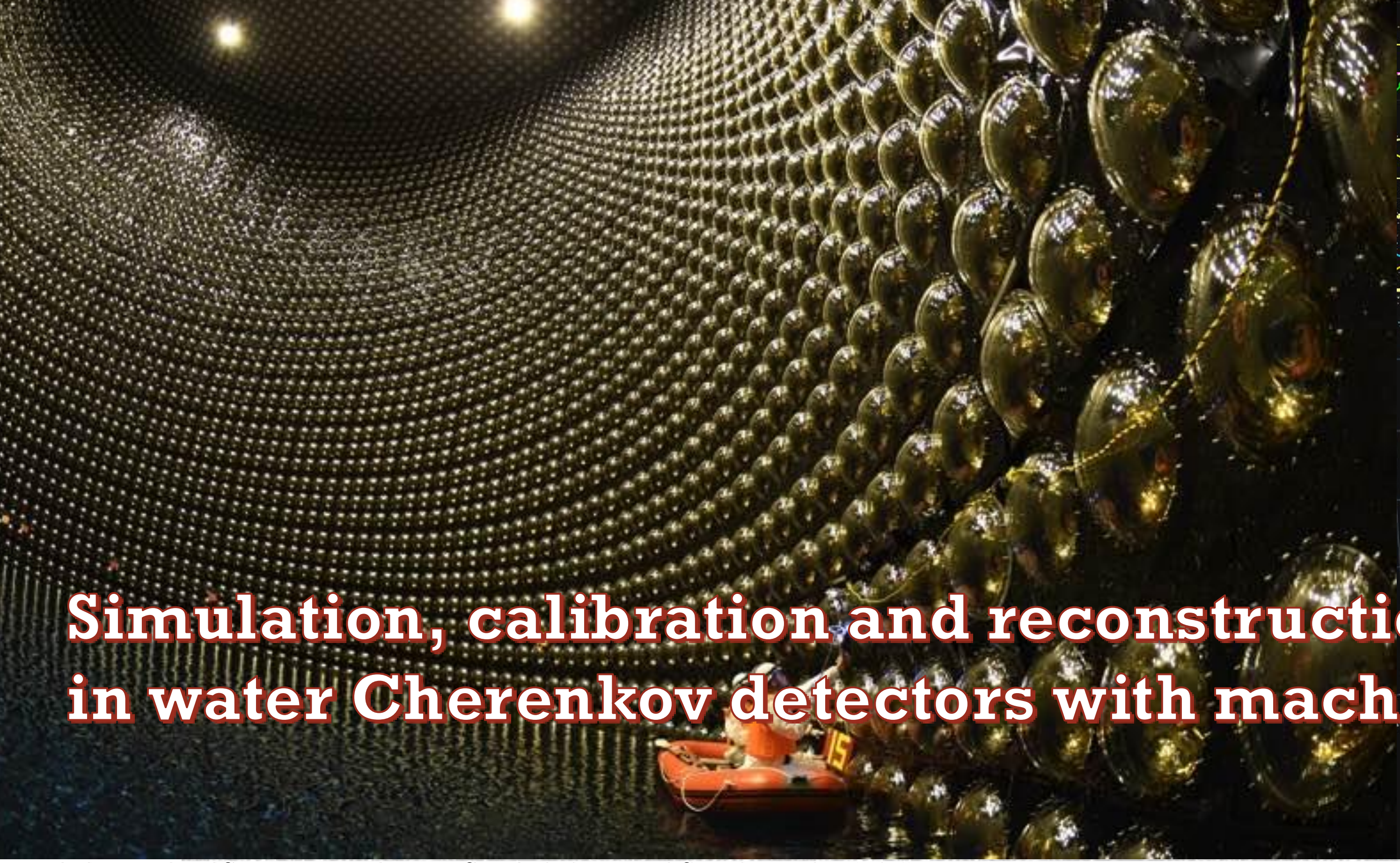
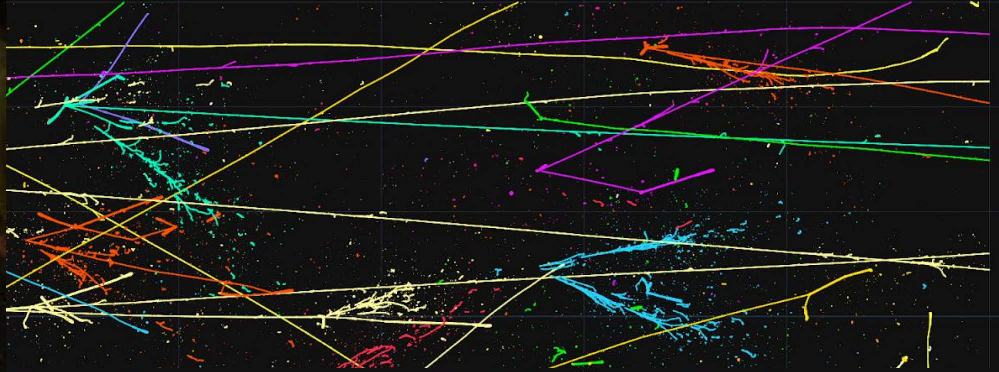




Simulation, calibration and reconstruction in water Cherenkov detectors with machine learning techniques

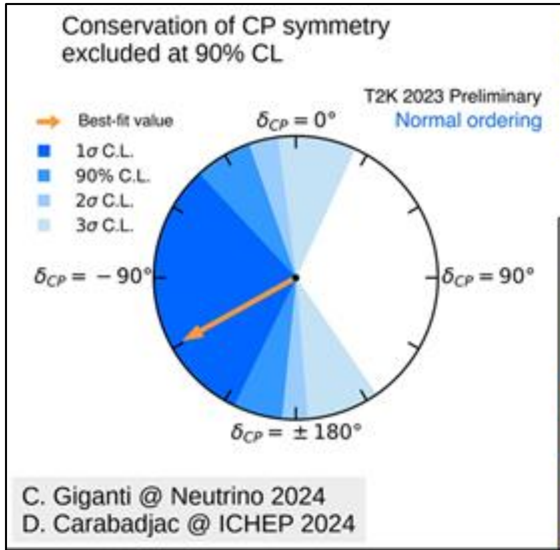
Ka Ming Tsui (Kavli IPMU)

for the CIDeR-ML Collaboration

Kavli IPMU · School of Science, The University of Tokyo · 東京理科大学 Tokyo University of Science · Institute of Science Tokyo · SLAC · University of Minnesota · Stanford University · Tufts University

Physics motivation

- CP violation in neutrinos → Potential source of the matter-antimatter asymmetry
- Hyper-K: 10 years to reach 5σ discovery potential at current systematic error understanding
- Systematic errors must be improved
  - New and upgraded near detectors

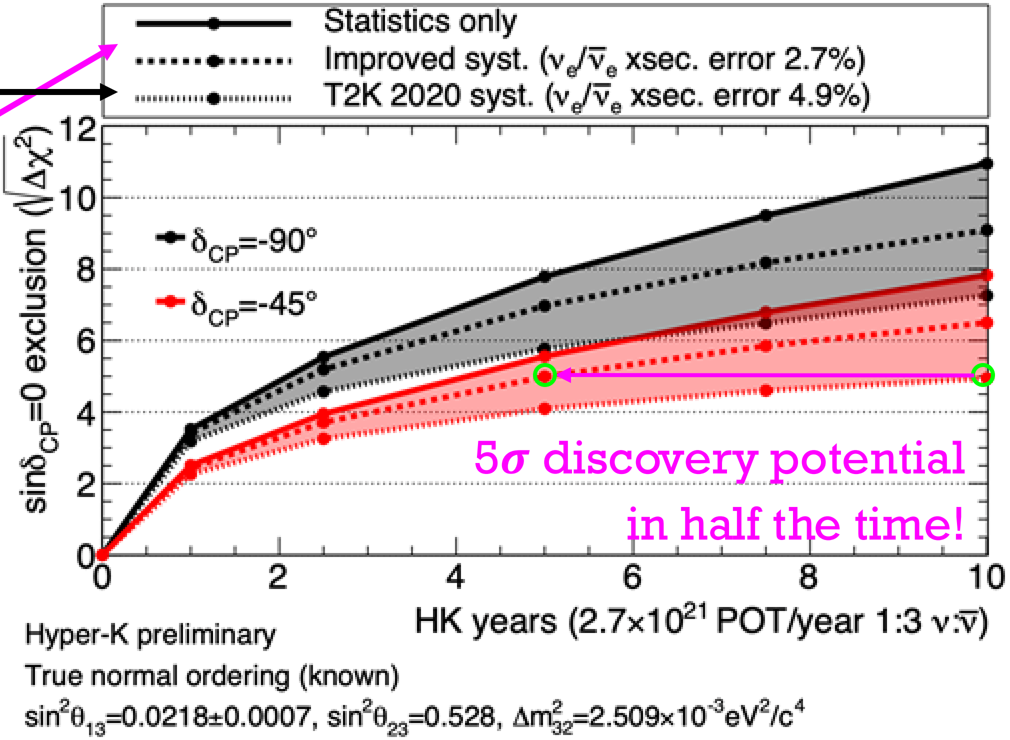


T2K Result

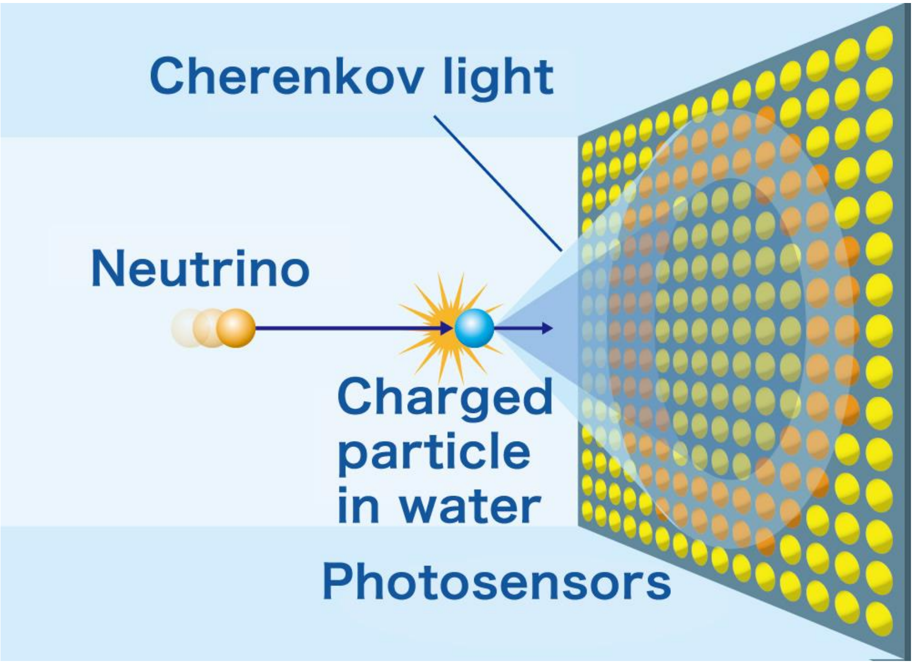
Far detector systematic uncertainty

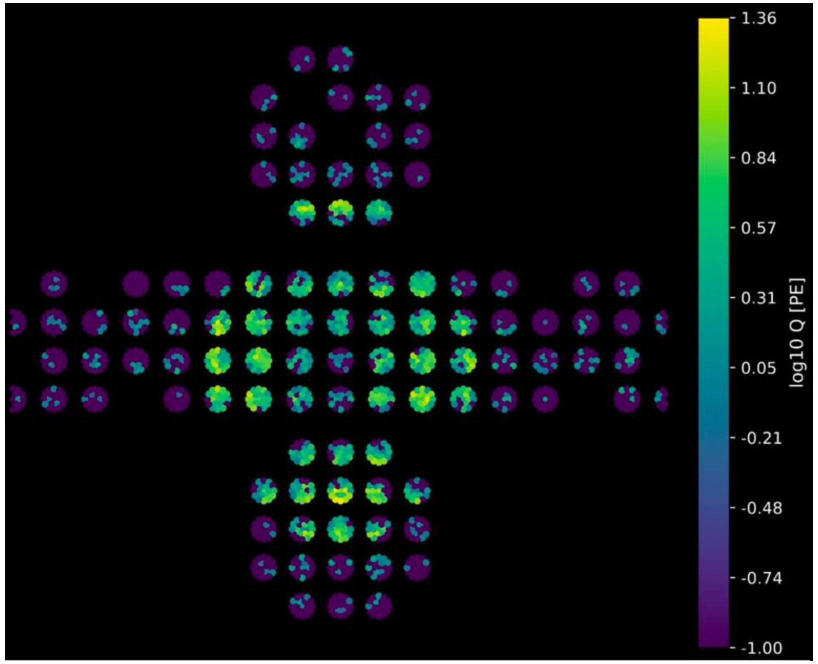
Sample	OA22	New results
ν-mode 1Rμ	3.4%	3.2%
ν-mode 1Re	5.2%	4.9%
ν-mode MR	4.9%	3.9%
ν-mode 1Re+d.e.	14.3%	6.3%
ν̄-mode 1Rμ	3.9%	5.0%
ν̄-mode 1Re	5.8%	6.7%

To HK



Water Cherenkov Detector principles





"T2K 2021 syst.": Phys. Rev. D 103, 112008

Error Source	% Error for CPV search
$\varphi + \sigma$ (ND constrained)	2.7
$\varphi + \sigma$ (ND unconstrained)	1.2
Nucleon removal energy	3.6
π re-interactions	1.6
$\sigma(\nu_e)$, $\sigma(\bar{\nu}_e)$	3.0
NC γ + other	1.5
SK far detector	1.5
Total	**6.0**

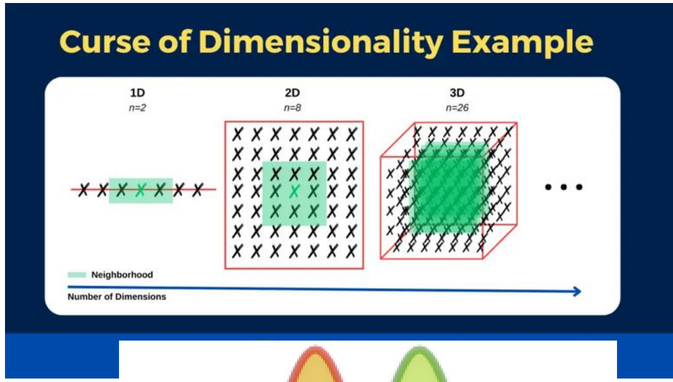
ND Fit

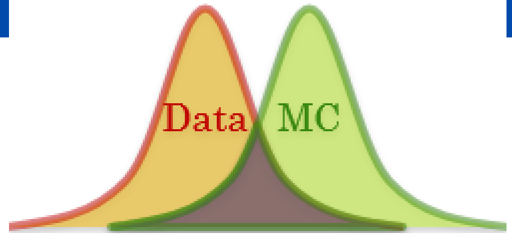
Goal for HK: 3%

- Traditionally use Monte Carlo (MC) simulations tuned with calibration data to generate observables with which we use to analyze real physics data → Hard but necessary to reduce detector uncertainties to <1%

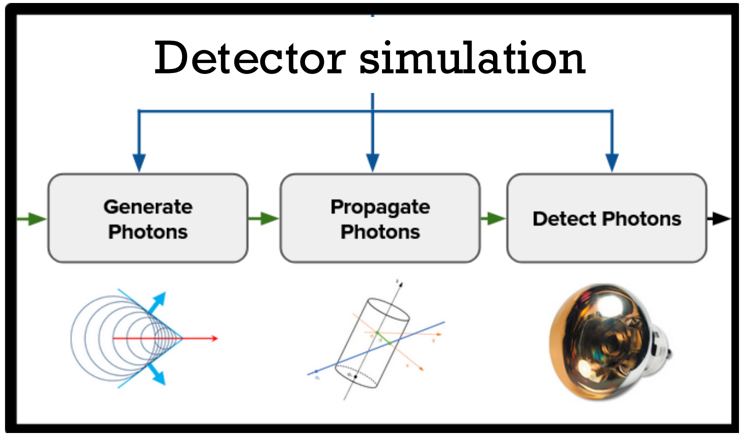
Detector Physics Modeling: Challenges

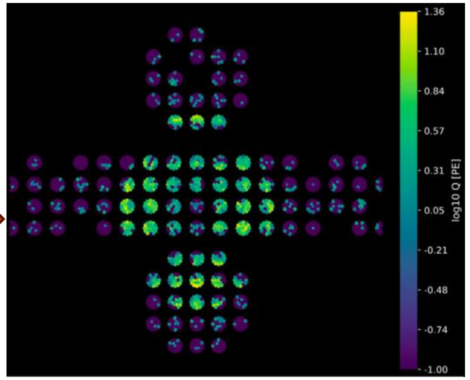
- **Quality** due to limitation in the traditional calibration
  - "one-by-one" (sequential) parameter calibration assumes weak/no correlations
  - Non-optimizable routines (e.g. look-up tables)
  - Not directly minimizing the "data-simulation discrepancy" metrics
- **Speed** for modeling a large and complex detectors without loss of quality
  - Seconds to minutes per event prevents high statistics $O(10^9)$ simulations
  - Faster = can unlock new analysis techniques (e.g. simulation-based inference)
- **Resources** for software development and maintenance
  - Same detector physics, separate software (simulation vs. calibration/reconstruction)



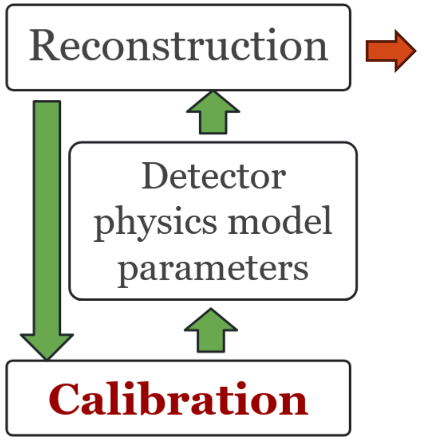


Track input





Detector output

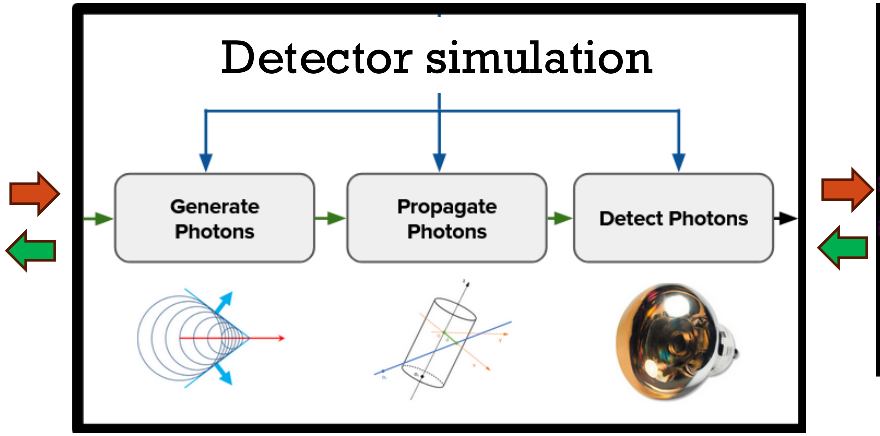


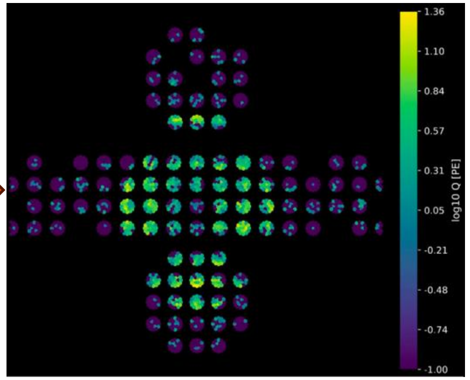
Reconstructed track

Proposal: Differentiable Detector Simulator (DDSim)

- **Automation of physics model tuning (via backpropagation)**
  - **“End-to-end”**: gradient-based optimization using calibration and physics datasets
  - **Interpretable**: analytical physics models for well-understood physics
  - **Flexible**: neural representations to incorporate complex features in real data
  - **Fast**: utilization of modern computing accelerators (e.g. GPUs)
- Proof of concept: **Differentiable Optical Detector Simulation**
  - OpticSiren (this talk), LUCiD (Omar’s talk on Thu)

Track input





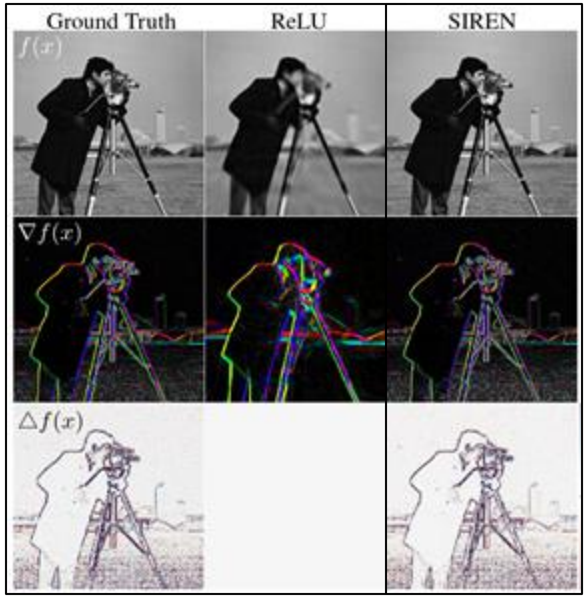
Detector output

OpticSiren

- **Differentiable surrogate model for Optical processes**
- ~ **Look-up Table**: a database of per-PMT photon "visibility" for binned positions/directions
  - Scales poorly for a large detector
  - Created with simulation. Difficult to tune with data
- **Siren**: fully connected deep neural network with sinusoidal activations
  - Implicitly defined, continuous, differentiable signal representations
  - "smooth" = model the underlying physics gradient
  - Optimizable directly on real calibration dataset!

Siren

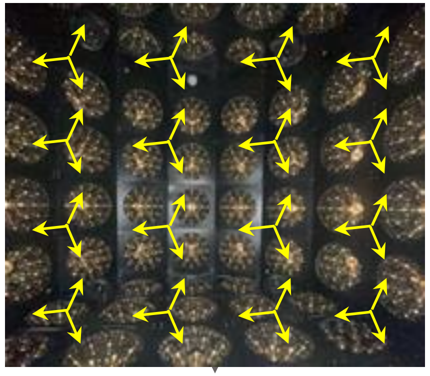
Sinusoidal Representation Network



Can better represent underlying distributions and their derivative compared to other networks

Training and validation

- Demonstration in WCTE (cylinder of length ~ 3 m)
- Simulate PMT response (~2000) with WCSim for *O*(1M) photon starting position×direction → Look-up table as training data
  - Spatial/angular resolution matched with PMT spacing
- Direct check of physical behavior



Photon Propagation

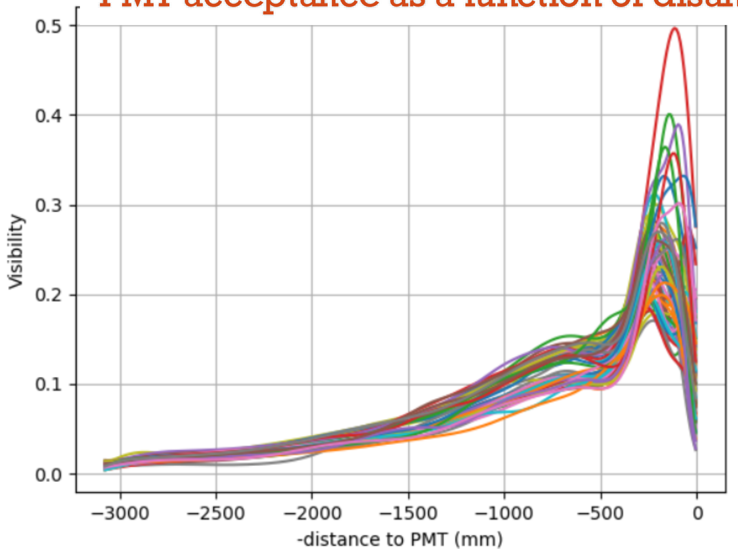
Conventional photon libraries are discrete and sampled

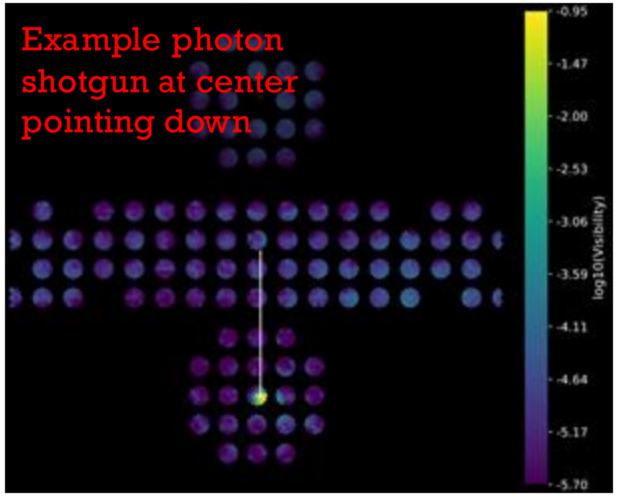
$\mathrm{vis}(\vec{x}, \vec{d})$

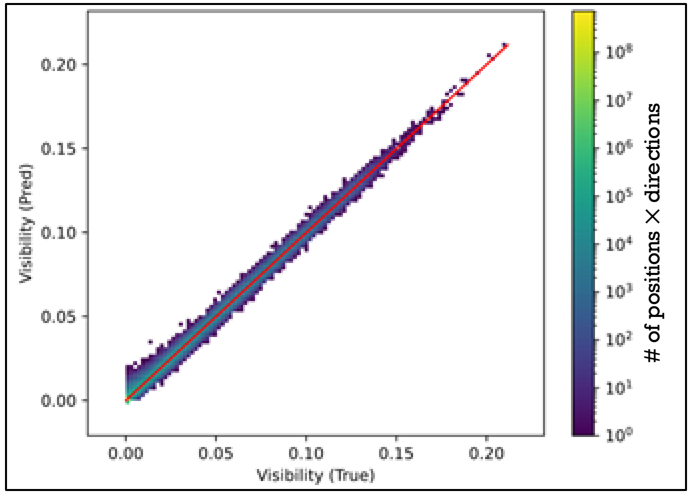
SIREN

Provides continuous and differentiable photon library

PMT acceptance as a function of disanrce







visibility: probability of observing 1 photon from some position in detector

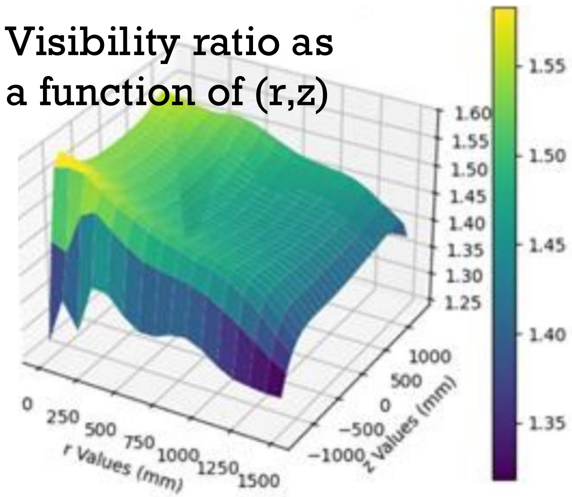
Calibration

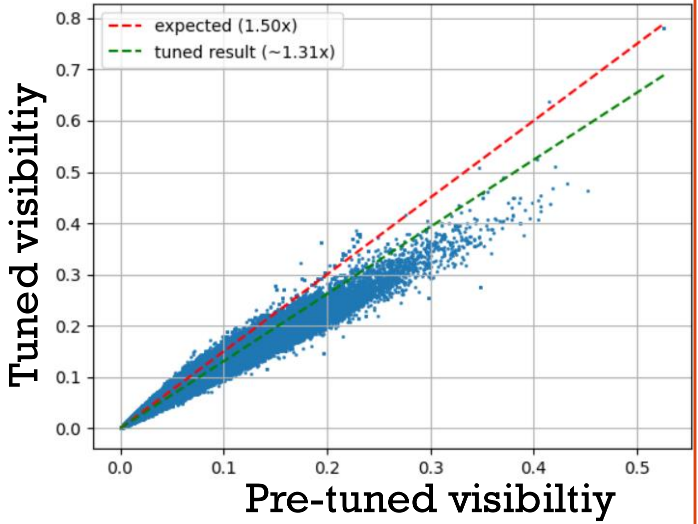
- Fine-tuning with (fake) calibration data
  - Fixed isotropic light sources: "overfit" around the given positions
  - On going work on cosmics → span whole detector

Fake calibration data with charge × 1.5

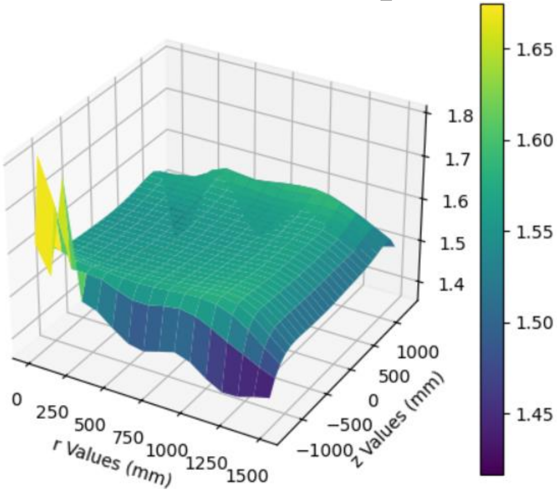
Case 1: 3 isotropic sources at (0,0,0), (0,0,-1m),(0,0,1m)

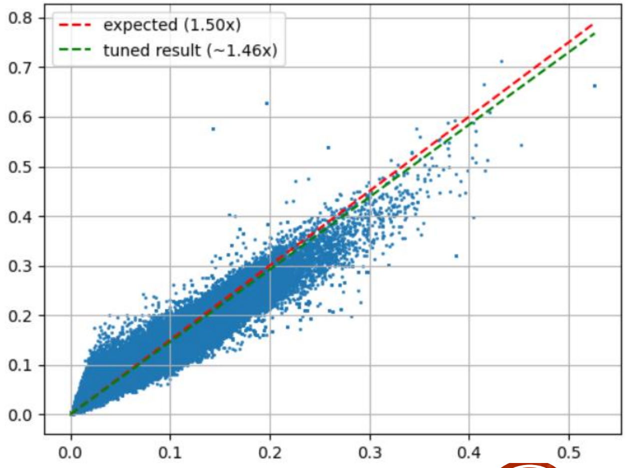
Visibility ratio as a function of (r,z)





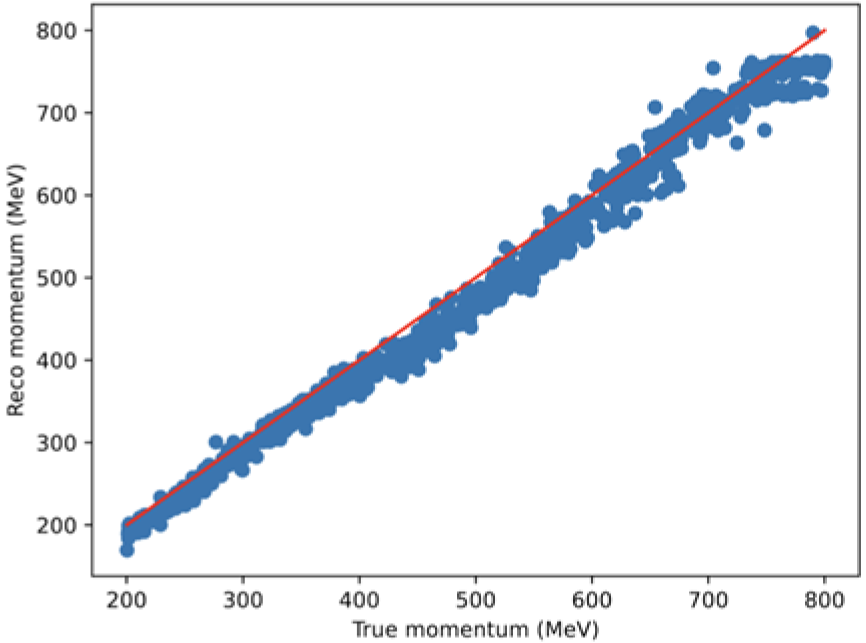
Case 2: 126 isotropic sources at 14 z-pos ⊗ 9 xy-pos

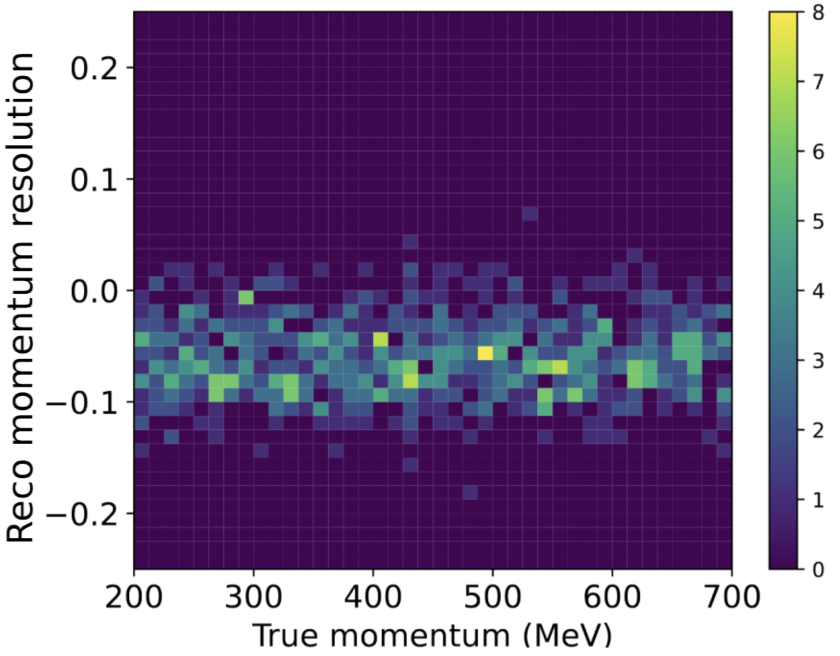




Reconstruction

- Connect the **Differentiable Optical Detector Simulation** with another differentiable model of Cherenkov photon production (Cherenkov Siren) → optimization of track parameters
- Muon momentum reconstruction with **OpticSiren**, on WCTE MC data

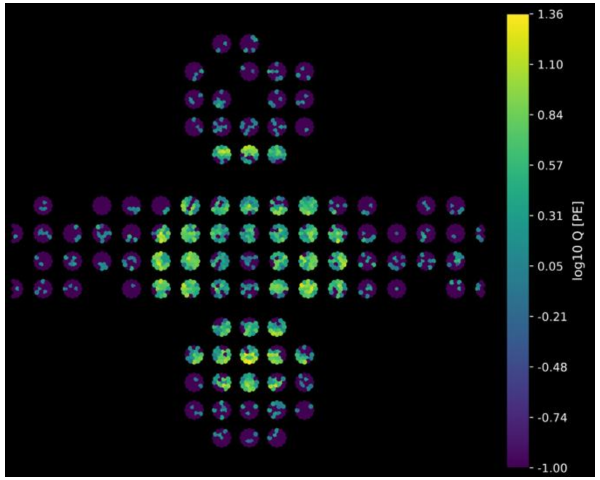




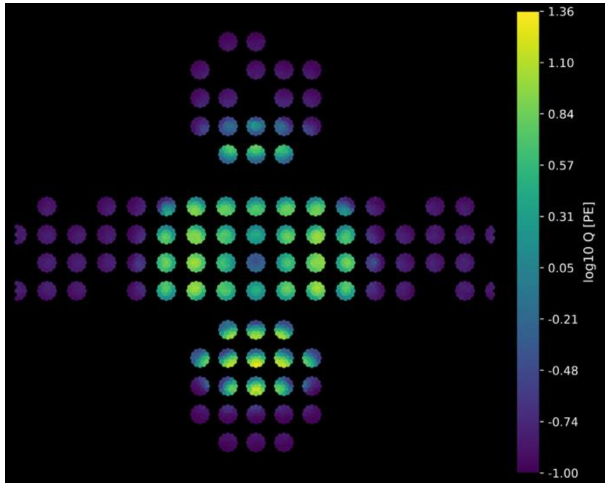
2.7% momentum resolution (assuming fixed position & direction, bias-corrected, but **charge only**), relative to:

- ResNet (WatChMaL, charge+time): 2.5%
- Traditional likelihood (fiTQun, charge+time): ~4%

MC



Muon of 500 MeV/c



DDSim

Summary

- Future long-baseline neutrino oscillation experiments will become systematics limited
  - New near detector technologies will mitigate flux+interaction systematics
    → Reducing detector systematics becomes crucial
- Differentiable detector simulator (DDSim) proof of concept demonstrated for both water Cherenkov and liquid argon detectors
  - Already competitive performance with current traditional algorithms
- Ongoing developments to improve performance and speed of the DDSim
- Aiming for applications to real data (WCTE, DUNE ND prototypes) this FY