Department of Physics
Technical University Munich

Master Thesis in Physics
submitted by

Katharina Korf

born in Bad Reichenhall (Germany)

February 2026

Measuring Belle II's track–finding efficiency with radiative Bhabha scattering $e^+e^- \to e^+e^-\gamma$

This Master Thesis has been carried out by Katharina Korf at the
Department of Physics at Technical University Munich (TUM)
under the supervision of
Dr. Daniel Greenwald, Lukas Grußbach, Prof. Dr. Stephan Paul

Abstract

This thesis determines the tracking efficiency of Belle II in radiative Bhabha scattering events, $e^+e^- \to e^+e^-\gamma$, as a function of charge, momentum magnitude, and polar angle. The events are reconstructed using a tag electron and two energy clusters reconstructed in the electromagnetic calorimeter, which do not inherit any information about the second track. From the well-known decay topology of radiative Bhabha events, the cluster originating from the charged particle is extracted. The purity and resolution of the dataset are improved by implementing a boosted decision tree that classifies events with the higher-energy cluster originating from the charged particle and those with the lower-energy cluster originating from the charged particle. The overall results are computed for a binning in the cluster energy and the cluster polar angle of the higher-energy cluster. The tracking efficiencies, calculated using an unfolding procedure that accounts for residual background contamination, are above 99 % with uncertainties below 0.01 %, yielding results from experimental data and simulation. For the correction factor between experimental data and simulation, the uncertainty is smaller than 0.01 % in all bins, which is a clear improvement over the previous tracking efficiency result at Belle II, which is 0.27 %. This result is determined for a lower energy region than our result from the radiative Bhabha study.

Zusammenfassung

Diese Arbeit bestimmt die Spurfindungseffizienz von Belle II in strahlenden Bhabha-Scattering-Ereignissen, $e^+e^- \to e^+e^-\gamma$, als Funktion von Ladung, Impulsgröße und Polarwinkel. Die rekonstruierten Events haben einen Tag-Elektron und zwei Energiecluster im elektromagnetischen Kalorimeter, welche keine Informationen über die zweite Spur enthalten. Aus der bekannten Zerfallstopologie der strahlenden Bhabha-Events extrahieren wir, welches der Cluster von einem geladenen Teilchen kommt. Die Reinheit und Auflösung des Datensets werden durch die Verwendung eines Boosted Decision Trees verbessert, indem Ereignisse, bei denen das höherenergetische Cluster vom geladenen Teilchen stammt, und solche, bei denen das niederenergetische Cluster vom geladenen Teilchen stammt, klassifiziert werden. Die Gesamtergebnisse werden für alle Bins im Phasenraum, aus der Clusterenergie und dem Polarwinkel des höherenergetischen Clusters berechnet. Für die Berechnung der Spurfindungseffizienz wird ein Entfaltungsverfahren verwendet, dabei werden Hintergrundkontaminationen berücksichtigt. Die Spurfindungseffizienz ist größer als 99 % mit einer Unsicherheit kleiner als 0.01 % sowohl für experimentelle Daten als auch für Simulationen. Der Korrekturfaktor zwischen experimentellen Daten und Simulationen hat eine Unsicherheit kleiner als 0.01 % in allen Bins. Dies ist eine klare Verbesserung des bisherigen Spurfindungseffizienzergebnisses von Belle II. Die bisherige Unsicherheit beträgt 0.27 % und wurde für einen niedrigeren Energiebereich berechnet als das Ergebnis aus der strahlenden Bahbha-Streuung.

Contents

Chapter 1

Introduction

In high-energy physics (HEP), a common way to perform experiments and measurements is to use a particle collider. At these facilities, charged particles are accelerated and collide at a specific interaction point with a certain collision energy. In those collisions, a large amount of charged and neutral particles is produced, decaying into different final state particles. The data taking and the monitoring of the created particles are done by several detectors placed around the collision point.

By using the detector outputs, the numbers of events of specific types are determined. This number is denoted as the yield Y_x, where x represents the analyzed process. The number of originally analyzed events of the type x is N_x. These two variables are related by

$$\mathrm{Y}_x = \epsilon_x N_x, \tag{1.0.1}$$

where ϵ_x is the efficiency of reconstructing an event of type x. This efficiency is a useful parameter for HEP experiments. One example is the branching fraction $\mathcal{B}$ of the scattering $e^+e^- \to e^+e^-\gamma$, which is determined by

$$\mathcal{B}_i(e^+e^- \to e^+e^-\gamma) = \frac{N(e^+e^- \to e^+e^-\gamma)}{N_{\text{total}}} = \frac{Y(e^+e^- \to e^+e^-\gamma)/\epsilon_i}{N_{\text{total}}}. \tag{1.0.2}$$

The branching fraction describes the probability of the particle collision producing a certain final state by dividing the number of expected final states of the specific decay by the total number of events within this decay channel N_{total}.

To determine the branching fraction, the efficiency for reconstructing events of interest is necessary. For most cases, this efficiency cannot be determined from the data as the number of expected events is unknown. Therefore, each HEP experiment has its own simulations, where particles, their different decay modes, and their behavior within a virtual model of the detector are simulated. The simulations are then used to determine the efficiency. To verify the reliability of these simulations, it is necessary to identify decay modes for which the number of expected events can be determined in experimental data. The events of those decay modes can be extracted by setting constraints during the decay reconstruction.

The number of expected decays is then extracted from the context of the decay. This means that from the decay reconstruction and the involved constraints, it becomes clear if the correct decay was reconstructed or not. This information is then used to determine the efficiency of reconstructing this decay in data. From calculating this efficiency in data and in simulation, a correction factor is determined. The efficiencies estimated from the simulation of other decays are then multiplied by this correction factor, so that the determined efficiency from the simulation matches the efficiency in the data.

The statistical uncertainty of the correction factor is determined, and when applying the correction factor to another efficiency, this statistical uncertainty becomes part of the systematic uncertainty of the estimated efficiency. The final results in HEP always have a statistical and a systematic uncertainty, where the systematic uncertainties come from detector inefficiencies or the statistical uncertainties of the used values within the analysis, reflecting the uncertainties of all quantities used in the calculation. To increase the precision of results, it is important to decrease the uncertainty on values like the efficiency for reconstructing a specific decay, and in conclusion, also the systematic uncertainty of other results decreases.

A usage of the efficiency for reconstructing events is the determination of the Weinberg-angle, also known as the weak-mixing angle $\sin^2\theta_\mathrm{w}$, or the forward-backward asymmetry for a certain decay, which is pursued by the doctoral candidate Lukas Grußbach at the Belle II experiment. The forward-backward asymmetry is determined by counting the number of charged particles ending up in the forward region of the detector N_F and those ending up in the backward region of the detector N_B. The forward direction is defined by $\cos\theta_\mathrm{cl,lab} > 0$ and the backward direction by $\cos\theta_\mathrm{cl,lab} < 0$. The forward-backward asymmetry is determined as

$$\mathcal{A}_\mathrm{FB} = \frac{N_\mathrm{F} - N_\mathrm{B}}{N_\mathrm{F} + N_\mathrm{B}}. \tag{1.0.3}$$

A tracking efficiency uncertainty binned in energy and momentum is here especially helpful to account for differences in the tracking algorithm based on where the particles hit the detector.

The Belle II experiment is a HEP experiment colliding an electron and a positron, located in Japan. The charged electrons collide at a center-of-mass energy of $\sqrt{s} = 10.58\,\mathrm{GeV}$ and produce various charged and neutral particles of which some are final-state particles while others decay further into other final state particles. At Belle II, the following final state particles can be detected: charged leptons, Kaons and pions, photons, long-lived hadrons like neutrons, protons, Deuterons, and K_L's.

We distinguish two groups of final state particles, the neutral particles and the charged ones. The neutral particles can only be detected by energy depositions within the detector. Reconstructing charged particles is divided into two parts at Belle II. The charged particles interact with the detectors, and therefore, an energy deposition within the detectors is measured. Secondly, the charged particles produce a track within the detector volume. The track is hereby the trajectory of the charged particle starting at the interaction point, or, respectively, the secondary decay vertex, and propagating outwards through the detector.

The determination of these trajectories and matching them to the corresponding hits in the detectors is called tracking. The tracking process itself is performed by algorithms, and the tracking efficiency describes their performance. The tracking efficiency is determined by dividing the number of reconstructed events by the number of generated events. As already mentioned, this efficiency can be determined from a simulation easily, but a calculation in data is only possible in certain contexts. For the determination in data, it is necessary to know exactly which particles were reconstructed and where the charged particles went without using tracking information. With this knowledge, the correction factor of the tracking efficiency can be determined by dividing the efficiency in data by the efficiency in simulation. The correction factor improves the scaling of the simulation towards the data, and its statistical uncertainty serves as an input to the systematic uncertainty of other analyses.

For Belle II, a result for the tracking efficiency already exists [1]. This result is one per-track uncertainty value independent of charge or direction. But for determining the forward-backward asymmetry and other physical analyses, a result that accounts for differences in the tracking depending on where the trajectory is going within the detector is more helpful. Therefore, it makes sense to recalculate the tracking efficiency in dependence on the polarity of the charge, the momentum magnitude in the lab frame, and the momentum direction, so the polar angle.

This thesis will determine the tracking efficiency of Belle II in radiative Bhabha scattering events in the dependence of the charge q, the momentum p_{lab}, and the polar angle θ_{lab}. We perform the decay reconstruction with the goal of finding a rule for where the charged particle went. To finalize this rule, a boosted decision tree is trained and applied. For the final data sample, resolution, purity, and data-simulation agreement are checked. Finally, the tracking efficiency is determined in dependence on charge, momentum, and polar angle, and in this dependence, also the calculation of the correction factor and the corresponding statistical uncertainties is carried out. The thesis concludes with a comparison of data and simulation, as well as the different charges.

Chapter 2

Tracking efficiency study using $e^+e^- \to \tau^+\tau^-$

The currently used systematic uncertainty to account for track finding inefficiencies is 0.27 %, which is the recommended systematic uncertainty from track finding per track, [1]. This result was achieved using $e^+e^- \to \tau^+\tau^-$ events in an energy range of 0.5 GeV to 4.5 GeV.

To determine the tracking efficiency, $\tau\tau$ events were used as they have a high cross section $\sigma_{\tau\tau} = (0.919 \pm 0.003)\,\mathrm{nb}$ at Belle II, leading to a large data set and therefore lower statistical uncertainties. Furthermore, the jet-like topology of the τ decay products suits well for a tracking performance study in low-multiplicity but high-density track environments [1]. Finally, the kinematics of the τ-pairs cover a wide track-momentum range.

The short-lived τ's decay either hadronically or leptonically. The final state particles from those decays are then reconstructed and used for the analysis. The study used two decay channels:

$$\tau^+ \to \ell^+\ell^+\ell^-\nu_\ell\bar{\nu_\tau}$$
$$\tau^+ \to \pi^+\pi^-\pi^+\bar{\nu_\tau} + n\pi^0$$

For better understanding, only positively charged τ's are listed here, but the same yields for the opposite charges. The n in front of the neutral pion is an integer number and gives the number of extra neutral pions involved in the decay.

For the τ study performed to determine the tracking efficiency, the tag-probe method is used. The tag-probe method is a standard tool used for decay reconstructions at Belle II. The reconstruction uses tight constraints on the tag side, ensuring it provides a high-confidence signal and that the reconstructed decay follows the decay topology we are aiming for. The probe side contains the particle(s) for which the tracking efficiency is to be measured. The selection criteria for the probe side can be quite loose so that a bias within the decay reconstruction is avoided.

When analyzing τ decays, it is important to keep in mind that the decay reconstruction consists of two steps. In the e^+e^- collision, $\tau^+\tau^-$

are created back-to-back, and both of them decay into other final state particles. This allows us to divide the phase space into two hemispheres. From the two τ decays, one is constrained as the tag side, and the other τ decay is used as the probe side. The tagged τ decay is constrained in such a way that the probed τ decays into three charged particles. This decay is also known as the probe side. Finally, for the probe side, the tag-probe method is applied again. The tag side of the probe is defined by having two high-quality tracks, which leads to a total known charge of ± 1.

A good quality track is defined by the vertex of tracks in z with respect to the interaction point (IP), the transverse distance of a vertex with respect to the IP, and the transverse momentum. The exact constraints are $|dz| < 3\,\mathrm{cm}$, $dr < 1\,\mathrm{cm}$, and $p_\mathrm{T} > 200\,\mathrm{MeV}$.

The charge conservation law requires the overall decay to have an even number of final-state charged particles. Therefore, the characterized τ-pair events need to have one additional track. This probe track has to fulfill: $|dz| < 3\,\mathrm{cm}$ and $dr < 1\,\mathrm{cm}$. Following this definition, the tracking efficiency ϵ_tr describes whether the probe track was reconstructed or not

$$\epsilon_\mathrm{tr} = \frac{N_4}{N_3 + N_4}, \tag{2.0.1}$$

N_3 and N_4 numbers of events with three or four reconstructed tracks. The tag selection differs depending on the particle species. The electrons need to have a ratio of cluster energy over momentum E_cl/p between 0.8 GeV and 1.2 GeV and no restriction on the muonID. The muonID refers to the muon identification probability, which combines information from all available detectors. Muons and tauons need to fulfill $E_\mathrm{cl}/p = (0 - 0.6)$ GeV. For muons μ, the muonID needs to be larger than 0.9, while for taus, the muonID needs to be smaller than 0.9.

For the analysis in [1], a significant number of background processes remain after the reconstruction process. These unwanted decays are mainly from $e^+e^- \to q\bar{q}$, $e^+e^- \to \ell^+\ell^-\gamma$, and $e^+e^- \to e^+e^-\mu^+\mu^-$. Their amounts decrease by applying further suppression criteria, like limiting the invariant mass of the three tag tracks, the CM-frame momentum of the tag lepton track, or the angular separation between the different tag tracks.

Determining the tracking efficiency from Equation (2.0.1) for data and MC individually, the data-MC discrepancy is

$$\delta = \frac{\epsilon_\mathrm{MC} - \epsilon_\mathrm{Data}}{\epsilon_\mathrm{MC}} = 1 - \frac{\epsilon_\mathrm{Data}}{\epsilon_\mathrm{MC}}. \tag{2.0.2}$$

The final calibrated data-MC discrepancy in [1] is given as

$$\delta^* = 0.049 \,\pm\, 0.019\,(\mathrm{stat.}) \,\pm\, 0.268\,(\mathrm{sys.})\,\%. \tag{2.0.3}$$

The statistical uncertainty (stat.) accounts here for statistical fluctuation and other statistical influences. The systematic uncertainty (sys.) describes the uncertainties coming from the detectors or from other results used within the analysis.

The systematic uncertainty is measured by varying different parameters. The largest contribution of 97 % comes from charge asymmetry. This asymmetry is determined by requiring that the probe track either has a positive or a negative charge.

Concluding from the result in Equation (2.0.3), the recommendation of using 0.27 % as systematic per track uncertainty for the tracking efficiency was given. As the systematic uncertainty is an additive component and the uncertainty is given per-track, this result will lead to large systematic uncertainties for further studies. In addition, it is important to mention that the energy range covered by the τ-study is relatively small, going from 0.5 GeV to 4.5 GeV, as it is determined from a three-body decay involving also several neutrinos.

2.1 Choice of radiative Bhabha scattering

As mentioned in the beginning, it is helpful to have the tracking efficiency and its systematic uncertainty determined in dependence on polar angle, momentum, and charge. The current result at Belle II accounts for neither of these. Concerning the charge dependence, it remains unclear why the result from the τ-study was not given separately for the different charges of the probe. Especially, when considering that the systematic uncertainty is dominated by charge asymmetry.

Further, the τ decay cannot be used to determine a result depending on momentum magnitude and momentum direction, as τ decays involve neutrinos ν. In the Belle II detector, neutrinos cannot be measured, and so the energy of those is unknown in the decay reconstruction. Therefore, the energy and the direction of the track cannot be completely resolved in the final state.

To address this problem, a decay without neutrinos is chosen, e.g., $e^+e^- \to e^+e^-$ to determine a new tracking result. Electrons are stable, and deposit almost their entire energy in the Electromagnetic calorimeter (ECL) of the Belle II detector, so that the energy detected in the ECL can be assumed to be the momentum of the electron. There can still be missing energy, e.g., due to low energetic photons which the detector cannot identify, but those are only minor contributions.

Bhabha scattering $e^+e^- \to e^+e^-$ has a cross section of $\sigma_{ee} = (294 \pm 2)\,\text{nb}$ which is significantly larger than the $e^+e^- \to \tau^+\tau^-$ cross section $\sigma_{\tau\tau} = 0.919\,\text{nb}$ [2]. This large cross-section of Bhabha scattering would allow for a minimization of the statistical uncertainty. However, due to its large cross-section, the Bhabha decay is significantly suppressed at Belle II by the triggers. Without a trigger application, Bhabha events would dominate all other processes. To avoid most events being rejected by the triggers, we use the radiative Bhabha decay, $e^+e^- \to e^+e^-\gamma$, instead.

Using radiative Bhabha decay also addresses the problem of the strong correlation between the momentum and the polar angle for Bhabha decays.

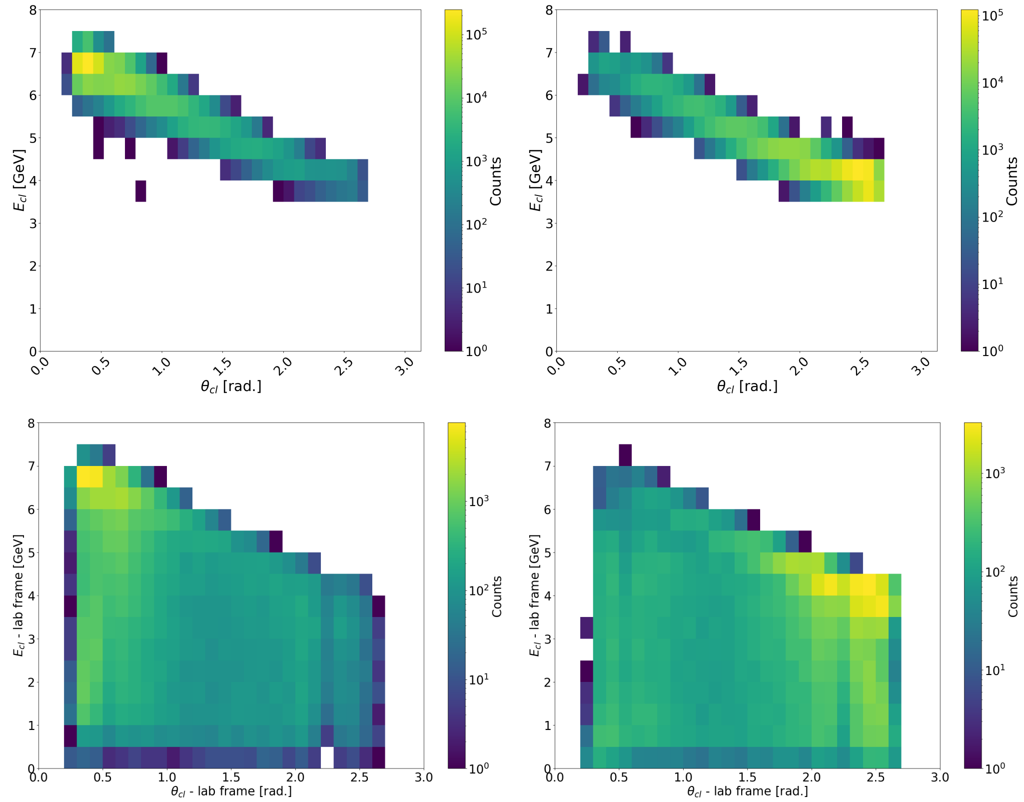

Figure 2.1.1: The polar angle of the ECL cluster versus ECL cluster energy, both in the lab frame, for particles in the Bhabha decay process $e^+e^- \to e^+e^-$ (top) and for particles in the radiative Bhabha decay $e^+e^- \to e^+e^-\gamma$ (bottom), for the reconstructed electron (left) and positron (right).

In the lab frame, the momenta and polar angles of the charged particles produced in Bhabha decays cover only an energy range of above 3 GeV, as seen in the top panels of Figure 2.1.1. When an additional photon is included in the decay process, the phase space coverage expands, as in the bottom panel of Figure 2.1.1.

For the radiative Bhabha decay, two different tag particles e^+ and e^- are used in the reconstruction, which leads to two separate charge-dependent values for the tracking efficiency. The charge dependence is especially important as the detectors consist of material and not of anti-material. Particles and anti-particles interact differently with those materials.

The radiative Bhabha decay $e^+e^- \to e^+e^-\gamma$ seems to be a good option for the determination of the tracking efficiency depending on charge, momentum, and polar angle. Further, in the previous result determined from $e^+e^- \to \tau^+\tau^-$, the method for the calculation of the tracking efficiency was not discussed. But the tracking efficiency is a statistical measure and is therefore handled like any other statistical efficiency.

2.2 Statistical treatment of efficiency

Statistically, efficiency is the probability of reconstructing events in relation to the number of generated events. For a statistical process with a total number of events N, we know how to determine the number of correctly reconstructed events K out of N. A binomial distribution describes the probability of getting K events

$$P(K|N,\epsilon) = \binom{N}{K}\epsilon^K(1-\epsilon)^{N-K}. \tag{2.2.1}$$

By using Bayes theorem, the probability of getting ϵ_{tr} for given K and N is

$$P(\epsilon_{\mathrm{tr}}|N,K) \;\propto\; P(K|N,\epsilon_{\mathrm{tr}})P_0(\epsilon_{\mathrm{tr}}). \tag{2.2.2}$$

Where $P_0(\epsilon_{\mathrm{tr}})$ is the prior knowledge of the efficiency. For the calculation of the tracking efficiency it yields: $\epsilon_{\mathrm{tr}} \in [0,1]$ with all values being equally possible. As there is no prior knowledge for the tracking efficiency we don't want to favor lower or higher probabilities and therefore a flat prior is chosen.

Using this information, the new probability function for the tracking efficiency ϵ_{tr} is

$$P(\epsilon_{\mathrm{tr}}|N,K) = P(K|N,\epsilon_{\mathrm{tr}}) = \binom{N}{K}\epsilon_{\mathrm{tr}}^K(1-\epsilon_{\mathrm{tr}})^{N-K}. \tag{2.2.3}$$

This probability density function has the form of a Beta function $B(\alpha,\beta)$

$$P(\epsilon_{\mathrm{tr}}|N,K) = Beta(\epsilon_{\mathrm{tr}};K+1,N-K+1) = \frac{\epsilon_{\mathrm{tr}}^K(1-\epsilon_{\mathrm{tr}})^{N-K}}{Beta(K+1,N-K+1)}. \tag{2.2.4}$$

For the Beta distribution, the mean and the variance are known.

$$\mathbb{E}[\epsilon_{\mathrm{tr}}] = \frac{\alpha}{\alpha+\beta} = \frac{K+1}{N+3} \tag{2.2.5}$$

$$\sigma^2_{\epsilon_{\mathrm{tr}}} = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \tag{2.2.6}$$

The variance is determined more easily using the mean value $\mathbb{E}[\epsilon_{\mathrm{tr}}]^2$:

$$\frac{\sigma^2_{\epsilon_{\mathrm{tr}}}}{\mathbb{E}[\epsilon_{\mathrm{tr}}]^2} = \frac{(\alpha+\beta)^2}{\alpha^2}\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}. \tag{2.2.7}$$

Inserting the correct terms for α and β and reshaping the function, the ratio of variance over mean can be written as

$$\sigma^2_{\epsilon_{\mathrm{tr}}} = \mathbb{E}[\epsilon_{\mathrm{tr}}]^2\frac{N-K+1}{(K+1)(N+3)}. \tag{2.2.8}$$

For the calculation of the efficiency at Belle II, large datasets are used, so we have large numbers for N and K. Further, approximations of the mean in Equation (2.2.5) and the variance in Equation (2.2.8) are

$$\mathbb{E}[\epsilon_{\text{tr}}] \xrightarrow[largeN,K]{} \frac{K}{N} \tag{2.2.9}$$

$$\sigma_{\epsilon_{\text{tr}}} \xrightarrow[largeN,K]{} \frac{K}{N}\sqrt{\frac{1}{K} - \frac{1}{N}}. \tag{2.2.10}$$

Another important aspect for calculating efficiencies is that, when K contains only a very small number of events, an upper limit calculation should be performed to estimate a reasonable reconstruction efficiency. This upper limit calculation is done by using the incomplete Beta function

$$Beta(\epsilon|k+1, N-k+1) = I_\epsilon(\alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^x u^{\alpha-1}(1-u)^{\beta-1}du \tag{2.2.11}$$

where for our case $\alpha = k+1$ and $\beta = N-k+1$. This gives two special cases: $k = 0$ (no reconstructed events) and $k = N$ (number of reconstructed events is equal to the total number of events). The incomplete Beta function results are

$$\text{for } k = 0 : I_\epsilon(1, N+1) = 1 - (1-\epsilon)^{N+1} \tag{2.2.12}$$

$$\text{for } k = N : I_\epsilon(N+1, 1) = \epsilon^{N+1}. \tag{2.2.13}$$

The upper limit efficiency is calculated by giving I_ϵ a fixed value and solving Equation (2.2.12) and Equation (2.2.13) for ϵ.

Chapter 3

Experimental framework

3.1 The SuperKEKB collider

The B-factory „SuperKEKB“ located in Tsukuba, Japan, is an electron-positron double-ring circular collider. The 7 GeV electron beam and the 4 GeV positron beam collide at center-of-mass (CM) energy, which is close to the energy of the $\Upsilon(4S)$ resonance [3], of $E_{\Upsilon(4\mathrm{S})} = 10.58\,\mathrm{GeV}$ [4]. The center-of-mass energy is chosen for operation as it is slightly higher than the rest energy of a B-meson pair, [3], leading to a low number of fragmentation particles when producing B-meson pairs. The SuperKEKB collider is operating with a so-called nano-beam, where the beam size at the collision point is only 50 nm, leading to a beampipe radius of max. 10 mm in the interaction region, [3] and [5]. The collider aims for a design luminosity of $8 \times 10^{35}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, the current peak luminosity of $5.1 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ was reached in December 2024 [6].

The asymmetric beam geometry creates a Lorentz boost in the forward direction within the CM frame, allowing for, e.g., time-dependent charge-parity symmetry violation measurements, [3]. The two accelerator rings at SuperKEKB have each a circumference of approximately 3 km and the two beams collide at an angle of $83\,\mathrm{mrad} = 4.75°$.

3.2 The Belle II detector

The Belle II detector, operating at the collision point of the SuperKEKB collider, is run by the Belle II collaboration. Its primary physics goal is to search for physics beyond the Standard Model in the flavor sector and to improve precision measurements of the Standard Model. The Belle II detector is a significantly upgraded version of the Belle detector, designed to improve the precision of measurements and to provide the possibility to perform a larger variety of analyses. Belle II started data taking in 2019, since then $575.47\,\mathrm{fb}^{-1}$ were accumulated, as of January 6, 2025 [7].

As the name B-factory already implies, B-physics play a major part in the Belle and Belle II physics programs, but with the upgrade to Belle

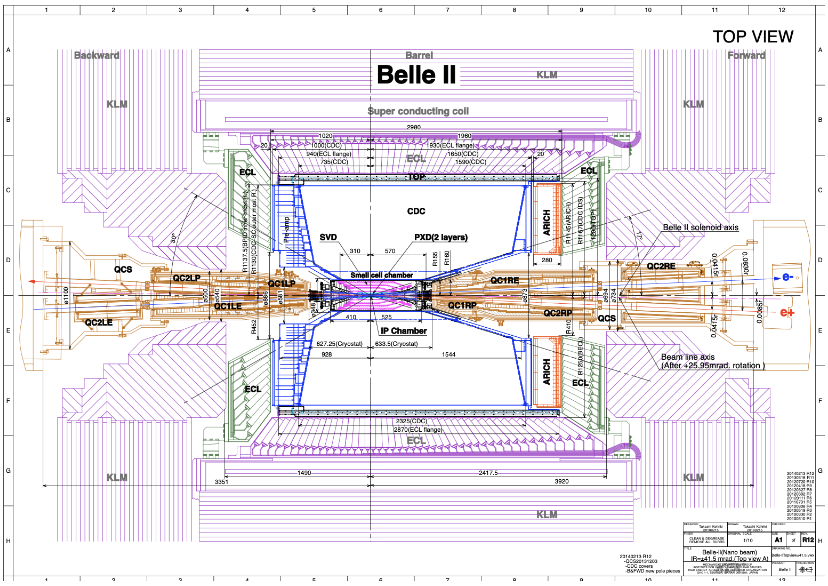

Figure 3.2.1: Setup of the Belle II detector [3].

II, also τ physics, dark sector searches, and precision measurements of low-multiplicity and initial-state radiation (ISR) processes are interesting studies to carry out. Some advantages of the new Belle II detector, in comparison to the Belle detector, are its excellent vertex resolution, good momentum resolution over the total kinematic range, and coverage of nearly the full solid angle. The overall setup of the Belle II detector is presented in Figure 3.2.1, and the subdetector systems are described in greater detail in the following, [3] and [5].

The region around the interaction point (IP) of the electron beams is marked as „IP chamber“ in Figure 3.2.1. All subdetectors are arranged in a cylinder around the IP region, where some detectors are then split into forward and backward endcaps and the barrel region.

The IP chamber is surrounded by the innermost detector, the **vertex detector (VXD)**, consisting of the silicon pixel detector (PXD) covering the two innermost layers and the silicon vertex detector (SVD) for the rest of the VXD. It is followed by the **central drift chamber (CDC)** and the particle identification (PID) system. The PID is divided into the **time-of-propagation (TOP) counter**, covering the barrel region, and the **Aerogel Ring-Imaging Cherenkov detector (ARICH)** covering the forward endcap region. After the PID system, the **electromagnetic calorimeter (ECL)** is placed, and it is followed by the **K_L^0-Muon Detector (KLM)**. Between the ECL and the KLM, the superconducting magnets are placed.

The central tracking device of the Belle II detector is the CDC. Here, the charged tracks are reconstructed, and their momenta are measured precisely. Further, the energy loss within the gas volume is measured and used for particle identification. The CDC provides reliable and efficient

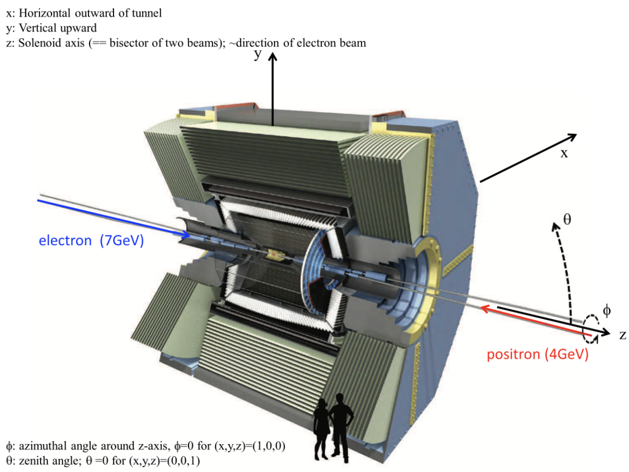

Figure 3.2.2: Belle II Coordinate system [8].

trigger signals for charged particles. The detector reaches from $r_{\text{min}} = 160\,\text{mm}$ to $r_{\text{max}} = 1130\,\text{mm}$. It is filled with a 50:50 mixture of $He-C_2H_6$, and is built out of 56 layers of sense wires. The wires are oriented in different ways so that a combination of the gained information gives the ability to reconstruct a full 3D helix for each (charged) particle candidate.

To handle the large amount of neutral particles produced in the collisions, a high-resolution ECL is required. It detects photons and determines their properties, identifies electrons, generates a proper signal for the trigger, and together with the K_L^0-Muon Detector, it detects K_L^0's. The calorimeter is built from 8736 thallium-doped caesium iodide scintillation crystals, which are assembled in the three different parts of the ECL, the barrel and the forward and backward endcaps. The ECL has an inner radius of 1.25 m and the region covered by the polar angle goes from 12.4° to 155.1°, with two gaps between each of the endcaps and the barrel: from 31.4° to 32.2° and from 128.7° to 130.7°. The total solid angle coverage in the CM-frame is 90 %. Due to the collision geometry, we expect a much higher background rate in the forward region of the detector than in the backward region.

The coordinate system of Belle II is defined by the z-axis going along the direction of the electron beam. This axis is also referred to as the solenoid axis. The y-axis points upwards, and the x-axis points towards the outside of the accelerator tunnel. The respective directions, the positron beam (red) and the electron beam (blue), are included in Figure 3.2.2.

The track finding at Belle II is done by a combination of different algorithms [9]. Those rely on information from the PXD, the SVD, and the CDC. The track-reconstruction software at Belle II is based on several independent algorithms, which are used to process the measurements from the different detectors individually. These individual algorithms allow us to consider the different specific properties of the individual detectors. Another algorithm is then used to integrate all the different results into a single final set of tracks. The output of the software is just one set of tracks, based on the individual measurements recorded by the detectors used for track reconstruction.

As the Belle II detector is a HEP experiment, one characteristic is the large amount of data gained within very small time frames. Therefore, it is necessary to have a trigger system that reduces the number of recorded events and, consequently, the amount of offline saved data. To decide whether an event is stored or not, the events of interest must be identified and classified by using a binary logic. The binary logic assigns an event either 0 or 1, where 0 denotes that the event is discarded, and 1 denotes that the event of interest is triggered and kept in the dataset.

At Belle II, this process is performed by a two-level system consisting of the **L1 Trigger (L1)** and the **High Level Trigger (HLT)**. The L1 is a hardware-based low-level trigger, which is limited by the read-in rate of the Data Acquisition system. It mainly suppresses the beam background, originating from outside the interaction point region. The L1 generates a trigger signal for the events of interest using readout data with low resolution coming from CDC, ECL, and KLM. The trigger decisions come with a delay of $4\,\mu s$, so nearly in real time. To allow for a separation between different decay channels, various trigger bits exist.

One decay channel, which is significantly reduced by the application of the L1, is the Bhabha scattering, as the cross section of Bhabha events is significantly higher than the cross section of other, physically more interesting, decays. The cross section of Bhabha events is $\sigma_{ee} = 295\,\text{nb}$ and the cross sections of the background processes are smaller than $39.74\,\text{nb}$ [2], this is the cross section for $e^+e^- \to e^+e^-e^+e^-$.

Further, the Bhabha scattering is a well-understood process and therefore less interesting to study, but it is a useful process to perform calibrations of the detectors. To be able to study the more interesting processes, the L1 trigger bit `bha_veto` uses a pre-scaling factor of 100, which reduces the amount of Bhabha events down to $1\,\%$. The various L1 trigger bits serve different purposes and have different pre-scaling factors.

The second level of the trigger system is the HLT, which is a software-based high-level trigger. Here, the computing cluster receives the full raw datasets from the different subdetectors for all events triggered by the L1. At the HLT level, a full event reconstruction is carried out, and the binary classification of events is done based on this reconstruction. In total, the HLT reduces the amount of data by at least $60\,\%$. For the HLT, different trigger lines, also referred to as filters, exist. Each trigger line consists of several constraints, like limiting the energies or the angles of reconstructed particles. The L1 bits and the HLT line are chosen based on the goal of the performed analysis.

3.3 Data at Belle II

For high-energy physics experiments it is quite common to simulate events and their behavior within a virtual detector. The creation of those simulated events is done by using the Monte Carlo method, therefore the simulations are often just referred to as Monte Carlo simulation (MC). The simulations itself are performed by using different physics simulation tools.

At Belle II, the first step is the generation of the events by the EVTGEN generator. This generator simulates the collision of the electrons at a given resonance energy, at Belle II: $e^+e^- \to \Upsilon(4S)$. The different decay channels for the final states are simulated by various generators, e.g., tauons are simulated by TAUOLA or leptons are simulated by MADGRAPH5. Further important is the PHOTOS generator, accounting for the electromagnetic final-state radiation. Finally, for all events, the response of the particles traveling through the Belle II detector is simulated by using GEANT4. The production of MC simulations is improved continuously. In this thesis, the newest available MC files from the MC16 campaign are used.

In contrast to the data taken at the experiment, the simulation of events is done separately for each decay channel. Where a generic simulation is done based on basic Standard Model processes. This means that, depending on the decay to be analyzed, a different MC file is used as a signal file. All other simulated decay channels are then referred to as background. Those background files are normally added at a later point during the analysis, as it is easier to classify the correct decay channel, using only signal MC. And it is assumed that from the background decays only very few events are reconstructed.

The decay reconstruction and the analysis of the different output variables are done using the Belle II Analysis Software Framework (basf2) [10]. This software framework is developed for analyses carried out at Belle II. It includes various important features and parameters, which are quite specific for the analysis of data taken at Belle II. As the software framework is under continuous development, several versions exist. For this thesis, the light release „light-2505-deimos“ is used.

One feature included in basf2 is that for events in simulated data, the MC truth matching can be performed by using the generator-level information of the simulated events. In this procedure, it is attempted to match each reconstructed particle, consisting of either only a cluster or if it is a charged particle of a cluster and a track, to a particle produced on the generator-level. Using the known properties of the generated particles, the topology of reconstructed events, as well as the misidentification of reconstructed particles, can be understood in greater detail, [9].

The MC matching returns two variables, the „mcPDG“ and the „mcErrors“. The MC PDG code describes the particle species of the particle on the generator-level matched to the reconstructed particle. Each particle species has a unique PDG code value. The mcErrors variable describes the

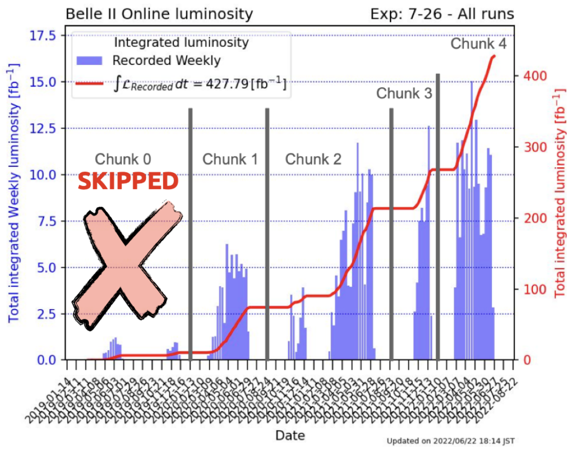

Figure 3.3.1: Total integrated luminosity and weekly recorded integrated luminosity for Belle II from 2019 to 2022, divided into different chunks of data, including experiments 7 to 26 [11].

correctness of the MC matching. Correctly reconstructed particles have the mcErrors value 0, meaning the reconstructed particle is of the same particle species as the particle on generator-level. Reconstructed particles, which got assigned a different particle species than the particle on generator-level, get an mcErrors value that is larger than 0.

All values of the mcErrors refer to the reason that caused the wrong match between data and MC, a detailed list of the different reasons for mismatching is found here [10]. It is important to note that the MC error code also reflects a combination of several issues. In this case, the value is a sum of the error values of the different issues. This sum is always unique, as the MC error values are assigned in such a way that all summed values are still individually reflecting only one error scenario.

When the error code is unequal to 0, the MC PDG code reflects the particle species of the generator-level particle. From the MC simulations and the truth-matching, other variables, including generator information about each matched particle, can also be accessed. Those variables are, e.g., the MC energy, the MC momentum, and the angles θ_{MC} and ϕ_{MC}. These variables are accessed after performing MC truth matching and are then used for comparisons between the reconstruction-level and generator-level.

At the Belle II experiment, data and MC simulations are grouped hierarchically. The individual events are first grouped into runs. Between runs, e.g., the beam conditions can be changed. One run is described as a period of continuous data-taking. Further, it is important to note that the MC simulations used here are computed as run-dependent data, which means the differences between the different experimental runs are taken into account when producing simulations.

Several runs during a specific data-taking period are then combined within one experiment. During one experiment, the run conditions stay the same. Finally, for easier handling of the data, several experiments are combined within one chunk. The different chunks recorded at Belle II, and the integrated luminosities of the different chunks are taken from Figure 3.3.1.

Chapter 4

Decay reconstruction

H. Haigh's study from 2023 [12] determined the tracking efficiency for high-momentum particles at Belle II. This study analyzed radiative Bhabha events $e^+e^- \to e^+e^-\gamma$ and used the tag-probe method for the event reconstruction. The presence of an electron and a photon tags the events. The probe consists of the second electron, where the electron is used as a synonym for both charges. If the track reconstruction for the probe was not successful, the probe is a reconstructed photon. In this case, this study treats the photon as a mis-reconstructed electron from the Bhabha event. This assumption seems odd, as all probe particles are treated as charged particles regardless of their actual charge.

Further interesting in this approach by Haigh is that following the selection criteria, the cluster energy of the probe particle has to be larger than the cluster energy of the tag photon. This is an important point, as for the calculation of the tracking efficiency, the hypothesis of the probe electron being more energetic than the probe photon is used.

The tracking efficiency in Haigh's analysis is then calculated using a two-dimensional binning, based on polar angle and transverse momentum. For the tracking efficiency itself, the number of correctly reconstructed tracks is compared to the total number of reconstructed particles. The calculation of the tracking efficiency and its uncertainty is done for data and simulation, so that both values are compared to each other.

An important detail about the data-MC study by Haigh is that there are data-MC discrepancies in the dependence of the charge, implying that a charge-dependent study leads to an improved final result. Finally, Haigh's study only provides preliminary results as the final evaluation of the uncertainties is still missing.

This previous study serves as a starting point for a new calculation of the tracking efficiency within this thesis. The first step of our analysis is to perform a similar decay reconstruction and to identify the fraction of events where the charged particles are more energetic than the neutral ones. Most importantly, we will continue using a phase-space binning and add charge dependence to the calculation.

4.1 Particle based reconstruction

The initial attempt to reconstruct the necessary radiative Bhabha decays is described as **particle based**. We use this approach and most of the selection criteria from H. Haigh's analysis, only the criterion that the energy of the charged particle is larger than that of the neutral particle is not used. The goal of our decay reconstruction is to quantify the fraction of events with the charged probe particle being more energetic than the probe photon.

We use the tag-probe method with the tag particle being an electron, which is extracted by applying several constraints. The probe side consists of the other electron and the photon. Table 4.1.1 lists all constraints used for the selection of the events.

The event-based criteria are that the event needs to have at least one good track, but not more than two good tracks. The total number of clusters within the ECL has to be exactly three, and the sum of the energies of these three clusters has to be within the range of $E_{\text{sum}} = (10 - 11)\,\text{GeV}$.

A good track is defined as a track originating directly from the interaction point (IP) region. This is assured by the IP-selection criterion [10]. This defines a cylinder around the actual IP of the beams. The size of this cylinder is given in Table 4.1.1 as IP-selection. All tracks originating in this region are treated as particles coming from the interaction of the colliding particles.

The tag side has very strict constraints to ensure that the tag particle is an electron. The charge of the tag particle has to be ± 1, the transverse momentum has to be larger than $0.3\,\text{GeV}$, and the electron cluster needs to be associated with a good track. Additionally, the ratio of the cluster energy of the tag electron over its momentum has to be within the range of 0.95 to 1.05. This parameter ensures that the reconstructed particle deposits all its energy in the ECL and that it can be treated as an electron.

The particles used for the probe side have much looser constraints. Most importantly, the probe electron must not be the same particle as the tag electron, and its cluster energy has to be larger than $0.3\,\text{GeV}$. The photon must only satisfy the latter condition.

Using the current reconstruction logic, only correct radiative Bhabha scattering events are reconstructed. However, for the determination of the tracking efficiency, not only the perfectly reconstructed radiative Bhabha events are of interest, but also reconstructed decays with three charged particles or two photons.

Of course, decays with an odd number of charged final-state particles do not directly satisfy charge conservation, but as electrons are particles with a comparably low mass, $m_{\text{e}} = 511\,\text{MeV}$ [4], and the range of the total reconstructed energy is $E_{\text{sum}} = (10 - 11)\,\text{GeV}$. The assumption of an additional existing, but unreconstructed, electron can be made, and with this extra electron, the charge conservation would be satisfied.

Using the described selection criteria for reconstructing the radiative

Cut	Value
Event based	
Number of good Tracks	1 or 2
Sum of all ECL cluster energies (CM frame)	[10,11] GeV
Tag track definition for electron	
IP-selection	dr <0.5 cm & $\|dz\|$ >2 cm
charge	± 1
Transverse momentum	$p_{\mathrm{t}} > 0.3\,\mathrm{GeV}$
Cluster energy over momentum ratio	$0,95 < E_{\mathrm{cl}}/p < 1,05$
Probe track definition for electron	
is part of tag track list	False
Cluster energy	$E_{\mathrm{cl}} > 0.3\,\mathrm{GeV}$
Photon list	
Cluster energy	$E_{\mathrm{cl}} > 0.3\,\mathrm{GeV}$

Table 4.1.1: Event selection criteria which are based on the analysis by H.Haigh [13]. Electrons and positrons are referred to interchangeably as electrons.

Bhabha decay ing MC leads to 16 % of events out of all generated $ee(\gamma)$-events, where a tag e^- and a probe e^+ exist, and only 5 % of all generated events have an additional correctly reconstructed photon γ.

The percentage of reconstructed radiative Bhabha events seems small, but the amount of correctly reconstructed decays is reasonable. This is checked by using the generator-level information. The MC dataset is dominated by non-radiative Bhabha events, and it contains only a very small number of events with one or more photons on the generator level. From the analysis of the statistics on the generator level, I conclude that a large number of produced MC particles, independent of the particle species, would hit the ECL and the CDC, outside their acceptance ranges, and are therefore not reconstructed by basf2. Further, there is a larger amount of photons created, but most of the photons are low energetic. Where the resolution of the detectors only allows for a correct reconstruction of photons with higher energies, $E_\gamma > 0.3\,\mathrm{GeV}$. These reasons explain the relatively low number of reconstructed events, compared to the number of generated events.

Using the events reconstructed with the particle based approach, the fractions of events where probed electrons have higher energy than the probed photons in the event are extracted. This is done by plotting the ECL cluster energy of the photons on the vertical axis and the ECL cluster energy of the probe electrons on the horizontal axis, as in Figure 4.1.1. The events in Figure 4.1.1 are all by MC truth-matching correctly reconstructed radiative Bhabha events.

The red line separates the events into two different categories: the events where the positron has a higher cluster energy and the events where the photon has a higher cluster energy. In relative numbers, the events

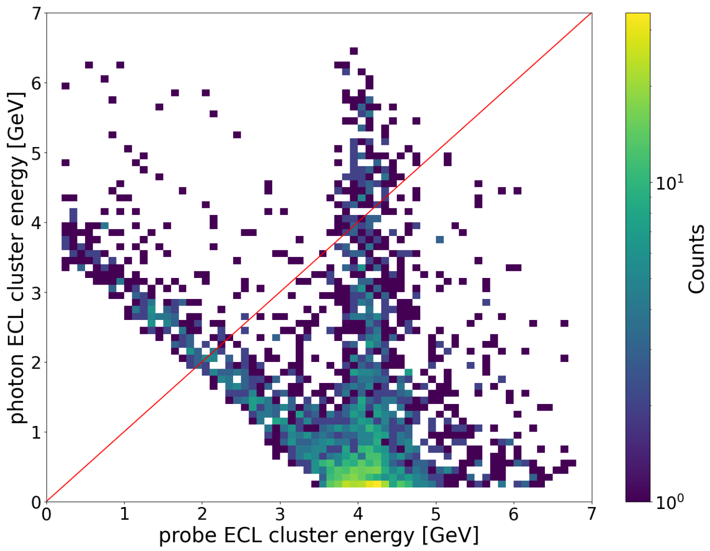

Figure 4.1.1: The cluster energy of the probed electron against the cluster energy of the probe photon, both in the lab frame, extracted by using an electron tag, and all events are by MC truth correctly reconstructed radiative Bhabha events.

where the positron is the more energetic particle make up 82.76 % of all events, while in 17.24 % of cases, the photon has more energy. As the percentage of events where photons have a higher energy is significant, it would not make much sense to assume that the more energetic cluster originates from an electron. This would lead to a dataset with very low purity and accuracy when assigning the track to the more energetic particle.

From this first analysis step, no reasonable conclusion about the ECL cluster, originating from a charged particle, can be made, and no acceptable purity of the data sample is reached. Therefore, a different approach for the decay reconstruction is described in the following.

4.2 Cluster based reconstruction

The more general approach to the reconstruction of the radiative Bhabha decay is called the **cluster based reconstruction**. For this reconstruction, only ECL clusters are used for the decay reconstruction, independent of whether a track is associated with the cluster or not. With this approach, the different decay modes, described in the previous Section 4.1, are reconstructed simultaneously, and a more particle-independent calculation of the tracking efficiency is possible.

The reconstruction is done by using the tag-probe method. The tag particle remains the electron with the constraints from Table 4.1.1 applying. The probe side consists of the two most energetic ECL clusters, in addition to the cluster from the tag particle. C_{most} describes the cluster with the most ECL energy, excluding the cluster of the tag particle. C_{sec} describes

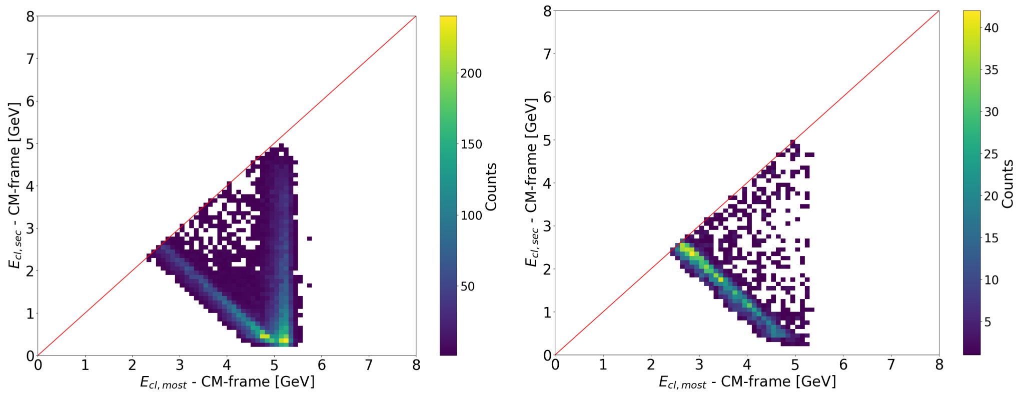

Figure 4.2.1: Cluster energy of the most energetic cluster versus cluster energy of the second most energetic cluster, both in the CM-frame reconstructed using an electron tag for events categorized as case A (left) and case B (right).

the cluster with less ECL energy than $\mathrm{C_{most}}$

$$e^+e^- \to e^\pm \, \mathrm{C_{most}} \, \mathrm{C_{sec}}. \tag{4.2.1}$$

Assuming that each ECL cluster can originate from either charged electrons or neutral photons. This decay reconstruction logic leads to four different combinations of particles within the reconstructed decay. The combinations are correct radiative Bhabha scattering events with the charged electron creating $\mathrm{C_{most}}$ or $\mathrm{C_{sec}}$. Further, we have the $e^+e^- \to e^+\gamma\gamma$ and the $e^+e^- \to e^+e^\mp e^\pm$ decays as combinations C and D. Table 4.2.1 lists all different combinations for an electron tag. For the positron tag, the combinations are the same, just the signs have to be flipped. The considered cases in Table 4.2.1 are all based on particles, which have the MC error flag 0. This is further referred to as correctly reconstructed. For further analysis in this chapter, the reconstruction with an electron tag is used.

At first, each case is analyzed quantitatively and qualitatively. This is done for a dataset containing only $ee(\gamma)$-signal events and for a second dataset including additionally all luminosity-scaled background processes.

To make a qualitative statement about the distribution of the events for the different cases, the cluster energy of the most energetic cluster is

Case	**tag**	$\mathrm{C_{most}}$	$\mathrm{C_{sec}}$
A	e^-	e^+	γ
B	e^-	γ	e^+
C	e^-	γ	γ
D	e^-	$e^\mp$	$e^\pm$

Table 4.2.1: Decay reconstruction cases based on particles correctly matched by MC truth for an electron tag.

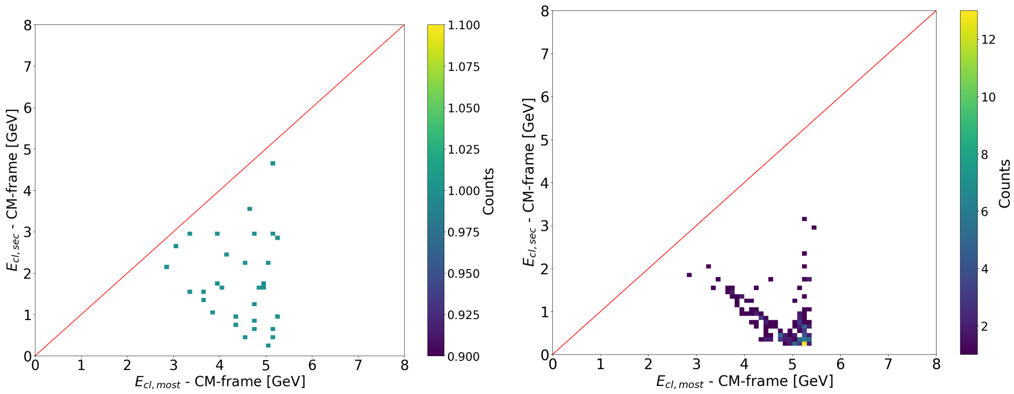

Figure 4.2.2: Cluster energy of the most energetic cluster versus cluster energy of the second most energetic cluster, both in the CM-frame reconstructed using an electron tag for events categorized as case C (left) and case D (right).

plotted versus the cluster energy of the second most energetic cluster, both in the CM-frame. For case A, the two main regions include the majority of events. One region at $E_{\text{cl,most}} = 5\,\text{GeV}$ includes events where the energy of the probe electron stays constant, and the energy of the photon varies. The second region includes events where the energy of the probe electron and photon change, while the energy of the tag electron stays the same, this is the diagonal between $E = 5\,\text{GeV}$, of both clusters visualized in the left panel of Figure 4.2.1. For case B, the main part of the events has a similar tag electron energy, while the energies of the probe electron and the photon vary, which are the events located on the diagonal in the right panel of Figure 4.2.1. The number of events categorized as cases C and D is quite small, but still, a similar trend for the distribution of the events is identified, as shown in Figure 4.2.2. I conclude that one of the reconstructed electrons always has an energy of 5 GeV in the CM-frame, while the rest of the energy is split between the second electron and the photon.

A quantitative analysis of the number of events for a dataset including luminosity-scaled background events is performed. The first important thing to notice is that the decay reconstruction algorithm only reconstructs a very small amount of background events, the fraction of non $ee(\gamma)$-signal events is 0.52 %. Further, I conclude that the events from the background processes mainly contribute to case B. This observation is quite helpful, as it means that, when considering case A, where the most energetic cluster is created by the charged particle, the contributions from background processes are negligible. Further, it is interesting that the majority of the background events come from $e^+e^- \to \gamma\gamma$ events. For those events, we assume that one of the photons splits into two electrons while flying through the detector. The full quantitative analysis of the dataset, including luminosity-scaled background processes, and the dataset containing only $ee(\gamma)$-signal is available in Table 4.2.2.

From Table 4.2.2 I extract that the percentage of events where both

dataset	parameters	total	case A	case B	case C	case D
only ee	N_{events}	58.368	17.900	2.593	31	156
	fraction [%]	100	30,67	4,44	0,05	0,27
Including background	N_{events}	58.672,2	17.900	2.871,4	33,9	156,1
	fraction [%]	100	30,51	4,89	0,06	0,27

Table 4.2.2: Number of events and their fraction from the total number of reconstructed events N_{events} for a dataset consisting of only ee processes and for a dataset including luminosity-scaled background processes.

energy clusters are correctly matched by MC truth is only 35 % of all reconstructed events. Meaning that, the larger part of reconstructed events contains candidates that could not be matched correctly to generator-level particles.

For the most energetic cluster, 86 % of the particle candidates are correctly matched, and the majority of the correctly matched particles are positrons: 95 %. The second most energetic cluster includes 47 % correctly matched candidates, with one third of those being positrons and the rest photons. For the second most energetic cluster 22.7 % of the particle candidates are misidentified final-state charged particles. This MC error flag characterizes, in our case, clusters probably affected by cluster splitting. In this thesis, these particle candidates are further referred to as cluster splitting candidates. In more detail, this error flag says that a final state charged particle is misidentified. Further, 29.5 % of the candidates of the second most energetic cluster have the error flag 512, corresponding to reconstructed clusters which could not be matched to any generator-level particles, indicating fake or background tracks or clusters.

From these numbers, I extract that a significant amount of events do not have three correctly reconstructed particles, and the events and particle candidates that could not be matched correctly by MC truth have to be analyzed with the goal of finding a possibility to include those events in the final dataset or to be able to remove those events in an efficient way.

4.3 Analysis of the key issues

When analyzing the reconstructed events, two key issues are encountered. The first one is the misidentification of charged final-state particles. The other one is the reconstruction of a large number of fake photon clusters. Both issues were identified in the particle-based and the cluster-based reconstruction.

Both approaches for the decay reconstruction provide distinct views of the same problems. As the cluster-based approach is used to carry out the final analysis, it is also used to classify the key issues. For this analysis, the information gained by MC truth matching plays an important role, as it gives the possibility to compare the variables of reconstructed and

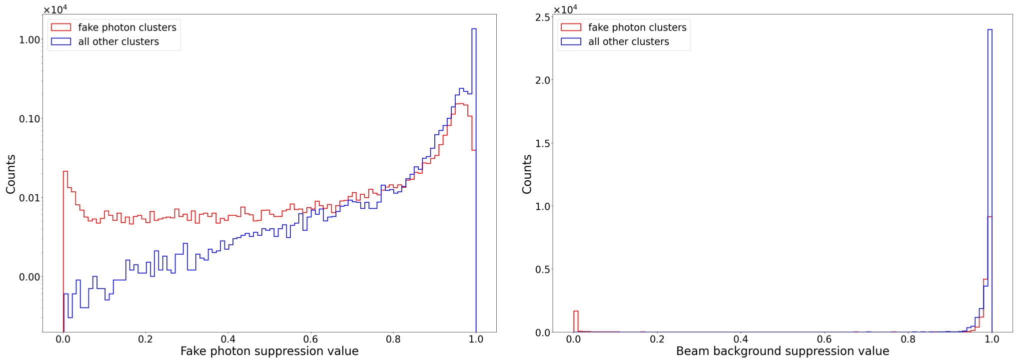

Figure 4.3.1: The fake photon suppression variable on the left and beam background suppression variable on the right, both for the second most energetic cluster.

generated particles.

The first issue to address are the candidates having the error flag 512 „No valid match was found“. This error flag indicates fake or background tracks and clusters [10]. Events with this error flag are further referred to as fake photon clusters. The large amount of unmatched candidates is mainly an issue for the second most energetic cluster. As for this cluster 29 % of the events are fake photon clusters. From the analysis of the MC truth information, I extract that most of the candidates indicating fake photon clusters are clusters reconstructed to be photons, but those photon candidates do not have any matching particle on the generator level. For the identification of fake photons and beam background clusters, two separate variables are available in basf2 [10].

The „beamBackgroundSuppression“ variable assigns each photon cluster a value between 0 and 1. Values close to 1 indicate a true photon cluster, while values close to 0 indicate a beam background cluster. A beam background cluster comes from photons emitted by the beam electrons. The „fakePhotonSuppression“ variable works in the same way: values near 1 indicate a true photon cluster and values near 0 indicate a fake photon cluster. As all float values between 0 and 1 are possible, the output reflects how strongly a cluster is consistent with a true photon cluster, with beam background, or fake photons.

Analyzing the beam background suppression variable in the right panel of Figure 4.3.1, the candidates categorized as fake photon clusters have a smaller, narrower peak at 0 and a narrow but larger peak at 1. All other candidates only have a peak at 1. For the fake photon suppression, the values for the fake photon clusters range between 0 and 1. All other clusters show a similar distribution, but with a clear peak at one, shown in the left panel of Figure 4.3.1.

From these observations, we conclude that the suppression variables help to separate fake photon clusters from true photon clusters, but not necessarily all clusters categorized by MC truth as fake photon clusters are identified correctly by using these two variables. Quantitatively, 34 %

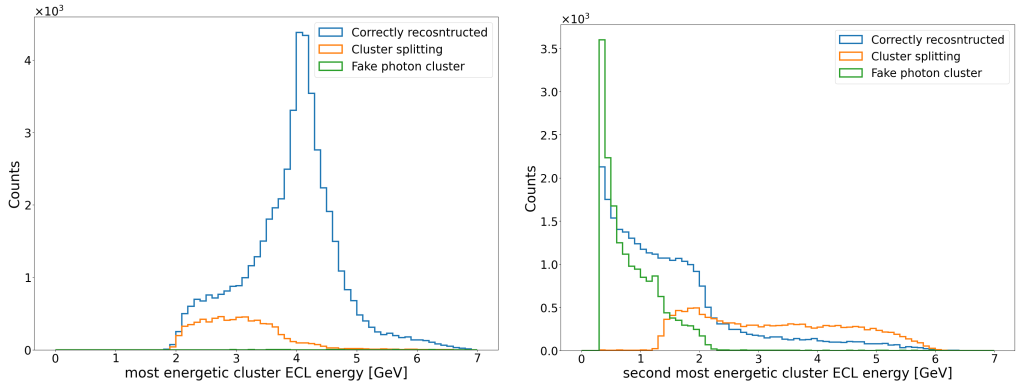

Figure 4.3.2: The ECL cluster energy plotted for different MC error flag values on the left for the most energetic cluster and on the right for the second most energetic cluster. All events are reconstructed by using an electron tag.

of the candidates reconstructed as fake photon clusters are removed when both suppression values are larger than 0.8.

Further interesting for fake photon clusters is the ECL cluster energy, shown in Figure 4.3.2. For the second most energetic cluster, I find that the majority of the fake photon clusters (green) have an energy below 1 GeV but also correctly reconstructed clusters (blue) peak at low energies, which makes it impossible to remove events involving fake photon clusters by setting a higher energy threshold without removing too many correct candidates. In the right plot of Figure 4.3.2, the ECL cluster energy distributions of the second most energetic cluster are plotted on the horizontal axis and the number of counts per bin on the vertical axis. If a cut on the energy threshold will be necessary for the final analysis, will be decided after the effect of the trigger selection on the dataset composition is analyzed.

The second issue to analyze are the reconstructed particle candidates classified by the MC error code 128: „One of the charged final state particles is misidentified“ [10]. This means that the reconstruction algorithm reconstructs photons, but on the MC truth level, those particles are actually electrons or other charged particles. Candidates with this error flag are referred to in this thesis as cluster splitting candidates. To classify this error and identify its cause, various parameters are analyzed.

The cluster splitting candidates are separated into two groups, depending on the PDG code of the matched generator-level particles. The reconstructed clusters, assumed to be cluster splitting candidates, either correspond to an electron (PDG code = 11) or to a positron (PDG code = −11). For the most energetic cluster 9.4 % of all events are probably affected by cluster splitting. But the mentioned separation is not relevant for this cluster, as less than 0.1 % of the events are associated with an electron, the rest are associated with a positron. For the second most energetic cluster, more than 16 % of all candidates are affected by the phenomena of cluster splitting. Here, the differentiation on the MC truth level between

	parameters	Error 128	Err. 128 & e^-	Err. 128 & e^+
C_{most}	N_{events}	7649	6	7643
	percentage [%]	9,42	0,0074	9,41
C_{sec}	N_{events}	13254	9739	3505
	percentage [%]	16,32	11,99	4,32

Table 4.3.1: The number of events and their fractions categorized by the different MC error flags for both clusters, with a total number of 81.189 events reconstructed from an electron tag. Positrons have the PDG code −11 and electrons have the PDG code 11.

electron and positron has to be made to be able to consider different effects. Table 4.3.1 gives the specific numbers and fractions.

Using the previously discussed distribution of the ECL cluster energy, for the most energetic cluster, the left plot in Figure 4.3.2 is of interest. The energies of candidates affected by cluster splitting (orange) range mainly between 2 GeV and 4 GeV. Here, clearly, the covered energy range overlaps with the energy range of the correctly reconstructed particles (blue), making an energy cut useless. For the second most energetic cluster (right panel), the energy distributions are similar. The correctly reconstructed events (blue) peak at energies below 2 GeV and the cluster splitting candidates (orange) are equally distributed between 1 GeV and 6 GeV. But there is still a major overlap in the energies, besides the peak of the correctly reconstructed ones.

From these observations, I conclude that setting a higher threshold on the ECL cluster energy threshold would remove events affected by cluster splitting, but at the same time, it would lead to an overall reduction of correctly reconstructed events, which is something we want to avoid due to the limited amount of available simulated events.

As from the energy distribution in dependence on the MC error flag, no useful conclusions were drawn, we now compare the reconstructed cluster energies to the generated energies, where the generated energies are taken from the generator-level particles matched to the reconstructed particles.

For correctly reconstructed particles of the most energetic cluster, most of the particle candidates have nearly the same value for the reconstructed cluster energy and the generated energy. Those candidates are located on the diagonal between the energies, which can be seen in the left panel of Figure 4.3.3. Further interesting is that a significant amount of events is located at $E_{\text{MC, most}} = 4.2\,\text{GeV}$. For those candidates, the reconstructed cluster energy is smaller than the MC energy. This missing reconstructed energy indicates that these particle candidates split their energy between several clusters. The energy comparison for cluster splitting candidates yields a similar conclusion. Their generated energy is mainly 4.2 GeV, and the reconstructed energy is smaller than the generated energy but always larger than 2 GeV. The allocation of candidates around $E_{\text{MC}} = 4.2\,\text{GeV}$ is interesting and could probably come from the fact that the positron beam

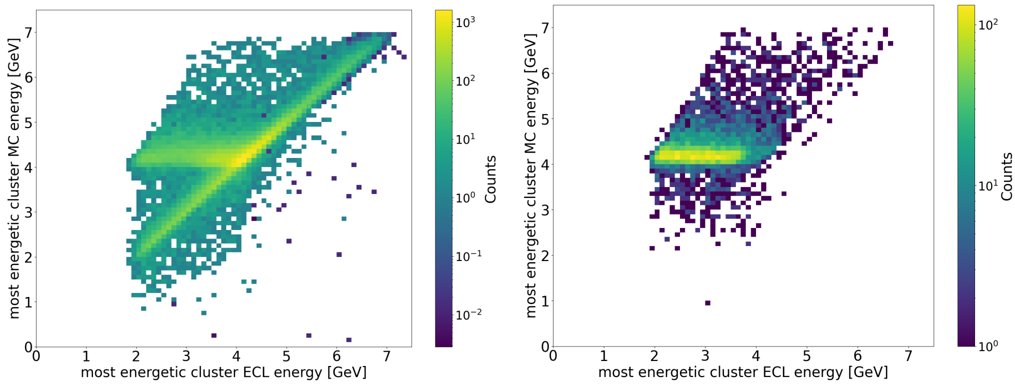

Figure 4.3.3: The cluster energy versus the generator-level energy of the most energetic cluster, both in the lab frame and for events reconstructed using an electron tag. The distribution for correctly matched candidates (left) and the candidates encountering cluster splitting (right).

has an energy of 4 GeV. But due to the limited timeframe of this thesis, it is not further investigated why this specific energy value occurs more often than other energy values.

The same comparison is also done for the candidates of the second most energetic cluster. Most of the correctly reconstructed candidates have the same value for reconstructed and generated energy, as it is expected. However, there are correctly reconstructed particle candidates with $E_{\mathrm{MC}} = 4.2\,\mathrm{GeV}$ but with a lower reconstructed cluster energy. In comparison to the most energetic cluster, the candidates for the second most energetic cluster have mainly reconstructed cluster energies below 2 GeV. This is shown in the left plot of Figure 4.3.4, the right plot visualizes the energy comparison for all candidates of the second most energetic cluster affected by cluster splitting, independent of the charge of the associated generator-level particle. Important for those candidates is that none of them has equal reconstructed and generated energy, but they are separated into two structures. One group has an energy on the generator-level of 4.2 GeV and reconstructed energies of 1 GeV to 2 GeV. Those candidates are matched to positrons on the generator-level. The other group has a generated energy of 6.8 GeV and their reconstructed energies range from 1 GeV up to 6.8 GeV. The events here are matched by MC truth to generator-level electrons.

From the energy comparison, we conclude that for a majority of candidates correctly matched to a generator-level particle, the reconstructed energy is similar or equal to the MC energy. But there is also a significant amount of candidates with a lower reconstructed energy than expected, both for the correctly reconstructed candidates and for those candidates indicating cluster splitting. This leads to the assumption that the reconstructed clusters do not represent the full particle hit and that maybe a combination of several clusters is necessary to analyze and reconstruct the particles correctly. This observation gives a first indication that the phenomenon of cluster splitting could be the root of those wrongly matched

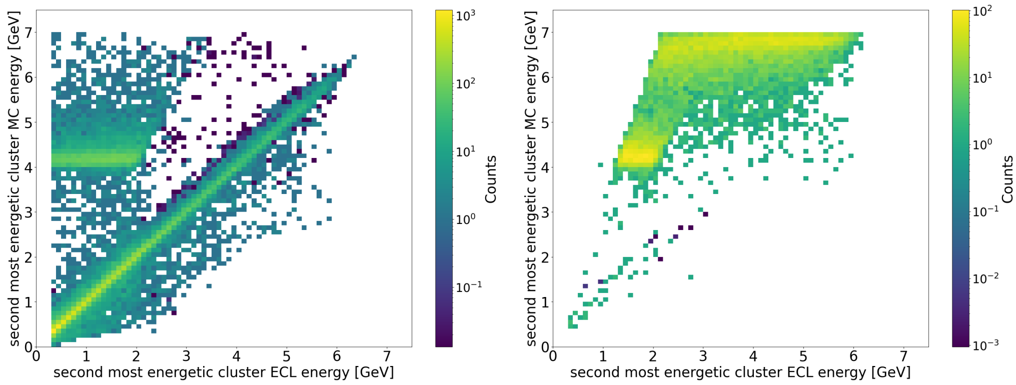

Figure 4.3.4: The cluster energy versus the generator-level energy of the second most energetic cluster, both in the lab frame and for events reconstructed using an electron tag. The distribution for correctly matched candidates (left) and the candidates encountering cluster splitting (right).

candidates. This means that one generator-level particle is reconstructed by two separate clusters, and a combination of the two corresponding clusters leads to the same parameters as those of the generator-level particle.

To validate this assumption, the opening angle α between the different clusters is determined. With this variable, we identify the relative location of the clusters to each other within the ECL.

The determination of α is done by using the vectors from the IP to the considered clusters. The vector of a single cluster is determined by using its ECL polar angle and azimuth:

$$\vec{r} = \begin{pmatrix} r\sin\theta_{\text{cl}}\cos\phi_{\text{cl}} \\ r\sin\theta_{\text{cl}}\sin\phi_{\text{cl}} \\ r\cos\theta_{\text{cl}} \end{pmatrix} . \tag{4.3.1}$$

With this vector definition, the dot product of the two clusters is used to determine the angle between them

$$\cos\alpha = \frac{\vec{r_1}\cdot\vec{r_2}}{|\vec{r_1}|\cdot|\vec{r_2}|} = \hat{r_1}\cdot\hat{r_2}. \tag{4.3.2}$$

The radius is set to one for this analysis, as only the relative location, and not the absolute difference between clusters, is relevant. The modified expression for the opening angle is:

$$\alpha = \arccos(\sin\theta_{\text{cl1}}\sin\theta_{\text{cl2}}(\cos\phi_{\text{cl1}}\cos\phi_{\text{cl2}} + \sin\phi_{\text{cl1}}\sin\phi_{\text{cl2}}) + \cos\theta_{\text{cl1}}\cos\theta_{\text{cl2}}) . \tag{4.3.3}$$

Using Equation (4.3.3), the opening angle is determined between the most and the second most energetic cluster, the most energetic cluster and the cluster of the tag electron, and between the second most energetic cluster and the cluster of the tag electron. For the analysis of the different opening angles, the events are split into two categories: those probably affected by cluster splitting and all other events. This helps us to identify trends for

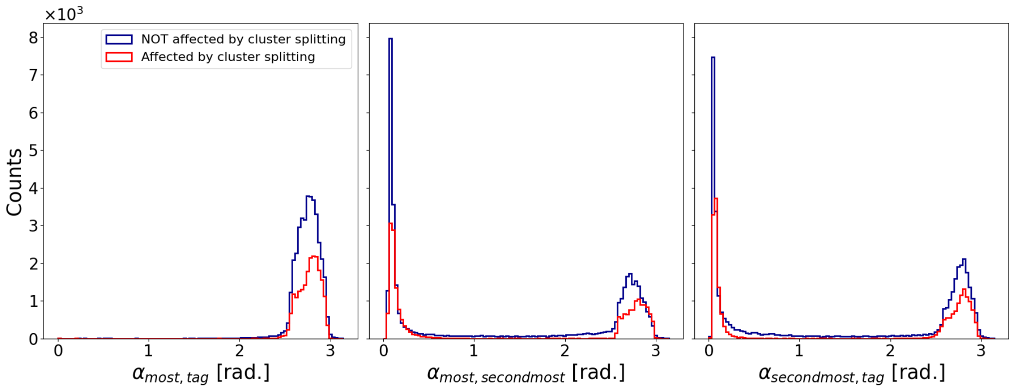

Figure 4.3.5: Opening angles between the different clusters, for a reconstruction using an electron tag in the lab frame.

the different types of events. The events affected by cluster splitting include events where either the most or the second most energetic cluster is a cluster splitting candidate.

For the opening angle between the most energetic cluster and the tag electron cluster, the majority of the events have an opening angle larger than 2.4 rad. This indicates that the most energetic cluster is always well separated from the cluster of the tag electron.

This does not yield for the opening angle between the most and the second most energetic cluster, and the opening angle between the second most energetic cluster and the cluster of the tag electron. For those opening angles, the separation between cluster splitting affected events and other events does not lead to any promising conclusions about a different behavior of those types of events. But I extract that, in addition to the peak at more than 2.4 rad, both opening angles have a large peak at 0.1 rad. All distributions are shown in Figure 4.3.5.

The small opening angles indicate that the two analyzed clusters are very close to each other within the detector, while larger opening angles tell us that the clusters are nicely separated. From a more detailed analysis, I extract that for each event affected by cluster splitting, the second most energetic cluster is located close to one of the other clusters within the event. The close location of two clusters within the ECL indicates that two clusters could originate from the same generator-level particle. However, based on the results from the opening angles, it is not fully proven that the events are actually affected by cluster splitting.

To further verify the assumption of cluster splitting, we use the variable „genParticleID“. This variable gives the array index of the related MC particle, [10]. The array index is a unique integer value assigned to every particle on generator-level, starting with the collision particles e^- and e^+ having the indices 0 and 1. All the collision products are then assigned the indices starting from 2. This variable is referred to as the particle ID. When two particle candidates have the same particle ID, it means that these two reconstructed particle candidates are associated with the same generator-level particle. If two candidates have different values for the particle ID, they are associated with different generator-level particles.

For events without candidates affected by cluster splitting less than 0.2 % share the same particle ID. Analyzing the events where one cluster indicates cluster splitting, less than 0.5 % of those events have different particle IDs. This leads to the conclusion that the majority of events having a cluster with an error flag indicating cluster splitting really are affected by cluster splitting.

Following all these different analysis steps, we check if combining the cluster energies of the clusters that encounter cluster splitting leads to the energy of the associated generator-level particles. For this analysis step, the two clusters corresponding to each other are identified by having the same assigned PDG value, and one of them has an error flag indicating cluster splitting.

For this analysis, only the events where the second most energetic cluster indicates cluster splitting are used, as they make up a majority of the events, but it works in the same way for the most energetic cluster. The reconstructed cluster energies of two associated clusters are added, where the associated cluster is either the most energetic cluster or the cluster of the tag electron. The summed reconstructed cluster energy is then compared to the MC energy of the particle associated with the most energetic cluster or the tag electron cluster.

Comparing eventwise the summed cluster energy to the MC energy leads, for a majority of the analyzed events, to similar values for both these energy values. The exact distributions are shown in Figure 4.3.6. In those plots, on the horizontal axis, the respective sum of the reconstructed cluster energies and on the vertical axis, the energy of the associated generator-level particle is plotted. This is done separately for events, where the most energetic cluster splits (left) and events where the tag electron cluster splits (right). The red line visualizes where the summed energy equals the MC energy.

Analyzing the difference between the summed energy and the MC energy, for both cases, for 60 % of the events affected by cluster splitting, the summed cluster energy and the MC energy differ by a maximum of 0.1 GeV.

From these two plots, I finally conclude that the candidates with an error flag indicating cluster splitting are not only misidentified final-state charged particles, but they are not even independent particles. The reconstructed decays, where one candidate indicates cluster splitting, are therefore only Bhabha decays, not radiative Bhabha decays. For determining the tracking efficiency per bin, the cluster energy measured in the ECL will be used. Therefore, we must ensure that the ECL cluster energy value accurately represents the total energy of the measured particle. In events where cluster splitting occurs, we cannot be certain that the cluster energy of a reconstructed cluster represents the total energy of a particle. Due to this, the events with cluster splitting need to be separated from other good

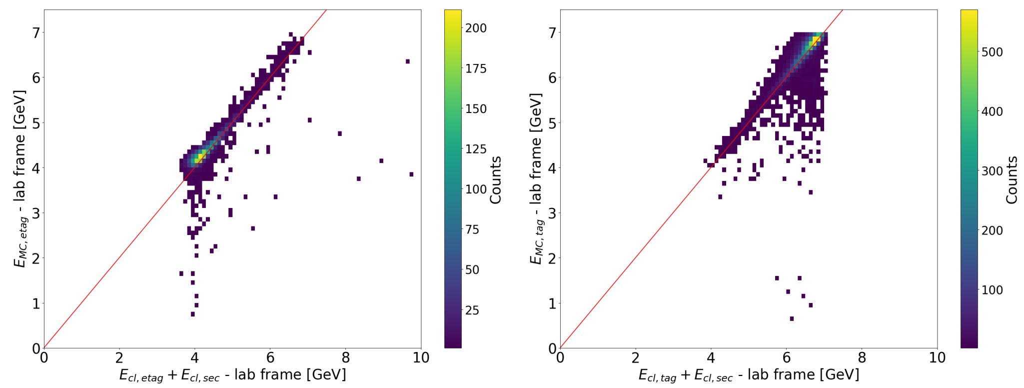

Figure 4.3.6: The summed reconstructed ECL cluster energies for the second most energetic cluster encountering cluster splitting plotted against the corresponding energy on generator level for the correctly reconstructed cluster, for the most energetic cluster being split (left), and the tag electron cluster being split (right).

events, so that the determination of the tracking efficiency is done with a very pure sample and with a good agreement between reconstructed and generated particles.

Now that it is relatively clear from where the encountered issues originate, one can use this information to define specific cuts, helping to create a very well-known and pure data sample. With the sources of the observed discrepancies now understood and the criteria for a pure event sample identified, our next step is to ensure consistency with experimental data. Since recorded datasets contain only triggered events, the corresponding trigger selections must be applied to the MC samples and evaluated within the decay reconstruction. We do this now before any other cuts are applied.

4.4 Trigger selection

Triggers must be applied to MC datasets to enable comparison with results from an experimental dataset. Those contain only events triggered by the defined trigger selection. The trigger selection is done separately for each analysis, as the configuration of the decay reconstruction and the goal of the analysis must be taken into consideration. For determining the tracking efficiency in this thesis, it is important that the trigger line does not explicitly require any tracks in the event, as this would bias the final result. Further, the performed decay reconstruction is based on information from the ECL. This aspect also needs to be covered by the chosen triggers.

The following triggers fulfill these conditions and are therefore used for our analysis: For the L1 trigger, the „hie“ trigger, and for the HLT, the trigger line: „ge1_Estargt2_GeV_neutral_clst_32130_not_gg2clst_ee1leg1clst_ee1leg1trk_eeBrem“.

The „hie“ trigger is a combination of the „ehigh“, the „!bha_veto“, and

L1 trigger bit	Definition
`ehigh`	ECL total energy in $\theta_{\mathrm{ID}} = 2 - 15$ is more than $1\,\mathrm{GeV}$.
`!bha_veto`	$160^\circ < \sum \theta_{\mathrm{CM}} < 90^\circ$ & $160^\circ < \Delta\phi_{\mathrm{CM}} < 200^\circ$ & $E_{\mathrm{CM}}^{\mathrm{low}} > 3\,\mathrm{GeV}$ & $E_{\mathrm{CM}}^{\mathrm{high}} > 4.5\,\mathrm{GeV}$
`!veto`	KEKB injection veto

Table 4.4.1: L1 trigger bit definitions [14].

the „!veto“ trigger bits. The last one is the KEKB injection veto, meaning that shortly after the injection of a new bunch into the accelerator, none of the collision events are saved. The exclamation mark in front of a trigger bit means that events that fulfill the trigger bit are rejected, instead of kept. The „!bha_veto“ ensures that events where in the CM-frame there are two high-energy clusters back-to-back in θ and ϕ are rejected. The „ehigh“ trigger bit ensures that the total ECL energy in a defined polar angle region, see Table 4.4.1, is larger than $1\,\mathrm{GeV}$.

The HLT trigger line consists of two main parts: constraints for the neutral cluster reconstructed in the event and the rejection of specific decay channels. The neutral cluster is constrained by having a cluster energy of $E_{\mathrm{CMS}} > 2\,\mathrm{GeV}$ and the polar angle being within the acceptance of the ECL: $\theta \in [32^\circ,\ 130^\circ]$ [15]. The second part of the trigger line includes different types of events, which are rejected by the trigger line. All these types of events have two particles, which are emitted back-to-back in the detector. A detailed list of the specific rejected event configurations is given in Table 4.4.2.

Applying this trigger selection to the MC data set leads to a total reduction of the number of events to only 1.1 % of the before-reconstructed events. The events surviving the trigger selection have a completely different composition concerning the MC error flags. More than 95 % of the reconstructed and triggered events have three correctly matched particle candidates. This means that the trigger application resolves the two issues discussed earlier, to some extent.

Filter	Selection criteria
`gg2clst`	Exactly 2 neutral clusters $E_{\mathrm{CMS}} > 3\,\mathrm{GeV}$ and at least one cluster $4.5\,\mathrm{GeV}$ $\Delta\phi > 178°$& $\lvert\theta_1 + \theta_2 - 180°\rvert < 15°$
`ee1leg1clst`	1 loose track with $E_{\mathrm{cl,\,CMS}} > 4.5\,\mathrm{GeV}$ 2 clusters with $E_{\mathrm{CMS}} > 3\,\mathrm{GeV}$ (at least one $> 4.5\,\mathrm{GeV}$) $\Delta\phi$ =160°- 178° & $\lvert\theta_1 + \theta_2$ - 180°$\rvert$ <15°
`ee1leg1trk`	2 loose tracks with $\Delta\phi > 175°$ $\lvert\theta_1 + \theta_2$ - 180°$\rvert$ <15° one track with $E_{\mathrm{cl,\,CMS}} > 4.5\,\mathrm{GeV}$ other track with p_{CMS} >3 GeV
`eeBrem`	2 loose tracks with $p_{\mathrm{CMS,\,max}} > 4.5\,\mathrm{GeV}$ at least one cluster with $E_{\mathrm{lab}} > 1\,\mathrm{GeV}$ missing momentum $p_{\mathrm{missing}} > 1\,\mathrm{GeV}$ $\Delta\phi > 175°$ & $\lvert\theta_1 + \theta_2$ - 180°$\rvert$ <5°

Table 4.4.2: Definition of the processes rejected by the HLT trigger line [15].

4.5 Purity of the data sample

It is now clear that including the trigger selections removes a large part of the wrongly reconstructed events. But it is still unknown if the assumption of the ECL cluster with higher energy originating from the charged particle is correct and will lead to a pure final data sample. To analyze and improve the purity of the data sample, different cuts, based on the analysis in the Section 4.3, are applied. For each of the cuts, the purity and the cut efficiency are evaluated. The purity describes the fraction of events where the most energetic cluster originates from a correctly matched positron. The cut efficiency is the fraction of remaining events after the cut application, relative to the total number of events. Table 4.5.1 lists both variables for different cuts.

From Table 4.5.1, two different conclusions are drawn. At first, I extract that the total number of events is reduced drastically to below 10 % by applying cuts on the opening angles between the different clusters in combination with higher energy thresholds. Further, the application of the „hie“ trigger and the HLT line leads to an even larger reduction in events. But when applying both trigger cuts and $E_{\mathrm{cl,\,all}} > 0.7\,\mathrm{GeV}$ & $\alpha > 15°$ the number of remaining events and the cut efficiency changes only slightly compared to only applying the trigger cut. Meaning the triggers remove a large part of the wrong reconstructed events without further cuts being necessary.

The other variable in Table 4.5.1 is the purity of the data sample. The purity in the original data sample without cuts is already quite high, but with 81.6 % still too low to say it is a pure data sample, a good purity would be above 95 %. Additionally, in the original data sample, the events where

cuts	N_{orig}	purity	cut efficiency
no cuts	58.672,16	81.64 %	-
$E_{\text{cl, all}} > 0.7\,\text{GeV}$	42.833,02	79.30 %	73.00 %
$E_{\text{cl, all}} > 1\,\text{GeV}$	35.772,95	74.90 %	60.97 %
$\alpha > 15°$	9.336,55	75.18 %	15.91 %
$\alpha > 30°$	5.370.46	83.03 %	9.15 %
$E_{\text{cl, all}} > 0.7\,\text{GeV}$ & $\alpha > 15°$	4.914,42	73.28 %	8.38 %
$E_{\text{cl, all}} > 0.7\,\text{GeV}$ & $\alpha > 30°$	2.780,59	77.43 %	4.74 %
hie	4.731,27	66.20 %	8.06 %
hie & *HLT*	672,7	56.49 %	1.15 %
hie & *HLT* & $E_{\text{cl, all}} > 0.7\,\text{GeV}$ & $\alpha > 15°$	599,97	57.84 %	1.02 %

Table 4.5.1: Number of events, purity, and cut efficiency for different cuts. „all“ means that this cut is applied for the most and the second most energetic cluster. α refers to all three opening angle variations.

cluster splitting occurs are also included, meaning the ECL energies do not necessarily match the energies of the generated particles, which would lead to wrong results for the tracking efficiency.

By applying the different energy and opening angle cuts, the purity ranges from 73 % up to 83 %, so all those values are too low. As the L1 and the HLT trigger have to be applied for the final analysis, their impact on the purity of the data sample is also checked. The dataset's purity decreases below 70 % when applying the trigger selection. Also, applying an energy and an opening angle cut in addition to the trigger cut does not lead to an increased purity.

From these results, I conclude that the application of the selected triggers already leads to an improvement of the configuration of the dataset, but the dataset is not pure enough to conclude which cluster is created from the charged particle. Also, the application of cuts based on different kinematic variables does not lead to a pure enough dataset.

Chapter 5

Boosted Decision Tree

To improve the purity of the analyzed data sample and determine which cluster to use for the final binning of the tracking efficiency, a boosted decision tree (BDT) is employed. The algorithm takes a multidimensional set of input variables characterizes each event and assigns it to predefined classes based on patterns learned from the training data. We define these classes using MC truth information. The goal of the BDT is to apply it to MC and experimental data. For the latter, the different classes cannot be assigned by their definitions, as there is no MC truth information available for experimental data.

5.1 Introduction to Boosted Decision Trees

Decision trees are used to divide a dataset into various predefined classes. In this thesis, these classes are the class where the more energetic cluster originates from the charged particle, the class where the less energetic cluster originates from the charged particle, and a class containing all other background events. The BDT used here features a multiclass definition, but it is also possible to use binary BDTs, where the BDT only distinguishes between signal and background classes.

The BDT is trained on a dataset completely independent of the dataset used for the final analysis. This is necessary to avoid an overperformance of the BDT. Further, the training of a BDT includes a cross-validation feature for which an extra validation dataset is necessary. We split the available simulated data into three different datasets. The data used for the final analysis is 40 % of all MC. For the training data set 42 % of the available MC data is used, and for the cross-validation set 18 % of the available MC data is used. For both the training and the validation dataset, the classes are defined by MC truth information, and the class labels of each event are provided to the BDT during training.

A decision tree is based on yes-or-no questions. It takes the input variables and defines the questions based on them. For each event, those questions are then answered individually. For example, if the energy is an input variable, then each event is evaluated by having an energy below or above a threshold, which is set by the algorithm when defining the yes-or-no question. Each event is categorized as yes or no relative to the defined question. In the next question, the decision tree uses another variable or another threshold and divides the subgroups again into yes and no groups. The variables used to train a BDT in this thesis are kinematic, particle identification, and detector variables.

Using a boosting framework to train the decision tree improves its performance. Boosting means that at the end of each tree, new per-event weights are calculated. These weights are applied to the individual events, creating a weighted dataset. Typically, events that are hard to classify get assigned a higher weight than those that are easily assigned to the correct class. The weighted dataset is then used to train the next decision tree. With this method, the decision trees are grown differently at each iteration, and the BDT's performance improves with every iteration. An iteration describes the growth of one decision tree.

The total number of iterations influences the duration of the training and the final result of the BDT. To select the ideal number of iterations, cross-validation is used. For this, the model created in the latest decision tree is applied to a validation dataset. When the prediction of the different classes works well on this validation set, meaning the class prediction is similar to the true classes, and when a new tree does not further improve those predictions, the training is stopped. The number of iterations executed at this point is the ideal number of iterations. It is important that the validation dataset, as well as the training dataset, inherits the true definition of the classes.

The learning rate λ of the BDT is a small positive number and is important for the training efficiency of the BDT. It describes the correction of the predictions of the BDT between the individual iterations. Smaller values yield smaller improvements between iterations, and larger values yield larger improvements. However, the learning rate and the number of iterations are correlated in the sense that for a smaller learning rate, a larger number of iterations is necessary to achieve a good performance. The other way round, for high learning rates, a lower number of iterations is sufficient.

Limiting the number of involved variables per tree reduces and controls the complexity of the boosting algorithm. Smaller trees have the ability to improve the overall performance of the BDT and reduce the risk of overtraining. When a model learns the training data too well, it is called overtraining. In a such case the performance of the BDT on testing data is rather poor while it performs extremly well on the training dataset.

5.2 LightGBM

To train our BDT, we use the LightGBM python package [16]. This package is a gradient boosting framework that is especially suitable for our

Figure 5.2.1: Visualization of level-wise (left) and leaf-wise (right) growth of decision trees [16].

case, as it has several characteristics that allow for a sufficient handling of large-scale data. Those characteristics are its low memory usage, fast training with high efficiency, and parallel and distributed creation of the decision trees.

Another unique characteristic of LightGBM is its approach to optimize the classification performance of the BDT. In most decision-tree learning algorithms, the trees are grown level-wise. For a level-wise trained BDT, all nodes of the tree are split at each intermediate step of the tree, as seen in the left part of Figure 5.2.1. But LightGBM instead uses a leaf-wise approach as displayed in the right part of Figure 5.2.1. This graphic shows in green the leaves with the most promising results. The algorithm continues training only with the most promising leaf. Meaning the next yes-or-no question is only applied to this leaf. This approach is used until the end of the decision tree. This strategy generally allows the model to reach a better performance faster than level-wise growth [16].

To control the number of grown trees and also the risk of overtraining, LightGBM uses early stopping. For this, the method of cross-validation is used. After each iteration, the accuracy of the trained model is evaluated against a validation dataset, and the loss function is computed. The loss function characterizes the performance of each individual tree, and it measures the difference between the predicted and the true classes. From the computed loss values, the delta loss is computed by subtracting the loss value of the current decision tree from the loss value of the tree trained in the previous iteration. By setting a limit for the early stopping function, the algorithm will stop improving the model when the delta loss does not improve within the early stopping limit. This feature allows us to find the best number of iterations and helps to reduce the overall training time and prevent overtraining.

The variables used for training our BDT are defined in Table 5.2.1. In general, nearly all available variables can be used to train a BDT, but they must be available for both the simulated dataset and the experimental data. This excludes all variables based on MC truth information. Further, excluded are also the energy of the most energetic cluster and the polar angle of this cluster, as the final tracking efficiency calculation will be binned in these two variables.

The prediction of the classes by the BDT is based on the per-event probability P. For a binary BDT, only the probability for the signal class

Category	Variables
Beam background suppression	Most / second most energetic cluster
Fake photon suppression	Most / second most energetic cluster
Cluster energy	$E_{\text{sum, CM}}$, $E_{\text{cl, etag}}$, $E_{\text{cl, secmost}}$
Cluster angles	θ_{cl}, ϕ_{cl} (etag, most, secmost)
Tag kinematics	p, p_t, p_x, p_y, p_z, θ, ϕ, E_{cl}/p
Shower shape	E1E9, E9E21 (etag, most, secmost)
Cluster size	clusterNHits (etag, most, secmost)
Angle differences	$\Delta\theta_{\text{cl}}$ and $\Delta\phi_{\text{cl}}$ between the three clusters

Table 5.2.1: Overview of input variables used in the BDT.

is determined, and then only events with a probability higher than a given threshold are assigned to the signal class, while all other events are assigned to the background class. For a multiclass BDT, for each event, the probability P_x for each class x is determined. The event is then assigned to the class with the highest probability. Here, the threshold for class acceptance is set individually. Events that do not meet this threshold are then discarded.

Often, the different, defined classes are not represented with the same fractions within the dataset, which leads to a bias within the performance of the BDT. This is especially important when the signal class is significantly larger or smaller than the background class. To resolve this issue of class imbalances, weights are included in the training of the BDT. The fraction of a class f_{c} and the total number of classes C determine the class weights accounting for class imbalances

$$w_c = \frac{1}{C \cdot f_{\text{c}}}. \tag{5.2.1}$$

Neglecting the weight of the different classes leads to an overprediction of the dominating classes. Overprediction means that the number of events predicted to be in a certain class is larger than the true number of events in this class. Including the class weights prevents the classes from overprediction.

The weight option in the BDT training is not only used to implement class weights, but also to include luminosity weights. The luminosity weights are necessary when using simulated data. They assure that the composition of the final data set reflects the correct proportions of each process, depending on the individual cross sections. The luminosity weight $w_{\mathcal{L}}$ is computed as

$$w_{\mathcal{L}} = \frac{\mathcal{L}_{ee}}{\mathcal{L}_{\text{proc}}}, \tag{5.2.2}$$

where $\mathcal{L}_{ee}$ is the integrated luminosity of the main process $e^+e^- \to e^+e^-(\gamma)$ and $\mathcal{L}_{\text{proc}}$ is the integrated luminosity of the different processes which are weighted relative to the main process. We use the class weights and the luminosity weights at the same time by computing a combined weighting factor. This is done by multiplying the corresponding weights for each event by each other

$$w = w_c \cdot w_{\mathcal{L}}. \tag{5.2.3}$$

5.3 BDT model selection and optimization

Various binary and multiclass BDTs were trained, tested, evaluated, and compared to get the best class definition and the best-performing BDT. The performance of a BDT is evaluated by using its confusion matrix, the purities of the different classes, and the overall accuracy.

A confusion matrix relates predicted classes and true classes, which are defined by MC truth information. From the confusion matrix, the purity of each class is determined by

$$P_c = \frac{N_{c,\mathrm{t}}}{N_c} \tag{5.3.1}$$

where N_c is the number of events predicted to be in class c, and $N_{c,\mathrm{t}}$ is the number of events predicted to be in c and assigned to the true class t. A combination of the purities of the various classes of a BDT gives the overall accuracy

$$P_{\mathrm{all}} = \frac{\sum P_c N_c}{\sum N_c} = \frac{\sum N_{c,\mathrm{t}}}{\sum N_c}. \tag{5.3.2}$$

For this thesis, the goal is to extract three classes of events: the higher-energy class, where the more energetic cluster originates from the charged particle. The lower-energy class, where the less energetic cluster originates from the charged particle, and the background class, containing all other events. To achieve pure samples for those classes, we test different MC-truth-based definitions for the classes of the BDT. This is done using only the reconstruction with an electron tag. But the training itself will later be performed separately for electron and positron tags. For now, we assume that the training of the BDT works equally well on both datasets. A good definition of the classes is reached when the classes contain only events useful for the final analysis and when the purity of those classes is very high. Any impurities in the classes lead to higher uncertainties in the final tracking efficiency results.

For the first BDT, the classes are defined by one of the clusters originating from a correctly reconstructed positron and the other cluster from a correctly reconstructed photon, so that only correctly reconstructed radiative Bhabha events are included in the two signal classes.

Higher class:

- Correctly reconstructed e^+ creates the higher-energy cluster
- Correctly reconstructed γ creates the lower-energy cluster

Lower class:

- Correctly reconstructed γ creates the higher-energy cluster
- Correctly reconstructed e^+ creates the lower-energy cluster

Background class: all other events

This multiclass BDT returns a purity for the higher class of 71 % and for the lower class of 44 %. Bboth purity values are too low for our analysis. Raising the threshold for the probability of each class increases the purity of the higher class to 95 % and for the lower class to 80 %. This is a good improvement, but it comes with the cost of discarding events.

To explore further possibilities to improve the BDT's performance, we use the same signal class definitions to train two separate binary BDTs, with the earlier-defined classes serving as signal classes and all other events classified as background. The expectation for those BDTs is that the purity of each class and the overall accuracy of the output data increase. But instead of a higher purity value, the purity of the signal classes and the overall accuracy are even lower than before. From these results, I conclude that using the same class definitions for training a binary BDT does not lead to an improvement in the performance or the purity of the different classes. The kind of BDT, binary or multiclass, is probably less important for a good performance than the exact definitions of the different classes. Therefore, binary BDTs are not considered any further.

As we are most interested in where the charged particle goes and less interested in whether the other cluster is a correctly reconstructed photon or not, the next BDT is using a more track-oriented class definition.

Higher class:

- Correctly reconstructed e^+ creates the higher-energy cluster
- The lower-energy cluster is not created by an electron or positron

Lower class:

- The higher-energy cluster is not created by an electron or positron
- Correctly reconstructed e^+ creates the lower-energy cluster

Background class: all other events

The intention of these class definitions is to assign all events affected by cluster splitting to the background class. This is important as for correctly and completely reconstructed electrons and photons, the reconstructed cluster energy is equal to the momentum of the particles. This is not the case for events affected by cluster splitting, but as the momentum will be used for the final binning of the tracking efficiency, we have to ensure the relation between reconstructed cluster energy and momentum is correct for all events used. Therefore, events affected by cluster splitting either need to be excluded from the final data set or clearly separated from correctly reconstructed radiative Bhabha events.

This BDT has a purity for the higher class of 96 % and for the lower class of 34 %. The overall accuracy of the BDT is 84 %. The purity for the higher class is already closer to an ideal purity of more than 97 %, but with the goal of using both signal classes for the tracking efficiency calculation, the purity of the lower class has to be improved significantly.

In the next BDT, all different configurations of clusters and matched particle species are considered for the class definition. This means that, in addition to the higher and lower classes, which include a correctly reconstructed positron and a correctly reconstructed photon, and the background class, we define two additional classes. Those classes represent the cases where either two correctly reconstructed photons or two correctly reconstructed electrons are part of the decay besides the tag electron.

Higher class:

- Correctly reconstructed e^+ creates the higher-energy cluster
- Correctly reconstructed γ creates the lower-energy cluster

Lower class:

- Correctly reconstructed γ creates the higher-energy cluster
- Correctly reconstructed e^+ creates the lower-energy cluster

Photon class:

- Correctly reconstructed γ creates the higher-energy cluster
- Correctly reconstructed γ creates the lower-energy cluster

Electron class:

- Correctly reconstructed e^+ creates the higher-energy cluster
- Correctly reconstructed $e^\pm$ creates the lower-energy cluster

Background class: all other events

The higher and the lower classes represent the correctly reconstructed radiative Bhabha decays, which are of interest for the final calculation. Whereas the photon and the electron classes are not further relevant for the final analysis, but are included here to provide a more precise definition of the classes. The goal of this BDT is a better performance in purity and overall accuracy than the previous BDTs by providing more information about the background decays. The overall accuracy of this BDT is 95 %. The purity of the higher class is 100 %, for the lower class it is 76 %, for the photon class it is 73 %, for the electron class it is 86 % and for the background class it is 99 %.

The numbers for the accuracy and the purity of the higher class and the electron class look very promising, but all other classes have large impurities. For the calculation of the tracking efficiency, the events in the higher class would be sufficient, as they cover quite a large range of the phase space and provide sufficient statistics, as it is also the largest class. The purity for the higher class did, as expected, improve with a more specific definition of the background classes. For the lower class, which could also be used for the tracking efficiency calculation, the purity is quite low for all trained BDTs and did not improve with more precise definitions of the classes. Due to the limited time for this thesis, we decided to set our focus on the higher class and will not analyze the lower class any further.

Analyzing the MC truth information of events assigned by the last BDT to the higher-energy class. I conclude that not all events within the predicted higher-energy class are correctly reconstructed radiative Bhabha events. There is a fraction of 16 % of events that are affected by cluster splitting. I conclude that the purity of this class concerning the correctly assigned classes is very good. But the definition of the higher-energy class allows for badly reconstructed events within the class.

This means the definition of the classes used for training the BDT has to be further adjusted so that the events affected by cluster splitting are really assigned to the background class.

From the investigation of the phenomena of cluster splitting in Section 4.3, we extract that the particle ID of the clusters helps to distinguish between events with and without cluster splitting on the MC truth level. In correctly reconstructed events, the two clusters have different particle IDs, whereas in events with cluster splitting, the two clusters share the same particle ID. We include this in the definition of the BDT classes.

Higher class:

- Correctly, primary reconstructed e^+ creates the higher-energy cluster
- Higher- and lower-energy clusters are not created by the same particle

Lower class:

- Correctly, primary reconstructed e^+ creates the lower-energy cluster

  – Higher- and lower-energy clusters are not created by the same particle

Background class: all other events

These class definitions assign nearly all events affected by cluster splitting to the background class.

The purity and accuracy values of this BDT are more meaningful than those from the previous BDTs, as the events assigned to the higher class are only those events to be used for the calculation of the tracking efficiency. This was also checked by using MC truth information. The purity of the higher class is 95 %, for the lower class it is 91 %, and the overall accuracy is 91 %. The purity of the higher-energy class increases further by setting a threshold on the per-event probability of the class.

5.4 The final BDT

The last defined classes are the final version for our BDT. For a positron tag, the charges are just flipped. Up to now, the BDT has only been optimized by changing the definition of the different classes. But the performance of the BDT improves further by changing its training parameters. The two most important parameters for training BDTs are the number of iterations and the learning rate, but also using the correct weighting factors is important.

For our BDT, two different weight configurations were used: the luminosity weight and the combined weight. The effect of overtraining is identified for BDTs, which only use a luminosity weighting and do not account for any class imbalances. This is quantified by using the percentage of events assigned to the higher-energy class by MC truth. The fraction of those events truly assigned to the higher-energy class is 58 %, visualized as the red line in Figure 5.4.1. For BDTs only using luminosity weights, this fraction is larger than 58 %. In Figure 5.4.1, the fraction of the higher-energy class f_0 is plotted against the purity of this class. The results from BDTs with only luminosity weight are all located above the red line of the true fraction of the higher class. From that, I conclude that the class weights are necessary for getting a good-performing BDT.

To find the ideal set of parameters for training the BDT, I tested different combinations and compared their results. The number of iterations is varied between 5,000 and 10,000, and the learning rate λ between 0.01 and 0.03. The performance of the different BDTs is compared by analyzing the purity of the higher-energy class, the efficiency of predicting the higher-energy class, and the potential for overtraining.

Figure 5.4.1 compares the performance of the BDTs on the training (green) and the validation (orange) dataset. The different types of data points refer to different parameter settings used in the BDT training. The performance of the BDT on the training dataset should be similar to the

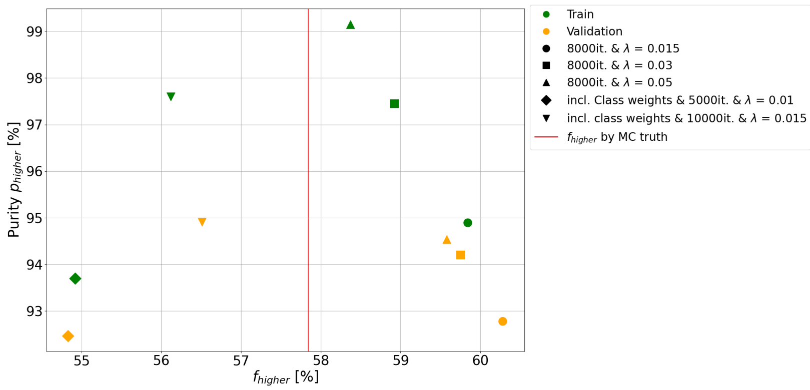

Figure 5.4.1: Purity and efficiency for the higher-energy class for different BDTs evaluated on training and validation datasets.

performance on the validation dataset, as the BDT should be able to work equally well on different datasets. A too good performance on the training dataset indicates over-training. For Figure 5.4.1, this means that the closer the green and orange datapoints for a BDT are, the less over-trained the BDT is. Overall, the differences for a BDT using the combined weight factors are not too large, so there is no clear sign of over-training.

To finalize the BDT parameters, the dataset is now expanded to the full available MC luminosity of $34\,\mathrm{fb}^{-1}$. The full MC dataset includes all available background processes scaled by luminosity. It is then split into training, validation, and testing datasets.

For the final training, the learning rate of 0.015 is used, the number of iterations is set to 8,000, and the early stopping limit is set to 200. This means, when the performance of the BDT doesn't improve for 200 iterations, the training process is stopped to avoid over-training. The now significantly larger dataset leads to the BDT already stopping training after less than 900 iterations, and the performance of the BDT is much better than with the smaller dataset.

This improved performance seems nice, but also a bit too good. Therefore further analysis steps are performed before finalizing the BDT. A first step for that is to reduce the set of training variables. This is necessary to prevent overtraining and to give the most meaningful variables a larger impact on the performance of the BDT. It also reduces the chance that the BDT is trained on distributions that don't match the experimental data. We use two characteristics to identify variables that can be removed from the set of training variables. Either several variables provide a similar input, so including only one of them in the set of training variables is sufficient, or the importance of a variable is very low. The BDT determines the importance of each of the training variables, indicating how often a variable was used to ask a yes-or-no question during the training process. From this parameter, the most and the least important variables in the training process of the BDT are extracted. For this, the exact values are less important, but the order and the relative values are analyzed qualita-

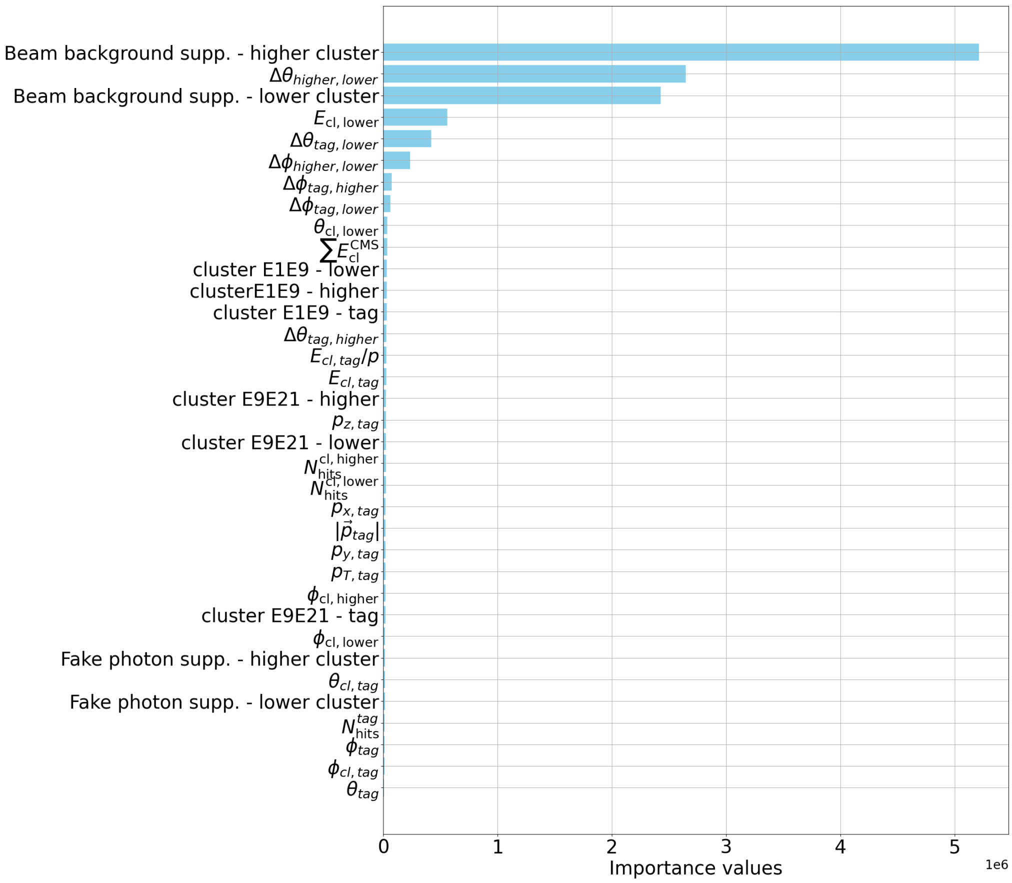

Figure 5.4.2: The importance of each variable used for training the BDT with a total of 35 variables.

tively. We identify variables with very low importance and remove them from the set of training variables. Figure 5.4.2 visualizes the importances of all originally used training variables. It becomes clear that quite a few variables have nearly no impact on the performance of the BDT relative to the variables with the largest importance.

In this first step of reducing the number of training variables, the variable describing the number of ECL crystals for one cluster hit for all three involved clusters, the tag electron cluster, the higher-energy cluster, and the lower-energy cluster, is removed. All of them have very low importance values. The same yields for the cluster energy over track momentum ratio of the tag electron. Further, for the tag electron, quite a few momentum variables were included, but as they all provide similar information, the p_x, p_y, and p_z variables are removed. The clusterE1E9 and cluster E9E21 variables are used for particle identification at Belle II, but as we are sure that the tag electrons are always electrons, those variables do not provide too much information for this analysis. This is also reflected by the low importance of those variables in the BDT as seen in Figure 5.4.2. Therefore, we remove the clusterE1E9 and clusterE9E21 variables of the tag electron from the set of training variables.

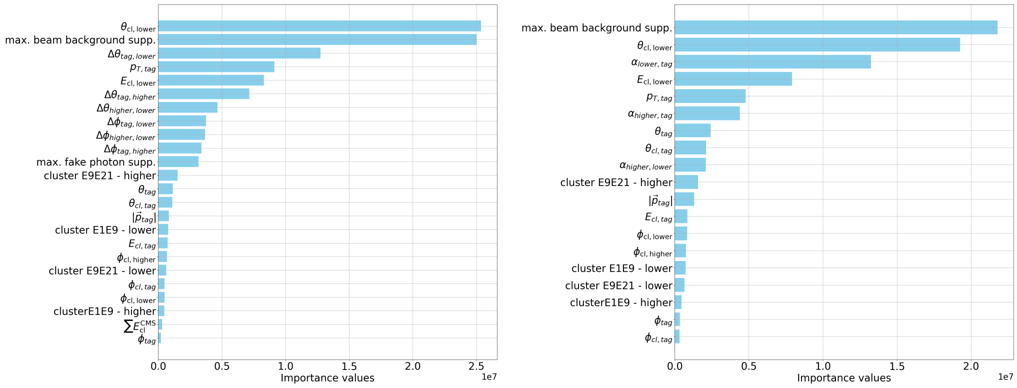

Figure 5.4.3: The importance of each variable used for training the BDT. The importances of a BDT using a larger training variable set (left), and the importances of the smaller and final set of training variables (right).

At this point, it became clear that the beam background suppression and fake photon suppression variables are very helpful for classifying real photons. But their large impact on the BDT stems mainly from the fact that those variables are calculated only for neutral clusters. This means using the suppression variables separately for each cluster leads to a bias in training the BDT. This bias arises because the BDT knows which cluster is the neutral cluster, since the clusters originating from tracks will have no value. To avoid this bias, a combined variable is defined. For each event, only the maximum suppression value of the two clusters is kept. This final variable contains one value for each event without inheriting the knowledge of which cluster the suppression value was calculated for. The new variables are „Beam background suppression“ and „Fake photon suppression“.

With the now reduced set of variables, we retrain the BDT and re-evaluate the importances of the different variables. The beam background suppression remains among the two most important variables. The importance of the fake photon suppression is only one-fifth of the importance of the most important variables, this is extracted from the left plot in Figure 5.4.3. I conclude that the fake photon suppression variable is not important for the training of the BDT, and the variable is therefore removed from the set of training variables.

At this point, not only is the importance of a variable relevant, but also the agreement of the individual training variables between data and simulation needs to be considered. Only well-simulated variables should be used as training variables, to ensure that the performance of the BDT on data and simulation is similar. From the comparison of data and simulation, for events assigned by the BDT to the higher-energy class, of the training variables mentioned in Table 5.2.1, variables with bad agreement are identified. The full analysis for the data-simulation agreement of all variables will be provided in Section 6.3.

One of the least important variables in the left panel of Figure 5.4.3 is the ECL energy sum of all three clusters. Therefore, and due to the

Category	Variables
Beam background suppression	Maximum value of both clusters
Cluster energy	$E_{\text{cl, tag}}$, $E_{\text{cl, lower}}$
Cluster angles	$\theta_{\text{cl,tag}}$, $\theta_{\text{cl,lower}}$, $\phi_{\text{cl,tag}}$, $\phi_{\text{cl,lower}}$
Tag kinematics	p, p_t, θ, ϕ
Shower shape	E1E9, E9E21 (higher, lower)
Opening angles	$\alpha_{\text{tag, higher}}$, $\alpha_{\text{tag, lower}}$ and $\alpha_{\text{higher, lower}}$

Table 5.4.1: Input variables used in the final training of the BDT.

non-ideal data-simulation agreement, this variable is removed from the set of training variables. The absolute differences of the polar angle and the azimuth between the different clusters all have a relatively poor agreement between data and simulation. But as their importances are relatively high, following the distribution in the left panel of Figure 5.4.3, removing them without a sufficient replacement weakens the performance of the BDT. Therefore, we determine the opening angles between the different clusters using Equation (4.3.3). The data-simulation agreement of the opening angles is better than the agreement of the pure angle differences. Therefore, the opening angles are added to the set of training variables, replacing the differences of the azimuth and the polar angle.

From the importances of the final set of variables, in the right plot of Figure 5.4.3, I extract that the beam background suppression, the polar angle of the lower-energy cluster, and the opening angle between the lower-energy cluster and the tag electron cluster are the most important variables for the BDT training. This agrees with the findings during the previous analysis steps. I also extracted that the importances are spread over all used training variables, and the most important variables are less dominant. The final list of variables, including the new suppression variable definition and the opening angles, is given in Table 5.4.1.

After changing and removing the training variables for the BDT, it is retrained using the same parameters as before: $\lambda = 0.015$ and 8,000 iterations. The complexity of the new BDT increases significantly compared to the BDT with the original set of variables. This increased complexity is identified from the increased number of iterations, which are necessary to achieve a similar result, and the early stopping function is not called anymore. The main difference in complexity comes from the new definition of the suppression variables.

Up to now, the purity of the higher-energy class has been discussed without setting any additional constraints on the BDT probabilities. To achieve a better purity, the threshold for accepting an event in a certain BDT class is adjusted. The performance of a BDT depends, in addition to the purity of the class of interest, also on the cut efficiency of this class.

To be able to extract the best parameters for training the BDT, we compare purity and cut efficiency for two different BDTs. This is done by applying the different threshold cuts to the BDT probability and eval-

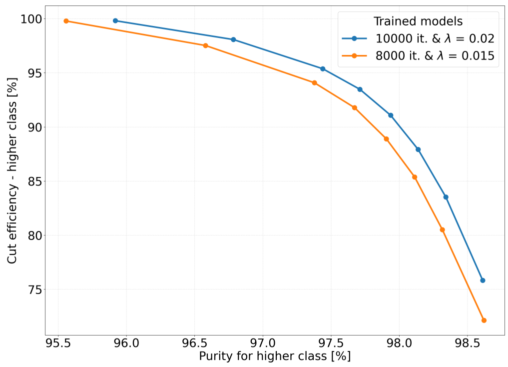

Figure 5.4.4: Purity of the higher-energy class versus the cut efficiency for this class, each datapoint represents one cut value on the probability of the BDT.

uating purity and cut efficiency of the higher-energy class for each of the thresholds. The cut efficiency of the higher-energy class is determined by dividing the number of events in this class after the threshold adjustment by the number of events in this class before any threshold adjustment.

The threshold values for the probability of the higher-energy class are 0.5, 0.6, 0.7, 0.75, 0.8, 0.85, 0.9, 0.95. For both BDTs, the expectation that purity increases with decreasing cut efficiency is true. In Figure 5.4.4, this distribution of the purity versus the cut efficiency of the higher-energy class is plotted, with the BDT with lower iterations and lower learning rate in orange, and the other BDT performance in blue. Both curves have a similar shape, while the purity is slightly higher for the blue data points. The same yields for the cut efficiency. From this analysis step, I extract that for the large dataset, and the slightly smaller set of variables, a higher number of iterations in combination with a slightly larger learning rate leads to a better-performing BDT.

From Figure 5.4.4 a second conclusion is drawn, concerning the ideal cut on the probability to accept events in the higher-energy class. For our analysis, it is important not to remove too many events, as we are already limited by the sample size of the available $ee(\gamma)$-signal simulation. Therefore, it is the goal to keep more than 90 % of the events assigned to the higher-energy class. To achieve this, the ideal cut on the probability is 0.8. Leading to a purity of 97 % for the higher-energy class and a total of 91 % of events remaining.

A final check before using the output of the BDT to calculate the tracking efficiency is to analyze whether the BDT is over-training or not. For this, the multiclass log loss of the training and the validation set is monitored during the training process. The multiclass log loss describes the improvement of the BDT predictions for each iteration compared to the previous iteration. The ideal behavior of those monitored values is that both values decrease in a similar way, and more importantly, the loss of the validation set is never increasing until the end of the training process.

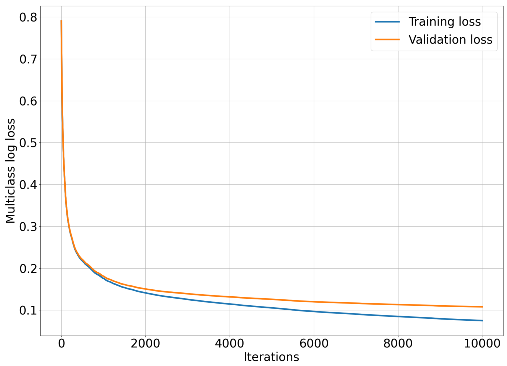

Figure 5.4.5: Multi log loss in dependence of the number of performed iterations for the training and the validation dataset for a BDT using an electron tag, $\lambda = 0.02$, and a maximum number of 10.000 iterations.

Figure 5.4.5 shows the log loss on the vertical axis, the number of iterations on the horizontal axis, the values for the training dataset in blue and those for the validation set in orange. The BDT trained here is not overtraining as the curves in Figure 5.4.5 behave ideally.

5.5 Application of the BDT

The performance of the BDT is now improved, and it is assured that it is not over-fitting. We now train two separate BDTs, one for an electron tag and one for a positron tag. For the higher-energy class, we accept only events with a probability larger than 0.8. Further, they are applied to the MC testing dataset and all available experimental data. The MC testing set has a luminosity of 13.7 fb^{-1} and the experimental dataset has a luminosity of 357.3 fb^{-1}.

For the MC dataset, the overall performance of the BDT is analyzed by a confusion matrix. For the electron tag, in the left panel of Figure 5.5.1, the purity for the higher-energy class is 97.7 % and events in this class make up 52.4 % of all events in the dataset. For the positron tag, right panel of Figure 5.5.1, the purity of the higher-energy class is 98.0 % and these events make up 54.7 % of all events in the dataset. The application of the BDT on data gives for the negative tag 45.5 % of events in the higher-energy class and 32.4 % of events in the lower-energy class. For the positron tag, it gives 45.5 % in the higher-energy class and 37.2 % in the lower-energy class. The large difference between the fractions of events in the higher-energy class for the MC and the data was not expected. One possibility to get similar fractions of events in the higher-energy class for both datasets is the application of a cut based on the trigger selection.

In Section 4.3, the trigger selection was discussed, but the relevance of the constraints on the neutral cluster within the HLT line definition was not evaluated in detail. The HLT line applies specific constraints on the neutral cluster in the event, those are $E_{\mathrm{CM}} > 2\,\mathrm{GeV}$ and $\theta \in [32°,130°]$. An

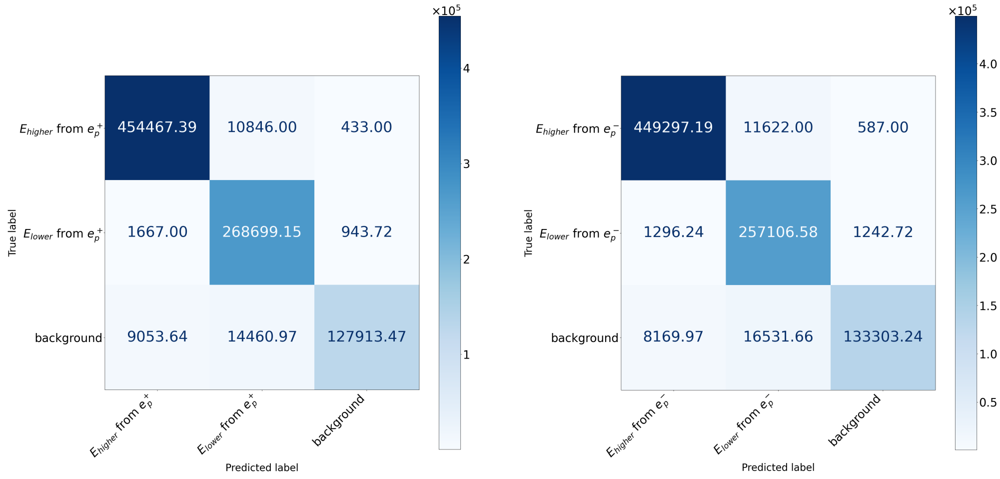

Figure 5.5.1: Confusion matrices for electron tag (left) and positron tag (right).

application of those constraints only on the neutral clusters creates a bias in our data, and especially in the final results. When the charged particle does not satisfy those cuts, it is clear which one of the clusters is originating from the charged particle. Due to this, the cuts on the energy and the polar angle are applied to the higher- and the lower-energy cluster, independent of the charge of the reconstructed particle.

Those cuts reduce the number of events in data and in MC. For MC 68 % of all events remain, and for data 67 % of all events remain. For the higher-energy class used for the final calculations, the fractions of remaining events are also quite similar. Here for the electron tag 79 % of the events in MC remain, and 77 % of the events in data remain. For the positron tag, the fractions of events categorized by the BDT as higher-energy class are slightly lower, for MC 73 % of those events remain, and for data 72 % remain.

The overall discrepancy in the fractions of events getting categorized as higher-energy class in data and MC is not resolved with this cut, as the fractions of events in the respective classes are reduced similarly in data and MC.

To verify that the BDT performs similarly on data and MC and to exclude the possibility that the different fractions of events come from a different performance of the BDT on data and MC, the distributions of the probabilities assigned by the BDT to each event are analyzed. We are mainly interested in the distributions of the probabilities determined for the higher-energy class. This probability is calculated for each event, and due to the cut on the probability, only events with a probability higher than 0.8 are assigned to the higher-energy class.

Comparing the shapes of the probability distributions for the individual classes between data and MC for the negative tag leads to the conclusion that the higher- and the lower-energy class events have very similar

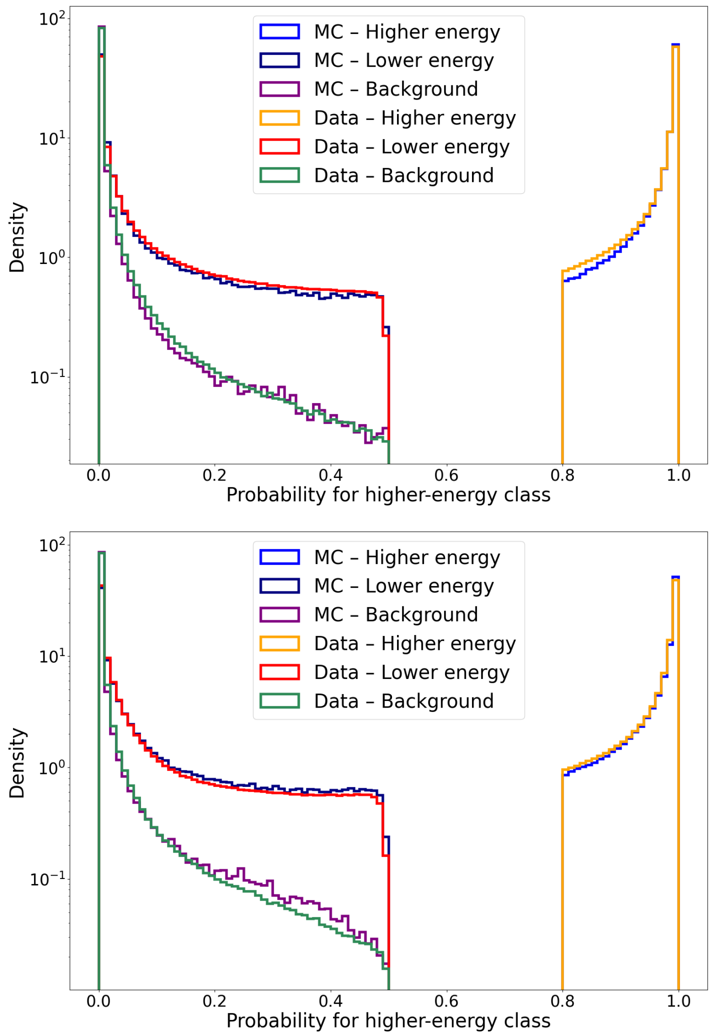

Figure 5.5.2: BDT output probabilities for the higher-energy class for an electron (top) and a positron (bottom) tags, including the trigger-based energy and polar angle cuts.

distribution shapes. The probability distributions after the application of the trigger-based energy and polar angle cuts are presented in Figure 5.5.2, where the probability of the higher-energy class is plotted as a density plot. The distribution of the background class is most interesting, as there are small deviations between the shape of MC and data.

The probabilities for the higher-energy class for the positron tag are shown in the bottom panel of Figure 5.5.2. The distributions for the higher- and lower-energy classes have the same shapes as for the electron tag, while the background class has an excess of MC events towards higher probability values. Overall, the differences are not significant, and no clear explanation for the differences in the fractions of the classes is found. The differences are therefore assumed to come from a slightly different configuration of the experimental data than the simulated data, which is expected as the MC is not perfectly simulating the experimental data taken at Belle II.

Chapter 6

The final data sample

In this chapter, the quality of the final dataset, consisting of events assigned to the higher-energy class by the BDT, is evaluated. Different variables are examined to assess the resolution of the selected sample. Based on this study, the final binning scheme is defined. The remaining background contributions are determined and evaluated. A detailed comparison between experimental data and MC is performed to validate the modeling of all relevant variables. The objective of this chapter is to confirm the composition and reliability of the final dataset.

6.1 Resolution of different variables

The resolution of the data sample is evaluated by comparing the reconstruction variables with the generator-level variables, which is only possible for MC. Based on the resolution of the reconstructed variables, a reasonable binning scheme for the final calculation of the tracking efficiency is defined. The tracking efficiency will be calculated for bins in a phase-space plot consisting of the ECL cluster energy and the ECL polar angle.

To check if the generated and reconstructed values are the same, the ratio of those is determined for each event. Ideally, this ratio is equal to one, whereas larger discrepancies between reconstructed and generated values have to be considered in the binning scheme or by applying further cuts to the dataset. In this thesis, the range of good resolution is defined as a 10 % interval centered on one.

A good resolution is necessary for the calculation of the tracking efficiency, as we assume that the values coming from the ECL describe the reconstructed particles correctly. A larger disagreement would not allow for the usage of this assumption, as it would place events in wrong bins.

The ECL cluster polar angle is compared to the MC polar angle by calculating their ratio: $\theta_{\mathrm{cl}}/\theta_{\mathrm{MC}}$. The ideal value for this ratio is one, as this means the reconstructed polar angles match the generated angles. For the ratio of the polar angles for both possible tags, more than 99.99 % of the analyzed events are located within the range of good agreement.

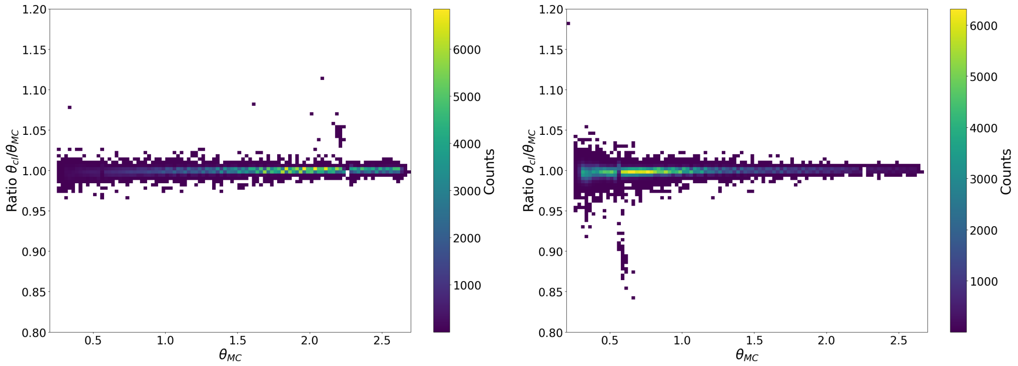

Figure 6.1.1: The generated polar angle and the ratio of reconstructed polar angle over the generated polar angle for the particle candidates of the higher-energy cluster. Only events predicted to be in the higher-energy class, for an electron tag (left) and a positron tag (right), are visualized.

In Figure 6.1.1, the ratio $\theta_{\text{cl}}/\theta_{\text{MC}}$ is plotted on the vertical axis and the polar angle of the generated particle on the horizontal axis, visualizing the resolution of the polar angle of the higher-energy cluster.

The majority of events reconstructed with an electron tag, in the left panel of Figure 6.1.1 have a ratio that deviates by at most 3 % from the ideal value of one. For the positron tag, the distribution, in the right panel of Figure 6.1.1, is slightly broader for small θ_{MC} values but still within the range of good resolution.

In both plots, a few events are clearly deviating from one, but their total number is so small that these are neglected. I conclude, therefore, that the reconstruction of the ECL polar angle works very well, and it is no problem to use this variable for the final binning of the tracking efficiency. For the final binning of the calculated efficiencies, the bin sizes of the polar angle as defined in [17] are used. The bin edges are: $\theta_{\text{cl}} = [0.22, 0.56, 1.13, 1.57, 1.88, 2.24, 2.71]$. To avoid misleading bins filled only with very few events, we customize the binning according to the applied trigger-based cut. The bin edge 0.56 is changed to 0.568, and the edge 2.24 is changed to 2.269.

The resolution of the reconstructed cluster energy is analyzed by comparing it to the momentum of the associated generated particle. For events reconstructed using an electron tag and assigned to the higher-energy class, more than 97 % of them are within the range of good resolution. In Figure 6.2.1, the ratio $E_{\text{cl}}/p_{\text{MC}}$ is on the vertical axis and the generated momentum on the horizontal axis for the electron tag (left) and the positron tag (right). For the positron tag, the fraction of events within [0.9, 1.1] is more than 98 %.

The final phase-space plots will therefore use an energy binning with a bin size of 1 GeV. With this bin size and the resolution of the energies, this leads to 9 % of events being placed in the wrong energy bin.

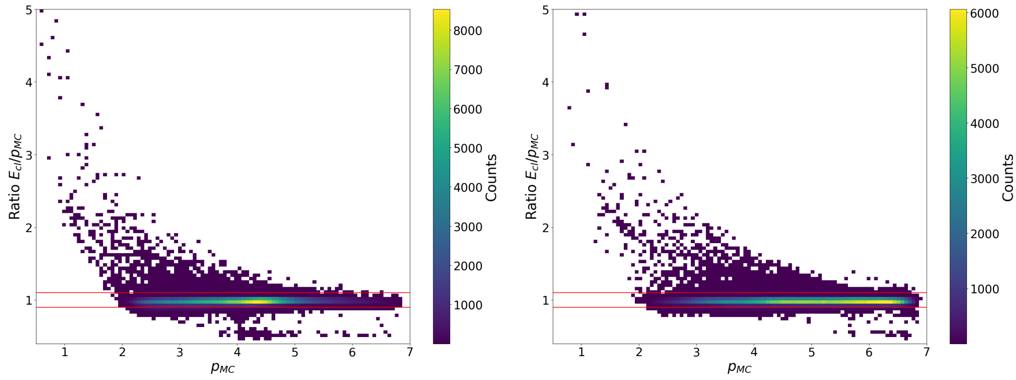

Figure 6.2.1: The generated momentum and the ratio of reconstructed energy over the generated momentum for the particle candidates of the higher-energy cluster. Only events predicted to be in the higher-energy class, for an electron tag (left) and a positron tag (right), are analyzed.

6.2 Background contributions

The resolution of the final MC data sample is good, but to ensure its high purity, the background contributions from other decays have to be examined. Here, we classify all events not coming from the $ee(\gamma)$-signal MC file as background contributions.

First, we determine the reconstruction efficiency for all MC processes, and the experimental data. Since two separate reconstructions for positron and electron tags are performed, the reconstruction efficiency is determined separately for those two tags. Both reconstructions are done using the same input data, but the definition of the reconstruction algorithms still leads to two independent data sets for the final analysis.

To quantify the success of the reconstruction, the number of generated events is compared to the number of reconstructed events by calculating the reconstruction efficiency using Equation (2.2.5), and the uncertainty of the efficiency is determined from Equation (2.2.8), but due to the large amount of events those are negligibly small.

For the processes with a very low number of reconstructed events, the efficiency is approximated by calculating the upper limit. Here, the 90 %-interval is determined by solving Equation (2.2.12) for the efficiency, using a fixed probability of getting zero events: $I_\epsilon(1, N+1) = 0.9$, the efficiency is then

$$\epsilon = 1 - \sqrt[N+1]{0.1}. \tag{6.2.1}$$

From the values in Table 6.2.1, I extract that for nearly two-thirds of the subprocesses, the upper limit calculation is necessary to determine the reconstruction efficiency. This and the low values for the reconstruction efficiency indicate that a large part of the background processes contribute only insignificant amounts of events to the analyzed MC dataset. The largest reconstruction efficiencies have the $e^+e^- \to e^+e^-(\gamma)$ and $e^+e^- \to \gamma\gamma$ MC datasets and the experimental dataset.

process	$N_{\mathrm{gen}}[10^6]$	$N_{\mathrm{rec}}^{\mathrm{neg}}$	ϵ^{neg} [%]	$N_{\mathrm{rec}}^{\mathrm{pos}}$	ϵ^{pos}[%]
data	33,300.0	24,508,803	0.07	23,916,642	0.07
e^+e^-	2,650.8	2,040,111	0.08	2,060,442	0.08
$\gamma\gamma$	254.2	596,066	0.23	569,107	0.22
$\tau^+\tau^-$	131.6	768	$5.8 \cdot 10^{-4}$	811	$6.2 \cdot 10^{-4}$
$e^+e^-e^+e^-$	1,416.2	537	$3.8 \cdot 10^{-5}$	543	$3.8 \cdot 10^{-5}$
$\mathrm{u\bar{u}}$	227.2	171	$7.6 \cdot 10^{-5}$	162	$7.2 \cdot 10^{-5}$
$\mathrm{d\bar{d}}$	56.8	34	$6.2 \cdot 10^{-5}$	42	$7.6 \cdot 10^{-5}$
$\mathrm{c\bar{c}}$	186.2	19	$1.1 \cdot 10^{-5}$	13	$7.5 \cdot 10^{-6}$
$\mathrm{s\bar{s}}$	51.9	14	$2.9 \cdot 10^{-5}$	4	$< 4.4 \cdot 10^{-6}$
$e^+e^-\mu^+\mu^-$	674.3	8	$< 3.4 \cdot 10^{-7}$	10	$1.6 \cdot 10^{-6}$
$\mathrm{B^+B^-}$	77.3	0	$< 3.0 \cdot 10^{-6}$	0	$< 3.0 \cdot 10^{-6}$
$\mathrm{B^0\bar{B}^0}$	73.0	0	$< 3.2 \cdot 10^{-6}$	0	$< 3.2 \cdot 10^{-6}$
$\mu^+\mu^-$	164.4	0	$< 1.4 \cdot 10^{-6}$	6	$< 1.4 \cdot 10^{-6}$
$\pi^+\pi^-\mathrm{ISR}$	4.5	48	$1.1 \cdot 10^{-9}$	47	$1.1 \cdot 10^{-3}$
$\mathrm{K^+K^-ISR}$	0.433	1	$< 5.3 \cdot 10^{-4}$	2	$< 5.3 \cdot 10^{-4}$
$\pi^+\pi^-\pi^0\mathrm{ISR}$	0.700	1	$< 3.3 \cdot 10^{-4}$	0	$< 3.3 \cdot 10^{-4}$
$\pi^+\pi^-\pi^0\pi^0\mathrm{ISR}$	1.041	1	$< 2.2 \cdot 10^{-4}$	2	$< 2.2 \cdot 10^{-4}$
$\mathrm{p\bar{p}ISR}$	0.016	1	< 0.01	1	< 0.01
$\mathrm{K^0\bar{K}^0ISR}$	0.237	0	$< 9.7 \cdot 10^{-4}$	0	$9.7 \cdot 10^{-4}$
$\pi^+\pi^-\eta\mathrm{ISR}$	0.010	0	< 0.02	0	< 0.02
$\pi^+\pi^-\pi^+\pi^-\mathrm{ISR}$	0.686	0	$< 3.4 \cdot 10^{-4}$	0	$< 3.4 \cdot 10^{-4}$
$e^+e^-\pi^+\pi^-$	271.4	2	$< 8.48 \cdot 10^{-7}$	9	$< 8.5 \cdot 10^{-7}$
$e^+e^-\mathrm{K^+K^-}$	11.4	1	$< 2.0 \cdot 10^{-5}$	1	$< 2.0 \cdot 10^{-5}$
$e^+e^-\mathrm{p\bar{p}}$	1.7	1	$1.4 \cdot 10^{-4}$	0	$< 1.4 \cdot 10^{-4}$
$e^+e^-\tau^+\tau^-$	2.6	1	$< 8.8 \cdot 10^{-5}$	4	$< 8.8 \cdot 10^{-5}$
$\mu^+\mu^-\mu^+\mu^-$	0.044	0	$< 5.1 \cdot 10^{-3}$	0	$< 5.1 \cdot 10^{-3}$
$\mu^+\mu^-\tau^+\tau^-$	0.015	0	< 0.02	0	0.02

Table 6.2.1: For each MC subprocess and the experimental data, the number of generated particles, the number of reconstructed particles for an electron and a positron tag, and the reconstruction efficiency are given. $<$ indicates that the upper limit was calculated for the reconstruction efficiency using Equation (6.2.1). The uncertainties on the reconstruction efficiencies are not given as they are negligibly small.

From the reconstruction efficiencies and the integrated luminosities of the different processes, the expected ratio of each background over the $ee(\gamma)$-signal is determined. This ratio quantifies whether the process needs to be included in the final analysis or if it is negligible. We neglect a process in the final data sample, if

$$\mathcal{R} = \frac{N_\mathrm{b}}{N_\mathrm{s}} < 0.001 \tag{6.2.2}$$

where N_b is the number of expected background events and N_s the number of expected signal events, both determined from

$$N_i = \mathcal{L}_\mathrm{s} \frac{N_{\mathrm{gen},i}}{\mathcal{L}_i} \epsilon_i. \tag{6.2.3}$$

The number of expected events N_i and the background-to-signal ratio are determined separately for both tag charges. From these calculations, only one background process is identified as non-negligible, this is the process of $e^+e^- \to \gamma\gamma$. All other background processes are negligible for our efficiency calculation. The respective background to signal ratios for the non-negligible process are $\mathcal{R}^{\gamma\gamma}_\mathrm{neg} = 0.094$ and $\mathcal{R}^{\gamma\gamma}_\mathrm{pos} = 0.088$.

We conclude that the decay reconstruction works well, and also the non-Bhabha background coming from $\gamma\gamma$ contributes only with less than 10 %.

6.3 Data-MC-comparison

To close up the analysis of the final dataset, a comparison between the data and MC for their individual variables is performed. For the comparison of experimental data and MC, another weighting factor is needed. All integrated luminosities of MC are scaled to the integrated luminosity of the experimental data set.

The luminosity of the experimental data sample used for the final analysis is $\mathcal{L}_\mathrm{data} = 357.3\,\mathrm{fb}^{-1}$. Here, the full available luminosity is used. For the MC data sample, only 40 % of all events are used for the calculations, as 60 % of this data is used to train the BDT.

We compare data and MC separately for reconstructions with electron (blue) and positron (red) tags, but present both in the same plots to make a reasonable comparison between both charges. Further, for the comparison, the MC data is analyzed as a complete dataset without distinguishing between the different subprocesses. A more detailed distinction is not necessary, as the majority of the MC events come from $ee(\gamma)$-signal and only some events come from $\gamma\gamma$-processes. Our main focus for this comparison is on the equalities and differences between the complete MC set and the experimental data.

To perform not only a qualitative comparison but also a quantitative one, the ratio of data over MC is determined by

$$\text{Data/MC} = \frac{N_{\text{Data}}}{N_{\text{MC}}}. \tag{6.3.1}$$

For calculating the uncertainties on the ratios, the uncertainty on the number of events is determined for a dataset x by

$$\Delta_x = \sqrt{\sum w_x^2} \tag{6.3.2}$$

using the luminosity weights w_x of each event. For the MC, we use different weighting factors for the different subprocesses, depending on the event type considered, while for data, the weighting factor for each event is one, and so the uncertainty is just $\Delta_{\text{data}} = \sqrt{N_{\text{data}}}$.

Using Gaussian error propagation, the uncertainty on the data-simulation ratio is determined as

$$\Delta\text{Data/MC} = \text{Data/MC}\sqrt{\left(\frac{\Delta_{\text{data}}}{N_{\text{data}}}\right)^2 + \left(\frac{\Delta_{\text{MC}}}{N_{\text{MC}}}\right)^2}. \tag{6.3.3}$$

We calculate the ratio and its uncertainty binwise, using the same binning scheme for data and MC. All data and simulation comparison plots are shown in Figure 6.3.1 to Figure 6.3.5.

For the cluster energy of the tag electron, both tag categories agree well between 2 GeV and 6 GeV. Below this range, the MC exceeds the experimental data for both tags, while at higher energies, an excess of events in data is observed for the positron tag and an excess of MC events for the electron tag. The distributions for the momentum of the tag electron are, as expected, similar to those of the cluster energy. The major differences are the slightly larger uncertainties on the ratio at the edges of the distributions and that for more than 4 GeV, the positron tag has an excess of data events. Overall, both variables have a solid agreement.

The polar angle distributions agree well in the ECL barrel, ranging from 0.56 rad to 2.24 rad. Outside of the ECL barrel, the agreement is rather poor, where larger discrepancies close to the gaps of the ECL and at the edges of the acceptance of the ECL endcaps occur. For the azimuth of the tag electron, all ratios have values very close to the ideal value of one. This lets me conclude that the overall luminosity scaling is correct, and data and MC distributions should be in good agreement for most of the variables.

We use the variables clusterE1E9 and clusterE9E21 for particle identification [10]. The clusterE1E9 variable returns the ratio of the energy of the central crystal over the energies of the 3×3 crystals around this central crystal. The ratio should be larger for photons, as they hit the ECL in a more compact way than for hadrons. The clusterE9E21 variable takes the energy sum of the 5×5 crystals around the central crystal without the corner crystals and relates it to the energy of the central crystal. The goal of those variables is to distinguish photons from hadrons, especially important are those variables for our two energy clusters.

The clusterE1E9 variable of the tag electron has a rather good agreement from 0.3 to 0.85, but the ratio drops significantly for larger values. For the clusterE9E21 variable, no trend is identified, except for an overall bad agreement. As we are, by the definition of the selection criteria for the tag electron, already quite sure to only select electrons as tag particles, those variables were removed from the training variables of the BDT and are considered less important for our analysis.

For the higher-energy cluster, the uncertainties on the data-MC ratio for clusterE1E9 are very large for values below 0.3. While from 0.3 to 0.85, data and MC are in very good agreement with smaller uncertainties on the ratio. Only for larger values, an excess of MC events is identified. The clusterE9E21 variable is from 0.8 to 0.96 within the range of good agreement between data and MC. While the agreement fluctuates at lower values, due to the low number of events in those bins. For the bins above 0.96, both tag configurations show an excess of MC events. I conclude that the overall agreement for both PID variables of the higher energy cluster is within the range of good agreement.

For the clusterE1E9 variable of the lower-energy cluster, all bins up to 0.85 are in good agreement. After this, the data-MC ratio values drop slightly. For the clusterE9E21 variable of the lower-energy cluster, the agreement is good for values between 0.9 and 1.0. For most of the bins at smaller values, an excess of MC events is found. But the disagreement is not so important here, as those bins have larger statistical uncertainties.

The data-MC ratio of the cluster energy over track momentum of the tag electron shows an interesting behavior. For both tags, the bins below 1 show an excess of experimental data events, while the bins larger than 1 show an excess of MC events. At both edges, the statistical uncertainties get quite large due to the low number of events per bin. The overall agreement of this variable is bad, so it is not used for training the BDT. The same yields for the summed ECL cluster energies. Here, for values smaller than 10.4 GeV an excess of data events exists, while larger values have an excess of MC events with much larger statistical uncertainties for the bins above 10.8 GeV. I conclude that this variable is helpful to restrict the event selection in general, but its data-MC agreement is not sufficient for a usage in the BDT.

All data-MC ratio values of the beam background suppression variable have relatively small statistical uncertainties and fluctuate around the ideal value of one. Only the rightmost bin has a data-MC ratio close to one with a very small statistical uncertainty. Following a detailed analysis of the number of entries in the different bins, the rightmost bin contains more than 95 % of all events, separately compared for both tags of data and MC. Therefore, the data-MC disagreement in the other bins is negligible. For the fake photon suppression, the last bin also contains a majority of the events, but in this case, an excess of events in the data is found in this bin. Further, for all bins at smaller values, an excess of MC events is found. Overall, data and simulation disagree, and the variable is therefore removed from the set of training variables for the BDT.

For the polar angle of the lower-energy cluster, the agreement or disagreement between data and MC is similar for both tags and fluctuates slightly. Having an excess of events in data between 1 rad and 1.5 rad, and an excess of MC events for bins at values higher than 2 rad. The rest of the bins are within the 10 % range of agreement. For the azimuth of the lower-energy cluster, all bins have a very good agreement.

The cluster energy of the lower-energy cluster has, for the electron tag, a good agreement up to 5 GeV and for higher energies an excess of MC events. For the positron tag, the agreement is good up to 4 GeV and then it also has a slight excess of MC events, but smaller than for the electron tag. As the number of events per bin decreases for higher energies, the disagreement does affect the overall agreement only slightly, and I conclude that the data-MC agreement is good enough. The opening angle between the tag electron cluster and the lower-energy cluster has a data-MC ratio fluctuating around the range of good agreement without any specific trends for either the electron or the positron tag. Concluding that the fluctuations of the agreement are still within a reasonable range, this variable is still used in the BDT training. Further, it has to be considered that the variable was computed from measured values, and the event-by-event uncertainties of these calculations are not reflected in these comparison plots.

Also, for the opening angles between the higher-energy cluster and the tag cluster and between the higher- and the lower-energy cluster, the uncertainties on the calculation of the opening angles are not included in the data-MC comparison. The data-MC ratio for the opening angle between the higher- and the lower-energy cluster looks similar to the ratio distribution of the previously analyzed opening angle. I conclude that those fluctuations are also within an agreement which is acceptable for our analysis. For the opening angle between the tag cluster and the higher-energy cluster, the ratio is in good agreement for bins above 2 rad, but for smaller values an excess of MC events is identified. But as the bins at values above 2 rad have only a low population, this disagreement is neglected.

As expected, the data-MC agreement for the azimuth of the higher-energy cluster is good for the positron and electron tag. The number of ECL crystal hits of a cluster returns the summed weight of all crystals included in one ECL hit [10]. For the higher-energy cluster, the number of ECL crystal hits has a clear disagreement between data and simulation for all bins. This behavior is also identified for the number of ECL crystal hits for the lower-energy cluster and the tag electron cluster. Therefore, we exclude all those variables from the training of the BDT.

Further variables removed from the set of training variables of the BDT are the polar angle differences and the azimuthal differences between the individual clusters. The azimuth difference between the higher and the lower cluster and between the tag and the lower cluster behaves similarly. The data-MC ratio peaks at 2 rad and at 4.2 rad, indicating an excess of events in the experimental data. For the rest, the ratio is close to the range of good agreement with an overall slight excess of MC events.

The data-MC ratio of the difference of the azimuth between the tag cluster and the lower-energy cluster is from 1.8 rad to 2.8 rad and from 3.2 rad to 4.2 rad within the range of good agreement. On both edges, an excess of events in the data is identified, while for values around 3 rad a slight excess of events in MC is identified. As the largest deviations from the ideal ratio occur for poorly populated bins, the overall agreement is good. The polar angle difference between the tag cluster and the lower energy cluster agrees well between data and simulation up to 1.3 rad. For higher values, the positron tag shows an excess of MC events, while for the electron tag, the ratio fluctuates around one and shows an excess for both MC events and for data events.

The polar angle differences between the higher- and the lower-energy cluster and between the lower-energy cluster and the tag cluster show the same behavior. Where for values higher than 1.3 rad the excess of MC events starts to increase. Due to this clearly visible disagreement for certain ranges, the angle differences were replaced by the opening angles in the set of training variables.

To complete the data-MC comparison, the cluster energy and the cluster polar angle of the higher-energy cluster are analyzed. For the electron tag, the cluster energy shows a good agreement up to 6.1 GeV and for higher values, where the bins are only poorly populated, a slight excess of MC events is observed. For the positron tag, the agreement is similar except for the range between 3 GeV and 4 GeV, where a slight excess of data events is visible. Overall, the agreement is good enough to use this variable in the final binning of the tracking efficiency. For the polar angle, a good agreement for both tags is extracted. I conclude that both variables are suitable for a usage in the final tracking efficiency plots.

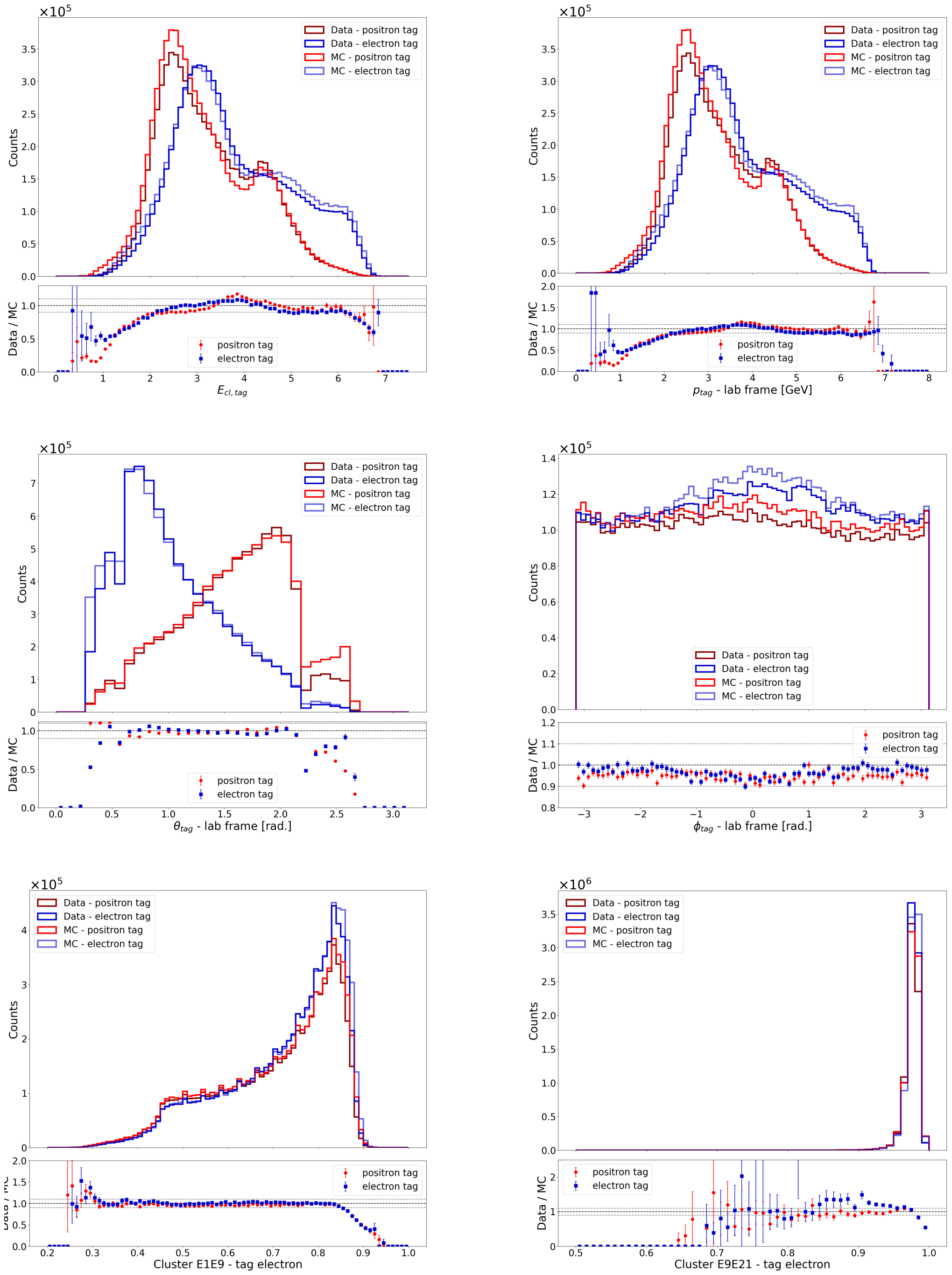

Figure 6.3.1: Data-MC comparison for positron and electron tags in different relevant variables. With the distributions for data and MC in the upper panels and the Data-MC ratio for each tag in the lower panels. The gray dashed lines indicate unity and a 10 % range around, visualizing good agreement.

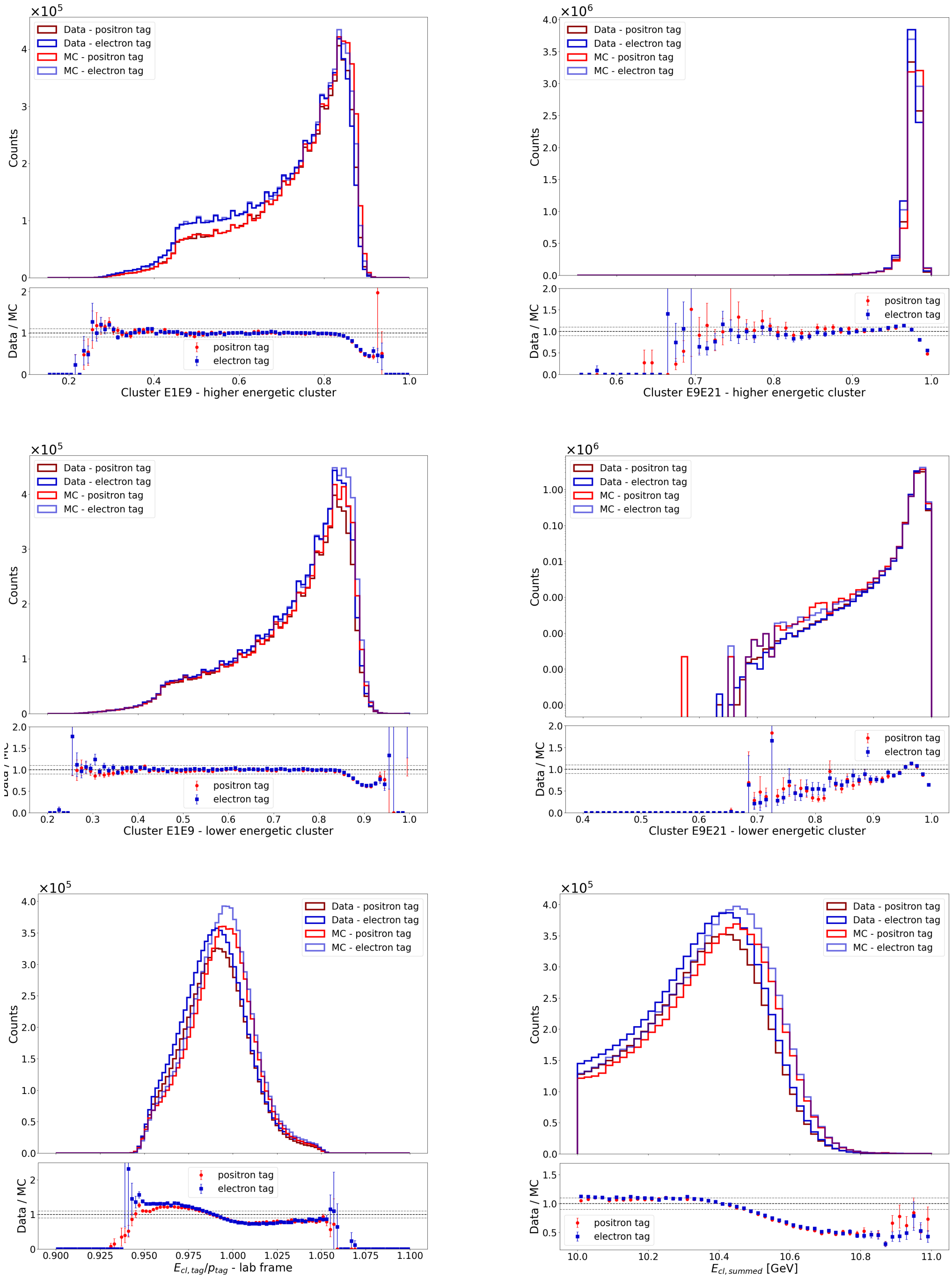

Figure 6.3.2: Data-MC comparison for positron and electron tags in different relevant variables (continued).

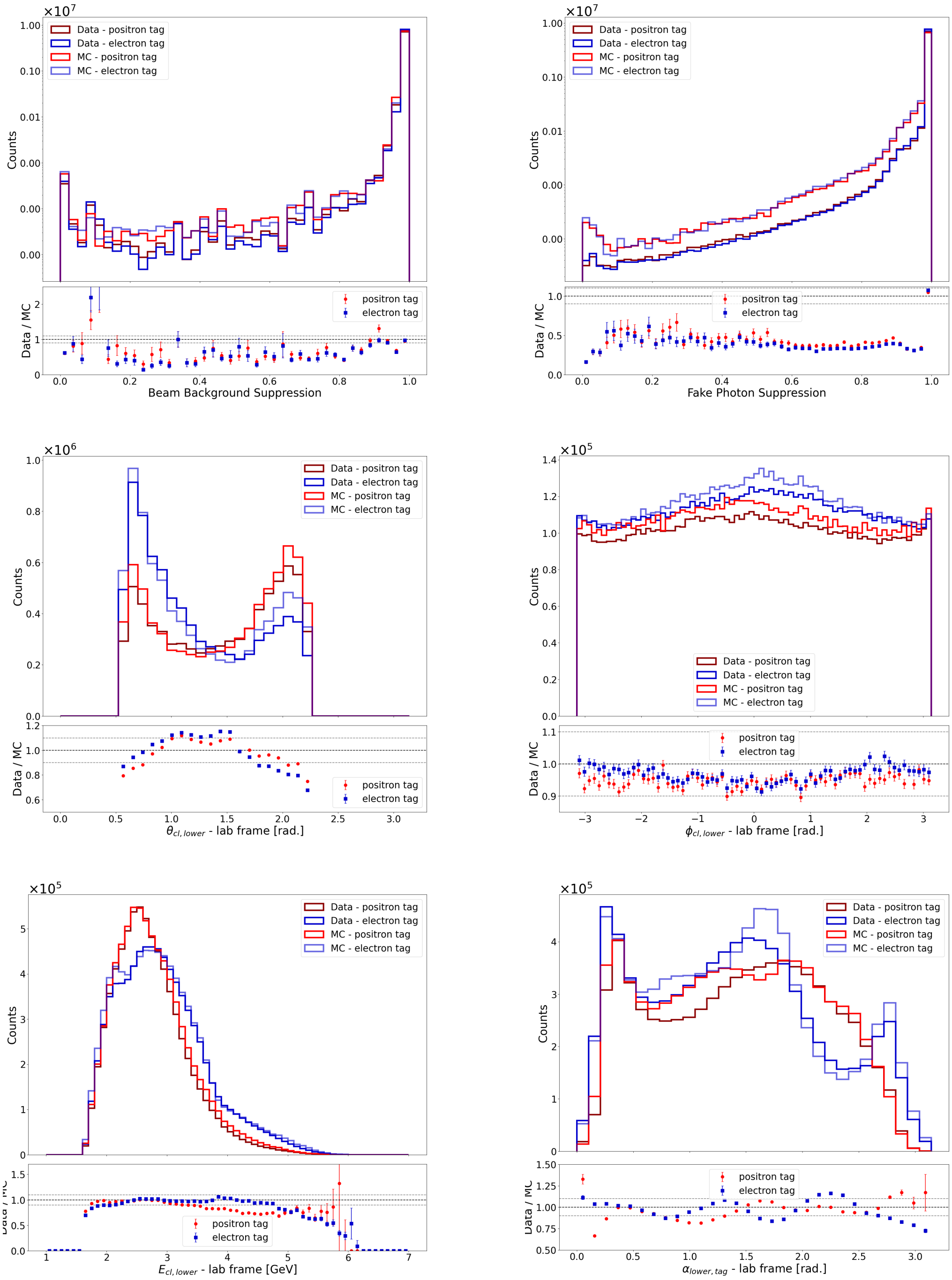

Figure 6.3.3: Data-MC comparison for positron and electron tags in different relevant variables (continued).

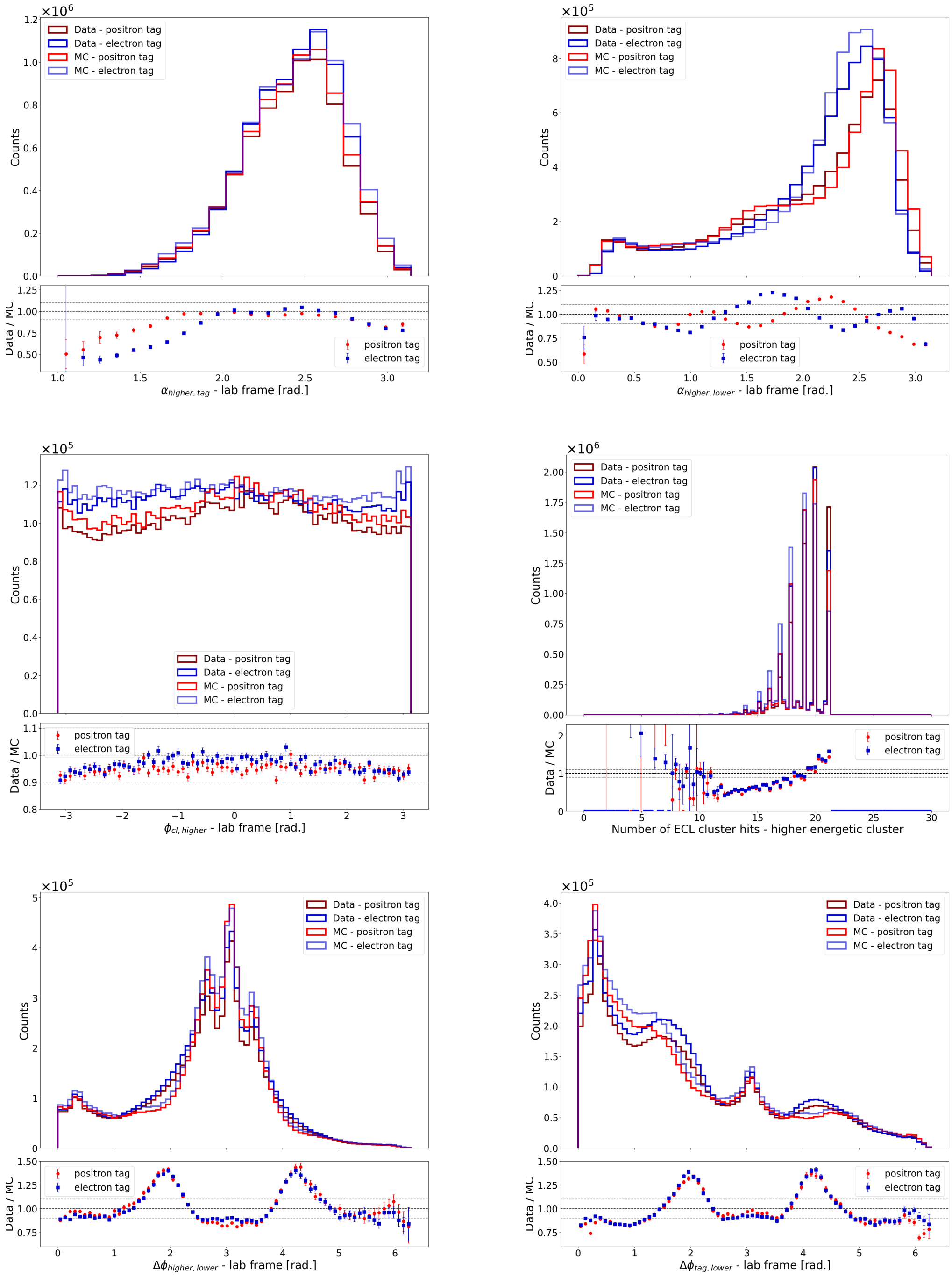

Figure 6.3.4: Data-MC comparison for positron and electron tags in different relevant variables (continued).

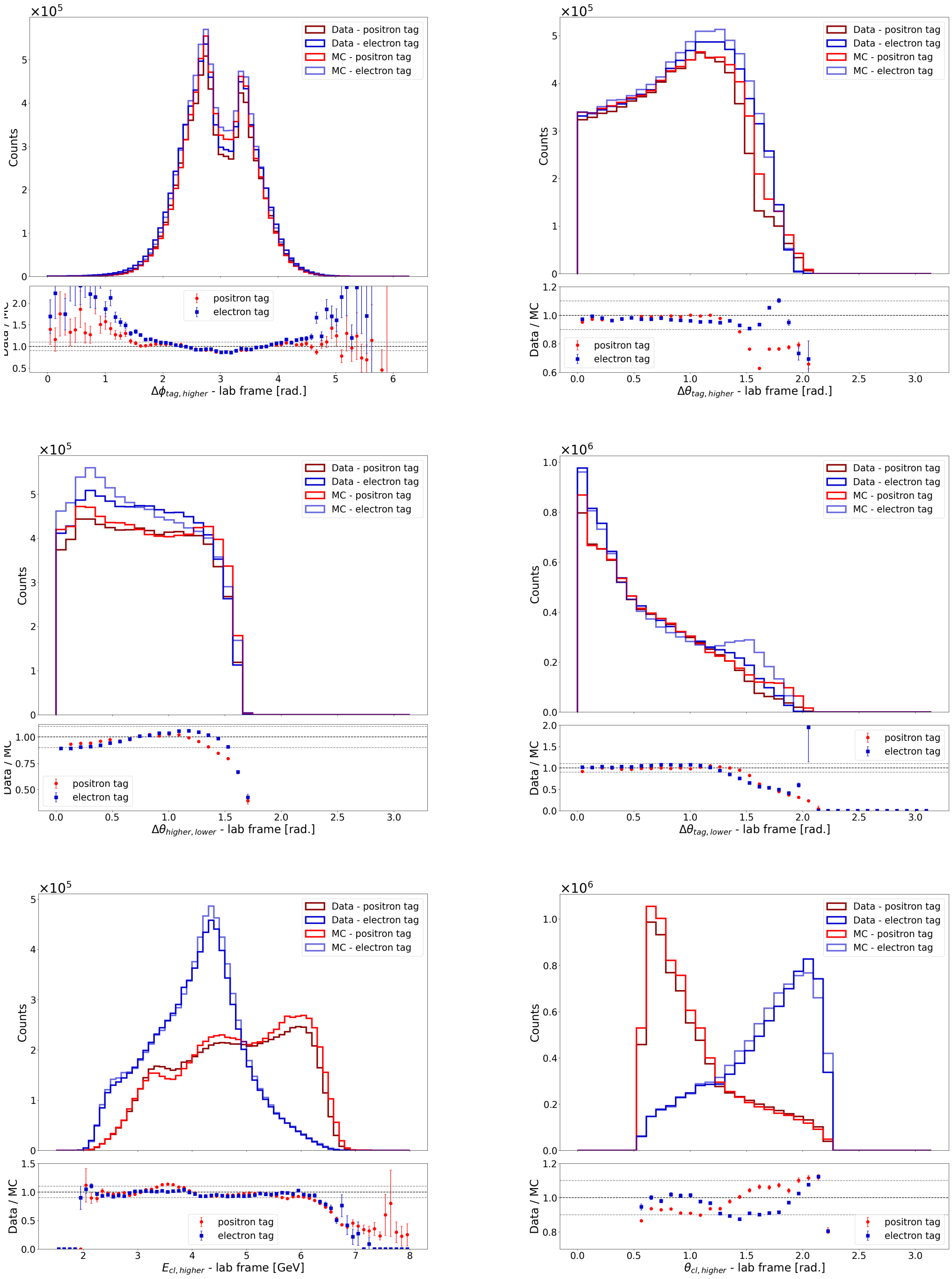

Figure 6.3.5: Data-MC comparison for positron and electron tags in different relevant variables (continued).

Chapter 7

Study of the tracking efficiency

7.1 Determination of the tracking efficiency

The tracking efficiency is calculated for each bin individually using the phase space distribution of the ECL cluster energy and the cluster polar angle of the higher-energy cluster. For a better understanding of the final results, the population of each bin in the phase space plot is evaluated first.

Overall, for both the electron and the positron tags in MC and data, the population of the phase space is similar, as shown in Figure 7.1.1. These plots map the performance of the tracking algorithm and the detector. For the electron tag of the MC (top-left), a larger number of events is located in the backward direction of the detector, close to the right edge of the ECL barrel (red line on the right). This is expected as the probed particle is a positive particle, and following the geometrical setup of Belle II, the majority of positively charged particles hit the detector in the backward direction, while the majority of negatively charged particles hit the detector in the forward direction. For the positron tag of the MC (top-right), we probe a negatively charged particle, and a peak of events is, as expected, located in the forward direction at 0.7 rad, close to the other end of the ECL barrel (left red line). For both tags, the bins with the least population are those with the highest energy in each angle bin. Further, there is only one bin with a cluster energy of 1 GeV to 2 GeV. The other bins in this energy range are not populated as the cuts applied due to the trigger selection remove all events with a cluster energy in the CM-frame smaller than 2 GeV. The phase space is shown here in the lab frame, and therefore, a few events remain in this energy region. Leading to one poorly populated bin in this energy range.

The population for the experimental dataset is shown in the lower panels of Figure 7.1.1. The coverage of the phase-space for MC and experimental data is similar. The major difference is that the bins with a high population include a significantly larger number of events. This is expected as the overall integrated luminosity of the experimental dataset is significantly larger. The most energetic bins for each polar angle bin are also poorly populated here. The number of events in those bins is mostly only

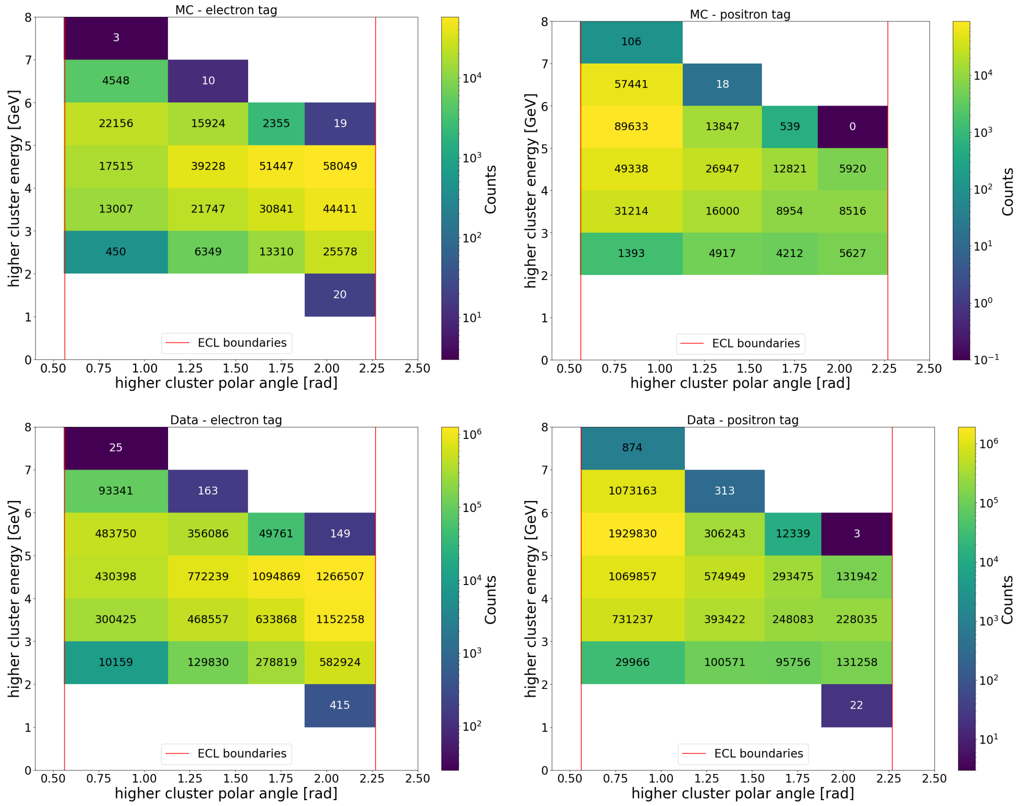

Figure 7.1.1: The number of events per bin as a function of the cluster energy and the polar angle of the higher-energy cluster, shown for the final binning scheme.

slightly larger in data than in simulation. Due to the low population in the mentioned bins, we expect larger statistical uncertainties there. Therefore, we exclude all events with $E_{\mathrm{cl}} < 2\,\mathrm{GeV}$ or $E_{\mathrm{cl}} > 7\,\mathrm{GeV}$ from the final analysis. Furthermore, the polar angle cut based on the trigger selection removes all events outside the ECL barrel.

For the calculation of the tracking efficiency, Equation (2.2.5) is used. Its statistical uncertainty is determined by taking the square root of Equation (2.2.8). Both are determined binwise based on the binning of the phase-space plot. The charge-dependent result is achieved by determining the tracking efficiency for each tag separately. This approach to calculating the efficiency is naive and does not include the effect of any background contributions.

The tracking efficiency for each bin in the phase space plot has, for all datasets, relatively uniform distributions. For the electron tag, the majority of bins in MC have a value larger than 99 %. For the data, the majority of bins have a value larger than 98 %. The distribution of the tracking efficiency values is shown in Figure 7.1.2, and the corresponding uncertainty distributions in Figure 7.1.3. As expected, well-populated bins exhibit small statistical uncertainties, whereas less populated bins show larger uncertainties. When these uncertainties are taken into account, the

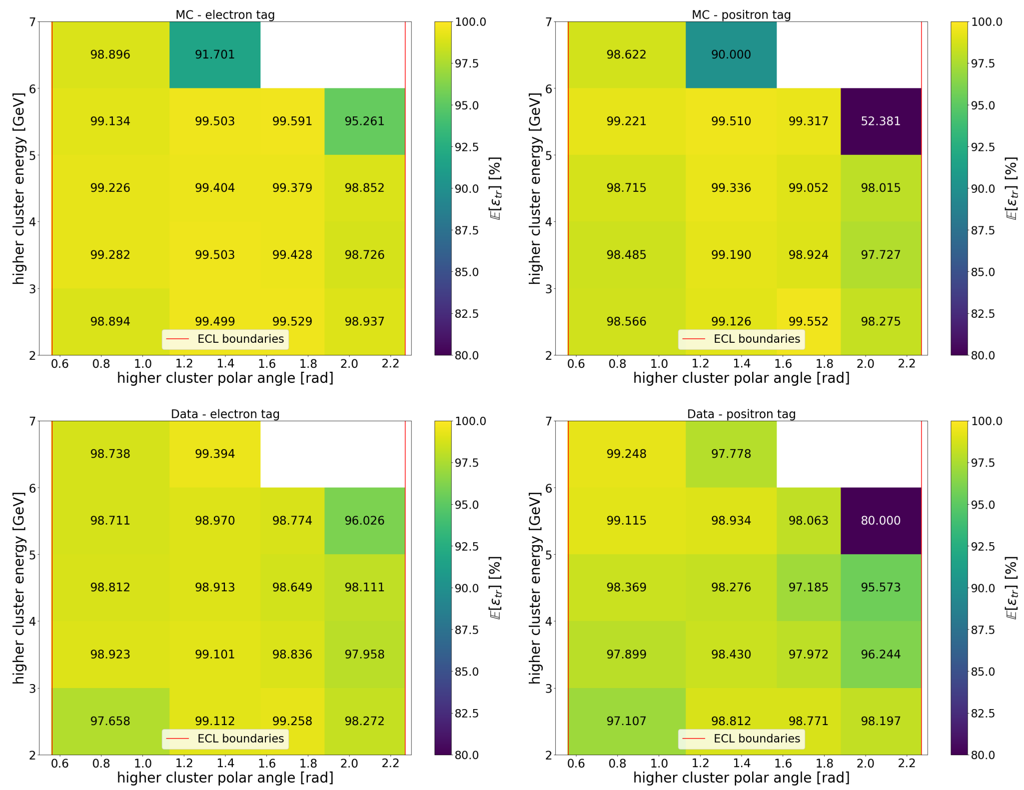

Figure 7.1.2: Tracking efficiencies as a function of the cluster energy and cluster polar angle. Calculated for MC in the upper row and for experimental data in the lower row, separately for electron (left) and positron (right) tags.

tracking efficiency measurements in the low-population regions are consistent with those of the well-populated bins. For the positron tag – so the tracking efficiency for negatively charged particles – the values of the tracking efficiency are overall slightly lower than for the electron tag.

The statistical uncertainties of the MC dataset are mainly below 0.1 %, while for the data, the statistical uncertainties are mainly below 0.05 %. This difference in the uncertainties by a factor of two comes from the available amount of data. This difference would lead to uncertainties in data being five times smaller than uncertainties in MC. The current improvement of the uncertainties, which is evaluated only by eye, by a factor of two for data with respect to MC is just an upper limit for the uncertainty.

The charge-dependent determination of the tracking efficiency is used to extract whether the tracking algorithm performs differently based on the charge of the probed particle. The charge-dependent comparison is done by determining the efficiency asymmetry

$$\mathcal{A}^x_{\epsilon_{\mathrm{tr}}} = \frac{\epsilon^x_+ - \epsilon^x_-}{\epsilon^x_+ + \epsilon^x_-}. \tag{7.1.1}$$

The signs describe the charge of the tag particle, and x refers to calculating the asymmetry for data or MC. The uncertainty on this asymmetry is

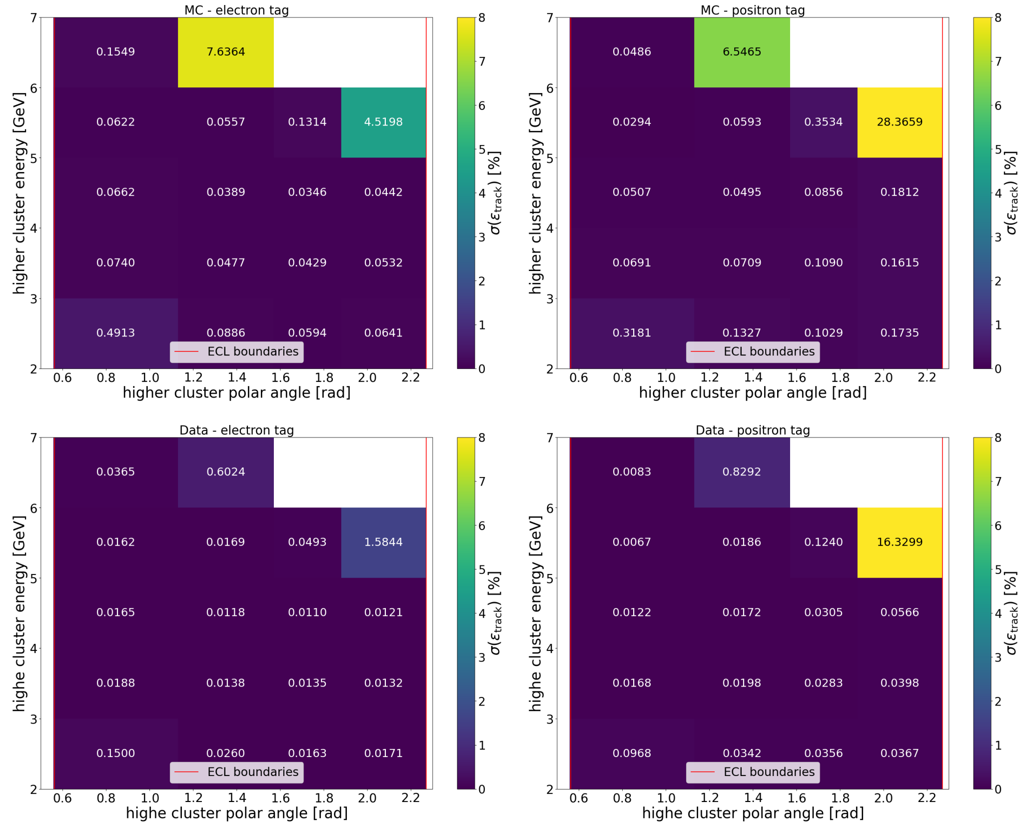

Figure 7.1.3: Statistical uncertainties on the tracking efficiency binned in cluster energy and cluster polar angle. Calculated for MC in the upper row and for experimental data in the lower row, separately for electron (left) and positron (right) tags.

determined by using the statistical uncertainty of the tracking efficiencies and the Gaussian error propagation

$$\Delta\mathcal{A}^{x}_{\epsilon_{\mathrm{tr}}} = \sqrt{\left(\frac{\partial\mathcal{A}_{\epsilon_{\mathrm{tr}}}}{\partial\epsilon_{+}}\sigma_{+}(\epsilon_{+})\right)^{2} + \left(\frac{\partial\mathcal{A}_{\epsilon_{\mathrm{tr}}}}{\partial\epsilon_{-}}\sigma_{-}(\epsilon_{-})\right)^{2}}. \tag{7.1.2}$$

Charge independence is achieved when the asymmetry is equal to zero. In this case, the tracking algorithm is insensitive to charge differences. A positive asymmetry indicates that the tracking algorithm and the detector perform better for negatively charged tracks, while a negative asymmetry indicates a better performance for positively charged tracks. In our case, the majority of the asymmetries are negative (Figure 7.1.4).

The charge dependence is assumed to depend on the direction of the particle within the detector. To check this, the tracking efficiency is calculated for polar angle bins, neglecting the difference in the cluster energies of the particles. The charge asymmetry in data and MC for those polar angle bins is shown in Figure 7.1.5.

The charge asymmetry in data has a linear behavior, with nearly no asymmetry in the forward direction of the experiment. The asymmetry

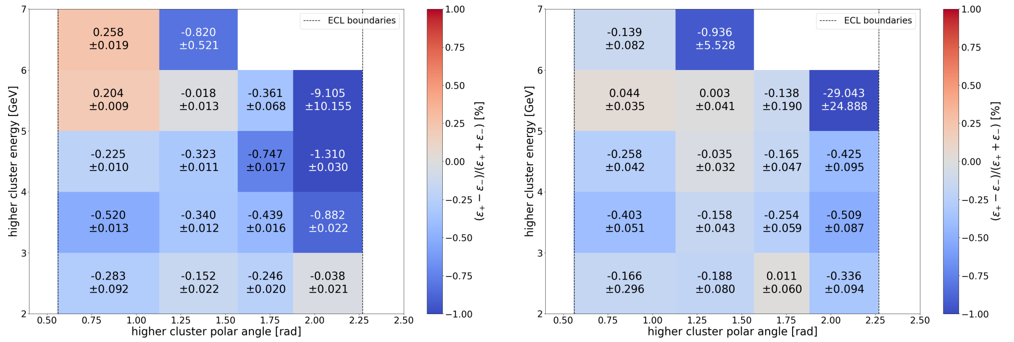

Figure 7.1.4: Charge Asymmetry binned in cluster polar angle and cluster energy for MC (left) and for data (right).

decreases slowly towards more negative values, ending at a minimum of $-0.8\,\%$.

For MC, the charge asymmetry behaves slightly different. The first three bins in the ECL barrel have values between $-0.1\,\%$ and $-0.2\,\%$. Only the last bin in the backward direction has a lower value of $-0.4\,\%$. Overall, the algorithm performs slightly better for positively charged tracks than for negatively charged tracks. This lets us conclude that the detector behavior is not perfectly modeled in the simulation and that the tracking algorithm works slightly better on positively charged tracks.

Further interesting for any charge-dependent or asymmetry studies is the charge asymmetry between the forward and the backward direction of the detector. The asymmetry is then calculated by

$$\mathcal{A}^{x}_{\epsilon,\mathrm{FB}} = \frac{\epsilon^{x}_{\mathrm{F}} - \epsilon^{x}_{\mathrm{B}}}{\epsilon^{x}_{\mathrm{F}} + \epsilon^{x}_{\mathrm{B}}}. \tag{7.1.3}$$

In data, this asymmetry is $\mathcal{A}^{\mathrm{data}}_{\epsilon,\mathrm{FB}} = (0.269 \pm 0.004)\ \%$ and in MC this asymmetry is $\mathcal{A}^{\mathrm{MC}}_{\epsilon,\mathrm{FB}} = (0.13 \pm 0.02)\ \%$. From these values, it is extracted that the tracking efficiency is higher in the background direction of the detector than in the forward direction. This indicates that either the reconstruction of charged particles works better in the backward direction or that the tracking algorithm performs better for particles in this region. This comparison gives an indication of the behavior in the forward and backward directions, but to use this result further, it has to be calculated in the CM-frame.

As most positively charged particles hit the detector in the backward region, the charge asymmetry values confirm that the positively charged tracks are better reconstructed than the negatively charged tracks.

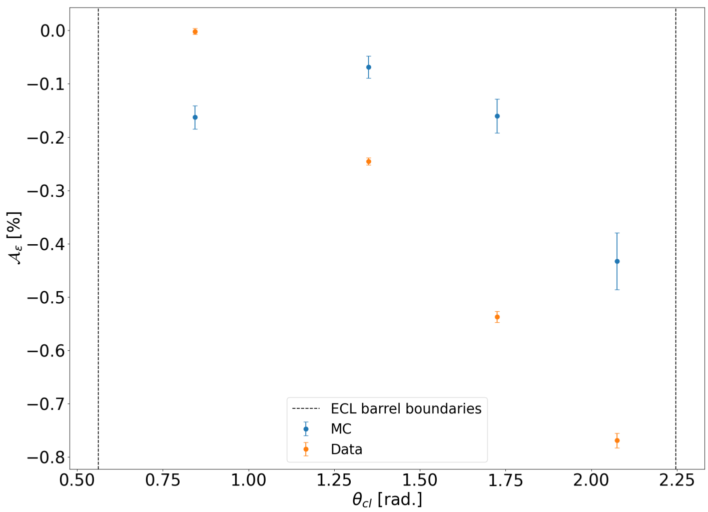

Figure 7.1.5: Tracking efficiency asymmetry per polar angle bin separately for MC and experimental data.

7.2 Correction factor

As already seen in the comparison of the charge asymmetries, there are differences between the performance of data and MC when it comes to the tracking efficiency. To account for those differences between data and MC, a correction factor is calculated separately for the different charges,

$$C_{\pm} = \frac{\epsilon_{\pm}^{\text{data}}}{\epsilon_{\pm}^{\text{MC}}}. \tag{7.2.1}$$

The statistical uncertainty on the correction factor is calculated by using the statistical uncertainties of the tracking efficiencies

$$\Delta C_{\pm} = C_{\pm} \cdot \sqrt{\left(\frac{\sigma(\epsilon_{\pm}^{\text{data}})}{\epsilon_{\pm}^{\text{data}}}\right)^2 + \left(\frac{\sigma(\epsilon_{\pm}^{\text{MC}})}{\epsilon_{\pm}^{\text{MC}}}\right)^2}. \tag{7.2.2}$$

The uncertainty of the correction factor is especially important, as it determines the precision of the correction. A small statistical uncertainty indicates a more reliable correction, while a large statistical uncertainty introduces systematic uncertainty into other analyses using charged particles and relying on the tracking efficiency.

A value of one for the correction factor means that the simulation reproduces the particle tracking in data very well, while deviations from one give the indication of a not perfect simulation of the MC processes. In these cases, a correction for the disagreement between data and MC is necessary. To visualize this even better, the plots in Figure 7.2.1 show the difference of the correction factor from one

$$\delta_{\pm} = 1 - C_{\pm}. \tag{7.2.3}$$

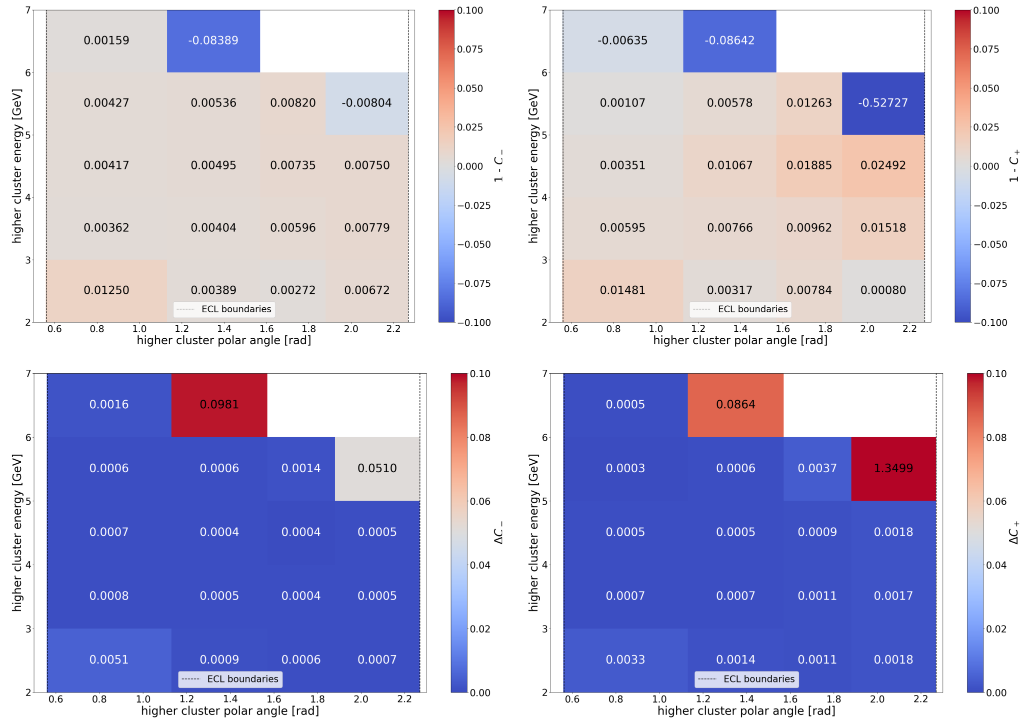

Figure 7.2.1: The difference of the correction factor between data and MC from one (top row) and the correction factor uncertainties (bottom row) binned in the phase space of the higher-energy cluster for an electron tag (left) and a positron tag (right).

The correction factors calculated in this thesis are all slightly smaller than one, as δ is larger than 0. This means that the tracking efficiency is overestimated in MC. All uncertainties on the correction factor for bins with sufficient statistical population are on a per mille or even sub-per-mille level, as seen in the lower panels of Figure 7.2.1. The current recommended uncertainty to use per track is 0.27 %, so we are in a similar range with our result.

To determine whether there is a residual charge-dependent bias between data and MC, the asymmetry between the correction factors is calculated

$$\mathcal{A}_{\mathrm{C}} = \frac{C_+ - C_-}{C_+ + C_-}. \tag{7.2.4}$$

An ideal asymmetry value would be zero, indicating that the charge-dependent efficiency is correctly modeled in the MC dataset.

All asymmetry values are on a sub-percent level, while their uncertainties are on a sub-per-mille level. In some cases, the asymmetry and its uncertainty are of the same order of magnitude. Similar to the calculations before, the results for the bins with very low population deviate from the rest of the bins. The asymmetry, shown in Figure 7.2.2, is negative for most of the bins. This means that in MC, ϵ_+ is underestimated compared to ϵ_-. For the few bins with a positive asymmetry value, this yields in the

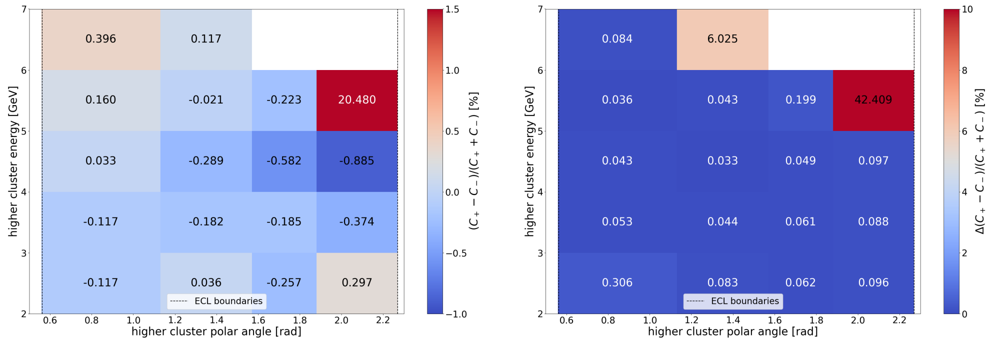

Figure 7.2.2: Asymmetry of the correction factors (left) and the corresponding uncertainty (right) binned in the phase space of the higher-energy cluster.

opposite way, so that the positive tracks need a larger correction than the negative ones.

7.3 Removing background contributions

The naively calculated tracking efficiency and its statistical uncertainty are based on all events predicted by the BDT to be in the higher-energy class. But this set of events also contains those events which are not radiative Bhabha scattering events.

For MC, the purity of the final dataset is known and analyzed by using MC truth information. For this analysis, the final dataset, containing only events predicted to be in the higher-energy class, is split into the three different categories of the true classes. We then calculate the tracking efficiency separately for each of the three groups.

The efficiency for the events being by prediction and by MC truth in the higher-energy class is assumed to be very high, as this group of events mainly contains correctly reconstructed radiative Bhabha events, which have exactly two tracks per event. Figure 7.3.1 shows the tracking efficiency for the different subsets of the final dataset as a function of the background fraction of the overall dataset. We determine the background fractions based on different cuts on the probability of the higher-energy class. I extract that the tracking efficiency for events truly being in the higher-energy class (orange) is independent of the background fraction, and the efficiency is larger than 99.99 %.

Events predicted to be in the higher-energy class but being in the lower-energy class by MC truth also have a high tracking efficiency, as the naive efficiency calculation is only targeting events with two tracks, independent of where the track is going. The distinction of where a track is going is made, in this thesis, by using the predictions from the BDT, so events truly being in the lower-energy class are events where the charged particle ends up in the wrong bin. For a larger background fraction, the

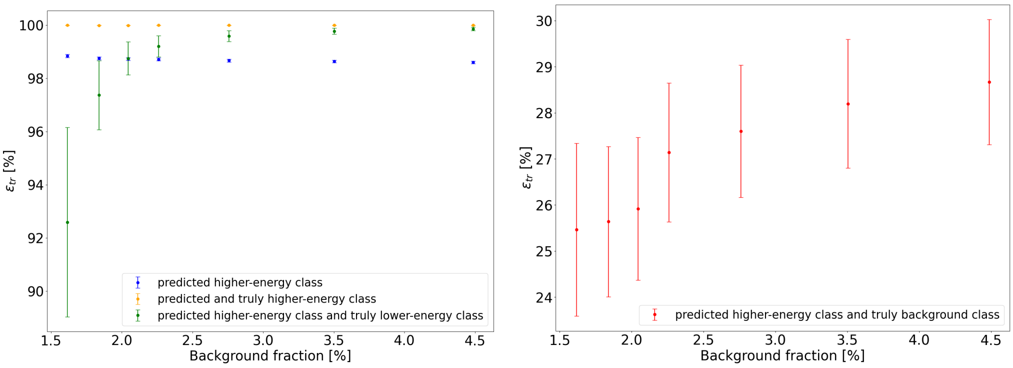

Figure 7.3.1: Tracking efficiency for the positron tag of MC for the bin going from 0.56 rad to 1.13 rad and from 4 GeV to 5 GeV, depending on the different overall background fractions. The tracking efficiency was calculated for various class combinations, with the higher efficiency range on the left and the lower efficiency range on the right.

efficiency of those wrongly binned events is similarly high as the efficiency for events truly assigned to the higher-energy class. But as the fraction of those events decreases, the efficiency (green) decreases towards 92 %. For the last two values, the statistical uncertainties also increase significantly.

The fraction of events being predicted to be in the higher-energy class, but being in the background class by MC truth (red), stays approximately at 1.7 % for all different cuts on the BDT probability. The tracking efficiency for this class of background events is 27 %. This low efficiency value is expected as the events in this class are not radiative Bhabha events, and therefore they don't necessarily have two tracks per event. But, this low efficiency contributes to the final tracking efficiency result, so that the final value is slightly lower than the efficiency of the events correctly assigned by MC truth to the higher-energy class.

I conclude that the final tracking efficiency result includes impurities originating from background processes, and those are not accounted for in the naive calculation of the tracking efficiency. To include those impurities for data and MC, we need to quantify all impurities. We do that by constructing a confusion matrix independent of MC truth information, so that we get a confusion matrix for MC and for data.

This confusion matrix is constructed by defining the three different BDT classes using tracking information. With those class definitions, we can only extract events with exactly two tracks for the higher and lower classes, as we do not know for which events the tracking algorithm failed. Therefore, our confusion matrix includes only the higher and the lower class events of the set K, which contains only events with two tracks. The confusion matrix is unknown for the set N containing all events with one or two tracks.

C^K_{ij} is the number of events in the set K having predicted category j and true category i, so each entry in the confusion matrix is described

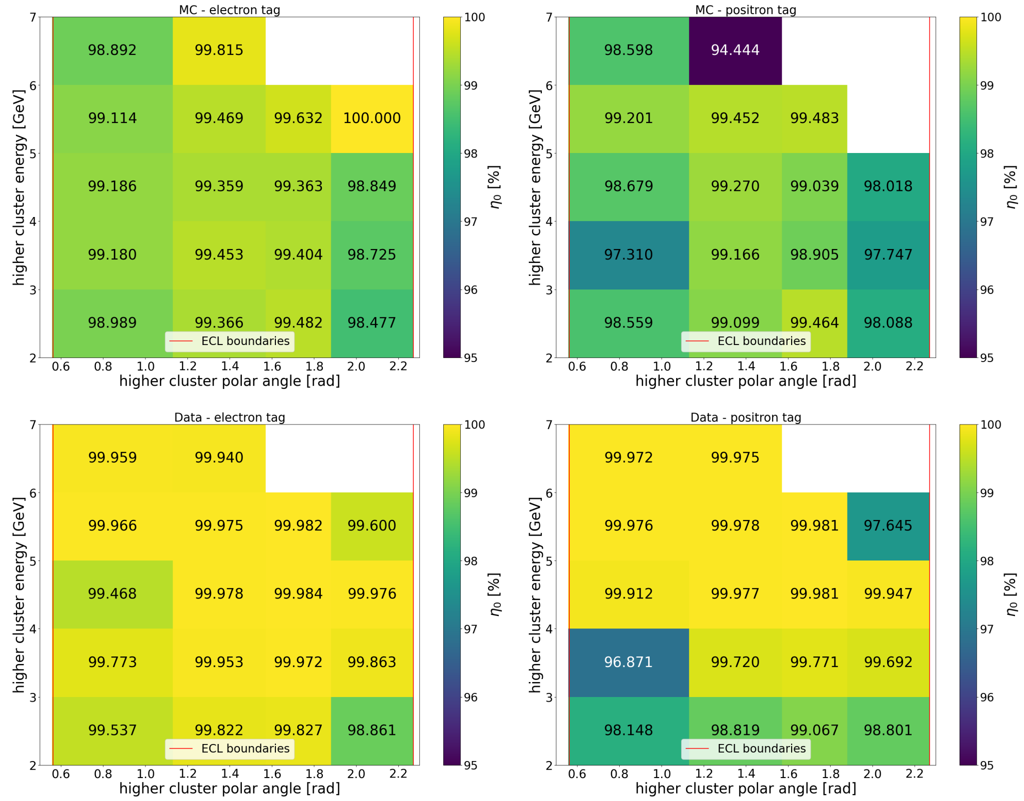

Figure 7.3.2: Unfolded tracking efficiency as a function of the cluster energy and the polar angle of the higher-energy cluster, shown for the final binning scheme, with MC results in the upper row and data results in the lower row, for the electron (left) and positron (right) tags.

by C_{ij}^K. The total number of events having true category i is T_i^X, where X is any set of events. The total number of events in a set X that have predicted category j is P_j^X.

With this knowledge, we determine the fraction of events with predicted category j and true category i as

$$f_{ij}^X \equiv \frac{C_{ij}^X}{P_j^X}. \tag{7.3.1}$$

where we do not know f_{ij}^X when C_{ij}^X is unknown. But as the BDT is not based on tracking information, we assume that our BDT performs identically for events in K and for events that are not in K, so we anticipate

$$f_{ij}^N \approx f_{ij}^K. \tag{7.3.2}$$

Using this approximation, we determine the number of events having true category i as

$$T_i^N \approx \sum_j f_{ij}^K P_j^N = \sum_j C_{ij}^K \epsilon_j^{-1}, \tag{7.3.3}$$

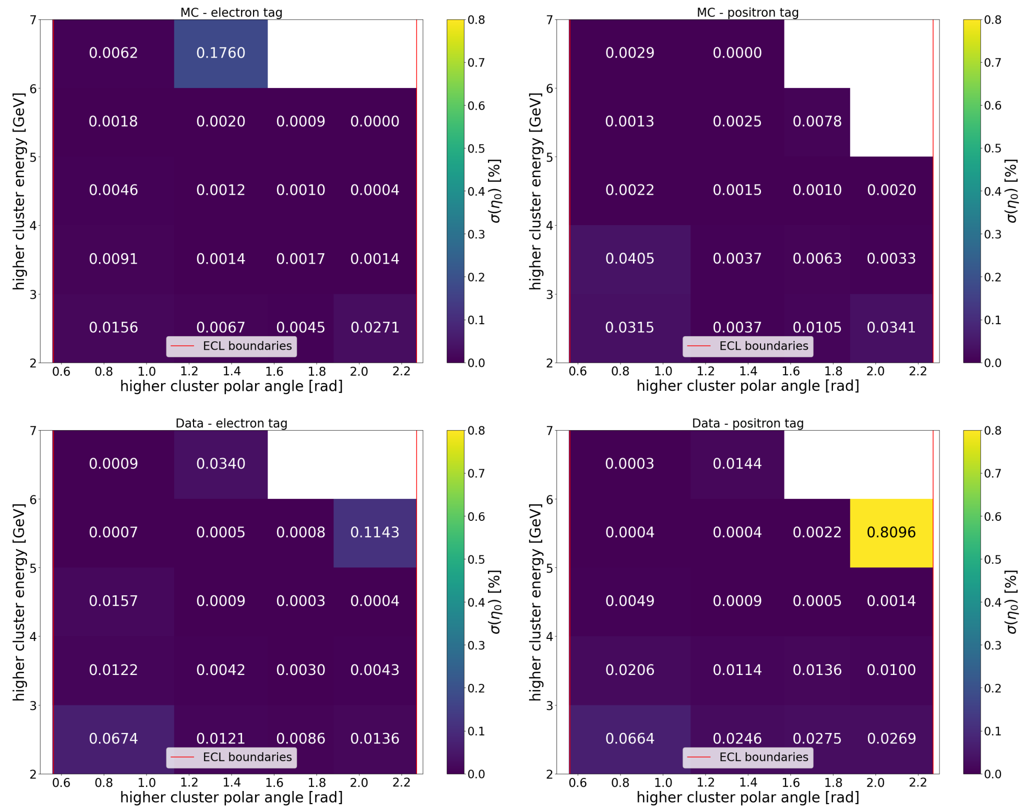

Figure 7.3.3: Statistical uncertainties of the unfolded tracking efficiency as a function of the cluster energy and the polar angle of the higher-energy cluster, shown for the final binning scheme, with MC results in the upper row and data results in the lower row, for the electron (left) and positron (right) tags.

where $\epsilon_j = \frac{P_j^K}{P_j^N}$ is naively calculated tracking efficiency. The tracking efficiency for any true category i, including the correction for any impurities, is

$$\eta_i = \frac{\sum_j C_{ij}^K}{\sum_j C_{ij}^K \frac{P_j^N}{P_j^K}} = \left[\left\langle \eta_j^{-1} \right\rangle_{C_{0j}^K} \right]^{-1}, \tag{7.3.4}$$

which is the unfolded approach to the efficiency calculation.

In Figure 7.3.2, we use this equation to calculate the corrected tracking efficiencies for all four datasets for the higher class $i = 0$. An important difference from the previously calculated efficiency is that the population of the bins changes slightly. For the naive efficiency calculation, only events predicted to be in the higher-energy class are used. While Equation (7.3.4) uses all events truly being in the higher-energy class, which means also events not predicted in $j = 0$ are used for the calculation. This difference in the events used for the calculations explains the different population of the bins.

The unfolded tracking efficiencies for each bin in the phase-space plot,

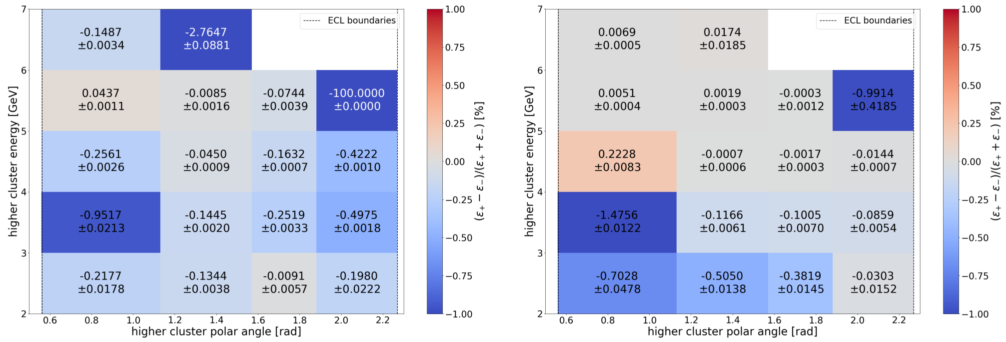

Figure 7.3.4: Charge Asymmetry binned in cluster polar angle and cluster energy of the higher-energy cluster for MC (left) and for data (right).

in Figure 7.3.2, have, for all datasets, relatively uniform distributions. For the electron tag, the majority of bins in MC and in experimental data have a value larger than 99 %. Bins with lower efficiencies correspond to regions with low event population and consequently large statistical uncertainties.

For the positron tag – so the tracking efficiency for negatively charged particles – the values of the tracking efficiency in MC are overall slightly lower than for the electron tag. Here, the majority of the values are larger than 98 %. Important to mention here is that the efficiencies are slightly smaller at the edges of the ECL barrel. For the experimental data, the efficiencies are mainly above 99 %, but also here the two low energetic bins in the forward region of the ECL barrel have a lower efficiency.

Overall, the bin-wise efficiencies increased compared to the naive calculation. This is expected as the calculation of η only considers true radiative Bhabha events, and for those events, we are certain that the charged particle creates the higher-energy cluster and that they have two tracks.

Using Gaussian error propagation, we calculate the uncertainties on the unfolded tracking efficiency. Overall, the uncertainties in Figure 7.3.3 are decreasing by one magnitude compared to the uncertainties from the naive calculation. This is expected as the efficiency values are much closer to one, leading to a small binomial variance, and using all predicted categories leads to increased statistics for each bin, resulting in lower statistical uncertainties. Considering that the uncertainties for bins with low population and therefore a bad efficiency are relatively large, the overall agreement of the tracking efficiencies between the different bins is given. The statistical uncertainties for the MC dataset are mainly below 0.01 %. While for the well-populated bins of the experimental dataset, the uncertainties are below 0.005 %. And as expected for those bins, the uncertainties for MC stay larger than those for data due to the larger integrated luminosity of the experimental dataset. For the bins with less population, the uncertainties in experimental data are even larger than in MC. This originates probably from a larger background contamination in experimental data and a lower population relative to the other bins.

With Equation (7.1.1), we calculate the charge asymmetry for data

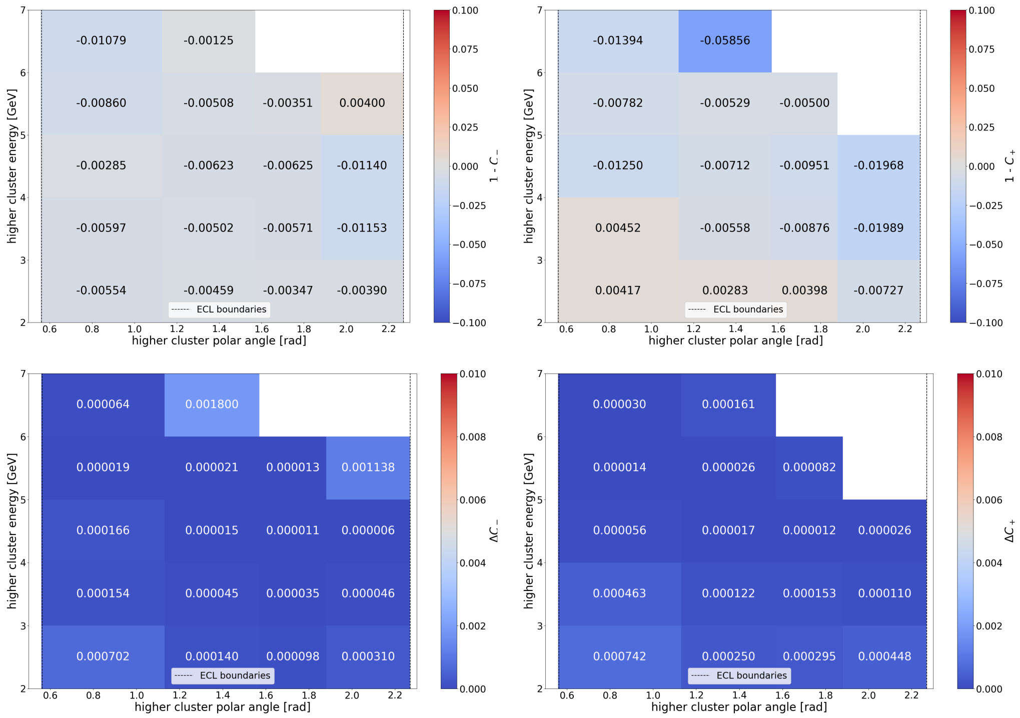

Figure 7.3.5: Correction factor for the corrected tracking efficiency (top row) and its uncertainty (bottom row) between data and MC binned in the phase space of the higher-energy cluster for an electron tag (left) and a positron tag (right).

and MC separately. Following the increased efficiencies, we also expect a smaller charge asymmetry. For the bins with reasonably high population, the charge asymmetry decreases by an order of magnitude. For the data, shown in the right panel of Figure 7.3.4, the majority of the bins in the forward region of the ECL barrel show larger deviations from zero. Also, for MC, the left panel in Figure 7.3.4, the asymmetry values are now closer to zero with even smaller uncertainties than before. But overall, they still deviate more from zero than the values for the experimental data. I conclude that correcting the efficiency for impurities removes a majority of charge-dependent differences in the tracking efficiency. But the bins in the forward direction behave slightly different than the rest.

Further, we calculate the correction factor with Equation (7.2.1) the difference of the correction factor from one is shown in Figure 7.3.5. All bins with high population have negative δ values, meaning the efficiency is larger in data than in MC. The correction factors of the naively calculated efficiencies were all smaller than one. From this change, we conclude that the calculation accounting for impurities improved the efficiency in data more than in MC, and the simulation is underestimating the tracking efficiency. The uncertainties on the correction factors improve by one or two orders of magnitude for data and for MC. As all uncertainties are smaller than 0.1 %, we conclude that the improved efficiency calculation makes our

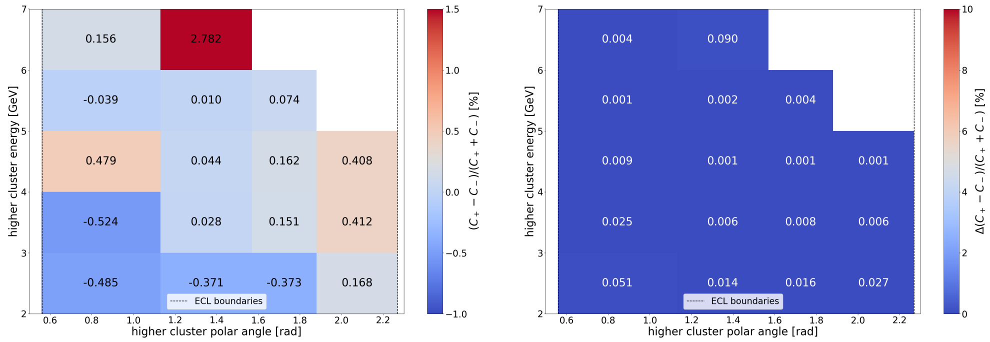

Figure 7.3.6: Asymmetry of the newly calculated correction factors (left) and the corresponding uncertainties (right) binned in the phase space of the higher-energy cluster.

overall result much more precise.

The correction factor asymmetry values, shown in Figure 7.3.6, are on a sub-percent level, indicating that there are only slight differences between the correction factors of the different charges. As the uncertainties are on the same sub-per-mille level, I conclude that the tracking algorithm works well for events with sufficient cluster energy and for events that really have two tracks. Overall, the tracking efficiency results have only small differences for different charges and different detector regions, except for the forward region of the ECL, where a higher efficiency for negatively charged tracks is found.

For a comparison of the calculated efficiencies with the defined track-finding efficiency of the track-finding algorithm at Belle II, the efficiencies are calculated solely as a function of the ECL cluster energy. We compare now the results for the different tags for data and MC, as well as the weighted average over both charges.

The track-finding efficiency for simulated events at the $\Upsilon(4S)$ resonance in dependence of the transverse momentum of the generated particles reaches, for particles with more than 2 GeV [9], more than 99 %, see

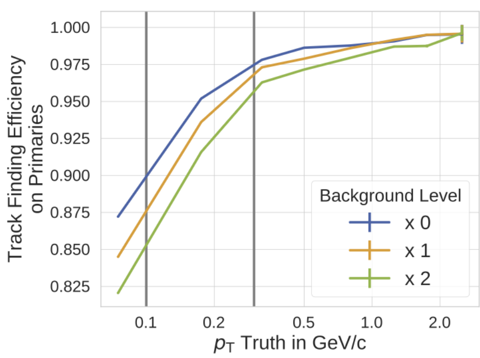

Figure 7.3.7: Track finding efficiency calculated for simulated $\Upsilon(4S)$ events for different fractions of beam-induced background [9].

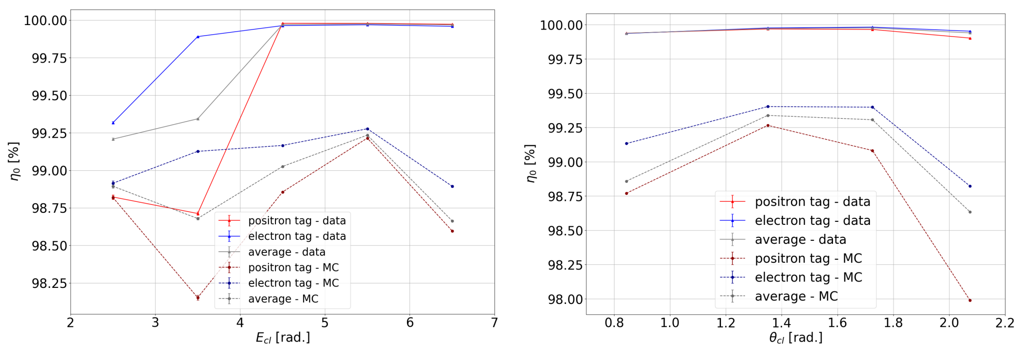

Figure 7.3.8: Corrected tracking efficiency for only cluster energy bins (left) and only polar angle bins (right), divided into the different datasets the different tags and the average of both tags.

Figure 7.3.7. We compare our tracking efficiency results across different cluster energy bins to this result by assuming that electrons have a momentum equal to the reconstructed ECL cluster energy. The efficiency values in data are all larger than 99 % with uncertainties smaller than 0.001 %. I also extracted that at lower energies, the efficiencies for the different tags differ slightly. In the simulation, also slight differences between the charges are visible, so the average over the charges is in the range of 98.5 % to 99.25 %. The uncertainties calculated for the simulation are also smaller than 0.001 %.

From comparing the distributions from Figure 7.3.7 and the left panel of Figure 7.3.8, I conclude that our calculated efficiency is slightly smaller than the ideal track-finding efficiencies determined for the track-finding algorithm. But for higher energies, the efficiencies calculated from data are within the range of expectation.

We also integrate the tracking efficiency values over the cluster energy so that we get an efficiency result only binned in the polar angle (right panel of Figure 7.3.8). The efficiencies for data are all above 99.75 % independent of the polar angle. For MC, the efficiencies range from 98 % to 99.5 %, where the bin in the most backward direction has the lowest efficiency for the positron tag with slightly less than 98 %. The uncertainties for all efficiencies are below 0.001 %.

7.4 Uncertainties

With the unfolded calculation of the tracking efficiency, the impurities coming from background contributions are accounted for and reflected in the statistical uncertainty of the efficiency. Further, the statistical uncertainty could be decreased due to a larger number of events being used for the calculation of the tracking efficiency.

For the unfolded efficiency calculation, we assume that the values and their uncertainties do not depend on the cut of the BDT probability. With a short qualitative check of the behavior of the per-bin efficiencies when adjusting the cut on the BDT probability, I conclude that there is no dependence of the final result on the BDT probability cut.

The trigger selection in this analysis is based solely on cluster information coming from the ECL, while the calculation of the tracking efficiency only uses tracking information, which comes from the CDC. Assuming that the inefficiencies of the tracking algorithm and those in the CDC are uncorrelated with the ECL-based trigger selection, we conclude that any contributions from the trigger to the systematic uncertainties cancel in the end, due to how the tracking efficiency is calculated.

The overall integrated tracking efficiency for a positron tag for MC is $(98.85 \pm 0,0007)$ % and for data it is $(99,95 \pm 0,0002)$ %. For the electron tag of the simulation, the efficiency is (99.16 ± 0.0003) %, and for data, this value is (99.97 ± 0.0002) %. The overall average for simulation is (99.07 ± 0.0004) % and for data (99.96 ± 0.0002) %.

This unfolded calculation of the tracking efficiency includes all impurities originating from events that are not radiative Bhabha events and from wrongly binned events. This and the independence of the BDT probability cut let us conclude that no further systematic uncertainties have to be included in the final result. Our uncertainties on the correction factor are all smaller than 0.1 %. This is an improvement compared to the previous per-track uncertainty [1], which is 0.27 %. This result of the τ-study covers a much smaller energy range than our result extracted from the radiative Bhabha events. So our result gives a more precise value, but is, in general, an extension to the previous tracking efficiency study. The uncertainties on the correction factors can now be used as an input to the systematic uncertainty in other analyses.

Chapter 8

Discussion & Conclusion

I determined the tracking efficiency for Belle II as a function of charge, lab-frame momentum magnitude, and polar angle from radiative Bhabha scattering events, $e^+e^- \to e^+e^-\gamma$. The tracking efficiency in dependence on different variables provides an input to many analyses, including charged particles, as differences in the track reconstruction and track finding are accounted for in a cleaner way. The radiative Bhabha events are reconstructed using the Belle II software framework. The reconstruction is based on a well-constrained tag electron and two additional ECL clusters. The decay reconstruction is performed for simulated and experimental data. To calculate charge-dependent results, the reconstruction is performed separately for an electron and a positron as a tag particle. For each of those four datasets, the ECL cluster originating from the other charged particle has to be identified.

This information is extracted by using a boosted decision tree. The BDT separates the events into three classes. The higher-energy class, where a charged particle creates the higher-energy cluster, the lower-energy class, and the background class. The resolution and the purity of the higher-energy class, extracted from the final BDT, are high enough to create a binned phase-space plot using the ECL cluster energy and the cluster polar angle of the higher-energy cluster.

The final tracking efficiency and its uncertainty are calculated using an unfolding technique, which accounts for any impurities by including all events truly being in the higher-energy class and not only those predicted to be in this class. The charge asymmetry is calculated from the efficiencies for the different tag particles. Positively charged tracks are slightly better reconstructed than negatively charged tracks. The correction factors between data and MC indicate a good overall agreement of data and MC, both differ only on a sub-per-mille level, with the tracking efficiency in MC being slightly underestimated. The uncertainties on the correction factor are all less than 0.01 %. In comparison to the systematic uncertainty recommended by the τ-study, our result is smaller and more granular due to the dependence on charge, momentum, and polar angle. The systematic uncertainty recommended by the τ-study is 0.27 % per track. Important here is that this result was calculated for a much smaller energy range than the energy range covered in this thesis. Therefore, our result provides a good extension to the result extracted from the τ-decay.

The efficiencies themselves are all larger than 99 % with uncertainties below 0.01 %. For further studies involving charged tracks at Belle II, the statistical uncertainty on the correction factor can be used as a systematic uncertainty.

For the future, a larger amount of $ee(\gamma)$-signal MC would help to decrease the statistical uncertainty on the MC values. Further, the implementation of other triggers will lead to a larger phase-space coverage of the analyzed events, so that the tracking efficiency is also calculated for the endcaps of the ECL and for lower energies.

References

[1] Petar Rados and Laura Zani. *Measurement of the Tracking Efficiency with $e^+e^- \to \tau^+\tau^-$ Events in Run1 Data and MC15rd*. BELLE2-NOTE-PH-2024-XXX. DRAFT Version 1. Belle II Collaboration, 2024. URL: https://gitlab.desy.de/belle2/software/tracking/systematics/tracking-efficiency-tau/-/tree/main?ref_type=heads.

[2] The Belle 2 Collaboration. *XWiki Belle 2: Physics Cross Sections Table*. URL: https://xwiki.desy.de/xwiki/bin/view/BI/Belle%20II%20Internal/Physics%20WebHome/Physics%20Cross%20Sections%20Table/?srid=hIUtQ0tr&srid=hIUtQ0tr (visited on 10/05/2025).

[3] E. Kou et al. „The Belle II Physics Book“. In: *Progress of Theoretical and Experimental Physics* 2019.12 (Dec. 2019). ISSN: 2050-3911. DOI: 10.1093/ptep/ptz106. URL: http://dx.doi.org/10.1093/ptep/ptz106.

[4] Particle Data Group. *Particle Data Group - Summary tables*. URL: https://pdg.lbl.gov/2025/tables/contents_tables.html (visited on 10/15/2025).

[5] Belle II Collaboration. *Belle II Technical Design Report*. 2010. arXiv: 1011.0352 [physics.ins-det]. URL: https://arxiv.org/abs/1011.0352.

[6] KEK High Energy Accelerator Research Organization. *SuperKEKB Project*. URL: https://www-superkekb.kek.jp (visited on 10/29/2025).

[7] The Belle 2 Collaboration. *Belle II - Luminosity*. URL: https://www.belle2.org/research/luminosity/ (visited on 10/15/2025).

[8] T. Hara, T. Kuhr, and Y. Ushiroda. „Belle II Coordinate System and Guideline of Belle II Numbering Scheme“. In: (2011). Belle II Note.

[9] Valerio Bertacchi et al. „Track finding at Belle II“. In: *Comput. Phys. Commun.* 259 (2021), p. 107610. DOI: 10.1016/j.cpc.2020.107610. arXiv: 2003.12466 [physics.ins-det].

[10] Belle II Software Group. *Belle II Software Documentation*. (Visited on 11/27/2025).

[11] Valerio Bertacchi for the skim team. *Data production and Skim*. Internal presentation (only available to Belle II members). URL: https://indico.belle2.org/event/14576/contributions/98788/attachments/36471/54093/presentation%2025%20-%20skim%20Belle%20II%20Germany.pdf (visited on 11/27/2025).

[12] H.Haigh. „Analysis of high momentum tracking efficiency using $e^+e^- \rightarrow e^+e^-\gamma$ events". In: *Belle II Note* (2023).

[13] Belle II Collaboration. *Analysis of high momentum tracking efficency using $e^+e^- \rightarrow e^+e^-\gamma$ events*. Belle II internal note. 2023. URL: https://docs.belle2.org/files/489/BELLE2-NOTE-TE-2023-008/1/BELLE2-NOTE-TE-2023-008.pdf.

[14] H. Nakazawa T.Koga. „Trigger Summary in experiment 24,25,26". In: *BELLE2-NOTE-TE-2022-011* (2022).

[15] Belle II Collaboration. *HLT filter documentation*. Accessed: 2025-05-28. URL: https://xwiki.desy.de/xwiki/bin/view/BI/Belle%20II%20Internal/Physics%20Performance%20Webhome/Trigger%20performance/HLT%20filter%20documentation/.

[16] LightGBM Contributors. *LightGBM Documentation*. 2025. URL: https://lightgbm.readthedocs.io/en/stable/index.html (visited on 12/17/2025).

[17] M.Milesi et.al. *Muon and electron identification performance with 189 fb^{-1} of Belle II data*. BELLE2-NOTE-TE-2021-011 (only available to Belle II members). 2022.