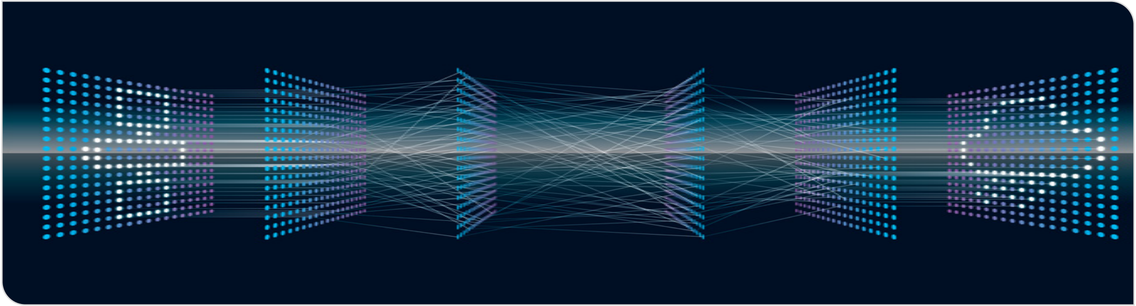
Karlsruhe Institute of Technology

Belle II

Institute of Experimental Particle Physics

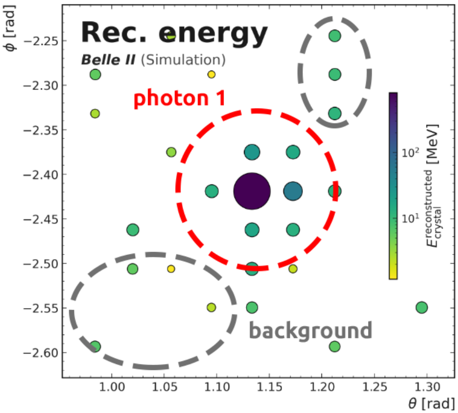
Clustering Energy Depositions in the ECL using Graph Neural Networks (GNNs)

Florian Wemmer, Torben Ferber | September 20, 2022



Setting and Objective

- 9×9 grid of crystals
- One photon cluster
- Nominal phase 3 beam background



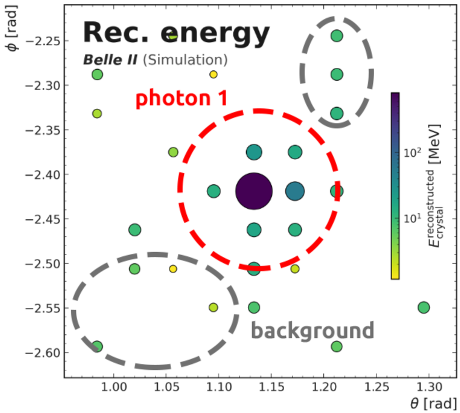
Setting and Objective

- 9×9 grid of crystals
- One photon cluster
- Nominal phase 3 beam background

- Cluster energy depositions
- Soft clustering including background

  $\Rightarrow$ Assign weights $w_i \in [0, 1]$ with $i \in \{\text{photon, bkg}\}$

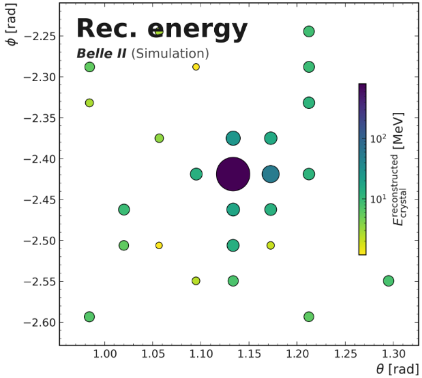
  $\Rightarrow \sum_i w_i = 1$ per crystal



Why (this) Graph Neural Network?

(and not a ‘regular’ Convolutional Neural Network)

- Additional valuable input features
- Learns representation space
- Can handle irregular detector geometry (endcaps)

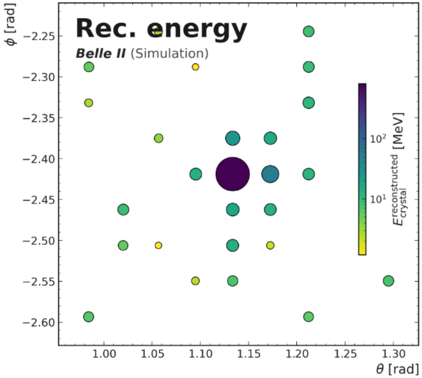


Why (this) Graph Neural Network?

(and not a ‘regular’ Convolutional Neural Network)

- Additional valuable input features
- Learns representation space
- Can handle irregular detector geometry (endcaps)
- Resources:
  - Only $\approx$ 16000 parameters and few computations
  - Small input: One hit $\hat{=}$ one node

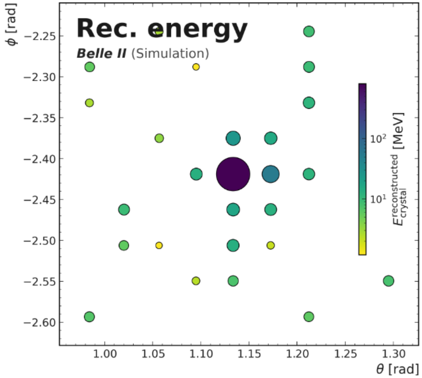
$\Rightarrow$ Potential real time application for L1 trigger



Approach

Conversion of an event to a graph

- Crystal hit in 9 x 9 view port becomes node
- Crystal measurements become node features
- No edges (yet)



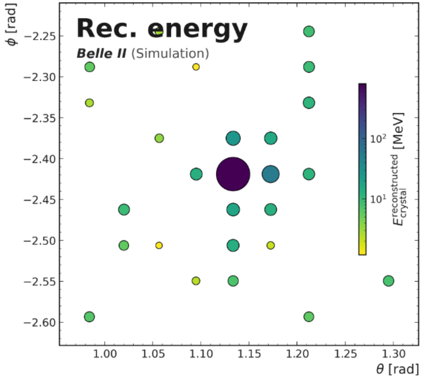
Approach

Conversion of an event to a graph

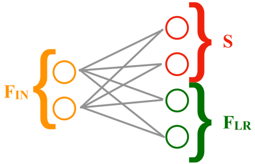
- Crystal hit in 9 x 9 view port becomes node
- Crystal measurements become node features
- No edges (yet)

Node features

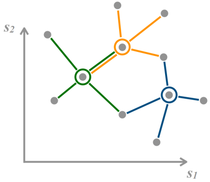
- Reconstructed Energy
- Reconstructed Time
- Pulse Shape Discrimination (PSD)
- Crystal coordinates (local and global)
- Crystal mass



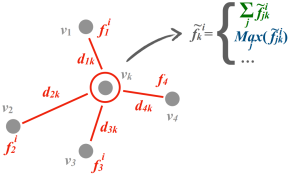
GravNet block



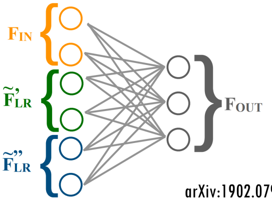
1. Learn representation and feature space



2. Connect *k* nearest neighbours



3. Message passing



arXiv:1902.07987

4. Concatenate messages

Model and Loss Function

Karlsruhe Institute of Technology

Model

- Stack three GravNet blocks
- Add batch normalization
- Fully connected layers into softmax

Model and Loss Function

Karlsruhe Institute of Technology

Model

- Stack three GravNet blocks
- Add batch normalization
- Fully connected layers into softmax

$$L_2 = \sum_{i,k} (p_{ik} - t_{ik})^2$$

- E_i reconstructed energy in node i
- t_{ik} true fraction of cluster k in node i
- p_{ik} pred. fraction of cluster k in node i

- i: number of nodes per event (varies)
- k $\in \{0, 1\}$: number of classes

Model and Loss Function

Karlsruhe Institute of Technology

Model

- Stack three GravNet blocks
- Add batch normalization
- Fully connected layers into softmax

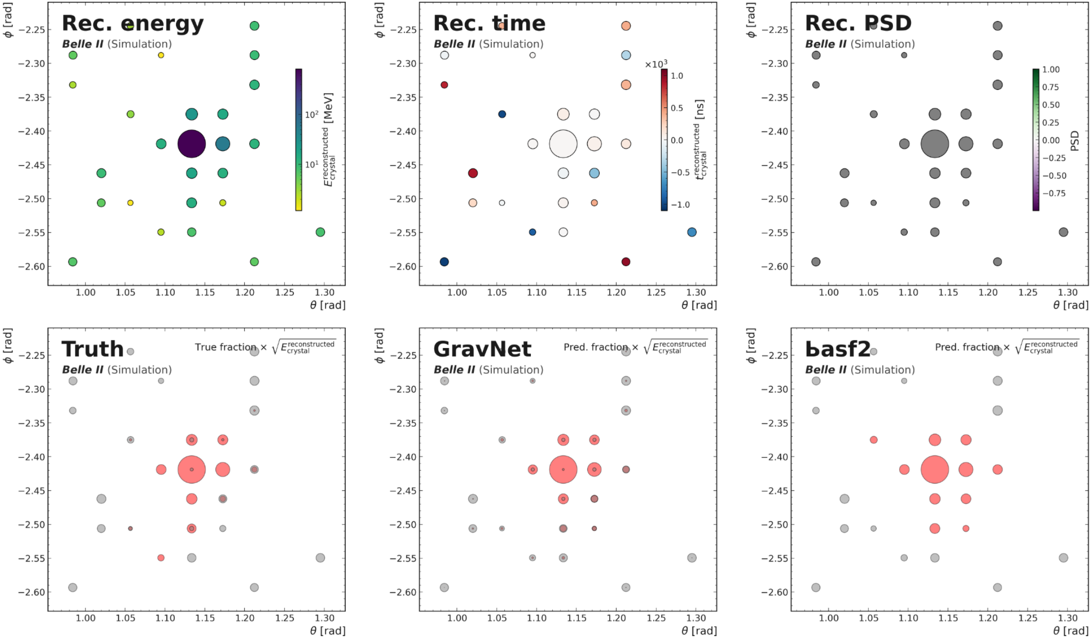
Machine Learning Settings

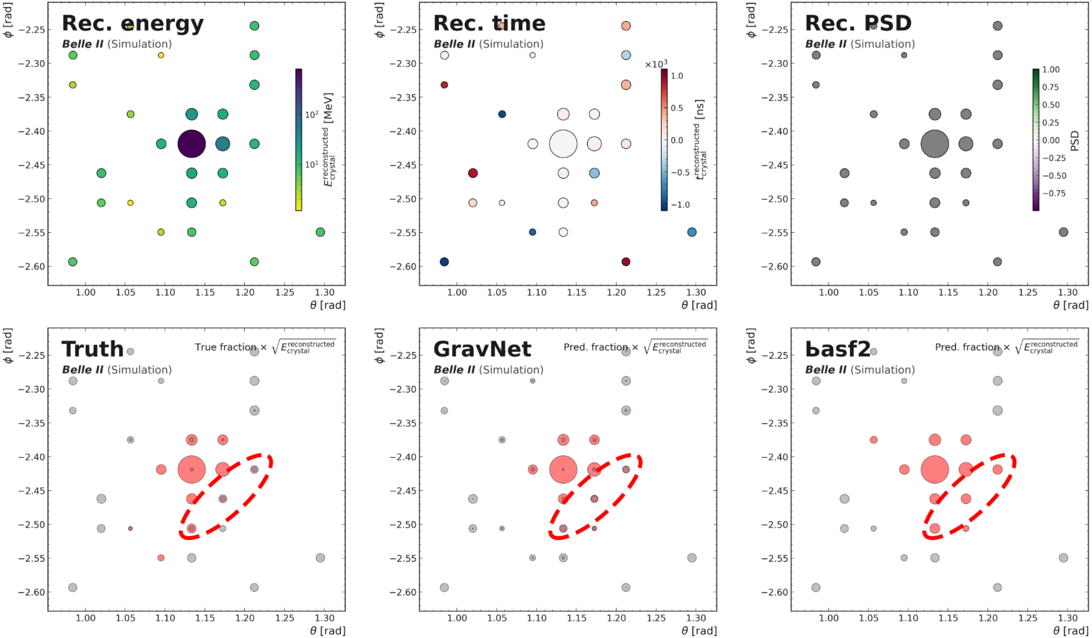
- Implemented in PyTorch Geometric
- 2 million MC events for training
- 200000 MC events for testing

$$L_2 = \sum_{i,k} (p_{ik} - t_{ik})^2$$

- E_i reconstructed energy in node i
- t_{ik} true fraction of cluster k in node i
- p_{ik} pred. fraction of cluster k in node i

- i: number of nodes per event (varies)
- k $\in \{0, 1\}$: number of classes


Rec. energy
Belle II (Simulation)
$E_{crystal}^{reconstructed}$ [MeV]
Rec. time
Belle II (Simulation)
$t_{crystal}^{reconstructed}$ [ns]
Rec. PSD
Belle II (Simulation)
PSD
Truth
Belle II (Simulation)
True fraction × $\sqrt{E_{crystal}^{reconstructed}}$
GravNet
Belle II (Simulation)
Pred. fraction × $\sqrt{E_{crystal}^{reconstructed}}$
basf2
Belle II (Simulation)
Pred. fraction × $\sqrt{E_{crystal}^{reconstructed}}$
ϕ [rad]
θ [rad]



Rec. energy
Belle II (Simulation)
$E_{crystal}^{reconstructed}$ [MeV]
Rec. time
Belle II (Simulation)
$\times 10^3$
$t_{crystal}^{reconstructed}$ [ns]
Rec. PSD
Belle II (Simulation)
PSD
Truth
Belle II (Simulation)
True fraction $\times \sqrt{E_{crystal}^{reconstructed}}$
GravNet
Belle II (Simulation)
Pred. fraction $\times \sqrt{E_{crystal}^{reconstructed}}$
basf2
Belle II (Simulation)
Pred. fraction $\times \sqrt{E_{crystal}^{reconstructed}}$
ϕ [rad]
θ [rad]

Generated Photon Energy Resolution

$$\frac{E_{\mathrm{pred}} - E_{\mathrm{gen}}}{E_{\mathrm{gen}}}$$

$$E_{\mathrm{pred}} = \begin{cases} \sum_i E_i p_i & \text{GravNet} \\ \mathrm{ClusterE} & \text{basf2} \end{cases}$$

$$E_{\mathrm{gen}} = E_\gamma$$

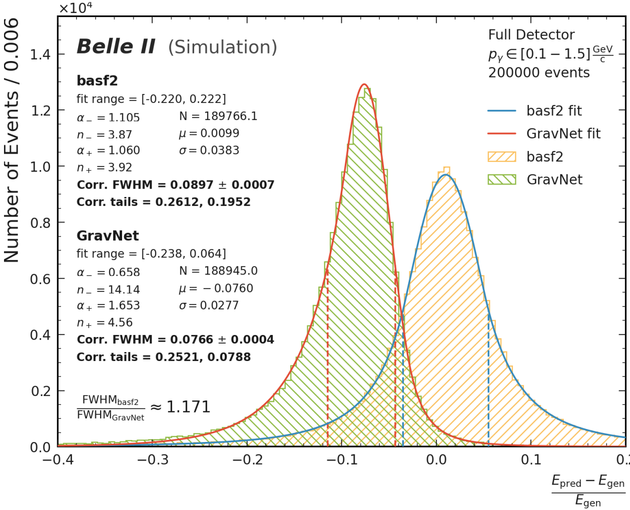
Generated Photon Energy Resolution

$$\frac{E_{\text{pred}} - E_{\text{gen}}}{E_{\text{gen}}}$$

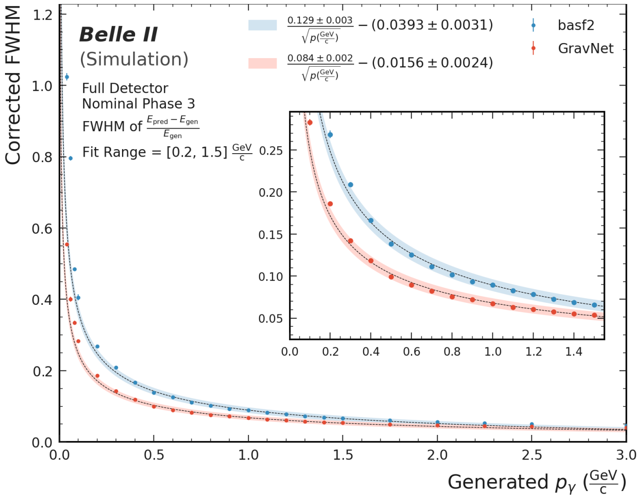
$$E_{\text{pred}} = \begin{cases} \sum_i E_i p_i & \text{GravNet} \\ \text{ClusterE} & \text{basf2} \end{cases}$$

$$E_{\text{gen}} = E_\gamma$$

$\Rightarrow$ 17.1 % improvement



Resolution Energy Dependence



π^0 Invariant Mass Resolution

- Shot π^0 particle gun with $p_{\pi^0} \in [0.2 - 2.0]\ \frac{\mathrm{GeV}}{\mathrm{c}}$
- Two separate photon signatures in detector

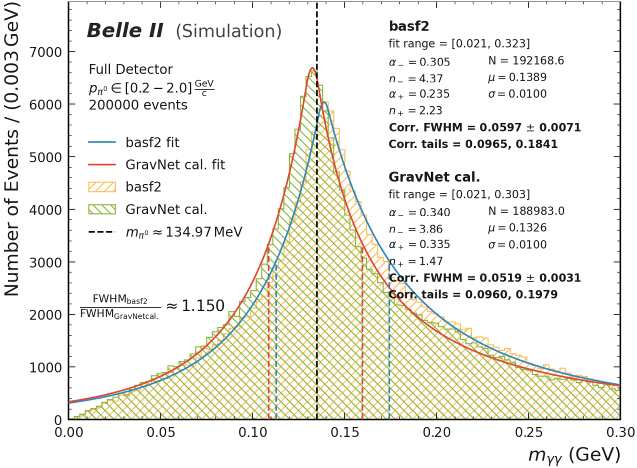
π^0 Invariant Mass Resolution

- Shot π^0 particle gun with $p_{\pi^0} \in [0.2 - 2.0]\ \frac{\mathrm{GeV}}{\mathrm{c}}$
- Two separate photon signatures in detector
- "Calibrate" GravNet by shifting generic photon peak
- Reconstruction of π^0 mass from $m_{\gamma\gamma}$

π^0 Invariant Mass Resolution

- Shot π^0 particle gun with $p_{\pi^0} \in [0.2 - 2.0]\ \frac{\mathrm{GeV}}{\mathrm{c}}$
- Two separate photon signatures in detector
- "Calibrate" GravNet by shifting generic photon peak
- Reconstruction of π^0 mass from $m_{\gamma\gamma}$

$\Rightarrow$ 15 % improvement



Summary and Outlook

Summary

- Few parameters and computations
- Well-suited for soft clustering with nominal phase 3 background
- Significant improvements to photon and π^0 resolution

Summary and Outlook

Summary

- Few parameters and computations
- Well-suited for soft clustering with nominal phase 3 background
- Significant improvements to photon and π^0 resolution

Further Work

- Early phase 3 background analysis
- Network for overlapping photons
- π^0 rec. from overlapping photons

Outlook

- Technical paper
- Evaluation on data
- Investigate feasibilty for basf2

Deposited Photon Energy Resolution

$$\frac{E_{\mathrm{pred}} - E_{\mathrm{dep}}}{E_{\mathrm{dep}}}$$

$$E_{\mathrm{pred}} = \begin{cases} \sum_i E_i p_i & \text{GravNet} \\ \mathrm{rawEnergy} & \text{basf2} \end{cases}$$

$$E_{\mathrm{dep}} = \sum_i E_i t_i$$

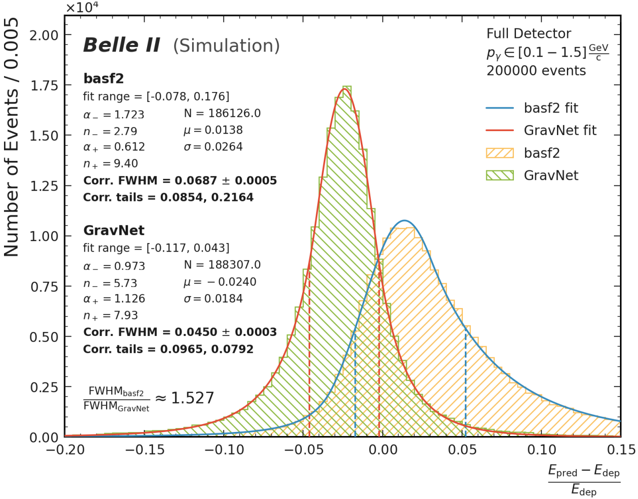
Deposited Photon Energy Resolution

$$\frac{E_{\text{pred}} - E_{\text{dep}}}{E_{\text{dep}}}$$

$$E_{\text{pred}} = \begin{cases} \sum_i E_i p_i & \text{GravNet} \\ \text{rawEnergy} & \text{basf2} \end{cases}$$

$$E_{\text{dep}} = \sum_i E_i t_i$$

$\Rightarrow$ 52.7 % improvement

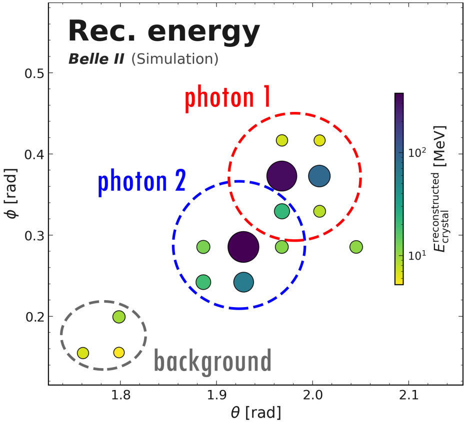


Setting and Objective Overlap

- Two clusters with overlap in 5×5 region
- 9×9 grid of crystals
- Beam background

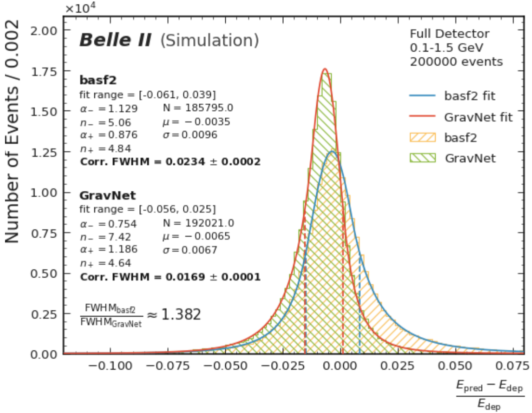
- Cluster overlapping energy depositions
- Soft clustering including background

  $\Rightarrow$ Assign weights $w_i \in [0, 1]$ with $i \in \{\text{c0, c1, bkg}\}$
  $\Rightarrow \sum_i w_i = 1$ per crystal



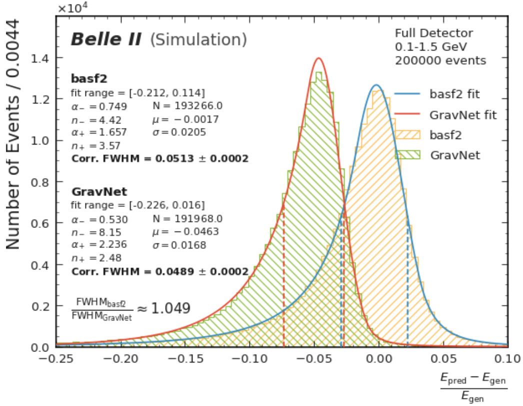
Early Phase 3

Deposited energy



⇒ 38.2 % improvement

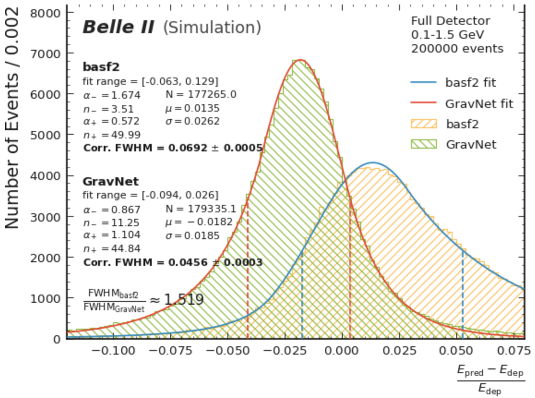
Generated energy



⇒ 4.9 % improvement

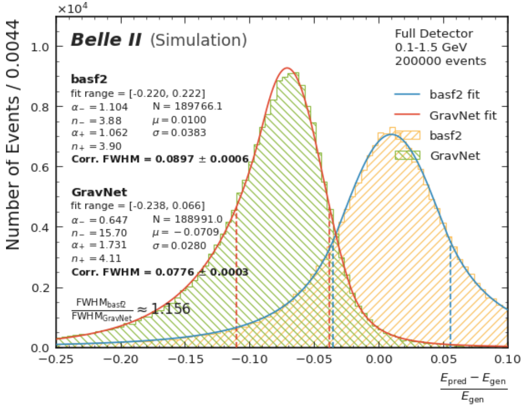
Nominal phase 3 Background

Deposited energy



$\Rightarrow$ 51.9 % improvement

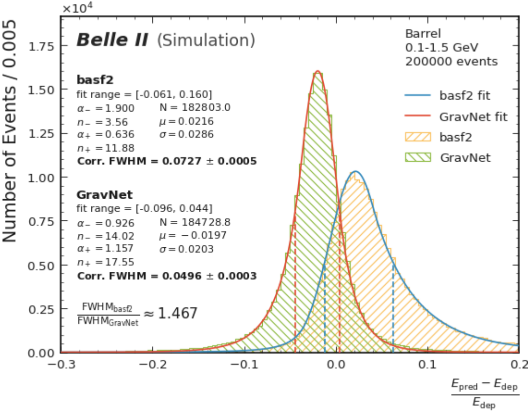
Generated energy



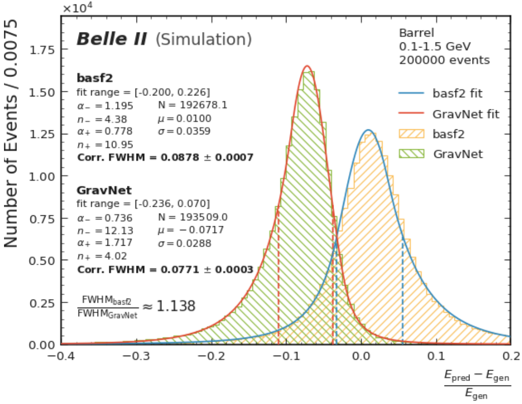
$\Rightarrow$ 15.6 % improvement

Barrel

Deposited Energy

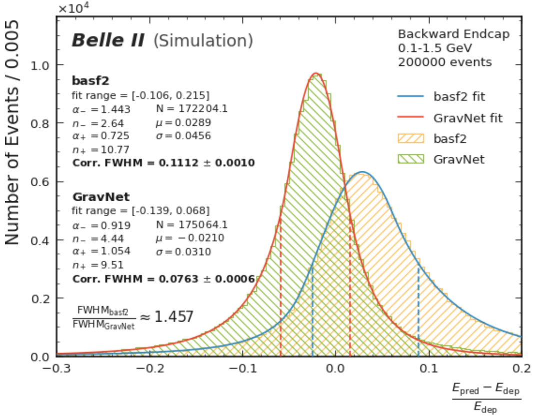


Generated Energy

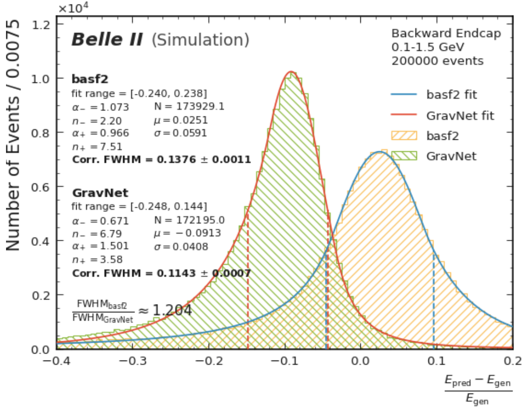


Backward Endcap

Deposited Energy

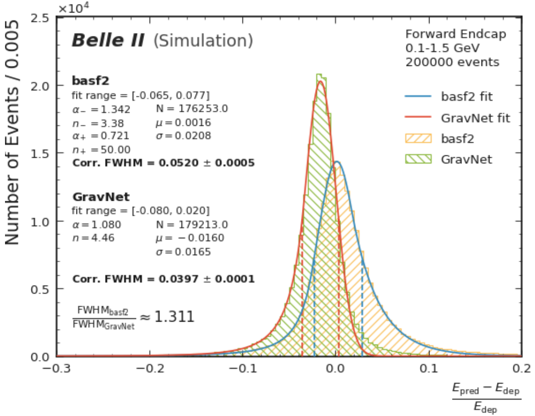


Generated Energy

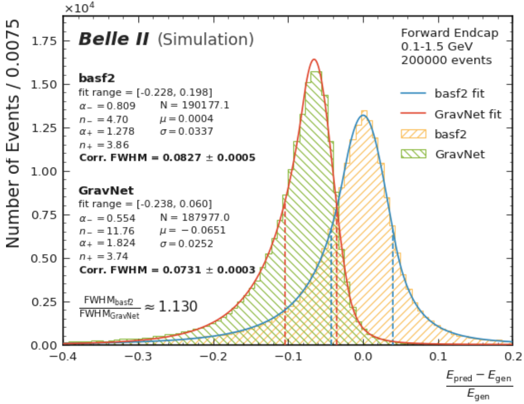


Forward Endcap

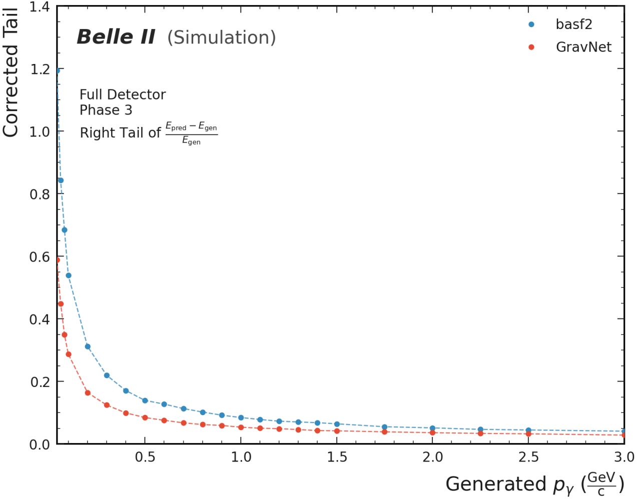
Deposited Energy

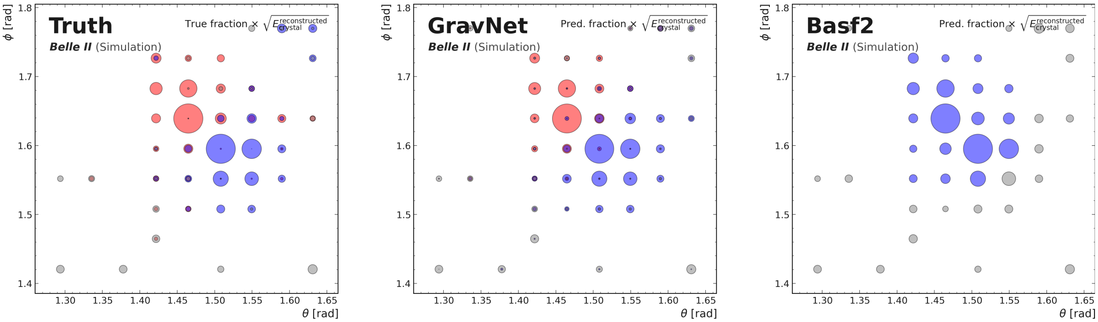


Generated Energy



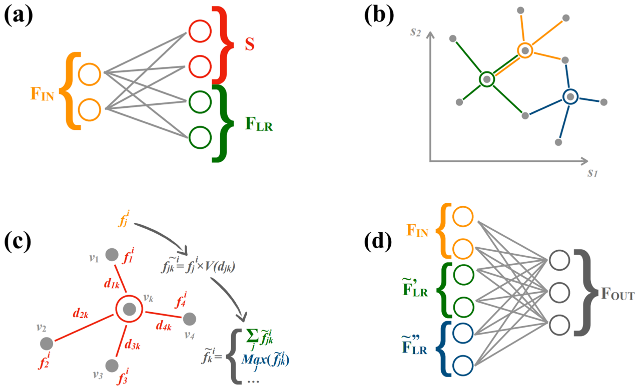
Tails Energy Dependence




Truth
Belle II (Simulation)
True fraction × $\sqrt{E_{\text{crystal}}^{\text{reconstructed}}}$
GravNet
Belle II (Simulation)
Pred. fraction × $\sqrt{E_{\text{crystal}}^{\text{reconstructed}}}$
Basf2
Belle II (Simulation)
Pred. fraction × $\sqrt{E_{\text{crystal}}^{\text{reconstructed}}}$
ϕ [rad]
θ [rad]
1.4
1.5
1.6
1.7
1.8
1.30
1.35
1.40
1.45
1.50
1.55
1.60
1.65

Full Model

Karlsruhe Institute of Technology



(a)

- Three fully connected layers
- Representation space dim = 3
- Learned node features = 16

(b)

- Nearest neighbours = 12

(c)

- Mean and maximum aggregation

(d)

- One fully connected layer

Input / Output Details

Features

- Rec. energy, rec. time, rec. PSD
- Crystal weight
- Global coordinates
- Local coordinates in 9 x 9 region

Softmax (sum fractions = 1)

- Fraction particle 1
- Fraction particle 2
- Fraction background

Full Machine Learning Settings

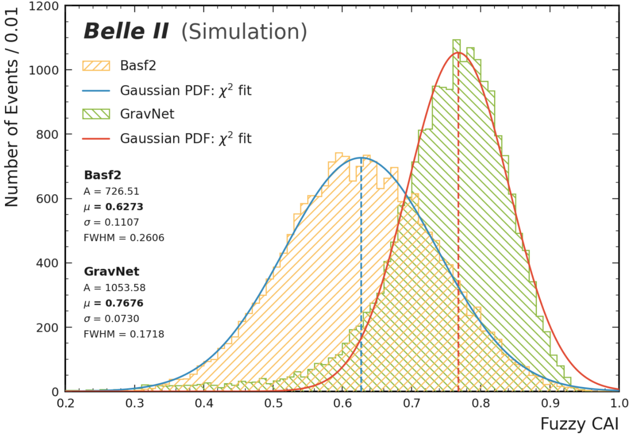
Training

- Random initialization
- Adam Optimizer
- Learning rate: 0.005
- Decaying learning rate on plateau

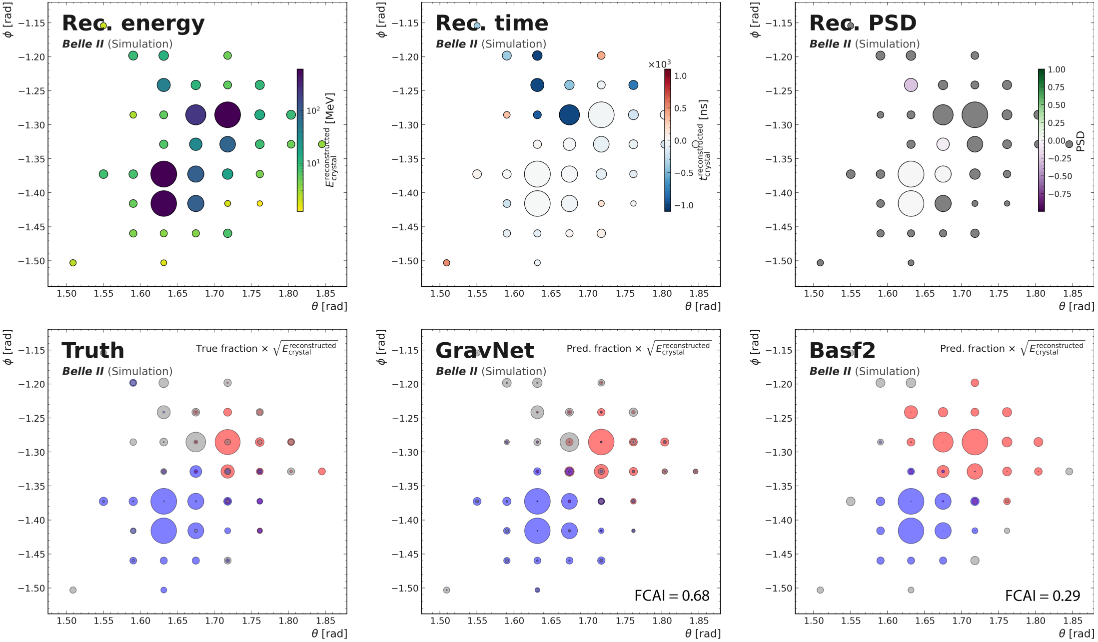
More details

- Batchsize: 512
- Number of epochs: $\approx$ 50 to convergence
- Batch normalization after each GravNet Layer
- (Mostly) Elu activation

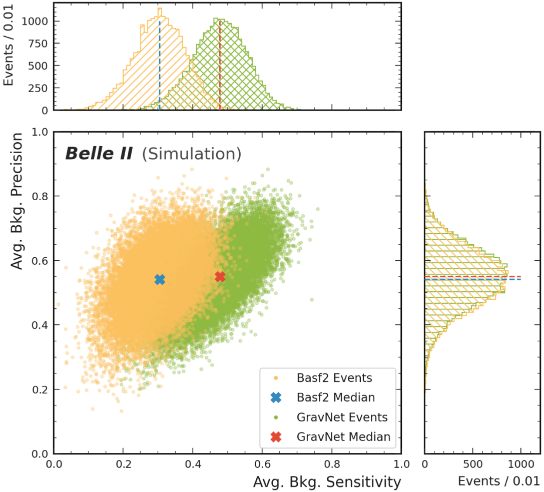
Clustering Metrics



- Evaluate quality of soft-clustering
- Use overlap between different classes at each data point as basis for pairwise comparison
- doi: 10.1609/aaai.v31i1.10905.


Rec. energy
Belle II (Simulation)
Rec. time
Belle II (Simulation)
Rec. PSD
Belle II (Simulation)
Truth
Belle II (Simulation)
True fraction × $\sqrt{E_{\text{crystal}}^{\text{reconstructed}}}$
GravNet
Belle II (Simulation)
Pred. fraction × $\sqrt{E_{\text{crystal}}^{\text{reconstructed}}}$
FCAI = 0.68
Basf2
Belle II (Simulation)
Pred. fraction × $\sqrt{E_{\text{crystal}}^{\text{reconstructed}}}$
FCAI = 0.29
ϕ [rad]
θ [rad]
$E_{\text{crystal}}^{\text{reconstructed}}$ [MeV]
$t_{\text{crystal}}^{\text{reconstructed}}$ [ns]
PSD

Clustering Metrics



- Sensitivity:

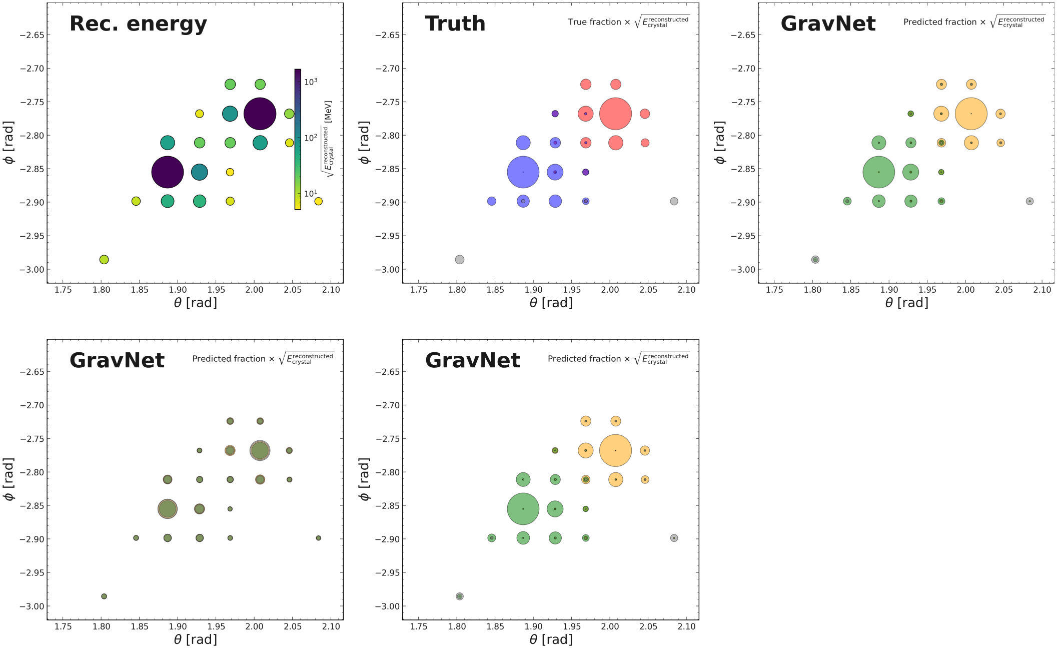
$$\frac{1}{n}\sum\frac{\text{(identified and correct background)}}{\text{(total true background)}}$$

- Precision:

$$\frac{1}{n}\sum\frac{\text{(identified and correct background)}}{\text{(total identified background)}}$$

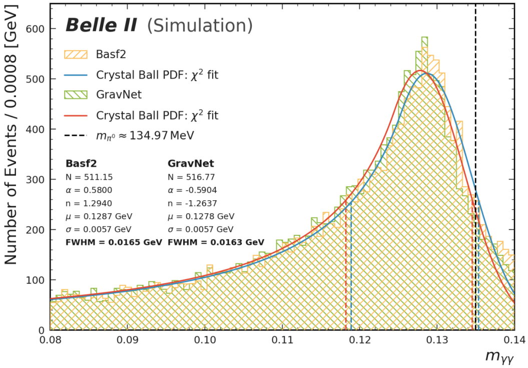
Data Selection

- Two separated local maxima
- At least 10 MeV measured energy
- At least 80% true physics deposition in local maxima
- Overlap in 5×5 area around local maxima

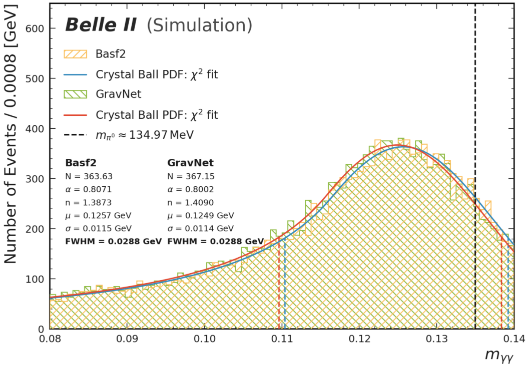

Rec. energy
Truth
True fraction × $\sqrt{E^{\text{reconstructed}}_{\text{crystal}}}$
GravNet
Predicted fraction × $\sqrt{E^{\text{reconstructed}}_{\text{crystal}}}$
$\sqrt{E^{\text{reconstructed}}_{\text{crystal}}}$ [MeV]
ϕ [rad]
θ [rad]

π^0 Invariant Mass Resolution

Reconstructed energy + MC position

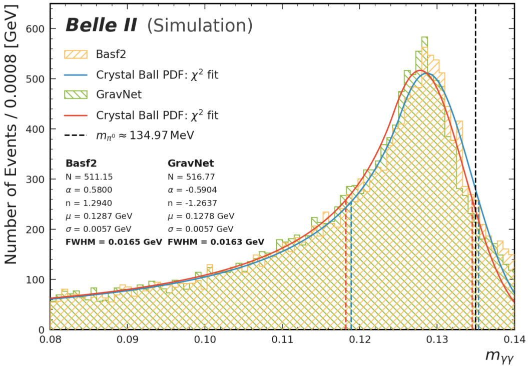


Reconstructed energy + reconstructed position

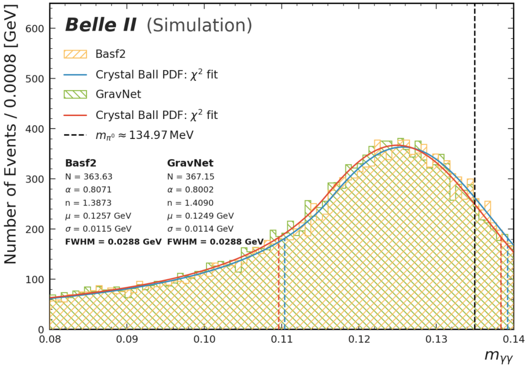


π^0 Invariant Mass Resolution

Reconstructed energy + MC position



Reconstructed energy + reconstructed position



$\Rightarrow$ Significant improvement in photon resolution does not transfer to invariant mass resolution:

$$\sigma^2_{m^2} \approx m^4_{\pi^0}\left(\frac{\sigma^2_{E\gamma_1}}{E^2_{\gamma_1}} + \frac{\sigma^2_{E\gamma_2}}{E^2_{\gamma_2}} + \frac{4\sigma^2_{\alpha}}{\alpha^2}\right)$$