Atmospheric Background Reduction using CNNs in DSNB Searches at SuperK-Gd

Soniya Samani

on behalf of the Super Kamiokande Collaboration

XVIII International Conference on Topics in Astroparticle and Underground Physics 2023

30th of August 2023

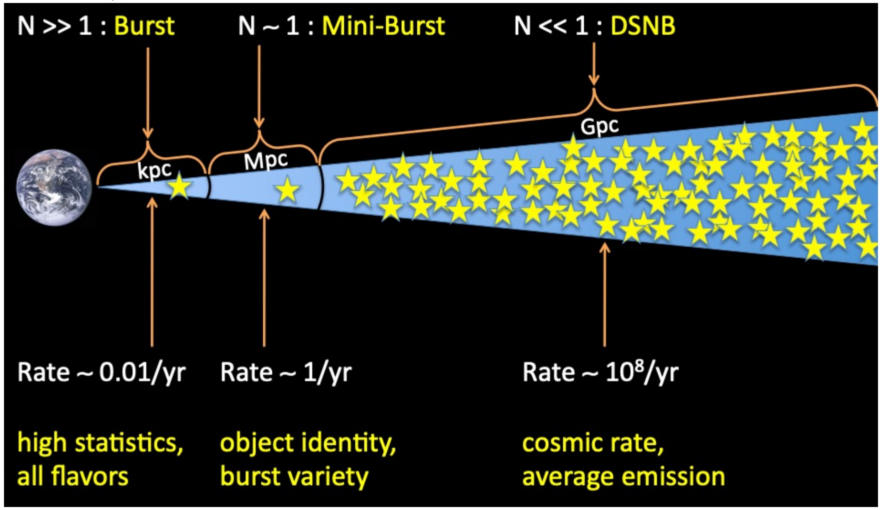
UNIVERSITY OF OXFORD

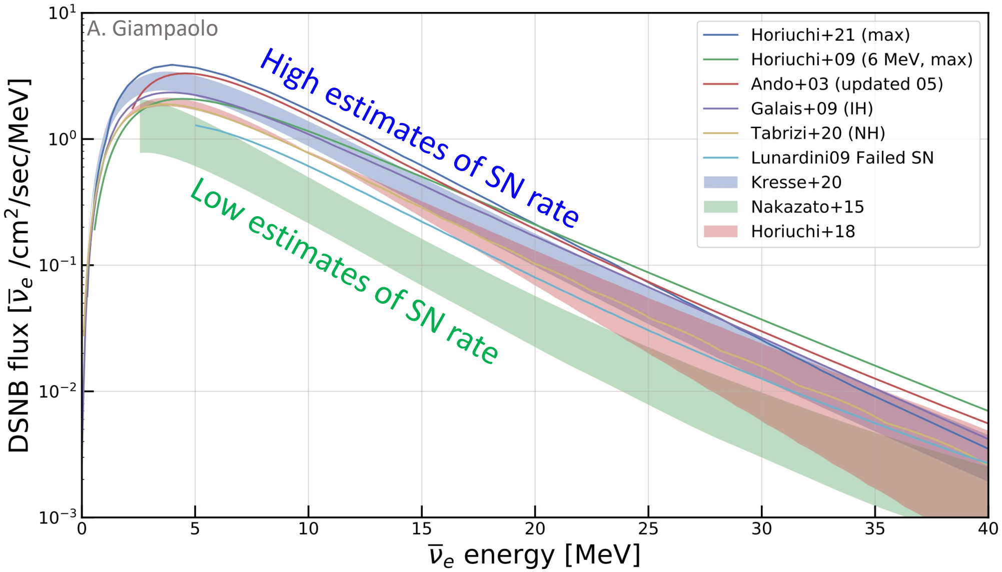
TAUP VIENNA 2023

SUPER K

The Diffuse Supernova Neutrino Background

J.Beacom, TAUP 2011





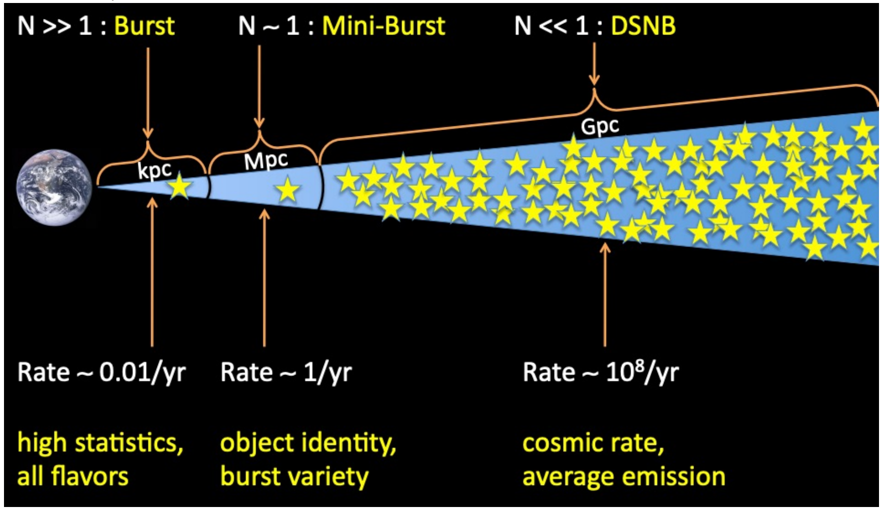
- Integrated neutrino flux from past core-collapse supernovae
- Detectable range within redshift $z \approx 1 - 2$
- Supernova physics, star formation and neutrino properties

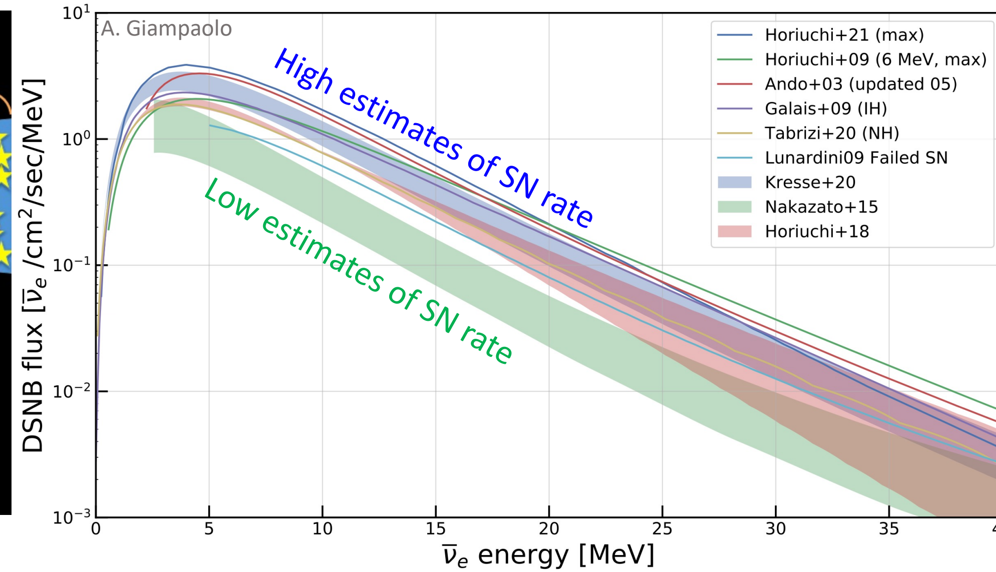
$$\Phi_{DSNB}(E,z) = \frac{c}{H_0}\int_0^{z_{max}} R_{SN}(z)\, F_\nu\,[E(1+z)]\,\frac{dz}{\sqrt{\Omega_M(1+z)^3+\Omega_\Lambda}}$$

- Elusive low energy signal!

The Diffuse Supernova Neutrino Background

J.Beacom, TAUP 2011



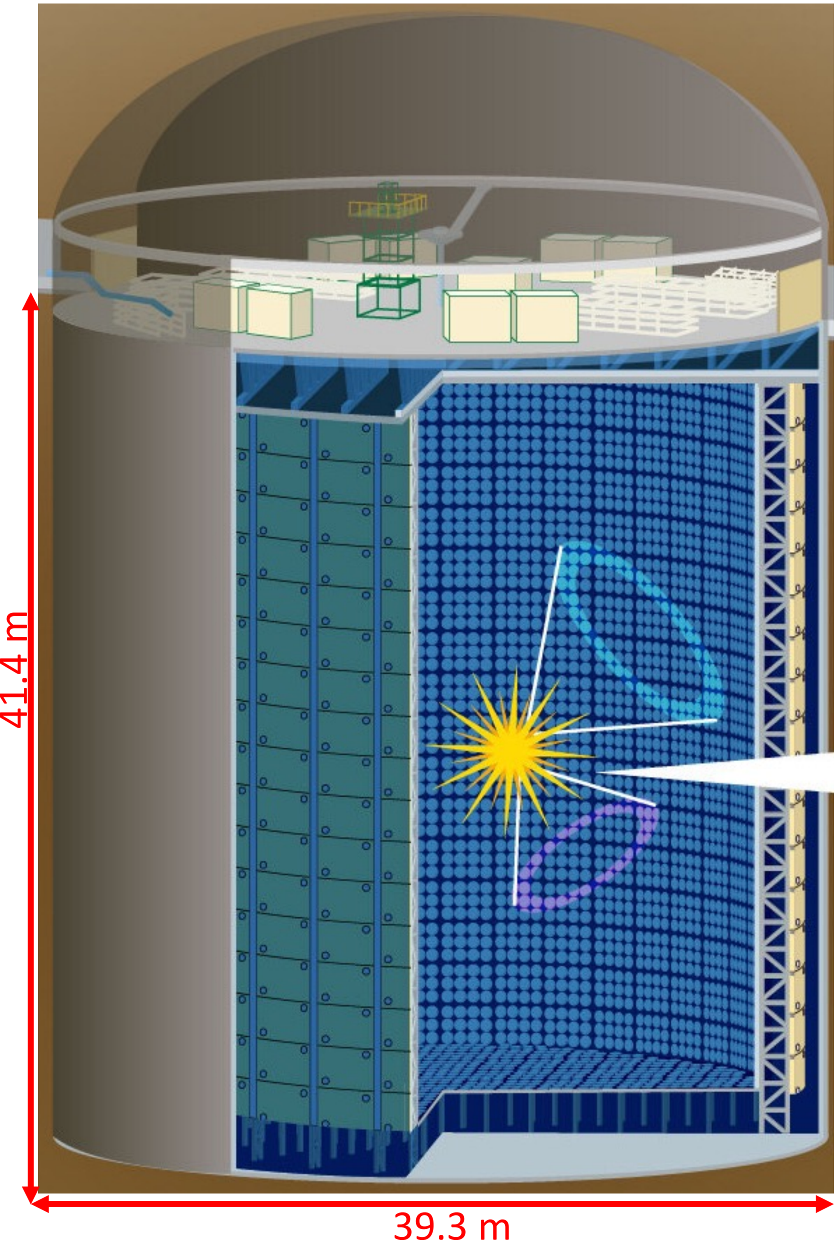


- Integrated neutrino flux from past core-collapse supernovae
- Detectable range within redshift $z \approx 1 - 2$
- Supernova physics, star formation and neutrino properties

$$\Phi_{DSNB} \propto \int \left[\text{SN Rate}\right] \otimes \left[\nu \text{ emission}\right] \otimes \left[\text{Cosmic Expansion}\right]$$

- Elusive low energy signal!

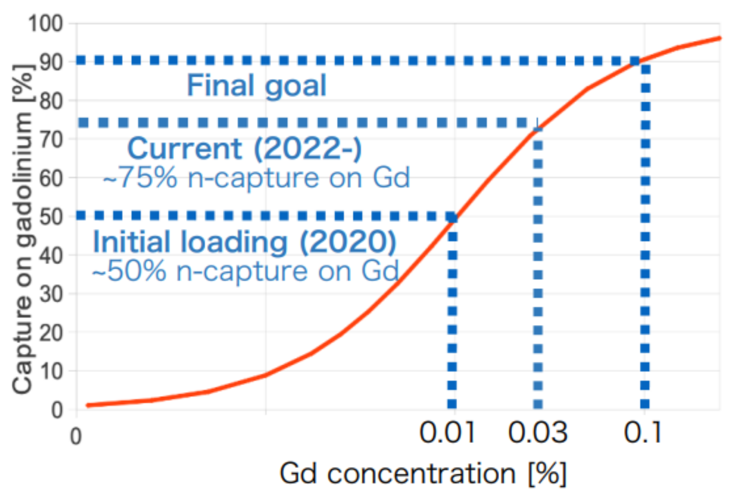
Super Kamiokande with Gd



- Located in Kamioka mine, Japan
- 50 kton water Cherenkov detector
- 22.5 kton fiducial volume
- **Inner Detector (ID)**: 11,129 50 cm PMTs
- **Outer Detector (OD)**: 1,885 20 cm PMTs + Wavelength Shifting Plates
- Low energy threshold ~ 4 MeV
- Neutron tagging via Gd loading

Enhanced capability of DSNB detection!

Neutron capture efficiency on Gd

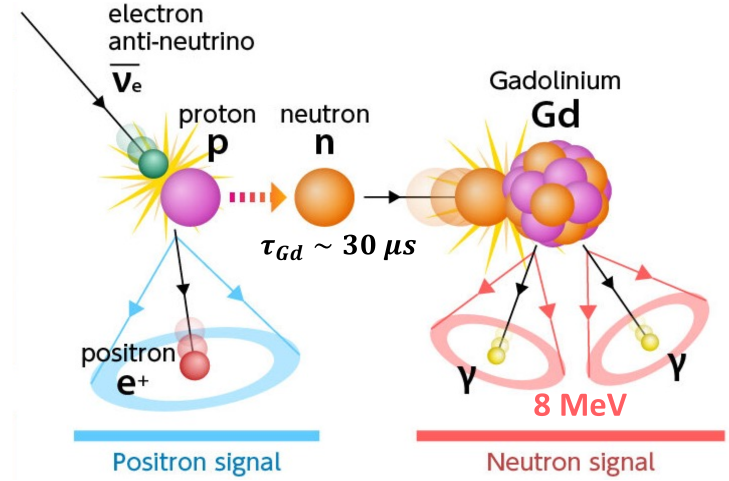


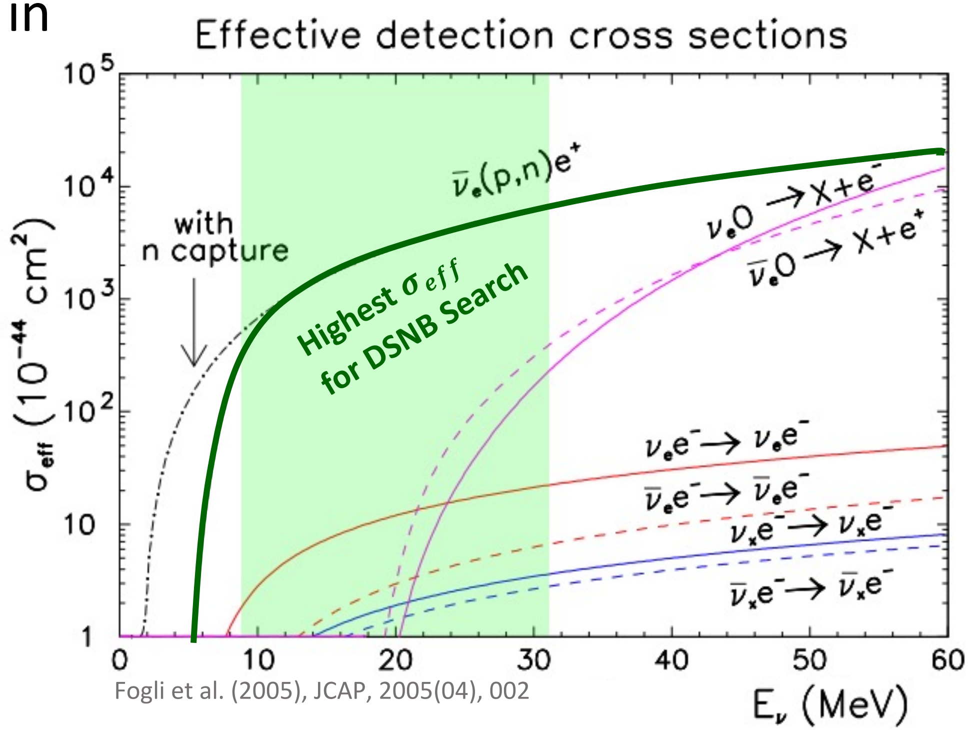
DSNB Detection in Super-Kamiokande

- **Inverse Beta Decay ($\beta + n$):**

$$\bar{\nu}_e + p \rightarrow e^+ + n$$

- **Largest interaction cross-section** for all other detection modes (2x)
- **8 – 30 MeV** DSNB Search Window in reconstructed visible energy





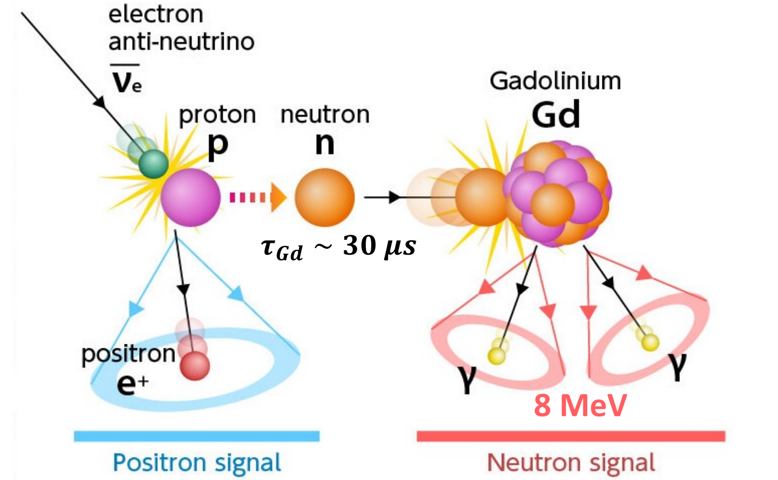
Fogli et al. (2005), JCAP, 2005(04), 002

DSNB Detection in Super-Kamiokande

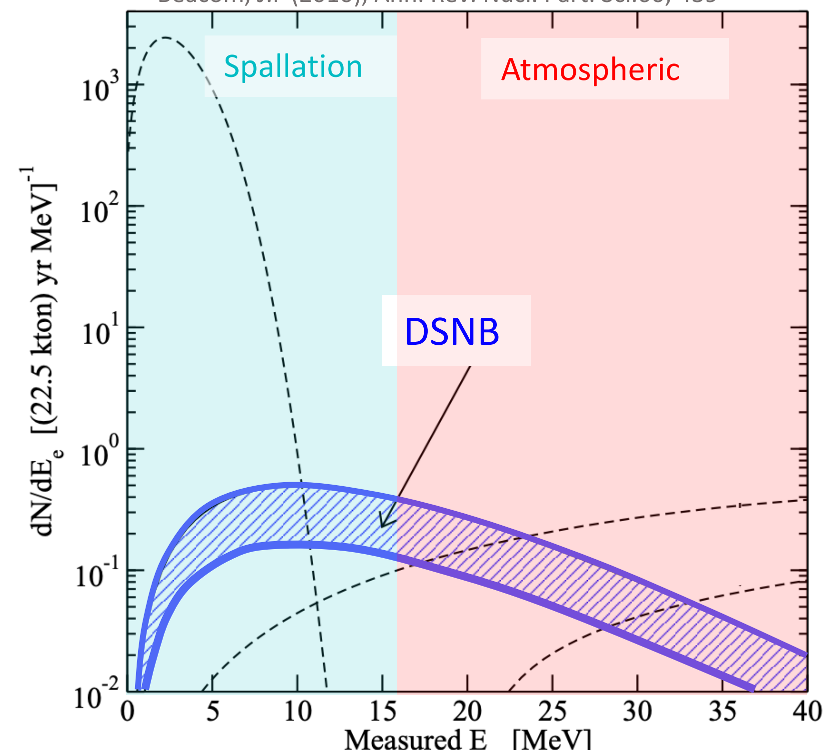
- **Inverse Beta Decay ($\beta + n$):**

$$\bar{\nu}_e + p \rightarrow e^+ + n$$

- **Largest interaction cross-section** for all other detection modes (2x)
- **8 – 30 MeV** DSNB Search Window in reconstructed visible energy
- **Few events** per year expected in the search window
- Dominated by backgrounds:
  - **Muon Spallation**
  - **Atmospheric Neutrinos**
- Require thorough background characterisation

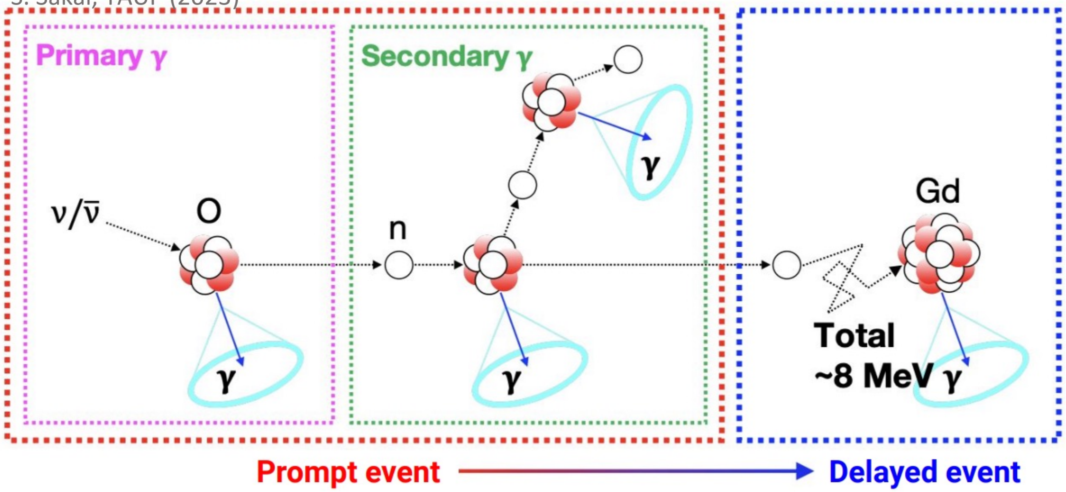


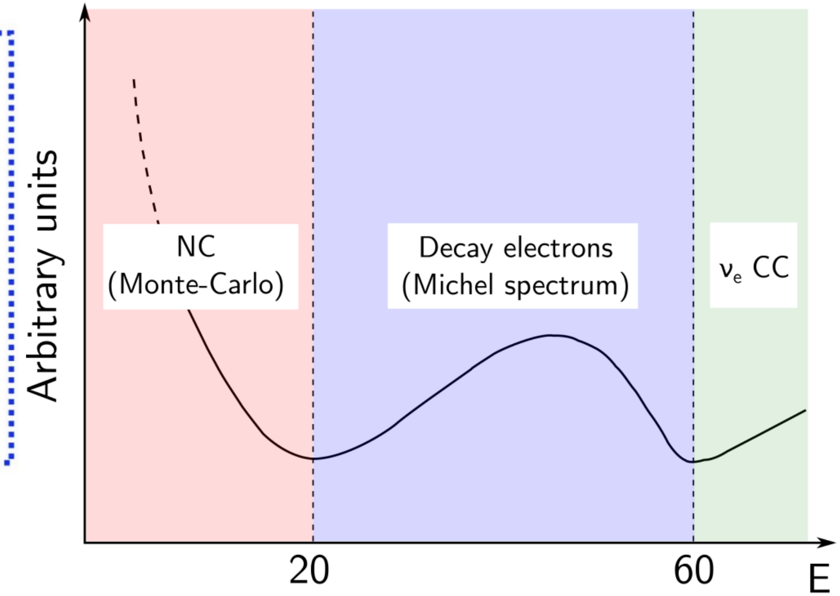
Beacom, J.F (2010), Ann. Rev. Nucl. Part. Sci.60, 439



NCQE Atmospheric Neutrino Interactions

S. Sakai, TAUP (2023)





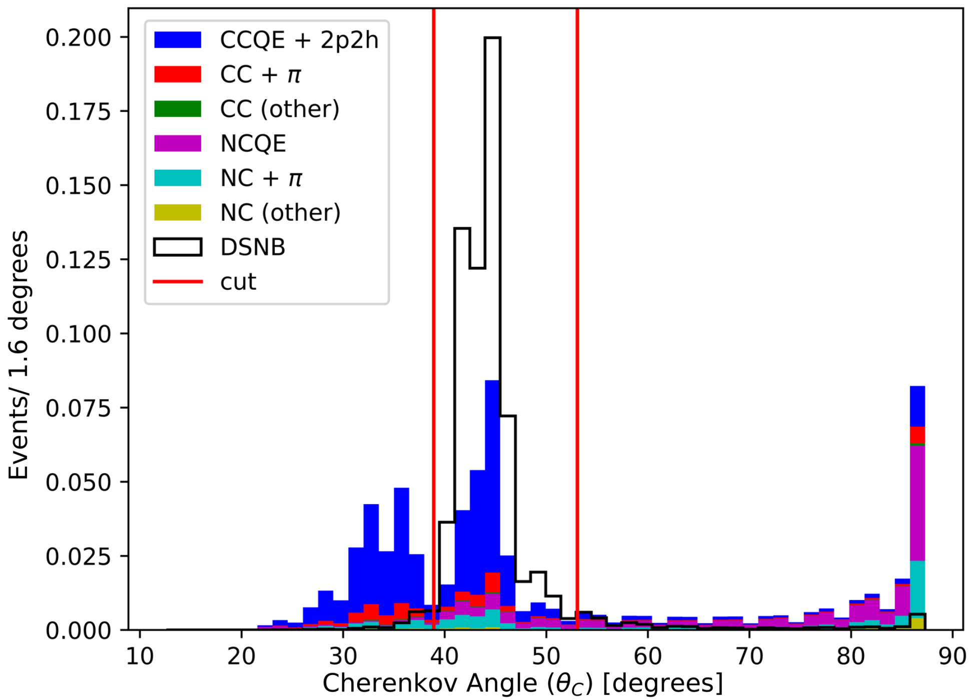
- Neutral Current Quasi-Elastic (**NCQE**) interactions:

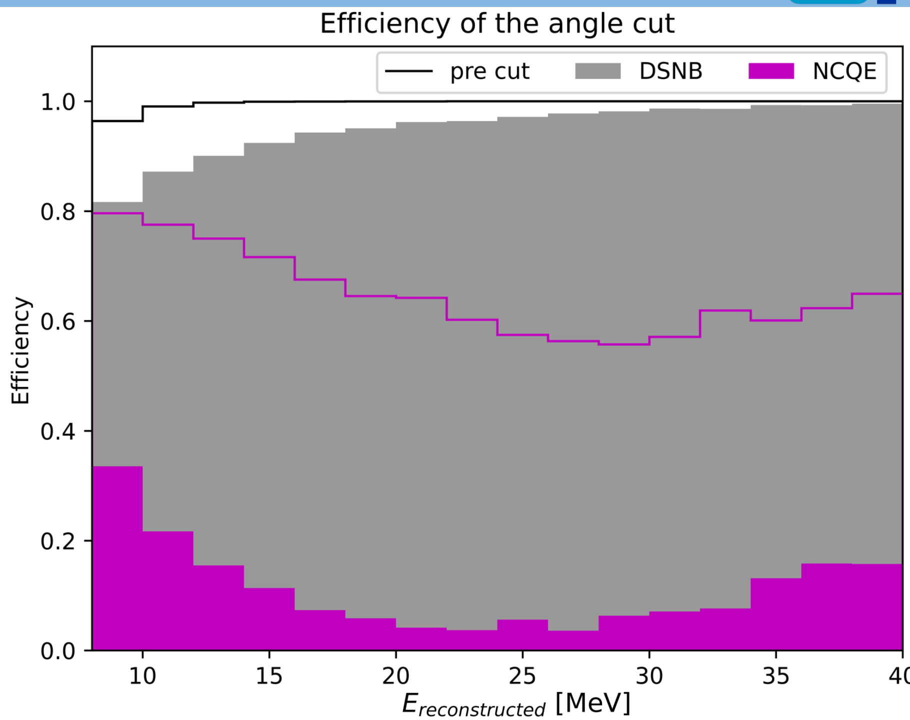
$$\nu(\bar{\nu}) + O^{16} \rightarrow \nu(\bar{\nu}) + \mathrm{n} + O^{15*}$$

$$\nu(\bar{\nu}) + O^{16} \rightarrow \nu(\bar{\nu}) + \mathrm{p} + N^{15*}$$

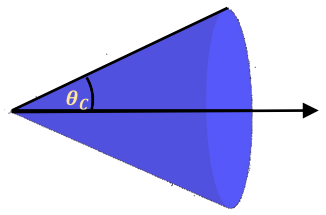
- De-excitation **primary γ-ray + neutron** (mimics IBD signal)
- Multiple **secondary γ-rays** from nucleon interactions in prompt window
- Dominant **below 20 MeV**
- Currently a **residual background** in the DSNB search

Reducing the NCQE Background



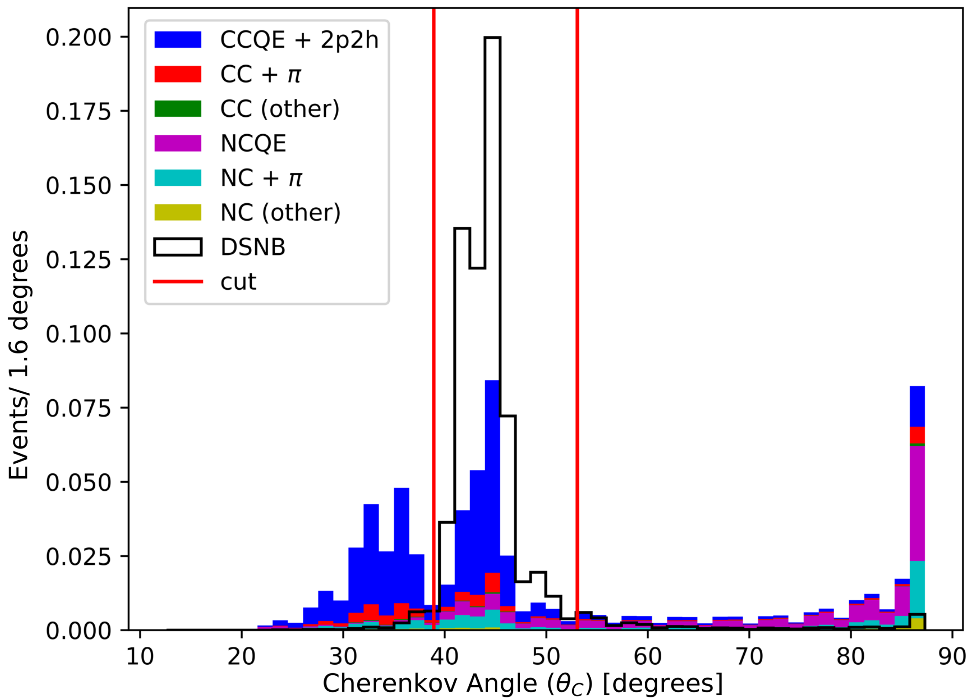


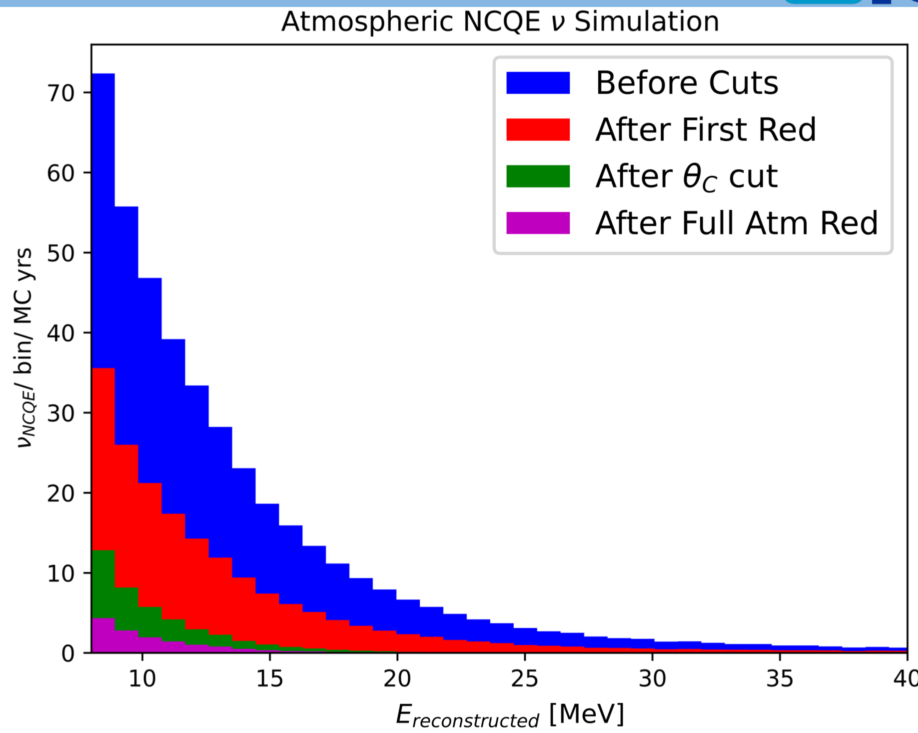
- Powerful discriminator based on the **Cherenkov opening angle** $(\boldsymbol{\theta_C})$
  - $\boldsymbol{\theta_C \sim 42^\circ}$ for IBD interactions (**DSNB**)
  - $\boldsymbol{\theta_C}$ **is large** for **NCQE** events due to multiple secondary γ-rays



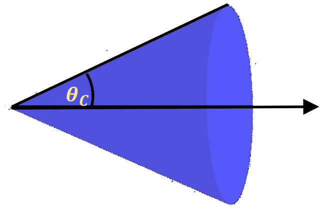
- DSNB search window (current analysis cut) $\boldsymbol{\theta_C}$: $[38^\circ, 53^\circ]$
- Large residual NCQE contribution after angle cut applied – challenging background requires new reduction techniques

Reducing the NCQE Background





- Powerful discriminator based on the **Cherenkov opening angle** $(\boldsymbol{\theta_C})$
  - $\boldsymbol{\theta_C \sim 42^{\circ}}$ for IBD interactions (**DSNB**)
  - $\boldsymbol{\theta_C}$ **is large** for **NCQE** events due to multiple secondary γ-rays



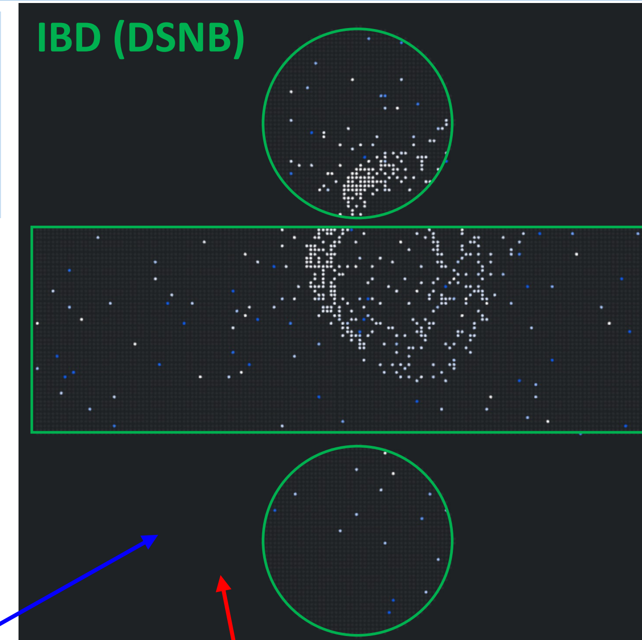
- DSNB search window (current analysis cut) $\boldsymbol{\theta_C}$: $[38^{\circ}, 53^{\circ}]$
- Large residual NCQE contribution after angle cut applied – challenging background requires new reduction techniques, **such as Convolutional Neural Networks!**

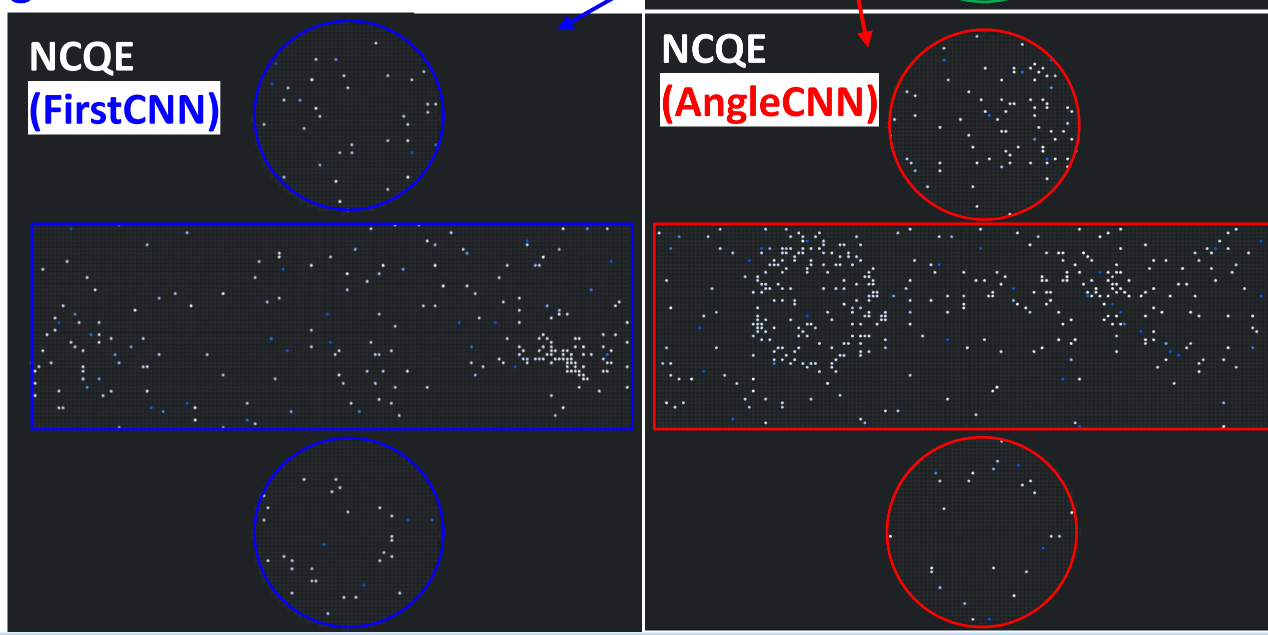
CNN Models: Event Selection for Training

Convolutional Neural Networks (CNNs) trained **with**/**without** **the θ_C cut** to investigate a new method for targeted NCQE reduction.

FirstCNN Model:

- First reduction of the DSNB analysis applied to IBD and NCQE Super-K (Gd) simulation
- This includes **only energy window, fiducial volume and noise cuts** to remove obvious and mis-reconstructed backgrounds
- Select events with **one tagged neutron**

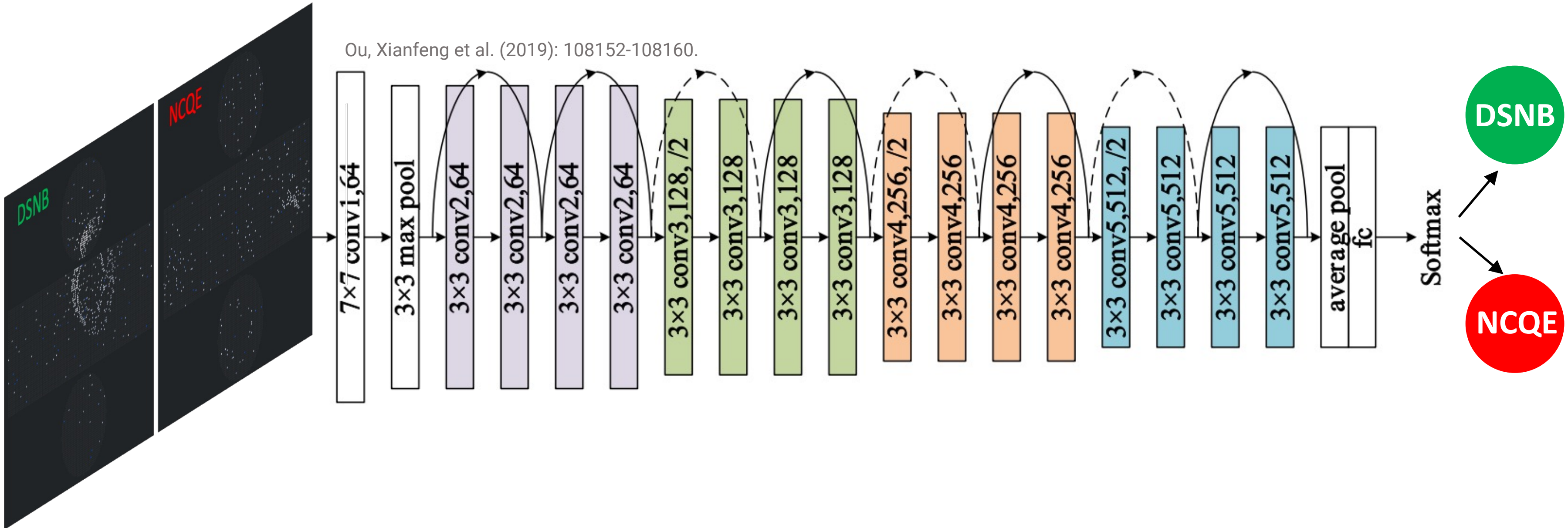




AngleCNN Model:

- Used FirstCNN event selection, then applied **the Cherenkov angle cut**

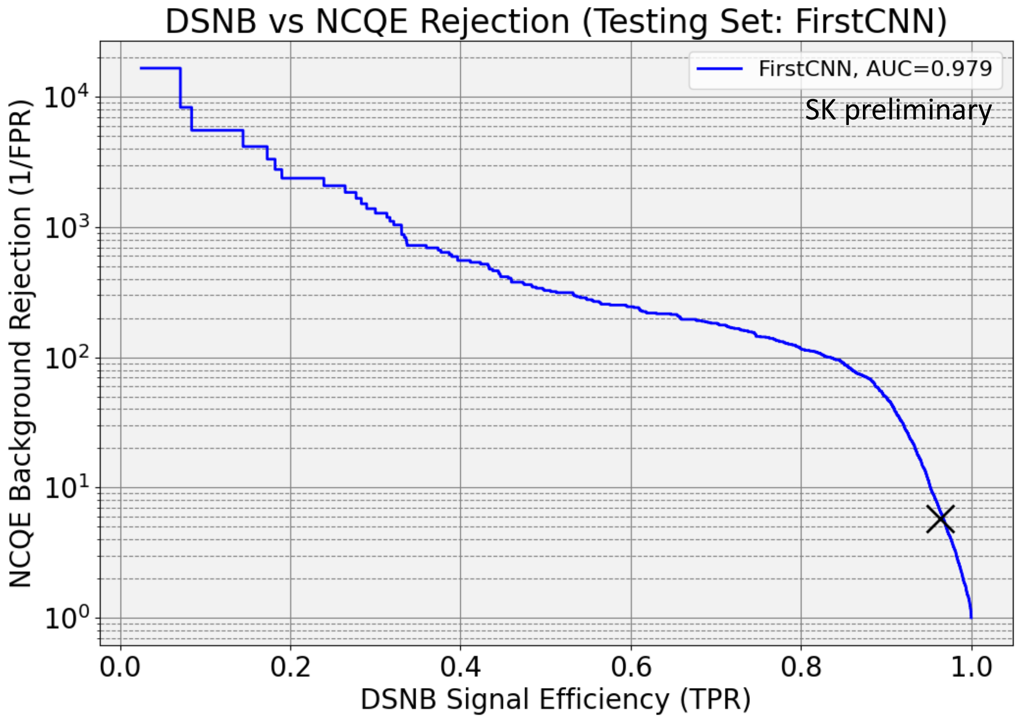
Convolutional Neural Networks

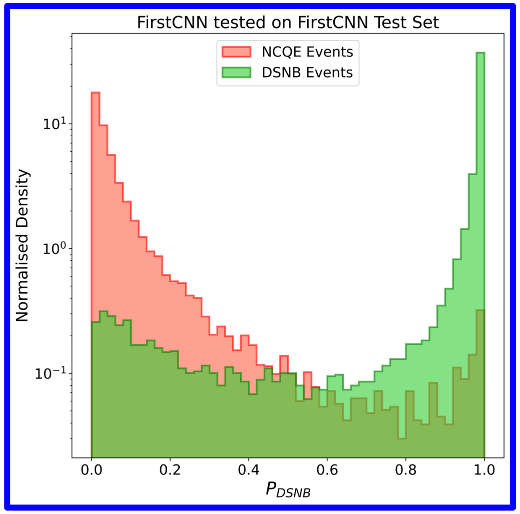


- **ResNet-18 Convolutional Neural Network (CNN):** 18 layers with residual connections to enable faster training for deeper networks.
- Using the Water Cherenkov Machine Learning (**WatChMaL**) framework 
- **Input:**
  - Standard SuperK event display (timing and charge) to 2D image mapping
  - Prompt event only for e/ γ classification
- **Training:**
  - Optimised hyper-parameters for training (see backup)

Model Performance on FirstCNN Test Set

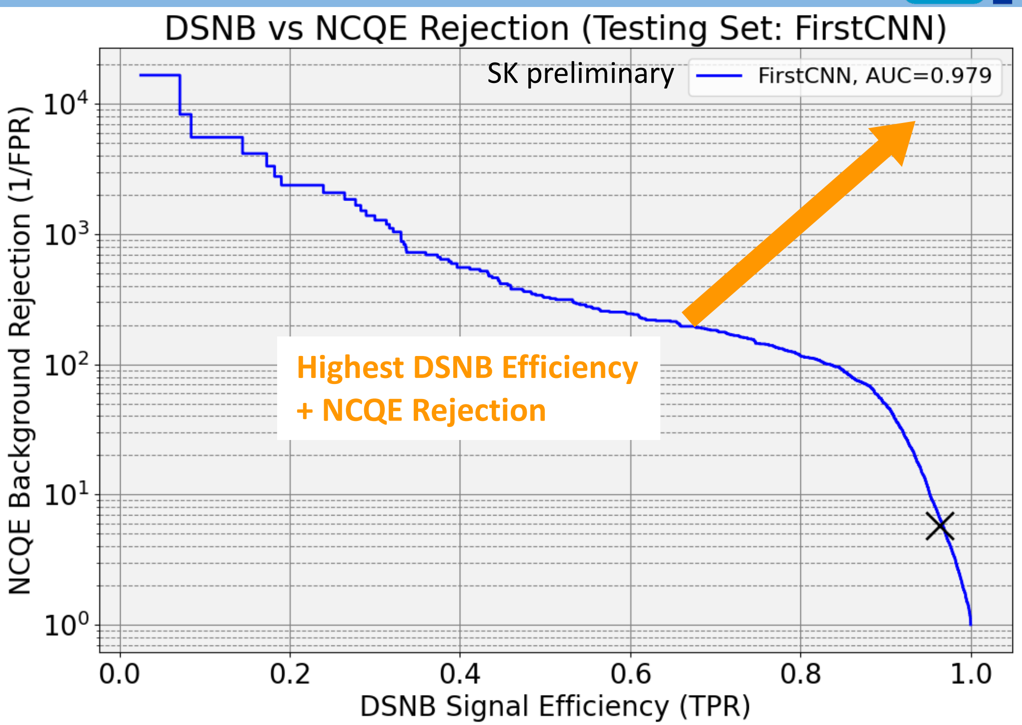
- **FirstCNN – Clear Backgrounds:**
  - Test set includes many NCQE events with ambiguous Cherenkov rings
  - Classifies apparent NCQE events that lack distinctive Cherenkov ring to enhance signal purity

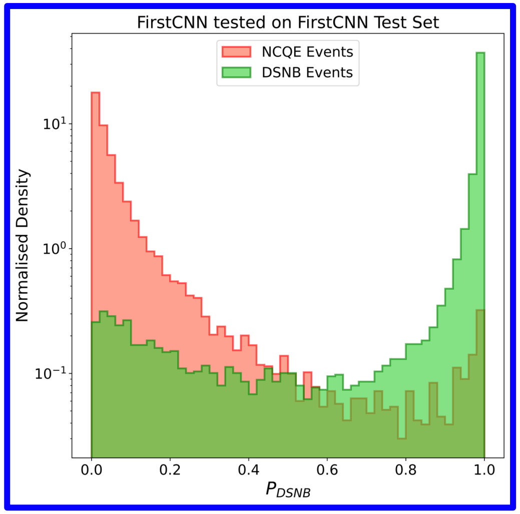




Model Performance on FirstCNN Test Set

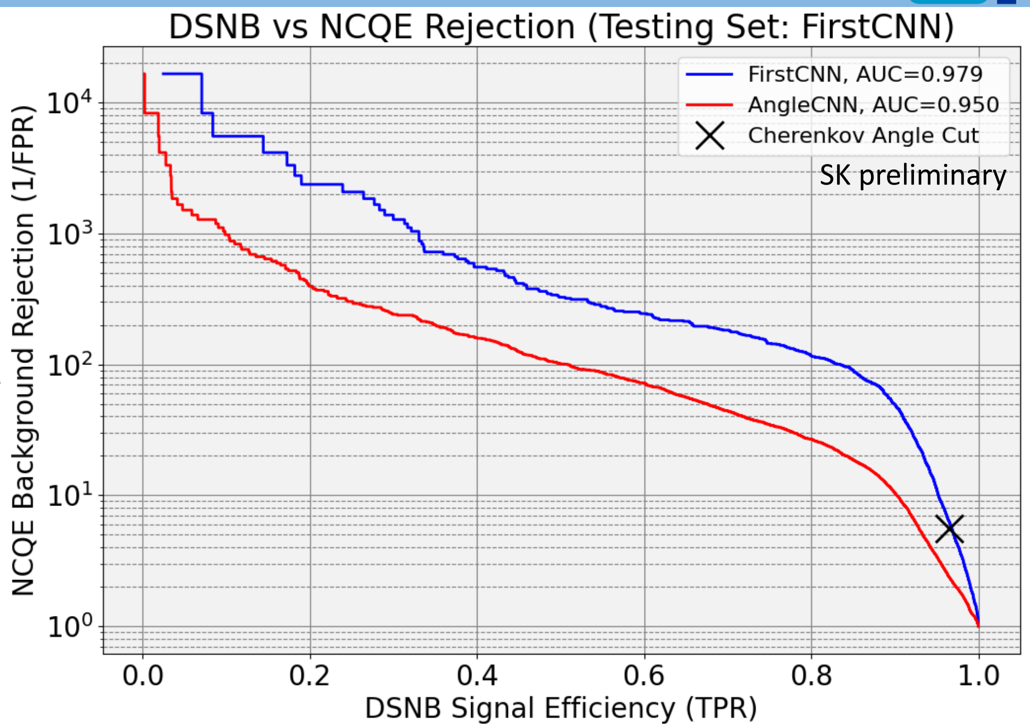
- **FirstCNN – Clear Backgrounds:**
  - Test set includes many NCQE events with ambiguous Cherenkov rings
  - Classifies apparent NCQE events that lack distinctive Cherenkov ring to enhance signal purity

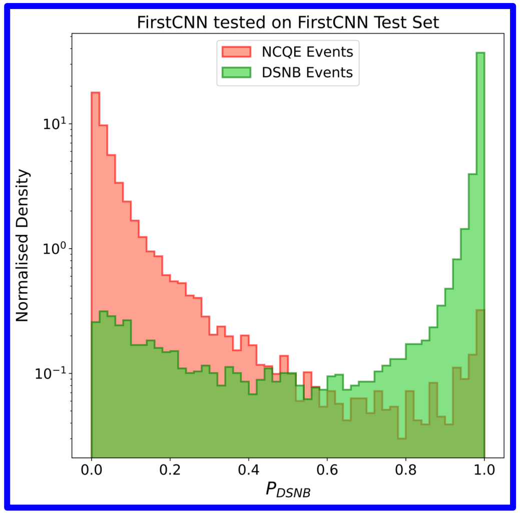


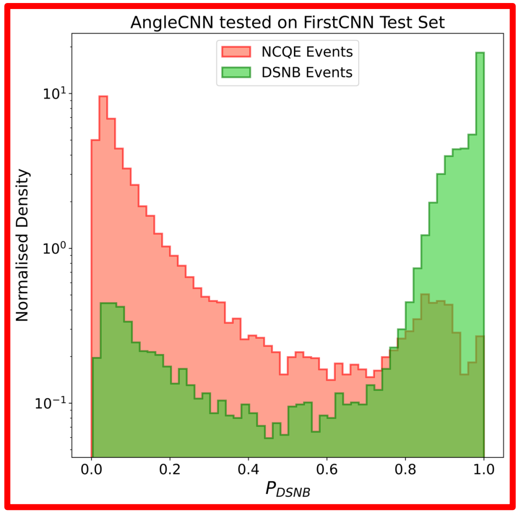


Model Performance on FirstCNN Test Set

- **FirstCNN – Clear Backgrounds:**
  - Test set includes many NCQE events with ambiguous Cherenkov rings
  - Classifies apparent NCQE events that lack distinctive Cherenkov ring to enhance signal purity
- **AngleCNN – Enhanced Reduction:**
  - Focused learning on subtle topological features of e and γ – induced Cherenkov ring
  - Distinguishes events with similar hit distributions, with Cherenkov ring characteristics

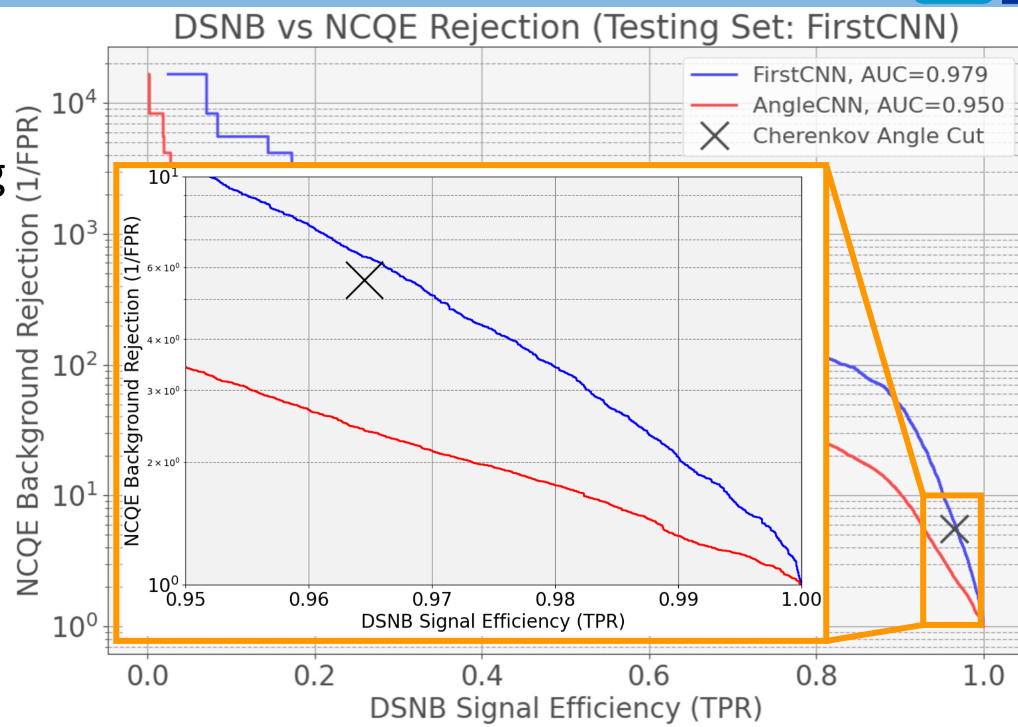


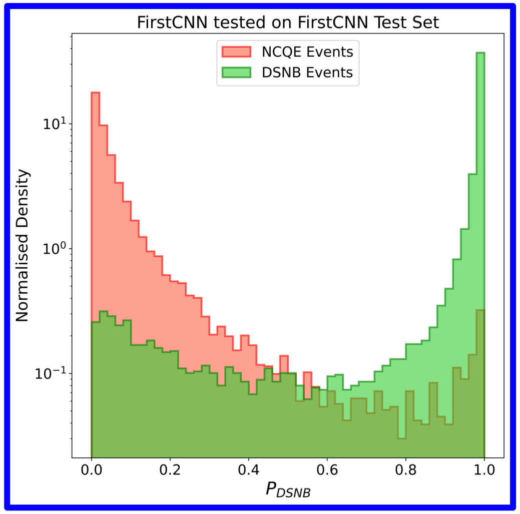


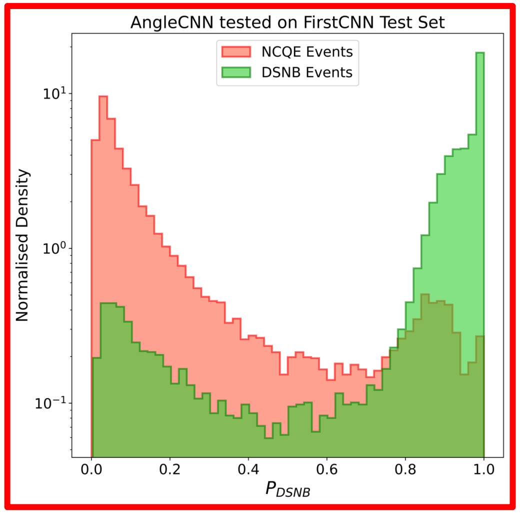


Model Performance on FirstCNN Test Set

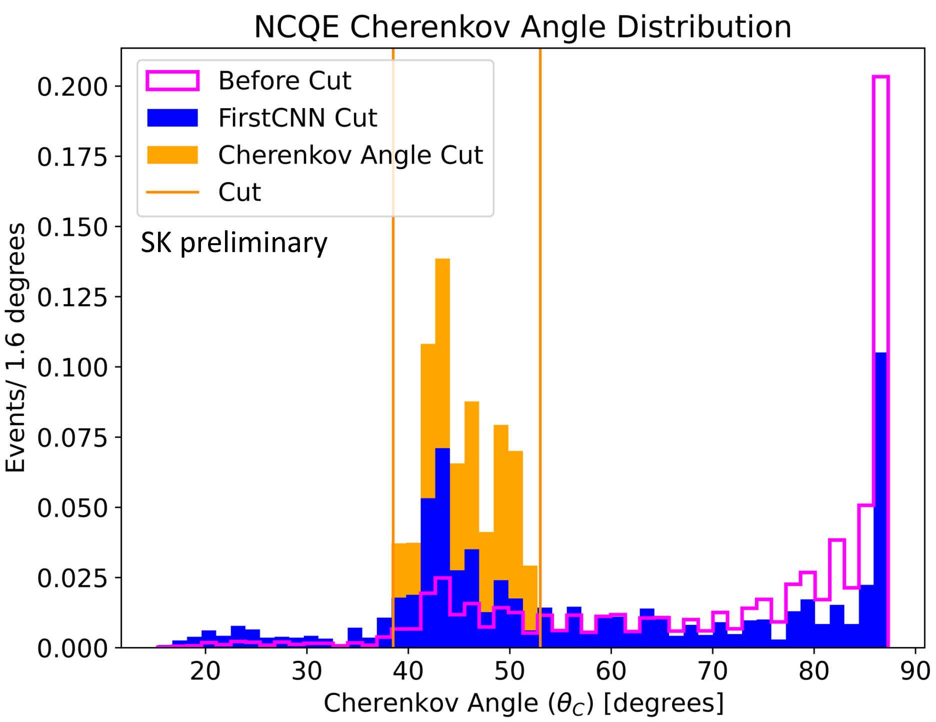
- **FirstCNN – Clear Backgrounds:**
  - Classifies apparent NCQE events that lack distinctive Cherenkov ring to enhance signal purity
  - Test set includes many NCQE events with ambiguous Cherenkov rings
- **AngleCNN – Enhanced Reduction:**
  - Focused learning on subtle topological features of e and γ – induced Cherenkov ring
  - Distinguishes events with similar hit distributions, with Cherenkov ring characteristics
- **FirstCNN > Angle Cut**
  - Marginal improvement at signal efficiency of Cherenkov angle cut

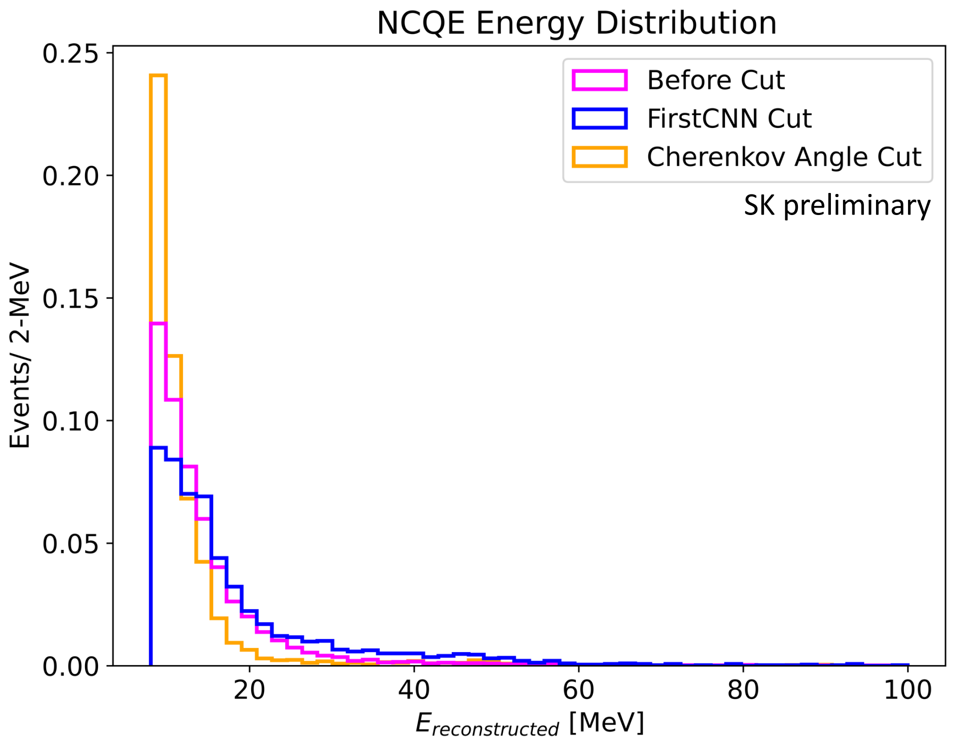






Understanding CNN Performance





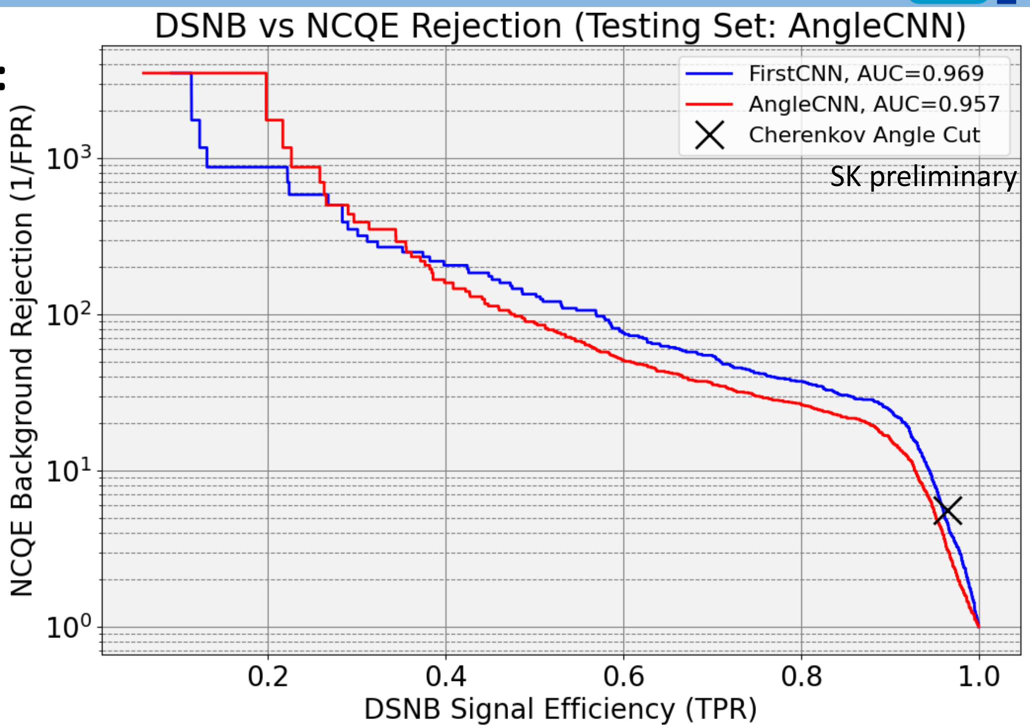
- **NCQE Angle and Energy distributions – Before any cuts**
- **Cherenkov Angle Cut:**
  - DSNB search window (current analysis cut) $\boldsymbol{\theta_C : [38^{\circ}, 53^{\circ}]}$
  - **Residual background at low energy**

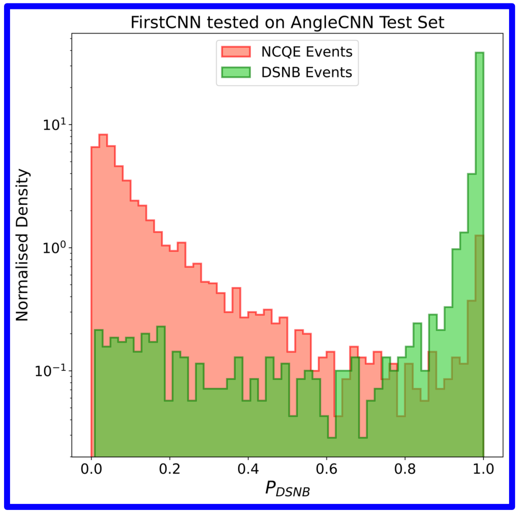
Evaluate at the Cherenkov Angle Signal Efficiency ($\sim$96%)

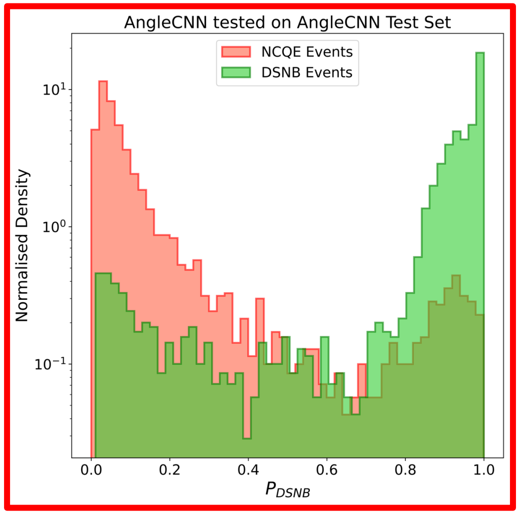
- **FirstCNN Cut:**
  - Applies an energy dependent cut – inferred from PMT hit distribution
  - **Combination of CNN and optimised Cherenkov Angle cut - effective reduction**

Model Performance on AngleCNN Test Set

- **FirstCNN outperforms AngleCNN:**
  - **Effective features**: Captures crucial event features of Cherenkov ring
  - **Generalisation**: Avoids overfitting to generalise well to unseen data
  - **Training Data**: Benefits from training data that includes both ambiguous and clear Cherenkov rings for NCQE events
  - Performance declines at high NCQE rejection
  - AngleCNN effective at discriminating challenging NCQE events at low signal efficiency





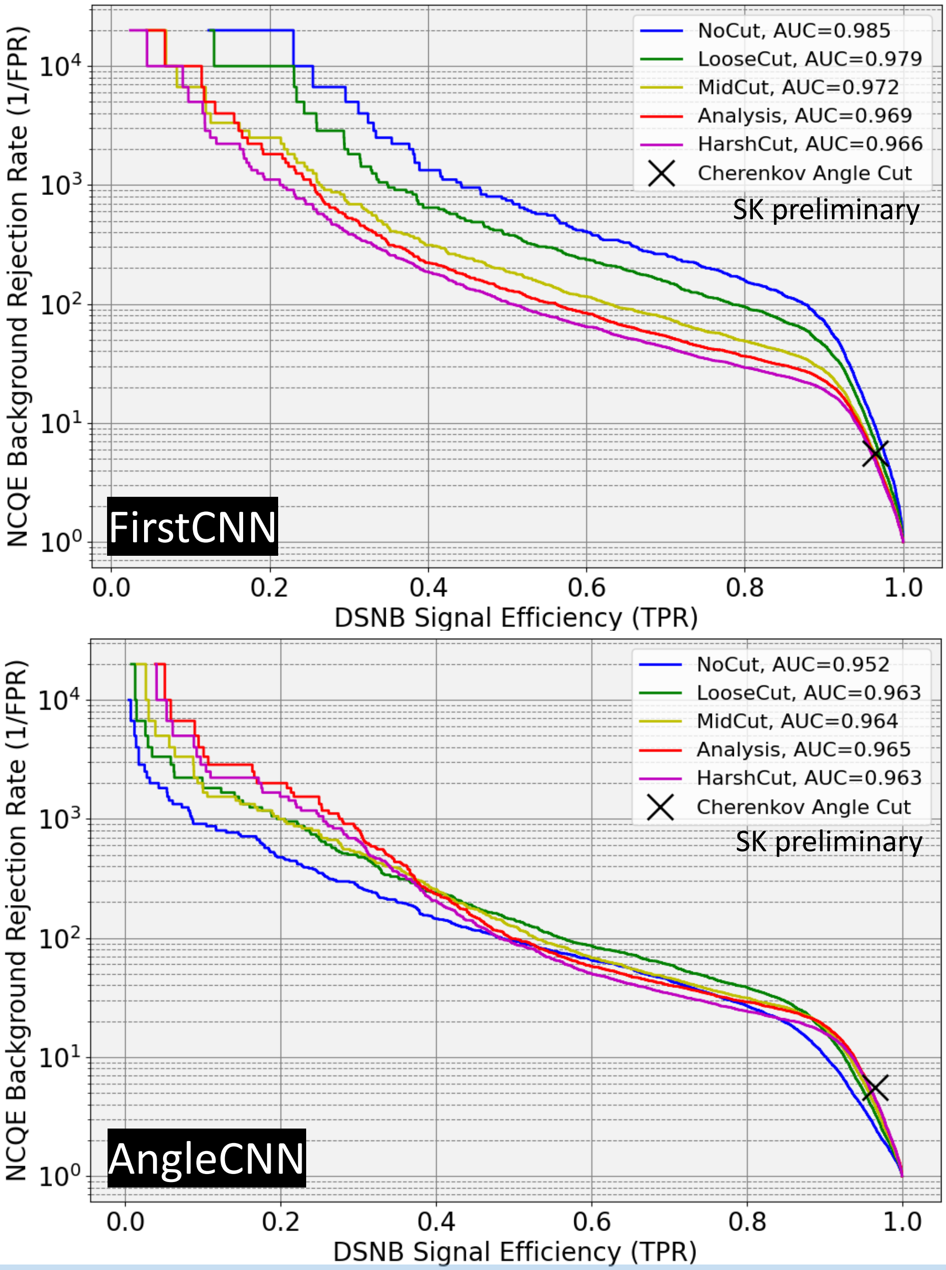


Robustness Analysis

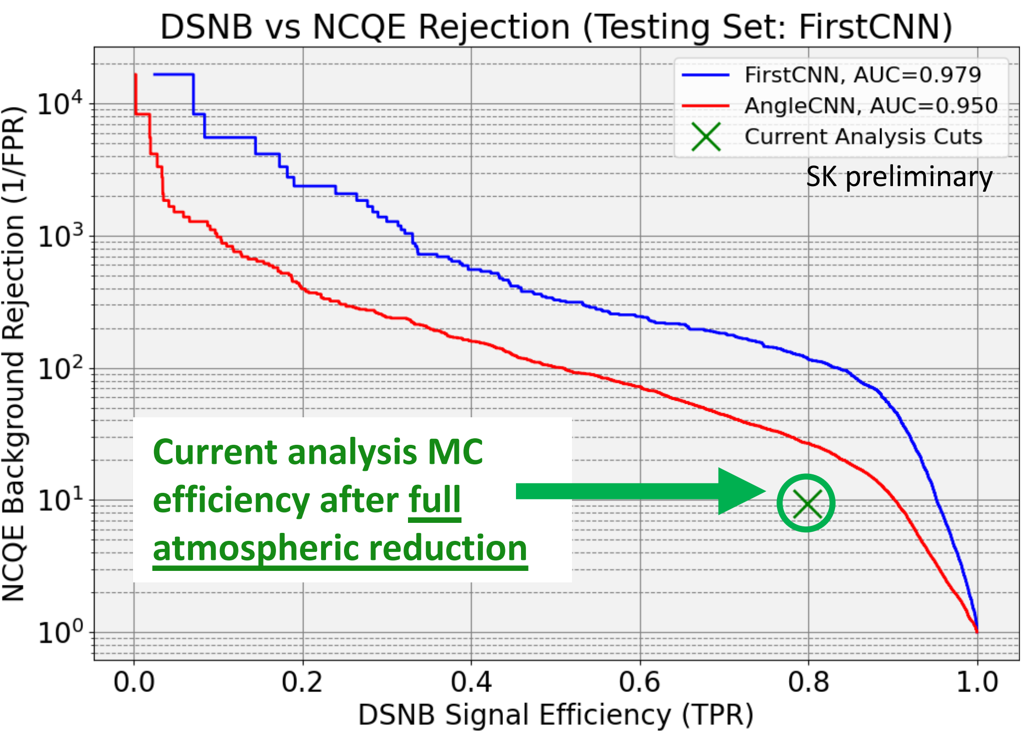
- **Testing on different angle cuts (θ_C)**
  - **NoCut⇔FirstCNN** test set $[0^\circ, 90^\circ]$
  - **LooseCut**: $[30^\circ, 80^\circ]$
  - **MidCut**: $[35^\circ, 60^\circ]$
  - **Analysis⇔AngleCNN** test set $[38^\circ, 53^\circ]$
  - **HarshCut**: $[40^\circ, 50^\circ]$

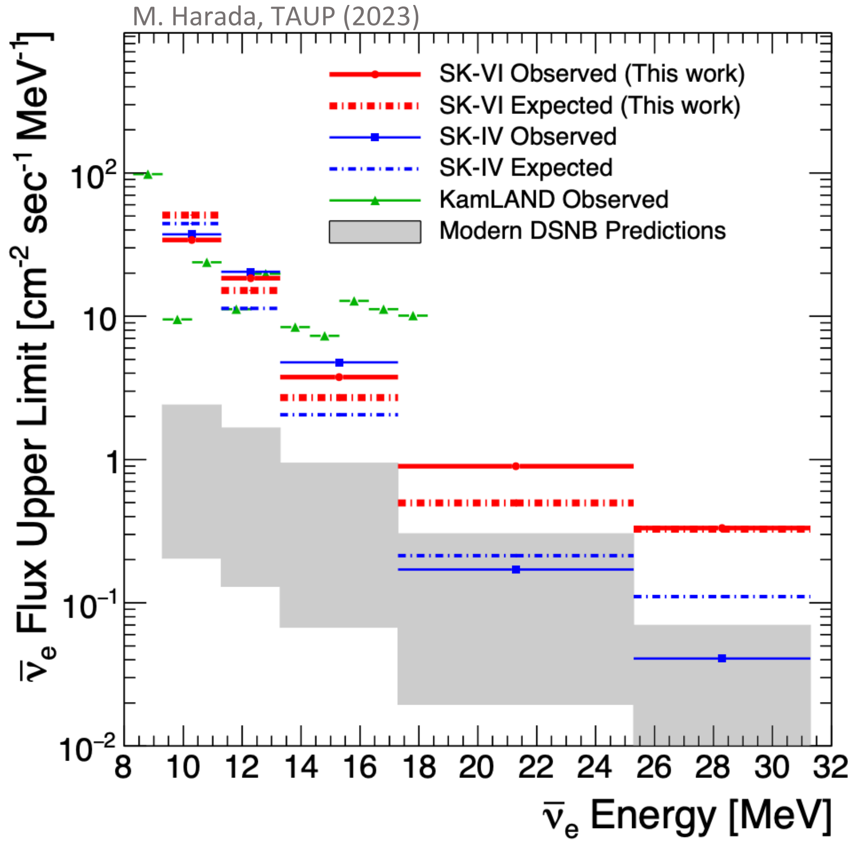
- **FirstCNN's Performance Decline:** Performance drops with more stringent cuts. Still outperforms AngleCNN at DSNB traditional analysis efficiency

- **AngleCNN's Robust Adaptability:** Maintains performance with stricter angle cut due to specialised training



DSNB Current Upper Limits





- Develop an effective reduction method to **enhance signal and background efficiencies beyond current level**
- Current **best upper limits achieved in SK-VI** (see talk by M.Harada, Thursday August 31st)

Summary

- **Dominant contribution of NCQE Background** in DSNB Search Window
- Current best handle is cut on **reconstructed Cherenkov angle**
- **Residual background at low energy** – in region of the highest flux prediction of DSNB models
- Investigated **new reduction techniques using CNNs** – ResNet-18 models trained on samples with/without Cherenkov cut applied
- **FirstCNN moderately improves reduction** at Cherenkov angle cut efficiency
- Combination of **CNN cut + optimised angle cut** will likely be an effective reduction method

Next Steps

- Explore deeper ML methods – **Graphical Neural Networks**
- **Integrate novel reduction approach into full energy bin-by-bin analysis**
- Aim to **improve DSNB upper limit constraints**

Thank You!

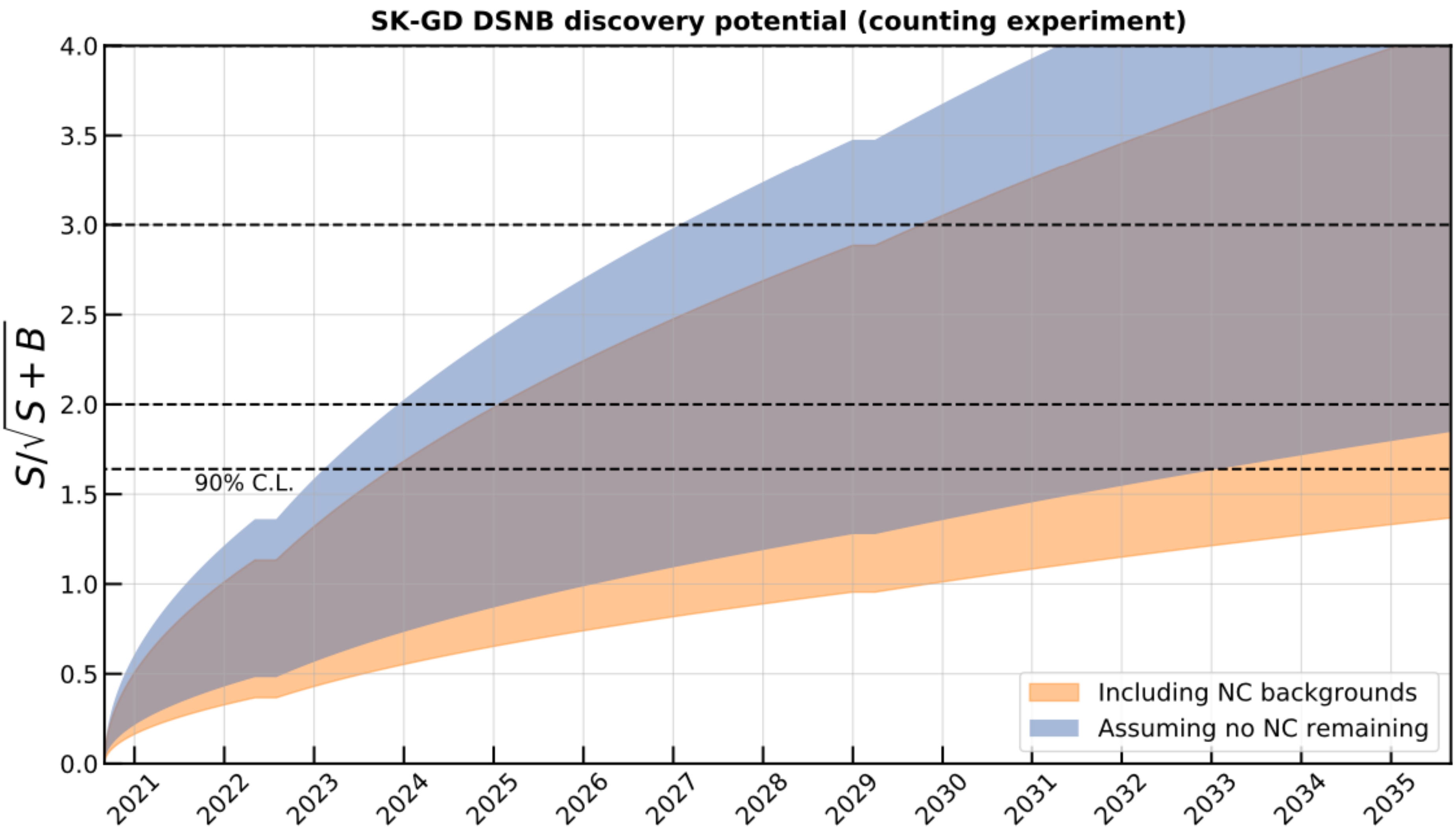
UNIVERSITY OF OXFORD

SUPER SK

TAUP VIENNA 2023

BACKUP – DISCOVERY POTENTIAL

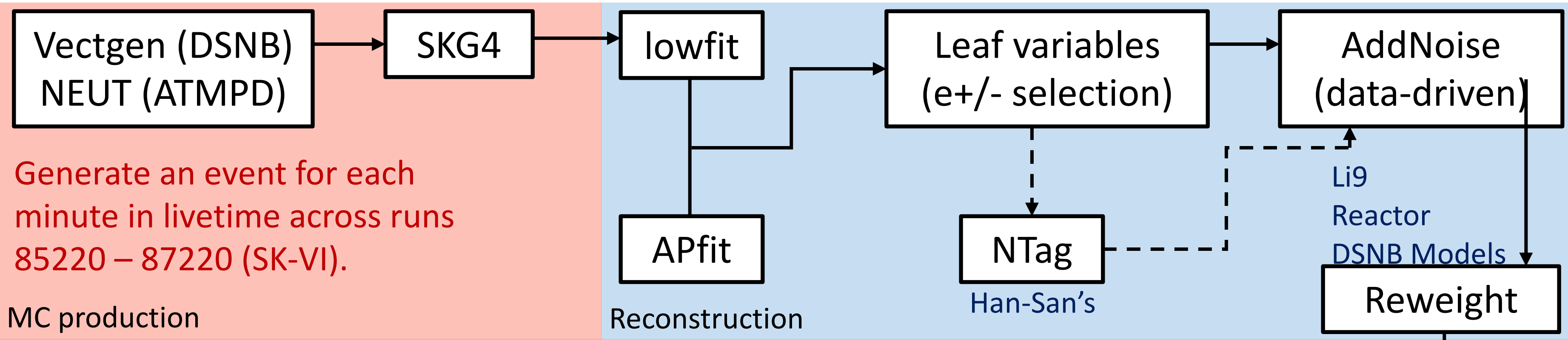
DSNB Discovery Potential



Study done by A. Giampaolo

BACKUP – TRADITIONAL ANALYSIS

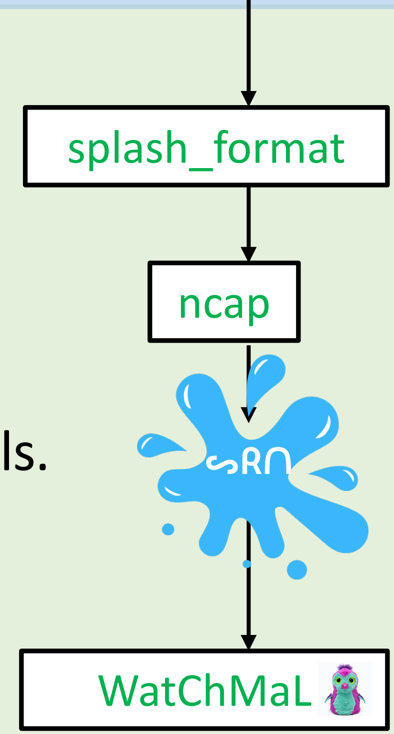
IBD and Atmospheric MC Generation



Analysis:

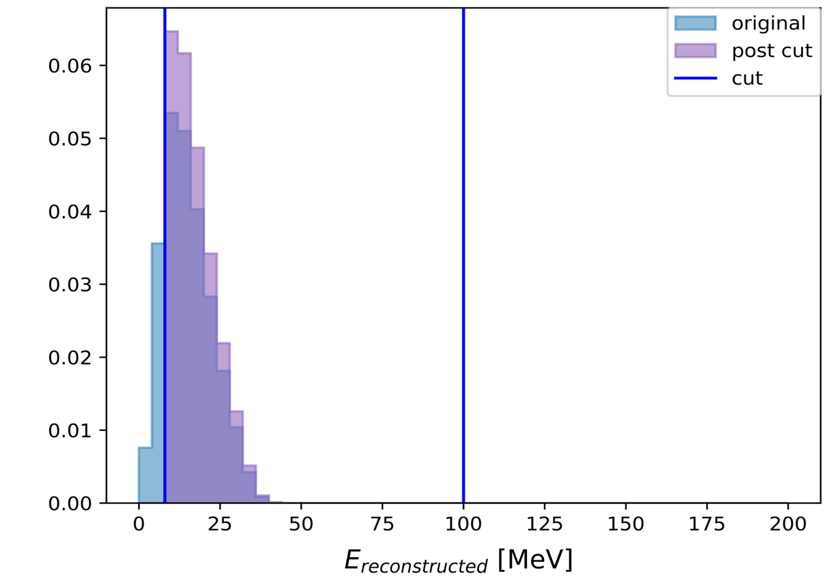
- DSNB MC: 522 livetime days of full SK-VI duration
- ATM MC: 500 years of atmospheric neutrino events, with an additional 5000 years equivalent NCQE MC.

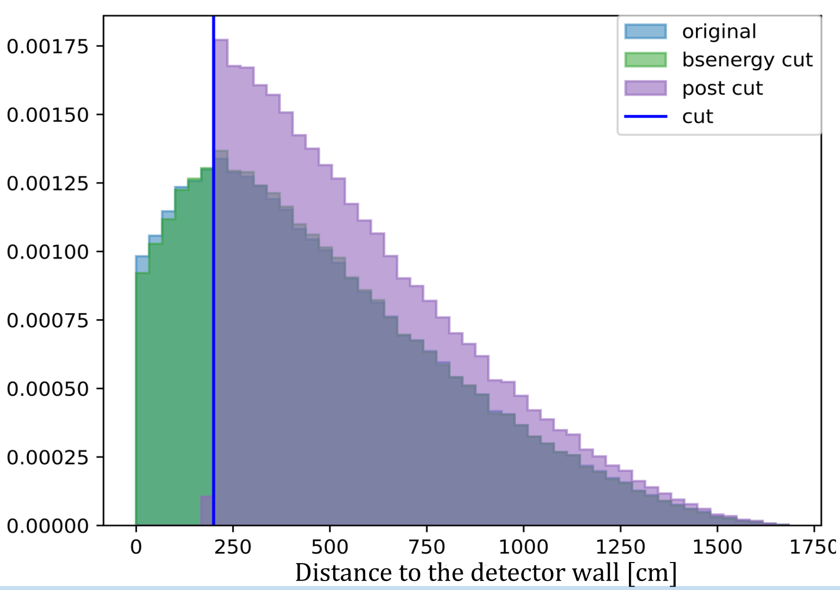
1. Atmospheric modelled with HKKM-2011 flux with NEUT.
2. Kresse spectrum for IBD MC – reweighting to DSNB models.
3. Low energy reconstruction
4. Inject data-driven noise
5. Apply neutron tagging
6. Event selection – efficiency calculations
7. Format for CNN input



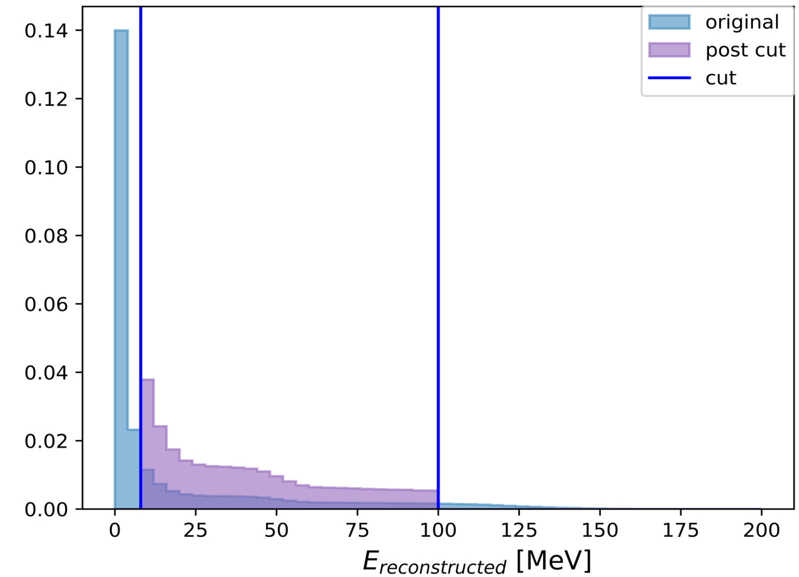
Event Selection – First Reduction

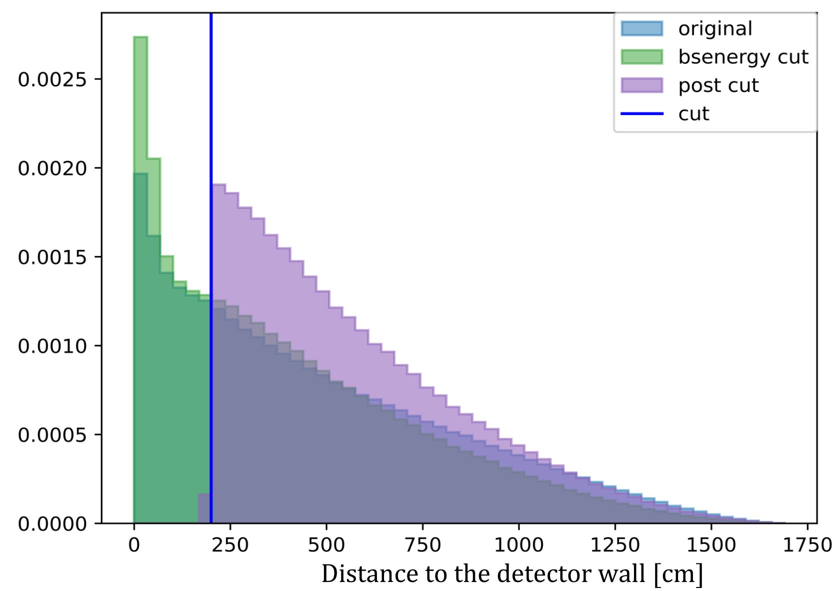
DSNB (Horiuchi)





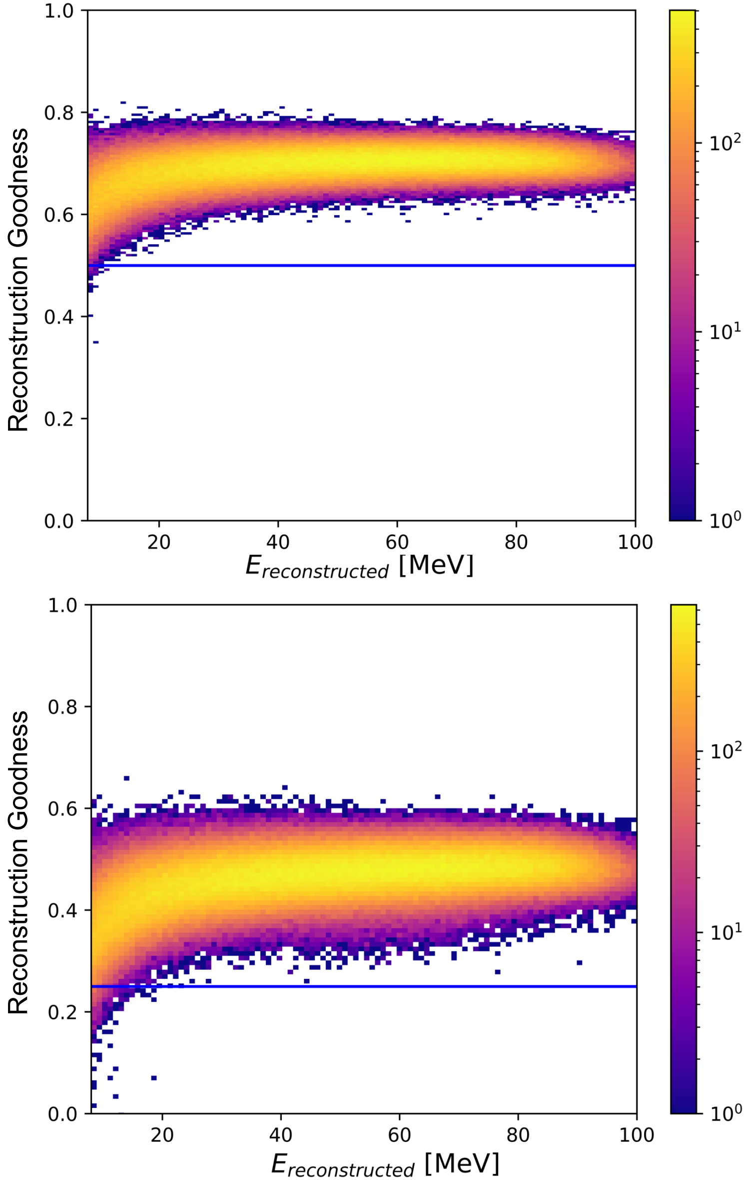
Atmospheric Background



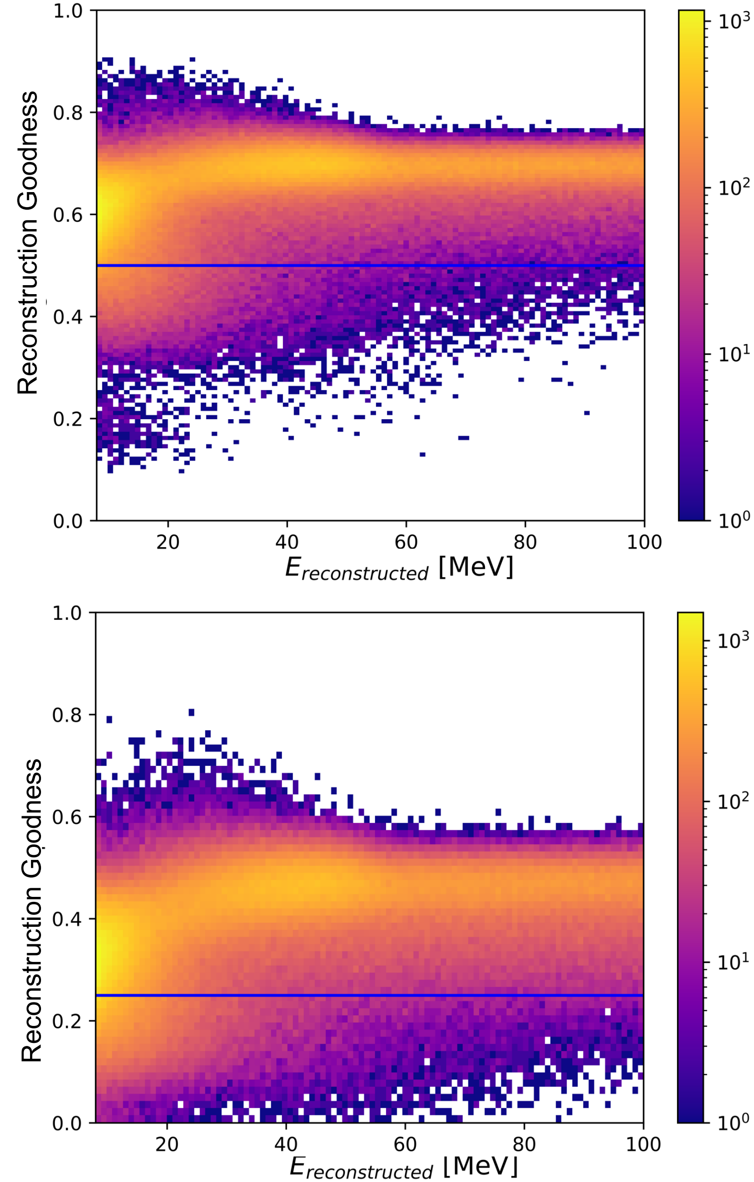


Event Selection – First Reduction

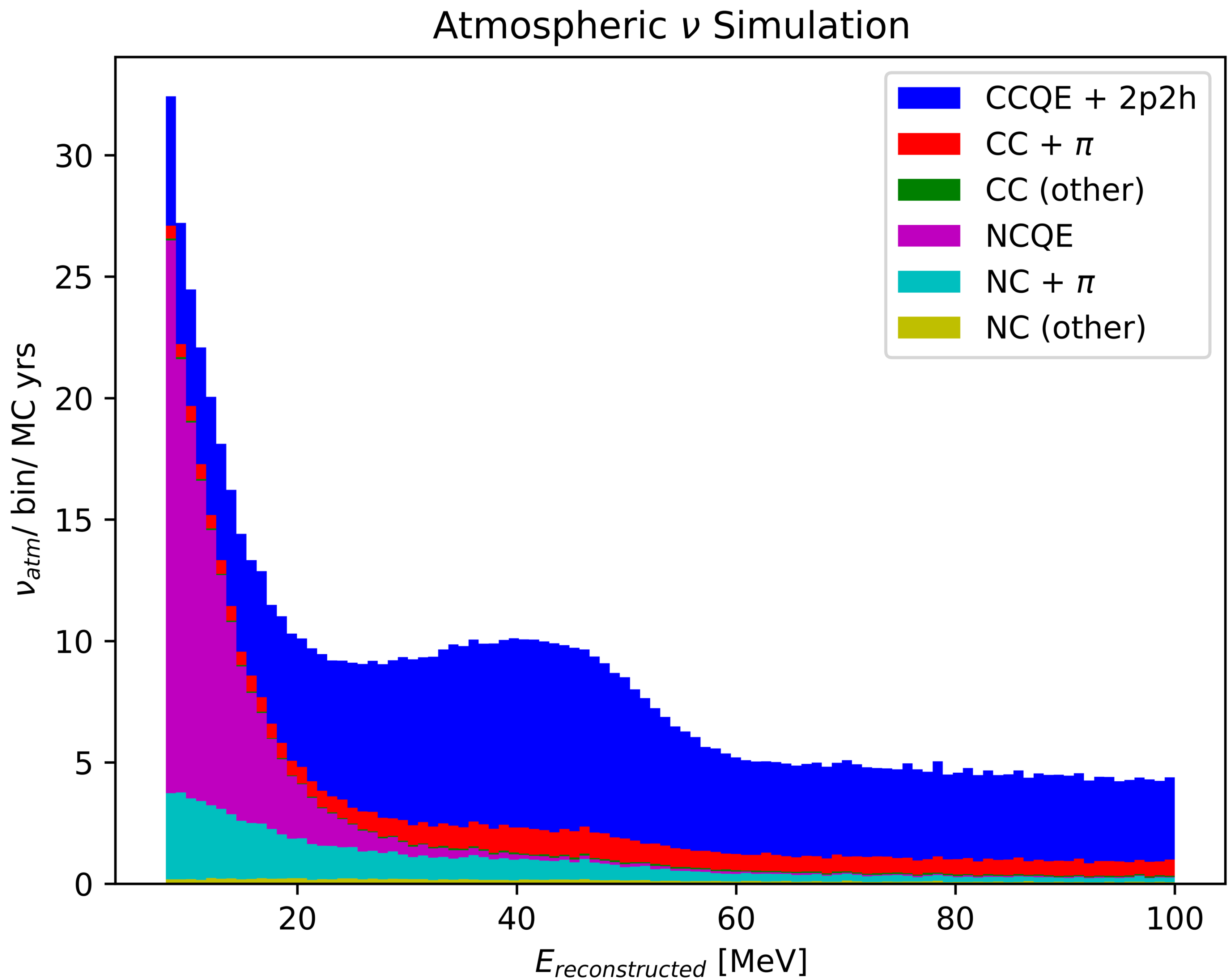
DSNB (Horiuchi)



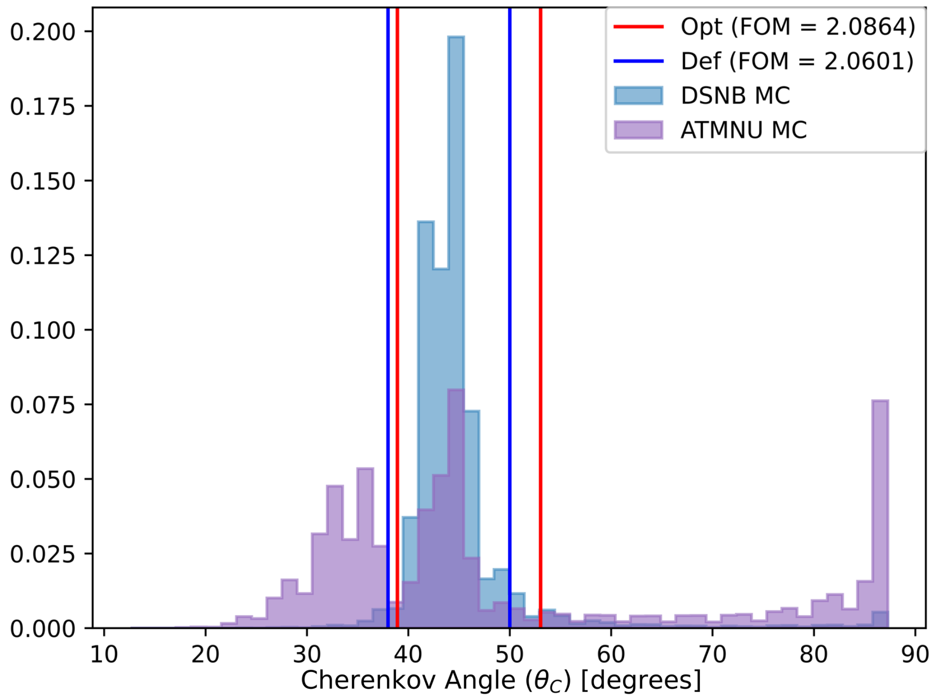
Atmospheric Background

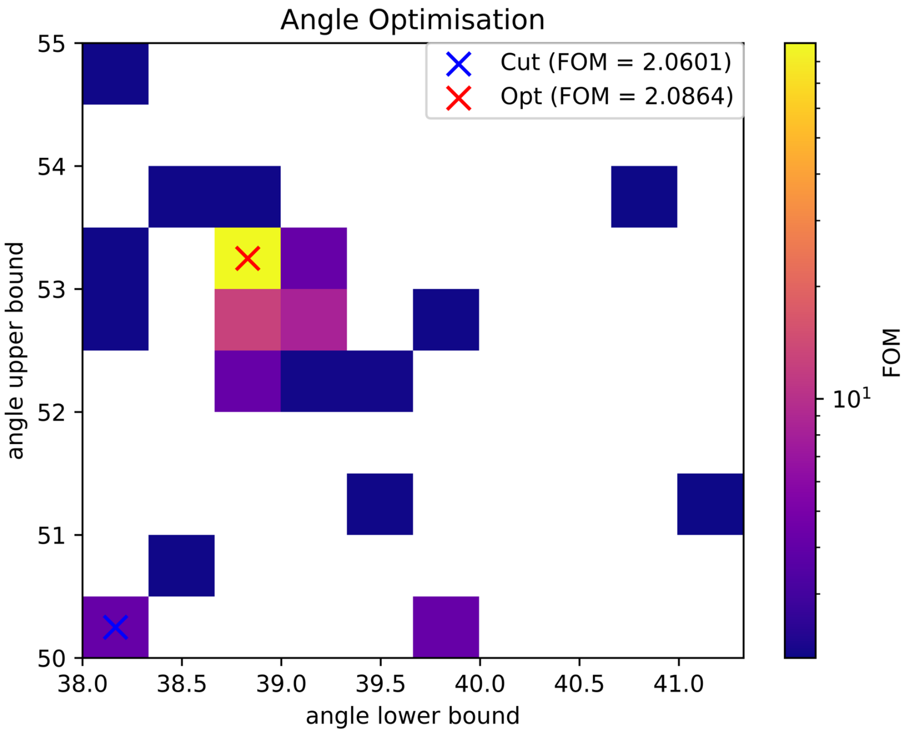


Efficiency after First Reduction



Cherenkov Angle (θ_C)

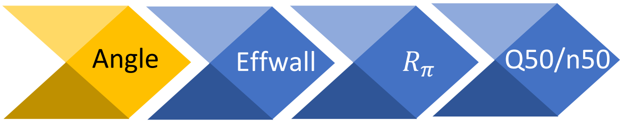




- Powerful discriminator based on the opening angle θ_C of the Cherenkov cone.
  - IBD $\theta_C \sim 42^\circ$
  - Heavier particles (μ/π) at low θ_C.
  - NC events mostly at higher θ_C due to multiple gamma emission.

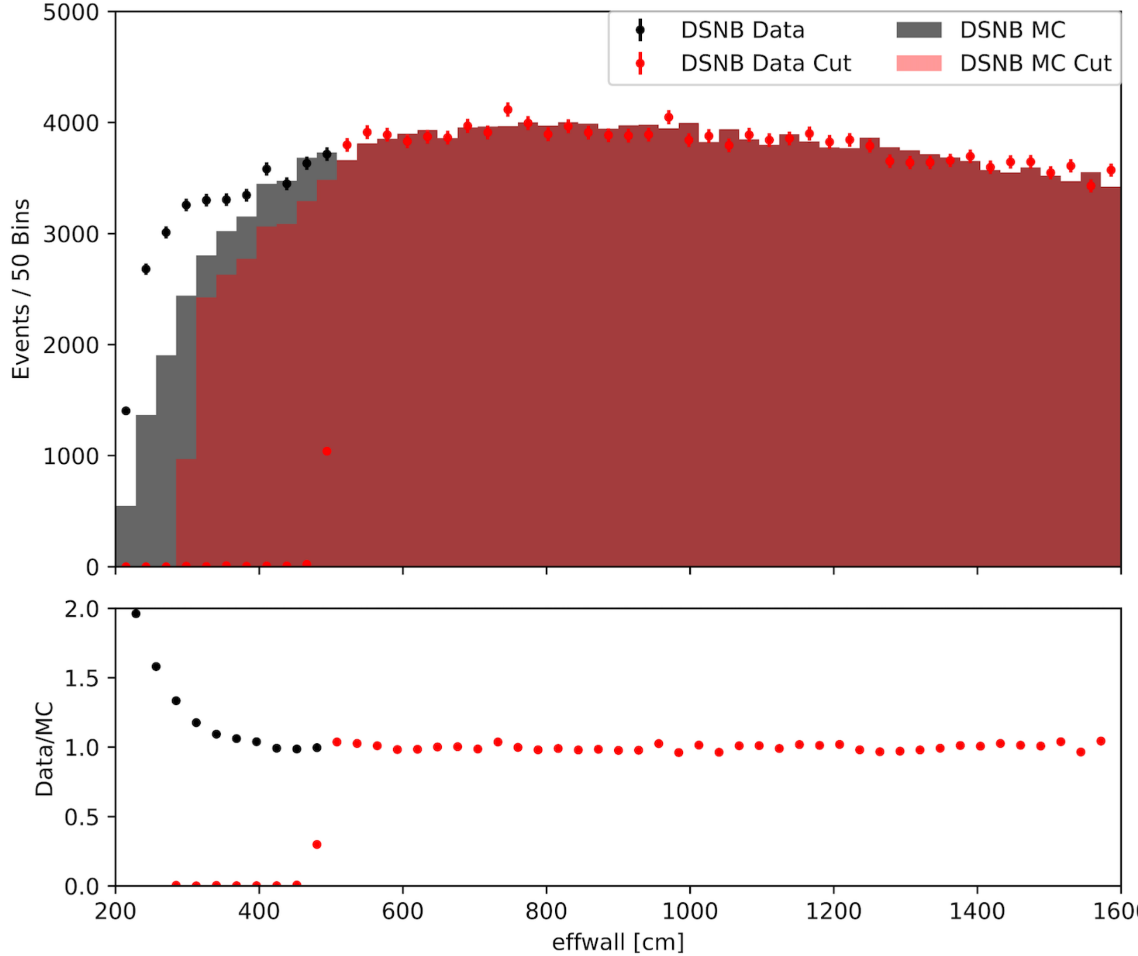
Default (SK-IV)	Optimised (SK-VI)
$38.0^\circ < \theta_C < 50.0^\circ$	$38.92^\circ < \theta_C < 53.05^\circ$

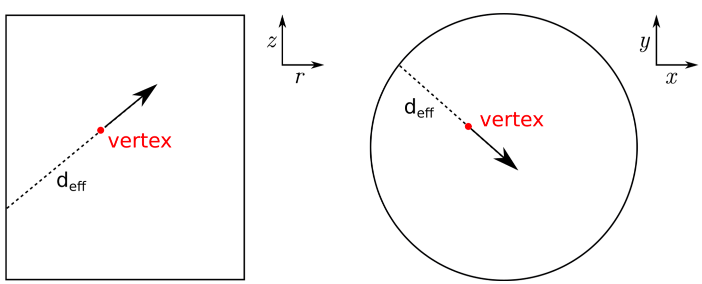
$$\text{FOM} = \frac{S}{\sqrt{S+B}}$$



Effwall (d_{eff})

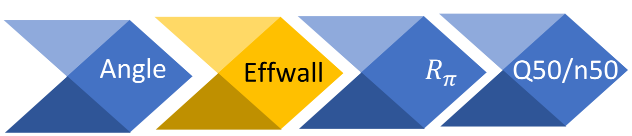
- Removes residual radioactive events from surrounding rock and detector walls.
- Since no available simulation, comparing data to IBD MC, normalised to the same area above 1000 cm, is evaluated .
- Radioactivity observed as an excess at low effwall values.
- Energy-dependent cut as background is dominant at low energy (same as SK-IV).



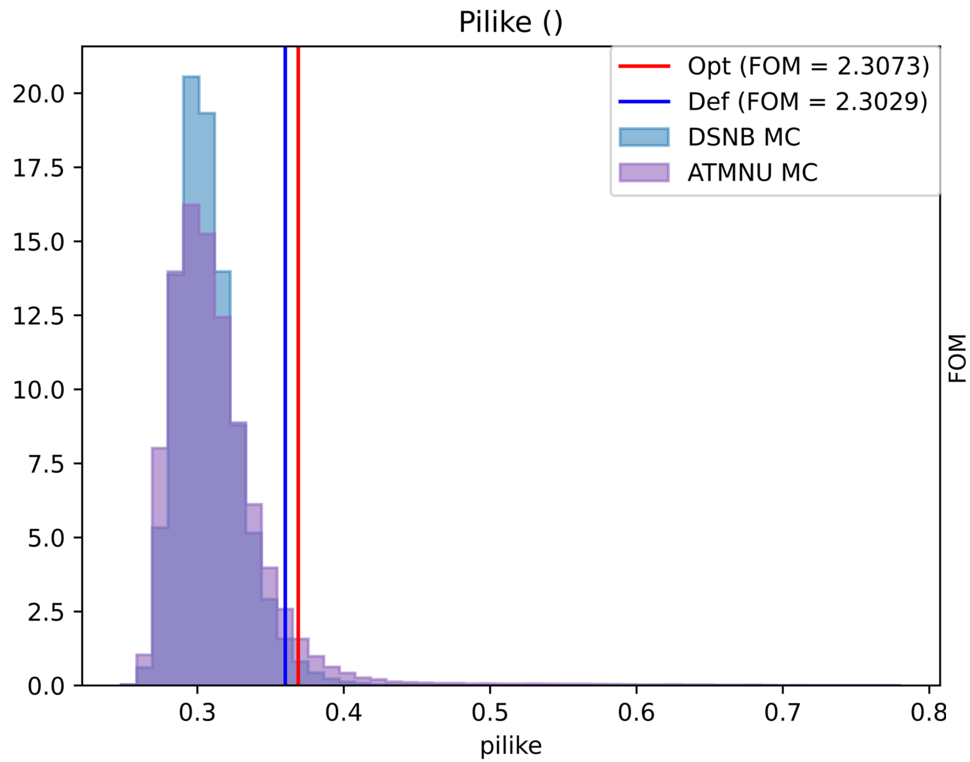


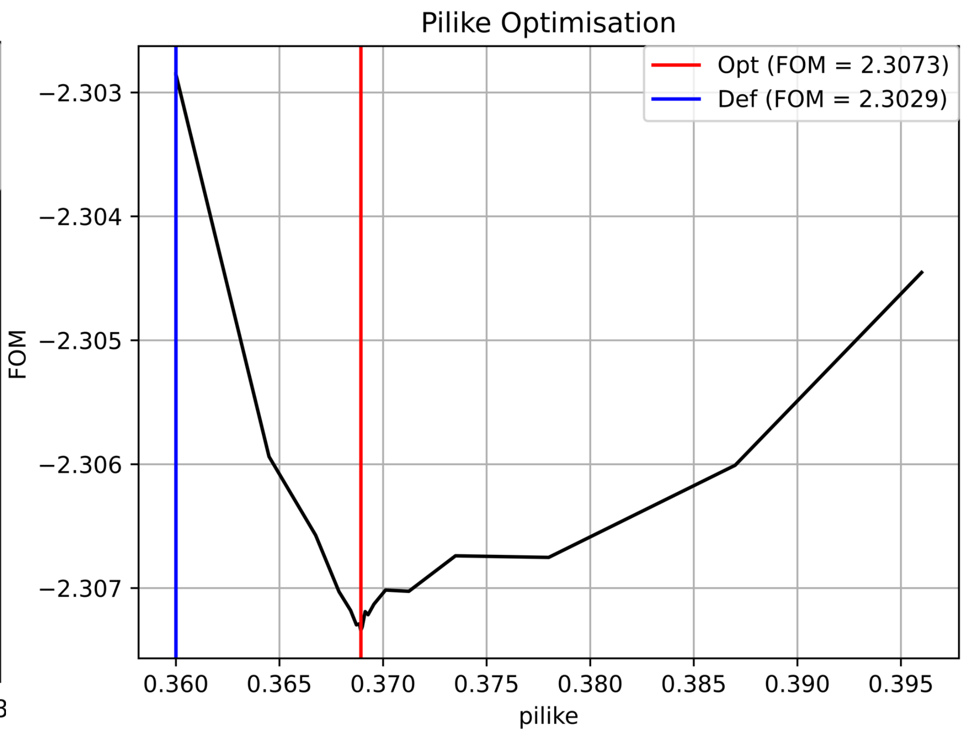
Default (SK-IV)

$$\max[300\ cm, 500\ cm\ - 50(E_{rec}\ - 16\ MeV)]$$



Ring Clearness (R_π)





- Ring clearness to remove pion-like events.

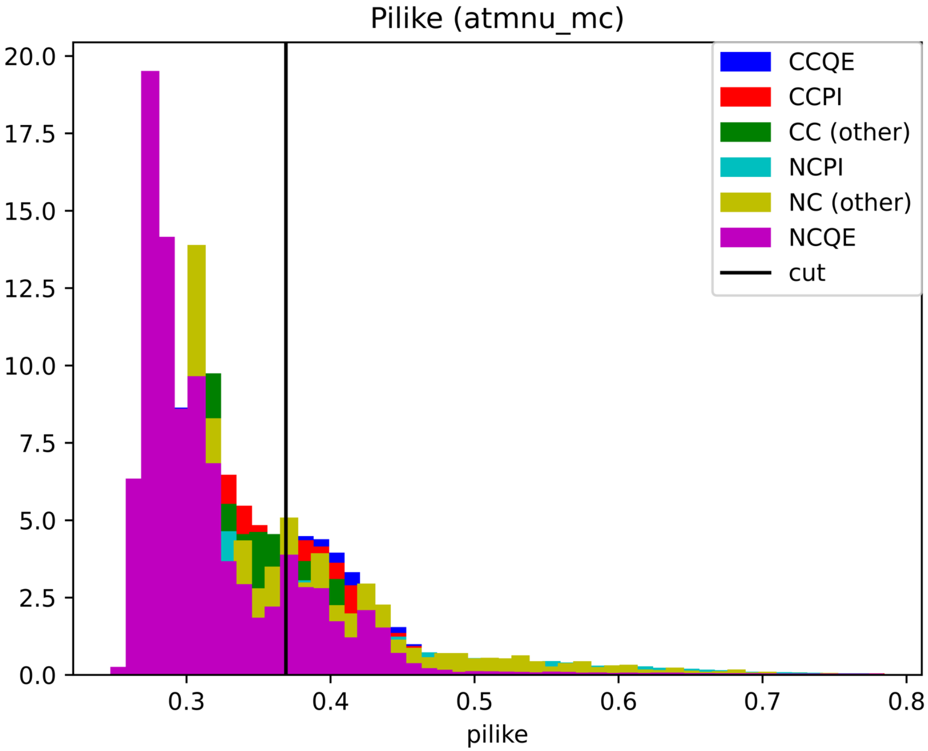
$$R_\pi = \frac{N_{triplets}(\theta_C \pm 3^\circ)}{N_{triplets}(\theta_C \pm 10^\circ)}$$

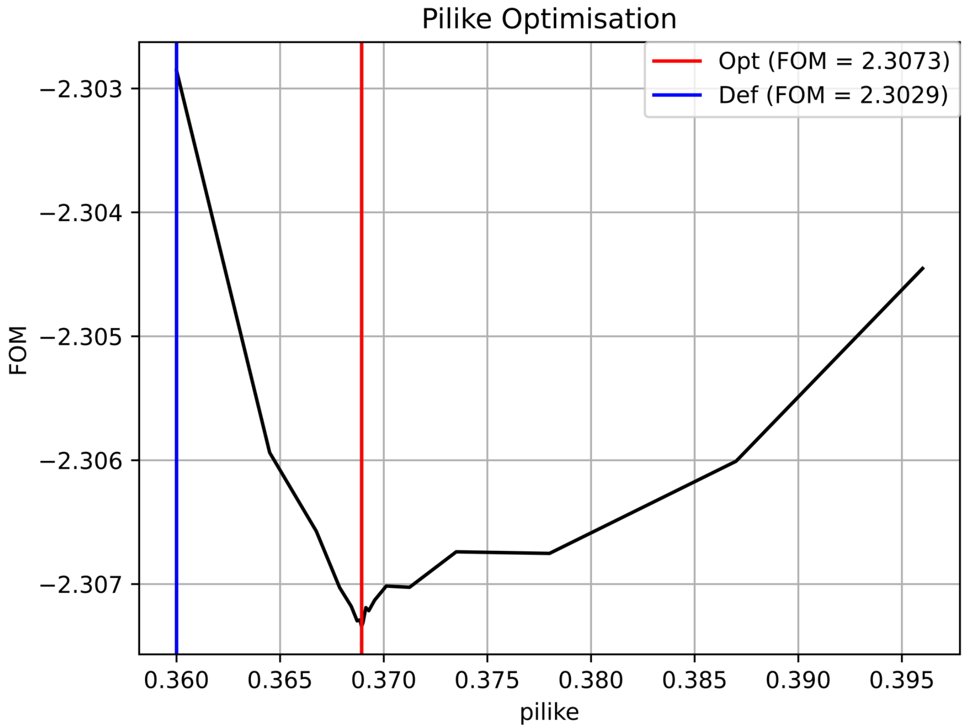
- High momentum particles tend to produce clearer Cherenkov ring.

Default (SK-IV)	Optimised (SK-VI)
< 0.36	< 0.369

$$\mathrm{FOM} = \frac{S}{\sqrt{S + B}}$$

Ring Clearness (R_π)





- Ring clearness to remove pion-like events.

$$R_\pi = \frac{N_{triplets}(\theta_C \pm 3^\circ)}{N_{triplets}(\theta_C \pm 10^\circ)}$$

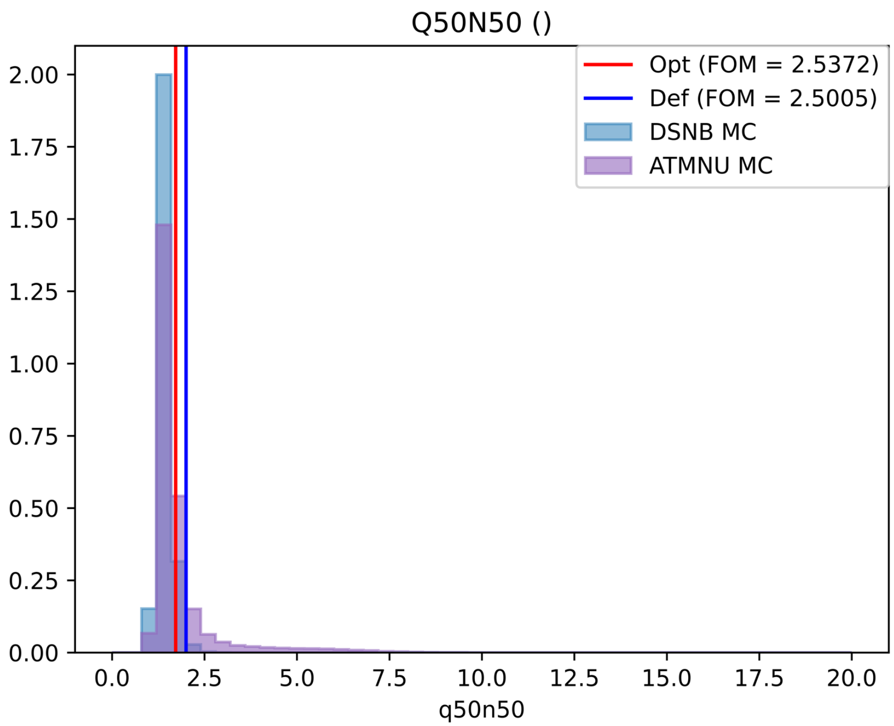
- High momentum particles tend to produce clearer Cherenkov ring.

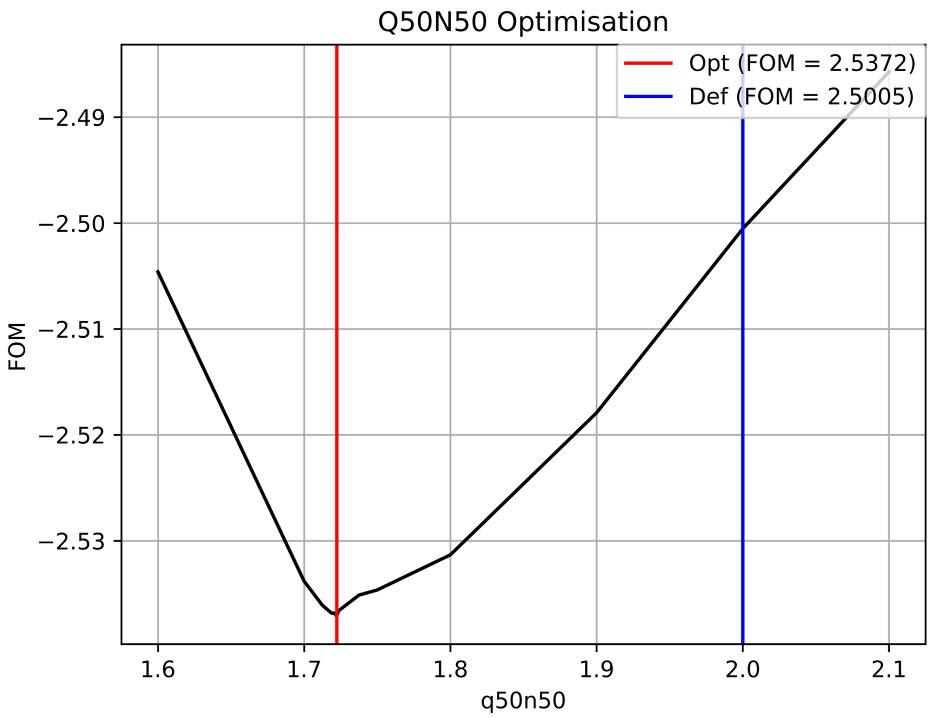
Default (SK-IV)	Optimised (SK-VI)
< 0.36	< 0.369

$$\text{FOM} = \frac{S}{\sqrt{S+B}}$$

Angle → Effwall → R_π → Q50/n50

Average Charge Deposition (q_{50}/n_{50})

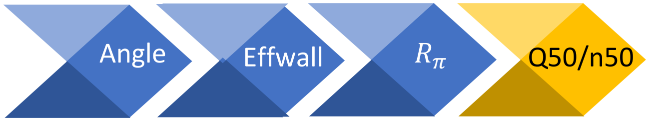




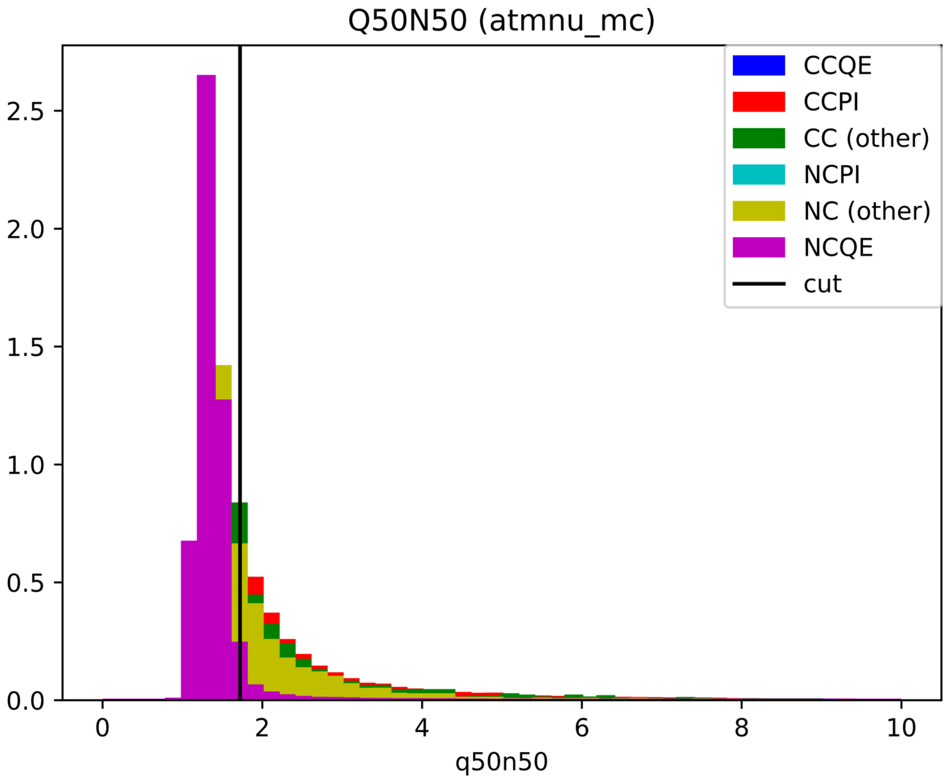
- Higher momentum particles tend to deposit more charge.
- Scan average charge deposited in 50 ns TOF window.

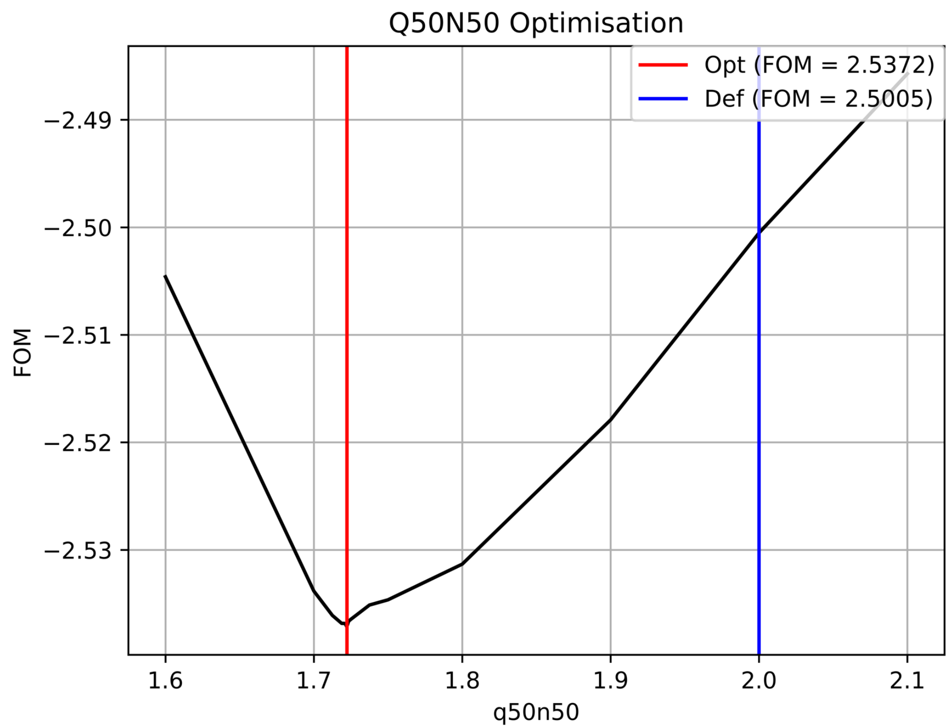
Default (SK-IV)	Optimised (SK-VI)
< 2.0	< 1.722

$$\mathrm{FOM} = \frac{\mathrm{S}}{\sqrt{\mathrm{S}+\mathrm{B}}}$$



Average Charge Deposition (q_{50}/n_{50})

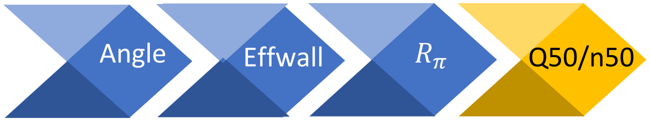




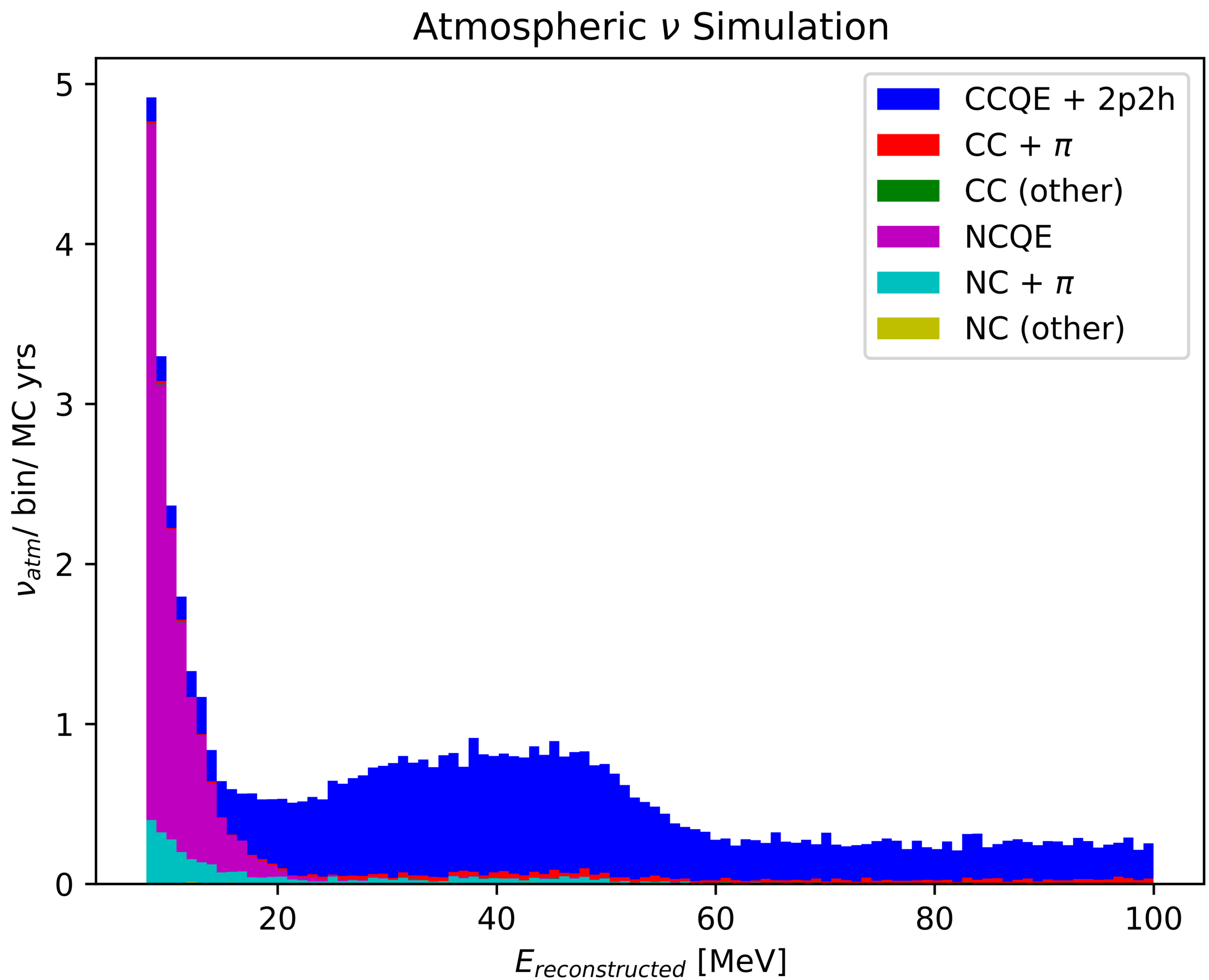
- Higher momentum particles tend to deposit more charge.
- Scan average charge deposited in 50 ns TOF window.

Default (SK-IV)	Optimised (SK-VI)
< 2.0	< 1.722

$$\mathrm{FOM} = \frac{\mathrm{S}}{\sqrt{\mathrm{S}+\mathrm{B}}}$$



Efficiency after Full Atmospheric Reduction

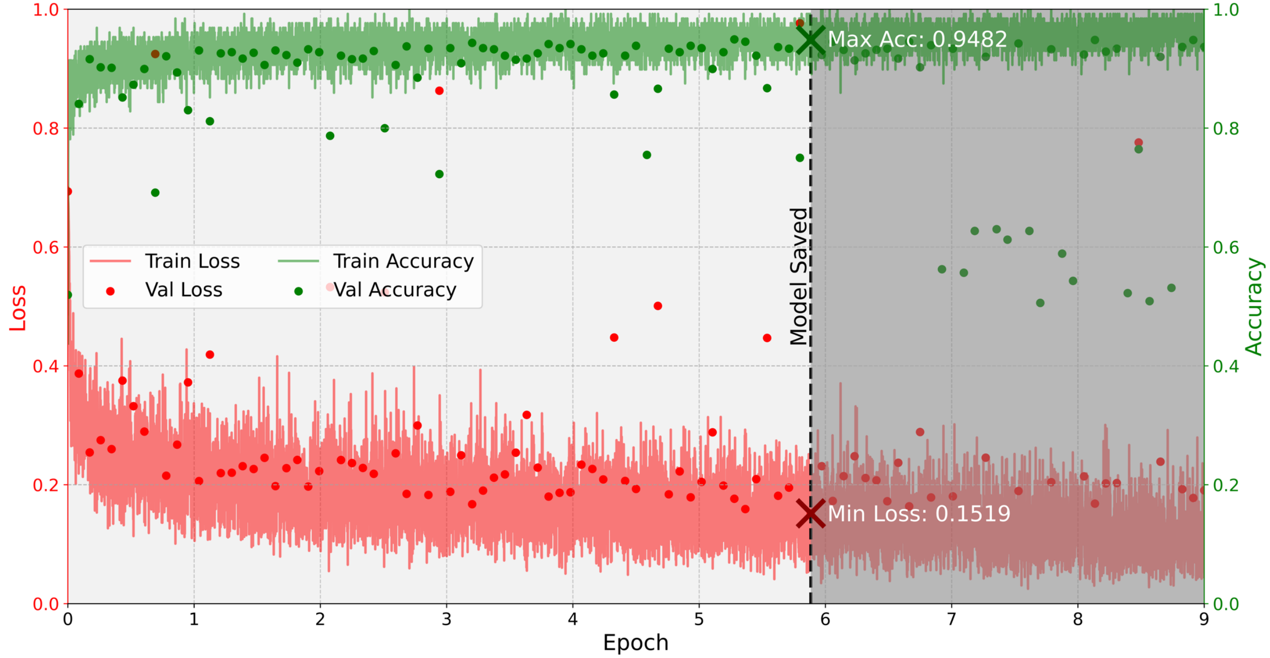


Efficiency after Atmospheric Reduction

Cut	Value	Signal Efficiency [PREV]	Signal Efficiency [TOTAL]
θ_C	$[\,38.92^{\circ}, 53.05^{\circ}]$	92.2%	92.2%
d_{eff}	Energy dependent	94.4%	86.9%
R_{π}	< 0.369	98.6%	85.6%
q_{50}/n_{50}	< 1.722	93.7%	80.2%
N_{pre}^{max}	< 12	100%	80.2%
$N_{pre}^{maxgate}$	< 5	100%	80.2%
$N_{decay-e}$	< 0	98.1%	78.6%

BACKUP – CNN ANALYSIS

FirstCNN Training

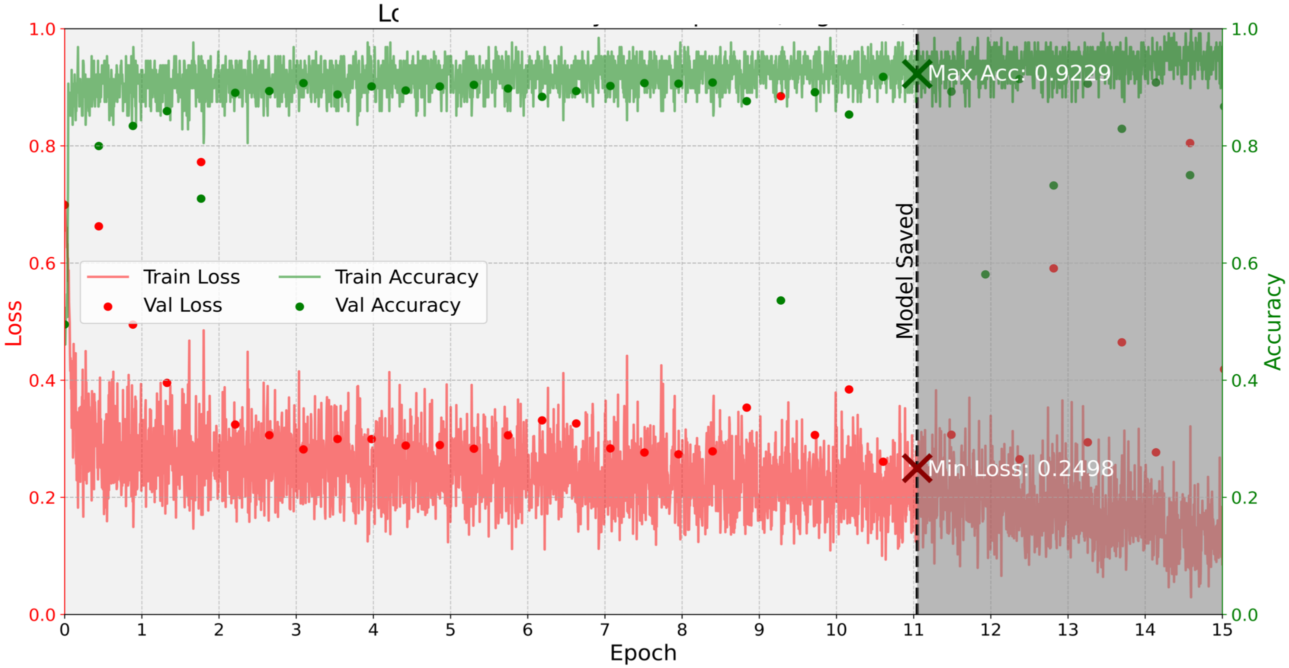


Training Parameters

Lr = 0.001
Epochs = 6
Batches= 256
Network = ResNet-18
Training Set = 400,000 events
Validation Set = 20,000 events
Testing Set = 20,000 events

- **CNN (Convolutional Neural Network):** Deep learning model designed using convolutional layers to automatically learn hierarchical patterns and features for classification of images.
- **ResNet-18:** Consists of 18 layers, utilises residual connections, enabling easier training for deeper networks.
- **SK event display** timing and charge as inputs.
- **Select 300 ns around the prompt window for e/ γ classification.**

AngleCNN Training



Training Parameters

Lr = 0.001
Epochs = 11
Batches= 256
Network = ResNet-18
Training Set = 100,000 events
Validation Set = 5,000 events
Testing Set = 5,000 events

- **CNN (Convolutional Neural Network):** Deep learning model designed using convolutional layers to automatically learn hierarchical patterns and features for classification of images.
- **ResNet-18:** Consists of 18 layers, utilises residual connections, enabling easier training for deeper networks.
- **SK event display** timing and charge as inputs.
- **Select 300 ns around the prompt window for e/ γ classification.**

CNN Model Performance Scaling

- **FirstCNN:**
  - Tested on FirstCNN test set
- **AngleCNN – Enhanced Reduction:**
  - Tested on AngleCNN test set
  - Scaled according to the signal and background efficiency of the Cherenkov angle cut
  - $\varepsilon_{sig} = 0.96$
  - $\varepsilon_{bkg} = 0.28$
- **AngleCNN > FirstCNN**
  - Marginal improvement for most of the ROC curve.
  - AngleCNN performance declines at the signal efficiency of Cherenkov angle cut