Jet characterization in Heavy Ion Collisions by QCD-Aware Graph Neural Networks

Yogesh Verma,[a,1] **Satyajit Jena,**[a,2]

[a]*Department of Physical Science, Indian Institute of Science Education and Research (IISER) Sector-81, Knowledge City, Mohali, India*

E-mail: ms16027@iisermohali.ac.in, sjena@iisermohali.ac.in

Abstract:

The identification of jets and their constituents is one of the key problems and challenging task in heavy ion experiments such as experiments at RHIC and LHC. The presence of huge background of soft particles pose a curse for jet finding techniques. The inabilities or lack of efficient techniques to filter out the background lead to a fake or combinatorial jet formation which may have an errorneous interpretation. In this article, we present Graph Reduction technique (GraphRed), a novel class of physics-aware and topology-based attention graph neural network built upon jet physics in heavy ion collisions. This approach directly works with the physical observables of variable-length set of final state particles on an event-by-event basis to find most likely jet-induced particles in an event. This technique demonstrate the robustness and applicability of this method for finding jet-induced particles and show that graph architectures are more efficient than previous frameworks. This technique exhibit foremost time a classifier working on particle-level in each heavy ion event produced at the LHC. We present the applicability and integration of the model with current jet finding algorithms such as FastJet.

[1]Corresponding author.

[2]Corresponding author.

Contents

1 Introduction

Jets are collimated sprays of particles in a narrow cone. Jets serves as the experimental signatures of Quark Gluon Plasma (QGP) as they are produced from fragmentation and hadronization of quarks and gluons given by Quantum Chromodynamics (QCD). An enormous amount of theoretical and experimental effort have been made around treatment of jets and to develop techniques that map theoretical predictions with experimental data while constraining some inherent properties. The main motivation for jet study is to yield a calibrated probe of perturbative QCD (pQCD) by providing measurements of observables with a cross-section which can be compared to its calculations. It is worth observing that many physical observables are responsive to non-perturbative effects like hadronization, thus affecting the measurements of numerous observables such as momentum spectra, transverse energy, flow etc.

Although, the jets are studied extensively in pp collisions for different physics searches, but the study of jets in heavy ion collisions is less explored due to the complexity of event structure, This may be due to the high multiplicity of particles present in heavy ion collisions as compared to pp collisions. When we move from pp collisions to AA collisions the choice to select final state particle particles for clustering into jet becomes more difficult. This is due to the dominated presence of large background of soft-QCD or hydro-dynamical parts originated from QGP in AA collisions. During the expansion of the QGP, there is a possibility that hard partons traverse through the medium which is expanding under the influence of temperature and density fluctuations. The mean transverse momentum of hadrons is around 680 MeV/c [1] giving an indication of formation of high p_T hadrons. However, it is not always convincing at what momentum regime the particle yield is dominated by jets rather than medium produced particles. Jets are modified by process such as strangeness enhancement or by hydro-dynamical flow and become quenched due to loss in energy when traversing through medium. Though, jet is dominated by high p_T particle but these low momentum particle carry critical information about the parent parton interaction with medium.

The interactions of the hard partons and its daughter particles complicate the selection procedure of what should belong to jet. This ambiguity makes the study of jets and interpretation of their result more challenging, rather qualitatively different from various other observables in characterizing heavy ion events. In addition, this gives rise to a finite contribution towards background to jets structure. The biggest source of background is due to collective flow of the particles originated from the QGP medium. Methods [2–5] which are employed to suppress and subtract the background from events depends on the nature of assumptions about the background which may change the sensitivity towards detection of medium modifications. These techniques may also bias the measured jet sample, for instance by selecting gluon jets at a higher rate than quark jets. For the jets in heavy ion collision, these analysis cuts serves as the definition of the jet and can not be disregarded. It is often assumed about the hidden interplay between hard and soft induced particles and about the distribution of background. Moreover, some observable are robust towards background assumptions but these are not always most responsive to energy loss mechanisms of jets happening with the medium.

In experiments, the final state particle are measured and observed which carries the traces of jets. The jets are determined using a recombination algorithm [6] from final state particles. These reconstructed jets are further used as the physics observable for characterizing the events shape and physics. These reconstruction techniques/algorithms are often mathematically challenging and computationally expensive tasks mainly due to the complex event structure. Such complex tasks can be tackle by using advanced computational techniques like imaging, machine learning, deep learning etc. In that context, numerous machine learning approaches have been developed upon the calorimeter-level considering them as images [7–14] and using recursive networks [15]. The graph networks have been proved to efficient in various task ranging from tracking [16] to jet finding [17]. The general approach of approach reduces the original jet classification to a image classification problem in the $\eta - \phi$ plane. Consequently, information loss occurs in discretization which may become a constraining factor in model performance. Most of these works deals with developing algorithms for tagging and characterizing jets in mostly pp collision events where backgrounds are comparatively smaller then AA collision. At the same time, when we scale pp collision to AA collision there is a subsequent change in the background distribution and signal (jets) distribution. In contrast to the pp collision there is a large background present in AA collision. These developed and stated algorithms have not been tested, applied and constructed in keeping the AA collision large background in consideration.

In this work, we propose instead a solution towards the large background problem prevalent in heavy ion collisions based on the physics aware and attention graph neural networks. Foremost, we propose to embed the full final state particle content in an event into a multi-dimensional topological space by geometrical embedding creating a web shaped graph. Now, this particle-level embedding can be fed into the graph neural network for particle-level classification into hydro and jet induced particles. This represent the first model operating on characterizing jets in heavy ion events in the lieu of large background distribution.

The paper is organized as follows: Section 2 describes the classification and prediction tasks at event level , 4 elucidates embedding of final state particles in topological space. Section 5 gives detail about model architecture and section 6 describes the results of our experiments on particle and jet-level and comparison with current frameworks. The paper concludes with Section 7.

2 Problem Statement

A Heavy Ion collision event is described by varying number of final state particles $N_{part} = [n_i | i = 1, 2...N]$ indexed by i, where each n_i is represented by its 4 momentum vector $v_i \in \mathbb{R}$. The 4-momenta of particles are clustered using sequential recombination jet algorithm like $anti - k_T$ [18] by recursively combining the particles that minimize

$$d_{ij} = min(1/p_{T,i}^2, 1/p_{T,j}^2)\frac{\Delta R_{ij}^2}{R^2} \tag{2.1}$$

The jets created by $anti - k_T$ algorithm are nearly symmetric in η & ϕ. Contrary to the pp collisions, AA collision suffer from a large background of soft particles originating from QGP. Therefore, it must be acknowledged that the jets reconstructed from a jet-finding algorithm may not be purely generated from hard processes thus making them no collinear safe. These will be fake or combinatorial jets and care must be taken to interpret the results from the jet finding algorithm. It may also happen that particles which are correlated through a hard process are grouped into jet candidates with background particles, hence, creating a difficult trade-off on how to search and characterize jets in heavy ion collisions.

This work builds on the above intersection by detecting the particles produced by hard processes in the event-level structure by incorporating the physics knowledge of hard and soft processes and their combined geometric structure with nearby particles giving topological information. In this framework, our goal is to build an intermediate model between experiment event-level output particles and jet finding algorithm to remove these fake and combinatorial jets.

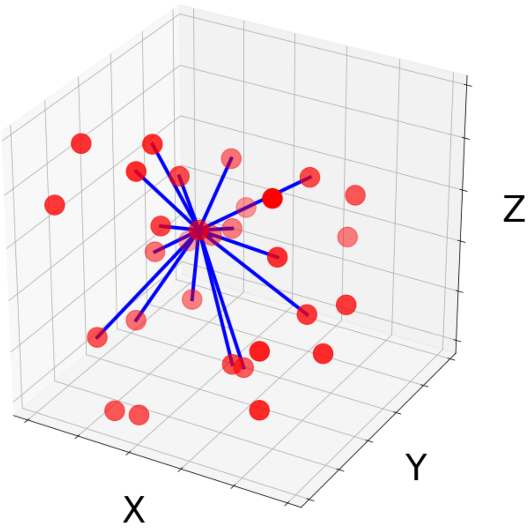
3 Dataset Generation

In order to focus attention on the generated jet particles and soft (hydro) particles, we used HYDJET++ [19] monte-carlo event generator for relativistic heavy ion collisions. The above model use PYTHIA [20] as a backend to generate parton shower and PYQUEN [21] model that modifies the "standard" jet event. The soft part is the thermal hadronic state generated on freeze-out hyper-surface on chemical and thermal freeze-out obtained from relativistic hydrodynamics.

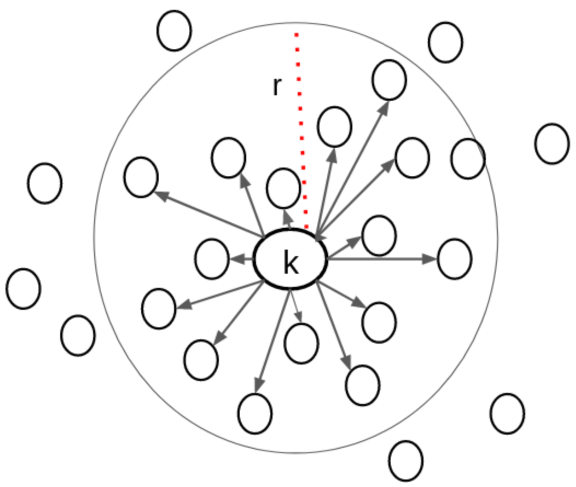
For this study, we generated $PbPb$ collision events at $\sqrt{s_{NN}} = 2.76$ TeV at $0 - 5\%$ centrality. The simulations provides us with various physical observable ranging from p_x, p_y to N_{jets}. A feature vector of 7 variables i.e. $(E, p_x, p_y, p_z, x, y, z)$ is chosen for each final state particle under $|\eta| < 0.5$ for geometrical embedding and then serving as a node features of the graph. The event generator provides us the explicit detail of all the particles arising from hydro part and the jet part which is used to train the model. Since, jets coming in heavy ion collisions consist of both quenched and non-quenched jets. We trained and evaluated our model on both quenched and non-quenched jets simulated data-sets explicitly so that it can also be extended for applications in pp collisions.

4 Geometrical Embedding

The main objective of executing geometrical embedding it to represent the particle level features in a multi-dimensional space to exploit the latent representations and hyper-dimensional connections between particles. Particle level physics observables are encoded into a hyper dimensional space where a graph is constructed over each particle to evaluate and its correlations are determined by the connections between particles. For each particle in an event, a web like graph is constructed as shown in Fig. 1 where edges (arrows) are the connections between individual particles and the nodes (circles) are the individual particles.



(a) XYZ system



(b) (η, ϕ) plane

Figure 1: (color online) Pictorial representation of web like graph construction for kth particle in (X, Y, Z) system and (η, ϕ) plane

The number of connections a particle can have is restricted by the r parameter which is a measure of distance between each particle in (η, ϕ) plane. The mathematical expression is

$$r_{ij} = \sqrt{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2} \tag{4.1}$$

Each node consist of particle-level feature vector of energy (E), momentum vector (p_x, p_y, p_z), position (x, y, z) which qualitatively transforms the embedding into a topological space of seven dimensions where each node has a definite projection in each plane and edges are connections in this multi dimension space. These features and connections are the input to our GraphRed model that which predicts the type of each particle by evaluating the connections between nearby particles in a multi-dimensional hyperspace.

5 GraphRed model

The GraphRed model is based on Attention based Graph Neural Network (AGNN) which proposes a modification to previously applied Message Passing Neural Networks (MPNN) described in Ref. [22, 23]. The attention based mechanism has been successfully used in various tasks such as machine translation [24, 25], machine reading [26] and many more [27]. An example of single graph attention layer is given in [27] where an arbitrary graph attention network is constructed and the layer computes the coefficient in attention mechanism for the node pair. They have also proved to outperform GCN [22] in several tasks, such as semi-supervised node classification, link prediction, and so on.

We have developed a new approach of application of AGNN for Jet finding. Once the geometrical embedding is executed and an input graph is created, the node features (multi dimensional vector) are encoded into a latent representation using a multi-Layer perceptron (Encoder Network) consisting of one fully connected layers having *tanh* activation function with trainable parameters (θ).

A recurrent set of {Node Network, Edge Network} iterating n_{iter} times is introduced after the *Encoder Network*. The latent outputs of the Encoder Network are passed in the Edge Network, where edge features are computed using the associated nodes features by fully connected network layers with *sigmoid* activation and a weight is computed for edges of the graph which is used as score to weight how the node features are aggregated. In this way, the graph is able to learn what attention it gives to each edge connected to each node. The edge score and features are input to the Node Network, where new node features are calculated and aggregated based on edge score, then it combines it with node's previous feature in a fully connected network to compute new features.

After n_{iter} iterations, the collective features of graph are passed through the Node Classification network which apply fully connected layers with *sigmoid* activation function to compute the final node scores. These node scores are interpreted as a probability of each node (particle) belonging to the hydro (soft) particle or the jet (hard) particle.

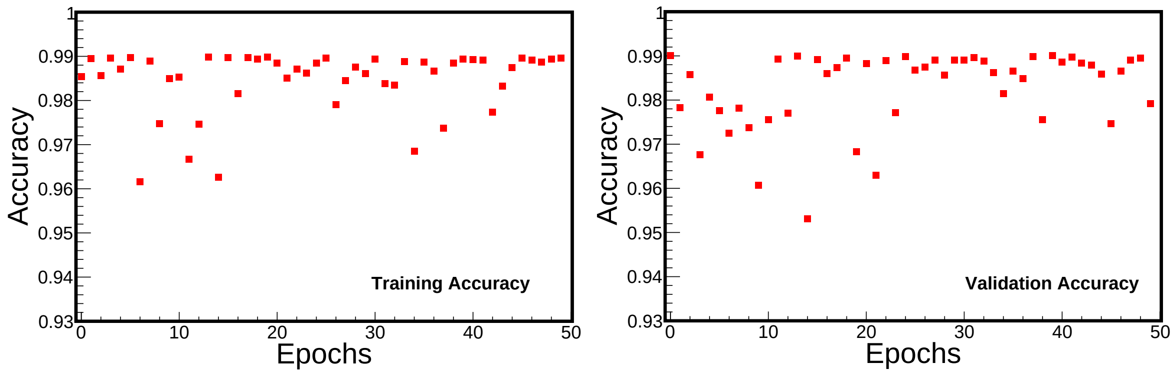
5.1 Node Classification

The GraphRed classification model described above performs binary classification of nodes of the graph to characterize the type of the particle. The graphs are constructed by geometrical embedding given in section 4. The model uses four graph recursions of network which is followed by a final classification layer that evaluates whether the node belongs to the soft or hard type.

Given the above architecture, we present the results of node classification to find jet (hard) particles. We present a hierarchical application of our graph network starting from particle level comparison to the jet level comparison.

6 Results

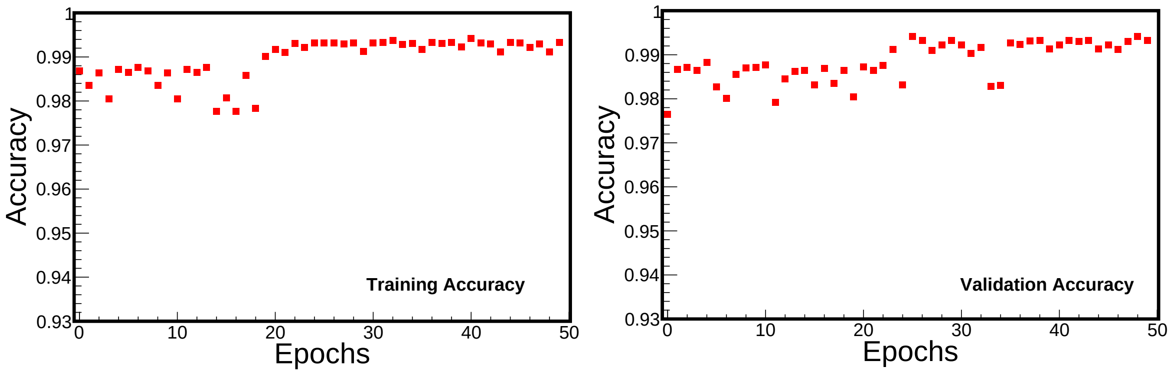
The model is trained over Quenched and Non-quenched dataset by embedding each event-level final state particles into its geometrical embedding collectively and then passing the whole event graph consisting of each particle geometrical embedding through the network for training. Training was performed for 50 epochs with training and validation set divided in a ratio of $80 : 20$ and value of r parameter was taken to be 0.6. This is due to the reason that if we consider high r value then there will be huge number of connections leading to computational intensive jobs and if we consider low r value then model will be not be able to find adequate features. The loss function used to train model is the binary cross entropy loss with ADAM optimizer. The training and validation accuracy for Quenched and Non-Quenched data-sets are shown in Fig. 2, 3.



(a) Training Accuracy (b) Validation Accuracy

Figure 2: (color online) Accuracy vs Epochs for Graph Neural Network for Non-Quenched Dataset

For both the data-sets, we achieve and accuracy $\sim$ 98-99% in finding soft and hard type particle. However, it is worthy to note that in an heavy ion collision the number of particles coming from hydro (soft) part are much more in number as compared to the jet (hard) part. Therefore, accuracy must not be considered as a sole measure of goodness of model. Hence, we have performed additional test for our model that will demonstrate its effectiveness in presence of large background.



(a) Training Accuracy

(b) Validation Accuracy

Figure 3: (color online) Accuracy vs Epochs for Graph Neural Network for Quenched Dataset

6.1 Particle-level Comparison

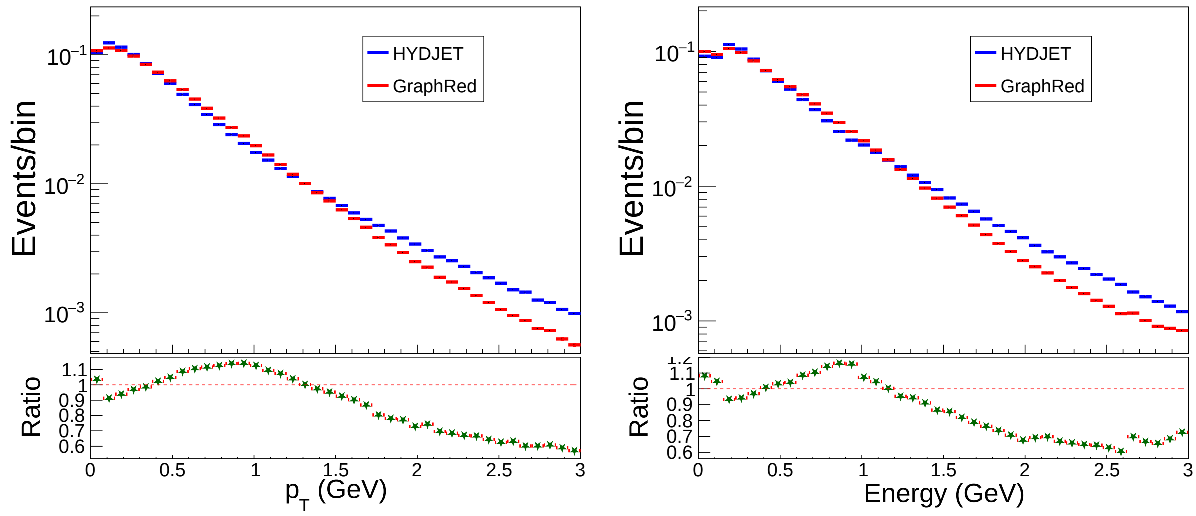
Accuracy is not the sole measure of the goodness of the model. Hence, we extend the evaluation of the model by comparison of various physical variables on particle-level like p_T spectra, Energy etc. We compared the particle-level physical observable distribution obtained from HYDJET++ (true) and our GraphRed (reconstructed) method. The method is evaluated for both Quenched Jets and Non-Quenched data-sets. The p_T and Energy distribution for Quenched Jets data-set is shown in Fig. 4 and for Non-Quenched Jets data-set in Fig. 5. A ratio plot of Data vs Reconstructed is also shown describing the effectiveness of the model.

We infer from the plots that the GraphRed is able to reconstruct the p_T and Energy distribution to a good extent which demonstrated the effectiveness in finding the most likely jet particles in an event. Moreover, it is worth noting that our proposed method is very much able to find the low p_T jet constituents as the ratio hovers around unity for low p_T range for both Quenched and Non-Quenched Datasets.

6.2 Jet-level Comparison

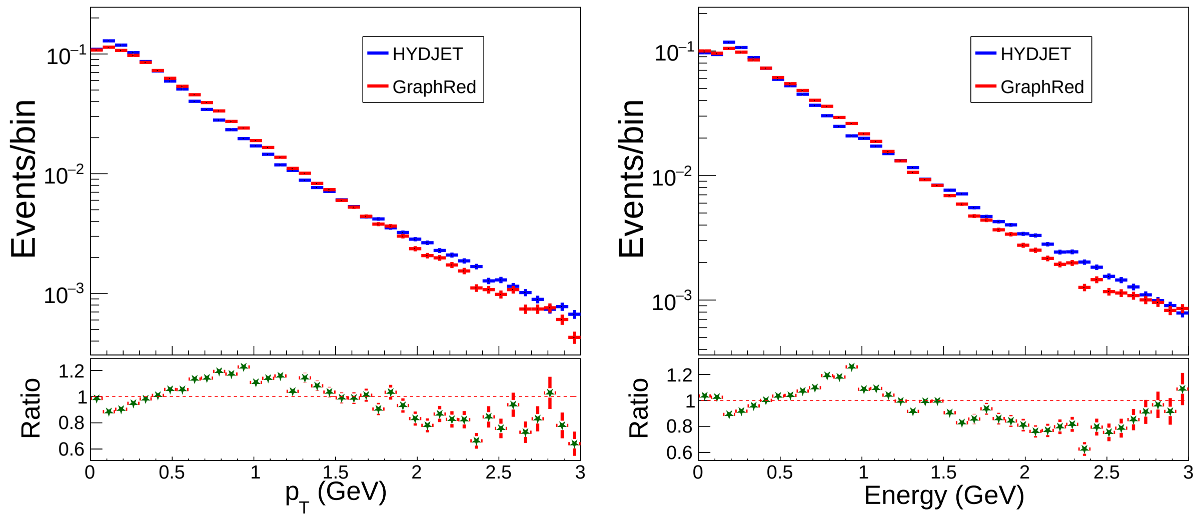
To demonstrate the effectiveness of our proposed method to find jet constituent particles, we used FastJet [28] to reconstruct jets from jet constituent particles by using anti-k_T algorithm in FastJet. We considered three cases to compare and contrast the our method from currently used frameworks. These are:

- The true Jet distribution from HYDJET++
- The Jet distribution obtained by passing HYDJET++ final state particles into FastJet
- The Jet distribution obtained by passing HYDJET++ final state particles into GraphRed and then passing GraphRed output into FastJet



(a) p_T Distribution (b) Energy Distribution

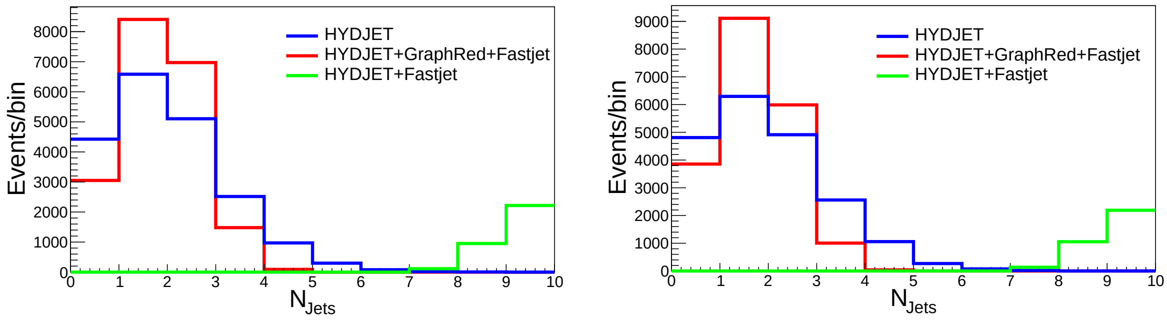
Figure 4: (color online) Distribution comparison between HYDJET++ (Data) jet constituent particles and GraphRed (Reco) most likely Non-Quenched jet constituent particles



(a) p_T Distribution (b) Energy Distribution

Figure 5: (color online) Distribution comparison between HYDJET++ (Data) jet constituent particles and GraphRed (Reco) most likely Quenched jet constituent particles

The comparison plots of number of jets identified by various cases is shown in Fig. 6 for the quenched and the non-quenched jet datasets. We infer that for the third case the FastJet is over-achieving by giving high number of jets in an event from the true distribution from HYDJET++. This is certainly due to fake or combinatorial jets and background particles which are grouped as jets by FastJet. However, the jet distribution obtained when using



(a) Non-quenched Jets

(b) Quenched Jets

Figure 6: (color online) Distribution comparison between Reconstructed jets for three cases described above.

our approach in intermediate steps is near to the true distribution of jets and able to remove the background picture and preserve the jet particle and its information.

7 Conclusion

In this article, we introduced GraphRed, a novel approach to find jet constituent particles in heavy ion collisions. The method is based on physics aware attention graph neural networks which manifest a significant improvement for jet finding and background reduction tasks in jet physics. Our studies have revealed that geometrical embedding and creation of web graph for each particle during the construction played an important role in performance of the model. By contrast, our model is able to work directly with the observables of variable length particle spectra, without information loss when considering them as discrete pixels in an image.

We showed that this method provides substantial improvements over existing frameworks when used in combination on the identification of key particles. The GraphRed models combined the power of physics with attention graph neural networks with an efficient geometrical embedding of the particle spectra. Our study indicates that this results in gain in terms of accuracy and background suppression as compared to traditional method. Moreover, for the first time we show a model working on particle-level exploiting the latent representation with the incorporation of domain knowledge from physics to segregate particles.

These results offers a solid direction for implementing essential machine learning based algorithms which are robust to the event by event fluctuations, model-dependent effects present in the data which is a notable application of machine learning at the LHC. The work presented in the article provides a nascent step towards a new paradigm of efficient, robust and tractable tools for jet physics in heavy ion collisions at LHC. In conclusion, we believe that there exists a great potential in advanced techniques which incorporate physics knowledge and exploit the power of artificial intelligence.

Acknowledgments

The simulation and training works were carried-out in the computing facility in EHEP Lab at IISER Mohali.

References

[1] ALICE Collaboration*. Multiplicity dependence of the average transverse momentum in pp, p-pb, and pb-pb collisions at the lhc. *Physics Letters B*, 727(4):371–380, 2013.

[2] Megan Connors, Christine Nattrass, Rosi Reed, and Sevil Salur. Jet measurements in heavy ion physics. *Reviews of Modern Physics*, 90(2), Jun 2018.

[3] B. Abelev and et al. Suppression of high transverse momentum d mesons in central pb-pb collisions at $\sqrt{s_{NN}}$=2.76 tev. *Journal of High Energy Physics*, 2012(9), Sep 2012.

[4] ATLAS Collaboration*. Measurement of the jet radius and transverse momentum dependence of inclusive jet suppression in pb-pb collisions at $\sqrt{s_{NN}}$=2.76 tev with the atlas detector. *Physics Letters B*, 719(4):220 – 241, 2013.

[5] Peter Berta, Martin Spousta, David W. Miller, and Rupert Leitner. Particle-level pileup subtraction for jets and jet shapes. *Journal of High Energy Physics*, 2014(6), Jun 2014.

[6] Gavin P. Salam. Towards jetography. *The European Physical Journal C*, 67(3-4):637–686, May 2010.

[7] Josh Cogan, Michael Kagan, Emanuel Strauss, and Ariel Schwarztman. Jet-images: computer vision inspired techniques for jet tagging. *Journal of High Energy Physics*, 2015(2), Feb 2015.

[8] Luke de Oliveira, Michael Kagan, Lester Mackey, Benjamin Nachman, and Ariel Schwartzman. Jet-images — deep learning edition. *Journal of High Energy Physics*, 2016(7), Jul 2016.

[9] Leandro G. Almeida, Mihailo Backovic, Mathieu Cliche, Seung J. Lee, and Maxim Perelstein. Playing tag with ann: Boosted top identification with pattern recognition, 2015.

[10] Pierre Baldi, Kevin Bauer, Clara Eng, Peter Sadowski, and Daniel Whiteson. Jet substructure classification in high-energy physics with deep neural networks. *Physical Review D*, 93(9), May 2016.

[11] Daniel Guest, Julian Collado, Pierre Baldi, Shih-Chieh Hsu, Gregor Urban, and Daniel Whiteson. Jet flavor classification in high-energy physics with deep neural networks. *Physical Review D*, 94(11), Dec 2016.

[12] James Barnard, Edmund Noel Dawe, Matthew J. Dolan, and Nina Rajcic. Parton shower uncertainties in jet substructure analyses with deep neural networks. *Physical Review D*, 95(1), Jan 2017.

[13] Patrick T. Komiske, Eric M. Metodiev, and Matthew D. Schwartz. Deep learning in color: towards automated quark/gluon jet discrimination. *Journal of High Energy Physics*, 2017(1), Jan 2017.

[14] Gregor Kasieczka, Tilman Plehn, Michael Russell, and Torben Schell. Deep-learning top taggers or the end of qcd? *Journal of High Energy Physics*, 2017(5), May 2017.

[15] Gilles Louppe, Kyunghyun Cho, Cyril Becot, and Kyle Cranmer. Qcd-aware recursive neural networks for jet physics. *Journal of High Energy Physics*, 2019(1), Jan 2019.

[16] Yogesh Verma and Satyajit Jena. Particle track reconstruction using geometric deep learning, 2020.

[17] Frédéric A. Dreyer and Huilin Qu. Jet tagging in the lund plane with graph networks, 2020.

[18] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. The anti-ktjet clustering algorithm. *Journal of High Energy Physics*, 2008(04):063–063, Apr 2008.

[19] I.P. Lokhtin, L.V. Malinina, S.V. Petrushanko, A.M. Snigirev, I. Arsene, and K. Tywoniuk. Heavy ion event generator HYDJET++ (HYDrodynamics plus JETs). *Comput. Phys. Commun.*, 180:779–799, 2009.

[20] Sjostrand et al. An introduction to pythia 8.2. *Computer Physics Communications*, 191:159–177, Jun 2015.

[21] Lokhtin, I. P. and Snigirev, A. M. A model of jet quenching in ultrarelativistic heavy ion collisionsand high-pt hadron spectra at rhic. *Eur. Phys. J. C*, 45(1):211–217, 2006.

[22] Thomas N. Kipf and Max Welling. Semi-supervised classification with graph convolutional networks, 2017.

[23] Nicholas Choma, Daniel Murnane, Xiangyang Ju, Paolo Calafiura, Sean Conlon, Steven Farrell, Prabhat, Giuseppe Cerati, Lindsey Gray, Thomas Klijnsma, Jim Kowalkowski, Panagiotis Spentzouris, Jean-Roch Vlimant, Maria Spiropulu, Adam Aurisano, Jeremy Hewes, Aristeidis Tsaris, Kazuhiro Terao, and Tracy Usher. Track seeding and labelling with embedded-space graph neural networks, 2020.

[24] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. Neural machine translation by jointly learning to align and translate, 2016.

[25] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need, 2017.

[26] Jianpeng Cheng, Li Dong, and Mirella Lapata. Long short-term memory-networks for machine reading, 2016.

[27] Petar Veli., Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. Graph attention networks, 2018.

[28] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. Fastjet user manual. *The European Physical Journal C*, 72(3), Mar 2012.