Tracking, Vertexing and b-tagging Performance in ATLAS

Mark J. Tibbetts[*]
On behalf of the ATLAS Collaboration
Lawrence Berkeley National Laboratory
E-mail: mjtibbetts@lbl.gov

Reconstruction of the trajectories of charged particles and the identification of their points of origin form a vital componant of the ATLAS experiment physics program at the LHC. The resulting reconstructed tracks and verticies can be used to identify jets originating from heavy flavour through the use of b-tagging algorithms. An overview of the current status of these techniques at ATLAS is presented.

36th International Conference on High Energy Physics,
July 4-11, 2012
Melbourne, Australia

---

[*]Speaker.

1. Introduction

The ATLAS experiment [1] records data from collisions produced by the Large Hadron Collider (LHC) [2] at CERN. In proton-proton collisions hundreds of primary and secondary stable[1] charged particles are produced and accurate measurement of their kinematics through track reconstruction is essential in all physics analyses. Vertex identification using reconstructed tracks is important for determining the position of primary interactions as well as those of displaced particle decays and material interactions. Multivariate b-tagging algorithms combine information based on the properties of reconstructed tracks and displaced vertices originating from the decays of heavy flavour hadrons to identify jets originating from heavy flavour partons in the hard scatter and resulting parton shower. Tracking, vertexting and b-tagging performance are determined by comparing the response of algorithms in data to the response of the same algorithms in detailed Monte Carlo (MC) simulations of the ATLAS detector.

2. Track and Vertex Reconstruction

The ATLAS Inner Detector (ID) tracking system surrounds the interaction point and consists of three sub-detectors: innermost is a silicon Pixel detector, this is surrounded by a silicon microstrip detector (SCT) and the combined silicon system is surrounded by a transition radiation tracker (TRT). The ID is immersed in a 2 T axial solenoid field and provides coverage to $|\eta| < 2.5$[2]. Track reconstruction uses combinatorial Kalman fitter pattern recognition algorithms. Most primary tracks are reconstructed with an inside out algorithm extending seeds from the silicon detectors out to the TRT [3]. Recovery of secondary tracks from conversions, material interactions and long-lived particle decays is achieved through a back-tracking algorithm taking TRT based seeds and extending inwards.
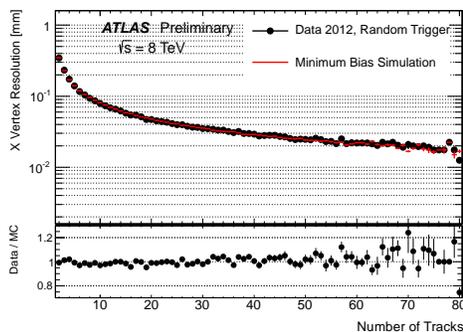

Figure 1: Vertex position resolution in data (black) and MC (red) [5]. The resolution is shown for the transverse coordinate as function of the number of tracks in the vertex fit.

Primary interaction vertices are reconstructed by iteratively fitting tracks consistent with the interaction region using a χ^2 based algorithm with a beam spot constraint [4]. Due to multiple

---

[1]Stable refers to charged particles with $\tau > 3 \times 10^{-11}$ s, i.e., pions, kaons, protons, electrons and muons.

[2]ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the centre of the detector and the z-axis along the beam pipe. The x-axis points from the IP to the centre of the LHC ring, and the y-axis points upward. Cylindrical coordinates (r, ϕ) are used in the transverse plane, ϕ being the azimuthal angle around the beam pipe. The pseudorapidity is defined in terms of the polar angle θ as $\eta = -\ln\tan(\theta/2)$.

proton-proton interactions in a single bunch crossing (pileup) events can have more than one reconstructed primary vertex. The primary vertex identified as the hard scatter in physics analyses is by default the vertex with the highest $\sum p_T^2$ of tracks associated to that vertex; however, analysers have the option to redefine that choice of vertex through physics object association. The reconstructed hard scatter vertex becomes the reference point for impact parameter and flight length measurement b-tagging algorithms. Primary vertex resolution in data is measured with a split vertex technique [4]. Figure 1 shows the primary vertex resolution in transverse coordinate *x* as a function of the number of tracks associated to that vertex in 8 TeV collision data collected during 2012. It can be seen that the MC simulation reproduces the resolution measured in data for all track multiplicities.
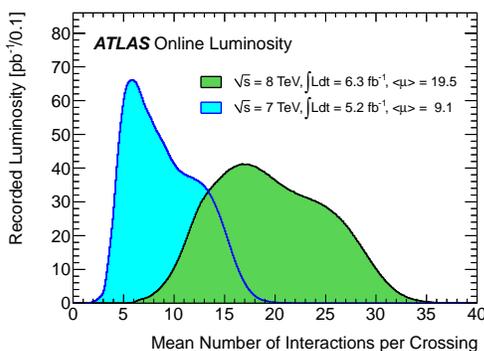
3. Impact of Pileup

Figure 2: Luminosity-weighted distribution of the mean number of interactions per crossing for the 2011 and 2012 data [6]. This shows the full 2011 run and 2012 data taken between April 4th and June 18th.
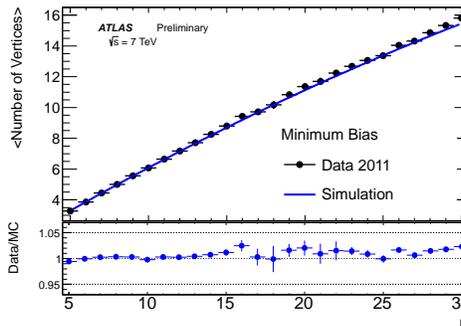
Figure 3: Average number of reconstructed primary vertices as a function of average number of pp interactions per bunch crossing measured for 2011 data [5].

Pileup presents a significant challenge for tracking and vertexing algorithms at ATLAS. Figure 2 compares the average number of interactions per bunch crossing, μ, in 5.2 fb^{-1} of 7 TeV data collected in 2011 to 6.3 fb^{-1} of 8 TeV data collected in 2012. The mean of this distribution is seen to more than double from 9.1 in 2011 to 19.5 in 2012. Figure 3 shows the number of primary vertices as a function of μ comparing 2011 data to simulation which is seen to reproduce the observed data distribution.

Pileup causes high occupancy in the ID which increases the rate of fake track reconstruction due to random combinations of tracking detector hits passing the reconstruction algorithms. This in turn leads to fake tracks contributing to primary vertex reconstruction and, at higher pileup, a non-negligible presence of fake primary vertices whose associated tracks are dominated by fakes. In 2012 data reconstruction, to mitigate these effects, a more robust track selection for primary vertex reconstruction has been developed. In simulation this is shown to lead to a moderate drop in primary vertex reconstruction efficiency of no more than 5% while significantly reducing the fake track fraction in primary vertices from around 30% to less than 10% and giving a negliglible rate fake vertex reconstruction up to μ=40 [7].
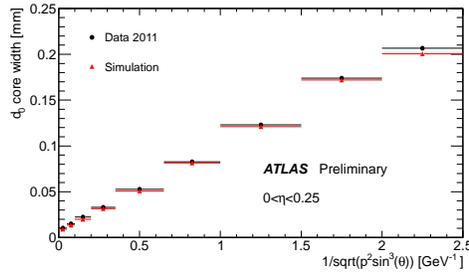
Figure 4: Transverse impact parameter, d_0, core width for data (black) and simulation (red) as a function of $1/p_T\sqrt{\sin\theta}$ for tracks with $0<\eta<0.25$. The core width is computed from a Gaussian fit to central region of the d_0 distribution [8].

For increased pileup a high density of charged particles in the Pixel detector can result in a single reconstructed cluster of hits originating from more than one charged particle interaction. At ATLAS pattern recognition from an artificial neural network (ANN) based clustering algorithm is used to identify and separate such merged clusters into distinct sub-clusters. This has the advantage of improving the pixel hit resolution resulting in an excellent impact parameter measurement. Figure 4 shows the transverse impact parameter resolution as a function of track kinematics in the central tracking region for 2011 data and simulation [8]. The high precision of the resolution in data is well modelled by the simulation.
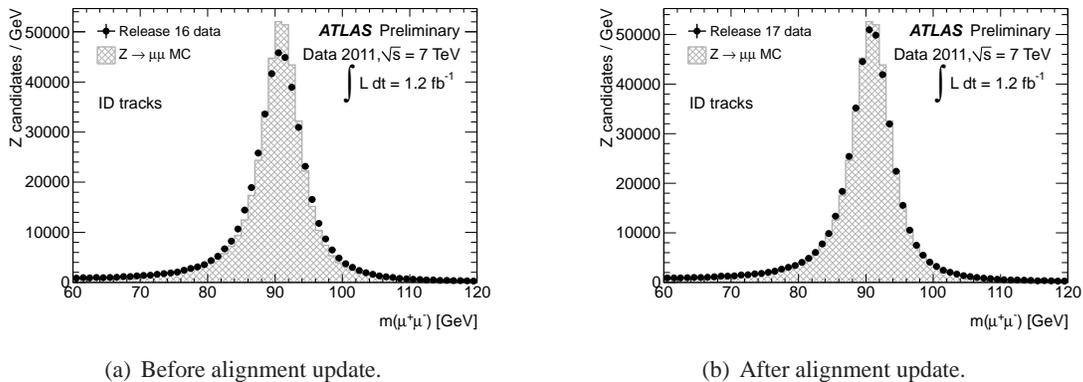
4. Detector Alignment

(a) Before alignment update.

(b) After alignment update.

Figure 5: Invariant mass distribution of $Z\to\mu\mu$ decays, where the mass is reconstructed using ID parameters of muon candidates only. Ideal alignment performance in MC is compared to observed performance of data before and after application of updated alignment calibration [9].

Critical to the performance in measuring track parameters is accurate knowledge of the ID alignment in the reconstruction algorithm. The alignment must be understood at different levels: the global alignment of the ID detectors, the alignment of barrel and endcaps, the alignment layer by layer and the alignment of individual modules. This is achieved at ATLAS through χ^2 minimisation of track-hit residuals; however, misalignment weak modes [9] exist which cannot be

constrained by the track residual method. Instead constraints are derived from analysing systematic bias in the invariant masses of well known particles reconstructed from pairs of opposite sign tracks (K_S^0, J/ψ, Z) as well as from the ratio of reconstructed electron energy measured in the AT-LAS calorimeters to the corresponding momentum measured in the ID [9]. Measurements of ID misalignment are propagated to the reconstruction geometry; Figure 5 shows the impact of updated alignment description on the Z mass resolution in $1.2\,\text{fb}^{-1}$ of 2011 data, a significant improvement is observed.
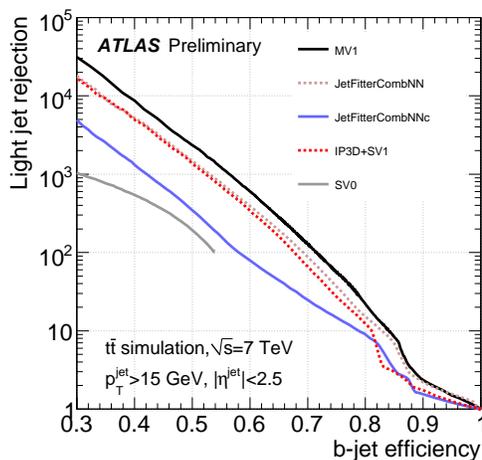
5. b-tagging Algorithms

Figure 6: Light-jet rejection as a function of the b-tagging efficiency in top pair MC simulation for all algorithms calibrated in 2011 data [11].

Many algorithms have been developed at ATLAS to identify jets originating from heavy flavour partons. These take advantage of the properties of reconstructed tracks from heavy flavour hadron decays by either directly reconstructing displaced secondary vertices or by exploiting the larger average impact parameter typical of such tracks with respect to the primary vertex. Figure 6 shows the performance of various ATLAS b-tagging algorithms in top MC simulation. The most commonly used algorithm in physics analyses is MV1, an ANN whose response is trained using both secondary vertex and track impact parameter properties as input. MV1 has the largest light jet rejection[3] for a given b-jet efficiency [10, 11].

6. b-tagging Performance

The performance of b-tagging algorithms is evaluated by measuring the efficiency to correctly identify jets originating from b partons in data as a function of jet p_T. Physics analyses then are required to correct the equivalent MC efficiency to match this measurement with an appropriate scale factor (SF). Efficiency measurements are performed primarily with two complementary methods

---

[3]Light jet rejection is the reciprocal of the fraction of jets labelled as light which pass the b-tagging algorithm [10].
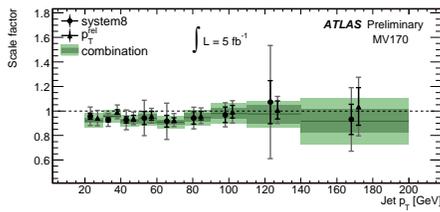
(a) SFs from muon based methods [11].
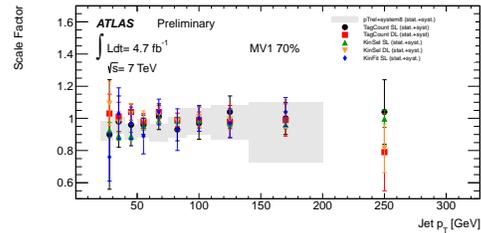
(b) SFs from top based methods [12].

Figure 7: Comparison of b-jet efficiency SFs measured in 2011 data for the MV1 70% b-tagging operating point

using tagged jets containing muons [11] and then cross-checked with methods using reconstructed top pair candidate events in the lepton plus jets channel [12]. Both the muon-jet and top samples have enhanced b-jet purity allowing efficiency measurements with an accuracy of 5 to 20% depending on b-tagging algorithm and jet p_T. Figure 7 compares the measured SFs from all b-jet efficiency methods in 2011 data for tagged jets identified using a MV1 operating point corresponding to 70% average b-jet efficiency in top MC events. It can be seen that within uncertainty the measured SFs are consistent with unity and that all methods are consistent with each other. The muon measurements are combined, accounting for correlations between the analyses, to provide a more accurate efficiency measurement.
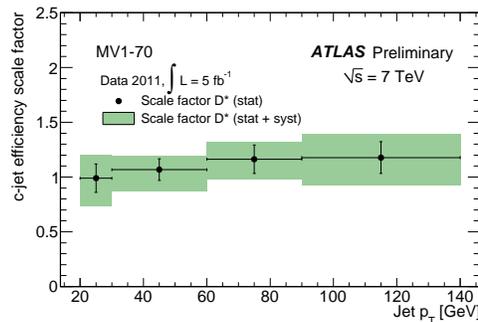

Figure 8: c-jet efficiency SFs measured with the D* method for the MV1 70% b-tagging working point [13].

In addition to b-jet efficiency, the algorithm efficiencies for jets originating from c quarks and the mistag rate for light jets are also measured in data. For the former a sample of jets containing exclusively reconstructed D* candidates provides efficiency measurements with accuracy of <25% [13]; for the latter two methods using inclusive jet samples are employed which show reasonable agreement and have uncertainties up to 100% depending on jet kinematics and b-tagging algorithm operating point [14]; such uncertainties are managable due to the high light jet rejection of b-tagging algorithms. Figure 8 shows the measured SFs for *c*-jets and Figure 9 the SFs for light jets when calibrating the same MV1 70% operating point; again consistentcy of the SFs with unity is observed within the uncertainties.
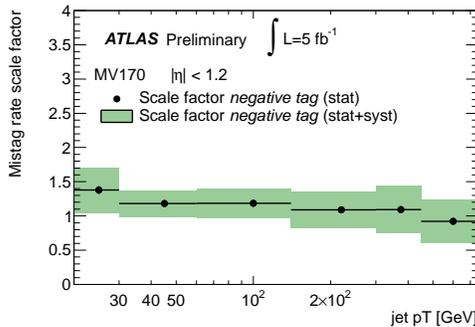
(a) Mistag rate SFs from negative tagging requirements.
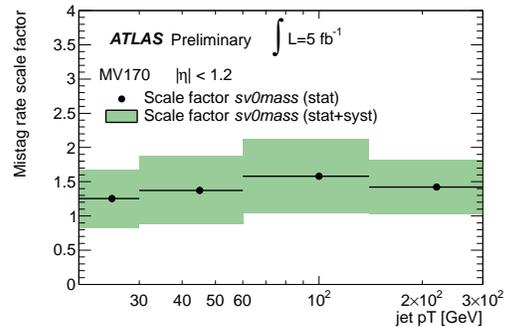
(b) Mistag rate SFs from template fits.

Figure 9: Mistag rate SFs measured in 2011 data for the MV1 70% b-tagging operating point for negative tag method which reverses tagger requirements to enrich the light jet sample and for a template fit method used as a cross-check [14].

7. Conclusions

ATLAS is maintaining excellent understanding of tracking, vertexing and b-tagging performance in the large integrated luminosities of proton-proton collisions recorded in 2011 and 2012. The impact of high pileup on tracking and primary vertex reconstruction has been studied and techniques implemented to maintain manageable fake rates are well described by simulation. Measurements of misalignment in the tracking detectors have been used to significantly improve physics performance. Multivariate b-tagging algorithms with excellent light jet rejection have been calibrated with the full 2011 data set and the first efficiency measurements from top events in 2011 data have been presented at this conference. Good consistency is observed between complementary methods for measuring the efficiencies and mistag rates of such algorithms in data.

References

[1] ATLAS Collaboration, 2008 JINST 3 S08004.

[2] L. Evans and P. Bryant (editors) 2008 JINST 3 S08001.

[3] ATLAS Collaboration, ATLAS-CONF-2010-072,
    http://cdsweb.cern.ch/record/1281363

[4] ATLAS Collaboration, ATLAS-CONF-2010-069,
    http://cdsweb.cern.ch/record/1281344

[5] http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/IDTRACKING/
    PublicPlots/ATL-COM-PHYS-2012-474/

[6] http://atlas.web.cern.ch/Atlas/GROUPS/DATAPREPARATION/
    InteractionsperCrossing/muplot/2012/mu_2011_2012.png

[7] ATLAS Collaboration, ATLAS-CONF-2012-042,
    http://cdsweb.cern.ch/record/1435196

[8] http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/IDTRACKING/
PublicPlots/ATL-COM-PHYS-2012-471/

[9] ATLAS Collaboration, ATLAS-CONF-2012-141,
http://cdsweb.cern.ch/record/1483518

[10] ATLAS Collaboration, ATLAS-CONF-2011-102,
http://cdsweb.cern.ch/record/1369219

[11] ATLAS Collaboration, ATLAS-CONF-2012-043,
http://cdsweb.cern.ch/record/1435197

[12] ATLAS Collaboration, ATLAS-CONF-2012-097,
http://cdsweb.cern.ch/record/1460443

[13] ATLAS Collaboration, ATLAS-CONF-2012-039,
http://cdsweb.cern.ch/record/1435193

[14] ATLAS Collaboration, ATLAS-CONF-2012-040,
http://cdsweb.cern.ch/record/1435194