Adapting high-granularity and compact silicon calorimeters from Higgs factories to dark-sector experiments

S. Huang, A Irles
AITANA group **IFIC, CSIC-UV (Valencia)**

MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES
CSIC CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS
VNIVERSITAT DE VALÈNCIA
IFIC INSTITUT DE FÍSICA CORPUSCULAR
EXCELENCIA SEVERO OCHOA

Financiado por la Unión Europea NextGenerationEU
Plan de Recuperación, Transformación y Resiliencia
GENERALITAT VALENCIANA
GVANEXT Fondos Next Generation en la Comunitat Valenciana
Gen-T
AITANA

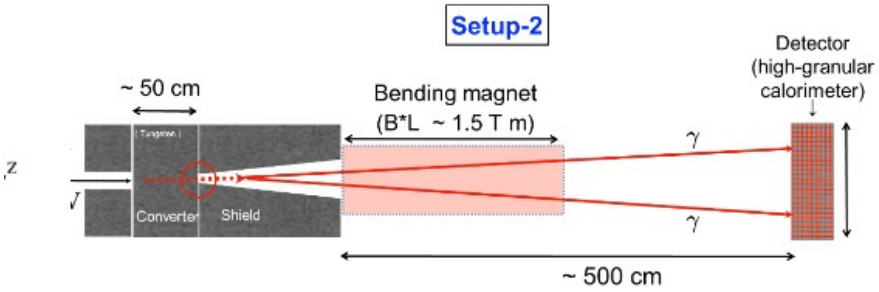
Intro

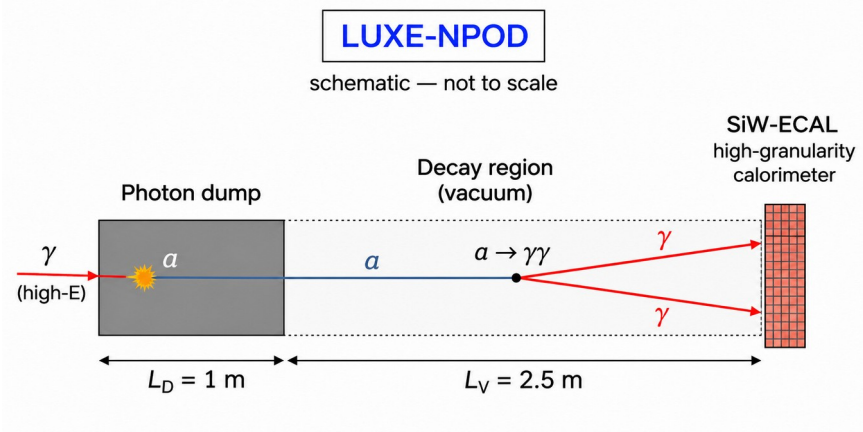
Highly-granular and compact silicon ECAL

for direct new physics searches

- Granular and compact Silicon-ECAL prototypes for Higgs Factories are proposed as the LUXE-NPOD and EBES at KEK (phase I-II) experiments.

EBES – talk at LCWS



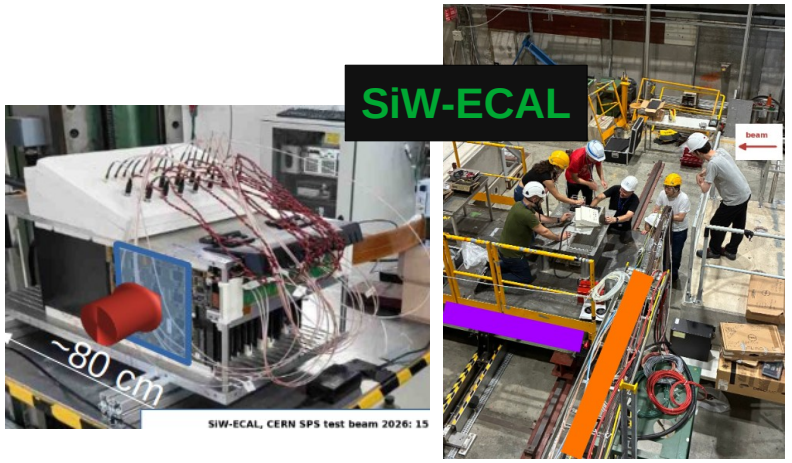


Beam energies are about 4-16 GeV , depending of the experiment.
Final state photons are expected to be of the order of few GeVs.

different challenge wrt the original application (Higgs Factories):
low energy photon kinematics (direction/momentum/entry point)

(potential) detector optimizations

- Assuming existing prototypes
  - The experiments should optimize the layout (beam dump material, positions, etc) matching the detector acceptance. An example in Phys.Rev.D 112 (2025) 11, 112014
- (expensive) optimizations for higher performance:
  - Pad size optimization would require a construction of new prototypes (new sensors, new adapted electronics).
- Further (unexpensive) optimizations:
  - Better algorithms.
  - Tungsten repartition.
  - These can be built around existing prototypes

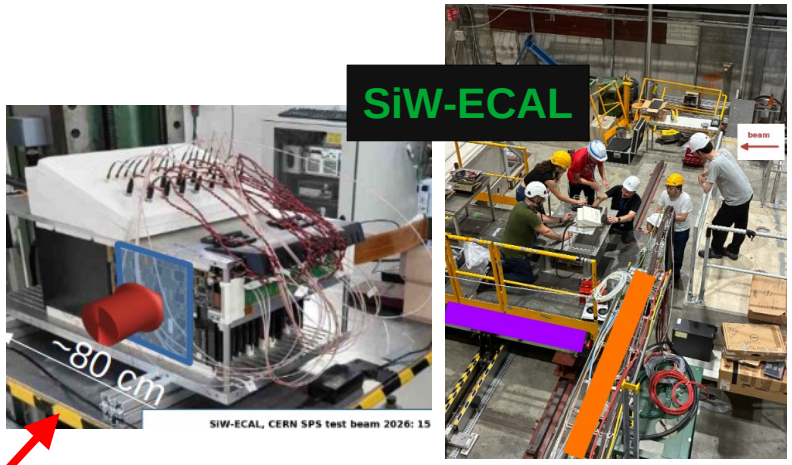




(potential) detector optimizations

- Assuming existing prototypes
  - The experiments should optimize the layout (beam dump material, positions, etc) matching the detector acceptance. An example in Phys.Rev.D 112 (2025) 11, 112014
- (expensive) optimizations for more higher performance:
  - Pad size optimization would require a construction of new prototypes (new sensors, new adapted electronics).
- Further (unexpensive) optimizations:
  - Better algorithms.
  - Tungsten repartition.
  - These can be built around existing prototypes

TODAY





Detector description – default geometry

Detector geometry simulated (the default for LUXE-NPOD)

SiW-ECAL

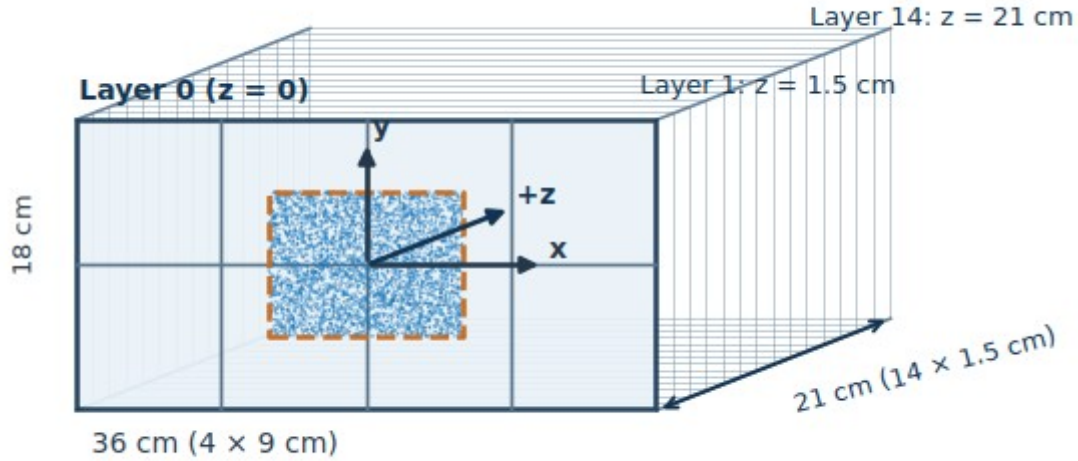


Detector geometry simulated (the default for LUXE-NPOD)

SiW-ECAL

ACTIVE PLANES

15 silicon layers



Twice the surface of current SiW-ECAL
(4x2 sensors = 2 ASUs)

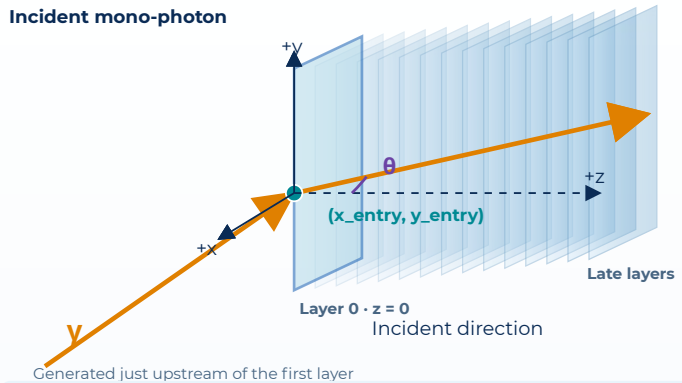
- Implementation in ddsim and LCIO
  - pixelization/digitization (noise smearing) / single cell mip calibration included in the simulation
  - Single hit required to be larger than 0.5MIP

EntryPoint-Direction resolution

Event simulations

A · Incident photon: entry point and direction

Reference ECAL: 15 instrumented SiW layers · schematic, not to scale



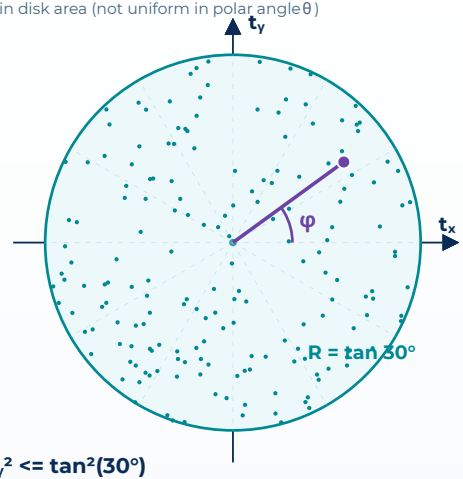
Generated just upstream of the first layer

$t_x = p_x / p_z$ $t_y = p_y / p_z$

$\theta = \text{atan}((t_x^2 + t_y^2)^{(1/2)})$ $\varphi = \text{atan2}(t_y, t_x)$

B · Generation in (t_x, t_y) slope space

Uniform in disk area (not uniform in polar angle θ)



$t_x^2 + t_y^2 <= \tan^2(30°)$

Uniform in area, not uniform in polar angle θ

Generated entry positions: x: -60 to +60 mm · y: -45 to +45 mm · generation plane z ~ -1 mm

Physical angle domain: θ from 0° to 30°; φ spans 0–360°. The shown ray/hits are illustrative.

A · Mono-photon energy grid

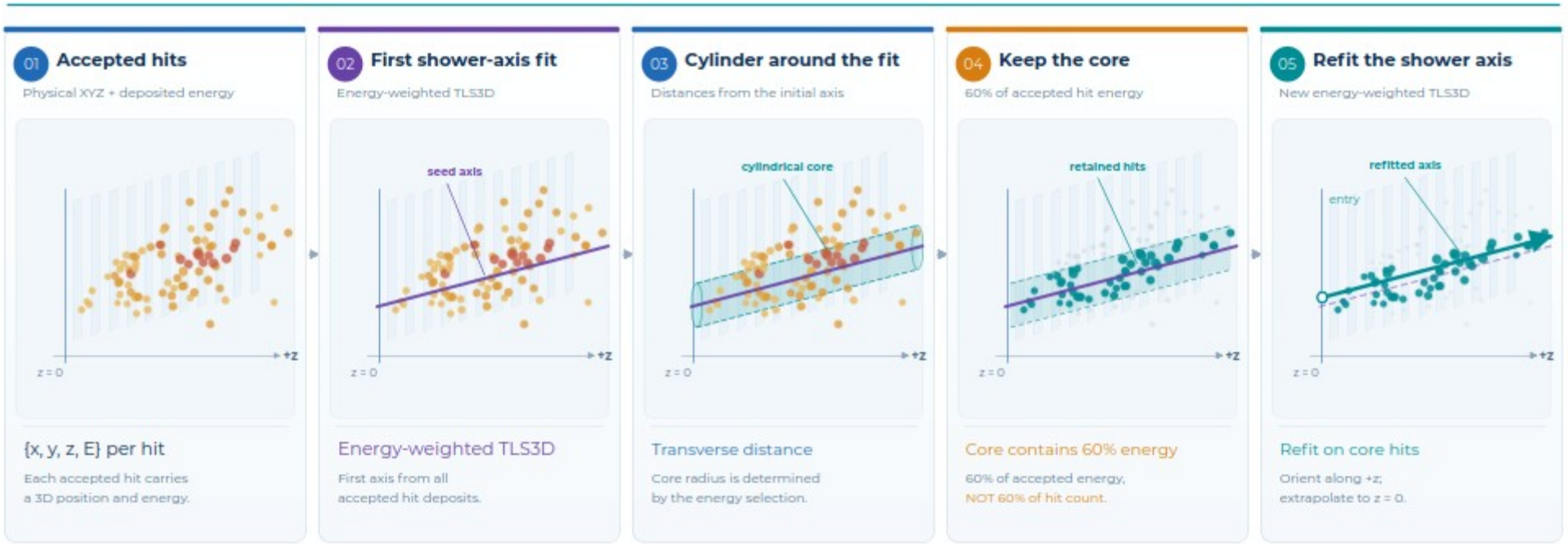
Reference geometry · Direction and Entry share the same simulated photon population



21 monoenergetic samples · ΔE = 0.2 GeV

0.2–4.2 GeV

Classic direction/entrypoint reconstruction



CNN-Direction & CNN-entry

two independent 3D conv networks for complementary observables

Input representation

Physical XYZ grid



5 × 64 × 32 × 15

Five channels:

E · M · Mx · My · Mz

Same event geometry and channel convention; separate network weights

CNN-Direction

96,708 trainable parameters

3D CNN encoder
Learned spatial shower features

Geometry descriptors
Layer-wise features from energy + XYZ
z | x/y | energy

Prediction head
vMF direction task-specific outputs

Reconstructed observable
Predicted unit direction
$\hat{u} = (u_x, u_y, u_z)$
Concentration κ

Directional concentration, not a guaranteed error

CNN-Entry

96,579 trainable parameters

3D CNN encoder
Learned spatial shower features

Geometry descriptors
Layer-wise features from energy + XYZ
z | x/y | energy

Prediction head
Gaussian-confidence task-specific outputs

Reconstructed observable
Predicted entry (z = 0)
(x_entry, y_entry) [mm]
Confidence score σ_event

σ_event: confidence ranking, not calibrated uncertainty

Training loss functions
(code-verified)

CNN-Direction · full vMF negative log-likelihood

$L_D = \text{mean}[-\log C_3(\kappa) - \kappa\,(\mu \cdot u_{true})]$

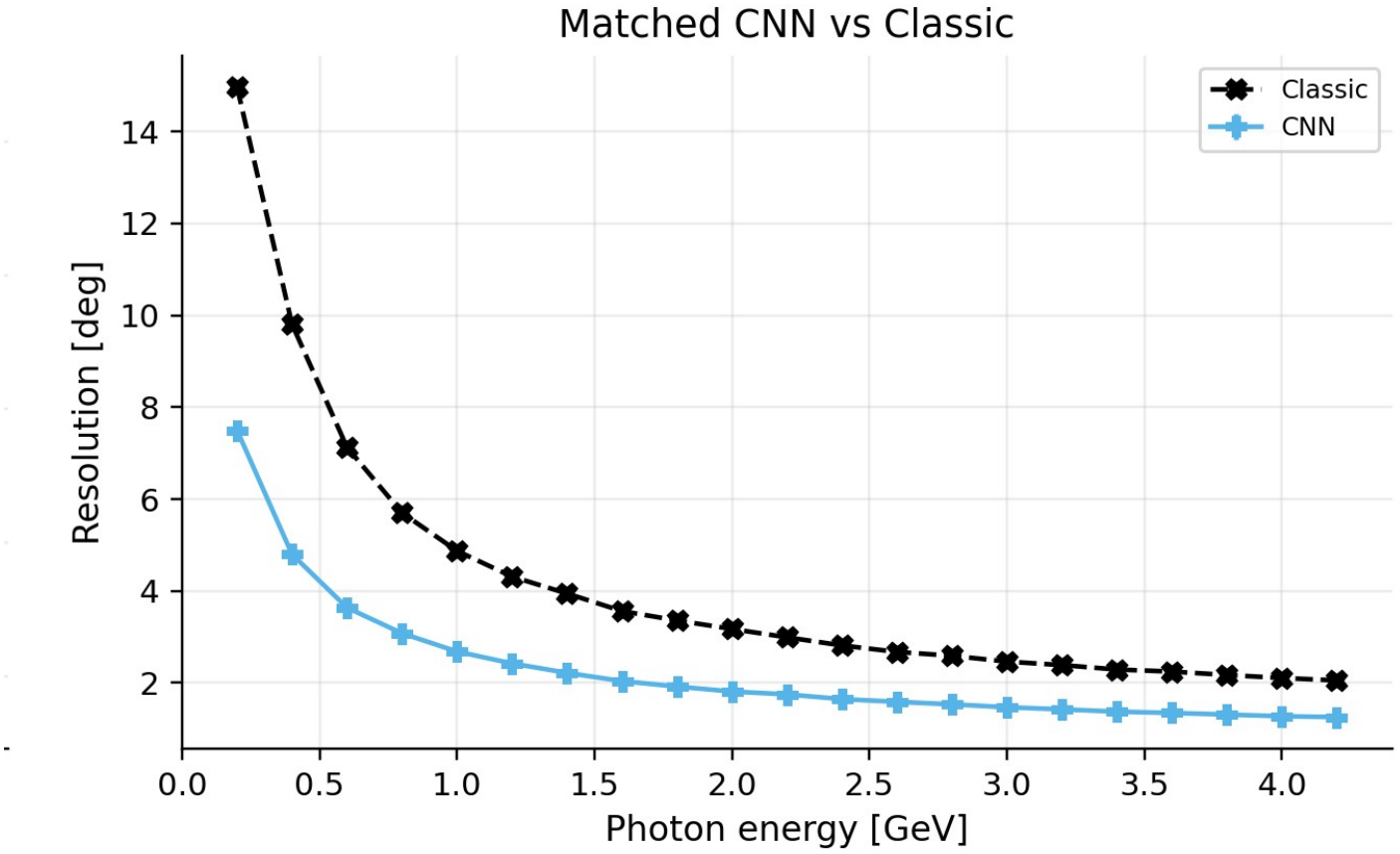
$C_3(\kappa) = \kappa/[4\pi \sinh(\kappa)] \cdot \kappa = \text{softplus}(\rho) + 10^{-6}$

CNN-Entry · 2D Gaussian-confidence NLL

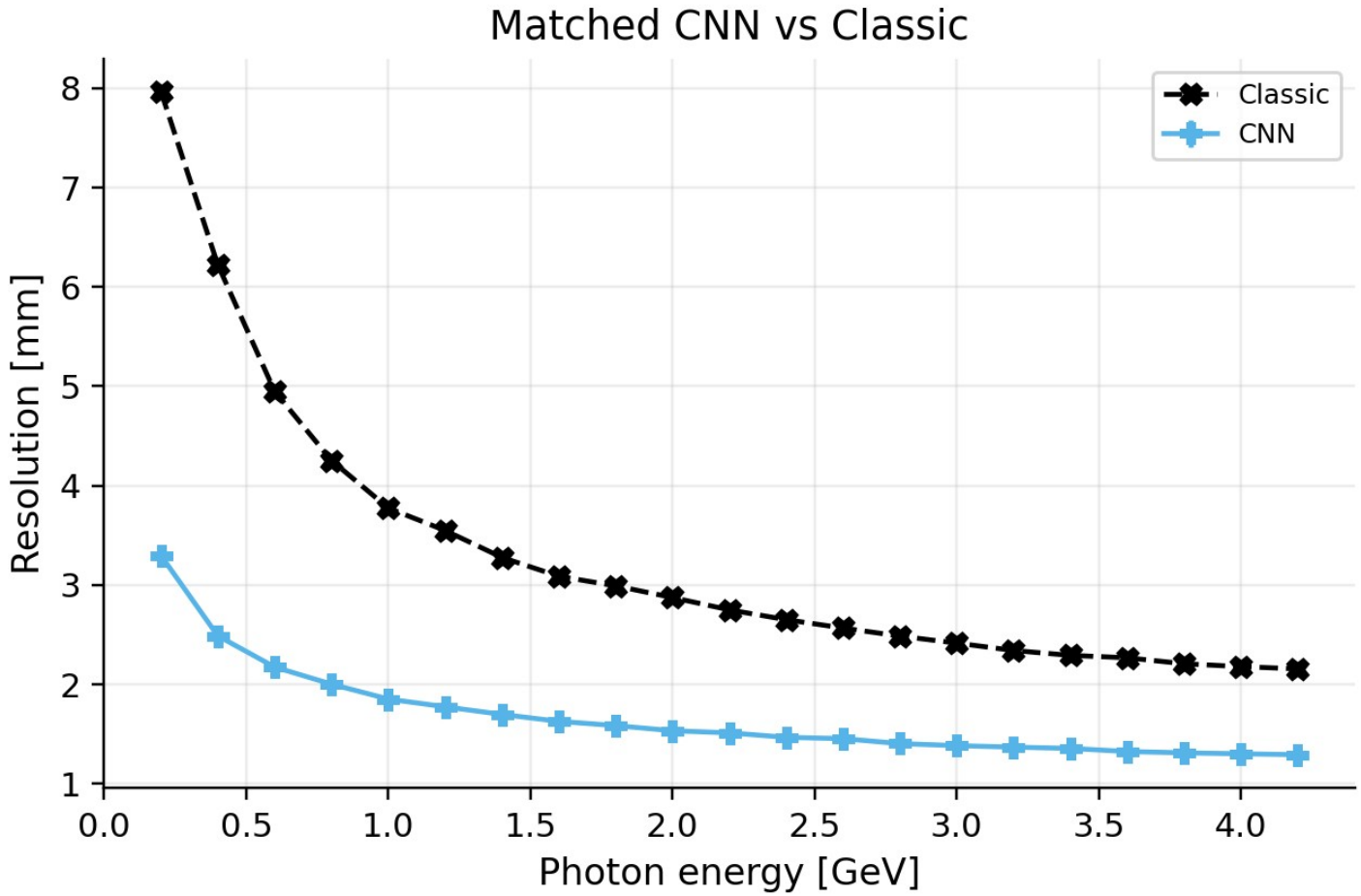
$L_E = \text{mean}[\tau\,||r - \mu||^2 / (2(5\text{ mm})^2) - \log \tau]$

$\tau = \text{softplus}(\rho) + 10^{-6} \cdot \sigma_{event} = (5\text{ mm})/\text{sqrt}(\tau)$

CNN-Direction - results



CNN-Entry - results



Energy resolution

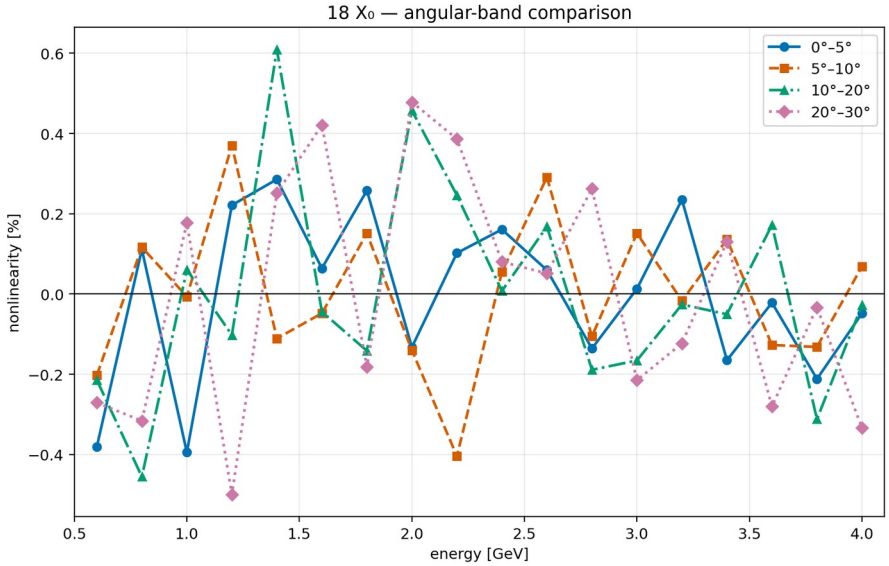
Energy resolution

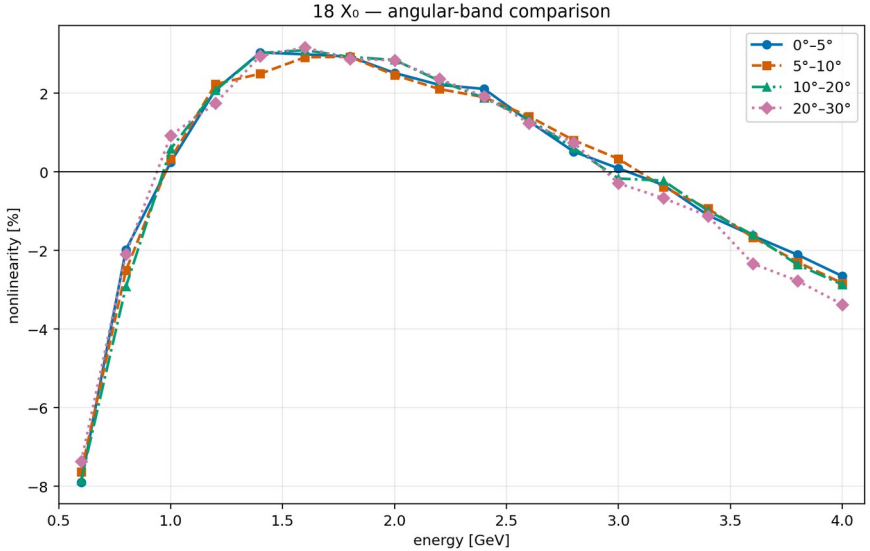
no methodology presented, only results

check backup slides for more

Classic Energy estimation

raw estimations





$$X_{\mathrm{A}} = \sum_{l} w_l \sum_{v \in H_l} E_v$$

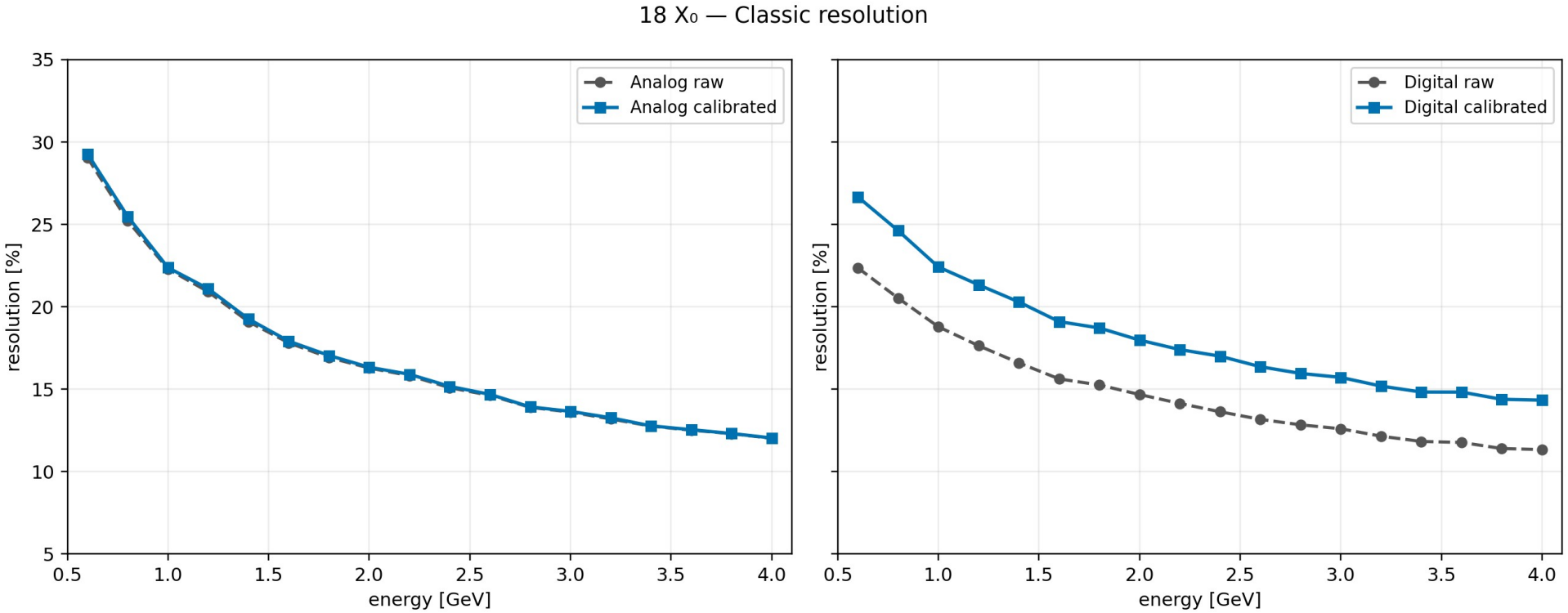
E_v : deposited energy in voxel v [MIP]
w_l : tungsten-layer weight (relative to 4.2 mm)

$$X_{\mathrm{D}} = \sum_{l} w_l \, N_l^{\mathrm{hits}}$$

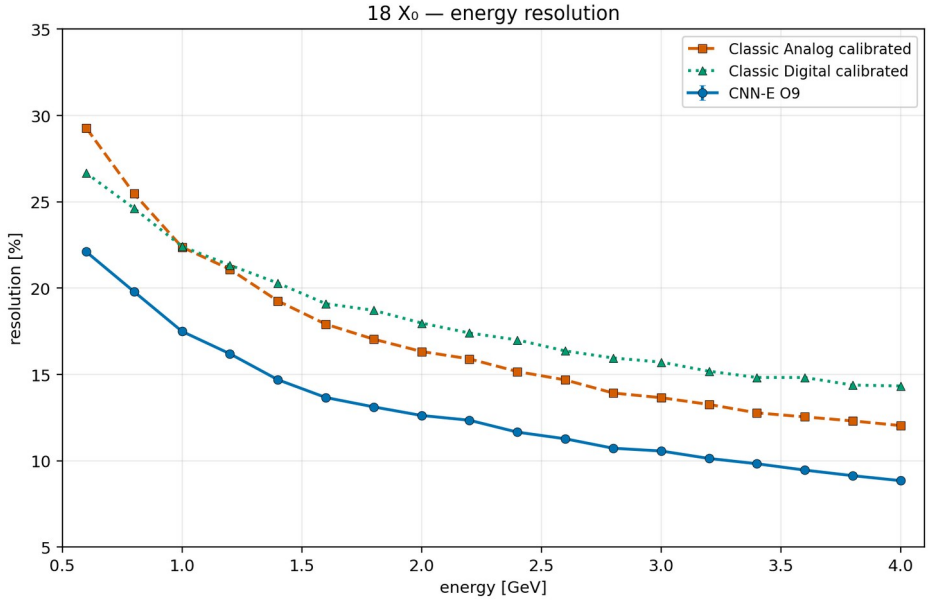
N_l^{hits} : number of hit voxels in layer l (unique voxels)
w_l : tungsten-layer weight (relative to 4.2 mm)

Classic Energy Resolution estimations

raw vs calibrated estimations



CNN-E vs classic



Resolution fit parameters

18 X_0 | 0°–5° | 0.6–4.0 GeV

Method	S [% $\sqrt{\text{GeV}}$]	N [% GeV]	C [%]
Classic Analog · 3-term	21.85	3.60	5.00
Classic Digital · 3-term	19.62	≈ 0	10.87
CNN-E · 2-term (original)	17.22	—	3.02

CNN-E: $R(E) = \sqrt{S^2/E + C^2}$

Classic: $R(E) = \sqrt{S^2/E + N^2/E^2 + C^2}$

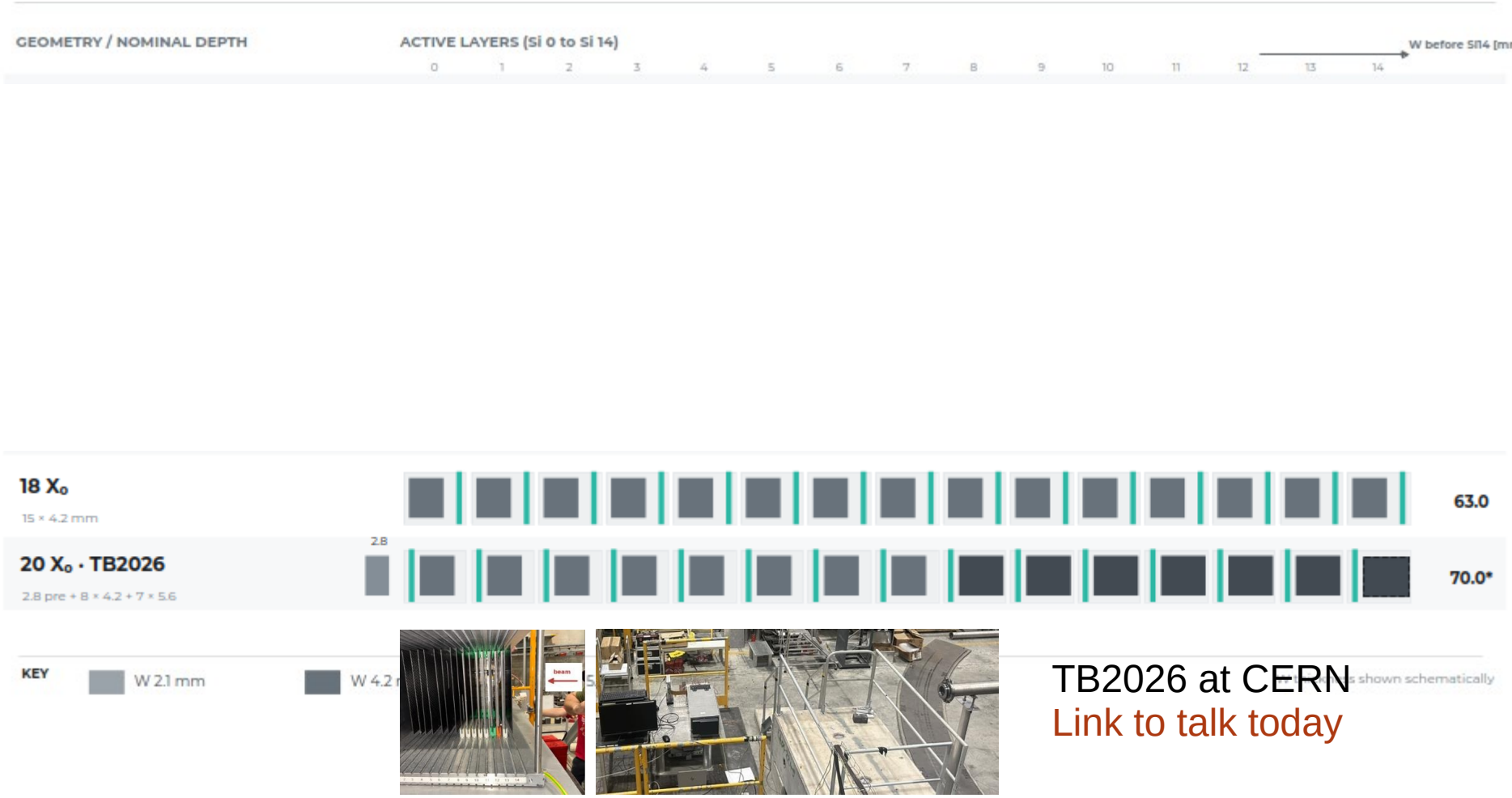
Classic parameters: descriptive unweighted fits to calibrated resolution; no parameter uncertainties estimated.

Geometry optimization

Geometry alternatives



Geometry alternatives



TB2026 at CERN

Link to talk today

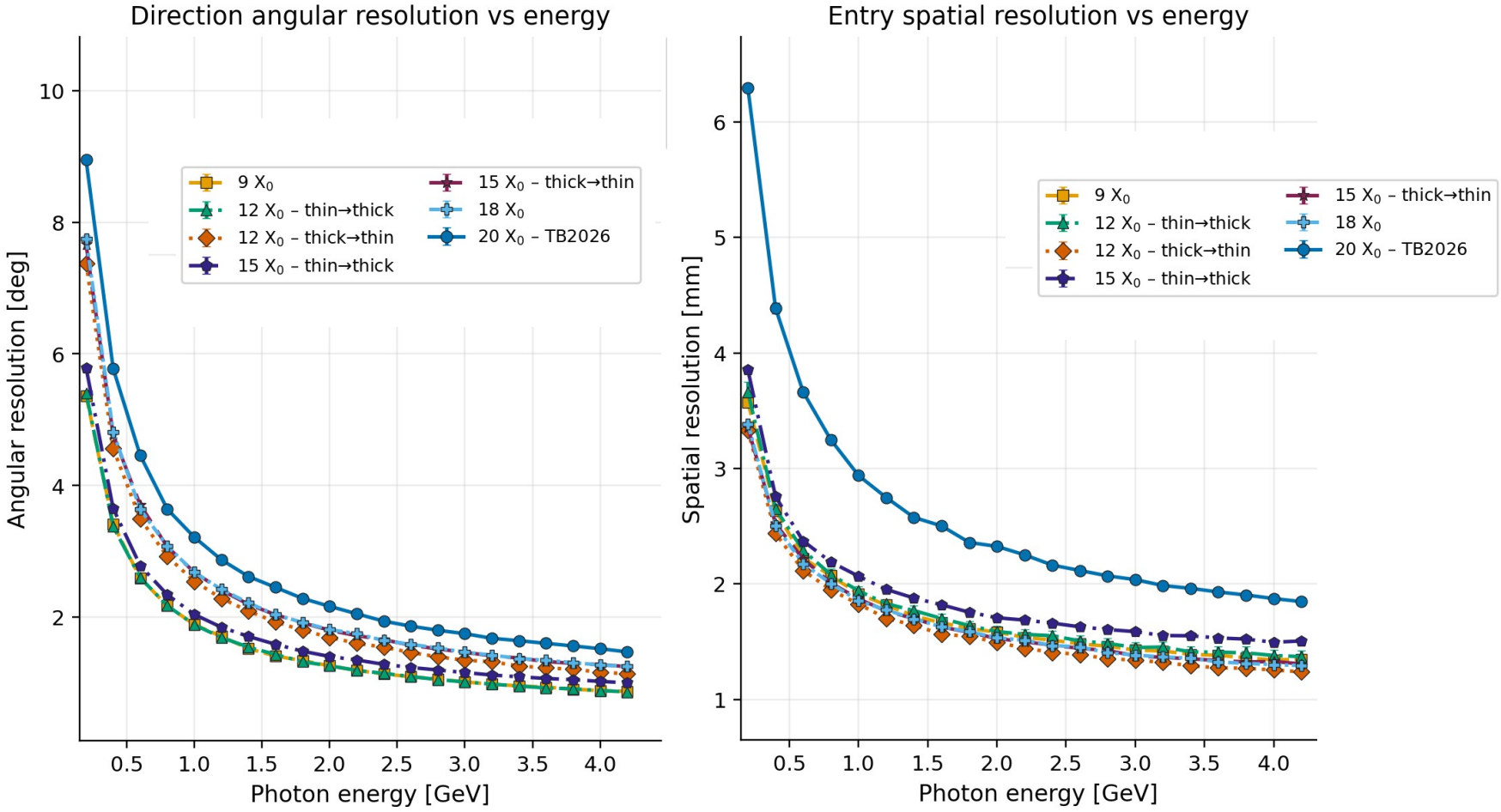
Geometry alternatives



All layouts: 15 active Si planes, same non-W material and 15 mm nominal pitch; TB2026 also reverses the Si/W order.

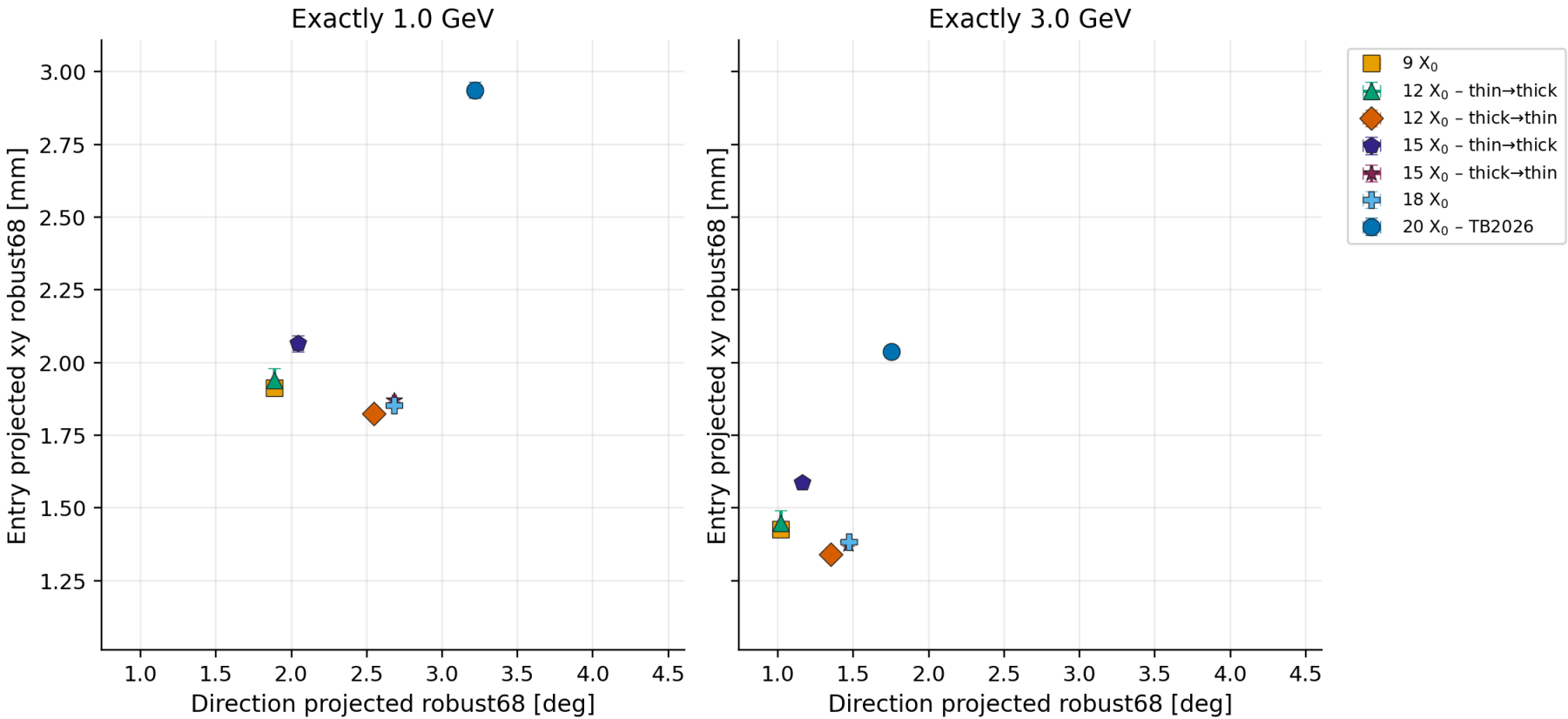
TB2026: 75.6 mm physical W, but only 70.0 mm before Si14 (*). First W 2.8 mm; final 5.6 mm lies downstream of Si14.

Results – direction & entry

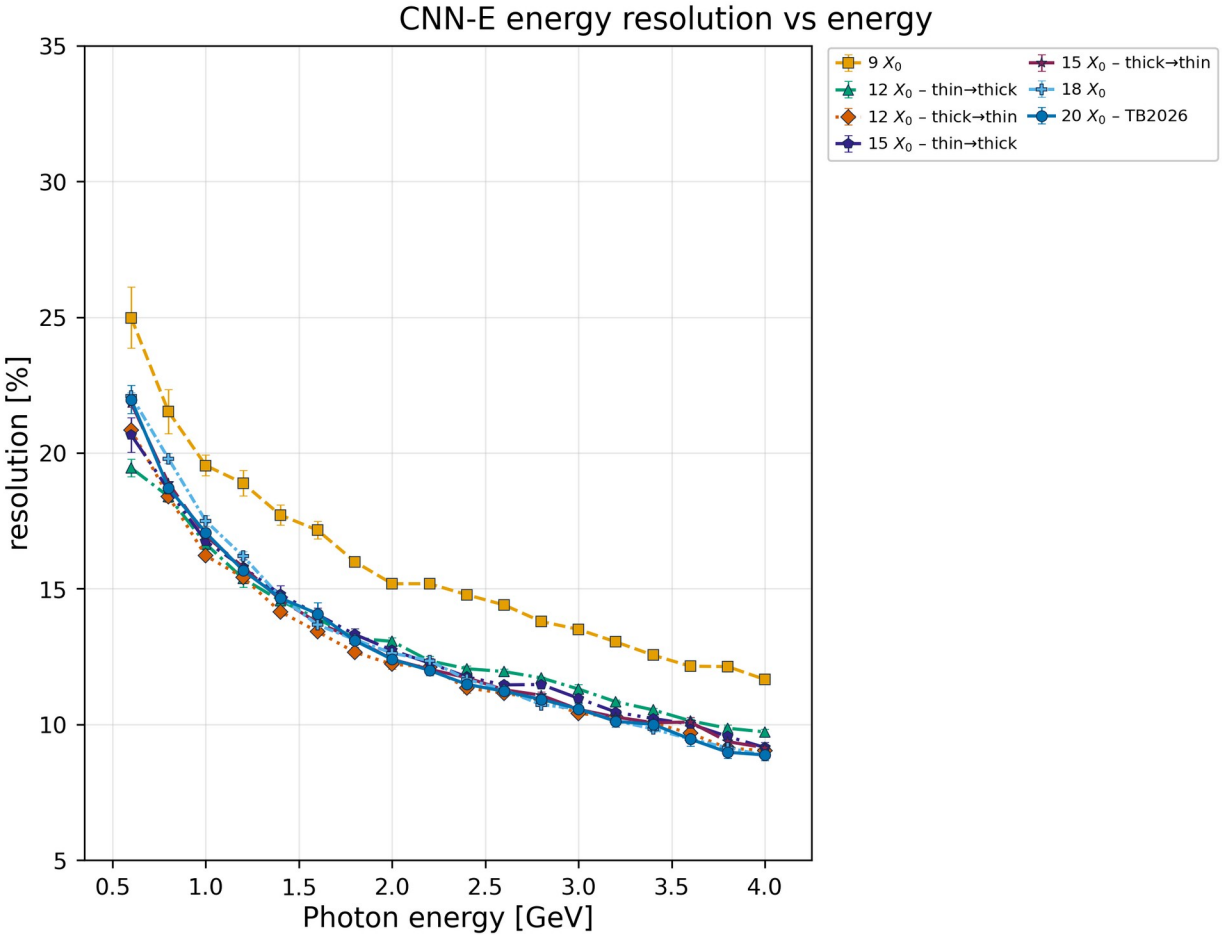


Arithmetic seed mean ± sample SD across three seeds (ddof=1)

Results – Direction vs Entry



Results - energy



Arithmetic seed mean ± sample SD across three seeds (ddof=1)

Summary

- With advanced algorithms with convolutional 3D networks we observe substantial improvements in kinematic reconstruction observables for photons:
  - Up to ~43% improvement in angle resolution
  - Up to ~46% improvement in entry point resolution
  - Up to ~33% improvement in energy resolution

- With affordable modifications of the detector layout, we can enhance different reconstruction capabilities.
  - Thin→thick layouts favour angular reconstruction, while thick→thin layouts slightly favour entry-point reconstruction.

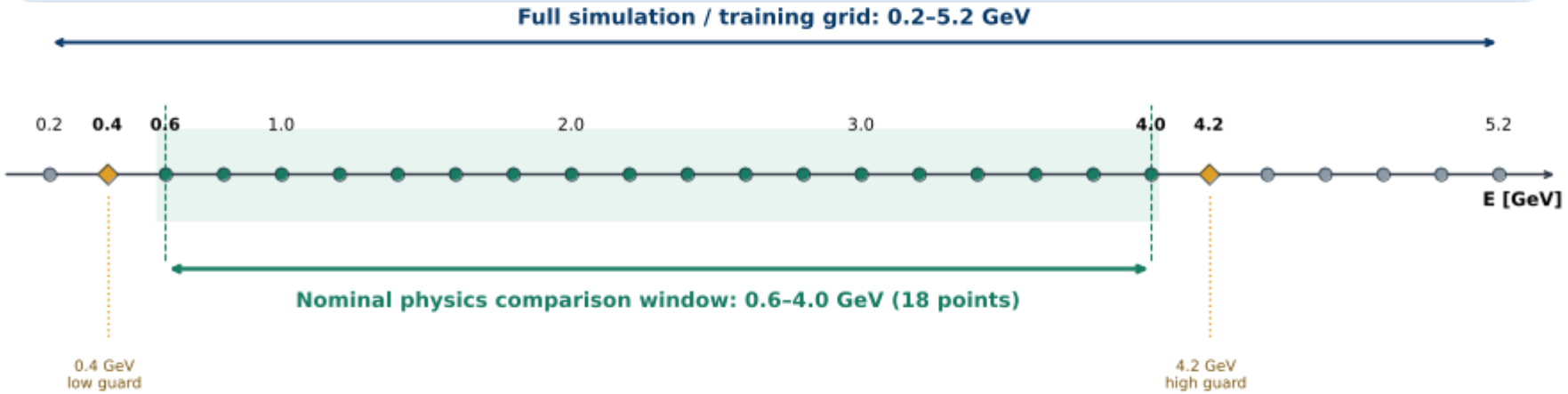
- With small (unexpensive) modifications, a flexible detector that can be adapted to different goals and perform in different experiments.

backup

Simulation for energy resolution studies

Simulated incident photons (γ) · Energy grid

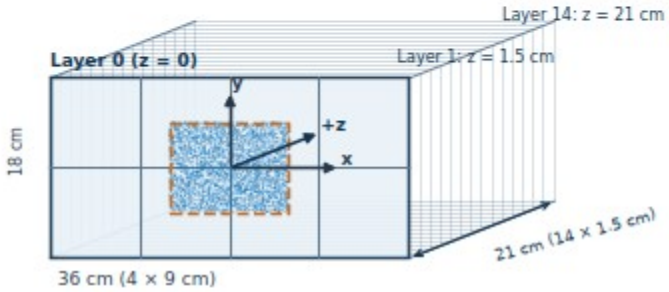
26 monoenergetic photon samples · ΔE = 0.2 GeV · 40,000 events per energy and angular band

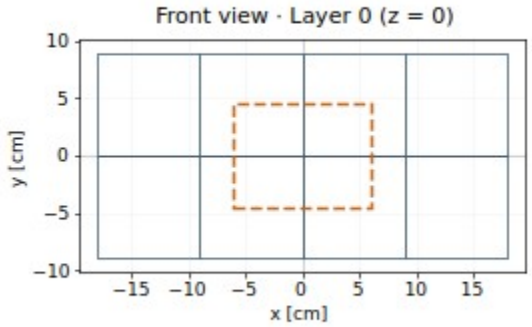


Extended energy coverage provides margins around the physics window to mitigate edge effects.

Simulation for energy resolution studies

A · ECAL entry points and active-layer geometry

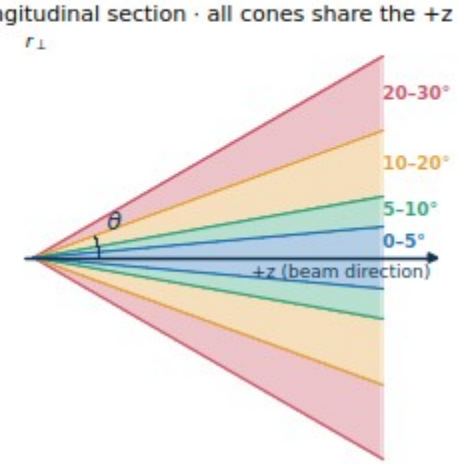


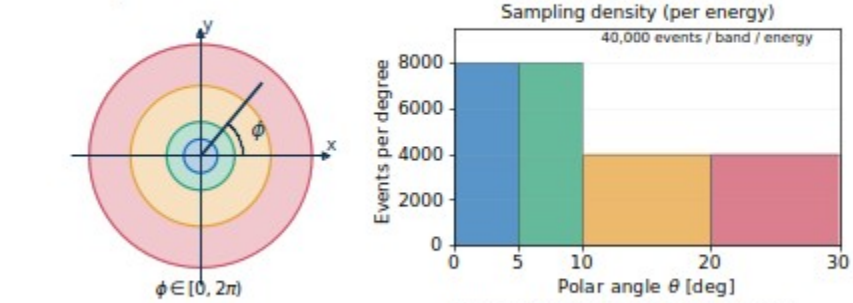


Blue dots: illustrative photon entries on Layer 0 (z = 0).

Generation plane $z = -1$ mm; $x \in [-6, 6]$ cm; $y \in [-4.5, 4.5]$ cm.

B · Angular bands around the beam axis (+z)





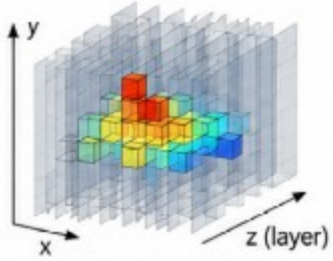
Uniform in θ per band; $\varphi \in [0, 2\pi)$ uniform.

40,000 events / energy / angular band.

Classic Energy Resolution estimations

Analog (energy sum)

1 Event observable (Analog)



Include only hits with energy > 0.5 MIP using tungsten-layer weights.

$$X_{\mathrm{A}} = \sum_{l} w_l \sum_{v \in H_l} E_v$$

E_v : deposited energy in voxel v [MIP]
w_l : tungsten-layer weight (relative to 4.2 mm)

2 Energy calibration (linear)

Fit the mean response in calibration bins (0.4 – 4.2 GeV):

$$\langle X_{\mathrm{A}} \rangle = a + b \cdot E_{\mathrm{true}}$$

Calibrated energy

$$E_{\mathrm{rec}}^{\mathrm{A}} = \frac{X_{\mathrm{A}} - a}{b}$$

3 Resolution (Analog)

Raw (no calibration)

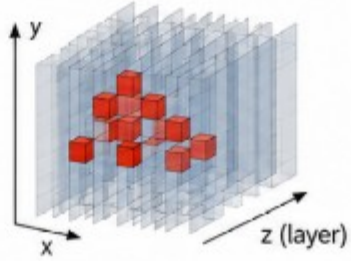
$$R_{\mathrm{raw}}^{\mathrm{A}} = \frac{\sigma(X_{\mathrm{A}})}{\langle X_{\mathrm{A}} \rangle}$$

Calibrated

$$R_{\mathrm{cal}}^{\mathrm{A}} = \frac{\sigma(E_{\mathrm{rec}}^{\mathrm{A}})}{E_{\mathrm{true}}}$$

Digital (hit counting)

1 Event observable (Digital)



Count unique voxels with energy > 0.5 MIP (identical hit threshold).

$$X_{\mathrm{D}} = \sum_{l} w_l \, N_l^{\mathrm{hits}}$$

N_l^{hits} : number of hit voxels in layer l (unique voxels)
w_l : tungsten-layer weight (relative to 4.2 mm)

2 Energy calibration (monotonic PCHIP)

Obtain a monotonic calibration from the mean response in calibration bins (0.4 – 4.2 GeV):

$$\langle X_{\mathrm{D}} \rangle \rightarrow E_{\mathrm{true}} \quad \text{(monotonic)}$$

Use linear secant extrapolation at the endpoints (no clipping).

Calibrated energy

$$E_{\mathrm{rec}}^{\mathrm{D}} = \mathcal{C}_{\mathrm{PCHIP}}^{-1}(X_{\mathrm{D}})$$

3 Resolution (Digital)

Raw (no calibration)

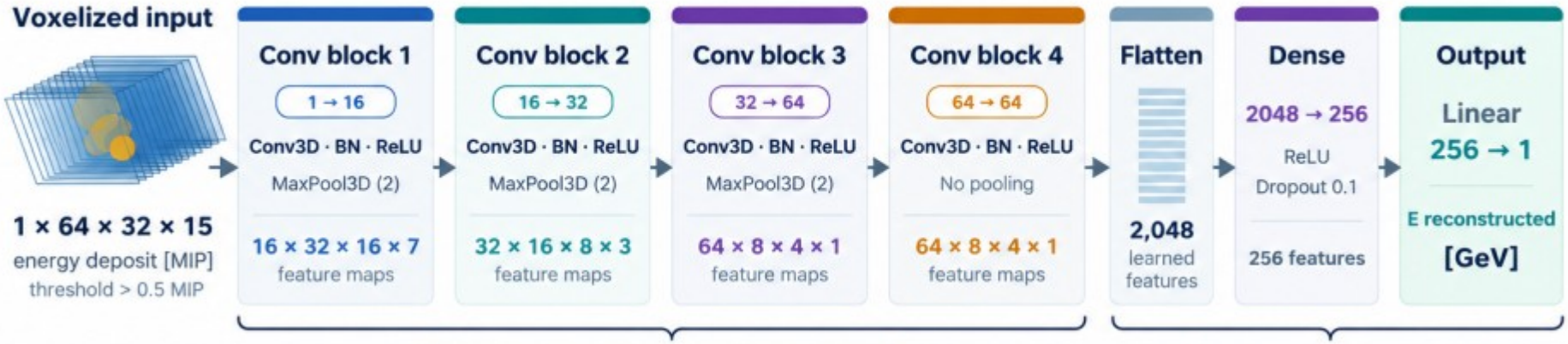
$$R_{\mathrm{raw}}^{\mathrm{D}} = \frac{\sigma(X_{\mathrm{D}})}{\langle X_{\mathrm{D}} \rangle}$$

Calibrated

$$R_{\mathrm{cal}}^{\mathrm{D}} = \frac{\sigma(E_{\mathrm{rec}}^{\mathrm{D}})}{E_{\mathrm{true}}}$$

CNN-E

3D convolutional regressor – for photon energy reconstruction by 15 layer SiW-ECAL like detector



INPUT REPRESENTATION

- 3D grid of energy deposits in the 15 instrumented SiW-ECAL layers
- Input tensor shape: **1 × 64 × 32 × 15** (C × X × Y × Z)
- Energies in MIP with threshold > 0.5 MIP
- No hand-crafted shower variables are provided (e.g. total energy, shower barycenter, depth of maximum, …).

LEARNED FEATURES AND OUTPUT

- Four 3D convolutional blocks with BatchNorm, ReLU and max pooling (only in blocks 1–3).
- The network learns **2,048 internal features** from the voxelized input.
- A fully connected layer (256 units, ReLU, dropout 0.1) is used for energy regression.
- Linear output (1 unit) providing the reconstructed photon energy in GeV.

CNN-E

Training, dataset split & checkpoint selection

01 DATASET SPLIT · PER ENERGY / GEOMETRY



Pairwise disjoint partitions · within each energy bin

02 TRAINING RECIPE

Optimizer	**Adam**	Early stopping	**patience 20**
Learning rate	**5×10^{-5}**	Improvement Δ	**0.001**
Weight decay	**1×10^{-5}**	Scheduler	**none**
Batch size	**100**	Dropout	**0.10**
Max. epochs	**80**	Independent seeds	**3**

03 CHECKPOINT SELECTION · PHYSICS-AWARE

1 TRAIN
For each seed

70% training split
Adam + hybrid loss
Store epoch candidates

→

2 VALIDATE
At each epoch

10% checkpoint split
Track loss & physics
No final-test access

→

3 RANK EPOCHS
Lexicographic priorities

1 Mean absolute energy bias
2 Maximum / RMS bias
3 Energy-response stability
4 Robust68 + width agreement
5 Val. loss / epoch tie-break

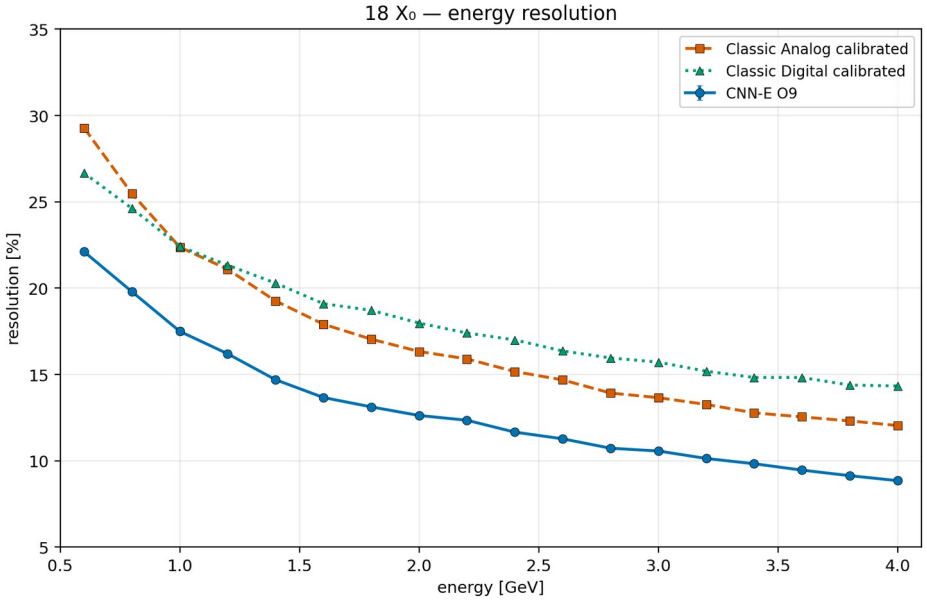
→

4 FREEZE + EVALUATE
After selection

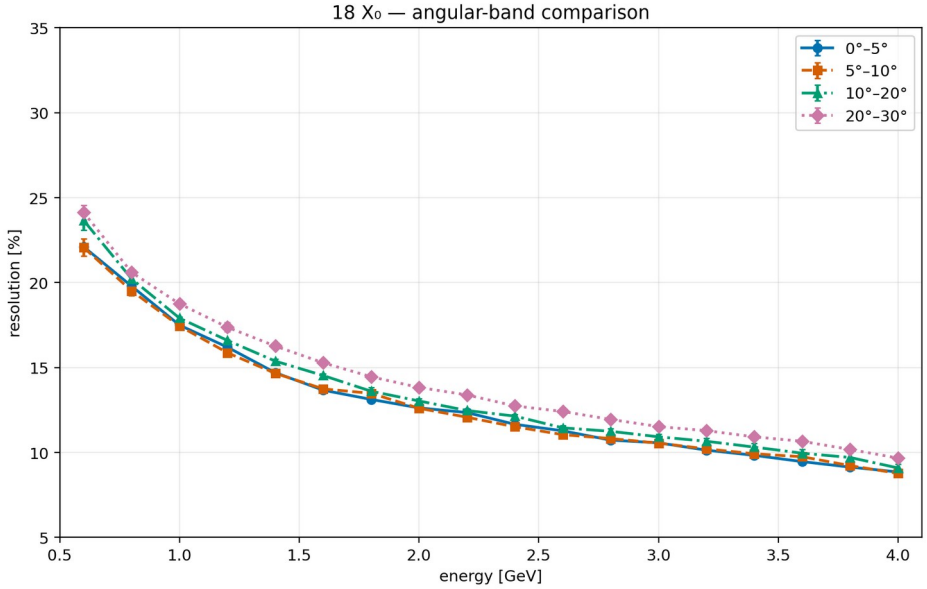
1 checkpoint per seed
Evaluate on held-out 20%
Report 3-seed mean ± SD

HYBRID LOSS $L = \langle \Delta E^2 \rangle + 0.5 \times \langle (\Delta E / E_{true})^2 \rangle$ with $\Delta E = E_{pred} - E_{true}$

CNN-E vs classic



CNN-E vs angle



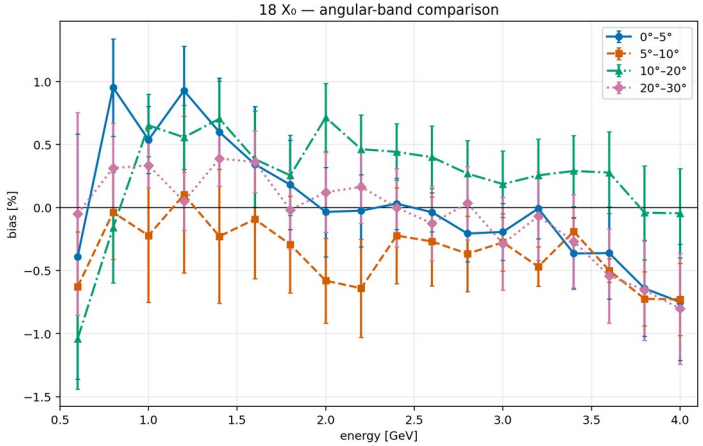
Resolution fit parameters — 18 X_0 | 0°–5° | 0.6–4.0 GeV

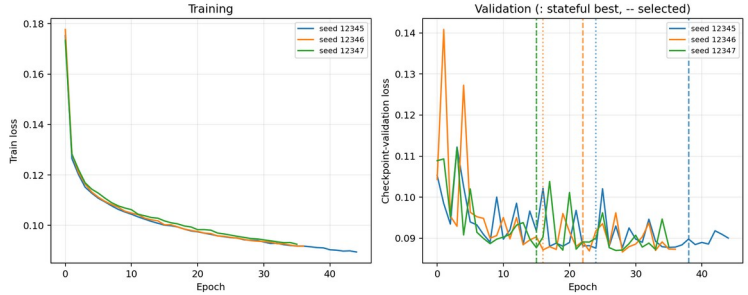
Method	S [% $\sqrt{\text{GeV}}$]	N [% GeV]	C [%]
Classic Analog · 3-term	21.85	3.60	5.00
Classic Digital · 3-term	19.62	≈ 0	10.87
CNN-E · 2-term (original)	17.22	—	3.02

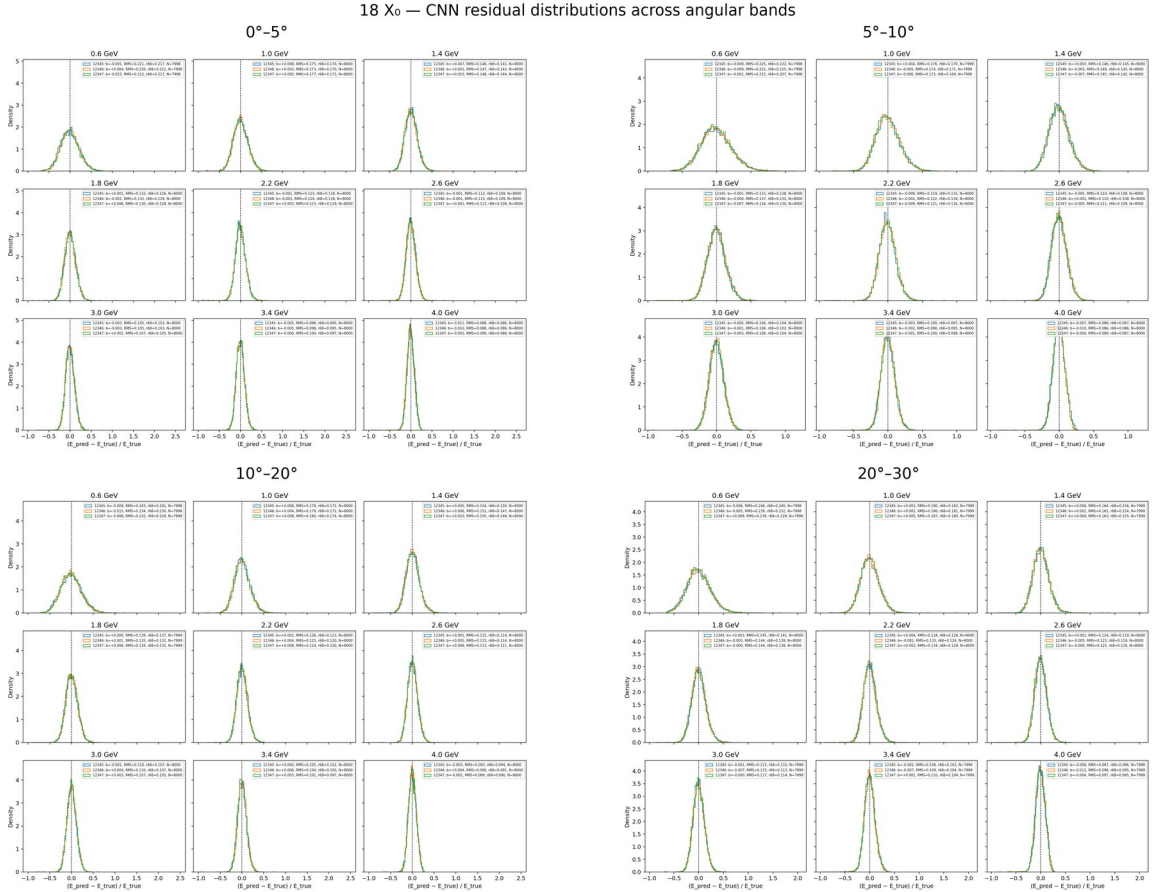
CNN-E: $R(E) = \sqrt{S^2/E + C^2}$ Classic: $R(E) = \sqrt{S^2/E + N^2/E^2 + C^2}$

Classic parameters: descriptive unweighted fits to calibrated resolution; no parameter uncertainties estimated.

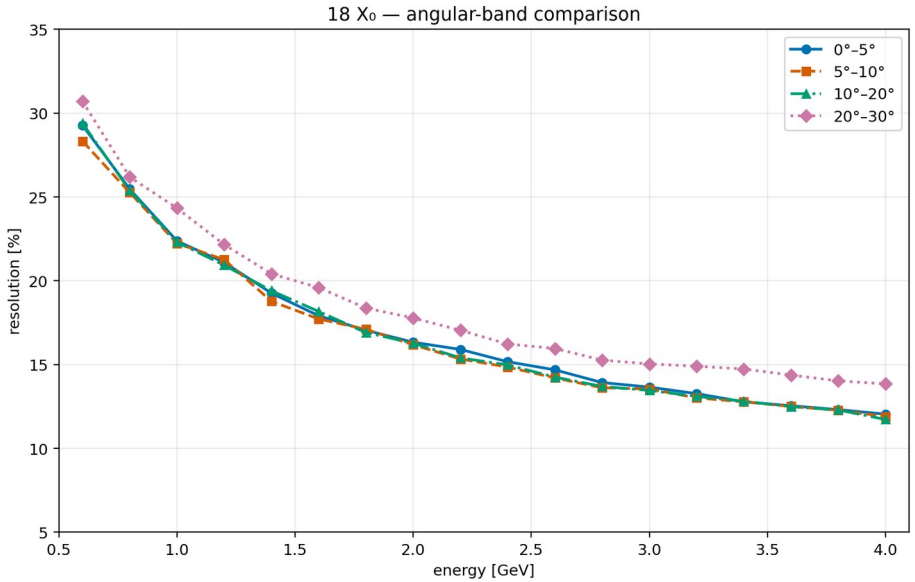
CNN-E


18 X₀ — angular-band comparison
0°–5°
5°–10°
10°–20°
20°–30°
bias [%]
energy [GeV]

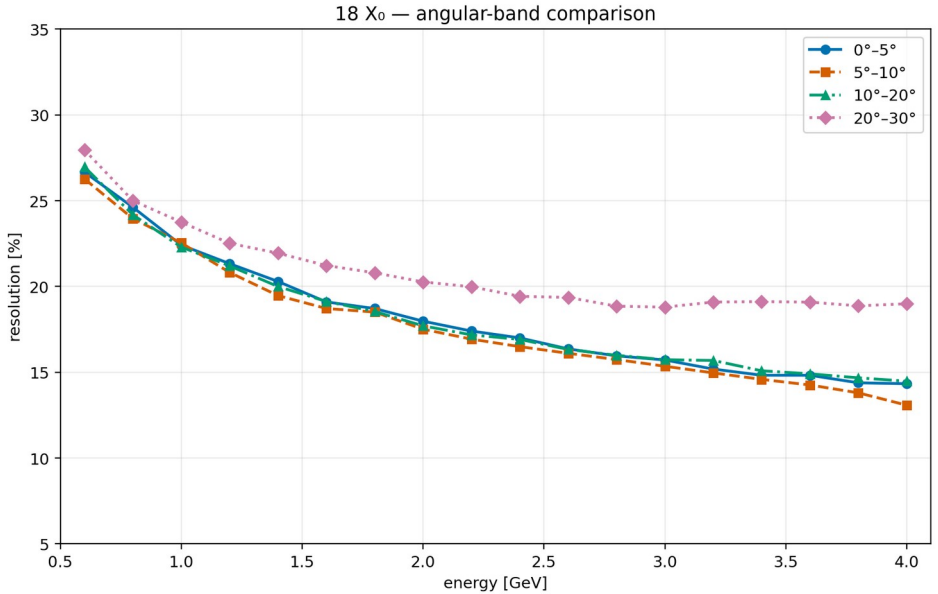

Training
Validation (: stateful best, -- selected)
seed 12345
seed 12346
seed 12347
Train loss
Checkpoint-validation loss
Epoch


18 X₀ — CNN residual distributions across angular bands
0°–5°
5°–10°
10°–20°
20°–30°
0.6 GeV
1.0 GeV
1.4 GeV
1.8 GeV
2.2 GeV
2.6 GeV
3.0 GeV
3.4 GeV
4.0 GeV
Density
(E_pred − E_true) / E_true

Energy res. vs angle



classic-analogue



classic-digital

CNN-E

3D convolutional regressor – for photon energy reconstruction by 15 layer SiW-ECAL like detector

CNN-E-15layerSiWECAL

3D convolutional neural network for ECAL energy reconstruction

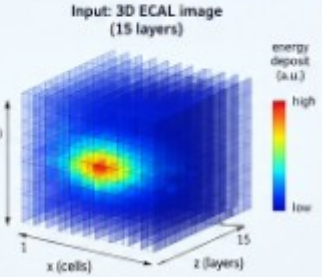
1. Goal and input

Objective

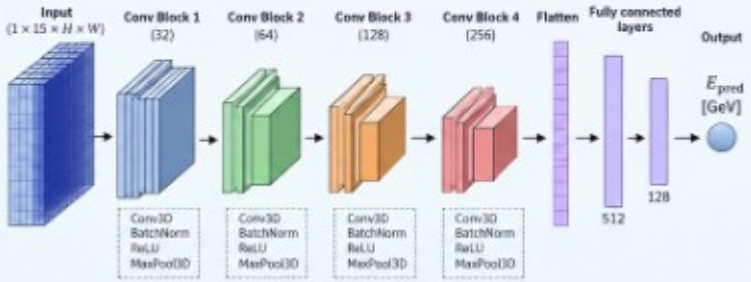
Predict the incident electron energy E_{true} from the 3D energy deposition pattern in the 15-layer SiW ECAL.

Input

- 3D voxelised ECAL image
- 15 longitudinal layers (SiW sampling)
- Transverse granularity defined by the detector geometry
- Single-channel input (deposited energy)
- Standard preprocessing (log-scaling and normalization, see recipe)



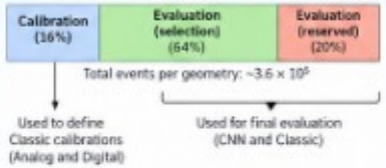
2. Network architecture



3. Training recipe

Training samples	Simulation (Oct 27 campaign) 8 detector geometries
Energy range (training)	0.4 – 4.2 GeV (9 discrete energies)
Angular selection	0° – 5° (00_05 band)
Train / validation split	80% / 20% (event level, no overlap)
Loss function	Mean squared error (MSE) on energy
Optimizer	Adam
Initial learning rate	1×10^{-3} (with scheduler)
Batch size	64
Number of epochs	200
Data augmentation	None (standard voxel input)
Input preprocessing	$\log(1 + E_{dep})$ and normalization (geometry dependent)
Checkpoint	Best validation loss (early stopping)
Random seeds	3 independent trainings (final results are averaged)

4. Dataset split (per geometry)



- The same split is used for all geometries.
- Calibration and evaluation samples are independent (no event overlap).
- CNN training uses only the training dataset (standard train/validation split).
- Final CNN results are obtained by averaging 3 independent trainings with different random seeds.

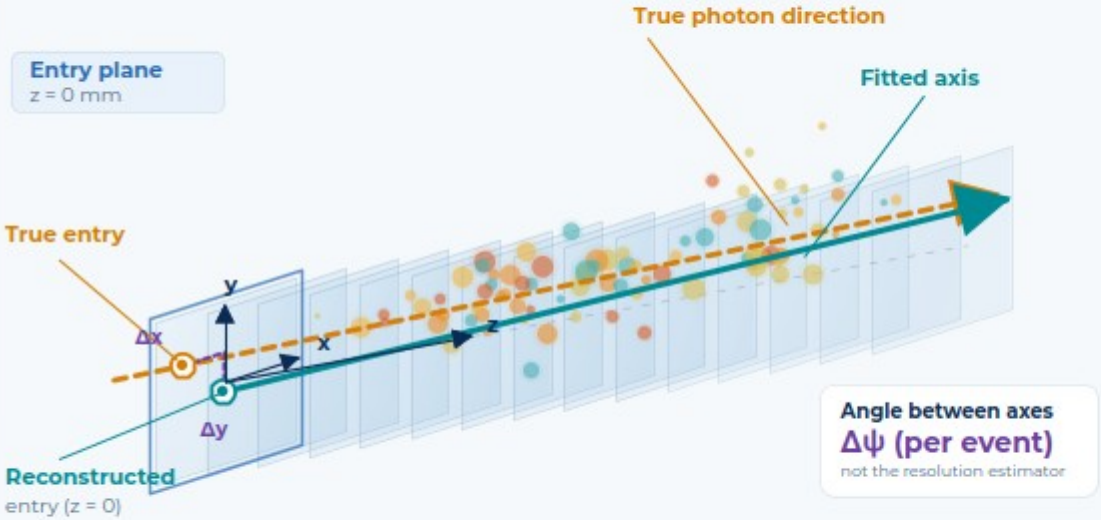
5. Model parameters and reference numbers

Quantity	Value
Input shape	$1 \times 15 \times H \times W$
Convolutional blocks	4 (32, 64, 128, 256 filters)
Kernel size	$3 \times 3 \times 3$
Pooling	$2 \times 2 \times 2$ (MaxPool3D)
Fully connected layers	$512 \rightarrow 128 \rightarrow 1$
Activation	ReLU (hidden layers)
Normalization	Batch Normalization (after each conv)
Total trainable parameters	$\approx 1.3 \times 10^6$
Output	Single value: E_{pred} [GeV]
Loss function	MSE
Optimizer	Adam
Batch size	64
Training epochs	200
Final model	Average of 3 seeds (CNN-E-15layerSiWECAL)

“Classical” measurements

A · 3D shower-axis reconstruction

Energy-weighted TLS3D refit · PseudoMoliere(60%) radius · fitted axis oriented along +z



INPUT Accepted 3D energy deposits → energy-weighted core60 axis refit

OUTPUT One fitted line gives both direction and intersection with z = 0

B · Two observables

From one fitted shower axis

01 · DIRECTION

Reconstructed orientation

ūrec from the refitted axis

Event angular separation: Δψ

Reported resolution: projected_robust68 [°]

Projected angular residuals · final metric

02 · ENTRY POINT

Intersection of the fitted axis with z = 0

Δx = xrec − xtrue

Δy = yrec − ytrue

Spatial resolution [mm]:

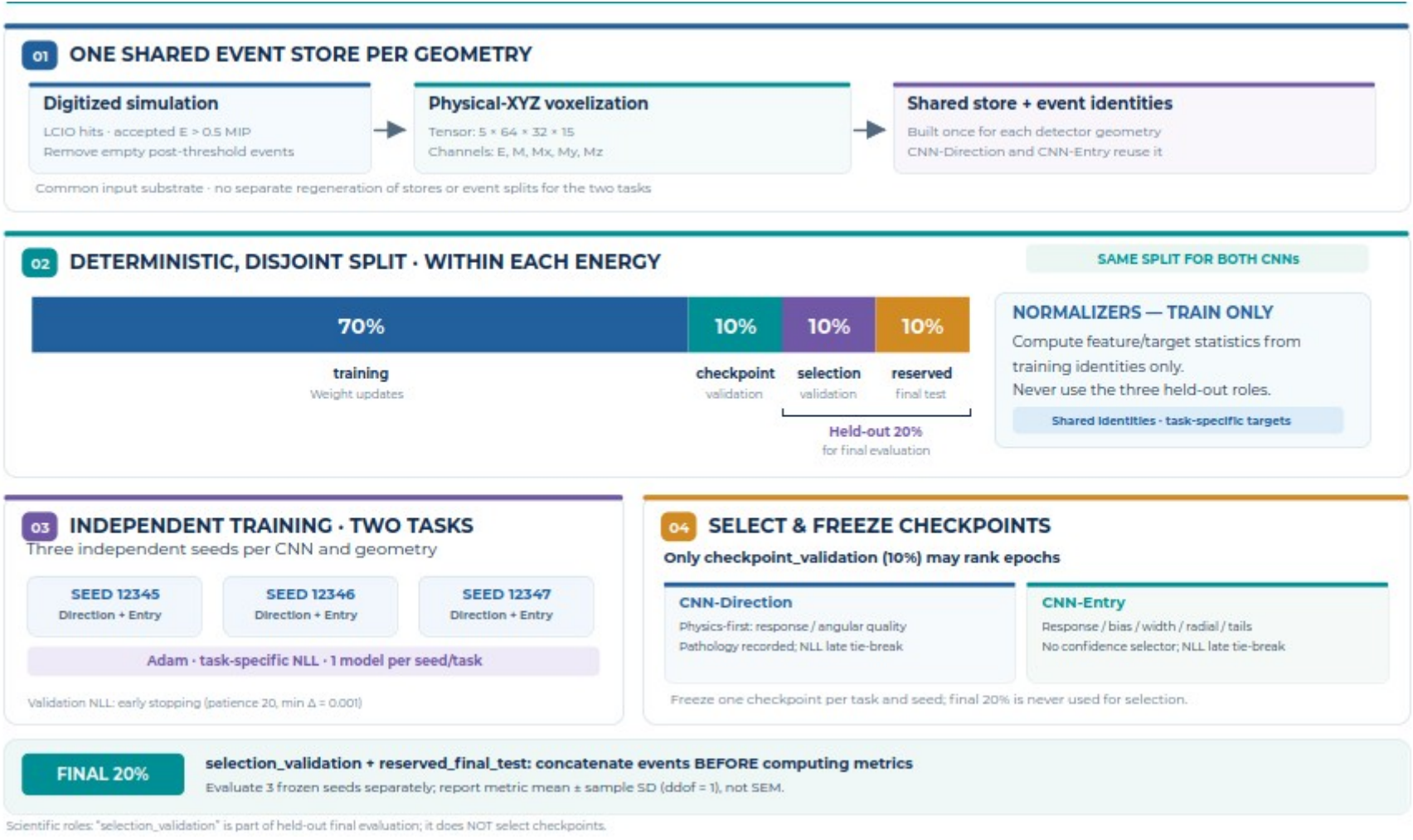
[robust68(Δx) + robust68(Δy)] / 2

NOT robust68 of the radial error Δr

Optimized from Phys.Rev.D 112 (2025) 11, 112014

CNN-Direction & CNN-entry

two independent 3D conv networks for complementary observables



CNN-Direction

01 3D ECAL input

Physical XYZ grid, 5 channels



$5 \times 64 \times 32 \times 15$

E · M · Mx · My · Mz

02 Shower features

3D CNN encoder + layer-wise geometry descriptors

Learned event features

Shared input convention; task-specific net

03 vMF prediction head

Predicts a mean direction and one concentration

Unit vector μ + κ

96,708 trainable parameters

04 Direction output

Reconstructed 3D axis

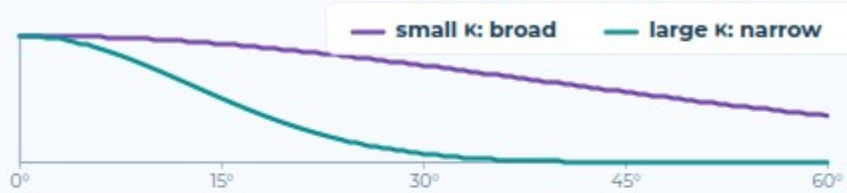
$\hat{u} = (u_x, u_y, u_z)$

κ: directional concentration score

What κ means in a vMF distribution

$p(u \mid \mu, \kappa) = C_3(\kappa) \cdot \exp(\kappa\, \mu \cdot u)$

κ controls concentration around the predicted unit direction μ.



Direction resolution: what is actually reported?

$\alpha_x = \text{atan2}(u_x, u_z)$ $\alpha_y = \text{atan2}(u_y, u_z)$

Signed residuals: $\Delta\alpha_x = \alpha_{x,\text{pred}} - \alpha_{x,\text{true}}$

$\Delta\alpha_y = \alpha_{y,\text{pred}} - \alpha_{y,\text{true}}$

$\text{robust68}(\Delta\alpha) = (q_{84} - q_{16}) / 2$

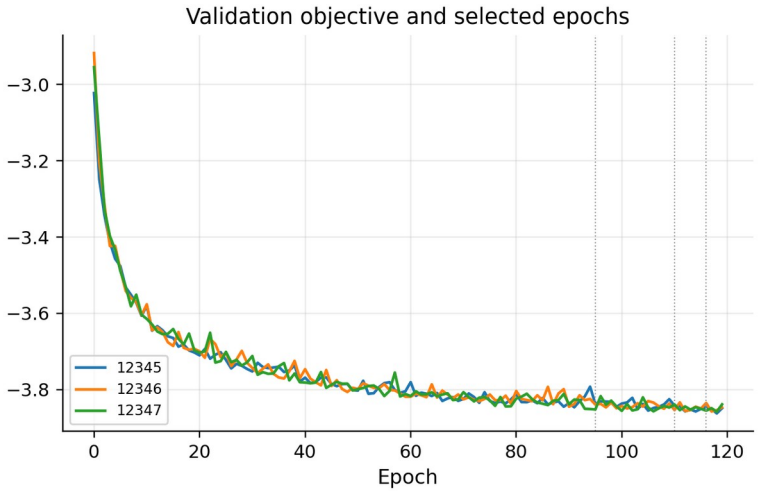
Primary: $[\text{robust68}(\Delta\alpha_x) + \text{robust68}(\Delta\alpha_y)] / 2$

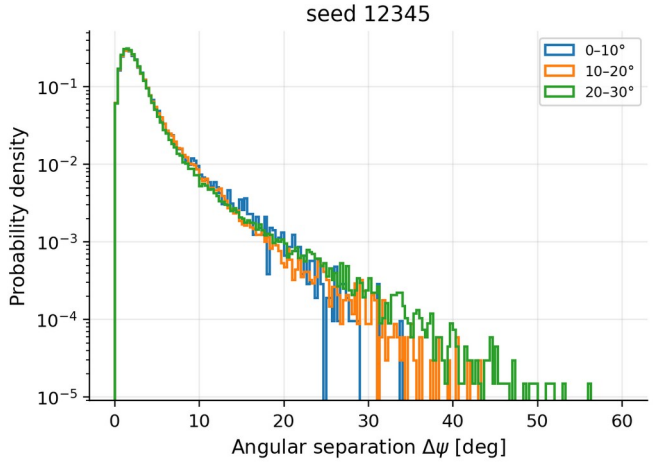
Supporting 3D error: $\Delta\psi = \arccos(\hat{u}_{\text{pred}} \cdot \hat{u}_{\text{true}})$

Training loss: $L = \text{mean}[-\log C_3(\kappa) - \kappa\, (\mu \cdot u_{\text{true}})]$; $C_3(\kappa) = \kappa / [4\pi \sinh(\kappa)]$.

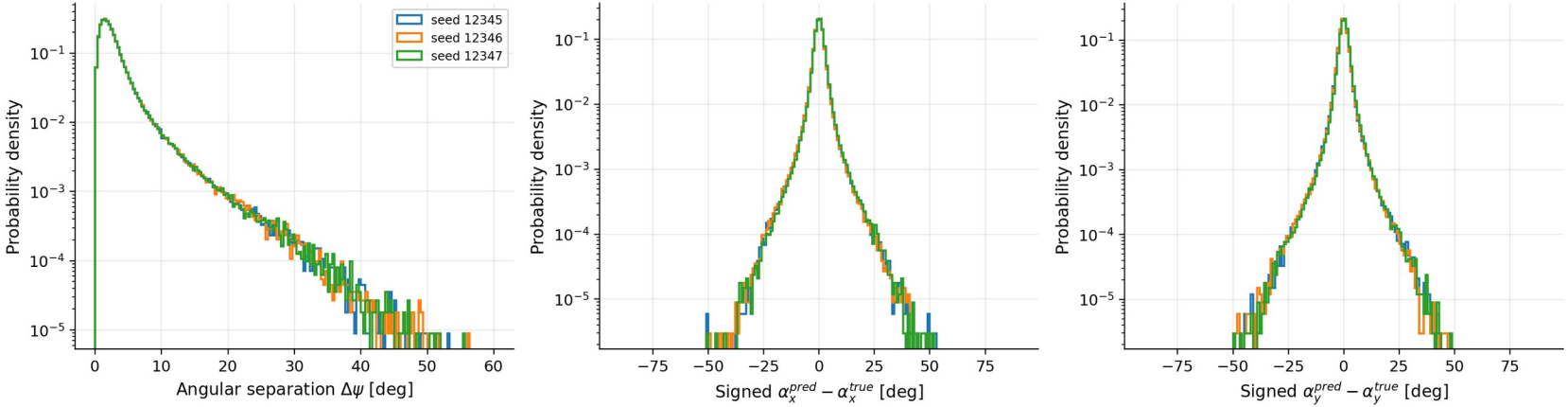
$\kappa = \text{softplus}(\rho) + 10^{-6}$; κ is not a calibrated angular error or the reported resolution.

CNN-Direction - results





18 X_0 Direction residual components — full saved range



CNN-Entry

01 3D ECAL input

Physical XYZ grid, 5 channels



$5 \times 64 \times 32 \times 15$

E · M · Mx · My · Mz

02 Shower geometry

3D CNN encoder + descriptors (including early reference)

Features for entry

Missing early anchor is flagged in input

03 Gaussian head

Two standardized means and a confidence parameter

μ_x,std · μ_y,std · ρ

96,579 trainable parameters

04 Entry output

Recover x and y at z = 0 plus a confidence score

(x, y) [mm] · σ_event

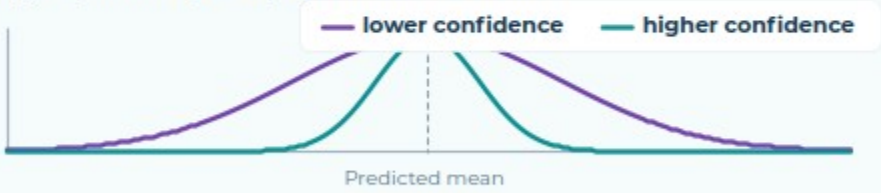
Not a calibrated per-event uncertainty

What "Gaussian confidence" means

$\tau = \text{softplus}(\rho) + 10^{-6}$

$\sigma_{event} = (5\ \text{mm}) / \sqrt{\tau}$

Larger τ | smaller σ_event | higher model confidence



Entry resolution and early-anchor information

Output coordinates: (xpred, ypred) at z = 0 mm

Signed residuals: Δx = xpred - xtrue

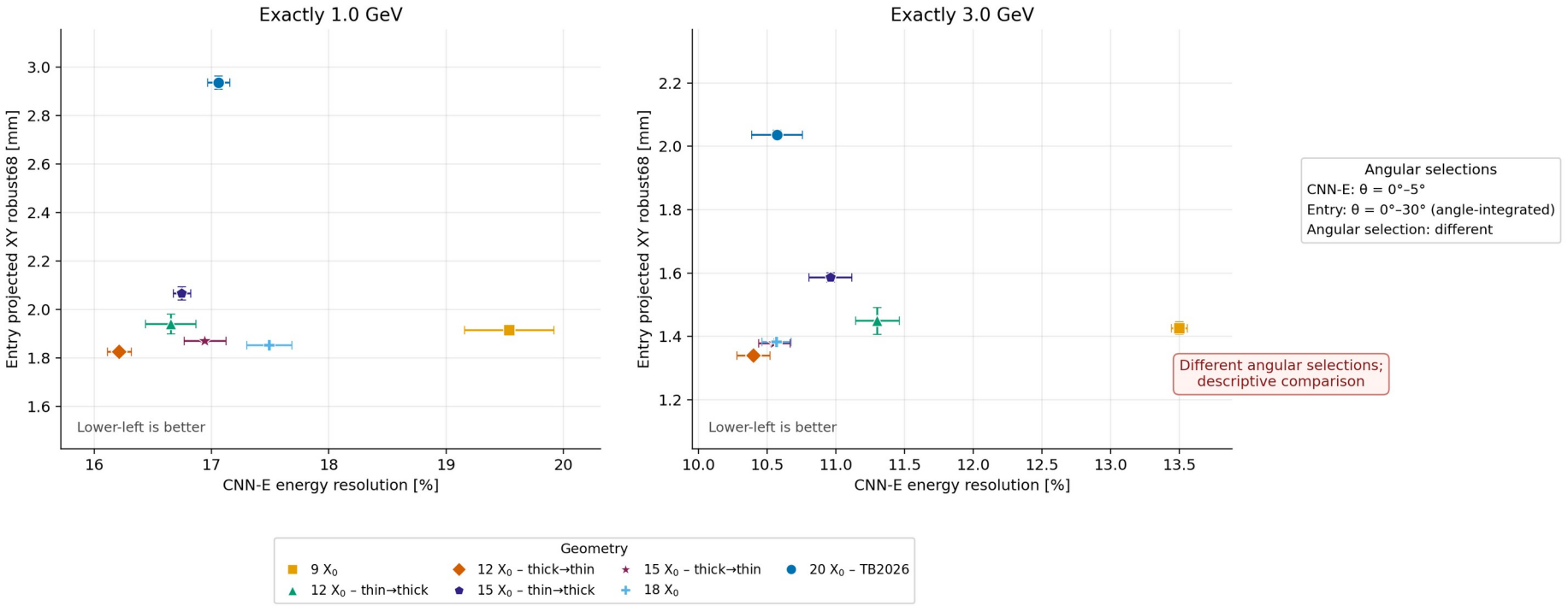
Δy = ypred - ytrue

Primary Entry resolution [mm]:
[robust68(Δx) + robust68(Δy)] / 2

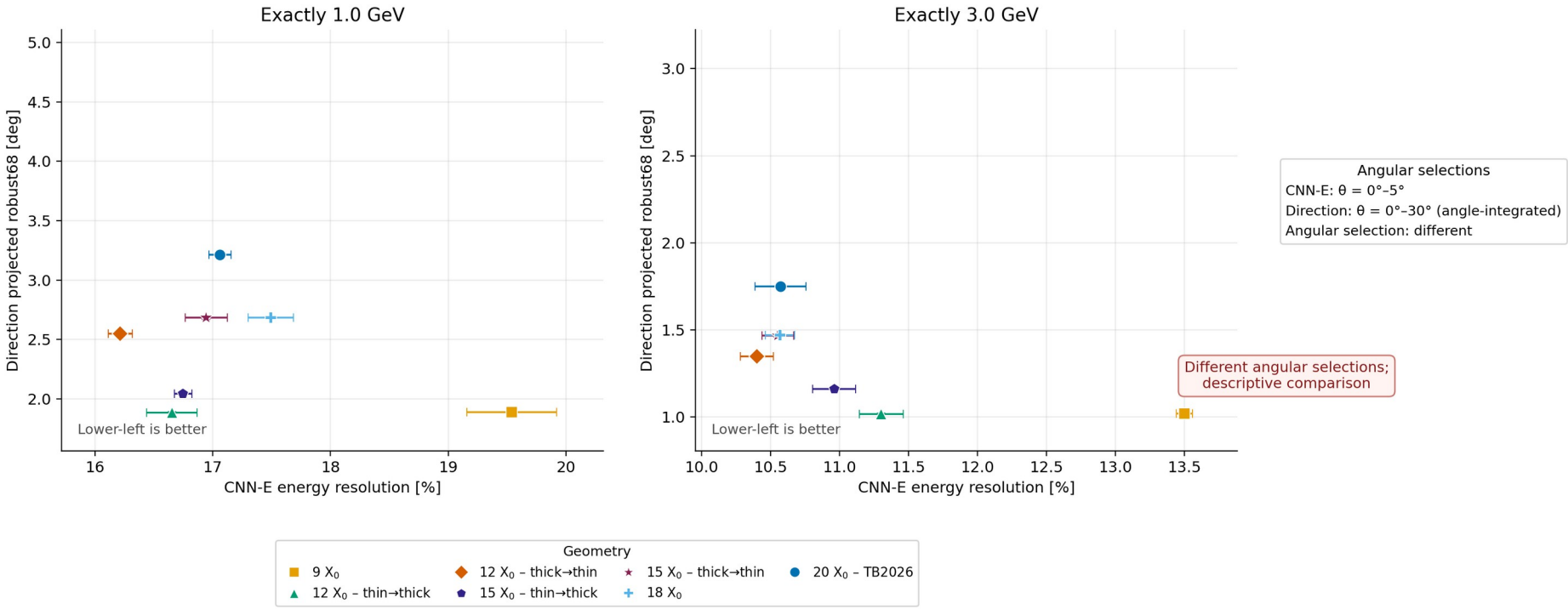
Early reference: E-weighted XY centroid in layers 0–2; if absent, use zero fallback + missingness indicator.

Training loss: $L = \text{mean}[\tau\, ||r - \mu||^2 / (2(5\ \text{mm})^2) - \log \tau]$ (constant omitted).
$\sigma_{event} = (5\ \text{mm})/\sqrt{\tau}$ is an event-confidence scale, not a calibrated error bar.

Results – Entrypoint vs Energy



Results – Direction vs Energy



CNN vs Classic

