CORE
Metadata, citation and similar papers at core.ac.uk
Provided by CERN Document Server

PoS PROCEEDINGS OF SCIENCE

Prospects for b-tagging in ATLAS and tracking commissioning results with cosmic rays

Laurent VACAVANT*, on behalf of the ATLAS Collaboration
CPPM, CNRS/IN2P3 et Aix-Marseille Université
F-13288 Marseille cedex 09, France
and
CERN, Physics Department
CH-1211 Genève 23, Switzerland
E-mail: vacavant@in2p3.fr

The ability to identify jets containing b-hadrons is important for the high-p_T physics program of a general-purpose experiment at the LHC such as ATLAS. This capability relies on the very accurate measurements of the parameters of charged tracks provided by the ATLAS Inner Detector. Using millions of cosmic-ray tracks collected during the automn 2008, the ATLAS Inner Detector has been aligned and its tracking performance assessed. Some of the very encouraging results which have been obtained and are relevant for b-tagging are discussed, notably the current level of alignment of the detector and the resolution on the transverse impact parameter of tracks. The various b-tagging algorithms are then described, and their anticipated performance discussed in the light of the cosmic-ray data results. Finaly, the expected accuracy with which the b-tagging performance will be measured in data is mentioned.

ATL-PHYS-PROC-2010-015
09 February 2010

XXth Hadron Collider Physics Symposium
November 16 – 20, 2009
Evian, France

*Speaker.

© Copyright owned by the author(s) under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike Licence.

1. Introduction

The ability to identify jets containing *b*-hadrons is important for the high-p_T physics program of a general-purpose experiment at the LHC such as ATLAS [1]. This is in particular useful to select very pure top quark samples, to search for new physics (supersymmetry, heavy gauge bosons, etc.), to search and study Standard Model or SUSY Higgs bosons and to veto the large $t\bar{t}$ background for many physics channels.

In 2010 the LHC is expected to deliver a sizable number of collisions at a 7 TeV center-of-mass energy. The lower the energy goes, the less favorable the signal over background ratio becomes for the top rediscovery in $t\bar{t}$ pairs. Requiring one jet to be *b*-tagged reduces significantly the background from $W+$ light jets at a modest cost in signal efficiency, typically improving the S/B ratio by a factor 2. It is thus particularly useful to commission the *b*-tagging at an early stage.

The identification of *b*-jets takes advantage of several of their properties which allow us to distinguish them from jets which contain only lighter quarks: hard fragmentation, high mass of *b*-hadrons and relatively long lifetime, of the order of 1.5 ps. A *b*-hadron in a jet with $p_T = 50$ GeV/c will therefore travel on average about 3 mm in the transverse plane before decaying. This can be identified either inclusively by measuring the impact parameters of the tracks (*i.e.* the distance between the location of the point of closest approach of the track to the collision point) from the *b*-hadron decay products or explicitely by reconstructing the displaced vertex. In both cases, the precise measurement of the parameters of charged tracks in the ATLAS Inner Detector is a key ingredient.

2. Tracking commissioning with cosmic rays

The ATLAS Inner Detector [1] surrounds the beam-pipe and extends up to about one meter in radius and 6 meters in length, covering pseudo-rapidities $|\eta|$ up to 2.5. It is enclosed inside a super-conducting solenoid providing a 2 T axial field. Three technologies are used: 80 million silicon pixels (50×400 μm) for the three innermost layers starting at $r = 5$ cm, four double-layers of silicon micro-strips (SCT, 80 μm pitch, 40 mrad stereo angle) and about 36 layers of 4 mm straw tubes (TRT). As described in Ref. [3], this detector is working extremely well, with more than 98% of its channels being operational and a noise occupancy within specifications (*e.g.* 10^{-10} for the pixel detector).

During autumn 2008, ATLAS recorded about 7.6 million cosmic-ray tracks with the Inner Detector fully integrated, in two configurations with and without magnetic field. Due to the geometrical acceptance for such tracks, the number of tracks crossing the silicon strip and the pixel detectors was reduced to 2 million and 420 thousand, respectively. Only a fraction of those are used in the following results.

The precision with which the positions and orientations of individual modules of the Inner Detector are known is limiting the accuracy of the track reconstruction and must be improved by an alignment procedure. One approach to constrain the 35000 degrees of freedom for the silicon modules of ATLAS is to use a large sample of tracks in order to minimize a χ^2 constructed from the differences between the hit positions and the track positions. A mixture of data recorded with

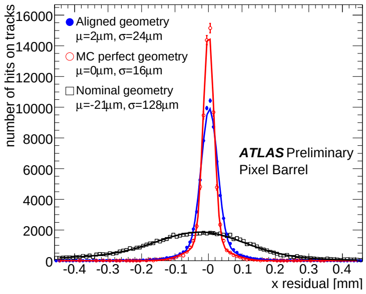

Figure 1: Distribution of the unbiased residual for pixel barrel hits, projected onto the local x coordinate (precision coordinate).

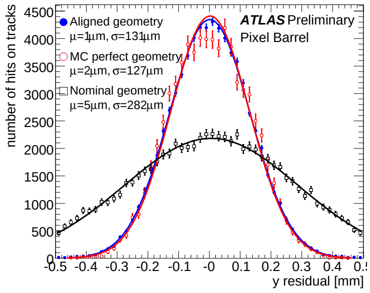

Figure 2: Distribution of the unbiased residual for pixel barrel hits, projected onto the local y coordinate (non-precision coordinate).

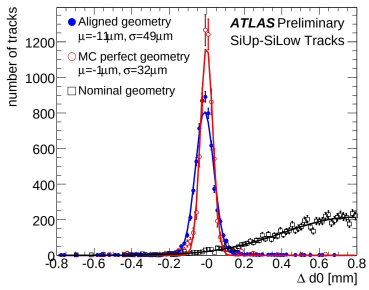

Figure 3: Distribution of the difference in transverse impact parameter between the upper and lower halves of a cosmic track.

and without solenoid magnetic field in 2008 was used to perform a first alignment and obtain a consistent set of alignment corrections.

Figures 1 and 2 show the distributions of the unbiased residuals for pixel barrel hits, projected onto the two local coordinates. The unbiased residual is defined as the distance between the measured hit position and the expected hit position from the track extrapolation, the track being refitted after having removed the hit under study. The data are shown before (black open squares) and after (blue solid markers) the alignment procedure: after alignment, the residuals are largely improved and are very close to the expectations from a simulation sample with perfect geometry (red open circles): the residual misalignment is of the order of 20 μm for the silicon detectors. Using the same alignment constants for data taken in 2009 led to similar results, indicating good stability of the detector over an extended period of time.

The cosmic-ray tracks crossing both upper and lower halves of the Inner Detector can be used to measure the track-parameter resolutions: by splitting the track into two parts one obtains two collision-like tracks. The comparison of the two half-tracks at the perigee provides information about the bias of individual track parameters (mean of the distribution), which is very sensitive to misalignment, and the resolution (width normalized by $\sqrt{2}$). Figure 3 shows such a distribution for the transverse impact parameter d_0 of tracks, showing again the huge improvement brought in by the alignment procedure. Figure 4 shows the distribution of the relative momentum resolution as a function of the transverse momentum of the track, using this splitting technique. Tracks are either reconstructed in the full Inner Detector (plain triangles) or only in the silicon detectors (open triangles) and are compared with simulated tracks in a perfectly aligned geometry. The relative momentum resolution increases with higher p_T due to stiffer tracks and a more difficult measurement of the sagitta: the effect is softened when including information from the TRT which extends the lever arm. Figure 5 shows the distribution of the transverse impact parameter resolution of tracks as a function of their transverse momentum p_T. In the low-p_T region, the d_0 resolution is worse due to multiple scattering effects while it reaches asymptotically the intrinsic detector resolution at high p_T. A typical track in a b-jet from a $t\bar{t}$ pair event at LHC has $p_T \approx 4$ GeV/c and an expected d_0 resolution of 44 μm in simulation: this resolution is measured to be 48 μm in the cosmic data, which is very encouraging. For both distributions, the difference to the simulation curve indicates the level of remaining misalignment which will be reduced once large samples of

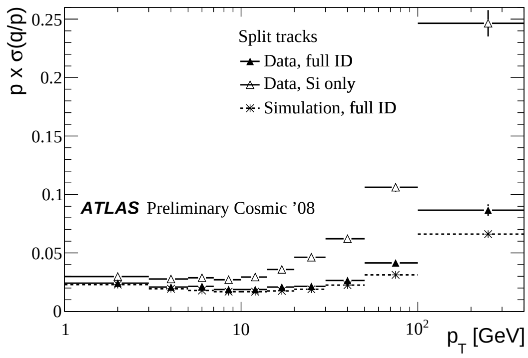

Figure 4: Distribution of the relative momentum resolution as a function of the transverse momentum of the track, for cosmic ray data (full tracker or silicon detectors only) and for simulation.

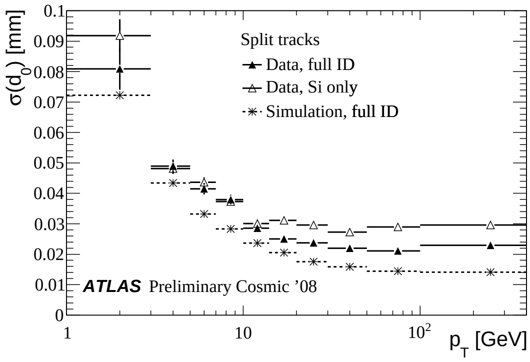

Figure 5: Distribution of the transverse impact parameter resolution as a function of the transverse momentum of the track, for cosmic ray data (full tracker or silicon detectors only) and for simulation.

tracks from collision data are collected and fed into the alignment procedure.

3. Algorithms for *b*-tagging

The transverse (d_0) and longitudinal (z_0) impact parameters of tracks are computed with respect to the primary vertex and are signed positively if the track crosses the jet axis in front of the primary vertex and negatively otherwise. To give more weight to well-measured tracks, the impact parameter significance d_0/σ_{d_0} is used for discriminating *b*- and light jets. The d_0 and d_0/σ_{d_0} distributions are shown on Figures 6 and 7 for jets of various flavors.

The simplest tagging algorithm, called TrackCounting, consists in counting tracks with large transverse impact parameter or impact parameter significance. Another algorithm, JetProb, compares for each track its d_0/σ_{d_0} to a resolution function for prompt tracks, measuring the probability that the track originates from the primary vertex. The track probabilities are then combined into a jet probability. The resolution function can be measured in data using the negative side of the signed impact parameter distribution, assuming there is no contribution from heavy-flavor particles.

To further increase the discrimination between *b*-jets and light jets, the inclusive vertex formed by the decay products of the bottom hadron, including the products of the eventual subsequent charm hadron decay, can be sought. Tracks leading to two-track vertices compatible with a K_s^0, Λ, photon conversion or material interaction are rejected. The distance between the primary vertex and the secondary vertex is used as a discriminant by the third tagging algorithm, called SV0.

These three algorithms are at the core of the *b*-tagging strategy for early data. In addition to them, more advanced algorithms are available based on a likelihood ratio approach: the measured value of a discriminating variable is compared to pre-defined distributions for both the *b*- and light jet hypotheses, obtained initially from Monte Carlo. Multi-dimensional probability density functions are also used by some algorithms. Another sophisticated algorithm, JetFitter, exploits the topological structure of *b*- and *c*-hadron decays inside the jet and provides some additional discrimination between *b*- and *c*-jets.

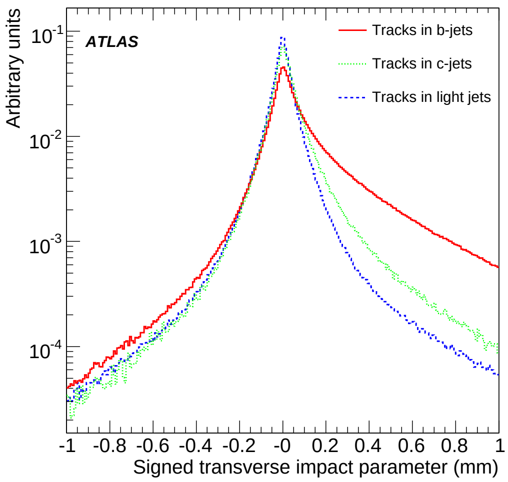

Figure 6: Distribution of the signed transverse impact parameter d_0 for *b*-tagging quality tracks in *b*-jets, *c*-jets and light jets (simulation).

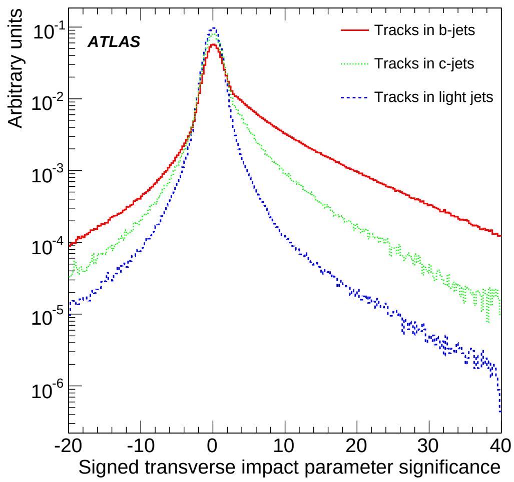

Figure 7: Distribution of the transverse impact parameter significance d_0/σ_{d_0} for *b*-tagging quality tracks in *b*-jets, *c*-jets and light jets (simulation).

4. Anticipated *b*-tagging performance

For performance studies, only jets fulfilling $p_T > 15$ GeV/c and $|\eta| < 2.5$ are considered. The expected *b*-tagging performance has been studied in $t\bar{t}$ simulated events and is estimated by looking at the rejection power against light jets ($1/\varepsilon_{light}$) versus the *b*-tagging efficiency ε_b. For top studies, $\varepsilon_b = 50\%$ is usually sufficient and the early TrackCounting, JetProb and SV0 algorithms can achieve rejections of 90, 110 and 170 respectively. The most advanced algorithms can reach rejections 2 to 5 times higher. For $\varepsilon_b = 60\%$, the light jet rejection ranges between 40 and 90 for the early taggers, and up to 300 for the advanced ones. For completeness, taking advantage of the semi-muonic decay of *b*-hadrons using a soft muon tagging algorithm provides a light jet rejection of 300 for a 10% efficiency on inclusive *b*-jets. The rejection of *c*-jets is limited by the lifetime of charm hadrons: a rejection of around 6 (20 with JetFitter and a dedicated tuning for charm) is obtained for $\varepsilon_b = 60\%$.

Among the various effects studied in Ref. [2], the impact on *b*-tagging of residual misalignments in the pixel detector was studied by running the actual ATLAS alignment procedures on a Monte Carlo sample in which the detector elements were slightly shifted and rotated according to actual surveys or known fabrication precisions. This is the most realistic case considered so far, and comprises many (but not all) systematic deformations including those caused by the alignment procedure itself. In this case, the light jet rejection is at most 25% lower for the same ε_b. In the unlikely case where no progress is made with respect to the current level of alignment exposed in Section 2, a simpler study based on random residual misalignments indicated that the rejection numbers shown above could be reduced by not more than a factor 2 for most algorithms.

It is also worth mentioning that the *b*-tagging performance depends strongly on the jet momentum and rapidity. At low p_T, performance is degraded mostly because of larger multiple scattering. This also holds for the high-$|\eta|$ region, where the amount of material in the tracking region increases significantly. In addition the z_0 resolution is degraded at large rapidities because of the pixel detector geometry. The optimal performance is achieved for central jets with $p_T \approx 120$ GeV/c. Several effects conspire to reduce the b-tagging performance as the jet p_T increases above this value because of the highly collimated jets and boosted b-hadrons.

5. Measurement of *b*-tagging performance in data

While a large effort is put into having a very accurate Monte Carlo simulation, the b-tagging performance must be measured in data. Several studies aiming at measuring the b-tagging efficiency in di-jet events or in $t\bar{t}$ events have been performed and are described in Ref. [2].

The QCD di-jet samples are enriched in heavy flavors by requiring that one of the jets contains a muon. A first method uses Monte Carlo-derived templates of the p_T of the muon relative to the jet+muon axis, for b-, c- and light jets. The second method employs two samples with different b-contents and two uncorrelated tagging algorithms: the soft muon tagger and a lifetime-based one. Both methods are working well for jets with $15 < p_T < 80$ GeV/c and can provide b-tagging efficiency binned in jet p_T and/or η. Studies indicates that it should be possible to control the absolute error on ε_b to 6%, dominated rapidly by systematic uncertainties.

Using the abundant production of $t\bar{t}$ events at LHC is complementary to the di-jet techniques: more data is needed but the tagging efficiency of jets of higher p_T can be measured. One method consists in counting the number of tagged jets: ε_b can be measured with a relative precision of $\pm$2.7(stat.)$\pm$3.4(syst.)% in the lepton+jets channel for 100 pb^{-1} of data. Another method relies on the identification of a very pure b-jet sample by fully reconstructing the $t\bar{t}$ decay chain in the lepton+jets channel. The b-jet on the hadronic side is tagged to improve purity, while the presumed b-jet on the leptonic side is unbiased and used as a probe to measure ε_b. With 200 pb^{-1} of data, a relative error on ε_b of $\pm$6.4(stat.)$\pm$3.4(syst.)% can be achieved.

The accuracy with which the light jet rejection can be measured deserves more studies.

6. Conclusion

A wide spectrum of algorithms has been developed for the identification of b-jets in ATLAS. The simpler ones should provide in early data a light jet rejection of around 50 for a b-tagging efficiency of 60%, while sophisticated algorithms will achieve rejections of 300 later on. The commissioning of the ATLAS Inner Detector and its tracking performance are very encouraging in this respect, showing promising results for instance for the impact parameter resolution, which will be quickly refined with a further alignment based on collision tracks.

References

[1] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST 3 S08003 (2008).

[2] ATLAS Collaboration, *Expected performance of the ATLAS Detector, Trigger and Physics*, CERN-OPEN-2008-020, arXiv:0901.0512 (2009).

[3] X. Wu, ATLAS detector global status, these proceedings.