Boosted W-Boson Identification at $\sqrt{s} = 8$ TeV with the ATLAS Detector

Chris Malena Delitzsch
on behalf of the ATLAS Collaboration
Département de Physique Nucléaire et Corpusculaire, Quai Ernest-Ansermet 24, 1211 Geneva, Switzerland

The performance of boosted hadronic W-boson identification has been studied with the data collected in 2012 with the ATLAS detector at the Large Hadron Collider and is presented here. Different jet grooming algorithms have been tested in simulated events to estimate the discrimination of signal jets coming from the decay of W-bosons with respect to light-quark and gluon jets. The performance of W-boson tagging based on the jet mass and different jet substructure variables is compared. Tagging variables with a high background rejection are compared in data and simulation using a pure sample of W-boson jets in top-quark pair events. Good agreement has been observed for all the variables considered.

1 Introduction

The high center-of-mass energy of the pp collisions at the LHC enables searches for new particles with masses at the TeV scale. These heavy resonances can decay to final states with high p_{T} W- and Z-bosons. The hadronic decay modes of these bosons hold the potential for increased sensitivity via use of jet substructure techniques. However the cross-section of background events originating from light-quark and gluon jets is orders of magnitude higher than the production of W/Z-bosons. At large transverse momentum p_{T}, the decay products of the boson are collimated into one individual large-radius jet. Due to the high-luminosity conditions, soft particles unrelated to the hard scattering can contaminate these jets resulting in a diminished mass resolution. To enhance the sensitivity to new physics processes and to mitigate the influence of pile-up, jet grooming algorithms such as trimming [1] and mass-drop filtering [2] (BDRS) have been designed. In addition, substructure techniques are used to distinguish a two-body decay of a boson from a jet originating from gluons or light-quarks. Within the ATLAS Collaboration [3] the combinations of these techniques have been extensively studied [4] to identify boosted W-bosons. The same techniques could be in principle used to identify boosted Z-bosons.

2 Samples and Event Selection

Simulations of Kaluza-Klein gravitons [5] G $\rightarrow WW \rightarrow \ell\nu jj$ with masses between 400 and 2000 GeV are used to provide a pure sample of jets from hadronically decaying W-bosons. For the background, jets produced in association with a $W(\rightarrow \ell\nu)$ and $Z(\rightarrow \ell\ell)$-boson are used. Jets originating from light-quarks and gluons are in the following referred to as "QCD jets".
To make an unbiased comparison of the different grooming techniques and jet clustering algorithms, the leading calorimeter jet has to be matched to the same C/A $R = 1.2$ ungroomed truth jet with $p_{\mathrm{T}} > 100$ GeV which is then used to divide the sample in different p_{T} bins.

3 Jet Algorithm and Grooming Algorithms

Jets used in the following studies are reconstructed either with the anti-k$_t$ [6] or the Cambridge-Aachen (C/A) [7,8] algorithms. To uncover the hard substructure of signal jets in dense pile-up conditions, several grooming algorithms have been studied but only the best performing are presented: trimming and the BDRS algorithm.

For the trimming algorithm, subjets of $R = 0.3$ are removed if they carry less than 5% of the parent anti-k$_t$ $R = 1.0$ jet p_{T}.

The BRDS algorithm decomposes a C/A jet with $R = 1.2$ (J) into two subjets $j_1,\ j_2$ ($m_{j_1} > m_{j_2}$) by undoing its last clustering step. If there is a significant mass drop $\mu_{12} = m_{j_1}/m_J < \frac{2}{3}$ and the subjets are balanced in momentum $\sqrt{y_f} = \frac{\min(p_{\mathrm{T}}^{j_1}, p_{\mathrm{T}}^{j_2})}{m_J} \times \Delta R_{12} > 0.3$, the jet J is presumed to have an underlying hard structure and is kept. Otherwise the procedure is repeated with $J \rightarrow j_1$.

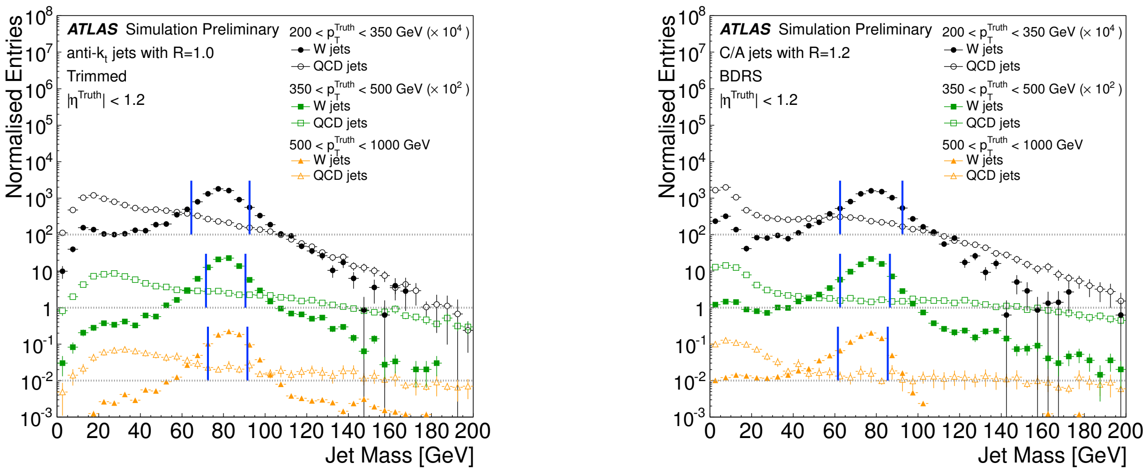

Figure 1 – Reconstructed mass distribution for anti-k$_t$ $R = 1.0$ trimmed jets (left) and C/A $R = 1.2$ BDRS jets (right). Shown are the jet mass distributions in signal (W-jets) and background (QCD-jets) for three different p_{T} bins. The smallest mass window containing 68% of the signal events is indicated by the blue vertical lines [4].

4 Grooming Algorithm Performance

The jet mass distributions are used to compare the performance of the different jet grooming algorithms and are shown in Fig. 1 for anti-k$_t$ $R = 1.0$ trimmed jets and C/A $R = 1.2$ BDRS jets in signal and background. A good grooming algorithm is required to have the most probable value close to the mass of the W-bosons, and a minimal background efficiency within a mass window containing 68% of the signal events. Furthermore, a good signal mass resolution is required. The fraction of QCD jets within the 68% mass window is smaller for the BDRS algorithm than the trimmed algorithm at high p_{T} as shown in Table 1.

Table 1: Background efficiency in the smallest mass window containing 68% of the signal events for anti-k$_t$ $R = 1.0$ trimmed jets and C/A $R = 1.2$ BDRS jets in different p_{T} bins.

p_{T} bin	anti-k$_t$ $R = 1.0$ trimmed	C/A $R = 1.2$ BDRS
$200 < p_{\mathrm{T}}^{\mathrm{truth}} < 350$ GeV	13.6 ± 0.1 %	14.8 ± 0.1 %
$350 < p_{\mathrm{T}}^{\mathrm{truth}} < 500$ GeV	9.9 ± 0.2 %	7.8 ± 0.2 %
$500 < p_{\mathrm{T}}^{\mathrm{truth}} < 1000$ GeV	8.4 ± 0.5 %	6.7 ± 0.5 %

5 Tagging Variable Performance

In addition to the jet mass, the background rejection can be further increased by considering variables that classify whether a jet is more likely to come from the hadronic W-boson decay or from light-quarks and gluons. Examples for these variables are the splitting scales $\sqrt{d_{12}}$ [9], the momentum balance and the N-subjettiness τ_N [10]. The splitting scale $\sqrt{d_{12}}$ distinguishes whether the energy distribution of a jet is symmetric (W-boson) or asymmetric (QCD jet). The N-subjettiness describes to what degree the substructure of a given jet is compatible with the hypothesis of the jet to consist of N or fewer subjets. To discriminate a jet containing N subjets, the ratio τ_N/τ_{N-1} is used. Hence for the two-body W-boson decay, the ratio is τ_2/τ_1. The discriminating power of $\sqrt{d_{12}}$, $\sqrt{y_f}$ and τ_{21} are shown in Fig. 2 for anti-k$_t$ trimmed and C/A BDRS jets.

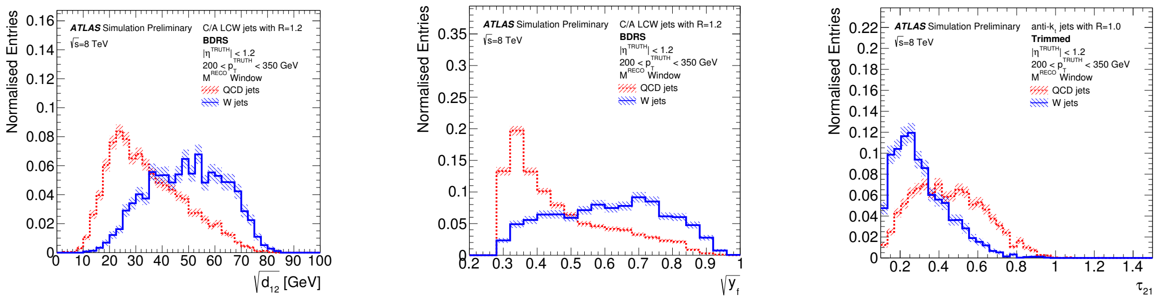

Figure 2 – Distributions of the jet splitting scale $\sqrt{d_{12}}$ (left) and momentum balance $\sqrt{y_f}$ (middle) for C/A $R = 1.2$ BDRS jets and the two-subjettiness τ_{21} (right) for anti-k$_t$ $R = 1.0$ trimmed jets. Shown are the distributions for signal (solid) and background (dashed) jets with $200 < p_{\mathrm{T}}^{\mathrm{Truth}} < 350$ GeV [4].

Within the 68% mass window, the signal efficiency against the background rejection (1 − background efficiency) is shown in Fig. 3, for each substructure variable considered. Only the grooming algorithm with the highest background rejection at a fixed signal efficiency of 50% is shown. The left side of Fig. 3 shows only the effect of the substructure tagging variable whereas the right side shows the combined effect of the mass window cut ($\epsilon^{\mathrm{Groom}}_{\mathrm{W\ Jets}} = 68\%$) and the tagging variable (ϵ^{Tag}). At low signal efficiencies, a high background rejection can be achieved with the BDRS algorithm and $\sqrt{y_f}$ as tagging variable whereas for high signal efficiencies, $\sqrt{d_{12}}$ has a higher background rejection.

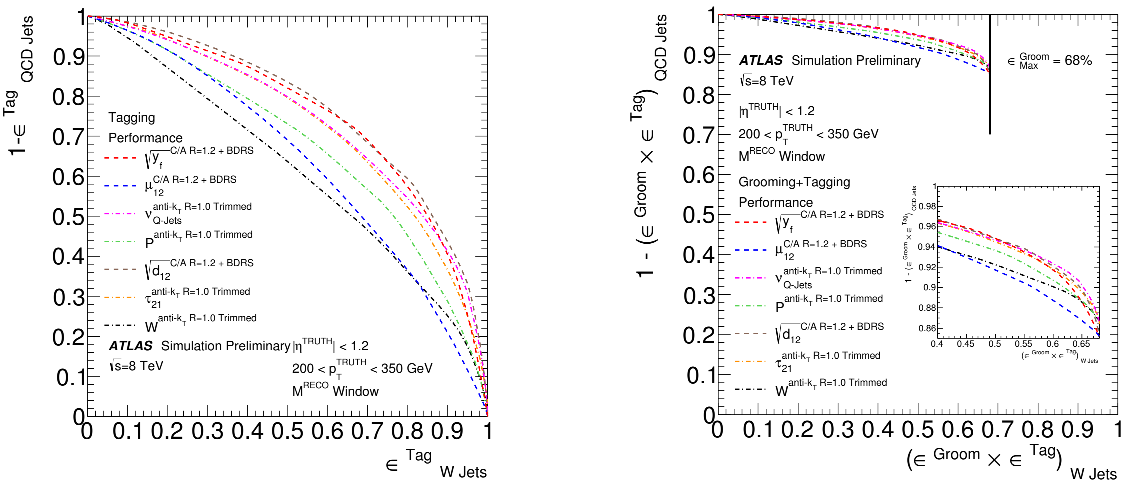

Figure 3 – Comparison of the background rejection (1− $\epsilon^{\mathrm{Tag}}_{\mathrm{QCD\ Jets}}$) as a function of the signal efficiency $\epsilon^{\mathrm{Tag}}_{\mathrm{W\ Jets}}$ for different substructure variables. For each variable only the jet algorithm with the highest background rejection is shown. The tagger performance (left) where the optimal mass window selection is applied in both the numerator and denominator is compared to the tagger+groomer performance (right) using the mass window cut only in the numerator. The grooming efficiency for signal jets $\epsilon^{\mathrm{Groom}}_{\mathrm{W\ Jets}}$ is 68% [4].

6 Substructure Variable Comparison in Data and Simulation

The modelling in the simulation of the variables employed in the tagging algorithms is as important as being able to achieve a large background rejection. In order to compare the distributions in simulation and data, the HEPTopTagger [11,12] algorithm is used to select hadronic boosted W-bosons from lepton+jets $t\bar{t}$ events. The dataset used for these studies corresponds to an integrated luminosity of $\mathcal{L} = 20.3$ fb^{-1} collected at a center-of-mass-energy of $\sqrt{s} = 8$ TeV. Fig. 4 depicts the jet mass, splitting scale $\sqrt{d_{12}}$ and two-subjettiness τ_{21} for anti-k$_t$ $R = 1.0$ trimmed jets in data and simulation. In addition, a Kolmogorov-Smirnov test is performed to test the compatibility between data and simulation. Overall good agreement is observed for the jet mass and the substructure variables.

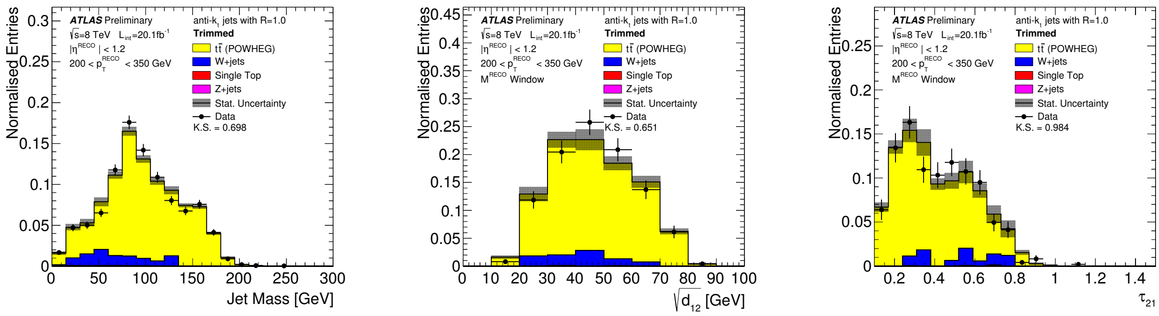

Figure 4 – Leading groomed jet mass (left), splitting scale $\sqrt{d_{12}}$ (middle) and two-subjettiness τ_{21} (right) distribution for anti-k$_t$ $R = 1.0$ trimmed jets in data and MC simulation events passing the HEPTopTagger selection [4].

7 Conclusions

The identification of hadronic W-bosons has been performed with the ATLAS data recorded in 2012 at a center-of-mass energy of $\sqrt{s} = 8$ TeV. To enhance the signal W-bosons over the QCD background (which is many orders of magnitude larger), different jet grooming algorithms and tagging variables have been explored and their combined performance has been studied. The full 2012 dataset corresponding to an integrated luminosity of $\mathcal{L} = 20.3$ fb^{-1} has been used to compare the tagging variables in data and simulation in top-quark pair events which provide a pure sample of hadronic W-bosons. Good agreement was observed for all the variables.

References

1. D. Krohn, J.Thaler, and L.-T. Wang *JHEP* **02**, 084 (2010)
2. J.M Butterworth et al., *Phys. Rev. Lett.* **100**, 242001 (2008)
3. ATLAS Collaboration, JINST **3**, S08003 (2008)
4. The ATLAS Collaboration, ATL-PHYS-PUB-2014-004 (2014)
5. K. Agashe, H. Davoudiasl, G. Perez, and A. Soni, *Phys. Rev.* D **76**, 036006 (2007)
6. M. Cacciari, G.P. Salam and G. Soyez *JHEP* **04**, 063 (2008)
7. Y. L. Dokshitzer, G. D. Leder, S. Moretti and B. R. Webber, *JHEP* **08**, 001 (1997)
8. M. Wobisch and T. Wengler, arXiv:hep-ph/9907280 (1999)
9. J.M Butterworth, B.E. Cox, and J.R. Forshaw, *Phys. Rev.* D **65**, 096014 (2002)
10. J.Tahler and K. Van Tilburg, *JHEP* **03**, 015 (2011)
11. T. Plehn, G.P. Salam, and M. Spannowsky, *Phys. Rev. Lett.* **104**, 111801 (2010)
12. T. Plehn et. al *JHEP* **10**, 078 (2010)