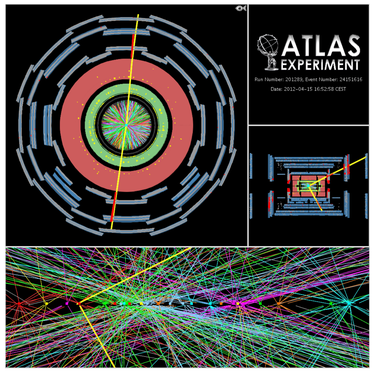
Neural Networks for calibrating ATLAS jets

Murtaza Safdari (murtazas@stanford.edu) and Nicole Hartman (nicole22@stanford.edu)

Motivation

- ATLAS is a particle detector analyzing proton-proton collisions from the LHC.
- Jets are collimated sprays of particles in the detector.
- Other low energy collisions in a beam crossing can add extra energy to the jet degrading our accuracy in reconstructing the jet's transverse momentum, p_T.
- Goal: Take energy depositions in the detector and use regression and neural networks to improve the p_T reconstruction that is uniform over a range of p_T values.



Features

- $p_T(j\ no\ area\ sub) = \sum_{clusters} p_T(cluster)$
- $p_T(j0) = \sum_{clusters} p_T(cluster) - \rho A_T$
- Rings of p_T summed in annuli of $\Delta R = 0.5$ (Jet radius: $R = \sqrt{\eta^2 + \phi^2} = 0.4$)

Currently ATLAS accounts for pile-up using $p_T^{\text{reco}} = p_T^{deposited} - \rho A_T - \alpha(N_{PV} - 1)$.
Then numerical inversion accounts for the non-uniformities of the detector.

Results II

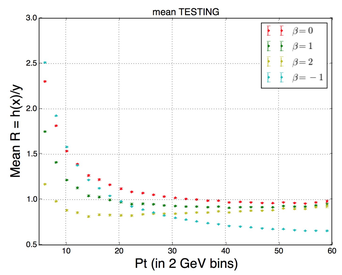
Since the neural networks aren't expected to do well at the high and low p_T limits, evaluate models using the error in p_T range [45,120] GeV.

Models	CV error
Inversion on $j0$	0.0327 ± 0.0006
Inversion on $j\ no\ area\ sub$	0.0311 ± 0.0006
NN with $j0$ input	0.0151 ± 0.0007
NN with j$j\ no\ area\ sub$	0.0138 ± 0.0007
NN: $j\ no\ area\ sub$, ρA_T	0.0153 ± 0.0007
NN: $j\ no\ area\ sub$, ρA_T, N_{PV}	0.0141 ± 0.0007
NN with 8 annuli and $j0$	0.0129 ± 0.0007

All the neural networks are sequential and have 1 hidden layer with 5 input nodes.
We prototyped a CNN on 50,000 jets with $(2 \times 8 \times 8) + (5 \times 5 \times 5) + (2 \times 2) + 5 + 1$ architecture, with a CV err over $p_T \in [20, 140]$ GeV of 0.076 ± 0.005

Cost Function

In initial linear regression studies we tried cost functions: $J_\alpha(\theta) = \sum_i w_i \frac{(y - h(x))^2}{y^\beta}$, $\beta = 2, 1, 0, -1$.



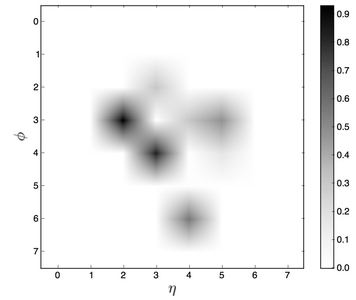
Optimize performance over entire p_T range by combining $\beta = 0$ and $\beta = 2$ cost functions:
$J(\theta) = \sum_i w_i (y - h(x))^2 \left(1 + \frac{\lambda}{y^2}\right)$, where $\lambda = 100$ to appropriately scale J_2 to J_0.

Dataset

Our dataset consists of ~ 6 million jets of detector level jets which contain:

- The true jet p_T
- The (η, ϕ) coordinates with the corresponding p_T for the clusters in a jet.
- Jet's transverse area: A_T
- Number of primary vertices in event, N_{PV}
- The event weights.

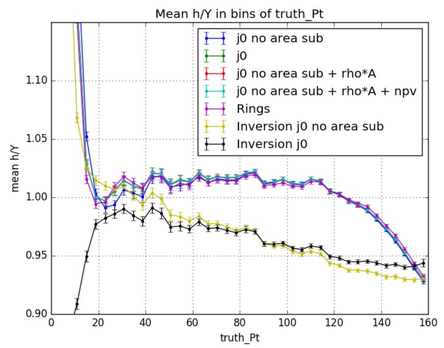
We divided our data into three samples: 80% as training set, 10% as CV set, and 10% as test set.
Since we had vastly more low p_T events, we revised the event weights by flattening the truth p_T distribution in 4 GeV bins.

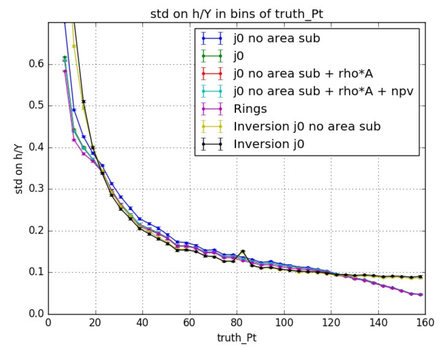


Pixel image of a jet in a 8×8 matrix binned in the (η, ϕ) coordinates relative to the jet's center with the bin's greyness set by the p_T.

Results I

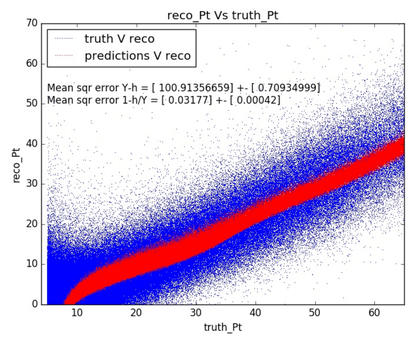
To ensure uniform performance over the entire p_T range, we looked at the closure $h_\theta(x)/y$ in 4 GeV bins and calculated the mean and standard deviation.





Discussion

- Every NN that we derived performs better than the current ATLAS standard, numerical inversion.
- The NN with annuli info does best, showing that jet substructure helps.
- All other NNs perform comparably.
  - The $j\ no\ area\ sub$ with ρA_T contains the same info as $j0$.
  - No new info appears to be gained by including the N_{PV} variable.



Future Steps

- Use a series of classifiers to predict which truth p_T bin a given jet is in, then train a regression model on the combined outputs.
- Improve and understand CNN architecture to better adapt to ATLAS jet reconstruction.

References

[1] The ATLAS Collaboration. Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ tev using the atlas detector. *arXiv 1510.03823*, 2015.

[2] Aviv Cukierman and Benjamin Nachman. Mathematical properties of numerical inversion for jet calibrations. *arXiv 1609.05195v1*, 2016.