Prediction of electromagnetic fraction in a hadronic shower using deep neural network

Marina Chadeeva (LPI), Sergey Korpachev (LPI, MIPT) and Anna Staritsyna (MIPT)

1. Motivation and goal
2. Simulations and preprocessing
3. DNN model and training
4. Preliminary results

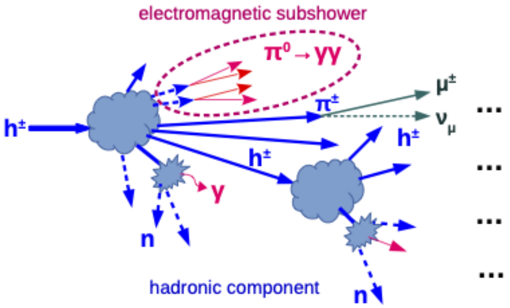
Motivation and goal

Percent-level precision of jet energy measurements is important for future collider experiments aimed at searches for New Physics beyond SM

- **modern trend**: particle flow (PF) reconstruction based on highly granular calorimeter systems and high-precision tracking
- **bottleneck for jet resolution**: contribution from neutral hadrons

Hadronic energy resolution

- complex structure of hadronic showers results in significant fluctuations of energy deposition due to fluctuations of electromagnetic fraction (with dominated contribution from π^0s)
- possible solutions for improvement:
  - hardware compensation (compensating calorimeters, e.g. ZEUS)
  - software compensation algorithms, e.g. in *[JINST 7 (2012) P09017]*



In this talk: software compensation using predictions of electromagnetic fraction by DNN

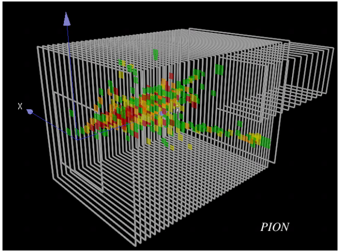
Machine-learning-based approach for software compensation

Software compensation (SC)

means event-by-event energy correction during reconstruction

Previous and ongoing studies:

- parameterised local (hit-based) and global (observables-based), up to 15% improvement in stochastic term for single pions for both data and MC
- Graph neural network trained on MC samples - considers hadronic shower as image (work in progress by CALICE)



Simulated 80-GeV pion shower in highly granular calorimeter

A novel approach proposed in this study

DNN training and optimisation on MC

- inputs: global calorimetric observables
- target: true electromagnetic fraction $f_{\mathrm{em}} = \sum_{over \pi^0 s} E_{\pi^0} / E_{\mathrm{initial}}$
- supervised learning

Inference (event-by-event)

- calculation of calorimetric observables
- prediction of f_{em} from DNN trained
- correction of event energy

Simulations, detector model and event selection

Simulations with Geant4 version 10.3

- **single negative pions @ 10–80 GeV**, about 500 kevt / energy point (raw samples centrally produced by CALICE DESY group)
- model of **highly granular analog hadron calorimeter AHCAL** (scintillator-SiPM, steel absorber, long. depth: ~4.3 nucl. int. length, transverse size: 72×72 cm^2)
- Physics lists: **FTFP_BERT_HP** and **QGSP_BERT_HP**
  HP (High Precision) – precise neutron models and cross sections below 20 MeV

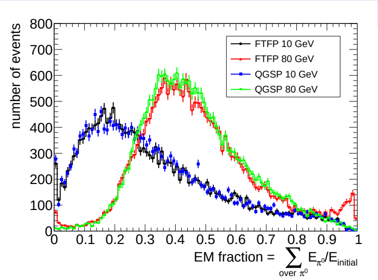
Calibration and digitisation

- MIP calibration with MC muons
- light collection and photon detection by SiPM, pixelisation and saturation are emulated in digitisation
- digitisation tuned with MC-to-Data comparisons for muons and electrons
- EM calibration factor, $C_{\mathrm{em}} = 0.0233$ GeV/MIP

Reconstruction chain and event selection

- cell signals above 0.5 MIP threshold — hits
- shower start finder algorithm tuned on MC
  for analysis: only events with found shower start at 3–6 AHCAL layers
- conversion factor to hadron energy scale: $C_{\mathrm{had}} = 1.2$

Event energy reconstruction



Standard hadron energy reconstruction

$$E^{\text{em}}_{\text{reco}} = C_{\text{em}} \cdot \sum_{i=1}^{N_{\text{hits}}} e_i , \qquad E^{\text{had}}_{\text{reco}} = C_{\text{had}} \cdot E^{\text{em}}_{\text{reco}}$$

where e_i - hit energy in MIP, N_{hits} - total number of hits

Energy reconstruction with correction from known electromagnetic fraction

$$E_{\text{measured}} = \langle em \rangle \cdot E_{\text{initial}} \cdot f_{\text{em}} + \langle h \rangle \cdot E_{\text{initial}} \cdot (1 - f_{\text{em}})$$

where $\langle em \rangle$ and $\langle h \rangle$ are mean reconstruction efficiencies for electromagnetic and hadronic subshowers and are assumed to be energy independent: $\langle em \rangle = 1$ for electromagnetic calibration and $\langle h \rangle$ can be estimated empirically for particular calorimeter ($\langle h \rangle \approx$ 0.7–0.75 for AHCAL)

In MC, f_{em} can be extracted as a sum of π^0 energies in event, and one can correct for f_{em} fluctuations:

$$E^{\text{cor}}_{\text{reco}} = \frac{E^{\text{em}}_{\text{reco}}}{f_{\text{em}} + \langle h \rangle \cdot (1 - f_{\text{em}})},$$

where $\langle h \rangle$ accounts also for C_{had}

f_{em} is not accessible in data but can be predicted by DNN from calorimetric observables!

Calorimetric observables for DNN input — 29 variables

Counting observables

- **Number of isolated hits**, $N_{\rm iso}$ [isolation – 0 neighbours in a cube of 3×3×3 cells around the hit]
- **Number of track hits**, $N_{\rm trk}$ [defined as having two in-line neighbours and MIP-like deposition]

Amplitude observables

- **Mean shower hit energy**, $\langle e_{\rm hit}\rangle$
- **Shower radius** $R_{\rm sh} = \frac{\sum_{i=1}^{N_{\rm sh}} e_i \cdot r_i}{\sum_{i=1}^{N_{\rm sh}} e_i}$, $r_i = \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2}$ - hit rad. dist. from sh. axis (x_0,y_0)
- **Longitudinal shower centre of gravity** $Z_{\rm CoG} = \frac{\sum_{i=1}^{N_{\rm sh}} e_i \cdot (z_i - z_{\rm start})}{\sum_{i=1}^{N_{\rm sh}} e_i}$,

  e_i - energy of hit with coordinates x_i, y_i, z_i; $N_{\rm sh}$ - number of shower hits
  z_i - hit longitudinal coordinate, $z_{\rm start}$ - longitudinal coordinate of shower start

Additional "ring" observables (integrated over longitudinal depth)

3-cm wide rings around shower axis, consistent with cell transverse size; 12 rings in total

- **number of isolated hits in a ring**, $N_{\rm iso}^{\rm ring}$
- **energy fraction in a ring**, $f_{\rm enr}^{\rm ring}$, w.r.t total energy sum

Deep Neural Network model, architecture and training

Target: electromagnetic fraction in event as a ratio of a sum of π^0 energies to the initial energy

Regression model with MSE loss function

$$L_i = (Y_i^{\text{predicted}} - Y_i^{\text{true}})^2 \quad \text{Loss} = \frac{1}{N} \cdot \sum_{i=1}^{N} W_i \cdot L_i$$

N - number of events for training

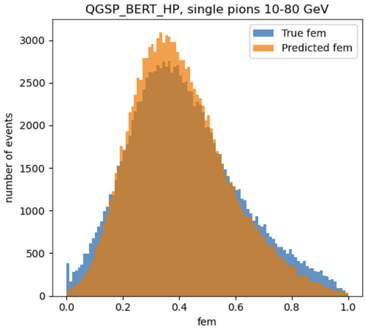
$\mathrm{W_i}$ – event weights from density-based weighting

Training, validation and test subsamples

- mixed sample for 10-80 GeV
- after selections: ∼130 kevt/ phys.list
- train:valid:test ≈ 60%:20%:20%

Example DNN architecture and output

- number of layers: 1 input, 6 hidden, 1 output
- number of neurons: 29/80/40/20/16/32/64/1
- activation function: ReLU $= \max(0, x)$ for hidden, linear $(\mathrm{f(y) = y})$ for output
- bias neurons and weighted loss
- number of training epochs: about 15
- **supervised learning**

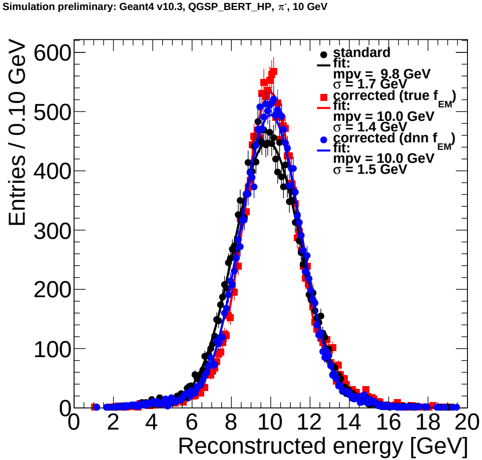


Tools: TensorFlow library, Keras framework, scikit-learn package

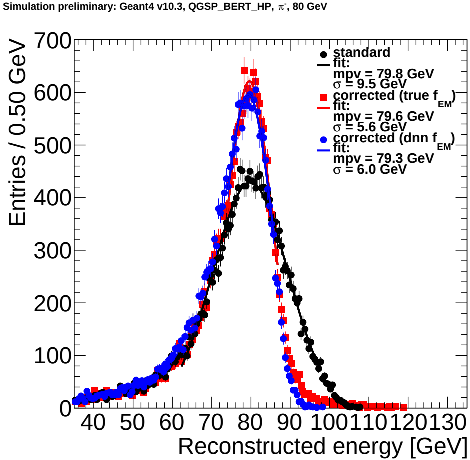
Example of corrected energy distributions for QGSP_BERT_HP

DNN trained on mixed sample of QGSP_BERT_HP and applied to QGSP_BERT_HP

10 GeV



80 GeV

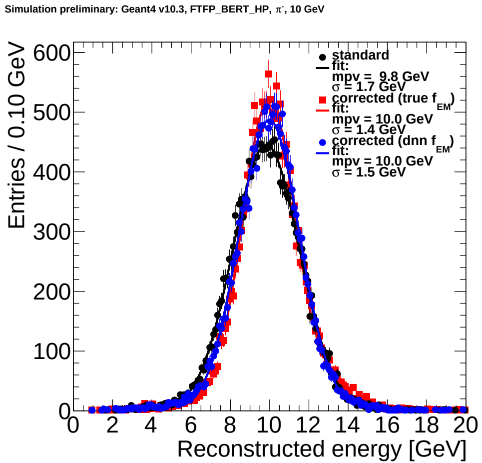


Comparable improvement by true and predicted f_{em} with $\langle h \rangle = 0.72$ and $\langle h \rangle = 0.75$.

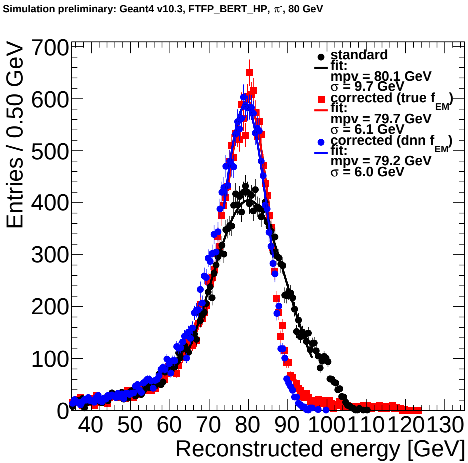
Example of corrected energy distributions for FTFP_BERT_HP

DNN trained on mixed sample of QGSP_BERT_HP and applied to FTFP_BERT_HP

10 GeV



80 GeV

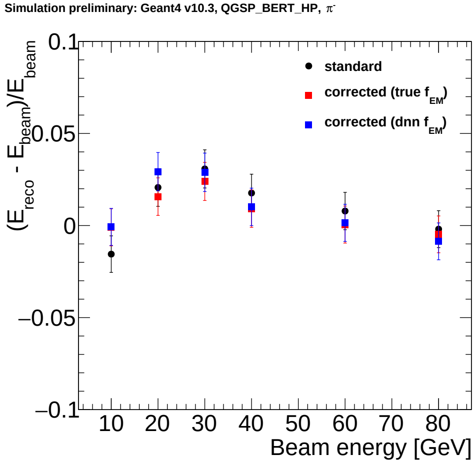


Comparable improvement by true and predicted f_{em} with $\langle h \rangle = 0.72$ and $\langle h \rangle = 0.75$.

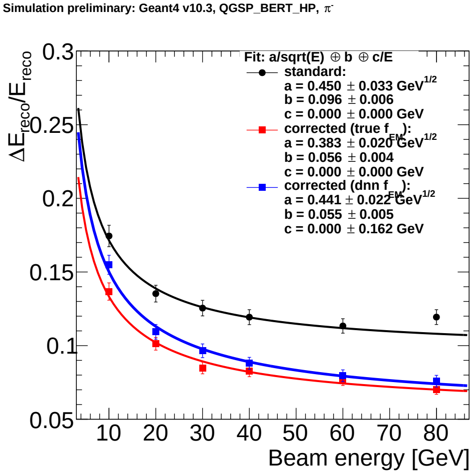
Energy correction using f_{em}

DNN trained on mixed sample QGSP_BERT_HP and applied to QGSP_BERT_HP

Linearity



Relative resolution

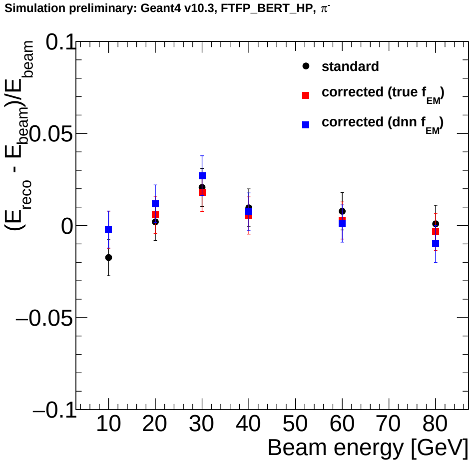


Better linearity and resolution with true f_{em}. Significant improvement in rel. resol. with predicted f_{em}.

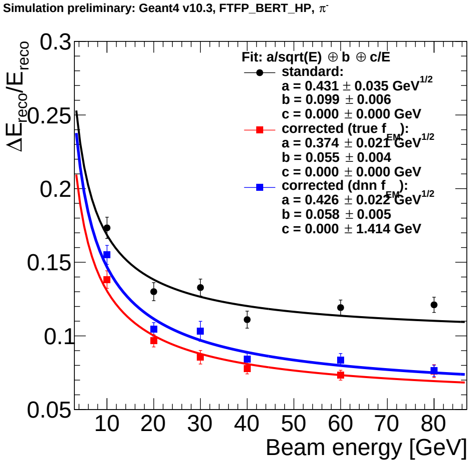
Energy correction using true and predicted f_{em}

DNN trained on mixed sample QGSP_BERT_HP and applied to FTFP_BERT_HP

Linearity



Relative resolution



Better linearity and resolution with true f_{em}. Significant improvement in rel. resol. with predicted f_{em}.

Summary

Prediction of em fraction in hadronic showers

tested on simulations with QGSP_BERT_HP and FTFP_BERT_HP physics lists from Geant4 v10.3

- **Technique:** regression model in Deep Neural Network trained using supervised learning
- **Inputs:** 29 calorimetric observables from a highly granular calorimeter
- **Target:** electromagnetic fraction in a hadronic shower
- **Preliminary results for single pion-induced showers**
  - DNN trained on mixed samples (10-80 GeV)
  - reasonable performance in event-by-event prediction of electromagnetic fraction

Application

- predicted em fraction was used to correct energy event-by-event
- **comparable correction effects are achieved with true and predicted em fraction**

Plans

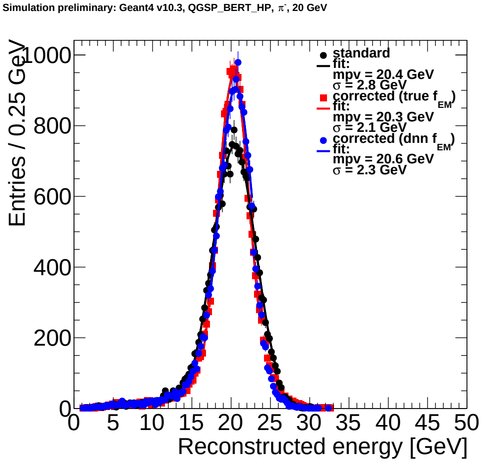
- optimise DNN hyperparameters to improve performance
- **apply trained model to data**

Backup slides

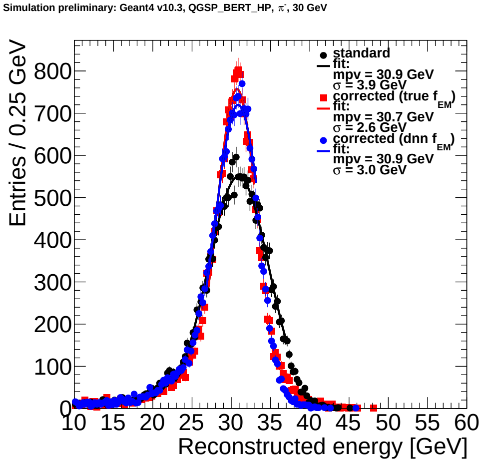
Example of corrected energy distributions for QGSP_BERT_HP

DNN trained on mixed sample of QGSP_BERT_HP and applied to QGSP_BERT_HP

20 GeV



30 GeV

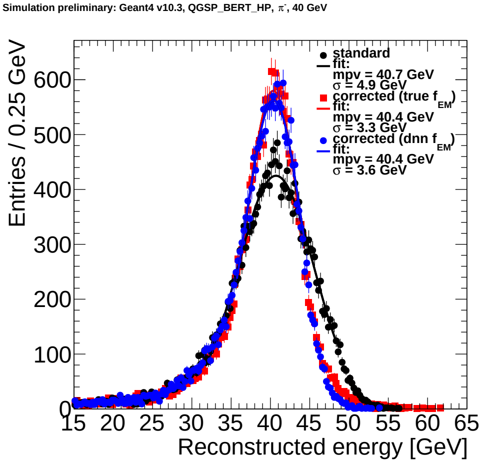


Comparable improvement by true and predicted $f_{\rm em}$ with $\langle h \rangle = 0.72$ and $\langle h \rangle = 0.75$.

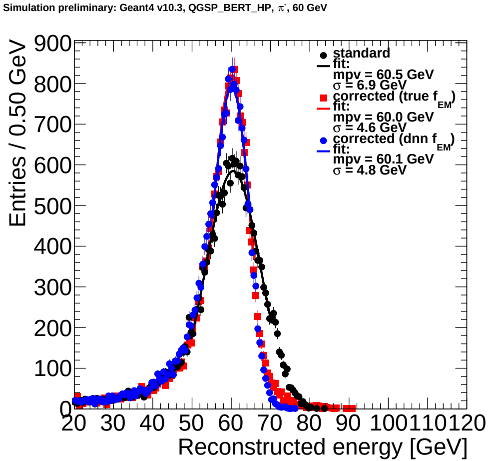
Example of corrected energy distributions for QGSP_BERT_HP

DNN trained on mixed sample of QGSP_BERT_HP and applied to QGSP_BERT_HP

40 GeV



60 GeV

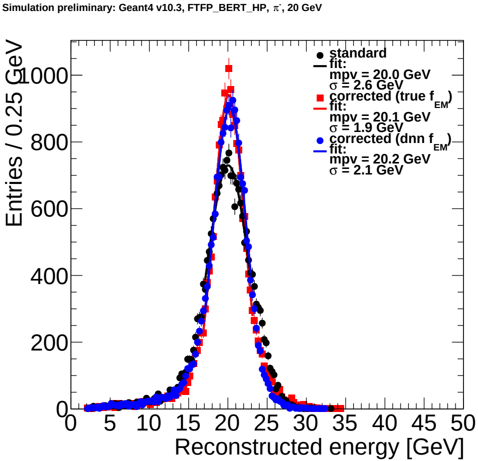


Comparable improvement by true and predicted f_{em} with $\langle h \rangle = 0.72$ and $\langle h \rangle = 0.75$.

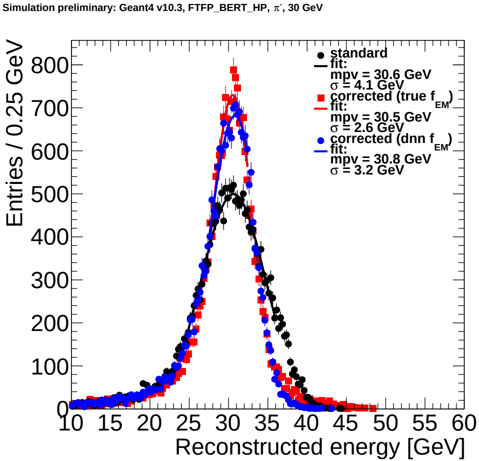
Example of corrected energy distributions for FTFP_BERT_HP

DNN trained on mixed sample of QGSP_BERT_HP and applied to FTFP_BERT_HP

20 GeV



30 GeV

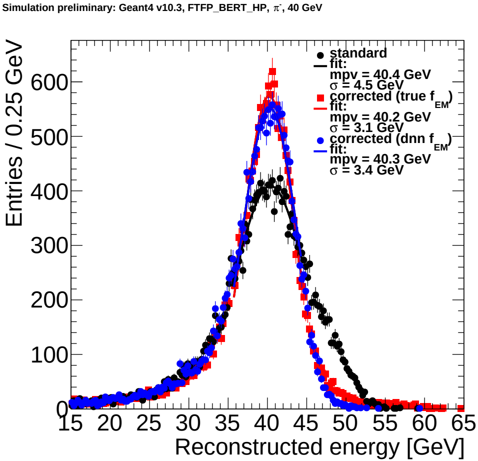


Comparable improvement by true and predicted f_{em} with $\langle h \rangle = 0.72$ and $\langle h \rangle = 0.75$.

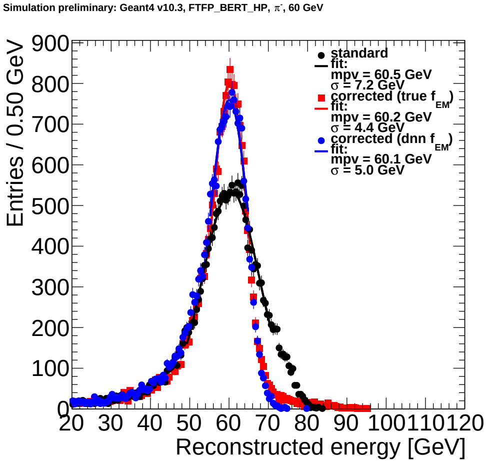
Example of corrected energy distributions for FTFP_BERT_HP

DNN trained on mixed sample of QGSP_BERT_HP and applied to FTFP_BERT_HP

40 GeV



60 GeV



Comparable improvement by true and predicted f_{em} with $\langle h \rangle = 0.72$ and $\langle h \rangle = 0.75$.