Deep learning for fast event reconstruction in the SNO+ scintillator phase and beyond

Mark Anderson[1] (anderson.mark@queensu.ca) **& Cal Hewitt**[2] (cal.hewitt@physics.ox.ac.uk)

on behalf of the SNO+ Collaboration

[1] Queen's University, Kingston, Canada

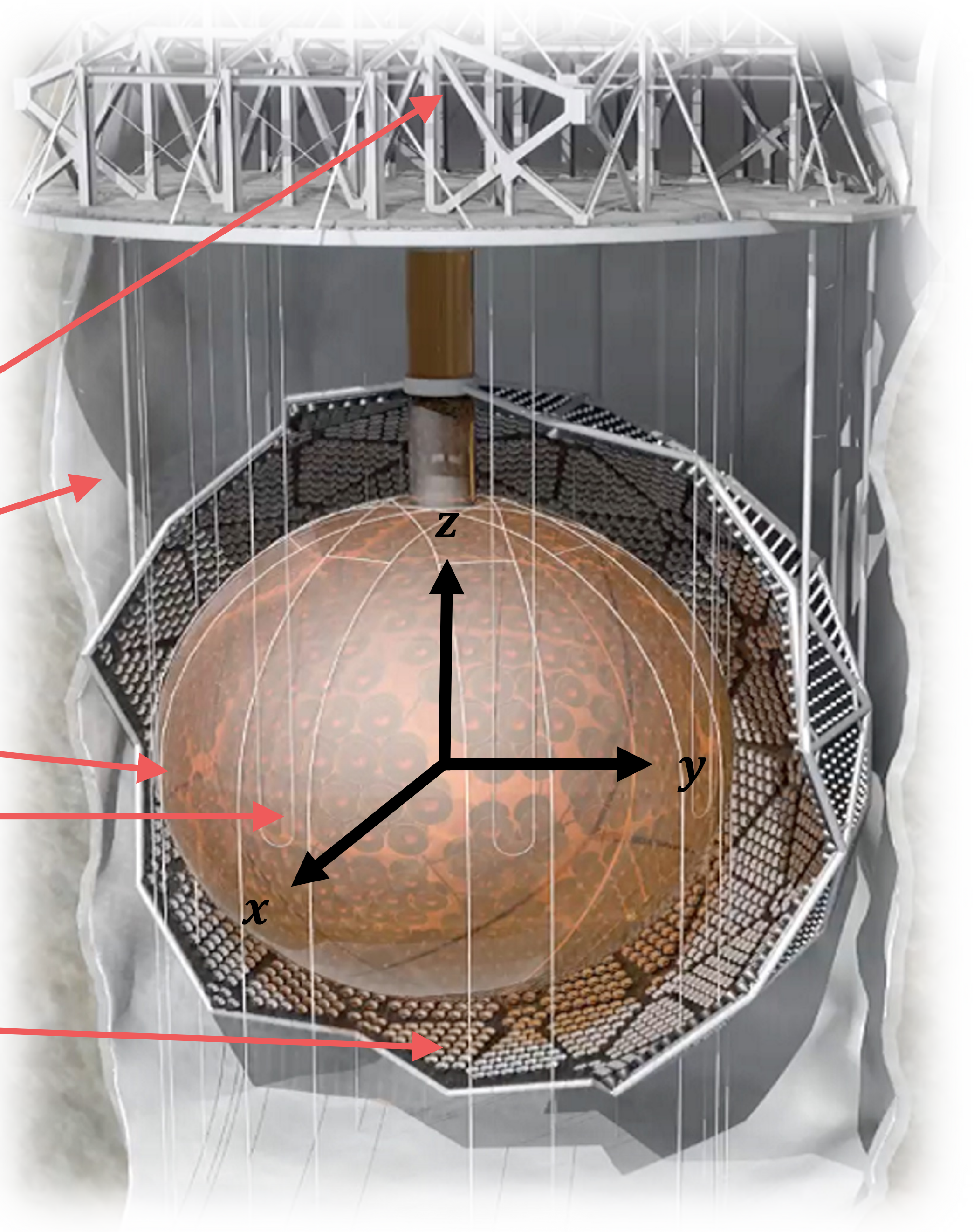
[2] University of Oxford, Oxford, United Kingdom

Neutrino Physics and Machine Learning 2024

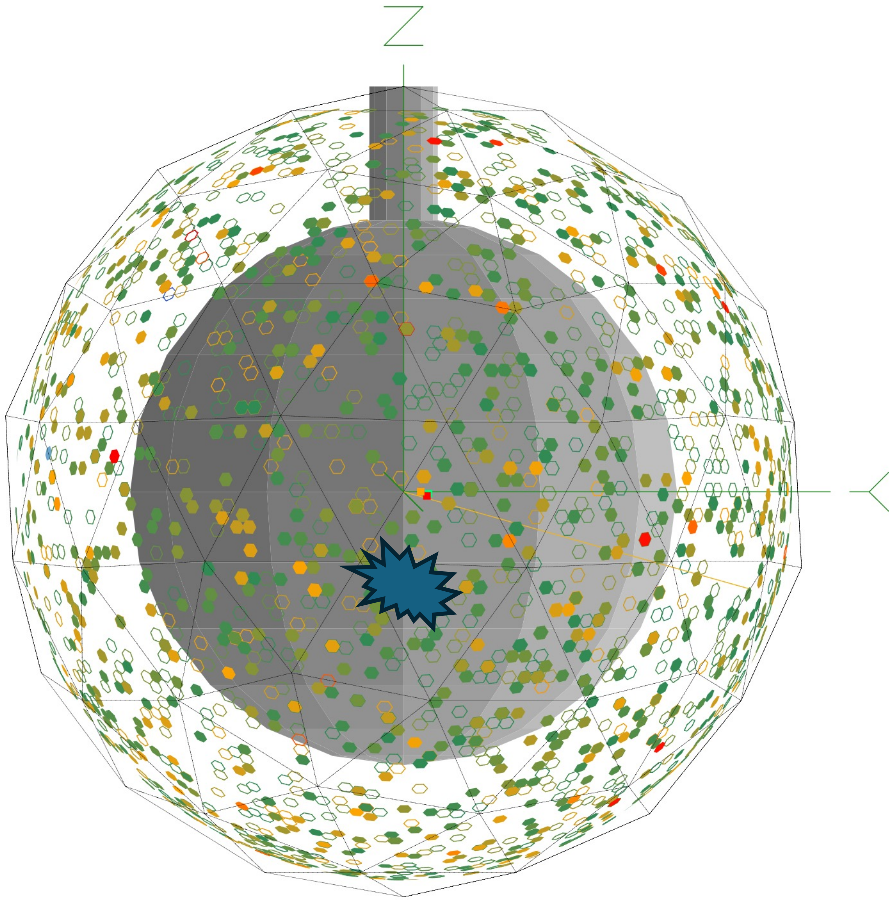
ETH Zürich

The SNO+ Detector

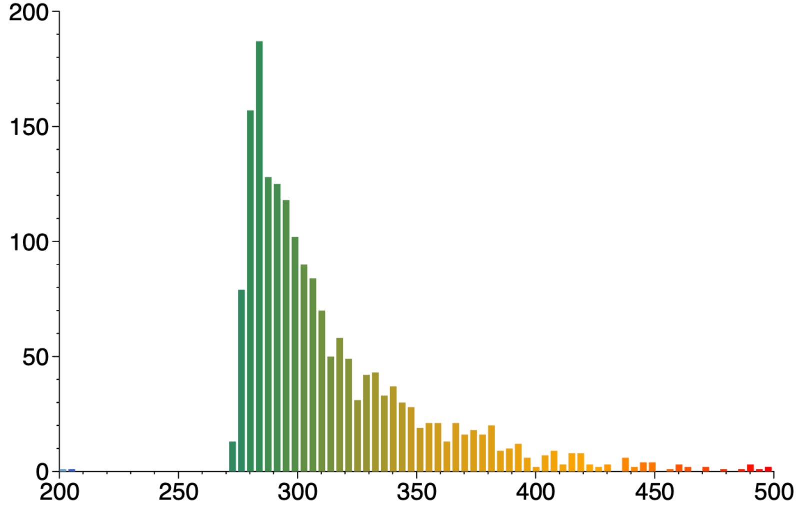
- Multipurpose neutrino physics experiment
- Located 2km underground at SNOLAB
  - ~6000m water equivalent, flat overburden
- Held within large cavity filled with 7kt of ultrapure water for shielding
- 12m diameter spherical acrylic vessel (AV)
- AV filled with 780t of liquid scintillator
  - LAB (bulk solvent) + PPO (fluor)
  - High light yield; ~250 hits/MeV
- Surrounded by ~9400 photomultiplier tubes (PMTs) to detect light from interactions
  - 18m diameter PMT support structure



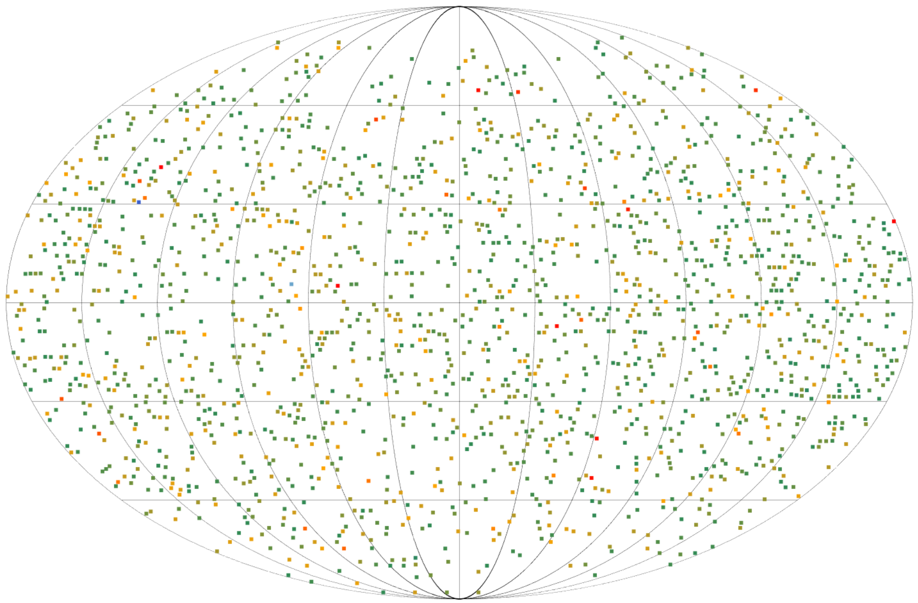
Position Reconstruction



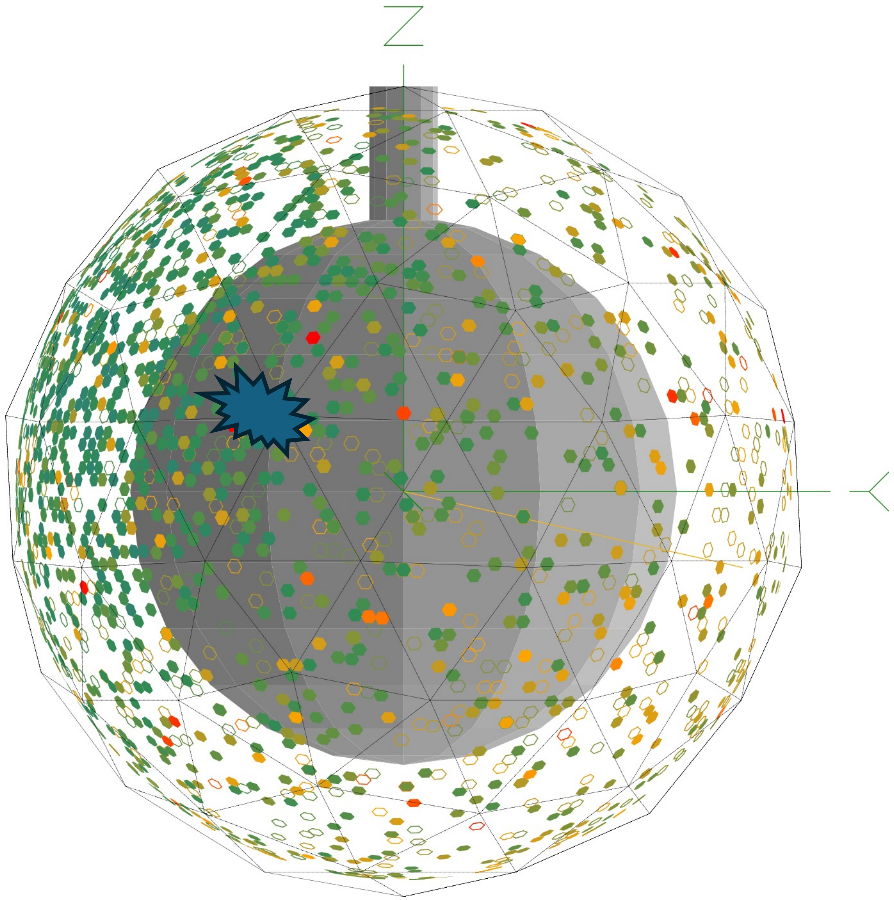
Event near the detector centre



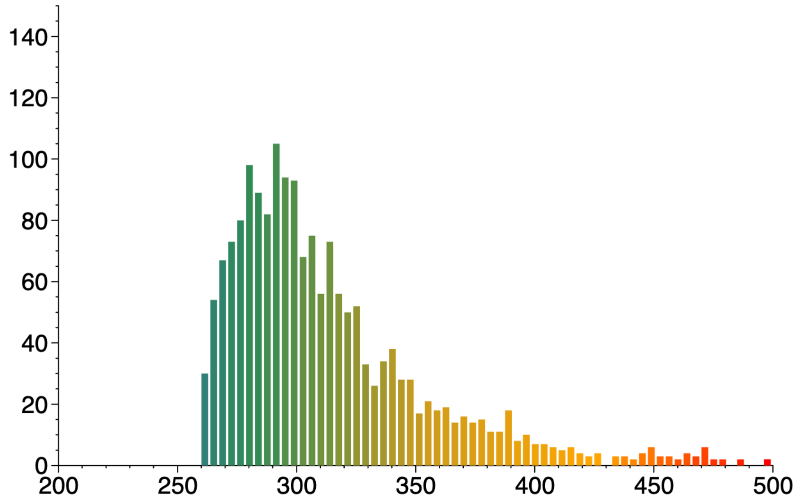
Calibrated PMT hit time (ns)



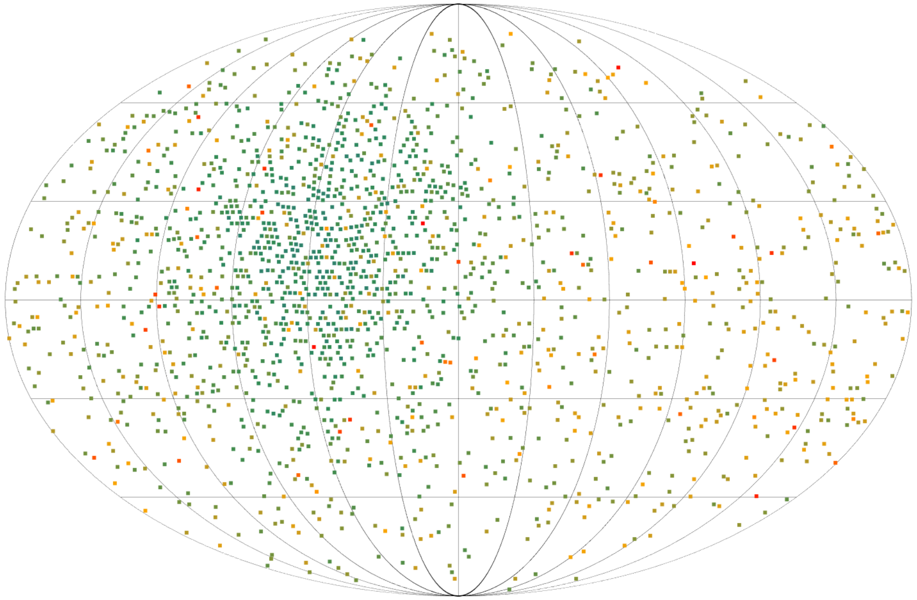
Position Reconstruction



Event near the detector edge



Calibrated PMT hit time (ns)

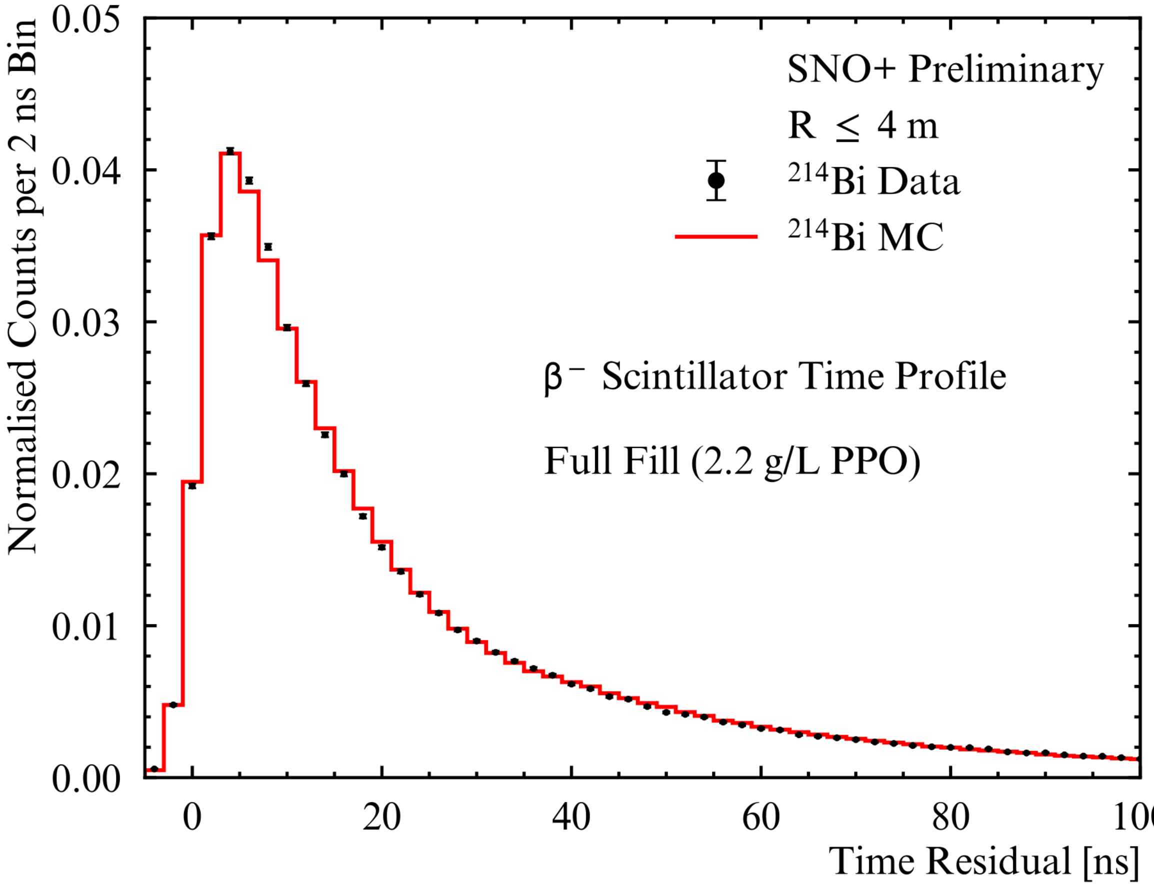


Position Likelihood

- Averaged distribution of **time residuals** derived from MC

$$t_{\text{res}} = t_{\text{hit}} - t_{\text{event}} - t_{\text{ToF}}$$

- **PMT hit time** is observable
- **Time of flight (ToF)** is from candidate event position, $\vec{r}_{\text{event}} = (x, y, z)$, to hit PMT



- **Optimize** event position and **event time** for likelihood that observed time residual distribution drawn from this PDF; product over the **number of hit PMTs (N_{hits})**

Position Likelihood

- Averaged distribution of **time residuals** derived from MC

$$\boldsymbol{t_{\text{res}} = t_{\text{hit}} - t_{\text{event}} - t_{\text{ToF}}}$$

- **PMT hit time** is observable
- **Time of flight (ToF)** is from candidate event position, $\vec{r}_{\text{event}} = (x, y, z)$, to hit PMT

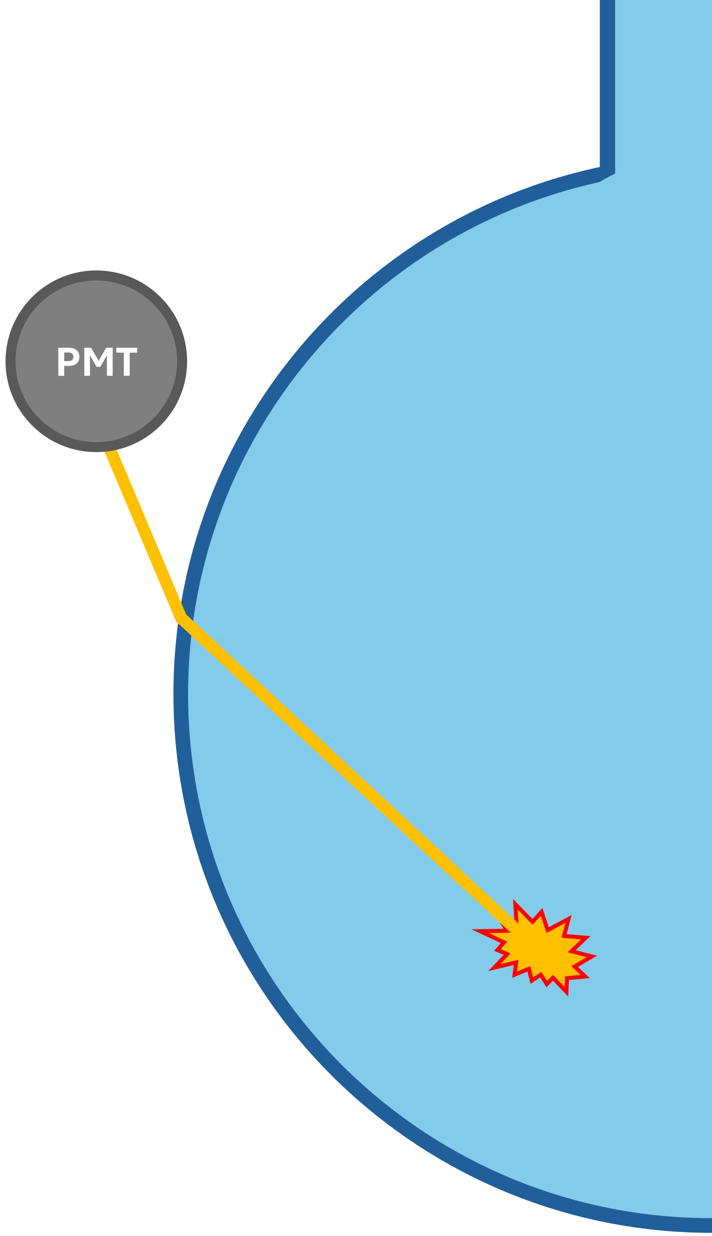
$$\mathcal{L}_{\text{vertex}} = \prod_{i=1}^{\text{Nhits}} P(t_{\text{res},i})$$

$$-\ln(\mathcal{L}_{\text{vertex}}) = -\sum_{i=1}^{\text{Nhits}} \ln P(t_{\text{res},i})$$

- **Optimize** event position and **event time** for likelihood that observed time residual distribution drawn from this PDF; product over the **number of hit PMTs (N_{hits})**

Position Likelihood

- Light travels through three materials in simplest case: time of flight must be estimated with **analytic light path calculator**
- Time residuals distribution has dependences on energy and position; mitigated with **effective corrections**
  - Effective speed of light which scales linearly with N_{hits}
  - Separate time residual PDFs used for different radial shells

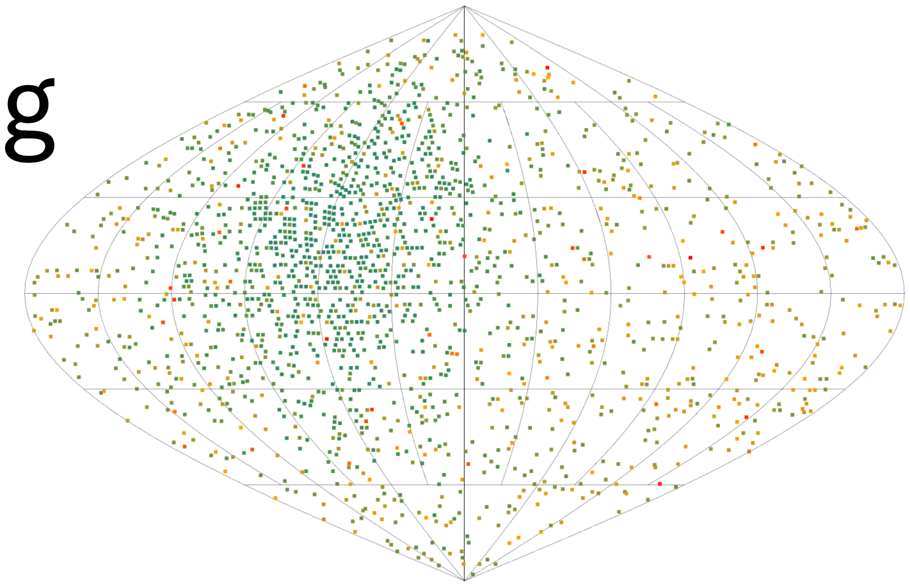


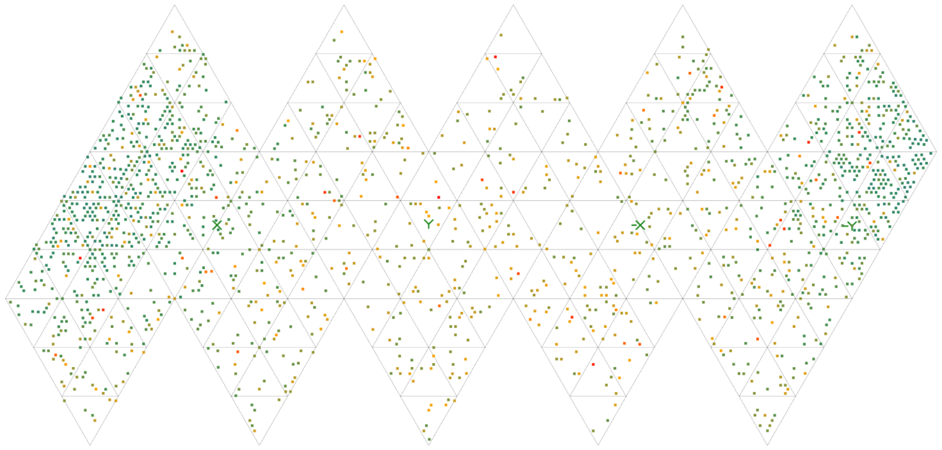
Motivation for Machine Learning Methods

- Likelihood method is comparatively slow (relative to trained model)
- Likelihood method can fail to converge
- Potential to improve upon areas where likelihood method does not do as well (near the AV, neck)
- Independent, complementary approaches are always welcome
  - Identify and correct problems (in either algorithm)
  - Could be used as a seed to the likelihood method

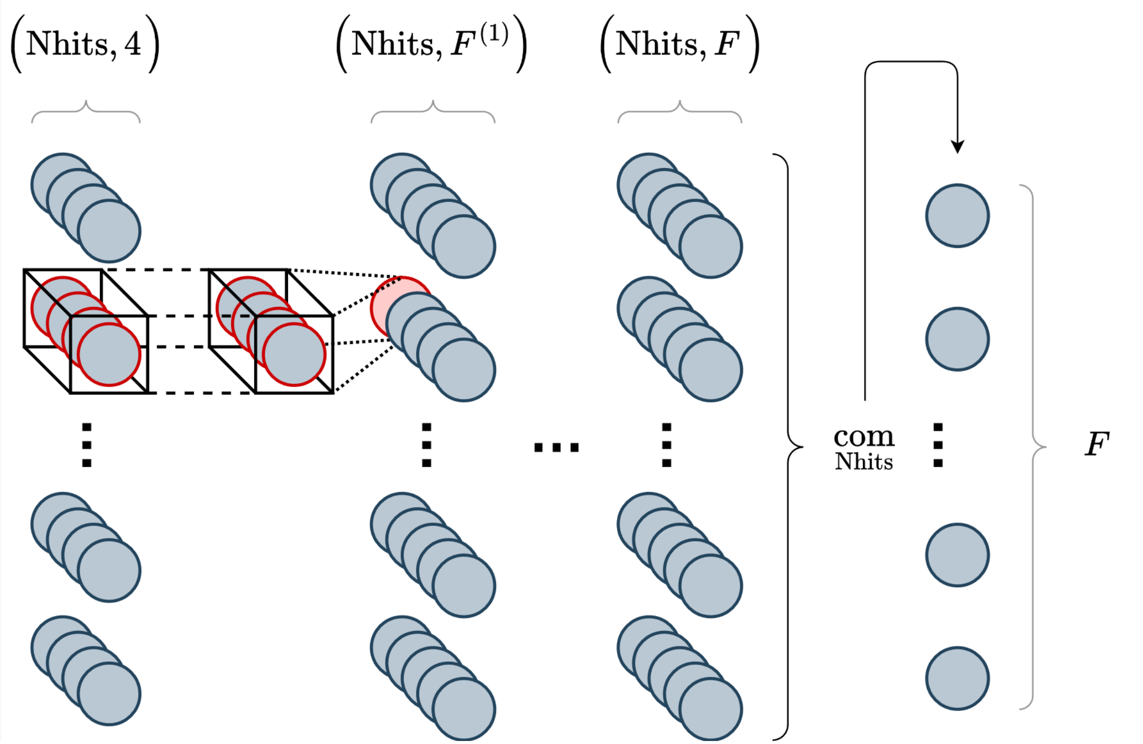
Challenges of Machine Learning

- Not straightforwardly applied to the data
  - Inputs are not the same length from event to event
  - Geometry needs to be provided in some way (either in the data itself or in the architecture)
- Naïve solutions fail
  - Feeding the network a length ~9400 vector of mostly zeroes: typical event is too sparse
  - Projecting the 3D spherical detector onto a 2D surface: no clear way to do this (all projections will distort the original)





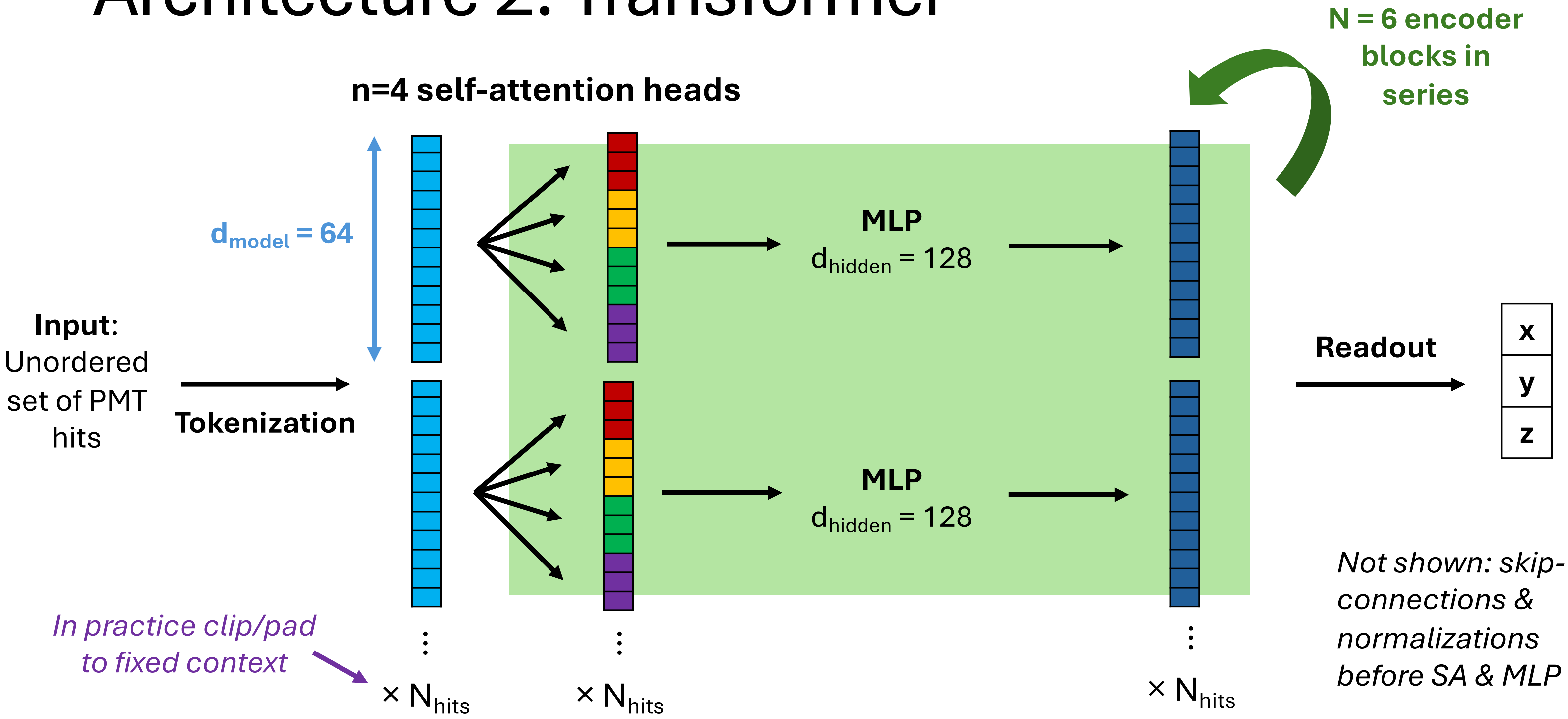
Architecture 1: Convolutional Neural Network



- Network consists of a **1×1 convolutional feature extractor**
- Input is the **set of PMT hit information,** $(x_{\text{hit}}, y_{\text{hit}}, z_{\text{hit}}, t_{\text{hit}})$, for each PMT
- **Commutative operation** (mean) applied over Nhits axis
- Outputs **fixed-length, permutation-invariant** representation of size F

- New representation is fed to a standard fully-connected neural network which predicts the **Cartesian coordinates of the event position,** $\vec{r}_{\text{event}}$

Architecture 2: Transformer



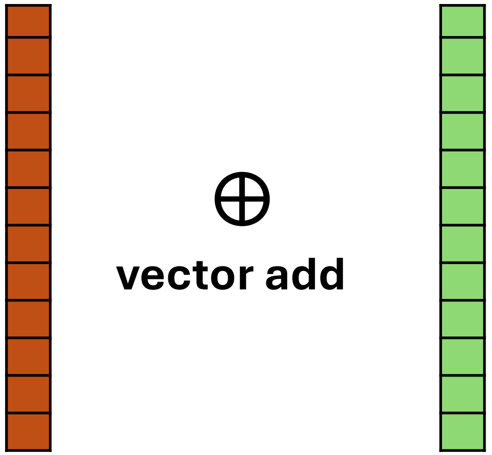
Not shown: skip-connections & normalizations before SA & MLP

Architecture 2: Transformer

Tokenization

Fully learnt features for each PMT ID

Outperforms mapping PMT (x,y,z) / Fourier embedding with MLP



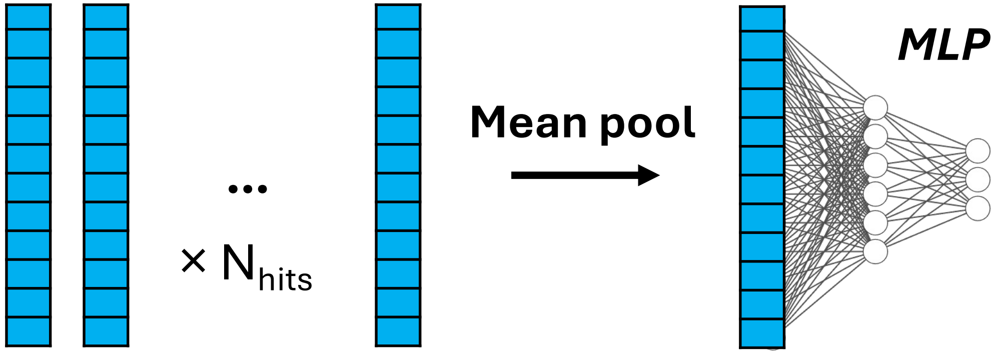
MLP

$d_{hidden} = 512$

Calibrated hit time

Any event-wise information mapped onto every token (N_{hits} …)

Readout



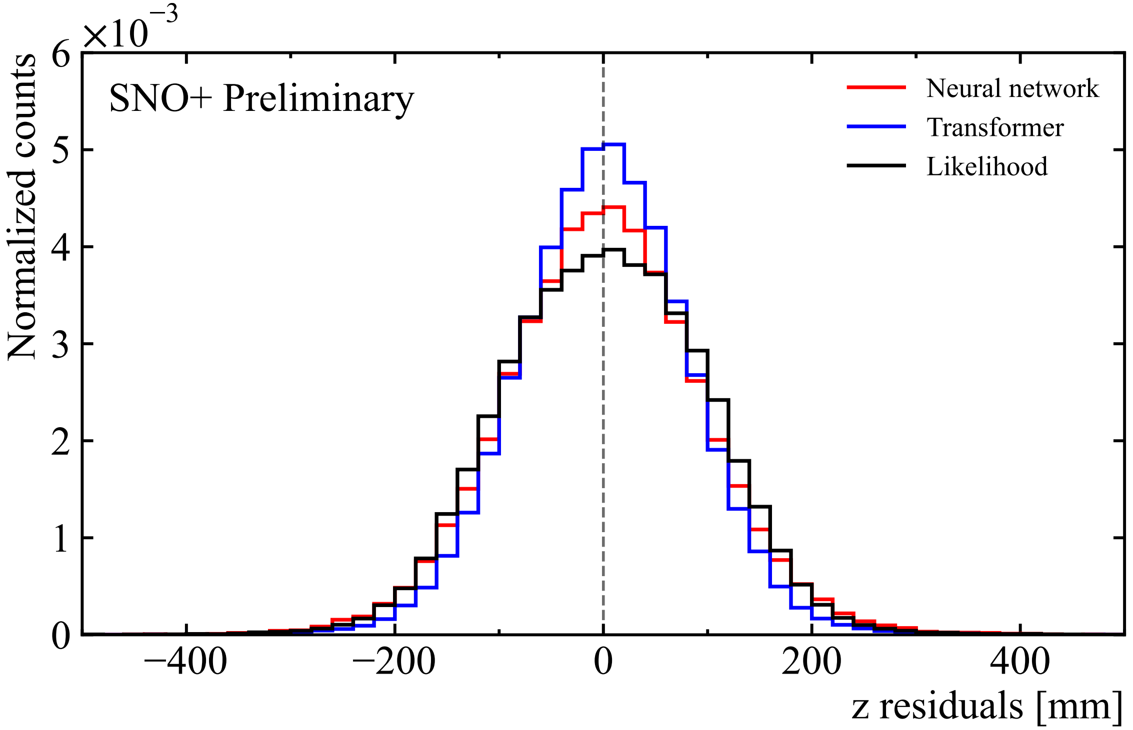
Slightly outperforms reading out from dummy token at top of stack

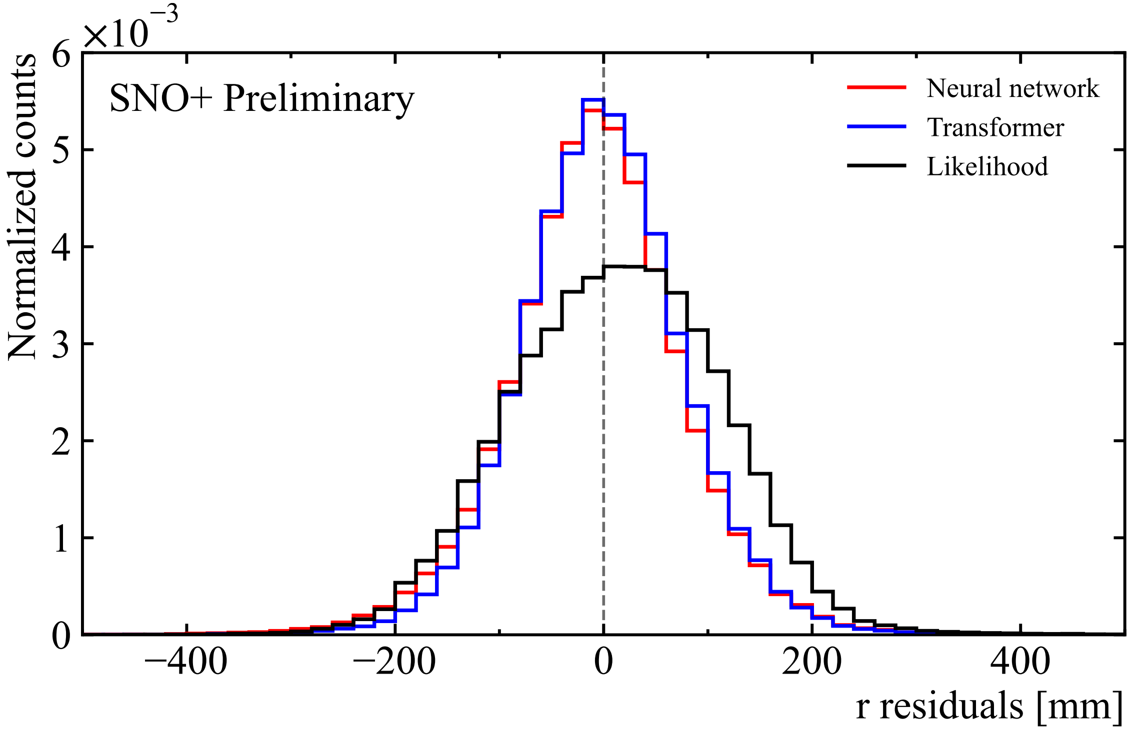
No improvement from supplying N_{hits} and regressing event energy

Datasets

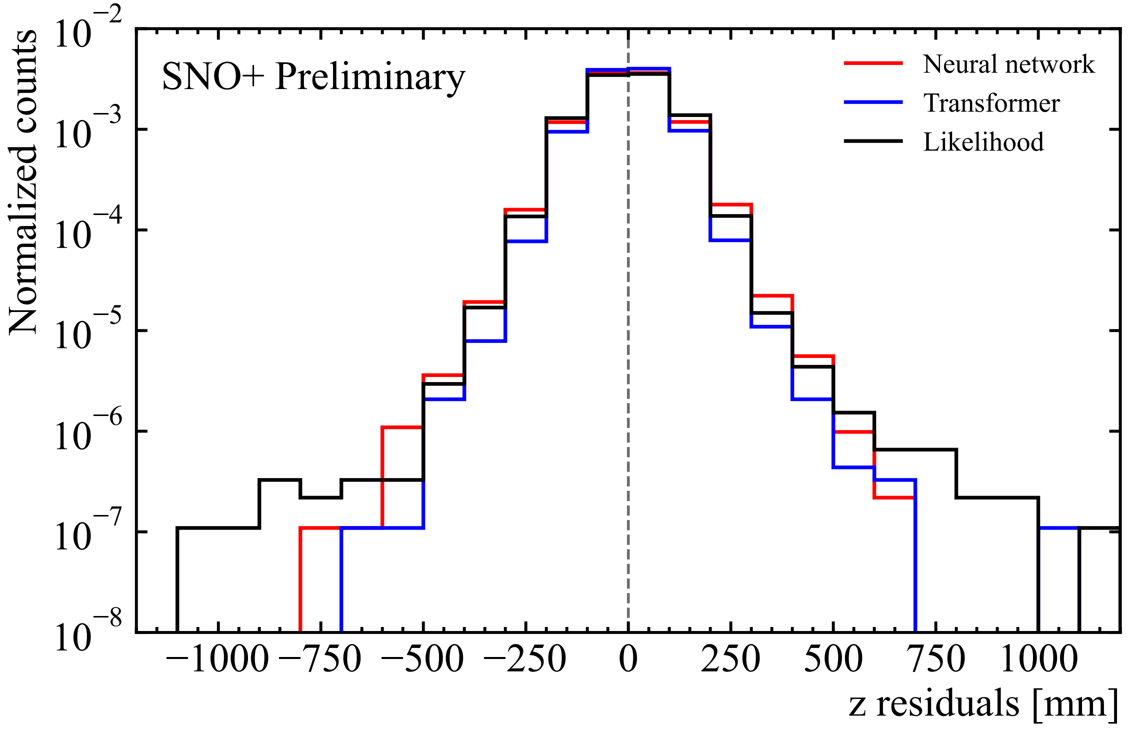
- Models trained and evaluated on simulated single electrons
  - Uniform in energy from 0.5 – 10 MeV: covers (almost) all events of physics interest in SNO+
  - Uniformly distributed inside the acrylic vessel
  - Isotropic in momentum
- 1 million events in dataset of which 900,000 used for training; 5,000 for validation; 95,000 for evaluation

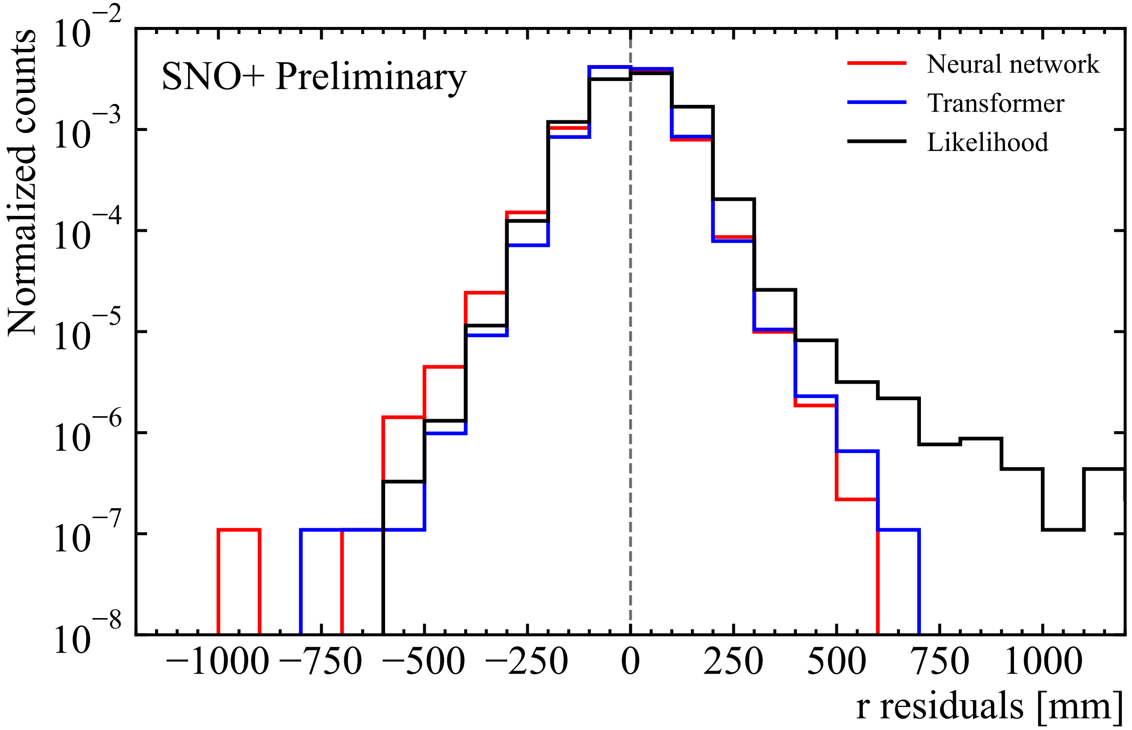
Results: residuals



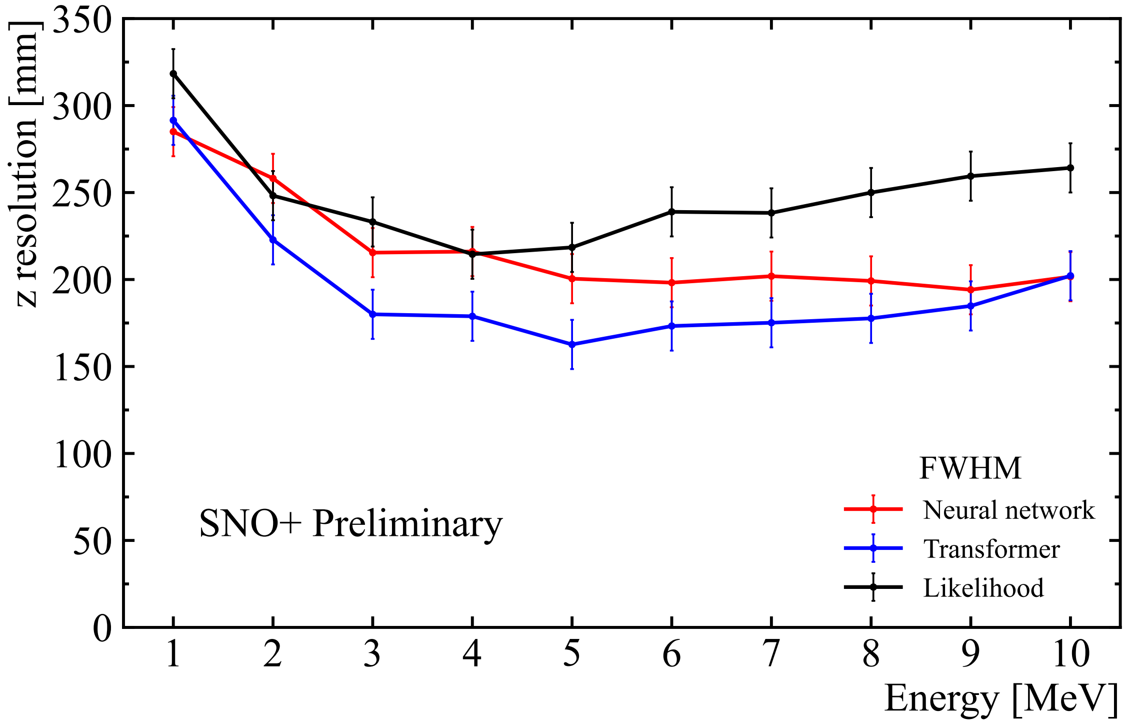


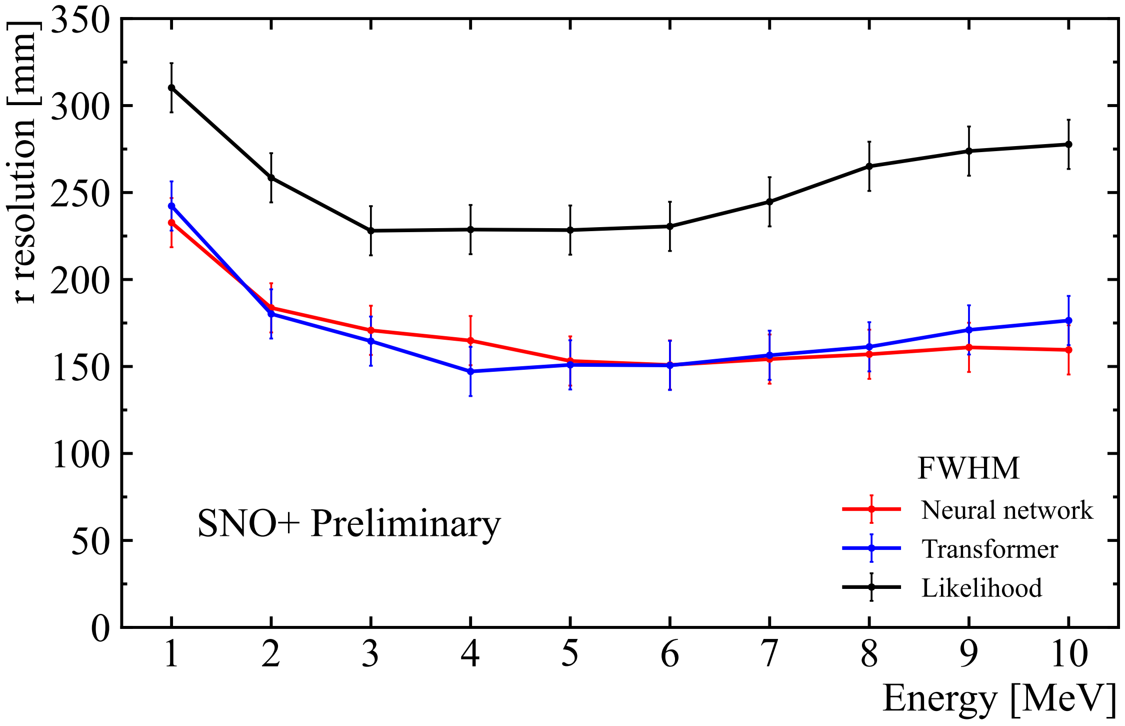
Results: residuals



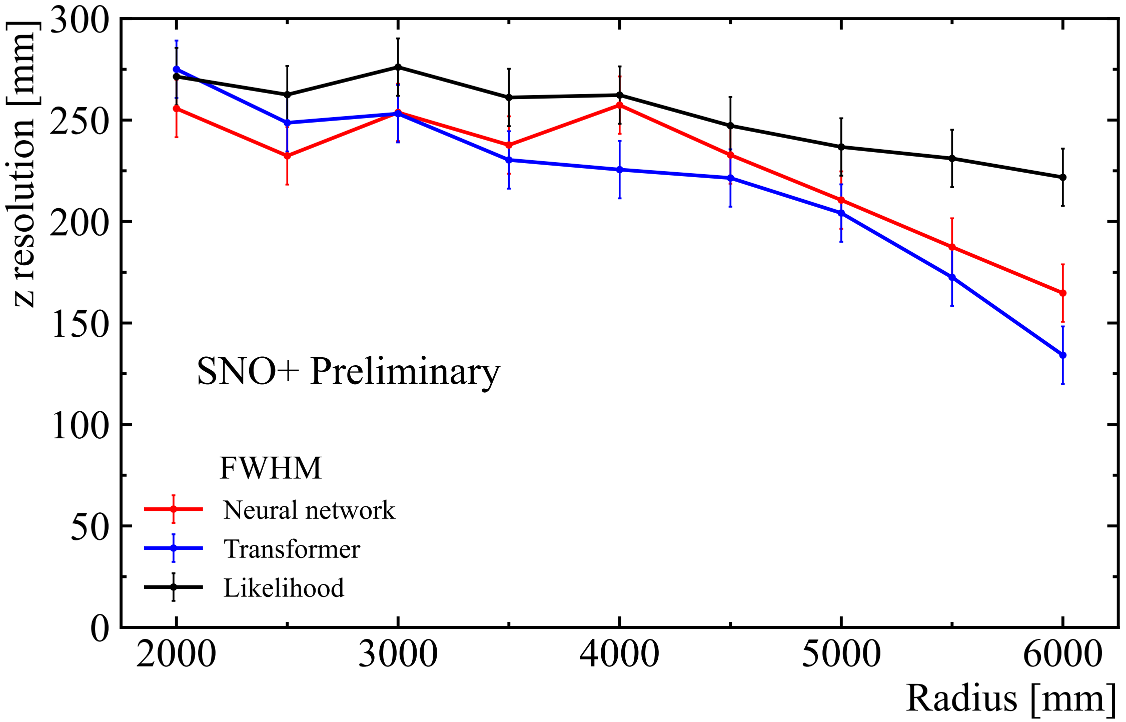


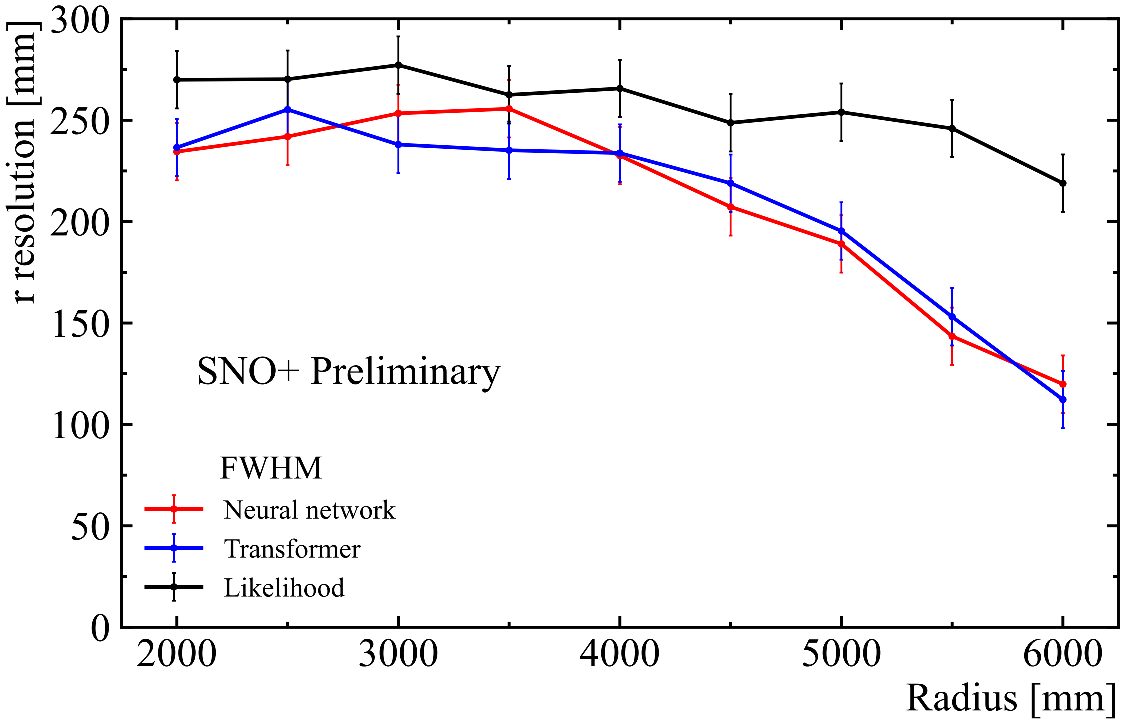
Results: resolution / energy



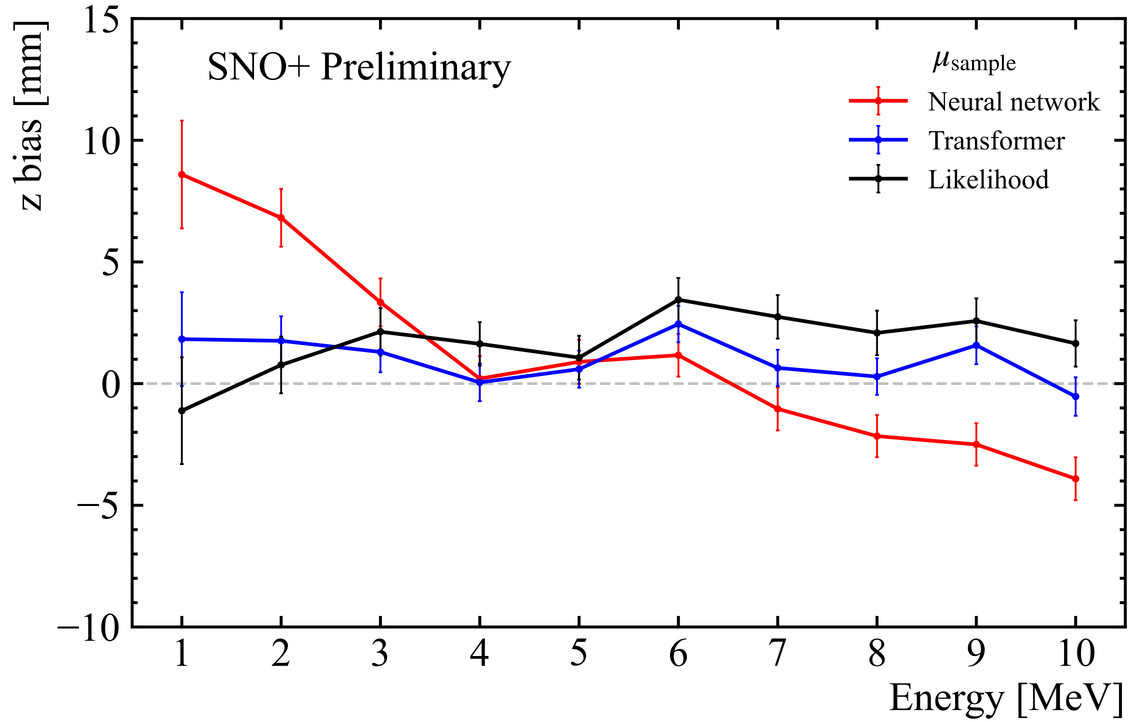


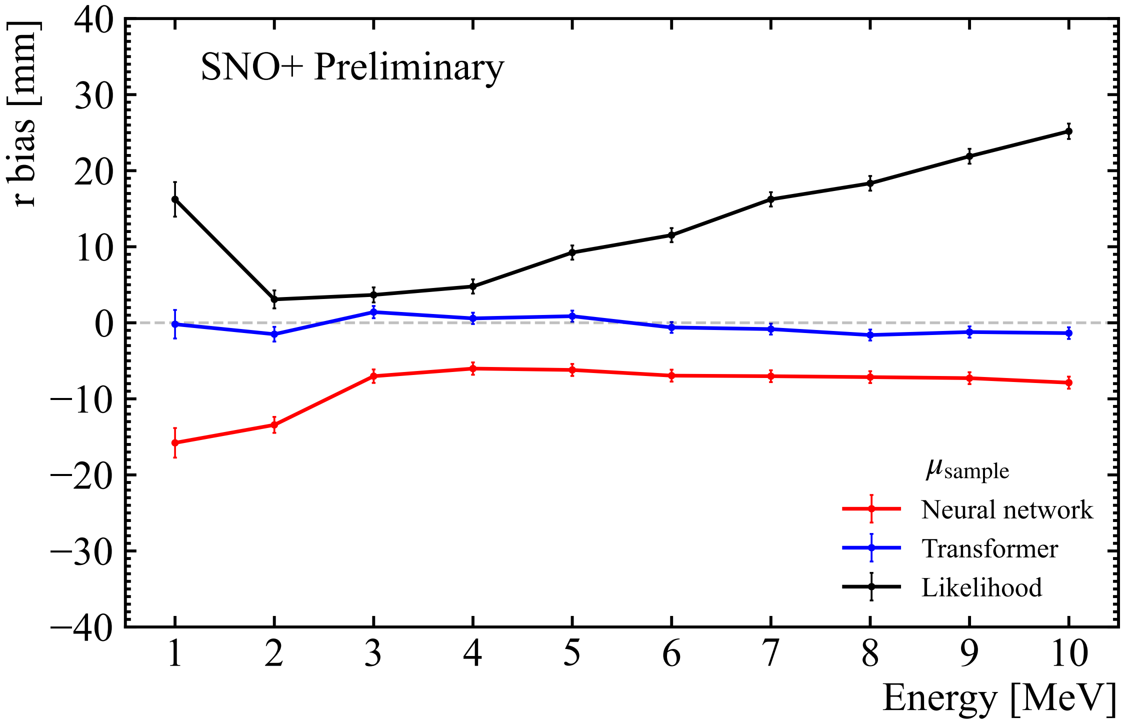
Results: resolution / radius



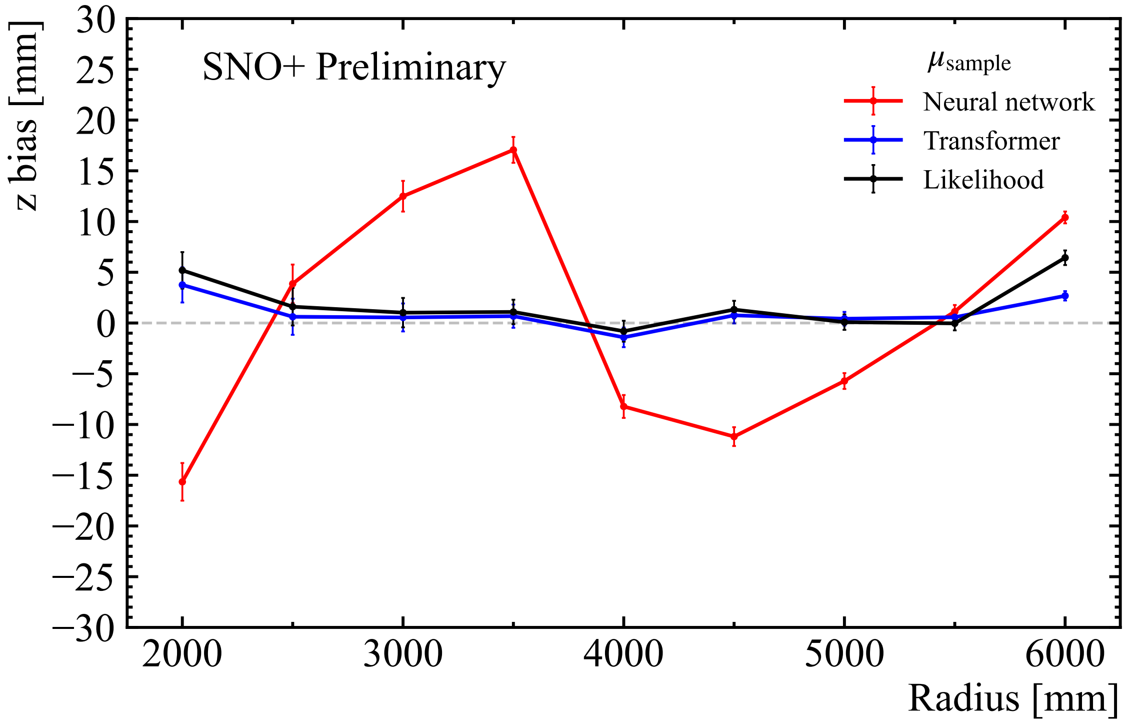


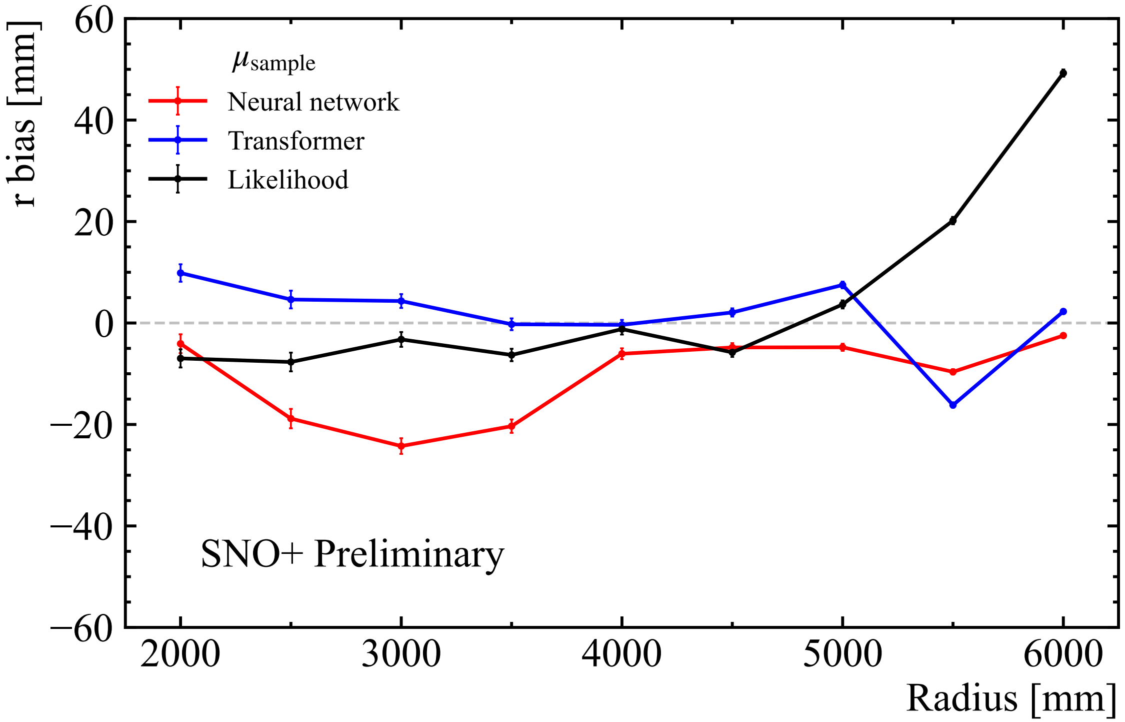
Results: bias / energy





Results: bias / radius





Results: inference time

Method	**Inference time (per event)**	
	CPU [event-by-event]	*GPU [batched]*
Neural network	~10 ms	< 1 ms
Transformer	~170 ms	< 1 ms
Likelihood	~150 ms	N/A

Conclusions

- We present two neural network architectures which effectively ingest SNO+ events; sparse, unordered sets of PMT hits with a variable number of nonzero channels
- Deep learning approaches can provide gains in resolution and a significant reduction in radial bias compared to a maximum likelihood-based method
  - Effectively learns dependencies and allows asymmetric and otherwise difficult regions to be modelled without complex corrections to the likelihood method

Ongoing and Future Work

- Investigating direction reconstruction with promising results
  - Difficult in liquid scintillator due to dominance of isotropic scintillation light
- Studying simultaneous position and direction reconstruction
  - PMT hit patterns depend on both position (mostly) and direction
  - Provides the network with more information
  - Should lead to improvements in both
- Other architectures (e.g., GNNs) show promise