AI for WCDA

Improve WCDA performance using transformer-backbone AI models

Shicong Hu, Zhiguo Yao, Zijie Huang, Yong Huang | IHEP

Weiwei Li | Z.AI Co., Ltd.

XiaoXiao Wang, Bingke Zhu | IA, CAS

Chenyuan Xu | ZHEJIANG LAB

1st LHAASO General Meeting, 2026

@NJU, Suzhou

Outline

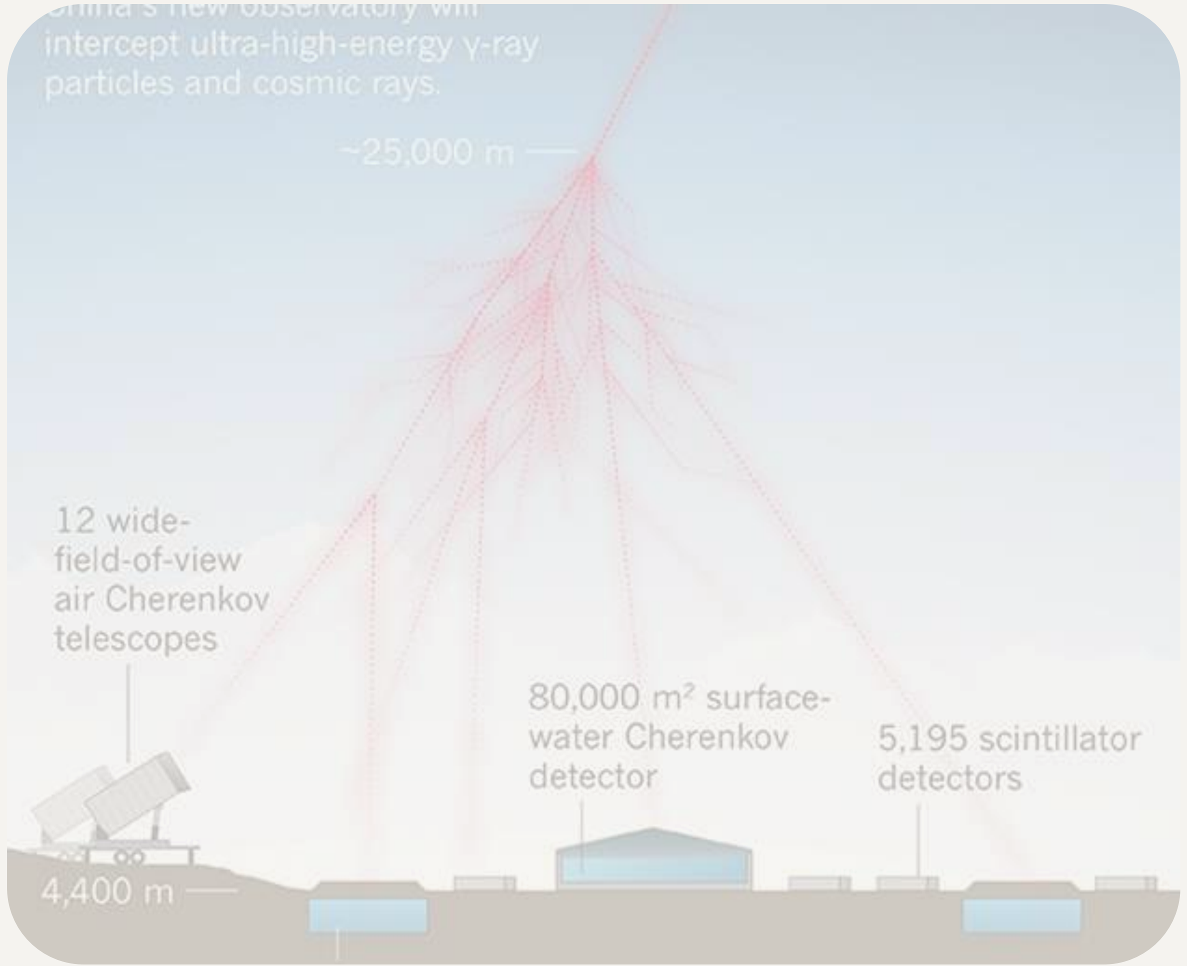
CONTENTS

01

CHAPTER 01

Background and Technical route

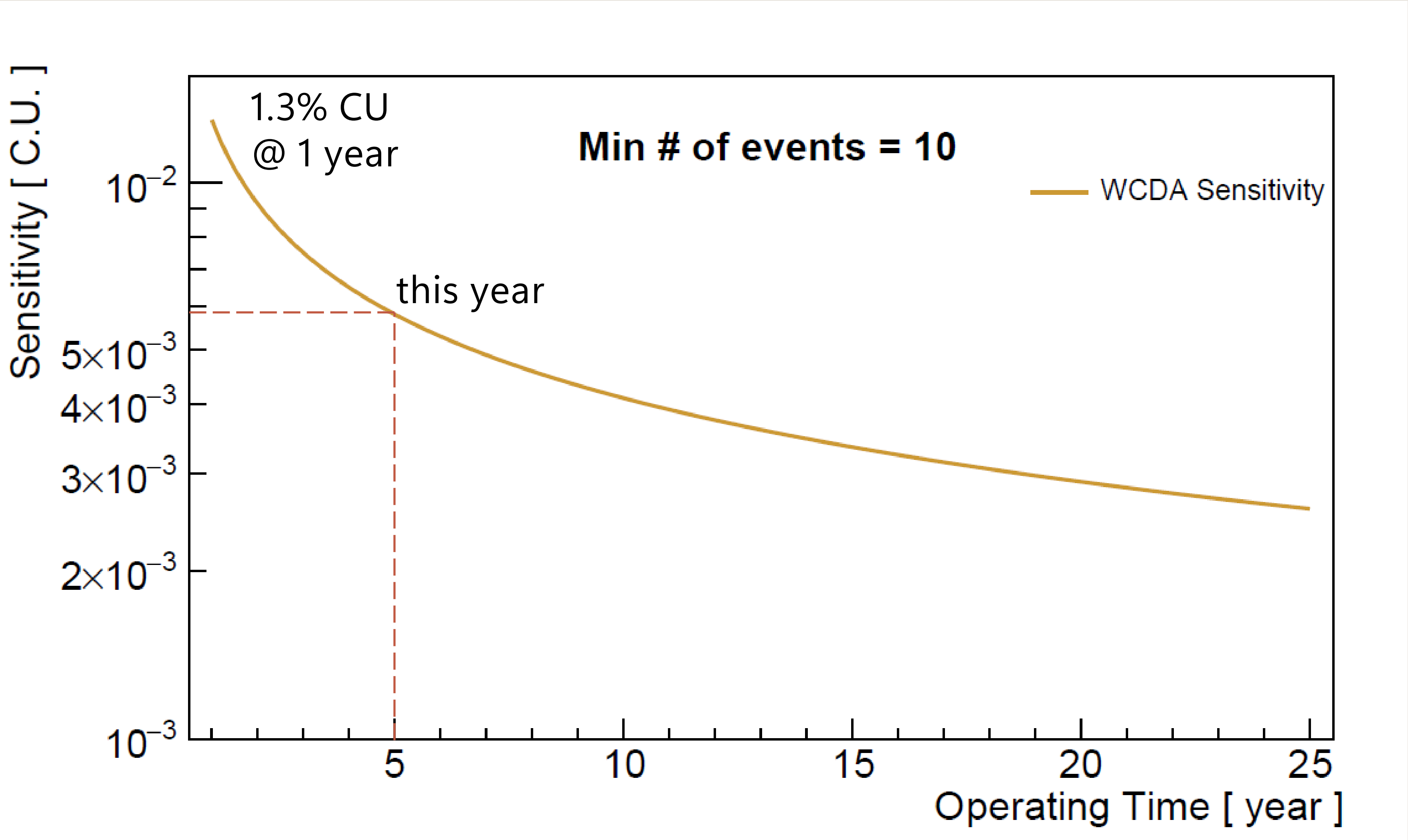
WCDA performance and AI



WCDA sensitivity

For steady source

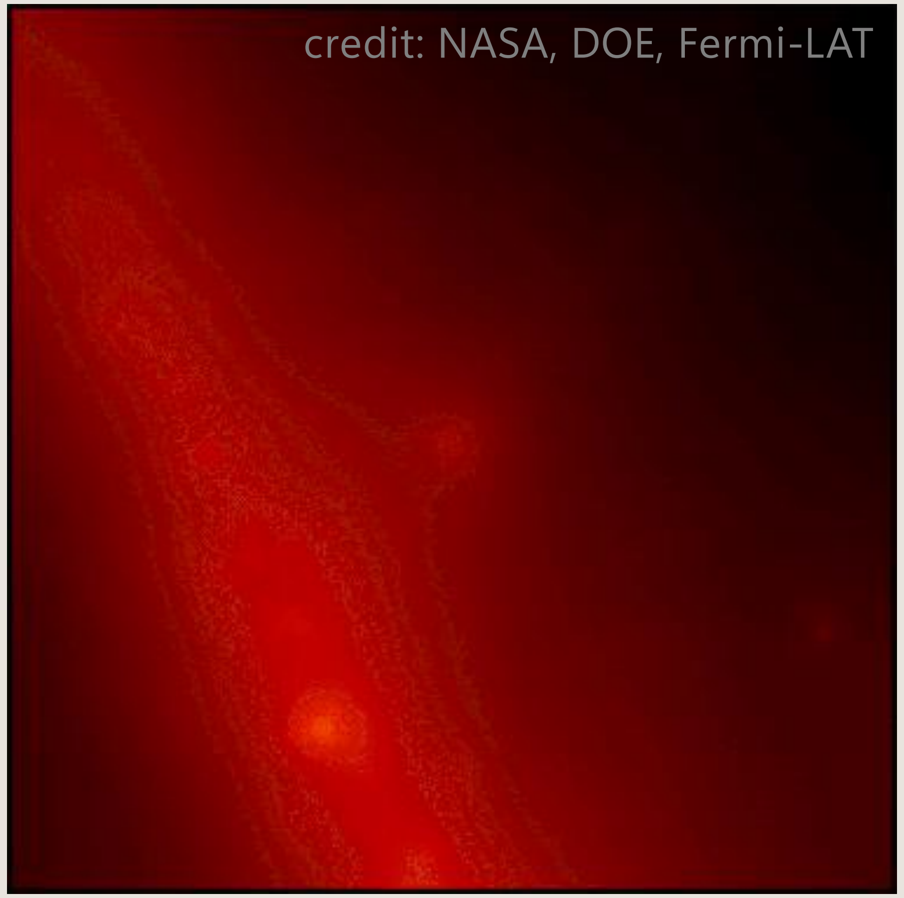
The increasement of sensitivity is slowing down as sqrt(T).



For transient source

Sensitivity for short time-scale transient source not particularly good

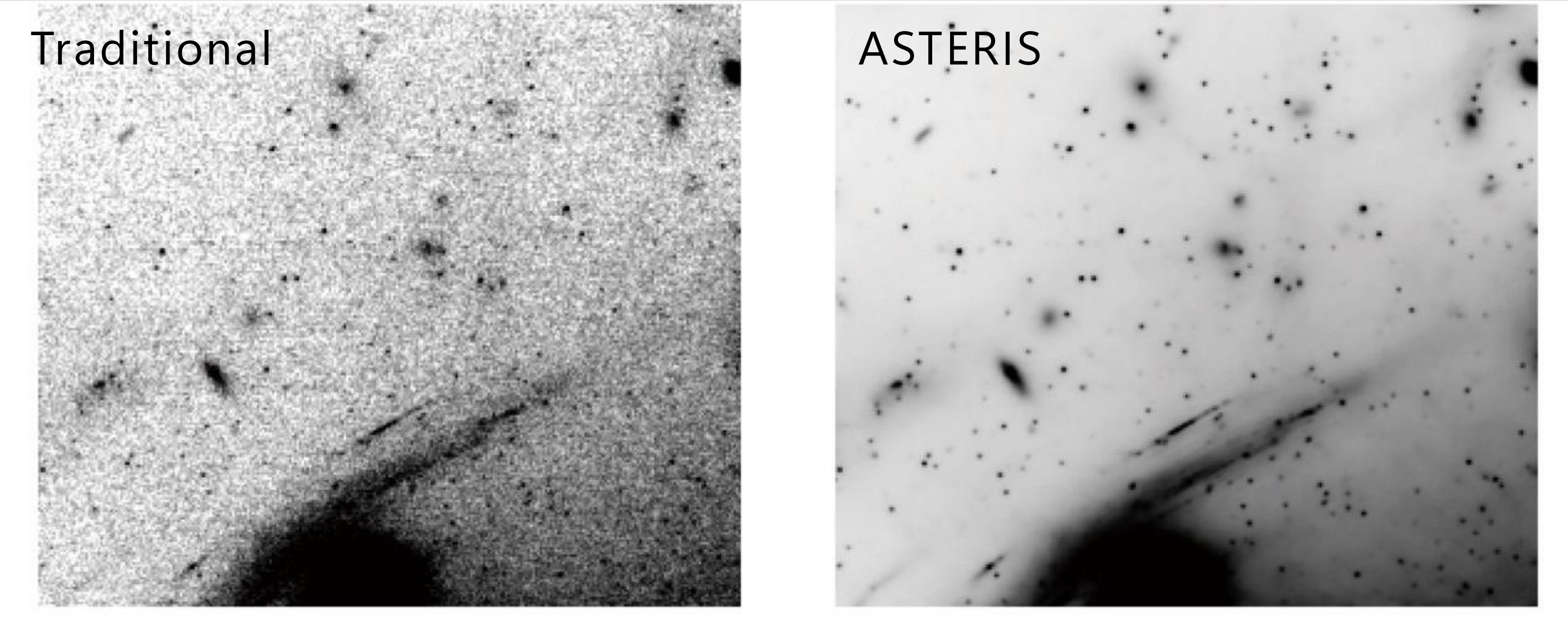
By now, 1 GRB, ~20 AGN (~100 at TeV)



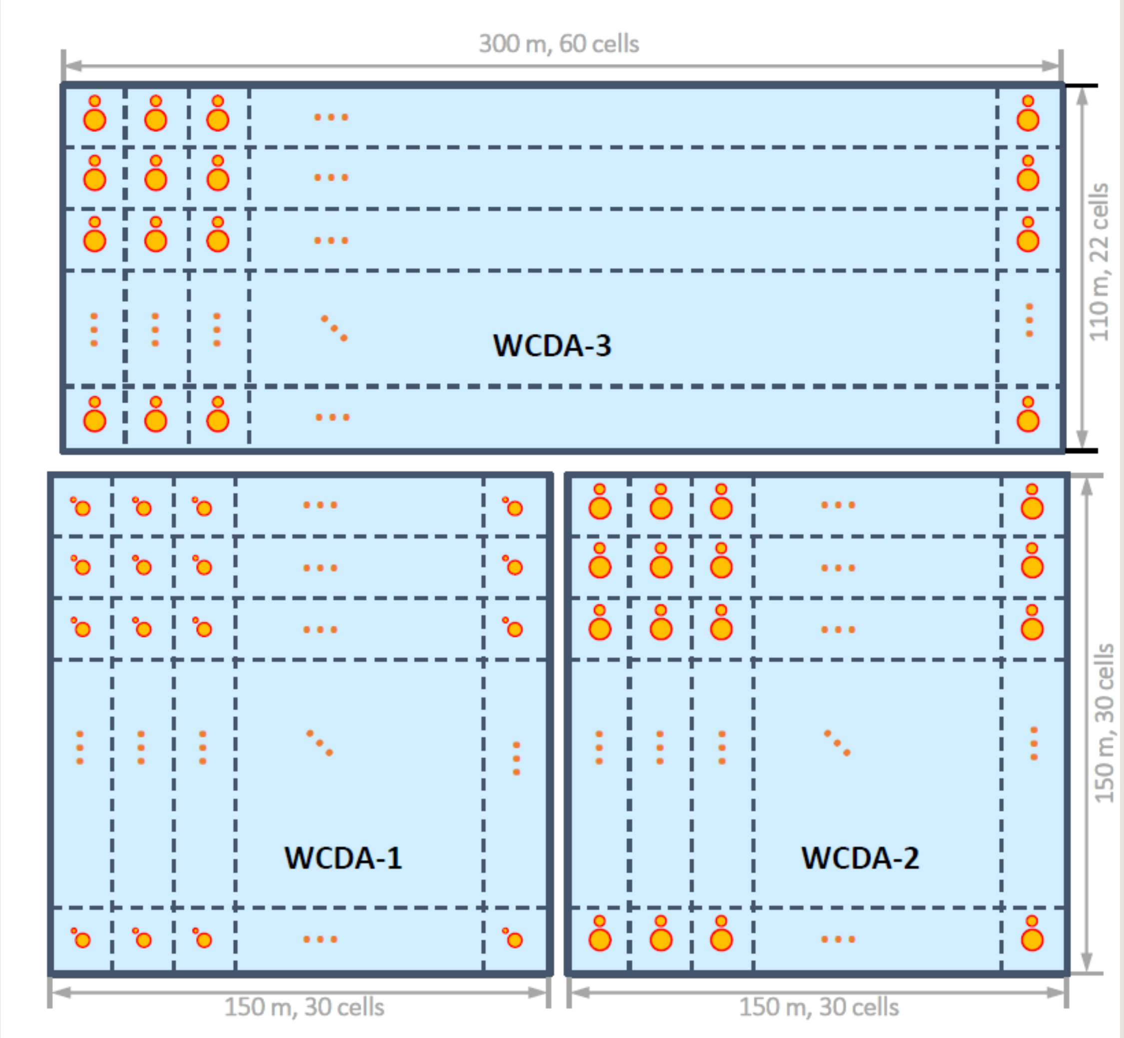
AI for LHAASO

AI for Science

AI significantly increased the detection depth of the James Webb Space Telescope by 1 magnitude (Y.D. Guo, et al, Science, 2026)



AI for WCDA

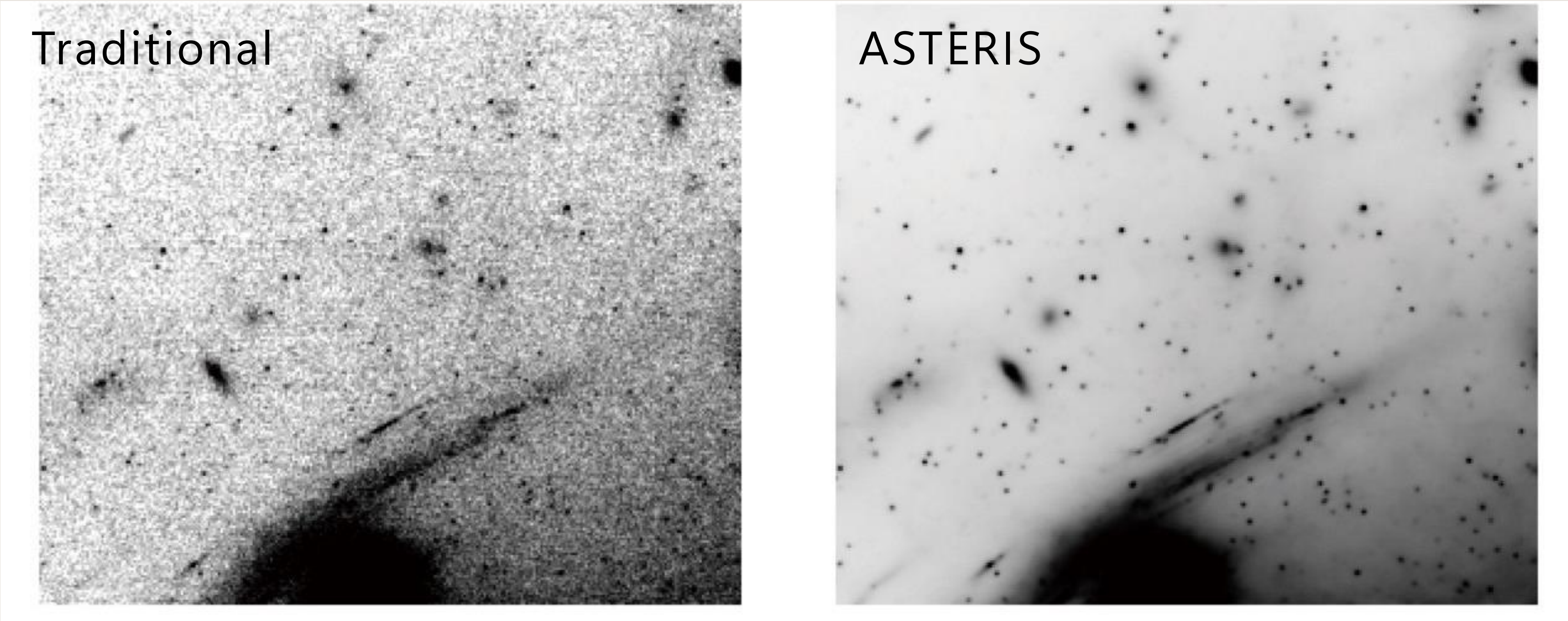


WCDA is a planar detector, and its events are suited to be represented as an image.

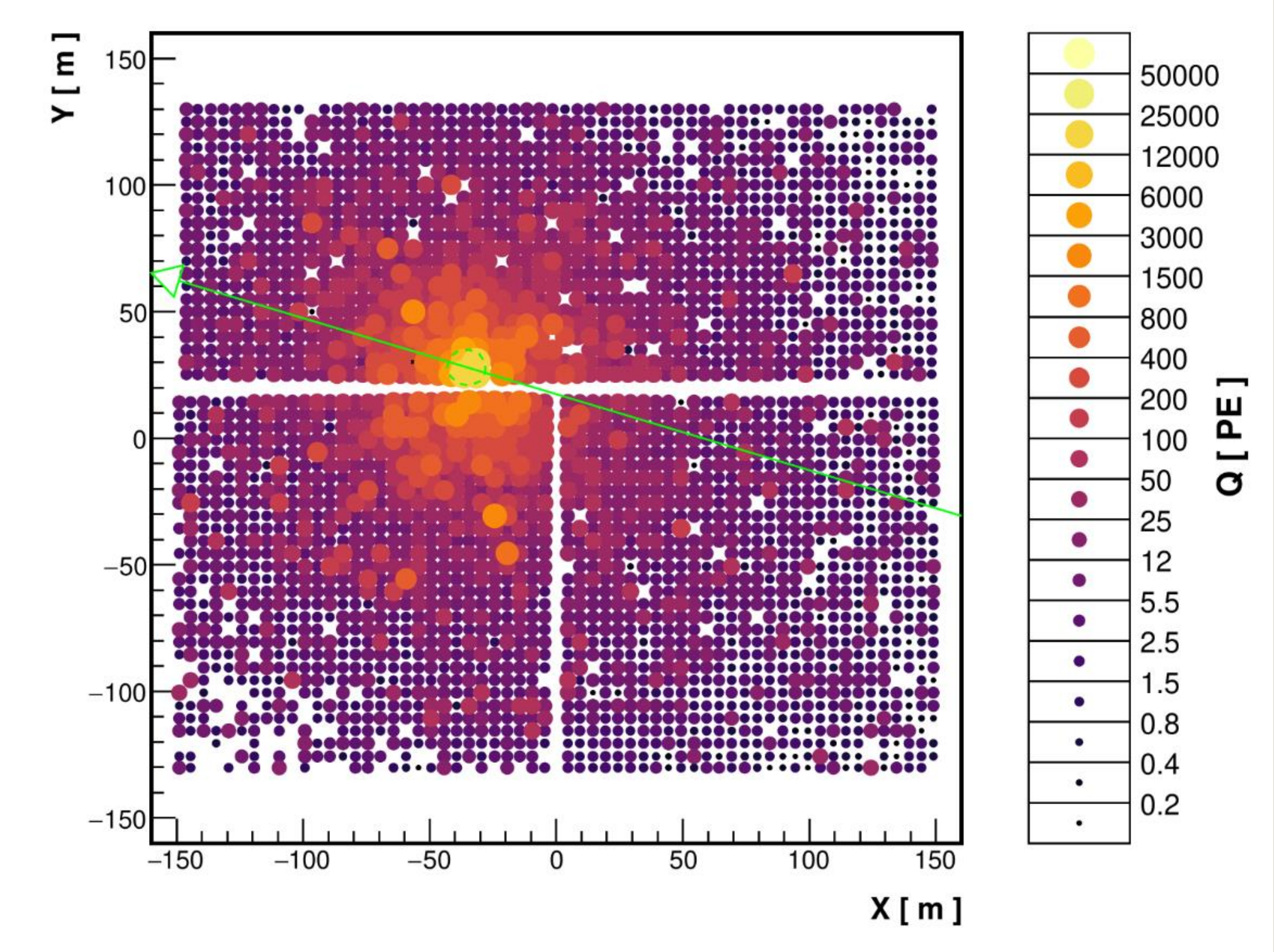
AI for LHAASO

AI for Science

AI significantly increased the detection depth of the James Webb Space Telescope by 1 magnitude (Y.D. Guo, et al, Science, 2026)



AI for WCDA

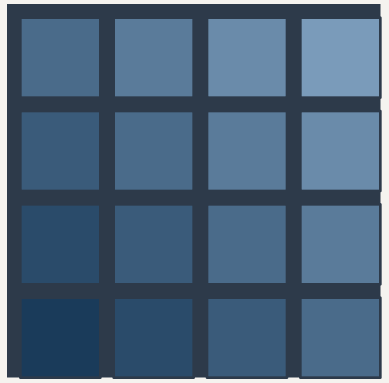


WCDA is a planar detector, and its events are suited to be represented as an image.

Vision Transformer

vit-base-patch16-224-in21k | ~80M parameters

Patch



Input Image

Patch -> 768-dimension vector space (token)

Patch Embedding
+ Position Embedding
+ [CLS] Token

Linear Projection

Merge the tokens and extract the features

Transformer
Encoder
Multi-Head Attention
MLP + LayerNorm

L Blocks

768-dimension -> 2-dimension

Gamma probability
CRs probability

MLP
Head

Classification

Key feature

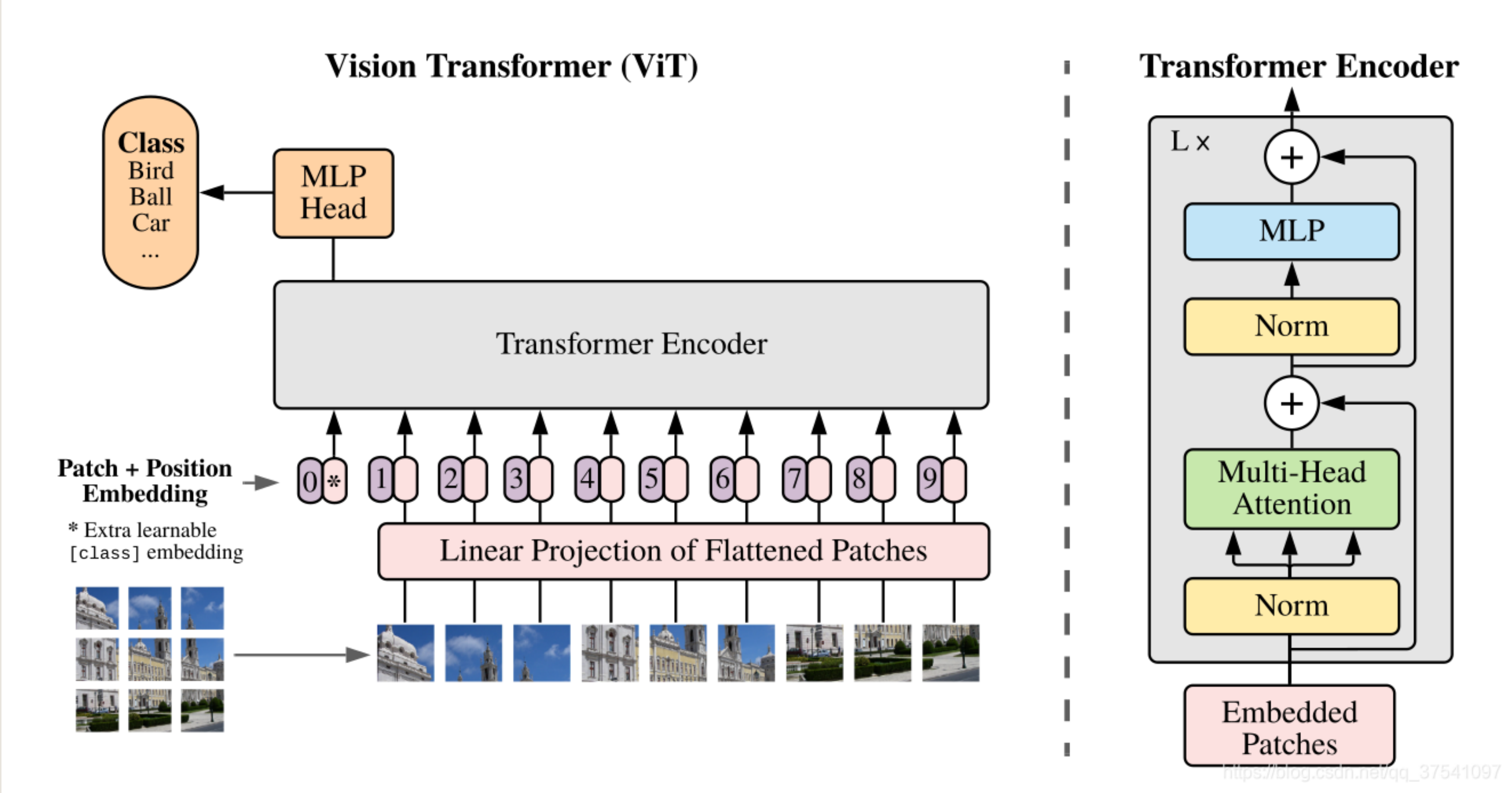
Base: model size ~86M parameters

Patch16: 16 16 pixel/patch

224: input image size 224 224 pixels

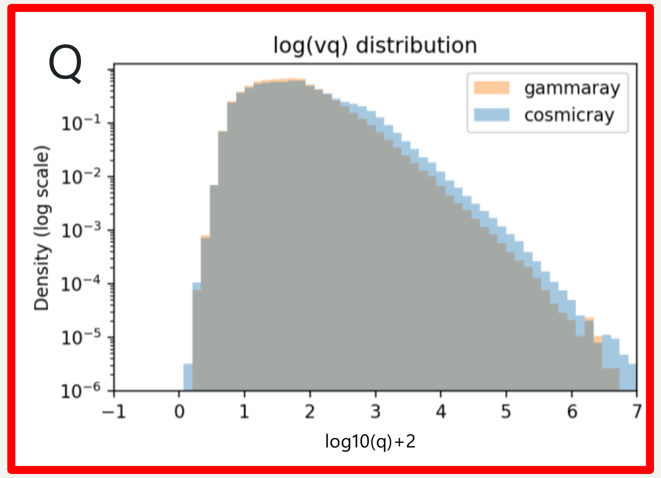
in21k: pre-training on ImageNet-21k

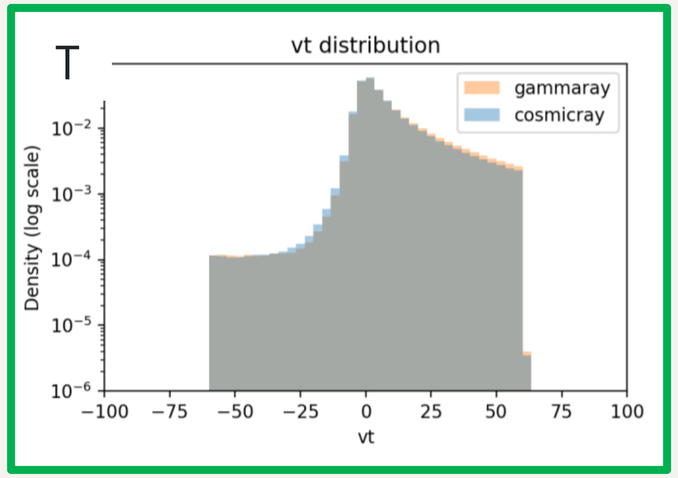
→ 224 224 image → 14 14 = 196 patches



LHAASO Data -> Image

Process of data projection

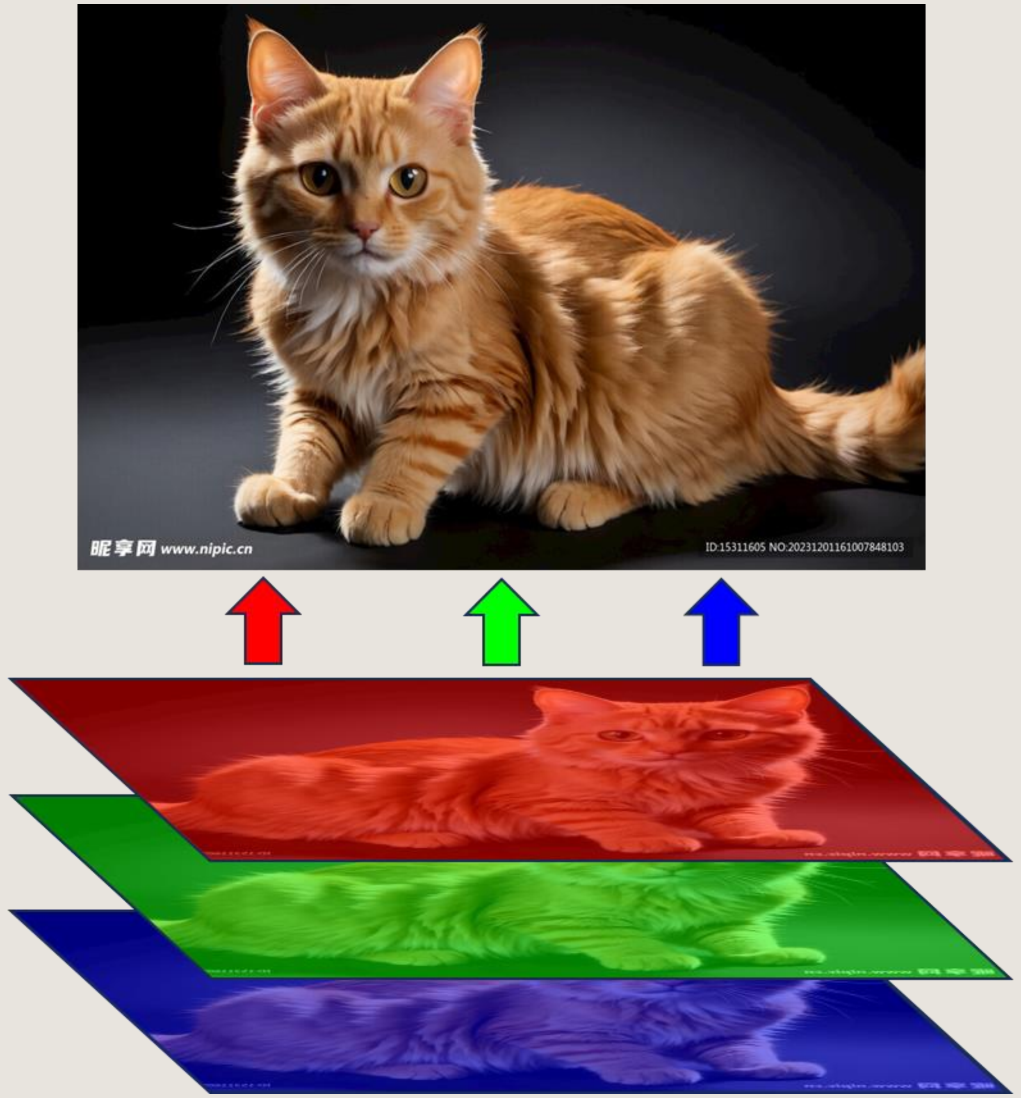




- Q → Red channel of image
- T → Green channel of image
- Linear projection and discretization to 0-255
- image size: 224×224

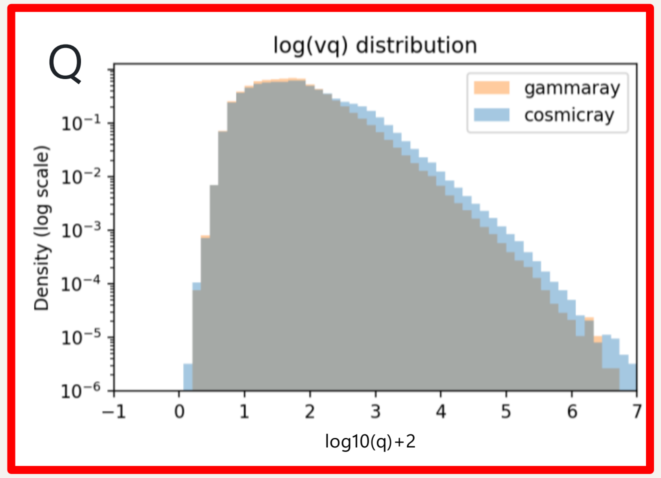
→ Q, T are aligned at the pixel level

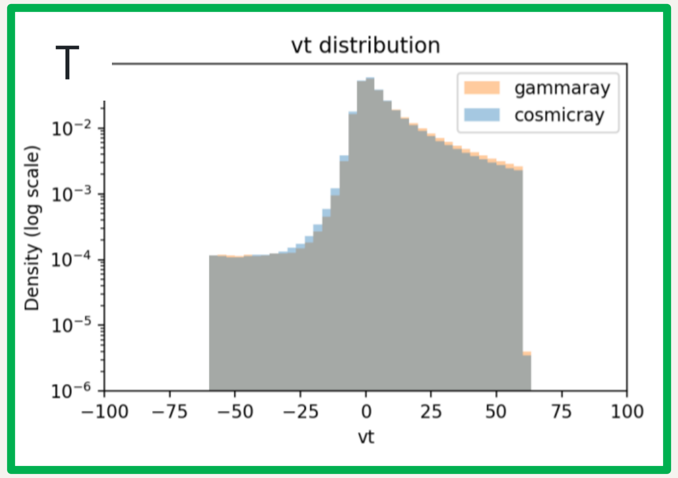
Q, T → Linear projection → 224 224



LHAASO Data -> Image

Process of data projection

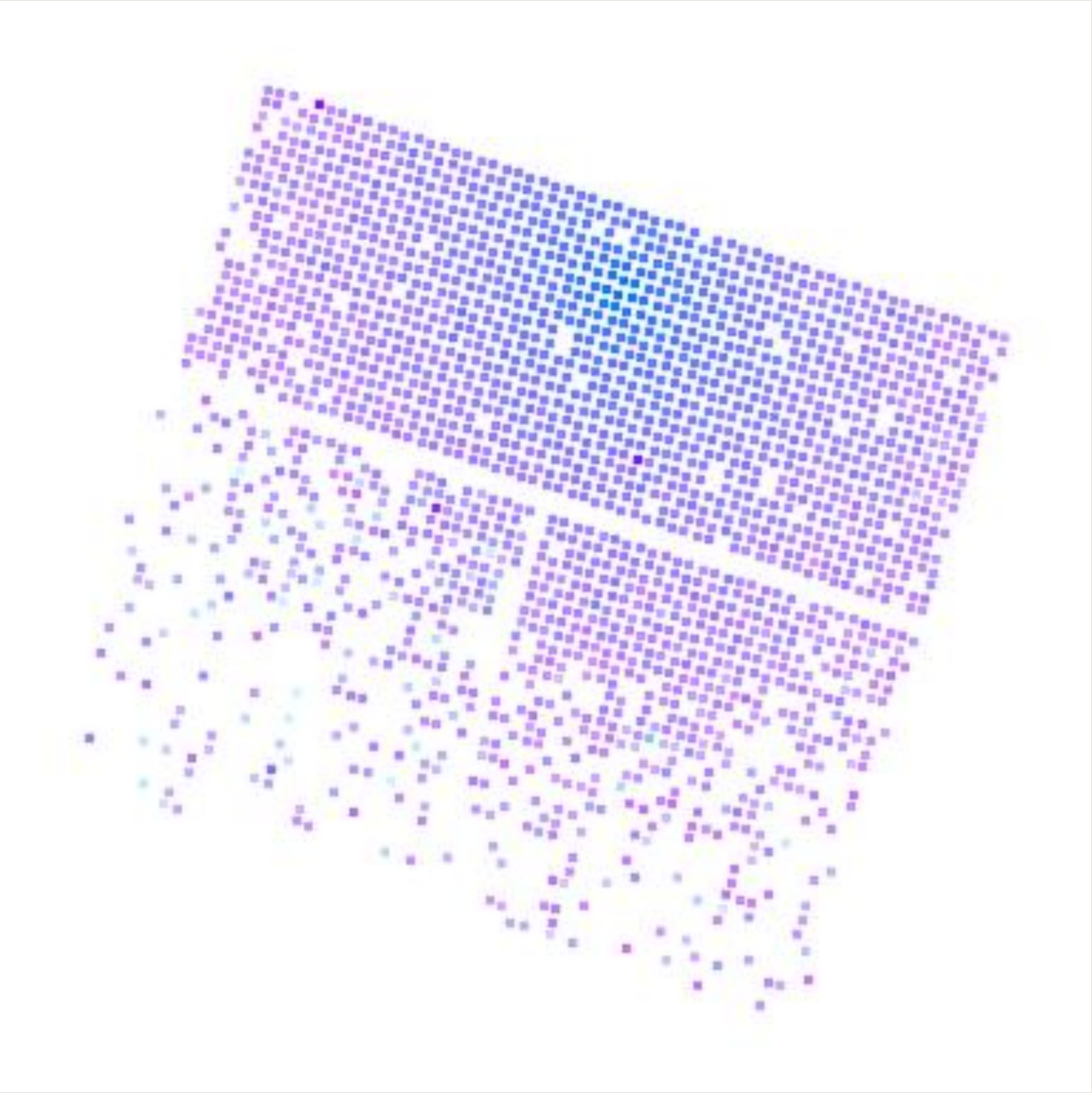




- Q → Red channel of image
- T → Green channel of image
- Linear projection and discretization to 0-255
- image size: 224×224

→ Q, T are aligned at the pixel level

Q, T → Linear projection → 224 224

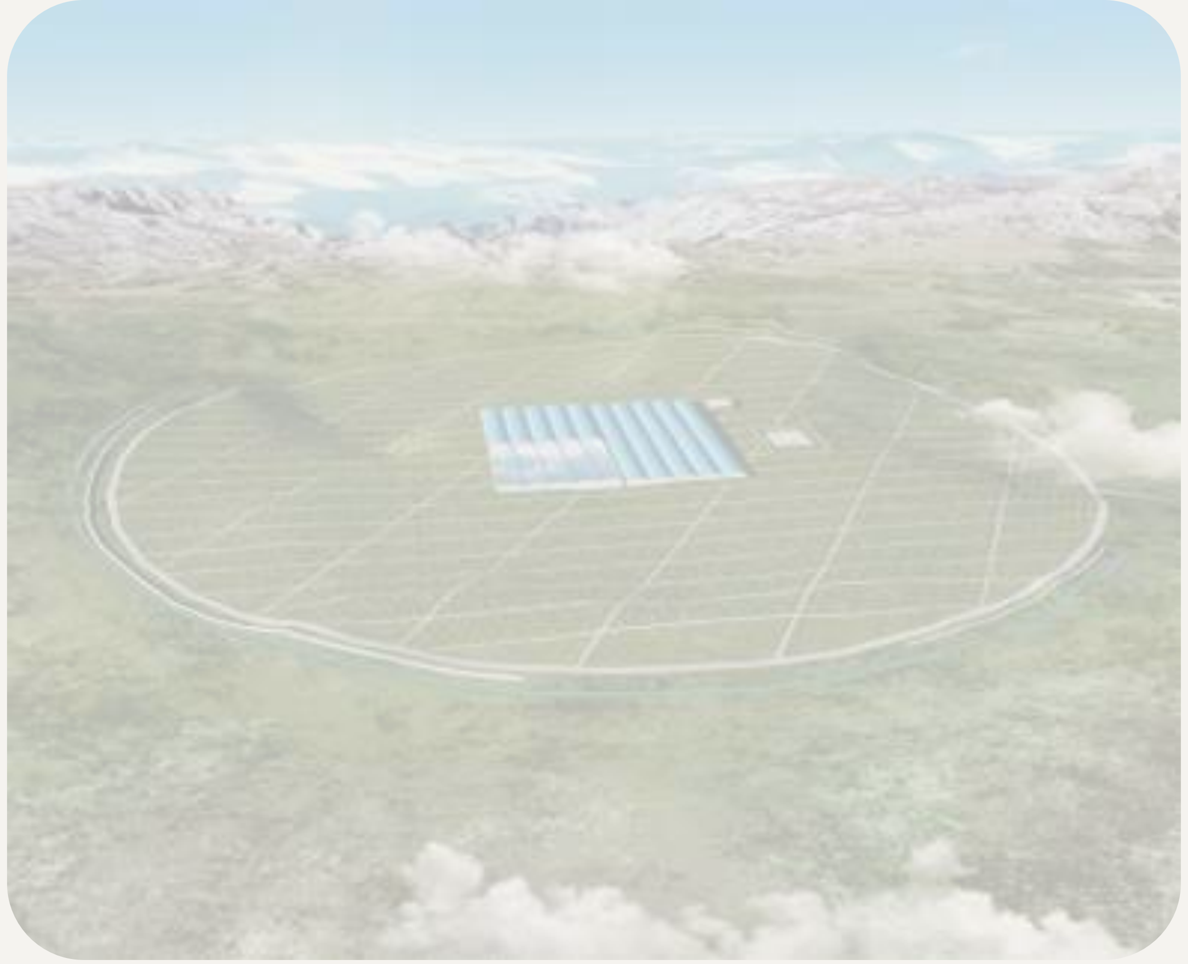


02

CHAPTER 02

Optimization performance of AI model

Effect of color augment, patch number and model architecture on AI performance



Optimization on classification of Gamma/CRs

Testing of image preprocessing

- remove color augment (~4%↑)
- Increase patch number (~10%↑)
  - Resize 1024→224/384
- Rotation, flip

Experiment setting

Train simulation dataset: ~1.3e7 events, Gamma:CRs 1:1

Test simulation dataset: ~3e6 events

Verify on real data: Crab, half month

The typical preprocessing includes color augment, and experiments have shown that removing it leads to better results.

Removing color augment **~4%** performance up

Crab significance @ N_{hit}>125 (~2 TeV)

Version	Crab Significance
with augment	1.29 times
without augment	1.34 times

Optimization on classification of Gamma/CRs

Testing of image preprocessing

- remove color augment (~4%↑)
- Increase patch number (~10%↑)
  - Resize 1024→224/384
- Rotation, flip

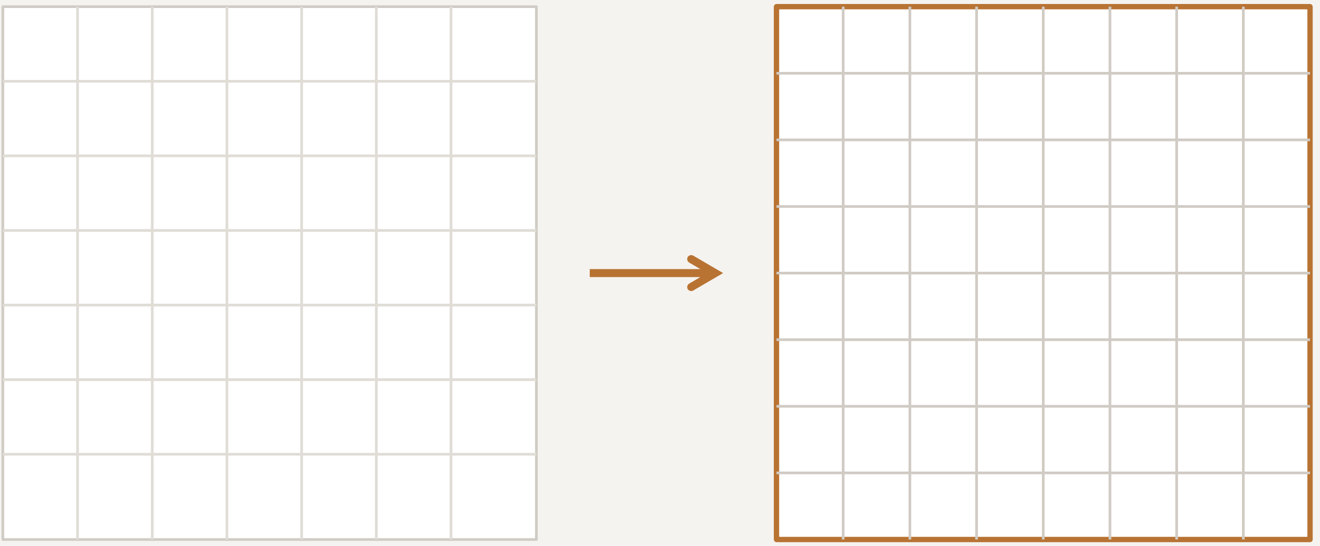
vit-base-patch16-224

- image size: 224 224
- Patch number: 14 14 = 196 patches
- Each patch: 16 16 pixels

vit-base-patch16-384

- image size 384 384
- Patch number: 24 24 = 576 patches
- Each patch: 16 16 pixels

Patch Division Illustration (14 14 vs 24 24)



14 14 = 196   24 24 = 576

Patch 14 14 → 24 24: Crab significance ~10% up

Crab significance: 1.34 → 1.47

Optimization on classification of Gamma/CRs

Testing of image preprocessing

- remove color augment (~4%↑)
- Increase patch number (~10%↑)
  - Resize 1024→224/384
- Rotation, flip

↓

Exploration of Model Architecture

- ViT → Swin Transformer
  - Multi-scale feature extraction
  - Relative position embedding

ViT -> Swin transformer

Characteristic	ViT	Swin
Attention mechanism	Global self-attention	Shifted Window Attention
Position embedding	Absolute embedding	Relative embedding
Feature extraction	Single scale	Multi scale
Computation complexity	$O(n^2)$	$O(n)$

Swin's multi-scale feature extraction and relative position embedding are more suitable for the local structural features of event images, and can further increase by ~7%

Crab Significance: 1.47 → 1.57

Optimization on classification of Gamma/CRs

Testing of image preprocessing

- remove color augment (~4%↑)
- Increase patch number (~10%↑)
  - Resize 1024→224/384
- Rotation, flip

↓

Exploration of Model Architecture

- ViT → Swin Transformer
  - Multi-scale feature extraction
  - Relative position embedding

ViT -> Swin transformer

Characteristic	ViT	Swin
Attention mechanism	Global self-attention	Shifted Window Attention
Position embedding	Absolute embedding	Relative embedding
Feature extraction	Single scale	Multi scale
Computation complexity	$O(n^2)$	$O(n)$

Swin's multi-scale feature extraction and relative position embedding are more suitable for the local structural features of event images, and can further increase by ~7%

Overall factor of Crab significance

×1.57

improved by 57% compared to traditional methods

Applied to direction reconstruction

Testing of image preprocessing

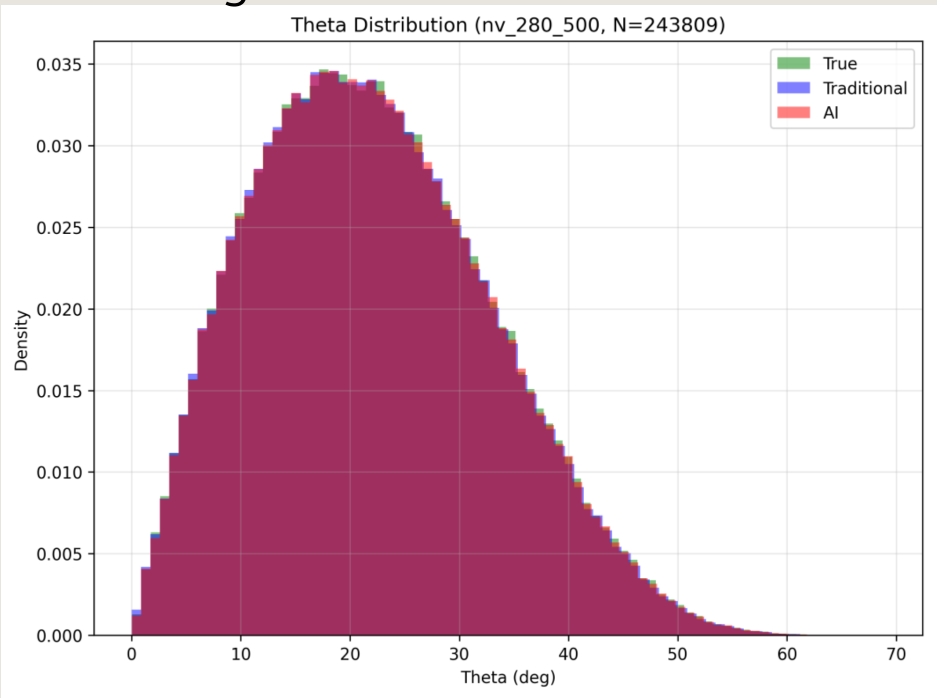
- remove color augment (~4%↑)
- Increase patch number (~10%↑)
  - Resize 1024→224/384
- Rotation, flip

↓

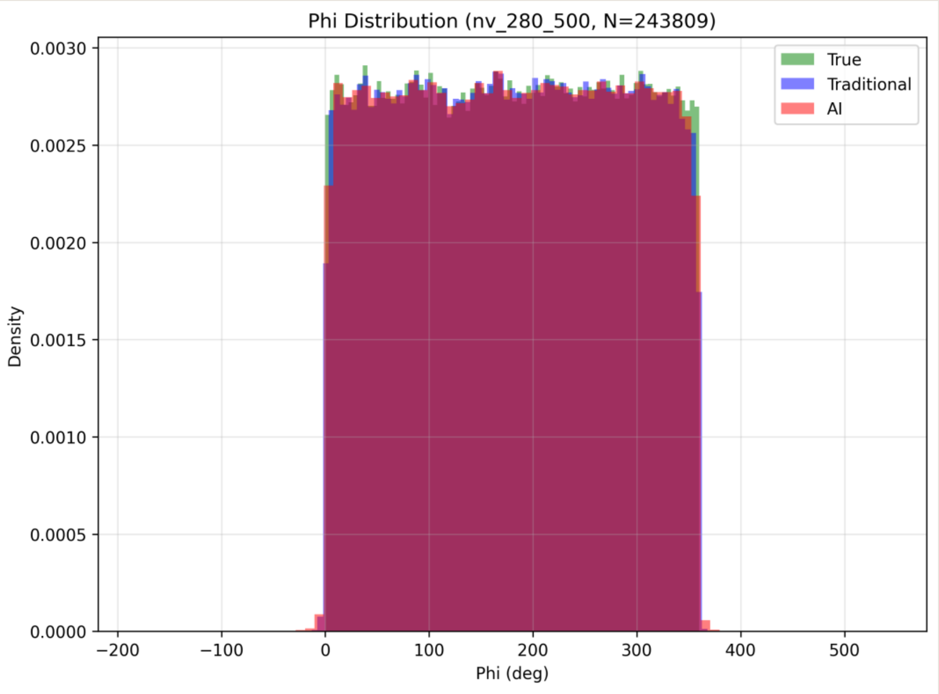
Exploration of Model Architecture

- ViT → Swin Transformer
  - Multi-scale feature extraction
  - Relative position embedding

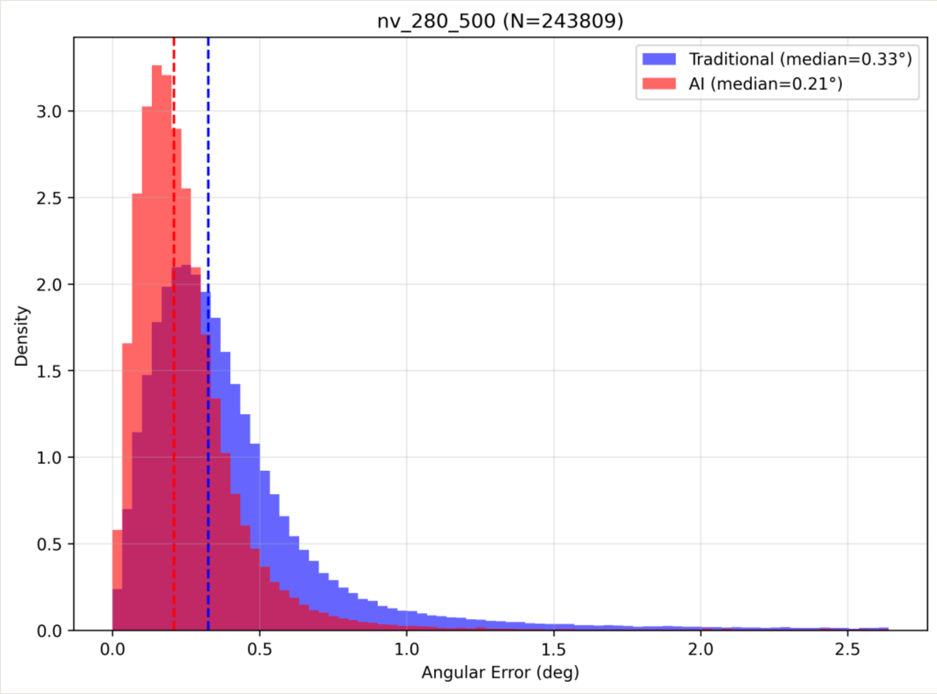
Zenith angle

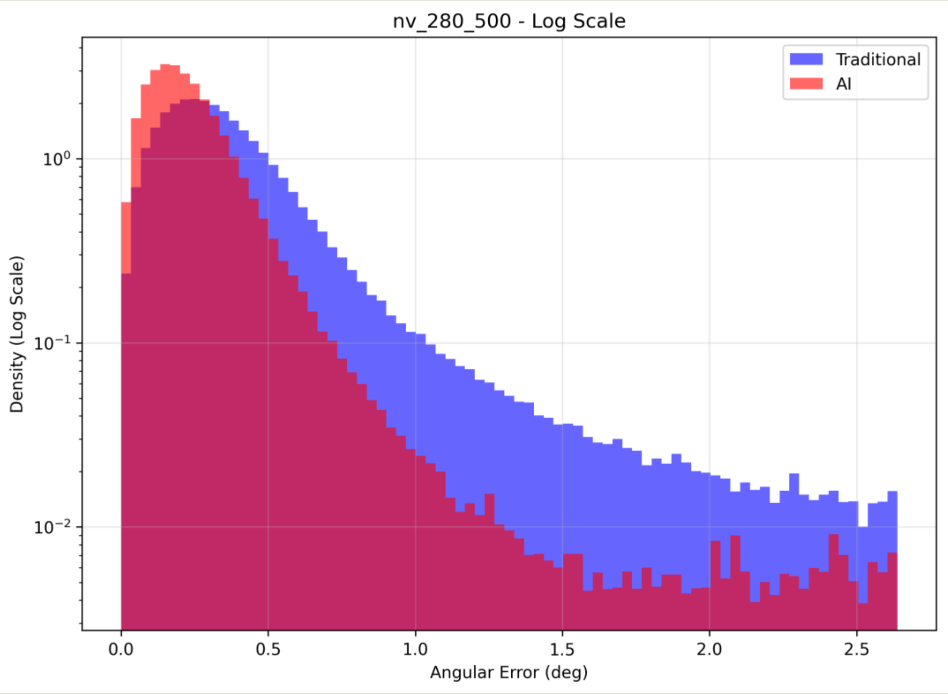


Azimuth angle



Angular distance between true and rec





Applied to direction reconstruction

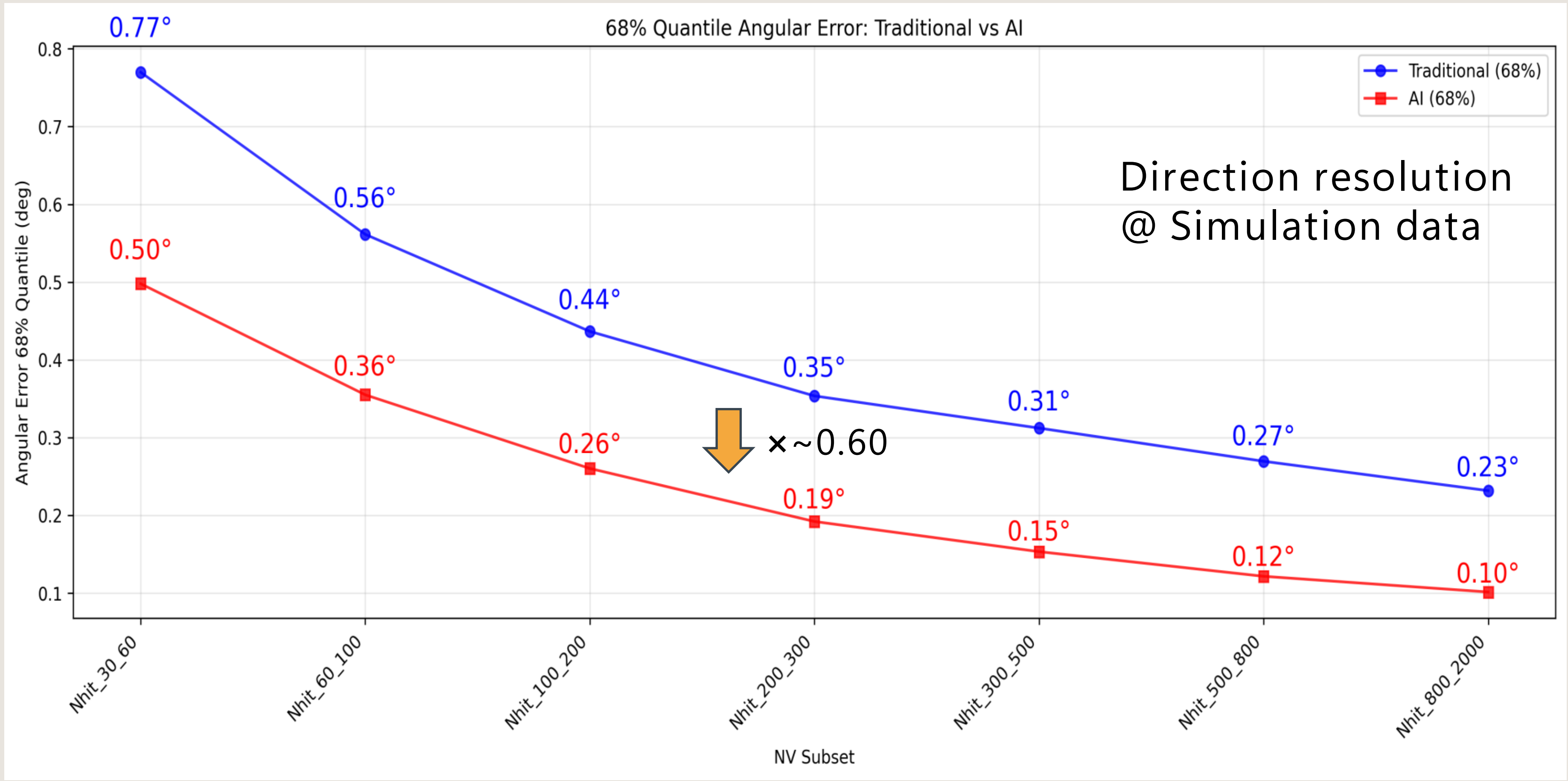
Testing of image preprocessing

- remove color augment (~4%↑)
- Increase patch number (~10%↑)
  - Resize 1024→224/384
- Rotation, flip

↓

Exploration of Model Architecture

- ViT → Swin Transformer
  - Multi-scale feature extraction
  - Relative position embedding



Crab significance improvement with AI direction

~1.5×

improved by ~50% compared to traditional methods

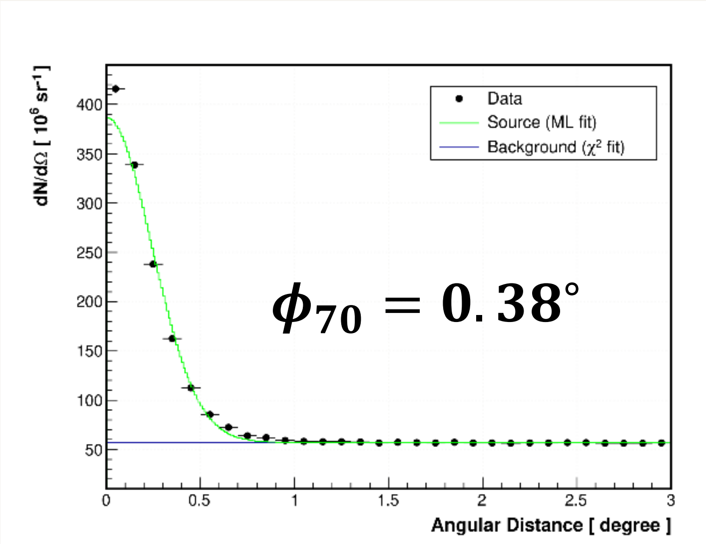
Overall Crab significance factor

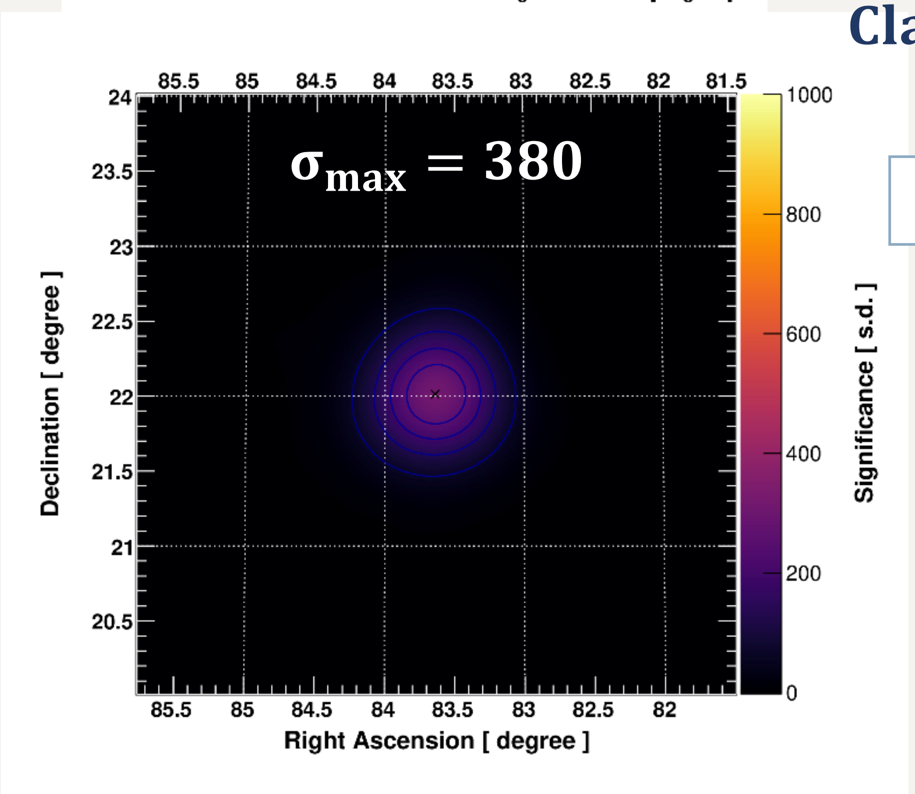
~2.35×

improved by ~135% compared to traditional methods

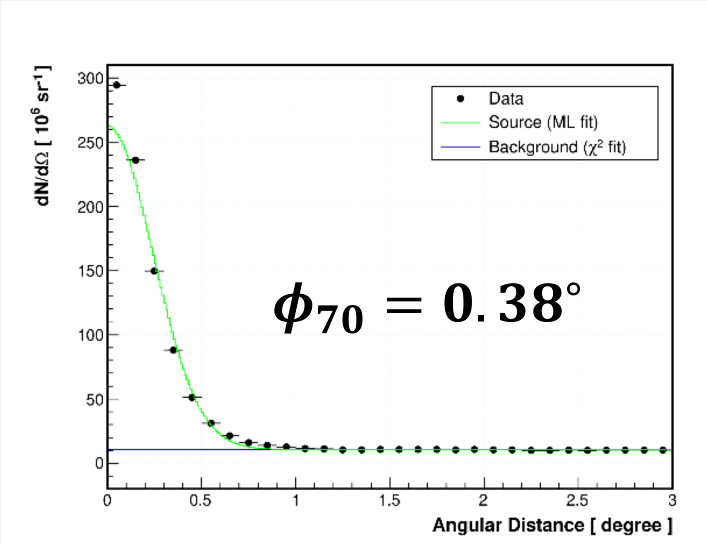
AI-powered WCDA: CRAB In 2022

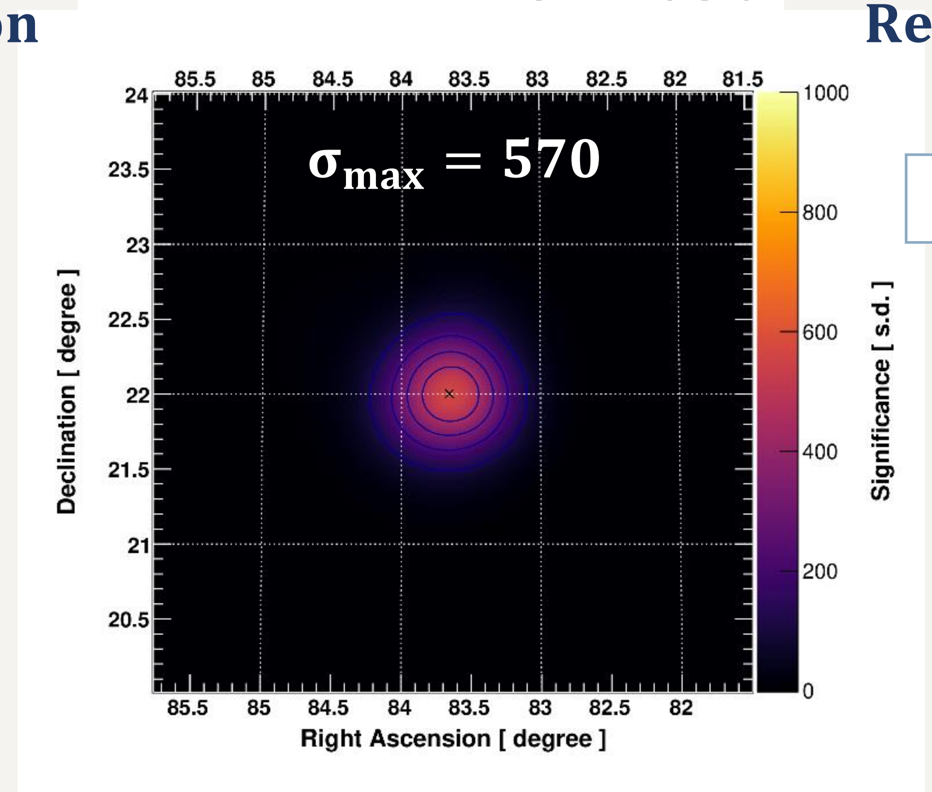
@N_{hit}≥125, ~2 TeV



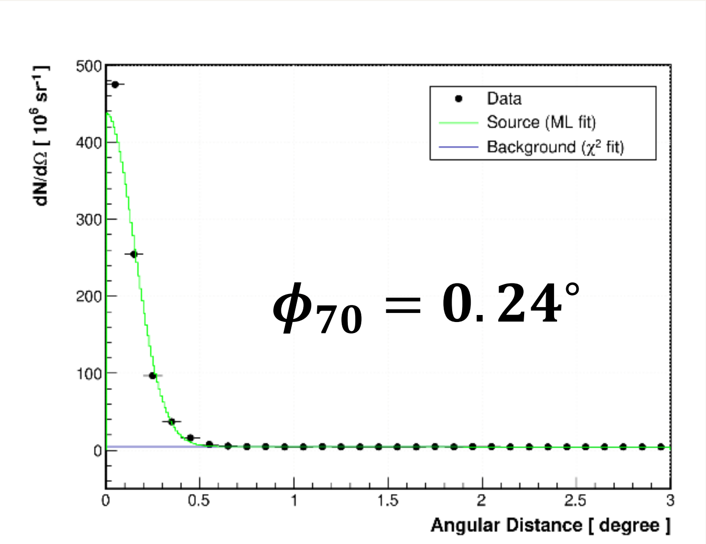


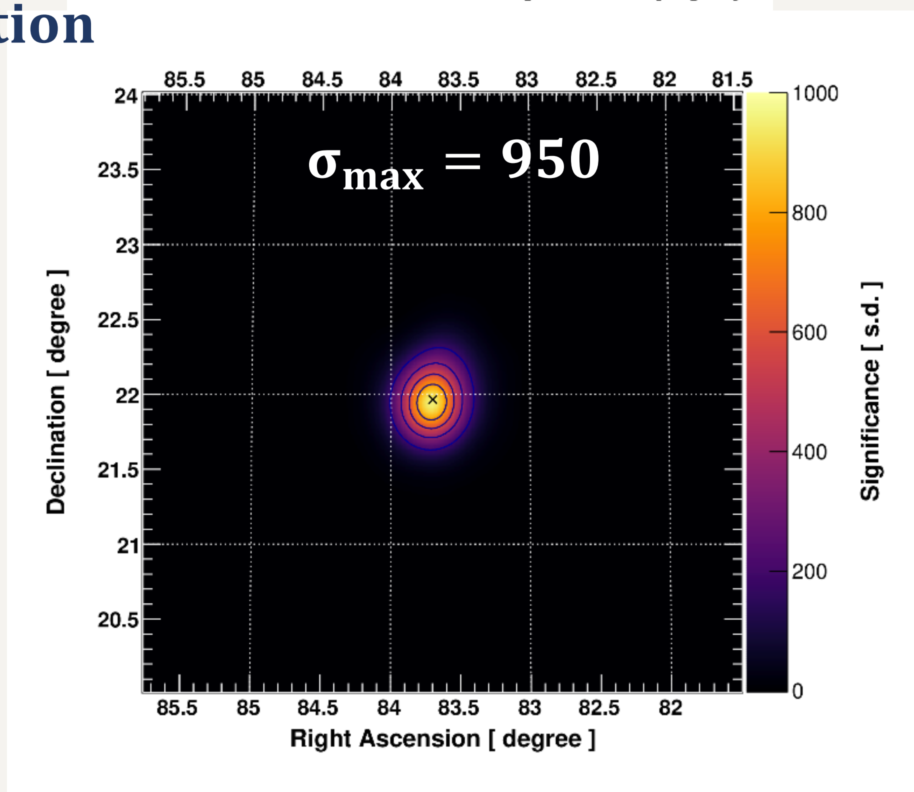
Classification



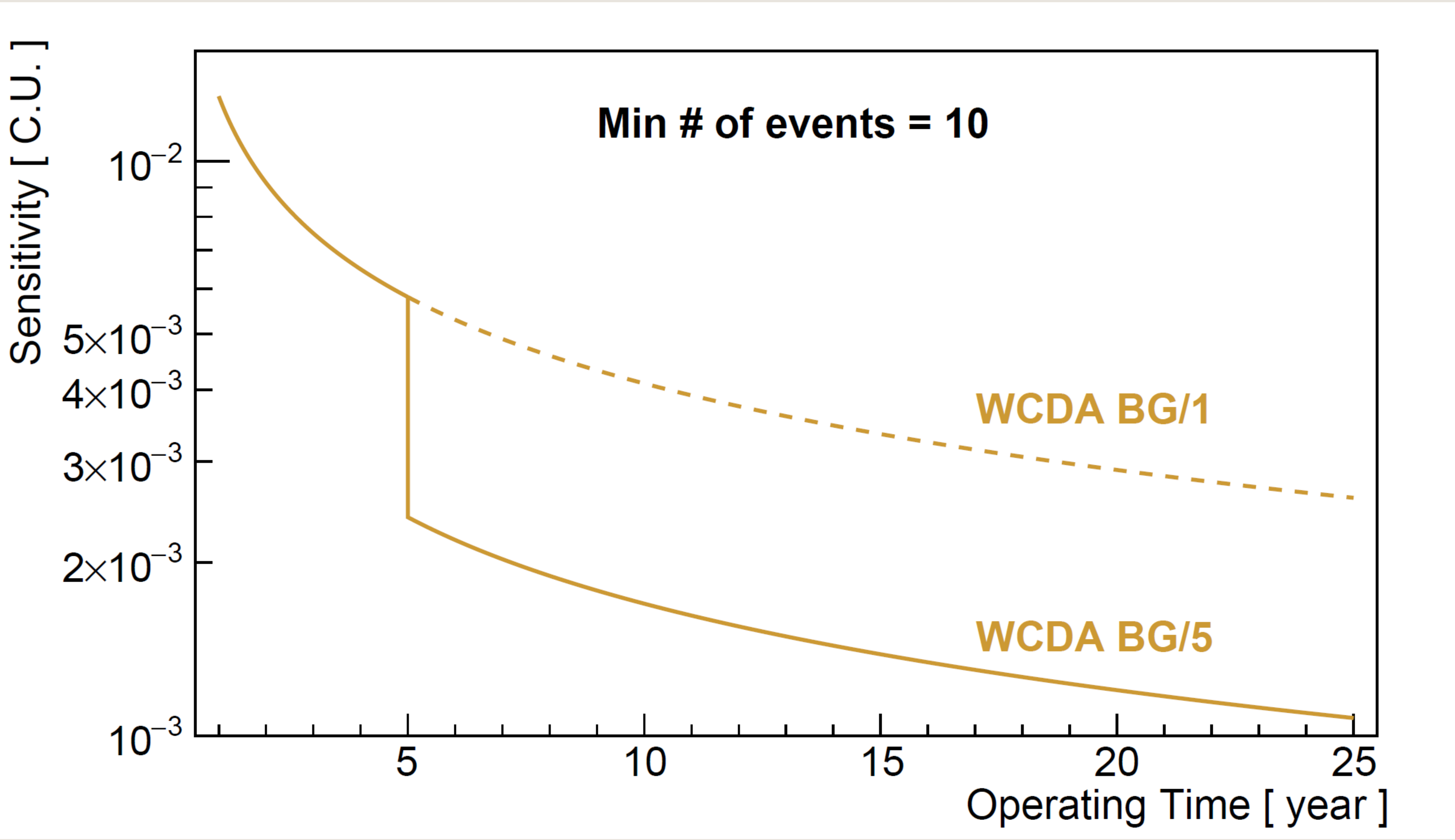


Reconstruction





AI-powered WCDA



03

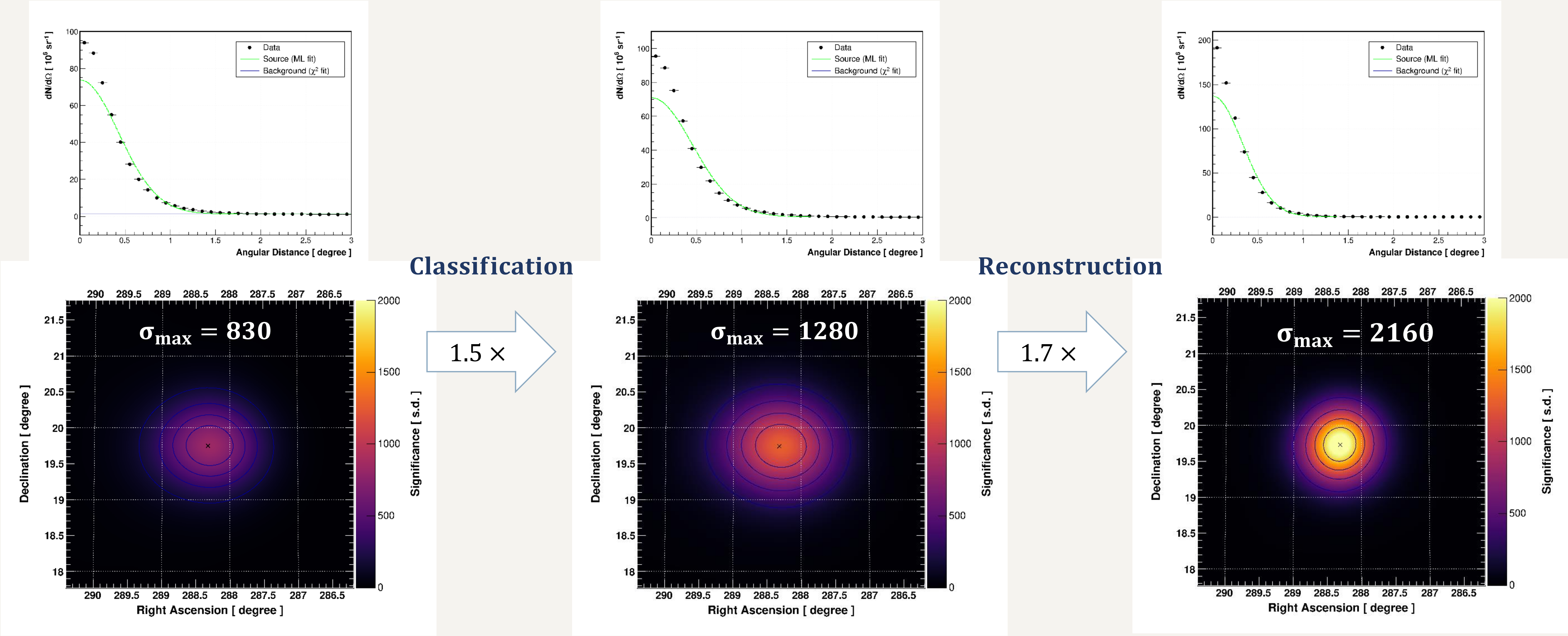
CHAPTER 03

Verification on BOAT GRB and Outlook

Q factor and Significance, Plan

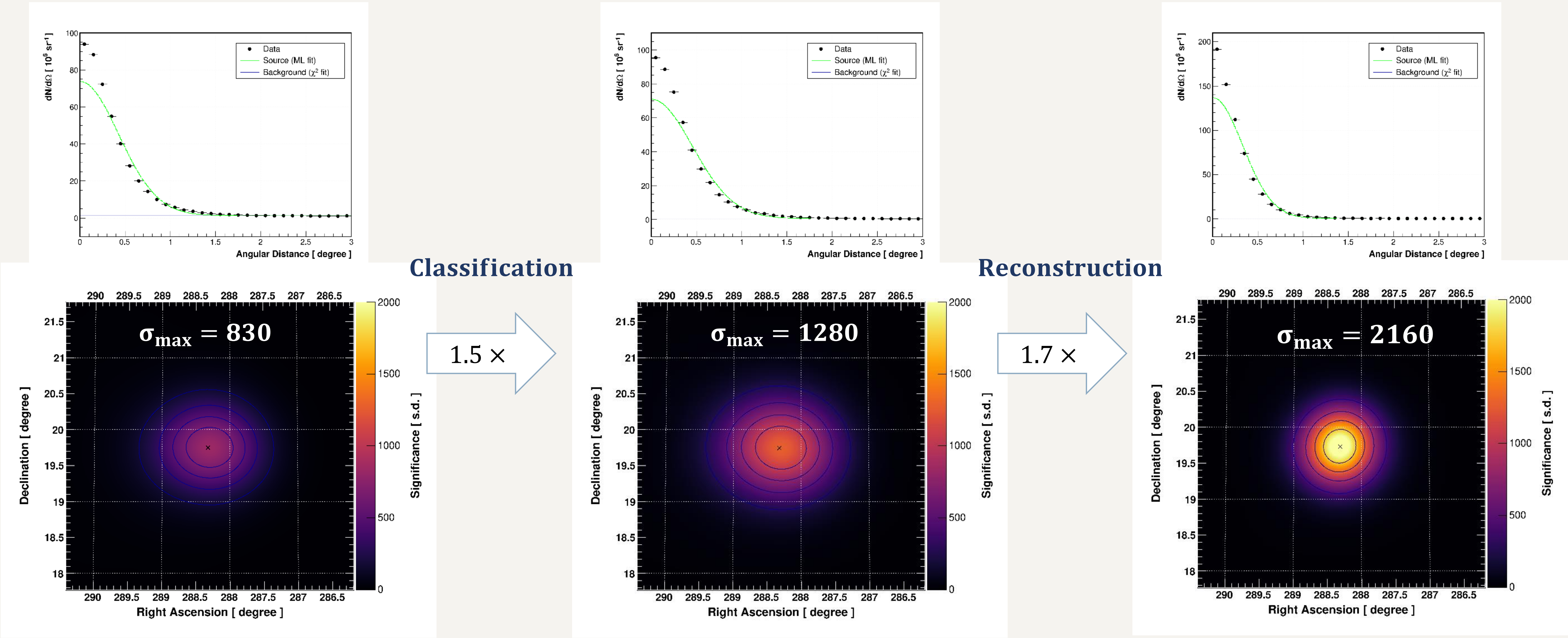
AI-powered WCDA: GRB 221009A (2 Hours)

$N_{\text{hit}} \geq 5$



AI-powered WCDA: GRB 221009A (2 Hours)

$N_{\text{hit}} \geq 5$



First Look At Light Curve of AI-Enhanced BOAT GRB will be presented in the last talk tomorrow morning.

Summary & Outlook

Key conclusion

- Gamma/CRs classification: Crab significance **×1.57**
- Direction reconstruction: Crab significance **~1.5×**
- BOAT GRB: background reduced by ~10×

Plan

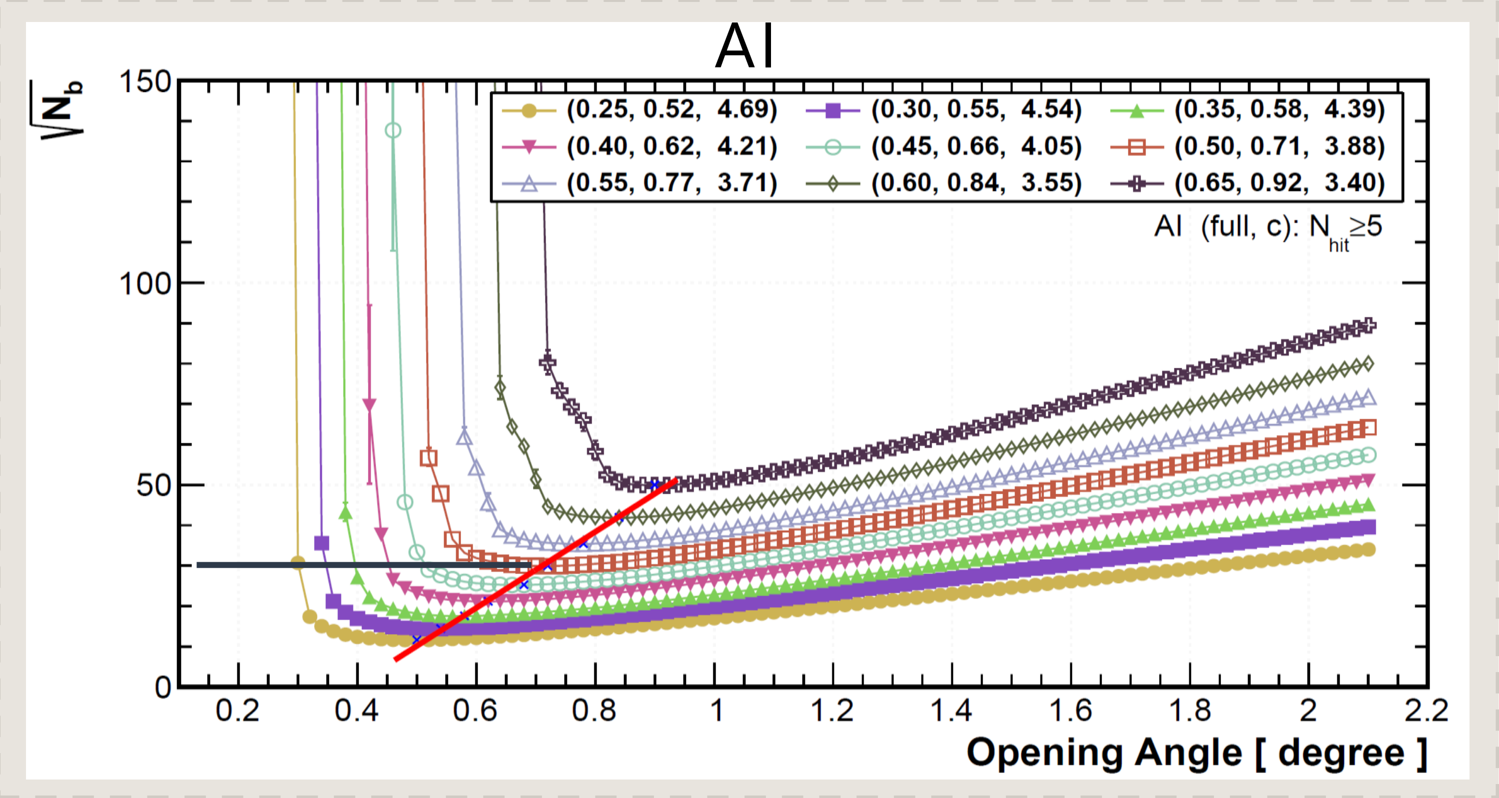
- ➢ Developing AI models for improving WCDA performance
  - Scaling on patch number and model size
  - Deep research of model architecture: CNN⊗transformer, MOE, ViT->Swin
  - Multi-tasks synergistic enhancement
  - More sources used for verification: Mrk 421, J2226+6057 etc.
- ➢ Simulation vs. Experiment
- ➢ Use AI-powered data to do physical analysis

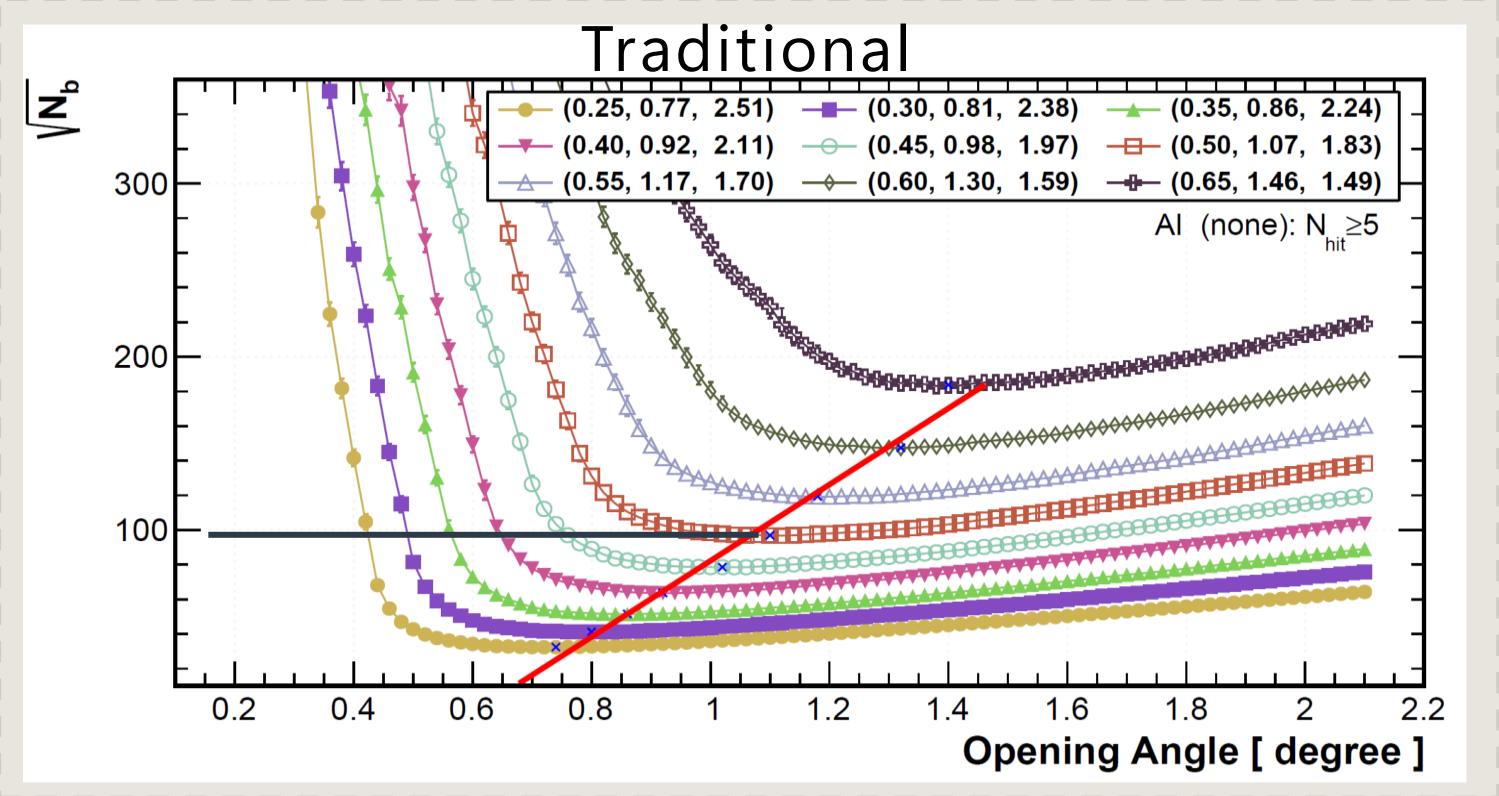
We are here now. →

Shicong Hu | IHEP

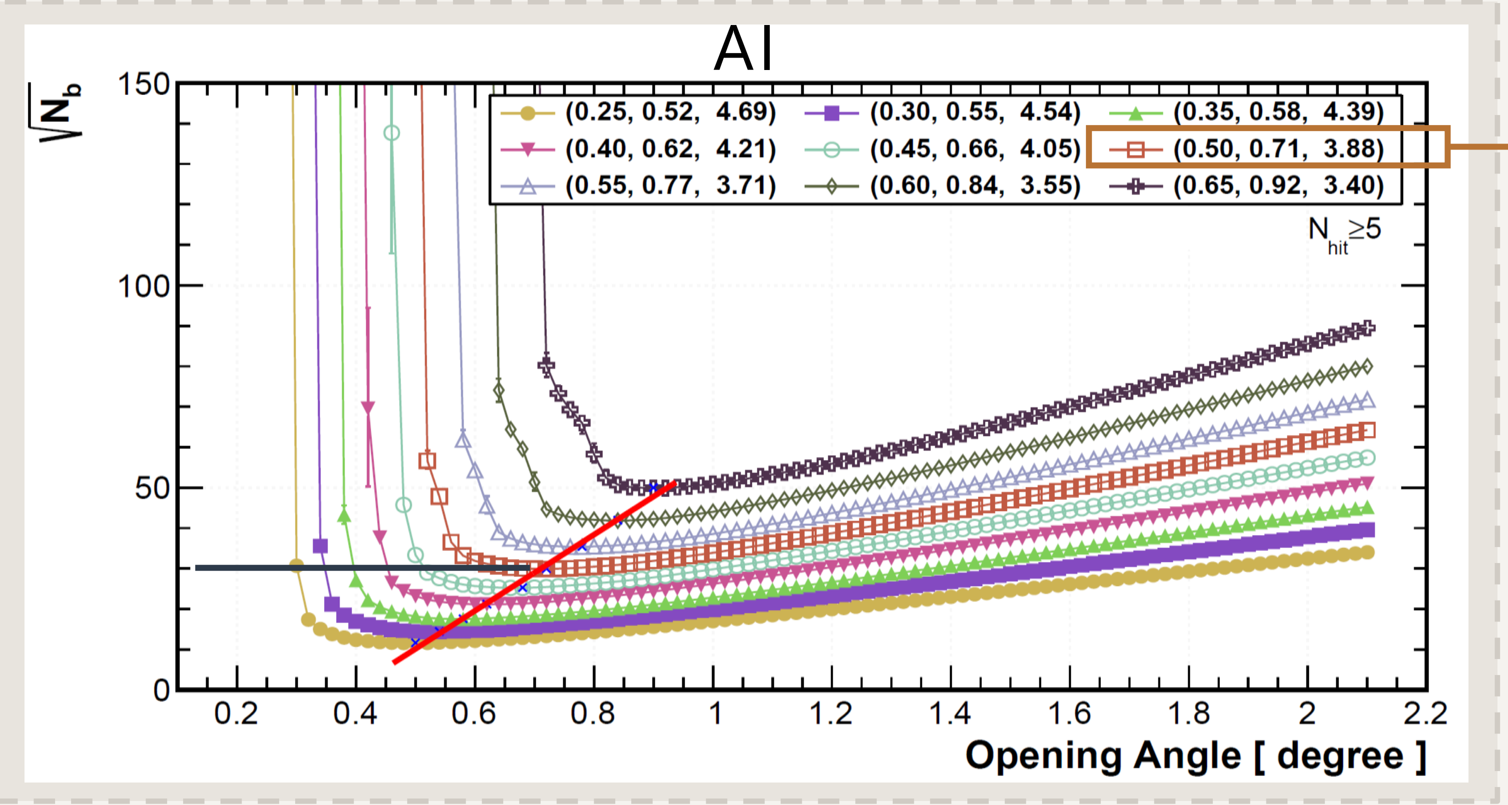
1st LHAASO General Meeting, 2026 @ NJU, Suzhou

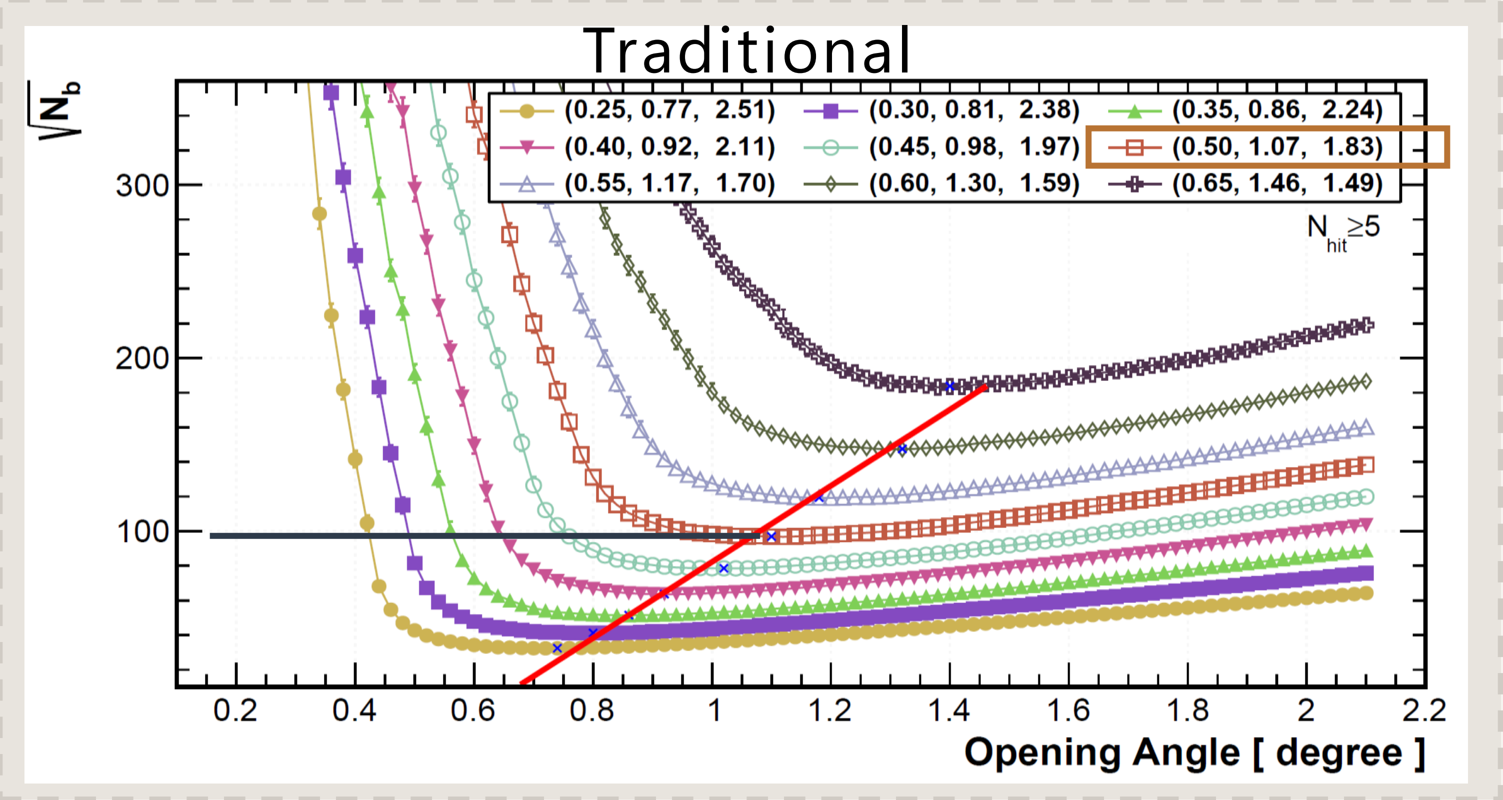
Q factor and angular resolution of BOAT GRB





Q factor and angular resolution of BOAT GRB



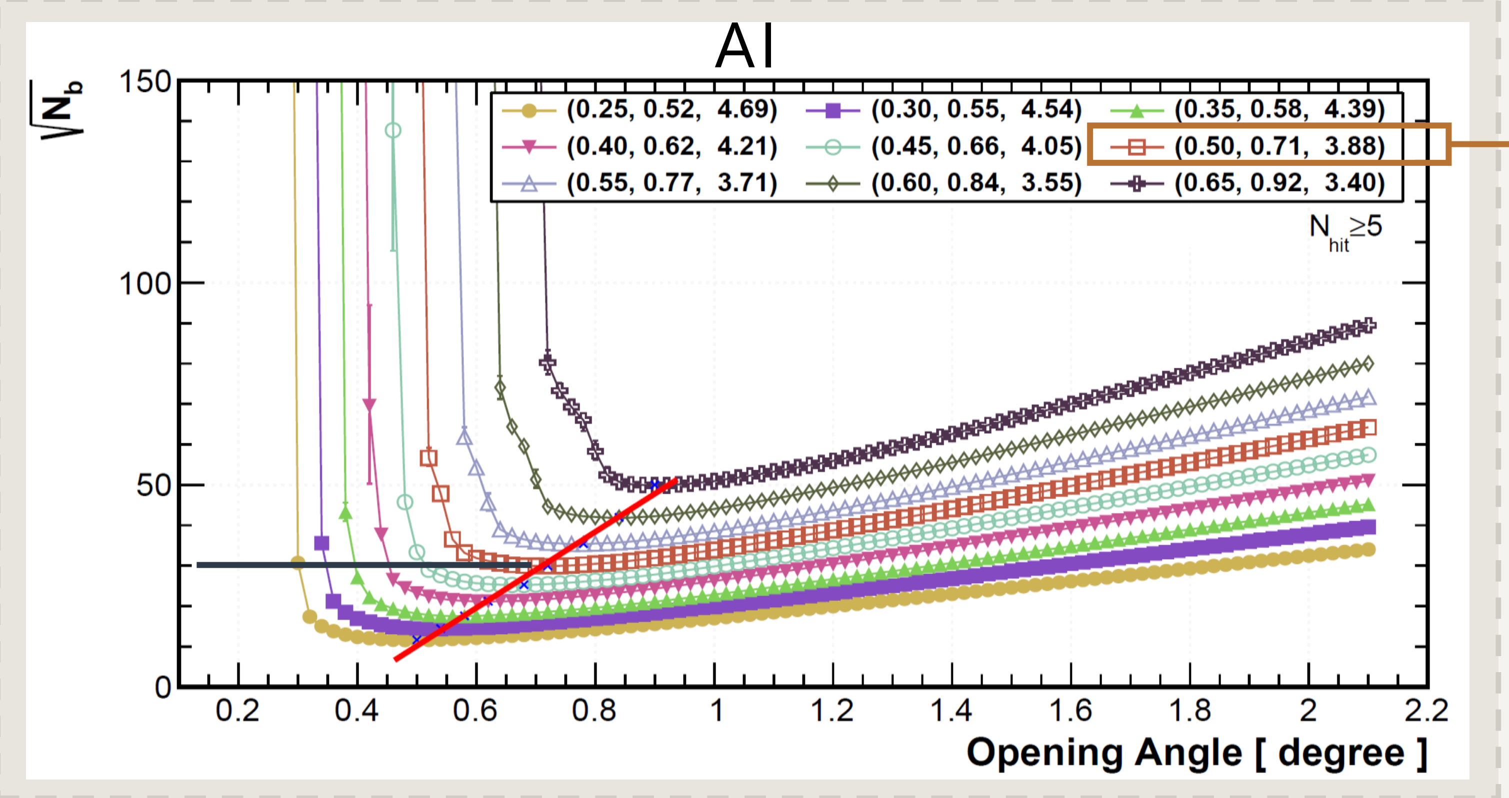


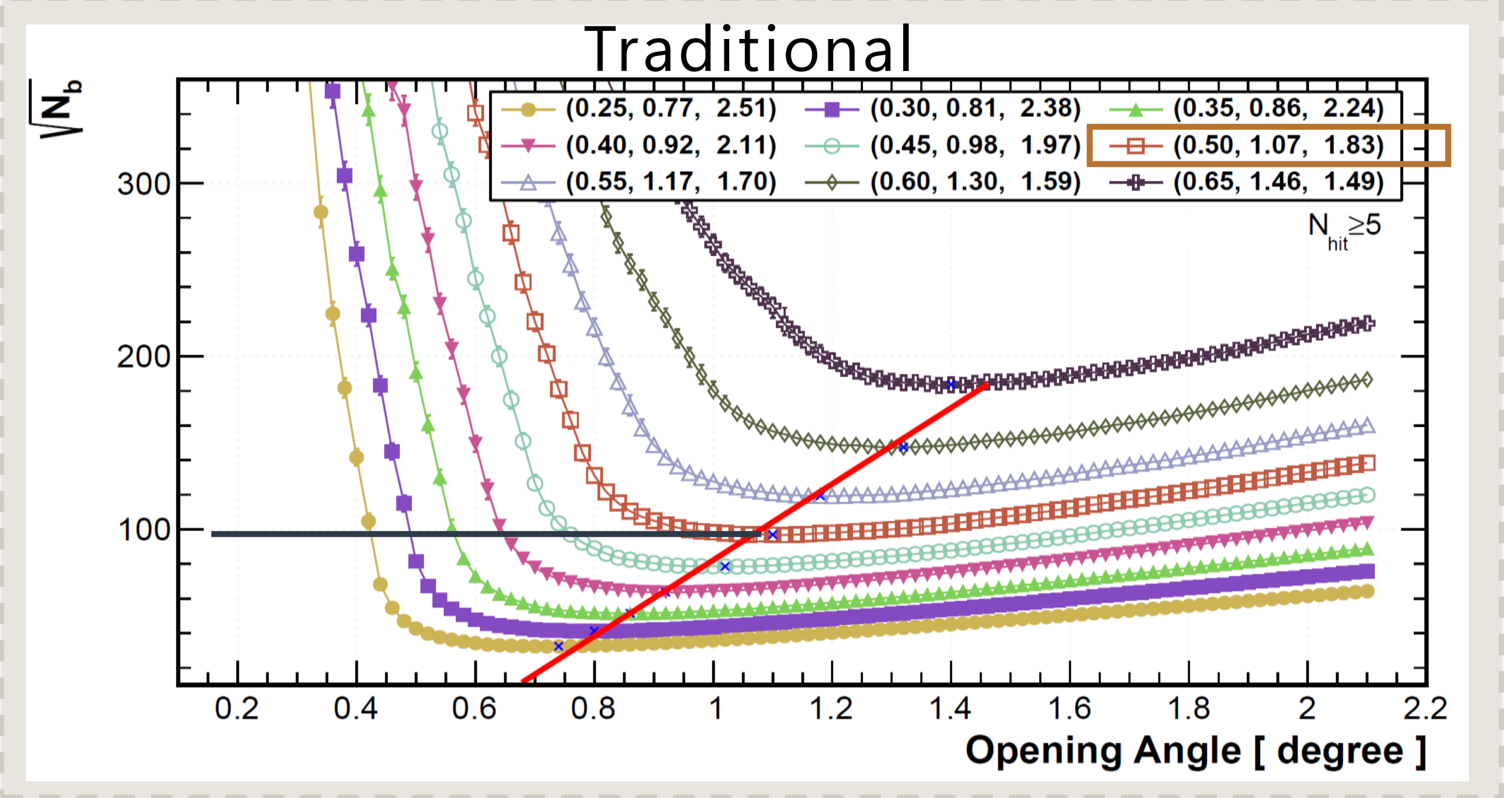
(Gamma ray retention ratio,
Optimal opening angle,
Q factor of Gamma/CRs classification)

	Optimal opening angle	Q factor
Traditional	1.07	1.83
AI	0.71	3.88
Ratio	1.51	2.12
Overall ratio	$3.2 = \sqrt{N_{b,0}}/\sqrt{N_{b,AI}}$	

→ With the same gamma retention ratio, the background level is reduced by 10 times.

Q factor and angular resolution of BOAT GRB





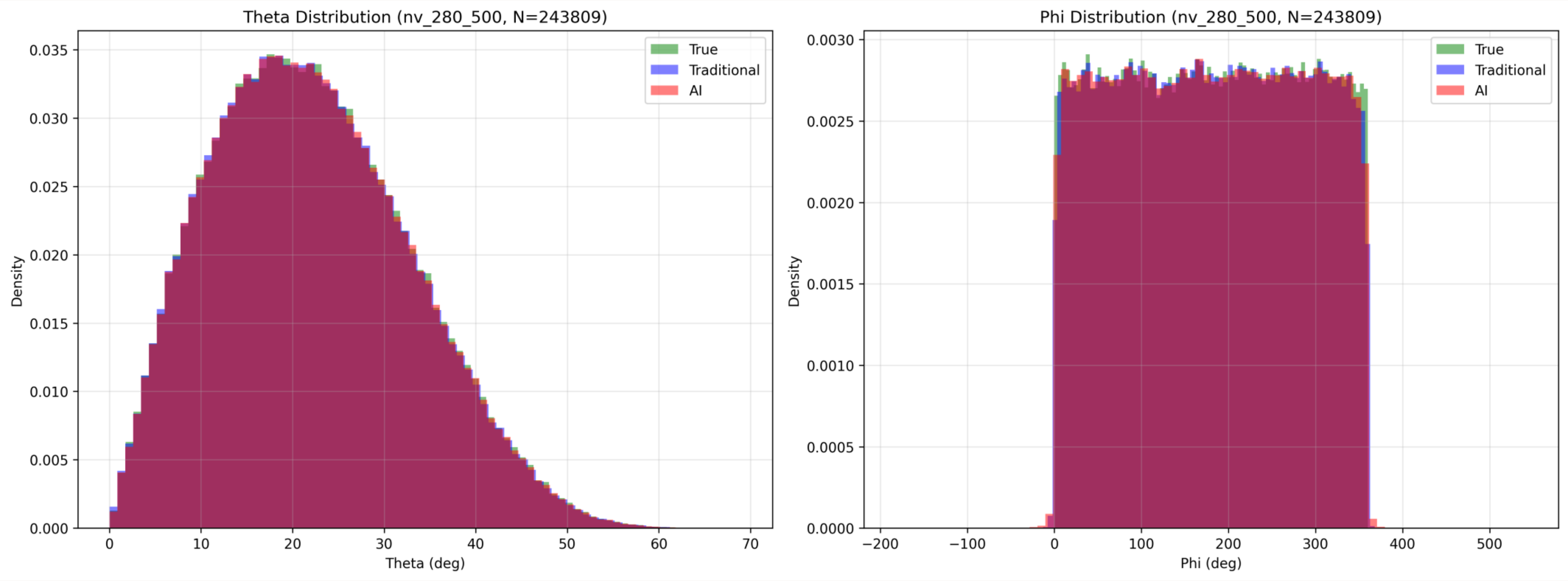
(Gamma ray retention ratio,
Optimal opening angle,
Q factor of Gamma/CRs classification)

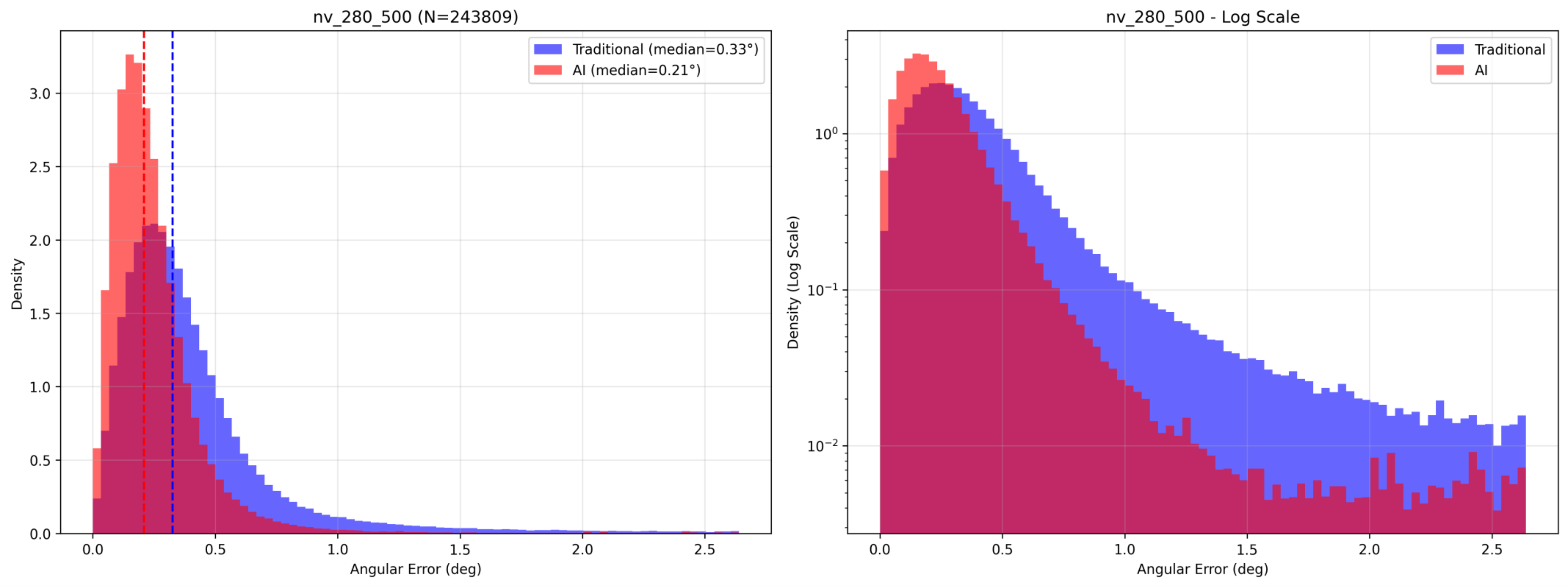
	Optimal opening angle	Q factor
Traditional	1.07	1.83
AI	0.71	3.88
Ratio	1.51	2.12
Overall ratio	$3.2 = \sqrt{N_{b,0}}/\sqrt{N_{b,AI}}$	

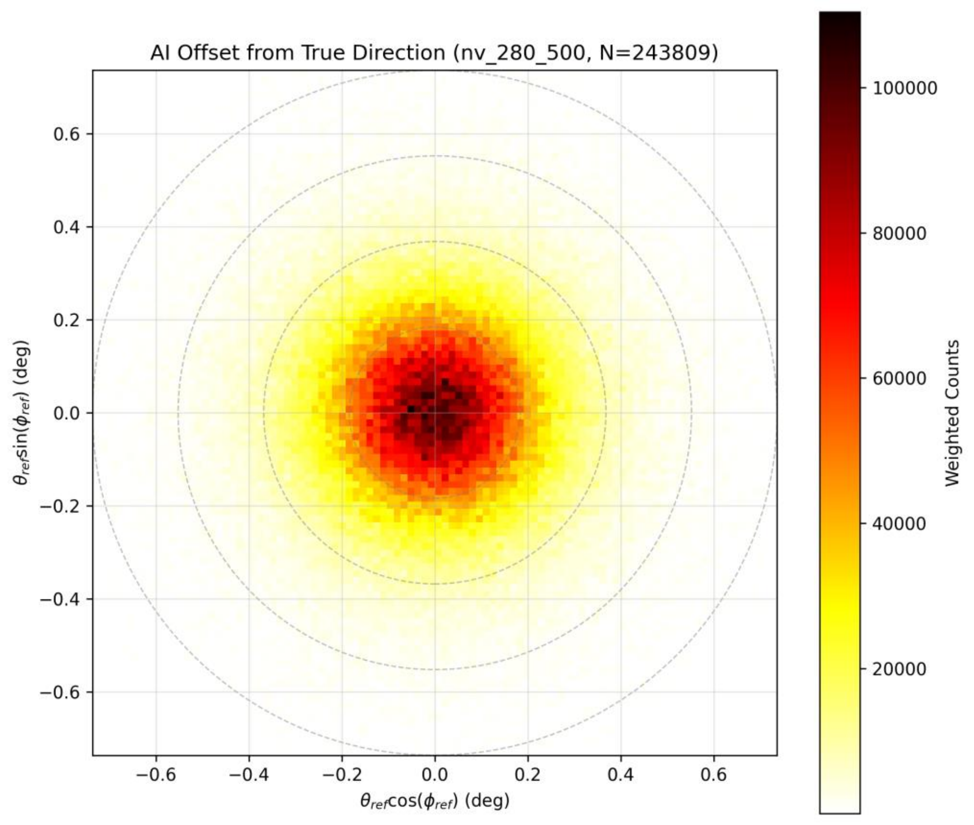
→ With the same gamma retention ratio, the background level is reduced by 10 times.

First Look At Light Curve of AI-Enhanced BOAT GRB will be presented in the last talk tomorrow morning.

Applied to direction reconstruction

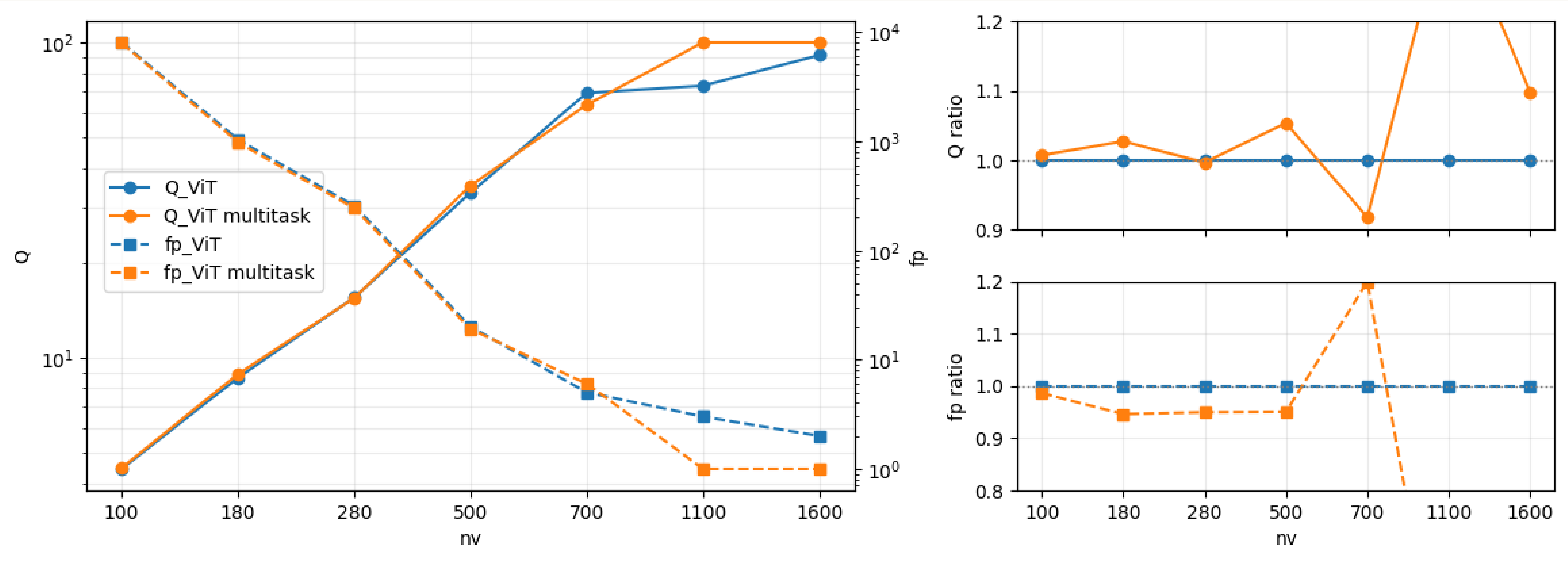






Multi task: Gamma/Proton & Energy reconstruction

Same Setting as optimal G/P



Preliminray energy reconstruction

Error bar: FWHM at linear scale, i.e. energy resolution, ~50%