Cosmic-Ray Composition Analysis
@ IceCube Observatory

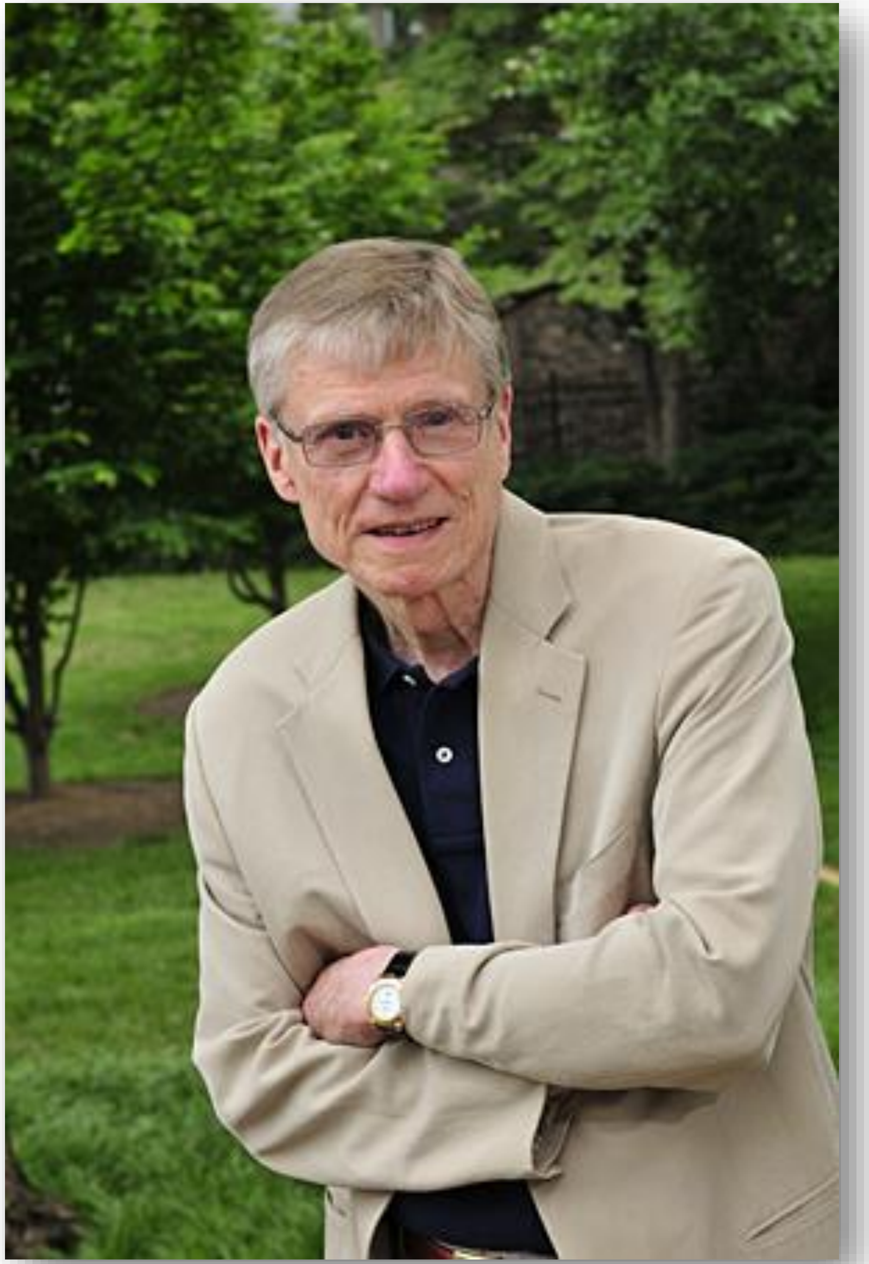
Paras Koundal
for IceCube Group @ KIT
paraskoundal.com

KSETA Plenary Workshop, 2022

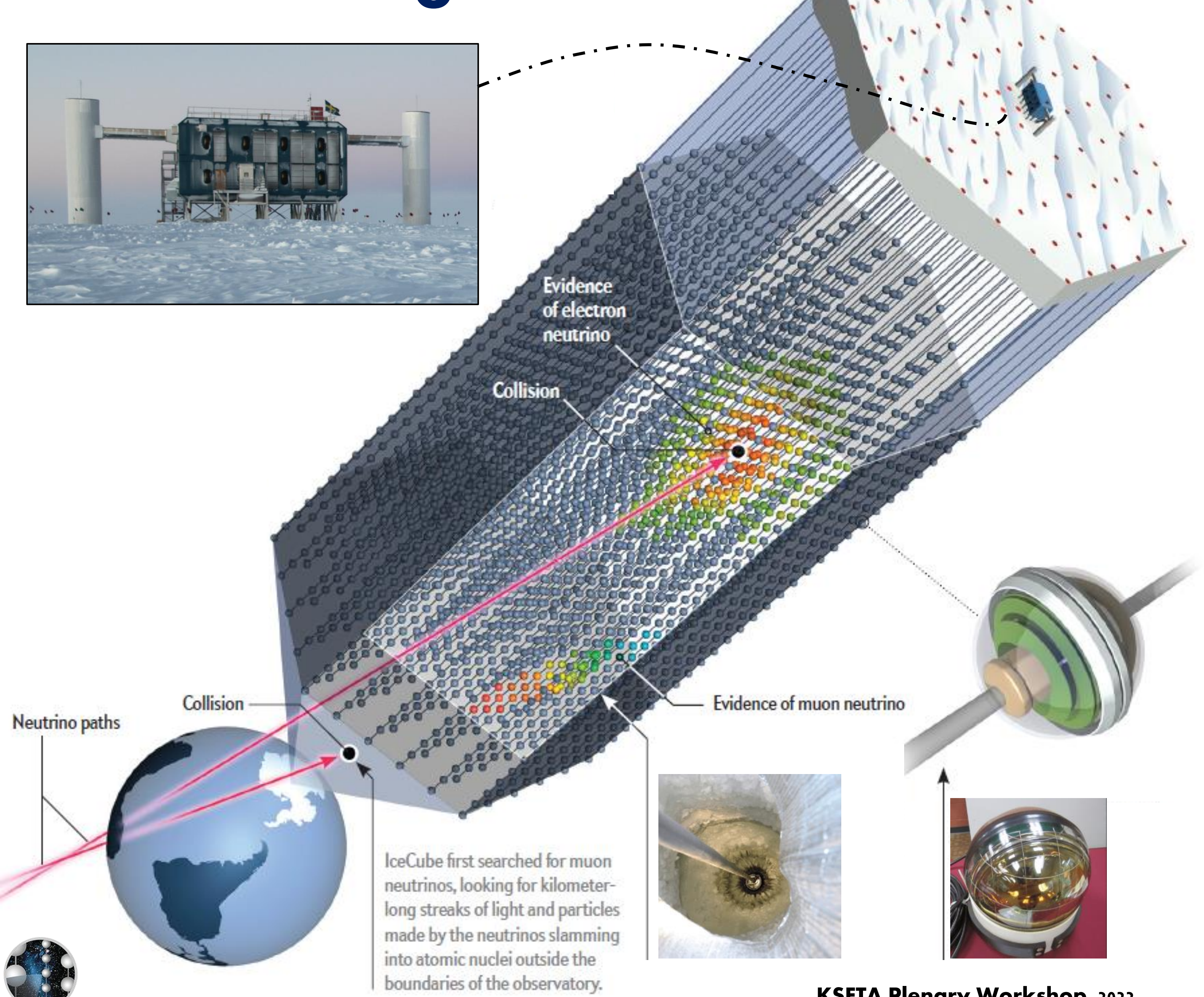
IceCube Collaboration by NSF

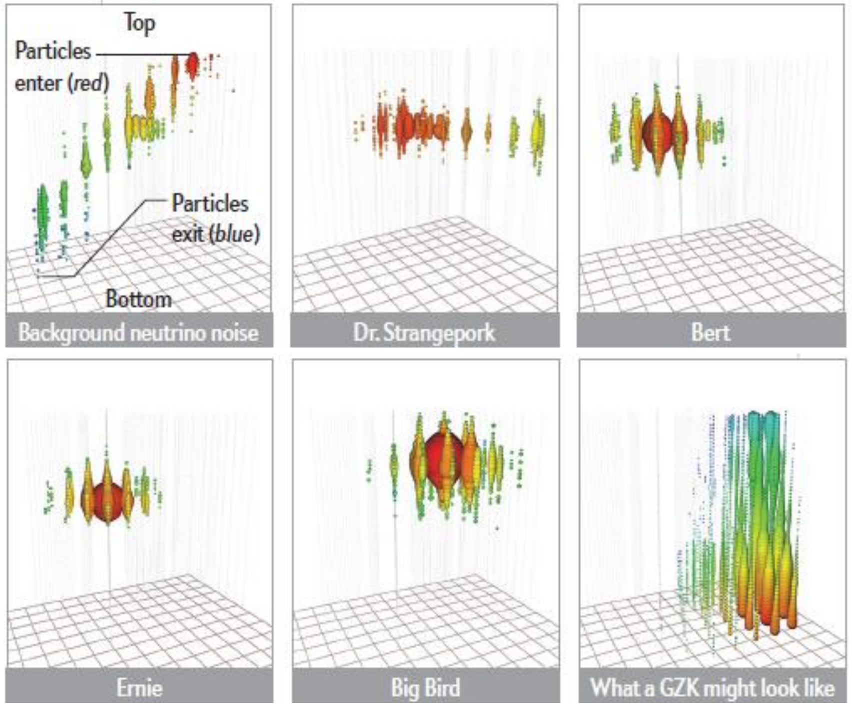
In Memory of

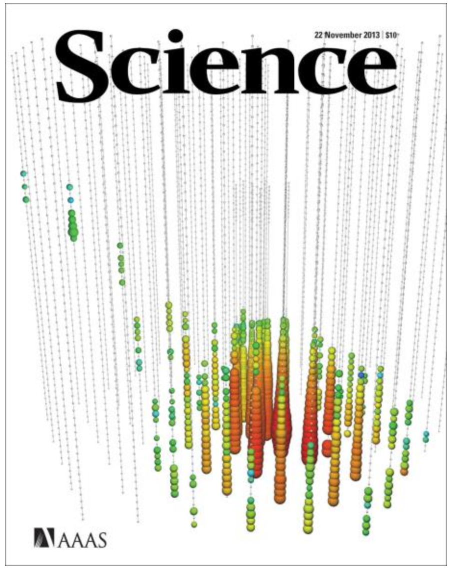
Thomas K. Gaisser



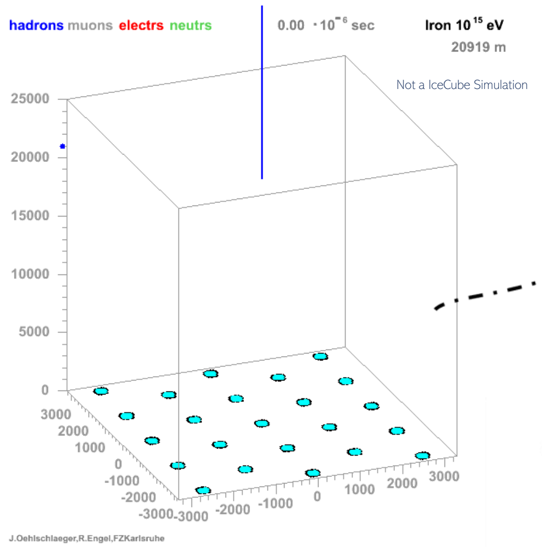
Ghost Hunting @ IceCube

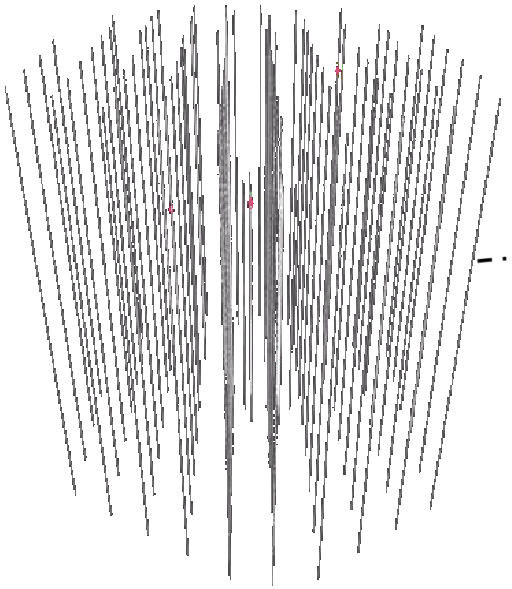


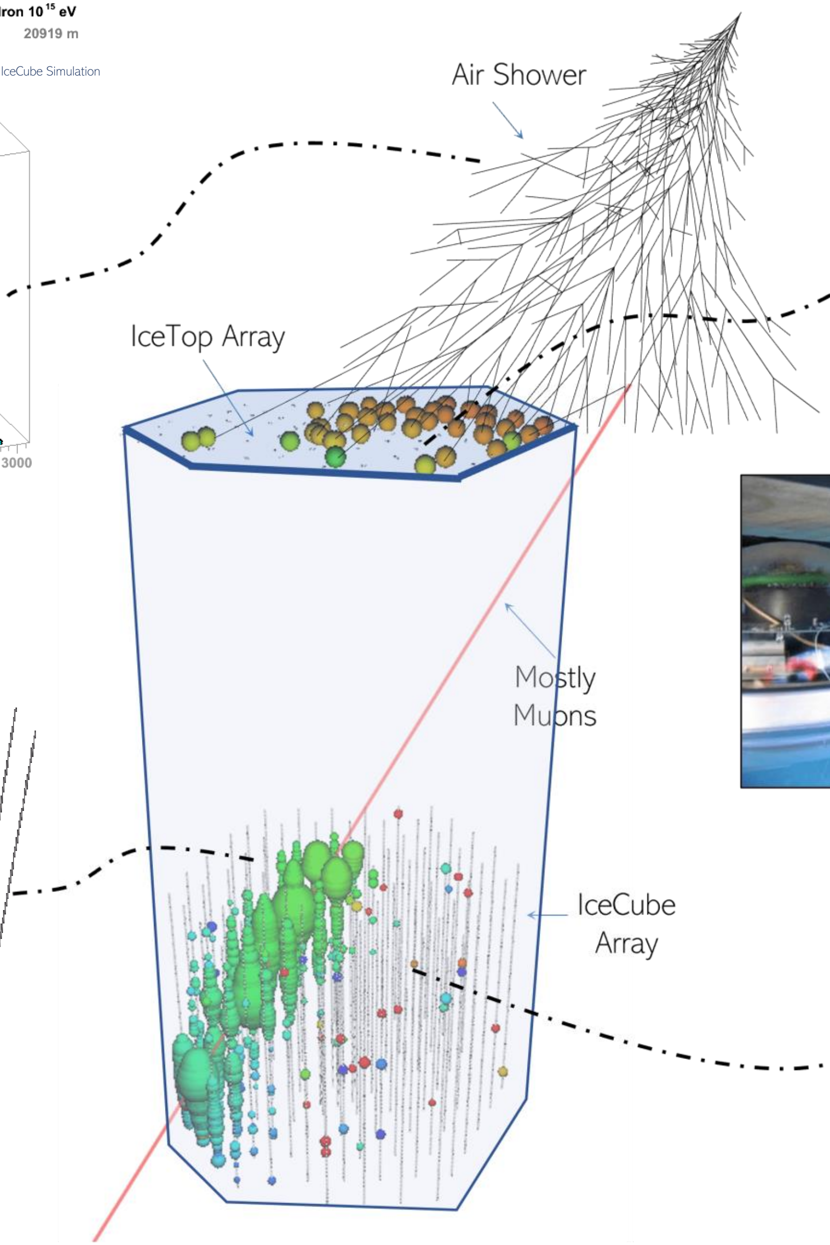


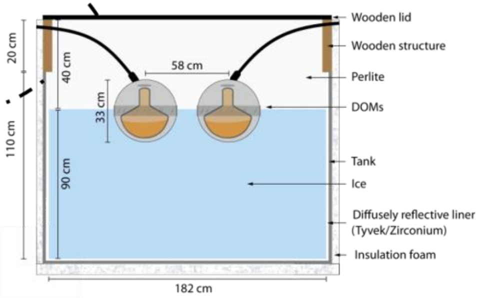


CR Analysis @ IceCube





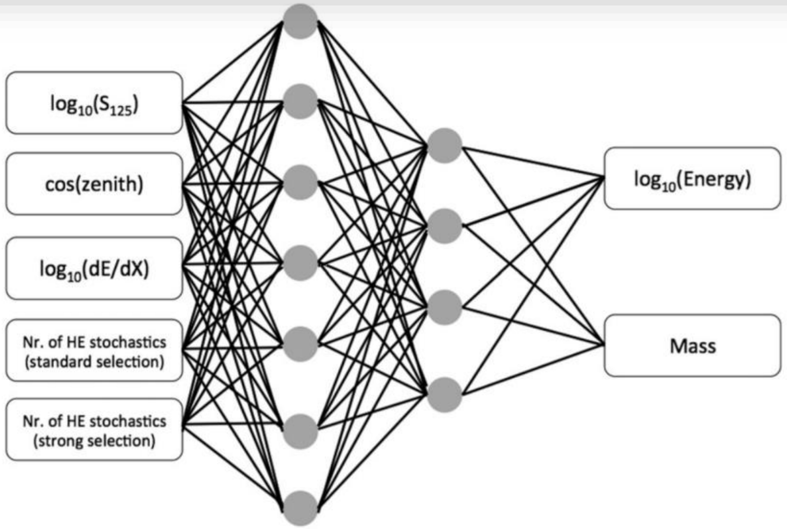


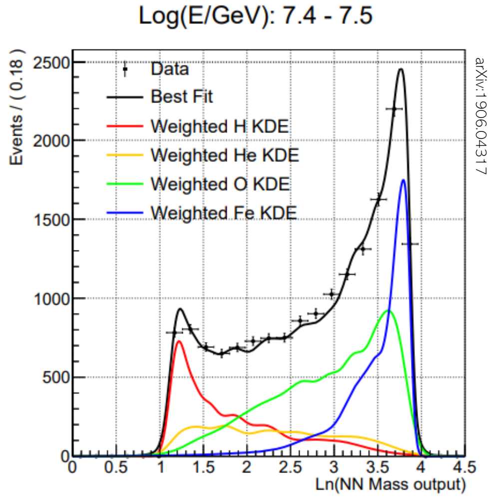


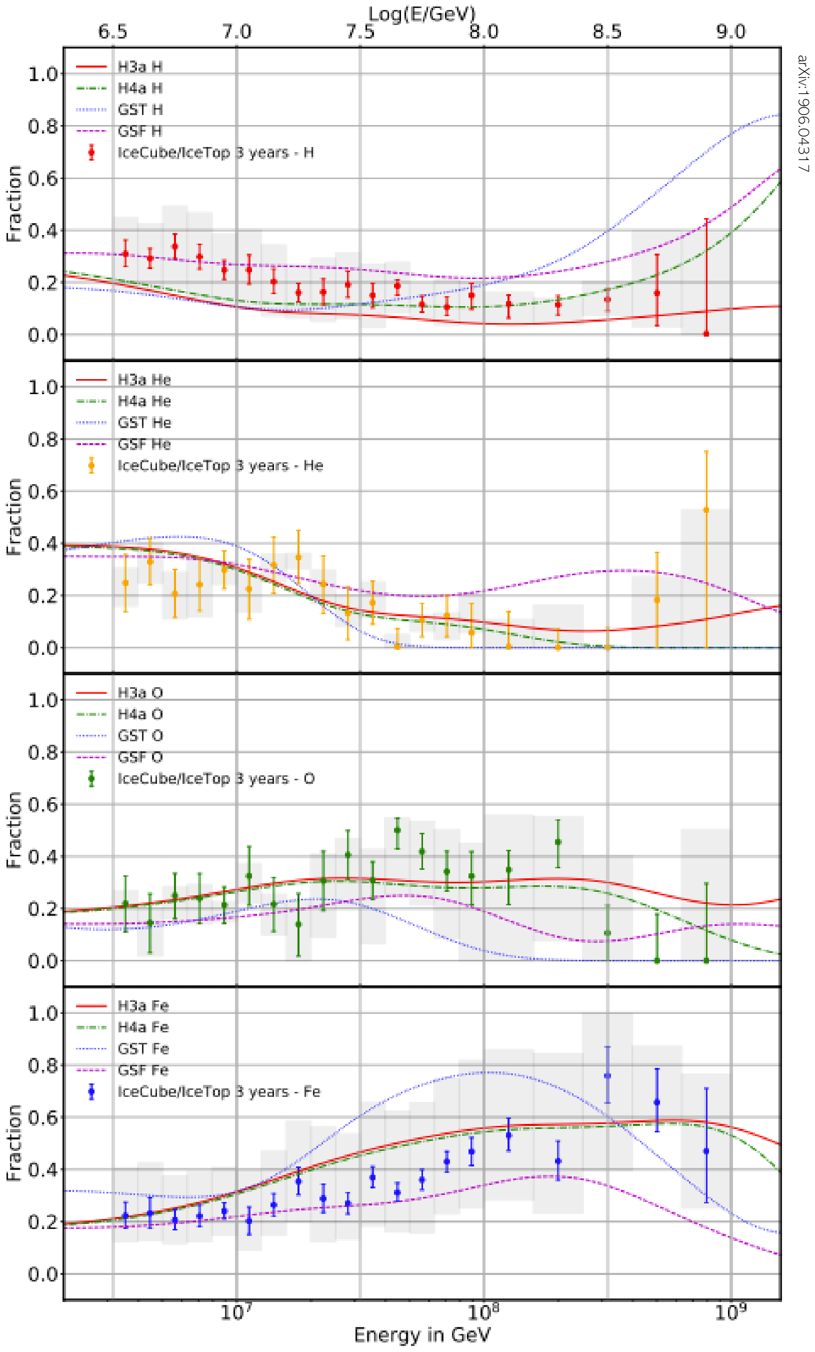
Previous Work

Cosmic ray spectrum and composition from PeV to EeV using 3 years of data from IceTop and IceCube

M. G. Aartsen *et al.* (IceCube Collaboration)
Phys. Rev. D **100**, 082002 – Published 23 October 2019



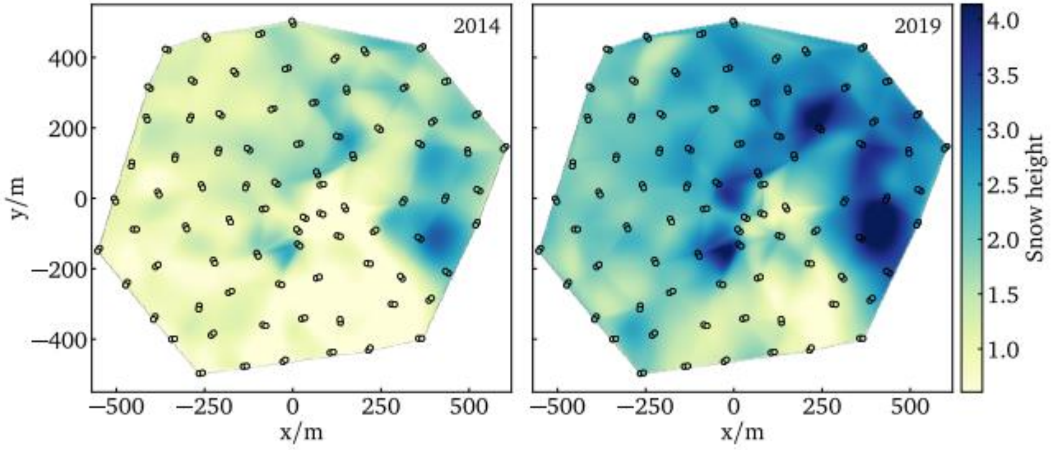




Enhancing CR Analysis: The Science Case

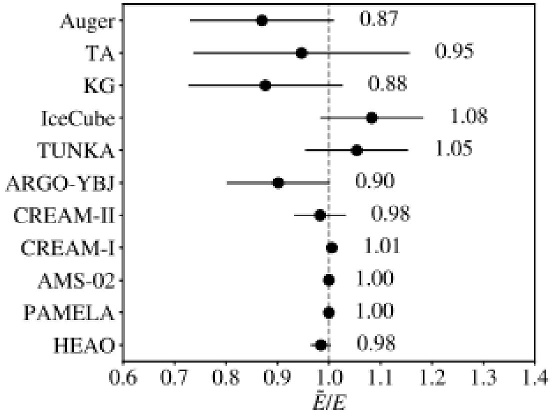
Attenuation by Snow-Coverage

KSETA Plenary Workshop, 2020: A. Leszczyńska



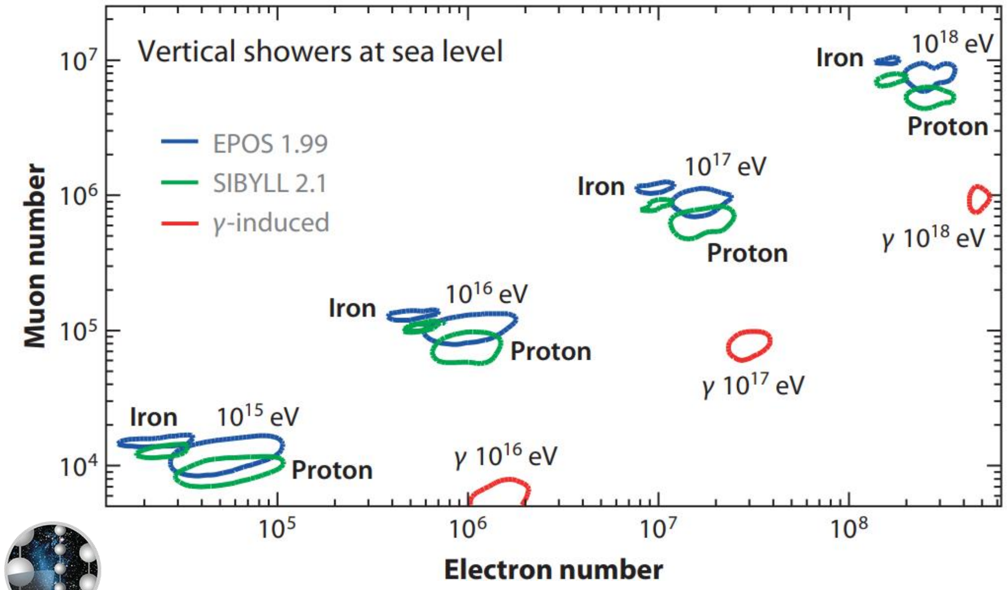
Differences between Experiments

PoS ICRC2017: H. Dembinski et al.

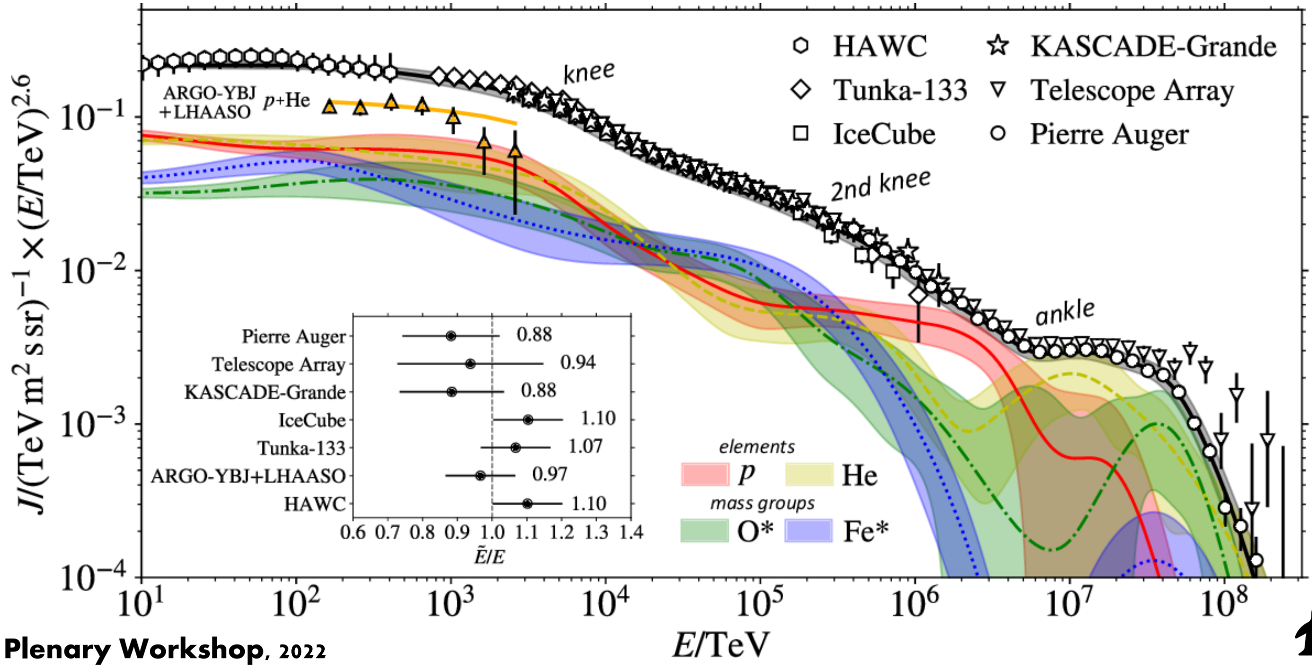


Hadronic Interaction Models

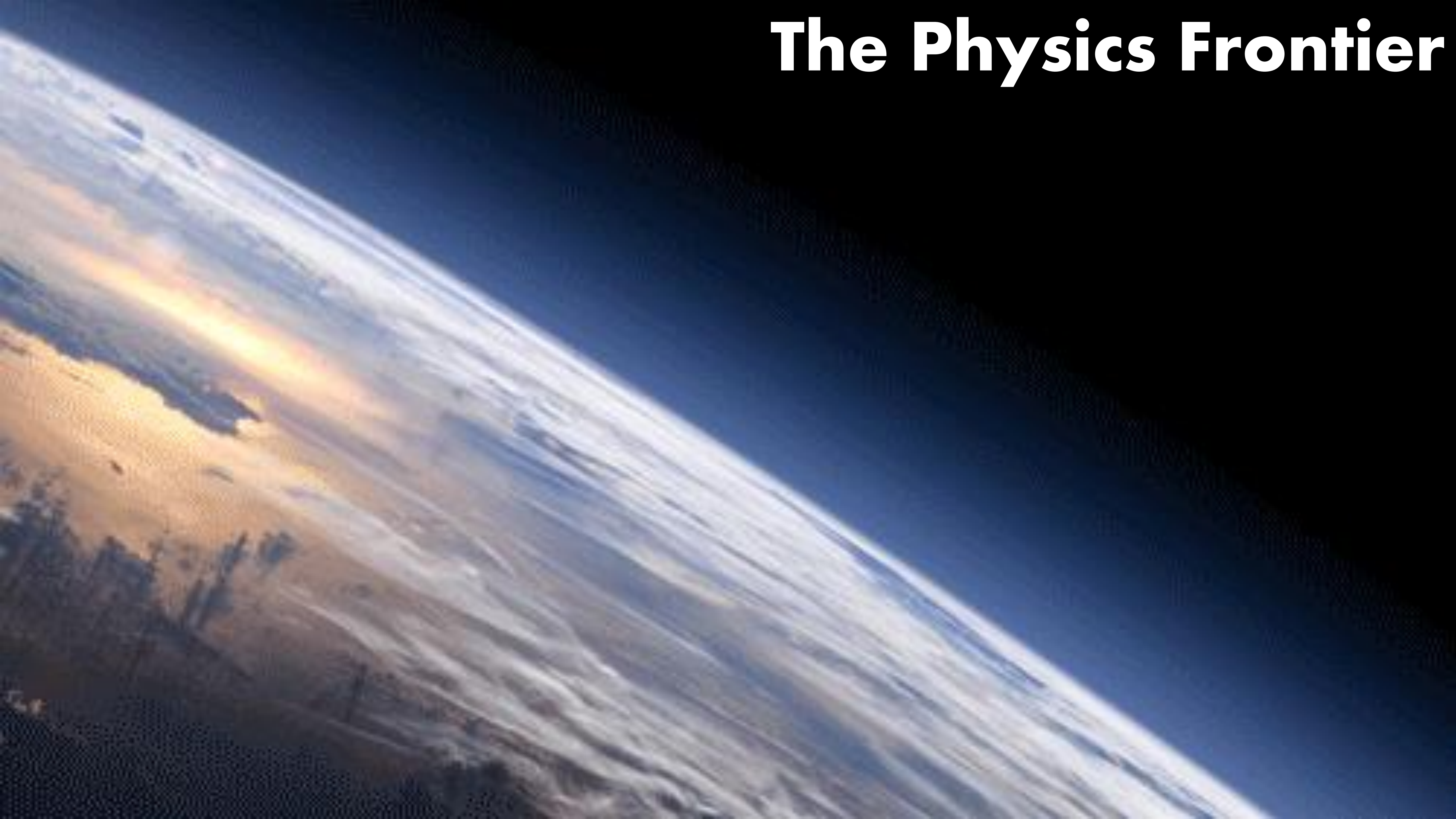
Annu. Rev. Nucl. Part. Sci. 2011: R. Engel et al.



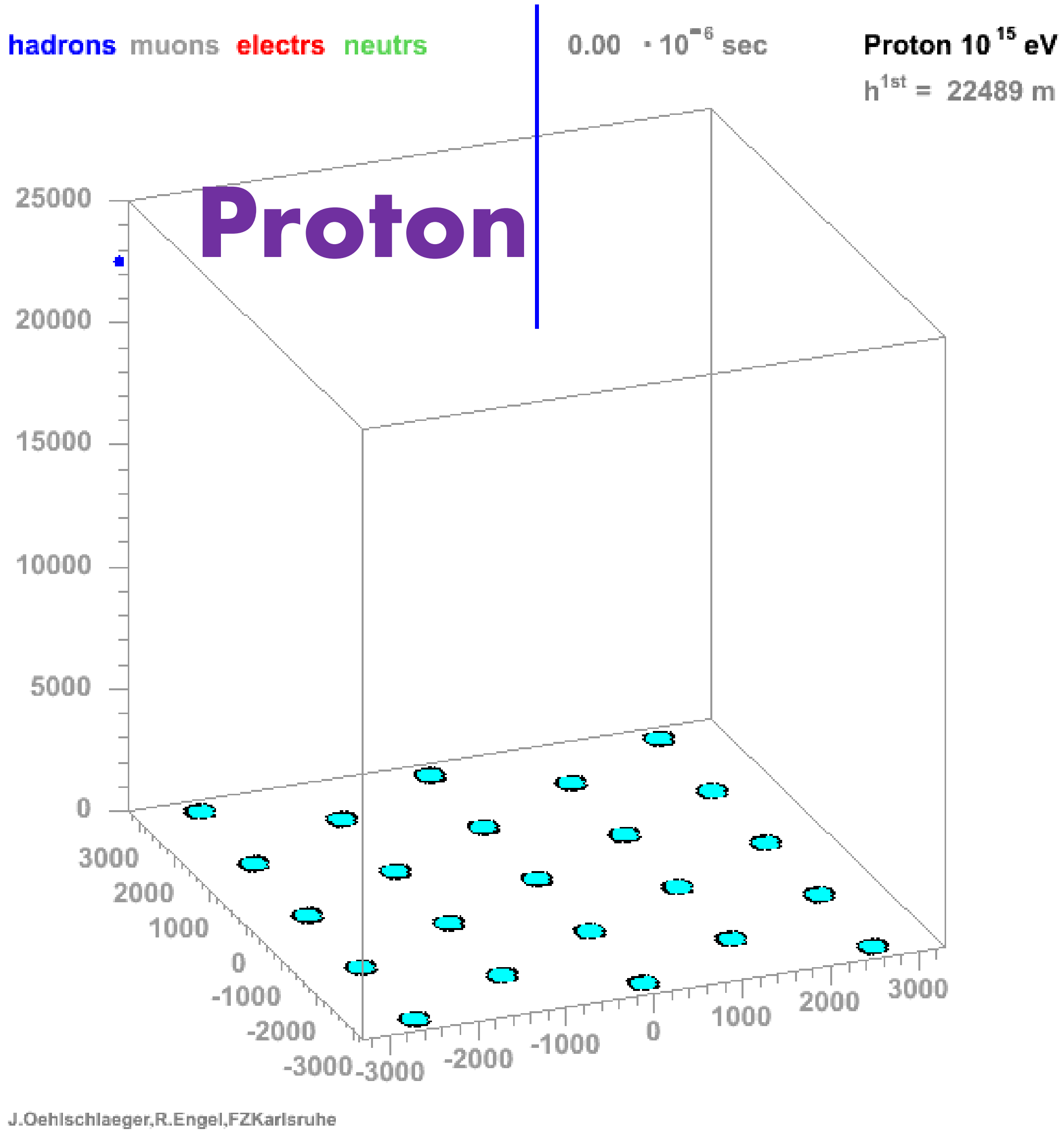
Astro2020 Science White Paper ; Frank G. Schröder et al.



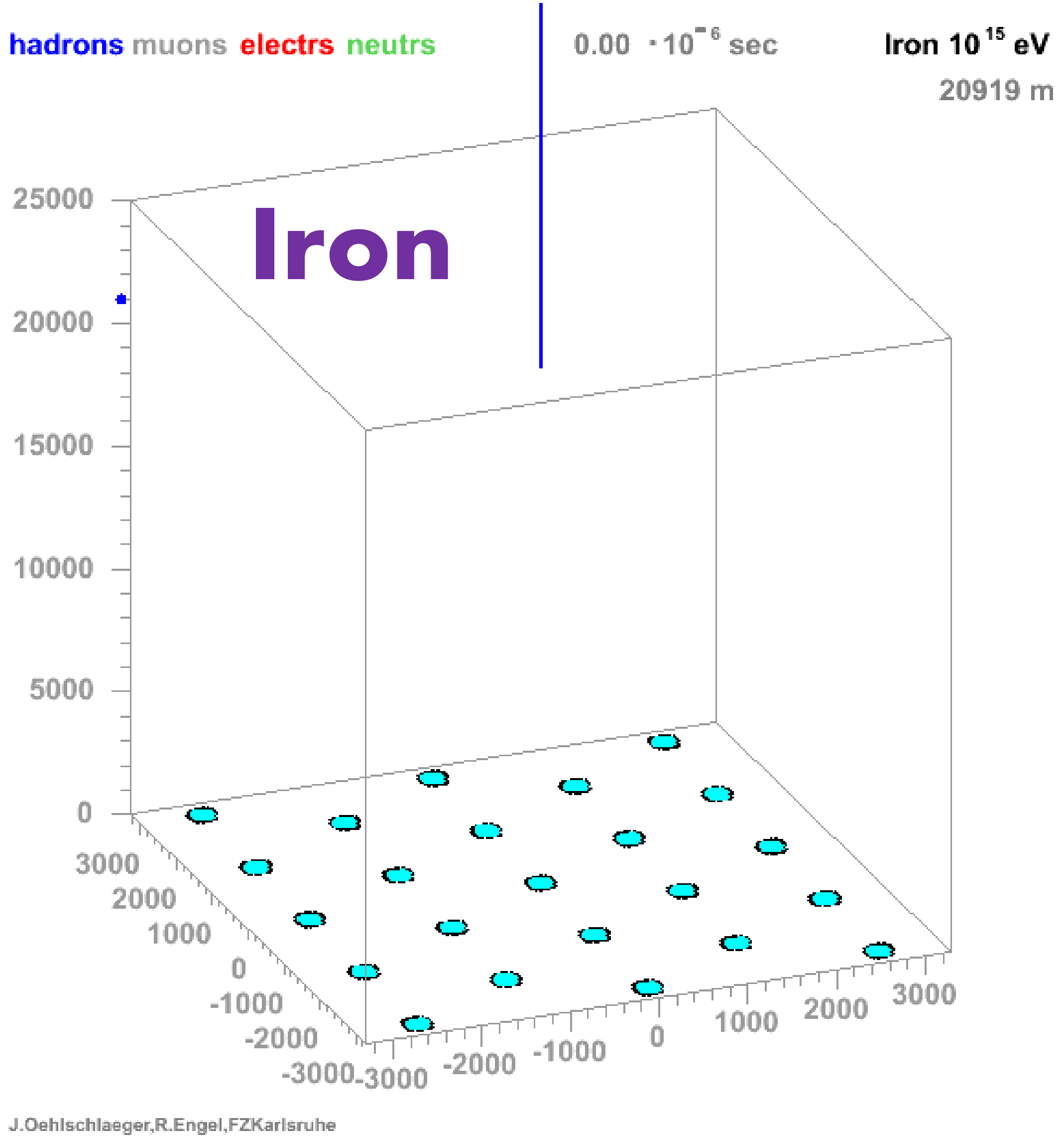
The Physics Frontier



Air-Shower



hadrons muons electrs neutrs
0.00 · 10⁻⁶ sec
Proton 10^15 eV
h^1st = 22489 m
Proton
25000
20000
15000
10000
5000
0
3000
2000
1000
0
-1000
-2000
-3000
-3000
-2000
-1000
0
1000
2000
3000
J.Oehlschlaeger,R.Engel,FZKarlsruhe



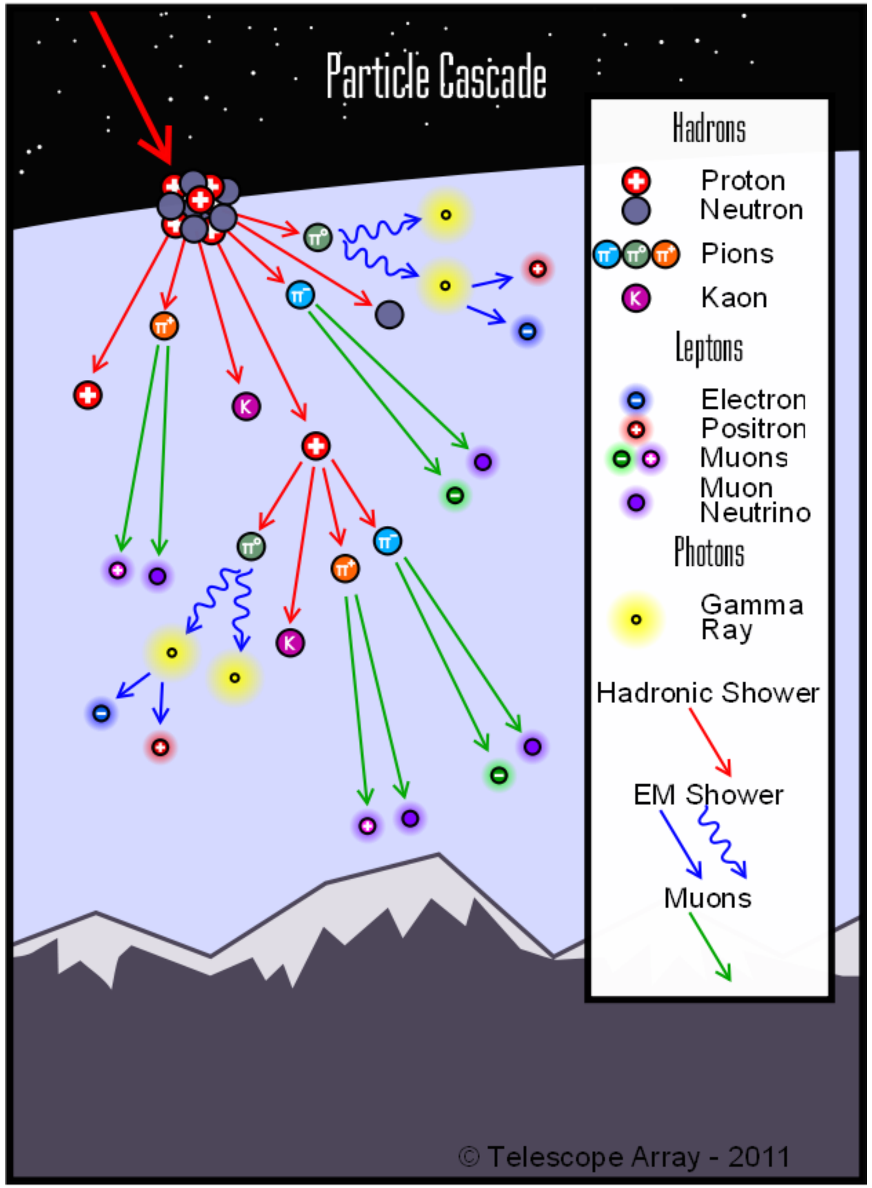
hadrons muons electrs neutrs
0.00 · 10⁻⁶ sec
Iron 10^15 eV
20919 m
Iron
25000
20000
15000
10000
5000
0
3000
2000
1000
0
-1000
-2000
-3000
-3000
-2000
-1000
0
1000
2000
3000
J.Oehlschlaeger,R.Engel,FZKarlsruhe

Muons as Probes

First Interaction

Particle Number

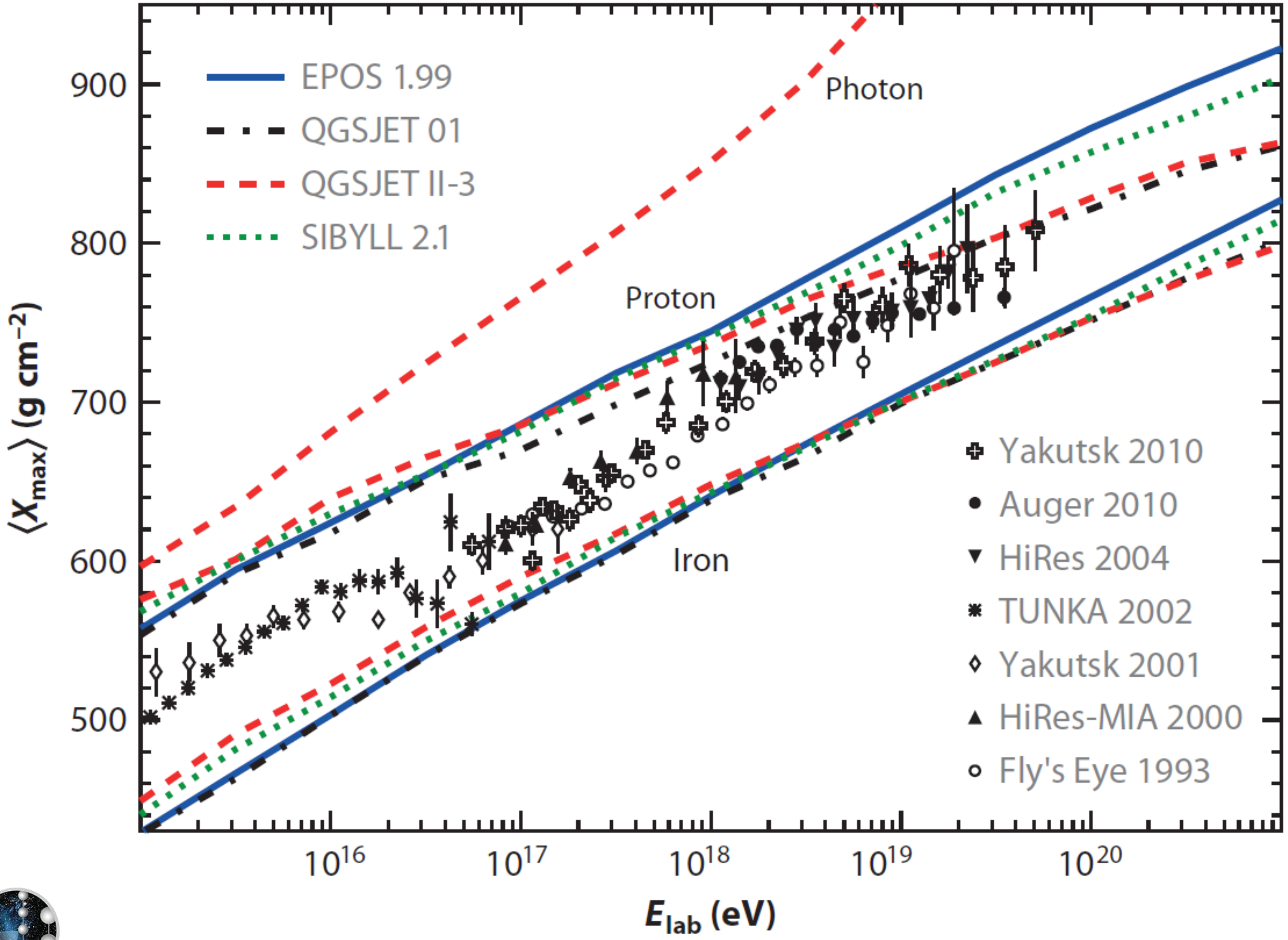
Energy-Deposit Rate



1. Dependence – First Interaction

Fe-shower interact earlier in the atmosphere

Annu. Rev. Nucl. Part. Sci. 2011: R. Engel et al.



Fe Bundles »

Larger muon multiplicity »

Creates wider muon bundles,

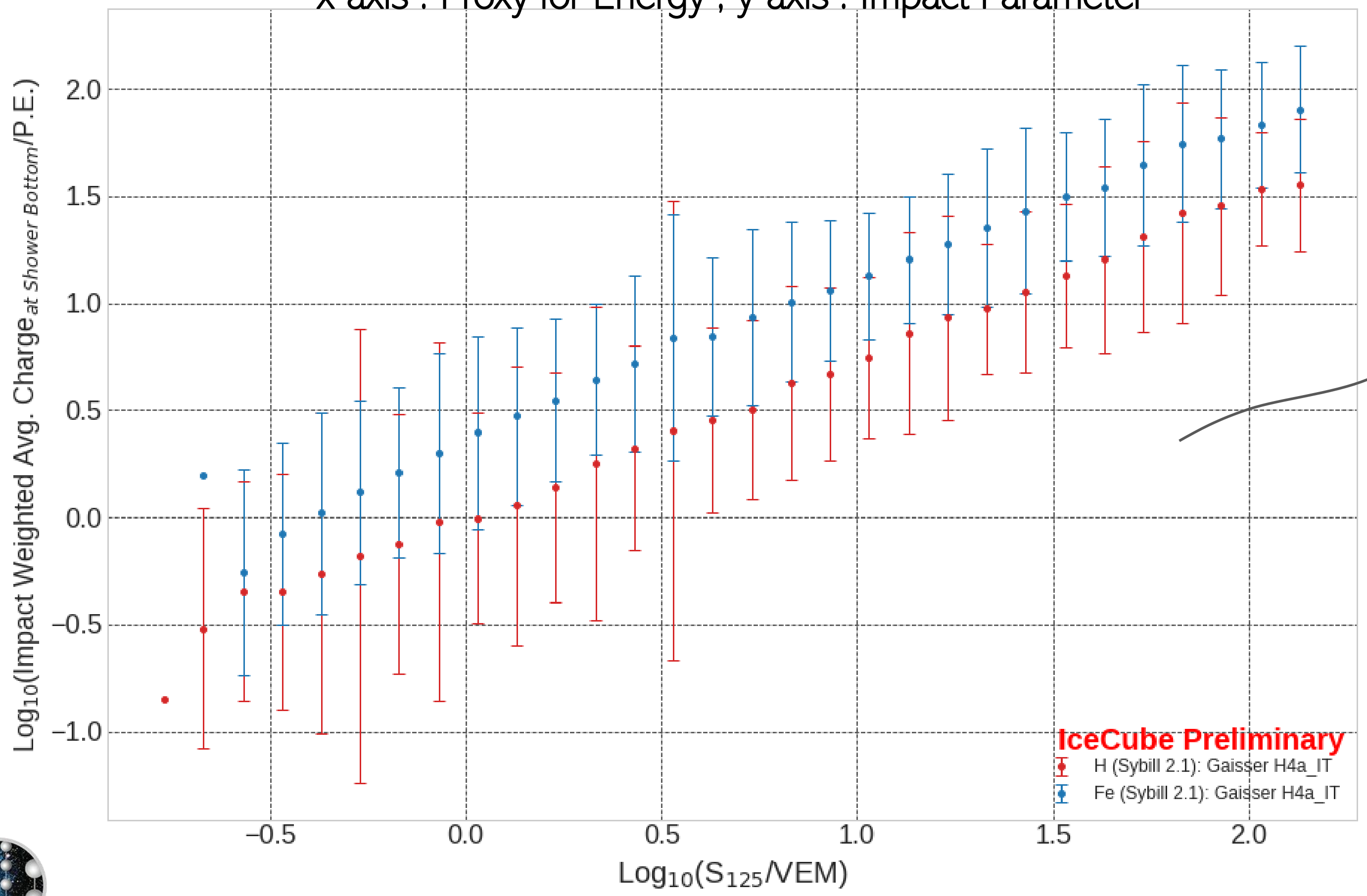
(Because lower energy muons have larger transversal momenta and are situated further from the shower axis).

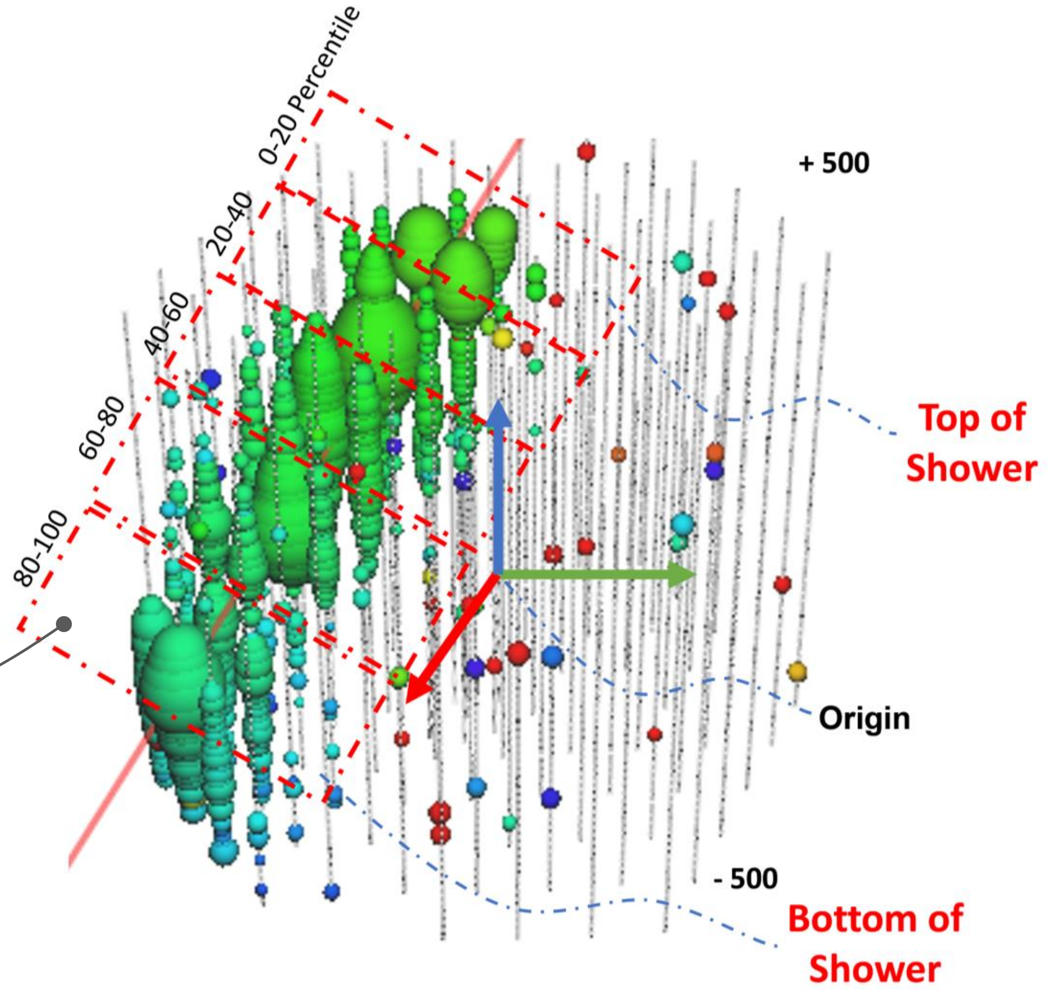
Impact Weighted Charge

$$\frac{\sum Charge_i * ri}{\sum r_i}$$

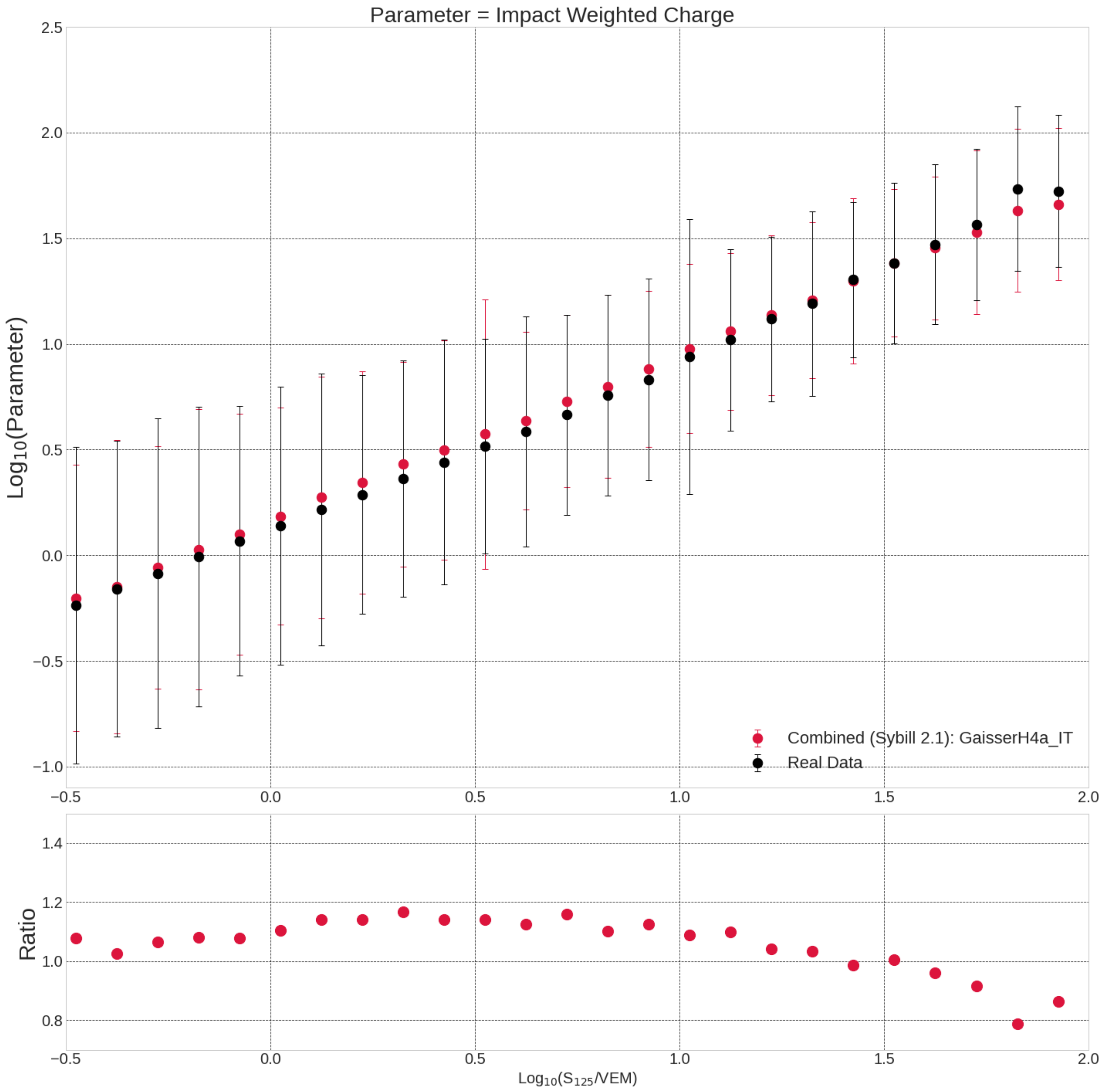
- Help understand shower attenuation in-ice and a dynamic parameter
- Possibly help in understanding photon propagation in-ice
    - Ongoing work
- Good Separation Between Primaries

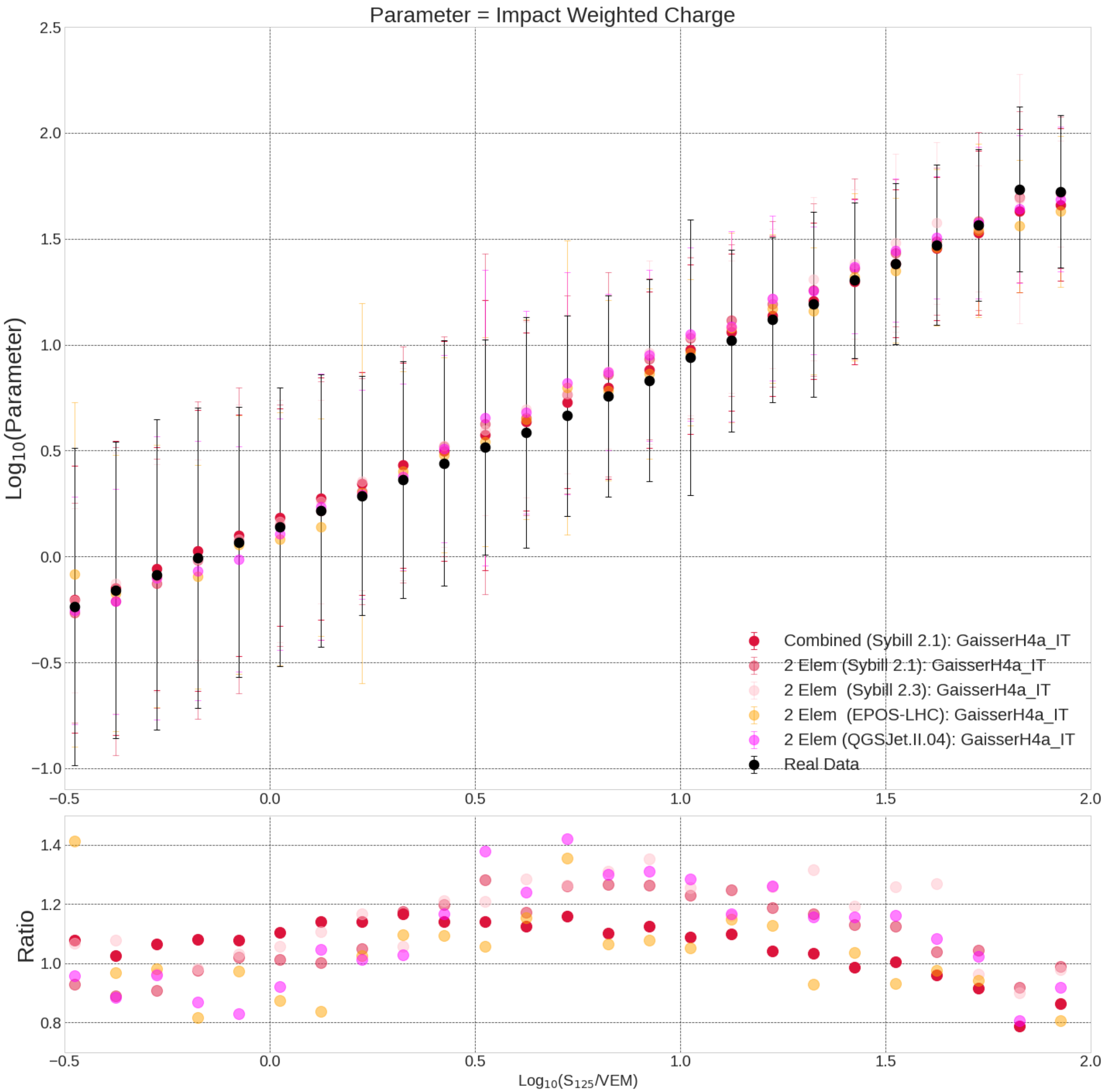
x-axis : Proxy for Energy ; y-axis : Impact Parameter





Data vs Simulations




Parameter = Impact Weighted Charge
Log10(Parameter)
Combined (Sybill 2.1): GaisserH4a_IT
2 Elem (Sybill 2.1): GaisserH4a_IT
2 Elem (Sybill 2.3): GaisserH4a_IT
2 Elem (EPOS-LHC): GaisserH4a_IT
2 Elem (QGSJet.II.04): GaisserH4a_IT
Real Data
Ratio
Log10(S125/VEM)

2. Dependence – Particle Number

Motivation

arXiv:1306.6283

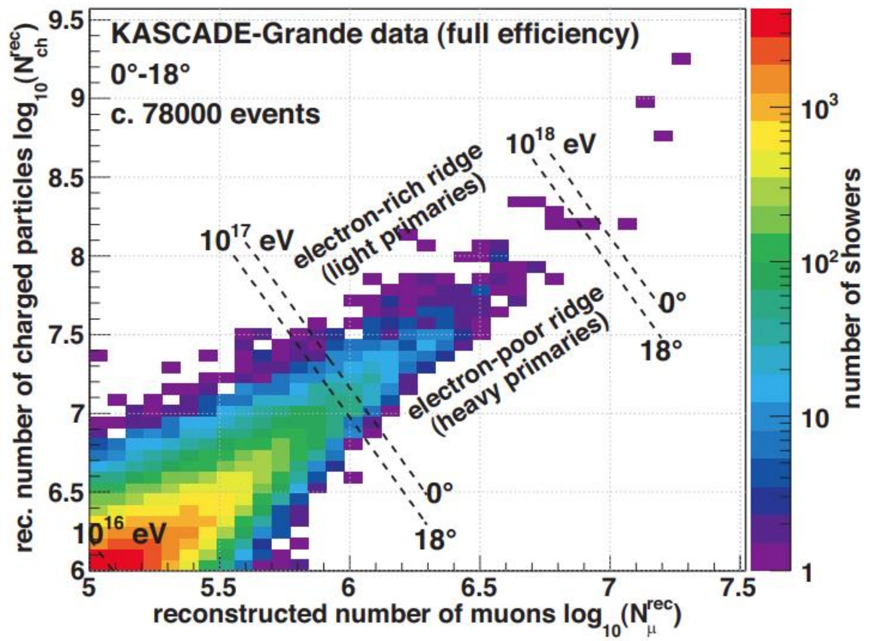


Figure 1: Two-dimensional distribution of the shower sizes (total number of charged particles and of muons) measured with KASCADE-Grande and used for this analysis. Diverse quality cuts are applied. Furthermore, only events with zenith angles $\theta \leq 18°$ and with shower sizes for which the experiment is fully efficient (above $\log_{10}(N_{\rm ch}^{\rm rec}) \approx 6.0$ and $\log_{10}(N_{\mu}^{\rm rec}) \approx 5.0$) are considered. In addition, a roughly estimated energy scale is indicated. Since KASCADE-Grande measures the shower sizes at atmospheric depths beyond the shower maximum, electron-rich showers are initiated preferentially by light primaries, and electron-poor showers by heavy ones, respectively (this is indicated in the figure, too).

$$\textbf{Ratio} = \frac{\log_{10} N\mu}{\log_{10} Ne}$$

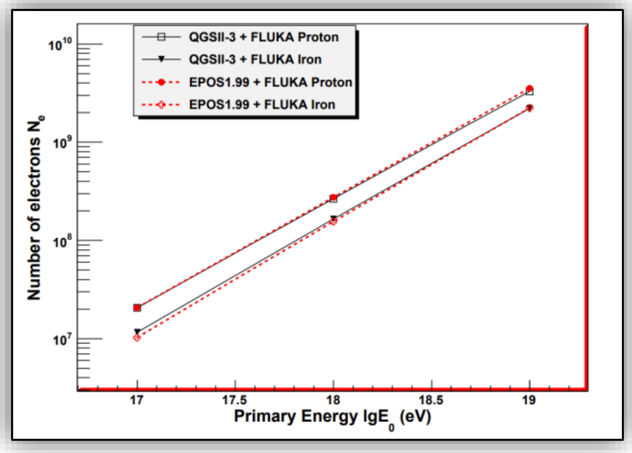
Proxy

$$\frac{\log_{10} N\mu}{\log_{10} Ne}$$

A mass dependent Parameter?

Total Energy?

arXiv:1202.3661



$$\textit{Ratio} \propto \frac{\log_{10}\mathrm{dEdX}_{1500}}{\log_{10}\mathrm{E}(\mathrm{S}_{125})}$$

$$\log_{10}(E/\mathrm{GeV}) = p_0 + p_1 \log_{10}(S_{125}/\mathrm{VEM}). \quad (2)$$

Using updated simulation and reconstruction algorithms, the mapping of $\log_{10}(S_{125})$ to $\log_{10}(E/\mathrm{GeV})$ from [5] has been re-optimized and applied to the 3-year data set. The updated fit parameters for Eqn. 2 are in the following table:

Zenith range	p_0	p_1
$0.95 < \cos(\theta) \leq 1.0$	6.011	0.933
$0.90 < \cos(\theta) \leq 0.95$	6.055	0.924
$0.85 < \cos(\theta) \leq 0.90$	6.110	0.915
$0.80 < \cos(\theta) \leq 0.85$	6.177	0.907

arXiv:1906.04317

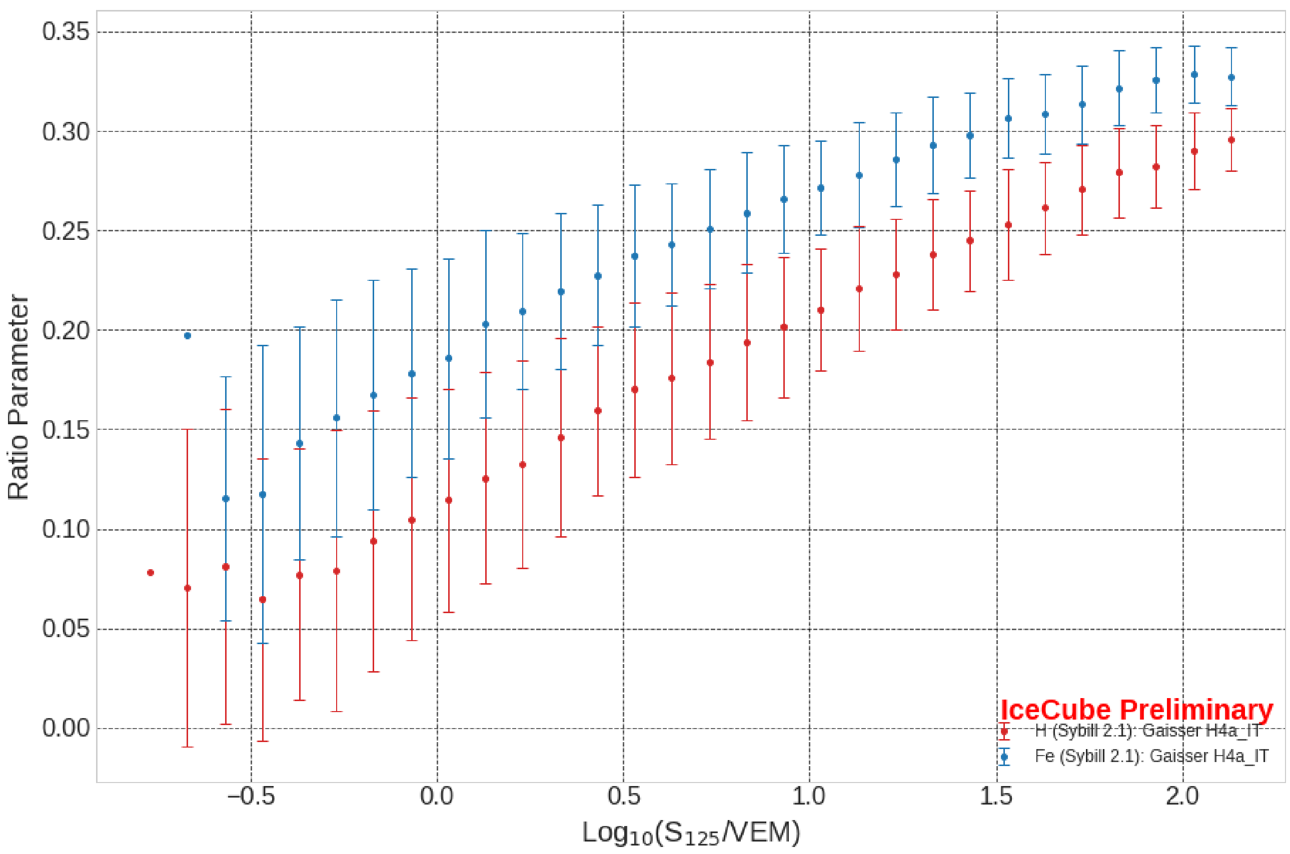
Expectations

For same Energy, ratio should be greater for Fe

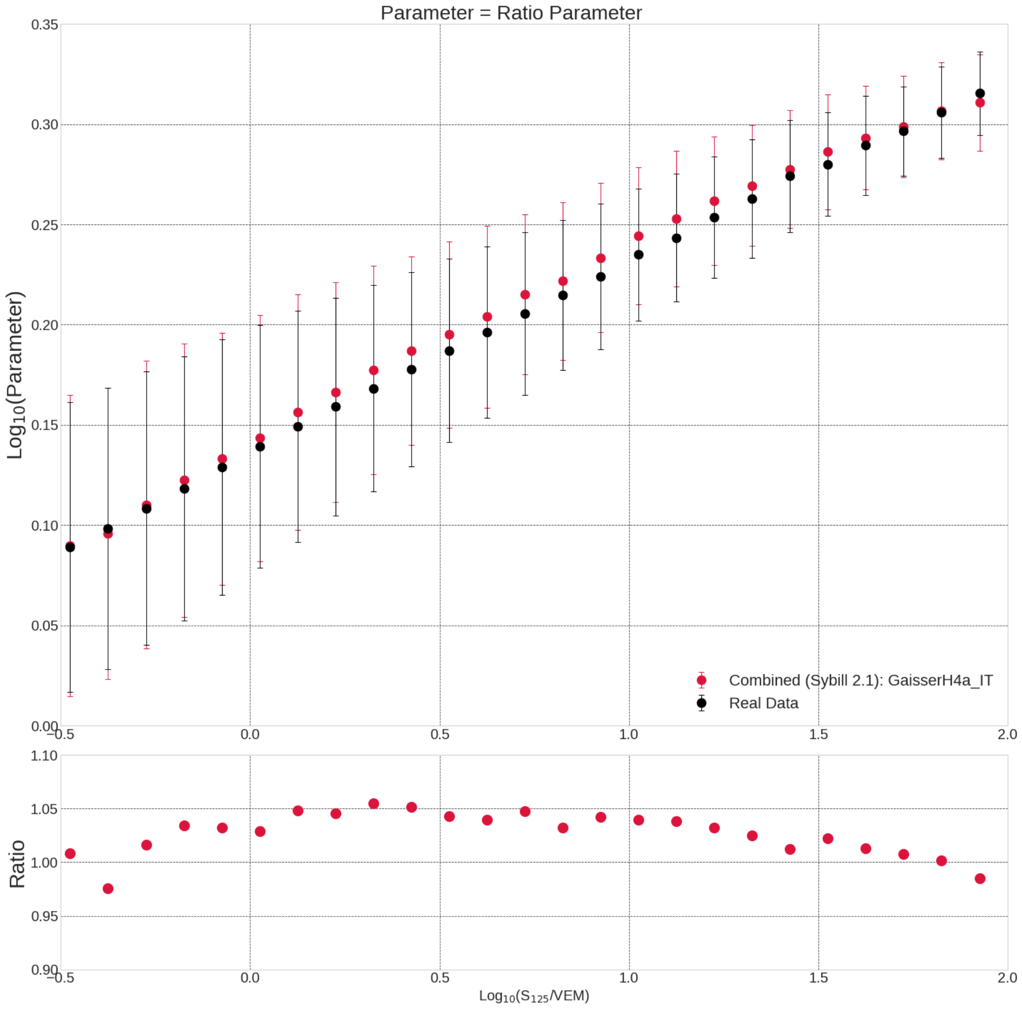
Ratio-Parameter

- Uses IceTop and In-Ice info.
- Proxy for muon-to-electron number ratio
  - Motivated by work of KASCADE-Grande (arXiv:1306.6283)

Composition-Sensitivity

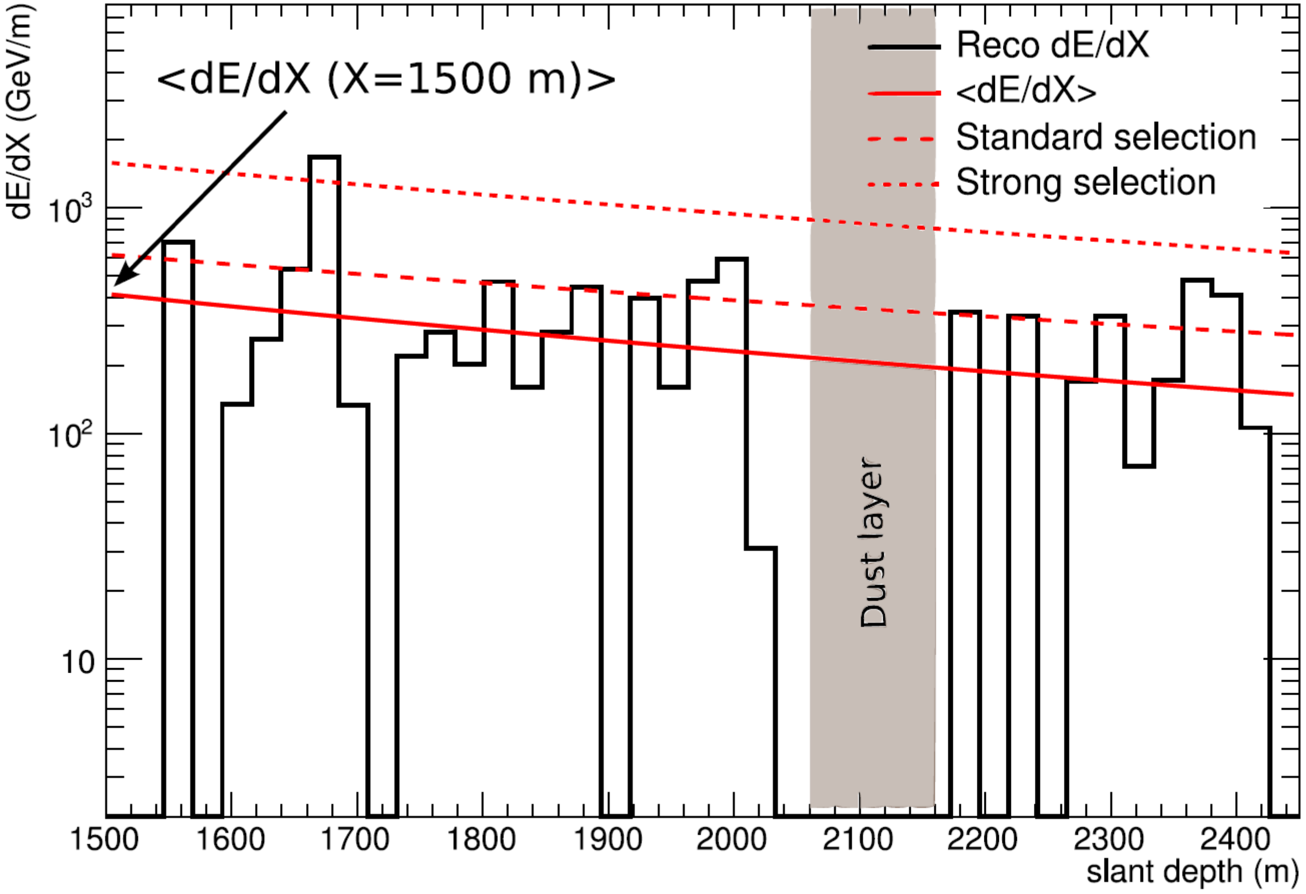


Data vs Simulations



3. Dependence – Energy-Deposit Rate

Motivation



X. Bai , ICRC 2009

$$\frac{dE_{\mu,B}(X)}{dX} = \omega \cdot b \cdot (p_1+1)\frac{1}{V+1}\left[\frac{1}{p_1+1}\left(\frac{a}{b}\right)^{-p_1} \cdot V^{-p_1-1} + \frac{1}{p_1}\left(\frac{a}{b}\right)^{-p_1} V^{-p_1} - \frac{1}{p_1+1}\left(\frac{a}{b}\right) \cdot \left(\frac{E_0}{A}\right)^{-p_1-1} - \frac{1}{p_1}\left(\frac{E_0}{A}\right)^{-p_1}\right]$$

For Same Energy: Numerator ∝ A

$$\text{Slope} \propto \frac{\text{dEdX}_{1500}}{\text{Approx. Slant Length}}$$

Expectations

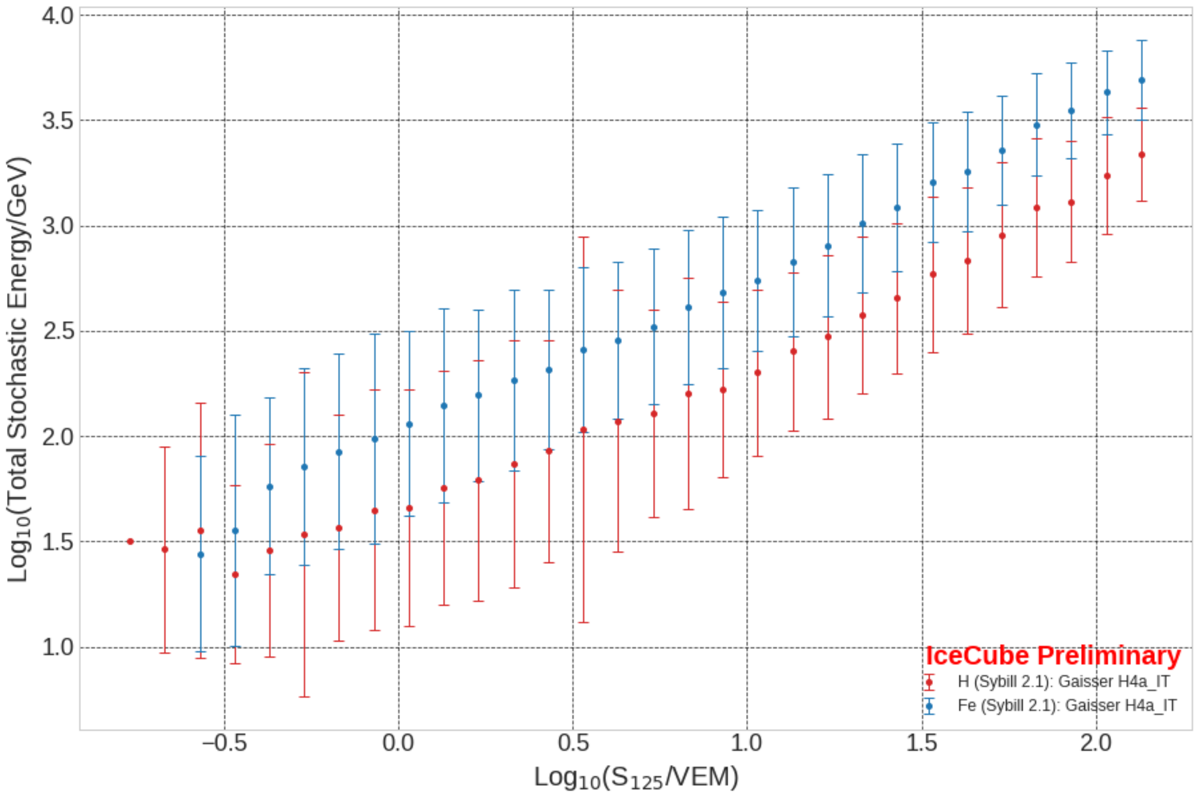
1. Fe(dEdX1500) > H(dEdX1500)
2. **At Same Energy:** Fe (Slant Length) < H (Slant Length)

For Fe

$$: \frac{\text{dEdX}_{1500}}{\text{Approx. Slant Length}}$$

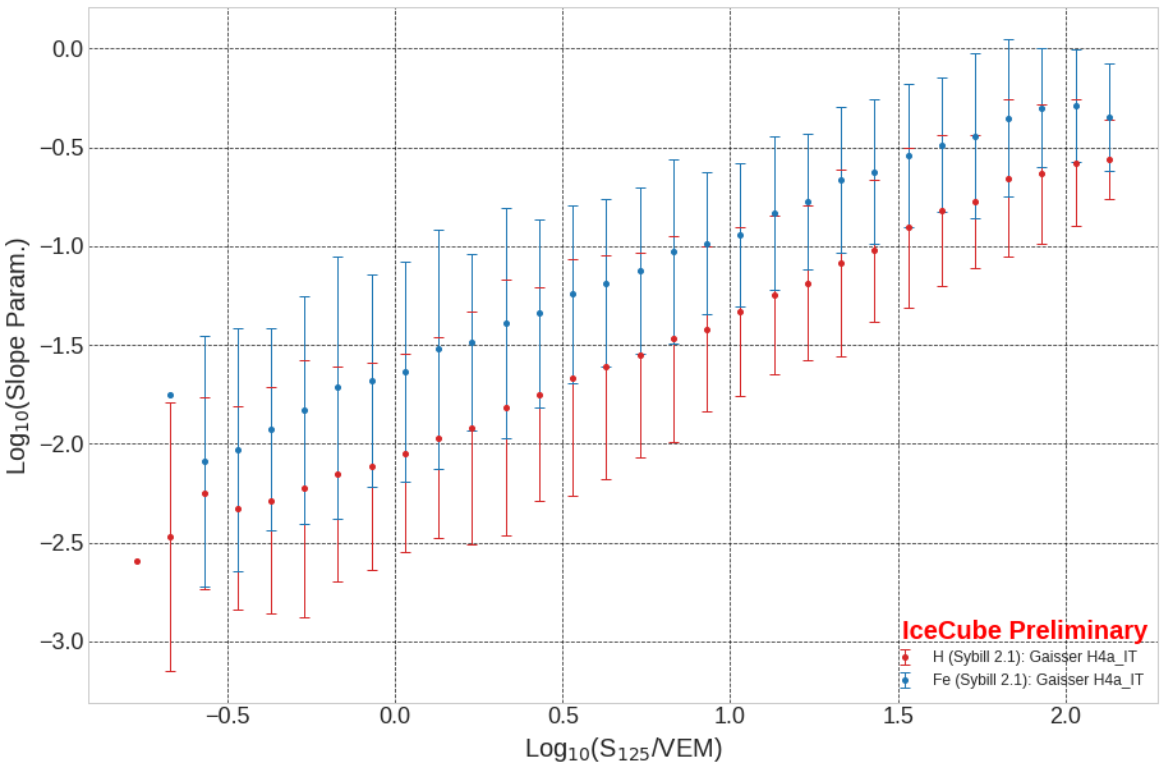
Total Stochastic Energy

- Uses In-Ice info.
- Captures local stochastic-deposits in-ice, primarily by muons.
- Good separation between primaries

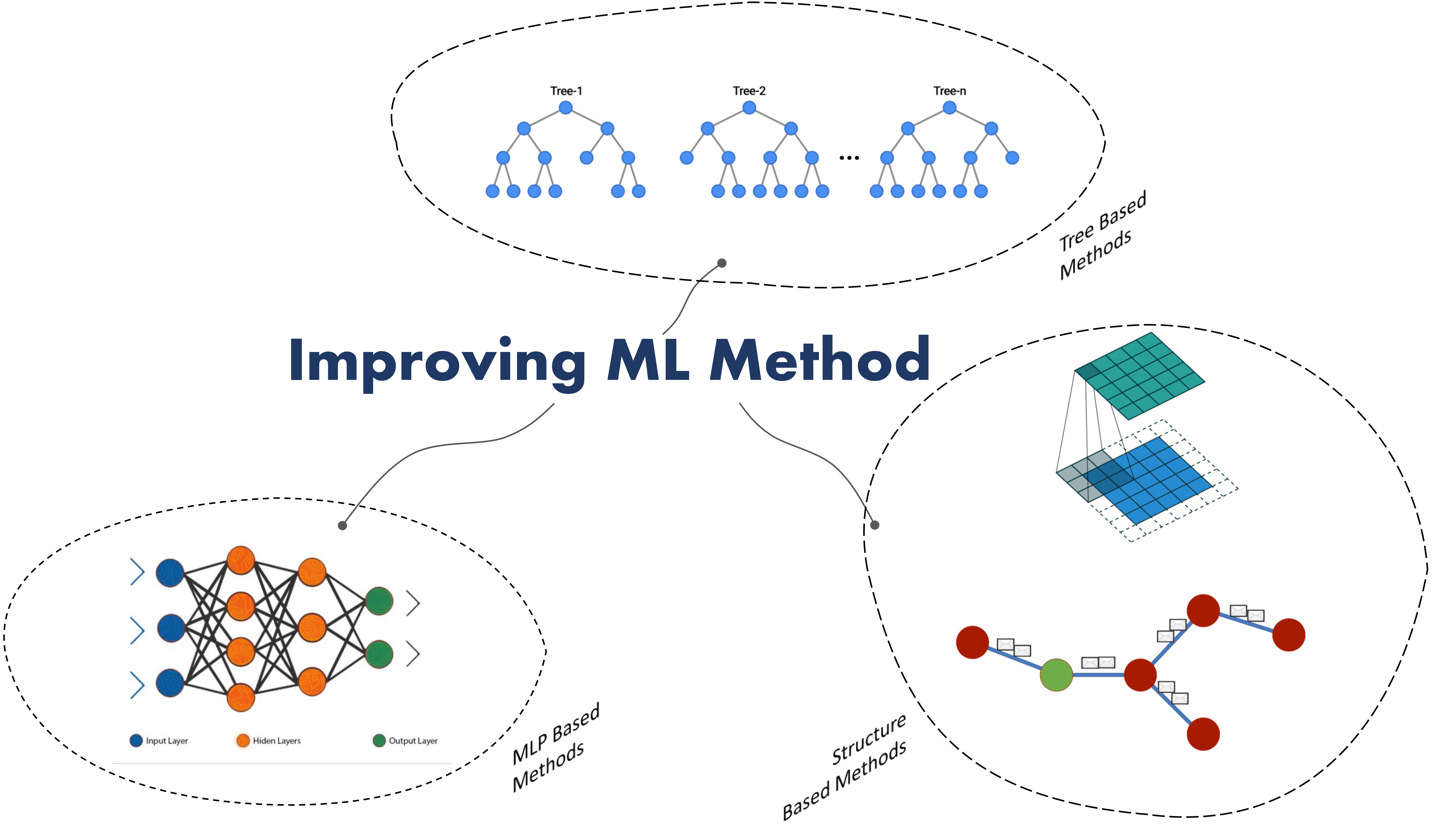


Slope-Parameter

- Uses In-Ice info.
- Captures the rate of in-ice charge deposit.
- Good separation between primaries



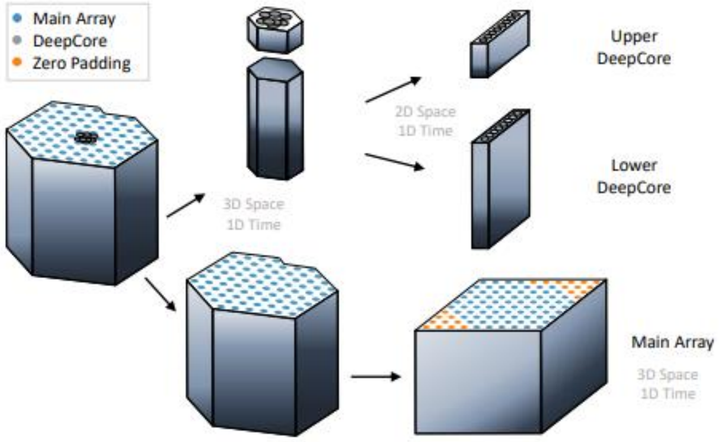
The ML Frontier

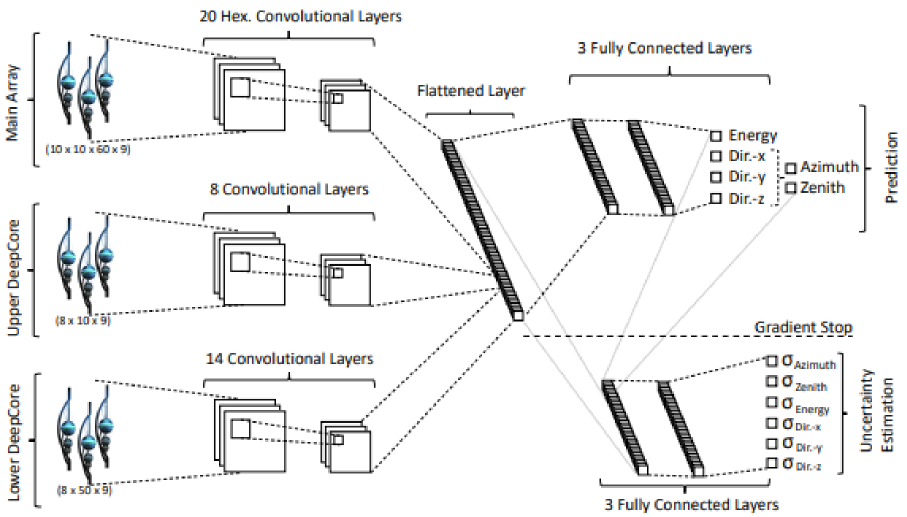

Improving ML Method
Tree-1
Tree-2
Tree-n
Tree Based Methods
MLP Based Methods
Input Layer
Hiden Layers
Output Layer
Structure Based Methods

Using In-Ice Signal Footprint

Current Implementation at IceCube

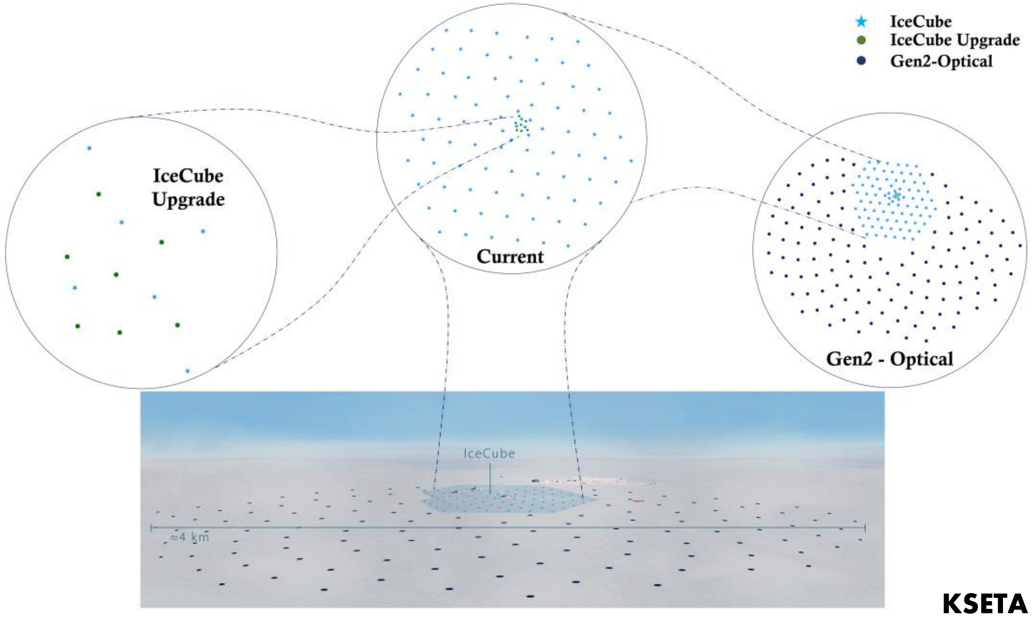
Credits: M. Huennefeld (arXiv:2101.11589)

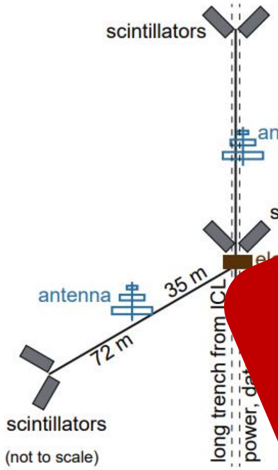




Moving Away from CNNs

Credits: IceCube-Gen2 (arXiv:2008.04323)





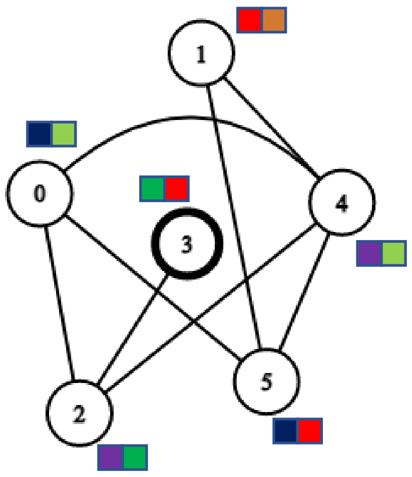
KIT – Major Player

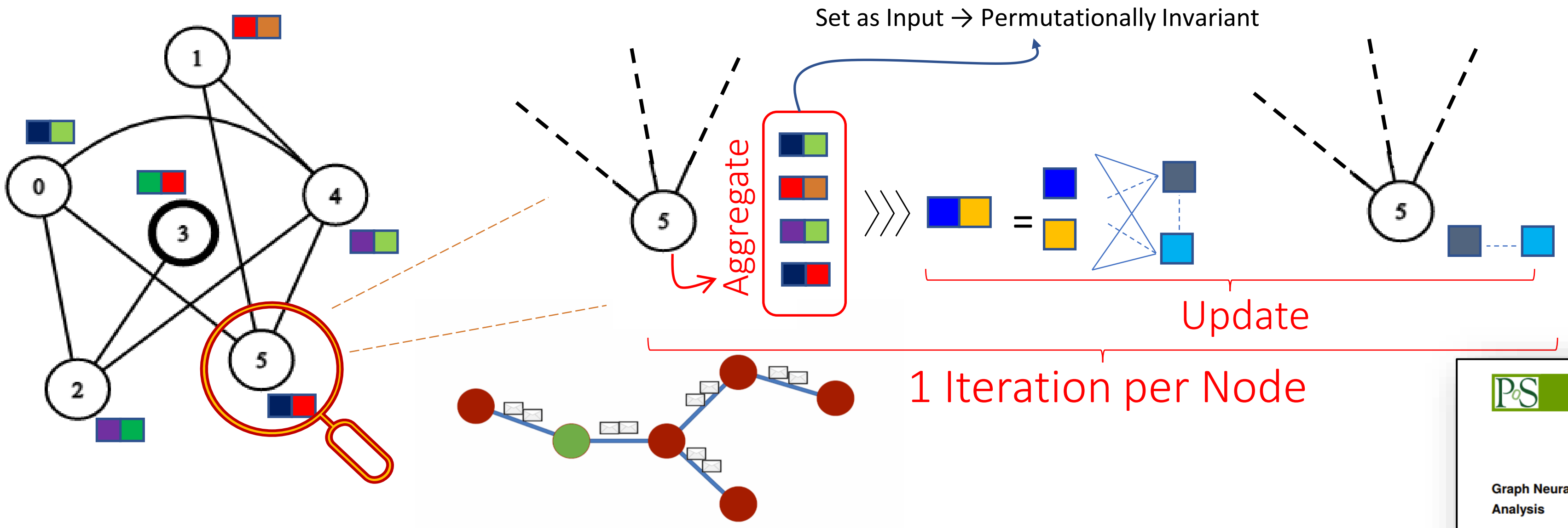
Learning on Graphs

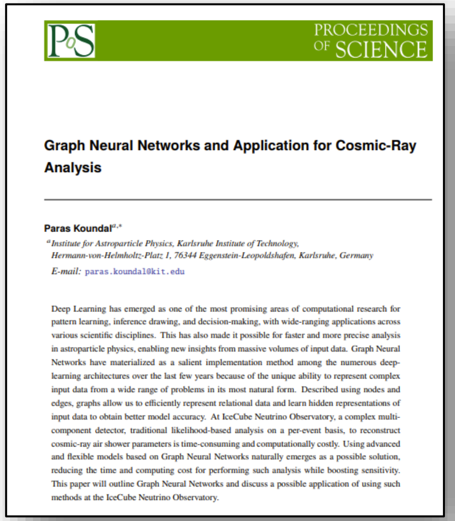
Defined by set of nodes (V) and set of edges (E) between the nodes

- Neighborhood and Connectivity & permutational invariance of Node Labelling

Undirected : Facebook Friends … ; Directed : Citation Graph … ; Bidirectional : Twitter Follows






PROCEEDINGS OF SCIENCE

Graph Neural Networks and Application for Cosmic-Ray Analysis

Paras Koundal[a,*]

[a] *Institute for Astroparticle Physics, Karlsruhe Institute of Technology, Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Karlsruhe, Germany*

E-mail: paras.koundal@kit.edu

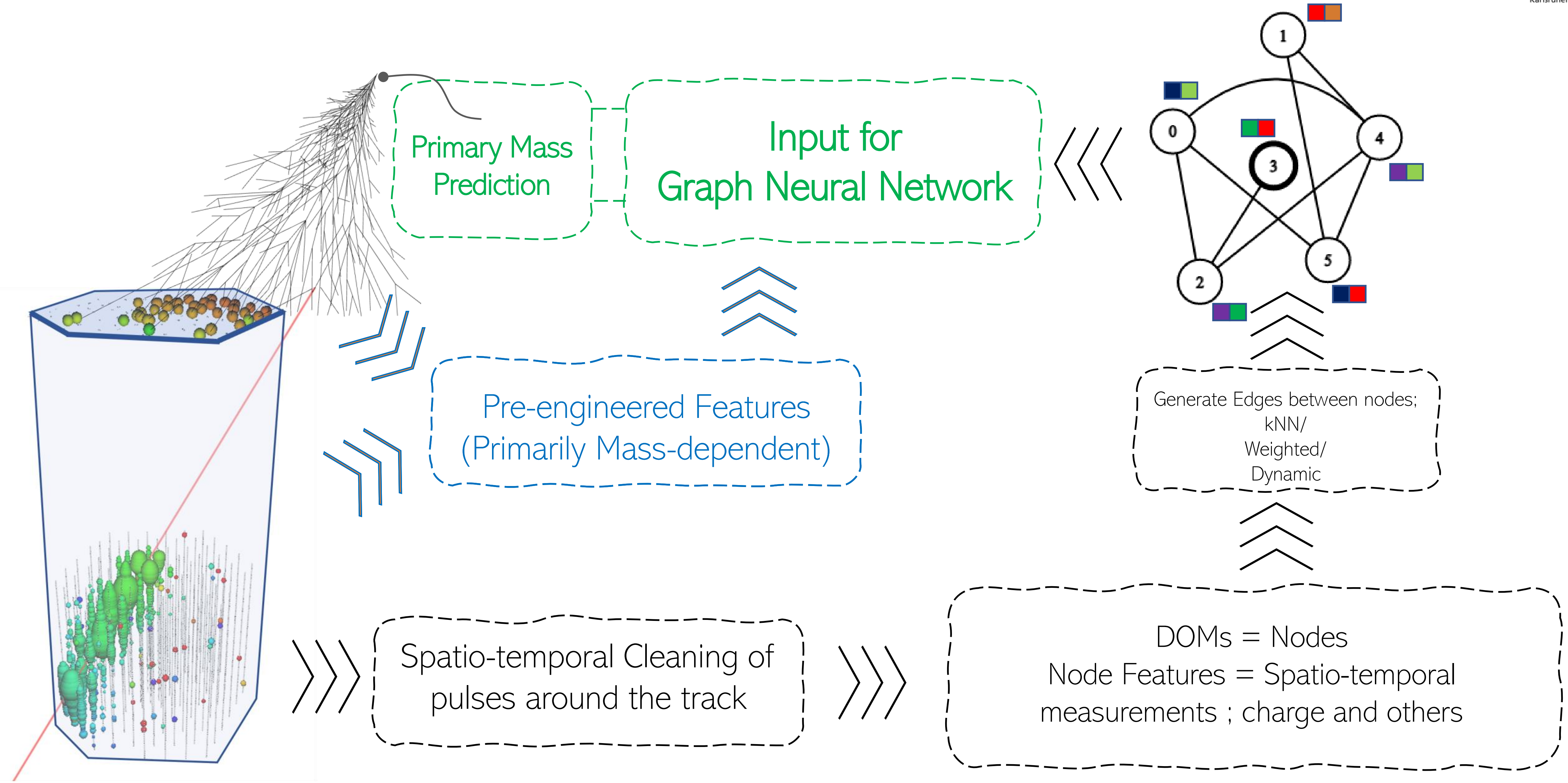
Deep Learning has emerged as one of the most promising areas of computational research for pattern learning, inference drawing, and decision-making, with wide-ranging applications across various scientific disciplines. This has also made it possible for faster and more precise analysis in astroparticle physics, enabling new insights from massive volumes of input data. Graph Neural Networks have materialized as a salient implementation method among the numerous deep-learning architectures over the last few years because of the unique ability to represent complex input data from a wide range of problems in its most natural form. Described using nodes and edges, graphs allow us to efficiently represent relational data and learn hidden representations of input data to obtain better model accuracy. At IceCube Neutrino Observatory, a complex multi-component detector, traditional likelihood-based analysis on a per-event basis, to reconstruct cosmic-ray air shower parameters is time-consuming and computationally costly. Using advanced and flexible models based on Graph Neural Networks naturally emerges as a possible solution, reducing the time and computing cost for performing such analysis while boosting sensitivity. This paper will outline Graph Neural Networks and discuss a possible application of using such methods at the IceCube Neutrino Observatory.

5th International Workshop on Deep Learning in Computational Physics, 2021

Other Material

- paraskoundal.com/dlcp21
- My ML & AI Paper List
- My Twitter (Paras Koundal)

GNNs at IceCube (for CR Analysis)



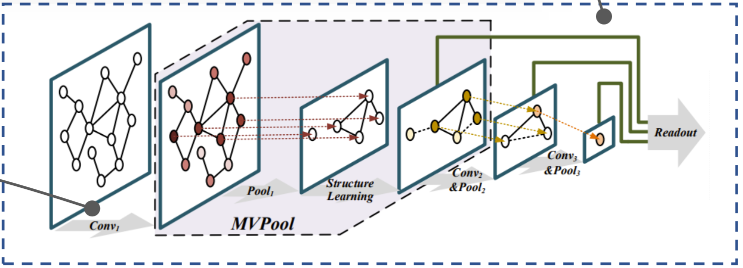
Architecture Details

- New Graph Message-Passing framework
  - Adapted from the work "**Hierarchical Multi-View Graph Pooling with Structure Learning**"
    - Improvement in Graph pooling (or downsampling) to learn hierarchical representations
    - Attention mechanism utilized to generate robust node ranks.
    - Preserve the underlying graph topological information, using a structure learning mechanism.

Date of Publication: 18 June 2021

Hierarchical Multi-View Graph Pooling with Structure Learning

Zhen Zhang, Jiajun Bu*, Member, IEEE, Martin Ester, Senior Member, IEEE, Jianfeng Zhang, Zhao Li*, Chengwei Yao, Huifen Dai, Zhi Yu, Can Wang, Member, IEEE



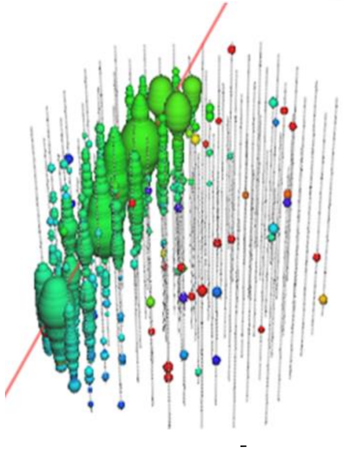
*3

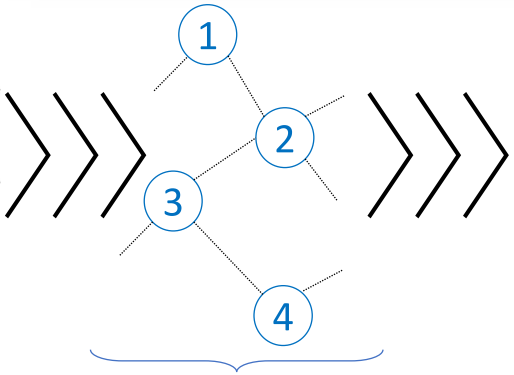
SEMI-SUPERVISED CLASSIFICATION WITH GRAPH CONVOLUTIONAL NETWORKS

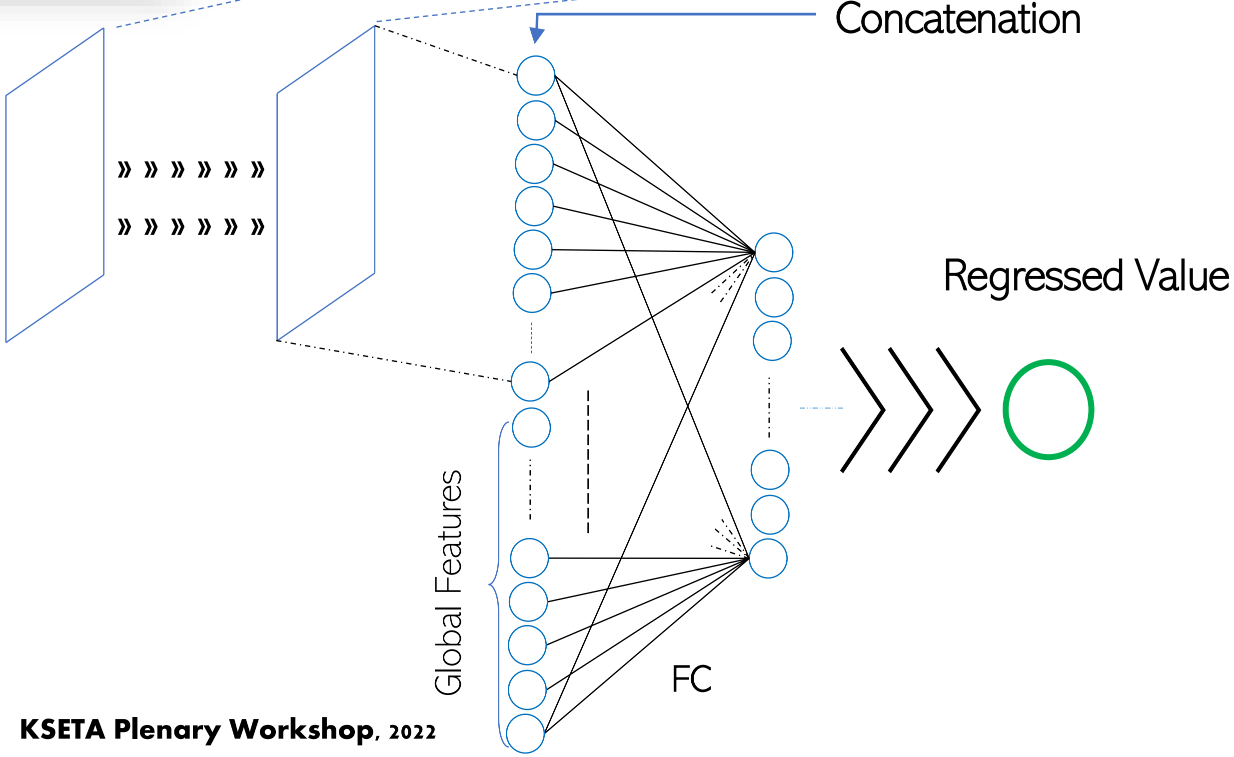
Thomas N. Kipf
University of Amsterdam
T.N.Kipf@uva.nl

Max Welling
University of Amsterdam
Canadian Institute for Advanced Research (CIFAR)
M.Welling@uva.nl

Date of Publication: 22 Feb 2017



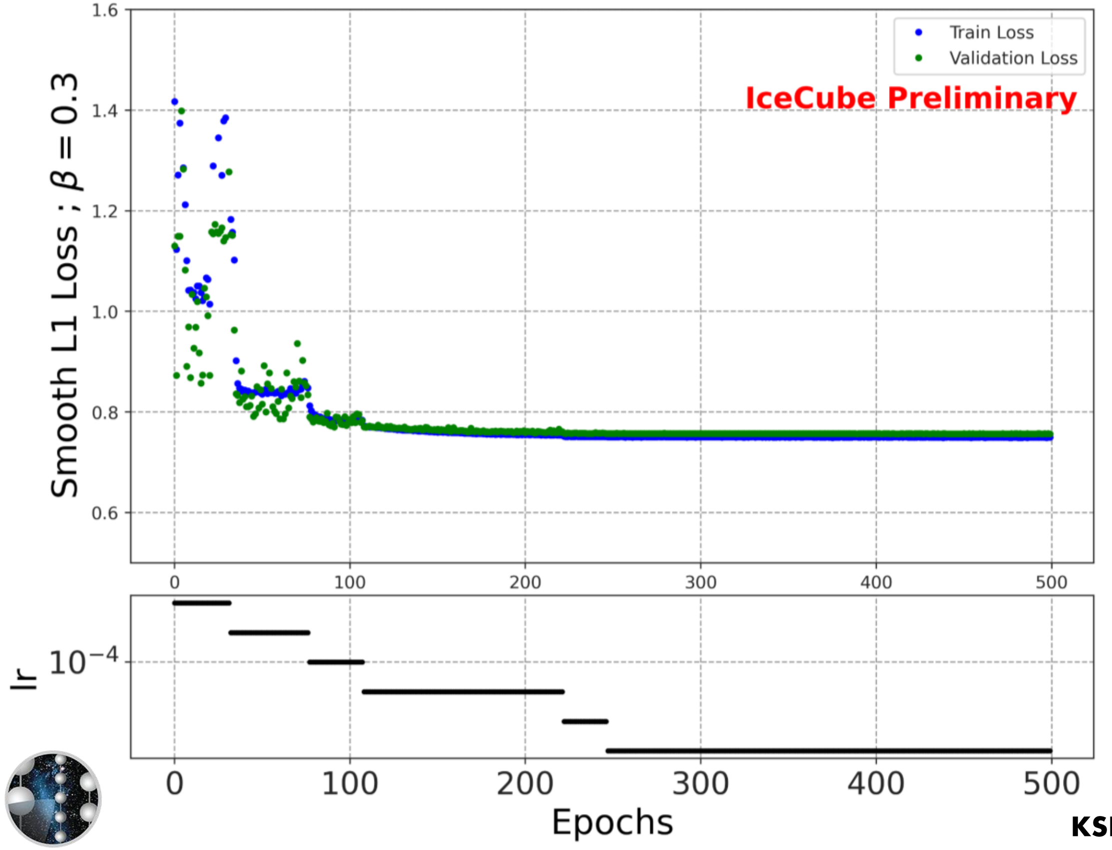


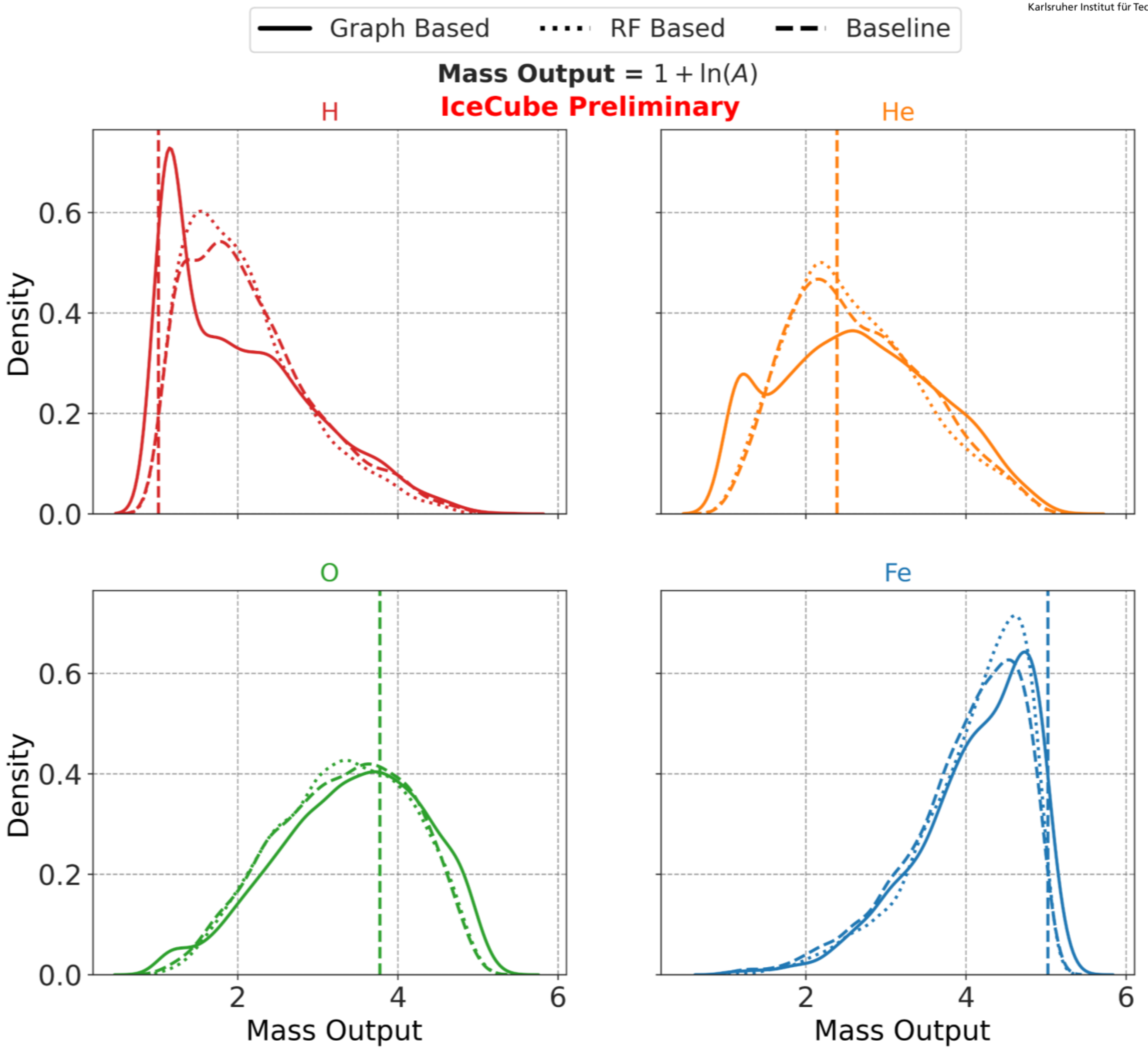


Many different architectures were tested

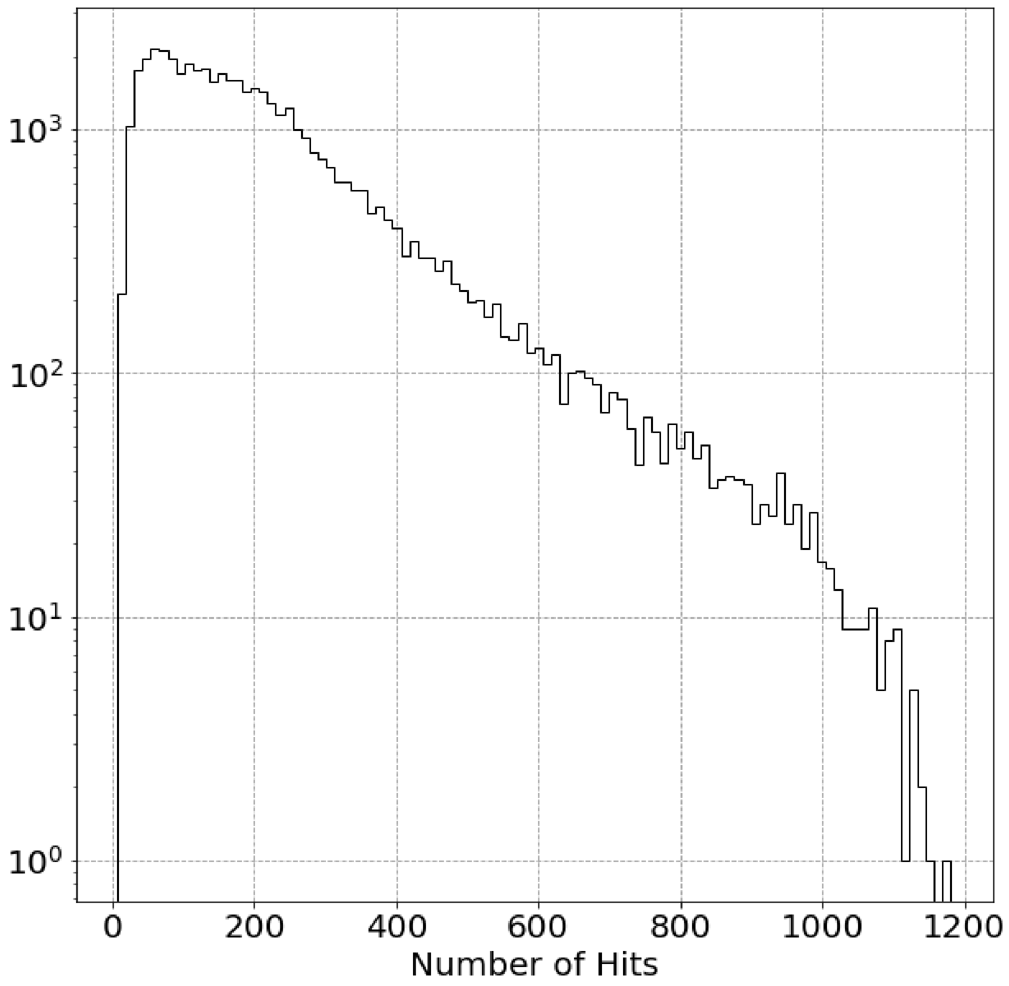
Results

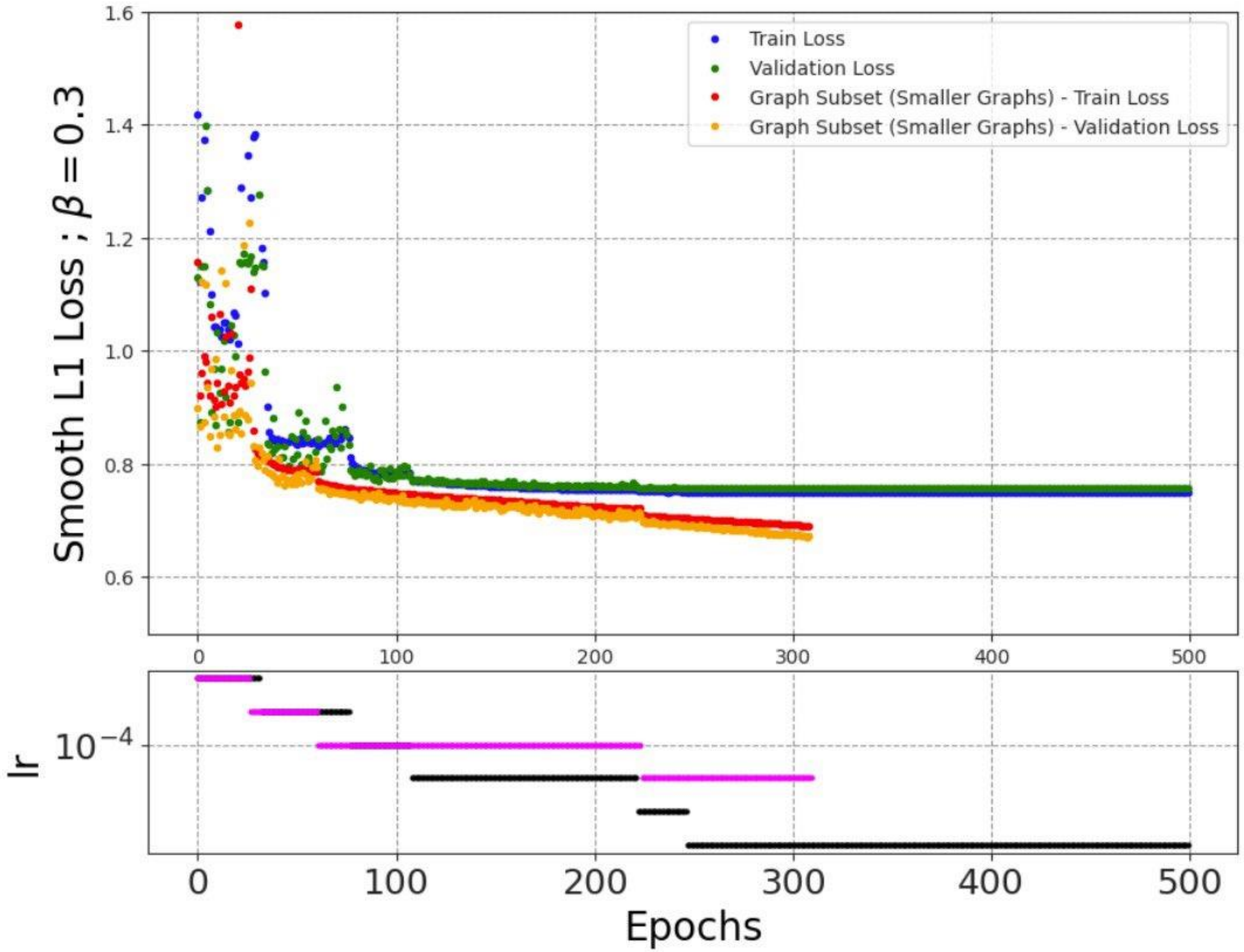
- Target Variable: 1+ln(A)
- Major Improvements for all Primary Types
  - Maximum at True value
  - Shift towards lighter elements for H and He
  - Shift towards heavier for O and Fe
- Loss:
  - Adaptive Learning rate
  - Very Gradual decrease in error
  - No-overfitting most of the times





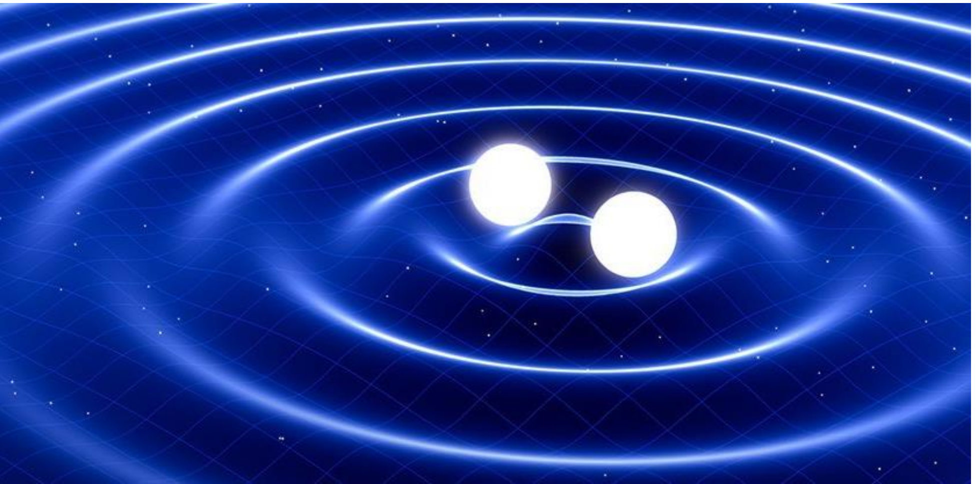
Divide and Tackle



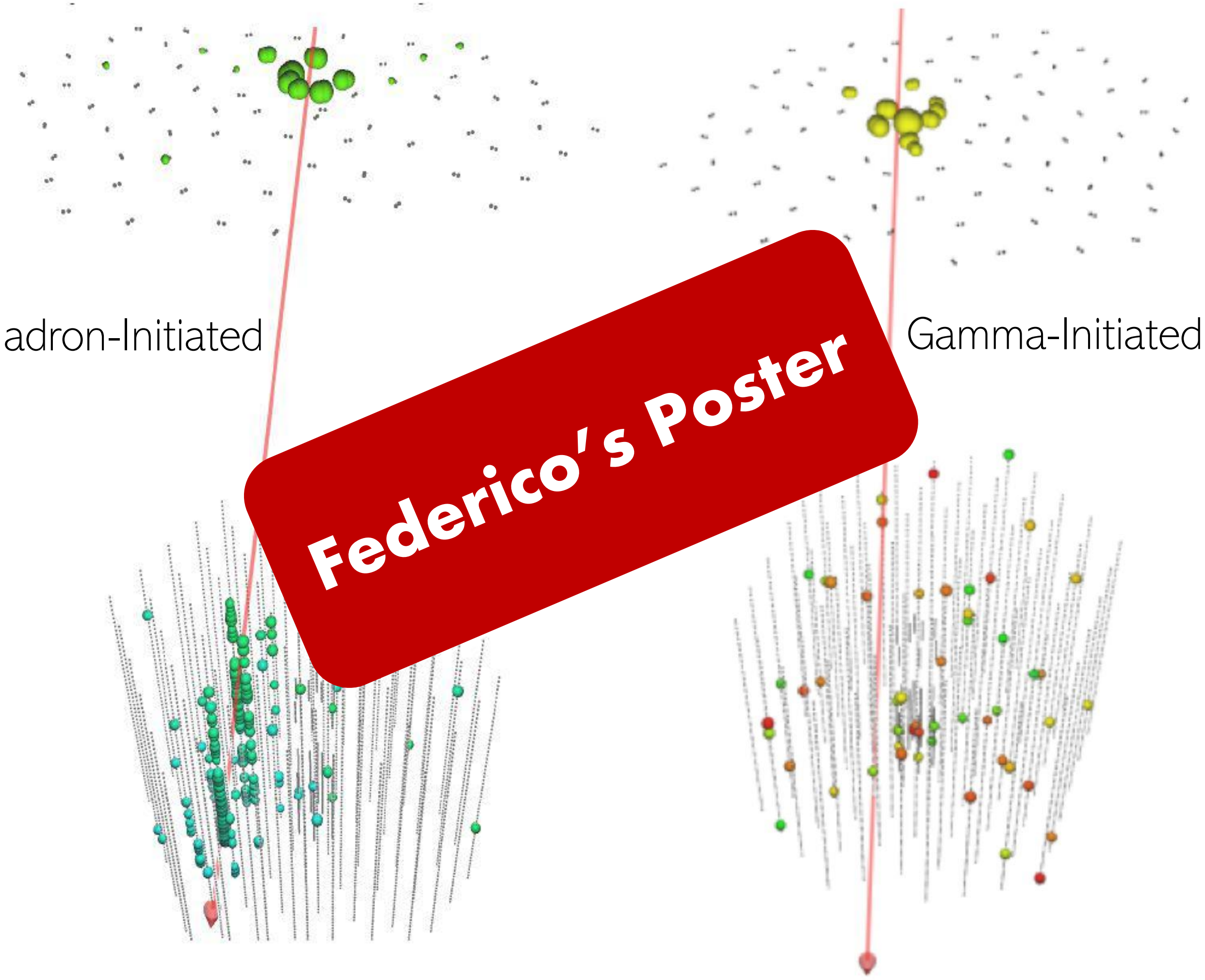


Stay Tuned...

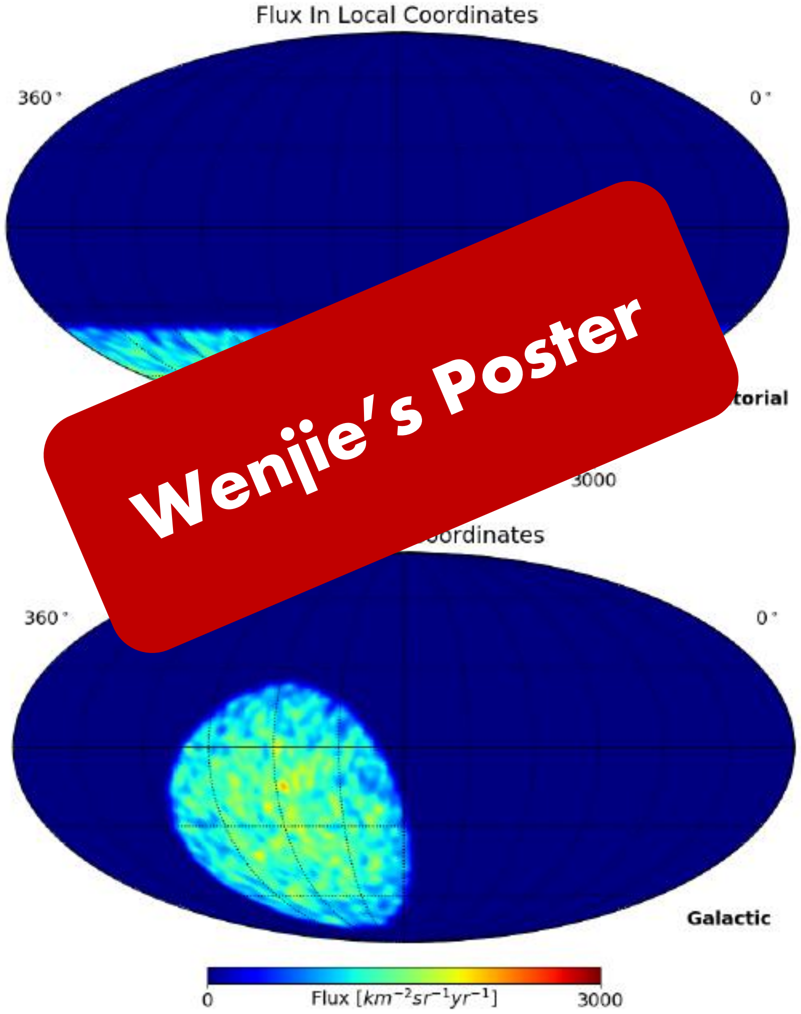



Gamma-Hadron Separation



IceCube-Gen2 Anisotropy Studies



Summary

- New composition-sensitive parameters
- Preliminary indication of lighter-composition at knee
- Promising results for composition analysis
- Heavy-involvement of KIT in detector R&D for hybrid surface array


Questions