Online pointing to SN bursts with machine learning algorithms in DUNE

IRN Neutrino Paris 2024

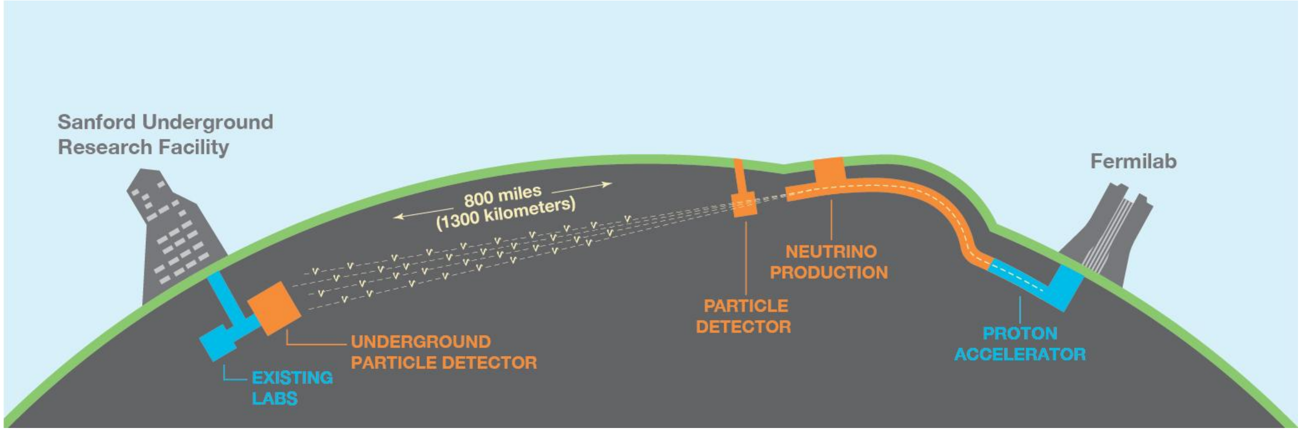
10/10/2024

Dario Pullia
On behalf of the DUNE collaboration
dario.pullia@cern.ch

Deep Underground Neutrino Experiment

DUNE is designed to study neutrino oscillations over a 1300 km baseline.

This requires building a Far Detector, which can also be used for various other physics purposes. The experiment is under construction and the beginning of data taking is expected in 2029.

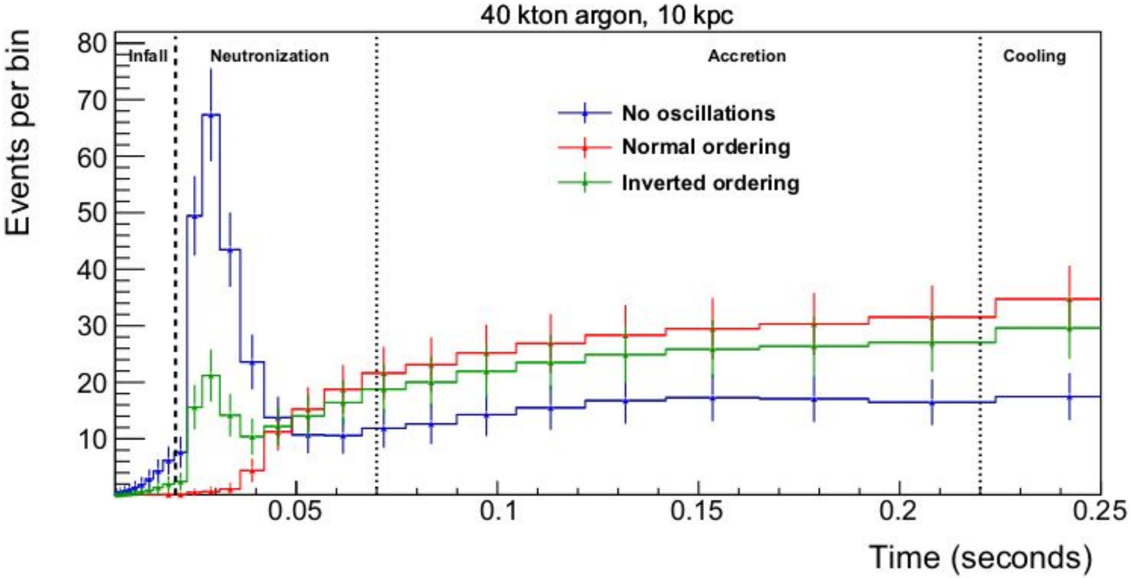


Supernova Neutrino Burst

[Eur. Phys. J. C 81, 423 (2021).]

A Supernova Neutrino Burst (SNB) happen at the end of a massive star's life.

- 99% of the gravitational energy is emitted in neutrinos.
- Expected 2-3 events per century, compatible with the expected ~30 years of data taking for DUNE.

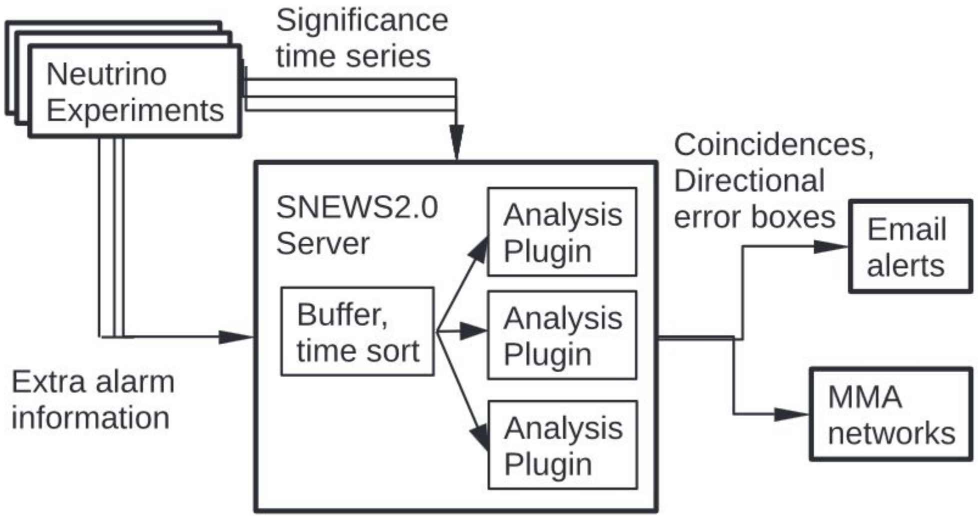


For these studies: **10 kpc** distance and **40 kton** of active LAr

SuperNova Early Warning System (SNEWS)

Neutrinos arrive on earth tens of minutes before the light.

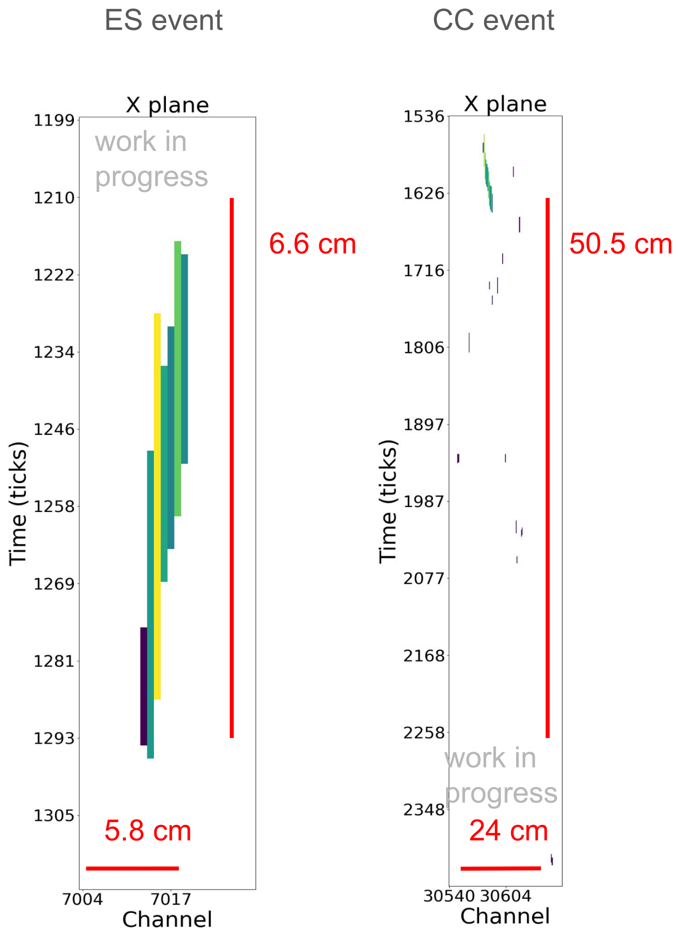
The SuperNova Early Warning System (SNEWS) is created to provide a prompt alert to telescopes if one SNB happens.



ES vs CC interaction properties

Channel	Liver-more	GKVM	Garching
$\nu_e +^{40}\text{Ar} \rightarrow e^- +^{40}\text{K}^*$	2648	3295	882
$\overline{\nu}_e +^{40}\text{Ar} \rightarrow e^+ +^{40}\text{Cl}^*$	224	155	23
$\nu_X + e^- \rightarrow \nu_X + e^-$	341	206	142
Total	3213	3656	1047

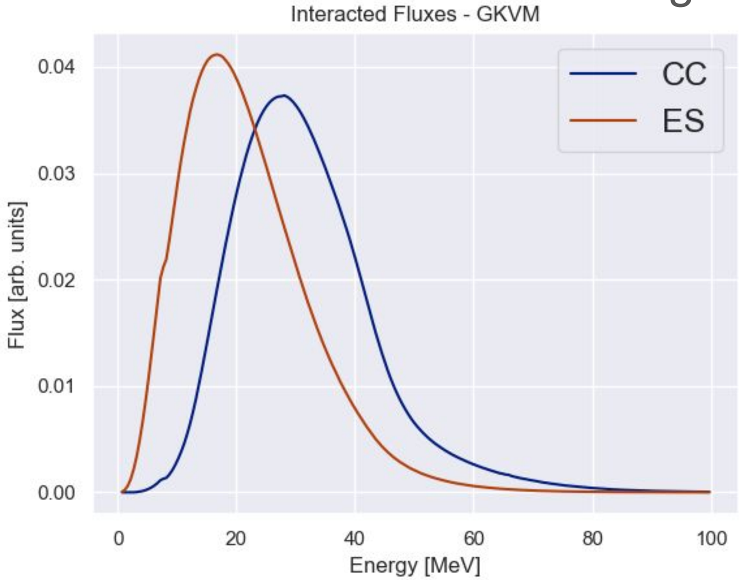
CC tracks are surrounded by **blips**, which are secondary small tracks resulting from the Compton scattering of gammas emitted during the nucleus's de-excitation.

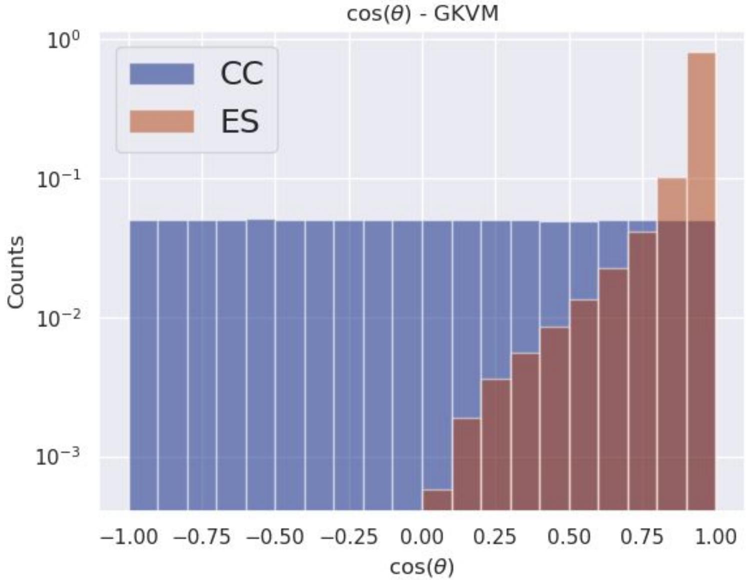


ES vs CC interaction properties

Only electrons from Elastic Scattering interaction carry directional information and can be used in this analysis.

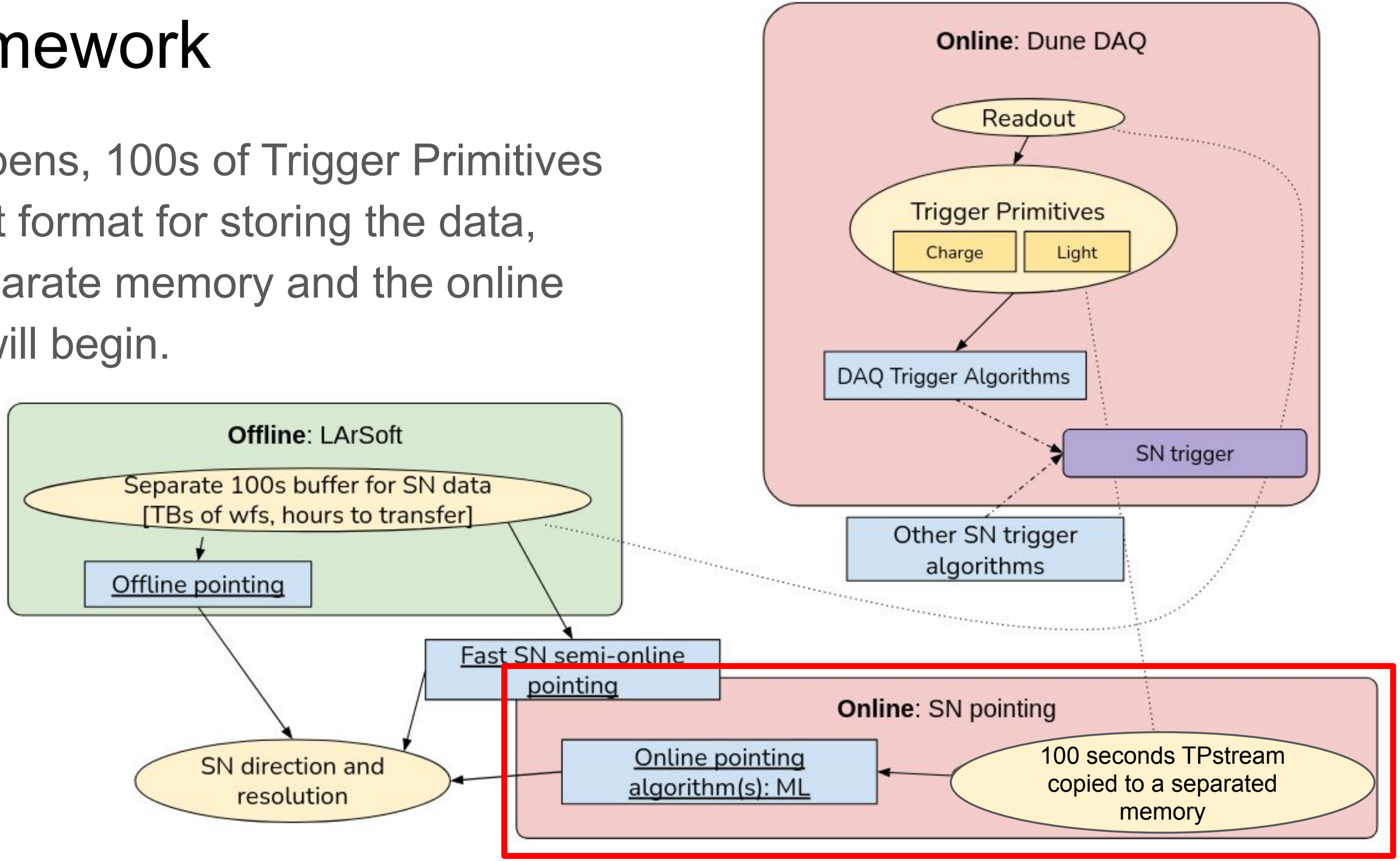
generated with MARLEY



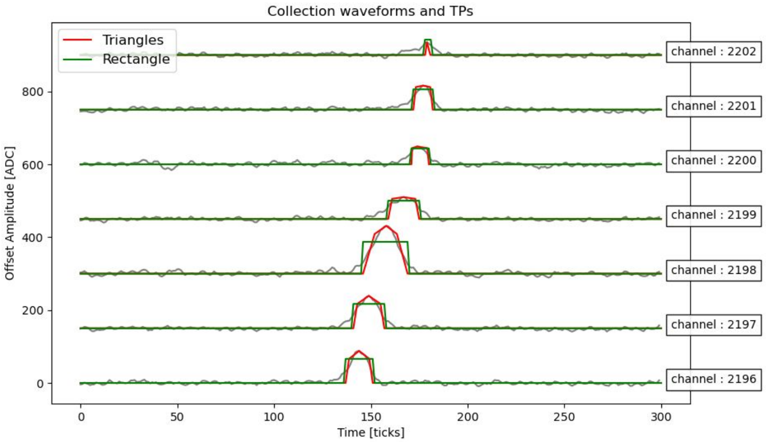


Analysis framework

If a SN trigger happens, 100s of Trigger Primitives (**TPs**), a lightweight format for storing the data, are copied to a separate memory and the online pointing workflow will begin.



Analysis framework



Start time
Time over threshold
Time of the ADC peak
Channel number
ADC integral
ADC peak

TPs advantages:

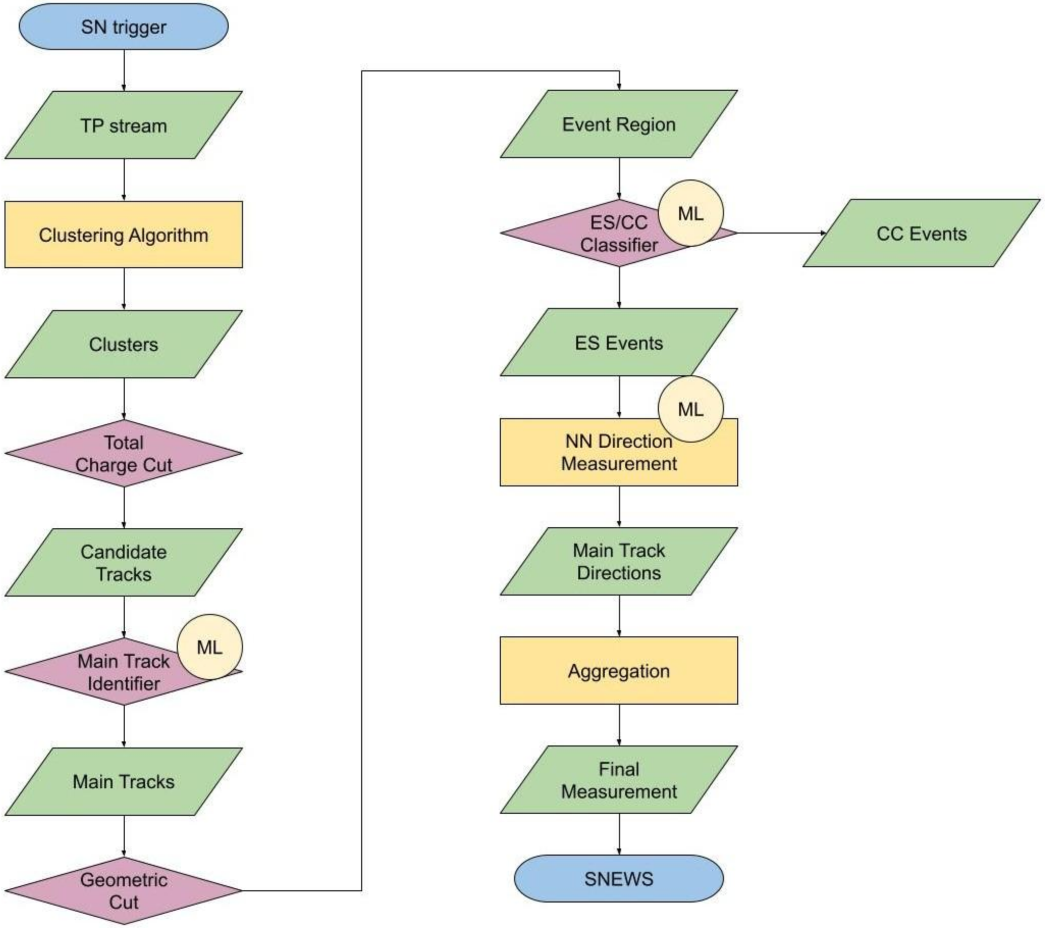
- Fast generation
- Lightweight format
- Accessibility during readout

TPs disadvantages:

- Lack of complete information about shape (e.g. double peaks due to two different hits)
- Might miss some waveform due to the fixed threshold

Workflow

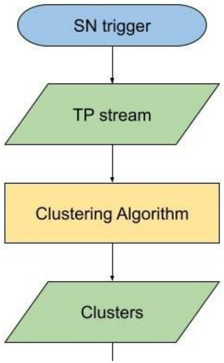
A multi-step approach is proposed, allowing flexibility on the single step solution and making possible to apply physics-based decisions.

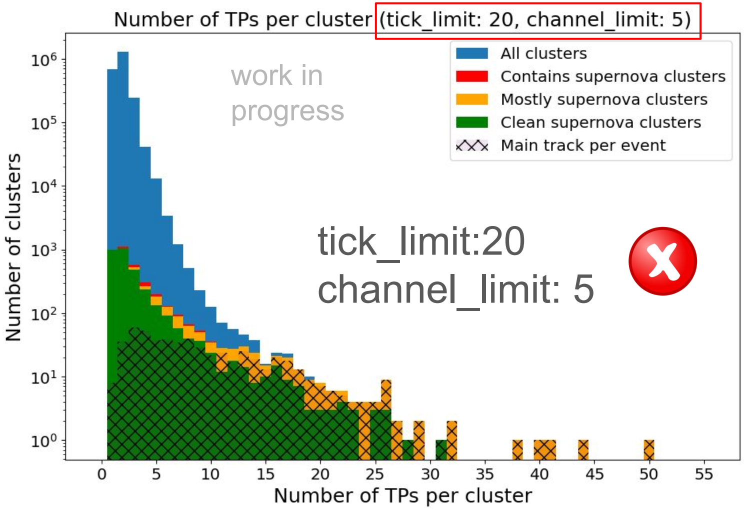


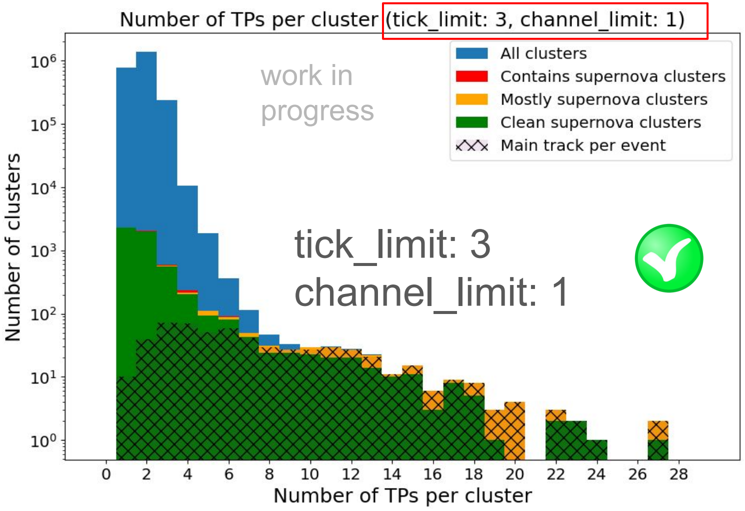
Clustering

All of these studies are based on **clusters** of TPs.

The clustering algorithm depends on two parameters that define the proximity conditions: **channel_limit** and **ticks_limit**.

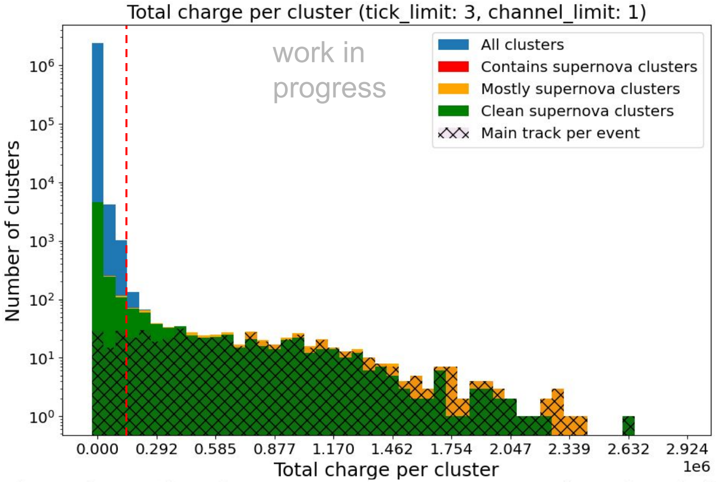


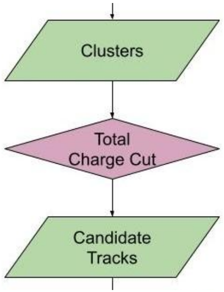


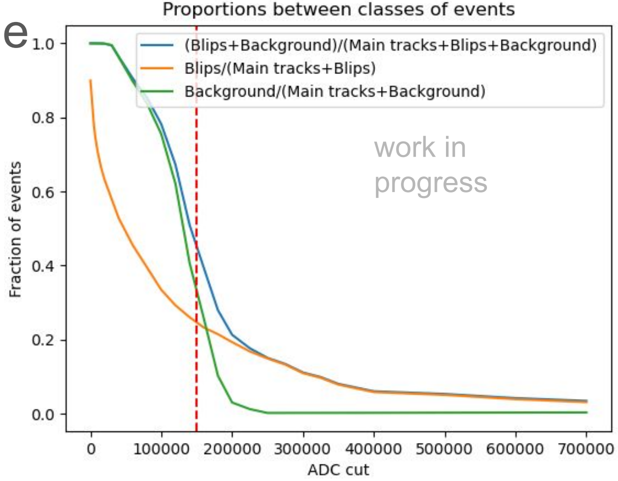


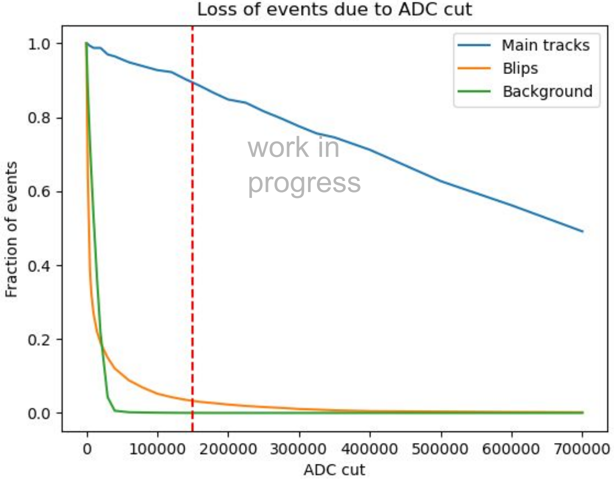
Total charge cut

The cut on the total charge is set at 150,000 ADC counts (~2.5 MeV) to remove most of the background events while maintaining ~90% of main electron tracks.





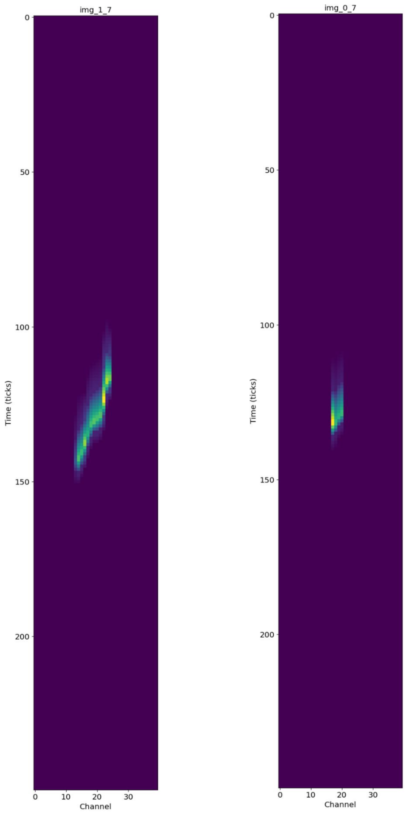




2D Convolutional Neural Network

Sample images

sig bkg



Pros:

- Established technique for DUNE analysis. (Phys. Rev. D 102, 092003 (2020))
- Reliable and widely used.

Cons:

- Not efficient on sparse images.

The NN architecture is optimized using the HyperOpt library to scan over many different combinations of the listed parameters.

Parameter	Values
Number of Convolutional Layers ($n_{\text{conv\_layers}}$)	1, 2, 3, 4
Number of Filters (n_{filters})	16, 32, 64
Kernel Size (kernel_size)	1, 3, 5
Number of Dense Layers ($n_{\text{dense\_layers}}$)	2, 3, 4
Number of Dense Units ($n_{\text{dense\_units}}$)	32, 64, 128
Learning Rate (learning_rate)	[0.0001, 0.001]
Decay Rate (decay_rate)	[0.90, 0.999]

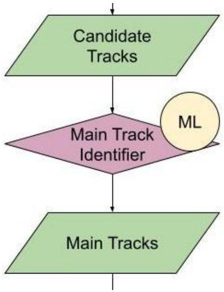
Main track identification (training)

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$



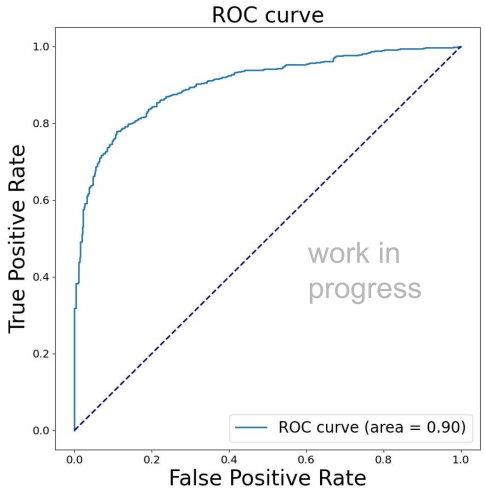
Accuracy	Precision	Recall	F1
0.821	0.824	0.823	0.820

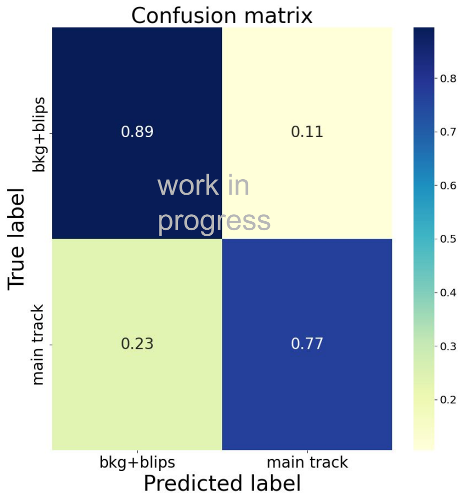
Conditions:

Total charge cut set to 80,000 to maximize the number of samples.

bkg and blips equally sampled to expose the network effectively to both topologies.

N main tracks	N bkg+blips
6634	6634





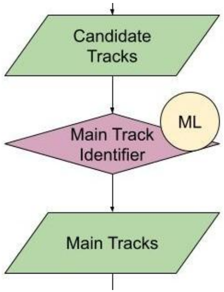
Main track identification (testing)

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

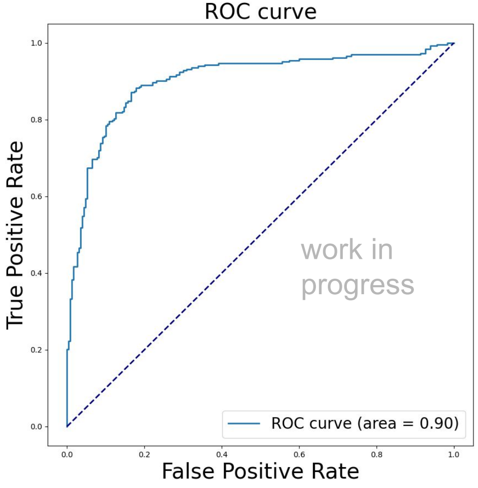


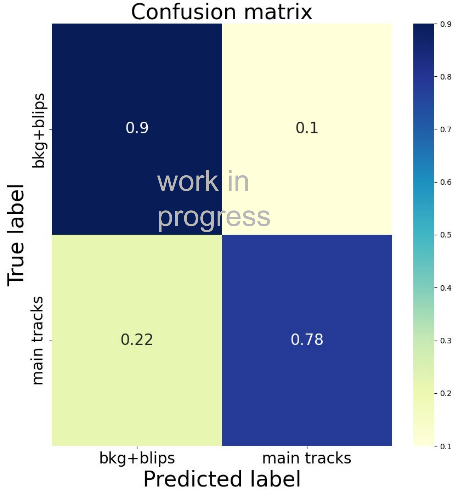
Accuracy	Precision	Recall	F1
0.838	0.842	0.842	0.838

Conditions:

Total charge cut set to 150,000 to mimic the conditions during operations.

Classification threshold is set at 0.715 to achieve 10% background in the main tracks.

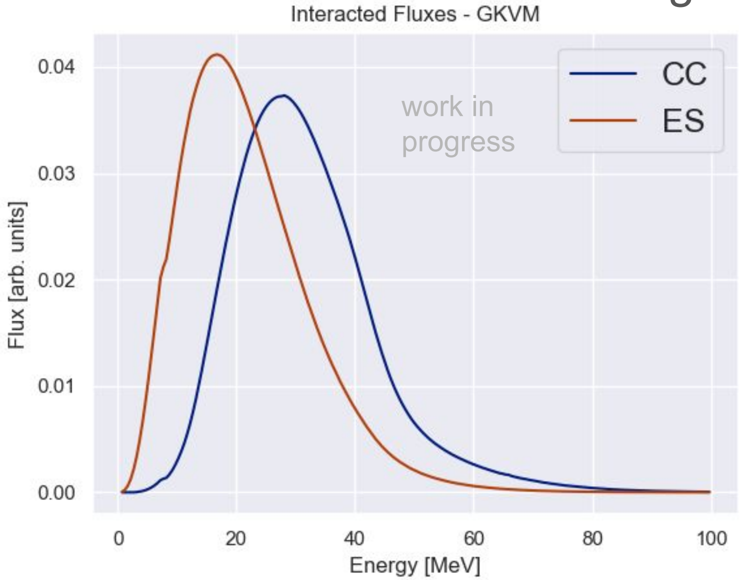


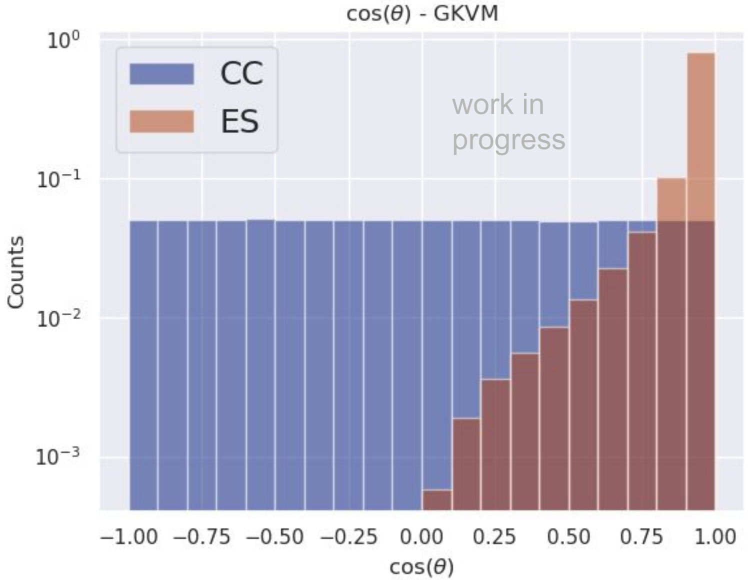


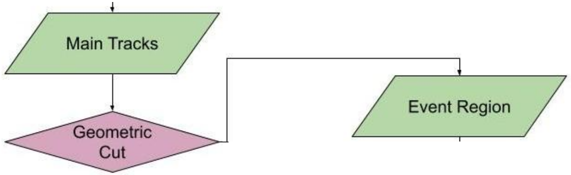
ES vs CC interaction properties

Only electrons from Elastic Scattering interaction carry directional information and can be used in this analysis.

generated with MARLEY

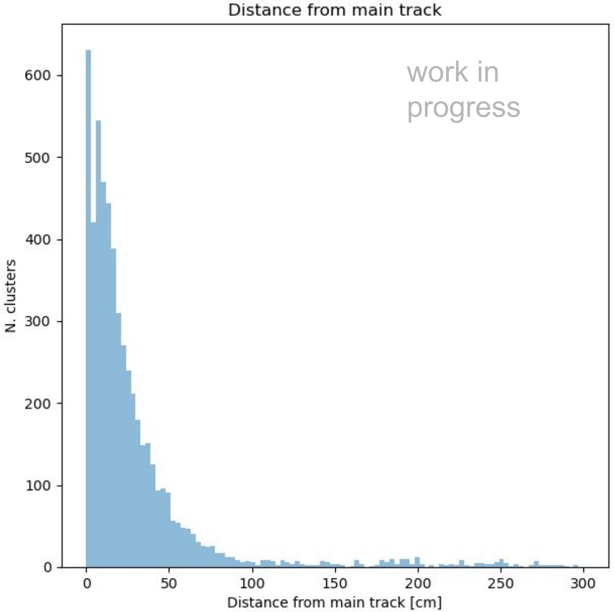


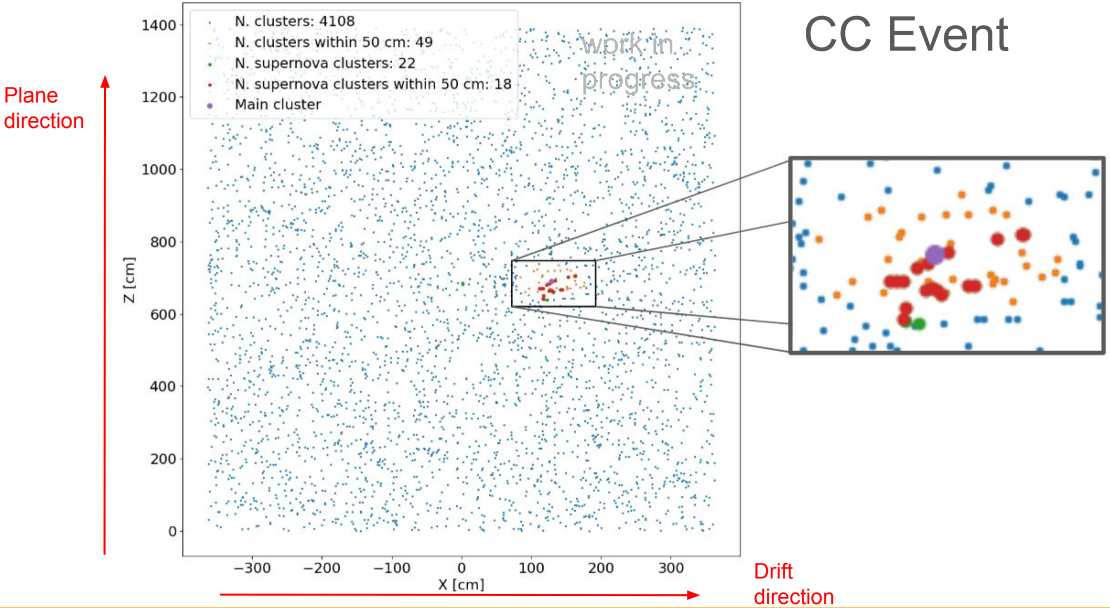




Cluster position

A fiducial volume is extracted around the selected main tracks to include the most blips while minimizing the included backgrounds.





CC Event

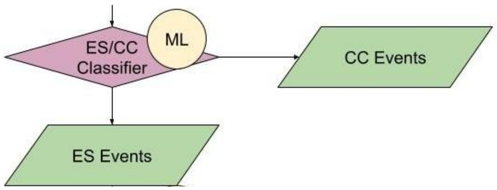
ES vs CC classification

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

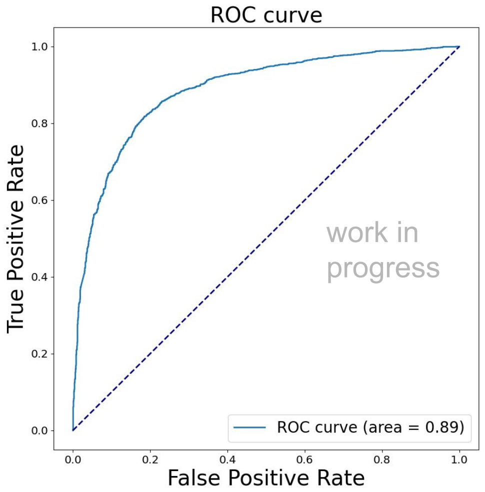


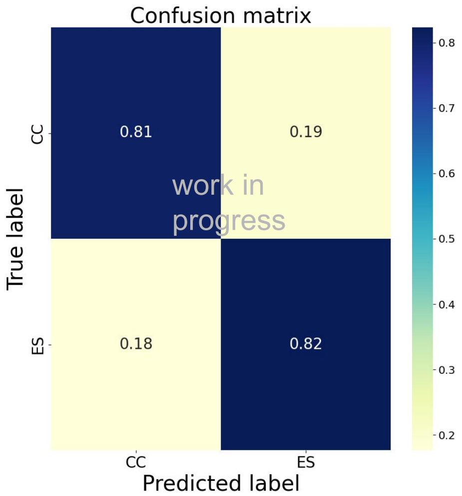
Conditions:

Test executed on an independent dataset from the training and validation one.

Accuracy	Precision	Recall	F1
0.815	0.815	0.815	0.815

N ES samples	N CC samples
14144	14144

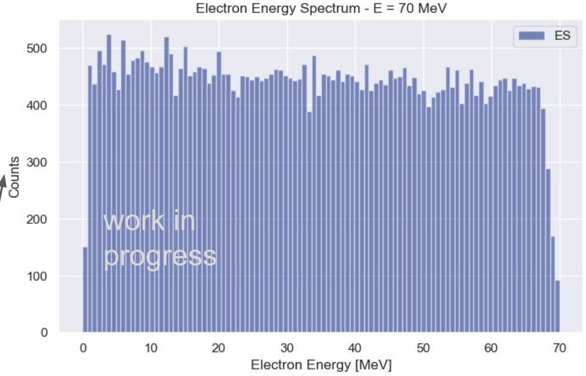


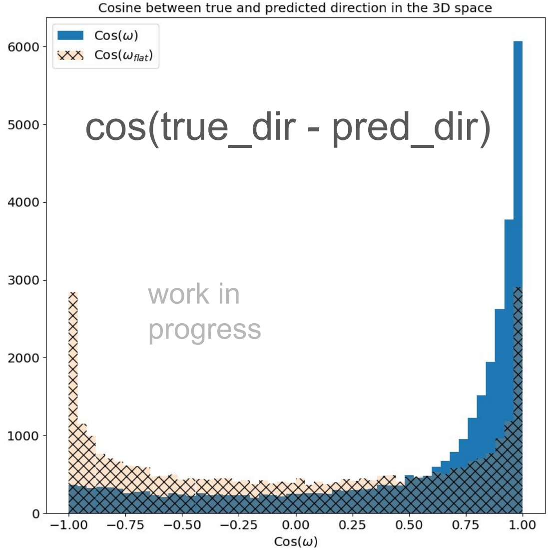


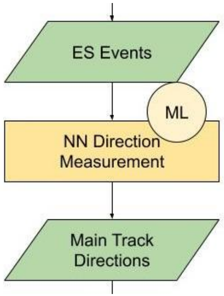
Pointing

- ~1.500 directions, ~200.000 individual tracks.
- Truth is the neutrino direction, not the individual electron one.
- Dataset generated using flat spectrum in the 30-70 MeV range. Electron energies are flat in the range (0, E_ν), therefore high energy neutrinos can be used to agnostically train the NN across all energies.

N samples	250889







Model

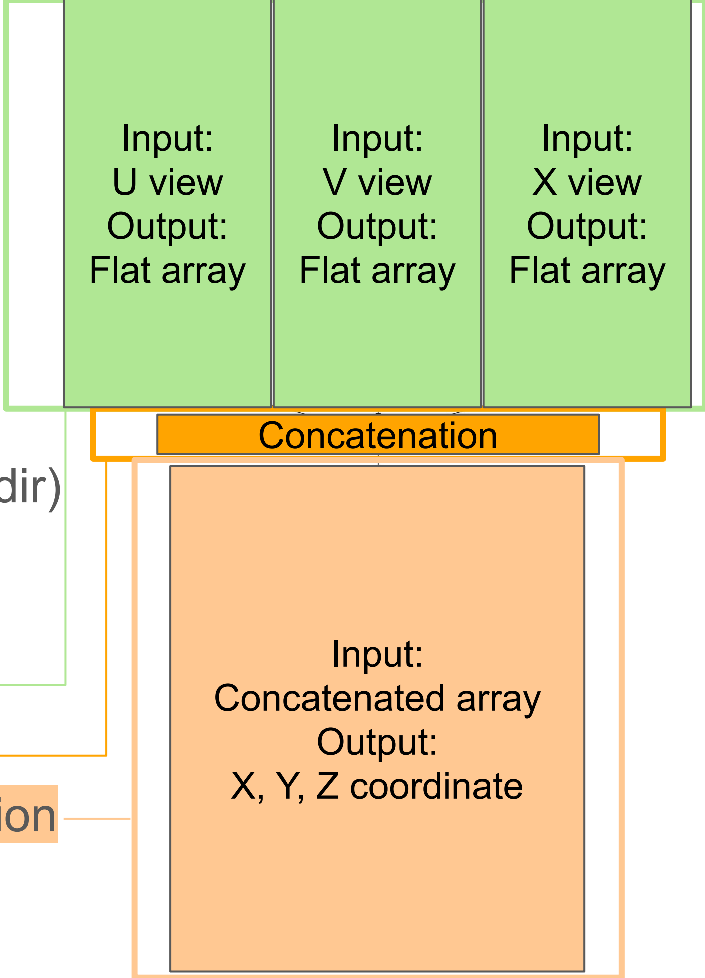
Input: 3 views x 250 ticks x 40 channels images.

Output: 3 numbers normalized to 1 (xyz)

Loss: 1-cos(ω) (ω≡the angle between true and reco dir)

Architecture:

- 3 independent CNN branches
- Flatten -> Concatenation
- Unique DNN section that returns the final prediction



Testing

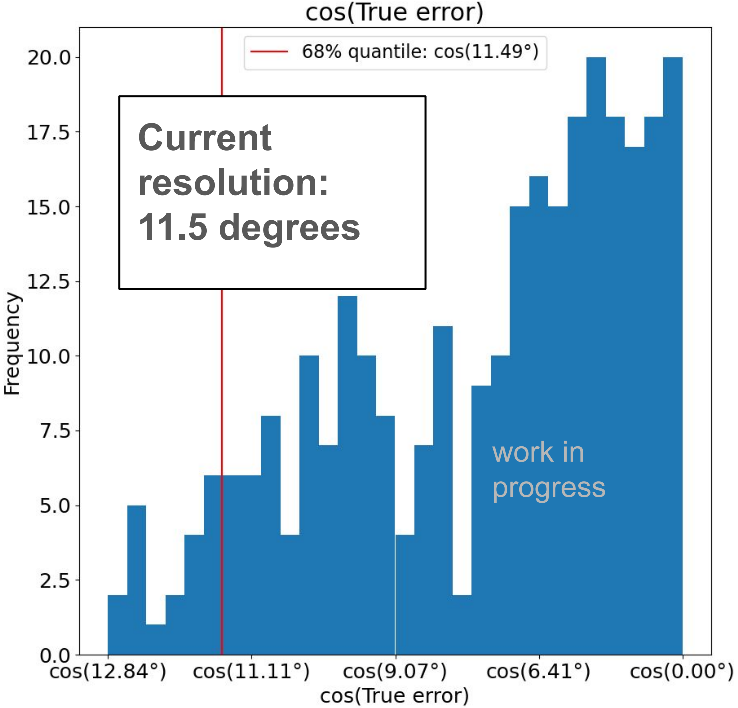
- 400 supernova burst
- Each burst contains 350 ES events, I can fully reconstruct ~220 (working for improving this).
- Per each burst I compute the predicted direction by sampling the log-likelihood with a MCMC method.

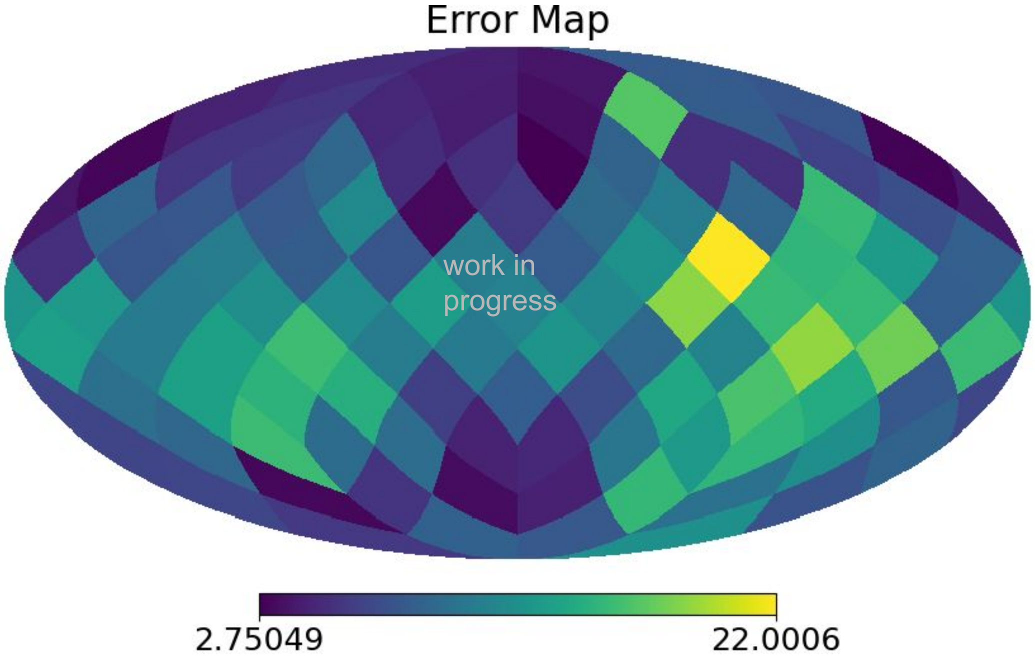
$$L = \prod_i e^{-\omega_i^2}$$

ω≡angle between measured and proposed direction

Results

Using the full waveform information and without time constraints:
4.3 degrees
https://arxiv.org/abs/2407.10339





True error = angle between the predicted and the MC true direction

Conclusions

- The first attempt of creating a supernova pointing system using TPs is proposed.
- Each component of the pipeline is developed and tested individually.

Next steps:

- Run the full pipeline over many burst simulations with background.
- Include electrical noise in simulations.
- Refine each step where possible (e.g use electron direction)

Thanks for the attention!

BACKUP

Model

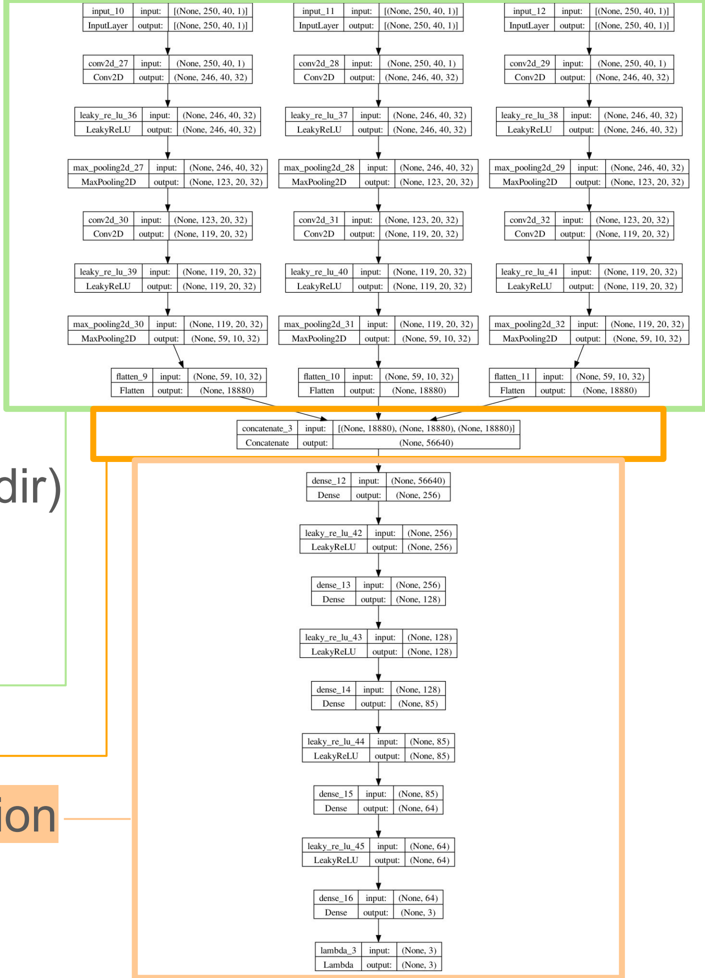
Input: 3 views x 250 ticks x 40 channels images.

Output: 3 numbers normalized to 1 (xyz)

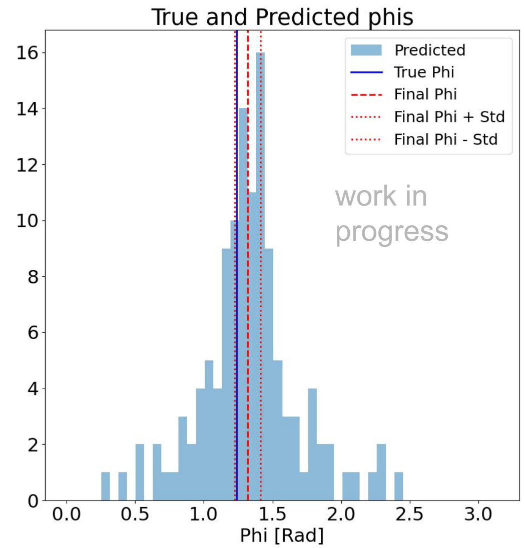
Loss: 1-cos(ω) (ω≡the angle between true and reco dir)

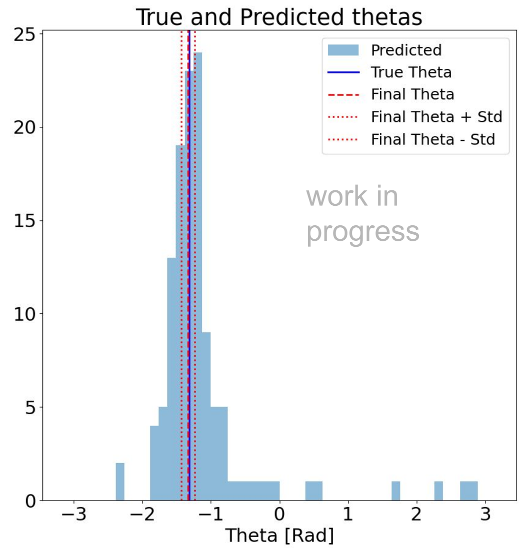
Architecture:

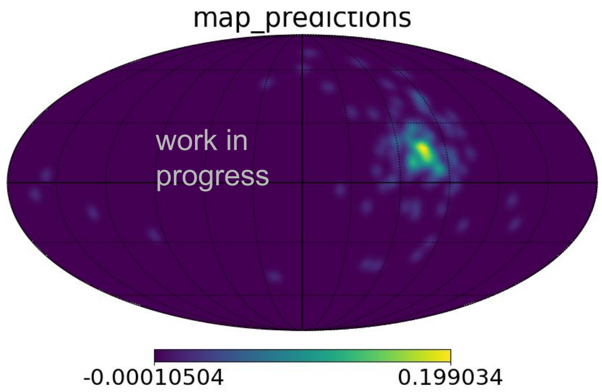
- 3 independent CNN branches
- Flatten -> Concatenation
- Unique DNN section that returns the final prediction

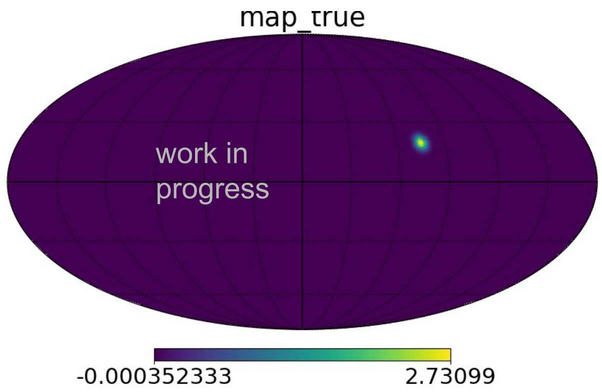


Some plots - A

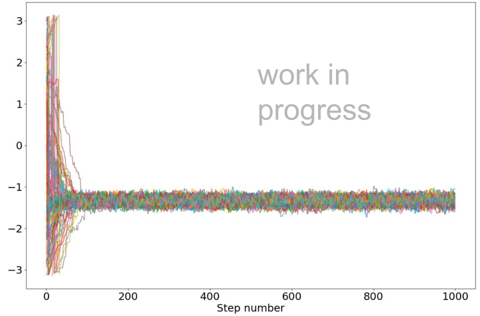

True and Predicted phis
Predicted
True Phi
Final Phi
Final Phi + Std
Final Phi - Std
work in progress
Phi [Rad]
0.0
0.5
1.0
1.5
2.0
2.5
3.0
0
2
4
6
8
10
12
14
16


True and Predicted thetas
Predicted
True Theta
Final Theta
Final Theta + Std
Final Theta - Std
work in progress
Theta [Rad]
−3
−2
−1
0
1
2
3
0
5
10
15
20
25

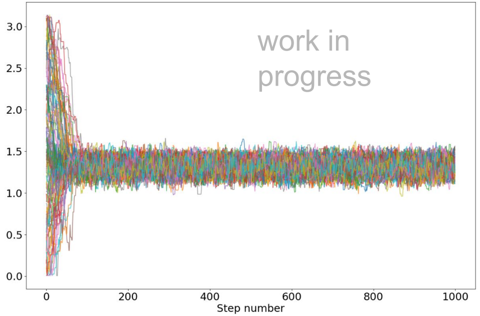

map_predictions
work in progress
-0.00010504
0.199034


map_true
work in progress
-0.000352333
2.73099

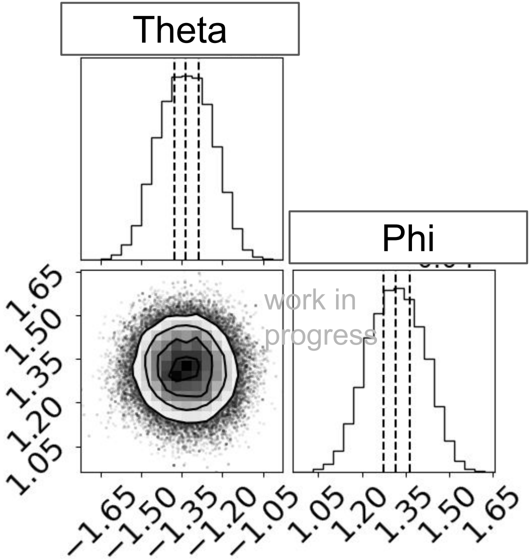
Some plots - A



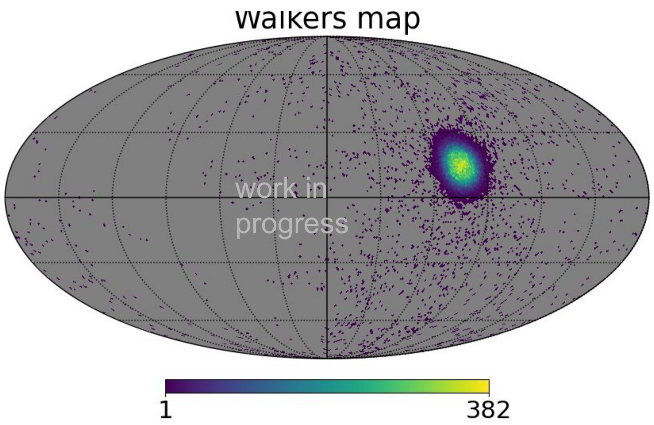
3
2
1
0
−1
−2
−3
work in
progress
0
200
400
600
800
1000
Step number



3.0
2.5
2.0
1.5
1.0
0.5
0.0
work in
progress
0
200
400
600
800
1000
Step number

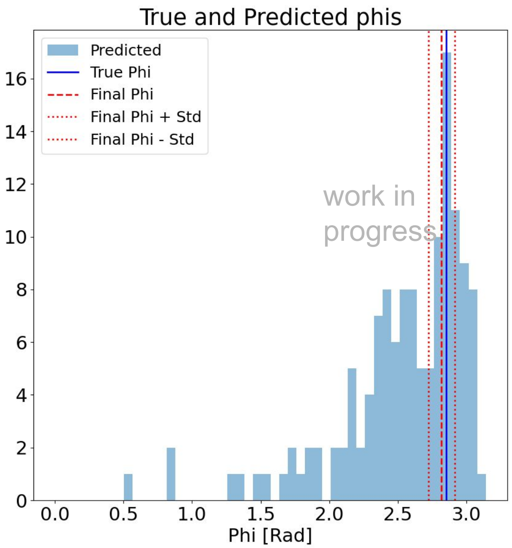


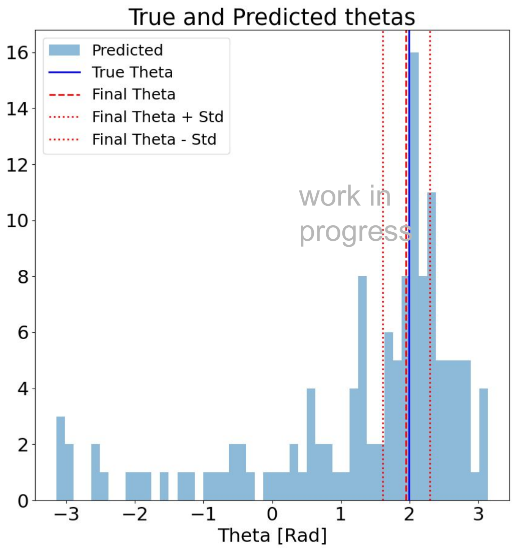
Theta
Phi
work in
progress
1.65
1.50
1.35
1.20
1.05
−1.65
−1.50
−1.35
−1.20
−1.05
1.05
1.20
1.35
1.50
1.65

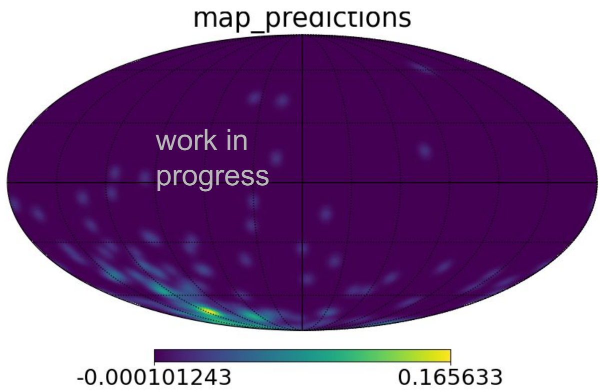


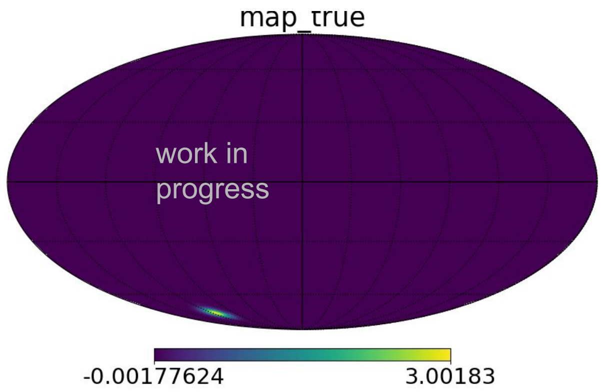
walkers map
work in
progress
1
382

Some plots - B

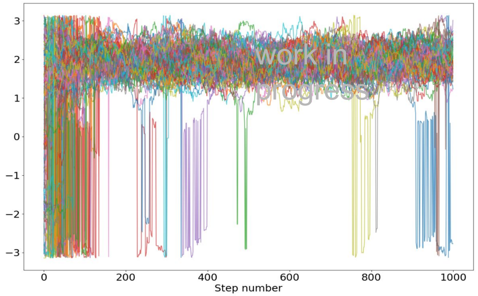

True and Predicted phis
Predicted
True Phi
Final Phi
Final Phi + Std
Final Phi - Std
work in progress
Phi [Rad]


True and Predicted thetas
Predicted
True Theta
Final Theta
Final Theta + Std
Final Theta - Std
work in progress
Theta [Rad]

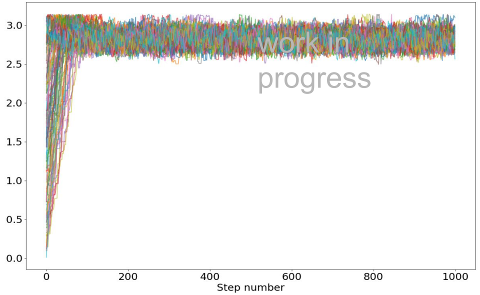

map_predictions
work in progress
-0.000101243
0.165633


map_true
work in progress
-0.00177624
3.00183

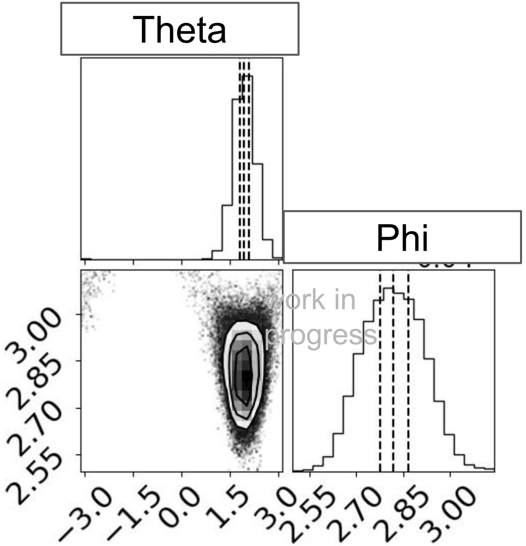
Some plots - B



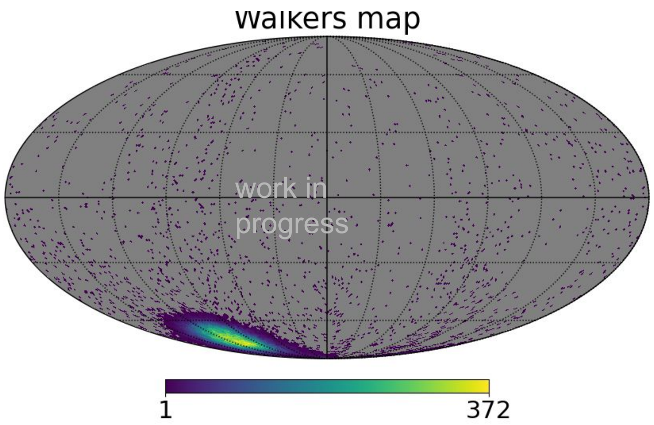
work in progress
Step number
0
200
400
600
800
1000



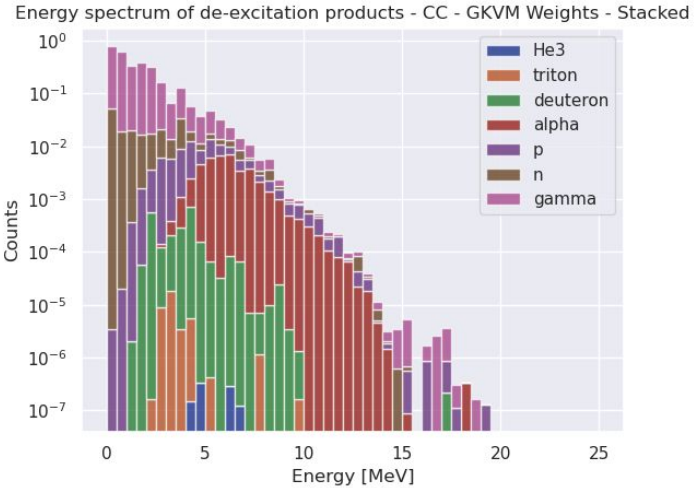
work in progress
Step number
0
200
400
600
800
1000

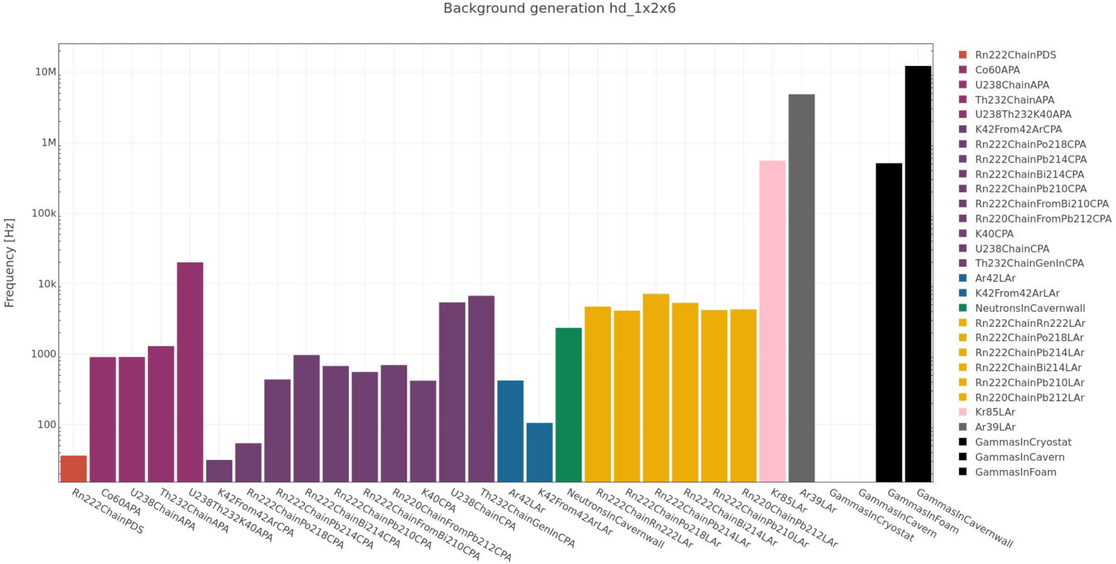


Theta
Phi
work in progress



walkers map
work in progress
1
372


Energy spectrum of de-excitation products - CC - GKVM Weights - Stacked
He3
triton
deuteron
alpha
p
n
gamma
Counts
Energy [MeV]


Background generation hd_1x2x6
Frequency [Hz]
Rn222ChainPDS
Co60APA
U238ChainAPA
Th232ChainAPA
U238Th232K40APA
K42From42ArCPA
Rn222ChainPo218CPA
Rn222ChainPb214CPA
Rn222ChainBi214CPA
Rn222ChainPb210CPA
Rn222ChainFromBi210CPA
Rn220ChainFromPb212CPA
K40CPA
U238ChainCPA
Th232ChainGenInCPA
Ar42LAr
K42From42ArLAr
NeutronsInCavernwall
Rn222ChainRn222LAr
Rn222ChainPo218LAr
Rn222ChainPb214LAr
Rn222ChainBi214LAr
Rn222ChainPb210LAr
Rn220ChainPb212LAr
Kr85LAr
Ar39LAr
GammasInCryostat
GammasInCavern
GammasInFoam
GammasInCavernwall

Available Information:

- Collection plane: X coordinate with sign, Z coordinate.
- Induction planes: Unsigned X coordinate (wires coiled around the APA), Y as a function of Z.

I test the possible combinations, computing the available information (using the Z from the collection cluster). If the clusters are within a tolerance in all directions, they can be considered compatible.

The true information is extracted from the collection plane cluster.

```
Compatible clusters
Event number: 9, Label: 100
Cluster U: X: -163, Y: -201.533148355818
Cluster V: X: -163, Y: -211.1213580950258
Cluster X: X: 163, Z: 69

Compatible clusters
Event number: 9, Label: 100
Cluster U: X: -162, Y: -209.53429704786925
Cluster V: X: -164, Y: -203.9203242721797
Cluster X: X: 163, Z: 69

Compatible clusters
Event number: 9, Label: 100
Cluster U: X: -162, Y: -209.53429704786925
Cluster V: X: -163, Y: -211.1213580950258
Cluster X: X: 163, Z: 69

Compatible clusters
Event number: 10, Label: 100
Cluster U: X: -270, Y: 118.31656937050184
Cluster V: X: -271, Y: 119.07985899628545
Cluster X: X: 272, Z: 348
```

Frac is set to 0.5 by default, is changed if not:

0<FH<CH

$$\overline{AH} = time\_peak - time\_start$$

$$\overline{BH} = time\_over\_threshold - \overline{AH}$$

$$\overline{CH} = adc\_peak$$

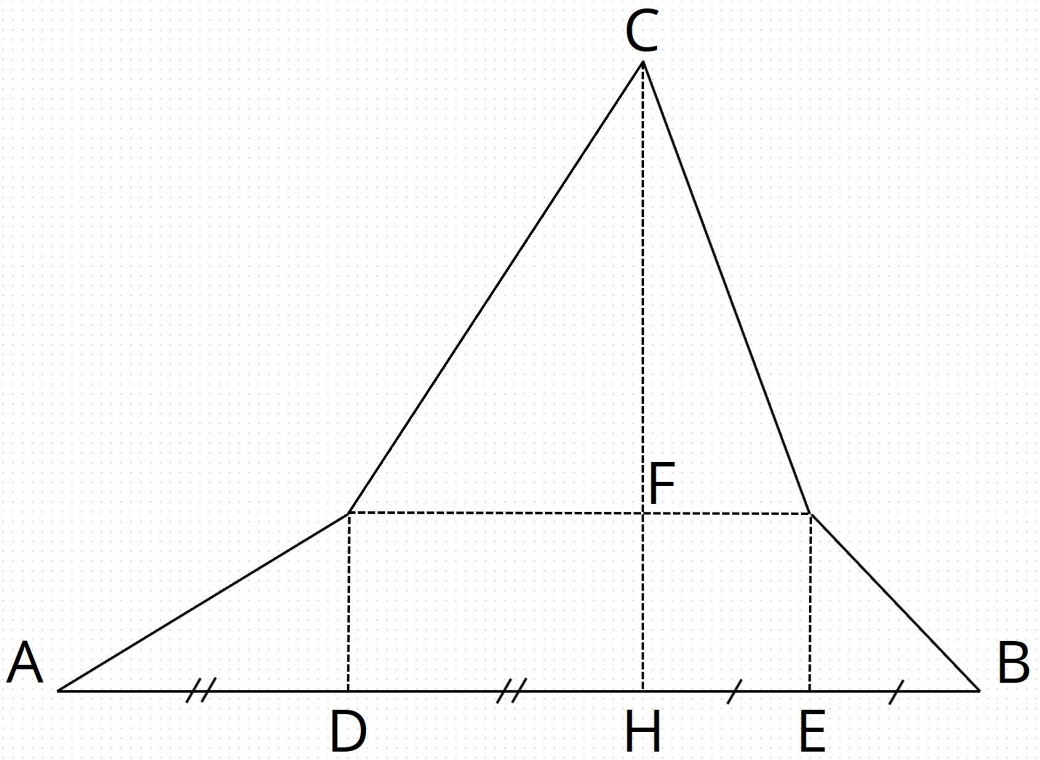
$$\overline{AD} = \overline{AH} \cdot frac$$

$$\overline{DH} = \overline{AH} - \overline{AD}$$

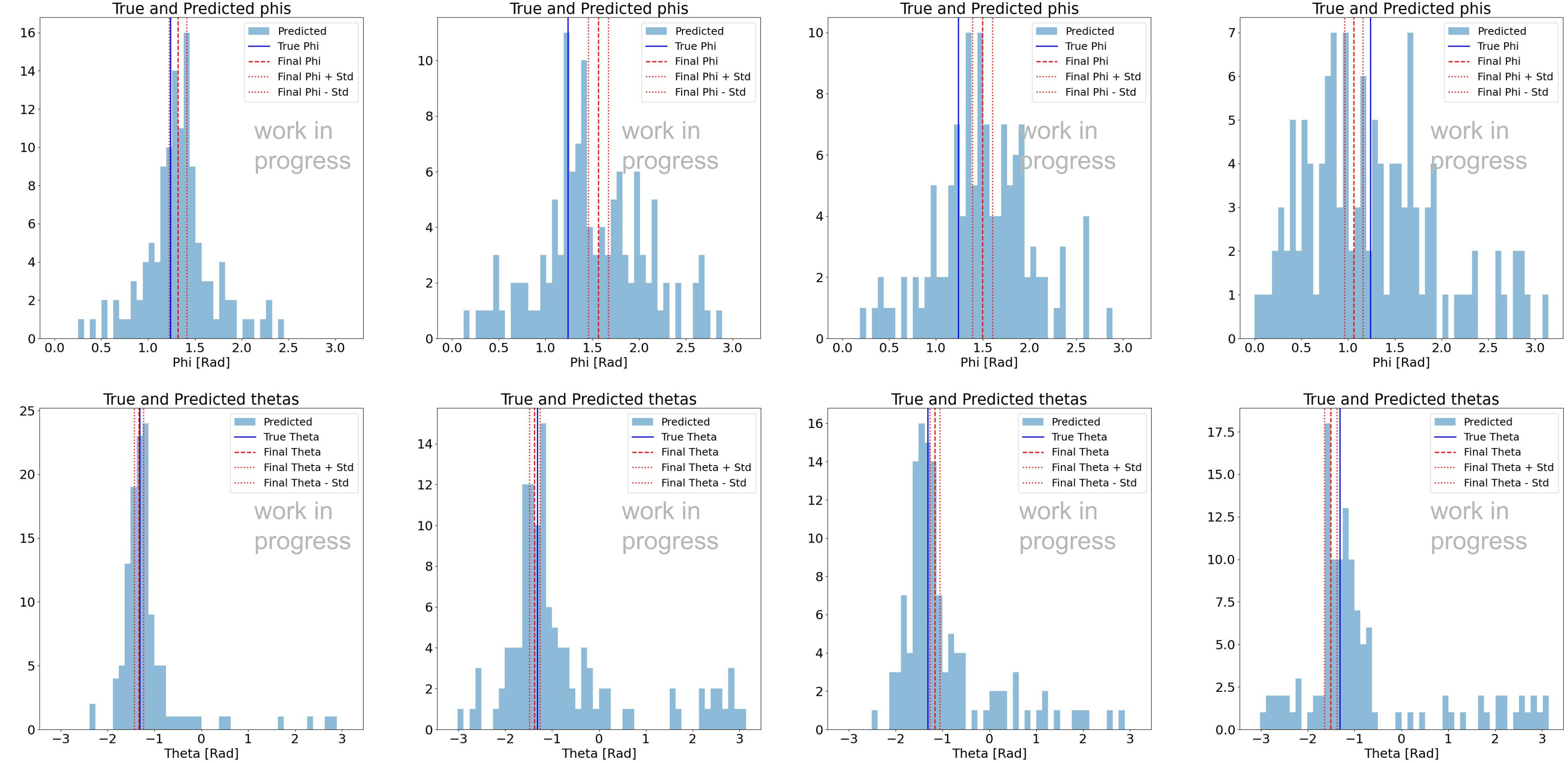
$$\overline{EB} = \overline{BH} \cdot frac$$

$$\overline{HE} = \overline{BH} - \overline{EB}$$

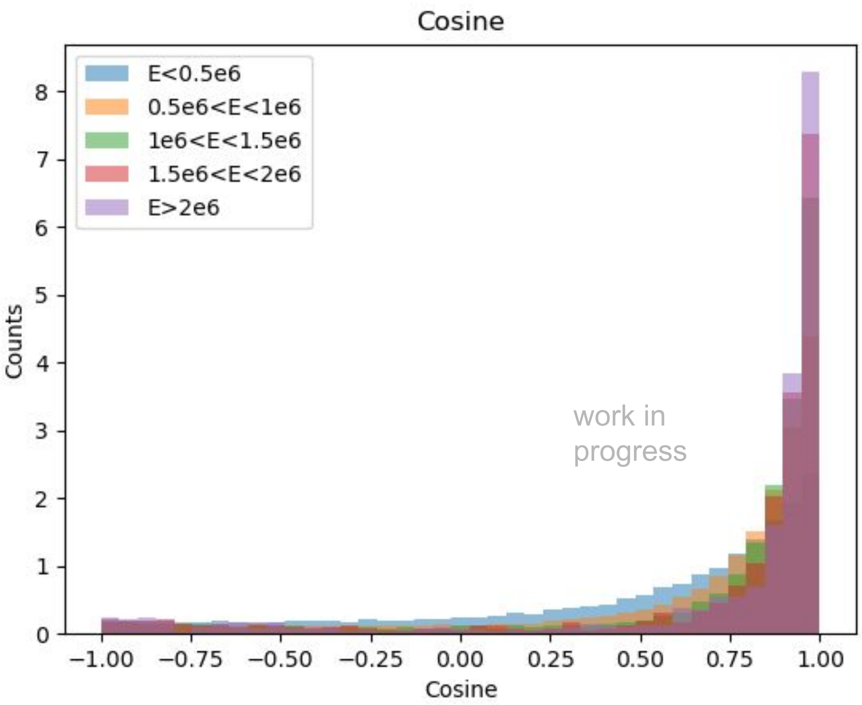
$$\overline{FH} = \frac{2 \cdot adc\_integral - \overline{CH} \cdot \overline{DH} - \overline{CH} \cdot \overline{HE}}{time\_over\_threshold}$$

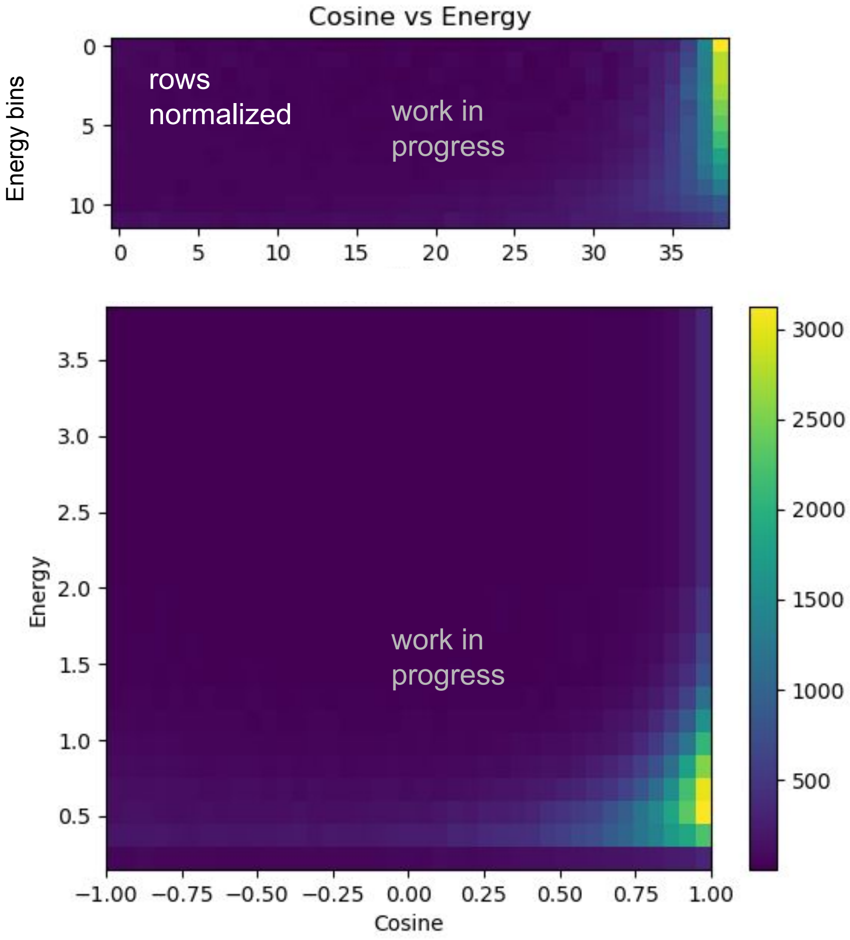


2442 all noU noV noX



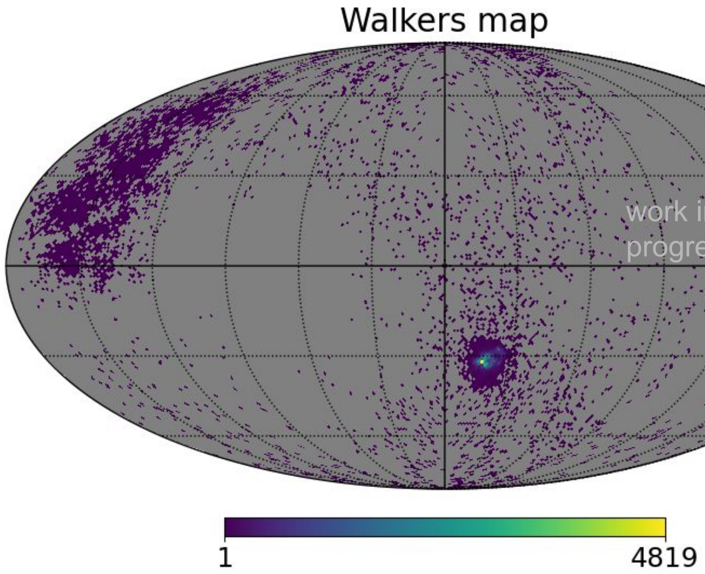
Likelihood with 2d pdf

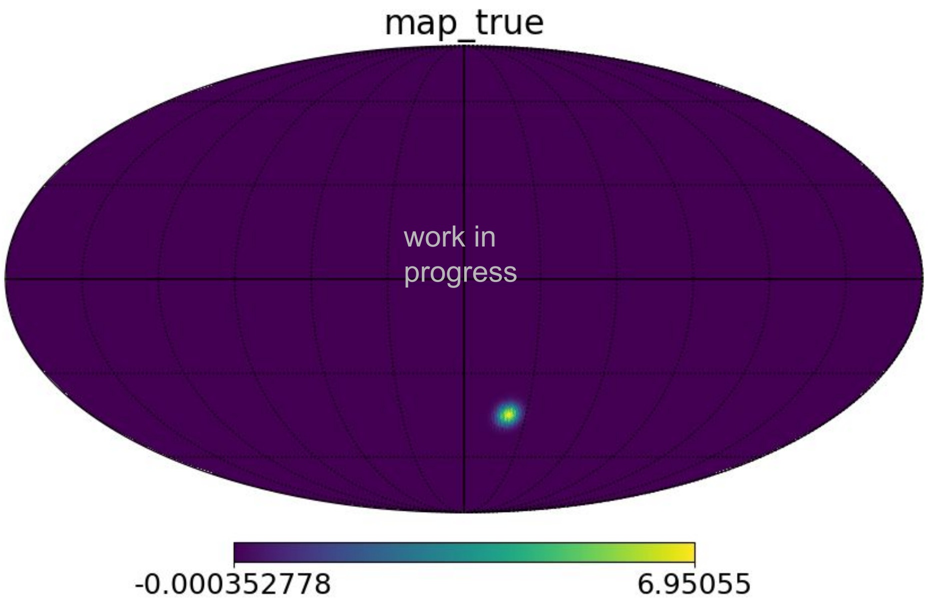




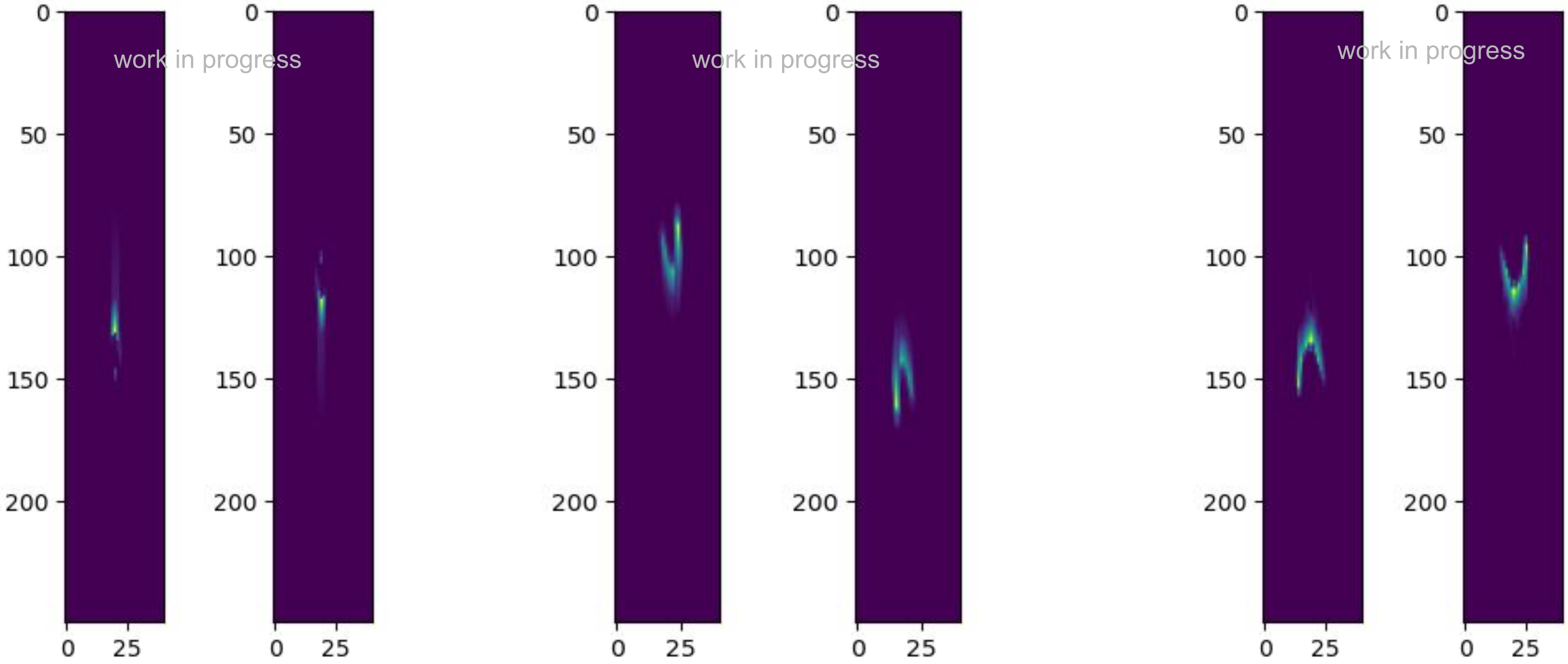
Likelihood with 2d pdf

Compute the log-likelihood as $log(L) = \sum_i log(pdf(E_i, cos_i))$

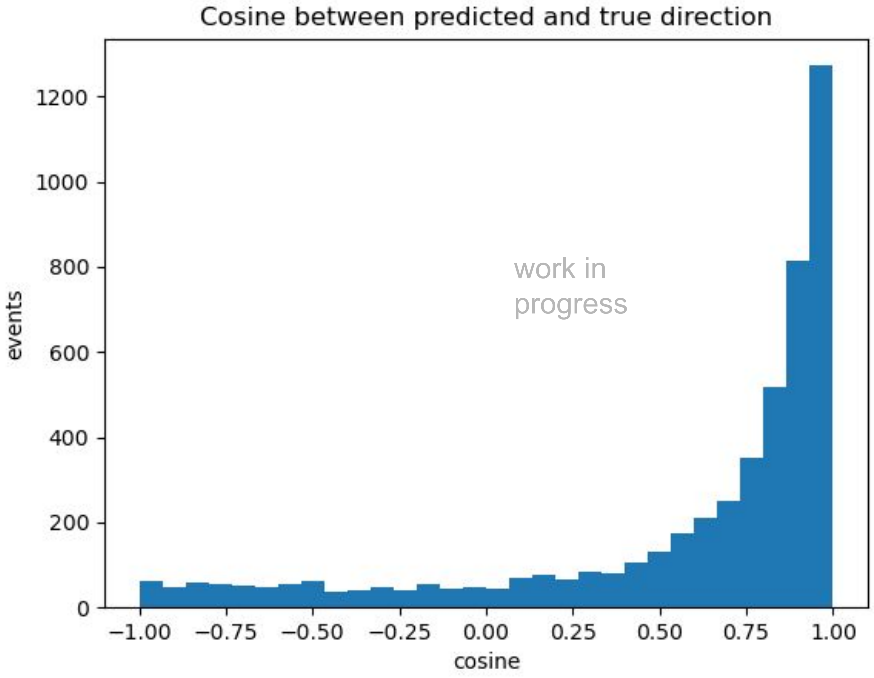


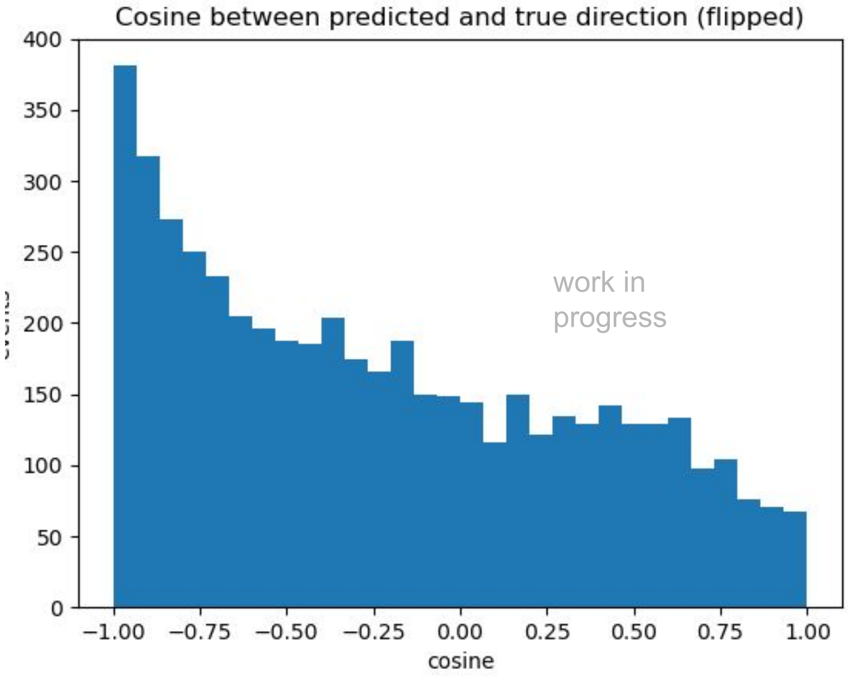


Is flipping a good metric to understand biases? No.

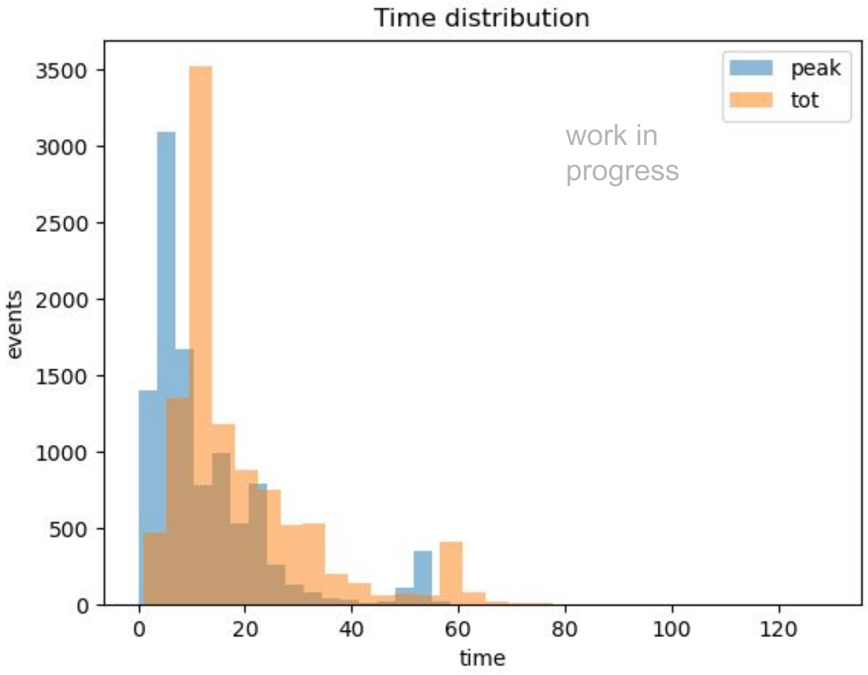


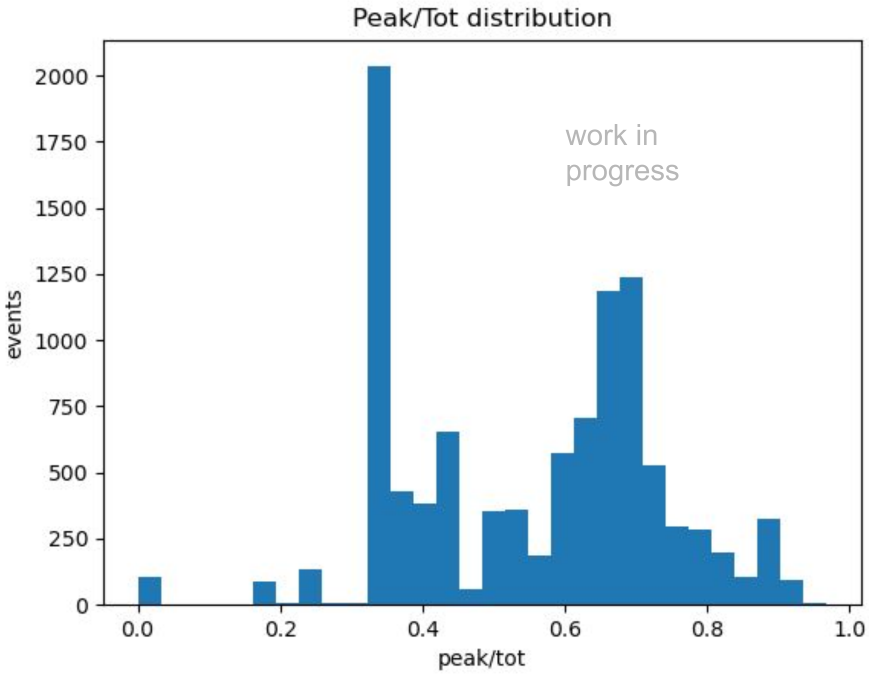
Is flipping a good metric to understand biases? No.





Why? Images not symmetric





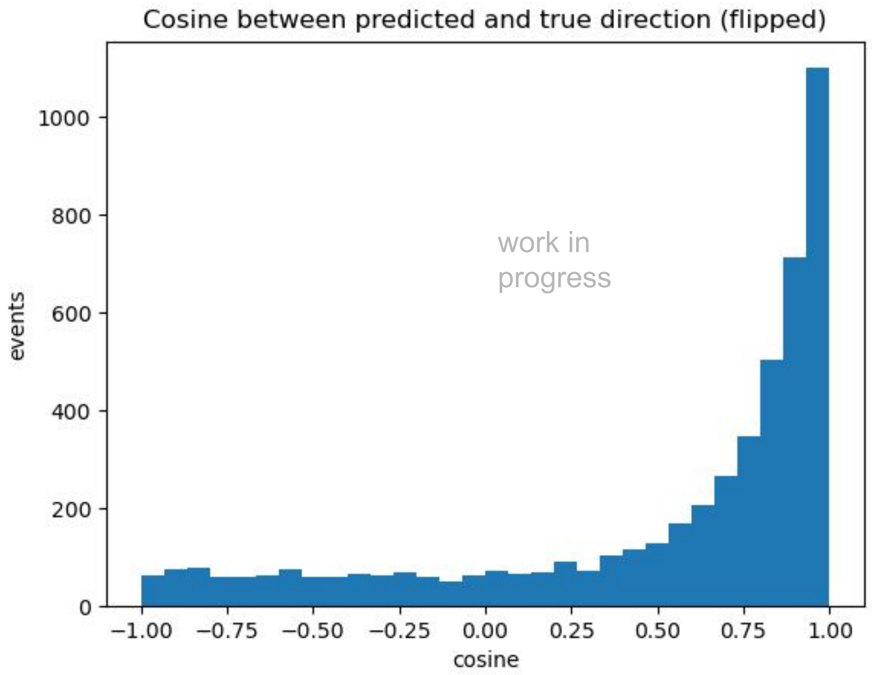
Flip only the channel axis

Flip only channel axis and change coordinates accordingly.

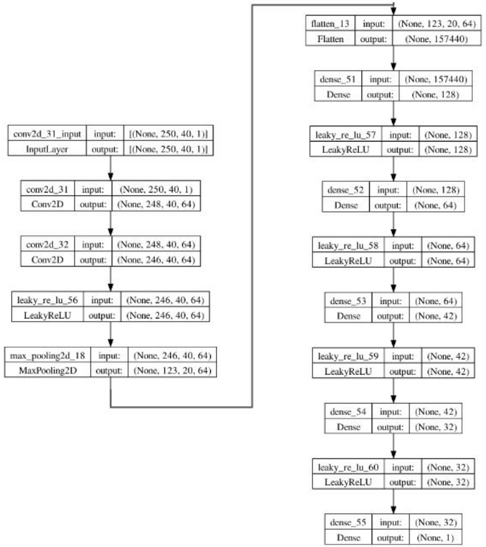
x -> x

y -> -y

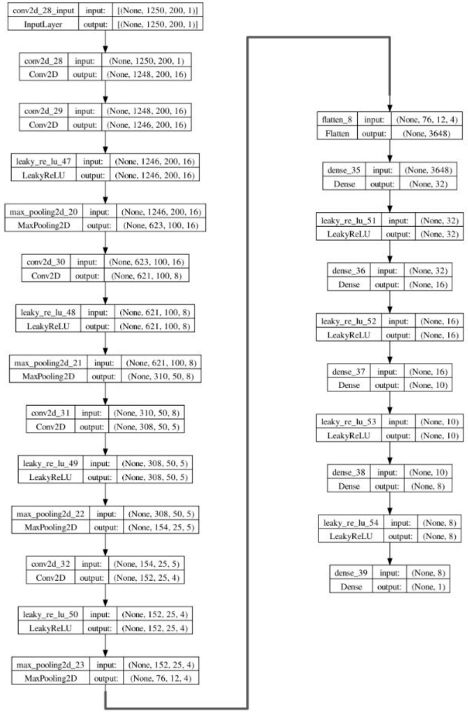
z -> -z

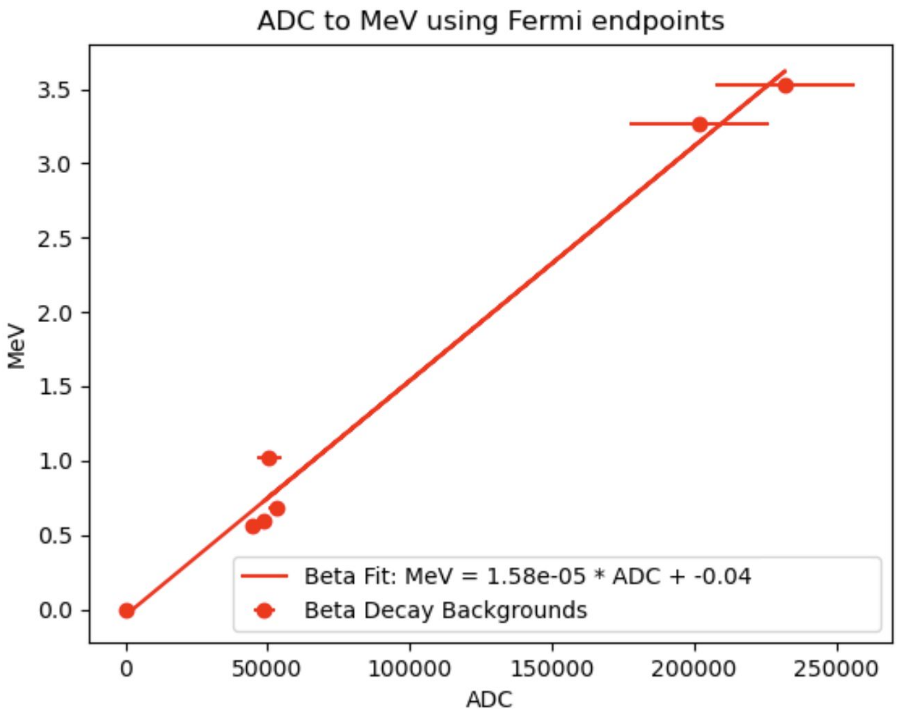


MT ID


conv2d_31_input | InputLayer | input: [(None, 250, 40, 1)] | output: [(None, 250, 40, 1)]
conv2d_31 | Conv2D | input: (None, 250, 40, 1) | output: (None, 248, 40, 64)
conv2d_32 | Conv2D | input: (None, 248, 40, 64) | output: (None, 246, 40, 64)
leaky_re_lu_56 | LeakyReLU | input: (None, 246, 40, 64) | output: (None, 246, 40, 64)
max_pooling2d_18 | MaxPooling2D | input: (None, 246, 40, 64) | output: (None, 123, 20, 64)
flatten_13 | Flatten | input: (None, 123, 20, 64) | output: (None, 157440)
dense_51 | Dense | input: (None, 157440) | output: (None, 128)
leaky_re_lu_57 | LeakyReLU | input: (None, 128) | output: (None, 128)
dense_52 | Dense | input: (None, 128) | output: (None, 64)
leaky_re_lu_58 | LeakyReLU | input: (None, 64) | output: (None, 64)
dense_53 | Dense | input: (None, 64) | output: (None, 42)
leaky_re_lu_59 | LeakyReLU | input: (None, 42) | output: (None, 42)
dense_54 | Dense | input: (None, 42) | output: (None, 32)
leaky_re_lu_60 | LeakyReLU | input: (None, 32) | output: (None, 32)
dense_55 | Dense | input: (None, 32) | output: (None, 1)

ES-CC


conv2d_28_input | InputLayer | input: [(None, 1250, 200, 1)] | output: [(None, 1250, 200, 1)]
conv2d_28 | Conv2D | input: (None, 1250, 200, 1) | output: (None, 1248, 200, 16)
conv2d_29 | Conv2D | input: (None, 1248, 200, 16) | output: (None, 1246, 200, 16)
leaky_re_lu_47 | LeakyReLU | input: (None, 1246, 200, 16) | output: (None, 1246, 200, 16)
max_pooling2d_20 | MaxPooling2D | input: (None, 1246, 200, 16) | output: (None, 623, 100, 16)
conv2d_30 | Conv2D | input: (None, 623, 100, 16) | output: (None, 621, 100, 8)
leaky_re_lu_48 | LeakyReLU | input: (None, 621, 100, 8) | output: (None, 621, 100, 8)
max_pooling2d_21 | MaxPooling2D | input: (None, 621, 100, 8) | output: (None, 310, 50, 8)
conv2d_31 | Conv2D | input: (None, 310, 50, 8) | output: (None, 308, 50, 5)
leaky_re_lu_49 | LeakyReLU | input: (None, 308, 50, 5) | output: (None, 308, 50, 5)
max_pooling2d_22 | MaxPooling2D | input: (None, 308, 50, 5) | output: (None, 154, 25, 5)
conv2d_32 | Conv2D | input: (None, 154, 25, 5) | output: (None, 152, 25, 4)
leaky_re_lu_50 | LeakyReLU | input: (None, 152, 25, 4) | output: (None, 152, 25, 4)
max_pooling2d_23 | MaxPooling2D | input: (None, 152, 25, 4) | output: (None, 76, 12, 4)
flatten_8 | Flatten | input: (None, 76, 12, 4) | output: (None, 3648)
dense_35 | Dense | input: (None, 3648) | output: (None, 32)
leaky_re_lu_51 | LeakyReLU | input: (None, 32) | output: (None, 32)
dense_36 | Dense | input: (None, 32) | output: (None, 16)
leaky_re_lu_52 | LeakyReLU | input: (None, 16) | output: (None, 16)
dense_37 | Dense | input: (None, 16) | output: (None, 10)
leaky_re_lu_53 | LeakyReLU | input: (None, 10) | output: (None, 10)
dense_38 | Dense | input: (None, 10) | output: (None, 8)
leaky_re_lu_54 | LeakyReLU | input: (None, 8) | output: (None, 8)
dense_39 | Dense | input: (None, 8) | output: (None, 1)



ADC to MeV using Fermi endpoints
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
MeV
0
50000
100000
150000
200000
250000
ADC
Beta Fit: MeV = 1.58e-05 * ADC + -0.04
Beta Decay Backgrounds