PROCEEDINGS OF SCIENCE

# Performance of charged-particle reconstruction within energetic jets in ATLAS for Run 3 data taking

**Otilia Anamaria Ducu** [a,*] **and On behalf of the ATLAS Collaboration**

[a] *Horia Hulubei National Institute of Physics and Nuclear Engineering, IFIN-HH, Bucharest, Romania*

*E-mail:* oducu@cern.ch

In preparation for LHC Run 3 data taking the track reconstruction algorithms for the ATLAS Inner Detector have been optimized with a particular focus on minimizing the number of wrongly reconstructed and low-quality (fake) tracks by rejecting them as early in the reconstruction chain as possible. This ensures the availability of high quality tracks for use in the reconstruction of higher-level objects and in physics analyses, a key aspect for the success of the ATLAS physics programme. Excellent reconstruction of charged particle tracks inside dense hadronic environments can improve the jet energy and mass resolutions, and it plays a crucial role in the hadronically decaying tau reconstruction, boosted object tagging (e.g. $H \to b\bar{b}$), or for the quark – gluon jets (or for light jets - bottom/charm quark jets) separation.

The performance of the improved track reconstruction algorithms has been validated on the full Run 2 dataset ($\sqrt{s}$ = 13 TeV, 139 fb$^{-1}$) collected with the ATLAS detector [1] and on simulated di-jet data generated with the Pythia 8 Monte Carlo (MC) generator with a particular focus on the track reconstruction efficiency and the rate of fake tracks in the dense environments found in the core of hadronic jets. The track reconstruction efficiency is measured in data and MC simulation using two variables of interest. The $r_{\text{track}}$ variable, which is a measure of the ratio of the charged to neutral energy inside a jet, and the $\zeta$ variable that quantifies the angular separation between pairs of jets. Both variables are measured as a function of the jet transverse momentum ($p_T$). These variables are sensitive to the reduction of track reconstruction efficiency that occurs in dense environments where tracks are more likely to overlap, thus to share hits in the pixel or silicon detectors. The rate of falsely reconstructed (fake) tracks in the core of jets is also measured in a dijet control region enriched in fake tracks. Using MC simulation, the origin of the fake tracks in the core of jets is studied as well.

The data and MC simulation used to get the results presented in this document were processed with the new ATLAS software release that will be used for Run 3 data taking. The methods employed are discussed in detail in Ref. [2].

ATL-PHYS-PROC-2022-055
20 August 2022

*41st International Conference on High Energy physics - ICHEP2022*
*6-13 July, 2022*
*Bologna, Italy*

[*] Speaker

## 1. The ATLAS detector and the charged-particle tracks reconstruction

The ATLAS detector is a multi-purpose particle detector with nearly $4\pi$ coverage in solid angle[1]. The tracks of charged particles are measured in a 2 T axial field generated by a solenoid magnet which surrounds the inner detector (ID) consisting of silicon pixels, silicon micro-strips tracker (SCT), and a transition radiation tracking detector (TRT). The innermost pixel layer is 33 mm from the beamline.

Charged-particle tracks are reconstructed using all three ID subsystems with a full coverage in $\phi$, $|\eta| < 2.5$ and $p_{\mathrm{T}} > 500$ MeV. It is using the so called 'inside-out' track reconstruction algorithm [3], designed to reconstruct primary tracks with high efficiency and low fake rate. Tracks are reconstructed using a fit with five parameters: the transverse ($d_0$) and longitudinal ($z_0$) impact parameters, $\phi$, $\theta$, and $Q/p$, where $Q$ is the charge and $p$ is the track momentum. To suppress the impact of multiple overlapping $pp$ collisions (pileup), tracks are required to originate from the primary collision vertex (PV). A series of additional criteria improves the quality of the selected tracks [4, 5]. The `Loose` track selection, used to get the results presented in this document, corresponds to the default track requirements applied during 'inside-out' track reconstruction and aims to obtain highly efficient charged particle reconstruction (~90% for pion tracks with $p_T \geq 5$ GeV) and low fraction of fake tracks.

Jets are clustered using the anti-$k_t$ jet algorithm with radius parameter $R = 0.4$, and it uses as inputs particle-flow objects, combining tracking and calorimetric information. An overall jet calibration corrects the detector-level jet to the particle-level jet. Jets are required to have $|\eta| < 2.5$, and tracks are matched to jets using the ghost-association technique.

## 2. Efficiency measurements: the $r_{\mathrm{track}}$ and $\zeta$ methods

The performance of charged particle track reconstruction can be quantified by four metrics: the reconstruction efficiency, the fake rate, and the defining parameter biases and resolutions [2]. One can define the reconstruction efficiency as the probability to have a charged particle reconstructed as a track, and contributions to an inefficiency in the track reconstruction are discussed further. Some sources are coming from the multiple scattering, and any other effects related to material interactions. These affect mainly the charged particles with $p_T$, and have a more significant effect in the detector end-caps (where there is more material).

Another contribution to the track reconstruction inefficiency is due to a high local density of charged particles. When the trajectories of the charged particle are close enough in space, the hits they leave in the detector components can result in a single reconstructed (merged) cluster. When outside of the jet core, clusters that are shared by too many tracks, as well as tracks that share too many clusters, are assumed to come from fake tracks and are rejected. However, e.g. inside the jet core the probability to have very nearby charged particles is significant. These leave real tracks that share clusters, and lead to reconstruction inefficiencies.

[1]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. Polar coordinates $(r, \phi)$ are used in the transverse plane, $\phi$ being the azimuthal angle around the beam pipe. The pseudorapidity is defined in terms of the polar angle $\theta$ as $\eta = -\ln \tan(\theta/2)$. The angular separation between two objects A and B is defined as $\Delta R(A, B) = \sqrt{\Delta\phi(A, B)^2 + \Delta\eta(A, B)^2}$. The transverse momentum and energy are defined in the $x$–$y$ plane as $p_{\mathrm{T}} = p \cdot \sin(\theta)$ and $E_{\mathrm{T}} = E \cdot \sin(\theta)$.

A measure of the track reconstruction efficiency is obtained with the $r_{\text{track}}$ method which relies on the observation that, at particle level, the ratio of the charged to neutral energy in the jet core is independent of the jet $p_T$. However, at detector level this ratio is $p_T$ dependent, as shown in Figure 1(a). To obtain these results, the $r_{\text{track}}$ variable is computed as the average ratio of the sum of the $p_T$ of tracks inside the jet and the jet $p_T$, and is shown as a function of the jet $p_T$ normalized to its value at 500 GeV. The significant decrease (of around 30%) of $\mathrm{r}_{\text{track}}$ with jet $p_T$ is due to the loss of track reconstruction efficiency in the jets core due to pixel and SCT cluster merging. Note that for this measurement only jets with $\mathrm{r}_{\text{track}} < 1$ are considered, and the "inside jet" condition is achieved by requiring the $\Delta R$ between the selected track and the jet, $\Delta R(\text{jet}, \text{track})$, to be less than 0.02. The observed data versus MC simulation disagreement is due to modeling of material interactions, and is consistent with the past results shown in Ref. [2].

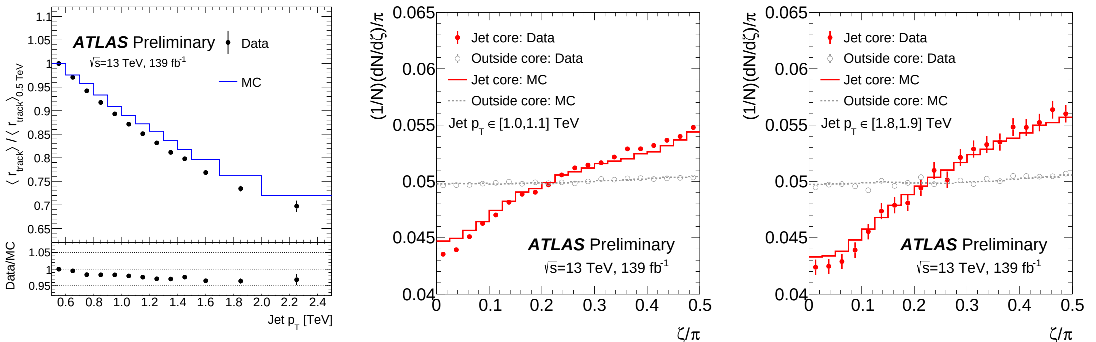


**Figure 1:** Left, 1(a): the $r_{\text{track}}$ distribution as a function of the jet $p_T$ normalised to its value at 500 GeV. Middle, 1(b): the distribution of $\zeta$ for 1 TeV $< p_T^{\text{jet}} <$ 1.1 TeV pairs of tracks that are inside ($\Delta R(\text{jet}, \text{track}) <$ 0.02) and outside ($\Delta R(\text{jet}, \text{track}) >$ 0.1) of the jet core, as a function of the jet $p_T$. Right, 1(c): same distribution as in Figure 1(b), but using jets in the [1.8, 1.9] TeV $p_T$ interval. Results are shown in data and MC simulation. No reweighting to correct the jet and track $p_T$ distributions in MC simulation to data is applied. Only the statistical uncertainties are shown. Figures taken from Ref. [6].

Another method used to measure the track reconstruction efficiency is the $\zeta$ method. It uses the angular separation $\zeta$ between track pairs, and exploits the asymmetry in the pixel dimensions. In the (innermost) pixel layer, the most common pitch in the direction transverse to the beam is 50 $\mu$m, where in the longitudinal direction is (400 $\mu$m) 250 $\mu$m. Thus, clusters from pairs of tracks that are a fixed distance apart are more likely to be merged if separated in the long direction compared to the short direction. The obtained results are shown in Figures 1(b) and 1(c) for two representative jet $p_T$ intervals, and illustrate the performance of the track reconstruction algorithm in dense environments. The quantity $\zeta$ is defined as $\zeta = |\tan^{-1}(\Delta\phi/\Delta\eta)|$, where $\Delta\phi$ and $\Delta\eta$ are the angular separation between any pair of tracks belonging to the same region of the jet. The radiation pattern inside a jet is approximately symmetric around the jet axis and hence follows a uniform distribution for track pairs outside the dense jet core ($\Delta R(\text{jet}, \text{track}) > 0.1$). Inside the jet core ($\Delta R(\text{jet}, \text{track}) < 0.02$), some particle pairs that share pixel clusters cannot be reconstructed. More track pairs are lost due to cluster merging if the particles are close-by in the $\phi$ direction due to the asymmetry in the pixel pitch, leading to a visible slope in $\zeta$ in the jet core [6].

## 3. A measure of the modelling of fake tracks in data, in the core of jets

A fake track can be defined as the probability to have the majority of energy deposits (or hits) used for a reconstructed track not coming from any single charged particle. In MC simulation, the real and fake tracks can be separated using a weighted matching probability $P_{match}$ that is defined using the ratio of the number of hits which are common to a given track and the corresponding truth particle ($\mathrm{N}_{\mathrm{common}}^{\mathrm{Pixel;SCT;TRT}}$) and the number of hits which form the track ($\mathrm{N}^{\mathrm{Pixel;SCT;TRT}}$) [7]:

$$P_{match} = \frac{10 \times \mathrm{N}_{\mathrm{common}}^{\mathrm{Pixel}} + 5 \times \mathrm{N}_{\mathrm{common}}^{\mathrm{SCT}} + \mathrm{N}_{\mathrm{common}}^{\mathrm{TRT}}}{10 \times \mathrm{N}^{\mathrm{Pixel}} + 5 \times \mathrm{N}^{\mathrm{SCT}} + \mathrm{N}^{\mathrm{TRT}}} \quad (1)$$

If $P_{match} > 50\%$ ($< 50\%$) the selected track is assumed to be real (fake).

In the dense environment of the jets core, is it found that the rate of fake tracks due to associating wrong hits from nearby high $p_T$ charged particles is not negligible, and an equally important source is given by the random combinations of hits (and their multiplicity increases with pileup). This is illustrated in Figures 2(a) and 2(b). These results (and the following ones) are obtained using a dijet event selection, where the leading jet $p_T$ is in the [460, 600] GeV interval. Only the tracks inside the leading jet are used ($\Delta R$(jet, track) $< 2$), with $p_T > 1$ GeV. To reject real tracks, cuts are applied on the number of TRT and SCT hits, and on track $\sin(\theta)$ variables. This is called pre-selection.

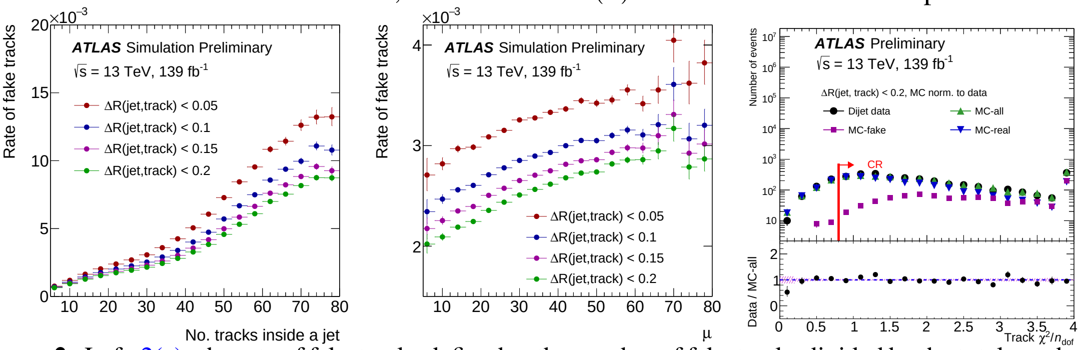


**Figure 2:** Left, 2(a): the rate of fake tracks defined as the number of fake tracks divided by the total number of tracks within the jet core as a function of the number of tracks forming the jet, and (middle, 2(b)) as a function of the average number of interactions per bunch crossing, $\mu$. Results are obtained using MC. Distributions obtained for several $\Delta R$(jet, track) intervals. Right 2(c): the distributions of the track $\chi^2/n_{\mathrm{dof}}$ parameter, close to and in the dijet CR. Hashed magenta area in the lower pad for the MC statistical uncertainty. Results are obtained using data and MC simulation. The MC distributions are also plotted separately for real (down-triangles) and fake (squares) tracks. Last bin includes overflow. Only the statistical uncertainties are shown. Figures taken from Ref. [6].

The rate of fake tracks in data is obtained from a fit performed in a dijet control region (CR) enriched in fake tracks, using the shape of the real and fake tracks distributions from MC simulations. The CR is build on top of the pre-selection requirements. To define it, dedicated studies using MC simulation are carried out to identify variables able to discriminate the fake tracks among the selected ones [2]. As a result of the optimization, the CR is defined with the following variables and requirements: track $d_0 > 0.2$ mm, $\sin\theta \times |z_0^{PV}| > 0.3$ mm, track $\chi^2/n_{\mathrm{dof}} > 0.8$[2], $n_{\mathrm{SCT}} < 8$ and $n_{\mathrm{TRT}} < 30$. Figure 2(c) show the track $\chi^2/n_{\mathrm{dof}}$ parameter distribution in the CR. Generally, data and MC simulation agree well, showing that the MC simulation models well the data in the CR, which represents only a small corner of the track phase space.

The fraction of fake and real tracks are fitted to the data histograms in the CR through a standard likelihood fit using Poisson statistics, as implemented in the `TFractionFitter` function of `ROOT` [2]. Rates of the fraction of fake tracks are obtained by fitting the distribution of the track

[2] $n_{\mathrm{dof}}$ stands for the number of degrees of freedom.

$\chi^2/n_{\text{dof}}$ variable, or number of SCT hits, or number of SCT holes, or the number of SCT shared hits. The obtained results in the CR are shown in Figure 3(a), and are found to agree within 30% with the fraction of fake tracks in the MC simulation, which is 29%.

These data estimations are subject to potential mis-modelling of the variables used to define the CR, as it selects only a small fraction of all tracks. Therefore, the final result is presented by extrapolating the fraction of fake tracks to the pre-selection sample, as detailed in Ref. [2]. Depending on the $f_{\text{CR}}^{\text{fakeTrk}}$ variable, the fraction of fake tracks in data at pre-selection level is found to vary between 0.21% and 0.24% (Figure 3(b)), compared to 0.77% in Ref. [2]. The decrease can be explained by a tighter pre-selection applied for this study, and certainly because of the improvements in the track reconstruction algorithms mentioned at beginning of this document.

| Selected variable | $f_{\text{CR}}^{\text{fakeTrk}}$ | | |
|---|---|---|---|
| | MLE | toys | histo $\chi^2/n_{\text{dof}}$ |
| track $\chi^2/n_{\text{dof}}$ | 0.25 ± 0.02 | 0.24 ± 0.03 | 0.23 ± 0.03 |
| $n_{\text{SCT}}$ hits | 0.29 ± 0.06 | 0.33 ± 0.06 | 0.33 ± 0.06 |
| $n_{\text{SCT}}$ holes | 0.28 ± 0.04 | 0.34 ± 0.05 | 0.34 ± 0.05 |
| $n_{\text{SCT}}$ shared hits | 0.29 ± 0.04 | 0.23 ± 0.04 | 0.23 ± 0.04 |

| Selected variable | $f_{\text{pre-Sel}}^{\text{fakeTrk}}$ | |
|---|---|---|
| | Data | MC |
| track $\chi^2/n_{\text{dof}}$ | 0.0021 | 0.0024 |
| $n_{\text{SCT}}$ hits | 0.0024 | 0.0024 |
| $n_{\text{SCT}}$ holes | 0.0023 | 0.0024 |
| $n_{\text{SCT}}$ shared hits | 0.0024 | 0.0024 |

**Figure 3:** Left, 3(a): fraction of fake tracks in the dijet CR ($f_{\text{CR}}^{\text{fakeTrk}}$), and the associated statistical uncertainty obtained after fitting the selected variables distributions. The results of the 1000 toy experiments are shown in the middle column. The minima of the $\chi^2$/DOF of the data and MC histograms of these variables are shown in the last column. MLE stands for the maximum likelihood estimation. Right, 3(b): Fraction of fake tracks at pre-selection level, $f_{\text{pre-Sel}}^{\text{fakeTrk}}$ obtained with the method from Ref. [2]. $f_{\text{CR}}^{\text{fakeTrk}}$ is taken from the fit of the selected variables distributions. Tables taken from Ref. [6].

## 4. Conclusions

For LHC Run 3 data taking, the track reconstruction algorithm used for the ATLAS Inner Detector has been optimized to ensure a collection of high quality tracks to downstream reconstruction and physics. This document, dedicated to modelling of track reconstruction in the core of the jets, presented new measurements of the charged particle reconstruction inefficiency and fake rate inside jets. As expected, the rate of fake tracks is seen to be several orders of magnitude when compared to the previous measurements.



## References

[1] ATLAS Collaboration, JINST **3** (2008) S08003

[2] ATLAS Collaboration, ATL-PHYS-PUB-2017-016 (2017)

[3] T Cornelissen et al., Journal of Physics: Conference Series **119** (2008) 032014

[4] ATLAS Collaboration, ATL-PHYS-PUB-2015-018 (2015)

[5] ATLAS Collaboration, ATL-PHYS-PUB-2021-012 (2021)

[6] ATLAS Collaboration, IDTR-2022-05 (2022)

[7] ATLAS Collaboration, ATL-PHYS-PUB-2015-051 (2015)