

Reconstruction software reconstructs basic objects from signals collected from the event: interaction vertices, tracks, topological clusters of energy deposits
 These quantities then used to reconstruct physics objects i.e. particles (electron, muon), jets, MET

0.1 Primary reconstruction

0.1.1 Topological clusters

[**reco:topocluster**][**reco:topocluster*2**]

Topological cluster (topo-cluster): Clusters of topologically connected cell signals in the calorimeter at the EM scale. This scale does not consider loss of signal from hadrons. Singular hits without hits from neighboring cells are considered noise.

Done in an effort to extract signal while minimizing electronic effects and physical fluctuations. Used to reconstruct hadronic objects and particles decaying hadronically i.e. τ leptons
 Signal hits with significance above a cell signal significance level $\zeta_{\text{cell}}^{\text{EM}}$ are seeded in as part of a proto-cluster. Neighboring cells satisfying a cluster growth threshold are collected into the cluster.

Two clusters are merged if a cell is matched to both

If a cluster has two or more local signal maxima satisfying $E_{\text{cell}}^{\text{EM}} > 500$ MeV, the cluster is split accordingly.

The process continues iteratively until all cells with significant signal efficiency have been matched to a cluster.

0.1.2 Tracks

[**reco:track**]

Charged particles deposit energy in different layers of the inner detector and muon spectrometer

ID reco software: inside-out and outside-in algorithms

- Inside-out: [**reco:io**]

Starts with seeded hits in the silicon detector in pixel & SCT

Loosely matched to an EM cluster to form a track candidate

Hits are added to form a track candidate using a pattern recognition algorithm based on a Kalman filter formalism [**reco:kalman**]

Track candidates are then fitted with a χ^2 filter [**reco:track*chi2**] and loosely matched to a fixed-sized EM cluster. Successfully matched track candidates are re-fitted with a Gaussian-sum filter (GSF) [**reco:track*gsf**]

This is followed by a track scoring strategy to resolve fake tracks & hit ambiguity between different tracks [**reco:track*ambiguity**]

Extend to TRT to form final tracks, filtered by threshold $p_{\text{T}} > 400$ MeV.

ATLAS simulation 2010

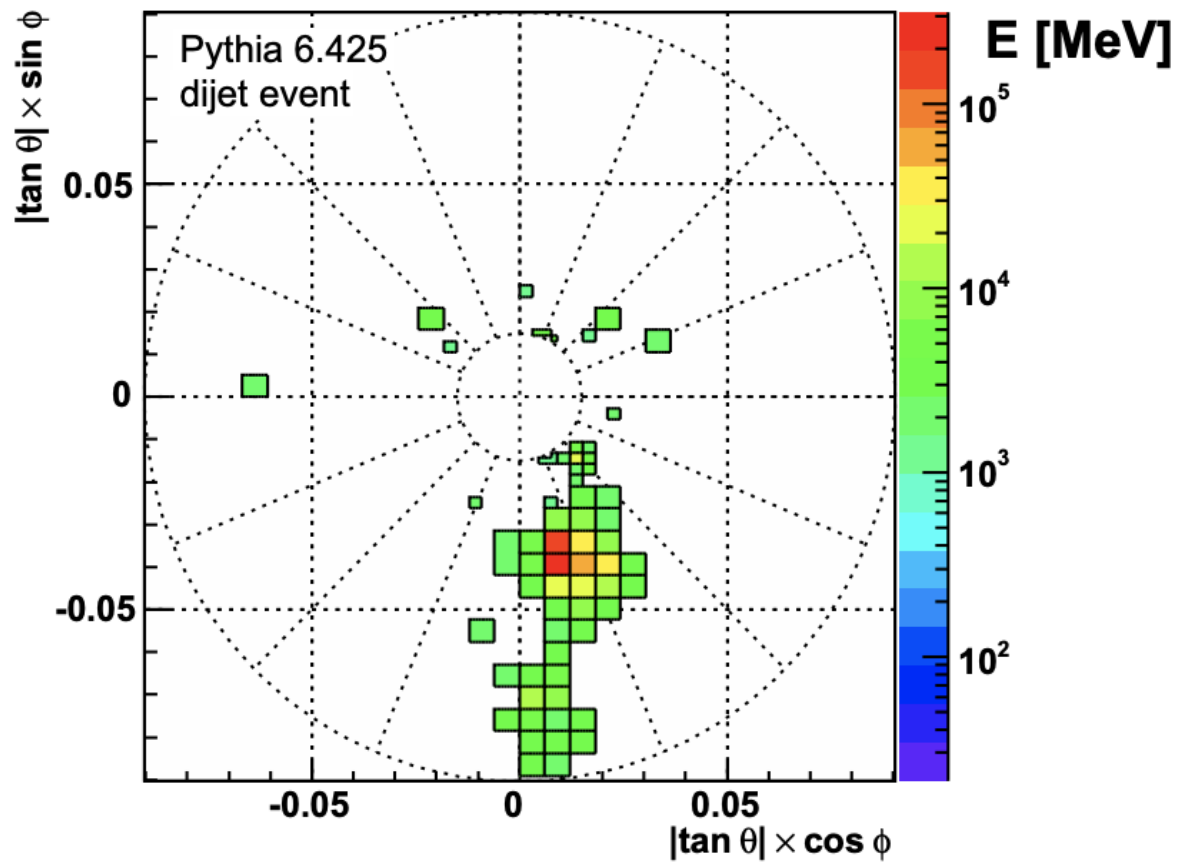


Figure 1

- Outside-in: [**reco:oi**]
Reverse, starts with segments in TRT extending inward to silicon hits in pixel & SCT
Targeting secondary tracks (decays/interactions of primary particles) or long-lived particles

0.1.3 Vertices

Vertices: interaction or decay point

Primary vertex: pp interaction point

Important for reconstruction of the hard scattering pp interaction, resulting trajectories and kinematic information of the event

- Vertex finding:
Uses the z-position of a track as input
Vertices require to have at least 2 tracks
Iterative χ^2 algorithm evaluate track-vertex compatibility, using the track as new seed for another vertex if large discrepancy
- Vertex fitting:
Adaptive multi-vertex fitter (AVF) algorithm assigns weights that depend on the track-vertex compatibility to each track to measure the probability of the track being an outlier vs inlier.
Vertex is then estimated by iteratively minimizing an objective function of these weights

0.2 Jets

- Quarks, gluons & other non-color-neutral hadrons cannot be observed individually due to QCD color confinement
- A non-color-neutral hadron will almost immediately undergo hadronization producing a cone of color-neutral hadrons also known as a jet
- Jet signals can be used to reconstruct and consequently indirectly observe the original quarks/gluons the jets originated from
- Jet reconstruction:
 - PFlow: energy deposited in the calorimeter systems by charged particles is removed and replaced by particle objects created with the remaining energy in the calorimeter and tracks matched to the topo-clusters. (include PFlow graphics)
 - anti- k_t algorithms: sequential recombination jet algorithms
 - pile-up jets: multiple interactions associated with one bunch crossing in addition to the hard scattering of interest and reconstructed as jets in the final states. Reconstructed pile-up jets can result from Pile-up jets are usually from soft interactions and can be distinguished with JVT algorithm using tracking information from the ID.

- JES/JER calibration: Jet reconstruction at EM scale does not accurately account for energy from QCD interactions and needs to be calibrated to jets reconstructed at particle level. This is done via a MC-based JES calibration sequence and additional JER calibration to match jet resolution in simulation to data using dijet events.

Flavor tagging

- Classification of hadronic jets is an important task for many LHC analyses especially ones studying final states (Higgs decay/4top)
- Flavor tagging is namely interested in identifying jets containing b -hadrons, c -hadrons, uds -hadrons (light-jets), and hadronic decays from τ .
- Of these, identifying b -jets is of particular interest due to their characteristically long lifetime (≈ 1.5 ps) from decay suppression by CKM factor, with a displaced secondary decay vertex and usually a tertiary vertex from c -hadron decays.

Efficiency calibration

- [ftag:calib]
- Performance of b -taggers are studied on MC simulated samples. However, the b -tagging efficiency predicted by simulation $\varepsilon_b^{\text{sim}}$ is usually not the same as the efficiency measured in data $\varepsilon_b^{\text{data}}$.
- The correction for the rate of events after applying a b -tagging requirement is calibrated and applied jet by jet in the form of data-to-simulation scale factors $\text{SF} = \varepsilon_b^{\text{data}}/\varepsilon_b^{\text{sim}}$.
- Usage of b -tagger in this analysis is done via five operating points (OPs), corresponding to 60%, 70%, 77%, 85% and 90% b -jet tagging efficiency ε_b in simulated $t\bar{t}$ events in order from loosest to tightest discriminant cut points. - OPs are defined by selection on the tagger output to provide a pre-defined level of ε_b , and act as a variable trade-off between b -tagging efficiency and c -/light-jet rejection i.e. b -jet purity
- A jet is considered b -tagged if it passes the efficiency criteria for a given OP. A pseudo-continuous b -tagging (PCBT) score is defined to summarize the OP criteria a jet passes into a variable. The PCBT score can take integer values between 1 and 6, where a score of 6 means a jet passes all four OP thresholds (passing 65% OP), a score of 2 for jets that pass only the 90% OP, and a score of 1 for jets that don't pass any OP. Additionally, PCBT defines a value of -1 for any jet that does not satisfy b -tagging criteria.
- (insert TPR/FPR discriminant trade-off figure)
- $t\bar{t}$ calibration [ftag:ttbar]
- ptrel and high pT calibration [ftag:multijet][ftag:highpt][ftag:ptrel]
- impact parameter \rightarrow signed transverse impact parameter significance
- calibration results

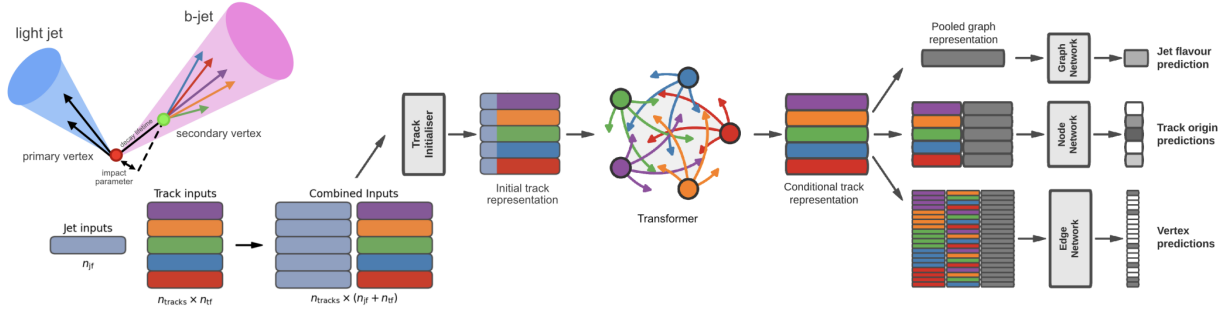


Figure 2: [ftag:gn1][ftag:gn1.1][ftag:gn2]

GN2 b -tagging algorithm

- GN2 transformer-based b -tagging algorithm, utilized for analysis of Run 2 and Run 3 data
- GN2 gives a factor of 1.5-4 improvement in experimental applications compared to the previous convolutional neural network-based standard b -tagging algorithm, DL1d, without dependence on the choice of MC event generator.
- Attention-based architecture, modified to incorporate domain knowledge and additional auxiliary physics objectives: grouping tracks originating from common vertices and prediction of the underlying process for each track
- MC simulated SM $t\bar{t}$ and BSM Z' events from pp collisions were used as training and evaluation samples. In order to minimize bias, both b - and light-jet samples are re-sampled to match c -jet distributions.
- GN2 concatenates 2 jet and 19 track reconstruction variables of up to 40 tracks to form the input feature vector, normalized to zero mean and unit variance.
- The output consists of a jet classification layer of size 4 consisting of p_b , p_c , p_u and p_τ for the probability of each jet being a b -, c -, light- or τ -jet respectively; a track-pairing output layer of size 2, and a track origin classification layer of 7 output categories.

0.3 Leptons

- Lepton reconstruction is concerned mainly with electron and muon construction, since tau decays quickly and can either be reconstructed using jets or light leptons. From here on out lepton will be used mostly to refer to electrons and muons
- Leptons can be classified into two categories: prompt leptons resulting from heavy particle decays, or non-prompt leptons resulting from detector or reconstruction effects, or from b - or c - hadron decays
- Reconstruction of leptons is therefore important to study the underlying physics and suppressing background

0.3.1 Electrons

- [reco:electron'id][reco:electron'meas]
- Electrons lose energy interacting with the detector materials via bremsstrahlung. The bremsstrahlung photon can then produce an electron-positron pair which can itself leave signals in the detector, creating a collimated object that can leave multiple tracks in the ID or EM showers in the calorimeter, all considered part of the same EM topo-cluster.
- Electron signal signature has three characteristic components: localized energy deposits in the calorimeter, multiple tracks in the ID and compatibility between the above tracks and energy clusters in the $\eta \times \phi$ plane. Electron reconstruction in ATLAS follows these steps accordingly - Seed-cluster reconstruction and track reconstruction are performed sequentially

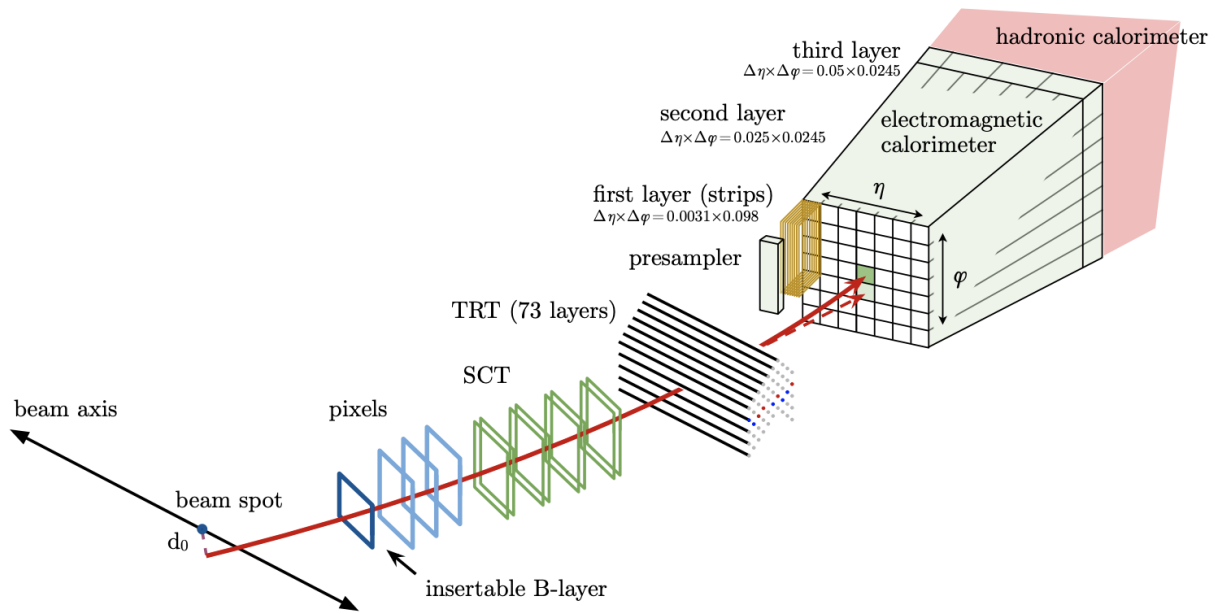


Figure 3: [reco:electron'id]

in accordance with the iterative clustering algorithm and track reconstruction method respectively, described in ??

- The seed-cluster and track candidate associated with a conversion vertex are then matched to form an electron candidate.
- A reconstructed cluster is expanded from the seed-cluster in either ϕ or η in the barrel or endcap region respectively
- The cluster energy is then calibrated to compute the original electron energy.

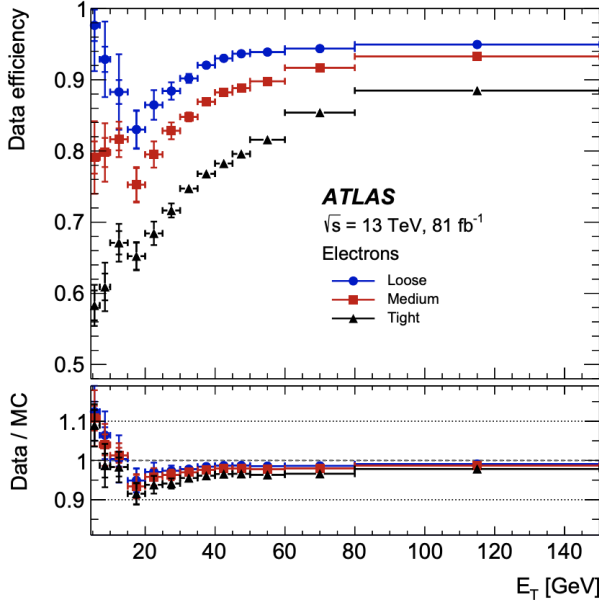
0.3.1.1 Electron identification

- Additional likelihood-based identification selections using ID and EM calorimeter information are implemented to further improve the purity of the reconstructed electrons and photons. These selections also help suppress background from hadronic jet deposits, photon

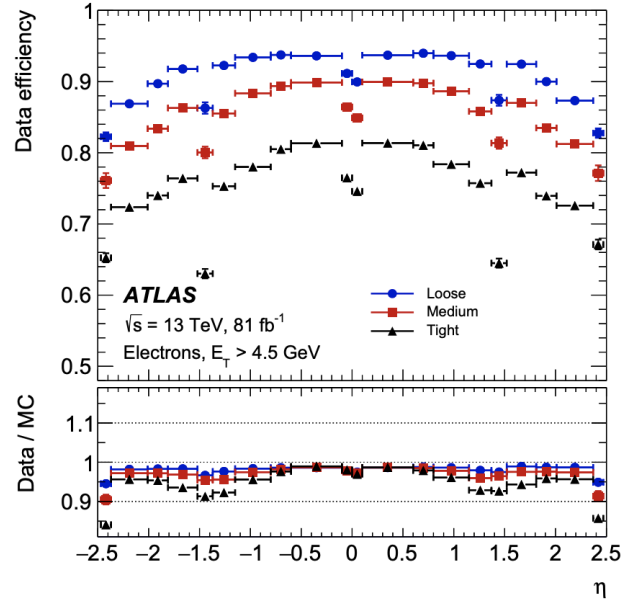
conversions or electrons from heavy-flavor decays.

- Three operating points are defined for physics analyses: Loose, Medium and Tight, optimized for 9 bins in $|\eta|$ and 12 bins in E_T with each corresponding to a fixed efficiency requirement for each bin. The target efficiencies for Loose, Medium and Tight start at 93%, 88% and 80% respectively for typical EW processes and increases with E_T

Similar to b -tagging OPs, the electron OPs represent a trade-off in signal efficiency and background rejection. The electron efficiency are estimated using tag-and-probe method [reco:electron'id] on samples of $J/\Psi \rightarrow ee$ and $Z \rightarrow ee$.



(a) Figure A



(b) Figure B

Figure 4: [reco:electron'meas]

0.3.1.2 Electron isolation

- A characteristic distinction between prompt electrons and electrons from background processes is the relative lack of activity in both the ID and calorimeter within an area of $\Delta\eta \times \Delta\phi$ surrounding the reconstruction candidate

- Electron isolation variables are needed to quantify the amount of activity around the electron candidate.

- Calorimeter-based isolation variables $E_T^{\text{cone}XX}$ is calculated by first summing the energy of topological clusters with barycenters falling within a cone of radius $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = XX/100$ around the direction of the electron candidate.

- The final isolation variable is then obtained by subtracting energy at the core of the cone belonging to the candidate electron from the sum, then applying corrections for energy leakage outside of the core and pile-up effects.

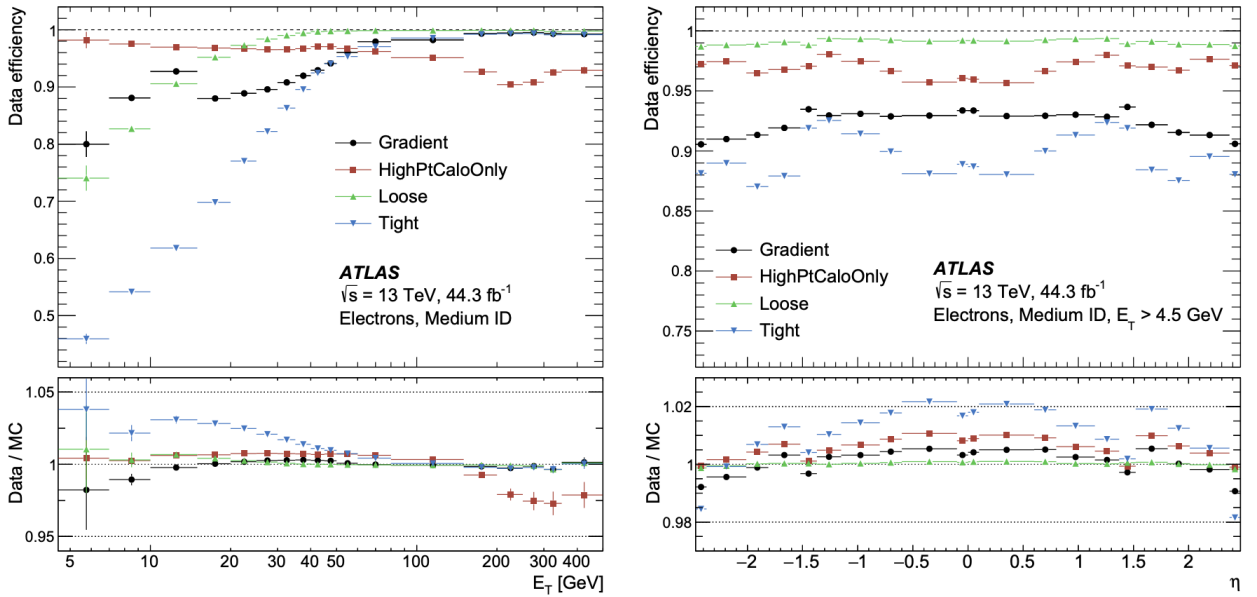
- Similar to calorimeter-based variables, track-based isolation variables $p_T^{\text{varcone}XX}$ are calculated by summing all track p_T within a cone of variable radius ΔR around the electron

candidate, minus the candidate's contribution. The cone radius is variable as a function of p_T

$$\Delta R = \min \left(\frac{10}{p_T [\text{GeV}]}, \Delta R_{\text{max}} \right)$$

with ΔR_{max} being the maximum cone size, to account for the closer proximity of decay products to the electron in high-momentum heavy particle decays.

- Four isolation operating points are implemented to satisfy specific needs by physics analyses: Loose, Tight, HighPtCaloOnly and Gradient. The first three OPs are fixed in isolation variables, while the Gradient OP fixes the isolation efficiency to a p_T dependent function defined as $\varepsilon = 0.1143 \times p_T + 92.14\%$ with p_T in GeV, using $\Delta R = 0.2$ for calorimeter isolation and $\Delta R_{\text{max}} = 0.2$ for track isolation.



(a) Figure A

(b) Figure B

Figure 5: [reco:electron'meas]

0.3.1.3 Non-prompt electrons

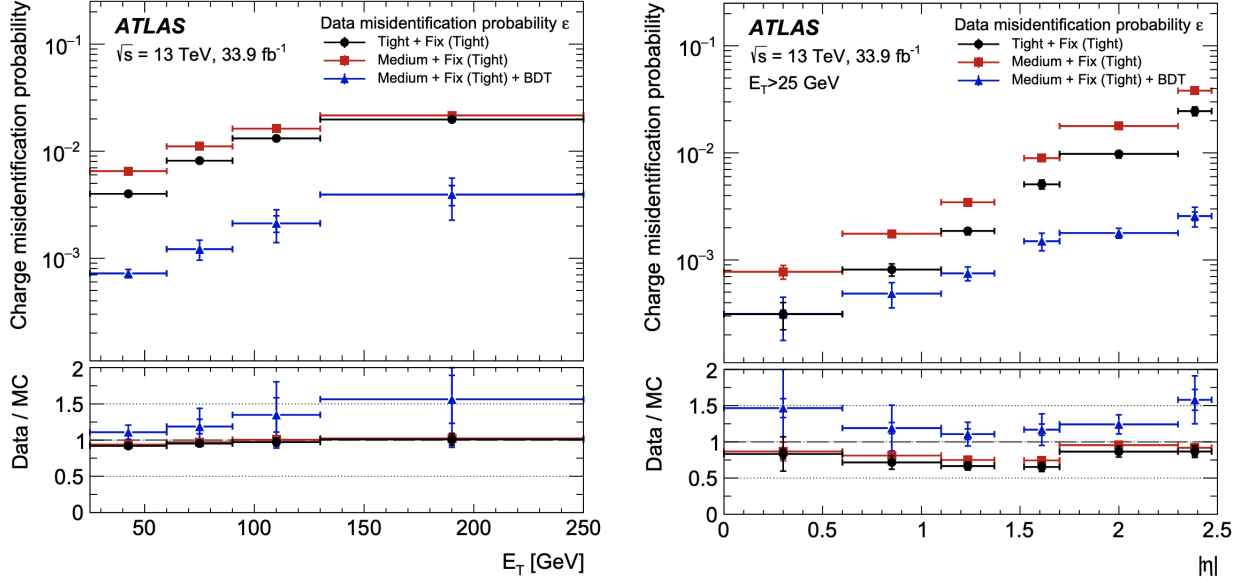
Electron charge misidentification [reco:electron'id][reco:qmisid'cnn]

Electron charge is determined by the curvature of the associated track. Misidentification of charge can then occur via either an incorrect curvature measurement, or an incorrectly matched track.

The former is more likely for electrons with high p_T due to the small curvature in track trajectories at such scale, while the latter usually results from bremsstrahlung pair-production, creating additional secondary tracks in the vicinity.

Charge misidentification is a crucial irreducible background for analyses with charge selection criteria, and suppression of this background is done via a boosted decision tree discriminant

known as the Electron Charge ID Selector (ECIDS) [`reco:electron`id`]. The addition of ECIDS removed 90% of electrons with incorrect charge while selecting 98% of electrons with correct charge from electrons in $Z \rightarrow ee$ events satisfying Medium/Tight identification and Tight isolation criteria.



(a) Figure A

(b) Figure B

Figure 6: [`reco:electron`meas`]

Photon conversion electrons

0.3.2 Muons

Muon identification

Muon isolation

0.4 Missing transverse momentum

0.5 Pile-up & overlap removal