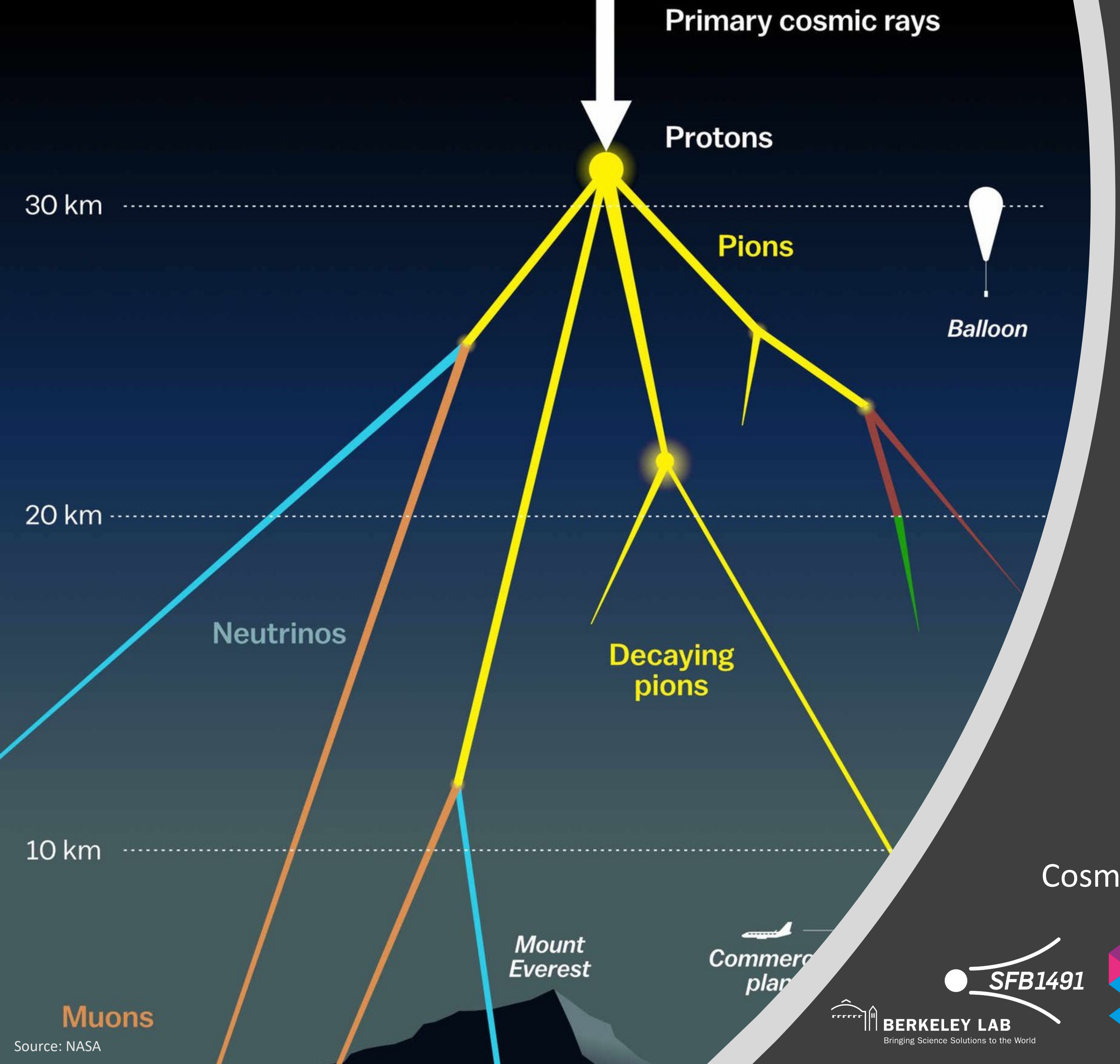


tu dortmund university

ICECUBE NEUTRINO OBSERVATORY

Unfolding the Atmospheric Muon Flux with IceCube

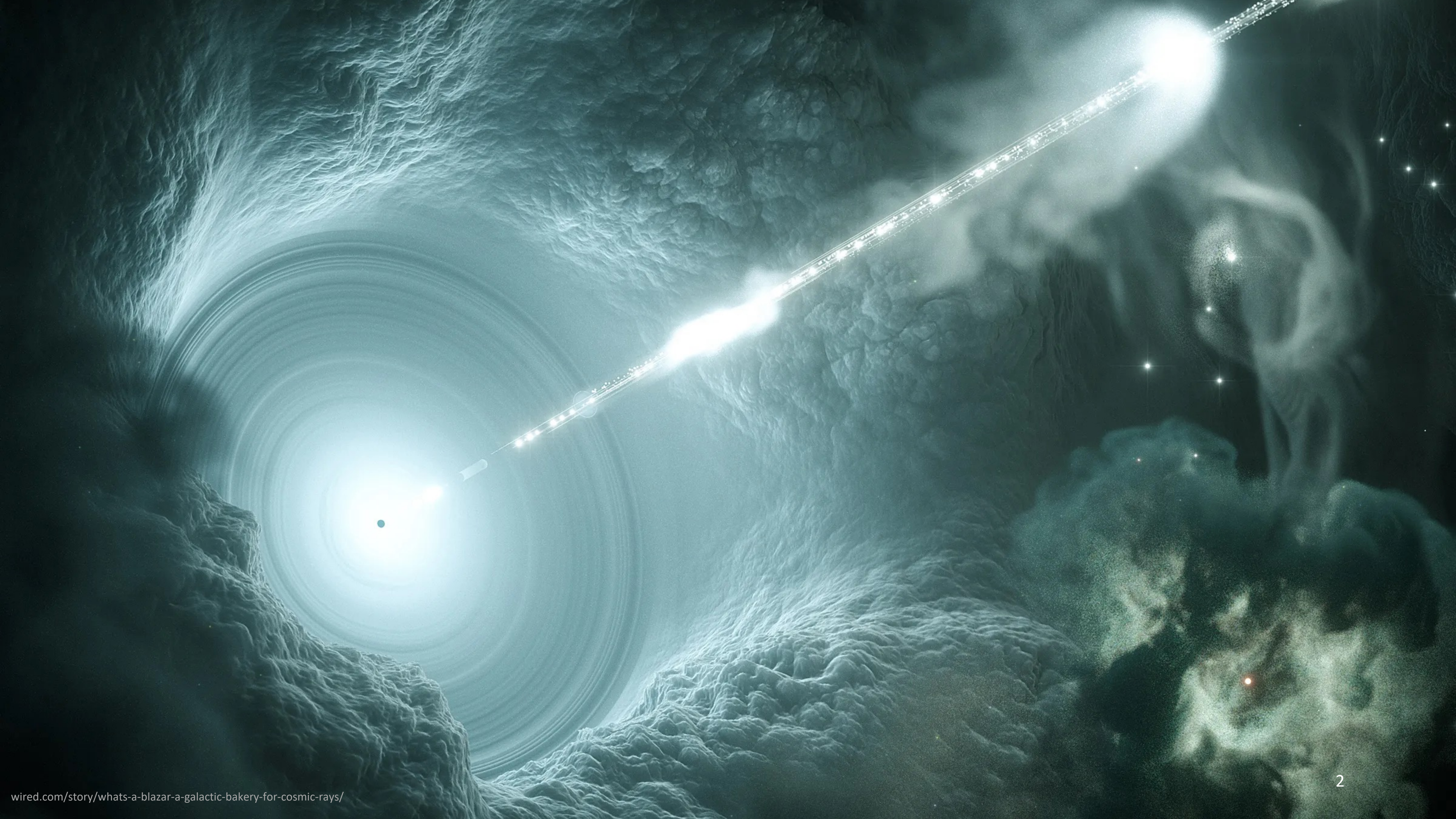
Pascal Gutjahr
for the IceCube Collaboration

Cosmic Ray physics with high-energy atmospheric muons

SFB1491 BERKELEY LAB Bringing Science Solutions to the World LAMARR INSTITUTE FOR MACHINE LEARNING AND ARTIFICIAL INTELLIGENCE NSF

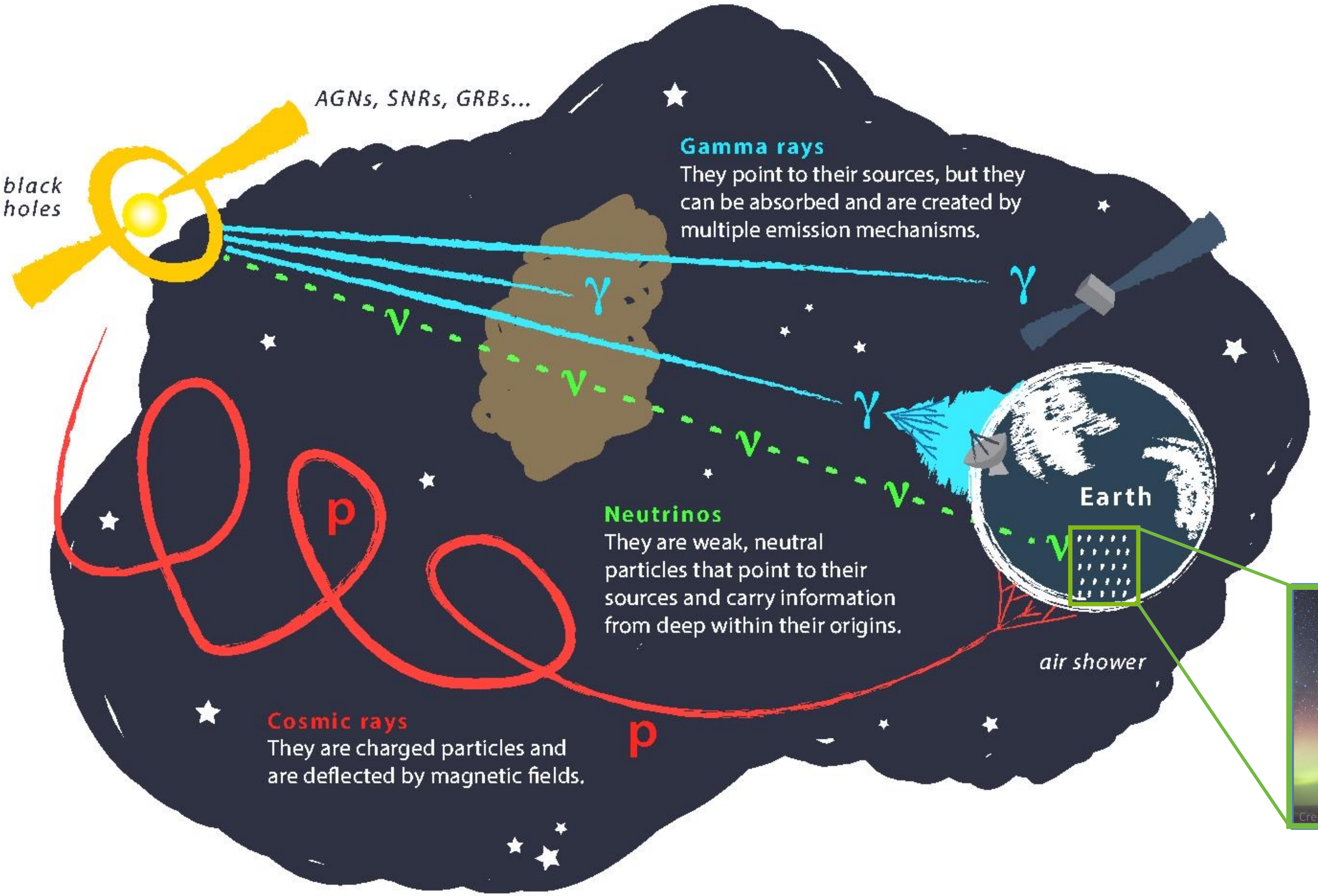
LPC Caen

September 17-19, 2025

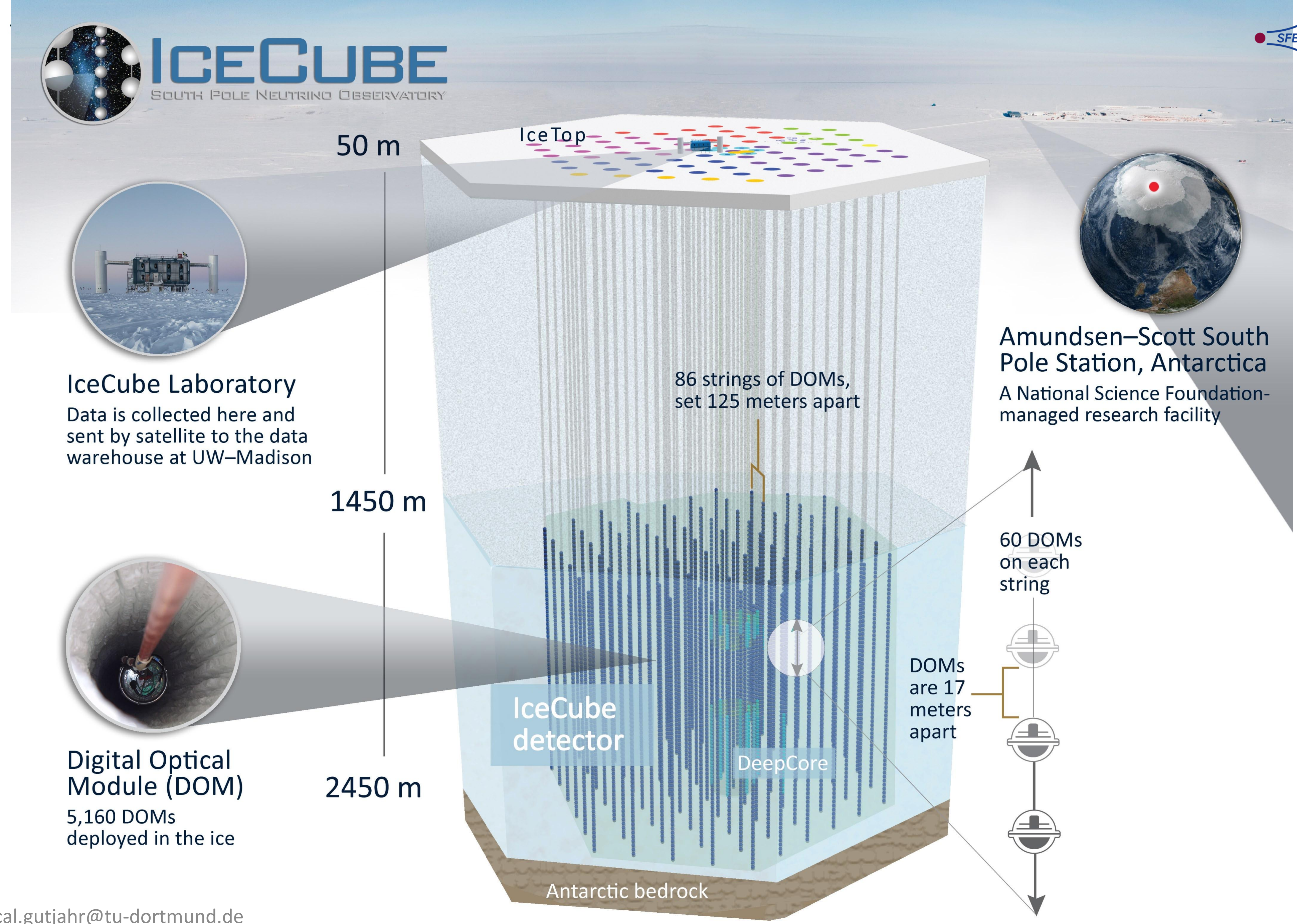


wired.com/story/whats-a-blazar-a-galactic-bakery-for-cosmic-rays/

Astroparticle physics



Source: gallery.icecube.wisc.edu



Credit: IceCube

Introduction

Motivation

- Characterize muon flux components and depth intensity
- Probe hadronic interactions in the atmosphere
- Constrain uncertainties and longstanding questions like the "muon puzzle"

Long term

- Enhance understanding of hadronic processes relevant in distant astrophysical sources (e.g., supernovae, AGNs, …)

Conventional Muon:
Parent is pion or kaon

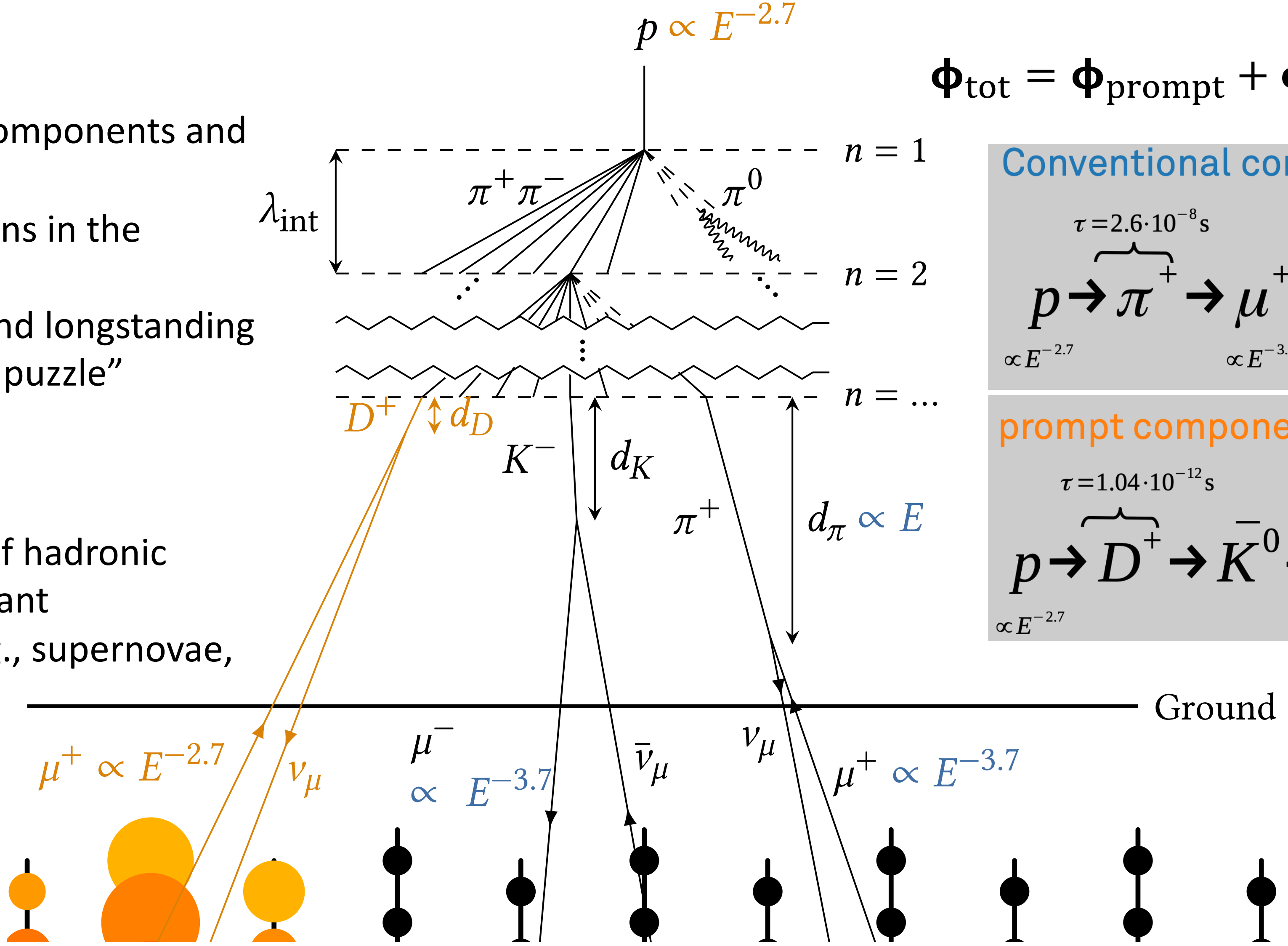
$$\Phi_{\text{tot}} = \Phi_{\text{prompt}} + \Phi_{\text{conventional}}$$

Conventional component:

$$\underset{\propto E^{-2.7}}{p} \rightarrow \overbrace{\pi^+}^{\tau = 2.6\cdot 10^{-8}\,\mathrm{s}} \rightarrow \underset{\propto E^{-3.7}}{\mu^+} + \nu_\mu$$

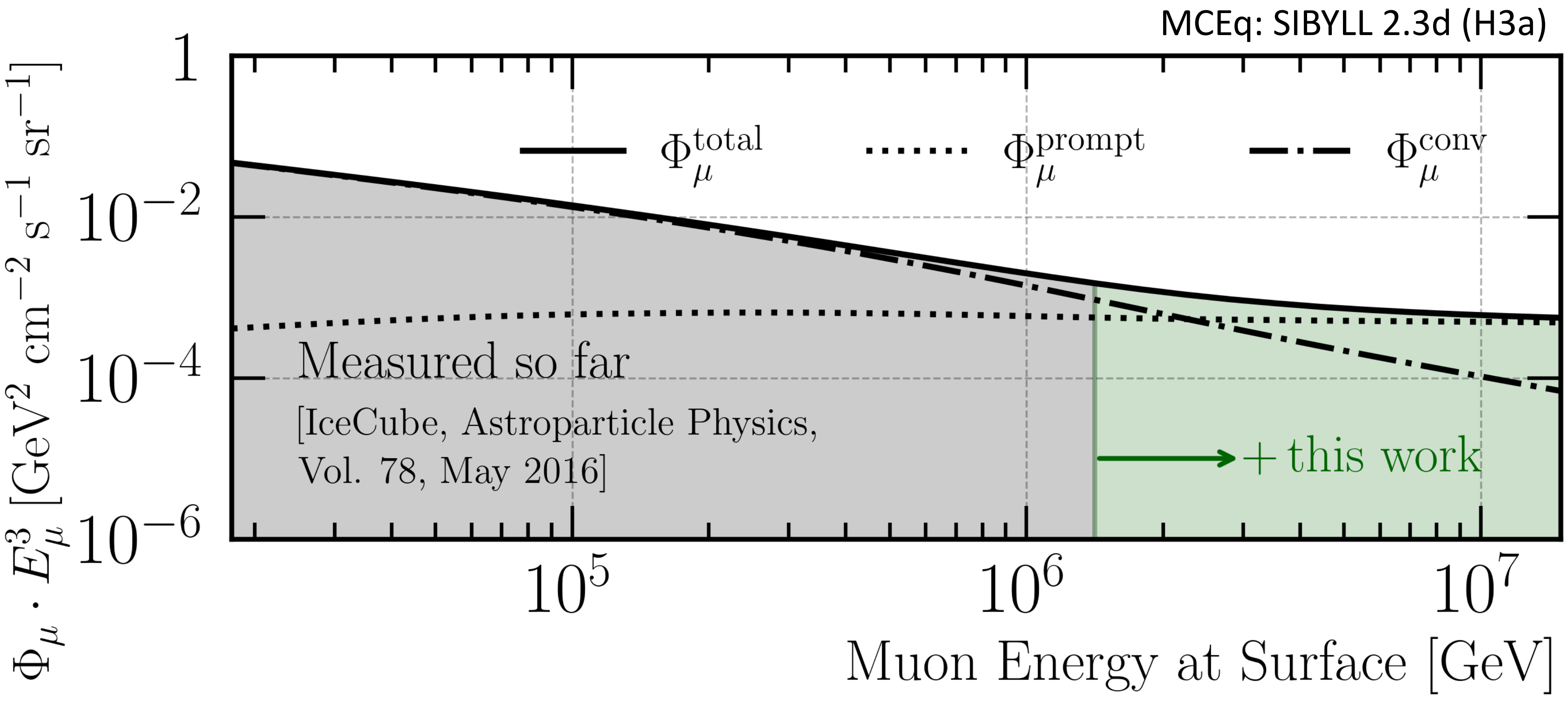
prompt component:

$$\underset{\propto E^{-2.7}}{p} \rightarrow \overbrace{D^+}^{\tau = 1.04\cdot 10^{-12}\,\mathrm{s}} \rightarrow \bar{K}^0 + \underset{\propto E^{-2.7}}{\mu^+} + \nu_\mu$$



Credit: Ludwig Neste

Goal: Measure Muon Flux at Surface



Overview

Simulation → **Reconstruction** → **Selection** → **Unfolding**

Simulation

Dataset
- CORSIKA 77500 Ehist
- SIBYLL 2.3d
- 10 TeV – 100 EeV

Validation
- Tag prompt / conv particles
- MCEq comparisons

Reconstruction

DNN reconstructions
- Muon energies
- Muon direction
- Track geometry

Selection

Level 3
- L2 muon filter
- 500 TeV bundle energy cut at surface
- Add labels

Level 4
- Add DNN reconstructions

Level 5
- Data-MC quality cuts

Final Level
- Leadingness cut

Unfolding

Muon spectrum
- Unfold muon flux at surface

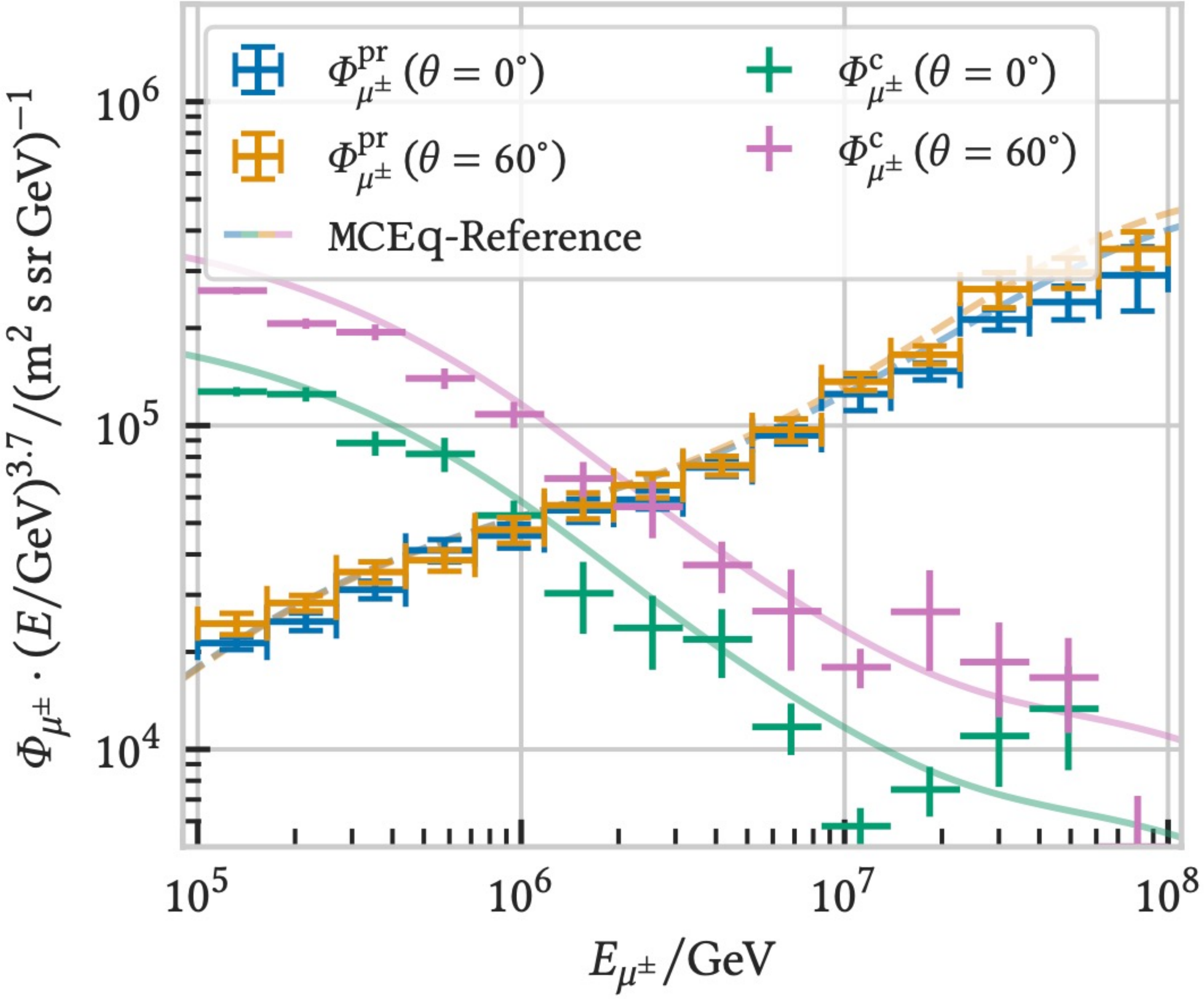
New CORSIKA simulation

with extended history option for information about the parent particles tu dortmund university BERKELEY LAB

ICECUBE NEUTRINO OBSERVATORY

CORSIKA 7 vs. MCEq

MCEq: tool to numerically solve the cascade equations that describes the evolution of particle densities as they propagate through a gaseous, dense medium
https://github.com/mceq-project/MCEq



➢ Good agreement for inclusive flux

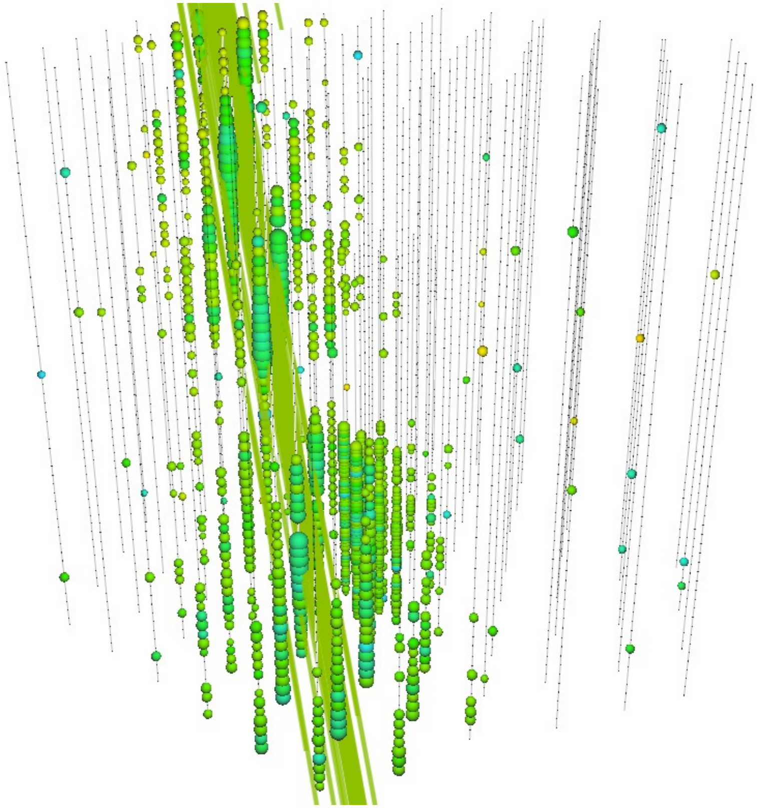
Python package developed – PANAMA

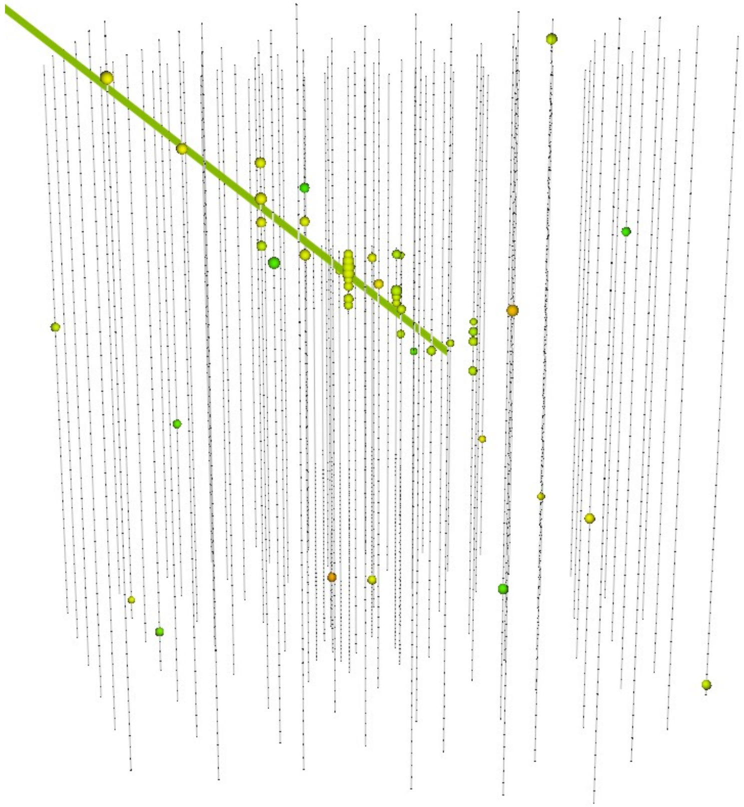
- Execute CORSIKA 7 (multi core)
- Read DAT files → pandas DataFrames
- Parse EHIST option
- Calculate primary weightings

https://github.com/The-Ludwig/PANAMA
Neste & Gutjahr et al., EPJ C, 85, 759, (2025)

Stopping and Leading Muons

- Leading muons:
  - Most energetic muon in the bundle
  - High energies above 10 TeV
  - Focus on muons carrying > 40% of the entire bundle





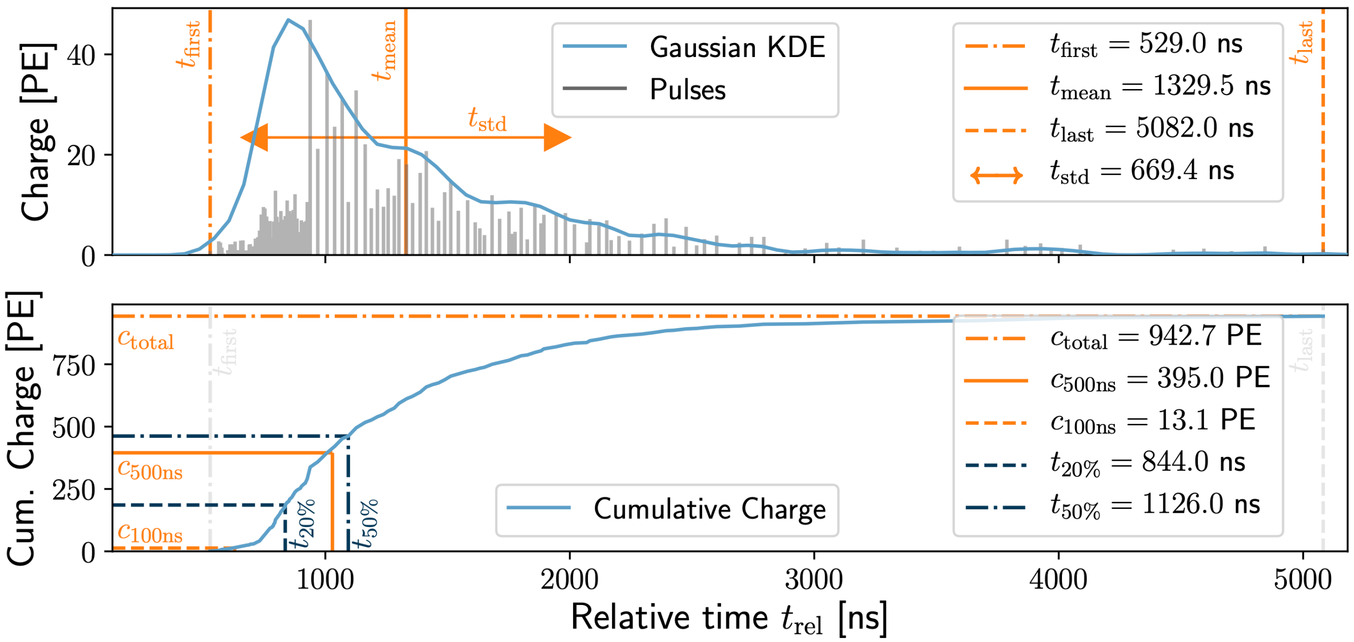
- Stopping muons:
  - Stop (decay) inside the in-ice array
  - Stopping point + direction → propagated length
  - Proxy to muon energy at the surface
  - Low-energy muons

Reconstruction

Machine Learning (CNN)

Input data per DOM

10.1088/1748-0221/16/07/P07041



3 inputs

- c_{total}: Total charge
  - Sum of charge
- t_{first}: Relative time of first pulse
  - Relative to total time offset, calculated as the charge weighted mean time of all pulses
- t_{std}: Standard deviation of first pulse
  - Charge weighted standard deviation of pulse times relative to total time offset

9 inputs

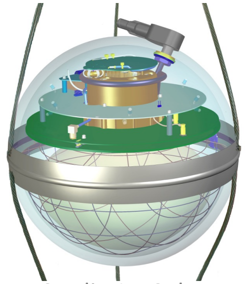
- t_{last}: Relative time of last pulse
  - Relative to total time offset, calculated as the charge weighted mean time of all pulses
- $t_{20\%}$: Relative time of 20% charge
  - Relative to total time offset, calculated as the charge weighted mean time of all pulses
- $t_{50\%}$: Relative time of 50% charge
  - Relative to total time offset, calculated as the charge weighted mean time of all pulses
- t_{mean}: Mean time
  - Charge weighted mean time of all pulses relative to total time offset
- $c_{500\mathrm{ns}}$: Charge at 500ns
  - Sum of charge after 500ns
- $c_{100\mathrm{ns}}$: Charge at 100ns
  - Sum of charge after 100ns

Input pulse cleaning

- 6000 ns

Training data

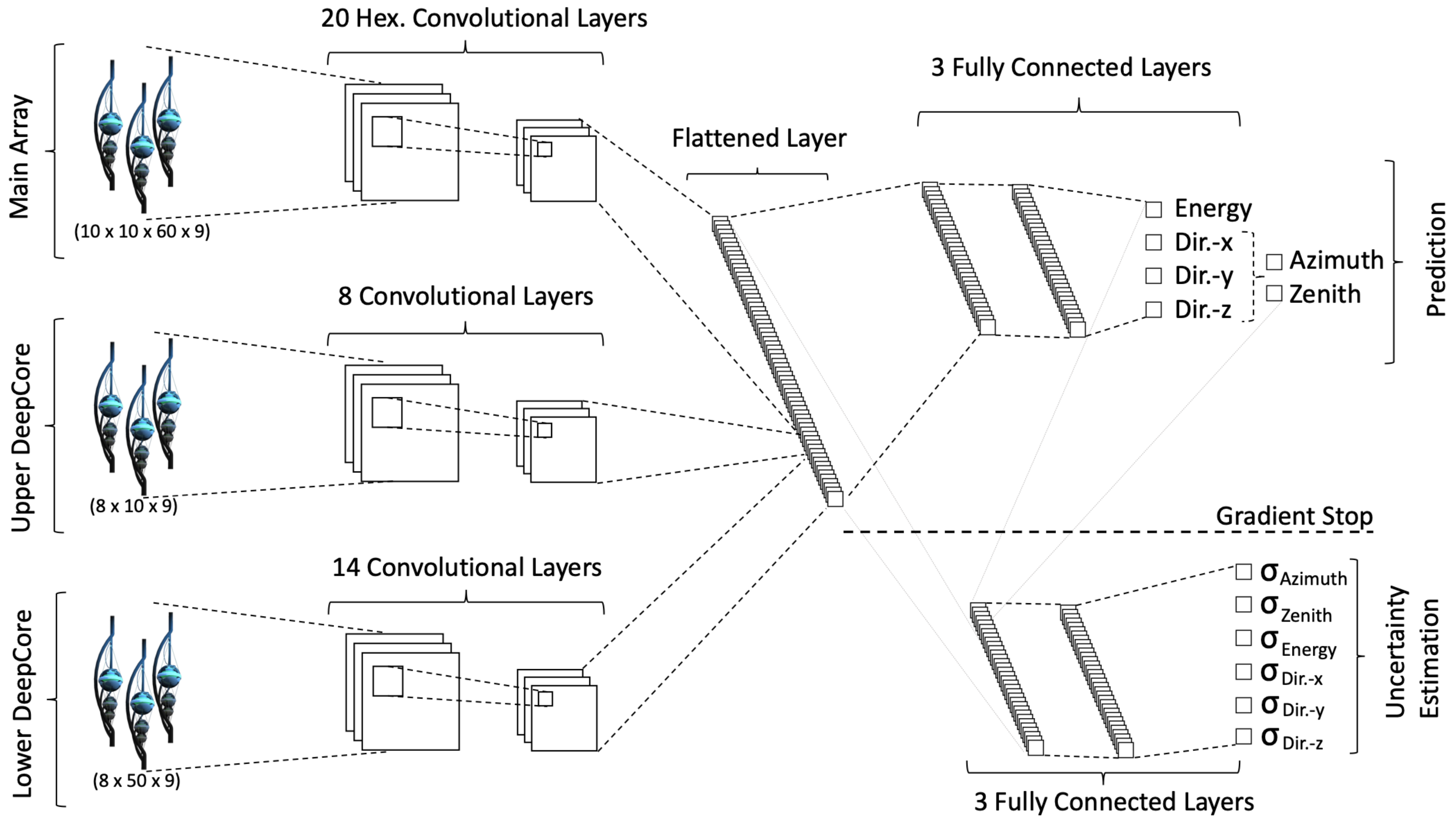
- Different simulations for robustness


Credit: IceCube

Convolutional Neural Network (CNN)

Machine learning approach:

- fast
- identifying spatial patterns
- better event reconstruction and classification



Reconstructed properties

Energy

- entry_energy: Leading muon energy at the detector entry
- bundle_energy_at_entry: Muon bundle energy at the detector entry
- muon_energy_first_mctree: Leading muon energy at surface
- bundle_energy_in_mctree: Muon bundle energy at surface

Network 1

Network 0: Muon bundle energy at surface (fast pre-cut)

Track geometry

- Length: Propagation length of muon in the ice
- LengthInDetector: Propagation length of muon in the detector
- center_pos_x: Closest x position of muon to center of the detector
- center_pos_y: Closest y position of muon to center of the detector
- center_pos_z: Closest z position of muon to center of the detector
- center_pos_t: Time of closest approach to the center of the detector
- entry_pos_x: x position of muon at the detector entry
- entry_pos_y: y position of muon at the detector entry
- entry_pos_z: z position of muon at the detector entry
- entry_pos_t: Time of muon at the detector entry

Network 2

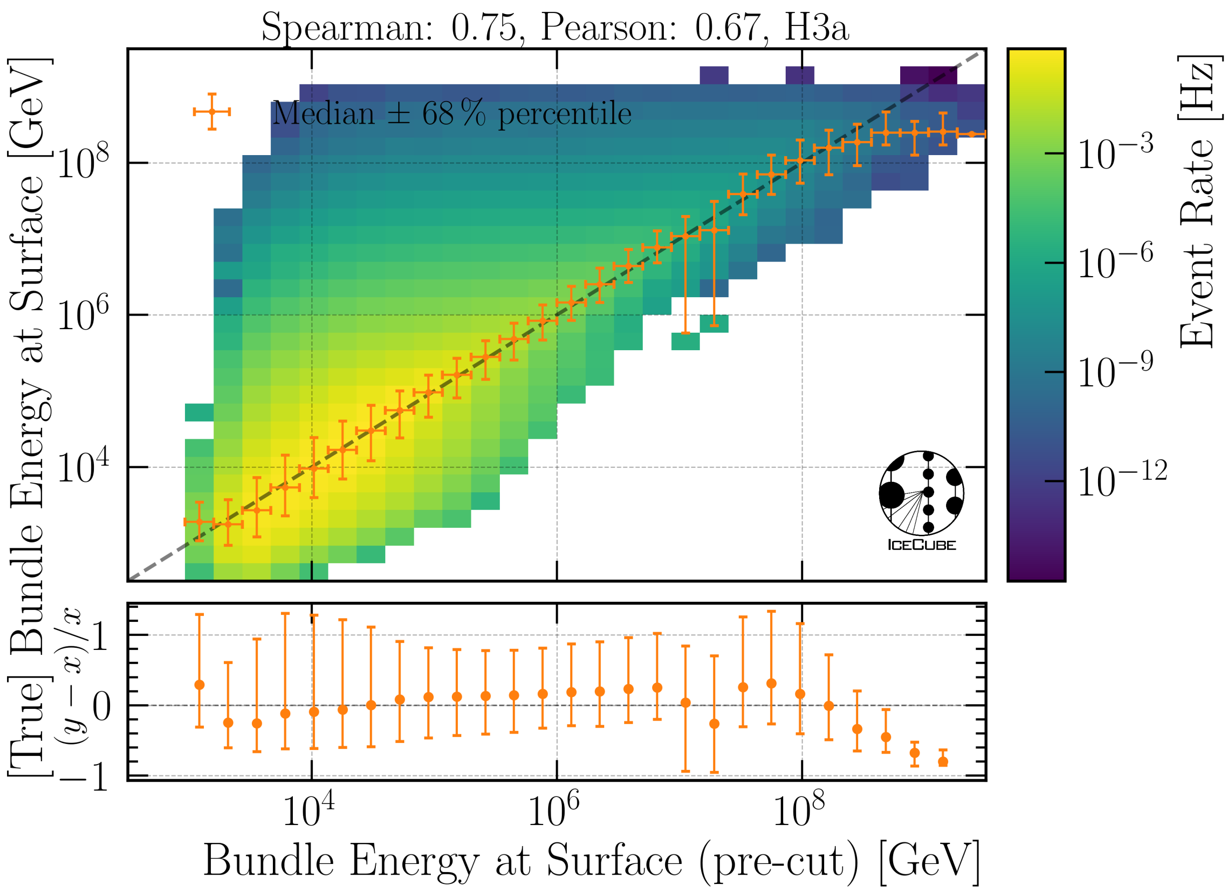
Direction

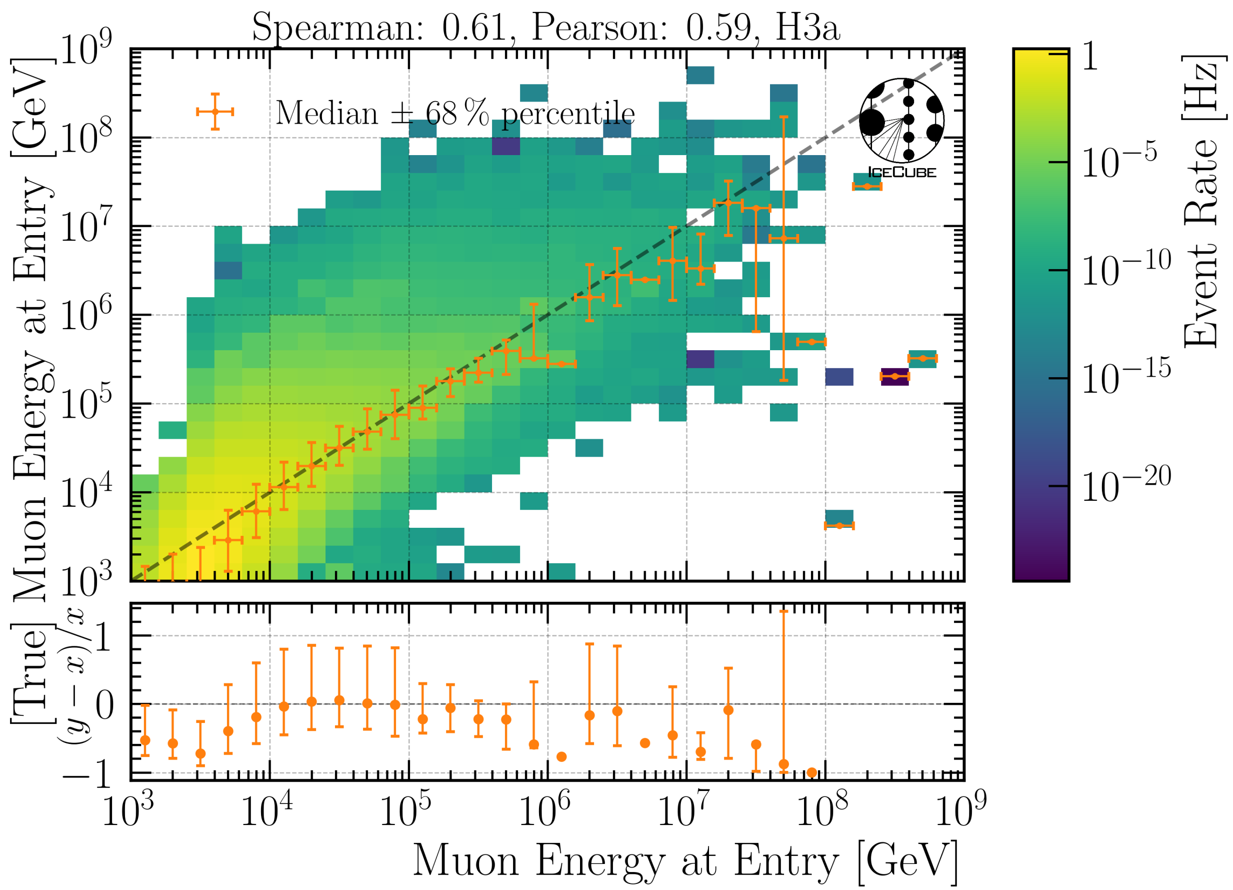
- zenith: Zenith angle of muon
- azimuth: Azimuth angle of muon

Network 3

performed on Level 3 (explained later)

Energy Reconstructions



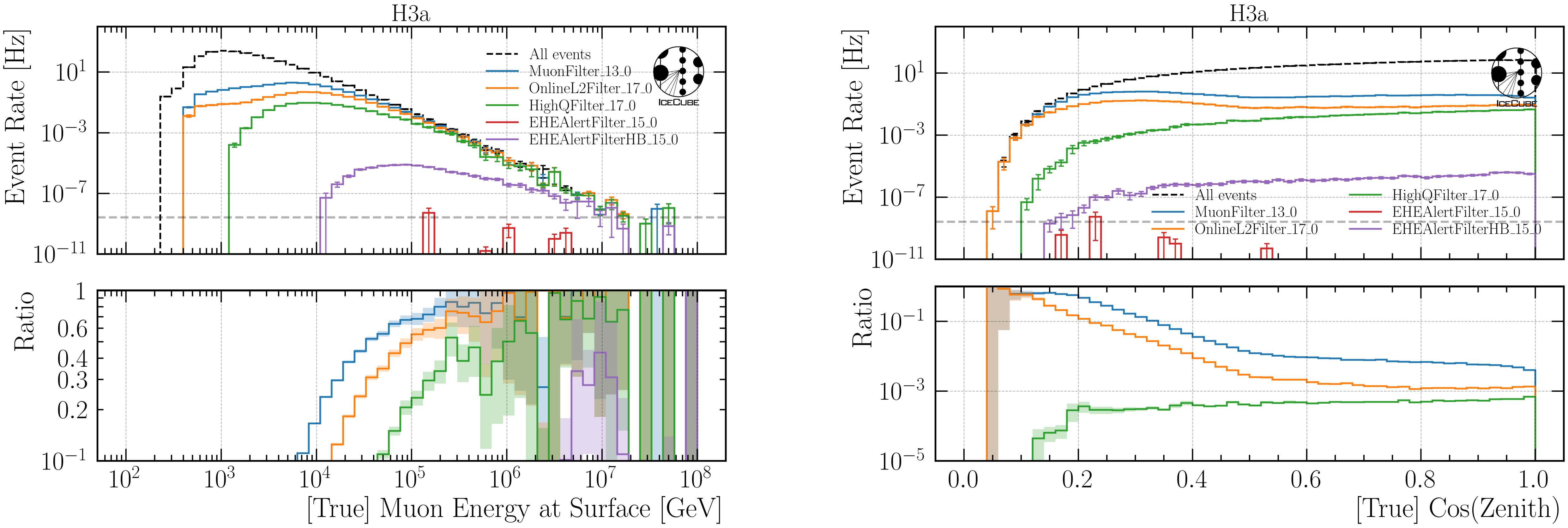


- ➢ Sufficient energy reconstructions
- ➢ Tight 68 % intervals, with outliers

pascal.gutjahr@tu-dortmund.de

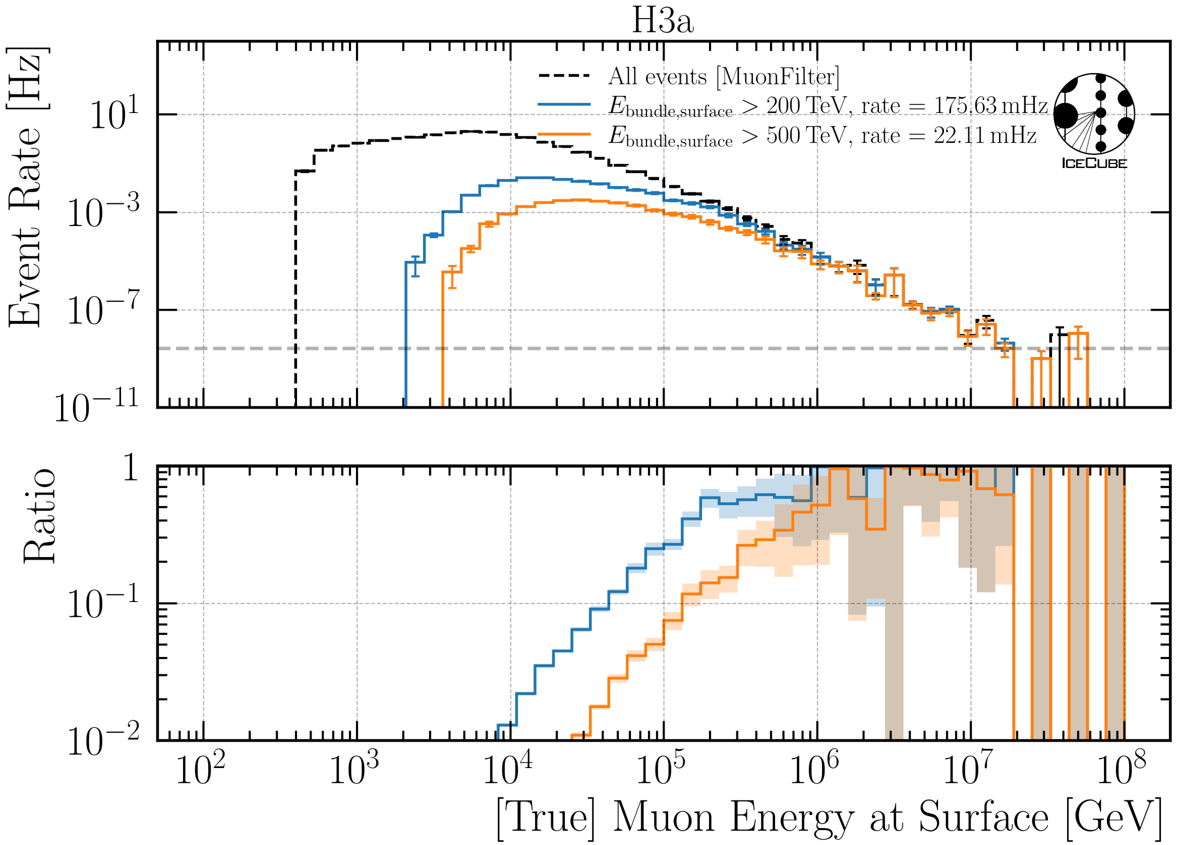
Selection

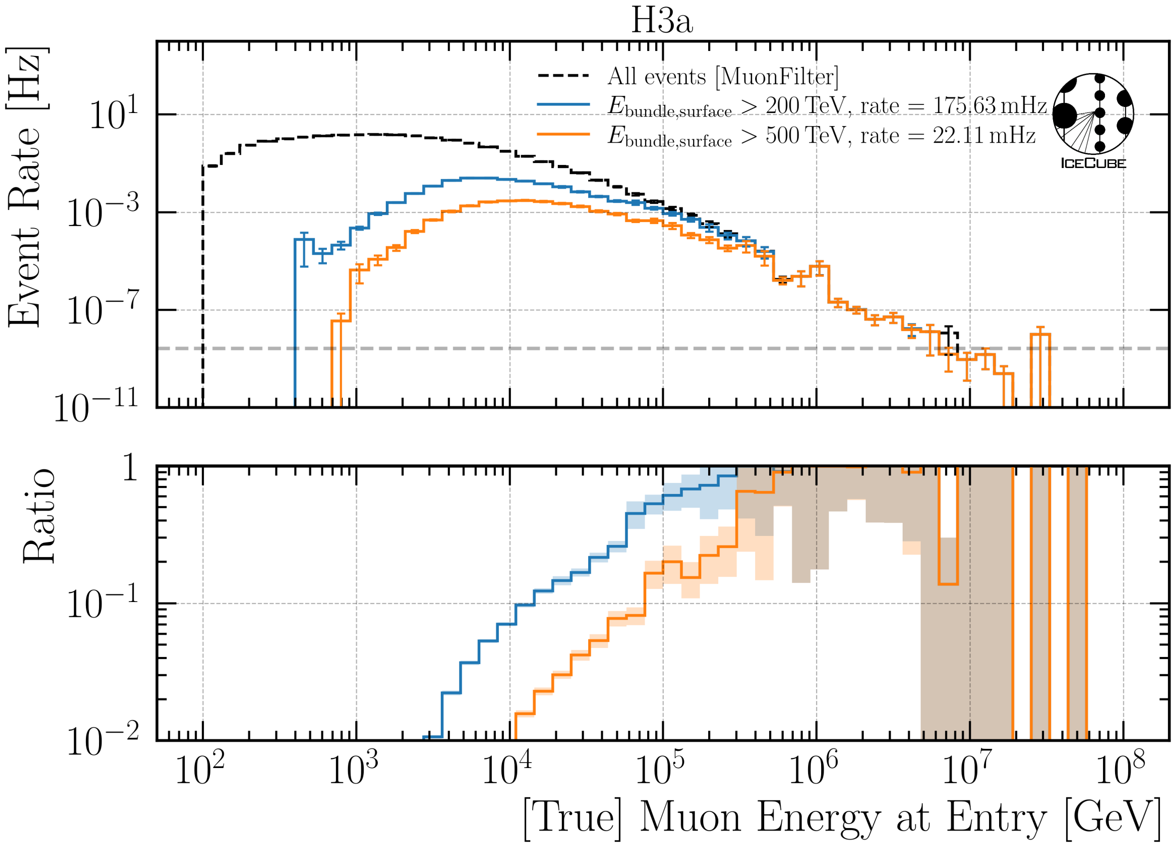
Level 3: Muon Filter



➢ Muon Filter to select as many high-energy muons as possible

Level 4: Energy Cut



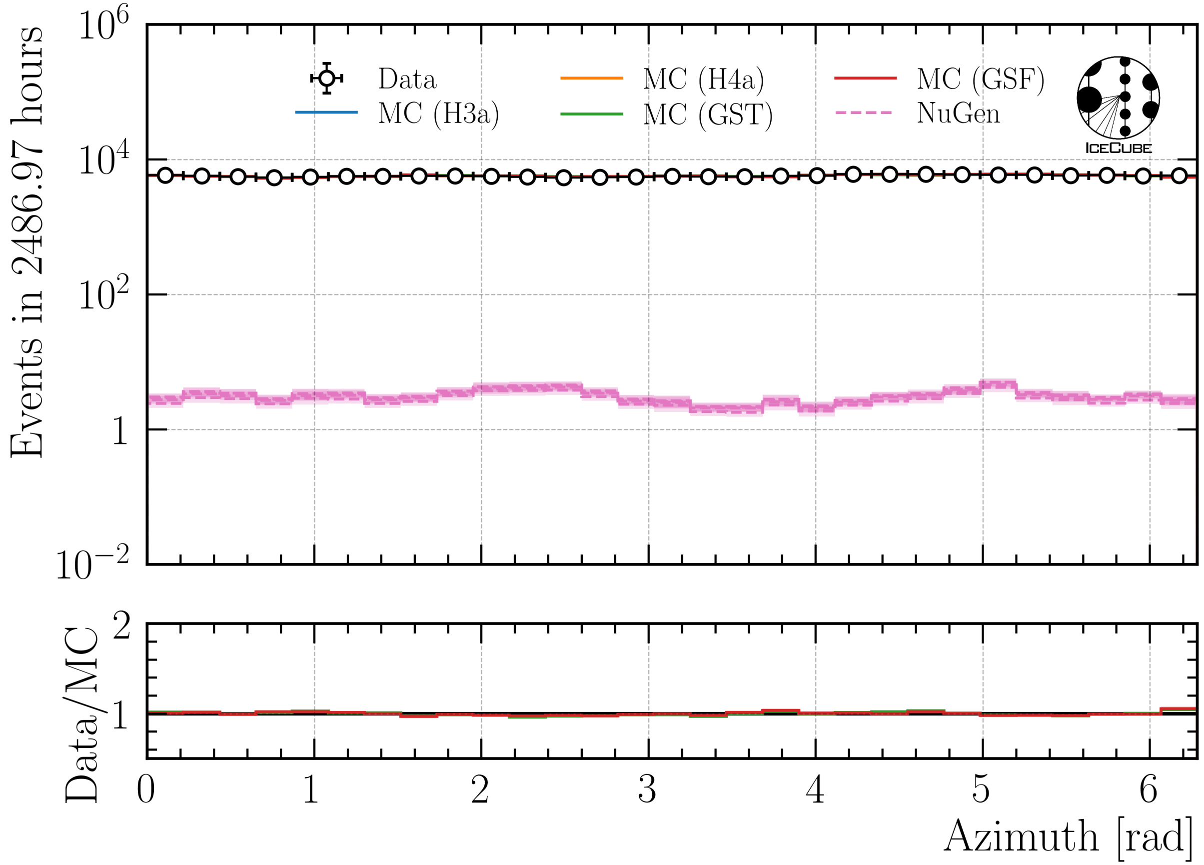


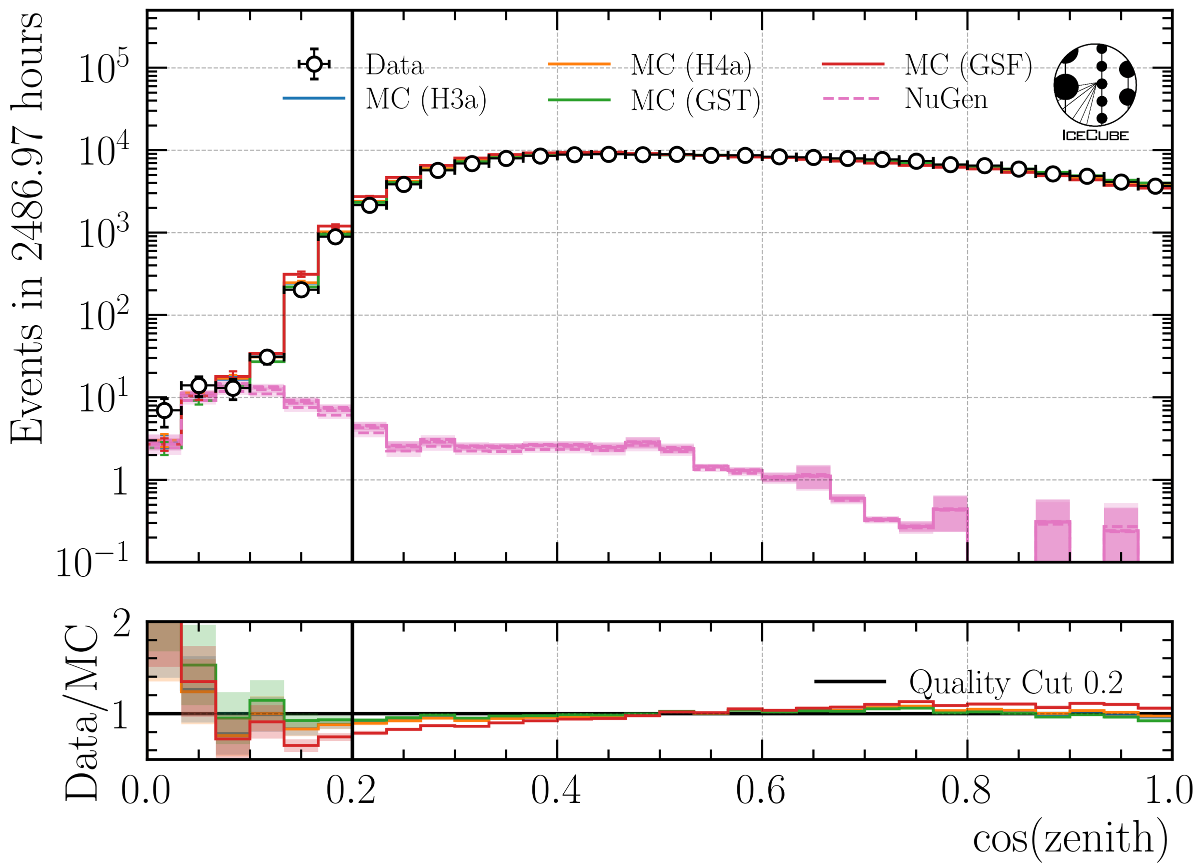
- 6 billion events expected in 10 years → computationally not feasible
- focus on high-energetic events
- Remove low-energy muons: bundle energy at surface > 500 TeV

Level 5: Data/MC Quality Cuts

Neutrino weighting:
SPL: n = 1.8, γ = 2.52

IceCube Collaboration PoS ICRC2023 1064

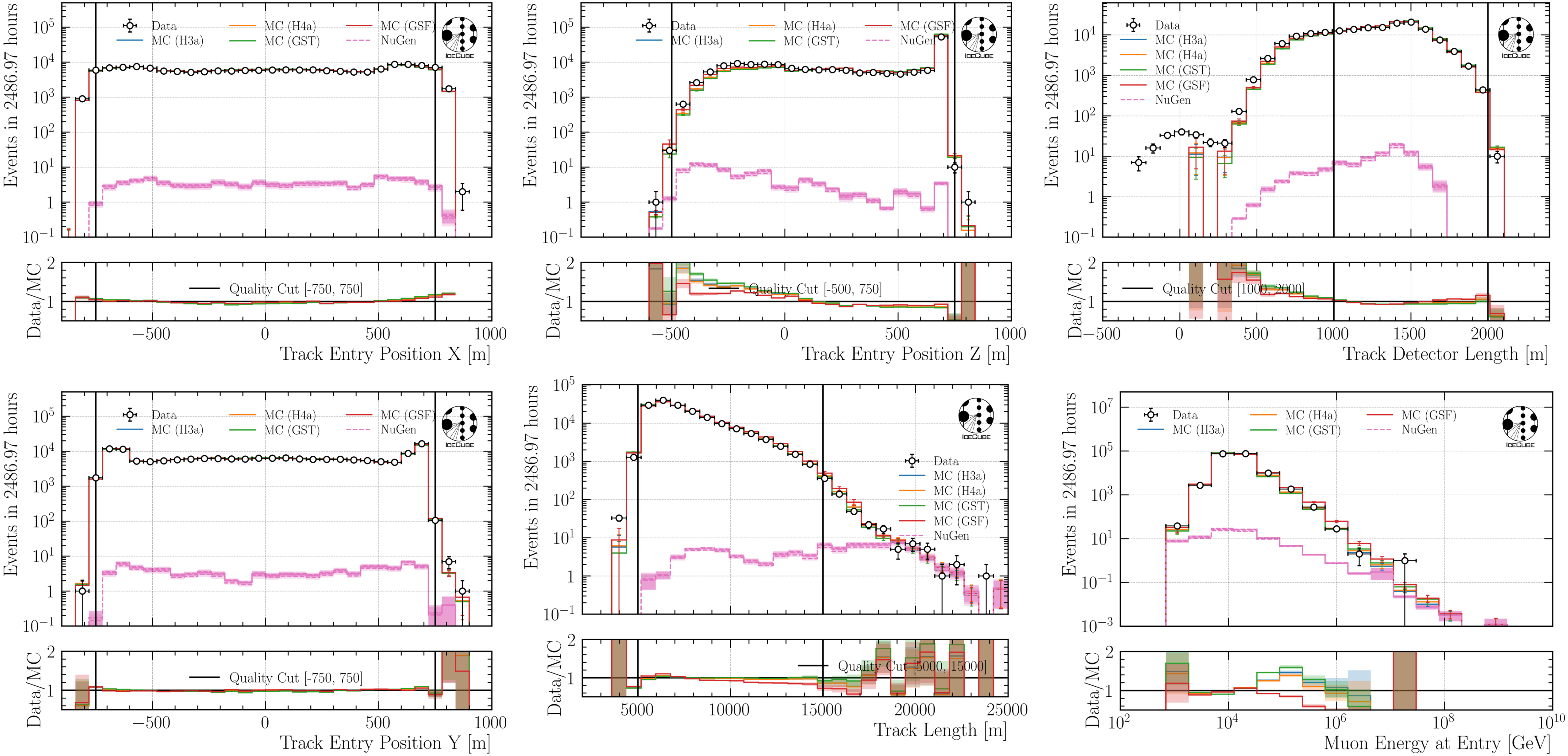




- Good directional reconstruction
- Cut: cos(zenith) > 0.2

Level 5: Data/MC Quality Cuts

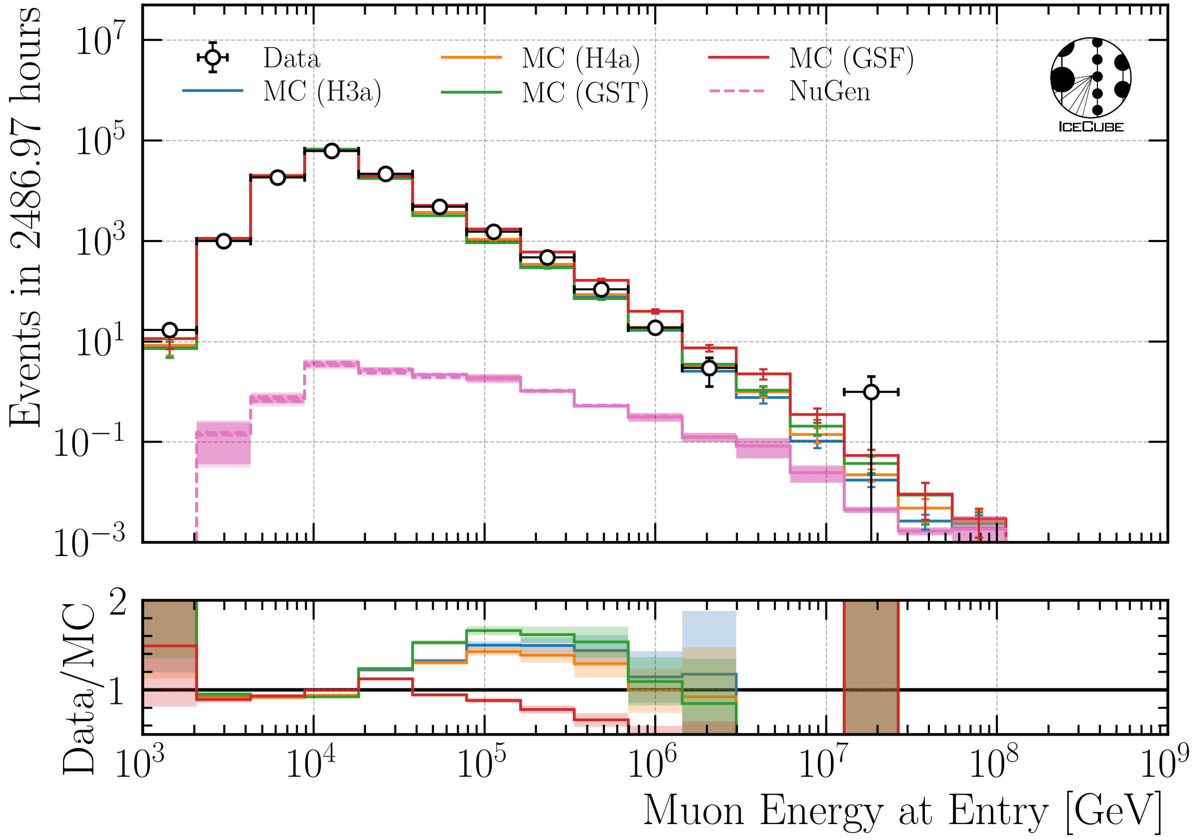
- Remove outliers
- Depth—dependent slope (no analysis relevance)
- CR—model impact on energy reconstruction
- 23 quality cuts in total



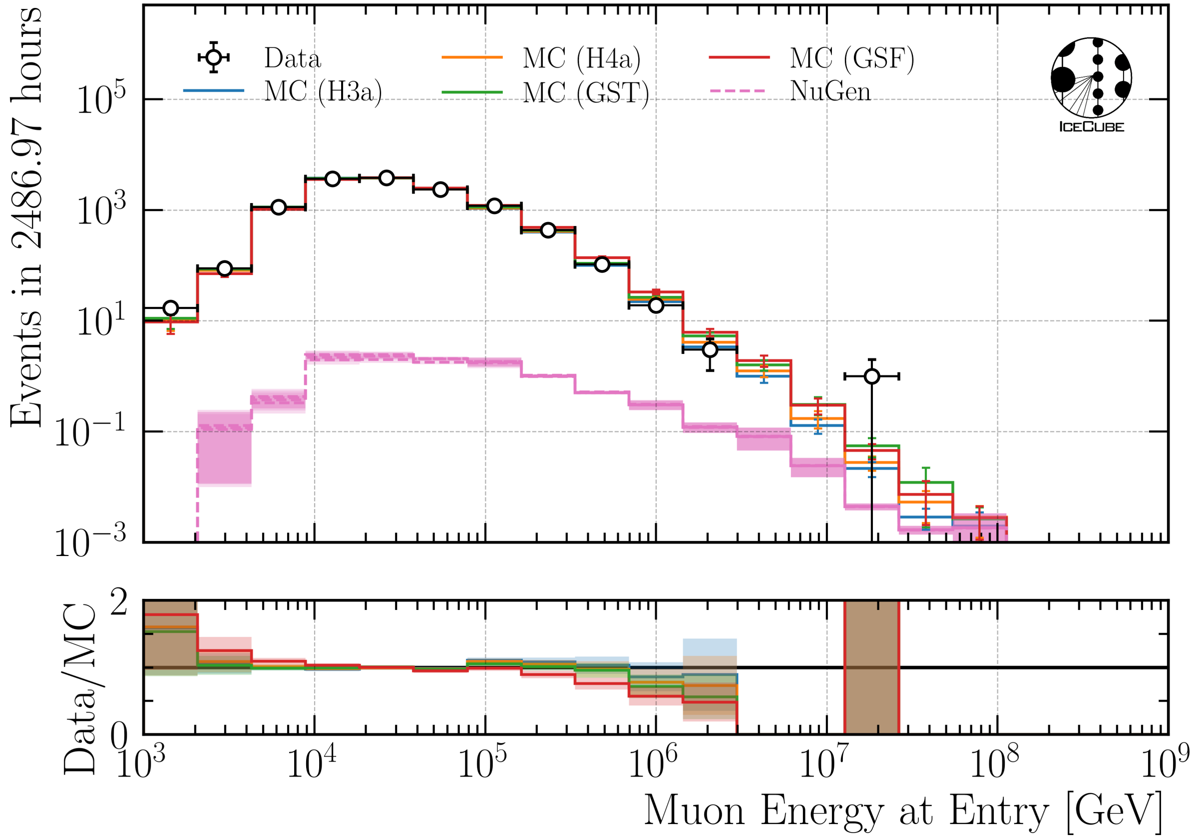
Final Level: Leadingness > 40 %

$$L = \frac{\text{Leading Energy at Entry}}{\text{Bundle Energy at Entry}}$$

No leadingness cut



Leadingness > 40 %



➢ Improve Data/MC by leadingness cut

Unfolding

Unfolding in a nutshell

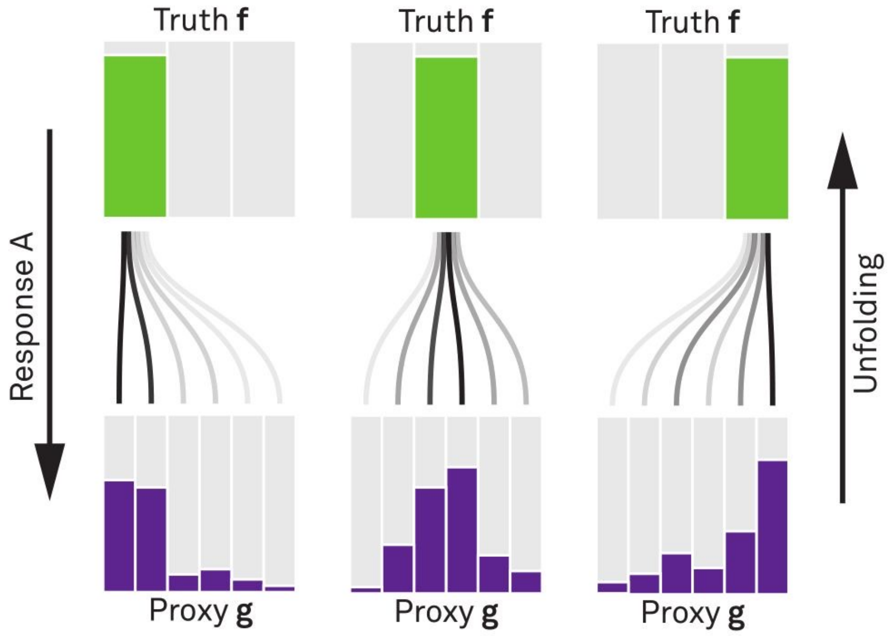
measured proxy

detector response

background

$$g(y) = \int_{E_0}^{E_1} A(E_\mu, y) f(E_\mu)\,\mathrm{d}E_\mu + b(y)$$

true energy distribution



Credit: T. Hoinka

folding unfolding

1. Discretized form: $\underbrace{\vec{g} = A\vec{f}}_{\text{folding}} \leftrightarrow \underbrace{\vec{f} = A^{-1}\vec{g}}_{\text{unfolding}}$

2. Maximum likelihod method:

$$\mathcal{L}(\vec{g}|\vec{f}) = \prod_{j=1}^{M} \frac{\lambda_j^{g_j}}{g_j!}\exp(-\lambda_j)$$
$$= \prod_{j=1}^{M} \frac{(A\vec{f})_j^{g_j}}{g_j!}\exp(-(A\vec{f})_j)$$

3. Thikonov regularization:

$$t(\vec{f}) = -\frac{1}{2}(C\vec{f})^T(\tau 1)^{-1}(C\vec{f})$$

write down full LLH

4. Maximize $\log(\mathcal{L}(\vec{g}|\vec{f})) + t(\vec{f})$ with respect to $\vec{f}$ using Markov Chain Monte Carlo (MCMC) or Minuit

GitHub **funfolding** by M. Börner

Muon Flux Unfolding

- ❑ Effective area
- ❑ Systematics
- ❑ Proxy that correlates with target
- ❑ Regularization
- ➢ Unfolding

Effective Area

- Unfolding estimates an event rate
- Transfer event rate to flux

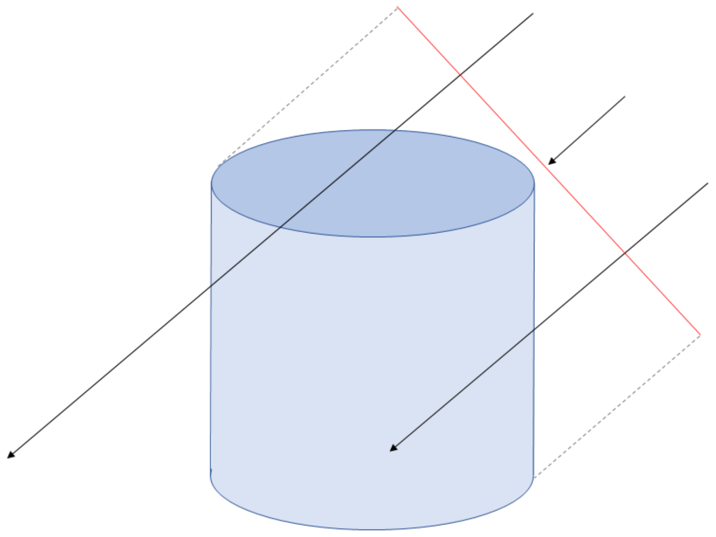
$$\Phi_i = \frac{N_i}{T \cdot \Delta L_i \cdot \Omega_i \cdot A_{\mathrm{eff},i}}$$

- with solid angle

$$\Omega_i = 2\pi \cdot (\cos \Theta_{\mathrm{min},i} - \cos \Theta_{\mathrm{max},i})$$

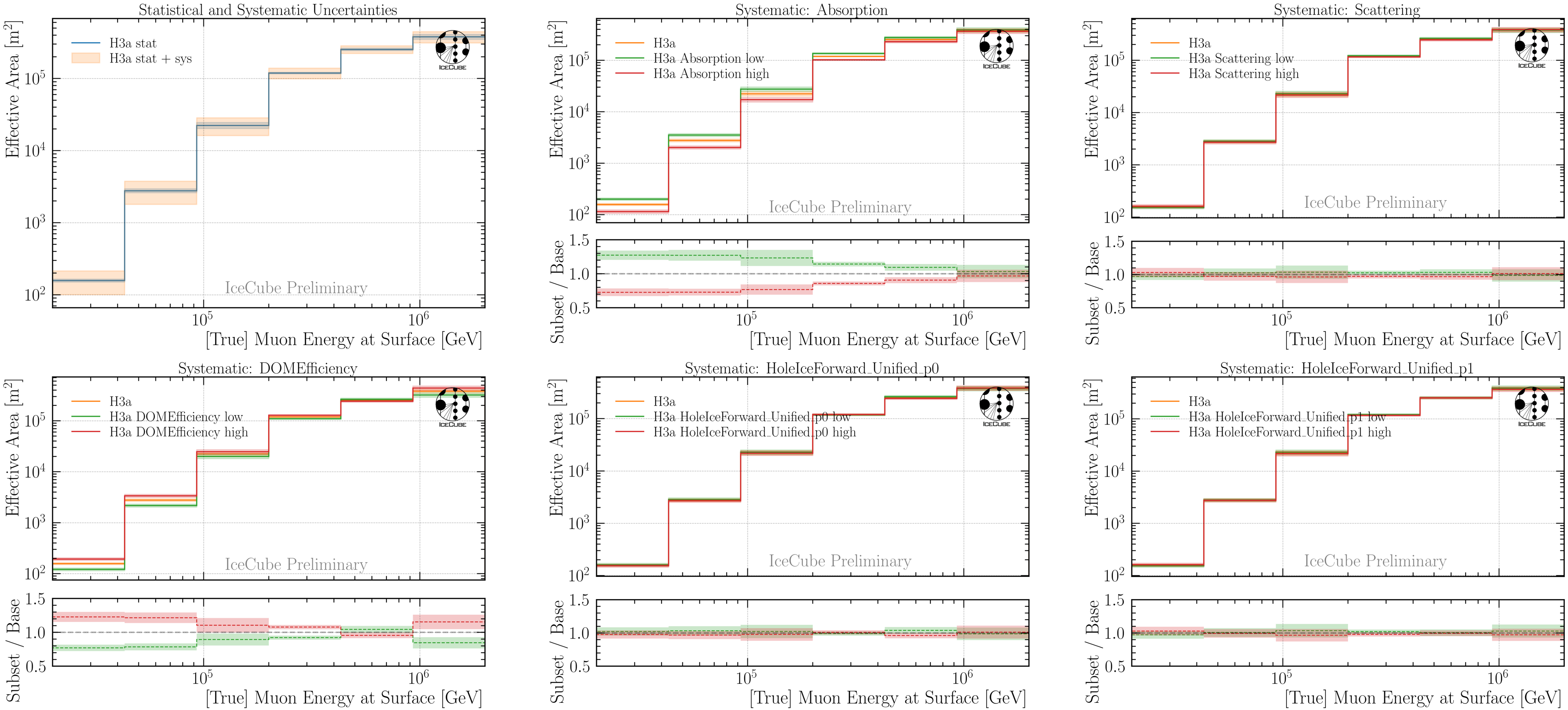
- and effective area

$$A_{\mathrm{eff}} = A_{\mathrm{sim}} \frac{N_{\mathrm{sel}}}{N_{\mathrm{gen}}}$$



Ice & Detector Systematics

$$\sigma_{sys,i} = \max(A_{\mathrm{low}} - \sigma, A_{\mathrm{high}} + \sigma) - A_{\mathrm{base}}$$

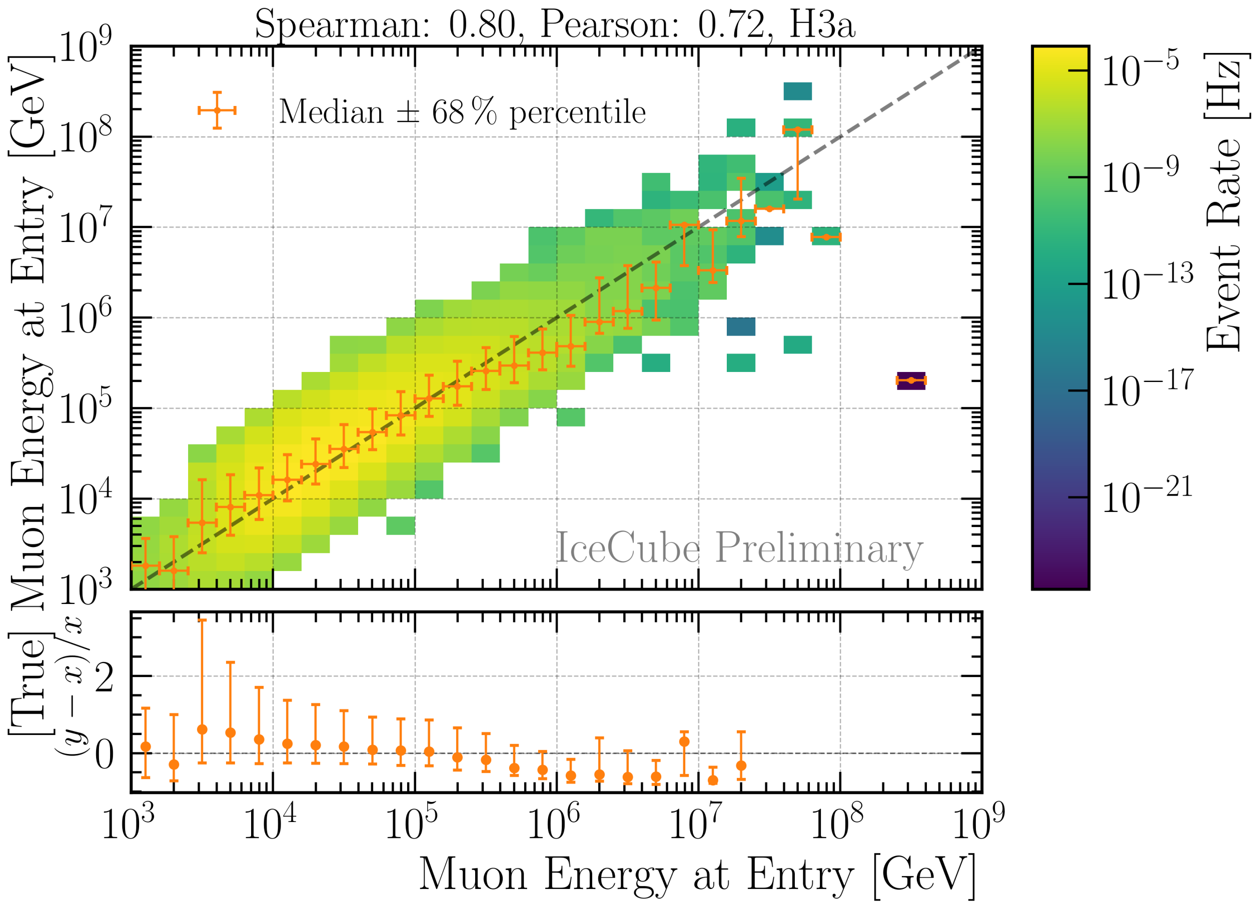


- Baseline: entire set
- Subset: above/below center

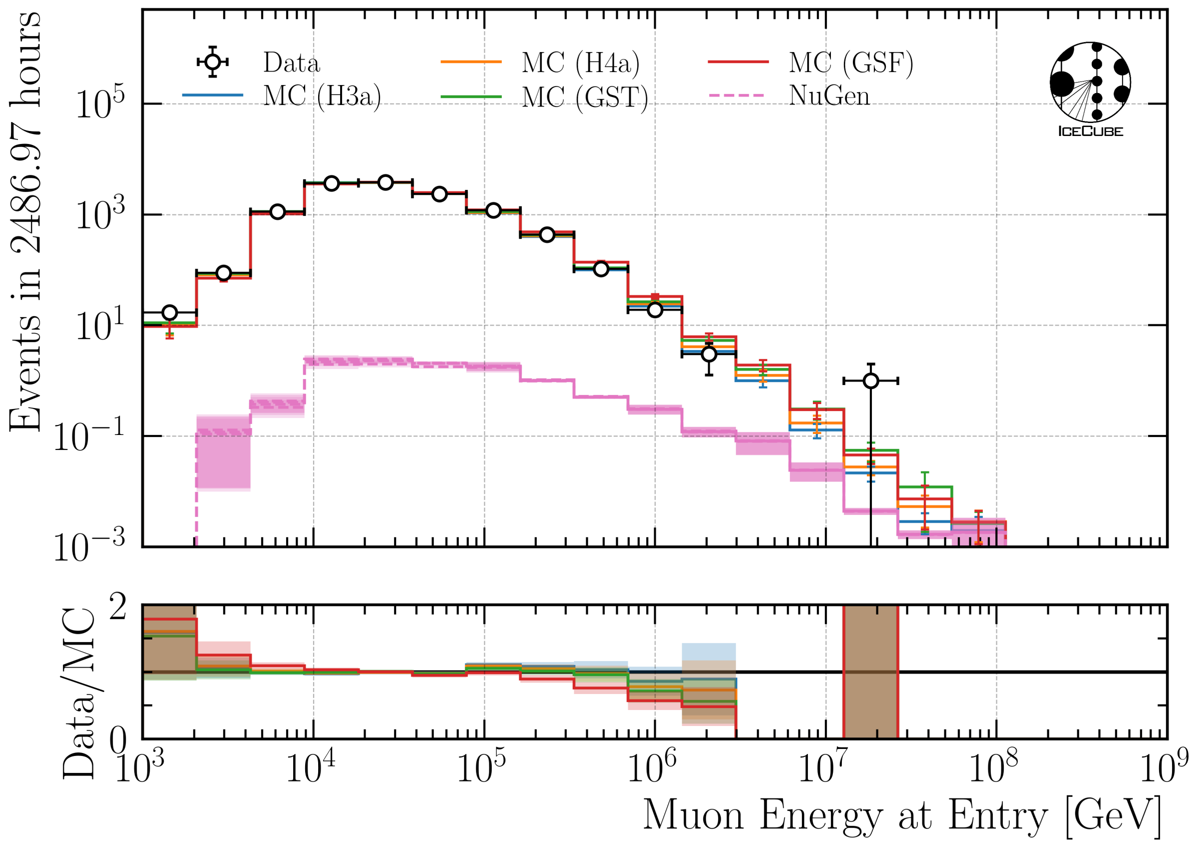
$$\sigma_{\mathrm{tot}} = \sqrt{\sigma_{\mathrm{stat}}^2 + \sum \sigma_i^2}\,, i\text{: scat, abs, DOME, holeice p0, p1}$$

Reconstruction and Data—MC: Leading Muons

- Good reconstruction of leading muon energy → proxy

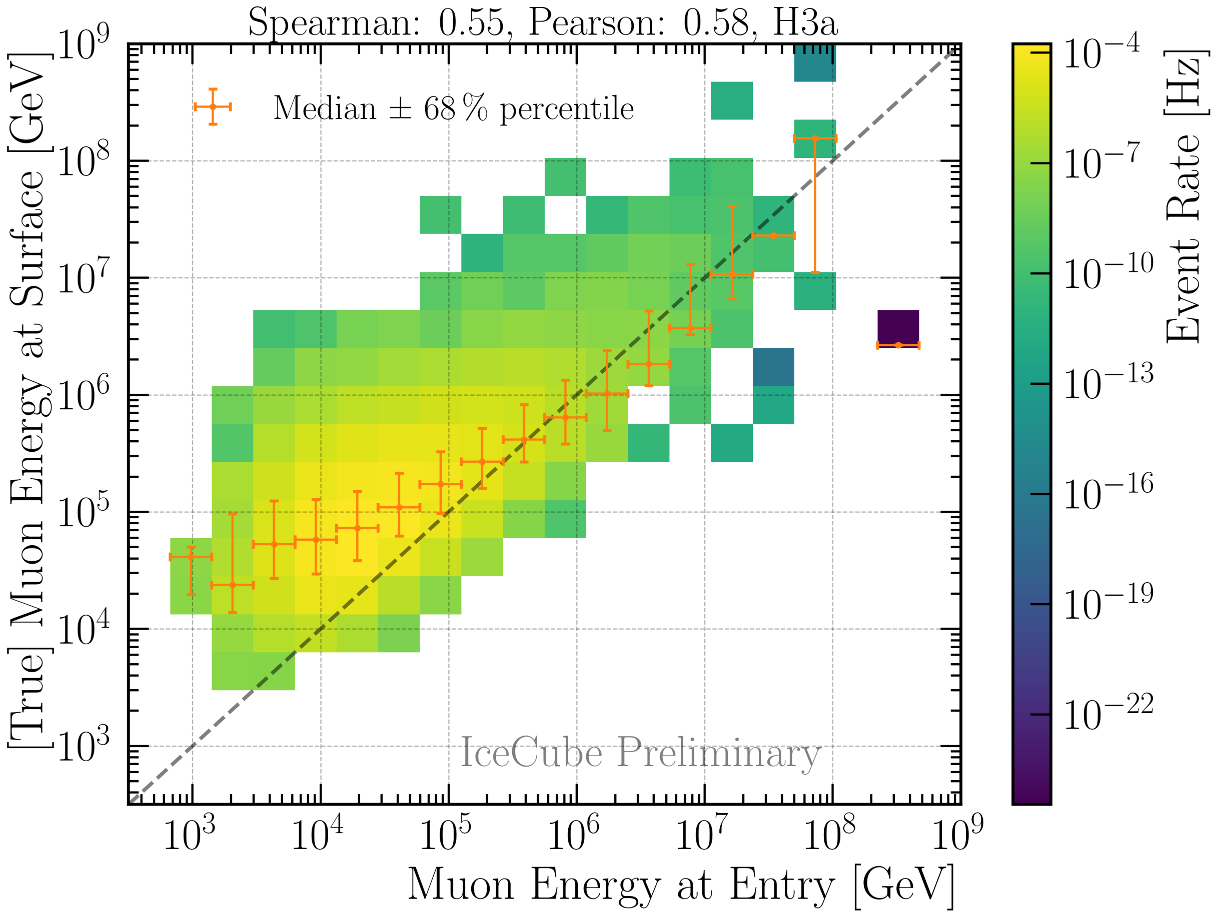


- Good data—MC agreement
- Global offset → upscale MC by 12%

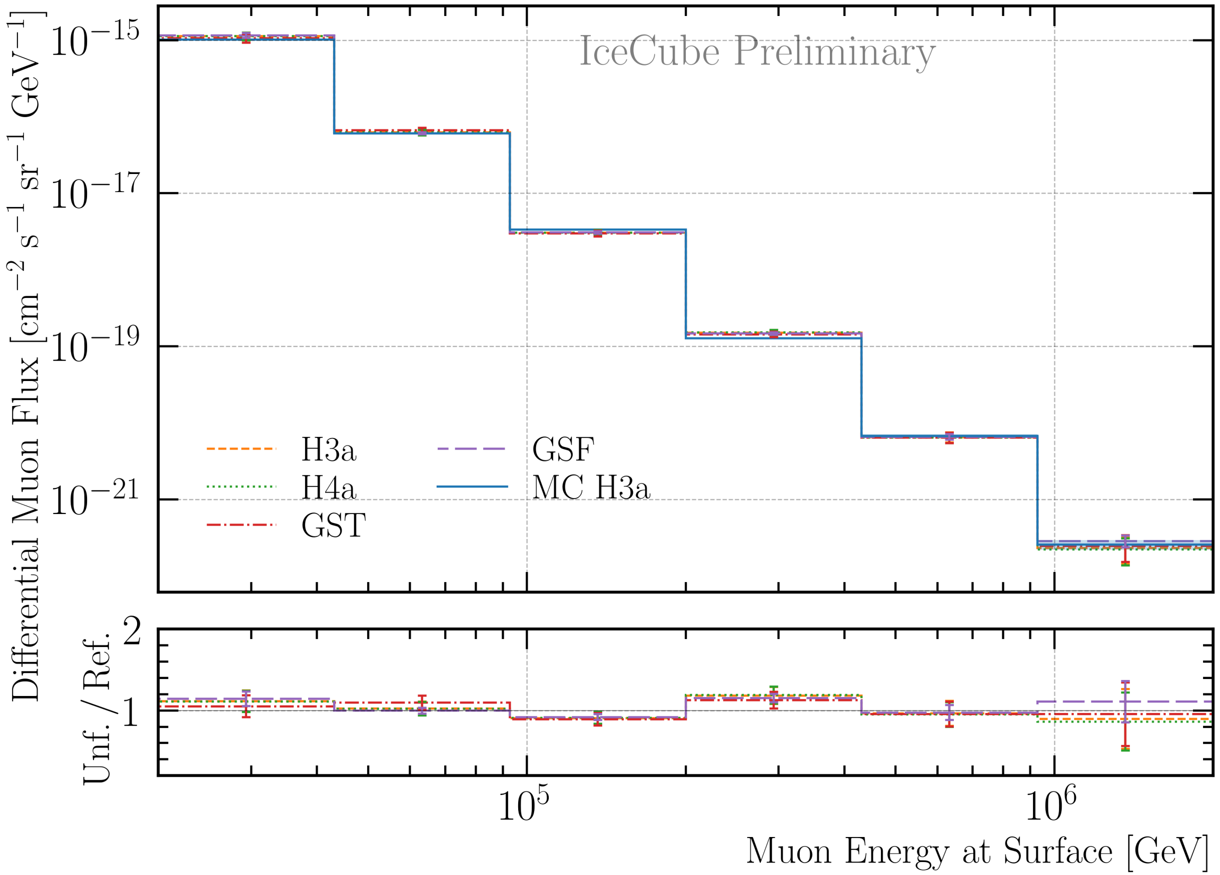


Correlation and Robustness Test: Leading Muons

- Correlation between proxy and target



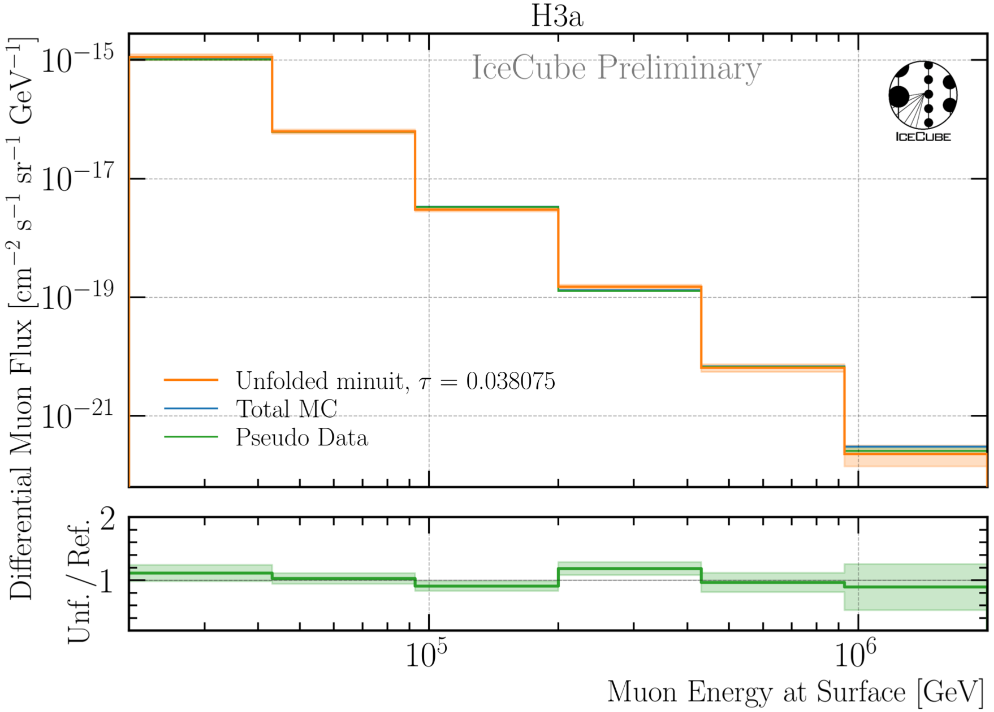
- Build unfolding matrix on H3a, H4a, GST & GSF
- Unfold H3a as test set
- Results are within uncertainties → robust

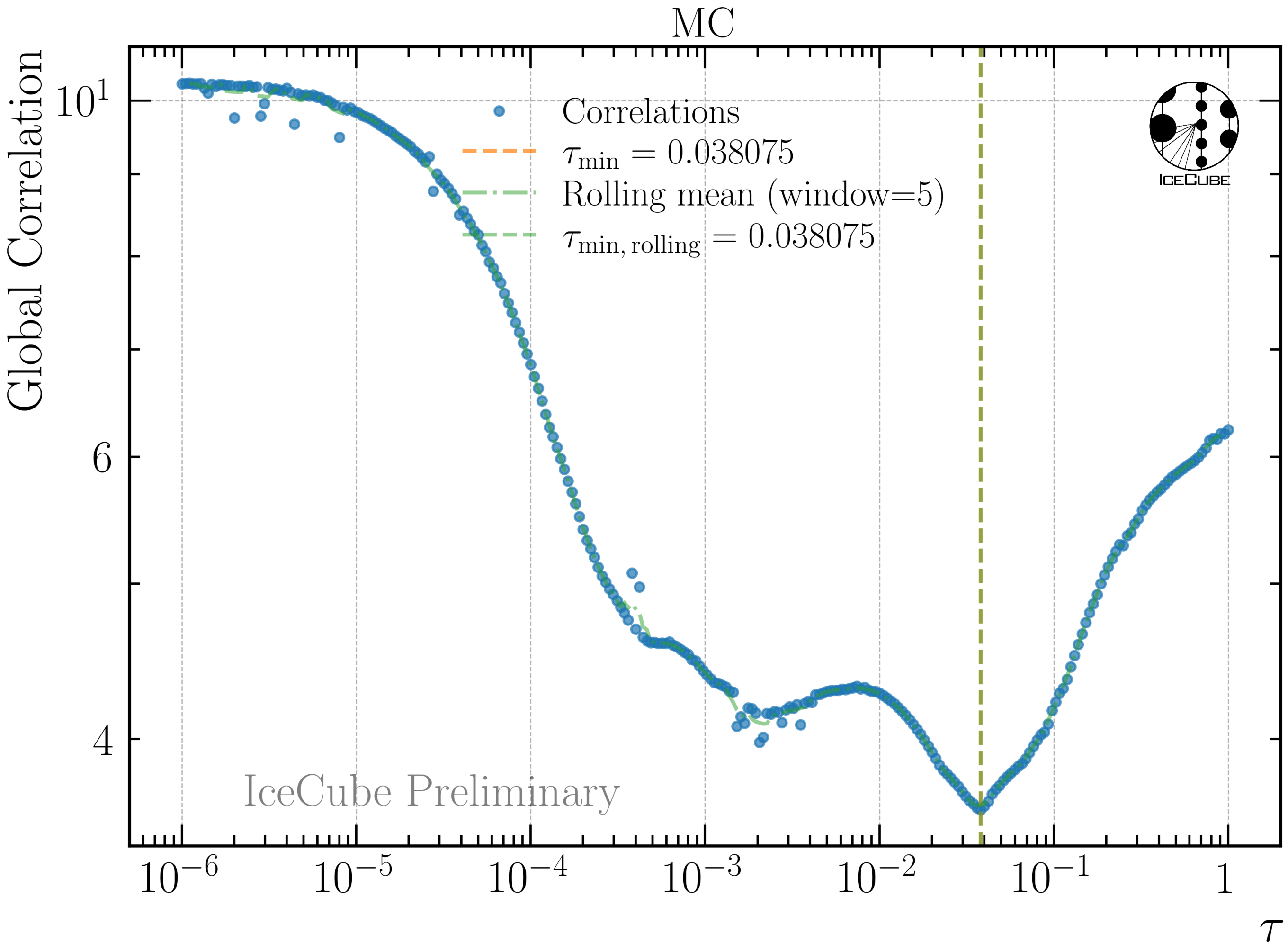


Determine Regularization

- Find regularization τ with minimal bin-to-bin correlation
- LLH minimization (unfolding) provides full covariance matrix V

➢ Minimize global correlation $\rho = \sum_{i > i} V_{ij}$



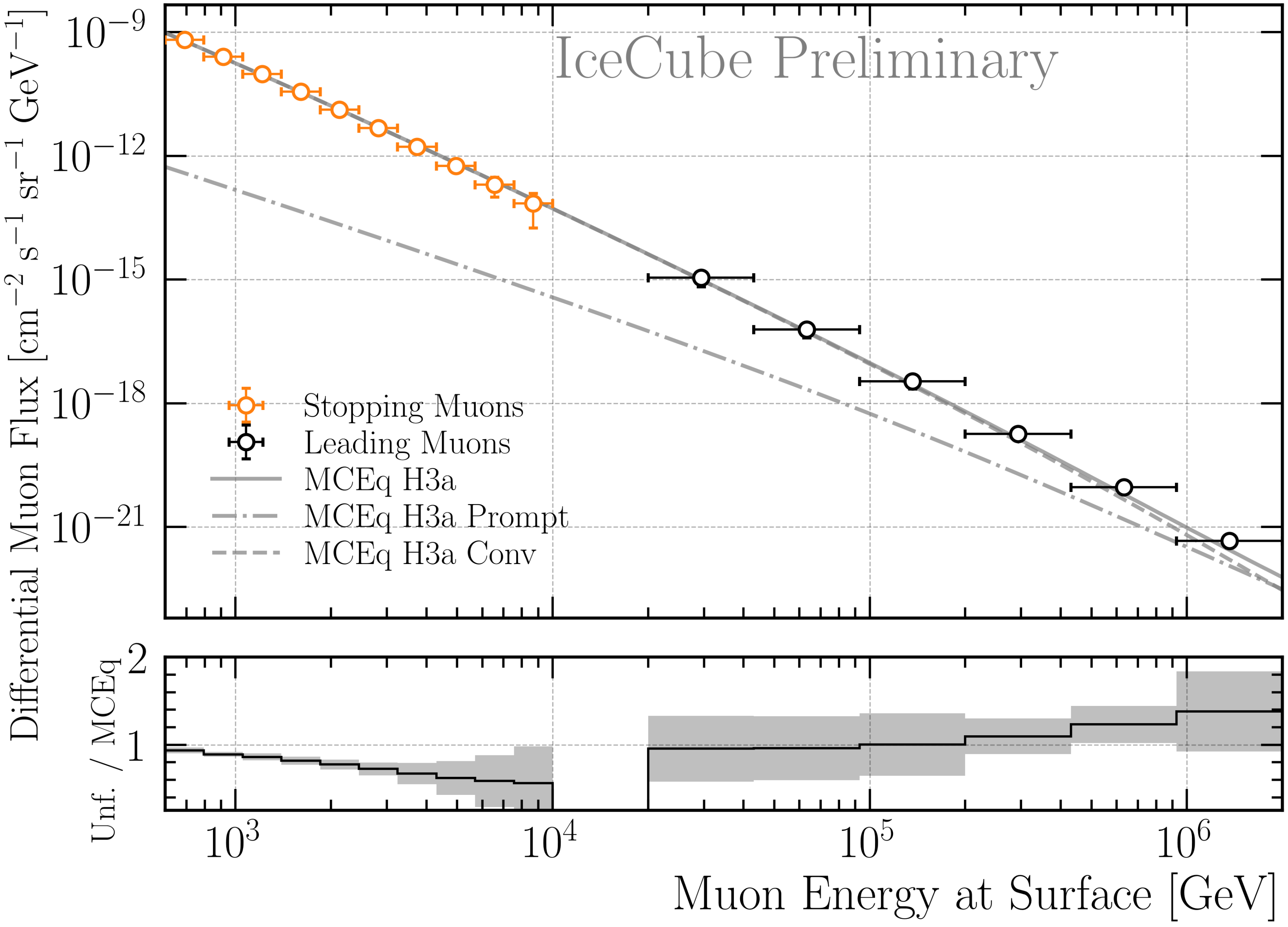


Unfold MC truth → ratio should be 1

Muon Flux Unfolding

Leading muons
- ➢ 2487 h IceCube data
- ➢ 12754 events
- ➢ Agrees with MCEq

Stopping muons
- ➢ 47 min IceCube data
- ➢ 32943 events
- ➢ Below MCEq prediction

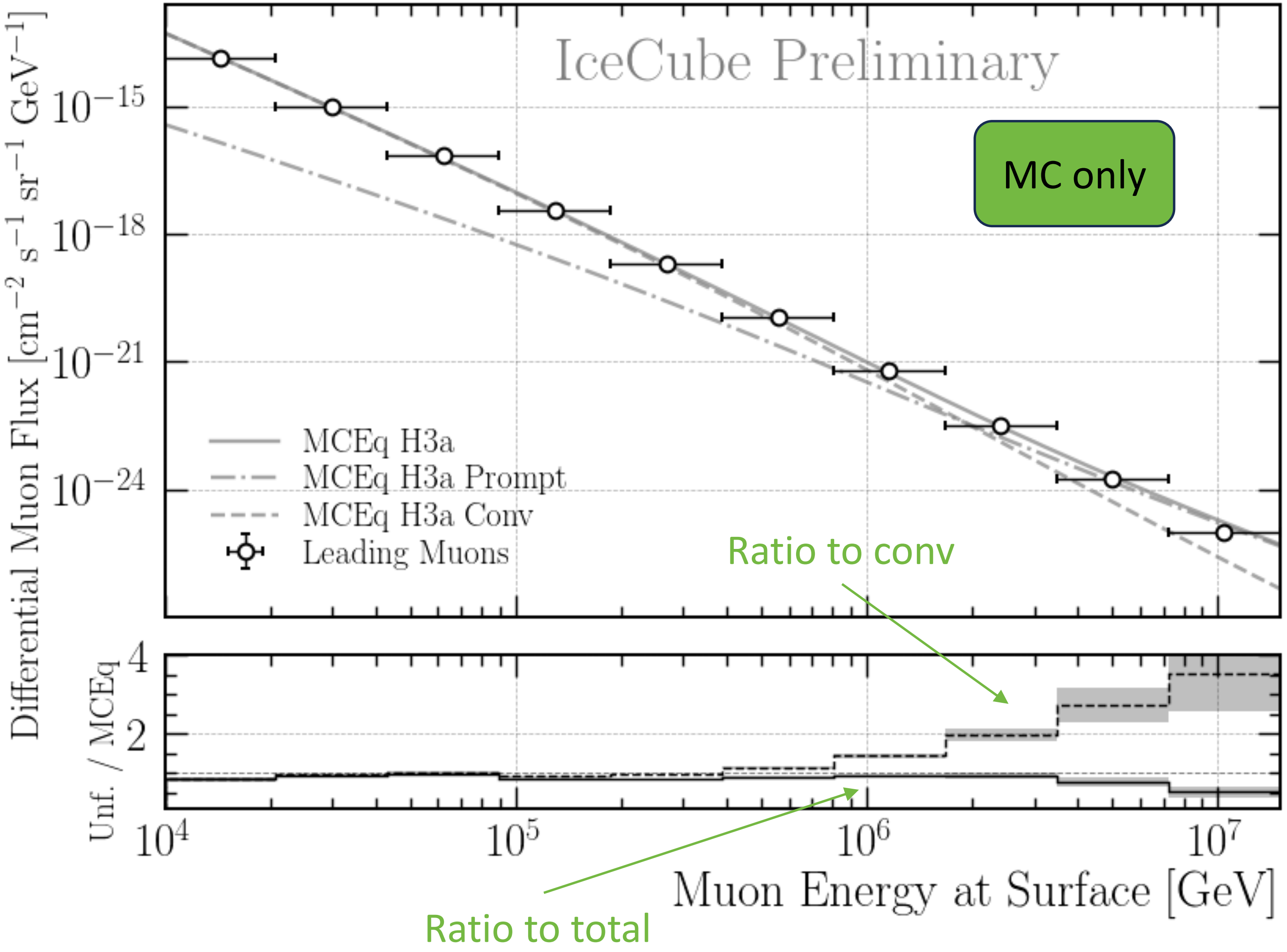


pascal.gutjahr@tu-dortmund.de

*Stopping muon analysis done by Lucas Witthaus

Future: 12 Years of IceCube Data Unfolding

- Finalize analysis
  - Proper systematic uncertainty estimation
- Unfold 12 years of data
- Extend energy range to 10 PeV
- Characterize the atmospheric muon flux
  - Determine prompt component

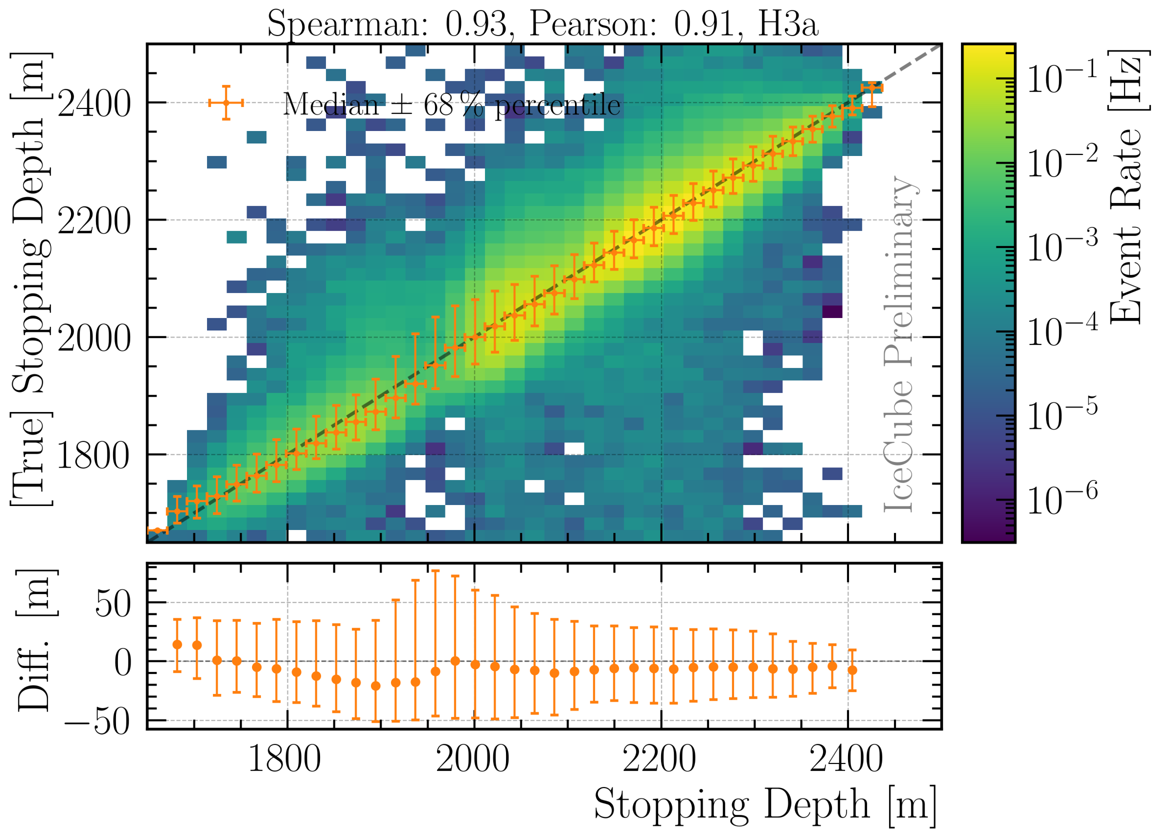


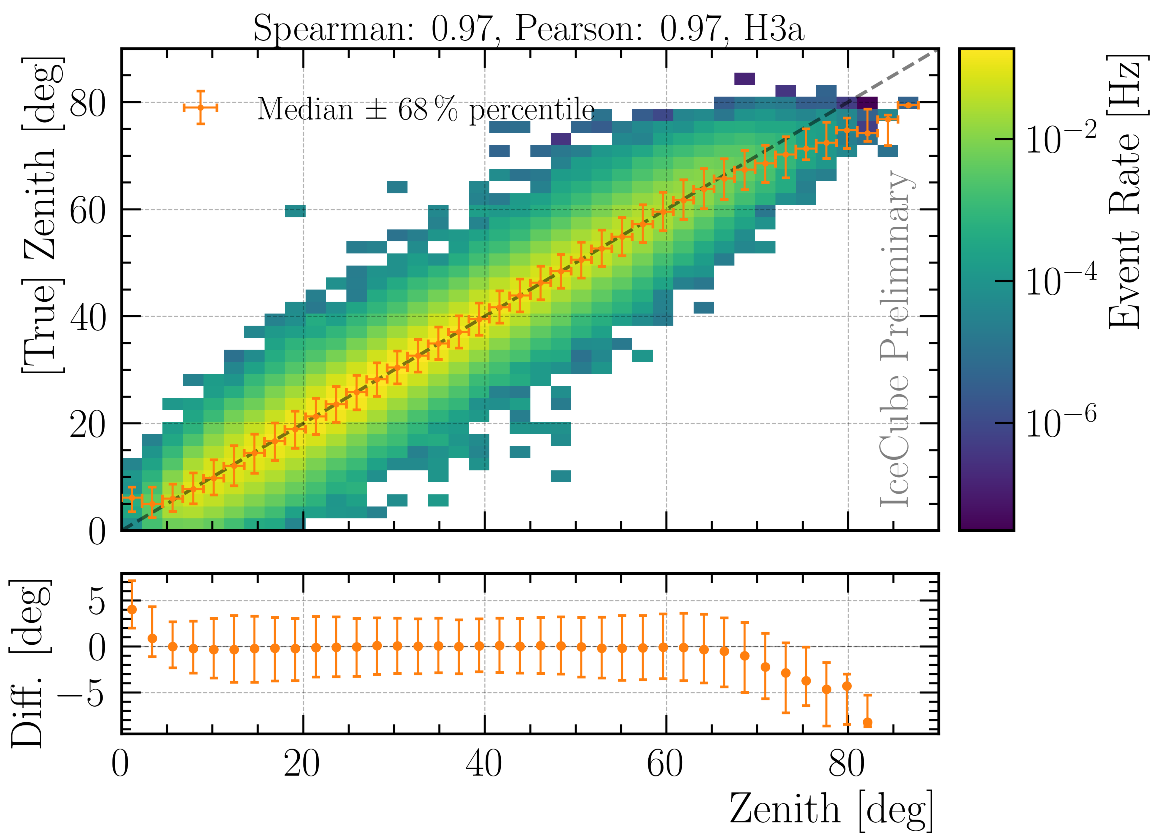
Depth Intensity Unfolding

Utilizing stopping muons

Reconstructions: Stopping Muons

- Good reconstruction of stopping depth and zenith angle
- Calculate propagation length → proxy variable for unfolding



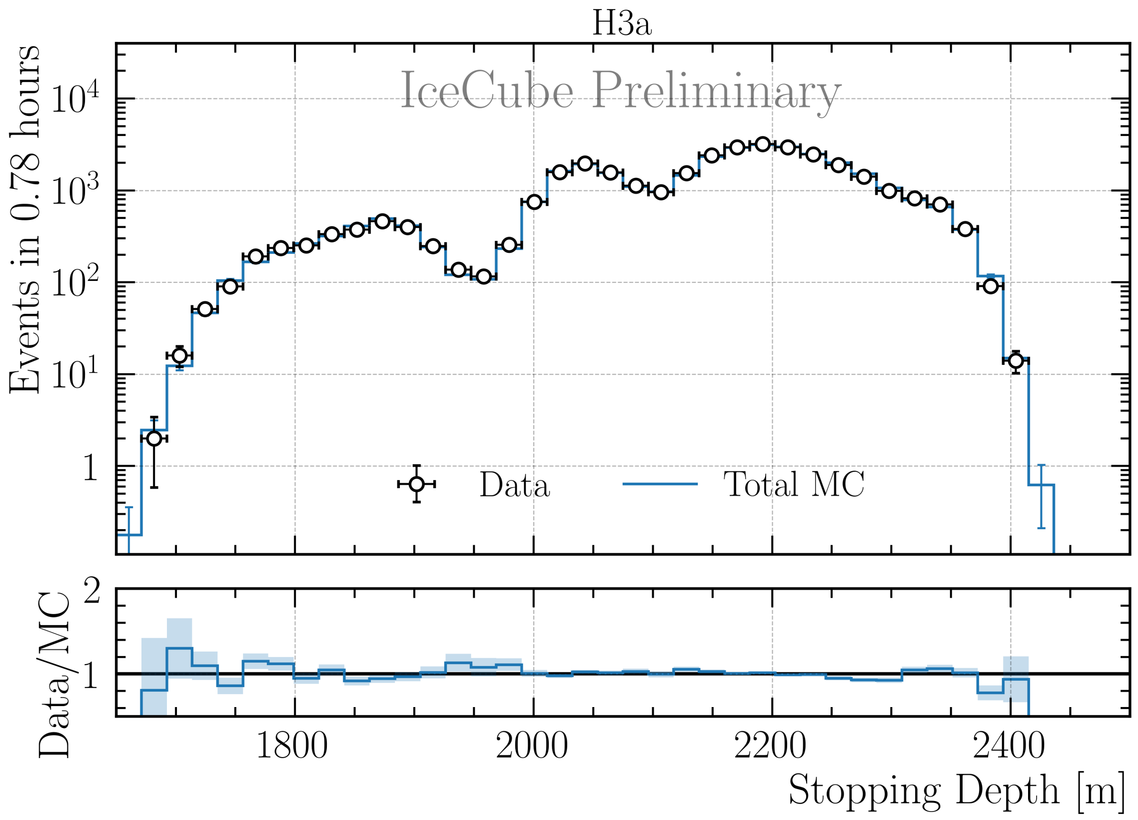


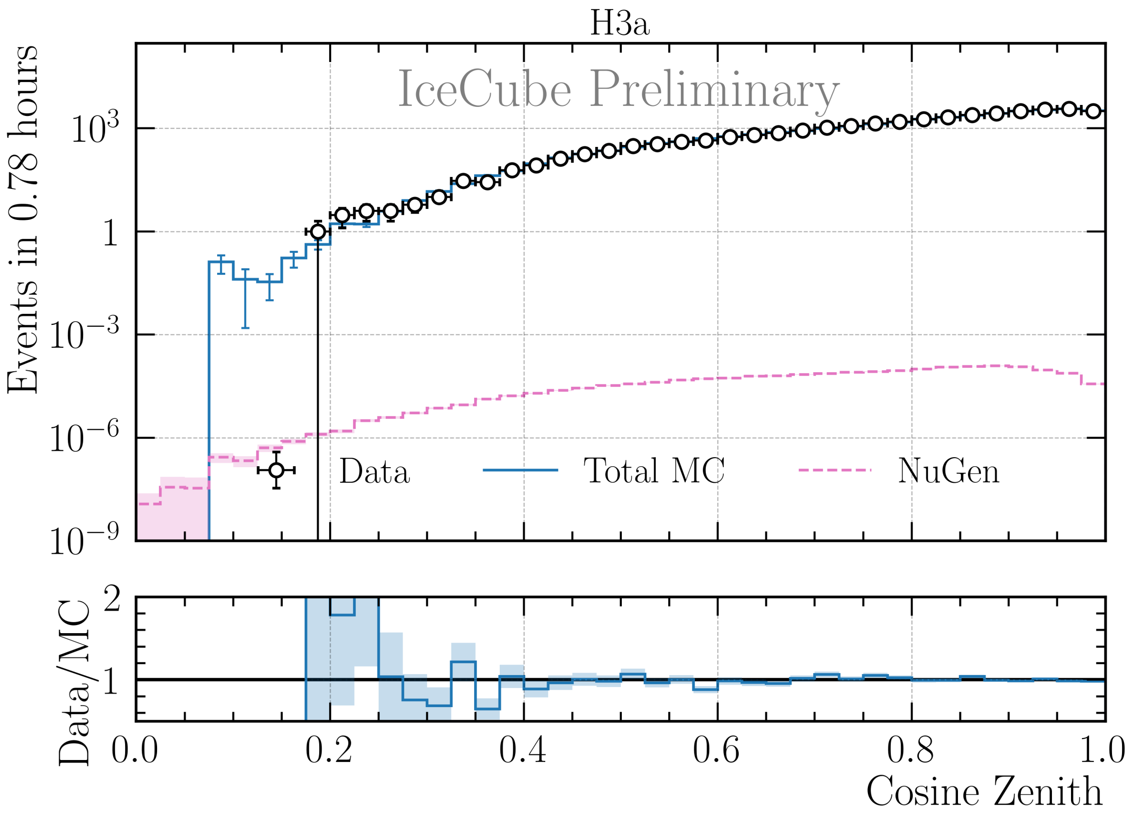
Data—MC: Stopping Muons

- Good data—MC agreement
- Global offset → upscale MC by 10%

Neutrino weighting:
SPL: n = 1.8, γ = 2.52

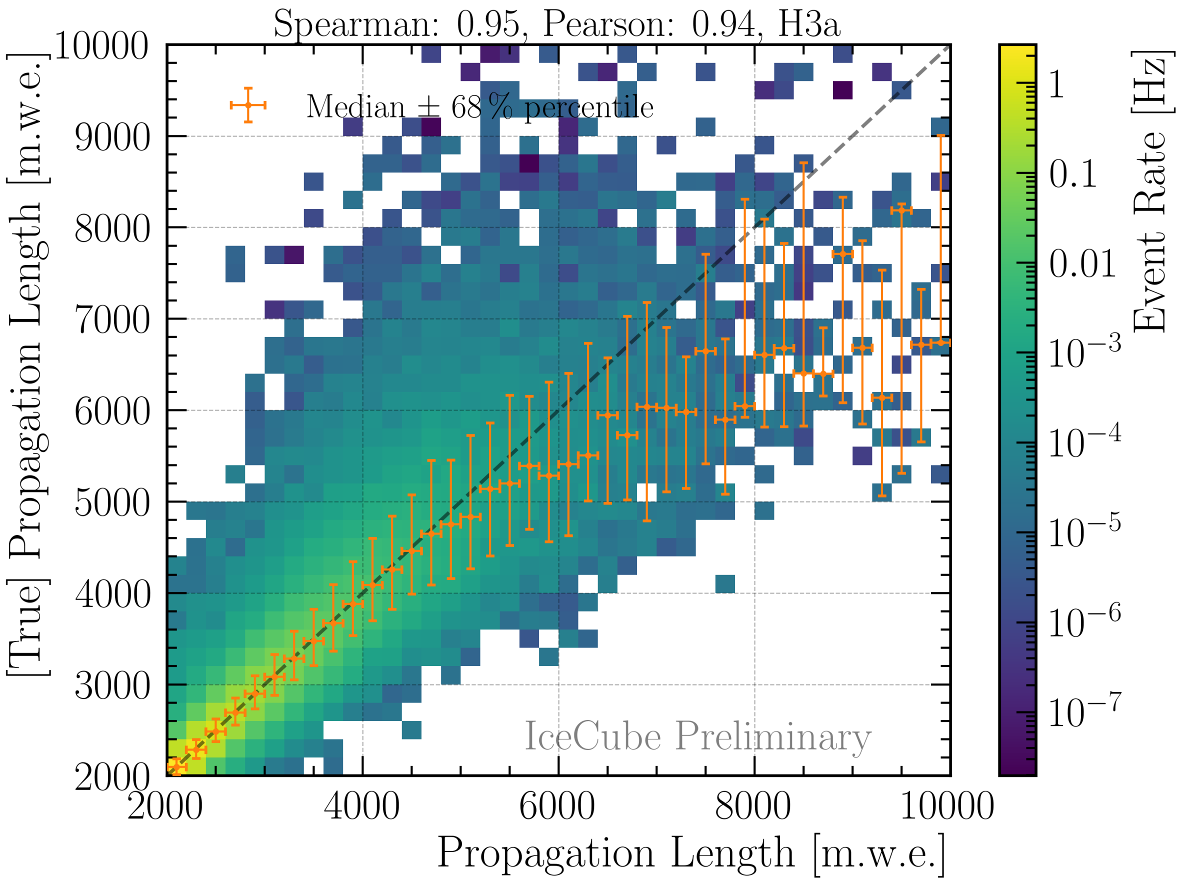
IceCube Collaboration PoS ICRC2023 1064



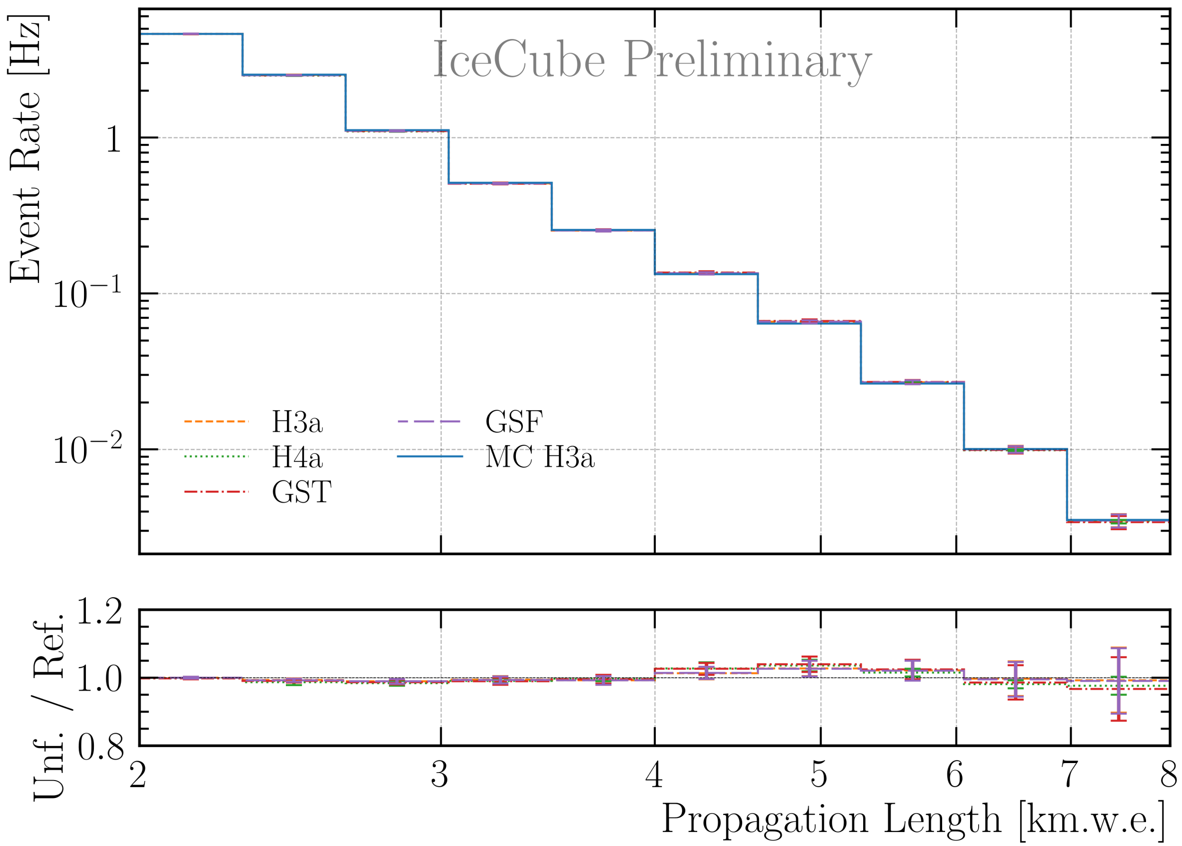


Correlation and Robustness Test: Stopping Muons

- Strong correlation between proxy and target



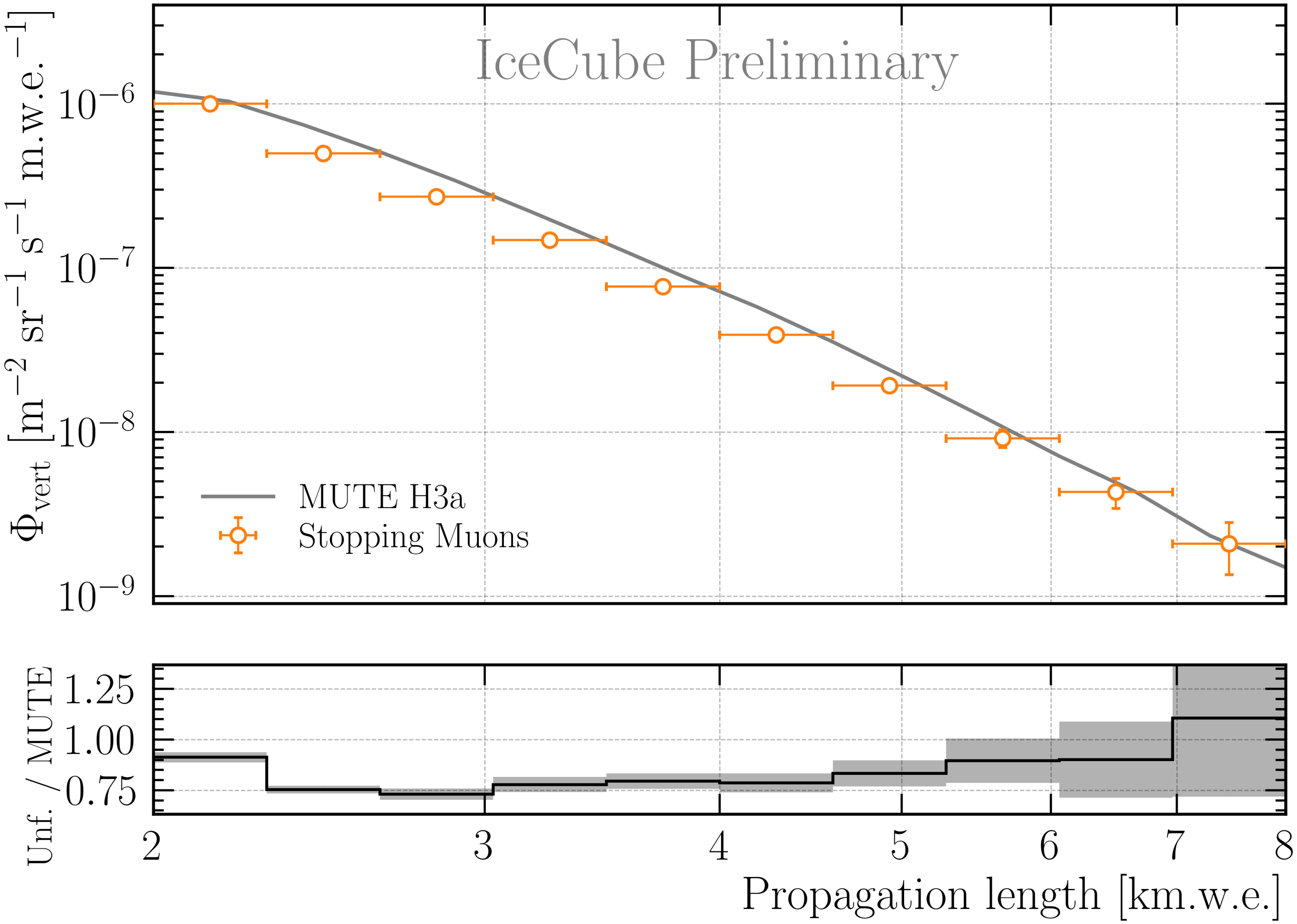
- Build unfolding matrix on H3a, H4a, GST & GSF
- Unfold H3a as test set
- Results are within uncertainties → robust



Depth Intensity Unfolding

- 47 min IceCube data
- 32943 events

- Comparison to MUTE[1]
- Up to 25 % below prediction at 2.5 – 5.5 km.w.e.
- Mismodelling in energy losses?

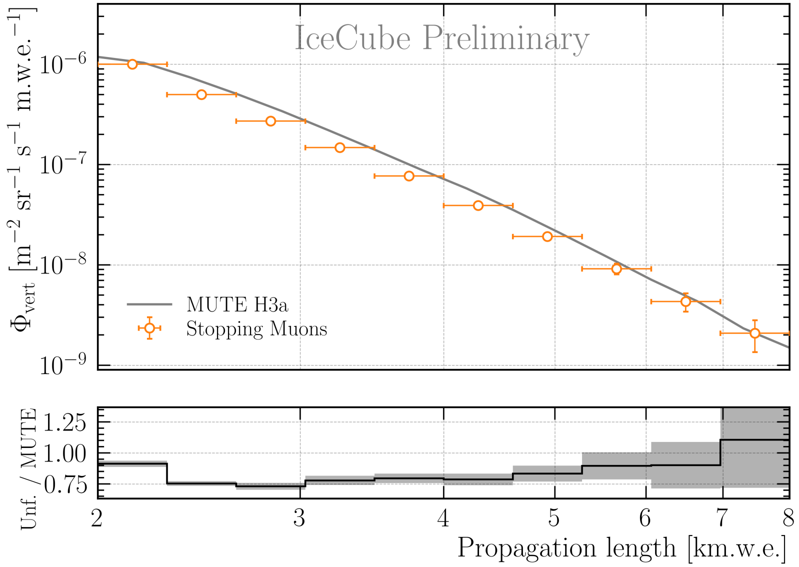


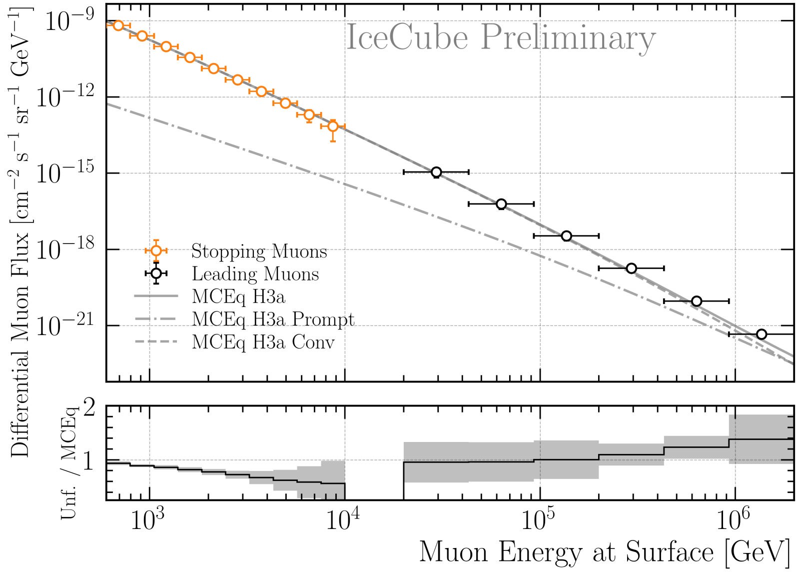
pascal.gutjahr@tu-dortmund.de

[1]W. Woodley et al. Phys. Rev. D 110 no. 6, (2024) 063006

Conclusion & Outlook

- DNN-based reconstructions performed
  - Stopping muon selection
  - Leading muon selection
  - Good Data—MC agreement
- Depth intensity unfolded using 47 min of IceCube data
  - Undershoots MUTE prediction up to 5.5 km.w.e.
- Muon flux at surface unfolded using stopping muons and 2487 h of leading muon IceCube data
  - Flux undershoots up to 10 TeV (MCEq prediction) – stopping muons
  - Flux agrees from 20 TeV to 2 PeV with MCEq – leading muons
    - Slope increases towards higher energies
    - Prompt and conventional flux contributes similar to the flux at ~2 PeV
- Finalize both analyses → unfold with more data
  - Extend energy range up to ~10 PeV





Backup

Muon Filter

MuonFilter The `I3MuonFilter_13` selects events based on reconstructed track quality and charge brightness, using criteria that depend on the reconstructed zenith angle. It requires three inputs: a cleaned pulse series map, a reconstructed track, and the log-likelihood fit parameters associated with the track. From the pulse series map, the number of hit DOMs N_{ch} and the total charge Q_{tot}, defined as the sum of the pulse charges over all DOMs, is determined. Then, the logarithmic charge $\ell = \log_{10}(Q_{\text{tot}})$ is calculated. The cosine of the reconstructed zenith angle is computed as $\cos(\theta)$, and the per-DOM log-likelihood value is retrieved from the fit parameters.

Events are classified into three angular zones based on the value of $\cos(\theta)$:

- Zone 1: $-1.0 < \cos\theta \leq 0.2$
- Zone 2: $0.2 < \cos\theta \leq 0.5$
- Zone 3: $0.5 < \cos\theta \leq 1.0$

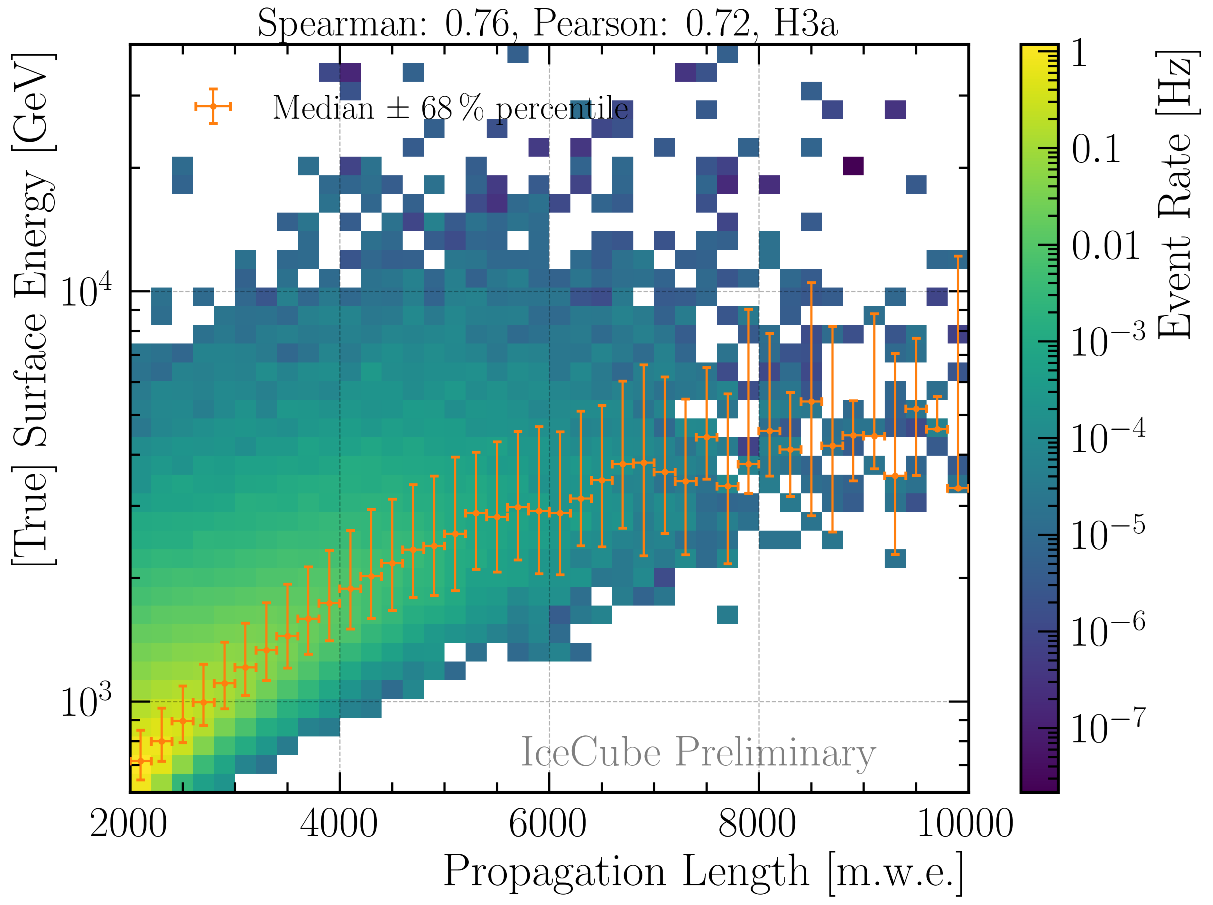
Each zone applies a different selection criterion:

- Zone 1: the filter computes a scaled log-likelihood value $\tilde{\lambda} = \text{rlogl} \cdot \frac{N_{\text{ch}}-5}{N_{\text{ch}}-3}$, and accepts the event if $\tilde{\lambda} \leq 8.7$.
- Zone 2: the filter applies a brightness-based cut and accepts the event if $\ell > 3.9 \cdot \cos\theta + 0.65$.
- Zone 3: a similar brightness cut is used with a shallower slope $\ell > 0.6 \cdot \cos\theta + 2.3$.

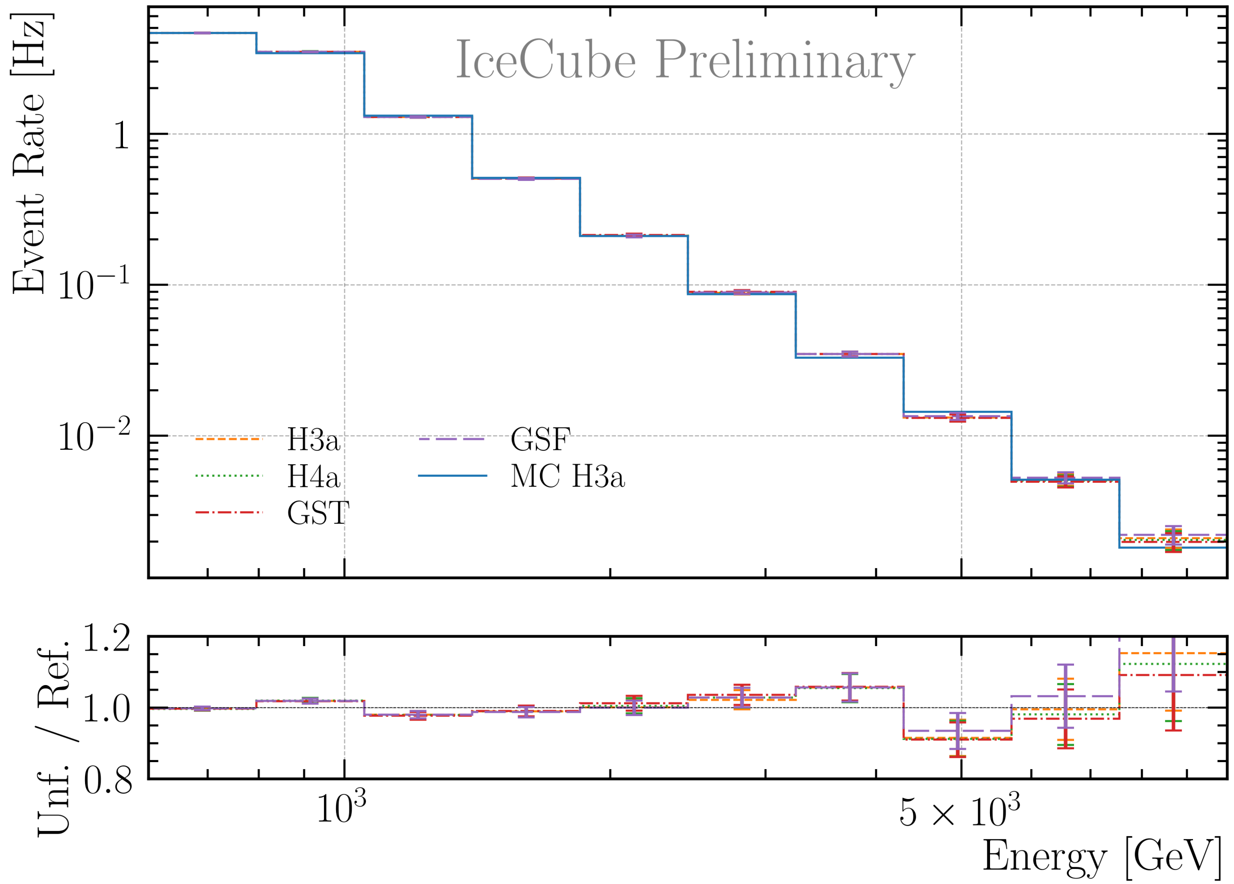
The event is accepted if it passes the selection criteria for any of the three zones. This design ensures that events are selected based on fit quality in the near-horizontal region and on brightness in the vertical direction, balancing background suppression with high signal efficiency across the full zenith range.

Correlation and Robustness Test: Stopping Muons

- Correlation between proxy and target

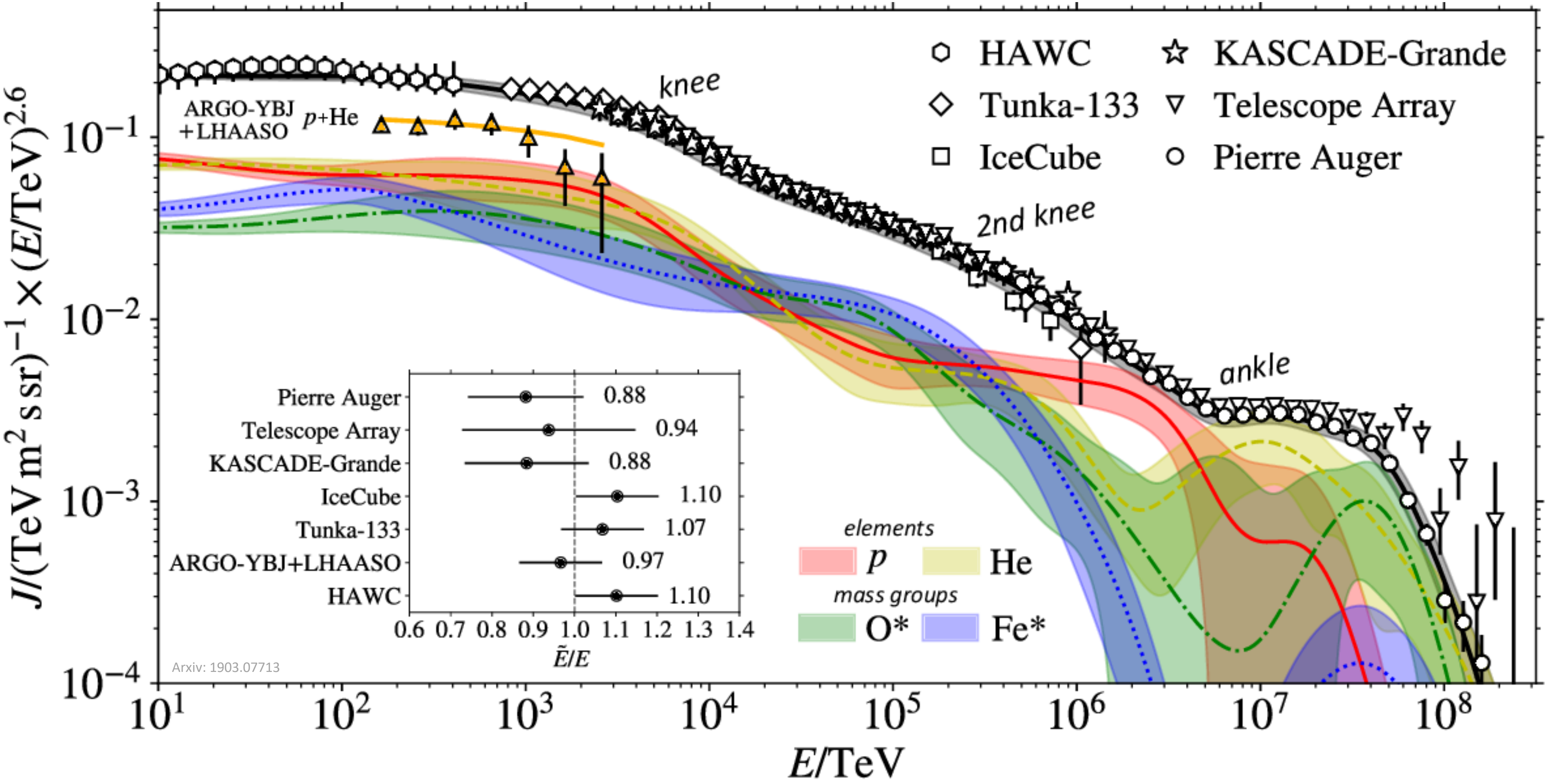


- Build unfolding matrix on H3a, H4a, GST & GSF
- Unfold H3a as test set
- Results are within uncertainties → robust

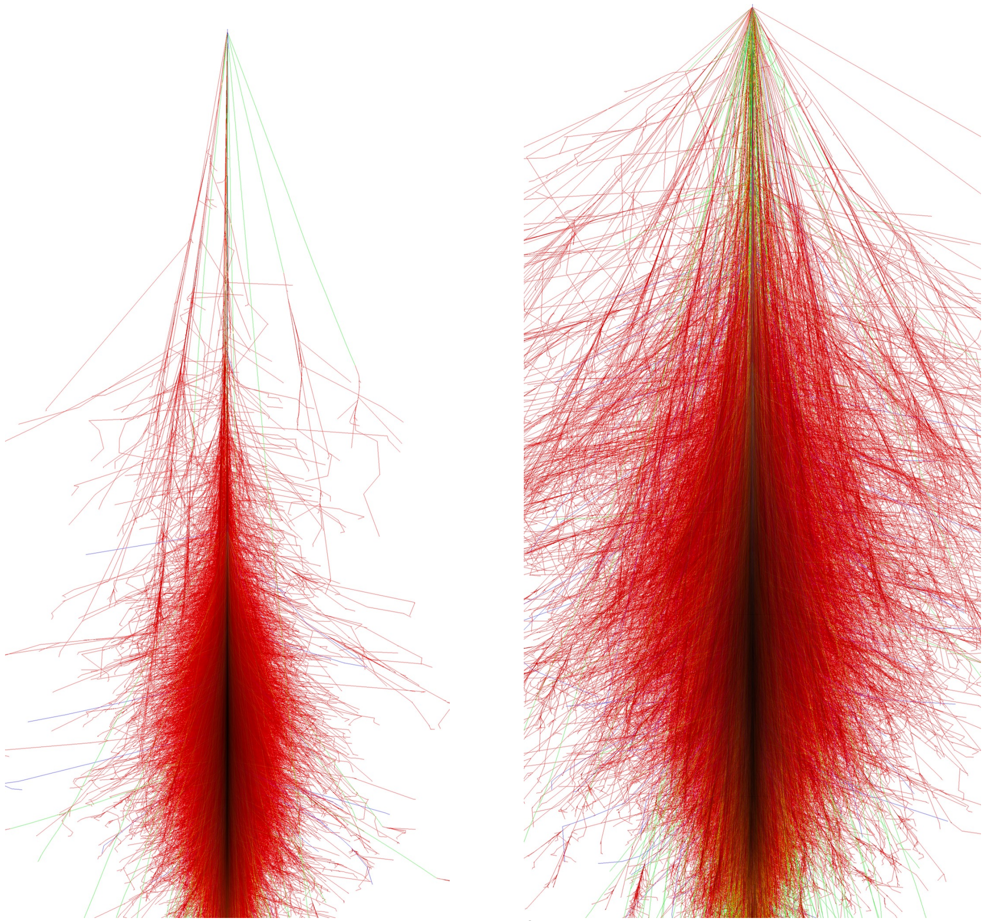


pascal.gutjahr@tu-dortmund.de

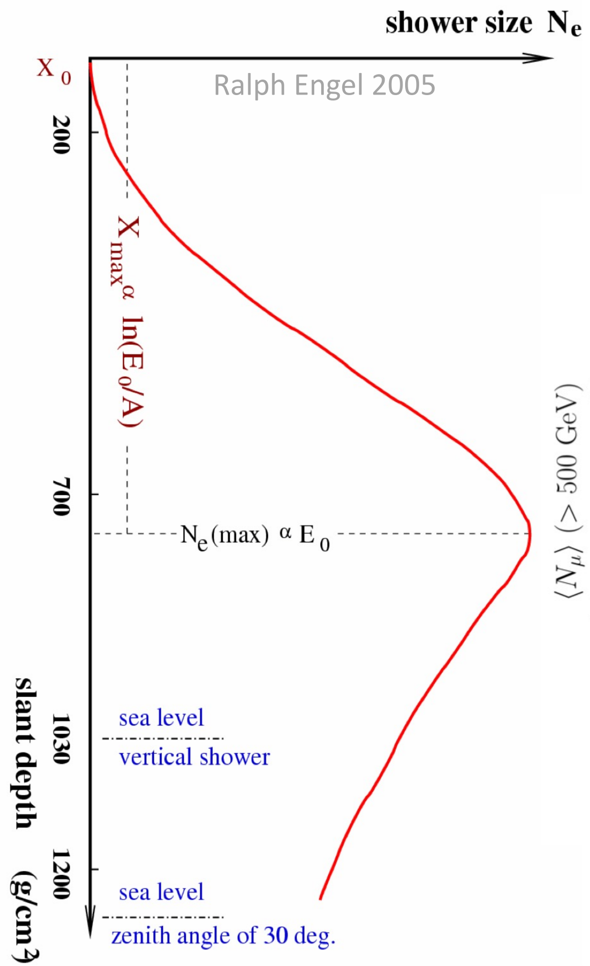
Cosmic ray flux

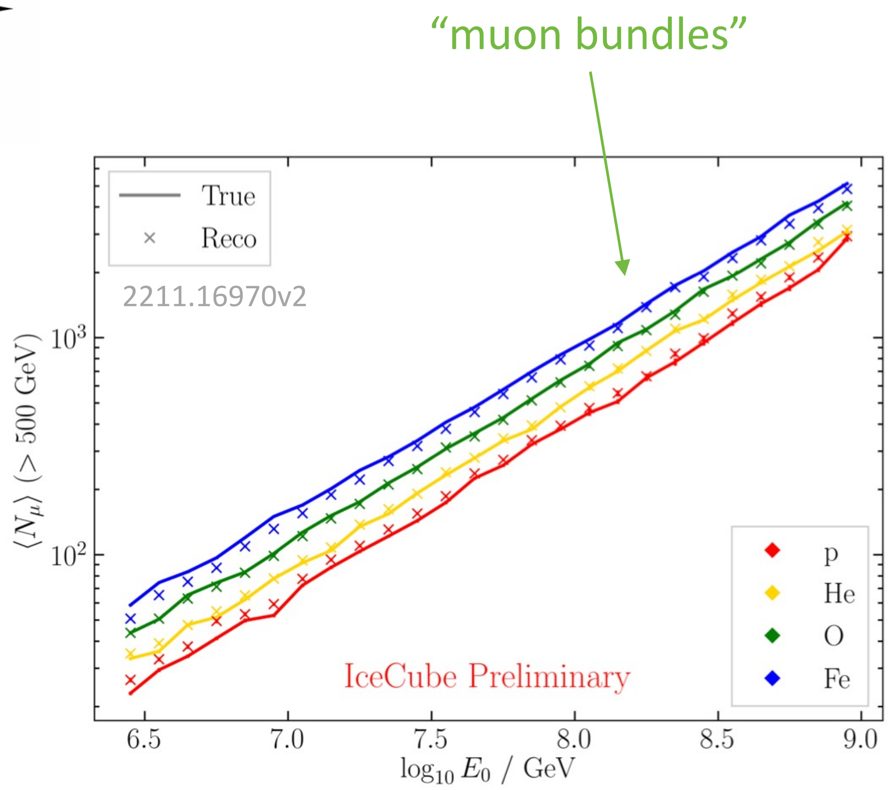


Air shower – 10 TeV



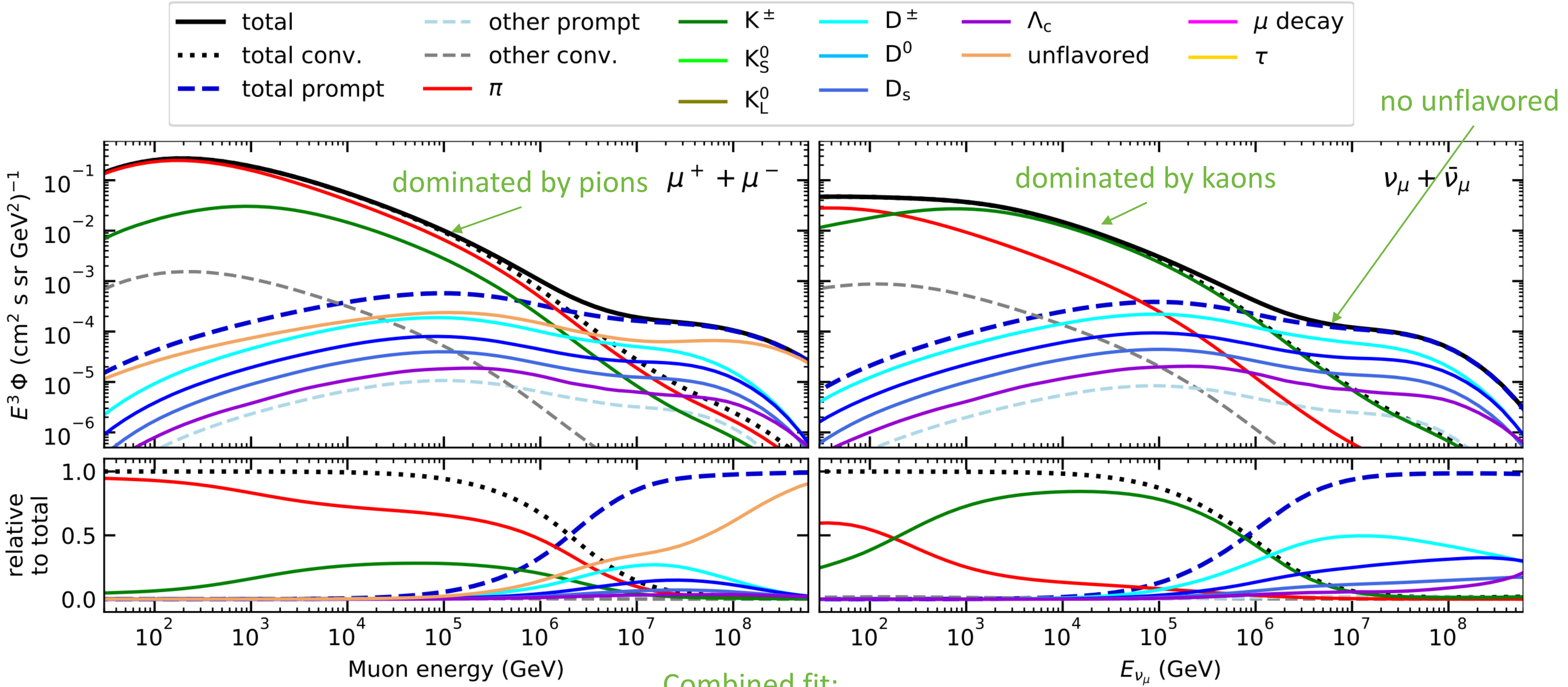
Proton Iron





Prompt atmospheric muons and neutrinos

10.1103/PhysRevD.100.103018



Combined fit:

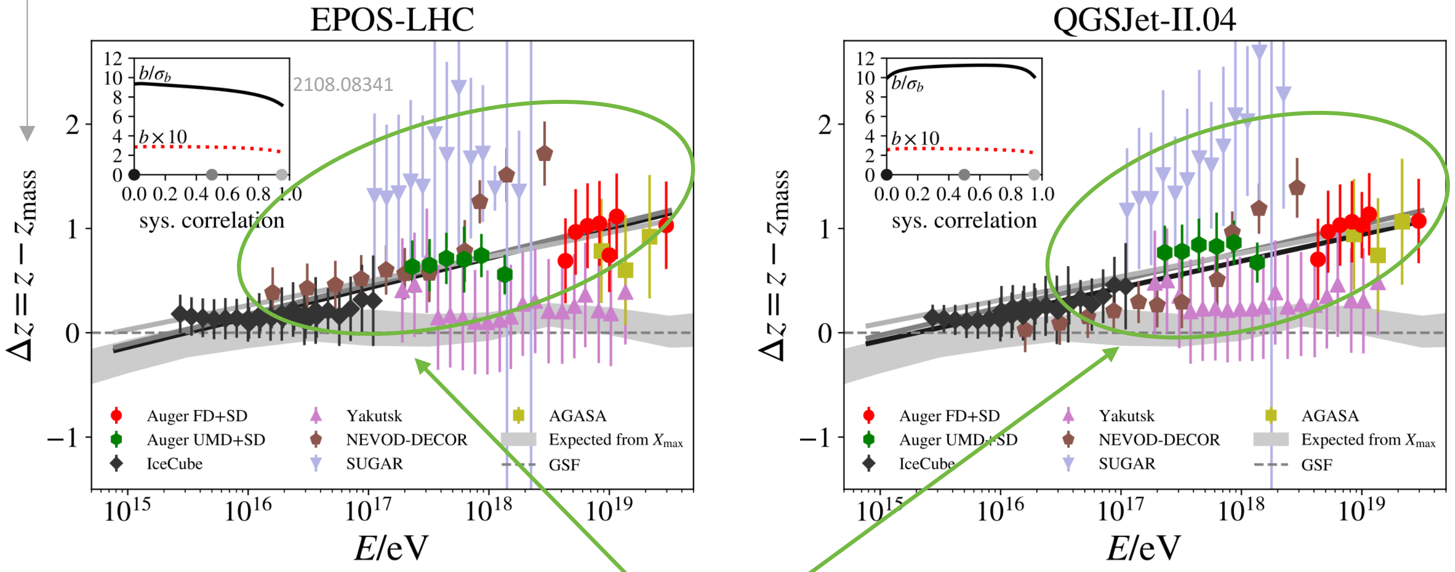
- handle on pion/kaon ratio
- handle on charmed mesons

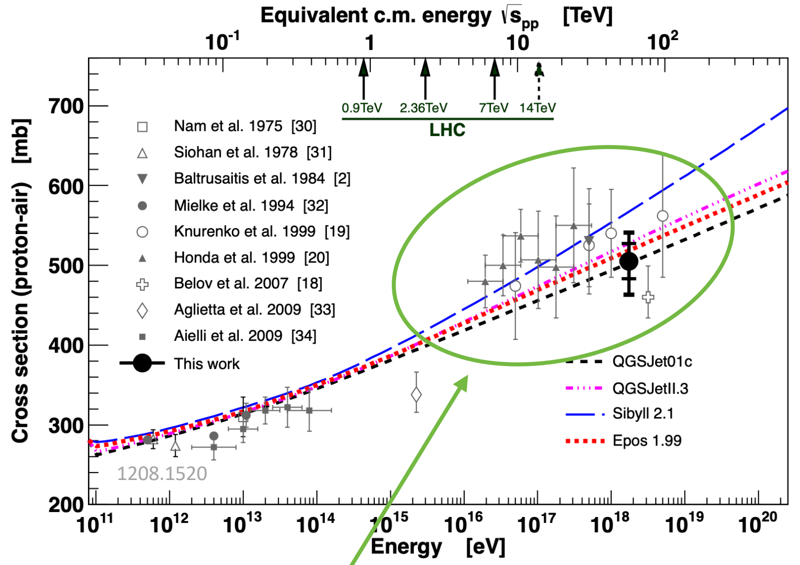
Muon puzzle and model uncertainties

"muon number"

Expected z ("muon number")

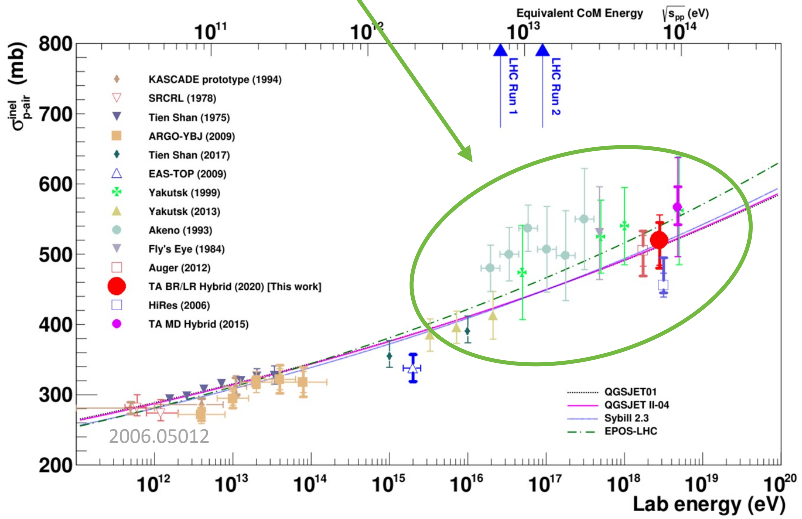
$$z = \frac{\ln\langle N_\mu \rangle - \ln\langle N_\mu \rangle_p}{\ln\langle N_\mu \rangle_{\mathrm{Fe}} - \ln\langle N_\mu \rangle_p}$$





2012

➢ Uncertainties at E > 10 PeV

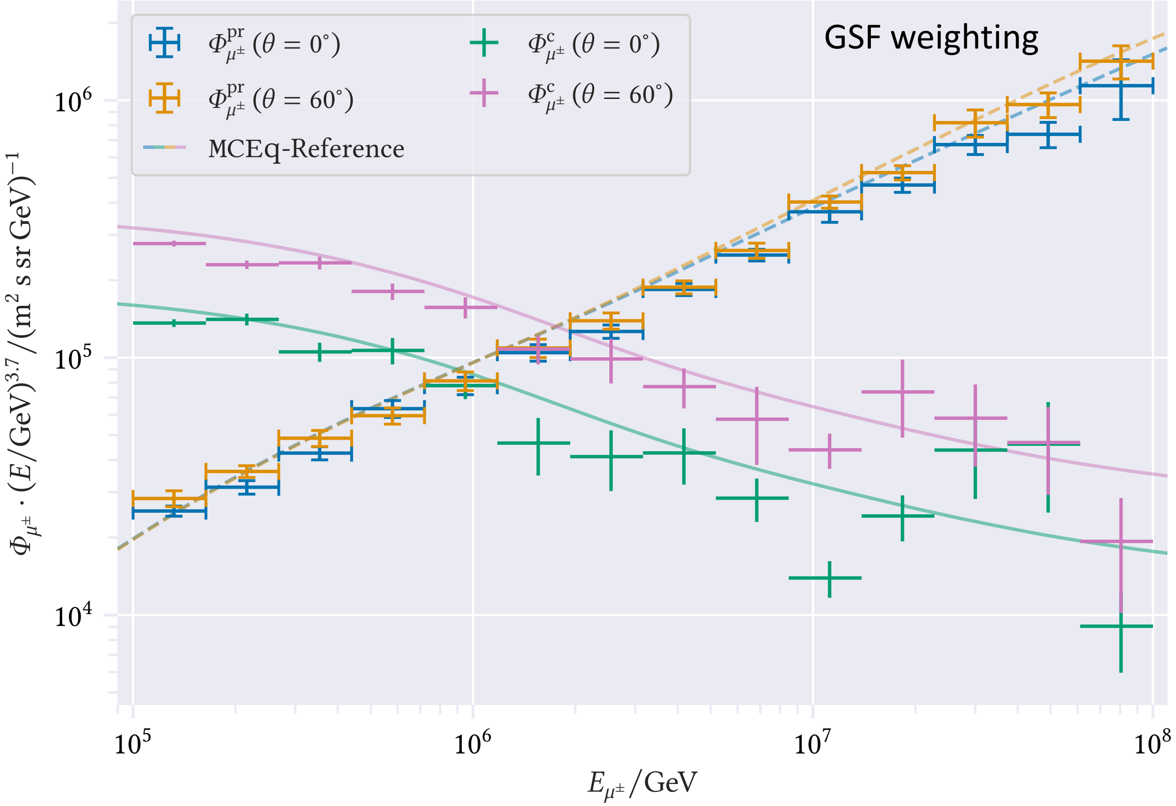


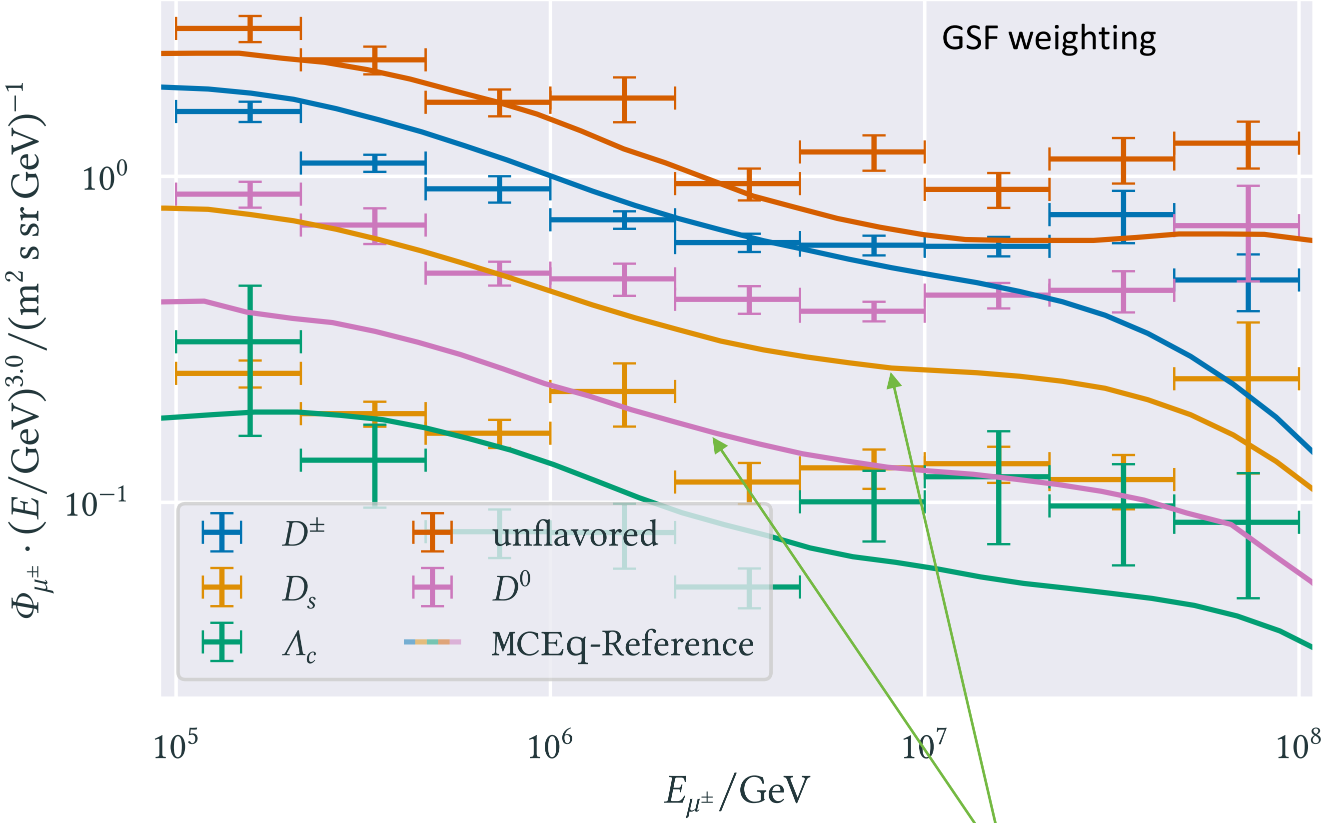
2020

➢ More muons measured than simulated for E > 40 PeV ~ cms 8 TeV

CORSIKA 7 tagging and MCEq comparison

MCEq: tool to numerically solve the cascade equations that describes the evolution of particle densities as they propagate through a gaseous, dense medium
https://github.com/mceq-project/MCEq

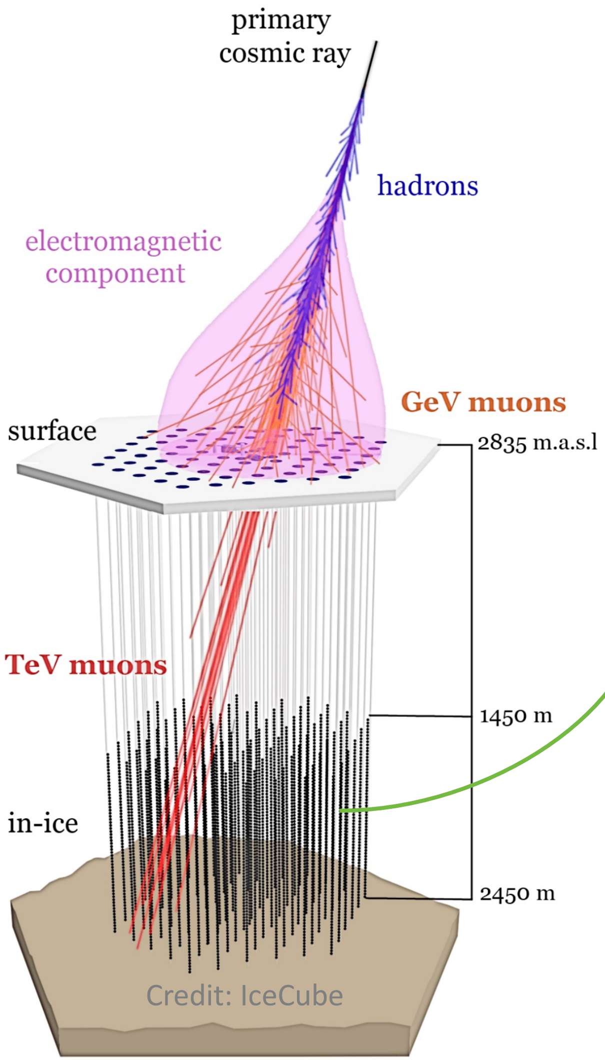


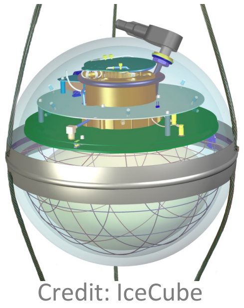


- Good agreement in total prompt and conv muon flux
- D^0 and D_s are swapped here but this is fixed in MCEq

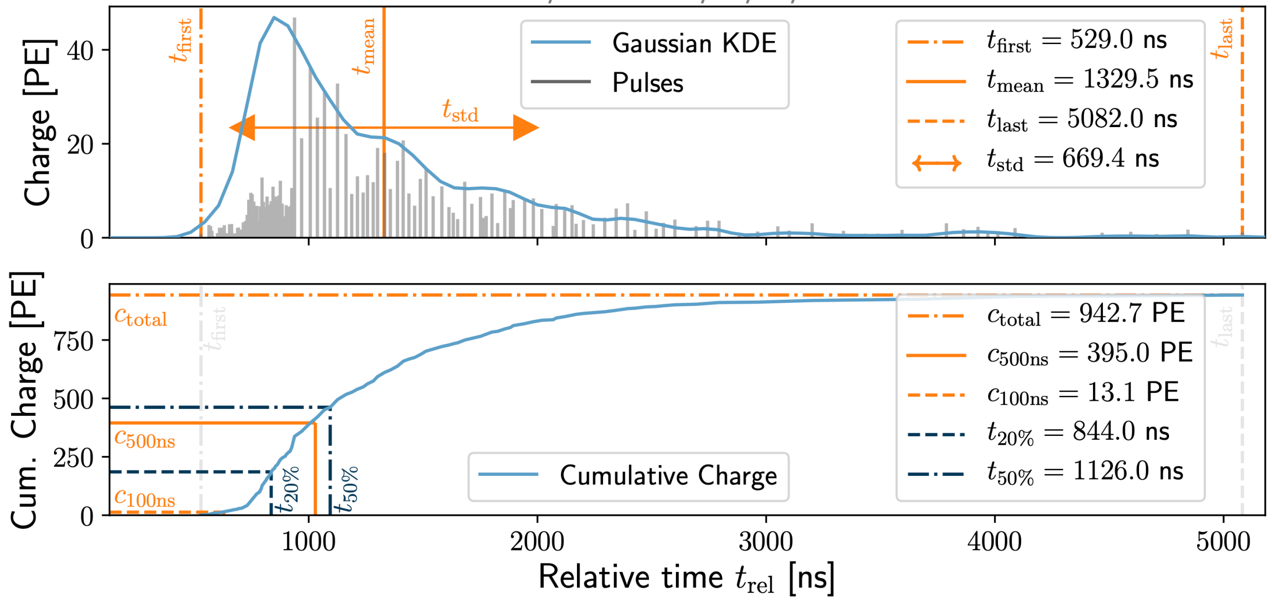
mention panama

Machine Learning Reconstructions





10.1088/1748-0221/16/07/P07041



Convolutional Neural Network

Inputs

- c_{total}: Total charge
  - Sum of charge
- t_{first}: Relative time of first pulse
  - Relative to total time offset
- t_{std}: Standard deviation of first pulse
  - Charge weighted standard deviation of pulse times

Outputs

- Direction
- Stopping point
- Entry point
- Energy at entry/surface
- …

Level5: quality cuts

containment cuts	>	<
length in detector	1000 m	2000 m
entry pos x, y	-750 m	750 m
entry pos z	-500 m	750 m
center pos x, y	-550 m	550 m
center pos z	-650 m	650 m

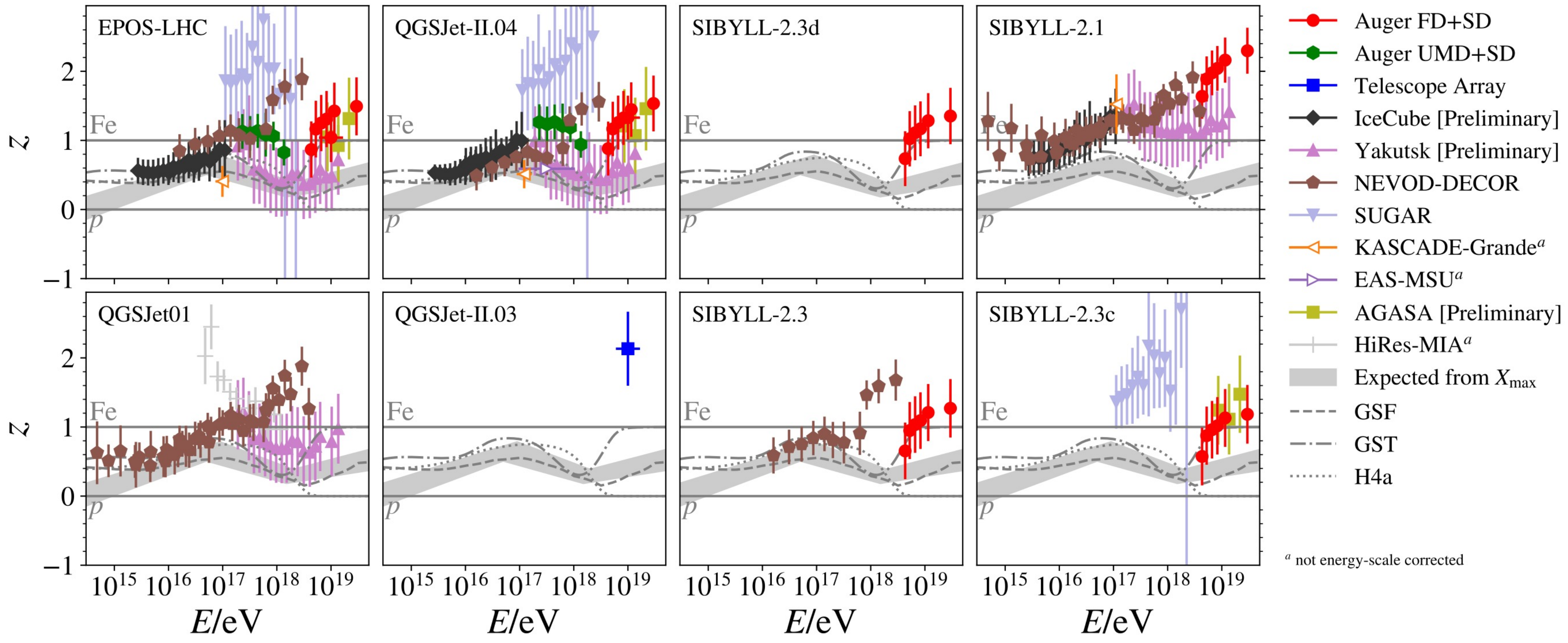
neutrino cuts	>	<
cos(zenith)	0.2	
length	5000 m	15000 m

uncertainty cuts	<
bundle energy at entry	0.9 log10(GeV)
bundle energy at surface	2.0 log10(GeV)
zenith	0.1 rad
azimuth	0.2 rad
entry pos x , y, z	42 m
center pos x, y, z	50 m
entry pos time	200 ns
center pos time	600 ns
length in detector	160 m
length	2000 m

The Muon Puzzle

”muon number”

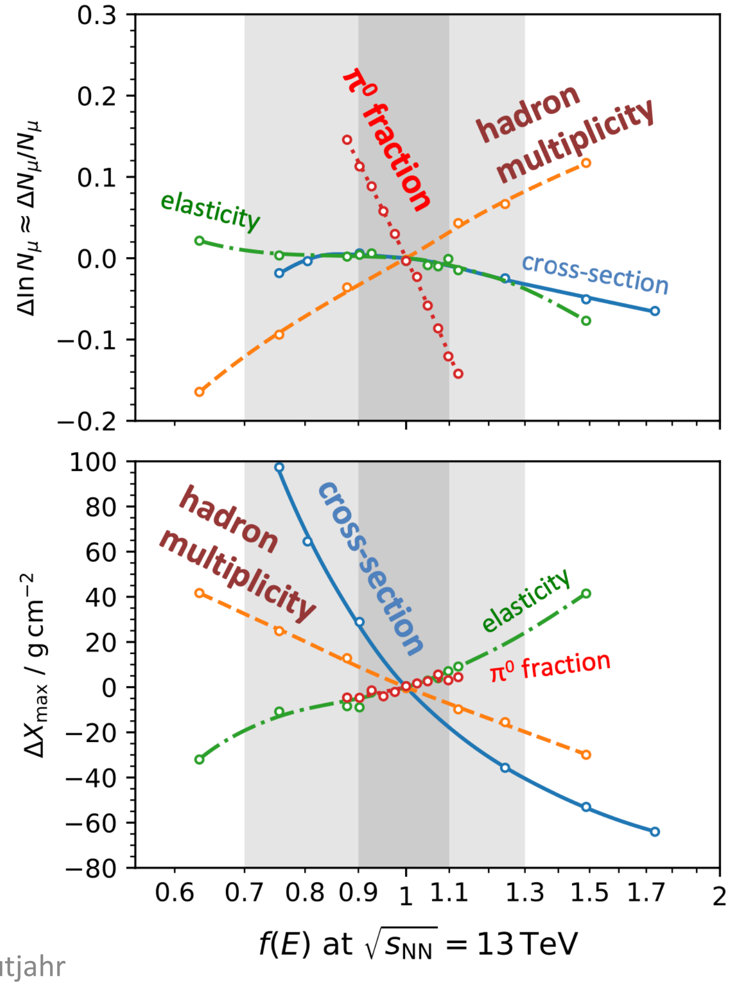
$$z = \frac{\ln\langle N_\mu\rangle - \ln\langle N_\mu\rangle_p}{\ln\langle N_\mu\rangle_{\mathrm{Fe}} - \ln\langle N_\mu\rangle_p}$$



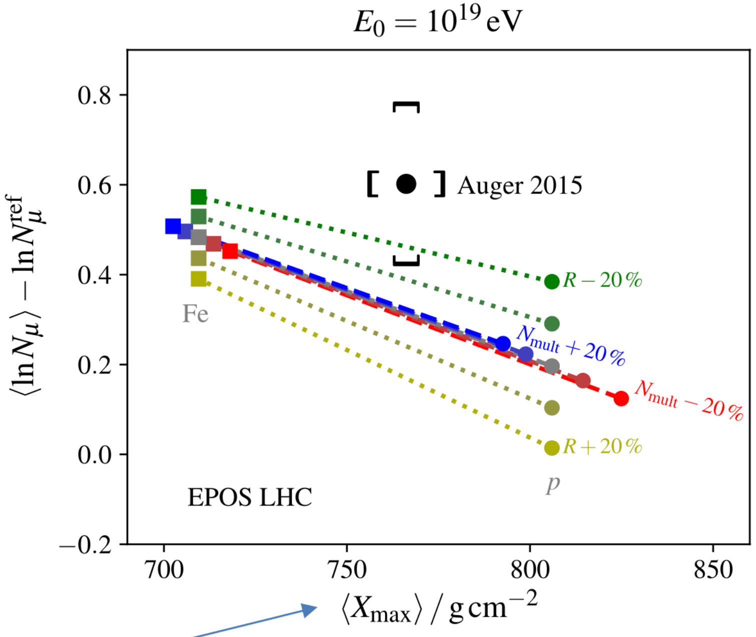
Arxiv: 2108.08341

Possible Solutions

R. Ulrich, R. Engel, M. Unger, PRD 83 (2011) 054026



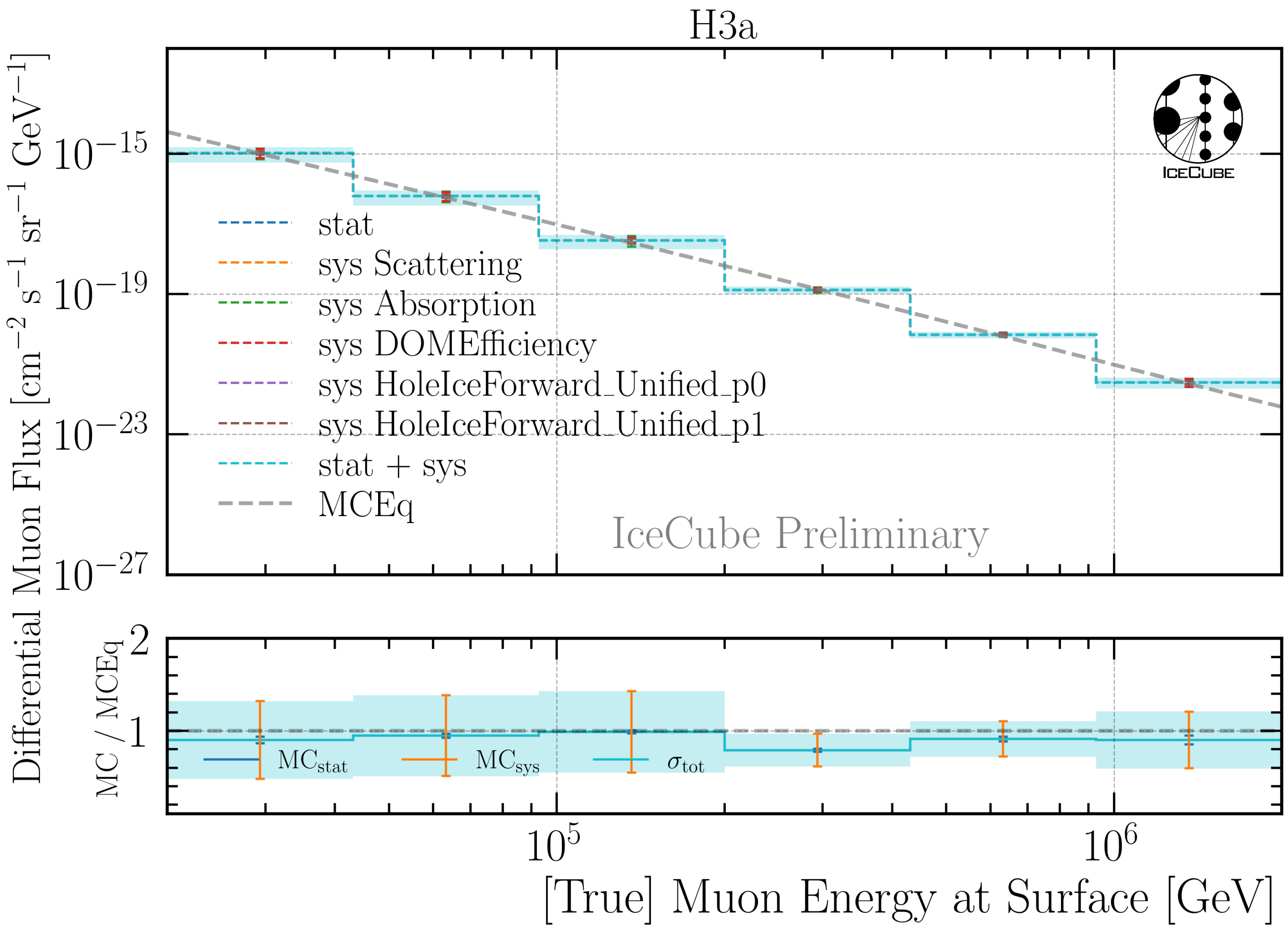
S. Baur, HD, M. Perlin, T. Pierog, R. Ulrich, K. Werner, arXiv:1902.09265



$$R = \frac{E_{\pi^0}}{E_{\text{other hadrons}}}$$

- Only changes to *R* can solve muon puzzle
- Small changes have large effect, *R* needs to be known to about 5 %

Slide taken from Hans Dembinski



H3a
Differential Muon Flux [cm^{-2} s^{-1} sr^{-1} GeV^{-1}]
10^{-15}
10^{-19}
10^{-23}
10^{-27}
stat
sys Scattering
sys Absorption
sys DOMEfficiency
sys HoleIceForward_Unified_p0
sys HoleIceForward_Unified_p1
stat + sys
MCEq
IceCube Preliminary
IceCube
MC / MCEq
2
1
MC_{stat}
MC_{sys}
σ_{tot}
10^5
10^6
[True] Muon Energy at Surface [GeV]

- present several reconstructions used for the selection
- pre-cut bundle energy, leading energy, direction, track geometry,…

- explain selection: what filters and what cuts?
- both for stopping muons and leading muons

- ludwigs thesis: different prompt definitions