Neutrino Physics with Deep Learning

A Summary of Applications, Successes, and Lessons.

Fernanda Psihas

Fermilab

Based on: "Review of Machine Learning for Neutrino Physics."
For more on my current research see www.fernandapsihas.com

Takeaways

Neutrino experiments use ML in a an increasingly **high-rates & rare-signal** environments where systematic uncertainties will soon dominate.

CNNs and Graph Nets for ID and reconstruction are common, and **new fast ML implementations** will address the data challenges of next generation experiments.

Neutrino experiments must **solve the bias and uncertainty problem in ML** to enable precision measurements for next-generation experiments.

Outline

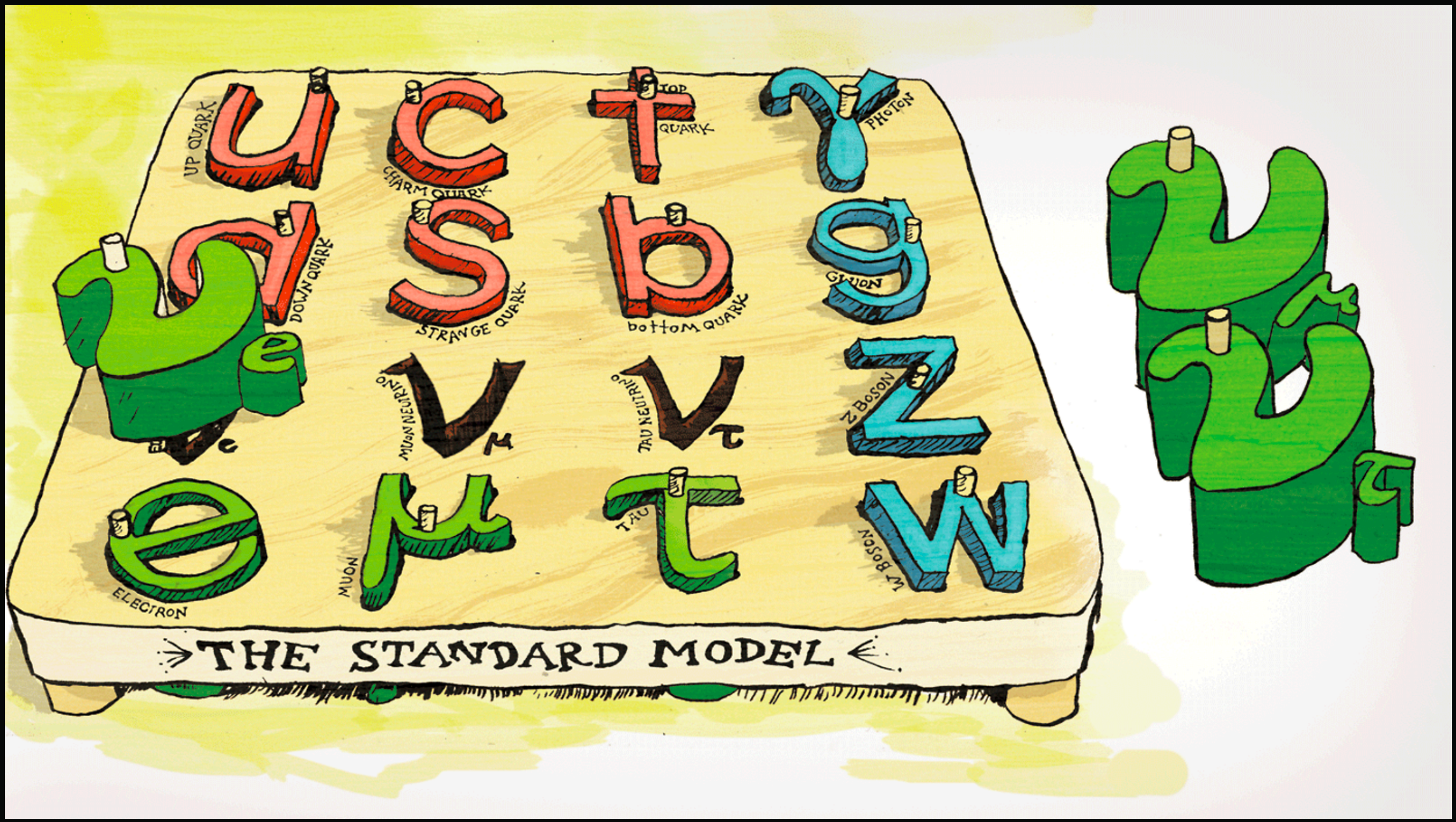
Experimental questions in neutrino physics

Neutrino signals & Neutrino detectors

Common applications & current trends

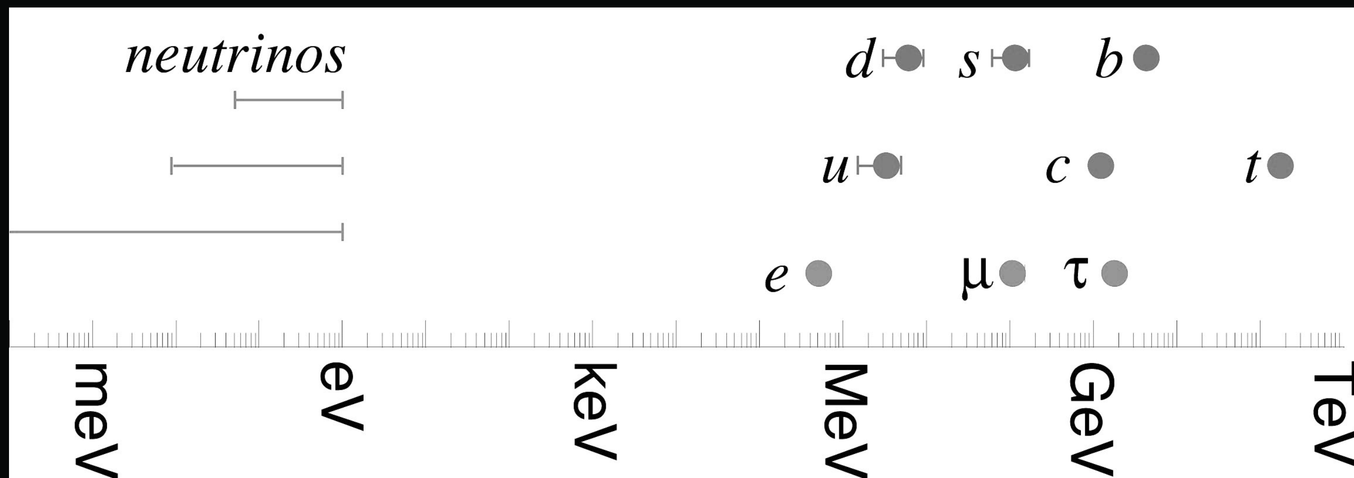
Addressing ν experiment challenges & ML challenges

The Standard Model of Particle Physics



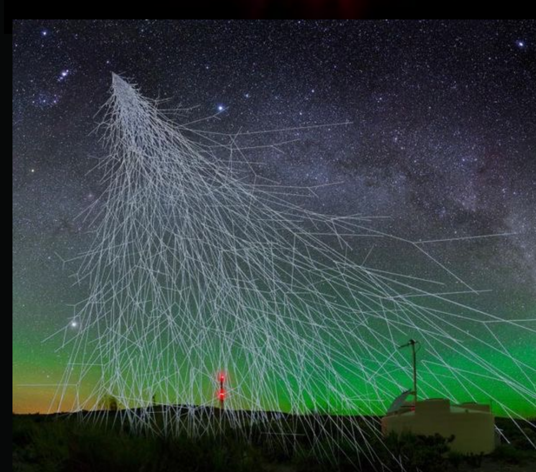
The evidence for non-zero masses of neutrinos is **evidence of physics beyond the Standard Model.**

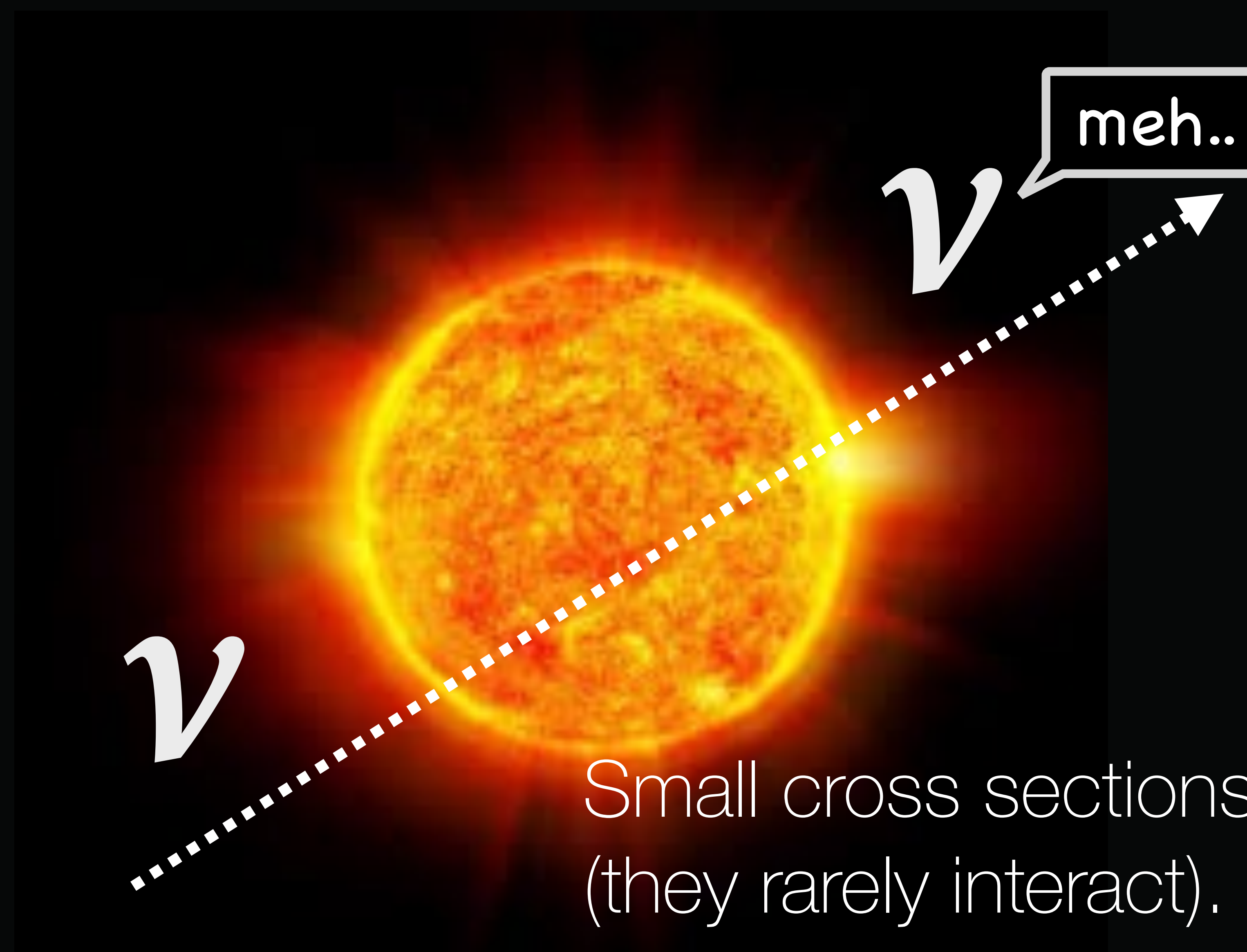
Ubiquitous & Elusive



Many orders of magnitude smaller than all other elementary particles.

Neutrinos are produced in the sun, supernovae and cosmic rays.

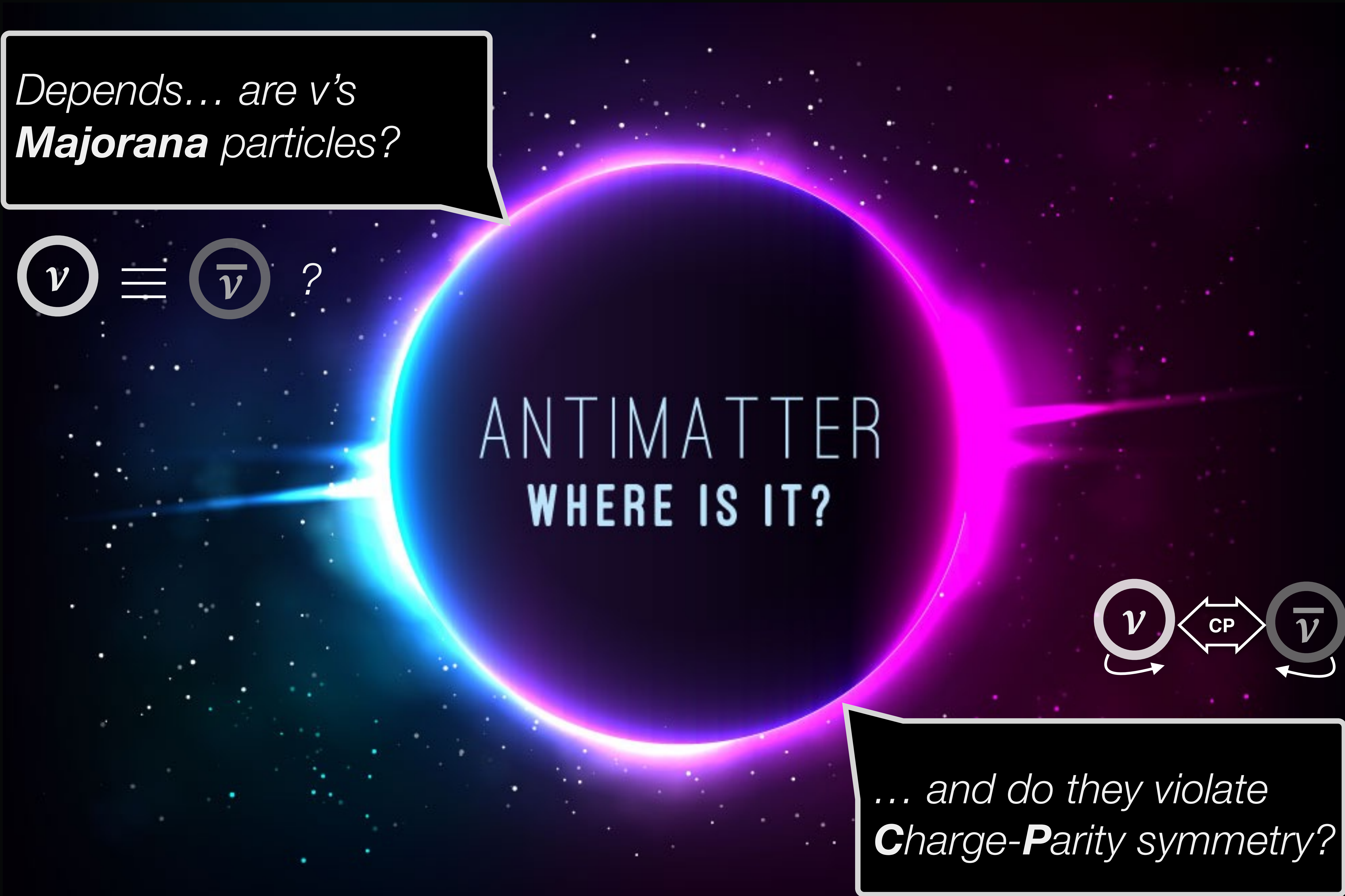




Small cross sections (they rarely interact).

Depends… are ν's ***Majorana*** *particles?*

$\nu \equiv \bar{\nu}$?



$\nu \overset{CP}{\Longleftrightarrow} \bar{\nu}$

… and do they violate ***C****harge-****P****arity symmetry?*

Remaining Questions in Neutrino Physics

ν

Do they violate Charge-Parity symmetry?

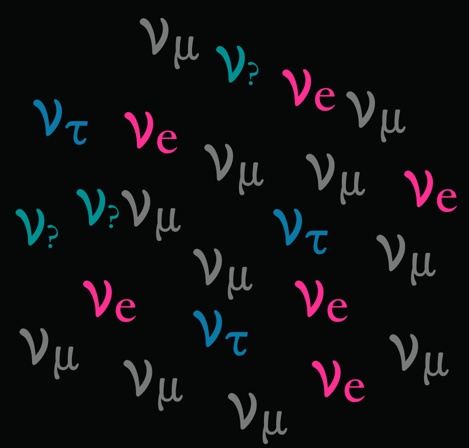
What is the complete picture of oscillations?

Are there more neutrinos beyond 3 flavors?

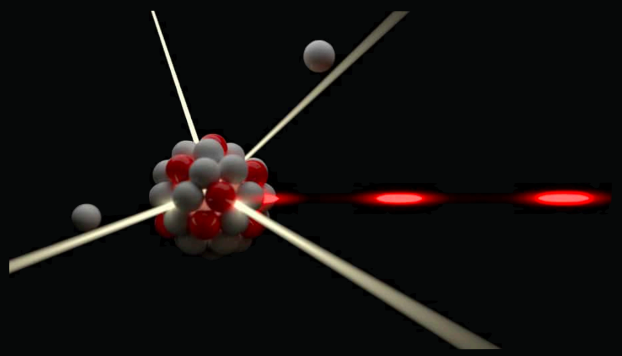
What is the ordering of the neutrino m**asses?**

Are neutrinos Dirac or Majorana?

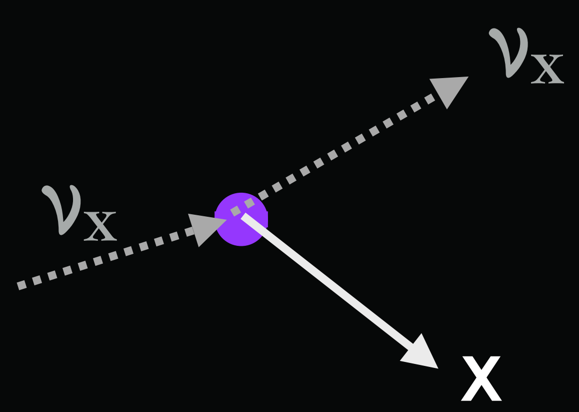
Experiments Studying the Remaining ν Mysteries



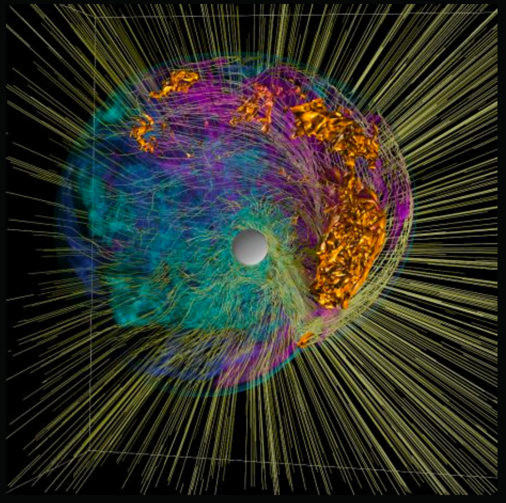
STUDY NEUTRINO OSCILLATIONS



SEARCH FOR RARE PROCESSES

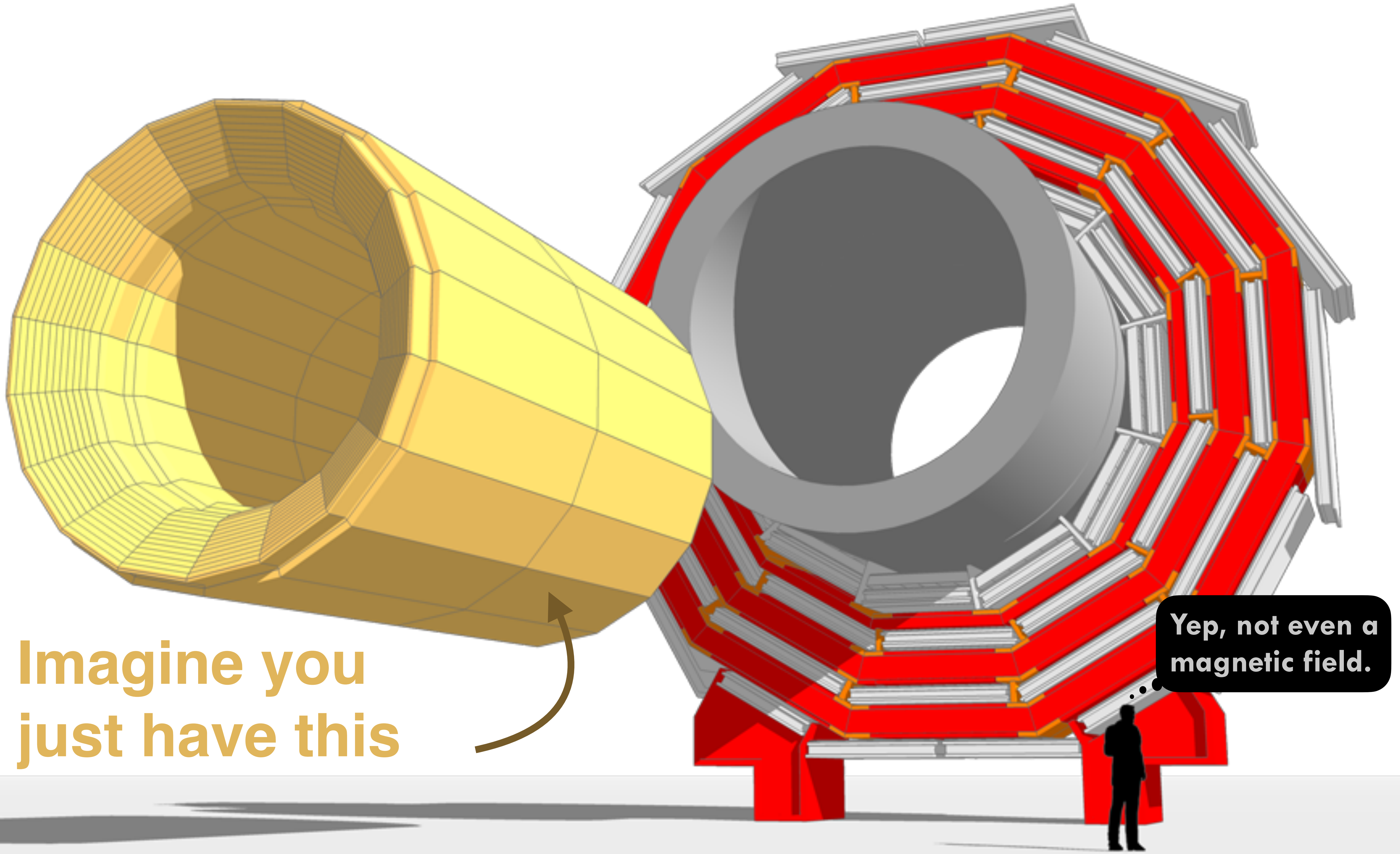


UNDERSTAND THEIR INTERACTIONS



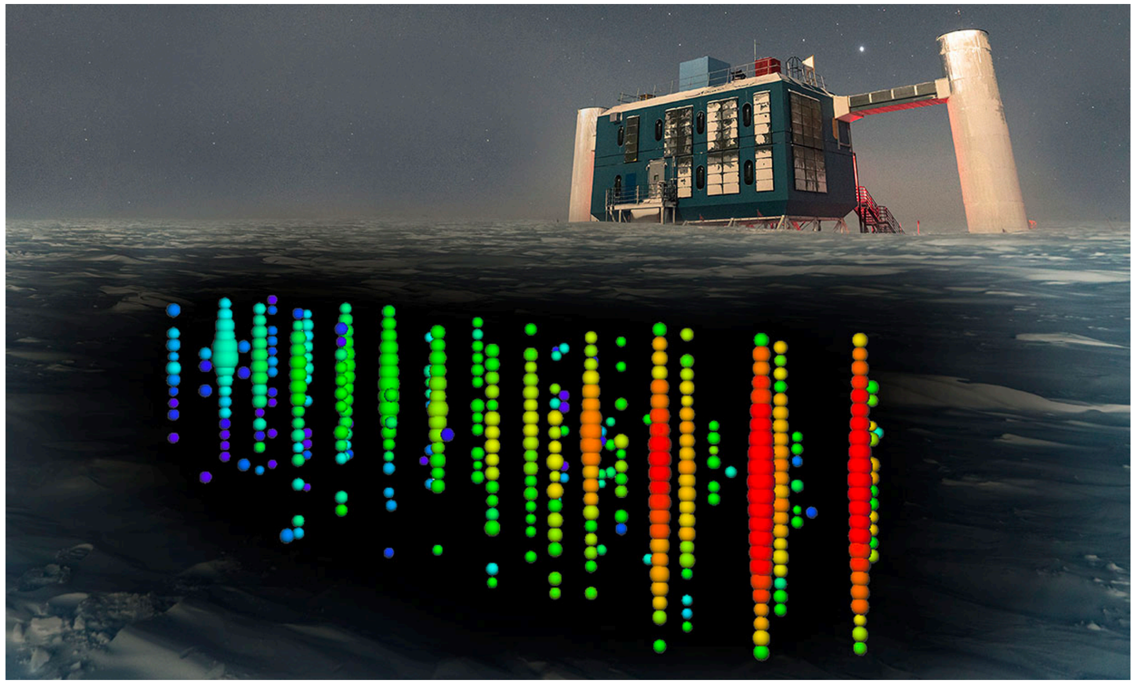
CONNECTIONS WITH ASTRO. PHENOMENA

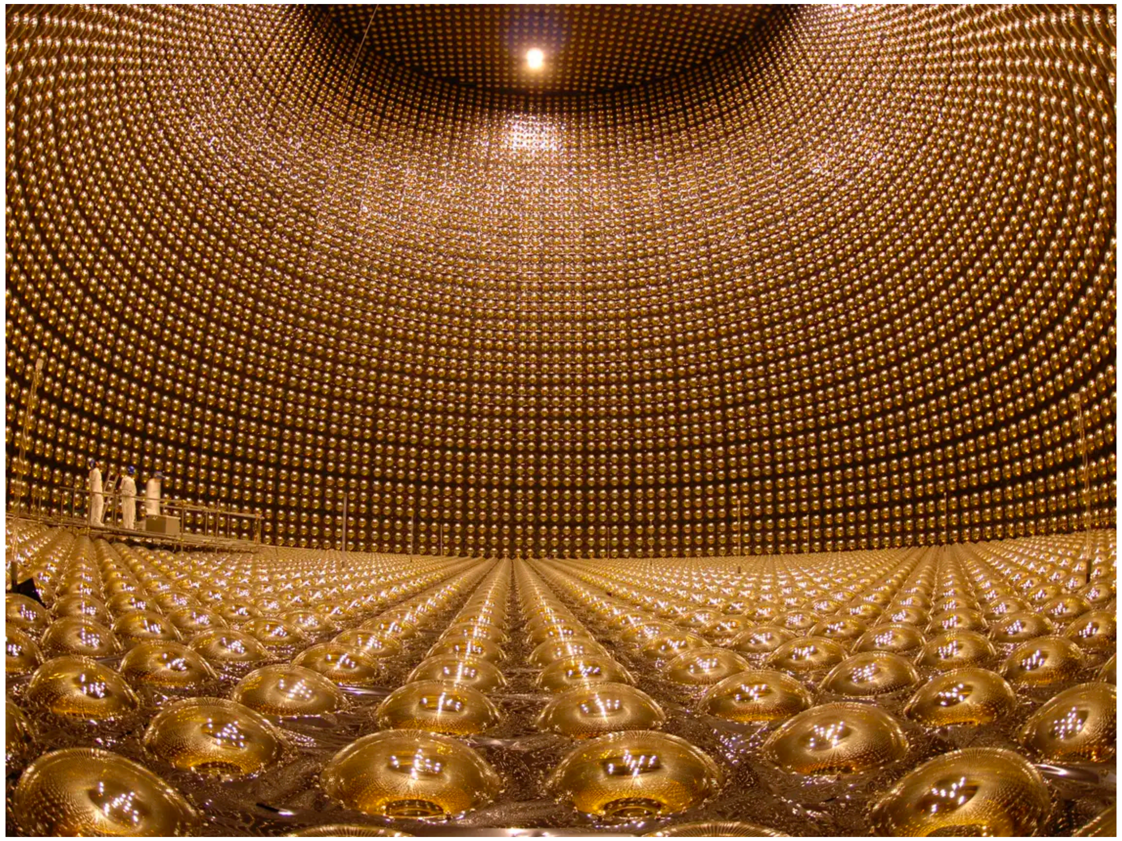
Some collider context

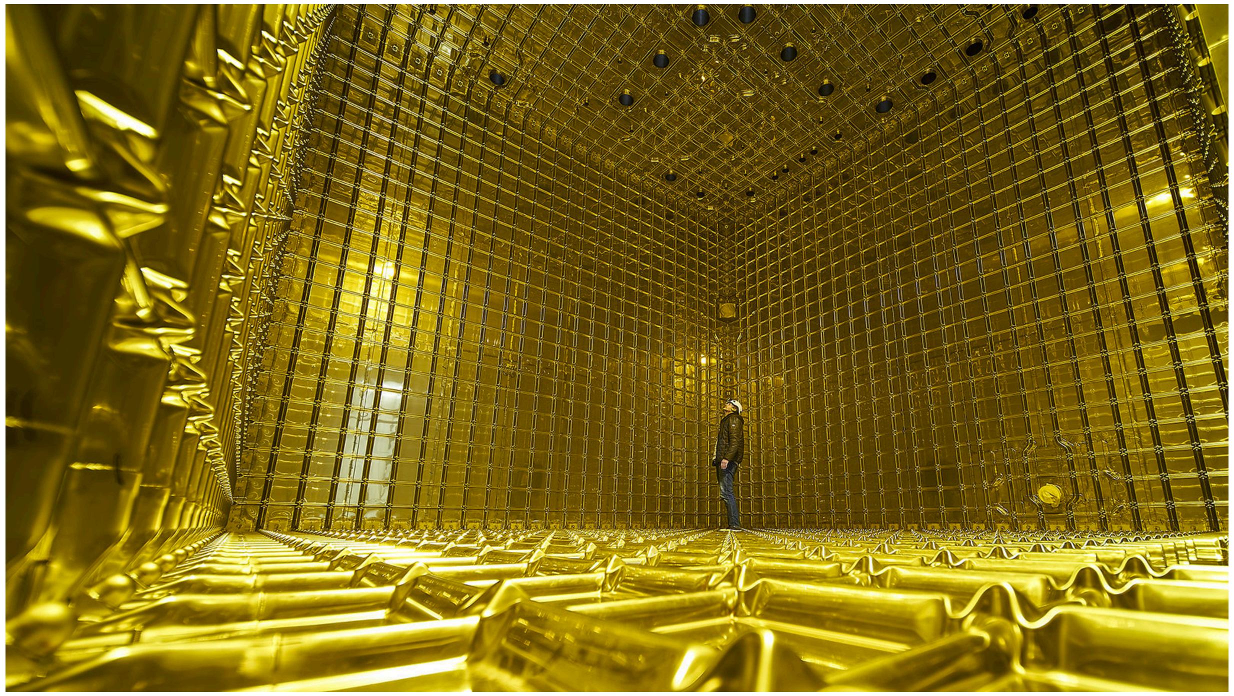


https://twiki.cern.ch/twiki/bin/view/CMSPublic/SketchUpCMS

Neutrino detectors







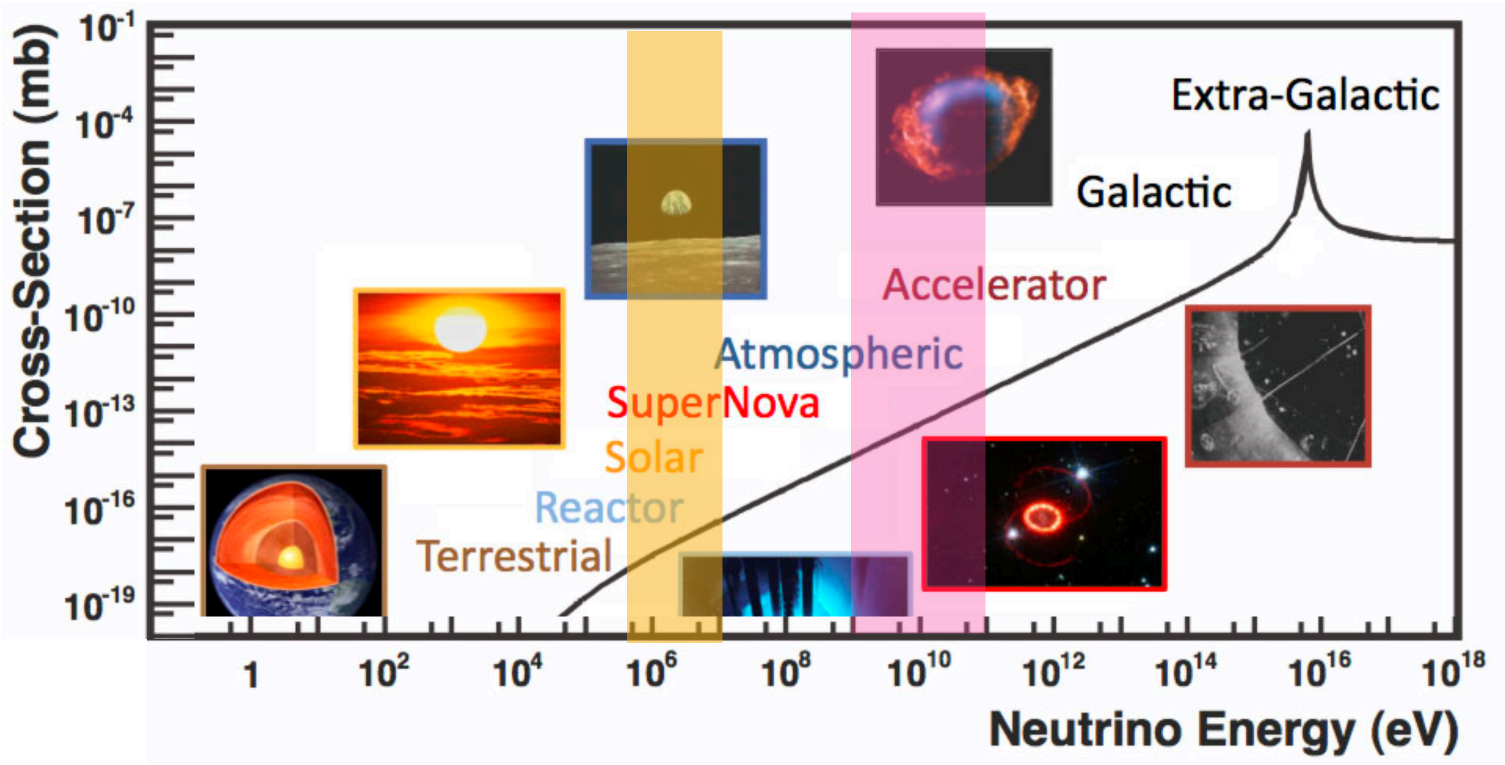
Large active volume.

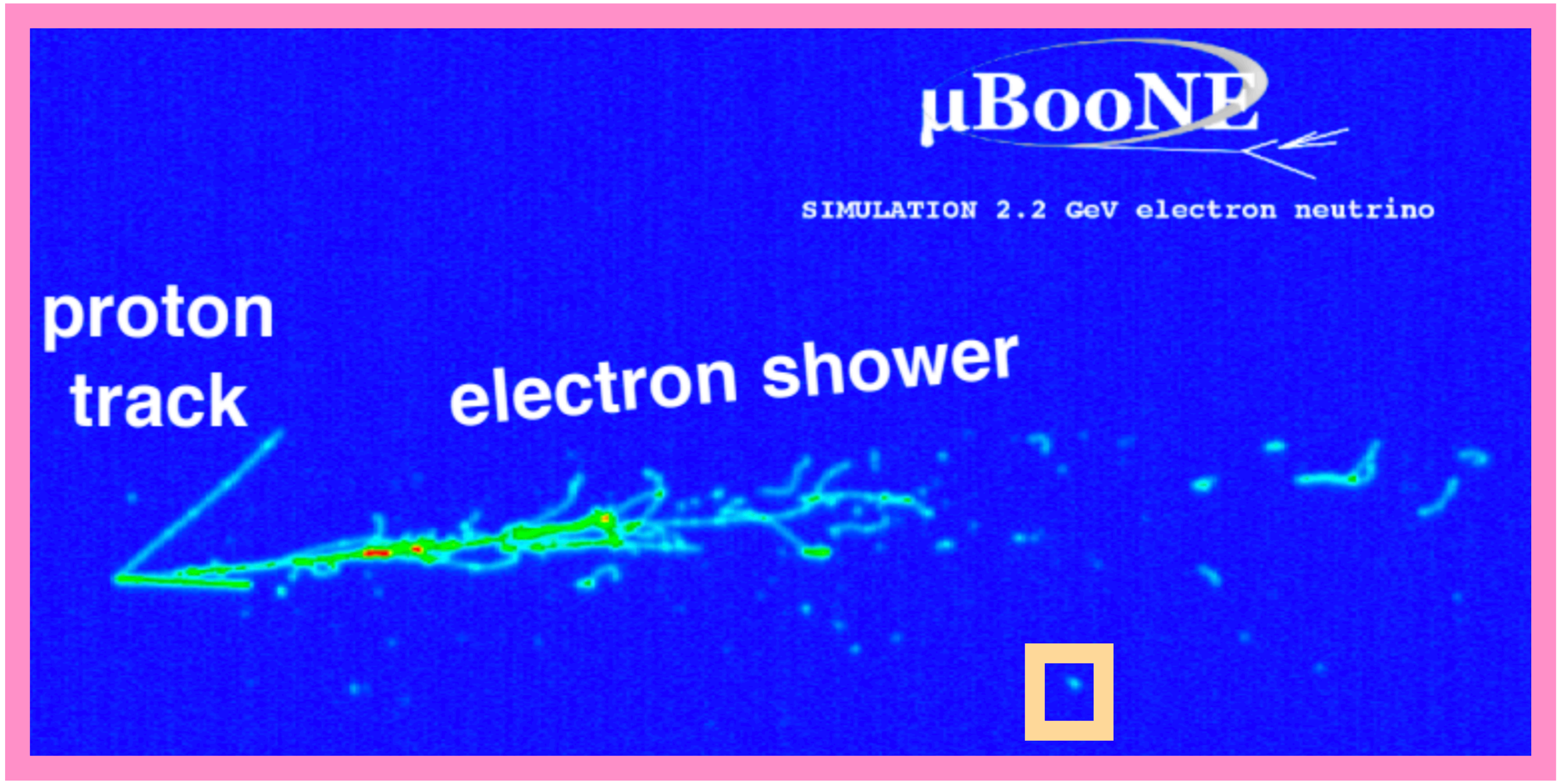
Often homogenous active material.

Neutrino events

Overall very small cross-section interactions

Wide range of energies, sometimes studied in the same detector.

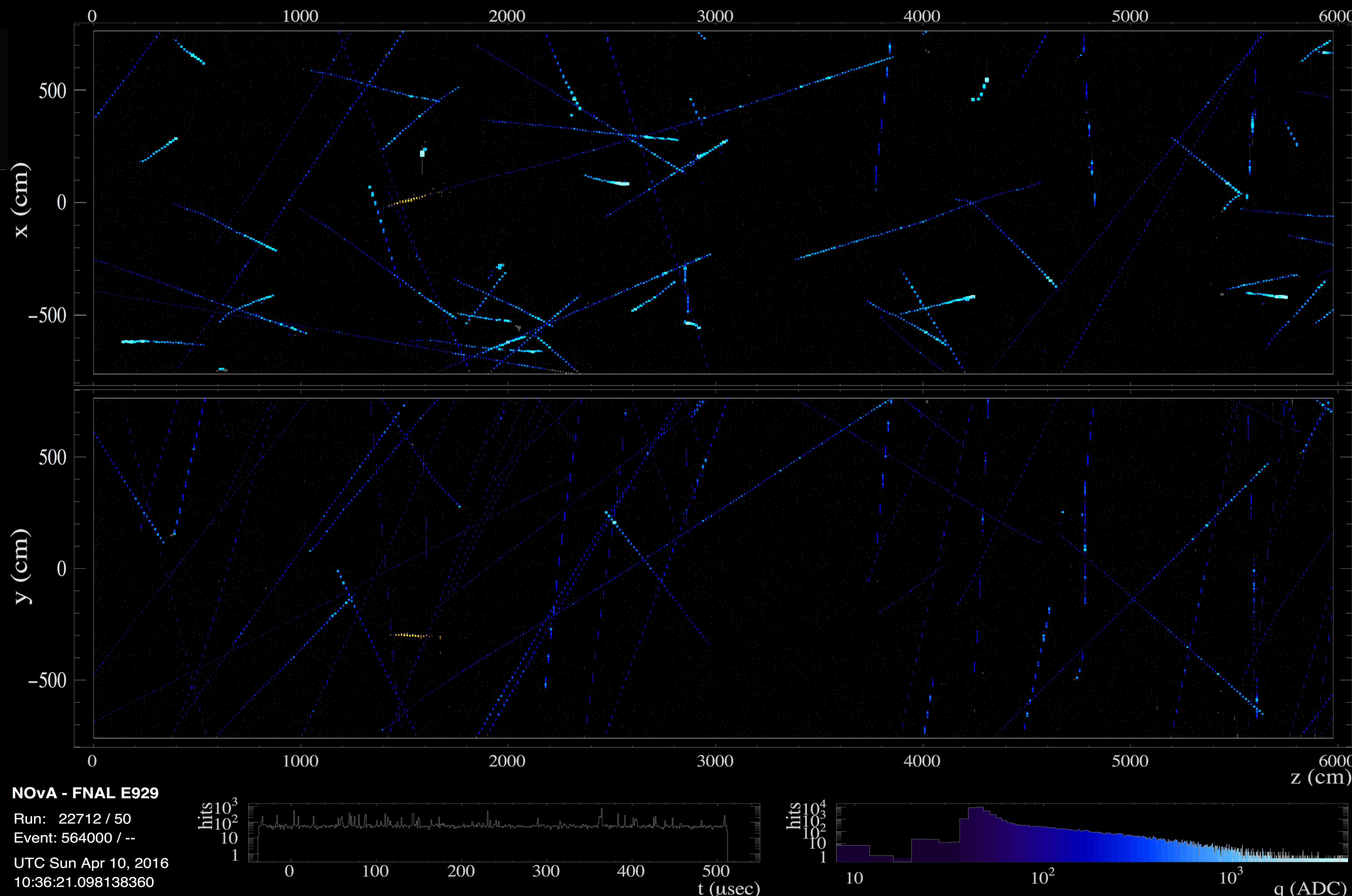




Increasing precision needs and data rates.

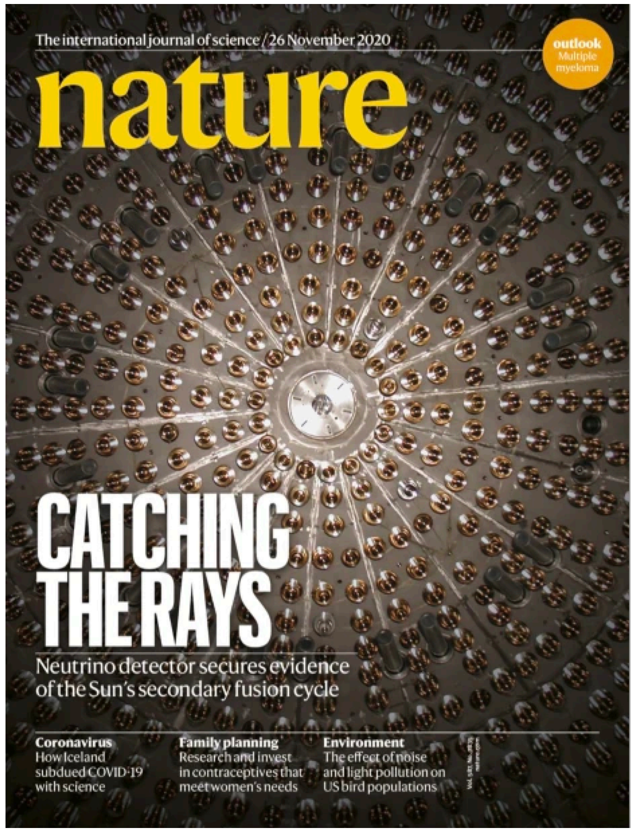
Low occupancy events.

NOvA Raw detector data



Physics applications & impact

Event-by-event tagging of the 210Po decay α **is essential to the CNO detection.**



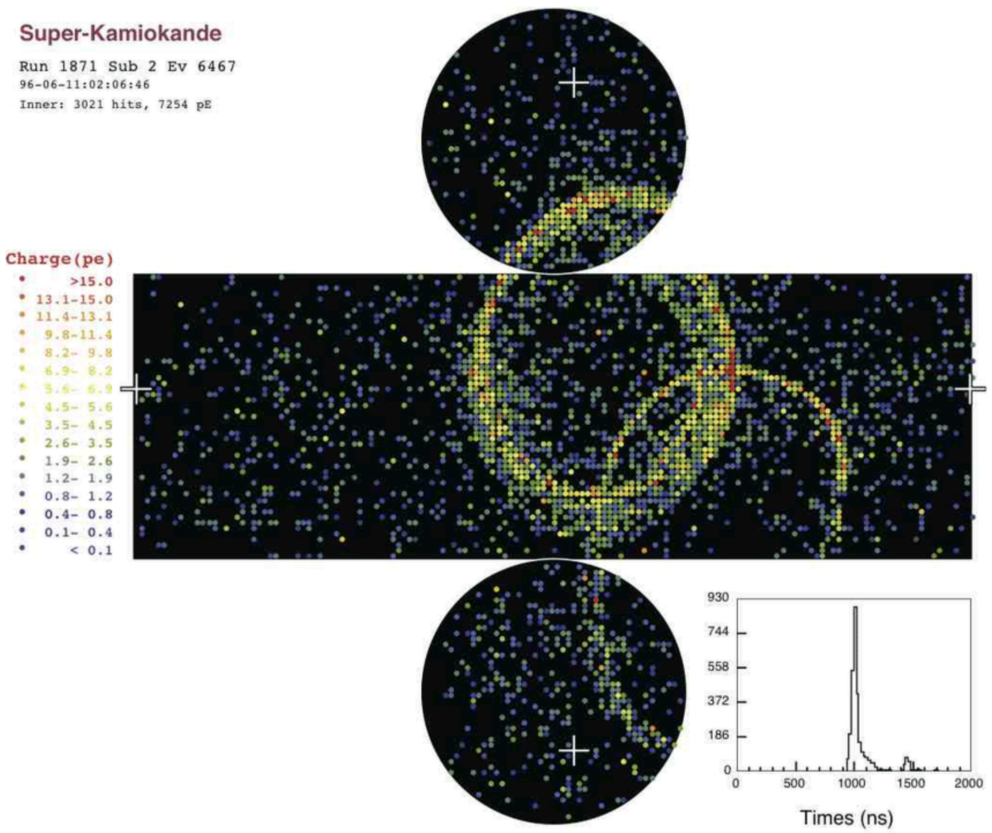
Experimental evidence of neutrinos produced in the CNO fusion cycle in the Sun

The Borexino Collaboration

Nature **587**, 577–582 (2020) | Cite this article

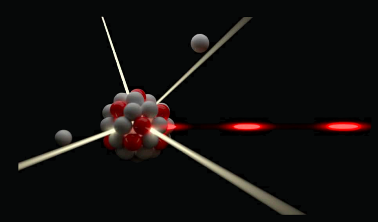
This is enabled by a Multi-layer perceptron. The **MLP exploits the scintillation time-decay differences from alpha and beta-like events.**

Adoption of BDT improves their multi-site tagging of e CCve-like events, from **34.4% to 46.7% in efficiency.**

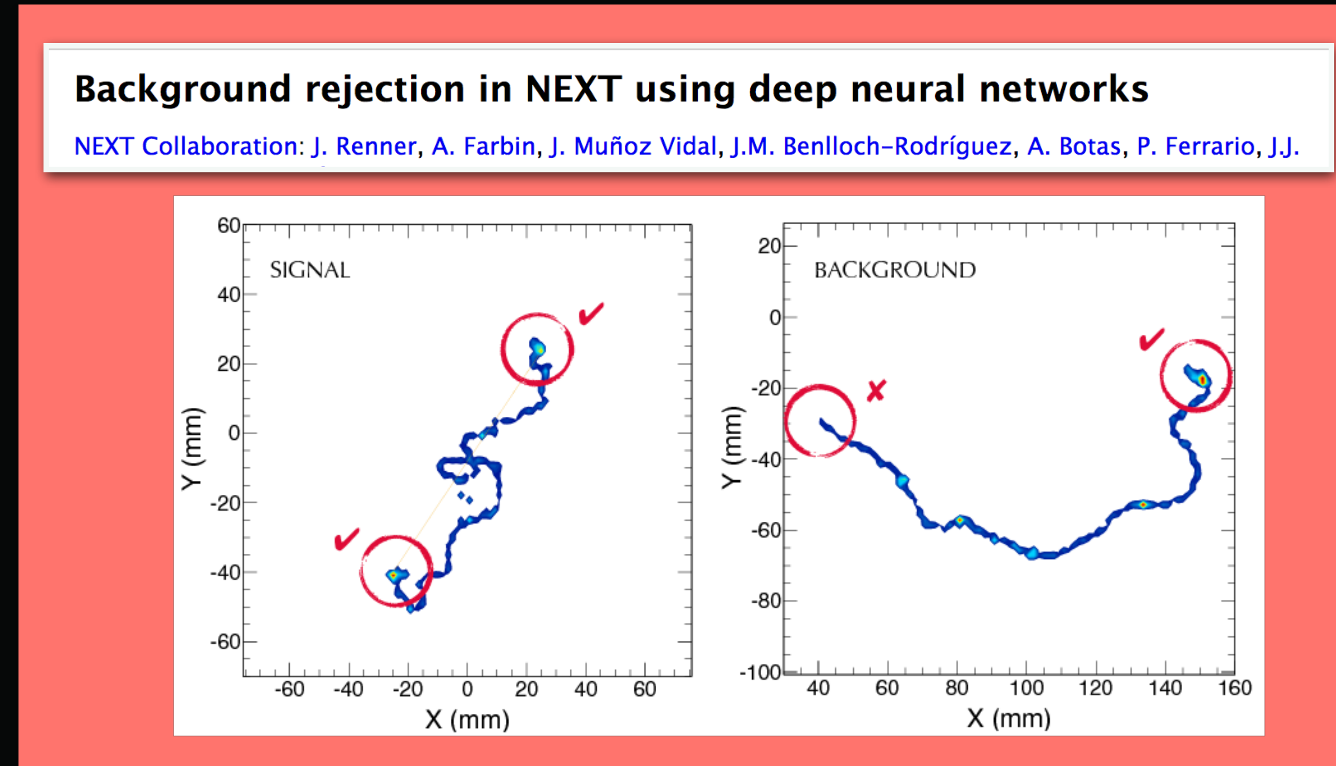


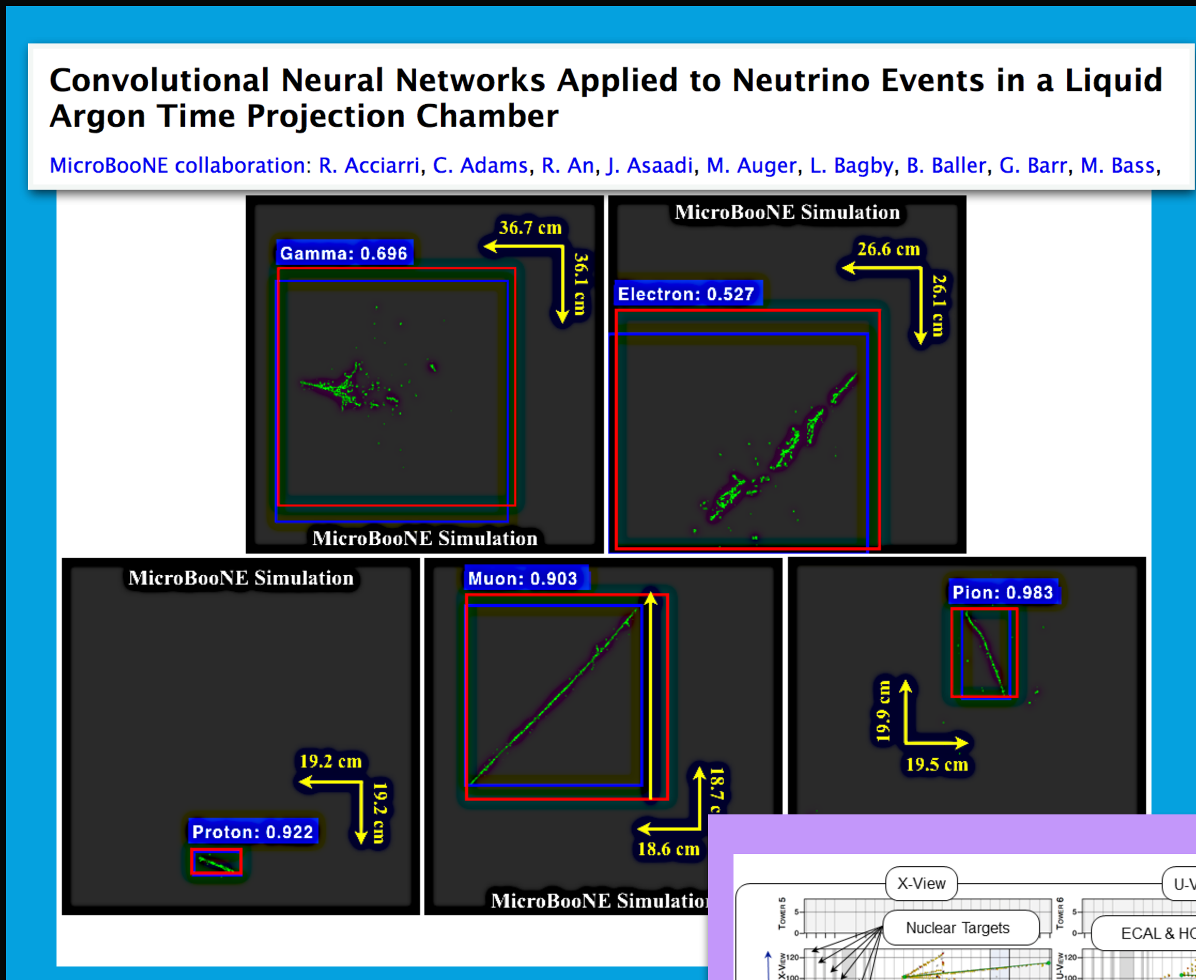
(ICRC2021) Atmospheric oscillations with Super-Kamiokande and prospects for SuperK-Gd - https://pos.sissa.it/395/008

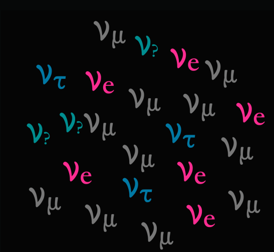
https://arxiv.org/abs/1109.3262



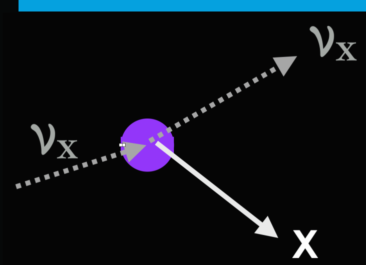
Rare Processes



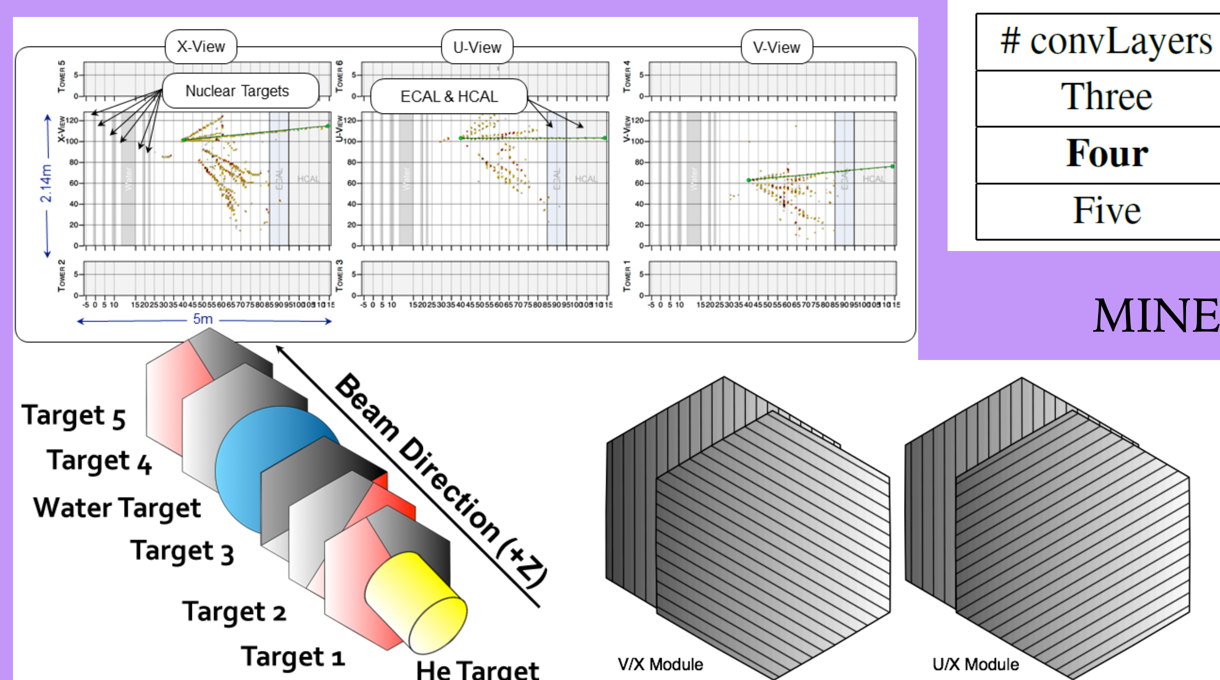




Neutrino Oscillations



Neutrino Interactions

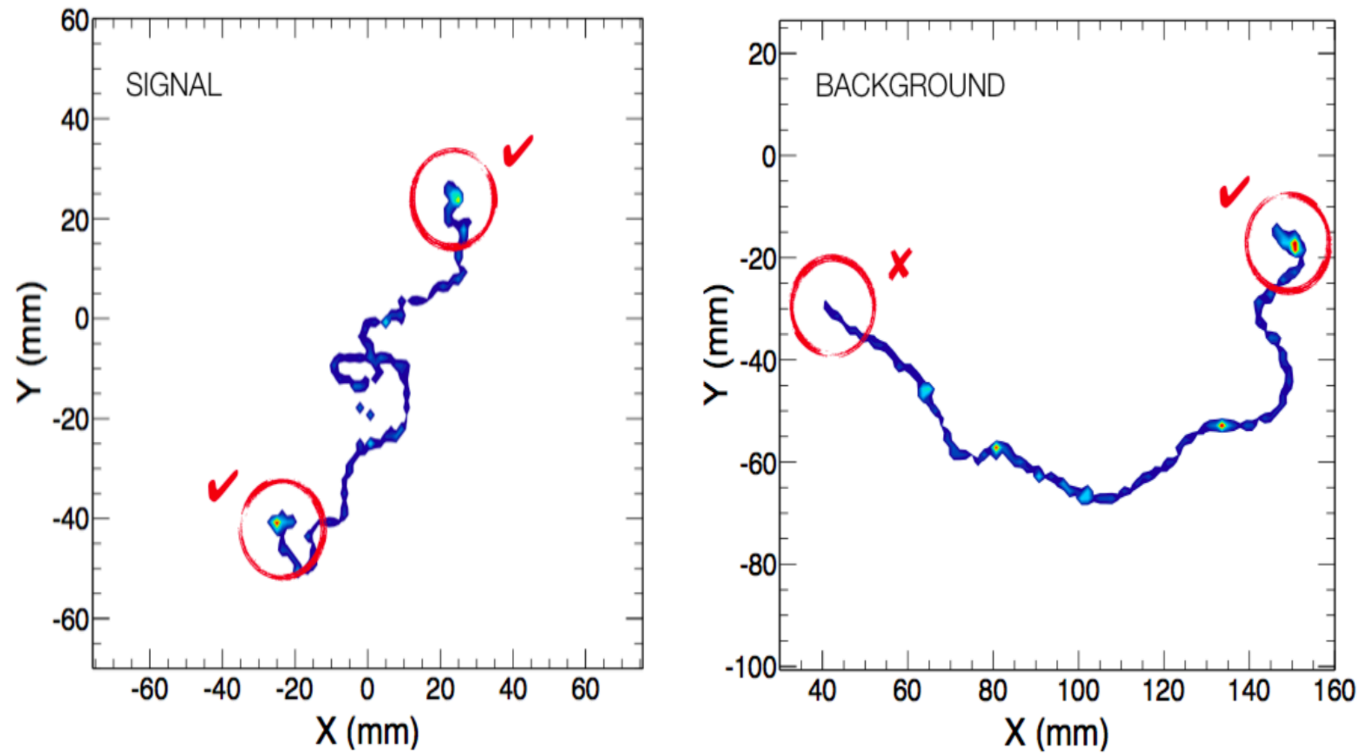


# convLayers	Kernel Sizes $(\{h\} \times w)$	Accuracy
Three	$\{6, 6, 3\} \times 3$	93.58%
Four	$\mathbf{\{8, 8, 7, 6\} \times 3}$	**94.09%**
Five	$\{8, 7, 7, 3, 3\} \times 3$	93.55%

MINERvA uses a CNN to localize the vertex/interaction point.

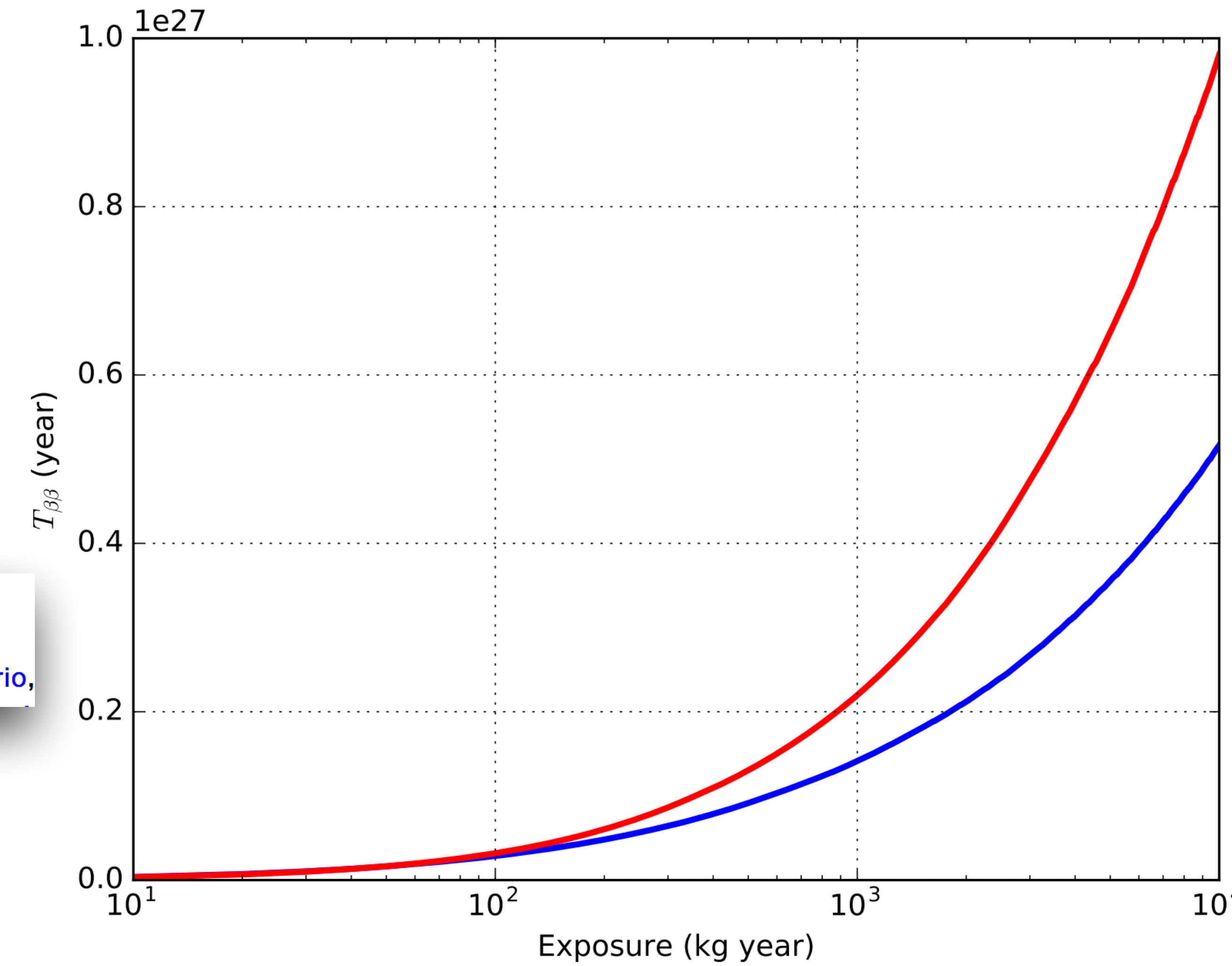
Varying network parameters they accomplish 94% accuracy for vertex location in the target material.

CNNs in Gas TPCs



Inputs are the 2D projections, with voxel size optimization.

Applied out of the box GoogLeNet architecture already promises large improvements in potential sensitivity.



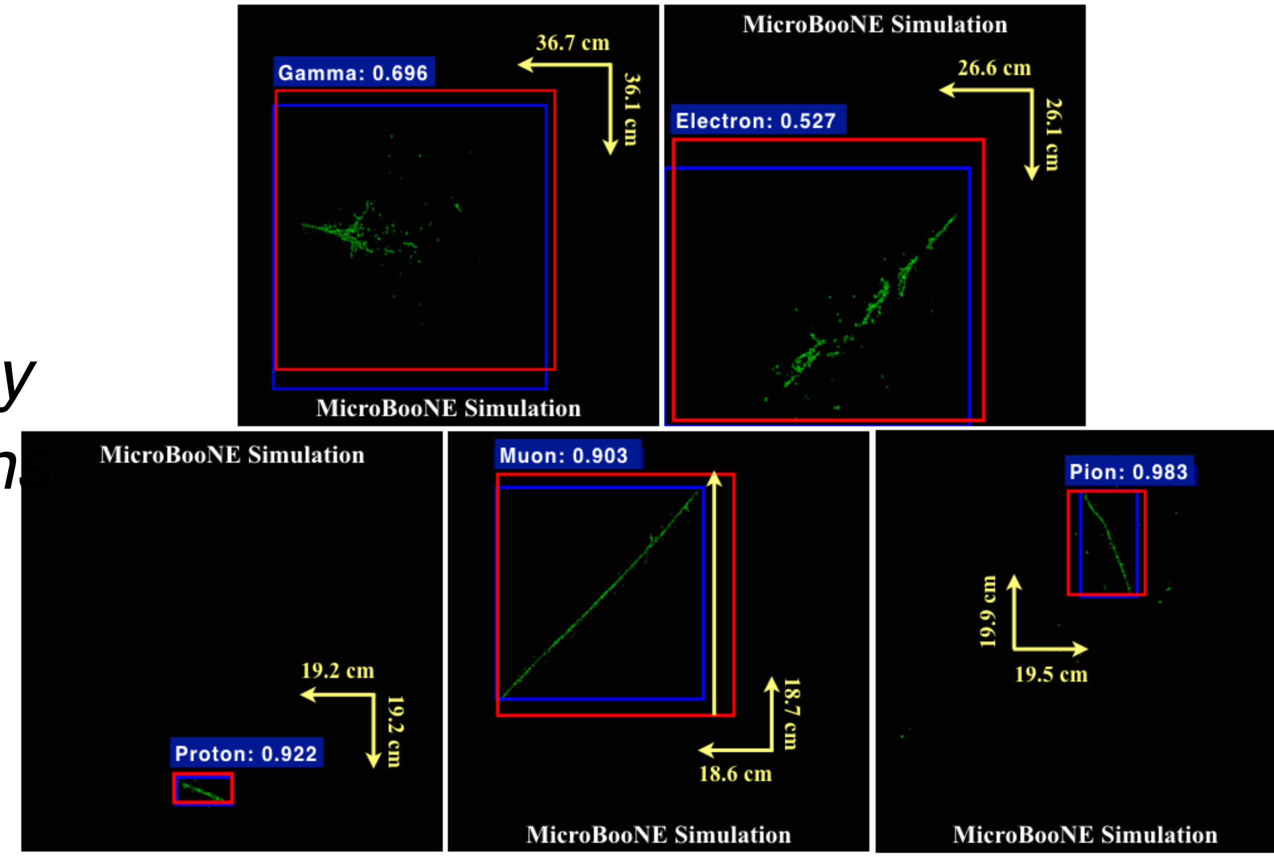
Background rejection in NEXT using deep neural networks

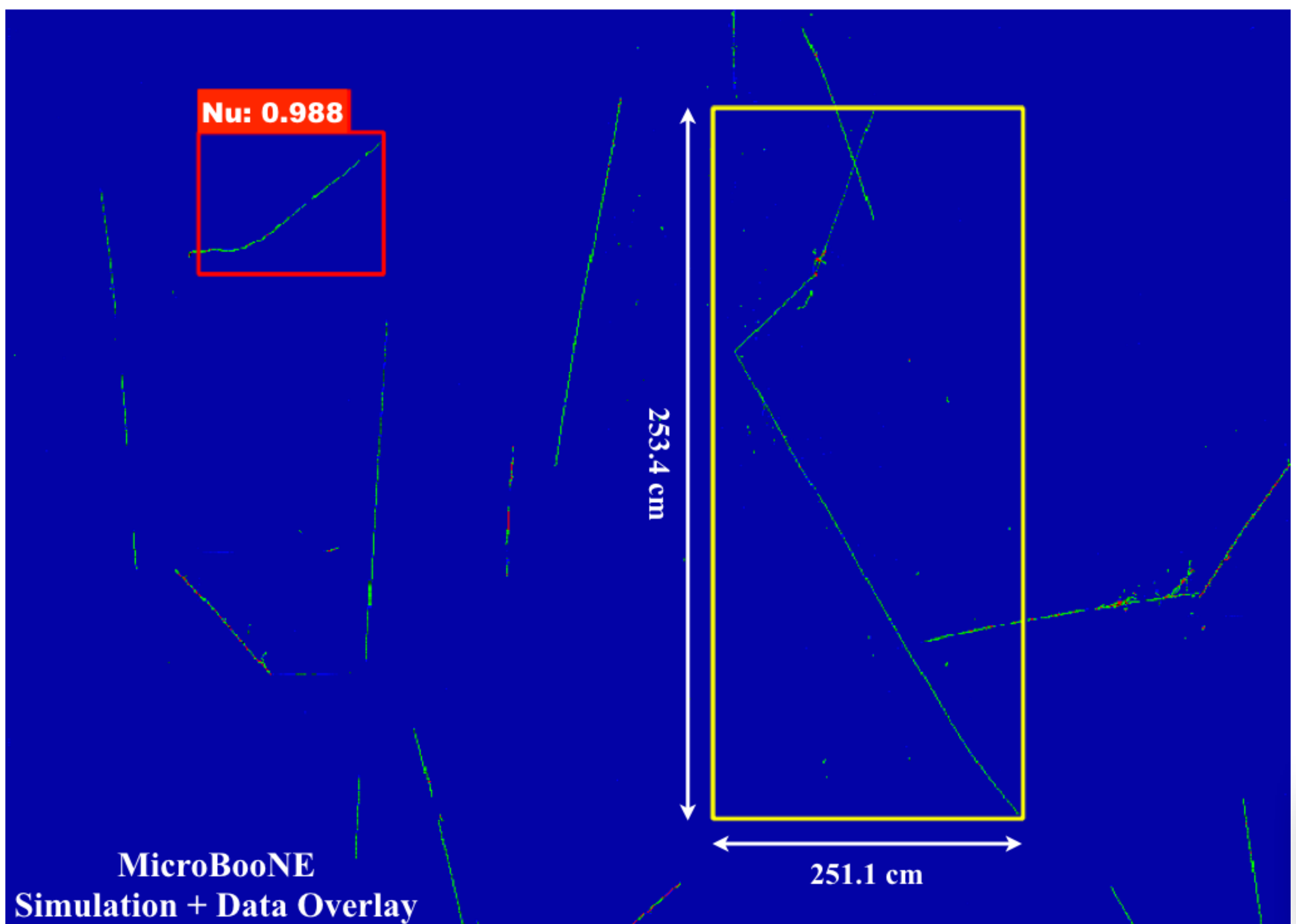
NEXT Collaboration: J. Renner, A. Farbin, J. Muñoz Vidal, J.M. Benlloch-Rodríguez, A. Botas, P. Ferrario,

CNN's For NEUTRINO EVENTS

MicroBooNE is exploring CNN implementations on LAr-TPC for:

- ★ *Neutrino interaction detection 85% efficiency*
- ★ *Multi-particle classification 83% efficiency for electrons and 95% efficiency for muons*



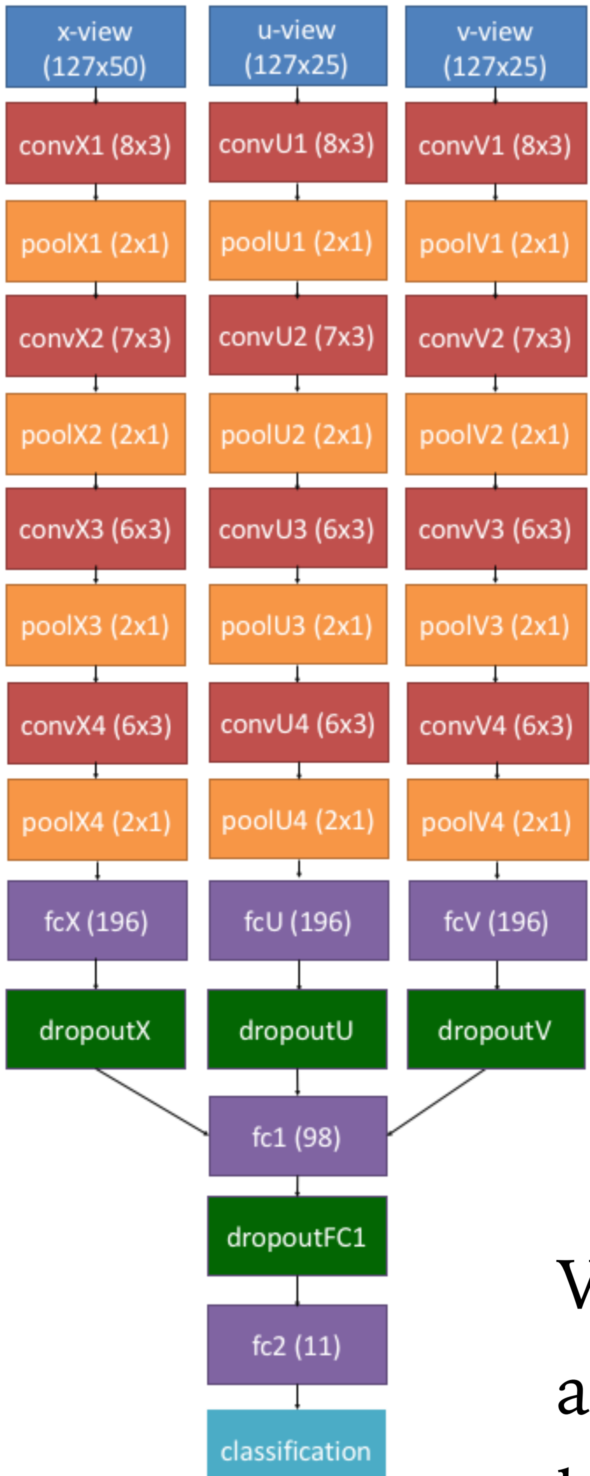


Explored challenges GPU performance vs downsampling effects for large LAr-TPCs

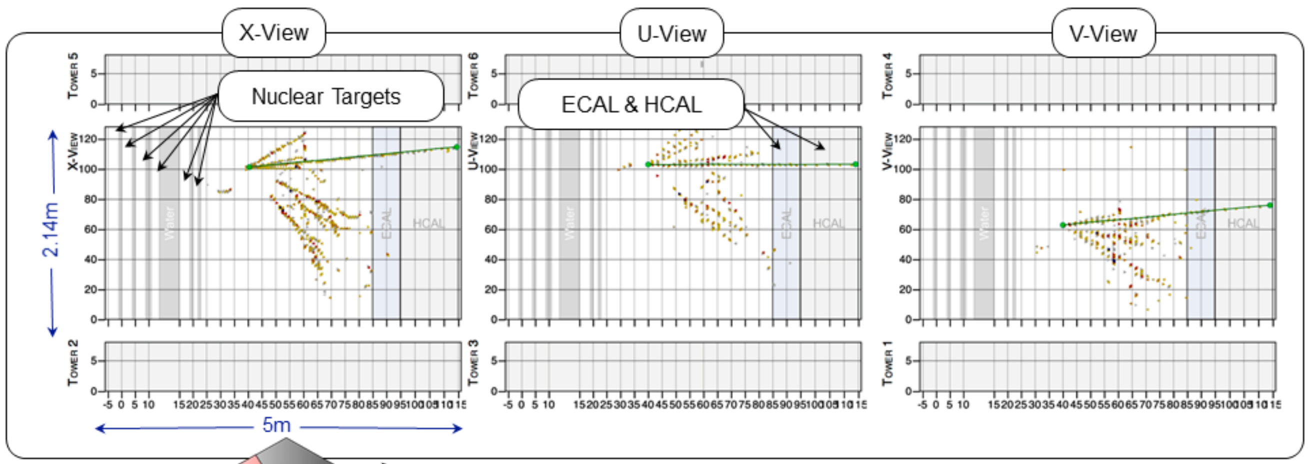
Convolutional Neural Networks Applied to Neutrino Events in a Liquid Argon Time Projection Chamber

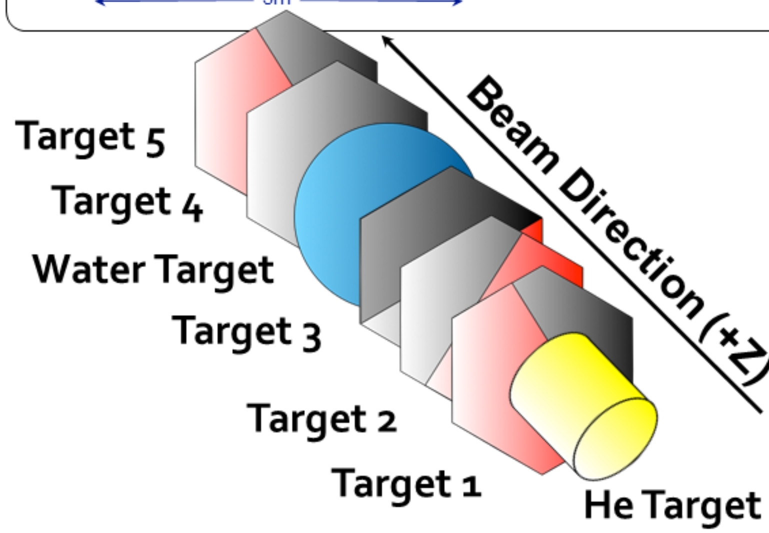
MicroBooNE collaboration: R. Acciarri, C. Adams, R. An, J. Asaadi, M. Auger, L. Bagby, B. Baller, G. Barr, M. Bass,

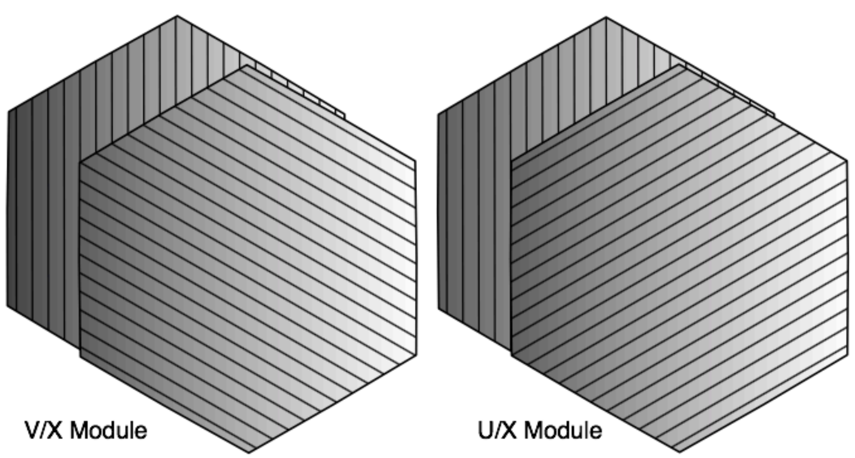
CNNs + Regression for Vertex Finding



MINERvA uses a CNN with 3 prongs in order to combine information from the X V & U views of the event.





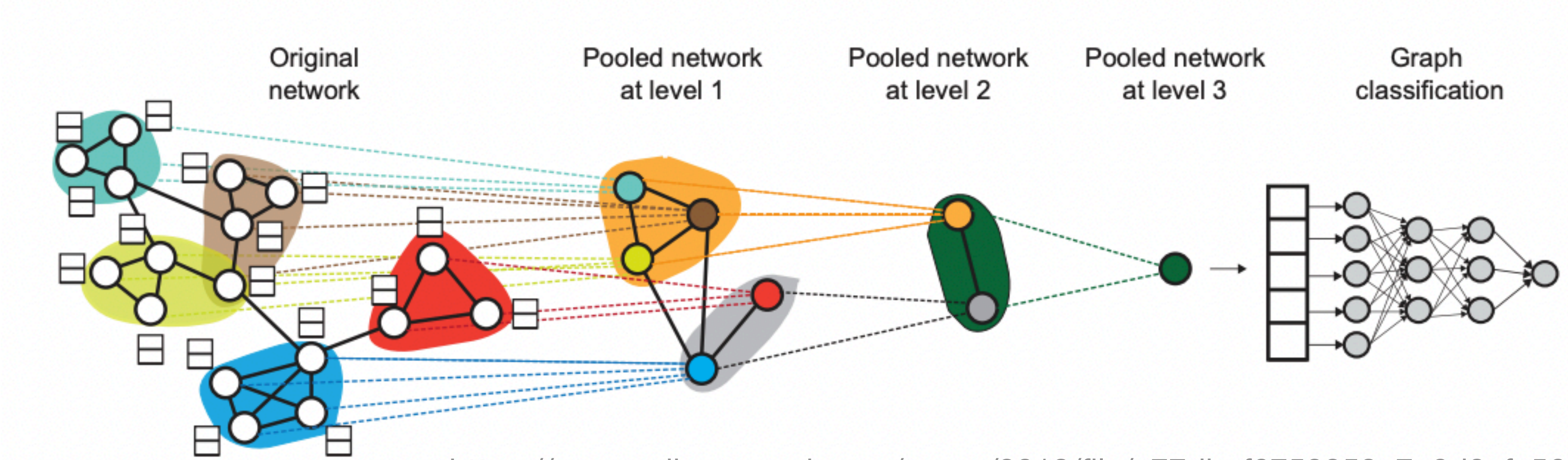


Varying network parameters they accomplish 94% accuracy for vertex location in the Z direction.

# convLayers	Kernel Sizes ($\{h\} \times w$)	Accuracy
Three	$\{6, 6, 3\} \times 3$	93.58%
Four	$\mathbf{\{8, 8, 7, 6\} \times 3}$	**94.09%**
Five	$\{8, 7, 7, 3, 3\} \times 3$	93.55%

Incorporating Detector Geometry

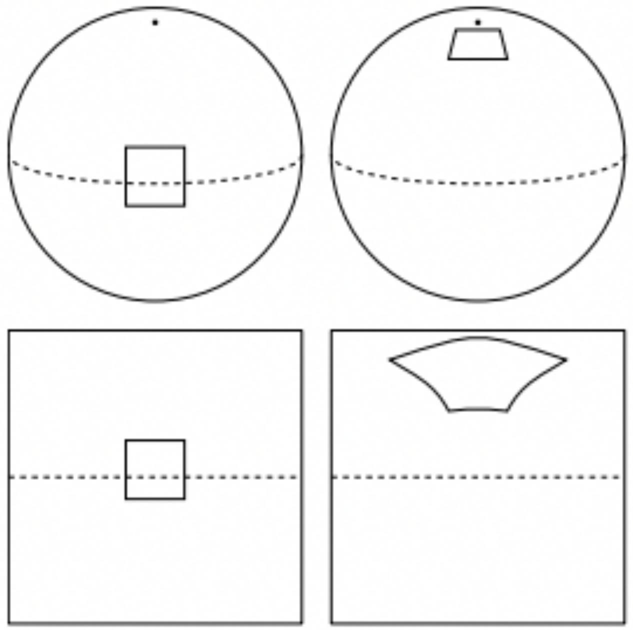
Graph Neural Networks infer the directional correlations from features and relative positioning of elements in the training data, which is useful for clustering populations.



https://proceedings.neurips.cc/paper/2018/file/e77dbaf6759253c7c6d0efc5690369c7-Paper.pdf

Spherical CNNs use projections of 2D arrays onto a spherical plane. Good example of potential for adapting CNNs to detector geometry.

https://arxiv.org/abs/1801.10130

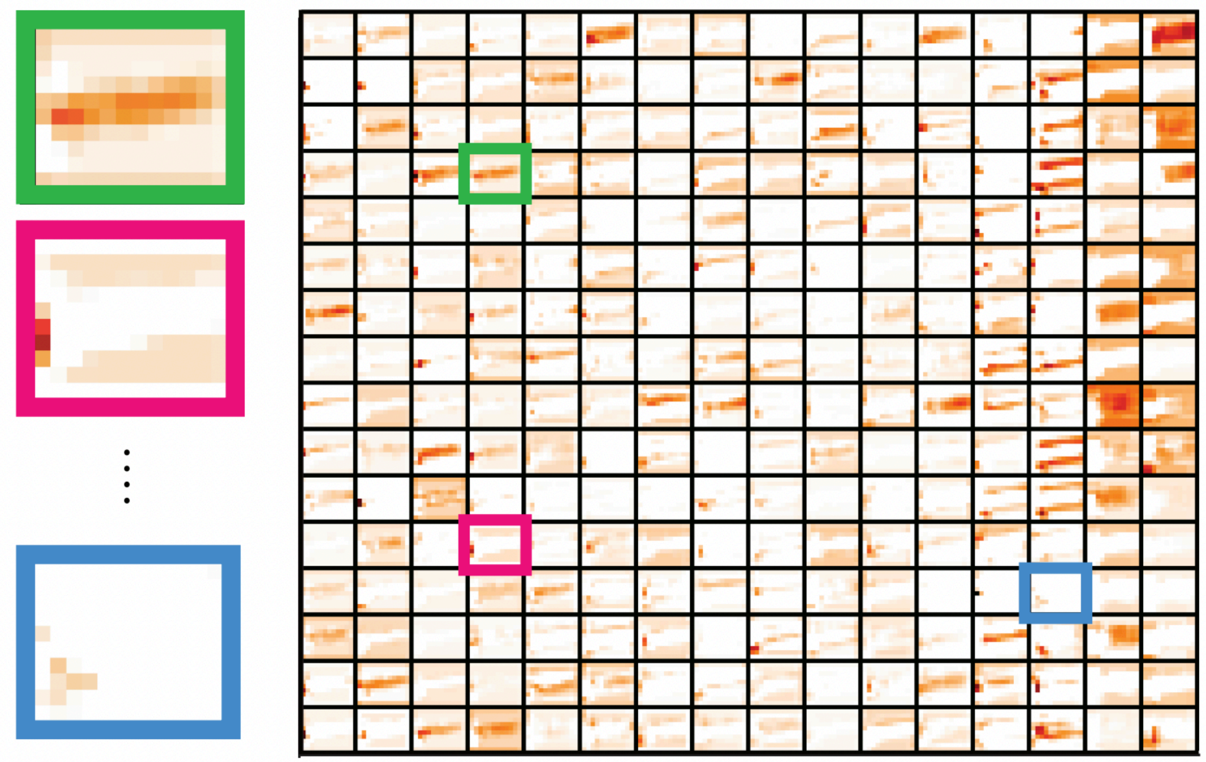


A convolutional neural network neutrino event classifier

A. Aurisano[1], A. Radovic[2], D. Rocco[3], A. Himmel[4], M.D. Messier[5], E. Niner[4], G. Pawloski[3], F. Psihas[5], A. Sousa[1] and P. Vahle[2]

Published 1 September 2016 • © 2016 IOP Publishing Ltd and Sissa Medialab srl
Journal of Instrumentation, Volume 11, September 2016

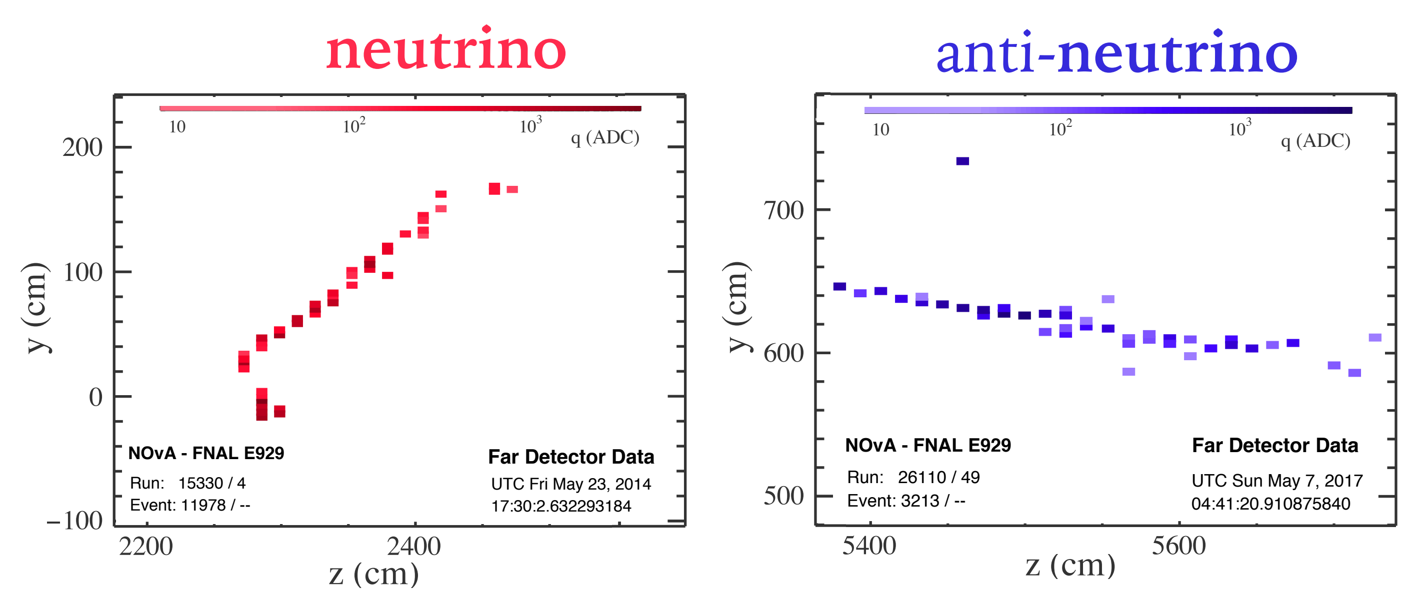
NOvA uses Convolutional Neural Networks to extract features and classify events.



CNNs increased effective exposure by 30% compared to traditional ID methods.

Training on neutrino beam and anti-neutrino beam simulations separately further increased their efficiency for anti-ve signal by 14%

$\bar{\nu}$ **Efficiency Improvement** Training Sample (ID > 0.9)		
$\bar{\nu}_e$ CC Signal	$\bar{\nu}_\mu$ CC Signal	$\bar{\nu}$ NC Signal
14%	6%	10%

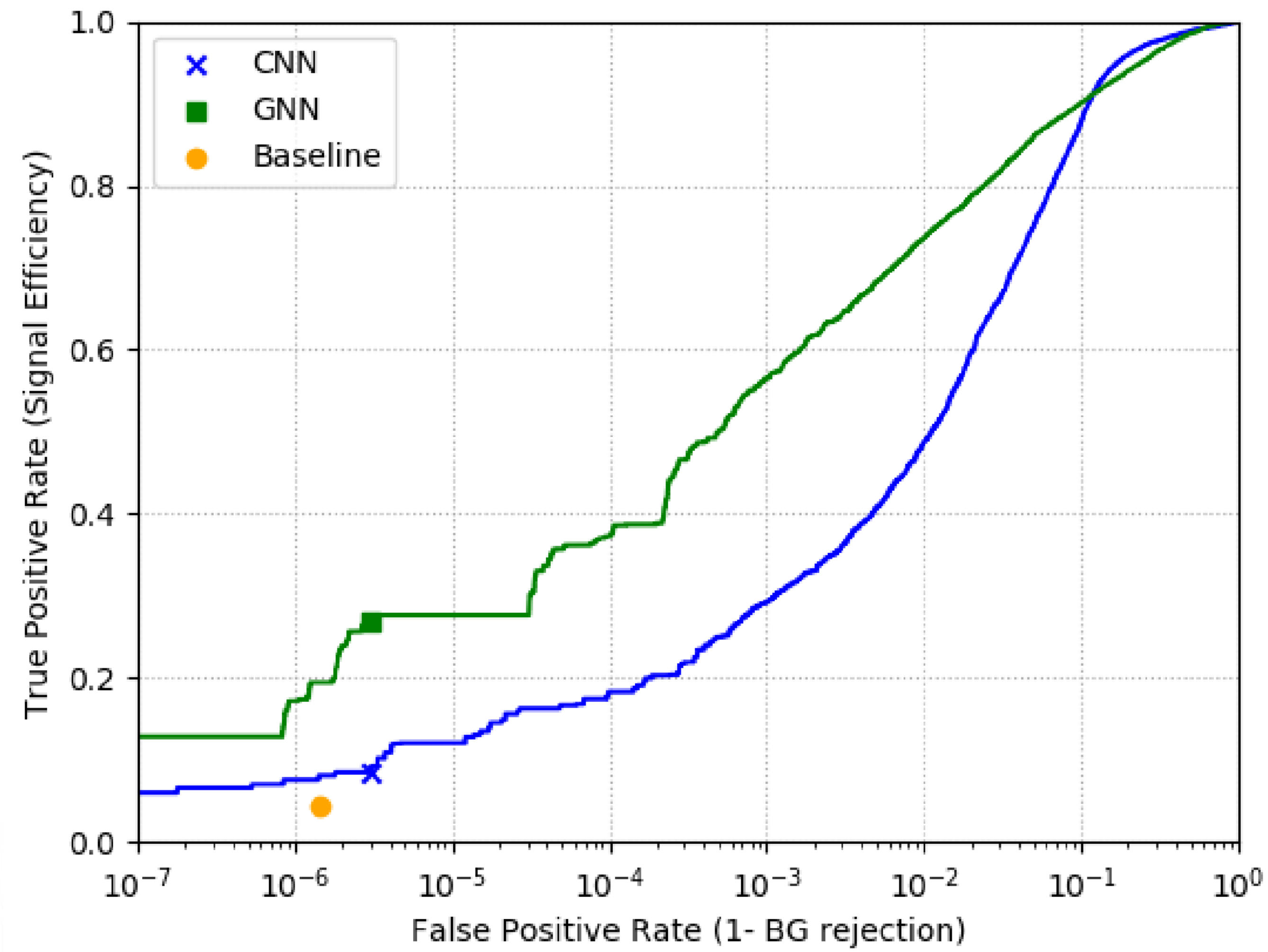


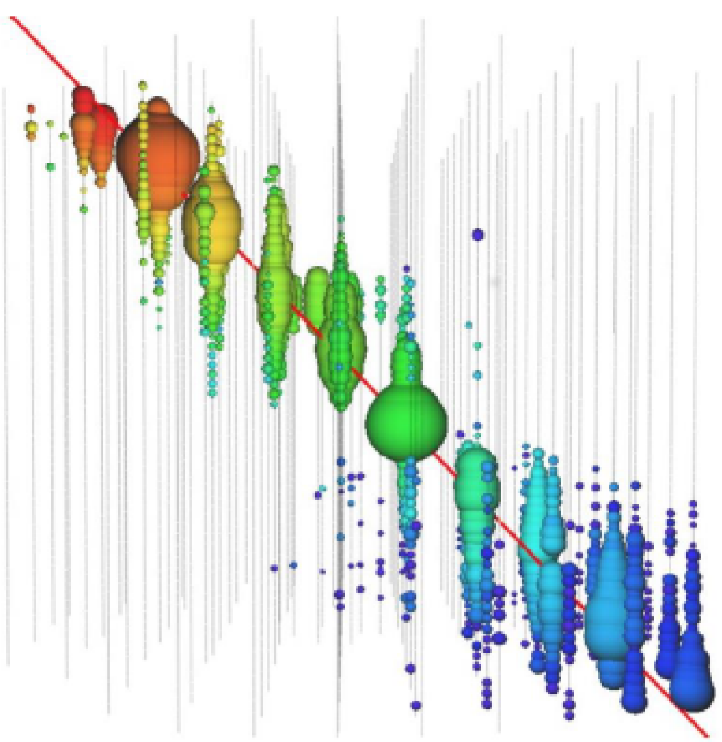
Measurement of Neutrino Oscillations and Improvements from Deep Learning. Fernanda Psihas https://inspirehep.net/literature/1672901

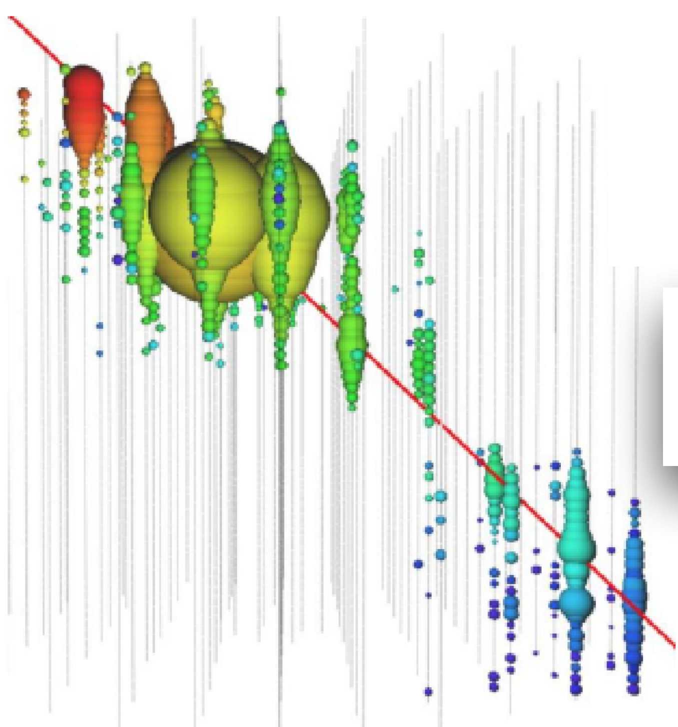
IceCube Graph Nets

3x improvement in signal-to-noise ratio.

6.3x more signal events than traditional physics-based methods.



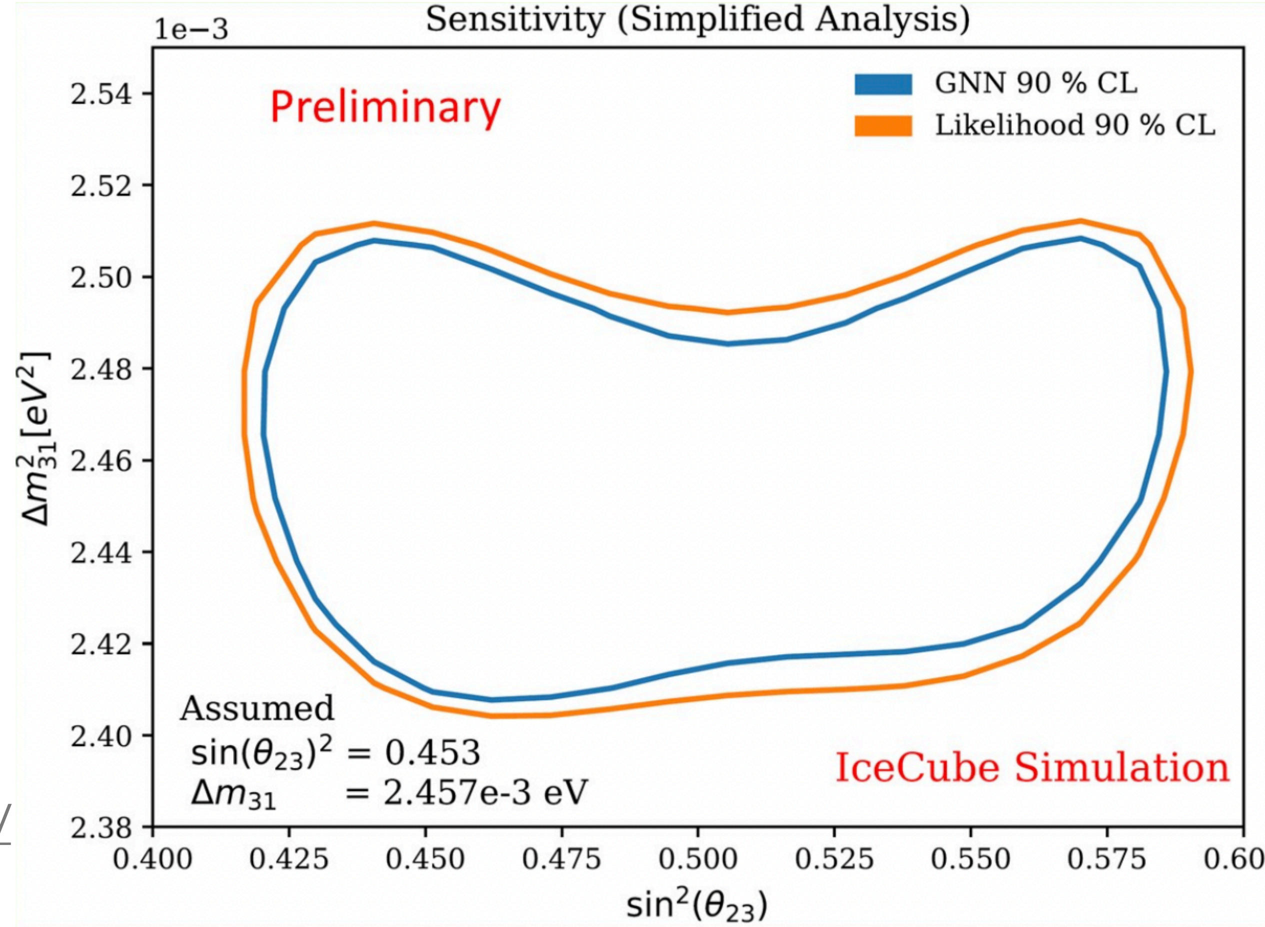




Graph Neural Networks for IceCube Signal Classification

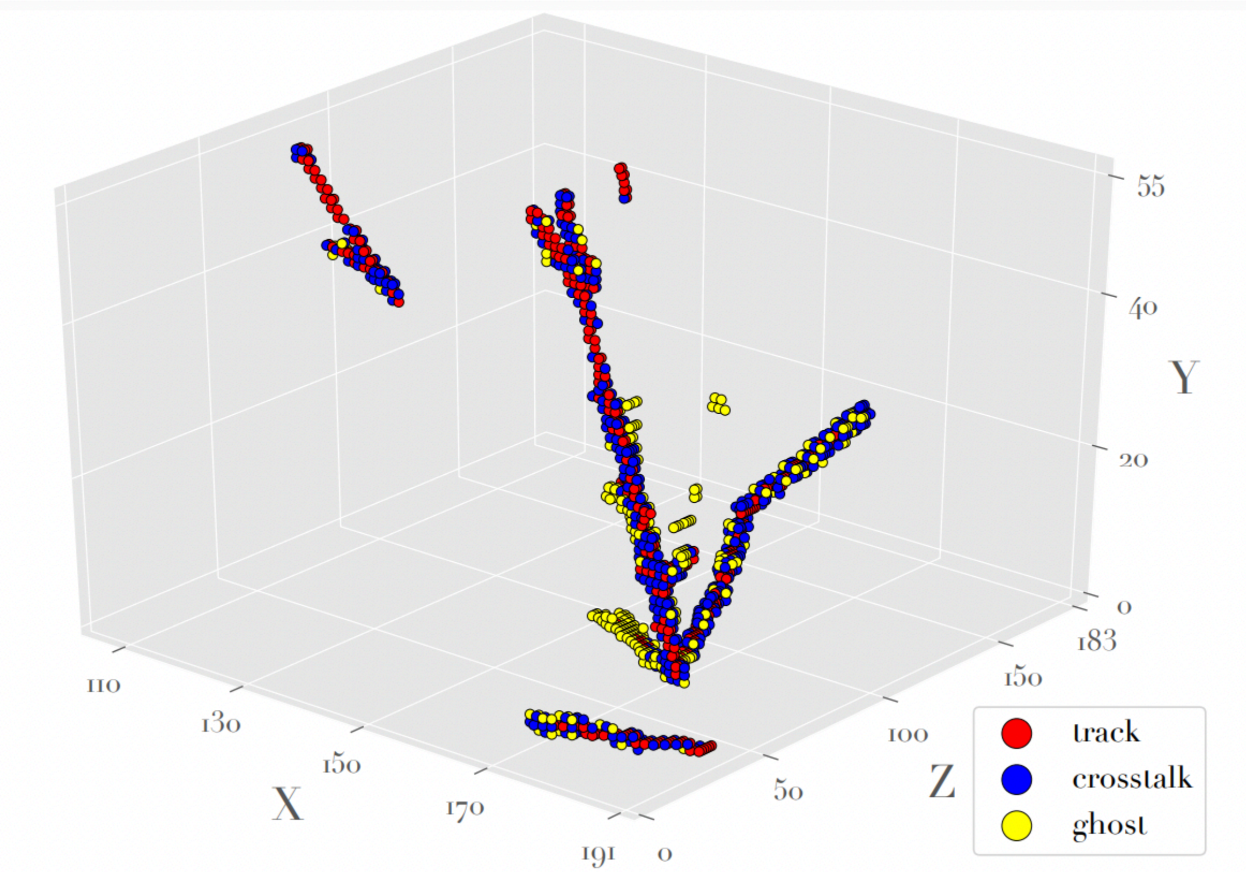
Nicholas Choma, Federico Monti, Lisa Gerhardt, Tomasz Palczewski, Zahra Ronaghi, Prabhat, Wahid Bhimji, Michael M. Bronstein, Spencer R. Klein, Joan Bruna

ICECUBE
SOUTH POLE NEUTRINO OBSERVATORY



Graph NNs are natural for clustering PMT signals. GNN based reco. Yields 20% + resolution in energy & zenith. **Expected sensitivity equivalent to 25% more statistics.**

GNNs Neutrino Event Reconstruction. Neutrino 2022 poster. Rasmus Ørsøe. https://indico.kps.or.kr/event/30/contributions/785/



(a) Prediction: voxels are colored based on the GNN predictions.

T2K is also using GNNs for removing cross-talk & ghost hits from tracks in preparation for The SuperFGD near detector for **improvements with respect to charge cuts.**

GNN			Charge Cut		
	Track	Other		Track	Other
Efficiency	94%	96%	Efficiency	93%	80%
Purity	96%	95%	Purity	80%	91%

Graph neural network for 3D classification of ambiguities and optical crosstalk in scintillator-based neutrino detectors Sa´ul Alonso-Monsalve, et.al. https://arxiv.org/pdf/2009.00688.pdf

THIS SLIDE: NOT YET IMPLEMENTED FOR OSCILLATIONS

ML for neutrino experiments

Reconstruction

Background rejection/ classification

Data quality selections

Data-size reduction

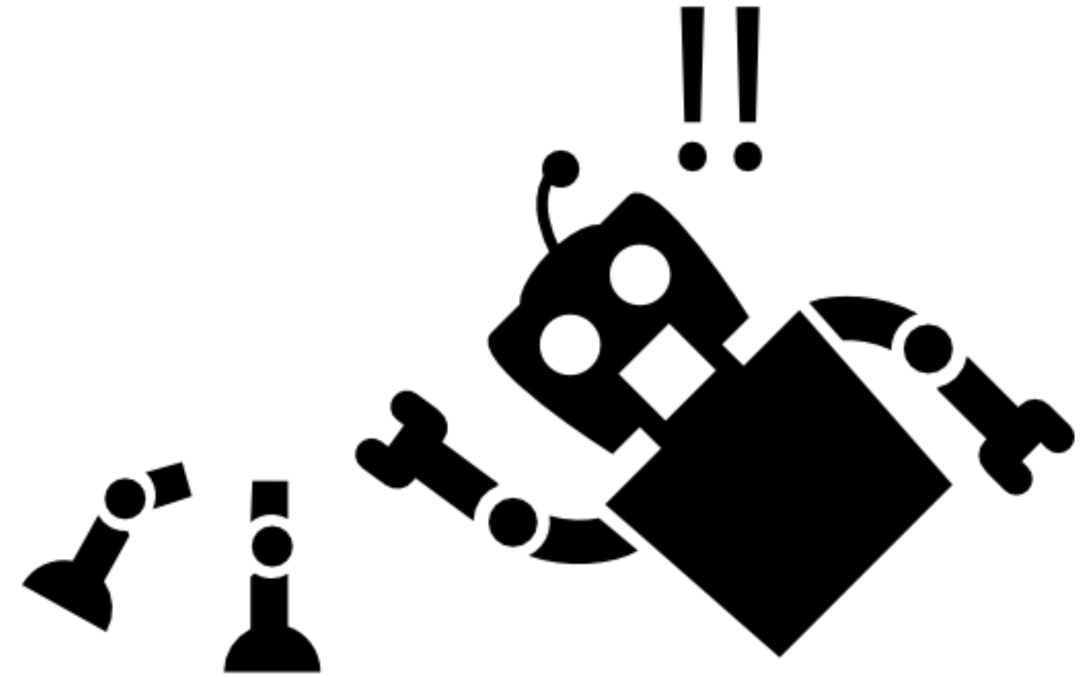
These techniques are improving our current sensitivities and informing experiment design for next-gen experiments, enabling physics analyses that would otherwise be impossible.

Machine learning in the search for new fundamental physics

Georgia Karagiorgi, Gregor Kasieczka, Scott Kravitz, Benjamin Nachman & David Shih

Nature Reviews Physics **4**, 399–412 (2022) | Cite this article

Challenges of applying ML



Model Dependence

Network Interpretability

Bias and Uncertainties

Computational Constrains

World Scientific
Connecting Great Minds

International Journal of Modern Physics A | Vol. 35, No. 33, 2043005 (2020) | Special Issue: "Learning t...

A review on machine learning for neutrino experiments

Fernanda Psihas, Micah Groh, Christopher Tunnell and Karl Warburton

https://doi.org/10.1142/S0217751X20430058

Model Dependence & Uncertainty

(A) **Cow: 0.99**, Pasture: 0.99, Grass: 0.99, No Person: 0.98, Mammal: 0.98

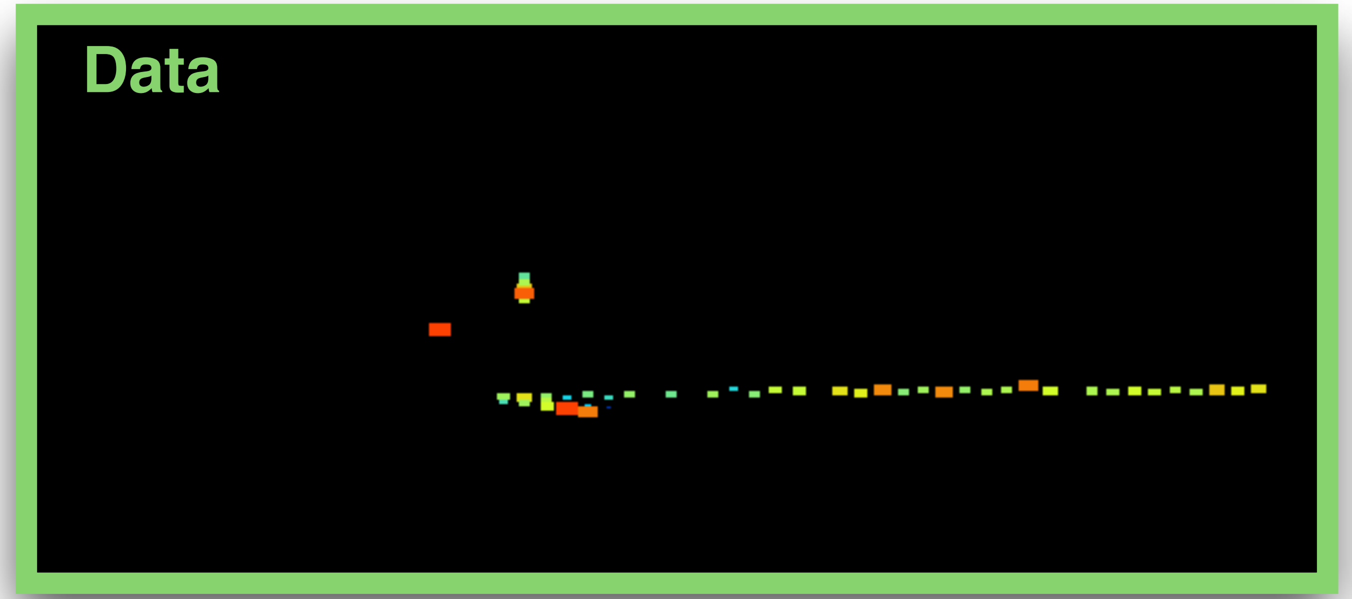
(B) No Person: 0.99, Water: 0.98, Beach: 0.97, Outdoors: 0.97, Seashore: 0.97

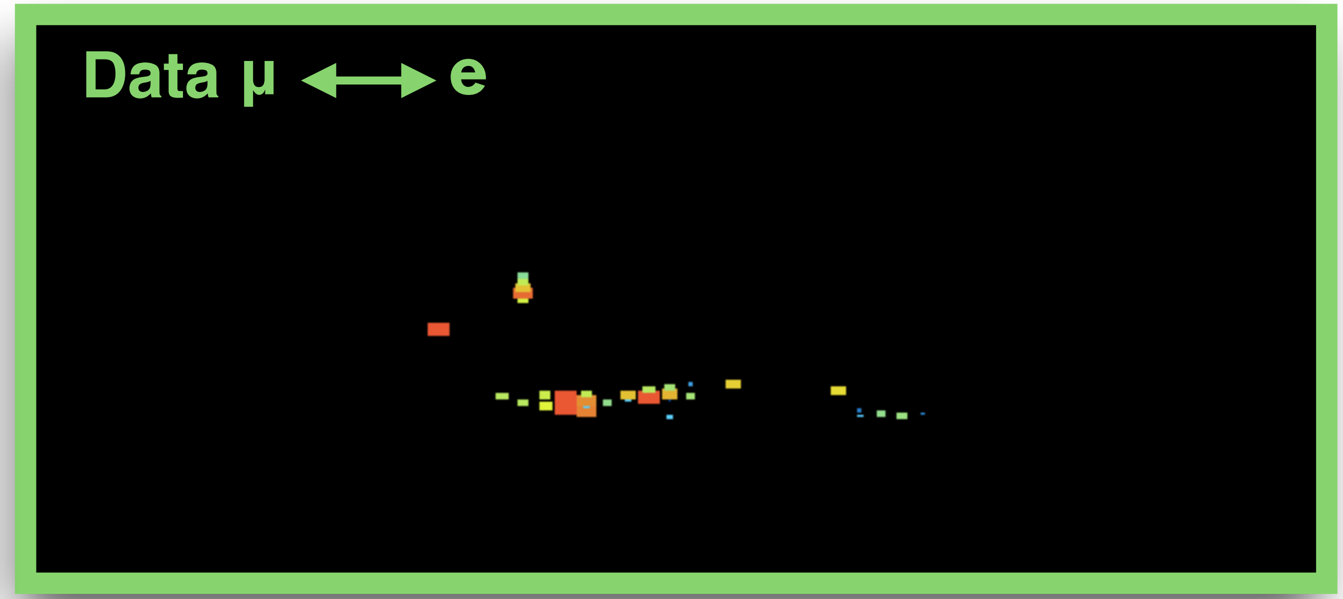
https://arxiv.org/pdf/1807.04975.pdf Recognition in Terra Incognita, October 2018. Conference: 15th European Conference on Computer Vision (ECCV 2018)

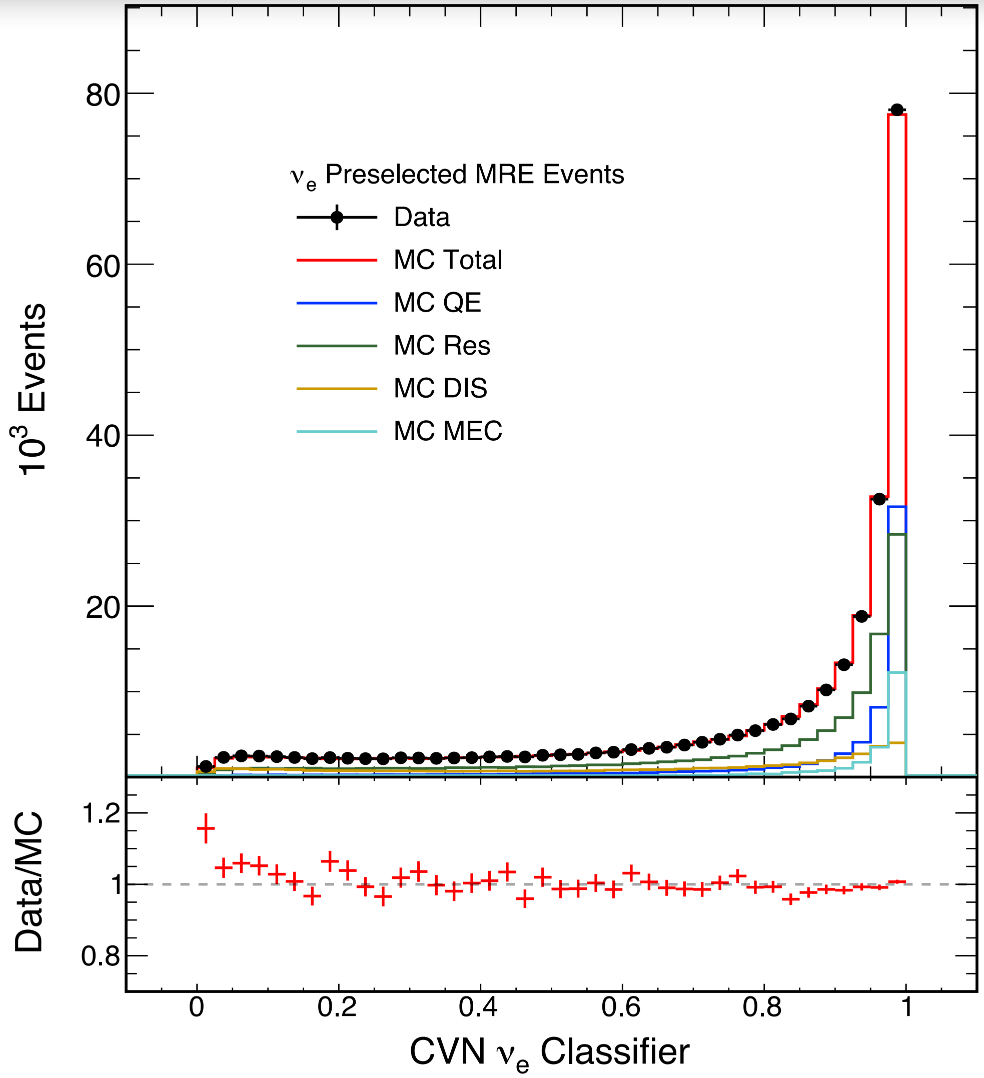
The **composition of the training samples** largely impacts network performance.

Are our algorithms **reproducing the model-based distributions** we train with?

Performance On Real Data







MRE (Muon Removed - Electron):

Select a muon neutrino interaction with traditional ID methods.

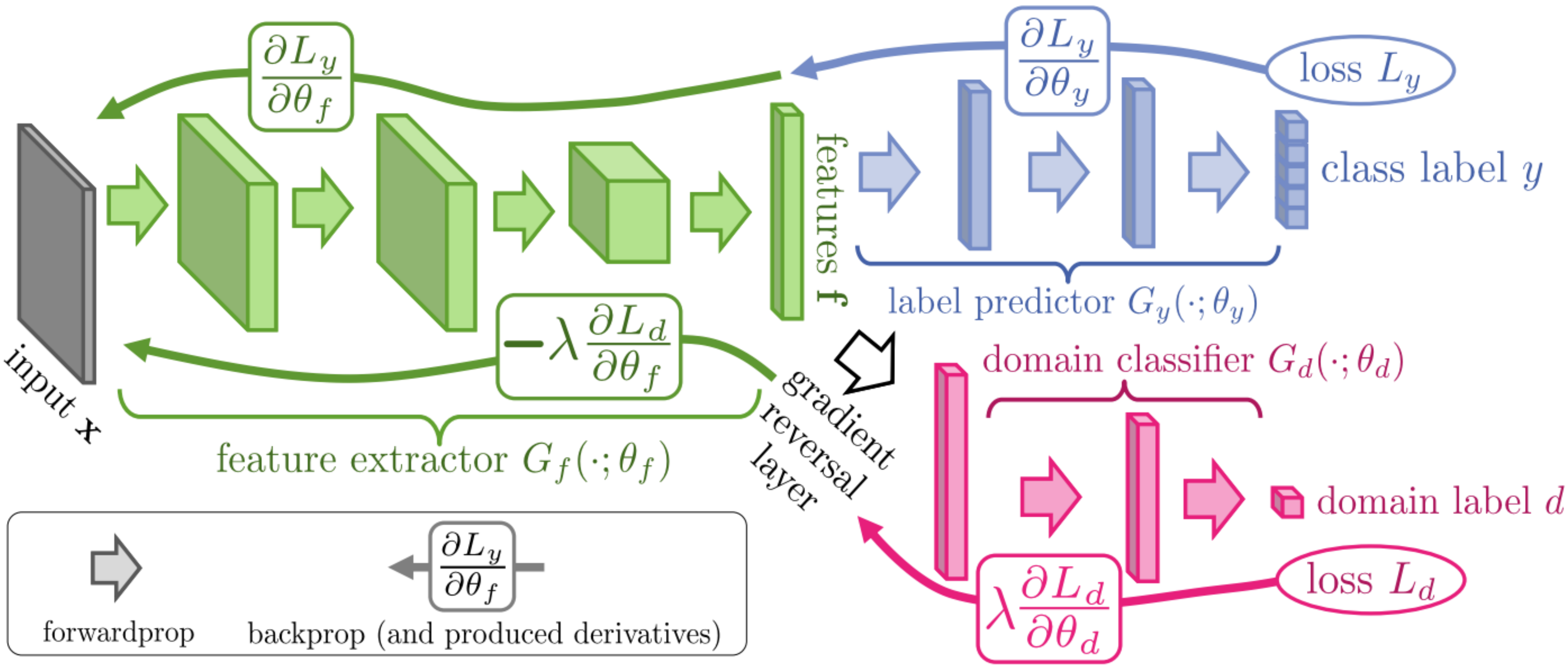
Remove the muon hits and replace them with a single simulated electron of matching momentum.

Data/MC comparisons show less than 1% difference in efficiency.

PID	Sample	Preselection	PID	Efficiency	Efficiency diff %
CVN	Data	262884	188809	0.718222	-0.36%
	MC	277320	199895	0.720809	

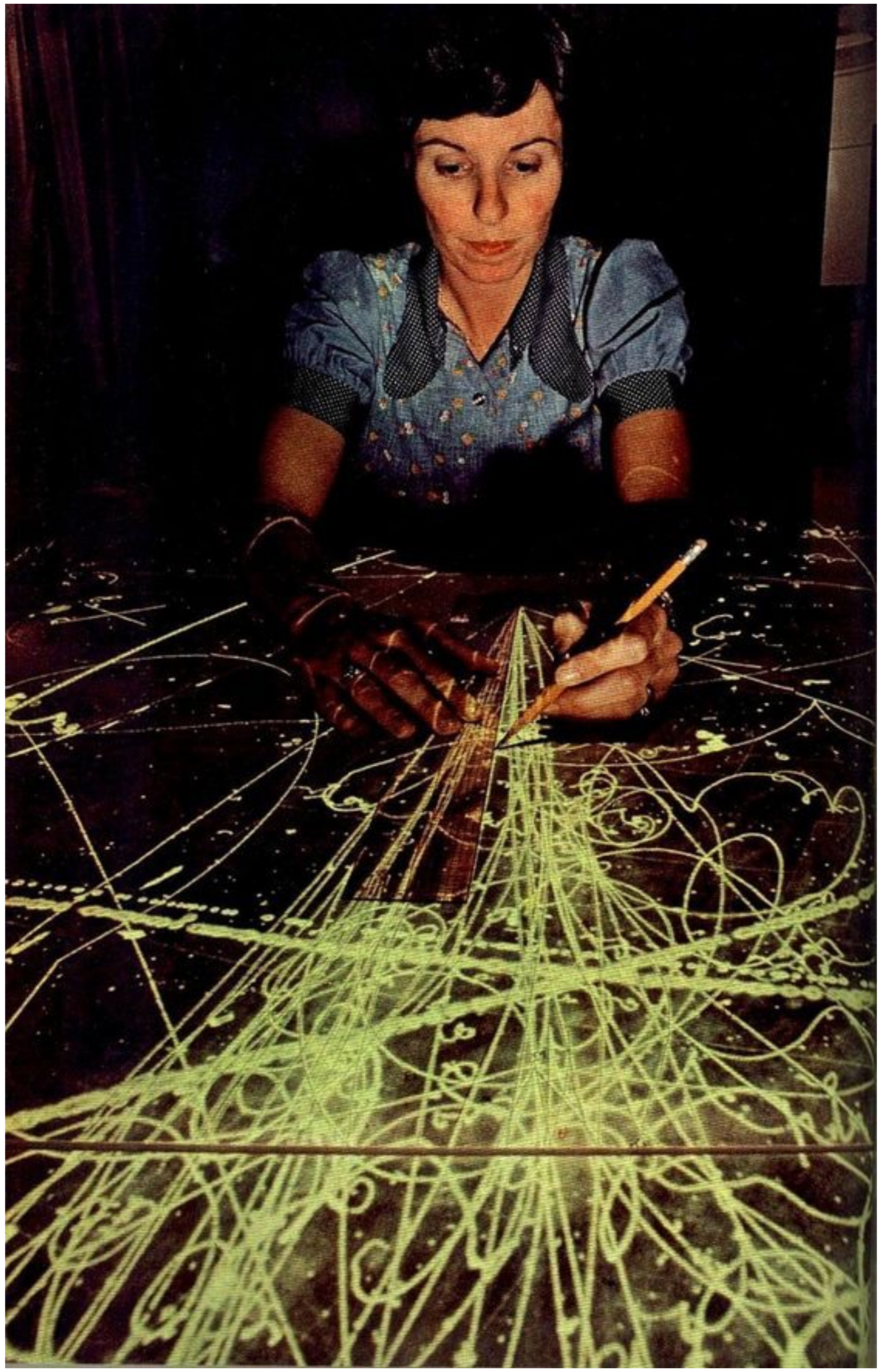
Domain Adversarial Networks

A tool for bias reduction

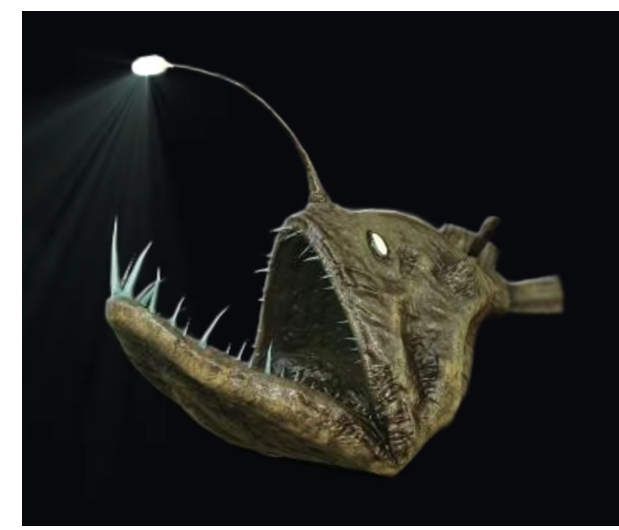


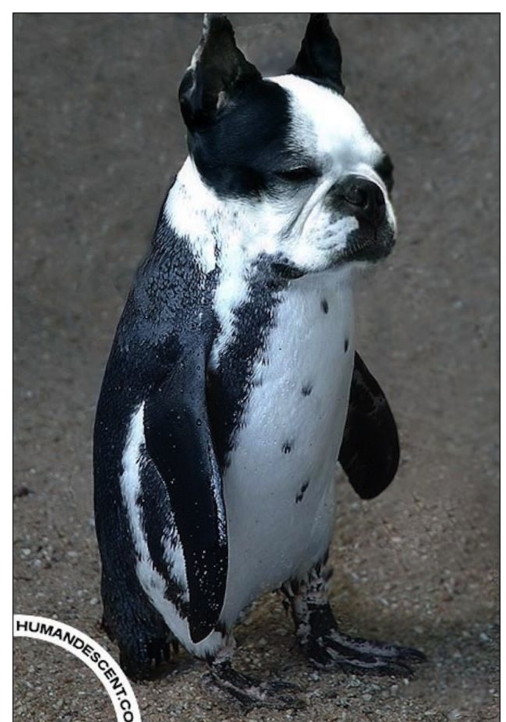
Reducing model bias in a deep learning classifier using domain adversarial neural networks in the MINERvA experiment. G. Perdue, et.al. https://doi.org/10.48550/arXiv.1808.0833

Maintaining sensitivity to new physics

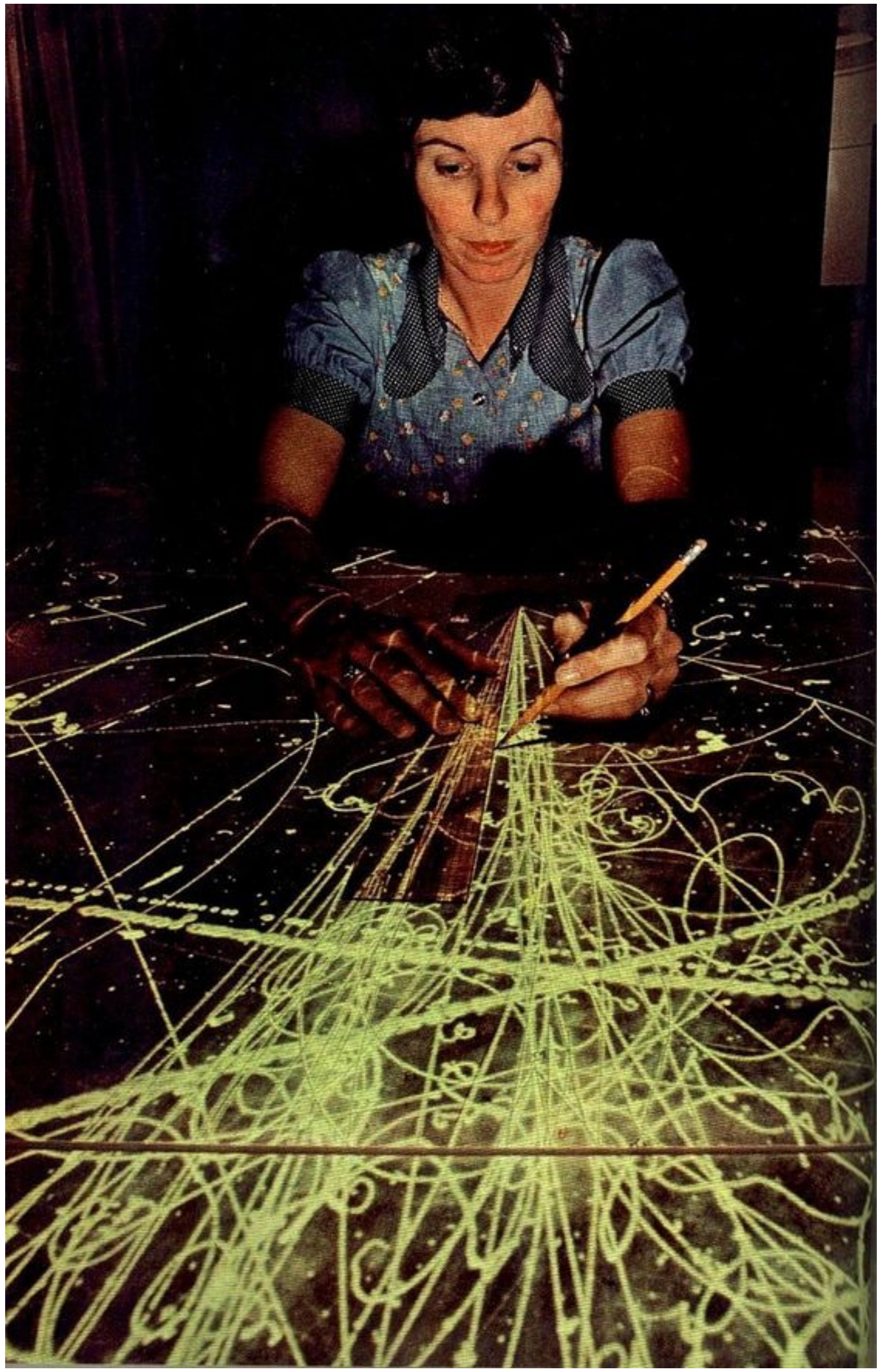


She would know this is **not** what animals look like in nature.

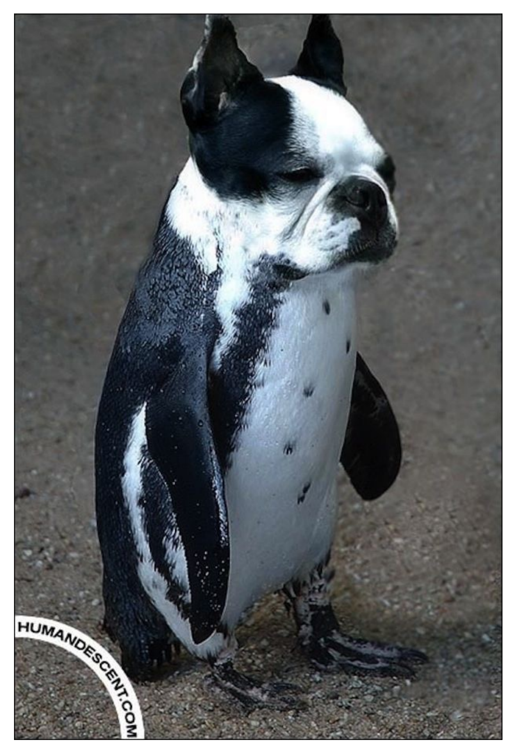


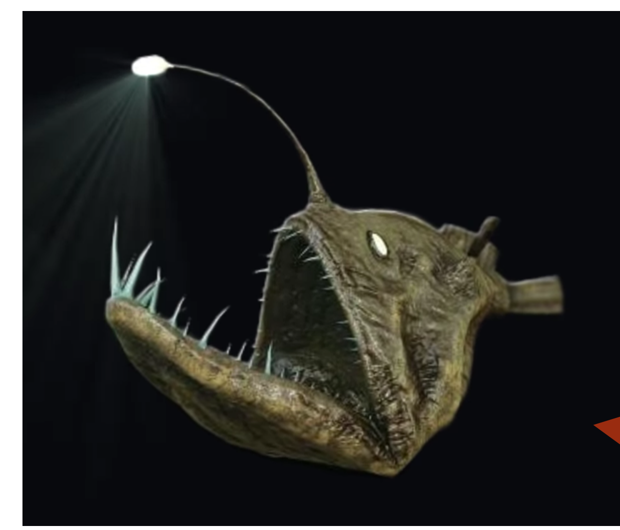


Maintaining sensitivity to new physics



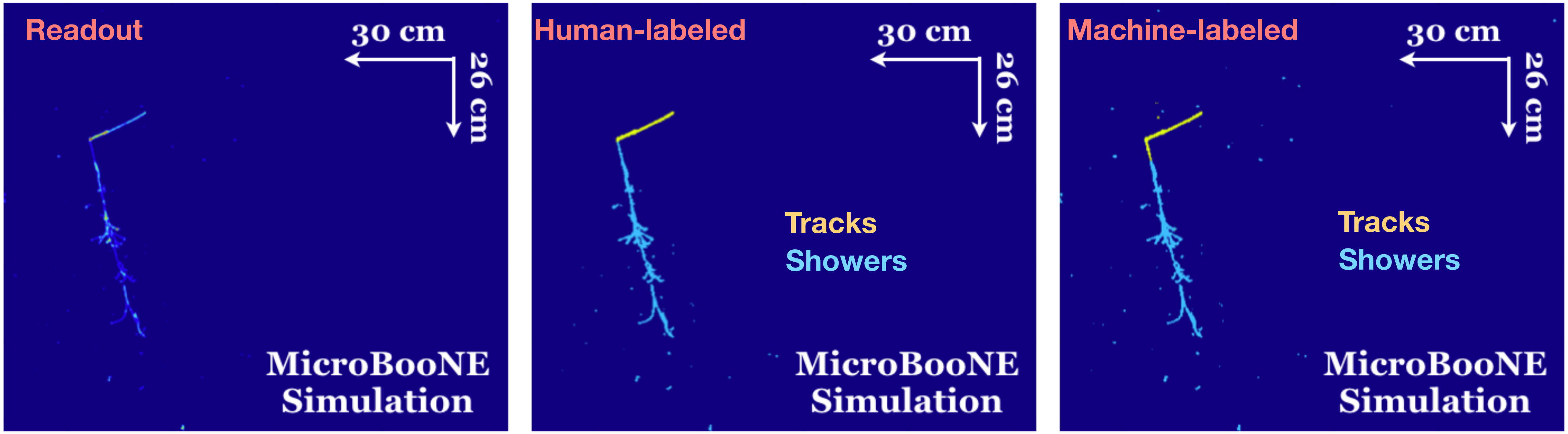
She would know this is **not** what animals look like in nature.

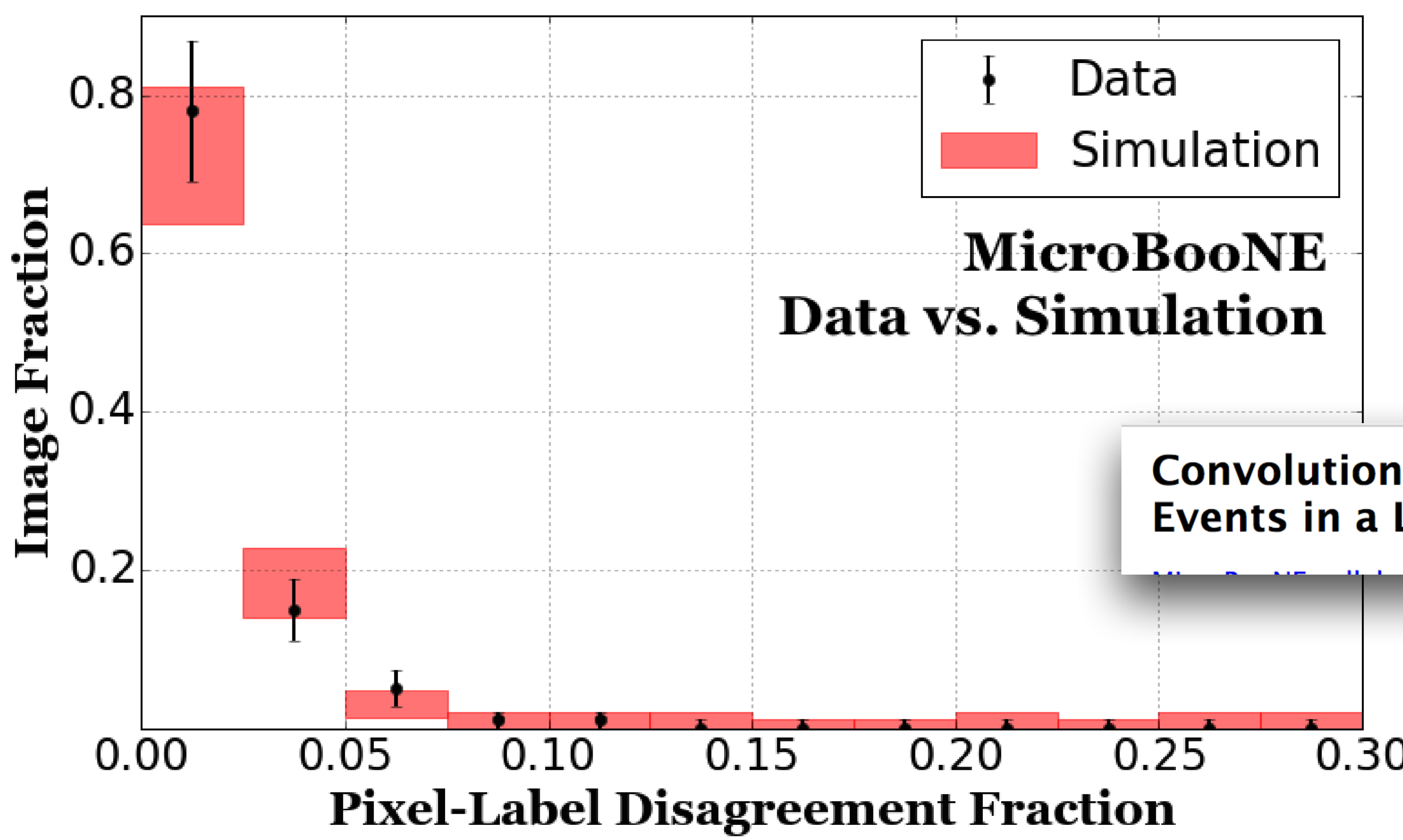




This is a real fish!
Training can enhance or suppress sensitivity to the unexpected!

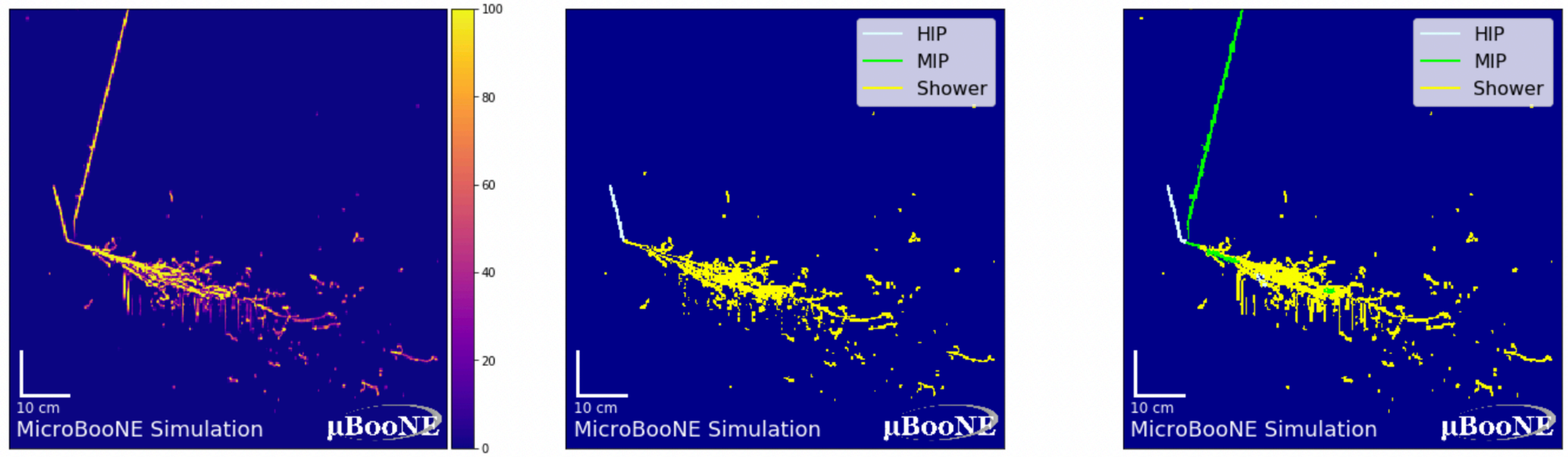
Human-labeled dataset

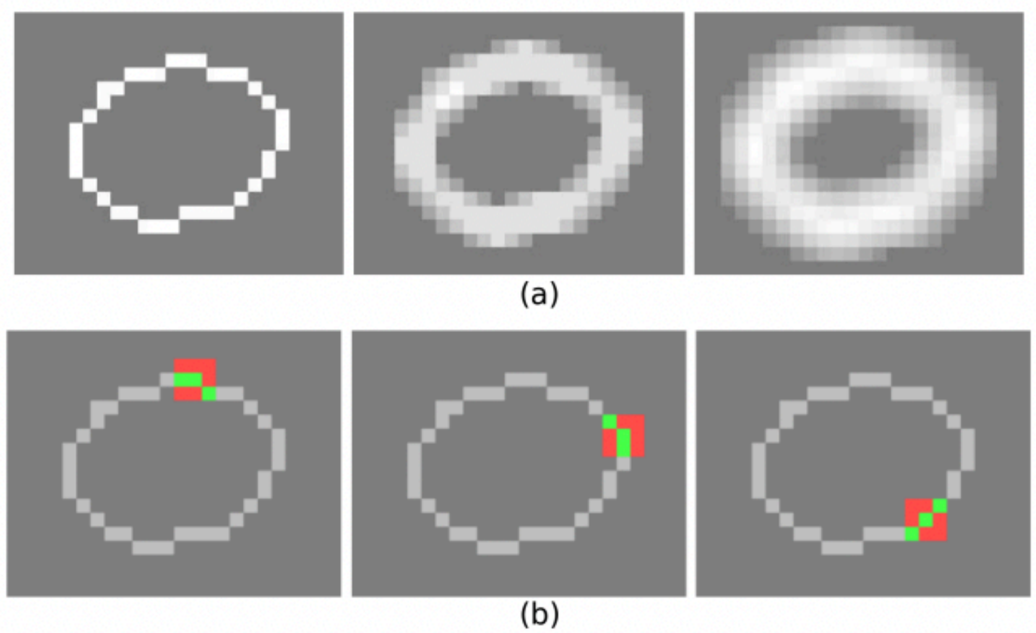




Convolutional Neural Networks Applied to Neutrino Events in a Liquid Argon Time Projection Chamber

Sparcity





Constraining only pixels which were non-zero in the input layer to be activated in hidden layers

Sparse convolutions cut inference time from ≈ 5 s to ≈ 0.5 s as well as the memory usage from ≈ 5 GB to ≈ 1 GB

	Test	Intrinsic ν_e	Full-BNB
Track	0.992	0.992	0.998
Shower	0.996	0.859	0.823

Semantic segmentation with a sparse convolutional neural network for event reconstruction in MicroBooNE

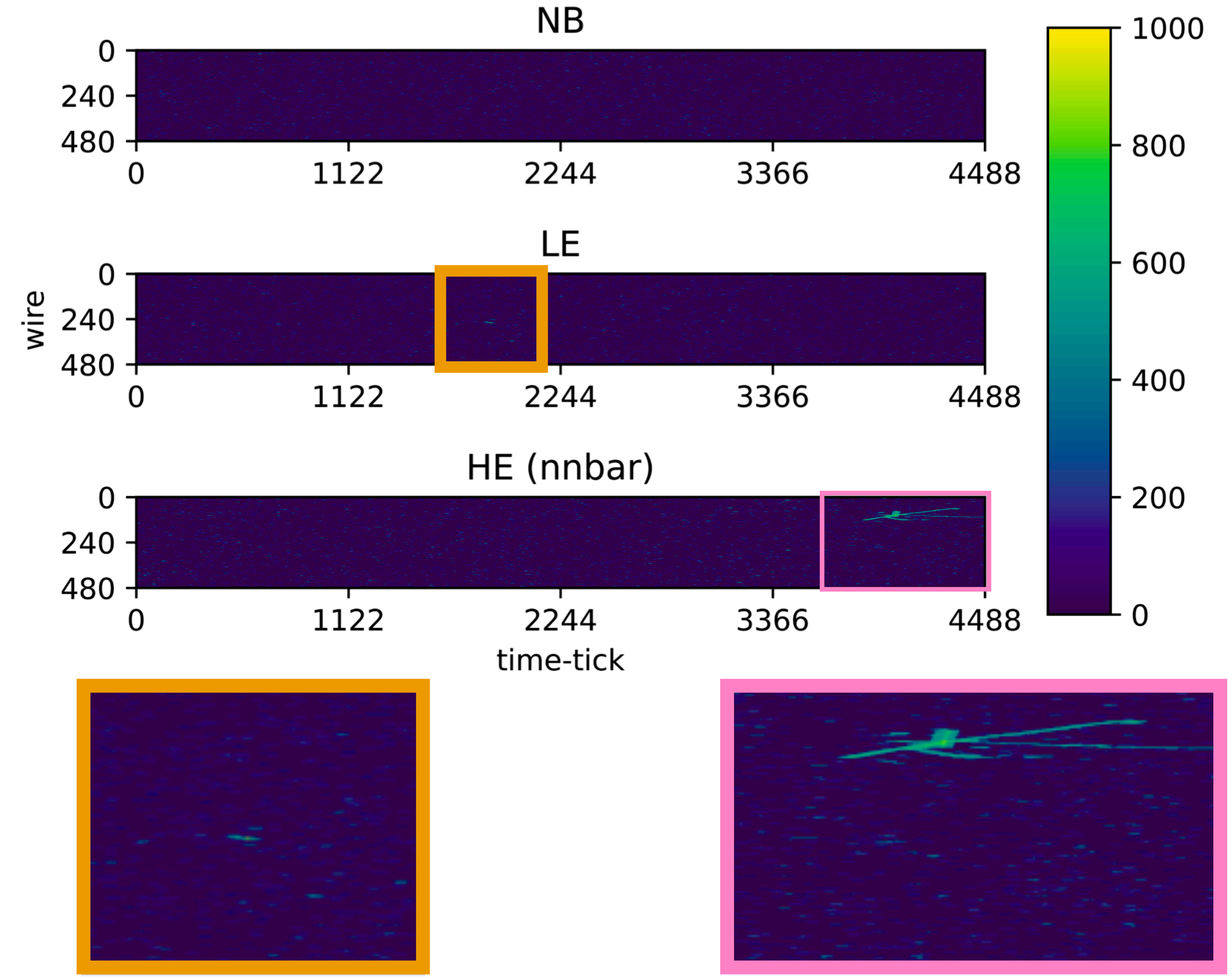
P. Abratenko *et al.* (The MicroBooNE Collaboration)
Phys. Rev. D **103**, 052012 – Published 26 March 2021

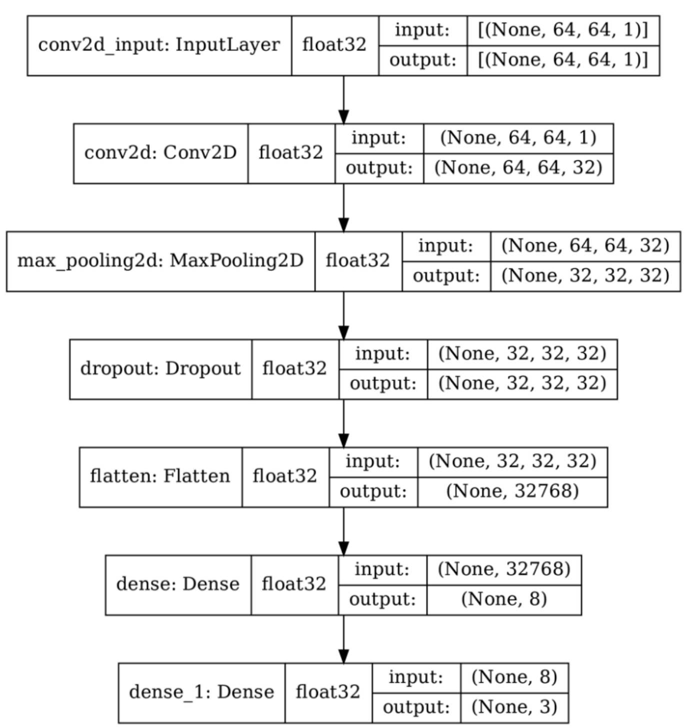
Fast ML

arXiv:2201.05638

The raw data rate expected for each DUNE module is 1.175 TB/s

Live data-processing can enable physics analyses for multiple signals across a large energy range.

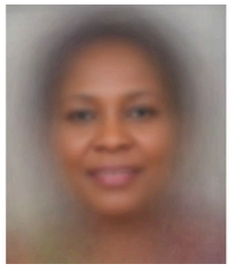
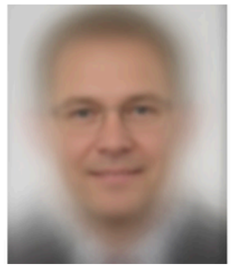
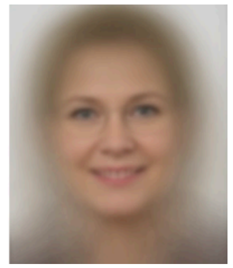




Simple CNNs for pre-processing yield target signal selection efficiencies that meet the DUNE FD physics requirements, **while also providing the needed 10^4 factor of overall data rate reduction**

Addressing ML challenges to neutrino experiments

… from the research

Bias

Find bias AND reduce bias AND quantify bias

There is NO "unbiased" training sample! (Bias to flat is still bias)


Model Dependence

There is no "model independent" sample! (Non-physical models are still models)

Propagate uncertainties through both model training AND model usage

Design algorithms that minimize across known systematic uncertainties.


Robust training

Compare algorithm performance in real data.

Design labeled-data training sets (test beams, known sources, etc.)

Design further tests of Data-MC robustness.


Sensitivity to new physics

Unsupervised learning to identify missing physics & unexpected learned features

Design tools for interpretability: test extracted features, principal component, etc.

Addressing ML challenges to neutrino experiments

… from the community

ML is part of the particle physics toolkit. With increasing complexity, **increasing scrutiny is required.** Teach the use and interpretation of ML as an essential skill of particle physics research.

Develop techniques for robustness metrics and systematic bias assessment that can become the standard for machine learning applications in particle physics.

Contribute to AI research by developing solutions to the bias and uncertainty questions of the industry broadly.

Particle Physics is Uniquely Applicable to A.I.

DETECTOR DATA

is **information-dense** & **un-labeled**
Many times includes **space correlations**/topology.

SIMULATIONS

Produced at **large-scale** and reproducible from physics principles.
Tunable to better/worse match **real data.**

MEASUREMENTS

Analyses that produce **high precision** measurements
Focus on **uncertainty quantification** and bias assessments.

Conclusions

Machine learning techniques have and will continue to improve our experimental sensitivities in neutrino physics.

Developing expertise as a community will enable us to **face the challenges** introduced by increasing algorithm complexity

Particle physics experiments are uniquely equipped to **solve the bias and uncertainty problem in ML** for next-gen oscillation experiments and the broader community.

Thank you.