

Deep Learning for the Classification of Quenched Jets

R. Pedro^{a,*}

^a*LIP - Laboratório de Instrumentação e Física Experimental de Partículas,
Av. Professor Gama Pinto 2, 1649-003 Lisboa, Portugal*

E-mail: rute@lip.pt, rute.pedro@cern.ch

An important aspect of the study of Quark-Gluon Plasma (QGP) in ultra-relativistic collisions of heavy ions is the ability to identify a subset of jets that were strongly modified by the interaction with the QGP. In this talk, we will show how deep learning techniques can be applied for this purpose. Samples of Z+jet events were simulated in vacuum and medium and used to train deep neural networks with the objective of discriminating between medium- and vacuum-like jets. Dedicated Convolutional Neural Networks, Dense Neural Networks and Recurrent Neural Networks were developed and trained, and their performance will be shown. The results show the potential of these techniques for the identification of jet quenching effects induced by the presence of the QGP.

*** *Particles and Nuclei International Conference - PANIC2021* ***

*** *5 - 10 September, 2021* ***

*** *Online* ***

*Speaker

1. Introduction

In this work we explore Deep Learning (DL) as a basis tool for quenched jet classification able to distinguish strongly modified jets from unmodified ones [1]. Such tool would be interesting to obtain pure samples of modified jets, where the jet-QGP interaction could be studied in greater detail. A second axis of the work is to explore jet simulated data in different representations, both to take advantage of the variety of DL algorithms and input processing capabilities and of the different features of the substructure of quenched jets that may be naturally implicit in each representation. Finally, we investigate what have the DL models learnt as the difference between quenched and non-quenched jets with first-order techniques of model interpretation.

2. Data Simulation and Data Representations

We use samples of Z+jets generated with JEWEL v2.2.0 and hadronised with PYTHIA for vacuum (pp) and medium (PbPb) collisions at a centre-of-mass of $\sqrt{s_{NN}} = 5.02$ TeV. JEWEL accounts for medium-induced effects during the QCD parton shower evolution and is parametrised as detailed in. The Z+jet events are reconstructed by the $Z \rightarrow \mu\mu$ decay recoiling against a jet of radius $R = 0.5$ clustered with the anti- k_T algorithm using the FastJet package. The event selection criteria are: Z boson transverse momentum $p_{T,Z}^{min} > 90$ GeV and mass $m_Z \in [75; 105]$ GeV; jet transverse momentum $p_{T,jet}^{min} > 30$ GeV and pseudo-rapidity of $|\eta_{jet}| < 1.0$; and the azimuthal separation between the jet and the Z-boson is required to be $|\Delta\phi| = |\phi_Z - \phi_{jet}| > 7\pi/8$.

This system allows to use the Z boson as a clean probe to access the effects of the interaction of the jet with the QCD medium. While in vacuum there is a large correlation between $p_{T,jet}$ and $p_{T,Z}$, see left panel of Fig. 1, the energy loss resulting from jet quenching in the medium partially destroys the correlation between the boson and jet transverse momenta (right panel of Fig. 1).

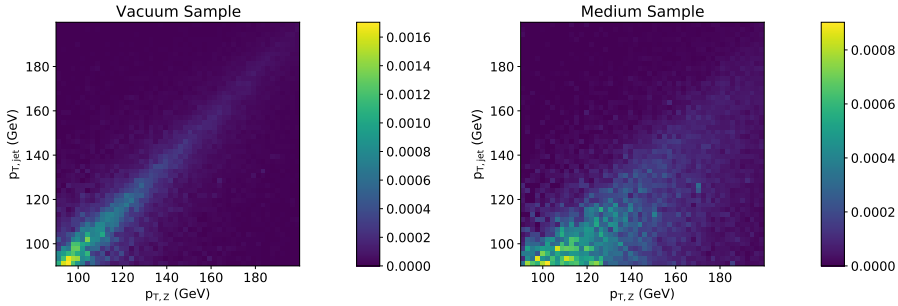


Figure 1: Bi-dimensional distributions of the Z-boson and jet transverse momenta, $p_{T,Z}$ and $p_{T,jet}$, for the JEWEL+PYTHIA (left) Vacuum and (right) Medium samples.

The aim of the work is to use DL techniques to develop a model for classification of quenched/non-quenched jets from supervised learning of the two example sets enclosed in the simulated samples of the Medium and Vacuum Z+jets simulation. DL is a versatile tool able to solve very different tasks, such as computer vision and natural language processing. Part of its flexibility comes from the ability to process data at different representation levels - images, sequences of words, tabular data, are a few examples. Here, we represented jets in three different formats to investigate what discriminating features of quenching effects are enhanced by each of them:

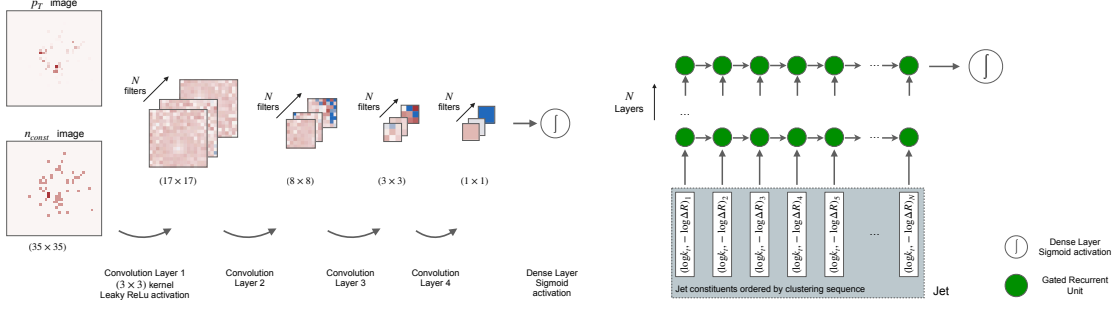


Figure 2: Diagram of the (left) Convolutional and (right) Recurrent Neural Networks used for jet classification from image representations and sequences of coordinates of the jet constituents in the Lund plane.

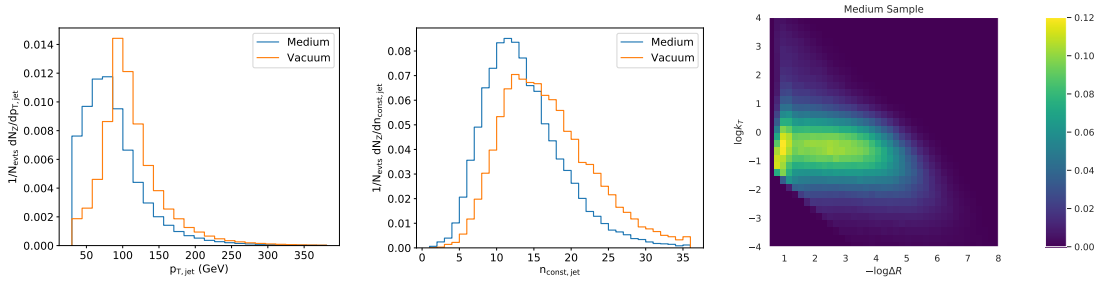


Figure 3: Distribution of the (left) $p_{T,jet}$ and (middle) number of jet constituents n_{const} for Vacuum and Medium. Representation of the jets in the primary Lund plane $(\log k_T, -\log \Delta R)$ for Medium.

- Jet Images: two $(\Delta\eta, \Delta\phi)$ grids centred in the jet axis with jet p_T and number n of constituents per pixel, see Fig. 2 (left). We define unnormalised and normalised images corresponding to the absolute jet information or relative to the total jet p_T and n , respectively;
- Lund-planes: $(\log k_T, \log \Delta R)$ coordinates of the jet constituents according to the Cambridge/Aachen (C/A) jet clustering sequence, see Fig. 3 (right);
- Tabular: total jet p_T and n constituents, see Fig. 3 (left), designed as the baseline to assess the gain with the jet substructure information implicit on Images and Lund planes.

The distributions of the total jet p_T and n constituents are shown in Fig. 3 (left) for the Medium and Vacuum samples. We see the distinctive signature of jet quenching by the QGP, with the modification of the momentum spectrum towards softer scales and suppression of constituents in jets simulated with the QCD medium. Fig. 3 (right) shows the average coordinates of jet constituents in the Lund plane for the Medium sample.

3. Deep Learning Models for the Classification of Quenched Jets

We train DL algorithms with the objective of classifying quenched and non-quenched jets using supervised learning on jets of the Medium and Vacuum samples. We highlight that the Medium sample is not pure on quenched jets, instead being composed of a set of jets modified by the QGP

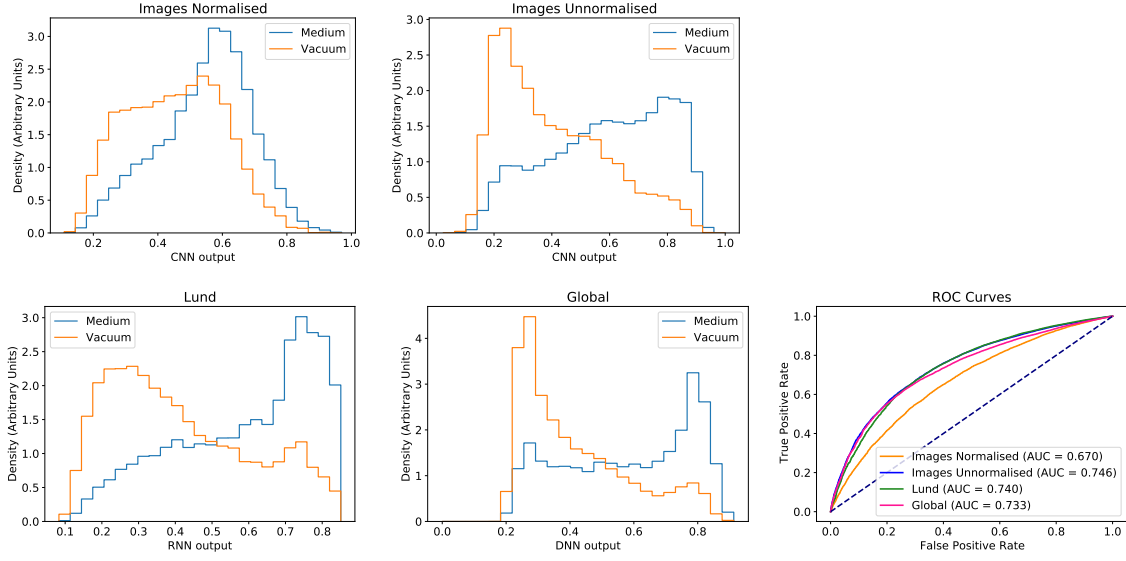


Figure 4: Distribution of the different DL outputs for the Vacuum and Medium samples and corresponding ROC curves for the separation of the two classes.

to very different extents. Despite this, the Medium *versus* Vacuum training will be effectively enhancing the underlying distinguishing quenching features.

We employ three DL models, each processing a specific jet data representation:

- Convolutional Neural Networks (CNNs) using jet Images as input. CNNs, see Fig. 2 (left), scan the images looking for successively detailed discriminant patterns in the image;
- Recurrent Neural Networks (RNNs) use Lund coordinates of the jet constituents as inputs, see Fig. 2 (right). RNNs are sensitive to causal ordering (eg. speech) and may exploit quenching effects present at the level of the C/A ordering sequence;
- Dense Neural Networks (DNNs) work on the tabular data.

To train and test the models, we split the data sample into a train, validation and test sets in 1:1:1 proportion. All the models were implemented in Keras and the free hyper parameters of the networks were optimised with a Bayesian loop within a user pre-defined range using Optuna. The DL discriminant is shown for the Medium and Vacuum test samples per model type in Fig. 4, along with the corresponding ROC curve for the two-class separation. We attain the best classification performance for CNNs trained on unnormalised images, where besides access to the jet absolute p_T and n constituents the network is effectively gaining from additional quenching effects in the jet substructure that are implicit in the image pattern. This is corroborated by the improvement of the performance relative to the DNN trained on tabular data. RNNs running over Lund coordinate sequences have also a good performance, indicating that both Lund plane representations and jet images bring in additional sensitivity to the quenching effects at the jet substructure level.

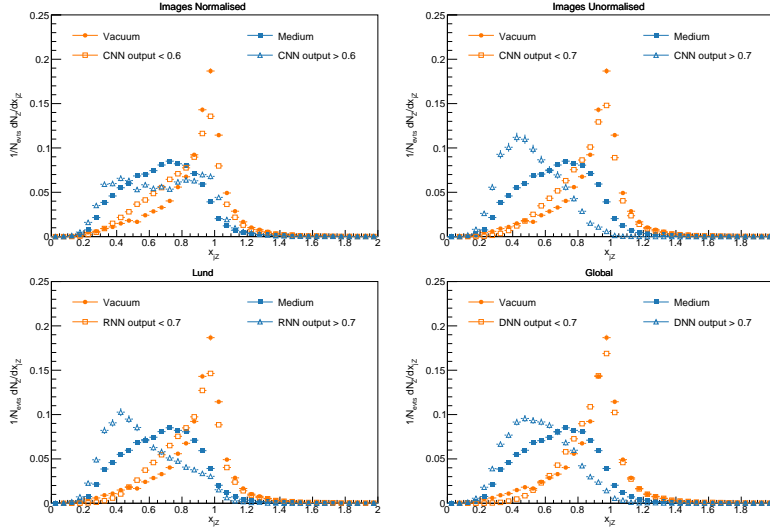


Figure 5: Distribution of x_{jZ} for Vacuum and Medium (solid symbols) and for samples of *vacuum-like* jets (open blue) and *medium-like* jets (open orange), selected by the different DL architectures.

4. Interpretation of the Deep Learning Models

To investigate the decision boundary of the DL models we selected two samples of *medium-like* and *vacuum-like* jets from the original Vacuum and Medium samples according to the output of each DL network - above or below a certain threshold, respectively. Fig. 5 shows the distribution of the p_T imbalance of the Z+jet system $x_{jZ} = p_{T,jet}/p_{T,Z}$ for the various samples. We observe that the *medium-like* selection shifts the x_{jZ} distribution towards smaller values where the Z+jet p_T imbalance is larger, a phenomenon associated to quenching. Thus, the DL classifiers are effectively able to enrich the sample with quenched jets.

5. Conclusions

We used DL to explore different representations of jets and build classifiers of quenched jets. There is an enhancement of the quenching features on distributions of *medium-like* jets selected with the DL quenching classifiers, indicating that the models' decision is triggered by actual quenching effects. In addition, we have found that classifiers based on Lund plane coordinates and jet images outperform discriminants from tabular data, concluding that these data formats encode additional information about jet fragmentation and that RNNs and CNNs are able to capture it.

References

- [1] Apolinário, L., Castro, N.F., Romão, M.C. et al. *Deep Learning for the classification of quenched jets*. J. High Energ. Phys. 2021, 219.