Laying the Foundation for Complete Automation in Particle Image Inference

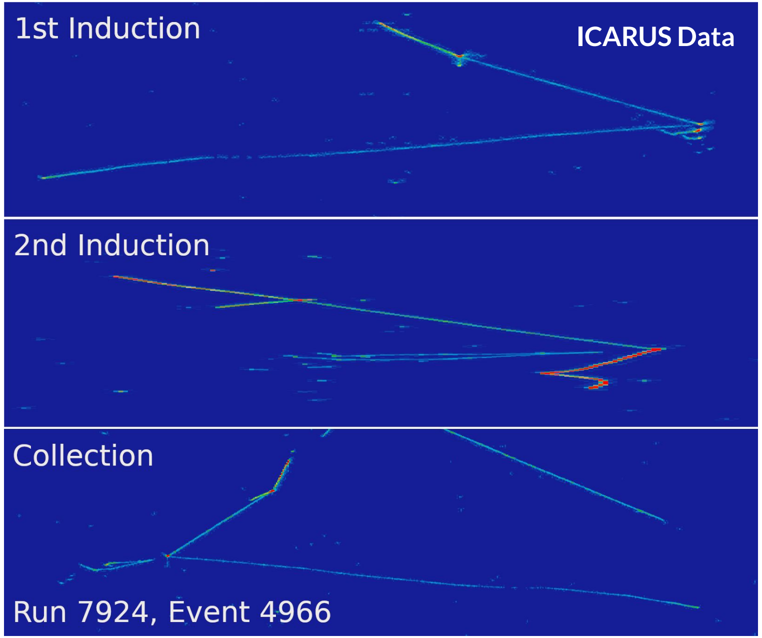
François Drielsma (SLAC)

Panofsky Seminar, Feb. 22nd 2024

SLAC NATIONAL ACCELERATOR LABORATORY
Stanford University
U.S. DEPARTMENT OF ENERGY

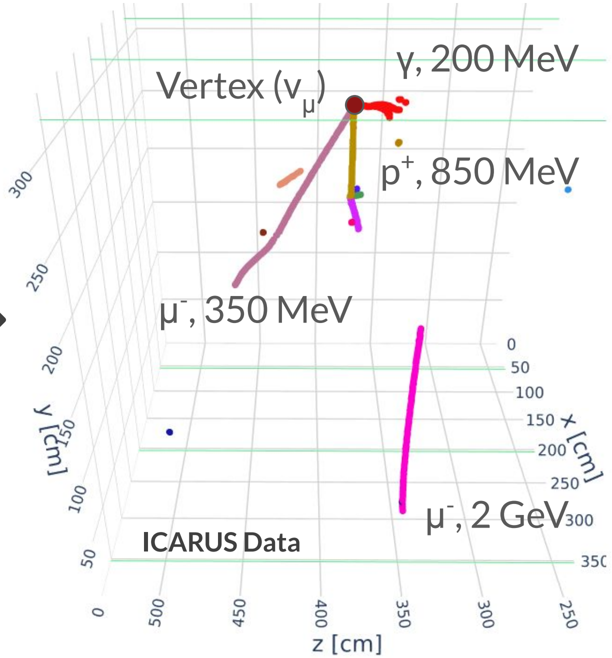
Ultimate Research Goal

What we get (raw data)



AI

What we want (“particle flow”)



→**Goal**: Leverage **Artificial Intelligence (AI)** to **automate** this task

Neutrino Oscillations

Neutrinos are **produced as different types**

- Neutrino types are a **superposition** of **quantum mass states**

Neutrino types

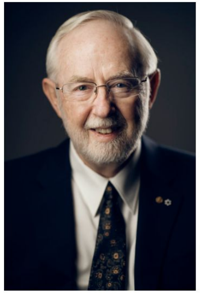
$$\begin{pmatrix} \nu_e \\ \nu_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix}$$

Mass states

Mixing matrix

2015

Takaaki Kajita



Arthur B. McDonald

Neutrino Oscillations

SLAC NATIONAL ACCELERATOR LABORATORY

Neutrinos are **produced as different types**

- Neutrino types are a **superposition** of **quantum mass states**

Neutrino types

$$\begin{pmatrix}\nu_e\\ \nu_\mu\end{pmatrix} = \begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}\nu_1\\ \nu_2\end{pmatrix}$$

Mass states

Mixing matrix

- Mass wavefunctions oscillate at different rate→ **mixture changes**

Amplitude

Mass splitting

$$P(\nu_e \rightarrow \nu_\mu) = \sin^2 2\theta \sin^2\left(\frac{\Delta m^2_{21} L}{4E}\right)$$

Baseline

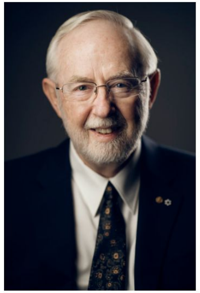
Appearance probability

Neutrino energy

2015

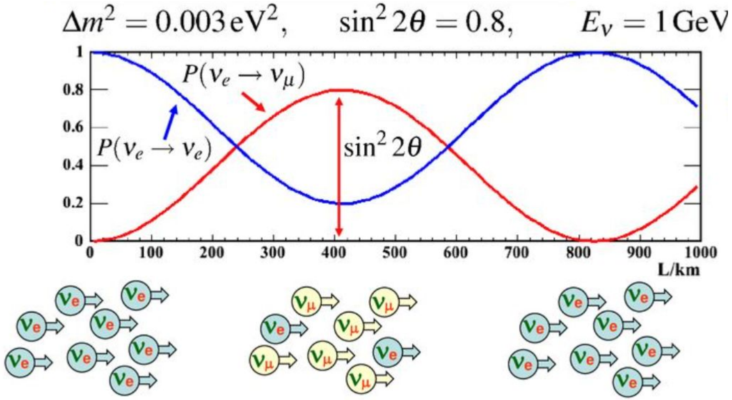
© Nobel Media AB. Photo: A. Mahmoud

Takaaki Kajita



© Nobel Media AB. Photo: A. Mahmoud

Arthur B. McDonald



Neutrino Oscillations

Neutrinos are **produced as different types**

- Neutrino types are a **superposition** of **quantum mass states**

Neutrino types

$$\begin{pmatrix} \nu_e \\ \nu_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix}$$

Mass states

Mixing matrix

- Mass wavefunctions oscillate at different rate→ **mixture changes**

Mass splitting

Amplitude

$$P(\nu_e \to \nu_\mu) = \sin^2 2\theta \sin^2\left(\frac{\Delta m^2_{21} L}{4E}\right)$$

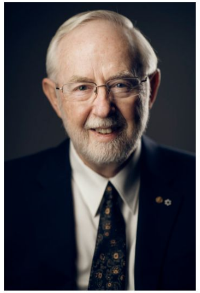
Baseline

Appearance probability

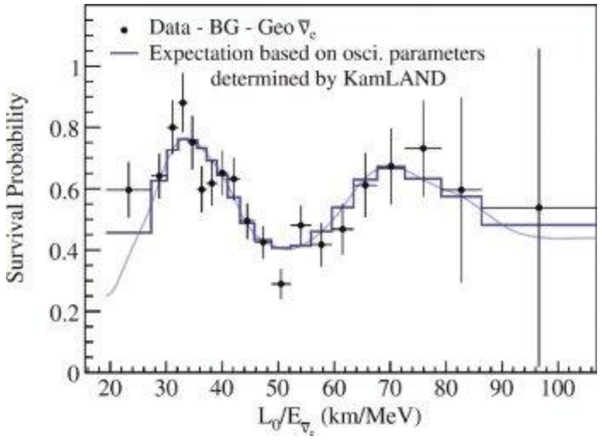
Neutrino energy

2015

Takaaki Kajita



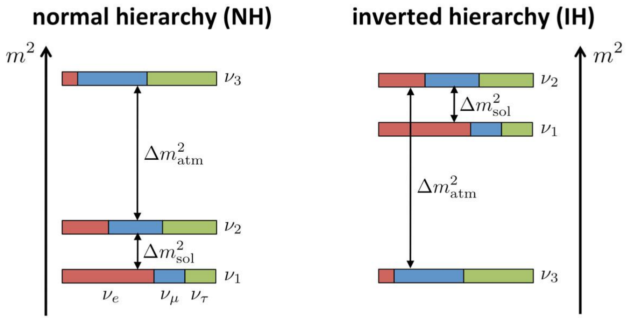
Arthur B. McDonald



What’s Missing?

Several neutrino properties **remain elusive**

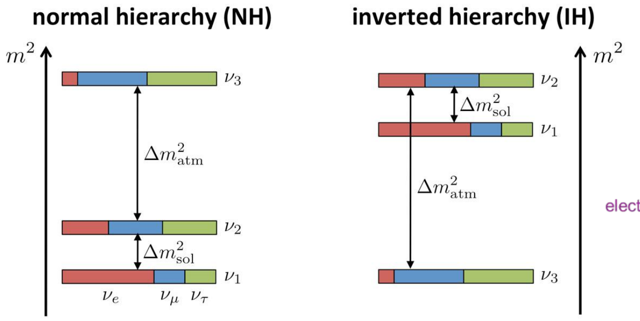
- Neutrino mass ordering

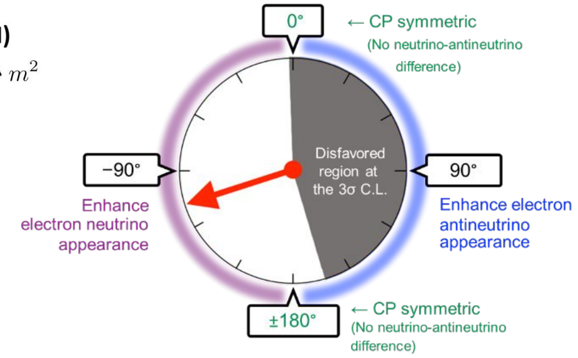


What's Missing?

Several neutrino properties **remain elusive**

- Neutrino mass ordering
- Leptonic CP-violation: **origin of matter-antimatter asymmetry?**





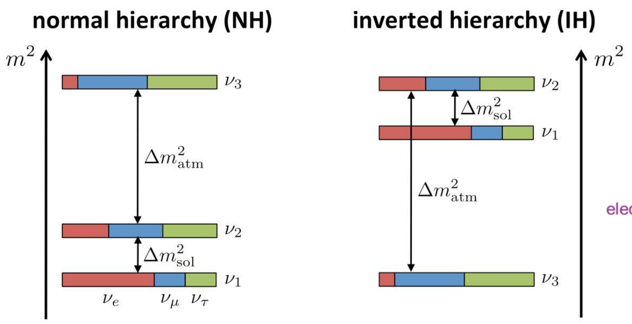
T2K

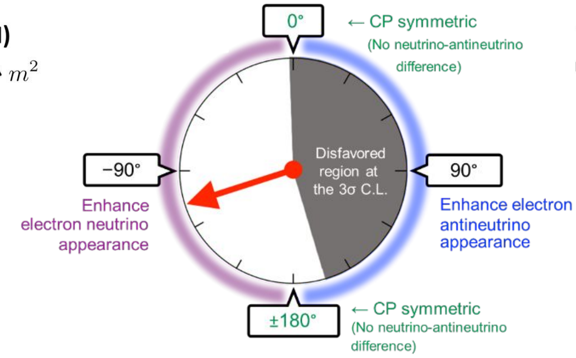
What’s Missing?

Several neutrino properties **remain elusive**

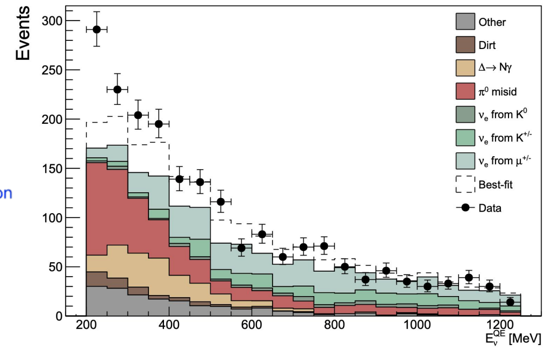
- Neutrino mass ordering
- Leptonic CP-violation: **origin of matter-antimatter asymmetry?**
- Low-energy ν_e excess at short baseline: **new type of neutrino?**

→ **My research** will help **answer these questions**





T2K

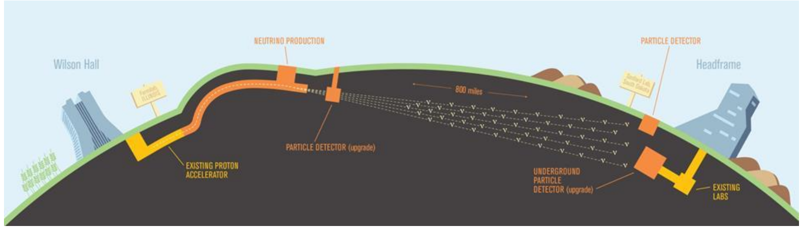


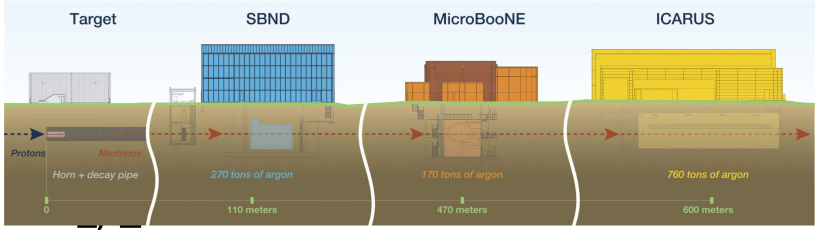
MiniBooNE

DUNE and SBN

Two **US-based neutrino oscillation experiments** to answer these questions

Deep Underground Neutrino Experiment (DUNE), 2028-?	**Short Baseline Neutrino (SBN) program, 2015-2027**
1300 km: **enhance matter effects**	0.6 km: **observe anomalies**
• Mass ordering, CP violation	• New type of neutrino?
• **DUNE**-FD rate: **O(10^3) ν / year**	• **SBN** S/B ratio: **~ O(10^{-5})**





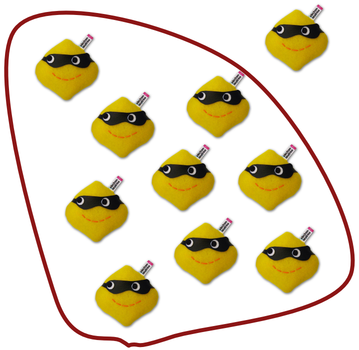
→ Shared detector technology: the **Liquid Argon Time Projection Chamber**

DUNE/SBN Requirements

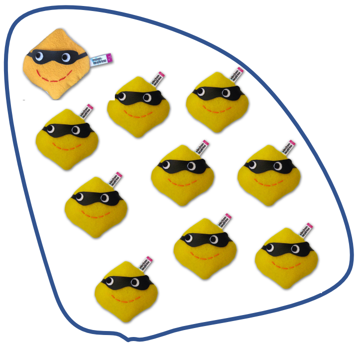
LArTPC requirements
(proposals circa 2015):

- Efficiency for ν_μ ID: **> 90 %**
- Efficiency for ν_e ID: **~ 80 %**
- Purity for both: ~ 85-90 %

ν_e efficiency



ν_e purity

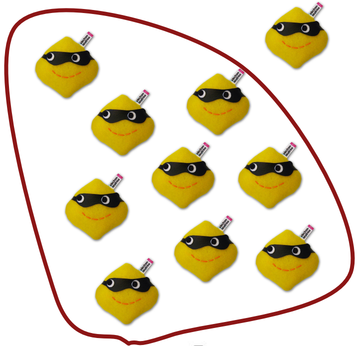


DUNE/SBN Requirements

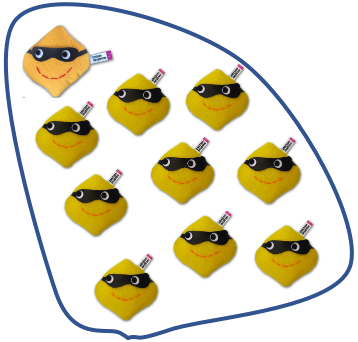
LArTPC requirements (proposals circa 2015):

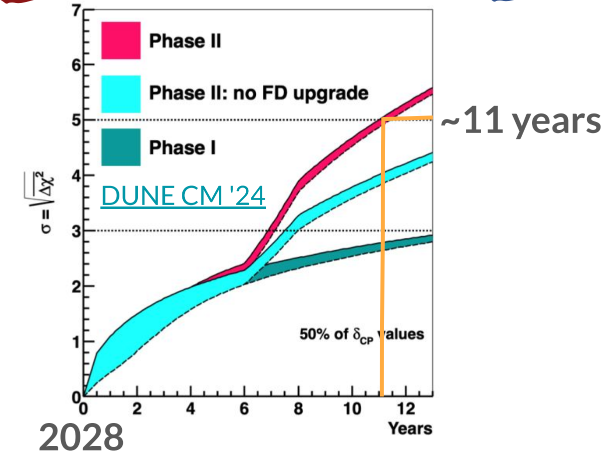
- Efficiency for ν_μ ID: **> 90 %**
- Efficiency for ν_e ID: **~ 80 %**
- Purity for both: ~ 85-90 %

ν_e efficiency



ν_e purity





DUNE/SBN Requirements

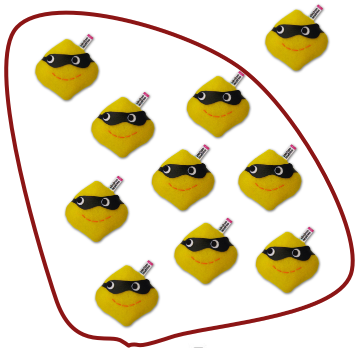
LArTPC requirements
(proposals circa 2015):

- Efficiency for ν_μ ID: **> 90 %**
- Efficiency for ν_e ID: **~ 80 %**
- Purity for both: ~ 85-90 %

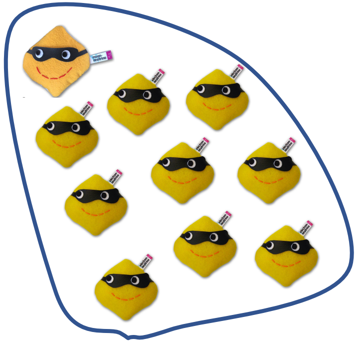
What if efficiency drops?

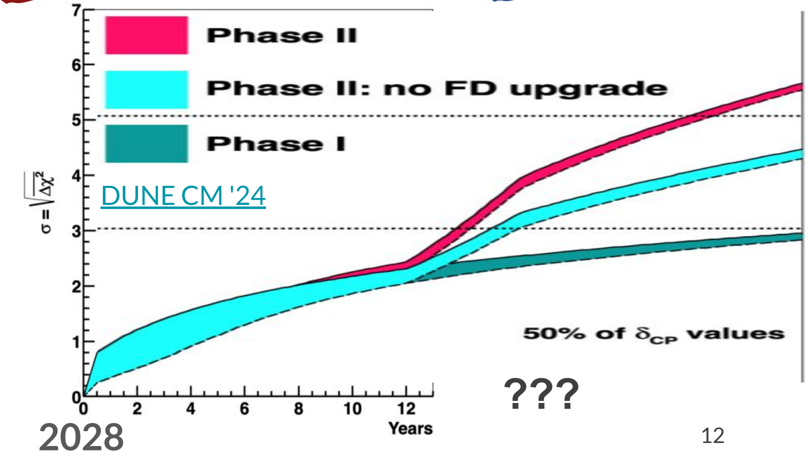
- **Less effective exposure**
- **Less sensitivity**

ν_e efficiency



ν_e purity





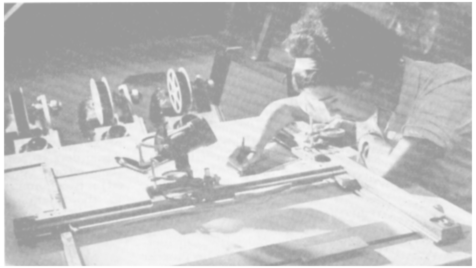
No Reconstruction, No Physics

DUNE/SBN **cannot deliver physics** without **a reliable reconstruction…**

No Reconstruction, No Physics

DUNE/SBN **cannot deliver physics** without **a reliable reconstruction…**

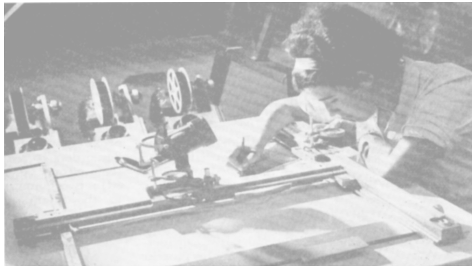
1. **The Era of Humans**
    - What has been the traditional approach in particle imaging detector? Has it been successful?



No Reconstruction, No Physics

DUNE/SBN **cannot deliver physics** without **a reliable reconstruction…**

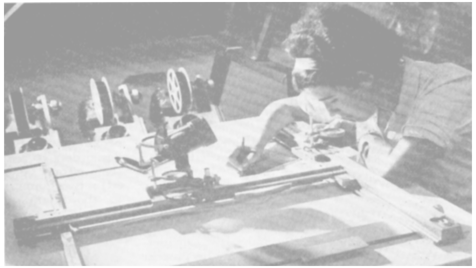
1. **The Era of Humans**
    - What has been the traditional approach in particle imaging detector? Has it been successful?
2. **Cybernetic Augmentation**
    - How have **Machine Learning** tools helped improve reconstruction so far?

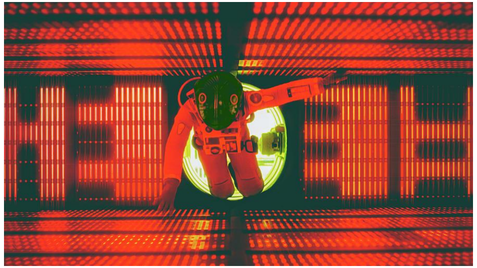


No Reconstruction, No Physics

DUNE/SBN **cannot deliver physics** without **a reliable reconstruction…**

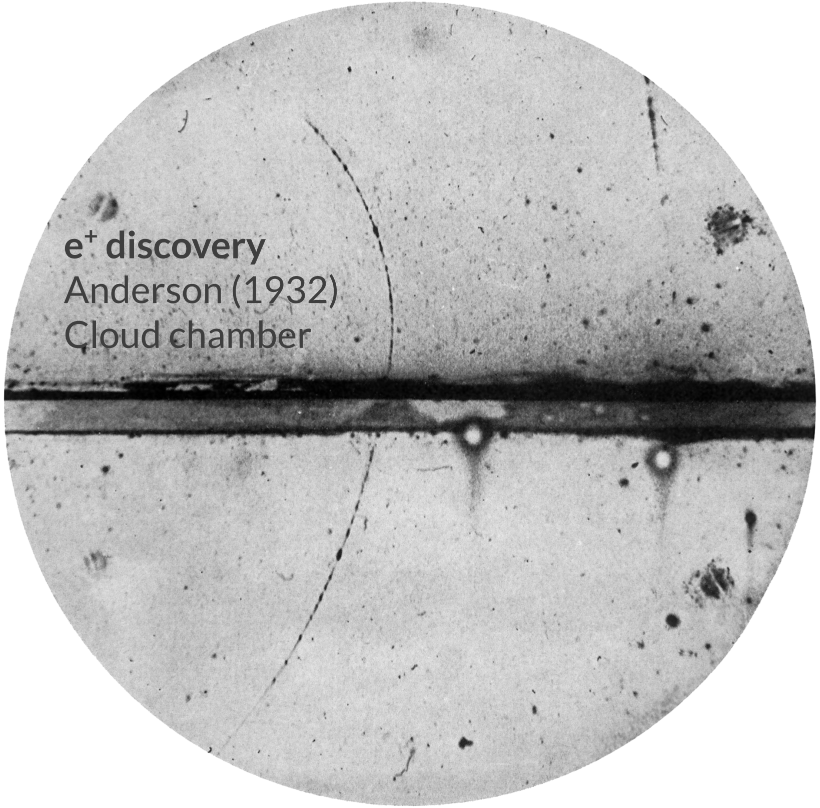
1. **The Era of Humans**
   - What has been the traditional approach in particle imaging detector? Has it been successful?
2. **Cybernetic Augmentation**
   - How have **Machine Learning** tools helped improve reconstruction so far?
3. **The Age of Machines**
   - How can **Artificial Intelligence** definitively solve the issue of automation in particle imaging inference?





1. The Era of Humans

Particle Imaging Detectors



imaging

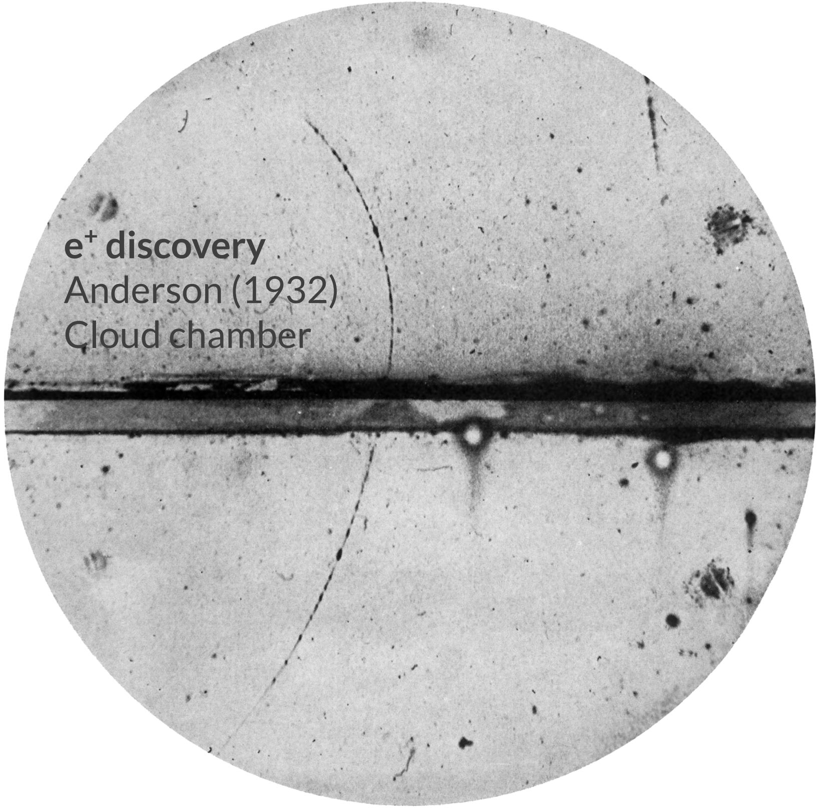
noun [U] • COMPUTING • specialized

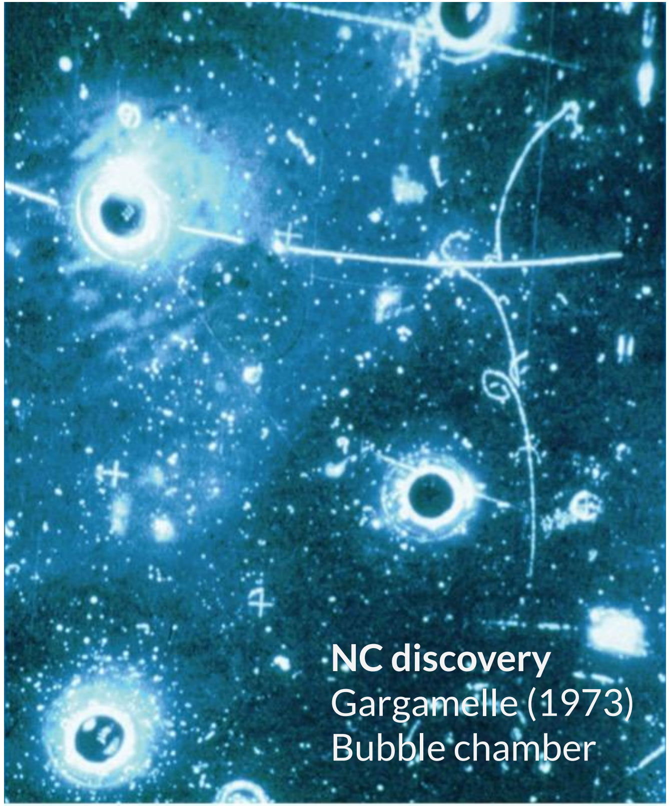
US /ɪˈmɪdʒ.ɪŋ/ UK /ɪˈmɪdʒ.ɪŋ/

the process of producing an exact picture of something

https://dictionary.cambridge.org/us/dictionary/english/imaging

Particle Imaging Detectors





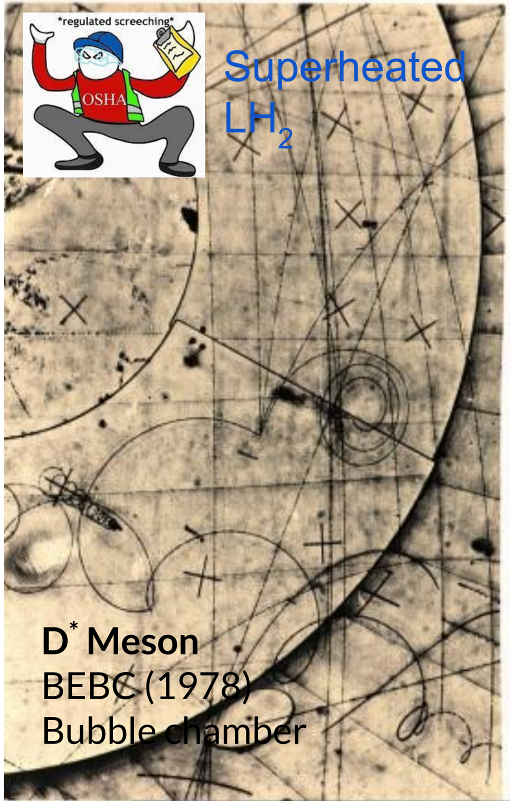
Particle Imaging Detectors Reconstruction

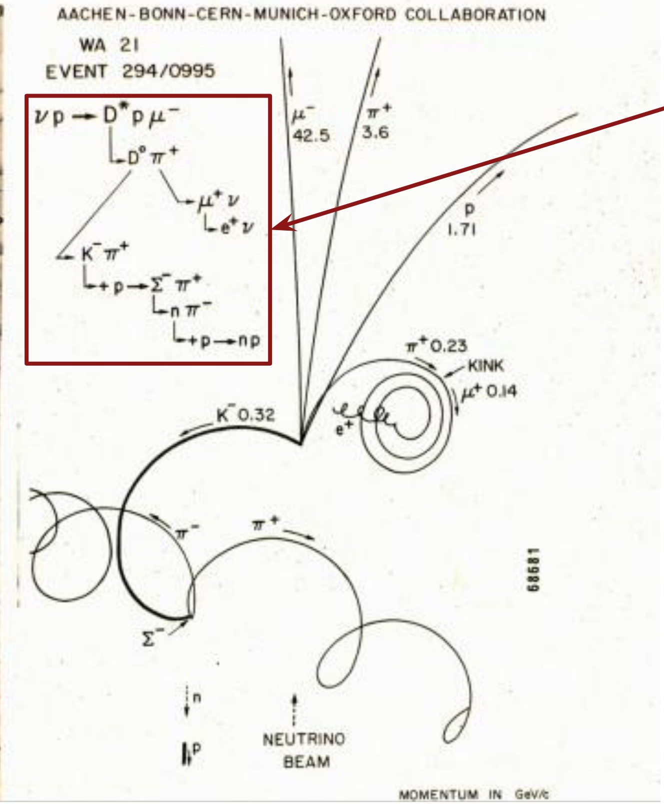
Step-by-step approach:

1. Identify interesting events in pictures **by eye**
   a. Scanning experts
   b. Grad students
2. Trace particles on paper **by hand**
3. Estimate particle kinematics
4. Review by senior physicist


Anita Bjorkebo
Gargamelle (1969)

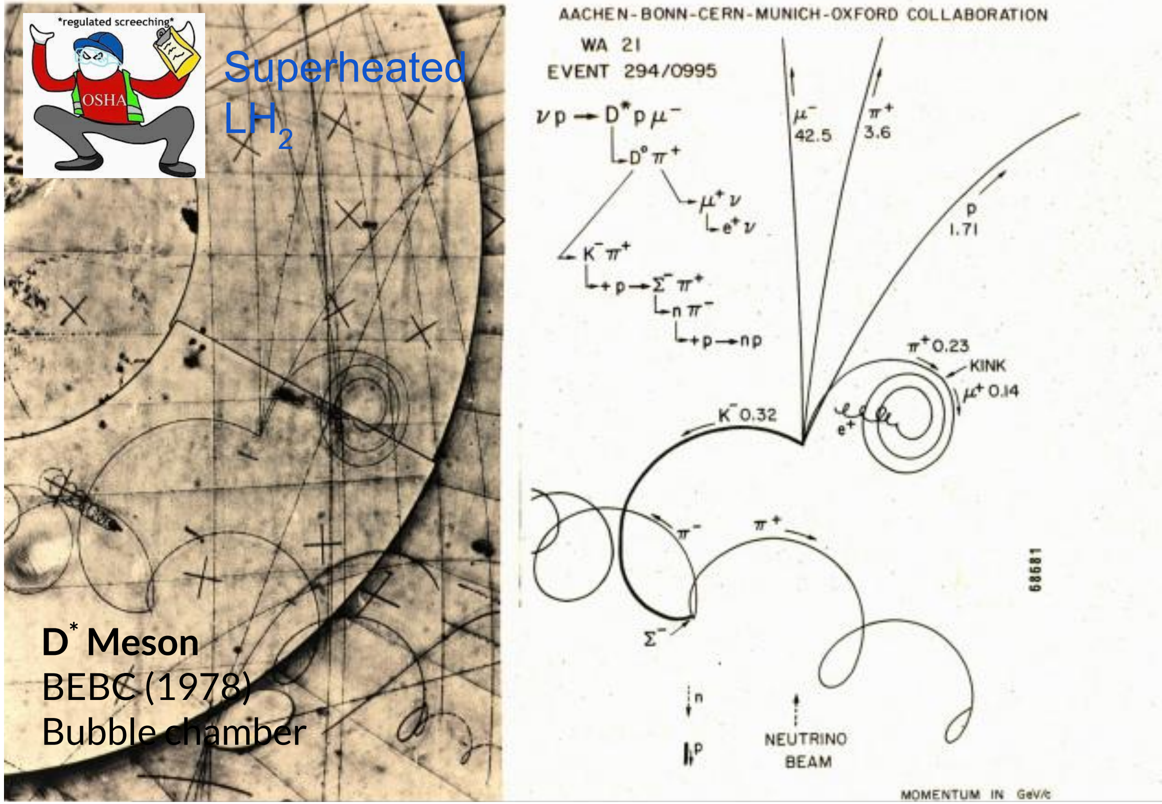
Particle Imaging Detectors Reconstruction





Full “particle flow”

Particle Imaging Detectors Reconstruction



Advantages:

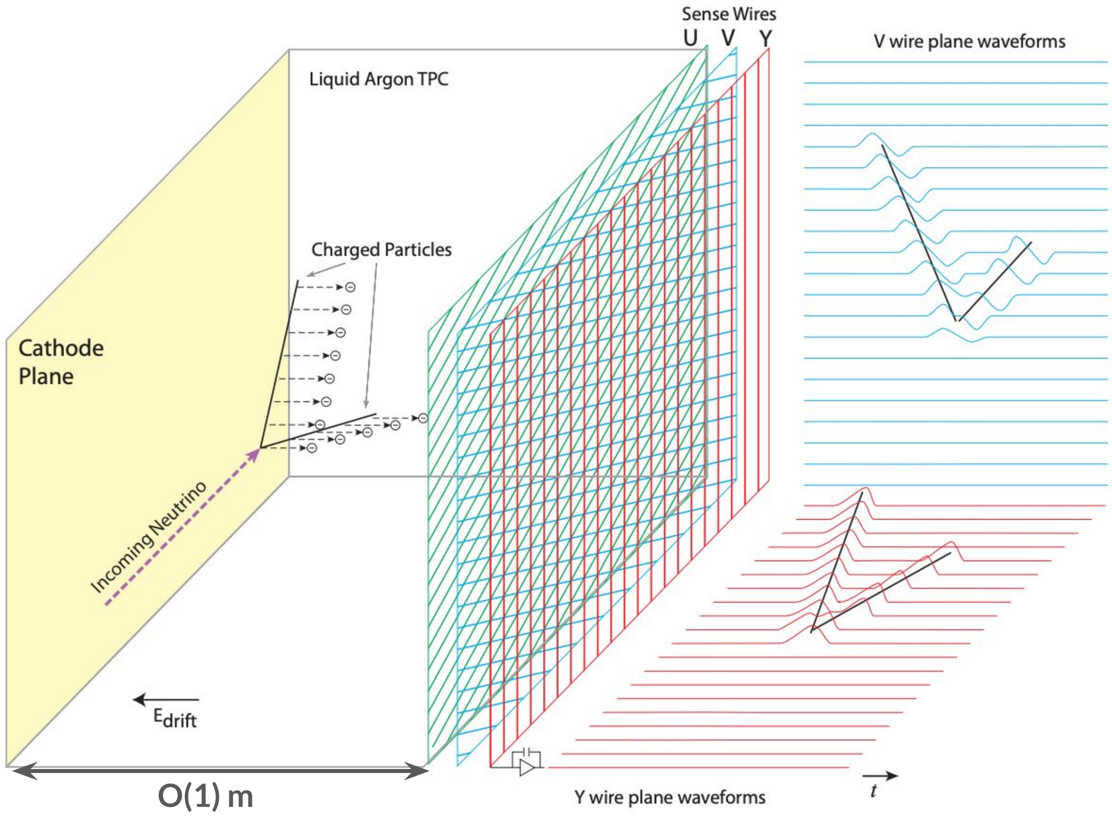
- Very detailed interaction

Limitations:

- **Time intensive reconstruction**
- **Hard to scale**
  - Largest: 15 m^3
  - Slow (~ seconds)

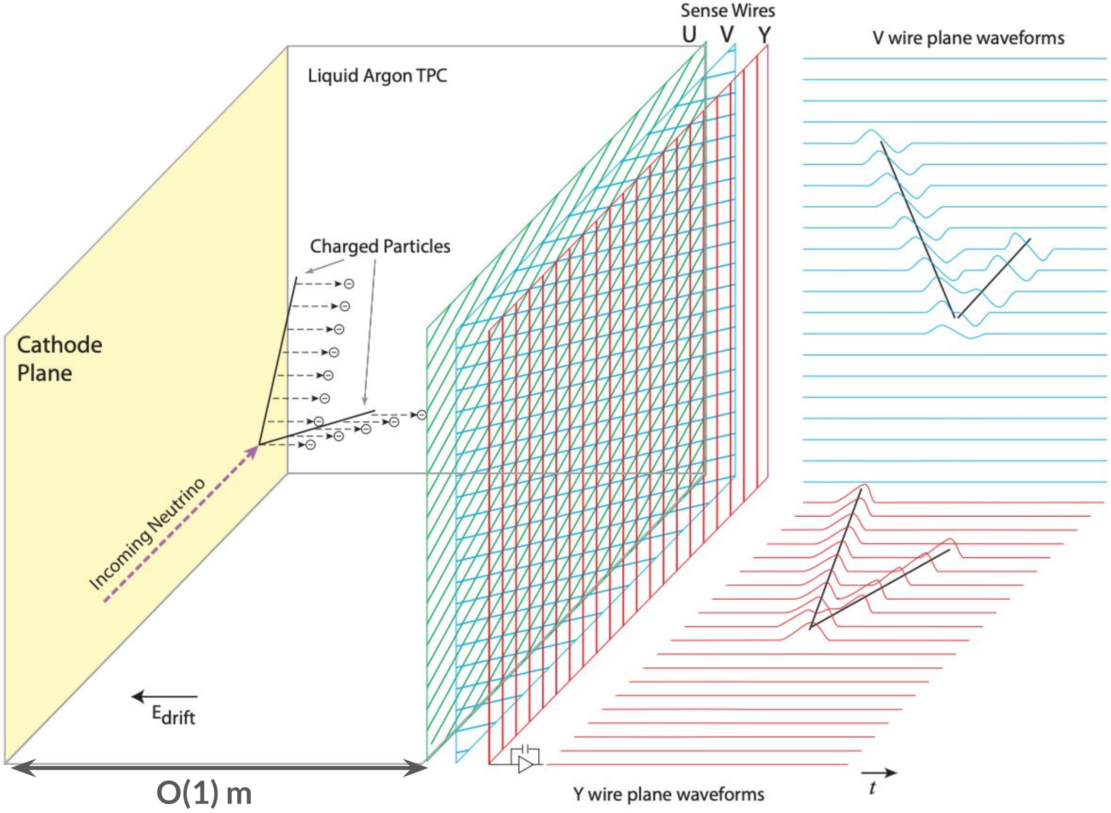
Neutrino interactions are **rare**… 1 LY of Pb for 50 % chance

Liquid Argon Time Projection Chamber



LArTPC: Detector used today in **DUNE** and **SBN**

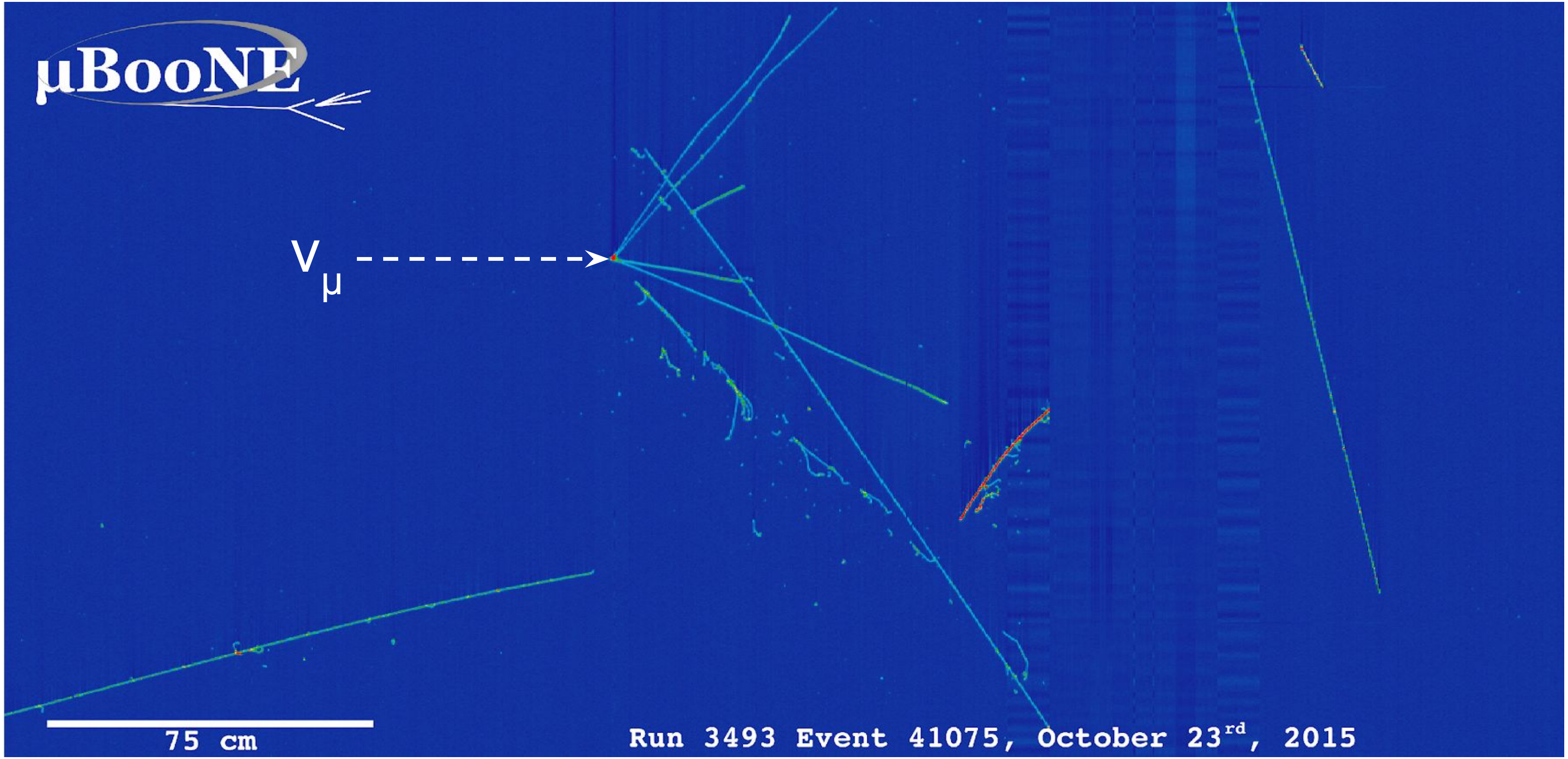
Liquid Argon Time Projection Chamber



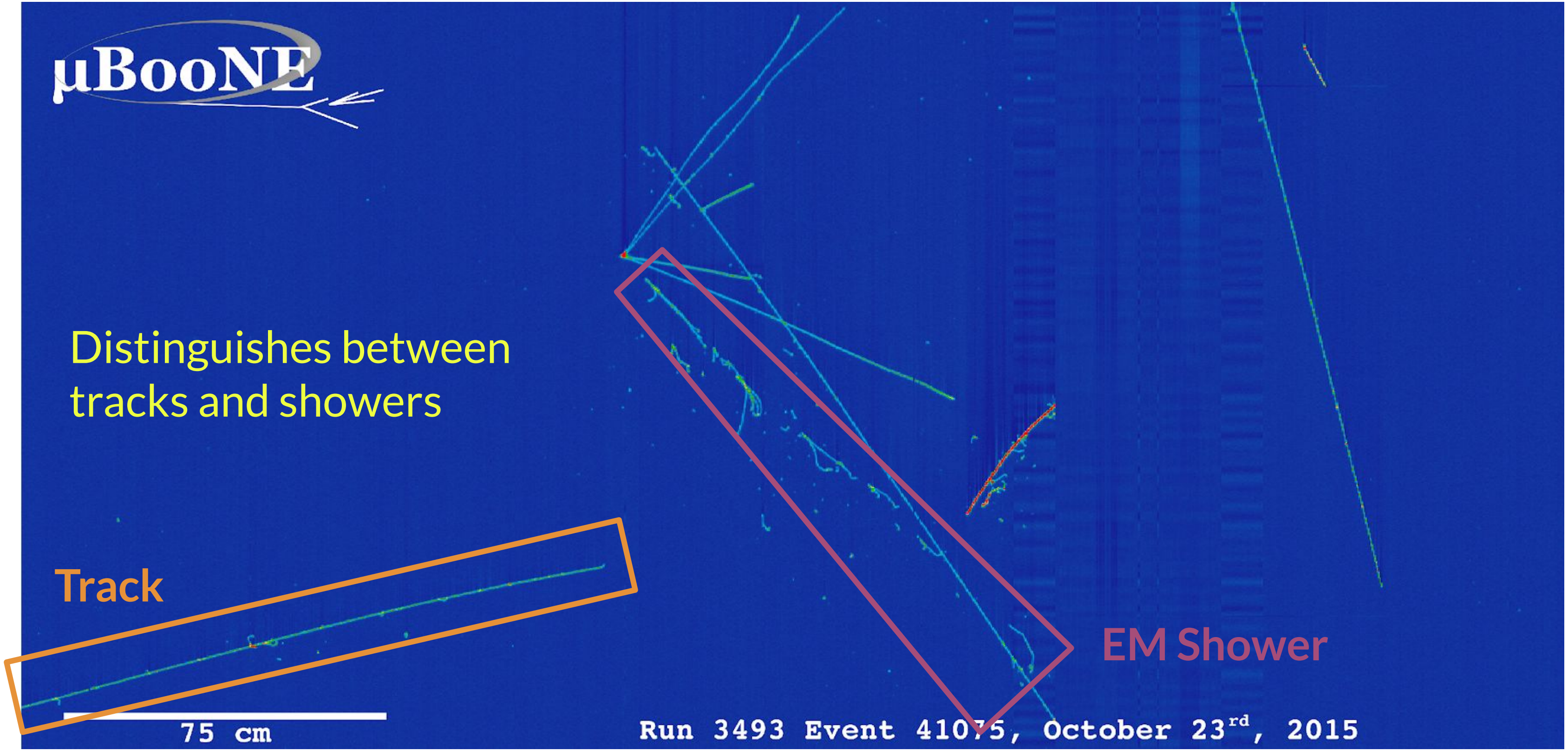
LArTPC: Detector used today in **DUNE** and **SBN**:

- Precision of its ancestors
- Dense (1.4 g/cm^3)
- Cheap (~ 1$/kg), **a.k.a. scalable**

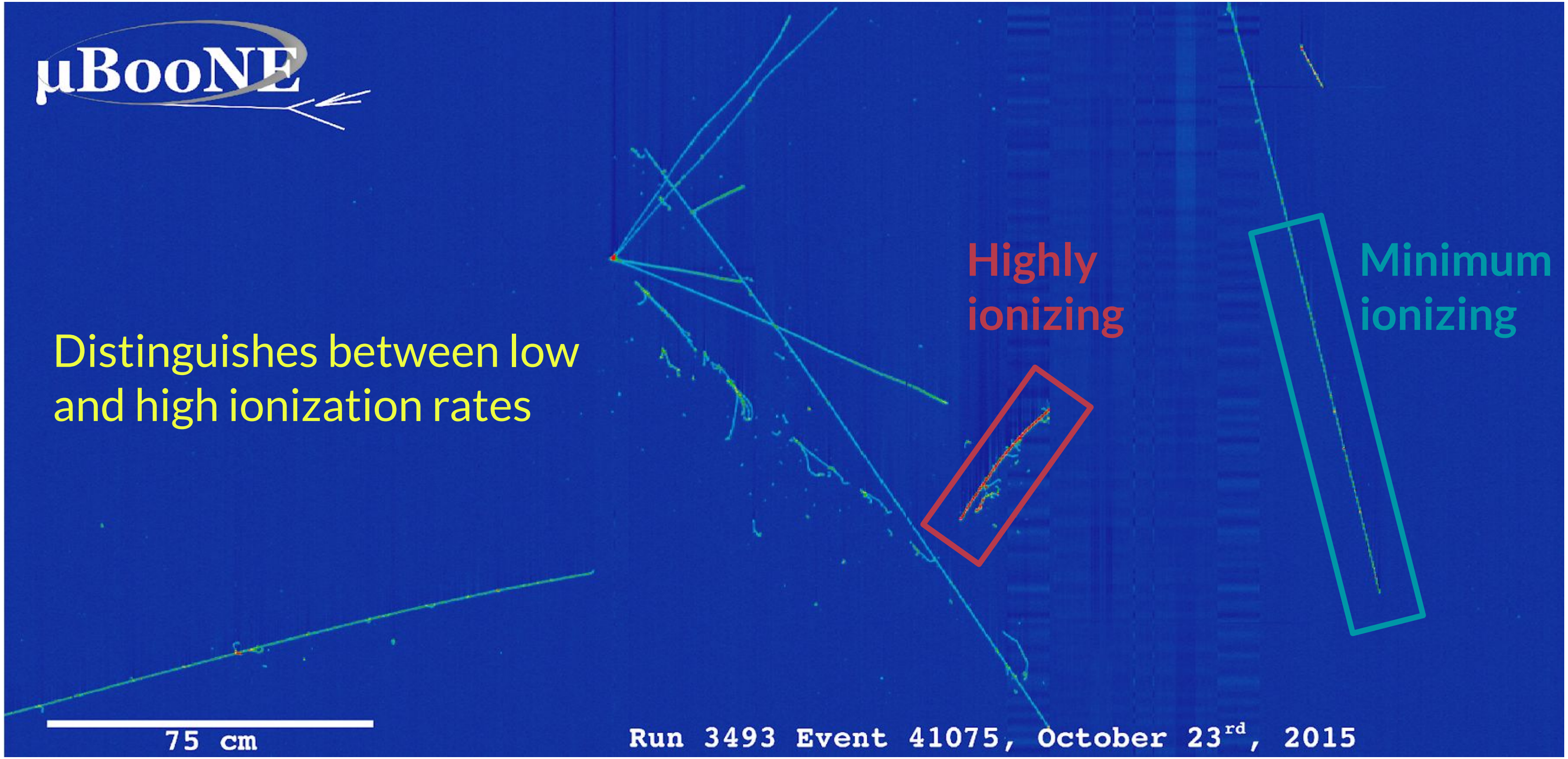
LArTPC Image



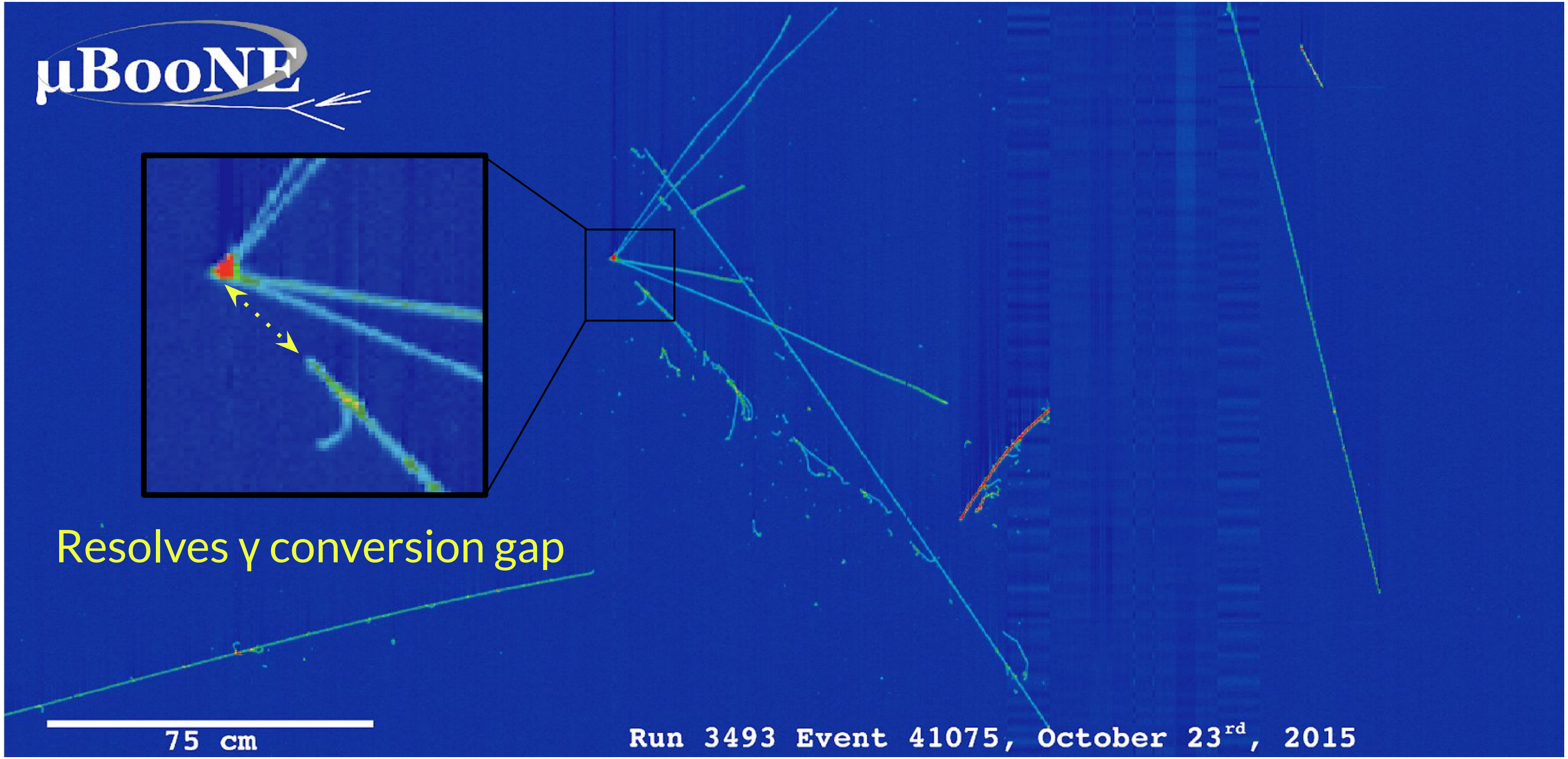
LArTPC Image



LArTPC Image



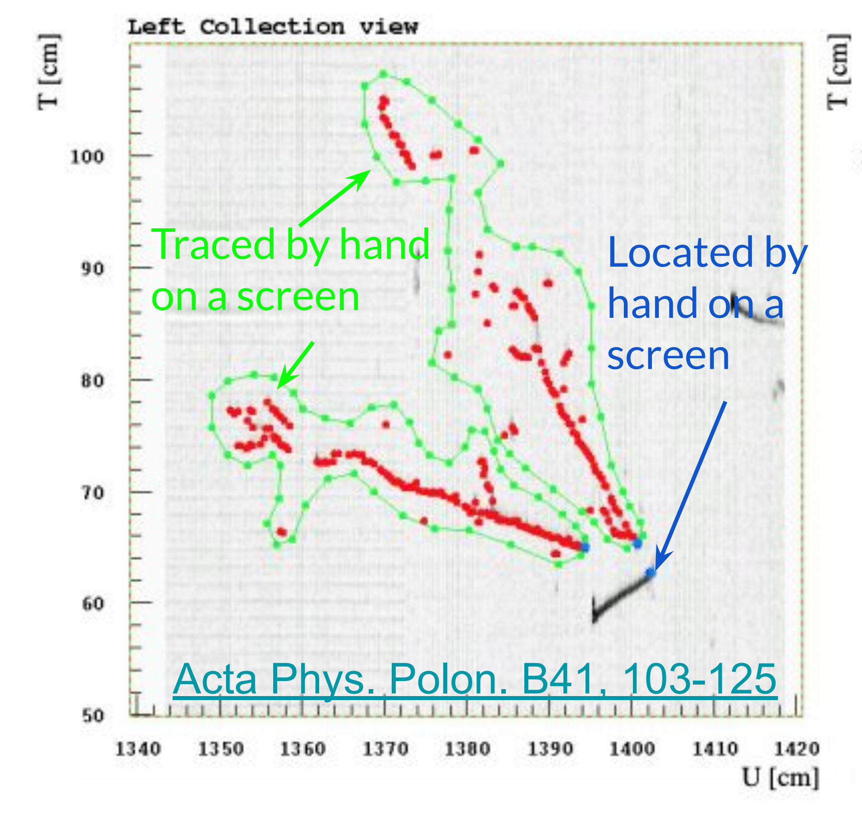
LArTPC Image



LArTPC Reconstruction

Moved away from hand-scanning in the XXI[st] century?

LArTPC Reconstruction

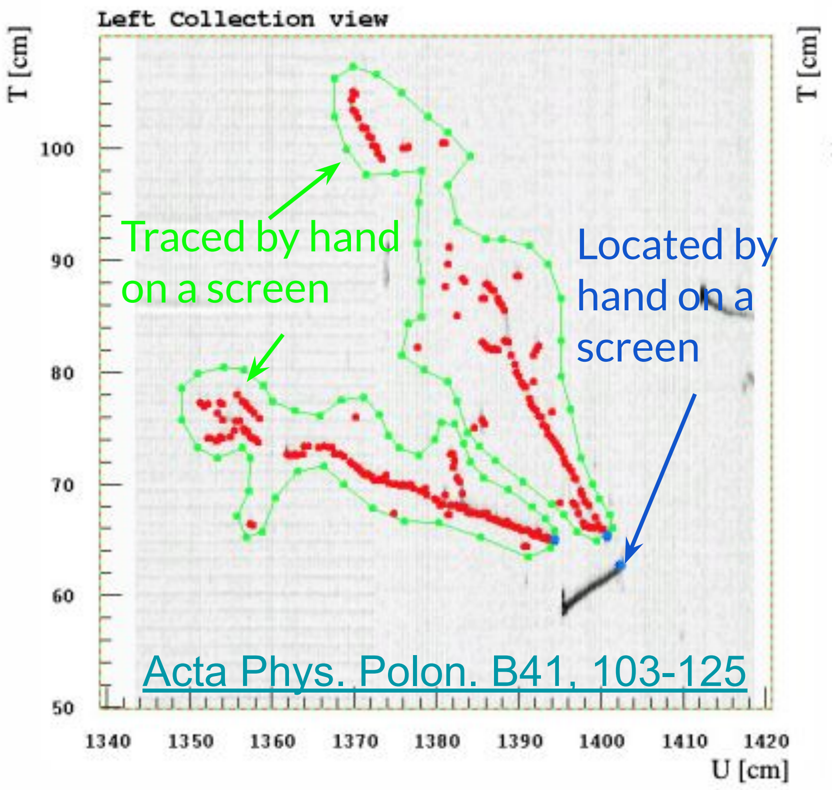


Not immediately…

- Papers still published using this technique until **2013**!
- Still **time intensive to reconstruct**

Viable with low rates…

LArTPC Reconstruction



Not immediately…

- Papers still published using this technique until **2013**!
- Still **time intensive to reconstruct**

Viable with low rates…

…dead on arrival at SBN and DUNE

- **ICARUS:** O(10^4) ν candidate / day!
- **DUNE-ND:** O(10^6) ν / day!

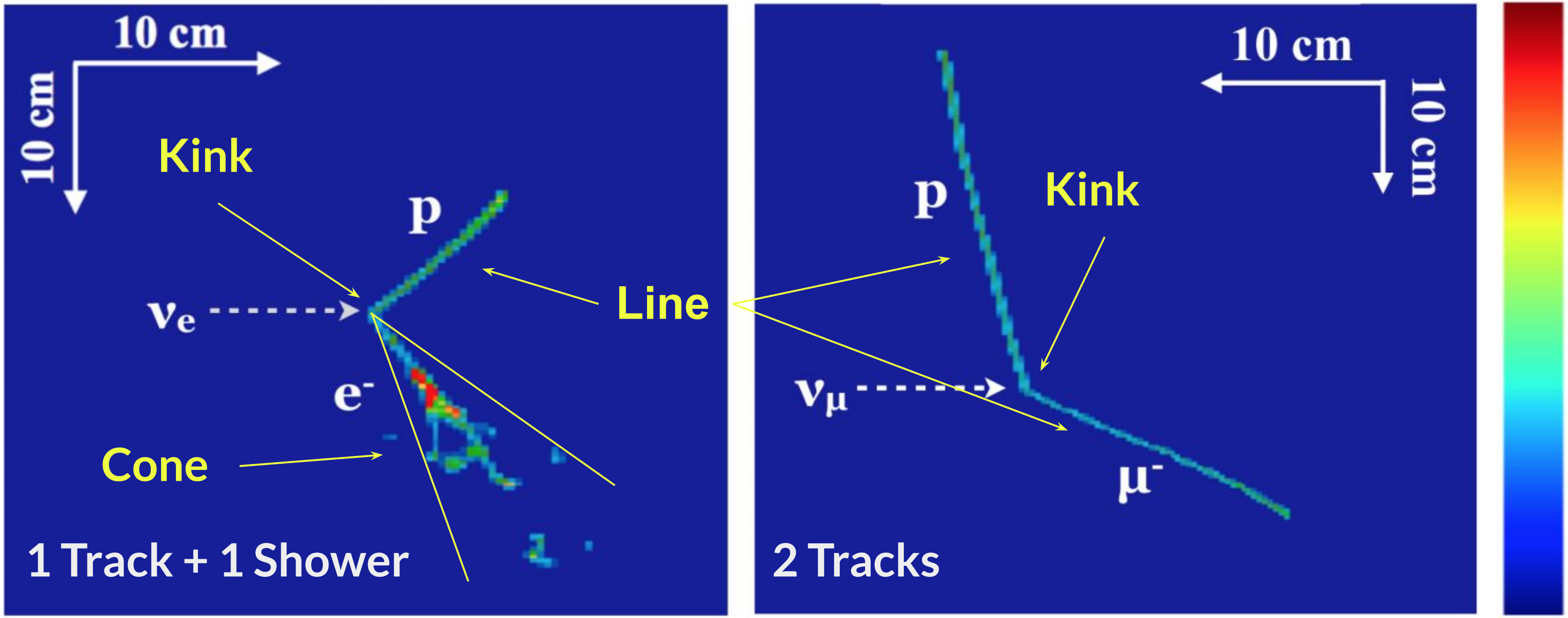
Grad. student nightmare…

LArTPC Reconstruction Challenge

SLAC NATIONAL ACCELERATOR LABORATORY

Why is it so **challenging to automate**?

- Write an algorithm based on **physics principles...**



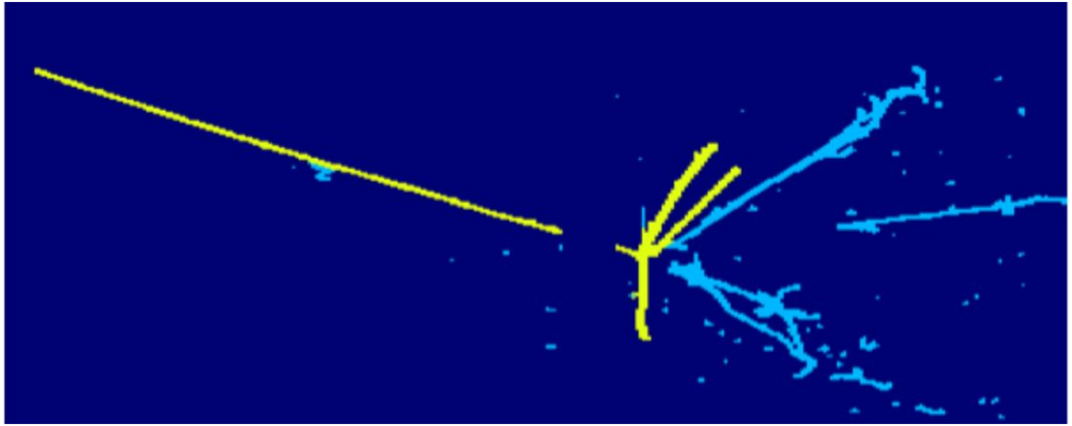
LArTPC Reconstruction Challenge

SLAC NATIONAL ACCELERATOR LABORATORY

Why is it so **challenging to automate**?

- Write an algorithm based on **physics principles…**
- Realize it **fails on harder topologies**
- **Add new rules** to handle new topology, **repeat**

Years of development



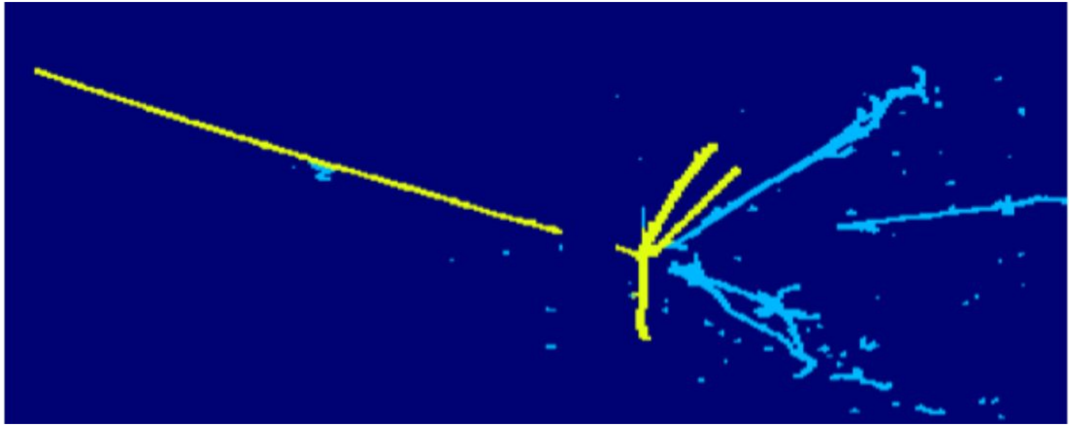
LArTPC Reconstruction Challenge

SLAC NATIONAL ACCELERATOR LABORATORY

Why is it so **challenging to automate**?

- Write an algorithm based on **physics principles…**
- Realize it **fails on harder topologies**
- **Add new rules** to handle new topology, **repeat**

Years of development



With great detail comes great responsibility

- **Variety** of possible **neutrino interactions topologies is huge**

The Weight of Expectations

PandoraPFA: particle flow algorithm developed for future e^+e^- colliders

- Adapted in the 2010s to be used in LArTPCs, **> 10 years of development**
- **Best performing traditional approach** in several LArTPC experiments

Does it live up to the DUNE/SBN requirements?

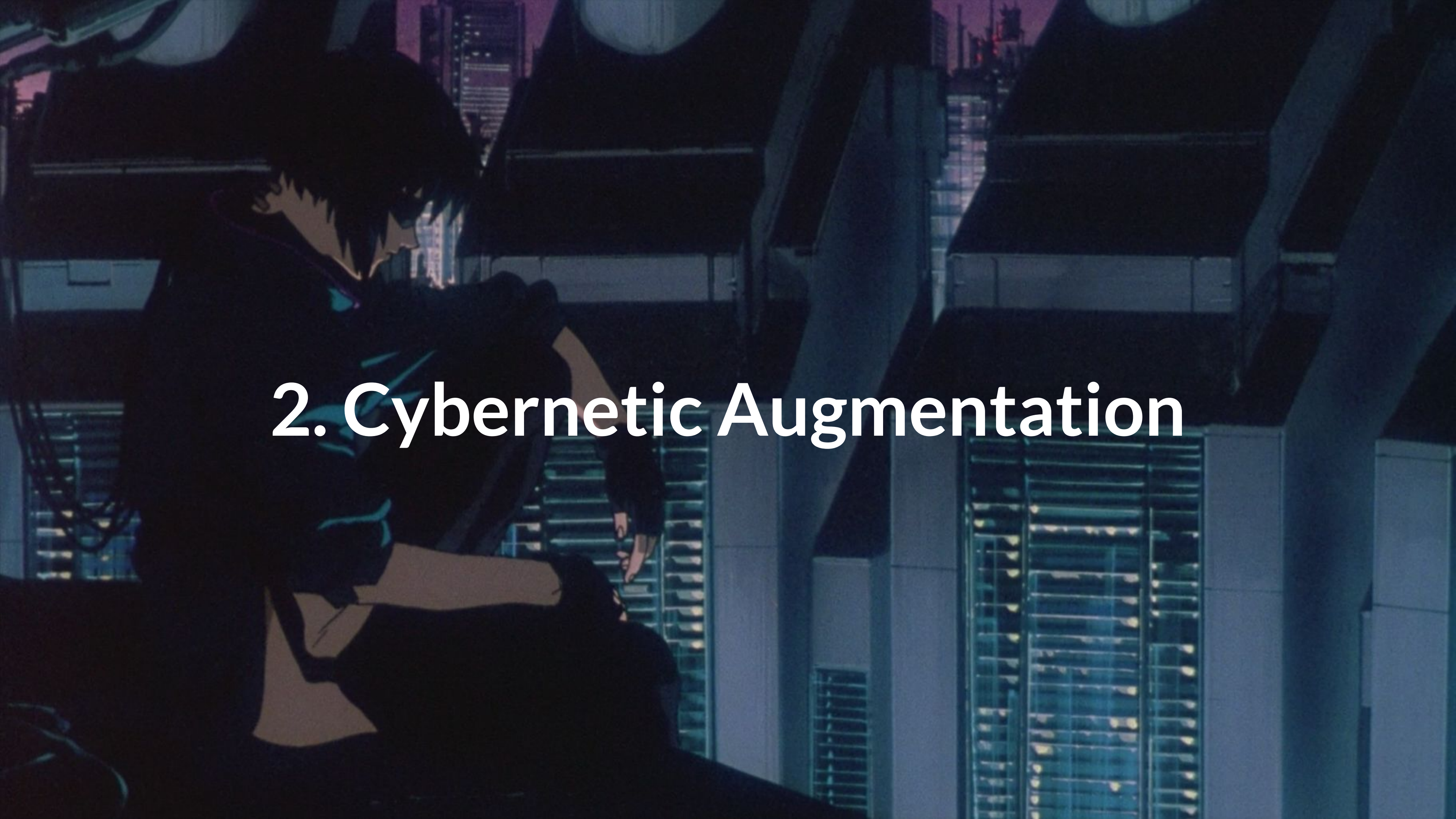
The Weight of Expectations

PandoraPFA: particle flow algorithm developed for future e^+e^- colliders

- Adapted in the 2010s to be used in LArTPCs, **> 10 years of development**
- **Best performing traditional approach** in several LArTPC experiments

Does it live up to the DUNE/SBN requirements? **Not quite…**

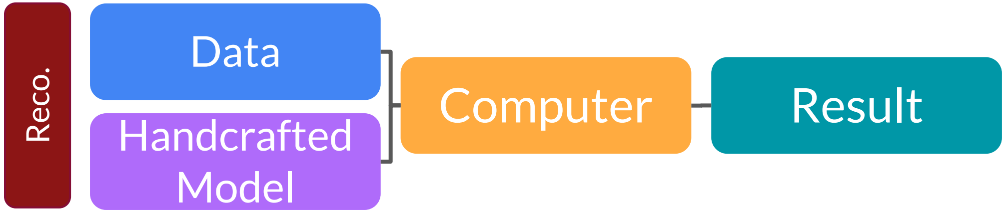
	SBN Proposal 1eX (hand-scanning)	MicroBooNE 1eNp0π paper (PandoraPFA)
Purity	**85 %**	**80 %**
Efficiency	**80 %**	**15 %**
	arXiv:1503.01520	arXiv:2110.14065



2. Cybernetic Augmentation

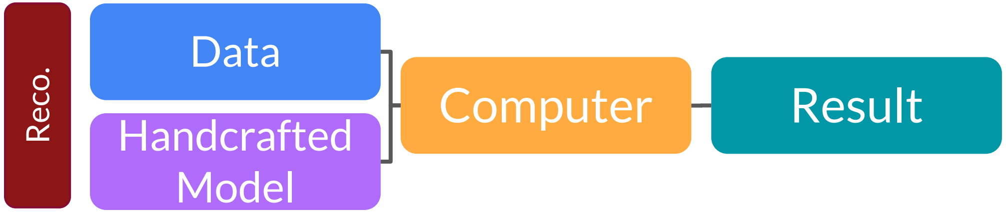
Machine Learning

Traditional Approach

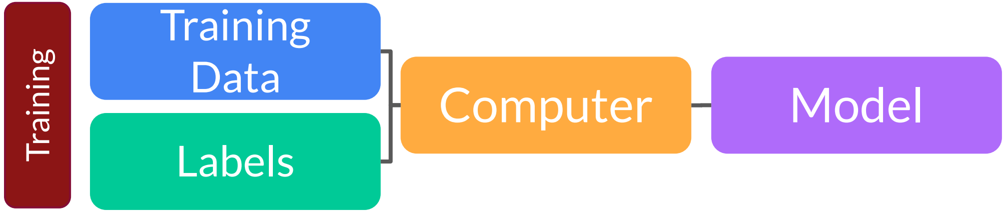


Machine Learning

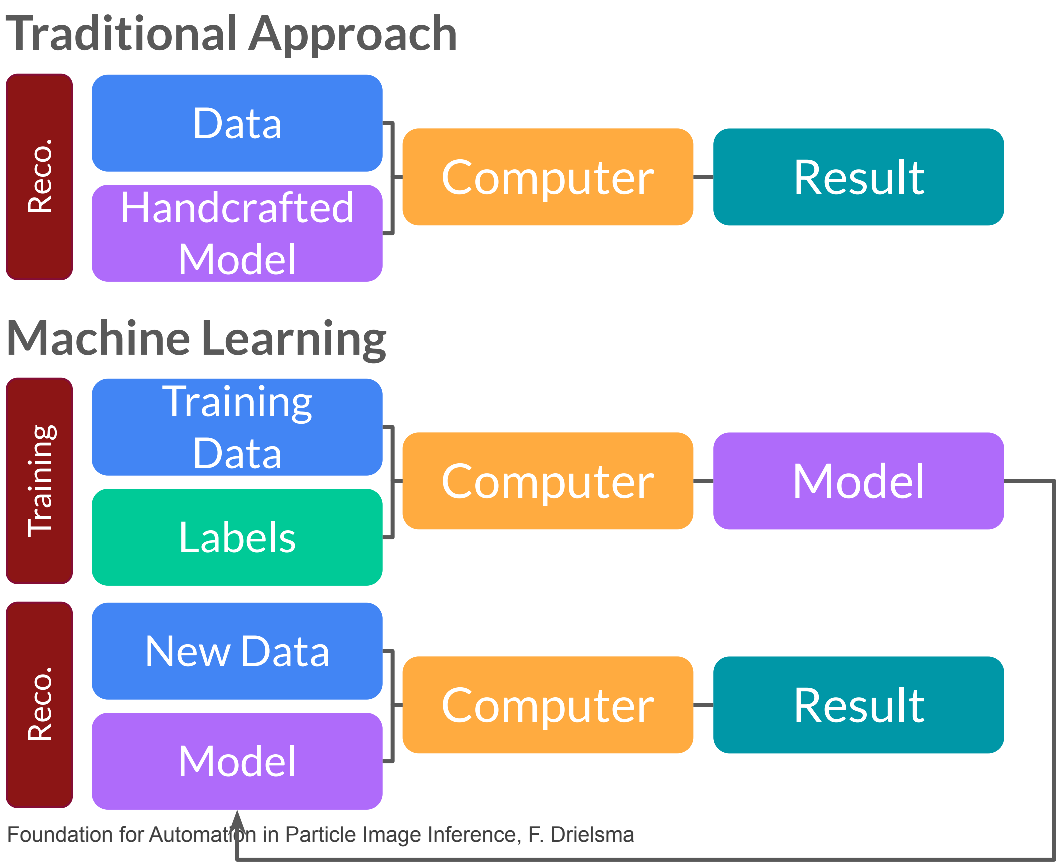
Traditional Approach



Machine Learning



Machine Learning



Machine Learning

Traditional Approach

Reco.

Data

Handcrafted Model

Computer

Result

Machine Learning

Training

Training Data

Labels

Computer

Model

Reco.

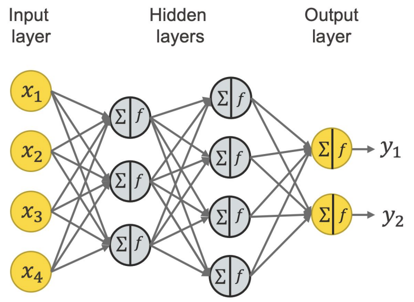
New Data

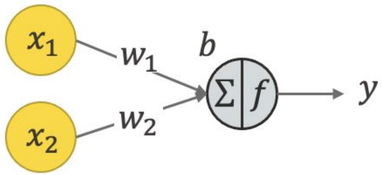
Model

Computer

Result

~ Universal function approximator



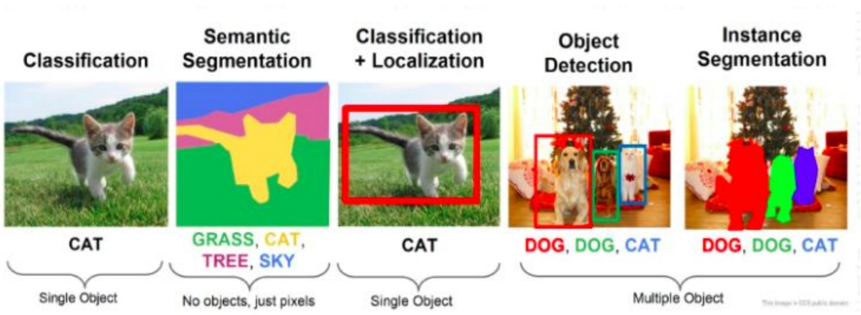


$$y = f(w_1x_1 + w_2x_2 + b)$$
$$f = \text{Tanh, ReLU, ...}$$

Machine Learning and Computer Vision

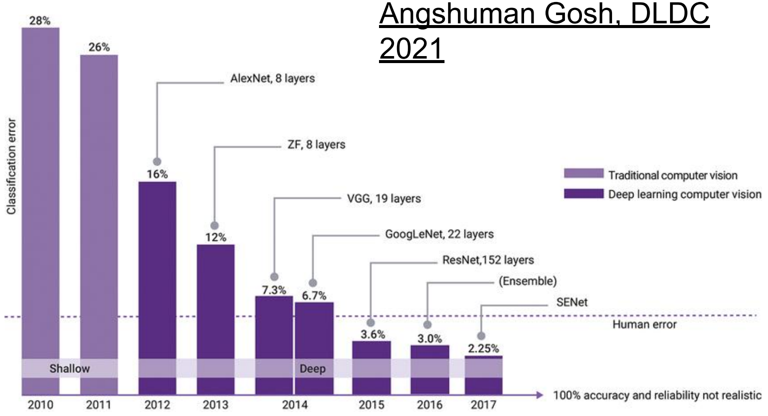
ML is the state-of-the-art in CV, i.e. extracting high-level information from **images**

- ML revolutionized accuracy on image processing tasks
- **AI/ML for science: leverage those techniques in LArTPCs** (image data)



Stanford, CS231

Angshuman Gosh, DLDC 2021

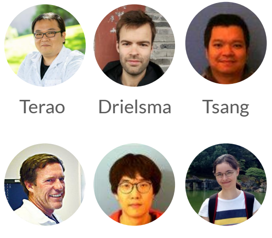


LArTPC ML “Network”

SLAC NATIONAL ACCELERATOR LABORATORY

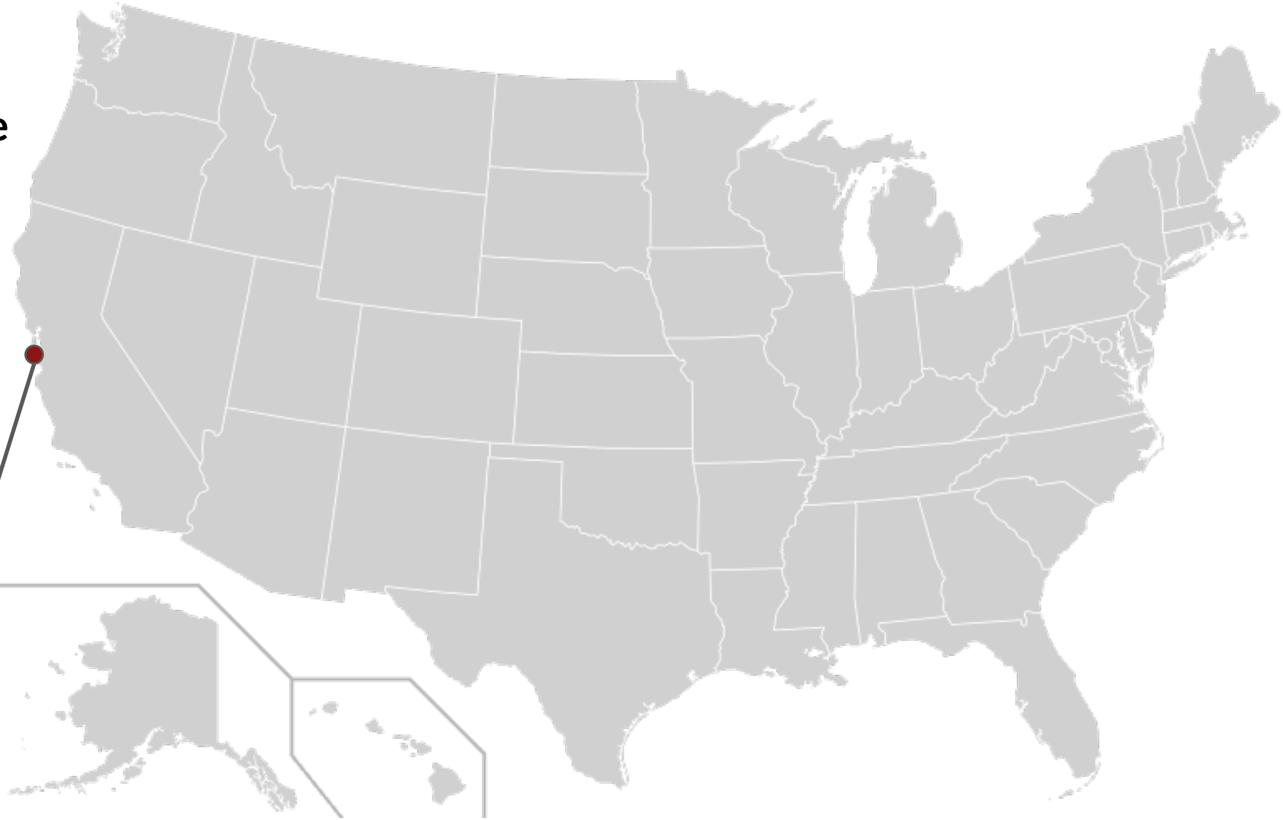
Joined SLAC in 2019:
FD responsible to deliver ML reco. chain to LArTPCs

SLAC NATIONAL ACCELERATOR LABORATORY



Terao Drielsma Tsang

Usher Koh Dominé



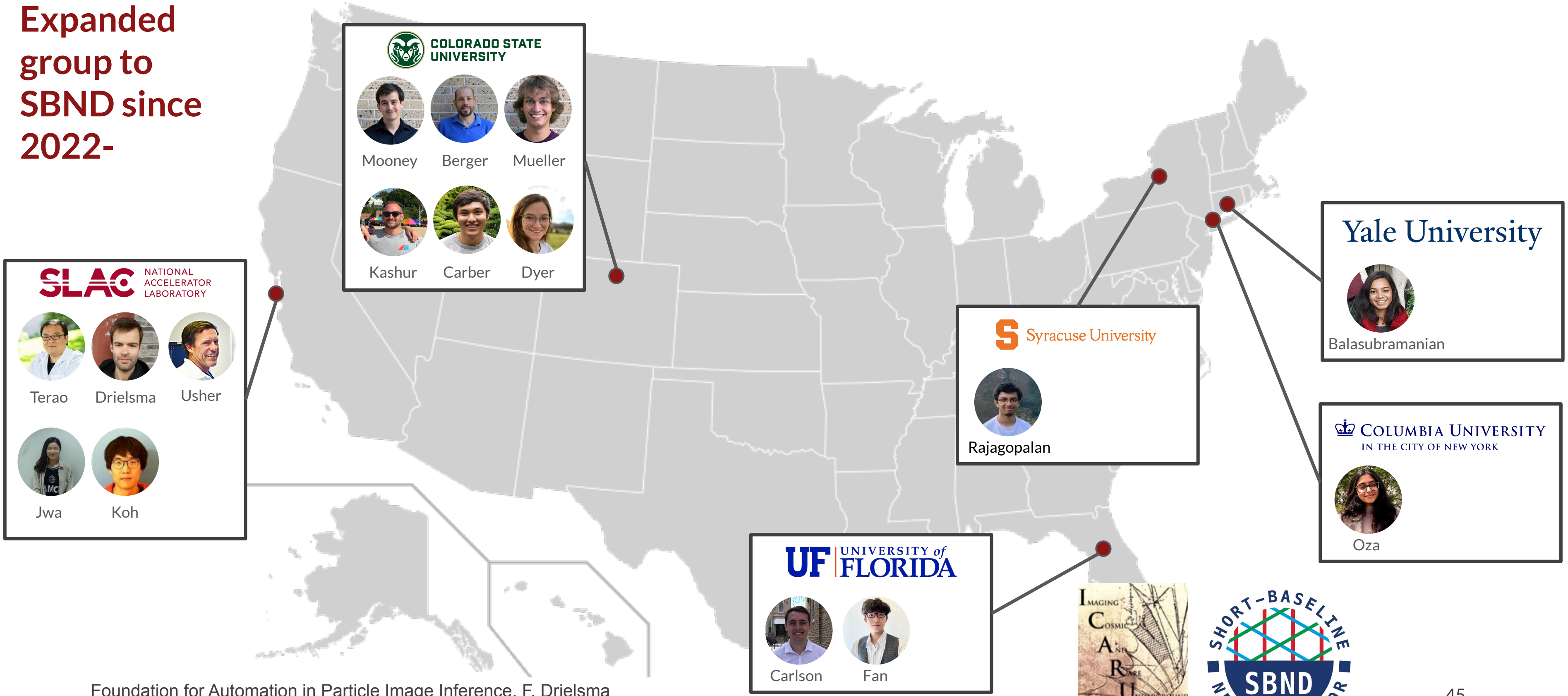
LArTPC ML "Network"

Formed and convene the ICARUS ML group since 2020-

LArTPC ML "Network"

SLAC NATIONAL ACCELERATOR LABORATORY

Expanded group to SBND since 2022-

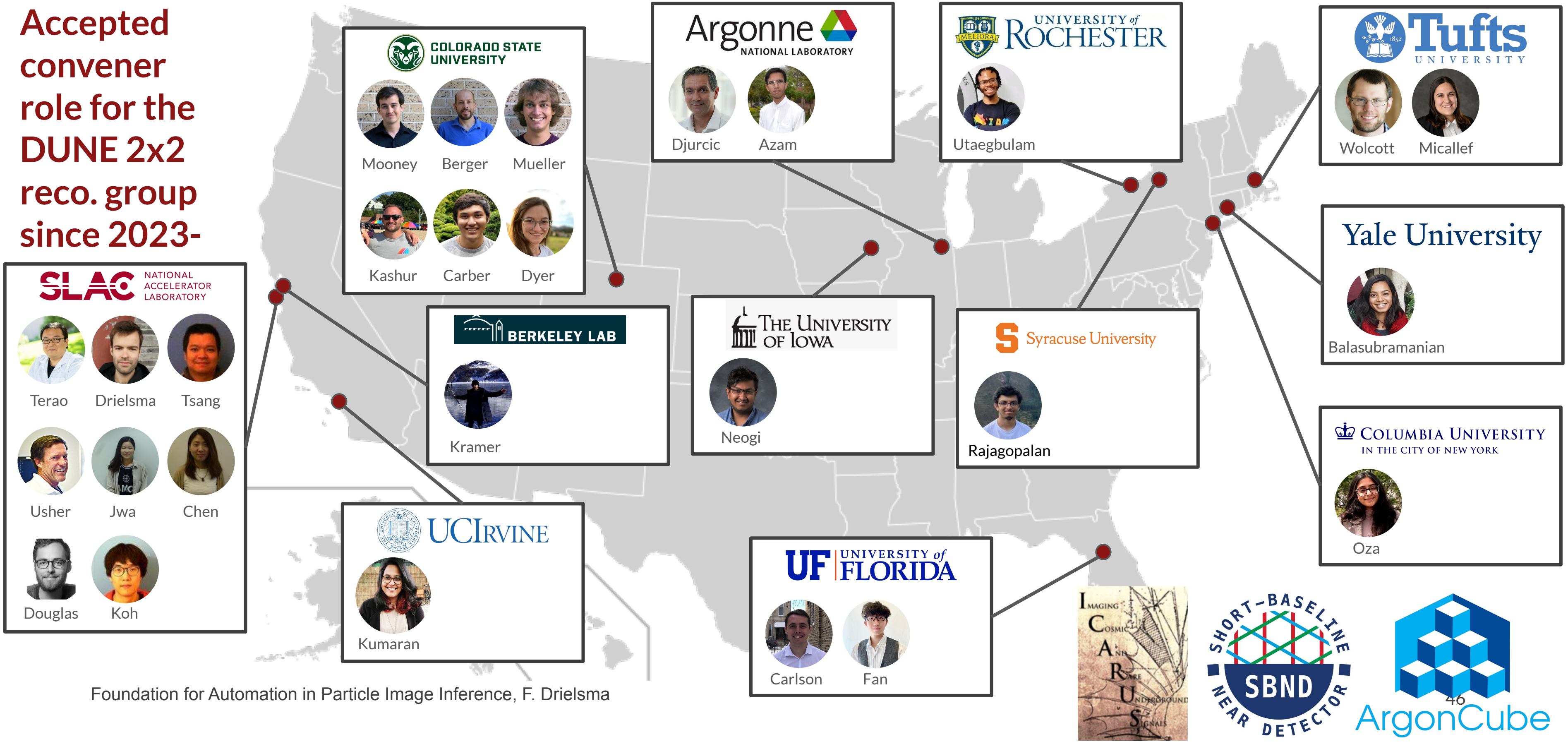


Imaging Cosmic And Rare Underground Signals

SHORT-BASELINE NEAR DETECTOR SBND

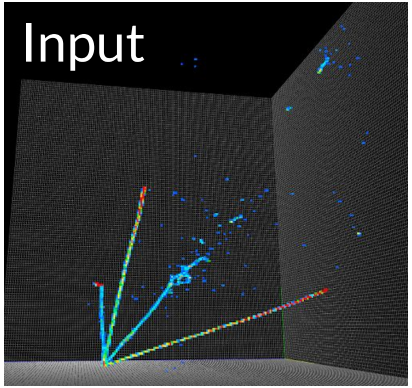
LArTPC ML "Network"

Accepted convener role for the DUNE 2x2 reco. group since 2023-



Physics-Informed ML Reconstruction

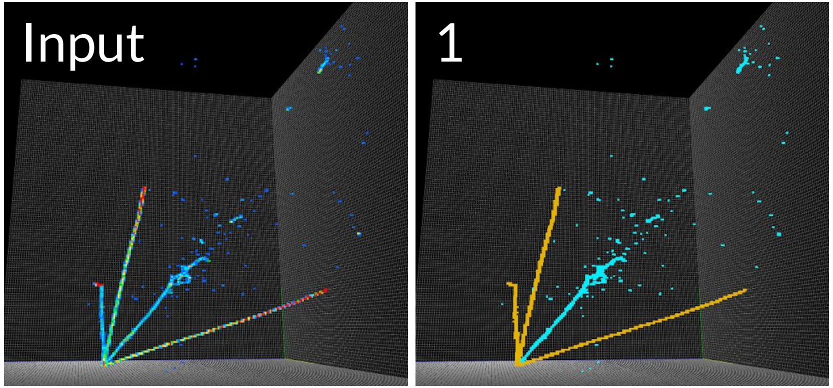
What is relevant to pattern recognition in a detailed interaction image?



Physics-Informed ML Reconstruction

What is relevant to pattern recognition in a detailed interaction image?

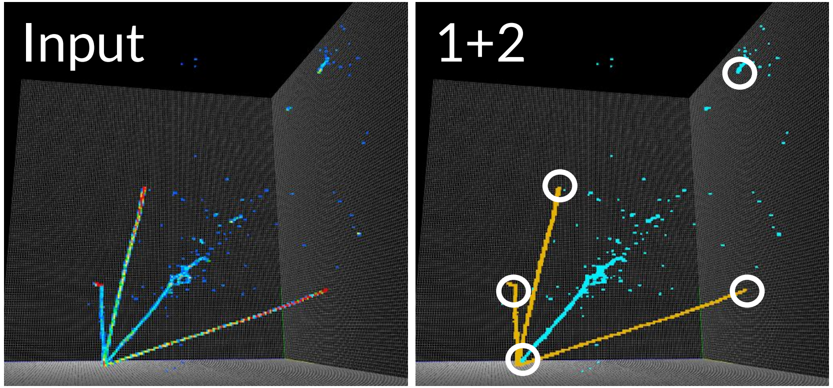
1. Separate topologically distinguishable **types of activity**



Physics-Informed ML Reconstruction

What is relevant to pattern recognition in a detailed interaction image?

1. Separate topologically distinguishable **types of activity**
2. Identify **important points** (vertex, start points, end points)

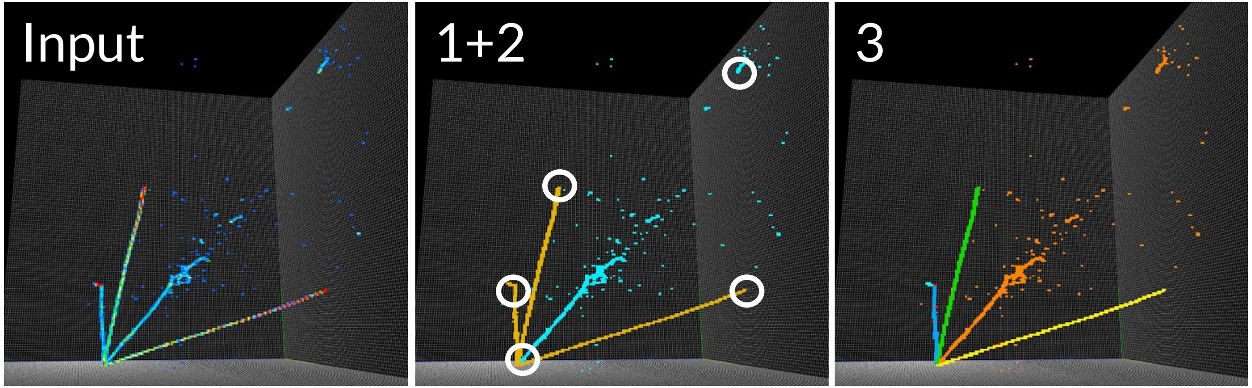


Foundation for Automation in Particle Image Inference, F. Drielsma

Physics-Informed ML Reconstruction

What is relevant to pattern recognition in a detailed interaction image?

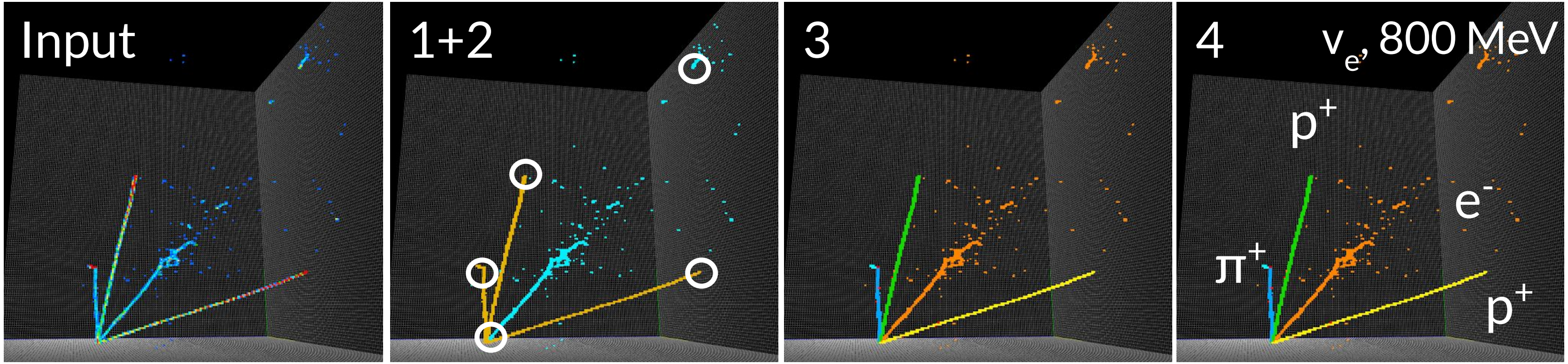
1. Separate topologically distinguishable **types of activity**
2. Identify **important points** (vertex, start points, end points)
3. Cluster individual **particles** (tracks and full showers)



Physics-Informed ML Reconstruction

What is relevant to pattern recognition in a detailed interaction image?

1. Separate topologically distinguishable **types of activity**
2. Identify **important points** (vertex, start points, end points)
3. Cluster individual **particles** (tracks and full showers)
4. Cluster **interactions**, identify **particle properties** in context



Physics-Informed ML Reconstruction

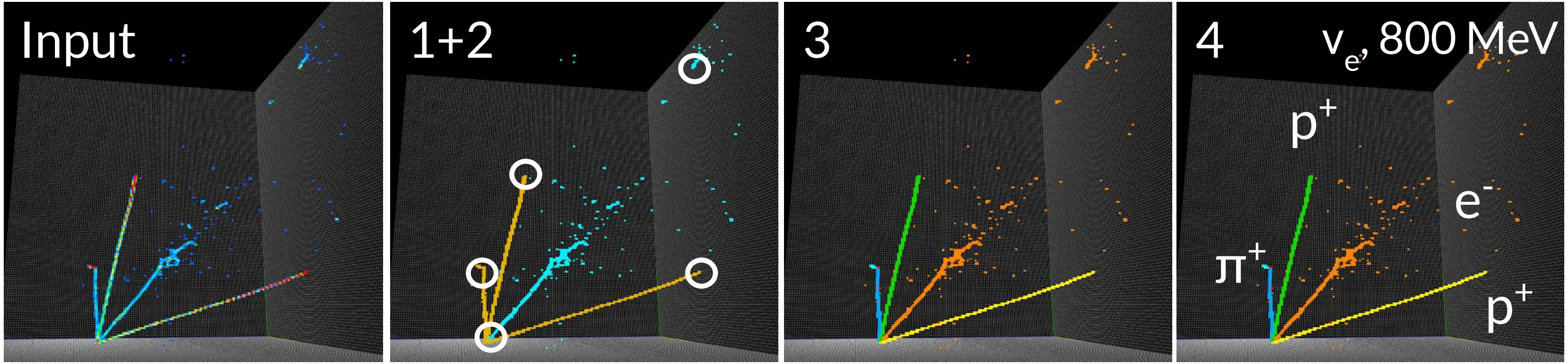
What is relevant to pattern recognition in a detailed interaction image?

1. Separate topologically distinguishable **types of activity**
2. Identify **important points** (vertex, start points, end points)

→ **Pixel-level**

3. Cluster individual **particles** (tracks and full showers)
4. Cluster **interactions**, identify **particle properties** in context

→ **Cluster-level**

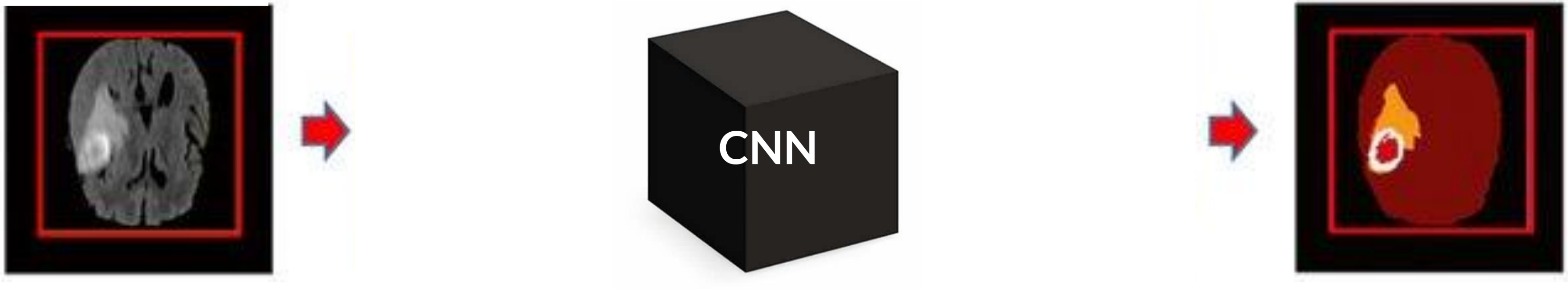


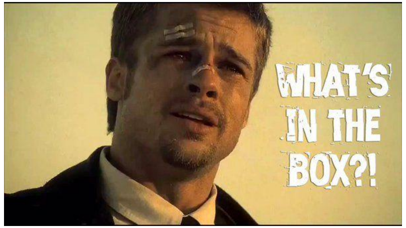
Pixel-level Feature Extraction

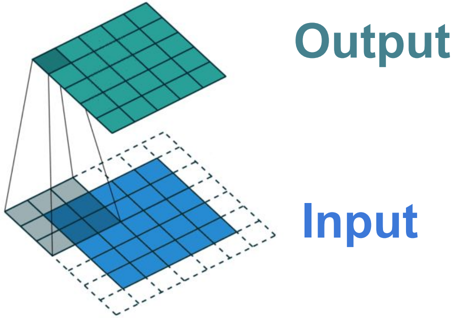
Convolutional Neural Networks (CNN, source)

Pixel-level Feature Extraction

Convolutional Neural Networks (source)

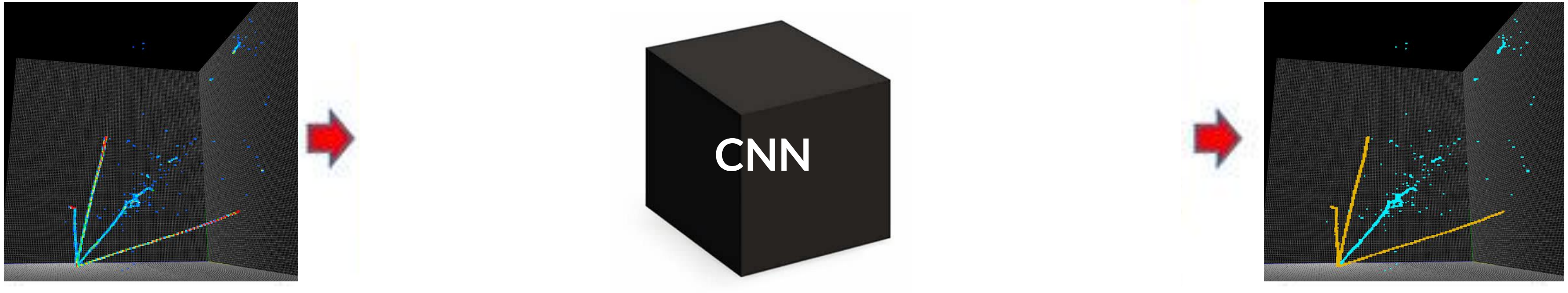






Pixel-level Feature Extraction

Does it work on LArTPC data?



Specificity:

- **3D**: ICARUS = O(10) Gigapixels
- **Occupancy**: ~ 10^{-4}, locally dense
    - Mostly meaningless space

Pixel-level Feature Extraction

Does it work on LArTPC data?

Specificity:

- **3D**: ICARUS = O(10) Gigapixels
- **Occupancy**: ~ 10^{-4}, locally dense
    - Mostly meaningless space

Problems:

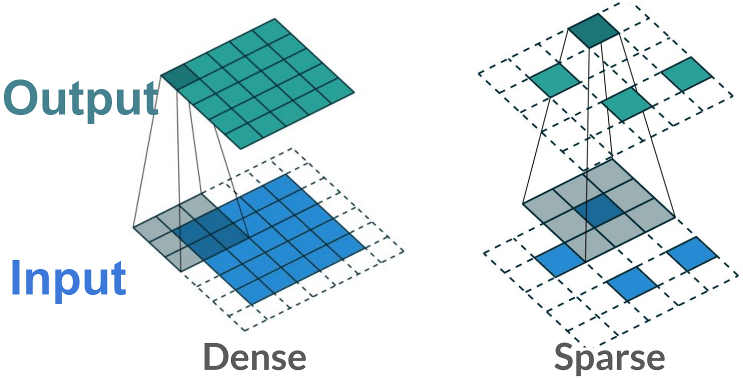
- **Memory:** ~ 100 GB per image…
- **Wasted computation**: 99.99% empty

→ **Not viable**

Pixel-level Feature Extraction

Solution? **Sparse Convolutions!**

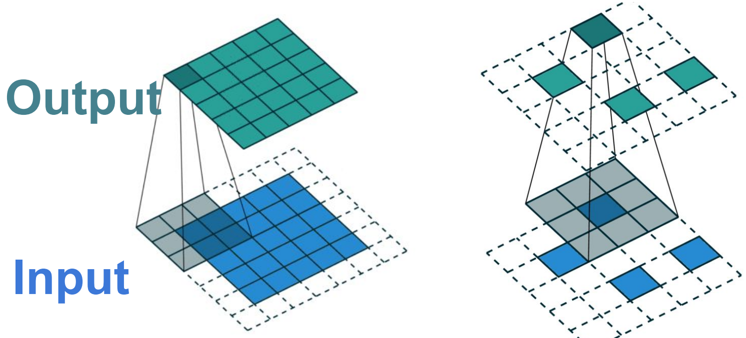
- Only operate on **active pixels**
- Technique (SCN) invented at ~~Facebook~~ Meta in 2017
- Pioneered use in Physics at SLAC: Quanta Magazine, PRD paper



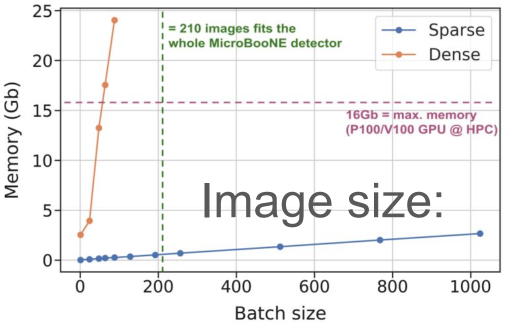
Pixel-level Feature Extraction

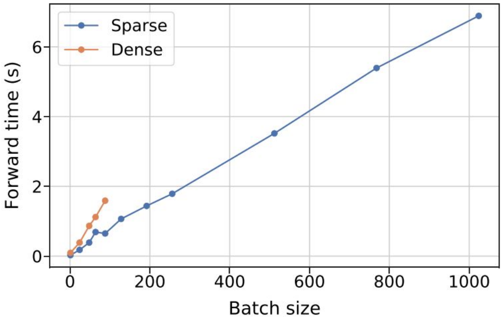
Solution? **Sparse Convolutions!**

- Only operate on **active pixels**
- Technique (SCN) invented at ~~Facebook~~ Meta in 2017
- Pioneered use in Physics at SLAC: Quanta Magazine, PRD paper
- **Scales with space point count only!** ~~O(10) GPix~~ $\rightarrow$ O(1) MPix



Dense Sparse

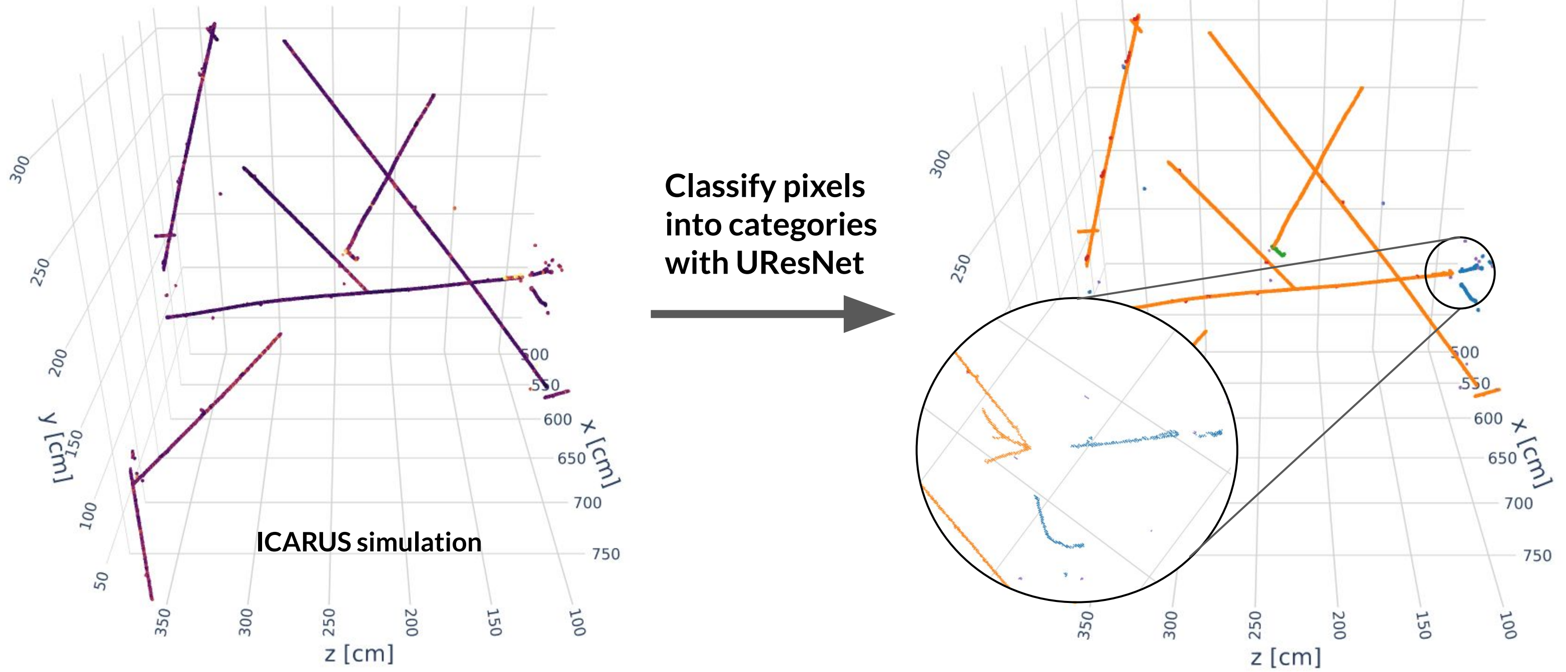




Semantic Segmentation

Separate **topologically different** types of activity

- **Tracks**, **Showers**, **delta rays**, **Michel electrons**, **low energy blips**

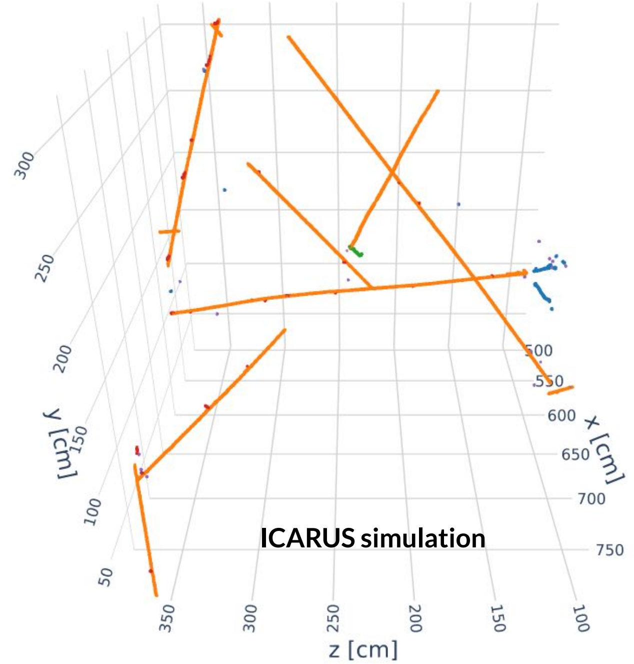


Paper: PhysRevD.102.012005

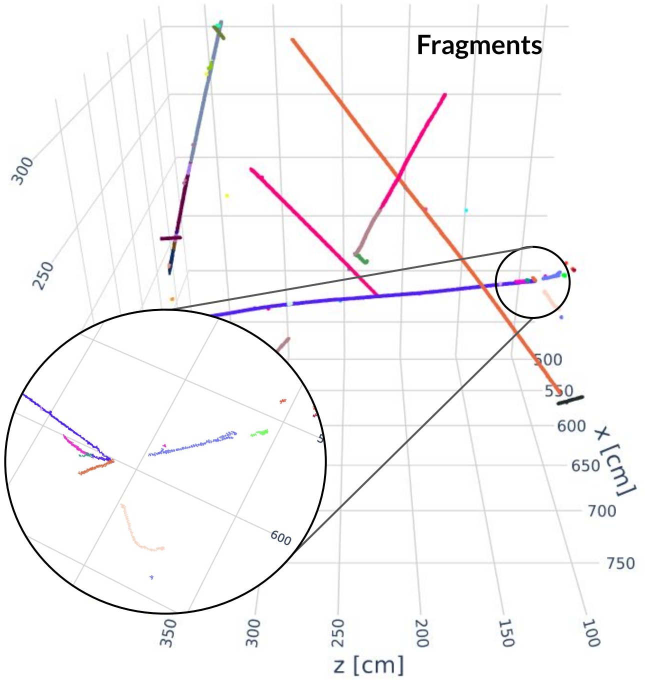
Dense Fragment Formation

Break track/shower fragment instances **where they touch**

- Cluster track/shower **fragments** at this stage



Classify pixels into dense clusters

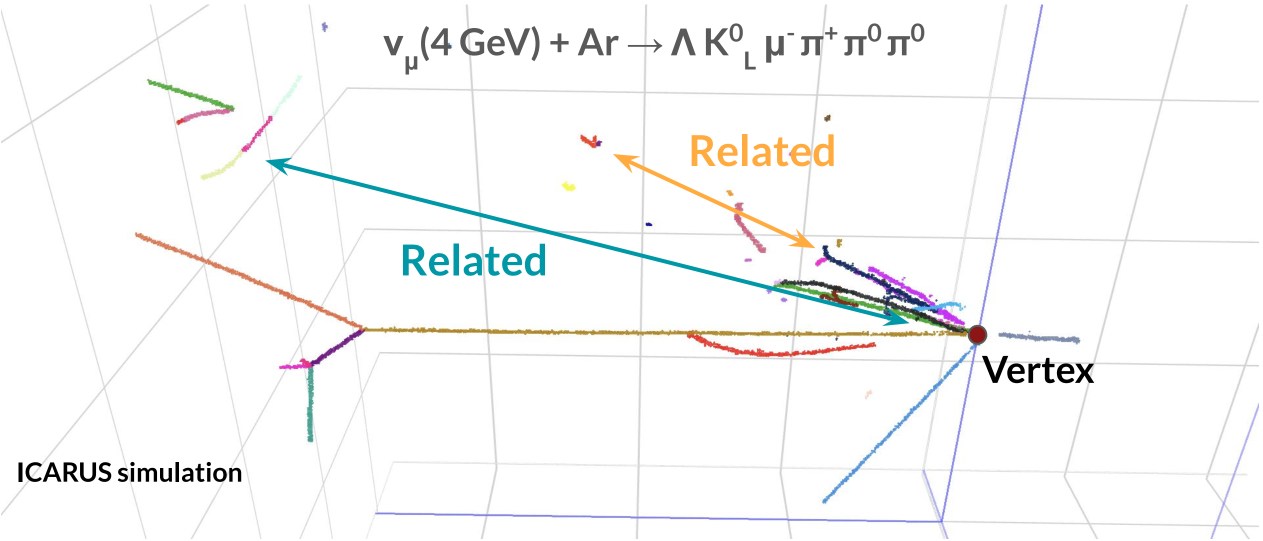


Paper: arXiv:2007.03083

Cluster-level feature extraction

CNN: mostly sensitive to **local neighborhood** of pixel, but…

- **EM showers**: photon mean free path in LAr = 18 cm (**60 pixels**)
- **Interactions**: π^0, K^0, Λ, neutrons



First issue I tackled at SLAC

Cluster-level feature extraction

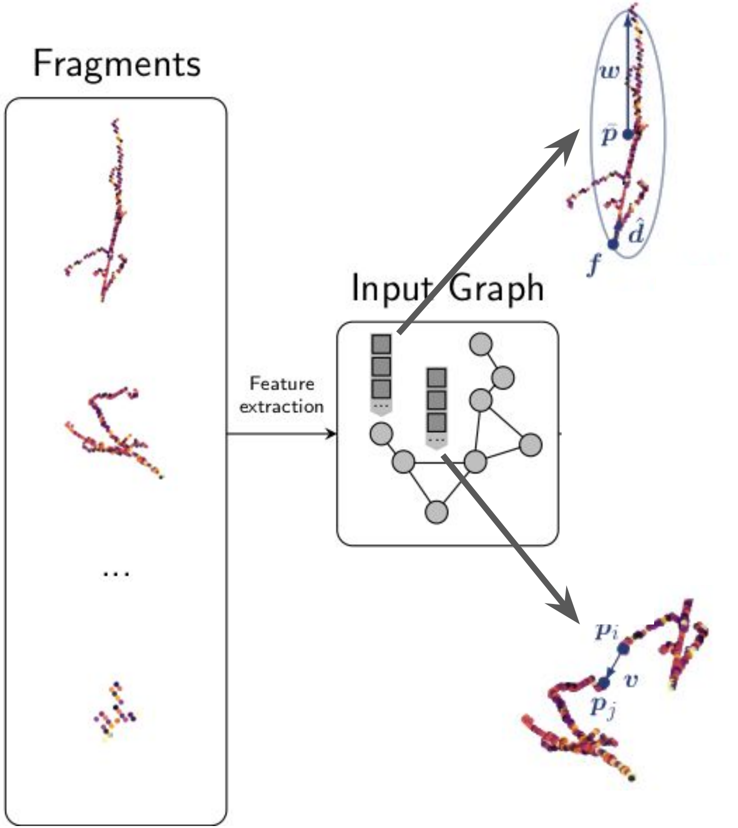
We now represent the set of fragments as a **set of nodes in a graph** where **edges represent correlations**

Node features:

- Centroid
- Covariance matrix
- Start point/direction
- . . .

Edge features:

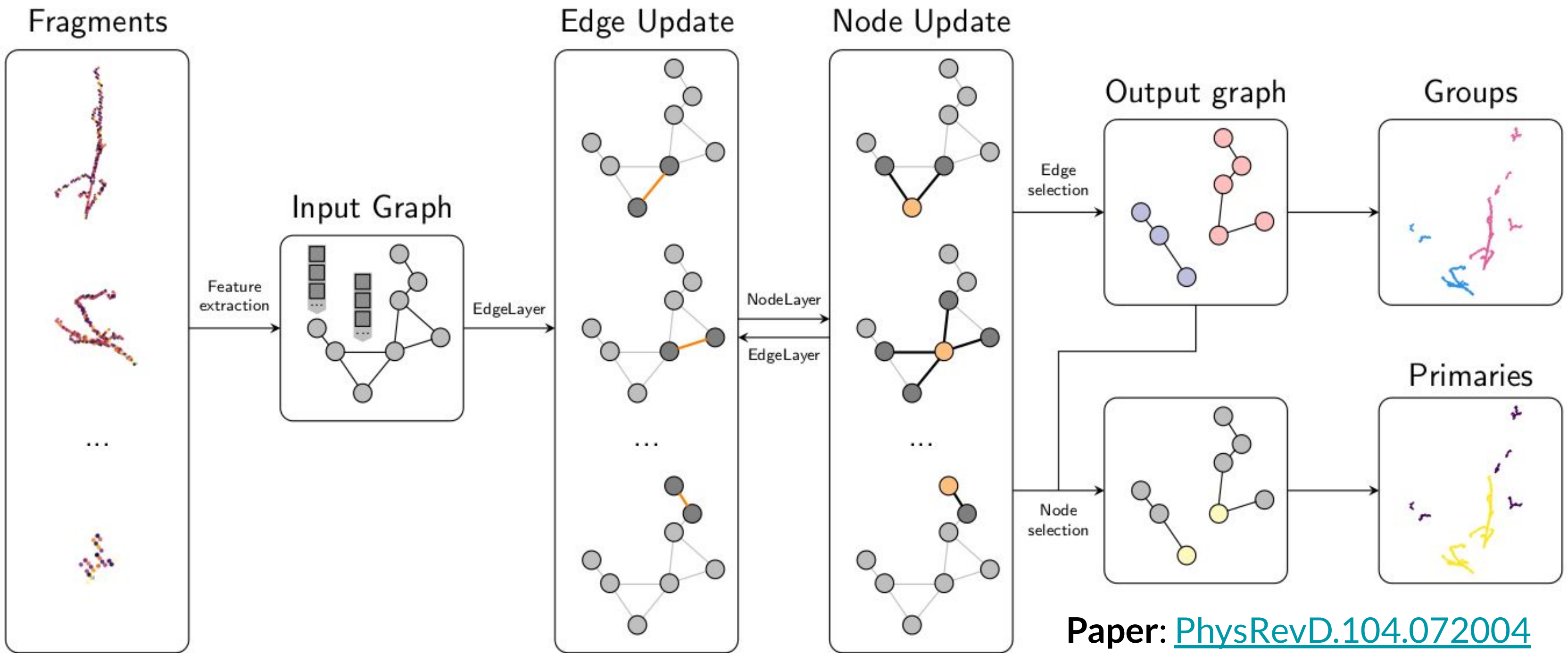
- Displacement vector
- . . .



Cluster-level feature extraction

SLAC NATIONAL ACCELERATOR LABORATORY

Graph Neural Network (2017): develop features useful to node/edge classification

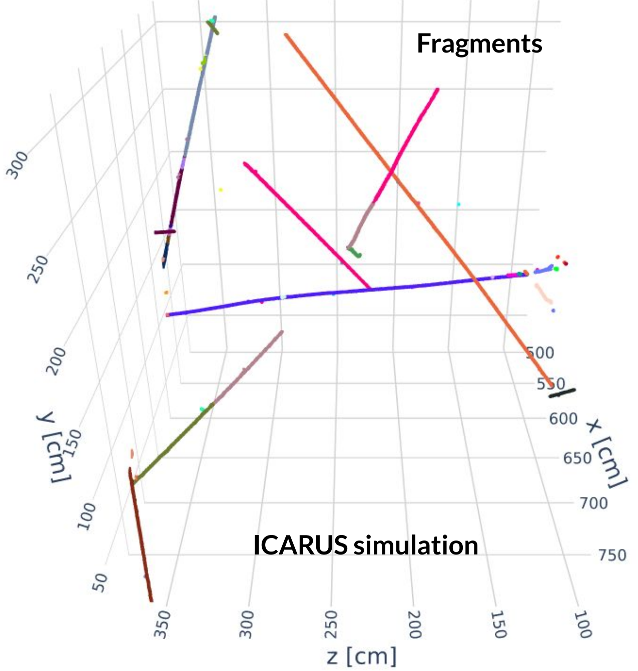


Paper: PhysRevD.104.072004
(F. Drielsma et al.)

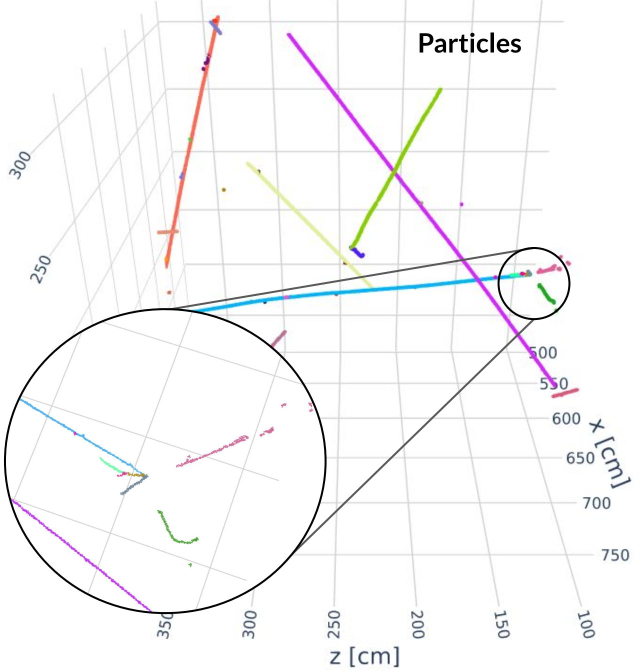
Particle Aggregation

Aggregate track/shower fragment instances into **particles**

- Find edges that connect **fragments that belong together**



Aggregate particle fragments

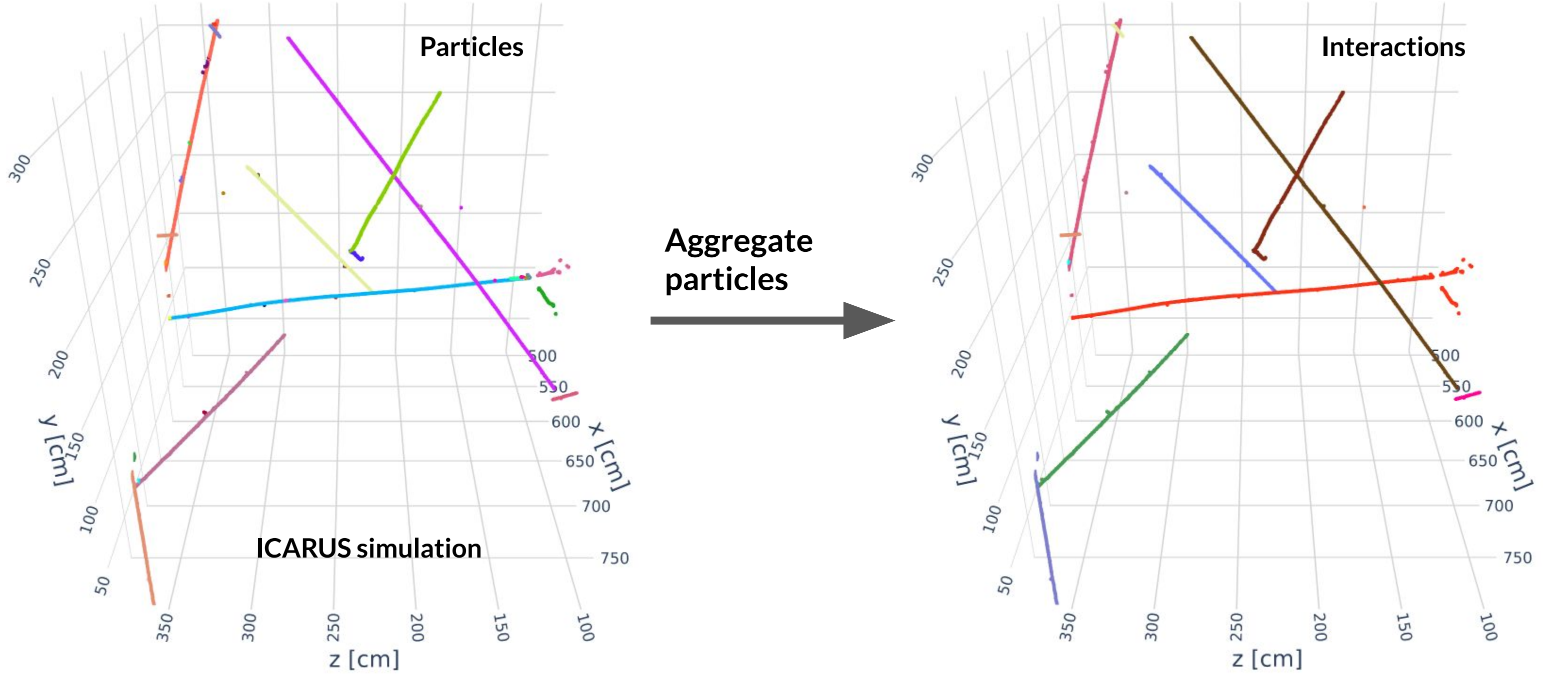


Paper: PhysRevD.104.072004

Particle Aggregation

Aggregate track/shower instances into **interactions**

- Find edges that connect **particles that belong together**



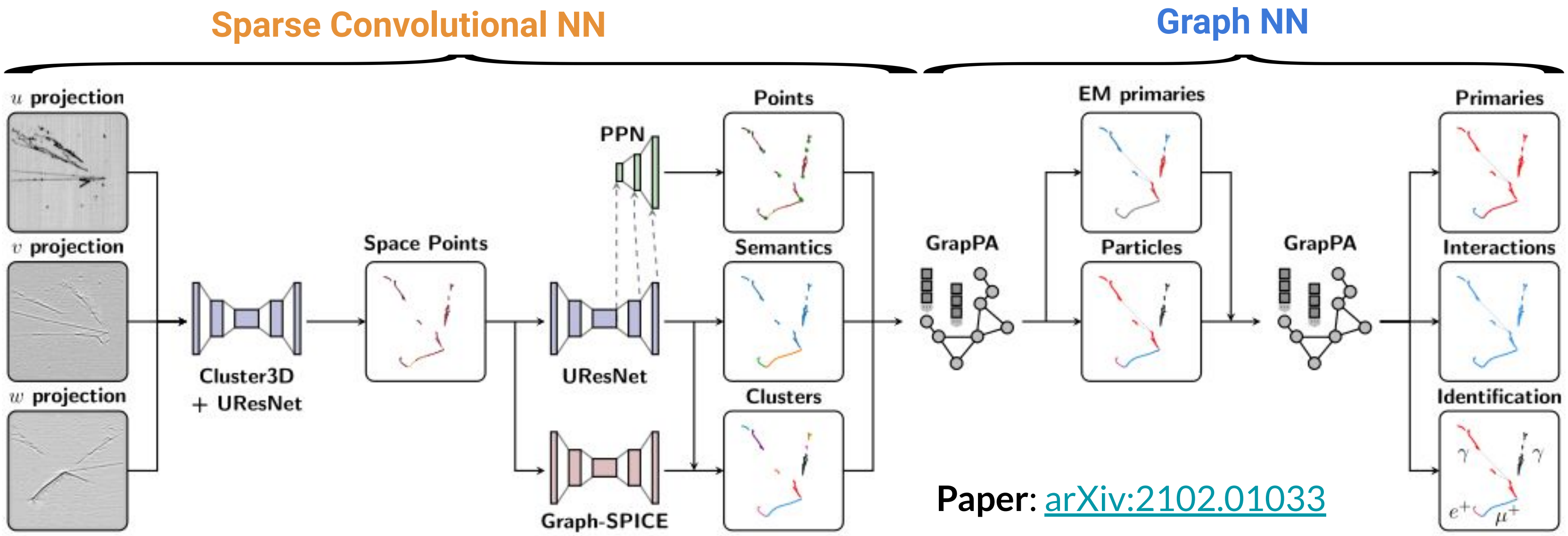
Paper: PhysRevD.104.072004

Reconstruction in LArTPCs

SLAC NATIONAL ACCELERATOR LABORATORY

End-to-end ML-based reconstruction chain

- **Sparse CNN** for pixel-level features, **GrapPA** for superstructure formation

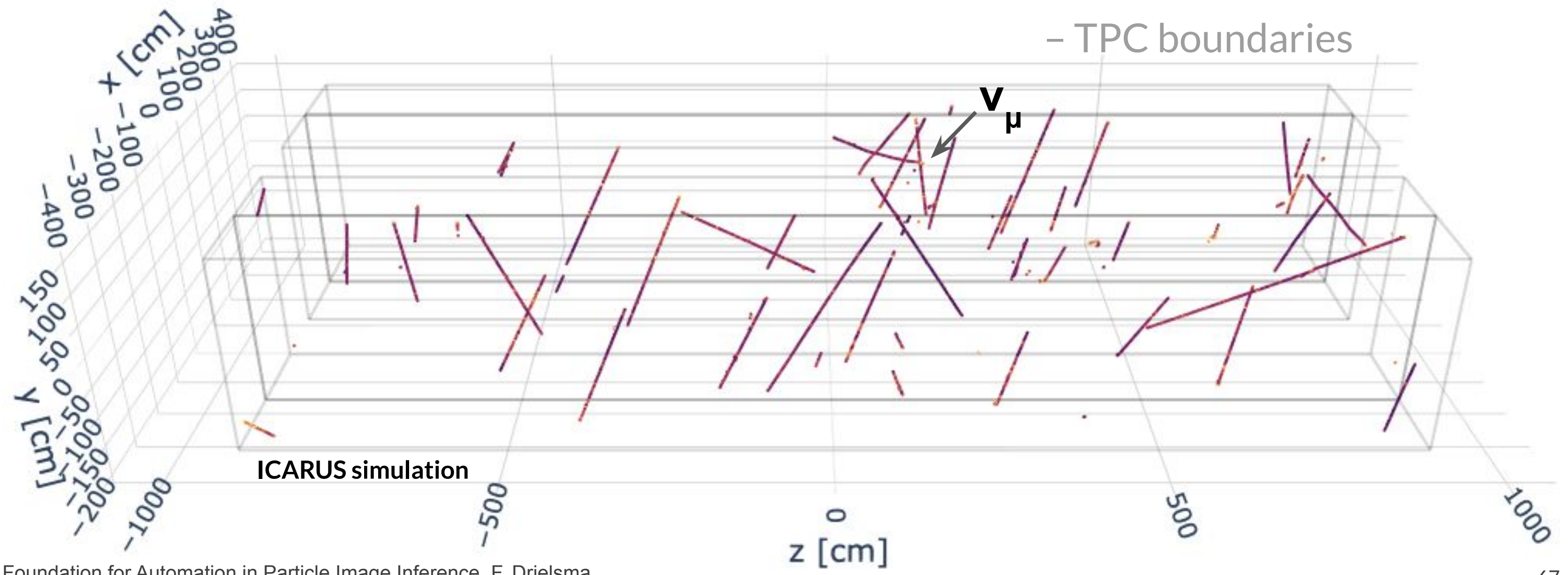


Paper: arXiv:2102.01033

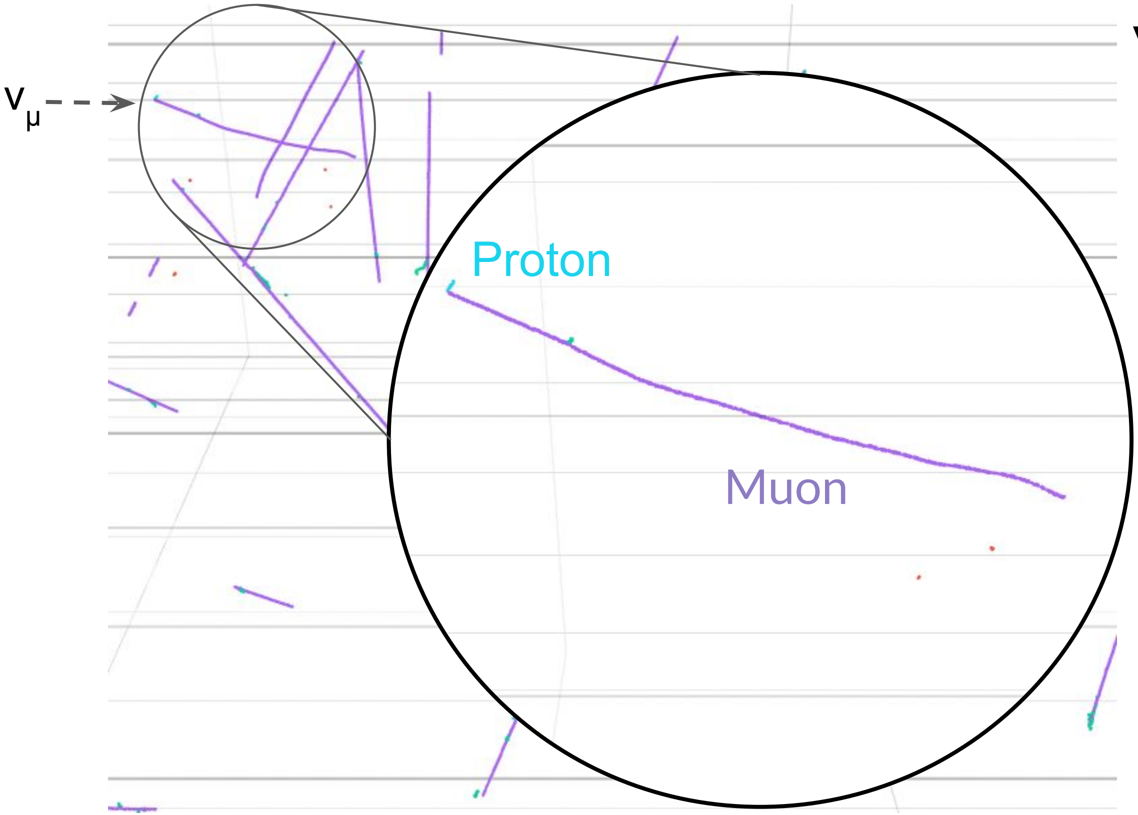
LArTPC Simulation Test Case

Realistic **Neutrino + Cosmic** ICARUS simulation as a **benchmark**

- **One (two) ν_μ** (/ν_e) + Ar interaction/image
- **~25 cosmic** interactions/image (surface detector)



ν_μ selection in ICARUS

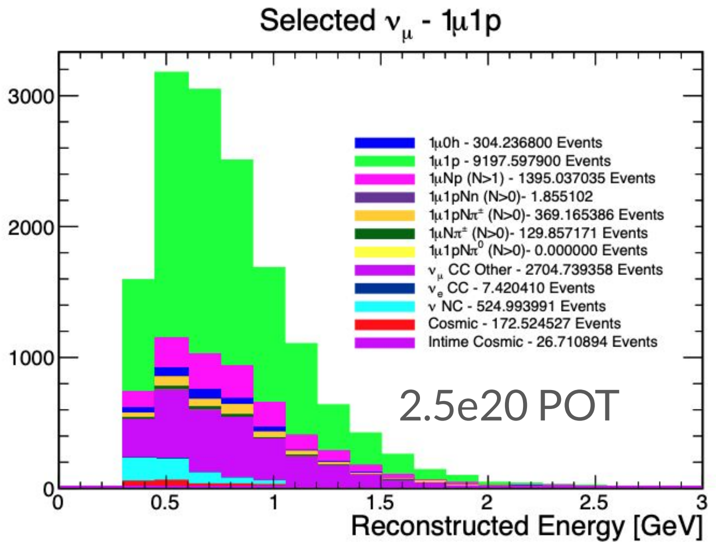


ν_μ-CCQE Selection

- Topology: **1μ1p**
- **Simplest topology** to reconstruct
- First test that **cosmic ray removal** works
- First test to ensure the **reconstruction is working at a basic level**

ν_μ selection in ICARUS

Traditional (PandoraPFA)	ML (ours)



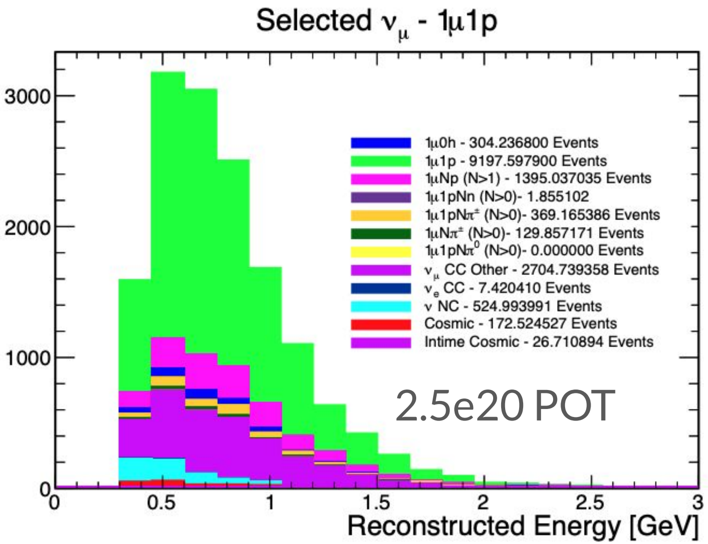
Purity: **62.2 %**

Efficiency: **40.0 % (9198 events)**

ν_μ selection in ICARUS

SLAC NATIONAL ACCELERATOR LABORATORY

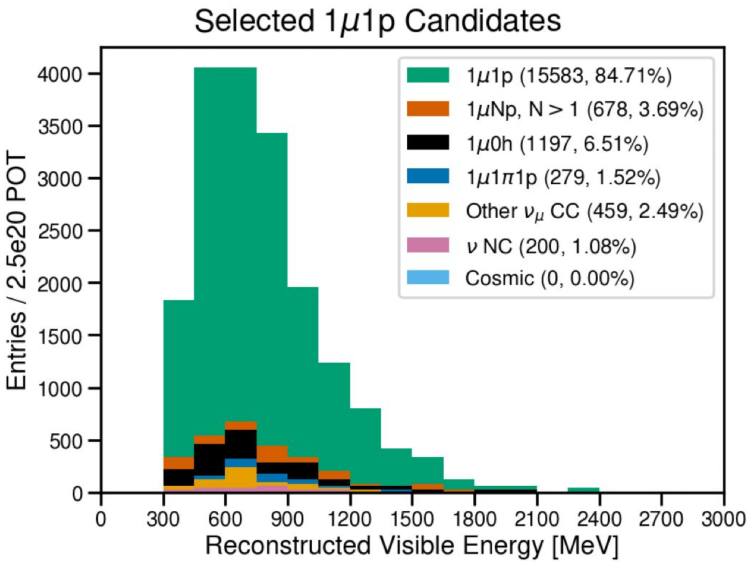
Traditional (PandoraPFA)



Purity: **62.2 %**

Efficiency: **40.0 %** (9198 events)

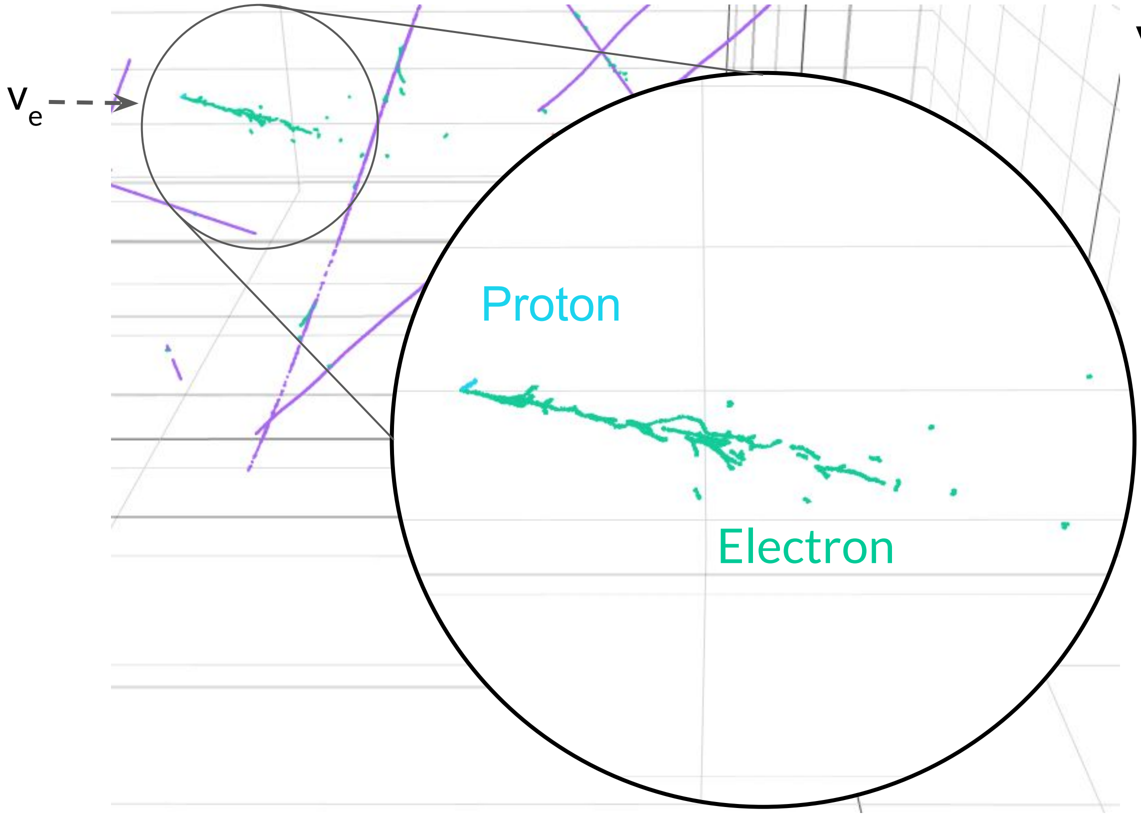
ML (ours)



Purity: **84.4 %**

Efficiency: **67.8 %** (15583 events)

ν_e selection in ICARUS

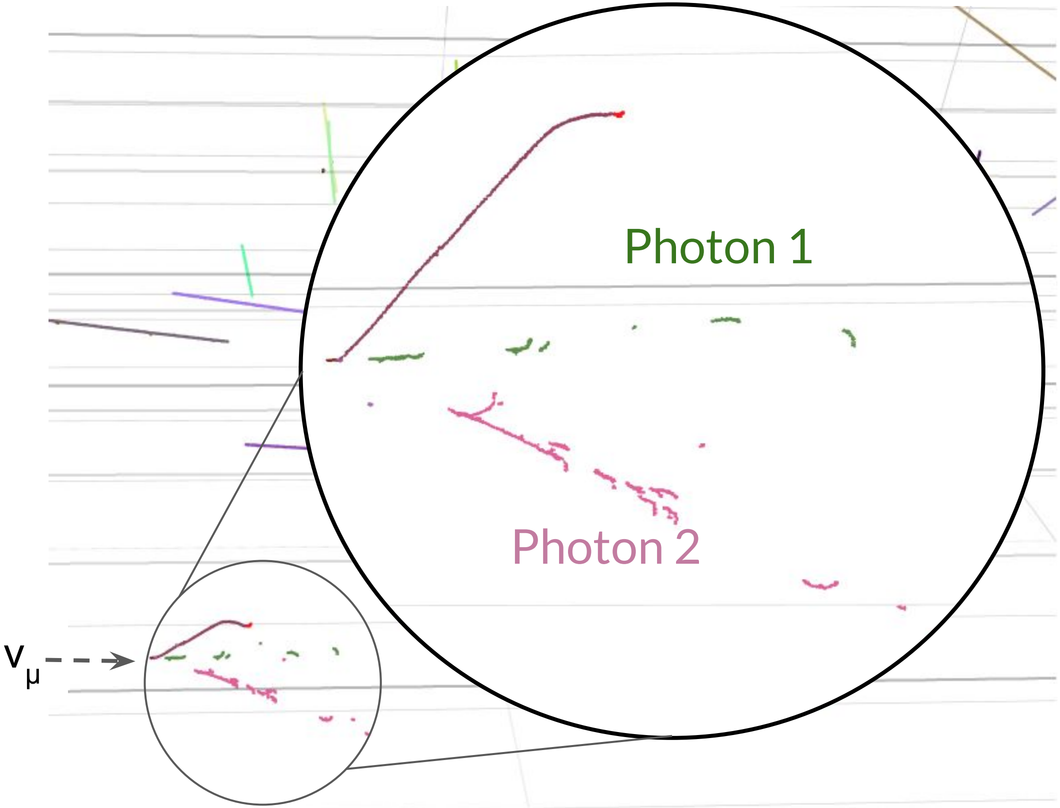


ν_e-CCQE Selection

- Topology: **1e1p**
- **Flashship measurement** in SBN (low-energy ν_e excess)
- Signal to background ratio in the beam: $O(10^{-5})$

PandoraPFA	ML (ours)
Purity: 67.0 % Efficiency: 25.3 %	Purity: 78.3 % Efficiency: 62.9 %

Invariant mass of π^0 in ICARUS



Neutral pion invariant mass:

- Standard candle for shower energy scale

$$m_{\pi^0} = \sqrt{2E_1E_2(1-\cos\theta)}$$

- Only **calibration source** for **EM shower energy**

PandoraPFA	ML (ours)
Resolution: 19.8 %	Resolution: 12.1 %

More Physics on the Horizon in SBN

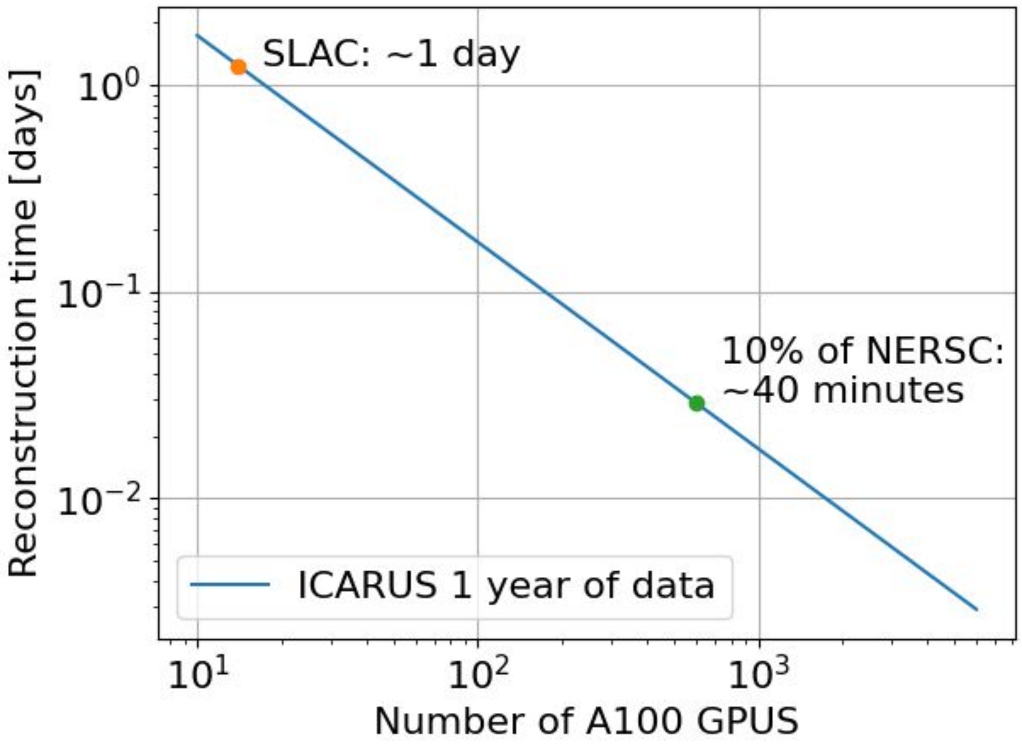
First DUNE-FD module still 4 years away…

SBN will produce plenty of interesting physics until then:

- Short baseline **oscillation** test (MiniBooNE anomaly)
- Rich **cross-section** program (arXiv:1903.04608):
    - NuMI off-axis @ ICARUS: 10 k ν_e / year, higher energy than BNB (up to 3 GeV)
    - BNB @ SBND: 2 M ν / year, O(1) k Λ^0 / Σ^+ hyperons, ~400 ν_e – e scattering

→ **ML chain essential to deliver on these physics goals**

Scalability



On **ICARUS**:

- **1 s / event,** leveraging **GPU acceleration**
  - Pandora: 40 s / event
- ~1.5 M beam events / yr

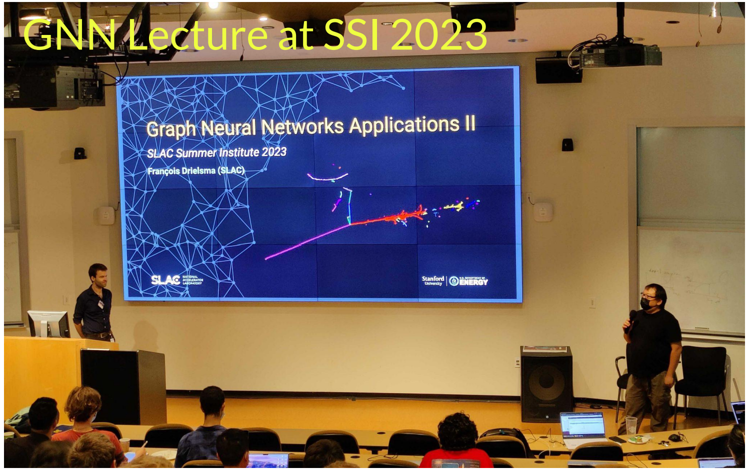
Implications:

- Fast **software development** (testing)
- Fast **turnaround**

Impact Beyond Neutrino Physics

Outreach: familiarize physics the community with ML tools

- Targeted ML workshops and schools
  - 2 ICARUS/SBND, 2 SLAC ML Schools, KMI Nagoya ML School (astrophysics), SSI
- Open source software stack and first public LArTPC neutrino dataset

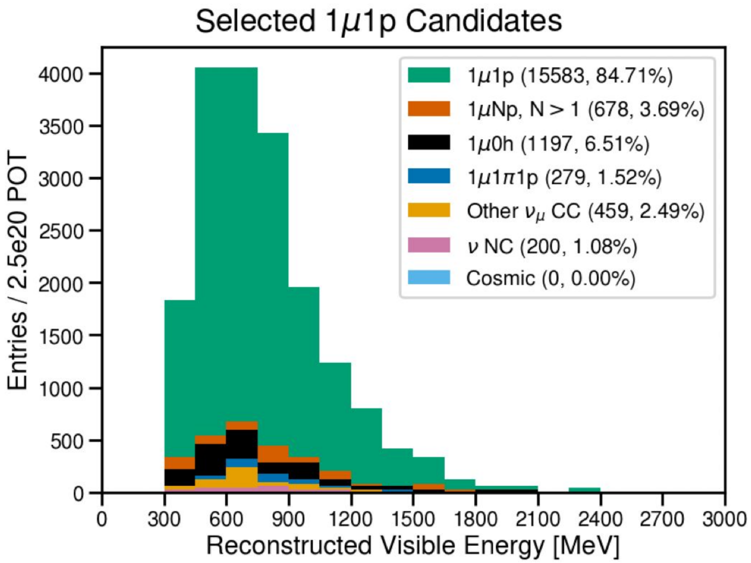



3. The Age of Machines

Performance Drop on Data

SLAC NATIONAL ACCELERATOR LABORATORY

ML on Simulation



Purity: 84.4 %

Efficiency: 67.8 %

ML on Data

Automatic selection T_p>50 MeV	selected as 1mu+1p
1mu + 1 p	109
1mu + 2p	9
1mu + 3p	
Neutrinos not contained	8
Other interactions	9
Cosmics	
TOTAL EVENT SELECTED	135
Purity	109/135 =>80.7%
FLAT SCAN	52

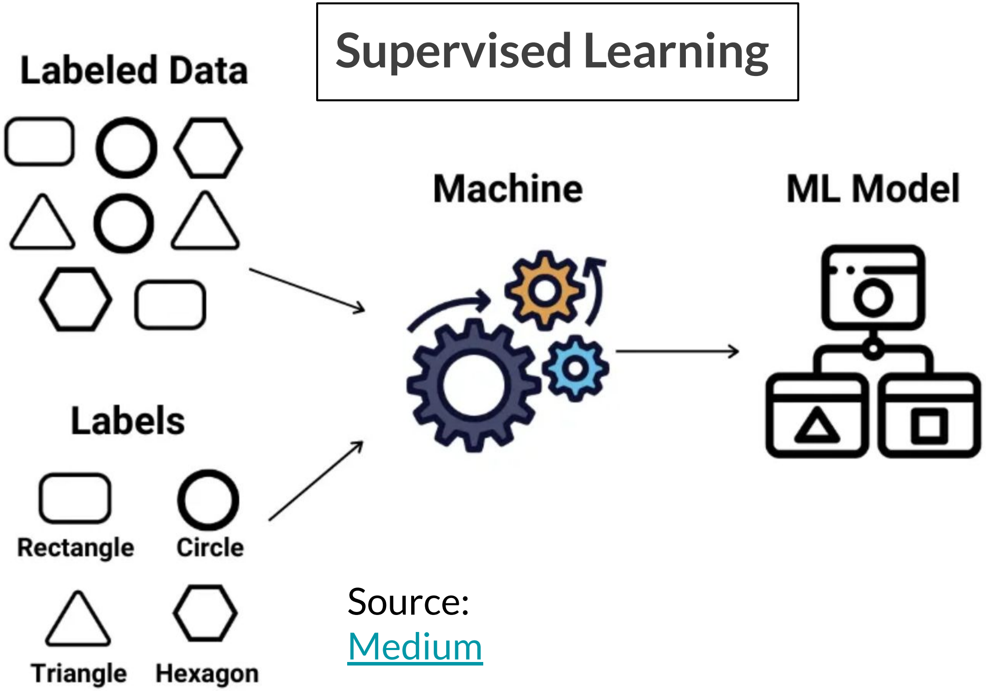
Purity:
- Handscan selected events

Efficiency:
- Reconstruct handscanned events

Purity: 80.7 %

Efficiency: 52 % (target: >90 %)

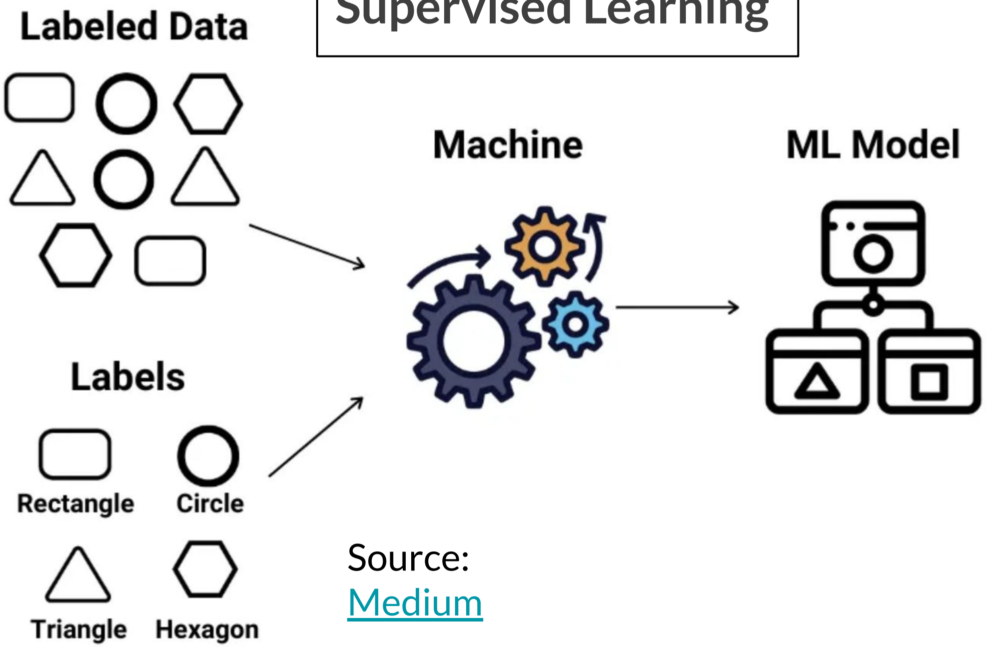
Domain Shift

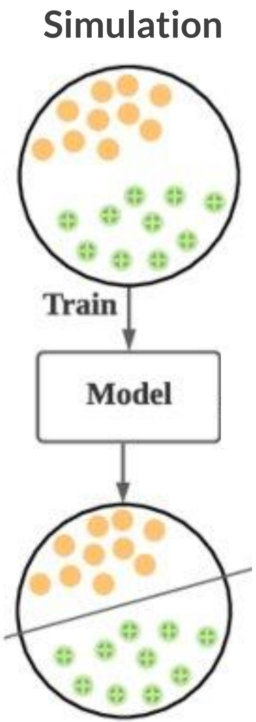


Source: [Medium]

Domain Shift

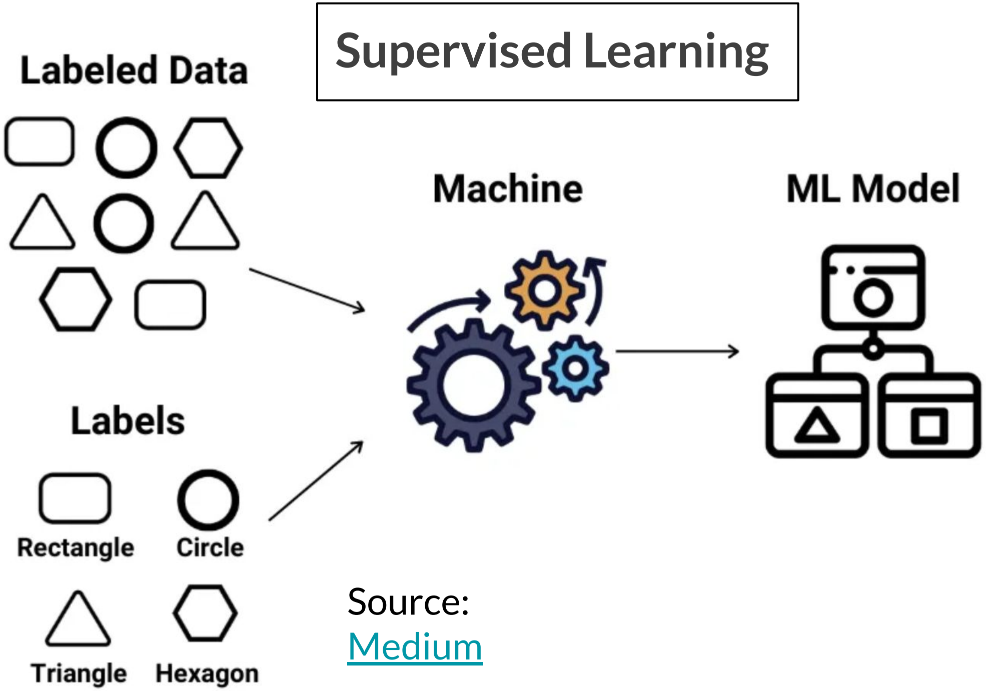
Supervised Learning



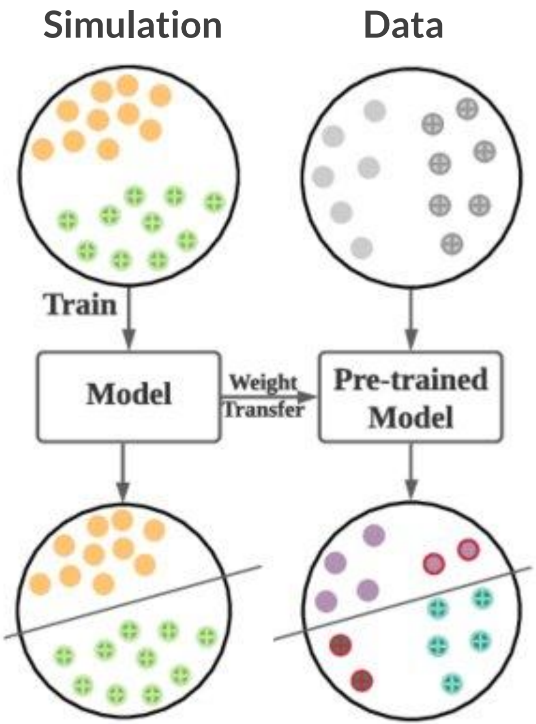


Source: [Medium]

Domain Shift



Source:
[Medium]

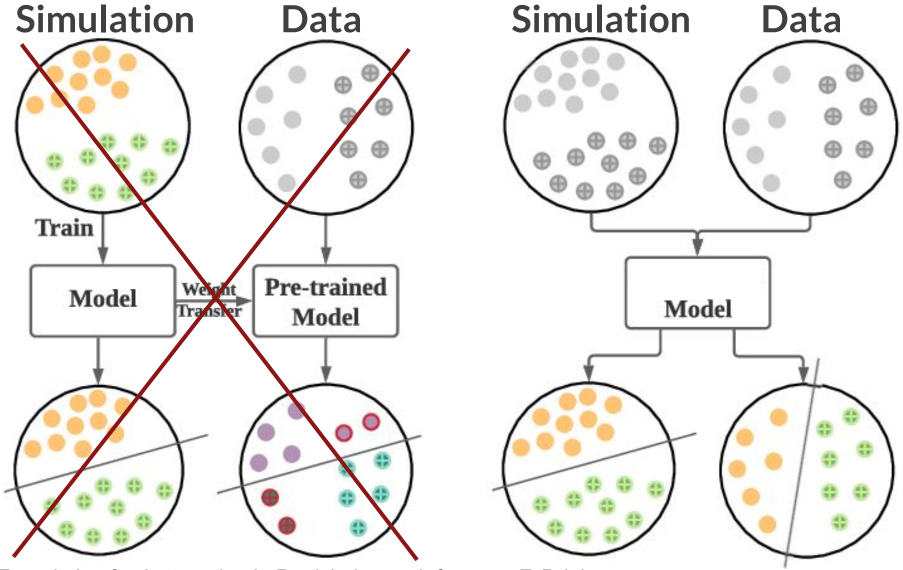


Domain shift!

Addressing the Domain Shift

What if we could train **directly on data?**

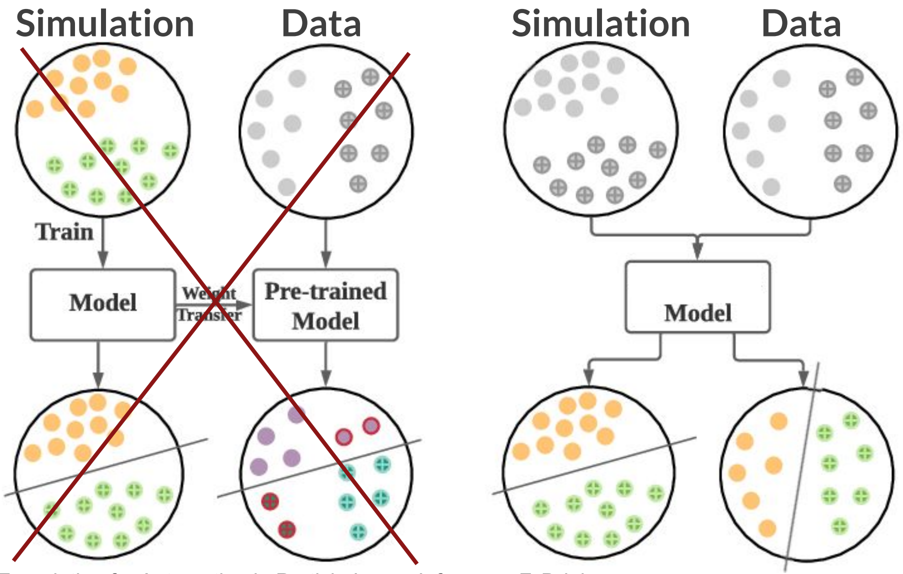
- Start from raw data = **target domain** → no domain shift!
- Reduces detector systematics



Addressing the Domain Shift

What if we could train **directly on data?**

- Start from raw data = **target domain** → no domain shift!
- Reduces detector systematics



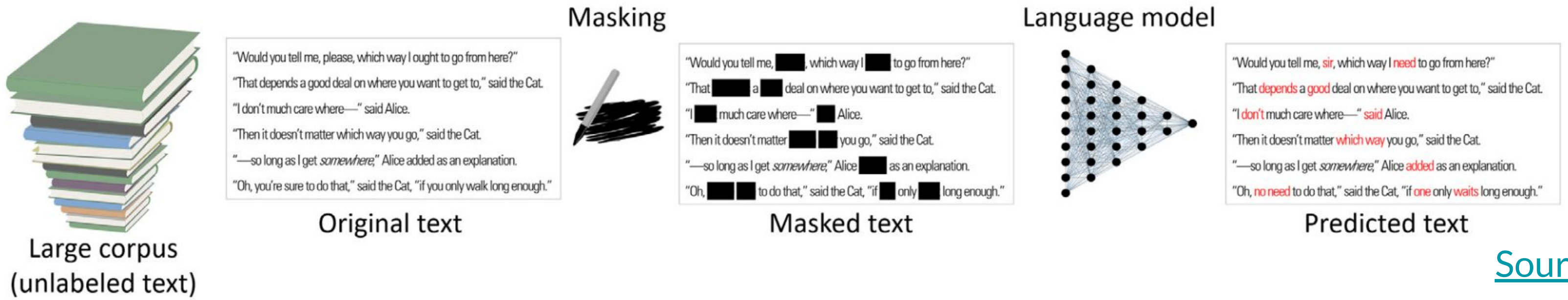
Elephant in the room:

- Raw data has **no obvious labels**, how can we train a network on it?

Foundation for Automation in Particle Image Inference, F. Drielsma

Self-Supervision in Large Language Models

Simple idea: **mask known areas** of a text, try to **reconstruct it**



Source

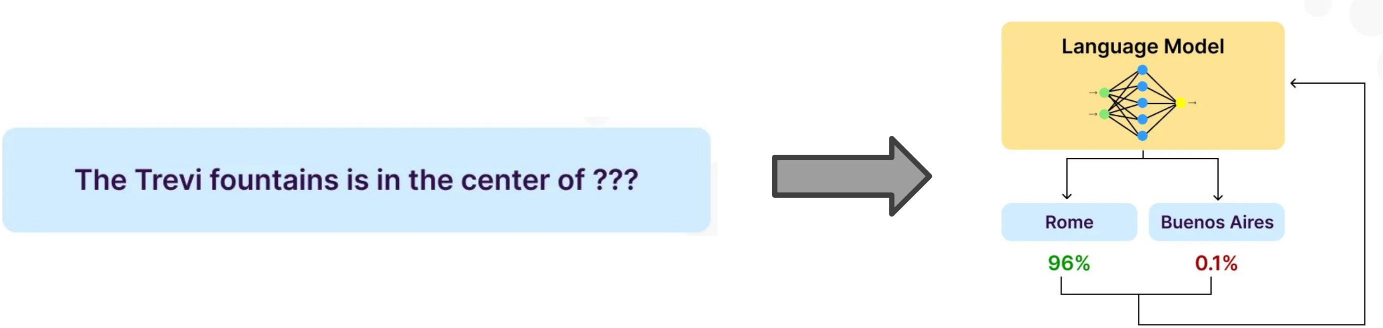
Masked Self-Supervision Learning (SSL)

Self-Supervision in Large Language Models

SLAC NATIONAL ACCELERATOR LABORATORY

Self-supervision works **incredibly well** for **Large Language Models (LLMs)**

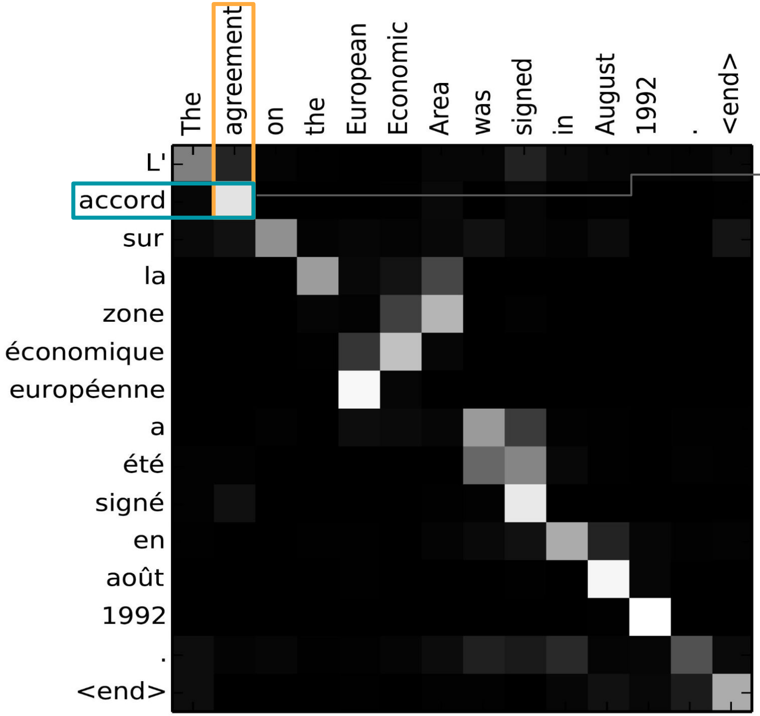
Source



Source

Transformers

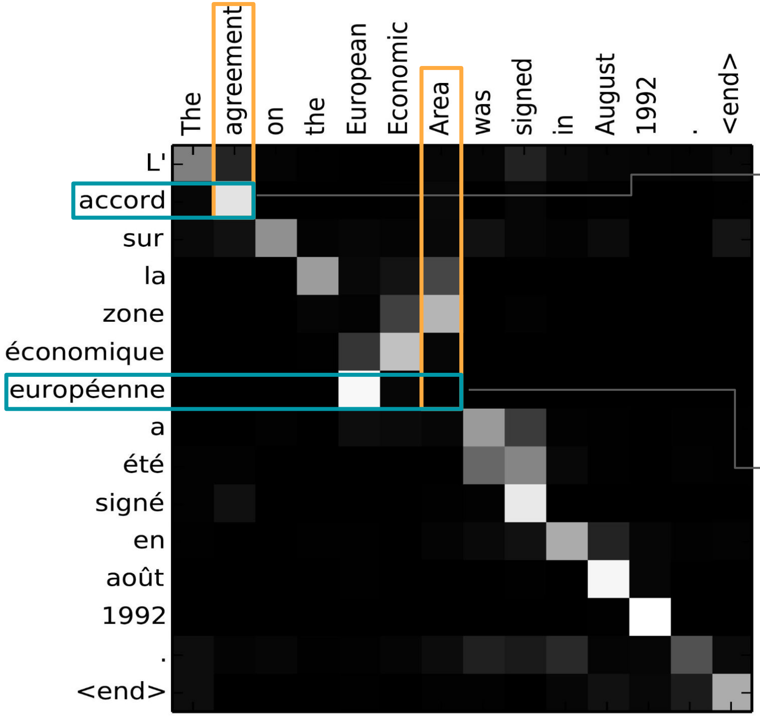
LLMs use **transformers** to learn **correlations between words** in a sentence



Words with the **same meaning** are **strongly correlated**

Transformers

LLMs use **transformers** to learn **correlations between words** in a sentence

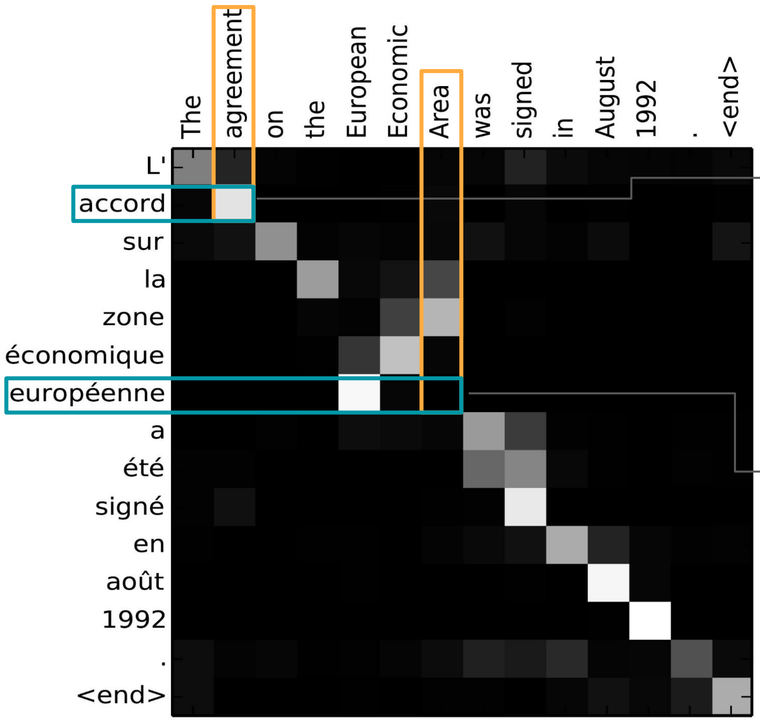


Words with the **same meaning** are **strongly correlated**

Unrelated Words are **not correlated**

Transformers

LLMs use **transformers** to learn **correlations between words** in a sentence



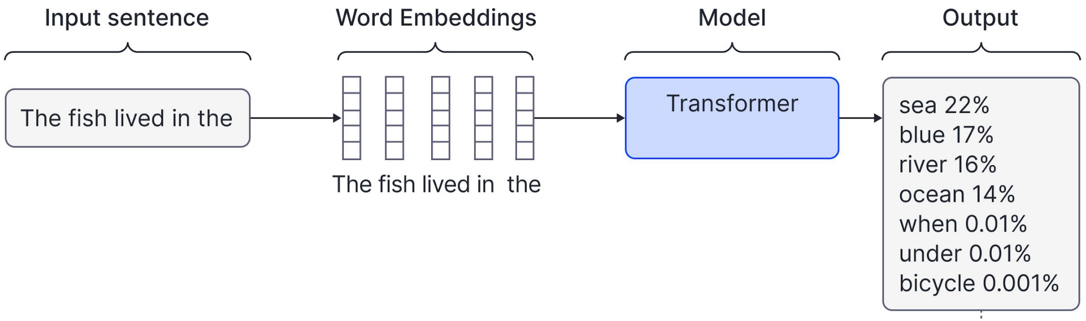
Words with the **same meaning** are **strongly correlated**

Unrelated Words are **not correlated**

Complexity: N^2

Transformers

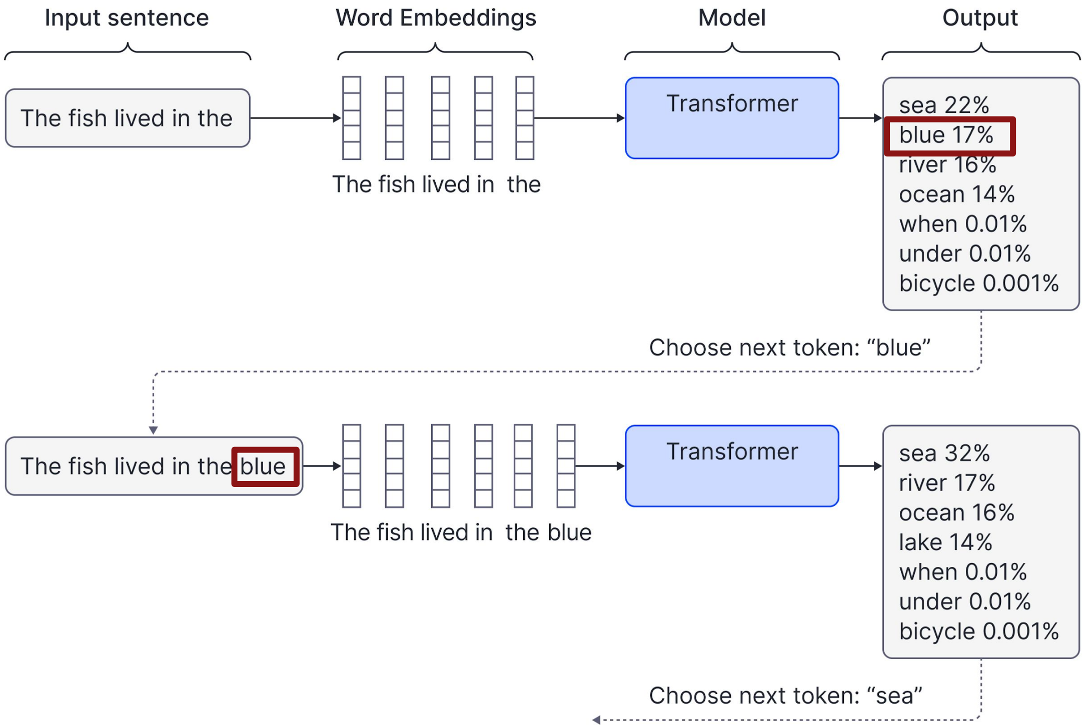
Transformers can use **learned correlations** to **guess the next word in a sequence**



Most likely next word: **sea**

Transformers

Transformers can use **learned correlations** to **guess the next word in a sequence**



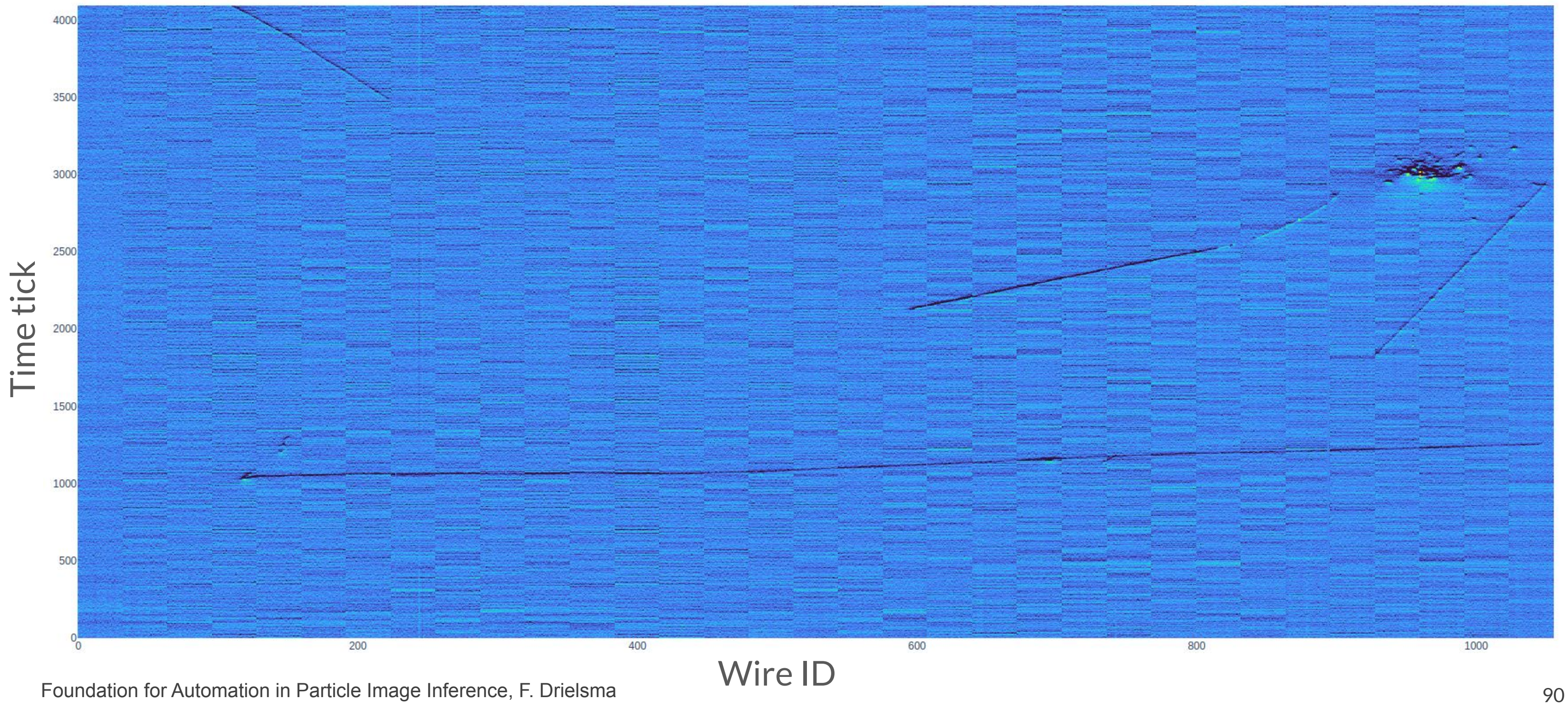
Most likely next word: **sea**, but let’s pick **blue** instead

Most likely next word: **sea**

…

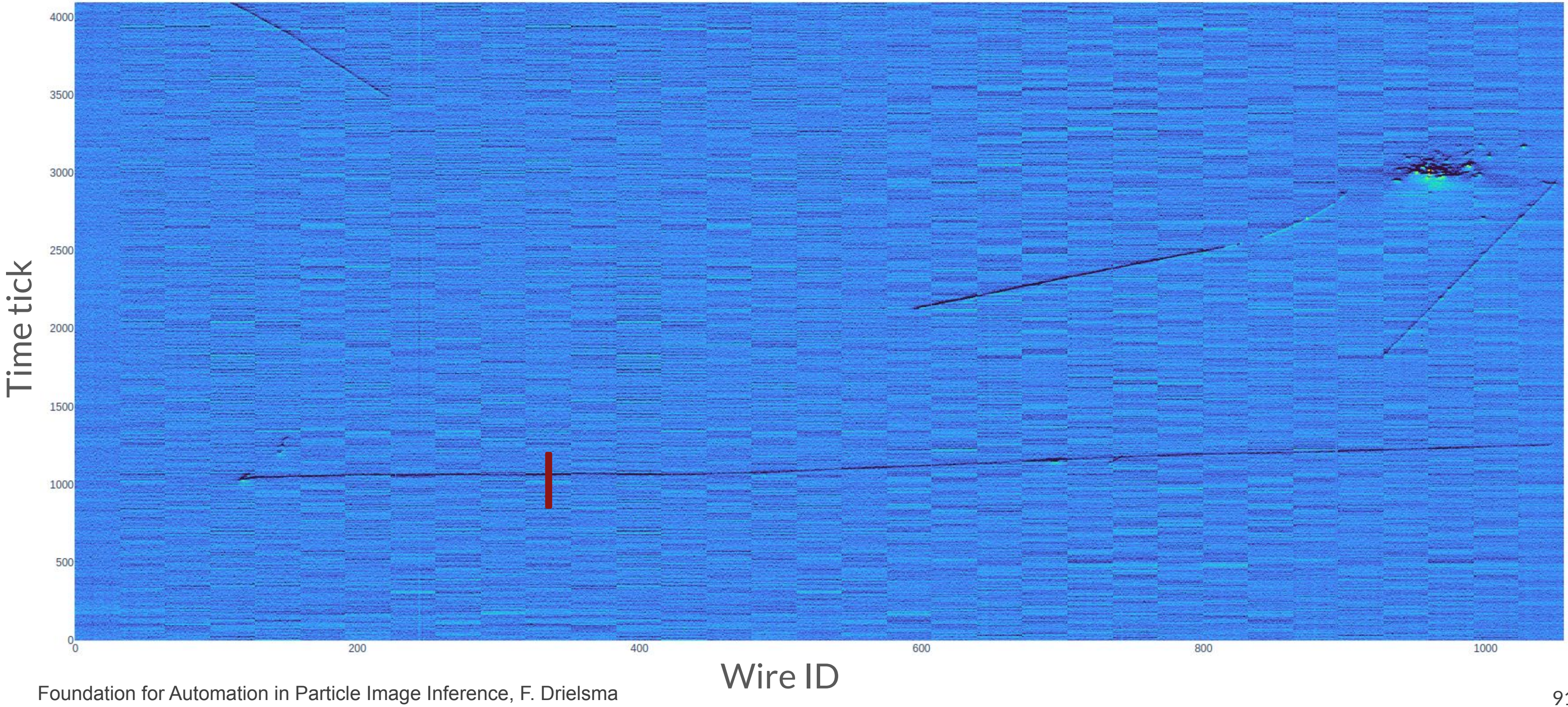
Raw LArTPC data

This is what **LArTPC data looks like raw…**



Raw LArTPC data

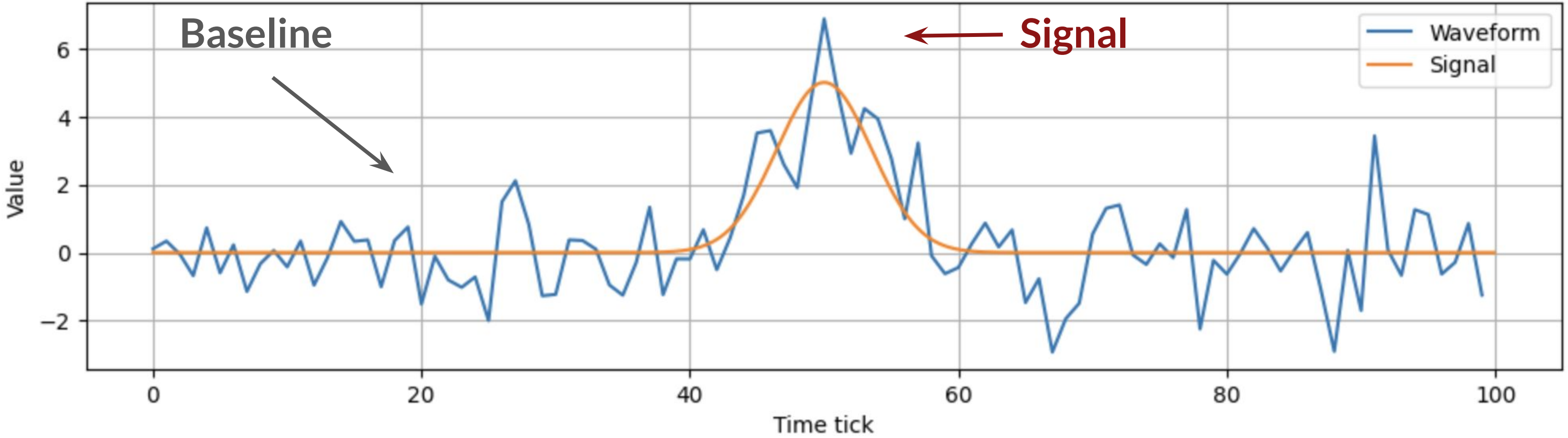
This is what **LArTPC data looks like raw…**



Language of Detectors: the Waveform

At **the most basic level**, most detectors produce **waveforms**

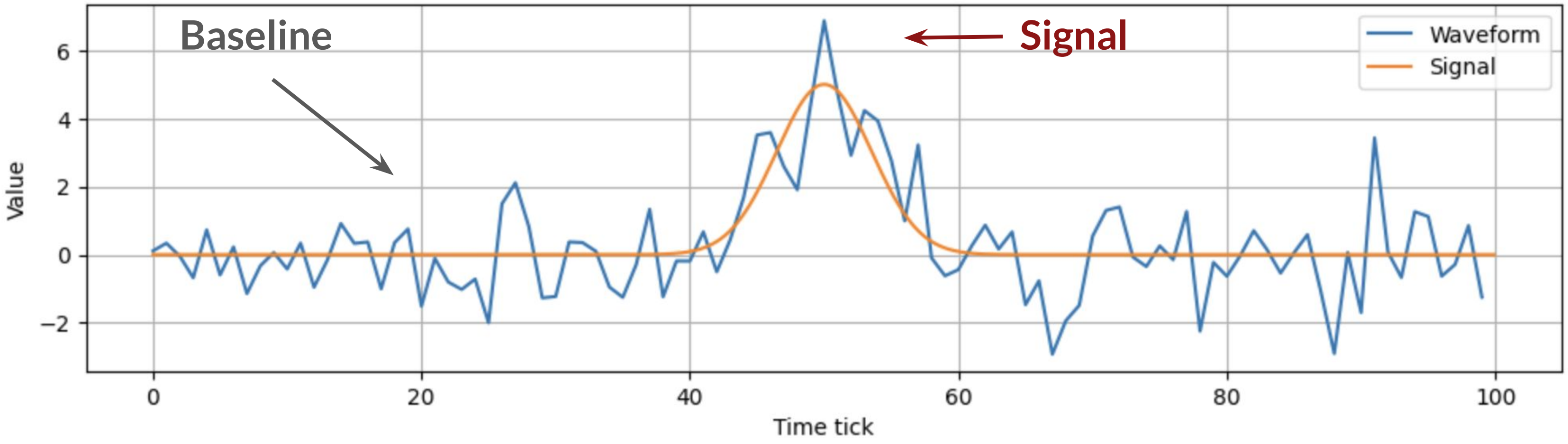
- Response of electronics to a charge signal (collected electrons)



Language of Detectors: the Waveform

At **the most basic level**, most detectors produce **waveforms**

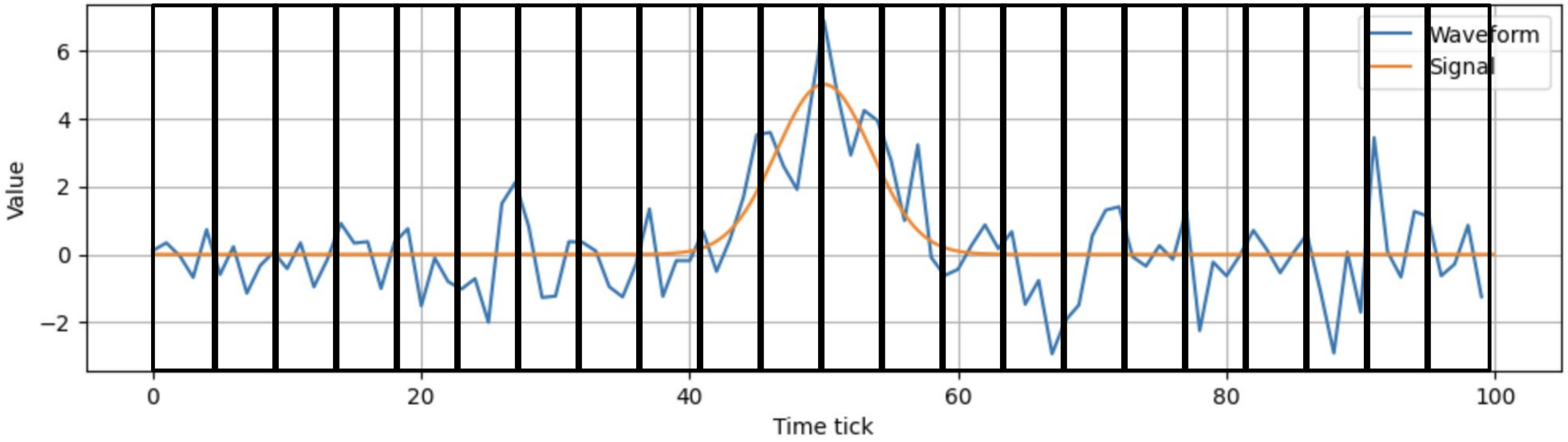
- Response of electronics to a charge signal (collected electrons)
- **Information sparse**: mostly meaningless noise
- **Long**: 4096 samples **per wire** in ICARUS, O(10^4) wires



Language of Detectors: the Waveform

Can we use **Transformers**?

- **Words**: waveform chunks
- **Word representation**: value for each tick in chunk
- **Information extracted**: correlations between waveform chunks

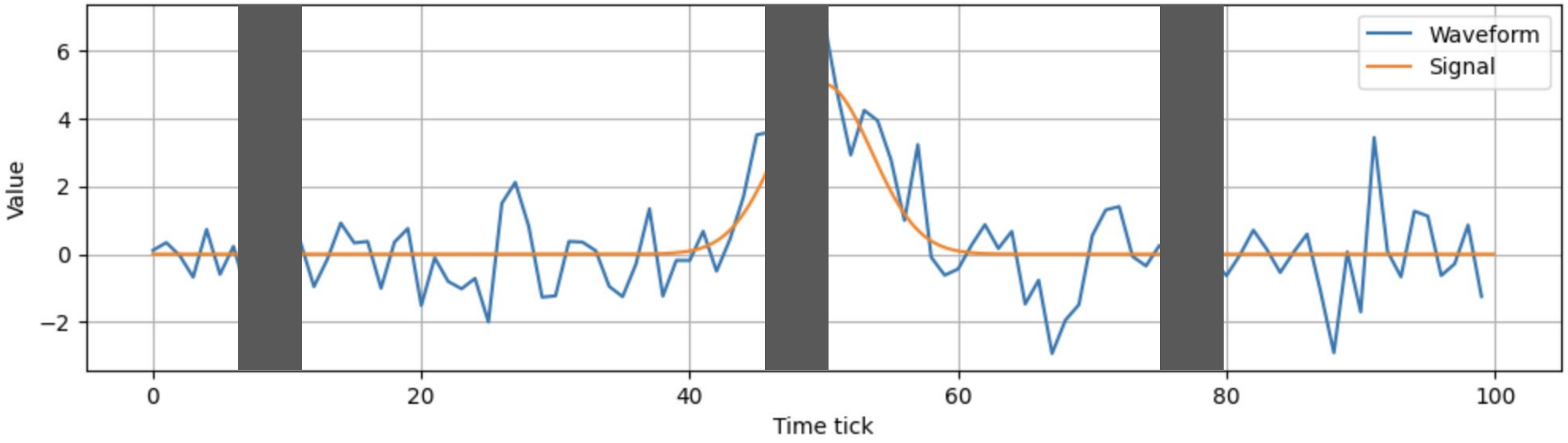


Denoising

Naive approach:

1. **Mask one to a few ticks** of the waveform
2. Train network to reproduce missing region(s), minimize: $\mathcal{L} = \sum_{i \in M} (\hat{y}_i - y_i)^2$

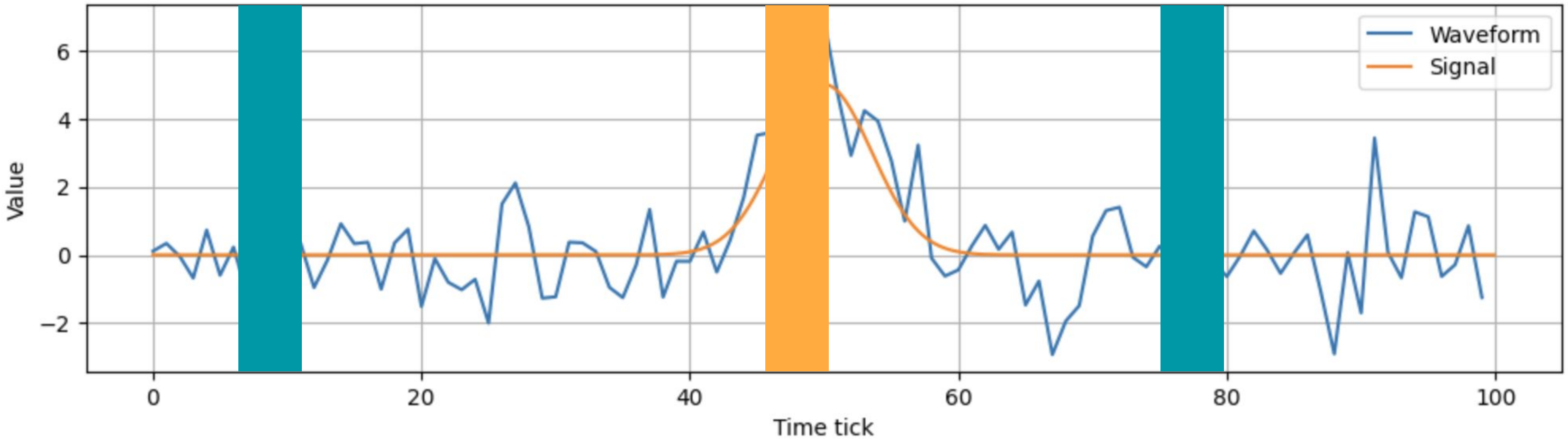
a.k.a. L2 loss



Denoising

What you get from $\mathcal{L} = \sum_{i \in M} (\hat{y}_i - y_i)^2$

1. Cannot reproduce random noise, **baseline is the best fit**
2. **Can reproduce signal**, if visible around the mask region

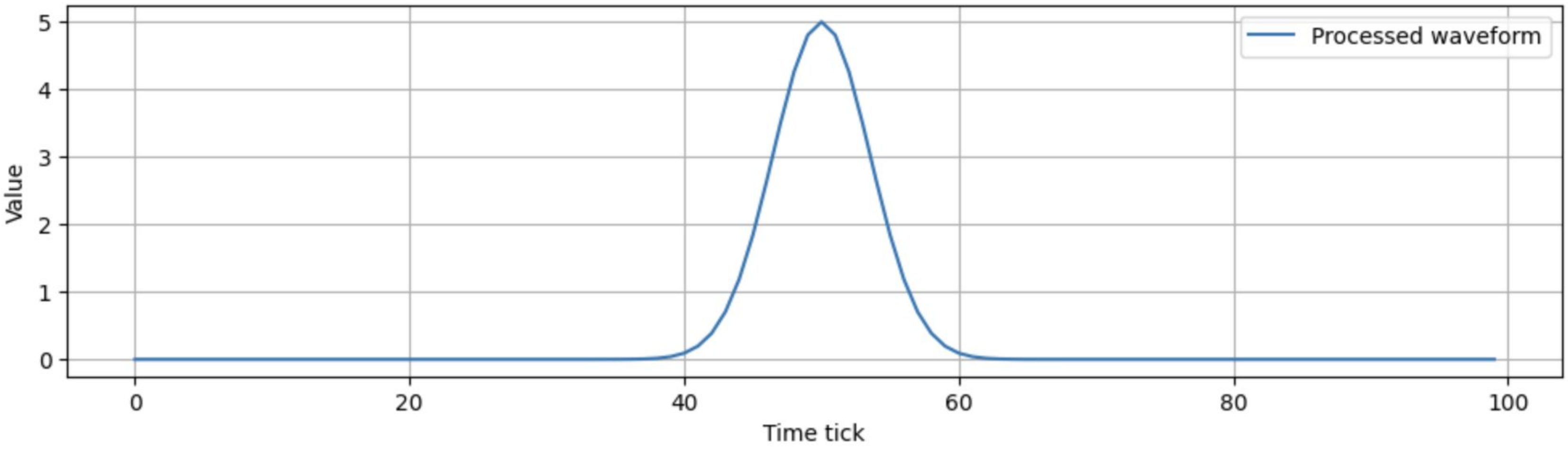


Denoising

What you get from $\mathcal{L} = \sum_{i \in M} (\hat{y}_i - y_i)^2$

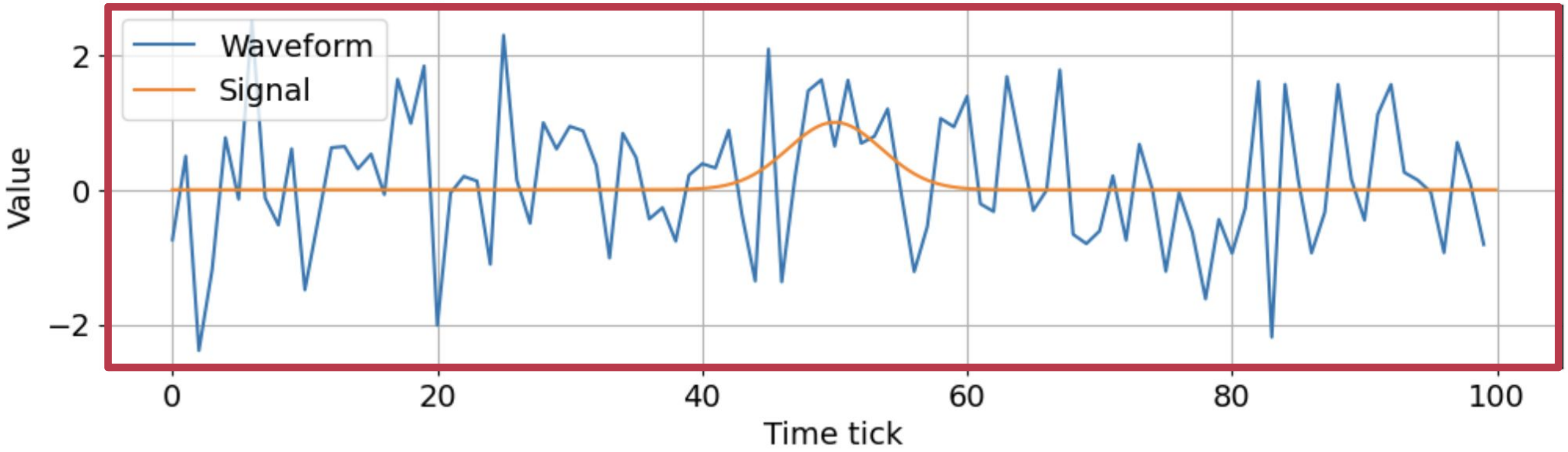
1. Cannot reproduce random noise, **baseline is the best fit**
2. **Can reproduce signal**, if visible around the mask region

→ Learn **signal shape and noise removal**



Denoising

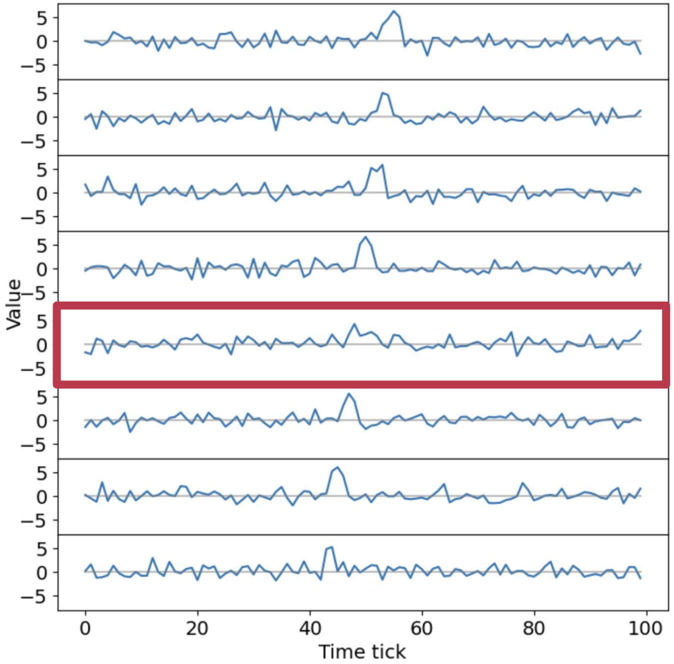
What if the **signal is buried in noise**? What if **noise looks like signal**?

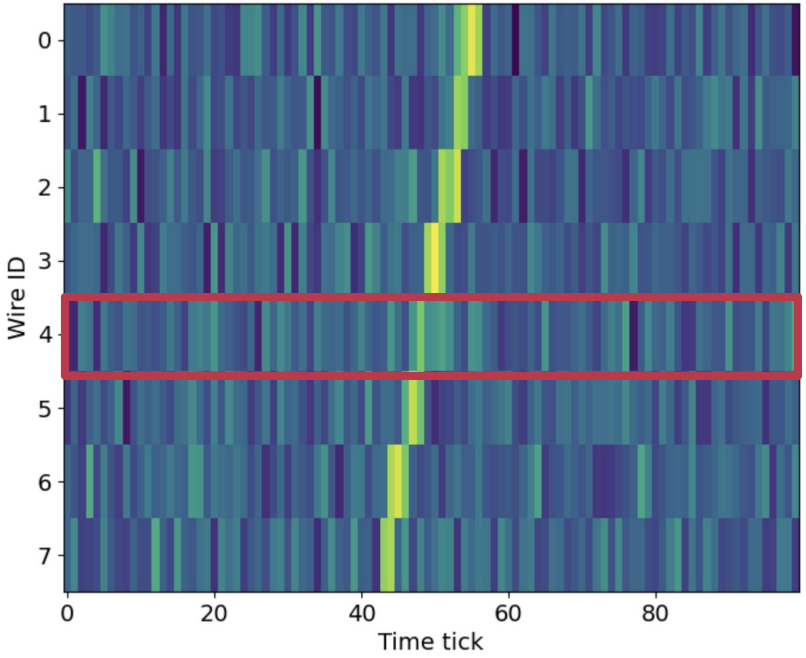


Denoising

What if the **signal is buried in noise**? What if **noise looks like signal**?

- We have **context!**

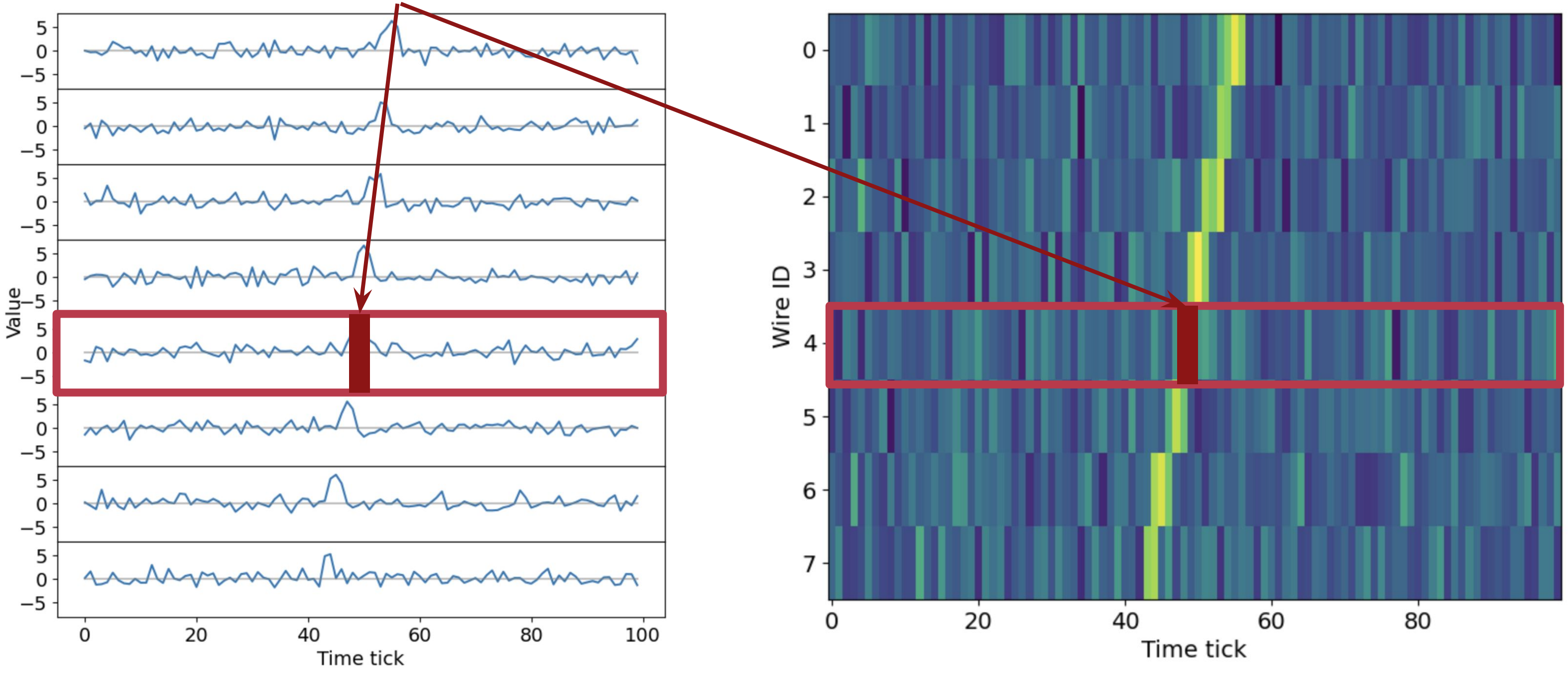




Denoising

What if the **signal is buried in noise**? What if **noise looks like signal**?

- We have **context!** It allows us to **infer dead regions effectively**

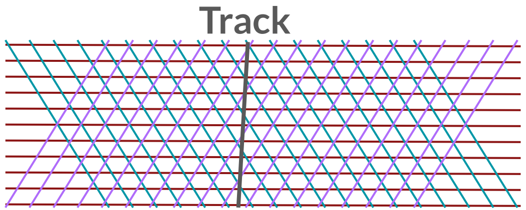


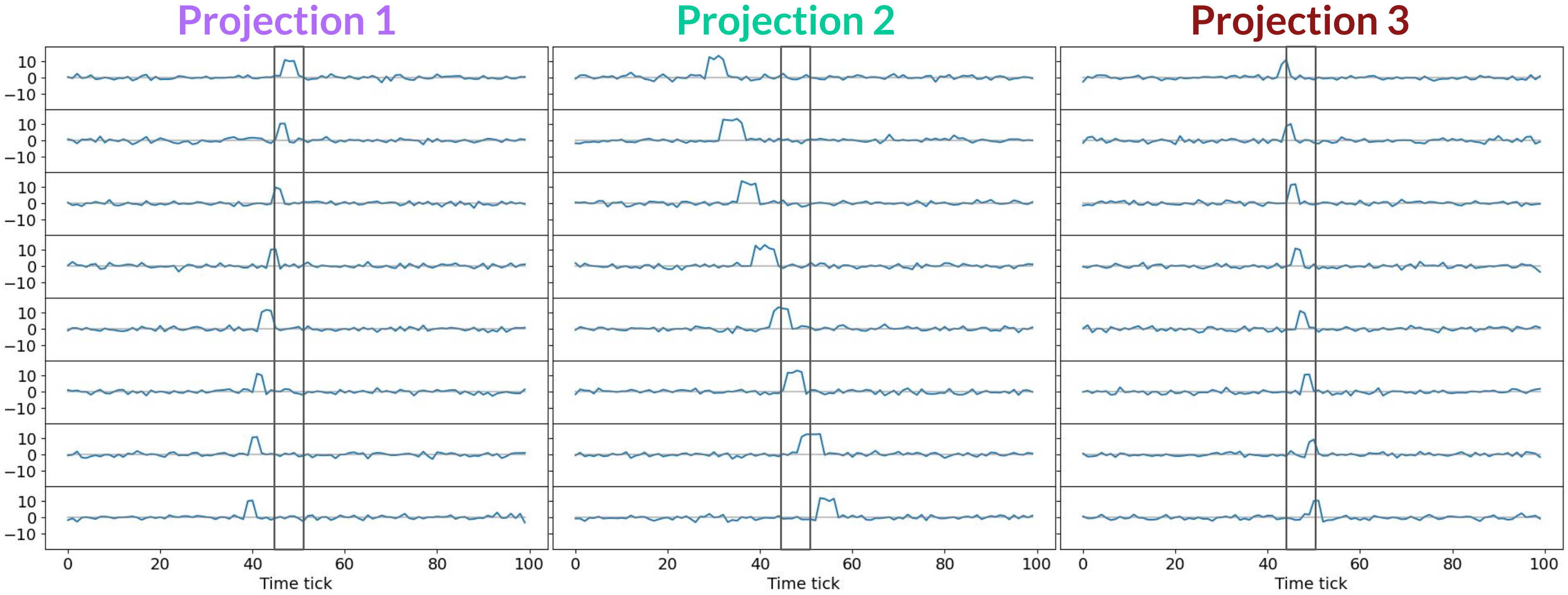
Foundation for Automation in Particle Image Inference, F. Drielsma

Tomographic reconstruction

In LArTPCs, **we have 3 projections available**

- **Game**: find correlated signals across 3 planes



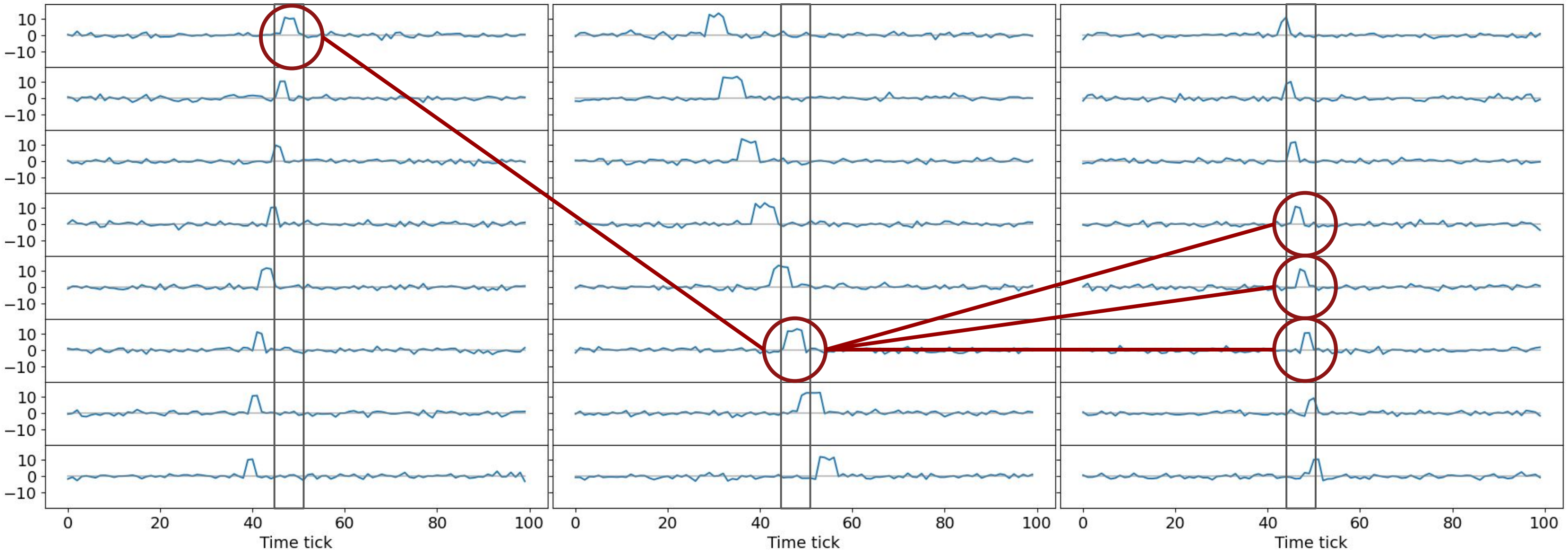


Tomographic reconstruction

In LArTPCs, **we have 3 projections available**

- **Game:** find correlated signals across 3 planes
- **Transformers are correlation machines**

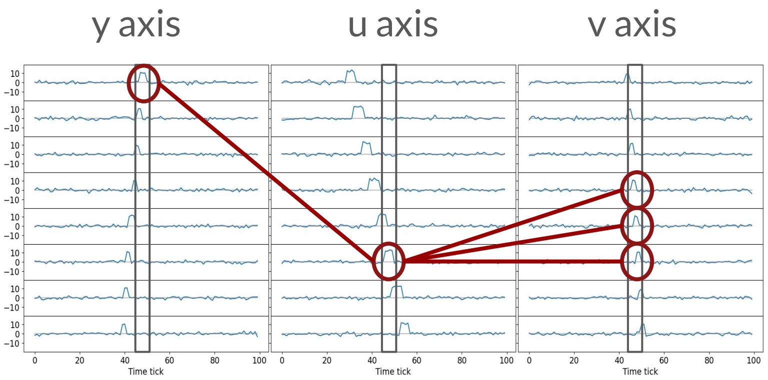
3 space points in 3D



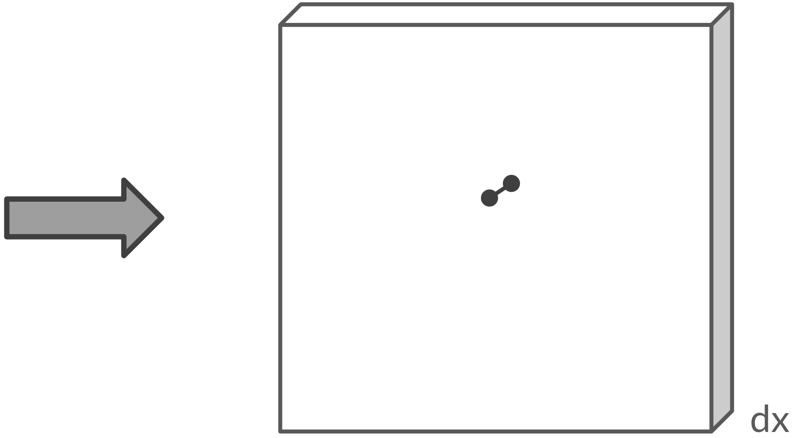
Tomographic reconstruction

In LArTPCs, **we have 3 projections available**

- **Game:** find correlated signals across 3 planes
- **Transformers are correlation machines**



Small time chunk = **slice in x**

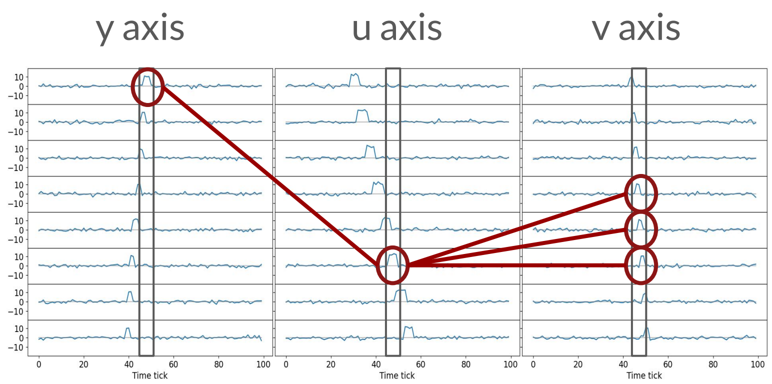


Infer space points/segment from correlated wire signals

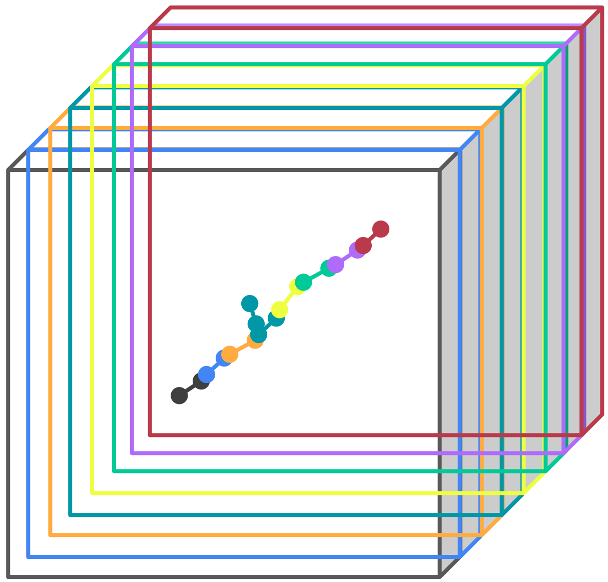
Tomographic reconstruction

In LArTPCs, **we have 3 projections available**

- **Game:** find correlated signals across 3 planes
- **Transformers are correlation machines**



Small time chunk = **slice in x**

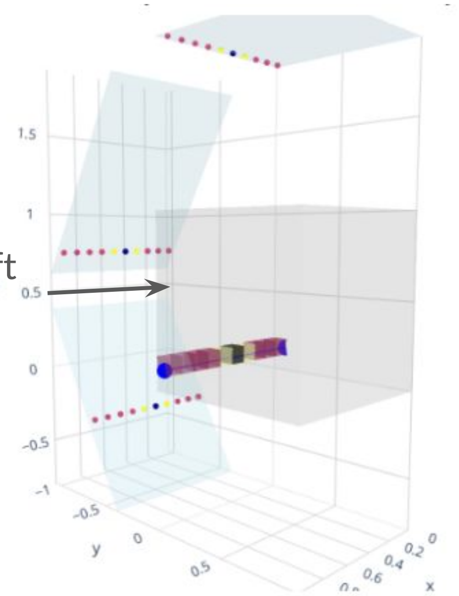


Infer space points/segment from correlated wire signals

Tomographic reconstruction

Feasibility study with GNNs: **it can work** (TomoGNN)

- Relied on 2D hits being built, **signal-based method would not**

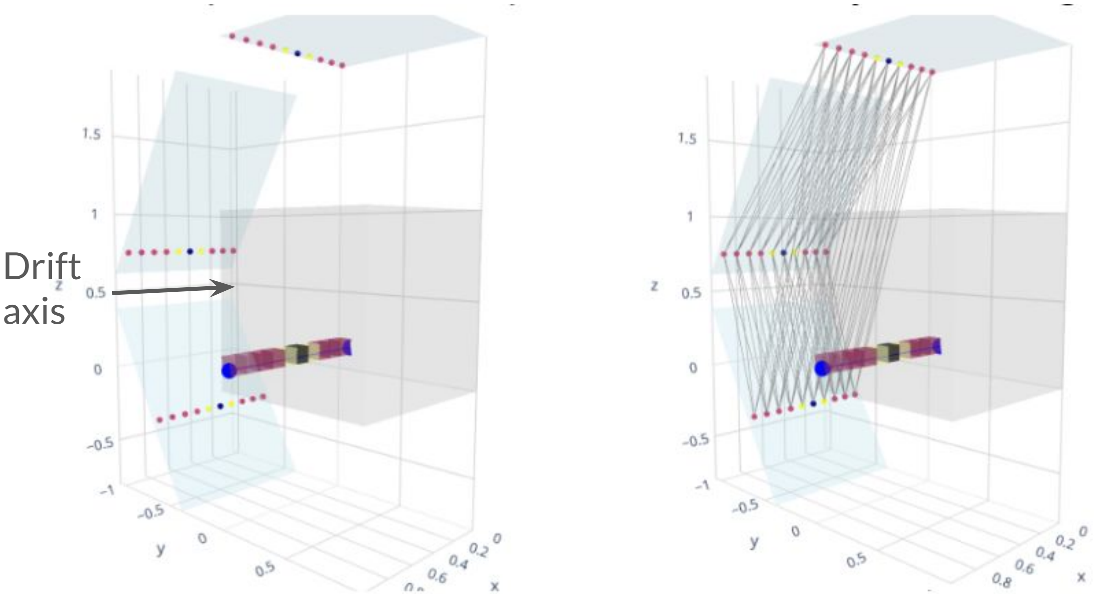


Drift axis

Tomographic reconstruction

Feasibility study with GNNs: **it can work** (TomoGNN)

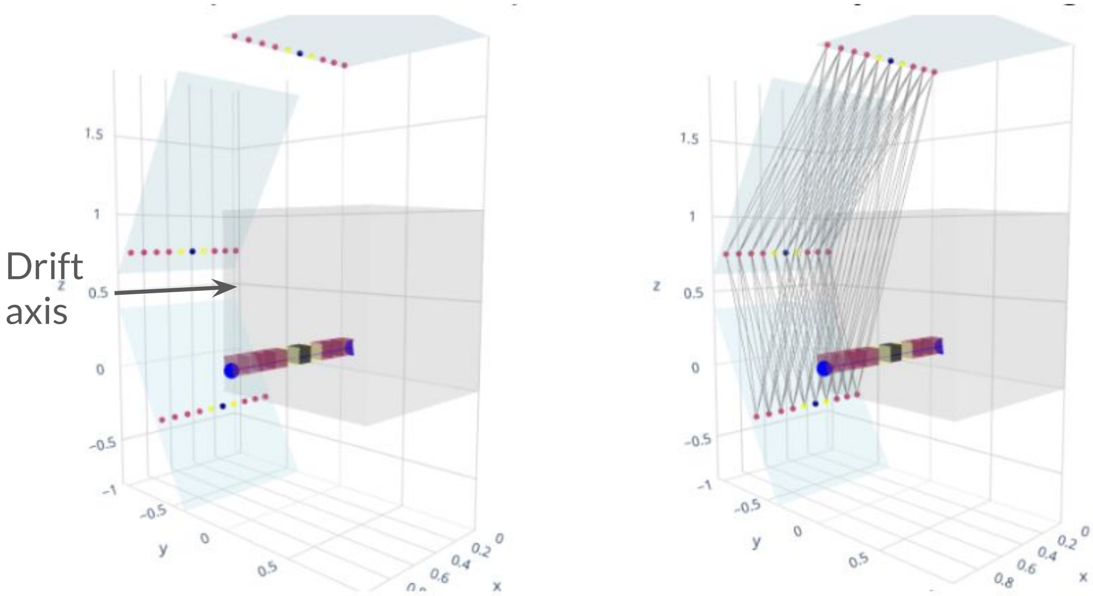
- Relied on 2D hits being built, **signal-based method would not**



Tomographic reconstruction

Feasibility study with GNNs: **it can work** (TomoGNN)

- Relied on 2D hits being built, **signal-based method would not**



Generic Artificial Intelligence

Self-supervised LLMs (e.g. GPT-4) have **very interesting properties if**

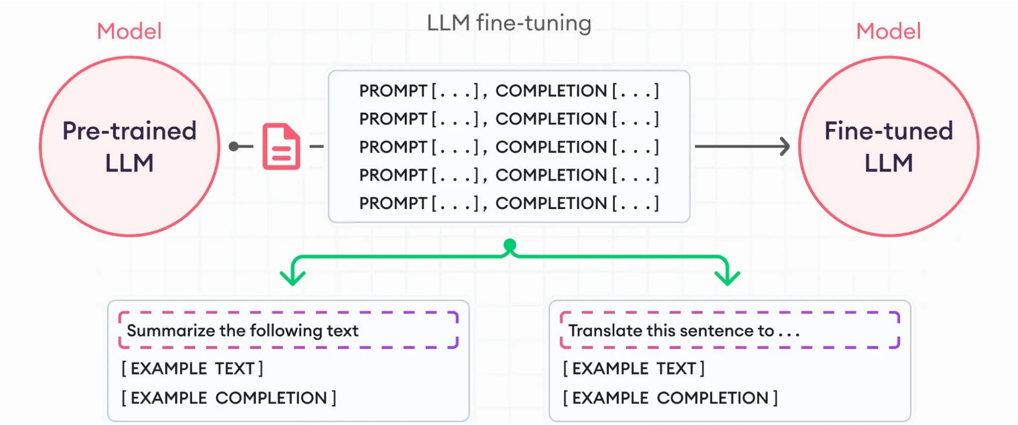
- Built on a **very complex model**: ~ 10^{12} parameters
- Given a **huge amount of raw training data**: ~ 5 x 10^{9} words (45 GB)

Generic Artificial Intelligence

Self-supervised LLMs (e.g. GPT-4) have **very interesting properties if**

- Built on a **very complex model:** ~ 10^{12} parameters
- Given a **huge amount of raw training data:** ~ 5 x 10^{9} words (45 GB)

→ Can be quickly **fine-tuned to perform specific tasks**!

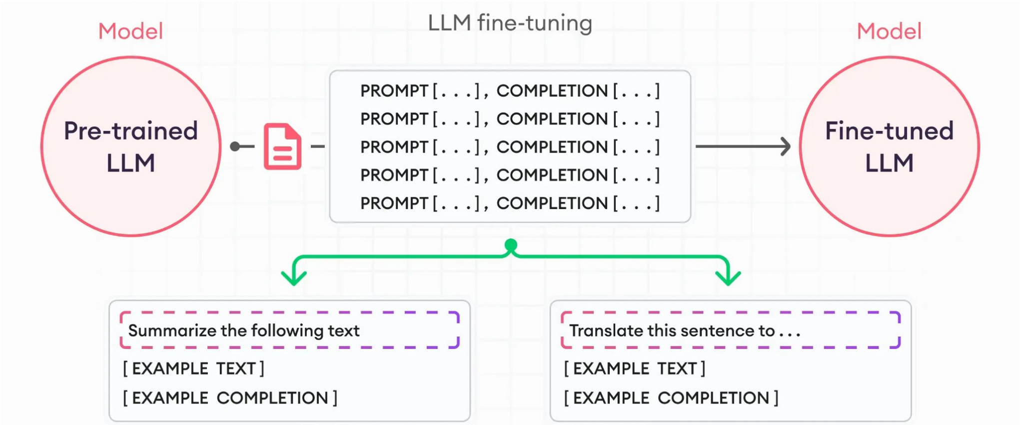


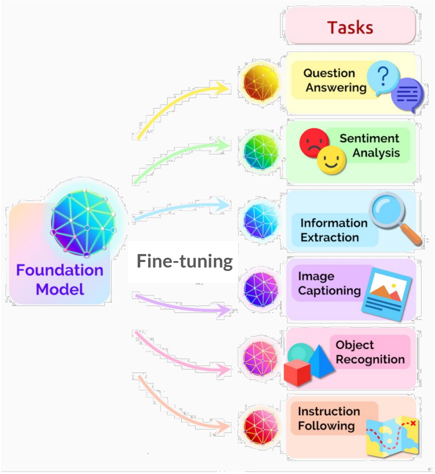
Generic Artificial Intelligence

Self-supervised LLMs (e.g. GPT-4) have **very interesting properties if**

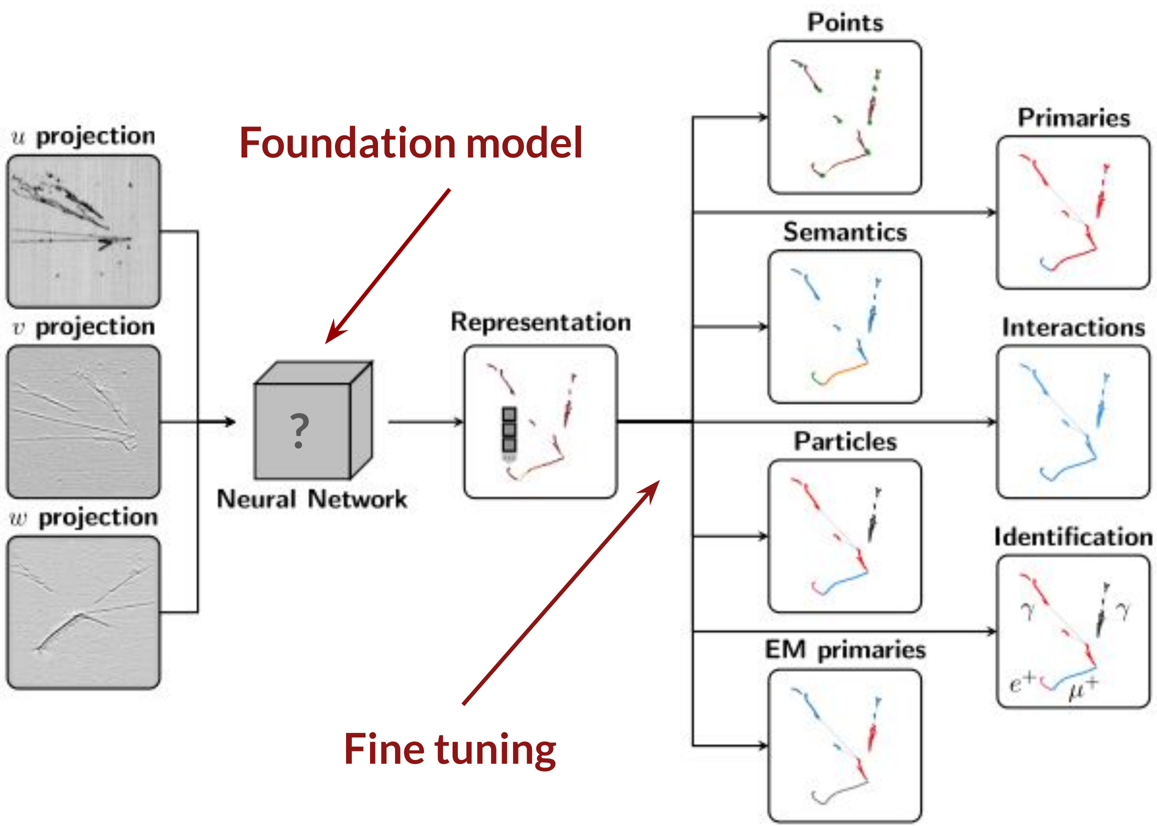
- Built on a **very complex model:** ~ 10^{12} parameters
- Given a **huge amount of raw training data:** ~ 5 x 10^{9} words (45 GB)

→ Can be quickly **fine-tuned to perform specific tasks**!





Generic Artificial Intelligence



Instead: **give a lot of raw waveform data to a large model**

- Learn a **general raw data representation** useful to all reco. tasks

Center for Research on Foundation Models

SLAC NATIONAL ACCELERATOR LABORATORY

Unique opportunity to collaborate with and **contribute to the CRFM**

- **Support and experts** to **mitigate the risks** associated with this endeavor

SLAC Colloquium - "Some Building Blocks for Foundation Model Systems" by Chris Re, Stanford University

Chris Re, Stanford University, Department of Computer Science

Center for Research on Foundation Models

HAI Stanford University Human-Centered Artificial Intelligence

Research on long sequence processing:

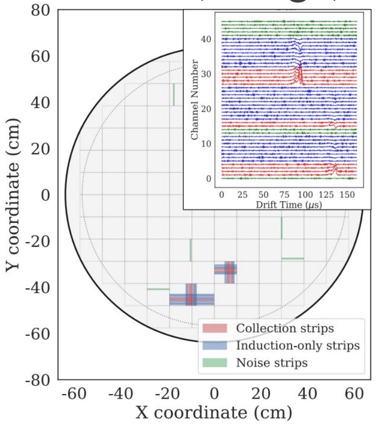
- **FlashAttention**: speed-up by localizing computations on GPU
- **S4/Mambda Model**: subquadratic implementation of a state-space model with FM properties
- **HyenaDNA**: 420k "words"

Self-Supervision beyond LArTPC

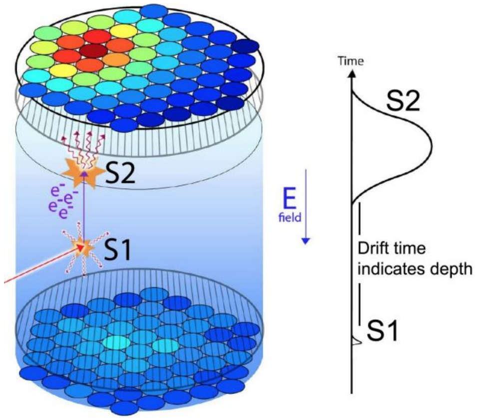
Waveform data is **ubiquitous in science and industry**

→ **Successful self-supervision** on raw waveform data has implications well beyond LArTPCs alone, **synergies at SLAC**

nEXO (charge)



LZ (light)



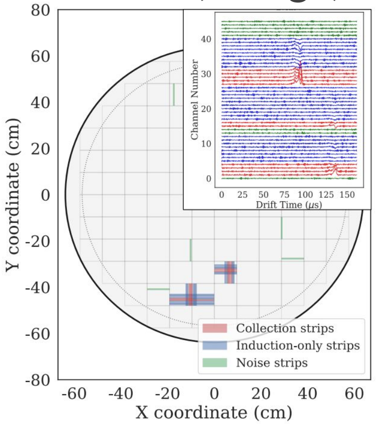
Self-Supervision beyond LArTPC

Waveform data is **ubiquitous in science and industry**

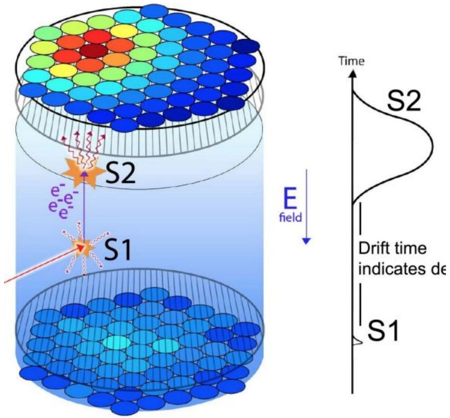
→ **Successful self-supervision** on raw waveform data has implications well beyond LArTPCs alone, **synergies at SLAC**

→ **Science for AI/ML**

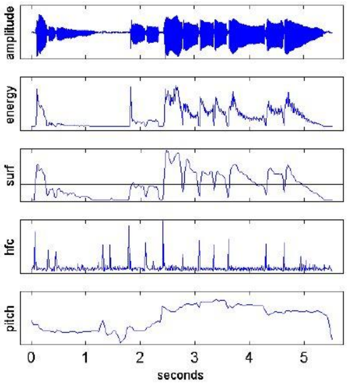
nEXO (charge)



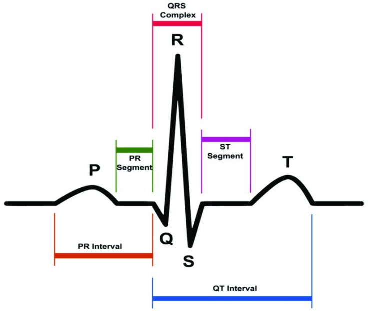
LZ (light)



Music (sound)



Heartbeat (EKG), brain (EEG)



Conclusions

LArTPCs are at the **core of the US-based accelerator neutrino program**

- **DUNE** and **SBN** cannot succeed without a **high-quality reconstruction**
- Partially **automated the reconstruction from space point to interactions**
  - New state-of-the-art on 3 LArTPC experiments
- Clear road-map towards **foundation models in waveform data**
  - Address the remaining challenges in LArTPCs and **open a new pole of research**