

Measurement of Electron Reconstruction Efficiencies in Run 2 and Run 3 at the ATLAS Experiment

Master-Arbeit
zur Erlangung des Hochschulgrades
Master of Science
im Master-Studiengang Physik

vorgelegt von

MARIUS MELCHER
geboren am 16.05.2000 in SPREMBERG

Institut für Kern- und Teilchenphysik
Fakultät Physik
Bereich Mathematik und Naturwissenschaften
Technische Universität Dresden
2025

Eingereicht am 29. Oktober 2025

1. Gutachter: Prof. Dr. Arno Straessner
2. Gutachter: Prof. Dr. Michael Kobel

Abstract

To uncover evidence for new laws of nature, the Standard Model of particle physics is probed at high-energy colliders with increasing precision. The wide range of surveyed processes has to be reconstructed from only a few directly observable final-state objects. A precise understanding of the behavior of these objects in the detector, as well as the performance of the employed algorithms, is therefore crucial.

This thesis presents the measurement of the electron reconstruction efficiency in LHC Run 2 and Run 3 at the ATLAS experiment, corresponding to a total integrated luminosity of 307.1 fb⁻¹. The efficiency is determined using the tag-and-probe method applied to $Z \rightarrow ee$ decays together with data-driven background estimation techniques. The comparison of efficiencies between data and simulation enables the derivation of scale factors, which are used by physics analyses involving electron final-states.

When integrated over all included kinematic regions, an overall reconstruction efficiency of 98% is measured with permille-level uncertainties. While in Run 2 the scale factors are close to unity, several deviations between efficiencies in data and simulation are observed in Run 3, which is subsequently investigated. To further improve the stability of the results and reduce their uncertainties, major changes in background estimation fitting procedure are evaluated.

Kurzfassung

Bei der Suche nach Hinweisen auf neue Naturgesetze wird das Standardmodell der Teilchenphysik an Hochenergiebeschleunigern mit immer höherer Präzision auf die Probe gestellt. Obwohl dabei eine Vielzahl verschiedener Prozesse betrachtet wird, können nur wenige Endzustände direkt im Detektor gemessen werden. Das Verhalten dieser Objekte im Detektor und die Performance der verwendeten Algorithmen muss daher präzise bestimmt werden.

In dieser Arbeit wird die Messung der Elektronen Rekonstruktionseffizienz in LHC Run 2 und Run 3 des ATLAS Experiments vorgestellt. Dabei wird eine integrierte Luminosität von 307.1 fb^{-1} ausgewertet. Die Tag-and-Probe-Methode wird zusammen mit Methoden datenbasierter Untergrundabschätzung verwendet, um aus $Z \rightarrow ee$ Zerfällen eine Effizienz zu ermitteln. Aus dem Verhältnis aus Effizienzen in Daten und in Simulation ergeben sich Skalenfaktoren, die in Analysen mit Elektronen im Endzustand verwendet werden, um die Simulation zu korrigieren.

Integriert über die gesamte betrachtete kinematische Region ergibt sich eine Rekonstruktionseffizienz von knapp 98 % mit Unsicherheiten im Promillebereich. Während in Run 2 die meisten Skalenfaktoren nahe Eins sind, zeigen sich in Run 3 mehrere Abweichungen, welche daraufhin näher untersucht werden. Um die Stabilität der Ergebnisse weiter zu erhöhen und die Unsicherheiten zu reduzieren, werden außerdem Änderungen der Untergrundabschätzung untersucht.

Contents

1	Introduction	1
2	The ATLAS Experiment at the LHC	3
2.1	Large Hadron Collider	3
2.2	ATLAS Detector	4
2.2.1	ATLAS Coordinate System	5
2.2.2	Inner Detector	6
2.2.3	Calorimeter	7
2.2.4	Muon Spectrometer	9
2.3	Data and Luminosity	10
2.4	Event Simulation	12
3	Electrons at ATLAS	15
3.1	Trigger	15
3.2	Reconstruction	15
3.3	Energy Calibration	18
3.4	Identification	19
3.5	Isolation	19
3.6	Relation to other High-Level Physics Objects	20
3.6.1	Photons	20
3.6.2	Taus	21
3.7	Electron Efficiencies	21
4	Measurement of Electron Reconstruction Efficiency	25
4.1	Methodology	25
4.1.1	Tag-and-Probe Method	25
4.1.2	Efficiency Definition	26
4.1.3	Event Selection and Binning	27
4.1.4	Background Estimation	28
4.2	Uncertainties	30
4.2.1	Statistical Uncertainties	30

4.2.2	Systematical Uncertainties	32
4.3	Efficiency and Scale Factor Results	37
4.3.1	Electron Reconstruction Efficiency as a Function of p_T and η	37
4.3.2	Run 2 Release 22 Results	40
4.3.3	Run 3 Results	42
4.3.4	Fast Simulation Efficiencies and Scale Factors	48
4.3.5	Improvement of photonic Background Estimation	51
4.4	Open Questions and Outlook	58
5	Conclusion	61
	Bibliography	63
	Appendix	71
A.1	List of Samples	71
A.2	List of Triggers	73
A.3	Reconstruction Efficiencies	74
A.3.1	Comparison between Run 2 and Run 3	74
A.3.2	Impact of Cluster Quality on Efficiencies	75
A.3.3	Comparison between Full and Fast Simulation	76
A.3.4	Run 3 Results using the new Fitting Setup	78

1 Introduction

Our current physical understanding of the universe at its most fundamental level is based on two theoretical frameworks: the Standard Model of particle physics and general relativity. To describe the interactions of elementary particles, the Standard Model includes three fundamental forces: electromagnetism [1, 2], the weak interaction [3, 4] and the strong interaction [5, 6]. Gravity, however, could not be integrated consistently, here general relativity [7] is currently the best available description. In addition to the absence of a quantum theory of gravity, several other open questions remain in the Standard Model, such as the observed matter-antimatter asymmetry [8] and the unknown nature of dark matter [9], which could explain discrepancies in astronomical observations.

Despite its apparent incompleteness, the predictions of the Standard Model have so far been confirmed experimentally without any statistically significant deviations, ruling out many possible theoretical extensions. To find leads for physics beyond the Standard Model, high energy particle colliders provide the opportunity to probe the Standard Model with high precision and at previously uncharted energy scales. Large machines are needed to accelerate particles close to the speed of light and to detect the particles created in their collision.

At the ATLAS experiment, one of two general-purpose detectors at the Large Hadron Collider, a wide range of physics processes is probed. Precision measurements of processes described by the Standard Model are complemented with direct searches for new particles like additional Higgs bosons, supersymmetric particles and other possible dark matter candidates. While these interactions of main physical interest include a variety of particles, there are only a few possible final-state objects that can be observed directly in experiments. Besides photons, muons and jets, electrons are among the key objects the ATLAS detector is designed to measure directly.

Electrons and positrons are either produced in pairs e^+e^- (e.g. produced through Z or γ) or in combination with a neutrino $e\nu_e$ (via weak decays, e.g. from τ), which can be inferred indirectly through missing momentum. Electrons and positrons are jointly denoted as e . This provides a clear experimental signature, which was, for example, instrumental in the original discovery of the Higgs boson at ATLAS [10], where $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$ and $H \rightarrow WW^{(*)} \rightarrow e\nu\mu\nu$ were two of

the first channels seeing evidence for the Higgs boson. From their production after the collision, several reconstruction steps are needed to arrive at electrons usable by physics analyses. It is crucial that these steps and their performance are well understood, since electrons appear in many precision measurements and searches, from electroweak processes to Higgs and new physics channels.

This thesis reports on the measurement of the electron reconstruction efficiency in Run 2 and Run 3 at the ATLAS experiment. The reconstruction efficiency describes the probability that an electron traversing the detector is successfully reconstructed as such by the corresponding algorithms. It is not only used to evaluate the performance of the reconstruction algorithms, but also to derive scale factors, which are applied to simulated results to compensate for any differences in electron reconstruction between data and simulation.

In chapter 2, the experimental setup at the ATLAS experiment is introduced. The algorithms used to reconstruct electrons and discriminate them from other particles are outlined in chapter 3, followed by the results for the reconstruction efficiency in chapter 4. Here, first the measurement methodology and the uncertainty estimation are explained. The results in Run 2 and Run 3 are discussed in detail, including several investigations into unexpected changes observed in Run 3. Finally, improvements to the background estimation are developed and evaluated. Throughout this thesis, natural units ($\hbar = c = 1$) are used.

2 The ATLAS Experiment at the LHC

To probe the Standard Model with high precision and search for signatures of physics beyond the Standard Model, a powerful experimental environment is required. Two parameters are of primary importance: energy and luminosity. Higher collision energies allow the exploration of new mass regimes and the investigation of properties of heavy particles such as the top quark. On the other hand, a high luminosity provides large statistical samples, which is necessary to study rare processes. The Large Hadron Collider (LHC) is the world's most powerful particle accelerator in terms of center-of-mass energy, while also achieving record luminosities.

2.1 Large Hadron Collider

Located at the border between France and Switzerland near Geneva, the Large Hadron Collider is a two-ring particle accelerator with a 26.7 km circumference. It is part of the European Organization for Nuclear Research, CERN, and its extensive accelerator complex, shown in Fig. 2.1. At the LHC and its experiments, protons and heavy ions are collided. This work focuses on proton-proton collisions recorded during Run 2 (2015–2018) and Run 3 (2022–2026).

Before being injected into the LHC, protons and heavy ions are accelerated through a series of smaller accelerators, known as the injector chain. For proton collisions, the beam first passes through a linear accelerator. Linac2 was in use for the first two data-taking Runs and its successor Linac4 is used since 2020. Unlike Linac2, which accelerated protons directly, Linac4 accelerates negatively charged hydrogen ions. Before injection into the next accelerator, the electrons are stripped from these ions to produce a pure proton beam. The protons are then subsequently accelerated in a sequence of synchrotrons: the Proton Synchrotron Booster, the Proton Synchrotron and Super Proton Synchrotron. In this last injector, protons reach a kinetic energy of 450 GeV and are then injected into the two counter-rotating rings of the LHC [12].

It is important to note that, in addition to serving the experiments at the LHC, the CERN accelerator complex supplies beams with various energies to numerous other experiments,

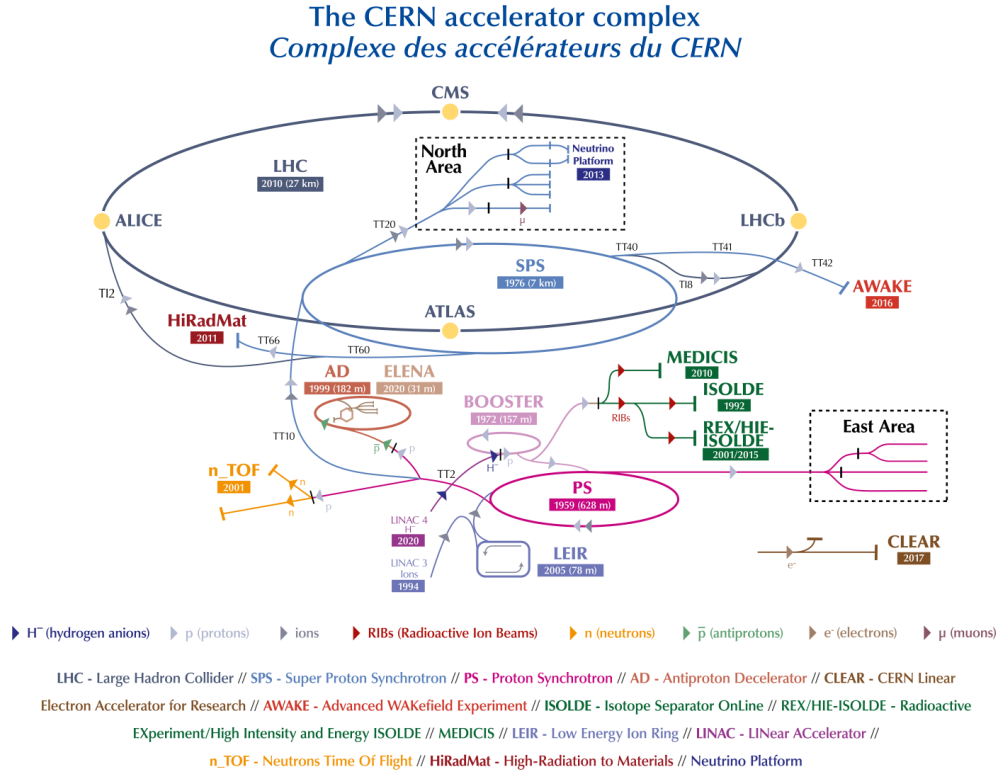


Figure 2.1: Schematic of CERN accelerator complex in 2022 [11].

enabling a diverse research program.

The beams in the LHC consist of bunches of approximately 10^{11} protons each, with a spacing of 25 ns between the bunches. The protons are further accelerated using radio-frequency cavities, reaching their final energy of 6.5 TeV in Run 2 and 6.8 TeV in Run 3, respectively. The two beams are then collided at four different interaction points, resulting in a center-of-mass energy $\sqrt{s}$ of 13 TeV in Run 2 and 13.6 TeV in Run 3 for proton-proton collisions [13].

At the interaction points, the four main experiments and their detectors are located: the general-purpose ATLAS and CMS experiments, ALICE, which is focused primarily on the study of quark-gluon plasma produced by heavy ion collisions, and LHCb, specialized on measuring the decays of hadrons involving b quarks.

2.2 ATLAS Detector

The ATLAS detector, short for A Large LHC ApparatuS, is one of two general-purpose detectors at the LHC. To study a wide range of physics processes, the ATLAS detector's objective

is not only to measure properties of produced particles such as energy, momentum and charge reliably with the best possible precision. In addition to that, it is crucial to provide powerful tools to discriminate between various different particle types. To achieve this, the detector consists of several layers surrounding the beam pipe with different specializations and geometries, enabling optimized spatial coverage around the interaction point.

2.2.1 ATLAS Coordinate System

At the ATLAS experiment, several coordinate systems are in use. With the interaction point as its origin, a Cartesian coordinate system is defined with the z -axis in beam direction, the x -axis pointing towards the center of the LHC and the y -axis pointing upwards. Since ATLAS is built symmetrically around the interaction point, it is in many cases convenient to use polar coordinates with radius r , the azimuthal angle ϕ around the beam axis and polar angle θ from the beam axis. Instead of using θ , the (pseudo-)rapidity is often a useful choice for an angular coordinate. The rapidity is the relativistic velocity in beam direction and can be written as:

$$y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z}. \quad (2.1)$$

For particles with negligible mass compared to the kinetic energy ($E \approx p$), the rapidity y approximates to the pseudorapidity η :

$$y \approx \eta = -\ln \tan \left(\frac{\theta}{2} \right). \quad (2.2)$$

Differences in rapidity Δy are invariant under Lorentz transformations along the beam axis, consequently differential cross sections $d\sigma/dy$ are invariant as well. For hadron colliders like the LHC, this is useful since, when colliding, the constituent partons, rather than the full protons interact with each other, each contributing only an a priori unknown momentum fraction x_i . Therefore, the system of interaction products observed is boosted along the beam axis, depending on x_1 and x_2 , leaving rapidity differences and derived quantities invariant. The angular distance is then measured in units of $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ [14].

Another important quantity for particle collisions is the momentum perpendicular to the beam axis, the transverse momentum $\vec{p}_T$ and its absolute value, which is denoted as p_T . The total vector sum of three momenta of the interaction products depends on the parton momentum fractions, but the total $\vec{p}_T$ of the system after interaction has to be close to zero, since the partons had negligible transverse momenta before the interaction. The missing transverse

momentum p_T^{miss} or, more commonly, the missing transverse energy E_T^{miss} gives some information about particles in the final-state that leave the detector without being measured, like neutrinos:

$$E_T^{\text{miss}} = \left| - \sum \vec{p}_T \right|. \quad (2.3)$$

2.2.2 Inner Detector

After the collision, particles first enter the inner detector, designed to measure the tracks of charged particles with high granularity in the range of $|\eta| < 2.5$. It consists of three different subsystems relying on three different technologies. In Fig. 2.2, the profile of these subdetectors can be seen, as well as their distance from the beamline. The inner detector is embedded in a 2T magnetic field produced by a superconducting solenoid. The resulting curvature of tracks due to the Lorentz force allows reconstruction of the particle momentum from the degree of curvature and the sign of the electrical charge from the direction of curvature.

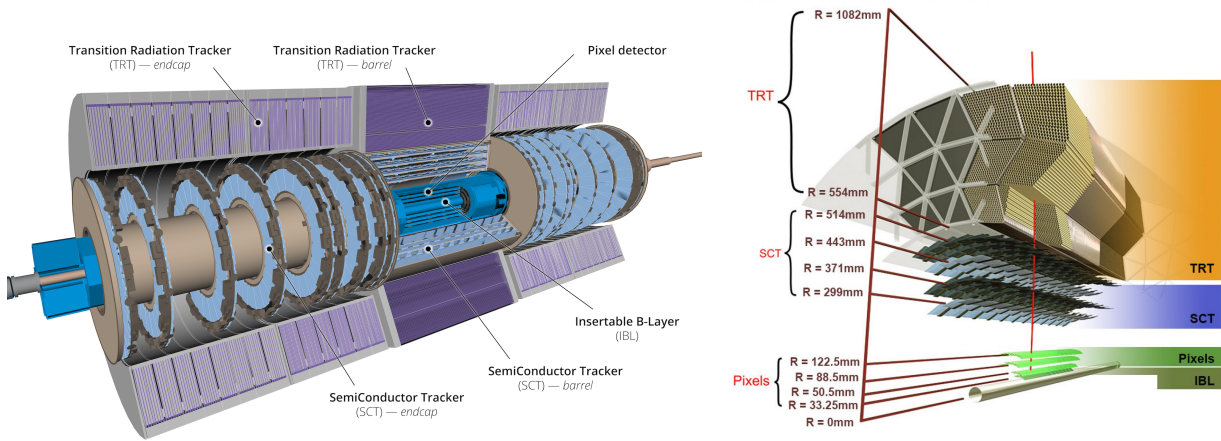


Figure 2.2: 3D schematic of the ATLAS inner detector and its subsections (left) as well as transverse view with distances from the beam line [15].

Pixel Detector: The part closest to the beam pipe is the pixel detector, initially composed of three layers in the barrel region and two endcaps containing three discs each. Additionally, after Run 1, another layer, called the insertable B-layer, of silicon pixel detectors was added between first layer and the beam pipe at 33.25 mm from the beam line, to improve robustness of track reconstruction. This is especially important close to the beam pipe to allow precise

reconstruction of primary and potentially secondary interaction vertices. The pixel detector consists of 2024 silicon pixel modules, each including front-end chips and a semiconductor sensor substrate, in which passing charged particles produce electron-hole pairs, inducing a signal. The pixel detector has an intrinsic resolution of $10\ \mu\text{m}$ in radius r and azimuthal angle ϕ [16,17].

Semiconductor Tracker (SCT): Like the pixel detector, the semiconductor tracker relies on silicon sensors, but a larger area is covered with a reduced intrinsic resolution of $17\ \mu\text{m}$ in the r - ϕ plane. It consists of 4088 silicon strip modules spread over four barrel layers and two endcaps with 9 disks each. The term silicon hits used in the following chapters refers to hits in both the SCT and the pixel detector [16].

Transition Radiation Tracker (TRT): Being the outermost subdetector, the TRT covers an even larger area with a resolution of $130\ \mu\text{m}$. It extends the reconstructed track outwards to a radius of 1082 mm from the beamline, which, among other things, is beneficial for reliable matching of clusters in the calorimeters to tracks in the inner detector, even though it comes with a reduced intrinsic resolution. Unlike the first two subdetectors, the TRT is based on gas filled drift tubes. Here, the signal is created by particles ionizing the gas. Transition radiation can further be used to discriminate between different particle types [14]. After Run 3, the current inner detector will be replaced by an all-silicon inner tracker, the ITk. The objective of this upgrade is to maintain or to improve tracking performance in the challenging high-pileup environment at the High Luminosity LHC [18].

2.2.3 Calorimeter

Beyond the inner detector, the calorimeter system is installed. Within the two calorimeters, most particles are stopped after depositing all their energy. While the relative momentum resolution from tracking detectors decreases with increasing momentum, the relative energy resolution increases with energy, enabling a complementary measurement. Additionally, electrically neutral particles can be measured in calorimeters, unlike in the inner detector.

For accurate energy measurement, full azimuthal angle coverage and good pseudorapidity coverage ($|\eta| < 4.9$) is achieved at ATLAS. All calorimeters at ATLAS are sampling calorimeters composed of alternating layers of active and passive detector material. Particles are interacting with passive absorbers, producing showers of secondary particles, which are then measured in the active detector material.

There are two different calorimeters specialized on different objects. Most electrons and photons are fully absorbed by the electromagnetic calorimeter (ECAL), while hadrons enter the

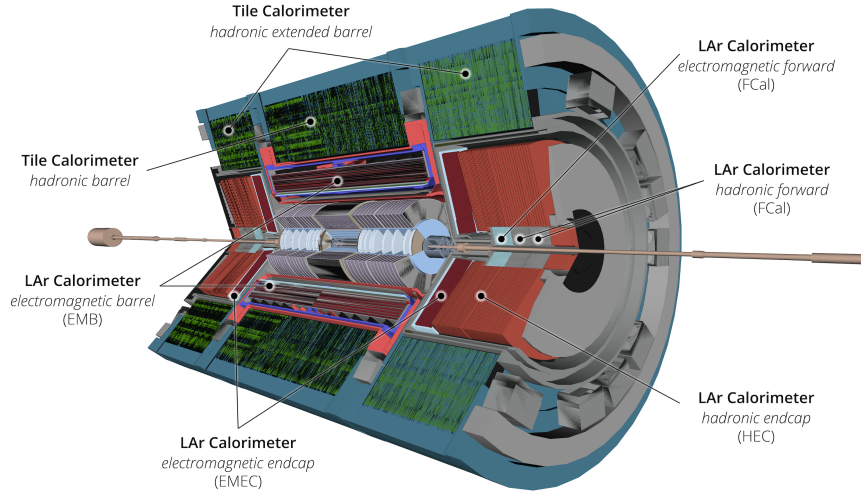


Figure 2.3: 3D schematic of the ATLAS calorimeter system [15].

hadronic calorimeter (HCAL) before being stopped. The components of the calorimeter system are shown in Fig. 2.3.

Electromagnetic Calorimeter: The electromagnetic calorimeter consists of a barrel ($|\eta| < 1.475$) and two endcap components ($1.375 < |\eta| < 3.2$). The region between barrel and endcap is also called the transition region. Lead is used as passive absorber, here, electrons and photons produce a shower of secondary particles, mainly through e^+e^- pair production and bremsstrahlung. Charged particles then induce a signal through ionization of liquid argon (LAr), the active detector material. To ensure full coverage without dead areas, an accordion shape is used [14].

In the barrel region the EM calorimeter is composed of three layers with differing granularity, which is depicted in Fig.2.4. The first layer provides a fine granularity in η , which varies depending on the η region. In the second layer, with coarser η but finer ϕ granularity, electrons and photons deposit most of their energy, since it is thicker than the other layers. The upper part of the EM showers reaches the third layer, which has a coarser granularity. Additionally, a separate thin liquid argon layer, called presampler, is installed in front of the three EM calorimeter layers. Covering the region of $|\eta| < 1.8$, it allows to correct for energy losses in material upstream of the calorimeter.

Hadronic Calorimeter: Jets and some electrons and photons, which did not deposit all their energy, reach the hadronic calorimeter. While the fine granularity in the EM calorimeter enables the measurement of electrons and photons with high precision, the coarser granularity in the hadronic calorimeter still satisfies the physics requirements for jet reconstruction and missing transverse energy measurements [15].

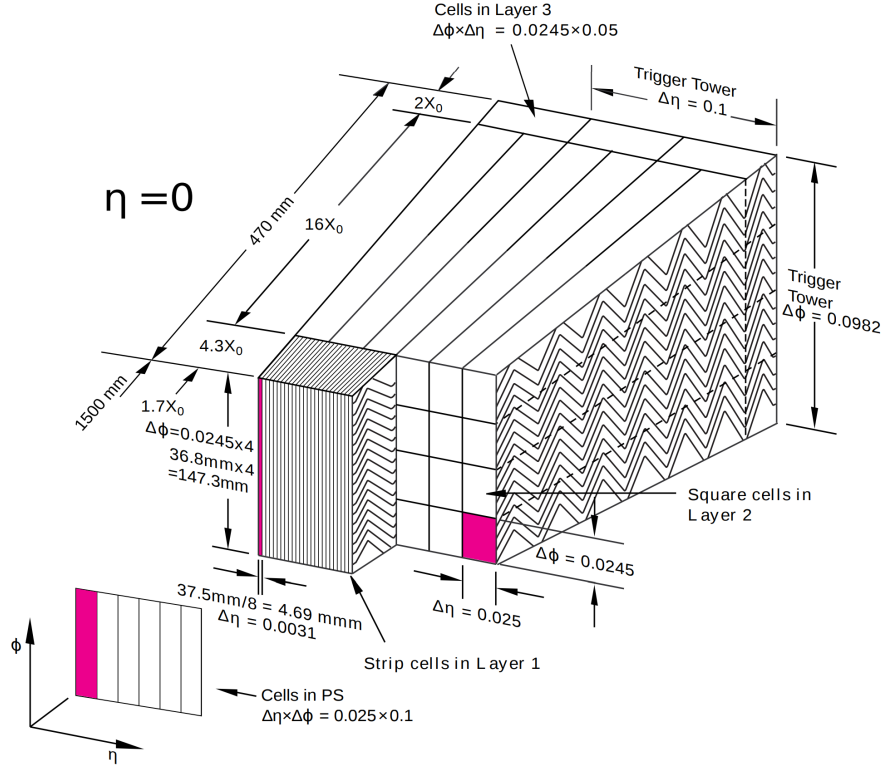


Figure 2.4: Sketch of a ($\eta = 0$) barrel module of the EM calorimeter including the different layers and their granularity in η and ϕ [14].

The barrel and extended barrel sections (see Fig. 2.3) cover $|\eta| < 1.7$. Here, steel is used as absorber and scintillating tiles are used as active material. In the endcaps, liquid argon is used as the active material and copper plates as the absorber. The LAr hadronic endcap calorimeter is located behind its EM counterpart, covering $1.5 < |\eta| < 3.2$ with two independent wheels per endcap. Additionally, the forward calorimeter covers the region $3.1 < |\eta| < 4.9$, completing the nearly full spatial coverage of the ATLAS detector. In this region, the coverage is only calorimetric. It consists of three components, the first uses copper absorbers and is optimised for EM measurements, while the other two parts use tungsten absorbers and measure mostly the energy of hadronic interactions. All three components use liquid argon as active detector material [14, 19].

2.2.4 Muon Spectrometer

Muons generally pass the previous detector layers without losing much of their energy. In the outermost section, the muon spectrometer, different technologies are used to measure the tracks of the muons precisely: Thin Gap Chambers, Resistive Plate Chambers, Monitored Drift Tubes, Small-Strip Thin Gap Chambers and Micromegas (since Run 3) [14, 20]. The setup in Run 3 is shown in Fig. 2.5

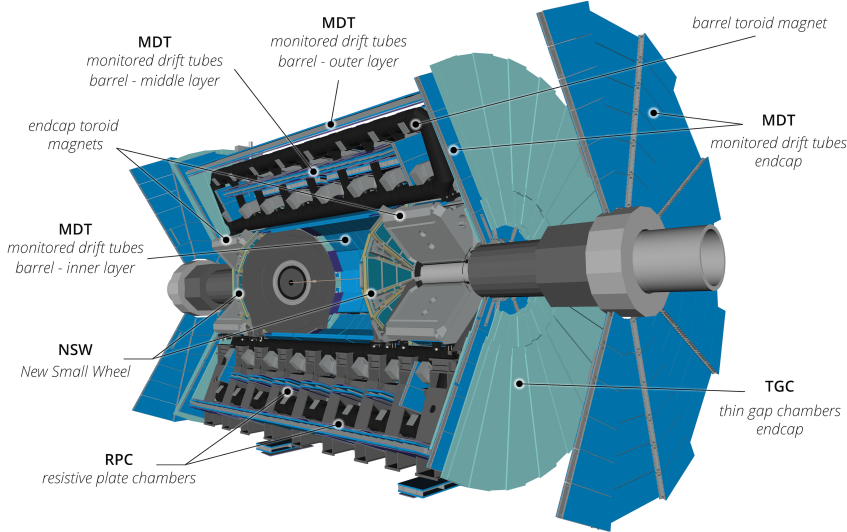


Figure 2.5: 3D schematic of the ATLAS muon spectrometer in Run 3 [15].

Using a toroidal magnet, a magnetic field is applied, allowing again the reconstruction of muon momentum and charge from the curvature. Since the muons are not stopped, a complementary calorimetric measurement is not possible.

The only (stable) Standard Model particles not covered by any of the detector layers are the neutrinos. Their interaction with matter has an extremely small cross section, causing them to pass the detector without interacting with it in most cases. Some information about them can be obtained indirectly by calculating the missing transverse energy E_T^{miss} of an event.

2.3 Data and Luminosity

In order to measure rare processes or achieve high precision, a large number of pp collisions is studied. Instantaneous luminosity $\mathcal{L}$ describes the number of collisions and is historically measured in cm^2s^{-1} and depends on collider properties like the number of protons in a bunch, the number of proton bunches and how focused the particles are within a bunch. The LHC has a design luminosity of $10^{34}\text{cm}^2\text{s}^{-1}$. By integrating over data-taking time, one obtains the integrated luminosity $L = \int \mathcal{L} dt$. This quantity is useful to calculate the expected number of events for a process with cross section σ :

$$N_{\text{events}} = L\sigma. \quad (2.4)$$

For this work, data from the full Run 2 (2015–2018) and the first three years of Run 3 (2022–

2024) is used. This results in a total integrated luminosity of 307.1 fb^{-1} analyzed in these studies. The total integrated luminosity recorded can be seen in Fig. 2.6. Not all the data delivered by the LHC or recorded at ATLAS can be used for physics analyses, e.g. due to some temporary defects. Datasets suitable for physics analysis are collected in Good Run Lists (GRLs). While most measurements are done separately for each data-taking year, the years 2015 and 2016 will be combined to ensure sufficient statistics.

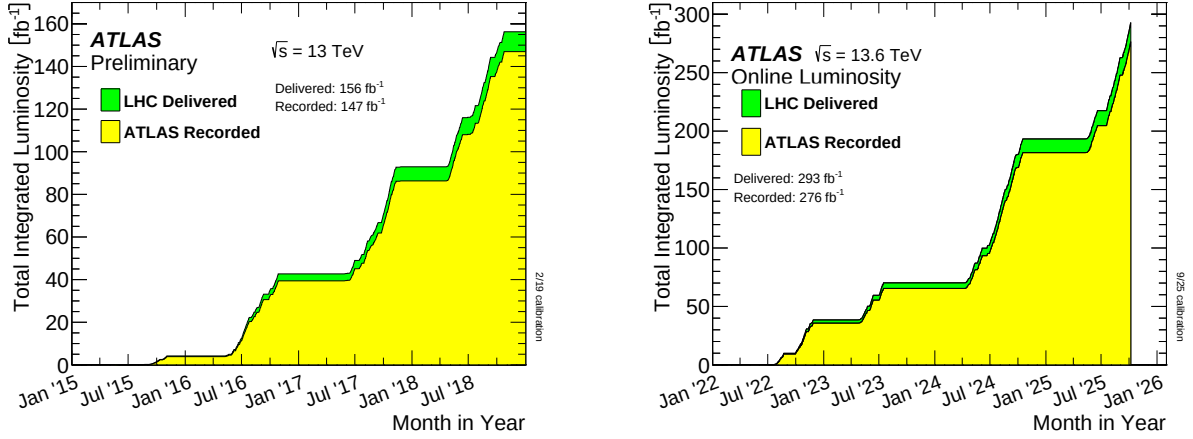


Figure 2.6: Integrated luminosity delivered by the LHC and recorded by ATLAS in Run 2 (left) and Run 3 (right) [21].

Compared to Run 2, for Run 3 the center-of-mass energy has been increased from 13 TeV to 13.6 TeV and several upgrades were performed on the detector. Especially the trigger system underwent major changes to keep up with rising luminosities. Of special interest for object reconstruction is also the number of interactions per bunch crossing, which increased significantly since the start of LHC operations. A higher number of interactions causes a busy experimental environment, with decay products of many near-simultaneous interactions close to each other. This enhances the problem of pileup, where products of different pp interactions are reconstructed in the same event. To suppress this, the primary interaction vertex has to be determined precisely, requiring good track reconstruction. In Fig. 2.7, the mean number of interactions per bunch crossing is shown per year. The impact this has on electron reconstruction is investigated in section 4.3.3.

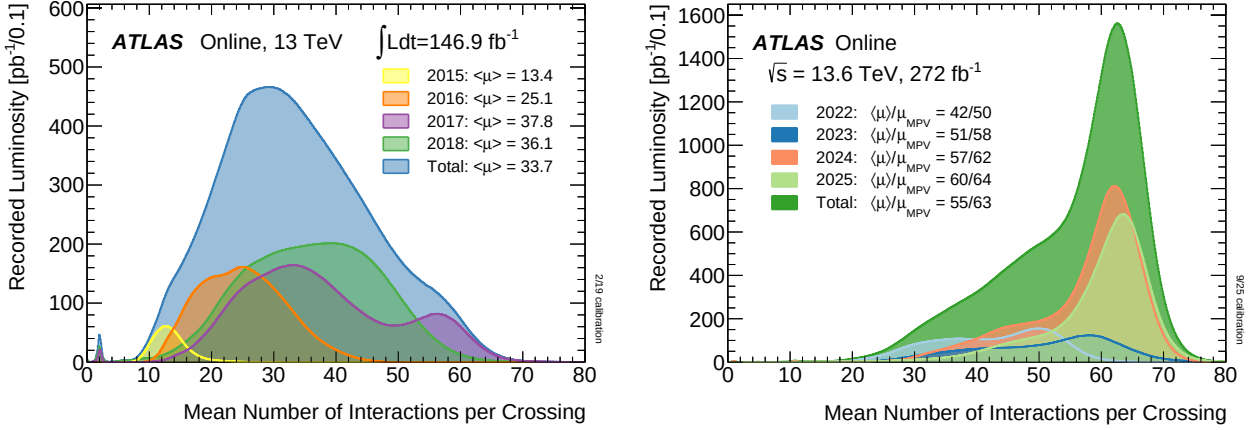


Figure 2.7: Mean number of interactions per bunch crossing per year for Run 2 (left) and Run 3 (right) [21].

2.4 Event Simulation

The interplay between collected data and simulated samples is crucial across the field of particle physics and beyond. At ATLAS, simulated events are, for instance, used to calibrate the measured data and to obtain the number of estimated events. The latter is the basis for defining hypothesis tests, evaluating if data is compatible with the background-only hypothesis or if there is enough evidence to prove the existence of a signal process. The simulation is based on the Monte Carlo (MC) method, where a large number of random samples is used to describe the behavior of the system.

The generation of simulated data at ATLAS is a multi-staged process. First, Monte Carlo event generators describe the underlying particle interactions. The initial state consists of two protons, which are made up of partons: valence quarks, virtual sea quarks and gluons. These partons, described by parton density functions (PDFs), then interact with each other, producing numerous possible final-states. Since many of the particles created after the initial interaction are not stable, event generators also need to produce showers of secondary particles, e.g. from hadronization. An example event produced by event generators is shown in Fig. 2.8. For this analysis, the signal process $Z \rightarrow ee$ was simulated using POWHEG [22] and PYTHIA version 8 [23], see also appendix A.1 for a list of used data and MC samples.

After event generators provide a description of all particles that can be produced, the second step is simulating how these particles interact with the detector, mostly using GEANT4 [25]. In addition to the detector description, which needs to be as accurate as possible, a good set of physics models needs to be included. They describe possible processes induced by passing particles, like ionization, bremsstrahlung or Compton scattering. This includes interactions with sensitive detectors, which are used to detect a signal [26].

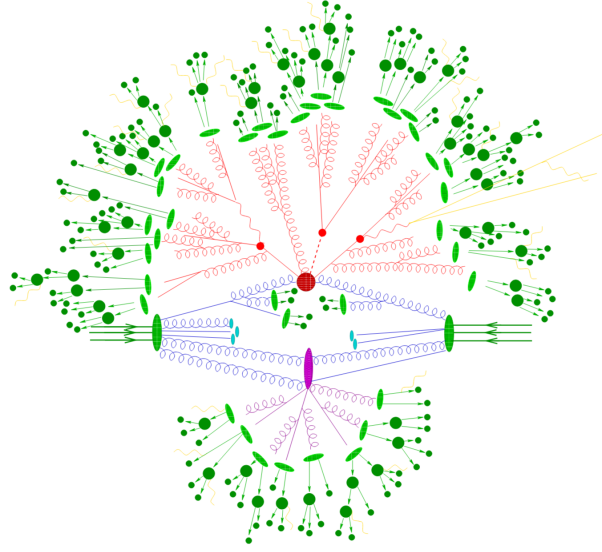


Figure 2.8: Schematic of a $t\bar{t}h$ event produced by an event generator. Starting from a hard interaction (large dark red dot), the top quarks and the Higgs boson decay (small red dots), with additional QCD radiation in red. Also, a secondary interaction takes place (purple). Hadronization and hadron decay is depicted in green as well as photon radiation in yellow [24].

The detector simulation is followed by digitization, converting the energy deposits with timings and locations from the detector simulation to read-outs like voltages and currents, producing an output similar to measured data. Simulated samples also include truth level information about the underlying event, which is not accessible via measured data.

With increasing amounts of data taken at ATLAS, more simulated events are needed as well to match the improved statistical power. This is computationally challenging. Detector simulation alone took up 38% of CPU hours consumed by the ATLAS experiment in 2018 [27]. Therefore, different tools are developed to save computational resources while still maintaining good accuracy. Especially the simulation of showers in calorimeters is complex to simulate due to the large number of secondary particles interacting with a complex detector geometry like the accordion structure in the EM calorimeter.

To address this, ATLAS fast simulation was introduced. The ATLFast3 setup for Run 3 can be seen in Figure 2.9. While for tracks in the inner detector GEANT4 is kept, two other methods were developed to simulate showers in the calorimeter system. FastCaloSim V2 uses a parametrization approach to describe shower development [28].

Additionally, FastCaloGAN V2 employs generative adversarial networks (GANs) to generate realistic showers. In a GAN, two deep networks compete against each other during training, one generating showers and one trying to discriminate between artificial and real showers. The kinetic energy of the particle is used as a conditional parameter of the network. 300 GANs

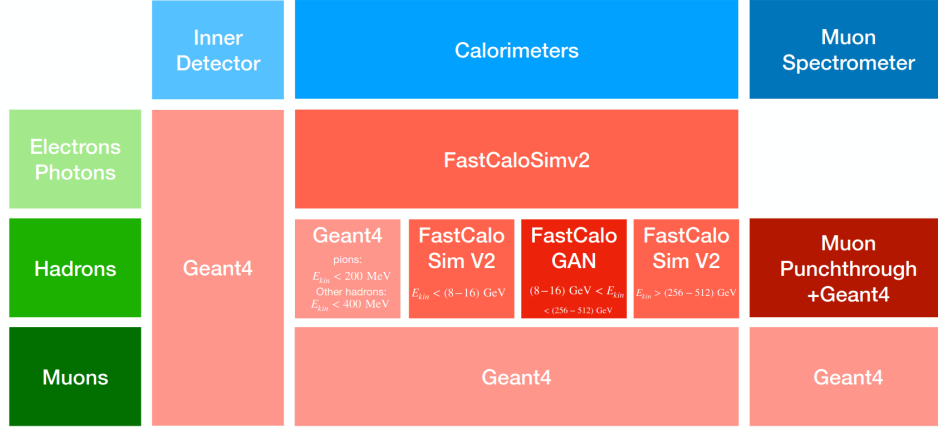


Figure 2.9: Run 2 configuration of different AF3 tools used, depending on the detector part, particle type and energy [28].

are used in total, each corresponding to a specific particle type and $|\eta|$ slice.

Also, secondary particles from hadronic showers entering the muon spectrometer need to be taken into account. This so-called muon punchthrough affects the background for reconstructed muons. The calorimeter response parametrization introduced by AF3 removes some information about the path of particles. Therefore, a dedicated treatment of these objects is needed to model the response of the muon spectrometer accurately.

ATLFAST3 allows accurate simulation of events, usually agreeing with full GEANT4 simulation within a few percent, while only requiring about 20 % of computational resources. How well AF3 approximates full simulation in the context of electron reconstruction is investigated in section 4.3.4.

3 Electrons at ATLAS

Between reading out and processing detector information and final objects used in physics analyses, several steps are needed, combining information from all sub-detectors. In this chapter, the sequence of steps needed to reconstruct and select objects is explained, with a focus on electrons.

3.1 Trigger

At the LHC, bunches collide every 25 ns, leading to an event rate of 40 MHz, which is several orders of magnitude higher than what can be fully processed and saved. A two-level trigger system is used to only select events of interest, reducing the event rate significantly. The first-level (L1) trigger decides within $2.5\,\mu\text{s}$ whether to keep an event or not, using only calorimeter and muon chamber data with reduced granularity. It achieves a reduction to below 100 kHz. Accepted events enter the high-level trigger (HLT), which uses information from all detector parts and performs an initial reconstruction of the objects, called online reconstruction, as well as preliminary identification. The measurements of reconstruction efficiencies carried out in this thesis, always refer to the final 'offline' reconstruction. With a few seconds of compute time per event, the high-level trigger reduces the final event rate to about 1 kHz. There are numerous different triggers corresponding to different particle types or particle combinations and certain p_T or E_T thresholds [29].

3.2 Reconstruction

The signature of an electron passing the ATLAS detector consists of hits in the inner detector and energy deposits in the EM calorimeter, with some possible leakage into the hadronic calorimeter. Information from both subdetectors needs to be combined to construct electron objects.

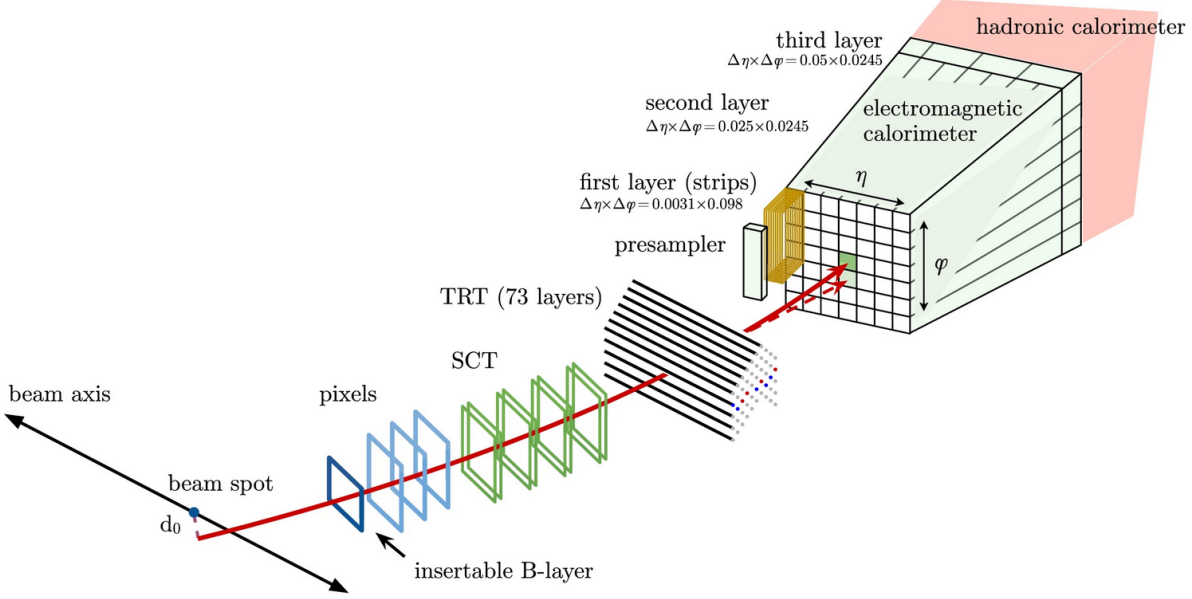


Figure 3.1: Schematic of a path of an electron (red solid line) through the ATLAS detector with emission of a photon due to bremsstrahlung (red dotted line) [30].

Cluster formation: The reconstruction begins with the combination of EM calorimeter cells into variable-sized clusters, called EM topological clusters. The clustering is based on a significance, defined as the ratio of cell energy $E_{\text{cell}}^{\text{EM}}$ and a noise threshold $\sigma_{\text{noise, cell}}^{\text{EM}}$:

$$\zeta_{\text{cell}}^{\text{EM}} = \frac{E_{\text{cell}}^{\text{EM}}}{\sigma_{\text{noise, cell}}^{\text{EM}}}. \quad (3.1)$$

To initiate the cluster, a cell needs a significance of $|\zeta_{\text{cell}}^{\text{EM}}| \geq 4$, then neighboring cells with $|\zeta_{\text{cell}}^{\text{EM}}| \geq 2$ are iteratively added to the cluster, collecting all connected cells with sufficient significance. Finally, a last iteration adds all neighboring cells without any energy requirements ($|\zeta_{\text{cell}}^{\text{EM}}| \geq 0$). These thresholds define the so-called '4-2-0' topological cluster reconstruction. Overlapping clusters are merged and if there are several local maxima with $E_{\text{cell}}^{\text{EM}} > 500 \text{ MeV}$ they are split. These initial topological clusters are the basis for reconstruction, but not the final clusters used [31].

Tracking: Initially independent from the clusters, tracks are reconstructed from individual hits in the inner detector. The granular hits in the pixel and SCT detector are used to construct track seeds from at least three space points. A pattern-recognition algorithm now extends all initial track seeds to full tracks of at least eight silicon hits, otherwise a second attempt with broader requirements is performed. The track is then extended into the TRT. If several tracks share hits, these ambiguities are resolved. A Gaussian-sum filter is applied to better account

for energy losses, especially due to bremsstrahlung [32].

Tracks are then loosely matched to EM clusters. If different tracks lead to one cluster, they are ranked, preferring tracks with hits in the pixel detector and small ΔR between track and cluster. Electrons radiating away a photon in the inner detector complicate electron reconstruction, since the photon can be measured as an independent object. While the initial unconverted photon has no track in the inner detector, it can create a separated cluster in the EM calorimeter. Since all energy deposits in the calorimeter need to be collected to achieve precise results, the clustering needs to take these radiated photons into account.

Superclusters: In a 3×5 window in the $\Delta\eta\text{-}\Delta\phi$ plane in units of 0.025, EM topological clusters in the proximity of an initial high E_T seed cluster are added as satellite clusters. The seed for an electron needs to have E_T of at least 1 GeV. The satellites can originate from bremsstrahlung or topological cluster splitting. Satellites in a 3×5 window are not required to have the same track matched, especially since due to photon conversion the satellite cluster might have its own track. If seed and satellite share the same track, the window is extended to a 5×12 $\Delta\eta\text{-}\Delta\phi$ window, see also Fig. 3.2. The seed and all satellites are combined into superclusters, which are then used for the determination of the electron energy [31].

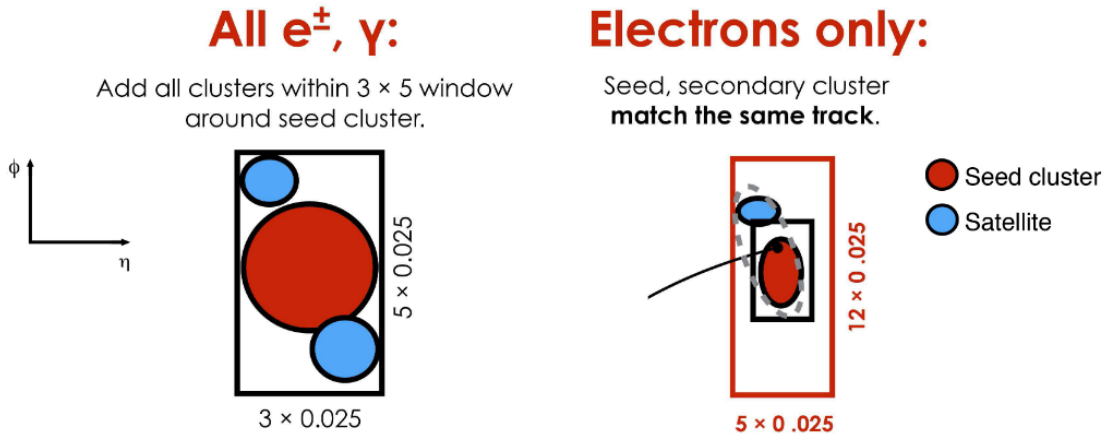


Figure 3.2: Drawing of supercluster algorithm concept, showing initial 3×5 $\eta\text{-}\phi$ window, expanded for electrons to 5×12 if seed and satellite clusters have same track as best matched track [30].

Finally, analysis-level electrons are created from tracks in the inner detector, superclusters in the EM calorimeter and potentially clusters in the hadronic calorimeter. The final η and ϕ directions are taken from the track best matched to the cluster [30].

3.3 Energy Calibration

While the reconstruction step relies on approximate EM-scale energy clusters, the final energy of electron candidates is calibrated after the reconstruction step, since it is done separately for electrons and photons. Photons are further divided into converted and unconverted photons, see section 3.6.1. The electron and photon calibration procedure is summarized in Fig. 3.3. Several scale corrections α_i and resolution corrections c_i are performed:

$$E_{\text{corr},i} = \frac{E_i}{1 + \alpha_i} \quad (3.2)$$

$$\left(\frac{\Delta E_i}{E_i} \right)_{\text{corr}} = \frac{\Delta E_i}{E_i} + c_i \quad (3.3)$$

Based on the ratio of data and MC energy deposits, the different layers in the calorimeter receive separate scale corrections α_{PS} for the presampler and α_{12} for the first and second layer. For data, additional corrections are applied, e.g. due to non uniformities in ϕ or biases. Finally, ee pairs from Z decays are used for final adjustments: after calibration the ee invariant mass distributions should match between data and MC. Via a scale correction α the data peak position is shifted towards MC. The resolution correction (c) is applied to MC to smear the simulated resolution towards the data resolution. The corrections obtained in $Z \rightarrow ee$ are then validated against J/ψ results. For more details, see [33].

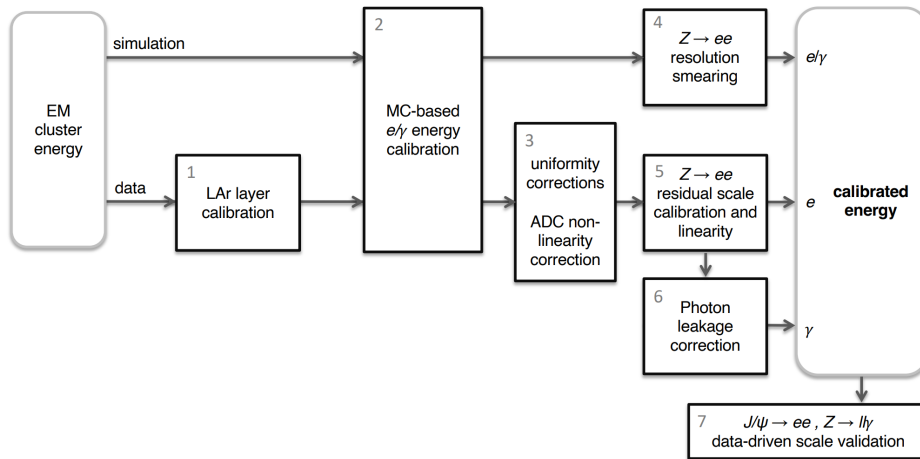


Figure 3.3: Overview of electron and photon energy calibration procedure at ATLAS [33].

3.4 Identification

Not every object reconstructed as an electron is a prompt electron required by analyses. Backgrounds include pile-up electrons and non-electron objects, so-called fakes. Here, other objects are reconstructed as electrons, like hadronic jets. This is especially challenging since jets can produce genuine electrons as secondary particles. To suppress these backgrounds while keeping as much prompt electrons as possible, identification and isolation are used, beside other analysis-specific cuts. The e/gamma combined performance (CP) group provides separate identification working points based on a likelihood-based (LH) or on a deep neural network (DNN) based approach. Several variables from tracks and clusters are used, like hadronic leakage, shape of the EM shower and transition radiation. For a more detailed description, see for instance [30].

3.5 Isolation

Many non-prompt backgrounds, like jets or electrons from photon conversions, are not as well separated from other objects in their proximity compared to electrons. Isolation uses this to further suppress backgrounds and can be based on tracks, clusters or a combination of both. For the isolation variables, a cone with a radius in ΔR is defined, then the transverse momenta for tracks and the transverse energy for clusters of the objects within the radius are summed up and the initial electron is subtracted:

$$p_{\text{T}}^{\text{cone}} = \sum_{i \in \text{cone}} p_{\text{T},i} - p_{\text{T},\text{Signal}} \quad (3.4)$$

$$E_{\text{T}}^{\text{cone}} = \sum_{i \in \text{cone}} E_{\text{T},i} - \sum_{j \in \text{core}} E_{\text{T},j} \quad (3.5)$$

Similar to identification, working points like *Tight* or *LooseTrackOnly* with different target efficiencies are defined, but it is also possible to apply selection cuts using these isolation variables.

3.6 Relation to other High-Level Physics Objects

3.6.1 Photons

Photons and electrons are included in a joint reconstruction algorithm due to their similarities, which is summarized in Fig. 3.4. Both deposit their energy mostly in the EM calorimeter. While photons initially do not leave a track in the inner detector, electron-positron pair-production can occur, resulting in tracks in the inner detector. This so called photon conversion results in the reconstruction into two separate categories: converted and unconverted photons. For converted photons, again, superclusters are used to collect all topological clusters corresponding to an initial photon.

Since electron and photon superclusters are built independently, it is possible that a seed cluster corresponds both to a photon and an electron candidate. An additional algorithm tries to resolve the resulting ambiguity. A supercluster is categorized as a photon, if only short tracks are attached to the clusters and a conversion vertex is reconstructed in the event. An electron is expected to have no good conversion vertex and a track with hits in the innermost parts of the inner detector. If a supercluster can not be clearly classified into electron or photon both objects are created for physics analyses [34].

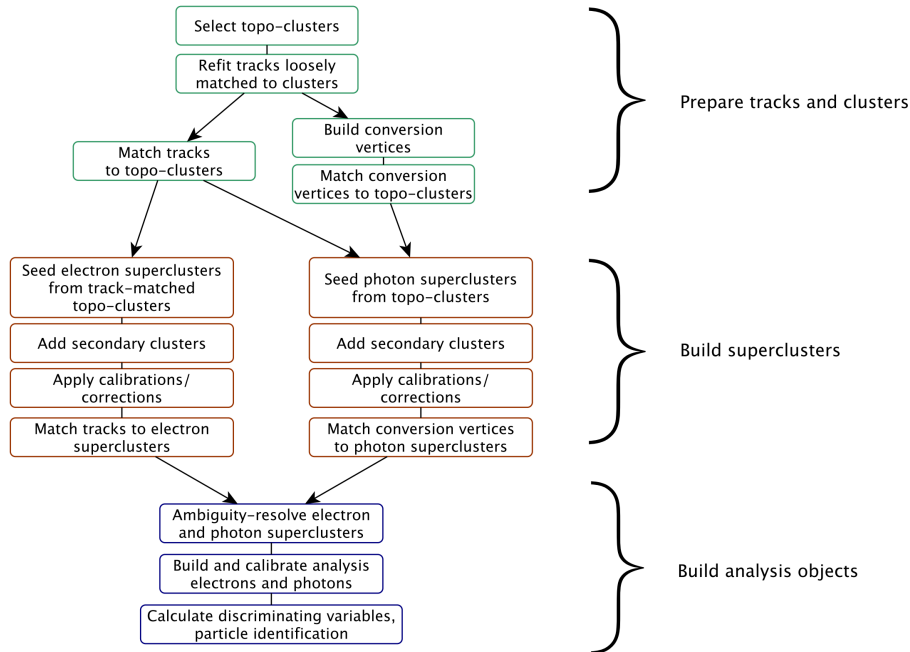


Figure 3.4: Algorithm flow for electron reconstruction (left) and photon reconstruction (right) at ATLAS [31].

3.6.2 Taus

Tau leptons are important signatures at ATLAS, but due to their short lifetime of 2.9×10^{-13} s, they decay mostly inside the beam-pipe and can not be measured directly. Approximately 35 % of the time, tau leptons decay leptonically and are usually reconstructed as prompt electrons or muons. Additional selection criteria are then needed to achieve separation between genuine prompt light leptons and leptons from τ decays, e.g. requiring E_T^{miss} .

It is possible, however, to reconstruct τ_{had} candidates from their hadronic decay modes, which correspond a total branching fraction of 65 %. The reconstruction is seeded from calorimeter jets and identification criteria allow the distinction between τ -jets and quark- and gluon-jets. The final-state of hadronic tau decays typically contains one or three charged pions. The decay into one charged pion creates a similar signature in the detector as electrons. Therefore, electrons can be misidentified as taus and a separate electron rejection algorithm is applied to suppress electrons for τ candidates [35].

3.7 Electron Efficiencies

Each of the mentioned steps to reconstruct an electron and reject backgrounds has the objective to keep as many genuine prompt electrons as possible. Efficiencies are useful quantities to characterize the performance of these steps. The total efficiency of triggering, reconstructing and selecting an electron passing identification and isolation can be factorized as:

$$\begin{aligned} \varepsilon_{\text{total}} &= \varepsilon_{\text{cluster}} \times \varepsilon_{\text{reco}} \times \varepsilon_{\text{id}} \times \varepsilon_{\text{iso}} \times \varepsilon_{\text{trigger}} \\ &= \frac{N_{\text{cluster}}}{N_{\text{all}}} \times \frac{N_{\text{reco}}}{N_{\text{cluster}}} \times \frac{N_{\text{id}}}{N_{\text{reco}}} \times \frac{N_{\text{iso}}}{N_{\text{id}}} \times \frac{N_{\text{trigger}}}{N_{\text{iso}}} \end{aligned} \quad (3.6)$$

The cluster efficiency $\varepsilon_{\text{cluster}}$ is the probability to successfully reconstruct an EM cluster from a genuine electron. Since EM clusters are the starting point for the reconstruction algorithm, this needs to be taken into account as well. From data, it is not possible to estimate this efficiency, since there is no truth information about electrons, so these studies are only done in Monte Carlo simulation. This efficiency has been found to be well above 99 % for $p_T > 10$ GeV, see also Fig. 3.5 [31].

Starting from EM clusters, the reconstruction efficiency $\varepsilon_{\text{reco}}$ is the probability of successfully

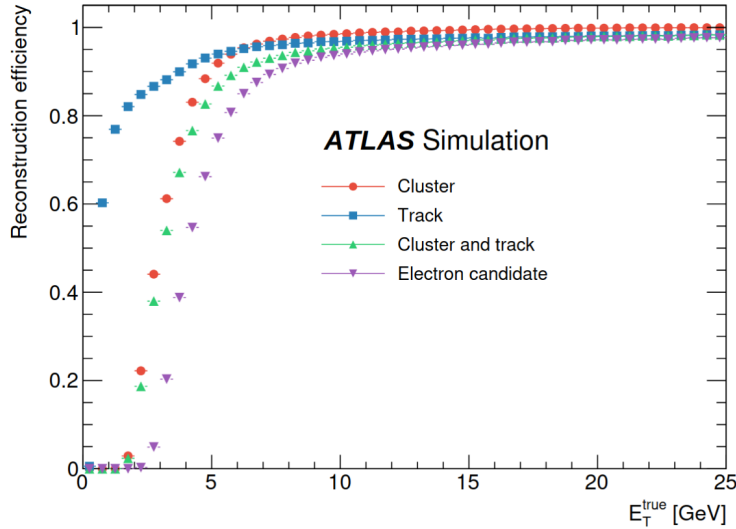


Figure 3.5: Electron cluster, track, cluster and track as well as overall reconstruction efficiency as a function of truth level electron E_T [31].

reconstructing electron objects. This consists of 1) track reconstruction and 2) matching of track and cluster. The identification efficiency ε_{id} represents how many reconstructed electrons pass a specific ID working point. ε_{iso} characterizes isolation performance accordingly.

Even though chronologically, triggering would be the first step, it is the last term in equation 3.7. In order to be able to estimate the trigger efficiency, one first needs the number of reconstructed, identified and isolated electrons after applying the offline algorithms N_{iso} and then measures how many of these electrons pass the trigger requirements N_{trigger} [29].

While for the first two efficiencies there is only one configuration, the identification, isolation and trigger efficiencies are measured with respect to each other, so all possible combinations of identification and isolation working points as well as triggers need to be measured separately.

A major motivation for calculating these efficiencies is the comparison between data and Monte Carlo. Physics analyses rely heavily on precise comparison between data and simulation, which should match as closely as possible. But since simulation is only an approximation, the electron efficiencies might differ significantly between data and Monte Carlo. Determining the efficiencies separately in data and MC allows the calculation of scale factors:

$$\text{SF}_i = \frac{\varepsilon_{i,\text{data}}}{\varepsilon_{i,\text{MC}}}. \quad (3.7)$$

These can then be multiplied to MC event yields to get corrected predictions, which are

comparable with data. These electron scale factors are calculated at ATLAS by the e/gamma combined performance (CP) group and then distributed to physics analyses using electrons in the final-state. Since not only the scale factors but also their uncertainties are propagated, it is crucial that the efficiency measurements are well understood with acceptable uncertainties. In the next chapter, the measurement of the electron reconstruction efficiencies in data and Monte Carlo as well as the derivation of scale factors is described.

4 Measurement of Electron Reconstruction Efficiency

The electron reconstruction efficiency was measured continuously since Run 1. The measurement framework used in this analysis is based on the first implementation for Run 1 [36], which was then successively improved and applied to more data [37–39]. The current analysis is embedded within the e/gamma **TagAndProbe** framework, to which several contributions were made during this work.

The latest measurements presented in [39] include initial Run 2 ATHENA software release 22 results. In release 22, several changes were introduced compared to release 21, including major updates to the reconstruction algorithms. Therefore, a new set of scale factors specific to the new software release is needed. The first objective of this work is to reproduce the initial findings, achieve stable and well-understood results and to provide new scale factors for recommendations with an updated uncertainty estimation. Regarding Run 3, only a first look at early 2022 data has been carried out in previous measurements, showing several changes compared to Run 2. The results presented in this work include measurements with the full datasets of 2022, 2023 and 2024 as well as some investigations into not yet understood changes and improvements of the analysis.

4.1 Methodology

4.1.1 Tag-and-Probe Method

Well-understood resonances like $Z \rightarrow e^+e^-$ or $J/\psi \rightarrow e^+e^-$ are used to obtain an unbiased sample of electron-positron pairs from their decays. In the following, both will just be referred to as electrons unless explicitly stated otherwise. The tag-and-probe method is then applied to estimate electron efficiencies. Here, a system of two electrons originating from the same decay is considered. One electron has to pass strict selection criteria, the so-called tag. A second

object is assigned to this tag electron by requiring a combined invariant mass close to the resonance, in this case m_Z . This loosely matches the two objects to the decay electrons. For the second object, as loose criteria as possible are chosen. This allows a unbiased determination of the rate in which the reconstruction, identification etc. is successful for the second electron, the probe object.

The choice of the resonance determines the energy range in which precise results are obtained. The electron reconstruction efficiency is measured using decays of the Z -boson with a mass of $m_Z = 91.2 \text{ GeV}$ in a transverse momentum range of $15 \text{ GeV} \leq p_T \leq 250 \text{ GeV}$. To get enough statistics for electrons with lower transverse momentum, lighter resonances like J/ψ with $m_{J/\psi} = 3.1 \text{ GeV}$ would be necessary.

For reconstruction, EM clusters are the starting point, therefore, all EM clusters are considered as possible probes. If the reconstruction was successful, the probe is a reconstructed electron candidate. Failed reconstruction corresponds to EM clusters without a matched track as discussed in section 3.7, which results in a reconstruction as a photon candidate.

4.1.2 Efficiency Definition

As discussed in the previous section, the reconstruction efficiency is the number of reconstructed electrons N_{reco} divided by the number of all EM clusters N_{clusters} corresponding to genuine electrons. While it is not strictly necessary to be reconstructed as an analysis-level electron candidate, additional track and cluster quality requirements are introduced for the efficiency calculation. A good quality track needs to be reconstructed with at least one hit in the pixel detector and at least eight silicon hits total (pixel+SCT). Additionally, the clusters need to pass certain data quality cuts, see sec. 4.3.3. Analyses using these scale factors are expected to apply these requirements as well. This results in the following calculation of the electron reconstruction efficiency:

$$\varepsilon_{\text{reco}} = \frac{N_{\text{Quality}} - B_{\text{Quality}}}{(N_{\text{Quality}} - B_{\text{Quality}}) + (N_{\text{NoQuality}} - B_{\text{NoQuality}}) + (N_{\text{NoTrack}} - B_{\text{NoTrack}})}. \quad (4.1)$$

In equation 4.1, N_X corresponds to the total number of probes in the category X from which the number of background probes B_X is subtracted to filter for signal electrons from Z decays. In the numerator, only successfully reconstructed *Quality* probes are considered, with a matched track and passing quality requirements. In the denominator there are additional terms for probes failing the requirements. They can be split into *NoQuality* probes, which are electron

candidates that fail track or cluster quality requirements and *NoTrack* probes. Here, the track matching failed and genuine electrons ended up being reconstructed as photon candidates instead.

4.1.3 Event Selection and Binning

Table 4.1 shows the object selection criteria motivated in the last sections. The tag has to fulfill stricter η and p_T requirements and, most notably, needs to pass Tight identification to ensure it is a genuine electron. Since the calorimeter transition region between the detector endcap and barrel is not homogeneously covered by the tracking and calorimeter detectors, it is excluded for the tag. Probes in this region are not vetoed, therefore, the scale factors are provided in the transition region, even though it is excluded in most physics analyses. Notably, the selection of electron objects for tag and electronic probes also includes the ambiguous category described in section 3.6.1, which collects objects where no clear classification into electron or converted photon was possible.

An invariant mass in a narrow window ± 10 GeV around m_Z is required in the nominal setup. Currently, no opposite charge requirements between the tag and electronic probe are applied. This is done mostly for consistency reasons, since photon candidates have no charge but should be treated as equally to electron probes as possible. Requiring an opposite charge has been shown to reduce the backgrounds, with only minor impact on efficiencies. If there are different possible tag and probe pairs in one event, the pair with the highest invariant mass is chosen to resolve the ambiguity.

Since the reconstruction efficiency varies greatly depending on different locations in the detector and electron momentum, it is measured separately in 11 p_T and 20 η bins to provide specialized scale factors. The binning is shown in Tables 4.2 and 4.3. In the context of increasing statistics in Run 3, a finer binning, especially in η , might be interesting in the future.

Table 4.1: Object level selection criteria for tag and three types of probes.

Tag Baseline	Probe		
	Quality	NoQuality	NoTrack
$ \eta \leq 2.47$ $p_T \geq 25$ GeV Tight Identification Matched to trigger electron Electron Pass Calo Quality Veto: $1.37 \leq \eta \leq 1.52$	$ \eta \leq 2.47$ $p_T \geq 15$ GeV Distinct from Tag Veto Jets with $\Delta R \leq 0.4$		
	Electron Pass Calo Quality Pass Track Quality	Electron No Calo Quality No Track Quality	Photon $\Delta R > 0.15$ to any e MC only: $\Delta R < 0.2$ to true electron

Instead of η itself, which is reconstructed mainly from the track (if there is a track), the pseudorapidity η_{BE2} of the primary cluster barycenter in the second layer of the EM calorimeter is used. This is a more suitable variable, since the reconstruction algorithms start with EM clusters. For simplicity, in the rest of the thesis, η_{BE2} is just referred to as η .

4.1.4 Background Estimation

While the total number of probes N_X can be measured directly, different techniques are needed to estimate the backgrounds B_X in equation 4.1. Misreconstructed objects play an important role in the backgrounds as well as the 'photonic' signal produced by genuine electrons. These fakes are usually hard to simulate precisely, so data-driven methods are used for background estimation. The invariant mass distribution close to the Z peak is the basis for the background estimation, exemplarily shown in Figure 4.1.

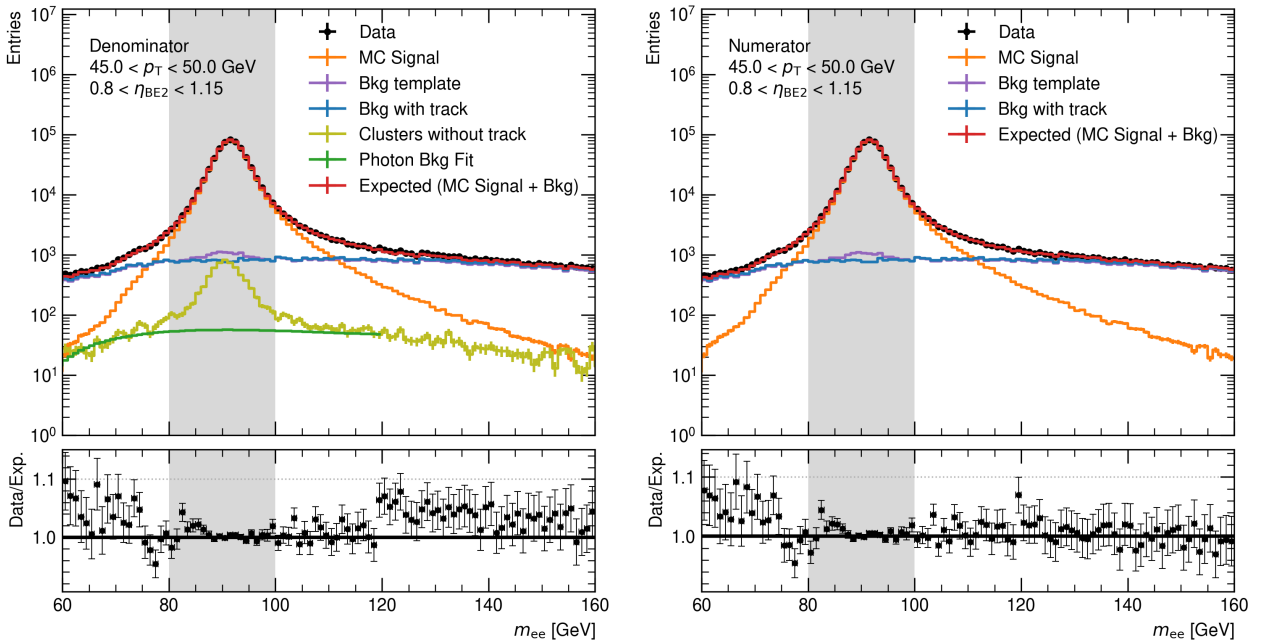


Figure 4.1: Invariant mass distribution of tag-and-probe pairs for probes in the reconstruction efficiency denominator (left) and numerator (right). The estimation for backgrounds with track (blue) are derived from a background template (purple). For clusters without track (light green), a polynomial fit is performed on the sidebands to estimate the background contribution (dark green), which is included in the total background and signal estimate (red) up to $m_{ee} = 120$ GeV. The signal region is shaded in grey.

Table 4.2: p_T bin boundaries used for electron reconstruction efficiency measurements.

p_T bin boundaries in GeV												
15	20	25	30	35	40	45	50	60	80	150	250	

Table 4.3: η bin boundaries used for electron reconstruction efficiency measurements, $+\eta$ and $-\eta$ are measured separately.

$ \eta $ bin boundaries									
0	0.1	0.6	0.8	1.15	1.37	1.81	2.01	2.37	2.47

Probes with associated track

Probes with a matched track are divided into probes passing the track and cluster quality requirements N_{Quality} and probes failing these requirements $N_{\text{NoQuality}}$. Their respective backgrounds B_{Quality} and $B_{\text{NoQuality}}$ are measured independently but following the same principle. In Fig. 4.1 both terms are included in the denominator (left), while only probes passing quality requirements contribute to the numerator (right).

First, a background-enriched region is defined, called template, by requiring the probes to fail loose electron identification criteria as well as loose isolation criteria. This template is shown in purple in the invariant mass distribution. It models the shape of the background with minimal electron signal. The remaining signal contamination in the template is removed using MC simulation. For each $p_{\text{T}}\text{-}\eta$ bin, a scale value is derived to normalize MC to data using probes passing tight identification with an invariant mass close to m_Z :

$$N_{\text{clean}} = N_{\text{data}} - N_{\text{signal,mc}} \frac{N_{\text{data,tight}}}{N_{\text{MC,tight}}}. \quad (4.2)$$

After estimating the background distribution shape using the template, the total amount of background events is estimated by normalizing the template to data. Here again, signal contamination is subtracted from data using simulation. This is done in the tail regions of the invariant mass distribution, $m_{ee} \in [120, 250]$ GeV in the nominal setup.

$$B = N_{\text{template,clean}}^{\text{peak}} \frac{N_{\text{data,clean}}^{\text{tail}}}{N_{\text{template,clean}}^{\text{tail}}}. \quad (4.3)$$

The final background estimation after signal contamination removal and normalization in the tail is shown in blue in Fig. 4.1. It is then integrated in the signal region around m_Z (grey shaded) to calculate B_{Quality} and $B_{\text{NoQuality}}$. In most $p_{\text{T}}\text{-}\eta$ bins, the normalization factor is close to 1.0, indicating that most backgrounds are already included by the fail ID and fail isolation template definition. Monte Carlo based studies have found that hadronic jets misreconstructed as electrons represent the most relevant background contribution [30].

Probes without track

For photon candidates, a template-based approach could not be established yet. The background consists of true photons and other objects misreconstructed as photons, like hadrons, which is hard to isolate via specific selection criteria. Additionally, the signal corresponds to electrons misreconstructed as photon candidates. For these 'fake' objects, Monte Carlo simulation is less reliable.

The background is estimated using a polynomial fit on the observed distribution of clusters without track, shown in light green in Fig. 4.1. Again, the signal contamination is subtracted from data using MC simulation. To reduce the reliance on simulation, the fit is only performed in low- and high-mass sidebands, excluding the signal region. In the nominal setup, data in $m_{ee} \in [60, 80] \cup [100, 120]$ GeV is included for a third order polynomial fit. The fit result (dark green) is then integrated over the signal region to calculate B_{NoTrack} . Due to divergences for large masses outside the fit range, only $m_{ee} \leq 120$ GeV fit results are included in invariant mass plots and their final signal plus background estimate, which explains incontinuity 120 GeV at the Data/Exp. curve in Fig. 4.1.

4.2 Uncertainties

4.2.1 Statistical Uncertainties

The statistical uncertainty can be obtained by error propagation from equation 4.1:

$$\Delta\varepsilon_{\text{stat}} = \sqrt{(1 - \varepsilon_{\text{reco}})^2 (\Delta S_{\text{stat,Quality}})^2 + \varepsilon_{\text{reco}}^2 ((\Delta S_{\text{stat,noQuality}})^2 + (\Delta\Gamma)^2)}. \quad (4.4)$$

The statistical uncertainties for signal with tracks S_X , with X corresponding to electronic probes passing or failing quality requirements are given from:

$$(\Delta S_{\text{stat},X})^2 = (\Delta N_{X,\text{Peak}})^2 + \left(\frac{T_{X,\text{clean,Peak}}}{T_{X,\text{clean,Tail}}} \Delta N_{X,\text{clean,Tail}} \right)^2 + \left(\frac{N_{X,\text{clean,Tail}}}{T_{X,\text{clean,Tail}}} \Delta T_{X,\text{clean,Peak}} \right)^2 \quad (4.5)$$

$$+ \left(\frac{T_{X,\text{clean,Peak}} N_{X,\text{clean,Tail}}}{T_{X,\text{clean,Tail}}^2} \Delta T_{X,\text{clean,Tail}} \right)^2 \quad (4.6)$$

with N_X corresponding to total events and T_X to events fulfilling the template definition in the region around the Z peak or in the tail region for normalization. Additionally, for most terms, the event number cleaned from signal contamination using MC is considered. ΔN and ΔT are Poisson uncertainties.

For probes without tracks, here denoted as Γ to avoid confusion with the previous terms, a slightly different uncertainty calculation is needed, since a fit is used to estimate backgrounds. The signal probes are obtained from $\Gamma = N_{\gamma,\text{Peak}} - B_{\gamma,\text{Peak}}$ with $(\Delta \Gamma^2) = (\Delta N_{\gamma,\text{Peak}})^2 + (\Delta B_{\gamma})^2$. The uncertainty on the background from the fit is obtained by error propagation on the number of events n_i in each bin i within the fit range. At leading order, this is calculated as:

$$\Delta B_{\gamma} = \sqrt{\sum_i \left(\frac{\partial B_{\gamma}}{\partial n_i} \right)^2 (\Delta n_{\text{stat},i})^2} \quad (4.7)$$

with the Poisson uncertainty $\Delta n_{\text{stat},i}$ for the event number in each bin. For a more detailed derivation of the fit uncertainty, see also [36]. The statistical uncertainty for data efficiencies for all p_T - η bins can be found in Figure 4.2 for 2018. With the efficiency in the order of one, the relative stat. uncertainty is at the permille-level for most bins, enhanced for low and high p_T as expected for electrons from Z decays.

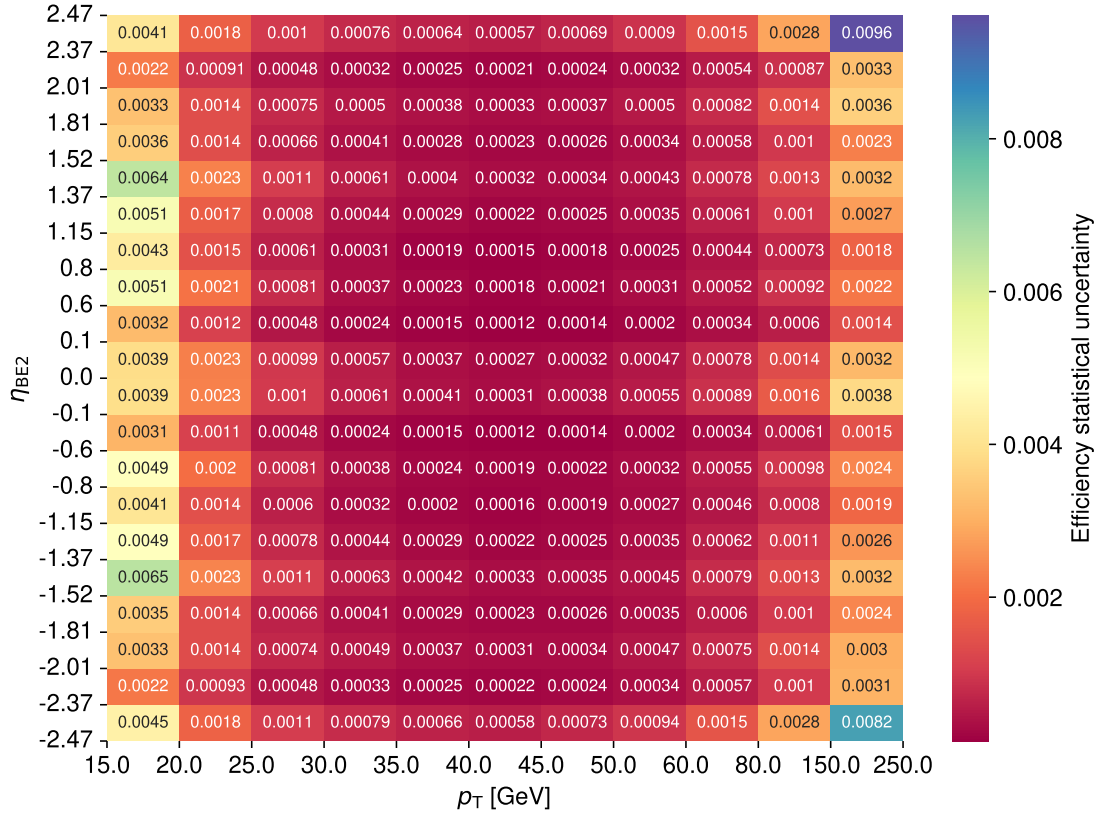


Figure 4.2: 2D p_T - η heatmap of statistical uncertainties on the reconstruction efficiency in 2018 data.

4.2.2 Systematical Uncertainties

Four different criteria are changed to define variations of the selection procedure, which is detailed in Table 4.4. The following analysis choices are affected:

- Template definition: variation of inverted ID and isolation cuts, which are used to define a background-enriched template and invariant mass region in which the template is normalized to data
- Tag definition: variation of identification and isolation cuts the tag needs to pass
- Integration range: variation of the invariant mass window around the Z peak, defining the signal region
- Photon sidebands: variation of low- and high-mass boundaries, in which the photonic background fit is performed

These variations allow the estimation of systematic uncertainties for this analysis. For previous measurements, all possible combinations of these four criteria were considered. Two template definitions, three tag definitions, three integration ranges and four photon sidebands resulted

in $2 \times 3 \times 3 \times 4 = 72$ different variations. The efficiency result and its statistical uncertainty are then the averages of these variations. The systematic uncertainty is then calculated as the root mean square (RMS) of the variations:

$$\Delta\epsilon_{\text{syst}} = \sqrt{\frac{1}{N} \sum_{i=0}^N (\bar{\epsilon} - \epsilon_i)^2}. \quad (4.8)$$

This approach has the disadvantage of introducing many strongly correlated variations, even though only a few variables were changed, complicating the interpretation of the results and the understanding of individual changes. That's why defining a new setup based on the initial variations was a major objective for this work. From the old variations, a nominal selection was chosen, which delivers stable results in all p_T - η bins. Then, only physically relevant variations are kept, differing from the nominal selection only in a specific definition. Efficiency result and statistical uncertainty are now measured directly using the nominal selection, while the systematic uncertainty loses the factor $\sqrt{1/N}$:

$$\Delta\epsilon_{\text{syst}} = \sqrt{\sum_{i=0}^N (\bar{\epsilon} - \epsilon_i)^2}. \quad (4.9)$$

An approach to systematics based on a nominal setup is more commonly used. Compared to the old setup, the number of variations is reduced from 72 to 6. Each variation now represents a change of one specific analysis choice. For integration range and photon fit range, up and down variations were defined, changing the same variable in opposing directions. The final result for recommendations is then symmetrized: from up-down variations, only the one with higher impact is kept and set as the uncertainty in both directions. The new setup is summarized in Table 4.4.

The deviation from the nominal efficiency and therefore the contribution to the systematic uncertainty is shown for all variations, exemplarily for 2018, in Figure 4.3. Variation of template and tag selection criteria only has a small uncertainty contribution with efficiency deviations up to 0.56 % at low p_T , well below 0.1 % above a p_T of 25 GeV and some fluctuations in the last p_T bin. Compared to other considered variations like tag medium ID, here, the efficiency trends in p_T and η are more structured and show less statistically dominated fluctuations.

Changing the invariant mass integration range (Z window) has a bigger impact, attaining efficiency deviations close to 1 % at low p_T . Here, any potential background mismodeling

Table 4.4: Description of selection criteria used in the nominal setup, together with up and down variations of the new measurement setup.

	nominal	up	down
Template	Fail LH ID UltraLoose Isolation $\frac{E_T^{\text{cone30}}}{E_T} > 0.02$ Normalization area: [120, 250] GeV for $p_T < 30$ GeV [120, 250] GeV for $p_T \geq 30$ GeV	Fail LH ID VeryLoose Isolation $\frac{E_T^{\text{cone30}}}{E_T} > 0.05$ Normalization area: [60, 70] GeV for $p_T < 30$ GeV [120, 250] GeV for $p_T \geq 30$ GeV	
Tag definition	Pass Tight ID	Pass Tight ID + Isolation	
Integration range	[80,100] GeV	[85,95] GeV	[75,105] GeV
Photon sidebands	[60,80] GeV + [100, 120] GeV	[70,80] GeV + [100, 110] GeV	[50,80] GeV + [100, 120] GeV

can contribute, including the photon fit. Especially narrowing down the window has a larger impact and is subject to heavier fluctuations due to reduced statistics.

The most relevant contribution to the systematic uncertainty comes from the photon background fit and the variation of its fit range. Fluctuations of up to 2.3% occur at low p_T , where the photon term itself has the highest impact. Since the fit is affected by the varied fit range, the up-down efficiency differences can become slightly unstable. Significant fluctuations indicate some instabilities in the fit in affected regions. For medium p_T in the range between 30 and 80 GeV, the contribution to the systematic uncertainty is well below 0.1% and in the same order of magnitude as the integration range variation.

Certain bins also show a difference of exactly zero, which is not expected for a changed fit. In the last p_T bin, there are no events to fit to in affected regions. Similarly, in the first p_T bin, in the central η region, there are almost no signal photons candidate to fit to which makes the background fit susceptible to overshooting the data slightly. If the background estimation exceeds the amount of total data, nonphysical efficiencies bigger than 1.0 would emerge, see also equation 4.1. To avoid this, the background event number is set to the data event number in affected p_T - η bins. If this happens for multiple variations, they will not show any efficiency difference between each other. This explains the lack of deviation for $p_T \in [15, 20]$ GeV $\eta \in [-0.1, 0.0]$ in Fig. 4.3e, which also results in a low total systematic uncertainty in this bin.

The total systematic uncertainty per p_T and η bin is shown in Fig. 4.4 for 2018 data. In comparison with the stat. uncertainty in Fig. 4.2 it is clear that the total uncertainty in this measurement is dominated by systematics at low p_T , reaching up to 3%. For $p_T > 25$ GeV, systematic and statistical uncertainty are in the same order of magnitude. Overall, the measurement of the electron efficiency reaches percent-level precision for $p_T < 25$ GeV and permille-level precision from 25 to 150 GeV.

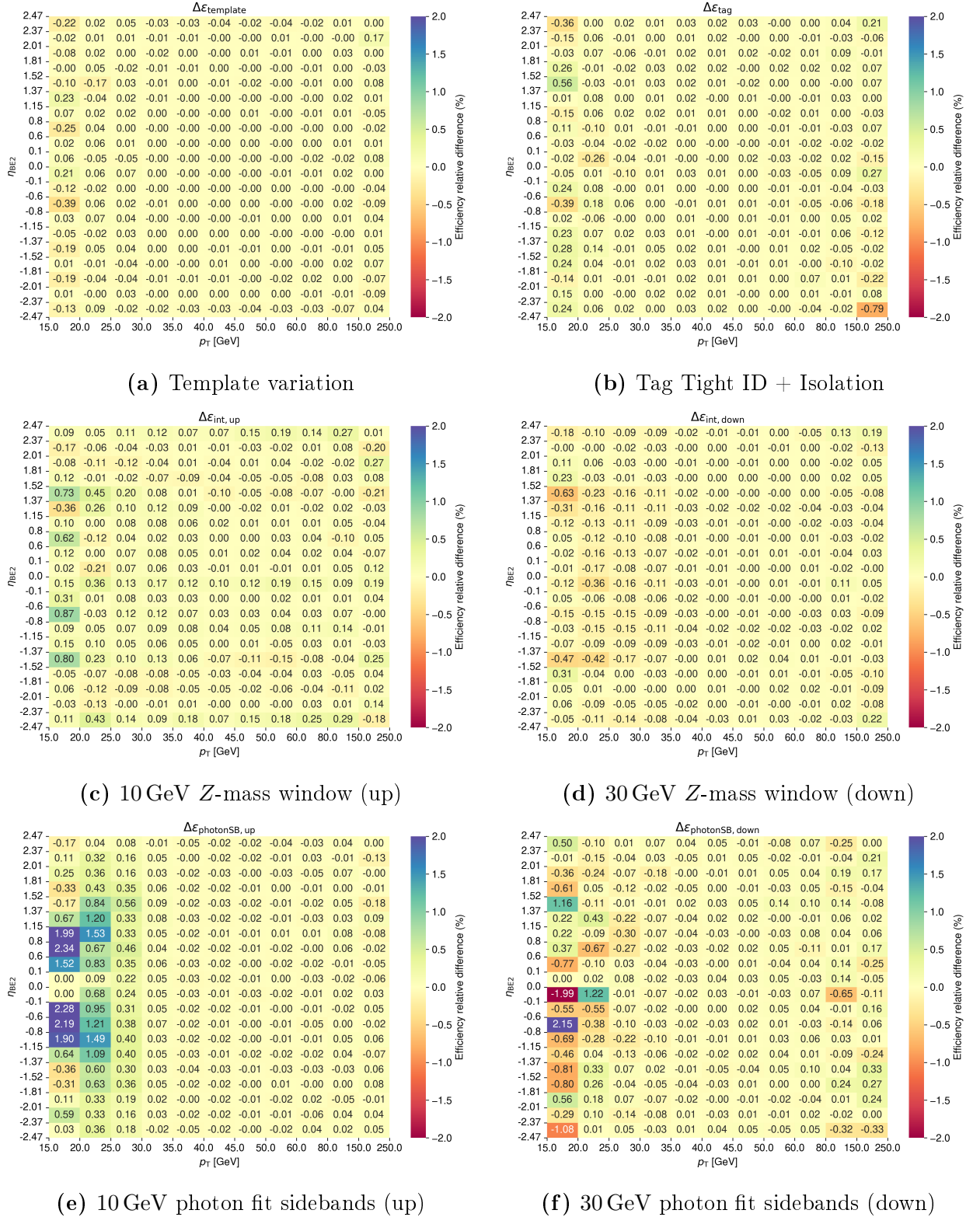


Figure 4.3: Heatmaps showing the efficiency difference of the reconstruction efficiency obtained in 2018 data in relation to the nominal setup for all p_T - η in %. All variations considered for the systematic uncertainty are depicted using the same color encoding.

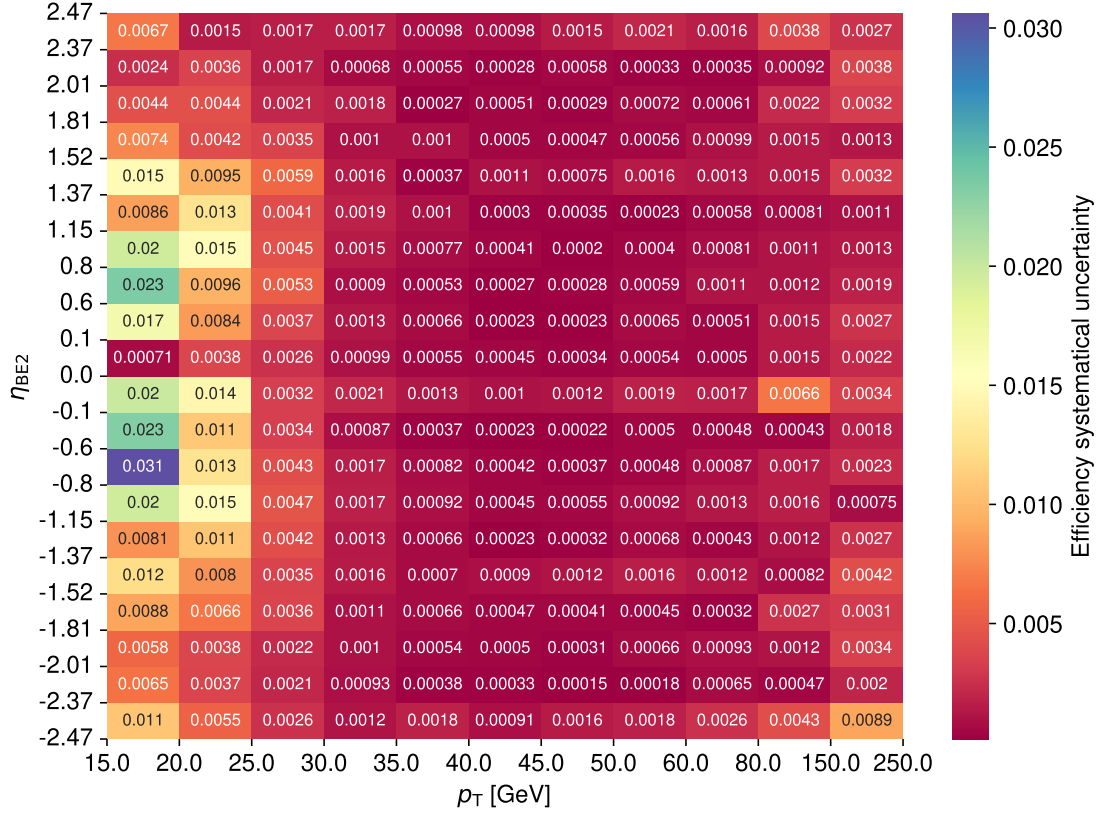


Figure 4.4: 2D p_T - η heatmap of systematical uncertainties on the reconstruction efficiency in 2018 data.

For Monte Carlo simulation, truth matching is used to identify signal probes. Therefore, systematics concerning the background estimation have no impact on MC, leaving currently only the variation on integration range, which has a small contribution in simulation. The scale factor and its uncertainty is computed using the same variations.

4.3 Efficiency and Scale Factor Results

4.3.1 Electron Reconstruction Efficiency as a Function of p_T and η

The final measurement of reconstruction efficiency is carried out for each data-taking year separately and divided into 220 kinematic p_T - η bins to provide more accurate scale factors. In this section, first the general p_T and η dependence of the efficiencies and scale factors is discussed based on the full Run 2 dataset. Figure 4.5 shows the electron reconstruction efficiency as a function of p_T (left) and η (right). Integrated over the whole kinematic phase space, a reconstruction efficiency of 98 % is obtained, showing excellent overall performance of the electron reconstruction algorithms.

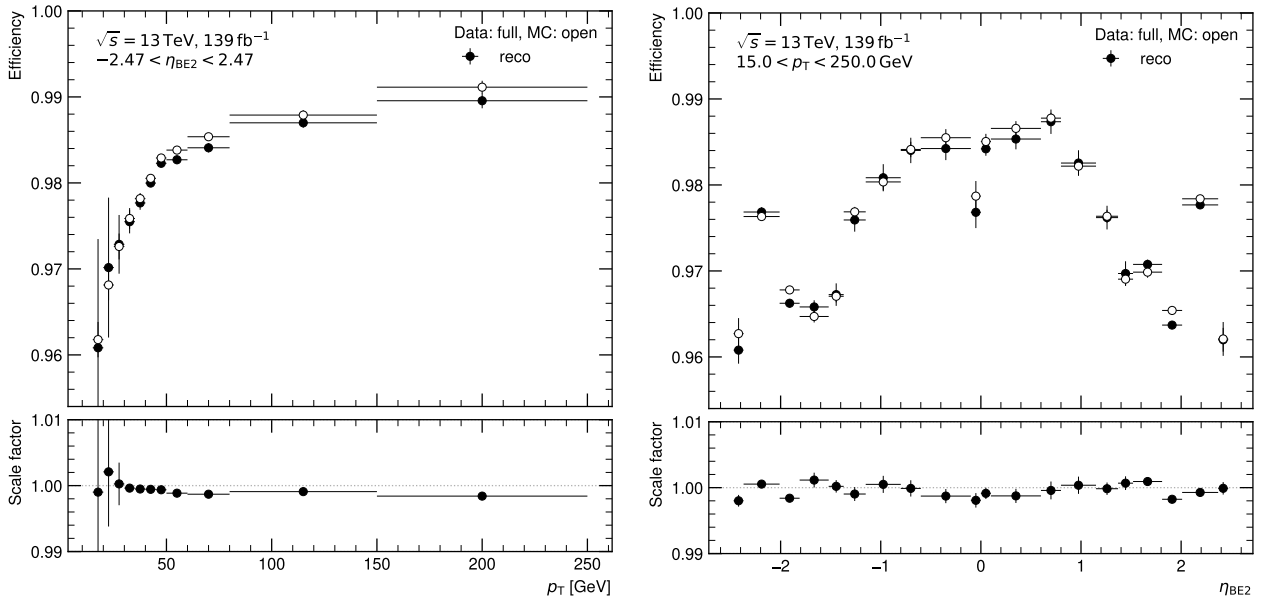


Figure 4.5: Electron reconstruction efficiency as a function of the transverse momentum integrated over the pseudorapidity range $-2.47 < \eta < 2.47$ (left) and as a function of η integrated over the p_T range $15.0 < p_T < 250.0 \text{ GeV}$ (right). The full Run 2 dataset was used. Data efficiencies correspond to filled dots and MC efficiencies correspond to open dots.

With decreasing p_T , the efficiency continuously decreases, approaching 96 % for a transverse momentum of 15 GeV. The reduction in efficiency is not evenly distributed across the η range, a characteristic shape can be observed. A plateau of efficiencies close to 99 % in the central $|\eta| < 1.0$ region is followed by minima around the calorimeter barrel-endcap region $1.37 < |\eta| < 1.57$. In the outermost eta bins ($2.37 < |\eta| < 2.47$) another minimum is reached.

To understand this behavior, it is useful to distinguish between the two possible reasons an electron is not considered successfully reconstructed in this measurement, which is presented in Fig. 4.6 for all 220 p_T - η bins. Here distributions in η are shown for every p_T bin side-by-side.

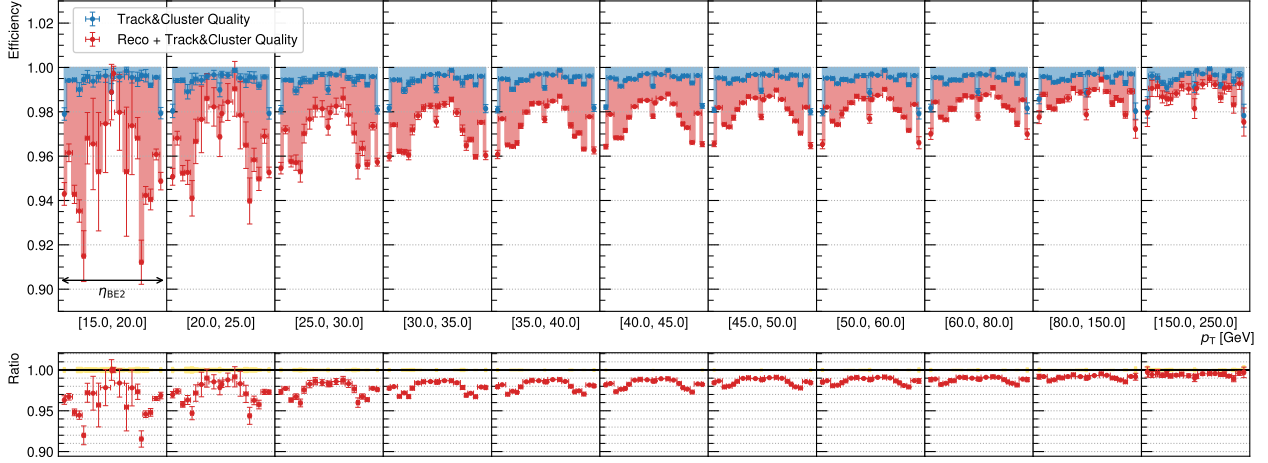


Figure 4.6: Electron reconstruction efficiency as a function of p_T and η for the full Run 2 dataset, with the distributions in η plotted side-by-side for each p_T bin. The full efficiency including track and cluster quality requirements is shown in red and the isolated impact of the track and cluster quality requirements is shown in blue.

Firstly, for an initial EM cluster, the track matching can fail, resulting in reconstruction as a photon candidate. The efficiency loss due to failed track matching corresponds to the red shaded area in Fig. 4.6. Additionally, a reconstructed electron candidate can fail the track and cluster quality requirements. The isolated impact these requirements have on the efficiency is shown in blue.

Failed track matching constitutes the dominant source of efficiency decrease in most kinematic bins. It causes the characteristic shape in η with a maximal efficiency decrease in the barrel-endcap transition region. Additionally, the probability of failed track matching is increased for decreasing transverse momentum. Physically, this behavior can be explained by bremsstrahlung. The radiation of a photon can complicate track matching, since the electron changes direction and additional clusters are created by the photon. Bremsstrahlung has a higher impact for low transverse momenta and depends on the amount of detector material passed by the electron, which is shown in Fig. 4.7 for a particle reaching the calorimeter at ATLAS.

The radiation length reaches its maximum in the barrel-endcap transition region at $|\eta| \approx 1.5$, which corresponds to the observed minima in the reconstruction efficiency. Therefore, bremsstrahlung causes both the p_T dependency as well as the shape in η of the electron reconstruction efficiency.

Superimposed to the failed track matching is the efficiency decrease due to track and cluster quality requirements in Figure 4.6. Here, an approximately p_T independent contribution is observed. In higher p_T regions, the quality requirements gain relevance due to decreased

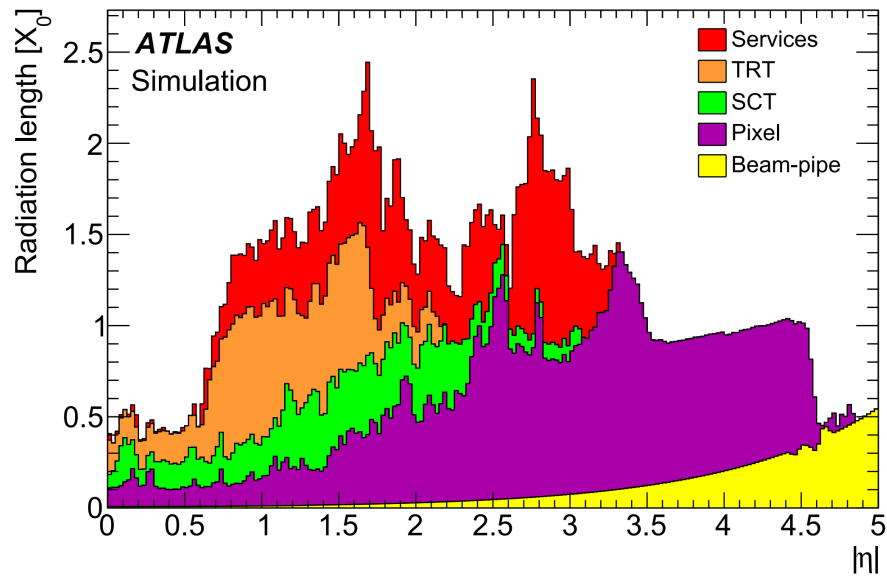


Figure 4.7: Material traversed by a particle up to the inner detector boundaries in units of radiation length X_0 as a function of η . Contributions of different detector elements are indicated by different colors [40].

impact of bremsstrahlung. While in most η bins only a small decrease of $\approx 0.5\%$ is observed, there are three discrete bins with higher impact, which is also visible in the total reconstruction efficiency: $|\eta| \in [2.37, 2.47]$ and $\eta \in [-0.1, 0.0]$. These discrete decreases are discussed in more detail in sec. 4.3.3.

4.3.2 Run 2 Release 22 Results

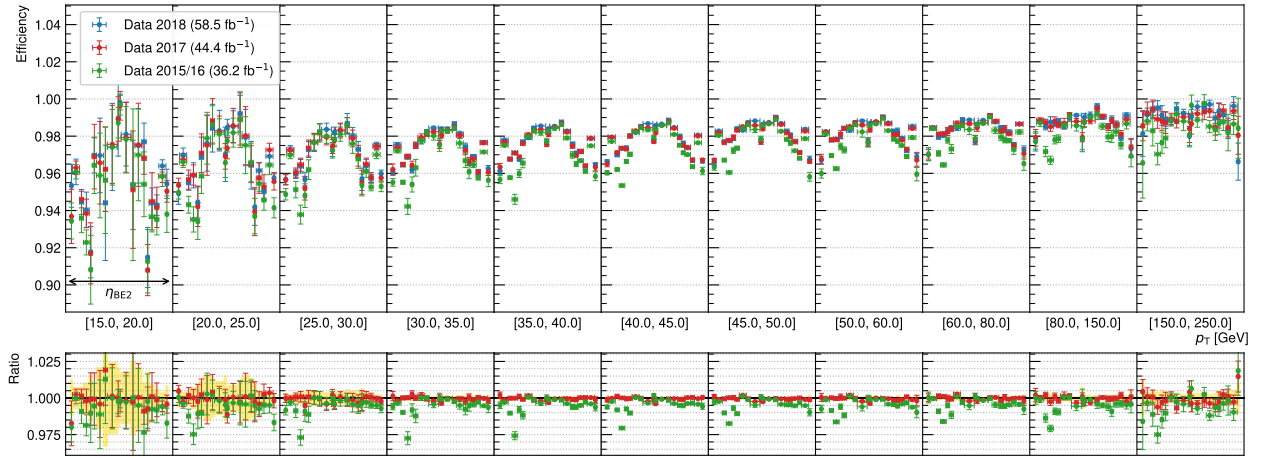
The Run 2 efficiency and scale factor results are included in Fig. 4.8, separated per data-taking year. The years 2015 and 2016 were merged since the 2015 data does not contain enough events to enable results with percent-level uncertainties or lower. Only 3.2 fb^{-1} of the 2015 data are used for this measurement, not including a bigger dataset with 50 ns bunch spacing instead of the usual 25 ns between pp collisions. The other years provide at least 30 fb^{-1} of integrated luminosity, allowing the partitioning into 220 kinematic p_T - η bins with sufficient statistics.

All efficiency distributions in Fig. 4.8 follow the expected shape: the efficiency increases with increasing p_T , while there is a characteristic shape in η with minima close to the transition region. In the first and last p_T bin, uncertainties due to low statistics are of similar magnitude as features in the η distribution, demonstrating that this method reaches its limitations. To expand the p_T range, other decays may be used, e.g. J/ψ to measure efficiencies for low p_T electrons, which is already done for identification efficiency measurements, but not yet for reconstruction.

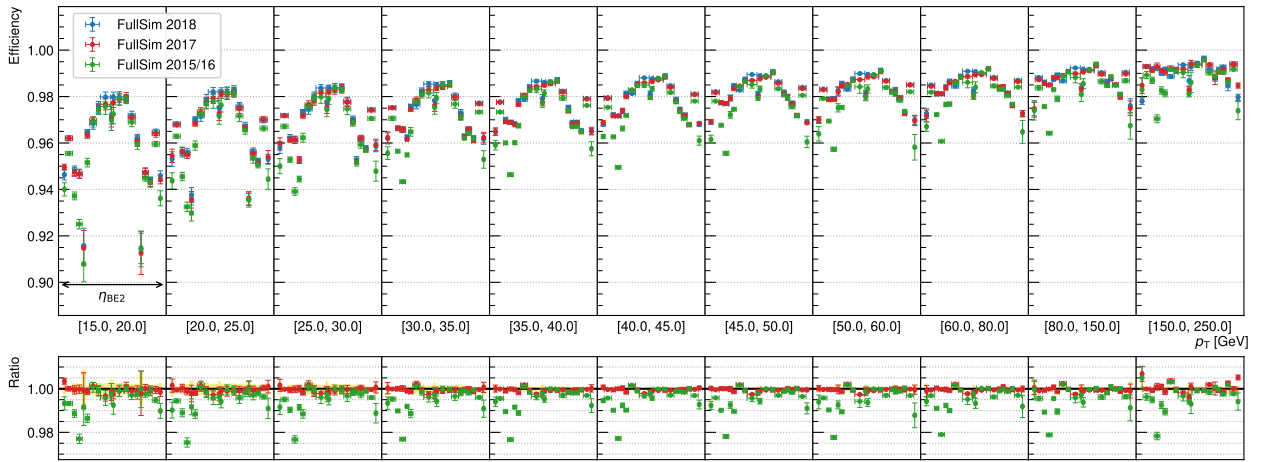
While good agreement between data and MC is observed in 2018 and 2017, 2015/16 shows a slightly different behavior. The efficiencies are lower over most parts of the η range, especially for negative η in the transition region $-1.57 < \eta < -1.37$ and neighboring bins. Over the whole p_T range, the efficiency decreases up to 2% in the affected transition region compared to other years and its $+\eta$ counterpart.

This is well described in MC, which results in scale factors close to one for all years. Deviations at low p_T are mostly included in the 1σ uncertainties while, especially for 2015/16, some significant excesses can be seen. Generally, deviations of the scale value from unity are not a problem, since it is not expected that the simulation matches data exactly, but it is important to understand anomalies. The physical reason for the 2015/16 efficiency drop and its slight mismatch between data and MC is discussed in section 4.3.3 together with similar occurrences in Run 3.

In Fig. 4.9, the obtained scale values are compared with the latest recommendations. The new measurement uses samples reprocessed with ATHENA software release 22, while the old results were produced using release 21. Towards release 22, the electron reconstruction algorithm was changed significantly, including reconstruction based fully on topological clusters, changes in tracking and conversion reconstruction. In comparison with release 21, the current scale factors shifted closer to unity, indicating improved agreement between data and MC. Since it produces results with acceptable uncertainties, the p_T range was expanded by one bin to 250 GeV. For more detailed comparisons with previous releases see [38, 39].



(a) Data efficiency



(b) FullSim MC efficiency



(c) Scale factor

Figure 4.8: Electron reconstruction efficiencies and scale factors as a function of p_T and η in Run 2. The 2018 results (blue) are compared with the results obtained in 2017 (red) and 2015/16 (green), including their ratio with respect to 2018.

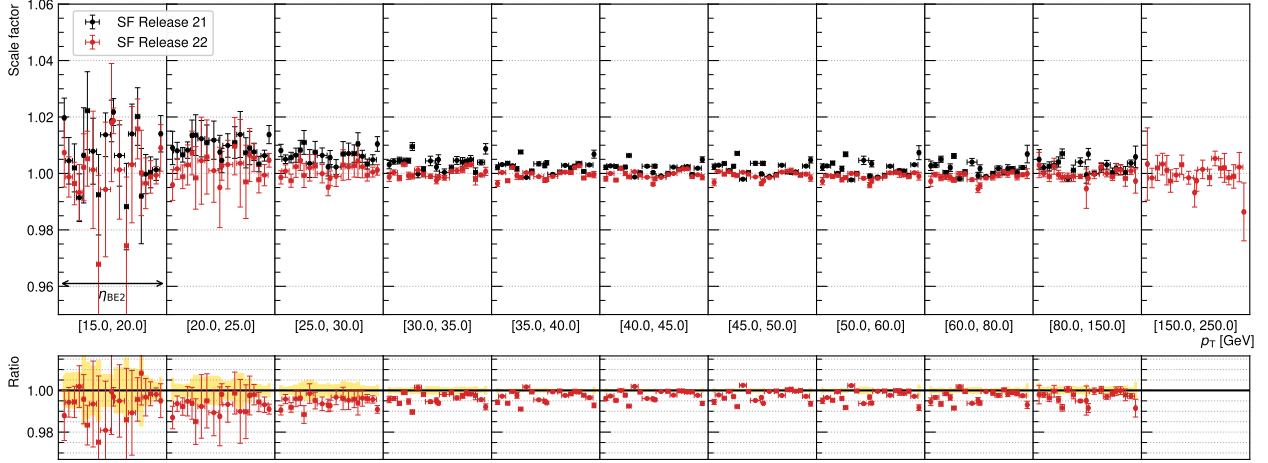


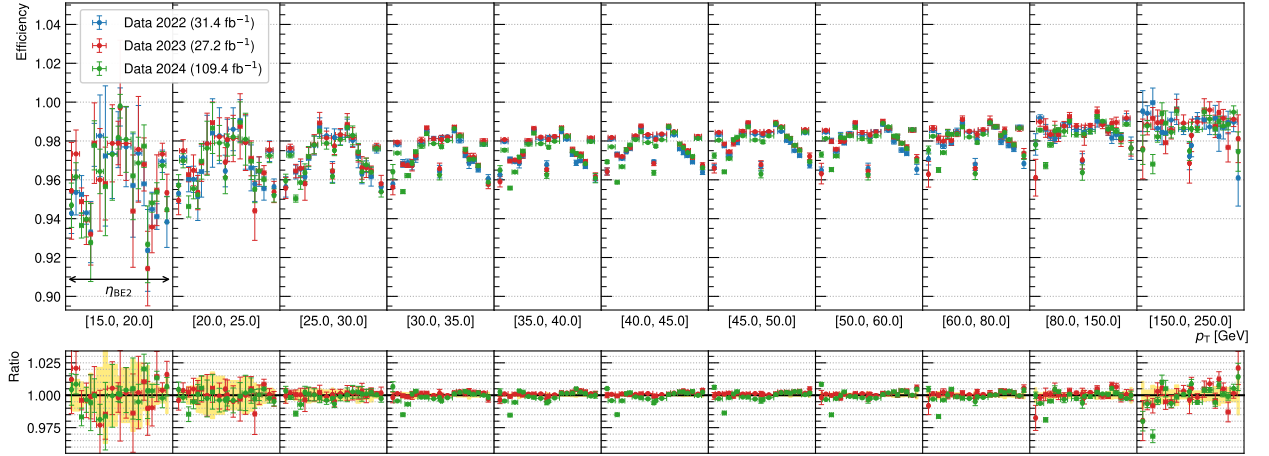
Figure 4.9: Electron reconstruction scale factors as a function of p_T and η for 2018 with release 21 in black and release 22 results in red.

The results obtained here were provided by the e/gamma CP group as a recommendation to physics analyses using electrons with the software release 22 and can be accessed over ATLAS GroupData with other electron scale factors as `2025_Run2Rel22_Recommendation_v2`. The usage is described in the documentation of the e/gamma CP group [41].

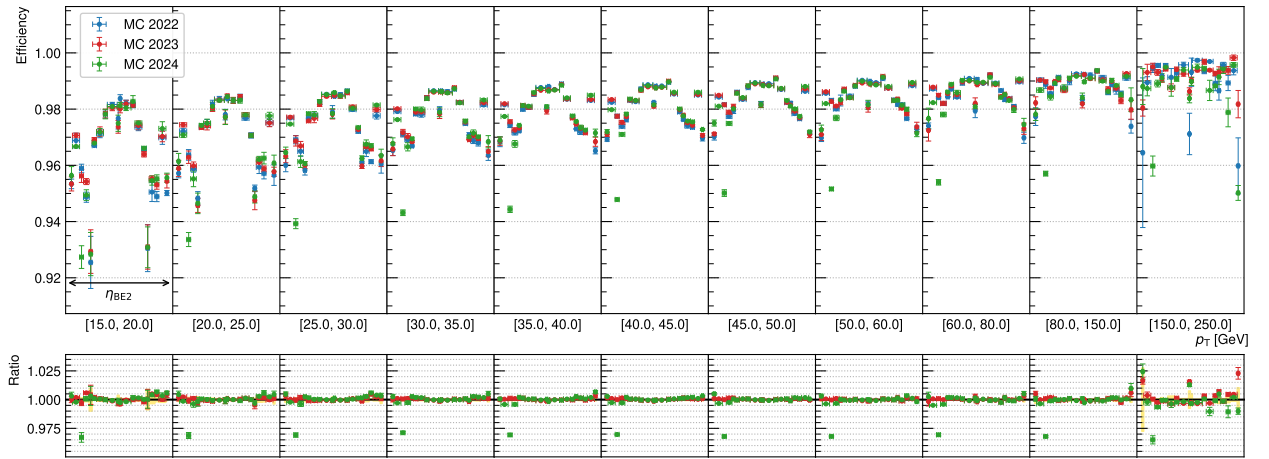
4.3.3 Run 3 Results

Comparing the efficiencies and scale factors in Run 2 (Fig. 4.8) with results from the first three years of Run 3 (Fig. 4.10) reveals some major differences. First of all, the scale factor baseline shifted slightly below unity in most bins. Excluding low p_T fluctuations, the data efficiency is decreased compared to simulation. Additionally, there are some p_T independent data efficiency reductions, which are not well described by MC, creating significant ($> 5\sigma$) deviations from a scale factor of one. There is an efficiency drop at $\eta \in [-0.1, 0.0]$ for all data-taking years of Run 3 of about 1.5%, which is present in MC with reduced magnitude ($\approx 0.75\%$). Exclusively in 2024, data and MC efficiencies drop sharply for $\eta \in [-2.01, -1.81]$, reaching a new minimal value in many p_T bins even though its neither the transition region nor the outermost η region. Since the scale factor deviations from unity are statistically significant, understanding and evaluating the newly appearing behavior in this region is crucial before they can be propagated to analyses. Except for the new efficiency decrease, the first three years of Run 3 are well compatible with each other. In the following, the decrease of data efficiency in Run 3 compared to Run 2 is investigated more closely.

The obtained scale factor results corresponding to 2022 and 2023 are included in the current set of consolidated Run 3 pre-recommendations [41].



(a) Data efficiency



(b) FullSim MC efficiency



(c) Scale factor

Figure 4.10: Electron reconstruction efficiencies and scale factors as a function of p_T and η in Run 3. The 2022 results (blue) are compared with the results obtained in 2023 (red) and 2024 (green), including their ratio with respect to 2022.

Investigation of Efficiency Decreases

To depict the origin of the efficiency decreases, which are sharply confined in η , analyzing the efficiency in ϕ - η instead of p_T - η is useful. This makes use of the observation that the appearance of the efficiency decrease is independent of p_T . Plotting the efficiencies in ϕ - η creates 2D maps of the detector, shown in Fig. 4.11. All p_T bins are merged and ϕ is binned in 20 equal width bins to ensure that enough statistics is contained in every bin.

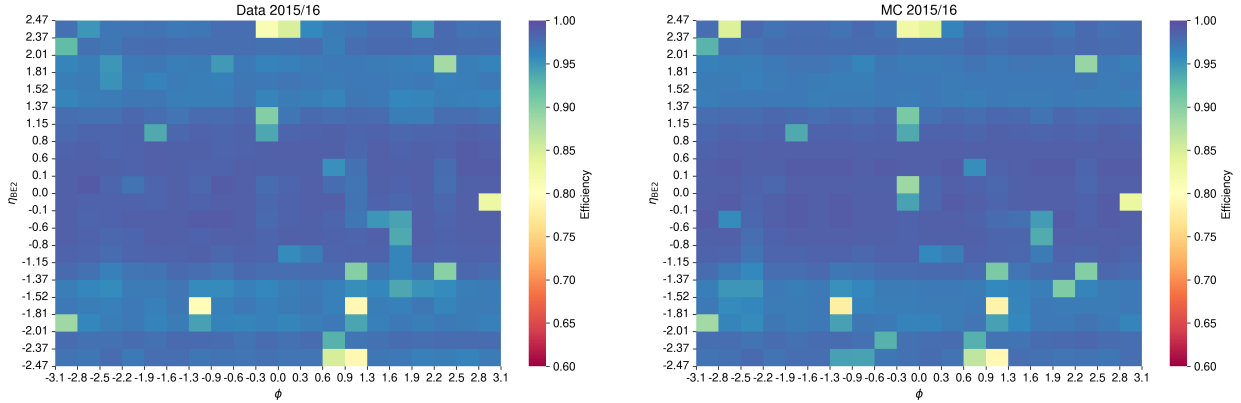


Figure 4.11: Electron reconstruction efficiencies in 2015/16 as a function of η and ϕ for data (left) and MC20 simulation (right).

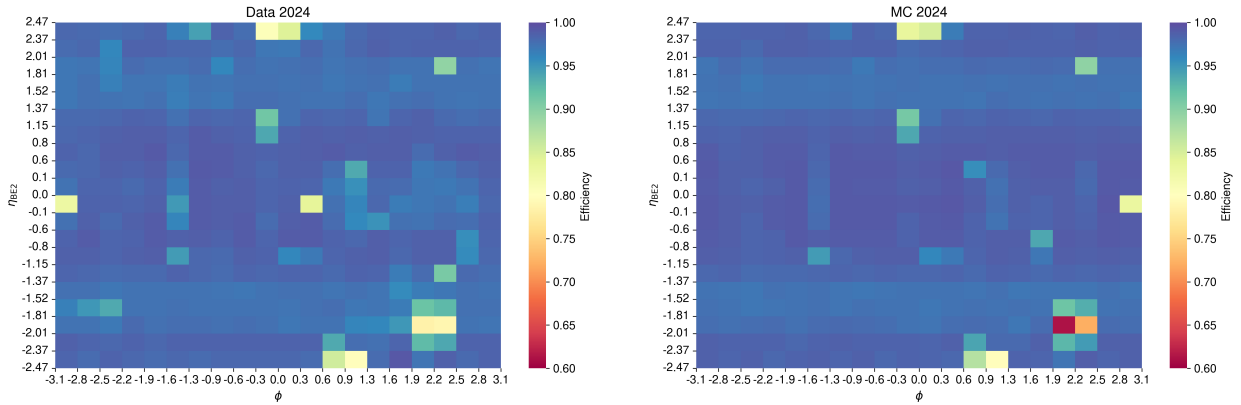


Figure 4.12: Electron reconstruction efficiencies in 2024 as a function of η and ϕ for data(left) and MC23 simulation (right).

The heatmaps show two distinct features. First, there is an underlying efficiency loss due to failed track cluster matching, peaking around the transition region. This part is homogeneous in ϕ , since the detector is, per design, symmetrical in this direction, as should all underlying physics processes. Superimposed are discrete efficiency decreases in confined η - ϕ regions of high magnitude. By varying cuts, it could be shown that for all relevant discrete drops ($\Delta\epsilon > 2\%$) the object quality setting is responsible. Besides track quality cuts with no visible impact, a cut called `BADCLUSELECTRON` is applied in the efficiency numerator only. It removes electron clusters in the following categories:

- MaskedCellCore - cell masked in the core of the cluster
- MissingFEBCell - missing front end board in core or edge of the cluster
- DeadHV - dead high voltage board in EM strips, middle or back affecting the cluster

Specifically, the cut on masked cells causes all visible reconstruction efficiency reductions. Masked cells are calorimeter cells intentionally excluded from further consideration, e.g. due to detector defects or otherwise impaired calorimeter spots. Plotting electron reconstruction efficiencies therefore produces a coarse map of detector defects and dead spots.

Fig. 4.12 shows numerous efficiency decreases of varying size and magnitude in 2024, similarly in Fig. 4.11 for 2015/16. The Run 2 MC20 samples were produced after these defects were identified, therefore, affected cells were precisely masked in MC as well, creating similar efficiency decreases and bringing the efficiency scale value close to unity. This is the case for the specific efficiency decrease close to the transition region in 2015/16 data, as mentioned in section 4.3.2, where only two ϕ bins dominate the efficiency decrease equally in data and MC.

For Run 3 this is more complicated, since new detector defects could not yet be included in the currently used MC23 campaign. Here, the dead spots observed during Run 2 were simply copied and implemented as masked cells to all Run 3 samples. The electron reconstruction scale factors thus reflect that the MC23 production does not yet contain all detector defects observed during Run 3. This can be seen in Fig. 4.12, where several regions with reduced data efficiency are not matched by MC, which creates deviations from unity in the scale value. One example is the $\eta \in [-0.1, 0.0]$ bin, where two confined ϕ regions dominate the loss of efficiency for the whole η bin for all Run 3 data-taking years. While the decrease close to $\phi = \pm\pi$ can be matched to MC, the reduction at $\phi \approx 0.3$ can not. The fact that masked cells seem to be located at $\phi = -\pi$ in data and at $\phi = +\pi$ in MC could indicate some alignment issues in ϕ .

For the 2024 exclusive efficiency decrease close to $\eta = -1.9$, a specific responsible defect could be identified. In August 2024, a front-end board in the electromagnetic calorimeter failed, effectively cutting all signals from 128 EM calorimeter cells. Since replacement of this board was only possible during the end-of-year shut-down, this created a significant dead region for the rest of 2024. The affected range is $\eta \in [-1.99, -1.81]$, $\phi \in [1.98, 2.35]$, which nicely matches the observed decrease. The binning does not cover the defect perfectly, however, spreading the impact over multiple bins in η and ϕ .

Initially, MC efficiencies did not show an effect in this region, but due to the magnitude of the defect, it was decided to mask all affected cells already for existing MC samples. Analyzing now the reconstruction efficiency, there seems to be an over-correction of the efficiency loss. In data, the efficiency decreases to 78 %, while for MC, it reaches 61 %, cutting more electron

candidates then necessary, resulting in scale factors significantly larger than one.

Therefore, the reconstruction scale factors are not only needed to correct for differing reconstruction success rates. They also play a role in transferring the impact of detector defects to simulation. While the impact of this is negligible in Run 2, where MC campaigns are already well adjusted to data, in Run 3 this currently has a very significant impact on the total correction by the scale factor. Future MC campaigns are expected to incorporate the dead spots, bringing the scale factors closer to unity again.

Electron Reconstruction Efficiency as a Function of Pileup

Another interesting quantity the reconstruction efficiency depends on is pileup. As mentioned in section 2.3, the average number of interactions per bunch-crossing varies each year of data-taking, generally increasing from year to year. While this is related to an increasing luminosity available for physics analyses, it also creates a more challenging environment for electron reconstruction.

Figure 4.13 shows the reconstruction efficiency as a function of the number of primary vertices in an event, which is closely related to the level of pileup. The efficiency steadily rises for increased pileup. Physically, this behavior can be explained by matching wrong tracks to electron clusters. An electron cluster, which otherwise would not have a successfully matched track, can be matched to tracks from pileup objects. While this effect causes the measured increase in reconstruction efficiency, this is not a desired behavior. Matching tracks and clusters originating from different particles creates an electron candidate with not well defined properties.

In release 22 and Run 2, this trend is well described by Monte Carlo for all years, resulting in only minor impacts on scale factors. Due to the previously explained decrease of the efficiencies for 2015/16, here also a slight negative offset compared to 2017 and 2018 can be observed.

In Run 3, especially in 2024, the pileup increased significantly. The efficiencies are presented in Fig. 4.14. While the MC efficiencies show the same trend as in Run 2, the data results do not follow the expected increase. 2024 data even shows a slight decrease for more than 20 primary vertices. The exact physical reason for this is not yet fully understood. Possible explanations include changes in reconstruction algorithms or detector performance, resulting in improved rejection of pileup tracks, especially in 2024. This behavior is not matched by MC and contributes to a global decrease of p_T - η dependent scale factors below unity in Run 3, as discussed above.

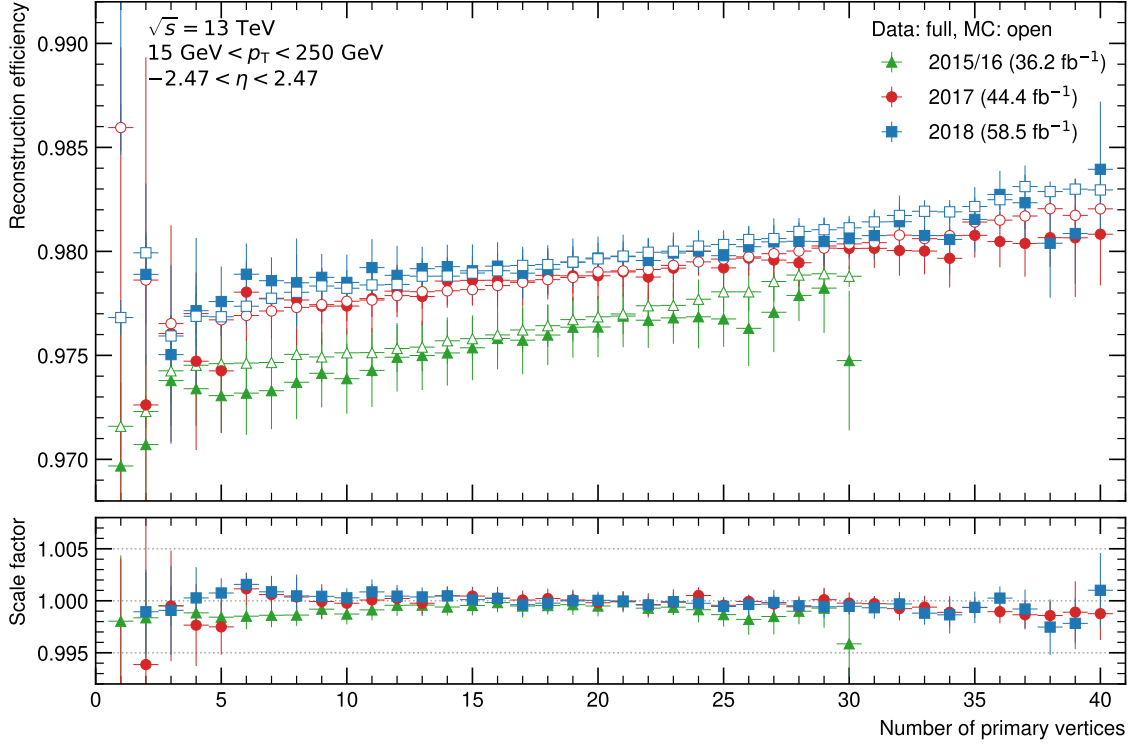


Figure 4.13: Electron reconstruction efficiencies and scale factors as a function of the number of reconstructed primary interaction vertices for Run 2. Data results correspond to filled dots and MC results correspond to open dots.

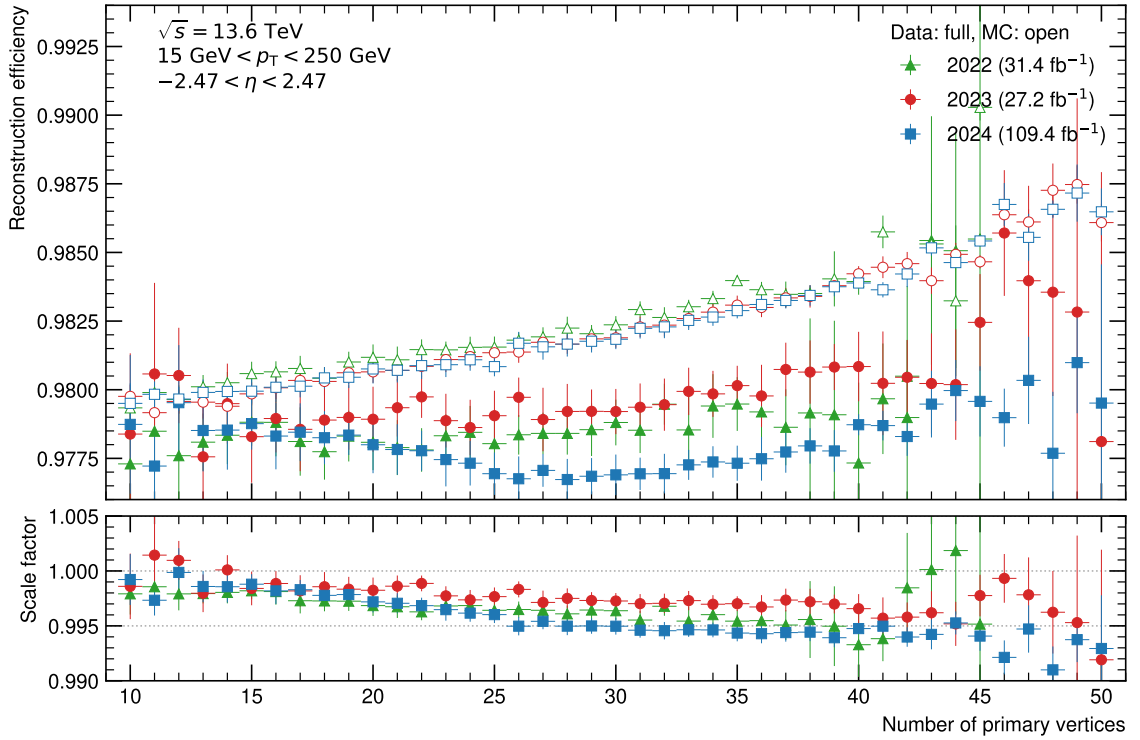


Figure 4.14: Electron reconstruction efficiencies and scale factors as a function of the number of reconstructed primary interaction vertices for Run 3. Data results correspond to filled dots and MC results correspond to open dots.

4.3.4 Fast Simulation Efficiencies and Scale Factors

As motivated in section 2.4, fast simulation can be an efficient approach to reduce compute load for increasingly big simulation samples. But fast simulation is only an approximation to the full GEANT4 simulation. Therefore, corrections can become necessary. Since the performance of the reconstruction algorithms might differ, MC-to-MC scale factors $SF_{AF3toFS}$ can be defined and measured in an analogous way. They are then multiplied with the FullSim-to-data scale factors $SF_{FStoData}$ to allow comparison between data and fast simulation samples:

$$SF_{AF3toData} = SF_{AF3toFS} \times SF_{FStoData} = \frac{\varepsilon_{FS}}{\varepsilon_{AF3}} \times \frac{\varepsilon_{Data}}{\varepsilon_{FS}}. \quad (4.10)$$

Full GEANT4 simulation (FS) efficiencies are compared to the current iteration of fast simulation, ATLFast3 (AF3). The same methodology and p_T - η binnings are used as in the FS scale factor measurement. Currently no systematics are run for the AF3 to FullSim correction, only the systematic uncertainty from $SF_{FStoData}$ and the statistical uncertainty from both SF measurements are included in the recommendations. There are two main reasons for this: 1) most current systematics refer to background estimation, which is obsolete for MC-only measurements using truth matching and 2) the remaining systematics would be heavily correlated with the $SF_{FStoData}$ systematics since the methodology is the same.

It is also important to note that for electrons, calibrated energies are used in the measurement, including fast simulation specific corrections in scale α and resolution c of reconstructed energies. The following energy calibration models were used: `es2023_R22_Run2_v1` for Run 2 and `es2024_Run3_v0` respectively for Run 3, see also [33] for more information. Especially in Run 3, these models should be considered as work in progress and may be subject to some relevant changes in the future. Especially energy scale corrections move the Z peak position in invariant mass distributions, impacting the efficiency measurement.

While the measurement and propagated results are done separately for each year, except the combined years 2015 and 2016, reconstruction efficiencies for merged samples of Run 2 and Run 3 are shown here. This is useful since qualitatively no significant changes can be observed within a run, as can be expected for samples of the same MC campaign. For a complete list of efficiencies, see appendix A.3.3.

Figure 4.15 shows a comparison of reconstruction efficiencies between full and fast simulation in Run 2. Their ratio corresponds to $SF_{AF3toFS}$. Generally, an exceptional agreement resulting in scale factors close to unity (deviation of less than 0.005) can be observed, AF3 approximates full

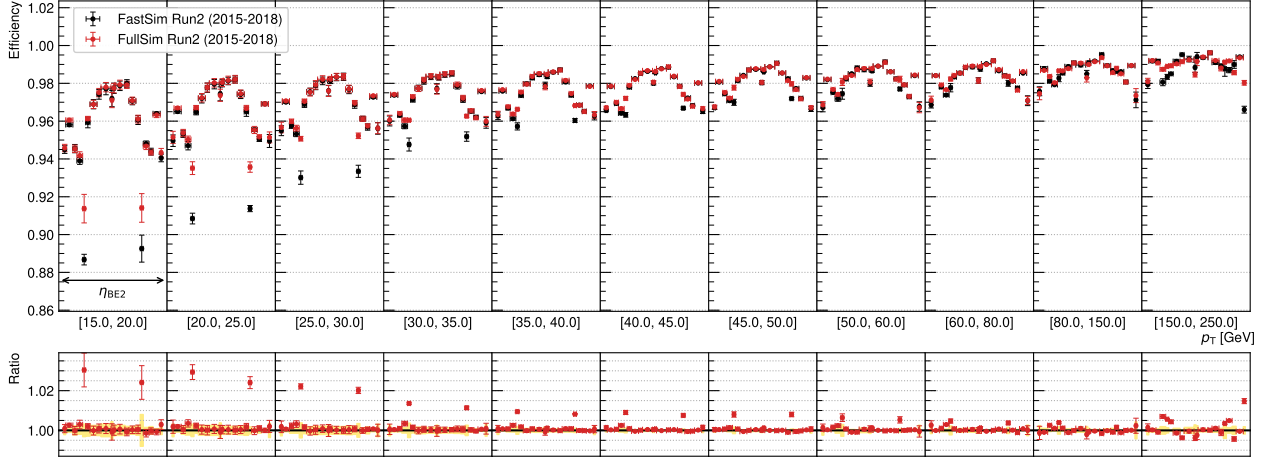


Figure 4.15: Electron reconstruction efficiencies in Run 2 as a function of p_T and η for full simulation (red) and AF3 fast simulation (black) and their ratio, which corresponds to a MC-to-MC scale factor.

GEANT4 simulations well in the context of electron reconstruction. An exception to that are the bins corresponding to the transition region, $1.37 < |\eta| < 1.53$ between barrel and endcap, where the efficiency decrease is enhanced for AF3 results consistently. Like in full simulation, this efficiency reduction corresponds to not successfully matched tracks and clusters, e.g. due to bremsstrahlung, which is mainly relevant for low p_T . A likely cause for the discrepancy is mismodeling of the transition region in AF3 itself, since similar disagreements were also recently seen in energy calibration.

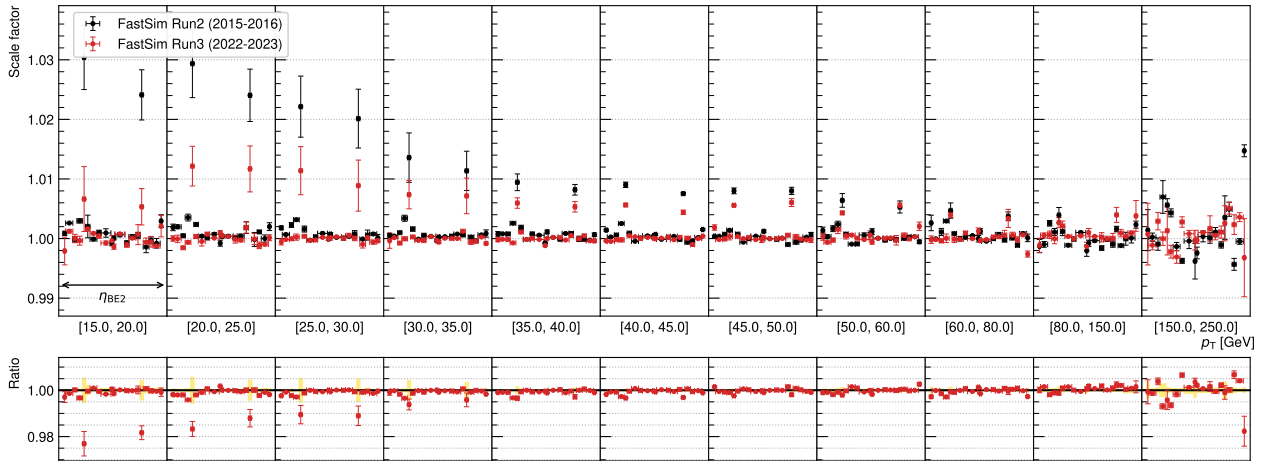


Figure 4.16: AF3 to FullSim electron reconstruction efficiency correction (scale factor) as a function of p_T and η for the first two years of Run 3 (red) and the full Run 2 (black) and their ratio.

In Run 3, the same close agreement in all regions except the transition region can be observed, but the maximal discrepancy shrinks from up to 3% to below 1.5%, showing some improvement. This can be seen in the scale factors in Fig. 4.16. Here, only the first two years of Run 3 contribute, since no 2024 AF3 sample was available. But for samples of the same MC

campaign (MC23), no major differences are expected. Since in both runs the discrepancy is sharply confined to the transition region η bin, this will not have an impact on most physics analyses, since they usually exclude this region. The scale factors allow some corrections to AF3 in terms of reconstruction efficiencies, but the likely presence of mismodeling further discourages the usage of this region with AF3.

Also interesting is a comparison of scale factors for different software releases in Run 2. While release 21 samples use AtlFast2 (AF2), reprocessed release 22 samples used in this work rely on ATLFast3. Compared to AF2, AF3 includes several improvements with the goal of resembling full GEANT4 better, while still being less computationally intensive. The parametrization based shower generation approach has been improved and neural networks have been introduced to predict showers in the calorimeters, see also [28].

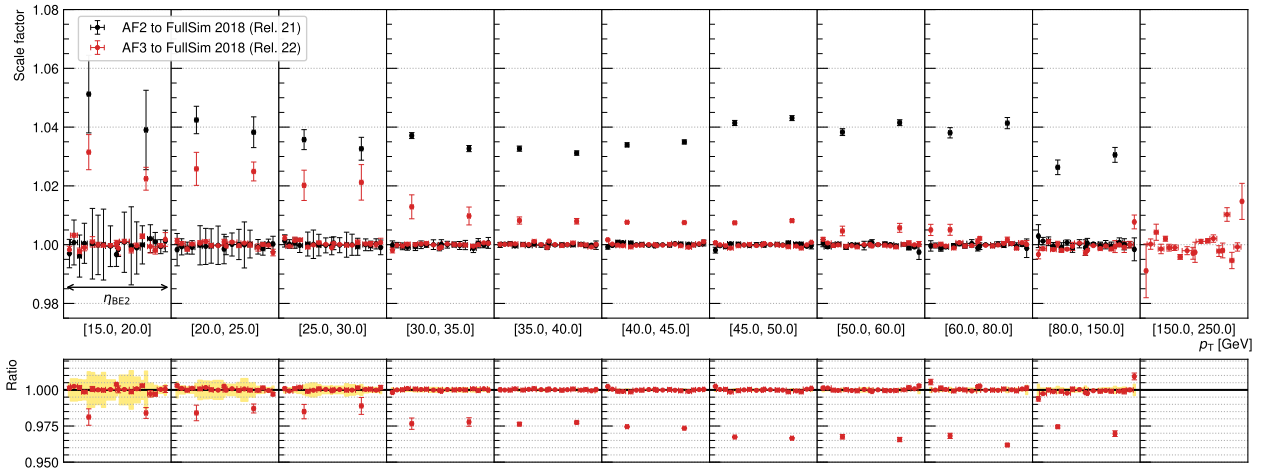


Figure 4.17: FastSim to FullSim electron reconstruction efficiency correction (scale factor) as a function of p_T and η in 2018 for AF3 (red) and AF2 (black) as well as their ratio.

The new AF3 scale factor correction $SF_{AF3toFS}$ is compared with the AF2 correction obtained from the latest release 21 recommendations in Fig. 4.17 for 2018. Like for the full simulation scale factors, the new measurement presented in this work adds a new p_T bin up to 250 GeV, which shows mainly statistical fluctuations. The uncertainties cannot be compared with each other, since the release 21 measurements were carried out with separate systematics and using the old systematic setup described in section 4.2.2.

Comparing the scale factors, it becomes clear, that the issue in the transition region was already present for AtlFast2, showing even higher discrepancies to full simulation of up to 5.5 %, which, in contrast to the AF3 results, remain relevant for high p_T . Aside from the transition region, similar agreement between fast and full simulation can be observed. In conclusion, fast simulation is generally an excellent approximation to full simulation in terms of electron reconstruction efficiencies, except in the transition region. Here, apparent mismodeling causes

discrepancies, which were reduced step by step from AF2 to AF3 in Run 2 and from Run 2 to Run 3. The Run 2 results obtained in this work were included in the last set of electron scale factor recommendations by the e/gamma CP group [41].

4.3.5 Improvement of photonic Background Estimation

As shown in section 4.2.2, the variations targeting the photonic background fit constitute the dominant source of systematic uncertainty in many p_T - η bins, especially at low p_T . Therefore, improving the estimation of this background and finding robust systematics to better estimate its uncertainty is crucial for improving the measurement. The goal is to find a good estimation for clusters without tracks, which are not originating from electrons with failed track matching. Since photon candidates are used, these backgrounds are commonly referred to as photons, but there is no recent study on the background composition and the fraction of fake photons.

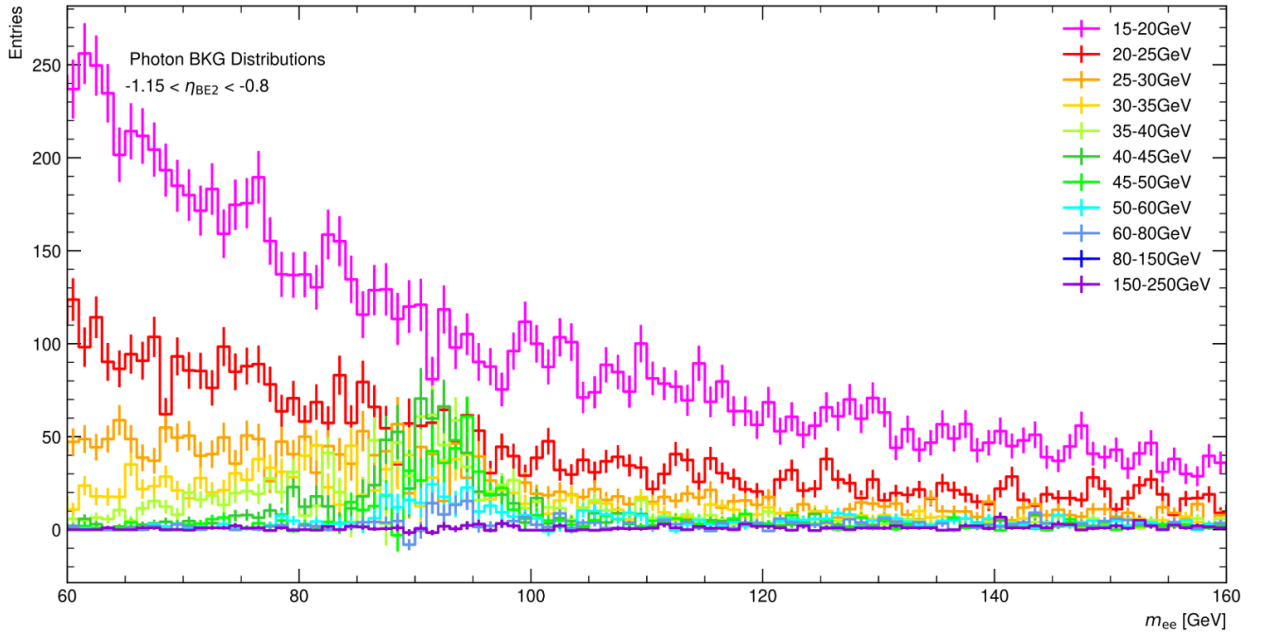


Figure 4.18: Invariant mass distribution for photon candidate backgrounds after signal contamination removal. For $-1.15 < \eta < -0.8$ distributions from all p_T bins are included.

Figure 4.18 shows the photon candidate backgrounds after removal of the $Z \rightarrow ee$ signal contamination in a specific η region. It highlights three major challenges for the background fit: 1) the shape of the invariant mass distribution varies depending on the kinematic bin, 2) the amount of events to fit to is drastically decreased for higher p_T and 3) in some bins a clear Z peak is visible, indicating that the MC-based signal contamination removal does not work as expected. A proposed update to the fitting procedure is introduced in this section.

Determination of new Fitting Setup

Compared to the old fitting setup, several changes were made. The first challenge is the varying number of background events considered in the fit, which depends heavily on the kinematic bin. To ensure enough events per bin, the width of the m_{ee} bins is determined dynamically based on the total number of events N in the fit region. Starting with a bin width of 1 GeV for $N > 2000$, the bins are merged successively for lower statistics up to a width of 5 GeV for $N < 30$. This stands in contrast to global 5 GeV bin widths in the old setup. The fit thus uses more information from a more detailed sampling of the spectrum.

A simple third-order polynomial fit has been used previously. In the new approach, more complex Bernstein polynomials are deployed. They have been found to model the background shape well, while also always being positive per definition. They are linear combinations of Bernstein base polynomials, which are defined as:

$$b_{i,n} = \binom{n}{i} x^i (1-x)^{n-i}. \quad (4.11)$$

For a given order n , the overall Bernstein polynomial is then given by:

$$B_n = \sum_{i=0}^n \alpha_i b_{i,n} \quad (4.12)$$

with coefficients α_i determined in the fit.

To determine the optimal polynomial order, the χ^2 divided by the number of degrees of freedom (d.o.f.) was taken into account, for each p_T - η bin. Figure 4.19 shows $\chi^2/\text{d.o.f.}$ for the orders 3, 4 and 5. While 3rd-order Bernstein polynomials often do not catch the complexity of the spectrum, 4th-order parametrization matches the data well. Increasing the order to 5 does not improve the fit considerably, therefore 4th-order Bernstein polynomials were chosen. Figure 4.19 also includes $\chi^2/\text{d.o.f.}$ for the old setup based on fixed bin widths and a simpler 3rd-order polynomial. Here a significant improvement can be seen for the new setup, indicating more accurate background estimation. An exemplary resulting invariant mass distribution is shown in Fig. 4.20, in comparison with the previously used fitting setup.

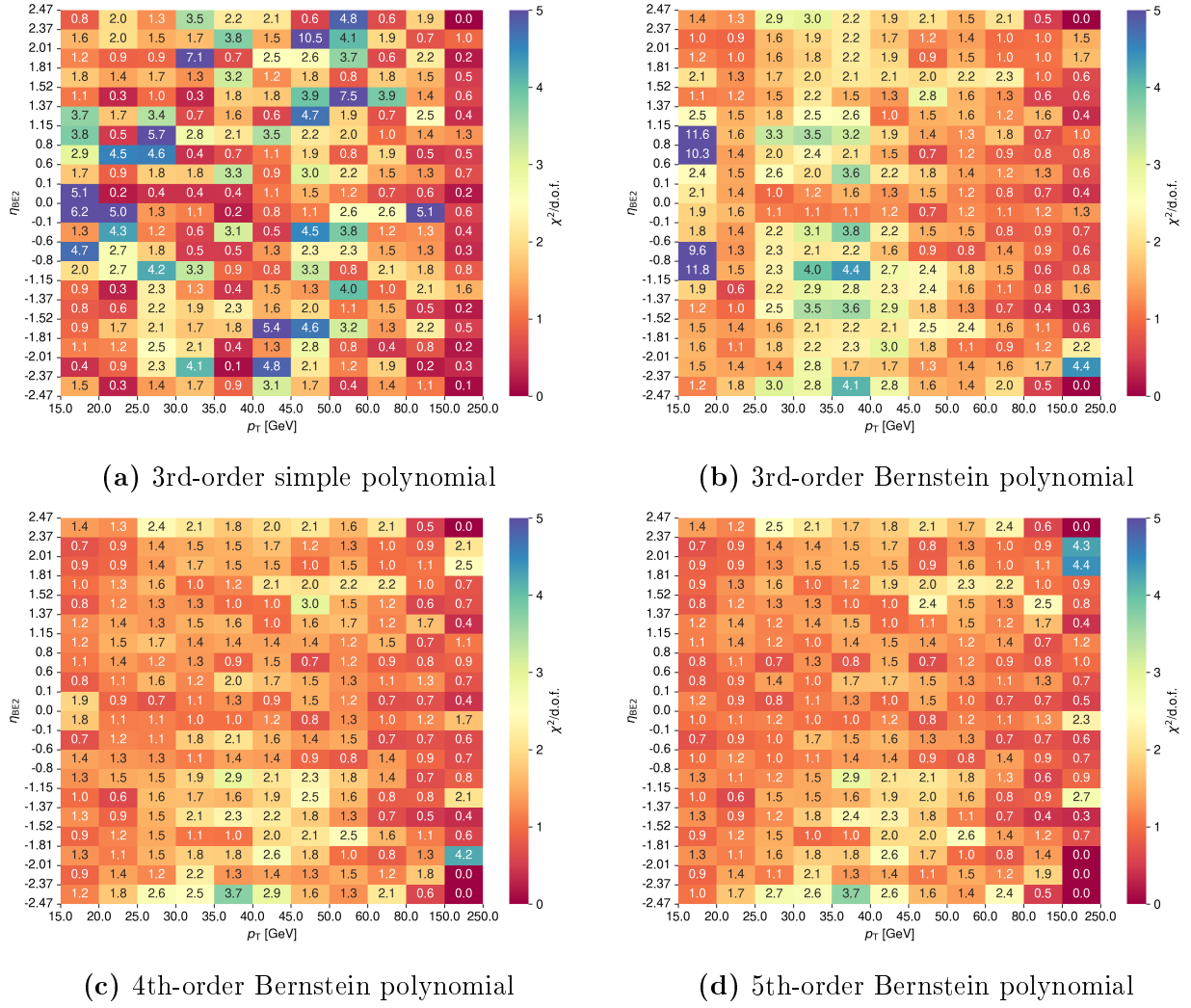


Figure 4.19: χ^2 divided by the number of degrees of freedom for the photon background fit in each p_T bin using 2018 data. Different polynomial orders are compared with the simple third-order polynomial fit of the old setup.

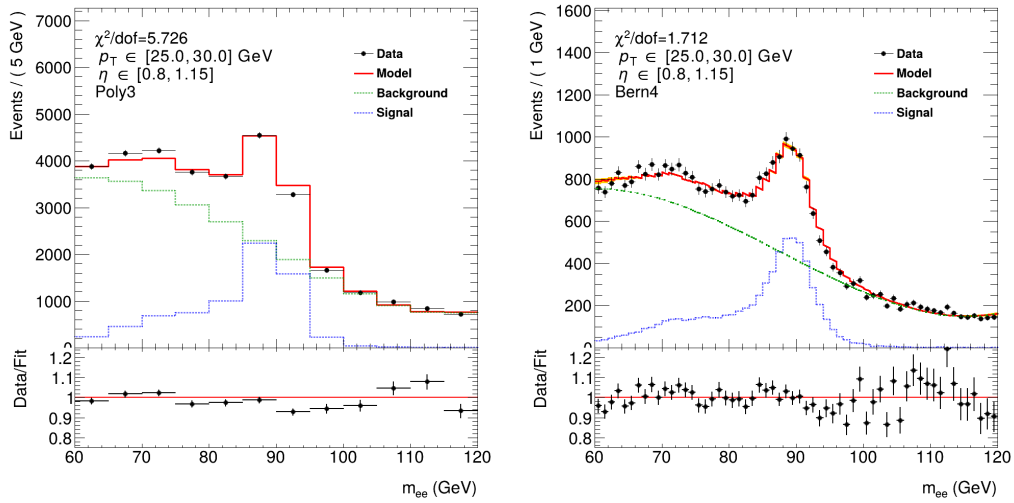


Figure 4.20: Invariant mass distribution with signal and fitted background compared to data, a simple 3rd-order polynomial fit in the old setup (left) is compared with a 4th-order Bernstein polynomial fit in the new setup (right).

Signal Rescaling

To perform a background-only fit, the signal is removed from the data using $Z \rightarrow ee$ MC simulation. To normalize MC to data, the same scale value as for electron probes is used: $s = N_{\text{data,tight}}/N_{\text{MC,tight}}$ using electron probes passing tight ID with an di-electron invariant mass close to the Z peak. This scale value is calculated separately per p_T - η bin. Applying it for photon probes as well assumes similar reconstruction efficiencies in data and MC, which is only approximately true, see the results presented earlier in section 4.3.2.

Therefore, signal MC is now rescaled specifically for photon candidates to improve the background fit. Lacking effective cuts to isolate the signal in data, this is done in the same fit as the background fit. The polynomial background model and the simulated signal with dynamic normalization are fitted simultaneously to data in the whole mass range without sidebands. From this first fit, the new MC scale value is extracted, as well as its uncertainty. Figure 4.21 shows the correction factor to the initial scale value in 2018. Most p_T - η bins receive only small corrections ($< 5\%$). This procedure has a bigger impact on certain bins, namely $0.6 < |\eta| < 0.8$ in the first p_T bin, as well as $-0.1 < \eta < 0.0$ across the p_T range and fluctuations in the highest p_T bin. This coincides nicely with the 2018 reconstruction scale factors presented in section 4.3.2, already hinting at data/MC differences in these regions.

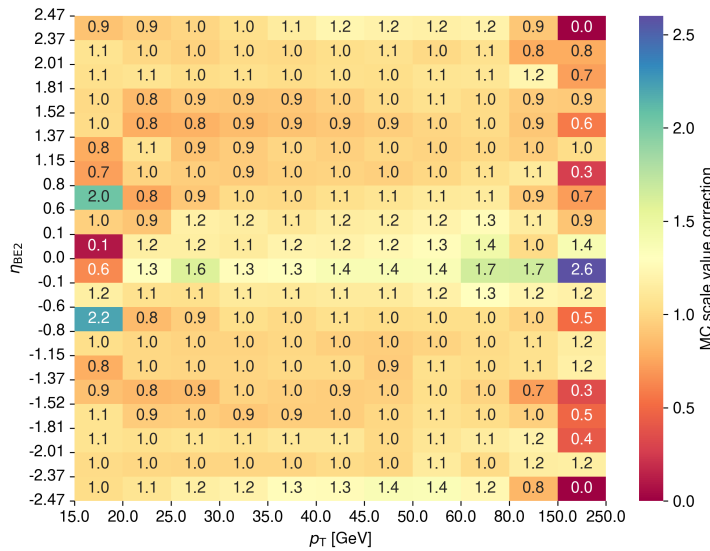


Figure 4.21: Correction factor to the MC-to-data scale value used for signal contamination removal for all p_T - η bins in 2018.

This procedure is illustrated in the Figures 4.22 and 4.23. Here, the left side corresponds to the old setup with a fixed MC scale value, which is used for signal contamination removal. The remaining signal in proximity of the Z peak complicates the background-only fit significantly. By rescaling the signal subtracted from data for the background-only fit (Fig. 4.23), a background sample of higher purity is achieved.

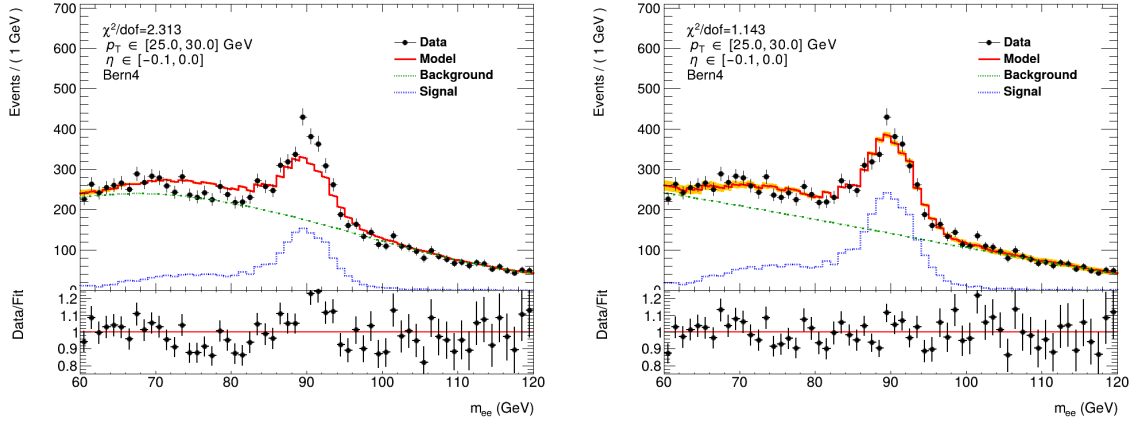


Figure 4.22: Invariant mass distribution for probes without a track, with MC signal and fitted background compared to data. Left: signal is scaled by a fixed scale value. Right: the MC signal normalization is determined by the fit.

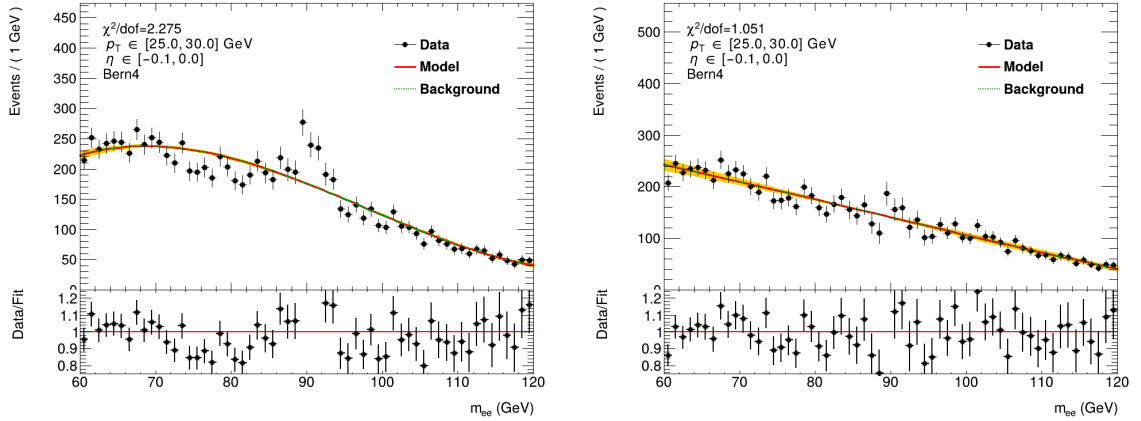


Figure 4.23: Invariant mass distribution for probes without a track. The background is fitted to data after signal contamination removal. Left: signal removed using fixed MC scale value (old setup). Right: rescaled MC signal removed (new setup).

In this way, the fitting of only sidebands to the mass peak can be avoided, which originally was motivated by significant disagreements in the signal region. To estimate the uncertainty of this newly added procedure, the second (background-only) fit is performed based on varying data spectra after signal contamination removal. Using the scale value obtained from the first fit and varying the fit parameters within their uncertainties defines an up and a down variation, allowing the estimation of systematic uncertainty for MC rescaling. The impact of these variations is shown in Fig. 4.24. The up and down variation produce symmetric deviations from the nominal setup, 0.1–0.8% in the first p_T bins and mostly well below 0.1% for higher p_T . While not being the most impactful variation, this is still a sizable effect and of the same magnitude as varying the signal mass range. That highlights the impact of a proper removal of $Z \rightarrow ee$ signal contamination.

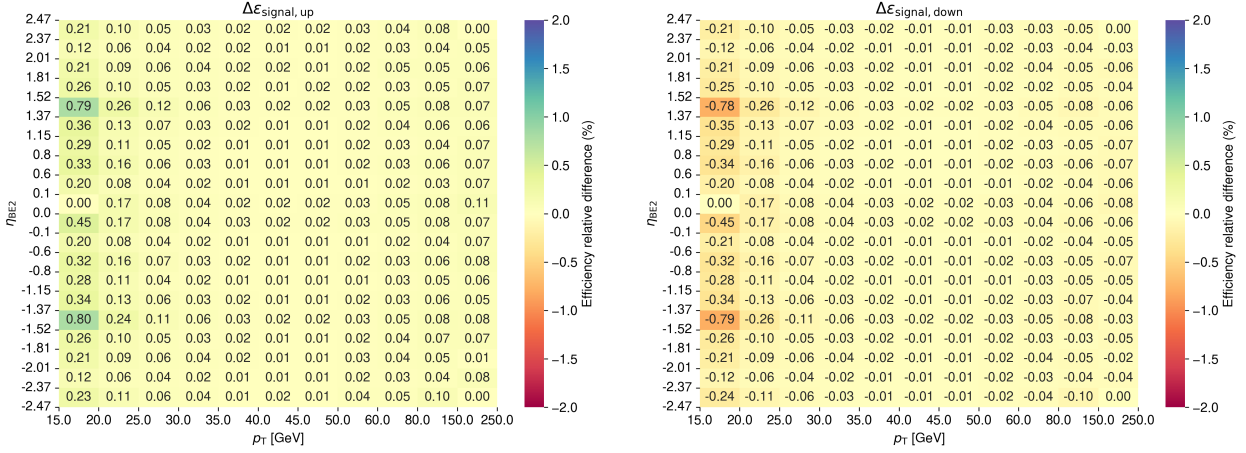


Figure 4.24: Efficiency difference to nominal variation for signal rescaling variations up (left) and down (right) in 2018 data.

First Results

The initial development and evaluation of this new procedure was performed on the 2018 dataset, since its results are well understood. In Figure 4.25, the efficiency results are compared with the old setup, based on a third order polynomial fit. The change in efficiency does not exceed 1% and is in every p_T - η bin within the statistical uncertainty. Additionally, no general trend can be observed. This makes the transition towards the new setup easier, since the efficiencies still show the same well understood behavior.

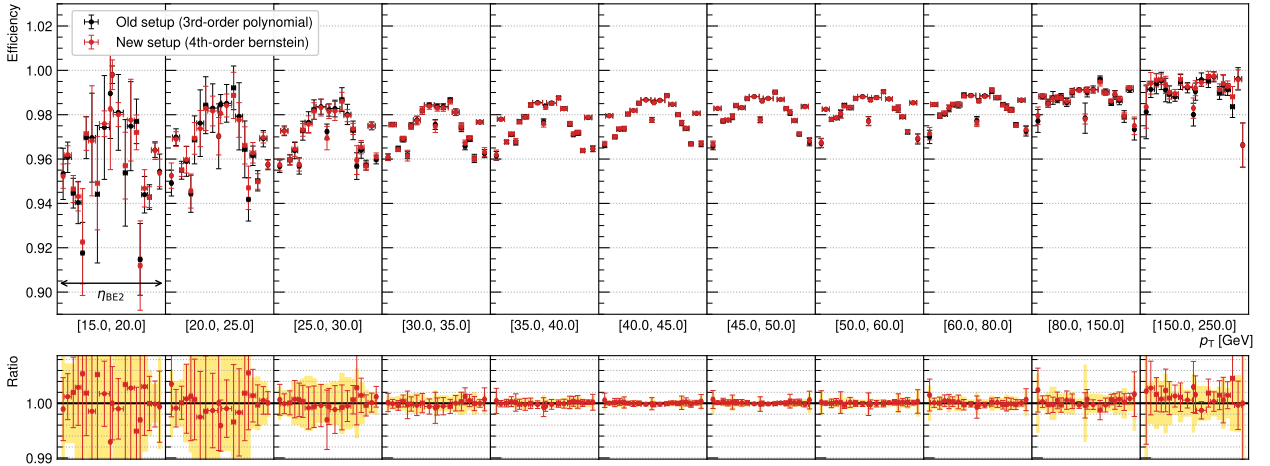


Figure 4.25: Electron reconstruction efficiency as a function of η and p_T for 2018 data. Results from the previously used setup are compared with the new approach.

A major objective is to include these changes in the next round of scale factor recommendations in Run 3. Therefore, it is also important that stable results are produced throughout all years. The efficiencies for the first three years of Run 3 are included in appendix A.3.4. A similar behavior as in 2018 is observed, with the changes mostly being covered by the statistical

uncertainty. The total uncertainty in 2018 for the new setup, as well as the difference in uncertainty compared to the old setup, is shown in Fig. 4.26.

The total uncertainty remains at the permille-level for $25 < p_T < 150 \text{ GeV}$. Since a fit is involved both in the new and the old setup, the change of uncertainty varies heavily between different p_T - η bins. In 80 out of 220 kinematic bins, the uncertainty increases, mostly in the central $30 < p_T < 60 \text{ GeV}$ transverse momentum region, where the absolute uncertainty is at 0.002 or lower. In the majority of bins, a moderate improvement of the uncertainty can be seen. Therefore, while offering compatible efficiency results, only a small improvement in terms of total uncertainties is achieved by the current setup.

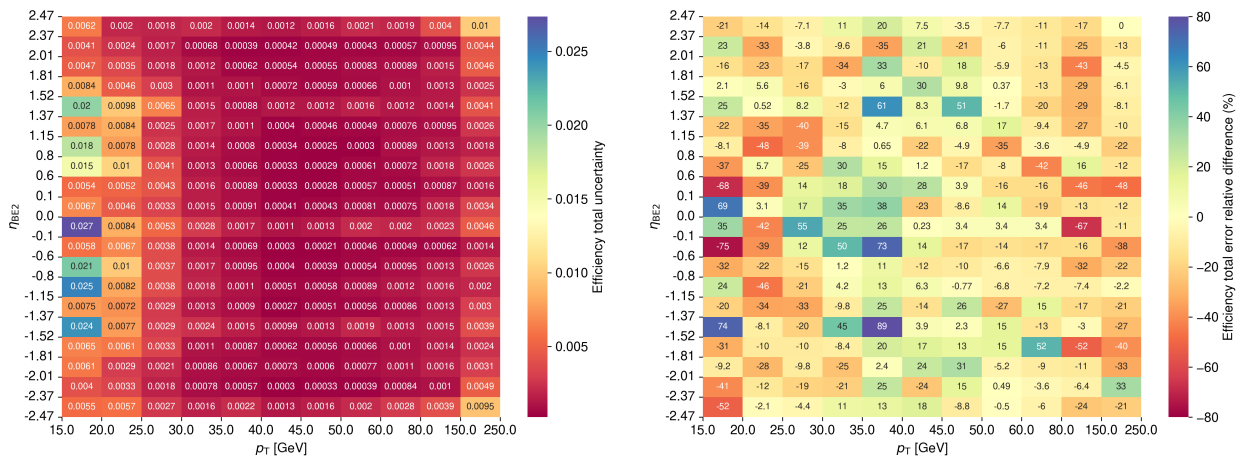


Figure 4.26: 2D p_T - η heatmap with total uncertainties of the 2018 reconstruction efficiency in data with the new fitting procedure (left), as well as the relative change of total uncertainty compared to the old setup (right).

Further tuning and evaluation is needed, before the new procedure can replace the well established setup. But the new fitting strategy demonstrated several conceptual advantages. A better agreement between data and modeled background is achieved and signal contamination removal is improved.

The variations currently employed primarily probe the numerical stability of the fit. In future iterations, it is planned to replace them by a set of variations designed to be less sensitive to statistical fluctuations, including the newly introduced systematic uncertainty related to signal contamination removal. In addition, exploring alternative fit functions or polynomial orders may help identify potential systematic deviations from the nominal setup.

4.4 Open Questions and Outlook

In preparation for the Run 3 recommendations, there are several possible promising improvements for the background estimation. Some of the changes introduced to the photonic backgrounds can be transferred to the electronic part. Especially determining the amount of MC signal via fitting and introducing a systematic on this rescaling proved to be a healthy procedure, improving the agreement with data significantly. Fitting the template to data in general could also decrease the reliance on correct signal simulation.

Furthermore, the background estimation currently does not use the sign difference between tag and probe electron charge. This was already briefly tested, not showing relevant changes in efficiency or its uncertainty. However, an improvement in the signal to background ratio of up to 30% could be observed, as can be seen in Fig. 4.27. A more thorough investigation is needed to determine if this leads to improved results. Additionally requiring same sign tag-probe pairs would produce a pure background sample containing, for example, fake electrons originating from jets.

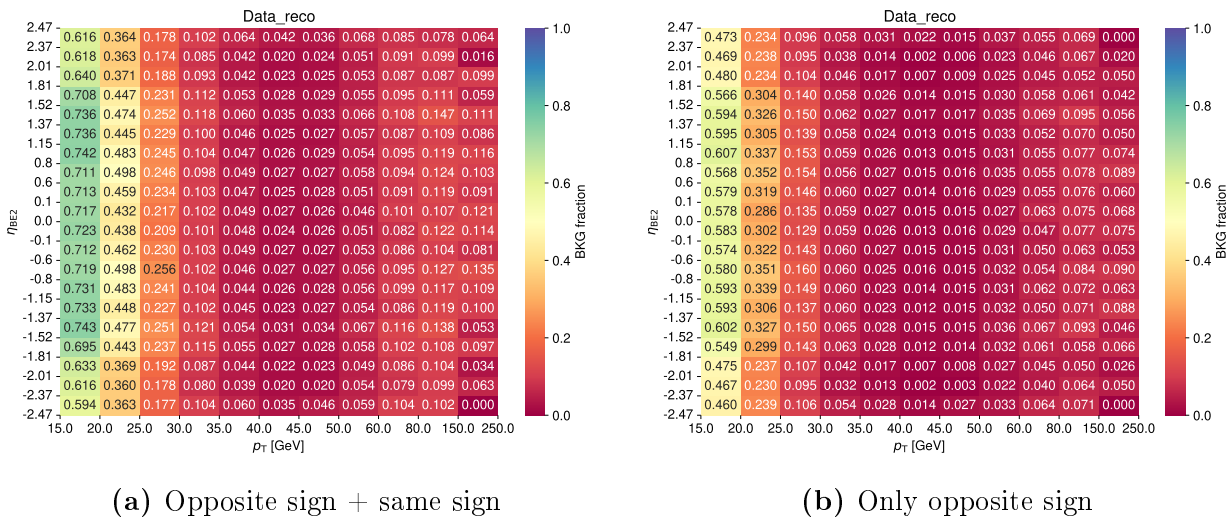


Figure 4.27: Estimated background events divided by total events for probes with a track in 2018. The current setup not using electron charge is compared with selecting only tag-and-probe pairs with opposite sign.

Additionally to these improvements, there are several not yet understood distributions, which could be further investigated. The unexpected decrease of efficiency with increasing pileup in 2024 was already mentioned in section 4.3.3. Also, the invariant mass distributions of MC signal and background do not match well to data in certain phase-space regions, especially for $|\eta|$ in the transition region and close to 2.47. Here, the tail of the z -peak has a different shape in data compared to Monte Carlo, hinting at not optimal energy calibration. In Fig. 4.28 this is shown exemplary for two bins in 2024. The other η regions show better agreement.

This issue already exists in Run 2, but is enhanced in Run 3. Mismodelling of the MC signal affects the background estimation, since simulation is used for signal contamination removal. A possible intermediate fix for this issue could be to introduce an a-posteriori energy scaling of the MC to improve the ratio of data to the signal and background modeling.

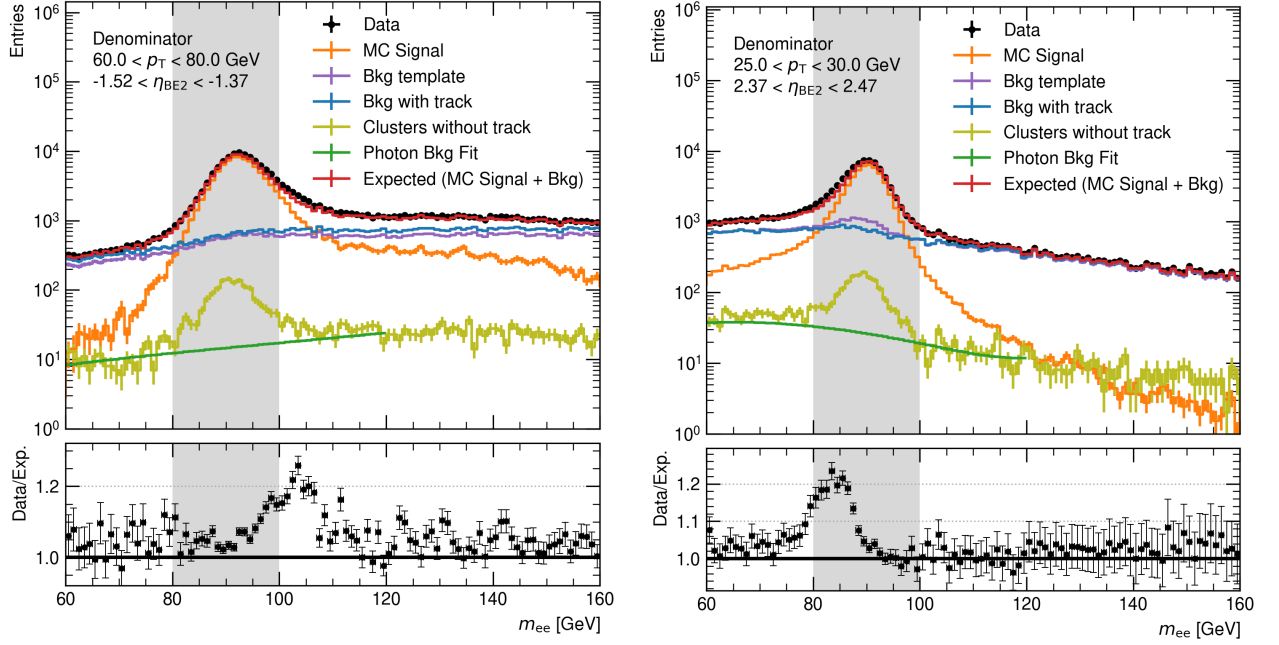


Figure 4.28: Tag-probe invariant mass distribution for all probes (denominator level), including MC signal, background contributions and total data.

5 Conclusion

Using the tag-and-probe method, the electron reconstruction efficiency was measured for the full Run 2 dataset and the first three years of Run 3. A new approach for systematic uncertainty estimation based on a variation of nominal tag-and-probe criteria was established.

The efficiency is generally high with approximately 98 % when integrating over the whole p_T - η range and is measured with percent to permille-level precision. The efficiency continuously decreases around the barrel-endcap transition region due to increased bremsstrahlung complicating track-cluster matching and also decreases at sharply confined space points due to dead spots in the detector. For Run 3, several not yet understood changes with respect to Run 2 were investigated. Most notably, the data efficiencies are heavily affected by detector defects, which are not yet included in current MC campaigns. Also in Run 3, the reconstruction efficiency does not show the expected increase with pileup.

From the measured efficiencies scale factors were derived, which are used to correct simulation to data. For full simulation, these were close to unity in Run 2, demonstrating good agreement between MC and data. For Run 3, more deviations were found, with the efficiency in data showing the tendency of being lower than in simulation. Additionally, fast simulation was compared with full GEANT4 simulation. Excellent agreement was observed, resulting in almost no corrections from the corresponding scale factors, except in the barrel-endcap transition region, where AF3 emulates significantly lower efficiencies than observed in data and full simulation.

The most relevant contribution to the uncertainty originates from the fit used to estimate backgrounds for clusters without a track. Several possible improvements to the fitting procedure were developed and evaluated, showing promising initial results. Especially the dynamic rescaling of the MC signal subtracted from data, improves the fit quality significantly. But further investigations are needed, before the new procedure can replace the current setup.

The scale factor results presented in this work are partially already distributed as recommendations to physics analyses using electron final-states with ATHENA software release 22 or newer in Run 2 and Run 3.

Bibliography

- [1] R. P. Feynman. Space-Time Approach to Non-Relativistic Quantum Mechanics. *Rev. Mod. Phys.*, 20:367–387, 1948.
- [2] J. Schwinger. On Quantum-Electrodynamics and the Magnetic Moment of the Electron. *Phys. Rev.*, 73:416–417, 1948.
- [3] S. L. Glashow. Partial-symmetries of weak interactions. *Nuclear Physics*, 22(4):579–588, 1961.
- [4] A. Salam. Weak and Electromagnetic Interactions. *Conf. Proc. C*, 680519:367–377, 1968.
- [5] S. Weinberg. Non-abelian gauge theories of the strong interactions. *Phys. Rev. Lett.*, 31:494–497, 1973.
- [6] H. Fritzsch, M. Gell-Mann, and H. Leutwyler. Advantages of the Color Octet Gluon Picture. *Phys. Lett. B*, 47:365–368, 1973.
- [7] A. Einstein. Die Grundlage der allgemeinen Relativitätstheorie. *Annalen Phys.*, 49(7):769–822, 1916.
- [8] M. Dine and A. Kusenko. Origin of the matter-antimatter asymmetry. *Rev. Mod. Phys.*, 76:1–30, 2003.
- [9] K. Garrett and G. Duda. Dark Matter: A Primer. *Adv. Astron.*, 2011:968283, 2011.
- [10] The ATLAS collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Physics Letters B*, 716(1):1–29, 2012.
- [11] E. Lopienska. The CERN accelerator complex, layout in 2022. 2022. General Photo.
- [12] G. Arduini *et al.* LHC Upgrades in preparation of Run 3. *Journal of Instrumentation*,

- 19(05):P05061, 2024.
- [13] L. Evans and P. Bryant. LHC Machine. *Journal of Instrumentation*, 3(08):S08001, 2008.
- [14] The ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *Journal of Instrumentation*, 3(08):S08003–S08003, 2008.
- [15] The ATLAS Collaboration. The ATLAS experiment at the CERN Large Hadron Collider: a description of the detector configuration for Run 3. *JINST*, 19(05):P05063, 2024.
- [16] The ATLAS Collaboration. The ATLAS Inner Detector commissioning and calibration. *The European Physical Journal C*, 70:787–821, 2010.
- [17] The ATLAS Collaboration. Alignment of the ATLAS Inner Detector in Run 2. *The European Physical Journal C*, 80:1194, 2020.
- [18] The ATLAS ITk Collaboration. The ATLAS ITk detector system for the Phase-II LHC upgrade. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 1045:167597, 2023.
- [19] A Artamonov *et al.* The ATLAS Forward Calorimeter. *Journal of Instrumentation*, 3(02):P02010, 2008.
- [20] I. Manthos *et al.* The micromegas project for the ATLAS new small wheel. *AIP Conference Proceedings*, 2075(1):080010, 2019.
- [21] The ATLAS Collaboration. LuminosityPublicResults . 2025.
- [22] S. Alioli *et al.* A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *JHEP*, 06:043, 2010.
- [23] Christian Bierlich *et al.* A comprehensive guide to the physics and usage of PYTHIA 8.3, 2022.
- [24] T. Gleisberg *et al.* Event generation with SHERPA 1.1. *Journal of High Energy Physics*, 2009(02):007, 2009.
- [25] The GEANT4 Collaboration. GEANT4 - A Simulation Toolkit. *Nucl. Instrum. Meth. A*, 506:250–303, 2003.
- [26] The ATLAS Collaboration. The ATLAS Simulation Infrastructure. *The European Phys-*

- ical Journal C*, 70(3), 2010.
- [27] The ATLAS Collaboration. ATLAS HL-LHC Computing Conceptual Design Report. Technical report, CERN, 2020.
- [28] The ATLAS Collaboration. AtlFast3: The Next Generation of Fast Simulation in ATLAS. *Computing and Software for Big Science*, 6(1), 2022.
- [29] The ATLAS Collaboration. Performance of electron and photon triggers in ATLAS during LHC Run 2. *The European Physical Journal C*, 80(1), 2020.
- [30] The ATLAS Collaboration. Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton-proton collision data at $\sqrt{s} = 13$ TeV. *The European Physical Journal C*, 79(8), 2019.
- [31] The ATLAS Collaboration. Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data. *Journal of Instrumentation*, 14(12):P12006, 2019.
- [32] The ATLAS Collaboration. Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung. Technical report, CERN, 2012.
- [33] The ATLAS Collaboration. Electron and photon energy calibration with the ATLAS detector using LHC Run 2 data. *Journal of Instrumentation*, 19(02):P02009, 2024.
- [34] The ATLAS collaboration. Electron and photon efficiencies in LHC Run 2 with the ATLAS experiment. *JHEP*, 05:162, 2024.
- [35] The ATLAS Collaboration. Reconstruction of hadronic decay products of tau leptons with the ATLAS experiment. *The European Physical Journal C*, 76(5), 2016.
- [36] J. Maurer. Mesure des performances de reconstruction des électrons et recherche de Supersymétrie dans les canaux avec deux leptons de même charge dans les données du détecteur ATLAS, PhD thesis, 2013.
- [37] A. Hadeef. Search for Higgs boson production in association with a top quark pair in two same-sign lepton final states with the ATLAS detector at LHC, PhD thesis, 2017.
- [38] G. Tarna. Studies of the Higgs boson properties and search for new physics beyond the standard model in the top sector with the ATLAS detector, PhD thesis, 2019.

- [39] H. Rasheed. Constraining the 19D phenomenological Minimal Supersymmetric Standard Model with the ATLAS Experiment, PhD thesis, 2024.
- [40] The ATLAS collaboration. Electron efficiency measurements with the ATLAS detector using 2012 LHC proton–proton collision data. *Eur. Phys. J. C*, 77(3):195, 2017.
- [41] ATLAS Electron ID sub-group documentation. Scale factor recommendations. <https://atlas-electron-id.docs.cern.ch/recommendations> (ATLAS internal).

Acknowledgments

I am deeply grateful to the many people who supported, guided and encouraged me throughout this journey. First and foremost, I would like to express my gratitude to Prof. Arno Straessner, who supported me over the last years, extending far beyond the scope of this work. I am very thankful to Asma Hadeef for introducing me to this fascinating research topic and for her invaluable advice and expertise, which enriched this work significantly.

I would also like to sincerely thank Julien Maurer, Asma Hadeef, Grigore Tarna and Hammad Rasheed, whose work on electron reconstruction efficiencies laid the groundwork for the present analysis. This thesis builds directly upon the foundations and results they established and I am grateful for the opportunity to continue and extend their research.

A special word of thanks goes to Christian Schmidt, David Dahiya and Manuel Gutsche for the many interesting discussions we shared in the office - both research-related and off-topic. Their company made the last year truly enjoyable. I also want to thank the entire e/gamma group for their warm welcome. In particular, I am deeply grateful to Martina Ojeda, Nghia Nguyen, and Filip Nechanský for their patient guidance and continuous support.

I would also like to extend my sincere appreciation to Manuel Gutsche and Tom Jungnickel for taking the time to read this thesis on such short notice and for providing valuable feedback. Finally, I want to thank my parents, who have always encouraged me to find my own path. Without them, none of this would have been possible.

Acronyms

AF3 ATLFast3

ALICE A Large Ion Collider Experiment

ATLAS A Toroidal LHC ApparatuS

CERN European Organization for Nuclear Research

CMS Compact Muon Solenoid

CP Combined Performance

DNN Deep Neural Network

ECAL Electromagnetic Calorimeter

EM Electromagnetic

FS Full GEANT4 Simulation

GAN Generative Adversarial Network

HCAL Hadronic Calorimeter

ID Identification Criteria

ITk ATLAS Inner tracker

LAr Liquid Argon

LH Likelihood

LHC Large Hadron Collider

MC Monte Carlo simulation

MS Muon Spectrometer

SCT Semiconductor Tracker

SF Scale Factor

TRT Transition Radiation Tracker

Appendix

A.1 List of Samples

To enable reproducibility, all used data and simulation samples are listed. A skimmed derivation called **EGAM1** is used, which selects events passing at least one of the following criteria:

- two electron candidates with opposite sign, one is central and passes tight LH ID and $p_T > 25$ GeV, the other passes medium LH ID and $p_T > 20$ GeV
- two central electron candidates with opposite sign, both pass medium LH ID and $p_T > 19.5$ GeV
- two central electron candidates, one passes tight LH ID and $p_T > 25$ GeV the other one passes $p_T > 4$ GeV
- one central electron candidate passing tight ID and $p_T > 25$ GeV, as well as one central photon candidate with $p_T > 15$ GeV

The term central here refers to electrons and photons in the central region of the ATLAS detector ($|\eta| < 2.47$). Additionally, only events with an invariant mass m_{ee} or $m_{e\gamma}$ of at least 50 GeV are kept. $Z \rightarrow ee$ signal Monte Carlo events are generated using POWHEG and PYTHIA version 8.

Run 2 reprocessed data:

```
data15_13TeV.periodAllYear.physics_Main.PhysCont.DAOD_EGAM1.grp15_v01_p5535
data16_13TeV.periodAllYear.physics_Main.PhysCont.DAOD_EGAM1.grp16_v01_p5535
data17_13TeV.periodAllYear.physics_Main.PhysCont.DAOD_EGAM1.grp17_v01_p5535
data18_13TeV.periodAllYear.physics_Main.PhysCont.DAOD_EGAM1.grp18_v01_p5535
```

Run 2 FullSim (MC20 a,d,e):

```
mc20_13TeV.361106.PowhegPythia8EvtGen_AZNLOCTEQ6L1_Zee.deriv.DAOD_EGAM1.e3601_s3681_r13167_p4940
mc20_13TeV.361106.PowhegPythia8EvtGen_AZNLOCTEQ6L1_Zee.deriv.DAOD_EGAM1.e3601_s3681_r13144_p4940
mc20_13TeV.361106.PowhegPythia8EvtGen_AZNLOCTEQ6L1_Zee.deriv.DAOD_EGAM1.e3601_s3681_r13145_p4940
```

Run 2 FastSim (MC20 a,d,e):

```
mc20_13TeV.361106.PowhegPythia8EvtGen_AZNLOCTEQ6L1_Zee.deriv.DAOD_EGAM1.e3601_e5984_a907_r14859_r14919_p6598
mc20_13TeV.361106.PowhegPythia8EvtGen_AZNLOCTEQ6L1_Zee.deriv.DAOD_EGAM1.e3601_e5984_a907_r14860_r14919_p6598
mc20_13TeV.361106.PowhegPythia8EvtGen_AZNLOCTEQ6L1_Zee.deriv.DAOD_EGAM1.e3601_e5984_a907_r14861_r14919_p6598
```

Run 3 data:

`data22_13p6TeV.periodAllYear.physics_Main.PhysCont.DAOD_EGAM1.grp22_v01_p5859`

`data23_13p6TeV.periodAllYear.physics_Main.PhysCont.DAOD_EGAM1.grp23_v01_p5859`

`data24_13p6TeV.periodAllYear.physics_Main.PhysCont.DAOD_EGAMX.grp24_v01_p6143`

Run 3 FullSim (MC23 a,d,e):

`mc23_13p6TeV.601189.PhPy8EG_AZNLO_Zee.deriv.DAOD_EGAM1.e8514_s4162_r15540_p6286`

`mc23_13p6TeV.601189.PhPy8EG_AZNLO_Zee.deriv.DAOD_EGAM1.e8514_s4159_r15530_p6286`

`mc23_13p6TeV.601189.PhPy8EG_AZNLO_Zee.deriv.DAOD_EGAM1.e8514_s4369_r16083_p6544`

Run 3 FastSim (MC23 a,d):

`mc23_13p6TeV.601189.PhPy8EG_AZNLO_Zee.deriv.DAOD_EGAM1.e8514_e8528_a910_s4114_r14932_r14663_p6598`

`mc23_13p6TeV.601189.PhPy8EG_AZNLO_Zee.deriv.DAOD_EGAM1.e8514_a911_r14908_p5857`

A.2 List of Triggers

The single electron triggers used to select an event are summarized in Table A.2.1. Additionally, tag electrons are required to match the trigger electron.

Table A.2.1: List of used triggers in Run 2 and Run 3.

year	trigger
2015	HLT_e24_lhmedium_L1EM20VH
	HLT_e60_lhmedium
	HLT_e120_lhloose
2016	HLT_e60_lhmedium
	HLT_e26_lhtight_nod0_ivarloose
	HLT_e60_lhmedium_nod0
	HLT_e140_lhloose_nod0
	HLT_e300_etcut
2017	HLT_e26_lhtight_nod0_ivarloose
	HLT_e60_lhmedium_nod0
	HLT_e140_lhloose_nod0
	HLT_e300_etcut
2018	HLT_e26_lhtight_nod0_ivarloose
	HLT_e60_lhmedium_nod0
	HLT_e140_lhloose_nod0
	HLT_e300_etcut
2022	HLT_e26_lhtight_ivarloose_L1EM22VHI
	HLT_e60_lhmedium_L1EM22VHI
	HLT_e140_lhloose_L1EM22VHI
	HLT_e300_etcut_L1EM22VHI
2023	HLT_e26_lhtight_ivarloose_L1eEM26M
	HLT_e60_lhmedium_L1eEM26M
	HLT_e140_lhloose_L1eEM26M
	HLT_e300_etcut_L1eEM26M
2024	HLT_e26_lhtight_ivarloose_L1eEM26M
	HLT_e60_lhmedium_L1eEM26M
	HLT_e140_lhloose_L1eEM26M
	HLT_e300_etcut_L1eEM26M

A.3 Reconstruction Efficiencies

A.3.1 Comparison between Run 2 and Run 3

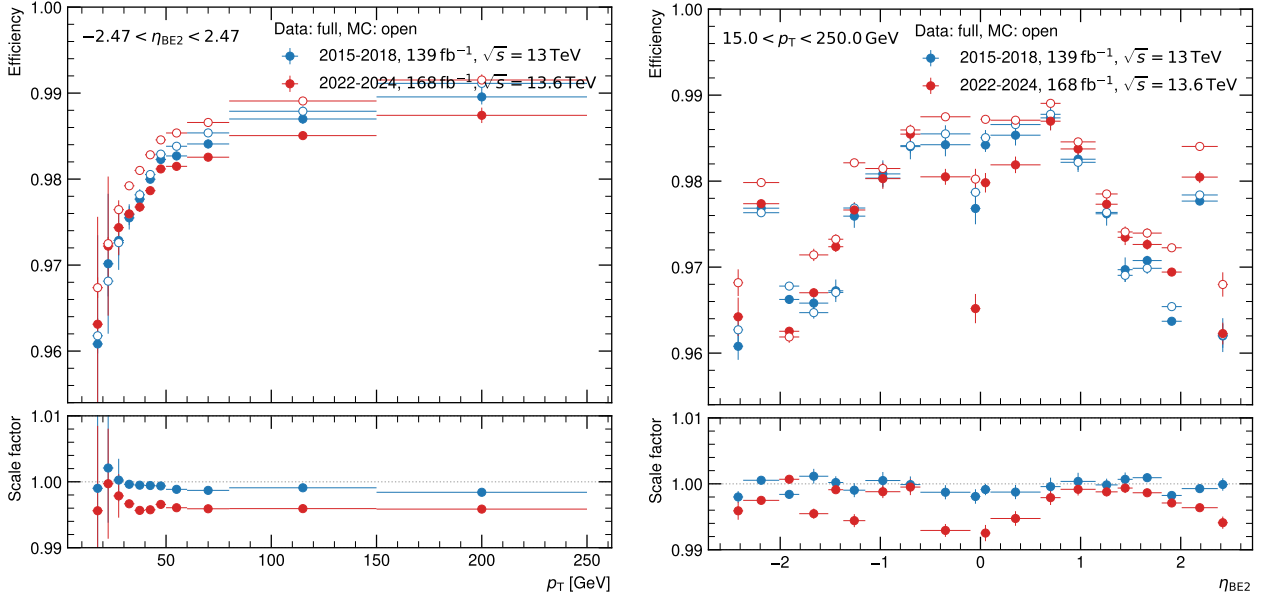


Figure A.3.1: Electron reconstruction efficiency as a function of the transverse momentum integrated over the pseudorapidity range $-2.47 < \eta < 2.47$ (left) and as a function of η integrated over the p_T range $15.0 < p_T < 250.0$ GeV (right). The full Run 2 dataset (blue) is compared with the currently available Run 3 results (red).

A.3.2 Impact of Cluster Quality on Efficiencies

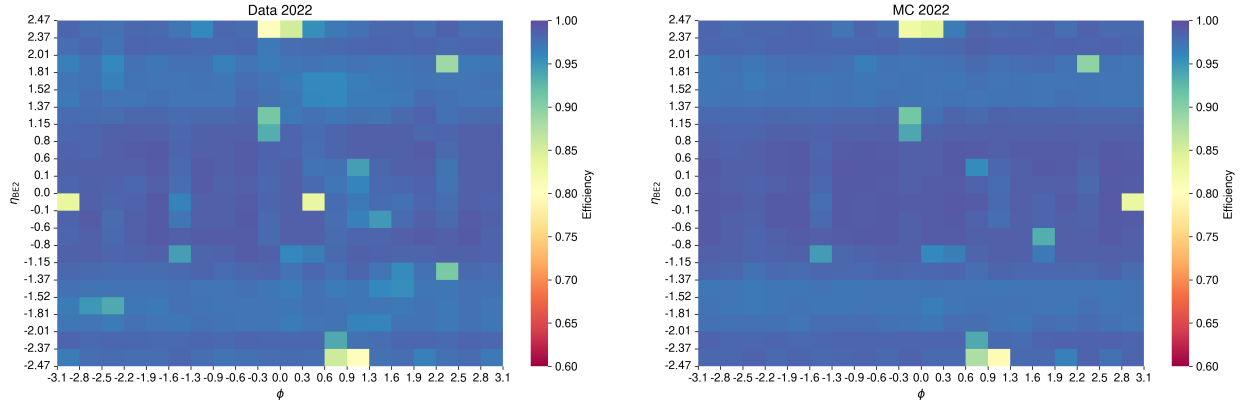


Figure A.3.2: Electron reconstruction efficiencies in 2022 as a function of η and ϕ for data (left) and simulation (right).

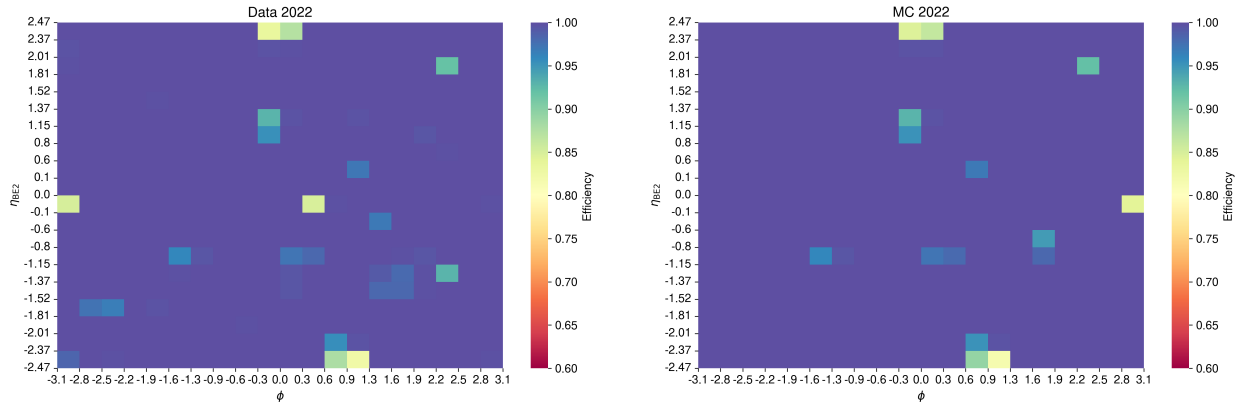
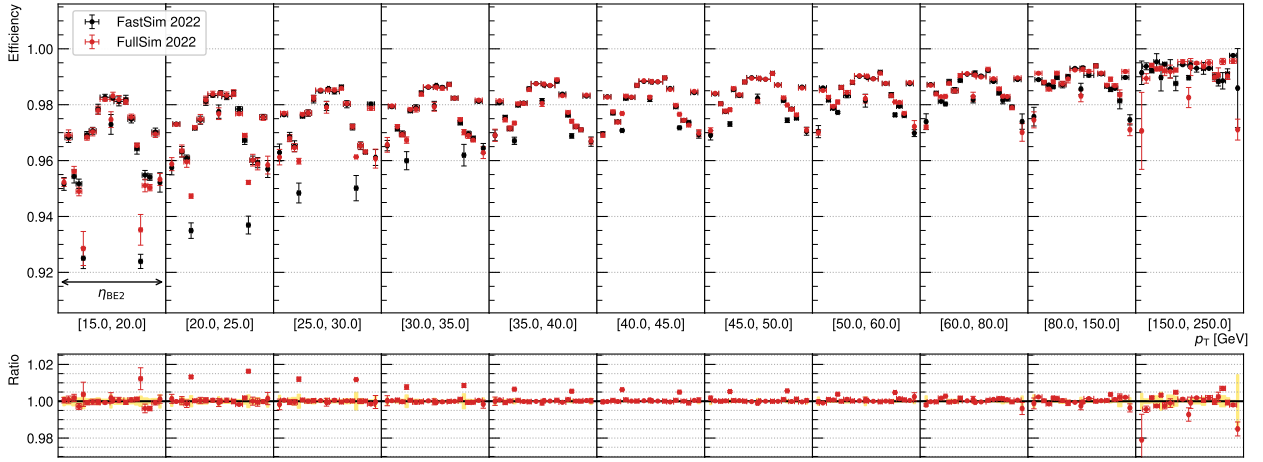
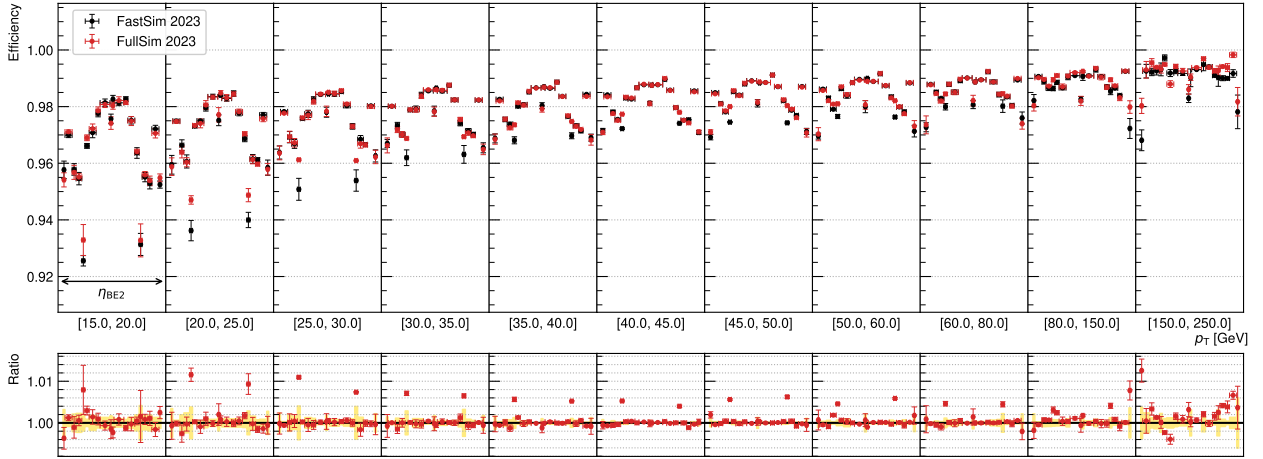


Figure A.3.3: Impact of only masked cell core cut on electron reconstruction efficiencies in 2022 as a function of in $\eta - \phi$ for data(left) and simulation (right).

A.3.3 Comparison between Full and Fast Simulation

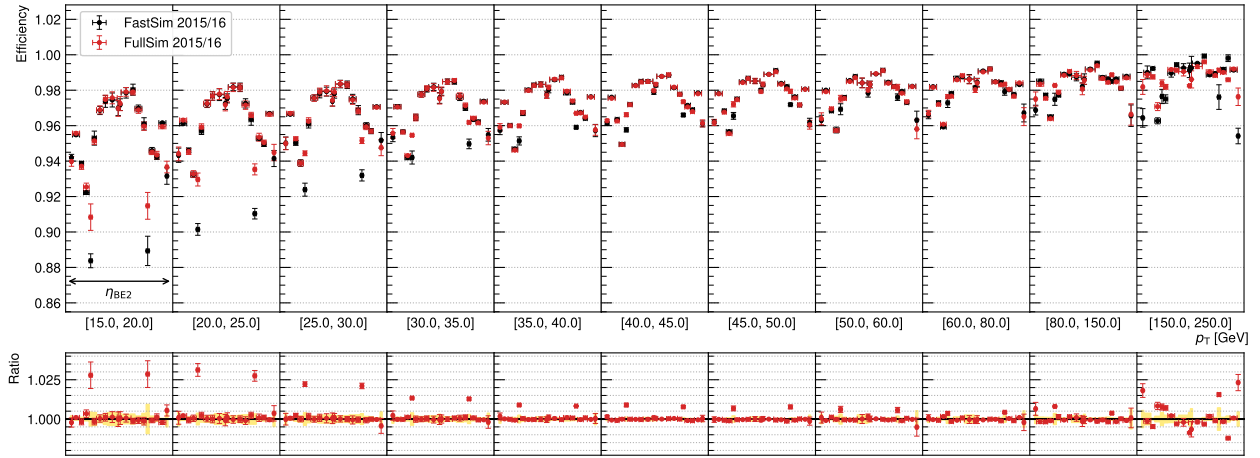


(a) 2022 (MC23a)

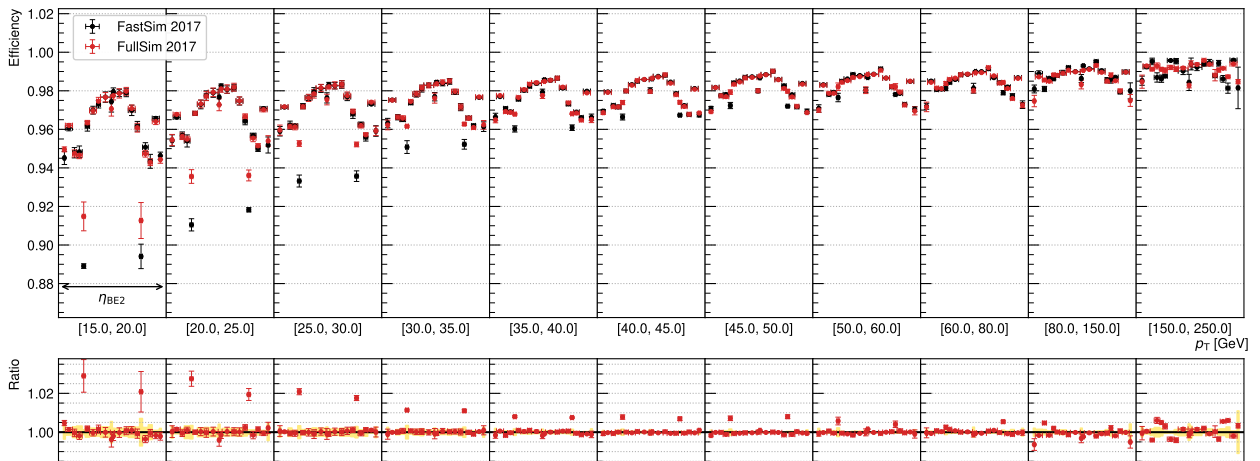


(b) 2023 (MC23d)

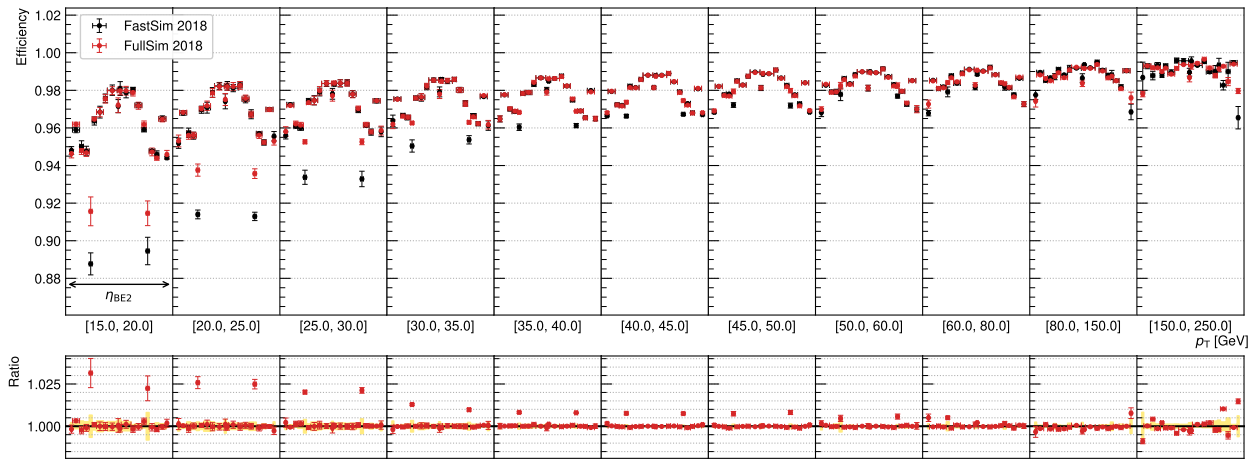
Figure A.3.4: Electron reconstruction efficiencies in Run 3 as a function of p_T and η for full simulation (red) and AF3 fast simulation (black) and their ratio, which corresponds to a MC-to-MC scale factor.



(a) 2015+2016 (MC20a)



(b) 2017 (MC20d)



(c) 2018 (MC20e)

Figure A.3.5: Electron reconstruction efficiencies in Run 2 as a function of p_T and η for full simulation (red) and AF3 fast simulation (black) and their ratio, which corresponds to a MC-to-MC scale factor.

A.3.4 Run 3 Results using the new Fitting Setup

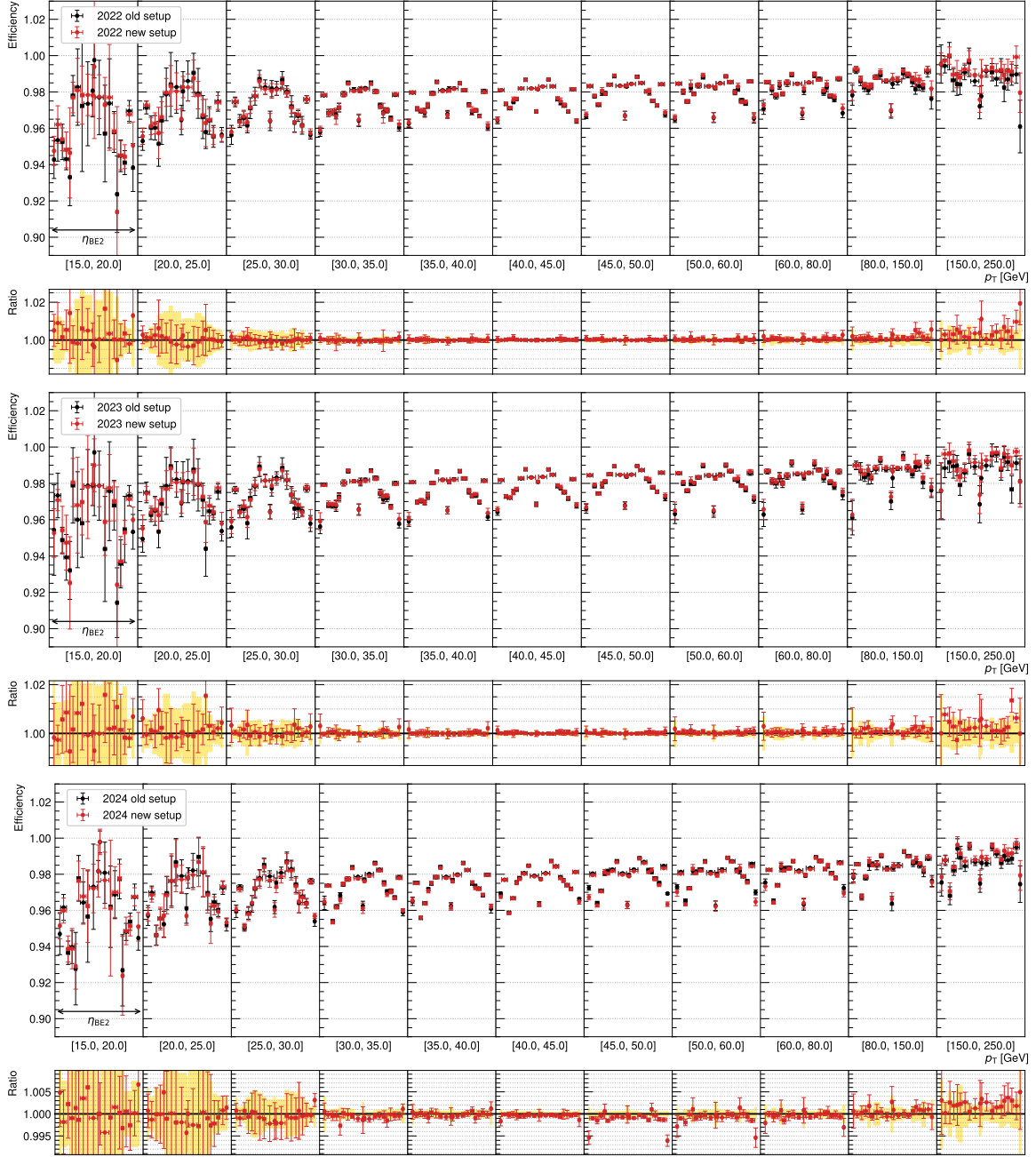


Figure A.3.6: Electron reconstruction efficiencies in Run 3 as a function of p_T and η for 2022 (top), 2023 (mid), 2024 (bot). Results using the previous fitting setup (black) are compared with results obtained with 4th-order Bernstein fits in the new setup (red).

Erklärung

Hiermit erkläre ich, dass ich diese Arbeit im Rahmen der Betreuung am Institut für Kern- und Teilchenphysik ohne unzulässige Hilfe Dritter verfasst und alle Quellen als solche gekennzeichnet habe.

Marius Melcher

Dresden, Oktober 2025