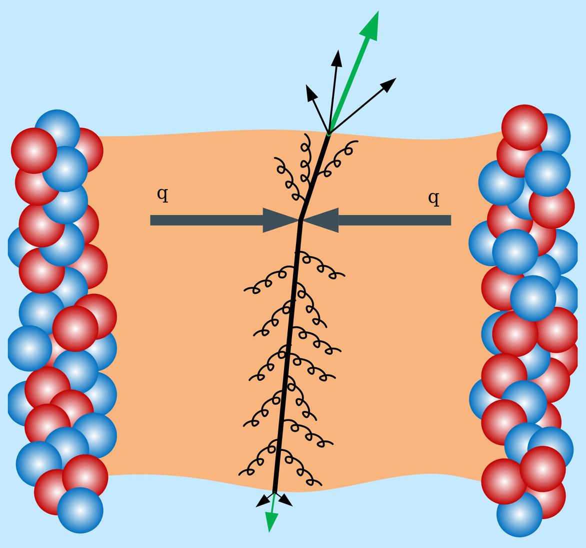
Machine Learning Based Jet p_T Reconstruction with Full Jets in ALICE

Hannah Bossi on behalf of the ALICE Collaboration

Quark Matter 2019 Wuhan, China

Jets and Heavy-Ion Collisions



Difficult to reconstruct jet p_T due to the large fluctuating background and combinatoric jets from the underlying event.

Area-Based Correction

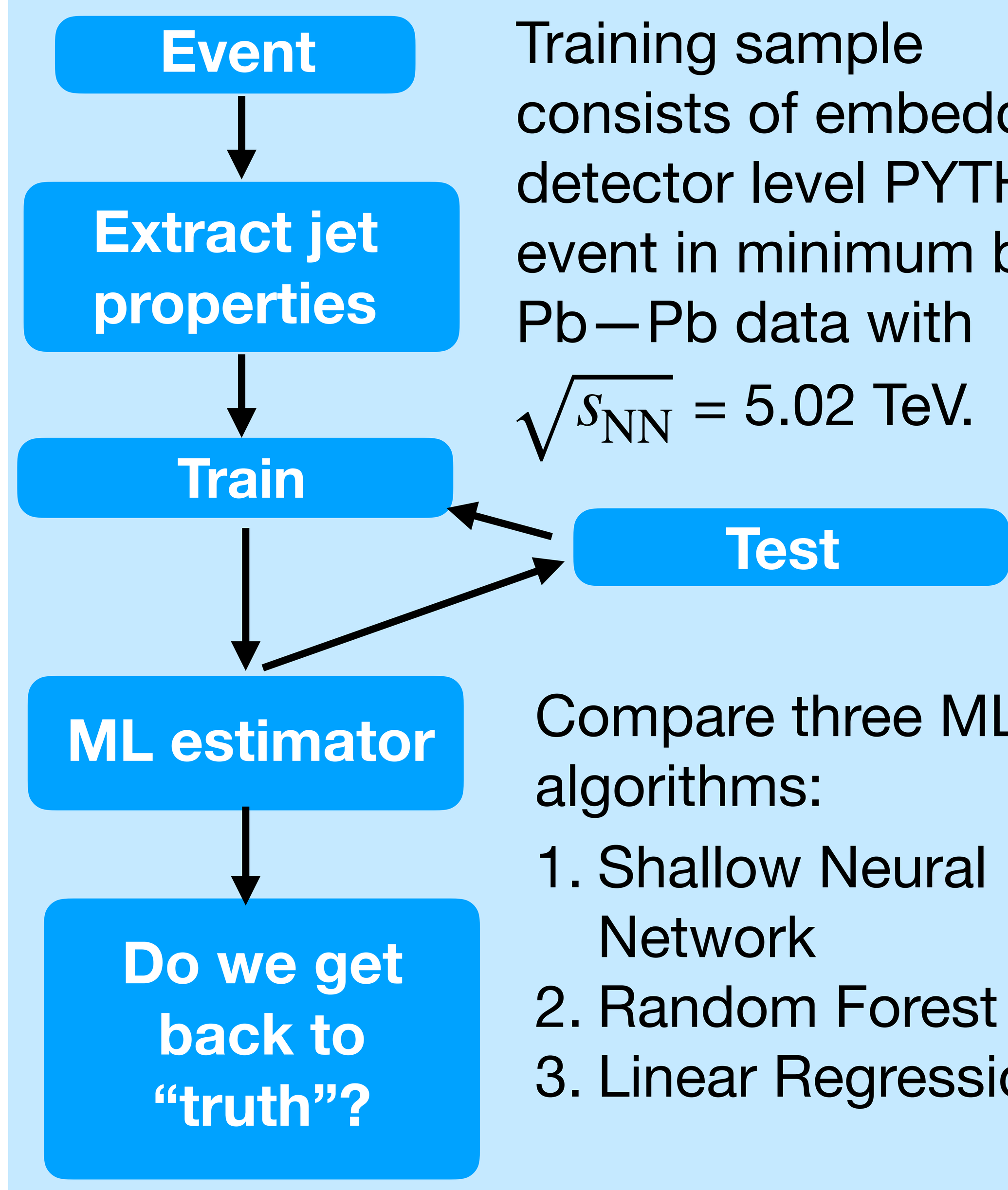
One method[1] is to correct for the average background on an *event-by-event* basis

$$p_{T,\,corr} = p_{T,\,raw} - \rho A$$

Following this pedestal subtraction, residual fluctuations remain. At low jet p_T residual fluctuations are on the order of the jet p_T itself, limiting measurements in this regime.

→**Can we do better?**

Process



Training sample consists of embedded detector level PYTHIA event in minimum bias Pb—Pb data with $\sqrt{s_{NN}}$ = 5.02 TeV.

Compare three ML algorithms:

1. Shallow Neural Network
2. Random Forest
3. Linear Regression

Geometrically matched PYTHIA jets serve as the target value of our ML estimator.

Input Parameters

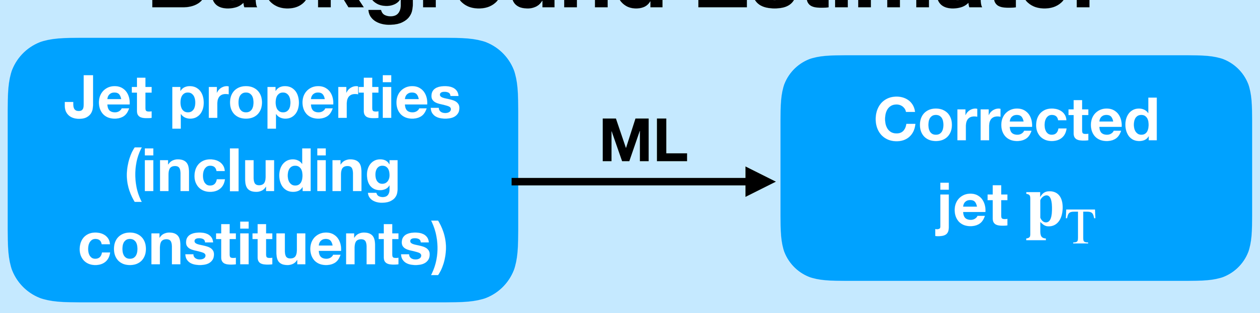
Feature	Importance	Feature	Importance
Jet p_T (area-based corr)	**0.7779**	$p^9_{T,track}$	0.0013
Number of tracks	0.0027	$p^{10}_{T,track}$	0.0015
Number of clusters	0.0007	$p^1_{T,cluster}$	**0.0266**
Mean of track p_T 's	**0.0775**	$p^2_{T,cluster}$	0.0074
Jet Angularity	0.0027	$p^3_{T,cluster}$	0.0021
$p^1_{T,track}$	0.0037	$p^4_{T,cluster}$	0.0010
$p^2_{T,track}$	**0.0664**	$p^5_{T,cluster}$	0.0010
$p^3_{T,track}$	**0.0160**	$p^6_{T,cluster}$	0.0009
$p^4_{T,track}$	0.0021	$p^7_{T,cluster}$	0.0006
$p^5_{T,track}$	0.0017	$p^8_{T,cluster}$	0.0006
$p^6_{T,track}$	0.0016	$p^9_{T,cluster}$	0.0006
$p^7_{T,track}$	0.0013	$p^{10}_{T,cluster}$	0.0006
$p^8_{T,track}$	0.0013		

Iteratively remove least important feature and check model performance.

When two or more features were highly correlated, keep the feature with the highest feature importance.

Want a simple and robust estimator!

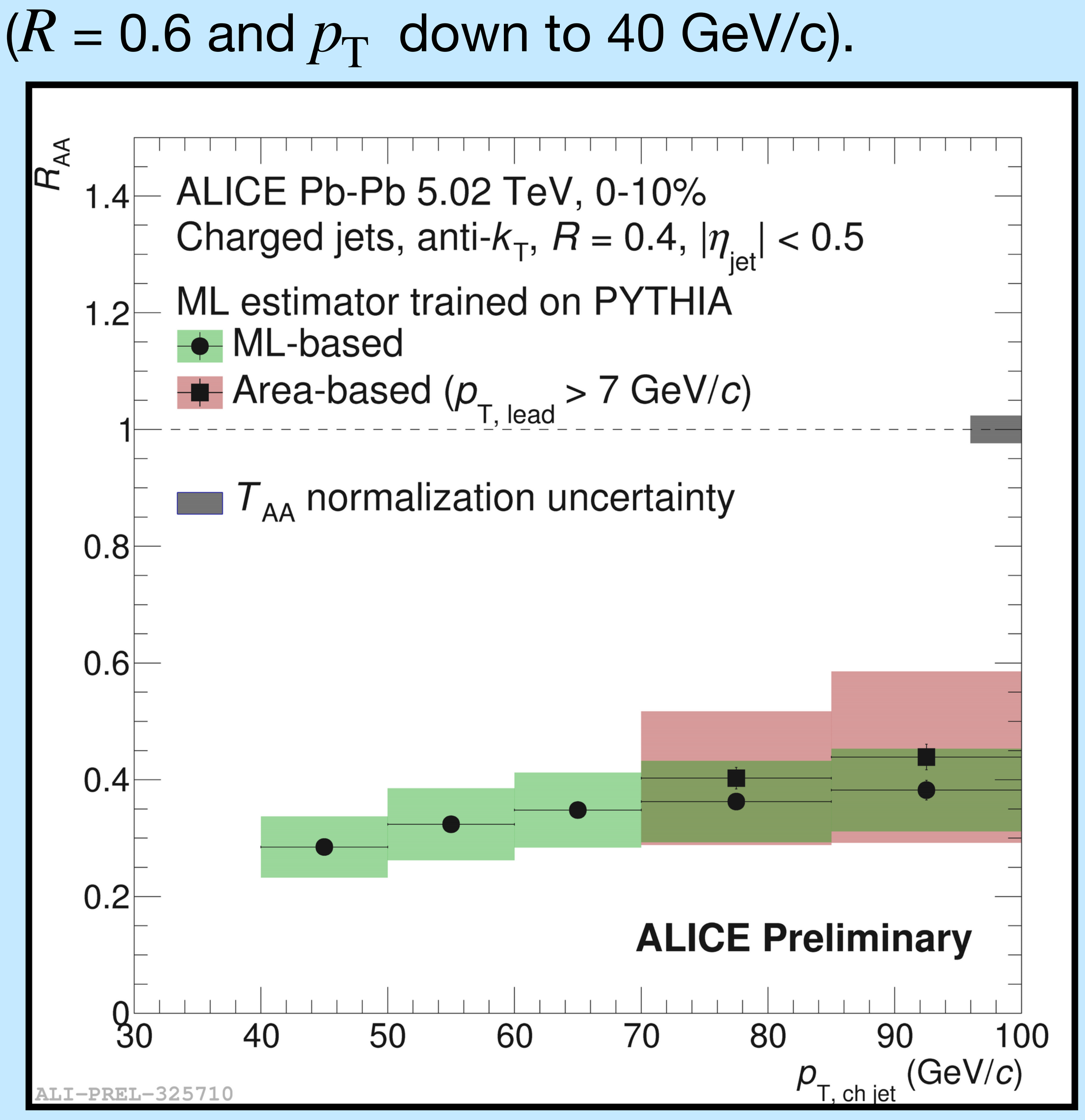
Machine Learning (ML) Based Background Estimator



ML corrects for the background on a ***jet-by jet*** basis by exploiting differences between each individual jet and the background particles which overlay it.

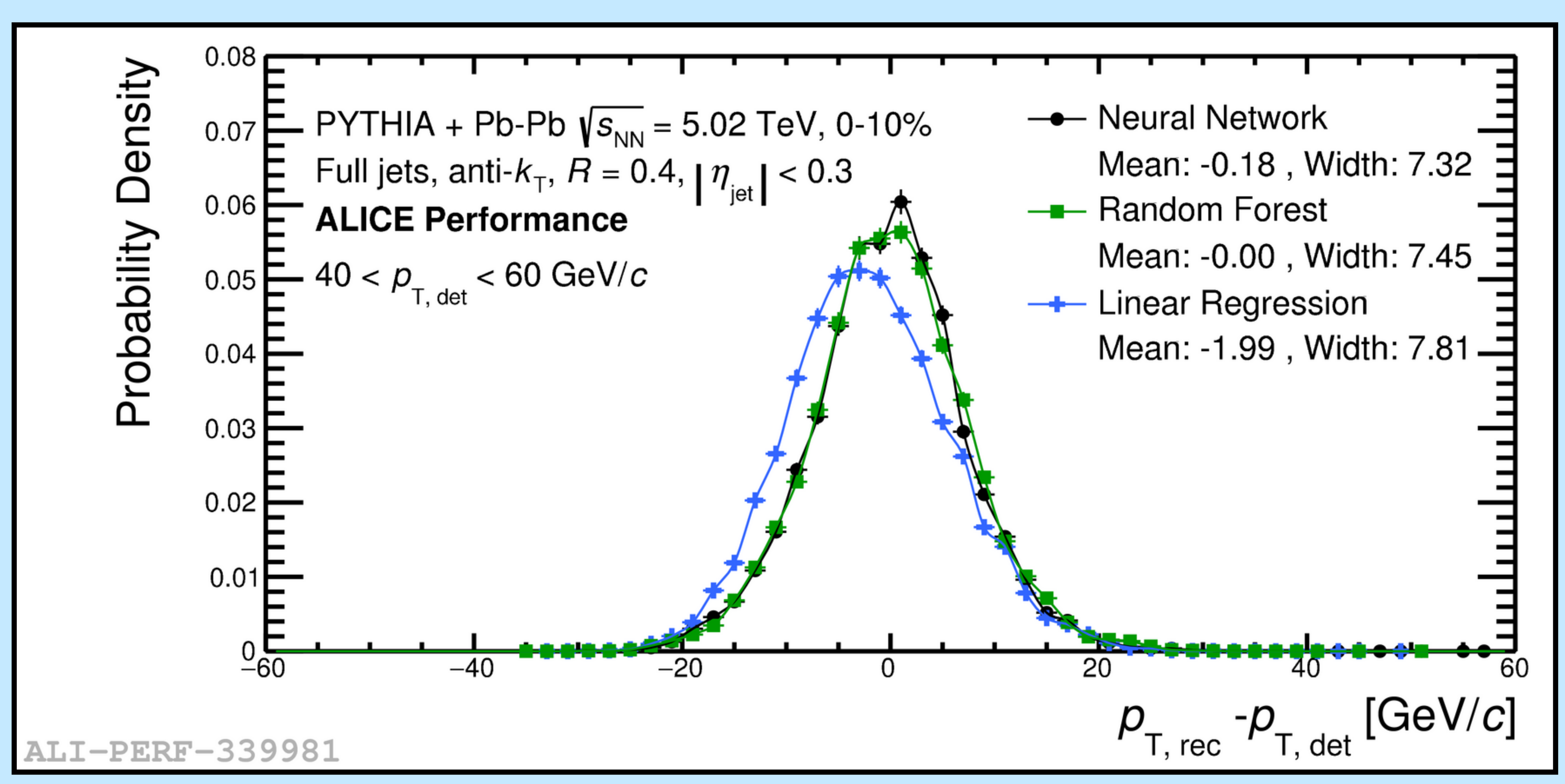
Reduced residual fluctuations and combinatoric background makes it easier to unfold.

Method[2] applied in ALICE to track based jets achieved unprecedented low p_T and R reach (R = 0.6 and p_T down to 40 GeV/c).



Extending to full jets aligns more with the traditional definition of a jet.

Results

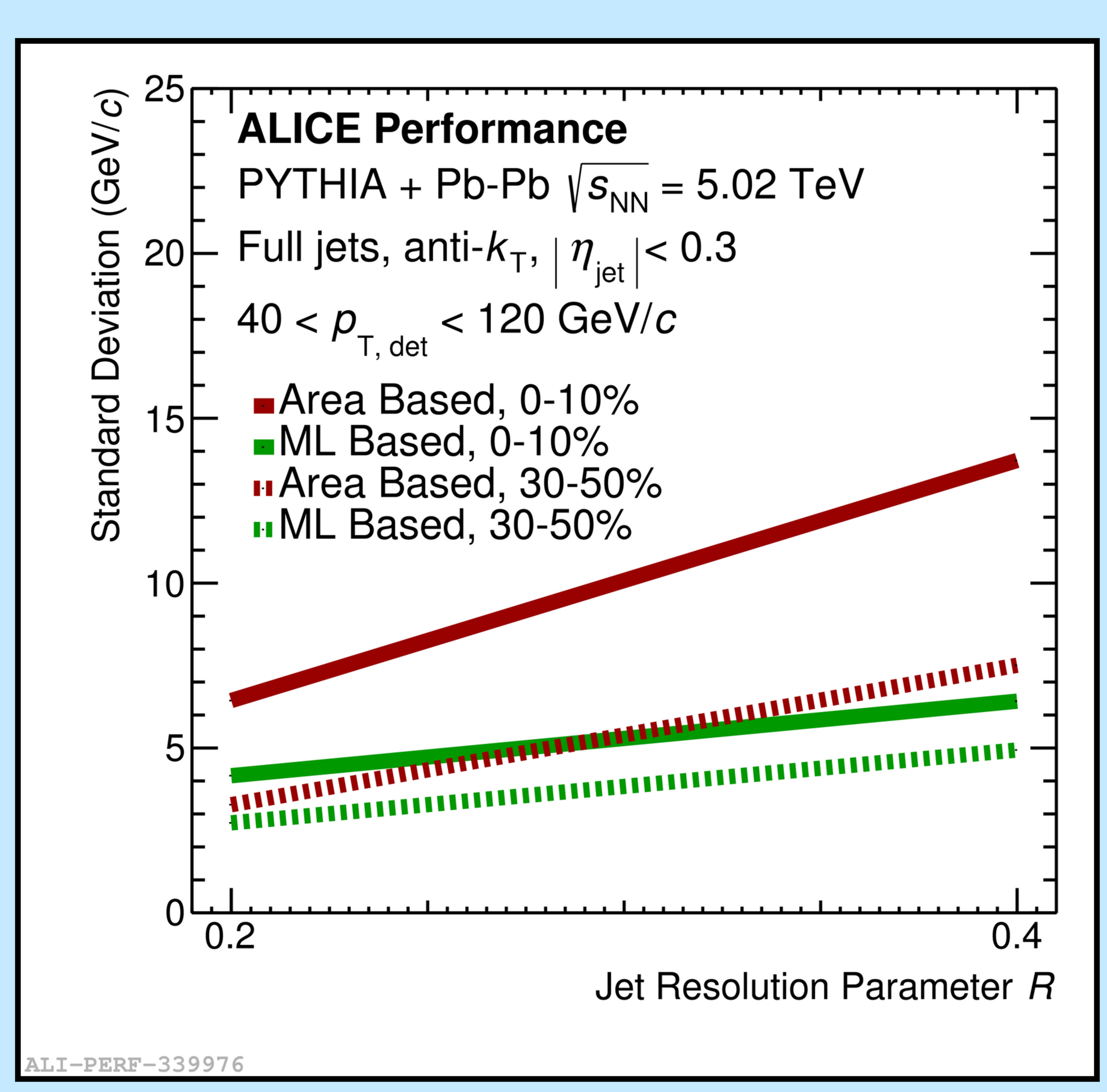


Width in δp_T distributions corresponds to the resolution of our background estimator.

Neural network performs the best amongst different ML algorithms used.

Largest gain occurs in central collisions for large jet resolution parameters.

Significant improvement in resolution over the area based method (Width = 14.70)!



Outlook

Further test fragmentation dependence of ML estimator.

→Use JEWEL or some other model.

Apply the method to make a full jet measurement!

Should allow for jet measurements to lower p_T and larger jet resolution parameters (enables comparisons to jet measurements at RHIC).

Goal: Gain information about parton energy loss in the QGP.

Wright Laboratory

U.S. DEPARTMENT OF ENERGY

References:

1. M.Cacciari and G.P. Salam, "Pileup subtraction using jet areas," *Phys. Lett.* **B659** (2008) 119-126, arXiv:0707.1378 [hep-ph].
2. R. Haake, C.Loizides (2018) Machine Learning based jet momentum reconstruction in heavy-ion collisions, arXiv:1810.06324[nucl-ex]

This work was supported by the US DOE grant number DE-SC004168

hannah.bossi@yale.edu