UNIVERSITY OF CAMBRIDGE

PANDORA

AI/ML and Neutrino Physics

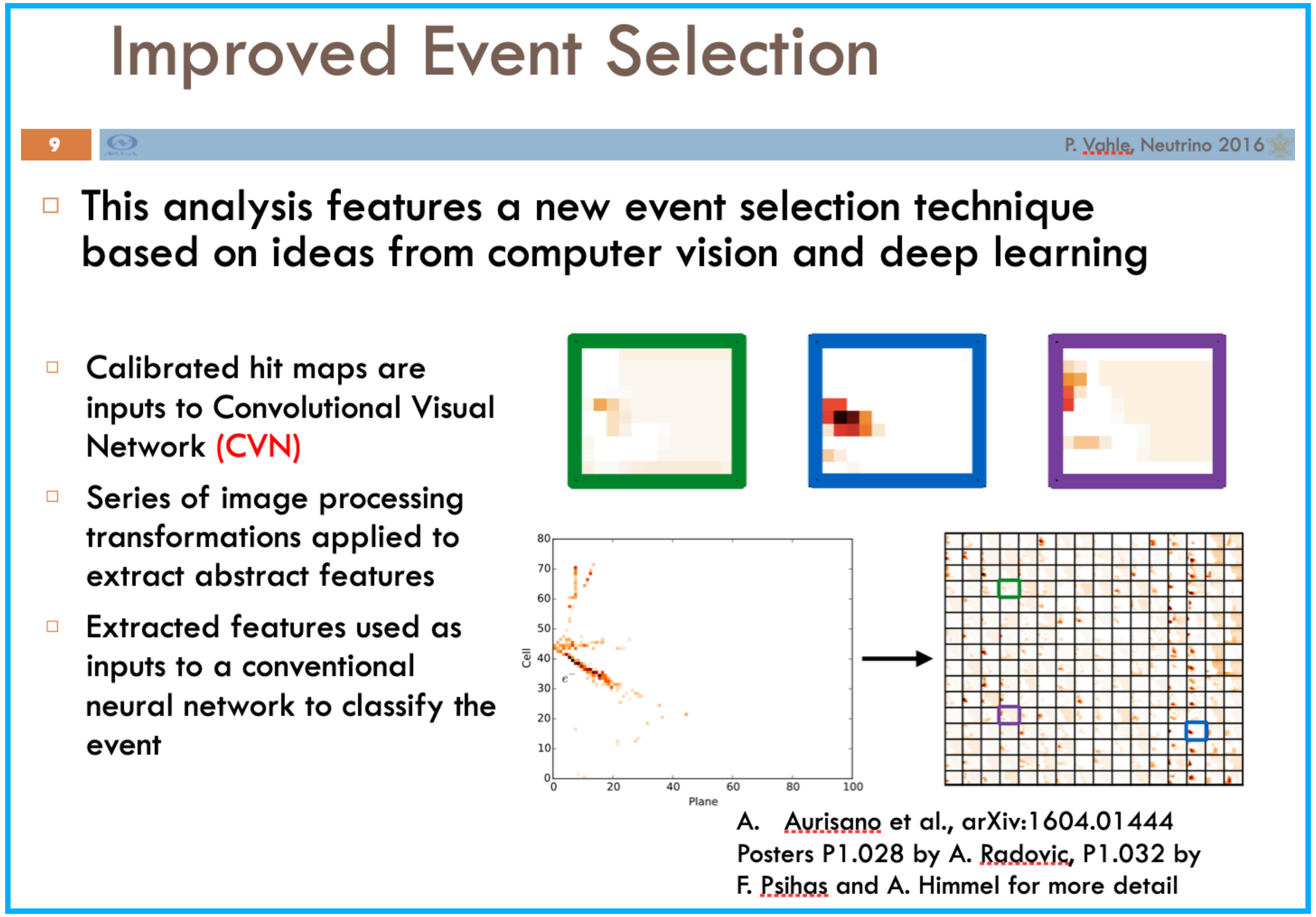
Leigh Whitehead

University of Cambridge

Neutrino XXXII, UC Irvine, June 2026

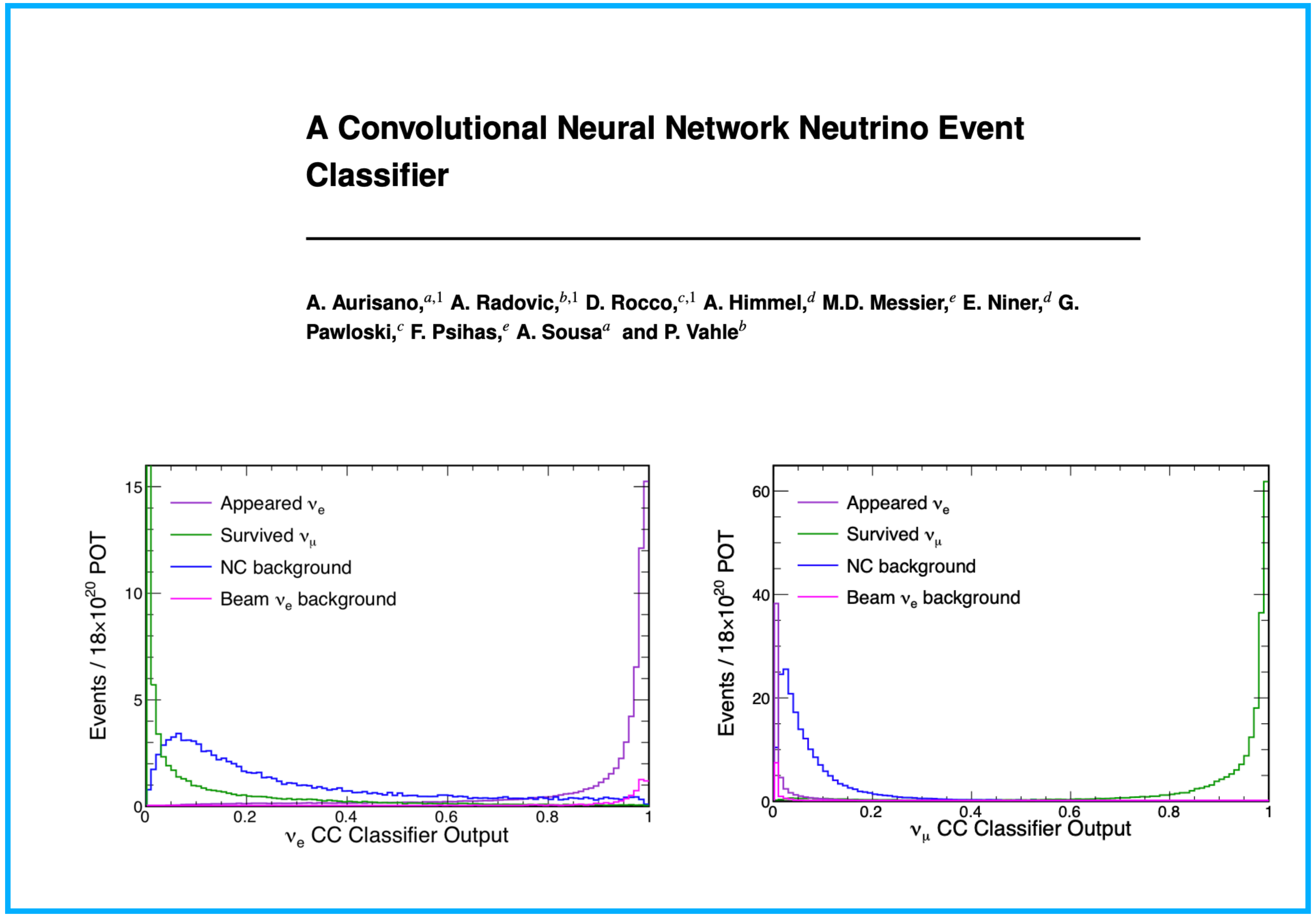
Introduction

- The only previous Neutrino conference I've been was 2016
- First presentation of Deep Learning
- NOvA introduced a CNN to identify the neutrino flavour
  - Improvement equivalent to 30% more exposure
  - Thanks to Trish for still having her original slides!
- Seemed like a game changer to me



Introduction

- The only previous Neutrino conference I've been was 2016
- First presentation of Deep Learning
- NOvA introduced a CNN to identify the neutrino flavour
  - Improvement equivalent to 30% more exposure
  - Thanks to Trish for still having her original slides!
- Seemed like a game changer to me

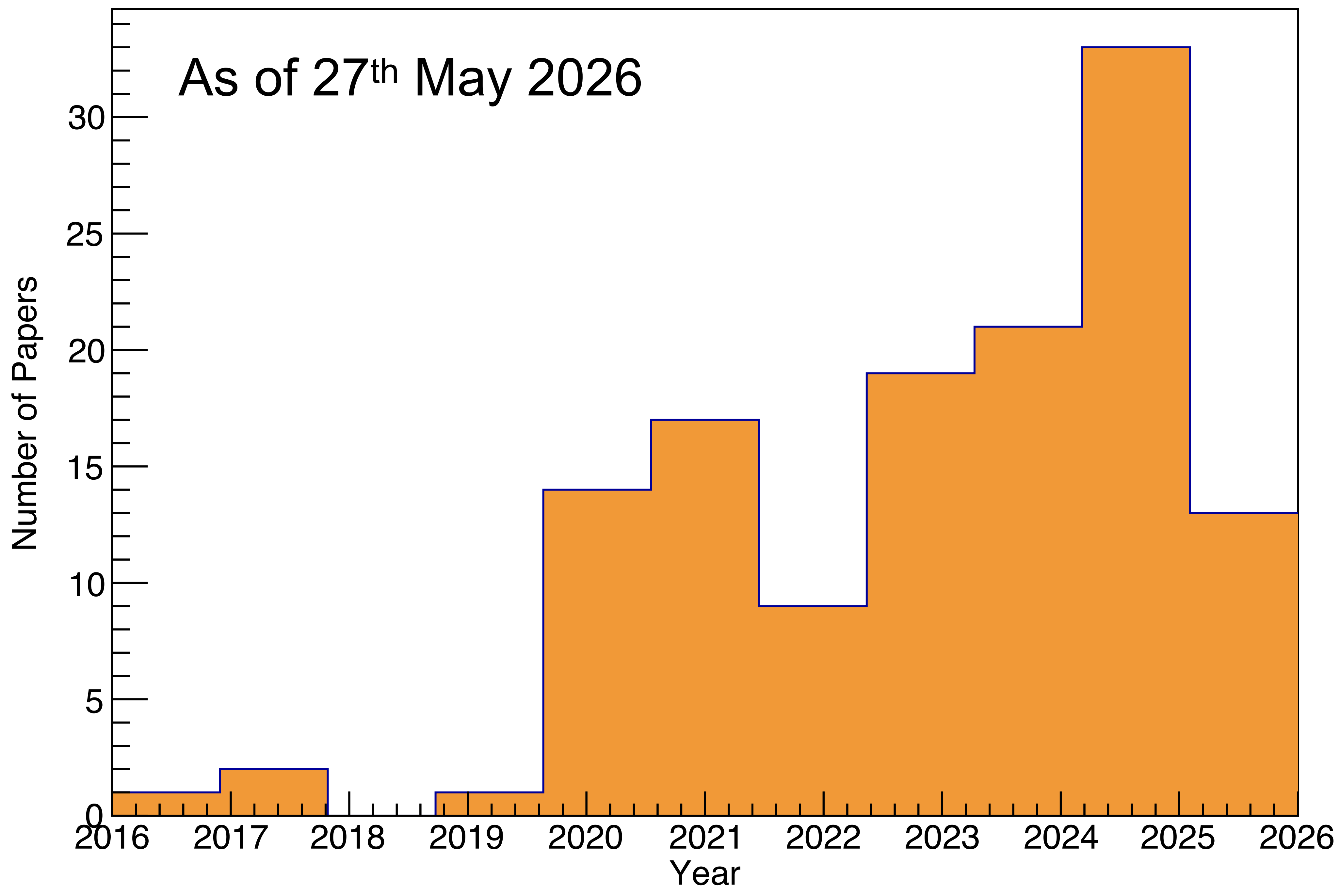

A Convolutional Neural Network Neutrino Event Classifier

A. Aurisano,[a,1] A. Radovic,[b,1] D. Rocco,[c,1] A. Himmel,[d] M.D. Messier,[e] E. Niner,[d] G. Pawloski,[c] F. Psihas,[e] A. Sousa[a] and P. Vahle[b]

A. Aurisano, et al., *A convolutional neural network neutrino event classifier,* Journal of Instrumentation 11 (2016) 09, P09001

AI/ML Adoption in Neutrino Physics

- To set the scene, I searched the HEP ML Living Review[1] for neutrino-related articles (includes both published and pre-prints)
  - Published articles have the year set from the journal publication date



[1] https://iml-wg.github.io/HEPML-LivingReview/

Overview

- I'm going to (try to) give a flavour of various modern techniques and applications over the last 10 years*
  - Convolutional Neural Networks, Graph Neural Networks, Transformers
  - Simulation and Generative models
- I'll also try to give a sense of what comes next
  - Including the US Genesis Mission
  - General thoughts about the future

*unfortunately I can't discuss all of the great work that has been done over the last 10 years. Any omissions are my fault and make no insinuation as to the quality of the work

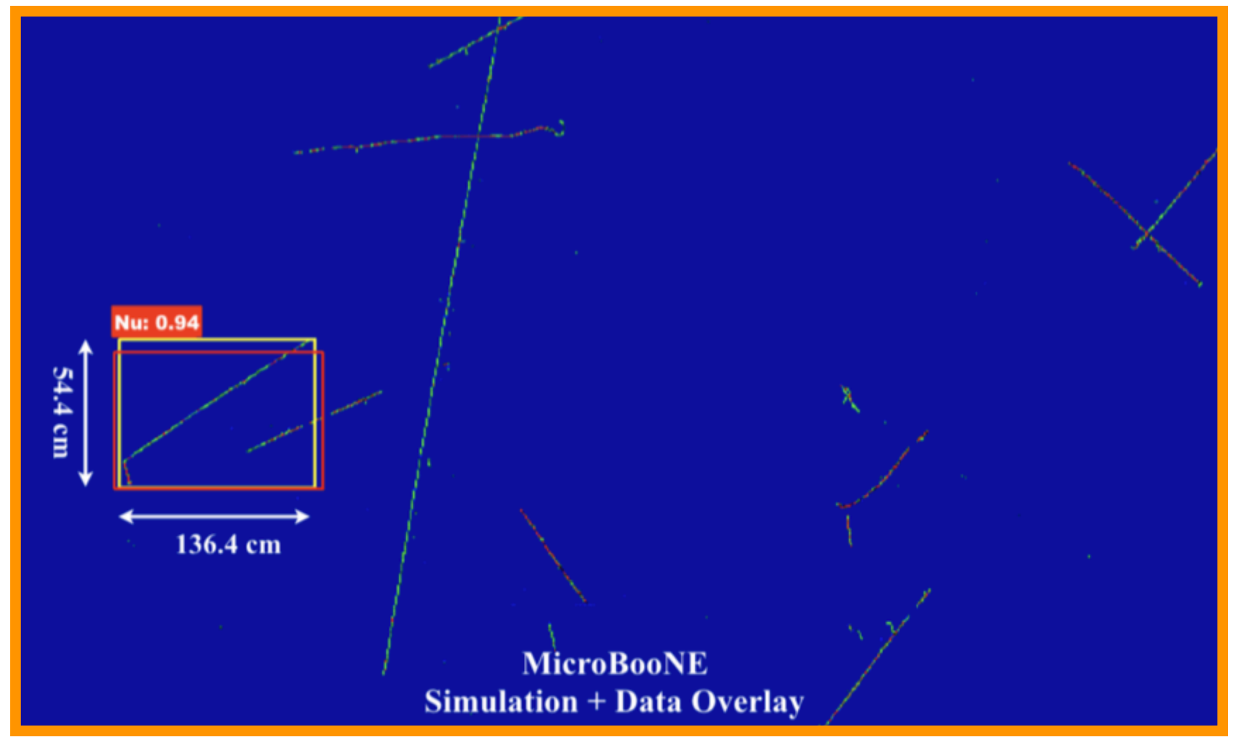
Also lots of great new work presented at NPML 2026 last week. See Kazu's talk after the coffee break for all the details!

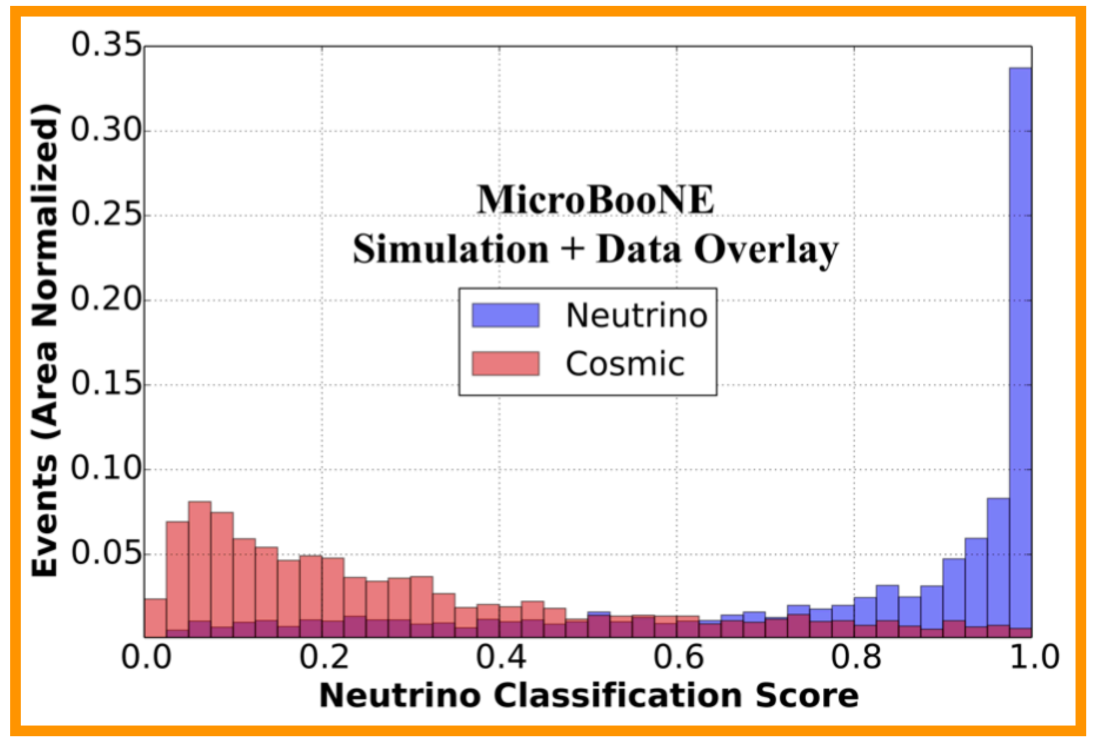
Convolutional Neural Networks

MicroBooNE

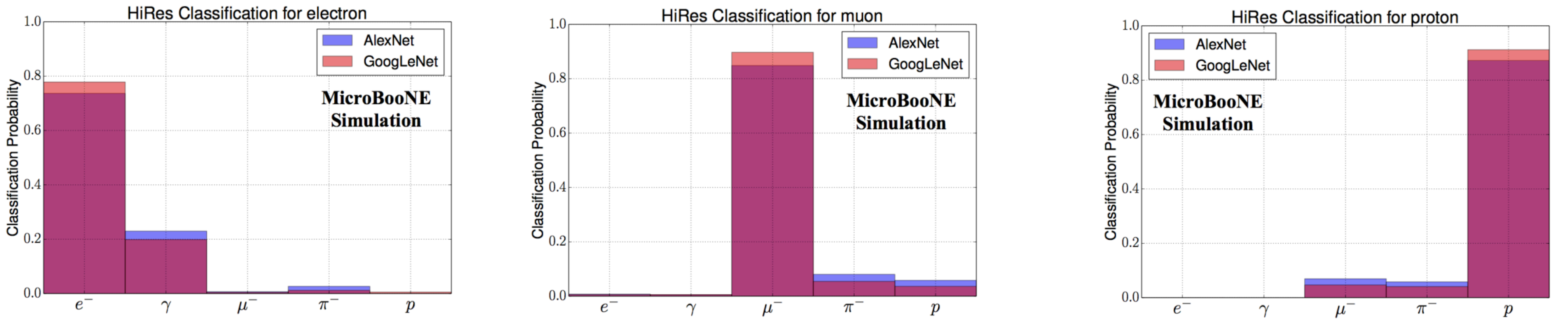
- MicroBooNE were hot on the heels of NOvA with their own CNN development
  - First use of Deep Learning in LArTPCs
  - There were a number of use cases

- CNN used to find the neutrino-like events and classify them



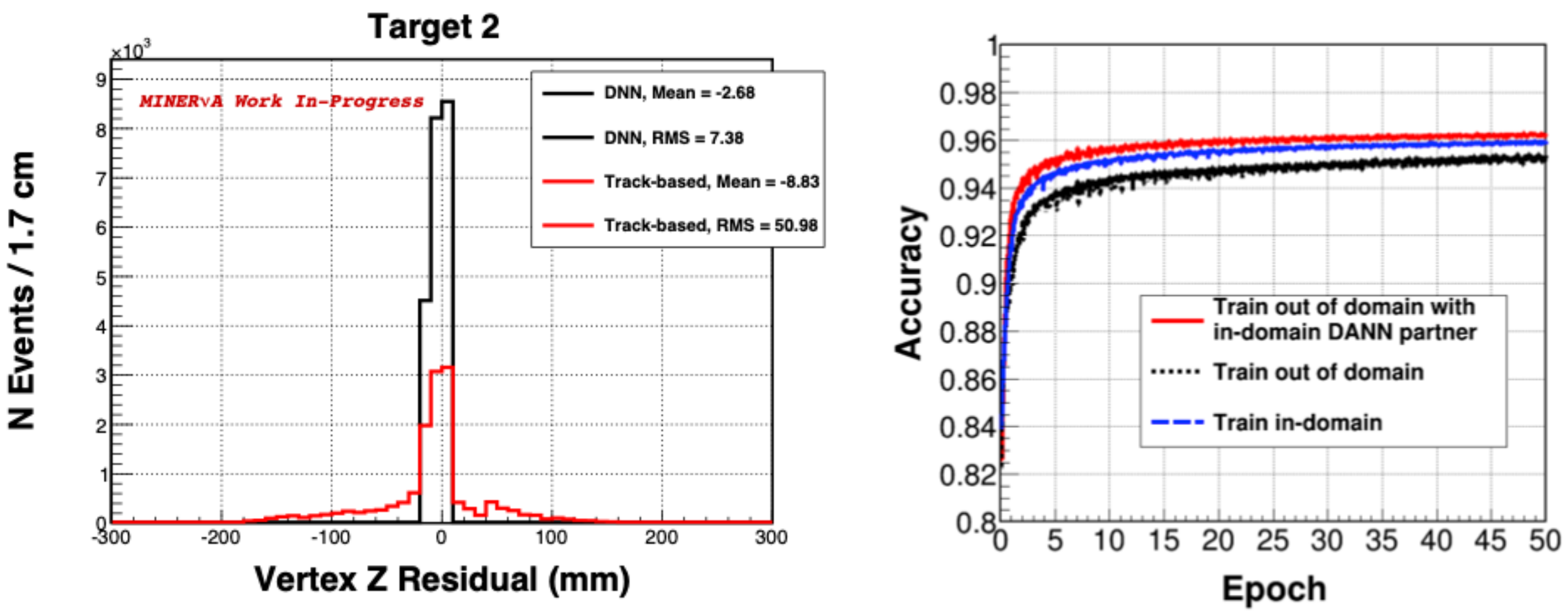


- Particle identification

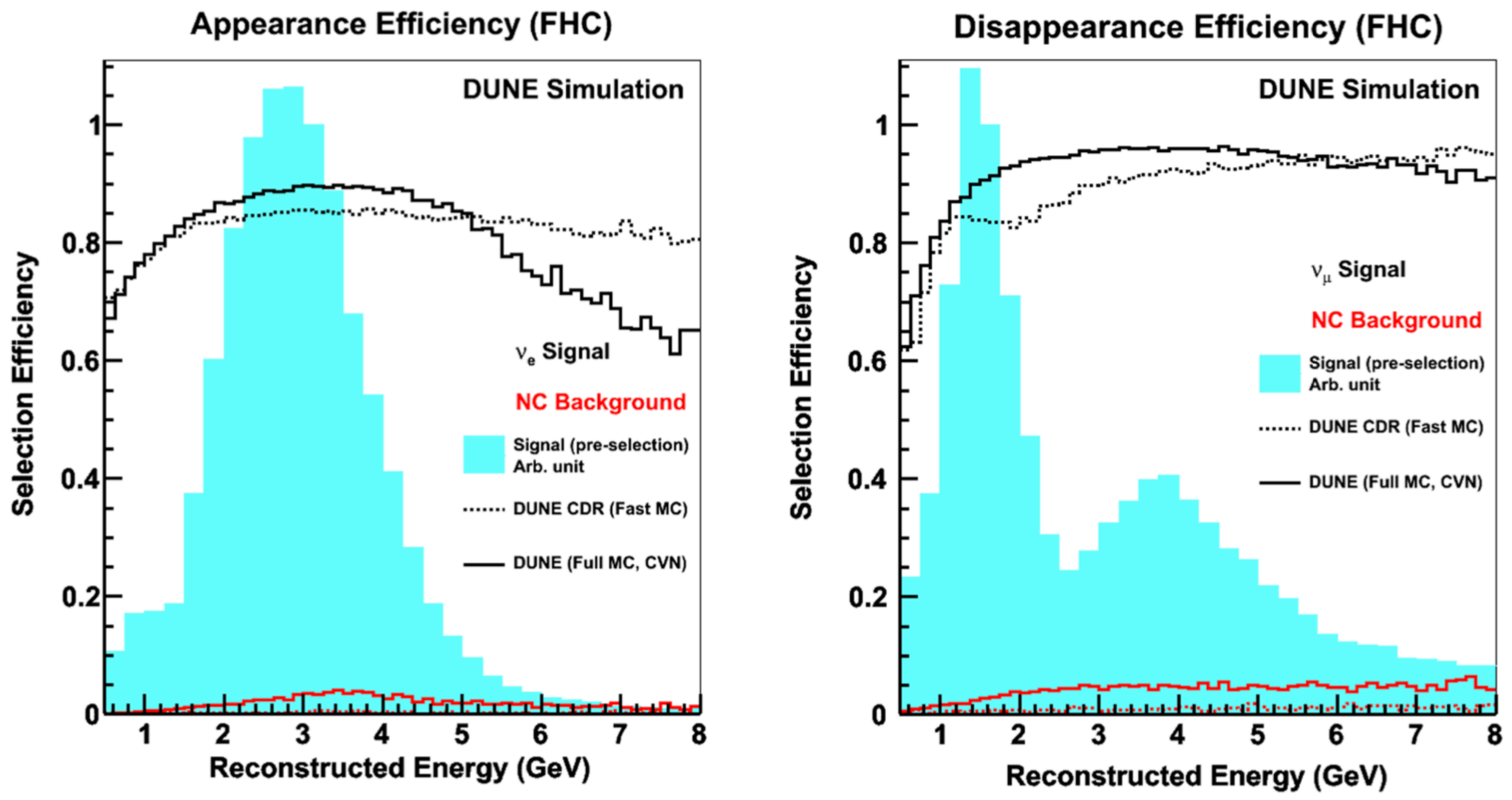


MicroBooNE Collaboration, *Convolutional Neural Networks Applied to Neutrino Events in a Liquid Argon Time Projection Chamber*, JINST 12 (2017) 03, P03011

More CNN use-cases I

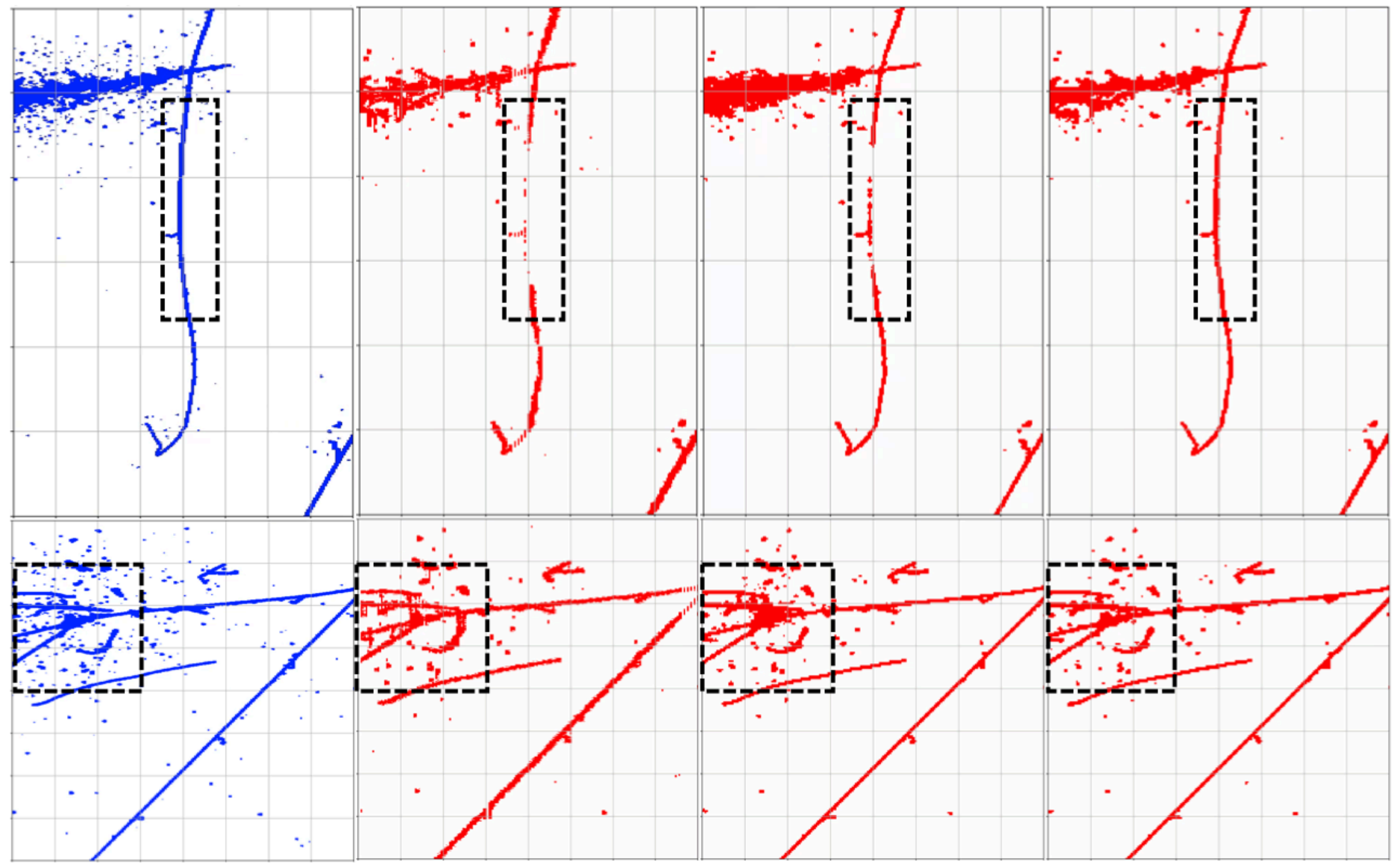


MINERvA Collaboration, *Reducing model bias in a deep learning classifier using domain adversarial neural networks in the MINERvA experiment*, JINST **13**, P11020 (2018)



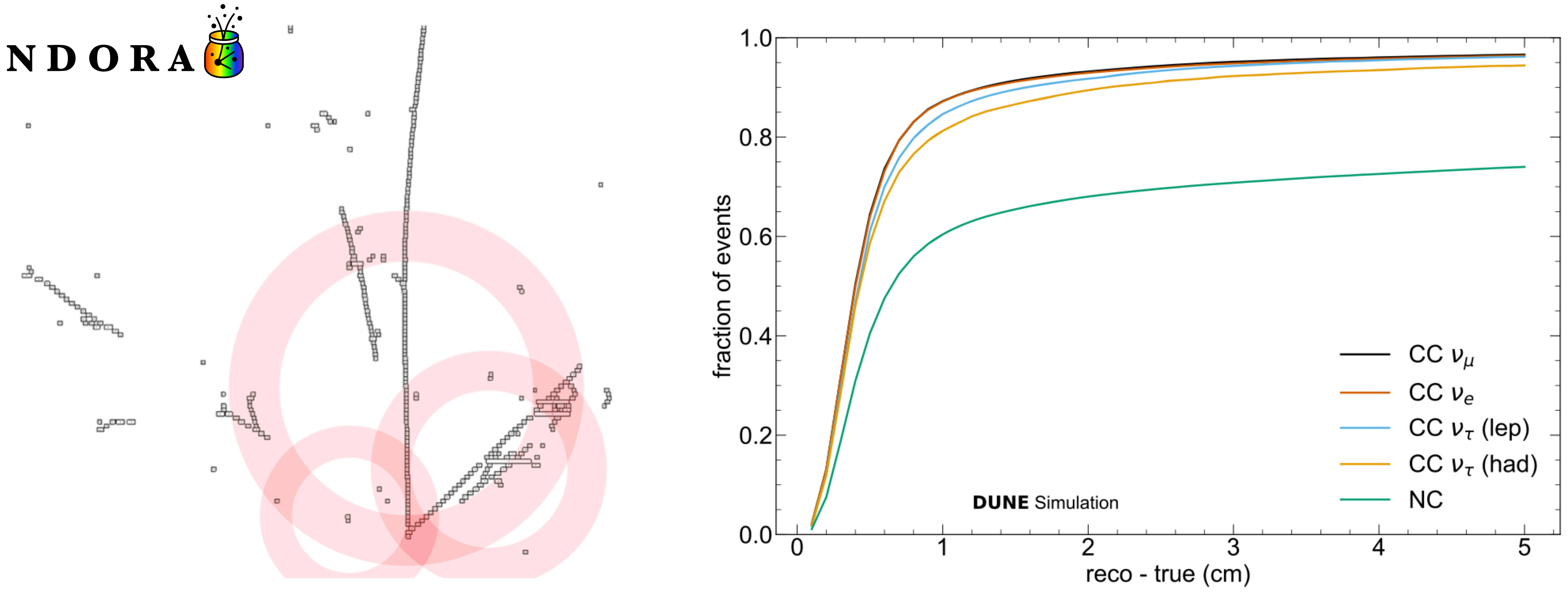
DUNE Collaboration, *Neutrino interaction classification with a convolutional neural network in the DUNE far detector,* Phys.Rev.D **102** 9, 092003 (2020)

Wire-Cell



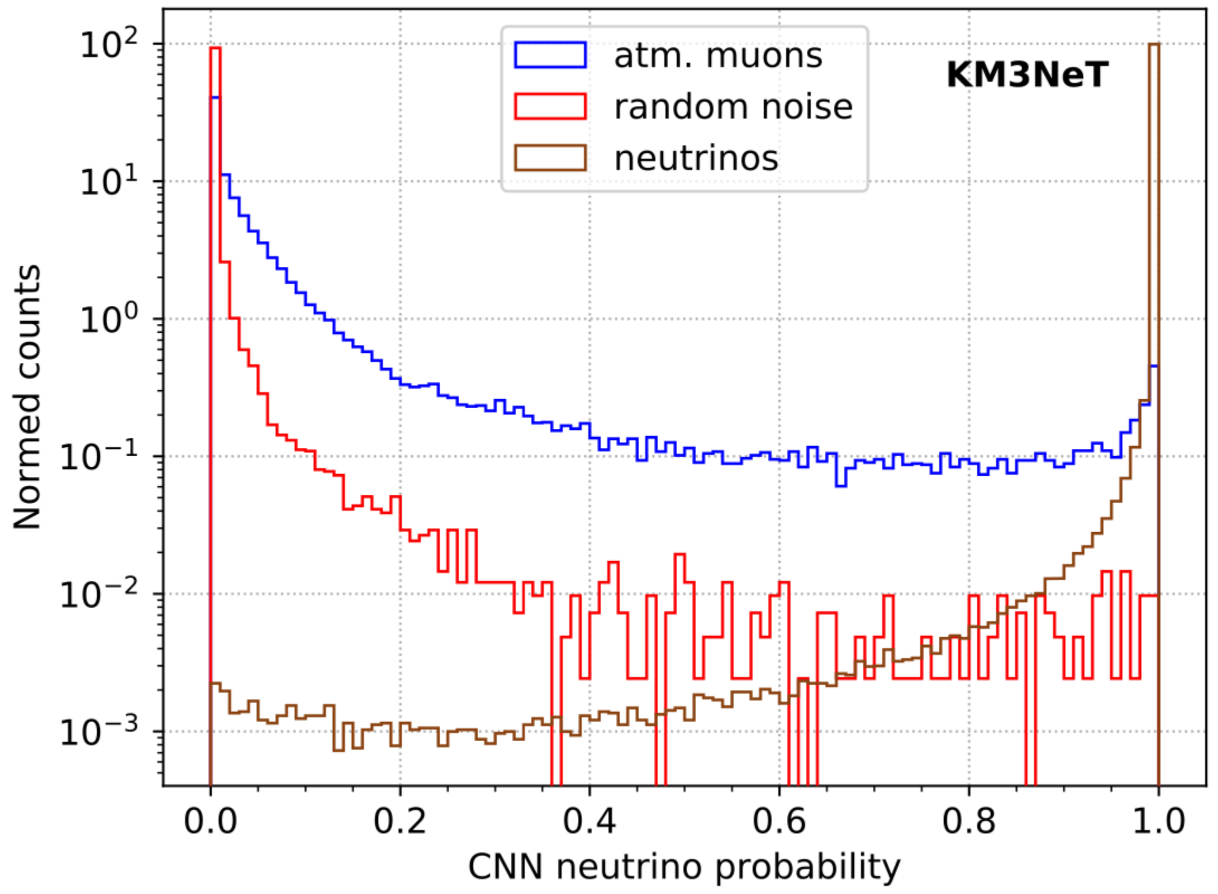
H. Yu, et al., Augmented signal processing in Liquid Argon Time Projection Chambers with a deep neural network, JINST **16** 01, P01036 (2021)

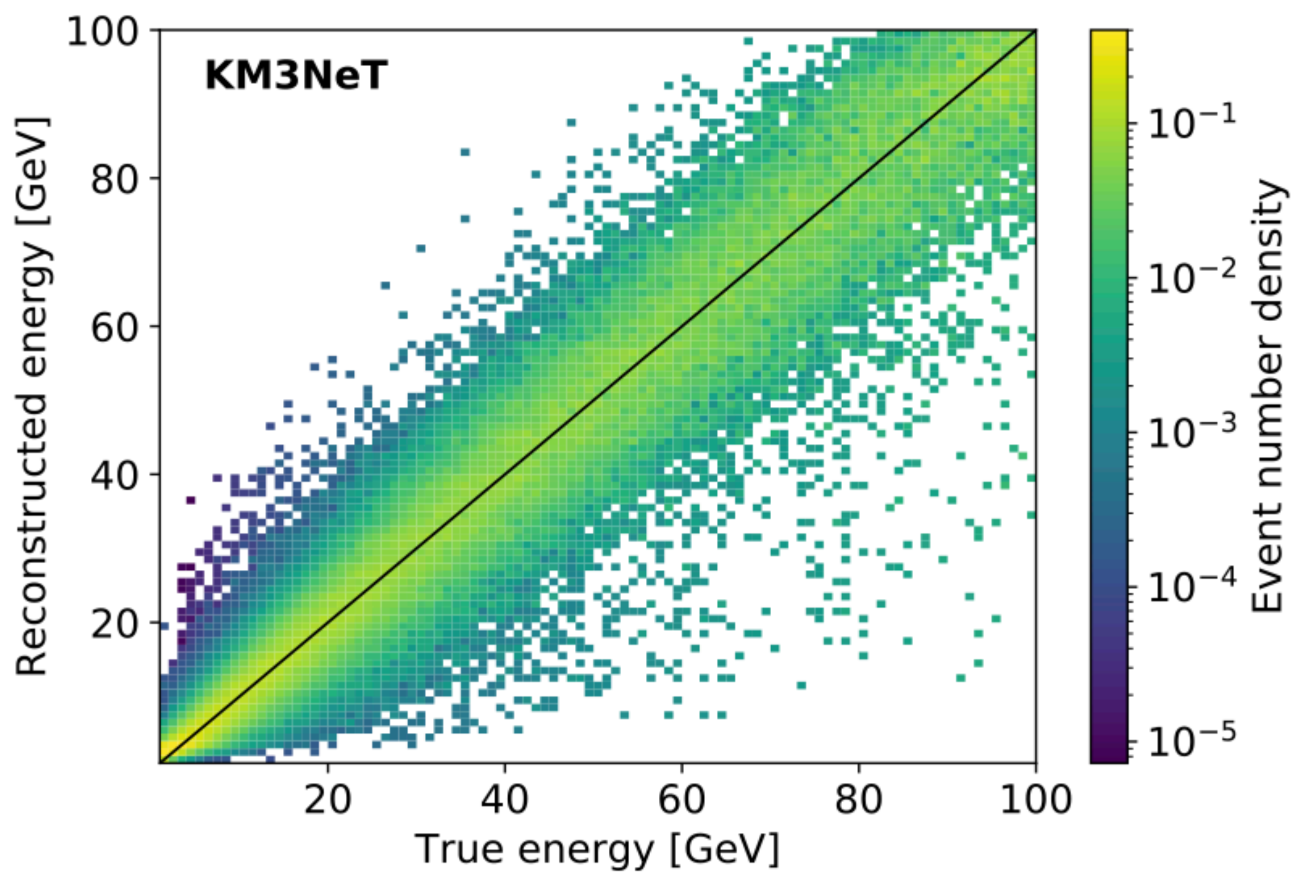
PANDORA



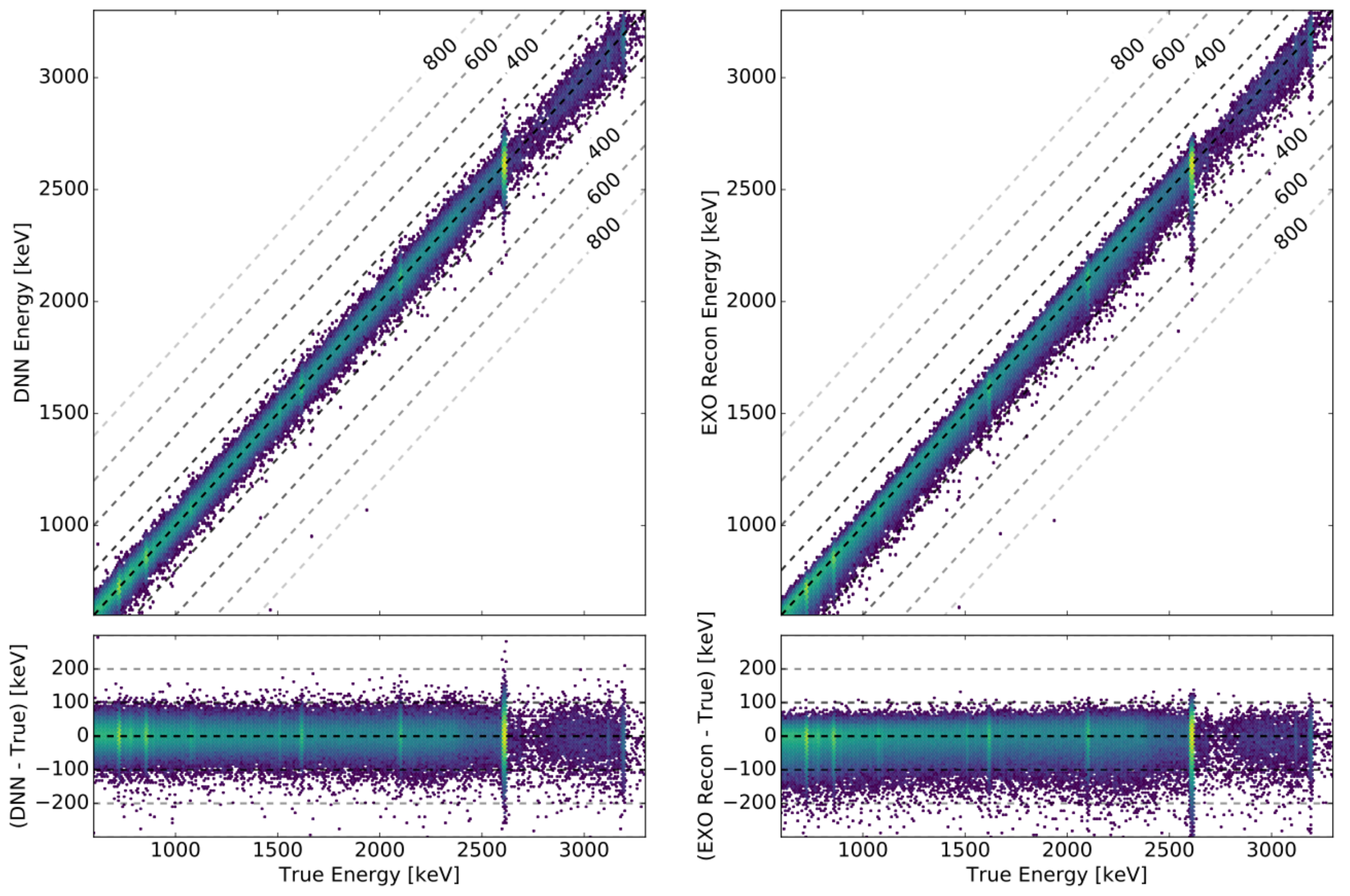
DUNE Collaboration, *Neutrino interaction vertex reconstruction in DUNE with Pandora deep learning,* Eur.Phys.J.C **85**, 697 (2025)

More CNN use-cases II

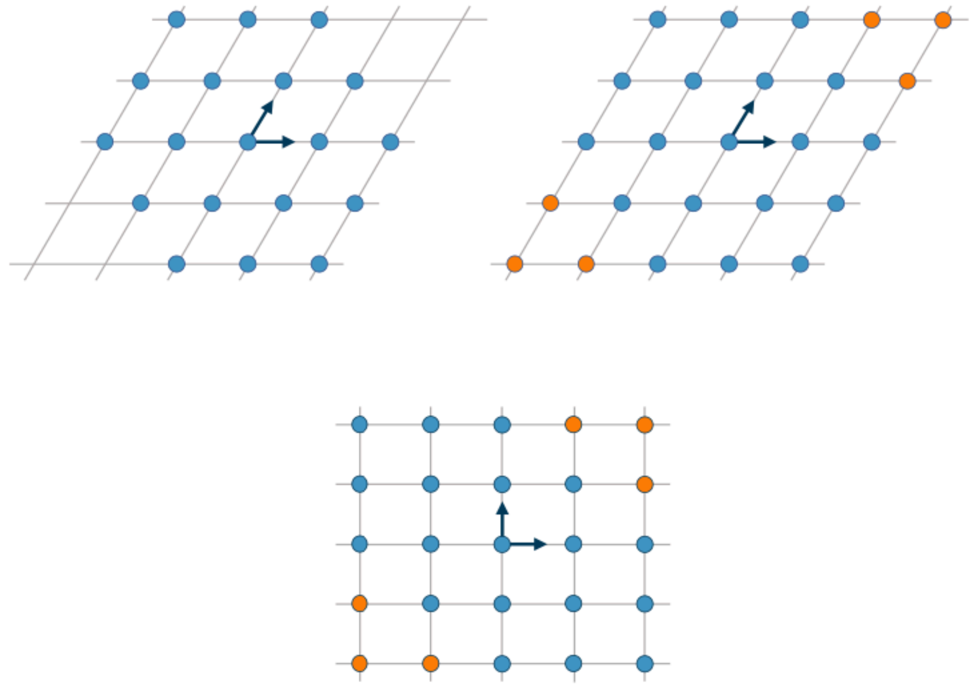


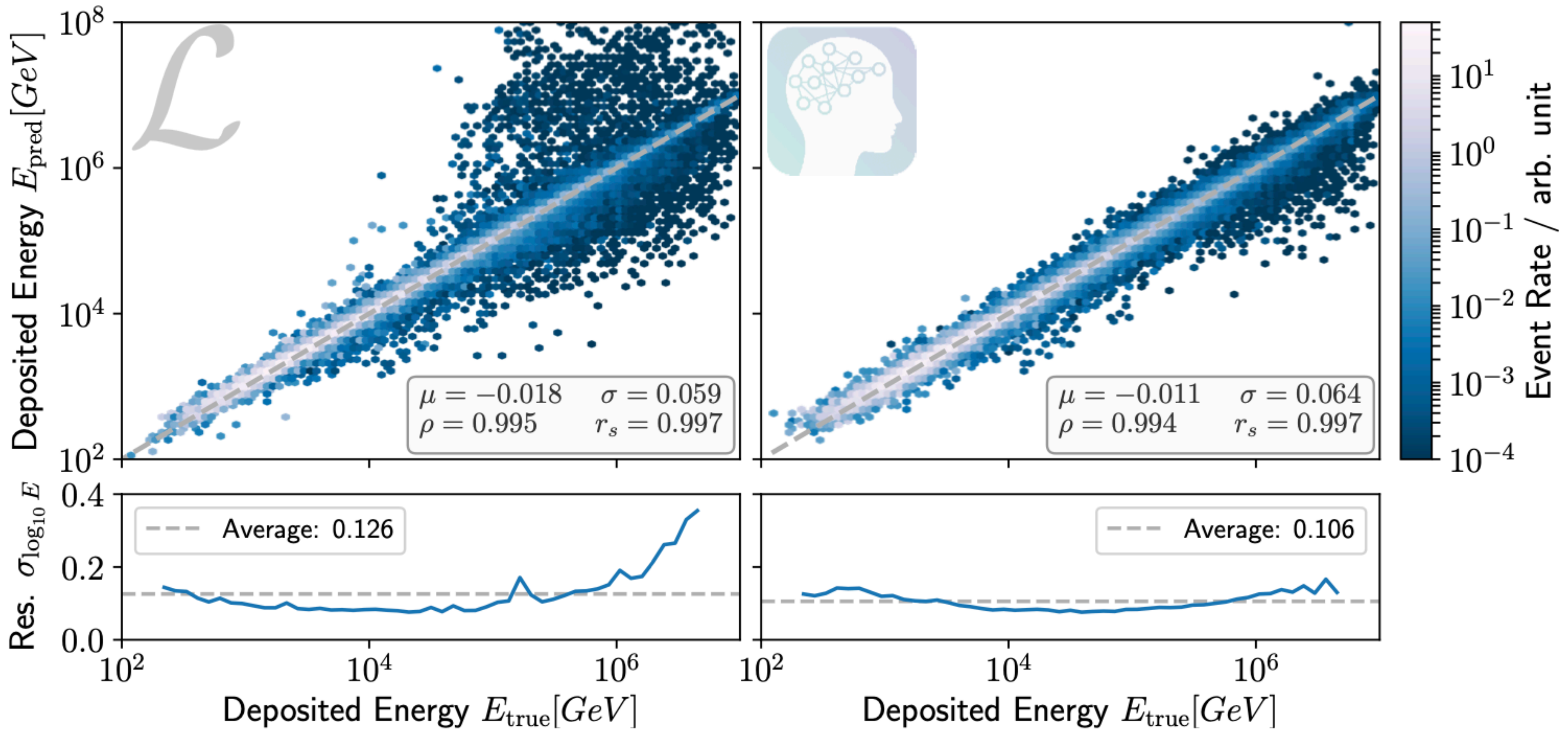


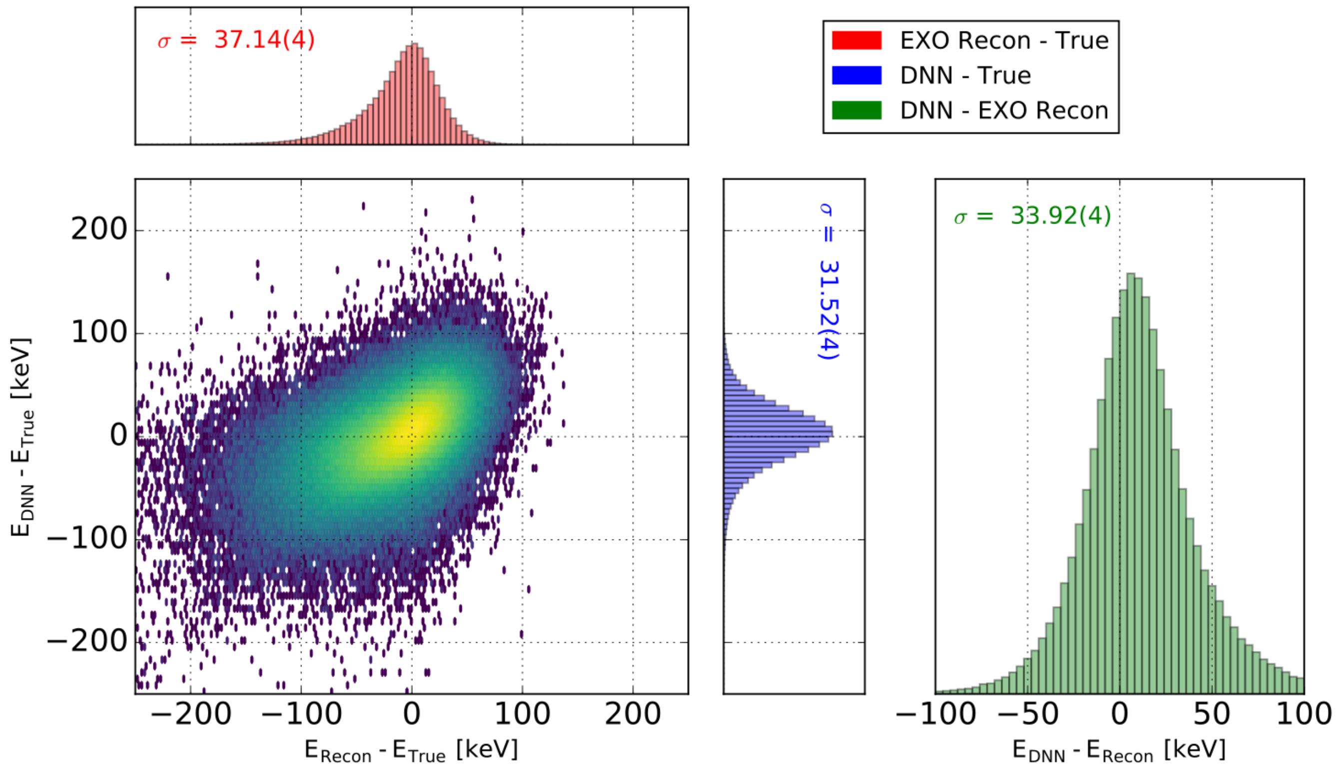
KM3NeT/ORCA Collaboration, *Event reconstruction for KM3NeT/ORCA using convolutional neural networks,* JINST **15**, P10005 (2020)



(a) DNN **(b)** EXO-recon







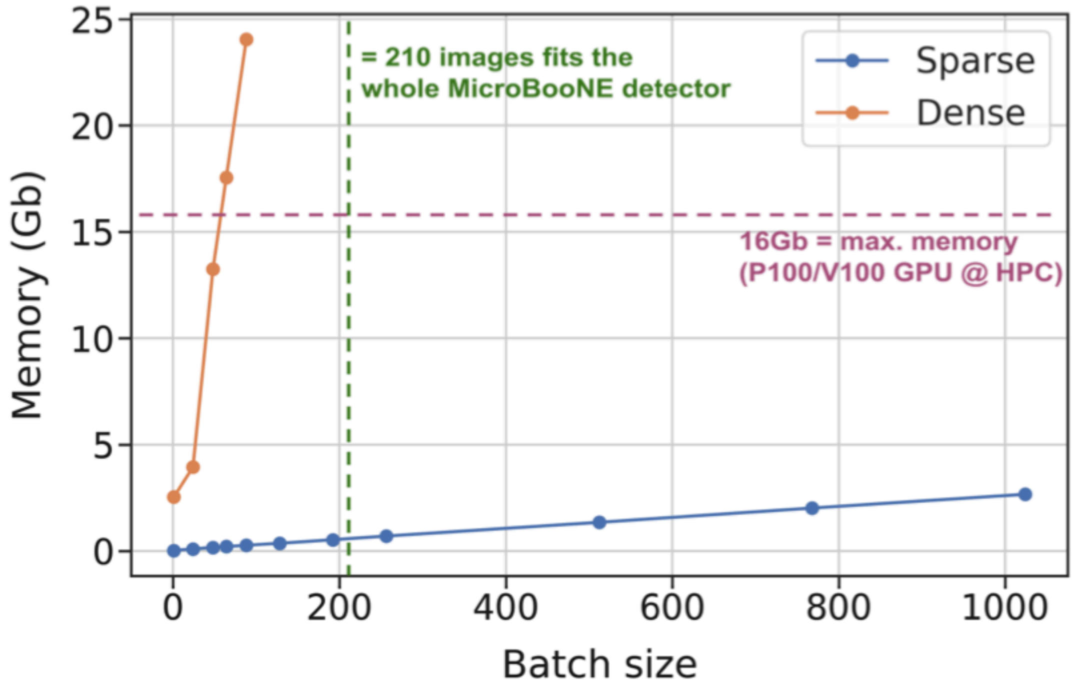
IceCube Collaboration, *A convolutional neural network based cascade reconstruction for the IceCube Neutrino Observatory, JINST* **16**, P07041 (2021)

S. Delaquis, et al., *Deep neural networks for energy and position reconstruction in EXO-200*, JINST **13** P08023 (2018)

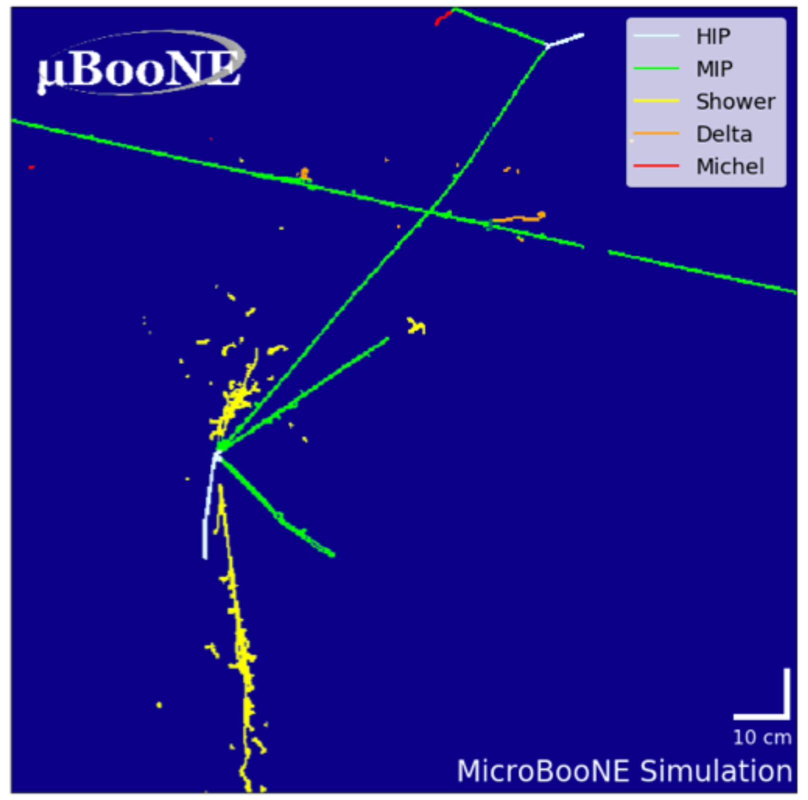
Going Sparse

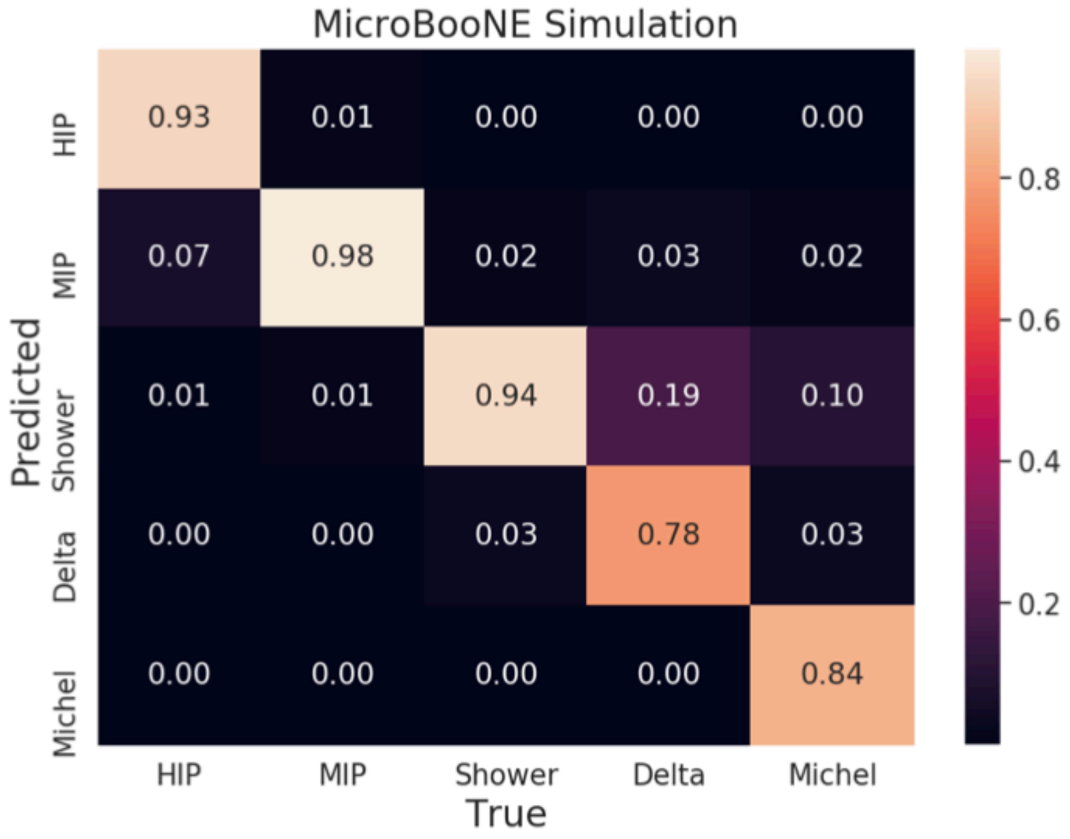
- You may have noticed that many of the images I’ve shown have lots of zeros
  - Images are usually sparse but locally dense: very inefficient for CNNs
- Sparse CNNs introduced
  - Significant reduction in memory use and run time, but also accuracy improvements

	Dense	Sparse
Batch size	4	4
Image size	192 px	192 px
Nonzero accuracy mean	92%	94%
Nonzero accuracy std	0.096	0.088
		Nvidia V100 GPU
Memory (test) [GB]	16	0.044
Memory (train) [GB]	26x4	0.21
Wall-time (test) [s]	3.3	0.10
Wall-time (train) [s]	25	0.21



L Dominé and K. Terao, Scalable deep convolutional neural networks for sparse, locally dense liquid argon time projection chamber data, Phys. Rev. D **102** 1, 012005 (2020)



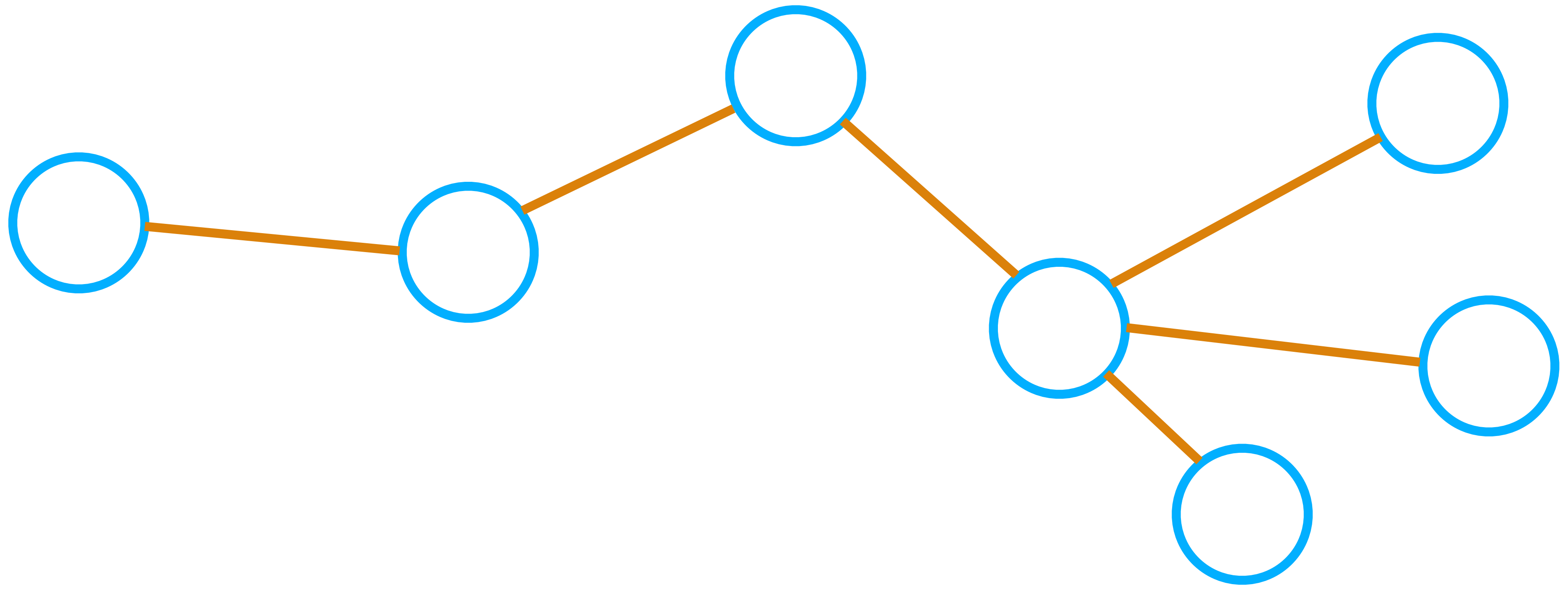


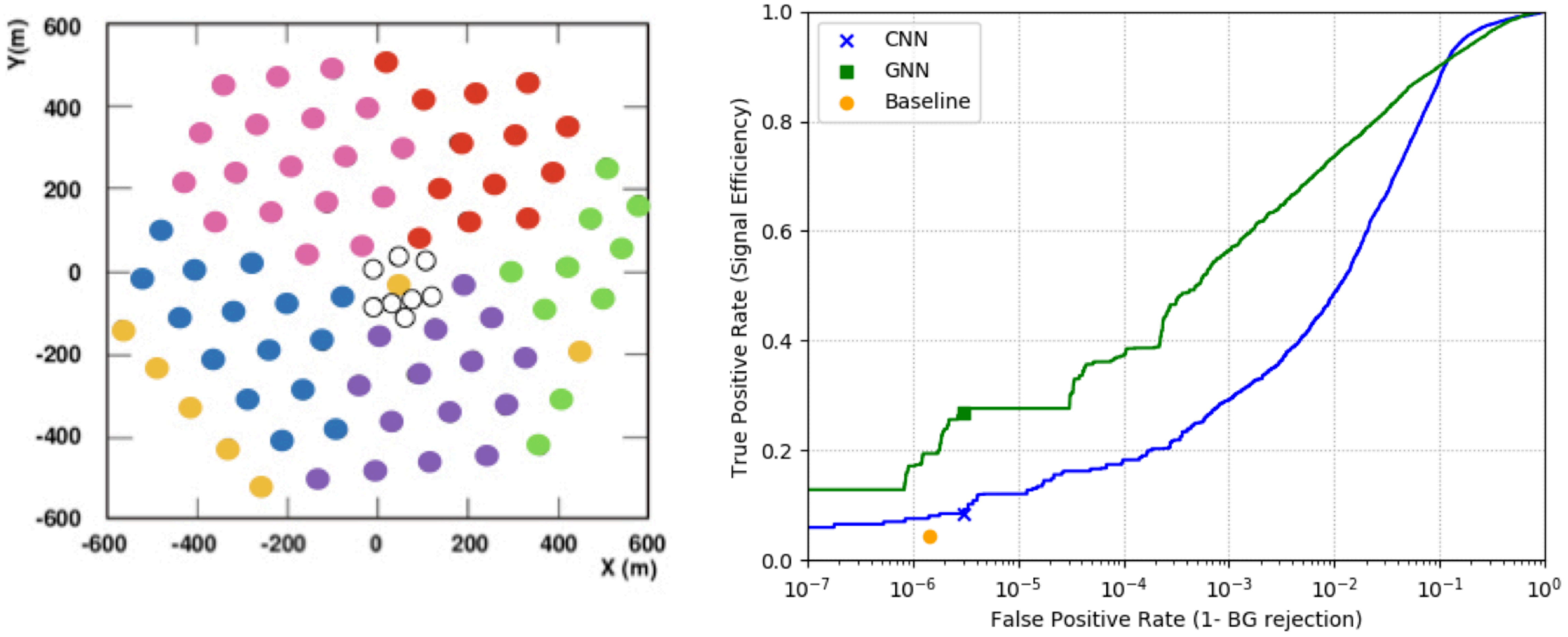
P. Abratenko et al., *Semantic segmentation with a sparse convolutional neural network for event reconstruction in MicroBooNE*, Phys. Rev. D **103**, 052012 (2021)

Graph Neural Networks

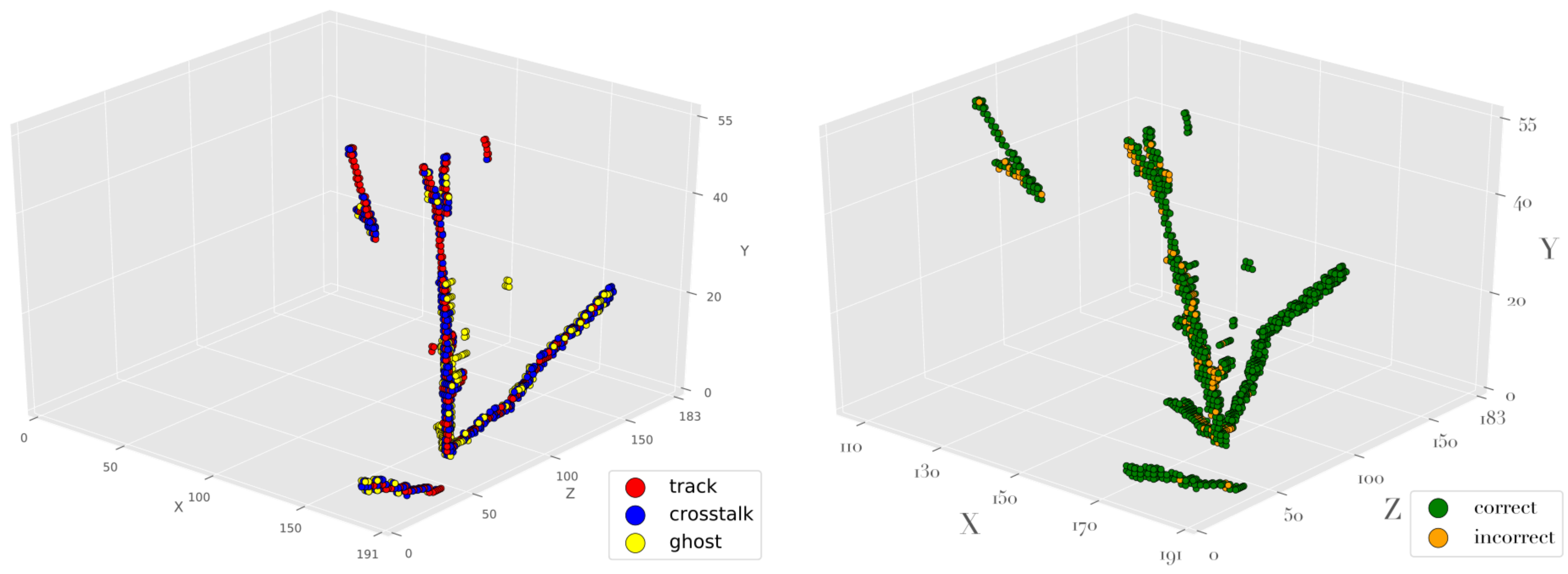
Graph Neural Networks

- Not all data is easily viewed as an image
- Graphs can provide a more natural data representation
  - They are also sparse by construction





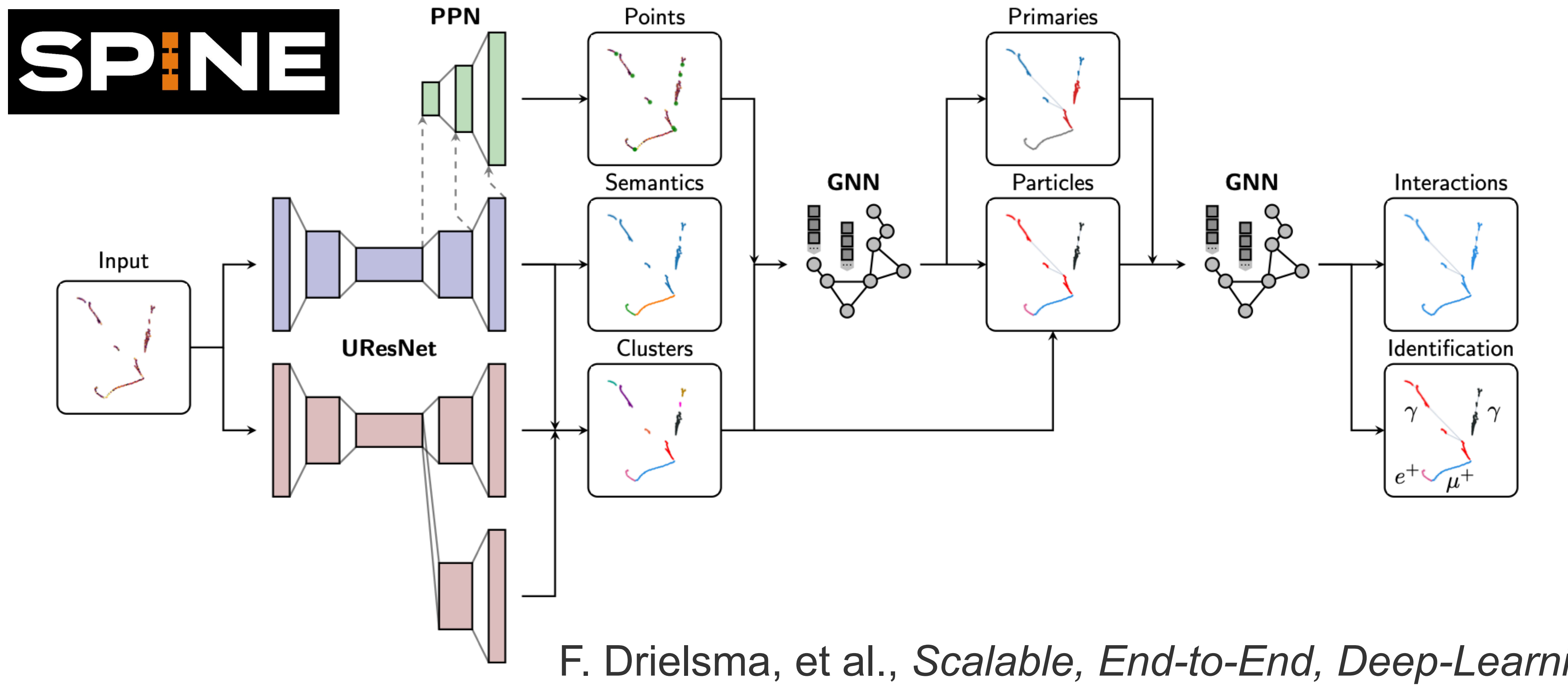
N. Choma, et al., *Graph Neural Networks for IceCube Signal Classification*, arXiv: 1809.06166 [cs.LG] (2018)

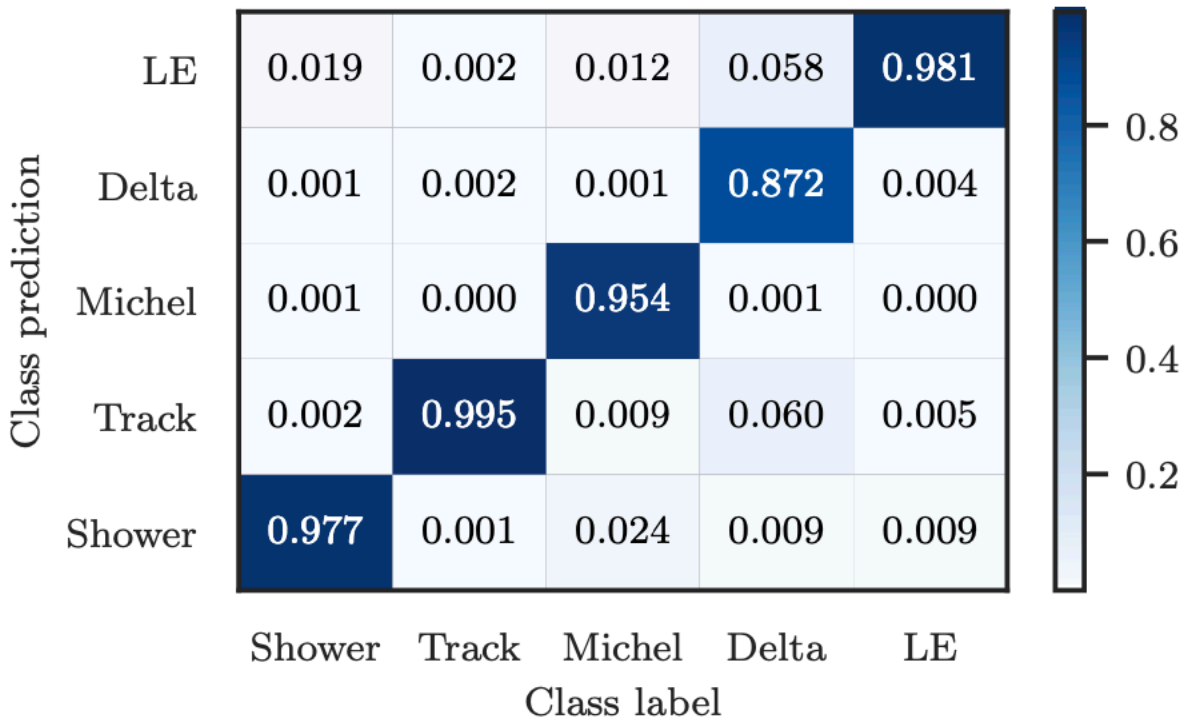


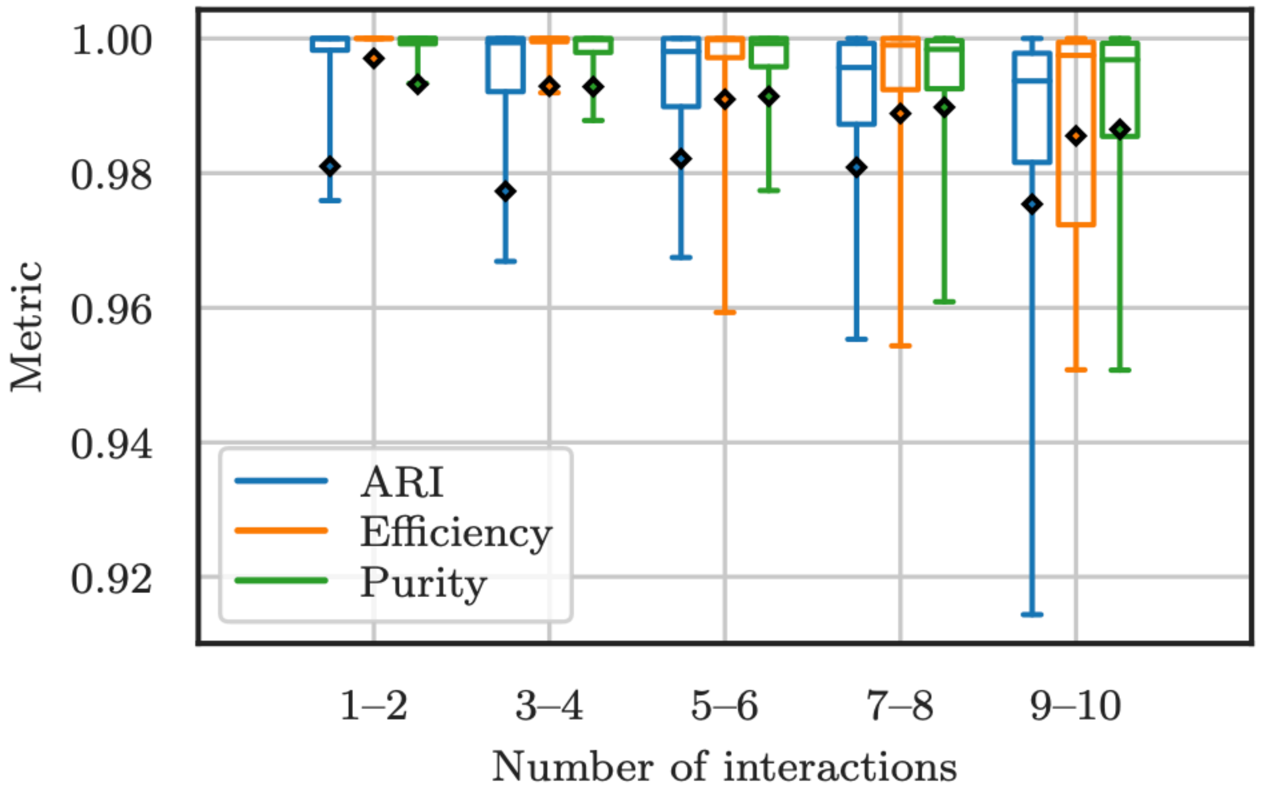
S. Alonso-Monsalve, et al., *Graph neural network for 3D classification of ambiguities and optical crosstalk in scintillator-based neutrino detectors*, Phys. Rev. D **103** 3, 032005 (2020)

Graph Neural Networks

- GNN-based end-to-end reconstruction

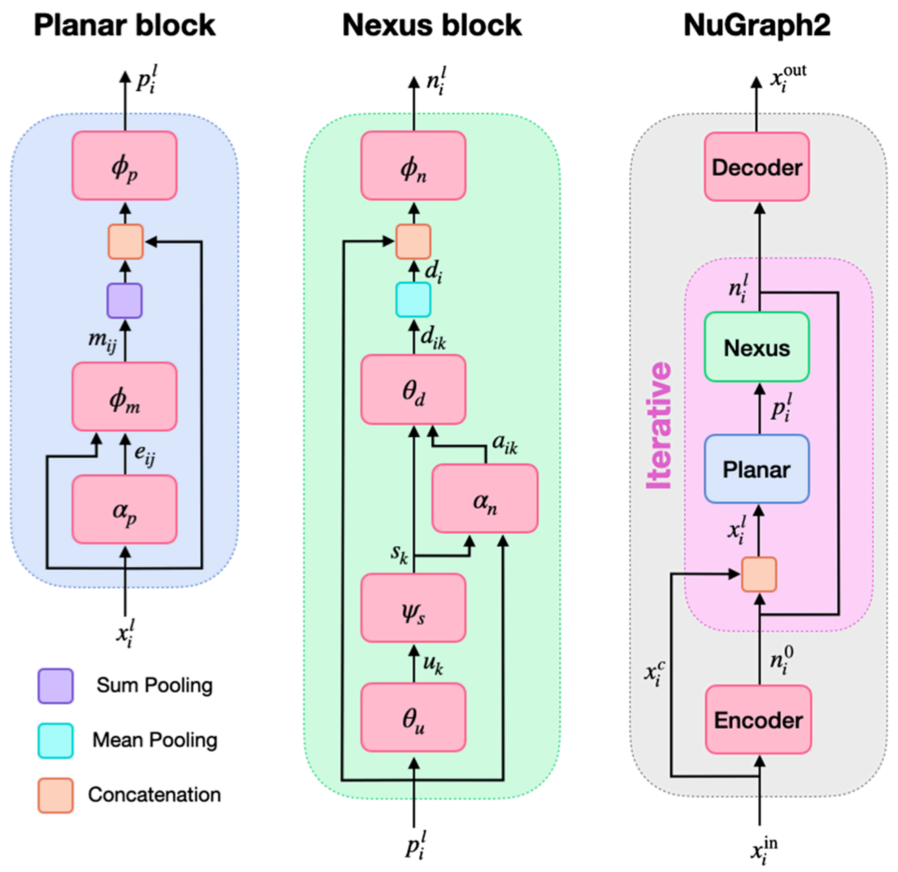


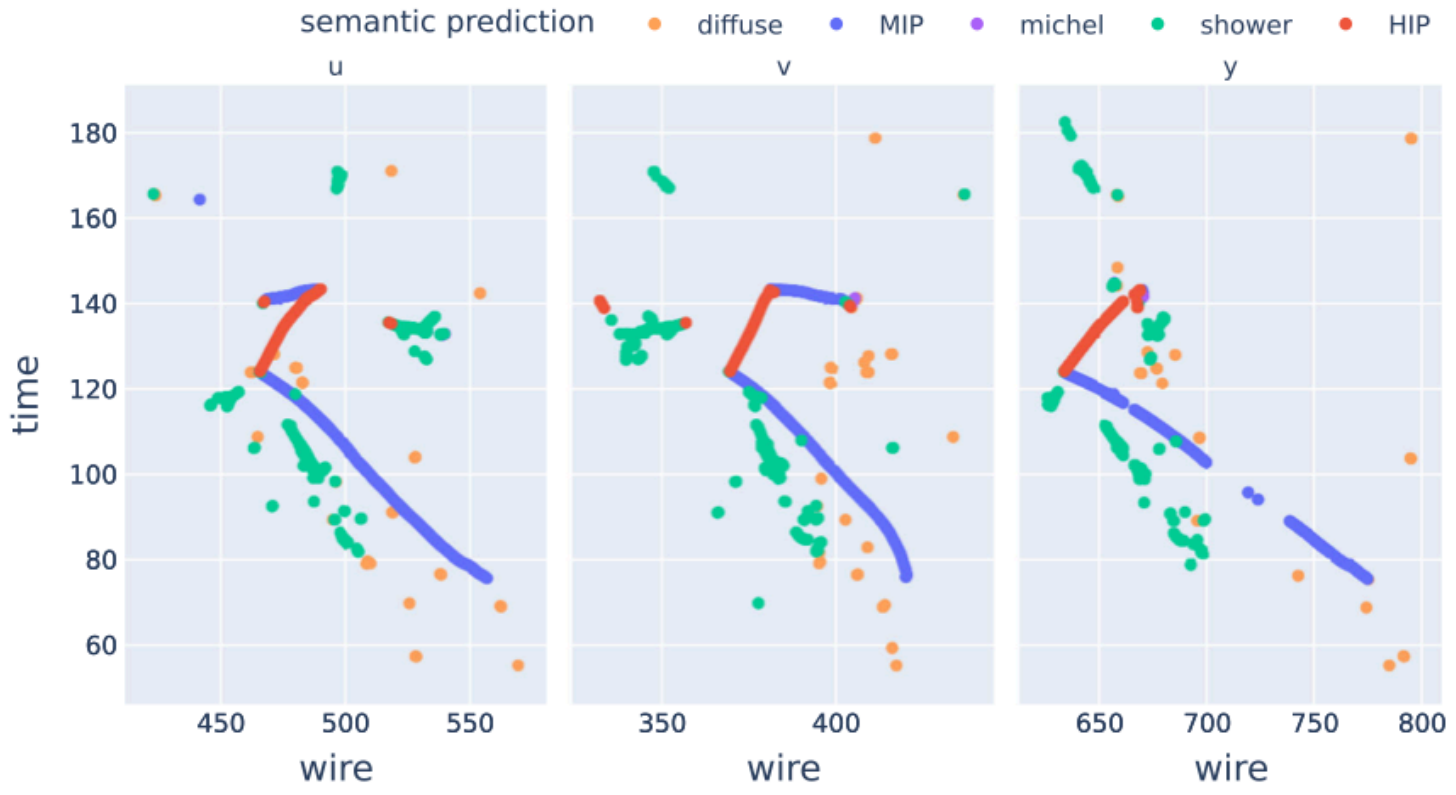


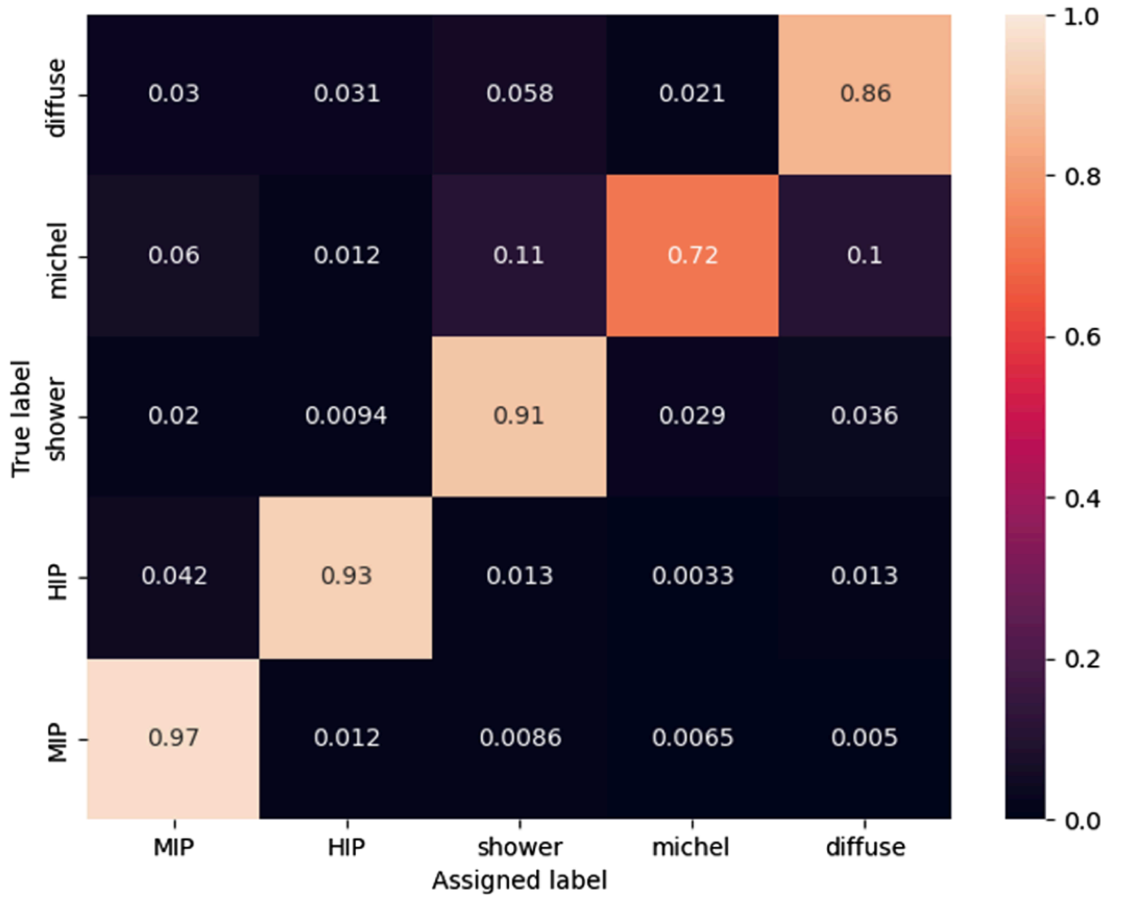


F. Drielsma, et al., *Scalable, End-to-End, Deep-Learning-Based Data Reconstruction Chain for Particle Imaging Detectors*, arXiv: 2102.01033 [hep-ex] (2021)

NuGraph2





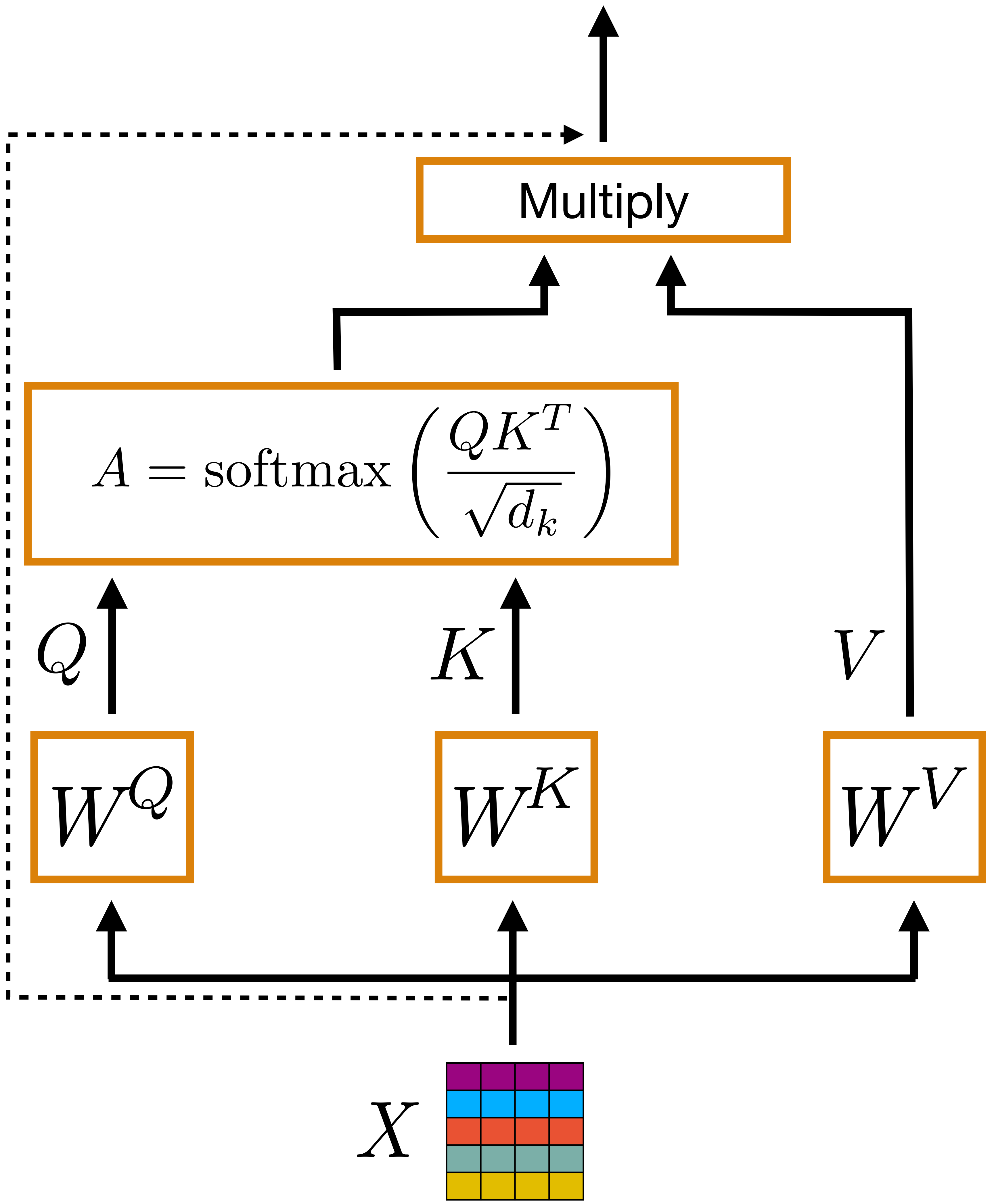


A. Aurisano, et al., *Graph neural network for neutrino physics event reconstruction*, Phys. Rev. D **110** 3, 032008 (2024)

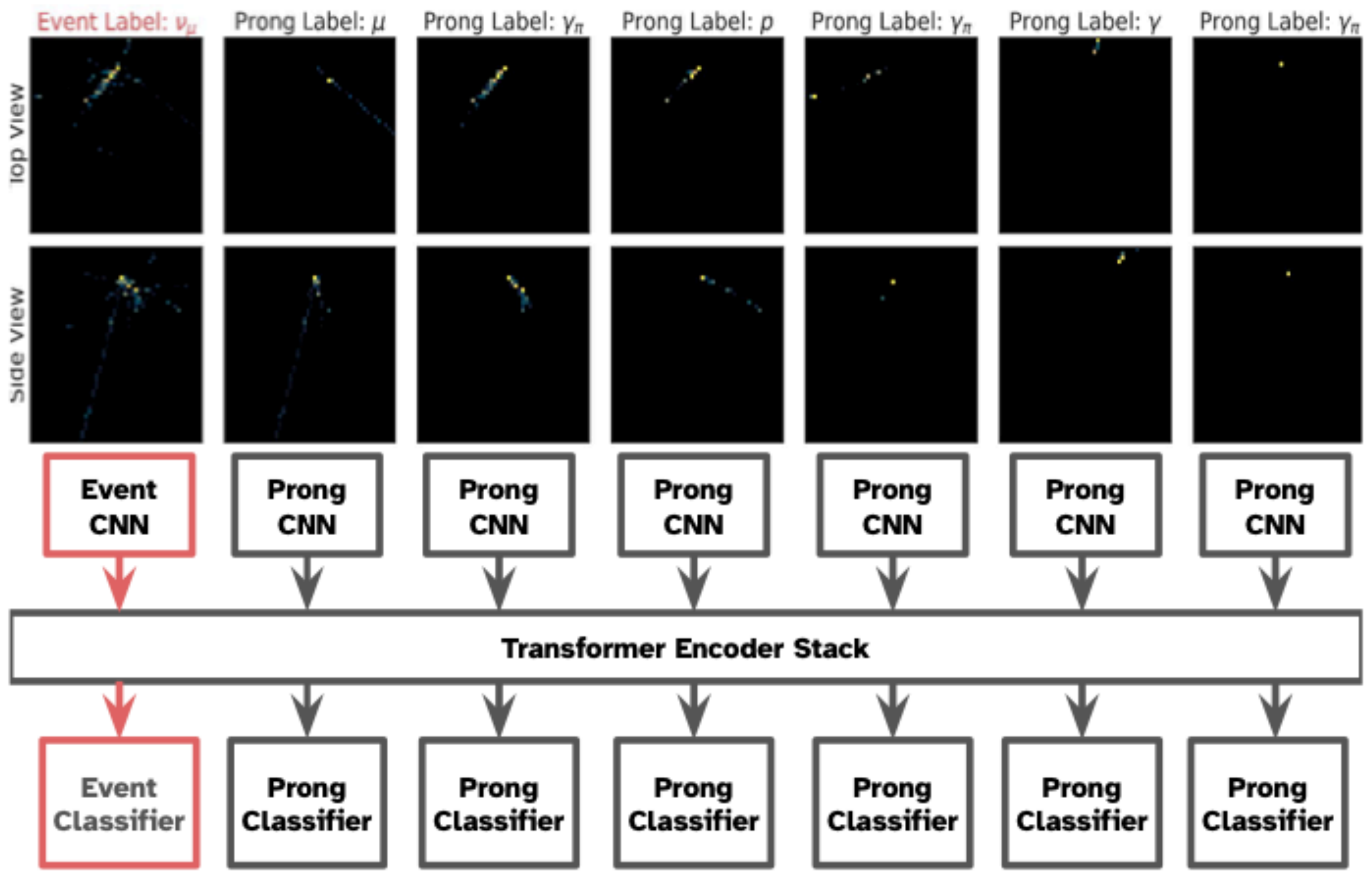
Transformers

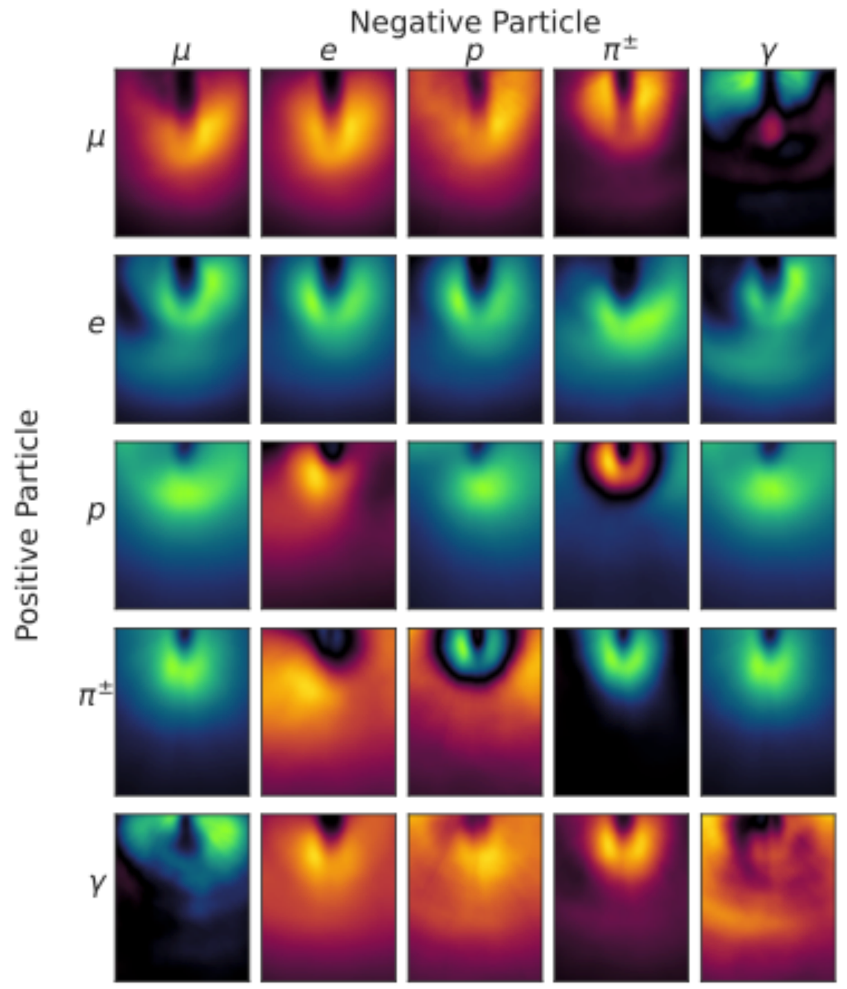
Transformers

- Introduced in 2017
  - Has become a leading technique in AI/ML
  - Form the backbone of LLMs
- The main feature of the transformer is the attention mechanism
  - It is very simple (in my opinion)
  - Three linear projections
  - Dot-product (typically) attention
  - Matrix multiplication
  - Two-layer feed-forward network (not shown)



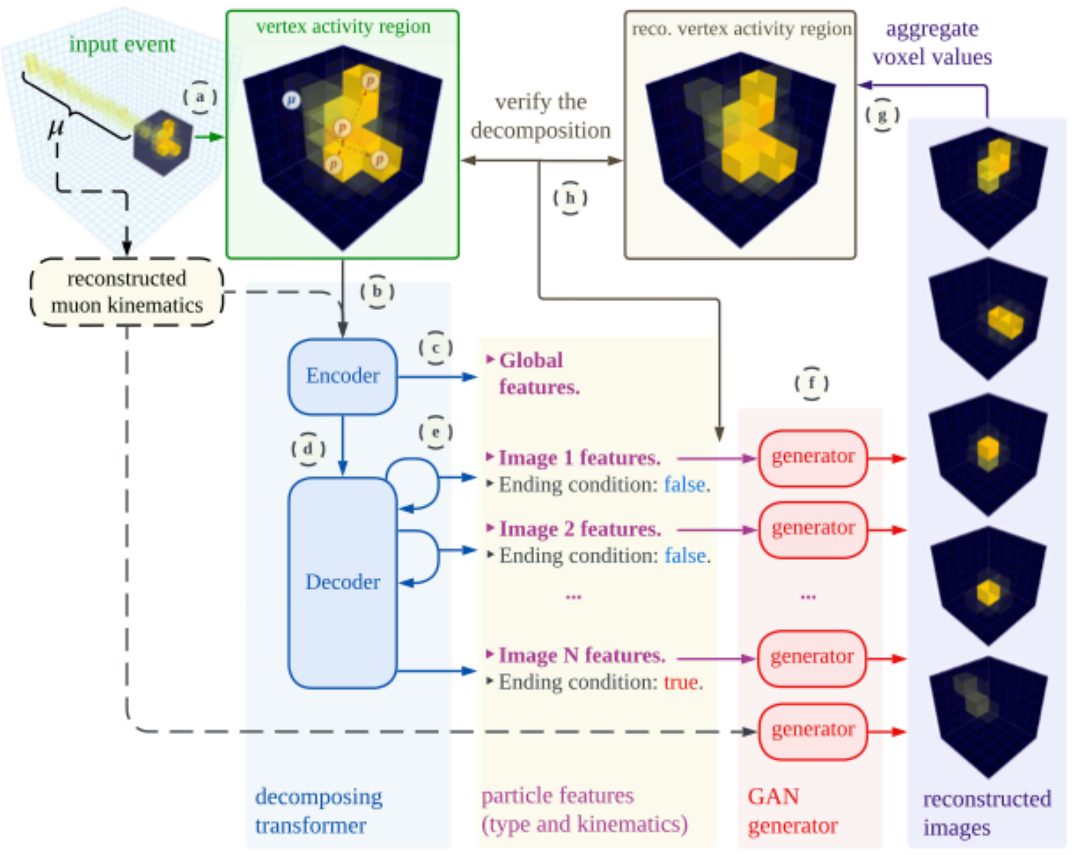
Transformers Use Cases I

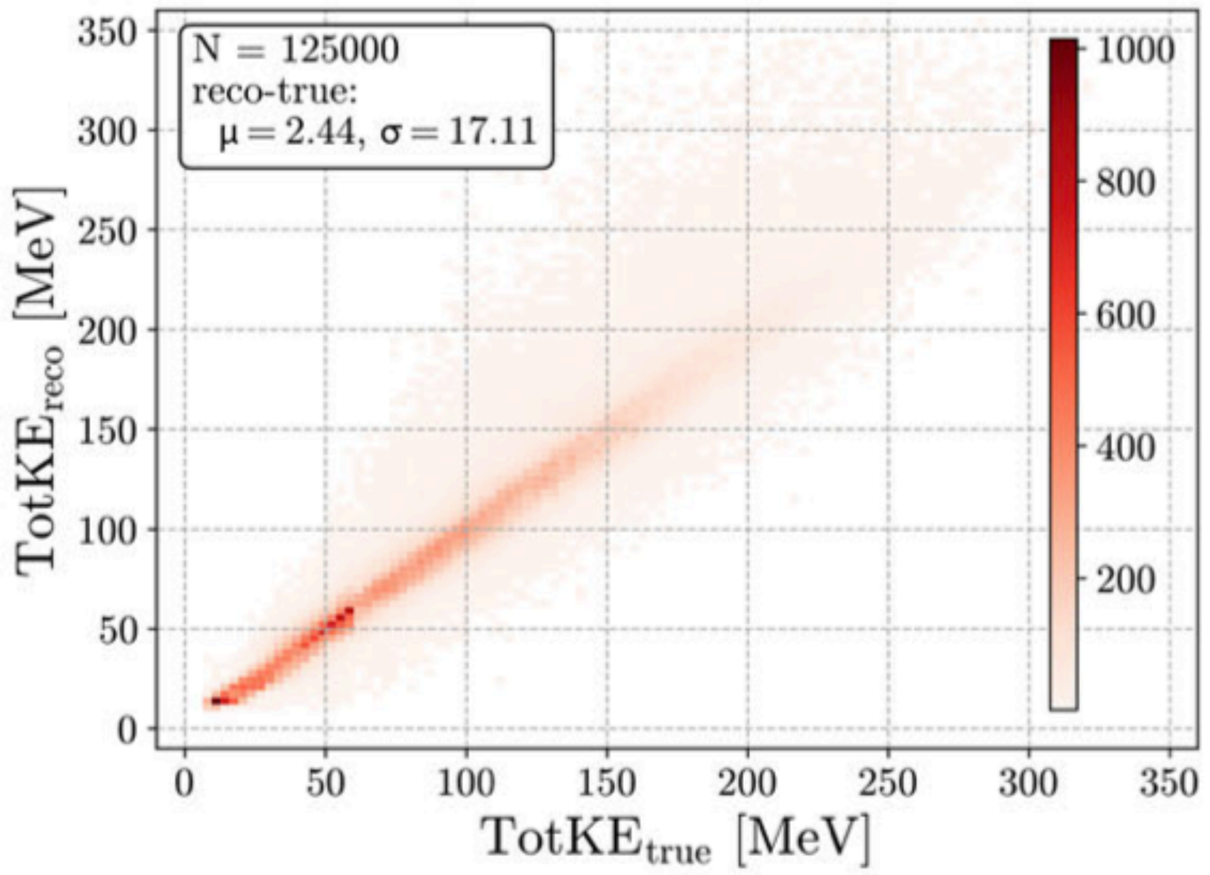


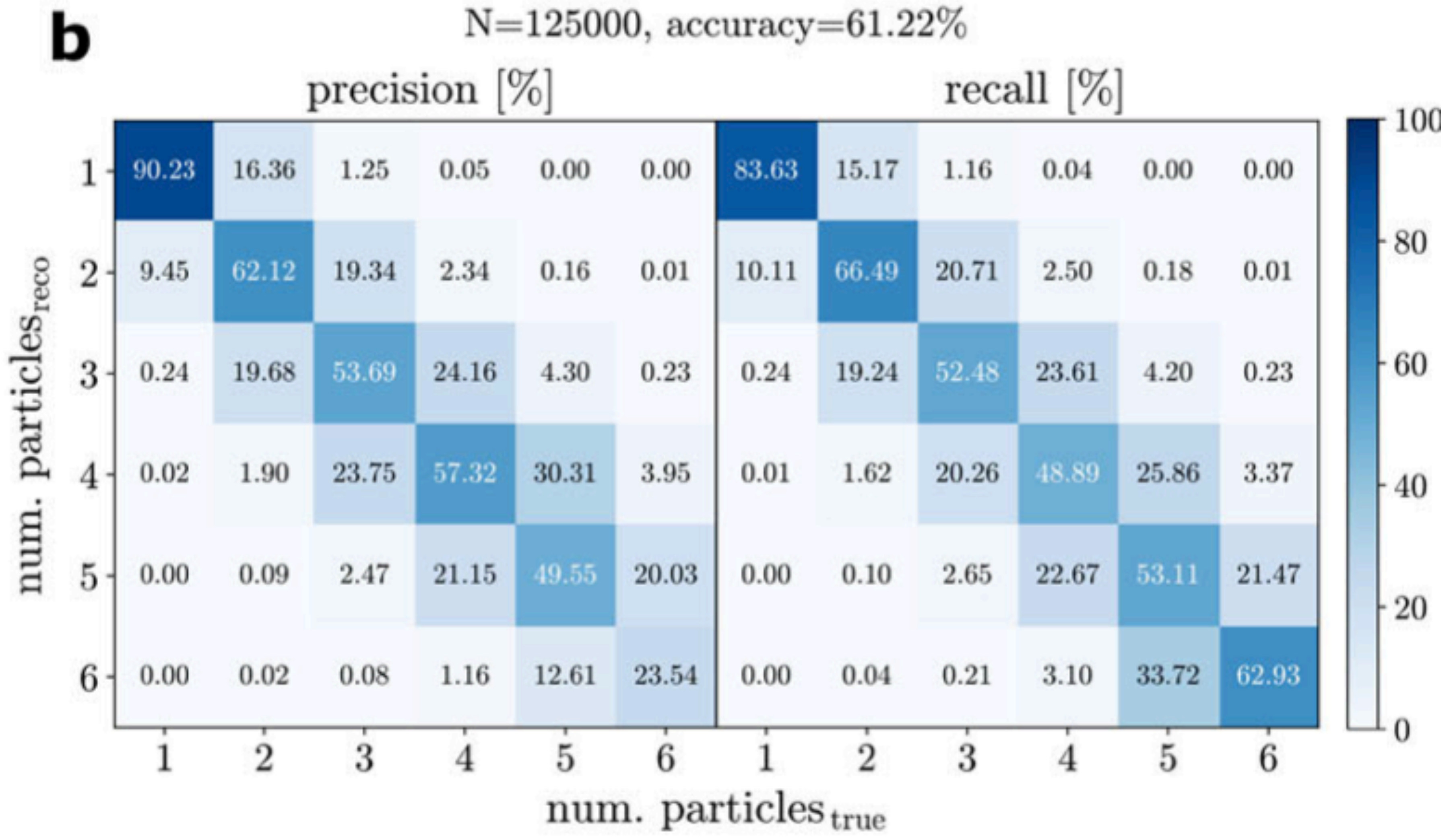


Metric	Transformer CVN	Prong CVN
Accuracy	0.783	0.726
Precision	0.783	0.760
Recall	0.783	0.726
ROC AUC	0.951	0.932

A. Shmakov, et al., *Interpretable Joint Event-Particle Reconstruction for Neutrino Physics at NOvA with Sparse CNNs and Transformers*, arXiv: 2303.06201 [cs.LG] (2023)

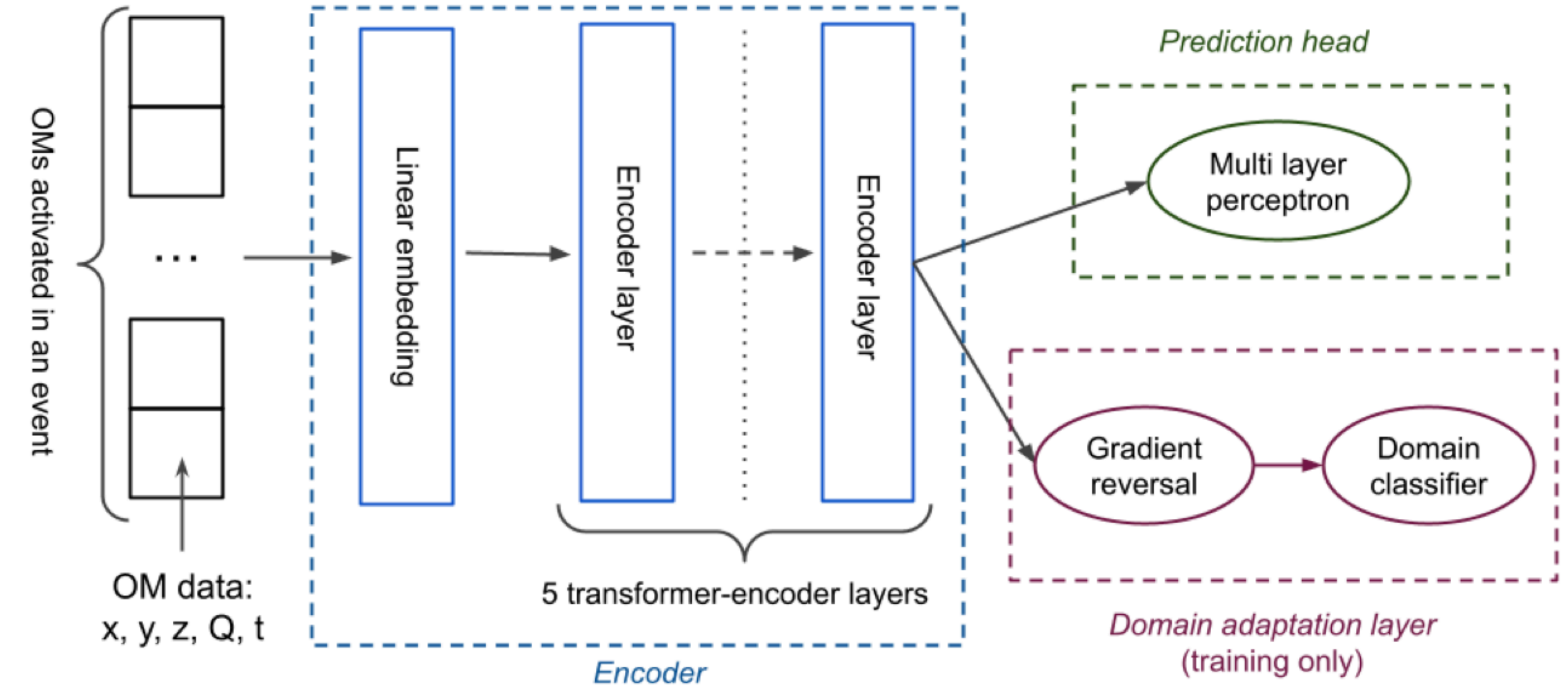


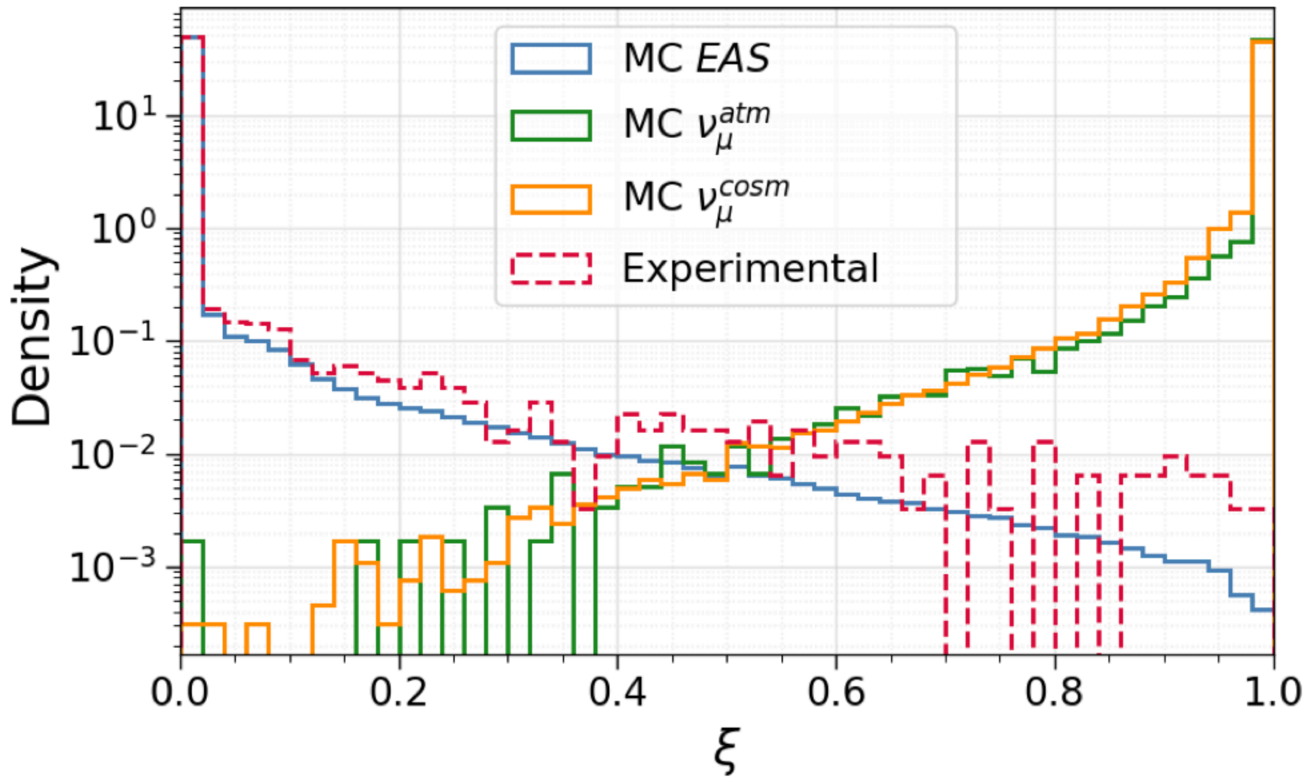


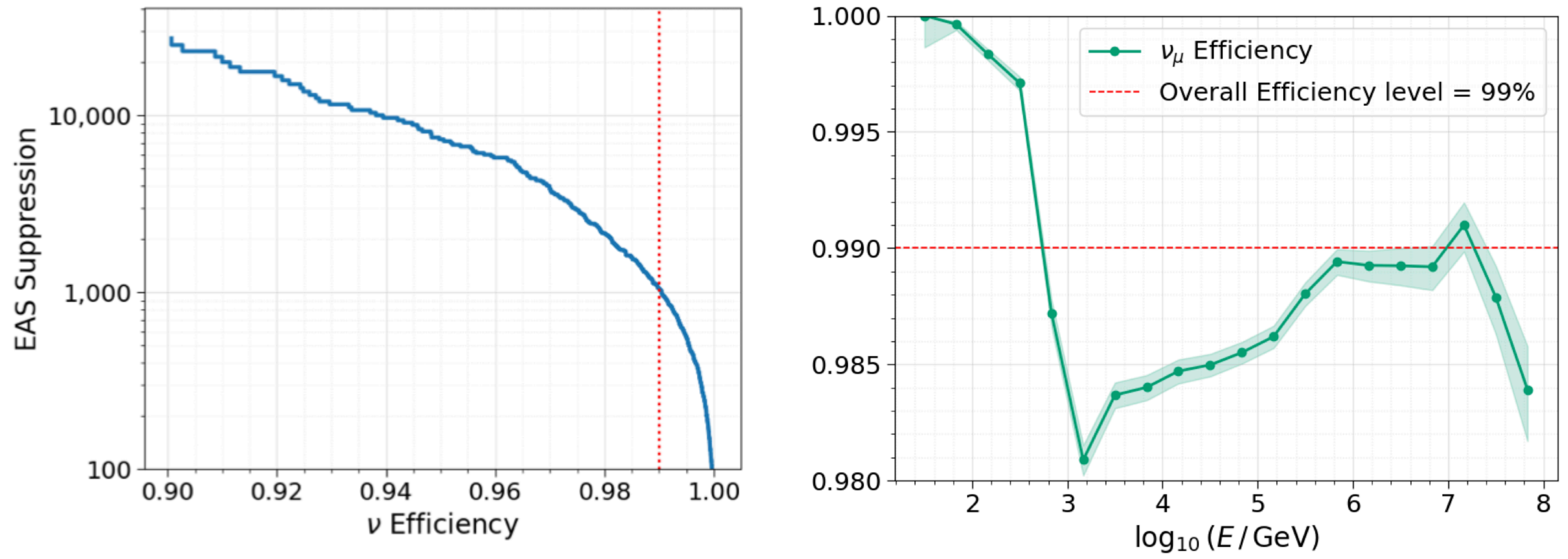


S. Alonso-Monsalve, *et al.*, *Deep-learning-based decomposition of overlapping-sparse images: application at the vertex of simulated neutrino interactions*, *Commun Phys* **7**, 173 (2024)

Transformer Use Cases II



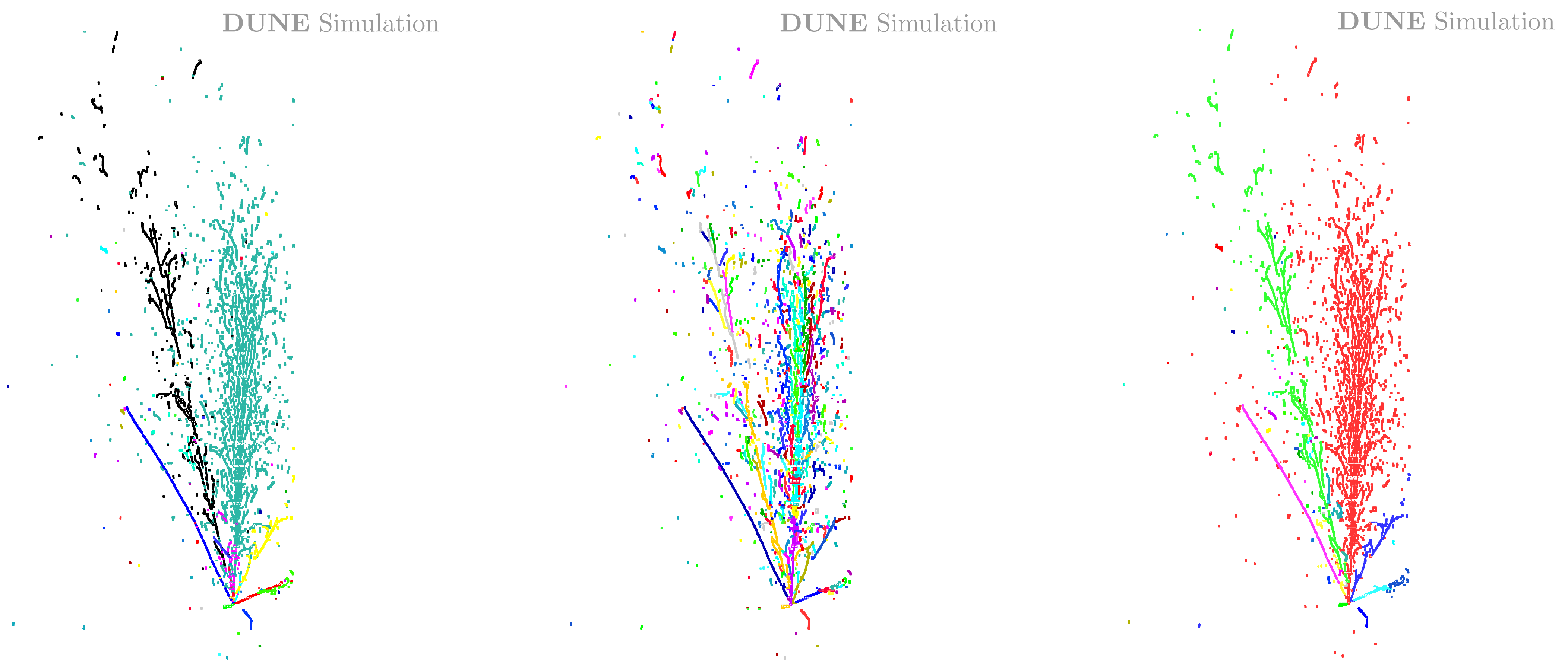


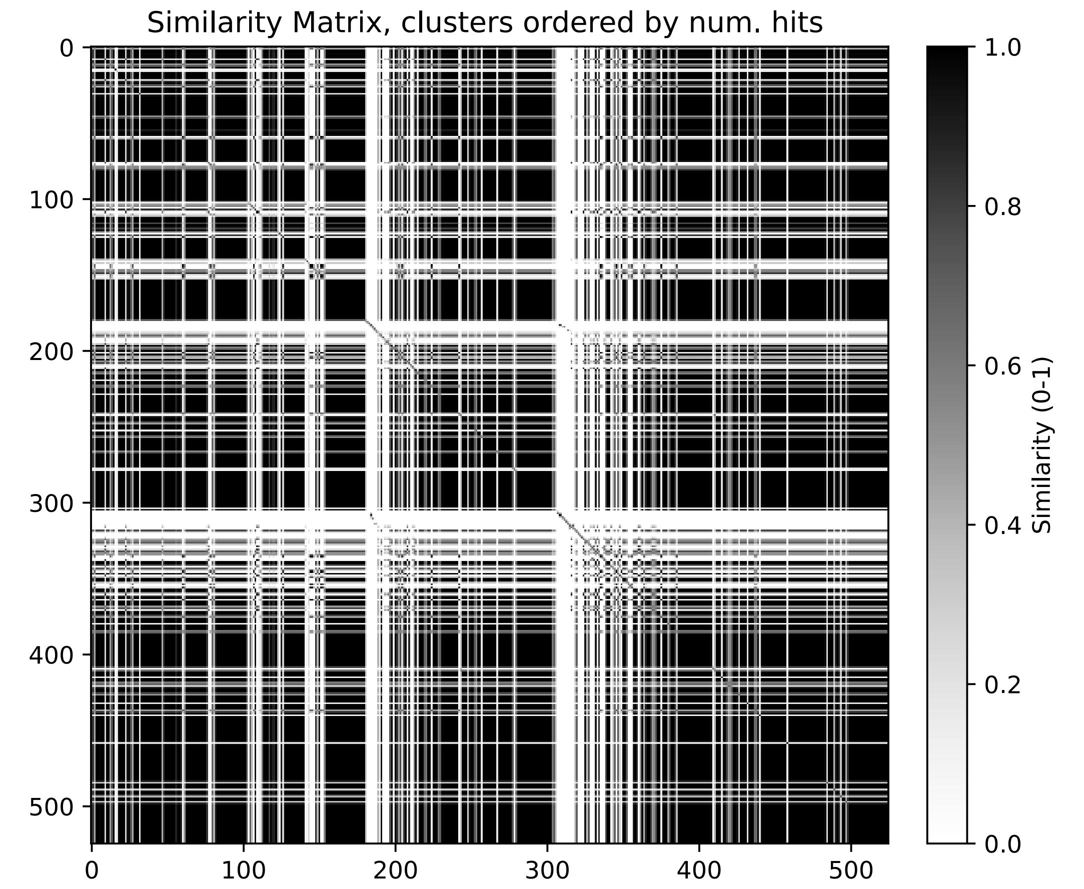


A. Matseiko, et al., *From raw data to neutrino candidates: a neural-network pipeline for Baikal-GVD*, arXiv:2605.11176 [astro-ph.IM]

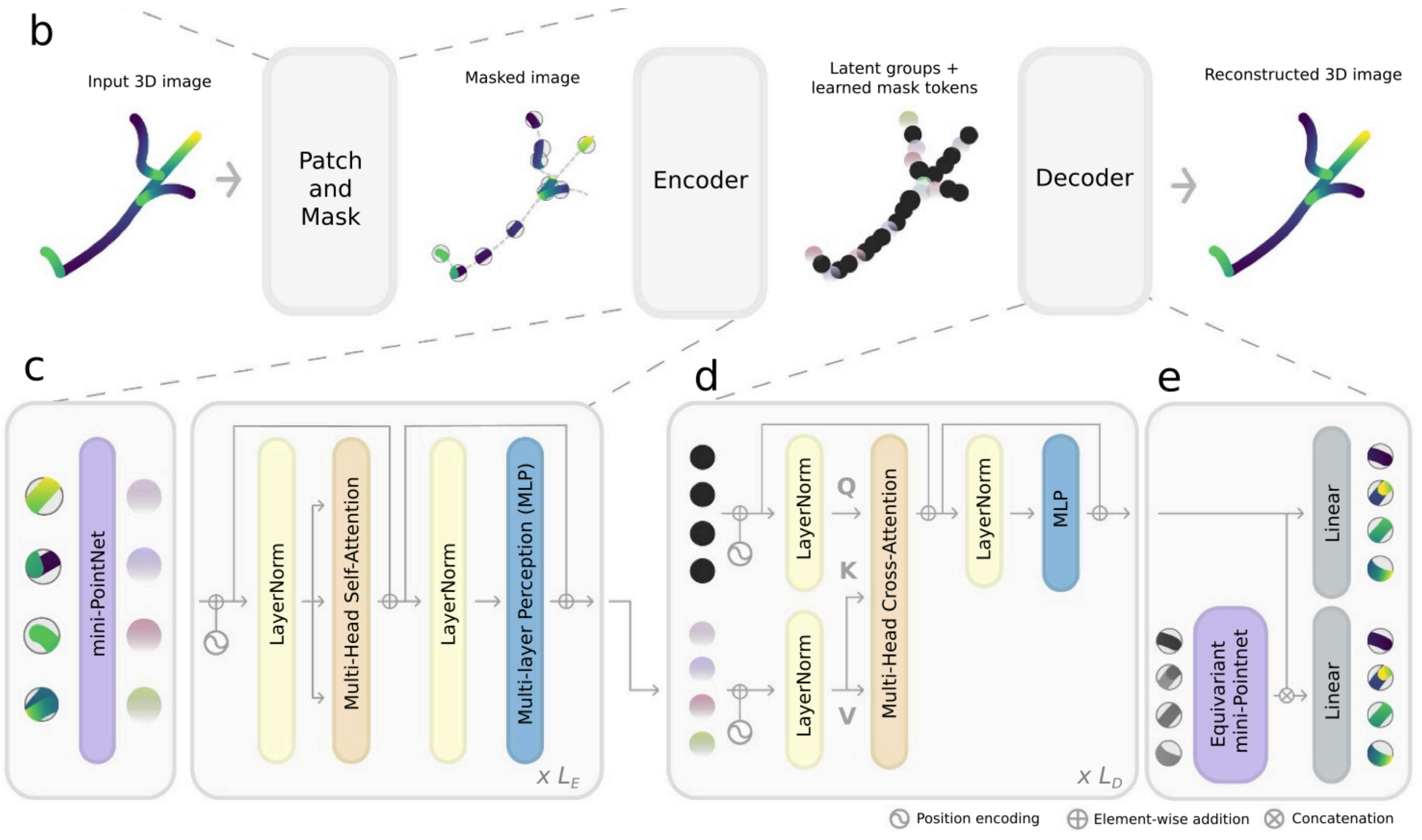
PANDORA

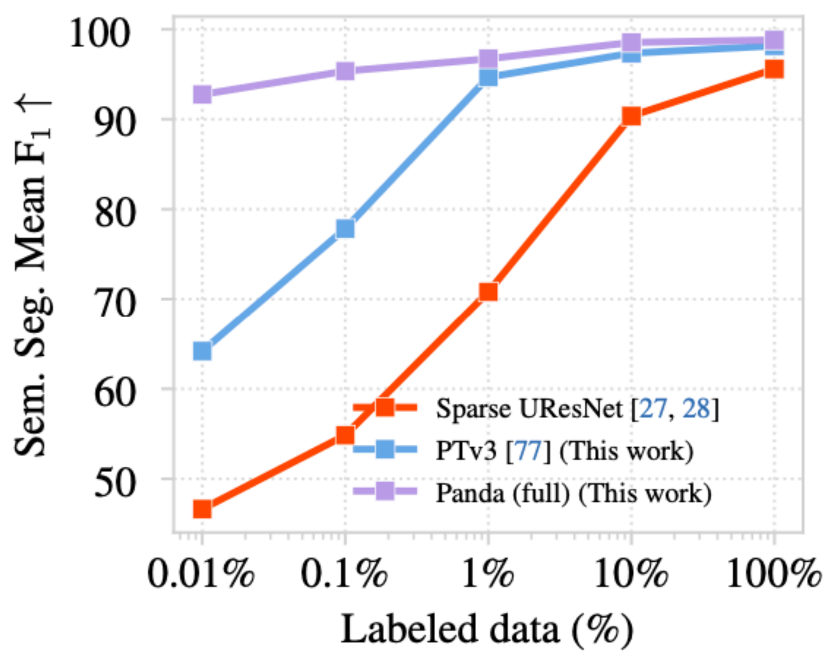
- Transformer encoder-based network to combine small clusters into EM showers
  - Predicts similarity matrix between clusters
- Significant performance increase over traditional methods



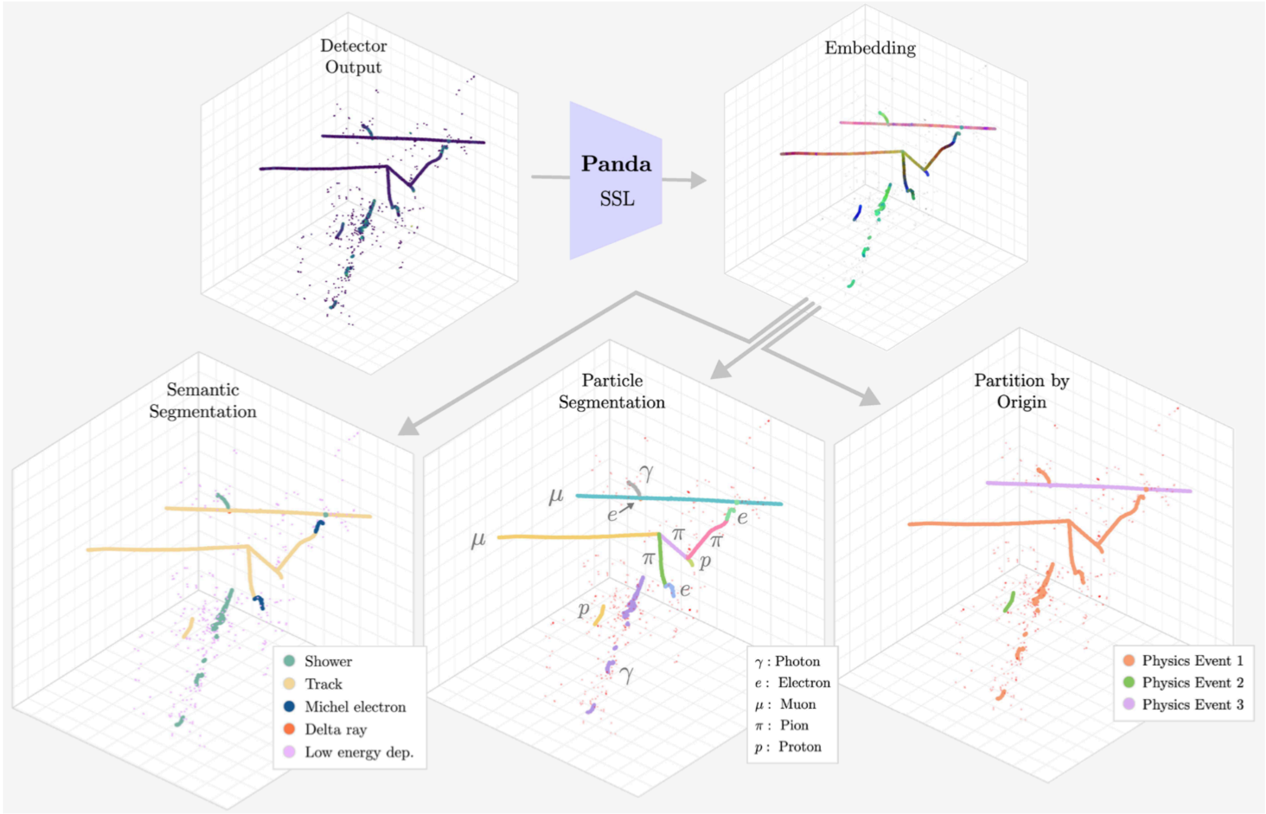


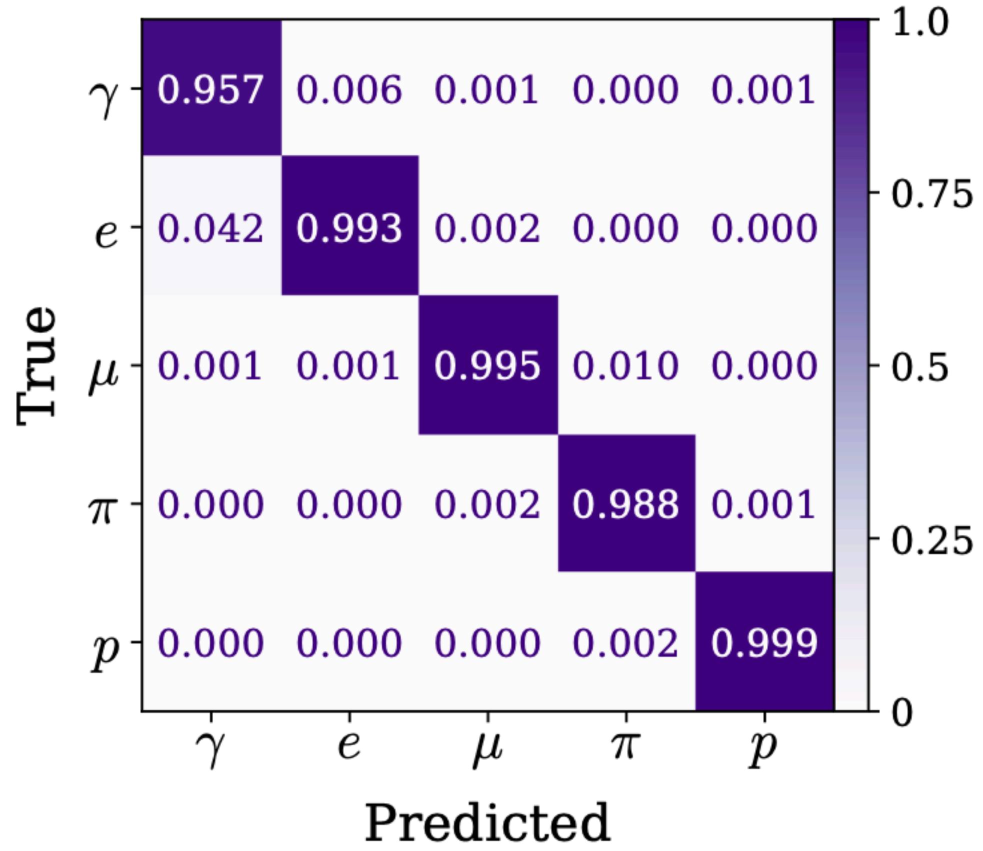
Towards Foundation Models





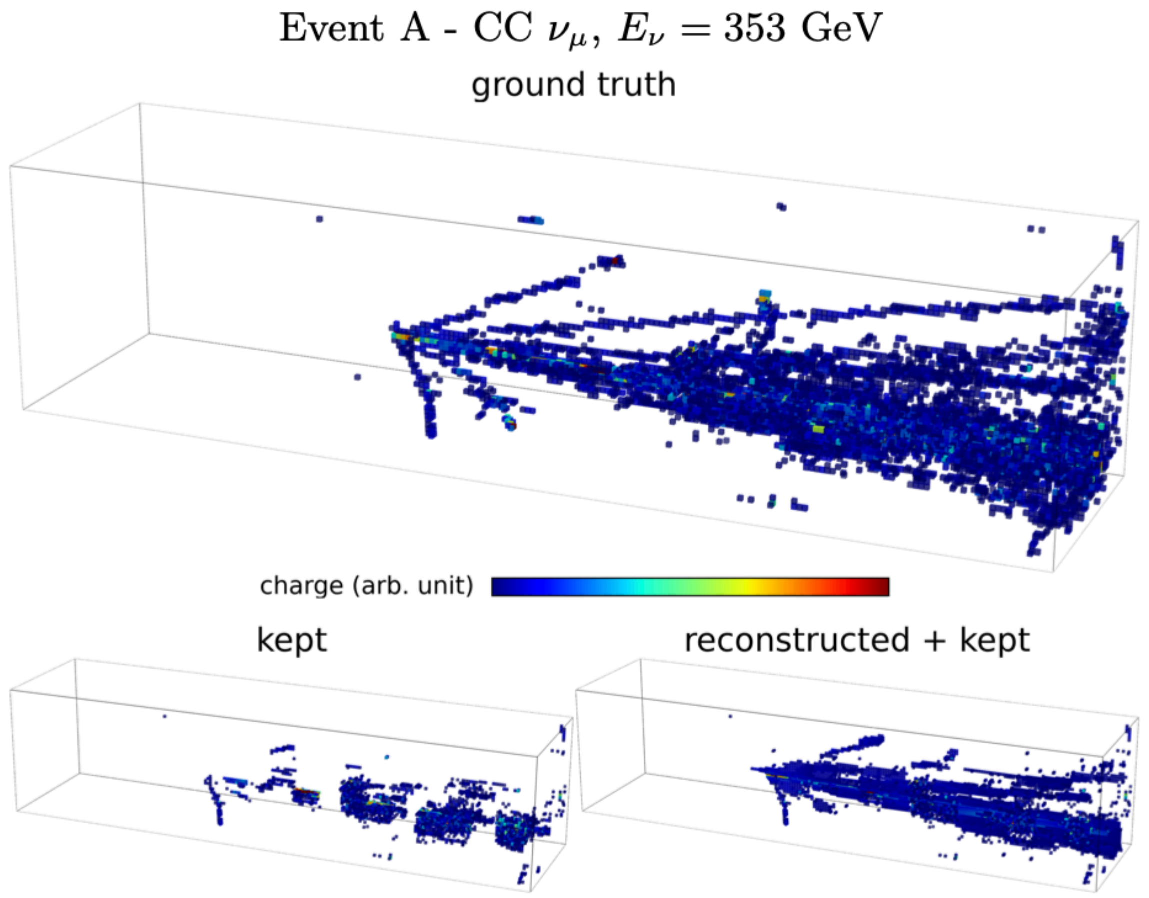
S. Young & K. Terao, *Panda: Self-distillation of Reusable Sensor-level Representations for High Energy Physics*, arXiv: 2512.01324 [hep-ex] (2025)

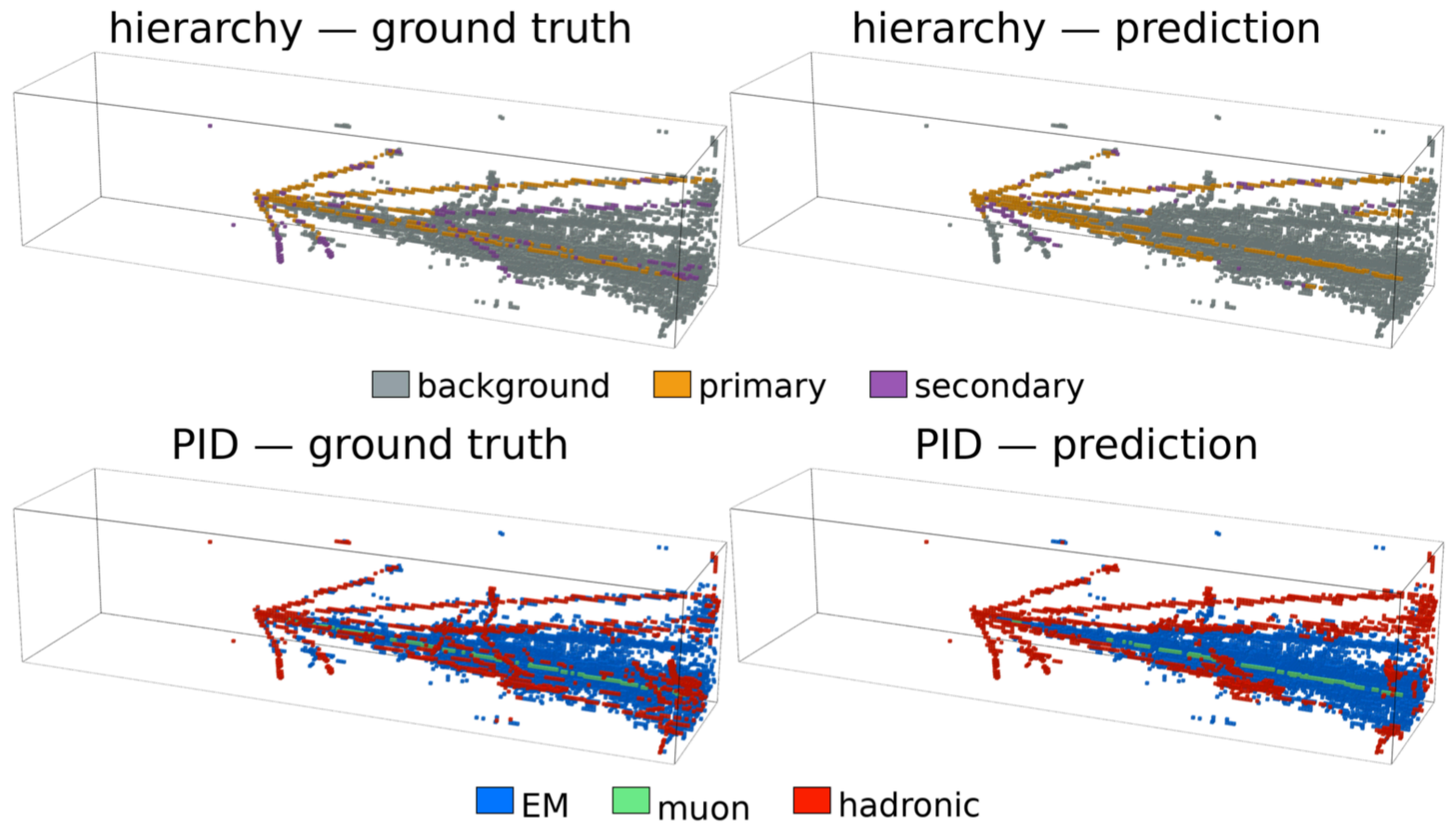




S. Young, et al., *Particle trajectory representation learning with masked point modeling*, Mach. Learn.: Sci. Technol. **7** 025023, (2026)

- Trained on TeV-scale neutrinos in a FASER-like detector.
- Very good performance when applied to other datasets:
  - SuperFGD-like detector
  - PILArNet LArTPC dataset

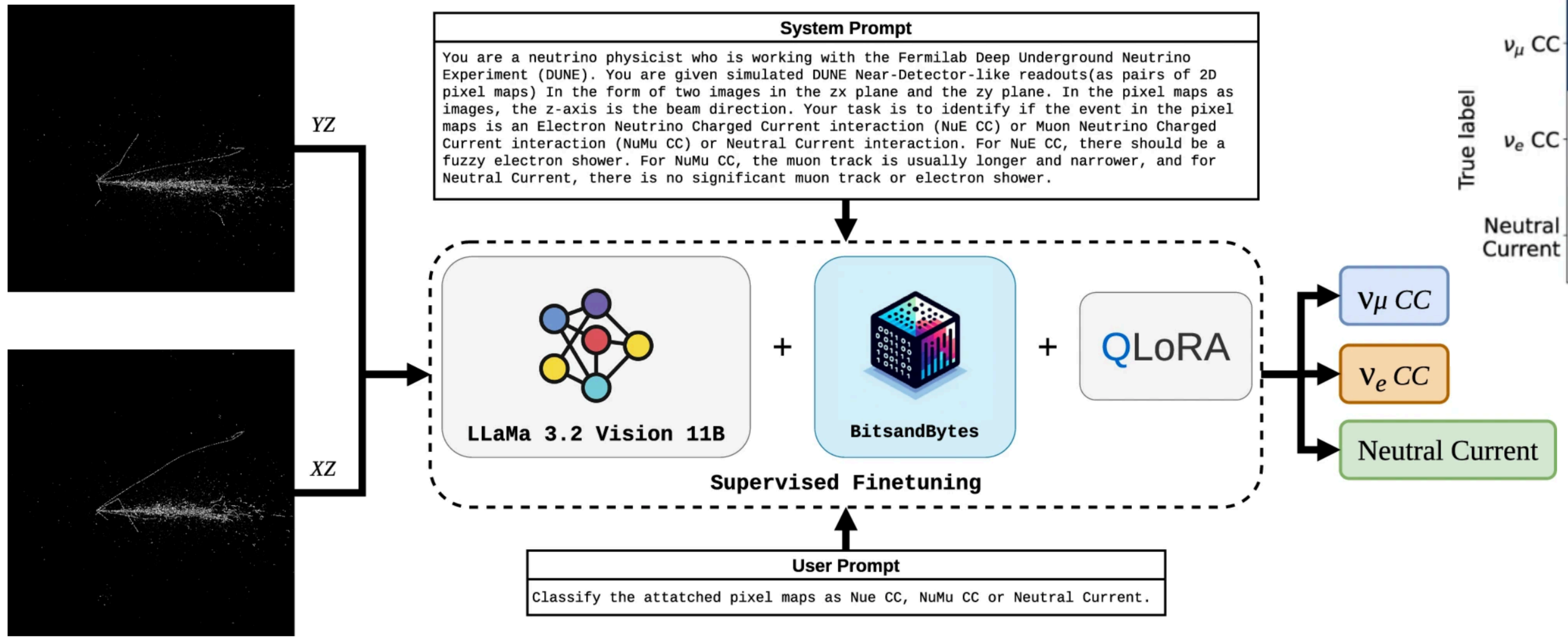


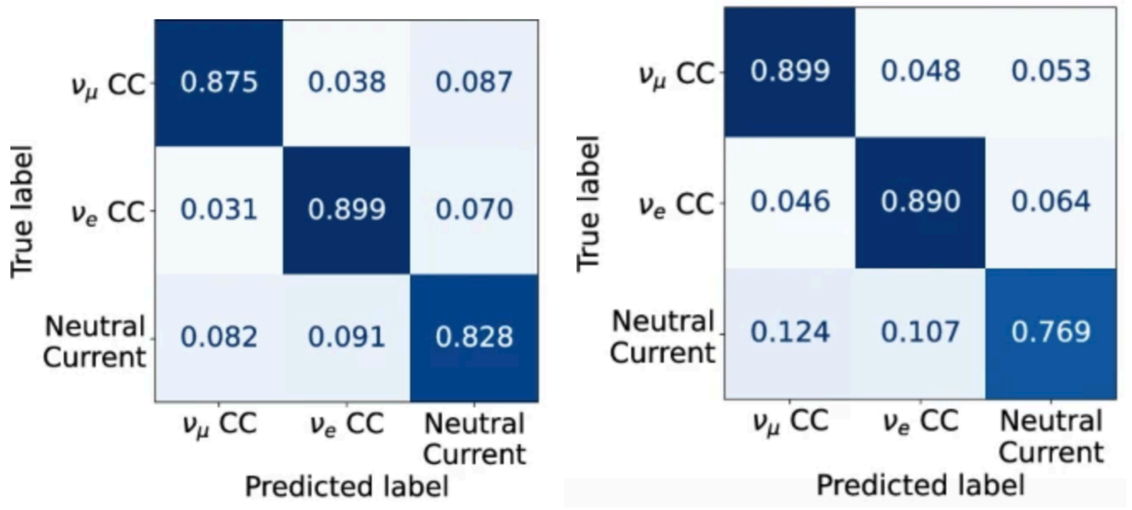


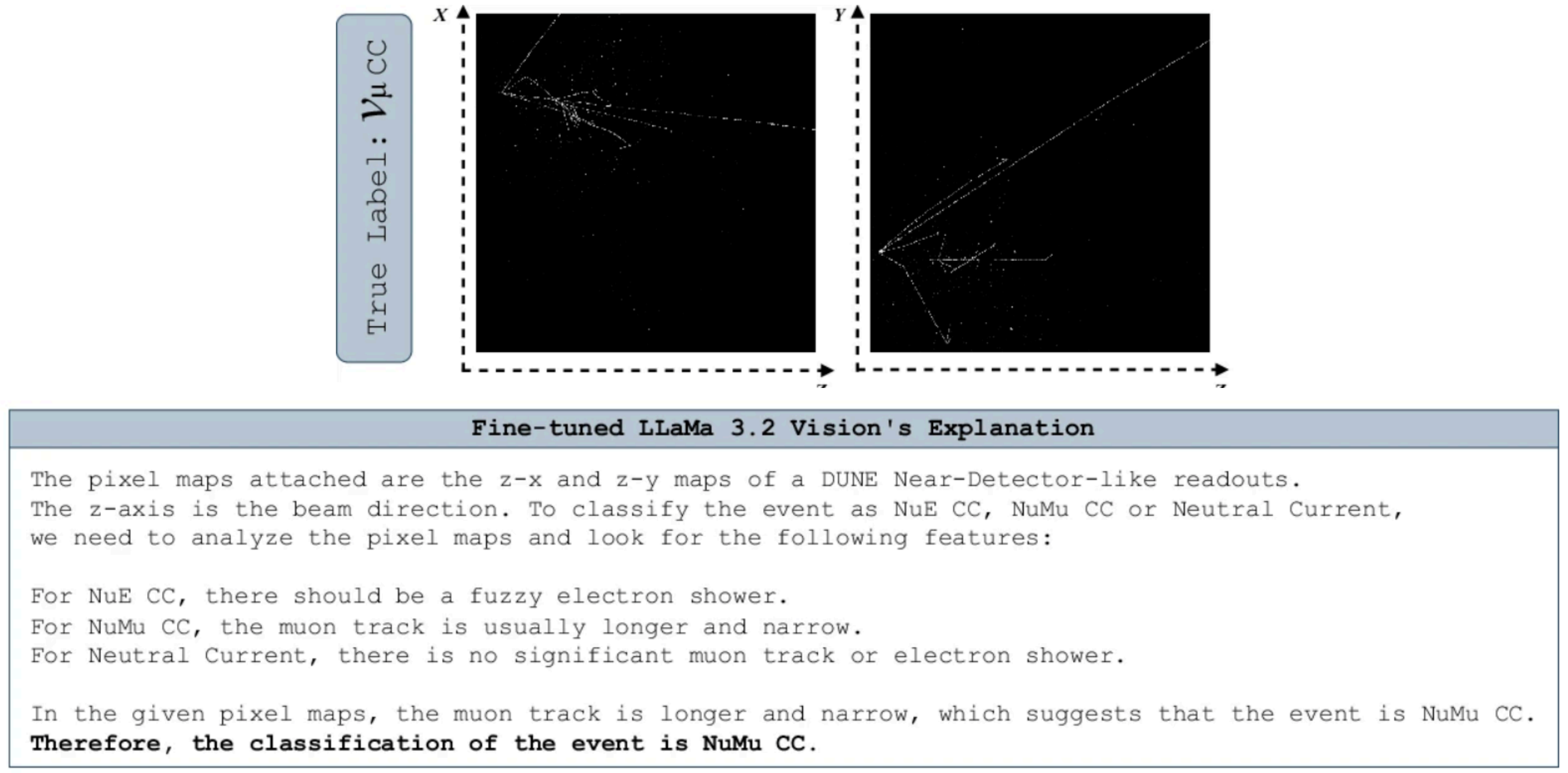
S. Alonso Monsalve, et al., *Towards foundation-style models for energy-frontier heterogeneous neutrino detectors via self-supervised pre-training*, arXiv: 2604.07037 [hep-ex] (2026)

LLMs

- Also starting to see use of LLMs and Multi-modal models
  - Can be fine-tuned for specific tasks







D. Sagar, et al., *Adapting vision-language models for neutrino event classification in high-energy physics*, Communications Physics **9,** 186 (2026)

- Also starting to see experiment specific models
  - For example, DUNEGPT that aggregates information across numerous repositories

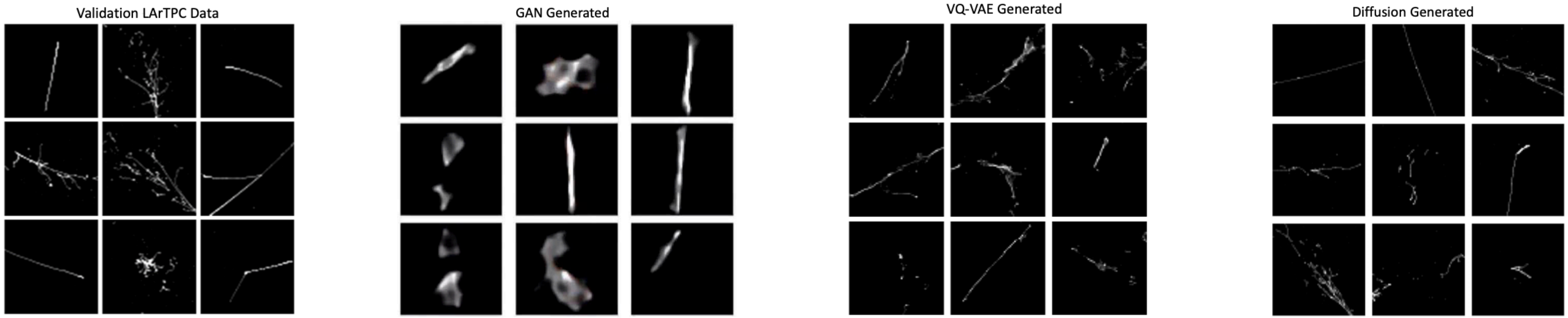
Simulation and Generative Models

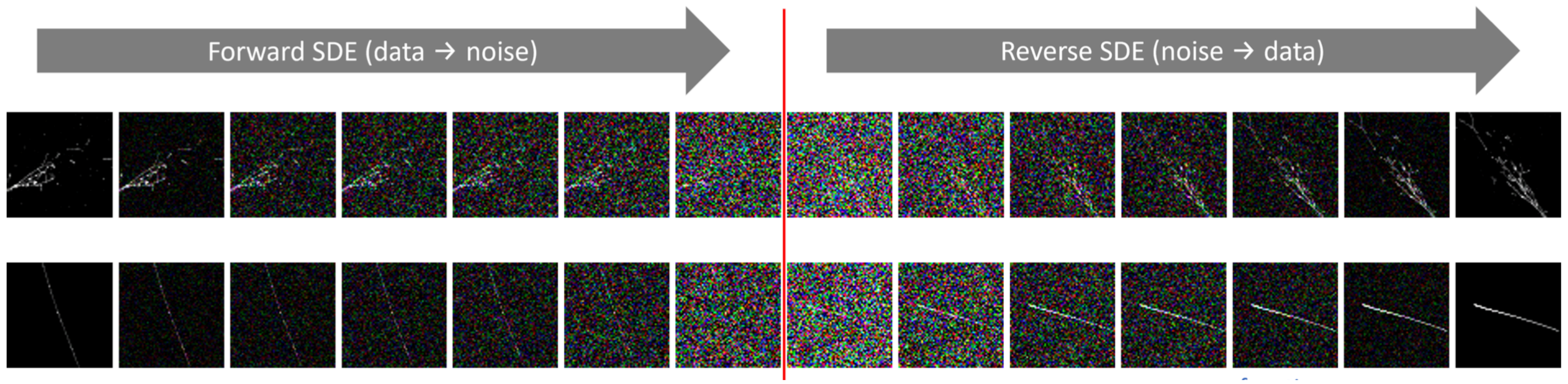
Simulation and Generative Models

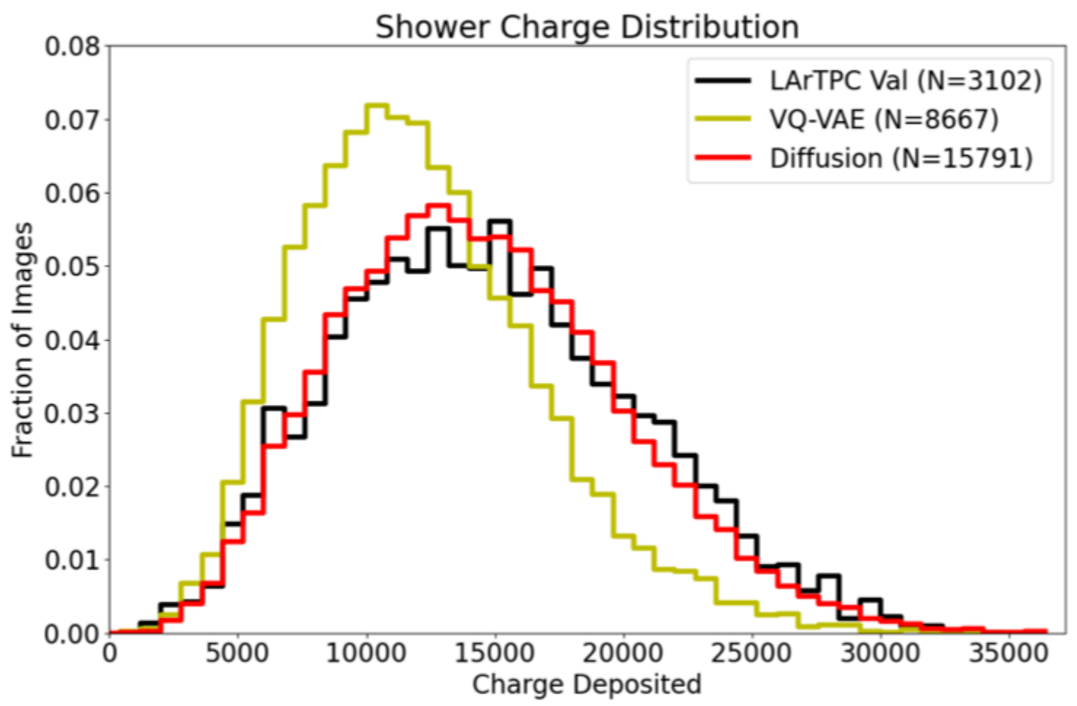
- Generative models naturally lend themselves to simulations
  - As the name suggests, these models produce data
  - Various algorithms exist:
    - Generative Adversarial Networks (GANs)
    - Variational auto-encoders (VAEs)
    - Diffusion models
- Will also give examples of differentiable simulations
  - These allow optimisation / calibration of physics parameters by fitting to data
  - Values are jointly optimised with the correct correlations

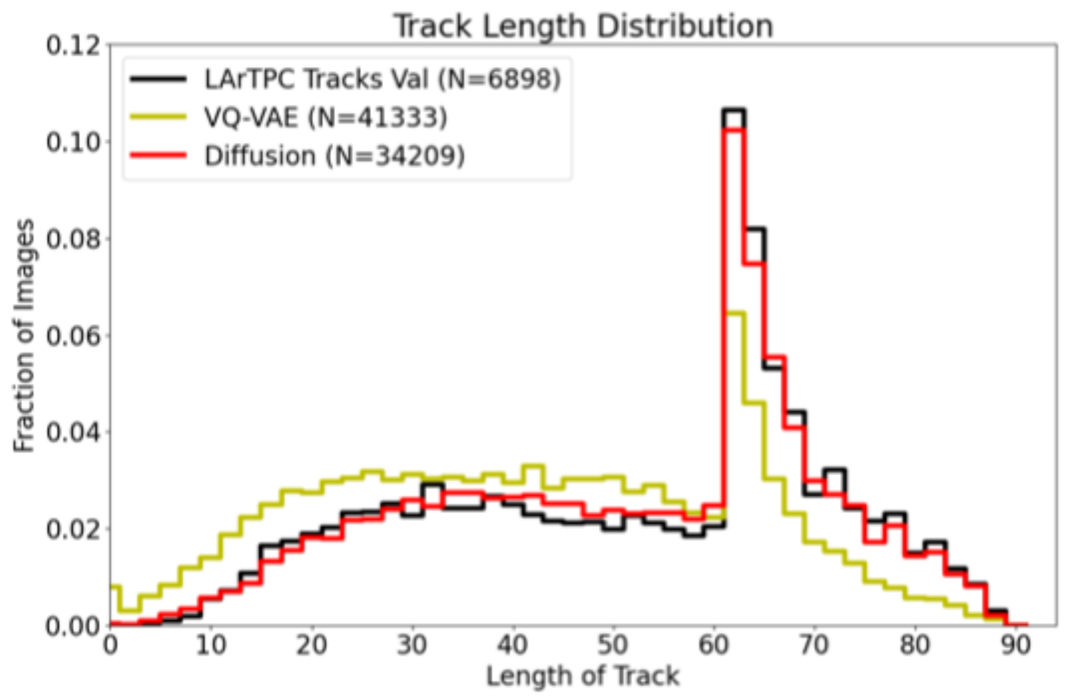
Generative Simulations

- Various techniques have been tried… GANs, VAEs, Diffusion models



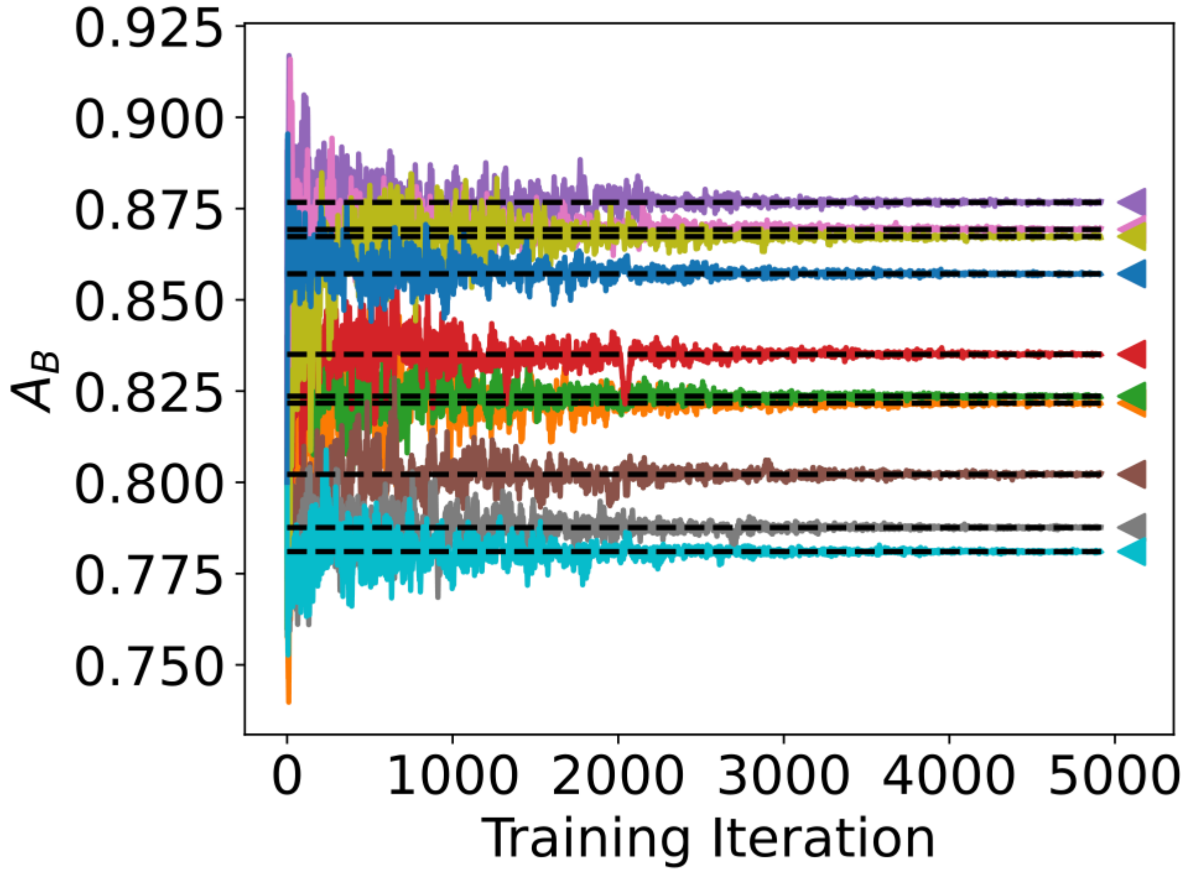


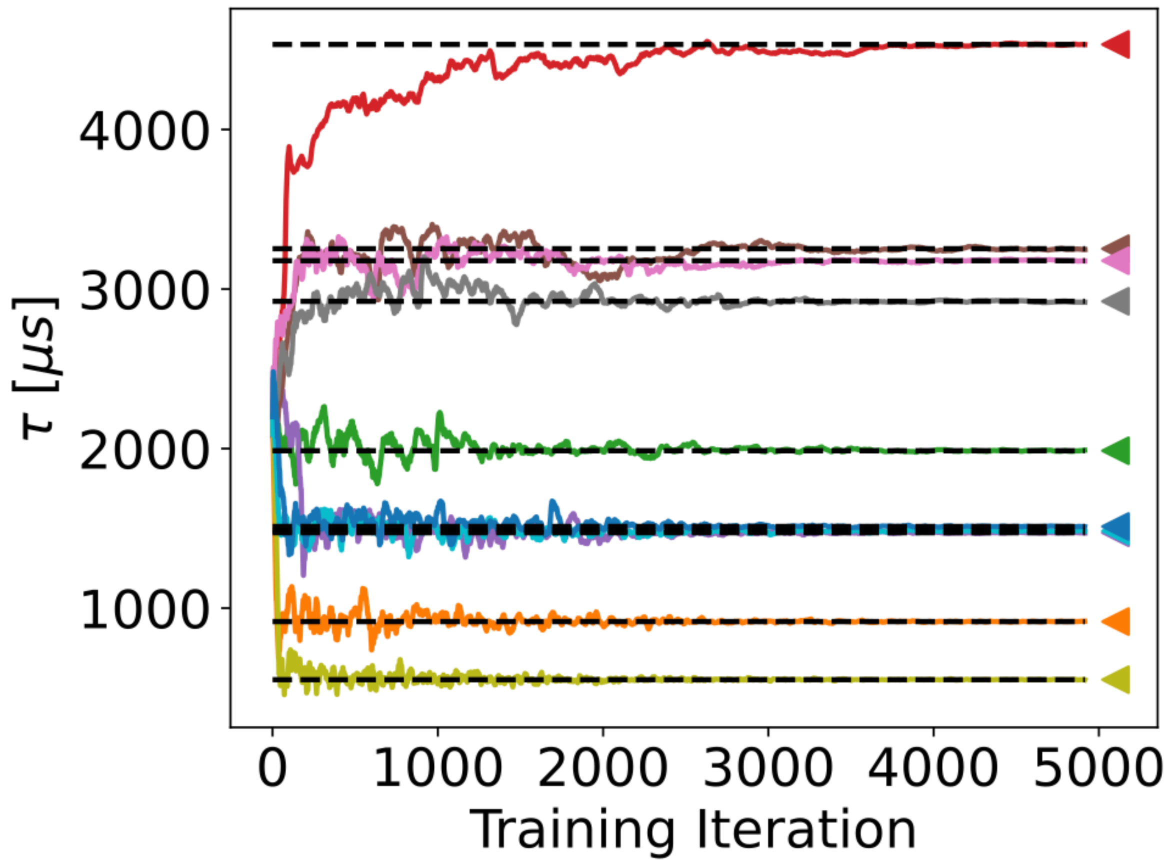


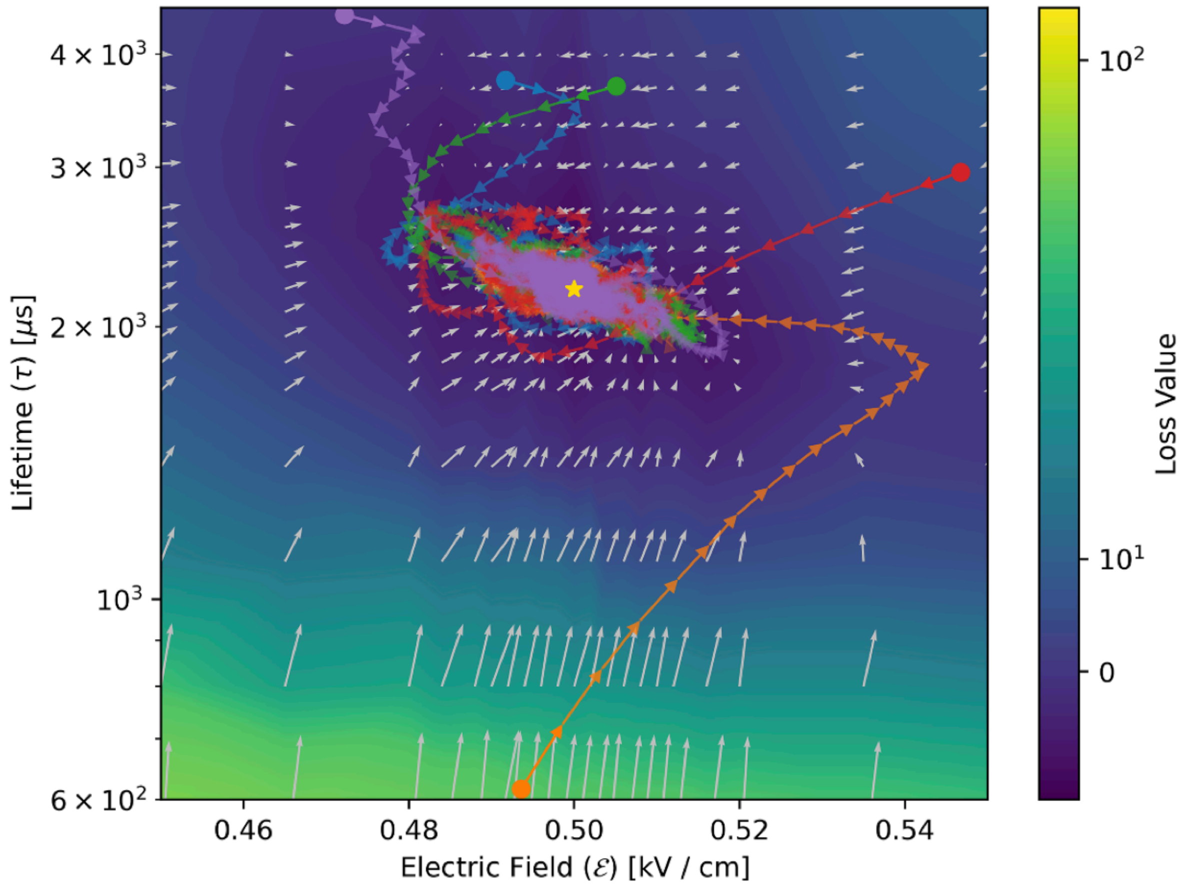


Z. Imani, et al., *Score-based diffusion models for generating liquid argon time projection chamber images*, Phys. Rev. D **109** 7, 072011 (2024)
Z. Imani, Generative Modeling for LArTPC Images, NPML 2024 (2024)

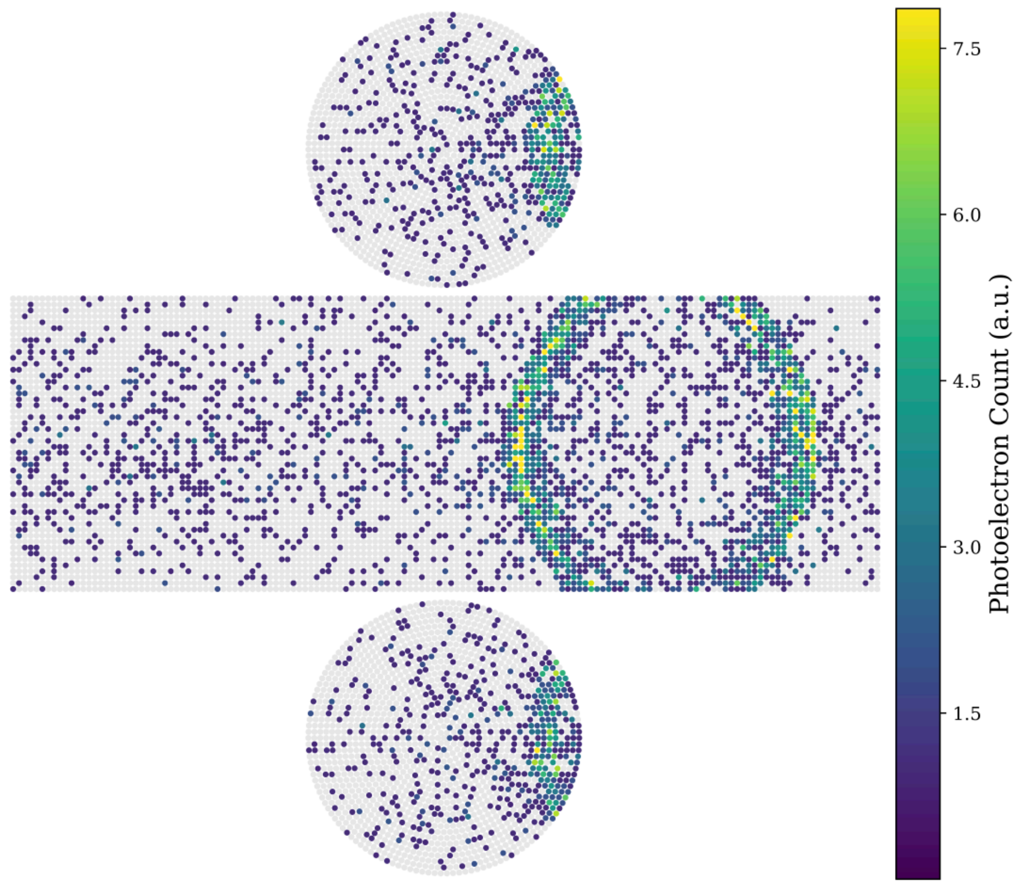
Differentiable Simulations

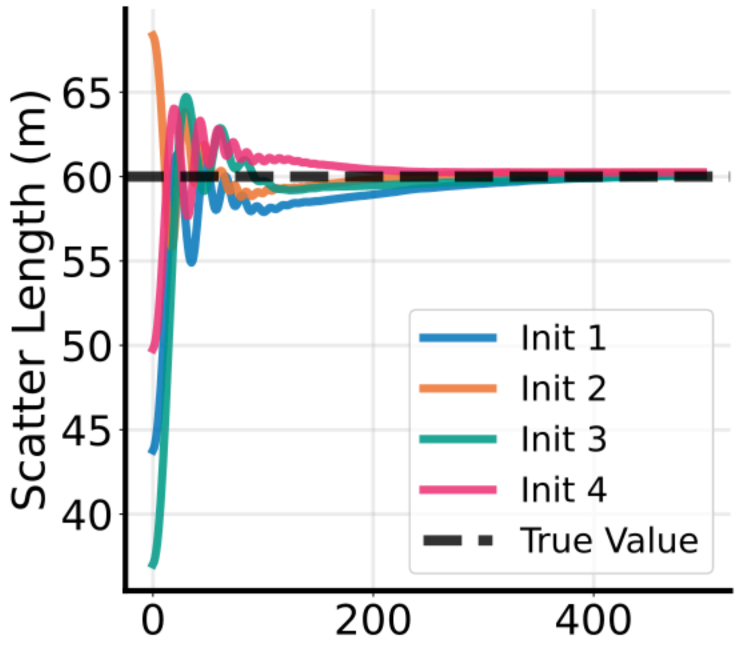


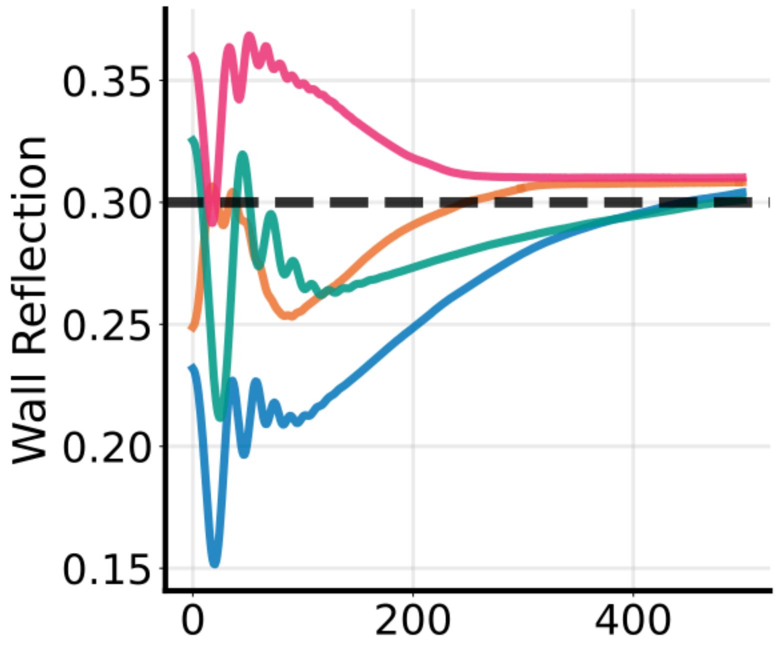


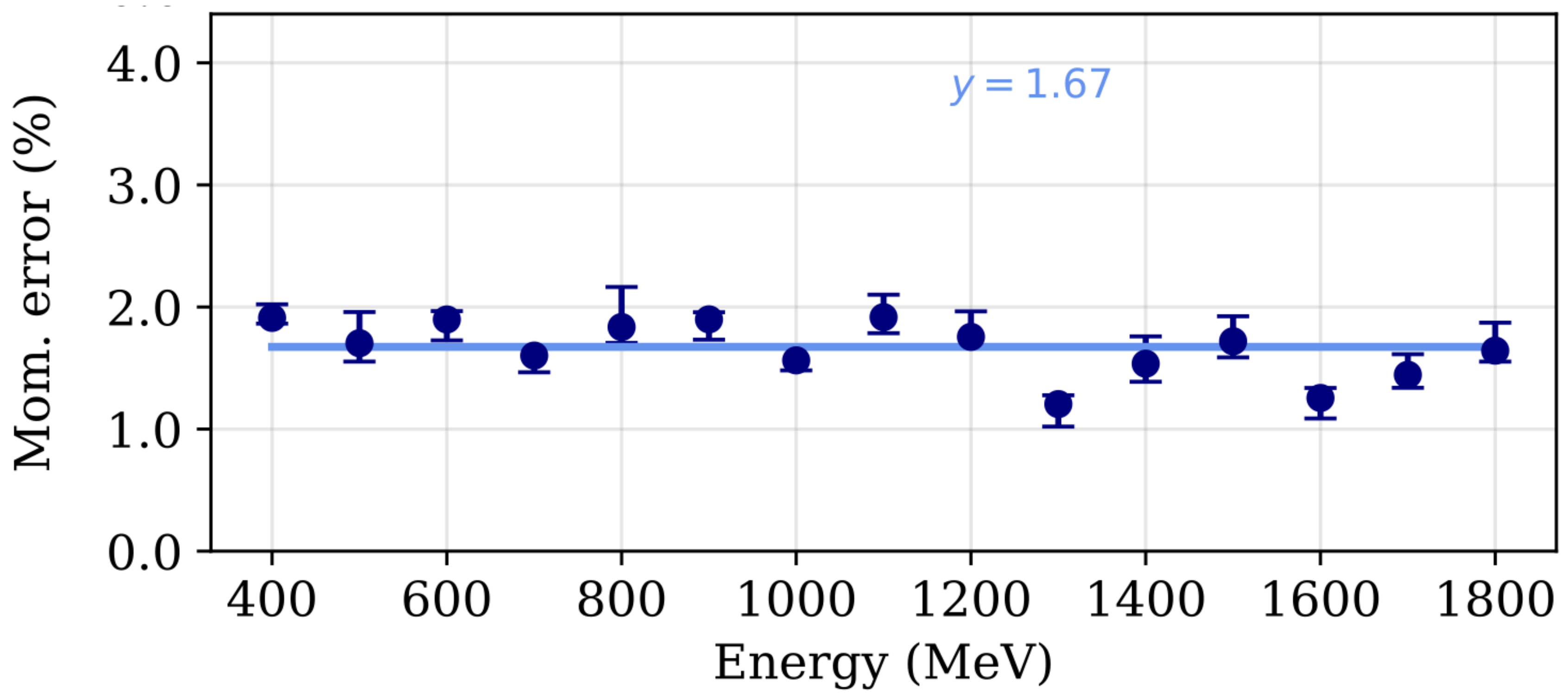


S. Gasiorowski, et al., *Differentiable simulation of a liquid argon time projection chamber*, Mach. Learn.: Sci. Technol. **5** 025012 (2024)









O. Alterkait, et al., *End-to-end Differentiable Calibration and Reconstruction for Optical Particle Detectors*, arXiv:2602.24129 [hep-ex] (2026)

Looking Ahead

Genesis Mission

- Recently announced US Government initiative
- Covers three main areas
  - Energy
  - Discovery Science (we fit in here)
  - National Security
- Components:
  - Collaboration between universities, DOE laboratories and industry
  - Big data and powerful computing resources
  - Training future leaders of AI/ML

Genesis Mission

A National Mission to Accelerate Science Through Artificial Intelligence

Genesis Mission will develop an integrated platform that connects the world's best supercomputers, experimental facilities, AI systems, and unique datasets across every major scientific domain to double the productivity and impact of American research and innovation within a decade.

https://genesis.energy.gov/

Genesis Mission

- Those components look familiar…
  - Many HEP experiments already tick these boxes!
  - This will be especially true for DUNE as a US-based mega-project
- Specific part of the Genesis Mission dedicated to DUNE

> **B. AI Accelerated DUNE Science (HEP)**
> Develop AI methods that significantly speed up and enhance the DUNE science program, reducing the time needed for the collaboration to publish neutrino oscillation measurements, significantly improving the sensitivity to neutrinos from core-collapse supernova, and developing new flagship measurements that will enhance the DUNE science goals.

- Awaiting the outcome of the first round of funding proposals in mid July

Where are we going?

- Always a dangerous game to make predictions*
- Foundation Models
  - One model to rule them all
  - Cross-experiment models
- Ensemble of expert models
- Prompt-driven agentic AI workflows
  - "Please make a full workflow for a numu disappearance analysis"

* such as England to win the World Cup

Where are we going?

- End-to-end differentiable analyses:
  - event generation and detector simulation
  - reconstruction, calibration and analysis

- Cross-collaboration cooperation [1]
  - Large open datasets (MicroBooNE, PILArNet, …)
  - Shared trained models and data challenges (e.g. IceCube Kaggle contest)
- HEP Community White Paper on these topics from earlier this year [2]
- However, don’t forget environmental impact concerns

[1] For example, the Doraemon project
[2] T.K. Aarrestad, et al., *Building an AI-native Research Ecosystem for Experimental Particle Physics: A Community Vision*, arXiv: 2602.17582 (2026)

Summary

- Many great examples of AI/ML have emerged over the last 10 years
  - It is here to stay and will likely become a key part of all stages of experiments
- So much more I couldn't cover here
  - More use-cases of the discussed architectures
  - Different topics such as uncertainty quantification and simulation-based inference
  - Robustness and "looking inside the black box"
  - See backup slides for some more examples of different methods
- Who knows where we'll be in another 10 years time?

Exciting Times Ahead!

Any Questions?

Backup: Simulation-based Inference

- Assume we have some simulation that maps physics parameters $\theta \rightarrow x$
  - Typically we don't know the function $p(x|\theta)$
  - However, for Bayesian inference we need $p(\theta|x) \propto p(x|\theta)p(\theta)$
- Solution: use some sort of neural network to learn $p(\theta|x)$ from the simulation
  - Can use this for inferring parameters from data and fast simulation
- Some examples:

K. Tame-Narvaez, et al., *Simulation-Based Inference for Neutrino Interaction Model Parameter Tuning*, NeurIPS 2025, arXiv: 2510.07454 [hep-ph] (2025)

A. Gavrikov, et al., *Simulation-based inference for precision neutrino physics through neural Monte Carlo tuning*, Commun. Phys. **9** 1, 63 (2026)

J. P. Woodward, et al., *A Simulation-Based Inference Evaluation of Tension Between MicroBooNE and MiniBooNE Results in a 3+1 Sterile Neutrino Global Fit*, arXiv: 2603.15322 [hep-ex] (2026)

Backup: Surrogate Models

- Surrogate models aim to replace a simulation with a neural network
  - In this case we want to learn p(x|θ)
- Typically used for creating fast (but often approximate) simulations

W. Mu, et al., *Photon detection probability prediction using one-dimensional generative neural network*, Mach. Learn.: Sci. Technol. **3** 015033 (2022)
M. Lei, et al., *Implicit Neural Representation as a Differentiable Surrogate for Photon Propagation in a Monolithic Neutrino Detector*, arXiv: 2211.01505 [physics.ins-det] (2022)
P. Pilar, et al., *A differentiable surrogate model for the generation of radio pulses from in-ice neutrino interactions*, Mach. Learn.: Sci. Technol. 7 025017 (2026)

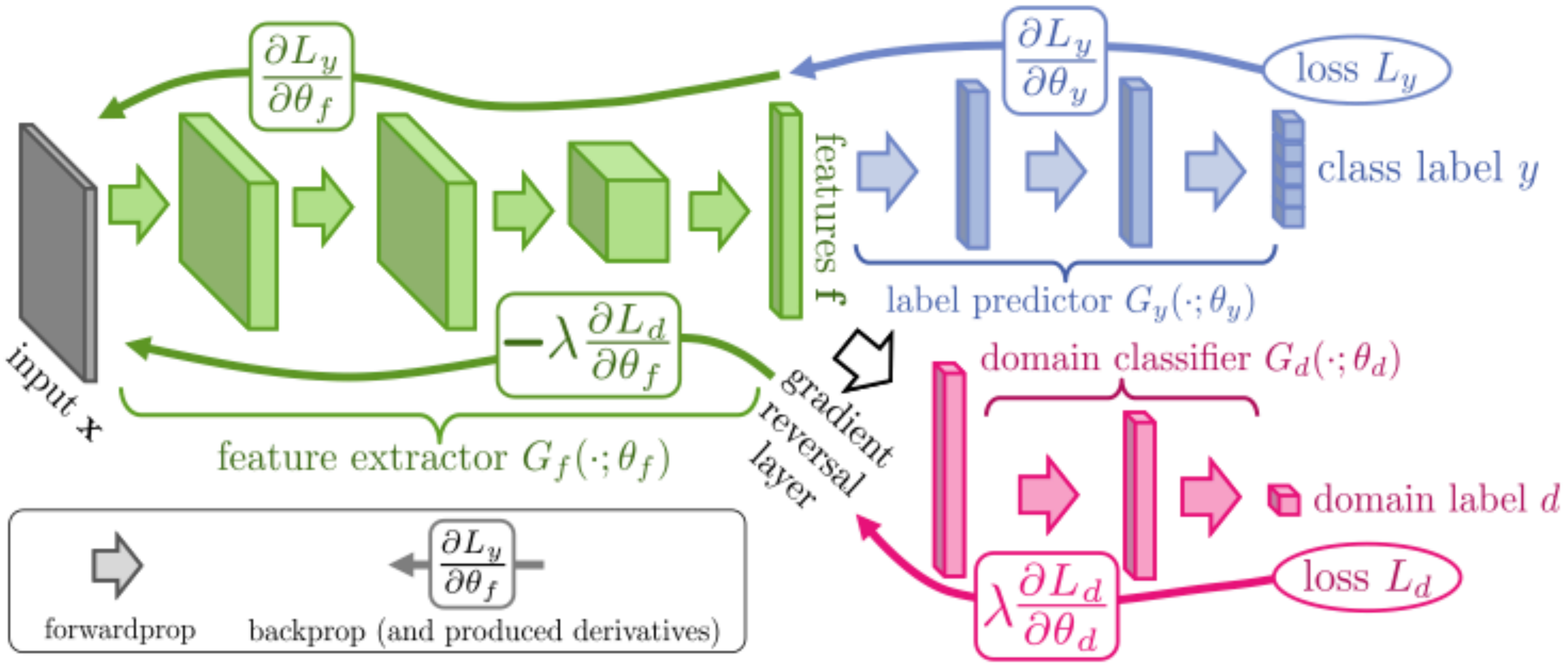
Backup: Transfer Learning

- Transfer learning can help reduce the computing burden of model training
  - Can re-use a pre-trained generic model and fine-tune it for your specific use case
- Often need a significantly smaller training sample
- More environmentally friendly!
- Often used in foundation models / LLMs, but I just give two specific examples:

A. Chappell and L. H. Whitehead, *Application of Transfer Learning to Neutrino Interaction Classification*, Eur. Phys. J. C **82** 12, 1099 (2022)
J. L. Bonilla, et al., *Transfer learning for neutrino scattering: Domain adaptation with generative adversarial networks*, Phys. Rev. D **113**, 053001 (2026)

Backup: Robustness: DANNs

- Domain Adversarial Neural Networks
  - Classifier output to predict if we have domain A (e.g. MC) or domain B (e.g. data)
  - Helps to learn only the generic features of the MC, not fine details of the model that could potentially bias results



Y. Ganin, et al., *Domain-Adversarial Training of Neural Networks*, JMLR **17** 59, p1-35 (2016)

Backup: Robustness: Occlusion Maps

- Occlusion tests allow us to investigate what the CNN is “looking” at
  - Or probe what it finds most important in the images
- Mask patches of the images and see how the classification score(s) change

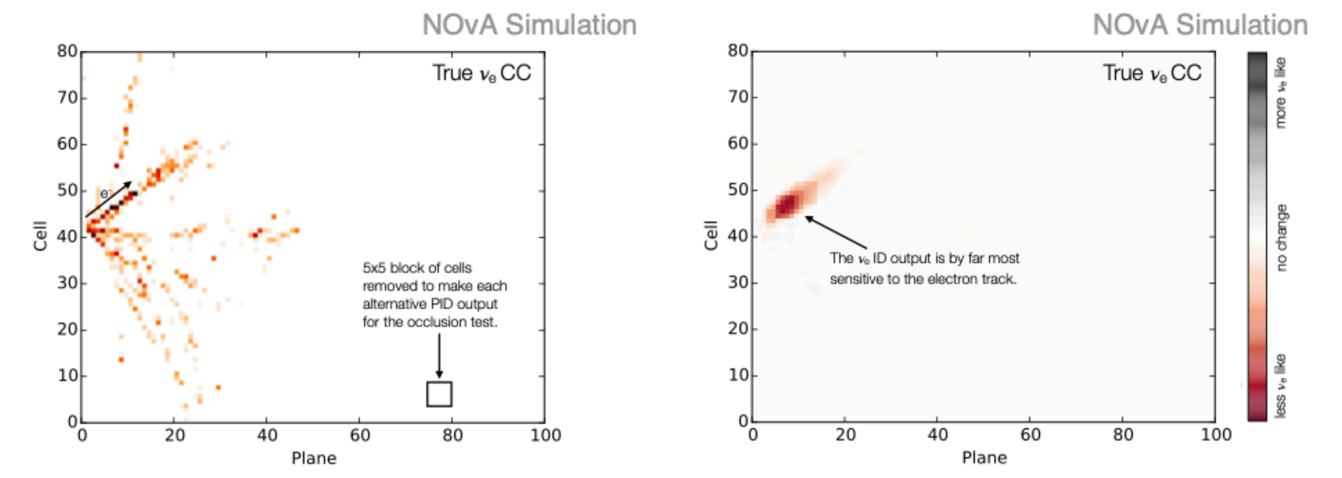
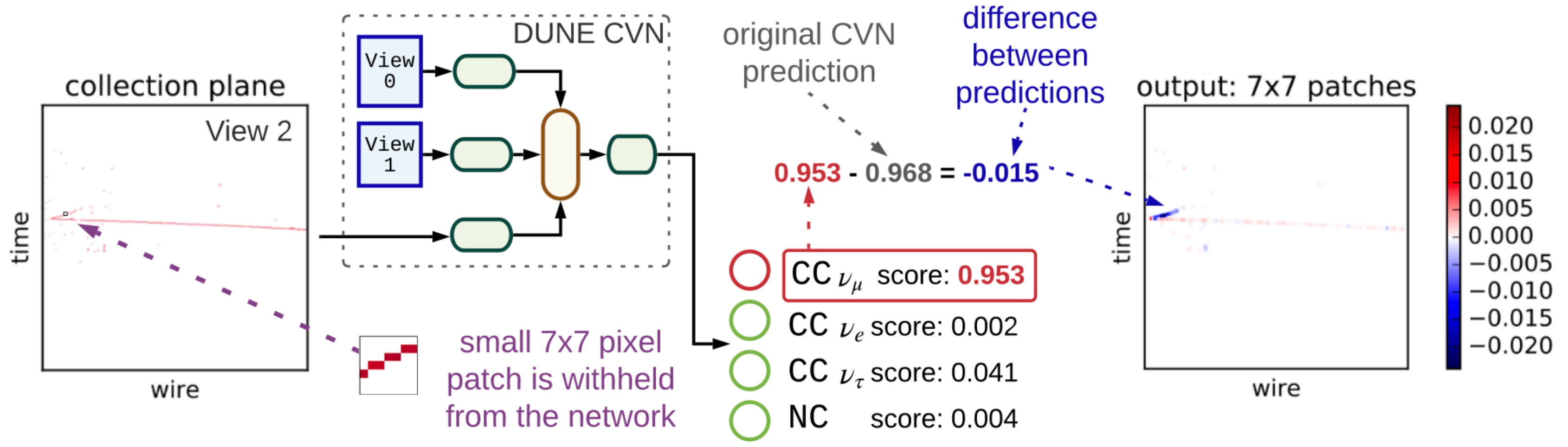
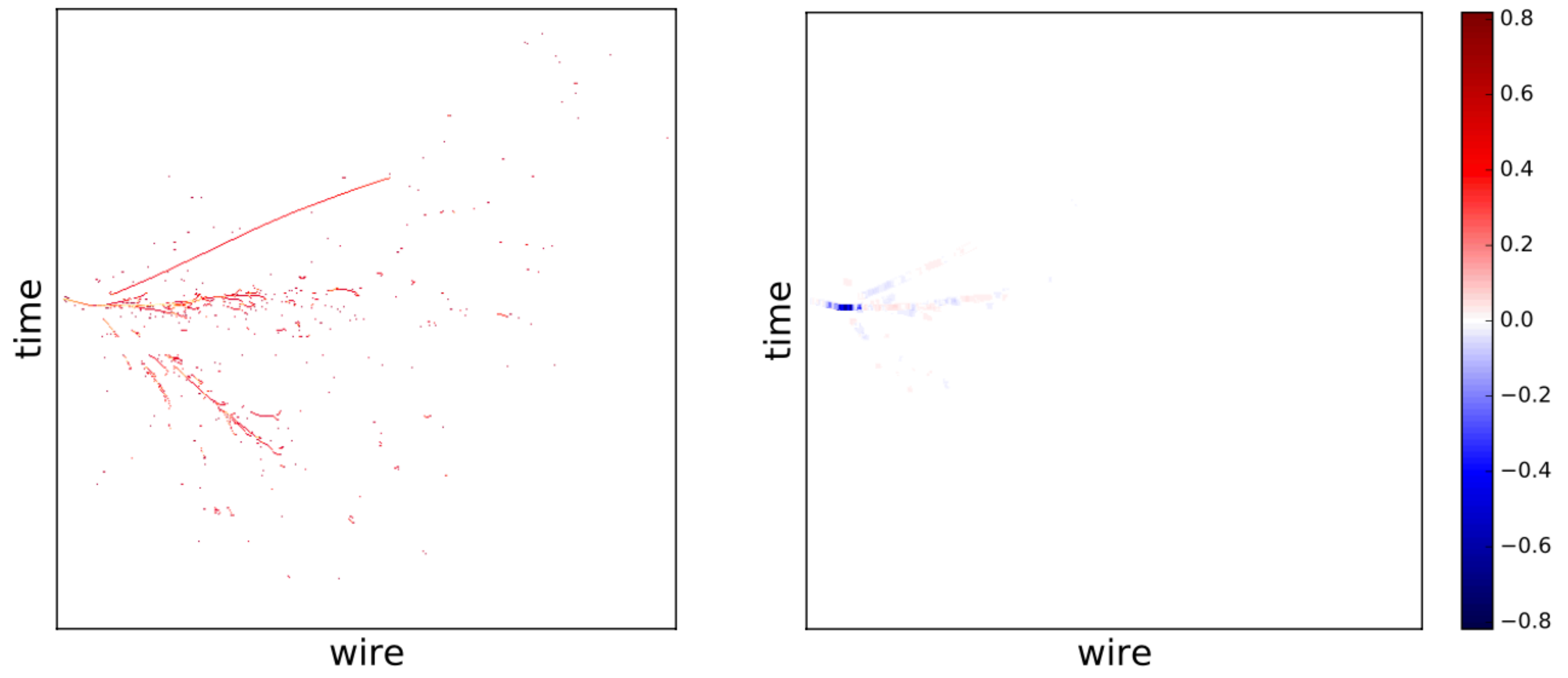


Fig. 20. An occlusion test demonstrating the most salient parts of an input image to the decision made by the NOvA CNN [8]. (left) A single view of a true ν_e-CC interaction. This interaction consists of a single electromagnetic shower from a primary electron along with several track-like objects from the hadrons produced by the nucleus. (right) The change in the ν_e-CC score as a function of the location of a 5×5 occluded region. Figures courtesy of the NOvA collaboration.



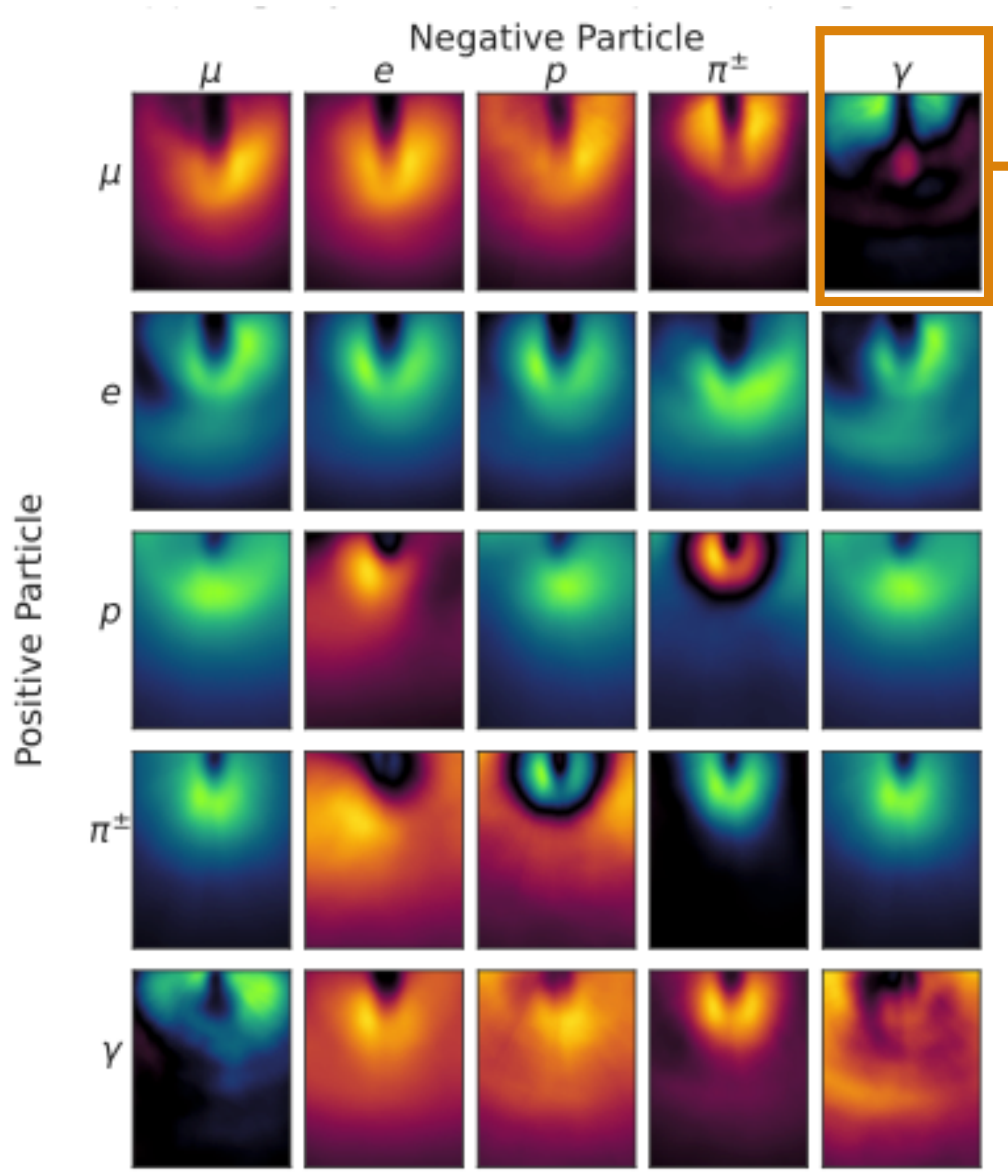


Actual flavour	Subfigure	CVN scores			
		CC ν_μ	CC ν_e	CC ν_τ	NC
CC ν_e	3.16a	0.0009	**0.9184**	0.0090	0.0717
	3.16b (largest difference)	0.0015	0.1003	0.0098	**0.8884**

S. Alonso Monsalve, *Novel usage of deep learning and high-performance computing in long-baseline neutrino oscillation experiments*, PhD Thesis, Carlos III U. Madrid (2021)

Backup: Robustness: Saliency Maps

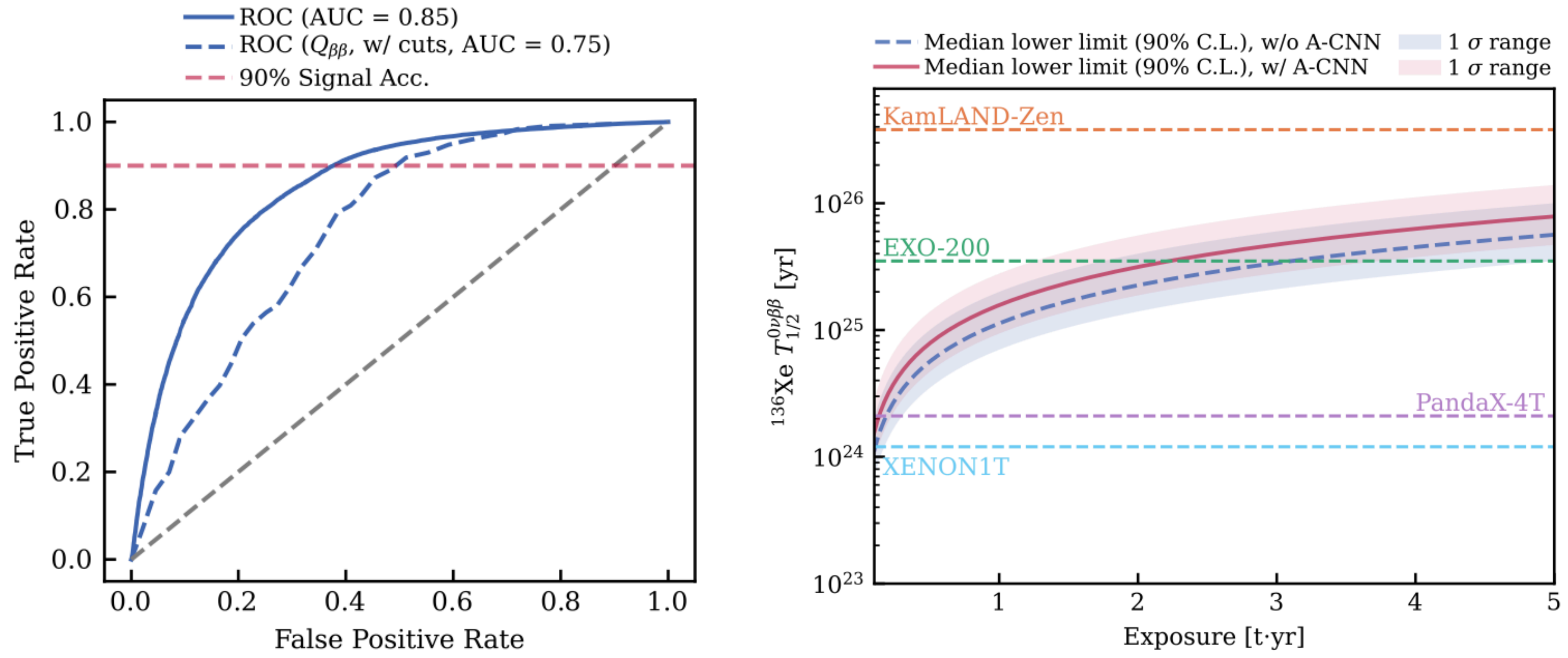
- Saliency maps show the gradient for each pixel in the images
  - A change in pixels with a large gradient would have significant impact on the results



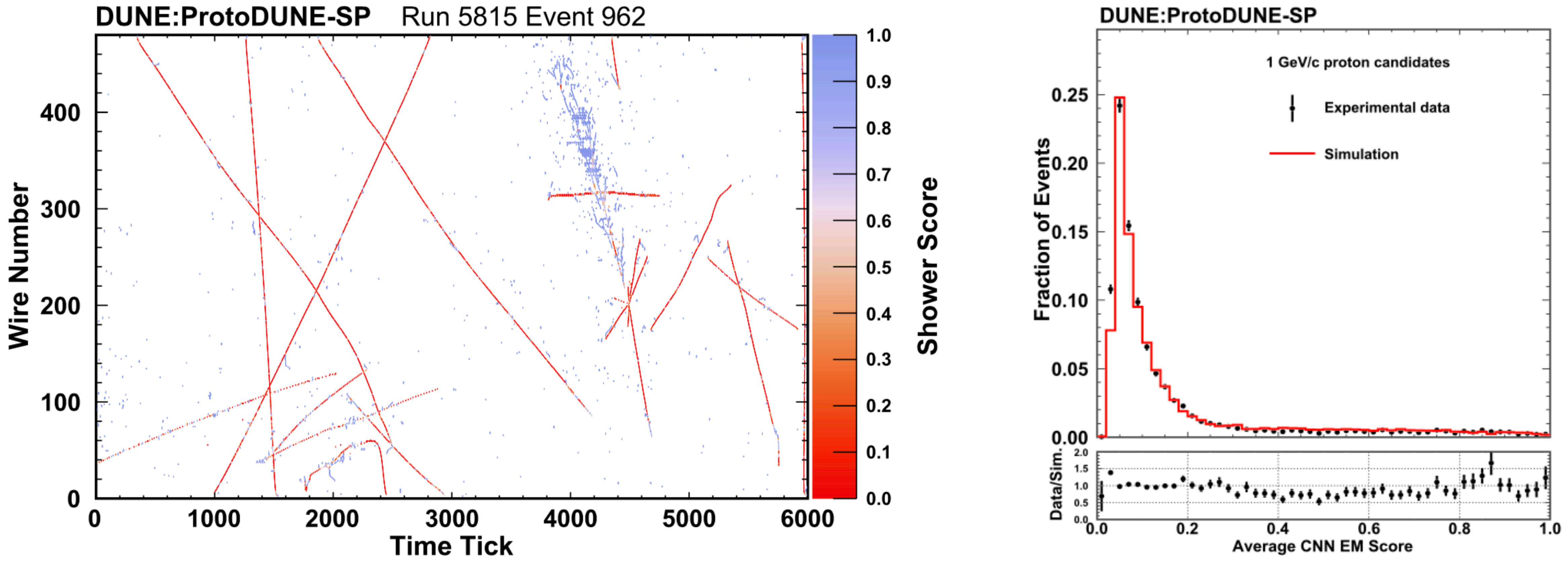
E.g: To make a photon look more muon like:
- Increase pixel values to make a straight line
- Remove peripheral hits

A. Shmakov, et al., *Interpretable Joint Event-Particle Reconstruction for Neutrino Physics at NOvA with Sparse CNNs and Transformers*, arXiv: 2303.06201 [cs.LG] (2023)

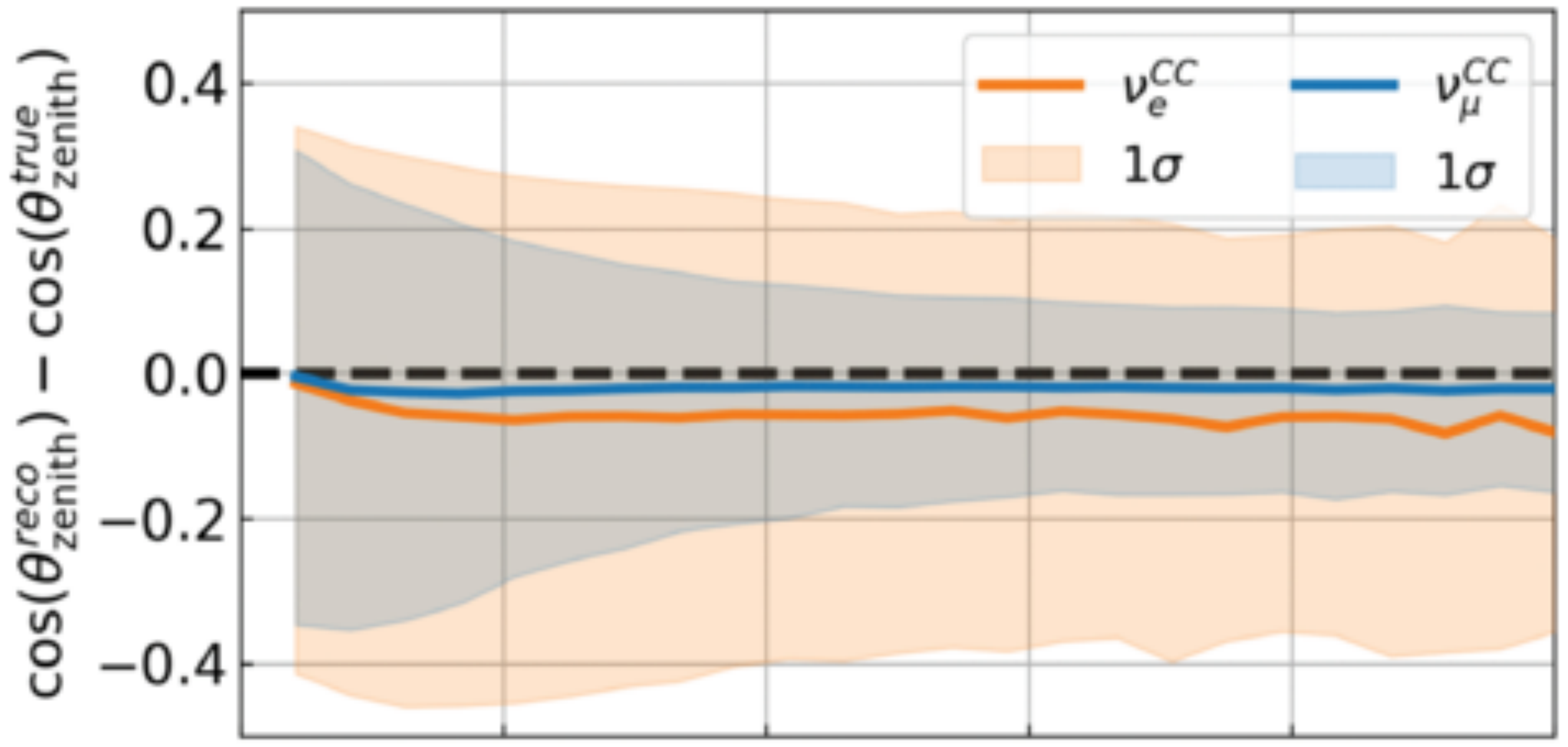
Backup: CNN use-cases III



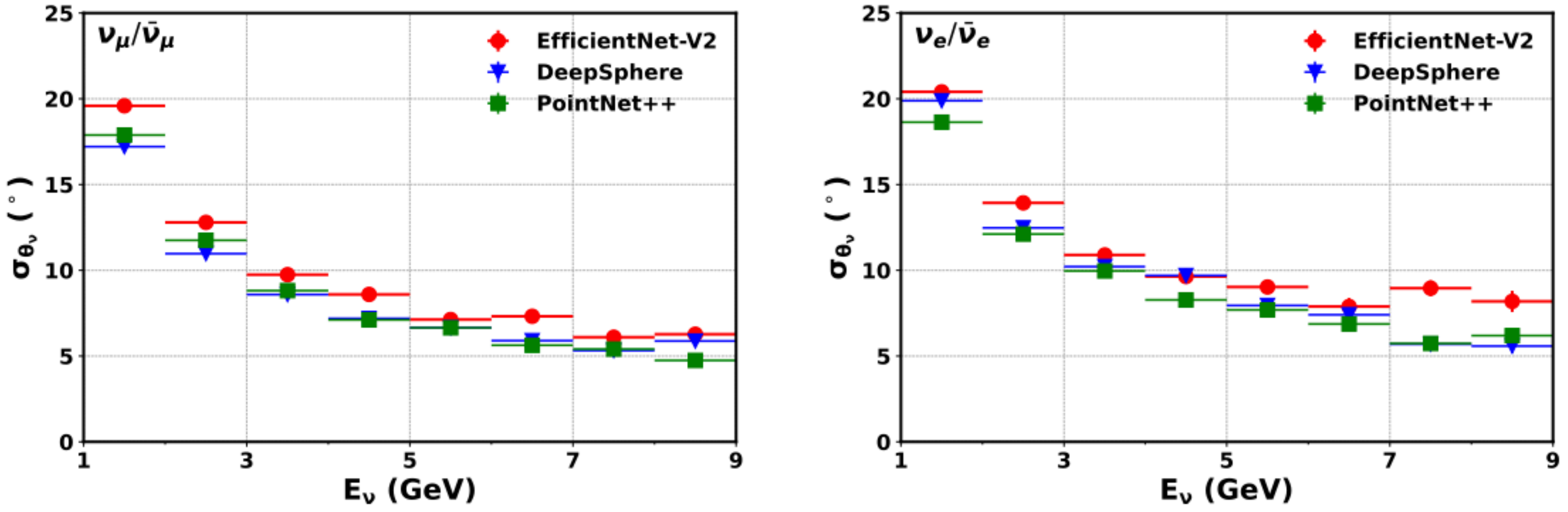
XENON Collaboration, *Enhancing Neutrinoless Double-Beta Decay Sensitivity of Liquid-Xenon Time Projection Chamber with Augmented Convolutional Neural Network*, arXiv: 2603.23549 [physics.ins-det] (2026)



DUNE Collaboration, *Separation of track- and shower-like energy deposits in ProtoDUNE-SP using a convolutional neural network*, Eur.Phys.J.C **82** 10, 903 (2022)

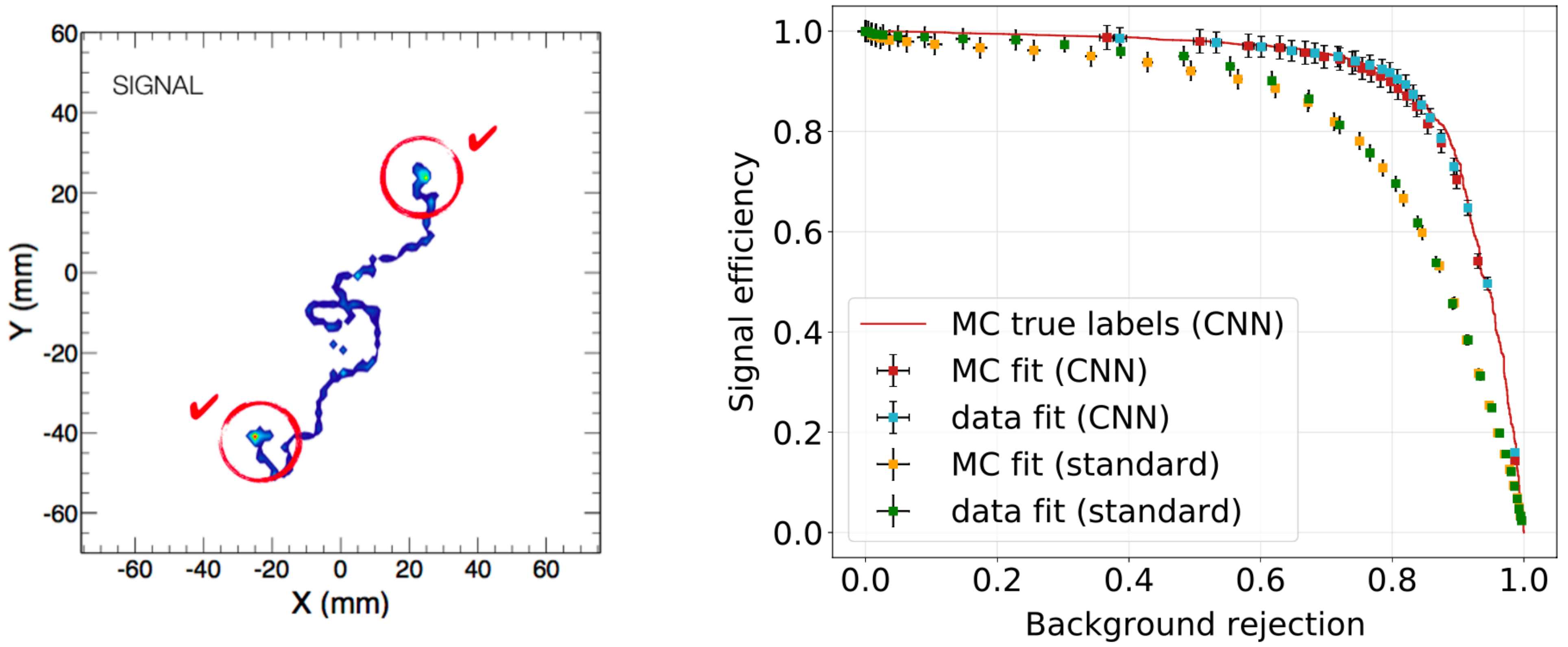


IceCube Collaboration, Measurement of Atmospheric Neutrino Oscillation Parameters Using Convolutional Neural Networks with 9.3 Years of Data in IceCube DeepCore, *Phys. Rev. Lett.* **134**, 091801 (2025)

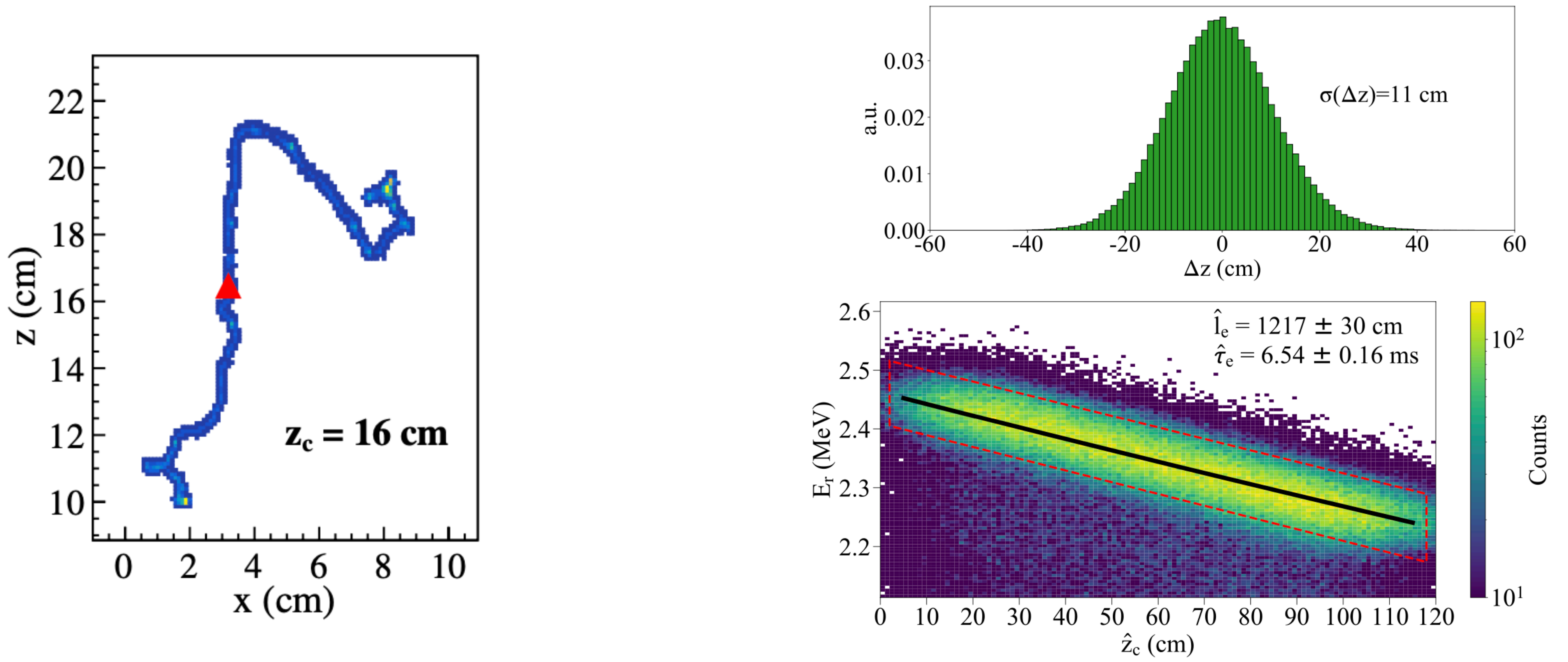


Z. Yang, et al., *First attempt of directionality reconstruction for atmospheric neutrinos in a large homogeneous liquid scintillator detector*, Phys. Rev. D 109, 052005 (2024)

Backup: CNN use-cases IV



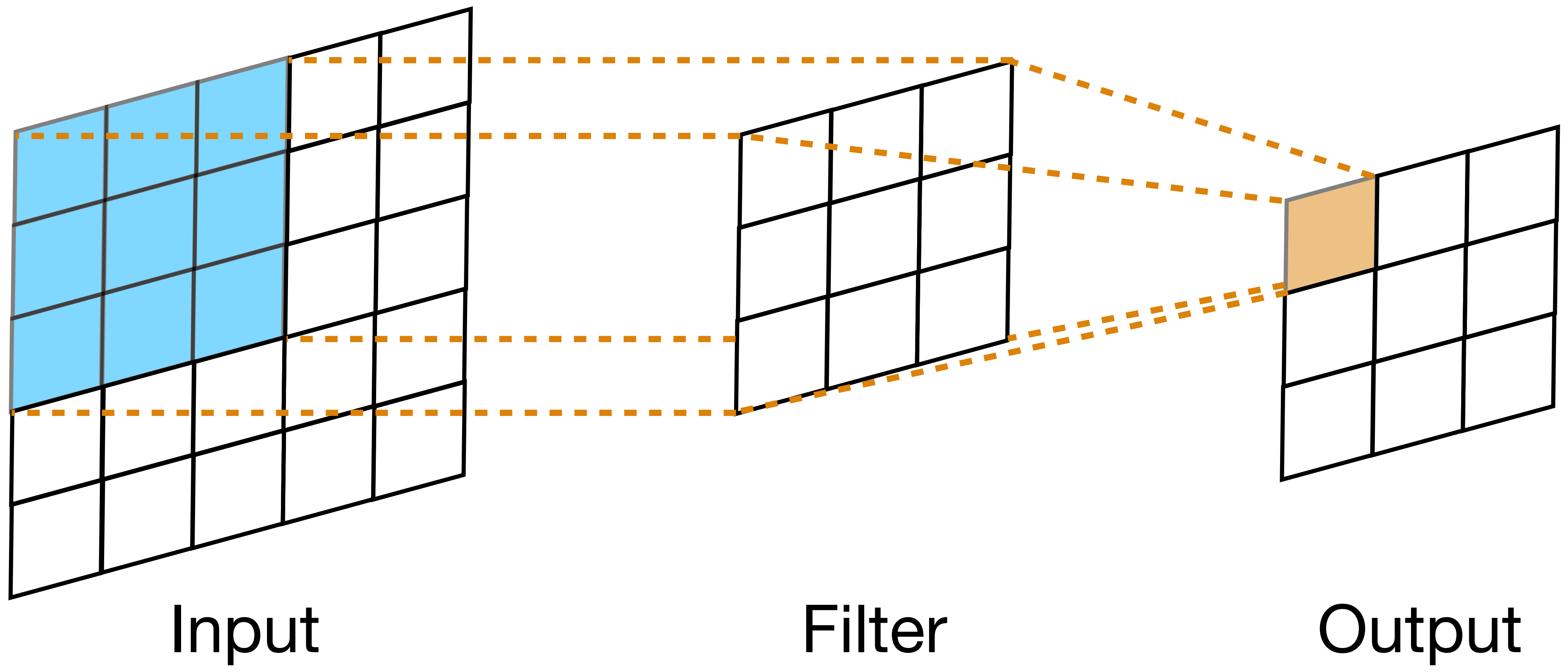
The NEXT Collaboration, *Demonstration of background rejection using deep convolutional neural networks in the NEXT experiment,* J. High Energ. Phys. **2021**, 189 (2021)



Li, T., et al., *Reconstruction of the event vertex in the PandaX-III experiment with convolution neural network*. J. High Energ. Phys. **2023**, 200 (2023)

How do CNNs work?

- CNNs themselves actually date back to 1989[1]
  - Started to dominate the image-processing field from 2012 with AlexNet[2]
- Effectively arrange neurons into 2D filters instead of linear layers (as in MLPs)
  - This filter is scanned across the input
  - At each point, perform an element-wise multiplication



[1] Y. Lecun, et al., Backpropagation Applied to Handwritten Zip Code Recognition, Neural Computation 1 (4): 541-551 (1989) https://doi.org/10.1162/neco.1989.1.4.541
[2] A. Krizhevsky, et al., ImageNet Classification with Deep Convolutional Neural Networks, Communications of the ACM. 60 (6): 84–90, 2012

Transformers

- The main feature of the transformer is the attention mechanism
  - It is a very simple architecture
  - Initial encoding of input sequence data
  - Three linear projections
  - Dot-product (typically) attent
  - Matrix multiplication
  - Two-layer feed-forward netw

$q_0.k_0$	$q_0.k_1$	$q_0.k_2$	$q_0.k_3$	$q_0.k_4$
$q_1.k_0$	$q_1.k_1$	$q_1.k_2$	$q_1.k_3$	$q_1.k_4$
$q_2.k_0$	$q_2.k_1$	$q_2.k_2$	$q_2.k_3$	$q_2.k_4$
$q_3.k_0$	$q_3.k_1$	$q_3.k_2$	$q_3.k_3$	$q_3.k_4$
$q_4.k_0$	$q_4.k_1$	$q_4.k_2$	$q_4.k_3$	$q_4.k_4$