Jet Energy Corrections with GNN Regression

Daniel Holmberg

Learning to Discover, 27.4.2022

HELSINKI INSTITUTE OF PHYSICS

CMS
Compact Muon Solenoid

Introduction

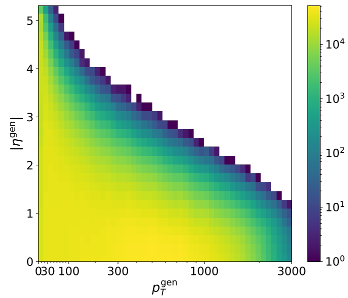
- The physical detector causes the jet transverse momentum p_T to be different from the true particle-level jet
- Corrected such that it agrees on average with the p_T of the particle level jet
  - Determined by using basic kinematic quantities of the jet
- Possible to include more information and get better corrections using machine learning
  - Has been done successfully for b-jets using a deep feed-forward neural network
- However, this study is about *generically* applicable ML-based corrections

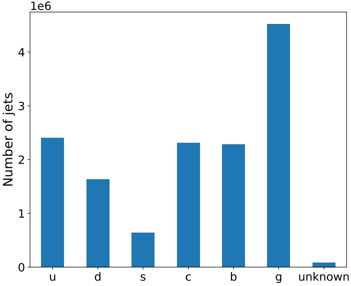
Dataset

- QCD H_T-binned samples, 2016 configuration
  - `/QCD_HT*_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/RunIISummer16MiniAODv3*/MINIAODSIM`
- Custom ML JEC dataset by A. Popov (ULB)
- Forked and added SV angles for initial coordinates in ParticleNet
- 14M jets: 60% training, 20% validation and 20% test

Data distribution

- Same shape for all jet flavours
- Flat in (p_T, η) at low p_T
- Steeply falling in p_T at high p_T
- Proportions of b, c, uds, and g jets fixed as $1 : 1 : 2 : 2$



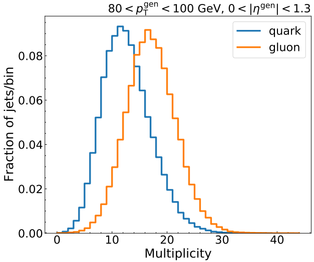


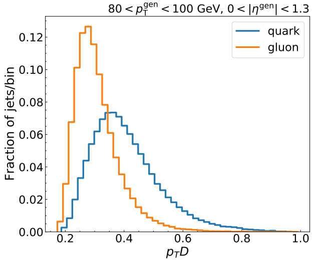
Training features

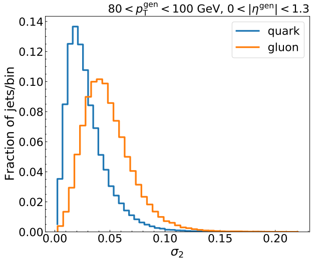
- Event level
  - p_T, $\log p_T$, η, ϕ
  - ρ, mass, area, num pv
- Charged PF candidates
  - p_T, η, ϕ
  - dxy, dz, dxy significance, normalized χ^2
  - num hits, num pixel hits, lost hits
  - particle id, pv association quality
- Neutral PF candidates
  - p_T, η, ϕ
  - particle id, hcal energy fraction
- Secondary vertices
  - p_T, η, ϕ, mass
  - flight distance, significance, num tracks

Feature engineering

Create event-level features: multiplicity, $p_T D$, σ_2 that helps with quark gluon discrimination







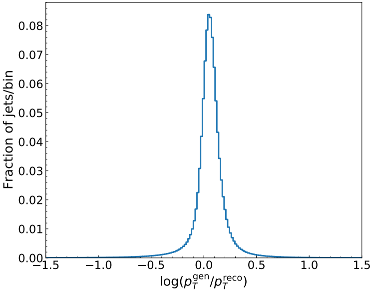
Feature engineering

- Relative features for all constituents
  - $\Delta p_{T,i} = p^{\text{pf}}_{T,i} / p^{\text{jet}}_{T}$
  - $\Delta \eta_i = \text{sgn}(\eta^{\text{jet}})(\eta^{\text{pf}}_i - \eta^{\text{jet}})$
  - $\Delta \phi_i = (\phi^{\text{pf}}_i - \phi^{\text{jet}} + \pi) \bmod 2\pi - \pi$
- One hot encode categorical features
  - particle id and primary vertex association quality
  - e.g. neutral pid:

```
[1, 2, 22, 130] -> [
  [1, 0, 0, 0], [0, 1, 0, 0],
  [0, 0, 1, 0], [0, 0, 0, 1]
]
```

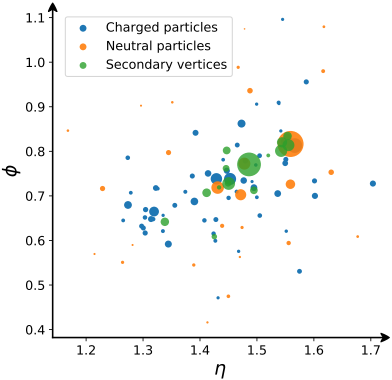
Target and loss

- Regression target $y = \log(p_T^{\text{gen}}/p_T)$
  - Correction factor is thus $e^{\hat{y}}$ where $\hat{y}$ is the NN output
- MAE loss function $L = \frac{1}{N}\sum_{i=1}^{N} |y_i - \hat{y}_i| I_{|y_i|<1}$
  - The last factor rejects 0.8% of jets where the target is way off



Choice of ML models

- For every jet there are global features as well as constituent features forming a particle cloud



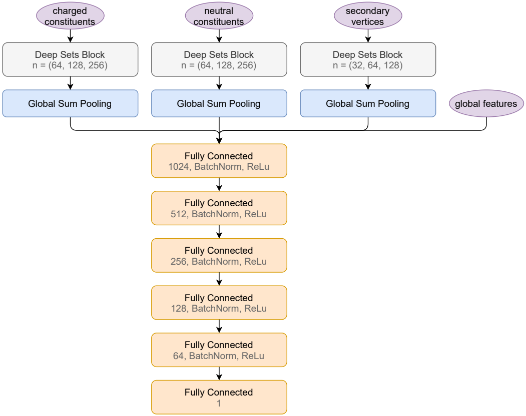
Choice of ML models

- Jet constituents can be represented as a *permutation invariant set*
  - Number of constituents varies from jet to jet
  - Order doesn't matter
  - $\Rightarrow$ Requires special treatment to use it for ML
- Deep Sets and Dynamic Graph CNN are examples of NN architectures allowing for unordered sets to be consumed
  - They have been used for *jet flavour tagging* in Energy Flow Networks and ParticleNet respectively

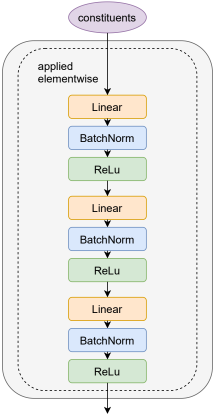
Particle Flow Network

- Shared MLP applied to every constituent
  - $\mathbf{h}_i = \phi(\mathbf{x}_i)$
- The learned parameters are aggregated using a permutation invariant operation
  - Global sum pooling chosen in line with the Deep Sets article
- Concatenate with global features and feed into MLP
  - $f(\mathbf{X}) = \rho\left(\sum_{i \in \mathcal{V}} \phi(\mathbf{x}_i)\right)$
- Any function $f(\{x_i\})$ invariant under permutations of its inputs can be approximated arbitrarily well as $\sum_i \phi(x_i)$

PFN-r



(a) Complete network

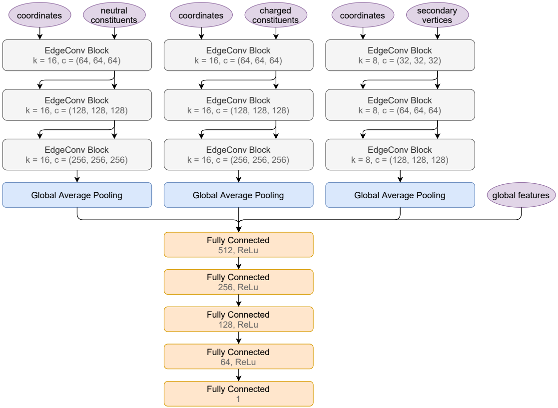


(b) Deep Sets block

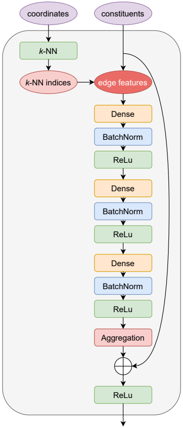
ParticleNet

- Edge convolution
  - Initial graph in $(\Delta\eta,\ \Delta\phi)$ space
  - Local patch for every particle using k-NN
  - Define *edge features* for each center-neighbor pair
    - $\mathbf{h}_i = \phi\left(\mathbf{x}_i, \frac{1}{k}\sum_{i\in\mathcal{N}_i^k}\psi(\mathbf{x}_i, \mathbf{x}_j - \mathbf{x}_i)\right)$
  - Perform permutation invariant aggregation
    - Chose *mean* which is used in the original article
  - Dynamic graph update
    - Edge features new coordinates in high-dim latent space
- Concatenate with global features and feed into MLP
  - $f(\mathbf{X}, \mathbf{A}) = \rho\left(\frac{1}{n}\sum_{i\in\mathcal{V}_i^n}\phi(\mathbf{x}_i, \mathbf{X}_{\mathcal{N}_i^k})\right)$

ParticleNet-r



(a) Complete network



(b) EdgeConv block

Training

- Four models are trained
  - PFN-r, PFN-r Lite, ParticleNet-r ParticleNet-r Lite
- Using TensorFlow 2.4.1
- MirroredStrategy on two Nvidia GeForce RTX 3090 cards
- Adam optimizer
- Batch size 1024
- Learning rate 2×10^{-3}, reduced by a factor of 5 when validation loss plateaus
- Regularization through early stopping callback
- Code: gitlab.cern.ch/dholmber/jec-gnn

Effective data pipeline

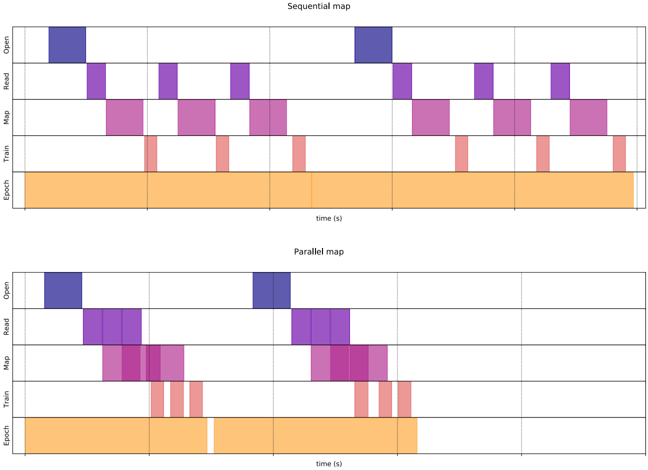
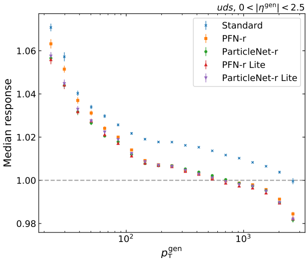
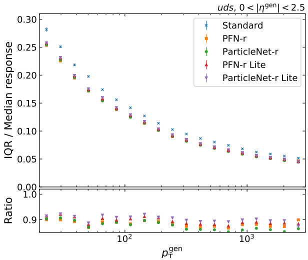


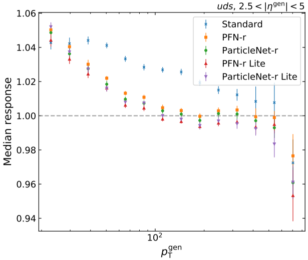
Figure: Naive and parallel data handling in TensorFlow.

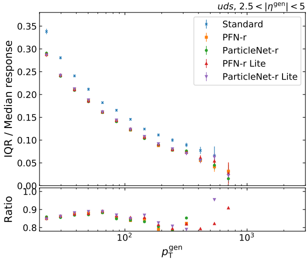
Results

uds jet response

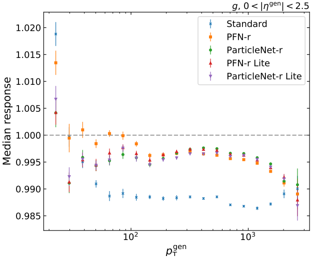


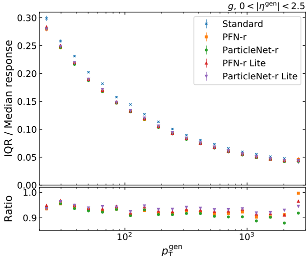


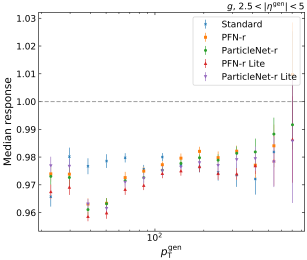


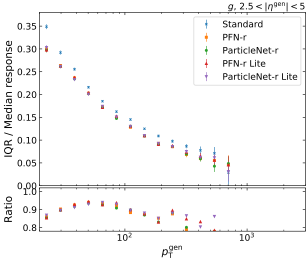


gluon jet response

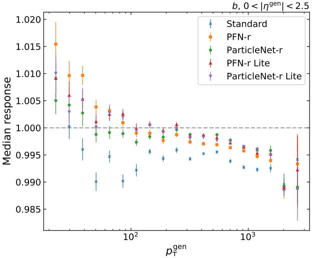


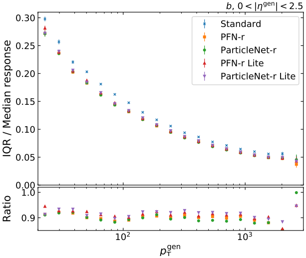


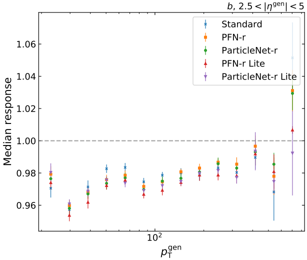


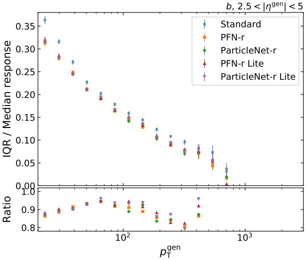


b jet response

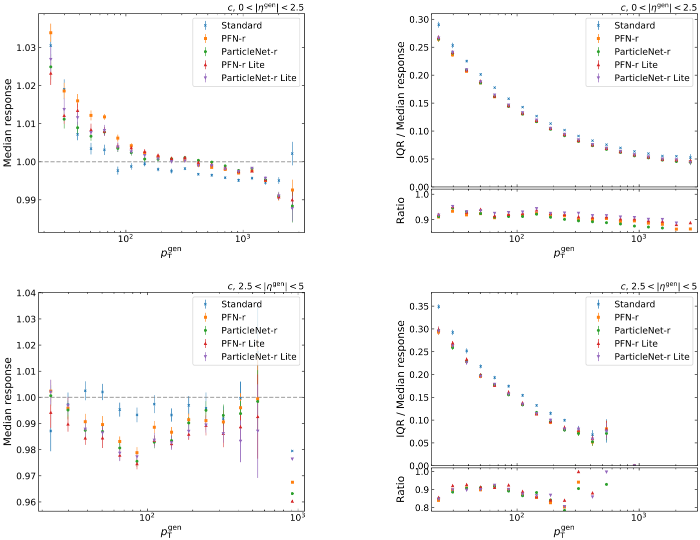




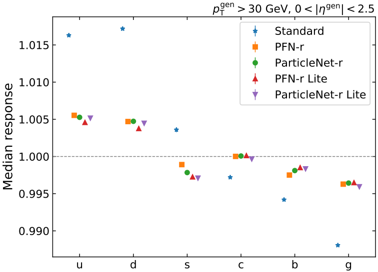


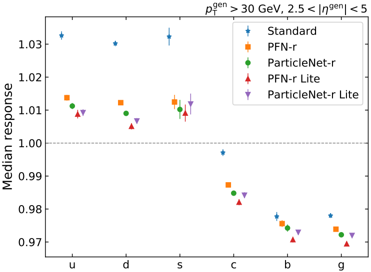


c jet response



flavour dependence





Summary

- Improved p_T resolution w.r.t standard corrections
  - 10-15% for uds jets, 10% for b & c jets and around 8% for g jets in the central region
  - 10-20% for uds jets and 5-20% for the rest of the jets in the forward region
- Reduced flavour dependence
  - 70% improvement in central region and 30% in forward region
- ParticleNet-r vs PFN-r
  - ParticleNet-r has 270k less parameters
  - Despite this it achieves slightly better resolution, especially for jets with higher p_T
  - Also slightly less flavour difference for the response
  - However, PFN-r is 9$\times$ faster to train and has 14$\times$ shorter inference time

Thank You!

Questions?

daniel.holmberg@cern.ch | www.danielholmberg.fi