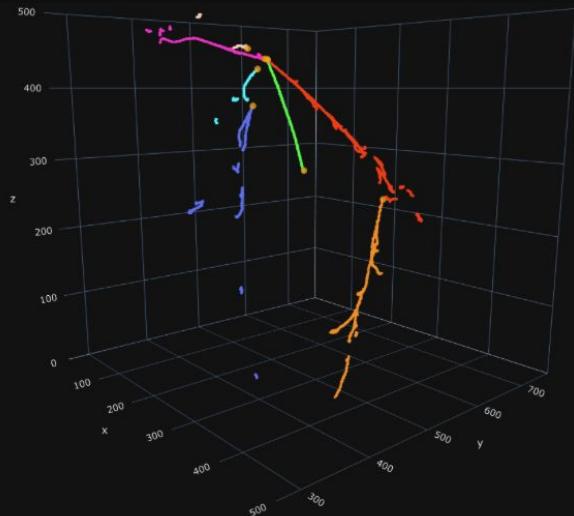
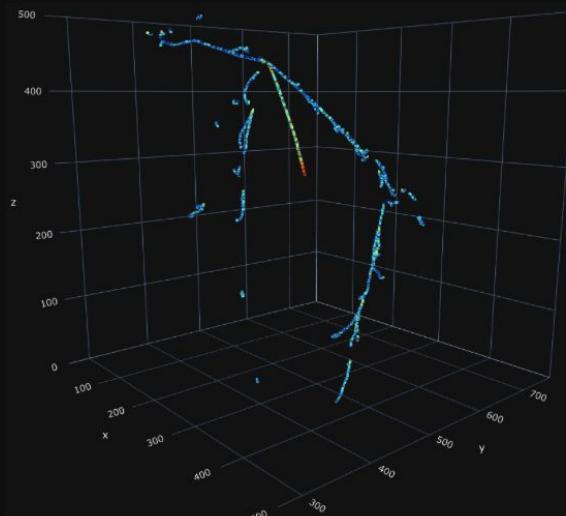
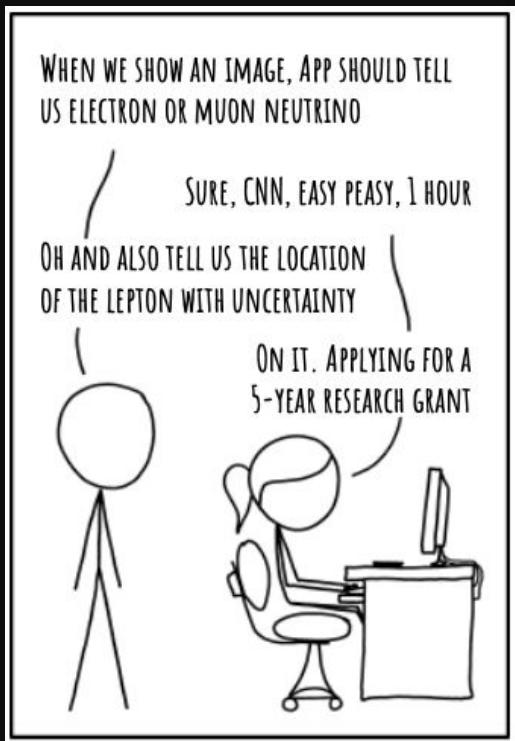


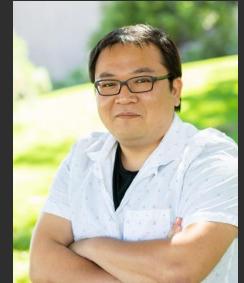
# Physics Inference Using Computer Vision and Machine Learning



Kazuhiro Terao  
Quantum Beam Imaging 2023  
Sep. 29th 2023 @ Osaka U.

# About me

- Life: 22 years in the U.S., family of 3 (a daughter 8 y.o.)
- Training: experimental neutrino physics (Ph.D, postdoc)
- Speciality: AI/ML for physics (scientist ~ now)
- Like: hiking, camping, backpacking, woodworking, research
- Ask me: AI/ML, studying/working in the U.S., great national parks, etc.



I was asked to talk about:

- How AI/ML is applied to physics research
- Challenges addressed by AI/ML (and not by traditional methods)

# About me

- Life: 22 years in the U.S., family of 3 (a daughter 8 y.o.)
- Training: experimental neutrino physics (Ph.D, postdoc)
- Speciality: AI/ML for physics (scientist ~ now)
- Like: hiking, camping, backpacking, woodworking, research
- Ask me: AI/ML, studying/working in the U.S., great national parks, etc.



## Outline

- Big picture: AI/ML from Computer Vision in physics
- Case study 1: multi-task image data analysis
- Case study 2: physics modeling using neural scene representation

# AI/ML @ the Frontier of HEP

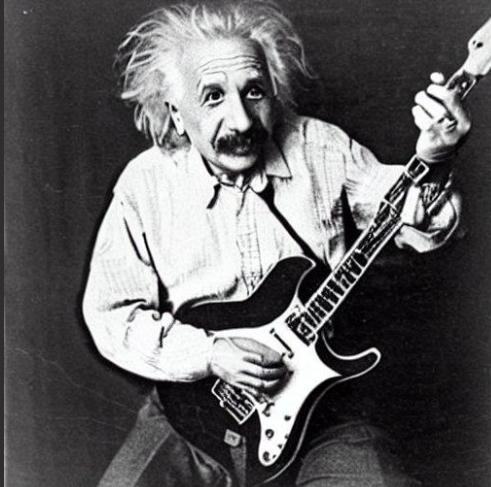
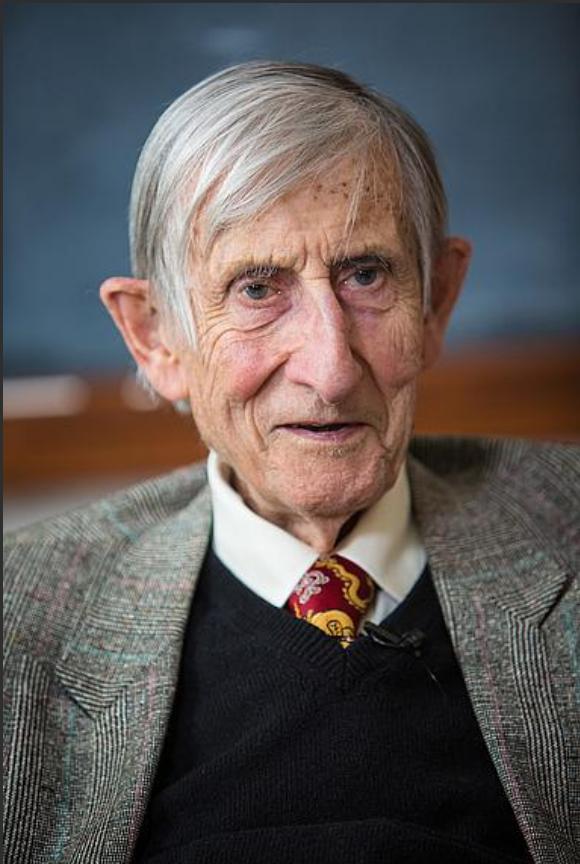


# Powerful Tools Accelerate Science

*“New directions in science are launched by new **tools** much more often than by new concepts.”*

Freeman Dyson (1923-2020)

*Truly exciting time for AI/ML + Physics!*

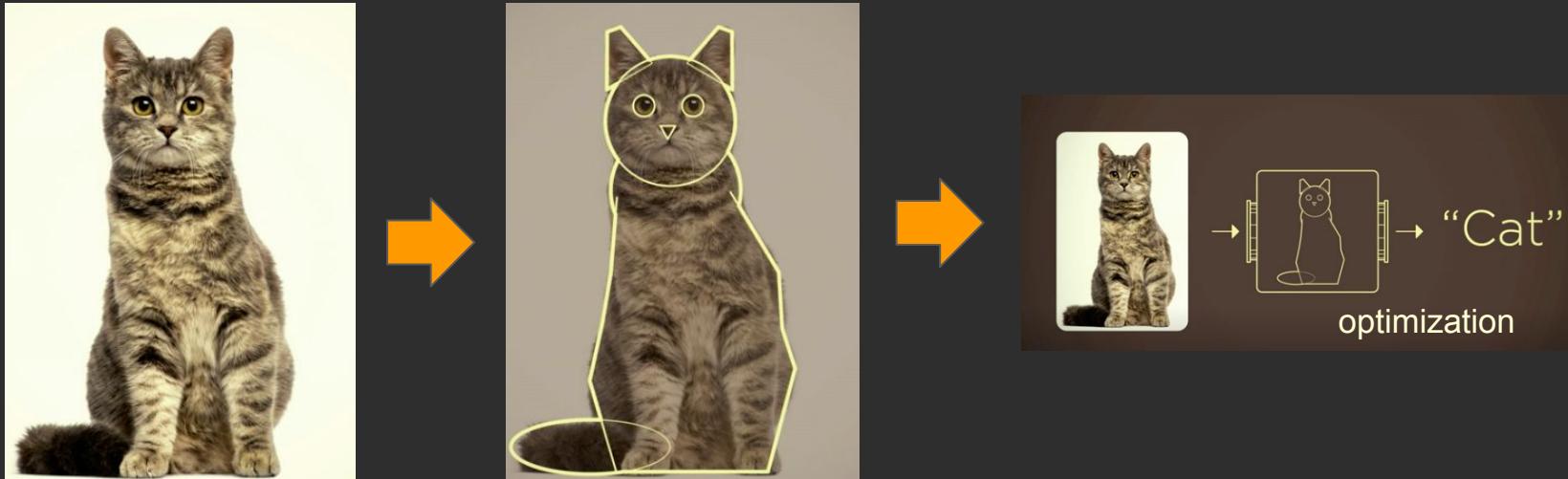


Teaser Q:  
Which is real?

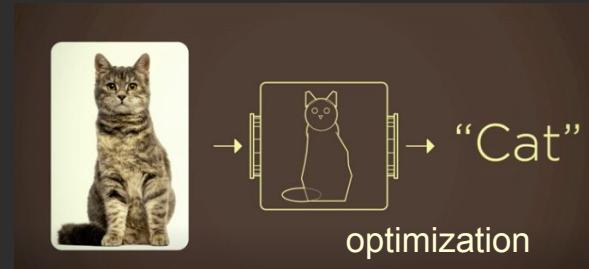
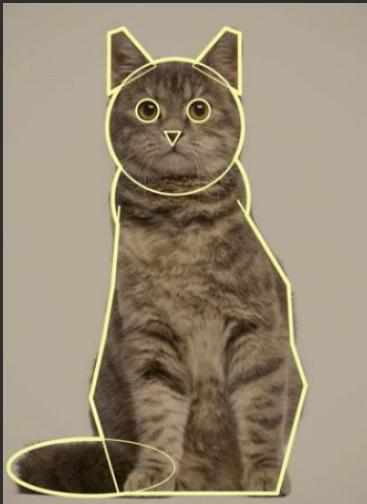
Left: Einstein w/  
an electric guitar

Right: a mugshot  
of my postdoc

# Particularly popular: Image Data Analysis

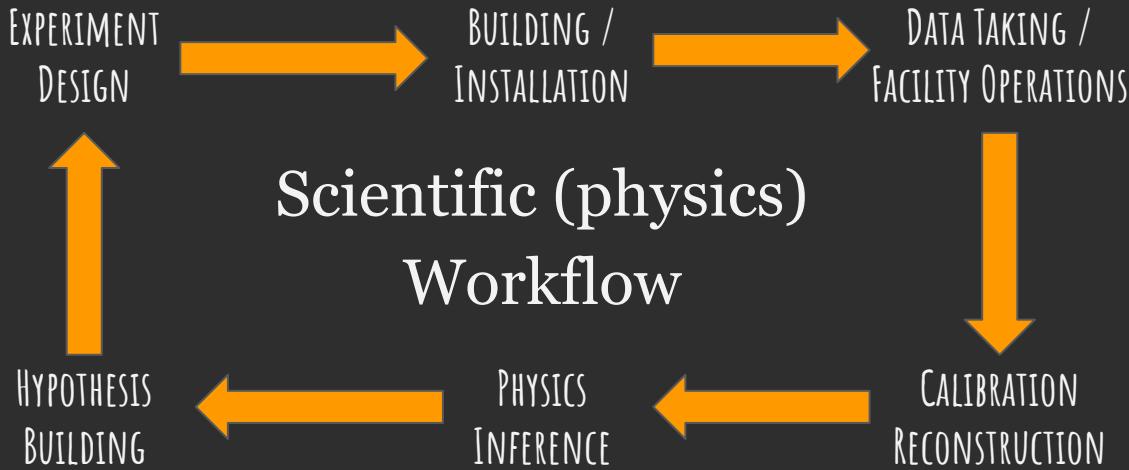


# Particularly popular: Image Data Analysis

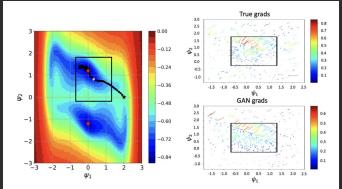


## Challenges

- Occlusion, masking, etc.
  - Many heuristics
- Algorithm optimization
- Reusability



## Aparatus design optimization ([2002.04632](#))



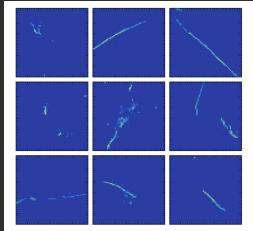
## AR for support (NASA)



## Fast-ML / Edge-ML ([HLS4ML](#))



## Fast Simulation ([2204.02496](#))



EXPERIMENT  
DESIGN

BUILDING /  
INSTALLATION

DATA TAKING /  
FACILITY OPERATIONS

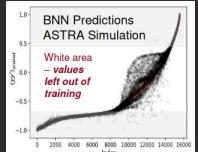
HYPOTHESIS  
BUILDING

AI/ML is  
Everywhere

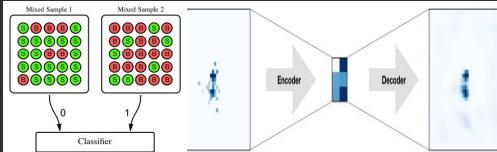
PHYSICS  
INFERENCE

CALIBRATION  
RECONSTRUCTION

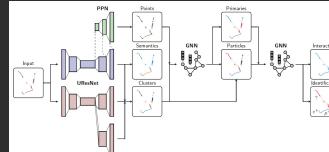
## Uncertainty Quantification ([PRAB 24.114601](#))



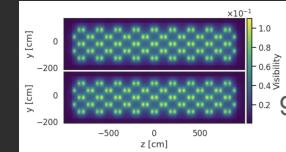
## New Physics Discovery ([2010.14554](#))



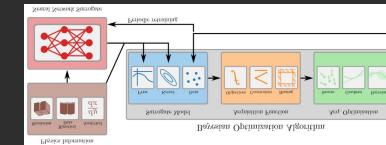
## Object Reconstruction ([2102.01033](#))



## Detector Response Calibration ([2102.01033](#))



## Bayesian Optimization Reinforcement Learning ([2010.09824](#), [2202.07747](#))

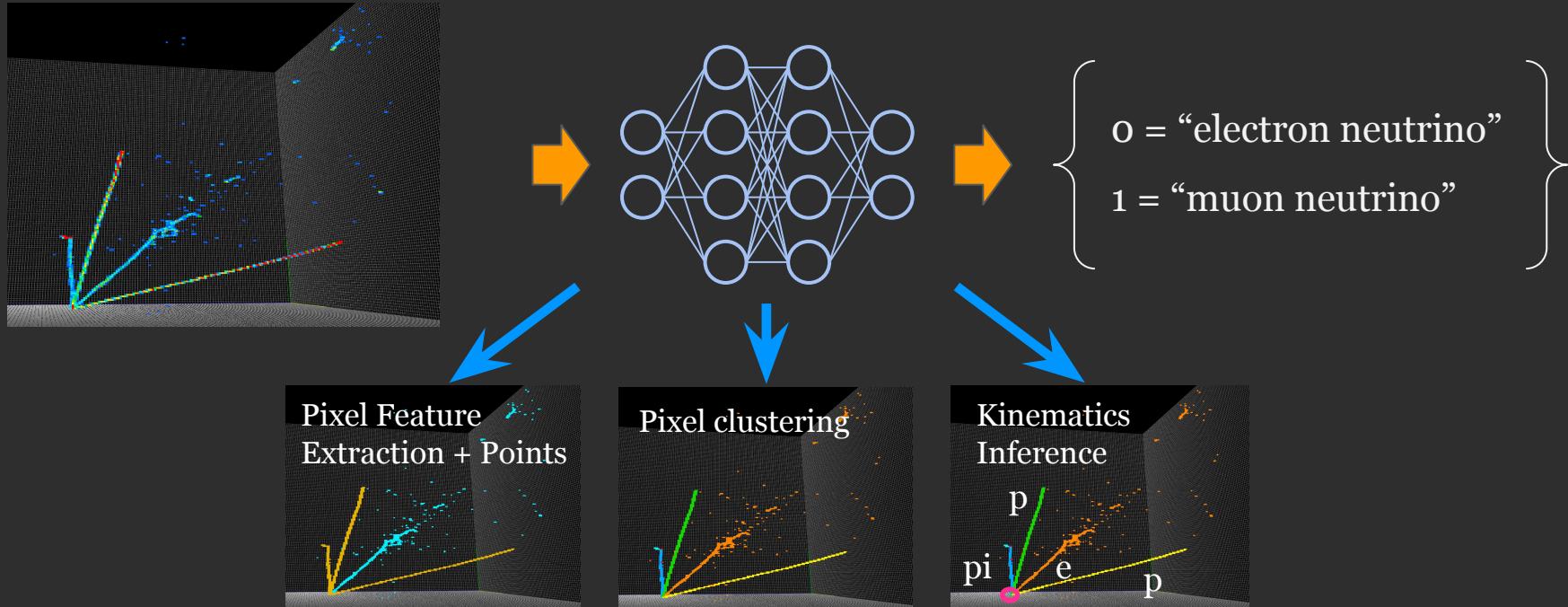


# AI/ML Applications in Physics Image Data

- AI/ML is everywhere in physics today
- Multi-task image data analysis in physics
- Physics modeling using neural scene representation
- Wrap-up



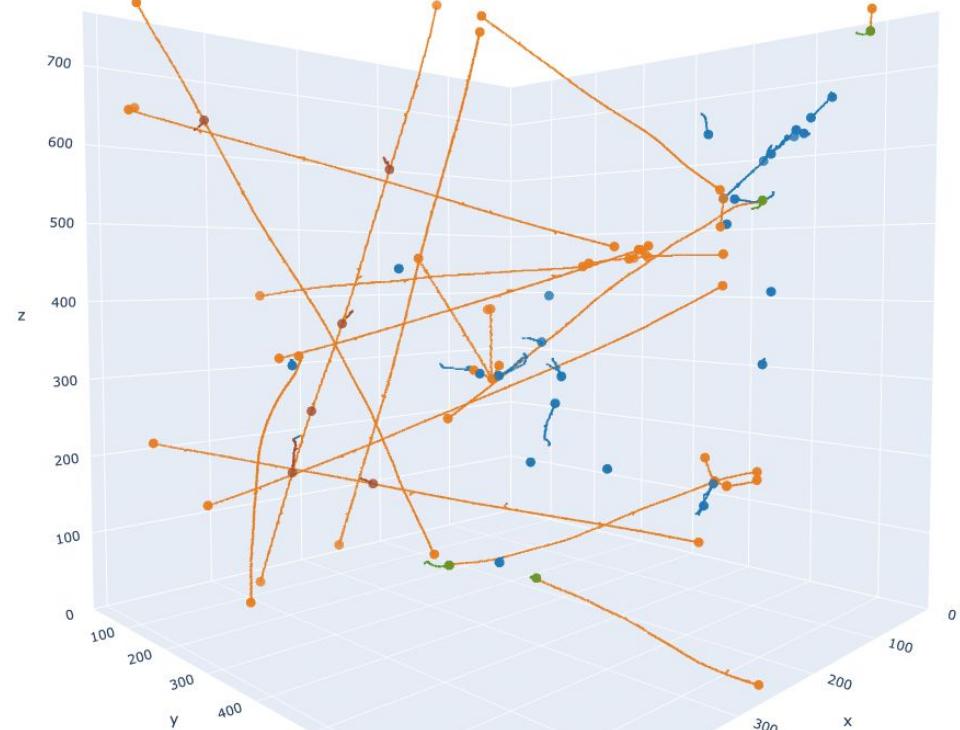
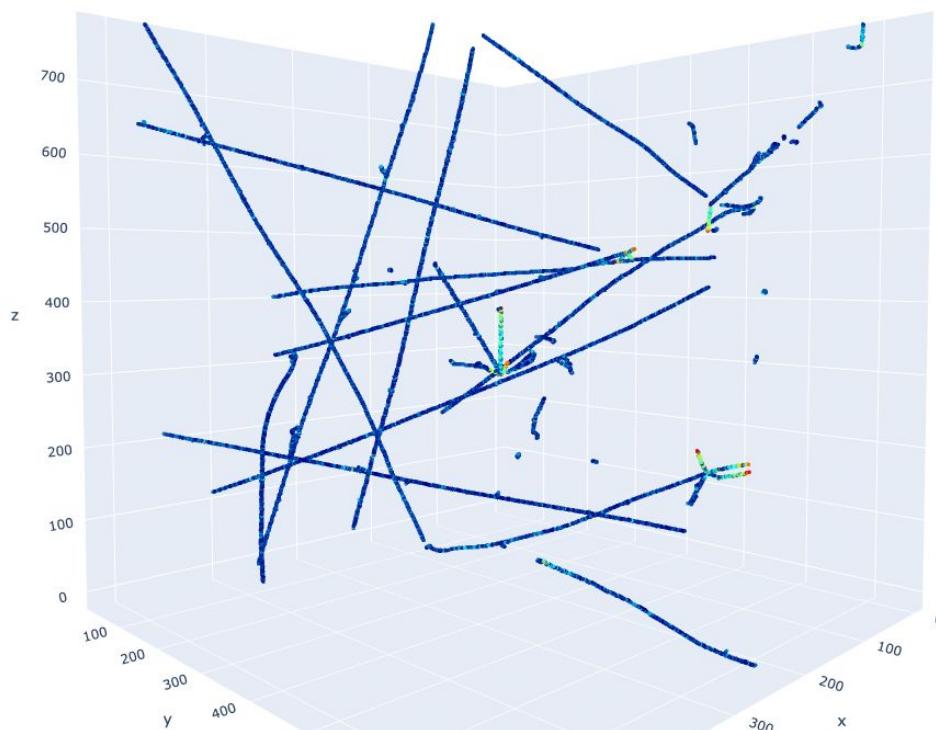
# Example 1: Object Reconstruction



Extract intermediate physical observables (“objects”) with sensible hierarchical correlations

# Example 1: Object Reconstruction

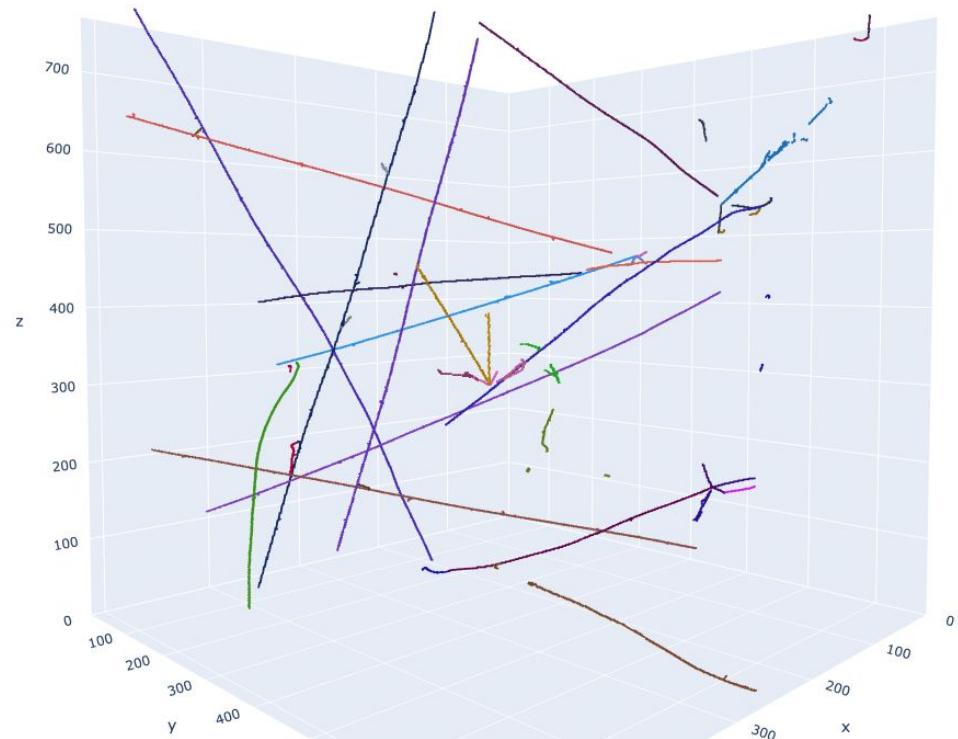
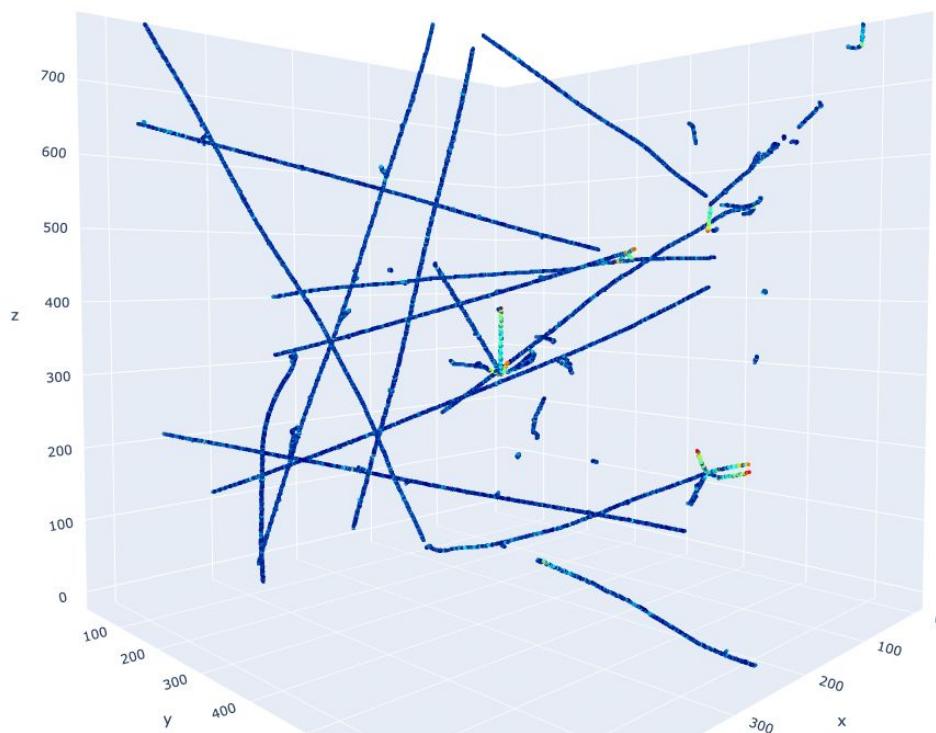
[Full chain \(NeurIPS WS\)](#)  
[Public dataset](#)  
[1](#), [2](#), [3](#), [4](#)



**Step 1:** Identifying pixel-level key features  
(Sparse-CNN for globally-sparse, locally-dense images)

# Example 1: Object Reconstruction

[Full chain \(NeurIPS WS\)](#)  
[Public dataset](#)  
[1](#), [2](#), [3](#), [4](#)

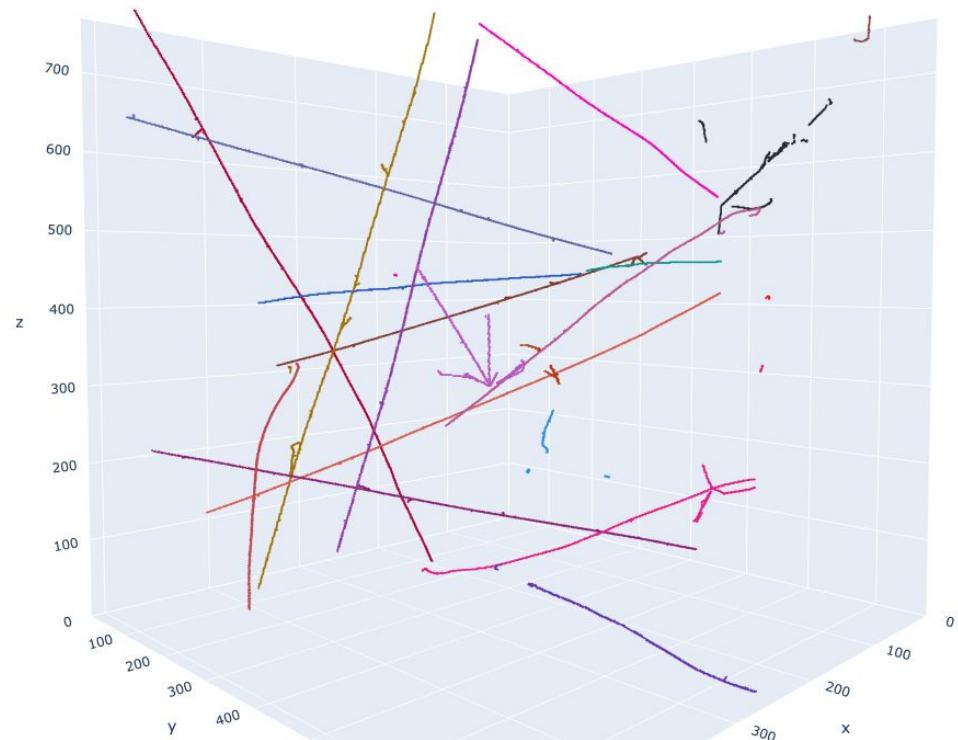
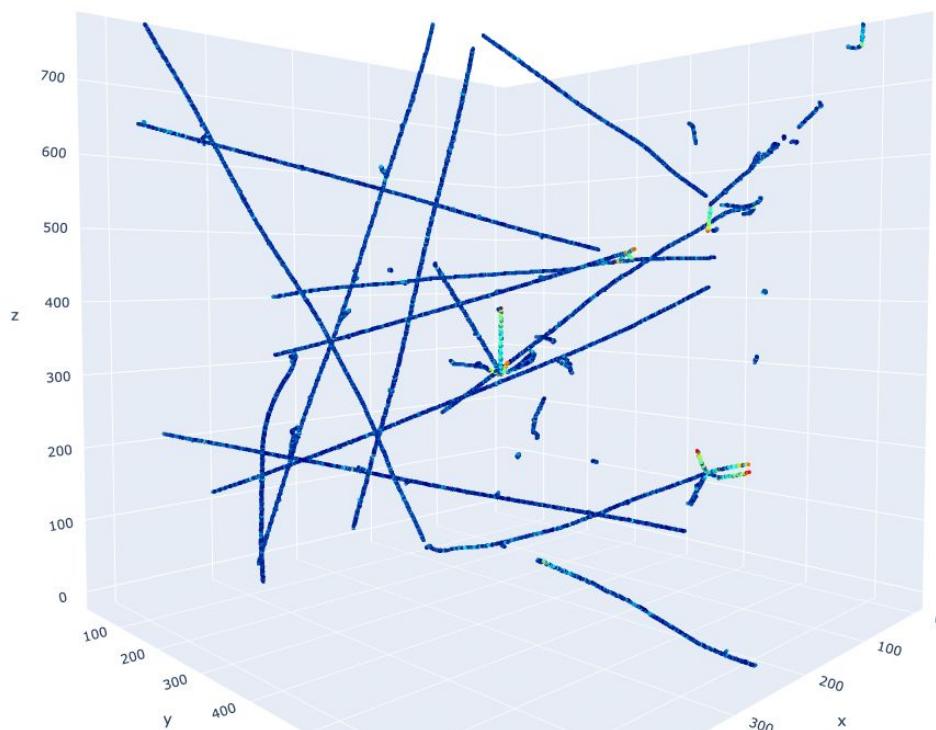


**Step 2:** Identifying individual particles

(CNN for dense-pixel clustering + GNN for scattered cluster aggregation)

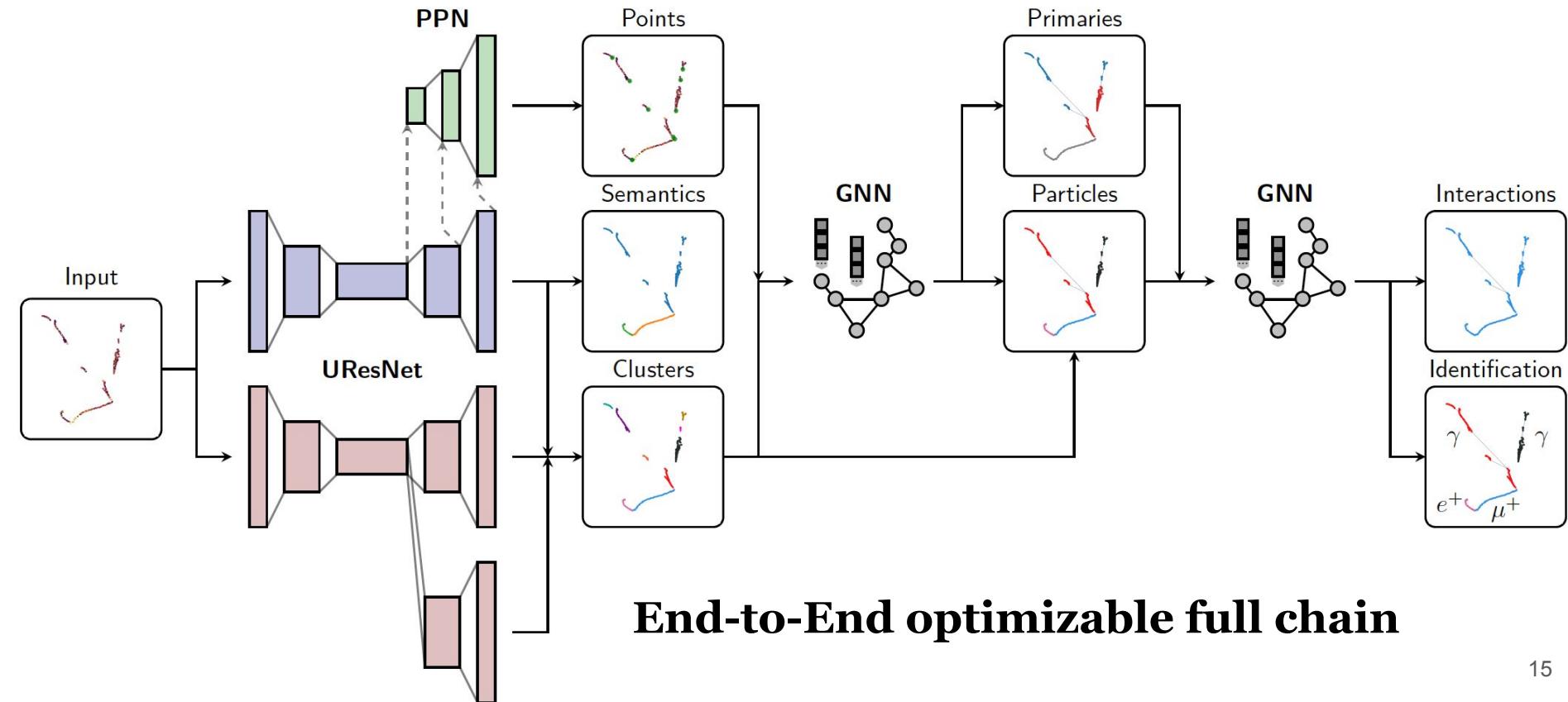
# Example 1: Object Reconstruction

[Full chain \(NeurIPS WS\)](#)  
[Public dataset](#)  
[1](#), [2](#), [3](#), [4](#)



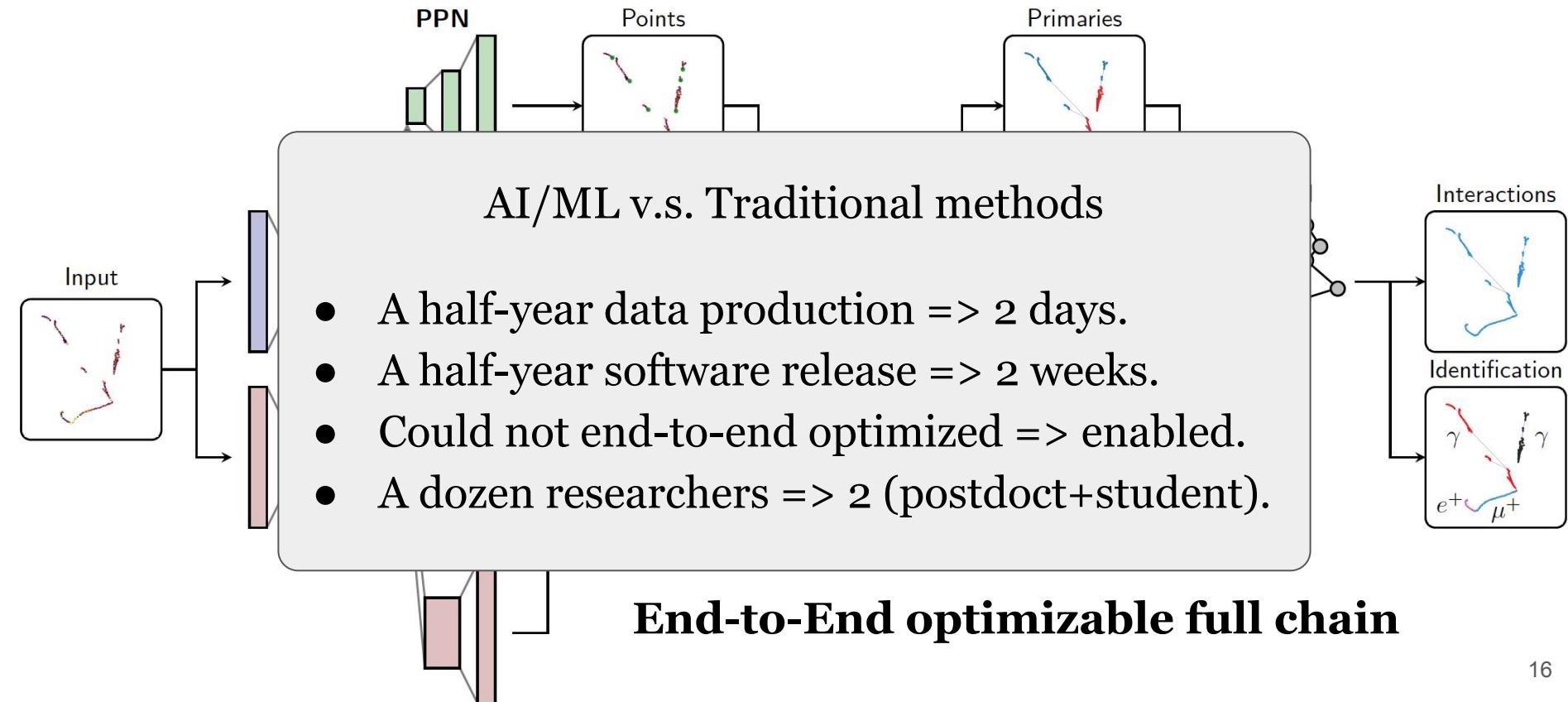
**Step 3:** Identifying particle-to-particle correlations  
(GNN with directed graph and node/edge aggregation layers)

# Full Data Reconstruction via Deep Learning

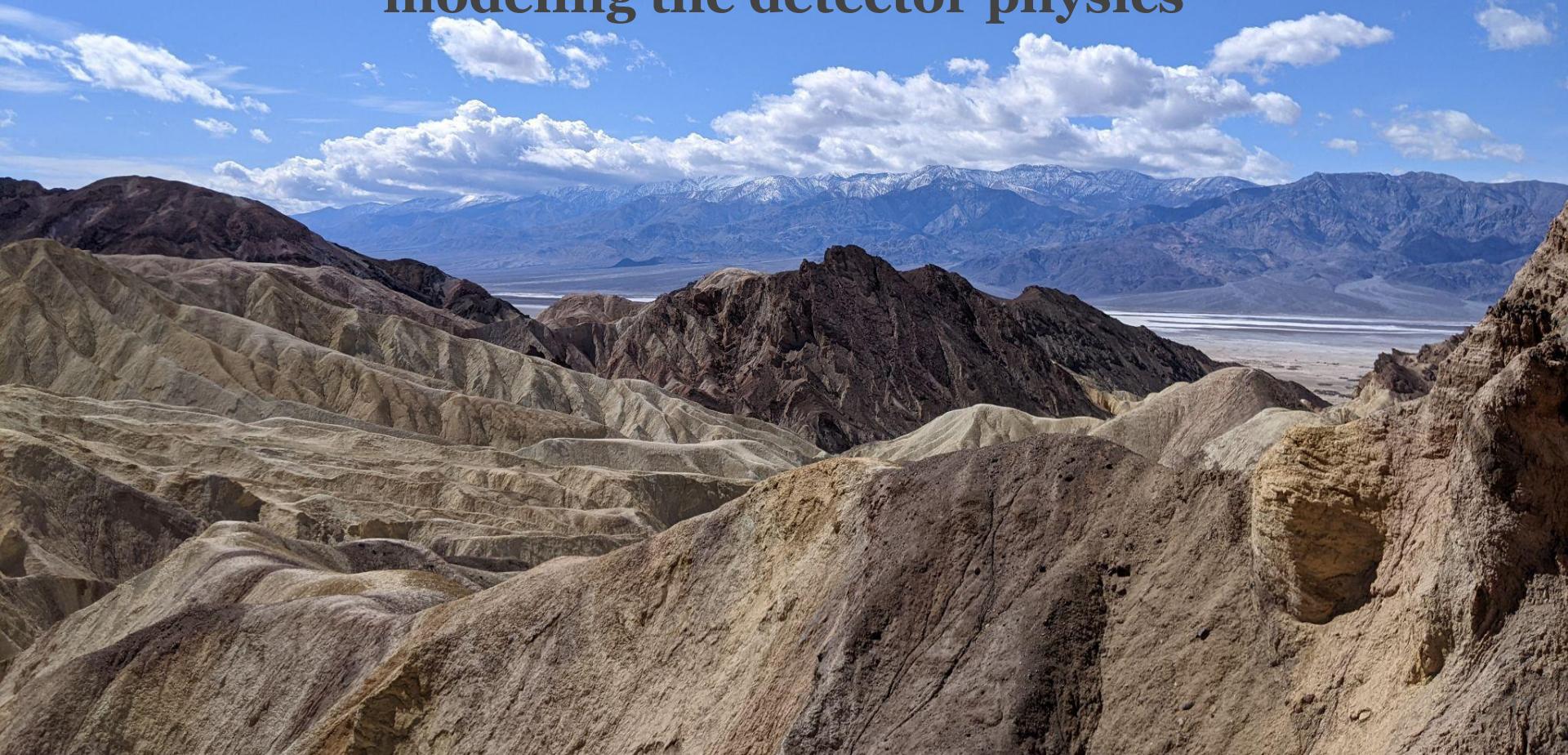


# Full Data Reconstruction via Deep Learning

[Full chain \(NeurIPS WS\)](#)  
[Public dataset](#)  
[1](#), [2](#), [3](#), [4](#)



# Neural Scene Representation for modeling the detector physics

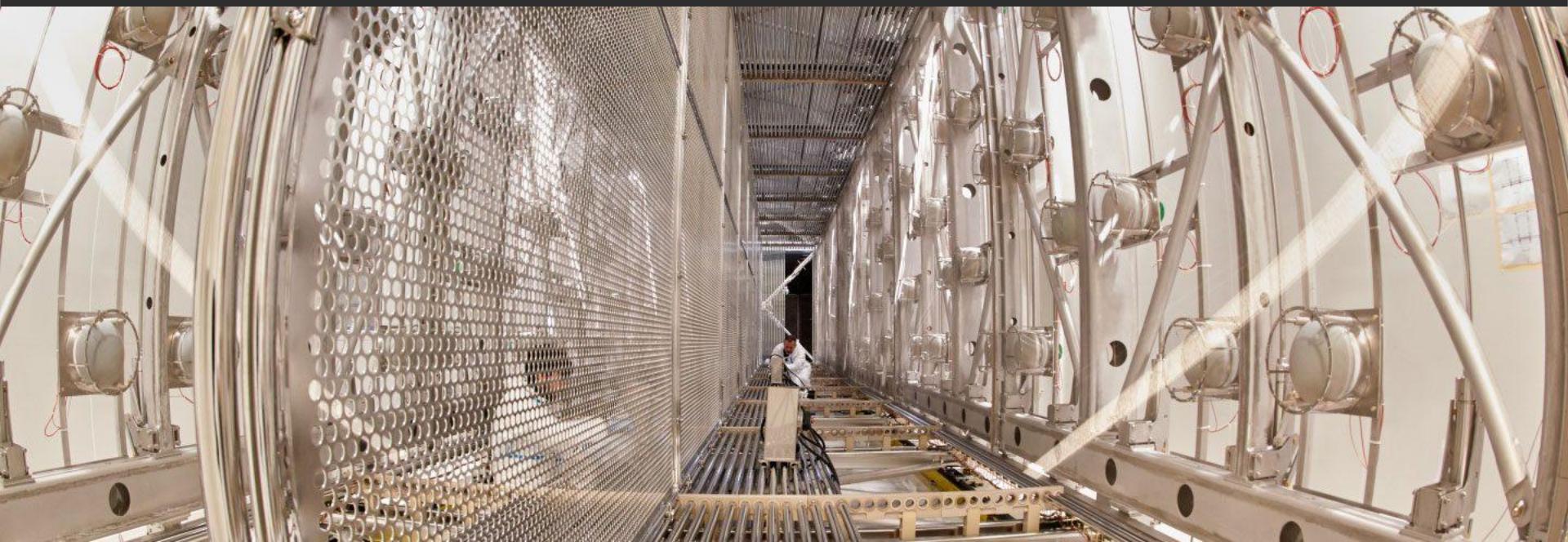


# ML for Detector Physics Modeling

## LAr scintillator light detection

SLAC

### Photo-multiplier tubes (PMTs)

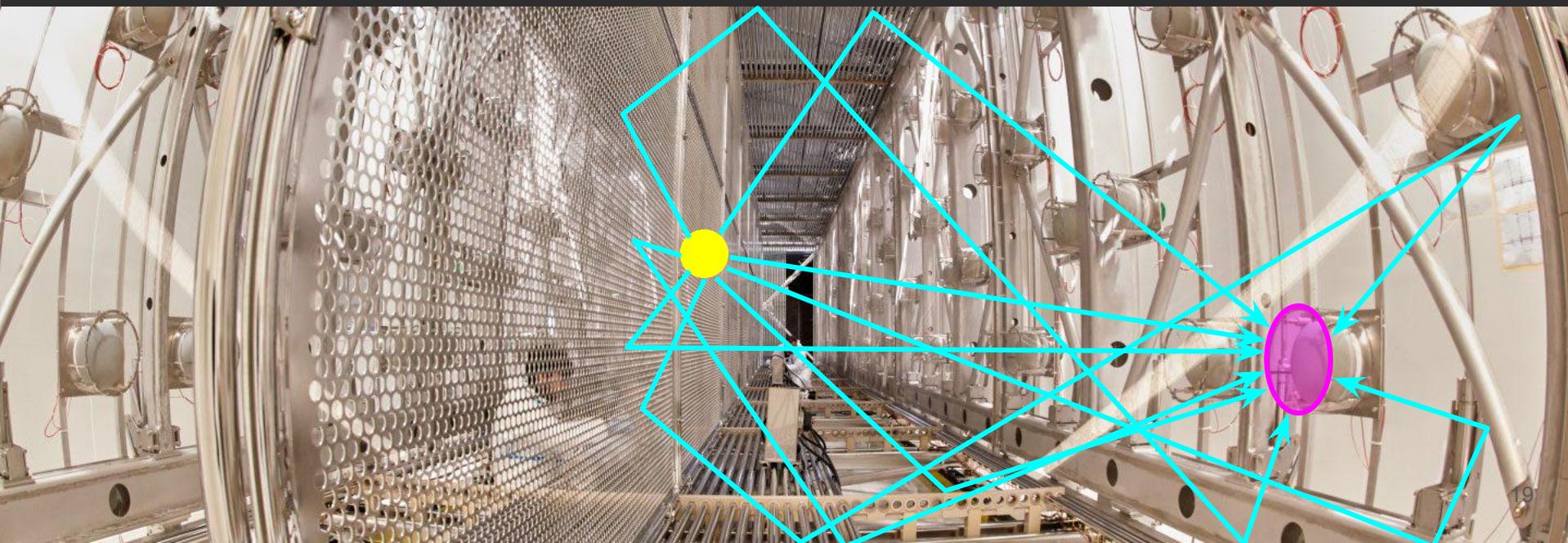


# ML for Detector Physics Modeling

## LAr scintillator light detection

SLAC

**Photo-multiplier tubes (PMTs)** detect scintillation photons produced isotropically from an Argon atom when charged particle deposit energy.



# ML for Detector Physics Modeling

## LAr scintillator light detection

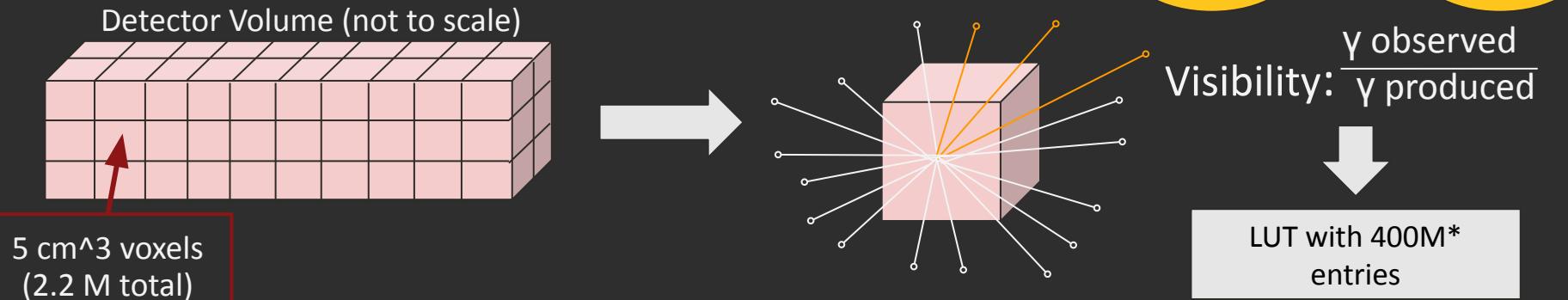
SLAC

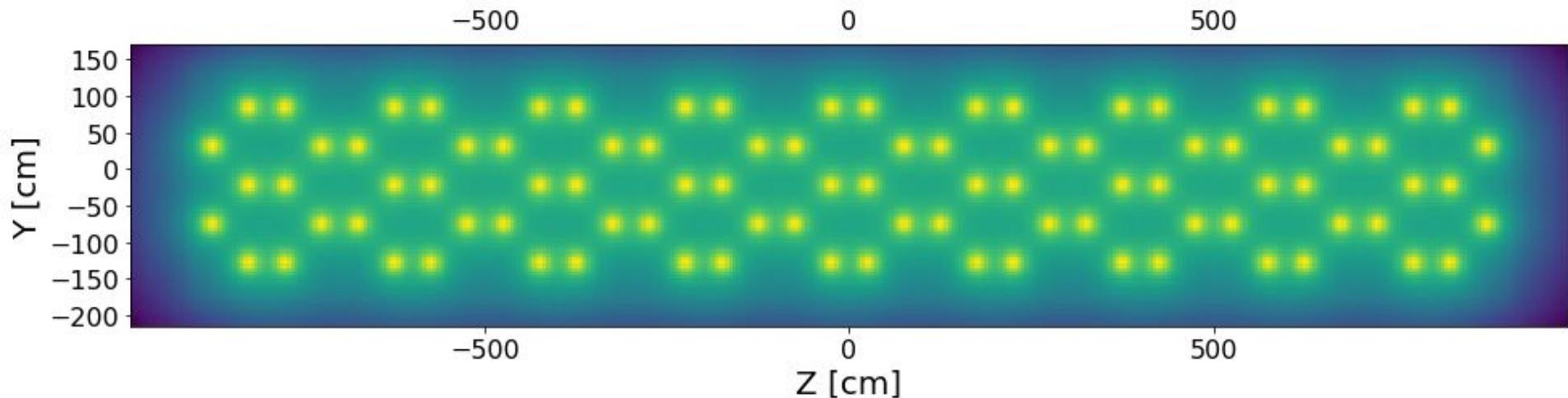
**Photo-multiplier tubes (PMTs)** detect scintillation photons produced isotropically from an Argon atom when charged particle deposit energy.

Challenge: modeling a LOT of photons

- Each event produces  $\sim 1E9$  photons ... too slow to simulate (track) each one

Traditional approach: a 4D Look-Up Table (LUT)



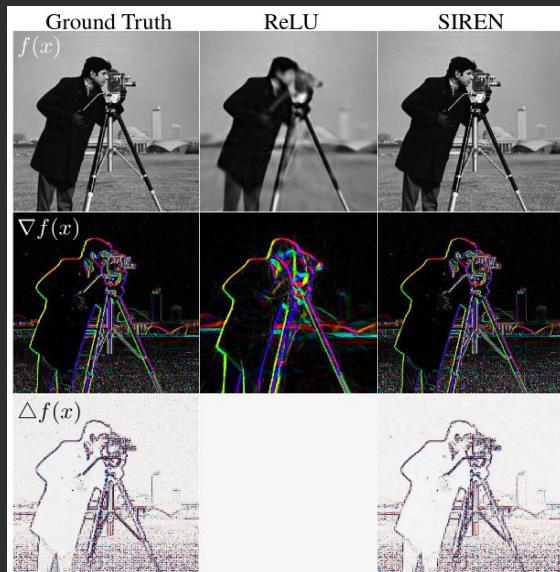


## Challenges for LUT

- Bad scaling
- Takes long time to generate (simulate)
- Generation process is subject to statistical fluctuation (MC statistics)
- Difficult to optimize on real data

*Could we machine learn  $P(x, y, z, pmt)$ ?*

### Differentiable Neural Scene Representation



### SIREN as a surrogate optical transport

- Designed as an implicit representation of a continuous function in space (suited to “visibility”, “E-field”, etc.)
- “Differentiable” = can directly optimize against “data v.s. simulation discrepancy” given control samples

**SIREN**: success of learning the 1st and 2nd order derivatives

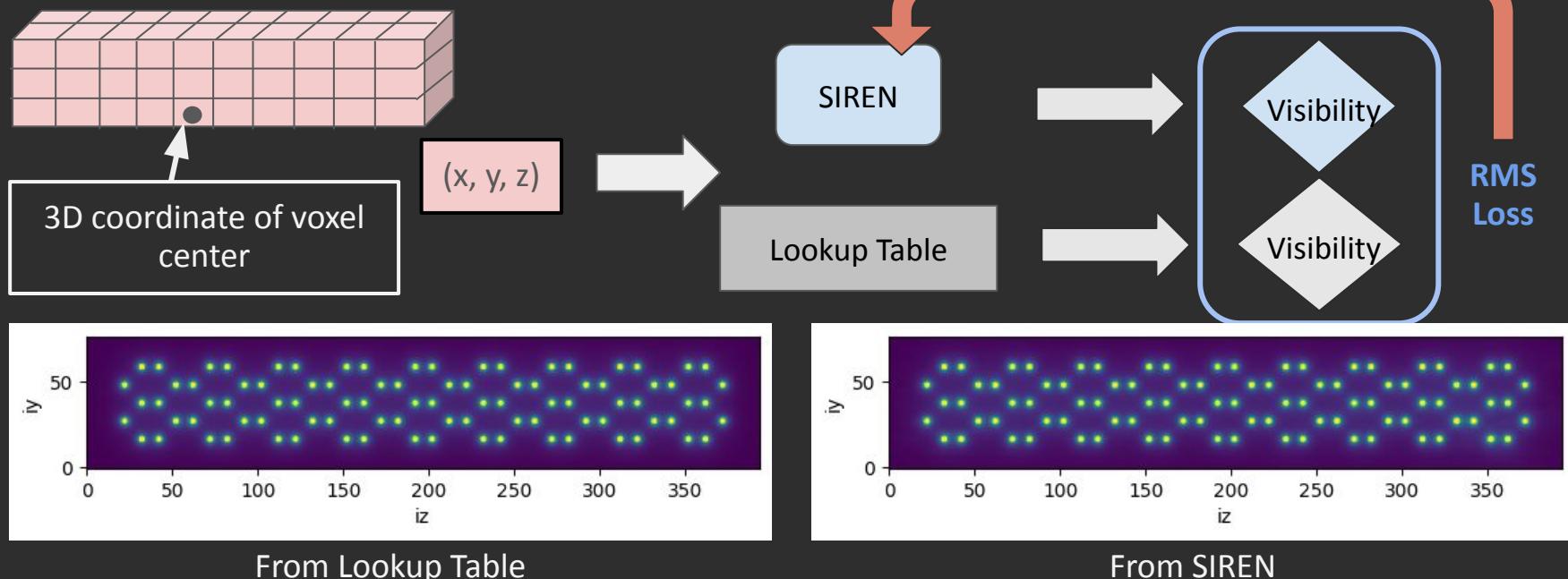
# ML for Detector Physics Modeling

## SIREN as a differentiable surrogate for optical detectors

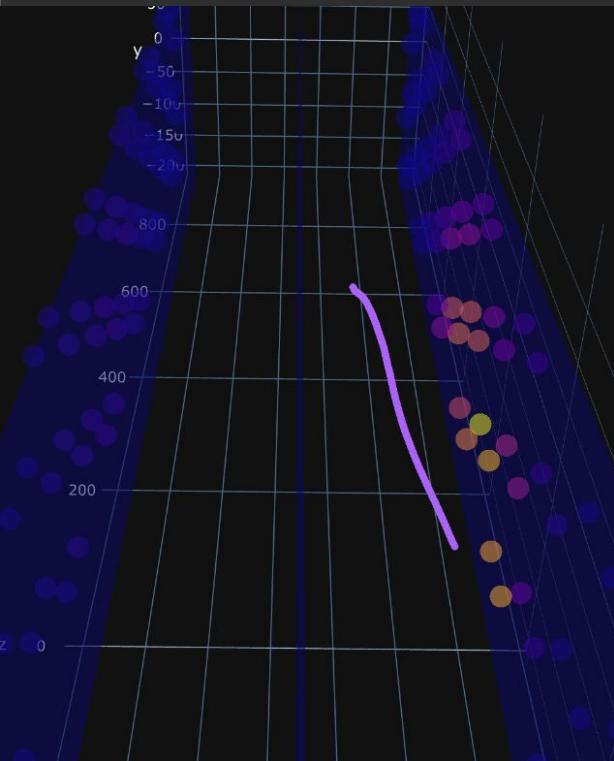
SLAC

### Optimize SIREN using LUT

#### Optimizing SIREN using Lookup Table



### Optimize SIREN using real data



**Control dataset:** 3D TPC trajectory for which XYZ position of space-points are accurately measured

$$P_j = \sum_i \text{Deposited charge at the point } i \times \text{Quantum efficiency of the PMT } j$$

Predicted P.E.      all points      Quantum efficiency  
 light yield      SIREN prediction for  
 the point i at the PMT j

$$\text{Loss} = \sum_j \frac{(P_j - O_j)^2}{P_j^2 + \epsilon}$$

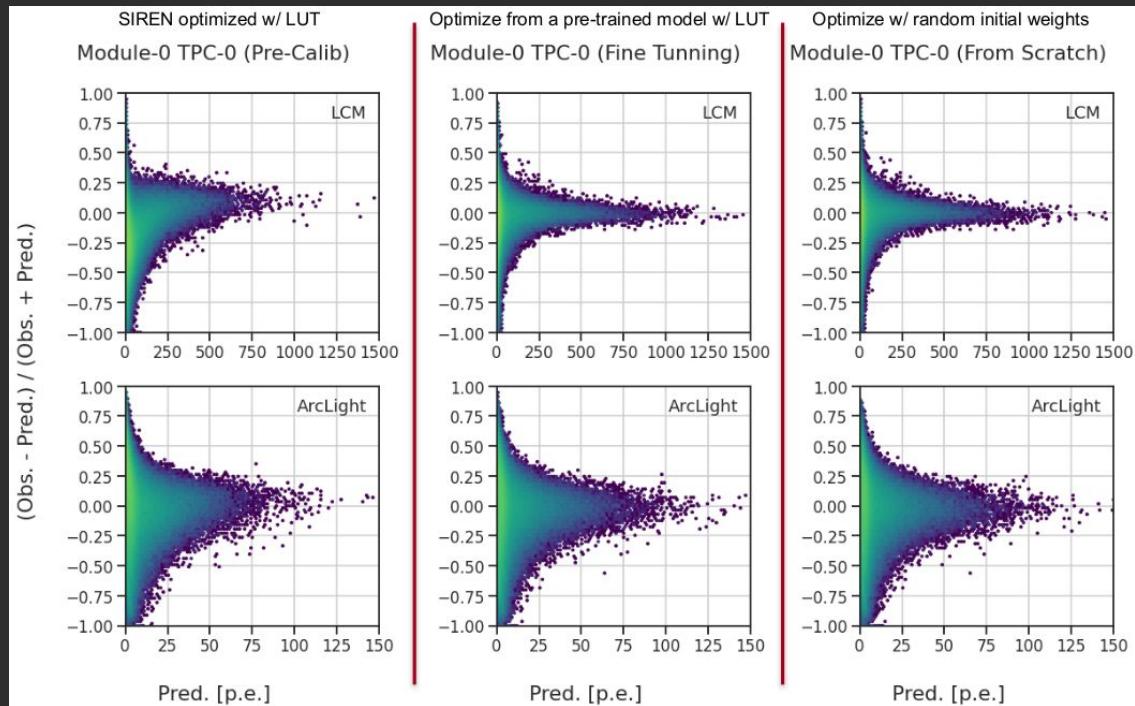
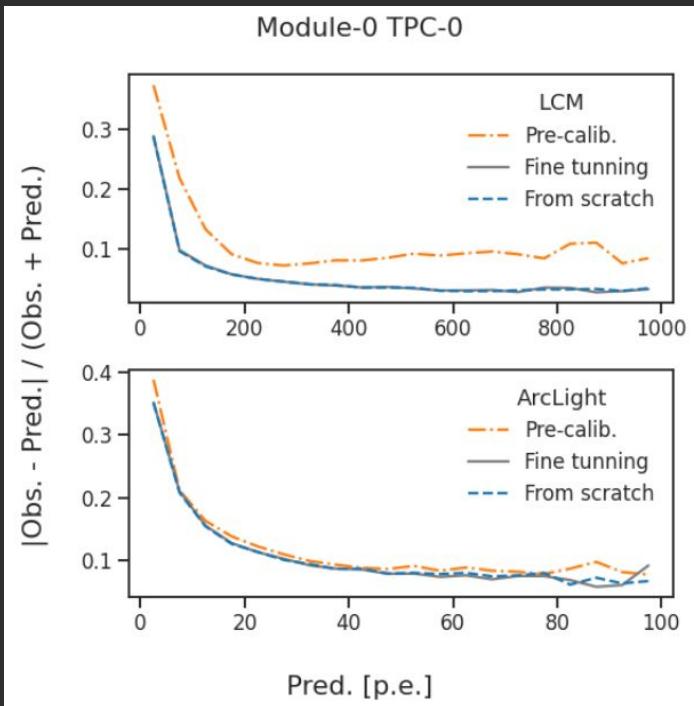
numerical stability,  
 ~25 P.E.-squared

# ML for Detector Physics Modeling

## SIREN as a differentiable surrogate for optical detectors

SLAC

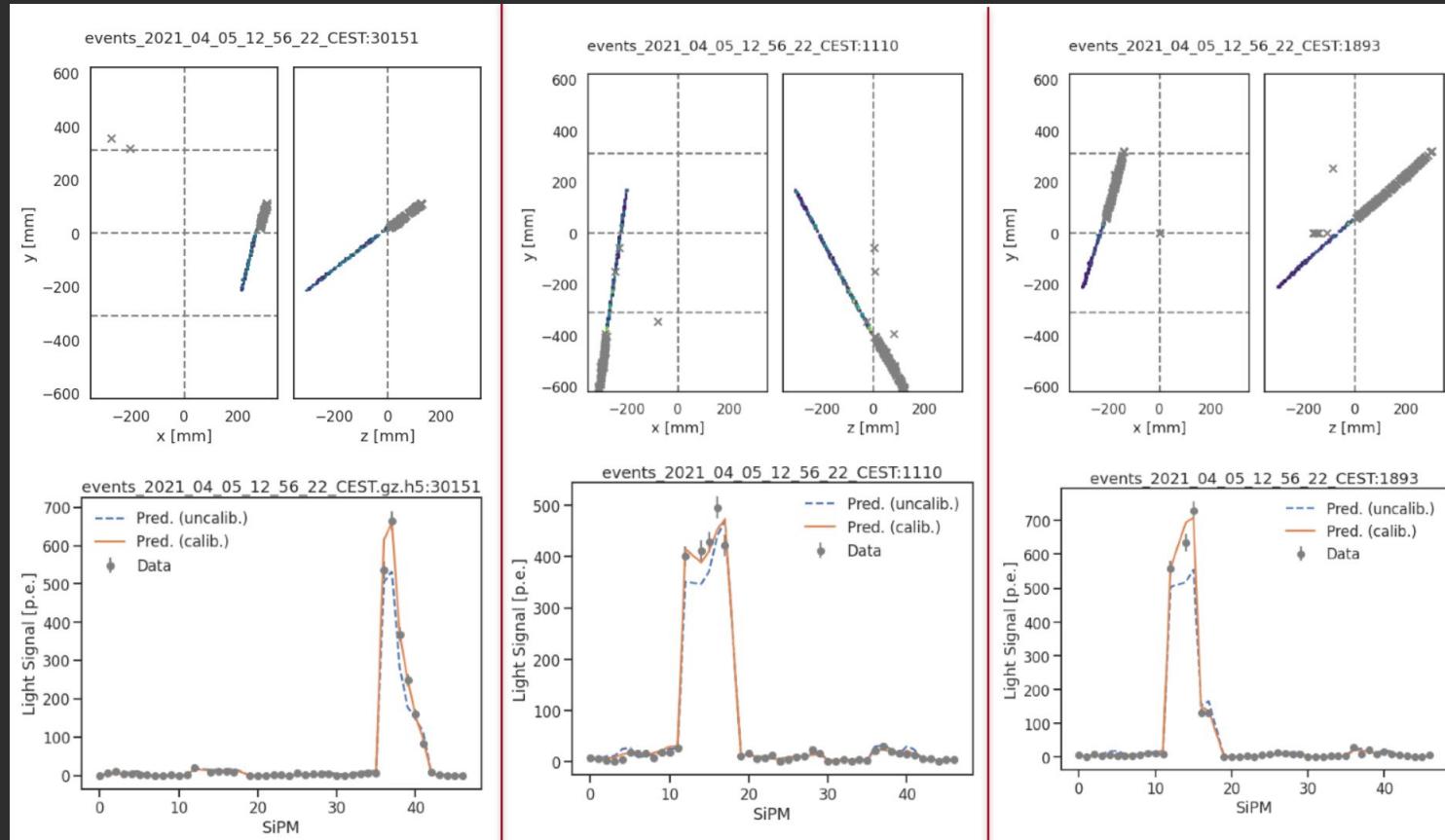
### Training SIREN on real data



# ML for Detector Physics Modeling

## SIREN as a differentiable surrogate for optical detectors

SLAC



A photograph of a vast, white salt flat under a dramatic sky. In the foreground, a person wearing a straw hat and a light-colored shirt walks away from the camera towards the horizon. The background features a range of mountains under a blue sky with scattered clouds.

Wrapping up...  
What's Next?

# Last Slide!

AI/ML application can make impact everywhere

It can fundamentally change how we design, build, operate, and analyze

AI/ML for physics or science is its own area of research

- “Off the shelf” solution often not optimal
- Need domain knowledge from both physics and AI/ML necessary
- Physics-informed AI/ML v.s. general AI?

Thank you for your attention!

Discussions, collaborations, organizing events all welcome :)

... also join our hiking journeys!



Fin.

# Back-up Slides



# Fun

## CatGPT

*What if ChatGPT were a cat?*



Meow, meow meow meow, meow meow?



What are the goals of the ICARUS experiment?



Meow meow.



I learned something. Thank you.



Meow meow.



# SciML: Applying AI/ML Hiking Skills For Physics Inference

A wide-angle photograph of a desert landscape, likely Death Valley National Park. In the foreground, there are rugged, layered rock formations with distinct sedimentary structures in shades of brown, tan, and light beige. To the right, a large, flat, light-colored valley floor stretches towards the horizon. In the background, a range of mountains is visible, with some peaks showing patches of snow or ice. The sky is a clear, vibrant blue with scattered white, puffy clouds.

# The Other View: Injecting AI/ML Into Physics

Success of AI is through gradient-based optimization that works for millions, billions, or even trillions of parameters.

Journal of Machine Learning Research 23 (2022) 1-40  
Submitted 8/21; Revised 3/22; Published 4/22

**Switch Transformers: Scaling to Trillion Parameter Models with Simple and Efficient Sparsity**

**William Fedus\***  
LIAMFEDUS@GOOGLE.COM

**Barret Zoph\***  
BARRETZOPH@GOOGLE.COM

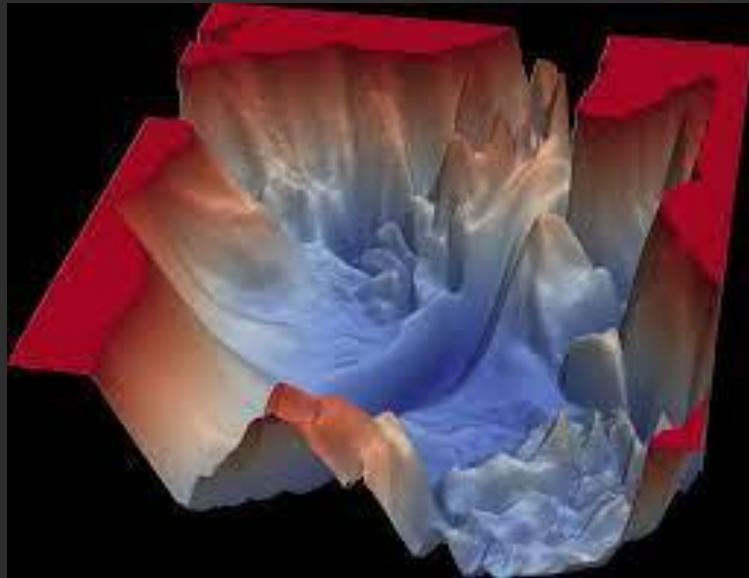
**Noam Shazeer**  
NOAM@GOOGLE.COM  
Google, Mountain View, CA 94043, USA

**Editor:** Alexander Clark

**Abstract**

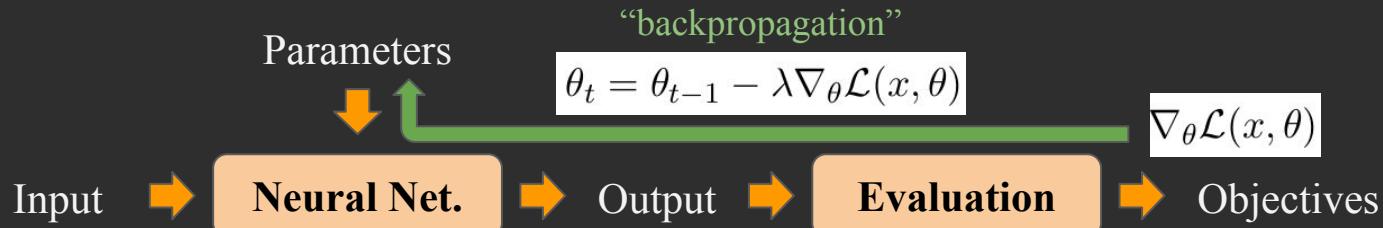
In deep learning, models typically reuse the same parameters for all inputs. Mixture of Experts (MoE) models defy this and instead select *different* parameters for each incoming example. The result is a sparsely-activated model—with an outrageous number of parameters—but a constant computational cost. However, despite several notable successes of MoE, widespread adoption has been hindered by complexity, communication costs, and training instability. We address these with the introduction of the Switch Transformer. We simplify the MoE routing algorithm and design intuitive improved models with reduced communication and computational costs. Our proposed training techniques mitigate the instabilities, and we show large sparse models may be trained, for the first time, with lower precision (bf16) formats. We design models based off T5-Base and T5-Large (Raffel et al., 2019) to obtain up to 7x increases in pre-training speed with the same computational resources. These improvements extend into multilingual settings where we measure gains over the mT5-Base version across all 101 languages. Finally, we advance the current scale of language models by pre-training up to trillion parameter models on the “Cocalico Clean Crawled Corpus”, and achieve a 4x speedup over the T5-XXL model.<sup>12</sup>

**Keywords:** mixture-of-experts, natural language processing, sparsity, large-scale machine learning, distributed computing

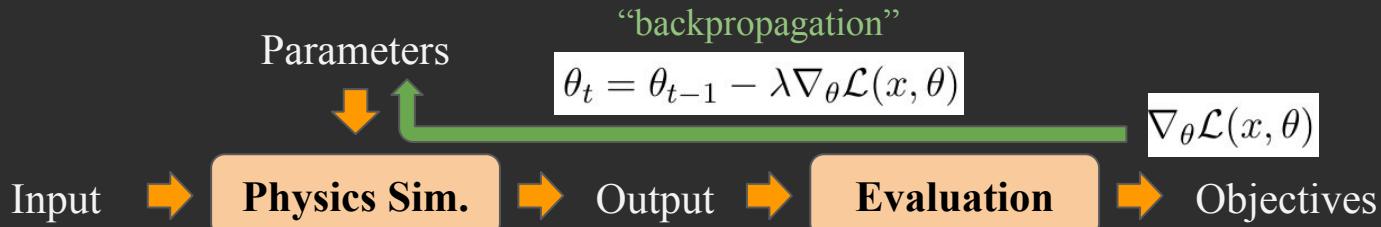


*“To deal with hyper-planes in a 14-dimensional space, visualize a 3-D space and say ‘fourteen’ to yourself very loudly. Everyone does it.” - Geoffrey Hinton*

# Gradient-based Optimization



# Differentiable Physics Models



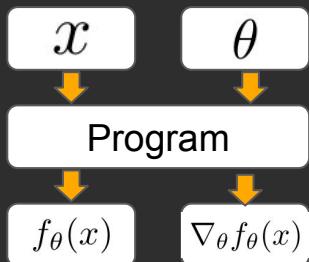
Class

TensorFlow

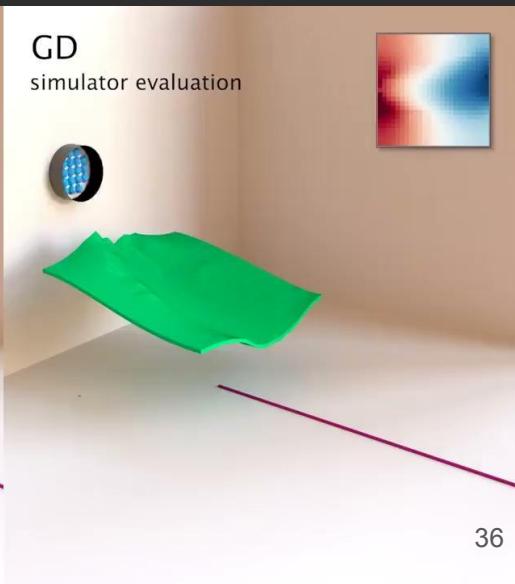
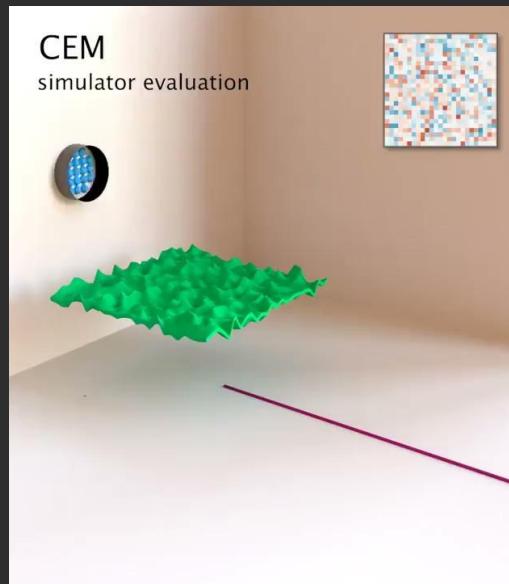


PyTorch

Zygote

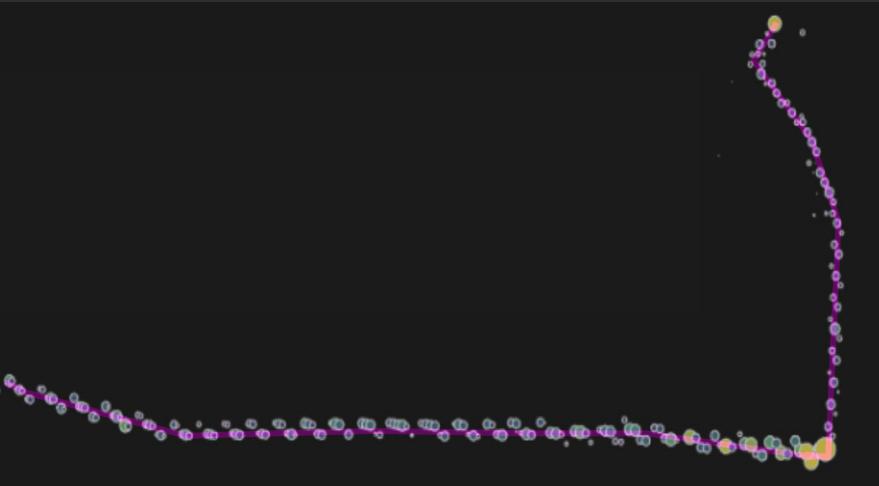


Physical Design  
using Differentiable  
Learned Simulators  
(DeepMind [2202.00728](#))

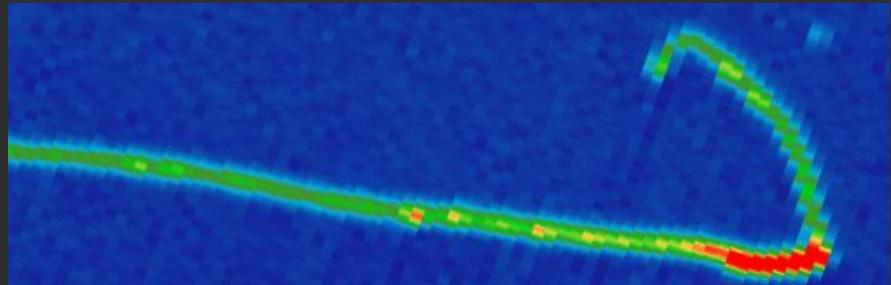


# Differentiable Physics Models

## Modeling Detector Physics

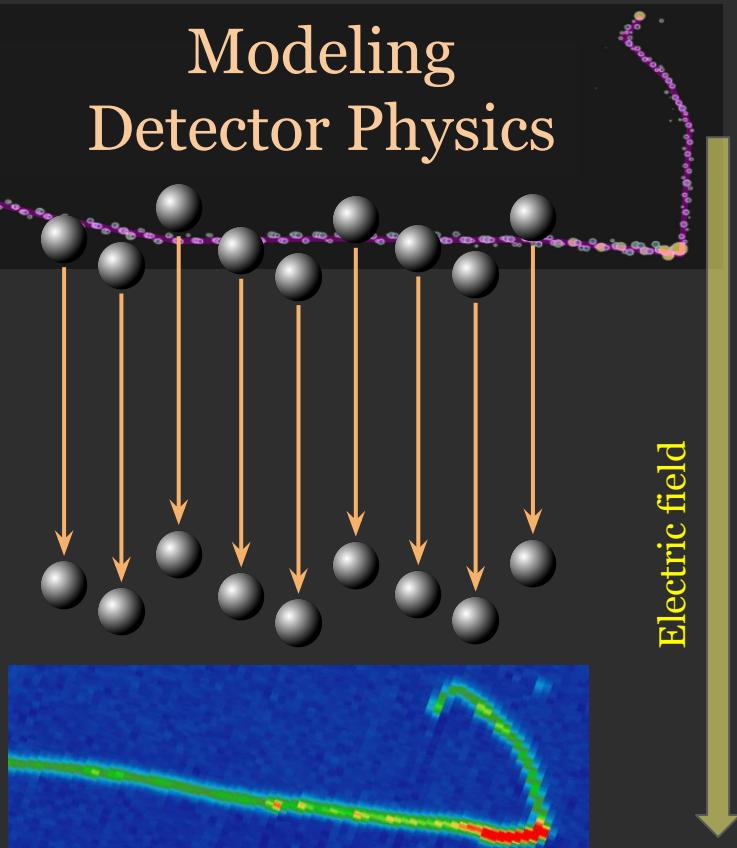


**Example:** Liquid Argon TPC  
**Objective:** given a calibration dataset (i.e. images of particle trajectories with approximated  $dE/dX$  values), “fit” the detector physics parameters



# Differentiable Physics Models

## Modeling Detector Physics



**Example:** Liquid Argon TPC

- Charged particle ionize electrons
- Electrons drifts under E-field
- Signal diffuse and attenuated

Detector Simulation

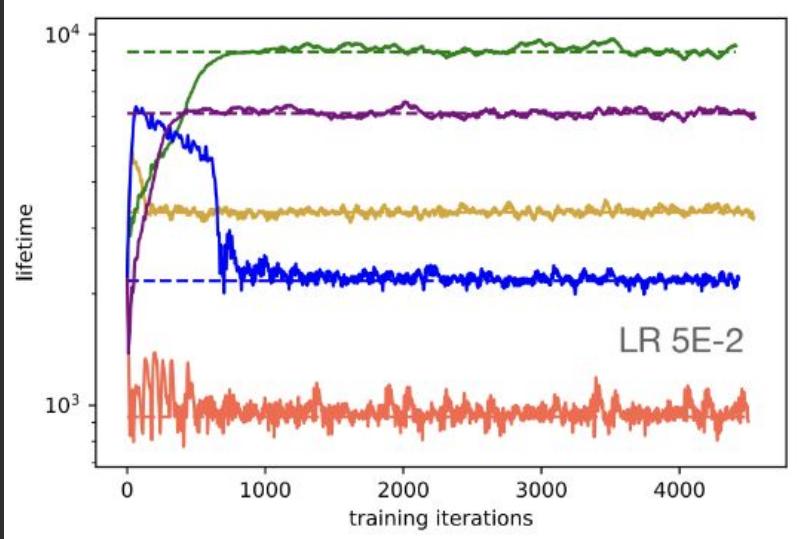
$$\mathcal{R}_{\text{ICARUS}} = \frac{A_B}{1 + k_B \cdot (dE/dx) / \mathcal{E}}$$

$$Q = Q_0 \exp(-v_{\text{drift}} t / \tau)$$

$$\sigma_t^2(t) \simeq \sigma_t^2(0) + \left( \frac{2D_L}{v_d^2} \right) t$$

...

# Differentiable Physics Models



Optimizing the “lifetime”  
physics parameter directly  
from calibration dataset

- Example:** Liquid Argon TPC
- Charged particle ionize electrons
  - Electrons drifts under E-field
  - Signal diffuse and attenuated

Detector Simulation

$$\mathcal{R}_{\text{ICARUS}} = \frac{A_B}{1 + k_B \cdot (dE/dx) / \mathcal{E}}$$

$$Q = Q_0 \exp(-v_{\text{drift}} t / \tau)$$

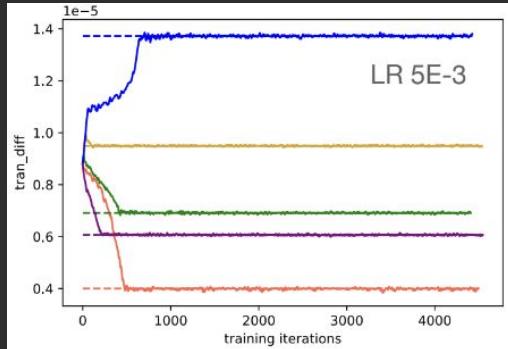
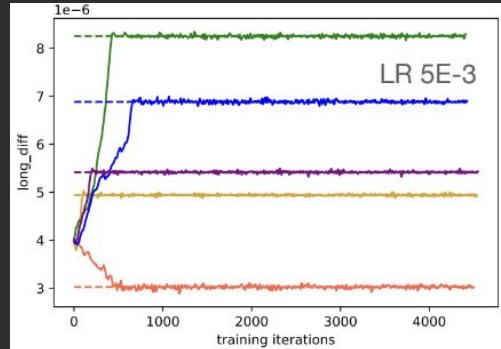
$$\sigma_t^2(t) \simeq \sigma_t^2(0) + \left( \frac{2D_L}{v_d^2} \right) t$$

...

# Differentiable Physics Models

Diffusion during the drift

$$\sigma_t^2(t) \simeq \sigma_t^2(0) + \left( \frac{2D_L}{v_d^2} \right) t$$

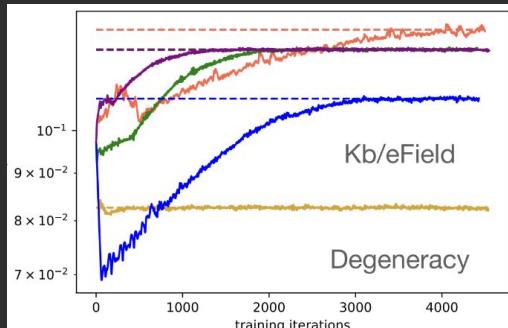
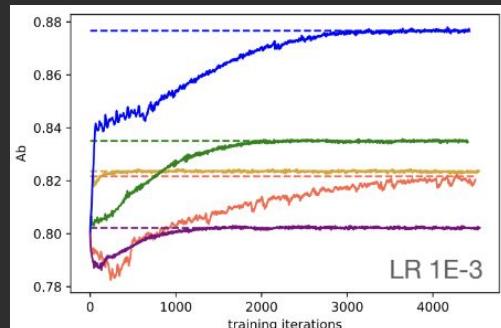


Work credit due (from left):

ML/Math: Youssef N., Sean G., Daniel R.  
neutrino: Yifan C., Roberto S.

Ionization (signal) yield

$$\mathcal{R}_{\text{ICARUS}} = \frac{A_B}{1 + k_B \cdot (dE/dx)/\mathcal{E}}$$



Lots of applications

- Simultaneous multi-parameter fit
- Inter-parameter dependency study
- Automation of calibration workflow
- Inverse imaging (i.e. reconstruction)

# SciML: Applying AI/ML Hiking Skills For Physics Inference

A wide-angle photograph of a desert landscape, likely Death Valley National Park. In the foreground, there are rugged, layered rock formations in shades of brown, tan, and light beige, showing clear sedimentary structures. To the right, a prominent, rounded rock formation with a distinct yellowish-brown hue stands out. In the middle ground, a vast, flat, light-colored valley floor stretches across the horizon. In the background, a range of mountains is visible, with their peaks covered in a layer of white snow, contrasting sharply with the warm tones of the surrounding terrain. The sky above is a bright, clear blue, dotted with scattered, fluffy white clouds.

# ML for Detector Physics Modeling

## SIREN as a differentiable surrogate for optical detectors

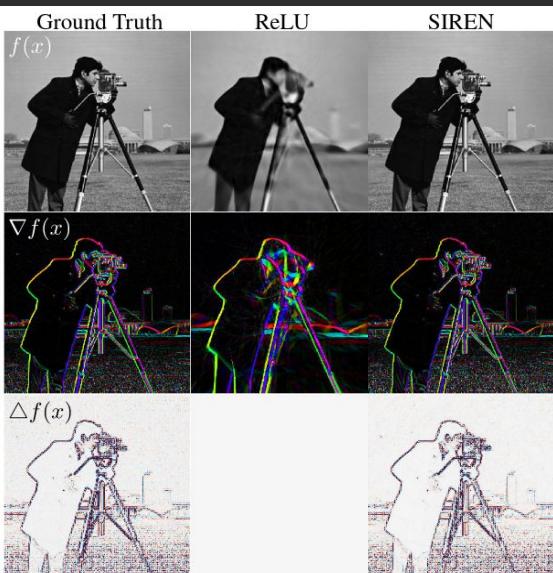
SLAC

### Differentiable Neural Scene Representation

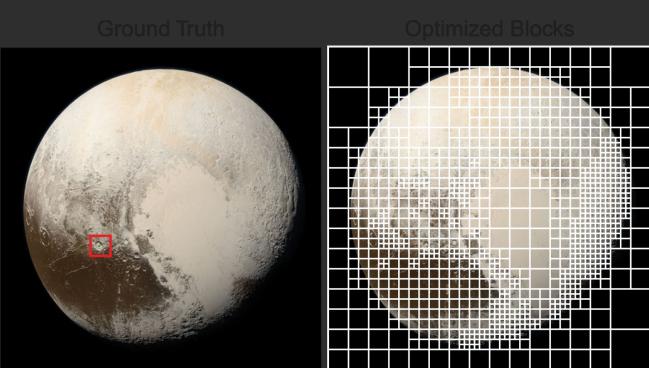


$$(x, y, z, \theta, \phi) \rightarrow \boxed{\text{III}} \rightarrow (RGB\sigma)$$
$$F_{\Theta}$$

NeRF: breakthrough on high resolution image representation by a very simple neural network



SIREN: success of learning the 1st and 2nd order derivatives



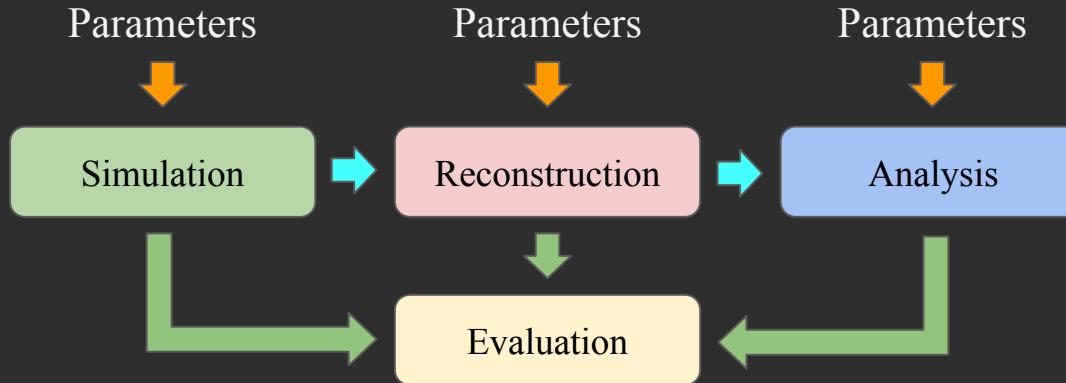
ACORN: scalable version of SIREN by adding spatial feature compression (essentially a learnable kd-tree)

... only a few examples

# Toward Fully AI/ML-based Scientific Workflow



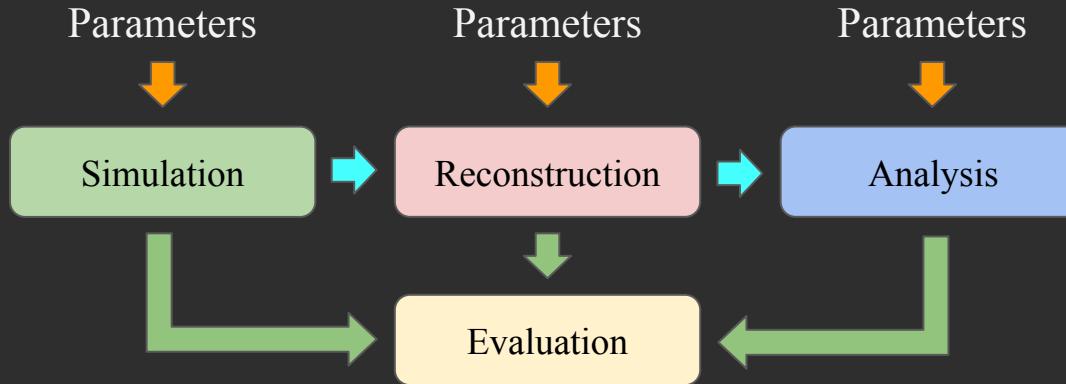
# Toward Fully SciML-based Pipeline



## Enables:

- Automated optimization of model parameters for the entire workflow
- Physics simulation and inference, design optimization, UQ study
- Explainable & interpretable by design where physics knowledge applied

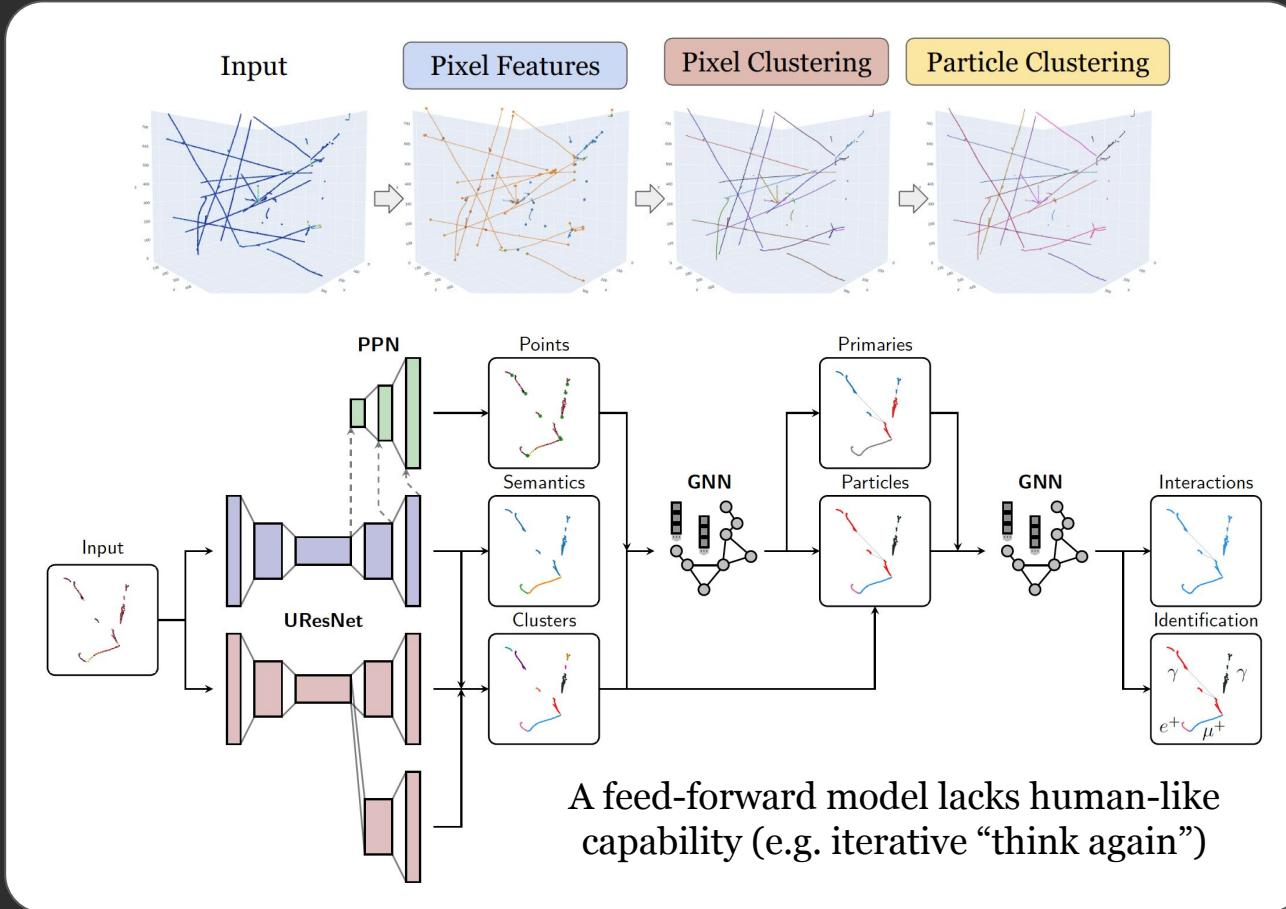
# Toward Fully SciML-based Pipeline



**... yet, lots of challenges!:**

- Non-differentiable operations (e.g. stochastic and/or discrete processes)
- Might suffer from a “missing (new) physics model” in data
- Feed-forward: the model architecture is fixed = cannot “discover”

# Toward Fully AI/ML-based Pipeline



# New AI/ML Tool: Foundation Models

## Self-supervision

Learn from co-occurrence patterns.

"The color of an apple is ."

Can utilize lots of unlabeled data!



## “Representation Learning”

Must learn the universe represented by data



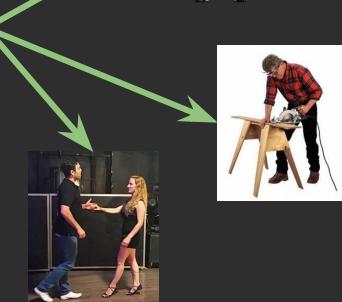
## Transfer Learning

Conceptualizing the world, the model can be adopted to perform various tasks (task-agnostic)

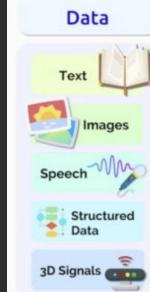
Transfer Learning  
(fine-tuning w/ labels)



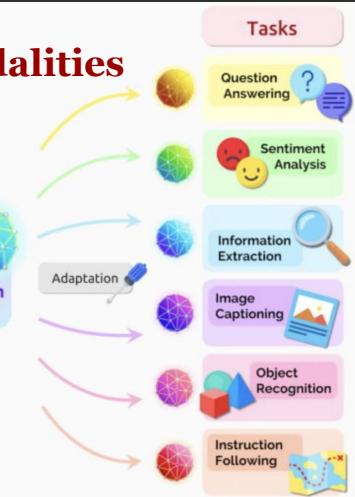
Rep. Learning  
(via self-supervision)



## Multiple data modalities

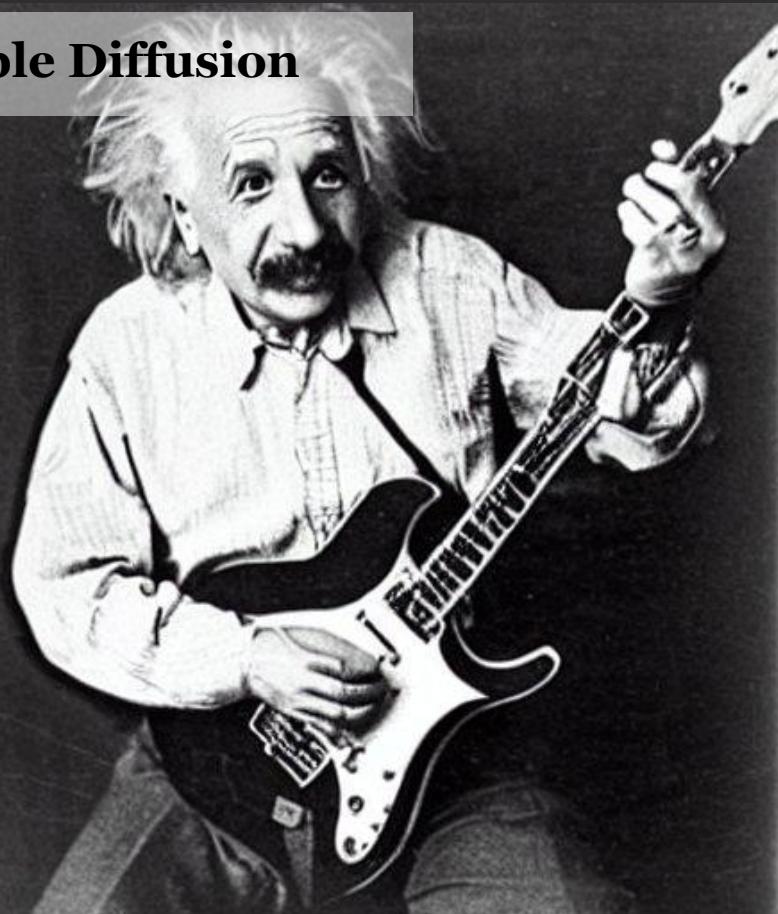


Training



# Foundation Models

## Stable Diffusion



Let me ask questions about logic and commonsense



Sure, I'll do my best to answer any questions you have about logic and common sense.  
Please go ahead and ask!



I give you a premise followed by a question.  
Tell me your answer, either Yes or No, followed by your reasoning.



Sure, I'll do my best to answer your questions based on the given premise. Please go ahead and provide the premise and question.



Premise: I have not had food.  
Question: Have I had fried rice?

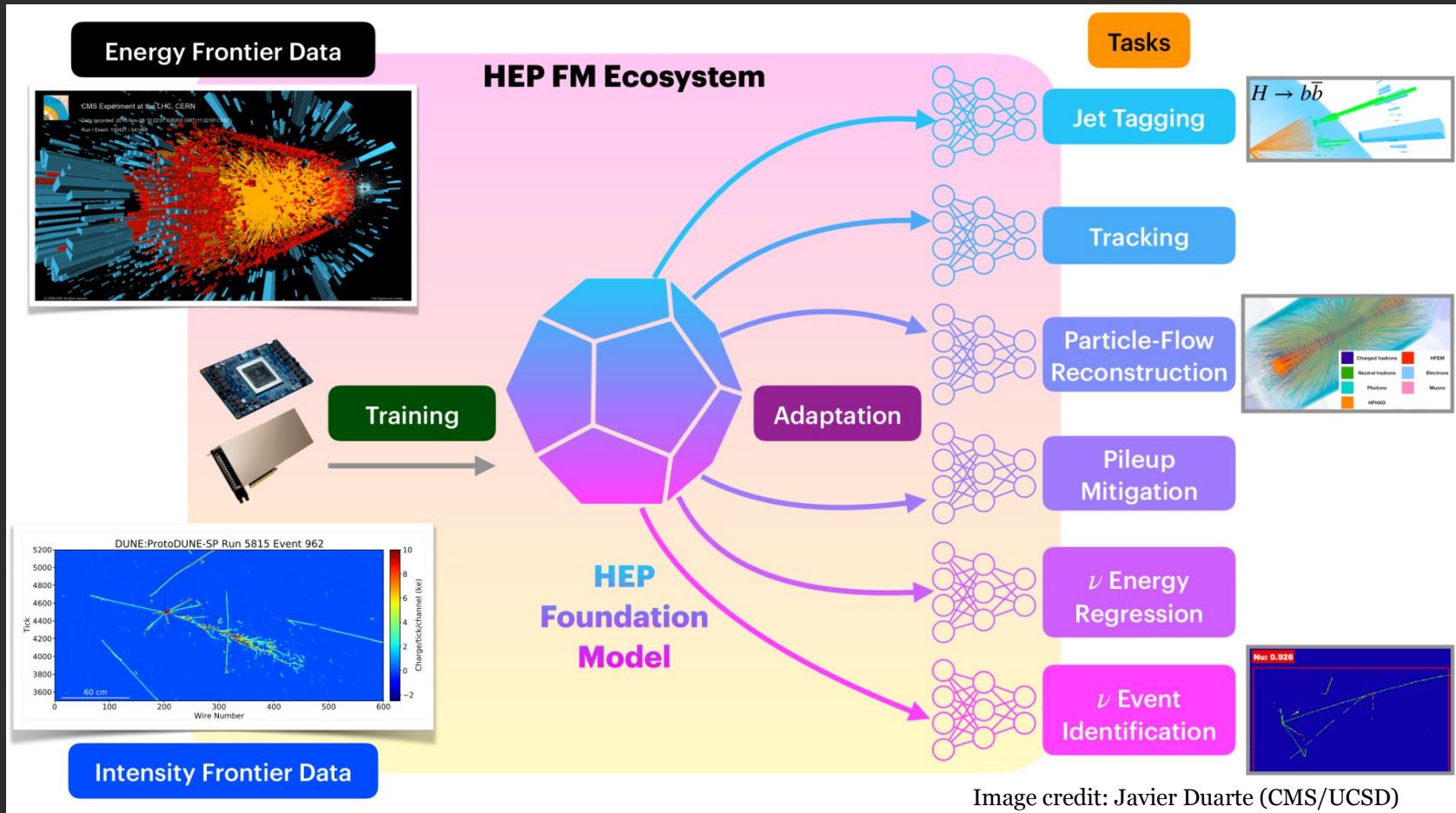


Based on the premise that you have not had any food, my answer would be "No, you have not had fried rice."

My reasoning is that if you have not had any food, it would be impossible for you to have had fried rice, which is a type of food.

## (Chat-) GPT

# Example: A shared “detector physics” model





# Landscape of AI/ML and Science

# Future Directions of AI/ML and Science

## **SciML: science for AI/ML**

- Encode within mathematical operations, loss functions, model architectures
- Differentiable physics models and inference methods

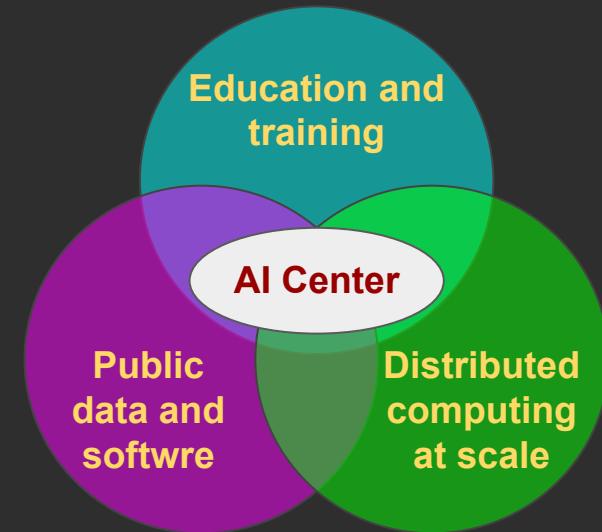
Physics projects (should) contribute to AI/ML research!

## **Foundation models: AI/ML for science**

- Learn, discover, and explain scientific insights in data

**What is the right way to pursue?**

# Closing Note: Ecosystem for AI/ML Research



## HEP Ecosystem for AI research

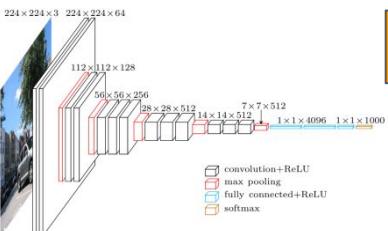
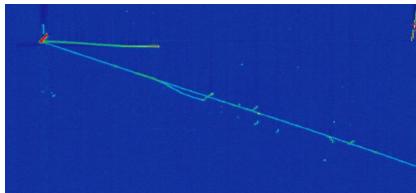
- Accessible **education and training** at all levels
  - **Interdisciplinary research environment** (e.g. workshops, hackathons, visiting scholars program)
  - **Shared computing resources** available always
  - **Open and reusable datasets and software** with documentation and performance metrics
- 
- **Ethics of AI**: how to ensure diversity, equity, and inclusion which is already terrible in STEM? How to ensure small and large projects both benefit from AI?

**AI is an accelerator.** It is coming. Don't avoid.  
**Participate to make sure the use is good.**

# Foundation Models

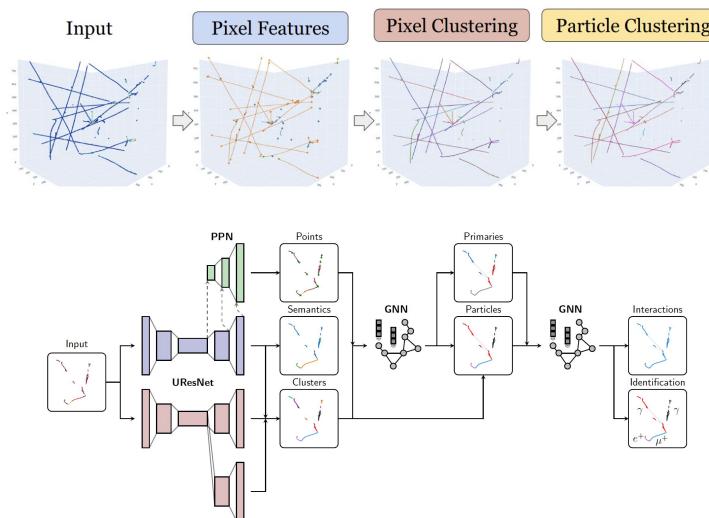
Can supersede multi-task cascade models (e.g. “end-to-end pipeline”) by learning the entire representation in data including hierarchy of features

Single Deep NN



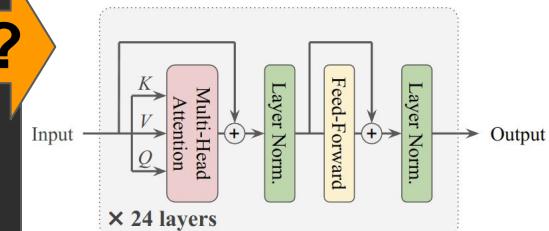
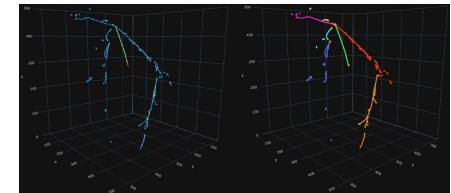
Single task & one model,  
a big black-box

Multi-task, Multi-modal, Composite DNNs



Structured by domain knowledge, some explainability, but limited by a fixed structure.

Transformer (Deep NN)



Task-agnostic, one big black-box but possible to learn more than a fixed model

# Human-in-the-Loop (RLHF)

[OpenAI ChatGPT blog post](#)

Step 1

**Collect demonstration data and train a supervised policy.**

A prompt is sampled from our prompt dataset.

Explain reinforcement learning to a 6 year old.



We give treats and punishments to teach...



SFT  
Once upon a time...

A labeler demonstrates the desired output behavior.

Once upon a time...

This data is used to fine-tune GPT-3.5 with supervised learning.

Step 2

**Collect comparison data and train a reward model.**

A prompt and several model outputs are sampled.

Explain reinforcement learning to a 6 year old.

A  
In reinforcement learning, the agent is...

B  
Explain rewards...

C  
In machine learning...

D  
We give treats and punishments to teach...



A labeler ranks the outputs from best to worst.

D > C > A > B

RM  
 $r_k$

This data is used to train our reward model.

Step 3

**Optimize a policy against the reward model using the PPO reinforcement learning algorithm.**

A new prompt is sampled from the dataset.

Write a story about otters.



PPO

The PPO model is initialized from the supervised policy.

The policy generates an output.

Once upon a time...



RM  
 $r_k$

The reward model calculates a reward for the output.

The reward is used to update the policy using PPO.

# Human-in-the-Loop (RLHF)

[OpenAI ChatGPT blog post](#)

Step 1

Collect demonstration data and train a supervised policy.

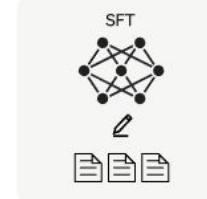
A prompt is sampled from our prompt dataset.

Explain reinforcement learning to a 6 year old.

Human binary classification (good/bad)



We give treats and punishments to teach...



This data is used to fine-tune GPT-3.5 with supervised learning.

Step 2

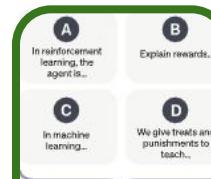
Collect comparison data and train a reward model.

A prompt and several model outputs are sampled.

Explain reinforcement learning to a 6 year old.

Ranking of samples by humans

A labeler ranks the outputs from best to worst.



D > C > A > B

This data is used to train our reward model.



D > C > A > B

Step 3

Optimize a policy against the reward model using the PPO reinforcement learning algorithm.

A new prompt is sampled from the dataset.

Write a story about otters.



Once upon a time...



r<sub>k</sub>

The PPO model is initialized from the supervised policy.

The policy generates an output.

The reward model calculates a reward for the output.

The reward is used to update the policy using PPO.

# Challenges for Differentiable Simulators

# Differentiable Physics Models

**Challenges:** physics models involve stochastic, discrete operations that are not differentiable as they are.

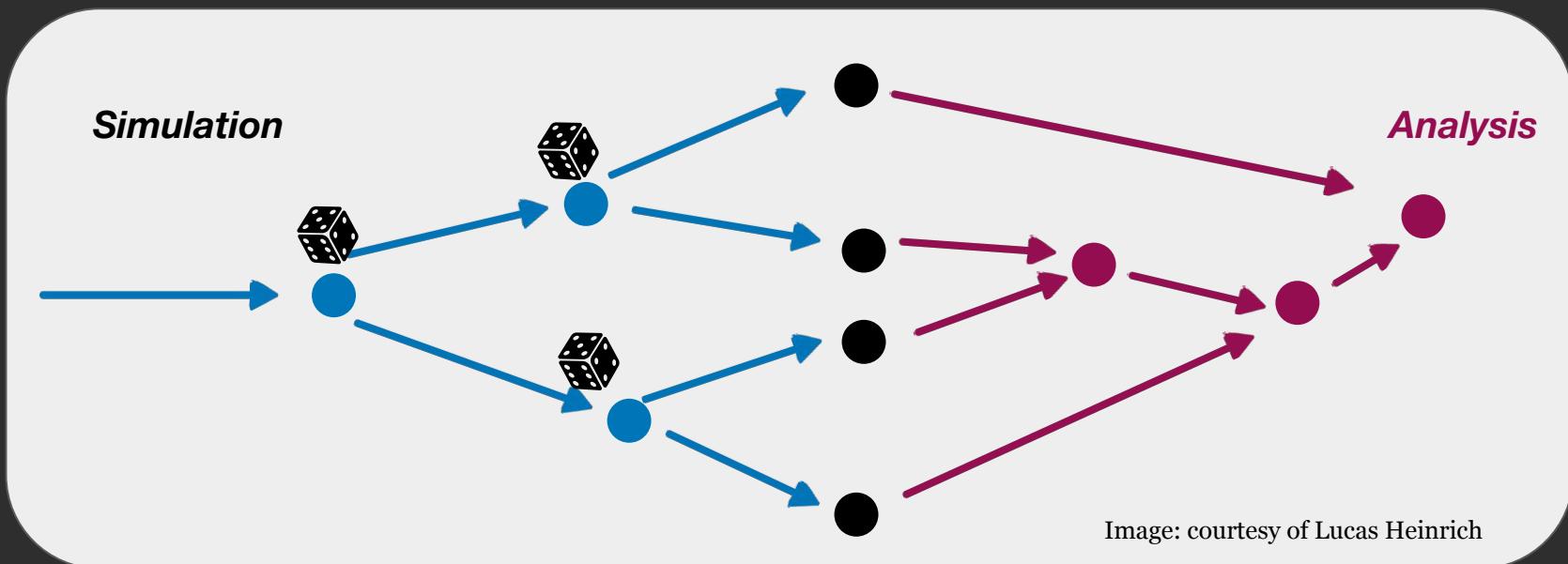


Image: courtesy of Lucas Heinrich

# Differentiable Physics Models

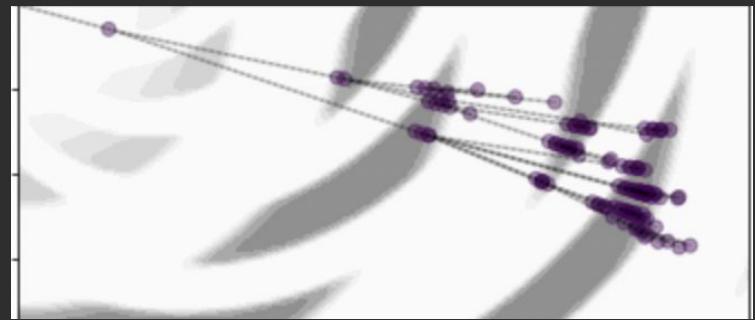
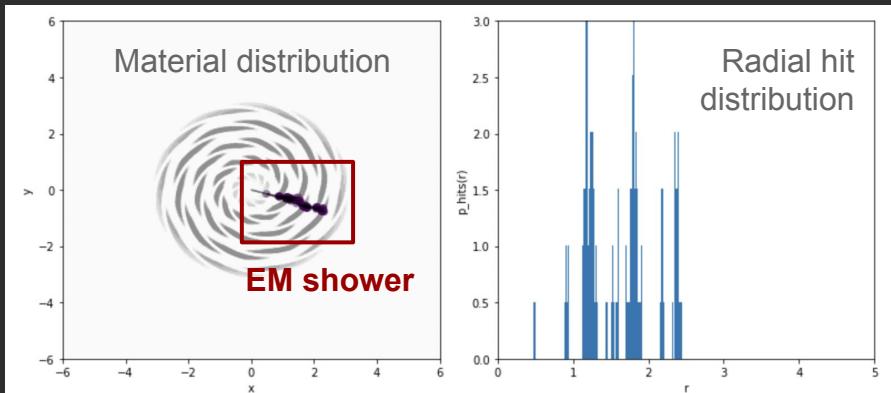
**Challenges:** physics models involve stochastic, discrete operations that are not differentiable as they are. But expectation values over statistics are usually smooth and differentiable (e.g. AI playing a game)



# Differentiable Physics Models

**Challenges:** physics models involve stochastic, discrete operations that are not differentiable as they are. **But expectation values over statistics are usually smooth and differentiable** (e.g. AI playing a game)

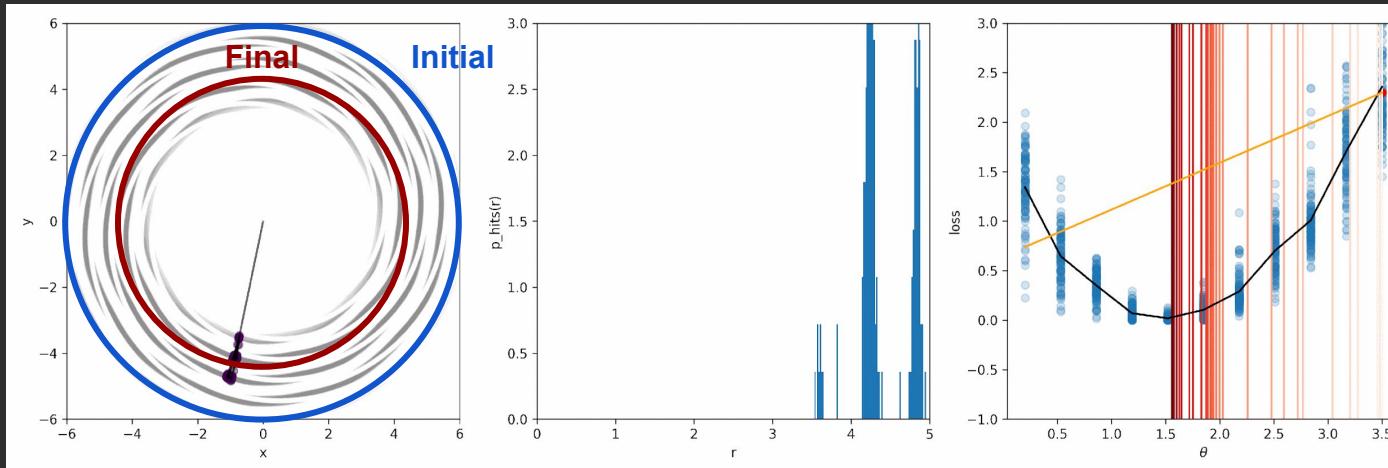
Simple experiment: optimize the calorimeter radius to contain a shower



# Differentiable Physics Models

**Challenges:** physics models involve stochastic, discrete operations that are not differentiable as they are. But expectation values over statistics are usually smooth and differentiable (e.g. AI playing a game)

Simple experiment: optimize the calorimeter radius to contain a shower



“Noisy gradient”  
But it works to  
find the optimal  
radius correctly.

# ML for Analyzing Big Image Data in Neutrino Experiments

## Reconstruction design principles

SLAC

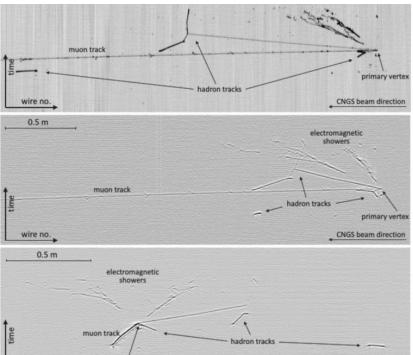
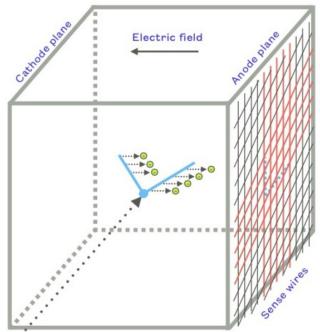
### More About LArTPC Data

# ML for Analyzing Big Image Data in Neutrino Experiments

## Challenges in particle imaging neutrino detectors

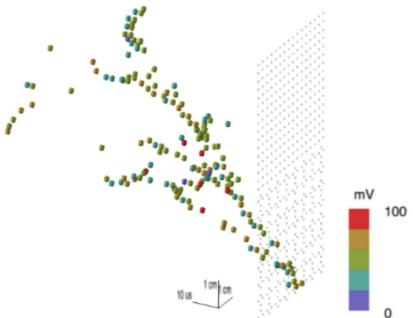
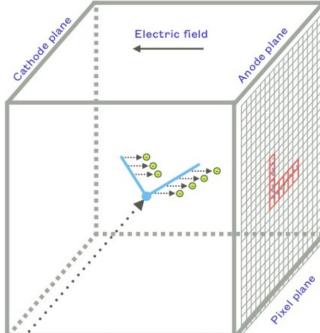
SLAC

Wire TPC (2D)



ICARUS, arXiv:1210.5089

Pixel TPC (3D)



LArPIX, arXiv:1808.02969

**LArTPC** are at the center stage of beam  $\nu$  physics in the US

**Short Baseline Neutrino** program

- $\mu$ BooNE, **ICARUS**, SBND

**DUNE** long-baseline experiment

- **Wire**: DUNE FD
- **Pixel**: DUNE ND-LAr

Advantages:

- **Detailed**:  $O(1)$  mm resolution, precise calorimetry
- **Scalable**: Up to tens of kt

# ML for Analyzing Big Image Data in Neutrino Experiments

## Challenges in particle imaging neutrino detectors

SLAC



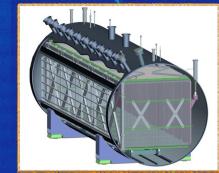
**high resolution,  
big image data  
100 M to giga-pixels**

75 cm

$\nu_\mu$

Run 3493 Event 41075, October 23<sup>rd</sup>, 2015

Liquid Argon TPC  
~mm/pixel spatial resolution  
~100 to 10,000 cubic-meters  
~MeV level sensitivity



MicroBooNE  
~87 ton (school bus)

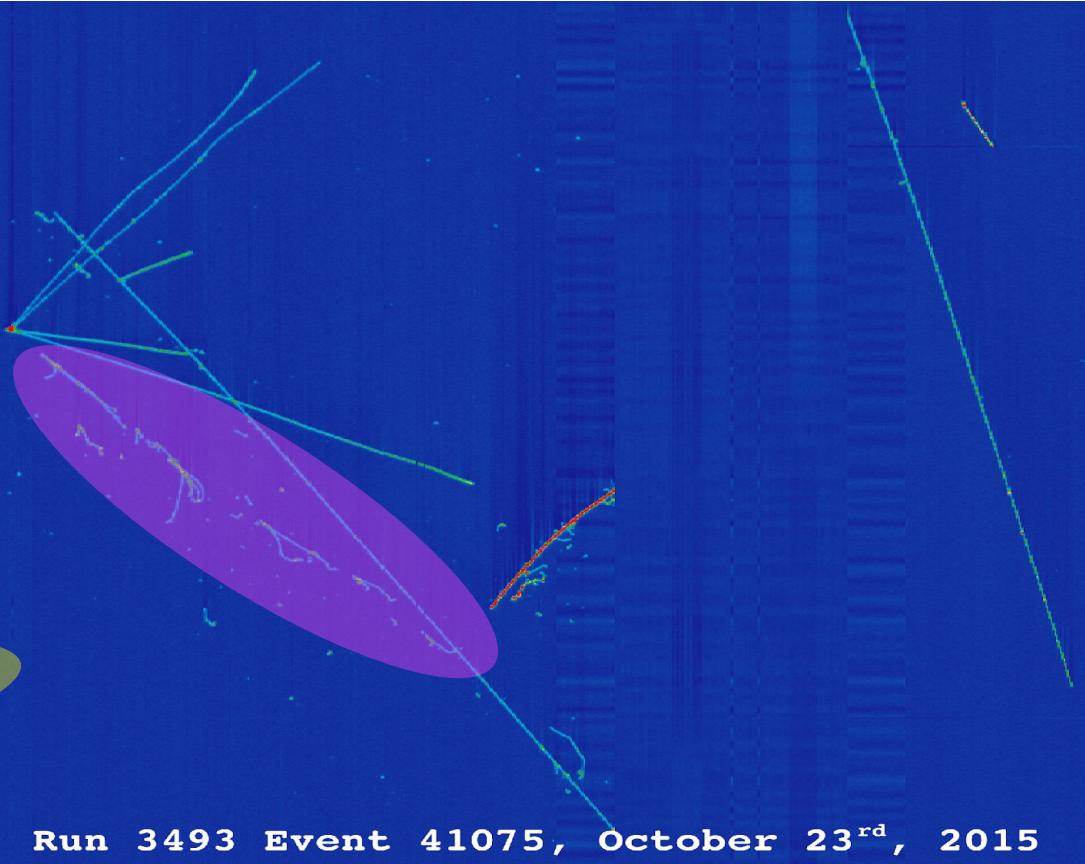
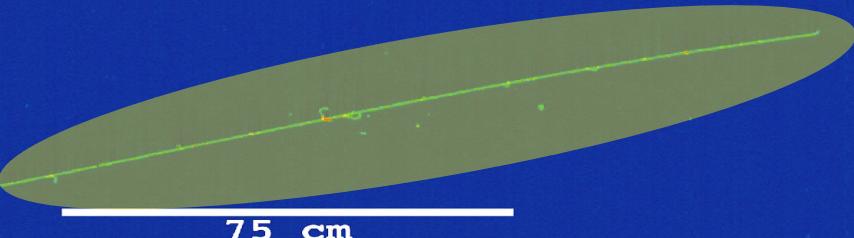
# ML for Analyzing Big Image Data in Neutrino Experiments

## Challenges in particle imaging neutrino detectors

SLAC

$\mu$ BooNE

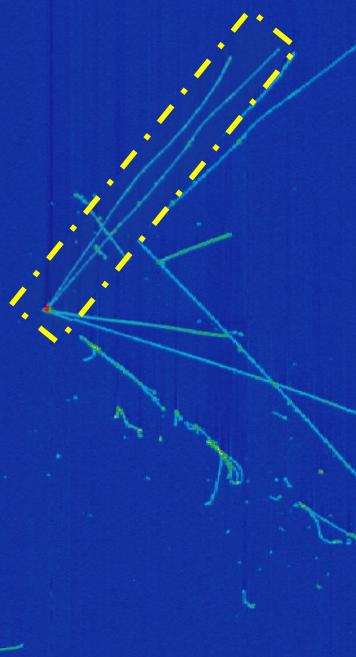
**Distinct shapes**  
“track” v.s. “shower”  
particle trajectories



# ML for Analyzing Big Image Data in Neutrino Experiments

## Challenges in particle imaging neutrino detectors

SLAC



**Kinks and wiggles**  
microscopic kinks tell  
particle momentum

75 cm

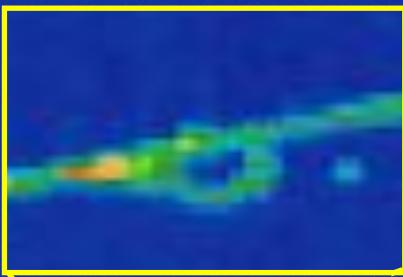
Run 3493 Event 41075, October 23<sup>rd</sup>, 2015

# ML for Analyzing Big Image Data in Neutrino Experiments

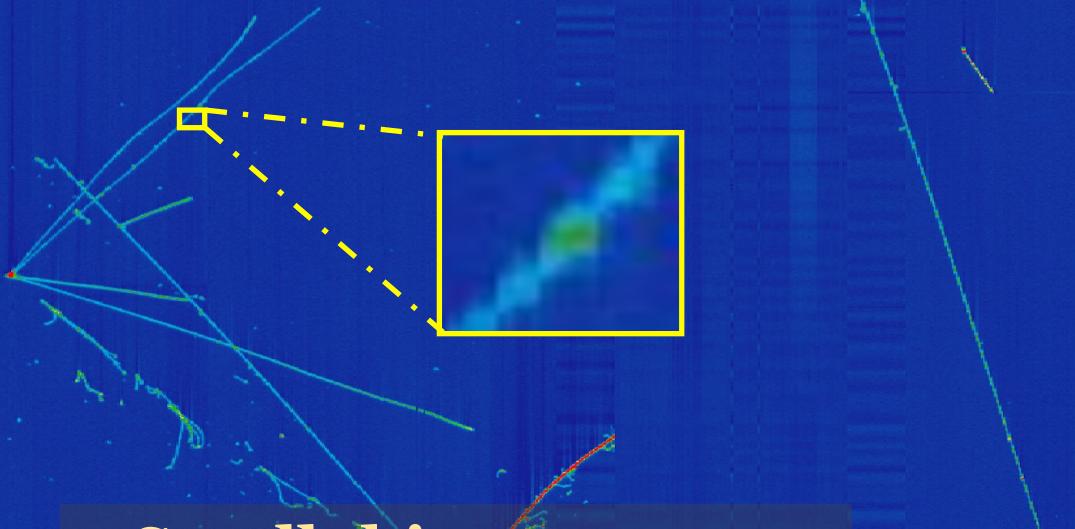
## Challenges in particle imaging neutrino detectors

SLAC

**μBooNE**



75 cm



**Small things matter**  
they inform directions and  
guide global topology

Run 3493 Event 41075, October 23<sup>rd</sup>, 2015

# ML for Analyzing Big Image Data in Neutrino Experiments

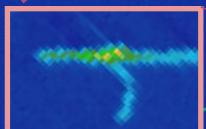
## Challenges in particle imaging neutrino detectors

SLAC



### Color = Energy

Both the absolute and the gradient of colors inform particle energy and type

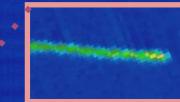


75 cm

Run 3493 Event 41075, October 23<sup>rd</sup>, 2015

e- vs.  $\gamma$   
using  $dE/dX$

Stopping  
particle



# ML for Analyzing Big Image Data in Neutrino Experiments

## Challenges in particle imaging neutrino detectors

SLAC

ArgonCube DUNE-ND 7x5 Modules Configuration Beam Spill

