telescopes into one parameter to cut on

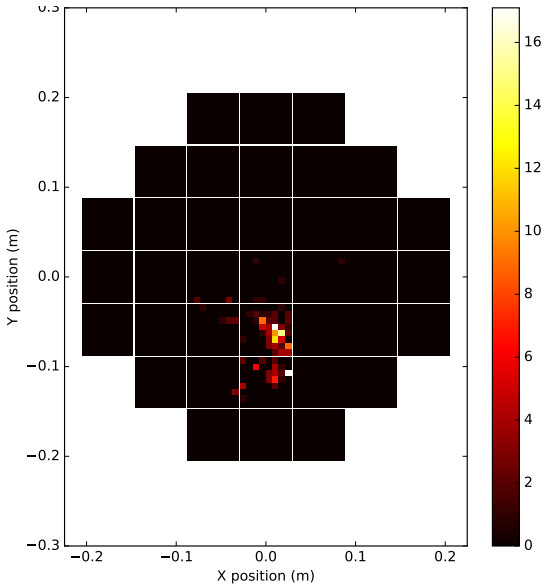
here instead:

- Discrimination with *RandomForestClassifier* fed with parameters from each camera image and the whole event

photon



proton

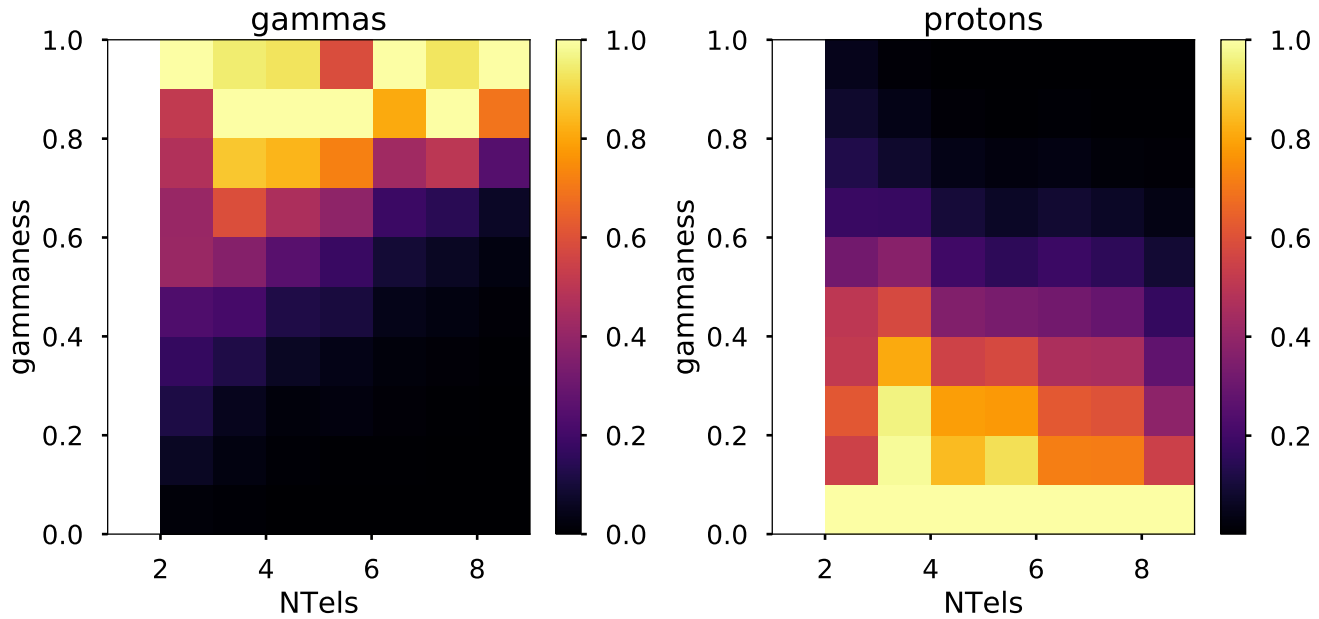


Discrimination

- using RandomForestClassifier implemented in *scikit-learn*
- data-mining approach: just throw all the data at it that we have
  - distance between telescope reconstructed impact position
  - error estimate on the impact position
  - Hillas parameters: width, length, skewness, kurtosis
  - total signal on camera
  - signal of the pixel with the highest count
  - total signal on all selected telescopes
  - number of selected telescopes

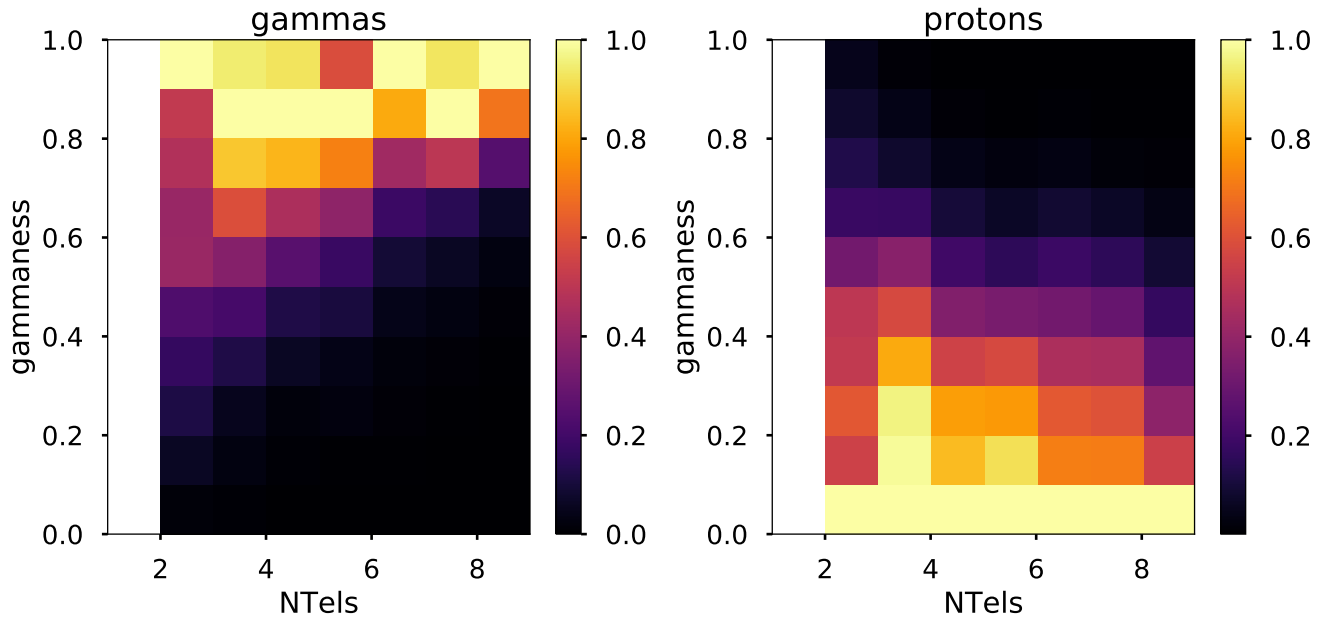
Discrimination

- using RandomForestClassifier implemented in *scikit-learn*
- data-mining approach: just throw all the data at it that we have
  - distance between telescope reconstructed impact position
  - error estimate on the impact position
  - Hillas parameters: width, length, skewness, kurtosis
  - total signal on camera
  - signal of the pixel with the highest count
  - total signal on all selected telescopes
  - number of selected telescopes



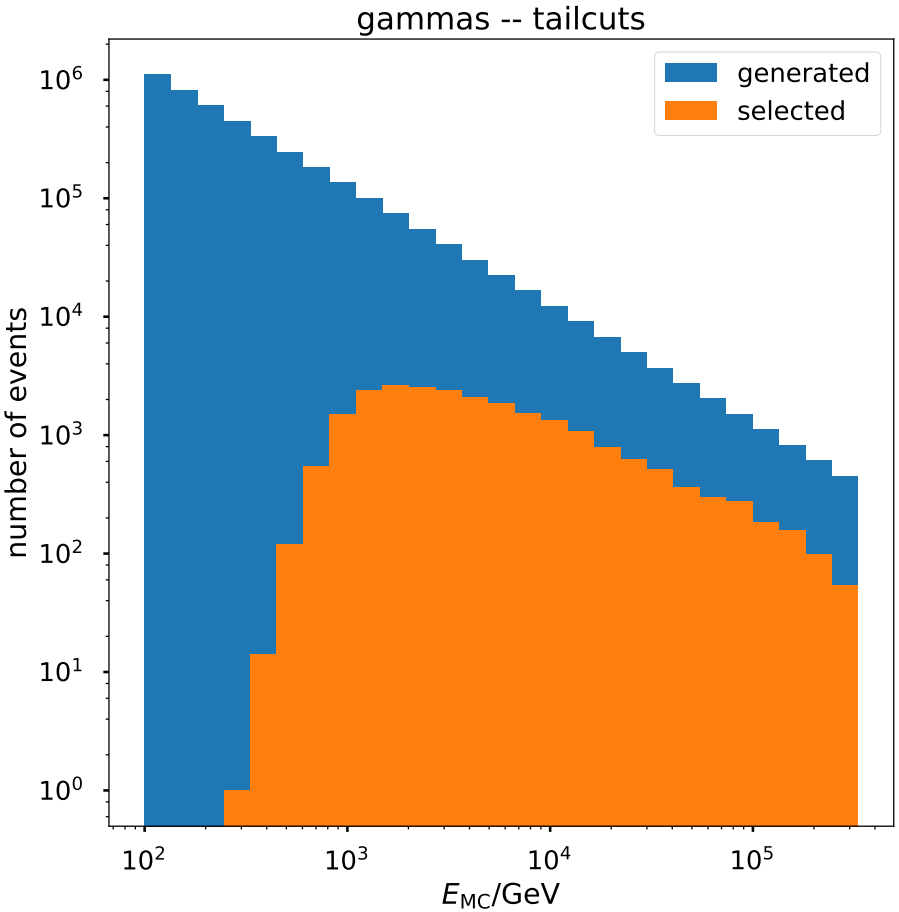
Discrimination

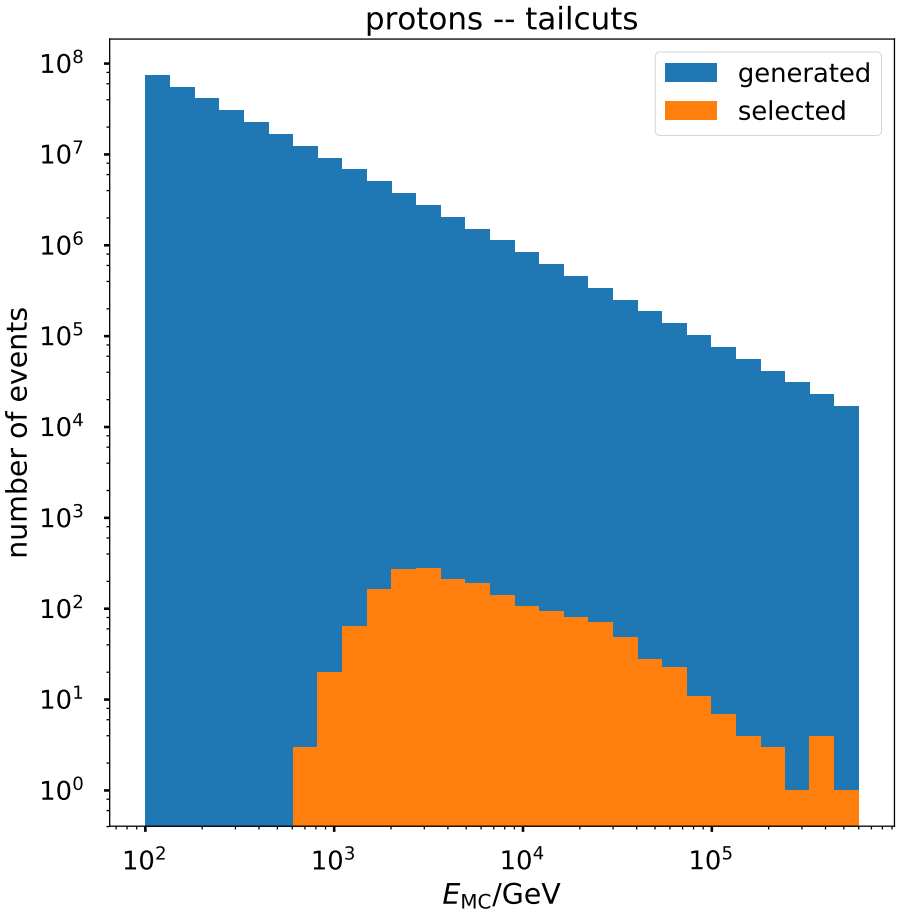
- using RandomForestClassifier implemented in *scikit-learn*
- data-mining approach: just throw all the data at it that we have
  - distance between telescope reconstructed impact position
  - error estimate on the impact position
  - Hillas parameters: width, length, skewness, kurtosis
  - total signal on camera
  - signal of the pixel with the highest count
  - total signal on all selected telescopes
  - number of selected telescopes



for now, cut on $NTels > 2$ & $gammaness > 0.75$

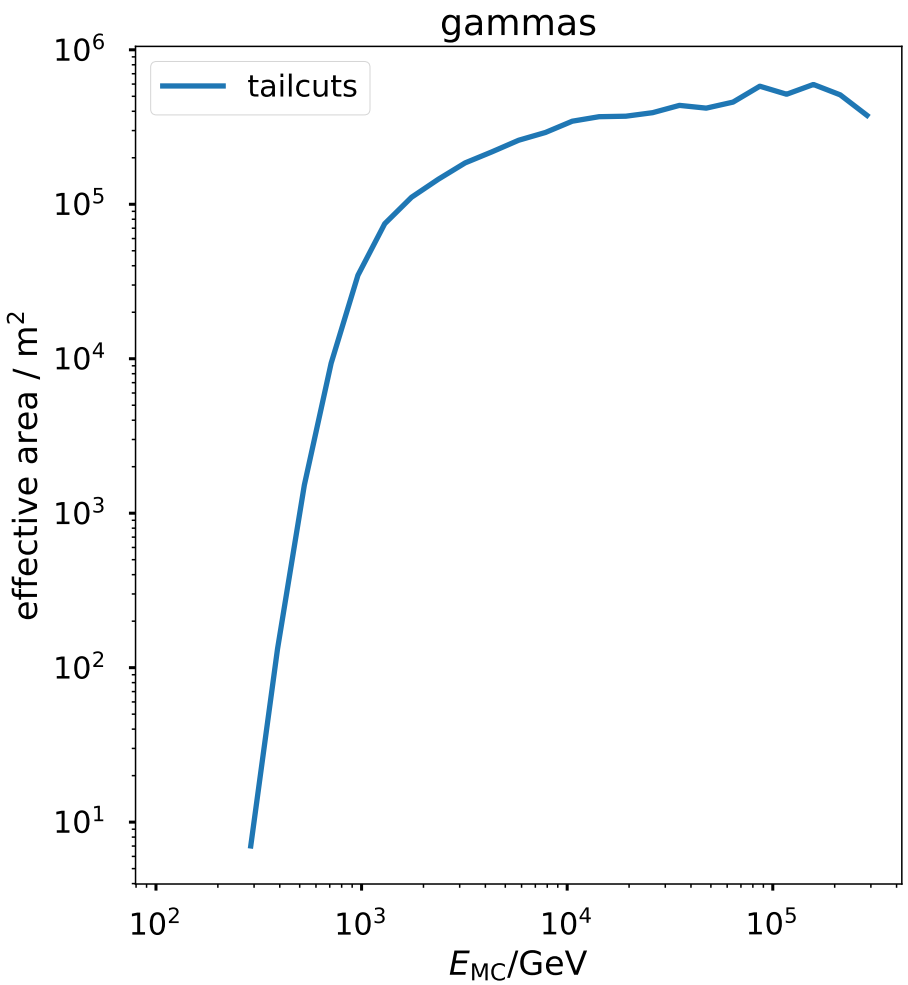
generated and selected MC events

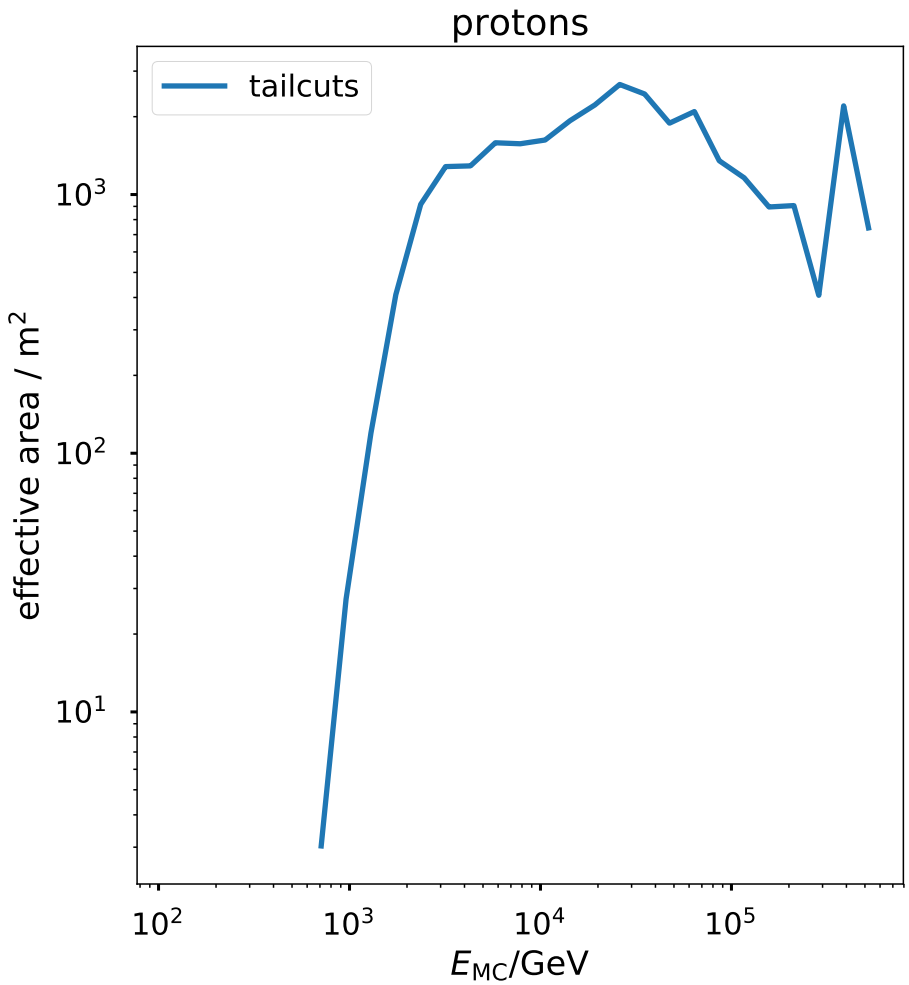




Effective Area

taking the ratio of the previous plots (i.e. the selection efficiency) and multiply every bin with the area in which the MC events have been generated in: *Effective Area*





- next step: reweighting of MC events to correspond to expected physical flux (e.g. Crab nebula)

- next step: reweighting of MC events to correspond to expected physical flux (e.g. Crab nebula)
- but how?

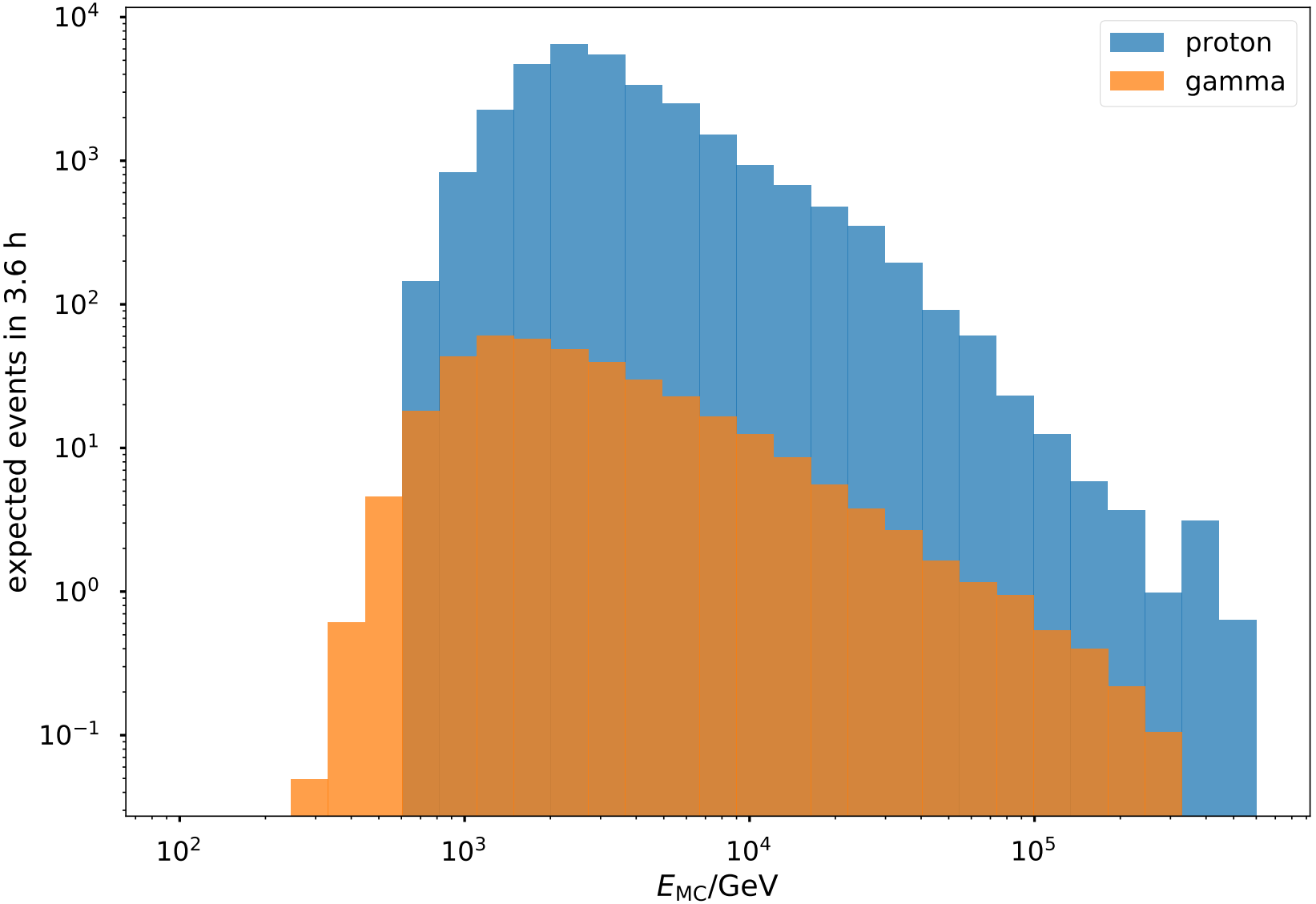
- next step: reweighting of MC events to correspond to expected physical flux (e.g. Crab nebula)
- simple binned approach:
  - already have the energy-binned efficiencies
  - apply these on the energy-binned histogram of expected *arriving* events from the source
  - $\longrightarrow$ get the number of expected selected events

- next step: reweighting of MC events to correspond to expected physical flux (e.g. Crab nebula)
- simple binned approach:
  - already have the energy-binned efficiencies
  - apply these on the energy-binned histogram of expected *arriving* events from the source
  - $\longrightarrow$ get the number of expected selected events
- **but:** it's binned... not nice

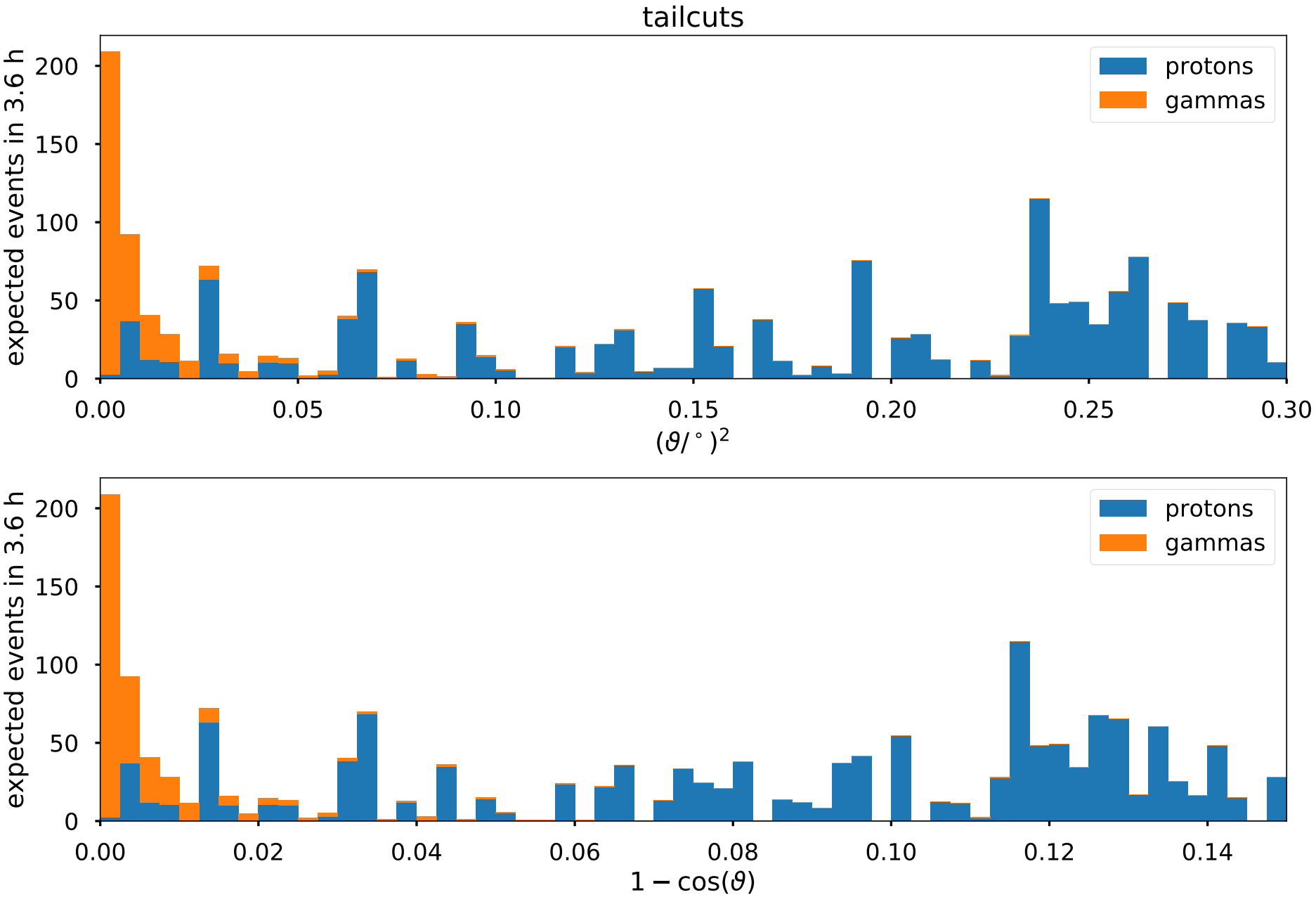
- next step: reweighting of MC events to correspond to expected physical flux (e.g. Crab nebula)
- instead: event-by-event weight that considers the generator spectrum:
- $w(E) = A_{\text{gen}} \times I_{\Theta} \times E^{\gamma} \times I_E \times T_{\text{obs}}/N_{\text{gen}}$
  with:
  - A_{gen}: MC generator Area
  - $I_{\Theta} = 2\pi(1 - \cos\vartheta)$: angular phase space factor for diffuse flux
  - E^{γ}: considers that MC events have been drawn with an E^{-2} spectrum
  - γ: spectral index of the MC generator (here equal 2)
  - $I_E = (E_{\text{max}}^{(1-\gamma)} - E_{\text{min}}^{(1-\gamma)})/(1-\gamma)$: energy phase space factor
  - T_{obs}: assumed observation time
  - N_{gen}: number of generated MC events
- $w(E) \times \Phi(E)$ gives weight for every MC event so that their energy distribution looks like the selected events from the assumed flux Φ

- next step: reweighting of MC events to correspond to expected physical flux (e.g. Crab nebula)
- instead: event-by-event weight that considers the generator spectrum:
- $w(E) = A_{\text{gen}} \times I_{\Theta} \times E^{\gamma} \times I_E \times T_{\text{obs}}/N_{\text{gen}}$
  with:
  - A_{gen}: MC generator Area
  - $I_{\Theta} = 2\pi(1 - \cos\vartheta)$: angular phase space factor for diffuse flux
  - E^{γ}: considers that MC events have been drawn with an E^{-2} spectrum
  - γ: spectral index of the MC generator (here equal 2)
  - $I_E = (E_{\text{max}}^{(1-\gamma)} - E_{\text{min}}^{(1-\gamma)})/(1-\gamma)$: energy phase space factor
  - T_{obs}: assumed observation time
  - N_{gen}: number of generated MC events
- $w(E) \times \Phi(E)$ gives weight for every MC event so that their energy distribution looks like the selected events from the assumed flux Φ
- described in old ANTARES internal note: ANTARES-SOFT-1999-003

Expected Events from Crab and Cosmic Rays



Expected Events from Crab and Cosmic Rays



calculating Sensitivity

- define on- and off-regions: 0.15° around MC source

  on-region count $N_{\text{on}} = N_\gamma + N_\text{p}$

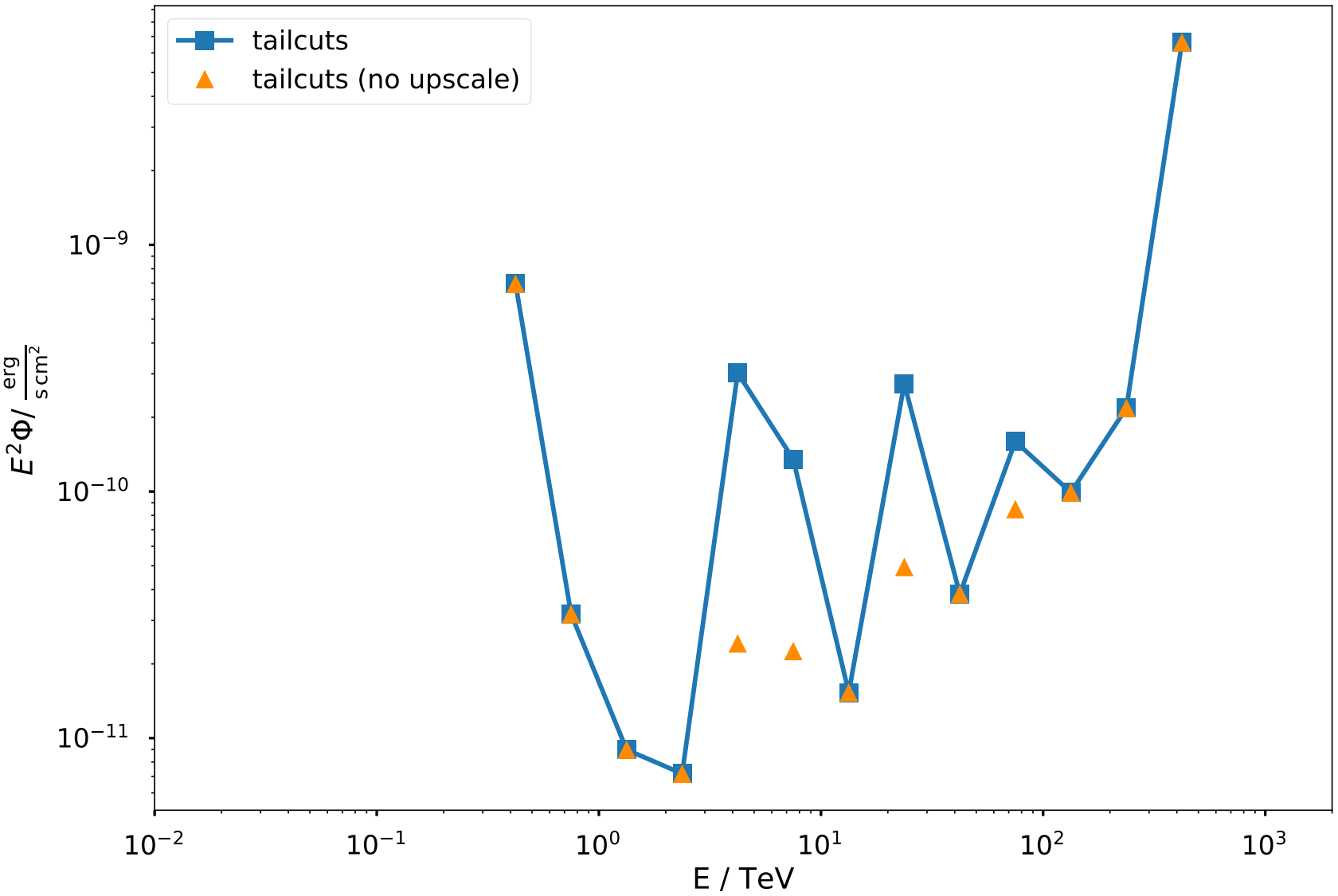
  off-region count $N_{\text{off}} = N_\text{p}$

- significance given by Li Ma (1983):

```
alpha1 = alpha + 1.0
sum    = Non + Noff
arg1   = Non / sum
arg2   = Noff / sum
term1  = Non * np.log((alpha1/alpha)*arg1)
if Noff == 0:
    term2 = 0
else:
    term2  = Noff * np.log(alpha1*arg2)
sigma  = np.sqrt(2.0 * (term1 + term2))
```

- given the expected N_γ from the assumed source, scale the flux up or down until $\mathit{sigma} = 5 \longrightarrow$ this is our sensitivity

Sensitivity of the ASTRI mini-array



that blue line...

- orange triangles represent the 5σ flux
- blue line due to the additional CTA requirements on the sensitivity:
- in every energy bin there need to be at least 10 events
- and a maximum background contribution of 5 %