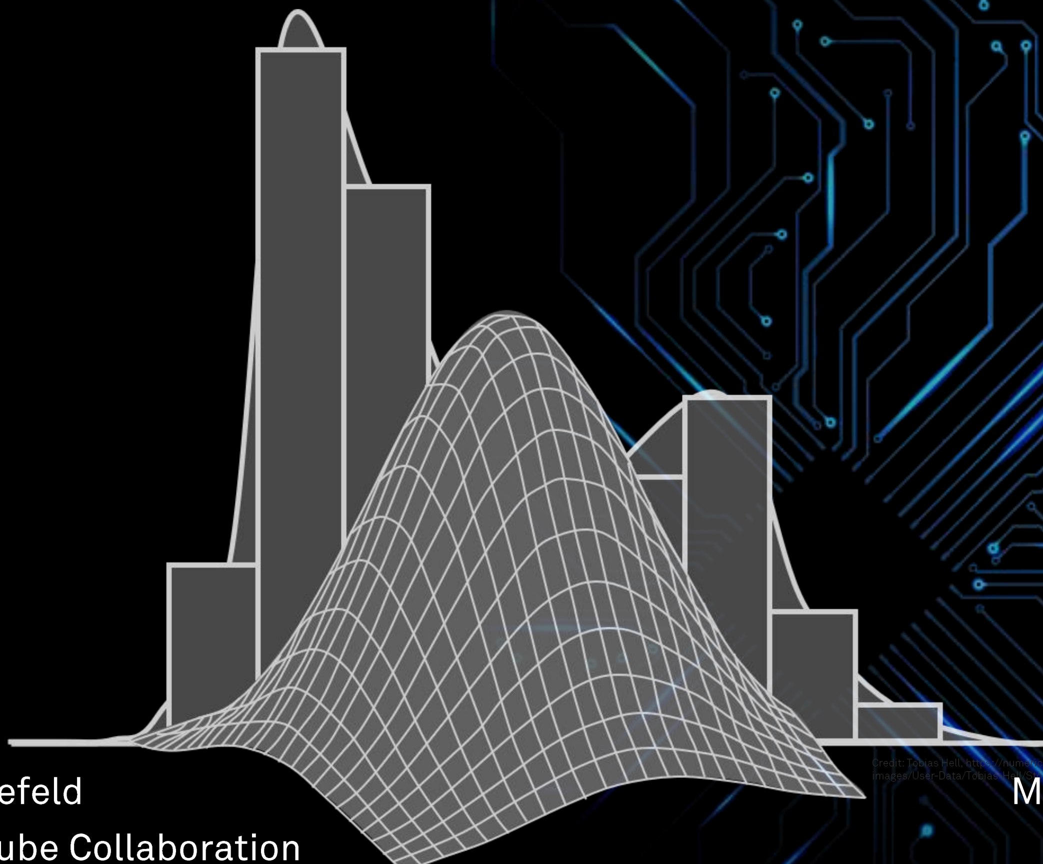
Interpretable deep learning for event reconstruction in IceCube

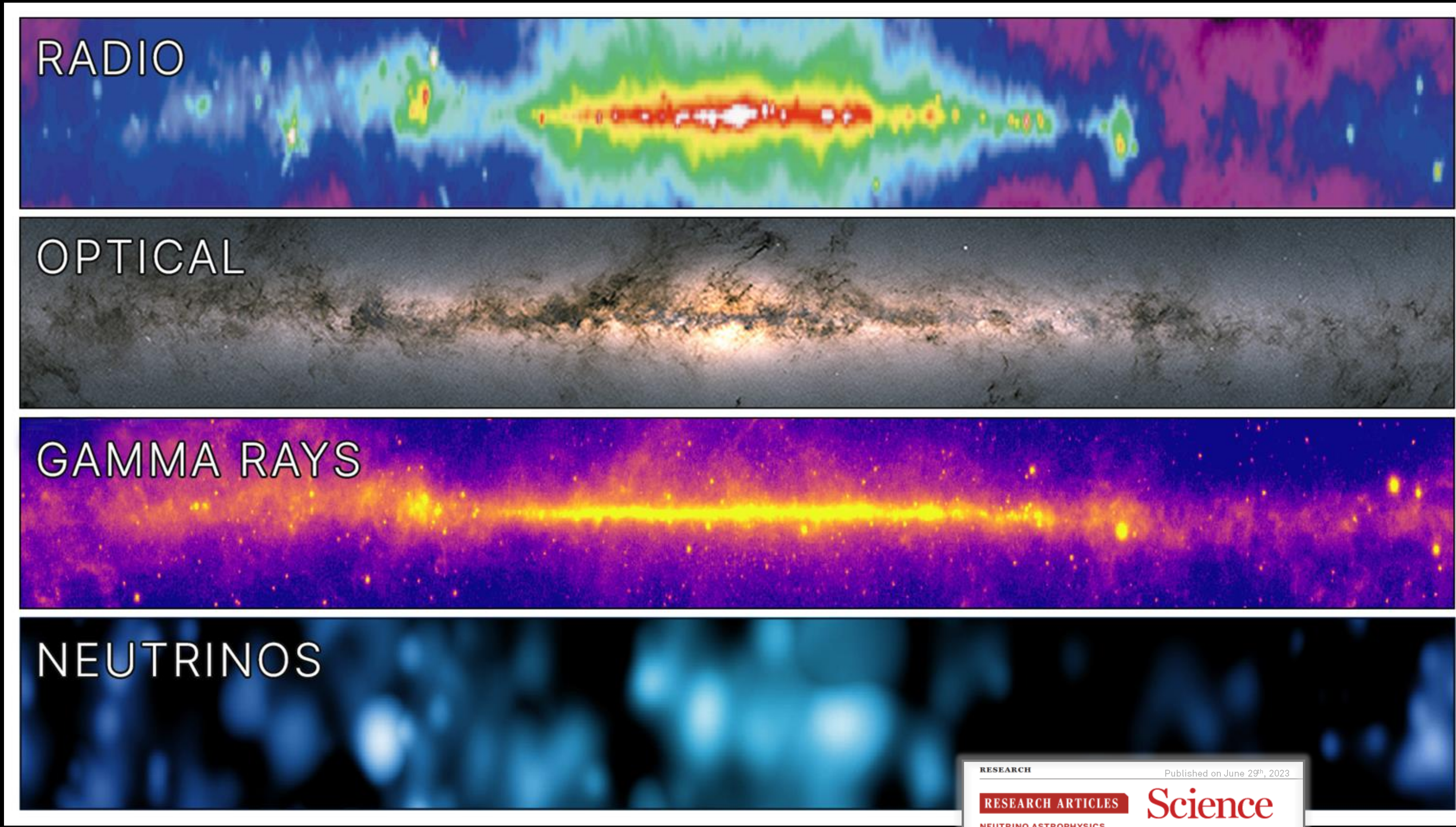


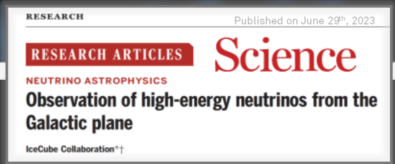
Credit: Tobias Hell, https://numerical-analysis.uibk.ac.at/images/User-Data/Tobias-Hell/StochastikHell.pdf

Mirco Hünnefeld
for the IceCube Collaboration
mhuennefeld@icecube.wisc.edu

ML for High-Energy Cosmic Particles
University of Delaware
January 28, 2025

Observation of high-energy neutrinos from the Galactic Plane




RESEARCH
Published on June 29th, 2023

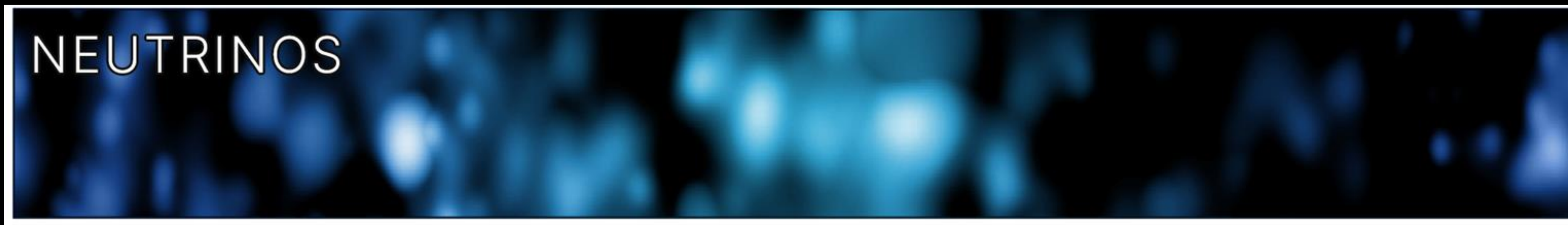
RESEARCH ARTICLES
Science

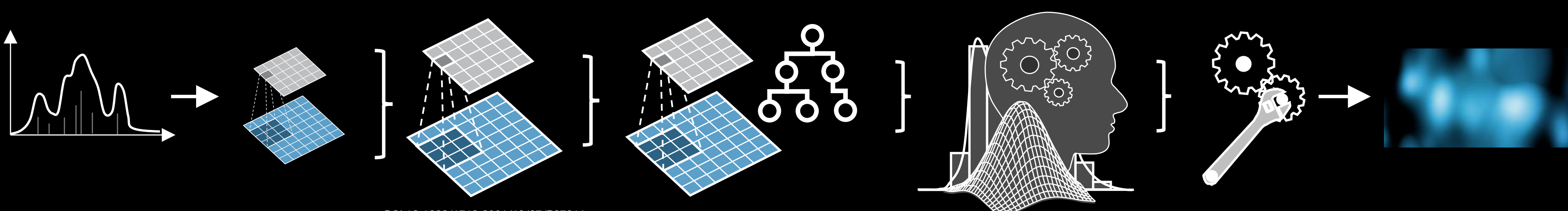
NEUTRINO ASTROPHYSICS

Observation of high-energy neutrinos from the Galactic plane

IceCube Collaboration*†

DOI: 10.1126/science.adc9818





DOI 10.1088/1748-0221/16/07/P07041

DOI: 10.22323/1.395.1065

~2700 / s

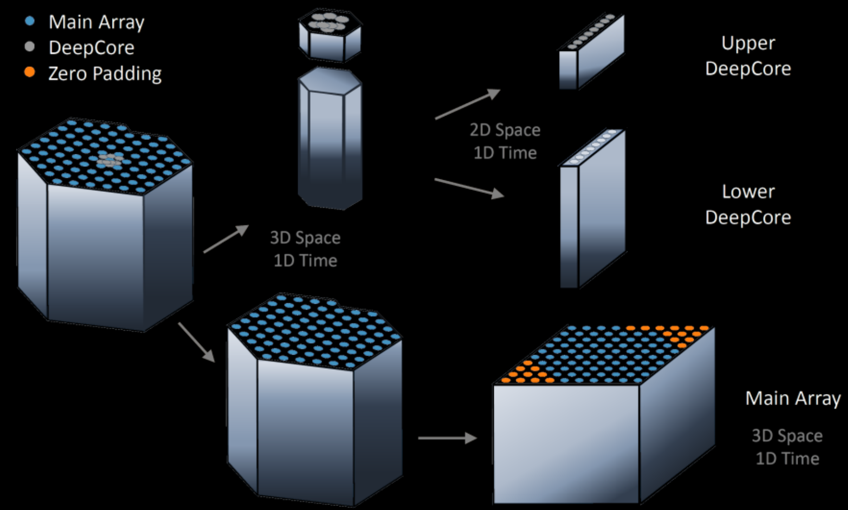
~2 / min

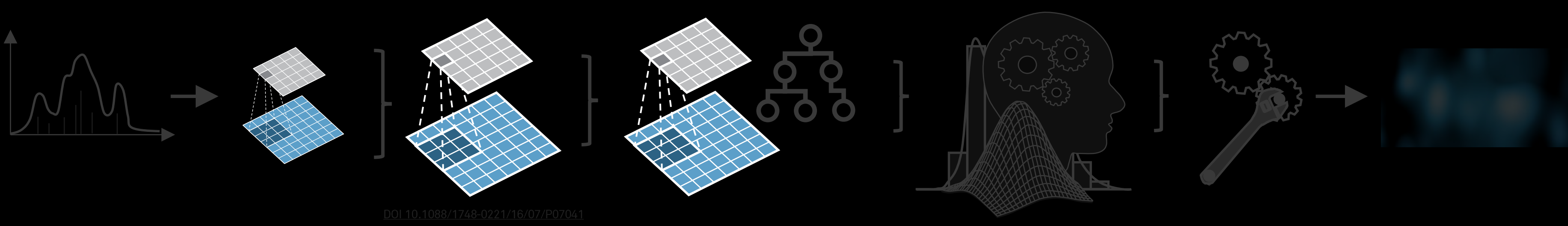
~17 / day

Convolutional Neural Networks

- Workhorse of event selection
- Classification & regression tasks
- Fast and reliable
- Hexagonal convolution kernels
- Uncertainty quantification

⟶ Exploit translational invariance

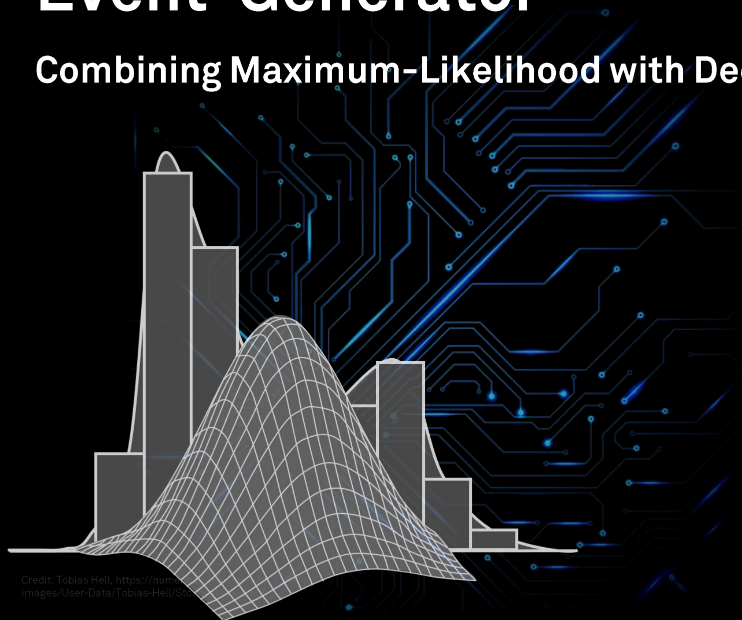




~2700 / s ~2 / min ~17 / day

Event-Generator

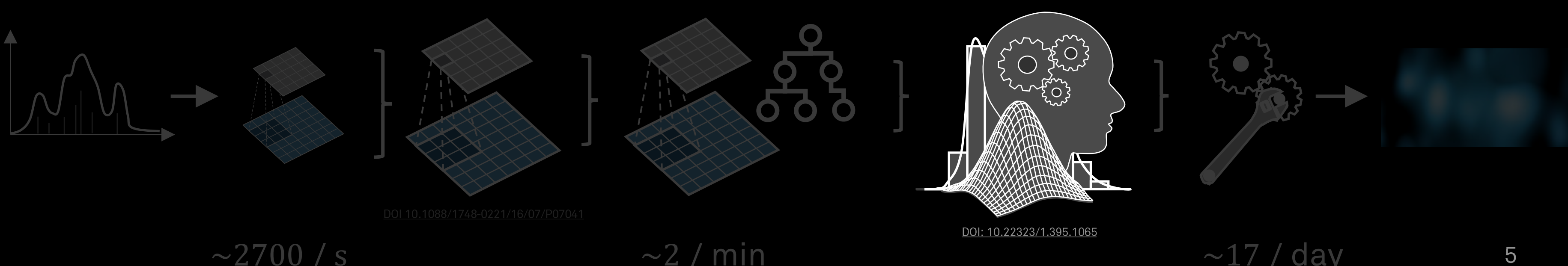
Combining Maximum-Likelihood with Deep Learning



DOI: 10.22323/1.395.1065

- Improved performance
- Robust
- Interpretable
- Uncertainty quantification
- Fast simulations

→ Leverage domain knowledge



Talk Outline

Importance of Domain Knowledge and Symmetries

Event Reconstruction in IceCube

- The IceCube Neutrino Observatory
- Data format and challenges
- Convolutional neural networks

Combining Maximum-Likelihood with Deep Learning

- Maximum-Likelihood Estimation
- Exploiting available symmetries and domain knowledge
- Interpretability and Generalization

Ongoing and future work

Conclusions

Talk Outline

Importance of Domain Knowledge and Symmetries

Event Reconstruction in IceCube

- The IceCube Neutrino Observatory
- Data format and challenges
- Convolutional neural networks

Combining Maximum-Likelihood with Deep Learning

- Maximum-Likelihood Estimation
- Exploiting available symmetries and domain knowledge
- Interpretability and Generalization

Ongoing and future work

Conclusions

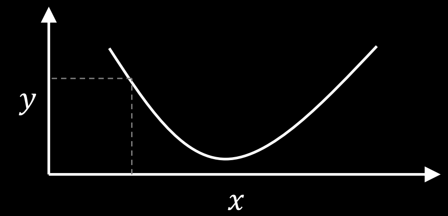
Credit: Tobias Hell, https://numerical-analysis.uibk.ac.at/images/User-Data/Tobias-Hell/StochastikHell.pdf

Deep Learning (DL) in a Nutshell

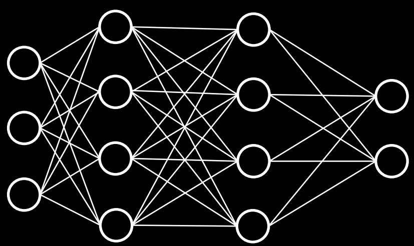
① DL performs a mapping from inputs to outputs

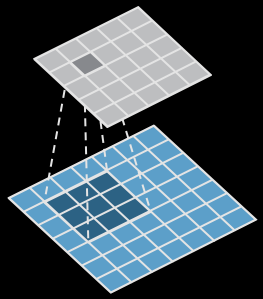
$$\boxed{f : I \rightarrow O}$$

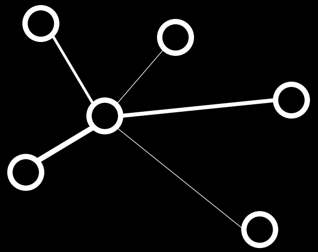
Input I	Output O
$x = 3$	$y = 12$
Can you explain deep learning in one sentence?	Deep learning is a subset of machine learning that uses artificial neural networks to model and understand complex patterns and relationships in data.



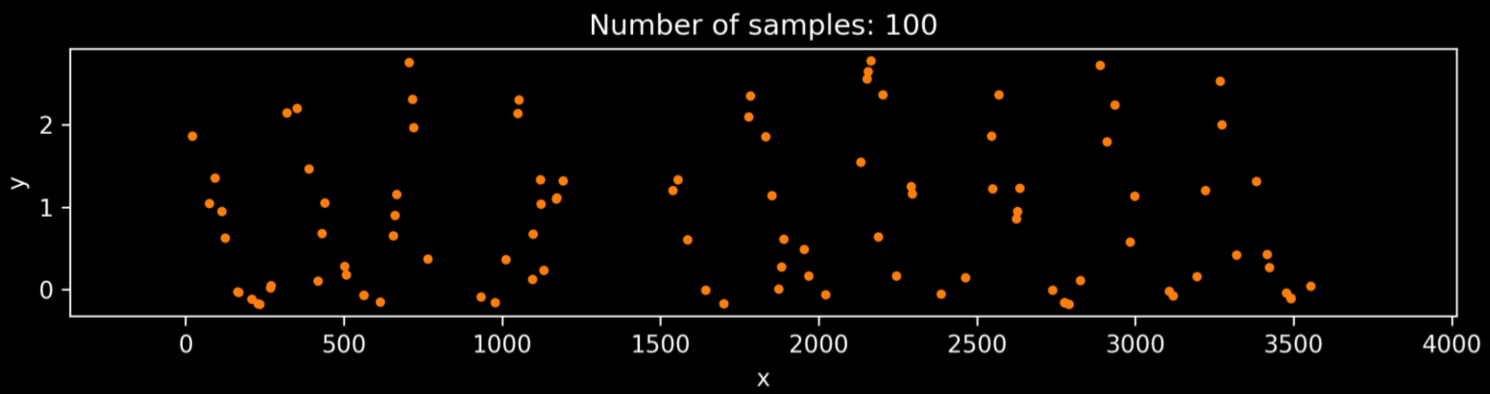
② Different architectures utilize different symmetries and domain knowledge



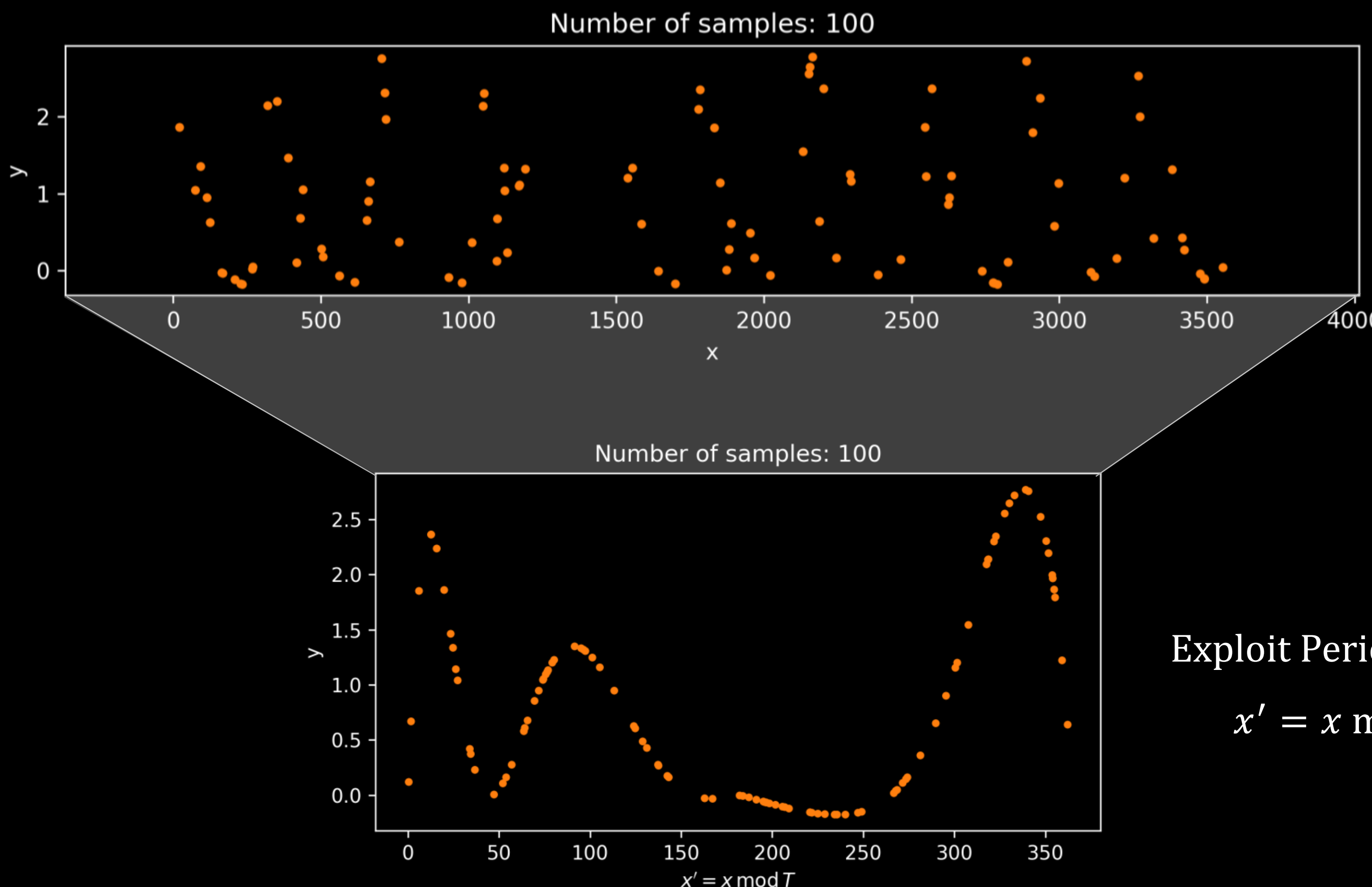




Utilizing Domain Knowledge



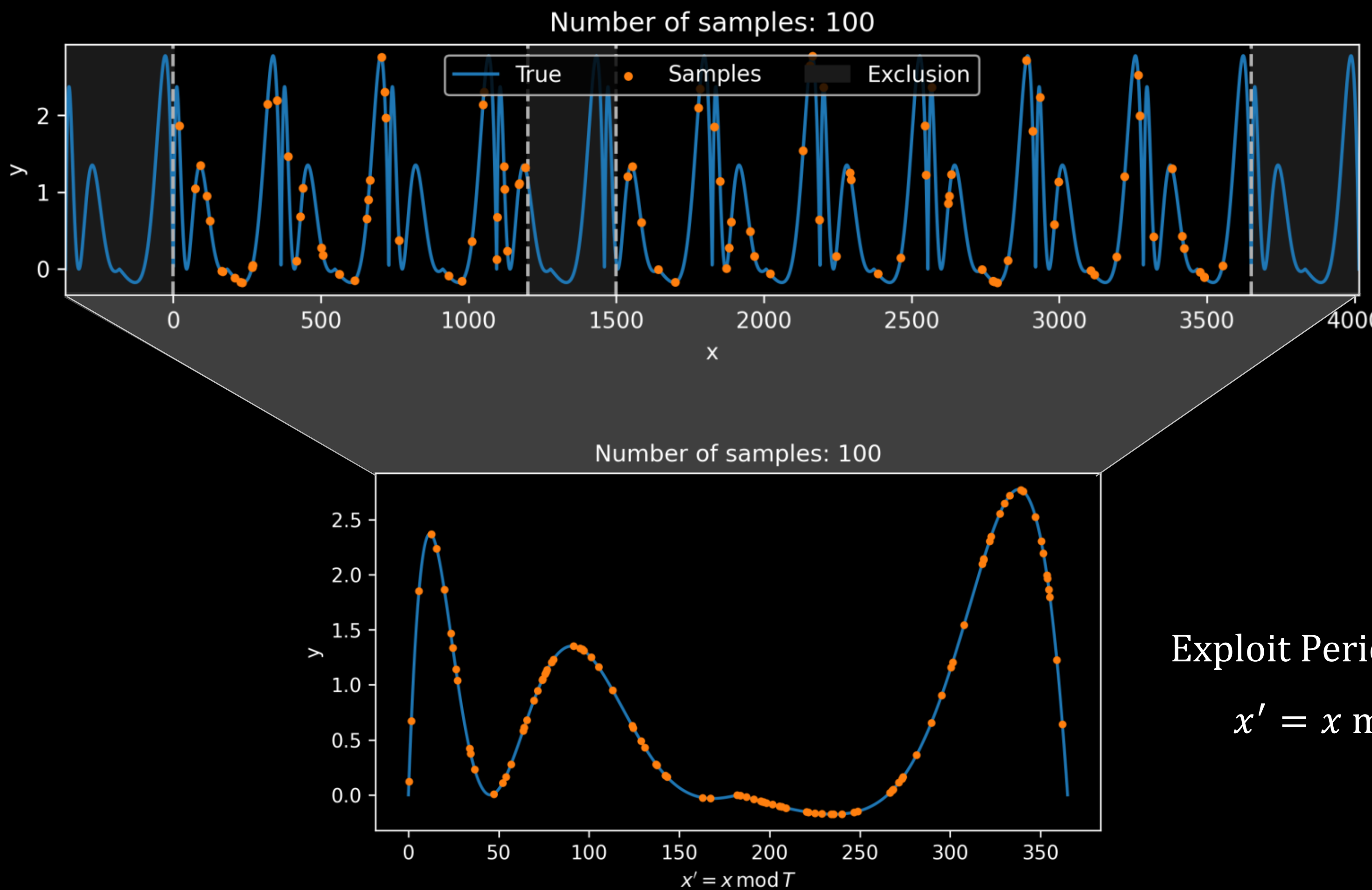
Utilizing Domain Knowledge



Exploit Periodicity:

$$x' = x \bmod T$$

Utilizing Domain Knowledge



Exploit Periodicity:

$$x' = x \bmod T$$

Talk Outline

Importance of Domain Knowledge and Symmetries

Event Reconstruction in IceCube

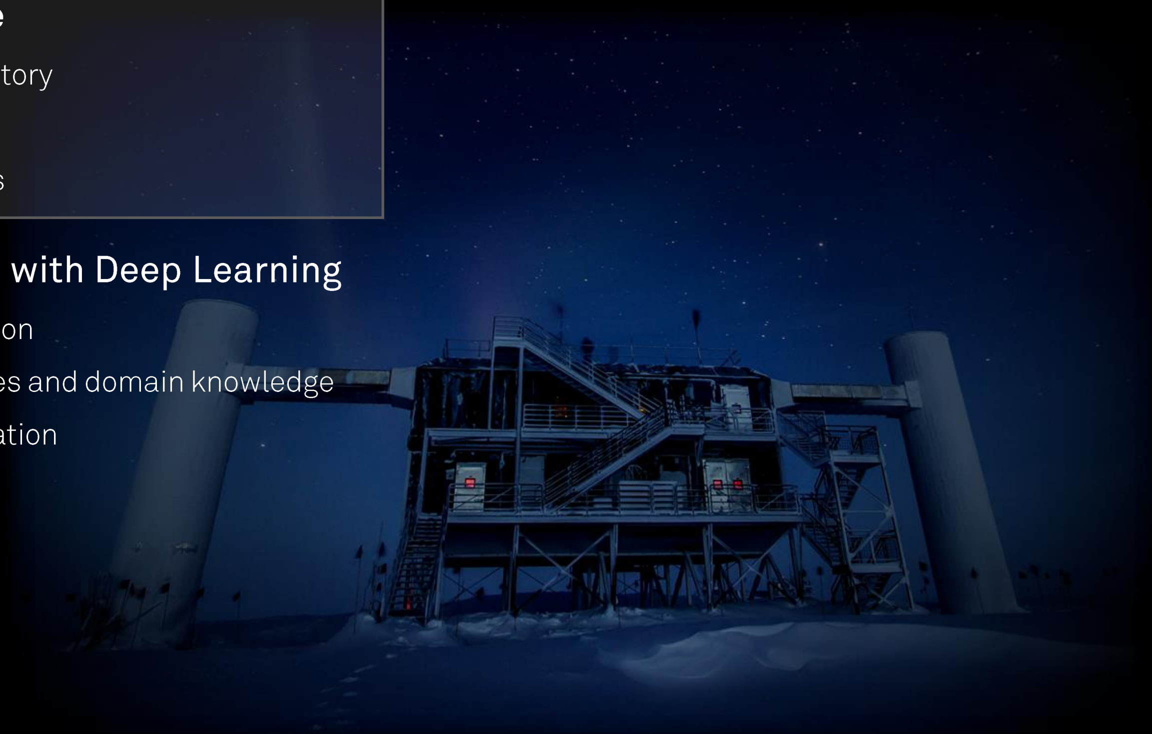
- The IceCube Neutrino Observatory
- Data format and challenges
- Convolutional neural networks

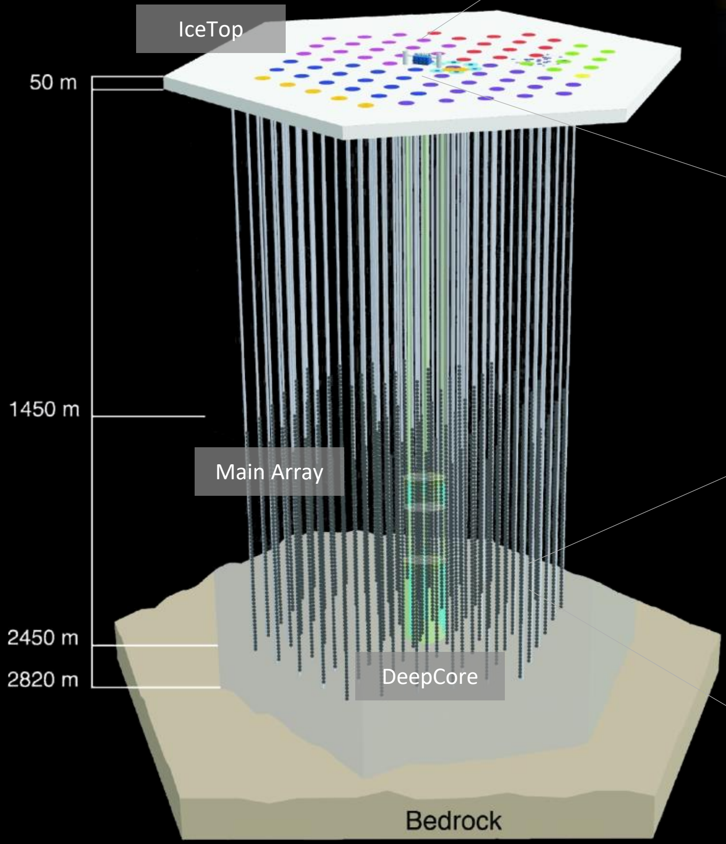
Combining Maximum-Likelihood with Deep Learning

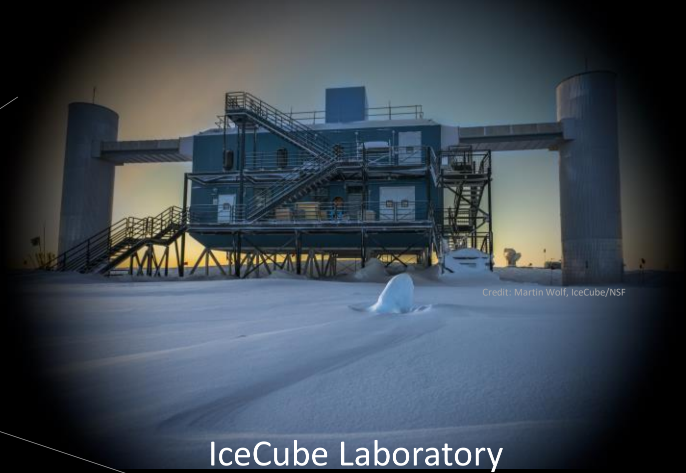
- Maximum-Likelihood Estimation
- Exploiting available symmetries and domain knowledge
- Interpretability and Generalization

Ongoing and future work

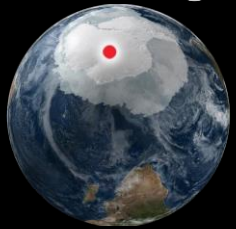
Conclusions



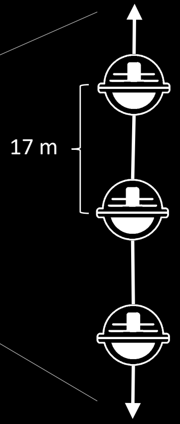


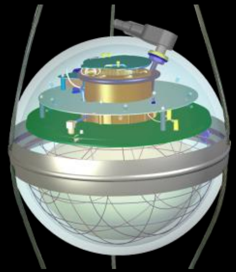


IceCube Laboratory



Amundsen-Scott South Pole Station, Antarctica

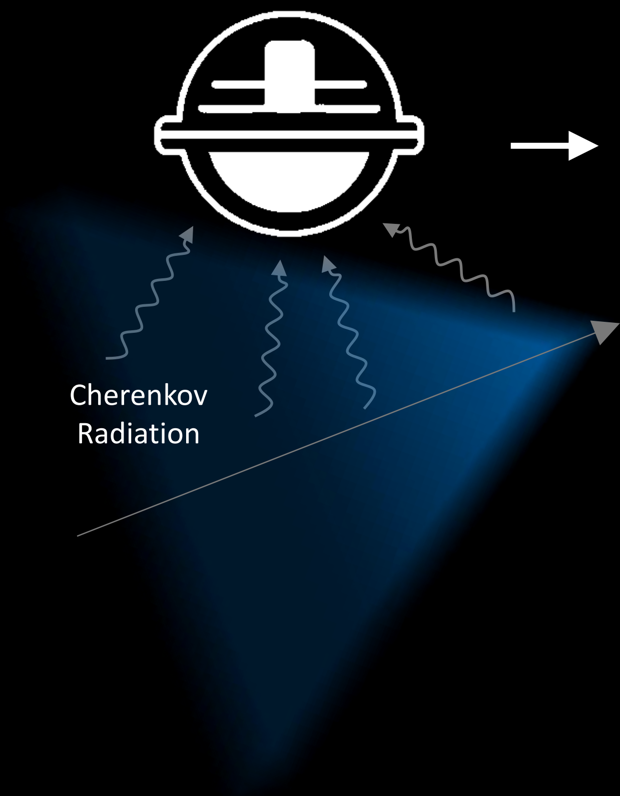




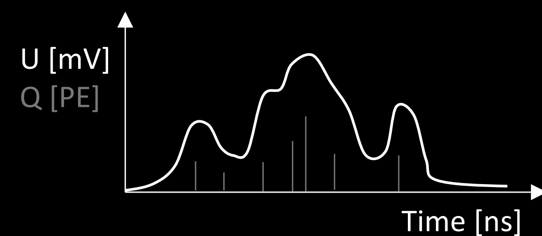
5160 Digital Optical Modules (DOMs)

86 Strings:
- 78 Main Array
- 8 DeepCore

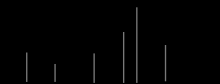
Detection Mechanism



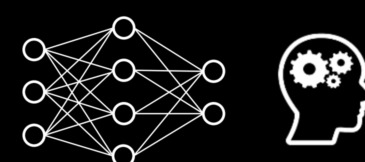
Data rate: ~1 TB/day



Pulse Series: (t_i, q_i)

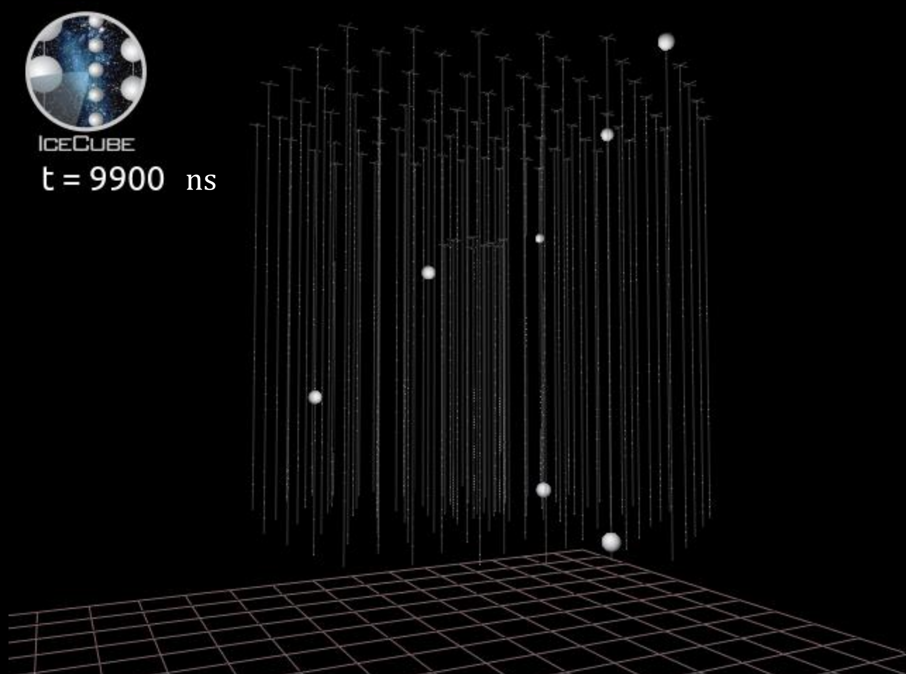


$$\mathcal{L}(\vec{x}|\vec{\theta}) = \prod_i p(x_i|\vec{\theta})$$

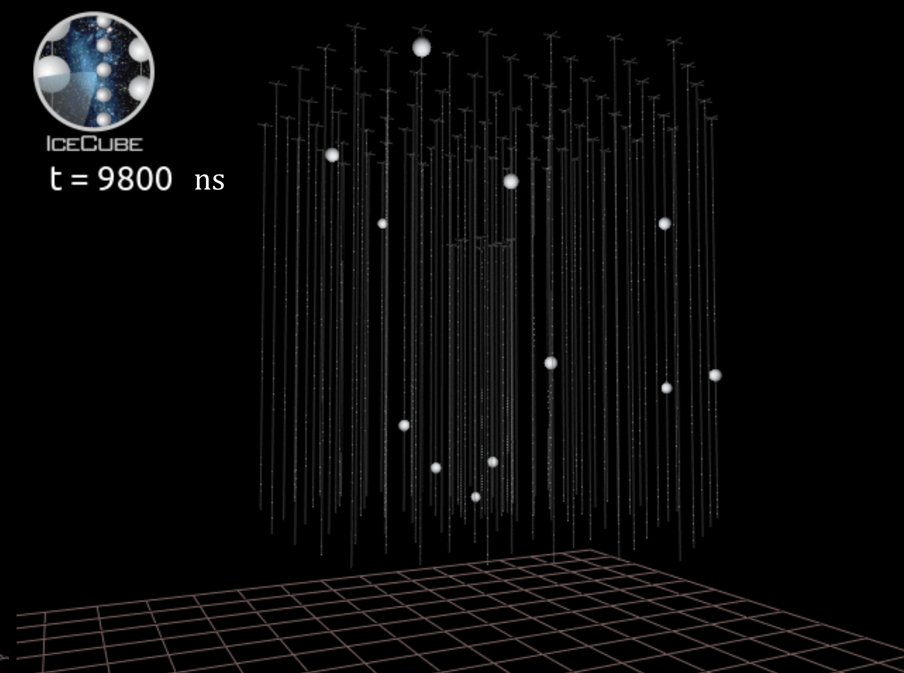


Event Topologies

Track Event



Cascade Event



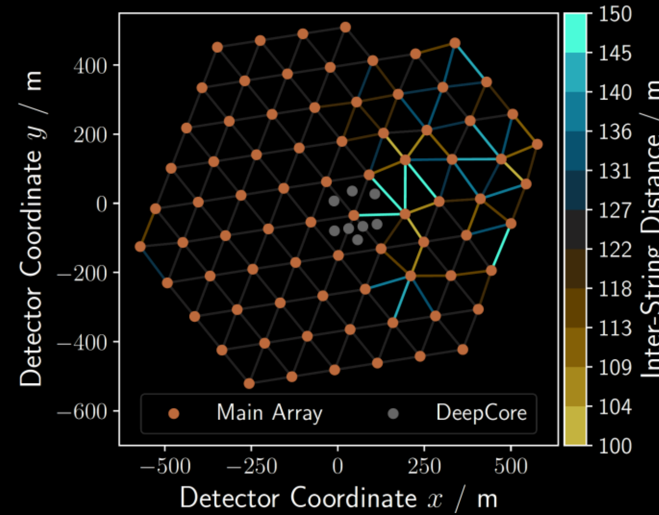
① Event Classification
Topology? Neutrino?

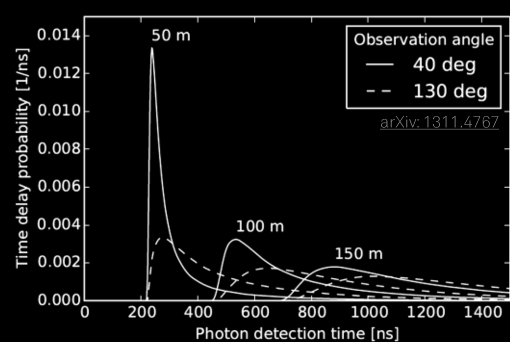
② Event Reconstruction
Direction? Energy? Vertex?

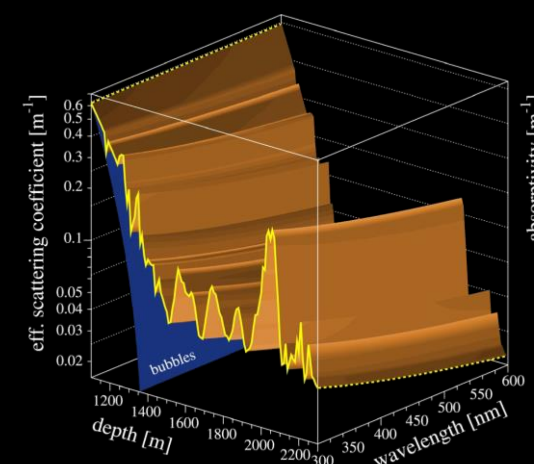
Ratio of signal to background: $\sim$ 1: 100 million

What do we know about our Data?

- Detector geometry:
  - 3 detector parts: main array, upper & lower DeepCore
  - Deviations from symmetric detector grid
- Underlying physics of neutrino interaction are invariant under translation and rotation
- Inhomogeneous photon propagation due to dust impurities and crystal structure of ice
- Light yield scales linearly with deposited energy
- General shape of photon arrival time PDF
- Photons (and in good approximation: the measured pulses) are independent of each other





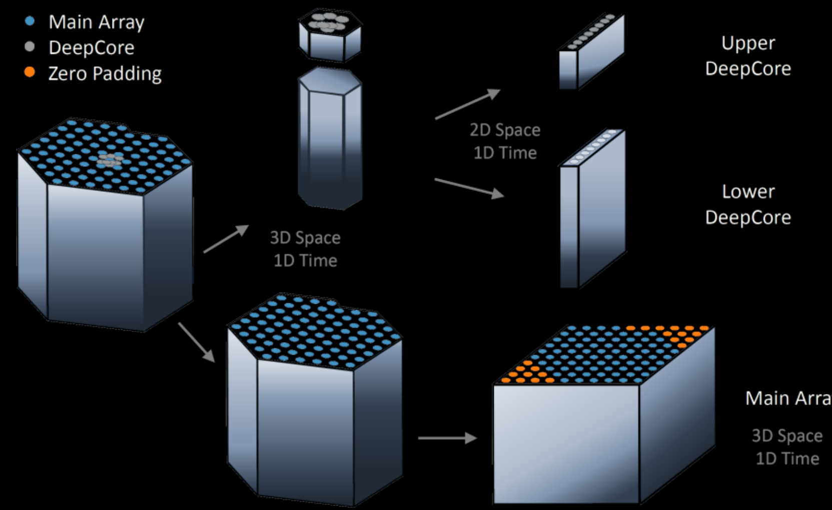


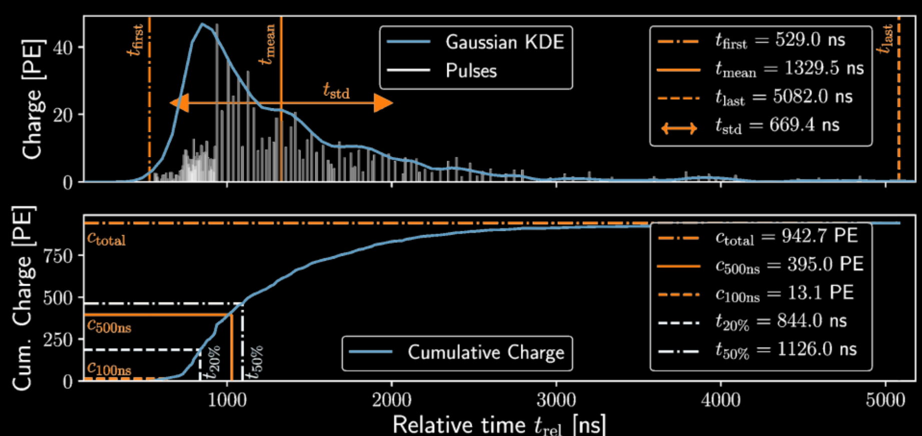
Source: 10.1029/2005JD006687 [4]

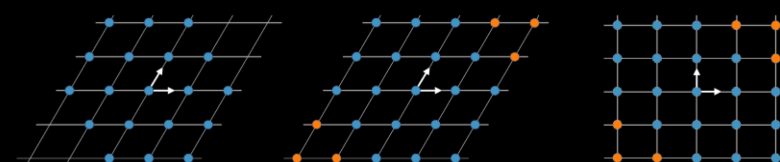
Convolutional Neural Networks

- Workhorse of event selection
- Classification & regression tasks
- Fast and reliable
- Hexagonal convolution kernels
- Uncertainty quantification

⟶ Exploit translational invariance







Hexagonal Convolution Kernels

Talk Outline

Importance of Domain Knowledge and Symmetries

Event Reconstruction in IceCube

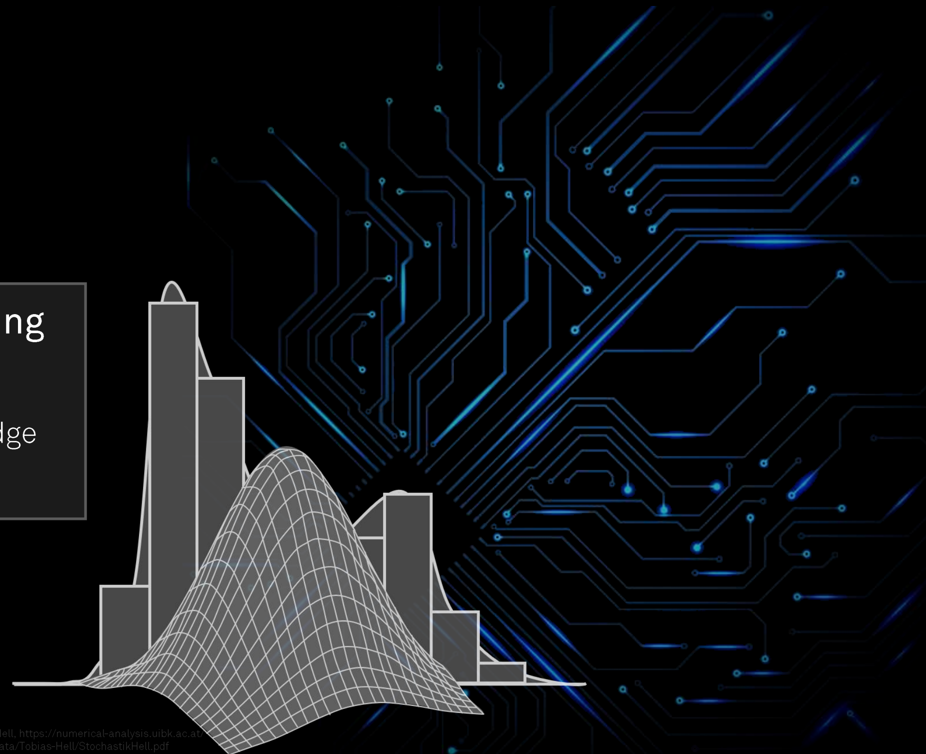
- The IceCube Neutrino Observatory
- Data format and challenges
- Convolutional neural networks

Combining Maximum-Likelihood with Deep Learning

- Maximum-Likelihood Estimation
- Exploiting available symmetries and domain knowledge
- Interpretability and Generalization

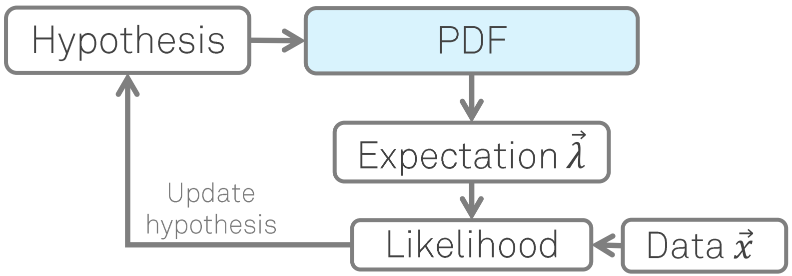
Ongoing and future work

Conclusions



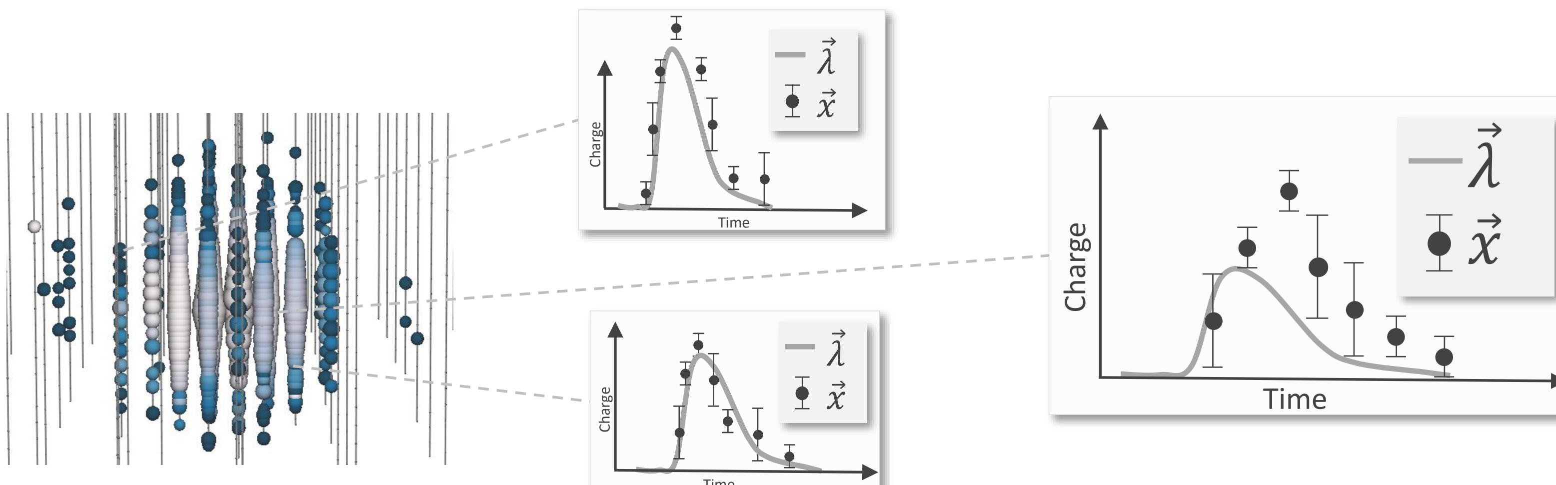
Credit: Tobias Hell, https://numerical-analysis.uibk.ac.at/images/User-Data/Tobias-Hell/StochastikHell.pdf

Event-Generator: Combining Maximum-Likelihood with DL

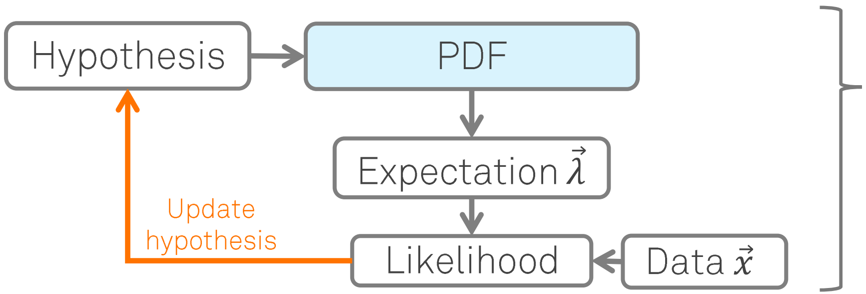


Hybrid reconstruction method:

- Combines maximum-likelihood estimation with deep learning
- Modeling of high-dimensional PDFs via generative model
- Exploits available information and symmetries
- **Robust and interpretable Deep Learning**

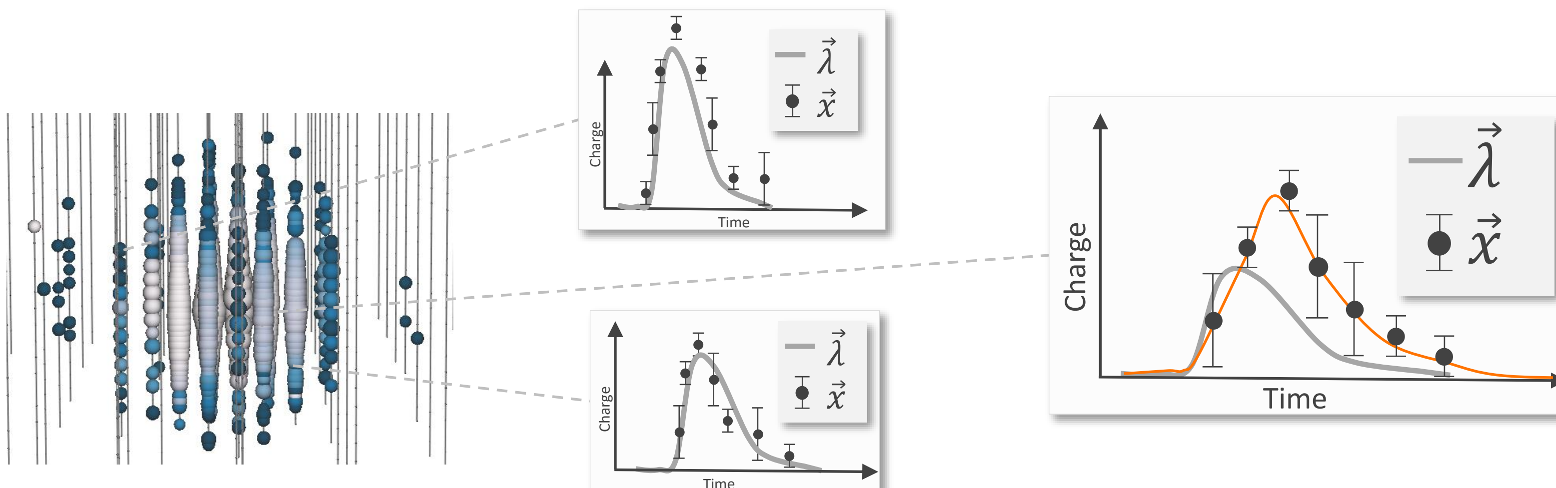


Event-Generator: Combining Maximum-Likelihood with DL

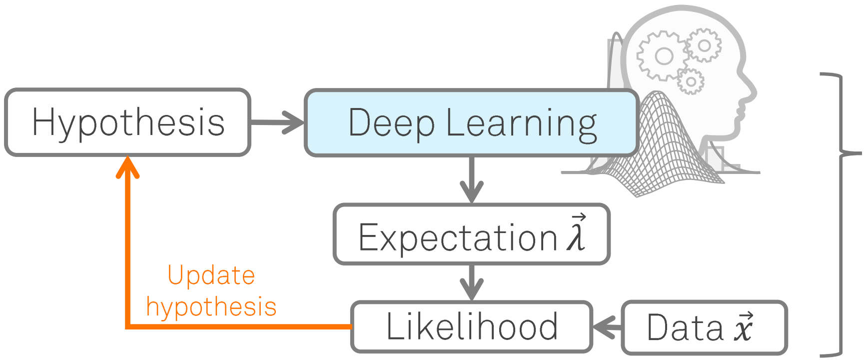


Hybrid reconstruction method:

- Combines maximum-likelihood estimation with deep learning
- Modeling of high-dimensional PDFs via generative model
- Exploits available information and symmetries
- **Robust and interpretable Deep Learning**

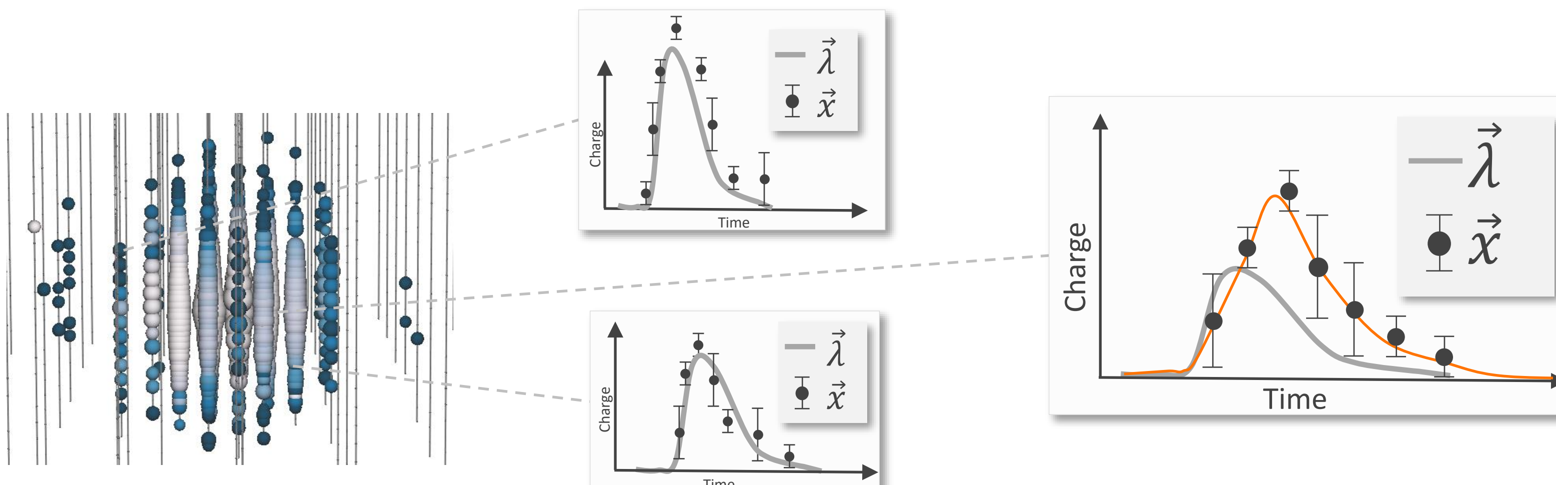


Event-Generator: Combining Maximum-Likelihood with DL



Hybrid reconstruction method:

- Combines maximum-likelihood estimation with deep learning
- Modeling of high-dimensional PDFs via generative model
- Exploits available information and symmetries
- **Robust and interpretable Deep Learning**



Event-Generator: Architecture

Generator NN learns mapping:

$$f : \vec{\theta} \rightarrow \vec{\lambda},\ \vec{p}(t_i|\vec{\theta})$$

$\vec{\theta}$: source parameters
$\vec{\lambda}$: expected DOM charge
$\vec{p}(t_i|\vec{\theta})$: pulse arrival PDFs

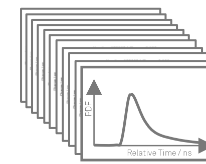
Example: Cascade Hypothesis

$$\vec{\theta} = (x, y, z, \varphi, \theta, E, t)$$

Neural Network that maps input [?, 7] to output [?, 5160, n] with number of mixture model components n

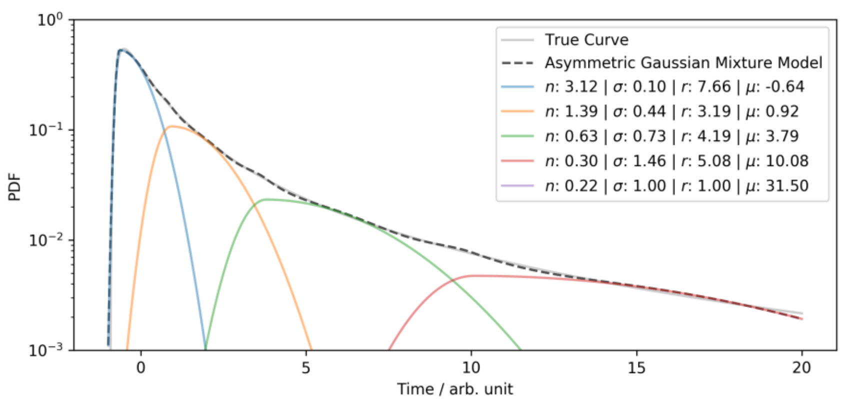
Generator

5160 PDFs for each of the DOMs



~200k output variables

Parameterization of pulse arrival time PDF:

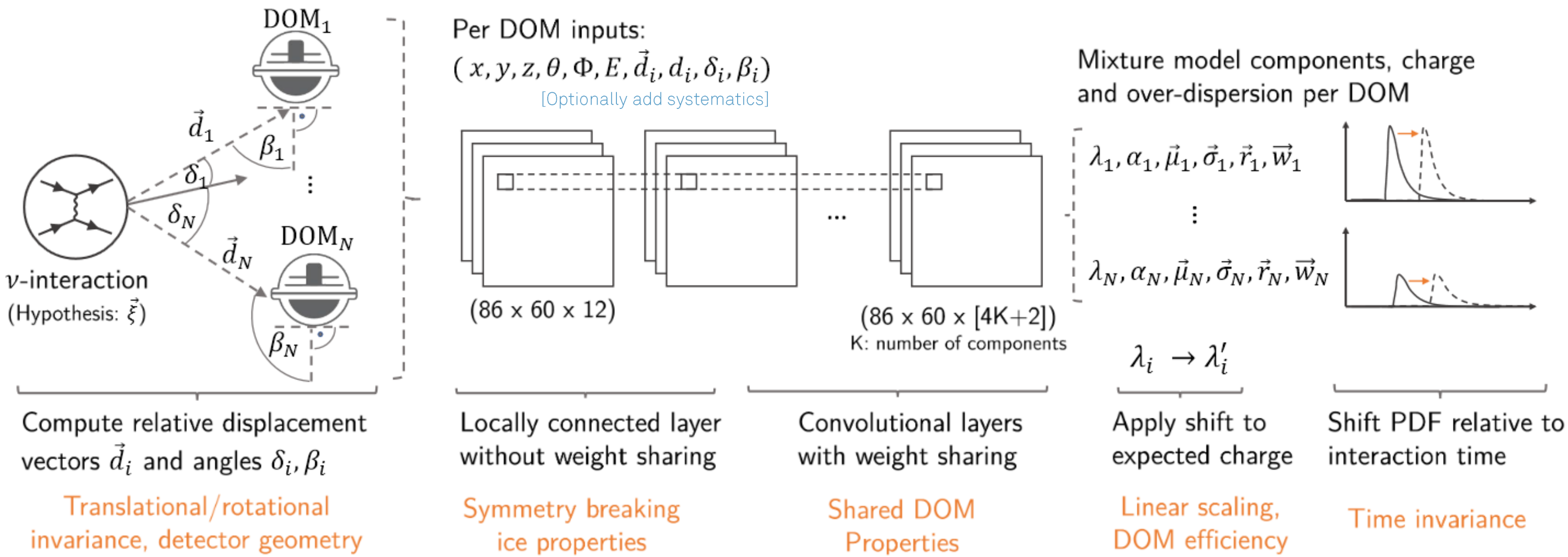


$$g(x|\mu,\sigma,r) = \begin{cases} N \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), & x \leq \mu \\ N \cdot \exp\left(-\frac{(x-\mu)^2}{2(\sigma r)^2}\right), & \text{otherwise} \end{cases}$$

$$N = \frac{2}{\sqrt{(2\pi)} \cdot \sigma (r+1)}$$

DOI: 10.1007/3-540-70659-3_42

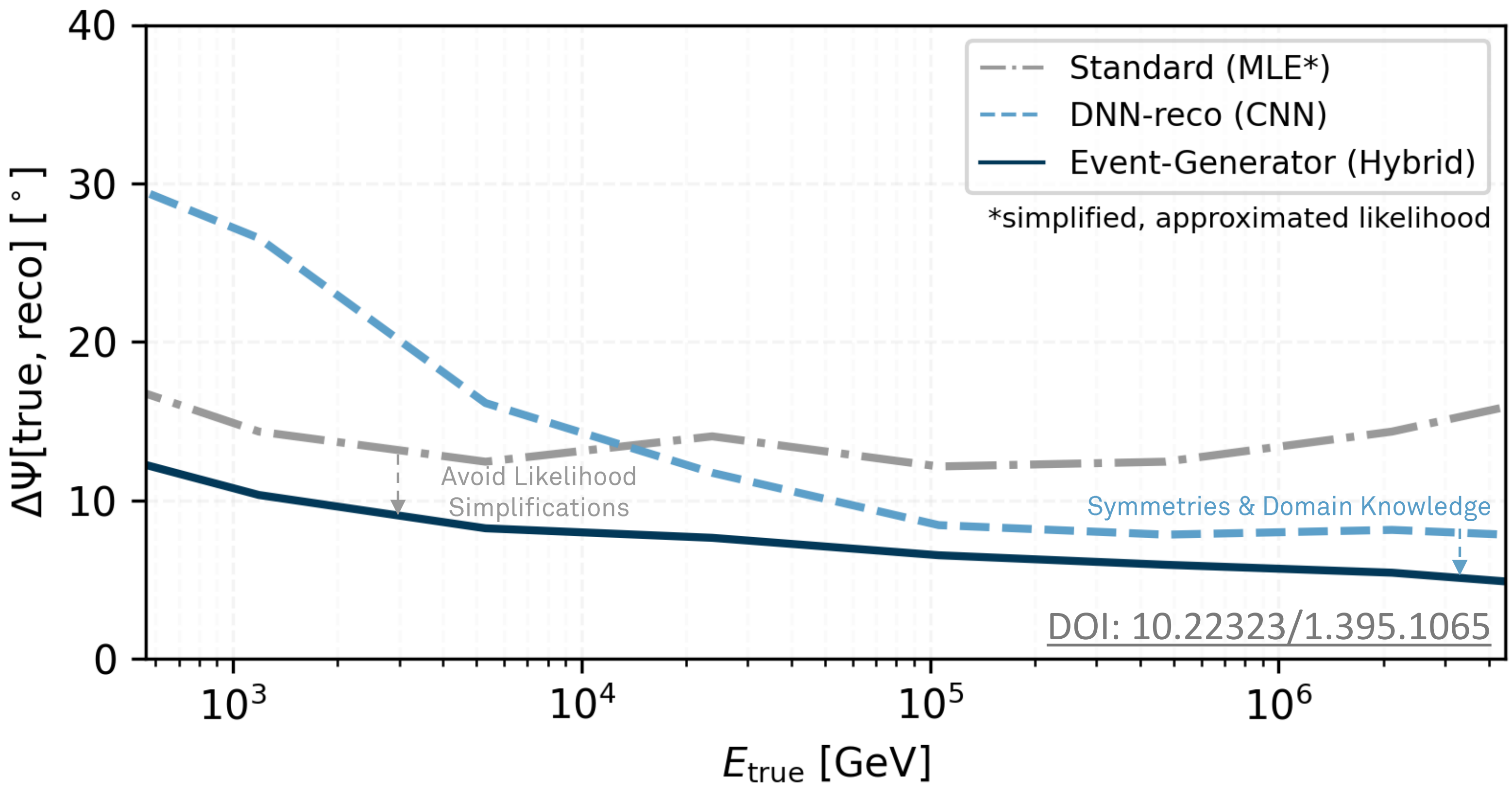
Event-Generator: Example Architecture for Cascades



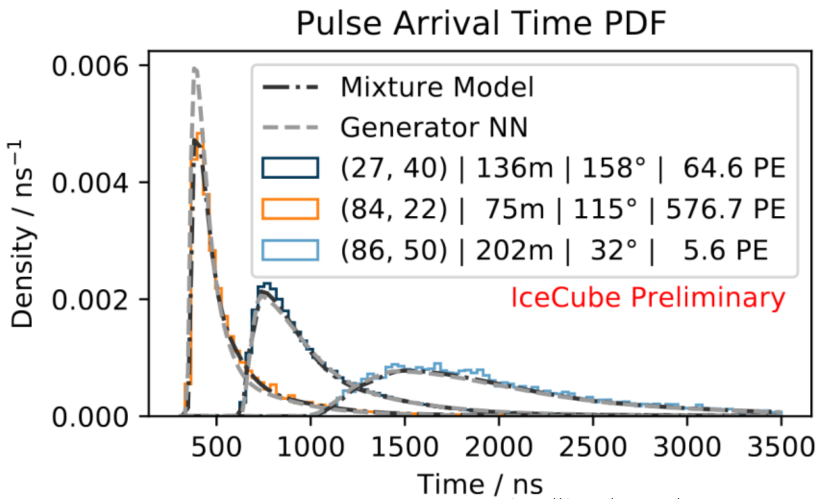
DOI: 10.22323/1.395.1065

- Can decouple physics and detector effects
- Easier to include information in forward direction when not convolved with detector response yet
- We know how to do this – we simulate the data!

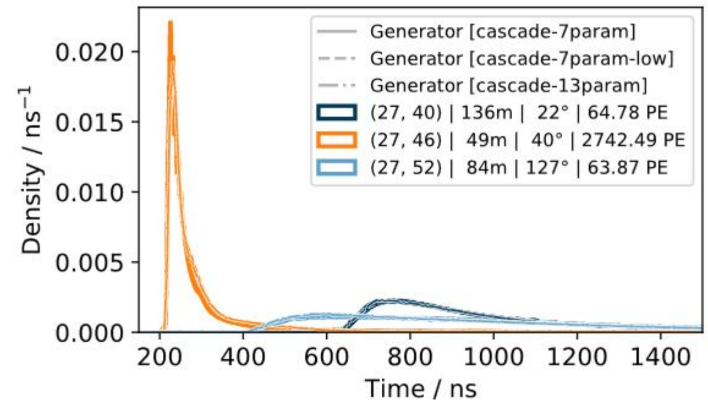
Event-Generator: Improved Performance

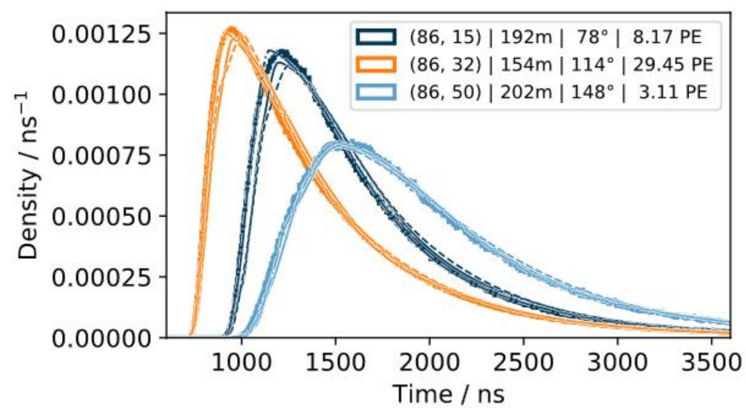


Event-Generator: Interpretability and Generalization

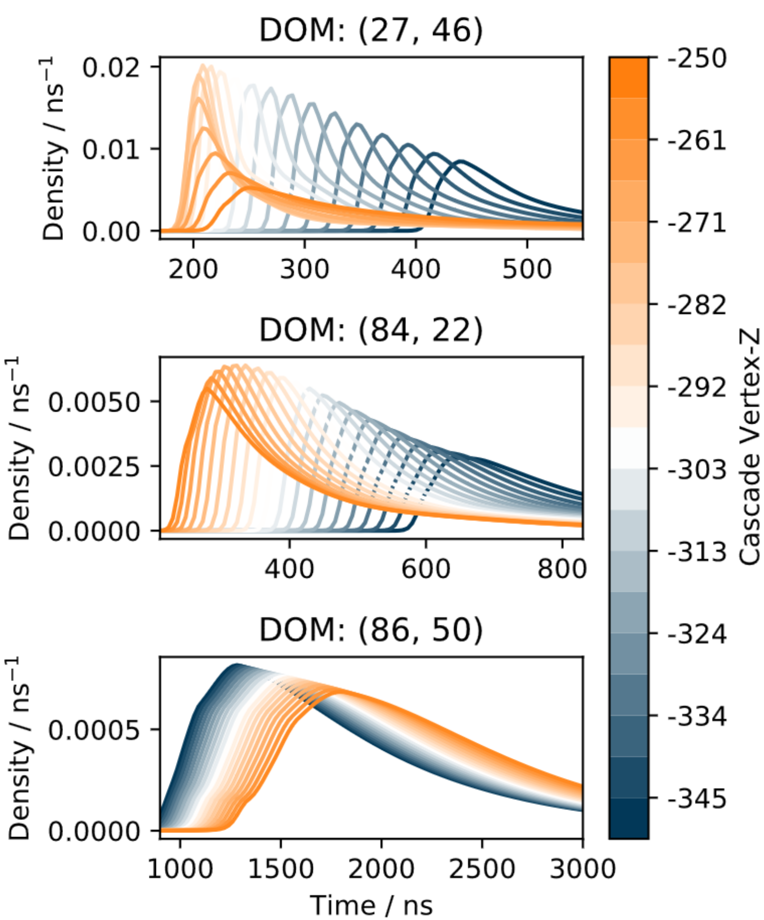


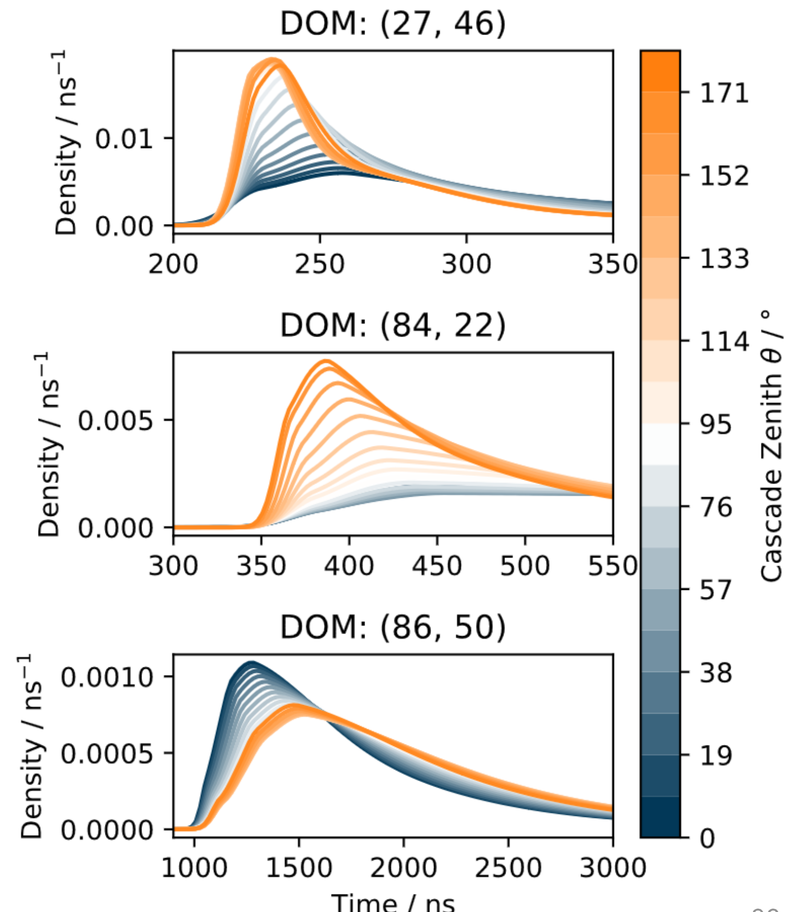
DOI: https://doi.org/10.22323/1.395.1065



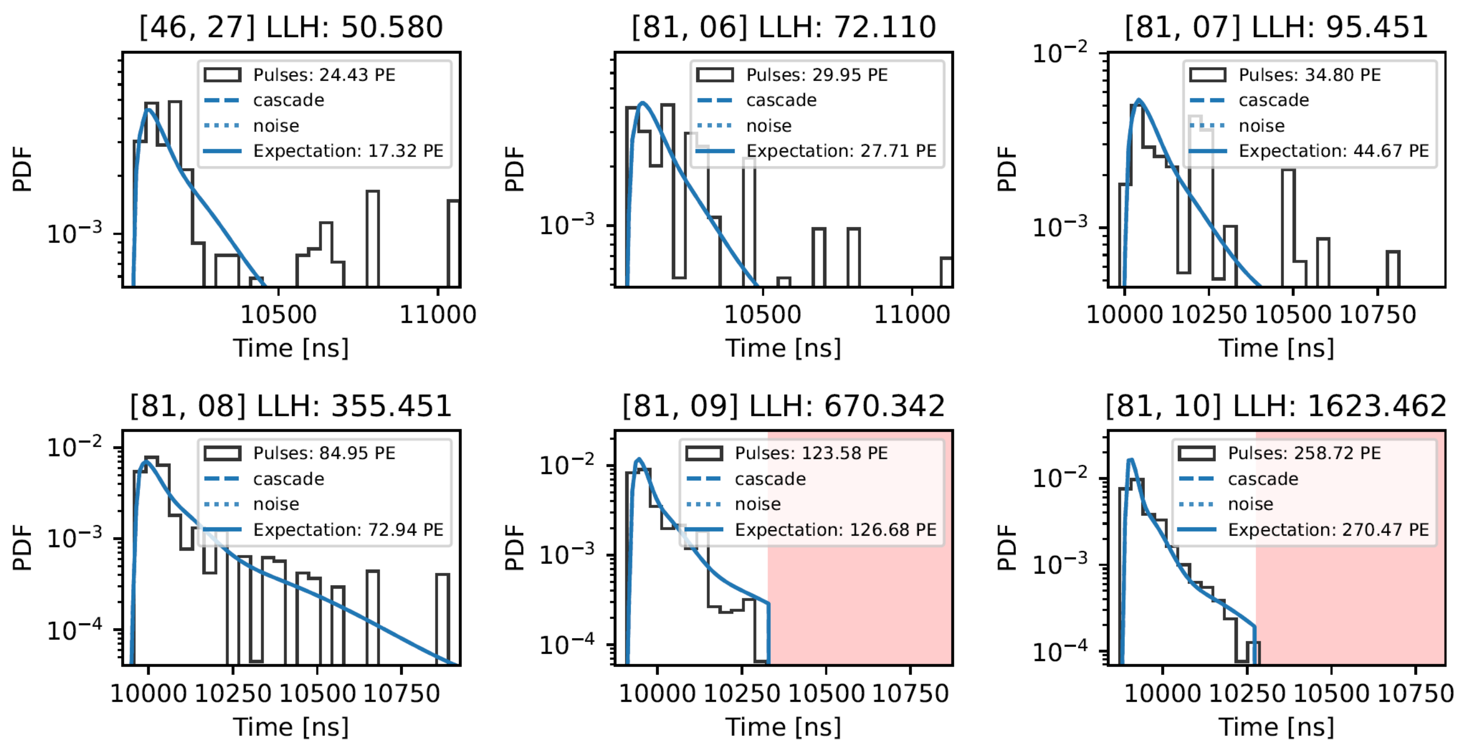


Event-Generator: Interpretability and Generalization

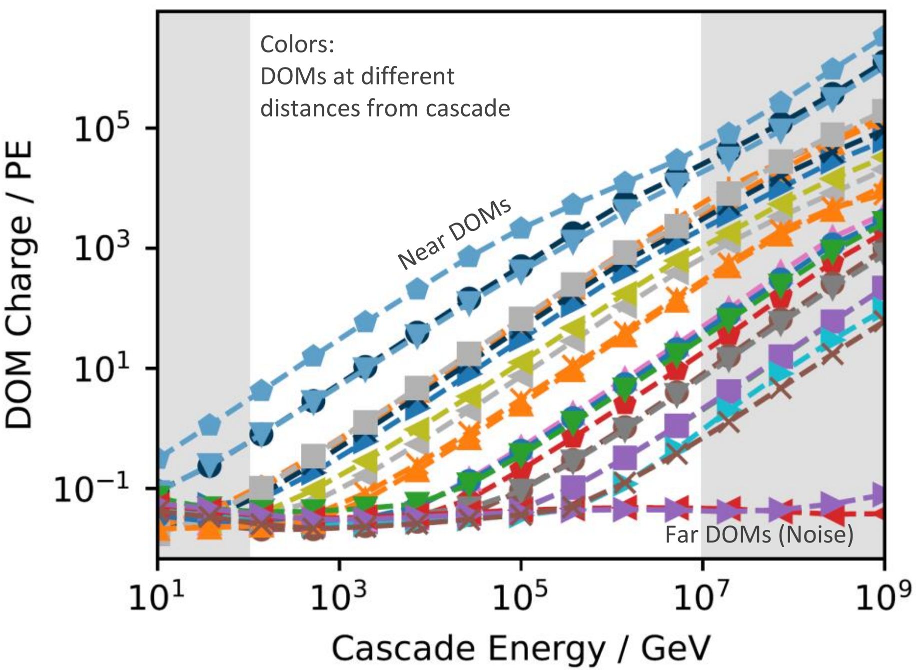


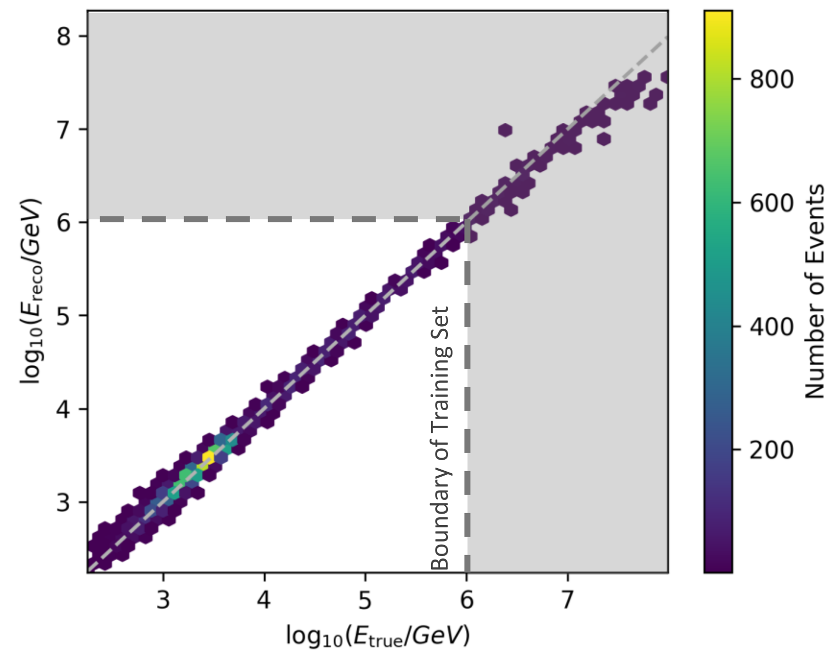


Event-Generator: Interpretability and Generalization



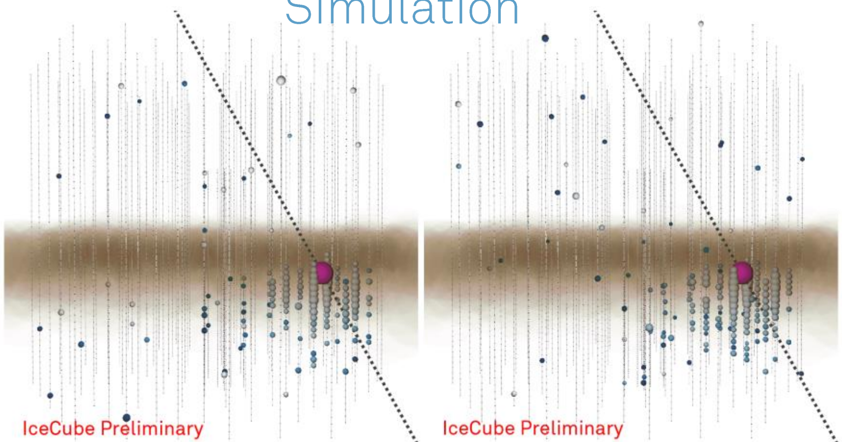
Event-Generator: extrapolating beyond the training data





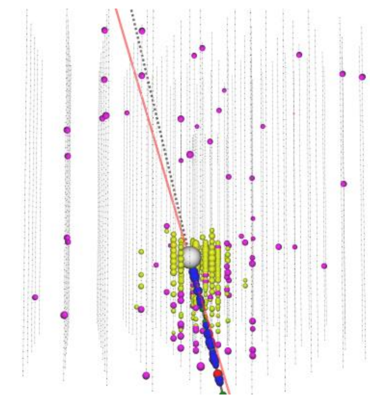
Event-Generator: Additional Applications

Simulation

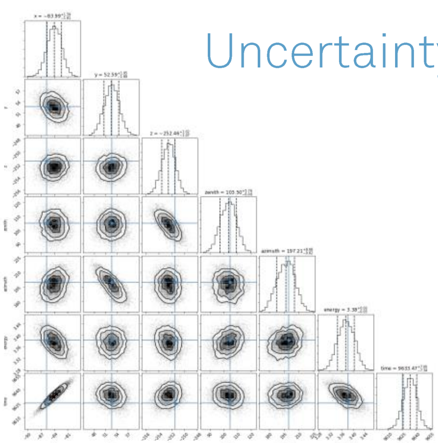


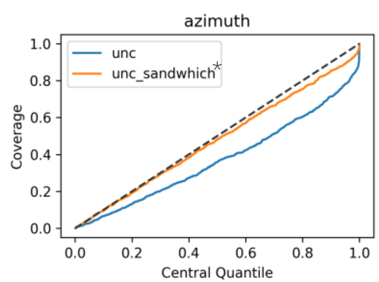
DOI: https://doi.org/10.22323/1.395.1065

Goodness-of-fit



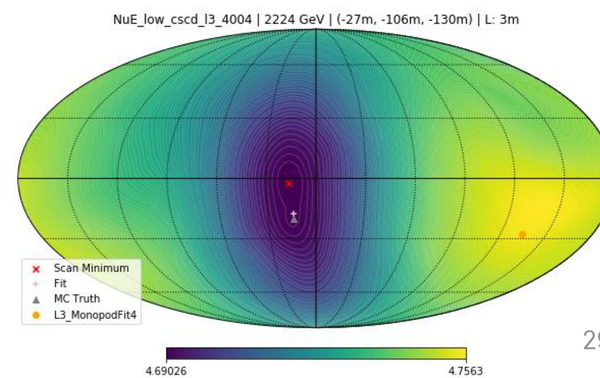
Uncertainty Estimation





*http://www.stat.umn.edu/geyer/5601/notes/sand.pdf

Likelihood Scans



Talk Outline

Importance of Domain Knowledge and Symmetries

Event Reconstruction in IceCube

- The IceCube Neutrino Observatory
- Data format and challenges
- Convolutional neural networks

Combining Maximum-Likelihood with Deep Learning

- Maximum-Likelihood Estimation
- Exploiting available symmetries and domain knowledge
- Interpretability and Generalization

Ongoing and future work

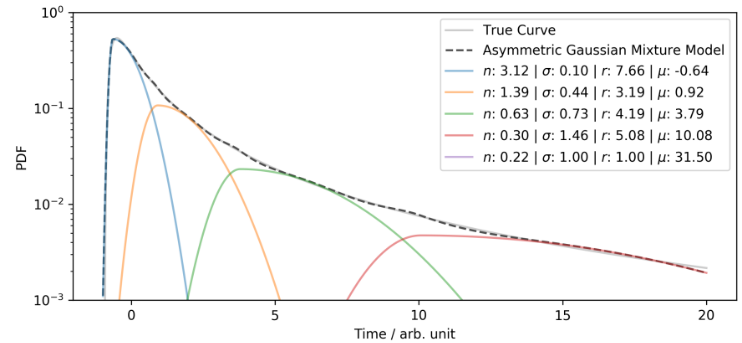
Conclusions

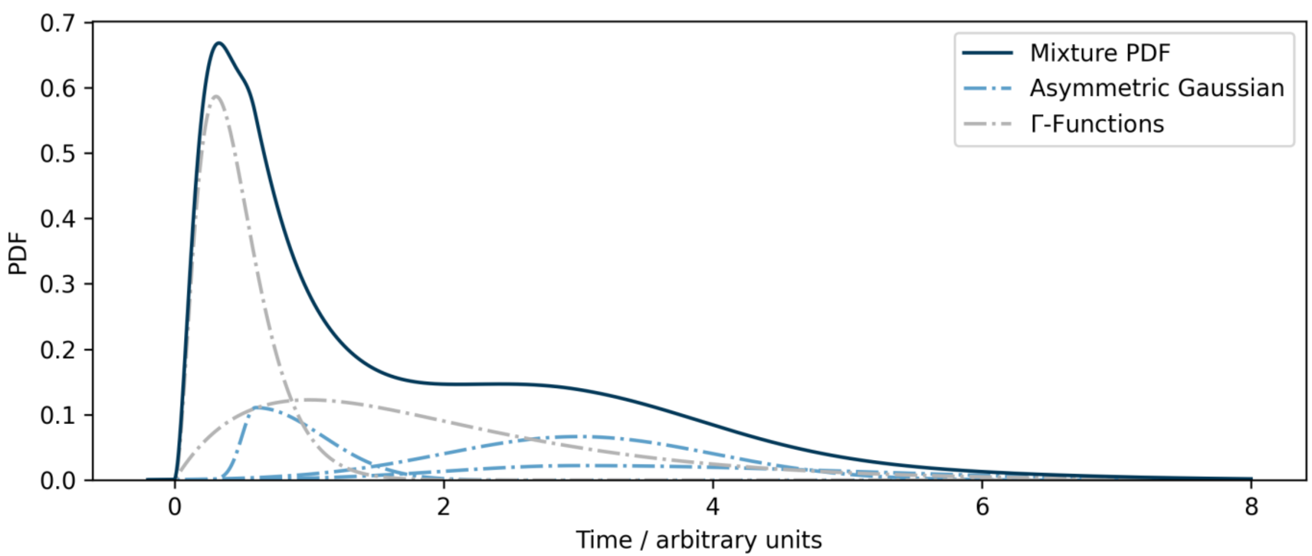
Event-Generator: more flexible time PDFs

Arrival time PDFs:

- Previously: mixture model of asymmetric Gaussians
- Now: mixture model of any basis PDF
- Option to include learnable time PDF parameterizations
  - Example: normalizing flow

→ More accurate modelling of arrival time PDFs





Event-Generator: nested event hypotheses

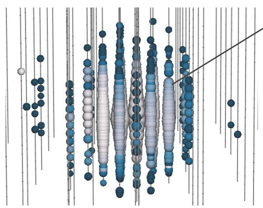
Source Class

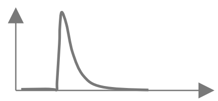
Defines an arbitrary light source

$$f: \vec{\theta} \rightarrow \vec{\lambda},\ \vec{p}(t_i|\vec{\theta})$$

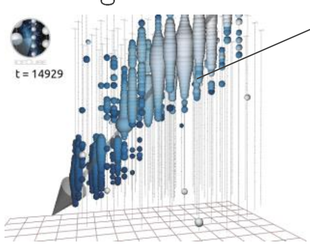
$\vec{\theta}$: source parameters
$\vec{\lambda}$: expected DOM charge
$\vec{p}(t_i|\vec{\theta})$: pulse arrival PDFs

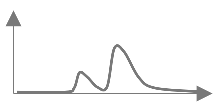
Cascade





Track segment/Muon





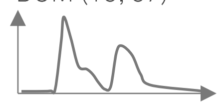
Double-bang

Flasher

...

MultiSource Class

Arbitrary collection of light sources defined by (nested) Source and/or MultiSource objects



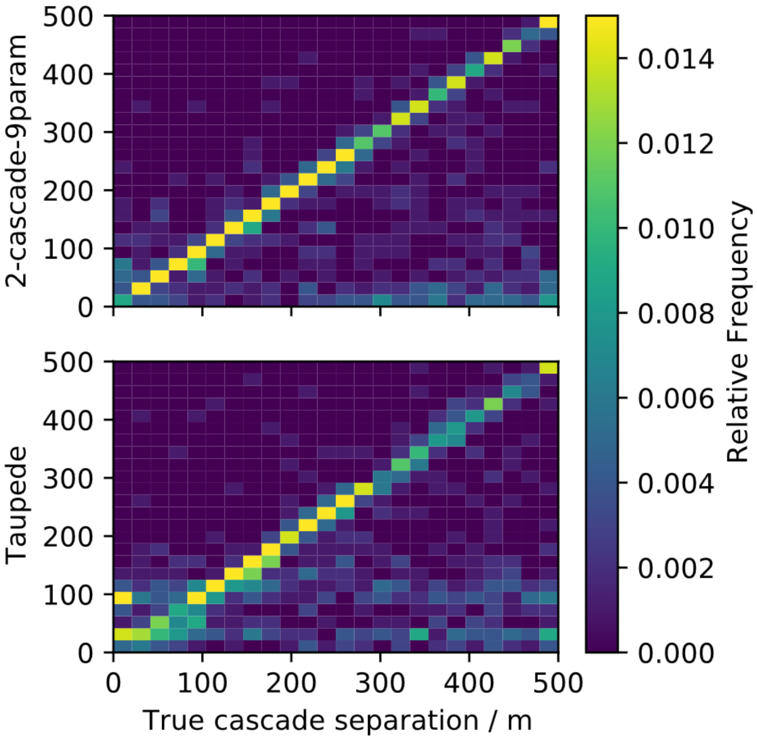
Examples:
- Muons (cascades + track segments)
- Muon bundle (muons)
- Coincident events

Event-Generator:

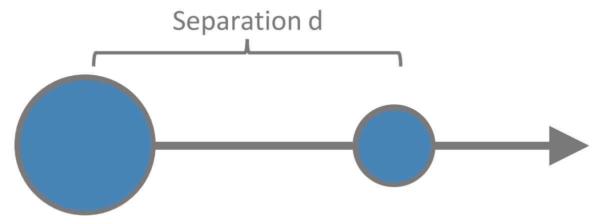
- Framework to compose arbitrary event hypotheses
- Framework ensures compatibility of objects, version control, and reproducibility
- Provides tools to reconstruct, perform scans, estimate uncertainties, simulate events, interpret and investigate results, create and train new light source models

Event-Generator: Double Cascade Reconstruction

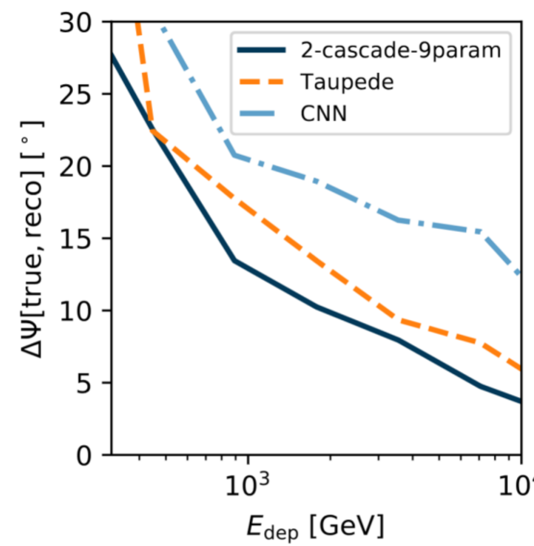
Double Cascade Separation



Double Cascade



Angular Resolution



Event-Generator: Track Reconstruction

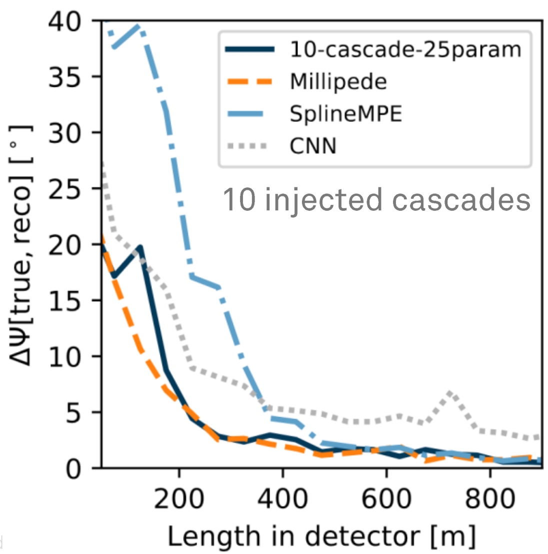
Track reconstruction:

- Hypothesis consists of:
  - Track vertex location and time, direction and energy/length
  - Energy depositions along track (e.g.: individual cascades)

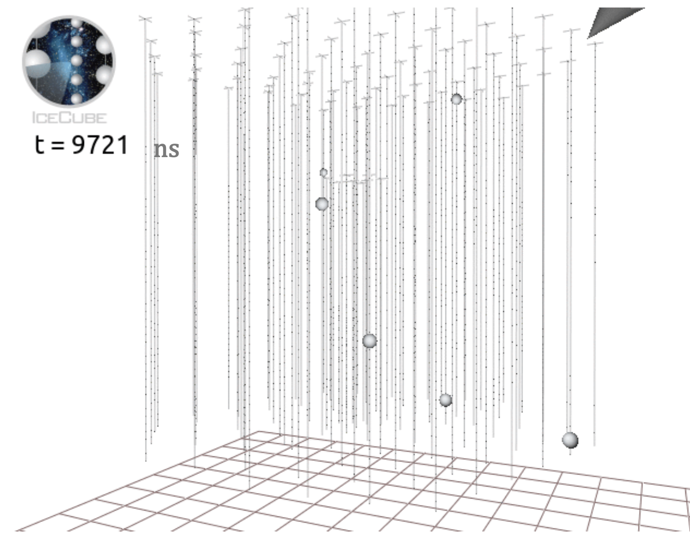
Challenges:

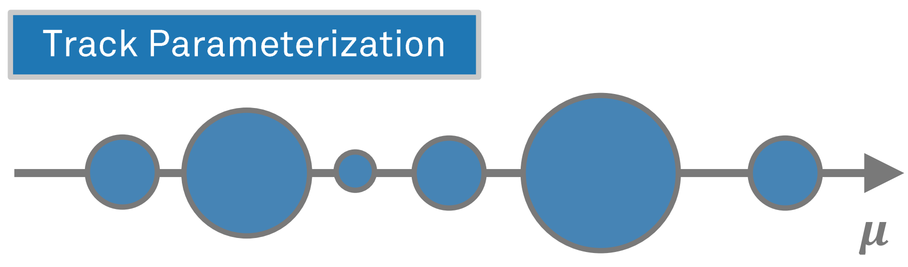
- For accurate description of event O(100) cascades are required!
- Current implementation runs into runtime and memory issues

⟶ High number of parameters: difficult minimization!



Entering μ





Talk Outline

Importance of Domain Knowledge and Symmetries

Event Reconstruction in IceCube

- The IceCube Neutrino Observatory
- Data format and challenges

Convolutional Neural Networks

- Exploit approximate translational invariance

Combining Maximum-Likelihood with Deep Learning

- Maximum-Likelihood Estimation
- Exploiting available symmetries and domain knowledge

Ongoing and future work

Conclusions

Conclusions

Utilization of Domain Knowledge and Symmetries:

- Reduces problem complexity
- Promotes robustness and generalizability
- Enhances interpretability

Event-Generator Framework:

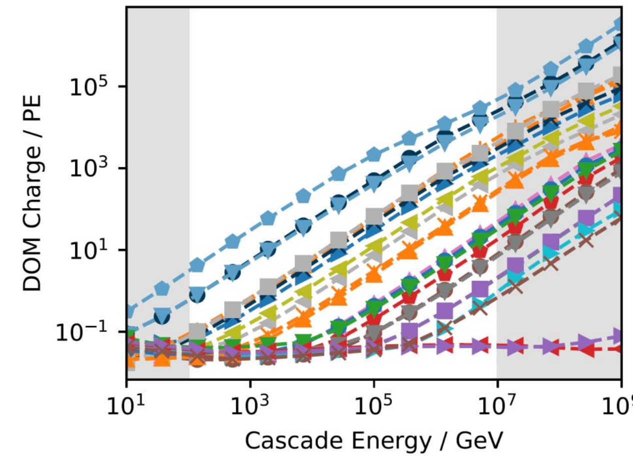
- Interpretable and robust DL via utilization of domain knowledge
- Support for arbitrary event hypotheses:
  - Mostly focused on cascades so far
  - Started to look deeper into tracks
- Versatile tool: reconstruction, stimulation, uncertainty estimation, sky scans, investigating/interpreting results

Ongoing and future work:

- Updates and improvements to Event-Generator framework
- Event selection for track-like event topologies

⟶ Analyses benefit from a combination of different DL methods

Extrapolating beyond training data



Improved performance