ATLAS NOTE

ATL-PHYS-PUB-2016-007

March 18, 2016

Measurement of track reconstruction inefficiencies in the core of jets via pixel dE/dx with the ATLAS experiment using $\sqrt{s}$ = 13 TeV *pp* collision data

The ATLAS Collaboration

Abstract

The inefficiency of the ATLAS charged particle reconstruction in the core of jets is studied as a function of the transverse momentum of the jet between 200 GeV and 1600 GeV, using proton-proton data collected by the ATLAS experiment from the LHC in 2015 at a center-of-mass energy of $\sqrt{s}$ = 13 TeV. Quantifying such an inefficiency is important for numerous performance studies and physics analyses, including calibration of the jet energy scale and mass. A fully data driven method to measure the fraction of lost tracks using the energy loss, dE/dx, of charged particles in the pixel detector is presented. The fraction of lost tracks is at most 5% as a function of jet transverse momentum. The fraction of lost tracks estimated from Monte Carlo simulation samples are compared to the results and found to agree within 25%.

1 Introduction

The characterization of the performance of the ATLAS [1] track reconstruction algorithms in dense environments, like the core of high p_{T} jets, is important for a number of ongoing performance studies and physics analyses including calibration of the jet energy scale using charged particle quantities [2], calibration of the jet mass in large radius jets [3], as well as topologies with boosted τ's [4]. In many cases, the leading sources of systematic uncertainty are due to uncertainties on track reconstruction.

Tracks in the core of high p_{T} jets are more likely to be lost than more isolated tracks. In these environments, it is not uncommon for multiple particles to deposit energy in the same or nearby pixels in the ATLAS pixel detector, which results in a single merged cluster during track reconstruction. Methods to identify and resolve such merged measurements are implemented in the ATLAS software [5][6]. A residual inefficiency in reconstructing tracks, due to high density and collimation of charged particles in high p_{T} jets, remains. This study uses the ionization energy loss ($\mathrm{d}E/\mathrm{d}x$) to deduce the probability of losing a track, measure the residual inefficiency, and estimate the systematic uncertainty on the track reconstruction efficiency as measured in simulation due to lost tracks.

The ATLAS inner detector system [1] is designed to reconstruct the trajectories of charged particles[1], measure their momentum, and to perform vertex reconstruction in the pseudorapidity range of $|\eta| < 2.5$. It is the innermost detector of ATLAS and consists of a silicon pixel detector, a silicon microstrip detector (SCT) and the straw tubes of the transition radiation tracker (TRT), surrounded by a solenoid magnet that generates a 2 T axial magnetic field. During the first long shutdown (2013-2015), a new pixel layer, the insertable B-layer (IBL) [7] was added along with a new beam pipe, in place of the original beam pipe. The IBL consists of 14 staves instrumented with silicon pixel sensors in planar and 3D technology along their length of 664 mm. The four layers of the barrel are at a radius of 33.2, 50.5, 88.5 and 122.5 mm, while the new beam pipe is located approximately at a radius of 24.3 mm. The typical sensor dimensions in the pixel detector are 50 μm in the transverse direction and 400 μm in the longitudinal direction, while the pixels of the IBL have the same transverse dimension (50 μm) and a smaller longitudinal dimension of 250 μm.

A measure of the collected charge is available from the pixel detector via the time over threshold (ToT) [8] technique. ToT is a measure of the time the pulse spends above a given threshold and is approximately proportional to the charge. The $\mathrm{d}E/\mathrm{d}x$ of a charged particle traversing the pixel sensor is measured from the charge collected in the clusters associated with the reconstructed track, where a cluster is a group of pixel sensors associated together by a clustering algorithm [6]. With single particles and thin layers, one expects the $\mathrm{d}E/\mathrm{d}x$ measurements to follow a Landau distribution [9]. A single particle of LHC energy is expected to be a minimum ionizing particle (MIP). Thus, two particles contributing to the same cluster are expected to deposit double the energy of a single MIP.

Due to the magnetic field, the spatial separation of charged particles increases in the transverse direction as they traverse the detector. Thus, the probability of merging clusters is greater in the inner pixel layers.

[1] ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the center of the detector and the z-axis along the beam pipe. The x-axis points from the IP to the center of the LHC ring, and the y-axis points upwards. Cylindrical coordinates (R, ϕ) are used in the transverse plane, ϕ being the azimuthal angle around the z-axis. The pseudorapidity is defined in terms of the polar angle θ as $\eta = -\ln\tan(\theta/2)$. The angular distance is computed as $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

The IBL only has 4 bits available to encode the ToT information, while the second barrel layer, the B-layer, has 8 bits available to encode ToT information and thus provides an enhanced ToT resolution. As a result, this study uses the cluster $\mathrm{d}E/\mathrm{d}x$ values corresponding to the B-layer of the pixel detector.

Section 2 describes the data and simulation samples used. The method to measure the inefficiency is described in Section 3. Section 4 presents the measurement of the fraction of lost tracks and the inefficiency for data and simulation. The results of the study are summarized in Section 5.

2 Data and Simulation Samples

This study uses a sample of proton-proton collisions at $\sqrt{s} = 13$ TeV taken during 2015 corresponding to an integrated luminosity of 2.8 fb^{-1}. Events were selected from data using single-jet-triggers. The minimum jet trigger p_{T} threshold is 100 GeV. Each trigger was prescaled by a reduction factor depending upon the instantaneous luminosity and the jet energy it triggered on. This suppressed the number of low p_{T} jets while keeping all high p_{T} jet events with at least one jet with $p_{\mathrm{T}} > 1$ TeV, resulting in a more uniform jet p_{T} spectrum. Appropriate data quality requirements are applied to all data sets. Events are required to have at least one reconstructed primary vertex with at least three tracks.

Data is compared to an inclusive di-jet Monte Carlo sample generated with Pythia 8.186 [10]. Generator parameters are set according to the A14 tune for the parton shower and hadronization are taken from the NNPDF2.3LO PDF set [11]. Monte Carlo samples generated with Herwig++ 2.7.1 [12], and Sherpa 2.1 [13] are also studied. For Herwig++, parameters corresponding to the UEEE5 tune are used with the CTEQ6L1 PDF set [14] and for Sherpa, parameters corresponding to the CT10 PDF set [15] are used. The ATLAS detector response is simulated [16] using the GEANT4 framework [17]. Events from Monte Carlo simulation are reweighed to match the number of events which were triggered on in data.

3 Track Reconstruction Inefficiency Measurement

A single charged particle at the LHC has the $\mathrm{d}E/\mathrm{d}x$ distribution of a MIP. When multiple particles traveling closely through the detector contribute to the same cluster, the measured cluster $\mathrm{d}E/\mathrm{d}x$ is compatible with multiple MIPs, resulting in a $\mathrm{d}E/\mathrm{d}x$ distribution that is distinctive compared to a single particle. By fitting the cluster $\mathrm{d}E/\mathrm{d}x$ for reconstructed tracks near the core of the jet, tracks reconstructed from single particles can be statistically separated from tracks reconstructed from multiple particle contributions. This is due to the fact, that near the jet core the charged particle density is high and particles can be heavily collimated. The tracks of these particles are thus more likely to have been reconstructed from merged clusters. The fraction of lost tracks can be inferred from the number of reconstructed tracks associated with a cluster $\mathrm{d}E/\mathrm{d}x$ compatible with multiple MIPs. By measuring the fraction of lost tracks in data and simulation, the residual track reconstruction inefficiency can be estimated.

3.1 Jet and Track Selection

Jets are reconstructed from topological clusters [18] using the anti-k_t algorithm [19] with a distance parameter $R = 0.4$. Jets are selected requiring a minimum jet p_{T} of 200 GeV and $|\eta^{\mathrm{jet}}| < 2.5$. These

jets are calibrated to the hadronic energy scale using $\sqrt{s} = 13$ TeV simulation [20]. It was previously demonstrated that the properties of jets in simulation agree well with data [21].

Tracks are reconstructed in the inner detector using an iterative algorithm seeded on combinations of measurements from the silicon detectors and combining a combinatorial Kalman filter [22, 23] with a stringent ambiguity solver. Reconstructed tracks are selected by the following requirements:

- exactly one pixel hit per layer
- $p_{\mathrm{T}}^{\mathrm{trk}} > 10$ GeV
- $|\eta^{\mathrm{trk}}| < 1.2$
- $|d_0^{\mathrm{BL}}| < 1.5$ mm
- $|z_0^{\mathrm{BL}} \sin\theta| < 1.5$ mm
- number of SCT hits ≥ 6
- number of pixel holes[2] ≤ 1

where $p_{\mathrm{T}}^{\mathrm{trk}}$ is the transverse track momentum, d_0^{BL} is the transverse impact parameter calculated with respect to the measured beam line position, z_0^{BL} is the difference between the longitudinal position of the track along the beam lime at the point where d_0^{BL} is measured and the longitudinal position of the primary vertex, and θ is the polar angle of the track.

3.2 Fit Method

A cluster $\mathrm{d}E/\mathrm{d}x$ distribution of tracks inside the jet core is fit using two $\mathrm{d}E/\mathrm{d}x$ template distributions: a single-track template of the $\mathrm{d}E/\mathrm{d}x$ distribution made from tracks reconstructed from a cluster where a single particle contributed, and a multiple-track template of the $\mathrm{d}E/\mathrm{d}x$ distribution made from tracks reconstructed from a merged cluster, to which multiple particles contributed.

The method is designed to create the templates in a purely data driven way. This is achieved by applying the selections illustrated in Figure 1, where a *Not-Multiply-Used* cluster is defined as a cluster associated with exactly one track candidate and a *Multiply-Used* cluster is defined as a cluster associated with two or more track candidates.

The data distribution is created from tracks inside the jet core (ΔR(jet,trk)< 0.05) that were reconstructed from a Not-Multiply-Used cluster. The single-track template is taken from tracks reconstructed from Not-Multiply-Used clusters that are outside the jet core (ΔR(jet,trk)> 0.1). It is expected that outside the jet core, the contribution of lost tracks is negligible. The multiple-track template is taken from tracks reconstructed from Multiply-Used pixel clusters that are inside the jet core.

Examples of the resulting distributions are shown in Figure 2. The single-track template, displayed as blue circles in Figure 2(a), contains a single peak at the $\mathrm{d}E/\mathrm{d}x$ value expected for a minimum ionizing particle traversing the B-layer of the pixel detector. The multiple-track template, displayed as green squares in the same figure, instead exhibits a peak in the $\mathrm{d}E/\mathrm{d}x$ range of two particles. A third, smaller peak occurs at $\mathrm{d}E/\mathrm{d}x > 3.2$ MeVg^{-1}cm^2 for clusters created by three particles.

[2] Holes are defined as intersections of the reconstructed track trajectory with a sensitive detector element that do not result in a hit. These are estimated by following closely the track trajectory and comparing the hits-on-track with the intersected modules. Inactive modules or regions such as edge areas on the silicon sensors are excluded from the hole definition.

The fraction of tracks containing a B-layer hit, which is consistent with multiple particles crossing but not used by other reconstructed tracks, is determined by fitting the data distribution, an example of which is shown in Figure 2(b), with both templates. Collimated particle pair simulation studies conducted using pseudo-tracking[3] have shown that the multiple track template is correlated with the dE/dx distribution in reconstructed tracks overlapping with lost tracks and that selecting tracks reconstructed from Multiply-Used clusters matching this template is consistent with selecting lost tracks. Thus, the fit fraction of the multiple-track template is called fraction of lost tracks, F^{lost}.

To minimize the effect of clusters created by three particle, the fit was performed over the range 0.67–3.07 (0.8–3.2) MeVg^{-1}cm^2 for data (simulation). The ranges were chosen to have the same fraction of clusters inside the fit range with respect to all clusters in the distribution.

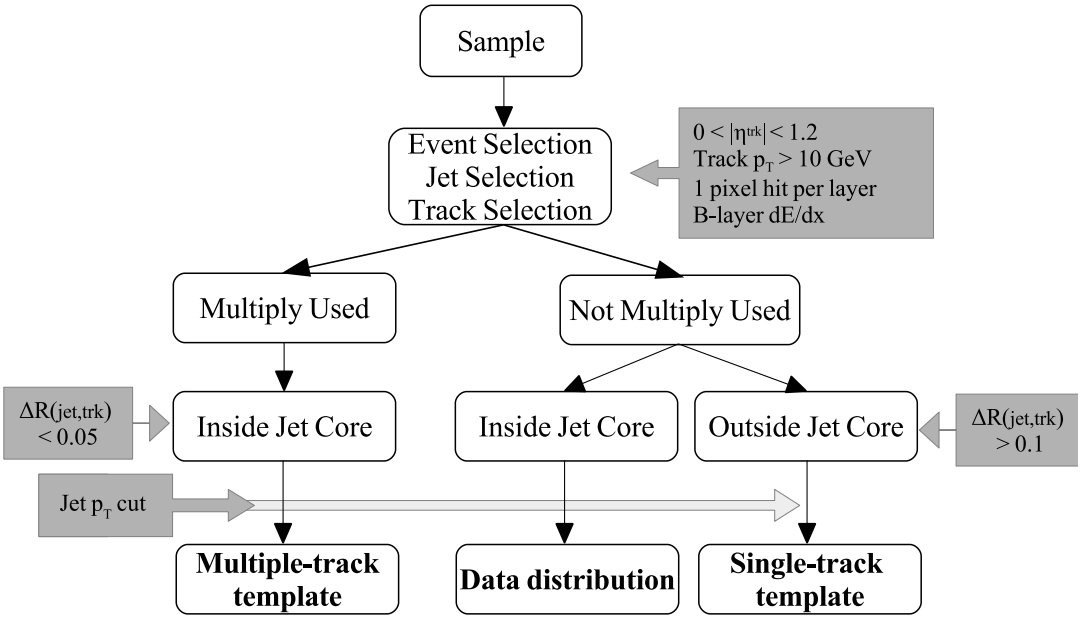
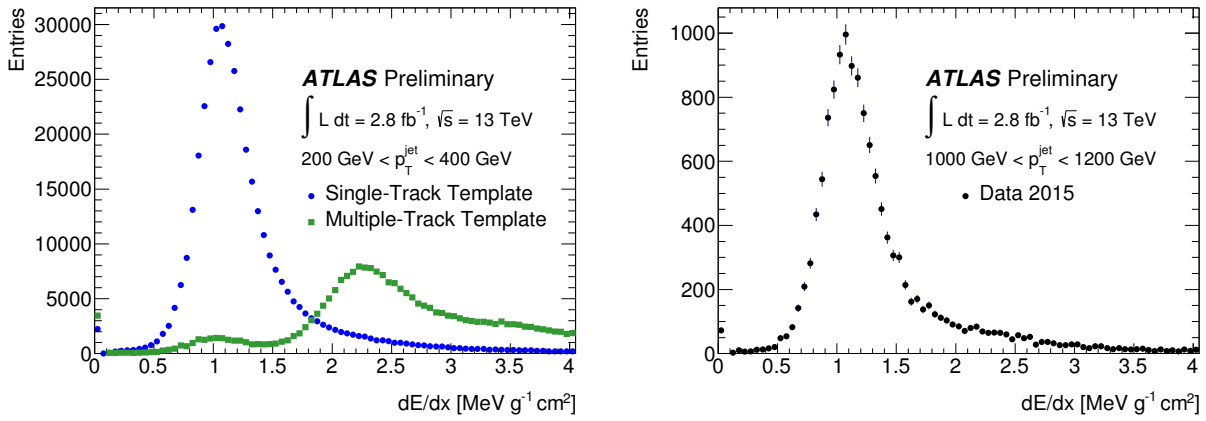


Figure 1: Definition of the templates and dE/dx distributions for data and simulation.

To study the dependence of lost tracks on jet p_{T}, the fit was performed in seven different bins of jet p_{T} ranging from 200 GeV to 1600 GeV in steps of 200 GeV in both data and simulation.

The measurement was performed both on data and simulation samples. For simulation, separate templates were constructed for each jet p_{T} bin. For data, the single-track and multiple-track templates were derived from the lowest jet p_{T} bin, shown in Figure 2(a), due to low statistics at higher jet p_{T}. It was verified that within the statistical uncertainty of the high p_{T} bins, the templates derived from the lowest jet p_{T} bin have the same shape within the fit range. An additional check using a high-statistics simulation is discussed in Section 3.3.

[3] Pseudo-tracking is a tool that reconstructs tracks with ideal performance by directly fitting the hits from the truth particle.



(a) (b)

Figure 2: (a) Single-track and multiple-track templates for data and (b) the dE/dx distribution for Not-Multiply-Used clusters near the jet core (ΔR(jet,trk)$<$ 0.05) for data with a jet p_{T} in the range 1000 GeV $< p_{\mathrm{T}}^{\mathrm{jet}} <$ 1200 GeV.

3.3 Systematic Uncertainty

Various potential sources of uncertainty were studied. For simulation, generator dependency provides the dominant systematic uncertainty. This uncertainty was estimated by comparing the fit results on simulated samples made with Pythia8, Sherpa and Herwig++ event generators as shown in Figure 3. For each jet p_{T} bin, the largest difference between the fit fractions of the three generators was taken to be the systematic uncertainty for that jet p_{T} bin. This systematic uncertainty was symmetrized and applied to all simulation results discussed in this note. The relative systematic uncertainties due to generator dependency on the fraction of lost tracks, F^{lost}, can be seen in Table 1.

The measured fraction of lost tracks, F^{lost}, varies as a function of the range in dE/dx for which the distribution is fit. The effect has been estimated by increasing the fit range beyond the baseline selection. The lower edge of the fit range was chosen to include all statistically significant bins and results are found to be stable under its variation. The fitting process was repeated for six different ranges with the upper edge increasing in 0.2 MeVg^{-1}cm^2 increments. For both data and simulation, F^{lost} changes as a function of the fit range, increasing on order 5% over its previous value with each increment. However a significant effect is observed only for the data, and a symmetric uncertainty, equal to the maximum change in F^{lost}, was applied to each jet p_T bin. Relative systematic uncertainty values are shown in Table 1.

The second leading systematic uncertainty for data is the result of fitting all data jet p_{T} bins with the templates from the lowest jet p_{T} bin, as mentioned in Section 3.2. The additional check performed with a high-statistics simulation showed a small bias in F^{lost} due to the fraction of tracks reconstructed from ≥ 3 contributing particles, relative to the two particle contribution in the multiple-track template, which varies with jet p_{T}. To account for this bias, a p_{T} dependent multiplicative correction was determined by comparing the F^{lost} values for simulation fit with templates from the corresponding jet p_{T} bin with the results of simulation fit with templates from the lowest jet p_{T} bin. This correction term was applied to data F^{lost} values after all fitting procedure was complete. In addition, the difference between the two simulation F^{lost} values compared for the correction factor was also included as a systematic uncertainty. The corresponding relative values of which can be seen in Table 1.

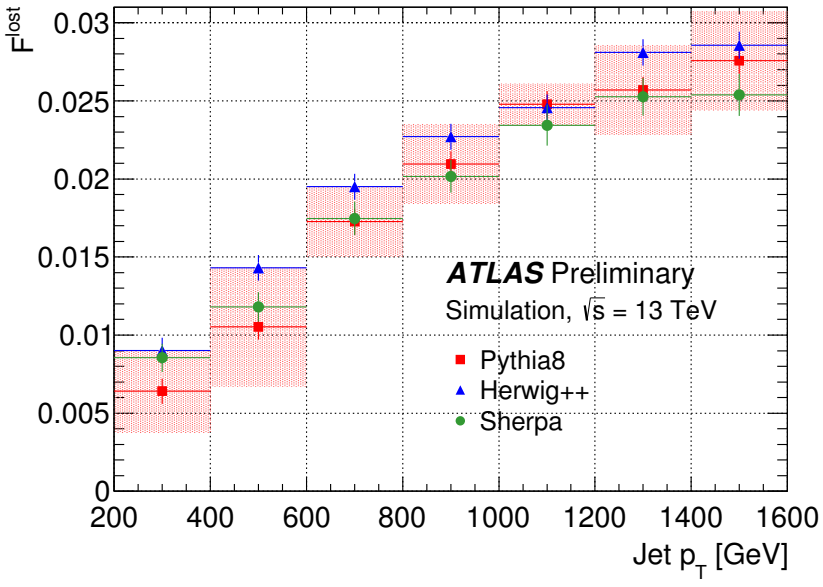


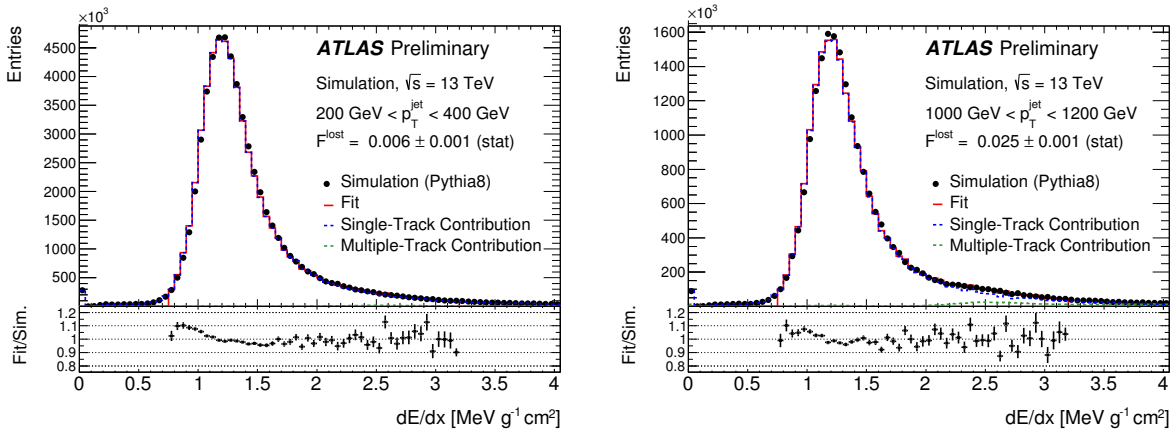
Figure 3: The fraction of lost tracks, F^{lost} , as a function of jet p_{T} for simulation with Pythia8 (red squares), Herwig++ (blue triangles), and Sherpa (green circles). Error bars indicate statistical uncertainty while the red error band indicates the jet p_{T} dependent systematic uncertainty applied to the Pythia8 simulation as a result of F^{lost} generator dependency.

	Data		Simulation
Jet p_{T} bin	Fit Range	Low p_{T} Templates	Generator Dependency
200–400 GeV	0.17	0.00	0.41
400–600 GeV	0.18	0.14	0.36
600–800 GeV	0.16	0.13	0.13
800–1000 GeV	0.23	0.10	0.12
1000–1200 GeV	0.12	0.11	0.05
1200–1400 GeV	0.13	0.17	0.11
1400–1600 GeV	0.25	0.11	0.12

Table 1: Relative values of leading systematic uncertainties on the fraction of lost tracks for data and simulation (Pythia8) in bins of jet p_{T}.

4 Results

Figures 4 and 5 show the fit results for simulation and data in two bins of jet p_{T} respectively. The single-track and multiple-track dE/dx templates provide a good description to the dE/dx distribution for both data and simulation.



(a) (b)

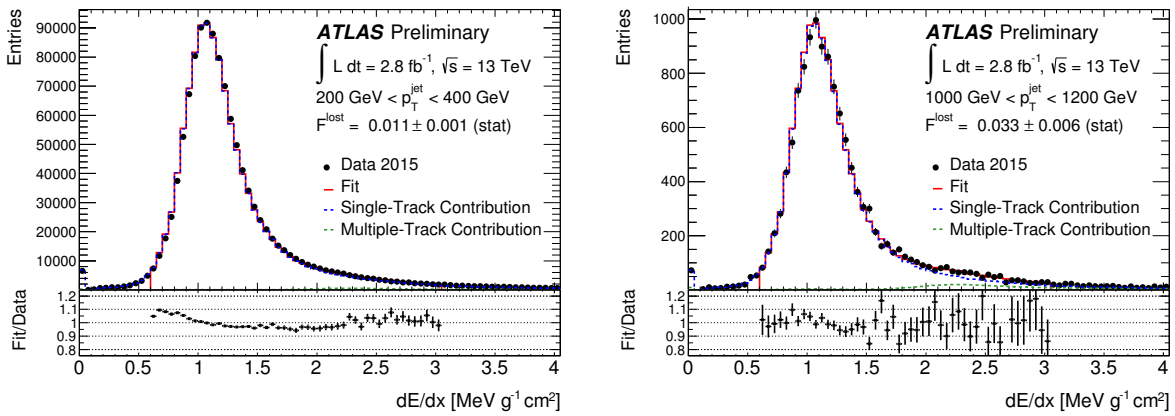
Figure 4: Simulation (Pythia8) dE/dx distributions (black circles) with fit results (red solid line) in two bins of jet p_{T}. The Single-Track template scaled by 1-F^{lost} is shown as the Single-Track Contribution (blue dashed line) and the Multiple-Track template scaled by F^{lost} is shown as the Multiple-Track Contribution (green dashed line). The bottom panel in each plot shows the ratio of fit/simulation within the fit range (0.8–3.2 MeVg^{-1}cm^2).

A comparison of F^{lost} as a function of jet p_{T} for data and simulation is shown in Figure 6. As the jet p_{T} bin increases, so does F^{lost} , with a similar trend observed in both data and simulation.

Figure 7 shows the ratio of F^{lost} as determined in data and simulation. Fitting a constant value results in a discrepancy between data and simulation of 25 ± 7% (stat) plus a systematic uncertainty in the range of ±15–75% depending upon jet p_{T} bin. The weighted average was calculated taking into account only the statistical uncertainties of both data and simulation.

5 Conclusions

A measurement of the fraction of lost tracks using a novel technique based on pixel cluster dE/dx and data from 13 TeV pp collisions was presented. The fraction of lost tracks in the core of jets was found to range from 1% to 5% as the jet p_{T} bin increases from 200 GeV to 1600 GeV. The relative discrepancy between data and simulation for the track reconstruction inefficiency was found to be a independent of jet p_{T} with a value of approximately 25%.



(a) (b)

Figure 5: Data dE/dx distributions (black circles) with fit results (red solid line) in two bins of jet p_{T}. The Single-Track template scaled by 1-F^{lost} is shown as the Single-Track Contribution (blue dashed line) and the Multiple-Track template scaled by F^{lost} is shown as the Multiple-Track Contribution (green dashed line). The bottom panel in each plot shows the ratio of fit/data within the fit range (0.67–3.07 MeVg^{-1}cm^2).

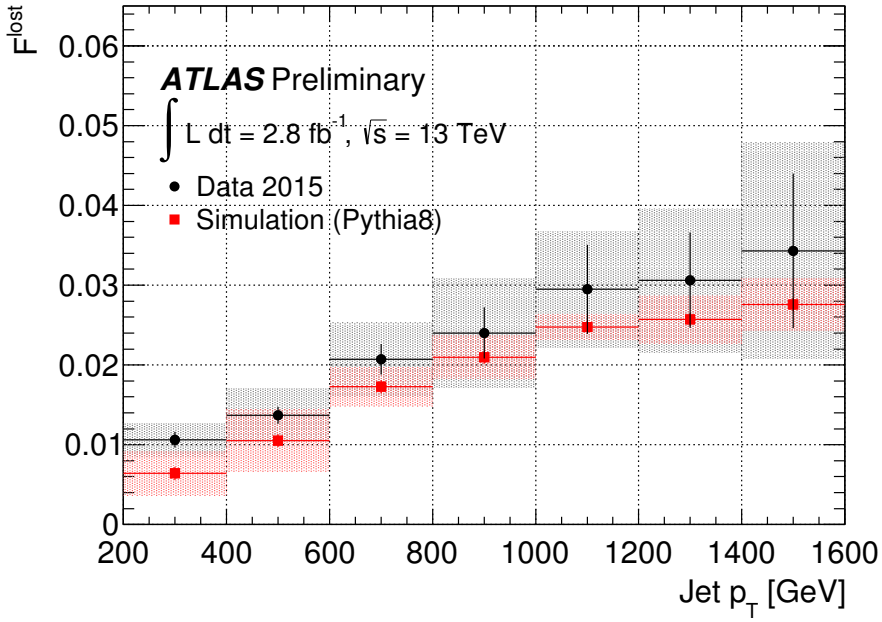


Figure 6: The measured fraction of lost tracks, F^{lost}, in the jet core (ΔR(jet,trk)$<$ 0.05) as a function of jet p_{T} for data (black circles) and simulation (red squares). Black error bars indicate statistical uncertainty while the grey and red error bands indicate the total uncertainty for data and simulation respectively.

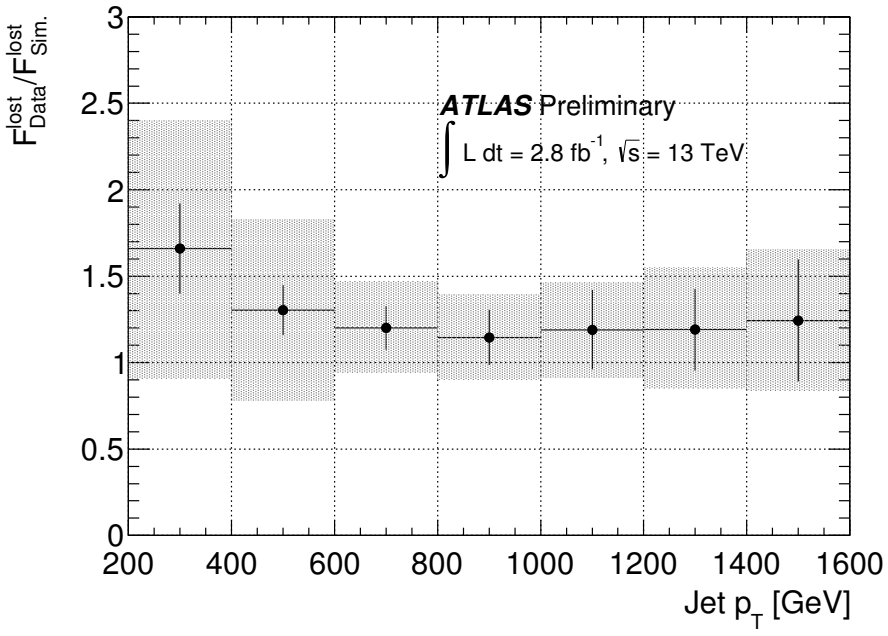


Figure 7: The ratio of the fraction of lost tracks, F^{lost}, in data with respect to simulation (Pythia8) as a function of jet p_{T}. Black error bars indicate the combined statistical uncertainty of data and simulation while the grey error band indicates the total uncertainty, taking into account the statistical and systematic uncertainties of both data and simulation.

References

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

[2] ATLAS Collaboration, *Jet energy resolution in proton-proton collisions at $\sqrt{s} = 7$ TeV recorded in 2010 with the ATLAS detector*, Eur. Phys. J. **C73** (2013) 2306, arXiv: 1210.6210 [hep-ex].

[3] ATLAS Collaboration, *Jet mass and substructure of inclusive jets in $\sqrt{s} = 7$ TeV pp collisions with the ATLAS experiment*, JHEP **05** (2012) 128, arXiv: 1203.4606 [hep-ex].

[4] ATLAS Collaboration, *A search for high-mass resonances decaying to $\tau^+\tau^-$ in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, JHEP **07** (2015) 157, arXiv: 1502.07177 [hep-ex].

[5] ATLAS Collaboration, *The Optimization of ATLAS Track Reconstruction in Dense Environments*, ATL-PHYS-PUB-2015-006, 2015, URL: https://cds.cern.ch/record/2002609.

[6] ATLAS Collaboration, *A neural network clustering algorithm for the ATLAS silicon pixel detector*, JINST **9** (2014) P09009, arXiv: 1406.7690 [hep-ex].

[7] ATLAS Collaboration, *ATLAS Insertable B-Layer Technical Design Report*, CERN-LHCC-2010-013. ATLAS-TDR-19, 2010, URL: https://cds.cern.ch/record/1291633.

[8] ATLAS Collaboration, *ATLAS pixel detector electronics and sensors*, JINST **3** (2008) P07007.

[9] K. A. Olive et al., *Review of Particle Physics*, Chin. Phys. **C38** (2014) 090001.

[10] T. Sjöstrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv: 0710.3820 [hep-ph].

[11] NNPDF Collaboration, *Parton distributions with LHC data*, Nucl. Phys. **B867** (2013) 244, arXiv: 1207.1303 [hep-ph].

[12] M. Bahr et al., *Herwig++ Physics and Manual*, Eur. Phys. J. **C58** (2008) 639, arXiv: 0803.0883 [hep-ph].

[13] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, arXiv: 0811.4622 [hep-ph].

[14] J. Pumplin et al., *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **07** (2002) 012, arXiv: hep-ph/0201195 [hep-ph].

[15] H.-L. Lai et al., *New parton distributions for collider physics*, Phys. Rev. **D82** (2010) 074024, arXiv: 1007.2241 [hep-ph].

[16] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. **C70** (2010) 823, arXiv: 1005.4568 [physics.ins-det].

[17] S. Agostinelli et al., *GEANT4: A Simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.

[18] ATLAS Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1* (2016), arXiv: 1603.02934 [hep-ex].

[19] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP **0804** (2008) 063, arXiv: 0802.1189 [hep-ph].

[20] ATLAS Collaboration, *Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at* $\sqrt{s} = 13$ *TeV*, ATL-PHYS-PUB-2015-015, 2015, URL: https://cds.cern.ch/record/2037613.

[21] ATLAS Collaboration, *Properties of Jets and Inputs to Jet Reconstruction and Calibration with the ATLAS Detector Using Proton-Proton Collisions at* $\sqrt{s} = 13$, ATL-PHYS-PUB-2015-036, 2015, URL: https://cdsweb.cern.ch/record/2044564.

[22] R. E. Kalman, *A New Approach to Linear Filtering and Prediction Problems*, Transactions of the ASME–Journal of Basic Engineering **82** (1960) 35.

[23] R. Fruhwirth, *Application of Kalman filtering to track and vertex fitting*, Nucl.Instrum.Meth. **A262** (1987) 444.