

Searches for photons above 1EeV: a novel approach using hybrid data from the Pierre Auger Observatory

Costanza Marinelli (XL cycle)

GSSI Passage of the year, October 2026

Advisors:

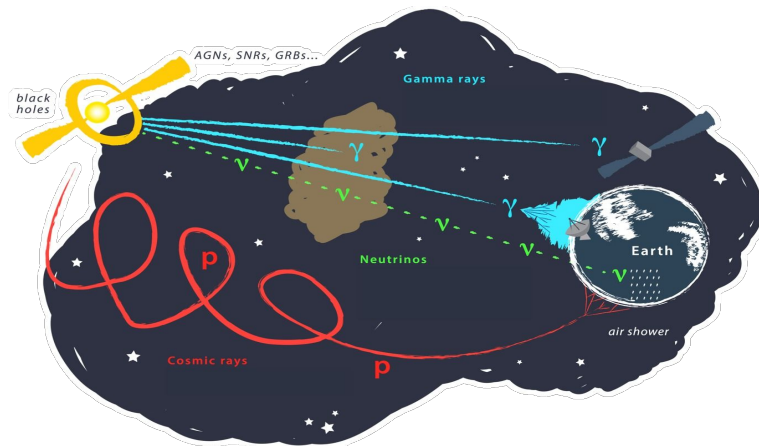
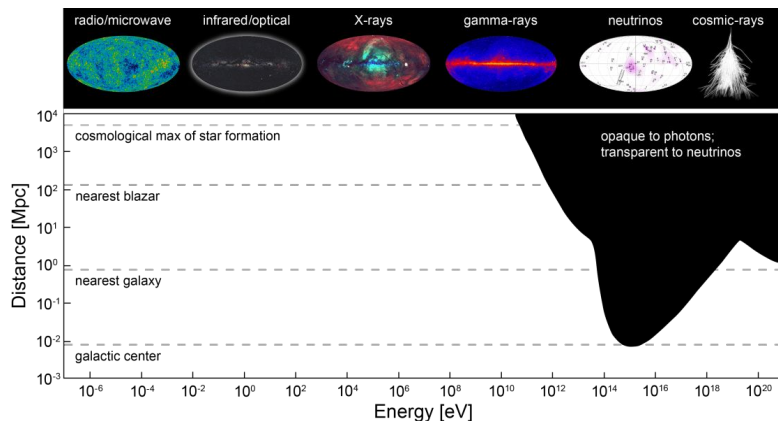
Carmelo Evoli, Pierpaolo Savina



Why looking for UHE photons?

Neutral particles unique probes:

- **Direct information** on their sources
- Smoking guns for fundamental physics
- Improving **understanding** of the **extreme Universe**



UHE-photons highly relevant for:

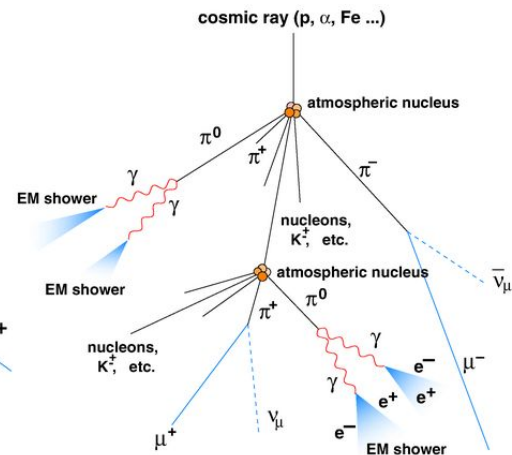
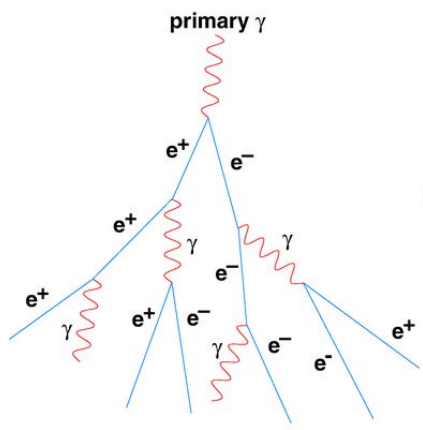
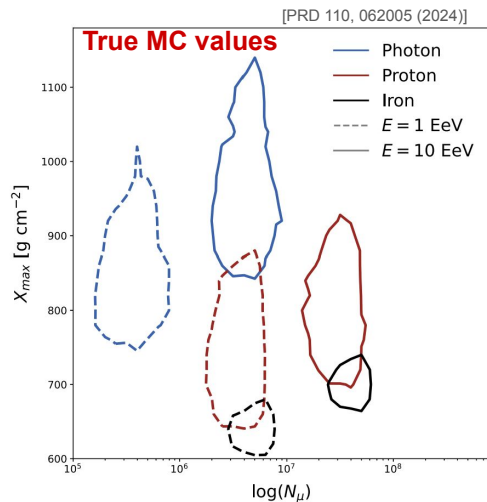
- **Probing UHECRs mass composition**
- Unveiling local universe sources
- **Source-targeted search**
- **SHDM annihilation** signatures search in the Galactic centre

Key signatures

The **steeply falling spectra require large ground detector areas**

Gamma rays reaching the atmosphere interact producing **extensive air showers (EAS)** of secondary particles. They are distinguished from hadronic EAS through some features:

- **Low muonic content**
- **Deeper $X_{\max}$**



The Pierre Auger Observatory

Hybrid ground-based **detector**:

Surface Detector (SD):

- SD1500: 1600 stations - 1.5 km grid
- SD750: 61 stations - 750 m grid
- SD433: 19 stations - 433 m grid

Sampling the secondary particles reaching the ground

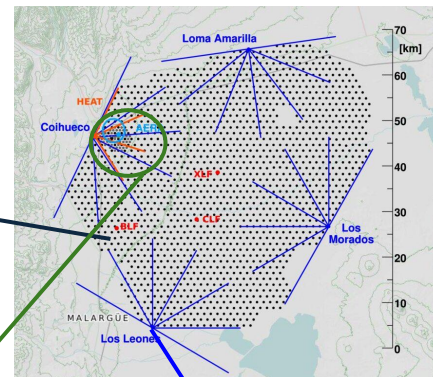
Fluorescence Detector (FD): measures the fluorescence light produced by de-excitation of air nitrogen molecules

Additional detector systems complement the main SD:

Radio antennas (RD)

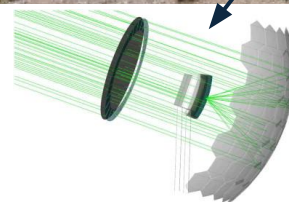
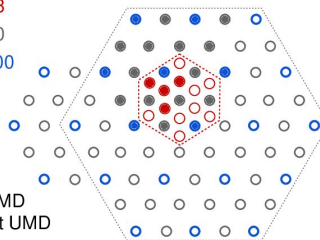
Scintillator planes (SSD)

Underground muon detectors (UMD)



- SD-433
- SD-750
- SD-1500

- With UMD
- Without UMD

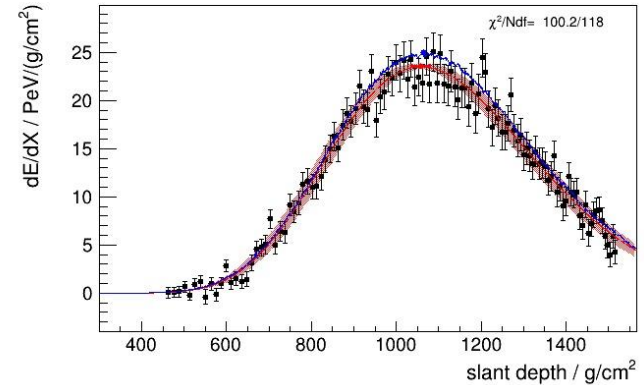
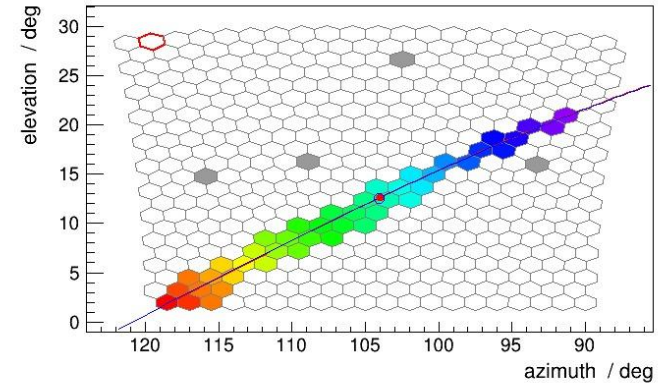
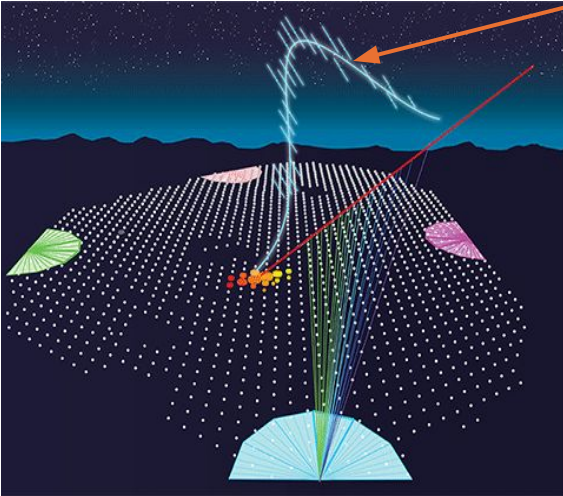


The observables

To observe photons we need to measure $X_{\max}$ and the **number of muons**

$X_{\max}$ directly measured by FD

APS/Carin Cain



The observables: universality model

Air shower universality:

Shower as a thermodynamic system

It describes the n. of particles and their distributions at different development stages through a few parameters

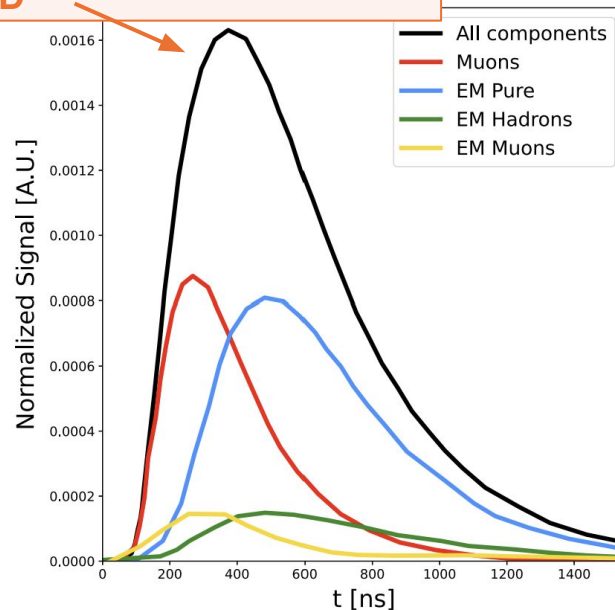
Total signal universality model:

Exploits this property to predict the **total signal in the tanks** as the **superposition of 4 components**

Their contribution depends on the primary particle mass through F_μ

$$S_{\text{pred}} = \sum_{i=1}^4 S_i = \sum_{i=1}^4 f_i(F_\mu) S_{i,\text{comp}}$$

The muon content is inferred from SD



Standard analysis vs new method

Standard research

Selection of suitable observables
MVA algorithms to combine them
Cut to look for candidates

Alternative approach

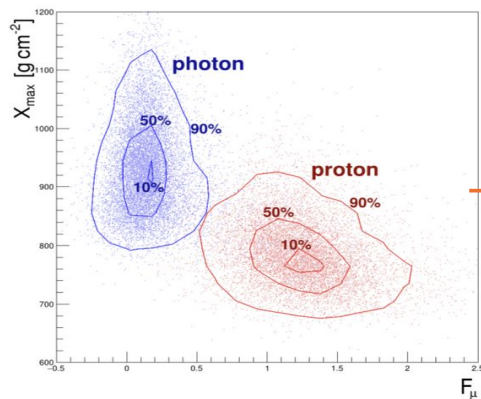
Search for UHE photons as **excesses** of the expected background

1. Optimisation of selection cuts

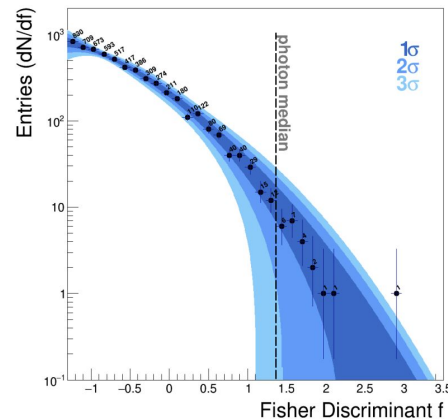
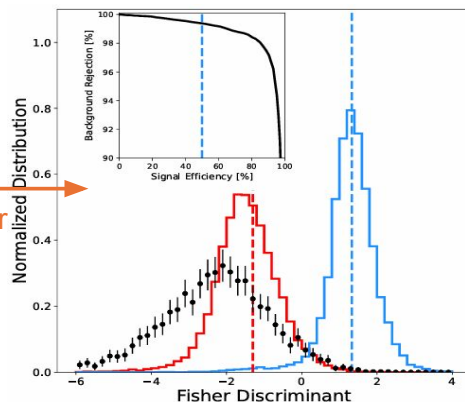
We want to apply cuts to reduce the tails of the background and **make the signal pop out**

2. Likelihood-ratio test

It is a powerful tool **to search for photon excesses**



Fisher



Standard analysis vs new method

Standard research

Selection of suitable observables

MVA algorithms to combine them

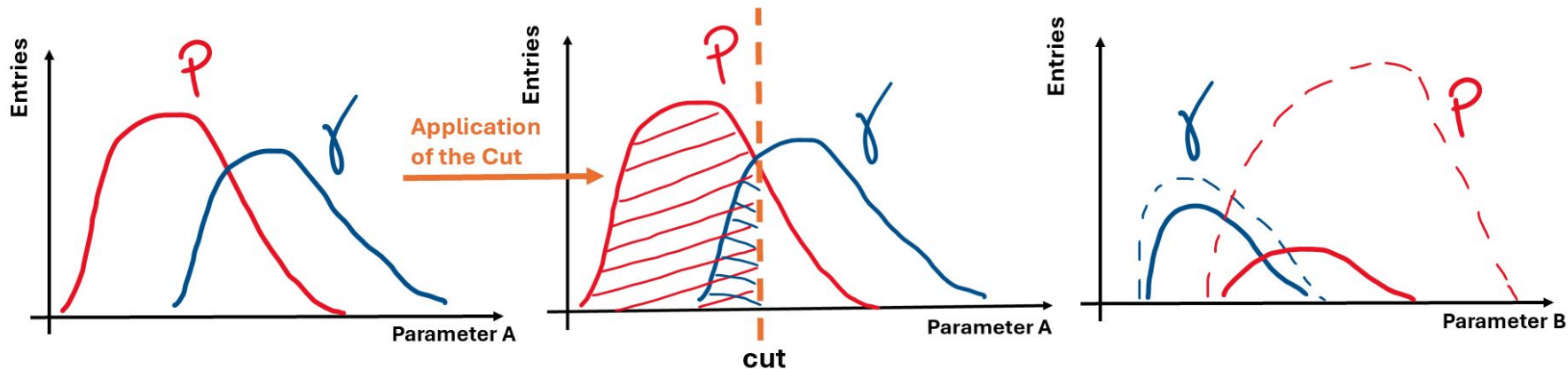
Cut to look for candidates

Alternative approach

Search for UHE photons as **excesses** of the expected background

1. Optimisation of selection cuts

We want to apply cuts to reduce the tails of the background and **make the signal pop out**



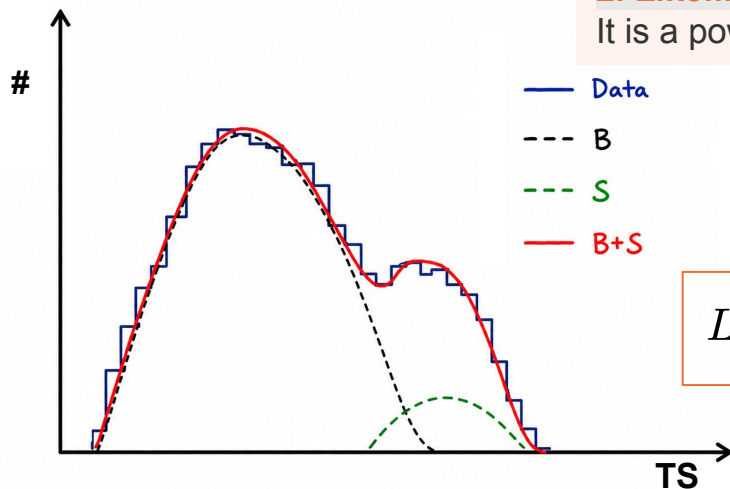
Standard analysis vs new method

Standard research

Selection of suitable observables

MVA algorithms to combine them

Cut to look for candidates



Alternative approach

Search for UHE photons as **excesses** of the expected background

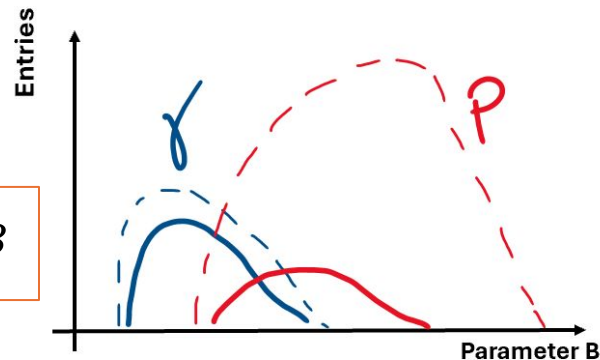
1. Optimisation of selection cuts

We want to apply cuts to reduce the tails of the background and **make the signal pop out**

2. Likelihood-ratio test

It is a powerful tool **to search for photon excesses**

$$L = \frac{n_s}{N} S + \left(1 - \frac{n_s}{N}\right) B$$



The dataset

Air-shower simulations

CORSIKA: FLUKA (low energy) | EPOS-LHC (high energy)

y: Signal **p:** Background

$\log(E / \text{eV}) \in [17.5, 19.5]$ $\theta \in [0-65^\circ]$

Realistic detector simulation for hybrid events

- **Surface detector:** Water-Cherenkov tanks
- **Fluorescence detector**

The dataset

Air-shower simulations

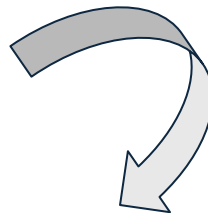
CORSIKA: FLUKA (low energy) | EPOS-LHC (high energy)

γ : Signal p : Background

$\log(E / \text{eV}) \in [17.5, 19.5]$ $\theta \in [0-65^\circ]$

Realistic detector simulation for hybrid events

- Surface detector: Water-Cherenkov tanks
- Fluorescence detector



Old and newer version of **EPOS-LHC** tested:
different features between the two **affect** the result
of our **analysis**

The dataset

Air-shower simulations

CORSIKA: FLUKA (low energy) | EPOS-LHC (high energy)

γ : Signal p : Background

$\log(E / \text{eV}) \in [17.5, 19.5]$ $\theta \in [0-65^\circ]$

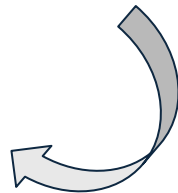
Realistic detector simulation for hybrid events

- Surface detector: Water-Cherenkov tanks
- Fluorescence detector

Old and newer version of **EPOS-LHC** tested:
different features between the two **affect** the result
of our **analysis**

5% burnt sample used in previous Auger analysis **exploited:**

- No assumption on background **mass composition**
- Lower model dependency



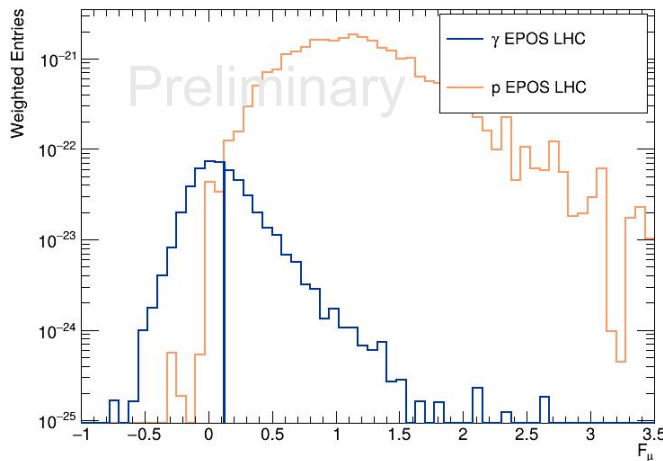
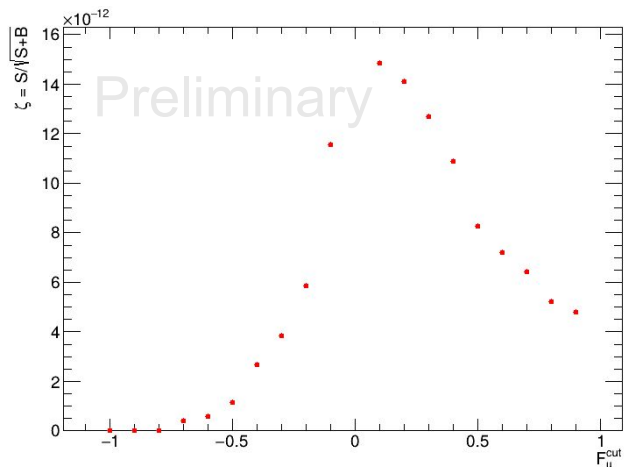
Step 1: F_μ -based cut optimisation

F_μ is nearly energy-independent

Optimal cut: F_μ^{cut} value giving maximum signal-to-noise ratio, ζ

We scan over different F_μ^{cut} values and compute ζ

$$\zeta = \frac{N_s(F_\mu < F_\mu^{\text{cut}})}{\sqrt{N_s(F_\mu < F_\mu^{\text{cut}}) + N_B(F_\mu < F_\mu^{\text{cut}})}}$$



Previous hadronic model:

$F_\mu^{\text{cut}} = 0.2$ with **signal efficiency ~78%** and **background contamination ~0.2%**

Optimal cut: $F_\mu^{\text{cut}} = 0.12$

Signal efficiency is ~58% and background contamination ~0.3%

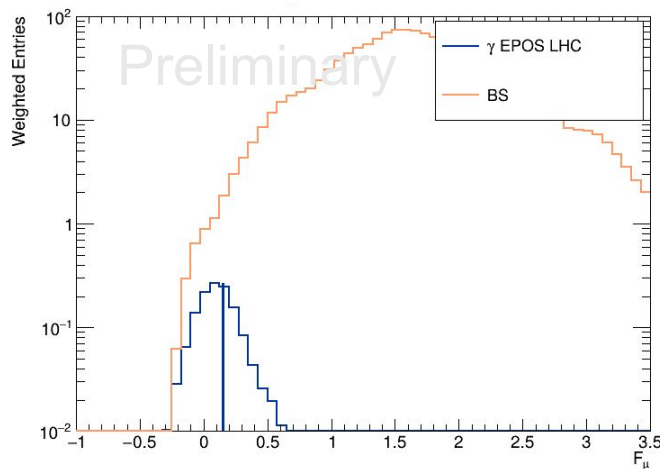
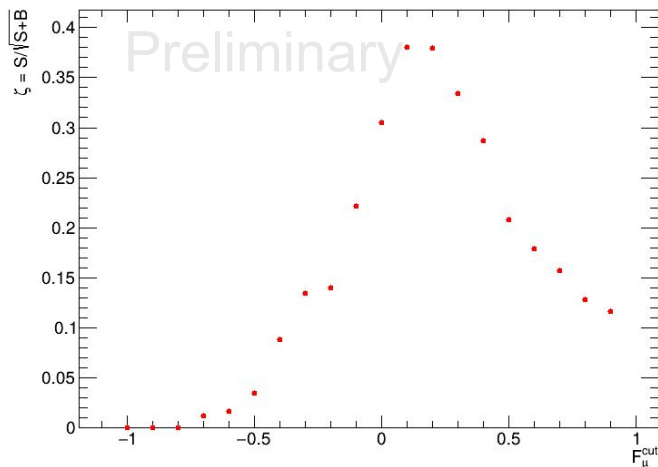
Step 1: F_μ -based cut optimisation on BS

We here use the **5% BS** in place of proton simulations

The **new optimal cut** is: $F_\mu^{\text{cut}} = 0.15$

Signal efficiency is **~72%** and background contamination **~0.4%**

$$\zeta = \frac{N_S(F_\mu < F_\mu^{\text{cut}})}{\sqrt{N_S(F_\mu < F_\mu^{\text{cut}}) + N_B(F_\mu < F_\mu^{\text{cut}})}}$$



With simulations:

$F_\mu^{\text{cut}} = 0.12$,

signal efficiency ~58%,

background

contamination **~0.3%**

Step 2: likelihood-ratio test

Aim: compare the new method with the latest method: [PRD 110, 062005 (2024)]

H_0 : Data described solely by background $\mathbf{B}$ ($n_s = 0$)

H_1 : Data described by background $\mathbf{B}$ + signal $\mathbf{S}$

Where:
$$L = \frac{n_s}{N} S(E) + \left(1 - \frac{n_s}{N}\right) B(E) \longrightarrow \text{TS} = -2\ln\left(\frac{L(n_s = 0)}{L(\hat{n}_s)}\right)$$

$\mathbf{S(E)}$ and $\mathbf{B(E)}$ described by **Gumbel** distributions: $\frac{1}{\beta} e^{-(z+e^{-z})}$, where: $z = \frac{x - \mu}{\beta}$

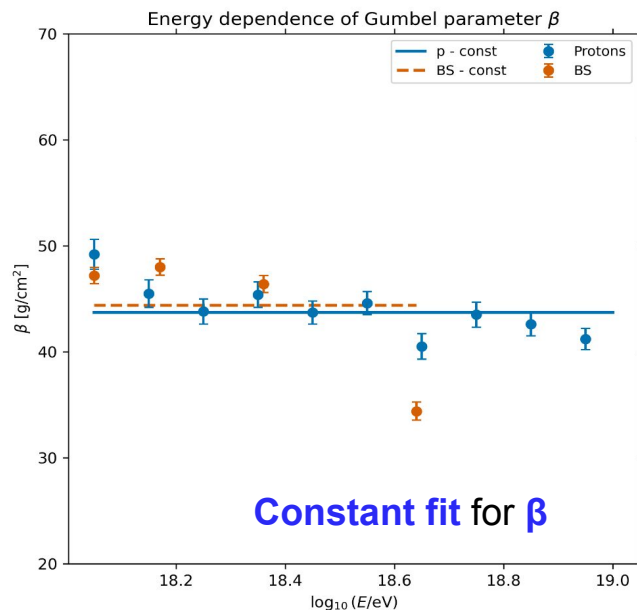
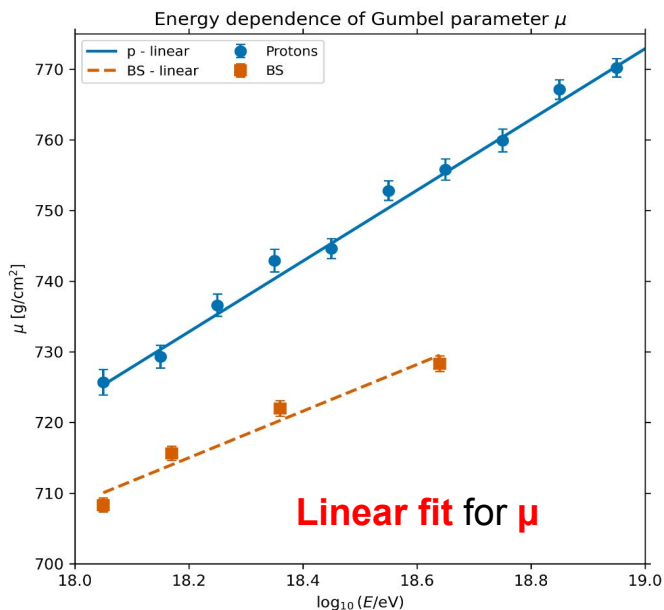
The energy dependence lies in the scale and location parameters of the distribution:

$$\mu = \mu(x) = mx + q \quad \beta = \beta(x) = \text{const} \longrightarrow x = \log_{10}(E)$$

Energy-dependent Gumbel

Fit Gumbel functions to the $X_{\max}$ distributions in different energy bins to infer parameter trend

- $\Delta \log(E / \text{eV}) = 0.1$ bin width for background **simulations**
- **Optimal bin** width for **BS** case



Photon trends
in backup!

$X_{\max}$ as analysis variable

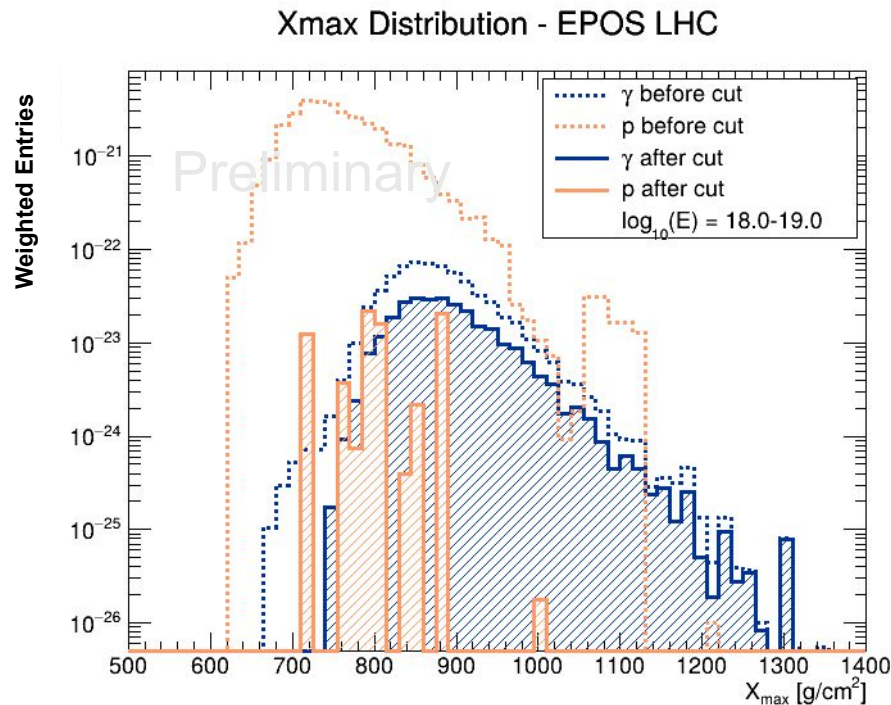
As expected the selection yields:

- **High signal efficiency**
- **Low background contamination**

Now $X_{\max}$ is **optimised** and can be used in the analysis

Additional validation: $X_{\max}$ is directly measured and its distribution is well known

Here the resulting distributions for
the whole energy range



$X_{\max}$ as analysis variable

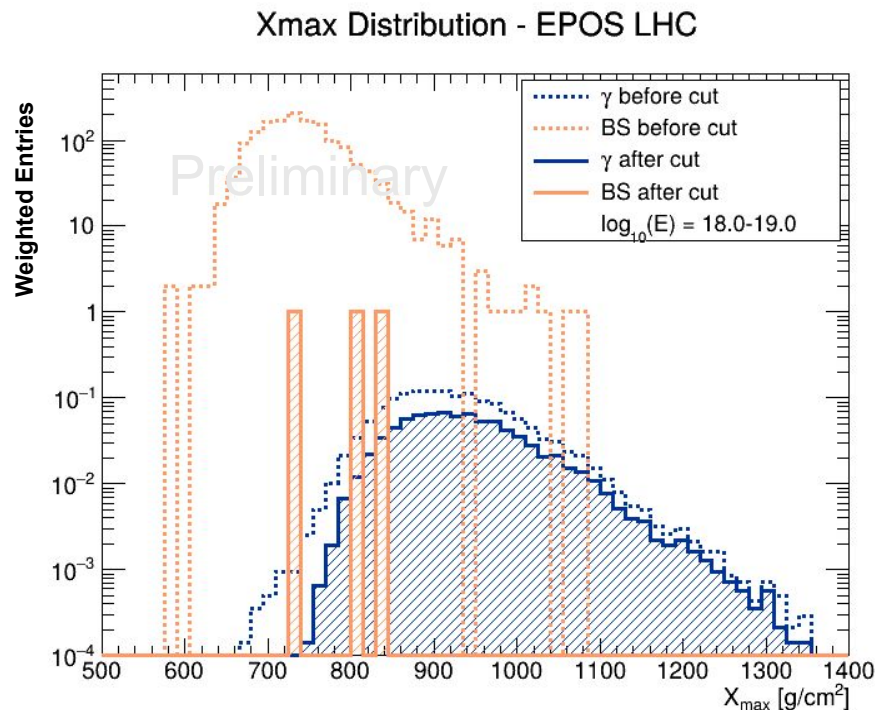
As expected the selection yields:

- **High signal efficiency**
- **Low background contamination**

Now $X_{\max}$ is **optimised** and can be used in the analysis

Additional validation: $X_{\max}$ is directly measured and its distribution is well known

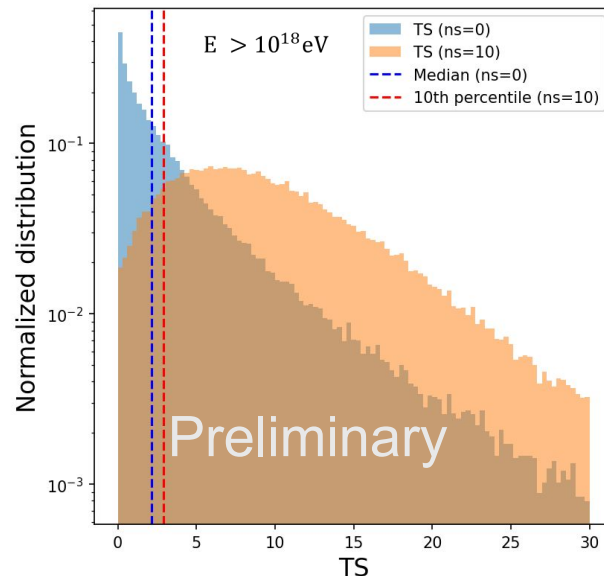
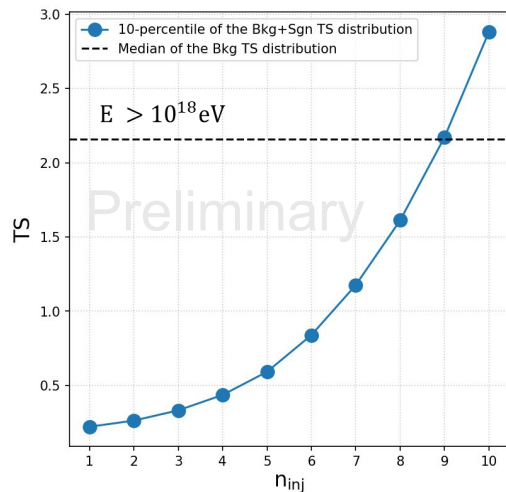
Here the resulting distributions for
the whole energy range



Sensitivity estimation

To determine the sensitivity:

- Compute **median of the bkg** TS distribution (H_0)
- Compute **10-percentile of each sig+bkg** TS distribution (H_1)

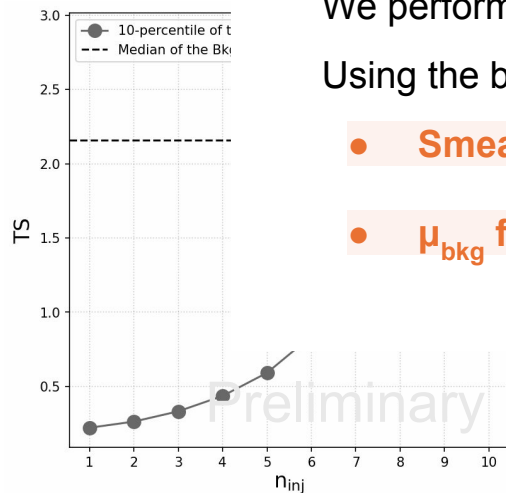


Analysis sensitivity: the n_{inj} whose **10-percentile** matches the **median**
Used to **estimate** the 90% confidence level (CL) **upper limits** (UL)

Sensitivity estimation

To determine the sensitivity:

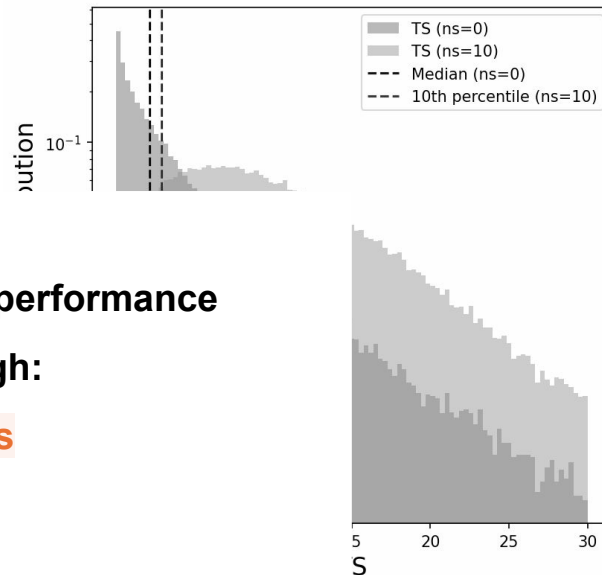
- Compute **median of the bkg** TS distribution (H_0)
- Compute



We performed different tests to select the **best performance**

Using the background BS, it is **obtained through**:

- **Smearing of bkg generation parameters**
- **μ_{bkg} fixed under both hypotheses**



Analysis sensitivity: the n_{inj} whose 10 percentile matches the median
Used to **estimate** the 90% confidence level (CL) **upper limits** (UL)

Exposure calculation

We have not applied the analysis to the data yet, but **we can evaluate the improvements on the ULs**

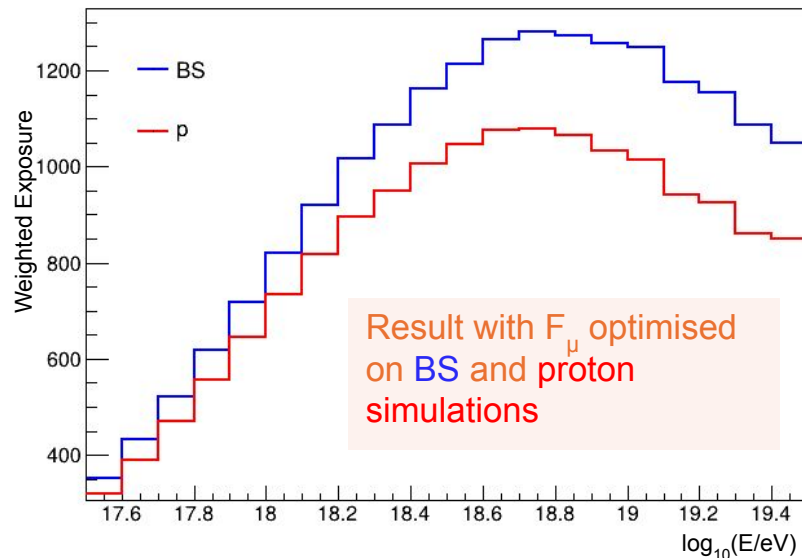
To extrapolate them we need the **exposure**

It is obtained through different steps:

- **γ air shower simulations:** shower library
- **Detector simulation:** generated events
 - Actual detector status
 - Detector time evolution
 - Atmospheric conditions
- **Analysis cuts:** selected events
 - Quality selection cuts
 - F_μ selection cut

$$\mathcal{E}_\gamma = \int dt \int d\Omega \cos \theta \int dS \mathcal{E}_\gamma(E_\gamma, t, \theta, \phi, x, y)$$

$$\mathcal{E}_\gamma^{\text{weighted}}(E_\gamma > E_\gamma^0) = \frac{\int_{E_\gamma^0}^{\infty} E_\gamma^{-\Gamma} \mathcal{E}_\gamma(E_\gamma) dE_\gamma}{\int_{E_\gamma^0}^{\infty} E_\gamma^{-\Gamma} dE_\gamma}$$



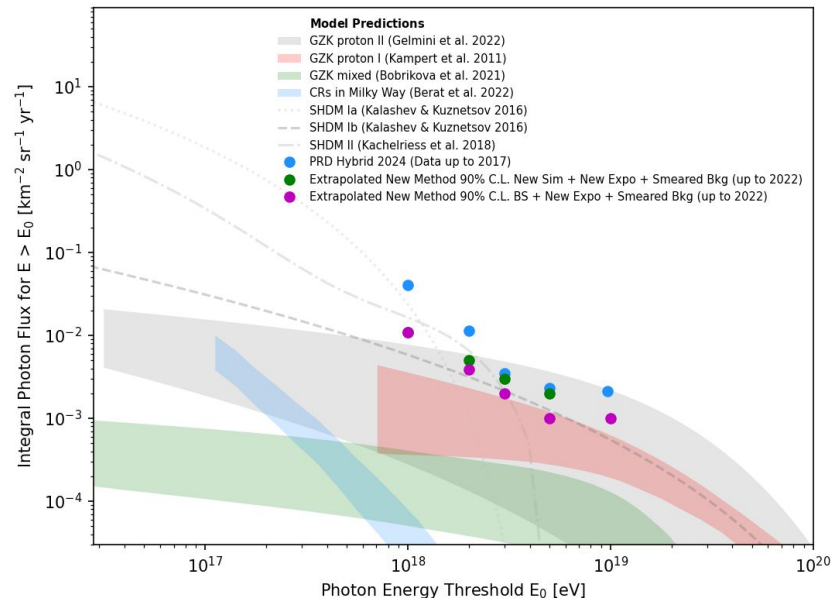
ULs extrapolation

- Old approach (data up to 2017)
- New approach with simulations
- New approach with BS = better performance

Improvement:

from factor ~ 2 at higher energies up to ~ 4 at lower energy with BS

The new method, jointly with the optimisation, allows for higher-sensitivity



Conclusions

The **change of paradigm** in photon searches proposed can lead to better results

It involves:

- **Optimizing the signal-to-noise ratio** on the observables
- Applying a **likelihood-ratio hypothesis test**

better performance than previous methods such as that of [PRD 110, 062005 (2024)]

Ongoing work

The use of a **data-driven approach increases the exposure** (less stringent F_{μ}^{cut}), next step is:

- Test the method on bigger control sample
- Systematic uncertainties estimation

Future perspectives...

- Move to targeted search: background reduction
- Analysis of the results/implications of previous steps



Production and experiences

Conferences and collaboration meetings

- CRIS-MAC, Sorrento, Italy, 13–17 July 2026. Talk: *Searching for UHE photons in background-dominated data*.
- ECRS, Santiago de Compostela, Spain, 29 June–3 July 2026. Talk on behalf of The Pierre Auger Collaboration: *Photon searches at the Pierre Auger Observatory*.
- Pierre Auger Collaboration meeting, Malargüe, Argentina, 16–21 November 2025. Talk: *Updating hybrid photon analysis with Phase I data*.
- Pierre Auger Collaboration meeting, Malargüe, Argentina, 22–27 March 2026. Online participation.
- Presentations at the regular monthly calls within the Pierre Auger Collaboration

Attended schools

- Scientific Communication in Astronomy School, Bertinoro, Italy, 12–17 October 2025.
- NBIA Summer School on Neutrinos: Here, There and Everywhere, Copenhagen, Denmark, 7–11 July 2025.

Outreach

- Sharper, European Research Night & Street Science, 25th September 2026, L'Aquila.
- Visitor Centre GSSI, tour guide, September 2026, L'Aquila.
- Street Science, 26th September 2025, L'Aquila.

Scientific publications

As member of the Pierre Auger Collaboration, since February 2026 I am a co-author of publications:

- The Pierre Auger Collaboration, *Long-term monitoring of the optical performance of the primary mirrors at the Coihueco and Loma Amarilla sites of the Pierre Auger Observatory*, JINST **21** (2026) no.06, P06040. doi: 10.1088/1748-0221/21/06/P06040.
- The Pierre Auger Collaboration, *Bounds on Lorentz Invariance Violation from Muon Fluctuations at the Pierre Auger Observatory*, Phys. Rev. Lett. **137** (2026) no.11, 111001. doi: 10.1103/t5k4-2m32.
- The Pierre Auger Collaboration, *Depth of maximum of air-shower profiles above $10^{17.7}$ eV measured with the fluorescence detector of the Pierre Auger Observatory*, Phys. Rev. D **114** (2026) no.4, 043016. doi: 10.1103/n616-15v5.

Conference proceedings

- C. Marinelli for The Pierre Auger Collaboration, *Photon searches at the Pierre Auger Observatory*, proceedings of CRIS-MAC 2026. *Submitted*.
- C. Marinelli, C. Evoli, P. Savina, *Searching for UHE photons in background-dominated data*, proceedings of ECRS 2026. *Submitted*.

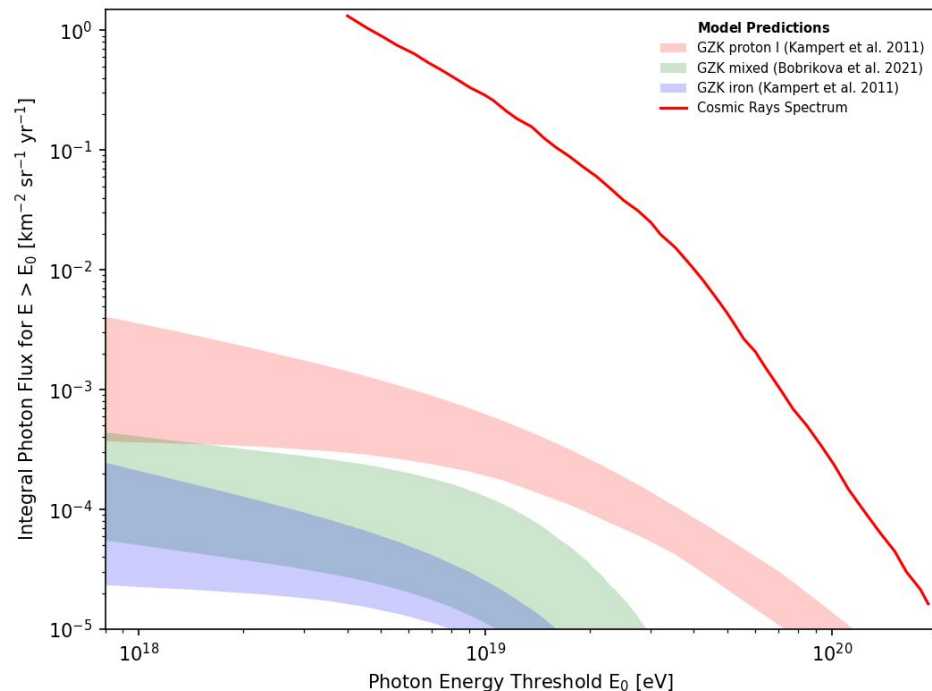
Internal collaboration notes

- C. Marinelli, P. Savina and C. Evoli, *Updates on hybrid photon search above 1 EeV: a signal-to-background driven optimization for a likelihood-based analysis*, (to be published October 2026).
- C. Marinelli, P. Savina and C. Evoli, *Hybrid photon search above 1 EeV: a signal-to-background driven optimization for a likelihood-based analysis*, April 2026.

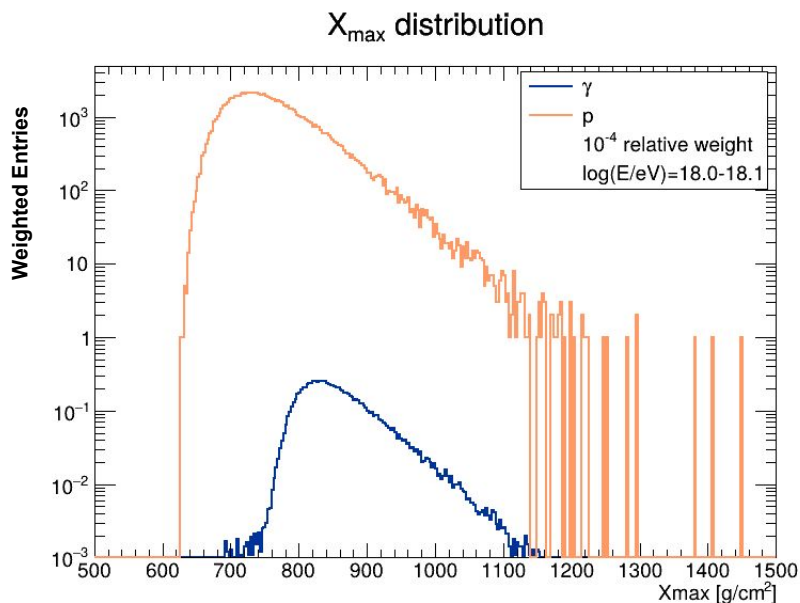
Backup slides

Needle in a haystack

Reality: the background is dominant



Powerful discrimination but low expected number of photon events



Simulation weights

The **simulated events** are **weighted** to mimic realistic fluxes

Protons are weighted using the **all-particle differential flux** from [\[https://link.springer.com/article/10.1140/epjc/s10052-021-09700-w\]](https://link.springer.com/article/10.1140/epjc/s10052-021-09700-w)

$$J(E, s) = J_0 \left(\frac{E}{E_0} \right)^{-\gamma_0} \frac{\prod_{i=0}^3 \left[1 + \left(\frac{E}{E_{ij}} \right)^{\frac{1}{\omega_{ij}}} \right]^{(\gamma_i - \gamma_j)\omega_{ij}}}{\prod_{i=0}^3 \left[1 + \left(\frac{E_0}{E_{ij}} \right)^{\frac{1}{\omega_{ij}}} \right]^{(\gamma_i - \gamma_j)\omega_{ij}}},$$

Where $E_0 = 10^{18.5} \text{ eV}$

Photons are weighted with:

$$J_{ph} = \phi_0(E_0^{ph}) \left(\frac{E}{E_0^{ph}} \right)^{-2}$$

$$\phi_0(E_0^{ph}) = 10^{-3} \cdot J(E_0^{ph}) = 10^{-3} \cdot J(10^{17.5} \text{ eV})$$

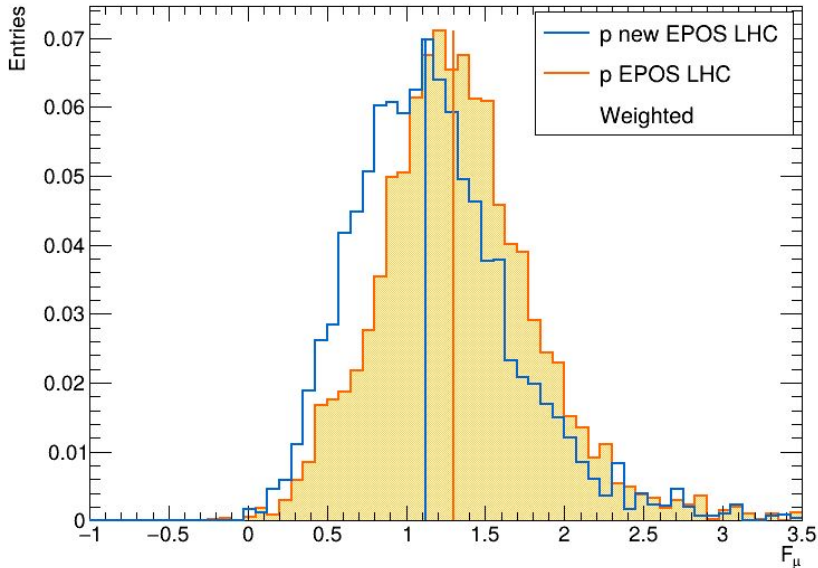
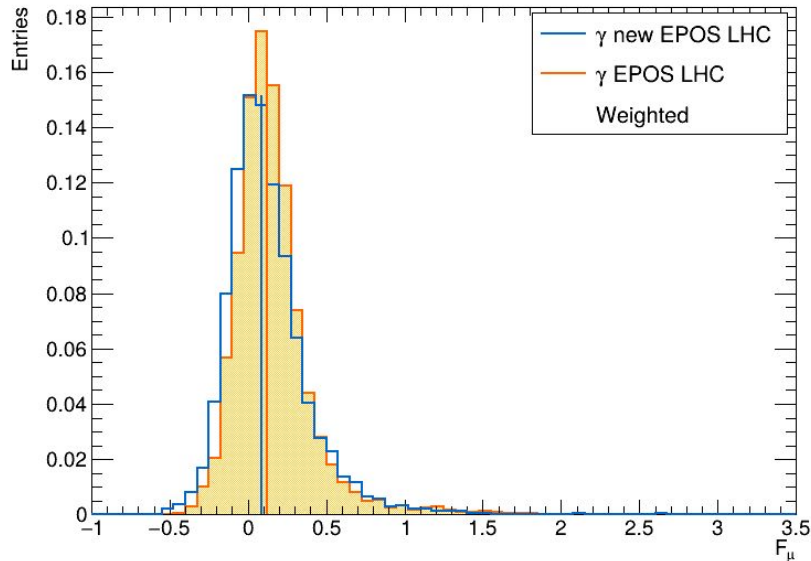
Parameter	Value $\pm \sigma_{\text{stat}} \pm \sigma_{\text{syst}}$
$J_0 / (\text{km}^2 \text{ yr sr eV})$	$(1.309 \pm 0.003 \pm 0.400) \times 10^{-18}$
ω_{01}	$0.43 \pm 0.04 \pm 0.34$
γ_1	$3.298 \pm 0.005 \pm 0.10$
E_{12}/eV	$(4.9 \pm 0.1 \pm 0.8) \times 10^{18}$
γ_2	$2.52 \pm 0.03 \pm 0.05$
E_{23}/eV	$(1.4 \pm 0.1 \pm 0.2) \times 10^{19}$
γ_3	$3.08 \pm 0.05 \pm 0.10$
E_{34}/eV	$(4.7 \pm 0.3 \pm 0.6) \times 10^{19}$
γ_4	$5.2 \pm 0.2 \pm 0.1$
γ_0	2.64 – fixed
E_{01}/eV	1.24×10^{17} – fixed
ω_{12}	0.05 – fixed
ω_{23}	0.05 – fixed
ω_{34}	0.05 – fixed

F_μ distributions: model comparison

Initial **old EPOS-LHC** simulation model vs **new EPOS-LHC** version

The two show **differences**:

New proton distribution shifted to the left → **Higher background contamination**



$X_{\max}$ as analysis variable

As expected the result is:

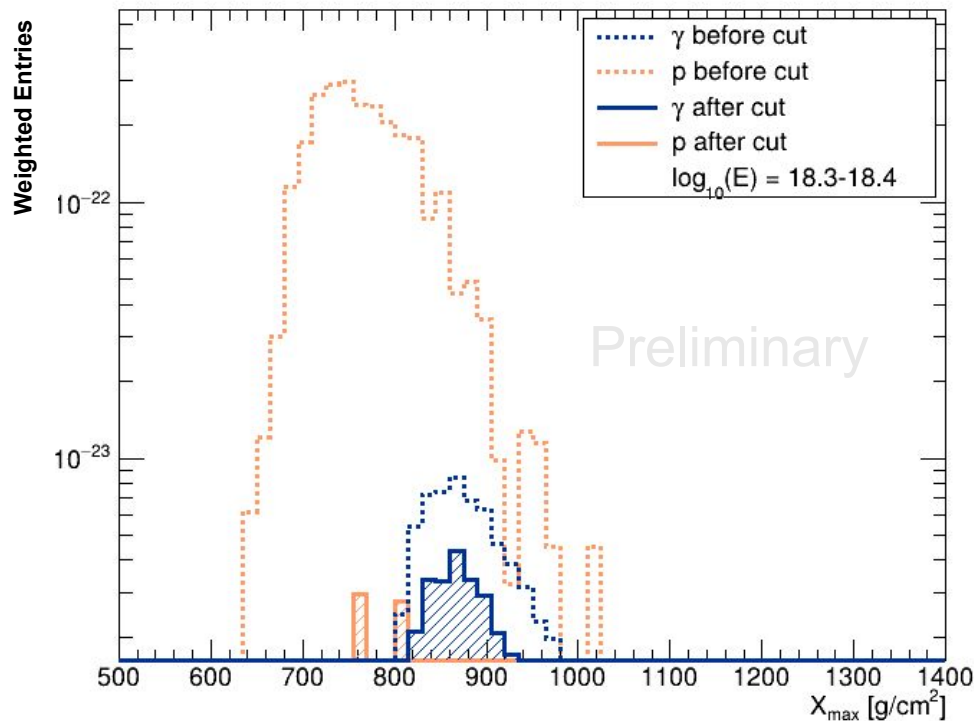
- **High signal efficiency**
- **Low background contamination**

Now $X_{\max}$ is **optimised** and can be used in the analysis

Additional validation: $X_{\max}$ is directly measured and its distribution is well known

Here the resulting distribution for $\log(E / \text{eV}) \in [18.3, 18.4]$

Xmax Distribution - EPOS LHC

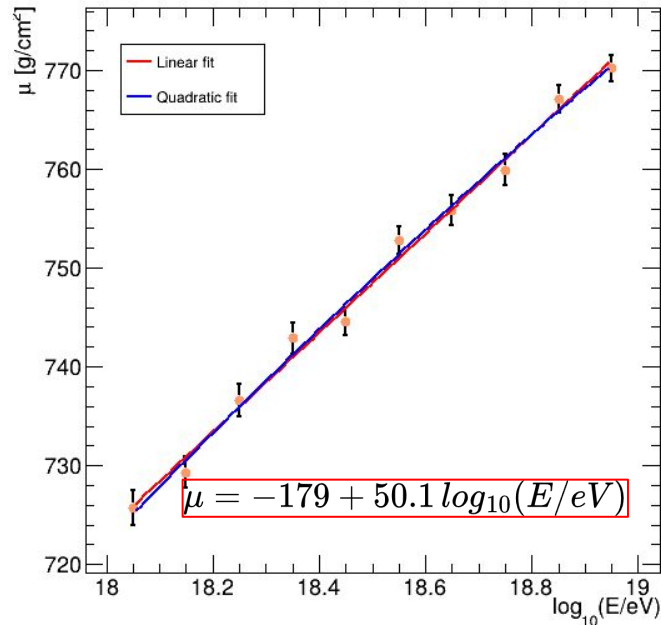


Energy-dependent Gumbel: protons

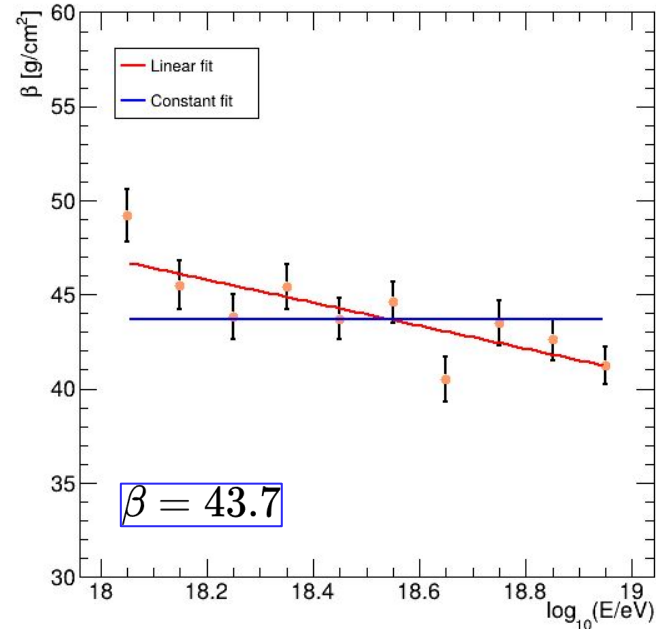
We fit Gumbel functions to the $X_{\max}$ distributions in different energy bins

$\Delta \log(E / \text{eV}) = 0.1$ bin width considered to infer the trends of the gumbel parameters

Linear and quadratic **fit** are compared for μ



Linear and **constant fit** are compared for β

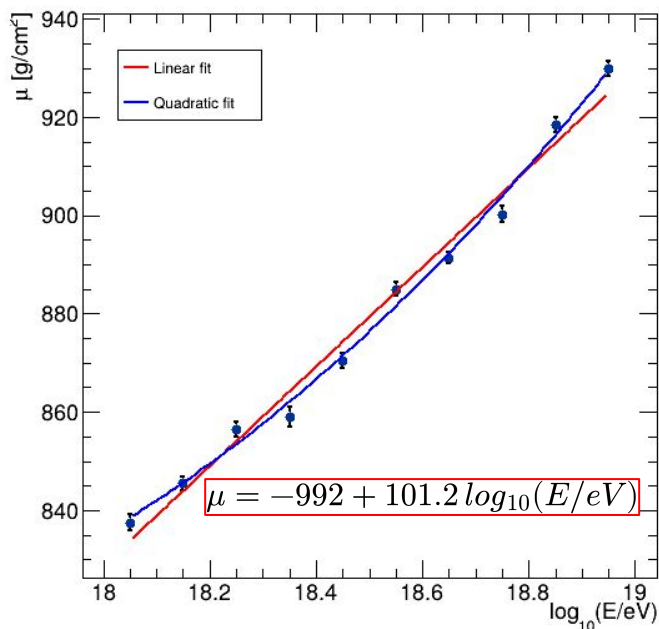


Energy-dependent Gumbel: photons

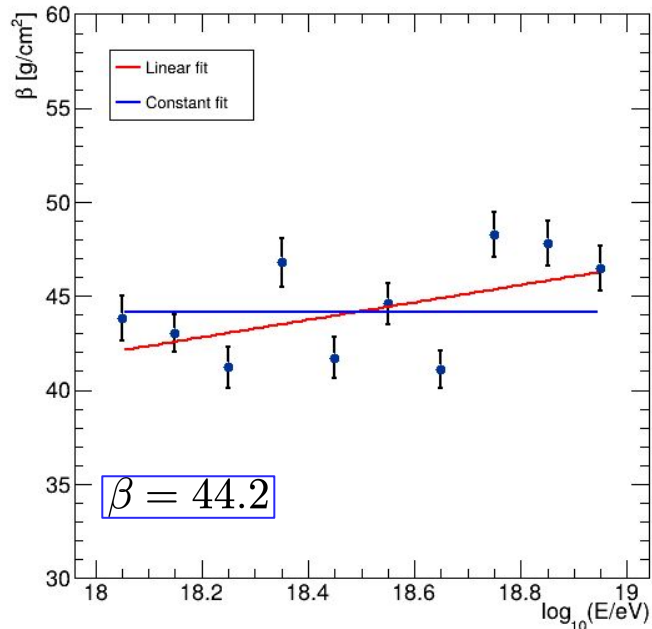
We fit Gumbel functions to the $X_{\max}$ distributions in different energy bins

$\Delta \log(E / \text{eV}) = 0.1$ bin width considered to infer the trends of the gumbel parameters

Linear and quadratic fit are compared for μ



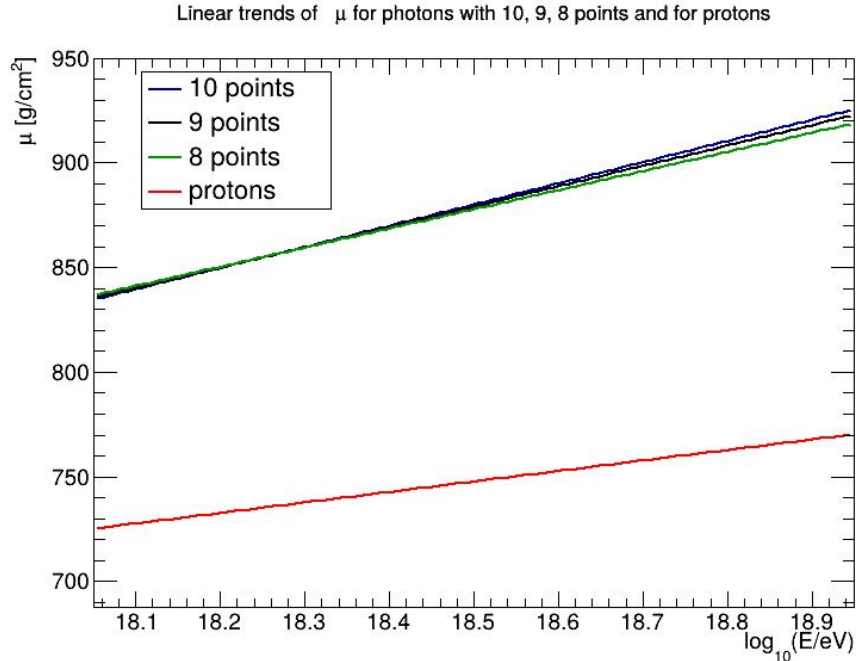
Linear and constant fit are compared for β



Energy-dependent Gumbel: photons

At **higher energies** photons are affected by **LPM and pre-showering effects**:

Different tests excluding highest energy points to choose most conservative option



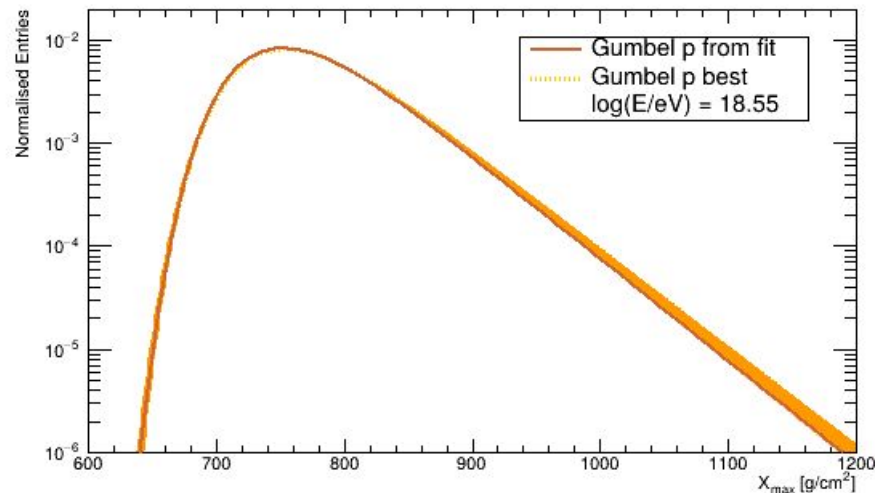
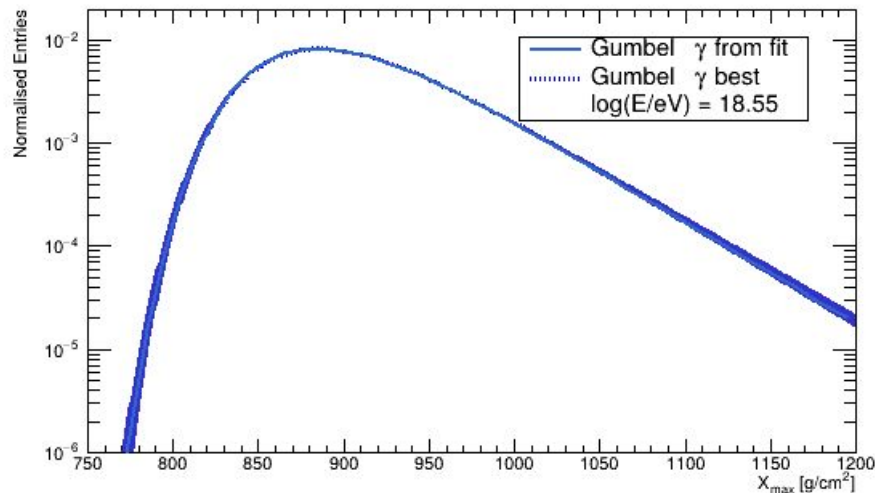
In reality they are equivalent

Gumbel parameters: trend validation

Bin-by-bin comparison between:

Original Gumbel distribution, $\Delta \log(E / \text{eV}) = 0.1$ bin width (best)

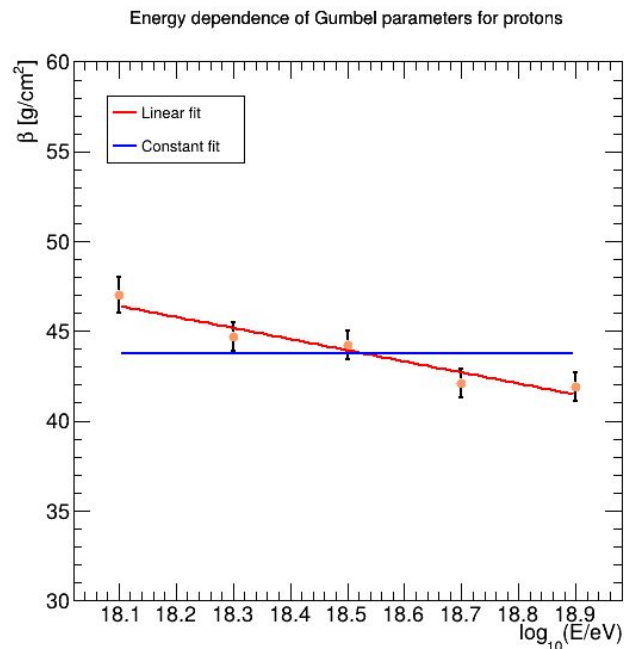
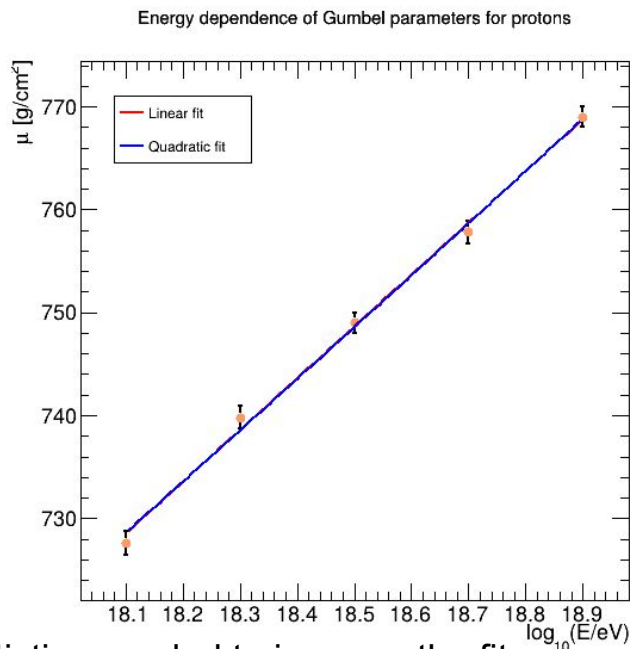
Derived Gumbel distribution, parameters inferred from the trends (from fit)



Gumbel parameters: trend validation

Trends of the gumbel parameters: $\Delta \log(E / \text{eV}) = 0.2$ bin width

Linear and quadratic fit are compared for μ Linear and **constant** fit are compared for β

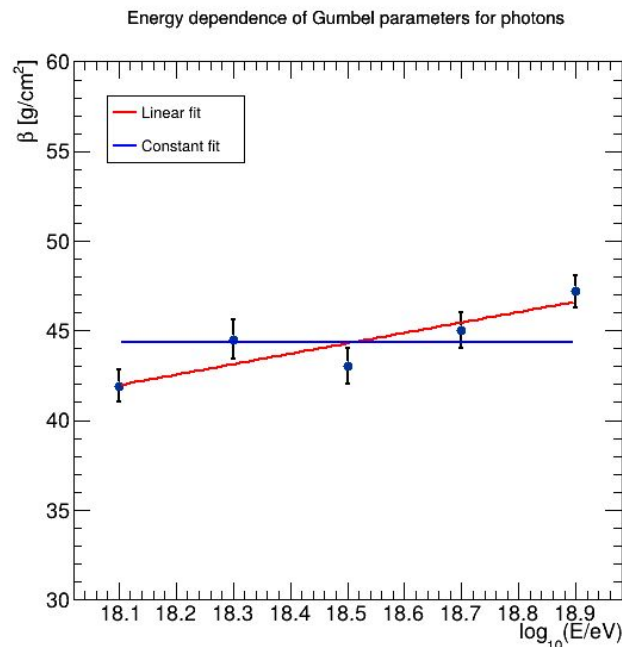
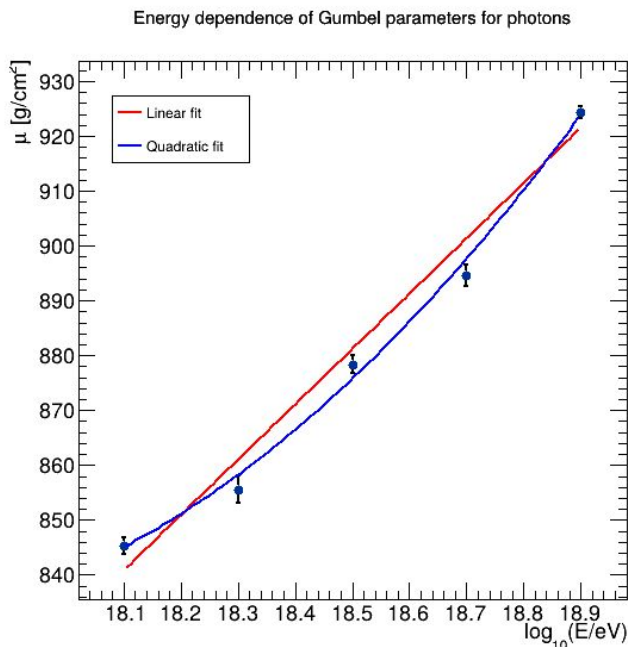


More statistics needed to improve the fit

Gumbel parameters: trend validation

Trends of the gumbel parameters: $\Delta \log(E / \text{eV}) = 0.2$ bin width

Linear and quadratic fit are compared for μ Linear and **constant** fit are compared for β



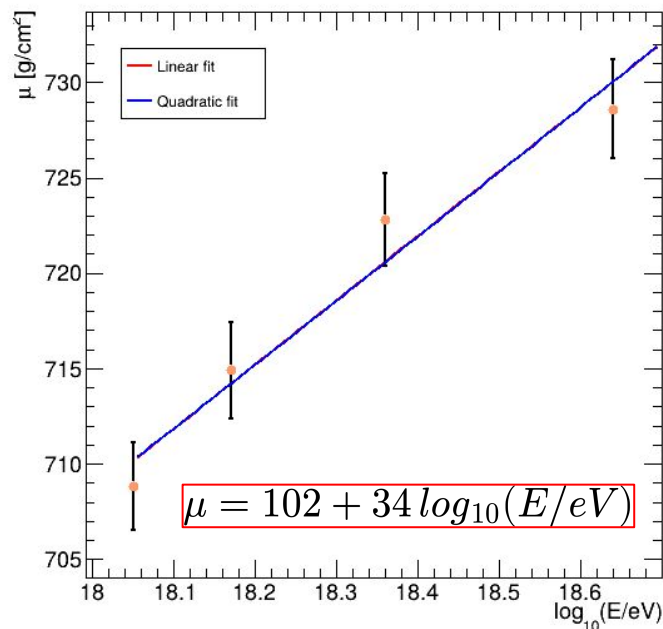
More statistics needed to improve the fit

Energy-dependent Gumbel: BS

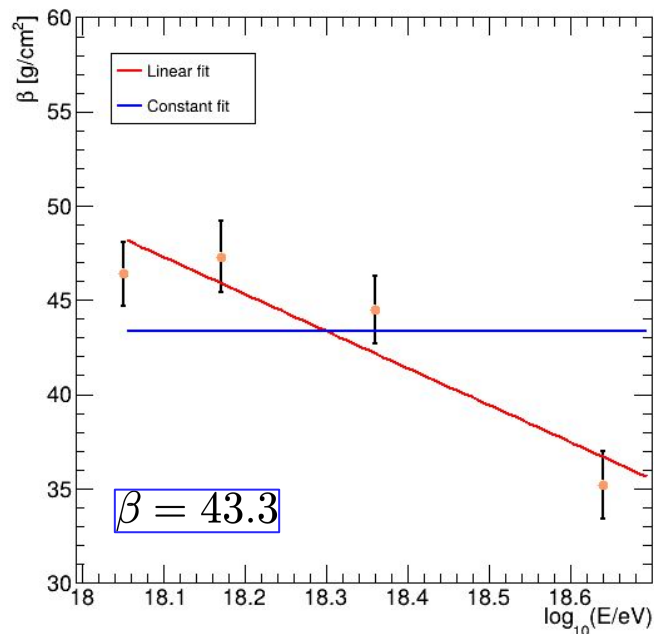
We fit Gumbel functions to the $X_{\max}$ distributions in different energy bins

4 optimal bins to get **about the same number of events**

Linear and quadratic **fit** are compared for μ



Linear and **constant fit** are compared for β

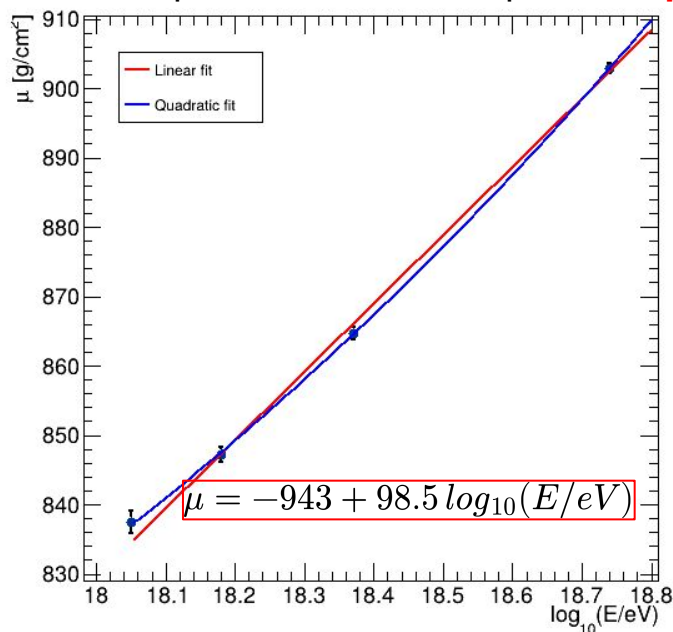


Energy-dependent Gumbel: photons

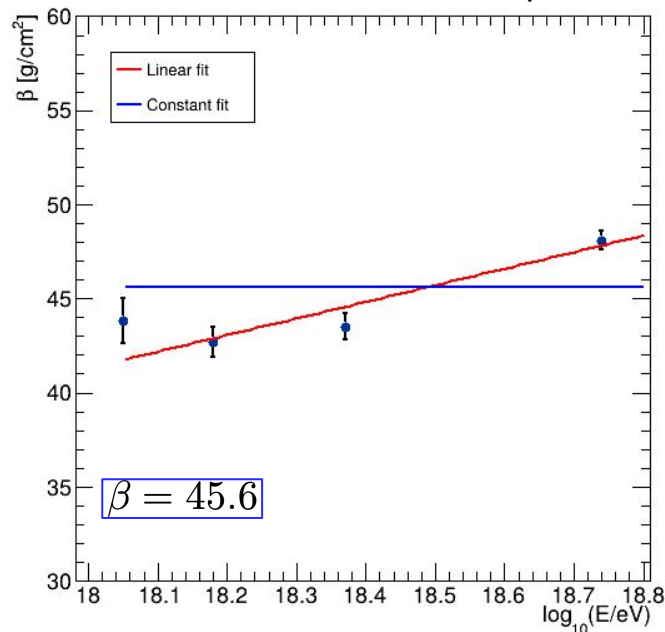
We fit Gumbel functions to the $X_{\max}$ distributions in different energy bins

4 optimal bins to get about the same number of events

Linear and quadratic **fit** are compared for μ



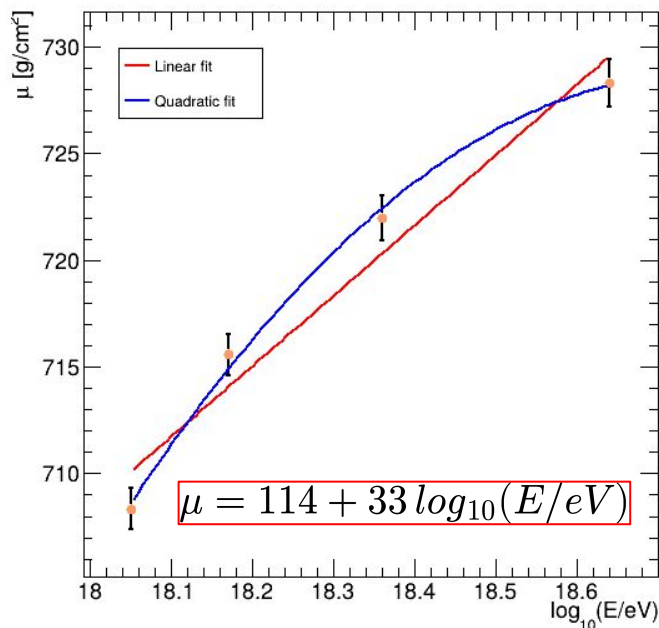
Linear and **constant fit** are compared for β



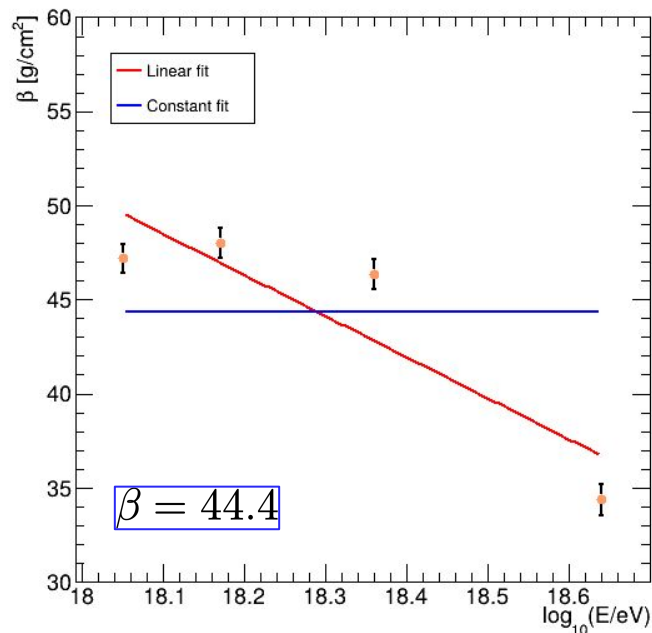
Energy-dependent Gumbel: BS 30%

To improve the uncertainty on the trend coefficients we perform the fit on an **extrapolated 30% BS**
4 optimal bins to get about the same number of events

Linear and quadratic **fit** are compared for μ



Linear and **constant fit** are compared for β



Sensitivity estimation: TS distribution

Background TS distribution:

- **100k pseudo-experiment** simulations under H_0 ($n_s = 0$)

Background and signal parameterised as **Gumbel**

Number of simulated events matching the expected number of background events

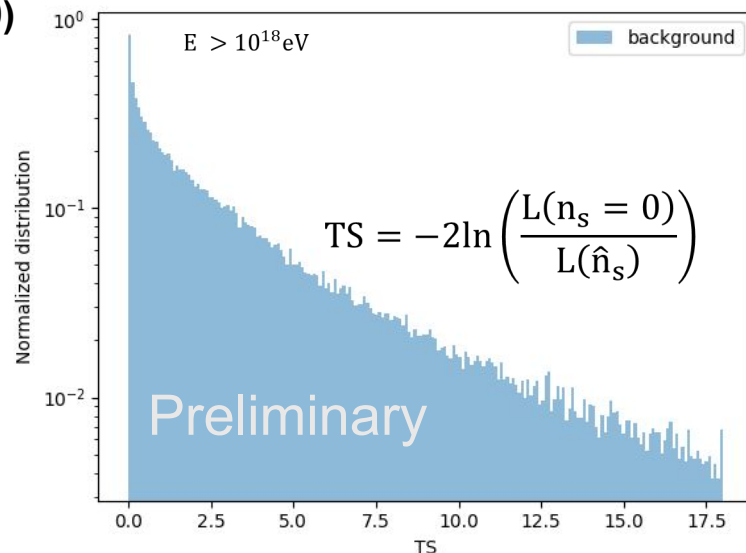
- Each pseudo-experiment is **fitted with null and alternative hypotheses**
- **TS distribution** built

Signal TS distribution:

- **Multiple 100k pseudo-experiments** under H_1 ($n_s \neq 0$)

Each with a **different** n_s injected (n_{inj})

- **TS distribution** built



Sensitivity estimation: tests

The same procedure is repeated under different conditions to define the (nominal) best configuration

Parameters available: $\beta_{\text{sig}}, \beta_{\text{bkg}}, \mu_{\text{sig}}, \mu_{\text{bkg}}, n_s, N$

Basis: n_s always fixed under H_0 , N always fixed

Fitting tests:

- Nothing fixed
- Sig & Bkg fixed under H_0 + Bkg fixed under H_1
- Same as before with β and μ trend from validation tests (also in generation)
- Sig fixed under H_0 + Bkg fixed under H_1
- β_{bkg} fixed under both hypotheses
- **μ_{bkg} fixed under both hypotheses**

Generation test:

- **Smearing of generation parameters**

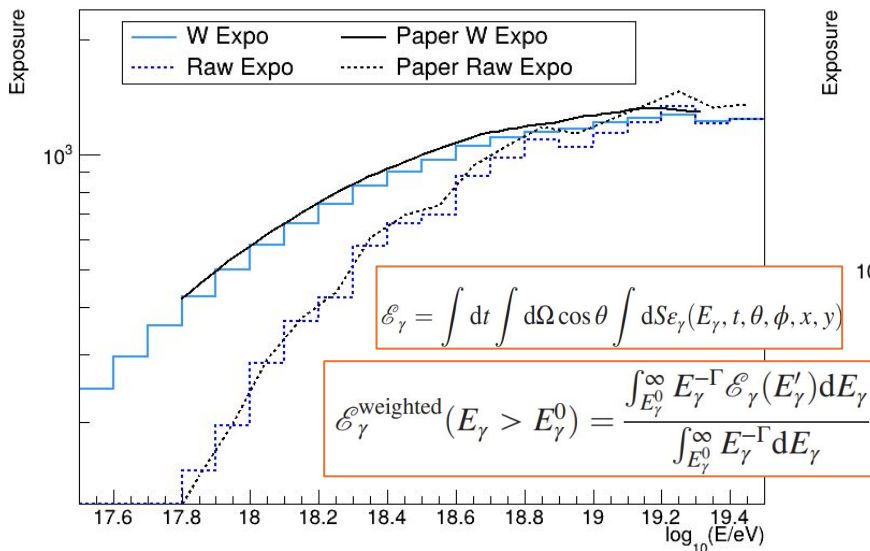
Best (nominal) **configuration**
Using **extrapolated 30% BS**

Exposure calculation

To determine the new upper limits (UL) we need the **exposure**

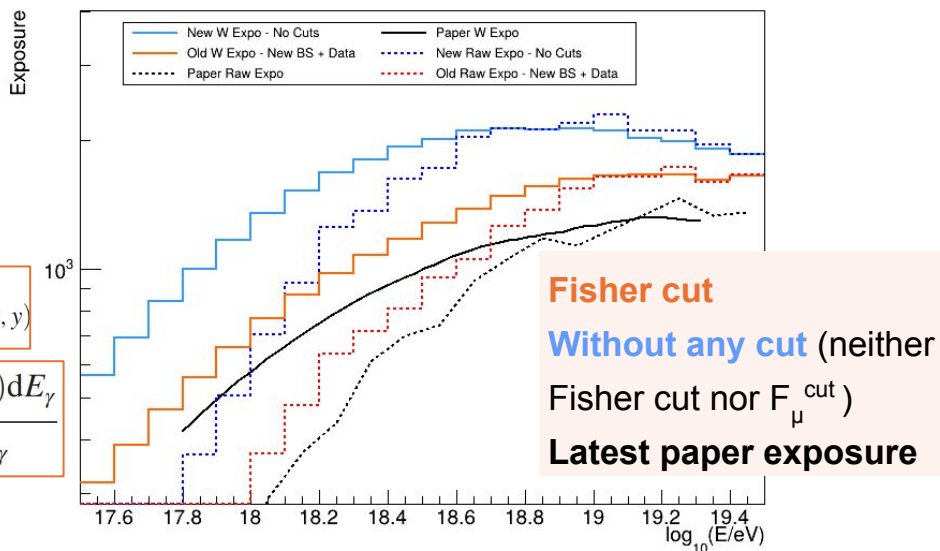
Step 1 :

- We verified that the **old Exposure** could be **reproduced**:



Step 2:

- Compute the exposure accounting for **BS change** and **temporal change**



Exposure calculation: new result

Finally the **NEW exposure** is derived

- **Entire new Real MC simulation done**

New events included

F_{μ}^{cut} on the best cut value considered

30% BS up to 2017 + 5% BS (2018 - 2022)

5% BS up to 2022

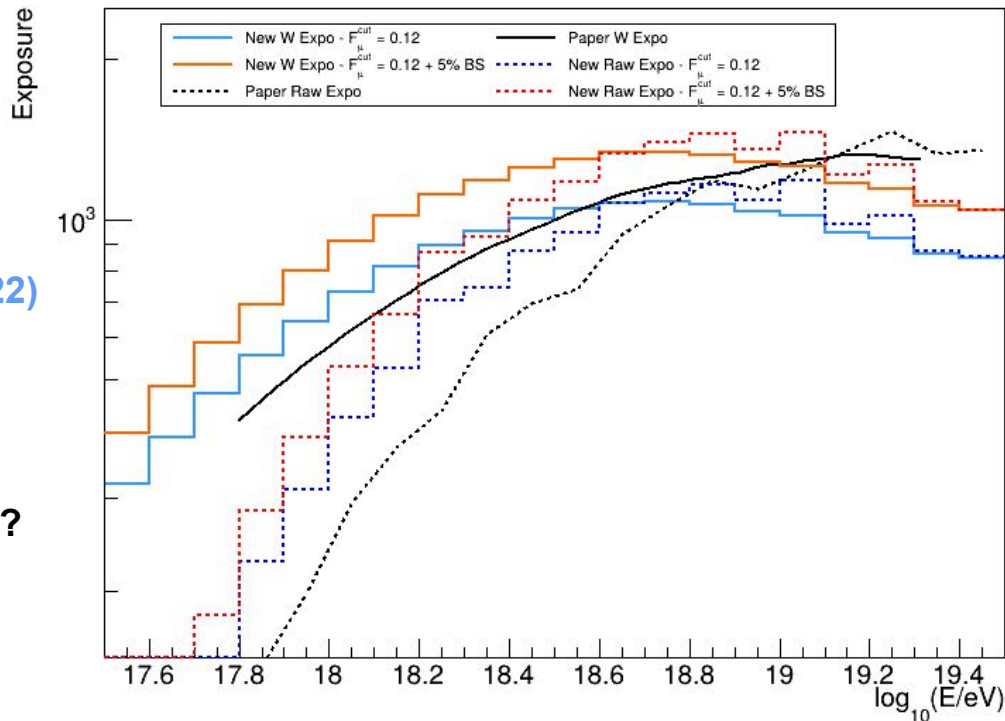
Why the decrease at higher energies?

Why the change of shape at higher energies?

Temporal component?

BS component?

No Fisher cut nor F_{μ}^{cut} ?



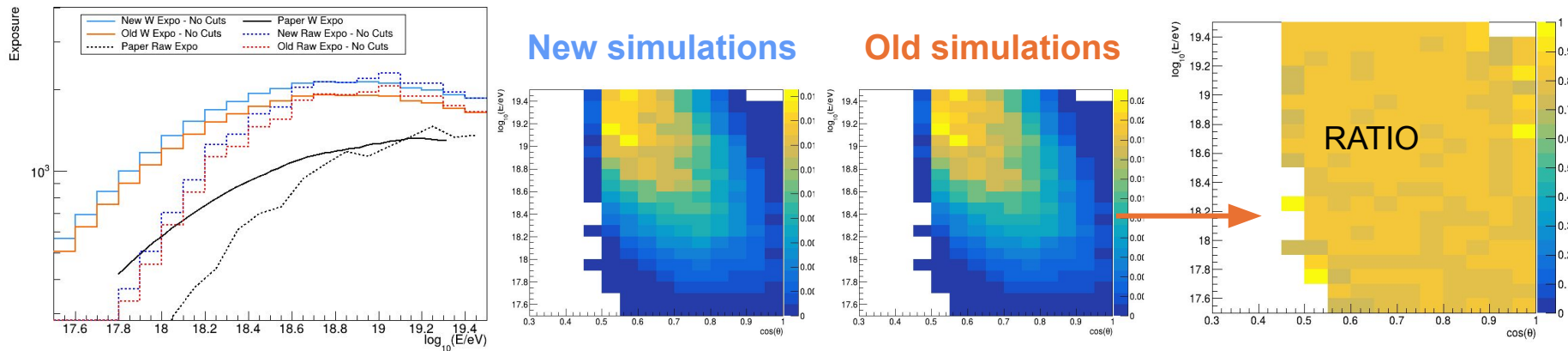
Exposure calculation: shape

Why the different shape at higher energies?

We only consider the change in the **temporal component**: **same calculation, without any cuts**, different simulation sets:

We use **old simulations (up to 2017)**, accounting for **5% BS**

We use **new simulations (up to 2022)**, accounting for **30% BS up to 2017 + 5% BS from 2017 to 2022**



We compare the **efficiency plots** (ratio of selected vs total number of events)

- ➔ **No evident change in the trend**: as far as the calculation is equal there are no changes in the exposure
- ➔ The **flattening at higher energies of the old exposure is due to the Fisher median cut**

Exposure calculation: decreasing trend

Why the decrease at higher energies?

The **value of the cut is hadronic model dependent**

New simulations: greater contamination bkg-signal → harder cut than the one before

- **Our evaluation is affected by the use of proton simulations** and not of a data-driven approach

A data-driven approach:

- **Provides** us a **lower contamination** (and consequently a **higher exposure**)
- **Affect** the n_s calculation → TS distributions are currently generated considering proton simulations

Therefore **we expect an even lower sensitivity**

And the **reduction of systematics related to the hadronic model**

AugerPrime: Observatory upgrade

Hybrid ground-based detector update:

Additional detector systems have recently been installed to complement the main SD:

Radio antennas (RD)

Scintillator planes (SSD)

Underground muon detectors (UMD)

Multi-hybrid events are collected with 100% duty cycle

Different shower components can be separated in different ways:

- WCD/SSD/RD → events up to $\sim 60^\circ$
- WCD/RD → inclined events ($> 60^\circ$)
- WCD/SSD/UMD → mass sensitivity extended to lower energies
+ improved photon/hadron separation

New electronics → enables **sensitivity enhancement** of triggers to electromagnetic signals: **photons and neutrinos**

