

# Enhancing the Analysis Pipeline of Future Gamma-ray Space Missions with Machine-Learning Approaches



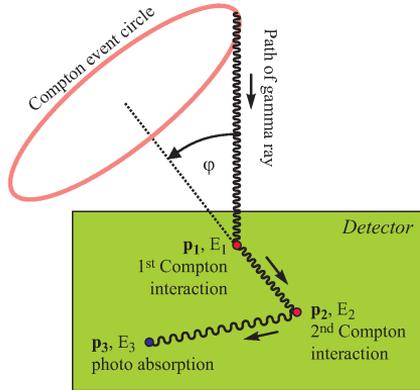
**Andreas Zoglauer**

Space Sciences Laboratory, UC Berkeley

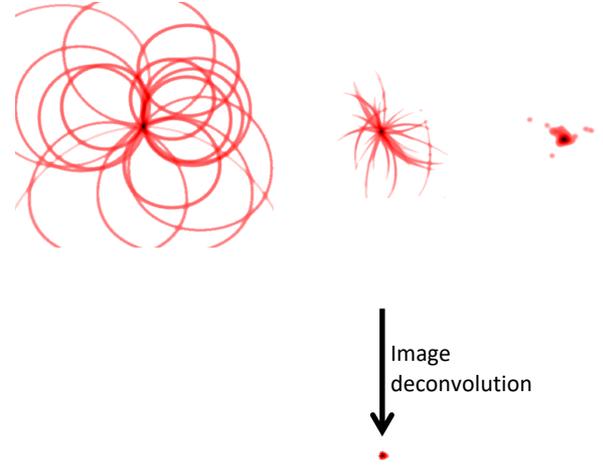
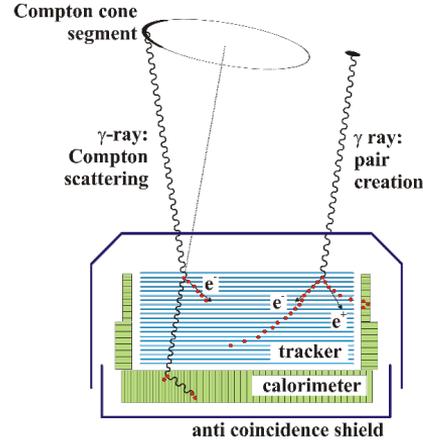


# Operating Principles of Future Gamma-ray Missions

Non-electron-trackers: COSI, GRAMS, GammaTPC, etc.



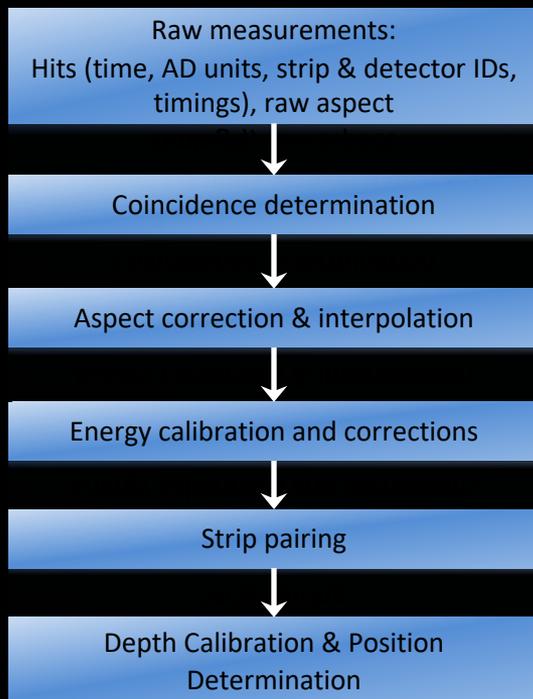
Electron-trackers: AMEGO(-X), ComPair, (e)ASTROGAM, SMILE, HARPO, etc.



- Gamma rays interact multiple times in the active detector(s) via Compton scatters (and/or pair creations).
- The interaction sequence has to be determined from information such as scatter angles, absorption probabilities, scatter probabilities.
- The origin of a single not-tracked Compton event can be restricted to the so called "event circle", an electron-tracked event to an arc, and a pair event to an extended "blob" which becomes a dot at higher energies.
- The position of peak overlap is the origin.
- Image deconvolution creates sky maps.

# Simplified Data-Analysis Pipeline

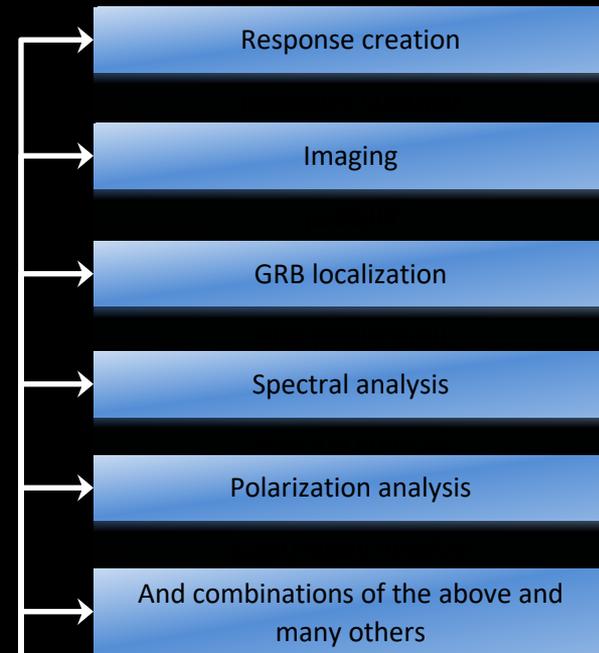
## Calibration



## Reconstruction



## Analysis



Many classification and regression tasks perfectly suited for machine learning

# Long-term Goal for Machine Learning

Go from:

- raw measurements to, for example, a localized transient (GRB)
- using a Compton and/or pair telescope
- on-board a space craft
- in close to real-time, and
- transmit data to other missions for follow ups.

Key Challenges:

- Need to develop radiation hard, powerful but power-efficient TPU-like system to go to space
- Need to develop multi-stage (i.e. deep) neural network architecture to do that
- Need a pipeline to robustly verify performance

# Development Steps of a ML approach

in a high-energy data-analysis pipeline

Stage	Description
1. Exploration	Play with different ML approaches and try to get one to work as well or better than already existing approaches.
2. Limited proof-of-concept	ML approach works on average for a limited simulated data set (e.g., toy-model, one energy, extremely cleaned data set, etc.)
3. Full proof-of-concept	ML approach works on average for a complete, simulated data set spanning the expected data range
4. Verification with simulations	Physics-based verification for all initial and measured parameters: Ensure there are no detectable biases with simulated datasets
5. Verification with measurements	Ensure there are no detectable biases in real data
6. Deployment	Ready for application to your astrophysical observations

# Software Tools (used for COSI)



## MEGALib

Medium-Energy  
Gamma-ray  
Astronomy library

*A. Zoglauer et al. 2006*



## ROOT

CERN's high-  
energy physics  
data analysis  
framework

*R. Brun & F.  
Rademakers, 1997*



## TMVA

Toolkit for  
Multivariate Data  
Analysis

*P. Speckmayer et al.  
2010*



## TensorFlow

Google's  
machine-learning  
library

*M. Abadi et al. 2016*



## Keras

User friendly  
front-end for  
TensorFlow

*F. Chollet et al. 2015*

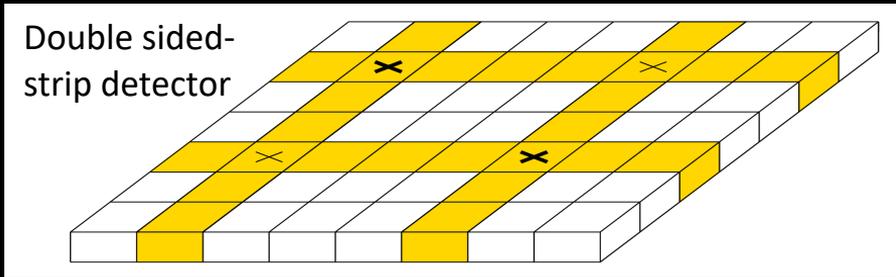
All are freely available & open source

# COSI: Strip Pairing

Stage 2: Limited  
proof-of-concept

## Task:

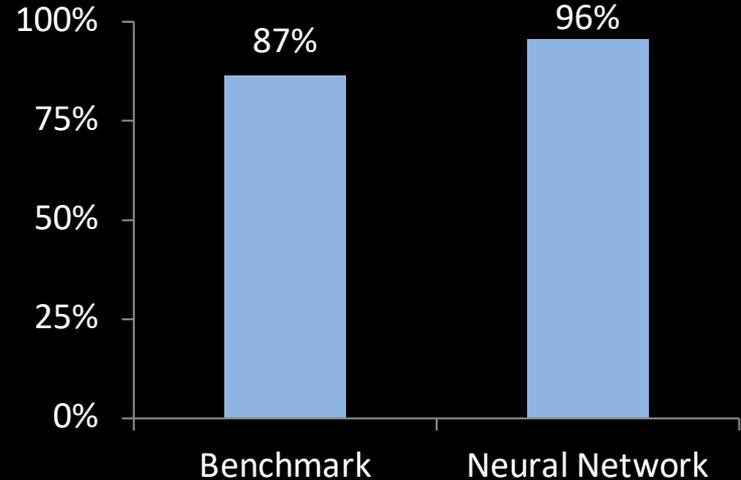
Find interaction locations in the (double-sided strip) detectors from the triggered strips



**Yellow:** Hit strips  
**x:** Possible interaction locations  
**X:** Real interaction locations

## Results:

Benchmark (chi-square approach) vs.  
4-layer fully connected neural  
network:

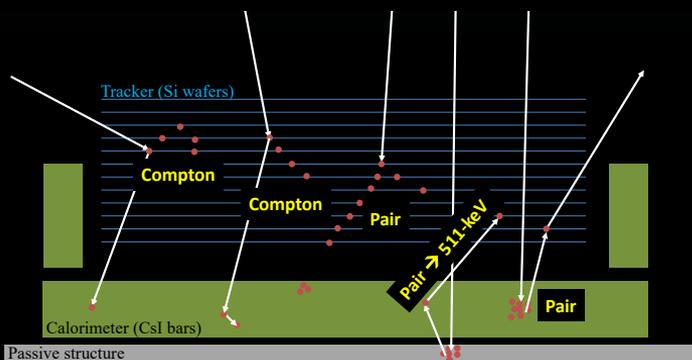


# Si-tracker: Event-type Identification

Stage 3: Full proof-of-concept

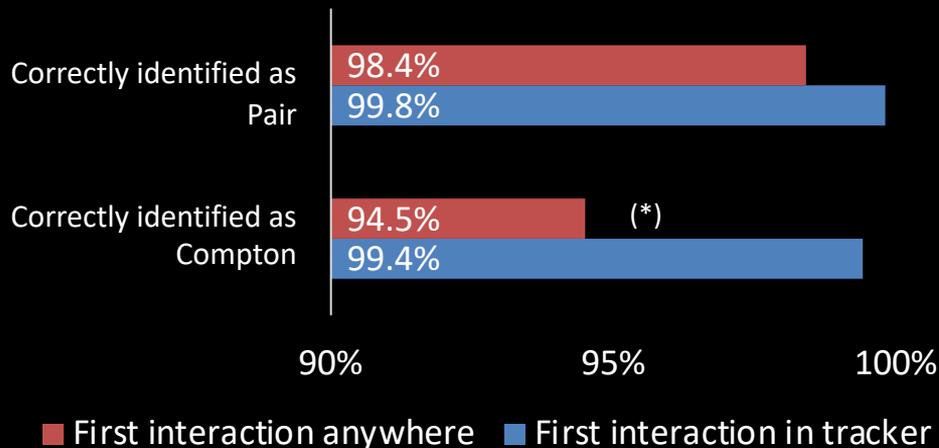
## Task:

Find the event type:  
Compton, pair, single-site,  
charged particle



## Results:

Apply a 3D convolutional neural network (CNN)  
with 4 convolutional + 2 dense layers:



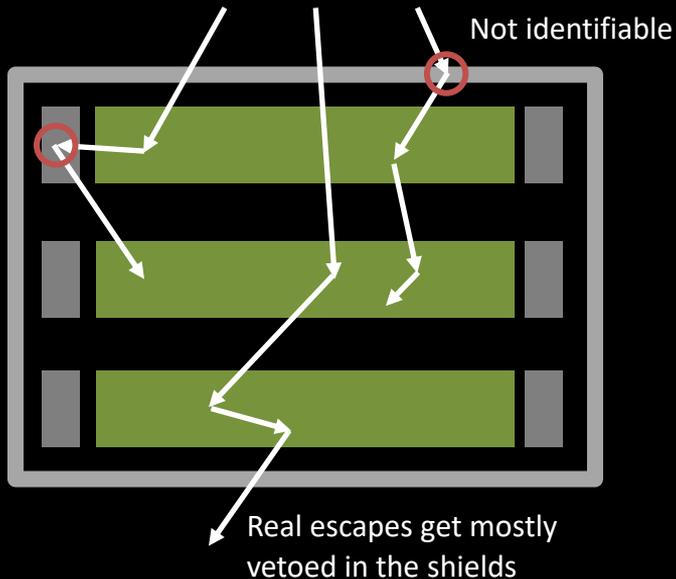
(\*) The majority of the missing 5.5% are mostly 511's from pair events in passive material (correctly) identified as Compton events

# COSI: Identification of Energy Leakage

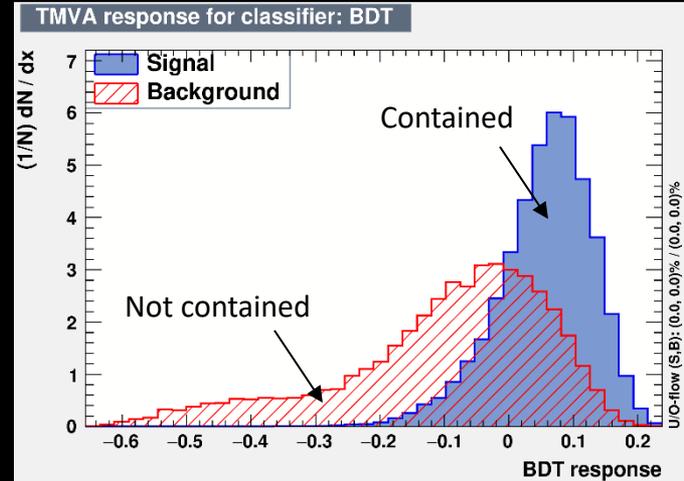
Stage 3: Full  
proof-of-concept

## Task:

Identify events which are not completely contained in active detector material (and do not raise a veto)



## Results:

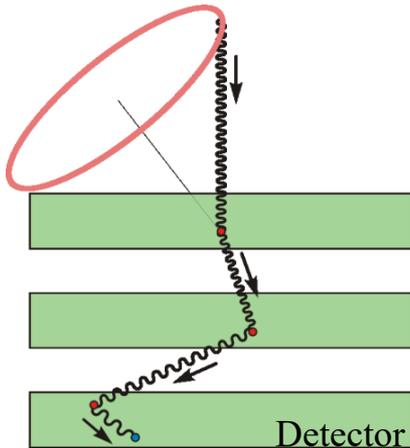
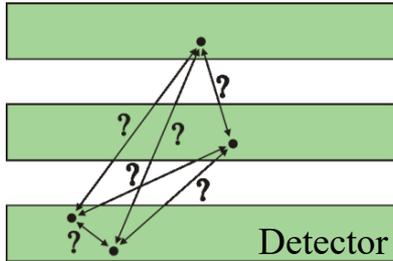


## Best separation:

- Random forest consisting of 500 boosted decision trees
- Neural network gave worse performance

# COSI: Event Reconstruction

Stage 5: Partially validated  
with measurements



## Problem:

- Detector only measures hits without time information  
→ Path of photon is unknown!

## Information available to determine path:

- The kinematics of the events, i.e., the redundant Compton scatter information (electron track and/or multiple Compton interactions)
- The known response of the detector to incident gamma-ray (absorption probabilities, scatter probabilities, etc.)

## Techniques used:

- Classic Compton sequence reconstruction (*Boggs+ 2000*)
- Naïve Bayesian approach (*Zoglauer+, 2005 & 2007*)
- Random forest of boosted decision trees (this)
- Neural network (this & *Zoglauer+ 2007*)

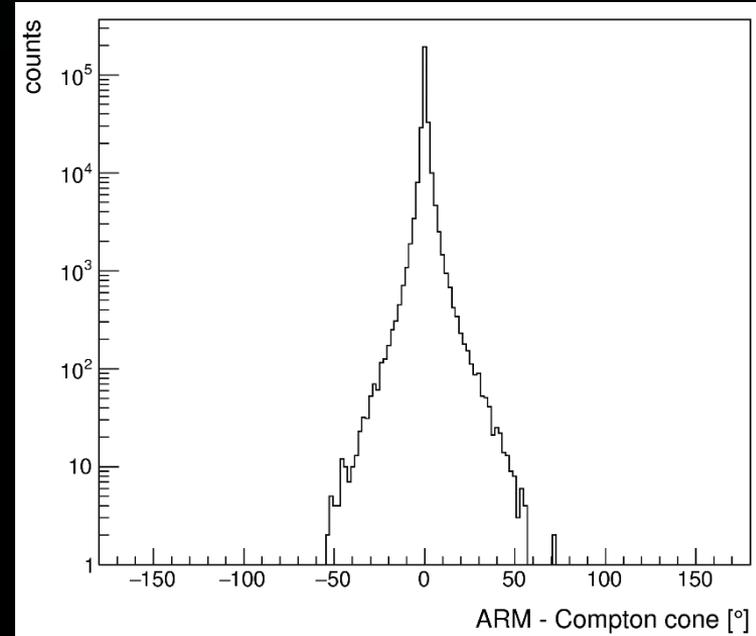
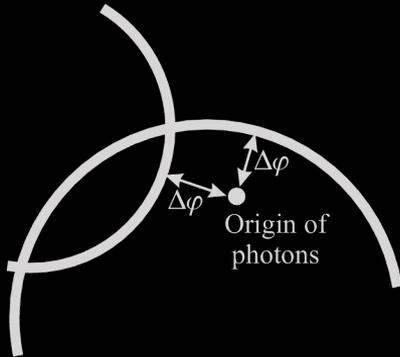
## Output:

- Correct interaction sequence

# Event Reconstruction Performance Metric: ARM

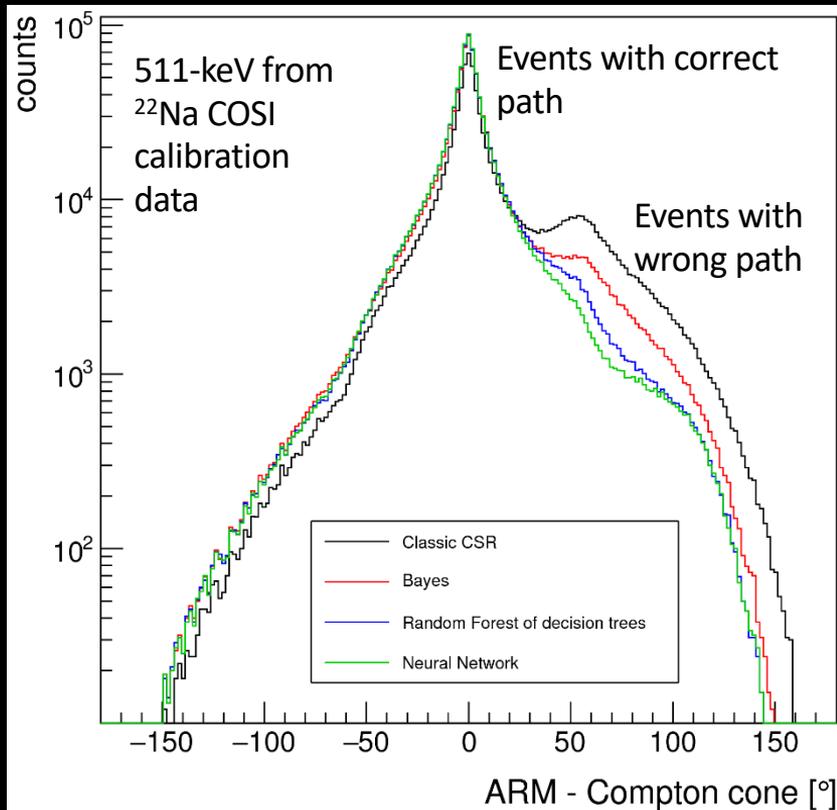
**ARM:**  
**Angular Resolution Measure**

Smallest distance between  
Compton cone and known origin  
of gamma rays



ARM of ideal detector: Only Doppler-broadening  
Reality: broadening by detector position and energy resolution

# Event Reconstruction Performance



## RMS improvement:

CSR: 0%

Bayes: 17.0%

RF-BDT: 24.9%

NN: 27.4%

## Neural network:

- Configuration: Dense, 1 hidden layer, 30% more hidden layer nodes than input layer nodes
- Testing: Look at each possible path
- Output: Highest rated path is chosen as correct one

# COSI: Imaging Response Approximation

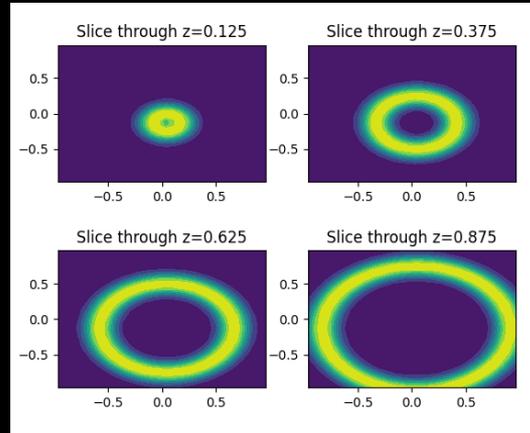
Stage 2: Limited  
proof-of-concept

## Task:

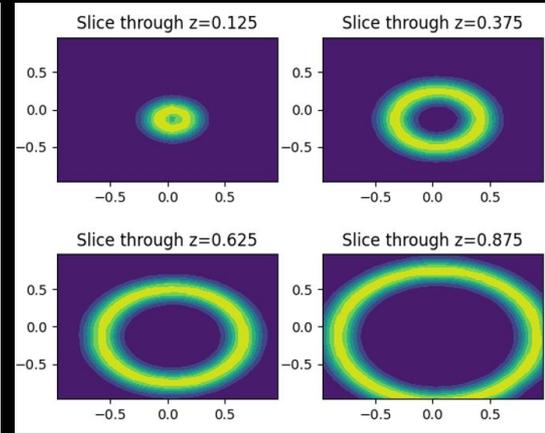
Simulating a 10-dimensional Compton imaging response requires significant supercomputing time. Can we approximate it with a neural network?

## Results:

Train  $\sim 1000$  locations in the sky with a 2-step CNN+autoencoder neural network. Test performance with a new, untrained location.

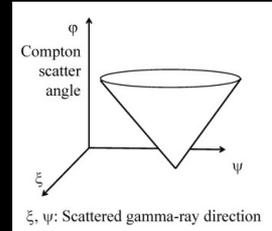


Test data: Gaussian approximated response at an untrained location



Neural network output for that new, untrained location on the sky

These are slices of the Compton cone



# Simple Compton Telescope: GRB Localization

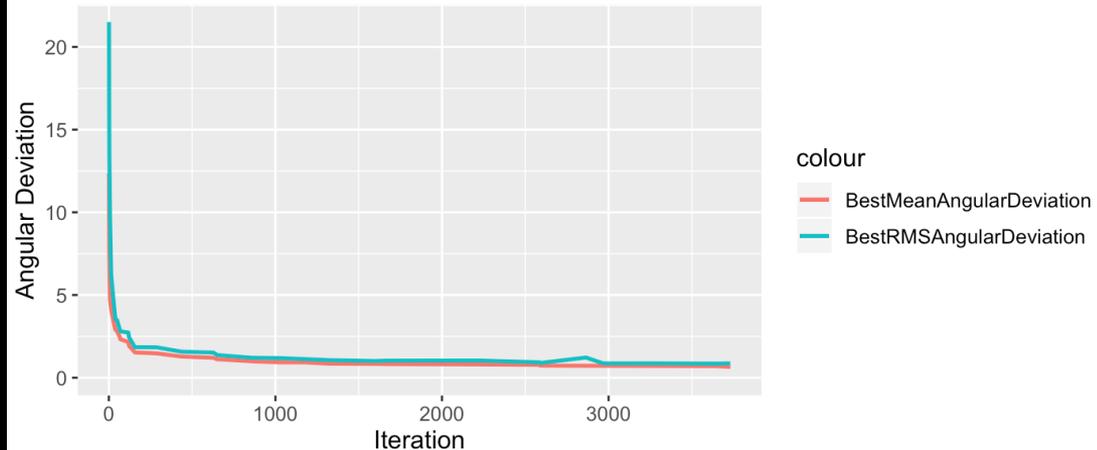
Stage 2: Limited  
proof-of-concept

## Task:

Localize GRBs without imaging using a 3D convolutional neural network with the data binned in the Compton data space in close to real-time, possibly onboard the space craft.

## Result:

Angular deviation (mean and RMS) as a function of iterations on the neural network

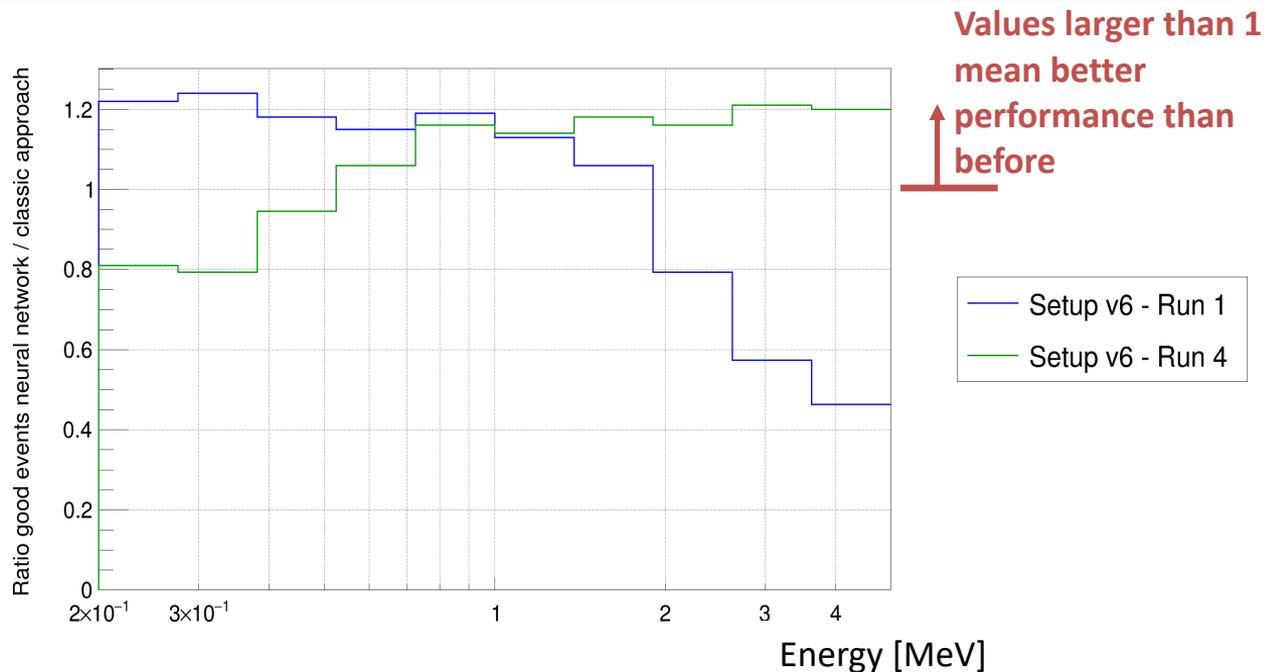


Best mean angular deviation:  
0.65 degrees for 5-degree data  
space bins

Limited proof-of-concept:

- no background
- mono-energetic

# Things can go wrong during training



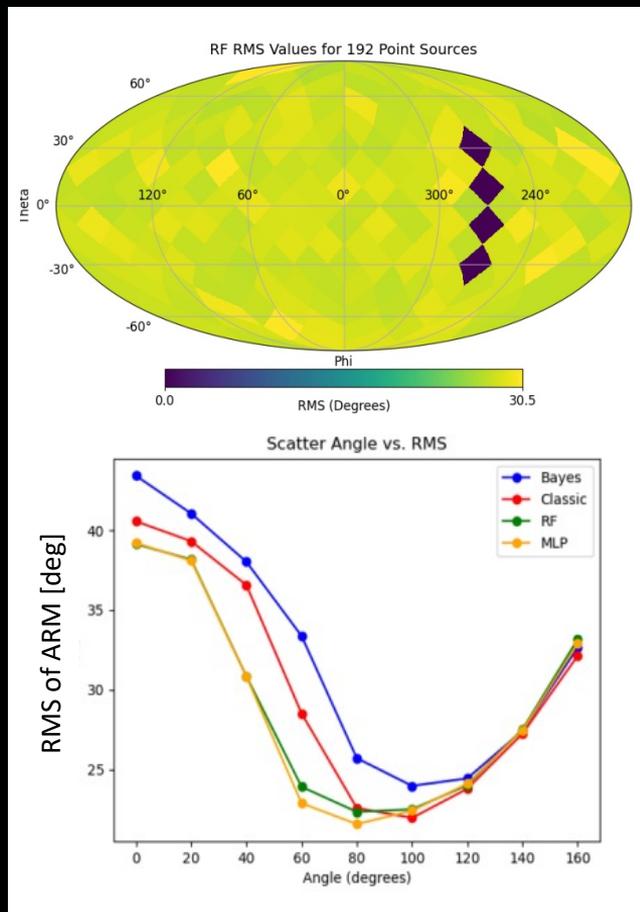
Event reconstruction performance during initial testing: Significant changes in the underlying cross-sections and not enough data make the networks either learn the high-energy or the low-energy part very well.

**Always perform stages 4 & 5:  
Full Verification with simulations and real measurements required  
to avoid biases = “blind spots” for gamma-ray astronomy**

# Verification with Simulations & Calibrations

Verify performance as a function of all physical parameters:

- the incoming gamma-ray parameters:
  - the energy of the incident gamma ray
  - the direction of the incoming gamma ray
- the measured gamma-ray parameters:
  - the number of interactions
  - the measured energies ranges
  - the Compton scatter angles
- other important physical effects in the detector and its material:
  - Rayleigh scatters in between Compton scatters
  - charged particle range in the detector
  - events with high Doppler broadening
- the origin of the gamma rays (astrophysical, internal decay, atmospheric)



Field-of-view validation: Identical performance within statistics (blue is software issue)

Scatter-angle validation: The overall shape of the curve is determined by Compton kinematics. Lower values indicate better performance.

# Some High-Level Lessons Learned For The Future

- 1. Machine learning approaches will be an essential part of future gamma-ray missions:** They show significant performance improvements compared to existing approaches in select elements of the pipeline
- 2. Start simple when developing an ML approach:** For example, one energy, simplified geometry, and then make it more complex / realistic step by step.
- 3. Reaching deployment is very significantly more time and resource intensive than the first proof of concept:** Reaching stage 6 is hard - many published results you might find are just “proof of concepts”
- 4. During the development phase of an instrument, it is better to use conventional (non ML) approaches:** You do not want to have to train and validate a whole pipeline of ML tools each time something changes in your detector
- 5. Data Cleaning And Data Selection:** Can be 1/3 of the work: Noise effects can throw the learning off, and it is sometimes hard for the approach to learn all the physics.
- 6. ML approaches can develop biases / blind spots:** Do not trust them until you have fully tested them in \*all\* aspects regarding all input, output, or other intermediate parameters
- 7. Simulation-calibration matching:** You need simulations to train the network. If your simulations are not well benchmarked to calibrations, you can not expect the ML approaches to give the same results
- 8. Verification:** Can be most of the work: You must ensure that the approach is working consistently with simulations, calibrations, and observations

# Technology Gaps

1. Machine-Learning approaches show very promising potential throughout the gamma-ray analysis pipeline: We need the resources to develop the ML approaches along with the verification pipeline to a stage that the ML approaches can be deployed - COSI will do part of it, but cannot do everything
2. If we want to do some of this in space: We need a dedicated radiation-hard, low-power AI-acceleration hardware (NPU/TPU+CPU) for that
3. Compile common gamma-ray analysis library of well tested non-ML & ML analysis tools to simplify development of future space missions. Since COSI will lead the development for the near future, I would vote for basing it on COSI's pipeline (MEGALib/COSIpy).