Monte Carlo simulation studies for unfolding hadron-in-jet multiplicity measurements at STAR

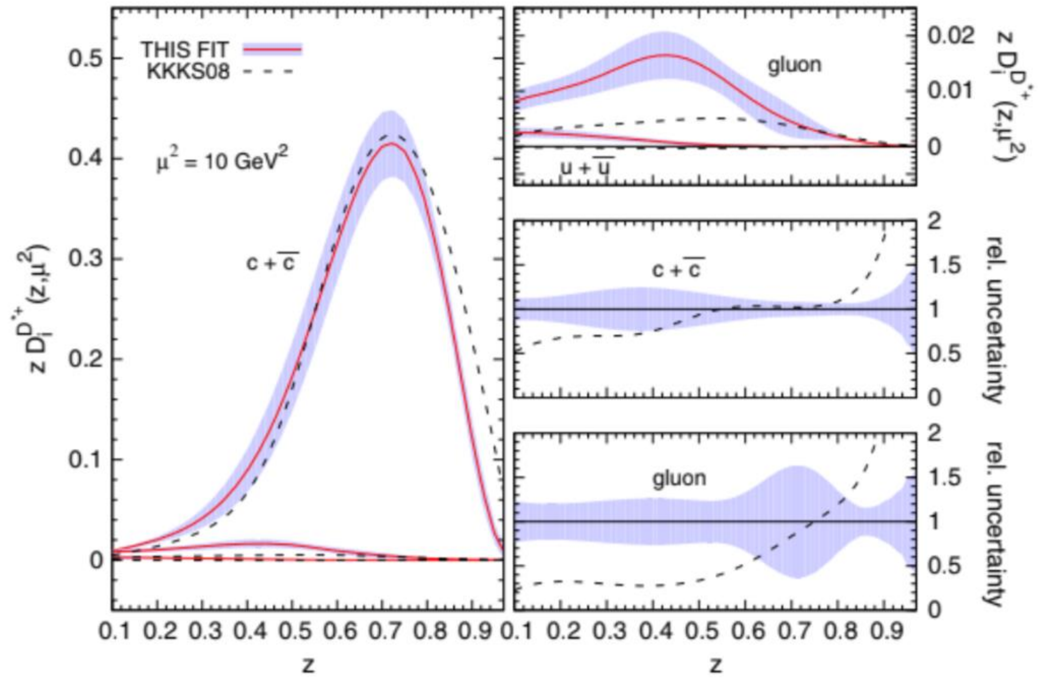
Roy Salinas for the STAR Collaboration
Abilene Christian University, USA

Overview

- Physics Motivation
- RooUnfold
- TMVA
- Conclusion

Physics Motivation

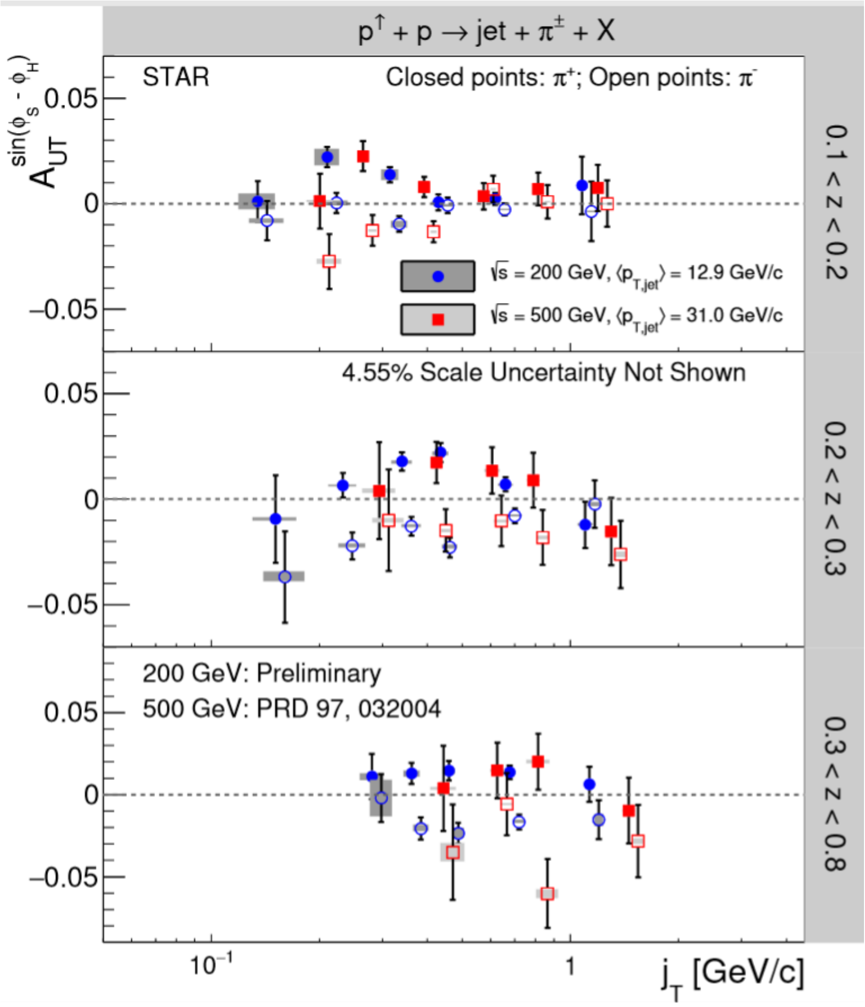
- **Measuring multiplicities for hadrons-in-jets can lead to a deeper insight into the gluon fragmentation function (GFF)**
- Several measurements have been made with the ATLAS & ALICE detector at the LHC
- Intriguing global analysis
  - ATLAS D*-in-jet data not well described by fits that only use e^+e^- and inclusive *pp* data (JHEP05 (2016) 125)
  - Anderle et al showed a global analysis that describes the D*-in-jet data with a larger gluon FF without compromising fits to e^+e^- and inclusive *pp* (*PRD 96, 034028 (2017)*)
- Their work provides an example of a framework that could incorporate future in-jet FF analyses



Anderle, Kaufmann, Stratmann, Ringer, and Vitev, PRD 96, 034028 (2017)

Physics Motivation

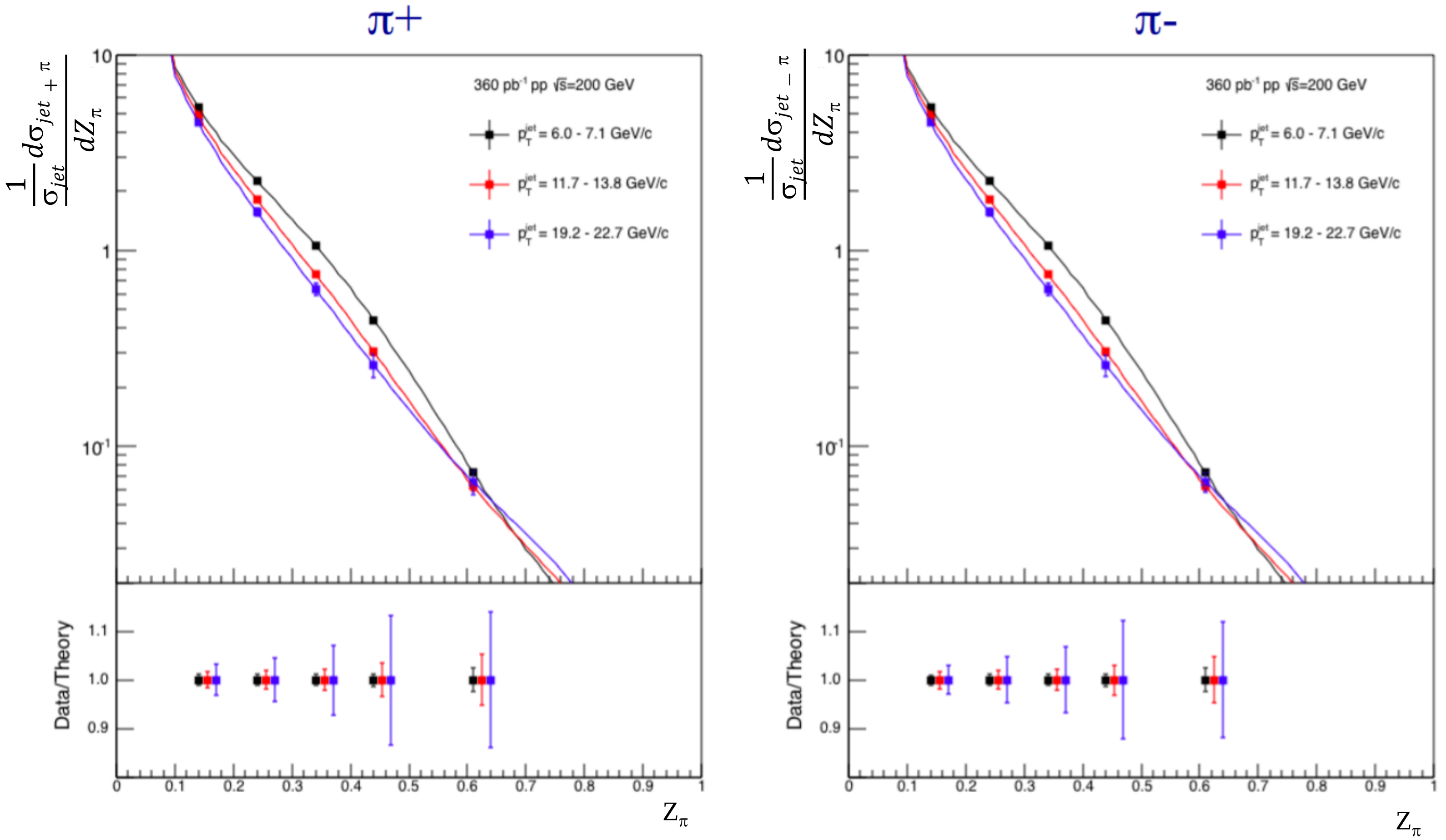
- **Hadrons-in-jets observables show promise in further investigation of spin asymmetries**
- Interesting structure in Collins asymmetry seen by binning in jet p_T, j_T, and Z
- **Models need data on unpolarized multiplicities**
- Need to understand the role of the GFF in denominator of Collins asymmetry
- Constraining the GFF via multiplicity measurements can lead to a better understanding of the Collins asymmetry



Preliminary 200 GeV: Int. J. Mod. Phys. Conf. Ser. 40, 1660040

Physics Motivation: From Polarized to Unpolarized

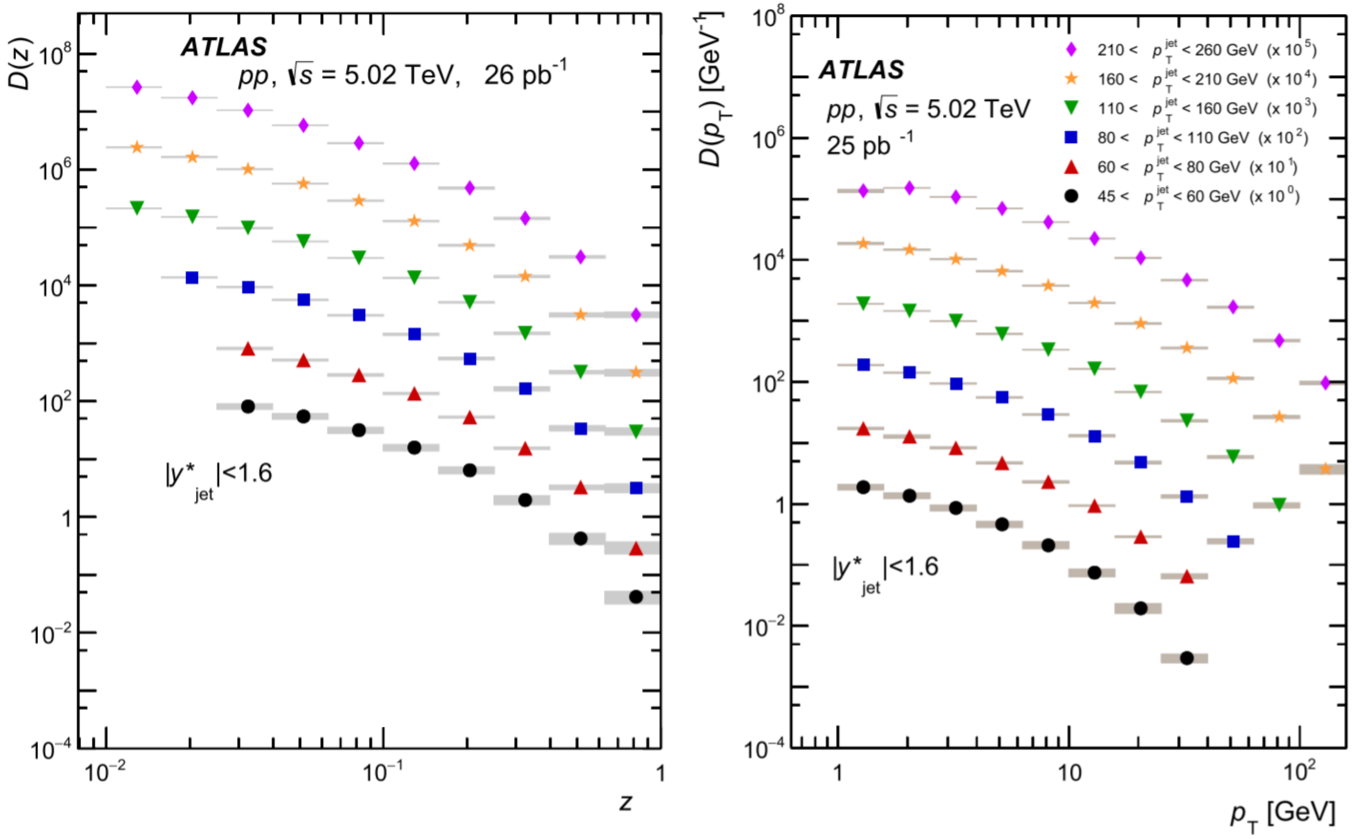
- **Efforts in understanding the GFF are being done at RHIC**
- To move from spin asymmetries to unpolarized multiplicities
    - Need detailed understanding of tracking efficiency corrections and systematics
    - Unfolding of kinematic variables (Jet p_T and hadron j_T and Z)
    - Unfolding could utilize RooUnfold and TMVA
    - Need detailed understanding of tracking efficiency
        - Significant work done to understand tracking efficiency at 200 GeV
        - Similar effort needed at 500 GeV



Anderle, Kaufmann, Stratmann, Ringer, and Vitev, PRD 96, 034028 (2017)

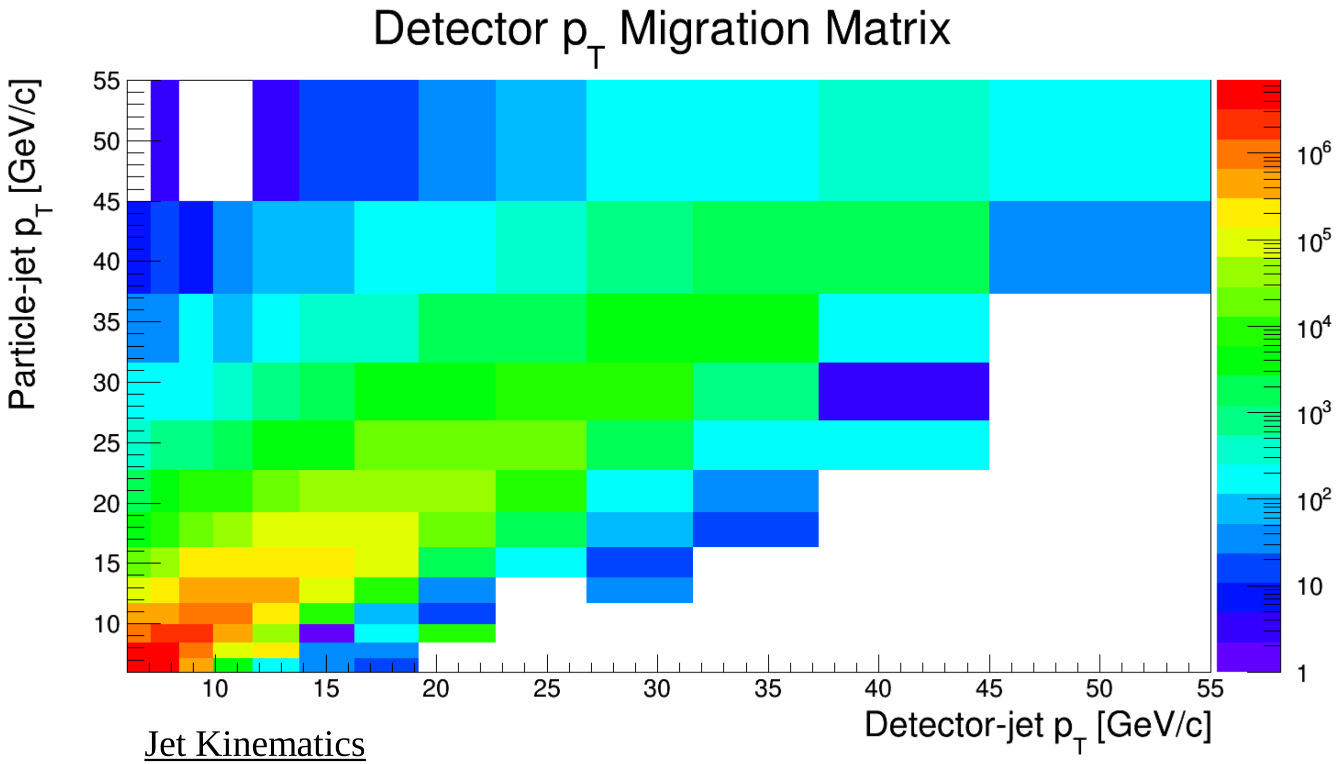
RooUnfold

- A native ROOT framework for unfolding
- Has been used by ATLAS and ALICE for a 2D unfolding
  - STAR regularly uses RooUnfold for cross-sections
- Is used for “deconvolutions” from detector read out
- Contains several unfolding algorithms
  - Iterative Bayesian
  - Singular Value decomposition
  - Bin-by-bin
  - Interface to TUnfold
- **Collins asymmetry binned in p_T, j_T, and Z**
  - RooUnfold can handle a 3D unfolding



NPA 978, 65 (2018)

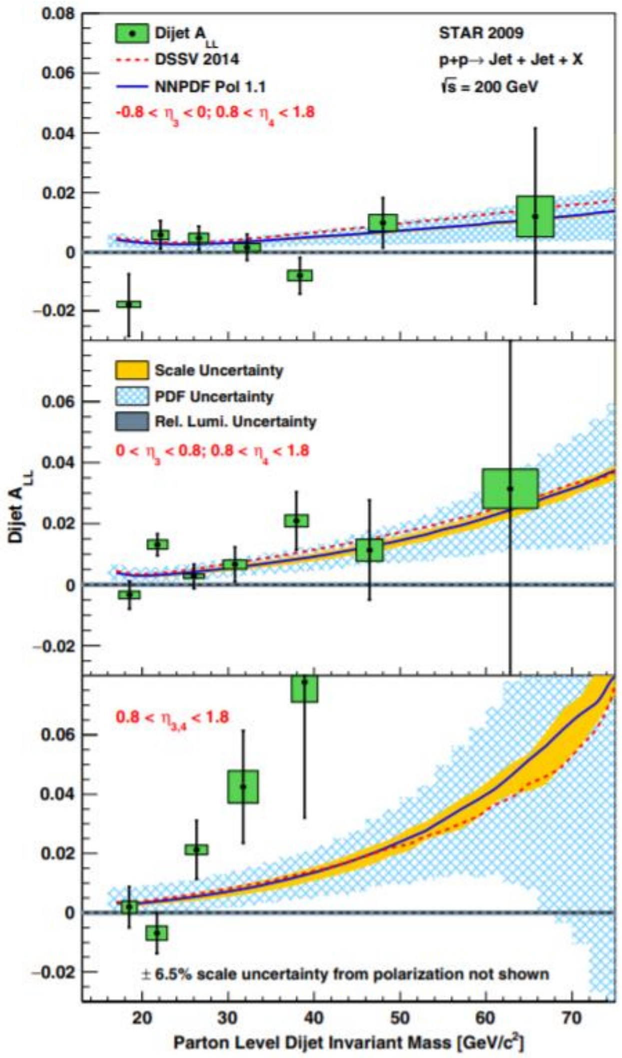
RooUnfold: Migration Matrix



<u>Jet Kinematics</u>
$\sqrt{s} = 500$ GeV
$|\eta| < 0.5$
$\Delta R < 0.5$
$p_T > 5$ GeV/c

- Matrix illustrating the migration seen at the detector from particle jets with RooUnfold
- Particle jet cuts were made in embedding file
- Similar matrices were filled for kinematic variables j_T & Z in bins of jet p_T
- Expected to see greater amounts of migration at high p_T
  - Due to detector resolution
- **Due to fine parsing of kinematic variables, large systematic uncertainties are prone to arise**
  - TMVA could assist in constraining uncertainties

TMVA

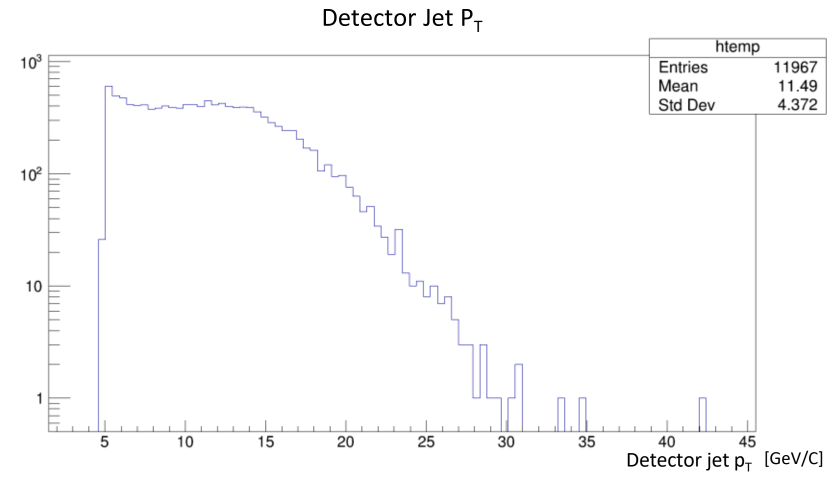


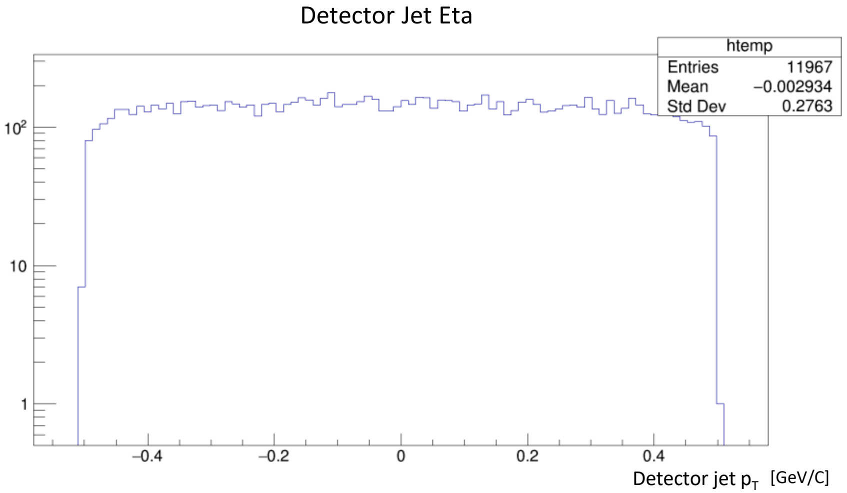
- TMVA
  - Toolkit for Multivariate Data Analysis with ROOT
- Is used for classification and regression (regression in our case)
- Has been used by STAR in asymmetry analyses
  - For A_{LL} di-jet asymmetries for $\sqrt{s} = 200$ GeV endcap analysis
  - 200 GeV: PHYSICAL REVIEW D 98, 032011 (2018)
- **Unfolding in 3D adds systematic uncertainties, TMVA handles this via machine learning methods**
  - Can help model the GFF
- Comes in two parts
  - Training
  - Application
- Has several methods available for regression
  - Multilayer perceptron, Deep Neural Network, Boosted decision tree,…

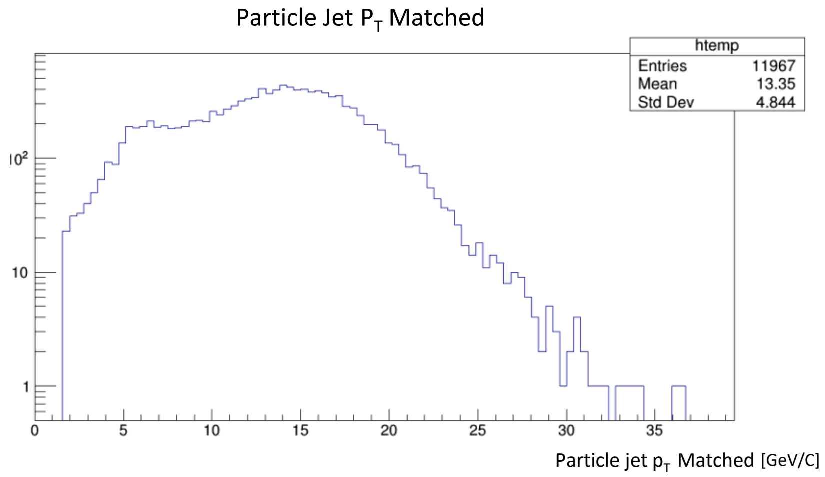
TMVA Regression : Input Variables

- **Used a small sample of Monte Carlo, generated with 15-20 GeV/c in parton-level p_T at $\sqrt{s}$ = 500 GeV as a first trial; "proof–of–principle"**
- Total of 22,967 events
  - Trained with 11000 events
  - Tested with 11967 events

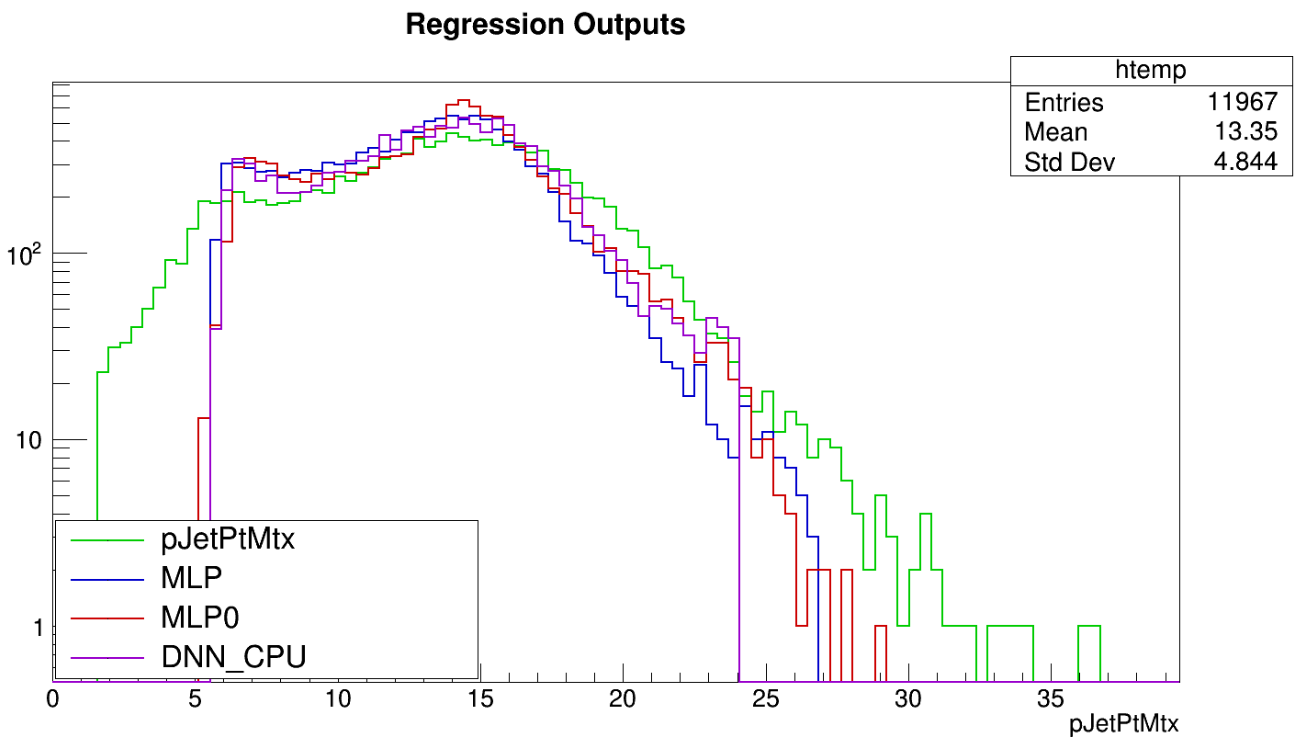
- **Note: Tight window of partonic p_T = not a continuous falling spectrum**
- <u>Regression target</u>
  - Particle jet p_T Matched to detector jet
  - Assures we get best match based on cuts
- <u>Input variables</u>
  - Detector jet p_T
  - Detector jet η







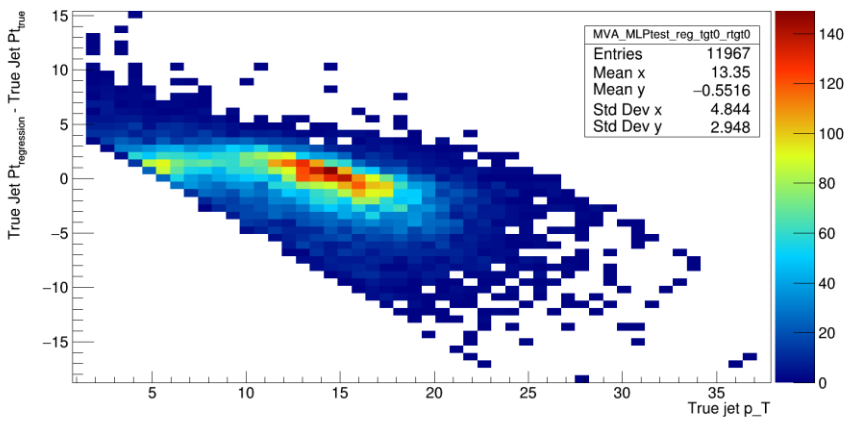
TMVA Regression : Output MLP, MLP0, DNN_CPU

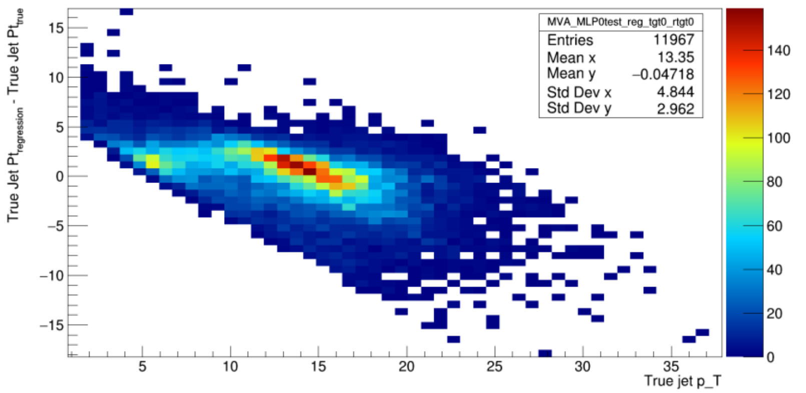


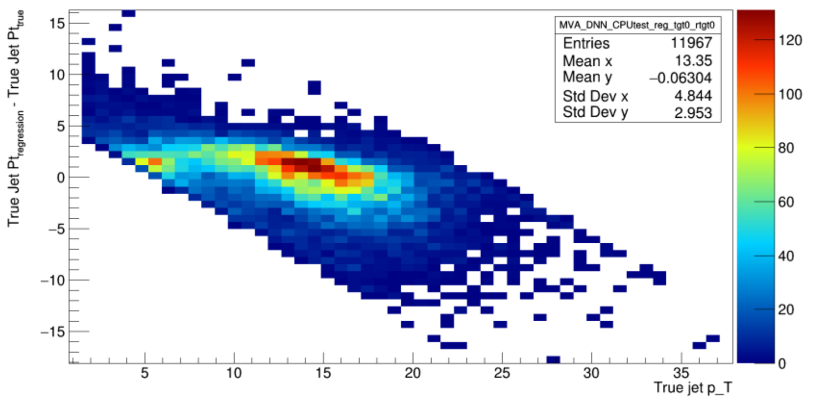
pJetPtMtx = Particle Jet p_T Matched to detector Jet
MLP = Multi-Layer Perceptron
MLP0 = Multi-Layer Perceptron 0
DNN_CPU = Deep Neural Network CPU

- **The long and winding regression…**
- The TMVA regression behaves well for Monte Carlo generated from partonic p_T level of 15-20 GeV/c
  - Small deviation from Monte Carlo from specified region
  - Sharp cut-off at 5 GeV/c for the regression from the 5 GeV/c threshold at the detector level

TMVA Regression : Output MLP, MLP0, DNN_CPU







- Utilized three different machine learning methods with TMVA
- **Important to note that all three different methods behave similarly towards the intended partonic bin of 15-20 GeV/c (small amounts of deviation)**
  - This could be improved on by, adjusting the neural network, trying different methods,…

Conclusion

<u>Conclusion</u>

- Need to optimize the variables for the TMVA input, e.g. including jet neutral energy fraction
- Modify the analysis infrastructure to handle the full set of Monte Carlo data
- Compare TMVA and RooUnfold
    - Compare systematic uncertainties between a 2D and 3D unfolding