Machine Learning Based Jet p_{T} Reconstruction in heavy-ion collisions with ALICE

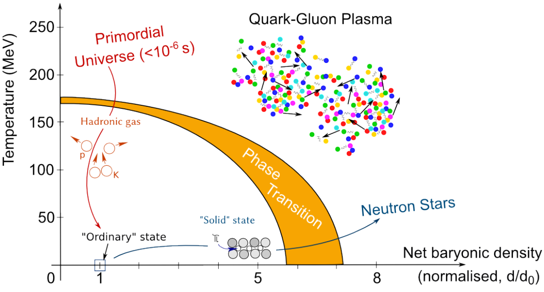
Hannah Bossi on behalf of the ALICE Collaboration

Yale University

ML4Jets 2020
1/17/2020
New York University

Funded by the U.S. DEPARTMENT OF ENERGY

Heavy-Ion Collisions and the QGP

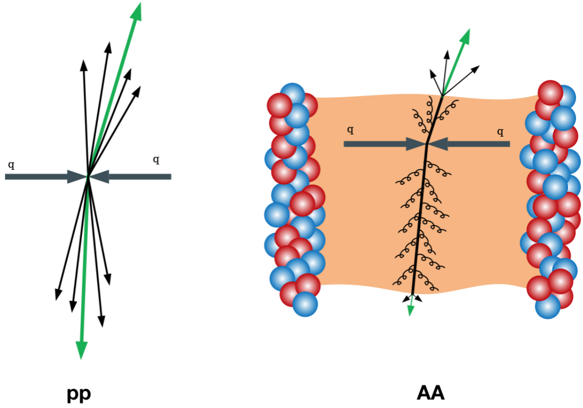


Phase diagram of strongly interacting matter.

- At extremely high temperatures and pressures, QCD matter becomes deconfined in a state referred to as the Quark-Gluon Plasma (QGP).
- These extreme conditions are reproduced in heavy-ion collisions.

- ALICE is optimized for such studies, measuring particles produced in collisions of Pb – Pb ions at $\sqrt{s_{\mathrm{NN}}}$ = 2.76 and 5.02 TeV.

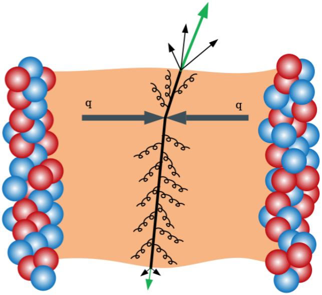
Jets as a Probe of the QGP

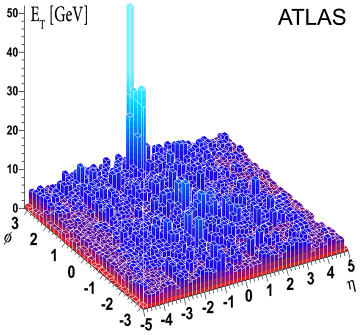


- Partons which make up jets are formed early in the collision before QGP formation.
- Expected to lose energy through interactions with the colored medium.
- Energy loss alters fragmentation.

- **We can use jets to probe the QGP!**

Challenges of Reconstructing the Jet p_{T}

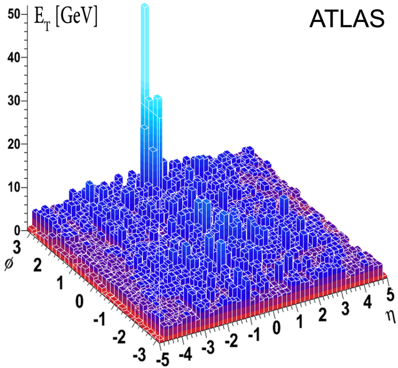




- Reconstructing the jet p_{T} is a challenging task in heavy-ion collisions due to the large uncorrelated background.

Foka, Panagiota, Janik, Małgorzata. (2016). An overview of experimental results from ultra-relativistic heavy-ion collisions at the CERN LHC: Hard probes. Reviews in Physics. 1. 10.1016/j.revip.2016.11.001.

Area Based Background Subtraction



Fluctuations $\sim$ 20 GeV/*c in R = 0.4,*
Central collisions $\rightarrow \rho A \sim$ 100 GeV/*c*

- Can we do better?

- ALICE standard is to correct jets for the average momentum density, ρ, calculated on an *event-by-event* basis.

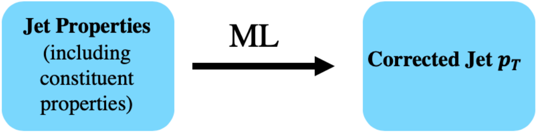
$$p_{\mathrm{T,corr}} = p_{\mathrm{T,raw}} - \rho A$$

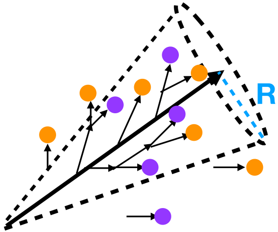
- Following this pedestal subtraction, residual fluctuations remain, causing problems for low p_{T} jets.
- *Possible problem for HL-LHC!*

Foka, Panagiota, Janik, Małgorzata. (2016). An overview of experimental results from ultra-relativistic heavy-ion collisions at the CERN LHC: Hard probes. Reviews in Physics. 1. 10.1016/j.revip.2016.11.001.

Machine Learning Based Background Estimator

- We use machine learning to create a data driven mapping that corrects for the background on a ***jet-by-jet basis***.



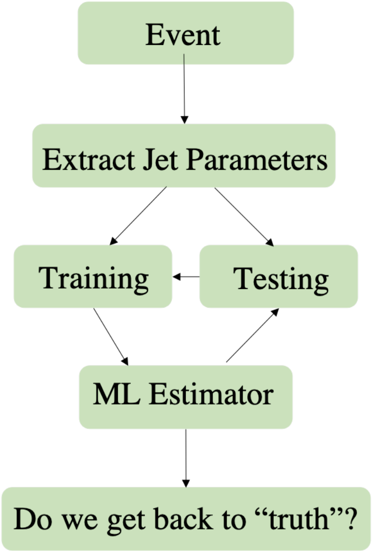


anti-k_T jets with various resolution parameters

- Exploit the difference of *each individual jet* and the background particles which overlay it.
- Aim is to reduce residual fluctuations, allowing for a better determination of the jet signal.
- Method can be applied to charged jets or full jets (which contain charged tracks and neutral clusters, measured in the TPC and EMCal, respectively).

Haake, Rüdiger, Loizides, Constantin. (2019). Machine-learning-based jet momentum reconstruction in heavy-ion collisions. Physical Review C. 99. 10.1103/PhysRevC.99.064904.

Process



- To create a suitable event for training, we embed a pp detector level PYTHIA event (truth) into a realistic background.
- Two options for the background
  1. Pb–Pb min bias data (more realistic)
  2. Simulated thermal heavy-ion background (easier to vary)
- Both yield quantitatively similar results!

ML Configurations

- Regression task → *predicting jet p_T*!
- We are prioritizing a **simple model!**
- Training is 10% of sample, testing 90%.
- Implemented in *scikit-learn*. Default parameters used unless otherwise specified.

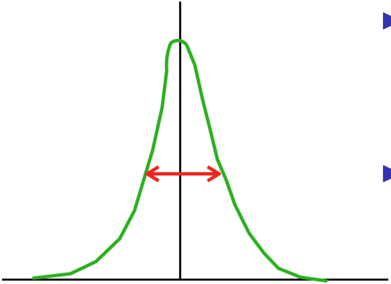
1. Shallow Neural Network
   - Shallow, three-layer network with [100,100,50] nodes.
   - ADAM optimizer, stochastic gradient descent algorithm.
   - Nodes/neurons are activated by a ReLU activation function.
2. Linear Regression
   - Normalization set to true by default.
3. Random Forest
   - Ensemble of 30 Decision Trees.
   - Maximum number of features set to 15.

Input Parameter Selection

- Ask two questions before selecting a feature
  1. How correlated is the feature with other features in the model?
  2. How important is the feature to the model's performance?

- Iteratively remove unimportant or highly correlated features (*ex: Uncorrected Jet p_{T} and area-based corrected jet p_{T}*).
- Simplified Input Parameters (charged jets): area-based corrected jet p_{T}, jet angularity, p_{T} of 8 leading tracks, number of constituents

How do we evaluate the performance?

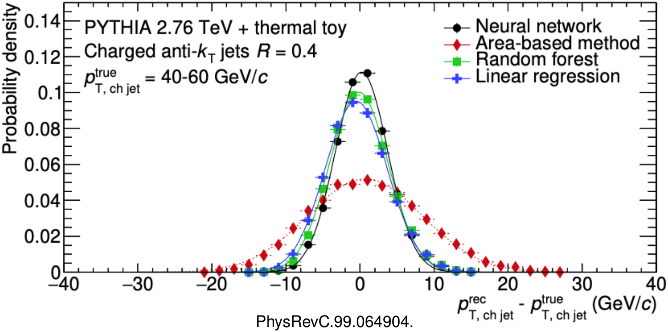


$\delta p_{\mathrm{T}} = p_{\mathrm{T,rec}} - p_{\mathrm{T,True}}$

- When we evaluate the performance of the ML based estimator, we are looking at the resolution.
- The narrower the peak in δp_{T}, the better the resolution of the background estimator.

- Residuals answer the question: Are we getting back to the true jet p_{T}?

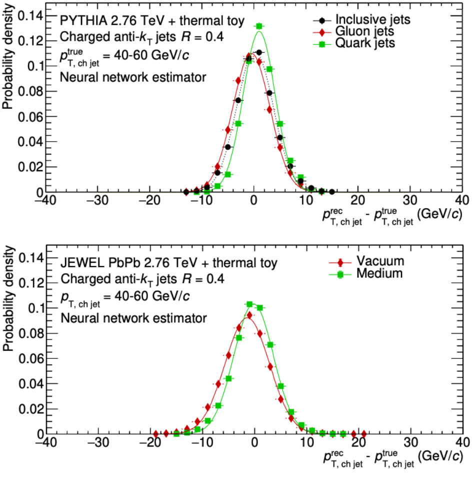
Model Performance (Charged Jets)



PhysRevC.99.064904.

- We see that ML methods show an increased performance over the area-based method!
- Different ML methods demonstrate a comparable performance, use neural network.

Investigations of Fragmentation Dependence



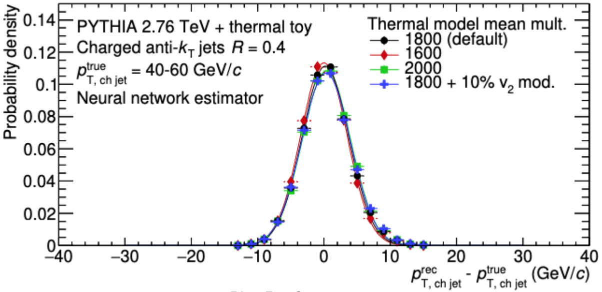
PhysRevC.99.064904.

- Introduced by learning from constituents.
- Investigate by checking model performance on two samples of jets with different fragmentation.
- Extreme variation: Quark vs. Gluon Jets
- Use JEWEL to test a fragmentation variation similar to that in heavy-ion collisions.

- A small bias observed.

Background Dependence

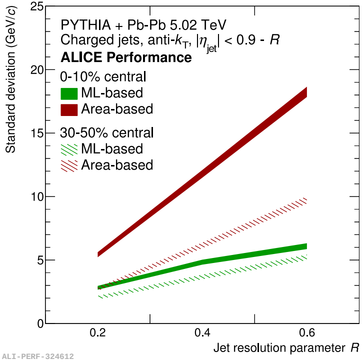
- We want to also look at the model dependence on the background used in training.
- We can test the effect of varying the mean multiplicity of the background using the toy model.



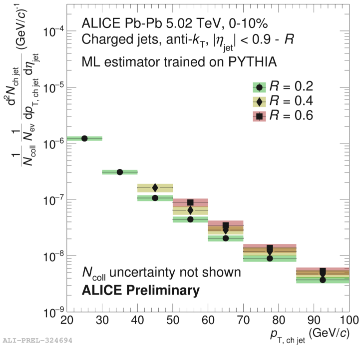
PhysRevC.99.064904.

- Model is robust to variations in the background!

Inclusive Charged Jet Spectra

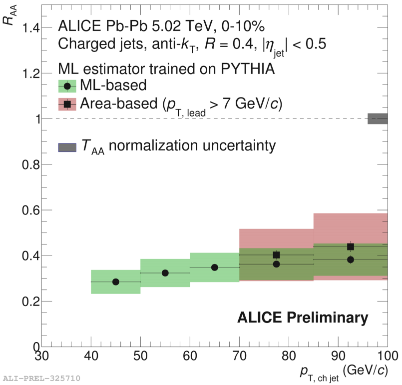


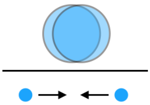
Central (0-10%) **Semi-Central (30-50%)**



- ML effectively reduces residual fluctuations, allowing us to greatly extend measurements.
- $R = 0.6$ is now possible!

Nuclear Modification Factor



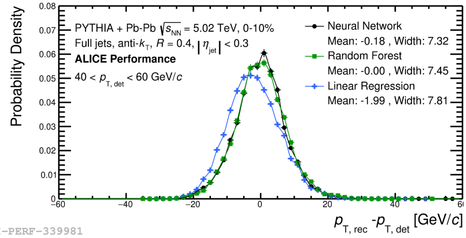


$$R_{AA} = \frac{\text{[Pb-Pb]}}{\text{[pp]}} = \frac{\frac{1}{N_{event}} \frac{d^2 N_{jet}^{PbPb}}{dp_T dy}\Big|_{cent}}{\langle T_{AA} \rangle \frac{d^2 \sigma_{jet}^{pp}}{dp_T dy}}$$

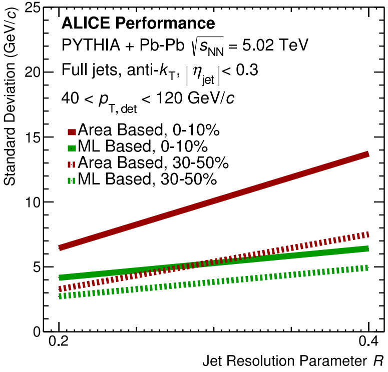
- ML agrees with area based results.
- Measurements over larger p_T ranges with smaller uncertainties.

ML-based correction provides opportunity to better understand jet quenching effects!

Adding in Full Jets



- Extension to full jets is advantageous, shows a greater alignment with the theoretical definition of a jet.



- Use a different set of input parameters, utilize neutral constituent information as well.
- ML demonstrates a similar improvement in full jets!!
- Measurements in progress!

Conclusions

- We introduce a novel method to reconstruct the jet p_{T} on a jet-by-jet basis using the properties of the jet and its constituents using machine learning techniques.
- This method shows a significantly improved performance over the area-based method.
  - Allows for measurements to lower p_{T} and larger R.
  - Can also compare to jet measurements at RHIC, *future home of the EIC*!!
- Full jet measurements in progress!
- Goal: Gain further information on parton energy loss in the QGP.

Backup

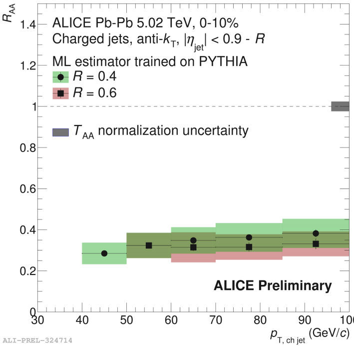
High Luminosity LHC

- At the High Lumiosity LHC, the average number of pileup interactions per bunch crossing will increase dramatically.
- Max Luminosity Run II: $\mu = 60$
- HL LHC: $\mu \approx 140 - 200$
- Area-median techniques still correct for the average background, but fluctuations are on the order of 10-15 GeV/*c*.
- This will cause issues for the measurement of low p_{T} jets!
- Could institute a low p_{T} cutoff, but many key measurements at the LHC (such as double Higgs production) rely on measurements to low p_{T}!

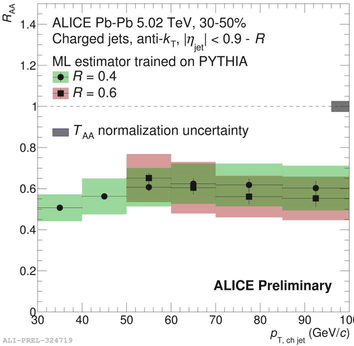
G. Soyez (2019) *Pileup mitigation at the LHC: A theorists view* (arXiv:1801.09721[hep-ph])

Nuclear Modification Factor, ML-Based Correction

- The nuclear modification factor, R_{AA}, is a ratio of the spectra in Pb–Pb collisions to the spectra expected if no QGP were present.



Central (0-10%)



Semi-Central (30-50%)

- ML-based correction allows for a greater understanding of jet quenching effects!

Regression Target

- We have two possibilities for the regression target which should represent the true jet p_{T}.

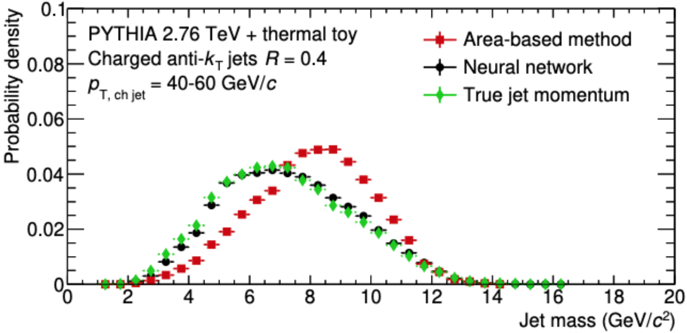
1. p_{T} of the Matched Detector Level Pythia Jet
   - Advantage: Has physical meaning.
   - Disadvantage: Adds another parameter, the matching radius, to the ML.
2. Use the true p_{T} fraction multiplied by the fully corrected jet p_{T}.
   - Advantage: No additional parameters.
   - Disadvantage: Difficult to exactly measure PYTHIA contribution to a hybrid cluster.

- Two regression targets yield similar results.

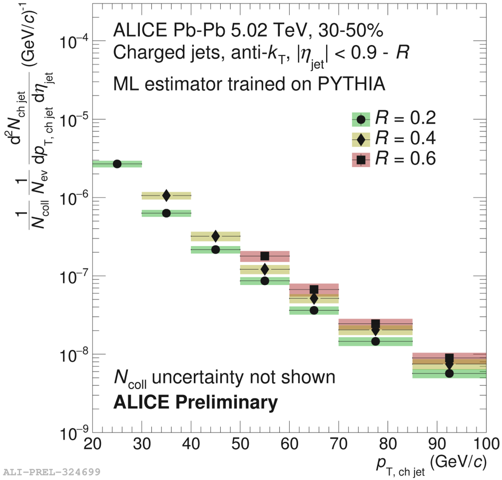
Feature Scores

Feature	Score	Feature	Score
Jet p_{T} (no corr.)	**0.1355**	$p^{1}_{\mathrm{T,\,const}}$	0.0012
Jet mass	0.0007	$p^{2}_{\mathrm{T,\,const}}$	**0.0039**
Jet area	0.0005	$p^{3}_{\mathrm{T,\,const}}$	0.0015
Jet p_{T} (area-based corr.)	**0.7876**	$p^{4}_{\mathrm{T,\,const}}$	0.0011
LeSub	0.0004	$p^{5}_{\mathrm{T,\,const}}$	0.0009
Radial moment	0.0005	$p^{6}_{\mathrm{T,\,const}}$	0.0009
Momentum dispersion	0.0007	$p^{7}_{\mathrm{T,\,const}}$	0.0008
Number of constituents	0.0008	$p^{8}_{\mathrm{T,\,const}}$	0.0007
Mean of const. p_{T}	**0.0585**	$p^{9}_{\mathrm{T,\,const}}$	0.0006
Median of const. p_{T}	0.0023	$p^{10}_{\mathrm{T,\,const}}$	0.0007

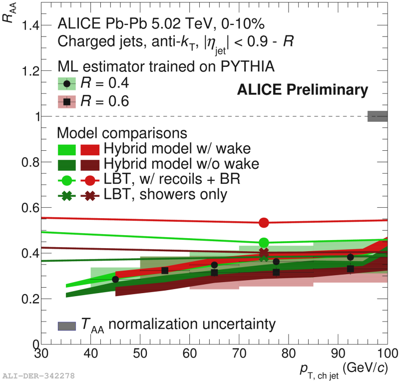
Jet Mass


PYTHIA 2.76 TeV + thermal toy
Charged anti-k_T jets $R = 0.4$
$p_{T,\text{ch jet}}$ = 40-60 GeV/c
Area-based method
Neural network
True jet momentum
Probability density
0.1
0.08
0.06
0.04
0.02
0
0
2
4
6
8
10
12
14
16
18
20
Jet mass (GeV/c^2)

Inclusive Charged Jet Spectra, Semi-Central



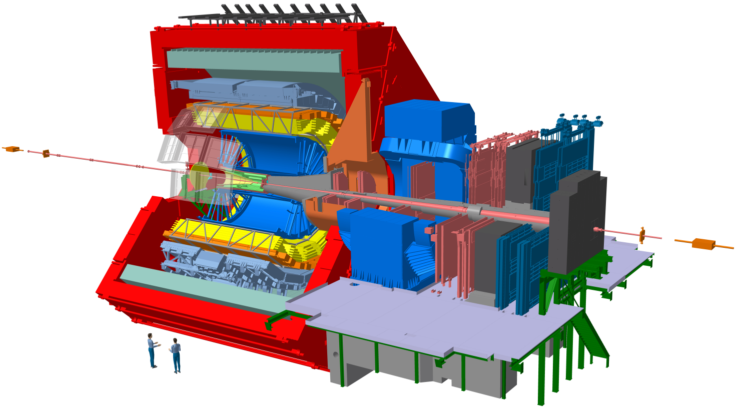
Model Comparisons

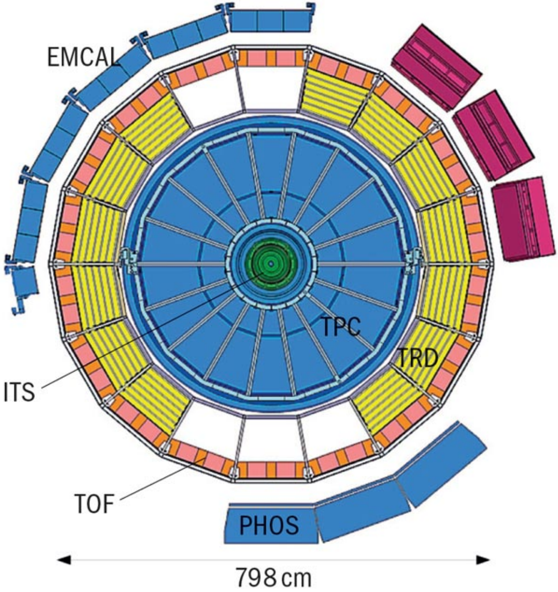


- Have to be careful with direct theory comparisons, ML algorithm is making a different set of ”cuts”.

Comparison of Algorithms Used

1. Neural Network
   - Utilize a series of nodes and layers, each with different weighting functions.
   - Each node carries a feature of the input data which is connected to every other node.
   - These nodes can either excite or inibit each other, allowing for complex relations.
2. Linear Regression
   - Creates a linear mapping between independent and dependent variables.
   - Iteratively tries to find the best fitting line which minimizes error.
   - Can use a gradient descent algorithm to find the best fitting line.
3. Random Forest
   - Uses an ensemble of decision trees, taking the mean prediction for the case of regression.
   - Possible to retrieve Gini index weightings or to look at individual decision trees $\rightarrow$ sheds light on the black box.




EMCAL
TPC
TRD
ITS
TOF
PHOS
798 cm