IL NUOVO CIMENTO **47 C** (2024) 127
DOI 10.1393/ncc/i2024-24127-7

Colloquia: IFAE 2023

Quantum Machine Learning for data analysis at LHCb(*)

A. Gianelle([1]), D. Lucchesi([1])([2]), S. Monaco([2]), D. Nicotra([3])([4]),
L. Sestini([1])([2]) and D. Zuliani([1])([2])

([1]) *INFN, Sezione di Padova - Padova, Italy*
([2]) *Dipartimento di Fisica, Università di Padova - Padova, Italy*
([3]) *Nikhef National Institute for Subatomic Physics - Amsterdam, The Netherlands*
([4]) *University of Maastricht - Maastricht, The Netherlands*

received 13 February 2024

Summary. — Machine learning (ML) algorithms have now become crucial in the field of High Energy Physics (HEP). An area where the application of such algorithms has proven particularly beneficial is the classification of hadronic jets produced at the Large Hadron Collider (LHC). Considering the complexity of the tasks in this field and the impending Run 3 data at higher luminosity, it is evident that a step-up in computational power is imperative. One potential candidate comes from the intersection between Quantum Computing (QC) and ML. Quantum Machine Learning (QML) algorithms leverage the intrinsic properties of QC, such as superposition and entanglement, to achieve better performance compared to their classical counterparts. This work provides an overview of these new learning models, with a focus in HEP. Specifically, we present studies of QML applications for the classification of jets produced (b *vs.* $\bar{b}$ and b *vs.* c) at the LHCb experiment. Notably, we discuss recent developments in measuring entanglement entropy between qubits to gain new insights from the jet events data.

1. – Heavy Jets at LHCb

Jets are narrow cones of particles resulting from the strong interactions between quarks in proton-proton collisions. These collisions, occurring at high energies, initiate a cascade of processes leading to the formation of collimated streams of particles.

The LHCb experiment stands out as a fitting tool for the study of jet formation, particularly those originating from heavy quarks. The unique geometry of LHCb in the forward region provides an advantageous and unique point of view for capturing events where heavy quark jets are relevant.

(*) IFAE 2023 - "Poster" session

Machine learning (ML) applications currently leverage the intricate structure and correlations among all reconstructed particles to accurately identify the flavor of the originating quark. These 'black box' implementations provide an ideal platform for this task, as finding an analytical algorithm that encompasses all detected particles is impossible.

The complexity of this task, prompts the introduction of a new computing framework in order to achieve a better performance. Quantum Machine Learning (QML) models emerge as a promising candidate. Given the intricate nature of jet formation and the wealth of information embedded in each collision event, quantum models can amplify the efficiency and precision of jet classification.

2. – Quantum Computing for Jets flavour identification

Quantum computers possess the capability to accurately evolve a quantum system comprising multiple quantum bits (*qubits*), gaining a computational advantage by harnessing quantum mechanical phenomena, including *entanglement*.

Operationally, we gain the ability to manipulate the wavefunction of the quantum system —a vector in the *Hilbert space*. Unlike a classical computer, which may be able to simulate the evolution of such a system up to ~ 20 qubits, an actual quantum implementation allows us to scale up to the order of hundreds and even thousands of qubits. This scalability enables us to explore a significantly larger solution space, providing the potential to capture more intricate relationships within the embedded input information.

QML derives its fundamental concepts from *classical* ML. For instance, a gradient-based technique is employed to adjust the parameters of the operators responsible for evolving the wavefunction. This procedure aims to reach a minimum of a loss function given by the specific task.

3. – Jets classification

In this section, we present the primary findings from two studies about *jet tagging* conducted by the LHCb Collaboration [1, 2].

3˙1. *b vs. $\bar{b}$ jet tagging.* – The first study delved into the development of quantum techniques for b *vs.* $\bar{b}$ jets classifications. It is notoriously a difficult task, and current models' efficiencies are affected by a high value of rejection rate.

Figure 1 shows a comparison in terms of tagging power $\epsilon_{\rm tag}$ between two different quantum models: "Angle Emb." and "Amplitude Emb.", as well as two classical methods: "Muon Tag" and Deep Neural Network (DNN).

3˙2. *b vs. c jet tagging.* – b *vs.* c classification is a relatively easier ask, in which models can better exploit the jet's substructure using the Secondary Vertex (SV) representation of the event. Figure 2 shows the bar-plots of the predictions of b-jets (marked in blue) and c-jets (marked in red) for the state-of-the-art classical architecture and a quantum model.

In both studies, the quantum methods showed slightly lower performances than their classical counterparts. It is crucial, however, to take into account that these novel methods are still in the early stages of development and face constraints in terms of system size (number of qubits). Once both software and hardware implementations are further developed, we expect to see the quantum models to show an advantage.

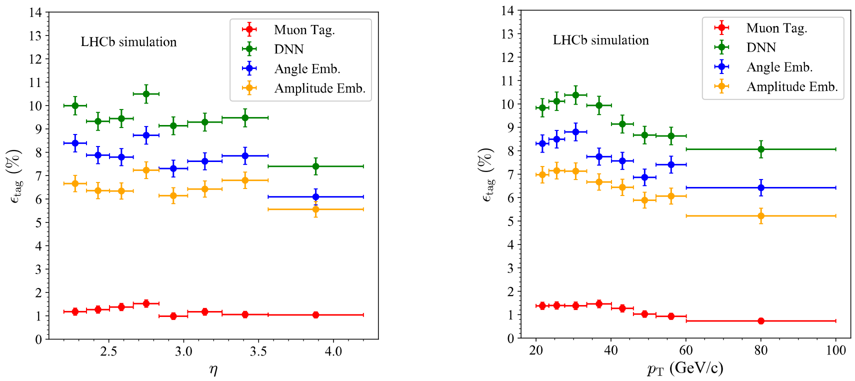


Fig. 1. – b *vs.* $\bar{b}$ jet tagging - $\epsilon_{\rm tag}$ with respect to (left) jet p_T and (right) jet η [1].

4. – Entanglement formation in a Quantum Circuit

An exclusive feature of quantum computation is the concept of *generated entanglement*. Throughout the training process, qubits exhibit varying degrees of entanglement among each other. In the context of QML, entanglement plays a crucial role in fostering cooperative behaviors between qubits, enabling the model to capture intricate patterns.

Earlier studies have delved into the analysis of entanglement formation during the training process [3, 4]. The examination of this quantity is significant, as it has been demonstrated to be correlated with the quality of the training process itself [5].

Additionally, the phenomenon of entanglement formation opens opportunities for the development of various models aimed at gaining insights into the input dataset. As an example, we present the results of a novel quantum model capable of getting the importance and correlations among the features of the input data set.

In fig. 3, a comparison is presented between the feature importances of 6 features from the SV dataset. These importances are computed using a classical model (*Permutation Importance*) and from the values of entanglement within the dedicated quantum model.

Conversely, fig. 4 displays the correlation results of features from the SV dataset. The correlations have been computed using the quantum method (left), and with two "classical" metrics, namely the normalized mutual information (middle) and the Pearson squared coefficient (right).

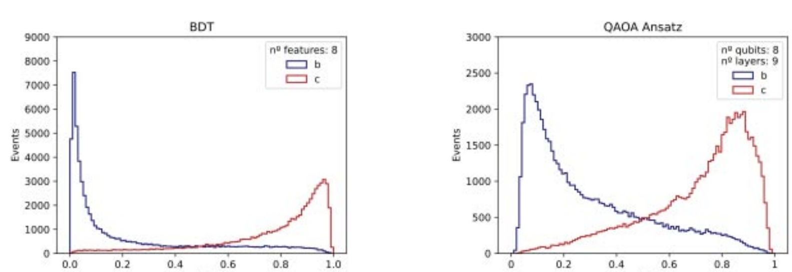


Fig. 2. – b *vs.* c jet tagging - predictions of the classical (left) and quantum model (right) [2].

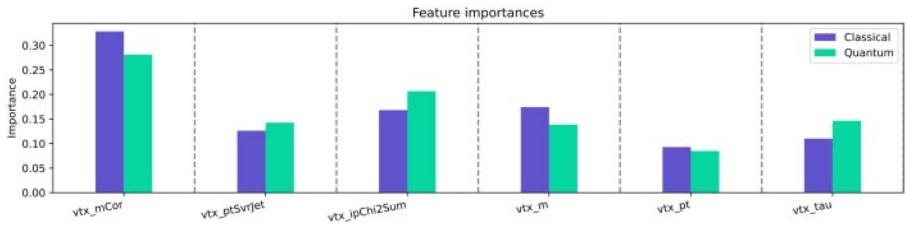
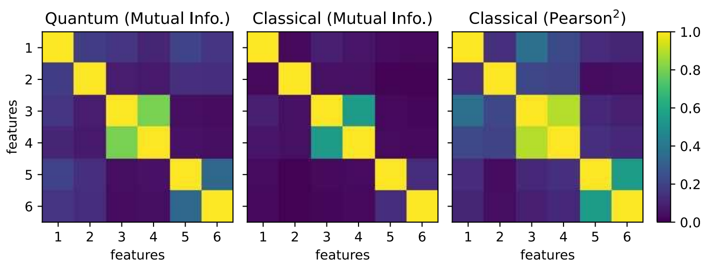


Fig. 3. – Feature importance of 6 features of the SV data set computed with the Quantum method and classical method [5].



Features:

1: `vtx_mCor`
2: `vtx_ipChi2Sum`
3: `vtx_nTrk`
4: `vtx_nTrkJet`
5: `vtx_ptSvrJet`
6: `vtx_pt`

Fig. 4. – Correlation of the features using the fully quantum method (left), classical mutual information (middle), and the Pearson coefficient (right) [5].

Since quantum methods exhibit greater expressiveness, entanglement-based models are potential candidates for revealing new hidden structures within the input data. The above results show similarities with their classical counterparts. However, further investigation is required to determine whether their differences might reveal new information about the jet formation phenomena.

REFERENCES

[1] Gianelle Alessio *et al.*, *Quantum Machine Learning for b-jet charge identification*, arXiv:2202.13943 (2022).
[2] Cocha Carlos *et al.*, *Study of b-and c-jets identification with quantum machine learning algorithms and application to the Higgs reconstruction*, Master's Thesis (2022) `https://hdl.handle.net/20.500.12608/32223`.
[3] Ballarin Marco *et al.*, *Entanglement entropy production in Quantum Neural Networks*, arXiv:2206.02474 (2022).
[4] Felser Timo *et al.*, *npj Quantum Inf.*, **7** (2021) 111.
[5] Monaco Saverio *et al.*, *Study of quantum correlations in LHCb simulated heavy flavour jets*, Master's Thesis (2022) `https://hdl.handle.net/20.500.12608/51029`.