University of Zurich[UZH]

Measurement of the $Z\gamma$ Production Cross Section at $\sqrt{s} = 8\,\mathrm{TeV}$ at the LHCb Experiment

Master Thesis

of

Moritz Küng

Mathematisch-naturwissenschaftliche Fakultät

der

Universität Zürich

Prof. Dr. U. Straumann

Dr. K. Müller

Ch. Elsasser

Zürich

2014

A measurement of the $Z\gamma$ production in proton-proton collisions at a centre-of-mass energy of 8 TeV is performed. The data is collected with the LHCb detector at the Large Hadron Collider (LHC), corresponding to an integrated luminosity of 2.06 ± 0.07 fb^{-1}. The Z candidates are reconstructed in the $\mu^+\mu^-$ decay channel. The muons must have transverse momenta, p_T, larger than 20 GeV/c and a pseudo-rapidity, η, between 2 and 4.5. The photon is selected via a multivariate analysis based on shower shape variables of the electromagnetic calorimeter and is required to have a p_T larger than 2 GeV/c and a pseudo-rapidity between 1.8 and 4.3. Between the muons and the photon there must be an isolation of $\Delta R \geq 0.1$.
The analysis is based on template fits with input templates from Monte Carlo simulations. Corrections for the efficiencies are included and the total production cross section is measured to be $\sigma_{pp\to Z\gamma+X} = 3652 \pm 351(\text{stat}) \pm 413(\text{sys})$ fb. A comparison with theoretical predictions of the Standard Model shows a 1.3σ difference.

Acknowledgement

First of all I would like to thank several people, who enabled this thesis and who supported me during the time I worked on this thesis. I thank Prof. Dr. Ulrich Straumann and Dr. Katharina Müller, who made it possible to write my master thesis in the LHCb Group. A special thank goes to my advisors Christian Elsasser and Dr. Katharina Müller. Their excellent support and the offer of their expertise have always been a great help when I had problems and questions. Further I would like to mention the people of our office, who helped me with computer issues and always made time for me: Roman Gredig, Arno Gadola, Nicola Chiapolini, Marco Tresch, Chris Marentini and Christian Elsasser. It has been a great time and I thank you for the support and the great atmosphere. Last but not least I thank Prof. Dr. Massimiliano Grazzini and Dirk Rathlev from the theory department who made the effort to calculate the cross section especially for my thesis.

Contents

1 Theory **11**
- 1.1 Standard Model . . . 11
  - 1.1.1 Forces and Bosons . . . 11
  - 1.1.2 Matter and Fermions . . . 12
  - 1.1.3 Higgs Boson . . . 12
- 1.2 $Z\gamma$ Production . . . 13
  - 1.2.1 Production Mechanisms . . . 13
  - 1.2.2 Theoretical Prediction . . . 14

2 LHCb Detector **17**
- 2.1 VELO and Tracking System . . . 19
- 2.2 Particle Identification . . . 19
  - 2.2.1 Electromagnetic Calorimeter . . . 19
- 2.3 Trigger . . . 20

3 Analysis **21**
- 3.1 Introduction . . . 21
  - 3.1.1 Variables . . . 21
- 3.2 Simulated Samples . . . 22
  - 3.2.1 Classification . . . 22
- 3.3 Data Sample . . . 23
- 3.4 Comparison of Data and Simulation . . . 23
- 3.5 Z Selection . . . 26
- 3.6 Photon Selection . . . 26
  - 3.6.1 Strategy . . . 29
  - 3.6.2 Multivariate Analysis: Boosted Decision Tree . . . 29
  - 3.6.3 Select hard and radiative Signal . . . 30
- 3.7 Muon Reconstruction Efficiency . . . 32
- 3.8 Photon Selection Efficiency . . . 33
  - 3.8.1 newIsPhoton Efficiency . . . 33
  - 3.8.2 BDT Efficiency . . . 34
  - 3.8.3 Photon Conversion Efficiency . . . 37
  - 3.8.4 ϕ Selection Efficiency . . . 38
  - 3.8.5 Total Photon Selection Efficiency . . . 38
- 3.9 Trigger Efficiency . . . 38
  - 3.9.1 Global Event Cut Efficiency . . . 39
- 3.10 Fit Procedure to determine the Signal Yield . . . 41
  - 3.10.1 *Radiative Signal* Fit . . . 42
  - 3.10.2 Toy Fit . . . 44
  - 3.10.3 *Hard Signal* Fit . . . 49

3.11 Systematic Uncertainties 56
3.12 Result 56

4 Conclusion and Outlook **59**
Bibliography 61

Introduction

This thesis reports the measurement of the $Z\gamma$ cross section in pp collisions at the LHCb experiment at CERN. It is organized in the following manner: in the first chapter, a brief review of the Standard Model of particle physics is presented and the $Z\gamma$ production in the forward direction as well as its theoretical predictions are discussed. In the following chapter, an illustration of the LHCb detector is given.
The main part covered in Chapter 3 is the description of the measurement of the cross section, its results and a comparison to the theoretical prediction It includes corrections for the reconstruction, selection, trigger and global event cut efficiency and the template fit that leads to the measured signal yield.
The conclusion at the end is a short summary of the results and a small outlook.

1 Theory

1.1 Standard Model

The Standard Model of particle physics (SM) describes the elementary particles and their interactions. The particles predicted by the SM contain three generations of fermions (half-integral spin) and the gauge bosons (integral spin).

1.1.1 Forces and Bosons

In nature four different interactions between particles are observable:

- gravitation,
- the electromagnetic interaction,
- the weak interaction and
- the strong interaction.

The SM does not include gravitation, as there is no known theory that describes gravitation consistent in the framework of the SM. But the gravitational force is very weak and is negligible on the small scales and masses studied in particle physics.
The carriers of the three remaining forces are the photon, $W^{\pm}$ and Z bosons as well as the gluons (Table 1.1).

Since process $p + p \to Z\gamma + X$ is analysed, the interesting particles are the Z boson and the photon.
While the photon is massless and stable, the Z is a massive vector boson and has a very short life time of $\tau = 1/\Gamma \approx 2.6 \times 10^{-25}\,\mathrm{s}$ [8]. It decays primarily in hadrons and neutrinos. A list of the most important decay modes and branching fractions is given in Table 1.2.

Interaction	**Electromagnetic**	**Weak force**		**Strong force**
Gauge bosons	photon (γ)	$W^{\pm}$	Z	gluons (g; 8 pieces)
charge Q	0	± 1	0	0
mass m	$0\,\mathrm{eV/c^2}$	$80.4\,\mathrm{GeV/c^2}$	$91.2\,\mathrm{GeV/c^2}$	$0\,\mathrm{eV/c^2}$
acting on	all charged particles	all left-handed fermions and $W^{\pm}$, Z		quarks and gluons

Table 1.1: Gauge bosons of the SM [8].

Mode ($Z \to$)	Fraction (Γ_i/Γ)
e^+e^-	(3.363±0.004)%
$\mu^+\mu^-$	(3.366±0.007)%
$\tau^+\tau^-$	(3.370±0.008)%
hadrons	(69.91 ±0.06)%
$\nu_i\overline{\nu_i}$	(20.00 ±0.06)%

Table 1.2: Decay modes of the Z boson with a branching fraction above 1% [8].

	1st generation	**2nd generation**	**3rd generation**
		Quarks	
	up (u)	charm (c)	top (t)
Q	+2/3	+2/3	+2/3
m	2.3 MeV/c^2	1.275 GeV/c^2	173.1 GeV/c^2
	down (d)	strange (s)	bottom or beauty (b)
Q	−1/3	−1/3	−1/3
m	4.8 MeV/c^2	95 MeV/c^2	4.18 GeV/c^2
		Leptons	
	electron (e^-)	muon (μ^-)	tau (τ^-)
Q	−1	−1	−1
m	511 keV/c^2	106 MeV/c^2	1.78 GeV/c^2
	electron-neutrino (ν_e)	muon-neutrino (ν_μ)	tau-neutrino (ν_τ)
Q	0	0	0
m	< 2 eV/c^2 @ 95% C.L.	< 0.19 MeV/c^2 @ 90% C.L.	< 18.2 MeV/c^2 @ 95% C.L.

Table 1.3: Fermions of the SM with charge Q and mass m [8].

1.1.2 Matter and Fermions

Matter and composite particles are built from the 12 fermions which are described by the SM. To each fermion exists an antiparticle having the same qualities, but opposite charge-like quantum numbers. Table 1.3 lists all known fermions f (the antifermions $\overline{f}$ are not listed).

1.1.3 Higgs Boson

The final elementary particle in the SM is the Higgs boson. The interaction between the Higgs boson and the elementary particles gives them mass.
The Higgs boson has a mass of (125.9 ± 0.4) GeV/c^2 [8] and has been discovered by ATLAS and CMS at CERN in 2012 [1], [12].

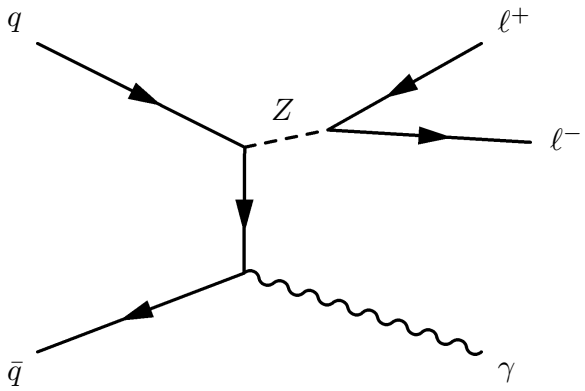

Figure 1.1: Feynman diagram of $q + \overline{q} \to Z(\to \ell^+ + \ell^-) + \gamma$.

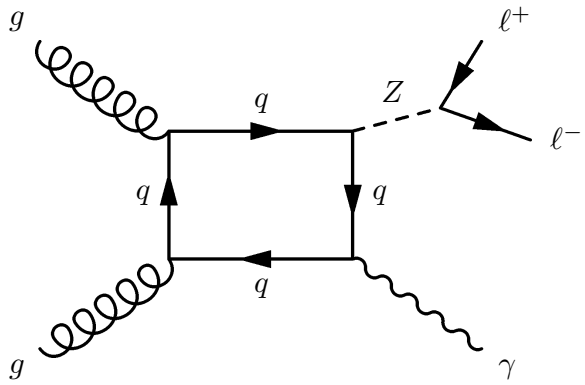

Figure 1.2: Example of a Feynman diagram contributing to $g + g \to Z(\to \ell^+ + \ell^-) + \gamma$ [11].

1.2 $Z\gamma$ Production

1.2.1 Production Mechanisms

In proton-proton collisions, the main production channel of $Z\gamma$ is the annihilation of a quark with an antiquark of the same flavour [11].

$$q + \overline{q} \to Z\gamma \tag{1.1}$$

The leading order Feynman diagram is shown in Fig. 1.1. A higher order contribution arises for example from gluon production (a possible Feynman diagram is shown in Fig. 1.2) [11]:

$$g + g \to Z\gamma \tag{1.2}$$

Another process which leads to a $\ell^+ + \ell^- + \gamma$ final state is Z production with final state radiation or Bremsstrahlung from a lepton of the Z decay:

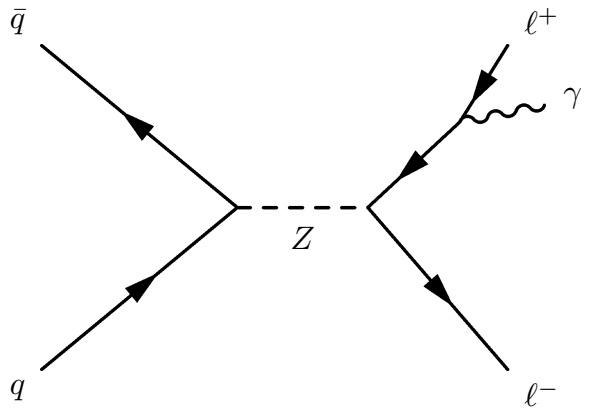

Figure 1.3: Feynman diagram of Z production with Bremsstrahlung radiated from the positive lepton of the Z decay.

$$q + \overline{q} \rightarrow Z \rightarrow \ell^{\pm}(\rightarrow \ell^{\pm} + \gamma) + \ell^{\mp} \tag{1.3}$$

The Feynman diagram for this process is depicted in Fig. 1.3

1.2.2 Theoretical Prediction

A theoretical prediction for $Z\gamma$ production at hadron colliders is computed in various papers (e.g. Ref. [16] which contains up to NNLO calculation), but the cross section for the phase space of LHCb has not been calculated so far. In Ref. [16] Grazzini et al. calculate that the difference between LO and NNLO is about a factor of 2 for the phase space of the ATLAS experiment (this calculation does not distinguish between photons emitted in the hard scatter or from the final state leptons). Grazzini et al. agreed to use the algorithm to calculate the cross section the LHCb phase space and the following cuts:

- $p_{T,\mu^{\pm}} \geq 20\,\mathrm{GeV/c}$
- $\eta_{\mu^{\pm}} \in [2.0, 4.5]$
- $m_{\mu\mu} \in [60\,\mathrm{Gev/c^2}, 120\,\mathrm{Gev/c^2}]$
- $\eta_{\gamma} \in [1.8, 4.3]$
- $p_{T,\gamma} \geq 2\,\mathrm{Gev/c}$
- $\Delta R(\mu) \geq 0.4$ which is an isolation between the photon and the muon (cf. Section 3.1.1).

The result for leading order (LO) is:

$$\sigma_{pp\rightarrow Z\gamma+X} = 2611.2 \pm 1.4 \text{ fb} \tag{1.4}$$

and for next to leading order (NLO):

$$\sigma_{pp\rightarrow Z\gamma+X} = 2806.7 \pm 1.7 \text{ fb} \tag{1.5}$$

The effects originating from higher order correction are much smaller in the LHCb phase space than in the one of ATLAS.

2 LHCb Detector

The LHCb experiment is a dedicated b- and c-physics experiment at the Large Hadron Collider (LHC). The total $b\bar{b}$ cross section at $\sqrt{s} = 7\,\text{TeV}$ has been measured to be $\sim 75\,\mu\text{b}$ inside the acceptance of LHCb and $\sim 300\,\mu\text{b}$ extrapolated to 4π solid angle [2]. Figure 2.1 shows the simulated polar angle distribution of the produced b and $\bar{b}$ quark. Due to the small angles of the b quarks with respect to the proton beams, the detector is designed as a single arm forward spectrometer and covers polar angles from $\sim 15\,\text{mrad} - 250\,\text{mrad}$ ($2.0 < \eta < 5.0$). A side view of the detector is shown in Fig. 2.2. The key features of the LHCb detector are:

- an excellent vertex and proper time resolution;
- a precise particle identification;
- a good momentum resolution and good opening angle determination, which implies precise invariant mass resolution;
- an efficient and flexible trigger system for leptonic and hadronic final states.

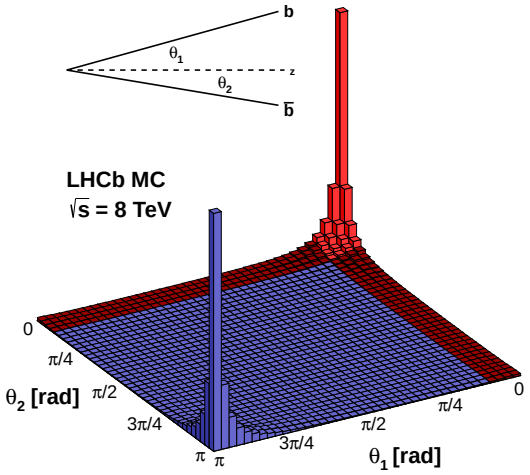

Figure 2.1: Pythia [25] simulation of the $b\bar{b}$ cross section in pp collisions at $\sqrt{s} = 8\,\text{TeV}$ as a function of the polar angle of the produced b and $\bar{b}$. The bright red part corresponds to phase space with both quarks in the LHCb acceptance, the dark red to the one with only one quark.

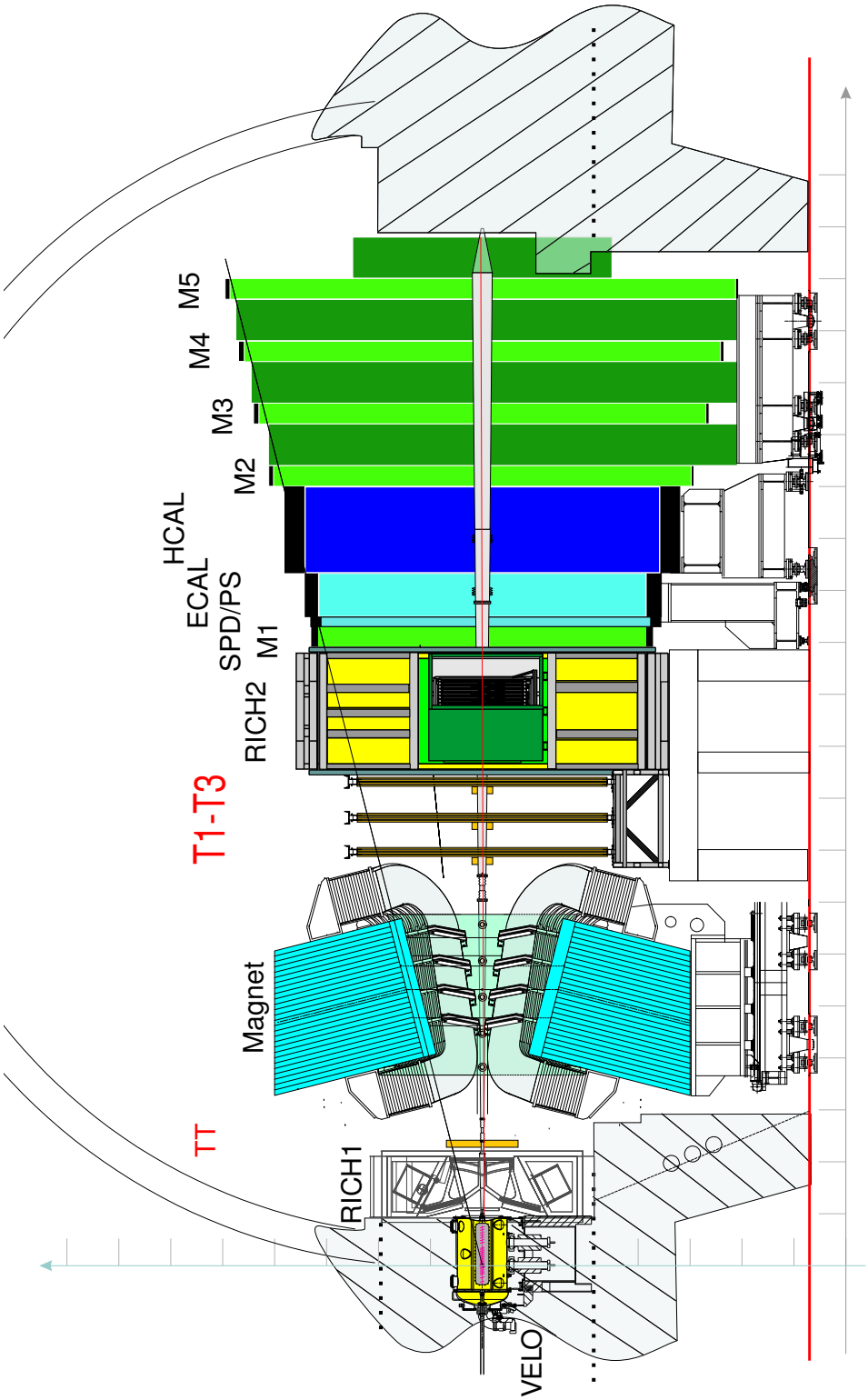

Figure 2.2: Side view of the LHCb detector showing the Vertex Locator (VELO), the dipole magnet, the two RICH detectors, the four tracking stations (TT and T1-T3), the Scintillating Pad Detector (SPD), the PreShower (PS), Electromagnetic (ECAL) and Hadronic (HCAL) calorimeters and the Muon system (M1 - M5)

2.1 VELO and Tracking System

The tracking system consists of a dipole magnet, four tracking stations and the vertex locator (VELO). The VELO is a silicon microstrip detector that is located at a very small distance (down to 8 mm) to the beam axis, providing an excellent vertex resolution. For a typical event, producing 25 tracks, the vertex resolution is 15 µm in the transverse plane and 75 µm along the beam axis.
The first tracking station (TT) is a micro-strip silicon detector. It is located upstream from the dipole magnet. The other tracking stations (T1-T3) are constructed of straw tubes in the outer region, and silicon strips close to the beam pipe. They are located downstream from the magnet. The achieved relative momentum resolution is $\sigma_p/p = 0.4\%$ for $p = 1\,\text{GeV/c}$ and $\sigma_p/p = 0.6\%$ for $p = 100\,\text{GeV/c}$ [18], [6], [7].

2.2 Particle Identification

The two Ring Imaging Cherenkov Detectors (RICH), one upstream from the magnet, one downstream, allow to distinguish between kaons and pions up to a momentum of 100 GeV/c.
The calorimeter system consists of a scintillator pad detector (SPD), pre-shower detector (PS), a lead-scintillator Electromagnetic Calorimeter (ECAL) of shashlik type and an iron-scintillator Hadron Calorimeter (HCAL). Electrons and photons are identified in the SPD/PS and ECAL. The SPD/PS plays an important role in the separation of single photons and merged pions. Due to their finer granularity compared with the ECAL, their information is essential in multivariate classifiers like the BDT described in Section 3.6.2 or newIsPhoton described in Section 3.6.3.
The muon system consists of five stations. They are equipped with Multiwire Proportional Chambers, interlaced with iron absorbers, with the exception of the central part of the first station, which uses triple-GEM (Gaseous Electron Multiplier) detectors.
The calorimeters and the muon system are responsible for the trigger decisions at Level 0 (cf. Section 2.3) [18], [6], [7].

2.2.1 Electromagnetic Calorimeter

For the identification of photons, the calorimeter system (especially SPD/PS and ECAL) plays an important role as photons do not leave tracks in the tracking system or signals in the RICH detectors. Because LHCb is a dedicated c and b physics experiment, the ECAL is not designed for high transverse energies and momenta. The ECAL readout electronics of a single calorimeter cell saturate at $p_{T,\gamma} = 10\,\text{GeV/c}$. For higher momenta, the energy cannot be reconstructed accurately. Figure 2.3 shows the measured versus the true transverse momentum of simulated photons (the simulation *hard photon signal MC* is used, see Section 3.2 for further information). At the threshold of 10 GeV/c measured $p_{T,\gamma}$ the ECAL is saturated. The second step for photons saturating two ECAL cells at $p_{T,\gamma} = 20\,\text{GeV/c}$ is visible, too.

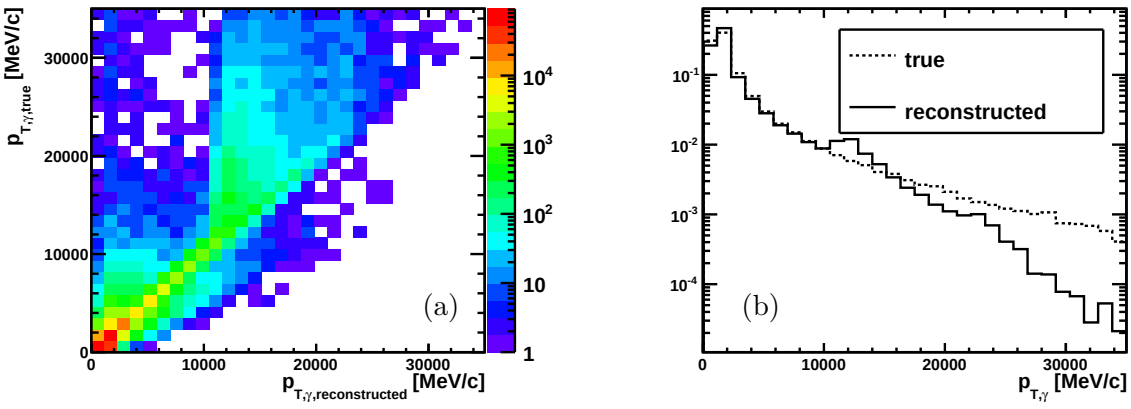

Figure 2.3: (a) Reconstructed versus the true transverse momentum of *hard photon signal MC* and (b) the distributions of true and reconstructed $p_{T,\gamma}$. The saturation of the ECAL at $p_{T,\gamma} = 10\,\mathrm{GeV/c}$ and $p_{T,\gamma} = 20\,\mathrm{GeV/c}$ for one and two calorimeter cells, respectively, is clearly visible.

2.3 Trigger

The design luminosity of the LHCb is $2 \times 10^{32}\,\mathrm{cm^{-2}s^{-1}}$. At this luminosity, the rate of events with at least two particles in the LHCb acceptance is $\sim 10\,\mathrm{MHz}$. The LHCb trigger reduces this rate by a factor 2000 to 5 kHz, which is the rate the events can be written to storage. Hence the trigger has to be highly selective and efficient. It consists of two levels: Level 0 (L0) and High Level Trigger (HLT). L0 is a hardware trigger, implemented on electronic boards. Its decision is based on calorimeter and muon chamber information and it selects muons, electrons, photons or hadrons above a given p_T or transverse energy E_T. It reduces the rate to about 1 MHz.
The HLT is a software algorithm, which is separated in two stages: HLT1 uses a partial reconstruction of the event and selects good-quality tracks. The reduction factor is about 30. HLT2 uses the fully reconstructed tracks to reduce the rate finally to 5 kHz [18], [6], [7].

3 Analysis

3.1 Introduction

The analysis of the data taken at the LHCb experiment was performed with the ROOT framework [10]. It is divided into the following steps:

1. Classification of the simulated events and searching for variables that discriminate between signal and background.
2. Selection of the photon and Z boson.
3. Calculation of the different efficiencies.
4. Fixing the fraction of final state radiation and Bremsstrahlung with a fit.
5. Estimation of the performance of the data fit with a toy model.
6. Fit to data with fixed radiative part and correction with the response matrix from the toy model.
7. Calculation of the production cross section.

3.1.1 Variables

Different kinematical and topological variables are used in this analysis. Their definition is given in the following list:

- p_T[MeV/c]: The transverse momentum of a particle in MeV/c.
- $\eta = -\ln(\tan\frac{\vartheta}{2})$: The pseudo-rapidity where ϑ is the polar angle between the beam axis and the measured momentum vector of the particle.
- ϕ: The azimuthal angle of the measured momentum vector of a particle.
- $\Delta R = \sqrt{(\phi_1 - \phi_2)^2 + (\eta_1 - \eta_2)^2}$: The distance in "solid angle" between two tracks from the same primary vertex.
- ΔR(all tracks): ΔR of the photon and the closest track from the same primary vertex.
- $\Delta R(\mu)$: ΔR of the photon and the closest muon track of the Z decay from the same primary vertex.
- $m_{\mu\mu}$[MeV/c^2]: The invariant mass of the two muons of the Z decay.
- $m_{\mu\mu\gamma}$[MeV/c^2]: The invariant mass of the two muons of the Z decay and the photon.

3.2 Simulated Samples

Two different Monte Carlo (MC) simulations are used for the analysis. The first is a $pp \to Z$+jet simulation, which describes the main part of the expected background (*background MC*). It contains $2 \times 4\,500\,000$ events. The reason for the factor 2 is that the events are simulated for the two magnet polarities separately.
The second MC sample is a $pp \to Z\gamma + X$ simulation representing the signal with $2 \times 525\,000$ events (*hard photon signal MC*).
The simulation of the final state radiation and Bremsstrahlung is obtained from events of the *hard photon signal MC* if a photon originating from a muon was found. If there are multiple photons in a single event, a candidate is created for each photon; e.g. an event with a photon from the hard scattering process and Bremsstrahlung photon from one muon results in a *hard signal* candidate, with the hard scattering photon and in a *radiative signal* candidate with the Bremsstrahlung photon.
The Monte Carlo samples are generated with Pythia6 [25] as event generator and the decays are simulated with EvtGen [19]. PHOTOS [15] is used to describe the final state radiation simulation. With GEANT4 [5] the LHCb detector is being simulated.

3.2.1 Classification

In the simulations four different classes for the photon candidates are defined:

- *hard signal* (usually displayed in black): The photon has to be truth matched (it is a real γ), the Z and the photon come from the same vertex and origin from the hard process (primary vertex).
- *radiative signal* (usually displayed in green): Final state radiation (FSR) and Bremsstrahlung, which are truth matched photons with a muon of the Z decay as mother particle.
- *background photons* (usually displayed in red): This sample contains the remaining truth matched photons which are no *hard signal* or *radiative signal*. Usually originating from neutral meson decays or Bremsstrahlung from other particles than the muons from the Z decay.
- *charged background* (usually displayed in blue): All candidates containing a detected misidentified photon (photon candidate truth matched to a non-photon particle).

The first two classes are basically the signal, the other two are background. Figure 3.1 shows the distributions of important variables for the different classes only applying the Z selection.
The $p_{T,\gamma}$ distribution shows that the signal classes have higher momenta than the background. The variable $\Delta R(\mu)$ has a good separation power for the *radiative signal* to all other classes (which have similar distributions in this variable). That is because the photons of the *radiative signal* origin from the muons of the Z decay and therefore should be close to the muon track. Also the invariant mass of the two muons ($m_{\mu\mu}$) has some separation power, as for the *radiative signal* a part of the energy of the two muons is radiated with the photon and therefore $m_{\mu\mu}$ tends to be smaller than the actual Z mass whereas the classes *hard signal* and *background photons* reproduce the Z resonance. Similar considerations apply for $m_{\mu\mu\gamma}$: The *radiative signal* peaks at the Z mass, the *hard signal* has higher invariant masses than the other classes due to the high-energetic photons and the background is in the middle.

ΔR(all tracks) separates between *background photons* and the other classes. The *background photons* are usually part of a jet and therefore have other particles nearby which is the reason for the small values of ΔR(all tracks).

3.3 Data Sample

A data sample recorded by LHCb in the data taking period 2012 with an integrated luminosity of 2.06 ± 0.07 fb^{-1} at a centre-of-mass energy of $\sqrt{s} = 8$ TeV is used for the analysis.
The following trigger lines were used:

Level	Trigger	Meaning
L0	L0Muon	SPD-hits ≤ 600 $p_{T,\mu} \geq 1.5\,\text{GeV/c}$
Hlt1	Hlt1SingleMuonHighPT	$p_{T,\mu} \geq 4.8\,\text{GeV/c}$ IsMuon [1]
Hlt2	Hlt2SingleMuonHighPT	$p_{T,\mu} \geq 10.0\,\text{GeV/c}$

Table 3.1: Triggers used for the data acquisition and for the calculation of the trigger efficiency.

It is required that at least one of the muons has fired all three levels (i.e. is TOS for all three levels, Trigger On Signal).
In the same way as in the simulations, if there are multiple photons in a single event, a candidate is created for each photon. This procedure could increase the cross section if there are a lot of events with multiple photons. Studies on simulation have shown, that less than 1% of the events passing the selection criteria have multiple photons. Therefore this effect is neglected.

3.4 Comparison of Data and Simulation

To assure that the simulation describes the data correctly, different distributions of variables between the simulations and data are compared. Because the simulations describe only a part of the data, cuts are applied in order to reduce the data to one specific class. Using the cut $\Delta R(\mu) \leq 0.2$ it is possible to isolate the *radiative signal* events (cf. Fig. 3.1(d)). The comparison of data with the applied cut $\Delta R(\mu) \leq 0.2$ and the class *radiative signal* is shown in Fig. 3.2.
Data and simulation are in good agreement for the important variables of the fit like $m_{\mu\mu\gamma}$ (Fig. 3.2(f)) and $m_{\mu\mu}$ (Fig. 3.2(e)) as well as ΔR(all tracks) (Fig. 3.2(c)). Small differences in the crucial variables for the photon selection like the BDT output and newIsPhoton may origin from the pollution of the data with *hard signal* and *background photons* and the input variables of the BDT are very sensitive to small uncertainties in simulation. As the efficiency for newIsPhoton is estimated from data in Section 3.8.1, these differences do not affect the result.

The data can be reduced to contain mainly *background photons* when selecting low energetic photons only. The cut $1\,\text{GeV/c} \leq p_{T,\gamma} \leq 1.5\,\text{GeV/c}$ is applied to the data and the distributions

[1] A particle is called "IsMuon", if there is a certain number of muon station hits in a Field of Interest defined by the track extrapolation. The number of required hits is momentum dependent.

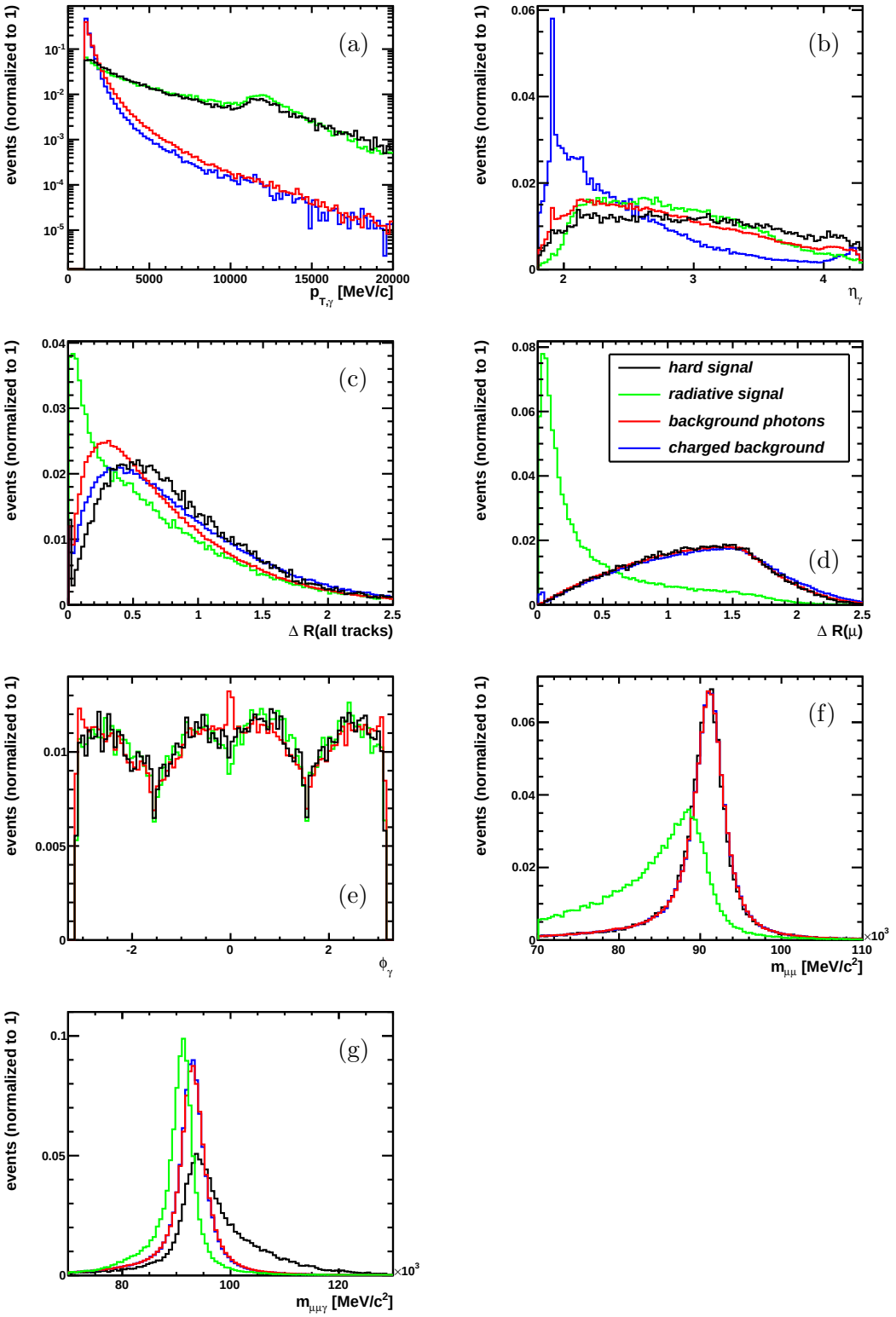

Figure 3.1: Normalized distributions of important variables for the different classes of the simulations only applying the Z selection.

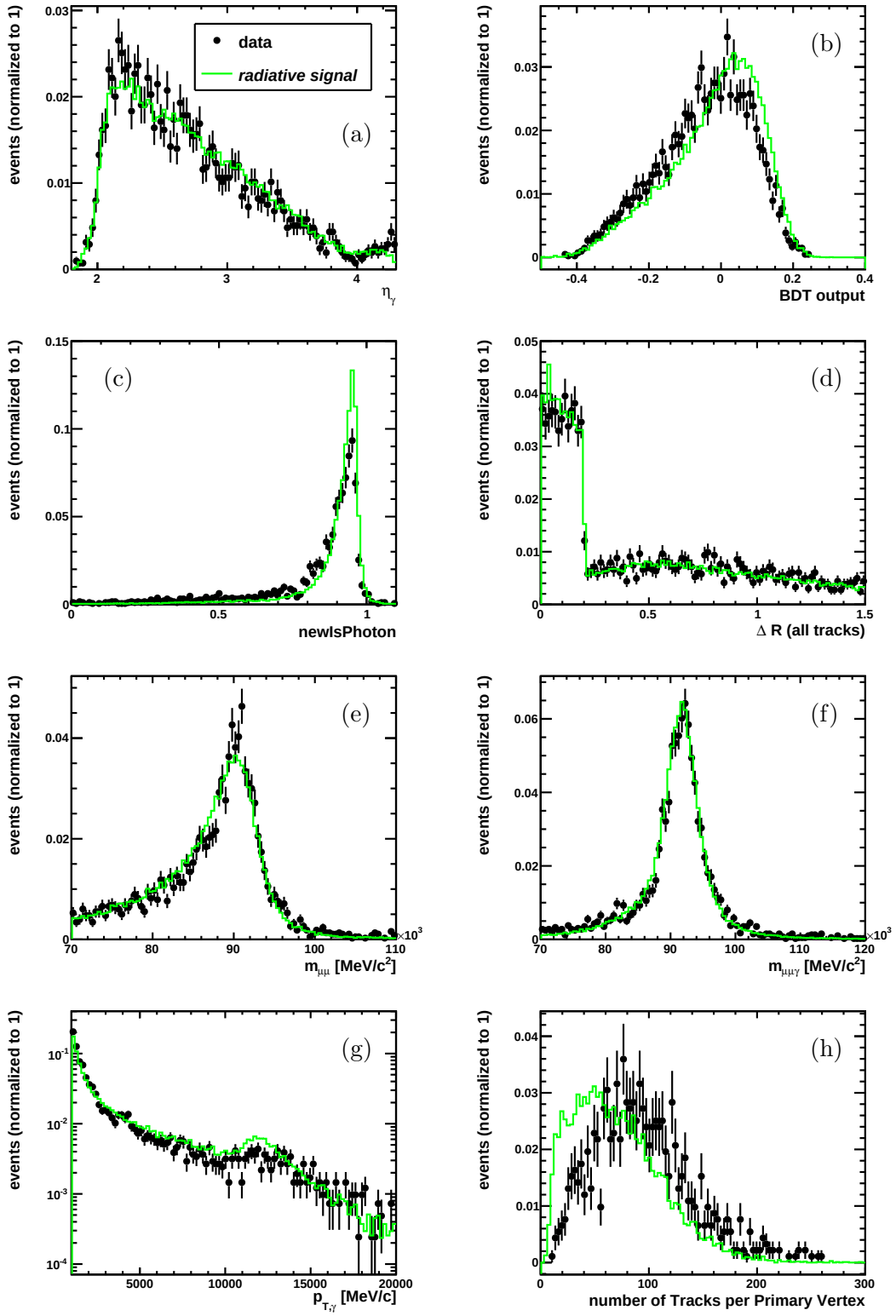

Figure 3.2: Comparison between data and simulation for the enhanced *radiative signal*.

of important variables are compared to the *background photons* from simulation (cf. Fig. 3.3) . The relevant variables for the fit are also very well reproduced by the simulations. $\Delta R(\mu)$ (Fig. 3.3(c)), ΔR(all tracks) (Fig. 3.3(d)), $m_{\mu\mu}$ (Fig. 3.3(e)) and $m_{\mu\mu\gamma}$ (Fig. 3.3(f)) show nearly perfect agreement. As for the *radiative signal* the BDT output shows differences between data and simulation originating from the same reasons. The variable newIsPhoton cannot be plotted as it is only defined for $p_{T,\gamma} \geq 2\,\mathrm{GeV/c}$.
For the *radiative signal* as well as for *background photons* differences are observed between data and simulation. As the BDT is crucial for the photon selection and its efficiency, these differences are taken into account in Section 3.8.2 with a correction factor.

3.5 Z Selection

The Z boson is selected via the decay muons. The positive and negative muons need to fulfill the following criteria:

- $p_{T,\mu^\pm} \geq 20\,\mathrm{GeV/c}$
- $\eta_{\mu^\pm} \in [2.0, 4.5]$
- $m_{\mu\mu} \in [60\,\mathrm{GeV/c^2}, 120\,\mathrm{GeV/c^2}]$
- $Prob(\chi^2_{\mathrm{tr}}/\mathrm{ndf}) \geq 0.01$ (χ^2 probability of the track fit)
- $\frac{\sigma_{p_{\mu^\pm}}}{p_{\mu^\pm}} \leq 0.1$ (the relative uncertainty on the momentum of the muons has to be smaller than 10%)

As large differences have been observed comparing data and simulation of the invariant masses in the Z resonance region, the invariant mass resolution is matched in MC by convoluting the distribution with a Gaussian with mean equal to zero and width $\sigma = 1104\,\mathrm{MeV/c^2}$ [14]. These effects origin from imperfect description of the detector in the simulations and a dependence of the invariant masses from the azimuthal angle ϕ [2].
The effect is plotted in Fig. 3.4 for the invariant mass of the muons of an inclusive $Z \to \mu\mu$ sample and the class *background photons*. At low masses the data is polluted with background and non resonant signal like contribution from Drell Yan $\gamma^* \to \mu\mu$. Due to this fact, the data is scaled with a factor of 1.16. This factor is obtained by a binned maximum likelihood fit to data with RooFit [26]. The Z signal is modeled with a Breit-Wigner distribution, the background and Drell Yan is described by an exponential.

3.6 Photon Selection

The photons need to pass the following cuts. The last two are going to be discussed in detail in Section 3.6.2 and Section 3.6.3.

- $\eta_\gamma \in [1.8, 4.3]$
- $p_{T,\gamma} \geq 2\,\mathrm{GeV/c}$
- $|\phi_\gamma| \in [0.1, 1.47] \bigcup [1.67, 3.04]$

[2] for more information compare Ref. [9] in the section "momentum scale correction"

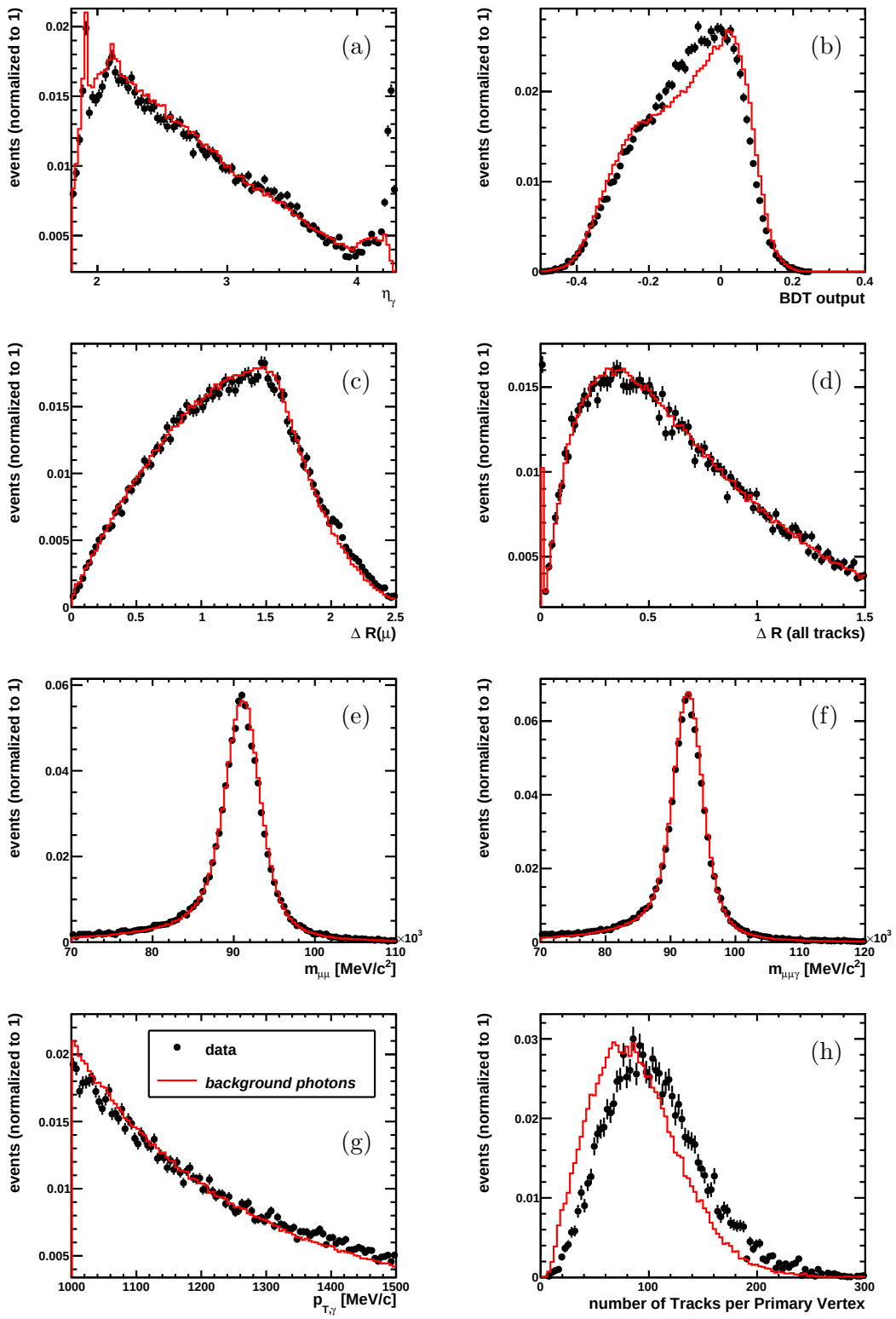

Figure 3.3: Comparison between data and simulation for a sample enhanced *background photons* with $1\,\mathrm{GeV/c} \leq p_{T,\gamma} \leq 1.5\,\mathrm{GeV/c}$.

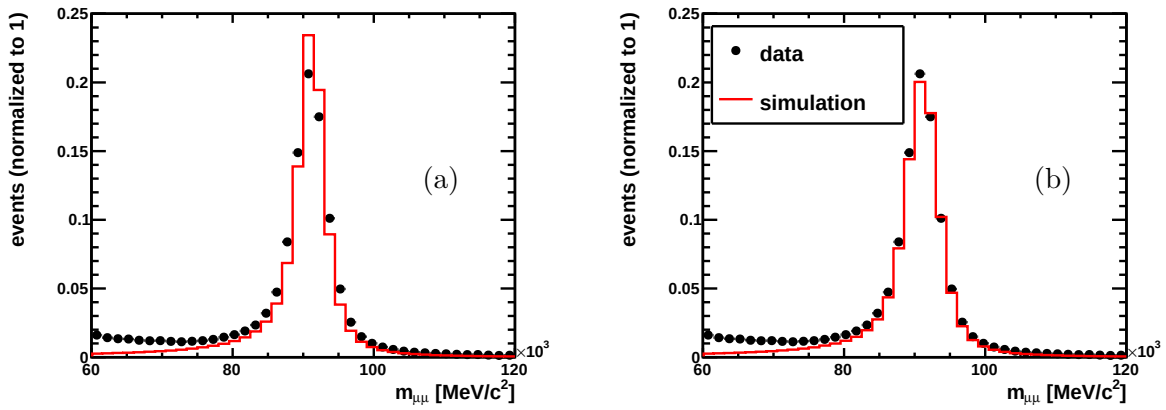

Figure 3.4: Comparison of data and simulation for the invariant masses of the muons. (a) without smearing (b) with smearing of the simulation. The simulation is normalized to 1. The data is scaled to 1.16.

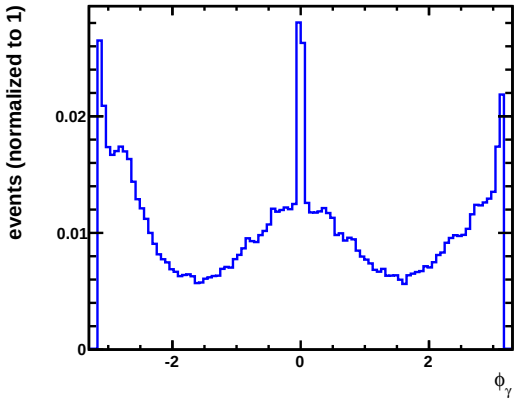

Figure 3.5: ϕ_γ distribution of the *charged background* with peaking mis-identification rate at $\phi_\gamma = 0, \pm\pi/2$ and $\pm\pi$.

- $\Delta R(\mu) \geq 0.1$
- BDT variable ≥ 0.05
- newIsPhoton ≥ 0.8

As the electromagnetic calorimeter only fully covers pseudo-rapidities between 1.8 and 4.3, photons in this range are selected. The cut of $p_{T,\gamma} \geq 2\,\mathrm{GeV/c}$ is applied, because *background photons* are usually low energetic (cf. Fig. 3.1(a)) and $\sim 85\%$ of the background is eliminated with this cut. The cut on ϕ is due to the high mis-identification rate at $\phi = 0, \pm\pi/2$ and $\pm\pi$ (cf. Fig. 3.5 and Fig. 3.1(e)). The reasons are detector effects at these azimuthal angles. The ϕ distribution should be flat and the regions where it is not are cut out.
After the Z and γ selection, 5331 ± 73 candidates remain and enter the fit procedure. The error is calculated as the Poisson uncertainty.

variable	explanation
CaloTrMatch	CaloID estimator : 2D chi2 for Track/CaloCluster matching (neutral + charged)
ShowerShape	CaloID estimator : 2nd order moment of the cluster (neutral)
ClusterMass	CaloID estimator : MergedPi0Alg cluster Mass (neutral)
CaloNeutralSpd	Spd Digits (0,1) associated to CaloHypo (neutral)
CaloNeutralPrs	Prs Digits (0,sum) associated to CaloHypo (neutral)
CaloNeutralEcal	Cluster energy associated to CaloHypo (neutral)
CaloNeutralHcal2Ecal	Hcal/Ecal energy ratio associated to CaloHypo (neutral)
CaloNeutralE49	2x2/3x3 energy ratio associated to CaloHypo (neutral)
PhotonID	Combined PDF for photonID (neutral)

Table 3.2: Explanation of the input variables of the BDT.

3.6.1 Strategy

1. Step: Select real photons from photon candidates. This means eliminating the *charged background*.
2. Step: Select the *hard signal* and the *radiative signal* and reduce the class *background photons*.

3.6.2 Multivariate Analysis: Boosted Decision Tree

For a proper analysis, the photons need to be selected carefully. A photon candidate is reconstructed from a cluster in the ECAL without any track associated. This is a very loose selection and in the *hard photon signal MC* only 74% of the selected photon candidates are real photons. To distinguish between fake and real photons, a multivariate classifier based on a boosted decision tree (BDT) of the TMVA framework (Toolkit for Multivariate Data Analysis) [17] is used. The BDT is optimized to reject charged particles. The input variables (Fig. 3.6) are shower shape variables and the matching of the cluster to tracks (explanations in Table 3.2) and the BDT is trained and tested on the simulations. The signal is the *hard signal* class, the background the *charged background* class which consists of charged hadrons and electrons.

In the variable CaloTrMatch the background is more often associated to a track, which is obvious because the background comes from charged particles. The ShowerShape is wider for the background because its origin are mainly hadronic particles. The variable CaloNeutralHcal2Ecal is the ratio of the energy deposit from the HCAL to the energy deposit in the ECAL. For the background, this variable is larger because the background are hadronic particles. In the variable CaloNeutralEcal the energy which is associated to the calorimeter hypothesis is stored. As the signal photons have higher momenta and energy, it is larger for the signal than background. CaloNeutralE49 is the ratio of the energy between 2×2 cells and 3×3 cells. Because the charged hadrons are often in a jet there is a higher energy deposit in the cells around the photon candidate and for the background this ratio tends to be smaller.
All these variables make use of the wider shower shape of hadrons, the isolation of photons, the track matching and the differences between the response in the SPD, PS, ECAL and HCAL for hadrons and photons.

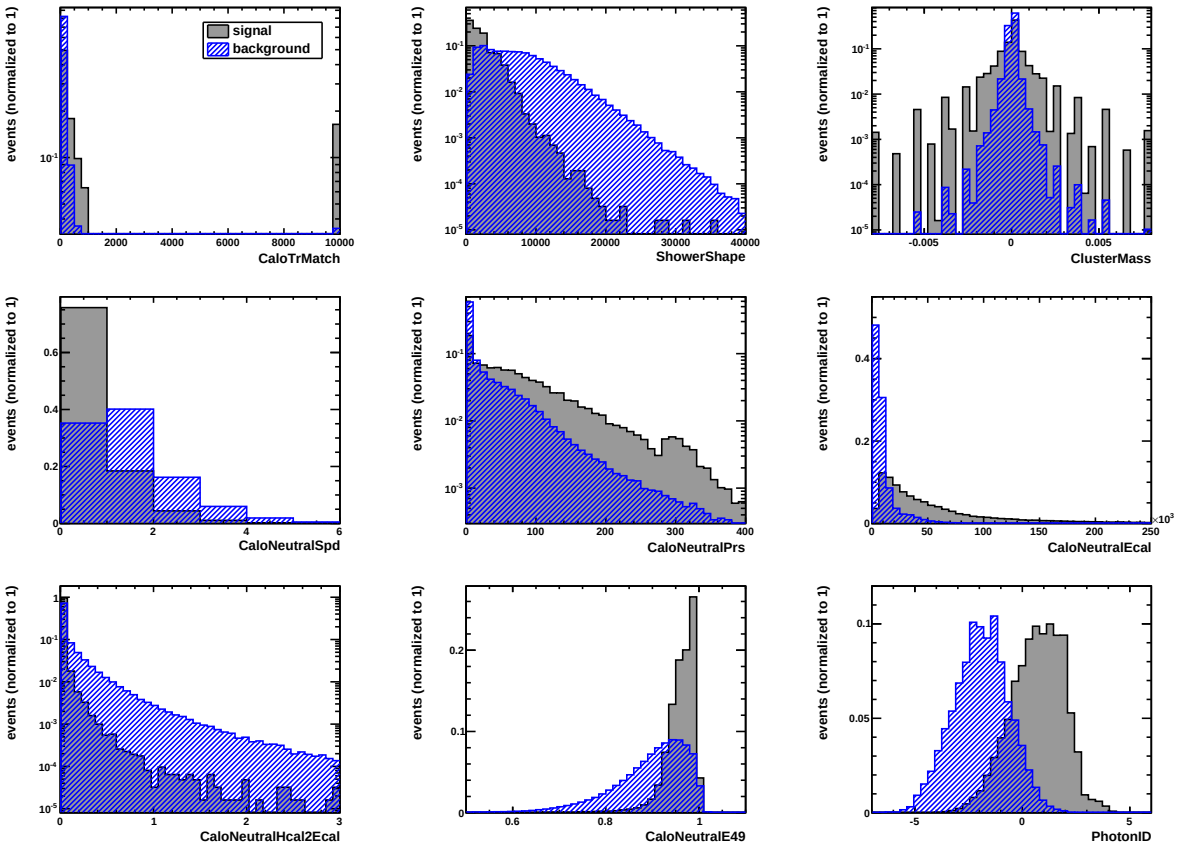

Figure 3.6: Input variables of the BDT (signal in gray, background in blue).

This BDT has a good discrimination power. The ROC-curve is plotted in Fig. 3.7(a). It illustrates the background rejection as a function of the signal efficiency. The cut efficiencies and purities are displayed for 100 times more background than signal in Fig. 3.7(b). Due to the small amount of $Z\gamma$ events the signal efficiency needs to be high. Therefore a cut value of the BDT output at > 0.05 is chosen maximizing the significance (cf. Fig. 3.7(b)). Based on an estimation in simulation, before the BDT cut approximately 28% of the background is *charged background*, the rest are *background photons*. After the selection this fraction is reduced to less than 0.5%. From now on, the *charged background* is ignored since it is even further reduced with the newIsPhoton selection discussed below.
The distributions for the BDT variable of the four classes from simulation is shown in Fig. 3.8

3.6.3 Select hard and radiative Signal

In a second step photons from decays of neutral mesons (mostly π^0) are rejected while still keeping the efficiency for the *hard signal* and the *radiative signal* photons high. If the transverse momentum of a π^0 is higher than 2 GeV/c, the two decay photons start to get merged in the same ECAL cluster and the cluster shape is very similar to that of a single photon. Therefore it is difficult to separate merged pions from single photons and a neutral pion can be easily identified as a photon or vice versa.

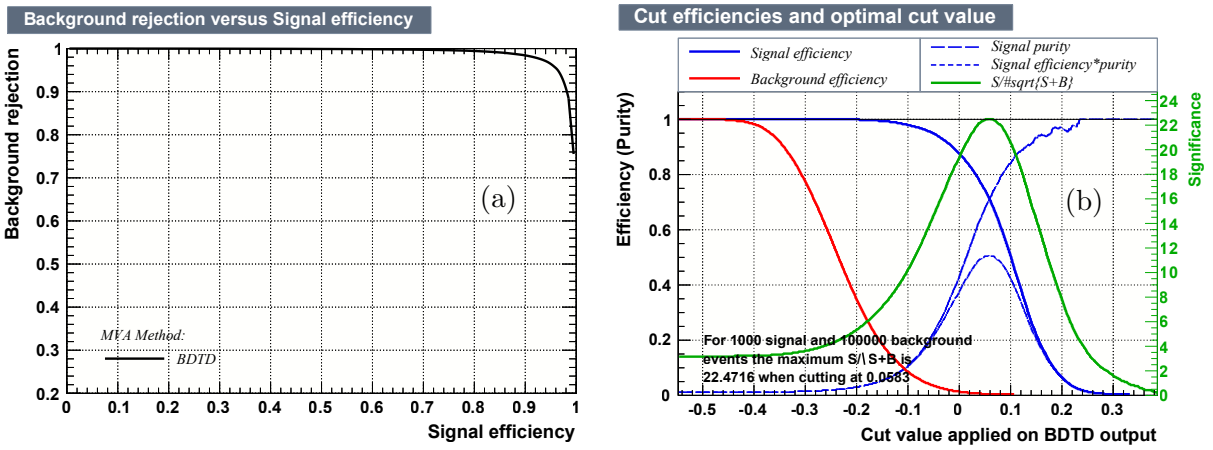

Figure 3.7: ROC curve of the BDT (a) and cut efficiencies of the output variable assuming 100 times more background than signal (b).

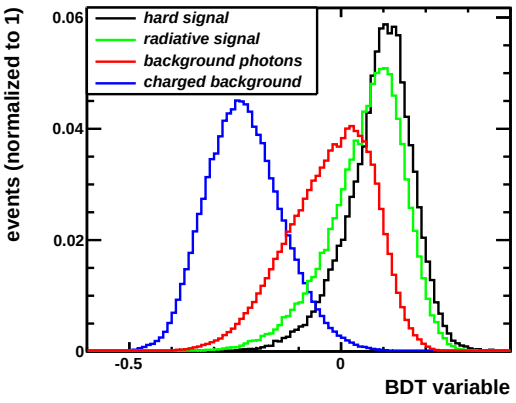

Figure 3.8: The distributions for the BDT variable of the four classes from simulation with Z selection only.

For this discrimination the variable newIsPhoton [22] [3] is used which is based on a TMVA MLP [17] (root tool for multivariate analysis with the method "Multilayer Perceptrons") with the shape of the cluster, its squashiness, the shape of the tails and other preshower and ECAL information as input [22]. It is defined for $p_T \geq 2\,\mathrm{GeV/c}$, because for lower p_T the photons from π^0 decays are reconstructed in separate clusters. The output is roughly between 0 and 1 and for $p_T < 2\,\mathrm{GeV/c}$ newIsPhoton has a value of -10000. A cut of newIsPhoton ≥ 0.8 is chosen, in order to keep the signal efficiency high while eliminating as much background as possible.

[3] So far, only internal LHCb documentation is available about this variable and no official paper.

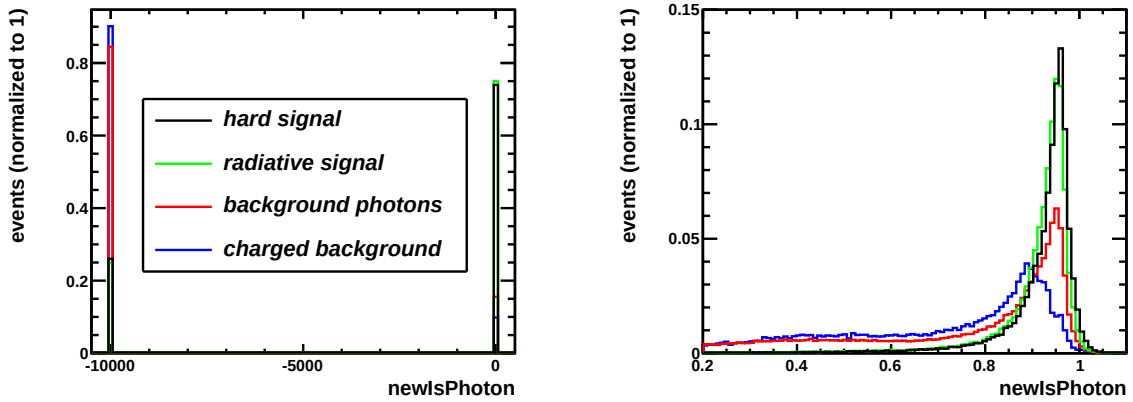

Figure 3.9: Normalized distributions of the variable newIsPhoton for the different classes of the simulations only applying the Z selection. If the $p_{T,\gamma}$ is smaller than 2 GeV/c, newIsPhoton has the value -10000. This explains the peak at -10000.

3.7 Muon Reconstruction Efficiency

The reconstruction efficiency for muons is the product of the tracking and the muon identification efficiency. These values are not calculated but taken from previous studies [4] [13], [20], [21]. The studies showed that the muon reconstruction efficiency depends on η. Dependencies on p_T or charge are small and can be neglected. The plots of the tracking- and muon identification efficiency as well as their product for a single muon are shown in Fig. 3.10(a). The efficiency for the Z reconstruction is basically the square of one muon reconstruction efficiency. Figure 3.10(b) shows the Z reconstruction efficiency as a function of η_{μ^+} and η_{μ^-}.
These efficiencies however depend on the event multiplicity. Therefore it was checked in simulation that the inclusive Z and $Z\gamma$ sample have the same multiplicity distribution. The reconstruction efficiency is calculated the following way for the two dimensional η-binning:

$$\varepsilon_{\mu\text{-rec}}(\eta_\mu) = \varepsilon_{\mu\text{-tracking}} \cdot \varepsilon_{\mu\text{-ID}} \tag{3.1}$$

$$\varepsilon_{Z\text{-rec}}(\eta_{\mu^+}, \eta_{\mu^-}) = \varepsilon_{\mu^+\text{-rec}}(\eta_{\mu^+}) \cdot \varepsilon_{\mu^-\text{-rec}}(\eta_{\mu^-}) \tag{3.2}$$

Considering the possibility of different η distributions for the two signal classes, the efficiency histograms are weighted with the number of events from simulation for each class. The resulting average Z reconstruction efficiency for the class *hard signal* is:

$$\varepsilon_{\text{rec, }\textit{hard signal}} = 0.939 \pm 0.010 \tag{3.3}$$

and for the *radiative signal*:

$$\varepsilon_{\text{rec, }\textit{radiative signal}} = 0.940 \pm 0.010\ . \tag{3.4}$$

[4] These studies use the tag and probe method with one muon satisfying all track and muon-id criteria and the other muon is selected with looser criteria that depends on the efficiency to be measured.

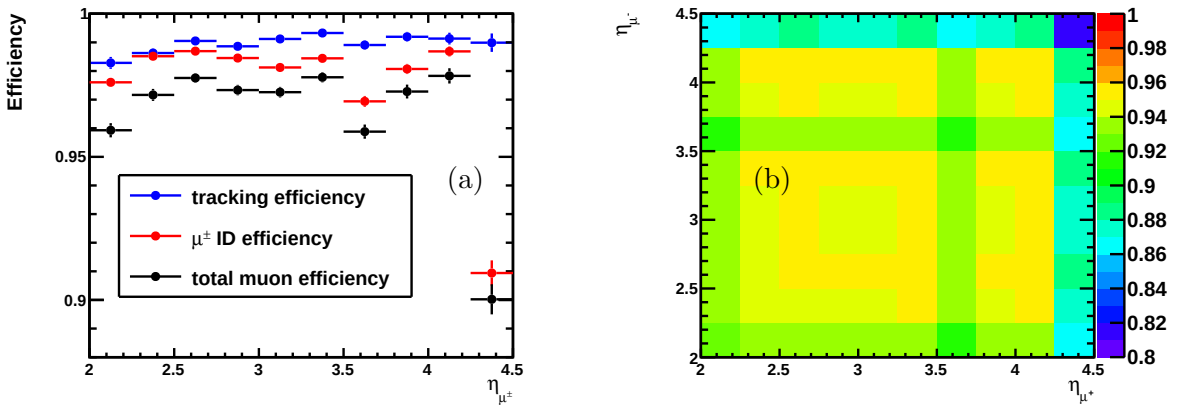

Figure 3.10: Tracking and muon identification efficiencies as a function of η (a) and the total Z reconstruction efficiency in bins of η of the two muons (b) [13], [20], [21].

3.8 Photon Selection Efficiency

The photon selection efficiency is defined as the fraction of reconstructed signal events that passed the photon selection.
Four stages lead to the calculation of the total photon selection efficiency:

1. Calculate the efficiency of the cut on newIsPhoton. An external tool [22] is used to calculate the efficiency in bins of $p_{T,\gamma}$.
2. Calculate the efficiency for the BDT selection for photons which are selected by the cut on newIsPhoton.
3. Calculate the fraction of photons which do not convert into an electron positron pair in simulation.
4. Calculate the geometrical efficiency for the cuts on ϕ assuming a flat distribution.

While the first efficiency is estimated from data all the remaining efficiencies are determined from simulation. The total selection efficiency is the product of these three efficiencies.

3.8.1 newIsPhoton Efficiency

The newIsPhoton efficiency $\varepsilon_{\text{newIsPhoton}}$ is calculated with an external tool [22], which is data driven. It uses samples of $B^0 \to K^*\gamma$ and $D^0 \to K\pi\pi^0$ candidates and separates residual background by the sPlot techniques (a statistical tool to unfold data distributions [24]) to calculate the efficiency in bins of $p_{T,\gamma}$. The efficiency for newIsPhoton is plotted in Fig. 3.11 as a function of $p_{T,\gamma}$. It drops by a few percent for high $p_{T,\gamma}$.
The average efficiency is obtained by weighting each bin with the number of events from simulation. For the two signal classes the average efficiency is:

$$\varepsilon_{\text{newIsPhoton, average, } \textit{hard signal}} = 0.865 \pm 0.014 \tag{3.5}$$

$$\varepsilon_{\text{newIsPhoton, average, } \textit{radiative signal}} = 0.863 \pm 0.014 \tag{3.6}$$

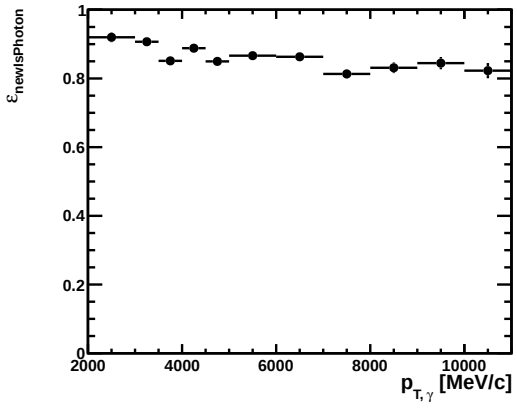

Figure 3.11: Efficiencies for the newIsPhoton cut in bins of $p_{T,\gamma}$.

3.8.2 BDT Efficiency

Simulations are used for the calculation of the BDT selection efficiency ε_{BDT}. First, a two-dimensional histogram in pseudo-rapidity (10 bins) and transverse momentum (10 bins) of the photon is filled with the Monte Carlo generated data and selected by newIsPhoton. Then the BDT selection cut is applied and a second histogram with the selected candidates is filled. The ratio of the two histograms is the BDT selection efficiency for each bin. Repeating this procedure for each event class leads to BDT selection efficiency for the different classes seperately (Fig. 3.12). The projections in pseudo-rapidity and transverse momentum are plotted in Fig. 3.13.

The histograms are weighted with the number of events per bin to get the average BDT selection efficiency. The results for the signal classes are

$$\varepsilon_{\text{BDT, }\textit{hard signal}} = 0.8432 \pm 0.0021 \tag{3.7}$$

$$\varepsilon_{\text{BDT, }\textit{radiative signal}} = 0.8071 \pm 0.0022 \tag{3.8}$$

and for the background

$$\varepsilon_{\text{BDT, }\textit{background photons}} = 0.5287 \pm 0.0018 \tag{3.9}$$

$$\varepsilon_{\text{BDT, }\textit{charged background}} = 0.0193 \pm 0.0012\ . \tag{3.10}$$

The efficiency for the *radiative signal* is slightly smaller than for the *hard signal* because of the muon which is usually close to the photon. Selecting only photons with a higher isolation to the muon leads to similar efficiencies (cf. Fig. 3.36 and Section 3.12). This selection reduces the *charged background* nearly completely while keeping the signal efficiency high. Although the BDT is only trained on *charged background* it also eliminates parts of the *background photons*. Because the input variables of the BDT are also shower shape variables (similar to the multivariate analysis of newIsPhoton), the difference of the shower shape in the calorimeter of two merged photons from π^0 decays and a single photon from the hard or radiative process

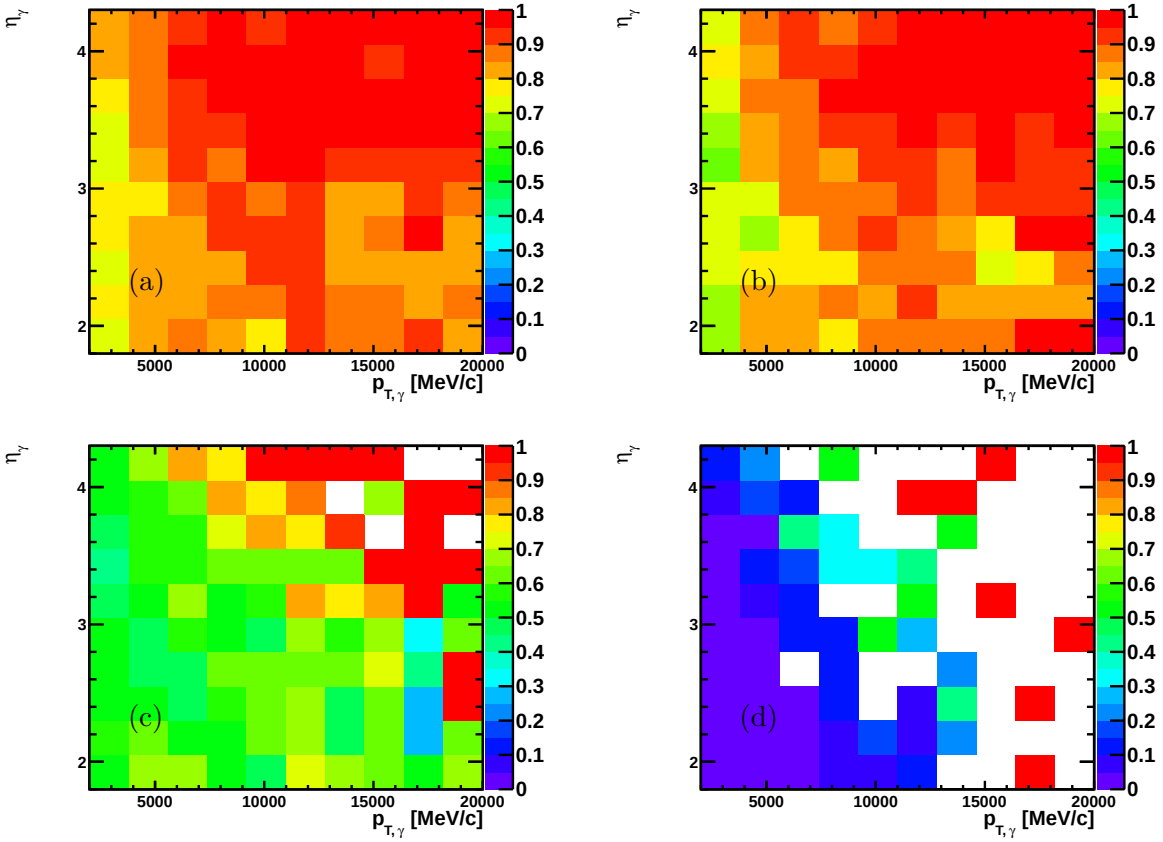

Figure 3.12: BDT selection efficiency for the *hard signal* (a), the *radiative signal* (b), *background photons* (c) and the *charged background* (d) from simulation.

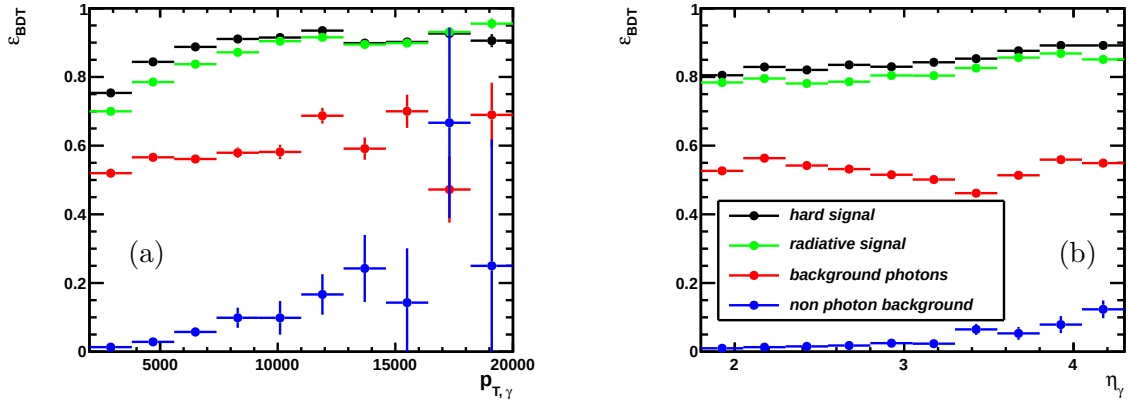

Figure 3.13: BDT selection efficiency as a function of $p_{T,\gamma}$ (a) and η_γ (b) from simulation.

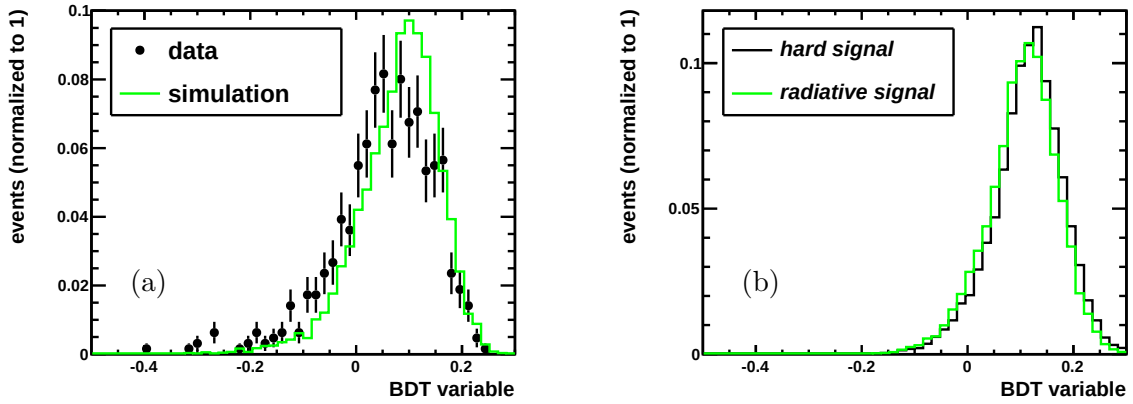

Figure 3.14: (a) BDT variable distribution for simulation and data with the cut $\Delta R(\mu) \leq 0.2$. (b) Comparison of the *hard signal* and the *radiative signal* from simulation.

is the reason for the rejection of the *background photons*. Furthermore π^0 photons are usually surrounded by other particles which could lead to a track matched to the calorimeter cluster.

Correction for Differences between Data and Simulation

As Fig. 3.2(b) and Fig. 3.3(b) show a significant difference between the distribution of the BDT variable between data and simulation, it is important to consider this fact. The distribution of the BDT variable after the newIsPhoton cut is plotted in Fig. 3.14(a) for data and *radiative signal* from simulation. The distributions are obtained by applying the cut $\Delta R(\mu) \leq 0.2$ as in Section 3.4. This selection still contains a small amount of *background photons* for data and this fraction is estimated from simulation to be $\sim 2\%$. Therefore 2% of simulated *background photons* are added to the simulation of the *radiative signal*.

Calculating the efficiency for data with the cut $\Delta R(\mu) \leq 0.2$ in the same way as before results in an average efficiency of

$$\varepsilon_{\text{BDT, } \textit{radiative signal}, \text{ data}} = 0.556 \pm 0.021 \; . \tag{3.11}$$

To take into account, that this efficiency is smaller due to the pollution through *background photons*, the efficiency in the simulation is also recalculated with the same pollution. This results in

$$\varepsilon_{\text{BDT, } \textit{radiative signal}, \text{ MC}} = 0.702 \pm 0.004 \; . \tag{3.12}$$

The difference between data and simulation is very big. The correction factor calculated from Eq. (3.12) and Eq. (3.11) turns out to be

$$f_{\text{correction}} = 0.79 \pm 0.03 \; . \tag{3.13}$$

The uncertainty is dominated by the small statistics from data.
As it is not possible to calculate $\varepsilon_{\text{BDT, } \textit{hard signal}}$ from data and because the distributions of

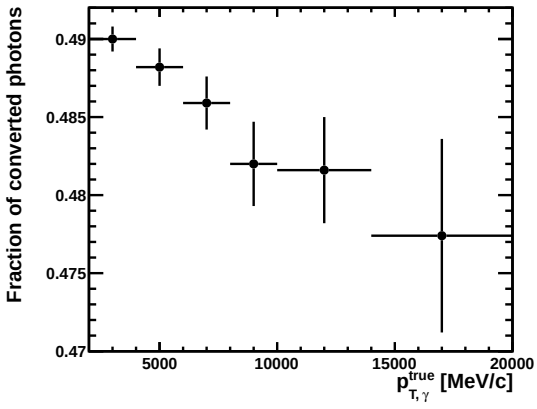

Figure 3.15: Probability of a photon converting into an electron positron pair before entering the ECAL, which is $1 - \varepsilon_{\text{unconverted}}$, in bins of $p_{T,\gamma}$.

the *hard* and the *radiative signal* of the BDT variable are very similar (cf. Fig. 3.14(b)), the differences between data and simulation are assumed to be identical for the two signal classes. The small statistics and the assumption of similar behavior for the *hard* and the *radiative signal* are the reasons for only applying half of the correction and to take half of the correction as a systematic uncertainty. This leads to a correction of

$$f_{\text{correction}} = 0.90 \pm 0.10(\text{sys}) \tag{3.14}$$

and the resulting BDT efficiencies are

$$\varepsilon_{\text{BDT, } \textit{hard signal}} = 0.758 \pm 0.002(\text{stat}) \pm 0.08(\text{sys}) \tag{3.15}$$

$$\varepsilon_{\text{BDT, } \textit{radiative signal}} = 0.726 \pm 0.002(\text{stat}) \pm 0.08(\text{sys}) \tag{3.16}$$

3.8.3 Photon Conversion Efficiency

A part of the photons which are produced at the primary vertex or nearby convert into an electron positron pair. They are not anymore detected in the ECAL as a photon, but as an electron and a positron. The BDT selection should remove this signature and therefore this part of the signal is not measured anymore. The fraction of the photons which convert are shown in Fig. 3.15 as a function of the true $p_{T,\gamma}$. This plot is taken from Ref. [23]. The histogram in Fig. 3.15 is weighted with the number of events per bin and the average photon conversion efficiency, $\varepsilon_{\text{unconverted}}$, is calculated for the signal classes. $\varepsilon_{\text{unconverted}}$ is the probability, that a photon does not convert. It follows that $\varepsilon_{\text{unconverted}} = 1 - P(\text{photon converts})$. The values are

$$\varepsilon_{\text{conversion, } \textit{hard signal}} = 0.521 \pm 0.002 \tag{3.17}$$

$$\varepsilon_{\text{conversion, } \textit{radiative signal}} = 0.517 \pm 0.002\ . \tag{3.18}$$

3.8.4 ϕ Selection Efficiency

Assuming a flat distribution in ϕ, the efficiency is the same for all different classes and it can be calculated from pure geometrical considerations:

$$\varepsilon_{\phi\text{-sel}} = \frac{1.47 - 0.1 + 3.04 - 1.67}{\pi} = 0.8722 \tag{3.19}$$

3.8.5 Total Photon Selection Efficiency

The total photon selection efficiency for the *hard signal* is the product of the four different parts. The errors are added in quadrature for the statistical and systematic uncertainty separately.

$$\varepsilon_{\gamma\text{-sel, }\textit{hard signal}} = \varepsilon_{\text{newIsPhoton}} \cdot \varepsilon_{\text{BDT}} \cdot \varepsilon_{\text{unconverted}} \cdot \varepsilon_{\phi\text{-sel}} = 0.298 \pm 0.001(\text{stat}) \pm 0.032(\text{sys}) \tag{3.20}$$

And for the *radiative signal* the following values are obtained:

$$\varepsilon_{\gamma\text{-sel, }\textit{radiative signal}} = \varepsilon_{\text{newIsPhoton}} \cdot \varepsilon_{\text{BDT}} \cdot \varepsilon_{\text{unconverted}} \cdot \varepsilon_{\phi\text{-sel}} = 0.283 \pm 0.001(\text{stat}) \pm 0.031(\text{sys}) \tag{3.21}$$

3.9 Trigger Efficiency

The trigger efficiency for selected and reconstructed muons is expressed as $\varepsilon_{\text{trig}} = N_{\text{triggered}}/N_{\text{all}}$ where N_{all} is the number of produced muons and $N_{\text{triggered}}$ the number of muons that passed the trigger requirements. It is measured for the positive and negative muon separately in bins of η.
Only triggering on a single muon trigger line on all three levels (cf. Table 3.1) allows to estimate the efficiency for the second muon by a tag and probe technique.
It is calculated by:

$$\varepsilon_{\text{trig}} = \frac{N_{\text{probe}}}{N_{\text{tag}}} \tag{3.22}$$

For the positive muon, candidates with triggered negative muons are used and these are called tagged events. The sum of these events is N_{tag}. In the set of the tagged events each event with a triggered positive muon is a probe event. The set of probe events is a subset of the tagged events and their sum is N_{probe}. The trigger efficiency for the positive muon is now calculated by Eq. (3.22). With this definition it is possible to evaluate the trigger efficiency for the positive and the negative muon separately (Fig. 3.16).

Because only one muon is required to trigger the event, the total trigger efficiency is one minus the probability, that no muon has fulfilled the trigger condition:

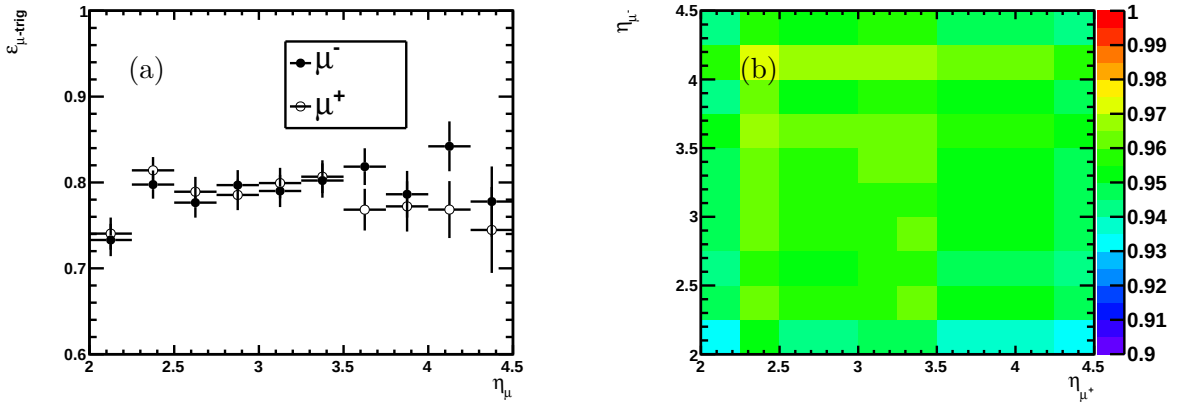

Figure 3.16: Trigger efficiencies for the positive and negative muon as a function of η (a) and Trigger efficiencies for $Z \to \mu^+\mu^-$ and $Z\gamma \to \mu^+\mu^-\gamma$ in bins of η_{μ^+} and η_{μ^-} (b).

$$\varepsilon_{\text{trig}} = 1 - (1 - \varepsilon_{\text{trig},\mu^+}) \cdot (1 - \varepsilon_{\text{trig},\mu^-}). \tag{3.23}$$

It can be calculated in bins of η_{μ^+} and η_{μ^-} (Fig. 3.16(b)). In the same way as in Section 3.7 and Section 3.8 each bin is weighted with the number of events per bin from simulation. The resulting average Z trigger efficiency for the class *hard signal* is:

$$\varepsilon_{\text{trigger, }\textit{hard signal}} = 0.95 \pm 0.04 \tag{3.24}$$

and for the *radiative signal*

$$\varepsilon_{\text{trigger, }\textit{radiative signal}} = 0.95 \pm 0.04\ . \tag{3.25}$$

3.9.1 Global Event Cut Efficiency

The L0 trigger uses a global event cut. In Table 3.1 the L0 cuts are explained and only events with less than 600 SPD hits are selected. This cut is not considered in the trigger efficiency since it does not depend on the muon which has fired the triggerline. It is always applied and that is why it is called a global event cut.
For this global selection it is important to estimate the efficiency as well. Since the multiplicity, which is described by the number of SPD hits, of the events is described insufficiently in the simulations (cf. Fig. 3.2(h) and Fig. 3.3(h) as well), this efficiency has to be calculated from data. It is difficult to isolate the *hard signal* events from the data sample, but the multiplicity in simulation is similar for the *radiative signal* and for the *hard signal* (cf. Fig. 3.17(a)). The global event cut efficiency, ε_{GEC}, can be calculated for the *radiative signal* by selecting events with $\Delta R(\mu) \leq 0.2$ and extrapolated to the *hard signal* with a correction factor arising from the small differences in simulation (cf. Fig. 3.17(a)). This procedure assumes identical difference between data and simulation for both classes.

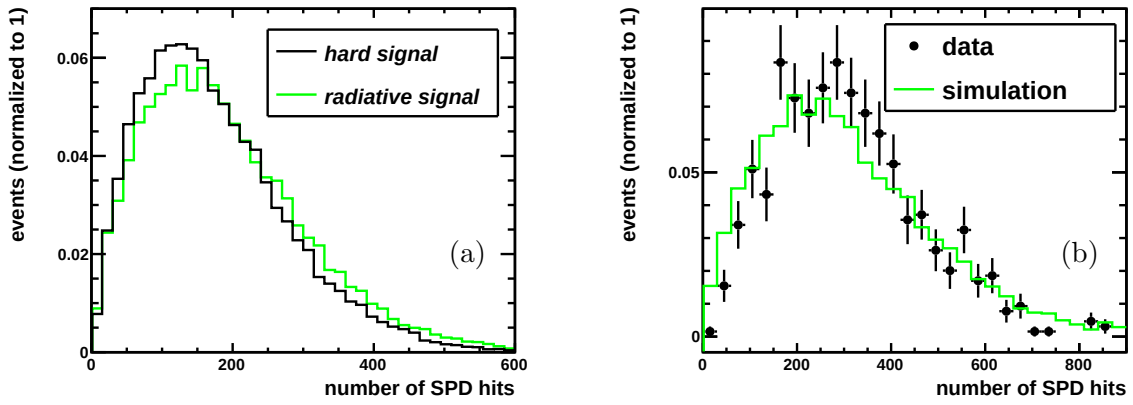

Figure 3.17: Comparison of the distribution of the number of SPD hits for the *radiative signal* and the *hard signal* for simulation (a) and the data compared to simulation (SPD hits × 1.6) for the *radiative signal* (b).

The number of SPD hits can be described by an analytic function called the gamma distribution. The general formula for the probability density function of the gamma distribution is

$$f(x) = \frac{\left(\frac{x-\mu}{\beta}\right)^{\gamma-1} \exp\left(-\frac{x-\mu}{\beta}\right)}{\beta\Gamma(\gamma)} \qquad x \geq \mu; \gamma, \beta > 0 \tag{3.26}$$

where γ is the shape parameter, μ is the location parameter, β is the scale parameter, and Γ is the gamma function which has the formula

$$\Gamma(a) = \int_0^\infty t^{a-1} e^{-t} \mathrm{d}t \; . \tag{3.27}$$

To determine the efficiency the gamma distribution $f(x)$ is fitted to the data selected by the cut $\Delta R(\mu) \leq 0.2$ in the range $0 - 600$ with a binned maximum likelihood fit with the RooFit tool [26] (cf. Fig. 3.18(a)). The efficiency is

$$\varepsilon_{\mathrm{GEC},\ \textit{radiative signal}} = \frac{\int_0^{600} f(x)\mathrm{d}x}{\int_0^{\infty} f(x)\mathrm{d}x} = 0.914 \tag{3.28}$$

In order to estimate the error, the data only requiring the trigger L0DiMuon to be fired (which has a global event cut SPD hits ≤ 900) is fitted with a gamma distribution in the range $0 - 900$ (cf. Fig. 3.18(b)). The efficiency, calculated analog to Eq. (3.28), is found to be 0.939. The difference between these two values is the uncertainty on $\varepsilon_{\mathrm{GEC},\ \textit{radiative signal}}$.

$$\varepsilon_{\mathrm{GEC},\ \textit{radiative signal}} = 0.914 \pm 0.025 \tag{3.29}$$

When looking at Fig. 3.17(b) it seems that the simulation fits the data if the number of SPD hits is multiplied by 1.6. This means the global event cut at 600 in the data is equivalent to a global event cut at 375 in the simulation. With this information the correction factor for the

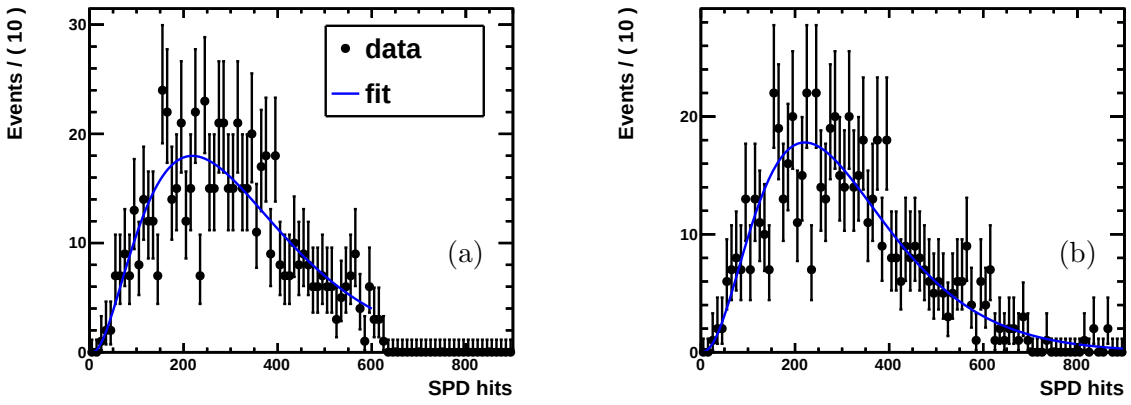

Figure 3.18: Data fit to the number of SPD hits with and without the global event cut at SPD hits ≥ 600 to estimate ε_{GEC}.

$\varepsilon_{\text{GEC, }\textit{hard signal}}$ is calculated as the ratio of the two integrals from 0 to 375 of the *hard signal* and the *radiative signal* from 3.17(a). This leads to the correction factor

$$f_{\text{correction}} = \frac{\int_0^{375} \textit{hard signal}\ \text{d(SPD hits)}}{\int_0^{375} \textit{radiative signal}\ \text{d(SPD hits)}} = 1.024 \pm 0.006 \tag{3.30}$$

where the uncertainty is obtained from varying the value of 375 as a global event cut in simulation from 350 to 400.
Combining $\varepsilon_{\text{GEC, }\textit{hard signal}}$ with $f_{\text{correction}}$ leads to

$$\varepsilon_{\text{GEC, }\textit{hard signal}} = 0.936 \pm 0.026 \tag{3.31}$$

3.10 Fit Procedure to determine the Signal Yield

The 5331 selected $Z\gamma$ candidates enter the fit procedure. The strategy for the fit consists of three steps. First, the fraction of the *radiative signal*, $f_{\textit{radiative signal}}$, is determined. Afterward this fraction is fixed. The second step is a toy model of the fit for the determination of the *hard signal* to understand the performance of the data fit. The last part is to get the *hard signal* fraction, $f_{\textit{hard signal}}$, with a fit to data and a correction with the response matrix from the toy fit. For these fits, variables which distinguish between the *hard signal*, the *radiative signal* and *background photons* are needed.
In Fig. 3.1(d) the variable $\Delta R(\mu)$ is shown. It has a good separation power for the *radiative signal* to all other classes, which have similar distributions in this variable. That is because the photons of the *radiative signal* origin from the muons of the Z decay and therefore should be close to the muon track.
Also the invariant mass of the two muons, $m_{\mu\mu}$ (cf. Fig. 3.1(f)), has some separation power, because for the *radiative signal* a part of the energy of the two muons is radiated with the photon and therefore $m_{\mu\mu}$ tends to be smaller than the actual Z mass whereas the classes *hard signal* and *background photons* reproduce the Z peak. $\Delta R(\mu)$ and $m_{\mu\mu}$ can both be used to separate

the *radiative signal*.
The next three variables make it possible to isolate the *hard signal*. The distance ΔR(all Tracks) should have separation power between the *hard* and the *radiative signal* on one side and *background photons* on the other side (cf. Fig. 3.1(c)):

- The *hard signal* is usually well isolated.
- The *radiative signal* photons are near to the muon tracks and therefore are dominated by the distance to the muons.
- The main part of *background photons* originate from π^0 or other neutral mesons in a jet and that is the reason why the ΔR to the next track tends to be smaller than for the *hard signal*.

The second variable which discriminates between the *hard signal* and background is the transverse momentum of the photon ($p_{T,\gamma}$, cf. Fig. 3.1(a)). The photons originating from decays of neutral mesons, the *background photons*, have lower momenta and energy than the photon from the *hard* or the *radiative signal* classes. Since the $p_{T,\gamma}$ distribution is very sensitive to higher order corrections and the saturation of the ECAL, this variable is not used for the fit, but only for cross-checks.
The third variable which is being used for the *hard signal* fit is the invariant mass of the two muons and the photon $m_{\mu\mu\gamma}$ (cf. Fig. 3.1(g)). It discriminates between all different classes. The *radiative signal* reproduces the Z peak. Since the *background photons* tend to have smaller momenta than the *hard signal* photons their invariant mass is smaller.
The fit methods are template fits with the ROOT tool TFractionFitter. As input templates, the variables in simulations of the different classes are used. The histograms are scaled in a way, that the χ^2 between the data and the sum of fitted templates is minimized.

3.10.1 Radiative Signal Fit

A template fit to $\Delta R(\mu)$ is used for the determination of $f_{\textit{radiative signal}}$. Two histograms from Monte Carlo simulations are used for the templates:

- The $\Delta R(\mu)$ distribution of the *radiative signal*.
- The $\Delta R(\mu)$ distribution of the *background photons*, as the distribution of the *hard signal* is exactly the same (cf. Fig. 3.1(d)), only one of them is needed.

The only free parameter is $f_{\textit{radiative signal}}$ ($f_{\textit{background photons and hard signal}} = 1 - f_{\textit{radiative signal}}$). The results are plotted in Fig. 3.19(a) and the values are:

$$\begin{aligned} f_{\textit{radiative signal}} &= 0.233 \pm 0.012 \\ f_{\textit{background photons and hard signal}} &= 0.767 \pm 0.016 \end{aligned} \tag{3.32}$$

The quality of the fit is good, the ratio χ^2/ndf is:

$$\frac{\chi^2}{\text{ndf}} = 1.419 \tag{3.33}$$

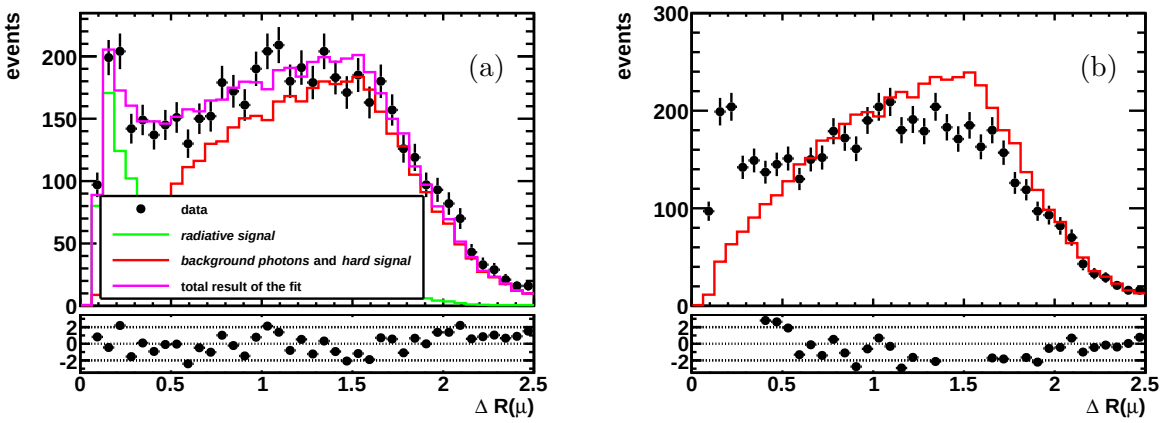

Figure 3.19: The template fit to $\Delta R(\mu)$ to determine $f_{radiative\ signal}$ with signal (a) and without signal (b).

In Fig. 3.19 the fit with and without the *radiative signal* is plotted. It is obvious that Fig. 3.19(b) does not describe the data. At low $\Delta R(\mu)$ the *radiative signal* is clearly missing and the fit compensates that by overshooting in the region of $\Delta R(\mu) \in [1, 1.8]$. While the fit with the *radiative signal* describes the data very well without any other difference from data except statistical fluctuations (see pulls).
The invariant mass of the muons is used as a cross-check. The distributions for the *radiative signal* and *background photons* in the variable $m_{\mu\mu}$ are scaled according to the fractions obtained from the fit to $\Delta R(\mu)$ in Eq. (3.32) and summed up. The result is plotted in Fig. 3.20(a) and the agreement of data and the sum of the templates is:

$$\frac{\chi^2}{\text{ndf}} = 1.500 \tag{3.34}$$

The χ^2 is slightly worse because the mass distribution is already not so well described for the inclusive Z sample [4].
As comparison the distribution without signal is shown again (Fig. 3.20(b)). The data is not described by *background photons* and the *hard signal* only. In the lower mass region the histogram undershoots the data and therefore is too high at the Z resonance. It is obvious that also in the cross-check the distribution with the *radiative signal* is preferred.

Toy Fit and Correction for the Radiative Signal Fit

The same procedure as in the next section [5] is used to correct $f_{radiative\ signal}$ from the previous section for a possible bias of the fit procedure. Simulation is used to do a toy study. The absolute value of the residuals as a function of $f_{radiative\ signal}$ and n (the number of toy candidates) is plotted in Fig. 3.21. The residuals are small and all fits are successful which means that the fit is stable and performs well.

The response matrix for $n = 5331$ candidates and 200 fits for $f_{radiative\ signal} = 0.150, 0.155, 0.160, ..., 0.395$ is shown in Fig. 3.22(a). It is used to correct the fit value from the previous section.

[5] for further information about the concept and ideas behind this study go to Section 3.10.2

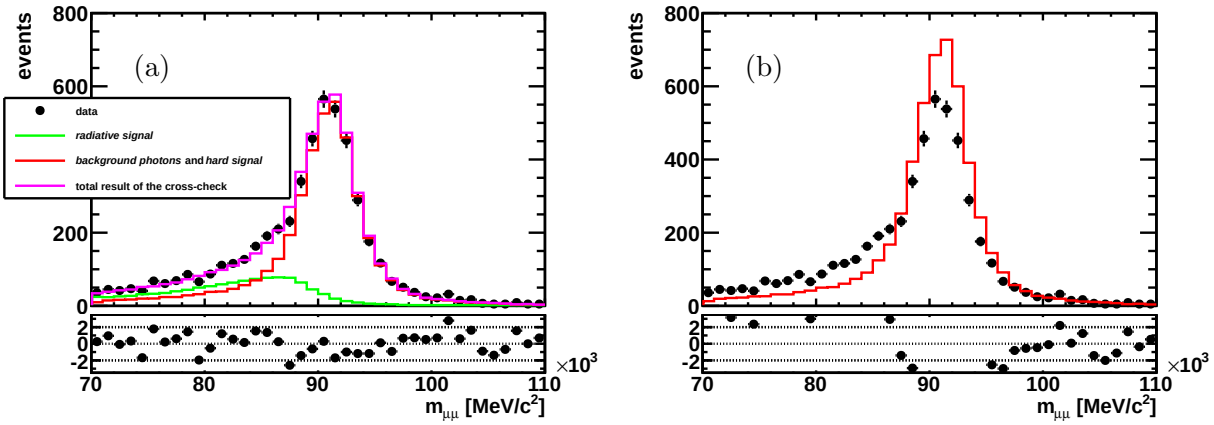

Figure 3.20: Cross-check with fractions determined by the fit to $\Delta R(\mu)$ for the invariant mass of the muons (Z mass) with signal (a) and without signal (b).

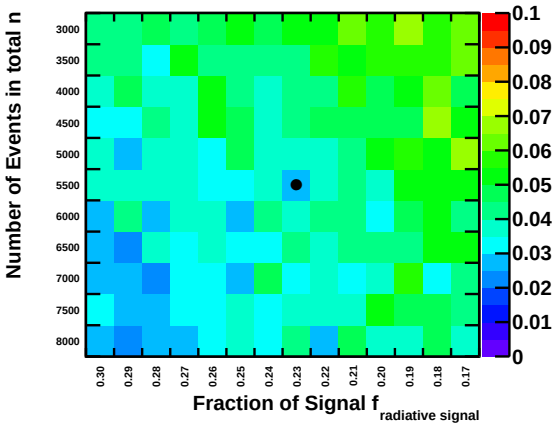

Figure 3.21: The absolute value of the residuals of the toy fits to $\Delta R(\mu)$ for 50 tries in total are shown. The black dot is the result from the fit to data.

Figure 3.22(b) is used to determine the errors on the corrected value.
The correction is almost negligible with only 1.5%. The corrected value is

$$f_{radiative\ signal} = 0.237 \pm 0.012 \tag{3.35}$$

and from now on this fraction is fixed.

3.10.2 Toy Fit

In order to understand the performance of the fit to $m_{\mu\mu\gamma}$ and ΔR(all tracks), toy studies with simulations are performed. The simulated samples are split into two parts. The first is used as data for the fit (toy data), the second to determine the templates. This splitting is necessary to avoid correlations between the toy data and the templates. For the toy data the distributions are normalized after the selection cuts for the *hard signal*, the *radiative signal* and *background*

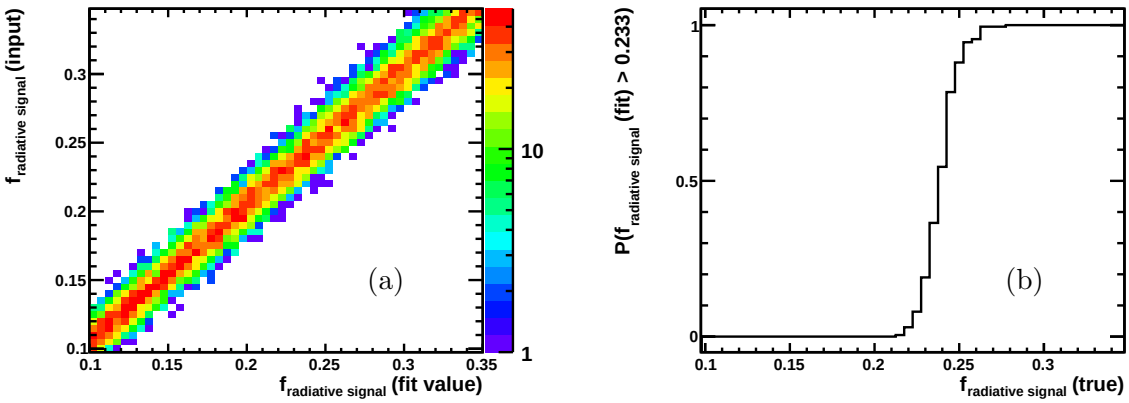

Figure 3.22: (a) The response matrix for the toy fit to $\Delta R(\mu)$. On the x-axis is the fit value, on the y-axis the fit input. (b) is the plot of the p-value.

photons and then added with different fractions [6] (e.g. $f_{hard\ signal} = 0.120, f_{radiative\ signal} = 0.237, f_{background\ photons} = 0.643$). The fraction of the *radiative signal* is always set to 0.237±0.012 (from Section 3.10.1). The resulting distribution is used to generate a toy data-set with n random events. This generated histogram is the toy data. The templates are then fitted to the toy data and compared to the input.
The only free parameter left is $f_{hard\ signal}$ ($f_{background\ photons} = 1 - (f_{radiative\ signal} + f_{hard\ signal})$) and $f_{radiative\ signal}$ can vary inside its uncertainty band. The procedure is executed for different signal fractions and total number of events to understand the performance of the fit as a function of these variables. Because the fit sometimes crashes, the fit is repeated 50 times for each number of events and signal fraction and the absolute value of the residuals ($|f_{\text{hard, input}} - f_{\text{hard, fit}}|/f_{\text{hard, input}}$) is averaged over all successful attempts. This toy fit study is repeated for a fit to $m_{\mu\mu\gamma}$ and ΔR(all tracks) in order to resolve the question which fit performs better and is more reliable and stable.

Toy Fit to $m_{\mu\mu\gamma}$

Figure 3.23 shows three examples for toy fits to $m_{\mu\mu\gamma}$ for *hard signal* fractions of 25, 12 and 6%. For all three the number of events is equal to the number of real data events ($n = 5331$) and $f_{hard\ signal}$ is varied. The left plot shows the toy data, the middle the fit without signal and the right the fit with signal. They should give an impression how this study works and performs. In Table 3.3 the results of these examples are presented. All three examples prefer the fit with signal according to the χ^2. The extracted signal fraction is however higher than the input.

In Fig. 3.24(a) the residuals as a function of the number of events and $f_{hard\ signal}$ are shown. There is an overview over possible values (Fig. 3.24(a) and (b)) and a magnification of the region where the result is expected to be (Fig. 3.24(c) and (d)). For all bins the fit procedure is successful in more than 80% of the cases (cf. Fig. 3.24(b) and (d)) which means it is stable. As long as the signal fraction is higher than 5% and there are more than 500 signal events, the fit has uncertainties below 30%.

[6] The *charged background* class is ignored, because the BDT selection should have removed them nearly completely.

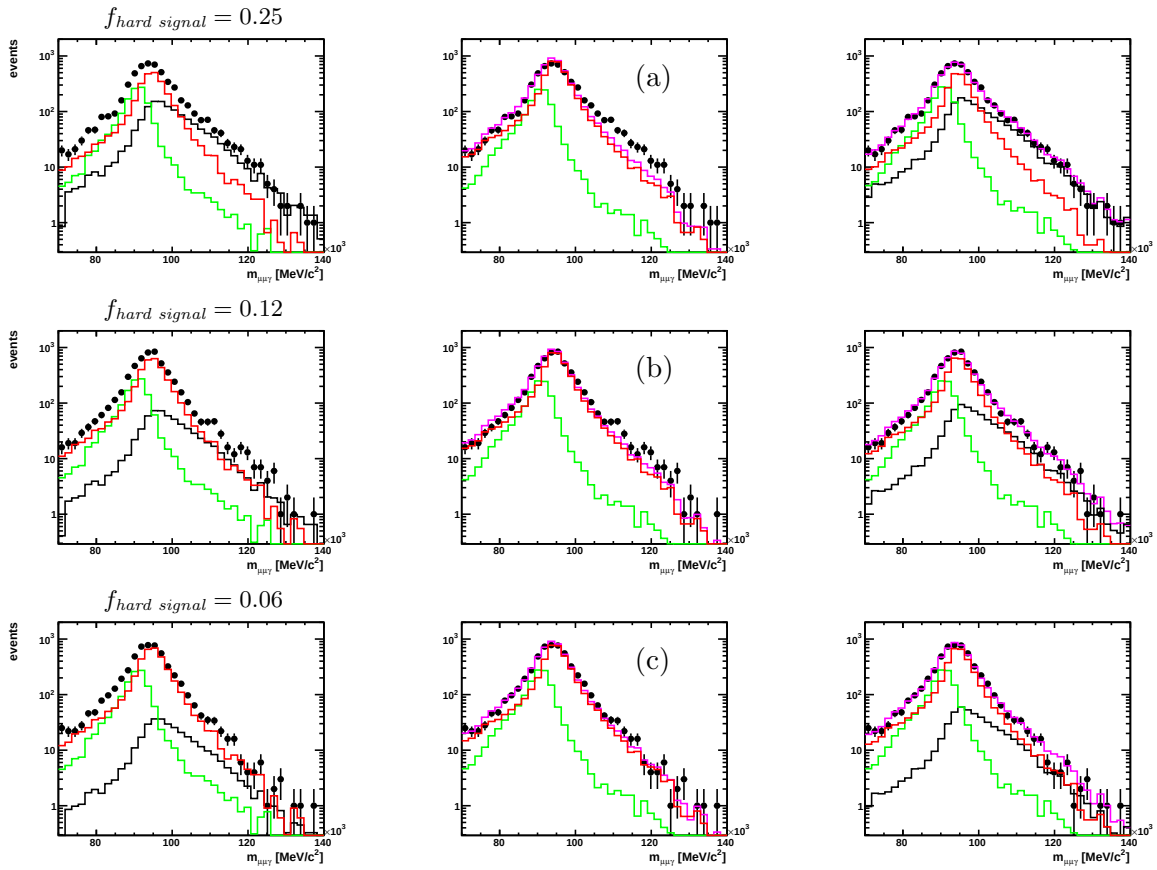

Figure 3.23: Toy model for the template fit to $m_{\mu\mu\gamma}$. On the left: Input distributions of the different classes and their sum which is the toy data. In the middle: Fit with background only. On the right: Fit with background and signal. The data points are the black dots, the *hard signal* is in black, the *radiative signal* in green, *background photons* in red and the result of the fit in magenta.

variable	true values	fit result without signal	fit result with signal
$f_{hard\ signal}$	0.250		0.281 ± 0.018
$f_{radiative\ signal}$	0.237	0.225 ± 0.012	0.249 ± 0.016
$f_{background\ photons}$	0.513	0.784 ± 0.014	0.468 ± 0.021
χ^2/ndf		9.185	1.177
$f_{hard\ signal}$	0.120		0.150 ± 0.016
$f_{radiative\ signal}$	0.237	0.225 ± 0.012	0.225 ± 0.015
$f_{background\ photons}$	0.643	0.787 ± 0.014	0.629 ± 0.021
χ^2/ndf		3.959	1.405
$f_{hard\ signal}$	0.060		0.095 ± 0.016
$f_{radiative\ signal}$	0.237	0.249 ± 0.015	0.249 ± 0.021
$f_{background\ photons}$	0.703	0.756 ± 0.014	0.655 ± 0.022
χ^2/ndf		2.558	1.496

Table 3.3: Results of examples of the toy fit to $m_{\mu\mu\gamma}$.

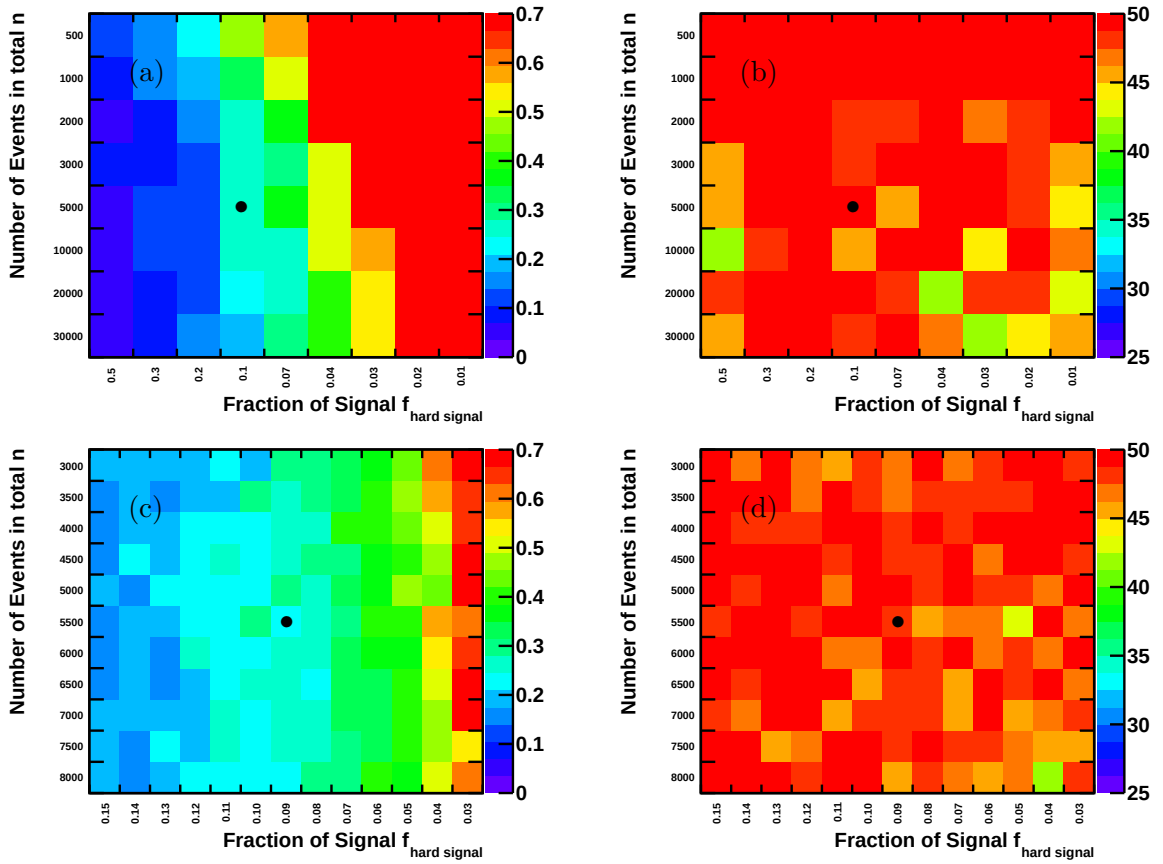

Figure 3.24: The residuals of the toy fits to $m_{\mu\mu\gamma}$ (left) and the number of successful fits for each bin (right) of 50 tries in total are shown. On top is an overview, below a magnification of the region where the fit values are expected to be. The black dot is the result from the fit to data.

It can be summarized that the fit is stable and performs as good as it is possible with the limited statistics.

If looking at various single fits (cf. Fig. 3.23 and Table 3.3) it seems likely that the fit has a bias and more often has too large or too small signal fractions, depending on the input variables. To check this the toy study is repeated with the number of events equal to the number of selected data candidates ($n = 5331$). To minimize the statistical uncertainties, the fit is repeated 200 times for each signal fraction $f_{hard\ signal} = 0.000, 0.005, 0.010, ..., 0.245$. The results are filled in a two dimensional histogram with $f_{hard\ signal}$ (fit value) on the x-axis and $f_{hard\ signal}$ (input) on the y-axis. This histogram is equal to a response matrix and can be used to correct the result in Section 3.10.3. It is plotted in Fig. 3.25.

Toy Fit to ΔR(all Tracks)

As in the previous section three examples for toy fits to ΔR(all Tracks) are presented first in Fig. 3.26. For all three the number of signal events is $n = 5331$ and $f_{hard\ signal}$ is varied. The left

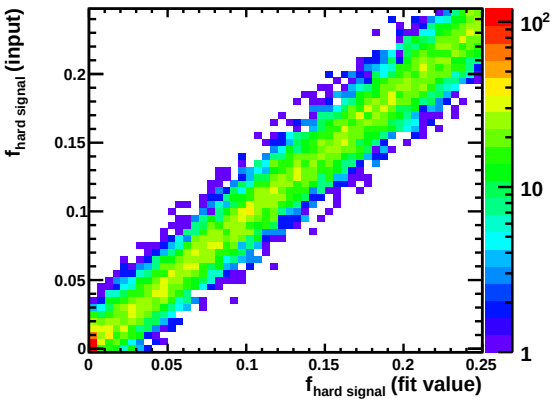

Figure 3.25: Response matrix for the toy fit to $m_{\mu\mu\gamma}$. On the x-axis is the fit value, on the y-axis the fit input.

variable	true values	fit result with signal	fit result without signal
$f_{hard\ signal}$	0.250		0.310 ± 0.033
$f_{radiative\ signal}$	0.237	0.249 ± 0.018	0.249 ± 0.021
$f_{background\ photons}$	0.513	0.749 ± 0.015	0.440 ± 0.034
χ^2/ndf		3.338	1.413
$f_{hard\ signal}$	0.120		0.088 ± 0.035
$f_{radiative\ signal}$	0.237	0.249 ± 0.021	0.249 ± 0.012
$f_{background\ photons}$	0.643	0.750 ± 0.016	0.663 ± 0.037
χ^2/ndf		1.039	0.902
$f_{hard\ signal}$	0.060		0.045 ± 0.043
$f_{radiative\ signal}$	0.237	0.249 ± 0.023	0.249 ± 0.012
$f_{background\ photons}$	0.703	0.750 ± 0.016	0.706 ± 0.042
χ^2/ndf		0987	0.970

Table 3.4: Results of examples of the toy fit to ΔR(all tracks).

plot shows the toy data, the middle the fit without signal and the right the fit with signal. In Table 3.4 the results of theses examples are presented. The first two examples prefer the fit with signal. The one with $f_{hard\ signal} = 6\%$ has a too small χ^2 and does not find the signal anymore reliably.

In Fig. 3.27 the residuals of the toy fit to ΔR(all Tracks) are shown. There is an overview over possible values (Fig. 3.27(a) and (b)) and a magnification of the region where the result is expected to be (Fig. 3.27(c) and (d)). If the signal fraction is below 0.05 the fit is not able to find the signal reliably and crashes regularly (cf. Fig. 3.27(b) and (d) where the number of successful fits decreases rapidly with smaller signal fractions). Nevertheless for all bins the fit procedure is successful in more than 50% of the cases (cf. Fig. 3.27(b) and (d)).
In general the performance is inferior to the toy fit to $m_{\mu\mu\gamma}$. The uncertainties are $10 - 20\%$ higher and the fit crashes more often.

on the $f_{hard\ signal}$. Thus the same study is applied and the results are plotted in Fig. 3.28 as a

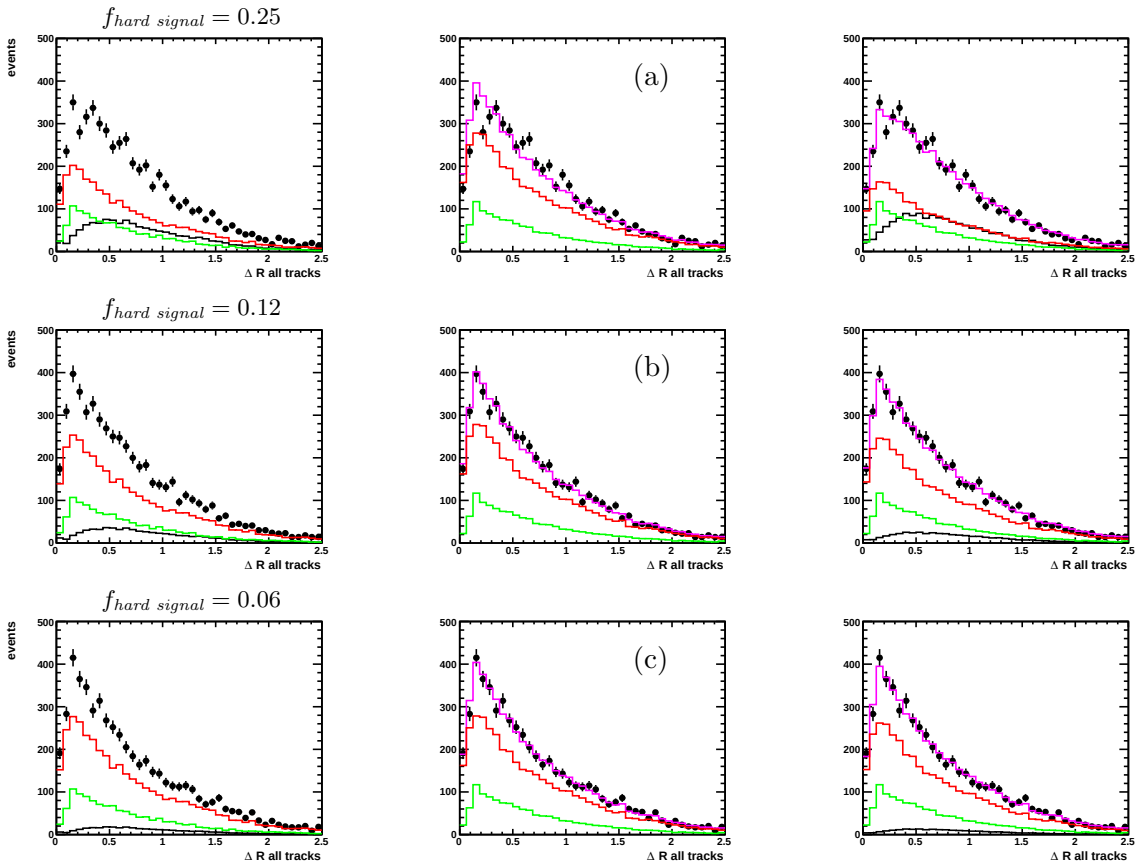

Figure 3.26: Toy model for the template fit to ΔR(all tracks). On the left: Input distributions of the different classes and their sum which is the toy data. In the middle: Fit with background only. On the right: Fit with background and signal. The data points are the black dots, the *hard signal* is in black, the *radiative signal* in green, *background photons* in red and the result of the fit in magenta.

response matrix.

A comparison between Fig. 3.25 and Fig. 3.28 shows again the better performance of the fit to $m_{\mu\mu\gamma}$. The fit values in Fig. 3.28 spread much more and therefore the band of possible results is much wider.

3.10.3 Hard Signal Fit

As a last step of the fit procedure, the *hard signal* fraction ($f_{hard\ signal}$) is obtained. Section 3.10.2 showed that the fit to $m_{\mu\mu\gamma}$ is more stable and performs better than the fit to ΔR(all tracks). However the fit is proceeded for both variables to cross-check and compare the results. The only free parameter is $f_{hard\ signal}$ and $f_{radiative\ signal}$ can vary inside its uncertainties, like in the toy fit.

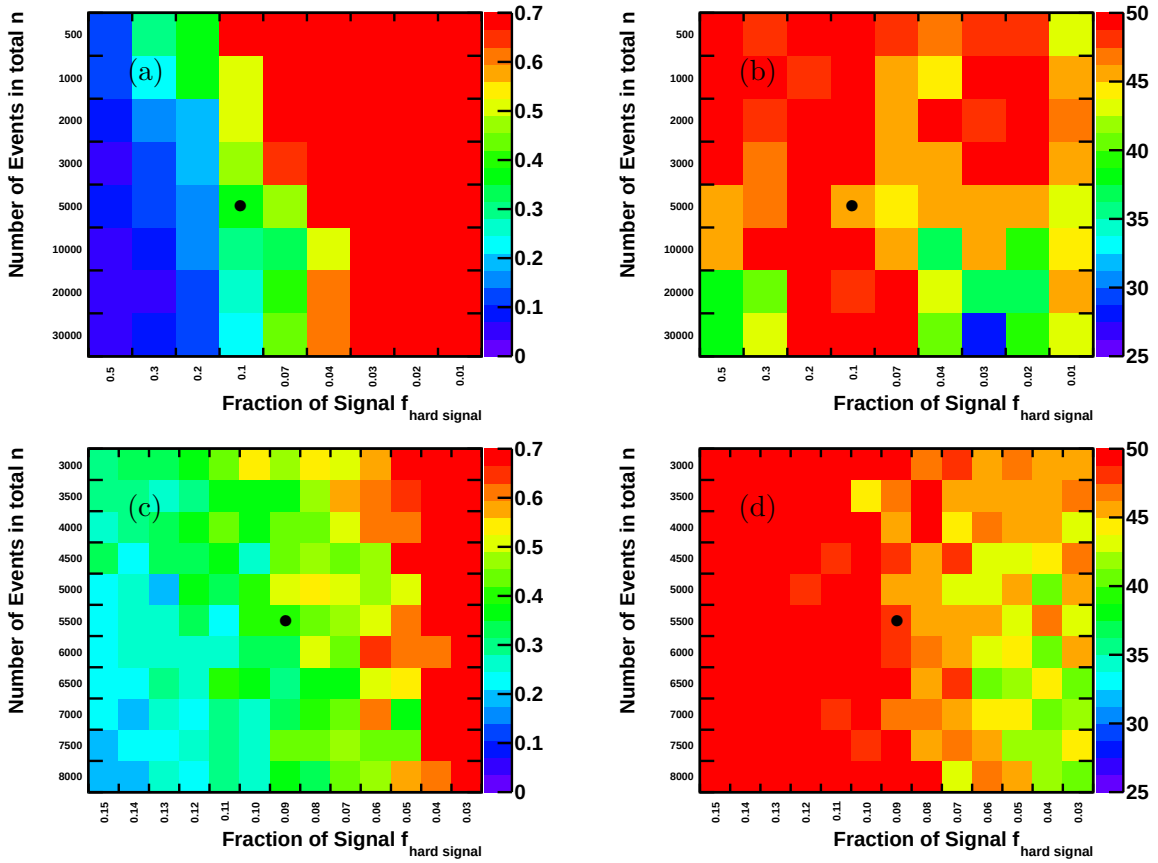

Figure 3.27: The residuals of the toy fits of ΔR(all Tracks) (left) and the number of successful fits for each bin (right) of 50 tries in total are shown. On top is an overview, below a magnification of the region where the fit values are expected to be. The black dot is the result from the fit to data.

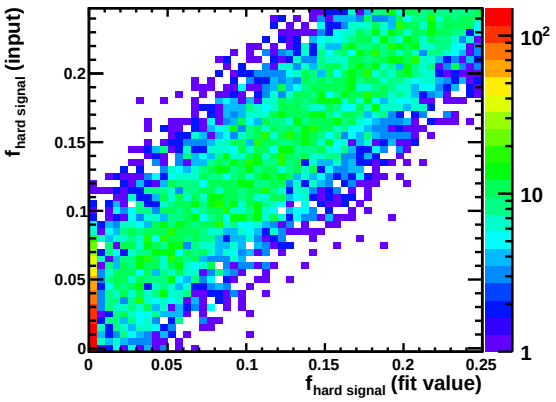

Figure 3.28: Response matrix for the toy fit to ΔR(all tracks). On the x-axis is the fit value, on the y-axis the fit input.

Hard Signal Fit to $m_{\mu\mu\gamma}$

The result of the template fit (Fig. 3.29(a)) is the following:

$$\begin{aligned} f_{hard\ signal} &= 0.095 \pm 0.016 \\ f_{radiative\ signal} &= 0.231 \pm 0.020 \\ f_{background\ photons} &= 0.673 \pm 0.026 \end{aligned} \tag{3.36}$$

The quality of the fit is good, the ratio χ^2/ndf is:

$$\frac{\chi^2}{\text{ndf}} = 1.168 \tag{3.37}$$

As comparison the fit without signal (Fig. 3.29(b), the only free parameter is $f_{radiative\ signal}$ inside its error band) leads to:

$$\begin{aligned} f_{radiative\ signal} &= 0.225 \pm 0.012 \\ f_{background\ photons} &= 0.779 \pm 0.014 \end{aligned} \tag{3.38}$$

The quality of the fit is inferior to the one with signal. The ratio χ^2/ndf is:

$$\frac{\chi^2}{\text{ndf}} = 2.684 \tag{3.39}$$

The plots of these two fits are shown in Fig. 3.29 with the pulls below. The pulls are calculated in the following way:

$$p = \frac{h_D - h_F}{\sigma_D} \tag{3.40}$$

Where h_D is the number of data events per bin, h_F the fit value per bin and σ_D the uncertainty on the data per bin.
At high masses, the fit without the *hard signal* undershoots the data due to the lack of the high energetic *hard signal* photons while the distribution is well described if the *hard signal* is included. The pulls clearly confirm this behavior. Until $100\,\text{GeV/c}^2$ the pulls in the fit without signal are negative, at higher masses they are positive while they are fluctuating around zero over the full mass range for the fit with the *radiative signal* included in Fig. 3.29(a).

The cross-check variables for the fit are $p_{T,\gamma}$ and ΔR(all tracks). The agreement (χ^2/ndf) of data and the sum of the templates, scaled with the parameters from the fit, is calculated for the fit with signal (Eq. (3.36)) and without signal (Eq. (3.38)). A comparison of the two values can give an estimate of the quality of the fit. They are shown in Table 3.5.
The plots with pulls are in Fig. 3.30 and Fig. 3.31.

The $p_{T,\gamma}$ is not described well for the cross-checks with and without signal, especially in Fig. 3.30 at the saturation of the ECAL around $p_{T,\gamma} \approx 10\,\text{GeV/c}$ the simulation describes the data

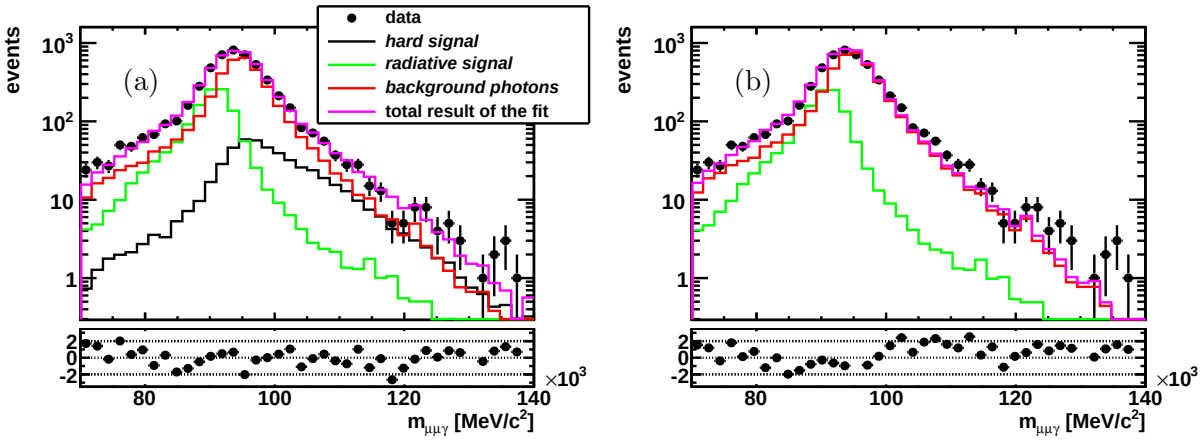

Figure 3.29: Template fit to $m_{\mu\mu\gamma}$ to determine $f_{hard\ signal}$. In (a) the fit with signal is plotted, in (b) without signal.

variable	χ^2/ndf for fit with signal	χ^2/ndf for fit without signal
$p_{T,\gamma}$	1.854	3.696
ΔR(all tracks)	1.119	1.195

Table 3.5: χ^2/ndf for the cross-check variables using the fractions as determined from the template fit to $m_{\mu\mu\gamma}$ with and without signal.

not adequately. At higher $p_{T,\gamma}$ the cross-check without the *hard signal* is always below data indicating that high high momentum photons from the hard process are missing.
In Fig. 3.31 the cross-check without the *hard signal* overshoots the data at small values. The reason for this is an overcompensation from the higher *background photons* fraction due to the lack of *hard signal* at large ΔR(all tracks).
Both cross-checks definitely prefer the fit with *hard signal*.

Hard Signal Fit to ΔR(all tracks)

The result of the template fit to ΔR(all tracks) (Fig. 3.32(a)) is:

$$\begin{aligned} f_{hard\ signal} &= 0.068 \pm 0.037 \\ f_{radiative\ signal} &= 0.249 \pm 0.012 \\ f_{background\ photons} &= 0.686 \pm 0.038 \end{aligned} \tag{3.41}$$

The quality of the fit is good, the ratio χ^2/ndf:

$$\frac{\chi^2}{\text{ndf}} = 1.142 \tag{3.42}$$

The comparison to the fit without signal (Fig. 3.32(b)), where the only free parameter is $f_{radiative\ signal}$ inside its error band, leads to:

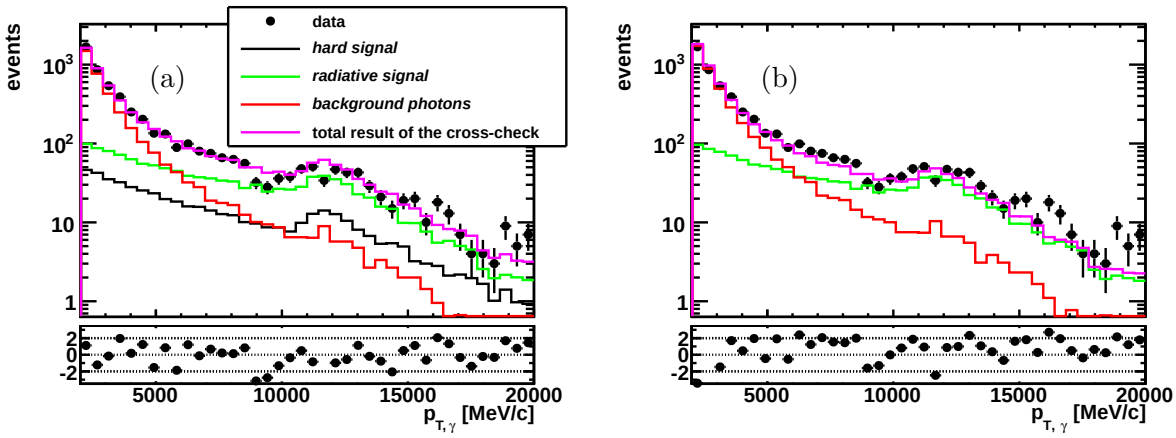

Figure 3.30: Cross-check with the fractions as determined from the template fit to $m_{\mu\mu\gamma}$ for the variable $p_{T,\gamma}$ with signal (a) and without signal (b).

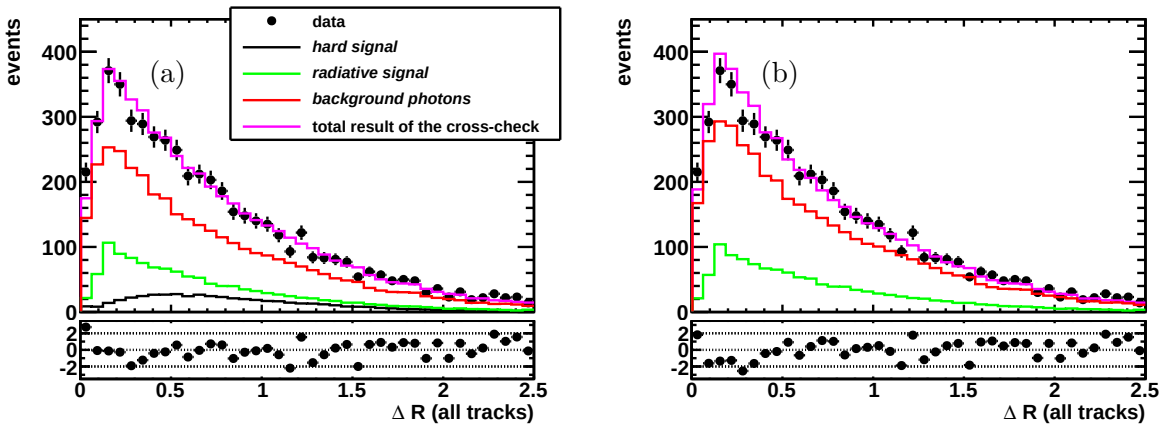

Figure 3.31: Cross-check with the fractions as determined from the template fit to $m_{\mu\mu\gamma}$ for the variable ΔR(all tracks) with signal (a) and without signal (b).

$$\begin{aligned} f_{radiative\ signal} &= 0.249 \pm 0.014 \\ f_{background\ photons} &= 0.753 \pm 0.015 \end{aligned} \tag{3.43}$$

The quality of the fit is inferior to the one with signal. The ratio χ^2/ndf is:

$$\frac{\chi^2}{\text{ndf}} = 1.235 \tag{3.44}$$

The difference in χ^2 is small for the two fits but differences are observed in the description of the data (cf. Fig. 3.32).
Due to the absence of the *hard signal* on the higher values of ΔR(all tracks) the fit raises the fraction of *background photons* in order to match the data-points at higher values of ΔR(all tracks). This leads to a slight overshoot at small ΔR(all tracks) because the distribution of the *background photons* is peaking at ΔR(all tracks) ≈ 0.2 while the shape of the *hard signal* is flatter

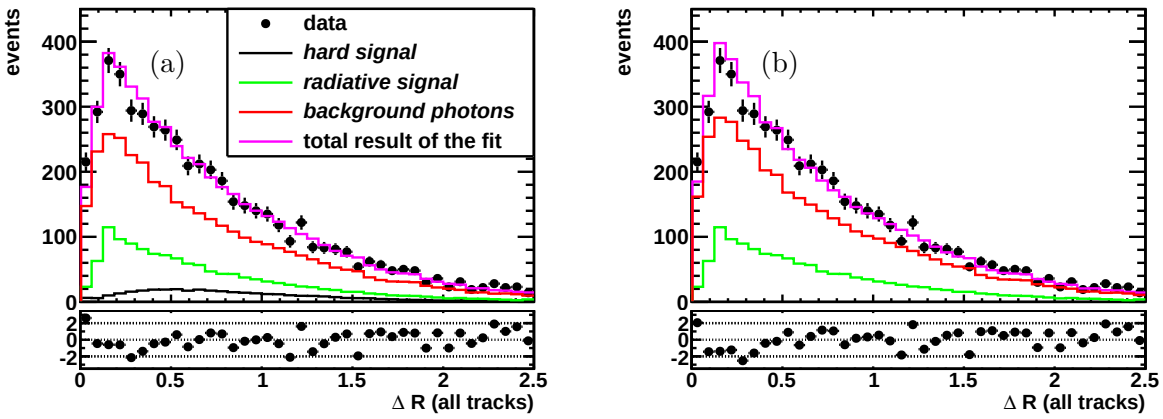

Figure 3.32: Template fit to ΔR(all tracks) to determine $f_{hard\ signal}$. In (a) the fit with signal is plotted, in (b) without signal.

variable	$\frac{\chi^2}{\text{ndf}}$ for fit with signal	$\frac{\chi^2}{\text{ndf}}$ for fit without signal
$p_{T,\gamma}$	1.888	2.854
$m_{\mu\mu\gamma}$	1.416	2.918

Table 3.6: χ^2/ndf for the cross-check variables using the fractions as determined from the template fit to ΔR(all tracks) with and without signal.

and peaks at ΔR(all tracks) ≈ 0.5. But this effect is not that significant as in the fit to $m_{\mu\mu\gamma}$. This also becomes clear when looking at the uncertainties of two fits in Eq. (3.41) and Eq. (3.36). The fit to $m_{\mu\mu\gamma}$ has much smaller uncertainties on the *hard signal* fraction.

The cross-check variables for the fit are $p_{T,\gamma}$ and $m_{\mu\mu\gamma}$. The same considerations as in the last preceding section apply. The agreement (χ^2/ndf) of data and the sum of the templates, scaled with the parameters of the fit, are shown in Table 3.6.

For this version of the fit both cross-checks prefer the fraction as determined by the fit with the *hard signal*, too.

Correction with the Response Matrix

The values from the two fits are corrected with the response matrices (Fig. 3.25 and Fig. 3.28). The lower uncertainty is obtained by finding the input value of the toy fit, that has a 16% probability to have a larger fit value than the value from the fit to data. The upper uncertainty is calculated equally, but the probability has to be 84% In Fig. 3.35 the probabilities for having a greater fit value than the value from the fit to data are shown as a function of the true input value. These plots are obtained from Fig. 3.25 and Fig. 3.28 by integrating from 0 to $f_{hard\ signal}$ (fit value) for the channel along the x-axis for each $f_{hard\ signal}$ (input) bin.
This plot confirms again that the fit to $m_{\mu\mu\gamma}$ performs much better than the fit to ΔR(all tracks). The procedure leads to a normal 1σ uncertainty level.

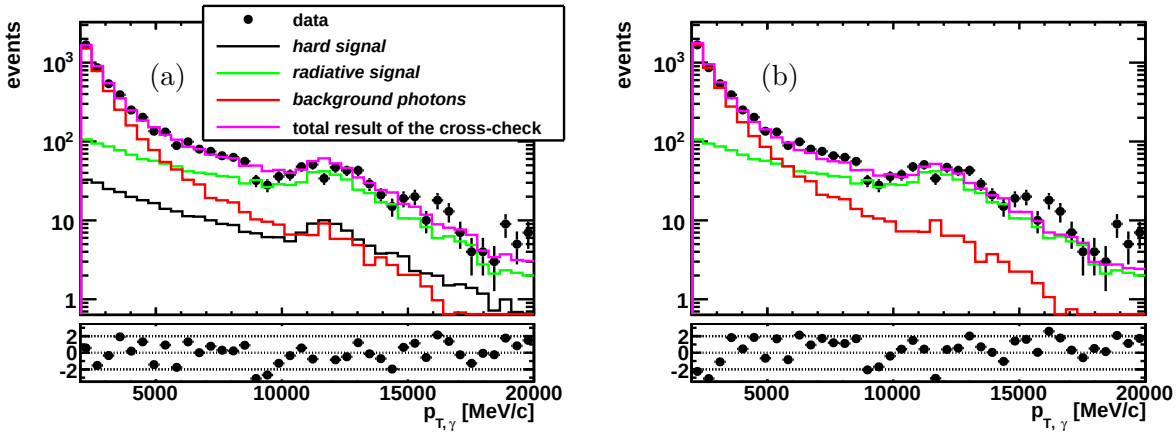

Figure 3.33: Cross-check with the fractions as determined from the template fit to ΔR(all tracks) for the variable $p_{T,\gamma}$ with signal (a) and without signal (b).

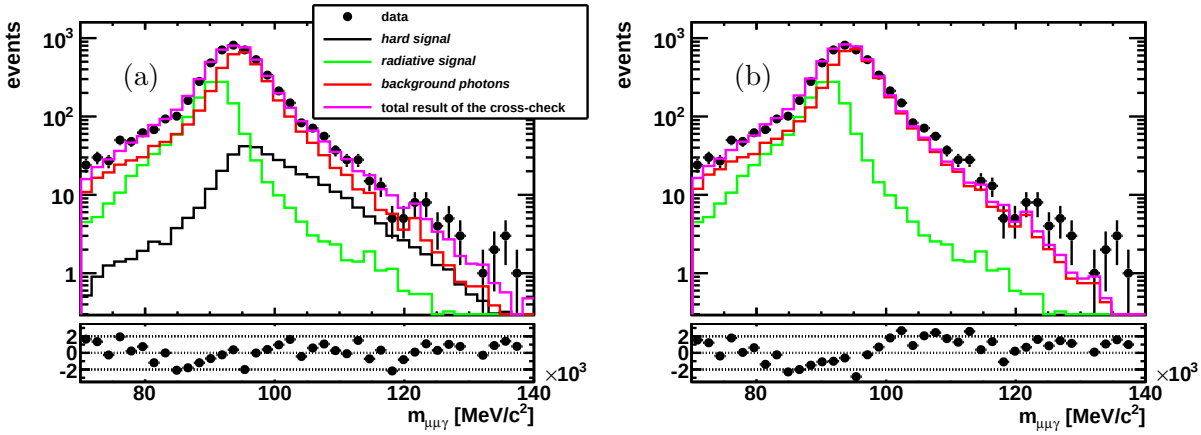

Figure 3.34: Cross-check with the fractions as determined from the template fit to ΔR(all tracks) for the variable $m_{\mu\mu\gamma}$ with signal (a) and without signal (b).

In Fig. 3.35(a) $P(f_{hard\ signal}(\text{fit}) > 0.095) = 0.5$ is the corrected hard signal fraction. $P(f_{hard\ signal}(\text{fit}) > 0.095) = 0.16$ the lower 1σ confidence level limit and $P(f_{hard\ signal}(\text{fit}) > 0.095) = 0.84$ the higher limit. In the case of the fit to ΔR(all tracks), Fig. 3.35(b), 0.095 has to be replaced by 0.068.
This leads to the following *hard signal* fraction for the fit to $m_{\mu\mu\gamma}$:

$$f_{hard\ signal} = 0.095 \pm 0.015 \tag{3.45}$$

and for the fit to ΔR(all tracks):

$$f_{hard\ signal} = 0.096 \pm 0.040\ . \tag{3.46}$$

The uncertainties on $f_{hard\ signal}$ from the plot of the p-value are in good agreement with the uncertainties from the fit to data (cf. Eq. (3.36) and Eq. (3.41)).

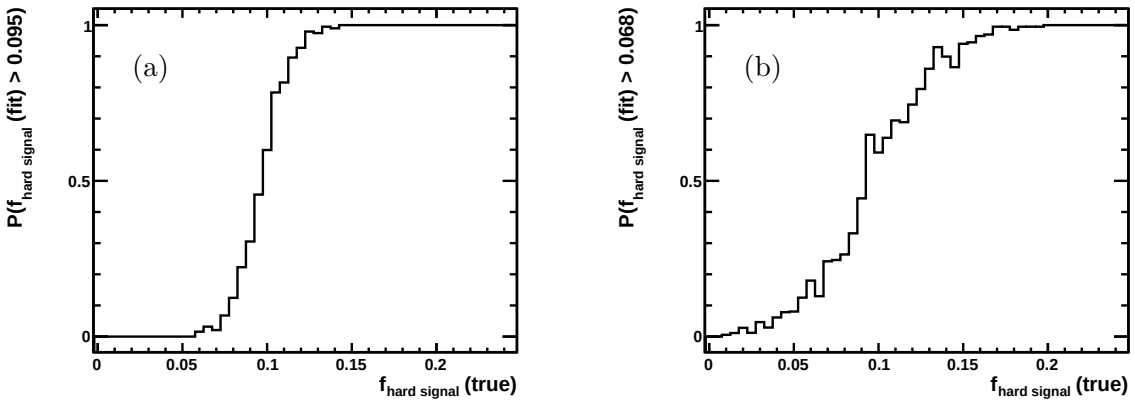

Figure 3.35: Probability from the toy fits to get a larger fit value than the value from the fit to data for $m_{\mu\mu\gamma}$ (a) and ΔR(all tracks) (b) which is basically the p-value.

Combination of the two Fits

The values for $f_{hard\ signal}$ are the weighted average of the fit to $m_{\mu\mu\gamma}$ and $\Delta R(all\ tracks)$. This leads to a signal fraction of

$$f_{hard\ signal} = 0.095 \pm 0.014\ . \tag{3.47}$$

3.11 Systematic Uncertainties

In this analysis, the statistical uncertainty should dominate the measurement for the *hard signal* except the systematic uncertainty on $\varepsilon_{\mathrm{BDT}}$. Due to the higher number of *radiative signal* candidates the systematic uncertainty is more important for this class and is higher than the statistical uncertainty. The errors contributing to the systematic uncertainty are summarized in Table 3.7 and their relative uncertainty is shown. The other uncertainties are statistical as they could be reduced with more data.
The most important factor is the uncertainty on $\varepsilon_{\mathrm{BDT}}$ originating from the problems in the calculation of the efficiency of the BDT selection as mentioned in Section 3.8.2. The second systematic uncertainty that plays an important role is the one on the integrated luminosity.
The systematic uncertainties are added in quadrature and the total systematic uncertainty for the *hard signal* is 11.3% and for the *radiative signal* it is 11.7%.

3.12 Result

Having the signal yields for the *radiative signal* and the *hard signal*, $f_{hard\ signal} = 0.095 \pm 0.014$ and $f_{radiative\ signal} = 0.237 \pm 0.012$, the total number of candidates, $n = 5331 \pm 73$, and the integrated luminosity of 2.06 ± 0.07 fb^{-1} it is possible to calculate the production cross section with the formula

quantity	relative uncertainty for *hard signal*	relative uncertainty for *radiative signal*
the integrated luminosity	3.5%	3.5%
ε_{rec}	1.1%	1.1%
$\varepsilon_{\text{newIsPhoton}}$	1.6%	1.6%
ε_{BDT}	10.6%	11.0%
$\varepsilon_{\text{unconverted}}$	0.4%	0.4%
total	11.3%	11.7%

Table 3.7: Summary of the systematic uncertainties.

efficiency	*hard signal*	*radiative signal*
$\boldsymbol{\varepsilon_{\text{rec}}}$	$\mathbf{0.939 \pm 0.010(sys)}$	$\mathbf{0.940 \pm 0.010(sys)}$
$\varepsilon_{\text{tracking}}$	$0.978 \pm 0.007(\text{sys})$	$0.978 \pm 0.007(\text{sys})$
$\varepsilon_{\mu-\text{ID}}$	$0.960 \pm 0.007(\text{sys})$	$0.961 \pm 0.007(\text{sys})$
$\boldsymbol{\varepsilon_{\gamma-\text{sel}}}$	$\mathbf{0.298 \pm 0.001(stat) \pm 0.032(sys)}$	$\mathbf{0.283 \pm 0.001(stat) \pm 0.031(sys)}$
$\varepsilon_{\text{newIsPhoton}}$	$0.865 \pm 0.014(\text{sys})$	$0.863 \pm 0.014(\text{sys})$
ε_{BDT}	$0.758 \pm 0.002(\text{stat}) \pm 0.08(\text{sys})$	$0.726 \pm 0.002(\text{stat}) \pm 0.08(\text{sys})$
$\varepsilon_{\phi-\text{sel}}$	0.8722	0.8722
$\varepsilon_{\text{unconverted}}$	$0.521 \pm 0.002(\text{sys})$	$0.517 \pm 0.002(\text{sys})$
$\boldsymbol{\varepsilon_{\text{trig}}}$	$\mathbf{0.95 \pm 0.04(stat)}$	$\mathbf{0.95 \pm 0.04(stat)}$
$\boldsymbol{\varepsilon_{\text{GEC}}}$	$\mathbf{0.936 \pm 0.026(stat)}$	$\mathbf{0.914 \pm 0.025(stat)}$
$\boldsymbol{\varepsilon_{\text{total}}}$	$\mathbf{0.249 \pm 0.013(stat) \pm 0.027(sys)}$	$\mathbf{0.231 \pm 0.012(stat) \pm 0.025(sys)}$

Table 3.8: Summary of the efficiencies for the signal classes and on the bottom the total efficiency.

$$\sigma = \frac{n \cdot f_{\text{signal}}}{\varepsilon_{\text{total}} \cdot \int \mathcal{L}\,\mathrm{d}t} \qquad \text{where} \qquad \varepsilon_{\text{total}} = \varepsilon_{\text{rec}} \cdot \varepsilon_{\gamma-\text{sel}} \cdot \varepsilon_{\text{trigger}} \cdot \varepsilon_{\text{GEC}} \tag{3.48}$$

and the uncertainties are added in quadrature for the statistical and systematic uncertainties separately. This results in

$$\begin{aligned} \sigma_{pp\to Z\gamma+X,\ \textit{hard signal}} &= 987 \pm 155(\text{stat}) \pm 112(\text{sys})\ \text{fb} \\ \sigma_{pp\to Z\gamma+X,\ \textit{radiative signal}} &= 2665 \pm 196(\text{stat}) \pm 301(\text{sys})\ \text{fb} \end{aligned} \tag{3.49}$$

for the two signal classes separately. The total production cross section of $Z\gamma$ is

$$\sigma_{pp\to Z\gamma+X} = 3652 \pm 351(\text{stat}) \pm 413(\text{sys})\ \text{fb}\,. \tag{3.50}$$

The theoretical predictions from Section 1.2.2 have an additional cut on the isolation between the photon and the muons of $\Delta R(\mu) \geq 0.4$. As this cut is not yet considered in the measurement a correction has to be applied. The correction factors are calculated from simulation for the

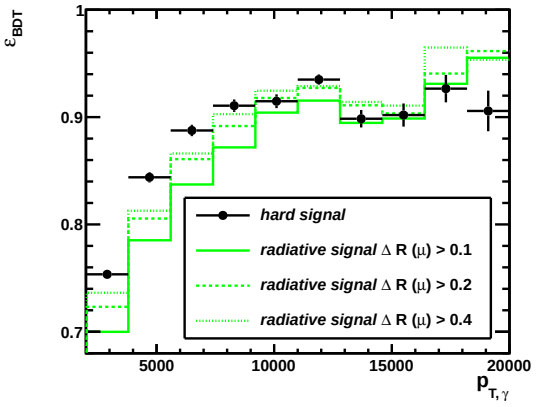

Figure 3.36: BDT efficiencies for the *radiative signal* for different photon isolation cuts.

two signal classes separately. Figure 3.1(d) shows that the *hard signal* is barely affected, while a large part of the *radiative signal* is eliminated with this cut. The results for the correction factors are

$$\begin{aligned} f_{\text{theory correction, } \textit{hard signal}} &= 0.9446 \pm 0.0015(\text{sys}) \\ f_{\text{theory correction, } \textit{radiative signal}} &= 0.5281 \pm 0.0032(\text{sys}) \ . \end{aligned} \tag{3.51}$$

The *radiative signal* seems to look similar to the *hard signal* for higher photon isolation cuts. The BDT efficiency, calculated from simulation and not yet applying the correction between data and simulation, raises from 0.806 ± 0.002 to 0.835 ± 0.003 (cf. Fig. 3.36) while the efficiency for the *hard signal* remains at 0.842 ± 0.002. Therefore the efficiencies from the *hard signal* are used for the *radiative signal* as well for the calculation of the cross section with the cut $\Delta R(\mu) \geq 0.4$.
The cross sections for the *hard signal* and the *radiative signal* are corrected with the factors in Eq. (3.51). The product of Eq. (3.49) and Eq. (3.51) is

$$\begin{aligned} \sigma_{pp\to Z\gamma+X,\ \textit{hard signal}} &= 932 \pm 146(\text{stat}) \pm 106(\text{sys})\ \text{fb} \\ \sigma_{pp\to Z\gamma+X,\ \textit{radiative signal}} &= 1301 \pm 95(\text{stat}) \pm 148(\text{sys})\ \text{fb} \end{aligned} \tag{3.52}$$

and the total measured production cross section of $Z\gamma$ corrected for the theory cuts is

$$\sigma_{pp\to Z\gamma+X} = 2233 \pm 201(\text{stat}) \pm 254(\text{sys})\ \text{fb} \ . \tag{3.53}$$

A comparison to the theoretical next to leading order prediction of

$$\sigma_{pp\to Z\gamma+X} = 2806.7 \pm 1.7\ \text{fb}\ [16] \tag{3.54}$$

shows a 1.3σ difference between the theory prediction and the experimental measurement. This is a reasonable agreement between the two values. The small difference could arise from overestimation of efficiencies as the fit seems to be very stable and should not bias the result significantly.

4 Conclusion and Outlook

The production cross section of $Z\gamma$ in proton proton collisions at the LHCb experiment is found to be

$$\sigma_{pp\to Z\gamma+X} = 3652 \pm 351(\text{stat}) \pm 413(\text{sys})\ \text{fb} \tag{4.1}$$

with the Z candidates reconstructed in the $\mu^+\mu^-$ decay channel. These muons must have a $p_T \geq 20\,\text{GeV/c}$ and η between 2 and 4.5. The photon is selected via a multivariate analysis and is required to have a $p_T \geq 2\,\text{GeV/c}$ and η between 1.8 and 4.3. Between the muons and the photon there must be an isolation of $\Delta R(\mu) \geq 0.1$.
Raising the photon isolation cut of $\Delta R(\mu) \geq 0.4$ reduces the production cross section to

$$\sigma_{pp\to Z\gamma+X} = 2233 \pm 201(\text{stat}) \pm 254(\text{sys})\ \text{fb} \tag{4.2}$$

due to the loss of the *radiative signal* where the photon is located close to the muons. The result is in reasonable agreement with the theory prediction.
For a better result, the efficiency of the BDT should be calculated completely from data, or the input variables of the BDT have to be reweighted from simulation to data in order to eliminate these differences. The differences from data and simulation in the variable newIsPhoton do not effect the efficiency as it is estimated from data and trigger- and reconstruction efficiencies for the muons are in good agreement with previous measurements.

Bibliography

[1] G. Aad *et al.* [ATLAS Collaboration], "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC," Phys. Lett. B **716** (2012) 1 [arXiv:1207.7214 [hep-ex]].

[2] R. Aaij *et al.* [LHCb Collaboration], "Measurement of $\sigma(pp \to b\bar{b}X)$ at $\sqrt{s} = 7$ TeV in the forward region," Phys. Lett. B **694** (2010) 209 [arXiv:1009.2731 [hep-ex]].

[3] R. Aaij *et al.* [LHCb Collaboration], "Strong constraints on the rare decays $B_s \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$," Phys. Rev. Lett. **108** (2012) 231801 [arXiv:1203.4493 [hep-ex]].

[4] R. Aaij *et al.* [LHCb Collaboration], "Measurement of the cross-section for $Z \to \mu\mu$ production with $1\,\mathrm{fb}^{-1}$ of pp collisions at $\sqrt{s} = 7\,\mathrm{TeV}$," LHCb-CONF-2013-007.

[5] S. Agostinelli *et al.* [GEANT4 Collaboration], "GEANT4: A Simulation toolkit," Nucl. Instrum. Meth. A **506** (2003) 250.

[6] M. P. Altarelli, "LHCb: detector performance and first physics results," arXiv:1105.5330 [hep-ex].

[7] A. A. Alves, Jr. *et al.* [LHCb Collaboration], "The LHCb Detector at the LHC," JINST **3** (2008) S08005.

[8] J. Beringer et al. (Particle Data Group), Phys. Rev. D86, 010001 (2012) and 2013 partial update for the 2014 edition.

[9] S. Bifani, LHCb-ANA-2014-049: Measurement of the W production cross-section in pp collisions at $\sqrt{s} = 7$ TeV with the 2011 data, Apr, 2014.

[10] R. Brun and F. Rademakers, "ROOT: An object oriented data analysis framework," Nucl. Instrum. Meth. A **389** (1997) 81.

[11] J. M. Campbell, R. K. Ellis and C. Williams, "Vector boson pair production at the LHC," JHEP **1107** (2011) 018 [arXiv:1105.0020 [hep-ph]].

[12] S. Chatrchyan *et al.* [CMS Collaboration], "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC," Phys. Lett. B **716** (2012) 30 [arXiv:1207.7235 [hep-ex]].

[13] M. De Cian *et al.*, "Measurement of the track finding efficiency," Tech. Rep. LHCb-PUB-2011-025. CERN-LHCb-PUB-2011-025, CERN, Geneva, Apr, 2012.

[14] Ch. Elsasser, "Determination of the Invariant Mass Distribution for $B^0_{(s)} \to \mu^+\mu^-$ at the LHCb Experiment," (2011).

[15] P. Golonka and Z. Was, "PHOTOS Monte Carlo: A Precision tool for QED corrections in Z and W decays," Eur. Phys. J. C **45** (2006) 97 [hep-ph/0506026].

[16] M. Grazzini, S. Kallweit, D. Rathlev and A. Torre, "$Z\gamma$ production at hadron colliders in NNLO QCD," Phys. Lett. B **731** (2014) 204 [arXiv:1309.7000 [hep-ph]].

[17] A. Hocker, J. Stelzer, F. Tegenfeldt, H. Voss, K. Voss, A. Christov, S. Henrot-Versille and M. Jachowski *et al.*, "TMVA - Toolkit for Multivariate Data Analysis," PoS ACAT (2007) 040 [physics/0703039 [PHYSICS]].

[18] E. Jans [LHCb Collaboration], "The LHCb detector," arXiv:0910.1740 [hep-ex].

[19] D. J. Lange, "The EvtGen particle decay simulation package," Nucl. Instrum. Meth. A **462** (2001) 152.

[20] [LHCb Collaboration], "A measurement of high-p_T muon reconstruction efficiencies in 2011 and 2012 data," LHCb-INT-2014-030, CERN, Geneva, Jun, 2014.

[21] [LHCb Collaboration], "Measurement of the track reconstruction efficiency at LHCb," LHCB-CERN-LHCB-DP-2013-002-001, CERN, Geneva, Jun, 2014.

[22] [LHCb Collaboration], "π^0/γ separation," https://twiki.cern.ch/twiki/bin/viewauth /LHCbPhysics/CalorimeterObjectsToolsGroupDOC#gamma_separation , Jan, 2014.

[23] J.-F. Marchand, private comunication.

[24] M. Pivk and F. R. Le Diberder, "SPlot: A Statistical tool to unfold data distributions," Nucl. Instrum. Meth. A **555** (2005) 356 [physics/0402083 [physics.data-an]].

[25] T. Sjöstrand, S. Mrenna, and P. Skands, A brief introduction to PYTHIA 8.1, Computer Physics Communications 178 (2008), no. 11 852

[26] W. Verkerke and D. P. Kirkby, "The RooFit toolkit for data modeling," eConf C **0303241** (2003) MOLT007 [physics/0306116].