Identification and tagging of double b-hadron jets from gluon splitting with the ATLAS Detector

Lic. María Laura González Silva

Tesis Doctoral en Ciencias Físicas

Facultad de Ciencias Exactas y Naturales

Universidad de Buenos Aires

Noviembre 2012

UNIVERSIDAD DE BUENOS AIRES

Facultad de Ciencias Exactas y Naturales

Departamento de Física

Identification and tagging of double b-hadron jets from gluon splitting with the ATLAS Detector

Trabajo de Tesis para optar por el título de
Doctor de la Universidad de Buenos Aires en el área Ciencias Físicas

por **María Laura González Silva**

Director de Tesis: Dr. Ricardo Piegaia

Consejero de estudios: Dr. Daniel Deflorian

Lugar de Trabajo: Departamento de Física (CONICET-UBA)

Buenos Aires, 2012

AGRADECIMIENTOS

Quiero agradecer a mi director, Ricardo Piegaia, y a todos aquellos que trabajaron junto conmigo en el experimento ATLAS, Gastón Romeo, Gustavo Otero y Garzón, Hernán Reisin y Sabrina Sacerdotti. Un especial agradecimiento a Ariel Schwartzman y su equipo.

Quiero agradecer también a mis compañeros de grupo y oficina, Javier Tiffenberg, Yann Guardincerri, Pablo Pieroni y Orel Gueta.

Quiero agradecer al Experimento ATLAS, al programa HELEN y al programa e-Planet. Quiero agradecer al CONICET y a la Fundación Exactas por hacer posible la realización de esta tesis.

Quiero agradecer el apoyo de mis compañeros de la carrera, especialmente a mis amigos Cecilia Bejarano y Tomas Teitelbaum.

Quiero agradecer a los amigos que hice a lo largo de estos años en mis visitas al Laboratorio CERN, y a mis colegas y amigos de la Universidad de la Plata. Un especial agradecimiento a Fernando Monticelli.

Quiero agradecer a mis amigos de la vida por continuar a mi lado a pesar de las ausencias.

Finalmente, quiero agradecer a mi familia por su apoyo y comprensión, especialmente a Cristina Silva, Lorena González y Juan Martín Alba.

Abstract

Esta tesis describe un método que permite la identificación de jets que contienen dos hadrones b, que se originan en la división de un gluon en un par $b\bar{b}$. La técnica desarrollada explota las diferencias cinemáticas entre los llamados jets "merged" y los genuinos jets b, usando variables que describen la estructura interna y la forma de los jets, construidas a partir de las trazas asociadas a los mismos. Las variables con mayor poder discriminador son combinadas en un análisis de multivariable. Poder identificar y remover jets b que provienen de la división de un gluon es importante para la estimación y la redución del fondo a señales de física dentro del Modelo Estándar y en nueva física. El algoritmo diseñado rechaza, en eventos simulados, el 95% (50%) de los jets "merged", mientras que retiene el 50% (90%) de los jets b genuinos.

Palabras clave: Experimento ATLAS, Jets, Subestructura de Jets, Etiquetado de Jets b, *Gluon Splitting*.

Abstract

This thesis describes a method that allows the identification of double B-hadron jets originating from gluon-splitting. The technique exploits the kinematic differences between the so called "merged" jets and single B-hadron jets using track-based jet shape and jet substructure variables combined in a multivariate likelihood analysis. The ability to reject b-jets from gluon splitting is important to reduce and to improve the estimation of the b-tag background in Standard Model analyses and in new physics searches involving b-jets in the final state. In the simulation, the algorithm rejects 95% (50%) of merged B-hadron jets while retaining 50% (90%) of the tagged b-jets, although the exact values depend on the jet p_T.

Keywords: ATLAS Experiment, Jets, Jet Substructure, b-tagging, Gluon Splitting.

Contents

Chapter 1

Multivariate Analysis

1.1 The multivariate classifiers

The following multivariate methods were explored:

- Likelihood ratio estimators
- Neural Networks (NN)
- Boosted decision Trees (BDTs)

And different trainings were tested:

- Inclusive, with p_T-weighting
- In bins of jet p_T

Signal and background jets were not weighted by the dijet samples cross-sections to allow the contribution of subleading lower p_T jets from high p_T events, and thus increase the statistics of merged jets in the low p_T bins.

Figure 1.1 and 1.2 show distributions of the MVA outputs in different bins of jet p_T for the two proposed trainings. In figures 1.3 and 1.4 a comparison of the performance of all methods, for inclusive and "in-bins", training is illustrated.

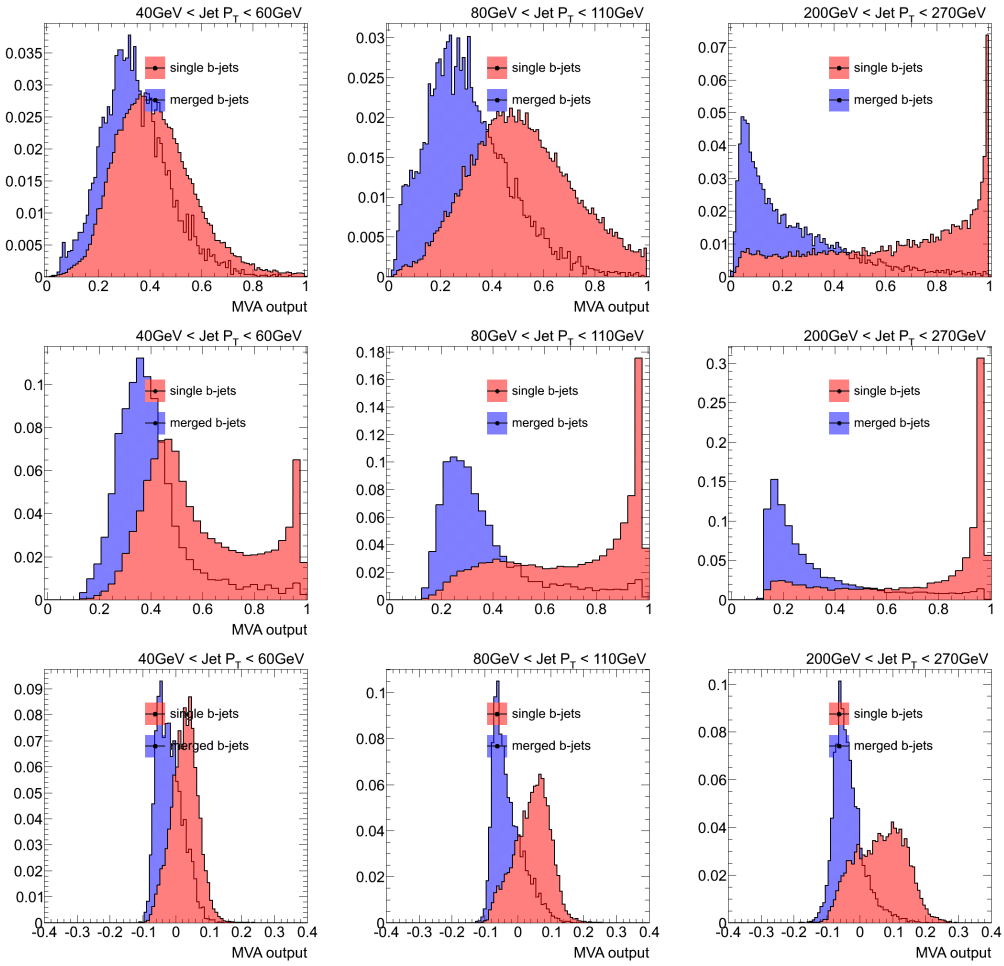

Figure 1.1: Distribution of the MVA discriminant outputs, for inclusive training, in single and merged b-jets, for low, medium and high jet p_T.

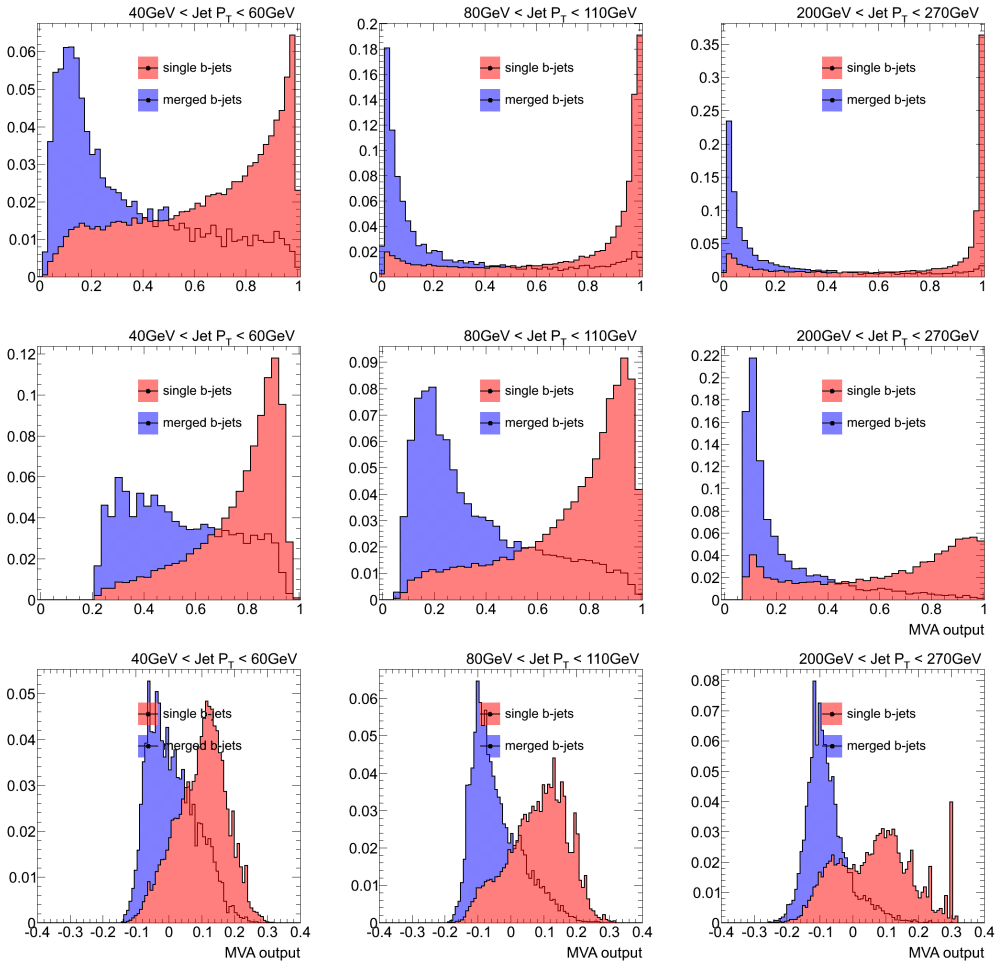

Figure 1.2: Distribution of the MVA discriminant outputs, for training in bins of jet p_T, in single and merged b-jets, for low, medium and high jet p_T.

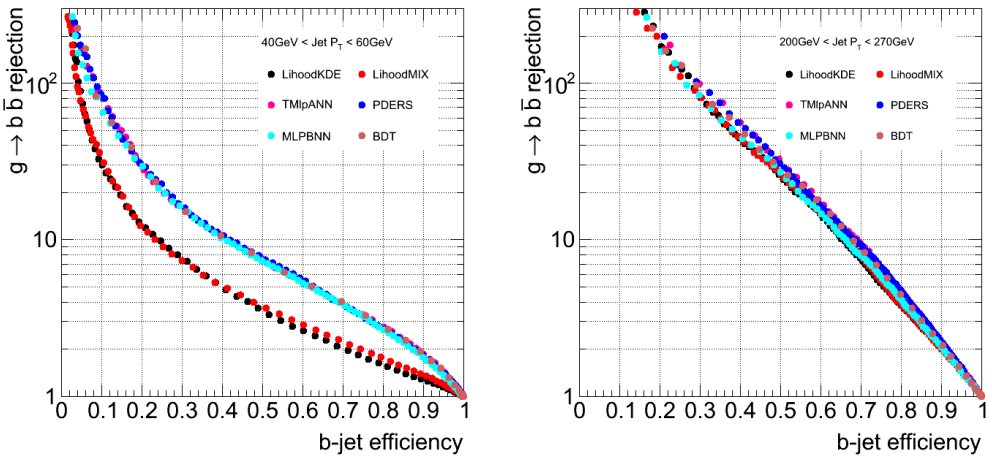

Figure 1.3: Distribution of the MVA discriminant performance for inclusive training, in single and merged b-jets, for low and high jet p_T.

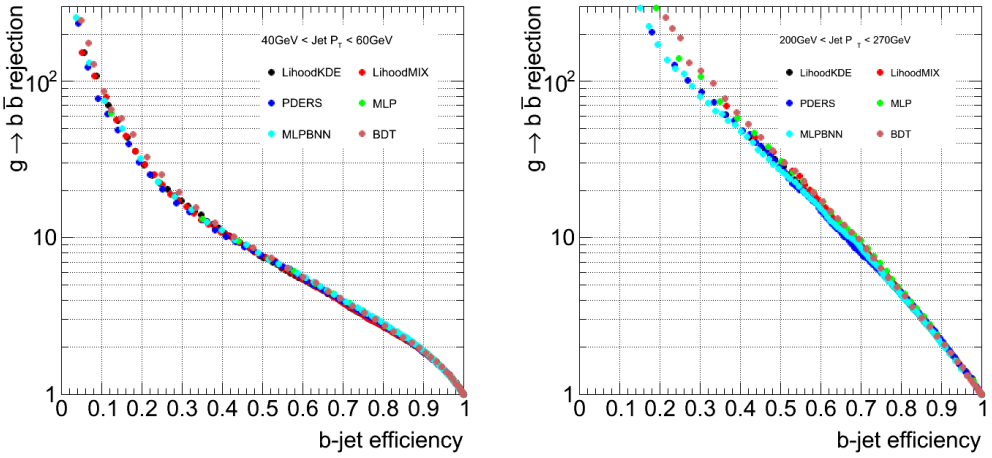

Figure 1.4: Distribution of the MVA discriminant performance for training in bins of jet p_T, in single and merged b-jets, for low and high jet p_T.

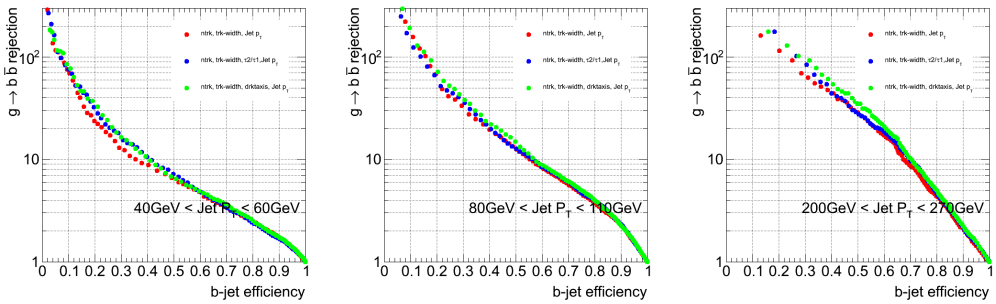

Figure 1.5: Distribution of the MVA discriminant performance for three sets of input variables, in single and merged b-jets, for low, medium and high jet p_T.

1.2 The input variables

Different groups of input variables were tested. Figure **??** shows the performance for three sets of variables for MVA classifier.

1.3 $g \to b\bar{b}$ likelihood training and performance

A discriminant between single b-jets and merged b-jets was built by training a simple likelihood estimator in the context of the Toolkit for Multivariate Data Analysis, TMVA [1].

A sub-set of the dijet Monte Carlo sample was used for training. After the event and jet selections were performed, the b-tagged jets with $|\eta| < 2.1$ were classified as signal (single b-jets) or background (merged b). The likelihood training was done in bins of calorimeter jet p_T. Signal and background jets were not weighted by the dijet samples cross-sections to allow the contribution of subleading lower p_T jets from high p_T events, and thus increase the statistics of merged jets in the low p_T bins. For the evaluation of the method the same procedure was followed.

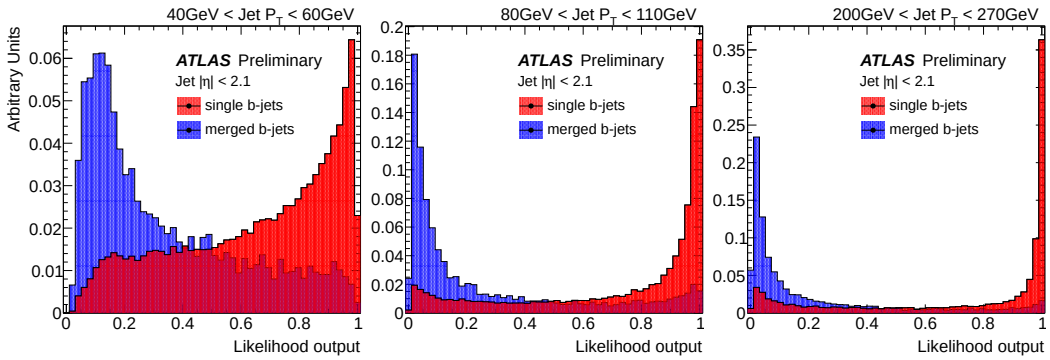

Figure 1.6: Distribution of the $g \to b\bar{b}$ likelihood output for single and merged b-jets for low, medium and high p_T jets.

Several combinations of the tracking and jet shape variables studied in the previous section were tested as input variables. We found that the following three offer the best performance:

1. Jet track multiplicity
2. Track-jet width
3. ΔR between the axes of 2 k_t subjets within the jet

A requirement of at least two matching tracks was imposed to all b-tagged jets in order to build the third variable listed. This cut was applied in both training and testing samples.

The distribution of the likelihood output for single and merged b-jets is shown in Fig. 1.6 for low, medium and high transverse momentum jets.

The performance of the $g \to b\bar{b}$ tagger in the simulation can be displayed in a plot of rejection $(1/\epsilon_{bkg})$ of merged b-jets as a function of single b-jet efficiency, where ϵ_{bkg} is the probability that a $b\bar{b}$-jet passes the tagger. This is shown in Fig. 1.7 for the eight bins of jet p_T mentioned in section **??**. The performance improves with p_T:

- $p_T > 40$ GeV: rejection above 8 at 50% eff.

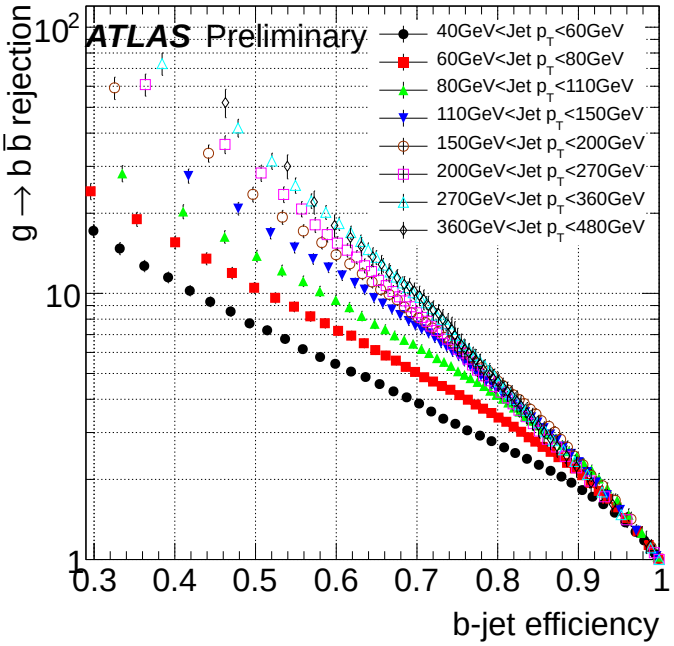

Figure 1.7: Rejection of $g \to b\bar{b}$ merged b-jets as a function of b-jet efficiency for dijet events in 8 jet p_T bins.

- $p_T > 60$ GeV: rejection above 10 at 50% eff.
- $p_T > 200$ GeV: rejection above 30 at 50% eff.

The likelihood was trained with jets that had been first tagged by the MV1 algorithm. In order to use the $g \to b\bar{b}$ classifier for jets tagged by another tagger a new training is required.

The rejection of merged jets attained as a function of p_T for the 50% and 60% efficiency working points are summarized in Table 1.1, together with their relative statistical error. These are propagated from the Poisson fluctuations of the number of events in the merged and single $b\bar{b}$ distributions. The error is slightly lower for the 60% efficiency working point because a higher efficiency allows for a greater number of Monte Carlo events to measure

the performance.

Jet p_T	single b-jet efficiency 50%		single b-jet efficiency 60%	
(GeV)	Rejection	stat.err.	Rejection	stat.err.
40 - 60	8	4%	5	3%
60 - 80	10	4%	7	4%
80 - 110	14	5%	9	4%
110 - 150	19	5%	12	4%
150 - 200	23	5%	14	5%
200 - 270	30	7%	16	6%
270 - 360	36	7%	19	6%
360 - 480	41	8%	18	8%

Table 1.1: The merged b-jet rejection for the 50% and 60% efficiency working points in bins of p_T.

1.4 Systematic uncertainties

The development, training and performance determination of the tagger is based on simulated events. Although the agreement between simulation and data explored in section **??** is a necessary validation condition, it is also important to investigate how the tagger performance depends on systematics relevant in the data. In particular we have considered:

- presence of additional interactions (pile-up)
- uncertainty in the b-jet tagging efficiency
- uncertainty in the track reconstruction efficiency

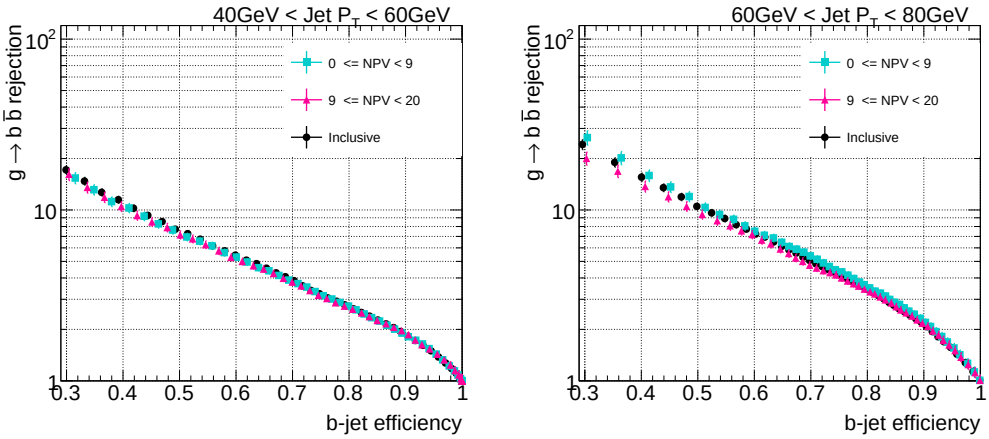

Figure 1.8: Rejection of $g \to b\bar{b}$ merged b-jets as a function of b-jet efficiency in bins of N_{vtx} for two low jet p_T bins.

- uncertainty in the track transverse momentum resolution
- uncertainty in the jet transverse momentum resolution

I. Pile-up

The size of this effect was studied by comparing the performance of the likelihood discriminant with b-jets in events with small (1-9) and large (9-20) number of primary vertices. The comparison of the performance in these two sub-samples can be seen in Fig. 1.8. As expected from the use of tracking (as opposed to calorimeter) variables no significant dependence with pile-up is observed within statistics. Of the 16 determinations (2 working points with 8 p_T bins each) of performance differences between high and low number of primary vertices events, it is observed that 6 of them are positive and 10 negative, with a global mean of 0.3%. We conclude that the effect is negligible compared to other source of uncertainties.

II. b-tagging efficiency

The performance of heavy-flavor tagging in Monte Carlo events is calibrated

to experimental data by means of the scale factors (SFs) measured by the b-tagging group. Such a measurement carries a systematic uncertainty, and in order to estimate its effect a conservative approach is followed: the SFs are varied in all the p_T bins simultaneously by one standard deviation both in the up and down directions. The result of this procedure for the distribution of two of the tracking variables used in our discriminant is illustrated in Fig. 1.9.

The effect of the b-tagging calibration uncertainty on the likelihood peformance is $< 1\%$, negligible with respect to the statistical uncertainty as it can be seen in Fig. 1.10. This was indeed expected. The scale factors depend on the true flavor of the jet and on its p_T, but these are basically constant in the performance determination, which is based on single flavor (true b-) jets classified in p_T-bins.

III. Track reconstruction efficiency

This uncertainty arises from the limit in the understanding of the material layout of the Inner Detector. To test its impact a fraction of tracks determined from the track efficiency uncertainty was randomly removed following the method in Ref. [2].

The tracking efficiency systematics are given in bins of track η. For tracks with $p_{\mathrm{T}}^{\mathrm{track}} > 500$ MeV the uncertainties are independent of p_T: 2% for $|\eta^{\mathrm{track}}| < 1.3$, 3% for $1.3 < |\eta^{\mathrm{track}}| < 1.9$, 4% for $1.9 < |\eta^{\mathrm{track}}| < 2.1$, 4% for $2.1 < |\eta^{\mathrm{track}}| < 2.3$ and 7% for $2.3 < |\eta^{\mathrm{track}}| < 2.5$ [3]. All numbers are relative to the corresponding tracking efficiencies.

The tracking variables were re-calculated and the performance of the nominal likelihood was evaluated in the new sample with worse tracking efficiency. The rejection-efficiency plots, shown in Fig. 1.11, show a small degradation

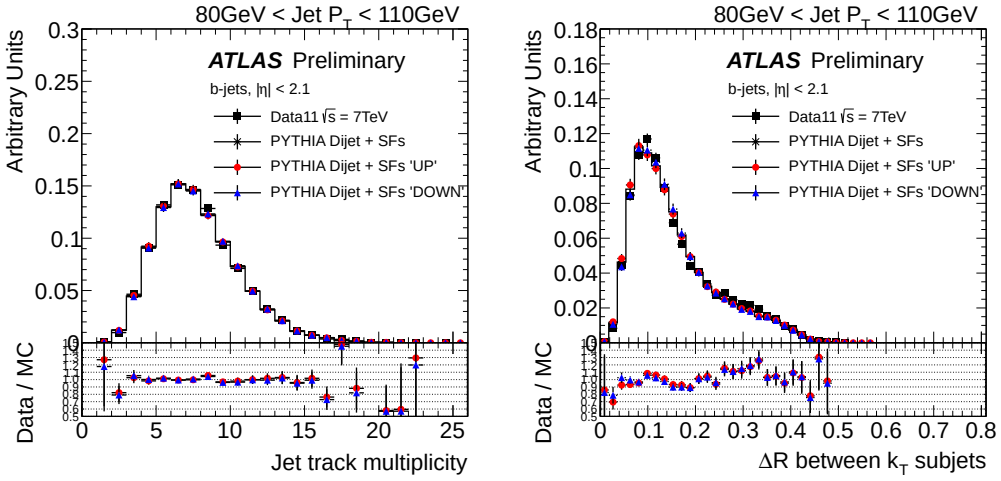

Figure 1.9: The effect of a variation in the b-tagging Scale Factors on the tracking variables distributions. Scale Factors were varied up (down) by 1-sigma to evaluate the systematic uncertainty from this source. The ratio data over MC is shown for MC PYTHIA with SFs varied up (circles) and down (triangles).

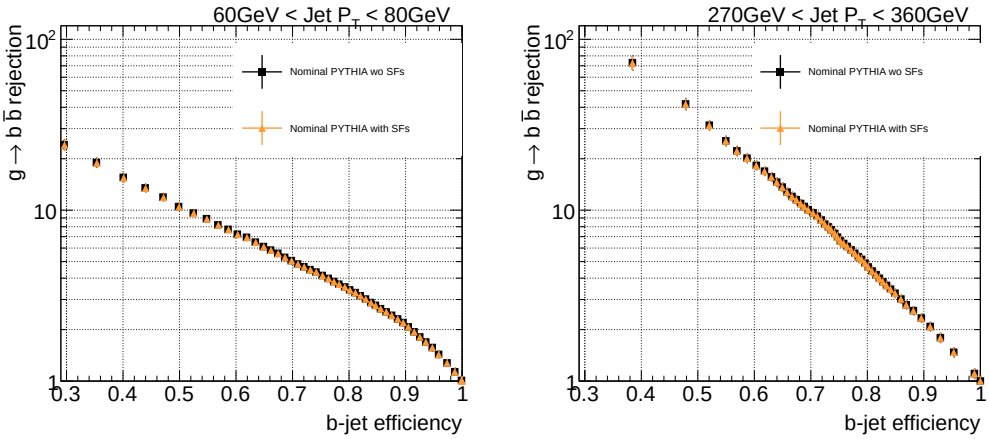

Figure 1.10: Rejection of $g \to b\bar{b}$ merged b-jets as a function of b-jet efficiency with and without scale factors.

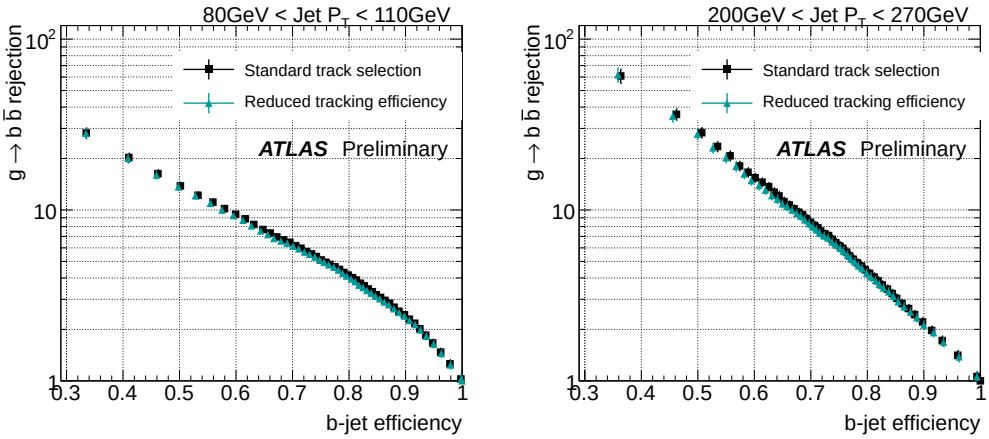

Figure 1.11: Rejection of $g \to b\bar{b}$ merged b-jets as a function of b-jet efficiency showing shift in likelihood performance caused by a reduction in the tracking efficiency .

of the performance which is comparable to the statistical uncertainty. The effect is however systematically present over all 16 p_T bin/working points, without a clear p_T dependence. We have thus taken the average over p_T, and obtained a global systematic uncertainty of 4% both for the 50% and 60% efficiency working points.

IV. Track momentum resolution

The knowledge of the track momentum resolution is limited by the precision both in the material description of the Inner Detector and in the mapping of the magnetic field. Its uncertainty propagates to the kinematic variables used in the $g \to b\bar{b}$ tagger. In order to study this effect, track momenta are over-smeared according to the measured resolution uncertainties before computing the rejection. The actual smearing is done in $1/p_T$, with an upper bound to the resolution uncertainty given by $\sigma(1/p_T)$=$0.02/p_T$ [4]. The effect is found to be negligible.

V. Jet transverse momentum resolution

The jet momentum resolution was measured for 2011 data and found to be in agreement with the predictions from the PYTHIA8-based simulation [5]. The precision of this measurement, determined in p_T and η bins,is typically 10%. The systematic uncertainty due to the calorimeter jet p_T resolution was estimated by over-smearing the jet 4-momentum in the simulated data, without changing jet η or ϕ angles. The performance is found to globally decrease by 6%, without a particular p_T dependence.

The different contributions to the systematic uncertainty on the $g \to b\bar{b}$ rejection are summarized in Table 1.2.

Systematic source	Uncertainty
pile-up	neglible
b-tagging efficiency	neglible
track reconstruction efficiency	4%
track p_T resolution	neglible
jet p_T resolution	6%

Table 1.2: Systematic uncertainities in the merged b-jet rejection (common to both the 50% and the 60% efficiency working points).

1.5 Isolation studies

Although the tagger was derived with isolated jets it can also be applied to non-isolated jets. Studies were performed to evaluate the likelihood rejection in b-jets with close-by jet with p_T between 7 GeV at electromagnetic scale

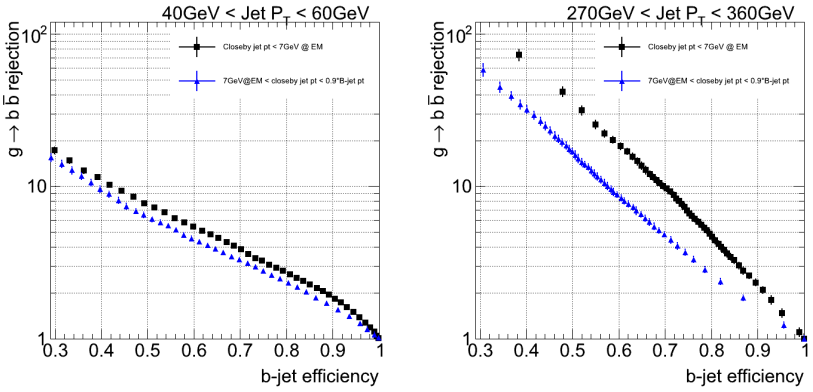

Figure 1.12: Rejection of $g \to b\bar{b}$ merged b-jets as a function of b-jet efficiency for two different isolation cuts.

scale and 90% of the b-jet p_T. The results can be seen in Fig. 1.12. The presence of close-by jets with a susbtancial fraction of the b-jet pt worsens the performance in more than 50% at very high p_T.

1.6 Other Monte Carlo generators

The development, training and performance determination of the tagger has been done using Monte Carlo events generated with the PYTHIA8 event simulator, interfaced to the GEANT4 based simulation of the ATLAS detector. An immediate question is what the performance would be if studied with a different simulation. In this section we investigate this question for the Perugia tune of PYTHIA8 and the HERWIG++ event generators.

Fig. 1.13 shows a comparison of the likelihood rejection, at the 50% efficiency working point, between nominal PYTHIA and the alternative simulations as a function of the jet p_T . The larger errors are due to the reduced statistics available, which are even lower for the Perugia case than for HERWIG.

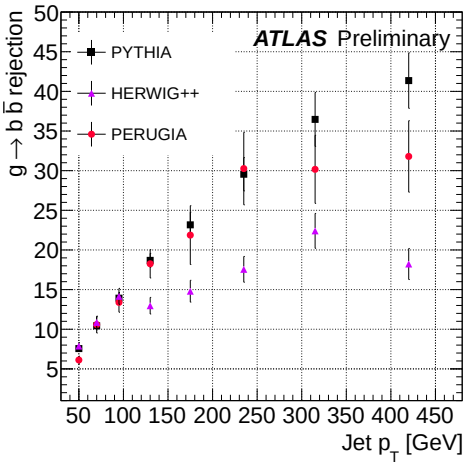

Figure 1.13: Rejection of $g \to b\bar{b}$ merged b-jets as a function of jet p_T for different Monte Carlo generators, at the 50% efficiency working point.

The performance in HERWIG shows a systematic trend, with agreement at low p_T and increasingly poor performances compared to PYTHIA as p_T grows. For the Perugia tune, on the other hand, there is no definite behavior, with the performance fluctuating above or below the nominal simulation for different p_T bins consistently with the statistical uncertainties.

The reason for the systematic difference observed between the performances of PYTHIA and HERWIG can be traced to the extent with which jets are accurately modelled. Fig. 1.14 compares the measured jet track multiplicity distributions in b-tagged jets and the prediction from both simulations, for low and high p_T jets. It is observed that indeed HERWIG++ does not correctly reproduce the data, particularly at high p_T. The level of agreement is found to be better for track-jet width and the ΔR between the axes of the two k_t subjets in the jet, the two other variables used for discrimination.

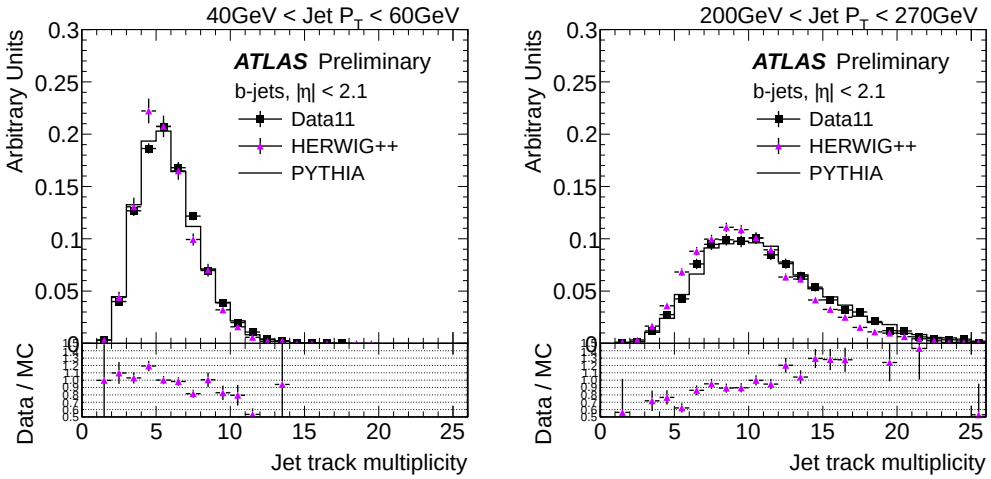

Figure 1.14: Distribution of the jet track multiplicity in 2 different jet p_T bins, for experimental data collected during 2011 (solid black points) and HERWIG++ events (solid violet triangules). The ratio data over HERWIG++ simulation is shown at the bottom of the plot. PYTHIA distribution is also shown for reference.

Bibliography

[1] Andreas Hoecker, Peter Speckmayer, Joerg Stelzer, Jan Therhaag, Eckhard von Toerne, and Helge Voss. TMVA: Toolkit for Multivariate Data Analysis. *PoS*, ACAT:040, 2007.

[2] R Alon et al. Backup Note for Measurement of Jet Mass and Substructure in QCD with the ATLAS Experiment. *ATL-COM-PHYS-2011-401*, 2011.

[3] G. Aad et al. Charged-particle multiplicities in pp interactions measured with the ATLAS detector at the LHC. *New J.Phys.*, 13:053033, 2011.

[4] ATLAS Collaboration. Estimating Track Momentum Resolution in Minimum Bias Events using Simulation and K_s in $\sqrt{s} = 900$ GeV collision data. *ATLAS-CONF-2010-009*, 2010.

[5] G. Romeo, A. Schwartzman, R. Piegaia, T. Carli, and R. Teuscher. Jet energy resolution from in-situ techniques with the atlas detector using proton-proton collisions at a center of mass energy $\sqrt{s} = 7$ tev. *ATL-COM-PHYS-2011-240*, 2011.