CERN

Quark and gluon jet tagging in Heavy Ion backgrounds

Author: José M. Penín Ascariz
University of Santiago de Compostela

Supervisor: Leticia Cunqueiro Méndez

September 4, 2014

1 Motivation

It is known both from QCD and for several experimental measurements [1], [2], [3], that the jets originated in pp collisions are different depending on the type of parton that originates them. A hierarchical relation between the energy radiated by gluons, light quarks and heavy quarks is seen, such that: E(gluon) > E(light quark) > E(heavy quark). The underlying principle that governs the first relation is the difference in the colour factors for quarks and gluons, that translates into the probability of the given parton to fragment. For gluons, $C_{\mathrm{A}} = 3$ and for quarks $C_{\mathrm{F}} = 4/3$. The second relation is due to a mass effect. The net result observed is that gluon jets are less collimated than quark jets, and that the fragmentation function of gluon jets is softer than that of quark jets.

The energy loss pattern in PbPb collisions may not be the same as in pp collisions, because of the formation of quark-gluon plasma (QGP). For this reason, it would be interesting to measure the R_{AA} quotient in pure quark and gluon jet subsamples. Hence, it is neccesary to find a set of jet observables able to perform the tagging task. Such a set is studied in this work.

2 Tagging capabilities of different jet observables

Many jet shapes are sensitive to differences in quark and gluon fragmentation:

- Girth (first radial moment):

$$g = \sum_{i \in \mathrm{jet}} \frac{p_{\mathrm{T}}^{i}}{p_{\mathrm{T}}^{\mathrm{jet}}} \mid \Delta R_{i,\mathrm{jet}} \mid$$

  - larger for gluons

- $p_{\mathrm{T}}D$:

$$p_{\mathrm{T}}D = \frac{\sqrt{\sum_i p_{\mathrm{T},i}^2}}{\sum_i p_{\mathrm{T},i}} \qquad \begin{cases} \text{if } n = 1, \ p_{\mathrm{T}}D \to 1 \\ \text{if } n \to \infty, \ p_{\mathrm{T}}D \to 0 \end{cases}$$

  - larger for quarks

- Number of constituents:
  - larger for gluons

- Leading constituent p_{T} -Subleading constituent p_{T}:
  - larger for quarks

- Circularity:
  - smallest eigenvalue of circularity matrix → larger for gluons.

3 Jet observables in vacuum and in presence of quenching

In order to analyze the tagging performance of the set of observables defined above in pp collisions and with the presence of quenching, 10^5 events of pp collisions were simulated using PYTHIA 6. The underlying event was simulated with the ACR tune A, modified with annealing. A cut in $p_{\mathrm{Tmin}}^{\mathrm{hard}} = 50$ GeV was placed. To introduce quenching effects, PYQUEN afterburner with default LHC parameters was used. Jets were found with Fastjet using anti-k_{T} clustering algorithm, with $R = 0.4$. The jets were tagged after requiring them to contain a parent parton within the distance $R = 0.2$. Quenching softens and broadens the jet fragmentation, as can be seen in figure 1, where a cut in $p_{\mathrm{T}}^{jet} > 40$ GeV was placed.

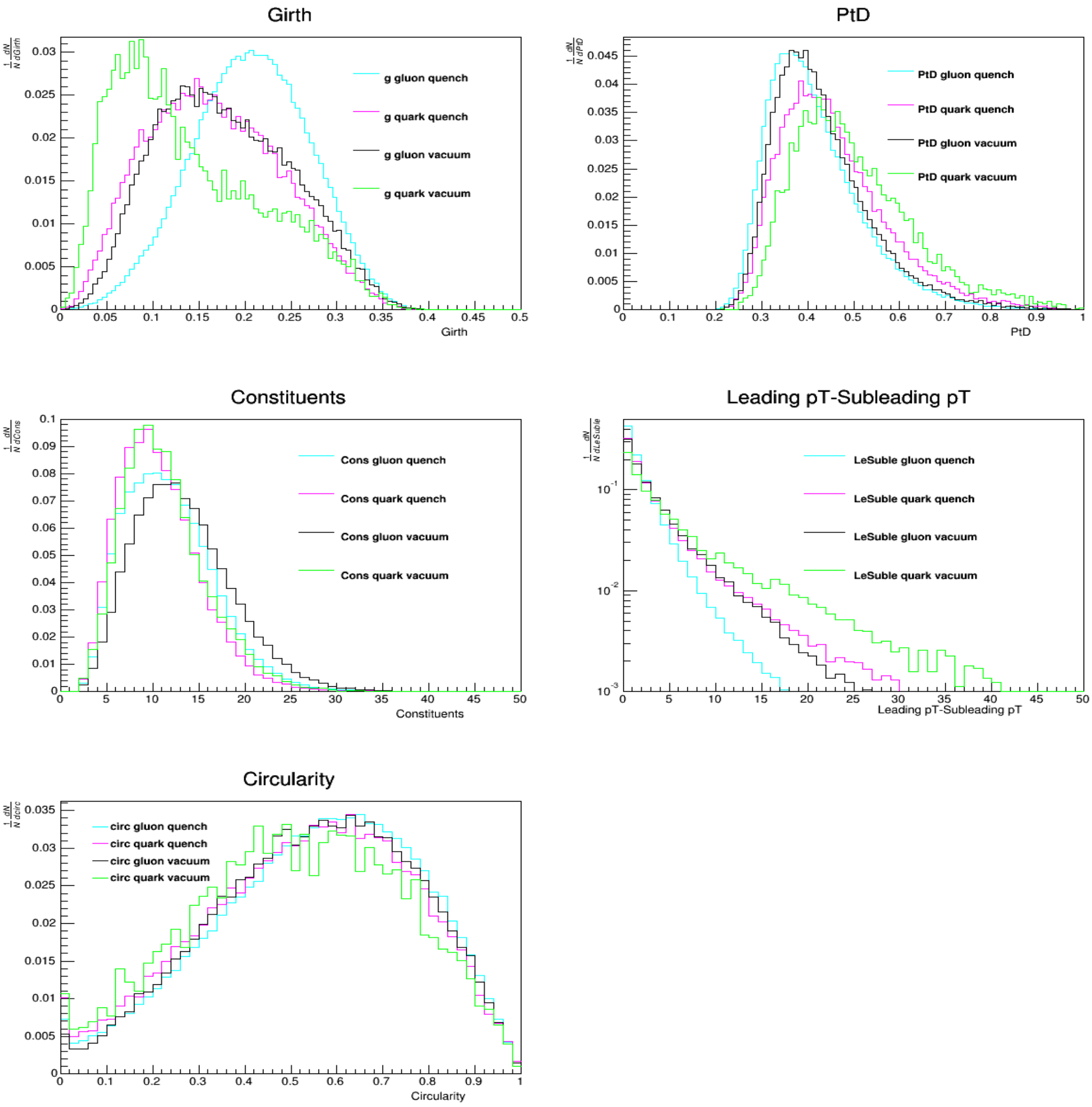


Figure 1: Jet observables in vacuum and quenching, with $p_{\mathrm{T}}^{jet} > 40$ GeV.

4 Multivariate techniques

It is important to find the optimal cuts in the set of observables to discriminate quark jets from gluon jets. For this, multivariate techniques were used, with the help of TMVA package [12]. The performance of three methods was studied:

- Likelihood classifier: it is built by the product of the jet observables defined above. It can be interpreted as the probability of a jet to be originated by a quark.
- BDT (Boosted Decision Trees): it consists on succesive decision nodes that classify events as either signal or background, and after that, cut criteria are defined to obtain maximum signal purity.
- Rectangular cuts: rectangular cuts are placed in the N-dimensional feature space to increase the performance to distinguish quark jets from gluon jets.

The performance of tagging is most understood by studying the output of the TMVA methods. The discriminant $y\ (x)$ is a projection of the N-dimensional feature space onto one dimensional axis (figure 2). Varying the value of $y\ (x)$, i.e., placing a cut in a given point, the working point along the ROC curve, i.e. (efficiency and purity) is moved (figure 3).

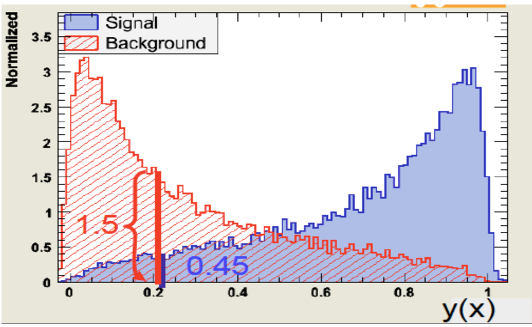


Figure 2: Classifier distribution.

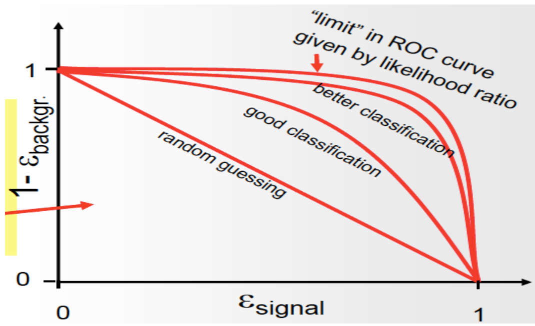
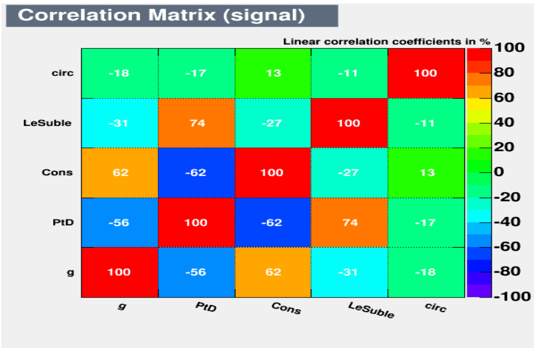


Figure 3: ROC curve.

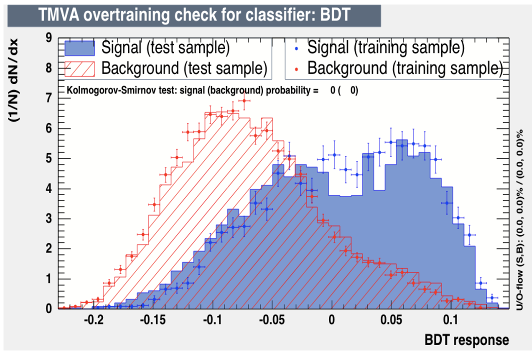
This techniques were applied to the simulations in vacuum. The results obtained are shown in figures 4 and 5. Only circularity is completely uncorrelated to the other set of observables. By studying the ROC curve, it can be seen that a significant tagging performance can be achieved. After applying quenching to the simulations, the tagging performance is not worsen (figure 5).

5 Quark and gluon tagging in Heavy Ions

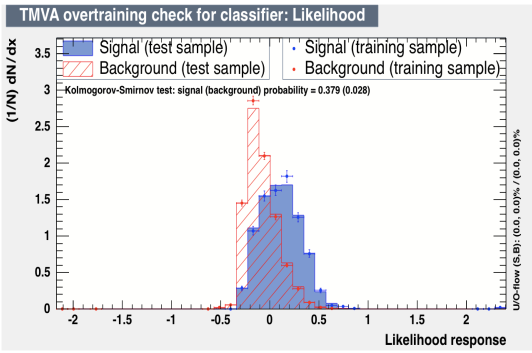
The next step is to analyze the performance of tagging in Heavy Ion backgrounds. Due to the high instantaneous luminosities achieved in LHC, the problem of pileup arises. Each hard pp collision is accompained by aditional $N \approx 600$ pp collisions. This causes an increase of low-p_T noise: $\rho \approx 150$ GeV/fm in central collisions, producing a smearing in the jet p_T, proportional to the noise level and the area of the jet. There are different methods to correct



(a) Correlation matrix.

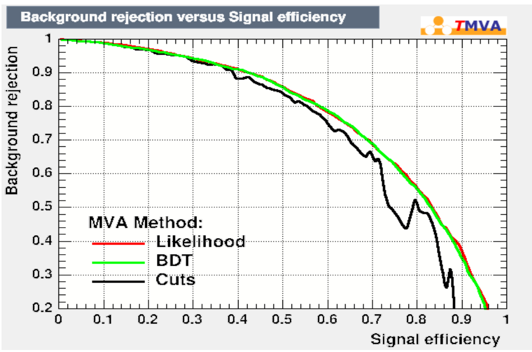


(b) BDT output distribution.

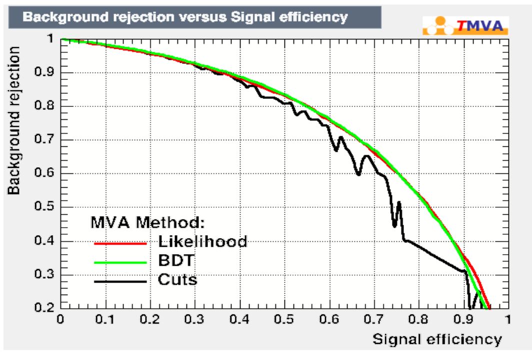


(c) Likelihood output distribution.

Figure 4: TMVA output for vacuum. ($p_{\mathrm{T}}^{jet} > 40$ GeV)



(a) Vacuum: good tagging performance.



(b) Quenching: tagging performance not worsen.

Figure 5: ROC curve for vacuum and quenching.

the jet p_{T}, such as the area-based subtraction, in which: $p_{\mathrm{T\ corrected}} = p_{\mathrm{T\ reconstructed}} - \rho \cdot A$. But this method has proven to be difficult to apply to jet shapes. For this reason, several new methods were introduced to correct jet shapes in general, which can be found in the Fastjet Contrib package, such as:

- Generic subtractor method [4]

- Constituent subtractor method [5]

For creating the simulations in medium with pileup, hard PYTHIA jets were embedded into a toy model for background: [10]. The hybrid jets were matched to the PYTHIA jets, by requiring them to be close in phase space and having at least 50 % of shared energy. After that, the jet shapes were corrected using the methods mentioned above.

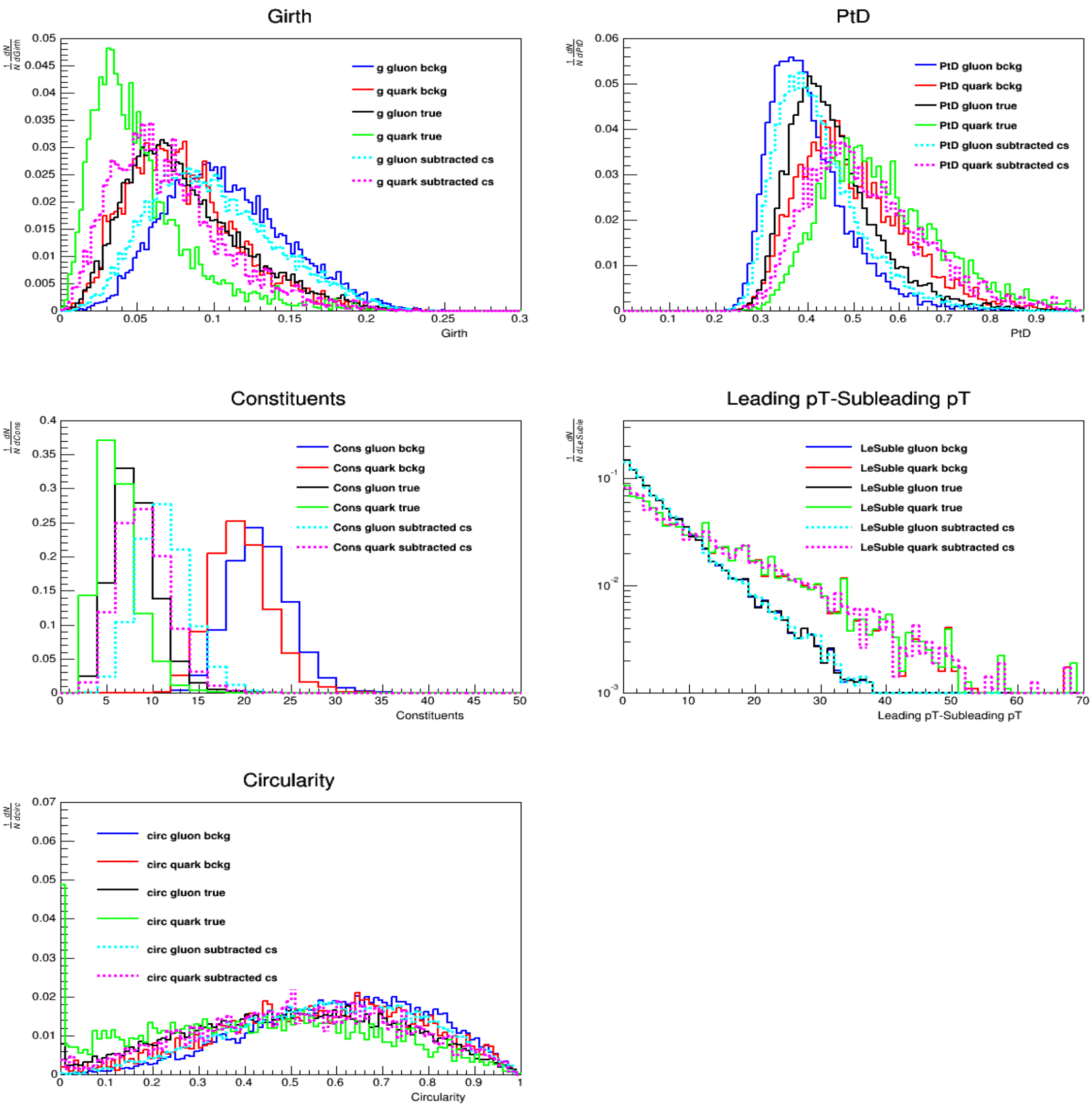
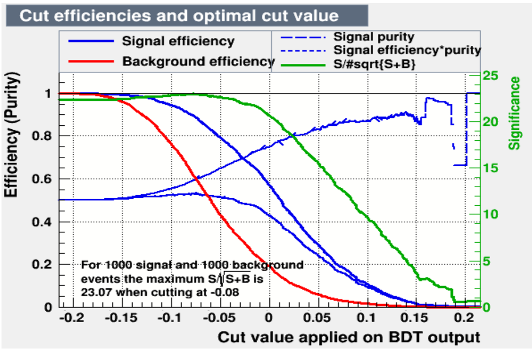


Figure 6: Subtracted shapes using Constituent subtractor method: significant performance of subtraction. ($p_{\mathrm{T}}^{true} > 40$ GeV and $p_{\mathrm{T}}^{cons} > 2$ GeV)

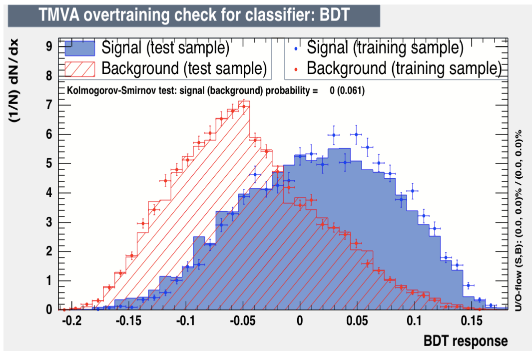
The results of these simulations show that there exist possibilities of tagging in Heavy Ion backgrounds. Both subtraction techniques applied have good performance, as can be seen in figure 6, in this case for the Constituent subtractor method.

A remarkable result obtained in the simulation is that the leading constituent p_T - subleading constituent p_T is a variable insensitive to background. The reason for this is that adding particles with low p_T do not change the roles of the particles with highest p_T. This seems to be an advantage to take into account in future studies. Another interesting result is that circularity is a new good observable, not correlated with the others, which improves the tagging task.

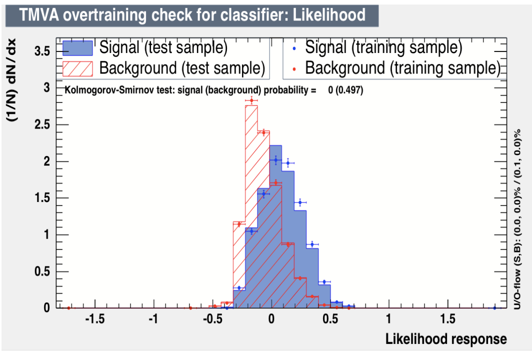
The output of TMVA methods reveals that in Heavy Ion background there are still possibilities of tagging. The output for the corrected shapes using the Constituent subtractor method is shown in figure 7.



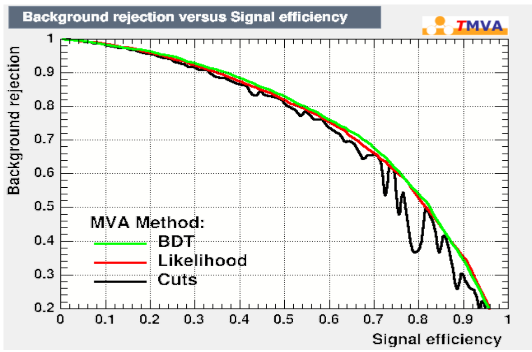
(a) BDT cut eficiencies.



(b) BDT output distribution.



(c) Likelihood output distribution.



(d) ROC.

Figure 7: Subtracted shapes using Constituent subtractor method: possible to discriminate. ($p_T^{cons} > 2$ GeV)

6 Conclusion

An exploratory study of the possibilities of quark-gluon jet tagging was presented. A promising set of observables which can perform the tagging task was found. Also new tools for background subtraction of jet shapes were studied, which may still be optimized. Future explorations may require the use of a more realistic PbPb background.

References

[1] P. D. Acton *et al.* [OPAL Collaboration], Z. Phys. C **58** (1993) 387.

[2] P. Abreu *et al.* [DELPHI Collaboration], Z. Phys. C **70** (1996) 179.

[3] R. Akers *et al.* [OPAL Collaboration], Z. Phys. C **68** (1995) 179.

[4] G. Soyez, G. P. Salam, J. Kim, S. Dutta and M. Cacciari, Phys. Rev. Lett. **110**, no. 16, 162001 (2013) [arXiv:1211.2811 [hep-ph]].

[5] P. Berta, M. Spousta, D. W. Miller and R. Leitner, JHEP **1406**, 092 (2014) [arXiv:1403.3108 [hep-ex]].

[6] M. Cacciari, G. P. Salam and G. Soyez, Eur. Phys. J. C **72** (2012) 1896 [arXiv:1111.6097 [hep-ph]].

[7] T. Sjostrand, S. Mrenna and P. Z. Skands, JHEP **0605** (2006) 026 [hep-ph/0603175].

[8] D. d'Enterria, arXiv:0902.2011 [nucl-ex].

[9] A. Altheimer, S. Arora, L. Asquith, G. Brooijmans, J. Butterworth, M. Campanelli, B. Chapleau and A. E. Cholakian *et al.*, J. Phys. G **39** (2012) 063001 [arXiv:1201.0008 [hep-ph]].

[10] N. S. Amelin, N. Armesto, C. Pajares and D. Sousa, Eur. Phys. J. C **22** (2001) 149 [hep-ph/0103060].

[11] I. P. Lokhtin and A. M. Snigirev, Eur. Phys. J. C **45** (2006) 211 [hep-ph/0506189].

[12] A. Hocker, J. Stelzer, F. Tegenfeldt, H. Voss, K. Voss, A. Christov, S. Henrot-Versille and M. Jachowski *et al.*, PoS ACAT (2007) 040 [physics/0703039 [PHYSICS]].