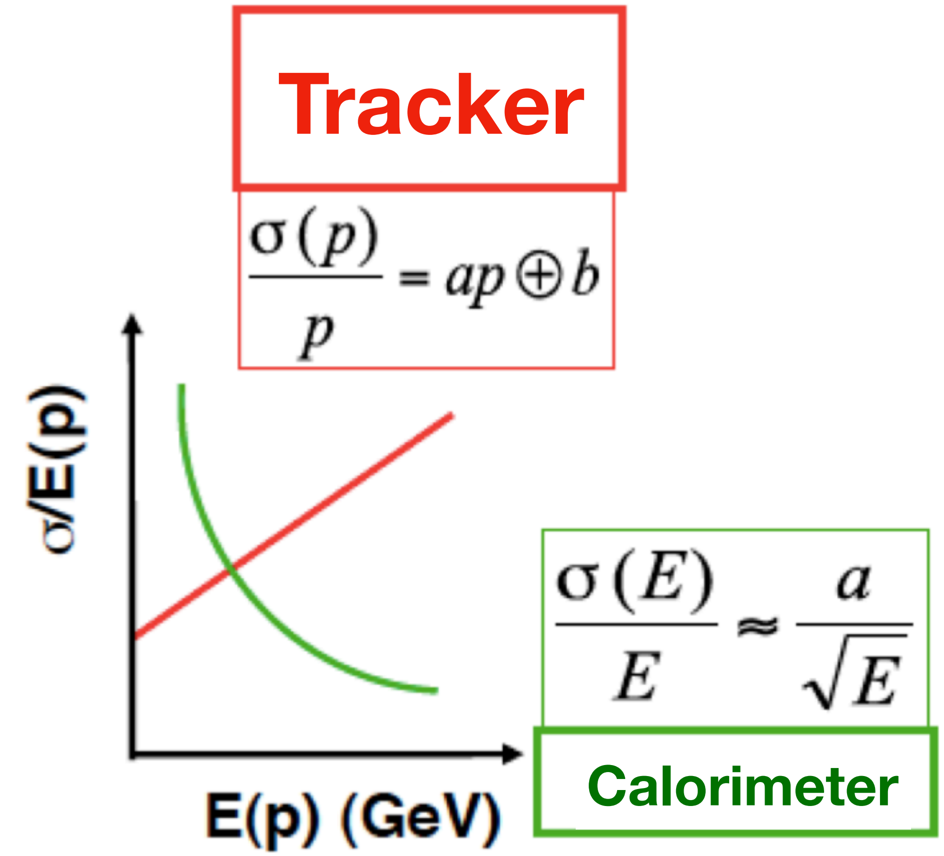
Muon Energy Regression from Radiative Losses in a Granular Calorimeter

ICNFP 2021

Jan Kieseler, Giles Strong, Filippo Chiandotto,
Tommaso Dorigo, Lukas Layer

Motivation

- **Muons** are important as a **clean probe for new physics searches & measurements**
- **Excellent energy resolution** will be crucial for the search of new physics phenomena at **future high-energy colliders**
- **Energy estimates** determined by **curvature** of the muon trajectory in **trackers**
- **But**: radius of curvature increases with energy → **tracker gives poor resolution at high energy** since relative uncertainty scales linearly with E
- Relative **uncertainty of calorimetric** measurements **decreases** with √ E if the particles are completely absorbed

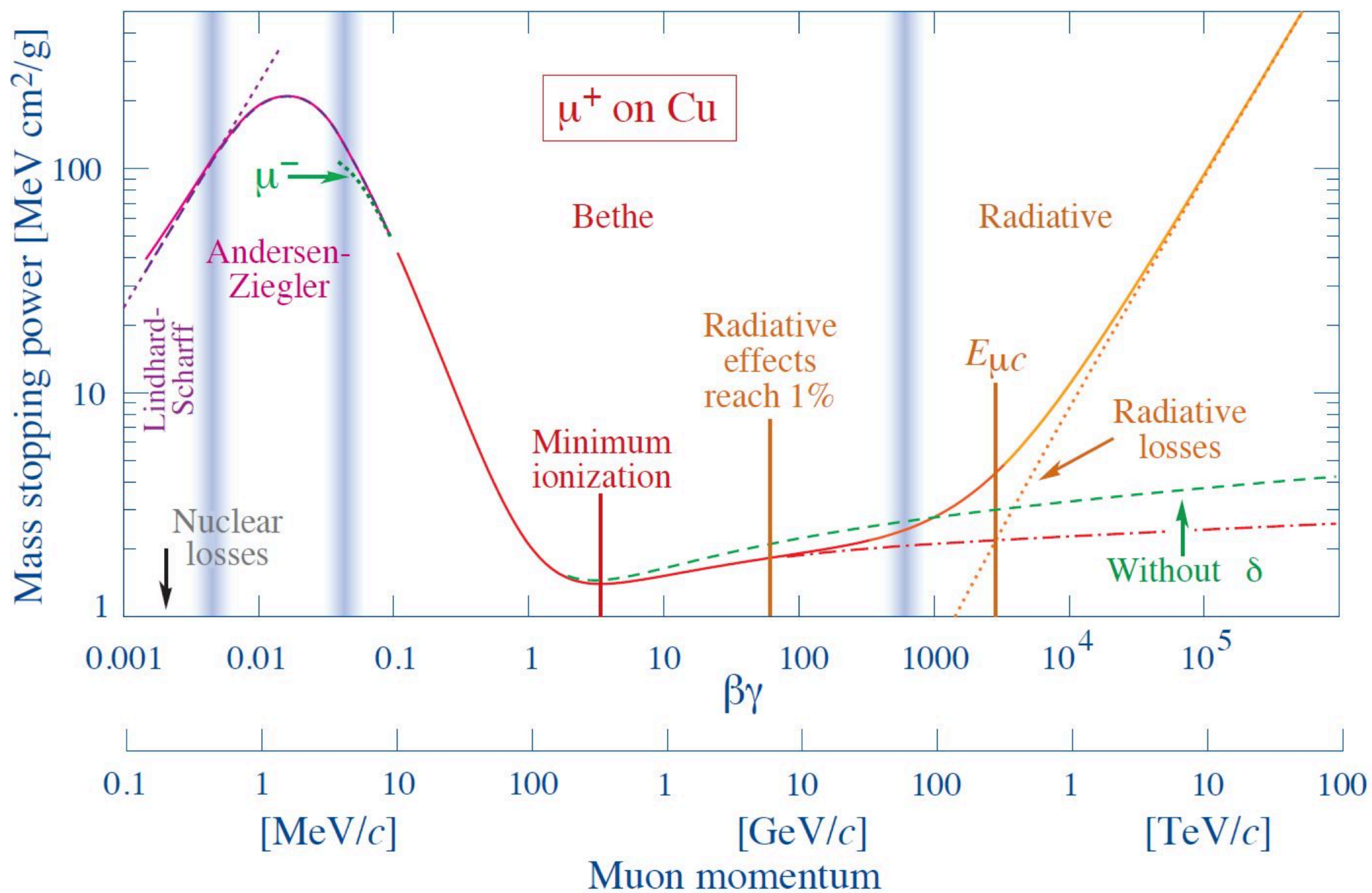


→ **Can we improve muon energy measurements with calorimeters?**

Energy loss of muons

- At **high energy**, muons do **not** behave as **minimum ionising** particles
- Rise of **radiative energy loss** above roughly **100GeV**
- But **radiation still low → radiative** losses **so far not exploited** to estimate energy in collider detectors
- Recording the **low-energy photons** in a **granular calorimeter** can provide a **complementary measurement** of the muon energy

Source: PDG Phys. Rev. D 98, 030001



Idea: regress muon energy from energy deposits in calorimeter using modern ML techniques

Simulated data

Detector

- Homogeneous **lead tungstate** cuboid **calorimeter**
- **Dimensions**: 2032x120x120 mm (z,x,y) subdivided into 39.6x3.75x3.75 mm cells
- **Number of cells**: 50x32x32 (z,x,y) = 51.200 cells
- Calorimeter embedded in a uniform ATLAS like **2-Tesla magnetic field**

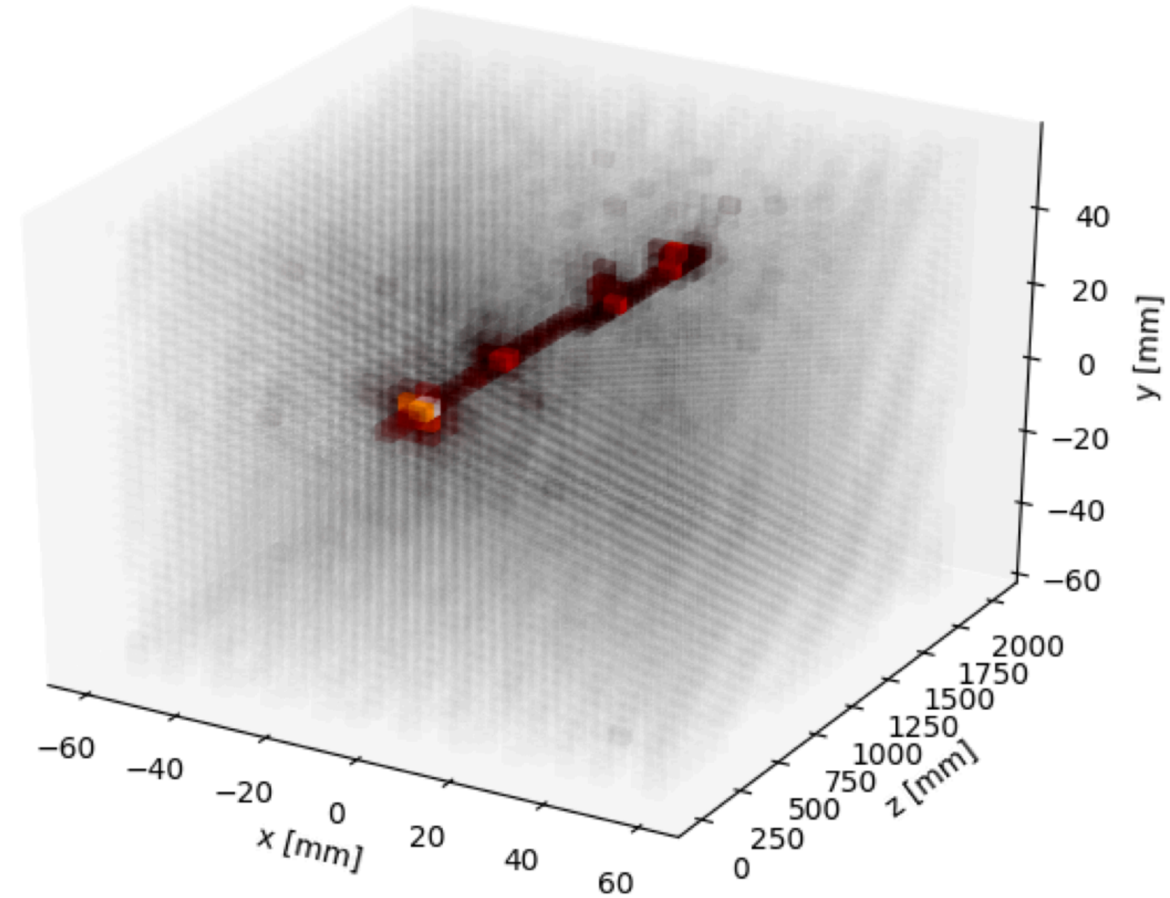
Muons

- **Interaction** of the muons with the detector material is **simulated using GEANT4**
- Muons are **unpolarized** with **uniform energy** in z-direction between **50GeV and 8TeV**
- **~850k** muons are generated for **training and validation**
- ~**430.000** muons are generated for **testing** with discrete energy in 400 GeV steps in 100-4100 GeV

Muon regression strategy

Muon in the detector

- Energy deposits concentrated along flight path
- Large number of low-energy deposits
- Relatively sparse hits

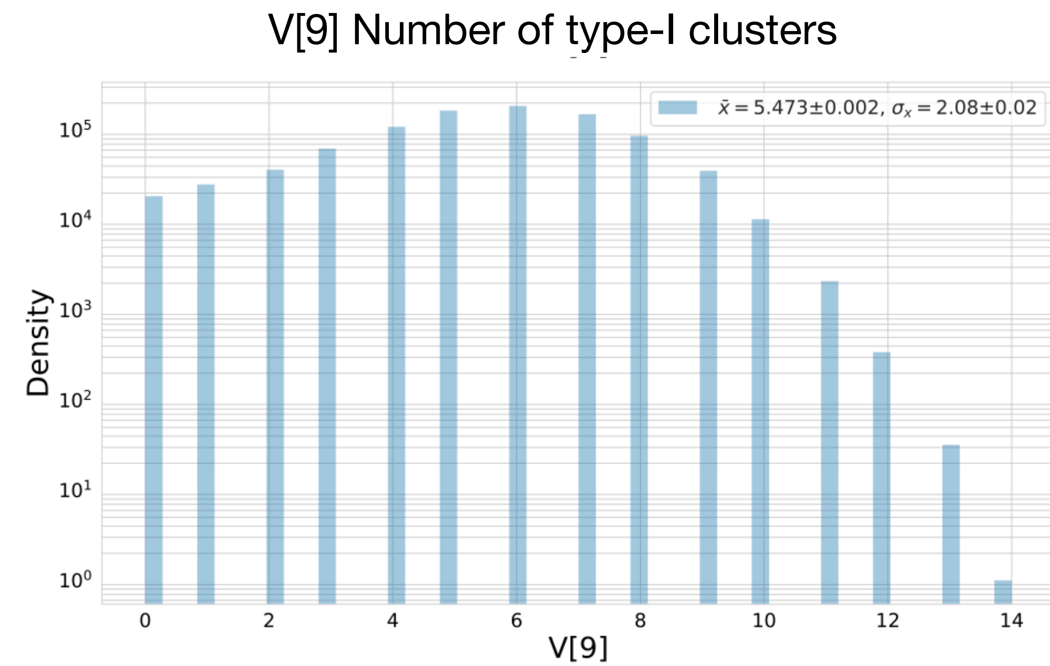


Input for Machine Learning

- Due to **stochastic nature of radiation** we **cannot** simply **rescale the reconstructed energy** sum to true energy
- → Build **high-level features** with our domain knowledge by using **spatial information** of the energy deposits
- → Exploit the **raw-hit data** of the calorimeter cells with **3D Convolutional Neural Networks** to learn to predict the muon energy - this can also be combined with the high-level features

High level features I: Clustering

Photons producing **showers** by pair production will produce a **signal in multiple cells** → decipher the pattern of emitted radiation by **aggregating** the granular cell-based **information into clusters**



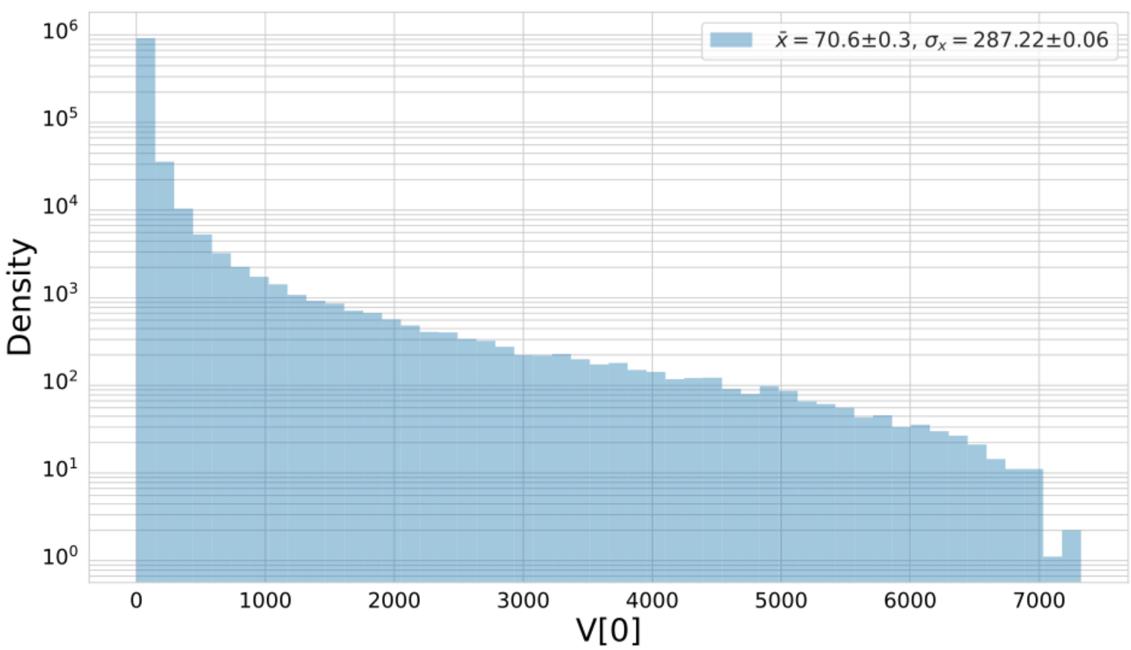
Clustering algorithm

- The algorithm starts with the **highest energy cell** as a seed belonging to the column of **same transverse coordinates** x and y of the **incident muon** if the cell fulfills $E_{thr} > 0.1$ GeV
- **Adjacent cells** with **non-null energy deposition** are progressively added to form a cluster
- The **final cluster** is formed when there are **no more adjacent cells** with non-null energy to be added and the algorithm is repeated with the remaining cells
- Once all clusters are formed, a **second set of clusters** is constructed using cells yet unassigned with **seeds irrespective of the x, y coordinates**
- **Compute features** like number of clusters per muon / maximum number of cells per cluster

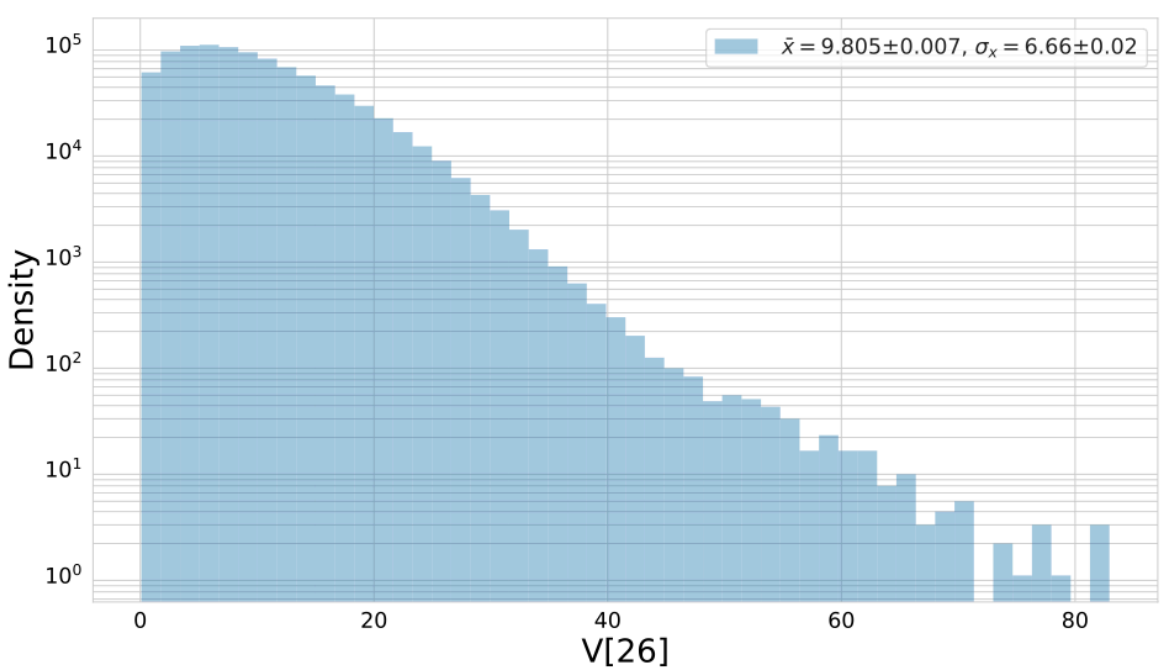
High level features II

- **E-sum features**: sum up energies in different intervals of cell energies
- **Energy 1st / 2nd moments** in x & y
- **Energy spread** for range of z depths
- Construct features **sensitive to curvature**
- **Missing transverse energy**
- In total **28 high level features**

V[0] Total energy in cells above 0.1 GeV

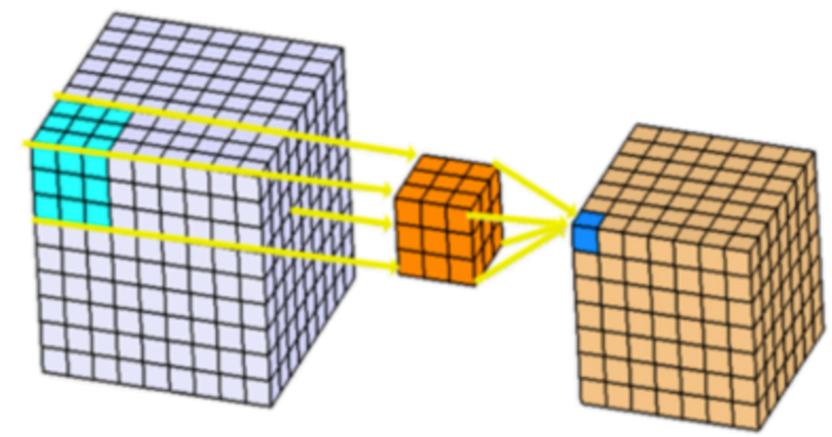


V[26] Total energy in cells below 0.01 GeV



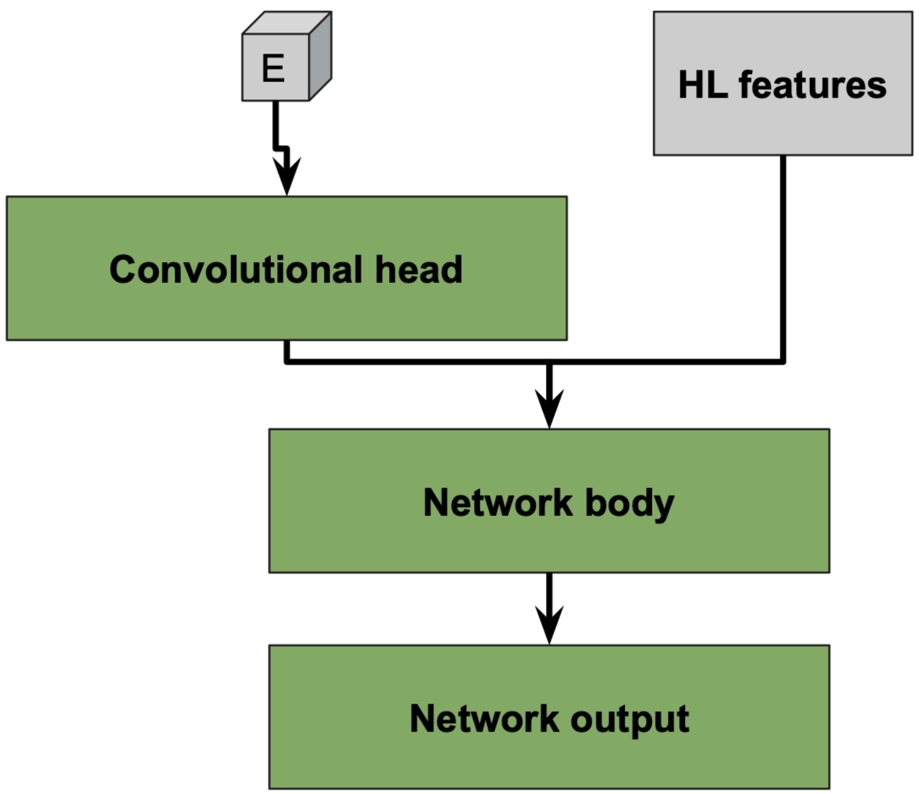
Muon regression model

3D Convolutional Neural Networks



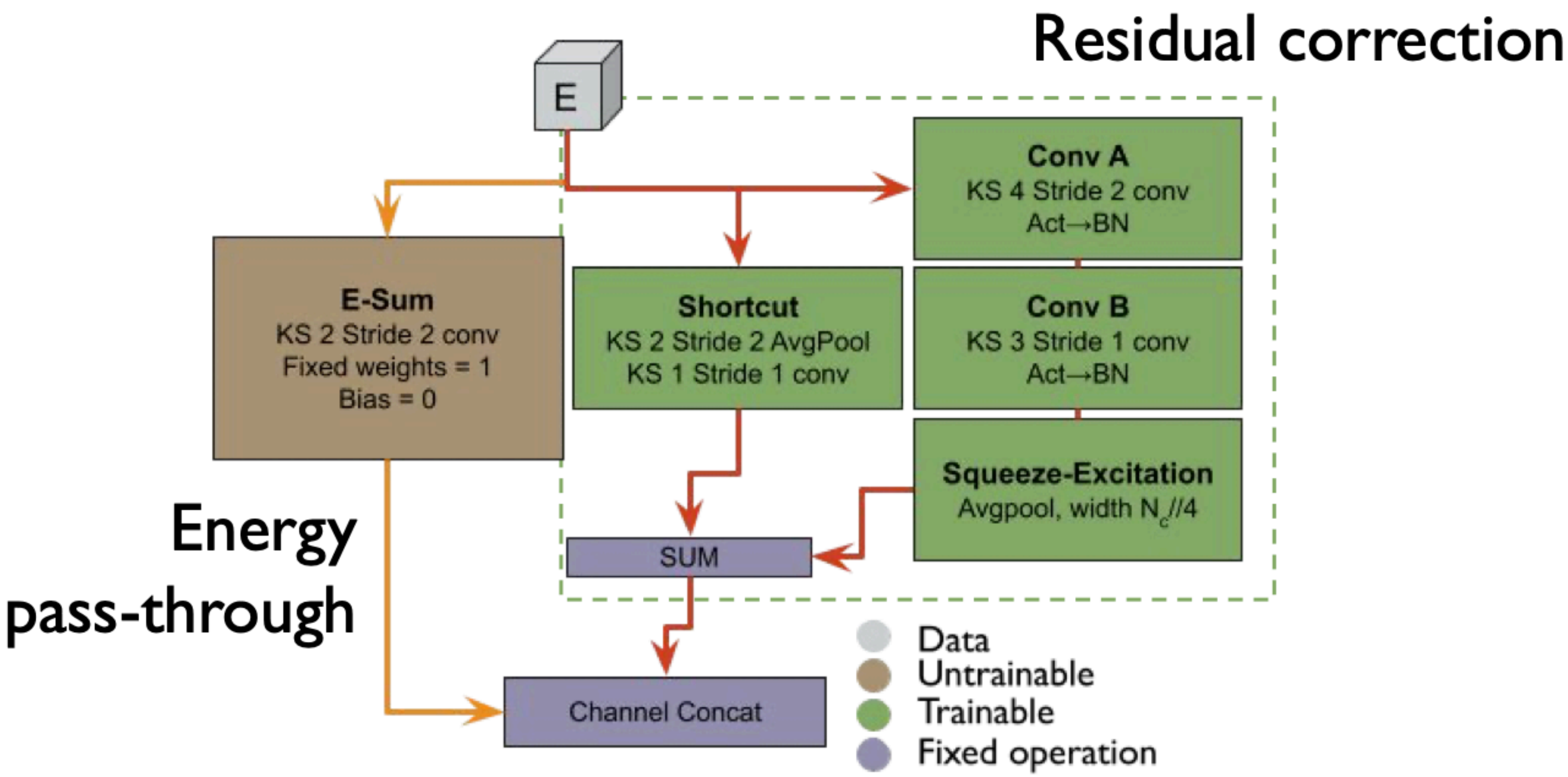
- Apply a **3 dimensional** filter to the data
- **Filter moves** in (x, y, z) direction to calculate the **feature representations**
- **Output** shape is a **3 dimensional volume** space

Muon regression model scheme



- **Convolutional head**: series of 3D convolutional layers (blocks) that exploits the 3D grid of energy deposits
- **Network body**: set of fully-connected layers with single neuron output
- Pre-computed **high-level features** are passed directly to the body

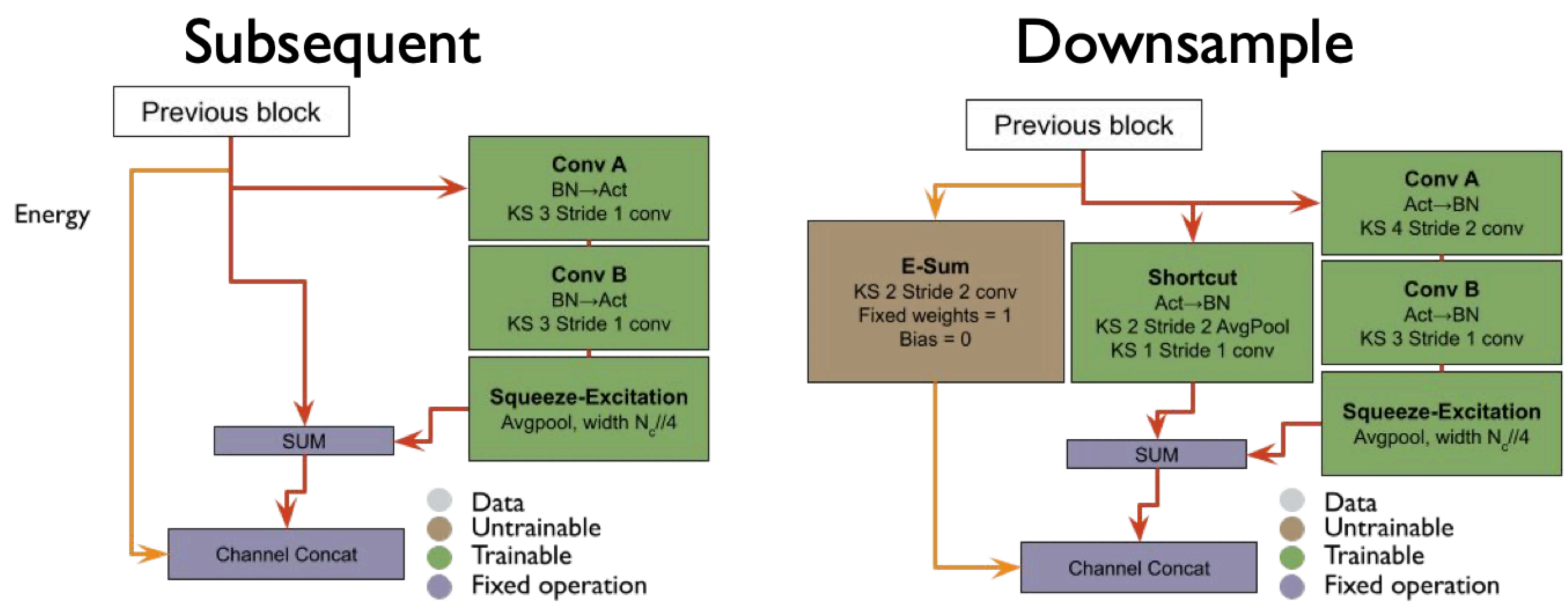
CNN block

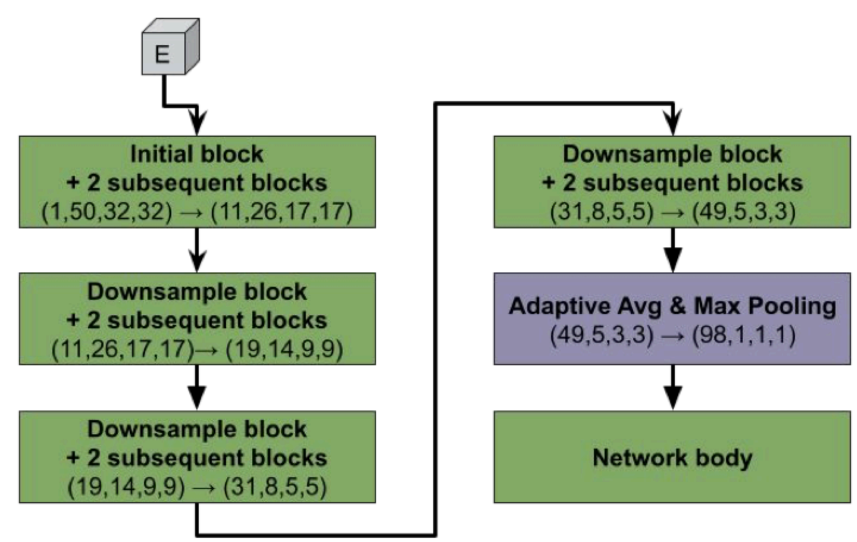


- **3D CNN** architecture aims to **learn small corrections to the reconstructed energy**
- Correction learnt by **residual convolutional layers**
- **Energy sum** is **concatenated** to output → always available to later layers
- **Running BatchNorm** helps with data sparsity
- **Squeeze-excitation block** further improves performance

CNN head

- Sets of **convolutional blocks** are used to construct the **full convolutional head**
- **Deeper networks** can be built by by adding in blocks which do **not downsample the grid**
- **Full CNN** contains **12 blocks**, followed by **mean and max aggregation**
- 51.200 inputs → **98 features**
- CNN head outputs **combined with HL features** and fed into the network body



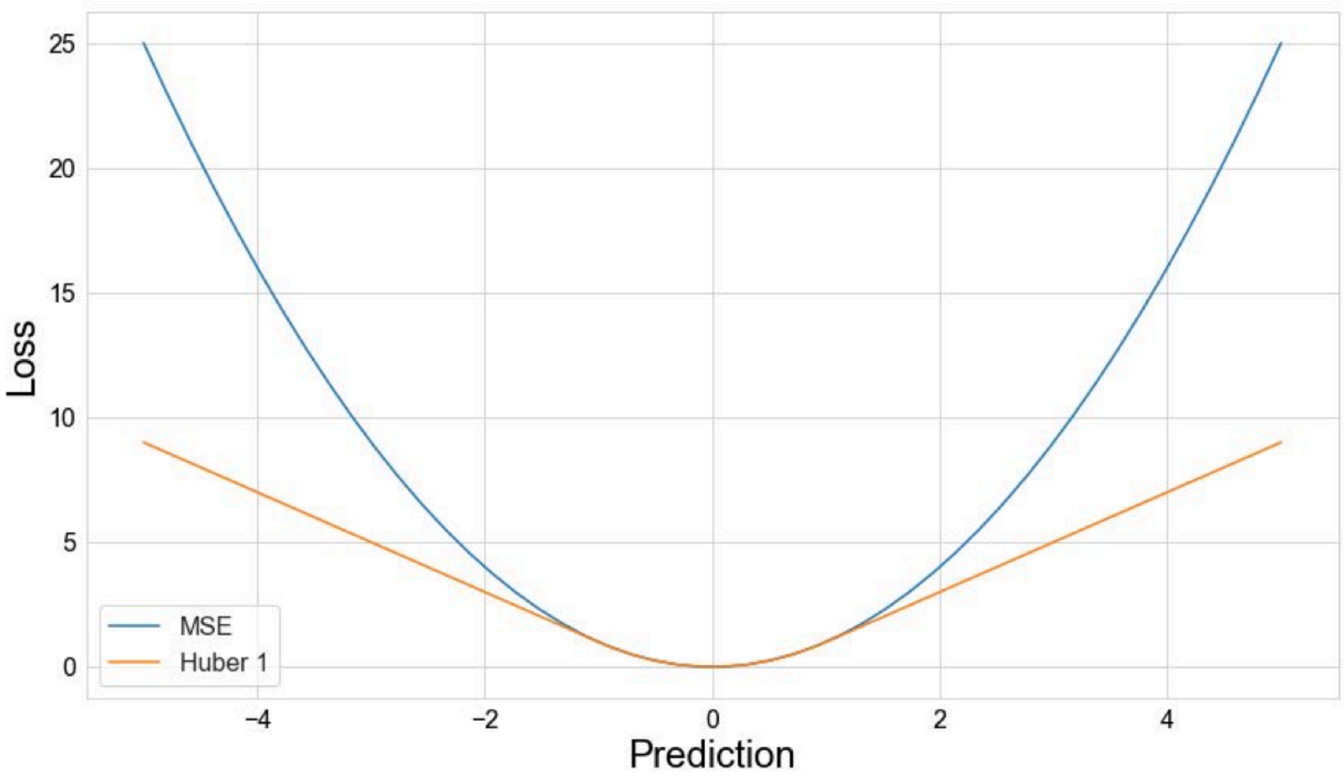


Loss function

MFSE

$$\mathcal{L}(y, \hat{y}) = \frac{1}{N} \sum_{n=1}^{N} \frac{(y_n - \hat{y_n})^2}{y_n}$$

Huberised MSE



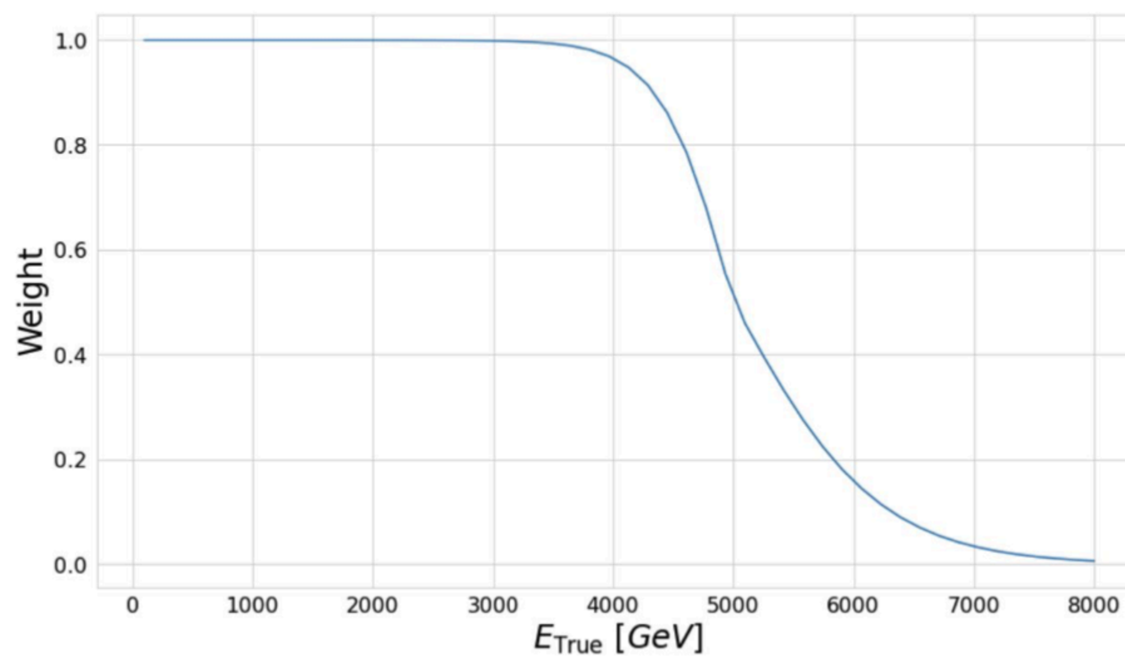
- Loss is **inspired** by linear scaling of the variance of the energy measurement with true energy, as is the case for **calorimeter showers**
- **Huberised version** of the **mean fractional squared error** (MFSE): **prevents outliers** from **dominating the loss** by modifying high-loss predictions above a threshold

Training

Data weighting

- **Hard boundaries** in training data → **biased regression** response due to **edge effects**

- Solution: **train** on **50-8000 GeV** muons - evaluate on **100-4000 GeV** muons and **compensate training** by **down-weighting** muons **above 4 TeV**

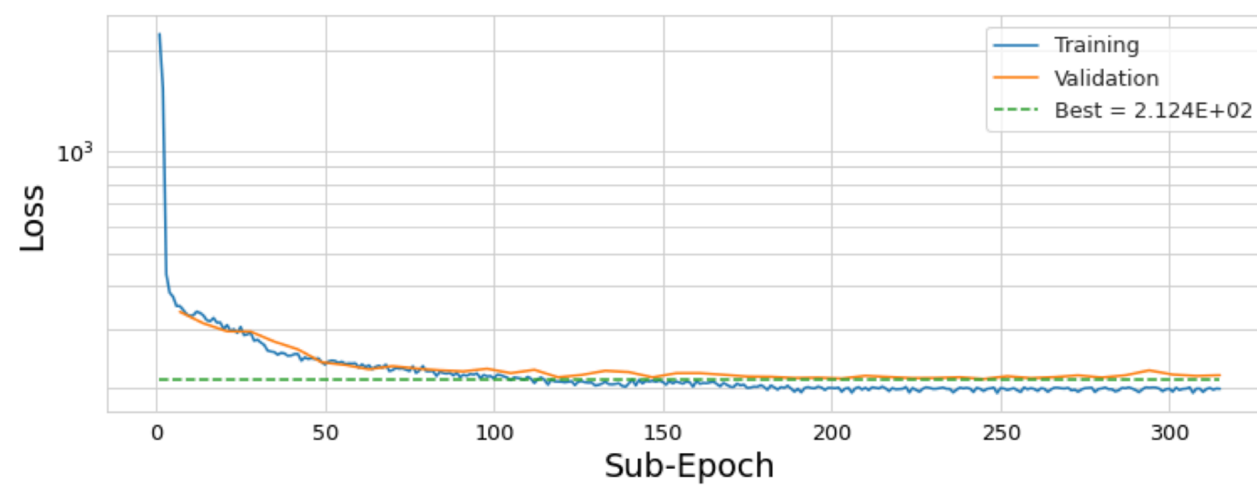


Ensemble training

- Train **ensemble of 5 models** with batchsize of 256 with **LUMIN**

- Uses **Adam** optimiser with**1-cycle schedule** with **cosine annealing** for 20 epochs

- Training time: 23*5 h on a Nvidia V100S

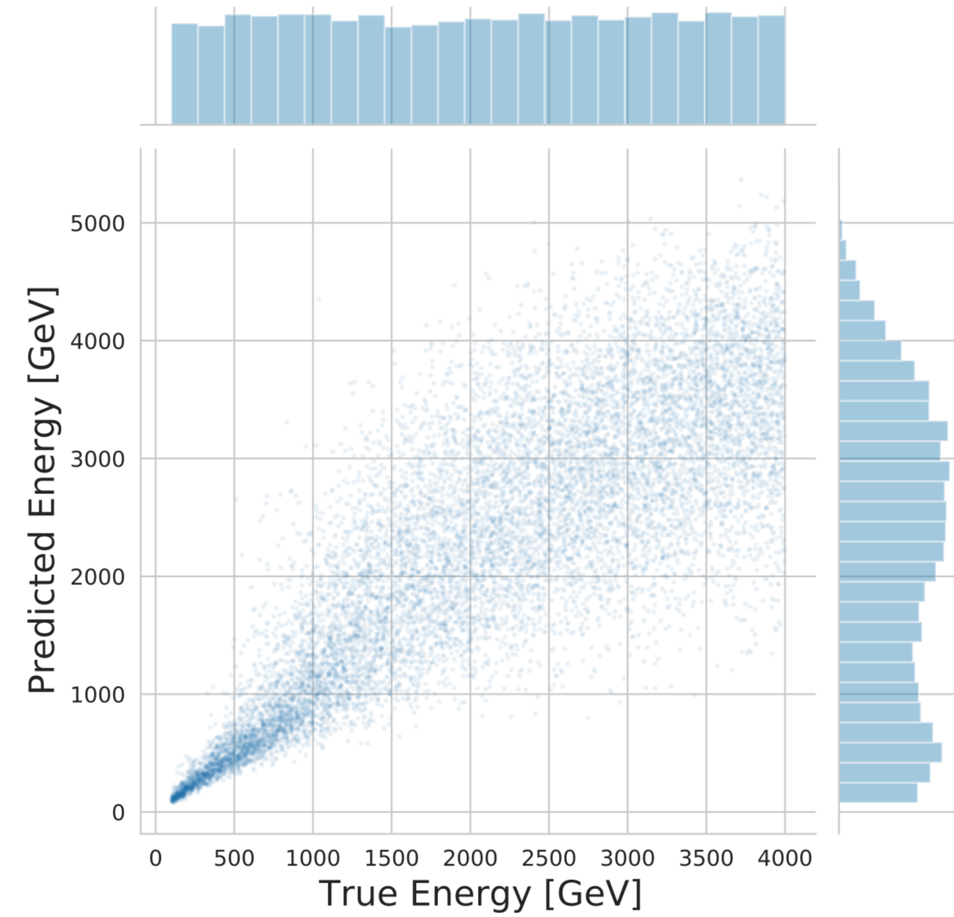


Regression results

Regressor response

- **Regression** predictions scale **quite linear** with true energy

- Tend to slightly **over-predict medium energy** and **under predict high energy**

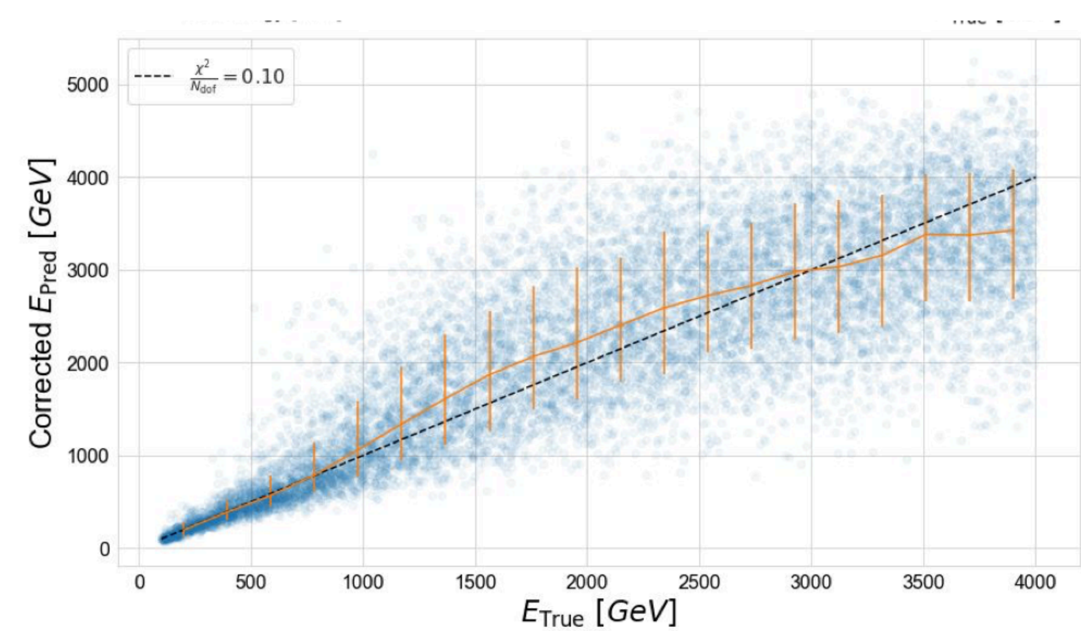


Bias correction

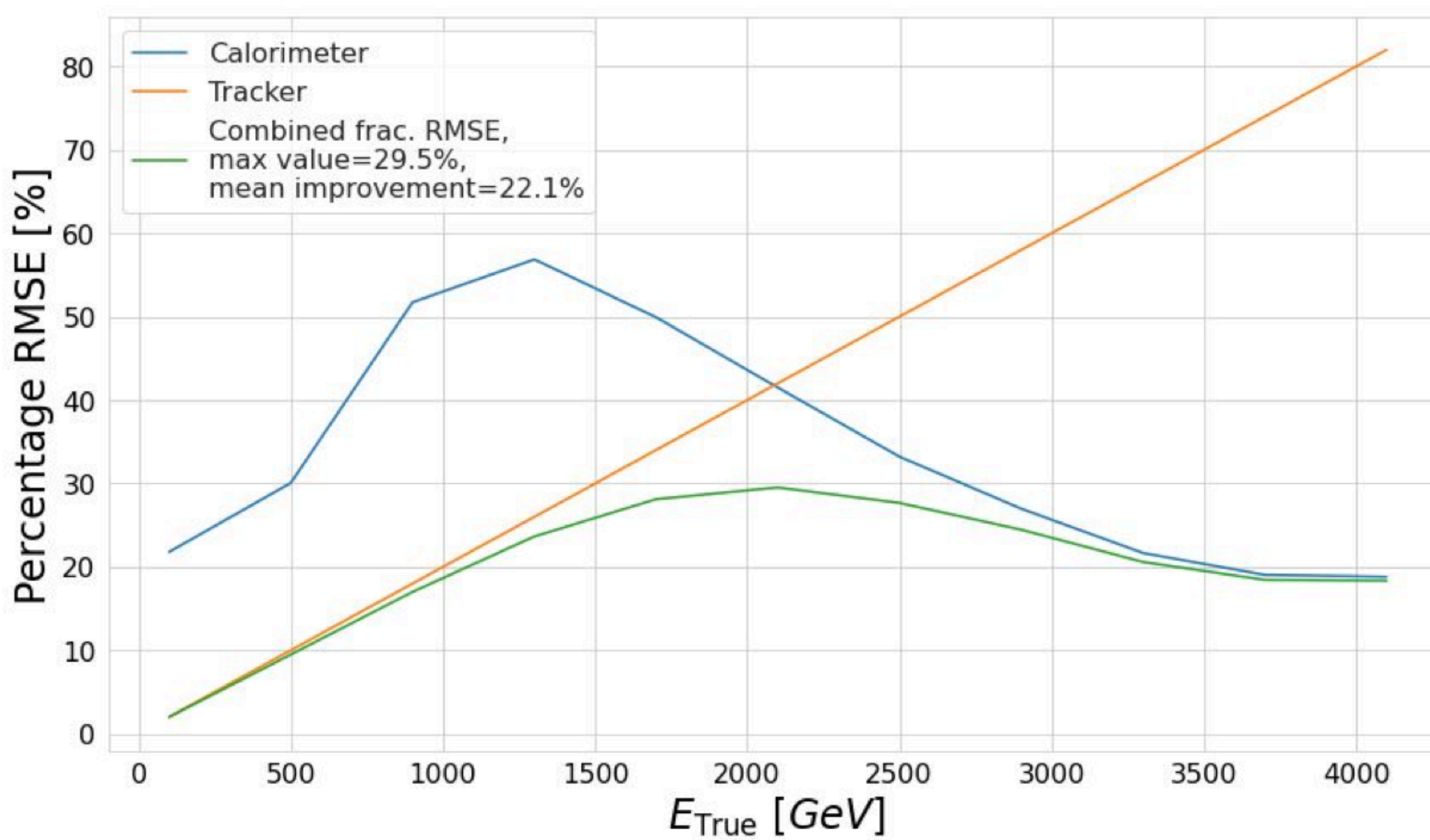
- **Correct bias** via **linear fit** to predictions in bins of true energy

- **Inversion of fit** provides a function to **correct the predictions**

- **Correction** is **fixed** using **validation data** and then applied to testing data

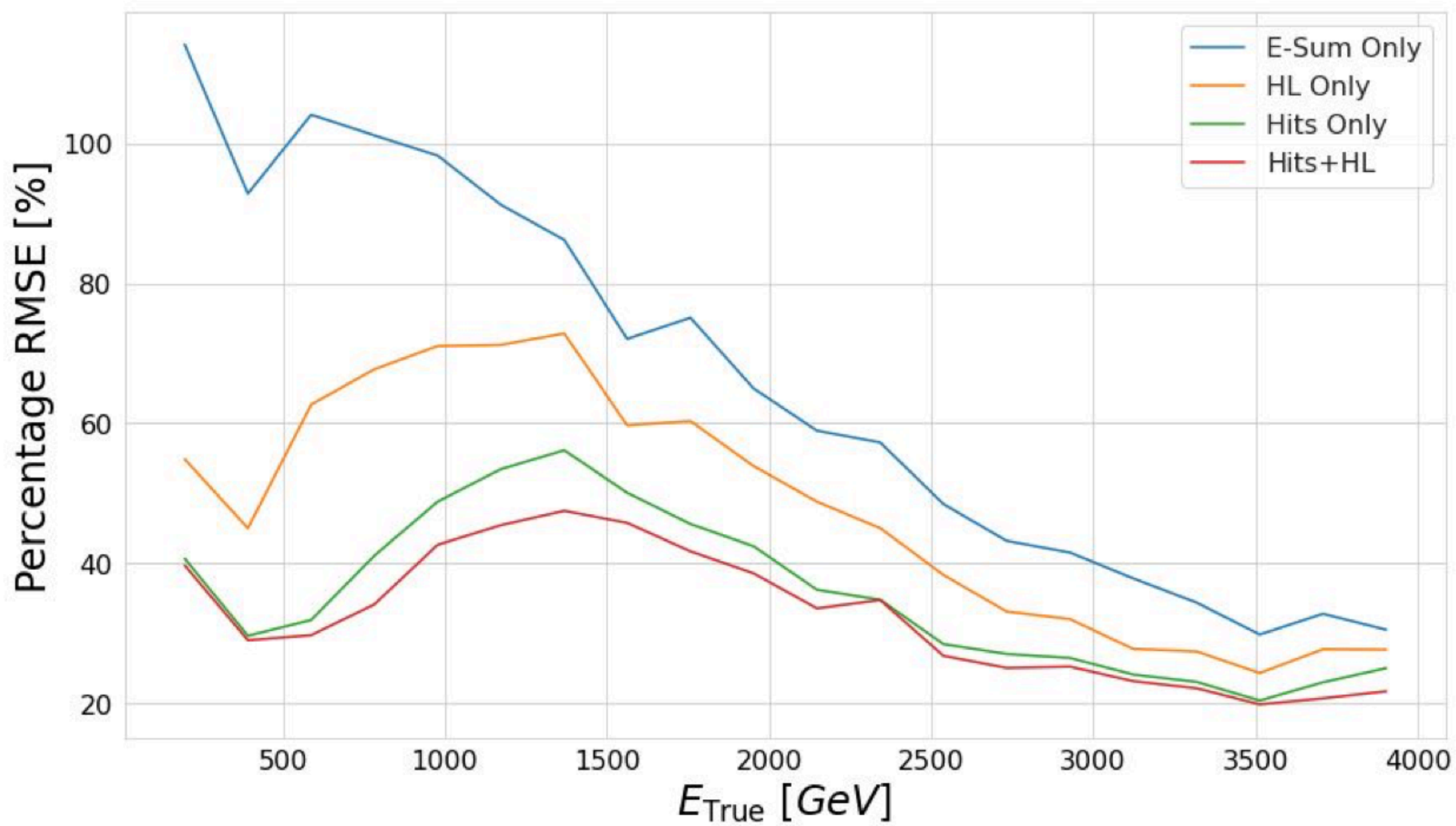


Resolution and combination with curvature measurement



- Compute **resolution** as **RMSE** quadrature of **sum of variance and residual bias**
- **Tracker measurement** assumed to have a resolution of 20% @ 1 TeV
- **Calorimeter** and **tracker** are **complementary** → best resolution is a **weighted average of both**
- The **improvement over tracker** due to calorimetric regression is useable as an **optimisation metric**

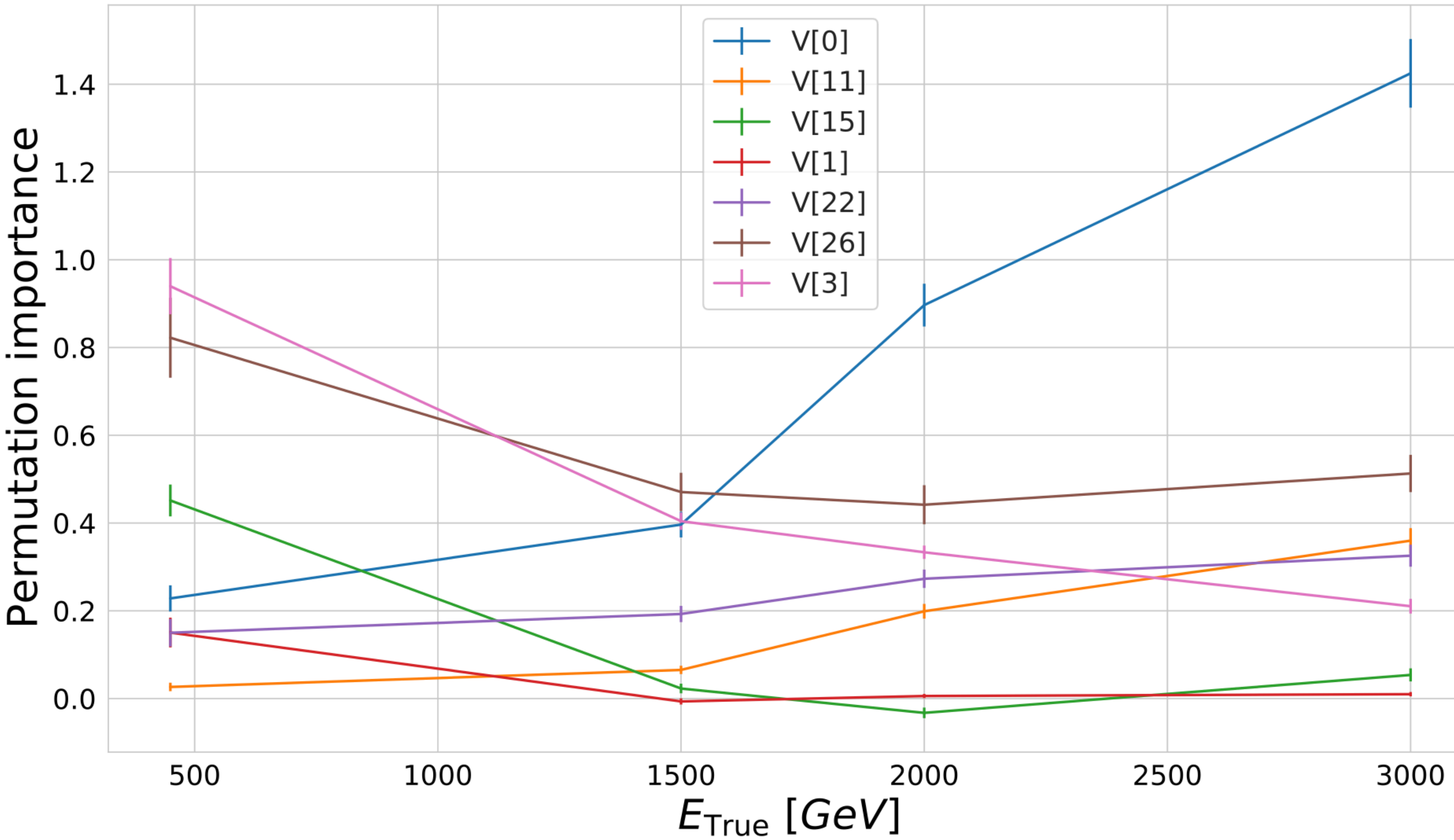
Comparison of the ML input



- **Quantify improvements** of Machine Learning by **comparing different models** that use only the reconstructed energy / only the HL features / only the raw 3D hits / the full model
- **High level features improve** significantly over using just the energy sum of recorded hits
- **3D CNN model** outperforms high-level features
- **High level features** still **beneficial at low energy** when the full CNN model is used

High level feature importance

- V[0] - E-sum in cells above 0.1 GeV
- V[1] - fractional MET
- V[3] - overall 2nd moment of transverse E distribution
- V[11] - maximum total E in clustered deposits
- V[15] - maximum energy of cells excluded from clustered deposits
- V[22] - relative 1st moment of E distribution along x-axis
- V[26] - E-sum in cells below 0.01 GeV



- **Feature importance** computed with **LUMIN**
- **Shift in importance** between **V[0] and V[26]** (the sum of energy in cells above 0.1 GeV and below 0.01 GeV)

Related work

Ice cube

- ***Abasi et al., 2013 & Aartsen et al., 2014***: measured muons but with much **larger calorimeter** (20-times larger in radiation-lengths)

- **Resolution 10x poorer** than ours **in the considered energy range,** however their focus is more on higher-energy muons

ATLAS

- ***Nikolopolous et al. 2007:*** ATLAS also demonstrated a **calorimeter-based method**

- **Sums up** the **deposited energy in cones**, rather than considering the spatial structure of hits

Conclusions

- We demonstrated that **radiative losses in calorimeters** can provide **muon-energy measurements** in collider experiments and achieve **good resolution** across the full muon-energy spectrum

- **First demonstration** in the context of particle colliders that utilises the **raw hit information of a calorimeter** to measure the **muon energy**

- **Calorimetric** measurements **improve with energy** and are **complementary** to existing **tracker** measurement approaches

- Can **improve** the **sensitivity** and **reach of searches for New Physics**

- **Preprint**: Jan Kieseler and Giles C. Strong and Filippo Chiandotto and Tommaso Dorigo and Lukas Layer, *Calorimetric Measurement of Multi-TeV Muons via Deep Regression*, **arxiv:2107.02119**