Department of Physics and Astronomy
Heidelberg University

Bachelor Thesis in Physics
submitted by

Tobias Heintz

born in Schweinfurt (Germany)

2023

Improvement of the template fit method for the fake photon background estimation at ATLAS

This Bachelor Thesis has been carried out by Tobias Heintz at the
Kirchhoff Institute for Physics in Heidelberg
under the supervision of
Prof. Dr. Hans-Christian Schultz-Coulon

Abstract

ATLAS employs a template fit method to estimate the relative contributions of promptly produced photons (so-called real photons) and photons arising within the fragmentation of jets (so-called fake photons) based on their isolation energy distributions. A generalization of this method is studied in this thesis using different photon final states. Templates for the photon isolation energy variable are extracted for real and fake photons from Monte Carlo simulations. The dependence of the templates on the transverse momentum of the photon, p_{T}, is studied and generic isolation templates are extracted for both photons of $Z\gamma\gamma$ final states. Those templates also describe the isolation energy distribution of photons in $Z\gamma$ final states. Good agreement is found by validating the Monte Carlo extracted isolation templates with control regions in proton-proton collisions at the Large Hadron Collider at $\sqrt{s} = 13\,\mathrm{TeV}$. The fake photon background in the $Z\gamma\gamma$ analysis for the full ATLAS Run-2 dataset is estimated to $N_{\mathrm{fakes}} = 41.3 \pm 18.9$ in agreement with a previous $Z\gamma\gamma$ analysis. The presented results demonstrate that a generalization of the isolation templates is feasible for both photons in $Z\gamma\gamma$ final states and for photons in $Z\gamma$ final states by taking into account the p_{T} distribution of the studied phase space.

Zusammenfassung

ATLAS verwendet die Template Fit Methode um die relativen Beiträge von prompt erzeugten Photonen (sogenannte reale Photonen) und Photonen, die bei der Fragmentierung von Jets entstehen (sogenannte fake Photonen), anhand der Verteilung der Isolationsenergie abzuschätzen. Eine Verallgemeinerung dieser Methode wird in dieser Arbeit unter Verwendung verschiedener Prozesse mit Photonen im Endzustand studiert. Templates für die Verteilung der Isolationsenergie von realen und von fake Photonen werden aus Monte Carlo Simulationen extrahiert. Die Abhängigkeit der Templates vom transversalen Impuls der Photonen, p_{T}, wird untersucht, und es werden generische Isolations-Templates extrahiert, welche die Verteilung der Isolationsenergie für beide Photonen in $Z\gamma\gamma$ Endzuständen und auch für Photonen in $Z\gamma$ Endzuständen beschreiben. Eine gute Übereinstimmung der aus Monte Carlo extrahierten Templates mit der Verteilung der Isolationsenergie innerhalb Kontrollregionen in Proton-Proton Kollisionen am Large Hadron Collider bei $\sqrt{s} = 13\,\mathrm{TeV}$ wird beobachtet. Der Untergrund der fake Photonen in der $Z\gamma\gamma$ Analyse der vollständigen ATLAS Run-2 Daten wird mit der Template Fit Methode auf $N_{\mathrm{fakes}} = 41.3 \pm 18.9$ geschätzt. Dies stimmt gut mit einer vorangegangenen Veröffentlichung überein, die den gleichen $Z\gamma\gamma$ Phasenraum untersucht hat. Die vorgestellten Ergebnisse zeigen, dass eine Verallgemeinerung der Isolations Templates für beide Photonen in $Z\gamma\gamma$ Endzuständen und für Photonen in $Z\gamma$ Endzuständen möglich ist, indem die p_{T} Verteilung des untersuchten Phasenraums berücksichtigt wird.

Contents

Chapter 1

Introduction

One fundamental question that has inspired mankind for centuries is the nature of the building blocks that make up the universe. The Standard Model (SM) [1] is *the* central theoretical framework of particle physics that summarizes the elementary particles and their interactions in the subatomic regime. It contains three generations of fermions, which are the elementary matter particles. Moreover, it includes the bosons, which mediate the electromagnetic, weak and strong nuclear forces between the particles.

Modern particle physics aims to test the predictions of the SM with high precision and to investigate new physics beyond the SM that may provide a more comprehensive understanding of the universe. For example, hypothetical supersymmetry particles, which are heavier partners of the known SM particles, could contribute to a unified theory of the fundamental forces [2].

Large particle accelerators and colliders are used to achieve the high energies and high luminosities which are necessary for the production of unknown particles and for the observation of rare processes. The most powerful particle collider to date is the Large Hadron Collider [3], where high-energetic protons collide and interact. The collision products are measured and analyzed with different particle detectors, for example the ATLAS detector [4].

One goal of the ATLAS experiment is testing the electroweak sector of the SM. For that reason, processes with photons (γ), W and Z bosons in the final state are studied, for instance $Z\gamma$, $Z\gamma\gamma$, $ZZ\gamma$, $W\gamma\gamma$ and more. A very important or even dominating background in such processes are so-called fake photons. These fake photons are not-promptly produced in the proton-proton collision, but arise in jets, mainly with a neutral pion decay into two photons, $\pi^0 \rightarrow \gamma\gamma$. Jets that are misidentified as photons also contribute to the fake photon background, but to a much lesser extent.

Photon isolation and identification selection criteria are used to distinguish between real and fake photons. Real photons are surrounded by less hadronic activity, whereas fake photons show additional activity in a cone around them, corresponding to the jet

remnant. Therefore, the distribution of the isolation energy of fake photons is shifted towards higher energies and is broader with respect to real photons. This is utilized in the so-called template fit method by fixing the shape of the isolation templates for real and fake photons and performing a simultaneous fit in order to determine the number of real and fake photons.

However, the shape of the isolation templates needs to be known and has to be extracted from Monte Carlo (MC) simulations in the case of limited statistics. The method which is currently in standard use exhibits a potential challenge, namely the dependence of the isolation templates on the transverse momentum of the photon, p_{T}. This is problematic if the p_{T} distribution of the MC simulation does not precisely model the p_{T} distribution in collision data. Another consequence for processes with more than one photon in the final state, e.g. $Z\gamma\gamma$, is that the isolation templates for the different final state photons vary due to their individual p_{T} distributions.

The main goals of this thesis are the extraction of isolation templates for real and fake photons and to investigate the generalization of those isolation templates to different photon final states. Therefore, it is examined if a combined template can be used for both photons in the $Z\gamma\gamma$ phase space and also for photons in other processes, exemplarily in $Z\gamma$ final states. This thesis also aims to validate the extracted isolation templates by estimating the fake photon background to $Z\gamma\gamma$ final states in the full ATLAS Run-2 dataset with the template fit method and by comparing the results to a previous analysis that studied the same $Z\gamma\gamma$ phase space [5, 6].

The thesis is structured as follows. In Chapter 2, a short overview of the ATLAS detector is given, followed by a description of photon reconstruction and identification in Chapter 3. In Chapter 4, isolation templates for real and fake photons are extracted and validated for $Z\gamma\gamma$ final states. In Chapter 5, the template fit is performed in a full Run-2 $Z\gamma\gamma$ phase space in order to estimate the fake photon background. Finally, in Chapter 6 the generalization of the $Z\gamma\gamma$ isolation templates is investigated, exemplarily for $Z\gamma$ final states. Chapter 7 serves as conclusion and summarizes the main results of this analysis.

Chapter 2

The ATLAS Experiment at the Large Hadron Collider

The Large Hadron Collider (LHC) [3] provides high-energetic proton-proton (pp) collisions and also heavy ion collisions with lead. It is located at the European Organization for Nuclear Research (CERN) near Geneva. The LHC started operations in 2008 and recently data taking of Run-3 has begun. The data used for this analysis was taken during Run-2 between 2015 and 2018 at the ATLAS detector (see below) with an integrated luminosity of $L_{\text{int}} = (140.1 \pm 1.2)\,\text{fb}^{-1}$ [7]. After pre-acceleration, two proton beams are accelerated in separate beam pipes of the LHC in opposite directions up to beam energies of $E_{\text{beam}} = 6.5\,\text{TeV}$. They are brought to collision at four interaction points, at which the experiments ALICE, ATLAS, CMS and LHCb are located.

The ATLAS Detector

The ATLAS[1] detector [4, 8] is a multipurpose detector with nearly 4π coverage and forward-backward symmetry. It has a cylindrical geometry with 44 m length, 25 m diameter and a weight of roughly 7000 t.

Cylindrical coordinates (r, ϕ) are chosen to describe the plane transverse to the beam (z direction), the polar angle is expressed with the pseudorapidity

$$\eta = -\log\left[\tan\left(\frac{\theta}{2}\right)\right], \tag{2.1}$$

and angular distances are given in terms of

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}. \tag{2.2}$$

[1] The name "ATLAS" originated as an acronym for "A Toroidal LHC ApparatuS", referring to the unique combination of a central solenoid for the magnetic field in the tracking system and three air-core toroids for the muon spectrometer.

An overall view of the ATLAS detector is given in Figure 2.1. It consists of different layers, which are arranged concentrically around the interaction point. Each subdetector follows a specific aim for particle detection and identification, as well as the measurement of fundamental quantities such as charge, momentum and energy.

- **Inner detector.** High granularity pixel detectors (including the insertable B-layer) and silicon microstrip detectors yield precise and discrete tracking and vertex measurements up to $|\eta| < 2.5$. More continuous tracking is achieved by the gas filled straw tubes of the transition radiation tracker, enabling track measurements up to $|\eta| < 2$. In addition, the transition radiation tracker allows the identification of electrons due to the detection of transition radiation photons. The central solenoid provides a 2 T axial magnetic field for momentum measurement of charged particles.

- **Calorimeter.** The central solenoid is surrounded by the electromagnetic (EM) calorimeter, which is a sampling calorimeter with accordion geometry. It contains liquid argon (LAr) as active material and lead as absorber. It consists of a barrel ($|\eta| < 1.475$) and two end-caps ($1.375 < |\eta| < 3.2$), which have a variable thickness of more than 22 and 24 radiation lengths X_0, respectively. The fine granularity of the EM calorimeter enables precision measurements of electrons and photons. Up to $|\eta| < 1.8$, the energy loss of electrons and photons within the inner detector is determined by the presampler. Up to $|\eta| < 2.5$, the EM calorimeter consists of three layers, which are shown in Figure 2.2. The first layer has a high granularity of 0.003×0.1 in the $\eta \times \phi$ space. This often allows two energy maxima of $\pi^0 \to \gamma\gamma$ to be resolved separately. It has a thickness of $4.4X_0$ and covers $|\eta| < 1.4$ as well as $1.5 < |\eta| < 2.4$. The second layer samples most of the deposited energy, since it has a relatively high thickness of $17X_0$. In the third layer, the leakage of electrons and photons behind the EM calorimeter is estimated.

  The hadronic calorimeter consists of a tile sampling calorimeter (scintillating tiles with steal absorbers) in the barrel region ($|\eta| < 1.7$) and LAr end-cap and forward calorimeters, covering $1.5 < |\eta| < 3.2$ and $3.1 < |\eta| < 4.9$, respectively.

- **Muon spectrometer.** Monitored drift tubes and high granularity cathode strip chambers are used for high precision tracking of muons. Three superconducting air-core toroid magnets (a barrel and two end-caps) provide 0.5 T and 1 T magnetic fields orthogonal to the trajectories of the muons for momentum measurement.

Due to the high collision rate of 40 MHz, a trigger system consisting of the level-1 trigger and the high-level trigger is used to minimize the rate to 1 kHz. The level-1 trigger is hardware based and operates on reduced granularity information of the calorimeters. It lowers the rate down to 100 kHz and defines regions of interest. The entire detector

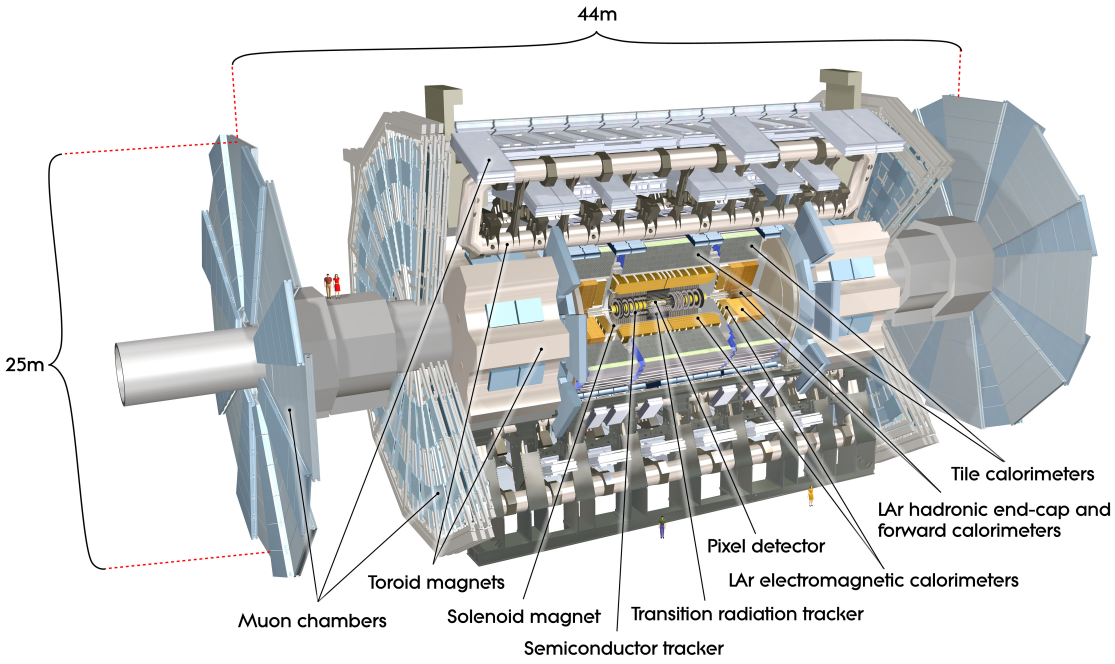


Figure 2.1: Schematic overview of the ATLAS detector [9].

data and the full granularity information within the regions of interest are used in the software based high-level trigger to reduce the rate for permanent storage to 1 kHz.

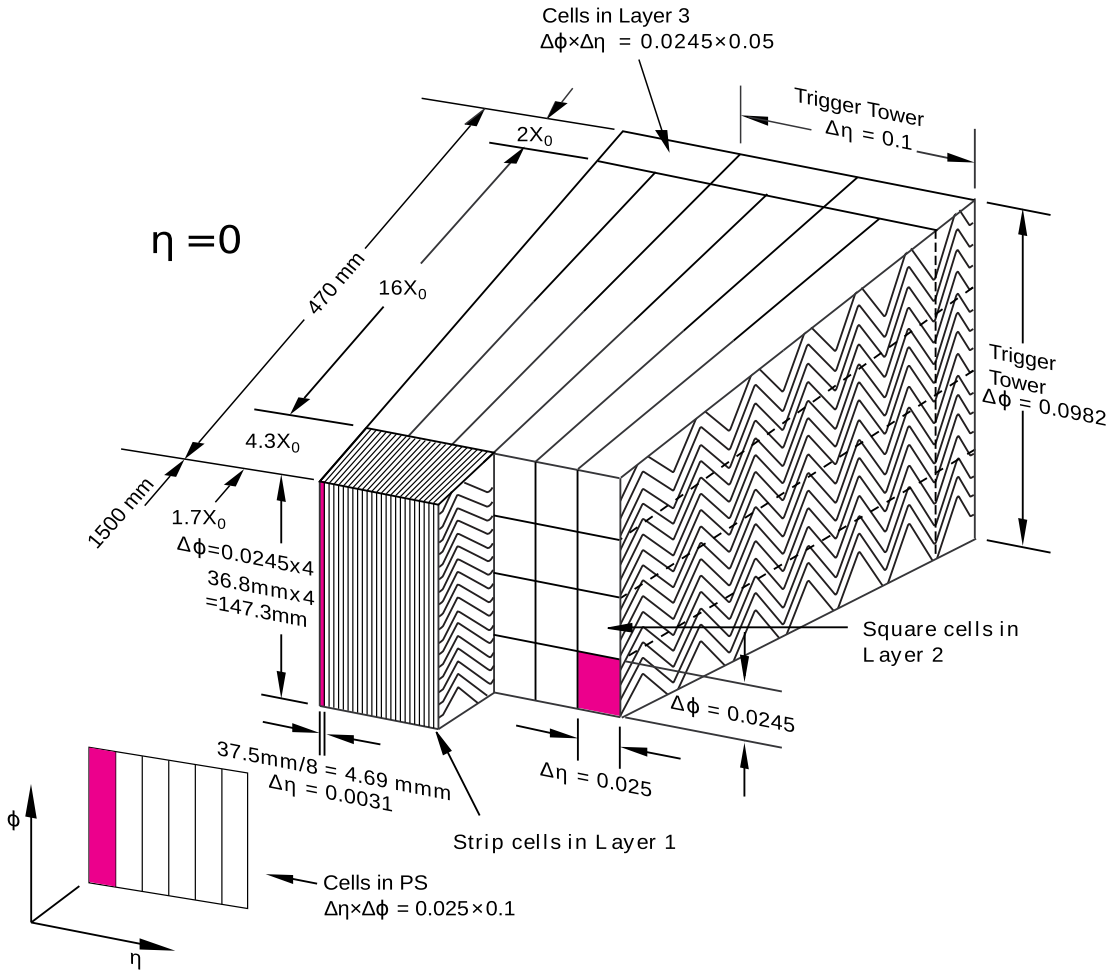


Figure 2.2: Segmentation of the ATLAS EM calorimeter (at $\eta = 0$) into presampler (PS), high granularity layer 1, layer 2 and layer 3. Moreover, the accordion geometry of the EM calorimeter can be seen at the side cut [10].

Chapter 3

Photon Reconstruction and Identification

Electrons and photons are reconstructed and identified based on their typical signatures in the subdetectors. Electrons interact with the material of the inner detector causing hits which are used for tracking. In contrast, prompt photons[2] cannot be tracked in the inner detector due to their neutral electrical charge. However, photons can interact with nuclei or electrons of atoms in the detector material via pair production and Compton scattering, respectively. So-called converted photons are characterized by (one or two) tracks starting at the conversion vertices. One track of the scattered electron is observed in the case of Compton scattering, and two tracks are found in the case of e^+e^- pair production. Electrons, photons and converted photons deposit their energy in the EM calorimeter in form of EM showers and hence produce clusters in the EM calorimeter. The typical signatures of electrons, prompt and converted photons are summarized in Figure 3.1.

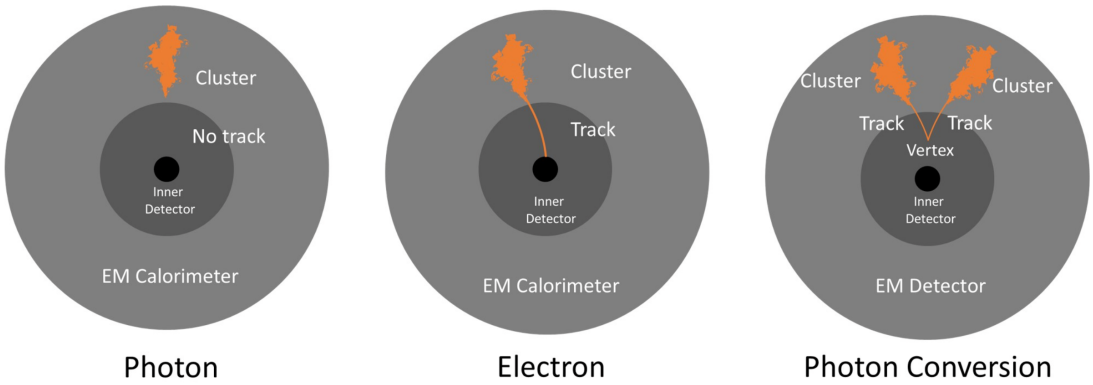


Figure 3.1: Characteristic signatures of prompt photons, electrons and converted photons in the inner detector and the EM calorimeter. Exemplarily, the photon conversion into an e^+e^- pair is visualized. In the case of Compton scattering, the photon conversion would be characterized by only one track in the inner detector and one cluster in the EM calorimeter [11].

[2]The terms "prompt photons" and "real photons" are used synonymously in this thesis.

Photon Reconstruction

The main aspects of the cluster-reconstruction algorithm are described in the following paragraphs, a more detailed description is given in [12].

- **Proto-clusters** are found by searching for cells in the EM and the hadronic calorimeters. In the first step of the *4-2-0 algorithm*, a cell with significance $|\zeta_{\text{cell}}| \geq 4$ is required. Here,

$$\zeta_{\text{cell}} = \frac{E_{\text{cell}}}{\sigma_{\text{noise,cell}}} \tag{3.1}$$

  is the ratio of the energy deposited in the cell, E_{cell}, and the expected cell noise, $\sigma_{\text{noise,cell}}$. Iteratively, adjacent cells with significances $|\zeta_{\text{cell}}| \geq 2$ are matched to the proto-cluster. In the final step, all neighboring cells with $|\zeta_{\text{cell}}| \geq 0$ are added to the proto-cluster.

- **Topological clusters (topo-clusters)** are proto-clusters with $E^{\text{EM}}_{\text{proto}} > 400\,\text{MeV}$ and with $f_{\text{EM}} = E^{\text{EM}}_{\text{proto}}/E^{\text{total}}_{\text{proto}} > 0.5$. Here, $E^{\text{EM}}_{\text{proto}}$ is the energy of the proto-cluster only in the EM calorimeter and $E^{\text{total}}_{\text{proto}}$ is the total energy of the proto-cluster.

- **Superclusters** are formed in two steps. At the first stage, the topo-clusters are sorted with respect to their transverse energy, E^{EM}_{T}. Starting by the topo-cluster with the highest energy, topo-clusters with $E^{\text{EM}}_{\text{T}} > 1.5\,\text{GeV}$ are chosen as seeds for the superclusters. In the second step, additional clusters within 0.075×0.125 relative to the seed center in the $\eta \times \phi$ space are added to the supercluster as satellite clusters. After energy calibration and position correction, the superclusters are matched with tracks and conversion vertices found in the inner detector.

Photon Identification

Promptly produced photons typically have more narrow energy depositions in the EM calorimeter and smaller leakage into the hadronic calorimeter with respect to fake photons. Due to the fine granularity of the first layer in the EM calorimeter, often two separated local energy maxima of fake photons, e.g. $\pi^0 \to \gamma\gamma$, can be resolved.

Loose and tight photon identification (ID) working points[3] are defined by different cuts on the *shower shape observables* [13]. This includes variables of the shower shape, different energy ratios and the shower width. An overview of the shower shape variables is given in Table A.1 and in Figure 3.2.

[3]The tight working point has a higher background rejection but also a lower signal yield with respect to the loose working point.

- The shower shape (first column in Figure 3.2) can be characterized by two variables. The front second maximum energy difference

$$\Delta E = E^{S1}_{\text{max},2} - E^{S1}_{\text{min}} \tag{3.2}$$

is the energy difference in the first layer of the EM calorimeter ("$S1$") between the minimum, E^{S1}_{min}, and the second-largest maximum, $E^{S1}_{\text{max},2}$. The front maxima relative energy ratio

$$E_{\text{ratio}} = \frac{E^{S1}_{\text{max},1} - E^{S1}_{\text{max},2}}{E^{S1}_{\text{max},1} + E^{S1}_{\text{max},2}} \tag{3.3}$$

measures the relative energy difference in the first layer between the first and the second-largest maximum.

- The lateral shower width in η direction (second column in Figure 3.2) is characterized by different shower shape observables. The front lateral width (3 strips/total), w_{s3} and w_{stot}, compare the energy centered around the highest energy deposition in the first layer to the energy deposited in a larger window. The middle lateral width, w_{η_2}, is the variance of the distribution of the energy deposition in η direction within the second layer.

- Different energy ratios (third column in Figure 3.2), including the leakage into the hadronic calorimeter, are used to discriminate the shower shape of real and fake photons further.

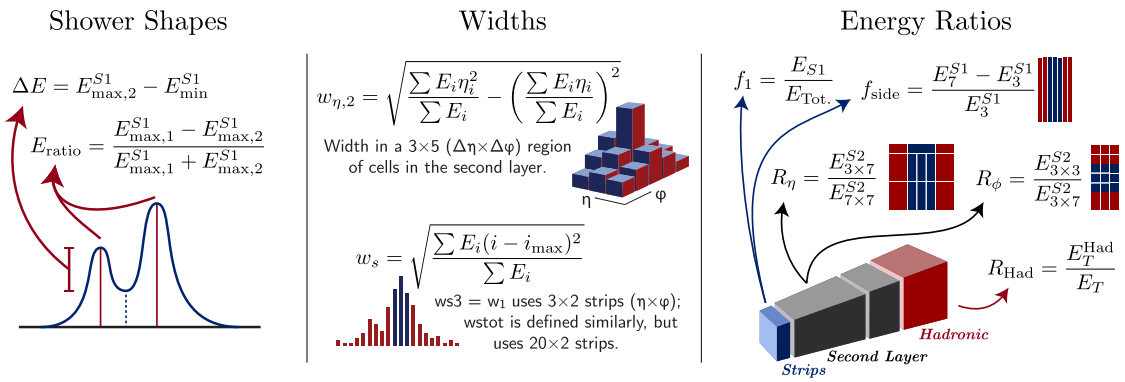


Figure 3.2: Discriminating shower shape observables that are used for photon identification. Adapted from [14].

Photon Isolation

Promptly produced photons show less hadronic activity in a cone around them, whereas fake photons are surrounded by additional activity from the jet remnant. Therefore, the isolation of the reconstructed tracks in the inner detector and the isolation of the clusters in the calorimeters are used to discriminate real and fake photons.

- **Calorimeter isolation.** The transverse energy $E_{\mathrm{T}}^{\mathrm{EM}}$ of topo-clusters with barycenter in a cone around the photon are summed to $E_{\mathrm{T,raw}}^{\mathrm{isol}} = \sum_i \left(E_{\mathrm{T}}^{\mathrm{EM}}\right)_i$. To get the additional energy in the cone, $E_{\mathrm{T}}^{\mathrm{cone}}$, the energy of the photon, $E_{\mathrm{T,core}}$, needs to be subtracted, and it needs to be corrected for leakage and pileup:

$$E_{\mathrm{T}}^{\mathrm{cone}} = E_{\mathrm{T,raw}}^{\mathrm{isol}} - E_{\mathrm{T,core}} - E_{\mathrm{T,leakage}} - E_{\mathrm{T,pileup}}. \tag{3.4}$$

  This is graphically visualized in Figure 3.3.

- **Track isolation.** The transverse momentum p_{T} of tracks in a cone around the photon are summed up. Here, only tracks with $p_{\mathrm{T}} > 1\,\mathrm{GeV}$ and $|\eta| < 2.5$ are considered. Additionally, the tracks must fulfill quality requirements, for instance to have at least seven hits in the silicon detector and at most one (two) hole(s), i.e. missing hits, in the pixel (silicon) detector.

For photon isolation, three ATLAS recommended [12] working points are defined in Table 3.1. The loose isolation working point, which is used in this analysis, requires

$$E_{\mathrm{iso}} := E_{\mathrm{T}}^{\mathrm{cone20}} - 0.065 \cdot p_{\mathrm{T}} < 0, \tag{3.5}$$

$$p_{\mathrm{iso}} := p_{\mathrm{T}}^{\mathrm{cone20}} - 0.05 \cdot p_{\mathrm{T}} < 0. \tag{3.6}$$

Here, "cone20" denotes that the cone has a radius of $\Delta R = 0.2$.

Table 3.1: Overview of ATLAS recommended nominal working points for photon isolation. In this analysis, the `FixedCutLoose` photon isolation working point is used.

Nominal Working Point	Calorimeter Isolation [GeV]	Track Isolation
`FixedCutTightCaloOnly`	$E_{\mathrm{T}}^{\mathrm{cone40}} - 0.022 \cdot p_{\mathrm{T}} < 2.45$	–
`FixedCutTight`	$E_{\mathrm{T}}^{\mathrm{cone40}} - 0.022 \cdot p_{\mathrm{T}} < 2.45$	$p_{\mathrm{T}}^{\mathrm{cone20}} - 0.05 p_{\mathrm{T}} < 0$
`FixedCutLoose`	$E_{\mathrm{T}}^{\mathrm{cone20}} - 0.065 \cdot p_{\mathrm{T}} < 0$	$p_{\mathrm{T}}^{\mathrm{cone20}} - 0.05 p_{\mathrm{T}} < 0$

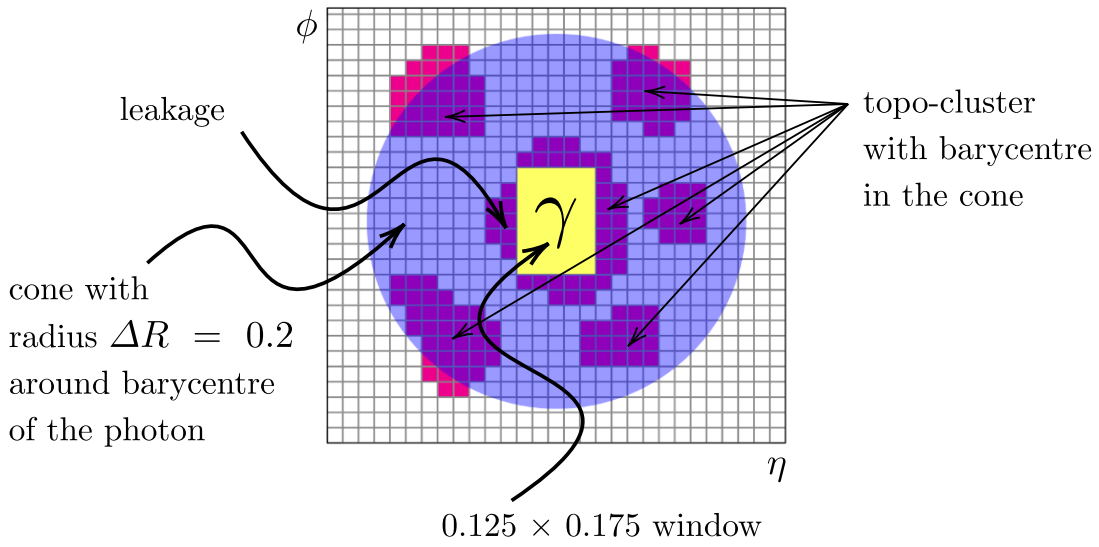


Figure 3.3: Sketch of the composition of additional energy in a cone around the photon. The energies of topo-clusters (magenta cells) whose barycenter lies in a cone with radius $\Delta R = 0.2$ (purple circle) around the photon are summed up. The energy of the photon itself is subtracted, by subtracting the energy of a 0.125×0.175 window (yellow rectangle) and correcting for leakage (magenta cells next to the yellow rectangle). Furthermore, a pileup correction is subtracted. The size of the cone in the sketch does not correspond to the scale of the remaining sketch. Adapted from [15].

Chapter 4

Extraction of Isolation Templates for Real and Fake Photons

In this chapter, templates for the $E_{\mathrm{T}}^{\mathrm{cone20}}$ distribution of real and fake photons are extracted. The template fit method was used in a recent publication [5, 6], which studied the $Z\gamma\gamma$ phase space of the ATLAS Run-2 dataset. Therefore, $Z\gamma\gamma$ final states are chosen as the starting point of this analysis. Later in this thesis, the generalization of the template fit method to other processes with photons in the final state is studied, exemplarily for $Z\gamma$ final states. Due to limited statistics in the $Z\gamma\gamma$ phase space, the isolation templates are extracted from MC simulations. $Z\gamma\gamma$ MC simulations are used to select prompt photons and $Z\gamma + \text{jets}$ MC simulations are used to select fake photons. In Section 4.1, an overview of the $Z\gamma\gamma$ selection is given. In Section 4.2, the general fit procedure is explained. In Section 4.3 and 4.4, the $E_{\mathrm{T}}^{\mathrm{cone20}}$ templates are extracted for real and fake photons from MC simulations and validated with control regions (CRs) of the Run-2 dataset.

4.1 $Z\gamma\gamma$ Selection

$Z\gamma\gamma$ final states with leptonic decays of the Z boson into a pair of e^+e^- or $\mu^+\mu^-$ are studied with $ee\gamma\gamma$ and $\mu\mu\gamma\gamma$ MC simulations generated by Sherpa 2.2.10 at next-to-leading order quantum chromodynamics (QCD) precision [16, 17], and within different signal and control regions in the ATLAS Run-2 dataset. The analysis is performed in an inital-state radiation (ISR)-enriched phase space, i.e. both photons are radiated from an initial quark line, as shown in Figure 4.1a. Therefore, events with one or two final-state radiation (FSR) photon(s), i.e. photons radiated by higher-order quantum electrodynamics processes from a lepton in the final state, as shown in Figure 4.1b, are not considered. The FSR cut is explained below.

If an event fired a single lepton trigger, candidates for electrons, muons and photons

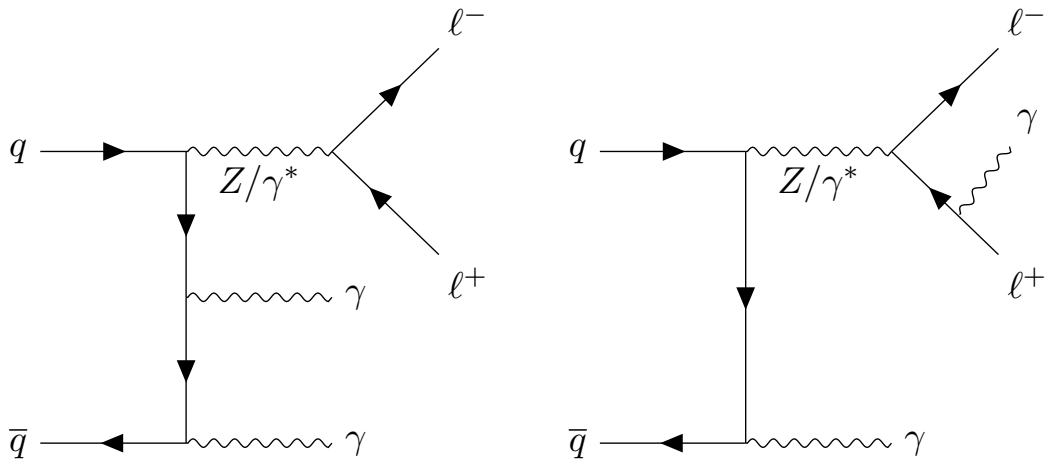


Figure 4.1: Feynman diagrams for $Z\gamma\gamma$ final states. The Z boson decays leptonically into a pair of e^+e^- or $\mu^+\mu^-$. Both photons are radiated by initial quark lines in (a). In contrast, one photon is radiated by a higher order quantum electrodynamics process from a lepton of the final state in (b).

Table 4.1: Overview of trigger and cuts on object level in the preselection process. Further cuts are required on event level and in the $Z\gamma\gamma$ selection. The transverse and longitudinal impact parameters, d_0 and z_0, are described in [18]. The identification (ID) and isolation (ISO) selection criteria are described in [12] and [19] for e/γ and μ, respectively.

Level	e	μ	γ
Trigger	single lepton trigger		–
Object Level	$p_{\mathrm{T}} > 20\,\mathrm{GeV}$	$p_{\mathrm{T}} > 20\,\mathrm{GeV}$	$p_{\mathrm{T}} > 15\,\mathrm{GeV}$
	$\lvert\eta\rvert < 1.37$ or $1.52 < \lvert\eta\rvert < 2.47$	$\lvert\eta\rvert < 2.5$	$\lvert\eta\rvert < 1.37$ or $1.52 < \lvert\eta\rvert < 2.37$
	$\left\lvert\frac{d_0}{\sigma_{d_0}}\right\rvert < 5$	$\left\lvert\frac{d_0}{\sigma_{d_0}}\right\rvert < 3$	–
	$\lvert z_0 \sin\theta\rvert < 0.5\,\mathrm{mm}$	$\lvert z_0 \sin\theta\rvert < 0.5\,\mathrm{mm}$	–
	medium e ID	medium μ ID	loose γ ID
	loose ISO	loose ISO	–

are searched for according to the object level selection as summarized in Table 4.1.

At least two such leptons ($e^{\pm}, \mu^{\pm}$) and two such photons are required, before so-called overlap removal cuts are applied:

$$\Delta R(\gamma, \ell) > 0.4 \quad \text{and} \quad \Delta R(e, \mu) > 0.2. \tag{4.1}$$

Finally, $Z\gamma\gamma$ events are selected according to the following event selection:

- At least two leptons with the same flavor and opposite charge and at least two photons are required. The two highest momentum leptons and photons are selected, requiring $p_{\mathrm{T}}(\ell_1) \geq 30\,\mathrm{GeV}$ and $p_{\mathrm{T}}(\ell_2, \gamma_{1,2}) \geq 20\,\mathrm{GeV}$. Here, the highest momentum photon (lepton) is denoted leading photon (lepton) and labelled with index 1.

Analogously the second-highest momentum photon (lepton) is denoted subleading photon (lepton) and labelled with index 2.

- The leading lepton must match the single lepton trigger requirements, by passing tight identification (electron) or tight isolation (muon) criteria.
- Both photons must pass tight identification criteria. Moreover, they must be separated by $\Delta R(\gamma_1, \gamma_2) > 0.4$.
- For the template fit method, no isolation cut is applied to the photons at the stage of the event selection. The loose isolation criteria, Equations 3.5 and 3.6, are taken into account by integrating the templates of E_{iso} and p_{iso} from $-\infty$ to 0 after performing the fit.
- The invariant mass of the lepton pair must be larger than 40 GeV, and in order to only select events with ISR photons, the FSR cut

$$m_{\ell\ell} + \min\left(m_{\ell\ell\gamma_1}, m_{\ell\ell\gamma_2}\right) > 2m_Z \tag{4.2}$$

is applied. Here, $m_{\ell\ell\gamma_i}$ is the three-body invariant mass of the lepton pair and the (sub)leading photon. If both photons are radiated from initial quark lines, Equation 4.2 is fulfilled, as $m_{\ell\ell} = m_Z$ and $\min_i m_{\ell\ell\gamma_i} > m_Z$. In contrast, if at least one photon is radiated from a lepton in the final state, Equation 4.2 is not satisfied, as $m_{\ell\ell} < m_Z$ and $\min_i m_{\ell\ell\gamma_i} \leq m_Z$. The FSR cut is visualized in Figure 4.2.

4.2 General Fit Procedure

As discussed in Chapter 3, fake photons are expected to be less isolated compared to prompt photons. Hence, the distribution of the isolation energy of fake photons is shifted towards higher values and is broader with respect to real photons. The basic idea of the template fit method, as visualized with toy data in Figure 4.3, is to fix the shape of the isolation energy distributions of real and fake photons and to fit the normalization factors for both templates simultaneously. After performing the fit, the loose isolation criteria are taken into account by integrating the isolation templates from $-\infty$ to 0.

As the subtraction of $6.5\% \cdot p_{\text{T}}$ in Equation 3.5 dominates for $p_{\text{T}} \gtrsim 50\,\text{GeV}$ with respect to $E_{\text{T}}^{\text{cone20}}$, E_{iso} is shifted to lower E_{iso} values for higher p_{T} values, as shown in Figure 4.4. Therefore, the isolation templates are mainly driven by the p_{T} spectrum of the MC simulation. To avoid this strong dependence on the p_{T} distribution of the MC simulation, templates are extracted for $E_{\text{T}}^{\text{cone20}}$ rather than E_{iso} in this chapter. The $E_{\text{T}}^{\text{cone20}}$ templates still show a p_{T} dependence, as discussed in the following two sections.

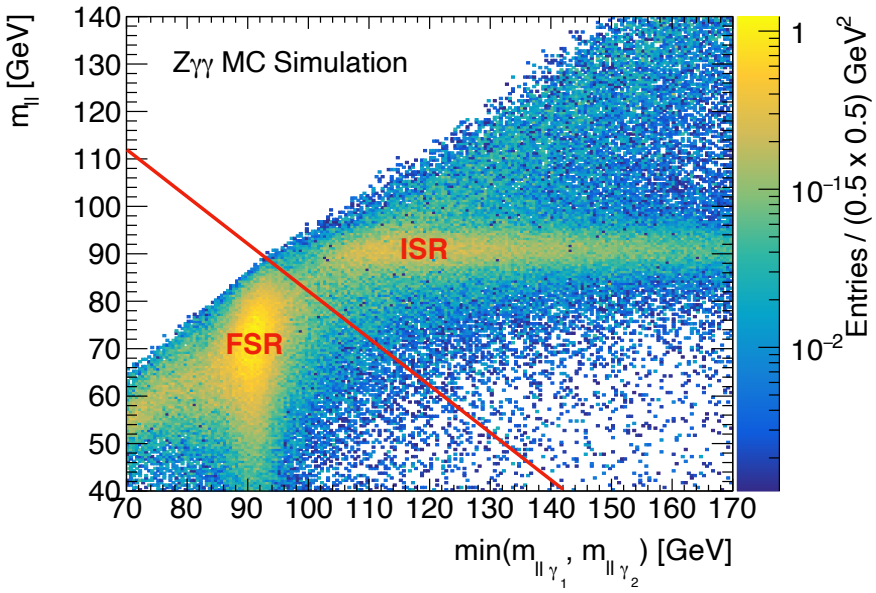


Figure 4.2: The correlations between the two-body invariant mass of the lepton pair, $m_{\ell\ell}$, and the minimum of the three-body invariant masses between the lepton pair and the (sub)leading photon, $\min_i m_{\ell\ell\gamma_i}$, are shown for a $Z\gamma\gamma$ MC simulation. An accumulation of events with initial state radiated (ISR) photons is visible at $m_{\ell\ell} \approx m_Z$ and $\min_i m_{\ell\ell\gamma_i} > m_Z$. In contrast, events with final state radiated (FSR) photons accumulate at $\min_i m_{\ell\ell\gamma_i} \approx m_Z$ and $m_{\ell\ell} < m_Z$. The red line visualizes the FSR cut, Equation 4.2, which is used for the ISR-enriched selection.

However, the dependence is much weaker. To tackle this p_{T} dependence, the currently used template fit method is modified by a new extraction procedure.

The procedure begins with one-dimensional models $M_{\gamma,j}(E_{\mathrm{T}}^{\mathrm{cone20}}|\vec{\theta}_0)$, which describe the $E_{\mathrm{T}}^{\mathrm{cone20}}$ distribution of prompt and fake photons, labelled with γ and j respectively. Here, $\vec{\theta}_0$ are the parameters of the one-dimensional model. In the next step, the p_{T} dependences of the parameters $\vec{\theta}_0 \equiv \vec{\theta}_0(p_{\mathrm{T}})$ are determined by extracting $M_{\gamma,j}(E_{\mathrm{T}}^{\mathrm{cone20}}|\vec{\theta}_0)$ for several p_{T} intervals. From this, two-dimensional $E_{\mathrm{T}}^{\mathrm{cone20}}$ templates

$$T_{\gamma,j}(E_{\mathrm{T}}^{\mathrm{cone20}}, p_{\mathrm{T}}|\vec{\theta}) \equiv M_{\gamma,j}(E_{\mathrm{T}}^{\mathrm{cone20}}|\vec{\theta}_0(p_{\mathrm{T}})) \tag{4.3}$$

are defined that include the p_{T} dependences of the parameters. Technically, $T_{\gamma,j}$ are two-dimensional conditional probability density functions (pdfs), i.e. the pdfs are normalized for each p_{T} value. By multiplying $T_{\gamma,j}$ with the corresponding p_{T} distribution, $f_X(p_{\mathrm{T}})$, the actual p_{T} distribution of the studied phase space X is taken into account. Therefore, the procedure is divided into two parts. The generic part, $T_{\gamma,j}$, can be used for leading and subleading photons in the $Z\gamma\gamma$ phase space and also for photons in different processes. The process-dependent part, $f_X(p_{\mathrm{T}})$, which models the p_{T} distribution of the studied phase space, needs to be extracted for each analysis individually.

The optimal parameters $\vec{\theta}_{\mathrm{opt}}$ are estimated by a two-dimensional binned maximum

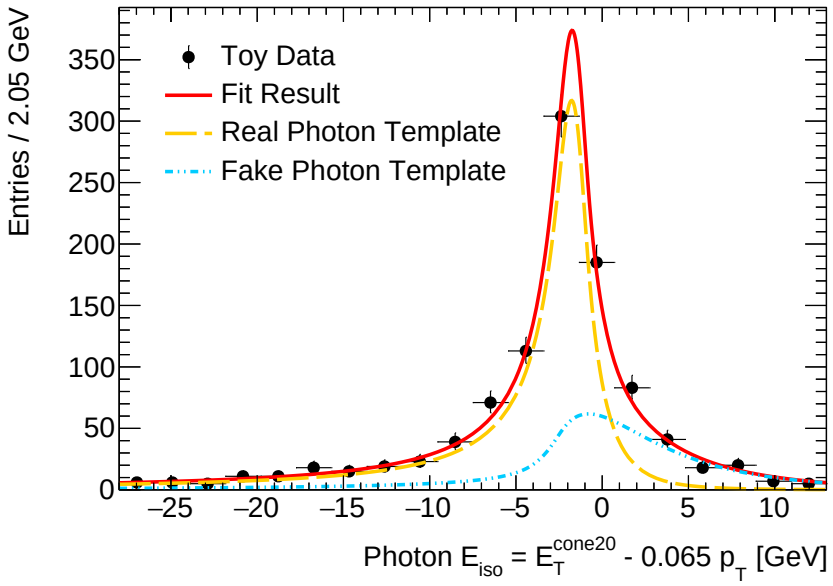


Figure 4.3: Sketch of the template fit approach with toy data. The isolation template of fake photons is broader and shifted to higher E_{iso} values with respect to the real photon isolation template. The shapes of the isolation templates are fixed, whereas the normalization factors are estimated by fitting the sum of real and fake photon isolation templates to data. The number of isolated photons is determined by integrating the isolation templates from $-\infty$ to 0.

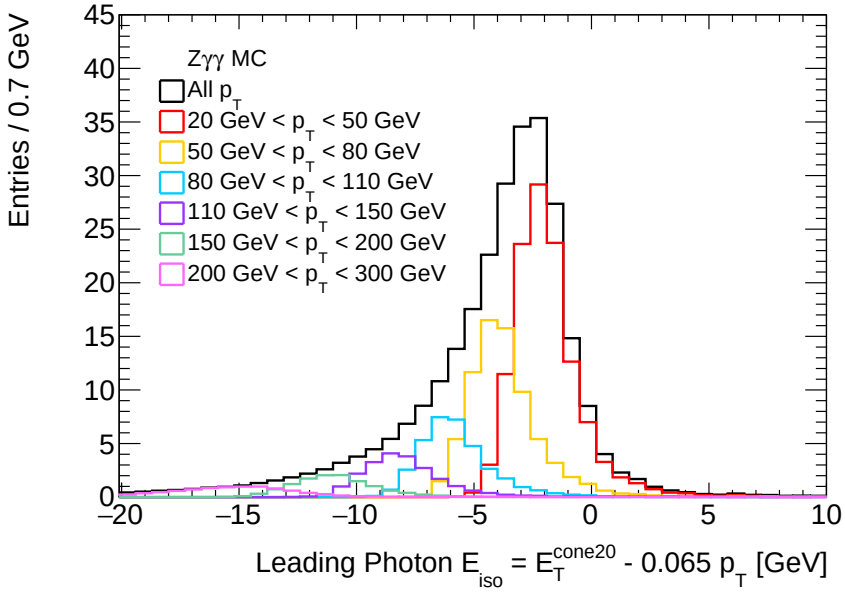


Figure 4.4: Photon p_{T} dependence of the E_{iso} template for leading photons in a $Z\gamma\gamma$ signal MC simulation.

likelihood fit of $T_{\gamma,j} \cdot f_{X_1}(p_{\mathrm{T}})$ within a phase space region X_1 which is chosen to yield a good compromise between a clean selection of real or fake photons and high statistics. The likelihood function reads

$$\mathcal{L}(\vec{\theta}|E_{\mathrm{T}}^{\mathrm{cone20}}, p_{\mathrm{T}}) = \prod_{i=1}^{N} T_{\gamma,j}(E_{\mathrm{T},i}^{\mathrm{cone20}}, p_{\mathrm{T},i}|\vec{\theta}) \cdot f_{X_1}(p_{\mathrm{T},i}), \tag{4.4}$$

$$\Rightarrow \vec{\theta}_{\mathrm{opt}} = \mathrm{argmax}_{\vec{\theta}}\, \mathcal{L}(\vec{\theta}|E_{\mathrm{T}}^{\mathrm{cone20}}, p_{\mathrm{T}}), \tag{4.5}$$

where $i = 1, ..., N$ runs over all events of X_1.

The $E_{\mathrm{T}}^{\mathrm{cone20}}$ template for another process or another region in the phase space, X_2, is given by

$$\mathcal{T}_{\gamma,j}(E_{\mathrm{T}}^{\mathrm{cone20}}) := \int_0^{\infty} T_{\gamma,j}(E_{\mathrm{T}}^{\mathrm{cone20}}, p_{\mathrm{T}}|\vec{\theta}_{\mathrm{opt}}) \cdot f_{X_2}(p_{\mathrm{T}})\, \mathrm{d}p_{\mathrm{T}}, \tag{4.6}$$

where the integration over p_{T} yields a projection of the two-dimensional model onto the $E_{\mathrm{T}}^{\mathrm{cone20}}$ axis. Here, $\mathcal{T}_{\gamma,j}(E_{\mathrm{T}}^{\mathrm{cone20}})$ already includes the optimal parameters $\vec{\theta}_{\mathrm{opt}}$ as well as the p_{T} distribution $f_{X_2}(p_{\mathrm{T}})$.

It should be mentioned that $T_{\gamma,j}(E_{\mathrm{T}}^{\mathrm{cone20}}, p_{\mathrm{T}}|\vec{\theta}_{\mathrm{opt}})$ is generically usable, i.e. independent from the sample X_2. Therefore, no new template extraction is needed and the same optimal parameters $\vec{\theta}_{\mathrm{opt}}$ are used that were extracted from X_1. In contrast, the p_{T} distribution $f_{X_2}(p_{\mathrm{T}})$ does obviously depend on X_2, but is extracted directly from X_2 as a binned histogram, which does not need a fit in the extraction process either.

4.3 Template for $E_{\mathrm{T}}^{\mathrm{cone20}}$ of Real Photons

The real photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ template, $T_{\gamma}(E_{\mathrm{T}}^{\mathrm{cone20}}, p_{\mathrm{T}}|\vec{\theta})$, is extracted from a $Z\gamma\gamma$ signal MC simulation. A double-sided crystal-ball (DSCB) function [20] is chosen for $M_{\gamma}(E_{\mathrm{T}}^{\mathrm{cone20}}|\vec{\theta}_0)$. The decision is driven based on the results of a previous analysis [5]. A DSCB is a central Gaussian with two power law tails on the left ("L") and right ("R") side:

$$M_{\gamma}(E_{\mathrm{T}}^{\mathrm{cone20}}|\vec{\theta}_0) = \begin{cases} A_{\mathrm{L}} \cdot \left(B_{\mathrm{L}} - \frac{E_{\mathrm{T}}^{\mathrm{cone20}} - \mu}{\sigma}\right)^{-n_{\mathrm{L}}}, & \text{for } \frac{E_{\mathrm{T}}^{\mathrm{cone20}} - \mu}{\sigma} < -\alpha_{\mathrm{L}} \\ \exp\left(-\frac{1}{2} \cdot \left[\frac{E_{\mathrm{T}}^{\mathrm{cone20}} - \mu}{\sigma}\right]^2\right), & \text{for } -\alpha_{\mathrm{L}} \leq \frac{E_{\mathrm{T}}^{\mathrm{cone20}} - \mu}{\sigma} \leq \alpha_{\mathrm{R}} \\ A_{\mathrm{R}} \cdot \left(B_{\mathrm{R}} + \frac{E_{\mathrm{T}}^{\mathrm{cone20}} - \mu}{\sigma}\right)^{-n_{\mathrm{R}}}, & \text{otherwise} \end{cases} \tag{4.7}$$

with

$$A_i = \left(\frac{n_i}{|\alpha_i|}\right)^{n_i} \cdot \exp\left(-\frac{|\alpha_i|^2}{2}\right) \quad \text{and} \quad B_i = \frac{n_i}{|\alpha_i|} - |\alpha_i|, \quad \text{for } i \in \{\mathrm{L}, \mathrm{R}\}, \tag{4.8}$$

and with $\vec{\theta}_0 = (\mu, \sigma, \alpha_{\mathrm{L}}, \alpha_{\mathrm{R}}, n_{\mathrm{L}}, n_{\mathrm{R}})$ being the free parameters of the DSCB. The shape of a DSCB function (with arbitrary values for the parameters) is shown in Figure 4.5.

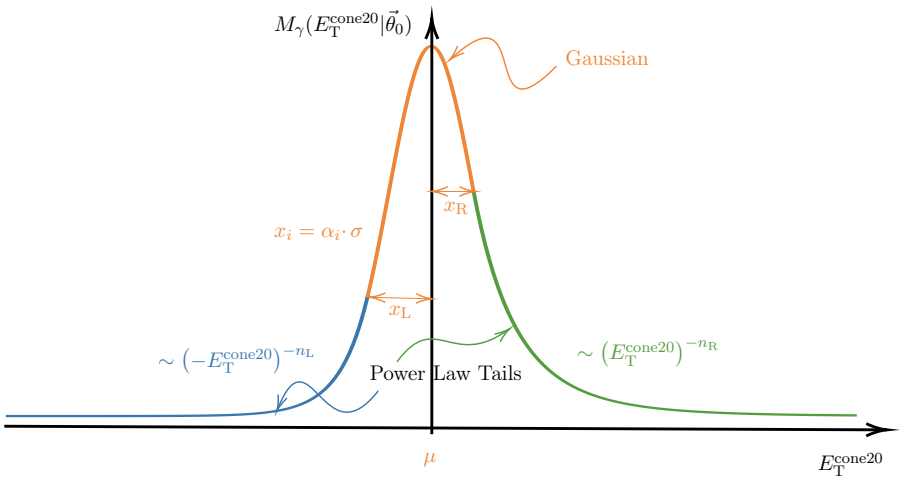


Figure 4.5: Shape of a DSCB function with arbitrary chosen values for the parameters. The parameters of a DSCB function are the mean μ and the standard deviation σ of the central Gaussian. Moreover, the exponents of the power law tails, n_i, as well as the transition points (where the central Gaussian merges into the power law tails), α_i in terms of σ, are free parameters.

Photon p_{T} Dependence of the Real Photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ Template

The p_{T} dependences of the paramters $\vec{\theta}_0$ are investigated with one-dimensional fits of $M_\gamma(E_{\mathrm{T}}^{\mathrm{cone20}}|\vec{\theta}_0)$ in several p_{T} intervals.

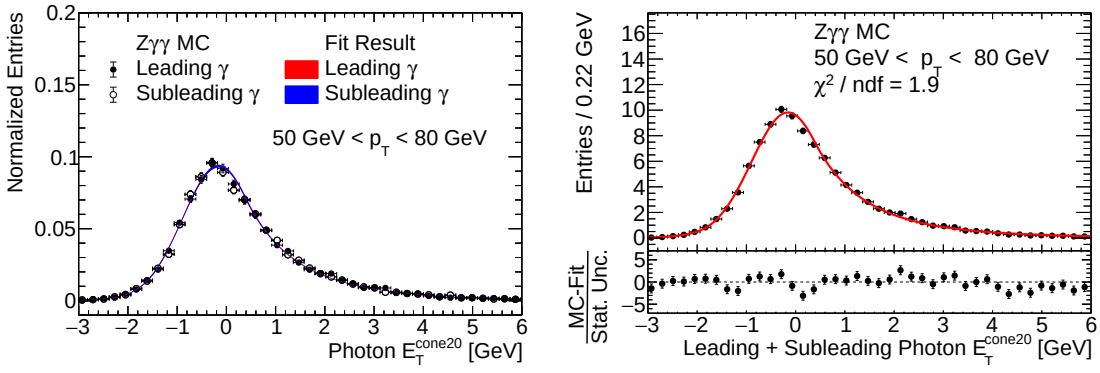
It is validated that the $E_{\mathrm{T}}^{\mathrm{cone20}}$ distributions of leading and subleading photons agree within their statistical uncertainties for each p_{T} interval. Figure 4.6a shows the comparison of leading and subleading photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ exemplarily for the p_{T} interval [50 GeV, 80 GeV]. The figure contains the bare comparison of the predicted distributions, as well as the comparison of the fit results that are visualized with their 1σ error bands. Good agreement is found between the predicted $E_{\mathrm{T}}^{\mathrm{cone20}}$ distributions of leading and subleading photons and also between the fitted $E_{\mathrm{T}}^{\mathrm{cone20}}$ templates of leading and subleading photons. Analogous figures for all p_{T} intervals are presented in Appendix B.1. As good agreement is found between the $E_{\mathrm{T}}^{\mathrm{cone20}}$ distribution of leading and subleading photons, the interval-wise one-dimensional fits of $M_\gamma(E_{\mathrm{T}}^{\mathrm{cone20}}|\vec{\theta}_0)$ are performed on a combined sample of leading and subleading photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ in order to increase statistics.

An often observed issue of fitting DSCB functions is the tendency of the exponents n_i to diverge to arbitrarily high values. In this analysis, n_{L} shows the problem just mentioned. It is validated that the differences of DSCB functions with various $n_{\mathrm{L}} \gtrsim 70$ are negligible, as $\left(-E_{\mathrm{T}}^{\mathrm{cone20}}\right)^{-70}$ converges fast enough to 0, and thus the relative differences are of permill order. Hence n_{L} is limited to $n_{\mathrm{L}} \in (0, 70)$. After performing the fit, n_i

are fixed and the fit is repeated. This has no influence on the optimal parameters, but improves the uncertainties of the remaining parameters.

Exemplarily, the fit for the p_{T} interval $[50\,\mathrm{GeV}, 80\,\mathrm{GeV}]$ is shown in Figure 4.6b. The fit results for all p_{T} intervals can be found in Appendix B.2. Especially the low-p_{T} intervals are reasonable well modeled by the DSCB function and no systematic trends are visible in the pull histogram (lower panel of Figure 4.6b), which visualizes the differences between the fit and the predicted distribution, normalized to the statistical uncertainties of the predicted distribution. In the higher-p_{T} intervals ($\gtrsim 150\,\mathrm{GeV}$), systematic deviations are seen in the central part. However, these p_{T} intervals are not relevant at current data statistics.

Furthermore, the normalized fit results of the different p_{T} intervals are visualized with their 1σ error bands in Figure 4.7. It clearly shows the p_{T} dependence of the $E_{\mathrm{T}}^{\mathrm{cone20}}$ template and especially that the $E_{\mathrm{T}}^{\mathrm{cone20}}$ template becomes broader with increasing p_{T}. Therefore, the parameters $\vec{\theta}_0$ are investigated as functions of p_{T}, which is shown in Figure 4.8. Since the p_{T} distribution within the p_{T} intervals is not uniform, as visualized in Figure 4.9, a bin-center correction is performed. The means of the interval-wise p_{T} distributions are taken as the corrected bin centers (which shifts the bin centers to lower p_{T} values) and the standard deviations are taken as the uncertainties of the corrected bins on the abscissa.



(a) (b)

Figure 4.6: Exemplary fit result for the p_{T} interval $[50\,\mathrm{GeV}, 80\,\mathrm{GeV}]$. The comparison of the $E_{\mathrm{T}}^{\mathrm{cone20}}$ variable of leading and subleading photons is shown in (a). The bare comparison of the predicted distributions is presented as well as the comparison of the fit results, which are visualized with their 1σ error band. Good agreement is observed between the $E_{\mathrm{T}}^{\mathrm{cone20}}$ distributions of leading and subleading photons. The fit result for the combined $E_{\mathrm{T}}^{\mathrm{cone20}}$ distribution of leading and subleading photons is shown in (b). The $E_{\mathrm{T}}^{\mathrm{cone20}}$ distribution is reasonably well described by the DSCB, no significant systematic trends are visible in the pull distribution (lower panel).

The standard deviation of the central Gaussian of the DSCB is found to be linearly dependent on p_{T}, parametrized by $\sigma(p_{\mathrm{T}}) = \sigma_0 + \sigma_1 \cdot p_{\mathrm{T}}$. The mean of the Gaussian can

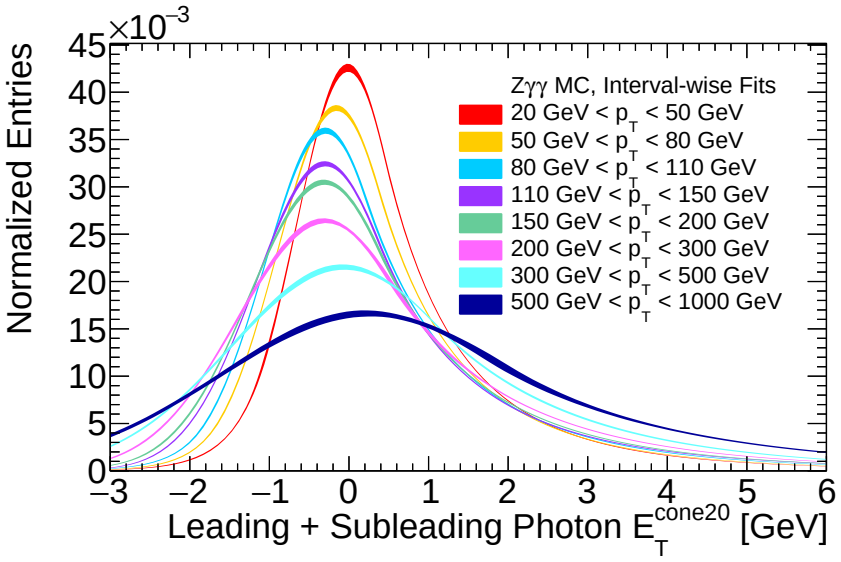


Figure 4.7: Comparison of the fit results for all p_{T} intervals visualized with their 1σ error band. The photon p_{T} dependence of the real photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ template is clearly visible, especially that the $E_{\mathrm{T}}^{\mathrm{cone20}}$ distribution becomes broader with increasing p_{T}.

be expressed with

$$\mu(p_{\mathrm{T}}) = D \cdot (1 - \exp\left[-a \cdot (p_{\mathrm{T}} - x_0)\right])^2 - c. \tag{4.9}$$

This dependence originates from the Morse potential [21] and is found empirically as it is more suitable than a third-order polynomial, which would have the same number of parameters. The physical interpretation of $\mu(p_{\mathrm{T}})$ is currently not fully understood. It is notable that the variations of μ are smaller than 1 GeV ($\mu \in (-0.3\,\mathrm{GeV}, 0.3\,\mathrm{GeV})$) and that the higher-$p_{\mathrm{T}}$ part ($p_{\mathrm{T}} \gtrsim 150\,\mathrm{GeV}$) is not relevant at current data statistics. However, Equation 4.9 is implemented into T_γ, since a comparable behavior is seen in the $Z\gamma$ MC simulation. It is validated that the $\mu(p_{\mathrm{T}})$ dependence is no artefact of the fit, but also observed for the actual means of the interval-wise $E_{\mathrm{T}}^{\mathrm{cone20}}$ distributions.

The transition points between the central Gaussian and the power law tails, α_i, are in first approximation linearly dependent on p_{T}, parametrized by $\alpha_i = \beta_i + \gamma_i \cdot p_{\mathrm{T}}$. The exponents n_i do not show a certain dependence on p_{T}.

After investigating the p_{T} dependences of the parameters as function of p_{T} and including them into a two-dimensional model (see Equation 4.3), a sanity check is performed to validate the approximated dependences. This is necessary mainly because the approximate linear dependence of α_i and the constancy of n_i are rough estimates. Therefore, $M_\gamma(E_{\mathrm{T}}^{\mathrm{cone20}}|\vec{\theta}_0(p_{\mathrm{T}}))$ is evaluated at the mean of the interval-wise p_{T} distributions (i.e. at the corrected bin center), and compared to the one-dimensional interval-wise fits. The

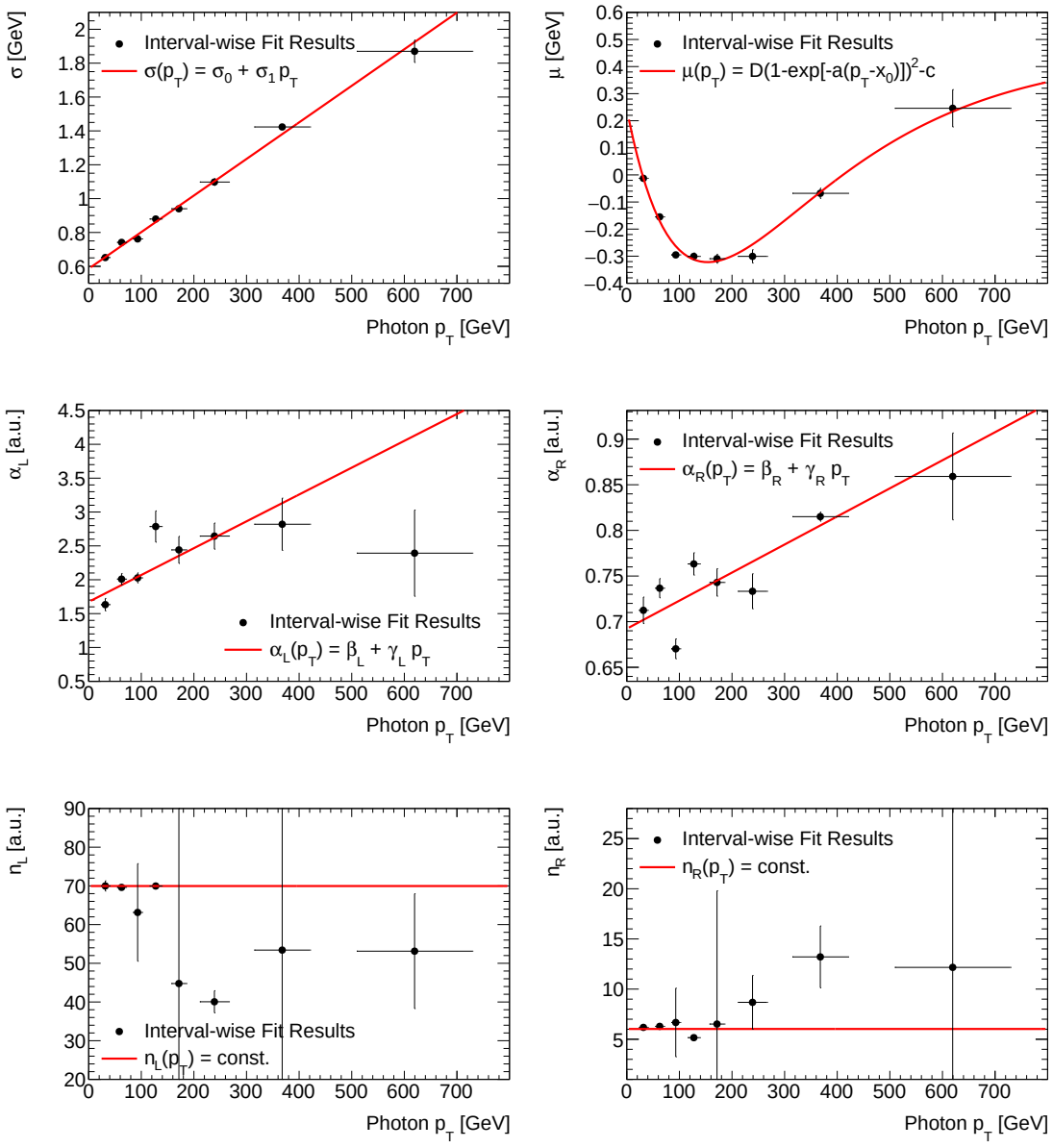


Figure 4.8: Parameters of the real photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ template as functions of photon p_{T}. Each point is obtained from the interval-wise one-dimensional fits. The values on the ordinate correspond to the fit parameters and their uncertainties. The values on the abscissa are obtained by a bin-center correction, and correspond to the mean and the standard deviation of the interval-wise p_{T} distributions.

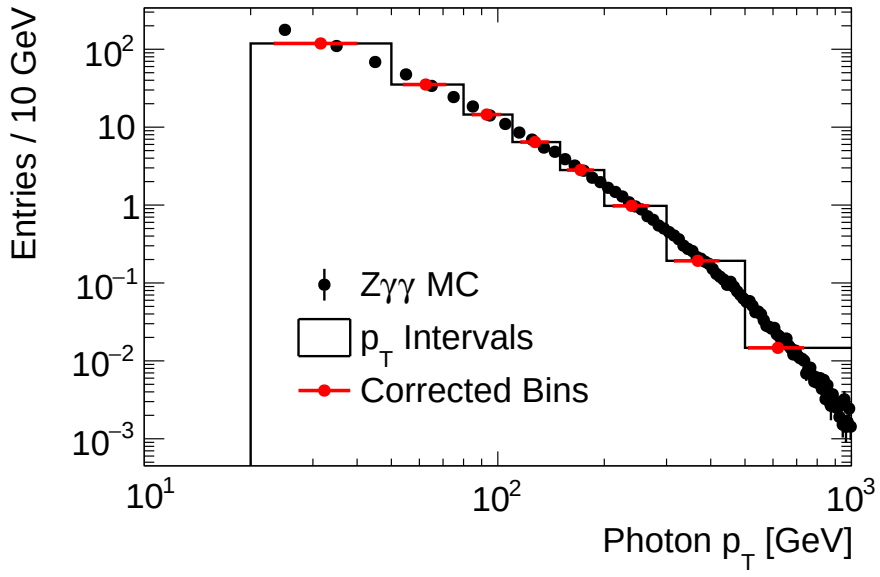


Figure 4.9: The non-uniform p_{T} distribution within the p_{T} intervals is taken into account by a bin-center correction. Therefore, the means of the interval-wise p_{T} distributions are taken as new bin centers, and the standard deviations are taken as half of the binwidth.

comparison is shown exemplarily for the p_{T} interval $[50\,\mathrm{GeV}, 80\,\mathrm{GeV}]$ in Figure 4.10 and all p_{T} intervals are presented in Appendix B.3. At all, the sanity checks agree well with the initial fits, which yields the conclusion that the approximated dependences describe the p_{T} dependence of the $E_{\mathrm{T}}^{\mathrm{cone20}}$ template reasonably well.

Optimal Parameter Extraction from the Leading Photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ Distribution in $Z\gamma\gamma$ Final States

To get a more precise estimate of the parameters $\vec{\theta} = (\sigma_0, \sigma_1, D, a, x_0, c, \beta_{\mathrm{L}}, \gamma_{\mathrm{L}}, \beta_{\mathrm{R}}, \gamma_{\mathrm{R}}, n_{\mathrm{L}}, n_{\mathrm{R}})$, a two-dimensional fit[4] of $T_\gamma(E_{\mathrm{T}}^{\mathrm{cone20}}, p_{\mathrm{T}}|\vec{\theta}) \cdot f_{X_1}(p_{\mathrm{T}})$, as described in Equation 4.4, is performed on the $E_{\mathrm{T}}^{\mathrm{cone20}}$ distribution of the leading photons in the $Z\gamma\gamma$ phase space.

The p_{T} distribution $f_{X_1}(p_{\mathrm{T}})$ is modeled with a histogram rather than a continuous fit model, as can be seen in Figure 4.11b. That is more flexible compared to describing the p_{T} distribution with an analytical model, and does not need determining optimal parameters by a fit.

The projection of the two-dimensional fit onto the $E_{\mathrm{T}}^{\mathrm{cone20}}$ and p_{T} axes are shown in Figure 4.11a and 4.11b, respectively. The optimal parameters $\vec{\theta}_{\mathrm{opt}}$ are given in Table 4.2. Good agreement is found between the real photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ template and the predicted

[4]It is observed that the parameters D, a, x_0 and c, which are describing the mean position of the central Gaussian, are strongly affected by the higher-p_{T} intervals, which are not relevant at current data statistics. To get a better estimation of the uncertainties of the remaining parameters, D, a, x_0 and c get fixed and the fit is repeated.

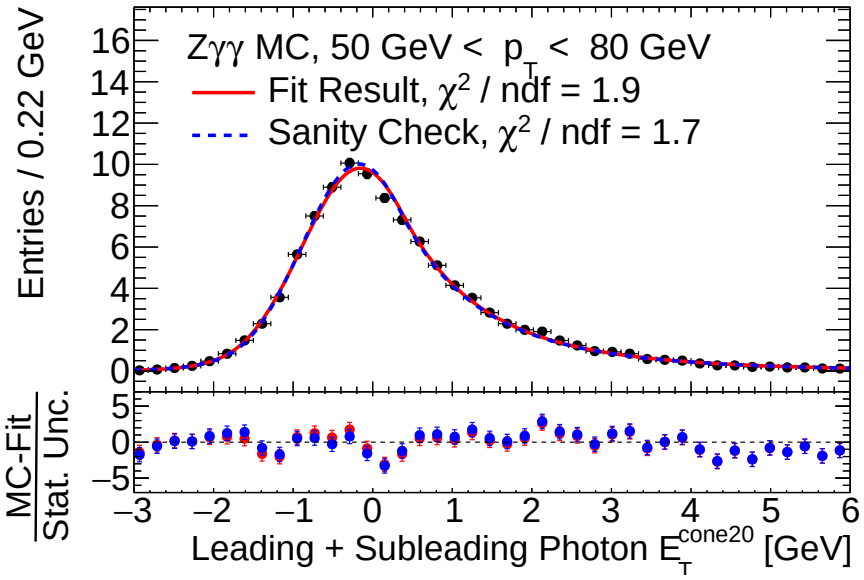


Figure 4.10: The one-dimensional fit within the p_{T} interval [50 GeV, 80 GeV] (red solid line) is compared to a sanity check (blue dashed line). This cross-check is performed to validate the approximated dependences $\vec{\theta}_0(p_{\mathrm{T}})$ by evaluating $M_\gamma(E_{\mathrm{T}}^{\mathrm{cone20}}|\vec{\theta}_0(p_{\mathrm{T}}))$ at the mean of the p_{T} distribution within [50 GeV, 80 GeV]. Good agreement is found between the initial fit and the sanity check.

$E_{\mathrm{T}}^{\mathrm{cone20}}$ distribution of leading photons in the $Z\gamma\gamma$ phase space. No significant systematic trend is visible in the pull histogram (lower panel).

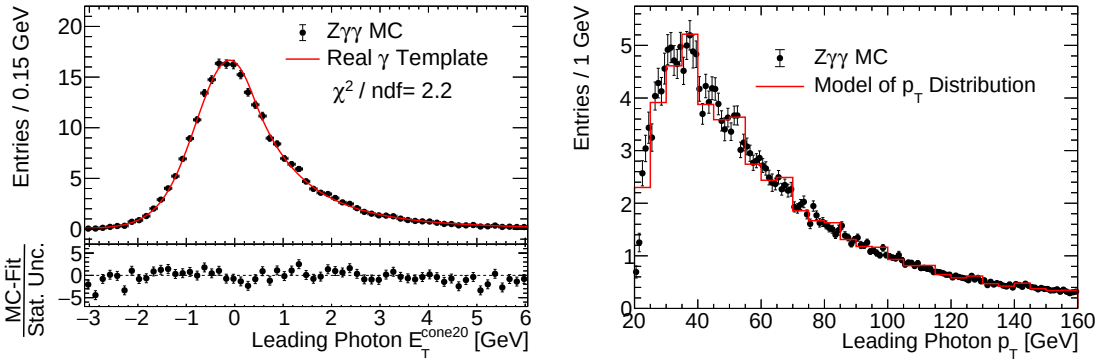
Comparison to the Subleading Photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ Distribution in $Z\gamma\gamma$ Final States

The real photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ template which is extracted from the leading photons of the $Z\gamma\gamma$ MC simulation is compared to the predicted $E_{\mathrm{T}}^{\mathrm{cone20}}$ distribution of subleading photons. Therefore, the real photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ template is multiplied with the p_{T} distribution of the subleading photons, according to Equation 4.6. Good agreement is found in the comparison, as shown in Figure 4.12. At this stage it is concluded that the real photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ template T_γ can be used for leading *and* subleading photons in $Z\gamma\gamma$ final states by multiplying it with the corresponding p_{T} distribution $f_X(p_{\mathrm{T}})$.

Validation of the Real Photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ Template in Data

In the last step, the MC-extracted real photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ template is validated with a control region (CR) in Run-2 data. The CR is FSR-enriched to select prompt photons. Therefore, the FSR cut is inverted and extended by a mass gap of $m_{\mathrm{g}} = 2\,\mathrm{GeV}$:

$$m_{\ell\ell} + \min\left(m_{\ell\ell\gamma_1}, m_{\ell\ell\gamma_2}\right) < 2m_Z - m_{\mathrm{g}}. \tag{4.10}$$



(a) (b)

Figure 4.11: Extraction of the real photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ template for the leading photons of the $Z\gamma\gamma$ signal MC simulation. The projections of the two-dimensional model onto the $E_{\mathrm{T}}^{\mathrm{cone20}}$ and onto the p_{T} axes are shown. From the projection onto the p_{T} axis, it can be seen that the p_{T} distribution $f_X(p_{\mathrm{T}})$ is modeled with a binned histogram rather than a continuous fit model.

Table 4.2: Optimal parameters of the extracted real photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ template.

$\vec{\theta}$	Optimal Parameter Value	Comment
σ_0	$(57.4 \pm 1.2) \times 10^{-2}\,\mathrm{GeV}$	
σ_1	$(230.6 \pm 8.4) \times 10^{-5}$	
β_{L}	$(158.9 \pm 3.9) \times 10^{-2}$	
γ_{L}	$(47.9 \pm 3.8) \times 10^{-4}\,\mathrm{GeV}^{-1}$	
β_{R}	$(700.5 \pm 3.8) \times 10^{-3}$	
γ_{R}	$(35.5 \pm 2.2) \times 10^{-5}\,\mathrm{GeV}^{-1}$	
n_{L}	$(699\,648.0 \pm 3.2) \times 10^{-4}$	at upper limit
n_{R}	$(608.5 \pm 2.7) \times 10^{-2}$	
D	$(7.8 \pm 2.4) \times 10^{-1}\,\mathrm{GeV}$	fixed
a	$(40.2 \pm 8.4) \times 10^{-4}\,\mathrm{GeV}^{-1}$	fixed
x_0	$(15.4 \pm 1.2) \times 10^{+1}\,\mathrm{GeV}$	fixed
c	$(32.2 \pm 1.1) \times 10^{-2}\,\mathrm{GeV}$	fixed

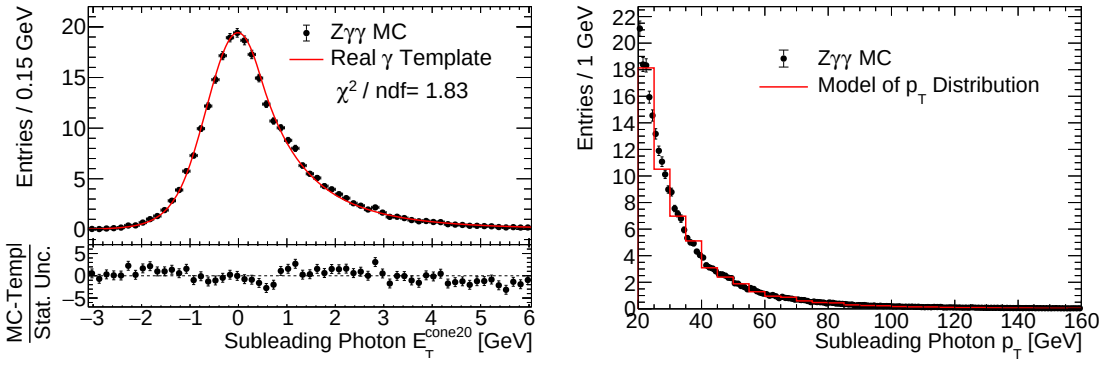


Figure 4.12: Comparison of the real photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ template (extracted from the leading photons of the $Z\gamma\gamma$ MC simulation) to the predicted $E_{\mathrm{T}}^{\mathrm{cone20}}$ distribution of subleading photons from the $Z\gamma\gamma$ MC simulation. It is notable that *no* fit is performed, but only the extracted $E_{\mathrm{T}}^{\mathrm{cone20}}$ template is multiplied with the p_{T} distribution of the subleading photons, according to Equation 4.6. Good agreement is observed between the template and the $E_{\mathrm{T}}^{\mathrm{cone20}}$ distribution of subleading photons.

The FSR selection is the cleanest way to select real photons in data. However, there can be still fake photons in the CR, and hence track isolation, Equation 3.6, is required to reject further fake photons from the CR. Again, no fit is performed, but T_γ is multiplied with the p_{T} distribution of the CR, according to Equation 4.6. The result is shown in Figure 4.13. The MC-extracted real photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ template agrees well with the $E_{\mathrm{T}}^{\mathrm{cone20}}$ distribution in data. Hence it is concluded that the $E_{\mathrm{T}}^{\mathrm{cone20}}$ distribution in the $Z\gamma\gamma$ MC simulation is modeled reasonably well.

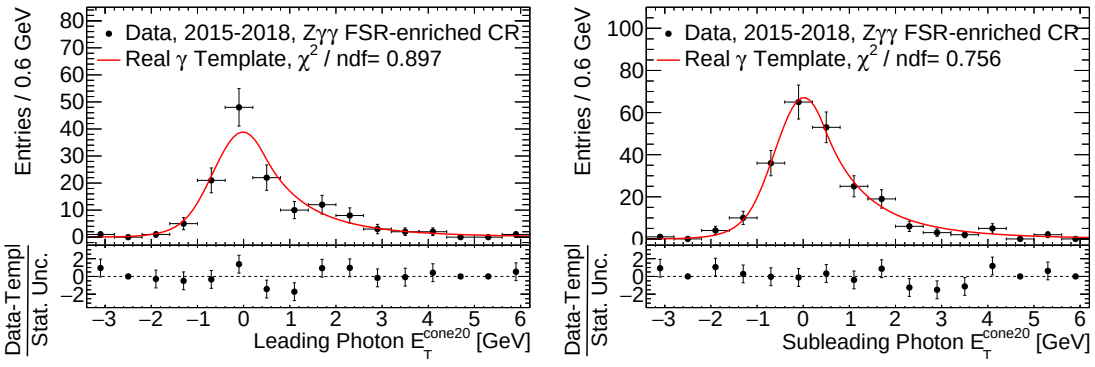


Figure 4.13: Validation of the extracted real photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ template within an FSR-enriched CR in Run-2 data.

4.4 Template for $E_{\mathrm{T}}^{\mathrm{cone20}}$ of Fake Photons

The fake photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ template, $T_j(E_{\mathrm{T}}^{\mathrm{cone20}}, p_{\mathrm{T}}|\vec{\theta})$, is extracted from $Z\gamma$ + jets MC simulations generated by MadGraphPythia8 at next-to-leading order QCD precision [17, 22, 23]. Fake photons are selected using photon identification selection criteria. More precisely, fake photons are selected requiring to fail the tight ID but pass the `LoosePrime4` working point [12]. The latter working point allows accessing shower shape observables, which are sensitive to prompt and fake photons. Different phase space regions are investigated:

- The leading photon is fake-matched[5] and the subleading photon passes tight ID (labelled as $Zj\gamma$).
- The leading photon passes tight ID and the subleading photon is fake-matched (labelled as $Z\gamma j$).
- Leading and subleading photons are fake-matched (labelled as Zjj).

The extraction of the fake photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ template is performed requiring the $Zj\gamma$ and $Z\gamma j$ selection. Based on the results of a previous analysis [5], the Bukin NovosibirskA function [24] is chosen to describe the $E_{\mathrm{T}}^{\mathrm{cone20}}$ distribution of fake photons. It is an asymmetric Gaussian stitched with two exponential tails on the left ("L") and right ("R") side:

$$M_j(E_{\mathrm{T}}^{\mathrm{cone20}}|\vec{\theta}_0) = \exp\left\{ \begin{array}{ll} \rho_{\mathrm{L}} \cdot \left(\frac{E_{\mathrm{T}}^{\mathrm{cone20}} - x_{\mathrm{L}}}{\mu - x_{\mathrm{L}}}\right)^2 - A_{\mathrm{L}}, & \text{for } E_{\mathrm{T}}^{\mathrm{cone20}} < x_{\mathrm{L}} \\ B \cdot \log^2\left[1 + C\frac{(E_{\mathrm{T}}^{\mathrm{cone20}} - \mu)}{\sigma}\right], & \text{for } x_{\mathrm{L}} \leq E_{\mathrm{T}}^{\mathrm{cone20}} < x_{\mathrm{R}} \\ \rho_{\mathrm{R}} \cdot \left(\frac{E_{\mathrm{T}}^{\mathrm{cone20}} - x_{\mathrm{R}}}{\mu - x_{\mathrm{R}}}\right)^2 - A_{\mathrm{R}}, & \text{for } x_{\mathrm{R}} \leq E_{\mathrm{T}}^{\mathrm{cone20}} \end{array} \right\} \quad (4.11)$$

with

$$x_i = \mu + \sqrt{2\log 2} \cdot \sigma \left(\frac{\xi}{a} \mp 1\right), \quad (4.12)$$

$$A_i = \log 2 \cdot \left(1 \mp \sqrt{\frac{2}{\log 2}} \frac{ab\left(E_{\mathrm{T}}^{\mathrm{cone20}} - x_i\right)}{\sigma\left(a \mp \xi\right)^2}\right), \quad (4.13)$$

$$B = -\frac{\log 2}{\log^2\left(1 + 2\xi \cdot (\xi - a)\right)}, \quad (4.14)$$

$$C = \sqrt{\frac{2}{\log 2}} a\xi, \quad (4.15)$$

[5] The term "fake-matched" denotes that the photon candidate passes the fake photon identification selection criteria as defined above.

where the negative (positive) sign is chosen for $i = \mathrm{L}$ ($i = \mathrm{R}$) and

$$a = \sqrt{\xi^2 + 1}, \quad b = \frac{\xi}{\log(a + \xi)}. \tag{4.16}$$

The Bukin function is parametrized by $\vec{\theta}_0 = (\mu, \sigma, \xi, \rho_\mathrm{L}, \rho_\mathrm{R})$. The overall shape of a Bukin function (with arbitrary chosen values for the parameters) is shown in Figure 4.14.

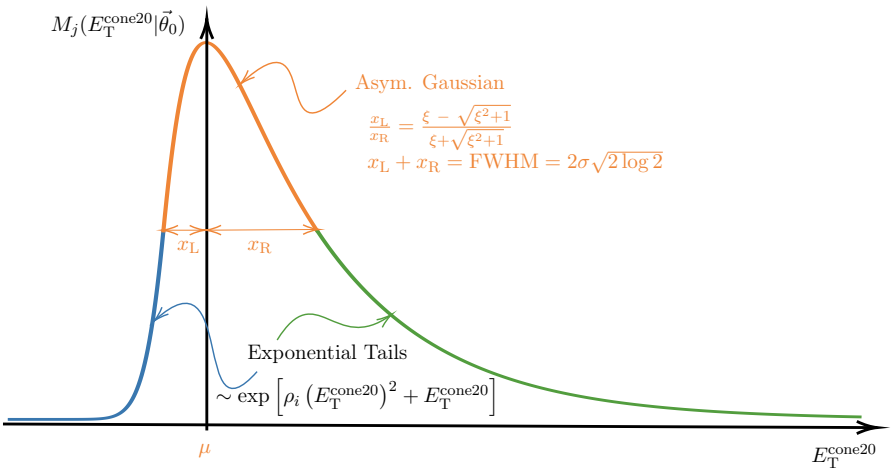
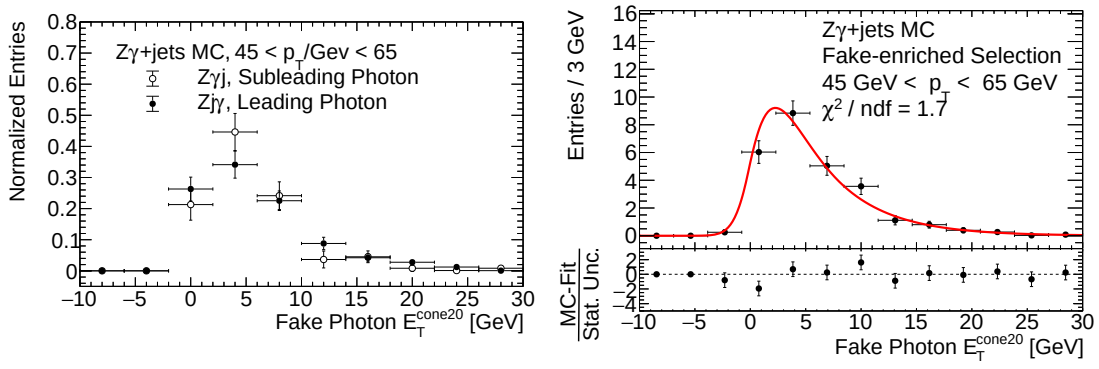


Figure 4.14: The shape of a Bukin NovosibirskA function is characterized by the peak position μ and the width $\mathrm{FWHM} = 2\sqrt{2\log 2}\sigma$ of the central Gaussian with asymmetry ξ. Furthermore, the tails are described by ρ_i.

Photon p_T Dependence of the Fake Photon $E_\mathrm{T}^{\mathrm{cone20}}$ Template

An analogous procedure as described in Section 4.3 is performed for fake photons in several p_T intervals. At first, it is validated that the $E_\mathrm{T}^{\mathrm{cone20}}$ distributions of leading and subleading fake photons in the $Zj\gamma$ and $Z\gamma j$ selections agree within their statistical uncertainties. The comparisons can be found in Appendix C.1 and exemplarily for the p_T interval $[45\,\mathrm{GeV}, 65\,\mathrm{GeV}]$ in Figure 4.15a. As no significant differences are found between the predicted $E_\mathrm{T}^{\mathrm{cone20}}$ distributions of leading and subleading fake photons, the interval-wise one-dimensional fits of $M_j(E_\mathrm{T}^{\mathrm{cone20}}|\vec{\theta}_0)$ are performed on a combined $E_\mathrm{T}^{\mathrm{cone20}}$ distribution of leading and subleading fake photons from the $Zj\gamma$ and $Z\gamma j$ selections, in order to increase statistics. The results of the interval-wise fits are presented in Appendix C.2 and exemplarily for the p_T interval $[45\,\mathrm{GeV}, 65\,\mathrm{GeV}]$ in Figure 4.15b.

To correct for the non-uniform p_T distribution within each p_T interval, a bin-center correction is performed analogous to the extraction process of the real photon $E_\mathrm{T}^{\mathrm{cone20}}$ template. A strong p_T dependence of the fake photon $E_\mathrm{T}^{\mathrm{cone20}}$ template is found, as visible in Figure 4.16. The peak position μ as well as the width σ of the Gaussian increase linearly with p_T, parametrized by $\mu(p_\mathrm{T}) = \mu_0 + \mu_1 \cdot p_\mathrm{T}$ and $\sigma(p_\mathrm{T}) = \sigma_0 + \sigma_1 \cdot p_\mathrm{T}$, respectively.



(a) (b)

Figure 4.15: The comparison of the predicted $E_{\mathrm{T}}^{\mathrm{cone20}}$ distributions of leading and subleading fake photons in the $Zj\gamma$ and $Z\gamma j$ selections is shown for the p_{T} interval $[45\,\mathrm{GeV}, 65\,\mathrm{GeV}]$ in (a). No significant differences are visible. The fit result for the combined $E_{\mathrm{T}}^{\mathrm{cone20}}$ distribution of leading and subleading fake photons in the $Zj\gamma$ and $Z\gamma j$ selections is shown exemplarily for the same p_{T} interval in (b).

The asymmetry ξ decreases in first approximation linearly, $\xi(p_{\mathrm{T}}) = \xi_0 + \xi_1 \cdot p_{\mathrm{T}}$, and ρ_i do not show a certain p_{T} dependence[6].

To validate the dependences $\vec{\theta_0}(p_{\mathrm{T}})$, a sanity check is performed analogous to the extraction process of the real photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ template. The results can be found in Appendix C.3 and exemplarily for the p_{T} interval $[45\,\mathrm{GeV}, 65\,\mathrm{GeV}]$ in Figure 4.17. The higher χ^2 value of the sanity check is mainly driven by a small overestimate at the right tail. In contrast, the central part agrees reasonably well with the $E_{\mathrm{T}}^{\mathrm{cone20}}$ distribution. Hence, it is concluded that the approximated p_{T} dependences describe the p_{T} dependences of the fake photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ template reasonable well.

Optimal Parameter Extraction from the Fake Photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ Distribution in the $Zj\gamma$ Selection

The parameters $\vec{\theta} = (\mu_0, \mu_1, \sigma_0, \sigma_1, \xi_0, \xi_1, \rho_{\mathrm{L}}, \rho_{\mathrm{R}})$ are extracted by a two-dimensional fit of $T_j(E_{\mathrm{T}}^{\mathrm{cone20}}, p_{\mathrm{T}}|\vec{\theta}) \cdot f_{X_1}(p_{\mathrm{T}})$ in the $Zj\gamma$ selection according to Equation 4.4. The fit procedure is analogous to Section 4.3. The projection of the two-dimensional template onto the $E_{\mathrm{T}}^{\mathrm{cone20}}$ axis as well as onto the p_{T} axis is shown in Figure 4.18 and the optimal parameters $\vec{\theta}_{\mathrm{opt}}$ are given in Table 4.3.

It is notable that $\mu_1 = (6.14 \pm 0.41)\%$ is likely the reason for the definition of the loose isolation working point. The subtraction of $6.5\% \cdot p_{\mathrm{T}}$ in Equation 3.5 compensates the shift of the peak position $\mu(p_{\mathrm{T}}) = \mu_0 + \mu_1 \cdot p_{\mathrm{T}}$. Hence, the peak of the fake photon isolation template is kept at a constant position $E_{\mathrm{iso}} \approx \mu_0$.

[6] Based on official RooFit recommendations [25], the parameters ρ_i are limited to $(-1, 0)$ and $(0, 1)$ for $i = \mathrm{L}$ and $i = \mathrm{R}$, respectively. However, they have the tendency to stay at the lower limit. Hence, the fits are repeated with fixed ρ_i, in order to estimate the uncertainties of the remaining parameters with higher accuracy.

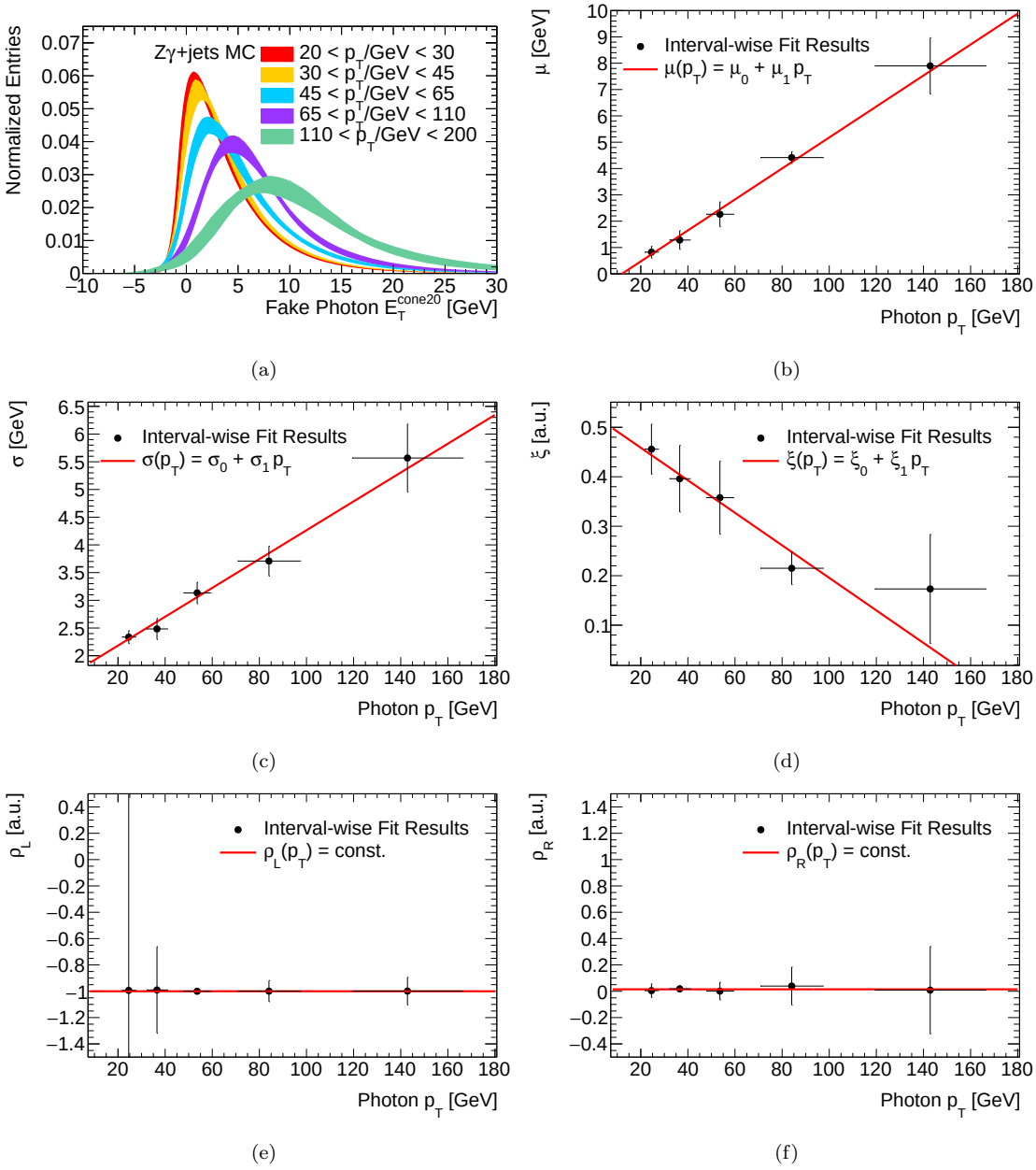


Figure 4.16: Investigation of the p_{T} dependence of the fake photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ template. A comparison of the fit results in several p_{T} intervals with their 1σ error bands is shown in (a). The strong p_{T} dependence is clearly visible. Especially that the fake photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ template gets broader and shifted towards higher $E_{\mathrm{T}}^{\mathrm{cone20}}$ values with increasing photon p_{T}. The parameters of the fake photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ template are taken as function of p_{T} in (b - f). Here, each point is obtained from the one-dimensional fits. The values on the ordinate correspond to the fit parameters and their uncertainties. The values on the abscissa are obtained by a bin-center correction, and correspond to the mean and the standard deviation of the interval-wise p_{T} distributions.

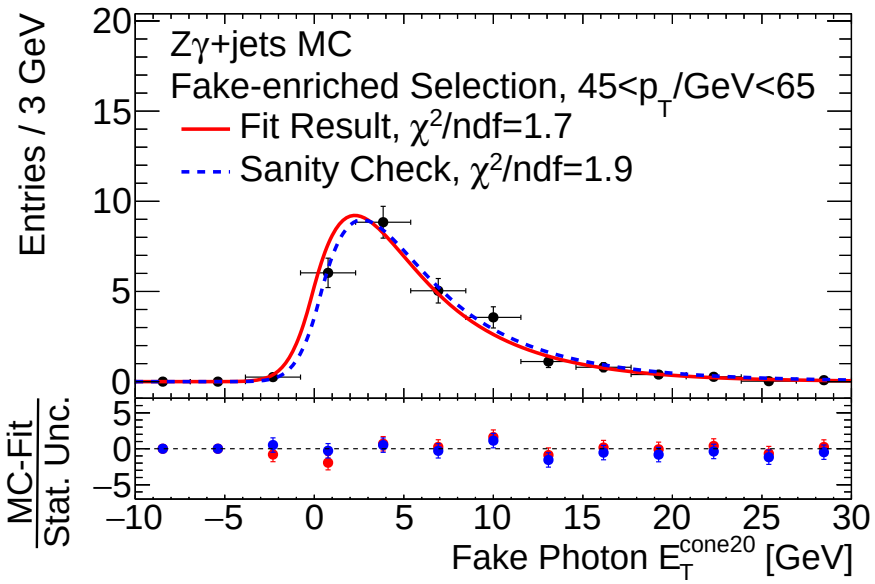


Figure 4.17: The one-dimensional fit within the p_{T} interval [45 GeV, 65 GeV] (red solid line) is compared to a sanity check (blue dashed line). This cross-check is performed to validate the approximated dependences $\vec{\theta}_0(p_{\mathrm{T}})$ by evaluating $M_j(E_{\mathrm{T}}^{\mathrm{cone20}}|\vec{\theta}_0(p_{\mathrm{T}}))$ at the mean of the p_{T} distribution within [45 GeV, 65 GeV]. The larger χ^2 value of the sanity check is driven by the small overestimate at the right tail. However, the central part is modeled more precisely by the sanity check. Hence, it is concluded that the approximated dependences describe the $E_{\mathrm{T}}^{\mathrm{cone20}}$ distribution reasonably well.

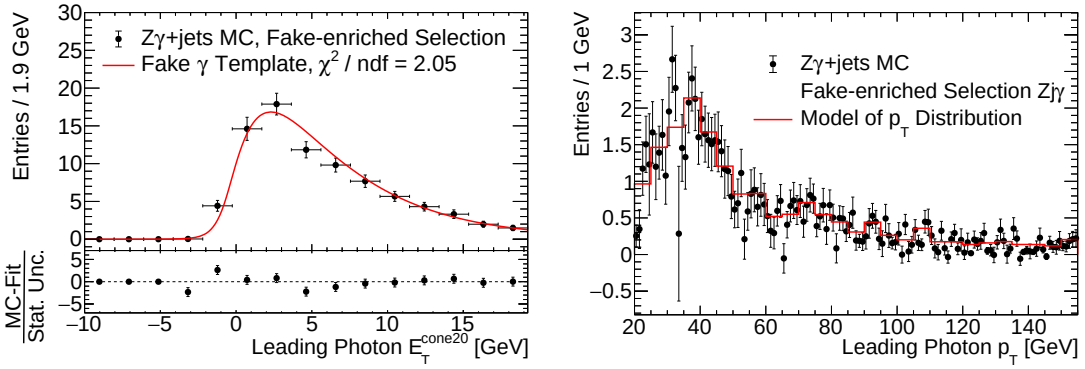


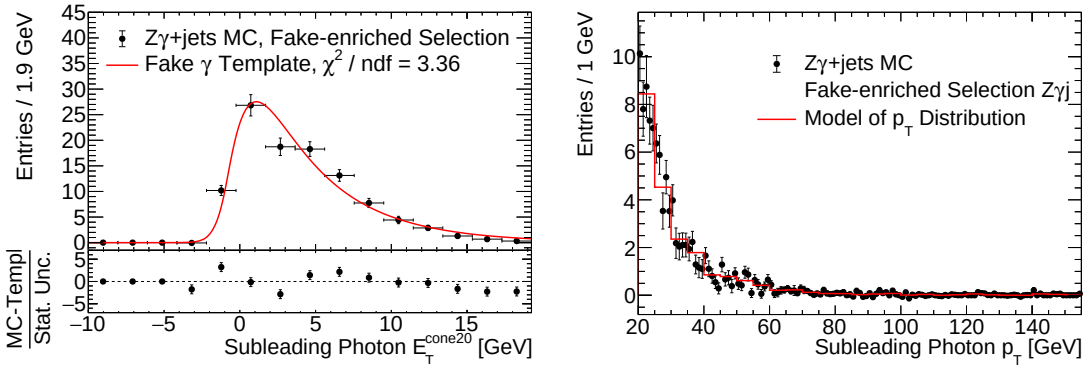
Figure 4.18: Extraction of the fake photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ template for the leading photon of the $Zj\gamma$ selection. The projections of the two-dimensional model onto the $E_{\mathrm{T}}^{\mathrm{cone20}}$ and onto the p_{T} axes are shown.

Table 4.3: Optimal parameters of the extracted fake photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ template.

$\vec{\theta}$	Optimal Parameter Value	Comment
μ_0	$-(6.0 \pm 1.8) \times 10^{-1}\,\mathrm{GeV}$	
μ_1	$(61.4 \pm 4.1) \times 10^{-3}$	
σ_0	$(171.8 \pm 9.0) \times 10^{-2}\,\mathrm{GeV}$	
σ_1	$(27.3 \pm 2.1) \times 10^{-3}$	
ξ_0	$(54.0 \pm 2.5) \times 10^{-2}$	
ξ_1	$-(30.5 \pm 3.9) \times 10^{-4}\,\mathrm{GeV}^{-1}$	
ρ_{L}	$-(9\,999.2 \pm 1.7) \times 10^{-4}$	at lower limit
ρ_{R}	$(2.5 \pm 2.1) \times 10^{-2}$	

Comparison to the Fake Photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ Distribution in the $Z\gamma j$ Selection

In the next step, the fake photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ template which is extracted from the fake photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ distribution in the $Zj\gamma$ selection is compared to the $E_{\mathrm{T}}^{\mathrm{cone20}}$ distribution of subleading fake photons in the $Z\gamma j$ selection. Therefore, the fake photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ template is multiplied with the p_{T} distribution of the subleading photons in the $Z\gamma j$ selection according to Equation 4.6. With respect to the extraction process of the real photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ template, the extraction of the fake photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ template is limited by low statistics. To take the larger uncertainties of the optimal parameters into account, the parameters can float within the uncertainties of the optimal parameters (see Table 4.3) while performing a fit in the comparison process. The results are shown in Figure 4.19. No deviations are seen within the scope of statistical uncertainties. Hence it is concluded that the same fake photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ template can be used for leading *and* subleading fake photons.

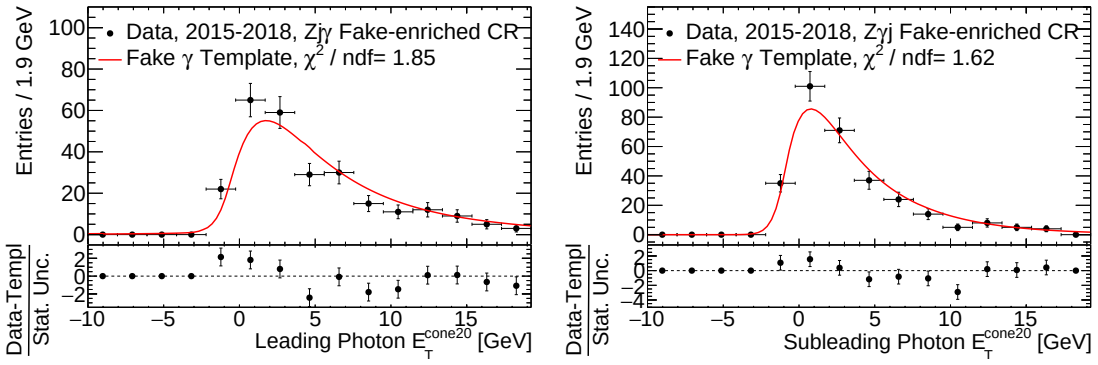


(a) (b)

Figure 4.19: Comparison of the fake photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ template (extracted from fake photons of the $Zj\gamma$ selection) with the predicted $E_{\mathrm{T}}^{\mathrm{cone20}}$ distribution of fake photons in the $Z\gamma j$ selection. The parameters of the fake photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ template are allowed to float within the uncertainties of $\vec{\theta}_{\mathrm{opt}}$. Good agreement is observed in the comparison.

Validation of the Fake Photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ Template in Data

In the last step, the MC-extracted fake photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ template is validated within a fake-enriched CR in the ATLAS Run-2 dataset. The same ID criteria as mentioned in the beginning of Section 4.4 are used to select fake photons in the Run-2 dataset. The results for the leading photon of the $Zj\gamma$ selection and the subleading photon of the $Z\gamma j$ selection are shown in Figure 4.20. Again, the parameters are found by a fit letting the parameters float within the uncertainties of the optimal parameters. No significant differences are found, and hence it is concluded that the $E_{\mathrm{T}}^{\mathrm{cone20}}$ distribution in the $Z\gamma$ + jets MC simulation is described adequately.



(a) (b)

Figure 4.20: Validation of the extracted fake photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ template within a fake-enriched CR in the ATLAS Run-2 dataset.

Chapter 5

Fake Photon Background Estimation for the $Z\gamma\gamma$ Analysis

In this chapter, the fake photon background of the full ATLAS Run-2 $Z\gamma\gamma$ phase space is determined using the template fit method. The results are compared to a previous analysis which studied the same $Z\gamma\gamma$ phase space [5, 6]. This is done to validate the methodology used in the template-extraction process.

$Z\gamma\gamma$ final states of the Run-2 dataset are selected according to the event selection described in Chapter 4.1. The basic idea of the template fit method, as introduced in Chapter 4.2, is fixing the shapes of the real and fake photon isolation templates while fitting the normalization factors for both templates simultaneously.

In the $Z\gamma\gamma$ analysis, it is sufficient to only separate real and fake photons by the isolation energy, and apply the track isolation with a hard cut [5], as described in Equation 3.6. This additional cut simplifies the procedure, since no isolation templates need to be extracted for p_{iso}.

However, in sense of the generalization-idea it is favorable for future analyses to include also not-track-isolated photons into the fake photon isolation template instead of rejecting them with the cut on the track isolation. This avoids biasing the data and leads to a better distinction between real and fake photons, as the cut on the track isolation shifts the fake photon isolation template closer to the real photon isolation template.

In the previous $Z\gamma\gamma$ analysis, the first approach was used, i.e. the hard cut on the track isolation was applied in the selection process. For better validability, this cut is also required in the selection used in this chapter.

Due to the additional cut, the shapes of the isolation templates are expected to change with respect to the isolation templates that are extracted in Chapter 4 without the cut on the track isolation. However, the dependences $\vec{\theta}_0(p_{\text{T}})$ remain unchanged, but the optimal parameters $\vec{\theta}_{\text{opt,trackiso}}$ change slightly, as shown in Table D.1.

The template fit is performed simultaneously for both photons of the $Z\gamma\gamma$ final states.

The two-dimensional isolation energy of leading and subleading photons is given by

$$I = N_{\gamma\gamma}\mathcal{T}'_{\gamma\gamma} + N_{\gamma j}\mathcal{T}'_{\gamma j} + N_{j\gamma}\mathcal{T}'_{j\gamma} + N_{jj}\mathcal{T}'_{jj}, \tag{5.1}$$

where N_{xy} are the normalization factors and $\mathcal{T}'_{xy}$ are the two-dimensional isolation templates. Here, the prime denotes that the $E_{\mathrm{T}}^{\mathrm{cone20}}$ templates are transformed to the isolation energy $E_{\mathrm{iso}} = E_{\mathrm{T}}^{\mathrm{cone20}} - 0.065 \cdot p_{\mathrm{T}}$. This is done to estimate the number of isolated photons by integrating the E_{iso} template from $-\infty$ to 0 according to Equation 3.5 after determining the normalization factors with the simultaneous fit.

The two-dimensional isolation templates $\mathcal{T}'_{xy}$ can be split into a product of the two corresponding one-dimensional isolation templates if the correlation between the leading and the subleading photon can be neglected. This is the case for $\mathcal{T}'_{\gamma\gamma}$, $\mathcal{T}'_{\gamma j}$ and $\mathcal{T}'_{j\gamma}$ but not for $\mathcal{T}'_{jj}$, due to color connection between the jets [5]. Hence,

$$\mathcal{T}'_{xy} = \mathcal{T}'_{x} \times \mathcal{T}'_{y} \quad \text{for } (x,y) \in \{(\gamma,\gamma),(\gamma,j),(j,\gamma)\}, \tag{5.2}$$

where $\mathcal{T}'_{\gamma,j}$ is defined similar to Equation 4.6:

$$\mathcal{T}'_{\gamma,j}(E_{\mathrm{iso}}) := \int_0^\infty T'_{\gamma,j}(E_{\mathrm{T}}^{\mathrm{cone20}}, p_{\mathrm{T}}|\vec{\theta}_{\mathrm{opt,trackiso}}) \cdot f(p_{\mathrm{T}})\mathrm{d}p_{\mathrm{T}}. \tag{5.3}$$

Here,

$$T_{\gamma,j}(E_{\mathrm{T}}^{\mathrm{cone20}}, p_{\mathrm{T}}) \to T'_{\gamma,j}(E_{\mathrm{T}}^{\mathrm{cone20}}, p_{\mathrm{T}}) \equiv T_{\gamma,j}(E_{\mathrm{T}}^{\mathrm{cone20}} - 0.065 p_{\mathrm{T}}, p_{\mathrm{T}}) \tag{5.4}$$

is the mentioned transformation of Equation 4.3 and $f(p_{\mathrm{T}})$ are the p_{T} distributions of the leading and subleading photons of the $Z\gamma\gamma$ final state, respectively.

The two-dimensional isolation template $\mathcal{T}'_{jj}$ is obtained data driven since the MC description is limited by low statistics and therefore can not be validated accurately. The correlations between the leading and the subleading photons are modeled with an adaptive kernel estimation [26], which is a superposition of Gaussian kernels with dynamic width. A smaller width is chosen for higher event density and vice versa. The kernel estimation is shown in Figure 5.1.

While the shapes of the two-dimensional isolation templates $\mathcal{T}'_{xy}$ are fixed, the normalization factors N_{xy} are estimated with a maximum likelihood fit of the two-dimensional isolation energy within the $Z\gamma\gamma$ signal region (SR). Before performing the two-dimensional fit, minor background sources are subtracted from the SR, including $t\bar{t}\gamma\gamma$, $ZZ(\to \ell\ell\ell\ell)$, $Z(\to \ell\ell)H(\to \gamma\gamma)$ and $WZ\gamma(\to \ell\nu\ell\ell\gamma)$. After performing the fit, the pile-up contribution

$$N_{\mathrm{PU}} = -8.15 \pm 0.27(\mathrm{stat}) \pm 1.36(\mathrm{sys}) \tag{5.5}$$

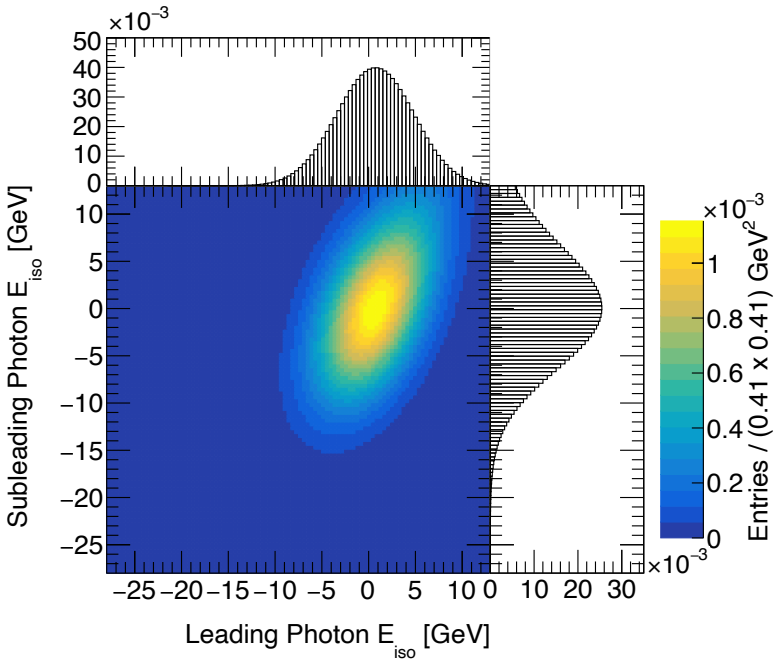


Figure 5.1: The leading and subleading isolation energy $E_{\text{iso}} = E_{\text{T}}^{\text{cone20}} - 0.065 p_{\text{T}}$ of fake photons in the Zjj selection are described by the the two-dimensional template $\mathcal{T}'_{jj}$. Due to limited statistics, $\mathcal{T}'_{jj}$ is obtained data driven. The correlations between leading and subleading photons in the Zjj phase space are modeled with an adaptive kernel estimation. This is a superposition of Gaussian kernels with dynamic width. The upper and the right panel show the projection of the kernel estimation onto the leading and subleading photon isolation energy, respectively.

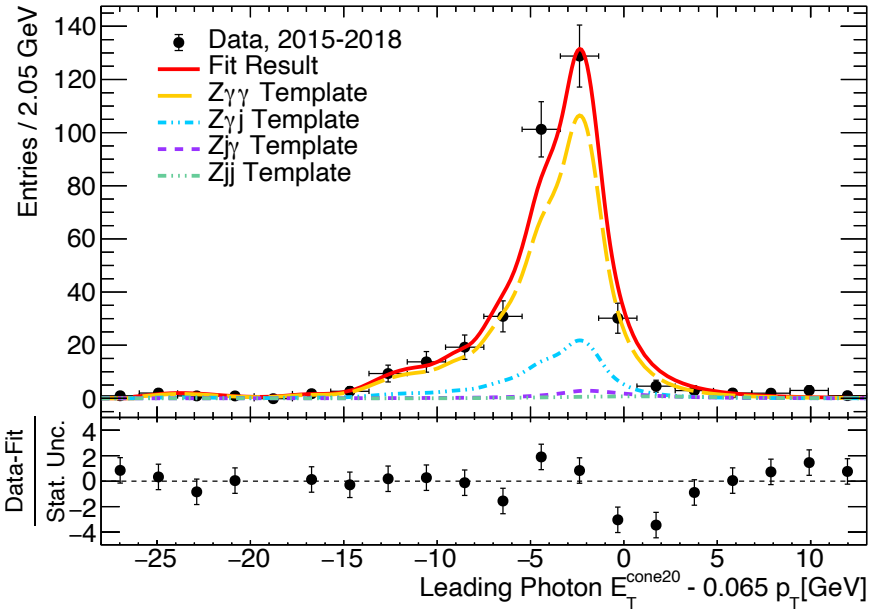
is subtracted from the number of isolated real photons. The background estimations are taken from [5]. The number of isolated photons is determined with

$$N_{xy}^{\text{isolated}} = N_{xy} \cdot \underbrace{\int_{-\infty}^{0} \mathcal{T}'_{xy} \mathrm{d}E_{\text{iso}}}_{\in[0,1]} . \tag{5.6}$$

The fit result is shown in Figure 5.2 and the background estimation is summarized in Table 5.1. The results are compared to a previous analysis that studied the same $Z\gamma\gamma$ phase space [5, 6]. In the previous analysis, the matrix method and the template fit method are used to estimate the fake photon background of the ATLAS Run-2 $Z\gamma\gamma$ phase space.

The left tail as well as the central part of the isolation templates are in good agreement with the data, as shown in Figure 5.2. The overestimation of the isolation templates at the transition to the right tail is also observed in [5].

Overall, the background estimation of the presented analysis agrees well with the previous analysis. It should be mentioned that the uncertainties are estimated with different methods. In [5], the statistical uncertainties are estimated by performing the maximum



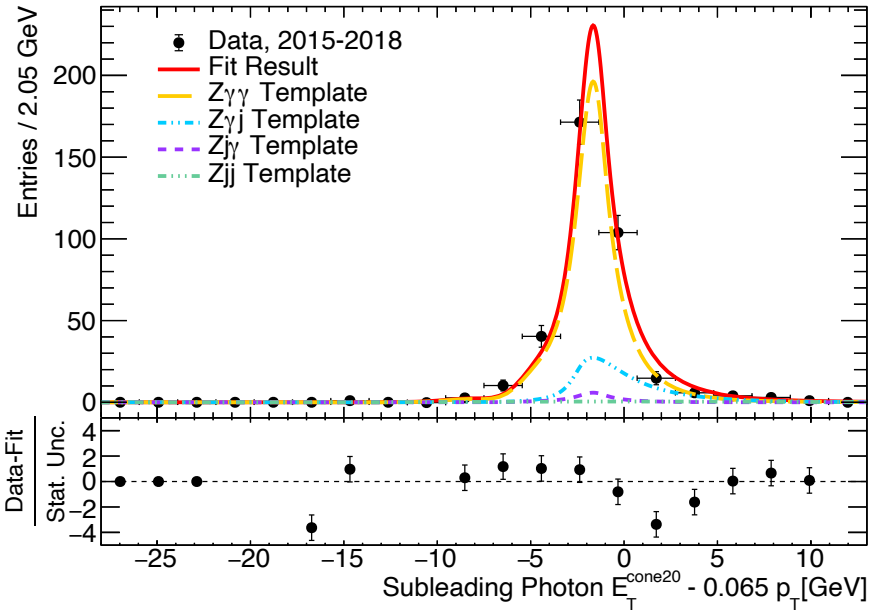


Figure 5.2: Projection of the two-dimensional template fit onto the leading and subleading isolation energy. The sum of the isolation templates agrees well with the data. The slight overestimation at the transition between the central part and the right tail can also be seen in [5]. Due to the background subtraction, there are several bins within the left tail that have a negative number of entries. However, these bins have small absolute numbers of entries and corresponding small statistical uncertainties, which causes large negative values in the pull histogram. These bins are *not* visible in the pull histogram (lower panel), as it is limited to the range $[-5, 5]$ for better visualization of the pull distribution of the remaining bins.

likelihood fit for 1000 replica datasets (so-called toy MC datasets). The statistical uncertainties are the root-mean-square of the obtained normalization factors. In contrast to that, the uncertainties presented in this analysis are the errors of the fit parameter directly obtained from the Hessian matrix.

Even by only taking the statistical uncertainties of [5] into account, the deviations are in the 1σ range for the comparison with the template fit in [5] except for $N_{\gamma j}$, where the results differ by 1.5σ. At this stage, correlations between the two methods are not taking into account. The comparison with the matrix method shows slightly larger deviations, especially $N_{j\gamma}$ is in tension when only considering the statistical uncertainties. However, the physical expectation of $N_{\gamma j} > N_{j\gamma}$ indicates that $N_{j\gamma}$ may be overestimated by the matrix method of [5]. Furthermore, by comparing the different methods (matrix method in [5] and template fit in this analysis) also the systematic uncertainties of [5] should be taken into account, as the correlations of the systematic uncertainties between the matrix method and the template fit are expected to be small [5]. In this case, all differences are are in the 1.7σ range.

Table 5.1: Number of isolated photons in the $Z\gamma\gamma$, $Z\gamma j$, $Zj\gamma$ and Zjj phase space. The numbers are obtained by integrating the corresponding isolation templates from $-\infty$ to 0. All backgrounds, including the pileup contribution, are already subtracted. The results are compared to a previous analysis, which used the matrix method and the template fit to determine the fake photon background of the ATLAS Run-2 $Z\gamma\gamma$ phase space. The number of fake photons, N_{fake}, is obtained by summing up $N_{\gamma j}$, $N_{j\gamma}$ and N_{jj}.

	This Analysis	Template Fit [5]	Matrix Method [5]
$N_{\gamma\gamma}$	220.97 ± 27.83	$239.3 \pm 20.4(\text{stat}) \pm 19.9(\text{sys})$	$225.9 \pm 21.9(\text{stat}) \pm 11.8(\text{sys})$
$N_{\gamma j}$	35.0 ± 17.4	$22.9 \pm 8.0(\text{stat}) \pm 8.4(\text{sys})$	$26.2 \pm 8.5(\text{stat}) \pm 8.3(\text{sys})$
$N_{j\gamma}$	5.3 ± 7.4	$12.5 \pm 6.8(\text{stat}) \pm 5.5(\text{sys})$	$29.0 \pm 9.8(\text{stat}) \pm 7.4(\text{sys})$
N_{jj}	1.0 ± 0.9	$1.4 \pm 1.3(\text{stat}) \pm 0.4(\text{sys})$	$9.0 \pm 3.6(\text{stat}) \pm 2.9(\text{sys})$
N_{fake}	41.3 ± 18.9	$36.8 \pm 9.9(\text{stat}) \pm 11.0(\text{sys})$	$64.2 \pm 12.3(\text{stat}) \pm 11.8(\text{sys})$

Further statistical and systematic studies with toy MC can be performed, which also can include a study of the correlations between the different methods. However, this would be beyond the scope of this bachelor thesis. Moreover, it is not the aim of the presented template fit to repeat an accurate $Z\gamma\gamma$ analysis, but to validate the modified template-extraction methodology. Since the fit results agree with the data (see Figure 5.2) and the determined numbers of real and fake photons agree with the previous analysis within the scope of uncertainties (see Table 5.1), it is concluded that the extracted isolation templates yield good results.

Chapter 6

Transfer of the Template Fit Method to the $Z\gamma$ Process

The $Z\gamma\gamma$ isolation templates can be divided into a generic part which can be used for leading *and* subleading photons in $Z\gamma\gamma$ final states, and a process-dependent part which describes the p_{T} distribution of the respective phase space. In this chapter, the further generalization of the generic part, $T_{\gamma,j}(E_{\mathrm{T}}^{\mathrm{cone20}}, p_{\mathrm{T}}|\vec{\theta}_{\mathrm{opt}})$, to another process is investigated.

This is done exemplarily for $Z\gamma$ final states. The Z boson decays leptonically into a pair of e^+e^- or $\mu^+\mu^-$ and the photon is radiated from an initial quark line, as shown in Figure 6.1. The event selection introduced in Chapter 4.1 is adapted to $Z\gamma$ final states. $Z\gamma$ MC simulations generated by Sherpa 2.2.4 at next-to-leading order QCD precision [16, 17] are used for the comparison of the real photon isolation template, and Z+jets MC simulations generated by MadGraphPythia8 at next-to-leading order QCD precision [17, 22, 23] are used for the comparison of the fake photon isolation template.

After the transformation of $E_{\mathrm{T}}^{\mathrm{cone20}} \rightarrow E_{\mathrm{iso}} = E_{\mathrm{T}}^{\mathrm{cone20}} - 0.065 \cdot p_{\mathrm{T}}$, the $Z\gamma\gamma$ isolation templates of Chapter 4 are multiplied with the corresponding p_{T} distributions $f_{Z\gamma,Zj}(p_{\mathrm{T}})$ and compared to the predicted E_{iso} distributions of $Z\gamma$ and Zj, respectively. Since the extraction of the $Z\gamma\gamma$ fake photon isolation template is limited by statistics, the parameters are allowed to float within the uncertainties of the optimal parameters (see Table 4.3), when comparing the fake photon isolation template with the Z + jets MC simulation.

The results are shown in Figure 6.2. Overall, the comparisons yield good results, which demonstrate that a generalization of the isolation templates is feasible for $Z\gamma\gamma$ and $Z\gamma$ processes by taking into account the p_{T} distribution of the corresponding phase space. To further study the systematic trend which is visible in the pull histogram (lower panel of Figure 6.2), the isolation templates can be extracted in a phase space with higher statistics ($Z\gamma$) and be compared to a phase space with lower statistics ($Z\gamma\gamma$).

The presented results demonstrate the strength of dividing $\mathcal{T}_{\gamma,j}(E_{\mathrm{T}}^{\mathrm{cone20}})$ into a generic

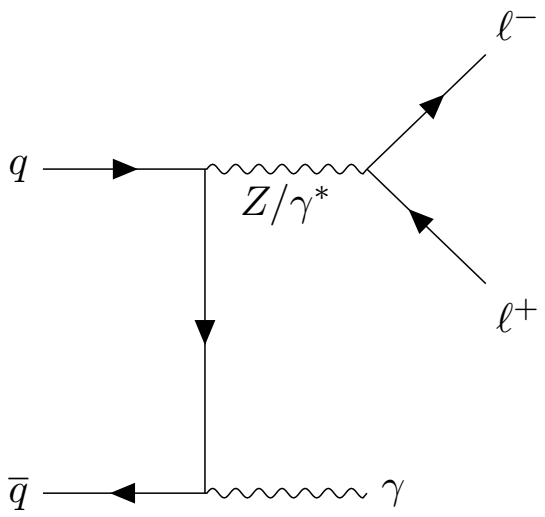
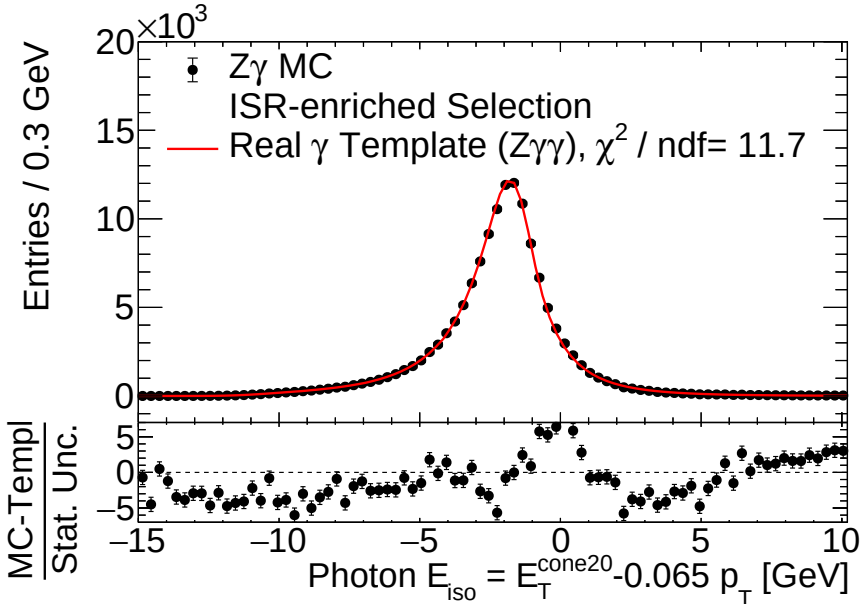


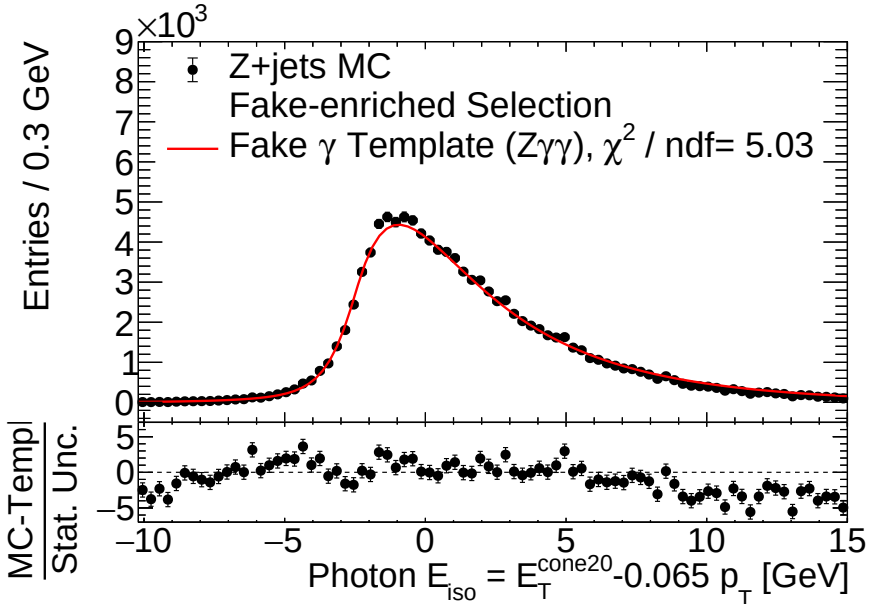
Figure 6.1: Feynman diagram of the studied $Z\gamma$ final state. The Z boson decays leptonically inton a pair of e^+e^- or $\mu^+\mu^-$ and the photon is radiated from an initial quark line.

part, $T_{\gamma,j}(E_{\mathrm{T}}^{\mathrm{cone20}}, p_{\mathrm{T}}|\vec{\theta}_{\mathrm{opt}})$, and a process-dependent part, $f_X(p_{\mathrm{T}})$. This method has the advantage that the optimal parameters need to be extracted only once, and that they can be used for different processes and different regions in the phase space by a multiplication with the corresponding p_{T} distribution. This is especially useful for low statistics processes, in which the fit of the template-extraction process does not yield stable results. Further studies with toy MC can be used to confirm the generalization of the isolation templates for real and fake photons.

It can be qualitatively observed that the isolation templates have a certain dependence on the pileup. This is especially important when using the isolation templates for Run-3 data. To tackle this dependence, either a pileup reweighting procedure has to be developed or a new set of optimal parameters has to be extracted.



(a)



(b)

Figure 6.2: Comparison of the $Z\gamma\gamma$ real and fake photon isolation templates to the isolation energy of real and fake photons in $Z\gamma$ and $Z+$jets MC simulations. The comparison of the real photon isolation is shown in (a) and of the fake photon isolation template in (b). The $Z\gamma\gamma$ real photon isolation template is multiplied with $f_{Z\gamma}(p_{\mathrm{T}})$ and plotted above the E_{iso} distribution of the $Z\gamma$ MC simulation without performing a fit. As the extraction of the $Z\gamma\gamma$ fake photon isolation template is limited by statistics, the uncertainties of the optimal parameters $\vec{\theta}_{\mathrm{opt}}$ are considered by letting the parameters float within the uncertainties of the optimal parameters and performing a fit after multiplying with $f_{Zj}(p_{\mathrm{T}})$.

Chapter 7

Conclusion

The template fit method can be used to estimate the relative contributions of promptly produced photons and fake photons based on their isolation energy distribution.

Templates for the $E_{\mathrm{T}}^{\mathrm{cone20}}$ distribution of real and fake photons are extracted from MC simulations in the $Z\gamma\gamma$ phase space. The photon p_{T} dependence of the isolation templates is studied by one-dimensional interval-wise fits. It is validated that the $E_{\mathrm{T}}^{\mathrm{cone20}}$ distributions of leading and subleading photons in $Z\gamma\gamma$ final states agree for each p_{T} interval. The p_{T} dependence of the parameters of the isolation templates are investigated, which is used to split the isolation template into a generic and a process-dependent part. It is verified that this methodology yields comparable results to the one-dimensional interval-wise fits. It is validated that the generic part can be used for both photons in $Z\gamma\gamma$ final states. The generic part also describes the isolation energy distribution of photons in $Z\gamma$ final states. The process-dependent part models the p_{T} distribution of the corresponding process. The isolation templates are also validated with control regions in ATLAS Run-2 data.

The fake photon contribution to $Z\gamma\gamma$ final states in the full ATLAS Run-2 $Z\gamma\gamma$ dataset is estimated by a template fit. For the fake photon background, $N_{\mathrm{fakes}} = 41.3 \pm 18.9$ is obtained in good agreement to the result of a previous analysis which estimated the fake photon background of the same phase space to be $N_{\mathrm{fakes}} = 36.8 \pm 9.9(\mathrm{stat}) \pm 11.0(\mathrm{sys})$. This validates the modified methodology used for the template extraction.

The $Z\gamma\gamma$ isolation templates are compared to the predicted distribution of the isolation energy of photons in $Z\gamma$ final states. The results demonstrate that a generalization of the isolation templates is feasible for photons in $Z\gamma\gamma$ and $Z\gamma$ final states by taking into account the p_{T} distribution of the studied phase space. The $Z\gamma\gamma$ isolation templates agree well with the isolation energy distribution of photons in $Z\gamma$ final states. Remaining differences are likely to be solved by performing the comparison vice versa, i.e. extracting the isolation templates in a phase space region with higher statistics (e.g. $Z\gamma$) and apply them to a phase space region with lower statistics (e.g. $Z\gamma\gamma$).

Overall, the template comparisons show good agreement, which demonstrates the strength of dividing the isolation templates into a generic part and a process-dependent part. This is especially useful for analyses, where the extraction of the isolation templates is limited by low statistics, as the isolation templates can be used with the same optimal parameters for different processes and phase space regions when taking into account the corresponding photon p_{T} distribution.

Appendix

Appendix A

Shower Shape Variables

Table A.1 summarizes the shower shape observables which are used for the identification of photons and electrons.

Table A.1: Summary of the shower shape observables which are used for the identification of photons and electrons [12].

Category	Description	Name	Usage
Hadronic leakage	Ratio of E_{T} in the first layer of the hadronic calorimeter to E_{T} of the EM cluster (used over the ranges $\lvert\eta\rvert < 0.8$ and $\lvert\eta\rvert > 1.37$)	R_{had_1}	e/γ
	Ratio of E_{T} in the hadronic calorimeter to E_{T} of the EM cluster (used over the range $0.8 < \lvert\eta\rvert < 1.37$)	R_{had}	e/γ
EM third layer	Ratio of the energy in the third layer to the total energy in the EM calorimeter	f_3	e
EM second layer	Ratio of the sum of the energies of the cells contained in a 3×7 $\eta \times \phi$ rectangle (measured in cell units) to the sum of the cell energies in a 7×7 rectangle, both centred around the most energetic cell	R_η	e/γ
	Lateral shower width, $\sqrt{(\Sigma E_i \eta_i^2)/(\Sigma E_i) - ((\Sigma E_i \eta_i)/(\Sigma E_i))^2}$, where E_i is the energy and η_i is the pseudorapidity of cell i and the sum is calculated within a window of 3×5 cells	w_{η_2}	e/γ
	Ratio of the sum of the energies of the cells contained in a 3×3 $\eta \times \phi$ rectangle (measured in cell units) to the sum of the cell energies in a 3×7 rectangle, both centred around the most energetic cell	R_ϕ	e/γ
EM first layer	Total lateral shower width, $\sqrt{(\Sigma E_i (i - i_{\mathrm{max}})^2)/(\Sigma E_i)}$, where i runs over all cells in a window of $\Delta\eta \approx 0.0625$ and i_{max} is the index of the highest-energy cell	$w_{s\,\mathrm{tot}}$	e/γ
	Lateral shower width, $\sqrt{(\Sigma E_i (i - i_{\mathrm{max}})^2)/(\Sigma E_i)}$, where i runs over all cells in a window of 3 cells around the highest-energy cell	$w_{s\,3}$	γ
	Energy fraction outside core of three central cells, within seven cells	f_{side}	γ
	Difference between the energy of the cell associated with the second maximum, and the energy reconstructed in the cell with the smallest value found between the first and second maxima	ΔE_s	γ
	Ratio of the energy difference between the maximum energy deposit and the energy deposit in a secondary maximum in the cluster to the sum of these energies	E_{ratio}	e/γ
	Ratio of the energy measured in the first layer of the electromagnetic calorimeter to the total energy of the EM cluster	f_1	e/γ
Track conditions	Number of hits in the innermost pixel layer	$n_{\mathrm{innermost}}$	e
	Number of hits in the pixel detector	n_{Pixel}	e
	Total number of hits in the pixel and SCT detectors	n_{Si}	e
	Transverse impact parameter relative to the beam-line	d_0	e
	Significance of transverse impact parameter defined as the ratio of d_0 to its uncertainty	$\lvert d_0/\sigma(d_0)\rvert$	e
	Momentum lost by the track between the perigee and the last measurement point divided by the momentum at perigee	$\Delta p/p$	e
	Likelihood probability based on transition radiation in the TRT	eProbabilityHT	e
Track-cluster matching	$\Delta\eta$ between the cluster position in the first layer of the EM calorimeter and the extrapolated track	$\Delta\eta_1$	e
	$\Delta\phi$ between the cluster position in the second layer of the EM calorimeter and the momentum-rescaled track, extrapolated from the perigee, times the charge q	$\Delta\phi_{\mathrm{res}}$	e
	Ratio of the cluster energy to the measured track momentum	E/p	e

Appendix B

Extraction of the Real Photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ Template

The real photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ template is extracted from a $Z\gamma\gamma$ MC simulation. The photon p_{T} dependence of the template is studied. For that reason, the template is extracted for several p_{T} intervals. To increase the available statistics, the $E_{\mathrm{T}}^{\mathrm{cone20}}$ distributions of leading and subleading photons are combined. This is validated in Figure B.1 by comparing the interval-wise $E_{\mathrm{T}}^{\mathrm{cone20}}$ distributions of leading and subleading photons of the $Z\gamma\gamma$ phase space. The results of the interval-wise template extraction is shown in Figure B.2. A generic template, which includes the p_{T} dependences of the parameters, is extracted. To validate this template, it is compared to the initial interval-wise fits in Figure B.3.

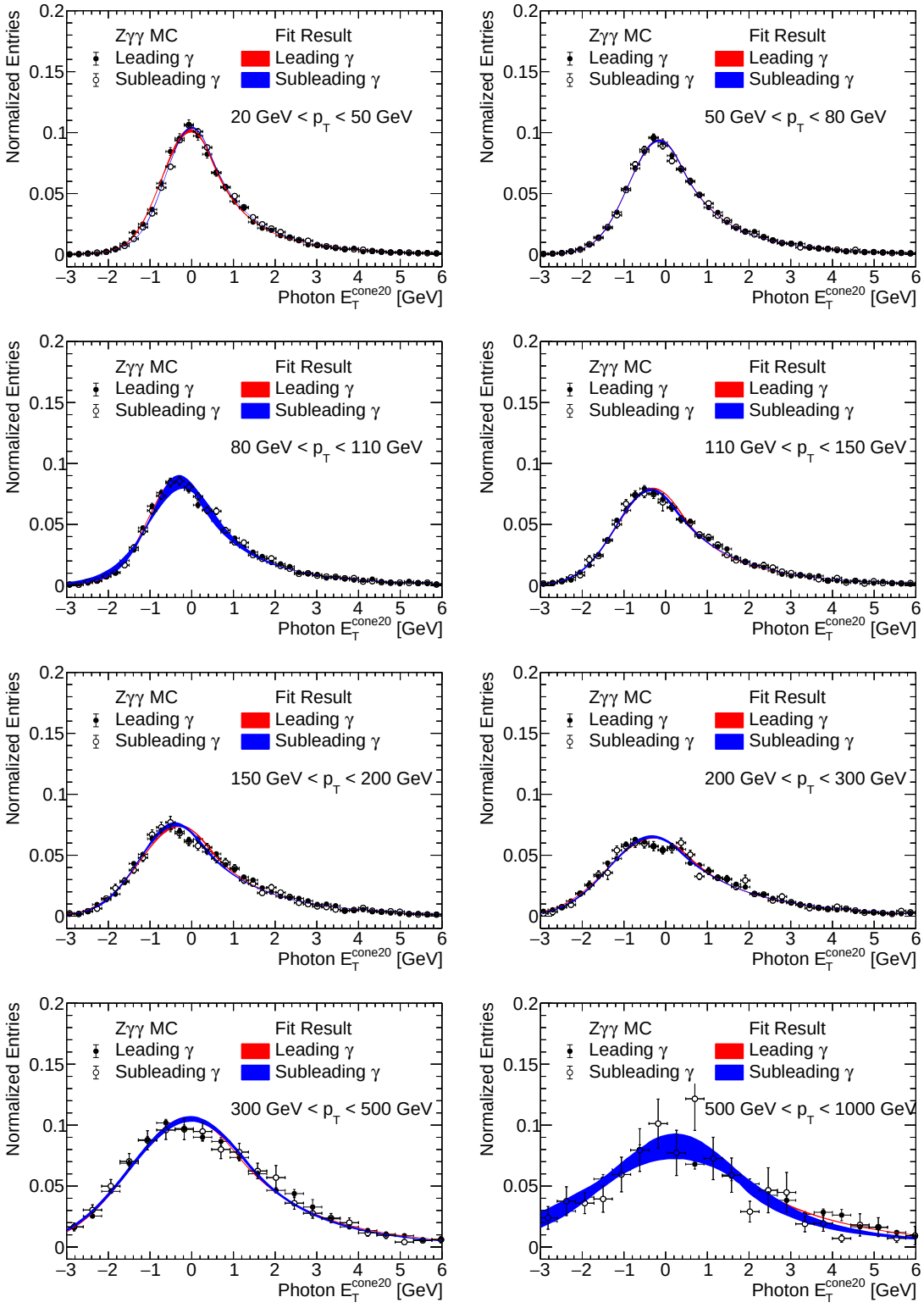


Figure B.1: Comparison of the predicted $E_{\mathrm{T}}^{\mathrm{cone20}}$ distributions of leading and subleading photons of a $Z\gamma\gamma$ MC simulation. The comparison of the bare distributions is shown, as well as the comparison of the fit results which are visualized with their 1σ error bands. Good agreement is observed between the $E_{\mathrm{T}}^{\mathrm{cone20}}$ distributions of leading and subleading photons in $Z\gamma\gamma$ final states.

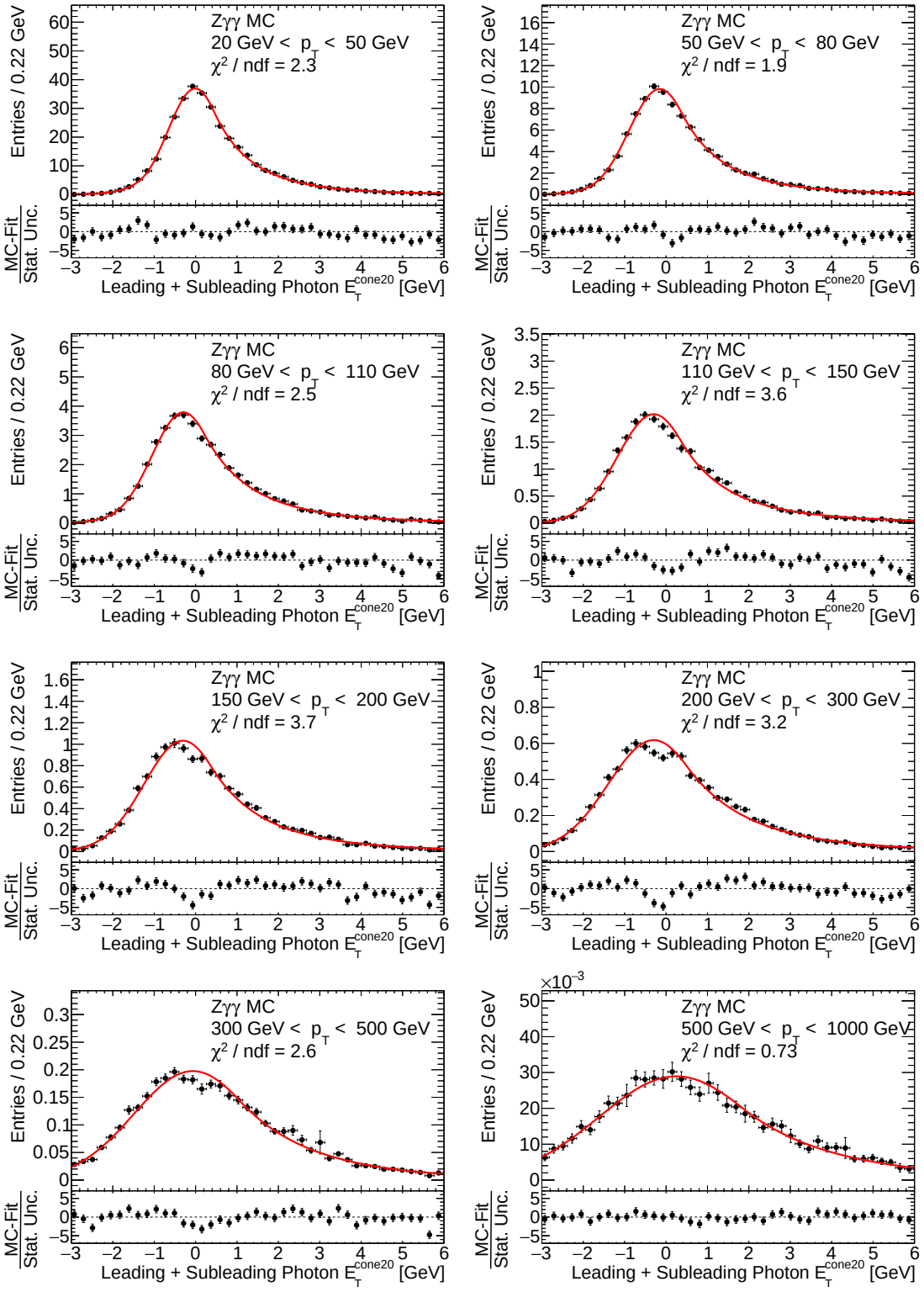


Figure B.2: The one-dimensional interval-wise fits within the combined $E_{\mathrm{T}}^{\mathrm{cone20}}$ distribution of leading and subleading photons of the $Z\gamma\gamma$ phase space are shown. The $E_{\mathrm{T}}^{\mathrm{cone20}}$ distribution is reasonably well described by the DSCB function. In the higher p_{T} intervals, a systematic deviation in the peak region is visible (see also the pull histogram in the lower panel which visualizes the differences between the fit and the predicted $E_{\mathrm{T}}^{\mathrm{cone20}}$ distribution). However, the p_{T} intervals above 150 GeV are not relevant at current data statistics.

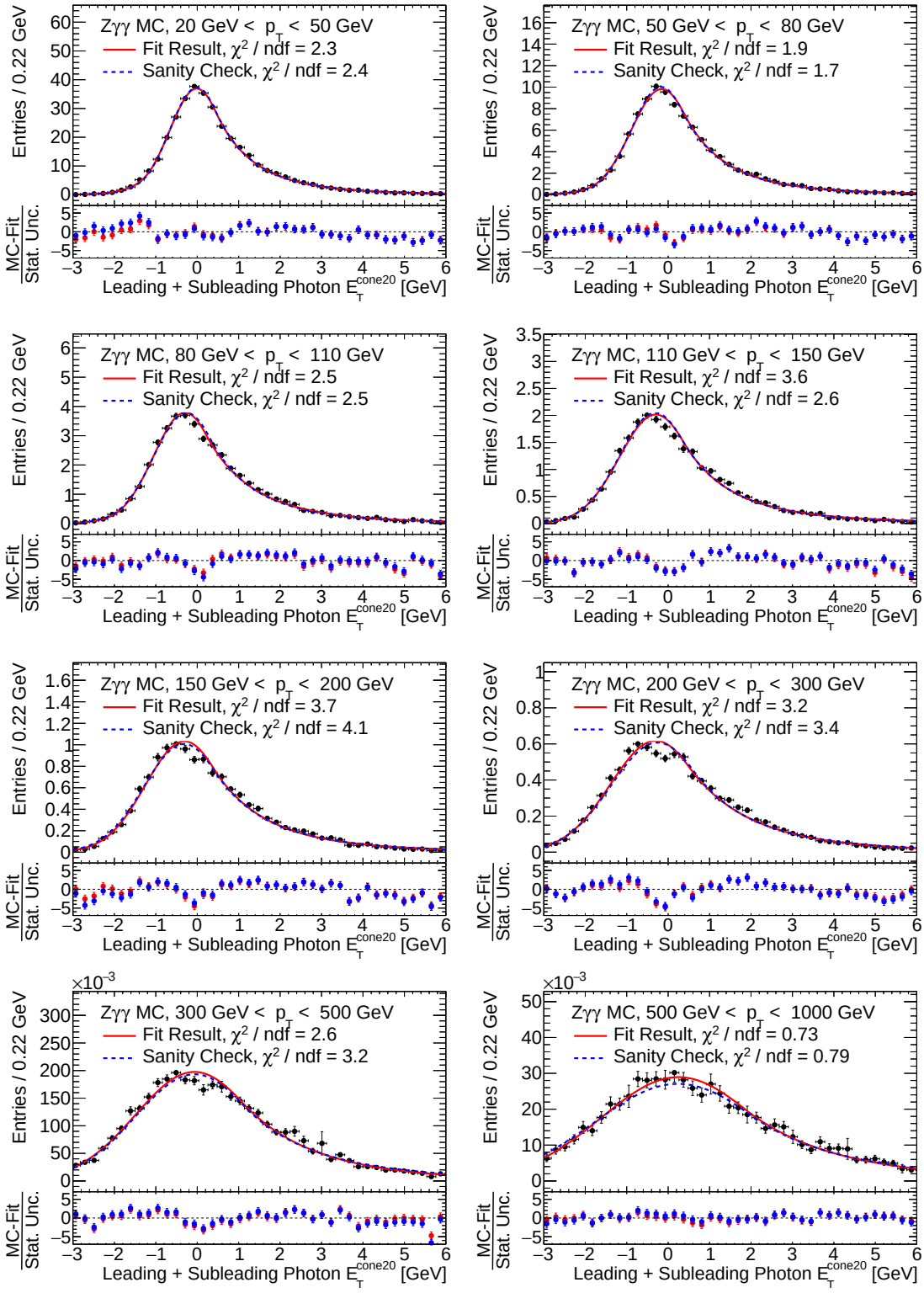


Figure B.3: The results of the one-dimensional interval-wise fits (red solid line) are compared to sanity checks (blue dashed line). This cross check is performed to validate the approximated dependences $\vec{\theta}_0(p_{\mathrm{T}})$ by evaluating $M_\gamma(E_{\mathrm{T}}^{\mathrm{cone20}}|\vec{\theta}_0(p_{\mathrm{T}}))$ at the mean of the interval-wise p_{T} distributions. Good agreement is found between the inital fits and the sanity checks.

Appendix C

Extraction of the Fake Photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ Template

An analogous procedure is performed to extract the fake photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ template and to study the photon p_{T} dependence of this template. In Figure C.1, the predicted $E_{\mathrm{T}}^{\mathrm{cone20}}$ distributions of leading and subleading fake photons in the $Zj\gamma$ and $Z\gamma j$ selections of a $Z\gamma +$ jets MC simulation are compared. In Figure C.2, the results of the one-dimensional interval-wise fits within the combined $E_{\mathrm{T}}^{\mathrm{cone20}}$ distribution of leading and subleading fake photons of the $Zj\gamma$ and $Z\gamma j$ selections are shown. A template which includes the p_{T} dependence of the parameters is extracted. To validate this template, it is compared to the initial one-dimensional interval-wise fits in Figure C.3.

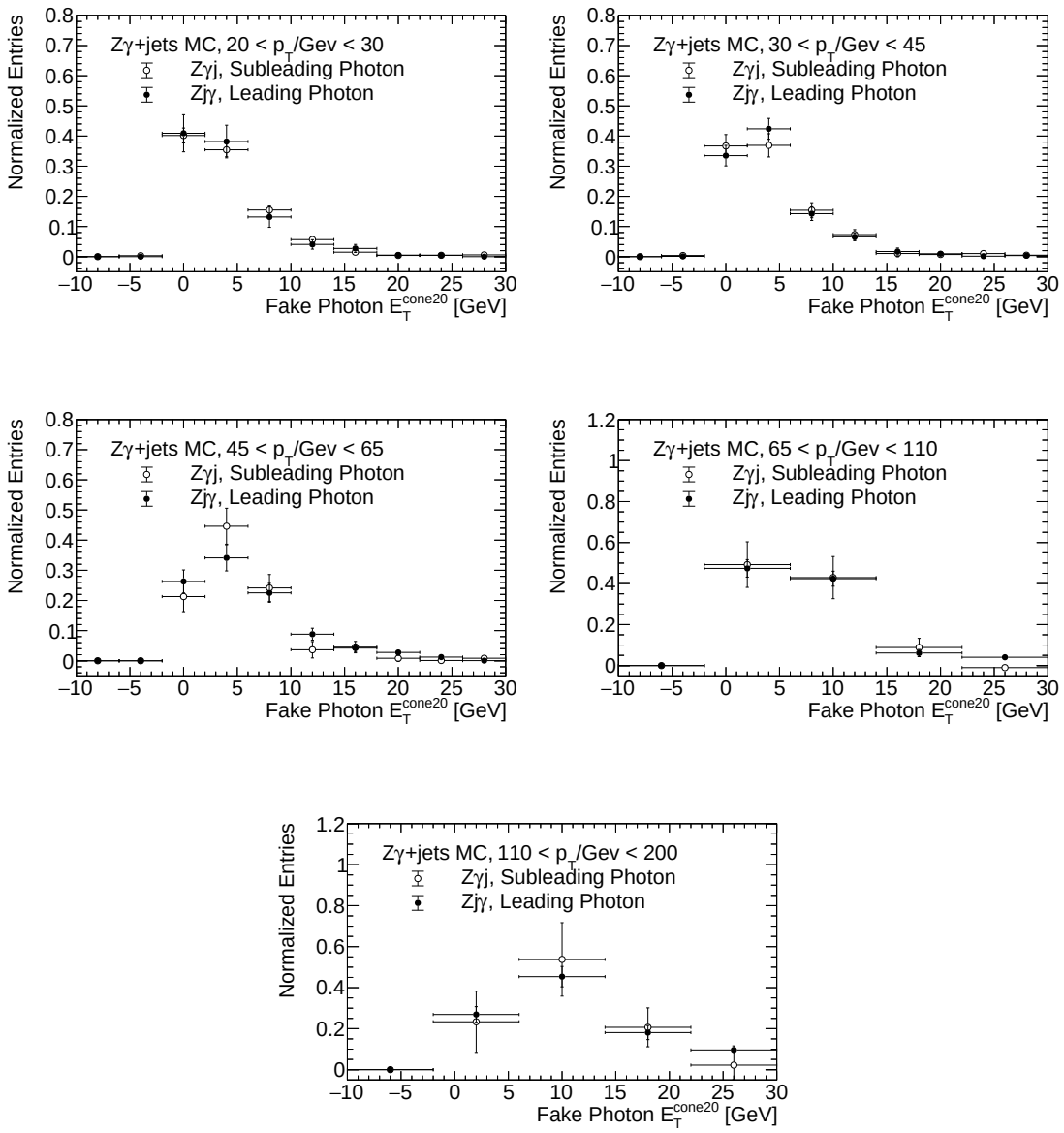


Figure C.1: Comparison of the predicted $E_{\mathrm{T}}^{\mathrm{cone20}}$ distributions of leading and subleading fake photons in the $Zj\gamma$ and $Z\gamma j$ selections of a $Z\gamma$ + jets MC. Good agreement is observed between the $E_{\mathrm{T}}^{\mathrm{cone20}}$ distributions of leading and subleading fake photons.

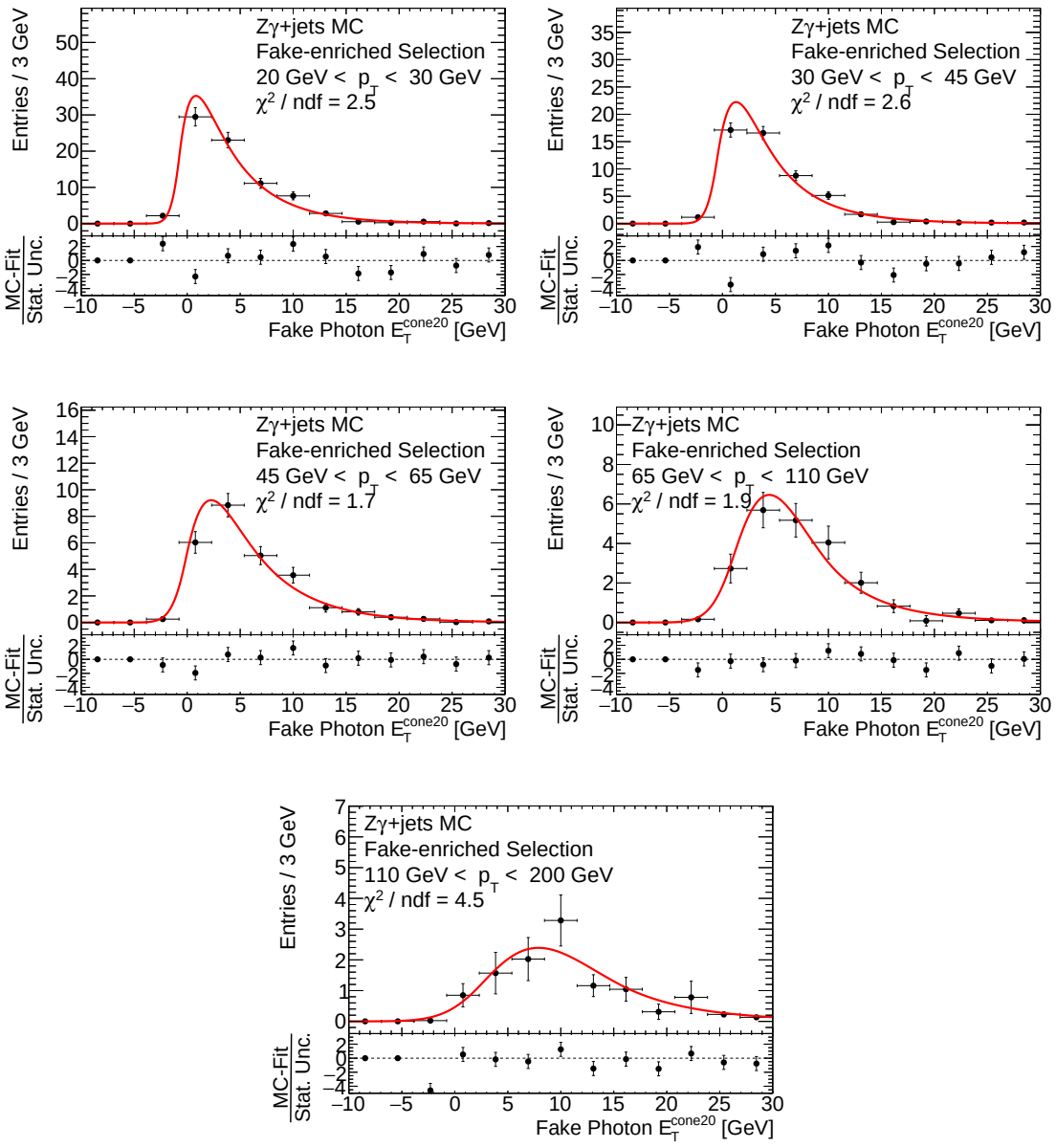


Figure C.2: The one-dimensional interval-wise fits within the combined $E_{\mathrm{T}}^{\mathrm{cone20}}$ distribution of fake photons of the $Zj\gamma$ and $Z\gamma j$ selections are shown. Under consideration of the statistical uncertainties, the $E_{\mathrm{T}}^{\mathrm{cone20}}$ distribution is reasonably described by the Bukin function.

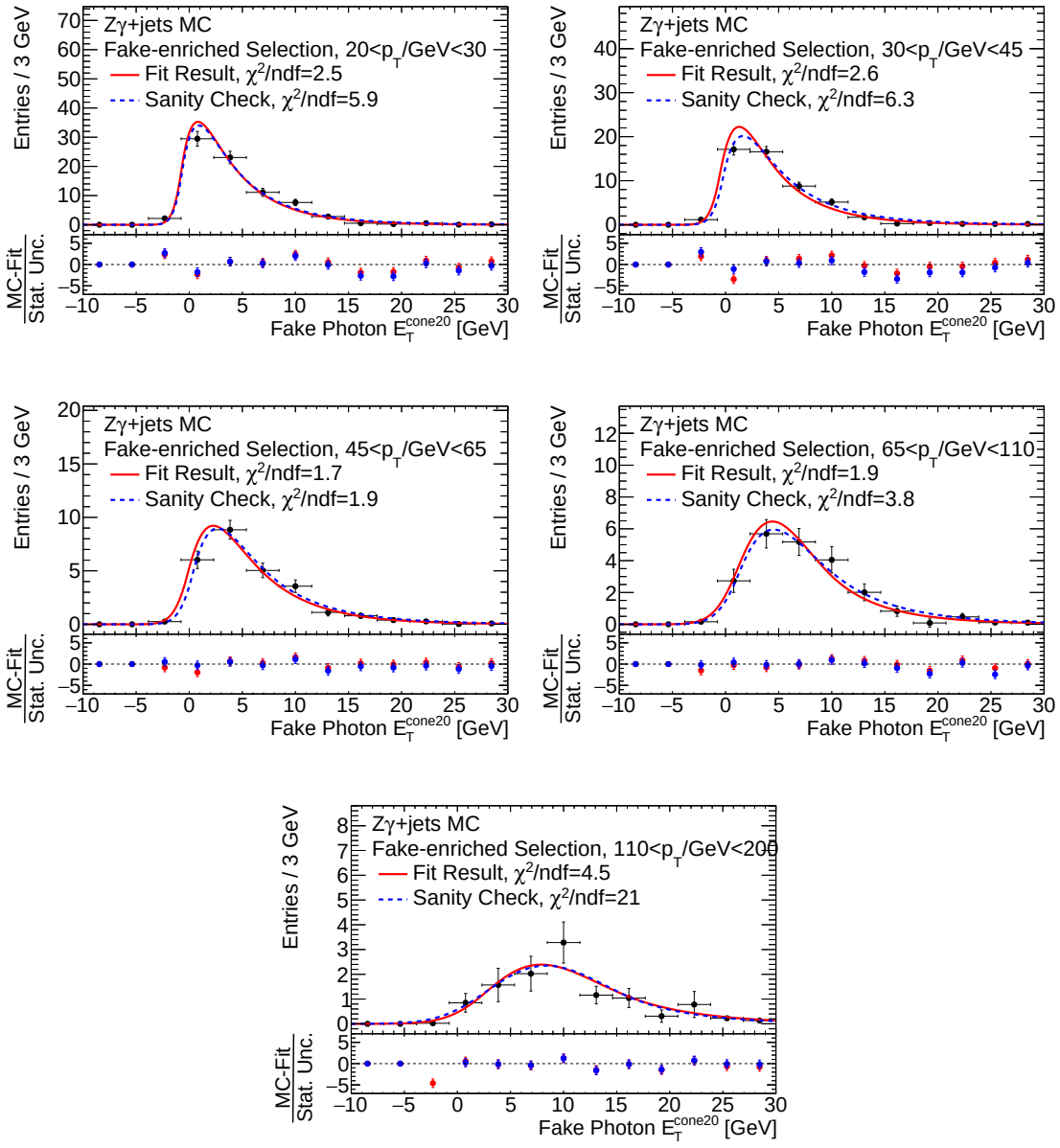


Figure C.3: The one-dimensional interval-wise fits (red solid line) are compared to sanity checks (blue dashed lines). This cross check is performed to validate the approximated dependences $\vec{\theta}_0(p_{\mathrm{T}})$ by evaluating $M_j(E_{\mathrm{T}}^{\mathrm{cone20}}|\vec{\theta}_0(p_{\mathrm{T}}))$ at the mean of the interval-wise p_{T} distributions. The higher χ^2 values in the sanity checks are mainly driven by the small overestimate at the right tail. However, the central part is better described by the sanity check. Therefore, it is concluded that the dependences $\vec{\theta}_0(p_{\mathrm{T}})$) hold in a good approximation.

Appendix D

Extraction of $E_{\mathrm{T}}^{\mathrm{cone20}}$ Templates requiring Track Isolation

The ATLAS recommended loose photon isolation working point [12] requires

$$E_{\mathrm{iso}} := E_{\mathrm{T}}^{\mathrm{cone20}} - 0.065 \cdot p_{\mathrm{T}} < 0, \tag{D.1}$$

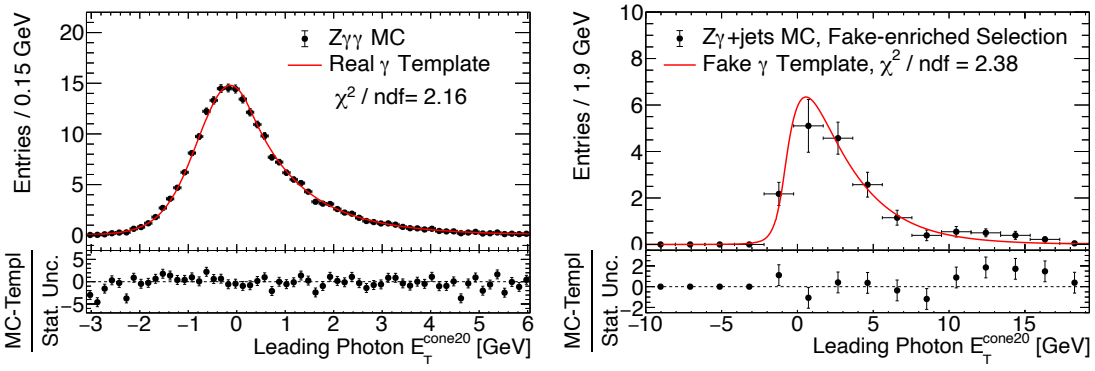
$$p_{\mathrm{iso}} := p_{\mathrm{T}}^{\mathrm{cone20}} - 0.05 \cdot p_{\mathrm{T}} < 0. \tag{D.2}$$

To simplify the fit procedure, the track isolation $p_{\mathrm{iso}} < 0$ can be applied by a hard cut, whereas the calorimetric isolation energy $E_{\mathrm{iso}} < 0$ is taken into account by integrating the templates from $-\infty$ to 0 after performing the template fit.

In sense of the generalization-idea, it is favorable to include also not-track-isolated photons to the fake-photon isolation template, as this avoids biasing the data. Moreover, real and fake photons are less separated when applying the cut on the track isolation.

The templates are extracted twice in this analysis. One time without applying the cut on the track isolation (detailed extraction in Chapter 4) and the second time requiring the cut on the track isolation. The second approach is used in Chapter 5 and the essential plots of the extraction process are shown in Figure D.1.

Despite the shapes of templates are expected to change when applying the cut on the track isolation, the dependences $\vec{\theta}_0(p_{\mathrm{T}})$ remain unchanged. However, the optimal parameters $\vec{\theta}$ change slightly when applying the additional cut, as shown in Table D.1.



(a) (b)

Figure D.1: Extraction of isolation templates requiring loose track isolation. The procedure of the extraction process is exactly the same as presented in Chapter 4, with the only difference of applying the cut on the track isolation before the template extraction. The result for the real photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ template is shown in (a) and for the fake photon $E_{\mathrm{T}}^{\mathrm{cone20}}$ template in (b).

Table D.1: Optimal parameters with and without the cut on the track isolation.

	$\vec{\theta}$	without cut	with cut
Real Photon Template	σ_0	$(57.4 \pm 1.2) \times 10^{-2}\,\mathrm{GeV}$	$(55.8 \pm 3.4) \times 10^{-2}\,\mathrm{GeV}$
	σ_1	$(230.6 \pm 8.4) \times 10^{-5}$	$(223 \pm 29) \times 10^{-5}$
	β_{L}	$(158.9 \pm 3.9) \times 10^{-2}$	$(155 \pm 16) \times 10^{-2}$
	γ_{L}	$(47.9 \pm 3.8) \times 10^{-4}\,\mathrm{GeV}^{-1}$	$(38.2 \pm 8.5) \times 10^{-4}\,\mathrm{GeV}^{-1}$
	β_{R}	$(700.5 \pm 3.8) \times 10^{-3}$	$(609.5 \pm 8.9) \times 10^{-3}$
	γ_{R}	$(35.5 \pm 2.2) \times 10^{-5}\,\mathrm{GeV}^{-1}$	$(41.7 \pm 7.1) \times 10^{-5}\,\mathrm{GeV}^{-1}$
	n_{L}	$(699\,648.0 \pm 3.2) \times 10^{-4}$	$(699\,651.0 \pm 5.2) \times 10^{-4}$
	n_{R}	$(608.5 \pm 2.7) \times 10^{-2}$	$(602.6 \pm 5.9) \times 10^{-2}$
	D	$(7.8 \pm 2.4) \times 10^{-1}\,\mathrm{GeV}$	$(8.6 \pm 2.7) \times 10^{-1}\,\mathrm{GeV}$
	a	$(40.2 \pm 8.4) \times 10^{-4}\,\mathrm{GeV}^{-1}$	$(40.0 \pm 9.0) \times 10^{-4}\,\mathrm{GeV}^{-1}$
	x_0	$(15.4 \pm 1.2) \times 10^{+1}\,\mathrm{GeV}$	$(15.5 \pm 1.2) \times 10^{+1}\,\mathrm{GeV}$
	c	$(32.2 \pm 1.1) \times 10^{-2}\,\mathrm{GeV}$	$(37.4 \pm 1.4) \times 10^{-2}\,\mathrm{GeV}$
Fake Photon Templ.	μ_0	$-(6.0 \pm 1.8) \times 10^{-1}\,\mathrm{GeV}$	$-(7.5 \pm 2.1) \times 10^{-1}\,\mathrm{GeV}$
	μ_1	$(61.4 \pm 4.1) \times 10^{-3}$	$(31.3 \pm 5.4) \times 10^{-3}$
	σ_0	$(171.8 \pm 9.0) \times 10^{-2}\,\mathrm{GeV}$	$(90 \pm 12) \times 10^{-2}\,\mathrm{GeV}$
	σ_1	$(27.3 \pm 2.1) \times 10^{-3}$	$(18.9 \pm 3.5) \times 10^{-3}$
	ξ_0	$(54.0 \pm 2.5) \times 10^{-2}$	$(48.6 \pm 4.7) \times 10^{-2}$
	ξ_1	$-(305 \pm 39) \times 10^{-5}\,\mathrm{GeV}^{-1}$	$-(311.7 \pm 9.2) \times 10^{-5}\,\mathrm{GeV}^{-1}$
	ρ_{L}	$-(9\,999.2 \pm 1.7) \times 10^{-4}$	$-(9\,432.2 \pm 3.6) \times 10^{-4}$
	ρ_{R}	$(2.5 \pm 2.1) \times 10^{-2}$	$(1.4 \pm 0.7) \times 10^{-2}$

List of Abbreviations

ATLAS A Toroidal LHC AppratuS

CERN European Organization for Nuclear Research

CR control region

DSCB double-sided crystal-ball

EM electromagnetic

FSR final-state radiation

ID identification

ISR inital-state radiation

LAr liquid argon

LHC Large Hadron Collider

MC Monte Carlo

pdf probability density function

QCD quantum chromodynamics

SM Standard Model

SR signal region

topo-cluster Topological cluster

Bibliography

[1] M. Thomson, Modern Particle Physics. Cambridge University Press, New York. 2013, ISBN 978-1-107-03426-6, URL http://dx.doi.org/10.1017/CBO9781139525367.

[2] W. de Boer, Grand unified theories and supersymmetry in particle physics and cosmology. Progress in Particle and Nuclear Physics, vol. 33 (1994), p. 201, URL http://dx.doi.org/10.1016/0146-6410(94)90045-0.

[3] R. Bruce, N. Fuster-Martínez, A. Mereghetti, D. Mirarchi and S. Redaelli, Review of LHC Run-2 machine configurations. 9th LHC Operations Evian Workshop, Geneva, Switzerland. 2019. p. 187, URL https://cds.cern.ch/record/2750415.

[4] ATLAS Collaboration, The ATLAS Experiment at the CERN Large Hadron Collider. Journal of Instrumentation, vol. 3 (2008), p. S08003, URL http://dx.doi.org/10.1088/1748-0221/3/08/S08003.

[5] P. Ott, Measurement of $Z\gamma\gamma$ production in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. Ph.D. thesis, Universität Heidelberg. 2023, URL https://www.kip.uni-heidelberg.de/Veroeffentlichungen/download.php/6894/temp/4601.pdf.

[6] ATLAS Collaboration, Measurement of $Z\gamma\gamma$ production in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. The European Physical Journal C, vol. 83(6) (2023), p. 539, URL http://dx.doi.org/10.1140/epjc/s10052-023-11579-8.

[7] ATLAS Collaboration, Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC. 2022, URL http://dx.doi.org/10.48550/arXiv.2212.09379.

[8] M. L. Dunford and P. Jenni, The ATLAS experiment. Scholarpedia. 2014, URL http://dx.doi.org/10.4249/scholarpedia.32147.

[9] J. Pequenao, Computer generated image of the whole ATLAS detector. URL https://cds.cern.ch/record/1095924. 2008.

[10] ATLAS Collaboration, Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run-2 data collected in 2015 and 2016. The European Physical Journal C, vol. 79(3) (2019), p. 205, URL http://dx.doi.org/10.1140/epjc/s10052-019-6650-6.

[11] N. Buatthaisong, Electron/Photon Ambiguity Resolution Using Neural networks For ATLAS Experiment. 2019, URL https://www.desy.de/f/students/2019/reports/nutthawara.buatthaisong.pdf.gz.

[12] ATLAS Collaboration, Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data. Journal of Instrumentation, vol. 14(12) (2019), p. P12006, URL http://dx.doi.org/10.1088/1748-0221/14/12/P12006.

[13] ATLAS Collaboration, Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run-1 data. The European Physical Journal C, vol. 76(12) (2016), p. 666, URL http://dx.doi.org/10.1140/epjc/s10052-016-4507-9.

[14] J. Saxon, Discovery of the higgs boson, measurements of its production, and a search for higgs boson pair production. 2014, URL http://cds.cern.ch/record/1746004.

[15] ATLAS Collaboration, Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at $\sqrt{s} = 13\,\text{TeV}$. The European Physical Journal C, vol. 79(8) (2019), p. 639, URL http://dx.doi.org/10.1140/epjc/s10052-019-7140-6.

[16] E. Bothmann et al., Event generation with Sherpa 2.2. SciPost Physics, vol. 7 (2019), p. 34, URL http://dx.doi.org/10.21468/SciPostPhys.7.3.034.

[17] R. D. Ball et al., Parton distributions for the LHC Run II. Journal of High Energy Physics, vol. 2015(4) (2015), p. 40, URL http://dx.doi.org/10.1007/JHEP04(2015)040.

[18] N. C. Benekos, R. Clifft, M. Elsing and A. Poppleton, ATLAS Inner Detector Performance. Tech. Rep., CERN, Geneva. 2003, URL https://cds.cern.ch/record/688762.

[19] ATLAS Collaboration, Muon reconstruction and identification efficiency in ATLAS using the full Run-2 pp collision data set at $\sqrt{s} = 13\,\text{TeV}$. The European Physical Journal C, vol. 81(7) (2021), p. 578, URL http://dx.doi.org/10.1140/epjc/s10052-021-09233-2.

[20] T. Skwarnicki, A study of the radiative CASCADE transitions between the Upsilon-Prime and Upsilon resonances. Ph.D. thesis, INP Cracow. 1986, URL https://inspirehep.net/literature/230779.

[21] P. M. Morse, Diatomic molecules according to the wave mechanics. II. Vibrational levels. Physical Review, vol. 34 (1929), p. 57, URL http://dx.doi.org/10.1103/PhysRev.34.57.

[22] J. Alwall et al., The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. Journal of High Energy Physics, vol. 2014(7) (2014), p. 79, URL http://dx.doi.org/10.1007/JHEP07(2014)079.

[23] T. Sjöstrand, S. Mrenna and P. Skands, A brief introduction to PYTHIA 8.1. Computer Physics Communications, vol. 178(11) (2008), p. 852, URL http://dx.doi.org/https://doi.org/10.1016/j.cpc.2008.01.036.

[24] A. D. Bukin, Fitting function for asymmetric peaks. 2007, URL http://dx.doi.org/10.48550/arXiv.0711.4449.

[25] R. William, RooBukinPdf.cxx source code. URL https://root.cern.ch/doc/master/RooBukinPdf_8cxx_source.html.

[26] K. Cranmer, Kernel estimation in high-energy physics. Computer Physics Communications, vol. 136(3) (2001), p. 198, URL http://dx.doi.org/https://doi.org/10.1016/S0010-4655(00)00243-5.