

A MEASUREMENT OF THE  $W\gamma$  CROSS SECTION AT  $\sqrt{S} = 8$  TEV IN  $PP$   
COLLISIONS WITH THE CMS DETECTOR

by

Ekaterina Avdeeva

A DISSERTATION

Presented to the Faculty of  
The Graduate College at the University of Nebraska  
In Partial Fulfilment of Requirements  
For the Degree of Doctor of Philosophy

Major: Physics & Astronomy

Under the Supervision of Professor Ilya Kravchenko

Lincoln, Nebraska

May, 2017

A MEASUREMENT OF THE  $W\gamma$  CROSS SECTION AT  $\sqrt{s} = 8$  TEV IN  $PP$   
COLLISIONS WITH THE CMS DETECTOR

Ekaterina Avdeeva, Ph.D.

University of Nebraska, 2017

Adviser: Ilya Kravchenko

The measurement of  $W\gamma \rightarrow l\nu\gamma$  total and differential cross section is reported. The measurement was performed with CMS at LHC center-of-mass energy of  $\sqrt{s} = 8$  TeV. The electron and the muon channels are considered. This is the first measurement of the  $W\gamma \rightarrow l\nu\gamma$  cross section at CMS. No deviation from the Standard Model has been observed.

## ACKNOWLEDGMENTS

This dissertation represents a work that required effort from a lot of different people in many different ways. The list of people includes but is not limited to CMS and UNL colleagues from all over the world, family and friends.

Specifically, first, I would like to thank my advisor Ilya Kravchenko for his help and support throughout my whole experience as a PhD student at UNL, from the admission process and to the process of writing this dissertation. We discussed every single step of the  $W\gamma$  measurements, and I could always rely on getting a guidance and valuable advice whether it is an issue in a measurement step, technical question or effective collaboration with CMS colleagues. Under Ilya's supervision, I raised from a first year graduate student who can only complete certain very specific tasks to a researcher who can complete her own project and get it endorsed for the PhD defense. Overall, the list of CMS colleagues who made direct contribution to the  $W\gamma$  measurement and/or provided a good number meaningful recommendations starts from Yurii Maravin, Lovedeep Saini, Joshua Kunkle, Senka Duric, Kuo Chia-Ming, Sachiko Toda McBride, Yutaro Iiyama, and never ends.

Additionally, I would like to express my deep gratitude Frank Meier for supervising another part of my work, the alignment of the CMS tracking system and to make my contribution to the performance of the CMS experiment. We spent hours discussing alignment algorithms, various technical tools, and results. Frank helped me to get involved in the project, get trained and tremendously improve in the level of understanding and capabilities. My work in the tracker alignment group was a very integrated team work, where it is very difficult to separate your contribution from the contributions of others. The list of colleagues who made our

collaboration so fruitful, besides Frank, includes Marco Musich, Matthias Schröder, Tapio Lampen, Patrick Connor, Nazar Stephaniuk, Candice You and many others.

I would like to thank UNL professors Ken Bloom, Brad Schadwick and Glenn Ledder for their time spent on serving as my graduate committee and providing set of questions and comments regarding the dissertation. Ken Bloom also, alongside with other UNLHEP professors, postdocs and students, participated in tricky discussions regarding various aspects of my research work during our group meetings as well as during individual meetings and over e-mails.

Finally, a special thanks to my parents Elena Tsukanova and Grigory Tsukanov, to my husband Sergey Avdeev, and to my daughters Oksana and Yana Avdeeva for being a source of support in my life and research career, for their constant patience and understanding. Family support has meant a lot to me, and these people's effect on any aspect of my personality cannot be overestimated.

# Contents

<b>Contents</b>	v
<b>List of Figures</b>	ix
<b>List of Tables</b>	xxiii
<b>1 Introduction</b>	1
1.1 Fundamental Particles and Interactions . . . . .	4
1.2 Electroweak Interactions . . . . .	7
1.3 Strong Interactions . . . . .	12
1.4 Physics of Proton-Proton Collisions . . . . .	14
1.5 Open Questions of the Standard Model . . . . .	16
<b>2 <math>W\gamma</math> Production Theory and Former Experimental Results</b>	18
2.1 Electroweak Theory of the Standard Model . . . . .	19
2.2 Cross Section and Luminosity . . . . .	28
2.3 Standard Model $W\gamma$ Production . . . . .	31
2.4 Anomalous $W\gamma$ Production . . . . .	36
2.5 A brief history of $W\gamma$ measurements . . . . .	40
<b>3 Experimental Setup</b>	45

3.1	Large Hadron Collider . . . . .	46
3.2	Compact Muon Solenoid . . . . .	52
3.2.1	Introduction . . . . .	52
3.2.2	Magnet . . . . .	57
3.2.3	Tracking System . . . . .	57
3.2.4	Electromagnetic Calorimeter . . . . .	59
3.2.5	Hadron Calorimeter . . . . .	60
3.2.6	Muon System . . . . .	61
3.2.7	Triggering and Data Acquisition . . . . .	62
3.2.8	Particle Flow Algorithm of Event Reconstruction . . . . .	64
<b>4</b>	<b>CMS Tracker Alignment</b>	<b>66</b>
4.1	Approach . . . . .	68
4.2	Selected Results on Alignment of the Tracking System with 2015 Data	72
4.2.1	Geometry Comparison . . . . .	73
4.2.2	Distributions of Medians of Unbiased Track-Hit Residuals . .	76
4.2.3	Cosmic Track Splitting Validation . . . . .	78
4.2.4	Primary Vertex Validation . . . . .	79
<b>5</b>	<b><math>W\gamma</math> Cross Section Measurement</b>	<b>81</b>
5.1	Measurement Strategy . . . . .	83
5.2	Data and Monte Carlo Samples . . . . .	85
5.3	Event and Object Selection . . . . .	90
5.3.1	Object Selection . . . . .	90
5.3.2	Event Level Selection . . . . .	92
5.3.3	Selected Events . . . . .	94
5.4	Background Estimation and Subtraction . . . . .	98

5.4.1	Background from Jets Faking Photons . . . . .	98
5.4.2	Background from Electrons Faking Photons in the Electron Channel . . . . .	101
5.4.3	Other Backgrounds . . . . .	103
5.4.4	$P_T^\gamma$ Spectra before and after the Background Subtraction . . .	104
5.5	Detector Resolution Unfolding . . . . .	110
5.6	Acceptance and Efficiency Correction . . . . .	116
5.7	Systematic Uncertainties . . . . .	118
5.7.1	Uncertainties Related to Jets $\rightarrow \gamma$ Background Estimation . .	118
5.7.2	Other Sources of the Systematic Uncertainties . . . . .	121
5.7.3	Summary of the Systematic Uncertainties . . . . .	123
5.8	Cross Section . . . . .	126
<b>6</b>	<b>Summary and Conclusions</b>	<b>130</b>
<b>A</b>	<b>Efficiency Scale Factors</b>	<b>131</b>
<b>B</b>	<b><math>Z\gamma</math> FSR and ISR Plots</b>	<b>134</b>
<b>C</b>	<b>Studies of <math>E_T^{miss}</math> Dependence of Templates</b>	<b>143</b>
<b>D</b>	<b>Template Fit Plots, <math>W\gamma</math>, Data</b>	<b>149</b>
<b>E</b>	<b>MC Closure Check</b>	<b>164</b>
<b>F</b>	<b>Fit Plots of <math>M_{e\gamma}</math></b>	<b>173</b>
<b>G</b>	<b>Fit Plots of <math>M_{e\gamma}</math> without Requirement on <math>M_T^W</math></b>	<b>184</b>
<b>H</b>	<b>Tables for <math>e \rightarrow \gamma</math> Background Estimation</b>	<b>195</b>

<b>I Data vs MC Plots of <math>M_{e\gamma}</math></b>	<b>198</b>
<b>J Correlation Matrices for Different Sources of the Systematic Uncertainties</b>	<b>205</b>
<b>K Code and Software</b>	<b>212</b>
<b>L <math>Z\gamma</math> Check</b>	<b>214</b>
<b>Bibliography</b>	<b>223</b>

# List of Figures

1.1	Standard Model Particles. . . . .	6
1.2	Electromagnetic interactions. Left: a photon radiation off a charged fermion, middle and right: Bhabha scattering. . . . .	7
1.3	Weak elementary processes and gauge couplings. Top left: a quark with charge $Q=+2/3$ enters, emits a $W$ boson, and a quark with charge $Q=-1/3$ escapes. Top middle: a charged lepton enters, emits a $W$ boson, and a neutrino or antineutrino escapes conserving a lepton flavor number. Top right: a fermion enters, emits a $Z$ boson and escapes. Bottom left: TGC couplings $WW\gamma$ and $WWZ$ . Bottom middle: QGC couplings $WW\gamma\gamma$ , $WWZ\gamma$ and $WWZZ$ . Bottom right: QGC coupling $WWWW$ . . . . .	8
1.4	The Higgs boson production and decay. Left: $H \rightarrow \gamma\gamma$ , right: $H \rightarrow ZZ \rightarrow 4l$ . . . . .	10
1.5	Elementary processes of strong interations . . . . .	12
1.6	The proton structure (left) and the proton-proton collision (right). . . . .	14
1.7	Parton distribution functions [1]. . . . .	15
2.1	Illustration of the differential cross section concept in the classical case. . . . .	29

2.2	Feynman diagrams of $W\gamma$ production. Top: LO diagrams, bottom: several examples of NLO in QCD. . . . .	32
2.3	Theory spectra. Top: NLO and NNLO $p_T^\gamma$ spectra of $W\gamma \rightarrow l\nu\gamma$ at $\sqrt{s} = 7$ TeV overlaid with ATLAS data for $N_{jet} \geq 0$ (left) and $N_{jet} = 0$ (right). Middle: LO, NLO and NNLO $m_T^{l\nu\gamma}$ spectra of $W\gamma \rightarrow l\nu\gamma$ at $\sqrt{s} = 7$ TeV for $P_T^\gamma > 15$ GeV (left) and $P_T^\gamma > 40$ GeV (right). Bottom: LO, NLO and NNLO $\Delta_{l\gamma}$ spectra of $W\gamma \rightarrow l\nu\gamma$ at $\sqrt{s} = 7$ TeV. . . . .	35
2.4	Charged TGC (first), neutral TGC (second), charged QGC (third and fourth), and neutral QGC (fifth) vertices. . . . .	36
2.5	Distributions of $P_T^\gamma$ in simulated $W\gamma \rightarrow \mu\nu\gamma$ events with different values of aTGC constants at LHC energy of $\sqrt{s} = 7$ TeV. Source of figure: [2]. . .	39
2.6	Summary of limits on the $WW\gamma$ aTGC coupling constants. Figure from [3].	40
2.7	The distribution of the $p_T^\gamma$ of $W\gamma$ candidates in the analysis of 7 TeV CMS data. Data vs signal simulation + background estimates. Left: $W\gamma \rightarrow e\nu\gamma$ , right: $W\gamma \rightarrow \mu\nu\gamma$ [4]. . . . .	42
2.8	The distribution of the photon transverse momentum (left) and missing transverse momentum (right) of $W\gamma$ candidates in the analysis of 7 TeV ATLAS data. Data vs signal simulation + background estimates [5]. . .	42
3.1	CERN's accelerator complex [6]. . . . .	47
3.2	Cross sections of different processes in $pp$ and $\bar{p}p$ collisions [7]. . . . .	49
3.3	LHC integrated luminosity by year [8]. . . . .	51
3.4	CMS coordinate system. . . . .	53

3.5 CMS detector, schematic view. Top: $r - z$ plane, bottom: $r - \phi$ plane at $z = 0$ [9]. The tracking system is shown in green, EB and EE of ECal are shown in gray, HB, HE and HF of HCal are shown in yellow. HCal is surrounded by a magnet which is shown in gray and white. Muon stations and return yokes are located outside of the magnet and are shown in blue and gray. A red line at the bottom plot is a muon trajectory demonstrating a typical muon to penetrate the whole CMS detector. People at the bottom illustrate the scale of the CMS detector. . . . .	55
3.6 CMS detector, a schematic view of a segment in the $r\phi$ plane at $z = 0$ . Traces left by muons, electrons, photons, charged and neutral hadrons in different subdetectors are shown. . . . .	56
3.7 Slice of the CMS tracking system in the $r-z$ plane. Pixel modules and strip modules shown in blue provide 3D position measurements. Strip modules shown in pink provide 2D position measurements. . . . .	59
3.8 Components of the CMS muon system. Left to right: drift tubes (DTs), cathode strip chambers (CSCs), resistive plate chambers (RPCs). . . . .	62

4.1	The alignment of a toy tracker, part 1. When a charged particle passes through a detector (top left), it crosses a toy tracker which consists of six flat equidistant modules (top right). If the modules were placed exactly at their design positions, we would observe the hits exactly at the points where the track crosses modules of ideal geometry (middle left). However, in reality, the positions and tilts of the modules are different from ones suggested by the ideal geometry (middle right). Hits, indeed, are recorded at the places where modules are mounted, not at the design ideal places (bottom left). If we assumed a tracker to be ideal and a track to be smooth, we would see that our hits are off-track (bottom right). Image by Frank Meier. . . . .	69
4.2	The alignment of a toy tracker, part 2. We record a large number of tracks and take into account them all to determine the alignment parameters by minimizing residuals between measured and predicted hits. Image by Frank Meier. . . . .	70
4.3	Local alignment parameters [10]. . . . .	70
4.4	Surface deformations [11]. . . . .	71
4.5	Comparison of Run II and Run I positions of the modules in the FPIX of the CMS tracking system. Positions are determined with the Millepede-II and HIP algorithms using cosmic ray data collected with the magnetic field of $B = 0\text{T}$ and $B = 3.8\text{T}$ magnetic field in the CMS solenoid. The difference $\Delta z$ (Run II - Run I) is plotted as a function of $z$ (left) and $\phi$ (right) in global coordinates. The plot shows the displacements of two pixel half disks by 4.5 and 5.5 mm. . . . .	74

4.6 Three-dimensional geometry comparison of Run II and Run I positions in the BPIX and FPIX of the CMS tracking system. Positions are determined with the Millepede-II and HIP algorithms using cosmic ray data collected with the magnetic field of $B = 0\text{T}$ and $B = 3.8\text{T}$ magnetic field in the CMS solenoid and collision data with $B = 0\text{T}$ at $\sqrt{s} = 13\text{ TeV}$ . The positions at the end of Run I are shown in gray. The module displacements between Run I and Run II are magnified by a factor of 5 for visualization purpose. The resulting positions are shown in red, yellow, or green, depending on the displacement magnitude. . . . .	75
4.7 DMRs for the local $x$ -direction (left) and for the local $y$ -direction (right) in the BPIX of the CMS tracking system, using 2 million cosmic ray tracks collected with the magnetic field of $B = 3.8\text{T}$ . The blue line shows the Run I geometry. The green line shows the alignment produced with the Millepede-II and HIP algorithms using cosmic ray data at $B = 0\text{T}$ and $B = 3.8\text{T}$ . The aligned geometry shows reasonable performance. . .	77
4.8 Results of the cosmic track splitting validation. The normalized differences between two parts of a cosmic track in the $xy$ distance between the track and the origin ( $d_{xy}$ , left), and in the distance in the $z$ direction between the track and the origin ( $d_z$ , right). Alignment is produced with the Millepede-II and HIP algorithms using cosmic ray data at the magnetic field of $B = 0\text{T}$ and $B = 3.8\text{T}$ of CMS solenoid. Geometry aligned with Run II data is shown in green, Run I geometry is shown in blue. Aligned geometry shows reasonable performance. . . . .	78

4.9	The results of the primary vertex validation. The distance of the track at its closest approach to a refit unbiased primary vertex ( $d_{xy}$ , left and $d_z$ , right) in the transverse plane. The validation is produced with $B = 0\text{T}$ collision data. The alignment is produced with the Millepede-II and HIP algorithms using $B = 0\text{T}$ and $B = 3.8\text{T}$ cosmic ray data and $B = 0\text{T}$ collision data. . . . .	79
5.1	Number of vertices of $Z\gamma$ candidates in the muon channel. Data vs MC. Left: no PU reweighting applied, right: PU reweighting applied. Ratio plot in the bottom shows data yields divided over total MC yields. EB+EE means that events with a final state photon reconstructed in the ECal barrel as well as events with a final state photon reconstructed in the ECal endcap are shown on the plots. . . . .	89
5.2	$M_T^W$ distribution of $W\gamma$ candidates. Data vs total MC. Left: muon channel, right: electron channel. All selection criteria except $M_T^W$ requirement are applied on all samples that are present on the plot. The $P_T^\gamma$ range where we do not expect any new physics is used: 15-45 GeV. The ratio plot is data divided by total MC. . . . .	95
5.3	$M_{l\gamma}$ distribution of $W\gamma$ candidates in the electron channel. Data vs total MC. All selection criteria except Z-mass window are applied on all samples that are present on the plot. The $P_T^\gamma$ range where we do not expect any new physics is used: 15-45 GeV. The ratio plot is data divided by total MC. . . . .	96
5.4	$P_T^\gamma$ distribution of $W\gamma$ candidates. Data vs total MC. Left column: muon channel, right column: electron channel. EB (top) and EE (bottom) photons. The ratio plots are data divided by total MC. . . . .	97

5.5 Top and middle: data vs fake- $\gamma$ background derived from the template method + real- $\gamma$ background predicted by dedicated MC samples + signal MC, with $I_{ch}$ (left) and $\sigma_{i\eta i\eta}$ (right) used as fit variables in EB (top) and EE (middle). Bottom: data yields after full background subtraction vs signal MC in EB (left) and EE (right). $I_{ch}$ vs $\sigma_{i\eta i\eta}$ fit results. Muon channel. . . . .	106
5.6 Top and middle: data vs fake- $\gamma$ background derived from the template method + real- $\gamma$ background predicted by dedicated MC samples + signal MC, with $I_{ch}$ (left) and $\sigma_{i\eta i\eta}$ (right) used as fit variables in EB (top) and EE (middle). Bottom: data yields after full background subtraction vs signal MC in EB (left) and EE (right). $I_{ch}$ vs $\sigma_{i\eta i\eta}$ fit results. Electron channel. . . . .	107
5.7 Response matrix derived from the signal MC. . . . .	113
5.8 Correlation matrices of statistical uncertainties on unfolded $W\gamma$ yields in the muon (top) and electron (bottom) channels. . . . .	114
5.9 $A \times \epsilon$ corrections in the muon (left) and electron (right) channels. Plots are produced with $W\gamma$ MC sample at $\sqrt{s} = 8$ TeV. . . . .	117
5.10 $W\gamma$ differential cross section. Top, left: the $W\gamma$ differential cross section; top, right: the ratio of measured over the MC-based $W\gamma$ differential cross section. Bottom: the $W\gamma$ measured differential cross section overlaid with the MC-based cross section in the muon channel (left) and in the electron channel (right). . . . .	129

B.1 Distributions of $M_{\mu\mu\gamma}$ (left) and $M_{\mu\mu}$ (right) in $Z\gamma \rightarrow \mu\mu\gamma$ -selected events, data vs MC. $P_T^\gamma : 15\text{-}500 \text{ GeV}$ . Left: $M_{\mu\mu\gamma}$ , right: $M_{\mu\mu}$ . Top: barrel photons, bottom: endcap photons. Peak highly dominated by $Z\gamma$ events corresponds to FSR.	135
B.2 Distributions of $\Delta R(\mu_1, \gamma)$ (left) and $\Delta R(\mu_2, \gamma)$ (right) in $Z\gamma \rightarrow \mu\mu\gamma$ -selected events, data vs MC. $P_T^\gamma : 15\text{-}500 \text{ GeV}$ . Left: $M_{\mu\mu\gamma}$ , right: $M_{\mu\mu}$ . Top: barrel photons, bottom: endcap. Peak highly dominated by $Z\gamma$ events corresponds to FSR.	136
B.3 $Z\gamma$ -selected FSR (left) and ISR (right) events, data vs MC.	137
B.4 $Z\gamma$ -selected FSR events, data vs MC. $P_T^\gamma > 15 \text{ GeV}$ . Distributions of $I_{chHad}^\gamma$ used for preparing real- $\gamma$ templates. Fake- $\gamma$ contribution to FSR region is subtracted based on DY+jets MC prediction to prepare real- $\gamma$ templates.	138
B.5 $Z\gamma$ -selected ISR events, data vs MC. $10 \text{ GeV} < P_T^\gamma < 15 \text{ GeV}$ . Distributions of $I_{chHad}^\gamma$ used for preparing fake- $\gamma$ templates. Real- $\gamma$ contribution to ISR region is subtracted based on $Z\gamma$ signal MC prediction to prepare fake- $\gamma$ templates.	138
B.6 $Z\gamma$ -selected ISR events, data vs MC. Distributions of $I_{chHad}^\gamma$ used for preparing fake- $\gamma$ templates. Real- $\gamma$ contribution to ISR region is subtracted based on $Z\gamma$ signal MC prediction to prepare fake- $\gamma$ templates. Ranges of $< P_T^\gamma$ are shown in the plot titles and cover the total range of $15 \text{ GeV} < P_T^\gamma < 500 \text{ GeV}$ . EB photons.	139

B.7	$Z\gamma$ -selected ISR events, data vs MC. Distributions of $I_{chHad}^\gamma$ used for preparing fake- $\gamma$ templates. Real- $\gamma$ contribution to ISR region is subtracted based on $Z\gamma$ signal MC prediction to prepare fake- $\gamma$ templates. Ranges of $P_T^\gamma$ are shown in the plot titles and cover the total range of $15 \text{ GeV} < P_T^\gamma < 500 \text{ GeV}$ . EE photons. . . . .	140
B.8	$Z\gamma$ -selected FSR events, data vs MC. Distributions of $\sigma_{i\eta i\eta}^\gamma$ are used for preparing real- $\gamma$ templates. Fake- $\gamma$ contribution to FSR region is subtracted based on DY+jets MC prediction to prepare real- $\gamma$ templates. The templates are prepared separately for barrel and endcap photons. . .	141
B.9	$Z\gamma$ -selected ISR events, data vs MC. Distributions of $\sigma_{i\eta i\eta}^\gamma$ are used for preparing real- $\gamma$ templates. Fake- $\gamma$ contribution to ISR region is subtracted based on DY+jets MC prediction to prepare real- $\gamma$ templates. The templates are prepared separately for barrel and endcap photons. . .	142
C.1	$Z\gamma$ ISR-selected data (includes DY+jets and $Z\gamma$ events), $I_{ch}^\gamma$ (top) and $\sigma_{i\eta i\eta}^\gamma$ (bottom) fake- $\gamma$ templates with and without $E_T^{miss}$ requirement. . .	144
C.2	$W\gamma$ MC, muon channel, $I_{ch}^\gamma$ templates. . . . .	145
C.3	$W$ +jets MC, muon channel, $I_{ch}^\gamma$ templates. . . . .	146
C.4	$W\gamma$ MC, muon channel, $\sigma_{i\eta i\eta}^\gamma$ templates. . . . .	147
C.5	$W$ +jets MC, muon channel, $\sigma_{i\eta i\eta}^\gamma$ templates. . . . .	148
D.1	Fits of $I_{ch}^\gamma$ templates, $W\gamma$ , muon channel, underflow bin ( $10 - 15 \text{ GeV}$ ). .	149
D.2	Fits of $I_{ch}^\gamma$ templates, $W\gamma$ , electron channel, underflow bin ( $10 - 15 \text{ GeV}$ ). .	150
D.3	Fits of $I_{ch}^\gamma$ templates, $W\gamma$ , muon channel. . . . .	151
D.4	Fits of $I_{ch}^\gamma$ templates, $W\gamma$ , muon channel. . . . .	152
D.5	Fits of $I_{ch}^\gamma$ templates, $W\gamma$ , muon channel. . . . .	153
D.6	Fits of $I_{ch}^\gamma$ templates, $W\gamma$ , electron channel. . . . .	154

D.7	Fits of $I_{ch}^{\gamma}$ templates, $W\gamma$ , electron channel. . . . .	155
D.8	Fits of $I_{ch}^{\gamma}$ templates, $W\gamma$ , electron channel. . . . .	156
D.9	Fits of $\sigma_{i\eta i\eta}$ templates, $W\gamma$ , muon channel, underflow bin (10 – 15 GeV). . . . .	157
D.10	Fits of $\sigma_{i\eta i\eta}$ templates, $W\gamma$ , electron channel, underflow bin (10 – 15 GeV). . . . .	157
D.11	Fits of $\sigma_{i\eta i\eta}$ templates, $W\gamma$ , muon channel. . . . .	158
D.12	Fits of $\sigma_{i\eta i\eta}$ templates, $W\gamma$ , muon channel. . . . .	159
D.13	Fits of $\sigma_{i\eta i\eta}$ templates, $W\gamma$ , muon channel. . . . .	160
D.14	Fits of $\sigma_{i\eta i\eta}$ templates, $W\gamma$ , electron channel. . . . .	161
D.15	Fits of $\sigma_{i\eta i\eta}$ templates, $W\gamma$ , electron channel. . . . .	162
D.16	Fits of $\sigma_{i\eta i\eta}$ templates, $W\gamma$ , electron channel. . . . .	163
E.1	Real- $\gamma$ yields derived from fits of pseudodata superimposed with $W\gamma$ MC. Left column - muon channel, right column - electron channel. Top to bottom: barrel and endcap photons. . . . .	165
E.2	$W+jets$ yields derived from fits of pseudodata superimposed with $W+jets$ MC. Left column - muon channel, right column - electron. Top to bottom: barrel and endcap photons. . . . .	166
E.3	Data (left) and pseudodata (right) vs background estimates and signal MC in bins of $P_T^{\gamma}$ . Jets $\rightarrow \gamma$ background estimated from fits of $I_{ch}^{\gamma}$ (top) and $\sigma_{i\eta i\eta}^{\gamma}$ (bottom). Muon channel. Barrel photons. . . . .	169
E.4	Data (left) and pseudodata (right) vs background estimates and signal MC in bins of $P_T^{\gamma}$ . Jets $\rightarrow \gamma$ background estimated from fits of $I_{ch}^{\gamma}$ (middle) and $\sigma_{i\eta i\eta}^{\gamma}$ (bottom). Muon channel. Endcap photons. . . . .	170
E.5	Data (left) and pseudodata (right) vs background estimates and signal MC in bins of $P_T^{\gamma}$ . Jets $\rightarrow \gamma$ background estimated from fits of $I_{ch}^{\gamma}$ (middle) and $\sigma_{i\eta i\eta}^{\gamma}$ (bottom). Electron channel. Barrel photons. . . . .	171

E.6 Data (left) and pseudodata (right) vs background estimates and signal MC in bins of $P_T^\gamma$ . Jets $\rightarrow \gamma$ background estimated from fits of $I_{ch}^\gamma$ (middle) and $\sigma_{inj\eta}^\gamma$ (bottom). Electron channel. Endcap photons.	172
F.1 $M_{e\gamma}$ fits, $W\gamma$ , electron channel, 15-20 GeV, 8 $\eta^\gamma$ bins.	174
F.2 $M_{e\gamma}$ fits, $W\gamma$ , electron channel, 20-25 GeV, 8 $\eta^\gamma$ bins.	175
F.3 $M_{e\gamma}$ fits, $W\gamma$ , electron channel, 25-30 GeV, 8 $\eta^\gamma$ bins.	176
F.4 $M_{e\gamma}$ fits, $W\gamma$ , electron channel, 30-35 GeV, 8 $\eta^\gamma$ bins.	177
F.5 $M_{e\gamma}$ fits, $W\gamma$ , electron channel, 35-45 GeV, 8 $\eta^\gamma$ bins.	178
F.6 $M_{e\gamma}$ fits, $W\gamma$ , electron channel, 45-55 GeV, 8 $\eta^\gamma$ bins.	179
F.7 $M_{e\gamma}$ fits, $W\gamma$ , electron channel, 55-65 GeV, 4 $\eta^\gamma$ bins.	180
F.8 $M_{e\gamma}$ fits, $W\gamma$ , electron channel, 65-75 GeV, 4 $\eta^\gamma$ bins.	181
F.9 $M_{e\gamma}$ fits, $W\gamma$ , electron channel, 75-85 GeV, 4 $\eta^\gamma$ bins.	182
F.10 $M_{e\gamma}$ fits, $W\gamma$ , electron channel, 85-95 GeV, 2 $\eta^\gamma$ bins.	182
F.11 $M_{e\gamma}$ fits, $W\gamma$ , electron channel, 95-120 GeV, 2 $\eta^\gamma$ bins.	183
F.12 $M_{e\gamma}$ fits, $W\gamma$ , electron channel, 120-500 GeV, 2 $\eta^\gamma$ bins.	183
G.1 $M_{e\gamma}$ fits, $W\gamma$ , electron channel, underflow bin (15-20 GeV), 8 $\eta^\gamma$ bins.	185
G.2 $M_{e\gamma}$ fits, $W\gamma$ , electron channel, underflow bin (20-25 GeV), 8 $\eta^\gamma$ bins.	186
G.3 $M_{e\gamma}$ fits, $W\gamma$ , electron channel, underflow bin (25-30 GeV), 8 $\eta^\gamma$ bins.	187
G.4 $M_{e\gamma}$ fits, $W\gamma$ , electron channel, underflow bin (30-35 GeV), 8 $\eta^\gamma$ bins.	188
G.5 $M_{e\gamma}$ fits, $W\gamma$ , electron channel, underflow bin (35-45 GeV), 8 $\eta^\gamma$ bins.	189
G.6 $M_{e\gamma}$ fits, $W\gamma$ , electron channel, underflow bin (45-55 GeV), 8 $\eta^\gamma$ bins.	190
G.7 $M_{e\gamma}$ fits, $W\gamma$ , electron channel, underflow bin (55-65 GeV), 4 $\eta^\gamma$ bins.	191
G.8 $M_{e\gamma}$ fits, $W\gamma$ , electron channel, underflow bin (65-75 GeV), 4 $\eta^\gamma$ bins.	192
G.9 $M_{e\gamma}$ fits, $W\gamma$ , electron channel, underflow bin (75-85 GeV), 4 $\eta^\gamma$ bins.	193
G.10 $M_{e\gamma}$ fits, $W\gamma$ , electron channel, underflow bin (85-95 GeV), 2 $\eta^\gamma$ bins.	193

G.11	$M_{e\gamma}$ fits, $W\gamma$ , electron channel, underflow bin (95–120 GeV), 2 $\eta^\gamma$ bins.	. . . . . 194
G.12	$M_{e\gamma}$ fits, $W\gamma$ , electron channel, underflow bin (120–500 GeV), 2 $\eta^\gamma$ bins.	194
I.1	$M_{e\gamma}$ distribution, data vs MC. Bins 15 – 20 – 25 GeV. Left: all MC samples are normalized to luminosity of data, PU weight and SFs, right: DY+jets( $e \rightarrow \gamma$ ) also normalized to $e \rightarrow \gamma$ background estimates. . . . .	199
I.2	$M_{e\gamma}$ distribution, data vs MC. Bins 25 – 30 – 35 GeV. Left: all MC samples are normalized to luminosity of data, PU weight and SFs, right: DY+jets( $e \rightarrow \gamma$ ) also normalized to $e \rightarrow \gamma$ background estimates. . . . .	200
I.3	$M_{e\gamma}$ distribution, data vs MC. Bins 35 – 45 – 55 GeV. Left: all MC samples are normalized to luminosity of data, PU weight and SFs, right: DY+jets( $e \rightarrow \gamma$ ) also normalized to $e \rightarrow \gamma$ background estimates. . . . .	201
I.4	$M_{e\gamma}$ distribution, data vs MC. Bins 55 – 65 – 75 GeV. Left: all MC samples are normalized to luminosity of data, PU weight and SFs, right: DY+jets( $e \rightarrow \gamma$ ) also normalized to $e \rightarrow \gamma$ background estimates. . . . .	202
I.5	$M_{e\gamma}$ distribution, data vs MC. Bins 75 – 85 – 95 GeV. Left: all MC samples are normalized to luminosity of data, PU weight and SFs, right: DY+jets( $e \rightarrow \gamma$ ) also normalized to $e \rightarrow \gamma$ background estimates. . . . .	203
I.6	$M_{e\gamma}$ distribution, data vs MC. Bins 95 – 120 – 500 GeV. Left: all MC samples are normalized to luminosity of data, PU weight and SFs, right: DY+jets( $e \rightarrow \gamma$ ) also normalized to $e \rightarrow \gamma$ background estimates. . . . .	204
J.1	Correlation Matrices for systematic error due to the difference between $I_{ch}^\gamma$ and $\sigma_{i\eta i\eta}^\gamma$ fit results. . . . .	206
J.2	Correlation Matrices for systematic error due to uncertainty on the $Z\gamma$ MC sample normalization. . . . .	207

J.3	Correlation Matrices for systematic error due to the template statistical power. . . . .	208
J.4	Correlation Matrices for systematic error due to real- $\gamma$ background subtraction. . . . .	209
J.5	Correlation Matrices for systematic error due to signal MC statistics for unfolding. . . . .	210
J.6	Correlation Matrix for systematic error due to statistics of different samples for $e \rightarrow \gamma$ background estimation. . . . .	211
J.7	Correlation Matrix for systematic error due to fit bias for $e \rightarrow \gamma$ background estimation. . . . .	211
L.1	Data vs MC plots. Left column - muon channel, right column - electron channel. Top to bottom: barrel and endcap photons . . . . .	215
L.2	$Z\gamma$ check. Muon channel. Top and middle: data vs fake- $\gamma$ background derived from the template method + real- $\gamma$ background predicted by dedicated MC samples + signal MC, with $I_{ch}$ and $\sigma_{i\eta}$ used as fit variables. Bottom: data yields after full background subtraction vs signal MC. $I_{ch}$ vs $\sigma_{i\eta}$ fit results. . . . .	216
L.3	$Z\gamma$ check. Electron channel. Top and middle: data vs fake- $\gamma$ background derived from the template method + real- $\gamma$ background predicted by dedicated MC samples + signal MC, with $I_{ch}$ and $\sigma_{i\eta}$ used as fit variables. Bottom: data yields after full background subtraction vs signal MC. $I_{ch}$ vs $\sigma_{i\eta}$ fit results. . . . .	217



# List of Tables

1.1	Lepton Flavor Number . . . . .	9
3.1	Main parameters of LHC [12] . . . . .	51
5.1	Measurement steps. The first column is the name of the step, the second and the third columns are algebraic representations of the steps for the differential and total cross section measurements, respectively. . . . .	83
5.2	Summary of simulated samples used in the measurement. . . . .	86
5.3	Fine $\eta^\gamma$ binning for fits for $e \rightarrow \gamma$ background estimation. . . . .	102
5.4	Data, signal and background yields. $W\gamma$ , muon channel. First column is $P_T^\gamma$ ranges, second column is numbers of selected events in data in different $P_T^\gamma$ regions, third and fourth columns are estimates of the jets $\rightarrow \gamma$ background using $I_{ch}^\gamma$ and $\sigma_{i\eta i\eta}^\gamma$ templates, respectively, fifth column is MC-based estimates of the real- $\gamma$ background, sixth column is yields after the full background subtraction, and seventh column is signal MC predictions. . . . .	108

5.5	Data, signal and background yields. $W\gamma$ , electron channel. First column is $P_T^\gamma$ ranges, second column is numbers of selected events in data in different $P_T^\gamma$ regions, third and fourth columns are estimates of the jets $\rightarrow \gamma$ background using $I_{ch}^\gamma$ and $\sigma_{inj}^\gamma$ templates, respectively, fifth column is $e \rightarrow \gamma$ background estimates, sixths column is MC-based estimates of the real- $\gamma$ background, seventh column is yields after the full background subtraction, and eighth column is signal MC predictions.	109
5.6	Unfolding Results in $W\gamma$ , muon channel. First column: $P_T^\gamma$ ranges, second column: $W\gamma$ differential yields after background subtraction before unfolding, third column: $W\gamma$ yields after unfolding. . . . .	111
5.7	Unfolding Results in $W\gamma$ , electron channel. First column: $P_T^\gamma$ ranges, second column: $W\gamma$ differential yields after background subtraction before unfolding, third column: $W\gamma$ yields after unfolding. . . . .	111
5.8	Unfolding, MC closure test. $W\gamma$ , muon channel. The first column is $P_T^\gamma$ range, the second column is true selected yields in the signal MC sample based on the generator level information, the third column is reconstructed and smeared yields in the signal MC sample, the fourth column is yields from the third column unfolded by the D'Agostini method, and the fifth column is yields from the third column unfolded by the matrix inversion method. . . . .	115

5.9 Unfolding, MC closure test. $W\gamma$ , electron channel. The first column is $P_T^\gamma$ range, the second column is true selected yields in the signal MC sample based on the generator level information, the third column is reconstructed and smeared yields in the signal MC sample, the fourth column is yields from the third column unfolded by the D'Agostini method, and the fifth column is yields from the third column unfolded by the matrix inversion method. . . . .	115
5.10 Results of the background subtraction based on fits of different variables. $W\gamma$ , muon channel. First column: $P_T^\gamma$ ranges, second column: signal MC prediction, third and fourth columns: signal yields extracted from fits of $I_{ch}^\gamma$ and $\sigma_{i\eta i\eta}^\gamma$ distributions of data, fifth and sixth columns: signal yields extracted from MC closure tests, seventh column: signal yield that is used in the next measurement steps. . . . .	119
5.11 Results of the background subtraction based on fits of different variables. $W\gamma$ , electron channel. First column: $P_T^\gamma$ ranges, second column: signal MC prediction, third and fourth columns: signal yields extracted from fits of $I_{ch}^\gamma$ and $\sigma_{i\eta i\eta}^\gamma$ distributions of data, fifth and sixth columns: signal yields extracted from MC closure tests, seventh column: signal yield that is used in the next measurement steps. . . . .	120
5.12 Relative uncertainties [%]. $W\gamma$ , muon channel. The details of the "syst other" column are provided in Tab. 5.14. . . . .	123
5.13 Relative uncertainties [%]. $W\gamma$ , electron channel. The details of the "syst other" and " $e \rightarrow \gamma$ " column are provided in Tab. 5.15 and 5.16 respectively. . . . .	124

- 5.14 Relative systematic uncertainties [%] of smaller contributions (details of the column “syst other” from Tab. 5.12).  $W\gamma$ , muon channel. First column:  $P_T^\gamma$  range, second column: total systematic uncertainties of smaller contributions, third column: uncertainties related to real- $\gamma$  background estimation, fourth column: uncertainties due to limited signal MC statistical power in  $A \times \epsilon$  correction, fifth column: uncertainties due to  $M_T^W$  selection requirement, sixth column: uncertainties related to PU reweighting, seventh column: uncertainties due to limited signal MC statistical power in the unfolding correction. . . . . 124

5.15 Relative systematic uncertainties [%] of smaller contributions.  $W\gamma$ , electron channel. First column:  $P_T^\gamma$  range, second column: total systematic uncertainties of smaller contributions, third column: uncertainties related to real- $\gamma$  background estimation, fourth column: uncertainties due to limited signal MC statistical power in  $A \times \epsilon$  correction, fifth column: uncertainties due to  $M_T^W$  selection requirement, sixth column: uncertainties related to PU reweighting, seventh column: uncertainties due to limited signal MC statistical power in the unfolding correction. . . . . 125

5.16 Relative systematic uncertainties [%] of  $e \rightarrow \gamma$  background estimation.  $W\gamma$ , electron channel. First column:  $P_T^\gamma$  ranges, second column: total systematic uncertainty related to  $e \rightarrow \gamma$  background estimation, third column: systematic uncertainty evaluated by comparing fit results of samples with and without  $M_T^W$  selection requirement, fourth column: systematic uncertainty due to statistical power of MC samples involved in  $e \rightarrow \gamma$  background estimation. . . . . 125

5.17 Cross section and errors.  $W\gamma$  at  $\sqrt{s} = 8$  TeV, muon channel. . . . . 128

5.18 Cross section and errors.  $W\gamma$  at  $\sqrt{s} = 8$  TeV, electron channel. . . . . 128

A.1	Scale Factors Applied . . . . .	131
A.2	Muon “Tight” ID Scale Factors as Recommended by POG . . . . .	132
A.3	Muon Isolation Scale Factors as Recommended by POG . . . . .	132
A.4	Electron “Tight” ID Scale Factors as Recommended by POG . . . . .	132
A.5	Photon “Medium” ID Scale Factors as Recommended by POG . . . . .	132
A.6	Additional Photon Scale Factors for “PixelSeedVeto” as Derived in $W\gamma\gamma$ Measurement . . . . .	133
H.1	$e \rightarrow \gamma$ background. Barrel. Fits with $M_T^W$ requirement applied . . . . .	196
H.2	$e \rightarrow \gamma$ background. Endcap. Fits with $M_T^W$ requirement applied . . . . .	196
H.3	$e \rightarrow \gamma$ background. Barrel. Fits without $M_T^W$ requirement applied . . . . .	197
H.4	$e \rightarrow \gamma$ background. Endcap. Fits without $M_T^W$ requirement applied . . . . .	197
L.1	Relative errors [%]. MUON $Z\gamma$ . . . . .	220
L.2	Relative errors [%]. ELECTRON $Z\gamma$ . . . . .	220
L.3	Cross section and errors. MUON $Z\gamma$ . . . . .	222
L.4	Cross section and errors. ELECTRON $Z\gamma$ . . . . .	222

# Chapter 1

## Introduction

Elementary particle physics describes fundamental particles and their interactions. Fundamental particles are the smallest constituents of our Universe. When examined at smaller scales, the substances around us consist of molecules, molecules consist of atoms. In an atom there is a nucleus made of neutrons and protons and some number of electrons occupying orbits around the nucleus. Protons and neutrons have a structure while an electron is not known to have any internal structure, therefore, an electron is an example of a particle which is considered to be fundamental.

Interactions of elementary particles are described by quantum field theories which incorporate principles of the quantum mechanics and the special theory of relativity. The set of such theories, including quantum electrodynamics (QED), quantum chromodynamics (QCD) and the theory of weak interactions is called the Standard Model (SM). Current observations have proved the SM to be an accurate description of elementary particle interactions.

However, there are several experimental observations that are not described by the SM such as effects of gravity, dark matter, dark energy, matter/antimatter

asymmetry and others. Therefore, the SM is not a complete theory of particle interactions. There are several SM extensions offered by theorists as well as radically new theories waiting for experimental confirmation or exclusion.

Some SM extensions and new theories predict the existence of heavy particles with masses lying beyond experimentally reachable energies. The search of these particles is a priority in particle physics. One source of highly energetic elementary particles is cosmic rays. The most energetic particles ever observed came from this source. However, cosmic rays are totally uncontrollable and such highly energetic particles are rare. If we want to produce a large number of particles in a given energy range, we need to use a particle accelerator. A large amount of data allows experimentalists to perform a statistical analysis and increase the probability of finding a new particle if it exists.

Symmetric colliding beams is the most effective way to produce as heavy particles as possible given the energies of the colliding particles. Compared to experiments colliding a single beam at a fixed target, in the case of a symmetric collision the total momentum of two colliding particles is zero and, therefore, much larger fraction of energy can transfer to a mass of a new particle. The Large Hadron Collider (LHC) is one such collider with the highest energy in the world. It can produce the most massive particles to probe physics beyond the SM (BSM).

The Compact Muon Solenoid (CMS) is one of two general-purpose detectors at the LHC. It is placed at one of four collision points. CMS has a broad physics program including searches for the BSM physics as well as the precision measurements of the parameters of the SM itself. The measurement of this dissertation is a SM measurement with CMS data collected in 2012 in proton-proton ( $pp$ ) collisions of LHC with beam energies of 4 TeV. The result can be compared to the SM prediction. Certain BSM theories predict a deviation of the result of this

measurement from its SM value, therefore, with this measurement, in addition to testing the SM, we also search for a new physics.

The rest of this chapter gives general introductory information about the SM while Ch. 2 concentrates on the theory of the SM and BSM  $W\gamma$  production and also discusses previous measurements of this process. Chapter 3 describes LHC and CMS in more details. Chapter 4 explains one specific detail of the CMS operation that is the spacial alignment of the tracking detector of charged particles. Finally, Ch. 5 describes the details of the measurement of this dissertation and reports the results.

## 1.1 Fundamental Particles and Interactions

The SM describes interactions of elementary particles. There are four fundamental interactions: electromagnetic, strong, weak and gravitational. The gravity is not included into the SM but its effect on particles is negligible compared to the other forces which makes it possible to develop a theory of the particle physics and conduct experiments even without having the gravity included into the model.

All fundamental elementary particles in the SM can be split into three categories by their spins. There are fermions which possess spin  $s=1/2$ , there are gauge bosons which are vector particles ( $s=1$ ) and there is the Higgs boson which is a scalar particle ( $s=0$ ).

The fermions are arranged into three generations, each generation consists of a quark with charge  $Q=+2/3$  (up, charm, and top quarks), a quark with  $Q=-1/3$  (down, strange, and bottom quarks), a charged lepton with  $Q=-1$  (electron, muon, and tau-lepton) and a neutrino (electron, muon, and tau neutrinos) which is electrically neutral. Each quark can carry any of three colors: red, blue, or green. Additionally, each fermion has its antiparticle. Therefore, the total number of fundamental fermions is  $(6(\text{leptons}) + 6(\text{quarks}) \cdot 3(\text{colors})) \cdot 2(\text{to include antiparticles}) = 48$ .

Corresponding particles in different generations have the same charges, spins and interaction properties but masses of particles increase with generation. These mass differences lead to different decay properties because a particle A can decay to particles B and C only if their masses relate as  $m_A > m_B + m_C$ . Thus, an electron is a stable particle, a muon decays as  $\mu^- \rightarrow e^- + \bar{\nu}_e + \nu_\mu$ , a tau-lepton, as the heaviest charged lepton, has the largest number of decay channels amongst the charged leptons:  $\tau^- \rightarrow \mu^- + \bar{\nu}_\mu + \nu_\tau$ ,  $\tau^- \rightarrow e^- + \bar{\nu}_e + \nu_\tau$ ,  $\tau^- \rightarrow \nu_\tau + \text{quarks}$ .

In addition to fermions, the SM includes gauge bosons which are interaction mediators. They are called mediators because fermions interact with each other by exchanging them. For example, two charged fermions can interact with each other by exchanging a photon. Such interaction is called electromagnetic interaction and a photon is a mediator for the electromagnetic interaction. Similarly, a gluon is a mediator for strong interactions, and  $W^\pm$  and  $Z^0$  bosons are mediators for weak interactions.  $W^\pm$  and  $Z^0$  bosons are massive while a photon and a gluon are massless particles.

The last SM particle is the Higgs boson. The Higgs boson is a scalar neutral particle which is playing a critical role in the electroweak symmetry breaking. The Higgs mechanism explains how  $W$  and  $Z$  bosons become massive particles.

All the particles are summarized in Fig. 1.1. These and only these fundamental particles and their antiparticles have been discovered by now. However, there are many composite particles which are called hadrons. Hadrons can consist of three quarks (baryons), quark and antiquark (meson), or three antiquarks (antibaryons). Hadrons always possess an integer charge.

Most of the particles are short-lived and decay within microseconds. The only stable particles are protons and antiprotons, electrons and positrons, neutrinos and antineutrinos, photons, and, in some sense, gluons. However, if a particle cannot decay, it does not mean that it would live forever. There are many different kinds of reactions in which particles can disappear. Antiprotons and positrons would immediately annihilate with protons and electrons, photons can be absorbed by charged particles, electrons and protons can scatter to produce neutrons and neutrinos and many other reactions are possible.

In this dissertation, the study of  $pp \rightarrow W\gamma + X \rightarrow l\nu\gamma$  process where  $\ell = e, \mu$  is presented. The  $W\gamma$  production with leptonic  $W$  decays proceeds through one of

the following three processes: the initial state radiation where a photon is emitted from one of the incoming partons, the final state radiation where a photon is radiated off the charged lepton from the  $W$  boson decay, and, finally, the triple gauge coupling (TGC) where a photon is emitted from the  $W$  boson. Many BSM theories predict an enhancement of the TGC production over the SM value and, therefore, the experimental search for such an enhancement is a good test for such theories.

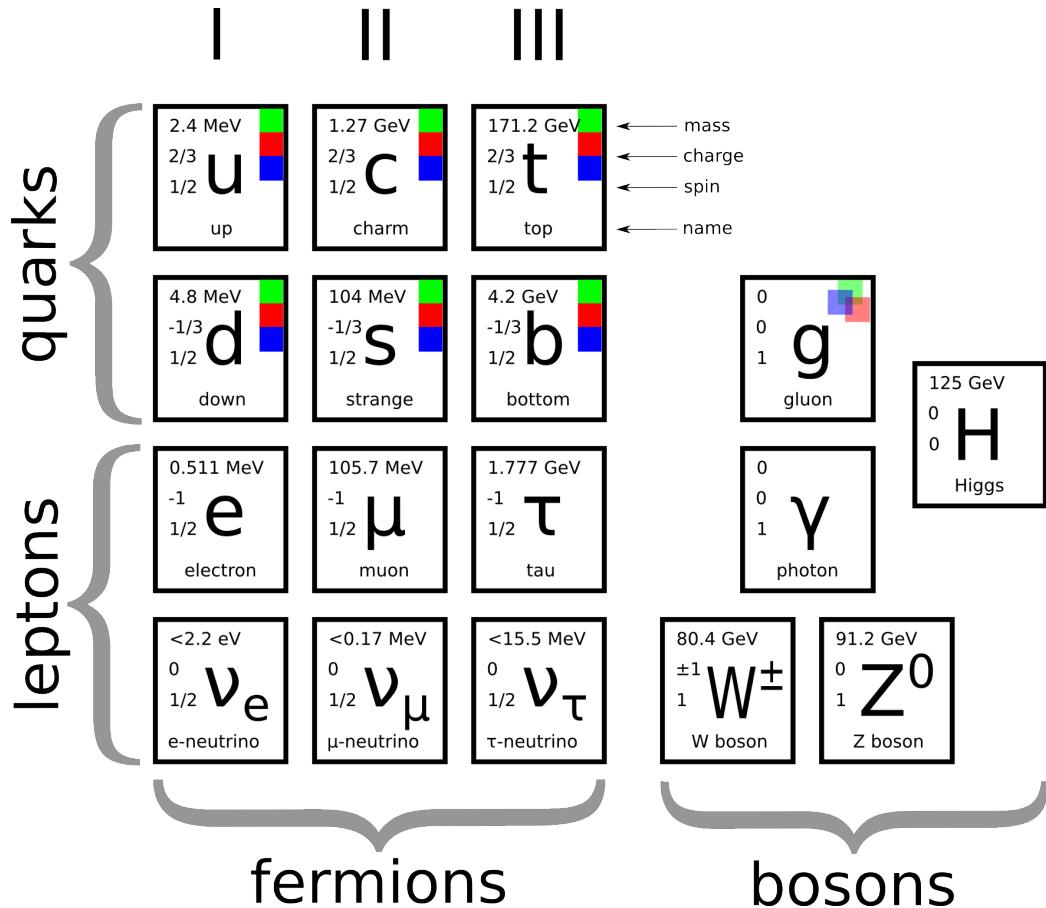


Figure 1.1: Standard Model Particles.

## 1.2 Electroweak Interactions

All electrically charged particles participate in electromagnetic interactions. The theory of electromagnetic interactions is called quantum electrodynamics (QED). All electromagnetic interactions are mediated by a photon, a spin-one electrically neutral massless particle, and can be reduced to one elementary process (Fig. 1.2, left). This process represents a charged fermion radiating or absorbing a photon. Such elementary process itself is forbidden by the energy conservation law but this element is a base of an actual process. For example, the Bhabha scattering,  $e^+e^- \rightarrow e^+e^-$ , occurs through  $e^+e^-$  annihilation with further production of a new  $e^+e^-$  pair (Fig. 1.2, middle) or through exchange of a photon between the positron and the electron (Fig. 1.2, right). Both cases involve nothing except the electromagnetic elementary process (Fig. 1.2, left). Such graphical representations of the particle physics processes are called Feynman diagrams.

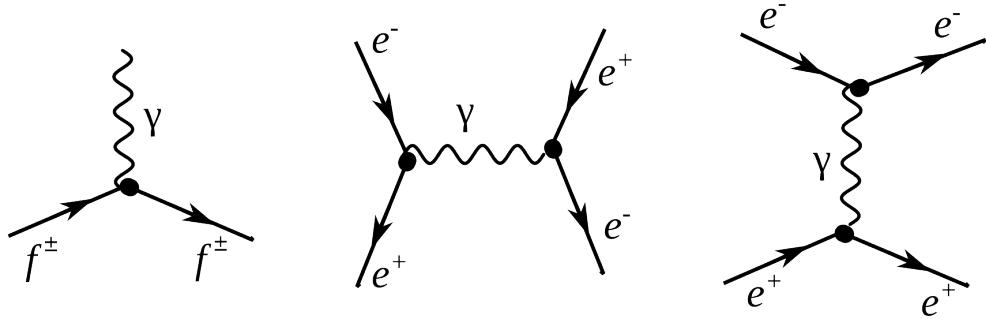


Figure 1.2: Electromagnetic interactions. Left: a photon radiation off a charged fermion, middle and right: Bhabha scattering.

As for the weak interactions, there are two kinds of them: neutral (mediated by a Z boson) and charged (mediated by a  $W^\pm$  boson). Elementary processes with W and Z bosons are shown in Fig. 1.3. Because the electric charge must be conserved

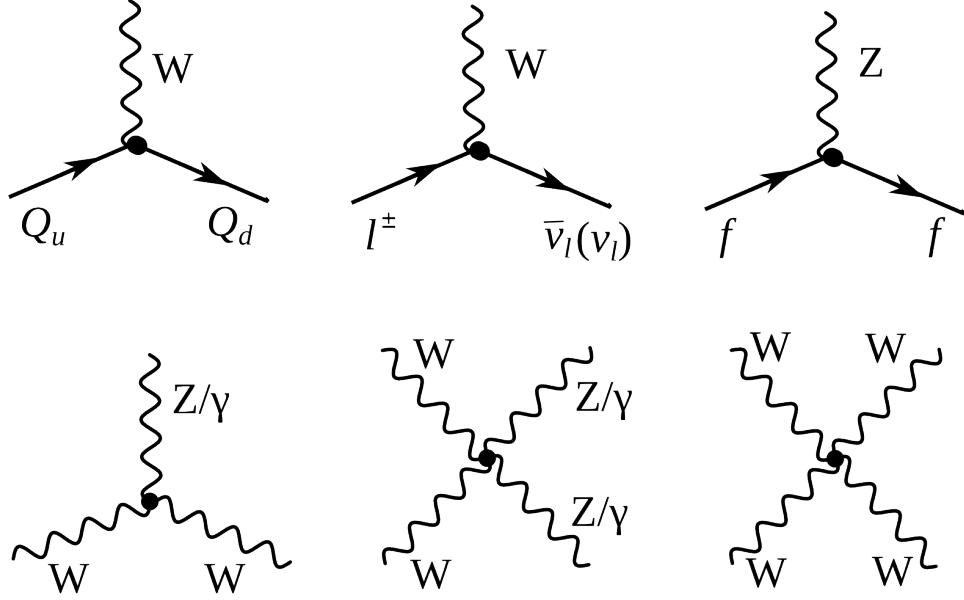


Figure 1.3: Weak elementary processes and gauge couplings. Top left: a quark with charge  $Q=+2/3$  enters, emits a  $W$  boson, and a quark with charge  $Q=-1/3$  escapes. Top middle: a charged lepton enters, emits a  $W$  boson, and a neutrino or antineutrino escapes conserving a lepton flavor number. Top right: a fermion enters, emits a  $Z$  boson and escapes. Bottom left: TGC couplings  $WW\gamma$  and  $WWZ$ . Bottom middle: QGC couplings  $WW\gamma\gamma$ ,  $WWZ\gamma$  and  $WWZZ$ . Bottom right: QGC coupling  $WWWW$ .

at any vertex, a particle radiating or absorbing a  $W$  boson converts to a different particle. Thus, a charged lepton converts to a neutrino (or vice versa) as shown in Fig. 1.3, top middle. Each lepton carries a lepton flavor number (Tab. 1.1). Lepton flavor is conserved in any interaction, thus an electron radiating a  $W$  boson always converts to an electron neutrino, a muon converts to a muon neutrino etc.

From top left diagram in Fig. 1.3 we see that if a quark with  $Q=+2/3$  enters, then a quark with  $Q=-1/3$  escapes and, therefore, the flavor of the quark is changed. The charged weak interaction is the only interaction which changes a quark flavor. The probability of each of three quarks with  $Q=-1/3$  to be born

Table 1.1: Lepton Flavor Number

particles	$L_e$	$L_\mu$	$L_\tau$
$e^-, \nu_e$	+1	0	0
$e^+, \bar{\nu}_e$	-1	0	0
$\mu^-, \nu_\mu$	0	+1	0
$\mu^+, \bar{\nu}_\mu$	0	-1	0
$\tau^-, \nu_\tau$	0	0	+1
$\tau^+, \bar{\nu}_\tau$	0	0	-1

is determined by the Cabibbo-Kobayashi-Maskawa matrix which relates mass eigenstates  $d, c$  and  $b$  to weak eigenstates  $d', c'$  and  $b'$  (Eq. 1.1). Absolute values of the matrix elements are all known (Eq. 1.2) and are the highest for the quark of the same generation as the initial state quark. In the particular case shown in the top left diagram in Fig. 1.3,  $u$  is the initial state quark and  $d$  has the highest probability to be produced after an interaction with a  $W$  boson but  $s$  and  $b$  can also be produced if there is enough energy.

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \quad (1.1)$$

$$\begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} = \begin{pmatrix} 0.97 & 0.23 & 0.00 \\ 0.23 & 0.97 & 0.04 \\ 0.01 & 0.04 & 1.00 \end{pmatrix} \quad (1.2)$$

An elementary process of a neutral weak interaction is an emission a  $Z$  boson off a fermion line (right top diagram in Fig. 1.3). Diagrams with a  $Z$  boson are very similar to ones with a photon except a photon can only be radiated off a charged particle but a  $Z$  boson can also be radiated off a neutrino or antineutrino.

The bottom diagrams in Fig. 1.3 are gauge bosons coupling diagrams including self-coupling of a  $W$  boson, its interaction with a  $Z$  boson and its electromagnetic radiation of a photon. Charge-conserving TGC and quartic gauge couplings (QGC) containing two or four  $W$  bosons are all possible in the SM:  $WWZ$ ,  $WW\gamma$ ,  $WWZZ$ ,  $WWZ\gamma$ ,  $WW\gamma\gamma$ , and  $WWWW$ .

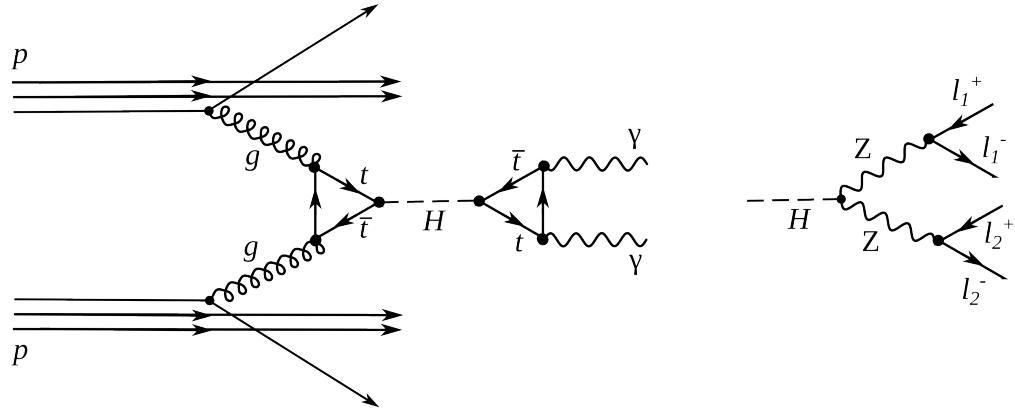


Figure 1.4: The Higgs boson production and decay. Left:  $H \rightarrow \gamma\gamma$ , right:  $H \rightarrow ZZ \rightarrow 4l$ .

Electromagnetic and weak interactions are unified by the electroweak Glashow-Weinberg-Salam (GWS) theory which is based on  $SU(2) \times U(1)$  symmetry.  $SU(2)$  is the symmetry of weak isospin which generates three bosons:  $W^1$ ,  $W^2$  and  $W^3$ .  $U(1)$  is the symmetry of the weak hypercharge and generate one neutral boson  $B$ .  $W^1$  and  $W^2$  are mixed to create  $W^+$  and  $W^-$  mediators while  $W^3$  and  $B$  are mixed to create a  $Z$  boson and a photon. Therefore, the GWS theory considers electromagnetic and weak forces as different manifestations of the electroweak force. The electroweak theory is discussed in greater details in Ch. 2.

Weak interactions are mediated by heavy bosons ( $M_W = 80$  GeV,  $M_Z = 91$  GeV) while electromagnetic interactions are mediated by a massless photon, thus, the electroweak symmetry is broken. To explain this phenomenon, the Higgs

mechanism was introduced. The mechanism predicted an existence of an additional boson: the Higgs boson. The Higgs boson was a missing piece of the SM for many years and was finally discovered in 2012 at LHC by ATLAS and CMS collaborations through the processes shown in Fig. 1.4 [13], [14].

The measurement in this dissertation is an electroweak measurement because the process involves a  $W$  boson. It includes an interaction of a  $W$  boson with leptons and quarks as well as the TGC  $WW\gamma$ . Thus, the measurement is a good test of the SM electroweak theory.

### 1.3 Strong Interactions

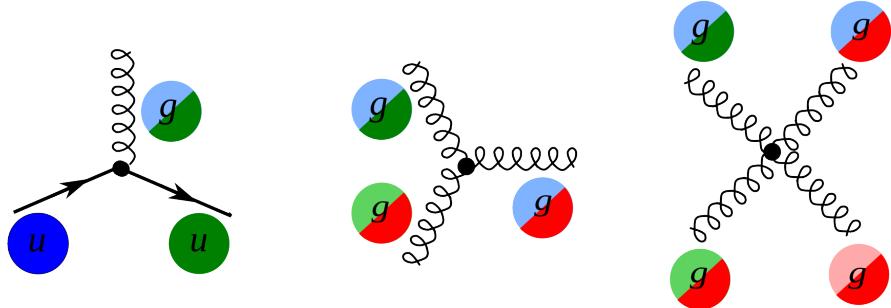


Figure 1.5: Elementary processes of strong interactions

The third fundamental force after the electromagnetic and weak ones is the strong force. The strong force is responsible for gluing protons and neutrons together in the nuclei as well as for forming protons and neutrons themselves. The strong interactions occur by exchanging gluons which are spin-one massless electrically neutral particles.

The elementary strong processes are shown in Fig. 1.5. There are three elementary processes:  $qqg$ ,  $ggg$  and  $gggg$ , all are involving particles with color charges. Thus, gluons couple to quarks and self-couple. Color charges must be conserved at each elementary process of the strong interaction. Each quark possesses one of three colors at a time, and there are eight types of gluons to cover all possible color exchanges.

The coupling constant of the strong interaction depends on the distance between interacting particles: it becomes larger as the distance becomes larger and smaller as the distance becomes smaller. As the distance approaches zero, the coupling constant approaches zero too, and, thus, in the asymptotic limit two quarks located at the same place do not interact. This property is called asymptotic freedom.

On the other hand, when the distance between quarks becomes larger, the coupling constant also becomes larger. This property confines quarks to always stay in the color neutral combinations (hadrons), it forbids the existence of free quarks. A combination becomes color neutral when there is the same amount of color and anticolor or if there is the same amount of each of the three colors. Thus, mesons are comprised of a quark and an antiquark with the opposite color charges, and baryons are composed of three quarks: red, green and blue one. Examples of baryons include such well-known particles as a proton and a neutron.

The asymptotic freedom and the confinement are properties that are specific to strong interactions. The theory of strong interactions is called the quantum chromodynamics (QCD) which is a quantum field theory invariant under  $SU(3)$  color transformations. When the coupling constant is much less than one  $\alpha_s \ll 1$ , the perturbative approach can be used to compute observables.

The  $W\gamma$  process being measured in this dissertation is not intended to test QCD, but a good understanding of QCD is essential for performing this measurement because the QCD corrections to the Feynman diagrams of the process are large. In addition, QCD describes the dynamics of quarks and gluons within colliding protons and predicts probabilities of one or another quark-antiquark pair to interact. Physics of proton-proton collisions is discussed in Ch. 1.4.

## 1.4 Physics of Proton-Proton Collisions

Consider a  $pp$  collision at LHC. The proton energies are so high that each proton behaves as a complex structure. A proton is a baryon, it consists of three quarks:  $uud$ . These three quarks are called valence quarks. They interact with each other by exchanging gluons which produce virtual  $q\bar{q}$  pairs (Fig. 1.6). Such virtual quarks are also called sea quarks.

Any parton, quark, antiquark or gluon, from one proton can interact with any parton from another proton. Probabilities  $f_i(x, Q^2)$  of any particular constituent  $i$  to interact are described partially by QCD and partially by experimental measurements and depend on the momentum transfer  $Q$  and the momentum fraction of a specific parton  $x$ . These probabilities are called parton distribution functions (PDFs).

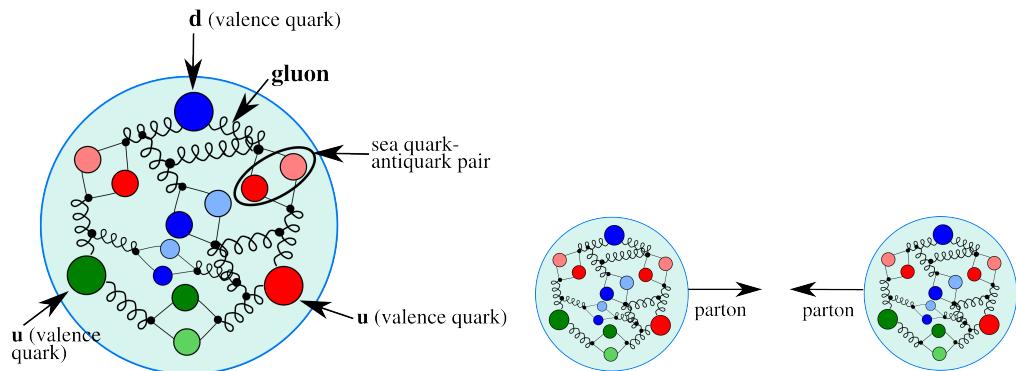


Figure 1.6: The proton structure (left) and the proton-proton collision (right).

For large  $Q^2$  and  $x$  gluon-gluon interactions have the largest probabilities to occur (Fig. 1.7). However, gluons do not couple directly to a  $W$  boson, thus, in the  $W\gamma$  measurement we are mostly interested in quark-antiquark pairs which would have a total charge corresponding to the charge of a  $W$  boson ( $\pm 1$ ). Since

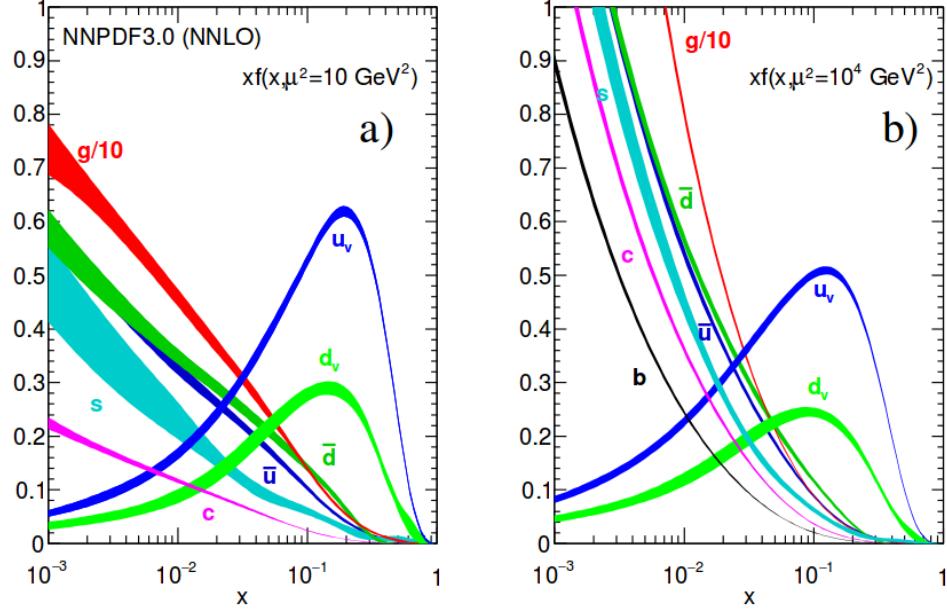


Figure 1.7: Parton distribution functions [1].

we have  $u$  and  $d$  as valence quarks and we know that the probability to couple to the same generation quark in charged weak interactions is the highest, most of the  $W$  bosons are created by  $u\bar{d}$  and  $d\bar{u}$  pairs however other  $q\bar{q}'$  combinations with the total charges of  $\pm 1$  are also possible.

## 1.5 Open Questions of the Standard Model

While the SM is an accurate description of all particle physics experimental results, there are certain phenomena which are not included into the SM. In this subsection we discuss some of them.

The gravitational interactions do not fit into the SM. It is the open question whether the quantum theory of gravity is possible and whether there is a mediator of the gravitational interactions. Also, it is not known why the gravitational force is so much weaker than the other forces. One possible explanation comes from a theory which predicts extra spatial dimensions beyond the three we experience (e.g. the string theory). In this case, it is possible that the gravitational force is shared with other dimensions and only a fraction is available in our three dimensions.

Another mystery of the universe is its composition: it is known from the studies of the gravitational effects that our universe consists of dark energy by 68%, of dark matter by 27% and of baryon matter only by 5% [15]. The dark energy resists the gravitational attraction and accelerates the expansion of the universe, and is not detectable by any effects except gravitational. The understanding of dark energy is a question of general relativity rather than particle physics. The dark matter, however, likely consists of particles and therefore is a subject of particle physics. It does not radiate and that is why it cannot be detected by telescopes. The nature of the dark matter is not known but its constituents must be very stable to remain since the Big Bang. The theory of the supersymmetry which is unifying fundamental particles and mediators predicts many of new heavy particles and the lightest supersymmetric particle, the neutralino, is a good candidate for dark matter.

One more open question is the reason for the matter/antimatter asymmetry.

Matter and antimatter should have been created in the same amount at the moment of the Big Bang. Most of it has annihilated but because of asymmetry, there was more matter than antimatter which led to the state of the Universe we observe now. There is a phenomenon of the CP-violation in weak interactions observed and described which predicts the asymmetry at a certain level. However, the effect of the CP-violation is not large enough to account for the observed amount of the matter and, therefore, the total matter/antimatter asymmetry remains unexplained.

The measurement of the  $W\gamma$  production in  $pp$  collisions has a goal to both test the SM and search for the BSM physics. We measure a cross section differential in the component of the photon momentum, transverse to the beamline (referred as photon transverse momentum, or  $P_T^\gamma$ ). The low  $P_T^\gamma$  region is not expected to be affected by any new physics and must agree well with the SM predictions while the high  $P_T^\gamma$  region may indicate an existence of new physics if there is an enhancement over the SM predictions. The excess would be indirect evidence of the BSM particles like supersymmetric particles or additional gauge bosons which could be part of the explanation of the dark matter presence or difference in magnitudes of different interactions. More theoretical details about the SM description of  $W\gamma$  process as well as possible BSM physics are given in Ch. 2.

## Chapter 2

# $W\gamma$ Production Theory and Former Experimental Results

Chapter 2 provides deeper theoretical background for the measurement of this dissertation and discusses former experimental results. The derivation of the electroweak Lagrangian is described in Ch. 2.1, including the appearance of triple gauge coupling (TGC) and quartic gauge coupling (QGC) terms. Then concepts of the cross section and the luminosity are discussed in Ch. 2.2. More specific details regarding the SM cross section of  $W\gamma$  are summarized in Ch. 2.3. Possible causes and potential effects of anomalous TGC (aTGC) are explained in Ch. 2.4. Finally, Ch. 2.5 lists former physics experiments which probed the same aTGC vertex which is probed in the measurement of this dissertation including measurements of exactly the same process at lower LHC beam energy.

## 2.1 Electroweak Theory of the Standard Model

To develop a quantum field theory, we start with the Lagrangian of free fermions. In order to describe a system with a conserved physical quantity, the Lagrangian is required to satisfy a local invariance with respect to a certain transformation. For instance, a conservation of electric charge requires local invariance under  $U(1)$  transformation for the QED Lagrangian [16]. The requirement of local invariance introduces an interaction between one or more new vector fields and our free fermions. The new vector fields are mediators of an interaction conserving the physical quantity. To provide a full description for a new boson field, in addition to the interaction term we introduce an invariant term for the kinetic energy of the boson. Such an approach allows us to derive a Lagrangian which is locally invariant with respect to a certain gauge transformation and contains interacting fermions as well as interaction mediators.

The SM is a quantum field theory invariant under the local  $SU(3)_C \times SU(2)_L \times U(1)_Y$  transformation [16]. The SM Lagrangian includes all observed quantum fields and their interactions.

The part of the SM Lagrangian based on the  $SU(3)_C$  symmetry is called QCD or the theory of strong interactions. QCD has three types of charges which are called colors: red, blue, and green. To be a subject of the strong interaction, a fermion must possess a color charge. Quarks and antiquarks are such fermions. The requirement to satisfy the gauge invariance with respect to  $SU(3)_C$  transformations generates eight massless gluons, and the non-abelian nature of the  $SU(3)$  group generates self-interactions of gluons including three-gluon and four-gluon vertices.

The part of the SM Lagrangian based on the  $SU(2)_L \times U(1)_Y$  symmetry is the foundation of the unified theory of electroweak interactions.  $SU(2)_L$  reflects

transformations in the weak isospin space of left-handed fermions ([17], Ch. 9) while  $U(1)_Y$  reflects transformations in a weak hypercharge space of all fermions. The requirement of the local gauge invariance generates four massless vector bosons which are mediators of electromagnetic and weak interactions. The non-abelian structure of the  $SU(2)$  group generates gauge boson self-couplings the same way as self-interactions of gluons appear in QCD.

Mass terms for the vector bosons would violate the gauge invariance of the electroweak Lagrangian, however it is experimentally known that the mediators of weak interactions are heavy particles with masses  $M_W = 80$  GeV and  $M_Z = 91$  GeV. A possible solution of this discrepancy is the mechanism of Spontaneous Symmetry Breaking.

The mechanism of Spontaneous Symmetry Breaking and the appearance of the mass terms for  $W$  and  $Z$  bosons is realized by introducing an additional doublet of scalar fields. After that, the Lagrangian is transformed in such a way that  $W$  and  $Z$  bosons acquire masses through their interactions with a new particle: the Higgs boson ( $H$ ). A photon does not couple to the Higgs boson remaining a massless particle and leaving QED symmetry group  $U(1)$  to be unbroken.

The measurement in this dissertation provides a test for the electroweak sector of the SM. We will retrace the steps of the derivation of the electroweak part of the SM Lagrangian starting from the terms for free fermions. The resulting Lagrangian accommodates electroweak gauge bosons and their self-couplings. One of these self-couplings,  $WW\gamma$ , is the primary focus of our measurement.

It is experimentally known that the dynamics of weak interactions depend on particle chirality ([17], chapter 4.4.1). In particular, a  $W$  boson couples to left-handed fermions and right-handed antifermions only. Given different properties of left-handed and right-handed fermions, they are treated differently by the

electroweak theory.  $SU(2)$  doublets are introduced for the wave functions of left-handed fermions while  $SU(2)$  singlets are introduced for the wave functions of right-handed fermions. Equations 2.1 and 2.2 show wave functions for the first generation fermions. Wave functions for the other two generations are constructed the same way.

$$\psi_1(x) = \begin{pmatrix} u \\ d' \end{pmatrix}_L, \psi_2(x) = u_R, \psi_3(x) = d'_R. \quad (2.1)$$

$$\psi_1(x) = \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L, \psi_2(x) = \nu_{eR}, \psi_3(x) = e^-_R. \quad (2.2)$$

The state  $d'$  in Eq. 2.1 is a weak eigenstate which is a linear combination of the mass eigenstates of the  $d$ ,  $c$  and  $b$  quark wave functions and is determined by the quark mixing matrix,  $V$ , which is also called Cabibbo-Kobayashi-Maskawa matrix [16]:

$$\begin{pmatrix} d' \\ c' \\ b' \end{pmatrix} = V \begin{pmatrix} d \\ c \\ b \end{pmatrix} \quad (2.3)$$

To derive the unified electroweak Lagrangian, we start with the free fermion terms:

$$L_0 = \sum_{j=1}^3 i\bar{\psi}_j(x)\gamma^\mu\partial_\mu\psi_j(x), \quad (2.4)$$

where  $\gamma^\mu$  are Dirac matrices ([17], chapter 7.1) and  $\psi_j(x)$  are wave functions determined by Eqs. 2.1 and 2.2.

The wave function  $\psi_1$  changes under the  $SU(2)_L \times U(1)_Y$  transformations in

the following way:

$$\psi_1(x) \rightarrow e^{iy_1\beta} U_L \psi_1(x), \quad (2.5)$$

while the wave functions  $\psi_{(2,3)}(x)$  are singlets of  $SU(2)_L$  and are affected only by  $U(1)$  transformations:

$$\psi_{(2,3)}(x) \rightarrow e^{iy_{(2,3)}\beta} \psi_{(2,3)}(x). \quad (2.6)$$

The transformation in the weak isospin space is defined as  $U_L \equiv e^{i\sigma_i \alpha_i / 2}$  where  $\sigma_i$  are Pauli matrices ([17], chapter 4.2.2). Phases  $\alpha_i(x)$  and  $\beta(x)$  in Eqs. 2.5 and 2.6 are arbitrary functions of  $x$ , and  $y_{(1,2,3)}$  are weak hypercharges which are named analogous to electric charges in QED.

In order for the Lagrangian to satisfy the local  $SU(2)_L \times U(1)_Y$  invariance, partial derivatives in Eq. 2.4 have to be substituted with covariant derivatives:

$$D_\mu \psi_1(x) = [\partial_\mu - ig\tilde{W}_\mu(x) - ig'y_1 B_\mu(x)] \psi_1(x) \quad (2.7)$$

$$D_\mu \psi_{(2,3)}(x) = [\partial_\mu - ig'y_{(2,3)} B_\mu(x)] \psi_{(2,3)}(x) \quad (2.8)$$

where  $g, g'$  are arbitrary constants,

$$\tilde{W}_\mu(x) \equiv \frac{\sigma_i}{2} W_\mu^i(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{2}W_\mu^3 & (W_\mu^1 - iW_\mu^2)/\sqrt{2} \\ (W_\mu^1 + iW_\mu^2)/\sqrt{2} & -W_\mu^3 \end{pmatrix}, \quad (2.9)$$

$B_\mu, W_\mu^1, W_\mu^2, W_\mu^3$  are four vector bosons that arise from the requirement that the Lagrangian is invariant under local  $SU(2)_L \times U(1)$  transformations.

The Lagrangian becomes:

$$L_0 \rightarrow L = \sum_{j=1}^3 i\bar{\psi}_j(x)\gamma^\mu D_\mu \psi_j(x) \quad (2.10)$$

To make new vector bosons physical fields it is necessary to add terms for their kinetic energies:

$$L_{KIN} = -\frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}W_{\mu\nu}^i W_i^{\mu\nu} \quad (2.11)$$

where  $B_{\mu\nu} \equiv \partial_\mu B_\nu - \partial_\nu B_\mu$ ,  $W_{\mu\nu}^i \equiv \partial_\mu W_\nu^i - \partial_\nu W_\mu^i + g\epsilon^{ijk}W_\mu^j W_\nu^k$

Off-diagonal terms of  $\tilde{W}_\mu$  are wave functions of charged vector bosons

$$W^\pm = (W_\mu^1 \mp iW_\mu^2)/\sqrt{2} \quad (2.12)$$

while  $W_\mu^3$  and  $B_\mu$  are neutral fields which are mixtures of a  $Z$  boson and a photon determined by:

$$\begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix} \equiv \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} \quad (2.13)$$

where  $\theta_W$  is the electroweak mixing angle and  $A_\mu$  is a photon field.

In order to be consistent with QED, terms involving  $A_\mu$  in the electroweak Lagrangian must be equal to the corresponding terms in the QED Lagrangian [16]:

$$L_{QED} = i\bar{\psi}(x)\gamma^\mu \partial_\mu \psi(x) - m\bar{\psi}(x)\psi(x) + qA_\mu(x)\bar{\psi}(x)\gamma^\mu \psi(x) - \frac{1}{4}F_{\mu\nu}(x)F^{\mu\nu}(x), \quad (2.14)$$

where  $q$  is electric charge of  $\psi(x)$  field,  $F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu$ .

This requirement relates  $g$ ,  $g'$ ,  $\theta_W$  and  $e$  as  $g \sin \theta_W = g' \cos \theta_W = e$  and provides an expression for weak hypercharges:  $y = q - t_3$ , where  $q$  is the electric charge and  $t_3$  is the  $z$ -component of the weak isospin. This results in  $y_1 = 1/6$ ,  $y_2 = 2/3$ , and  $y_3 = -1/3$  for quarks and  $y_1 = -1/2$ ,  $y_2 = 0$ , and  $y_3 = -1$  for leptons. A right-handed neutrino has a weak hypercharge of  $y_2 = 0$ . It also does not have an electric charge, and as a right-handed fermion has  $t_3 = 0$ , therefore, it does not couple to a  $W$  boson. Thus, a right-handed neutrino does not participate in any SM interaction.

Writing  $\tilde{W}_\mu$  in Eq. 2.11 explicitly, we obtain triple gauge coupling (TGC) and quartic gauge coupling (QGC) terms:

$$L_{TGC} = -\frac{g}{4}(\partial_\mu W_\nu^i - \partial_\nu W_\mu^i)\epsilon^{ijk}W^{\mu j}W^{\nu k} - \frac{g}{4}\epsilon^{ijk}W_\mu^jW_\nu^k(\partial^\mu W^{\nu i} - \partial^\nu W^{\mu i}) \quad (2.15)$$

$$L_{QGC} = -\frac{g^2}{4}\epsilon^{ijk}\epsilon^{ilm}W_\mu^jW_\nu^kW_\lambda^{\mu l}W^{\nu m} \quad (2.16)$$

Substituting expressions for  $W_\mu^i$  and  $B_\mu$  determined by Eqs. 2.12 and 2.13 into Eqs. 2.15 and 2.16 we receive charged TGC and QGC terms in the Lagrangian (those involving two or four  $W$  bosons) in the forms of Eqs. 2.17 and 2.20, but all neutral TGC and QGC terms (those not involving any  $W$  bosons) cancel out.

Equation 2.17 involves  $WWZ$  (Eq. 2.18) and  $WW\gamma$  (Eq. 2.19) interactions:

$$L_{TGC} = L_{TGC}^{(1)} + L_{TGC}^{(2)}, \quad (2.17)$$

$$L_{TGC}^{(1)} = -ie \cot \theta_W (W^{-\mu\nu} W_\mu^+ Z_\nu - W^{+\mu\nu} W_\mu^- Z_\nu + W_\mu^- W_\nu^+ Z^{\mu\nu}), \quad (2.18)$$

$$L_{TGC}^{(2)} = -ie (W^{-\mu\nu} W_\mu^+ A_\nu - W^{+\mu\nu} W_\mu^- A_\nu + W_\mu^- W_\nu^+ A^{\mu\nu}). \quad (2.19)$$

Equation 2.20 involves  $WWWW$  (Eq. 2.21),  $WWZZ$  (Eq. 2.22),  $WWZ\gamma$  (Eq. 2.23), and  $WW\gamma\gamma$  (Eq. 2.24) interactions:

$$L_{QGC} = L_{QGC}^{(1)} + L_{QGC}^{(2)} + L_{QGC}^{(3)} + L_{QGC}^{(4)}, \quad (2.20)$$

$$L_{QGC}^{(1)} = -\frac{e^2}{2 \sin^2 \theta_W} (W_\mu^+ W^{-\mu} W_\nu^+ W^{-\nu} - W_\mu^+ W^{\mu} W_\nu^- W^{-\nu}), \quad (2.21)$$

$$L_{QGC}^{(2)} = -e^2 \cot^2 \theta_W (W_\mu^+ W^{-\mu} Z_\nu Z^\nu - W_\mu^+ Z^\mu W_\nu^- Z^\nu), \quad (2.22)$$

$$L_{QGC}^{(3)} = -e^2 \cot \theta_W (2W_\mu^+ W^{-\mu} Z_\nu A^\nu - W_\mu^+ Z^\mu W_\nu^- A^\nu - W_\mu^+ A^\mu W_\nu^- Z^\nu), \quad (2.23)$$

$$L_{QGC}^{(4)} = -e^2 (W_\mu^+ W^{-\mu} A_\nu A^\nu - W_\mu^+ A^\mu W_\nu^- A^\nu). \quad (2.24)$$

In the measurement of this dissertation we probe  $WW\gamma$  coupling (Eq. 2.19).

The unified electroweak Lagrangian discussed above involves kinetic energy terms for fermions and gauge bosons as well as interactions of fermions with gauge bosons, TGC, and QGC. However, this Lagrangian does not contain any

mass terms. Because left-handed and right-handed wave functions transform differently under the electroweak symmetry, adding fermion mass terms of  $\frac{1}{2}m_f^2\bar{\psi}\psi$  would violate the Lagrangian invariance and, therefore, fermion mass terms are forbidden by the  $SU(2) \times U(1)$  symmetry requirement. Mass terms for gauge bosons also would violate the Lagrangian invariance just as a photon mass term  $\frac{1}{2}m^2A^\mu A_\mu$  would violate  $U(1)$  invariance of  $L_{QED}$  [17]. Therefore, Lagrangian  $L$  in Eq. 2.10 contains massless particles only.

However, it is known from experiments that the  $Z$  and  $W$  bosons as well as fermions are massive particles and, therefore, our theory should accommodate their masses. To introduce masses into the electroweak Lagrangian, an  $SU(2)_L$  doublet of complex scalar fields  $\phi(x)$  is added to the Lagrangian:

$$\phi(x) \equiv \begin{pmatrix} \phi^{(+)}(x) \\ \phi^{(0)}(x) \end{pmatrix} \quad (2.25)$$

By selecting a special gauge of  $\phi(x)$  it is possible to spontaneously break electroweak symmetry, generate a new scalar particle, the Higgs boson [16], and introduce mass terms for  $W$  and  $Z$  bosons and charged fermions through their couplings to the Higgs boson. The strength of the coupling constant is proportional to the square of the particle's mass, therefore, heavier particles are more likely to interact with  $H$ , and massless particles do not couple to  $H$ .

The mechanism of generating a fermion's mass involves both left-handed and right-handed components of the fermion. If our hypothesis that right-handed neutrinos do not exist is right, then the Higgs mechanism does not generate neutrino masses. However, from the experiments of neutrino oscillations, neutrinos are known to have masses even though they are orders of magnitude smaller than those of other fermions. Several hypotheses have been offered to resolve this

contradiction however at the moment the mechanism for neutrinos to acquire masses remain unknown [1].

In this dissertation, we study an electroweak process  $W\gamma \rightarrow l\nu_l\gamma$  and probe the TGC vertex  $WW\gamma$  (Eq. 2.19). To do that, we measure the differential cross section of  $W\gamma \rightarrow l\nu_l\gamma$  with respect to the photon transverse momentum. The concept of the cross section in particle physics is discussed in the next chapter.

## 2.2 Cross Section and Luminosity

In this dissertation we measure the total cross section of the process  $pp \rightarrow l\nu_l\gamma + X$  and its differential cross section in transverse momentum of the photon. A cross section in particle physics is an interaction probability per unit flux of incident particles [18]. It can be interpreted as an area which must be crossed by an incident particle in order to interact with a scattering center, or, in case of a differential cross section, area  $d\sigma$  within which an incident particle must appear to be scattered off by an angle  $d\theta$  (Fig. 2.1). The relationship between  $d\sigma$  and  $d\theta$  gives us the expression for a differential cross section  $d\sigma/d\theta$ . Integrating over  $d\theta$ , we obtain the total cross section  $\sigma$ . The cross section concept illustrated in Fig. 2.1 is generalized to be an effective area, and is generalized for two (or more) particle interactions rather than a light particle scattering off a stationary center.

The angle  $\theta$  here is used only as an illustration of a concept of differential cross section. In particle physics we measure a differential cross section with respect to a parameter  $X$  which can be a parameter of one of final state particles or of a system of final state particles. For example, a cross section could be measured as a function of the transverse momentum of a final state photon  $P_T^\gamma$ , the invariant mass of two final state leptons  $m_{ll}$ , or even discreet observables such as the number of jets associated with the process  $N_{jets}$ .

In the scenario illustrated in Fig. 2.1, the number of particles passing through the area  $\sigma$  per unit time is

$$N = L \cdot \sigma, \quad (2.26)$$

where  $L$  is the flux of incident particles and is called luminosity. For colliding beams, the luminosity is determined by collisions frequency, the number of col-

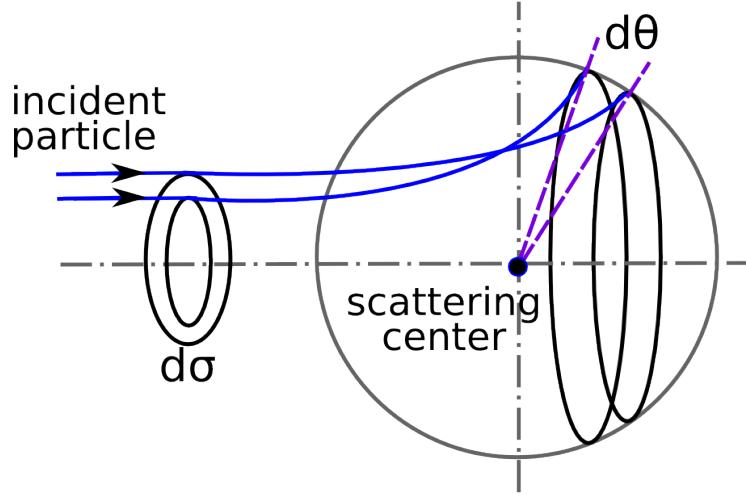


Figure 2.1: Illustration of the differential cross section concept in the classical case.

liding particles in each beam, and beams cross sections. The cross section  $\sigma$  of a specific process can be determined from an experiment as  $\sigma = N/L$ .

A cross section can be computed theoretically using the following expression:

$$\sigma = \frac{W_{fi}}{F} N_{fs}, \quad (2.27)$$

where  $W_{fi}$  is a transition probability between final and initial states of the system per unit spatial volume,  $F$  is the initial flux, and  $N_{fs}$  is the density of final states ([19], chapter 4.3). The initial flux in this expression is determined as number of incident particles per unit volume multiplied by their velocity and by the number of target particles per unit volume.

The formula for the cross section relevant for our measurement, two particles to three final state particles scattering  $1 + 2 \rightarrow 3 + 4 + 5$ , is determined by the Fermi's Golden Rule [17]:

$$\sigma = \frac{1}{4\sqrt{(p_1 p_2)^2 - (m_1 m_2)^2}} \int |M|^2 (2\pi)^4 \delta^4(p_1 + p_2 - p_3 - p_4 - p_5) \prod_{j=3}^5 \frac{1}{2\sqrt{\bar{p}_j^2 + m_j^2}} \frac{d^3 \bar{p}_j}{(2\pi)^3}, \quad (2.28)$$

where  $p_i$  are four-momenta and  $\bar{p}_i$  are three momenta of the initial state and the final state particles,  $m_i$  are masses of particles,  $M$  is the process amplitude determined by the dynamics of the particles interaction. All possible momenta of the final state particles is called the phase space.

During proton-proton collisions at high energy, the hard scattering process occurs between partons in the protons, as discussed in Ch. 1.4. Therefore, the cross section of a process  $pp \rightarrow X + Y$  has two ingredients: PDFs and a partonic cross section  $\sigma_{ab \rightarrow X}$ . The partonic cross section is described by perturbative QCD while PDFs require non-perturbative computations and are determined, in part, from experiments (Fig. 1.7). According to the QCD factorization theorem [20]:

$$\sigma(pp \rightarrow X + Y) = \sum_{a,b} \int dx_a dx_b f_a(x_a, Q^2) f_b(x_b, Q^2) \sigma(ab \rightarrow X). \quad (2.29)$$

In the case of  $W\gamma$  process,  $X$  is  $l\nu\gamma$ ,  $ab$  are  $q_i\bar{q}_j$  or  $q_j\bar{q}_i$ .  $Q^2$  is the large momentum scale that characterizes hard scattering,  $f_a$  and  $f_b$  are PDFs,  $x_a$  and  $x_b$  are fractions of momenta of the partons. In the next sections we will discuss the computation of partonic cross sections of the  $W\gamma$  process and possible BSM effects.

## 2.3 Standard Model $W\gamma$ Production

A W boson in proton-proton collisions can be produced in the processes  $q\bar{q}' \rightarrow W$  where  $q$  and  $\bar{q}'$  are a quark and an antiquark which have a total charge of +1 if producing a  $W^+$  boson or -1 if producing a  $W^-$  boson. The processes  $u\bar{d} \rightarrow W^+$  and  $d\bar{u} \rightarrow W^-$  are the most likely to occur because  $u$  and  $d$  are valence quarks in a proton. There are twice as many  $u$  quarks in a proton as  $d$  quarks, therefore,  $W^+$  is produced twice more frequently than  $W^-$ . Antiquarks  $\bar{d}$  and  $\bar{u}$  come from the sea  $q\bar{q}$  pairs of the other proton.

Once created, a W boson decays immediately, its lifetime is  $\simeq 10^{-25}$  s. In an experiment one detects its decay products rather than the W boson itself. Decay modes of a W boson include  $W^\pm \rightarrow l^\pm \nu_l (\bar{\nu}_l)$  where  $l^\pm = e^\pm, \mu^\pm$  or  $\tau^\pm$  with branching fractions of 11% per a leptonic channel [1]. The remaining 67% account for various  $W \rightarrow q\bar{q}'$  decays. In this dissertation we only consider  $W^\pm \rightarrow \mu^\pm \nu_\mu (\bar{\nu}_\mu)$  and  $W^\pm \rightarrow e^\pm \nu_e (\bar{\nu}_e)$  channels.

A photon can be emitted from any charged particle of the process: a quark, an antiquark, a charged lepton or a W boson (Fig. 2.2, top). A quark and an antiquark are initial state particles and, therefore, if one of them radiates a photon, we refer to the process as initial state radiation (ISR). A muon or an electron is a final state particle and if it radiates a photon, we call such a process final state radiation (FSR). Finally, a W boson is a gauge boson and if it radiates a photon, the process has a vertex with three gauge bosons:  $WW\gamma$ , and we call such process the triple gauge coupling (TGC). We cannot distinguish between these processes experimentally because we detect final state particles only.

The electroweak Lagrangian is described in Chapter 2.1. It is possible to derive equations of motion from the Lagrangian for any fields involved [17]. However, in

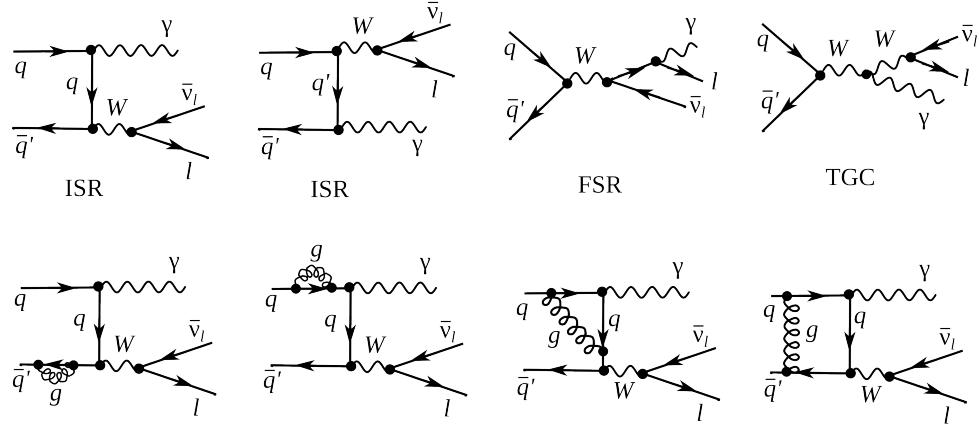


Figure 2.2: Feynman diagrams of  $W\gamma$  production. Top: LO diagrams, bottom: several examples of NLO in QCD.

a quantum field theory equations of motion cannot be solved exactly and, therefore, the perturbative approach is used if a coupling constants is  $g \ll 1$ .

To represent the process graphically Feynman diagrams were invented. Also the diagrams can be used to calculate the process amplitude  $M$  in Eq. 2.28 because they are determined by Lagrangian terms relevant to the process. There are an infinite number of Feynman diagrams corresponding to any specific process and the total amplitude of the process is a sum of individual amplitudes of each diagram and it is not technically possible to take into account all of them. Each vertex introduces a factor in the amplitude of the process that is proportional to the coupling constant. If the coupling constant is  $g \ll 1$ , the perturbative approach arranges all the diagrams by orders of contribution, and, therefore, the Feynman diagrams with fewer vertices would give a significantly larger contribution to the amplitude. In Fig. 2.2 examples of the Leading Order (LO) and the Next-to-Leading Order (NLO) Feynman diagrams are shown (top and bottom diagrams respectively).

At LO, the  $W\gamma$  process is represented by four Feynman diagrams including one FSR, one TGC and two ISR diagrams. Each LO diagram has three vertices. The first calculation of the  $W\gamma$  process with necessary expressions can be found in [21].

The NLO corrections to the amplitude of the  $W\gamma$  process that are shown in Fig. 2.2 are QCD corrections only, which include gluon loops at the same quark line and exchange of a gluon between two different quark lines, however, QED and weak NLO diagrams are also possible. QED corrections involve radiations of extra photons by charged particles, exchange of photons between different charged particles or a photon can be radiated and absorbed by the same charged particle forming a loop. Similarly, weak corrections involve extra virtual  $W$  or  $Z$  bosons. The QCD corrections are the largest among the discussed correction types because the QCD coupling constant is the largest.

A theoretical cross section in particle physics is compared to a measurement result to test the predictions of the model. Also the theoretical cross section is used for producing simulated data. In a simulation (often referred as Monte Carlo or MC), a large set of  $pp$  collisions resulting in a physics process of interest is modeled to create a data set that mimics real data. A typical simulation consists of two parts: the generation of the process and the simulation of particles paths through the detector. The first stage contains a collection of events with final state particles with kinematic quantities distributed according to theoretical predictions for a given process. This stage relies on the theory including the cross section and also all dynamics of the process. The second stage simulates the interaction with media during propagation of particles through the model of the detector as well as the response of detector electronics. In its final form, a simulated dataset has the same format and content of detector signals for each event as real data, and

can undergo the same reconstruction and analysis procedure as real data would.

The most precise theoretical  $W\gamma$  cross section available is the Next-to-Next-to-Leading Order (NNLO) cross section in QCD [22]. The effects of the NNLO correction over the NLO correction and over the LO result are shown in Fig. 2.3 for the transverse mass of the final state particles  $m_T^{l\nu\gamma}$  and for the rapidity difference between a charged lepton and a photon  $\Delta_{l\gamma}$ . The NNLO and NLO theoretical predictions for the photon transverse momentum  $p_T^\gamma$  are overlaid with the 7 TeV ATLAS result. The contribution from higher order corrections is estimated to be  $\pm 4\%$ . However, the NNLO theoretical result was published only recently, in 2015, and no NNLO  $W\gamma$  simulation is available at this time. The simulation used in this analysis is LO + up to two hadronic jets simulation which was found to give the same predictions as the NLO result.

Certain BSM theories predict an enhancement of the contribution from the TGC diagram over the SM prediction. The discussion of these BSM effects and how they affect the  $W\gamma$  process takes place in Ch. 2.4.

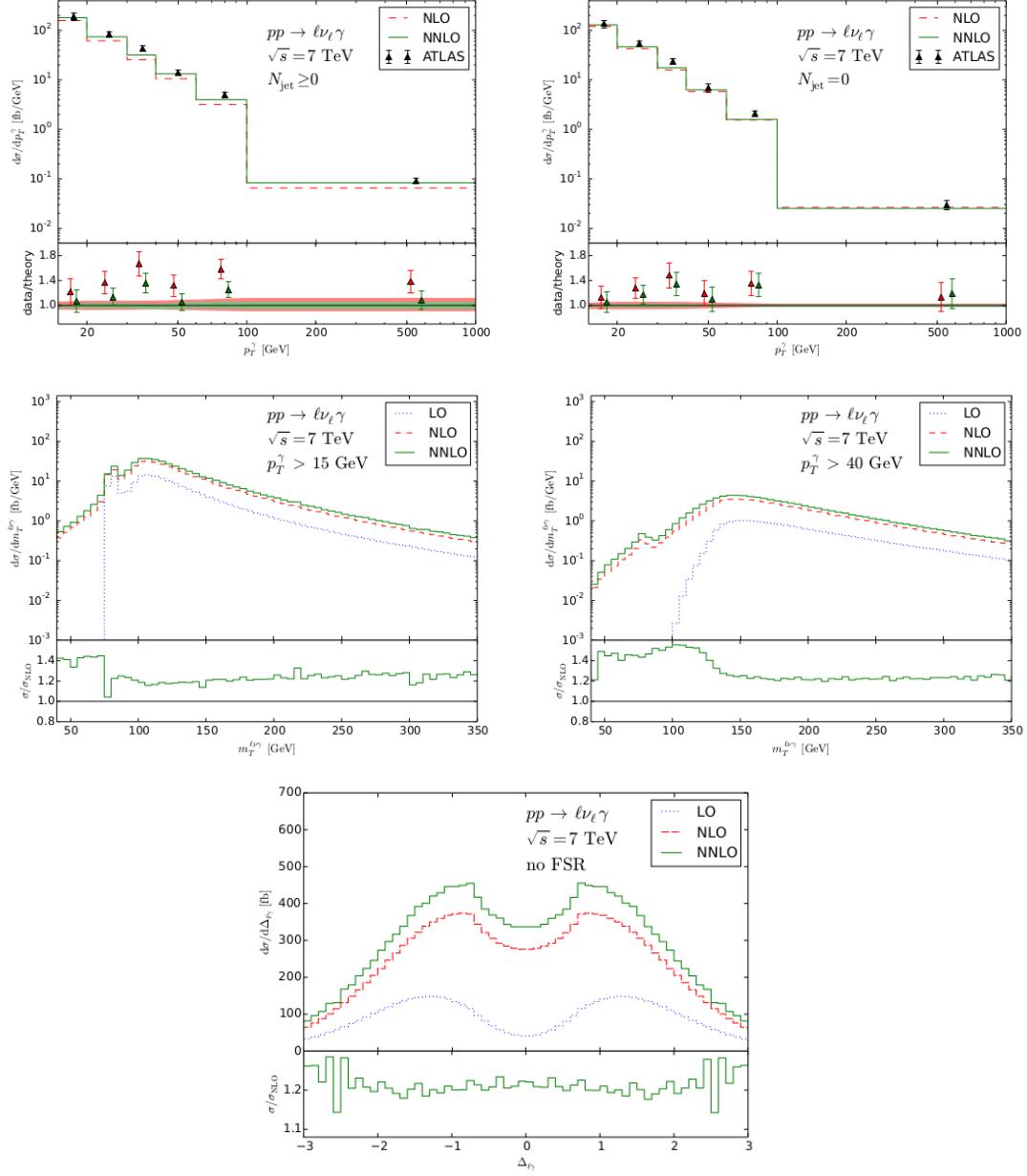


Figure 2.3: Theory spectra. Top: NLO and NNLO  $p_T^\gamma$  spectra of  $W\gamma \rightarrow l\nu\gamma$  at  $\sqrt{s} = 7$  TeV overlaid with ATLAS data for  $N_{jet} \geq 0$  (left) and  $N_{jet} = 0$  (right). Middle: LO, NLO and NNLO  $m_T^{l\nu\gamma}$  spectra of  $W\gamma \rightarrow l\nu\gamma$  at  $\sqrt{s} = 7$  TeV for  $P_T^\gamma > 15$  GeV (left) and  $P_T^\gamma > 40$  GeV (right). Bottom: LO, NLO and NNLO  $\Delta_{l\gamma}$  spectra of  $W\gamma \rightarrow l\nu\gamma$  at  $\sqrt{s} = 7$  TeV.

## 2.4 Anomalous $W\gamma$ Production

Most BSM physics theories predict the existence of particles with masses lying beyond the discovered energy range. If their masses are not accessible even at the accelerators with the highest energies, the direct detection of such particles is not possible. However, loops of heavy particles can affect diagrams of productions of lighter particles. They would give additional contributions to TGC and QGC couplings and, therefore, to the amplitudes to the processes involving TGC and QGC productions. There would be a different number of events produced in the process than one would expect based on SM predictions as shown in Fig. 2.5.

TGC and QGC couplings can be probed by precision measurements of SM processes of diboson and triboson productions because these processes can occur through TGC and QGC. TGC and QGC are represented by vertices with three and four bosons (Fig. 2.4). As discussed in Ch. 2.1, charged TGC and QGC are possible at tree level in the SM while neutral TGC and QGC are not.

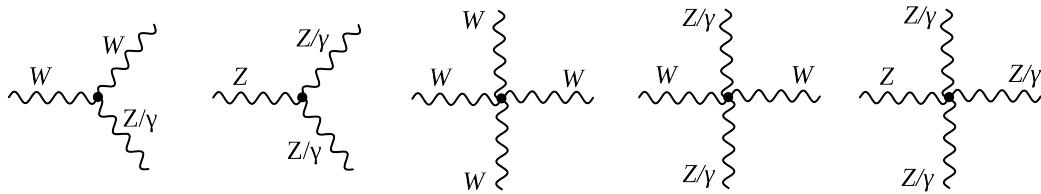


Figure 2.4: Charged TGC (first), neutral TGC (second), charged QGC (third and fourth), and neutral QGC (fifth) vertices.

To account for the effects from the potential loops of heavy particles, we introduce an effective Lagrangian with arbitrary values of coupling constants which can be reduced to the SM Lagrangian if these constants would have their SM values. Introducing the effective Lagrangian makes searches model-independent

because we do not specify particles that form the loops but instead just check whether there is a deviation from the SM prediction in measured observables.

In  $W\gamma$  measurement we can probe  $WW\gamma$  vertex. The most general Lorentz invariant Lagrangian terms of  $WW\gamma$  interaction takes the following form [23]:

$$iL_{eff}^{WW\gamma} = iL_{eff(1)}^{WW\gamma} + iL_{eff(2)}^{WW\gamma} + iL_{eff(3)}^{WW\gamma}, \quad (2.30)$$

where

$$iL_{eff(1)}^{WW\gamma} = e[g_1^\gamma A^\mu (W_{\mu\nu}^- W^{+\nu} - W_{\mu\nu}^+ W^{-\nu}) + \kappa_\gamma W_\mu^+ W_\nu^- A^{\mu\nu} + \frac{\lambda_\gamma}{m_W^2} A^{\mu\nu} W_\nu^{+\rho} W_{\rho\mu}^-], \quad (2.31)$$

$$iL_{eff(2)}^{WW\gamma} = e[ig_5^\gamma \epsilon_{\mu\nu\rho\sigma}((\partial^\rho W^{-\mu}) W^{+\nu} - W^{-\mu} (\partial^\rho W^{+\nu})) A^\sigma + ig_4^\gamma W_\mu^- W_\nu^+ (\partial^\mu A^\nu + \partial^\nu A^\mu)], \quad (2.32)$$

$$iL_{eff(3)}^{WW\gamma} = e[\frac{\tilde{\kappa}_\gamma}{2} W_\mu^- W_\nu^+ \epsilon^{\mu\nu\rho\sigma} A_{\rho\sigma} - \frac{\tilde{\lambda}_\gamma}{2m_W^2} W_{\rho\mu}^- W_\nu^{+\mu} \epsilon^{\nu\rho\alpha\beta} A_{\alpha\beta}], \quad (2.33)$$

where  $e$  is the absolute value of the electron charge,  $A^\mu$  is the photon field,  $W^{\pm\mu}$  are the fields of the  $W^\pm$  bosons,  $W_{\mu\nu} = \partial_\mu W_\nu - \partial_\nu W_\mu$ ,  $A_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ ,  $m_W$  is the mass of the  $W$  boson,  $g_1^\gamma$ ,  $\kappa_\gamma$ ,  $\lambda_\gamma$ ,  $g_5^\gamma$ ,  $g_4^\gamma$ ,  $\tilde{\kappa}_\gamma$ , and  $\tilde{\lambda}_\gamma$  are constants.

Despite seven constants in the extended Lagrangian, only  $\lambda_\gamma$  and  $\kappa_\gamma$  are considered in the aTGC searches. The rest of the constants are fixed to their SM values based on the following considerations. The constants  $g_1^\gamma = 1$  and  $g_5^\gamma = 0$  are fixed to make the Lagrangian obey the electromagnetic gauge invariance for the on-shell photons. The non-zero value of  $g_4^\gamma$  also violates C and P conserva-

tions, and non-zero values of  $g_4^\gamma$ ,  $\kappa_\gamma$ ,  $\tilde{\lambda}_\gamma$  violate the CP conservation law. Such violation parametrizations are not considered in charged TGC measurements, thus, constants  $g_4^\gamma$ ,  $\kappa_\gamma$ , and  $\tilde{\lambda}_\gamma$  are fixed to zero.

The SM values of  $\lambda_\gamma$  and  $\kappa_\gamma$  are  $\lambda_\gamma = 0$  and  $\kappa_\gamma = 1$ . For convenience, the deviation from the SM value is introduced  $\Delta\kappa_\gamma \equiv \kappa_\gamma - 1$ . These two parameters are tested in  $WW\gamma$  aTGC searches because non-zero values of these parameters would not violate any fundamental law.

The most significant effects of aTGC would appear at high energy scales. Figure 2.5 shows this effect in  $P_T^\gamma$  spectrum of 7 TeV  $W\gamma \rightarrow \mu\nu\gamma$  measurement. As seen in Fig. 2.5, the spectrum with non-zero values of aTGC constants at low  $P_T^\gamma$  coincides with the SM prediction but for higher  $P_T^\gamma$  the disagreement appears.

A common approach to aTGC searches is to measure the spectrum of a kinematic parameter highly correlated with the energy of a final state particle or a system of final state particles. For  $W\gamma$  process, the most sensitive variable is  $P_T^\gamma$ . Examining this spectrum allows us to probe and constrain aTGC coupling constants. Chapter 2.5 reviews the experimental results to date on constraining aTGC coupling constants of the  $WW\gamma$  vertex.

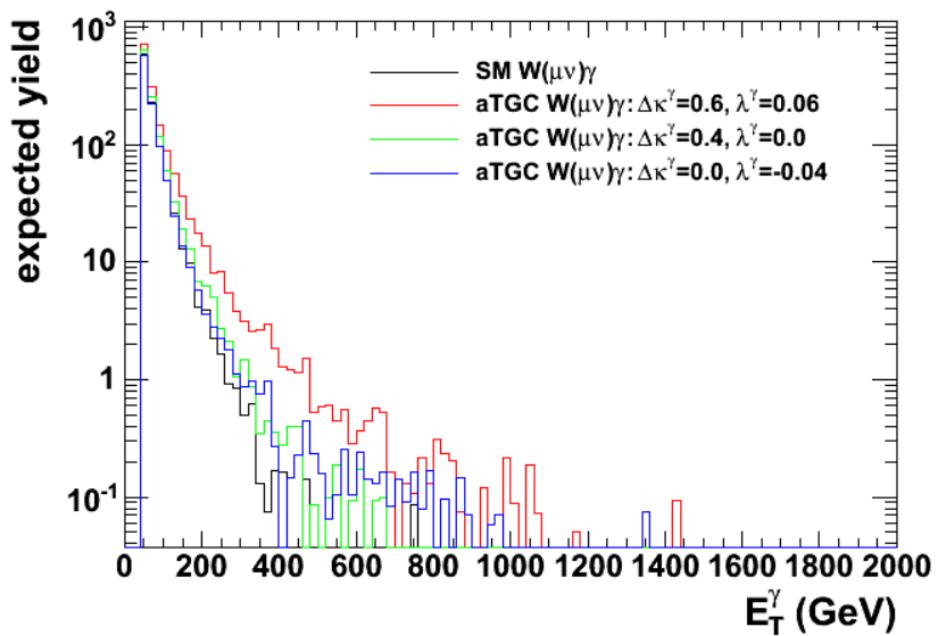


Figure 2.5: Distributions of  $P_T^\gamma$  in simulated  $W\gamma \rightarrow \mu\nu\gamma$  events with different values of aTGC constants at LHC energy of  $\sqrt{s} = 7$  TeV. Source of figure: [2].

## 2.5 A brief history of $W\gamma$ measurements

aTGC parameters of the  $WW\gamma$  vertex can be probed in measurements of  $W\gamma$ ,  $WW$ ,  $WZ$  processes. Limits on the  $\Delta\kappa_\gamma$  and  $\lambda_\gamma$  constants obtained by different experiments are summarized in Fig. 2.6. The summary includes the combination results from Do [24] and LEP [25] as well as results of several individual measurements by ATLAS and CMS including  $W\gamma$  at  $\sqrt{s} = 7$  TeV [5], [4],  $WW$  at  $\sqrt{s} = 7$  and 8 TeV [26], [27], [28], and  $WV$  at  $\sqrt{s} = 7$  and 8 TeV [29], [30] measurements.

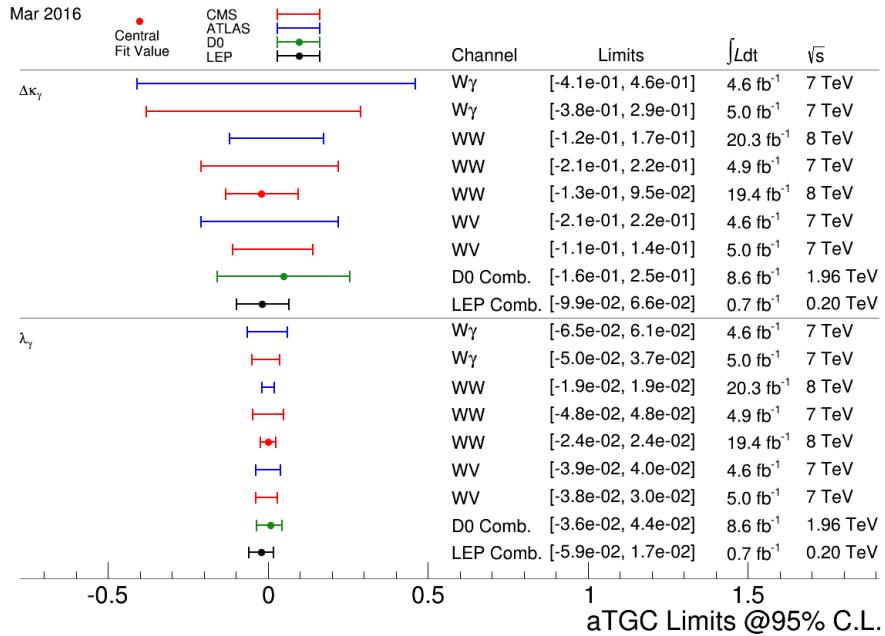


Figure 2.6: Summary of limits on the  $WW\gamma$  aTGC coupling constants. Figure from [3].

The most recent measurements of  $W\gamma$  production were performed by CMS [4] and ATLAS [5] collaborations with  $pp$  collisions at  $\sqrt{s} = 7$  GeV collected in 2011. Both collaborations considered two channels:  $W\gamma \rightarrow \mu\nu\gamma$  and  $W\gamma \rightarrow e\nu\gamma$ .

Diboson processes are rare in  $pp$ -collisions and analysts have to filter out events of their interest from many processes which are more likely to happen. To do

that, a variety of selection criteria are applied which reject most of the background events to increase the signal fraction in the selected sample as much as possible. However, even after all possible selection criteria are applied, the majority of selected events are still background events and it is not possible to reduce the background any further without also significantly reducing signal.

The major source of such background is the fake photon background where hadronic jets are misidentified as photons. Such events originate mostly from  $W+jets$ , but  $Z+jets$  and  $t\bar{t}+jets$  events contribute to this source of background as well. In the electron channel there is one more significant background that is the fake photon background where electron is misidentified as a photon. Such events are coming from  $Z+jets$  events. For the muon channels this background is small. Other sources of backgrounds for both channels include real- $\gamma$ , fake lepton + real photon and fake lepton + fake photon backgrounds. The major source of real- $\gamma$  background is the  $Z\gamma$  process where a final state lepton and a photon mimics the  $W\gamma$  final state. Fake lepton + real photon background originates from the  $\gamma+jets$  process where a jet is misidentified as a lepton. Fake lepton + fake photon backgrounds come from dijet and multijet events where one of the jets is misidentified as a lepton and the other one is misidentified as a photon. The probability of a jet to be misidentified as a lepton is very small, therefore fake lepton + real photon and fake lepton + fake photon backgrounds are negligible.

$P_T^\gamma$  spectra are measured because this variable is the most sensitive to the potential aTGC. The  $P_T^\gamma$  spectra of the selected events in data superimposed with selected events in the simulation of the signal and estimated background contribution for the muon and electron channels are shown in Fig. 2.7 for CMS and in Fig. 2.8 for ATLAS measurement. Both measurements show a good agreement between data and the simulation.

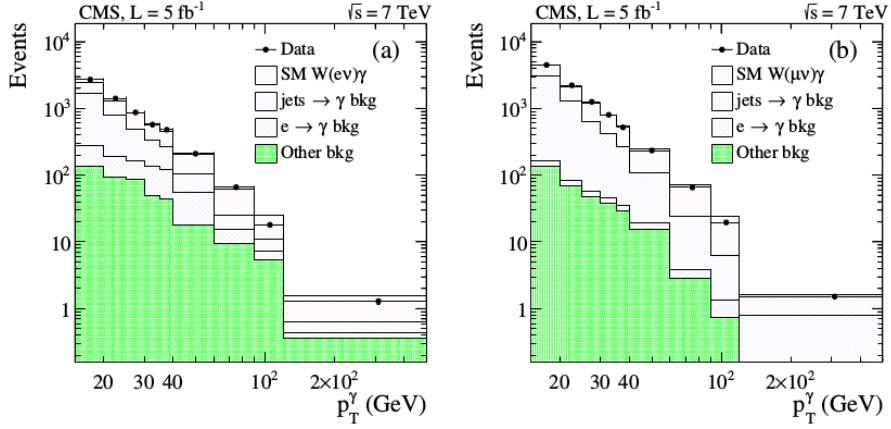


Figure 2.7: The distribution of the  $p_T^\gamma$  of  $W\gamma$  candidates in the analysis of 7 TeV CMS data. Data vs signal simulation + background estimates. Left:  $W\gamma \rightarrow e\nu\gamma$ , right:  $W\gamma \rightarrow \mu\nu\gamma$  [4].

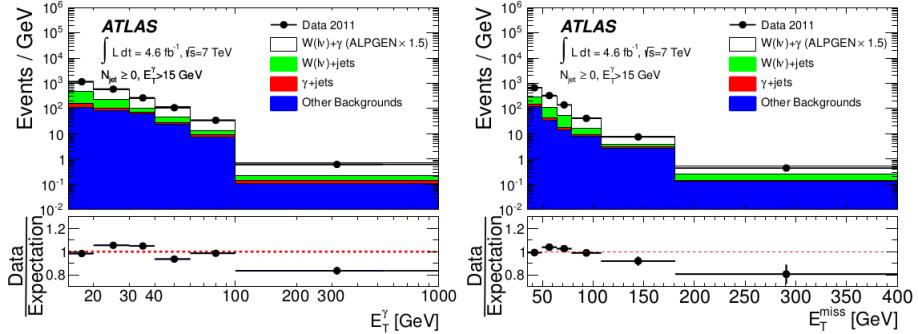


Figure 2.8: The distribution of the photon transverse momentum (left) and missing transverse momentum (right) of  $W\gamma$  candidates in the analysis of 7 TeV ATLAS data. Data vs signal simulation + background estimates [5].

The phase space restrictions of  $W\gamma$  measurements come from the considerations of the detector acceptance (Ch. 3.2), reducing heavily background-dominated regions and theoretical considerations such as to avoid divergence of the cross section and to reduce ISR and FSR contributions to the cross section.

CMS provides measurements of the  $P_T^\gamma$  spectrum, the total cross section within the phase spaces of  $\Delta R > 0.7$ ,  $P_T^\gamma > 15$  GeV,  $P_T^\gamma > 60$  GeV and  $P_T^\gamma > 90$  GeV.

ATLAS, in addition to the  $P_T^\gamma$  spectrum, total cross section and limits, provides the differential cross section and cross section with different number of associated jets. The phase space restrictions for ATLAS measurement include requirements on charged lepton kinematics  $P_T^l > 25$  GeV,  $|\eta_l| < 2.47$ , requirements on the transverse momentum of a neutrino  $P_T^\nu > 35$  GeV, photon kinematics  $P_T^\gamma > 15$  GeV,  $|\eta^\gamma| < 2.37$ , photon isolation fraction  $\epsilon_h^P < 0.5$  and lepton-photon separation  $\Delta R(l, \gamma) > 0.7$ . For the differential cross section in number of associated jets, the requirements on jets kinematics and jets separation from leptons and photons are also applied:  $E_T^{jet} > 30$  GeV,  $|\eta^{jet}| < 4.4$ ,  $\Delta R(e/\mu/\gamma, jet) > 0.3$ . No evidence of new physics is observed.

The estimated cross sections with any number of associated jets for  $P_T^\gamma > 15$  GeV are

$$\sigma(pp \rightarrow W\gamma \rightarrow l\nu\gamma) = 37.0 \pm 0.8 \text{ (stat.)} \pm 4.0 \text{ (syst.)} \pm 0.8 \text{ (lumi.) pb} \quad (2.34)$$

and

$$\sigma(pp \rightarrow W\gamma \rightarrow l\nu\gamma) = 2.77 \pm 0.03 \text{ (stat.)} \pm 0.33 \text{ (syst.)} \pm 0.14 \text{ (lumi.) pb} \quad (2.35)$$

for CMS and ATLAS respectively. The results agree with NLO MCFM [31] predictions of  $31.81 \pm 1.8$  pb for the phase space used by CMS and of  $1.96 \pm 0.17$  pb for the phase space used by ATLAS.

In addition to the cross sections, both CMS and ATLAS provide limits on aTGC coupling constants  $\Delta\kappa_\gamma$  and  $\lambda_\gamma$ . To do that, samples with non-zero aTGC coupling

constants are generated, run through the whole reconstruction and selection procedures, and compared to the measured results of  $P_T^\gamma$  spectra. The results on one-dimensional limits are quoted in Fig. 2.6 while the results on two-dimensional limits can be found in [5], [4].

In this dissertation we are measuring total and differential  $d\sigma/dP_T^\gamma$  cross section. While the aTGC limits are not derived in this dissertation, the measured differential cross section can be used to derive them. The measurement details and results are described in Chapter 5.

# Chapter 3

## Experimental Setup

The measurement reported in this dissertation is based on data collected by the CMS detector from the LHC  $p p$  collisions in 2012 at  $\sqrt{s} = 8$  TeV. The experimental setup for this measurement includes the LHC and the CMS detector that are described in Ch. 3.1 and Ch. 3.2 respectively.

### 3.1 Large Hadron Collider

The LHC [12, 32, 33] is the largest particle accelerator and the most ambitious particle physics research facility ever built. The LHC accelerates two particle beams up to near the speed of light. The beams travel in opposite directions, each in its own beam pipe, in ultrahigh vacuum. The beam is made up of protons which are grouped as bunches separated by several meters from each other. Each bunch contains  $10^{11}$  protons. The bunches of protons are accelerated by varying electromagnetic fields, focused by superconducting quadrupole magnets and steered by dipole magnets. The bunches collide at fixed collision points where particle detectors are placed. Particles are produced in the collisions and registered by the detectors to be subsequently used to accomplish physics goals of the experiments.

The LHC is located in the tunnel at the France-Switzerland border. The tunnel is located as deep as 175 meters underground, its circumference is about 27 km.

Before entering LHC, particle beams go through several stages of acceleration, and the LHC is the final machine of the chain of the CERN's accelerator complex (Fig. 3.1). Protons are extracted from hydrogen atoms, are accelerated by Linac2 to energies of 5 MeV, and are then injected into the Proton Synchrotron Booster (PSB) where they reach energies of 1.4 GeV. After that, protons are sent to PS and then to Super PS (SPS) where they are accelerated up to 25 GeV and 450 GeV respectively. Finally, protons enter the LHC and are accelerated to reach their collision energies of several TeV per beam. Besides protons, the complex also accelerates and collides lead ions. However, in this dissertation we analyze data from  $pp$  collisions only.

Six detectors are installed at the LHC to detect products of hadron collisions and to perform the measurements of the LHC physics program. ATLAS and CMS

## CERN's Accelerator Complex

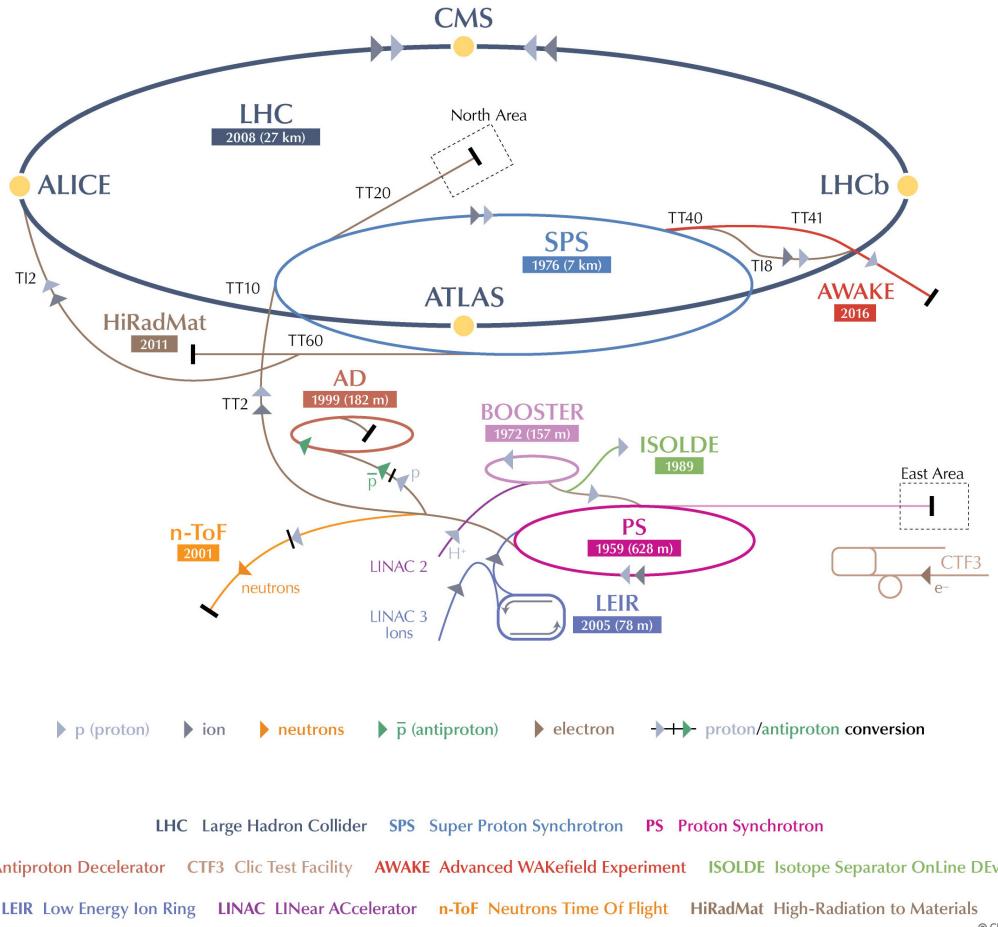


Figure 3.1: CERN's accelerator complex [6].

are general purpose detectors designed to explore a broad spectrum of particle physics questions within and beyond the SM, LHCb specializes in the physics of B mesons, ALICE is designed to detect products of heavy ion collisions, LHCf and TOTEM are small detectors with very specialized research goals. LHCf and TOTEM are installed close to the ATLAS and CMS collision points respectively and are referred as forward detectors.

The design collision energy of the LHC is  $\sqrt{s} = 14$  TeV which corresponds to 7 TeV per beam. However, several lower energy points were probed. In 2010-2011 the LHC operated at an energy of 3.5 TeV per beam which was already higher than the energy of any other collider. In 2012 the beam energy was increased up to 4 TeV. In 2013-2014 the LHC was shut down for upgrades. Collisions were restarted at 6.5 TeV in 2015 and continued at this energy in 2016. At any energy point, both LHC beams have equal energies.

All critical measurements performed at lower energies are also repeated at higher energies. For the BSM searches, the ability to probe higher energy scales increases our chances for a discovery. For a SM cross section measurement, it needs to be done at all energies and compared to the theory since cross sections evolve with energy (Fig. 3.2). While cross sections of parton-parton collisions typically decrease with energy,  $p p$  or  $\bar{p} p$  cross sections increase because as we go higher in  $\sqrt{s}$ , more partons in a given protons have enough energy to produce a certain type of interaction. This enables the observation of rarer processes as we increase energy.

In addition to the beam energy, there are many other collider parameters which reflect the ability of the collider to achieve stated goals. A brief summary of them is available in Tab. 3.1. One of the most critical parameters of an accelerator is the luminosity which determines how many interesting events can be produced (Ch. 2.2). The instantaneous luminosity is determined by the following expression [1]:

$$L^{inst} = f \frac{n_1 n_2}{4\pi\sigma_x\sigma_y} \quad (3.1)$$

where  $n_1$  and  $n_2$  are numbers of particles in colliding bunches,  $f$  is a frequency of collisions,  $\sigma_x$  and  $\sigma_y$  characterize sizes of overlapping parts of colliding beams in

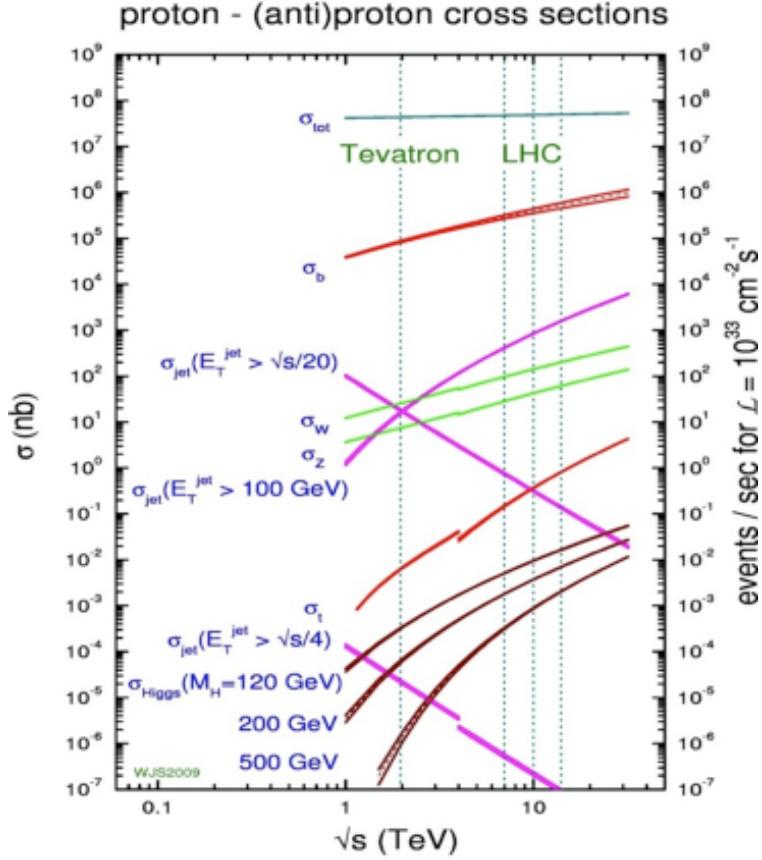


Figure 3.2: Cross sections of different processes in  $p\bar{p}$  and  $\bar{p}p$  collisions [7].

horizontal and vertical directions. The instantaneous luminosity multiplied by a cross section of a process gives an event rate (Eq. 2.26) of this specific process. If we know the instantaneous luminosity and the theoretically predicted cross section of the process, we can estimate how many events per unit time of this particular process will be produced by our experiment. To estimate how many events of the process will be produced during a certain time period, we have to use the integrated luminosity which is an integral of the instantaneous luminosity over time:

$$L = \int L^{inst} dt \quad (3.2)$$

The integrated luminosity of a data sample is a measure of the size of the data sample.

The integrated luminosity of the LHC for  $pp$  collisions for different years of the operation is shown in Fig. 3.3. Run I of the LHC operation covers run periods of 2010-2012. While running at the energy of  $\sqrt{s} = 7$  TeV, LHC delivered  $L = 45 \text{ pb}^{-1}$  and  $L = 6.1 \text{ fb}^{-1}$  of data in 2010 and 2011 year respectively. In 2012 the working energy of LHC was  $\sqrt{s} = 8$  TeV, and the integrated luminosity was  $L_{int} = 23.3 \text{ fb}^{-1}$ . After a long shutdown, LHC was upgraded for Run II, to operate at  $\sqrt{s} = 13$  TeV in 2015 and delivered  $L_{int} = 4.2 \text{ fb}^{-1}$  of data by the end of 2015. In 2016 LHC continued operating at  $\sqrt{s} = 13$  TeV and delivered the integrated luminosity of  $L_{int} = 41.1 \text{ fb}^{-1}$  [34].

The measurement of this dissertation is performed at the energy of 4 TeV per beam or the center of mass energy  $\sqrt{s} = 8$  TeV with  $19.6 \text{ fb}^{-1}$  of data collected in 2012. The same process was measured at  $\sqrt{s} = 7$  TeV with about four times less data by both CMS and ATLAS. These measurements are discussed in greater detail in Ch. 2.5.

Table 3.1: Main parameters of LHC [12]

Circumference	27 km
Dipole operating temperature	1.9 K
Number of magnets	9593
Number of main dipoles	1232
Number of main quadrupoles	392
Number of RF cavities	8 per beam
Nominal energy, protons	7 TeV
Nominal energy, lead ions	2.76 TeV per nucleon
Peak magnetic dipole field	8.33 T
Min. distance between bunches	7 m
Design luminosity	$10^{34} \text{ cm}^{-2} \text{ s}^{-1}$
No. of bunches per proton beam	2808
No. of protons per bunch (at start)	$1.1 \times 10^{11}$
No. of collisions per second	600 millions

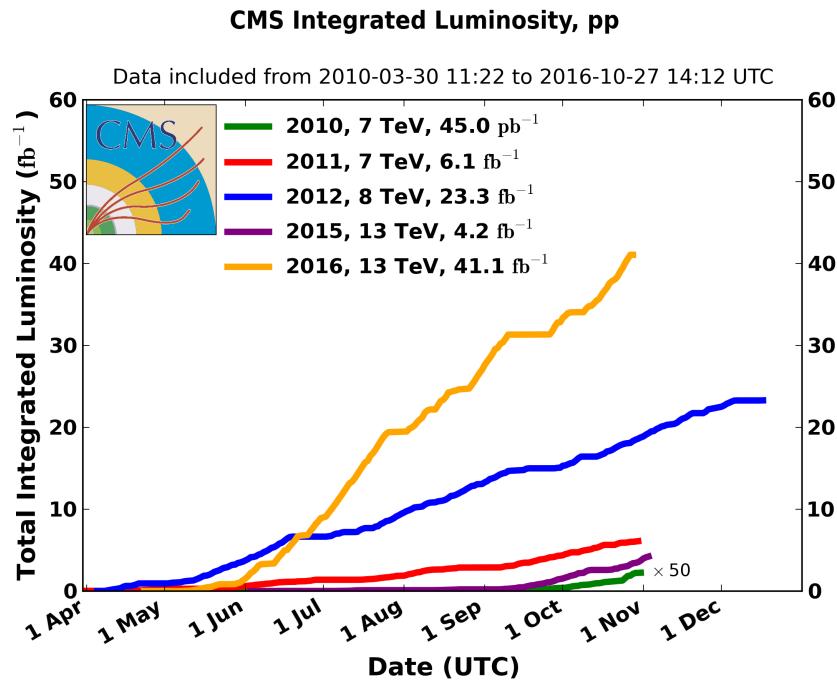


Figure 3.3: LHC integrated luminosity by year [8].

## 3.2 Compact Muon Solenoid

### 3.2.1 Introduction

CMS is a general-purpose detector designed to register particles with energies of tens and hundreds GeV that are produced in  $pp$  collisions at the LHC [35]. The CMS detector is cylindrically symmetric with the particle beam as the axis. Cartesian, cylindrical and spherical coordinates are all used to describe the CMS geometry, depending on the context. The  $x$ -axis of the CMS points towards the center of the LHC ring while the  $y$ -axis points vertically up. The orientation of the  $z$ -axis corresponds to the counterclockwise direction of the LHC beam (Fig. 3.4). Cylindrical coordinates are defined as  $r = \sqrt{x^2 + y^2}$ ,  $\phi = \arctan(y/x)$ . Instead of the polar angle  $\theta$ , it is more convenient to use the pseudorapidity  $\eta = -\ln \tan \theta/2$ . A pseudorapidity ranges from  $\eta = -\infty$  to  $\eta = +\infty$  with  $\eta = \pm\infty$  for directions parallel to the beam axis and  $\eta = 0$  for a direction perpendicular to the beamline. This variable is convenient for measurements because for typical physics process in  $pp$  collisions the created particles tend to be distributed uniformly in  $\eta$ . Another important feature making  $\eta$  a convenient variable is that  $\Delta\eta$  values are Lorentz invariants.

Certain particles produced in a collision cannot be registered by CMS due to geometrical limitations of the detector. Charged particles with very low momenta have very large track curvatures and cannot leave the beam pipe. Particles that have trajectories close to parallel to the beamline also cannot be registered by CMS. The range of geometrical and kinematic parameters of a particle that allows it to be registered by the detector is called the detector acceptance. The acceptance of the CMS in  $\eta$  is limited and varies from  $|\eta| < 2.4$  to  $|\eta| < 5.3$  depending on a subdetector (Fig. 3.5, top).

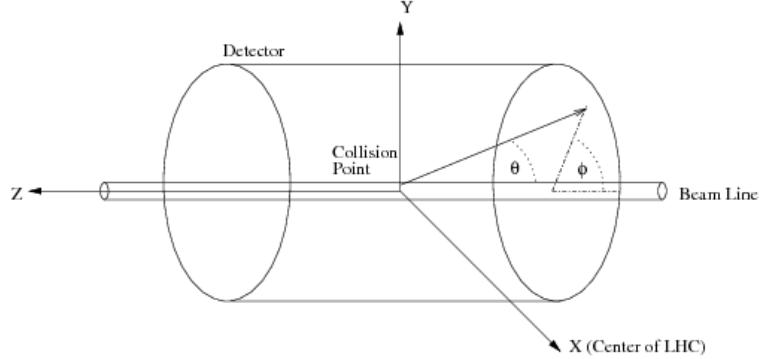


Figure 3.4: CMS coordinate system.

The detector consists, from the inner to the outer layer, of a tracking system, an electromagnetic calorimeter (ECal), a hadronic calorimeter (HCal), a magnet and a muon system. A slice of CMS in the  $r$ - $\phi$  plane is shown in Fig. 3.6. Most subdetectors have three geometrically distinct components: the cylindrical part at the central region (barrel) and the disk-like structures at each end (endcaps). Barrel and endcap regions vary depending a subdetector but approximately barrel covers  $|\eta| < 1.5$ , and endcap  $|\eta| > 1.5$ .

Most heavy particles produced in a collision decay immediately, and we detect their long-lived decay products including electrons, photons, muons, neutral or charged hadrons. Particles which can be detected by CMS are referred as “visible” particles in contrast to “invisible” particles which cannot be detected by CMS because their probabilities to interact with any part of the detector are very low. The SM example of an invisible particle is a neutrino.

We can identify the type of particle by the trace it leaves in different subdetectors. Charged particles interact with the substance of the tracking system which performs several position measurements of the particles. The sequence of these position measurements is called a track. Neutral particles do not leave any trace in

the tracking system because they do not ionize atoms.

Thus, electrons and positrons leave tracks in the tracking system while photons do not. Both these types of particles induce showers in the ECal of the same shapes, and are distinguished by having or not having a spatially matching track. Hadrons normally travel through the ECal undisturbed and induce a hadronic shower in the HCal (Ch. 3.2.5). Charged and neutral hadrons are distinguished from each other by linking or not linking to the tracks, similarly to how electrons are distinguished from photons. Muons are the only particles that penetrate the ECal, the HCal and the magnet and leave tracks in the CMS muon system. Neutrinos are not directly detected by CMS.

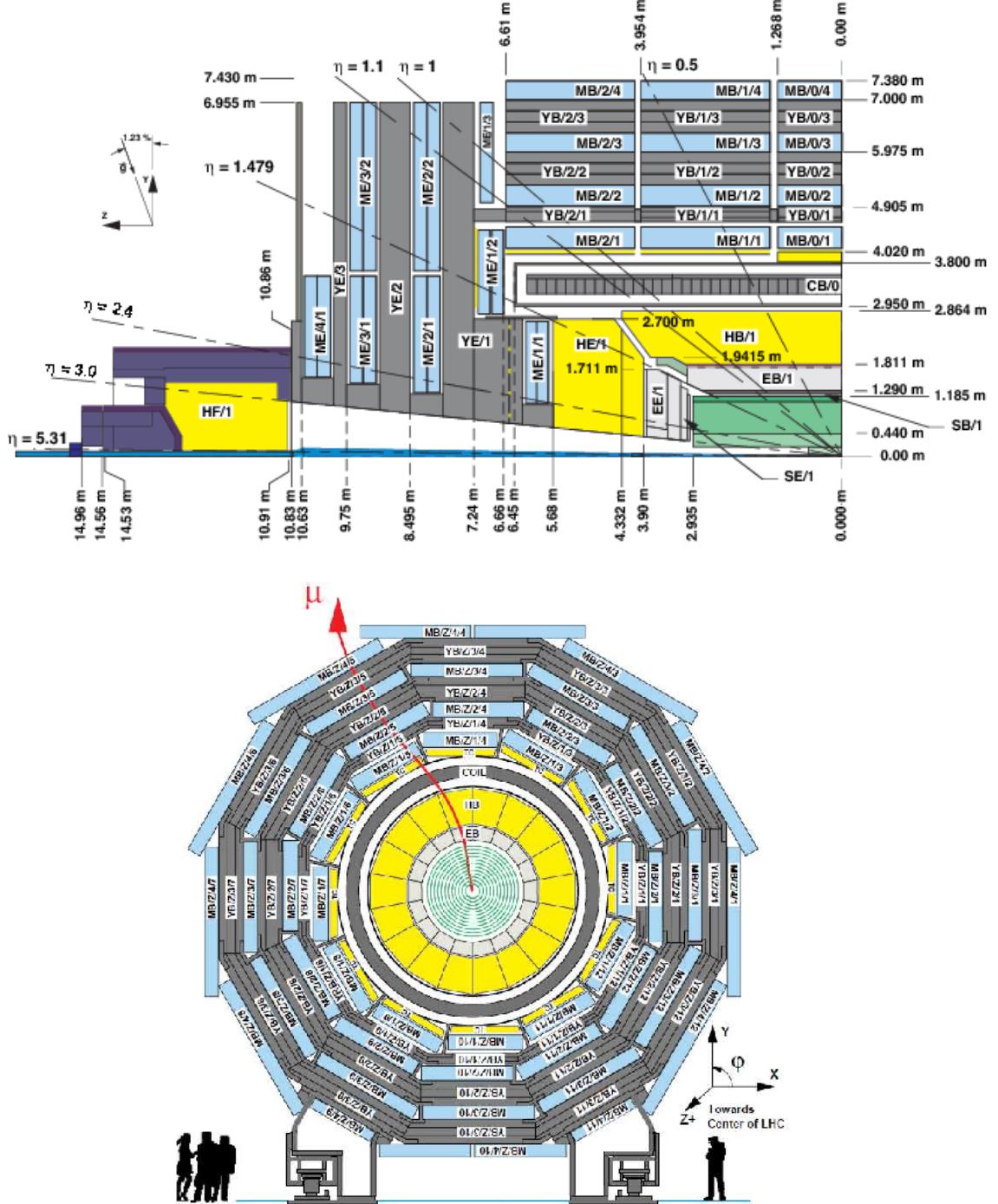


Figure 3.5: CMS detector, schematic view. Top:  $r - z$  plane, bottom:  $r - \phi$  plane at  $z = 0$  [9]. The tracking system is shown in green, EB and EE of ECal are shown in gray, HB, HE and HF of HCal are shown in yellow. HCal is surrounded by a magnet which is shown in gray and white. Muon stations and return yokes are located outside of the magnet and are shown in blue and gray. A red line at the bottom plot is a muon trajectory demonstrating a typical muon to penetrate the whole CMS detector. People at the bottom illustrate the scale of the CMS detector.

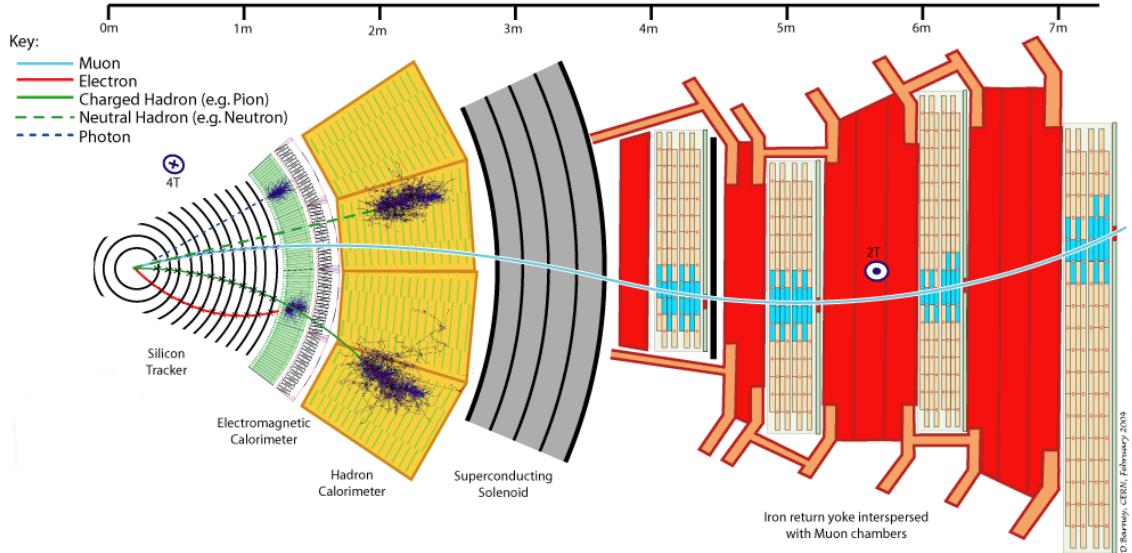


Figure 3.6: CMS detector, a schematic view of a segment in the  $r\phi$  plane at  $z = 0$ . Traces left by muons, electrons, photons, charged and neutral hadrons in different subdetectors are shown.

All subdetectors are essential for the  $W\gamma$  measurement, and the remainder of this chapter describes the subdetectors in greater detail. Muons and electrons, which we have as final state particles in the  $W\gamma$  measurement, are both affected by the CMS magnetic field, allowing the tracking system and the muon system to measure their trajectory parameters and momenta. ECal measures energy of electrons and photons and is also used to determine a photon's trajectory. The HCal is essential to determine the missing transverse energy which is a measure of a neutrino transverse momentum.

### 3.2.2 Magnet

A magnetic field in a particle detector is necessary to measure momenta of charged particles by track curvatures. The higher the momentum is, the less a particle trajectory is affected by the magnetic field. In the plane, transverse to the beamline, this relation is:

$$R = \frac{P_T}{qB}, \quad (3.3)$$

where  $R$  is a radius of a projection of a charged particle's trajectory to the transverse plane,  $P_T$  and  $q$  are the particle's transverse momentum and electric charge, and  $B$  is the magnetic field. In CMS, the tracking system measures momenta of all charged particles. Also, the muon system measures momenta of muons.

The CMS magnet is placed between the HCal and the muon system. The magnet is made of superconducting wires that are cooled to  $-268.5^0\text{C}$  by a cryogenic system based on a liquid helium flow [36]. An electric current flowing in the wires creates a uniform field of  $B = 4\text{T}$  inside the solenoid, for the tracking system, and also provides a smaller magnetic field of a certain configuration outside the solenoid, for the muon system. The stronger field in the tracking system is necessary because of higher track density and smaller size relative to the muon system.

### 3.2.3 Tracking System

The tracking system measures parameters of charged particle trajectories and their momenta, reconstructs primary and secondary vertices. The tracking system is designed to disturb a particle as little as possible when it passes through to be able to accurately measure its energy deposit in the ECal or HCal or, in case of a muon, accurately reconstruct a track in the muon system. The goals of little

distortion and high precision track reconstruction at the same time are achieved by CMS algorithms being capable of reconstructing a trajectory with just a few position measurements (“hits”), each as accurate as  $\sim 10 \mu\text{m}$  in the transverse plane and  $\sim 30 \mu\text{m}$  in the longitudinal direction [37].

Tracks that originate from proton collisions, collision tracks, start at the center and then cross the layers of the tracking system. Charged particles take helical paths in the magnetic field. Tracks are straight in the  $r - z$  plane and curved by the magnetic field in the  $r - \phi$  plane. The acceptance of the tracker system in the  $r - z$  plane is geometrically limited by the absolute value of the pseudorapidity  $|\eta| \leq 2.5$ .

The tracking system consists of silicon pixels and silicon strips (Fig. 3.7). The pixel tracker is the closest subsystem of CMS to the collision point. Thus, it experiences the largest particle flux: at 8 cm from the collision point the flux is about 10 million/(cm<sup>2</sup>s), and the pixel detector with its 65 million pixels is capable of reconstructing all these tracks. It consists of three cylindrical layers of pixel sensors in the barrel with radii of 4 cm, 7 cm and 11 cm which are referred as barrel pixel subdetectors (BPIX) and four disks in the endcap, two disks at each side, which are referred as forward pixel subdetector (FPIX). Pixel modules provide 3D position measurements as well as some of the strip modules while the other strip modules provide 2D position measurements.

The strip tracker is placed right outside the pixel tracker and occupies the detector volume up to 130 cm from the beam axis. The strip tracker consists of four parts: the tracker inner barrel (TIB), the tracker inner disks (TID), the tracker outer barrel (TOB) and the tracker endcap (TEC) as shown in Fig. 3.7.

The resolution of track parameters depends on a type of the reconstructed particle, its transverse momentum and pseudorapidity. For example, a momentum resolution of an isolated muon with  $P_T^\mu = 100 \text{ GeV}$  and  $\eta^\mu = 0$  is 2% and increases

with  $|\eta|$  as shown in [37], Fig. 14.

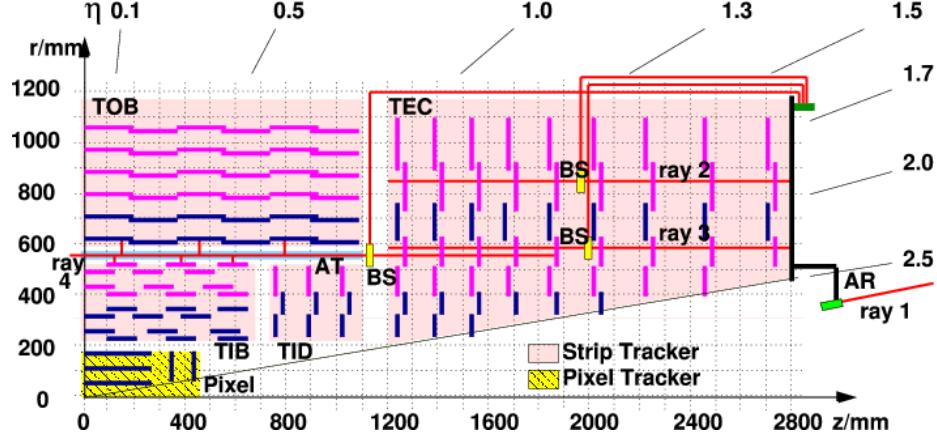


Figure 3.7: Slice of the CMS tracking system in the  $r$ - $z$  plane. Pixel modules and strip modules shown in blue provide 3D position measurements. Strip modules shown in pink provide 2D position measurements.

### 3.2.4 Electromagnetic Calorimeter

The ECal is placed between the tracking system and the HCal. It is made of high-density lead tungstate crystals arranged in a barrel section and two endcap sections. The crystals are scintillators. When an electron or a photon passes through the scintillators, it initiates an electromagnetic shower. A photon produces an  $e^+e^-$  pair while an electron emits a photon. These processes continue as long as photon has enough energy to produce an  $e^+e^-$  pair. With this mechanism, the whole energy of the entering particle converts to light. The scintillated light is then amplified by photomultipliers. After that, signals are digitized and taken away by fiber optic cables.

The ECal measures the energy of electrons and photons and parameters of their trajectories. In order to distinguish between electrons and photons, it is necessary

to perform spatial matching to the track in the tracking system. If there is a track, then the particle is an electron (or positron), otherwise the particle is a photon.

It is important for the ECal to be able to distinguish between a single energy photons of high energy and pairs of almost collinear photons of lower energy e.g. from a  $\pi^0$  decay. It is especially difficult in the endcap sections where the angle between two photon trajectories is small. It is achieved with ECal preshower detectors (PS) which are located in front of the endcaps and have  $\sim 15$  times smaller granularity. Such a small granularity is achieved by making preshower of two lead planes followed by silicon sensors. The ECal PS provide extra spatial precision.

The ECal energy resolution depends on photon or electron energy and of the ECal pseudorapidity region. The resolution is 2%-5% for electrons from  $Z \rightarrow ee$ , and 1%-5% for photons from  $H \rightarrow \gamma\gamma$  [38].

### 3.2.5 Hadron Calorimeter

The HCal measures the energy of charged and neutral hadrons. It consists of the barrel, endcap and forward parts: HB, HE and HF in Fig. 3.5, top, respectively. HCal stops all hadrons passing through, thus, it extends to  $|\eta| = 5.3$  for HF.

The HCal is a sampling calorimeter. It consists of alternating layers of brass absorbers and plastic scintillators. When a hadron hits an absorber, it induces a hadronic shower. Interacting strongly with the absorber's nucleons, the hadron produces secondary hadrons. When hadrons reaches the layer of the scintillator, they interact with the scintillator's nucleons, exciting the atoms. Then atoms in the scintillator release light that is collected on optic fibers and passed to the readout system. The total amount of light released in a certain region of the HCal is a measure of hadron's energy.

### 3.2.6 Muon System

Muons, unlike other visible particles, are not stopped by CMS calorimeters because they neither induce an electromagnetic shower in the ECal nor a hadronic shower in the HCal. The muon system, which is placed outside the magnet and which is the largest in spatial size part of the CMS detector, is designed to register muons.

There are four concentric layers of muon detectors (stations) and the iron return yoke between them. Muons induce several hits in the muon stations which are later fitted and matched to the tracking system measurements to provide the best possible resolution in the measurements of the muon's trajectory and momentum.

There are three types of muon chambers used in the CMS muon system: drift tubes (DTs), cathode strip chambers (CSCs) and resistive plate chambers (RPCs) (Fig. 3.8). Overall, there are 1400 muon chambers including 250 DTs, 540 CSCs and 610 RPCs.

The system of DTs measures positions of muons in the barrel. Each DT chamber is about 2 m by 2.5 m in size. A chamber consists of 12 layers of aluminum which are arranged in groups of four. There are up to 60 DTs in a layer. The middle group of layers measures z-coordinate and two other groups determine the perpendicular coordinate. The DT's volume is filled with a gas, and there is a wire inside. When a charged particle passes through the volume, it ionizes atoms. Released electrons drift in the electric field to the positively-charged wire. The position along the wire is registered, and the distance of the muon away from the wire is calculated providing measurements of two coordinates of the position of the muon.

CSCs are placed in the endcap regions. CSCs are arrays of anode wires crossed by copper cathode strips placed in a gas volume. When a charged particle penetrates the gas volume, it ionizes the gas. Electrons drift to the wires while ions

move to the strips, and charge pulses are induced on wires as well as on strips. Strips are perpendicular to wires. Thus, we measure two coordinates for each particle.

RPCs are parallel capacitors made of high-resistivity plastic plates with a space between them filled with gas. RPCs provide quick measurements of muon momenta. A muon passing through the RPC ionizes gas atoms. Released electrons ionize more atoms inducing an avalanche in the electric field. Electrodes receive signal and pass it to external strips that provide a quick measurement of the muon's position which is subsequently transformed to the momentum measurement by the trigger's electronics.

The momentum resolution of the muon system is  $\sim 10\%$  for muons with  $P_T^\gamma = 10$  GeV, however, in a conjunction with the tracking system it improves to  $\sim 1\%$ . A detailed documentation of the muon system performance is available at [39].

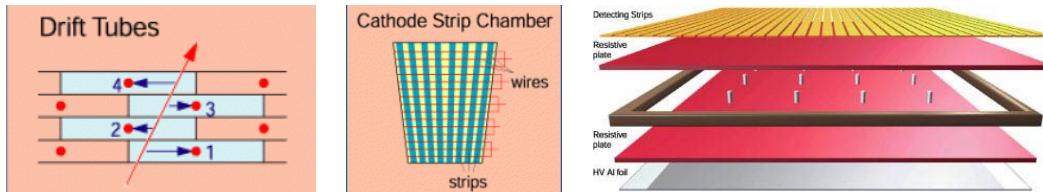


Figure 3.8: Components of the CMS muon system. Left to right: drift tubes (DTs), cathode strip chambers (CSCs), resistive plate chambers (RPCs).

### 3.2.7 Triggering and Data Acquisition

At peak luminosity, CMS experiences 40 million proton-proton bunch crossings per second that come in bunches separated by 25 ns. It is not technically feasible to read out all these events. Moreover, we do not need most of these events for a

physics measurement because most of them have not resulted from an interesting physics process. We have resources to store about 400 events out of 40 million, and that is why we need a trigger system that quickly decides what the best 400 events are.

If the triggers were not strict enough, and we would select 400 events too quickly, e.g., in 1/10 s, then we would not process the remaining 90% of events provided by LHC in a given second and we would lose 90% of potentially interesting events.

If the triggers were too strict, we would select, e.g., 40 events per second, not 400 and lose the potential to store and process data by 90% which would significantly reduce our chances for discovery and increase statistical uncertainties for precision measurements.

Thus, the challenge of the trigger system is to select the best 400 events per second and to do this quickly to be able to process every single event. To achieve this goal, a two-level trigger system was developed consisting of the Level 1 trigger (L1T) and the High Level Trigger (HLT) [40].

L1T is a hardware based trigger. It uses information from the ECal, HCal and muon system. L1T reduces the frequency of coming events from 40 MHz to 100 kHz. Events that did not pass the L1T are lost forever while events that pass the L1T are temporarily stored to be checked by the HLT.

HLT is a software-based trigger. It uses information from all subdetectors and runs fast reconstruction and identification algorithms to determine types of particles and their kinematics. It reduces the event rate to 400 Hz. Events that did not pass HLT are lost forever. Events that pass HLT are arranged into appropriate datasets depending on HLT selection criteria they passed and stored for physics measurements.

There is a large variety of triggers to capture the SM and new physics processes

of our interests. Typical trigger examples include at least one or two leptons or one or two jets with  $P_T$  higher than a certain threshold.

### 3.2.8 Particle Flow Algorithm of Event Reconstruction

A particle flow (PF) algorithm is used by CMS to identify and reconstruct stable particles [41]. It processes the information from all CMS subdetectors and identifies and reconstructs each stable particle in an event individually. The list of particles include muons, electrons, photons, charged and neutral hadrons. Each type of particles leaves its specific trace in the CMS detector as shown in Fig. 3.6. After reconstruction of individual stable particles, jets are built, missing transverse energy  $E_T^{miss}$  is determined, and certain short-lived particles are reconstructed based on the list of individual stable particles in the event.

One particle can induce several different particle-flow elements in different subdetectors. Examples of elements include a track in the tracking or muon systems, or a calorimeter cluster. The linking algorithm checks each pairs of elements in an event and produces blocks of elements if the distance between two elements is small and, therefore, they are considered to be linked. Usually, a block has between one and three elements. Links can be connections between the tracking system and PS, ECal or HCal, between PS and ECal, between ECal and HCal, and between a tracking system and a muon system.

In each block, muons are considered first. A link between charged tracks in the tracking and muon systems produces one “particle-flow muon”. The corresponding track in the tracking system is removed from the block and corresponding energy deposits are subtracted from ECal and HCal. Then electrons are reconstructed and identified using the tracking system and ECal. The corresponding

tracks and ECal clusters are removed from the block. Remaining tracks and clusters are considered more carefully to identify charged hadrons, neutral hadrons, and photons.

When all particles in the event are reconstructed and identified, the algorithm determines missing transverse energy  $E_T^{miss}$  as

$$E_T^{miss} = -|\sum \mathbf{P}_T|, \quad (3.4)$$

where the summation covers all visible particles in the event. For precise measurement of  $E_T^{miss}$  it is important to capture the full energy release of all visible particles.

$E_T^{miss}$  is used in physics measurements as a measure of  $P_T$  of neutrinos and other invisible particles in the event. Fake  $E_T^{miss}$  can originate from particles that did not fall into the detector acceptance, particles that did not reach the tracking system because their momenta was too low and, therefore, track curvature was too high, momenta mismeasurement, particle misidentification, cosmic rays particles, and machine background. Fake  $E_T^{miss}$  is a cause of background for the processes with real  $E_T^{miss}$ . For instance, in  $W\gamma$  measurement we have backgrounds from  $Z+jets$  and  $Z\gamma$  which do not have any real  $E_T^{miss}$ .

$E_T^{miss}$  is corrected through the propagation of corrections applied on kinematic parameters of jets. Additionally,  $E_T^{miss}$  is corrected for the PU effect.

In the measurement of this dissertation PF muons, electrons, photons, and  $E_T^{miss}$  are used for all the major steps of the cross section measurement including event selection, background subtraction, various corrections, and determination of phase space restrictions and bin boundaries. Each step is described in greater detail in Ch. 5.

# Chapter 4

## CMS Tracker Alignment

In the presence of a constant magnetic field, a charged particle has a helical trajectory which can be parametrized by five constants in three dimensions. While a charged particle travels through a tracking system, the tracking system detects hits. A reconstruction algorithm determines the track parameters by fitting the positions of hits assuming a helical trajectory. That allows the reconstruction of the full geometry of the track as well as the corresponding particle momentum, and to determine whether the particle came from the point of the  $pp$  collision or decay of a secondary particle.

High precision track reconstruction is necessary for accurate measurements of particle kinematics. Better location uncertainty leads to higher precision of a measurement of the track parameters. The location uncertainty depends on our knowledge of the positions and orientations in the space of the tracking system modules. For example, the hit resolution in the CMS pixel detector is  $\sim 10 \mu\text{m}$  in the  $r\text{-}\phi$  plane and  $\sim 30 \mu\text{m}$  in the  $r\text{-}z$  plane [37].

When the modules of the pixel detector are mounted, their positions are known with precision of  $\sim 200 \mu\text{m}$ . To take full advantage of the resolution of  $10 \mu\text{m}$ , we

need to know positions of modules at the accuracy of the single hit resolution. The procedure for the determination of the module locations and orientations is called the tracker alignment. The approach used for the tracking alignment in CMS is described in Ch. 4.1.

The procedure of tracker alignment is essential for the momentum measurement of all charged particles including electrons and muons that are the final state particles of the measurement of this dissertation as well as for the determination of the position of the primary vertex. The measurement of this dissertation is based on data collected in 2012 while the author of this dissertation participated in the alignment of the tracking system in 2015 (Ch. 4.2). The results of 2015 alignment are not used for the measurement of this dissertation but are used for all CMS physics measurements of 2015 data including  $W\gamma$  measurement at  $\sqrt{s} = 13$  TeV.

## 4.1 Approach

It is necessary to align a part of the tracking system whenever we suspect a physical change in a location or an orientation of this part. First of all, whenever a part of the CMS tracker is taken out and placed back, we need to realign it. Also whenever a magnet is turned on and off, different parts of the tracking system shift with respect to one another. Pixel half barrels are not screwed firmly, and are moving along each other on rails, therefore, they need to be aligned frequently.

The concept of track-based alignment can be illustrated in the example of the alignment of a toy tracker (Fig. 4.1-4.2). A charged particle crosses a toy tracker of six flat equidistant modules. Because real geometry of the tracker differs from the ideal one, hits are recorded at the places different from the design ideal places. We record and process a large number of tracks to determine positions and orientations of the modules.

The tracker alignment problem is a least squares problem. The expression to minimize is the following:

$$\chi^2(\mathbf{p}, \mathbf{q}) = \sum_j^{tracks} \sum_i^{hits} \left( \frac{m_{ij} - f_{ij}(\mathbf{p}, \mathbf{q}_j)}{\sigma_{ij}} \right)^2, \quad (4.1)$$

where  $\mathbf{p}$  are parameters describing the tracker geometry,  $\mathbf{q}_j$  are parameters of the  $j^{th}$  track,  $m_{ij} - f_{ij}$  are distances between the measured hit and a position predicted by the track fit (“residuals”),  $\sigma_{ij}$  is the Gaussian error of the measurement.

We can align the large substructures (like pixel half barrels, pixel endcap disks and other) with respect to the global CMS coordinate system and individual modules with respect to the coordinate systems of their substructures. The parameters to align large substructures include three coordinates to determine location and three angles to determine orientation of the substructure. At the module level,

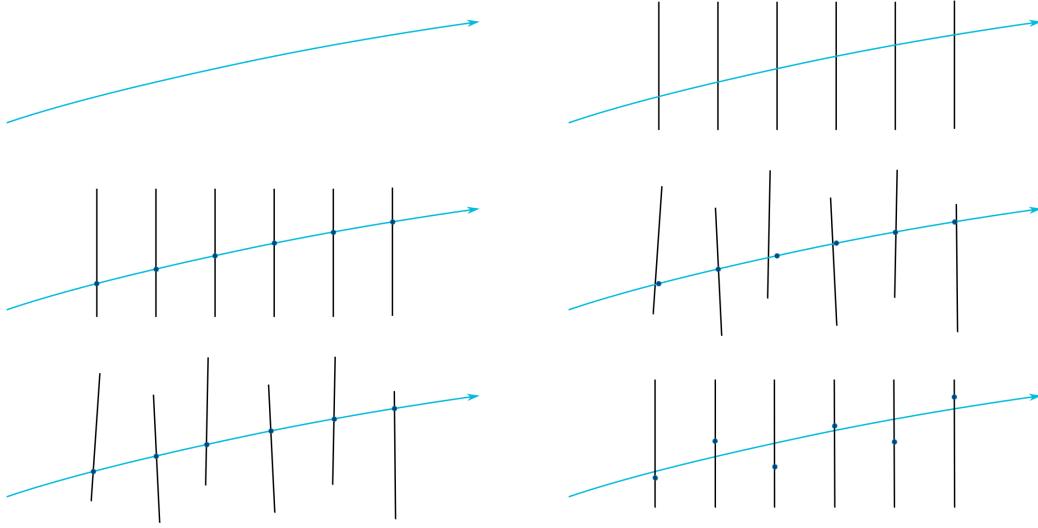


Figure 4.1: The alignment of a toy tracker, part 1. When a charged particle passes through a detector (top left), it crosses a toy tracker which consists of six flat equidistant modules (top right). If the modules were placed exactly at their design positions, we would observe the hits exactly at the points where the track crosses modules of ideal geometry (middle left). However, in reality, the positions and tilts of the modules are different from ones suggested by the ideal geometry (middle right). Hits, indeed, are recorded at the places where modules are mounted, not at the design ideal places (bottom left). If we assumed a tracker to be ideal and a track to be smooth, we would see that our hits are off-track (bottom right). Image by Frank Meier.

we align positions and rotations with respect to the positions and angles of the corresponding large structure (Fig. 4.3). Also at the module level, we align for surface deformations which are described by three parameters per sensor (Fig. 4.4).

The alignment process appears to require the inversion of giant matrices, of millions by millions. We have two alignment algorithms that use their ways to simplify the computation: Hits and Impact Parameter (HIP) [43] and Millepede-II [42]. HIP performs a minimization for one module at a time processing tracks that pass through this particular module. After the minimization is done for all modules, HIP performs the second iteration. The iterations are stopped after module positions

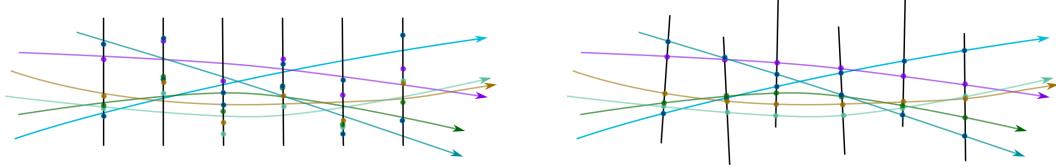


Figure 4.2: The alignment of a toy tracker, part 2. We record a large number of tracks and take into account them all to determine the alignment parameters by minimizing residuals between measured and predicted hits. Image by Frank Meier.

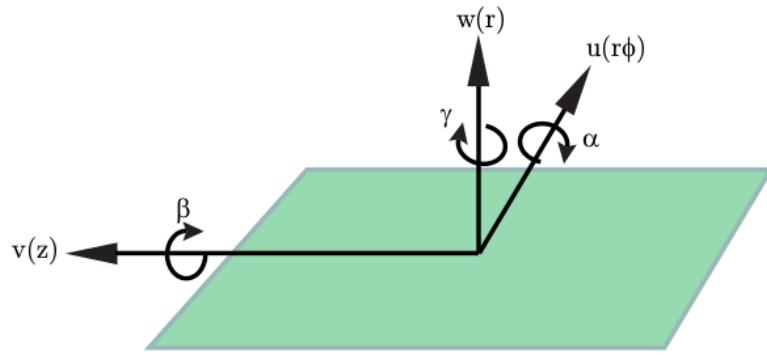


Figure 4.3: Local alignment parameters [10].

converge. Millepede-II performs a simultaneous fit of all alignment parameters at a time, and uses linear algebra tricks for sparse matrices to avoid dealing with most track parameters.

After the procedure of the tracking system alignment is performed, we validate the results. Chapter 4.2 discusses various tools of alignment validation using the example of the tracking system alignment based on the 2015 data.

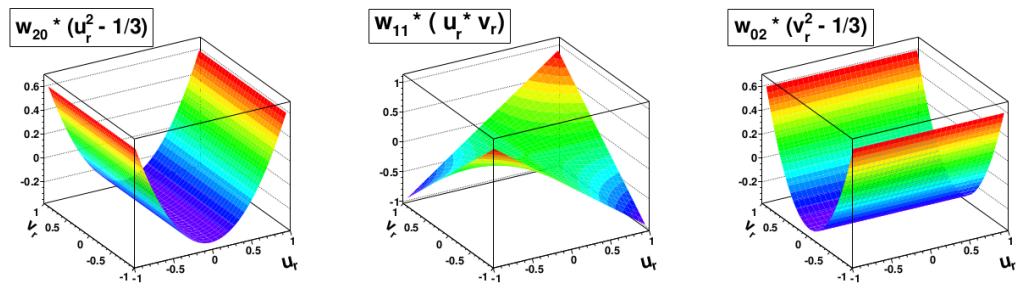


Figure 4.4: Surface deformations [11].

## 4.2 Selected Results on Alignment of the Tracking System with 2015 Data

After the long shutdown in 2013-2014, LHC restarted collisions in 2015. Data collection periods in 2015 include the following:

- cosmic ray data with CMS magnetic field of  $B = 0\text{T}$  (prior collisions);
- cosmic ray data at  $B = 3.8\text{T}$  (prior collisions);
- $p p$  collision data at  $B = 0\text{T}$  (interfilled with cosmic ray data);
- $p p$  collision data at  $B = 3.8\text{T}$  (interfilled with cosmic ray data) .

Only  $p p$  collision data at  $B = 3.8\text{T}$  are used for physics measurements and the three other periods are preliminary. During the preliminary periods we make sure that all parts of the detector work properly and also perform the preliminary alignment of the tracking system. Collision data are interfilled with cosmic ray data when LHC does not provide any collisions. This interfill cosmic ray data are also used for the tracker alignment.

Different data collection periods correspond to different detector geometries particularly due to changes of the magnetic field. Thus, alignment constants were derived separately for each of the data collection periods using the alignment results of the previous period of data collection as a starting point.

The modules in certain parts of BPIX were repaired during the shutdown, and all pixel subdetectors were moved within the tracker. That caused one of the largest differences between Run I and Run II geometries. The first alignment of the tracker corrected for these displacements using cosmic ray data with magnetic field turned on ( $B = 3.8\text{T}$ ) and off ( $B = 0\text{T}$ ).

After the cosmic ray data collection periods, the magnetic field was turned off again due to problems with the cryogenic system, and the first collisions were detected with  $B = 0\text{T}$ . This change in the magnetic field expectedly caused movements in the tracking system. The effect is the strongest in the pixel subdetectors. The alignment performed using  $B = 0\text{T}$  collisions and cosmic ray data recovers the tracker performance in reconstructing kinematic parameters of charged particles. When the magnetic field was turned back on, the large substructures of BPIX and FPIX have been displaced again, and, thus, the tracking system was aligned again to recover these displacements.

To validate results of tracking system alignment, the following tools are used:

- geometry comparison tool (Ch. 4.2.1);
- validation using distribution of median residuals (Ch. 4.2.2);
- cosmic track splitting validation (Ch. 4.2.3);
- primary vertex validation (Ch. 4.2.4).

The full results of the first alignment with Run II data are available at [44].

### 4.2.1 Geometry Comparison

The geometry comparison visualizes differences in positions of modules between two different geometries of the CMS tracking system. Figure 4.5 shows the comparison between positions of the FPIX modules between Run I and Run II geometries. Each dot in the figure corresponds to one module. Four clusters of red dots (Fig. 4.5, left) and shifted parts at  $(\phi < -\pi/2, \phi > \pi/2)$  and  $(-\pi/2 < \phi < \pi/2)$  (Fig. 4.5, right) represent displacements of four half-disks by 4.5 and 5.5 mm at the  $-z$  side of the FPIX. At the  $+z$  side of the FPIX, small relative movements

of individual modules are observed only. For more intuitive visualization, the three-dimensional plot of the pixel detector is produced (Fig. 4.6).

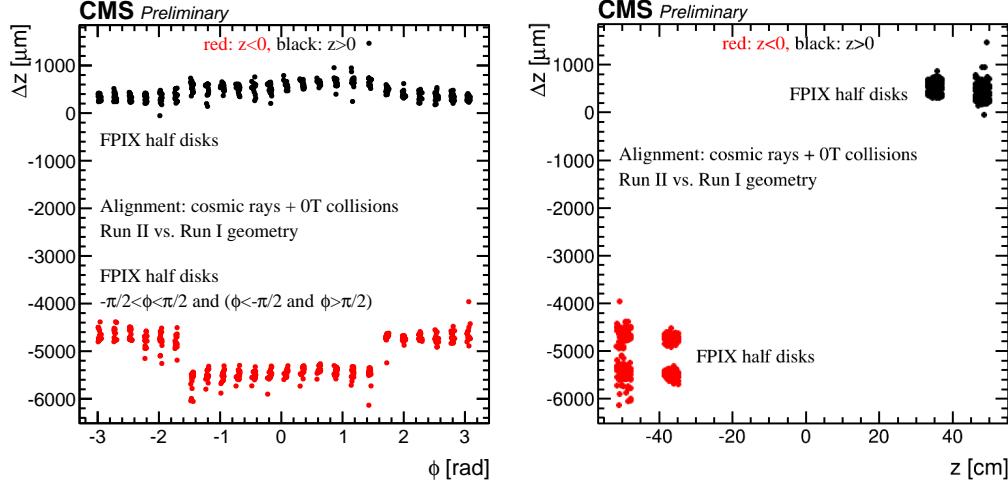


Figure 4.5: Comparison of Run II and Run I positions of the modules in the FPIX of the CMS tracking system. Positions are determined with the Millepede-II and HIP algorithms using cosmic ray data collected with the magnetic field of  $B = 0\text{T}$  and  $B = 3.8\text{T}$  magnetic field in the CMS solenoid. The difference  $\Delta z$  (Run II - Run I) is plotted as a function of  $z$  (left) and  $\phi$  (right) in global coordinates. The plot shows the displacements of two pixel half disks by 4.5 and 5.5 mm.

## CMS Preliminary

Alignment: cosmic rays + 0T collisions

Run II vs. Run I geometry, shift x 5

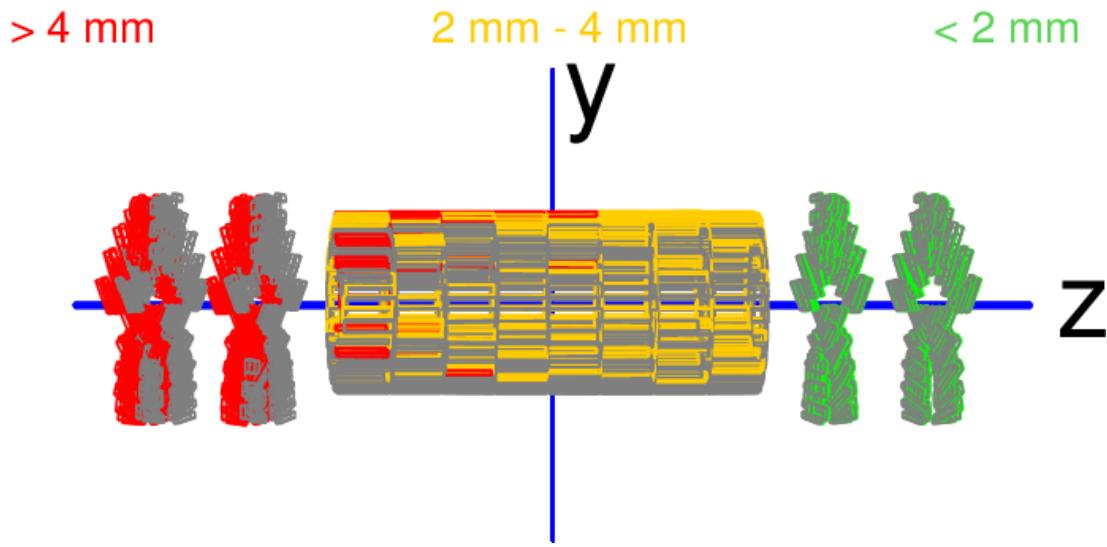


Figure 4.6: Three-dimensional geometry comparison of Run II and Run I positions in the BPIX and FPIX of the CMS tracking system. Positions are determined with the Millepede-II and HIP algorithms using cosmic ray data collected with the magnetic field of  $B = 0\text{T}$  and  $B = 3.8\text{T}$  magnetic field in the CMS solenoid and collision data with  $B = 0\text{T}$  at  $\sqrt{s} = 13\text{ TeV}$ . The positions at the end of Run I are shown in gray. The module displacements between Run I and Run II are magnified by a factor of 5 for visualization purpose. The resulting positions are shown in red, yellow, or green, depending on the displacement magnitude.

### 4.2.2 Distributions of Medians of Unbiased Track-Hit Residuals

Besides the geometry comparison, we also have the distributions of medians of unbiased track-hit residuals (DMR) validation tool. Each track from a given dataset is refitted using prepared alignment constants, and the hit for each module is predicted from all other hits of the track. After that, DMRs of all modules in a given subdetector are plotted on the same histogram. The width of the prepared DMR is a measure of the statistical precision of the derived alignment results.

The DMRs are plotted for the local  $x$ - (Fig. 4.7, left) and  $y$ -directions (Fig. 4.7, right) in the BPIX. The blue line shows the DMR for Run I while the green line shows the geometry aligned with Run II data. The RMS values show that performance of the aligned geometry is improved by a factor of 10 over the Run I geometry. Because of physical changes in the tracking system, including taking off and putting back on the pixel detector, replacing certain modules, and changes in the magnetic field, the Run I geometry is not expected to work well with Run II data.

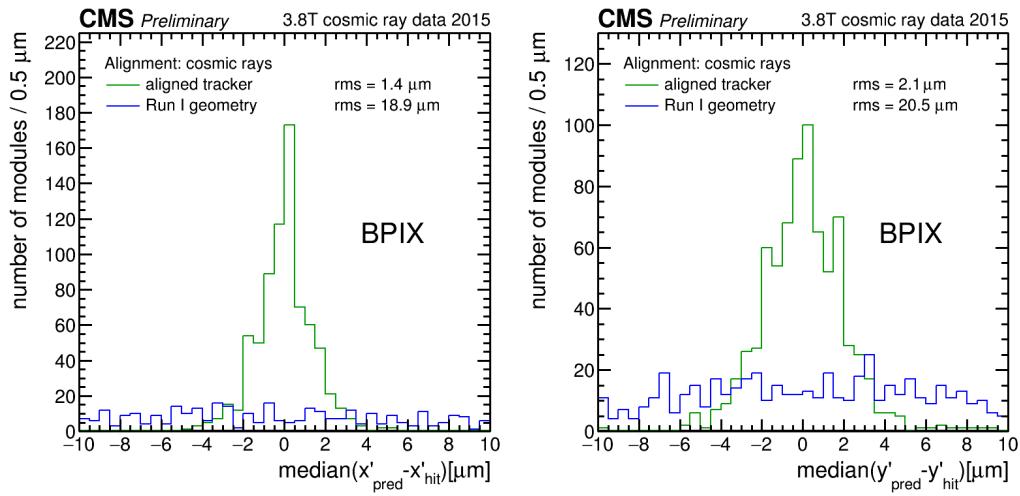


Figure 4.7: DMRs for the local  $x$ -direction (left) and for the local  $y$ -direction (right) in the BPIX of the CMS tracking system, using 2 million cosmic ray tracks collected with the magnetic field of  $B = 3.8\text{T}$ . The blue line shows the Run I geometry. The green line shows the alignment produced with the Millepede-II and HIP algorithms using cosmic ray data at  $B = 0\text{T}$  and  $B = 3.8\text{T}$ . The aligned geometry shows reasonable performance.

### 4.2.3 Cosmic Track Splitting Validation

To perform the cosmic track splitting validation, cosmic tracks are split into two parts at the hit closest to the center of the detector and both parts are reconstructed separately using alignment results. After that, the distributions of the differences in track parameters are prepared. The RMS values of the distributions are measures of the precision of the alignment constants. A deviation of a central value from zero would indicate a bias. The results of this validation for 2015 alignment are shown in Fig. 4.8.

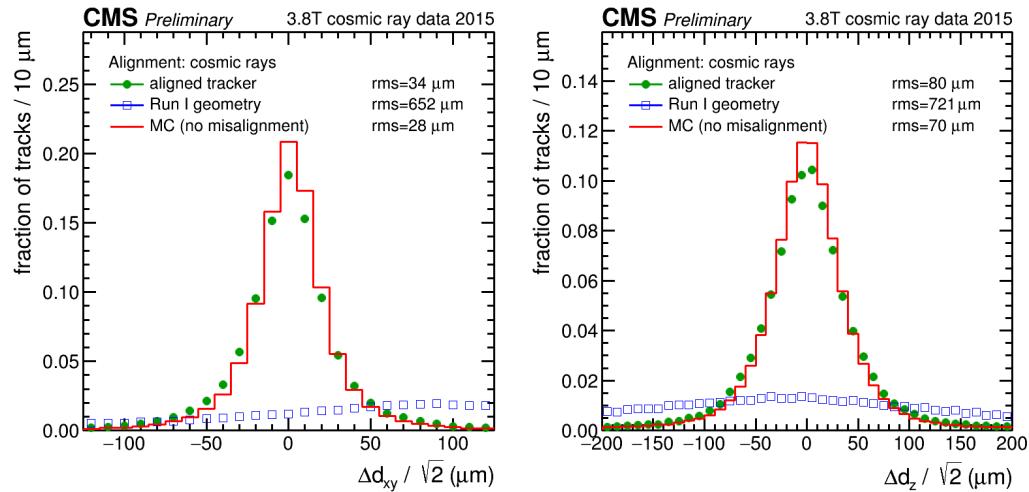


Figure 4.8: Results of the cosmic track splitting validation. The normalized differences between two parts of a cosmic track in the  $xy$  distance between the track and the origin ( $d_{xy}$ , left), and in the distance in the  $z$  direction between the track and the origin ( $d_z$ , right). Alignment is produced with the Millepede-II and HIP algorithms using cosmic ray data at the magnetic field of  $B = 0\text{T}$  and  $B = 3.8\text{T}$  of CMS solenoid. Geometry aligned with Run II data is shown in green, Run I geometry is shown in blue. Aligned geometry shows reasonable performance.

#### 4.2.4 Primary Vertex Validation

The resolution of the reconstructed vertex position is driven by the pixel subdetectors as the closest subdetectors to the interaction point which also have the best hit resolution. The primary vertex validation is based on the study the distances between tracks and the reconstructed vertex.

Figure 4.9 compares the previous alignment to the one of a previous alignment reached during the commissioning phase with cosmic ray tracks at full magnetic field and to a detailed detector simulation with perfect alignment and calibration. The structures of the green curve indicate relative movements of the pixel half-barrels.

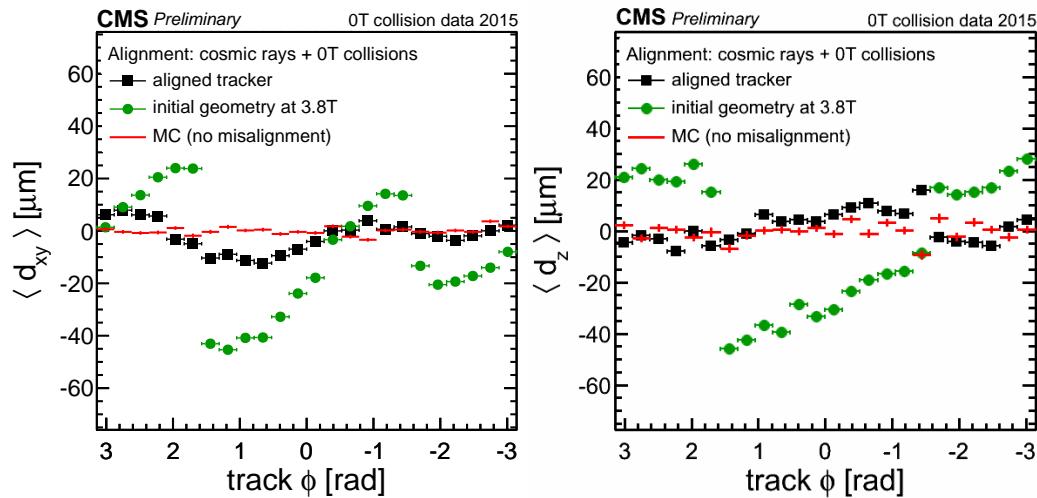


Figure 4.9: The results of the primary vertex validation. The distance of the track at its closest approach to a refit unbiased primary vertex ( $d_{xy}$ , left and  $d_z$ , right) in the transverse plane. The validation is produced with  $B = 0\text{T}$  collision data. The alignment is produced with the Millepede-II and HIP algorithms using  $B = 0\text{T}$  and  $B = 3.8\text{T}$  cosmic ray data and  $B = 0\text{T}$  collision data.

Given the complexity of the CMS detector, any single measurement based on CMS data requires an excellent understanding of the geometry and response of

all systems to particles of all types. The CMS Alignment and Calibration team coordinates hundreds of CMS physicists who are working on various aspects of this. The Chapter 4 of this dissertation presented one aspect of this work that concerns alignment of one system of CMS, the part in which the author of this dissertation played an important role that included running alignment and validation for various studies of the procedure performance as well as for actual alignment of the tracking system.

# Chapter 5

## $W\gamma$ Cross Section Measurement

The goal of this research is to measure the total and differential cross section of the  $W\gamma$  production in  $pp$  collisions as a function of the photon transverse momentum  $P_T^\gamma$  at  $\sqrt{s} = 8$  TeV center-of-mass collision energy. Decay channels  $W \rightarrow \mu\nu$  and  $W \rightarrow e\nu$  are considered. The measurement is performed using CMS data collected in 2012.

The phase space for the cross section was chosen from the theoretical considerations of a divergence at  $P_T^\gamma = 0$  and  $\Delta R(\gamma, l) = 0$ , from the consideration of the trigger requirements (Ch. 5.2) and detector acceptance. More specifically, the phase space requirements on the final state photon and lepton were chosen to match those of CMS  $Z\gamma$  measurement at 8 TeV for the comparison purpose. The full list of the phase space requirements include:

- $P_T^\gamma > 15$  GeV;
- $\Delta R(\gamma, l) > 0.7$ ;
- $|\eta^\gamma| < 2.5, |\eta^l| < 2.5$ ;
- $P_T^l > 20$  GeV;

- $I^\gamma < 5 \text{ GeV}$ , where  $I_\gamma$  is a sum of  $P_T$  of all MC generated particles particles  $p$  in the event within  $\Delta R(p, \gamma) < 0.3$ .

$P_T^\gamma$  binning for the differential cross section measurement was also chosen to match CMS  $Z\gamma$  measurement.  $P_T^\gamma$  bin boundaries are 15-20-25-30-35-45-55-65-75-85-95-120-500 GeV.

## 5.1 Measurement Strategy

The measurement strategy can be represented as a chain of steps that is summarized in Tab. 5.1. First, we select events and obtain number of selected events. In Tab. 5.1 the total number of selected events is  $N_{sel}$  and numbers of selected events in  $P_T^\gamma$  bins are  $N_{sel}^j$ , where  $j$  is a bin number. In the selected sample, we have events of our signal as well as background events. Our next step is to subtract the background. This step gives us background subtracted yields of signal events  $N_{sign}$  and  $N_{sign}^j$ . After that, we perform the detector resolution unfolding (Ch. 5.5) and obtain yields within acceptance and selection restrictions:  $N_{A \times \epsilon}^i = U_{ij} \cdot N_{sign}^j$ , where  $U_{ij}$  is the unfolding operator. Then we apply corrections for kinematic and geometrical acceptance and reconstruction and selection efficiency (Ch. 5.6). Finally, we divide measured number of events by integrated luminosity recorded by CMS and, in case of the differential cross section, by the width of  $P_T^\gamma$  bin. This gives us total and differential cross section. Each step has its systematic uncertainties associated with it and we estimate their contributions to the final results. The remainder of this Chapter describe each measurement step in detail.

Table 5.1: Measurement steps. The first column is the name of the step, the second and the third columns are algebraic representations of the steps for the differential and total cross section measurements, respectively.

Step	$d\sigma/dP_T$	$\sigma$
select events	$N_{sel}^j$	$N_{sel}$
subtract background	$N_{sign}^j = N_{sel}^j - N_{bkg}^j$	$N_{sign} = N_{sel} - N_{bkg}$
unfold	$N_{A \times \epsilon}^i = U_{ij} \cdot N_{sign}^j$	—
correct for the acceptance and efficiency	$N_{true}^i = \frac{N_{A \times \epsilon}^i}{(A \times \epsilon)^i}$	$N_{true} = \frac{N_{sign}}{A \times \epsilon}$
divide by luminosity and bin width	$\left(\frac{d\sigma}{dP_T^\gamma}\right)^i = \frac{N_{true}^i}{L \cdot (\Delta P_T^\gamma)^i}$	$\sigma = N_{true}/L$
estimate systematic uncertainties		

At first, we perform the measurement in a blinded way. It is done to avoid unintended biasing of our results in any direction. Our blinded strategy is the following:

- for  $p_T^\gamma < 45 \text{ GeV}$ : use 100% of data; and
- for  $p_T^\gamma > 45 \text{ GeV}$ : use 5% of data (every 20<sup>th</sup> event).

The threshold of  $p_T^\gamma = 45 \text{ GeV}$  is chosen because below that we do not expect any new physics, and the percentage of 5% is chosen because these amount of data gives us such a large statistical uncertainty so that we would not notice any new physics if it was there. After the measurement procedure is fully established, we perform the measurement using our full dataset (“unblinded” measurement). All plots shown in this dissertation are prepared in the unblinded manner.

## 5.2 Data and Monte Carlo Samples

The data sample we use in this measurement was recorded by the CMS experiment in 2012 in LHC  $pp$  collisions at  $\sqrt{s} = 8$  TeV. Integrated luminosity of the dataset is  $L = 19.6 \text{ fb}^{-1}$ . To select  $W\gamma$  events, we use data collected by single muon and single electron triggers. The single muon trigger requires that in each event there is at least one reconstructed isolated muon with  $P_T^\mu > 24 \text{ GeV}$  and  $|\eta| < 2.1$ . The single electron trigger requires at least one reconstructed electron with  $P_T^e > 27 \text{ GeV}$  which also passes a certain set of identification requirements, including isolation. Such trigger choice maximize our chances to select  $W\gamma$  events out of dominant multijets events.

In addition to  $W\gamma$ -selected data sample, we also prepare  $Z\gamma$ -selected data sample which is used for the background estimation (Ch. 5.4) and for cross checking purpose (App. L). To select  $Z\gamma$  events, we use double muon and double electron triggers. The double muon trigger requires a presence of at least two reconstructed muons with  $P_T^\mu > 17 \text{ GeV}$  and  $P_T^\mu > 8 \text{ GeV}$  per event. The double electron trigger requires a presence of at least two reconstructed electrons with  $P_T^e > 17 \text{ GeV}$  and  $P_T^e > 8 \text{ GeV}$  which also satisfy several other criteria of electron's quality.

All simulated samples (often referred as Monte Carlo or MC samples) used in this measurement are generated with MadGraph [45] and reconstructed centrally by the CMS simulation team. Information regarding MC samples used for our measurement is given in Tab. 5.2 alongside with the corresponding cross sections at 8 TeV. All cross sections are calculated with kinematic restrictions matching to the kinematic restrictions of the samples.

When we select  $W\gamma$  events, certain number of events from other processes pass

the selection criteria too. Tab. 5.2 contains all sources that significantly contribute to the selected sample.  $W\gamma \rightarrow l\nu\gamma$  contains  $W\gamma \rightarrow \mu\nu\gamma$  and  $W\gamma \rightarrow e\nu\gamma$  which are our signal samples and  $W\gamma \rightarrow \tau\nu\gamma$  which is a background for both channels. The other samples listed in Tab. 5.2 are background samples. They are used for the background estimation and cross checking as explained in detail in the remainder of the chapter.

Table 5.2: Summary of simulated samples used in the measurement.

Process	Type	$\sigma, \text{pb}$
$W\gamma \rightarrow l\nu\gamma$	signal	554
$W+\text{jets} \rightarrow l\nu+\text{jets}$	background	36257
$DY+\text{jets} \rightarrow ll+\text{jets}$	background	3504
$t\bar{t}+\text{jets} \rightarrow 1l+X$	background	99
$t\bar{t}+\text{jets} \rightarrow 2l+X$	background	24
$Z\gamma \rightarrow ll\gamma$	background	172

The  $W+\text{jets}$  process is simulated with the matrix element (ME) generator MADGRAPH 5.1.1 interfaced with PYTHIA 6.426 using the  $Z2^*$  tune for parton showering and hadronization. This sample of events is produced with the CTEQ6L1 parton distribution function (PDF) set and is normalized to the inclusive NNLO cross section calculated with FEWZ 3.1. The MADGRAPH 5 + PYTHIA 6 calculation includes the production of up to four partons at LO. The jets from matrix elements are matched to parton showers following the  $k_T$ -jet MLM prescription, where partons are clustered using the  $k$  algorithm with a distance parameter of 1. The merging of parton showers and matrix elements with the MLM scheme uses a matching scale of 20 GeV. The factorization and renormalization scales for the  $2 \rightarrow 2$  hard process in the event are chosen to be the transverse mass of the  $W$  boson produced in the central process. The  $k_T$  computed for each QCD emission vertex is used as renormalization scale for the calculation of the strong coupling

constant  $\alpha_S$  of that vertex.

Another significant background process for  $W\gamma$  is DY+jets where DY is a notation for the Drell Yan process which is  $pp \rightarrow / \gamma \rightarrow ll$ . The simulated sample of DY+jets events is generated with MADGRAPH version 5.1.1. This simulation uses the CTEQ6L1 PDFs. The DY+jets sample is normalized to the NNLO inclusive cross section calculated with FEWZ 3.1. The requirement on the invariant mass of the final state lepton pair is  $M_{ll} > 50$  GeV. Samples  $W$ +jets and DY+jets are prepared in a way they do not overlap with  $W\gamma$  and  $Z\gamma$  samples.

All MC samples are normalized to the luminosity of the dataset  $L = 19.6 \text{ pb}^{-1}$ . To perform the normalization, appropriate weights are applied to each event in each MC sample. Such weighted MC samples are often used for various MC estimates and for plotting histograms for data vs MC comparisons.

The NLO cross section of  $W\gamma$  was calculated with the MCFM in the same phase space for which the  $W\gamma$  sample was generated. The NNLO contribution is estimated to be 19%-26% of the NLO value [22]. We use an NLO cross section value, and the NNLO estimate is used as an systematic uncertainty of the normalization of the  $W\gamma$  sample when it is applicable.

The uncertainty on normalization of the  $Z\gamma$  sample gives a significant contribution to the uncertainty of a the measurement because  $Z\gamma$  MC sample is used to estimate the most significant background (Ch. 5.4.1). MCFM provides a value of the cross section with uncertainty of 20%. To minimize the uncertainty, we use a cross section of  $Z\gamma$  measured at 8 TeV by CMS [46] and recalculated it for the phase space of the generated  $Z\gamma$  MC sample.

The  $Z\gamma$  cross section of  $\sigma = 2073 \pm 95 \pm 11 \pm 53 \text{ fb}$  has been quoted in the phase space described in [46]. To determine the measured cross section in the generator

phase space, the following formula was used:

$$\sigma_{ps1} = \sigma_{ps2}^{meas.} \cdot \frac{N_{ps1}^{MC}}{N_{ps2}^{MC}}, \quad (5.1)$$

where  $\sigma_{ps2}^{meas.}$  is the 8 TeV cross section measured by CMS,  $N_{ps1}^{MC}$  and  $N_{ps2}^{MC}$  are numbers of events in the full phase space of  $Z\gamma$  MC samples and in the phase space corresponding to the measured cross section  $\sigma_{ps2}^{meas.}$ ,  $\sigma_{ps1}$  is the resulting cross section of the  $Z\gamma$  sample in its full phase space.

The resulting  $Z\gamma$  cross section is found to be 172 pb. Uncertainties on normalizations of other samples do not contribute significantly to the uncertainty of the measurement, therefore, we use MCFM values for them.

At the instantaneous luminosities of LHC in 2012, as a rule, multiple  $pp$  interactions occurred per bunch crossing. Multiple interactions are also simulated in the MC samples. However, MC samples are usually produced before data collection is finished, and in the end have to be reweighted so that the distribution of the number of interactions (pileup or PU) in a simulated sample matches the data. The PU weights are assigned on each event in each MC sample to make the PU distribution in MC accurately describe PU in data.

To validate the procedure of the PU reweighting, we check the agreement between data and MC in the distribution of the number of vertices in  $Z\gamma \rightarrow \mu\mu\gamma$ -selected datasets (Fig. 5.1). We choose the  $Z\gamma$  selected dataset instead of the  $W\gamma$ -selected dataset because the sample composition for  $Z\gamma$  selection is understood better and normalizations of the MC samples that pass  $Z\gamma$  selection are known better. The  $Z\gamma$  selection is explained in Ch. 5.3 alongside with the  $W\gamma$  selection.

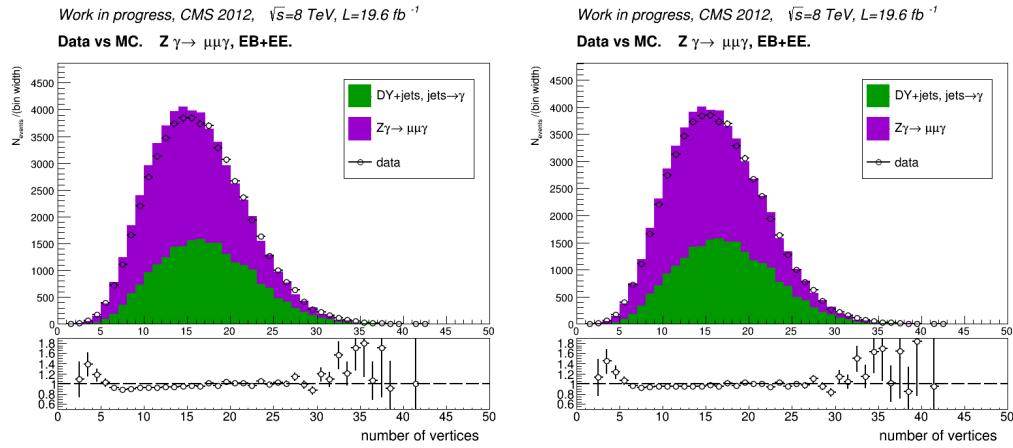


Figure 5.1: Number of vertices of  $Z\gamma$  candidates in the muon channel. Data vs MC. Left: no PU reweighting applied, right: PU reweighting applied. Ratio plot in the bottom shows data yields divided over total MC yields. EB+EE means that events with a final state photon reconstructed in the ECal barrel as well as events with a final state photon reconstructed in the ECal endcap are shown on the plots.

## 5.3 Event and Object Selection

### 5.3.1 Object Selection

We select events with a muon and a photon in the final state for the muon channel and events with an electron and a photon in the final state for the electron channel.

We apply selection requirements on transverse momenta of  $P_T^\mu > 25$  GeV on muons,  $P_T^e > 30$  GeV on electrons and  $P_T^\gamma > 15$  GeV on photons. In addition, electrons and photons must be within barrel (EB) or endcap (EE) sections of the ECal which corresponds to pseudorapidity ranges of  $|\eta^{e,\gamma}| < 1.4442$  and  $1.566 < |\eta^{e,\gamma}| < 2.5$ , respectively. Muons must be within  $|\eta^\mu| < 2.1$ . Selection requirements on  $P_T^\mu$ ,  $\eta^\mu$ , and  $P_T^e$  are determined by the trigger requirements,  $\eta^{e,\gamma}$  criteria are determined by the geometrical limitations of the detector acceptance, and  $P_T^\gamma > 15$  GeV is the phase space requirement.

CMS Particle Object Group (POG) provides their recommendations for object identification (ID) criteria for any given period of data collection. To satisfy a muon ID criteria, objects, first of all, must be reconstructed as muons by the PF algorithm. Quality requirements are applied on tracks reconstructed in both the tracking system and the muon system. These two tracks must match. An isolation from the other nearby PF objects is also required.

The electron ID and photon ID criteria include requirements on the shower shape, on ratio of energies released in ECal and HCal. The electron ID also includes requirements on the track quality. Similarly to muons, electrons and photons must be isolated from the nearby PF objects. For the photon ID, the isolation consists on three parts: charged hadron, neutral hadron, and photon isolation. To reject electrons reconstructed as photons, “conversion safe electron veto” (CSEV) is recommended as a part of any photon ID. CSEV rejects photons

with associated track originated from a converted photon. Such tracks are not identified, thus, such particle is reconstructed as a photon rather than an electron. But some of such tracks can originate from real electrons, and by requiring CSEV, we reject this background.

In  $W\gamma$  measurement, we applied object ID criteria as recommended by POG with one exception: in the electron channel, the CSEV is substituted with the “pixel seed veto” (PSV) as recommended by CMS  $W\gamma\gamma \rightarrow l\nu\gamma\gamma$  measurement team [47]. PSV rejects photons that have any track seed in the pixel detector that can match to the measured ECal supercluster. PSV is tighter requirement than CSEV and, therefore, is used in the electron channel only where we have much larger background from electrons misidentified as photons.

To reduce backgrounds from the process with two or more leptons, such as  $Z\gamma \rightarrow ll\gamma$  process, in the muon channel, we reject all events that have the second reconstructed muon candidate with  $P_T^\mu > 10$  GeV and  $|\eta|^\mu < 2.4$ , and, in the electron channel, we reject events that have the second reconstructed electron candidate with  $p_T^e > 10$  GeV and satisfying the “Veto” ID criteria.

Selection criteria are applied on the data sample as well as on all MC samples. The selection efficiency may differ between data and MC. The ratios between data and MC efficiencies are called the scale factors (SF). The SF for the selection criteria recommended are provided by CMS POG. For the PSV criterion in the photon selection in the electron channel, additional SF are applied as derived by the  $W\gamma\gamma$  team [47]. All SF are listed in App. A.

### 5.3.2 Event Level Selection

In the final state of the  $W\gamma \rightarrow l\nu\gamma$  process, there is a lepton, a photon, and a neutrino. To select events with such signature, we require events with exactly one lepton (muon or electron), a photon, both originating from the primary vertex, and with the significant missing transverse energy  $E_T^{miss}$ . The selection criteria for the individual electrons, muons and photons are described in Ch. 5.3.1.

The standard tool to detect a particle that decays is to reconstruct its invariant mass out of its decay products. Decay products of a  $W$  boson are a charged lepton and a neutrino. CMS does not detect neutrino, it only measures the missing momentum in the plane, transverse to the beamline, which can be partially associated with a neutrino. The transverse momentum is described by two parameters:  $E_T^{miss}$  and  $\phi^{miss}$ . Because we do not have an estimate of the longitudinal component of a neutrino, we cannot construct an invariant mass of a  $W$  boson. Instead, we construct its transverse mass:

$$M_T^W = \sqrt{(2P_T^l E_T^{miss} (1 - \cos(\phi^l - \phi^{miss}))),} \quad (5.2)$$

where  $P_T^l$  is a lepton transverse momentum,  $\phi^l$  is an azimuthal angle of the lepton momentum, and  $\phi^{miss}$  is an azimuthal angle of the missing transverse momentum. To enhance contribution from  $W\gamma$  compared to background processes without a final state neutrino, we require  $M_T^W > 40$  GeV. Value of 40 GeV was recommended by the CMS Standard Model Physics (SMP) group because the same requirement was used in  $W\gamma\gamma$  measurement. The  $M_T^W$  distribution is shown in Fig. 5.2. Photons with  $P_T^\gamma < 45$  GeV are selected for this plot because for such photons we do not expect them to be result of a new physics process.

After the listed selection criteria are applied, significant background from

DY+jets in the electron channel remains. This background is caused by one of the electrons misidentified as a photon. Its contribution is the most significant around the invariant mass of the electron-photon system  $M_{e\gamma}$  close to the mass of the Z boson (Fig. 5.3) because the distribution of  $M_{ee}$  in the  $Z \rightarrow ee$  decay is peaking at the value of the Z boson mass. To reduce this background, we apply Z-mass window selection criterion, more specifically, events with  $70 \text{ GeV} < M_{e\gamma} < 110 \text{ GeV}$  are rejected.

In addition to  $W\gamma$ -selected datasets, we also prepare  $Z\gamma$ -selected datasets in muon and electron channels. Selection requirements include at least two muons (or electrons) and at least one photon in the final state. Kinematics and identification requirements on the objects are the same as for the  $W\gamma$  selection with the only exception: unlike  $W\gamma$ , in the  $Z\gamma$  selection in the electron channel, photons are required to pass CSEV rather than PSV.

Invariant mass of the final state lepton pair is required to be  $M_{ll} > 50 \text{ GeV}$ . Finally, a separation between photon and each lepton must be  $\Delta R > 0.7$ .

Finally, the separation  $\Delta R = \sqrt{(\Delta\phi^2 + \Delta\eta^2)}$  between the final state lepton and photon is required to be  $\Delta R(l, \gamma) > 0.7$  to enhance the TGC contribution. In case if there is more than one photon in the selected event, the candidate with the photon of the highest  $P_T^\gamma$  is selected.

In addition to  $W\gamma$ -selected datasets, we also prepare  $Z\gamma$ -selected datasets in muon and electron channels. Selection requirements include at least two muons (or electrons) and at least one photon in the final state. Kinematics and identification requirements on the objects are the same as for the  $W\gamma$  selection. Unlike in  $W\gamma$ , in  $Z\gamma$  selection in the electron channel, photons are required to pass CSEV, not PSV.

Invariant mass of the final state lepton pair is required to be  $M_{ll} > 50 \text{ GeV}$ . Finally, a separation between photon and each lepton is required to be the same as

in the  $W\gamma$  selection:  $\Delta R > 0.7$ .

### 5.3.3 Selected Events

After the  $W\gamma$  selection procedure applied, 175889 and 85643 events survived in the muon and electron channels, respectively. These events are used for the total and differential cross section measurements with respect to  $P_T^\gamma$ . Distributions of  $P_T^\gamma$  of the selected events are shown in Fig. 5.4 and documented in Tab. 5.4-5.5. The plots and tables include information about the underflow  $P_T^\gamma$  bin (10-15 GeV). The measurement in this bin is used for the detector resolution unfolding (Ch. 5.5).

Selected samples are dominated by  $W+jets$  events because of jets misidentified as photons. The main mechanism of a jet to be misidentified as a photon is a  $\pi^0 \rightarrow \gamma\gamma$  decay. Photon ID requirements reject most of such fragmentation photons, however, a  $W+jets$  event has much larger probability to be produced in a  $pp$  collision than a  $W\gamma$  event, and, therefore, even a small fraction of fragmentation photons from  $W+jets$  events becomes a significant background to  $W\gamma$ .

DY+jets background in the electron channel consists of two parts: jets misidentified as photons and electron misidentified as photons. The DY+jets MC samples in the the electron channel is split into two parts based on the gen-level information, and these two parts are denotes as  $jets \rightarrow \gamma$  and  $e \rightarrow \gamma$ .

There are large discrepancies between data and MC predictions in all the distributions as shown in Fig. 5.2-5.4. Possible reasons for the discrepancies include but are not limited to uncerntainties in the normalizations of all MC samples involved and difficulties in modeling jet fragmentation. Therefore, the data-driven background estimates are necessary (Ch. 5.4).

MC samples in all the plots are reweighted to the luminosity of data. PU

weights and efficiency scale factors are applied as well.

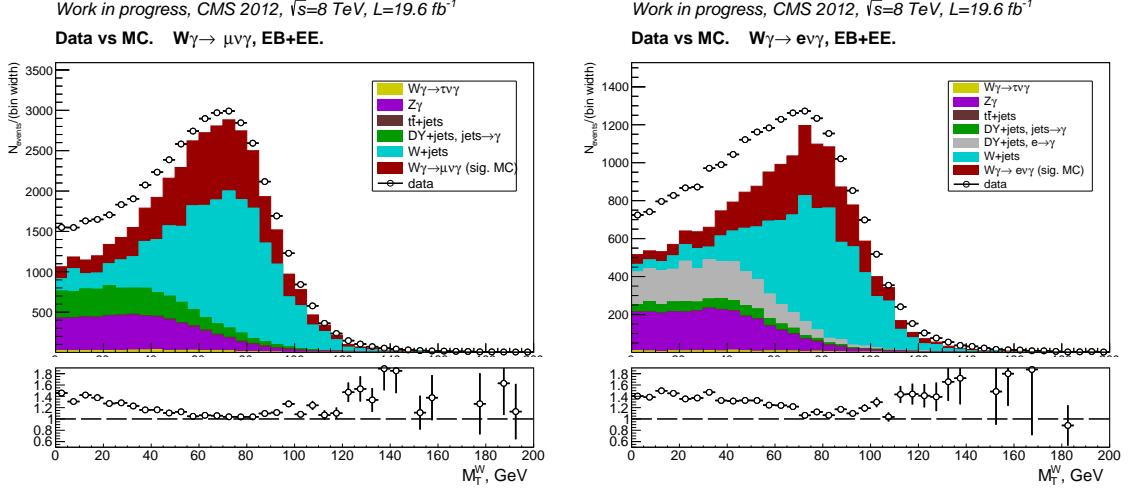


Figure 5.2:  $M_T^W$  distribution of  $W\gamma$  candidates. Data vs total MC. Left: muon channel, right: electron channel. All selection criteria except  $M_T^W$  requirement are applied on all samples that are present on the plot. The  $P_T^\gamma$  range where we do not expect any new physics is used: 15-45 GeV. The ratio plot is data divided by total MC.

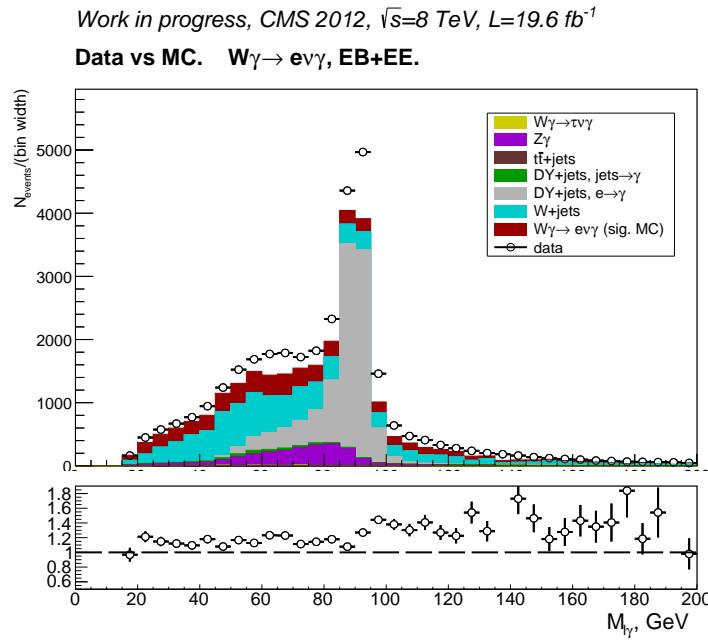


Figure 5.3:  $M_{l\gamma}$  distribution of  $W\gamma$  candidates in the electron channel. Data vs total MC. All selection criteria except Z-mass window are applied on all samples that are present on the plot. The  $P_T^\gamma$  range where we do not expect any new physics is used: 15–45 GeV. The ratio plot is data divided by total MC.

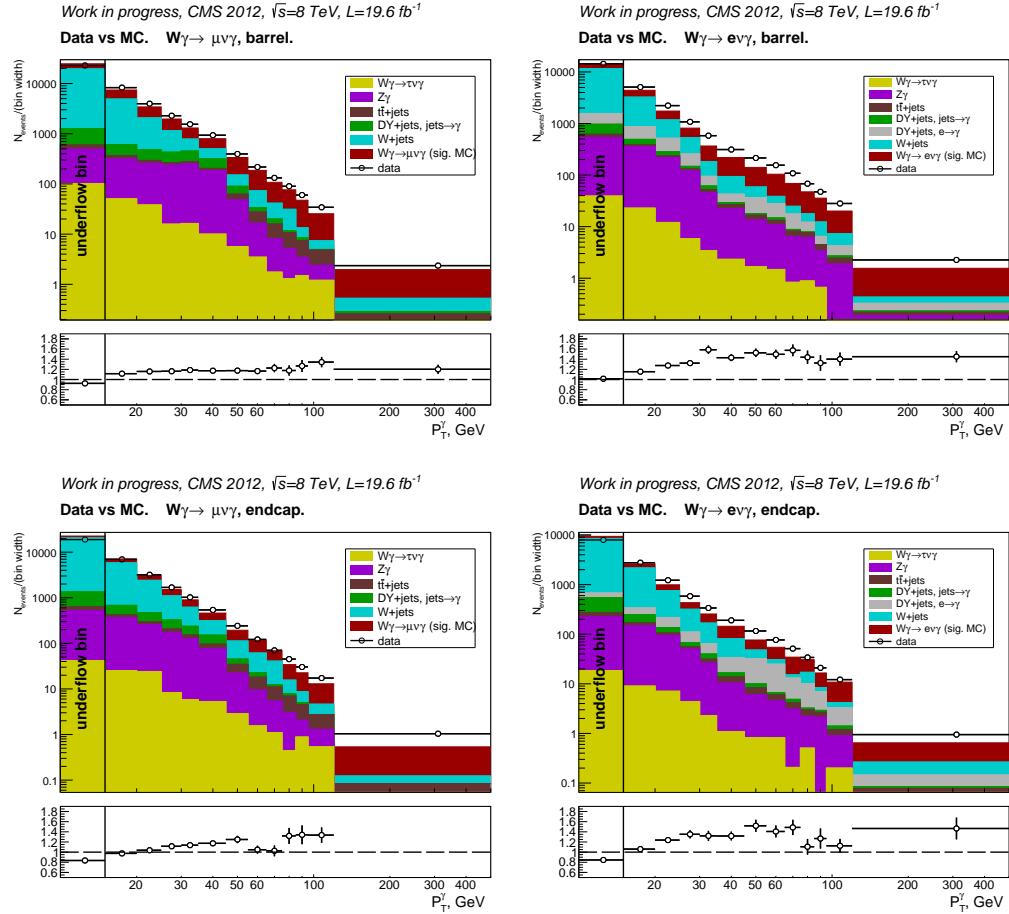


Figure 5.4:  $P_T^\gamma$  distribution of  $W\gamma$  candidates. Data vs total MC. Left column: muon channel, right column: electron channel. EB (top) and EE (bottom) photons. The ratio plots are data divided by total MC.

## 5.4 Background Estimation and Subtraction

The selected sample contains signal events as well as events coming from various backgrounds. To compute the cross section, we need to subtract background and estimate how many events in each  $P_T^\gamma$  bin originate from the  $W\gamma$  process. Main sources of backgrounds include jets misidentified as photons denoted as the jets $\rightarrow \gamma$  background, electrons misidentified as photons denoted as the  $e \rightarrow \gamma$  background, and backgrounds with real photons denoted as the real- $\gamma$  background. Jets $\rightarrow \gamma$  and real- $\gamma$  backgrounds are present in both channels while  $e \rightarrow \gamma$  background is only present in the electron channel. The remainder of this chapter describes the procedure of the background estimation and provides the results.

### 5.4.1 Background from Jets Faking Photons

The selected sample is dominated by the  $W$ +jets background which cannot be further significantly reduced without reducing our signal sample  $W\gamma$  as well. The photon ID selection criteria help to reduce  $W$ +jets background to a certain level but there is still a significant number of jets that are reconstructed as photons and pass all the photon ID criteria. DY+jets is another source of the jets $\rightarrow \gamma$  background but this source is significantly suppressed by the  $M_W^T$  selection criterion in both channels and by the Z-mass window requirement in the electron channel.

The template method is used to estimate jets $\rightarrow \gamma$  background. First of all, we choose a variable that has a significant discriminative power between the true and fake photon candidates  $V_{fit}$ . After that, we prepare real- $\gamma$  ( $T_{true}$ ) and fake- $\gamma$  ( $T_{fake}$ ) templates which should be accurate representations of  $V_{fit}$  distributions of real and fake photons in the  $W\gamma$ -selected dataset. The  $V_{fit}$  distribution in data is fitted

by the following function:

$$F(V_{fit}) = N_{true} \cdot T_{true}(V_{fit}) + N_{fake} \cdot T_{fake}(V_{fit}), \quad (5.3)$$

where  $V_{fit}$  is a fit variable,  $N_{true}$  and  $N_{fake}$  are numbers of real and fake photons in the data sample, respectively, and  $F(V_{fit})$  is a fit function.  $N_{true}$  and  $N_{fake}$  are fit parameters. We use the charged hadron isolation  $I_{ch}^{\gamma}$  and a variable representing ECal shower shape width,  $\sigma_{i\eta i\eta}^{\gamma}$ , as  $V_{fit}$ . Results of  $I_{ch}^{\gamma}$  fits are further propagated for the cross section calculation, and results of  $\sigma_{i\eta i\eta}^{\gamma}$  fits are used for the estimation of the systematic uncertainty.

The  $I_{ch}^{\gamma}$  is defined as

$$I_{ch}^{\gamma} = \sum_{ch} P_T, \quad (5.4)$$

where the sum runs over charged hadrons candidates reconstructed by the particle flow algorithm within  $\Delta R < 0.3$  from the photon.

The  $\sigma_{i\eta i\eta}^{\gamma}$  is defined as

$$\sigma_{i\eta i\eta}^{\gamma} = \frac{\sum (\eta_i - \eta)^2 w_i}{\sum (w_i)}, \quad (5.5)$$

where the sum runs over  $5 \times 5$  matrix of crystals around the crystal with the largest  $P_T$ , and  $w_i$  is the weight that has a logarithmic dependence on energy released by the photon.

To prepare templates, we use  $Z\gamma$ -selected dataset.  $Z\gamma$  goes through two different mechanisms: FSR, when a photon is radiated from one of the final state leptons, and ISR, when a photon is radiated from the initial state quark or antiquark. Real- $\gamma$  templates  $T_{true}$  are taken from the FSR events of  $Z\gamma \rightarrow \mu\mu\gamma$  while fake- $\gamma$  templates are taken from the ISR events of  $Z\gamma \rightarrow \mu\mu\gamma$ .

The FSR  $Z\gamma$  selection has muon-photon separation requirement of  $\Delta R_{min}(\mu, \gamma) >$

0.4. This requirement is chosen to be smaller than nominal separation requirement of  $\Delta R_{min}(\mu, \gamma) > 0.7$  because FSR events typically have smaller separation than ISR events and, therefore, smaller separation increases a fraction of FSR events. Three-particle invariant mass required to be  $M_{\gamma\mu\mu} < 101$  GeV.

The FSR contribution drops fast as a function of  $P_T^\gamma$ , therefore, the FSR sample has a small statistical power in high  $P_T^\gamma$  bins. To increase the statistical power, we use events from several  $P_T^\gamma$  ranges to prepare templates for each of them. The distribution of  $I_{ch}^\gamma$  of real photons does not depend on  $P_T^\gamma$  and, therefore, all events with  $P_T^\gamma > 15$  GeV are used to prepare  $I_{ch}^\gamma$  templates for all  $P_T^\gamma$  bins. Distributions of  $\sigma_{i\eta i\eta}^\gamma$  depend on  $P_T^\gamma$ . Only events of  $P_T^\gamma > 30$  GeV are combined together to prepare templates for all  $P_T^\gamma > 30$  GeV bins. Both  $I_{ch}^\gamma$  and  $\sigma_{i\eta i\eta}^\gamma$  templates are prepared separately for barrel and endcap photons.

To prepare fake- $\gamma$  templates, we need a sample that consists of jets reconstructed and identified as photons. To prepare such sample, we apply ISR  $Z\gamma$  selection conditions on double muon dataset. These selection conditions include nominal  $Z\gamma$  selection described in Ch. 5.3.2 with tighter requirement on the lepton-photon separation  $\Delta R_{min}(\mu, \gamma) > 1.0$ , and a requirement on the invariant mass of the three final state particles  $M_{ll\gamma} > 101$  GeV. A sample prepared in such a way consists of  $Z\gamma$ +(DY+jets) events, and jets from the DY+jets are reconstructed and identified as photons same as jets from  $W$ +jets, DY+jets and  $t\bar{t}$ +jets in a  $W\gamma$ -selected sample.

Therefore, to construct a fake- $\gamma$  template, we can use DY+jets events from the ISR  $Z\gamma$ -selected sample. However, in addition to DY+jets events, this sample also contains non-negligible amount of  $Z\gamma$  events. This real- $\gamma$  contribution is subtracted using  $Z\gamma$  MC predictions.

FSR and ISR selections are illustrated in App. B. Distributions of  $M_{ll\gamma}$  and  $M_{ll}$  for nominally selected  $Z\gamma$  dataset are shown in Fig. B.1. Distributions of  $\Delta R(l, \gamma)$

for ISR and FSR  $Z\gamma$  events are shown in Fig. B.2. Distributions of  $P_T^\gamma$  for ISR and FSR  $Z\gamma$  events are shown in Fig. B.3.

Fits are performed in the extended binned maximum likelihood manner separately in each  $P_T^\gamma$  bin, separately for EB and EE photons. Plots of the template fits are available in App. D. The outcomes of the fits are parameters  $N_{true}$  and  $N_{fake}$  from Eq. 5.3.

$N_{true}$  appears to be a number of real- $\gamma$  events in  $W\gamma$  dataset after all selection criteria applied except the selection condition on  $V_{fit}$  which is either  $I_{ch}^\gamma$  or  $\sigma_{inj\eta}^\gamma$ . However, our goal is to extract number of real- $\gamma$  events in  $W\gamma$  dataset after all selection criteria applied including the selection condition on  $V_{fit}$ . To extract real- $\gamma$  yield,  $N_{true}$  is multiplied by the efficiency of the selection condition on  $V_{fit}$ . The efficiency is estimated using the  $Z\gamma$ -selected FSR sample as

$$\epsilon_{V_{fit}} = \frac{N_{passed\_V_{fit\_condition}}}{N_{failed\_V_{fit\_condition}} + N_{passed\_V_{fit\_condition}}}, \quad (5.6)$$

where  $N_{passed\_V_{fit\_condition}}$  is a number of events in a specific  $P_T^\gamma$  range in the FSR sample which pass all  $Z\gamma$  FSR selection criteria including the selection condition on  $V_{fit}$ , and  $N_{failed\_V_{fit\_condition}}$  is a number of events in a specific  $P_T^\gamma$  range in the FSR sample which pass all  $Z\gamma$  FSR selection criteria except the selection condition on  $V_{fit}$ .

### 5.4.2 Background from Electrons Faking Photons in the Electron Channel

For the electron channel, DY+jets is the main source of the  $e \rightarrow \gamma$  background. The Z-mass window requirement of ( $M_{e\gamma} < 70$  GeV or  $M_{e\gamma} > 110$  GeV) significantly suppresses this background, however, the remaining contribution is non-negligible.

The contribution of  $e \rightarrow \gamma$   $N_{data-nom}^{e \rightarrow \gamma}$  is estimated separately for each  $P_T^\gamma$  bin and separately for barrel and endcap photons by scaling number of the nominally selected events in DY+jets MC sample  $N_{MC-nom}^{e \rightarrow \gamma}$  to the relation of numbers of events in the  $e \rightarrow \gamma$ -enriched data ( $N_{data-Zpeak}^{e \rightarrow \gamma}$ ) and DY+jets MC ( $N_{MC-Zpeak}^{e \rightarrow \gamma}$ ) samples under the Z-peak:

$$N_{data-nom}^{e \rightarrow \gamma} = N_{MC-nom}^{e \rightarrow \gamma} \cdot \frac{N_{data-Zpeak}^{e \rightarrow \gamma}}{N_{MC-Zpeak}^{e \rightarrow \gamma}}. \quad (5.7)$$

To estimate  $N_{data-Zpeak}^{e \rightarrow \gamma}$ ,  $e \rightarrow \gamma$ -enriched data and DY+jets MC samples are prepared by applying all  $W\gamma$  selection requirements except the Z-mass window requirement. After that, numbers of events in DY+jets MC samples  $N_{MC-Zpeak}^{e \rightarrow \gamma}$  and  $N_{MC-nom}^{e \rightarrow \gamma}$  are estimated by counting. The number  $N_{data-Zpeak}^{e \rightarrow \gamma}$  is extracted from fitting the  $M_{e\gamma}$  distribution in the Z-peak region.

The fits are performed in an extended unbinned maximum likelihood manner, separately in each  $P_T^\gamma$  bin in fine  $\eta^\gamma$  binning. The  $\eta^\gamma$  binning for different  $P_T^\gamma$  ranges is described in Tab. 5.3:

Table 5.3: Fine  $\eta^\gamma$  binning for fits for  $e \rightarrow \gamma$  background estimation.

$P_T^\gamma$ ranges, GeV	$\eta^\gamma$ binning in barrel	$\eta^\gamma$ binning in endcap
15-20-25-30-35-45-55-65	0.00-0.10-0.50-1.00-1.44	1.56-2.10-2.20-2.40-2.50
65-75-85-95	0.00-0.50-1.44	1.56-2.20-2.50
95-120-500	0.00-1.44	1.56-2.50
10-15 (underflow)	no fits; MC prediction used	

The fit model is composed by the template-based function *RooNDKeysPdf* [49] convoluted with the Gaussian distribution multiplied by the number of events from the  $e \rightarrow \gamma$  background  $N_{e \rightarrow \gamma}$  and the function describing a cumulative contribution of all other processes *RooCMSShapePdf* [50] multiplied by the number of events from other sources that have  $M_{e\gamma}$  distributions without a peak at the values of Z

mass  $N_{else}$ . The expression for the fit function is:

$$F_{fit}^{e \rightarrow \gamma} = N_{e \rightarrow \gamma} \cdot (RooNDKeysPdf * Gaussian) + N_{else} \cdot (RooCMSShapePdf). \quad (5.8)$$

The function *RooNDKeysPdf* is a function of the RooFit package [49] that creates a continuous probability distribution function out of a template. The templates are prepared from  $e \rightarrow \gamma$ -enriched DY+jets MC sample, separately for each  $P_T^\gamma$  and  $\eta^\gamma$  range. *RooNDKeysPdf* is convoluted with the Gaussian distribution, the two parameters of the Gaussian distribution are fit parameters.

The function *RooCMSShapePdf* [50] is a product of an exponential decay and a constant that smoothly turns over due to a threshold effect. *RooCMSShapePdf* is described by four parameters, and they all are used as fit parameters in  $F_{fit}^{e \rightarrow \gamma}$ . Overall,  $F_{fit}^{e \rightarrow \gamma}$  has eight fit parameters, including two parameters of the Gaussian distribution, four parameters of *RooCMSShapePdf*,  $N_{e \rightarrow \gamma}$  and  $N_{else}$ .

The fit plots are provided in App. F and the tables with numbers from Eq. 5.7 in different  $P_T^\gamma$  ranges separately for EB and EE final state photons are provided in App. H. The distributions of  $M_{e\gamma}$  in different  $P_T^\gamma$  bins are shown in App. I. After the  $M_{e\gamma}$  distribution in the DY+jets sample is scaled by

$$scale = \frac{N_{data-Zpeak}^{e \rightarrow \gamma}}{N_{MC-Zpeak}^{e \rightarrow \gamma}}, \quad (5.9)$$

the data vs MC agreement significantly improves.

### 5.4.3 Other Backgrounds

In addition to the backgrounds discussed before, there is also real- $\gamma$  background. The contributions to this background from  $Z\gamma$  and  $W\gamma \rightarrow \tau\nu\gamma$  are estimated based

on MC predictions.

- real- $\gamma$  background from  $WW\gamma$ ;
- $e \rightarrow \gamma$  background in the muon channel. Sources of these backgrounds are  $WW$  ( $W \rightarrow \mu\nu_\mu + W \rightarrow e\nu_e$ ),  $WZ$  ( $W \rightarrow \mu\nu_\mu + Z \rightarrow ee$  or  $W \rightarrow e\nu_\mu + Z \rightarrow \mu\mu$ ) and  $ZZ$  ( $Z \rightarrow \mu\mu + Z \rightarrow ee$ );
- fake lepton + real- $\gamma$  ( $\gamma$ +jets and  $\gamma\gamma$ +jets events);
- fake lepton + fake- $\gamma$  (multijets events).

#### 5.4.4 $P_T^\gamma$ Spectra before and after the Background Subtraction

The results of the background estimation and subtraction procedure are summarized in Fig. 5.5-5.6 and in Tab. 5.4-5.5. Figure 5.5 and Tab. 5.4 stand for the muon channel while Fig. 5.6 and Tab. 5.5 stand for the electron channel. Top and middle plots in Fig. 5.5-5.6 shows the  $P_T^\gamma$  spectrum in data superimposed with the signal MC and background estimates that includes jets $\rightarrow \gamma$  and real- $\gamma$  backgrounds in both channels and  $e \rightarrow \gamma$  background in the electron channel. The bottom plots show data yields after full background subtraction superimposed with signal MC.

Jets $\rightarrow \gamma$  background is estimated by two methods: by fits of  $I_{ch}^\gamma$  and  $\sigma_{inj\eta}^\gamma$  templates. The results provided by two methods differ significantly, as well as both show significant disagreement with the MC prediction. To understand the disagreements and validate the procedure, we perform two MC closure checks. The results of the checks are reported in App. E.

For the first MC closure check, the pseudodata sample is prepared from  $W+jets$  and  $W\gamma$  samples and mixed together to mimic real data. Templates are prepared from exactly the same  $W\gamma$  and  $W+jets$  samples that constitute pseudodata. Then

fits on this sample of pseudodata are performed and real- $\gamma$  and fake- $\gamma$  fit results are compared to  $W\gamma$  and  $W+\text{jets}$  MC predictions. Such approach eliminates effects of possible wrong normalizations of MC samples and the difference between real- $\gamma$  and fake- $\gamma$   $I_{ch}^\gamma$  and  $\sigma_{i\eta i\eta}^\gamma$  distributions in  $W\gamma/W+\text{jets}$  and  $Z\gamma/\text{DY+jets}$  samples. The agreement is quite good between real- $\gamma$  yields extracted from fits and the  $W\gamma$  MC prediction. However, the fake- $\gamma$  yields extracted from fits do not always agree that well with  $W+\text{jets}$  MC predictions.

The second check is more realistic: MC samples  $W\gamma$ ,  $W+\text{jets}$ ,  $Z\gamma$ ,  $Z+\text{jets}$ , and  $t\bar{t}+\text{jets}$  run through the  $W\gamma$  selection conditions and mixed together. To prepare templates,  $Z\gamma$  and  $DY+\text{jets}$  MC samples are mixed together to constitute a  $Z\gamma$ -selected pseudodataset which is used the same way as  $Z\gamma$ -selected dataset is used to prepare templates for the background estimation in the data analysis. Then the pseudodata histograms are fitted and the fit results are superimposed with MC predictions same as it is done for the real data. Such closure check eliminates the effects of possible wrong normalizations of MC samples but leaves the effect of the possible difference between real- $\gamma$  and fake- $\gamma$   $I_{ch}^\gamma$  and  $\sigma_{i\eta i\eta}^\gamma$  distributions in  $W\gamma/W+\text{jets}$  and  $Z\gamma/\text{DY+jets}$  samples. The results of this closure check show better agreement of data vs estimated background + signal MC however the disagreement in certain  $P_T^\gamma$  bins remain significant. Presumably, there is a systematic bias in our fit procedure, and it is taken into account as a difference between fit outcomes of the  $I_{ch}^\gamma$  and  $\sigma_{i\eta i\eta}^\gamma$  templates as described in detail in Ch. 5.7.

In addition to the checks described above, we also perform  $Z\gamma$  check on data and MC mixture (App. L). Both  $Z\gamma$  data analysis and  $Z\gamma$  MC closure check show very good agreement between data and background estimates + signal MC as well as between two methods of the background estimation.

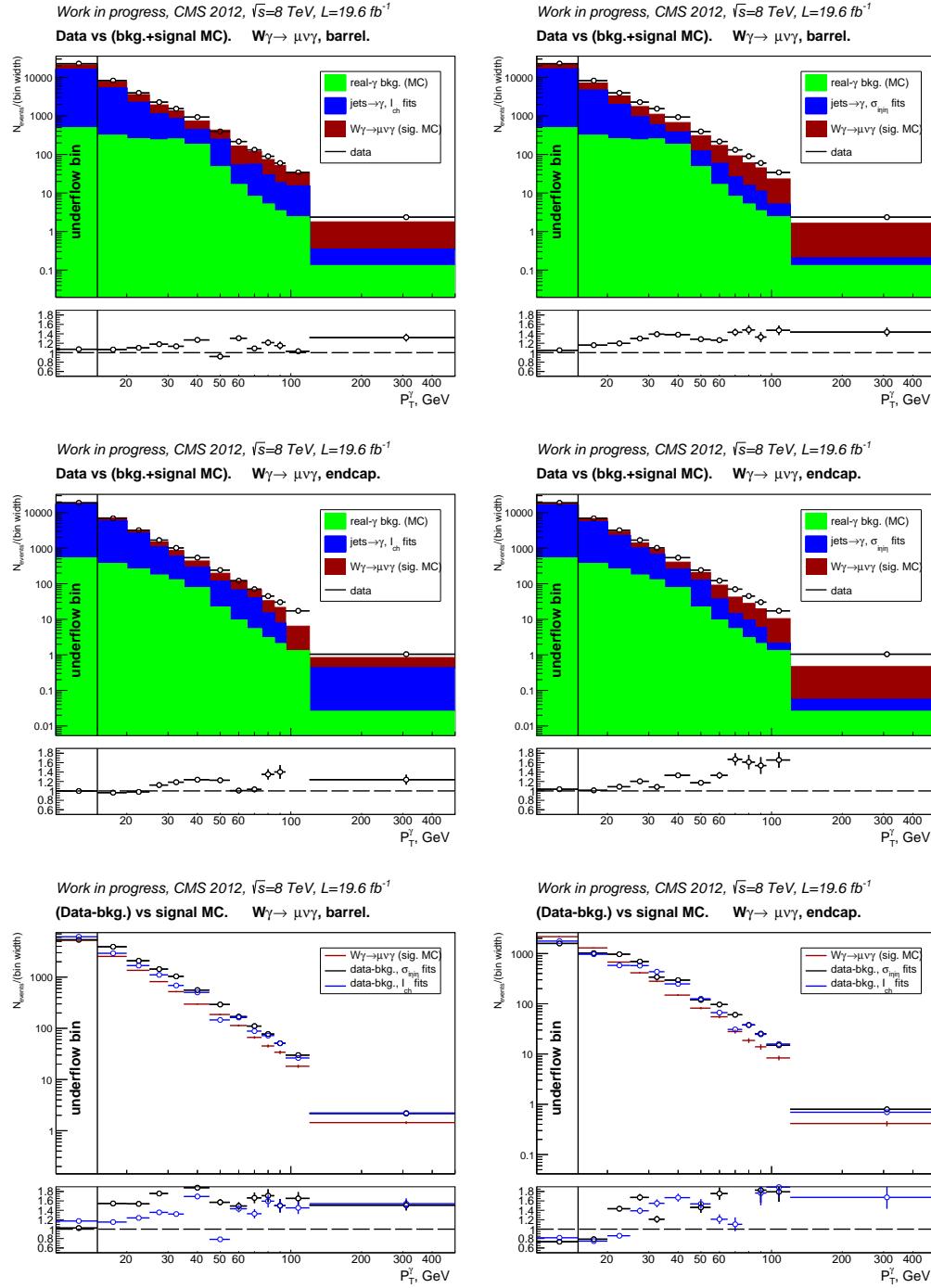


Figure 5.5: Top and middle: data vs fake- $\gamma$  background derived from the template method + real- $\gamma$  background predicted by dedicated MC samples + signal MC, with  $I_{ch}$  (left) and  $\sigma_{inj\eta}$  (right) used as fit variables in EB (top) and EE (middle). Bottom: data yields after full background subtraction vs signal MC in EB (left) and EE (right).  $I_{ch}$  vs  $\sigma_{inj\eta}$  fit results. Muon channel.

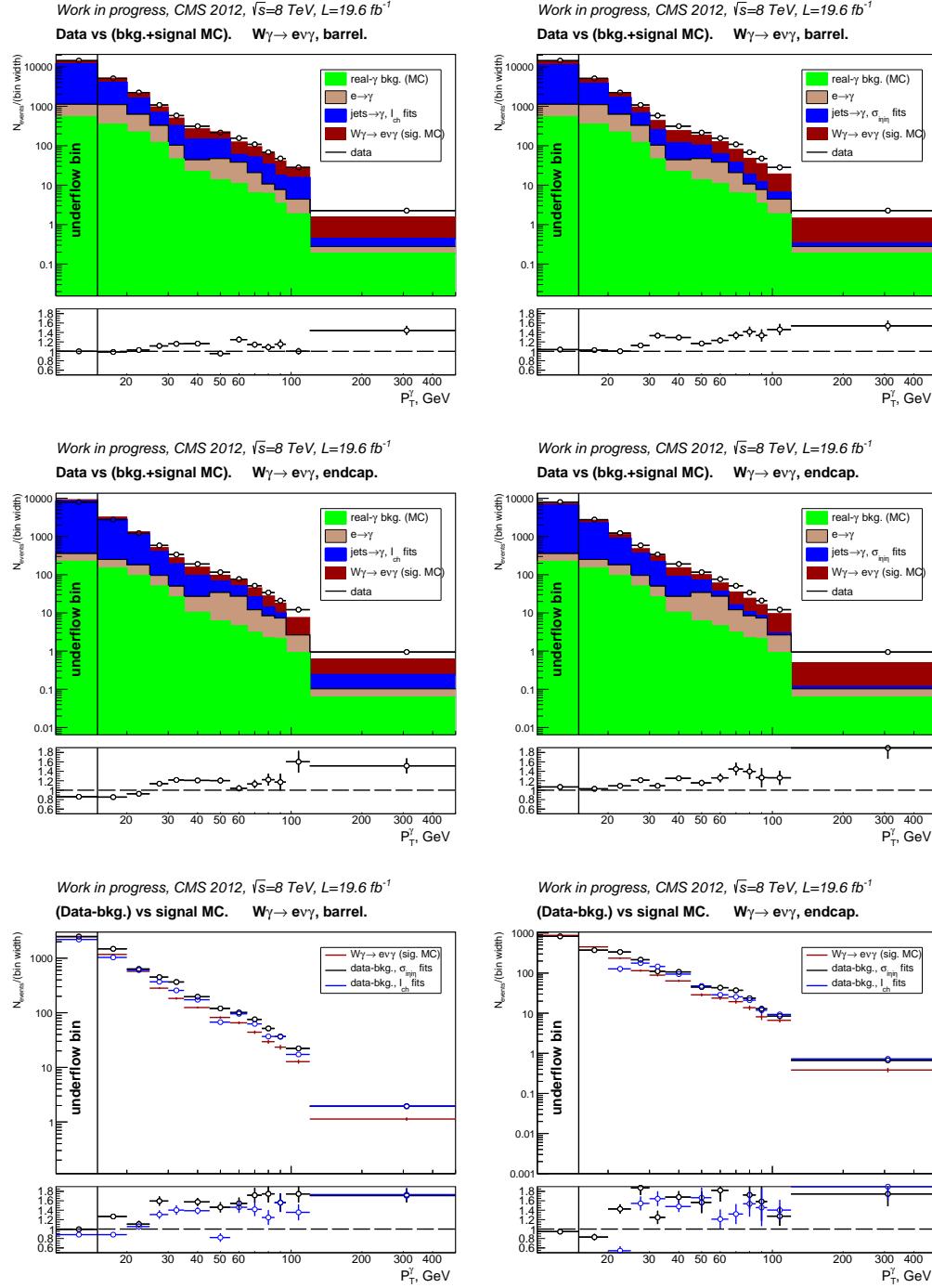


Figure 5.6: Top and middle: data vs fake- $\gamma$  background derived from the template method + real- $\gamma$  background predicted by dedicated MC samples + signal MC, with  $I_{ch}$  (left) and  $\sigma_{inj}$  (right) used as fit variables in EB (top) and EE (middle). Bottom: data yields after full background subtraction vs signal MC in EB (left) and EE (right).  $I_{ch}$  vs  $\sigma_{inj}$  fit results. Electron channel.

Table 5.4: Data, signal and background yields.  $W\gamma$ , muon channel. First column is  $P_T^\gamma$  ranges, second column is numbers of selected events in data in different  $P_T^\gamma$  regions, third and fourth columns are estimates of the jets  $\rightarrow \gamma$  background using  $I_{ch}^\gamma$  and  $\sigma_{inj}^\gamma$  templates, respectively, fifth column is MC-based estimates of the real- $\gamma$  background, sixth column is yields after the full background subtraction, and seventh column is signal MC predictions.

$P_T^\gamma$ , GeV	data yields	background estimates			data-bkg	signal MC
		$I_{ch}^\gamma$	jets $\rightarrow \gamma$	$\sigma_{inj}^\gamma$		
barrel photons						
10-15	114047 $\pm$ 338	79833 $\pm$ 251	77612 $\pm$ 322	2453 $\pm$ 46	26867 $\pm$ 385	26251 $\pm$ 240
15-20	41411 $\pm$ 203	21434 $\pm$ 150	24543 $\pm$ 143	1566 $\pm$ 34	19642 $\pm$ 245	12707 $\pm$ 164
20-25	19801 $\pm$ 141	8466 $\pm$ 101	9878 $\pm$ 78	1283 $\pm$ 31	10446 $\pm$ 168	6793 $\pm$ 120
25-30	11409 $\pm$ 107	3509 $\pm$ 68	4402 $\pm$ 46	1180 $\pm$ 27	7180 $\pm$ 123	4087 $\pm$ 93
30-35	7717 $\pm$ 88	1687 $\pm$ 49	2944 $\pm$ 40	1253 $\pm$ 27	5181 $\pm$ 99	2604 $\pm$ 75
35-45	9339 $\pm$ 97	1947 $\pm$ 56	2540 $\pm$ 33	1840 $\pm$ 33	5587 $\pm$ 114	2971 $\pm$ 80
45-55	3950 $\pm$ 63	731 $\pm$ 40	1964 $\pm$ 40	477 $\pm$ 18	2923 $\pm$ 75	1861 $\pm$ 63
55-65	2172 $\pm$ 47	415 $\pm$ 25	363 $\pm$ 11	164 $\pm$ 12	1694 $\pm$ 50	1135 $\pm$ 50
65-75	1320 $\pm$ 36	177 $\pm$ 17	466 $\pm$ 19	81 $\pm$ 8	1106 $\pm$ 39	664 $\pm$ 38
75-85	899 $\pm$ 30	103 $\pm$ 13	238 $\pm$ 16	51 $\pm$ 7	773 $\pm$ 32	452 $\pm$ 31
85-95	600 $\pm$ 24	76 $\pm$ 11	146 $\pm$ 11	34 $\pm$ 6	510 $\pm$ 26	341 $\pm$ 27
95-120	856 $\pm$ 29	67 $\pm$ 11	319 $\pm$ 25	60 $\pm$ 9	750 $\pm$ 31	454 $\pm$ 31
120-500	897 $\pm$ 30	28 $\pm$ 7	83 $\pm$ 7	50 $\pm$ 8	825 $\pm$ 31	547 $\pm$ 34
endcap photons						
10-15	94370 $\pm$ 307	77649 $\pm$ 215	81161 $\pm$ 426	2632 $\pm$ 41	7937 $\pm$ 306	10823 $\pm$ 154
15-20	34643 $\pm$ 186	25902 $\pm$ 142	27661 $\pm$ 184	1835 $\pm$ 34	5089 $\pm$ 213	6474 $\pm$ 120
20-25	15988 $\pm$ 126	10018 $\pm$ 93	11659 $\pm$ 102	1294 $\pm$ 29	4842 $\pm$ 138	3377 $\pm$ 86
25-30	8429 $\pm$ 92	4061 $\pm$ 58	4558 $\pm$ 51	871 $\pm$ 23	3460 $\pm$ 98	2068 $\pm$ 67
30-35	5110 $\pm$ 71	2669 $\pm$ 54	2268 $\pm$ 32	641 $\pm$ 19	1700 $\pm$ 81	1404 $\pm$ 56
35-45	5414 $\pm$ 74	1807 $\pm$ 44	2113 $\pm$ 29	771 $\pm$ 22	2957 $\pm$ 83	1489 $\pm$ 57
45-55	2422 $\pm$ 49	1025 $\pm$ 50	936 $\pm$ 19	222 $\pm$ 13	1196 $\pm$ 68	819 $\pm$ 43
55-65	1217 $\pm$ 35	270 $\pm$ 19	564 $\pm$ 16	94 $\pm$ 9	966 $\pm$ 36	551 $\pm$ 35
65-75	703 $\pm$ 27	87 $\pm$ 11	346 $\pm$ 13	54 $\pm$ 7	604 $\pm$ 28	280 $\pm$ 25
75-85	451 $\pm$ 21	63 $\pm$ 9	117 $\pm$ 6	30 $\pm$ 5	379 $\pm$ 22	186 $\pm$ 20
85-95	303 $\pm$ 17	37 $\pm$ 7	56 $\pm$ 4	21 $\pm$ 5	255 $\pm$ 18	139 $\pm$ 18
95-120	433 $\pm$ 21	20 $\pm$ 5	-81 $\pm$ 5	32 $\pm$ 6	374 $\pm$ 21	209 $\pm$ 22
120-500	396 $\pm$ 20	11 $\pm$ 4	153 $\pm$ 12	10 $\pm$ 2	302 $\pm$ 18	157 $\pm$ 19

Table 5.5: Data, signal and background yields.  $W\gamma$ , electron channel. First column is  $P_T^\gamma$  ranges, second column is numbers of selected events in data in different  $P_T^\gamma$  regions, third and fourth columns are estimates of the jets  $\rightarrow \gamma$  background using  $I_{ch}^\gamma$  and  $\sigma_{inj}^\gamma$  templates, respectively, fifth column is  $e \rightarrow \gamma$  background estimates, sixths column is MC-based estimates of the real- $\gamma$  background, seventh column is yields after the full background subtraction, and eighth column is signal MC predictions.

$P_T^\gamma$ , GeV	data yields	background estimates				MC real- $\gamma$	data-bkg	signal MC
		$I_{ch}^\gamma$	jets $\rightarrow \gamma$	$\sigma_{inj}^\gamma$	$e \rightarrow \gamma$			
barrel photons								
10-15	$71649 \pm 268$	$51004 \pm 200$	$53577 \pm 266$	$2923 \pm 80$	$2688 \pm 41$	$12425 \pm 316$	$12480 \pm 164$	
15-20	$25455 \pm 160$	$13487 \pm 118$	$14474 \pm 110$	$3715 \pm 178$	$1779 \pm 32$	$7422 \pm 262$	$5858 \pm 110$	
20-25	$11130 \pm 105$	$5112 \pm 78$	$4846 \pm 55$	$2023 \pm 137$	$1101 \pm 25$	$3168 \pm 186$	$2869 \pm 77$	
25-30	$5388 \pm 73$	$1748 \pm 47$	$1790 \pm 29$	$1031 \pm 72$	$603 \pm 18$	$2251 \pm 111$	$1412 \pm 54$	
30-35	$2907 \pm 54$	$752 \pm 32$	$1079 \pm 24$	$286 \pm 33$	$229 \pm 12$	$1831 \pm 68$	$916 \pm 44$	
35-45	$3128 \pm 56$	$735 \pm 34$	$1003 \pm 21$	$215 \pm 27$	$223 \pm 12$	$1965 \pm 70$	$1248 \pm 51$	
45-55	$2147 \pm 46$	$551 \pm 31$	$964 \pm 28$	$335 \pm 37$	$134 \pm 10$	$1200 \pm 65$	$821 \pm 42$	
55-65	$1556 \pm 39$	$228 \pm 19$	$211 \pm 8$	$272 \pm 39$	$108 \pm 9$	$1011 \pm 57$	$654 \pm 37$	
65-75	$1083 \pm 33$	$163 \pm 16$	$300 \pm 15$	$143 \pm 27$	$64 \pm 7$	$757 \pm 44$	$441 \pm 31$	
75-85	$680 \pm 26$	$79 \pm 11$	$224 \pm 15$	$45 \pm 13$	$62 \pm 7$	$516 \pm 31$	$295 \pm 26$	
85-95	$473 \pm 22$	$43 \pm 9$	$99 \pm 9$	$43 \pm 17$	$34 \pm 5$	$366 \pm 29$	$234 \pm 23$	
95-120	$703 \pm 27$	$53 \pm 9$	$274 \pm 24$	$63 \pm 19$	$47 \pm 5$	$555 \pm 34$	$318 \pm 27$	
120-500	$859 \pm 29$	$23 \pm 6$	$61 \pm 6$	$34 \pm 12$	$71 \pm 8$	$735 \pm 33$	$430 \pm 31$	
endcap photons								
10-15	$39746 \pm 199$	$31043 \pm 138$	$40022 \pm 13$	$666 \pm 38$	$1120 \pm 27$	$4130 \pm 204$	$4368 \pm 97$	
15-20	$13818 \pm 118$	$9920 \pm 88$	$12692 \pm 124$	$509 \pm 56$	$744 \pm 21$	$1870 \pm 145$	$2253 \pm 68$	
20-25	$6133 \pm 78$	$3538 \pm 56$	$4558 \pm 63$	$433 \pm 36$	$473 \pm 17$	$1680 \pm 92$	$1177 \pm 49$	
25-30	$2924 \pm 54$	$1358 \pm 34$	$1516 \pm 29$	$229 \pm 24$	$250 \pm 12$	$1079 \pm 62$	$575 \pm 34$	
30-35	$1690 \pm 41$	$850 \pm 31$	$694 \pm 18$	$120 \pm 16$	$130 \pm 9$	$555 \pm 49$	$445 \pm 31$	
35-45	$1905 \pm 44$	$613 \pm 26$	$670 \pm 16$	$167 \pm 19$	$103 \pm 8$	$1071 \pm 51$	$638 \pm 37$	
45-55	$1162 \pm 34$	$377 \pm 30$	$337 \pm 11$	$281 \pm 28$	$61 \pm 6$	$450 \pm 53$	$287 \pm 24$	
55-65	$767 \pm 28$	$98 \pm 12$	$228 \pm 11$	$227 \pm 28$	$46 \pm 6$	$433 \pm 40$	$238 \pm 22$	
65-75	$513 \pm 23$	$40 \pm 8$	$139 \pm 9$	$90 \pm 18$	$31 \pm 4$	$372 \pm 29$	$194 \pm 21$	
75-85	$340 \pm 18$	$22 \pm 6$	$57 \pm 5$	$62 \pm 15$	$22 \pm 5$	$236 \pm 24$	$137 \pm 18$	
85-95	$210 \pm 14$	$11 \pm 4$	$25 \pm 3$	$52 \pm 19$	$21 \pm 3$	$129 \pm 24$	$81 \pm 14$	
95-120	$304 \pm 17$	$8 \pm 3$	$-43 \pm 4$	$43 \pm 13$	$23 \pm 5$	$212 \pm 22$	$166 \pm 20$	
120-500	$360 \pm 19$	$5 \pm 3$	$53 \pm 7$	$15 \pm 7$	$24 \pm 4$	$254 \pm 18$	$146 \pm 19$	

## 5.5 Detector Resolution Unfolding

The finite detector resolution in  $P_T^\gamma$  causes the bin-by-bin migration in the  $P_T^\gamma$  spectrum. The reconstructed  $P_T^{\gamma(reco)}$  may not coincide with the true  $P_T^{\gamma(true)}$ , and, therefore, the event reconstructed in a  $P_T^\gamma$  bin  $j$  may, in fact, belong to the bin  $i \neq j$ . To recover the true  $P_T^\gamma$  spectrum, we apply the procedure of the detector resolution unfolding.

The reconstructed  $P_T^\gamma$  spectrum is related to true  $P_T^\gamma$  spectrum as:

$$N_j^{reco} = R_{ji} N_i^{true}, \quad (5.10)$$

where  $N_j$  and  $N_i$  are numbers of events in a given  $P_T^{\gamma(reco)}$  and  $P_T^{\gamma(true)}$  bins, respectively,  $R_{ji}$  is the “response matrix” where each element is the probability of an event with true  $P_T^\gamma$  in the bin “ $i$ ” to be reconstructed with  $P_T^\gamma$  in the bin “ $j$ ”.

To recover the true  $P_T^{\gamma(true)}$  spectrum, we use the D’Agostini method [48] which is based on Bayes’ theorem and unfolds the reconstructed spectrum iteratively. D’Agostini method was chosen because it is recommended by the CMS SMP group in contrast to the matrix inversion method ( $N_i^{true} = (R_{ji})^{-1} N_j^{reco}$ ) to avoid problems with possible matrix singularity and large statistical fluctuations.

Unfolding constants are derived from the signal MC sample ( $W\gamma \rightarrow \mu\nu_\mu\gamma/W\gamma \rightarrow e\nu_e\gamma$ ) where both true (gen-level) and reconstructed  $P_T^\gamma$  spectra are known. Using the signal MC sample, we prepare the migration matrix  $M_{ji}$  which has the number of selected signal events in each  $[j,i]$  bin.

After that, we provide the migration matrix  $M_{ji}$ , generated and reconstructed yields from signal MC  $N_i^{gen-MC}$  and  $N_j^{reco-MC}$  and reconstructed yields from data to RooUnfold class [54] which performs unfolding using D’Agostini method with five iterations. Yields before and after detector resolution unfolding are compared

in Tab. 5.6 for the muon channel and in Tab. 5.7 for the electron channel.

Table 5.6: Unfolding Results in  $W\gamma$ , muon channel. First column:  $P_T^\gamma$  ranges, second column:  $W\gamma$  differential yields after background subtraction before unfolding, third column:  $W\gamma$  yields after unfolding.

$P_T^\gamma$ , GeV	yields data-bkg	yields unfolded
10 - 15	$39621 \pm 678$	$38843 \pm 718$
15 - 20	$19449 \pm 361$	$19662 \pm 438$
20 - 25	$11315 \pm 230$	$11443 \pm 275$
25 - 30	$8417 \pm 160$	$8714 \pm 196$
30 - 35	$5613 \pm 128$	$5529 \pm 155$
35 - 45	$7518 \pm 133$	$7895 \pm 149$
45 - 55	$2716 \pm 95$	$2605 \pm 108$
55 - 65	$2293 \pm 62$	$2307 \pm 70$
65 - 75	$1191 \pm 53$	$1198 \pm 62$
75 - 85	$1101 \pm 41$	$1165 \pm 48$
85 - 95	$757 \pm 33$	$776 \pm 41$
95 - 120	$1054 \pm 44$	$1064 \pm 50$
120 - 500	$1107 \pm 39$	$1141 \pm 40$

Table 5.7: Unfolding Results in  $W\gamma$ , electron channel. First column:  $P_T^\gamma$  ranges, second column:  $W\gamma$  differential yields after background subtraction before unfolding, third column:  $W\gamma$  yields after unfolding.

$P_T^\gamma$ , GeV	yields data-bkg	yields unfolded
10 - 15	$9209 \pm 378$	$9192 \pm 413$
15 - 20	$4920 \pm 319$	$4850 \pm 380$
20 - 25	$3660 \pm 212$	$3698 \pm 249$
25 - 30	$2734 \pm 127$	$2948 \pm 159$
30 - 35	$2015 \pm 84$	$2075 \pm 102$
35 - 45	$2677 \pm 83$	$2770 \pm 91$
45 - 55	$1152 \pm 81$	$1116 \pm 93$
55 - 65	$1244 \pm 70$	$1214 \pm 80$
65 - 75	$881 \pm 56$	$911 \pm 63$
75 - 85	$579 \pm 42$	$590 \pm 48$
85 - 95	$483 \pm 38$	$490 \pm 45$
95 - 120	$664 \pm 46$	$692 \pm 50$
120 - 500	$1020 \pm 40$	$1052 \pm 40$

After  $P_T^\gamma$  spectrum is unfolded, measurements in different  $P_T^\gamma$  bins become correlated. Correlation matrices are shown in Fig. 5.8.

For illustration purpose, we in addition to the migration matrix we also prepare the response matrix  $R_{ji}$  (Fig. 5.7) by normalizing the migration matrix in each  $j$  bin

to all events reconstructed in this bin such as

$$R_{ji} = \frac{M_{ji}}{\sum_i M_{ji}}. \quad (5.11)$$

The response matrix is shown in Fig. 5.7.

To validate the detector resolution unfolding procedure, we perform the MC closure check. Gen-level and reconstructed yields are prepared using the signal MC. Then reconstructed yields are smeared by the Gaussian distribution according to the errors on the yields. The smeared yields are unfolded and compared to the gen-level yields. In addition to the D'Agostini method, we check the performance of the matrix inversion method for the unfolding which recovers the true yields as  $N_i^{true} = (R_{ji})^{-1} N_j^{reco}$ .

The results of the MC closure check are summarized in Tab. 5.8-5.9 for the muon and electron channels respectively. The unfolded yields show reasonable agreement to the gen-level yields except the underflow bin (10 – 15 GeV). The disagreement in the underflow bin may be caused by migration between  $P_T^\gamma < 10$  GeV and  $10 < P_T^\gamma < 15$  GeV ranges because events with  $P_T^\gamma < 10$  GeV are not available to us.

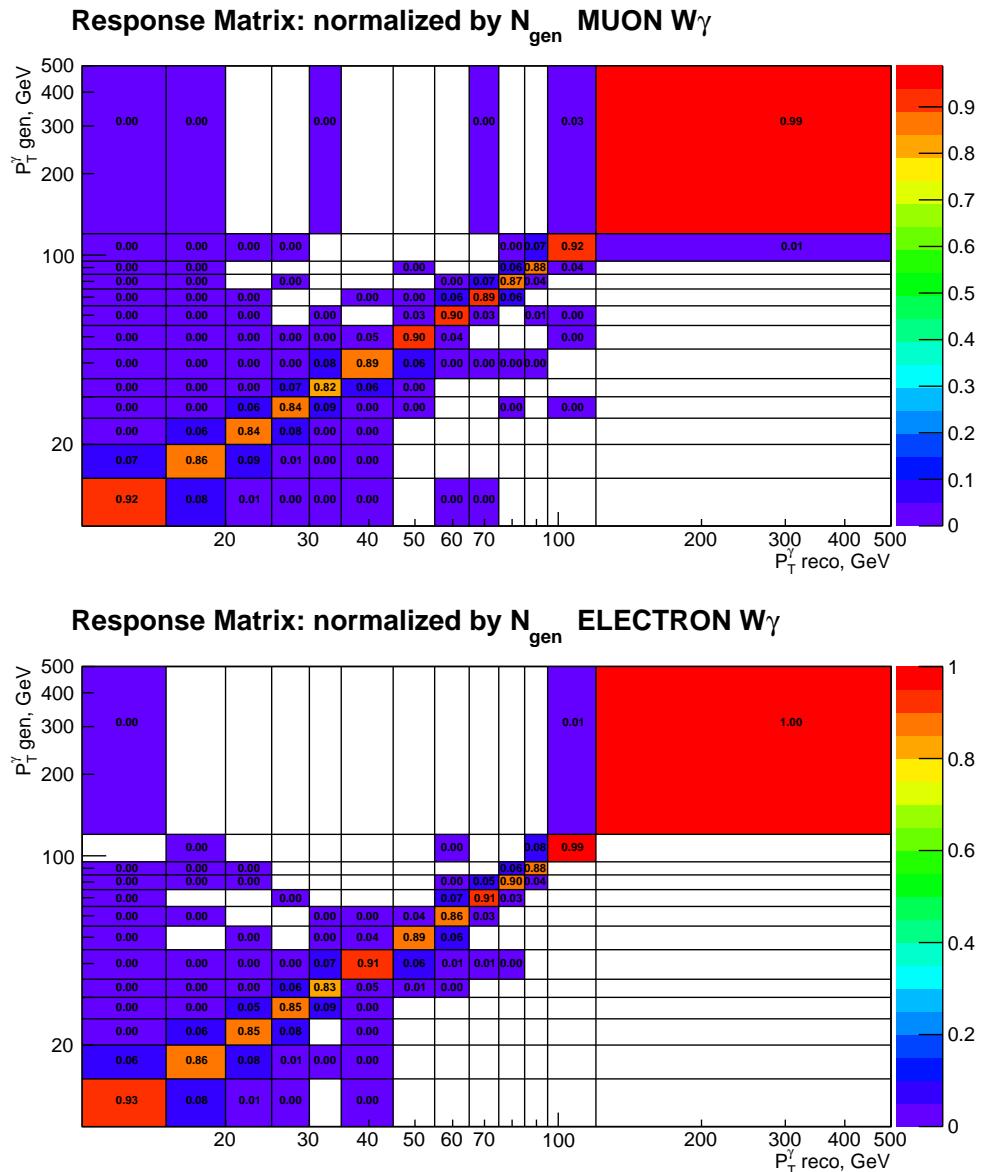


Figure 5.7: Response matrix derived from the signal MC.

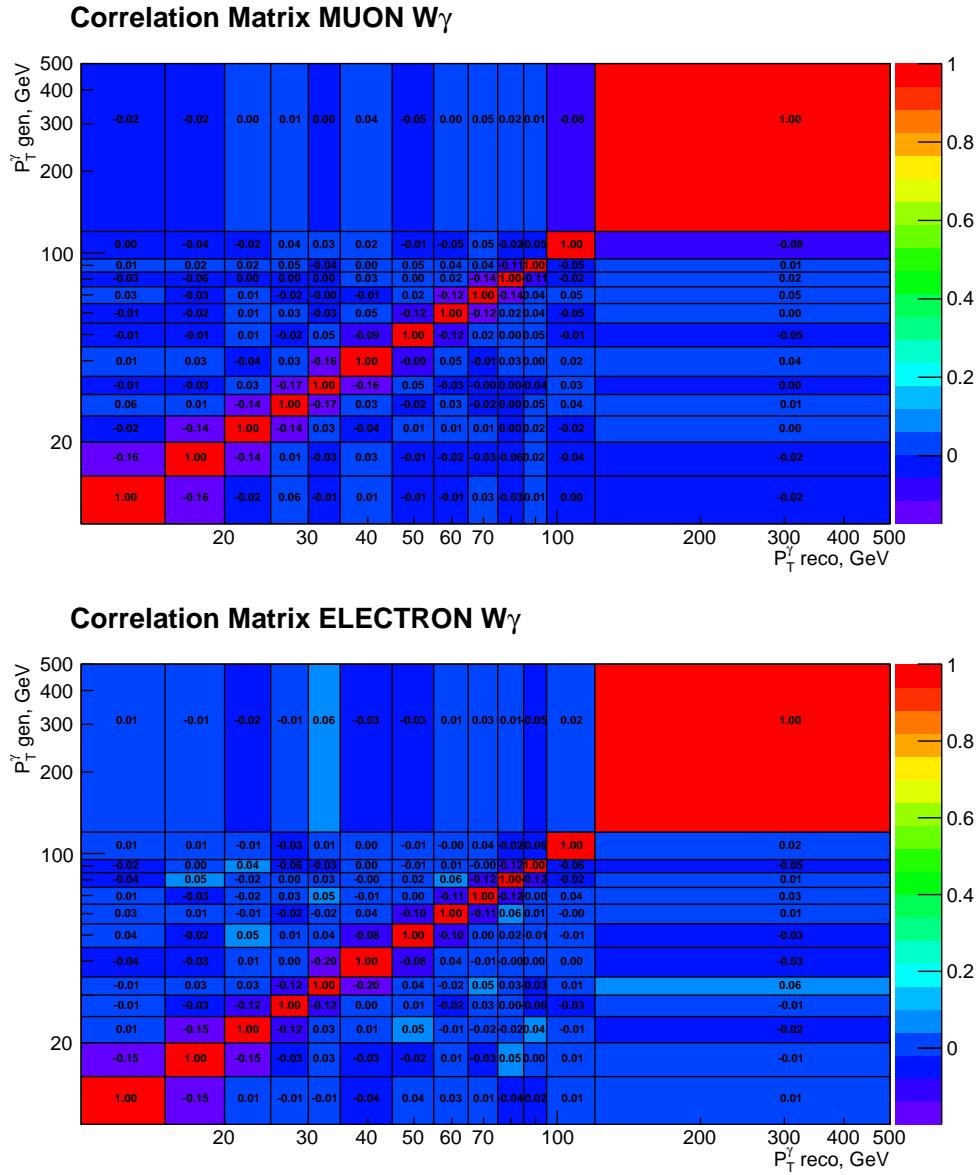


Figure 5.8: Correlation matrices of statistical uncertainties on unfolded  $W\gamma$  yields in the muon (top) and electron (bottom) channels.

Table 5.8: Unfolding, MC closure test.  $W\gamma$ , muon channel. The first column is  $P_T^\gamma$  range, the second column is true selected yields in the signal MC sample based on the generator level information, the third column is reconstructed and smeared yields in the signal MC sample, the fourth column is yields from the third column unfolded by the D'Agostini method, and the fifth column is yields from the third column unfolded by the matrix inversion method.

$P_T^\gamma$ , GeV	yields gen-level	yields rec	unfolded inversion	unfolded D'Agostini
10 - 15	$33888 \pm 273$	$37074 \pm 286$	$36226 \pm 206$	$36222 \pm 204$
15 - 20	$19736 \pm 207$	$19181 \pm 203$	$19612 \pm 171$	$19619 \pm 169$
20 - 25	$10364 \pm 149$	$10171 \pm 148$	$10358 \pm 122$	$10354 \pm 119$
25 - 30	$6254 \pm 116$	$6156 \pm 115$	$6233 \pm 96$	$6234 \pm 96$
30 - 35	$4026 \pm 93$	$4007 \pm 93$	$4010 \pm 81$	$4010 \pm 78$
35 - 45	$4516 \pm 99$	$4461 \pm 98$	$4502 \pm 79$	$4502 \pm 79$
45 - 55	$2731 \pm 77$	$2680 \pm 76$	$2724 \pm 57$	$2724 \pm 60$
55 - 65	$1662 \pm 60$	$1686 \pm 61$	$1655 \pm 45$	$1655 \pm 46$
65 - 75	$987 \pm 46$	$945 \pm 45$	$979 \pm 38$	$979 \pm 35$
75 - 85	$659 \pm 38$	$638 \pm 37$	$654 \pm 30$	$653 \pm 30$
85 - 95	$495 \pm 33$	$480 \pm 32$	$489 \pm 27$	$489 \pm 25$
95 - 120	$664 \pm 38$	$663 \pm 38$	$661 \pm 28$	$661 \pm 28$
120 - 500	$726 \pm 40$	$704 \pm 39$	$720 \pm 26$	$720 \pm 27$
500 - 2000	$2 \pm 2$	$2 \pm 2$	$2 \pm 1$	$2 \pm 1$

Table 5.9: Unfolding, MC closure test.  $W\gamma$ , electron channel. The first column is  $P_T^\gamma$  range, the second column is true selected yields in the signal MC sample based on the generator level information, the third column is reconstructed and smeared yields in the signal MC sample, the fourth column is yields from the third column unfolded by the D'Agostini method, and the fifth column is yields from the third column unfolded by the matrix inversion method.

$P_T^\gamma$ , GeV	yields gen-level	yields rec	unfolded inversion	unfolded D'Agostini
10 - 15	$16025 \pm 185$	$16849 \pm 190$	$17117 \pm 143$	$17116 \pm 141$
15 - 20	$8246 \pm 131$	$8111 \pm 130$	$8194 \pm 109$	$8196 \pm 108$
20 - 25	$4093 \pm 92$	$4046 \pm 92$	$4083 \pm 75$	$4082 \pm 74$
25 - 30	$2080 \pm 66$	$1987 \pm 64$	$2072 \pm 55$	$2072 \pm 55$
30 - 35	$1387 \pm 54$	$1361 \pm 54$	$1378 \pm 47$	$1378 \pm 46$
35 - 45	$1925 \pm 64$	$1886 \pm 63$	$1915 \pm 51$	$1915 \pm 50$
45 - 55	$1124 \pm 49$	$1108 \pm 48$	$1116 \pm 37$	$1116 \pm 38$
55 - 65	$855 \pm 42$	$892 \pm 43$	$848 \pm 33$	$848 \pm 34$
65 - 75	$655 \pm 38$	$635 \pm 37$	$649 \pm 30$	$649 \pm 28$
75 - 85	$447 \pm 32$	$433 \pm 32$	$442 \pm 24$	$442 \pm 24$
85 - 95	$316 \pm 27$	$316 \pm 27$	$311 \pm 21$	$311 \pm 20$
95 - 120	$507 \pm 34$	$484 \pm 33$	$501 \pm 23$	$501 \pm 23$
120 - 500	$593 \pm 37$	$575 \pm 36$	$587 \pm 23$	$587 \pm 24$
500 - 2000	$4 \pm 3$	$4 \pm 3$	$4 \pm 2$	$4 \pm 2$

## 5.6 Acceptance and Efficiency Correction

The selection requirements in the  $W\gamma$  measurement are stricter than the phase space requirements, thus, during the selection procedure, we lose a large number of signal events that are within our phase space. Our selection criteria intend to reduce various backgrounds, however, at the same time, they remove many signal events as well. The ratio between the number of selected signal events and the number of signal events reconstructed within the phase space is called a “selection efficiency”. In addition to the effect described above, a certain number of events are truly within our phase space but are reconstructed outside of the phase space and vice versa. A ratio between the number of signal events that are reconstructed within our phase space and the number of events that truly appear within our phase space is called a “reconstruction efficiency”. Finally, certain events that are truly within the phase space may not be caught by the detector due to the detector acceptance restrictions. Examples of such events include events with final state photons or electrons that goes into the gap between the EB and EE, with corresponding  $1.44 < |\eta^{\gamma,e}| < 1.56$ . A ratio between the number of events truly reconstructed within the phase space and the number of events that are also caught by the detector is called “acceptance”.

To correct our selected, background-subtracted, unfolded yields from Tab. 5.6-5.7 for these effects, we introduce a correction  $A \times \epsilon$  that accumulates all three effects. The correction is estimated using the signal MC sample, separately for the total yield and for each  $P_T^\gamma$  bin of the differential yields.

The numerator  $N^{A\epsilon}$  for the correction of the total yield is determined as the number of selected events in the signal MC with PU weight applied. The numerator  $N_i^{A\epsilon}$  for the correction of the differential yields is determined as selected signal

MC yields with PU weight applied in  $P_T^{\gamma\text{-}gen}$  bins at the gen-level. Index  $i$  stands for  $P_T^\gamma$  bins.

The denominator  $D^{A\epsilon}$  of the  $A \times \epsilon$  correction is determined as the number of events that are within the phase space based on their gen-level kinematic values. For the correction  $(A \times \epsilon)_i$  of the differential yields, the numbers  $D_i^{A\epsilon}$  are determined separately for each  $P_T^\gamma$  bin.

The  $A \times \epsilon$  correction is determined then as  $A \times \epsilon = N^{A\epsilon}/D^{A\epsilon}$  for the total cross section and as  $(A \times \epsilon)_i = N_i^{A\epsilon}/D_i^{A\epsilon}$  for the differential cross section where index  $i$  stands for a  $P_T^\gamma$  bin. The  $A \times \epsilon$  for the total cross section are  $0.2891 \pm 0.0006$  for the muon channel and  $0.1229 \pm 0.0004$  for the electron channel. The uncertainties are determined by the statistical power of the  $W\gamma$  MC sample. The values of the  $(A \times \epsilon)_i$  correction for the differential yields are plotted in Fig. 5.6.

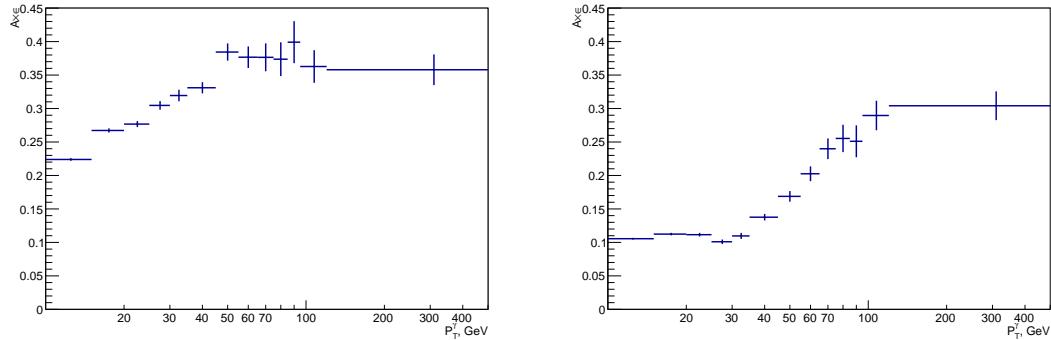


Figure 5.9:  $A \times \epsilon$  corrections in the muon (left) and electron (right) channels. Plots are produced with  $W\gamma$  MC sample at  $\sqrt{s} = 8$  TeV.

## 5.7 Systematic Uncertainties

### 5.7.1 Uncertainties Related to Jets $\rightarrow \gamma$ Background Estimation

The dominant systematic uncertainty is uncertainty due to bias in the procedure of the jets $\rightarrow \gamma$  background estimation. The following sources contribute to the uncertainty of the jets $\rightarrow \gamma$  background estimation:

- the biases in the template shapes and the fit procedure;
- the uncertainty on the normalizations of  $Z\gamma$  and DY+jets MC samples when the real- $\gamma$  (fake- $\gamma$ ) portions are subtracted from the ISR (FSR) templates.
- the limited statistical power of the fake- $\gamma$  and real- $\gamma$  templates.

The systematic uncertainty due to the biases in the template shapes and the fit procedure is computed as the difference between fit results of  $I_{ch}^{\gamma}$  and  $\sigma_{i\eta i\eta}^{\gamma}$  distributions

$$err = |N_{Ich} - N_{\sigma i\eta i\eta}|. \quad (5.12)$$

The uncertainty due to template statistics is computed by separately randomizing the real- $\gamma$  and the fake- $\gamma$  templates, we prepare 20 real- $\gamma$  and 100 fake- $\gamma$  templates by randomizing our nominal templates with the Gaussian distribution. Then we perform fits with new templates and take the standard deviation of the fit results as an uncertainty. The uncertainties are computed separately for the real- $\gamma$  and the fake- $\gamma$  templates. The statistical uncertainty of the fake- $\gamma$  template is larger.

The results of the systematic uncertainty of  $|N_{Ich} - N_{\sigma i\eta i\eta}|$  and the template statistical uncertainty are summarized in the Tab. 5.10-5.11. The column “yield

“data-bkg” is the background subtracted yield which is used for the cross section measurement. The central value of this yields is taken from the “data  $I_{ch}^\gamma$ ” column.

Table 5.10: Results of the background subtraction based on fits of different variables.  $W\gamma$ , muon channel. First column:  $P_T^\gamma$  ranges, second column: signal MC prediction, third and fourth columns: signal yields extracted from fits of  $I_{ch}^\gamma$  and  $\sigma_{ij\eta i\eta}^\gamma$  distributions of data, fifth and sixth columns: signal yields extracted from MC closure tests, seventh column: signal yield that is used in the next measurement steps.

$P_T^\gamma$ , GeV	sig. MC ( $W\gamma \rightarrow \mu\nu\gamma$ )	data		MC closure		yield data-bkg.
		$I_{ch}^\gamma$	$\sigma_{ij\eta i\eta}^\gamma$	$I_{ch}^\gamma$	$\sigma_{ij\eta i\eta}^\gamma$	
barrel photons						
10-15	$26250 \pm 240$	$30683 \pm 1919$	$26770 \pm 3134$	$29657 \pm 2476$	$35073 \pm 3726$	$30683 \pm 3913 \pm 1865$
15-20	$12706 \pm 164$	$14554 \pm 1070$	$19575 \pm 1771$	$10742 \pm 1079$	$14924 \pm 2123$	$14554 \pm 5021 \pm 1041$
20-25	$6793 \pm 120$	$8364 \pm 711$	$10398 \pm 4313$	$7698 \pm 626$	$9399 \pm 1741$	$8364 \pm 2033 \pm 693$
25-30	$4087 \pm 93$	$5505 \pm 685$	$7141 \pm 3437$	$4421 \pm 532$	$5023 \pm 2094$	$5505 \pm 1636 \pm 675$
30-35	$2603 \pm 74$	$3409 \pm 500$	$5152 \pm 2581$	$3422 \pm 197$	$3266 \pm 1156$	$3409 \pm 1742 \pm 490$
35-45	$2971 \pm 80$	$4990 \pm 466$	$5544 \pm 3366$	$4472 \pm 308$	$4351 \pm 1632$	$4990 \pm 554 \pm 454$
45-55	$1861 \pm 63$	$1428 \pm 410$	$2893 \pm 593$	$1799 \pm 277$	$2464 \pm 146$	$1428 \pm 1464 \pm 402$
55-65	$1135 \pm 49$	$1604 \pm 207$	$1671 \pm 501$	$1143 \pm 214$	$1475 \pm 311$	$1604 \pm 67 \pm 201$
65-75	$664 \pm 37$	$866 \pm 43$	$1090 \pm 271$	$773 \pm 193$	$680 \pm 162$	$866 \pm 223 \pm 7$
75-85	$451 \pm 31$	$710 \pm 34$	$762 \pm 88$	$621 \pm 106$	$701 \pm 143$	$710 \pm 52 \pm 0$
85-95	$340 \pm 27$	$502 \pm 139$	$502 \pm 175$	$455 \pm 62$	$443 \pm 98$	$502 \pm 0 \pm 136$
95-120	$453 \pm 31$	$641 \pm 105$	$732 \pm 31$	$714 \pm 113$	$577 \pm 83$	$641 \pm 91 \pm 98$
120-500	$546 \pm 34$	$810 \pm 214$	$792 \pm 63$	$777 \pm 105$	$678 \pm 191$	$810 \pm 18 \pm 211$
endcap photons						
10-15	$10823 \pm 154$	$8797 \pm 2242$	$7893 \pm 2947$	$16512 \pm 2900$	$-2675 \pm 1967$	$8797 \pm 903 \pm 2184$
15-20	$6474 \pm 119$	$4793 \pm 1132$	$5053 \pm 1518$	$6454 \pm 1142$	$2650 \pm 2124$	$4793 \pm 260 \pm 1101$
20-25	$3377 \pm 86$	$2877 \pm 729$	$4817 \pm 1329$	$5393 \pm 578$	$4458 \pm 1291$	$2877 \pm 1939 \pm 710$
25-30	$2068 \pm 67$	$2856 \pm 408$	$3443 \pm 1514$	$3137 \pm 356$	$3904 \pm 975$	$2856 \pm 586 \pm 394$
30-35	$1403 \pm 55$	$2160 \pm 306$	$1686 \pm 693$	$1808 \pm 401$	$1482 \pm 545$	$2160 \pm 474 \pm 295$
35-45	$1489 \pm 57$	$2465 \pm 339$	$2937 \pm 1009$	$2385 \pm 279$	$2185 \pm 935$	$2465 \pm 471 \pm 329$
45-55	$818 \pm 42$	$1246 \pm 243$	$1184 \pm 595$	$894 \pm 176$	$1139 \pm 226$	$1246 \pm 61 \pm 237$
55-65	$550 \pm 34$	$657 \pm 208$	$957 \pm 375$	$572 \pm 219$	$320 \pm 260$	$657 \pm 299 \pm 204$
65-75	$280 \pm 24$	$304 \pm 169$	$599 \pm 206$	$471 \pm 80$	$452 \pm 141$	$304 \pm 295 \pm 166$
75-85	$186 \pm 20$	$375 \pm 162$	$374 \pm 91$	$244 \pm 69$	$195 \pm 66$	$375 \pm 1 \pm 161$
85-95	$139 \pm 17$	$242 \pm 60$	$250 \pm 28$	$318 \pm 19$	$199 \pm 30$	$242 \pm 8 \pm 57$
95-120	$208 \pm 21$	$390 \pm 55$	$369 \pm 195$	$348 \pm 41$	$168 \pm 54$	$390 \pm 21 \pm 51$
120-500	$157 \pm 18$	$255 \pm 88$	$294 \pm 17$	$257 \pm 53$	$181 \pm 30$	$255 \pm 38 \pm 85$

Another source of the systematic uncertainty related to jets $\rightarrow\gamma$  background estimation  $\Delta\sigma^{Z\gamma-Norm}$  originates from the uncertainty on the  $Z\gamma$  MC normalization. The  $Z\gamma$  MC sample is used to prepare fake- $\gamma$  template, and the normalization of this sample significantly affects the template shape.

Table 5.11: Results of the background subtraction based on fits of different variables.  $W\gamma$ , electron channel. First column:  $P_T^\gamma$  ranges, second column: signal MC prediction, third and fourth columns: signal yields extracted from fits of  $I_{ch}^\gamma$  and  $\sigma_{inj}^\gamma$  distributions of data, fifth and sixth columns: signal yields extracted from MC closure tests, seventh column: signal yield that is used in the next measurement steps.

$P_T^\gamma$ , GeV	sig. MC ( $W\gamma \rightarrow e\nu\gamma$ )	data		MC closure		yield data-bkg.
		$I_{ch}^\gamma$	$\sigma_{inj}^\gamma$	$I_{ch}^\gamma$	$\sigma_{inj}^\gamma$	
barrel photons						
10-15	$12480 \pm 163$	$10994 \pm 1331$	$12425 \pm 2000$	$10640 \pm 1500$	$14995 \pm 2225$	$10994 \pm 1430 \pm 1277$
15-20	$5857 \pm 110$	$5160 \pm 668$	$7421 \pm 1173$	$4124 \pm 602$	$5721 \pm 1927$	$5160 \pm 2261 \pm 613$
20-25	$2868 \pm 77$	$3022 \pm 384$	$3168 \pm 2937$	$3390 \pm 258$	$3699 \pm 1261$	$3022 \pm 145 \pm 338$
25-30	$1411 \pm 54$	$1846 \pm 293$	$2250 \pm 1984$	$1365 \pm 152$	$1339 \pm 1167$	$1846 \pm 404 \pm 273$
30-35	$915 \pm 43$	$1283 \pm 193$	$1831 \pm 971$	$877 \pm 111$	$891 \pm 278$	$1283 \pm 547 \pm 180$
35-45	$1247 \pm 51$	$1732 \pm 190$	$1965 \pm 882$	$1359 \pm 111$	$1330 \pm 277$	$1732 \pm 232 \pm 178$
45-55	$820 \pm 41$	$673 \pm 207$	$1199 \pm 485$	$698 \pm 118$	$933 \pm 65$	$673 \pm 526 \pm 196$
55-65	$654 \pm 37$	$956 \pm 302$	$1010 \pm 157$	$566 \pm 95$	$666 \pm 152$	$956 \pm 53 \pm 296$
65-75	$440 \pm 30$	$625 \pm 252$	$756 \pm 47$	$357 \pm 99$	$458 \pm 123$	$625 \pm 131 \pm 248$
75-85	$295 \pm 25$	$367 \pm 137$	$516 \pm 134$	$339 \pm 45$	$285 \pm 84$	$367 \pm 148 \pm 132$
85-95	$234 \pm 22$	$364 \pm 29$	$366 \pm 33$	$315 \pm 63$	$283 \pm 83$	$364 \pm 1 \pm 2$
95-120	$318 \pm 26$	$430 \pm 88$	$555 \pm 66$	$397 \pm 77$	$400 \pm 135$	$430 \pm 124 \pm 78$
120-500	$429 \pm 30$	$743 \pm 234$	$734 \pm 40$	$568 \pm 54$	$537 \pm 236$	$743 \pm 9 \pm 231$
endcap photons						
10-15	$4368 \pm 96$	$-1785 \pm 122$	$4129 \pm 1180$	$2286 \pm 1356$	$-1502 \pm 1196$	$-1785 \pm 5915 \pm 108$
15-20	$2253 \pm 68$	$-241 \pm 537$	$1869 \pm 762$	$1541 \pm 483$	$352 \pm 759$	$-241 \pm 2110 \pm 506$
20-25	$1177 \pm 49$	$637 \pm 298$	$1679 \pm 534$	$1308 \pm 192$	$1414 \pm 481$	$637 \pm 1042 \pm 277$
25-30	$574 \pm 34$	$887 \pm 147$	$1078 \pm 646$	$674 \pm 117$	$1125 \pm 370$	$887 \pm 190 \pm 131$
30-35	$445 \pm 31$	$731 \pm 107$	$555 \pm 249$	$451 \pm 119$	$355 \pm 155$	$731 \pm 176 \pm 96$
35-45	$638 \pm 37$	$943 \pm 116$	$1071 \pm 326$	$773 \pm 76$	$789 \pm 189$	$943 \pm 127 \pm 104$
45-55	$287 \pm 24$	$478 \pm 106$	$449 \pm 449$	$307 \pm 67$	$347 \pm 78$	$478 \pm 28 \pm 95$
55-65	$237 \pm 22$	$287 \pm 155$	$433 \pm 44$	$225 \pm 51$	$220 \pm 114$	$287 \pm 145 \pm 150$
65-75	$194 \pm 21$	$255 \pm 73$	$372 \pm 38$	$154 \pm 45$	$37 \pm 87$	$255 \pm 116 \pm 67$
75-85	$137 \pm 18$	$210 \pm 47$	$236 \pm 28$	$201 \pm 59$	$155 \pm 73$	$210 \pm 25 \pm 40$
85-95	$81 \pm 14$	$118 \pm 47$	$128 \pm 30$	$146 \pm 39$	$44 \pm 40$	$118 \pm 10 \pm 40$
95-120	$166 \pm 20$	$233 \pm 51$	$211 \pm 21$	$224 \pm 21$	$192 \pm 49$	$233 \pm 21 \pm 46$
120-500	$145 \pm 18$	$276 \pm 21$	$254 \pm 24$	$227 \pm 31$	$194 \pm 46$	$276 \pm 22 \pm 3$

The uncertainty on the  $Z\gamma$  normalization is set to be 4.6% as reported by CMS  $Z\gamma$  measurement at  $\sqrt{s} = 8$  TeV [46]. To estimate  $\Delta\sigma^{Z\gamma-Norm}$ , we prepare templates with  $Z\gamma$  normalizations deviated by  $\pm 4.6\%$  from the nominal value. After that, we perform fits with such deviated templates, and compare results among the fits with templates of nominal normalization and with two deviated ones. The spread among three results is a systematic uncertainty on the signal yields.

### 5.7.2 Other Sources of the Systematic Uncertainties

The next significant uncertainty after the uncertainty due to jets $\rightarrow\gamma$  background estimation in the electron channel is the uncertainty due to  $e\rightarrow\gamma$  background estimation. This uncertainty is comprised of a fit bias and a statistical power of the MC samples used for this background estimation.

The uncertainty due to fit bias is evaluated by fitting Z-peak before and after  $M_T^W$  requirement is applied. Other selection requirements are applied except the Z-mass window requirement. The plots with the fit results of the datasets before and after  $M_W^T$  requirement applied are shown in App. G and F respectively.

Another source of uncertainty originates from the limited statistics of all MC samples involved in the  $e\rightarrow\gamma$  background estimation. Values of  $e\rightarrow\gamma$  uncertainties from both sources are summarized in Tab. 5.16.

For the real- $\gamma$  background subtraction, the statistical uncertainties of  $Z\gamma$  and  $W\gamma\rightarrow\tau\nu\gamma$  samples and their normalization uncertainties are taken into account. The normalization uncertainty applied for the  $Z\gamma$  sample is 4.6% and for the  $W\gamma\rightarrow\tau\nu\gamma$  is 20%.

All uncertainties related to the background subtraction procedure are estimated as uncertainties on “data-bkg. yields”. To estimate uncerntainties of the effects listed above on the cross section, the uncertainties on yields are propagated through the unfolding,  $A\times\epsilon$  correction, divided over the luminosity and the bin width (for the differential cross section).

Limited statistical power of the signal MC sample contributes to the systematic uncertainty of the differential cross section through the unfolding procedure. To evaluate this uncertainty, first, we randomize the migration matrix 100 times by Gaussian distribution as  $M_{ji}\rightarrow Gaus(M_{ji}, \sigma_{ji})$  where  $\sigma_{ji}$  is the signal MC statistical

uncertainties in particular  $[j, i]$  bin. After that, the procedure of unfolding is repeated for each migration matrix. The standard deviation out of all unfolding outputs is taken as an uncertainty in each  $P_T^\gamma$  bin, and, finally, the uncertainty is propagated through the  $A \times \epsilon$  correction and is divided over the luminosity and the bin width to estimate the uncertainty on the cross section.

For the uncertainty due to the  $A \times \epsilon$  correction, the signal MC statistical power is also taken into account. This uncertainty is estimated as

$$\Delta N_{true}^i = N_{acc}^i \cdot \frac{\Delta(A \times \epsilon)^i}{((A \times \epsilon)^i)^2}, \quad (5.13)$$

where  $N_{acc}^i$  are unfolded yields, before  $A \times \epsilon$  correction. To estimate uncertainty on the cross section,  $\Delta N_{true}^i$  is divided by the luminosity and by the bin width.

Another source of the systematic uncertainty originates from biases in  $E_T^{miss}$  modeling in the MC. To estimate this uncertainty, we, first of all, vary the  $P_T$  of the photons, electrons (for the electron channel) and jets in the event by their uncertainties as prescribed by CMS JetMET POG. Then sum up all the listed contributions as the Lorentz vectors and propagate their variations to the variation of  $E_T^{miss}$ . After that, we recalculate  $M_T^W$  and recompute  $A \times \epsilon$  and unfolding constants with the new  $M_T^W$  values. Then recompute the cross section values and take the spread in the cross section among the lower, upper and nominal  $M_T^W$  values as a systematic uncertainty.

The contribution from the uncertainties of the efficiency scale factors are also estimated. The scale factors are varied by  $\pm 1\sigma$ , then the  $A \times \epsilon$  and unfolding constants are recomputed and new values of the cross section are found. The spread in the cross section among the  $+1\sigma$ ,  $-1\sigma$  and the nominal scale factor values is taken as a systematic uncertainty.

The systematic uncertainty due to pileup reweighting is estimated by varying the pileup cross section by  $\pm 5\%$ . The luminosity uncertainty is 2.6%.

### 5.7.3 Summary of the Systematic Uncertainties

The relative systematic uncertainties are summarized in Tab. 5.12 and Tab. 5.13 for the muon and electron channels respectively. The systematic uncertainties related to pre-unfolding measurement steps have to be propagated through unfolding. For each of such uncertainties, a correlation matrix appears. All these correlation matrices are plotted in App. J.

Table 5.12: Relative uncertainties [%].  $W\gamma$ , muon channel. The details of the “syst other” column are provided in Tab. 5.14.

$P_T^\gamma$ , GeV	err stat	syst $ N_{I\gamma h} - N_{\sigma i \eta j \eta} $	$Z\gamma$ MC norm	templ stat	SFs err	syst lumi	syst other	syst total
total	1	10	24	4	2	3	4	27
15-20	2	31	12	10	3	3	6	35
20-25	2	29	13	11	1	3	6	34
25-30	2	24	13	11	1	3	5	30
30-35	3	40	15	13	2	3	7	45
35-45	2	11	12	8	2	3	6	19
45-55	4	62	19	20	2	3	8	68
55-65	3	15	12	14	1	3	7	24
65-75	6	36	19	17	1	3	10	44
75-85	4	6	11	16	1	3	10	21
85-95	5	2	9	23	1	3	13	25
95-120	5	10	8	12	1	3	9	18
120-500	3	4	11	21	2	3	9	24

Table 5.13: Relative uncertainties [%].  $W\gamma$ , electron channel. The details of the “syst other” and “ $e \rightarrow \gamma$ ” column are provided in Tab. 5.15 and 5.16 respectively.

$P_T^\gamma$ , GeV	err stat	syst $ N_{Ich} - N_{orig\eta} $	$Z\gamma$ MC norm	templ stat	SFs err	syst lumi	$e \rightarrow \gamma$	syst other	syst total
total	2	15	35	5	19	3	4	5	44
15-20	8	80	27	19	17	3	18	11	90
20-25	7	38	20	14	12	3	11	10	48
25-30	5	25	16	12	14	3	8	8	36
30-35	5	35	14	12	14	3	3	8	42
35-45	3	14	13	8	18	3	2	7	28
45-55	8	53	20	22	36	3	7	11	71
55-65	7	17	12	30	44	3	5	10	58
65-75	7	23	15	32	44	3	4	11	61
75-85	8	32	17	27	44	3	6	13	64
85-95	9	9	7	9	40	3	8	14	44
95-120	7	19	9	14	44	3	5	11	51
120-500	4	12	6	24	39	3	1	9	48

Table 5.14: Relative systematic uncertainties [%] of smaller contributions (details of the column “syst other” from Tab. 5.12).  $W\gamma$ , muon channel. First column:  $P_T^\gamma$  range, second column: total systematic uncertainties of smaller contributions, third column: uncertainties related to real- $\gamma$  background estimation, fourth column: uncertainties due to limited signal MC statistical power in  $A \times \epsilon$  correction, fifth column: uncertainties due to  $M_T^W$  selection requirement, sixth column: uncertainties related to PU reweighting, seventh column: uncertainties due to limited signal MC statistical power in the unfolding correction.

$P_T^\gamma$ , GeV	syst other	real- $\gamma$ bkg	$A \times \epsilon$ MC stat	$M_T^W$ req.	PU weight	unf MC stat
total	4	1	0	1	4	1
15-20	6	2	1	1	4	2
20-25	6	3	2	2	4	3
25-30	5	3	2	2	2	2
30-35	7	4	3	1	4	3
35-45	6	3	3	2	3	2
45-55	8	3	3	1	4	5
55-65	7	2	4	2	4	3
65-75	10	2	6	3	5	6
75-85	10	1	7	3	3	5
85-95	13	2	8	4	6	7
95-120	9	2	7	2	2	6
120-500	9	1	6	1	4	4

Table 5.15: Relative systematic uncertainties [%] of smaller contributions.  $W\gamma$ , electron channel. First column:  $P_T^\gamma$  range, second column: total systematic uncertainties of smaller contributions, third column: uncertainties related to real- $\gamma$  background estimation, fourth column: uncertainties due to limited signal MC statistical power in  $A \times \epsilon$  correction, fifth column: uncertainties due to  $M_T^W$  selection requirement, sixth column: uncertainties related to PU reweighting, seventh column: uncertainties due to limited signal MC statistical power in the unfolding correction.

$P_T^\gamma$ , GeV	syst other	real- $\gamma$ bkg	$A \times \epsilon$ MC stat	$M_T^W$ req.	PU weight	unf MC stat
total	5	2	0	1	4	2
15-20	11	6	2	1	4	8
20-25	10	5	2	1	4	7
25-30	8	3	3	1	3	6
30-35	8	2	4	1	3	6
35-45	7	1	4	1	4	4
45-55	11	2	5	3	4	9
55-65	10	2	5	3	5	7
65-75	11	1	6	1	4	8
75-85	13	2	8	2	3	9
85-95	14	2	9	2	2	9
95-120	11	1	8	1	4	7
120-500	9	1	7	2	3	4

Table 5.16: Relative systematic uncertainties [%] of  $e \rightarrow \gamma$  background estimation.  $W\gamma$ , electron channel. First column:  $P_T^\gamma$  ranges, second column: total systematic uncertainty related to  $e \rightarrow \gamma$  background estimation, third column: systematic uncertainty evaluated by comparing fit results of samples with and without  $M_T^W$  selection requirement, fourth column: systematic uncertainty due to statistical power of MC samples involved in  $e \rightarrow \gamma$  background estimation.

$P_T^\gamma$ , GeV	$e \rightarrow \gamma$	$e \rightarrow \gamma$ y/n $M_T^W$ req.	$e \rightarrow \gamma$ samp. stat
total	4	4	1
15-20	18	17	4
20-25	11	10	4
25-30	8	7	3
30-35	3	1	2
35-45	2	1	1
45-55	7	4	5
55-65	5	3	4
65-75	4	1	4
75-85	6	4	4
85-95	8	5	6
95-120	5	3	4
120-500	1	0	1

## 5.8 Cross Section

The cross section of  $pp \rightarrow W\gamma \rightarrow l\nu\gamma$ , where  $l = \mu, e$ , is measured as described in the beginning of Ch. 5.1 based on CMS data at  $\sqrt{s} = 8$  TeV in the phase space defined in Ch. 5. The measured cross section is compared to phase-space-corrected MCFM calculation at NLO which is referred as “NLO theory”.

The cross section of the simulated sample was computed with MCFM at NLO and is for the dedicated signal MC sample  $\sigma_1 = 553.92$  pb. NNLO and higher corrections are expected to have an effect of 20%. The MC sample was generated with MadGraph, and the cross section in our selected phase space was computed as

$$\sigma_2 = \sigma_1 \cdot \frac{N_2}{N_1},$$

where  $N_2$  and  $N_1$  are numbers of events falling into selected phase space and generated in the whole MC sample respectively. For the differential cross section,  $N_2$  is number of events falling into specific  $P_T^\gamma$  bin and to compute  $\frac{d\sigma}{dP_T^\gamma}$ , we divide by the bin width.

The total cross section is

$$\sigma(W\gamma \rightarrow \mu\nu\gamma) = 10949 \pm 91 \text{ (stat.)} \pm 2959 \text{ (syst.) fb}$$

for the muon channel, and

$$\sigma(W\gamma \rightarrow \mu\nu\gamma) = 9146 \pm 185 \text{ (stat.)} \pm 3981 \text{ (syst.) fb}$$

for the electron channel. the corresponding NLO theory values are 9139 and 9064 fb.

Tables 5.17-5.18 and Fig. 5.10 summarize the results of the differential cross section. The measured cross sections in different channels agree between each other as well as with the NLO theory cross section provided its accuracy.

The measurement is systematically-dominated. For the muon channel, the most significant sources of the systematic uncertainty are sources associated with jets $\rightarrow\gamma$  background estimation. For the electron channel, uncertainties on photon efficiency scale factors are more significant for certain bins. The relative systematic uncertainties on the total cross section are larger than those reported by the CMS measurement at  $\sqrt{s}=7$  TeV [4].

For validation of the measurement procedure, we measure the cross section of  $Z\gamma$  and compare the result with the published CMS result for  $Z\gamma$  at 8 TeV. We measure cross section in the muon and electron channels, in the same phase space as the published CMS measurement.

$Z\gamma \rightarrow \mu\mu\gamma$  FSR and ISR datasets which are used to prepare real- $\gamma$  and fake- $\gamma$  templates for jets $\rightarrow\gamma$  background estimation largely overlap with nominally selected  $Z\gamma$  dataset. Therefore, the measurement of  $Z\gamma$  cross section in the muon channel is a closure check while the measurement of  $Z\gamma$  cross section in the electron channel is a fully valid physics measurement. The results of our  $Z\gamma$  measurement agree well with the published results (App. L).

The ongoing  $W\gamma$  measurement based on 2015 and 2016 datasets has higher chances to discover a new physics if it presents because of higher energy of  $\sqrt{s}=13$  TeV and higher statistical power of over  $30\text{ fb}^{-1}$ . Although the largest uncertainties of the 8 TeV measurement are systematic uncertainties, many of them depend on amount of data in control samples, thus, the increased size of the data sample will help to reduce those uncertainties. Higher collision energy allows to observe more signal events in high  $P_T^\gamma$  ranges where the effect of potential aTGC is the largest.

Table 5.17: Cross section and errors.  $W\gamma$  at  $\sqrt{s} = 8$  TeV, muon channel.

$P_T^\gamma$ , GeV		$ \mathrm{d}\sigma/dP_T $ , [fb/GeV]
	theory	meas.
15-20	754	$747 \pm 17 \pm 260$
20-25	382	$419 \pm 10 \pm 142$
25-30	210	$290 \pm 6 \pm 87$
30-35	129	$175 \pm 5 \pm 78$
35-45	70	$121 \pm 2 \pm 22$
45-55	36	$34 \pm 1 \pm 23$
55-65	23	$31 \pm 1 \pm 7$
65-75	13	$16 \pm 1 \pm 7$
75-85	9	$16 \pm 1 \pm 3$
85-95	6.3	$9.8 \pm 0.5 \pm 2.4$
95-120	3.7	$5.9 \pm 0.3 \pm 1.0$
120-500	0.27	$0.41 \pm 0.01 \pm 0.10$

Table 5.18: Cross section and errors.  $W\gamma$  at  $\sqrt{s} = 8$  TeV, electron channel.

$P_T^\gamma$ , GeV		$ \mathrm{d}\sigma/dP_T $ , [fb/GeV]
	theory	meas.
15-20	749	$440 \pm 35 \pm 396$
20-25	375	$338 \pm 23 \pm 163$
25-30	210	$298 \pm 16 \pm 107$
30-35	129	$193 \pm 9 \pm 82$
35-45	71	$103 \pm 3 \pm 29$
45-55	34	$34 \pm 3 \pm 24$
55-65	22	$31 \pm 2 \pm 18$
65-75	14	$19 \pm 1 \pm 12$
75-85	9	$12 \pm 1 \pm 8$
85-95	6.4	$10.0 \pm 0.9 \pm 4.3$
95-120	3.6	$4.9 \pm 0.4 \pm 2.5$
120-500	0.26	$0.46 \pm 0.02 \pm 0.22$

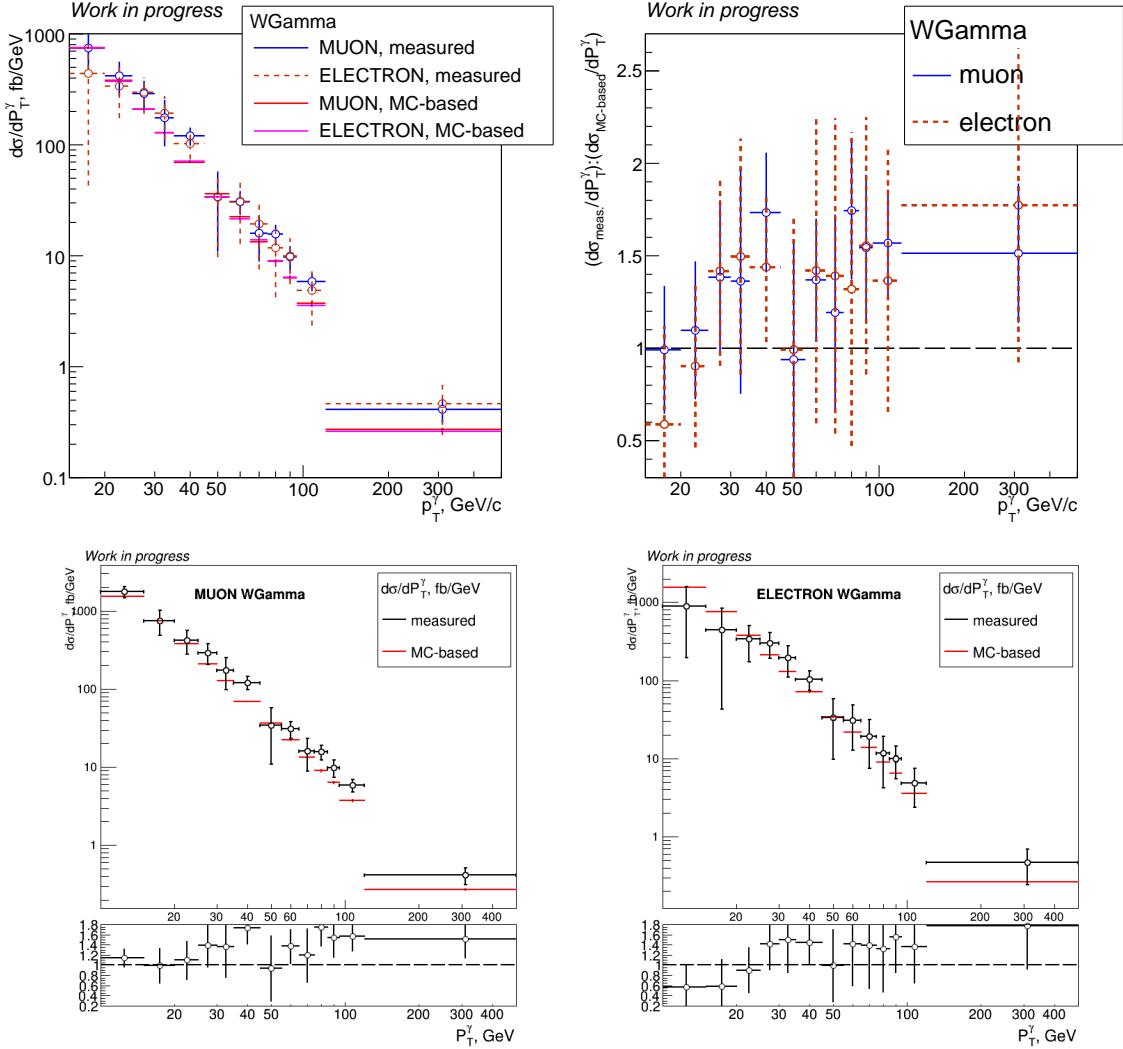


Figure 5.10:  $W\gamma$  differential cross section. Top, left: the  $W\gamma$  differential cross section; top, right: the ratio of measured over the MC-based  $W\gamma$  differential cross section. Bottom: the  $W\gamma$  measured differential cross section overlaid with the MC-based cross section in the muon channel (left) and in the electron channel (right).

# Chapter 6

## Summary and Conclusions

This dissertation reports a measurement of the total and differential in  $P_T^\gamma$  bins cross sections of the  $W\gamma \rightarrow \mu\nu\gamma$  and  $W\gamma \rightarrow e\nu\gamma$  processes using full 2012 dataset of  $L = 19.6 \text{ fb}^{-1}$  collected by CMS at  $\sqrt{s} = 8 \text{ TeV}$ . This is the first measurement of the differential  $W\gamma$  cross section with CMS. The results are in agreement between two channels and also agree with the MC-based cross section.

In addition to  $W\gamma$  cross section, we also measure  $Z\gamma$  cross section and compare results with the published  $Z\gamma$  CMS measurement at  $\sqrt{s} = 8 \text{ TeV}$ . The good agreement between our and published results on  $Z\gamma$  cross section validates parts of our  $W\gamma$  measurement that are the same between  $Z\gamma$  and  $W\gamma$  measurements including lepton and photon selection, jets $\rightarrow\gamma$  background estimation, detector resolution unfolding, acceptance and efficiency corrections.

Measurements of  $W\gamma$  at higher energies and luminosities and more likely to bring up a discovery if there is a new physics will have more opportunities to discover a new physics if it is there. Diboson and triboson studies remain a significant part of CMS physics program for studies at  $\sqrt{s} = 13 \text{ TeV}$ .

# Appendix A

## Efficiency Scale Factors

This appendix summarizes efficiency SF that were applied on MC events to make selection efficiency in MC match the selection efficiency in data. Tables A.2-A.3 contain muon ID ( $\rho_{ID}^\mu$ ) and muon isolation SF ( $\rho_{iso}^\mu$ ). A full SF on a muon object is a multiplication of ID and isolation SF. Electron ID SF are listed in Tab. A.4 ( $\rho_{ID}^e$ ). Photon ID ( $\rho_{ID}^\gamma$ ) and PSV ( $\rho_{PSV}^\gamma$ ) SF are summarized in Tab. A.5 and Tab. A.6.

For each  $W\gamma$  candidate in MC we apply a lepton and a photon SF. PSV SF is used in the electron channel only. For each  $Z\gamma$  candidate in MC we apply two lepton and one photon SF. Full SF for each type of candidate are summarized in Tab. A.1.

Table A.1: Scale Factors Applied

candidates	SF
$W\gamma \rightarrow \mu\nu\gamma$	$\rho_{ID}^\mu \times \rho_{iso}^\mu \times \rho_{ID}^\gamma$
$W\gamma \rightarrow e\nu\gamma$	$\rho_{ID}^e \times \rho_{ID}^\gamma \times \rho_{PSV}^\gamma$
$Z\gamma \rightarrow \mu\mu\gamma$	$\rho_{ID}^{\mu 1} \times \rho_{iso}^{\mu 1} \times \rho_{ID}^{\mu 2} \times \rho_{iso}^{\mu 2} \times \rho_{ID}^\gamma$
$Z\gamma \rightarrow ee\gamma$	$\rho_{ID}^1 \times \rho_{ID}^2 \times \rho_{ID}^\gamma$

Table A.2: Muon “Tight” ID Scale Factors as Recommended by POG

$P_T^\gamma$	$ \eta  < 0.9$	$0.9 <  \eta  < 1.2$	$1.2 <  \eta  < 2.1$
25-30	$0.992 \pm 0.001$	$0.995 \pm 0.001$	$0.998 \pm 0.001$
30-35	$0.993 \pm 0.001$	$0.993 \pm 0.001$	$0.997 \pm 0.001$
35-40	$0.994 \pm 0.000$	$0.992 \pm 0.001$	$0.997 \pm 0.001$
40-50	$0.992 \pm 0.000$	$0.992 \pm 0.000$	$0.997 \pm 0.000$
50-60	$0.992 \pm 0.001$	$0.995 \pm 0.001$	$0.995 \pm 0.001$
60-90	$0.989 \pm 0.001$	$0.990 \pm 0.002$	$0.992 \pm 0.002$
90-140	$1.004 \pm 0.003$	$1.009 \pm 0.006$	$1.023 \pm 0.005$
$>140$	$1.004 \pm 0.017$	$1.009 \pm 0.035$	$1.023 \pm 0.030$

Table A.3: Muon Isolation Scale Factors as Recommended by POG

$P_T^\gamma$	$ \eta  < 0.9$	$0.9 <  \eta  < 1.2$	$1.2 <  \eta  < 2.1$
25-30	$0.999 \pm 0.001$	$1.002 \pm 0.001$	$1.002 \pm 0.001$
30-35	$0.999 \pm 0.000$	$1.002 \pm 0.001$	$1.003 \pm 0.000$
35-40	$0.999 \pm 0.000$	$1.001 \pm 0.001$	$1.002 \pm 0.000$
40-45	$0.998 \pm 0.000$	$1.000 \pm 0.000$	$1.000 \pm 0.000$
45-50	$1.000 \pm 0.000$	$1.000 \pm 0.000$	$1.000 \pm 0.000$
50-60	$0.999 \pm 0.000$	$1.000 \pm 0.000$	$1.000 \pm 0.000$
60-90	$1.000 \pm 0.000$	$1.001 \pm 0.001$	$1.000 \pm 0.000$
90-140	$1.001 \pm 0.001$	$1.001 \pm 0.001$	$1.000 \pm 0.001$
$>140$	$1.001 \pm 0.002$	$1.004 \pm 0.005$	$0.997 \pm 0.002$

Table A.4: Electron “Tight” ID Scale Factors as Recommended by POG

$P_T^e$	$ \eta  \leq 0.80$	$0.80 <  \eta  \leq 1.44$	$1.57 <  \eta  \leq 2.00$	$ \eta  > 2.00$
$\leq 40$	$0.978 \pm 0.001$	$0.958 \pm 0.002$	$0.909 \pm 0.003$	$0.987 \pm 0.004$
40-50	$0.981 \pm 0.001$	$0.969 \pm 0.001$	$0.942 \pm 0.002$	$0.991 \pm 0.003$
$>50$	$0.982 \pm 0.002$	$0.969 \pm 0.002$	$0.957 \pm 0.004$	$0.999 \pm 0.005$

Table A.5: Photon “Medium” ID Scale Factors as Recommended by POG

$P_T^\gamma$	$ \eta  \leq 0.80$	$0.80 <  \eta  \leq 1.44$	$1.57 <  \eta  \leq 2.00$	$ \eta  > 2.00$
15-20	$0.95 \pm 0.02$	$0.99 \pm 0.02$	$1.00 \pm 0.02$	$1.02 \pm 0.02$
20-30	$0.96 \pm 0.01$	$0.97 \pm 0.01$	$0.98 \pm 0.01$	$1.00 \pm 0.01$
30-40	$0.98 \pm 0.01$	$0.98 \pm 0.01$	$0.99 \pm 0.01$	$1.00 \pm 0.01$
40-50	$0.98 \pm 0.01$	$0.98 \pm 0.01$	$1.00 \pm 0.01$	$1.01 \pm 0.01$
$>50$	$0.98 \pm 0.01$	$0.98 \pm 0.01$	$1.00 \pm 0.01$	$1.01 \pm 0.01$

Table A.6: Additional Photon Scale Factors for “PixelSeedVeto” as Derived in  $W\gamma\gamma$  Measurement

$P_T^\gamma$	barrel	endcap
15-20	$0.996 \pm 0.020$	$0.960 \pm 0.041$
20-25	$0.994 \pm 0.024$	$0.977 \pm 0.051$
25-30	$0.996 \pm 0.030$	$0.951 \pm 0.062$
30-40	$0.999 \pm 0.033$	$1.029 \pm 0.081$
40-50	$1.009 \pm 0.073$	$0.971 \pm 0.150$
50-70	$0.993 \pm 0.128$	$0.965 \pm 0.294$
$>70$	$1.047 \pm 0.111$	$1.145 \pm 0.371$

## Appendix B

### $Z\gamma$ FSR and ISR Plots

Data sample selected in  $Z\gamma \rightarrow \mu\mu\gamma$  conditions where photon selection is the same as for  $W\gamma$  selection is used to prepare real- $\gamma$  and fake- $\gamma$  templates for the jets $\rightarrow\gamma$  background estimation. FSR selection is used to prepare real- $\gamma$  templates:  $81 \text{ GeV} < M_{\mu\mu\gamma} < 101 \text{ GeV}$ ,  $\Delta R(\mu_{1,2}, \gamma) > 0.4$ ,  $\Delta R(\mu, \gamma)_{min} < 1.0$ ,  $M_{\mu\mu} < 80 \text{ GeV}$ . ISR selection is used to prepare fake- $\gamma$  templates:  $80 \text{ GeV} < M_{\mu\mu} < 100 \text{ GeV}$ ,  $\Delta R(\mu_{1,2}, \gamma) > 1.0$ .

The distributions of three-particle invariant mass  $M_{\mu\mu\gamma}$ , of the invariant mass of the dimuon system  $M_{\mu\mu}$  (Fig. B.1), and of the separation between the muon and the photon  $\Delta R(\mu, \gamma)$  (Fig. B.2) are used to split the dataset into the FSR sample which is dominated by real- $\gamma$  events and the ISR sample which is a mixture of real- $\gamma$  and fake- $\gamma$  events. The FSR sample is used to prepare real- $\gamma$  templates, and the fake- $\gamma$  contribution into the region is subtracted based on DY+jets MC predictions. The ISR sample is used to prepare fake- $\gamma$  templates, and the real- $\gamma$  contribution into the region is subtracted based on  $Z\gamma$  MC predictions. The number of real- $\gamma$  and fake- $\gamma$  events in different  $P_T^\gamma$  bins is shown in Fig. B.3. Various  $I_{ch}^\gamma$  and  $\sigma_{inj\eta}^\gamma$  templates are shown in Fig. B.4-B.9.

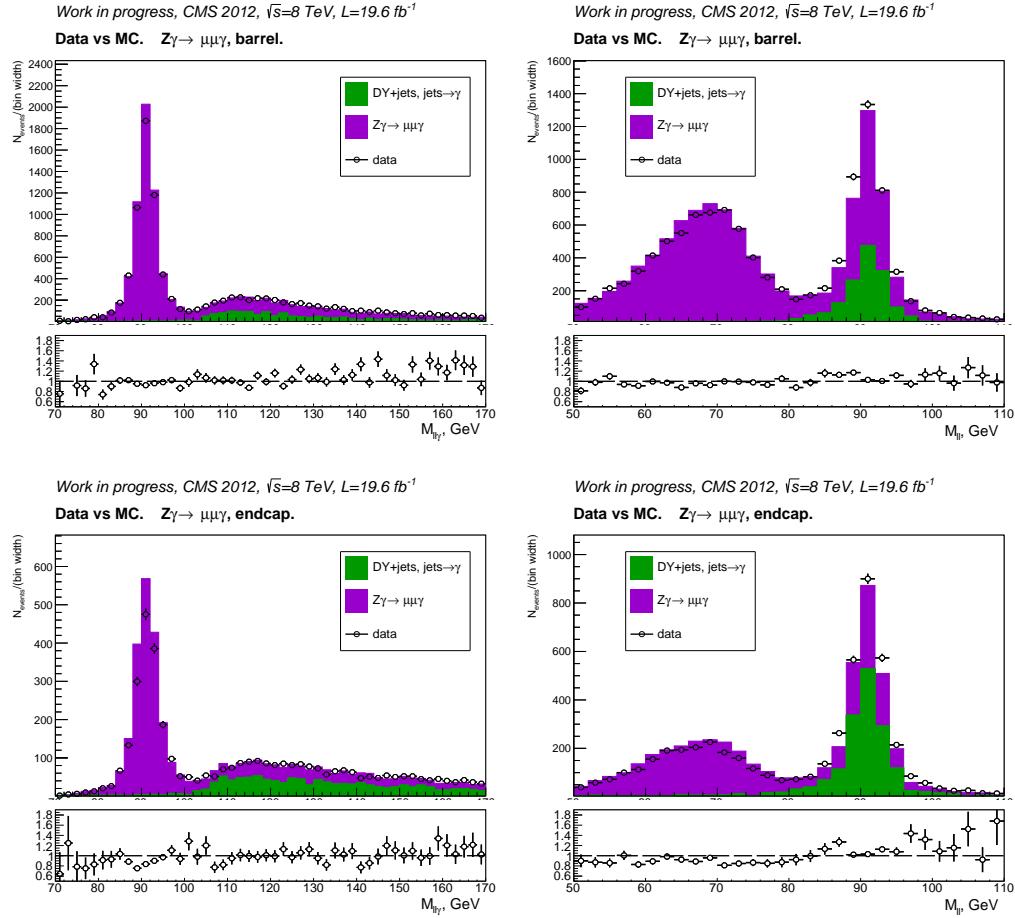


Figure B.1: Distributions of  $M_{\mu\mu\gamma}$  (left) and  $M_{\mu\mu}$  (right) in  $Z\gamma \rightarrow \mu\mu\gamma$ -selected events, data vs MC.  $P_T^\gamma : 15\text{-}500 \text{ GeV}$ . Left:  $M_{\mu\mu\gamma}$ , right:  $M_{\mu\mu}$ . Top: barrel photons, bottom: endcap photons. Peak highly dominated by  $Z\gamma$  events corresponds to FSR.

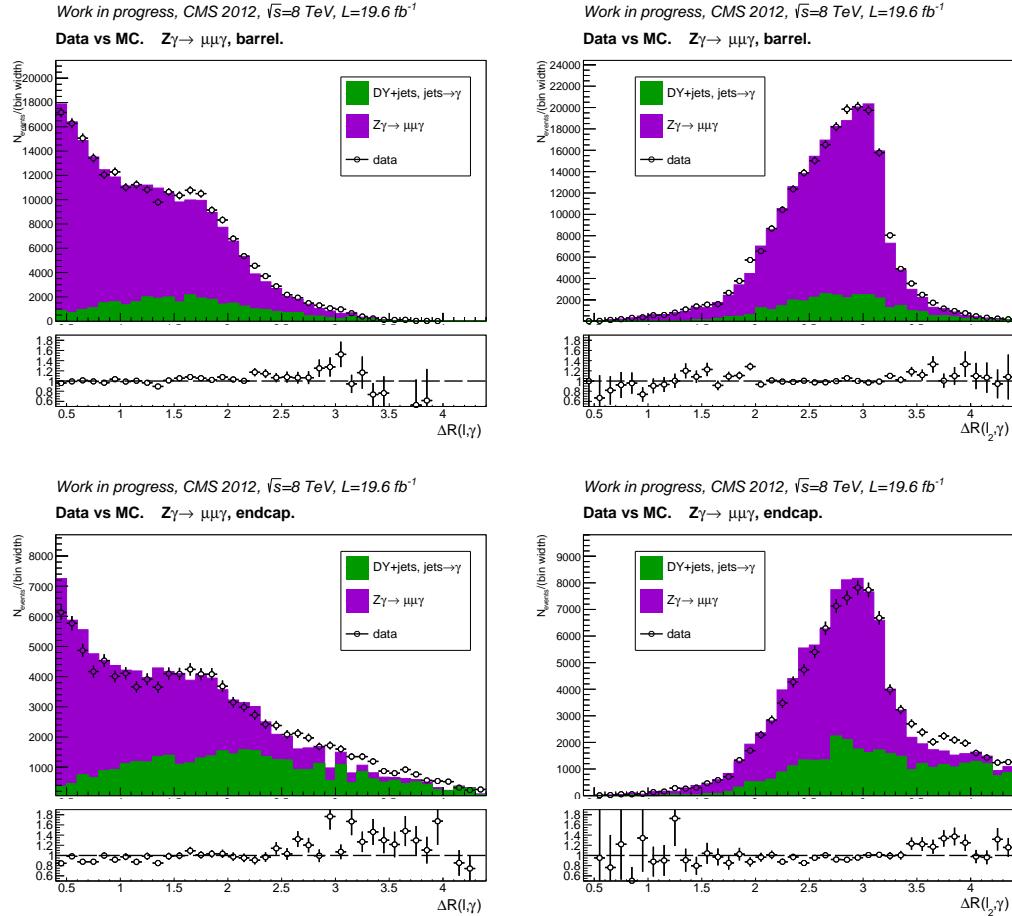


Figure B.2: Distributions of  $\Delta R(\mu_1, \gamma)$  (left) and  $\Delta R(\mu_2, \gamma)$  (right) in  $Z\gamma \rightarrow \mu\mu\gamma$ -selected events, data vs MC.  $P_T^\gamma : 15\text{-}500$  GeV. Left:  $M_{\mu\mu\gamma}$ , right:  $M_{\mu\mu}$ . Top: barrel photons, bottom: endcap. Peak highly dominated by  $Z\gamma$  events corresponds to FSR.

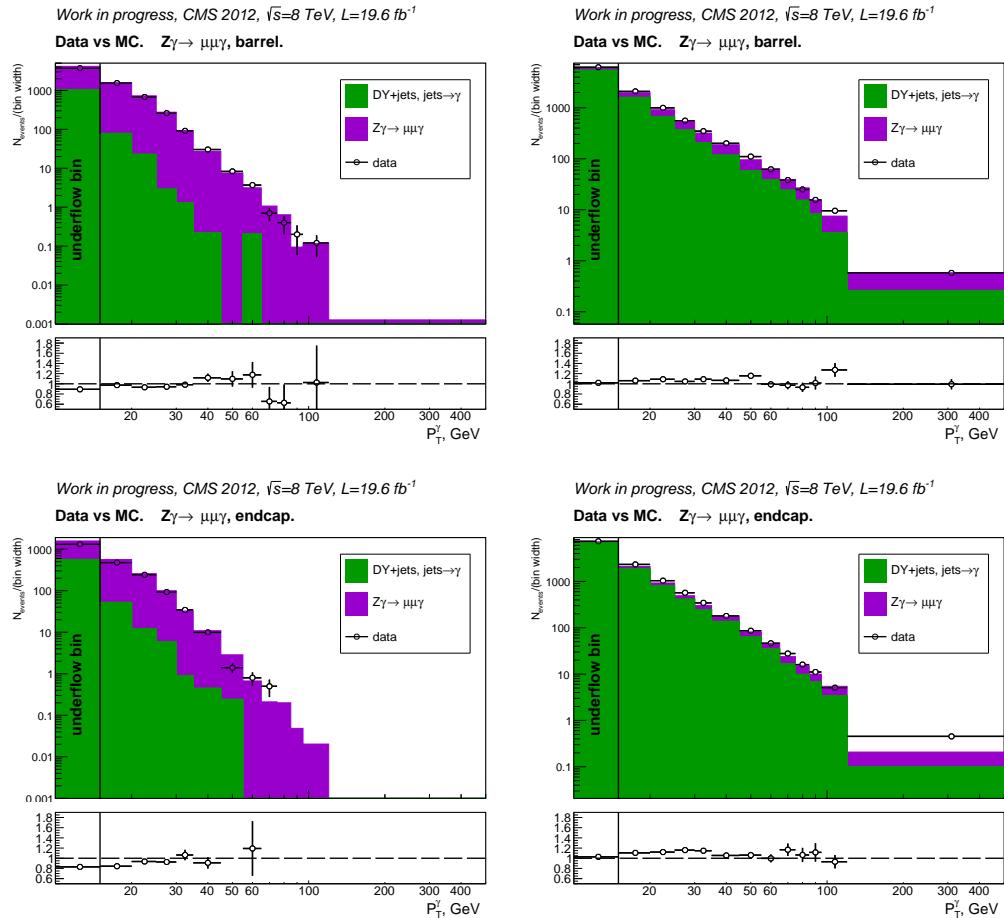


Figure B.3:  $Z\gamma$ -selected FSR (left) and ISR (right) events, data vs MC.

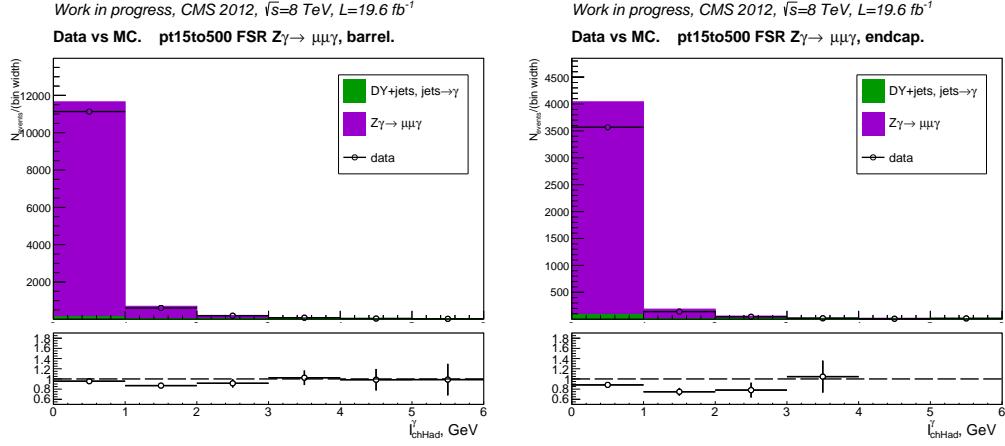


Figure B.4:  $Z\gamma$ -selected FSR events, data vs MC.  $P_T^\gamma > 15$  GeV. Distributions of  $I_{chHad}^\gamma$  used for preparing real- $\gamma$  templates. Fake- $\gamma$  contribution to FSR region is subtracted based on DY+jets MC prediction to prepare real- $\gamma$  templates.

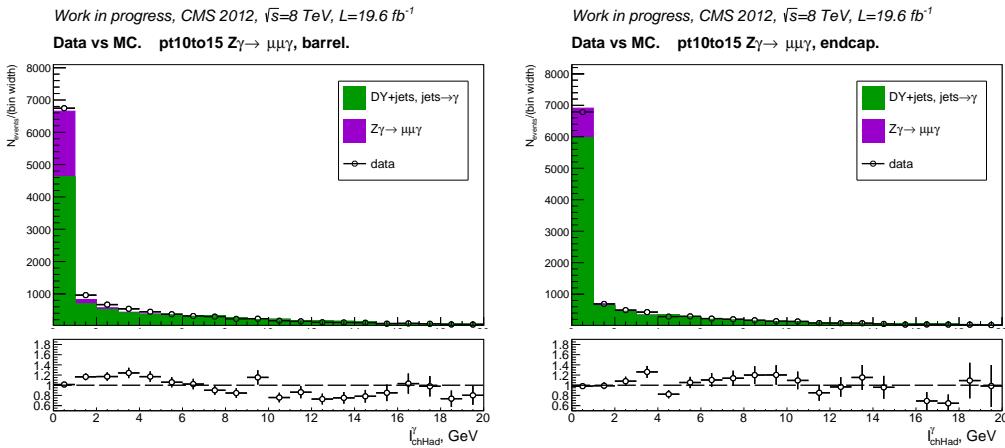


Figure B.5:  $Z\gamma$ -selected ISR events, data vs MC.  $10 \text{ GeV} < P_T^\gamma < 15$  GeV. Distributions of  $I_{chHad}^\gamma$  used for preparing fake- $\gamma$  templates. Real- $\gamma$  contribution to ISR region is subtracted based on  $Z\gamma$  signal MC prediction to prepare fake- $\gamma$  templates.

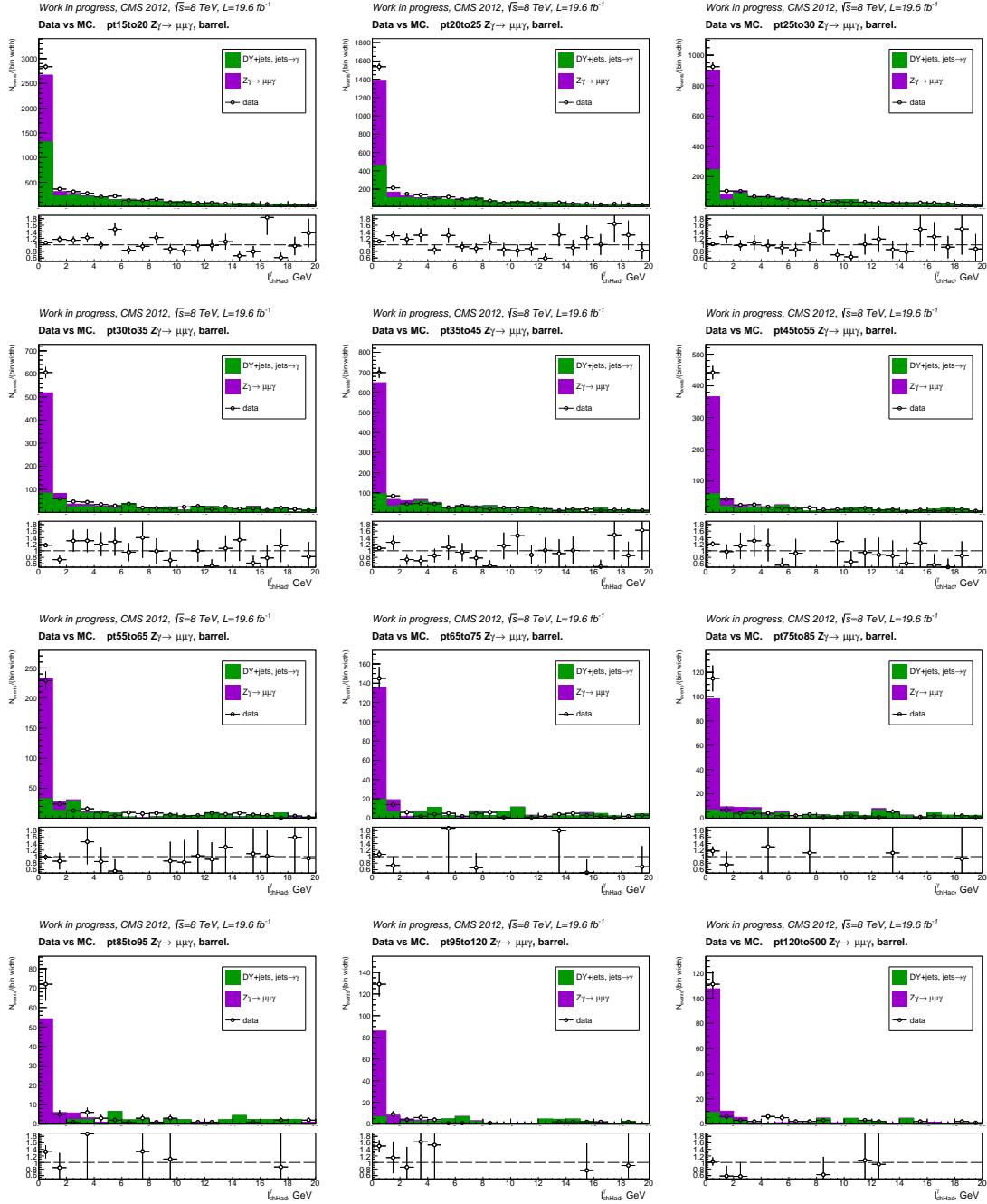


Figure B.6: Z $\gamma$ -selected ISR events, data vs MC. Distributions of  $I_{\text{chHad}}^\gamma$  used for preparing fake- $\gamma$  templates. Real- $\gamma$  contribution to ISR region is subtracted based on Z $\gamma$  signal MC prediction to prepare fake- $\gamma$  templates. Ranges of  $< P_T^\gamma$  are shown in the plot titles and cover the total range of  $15 \text{ GeV} < P_T^\gamma < 500 \text{ GeV}$ . EB photons.

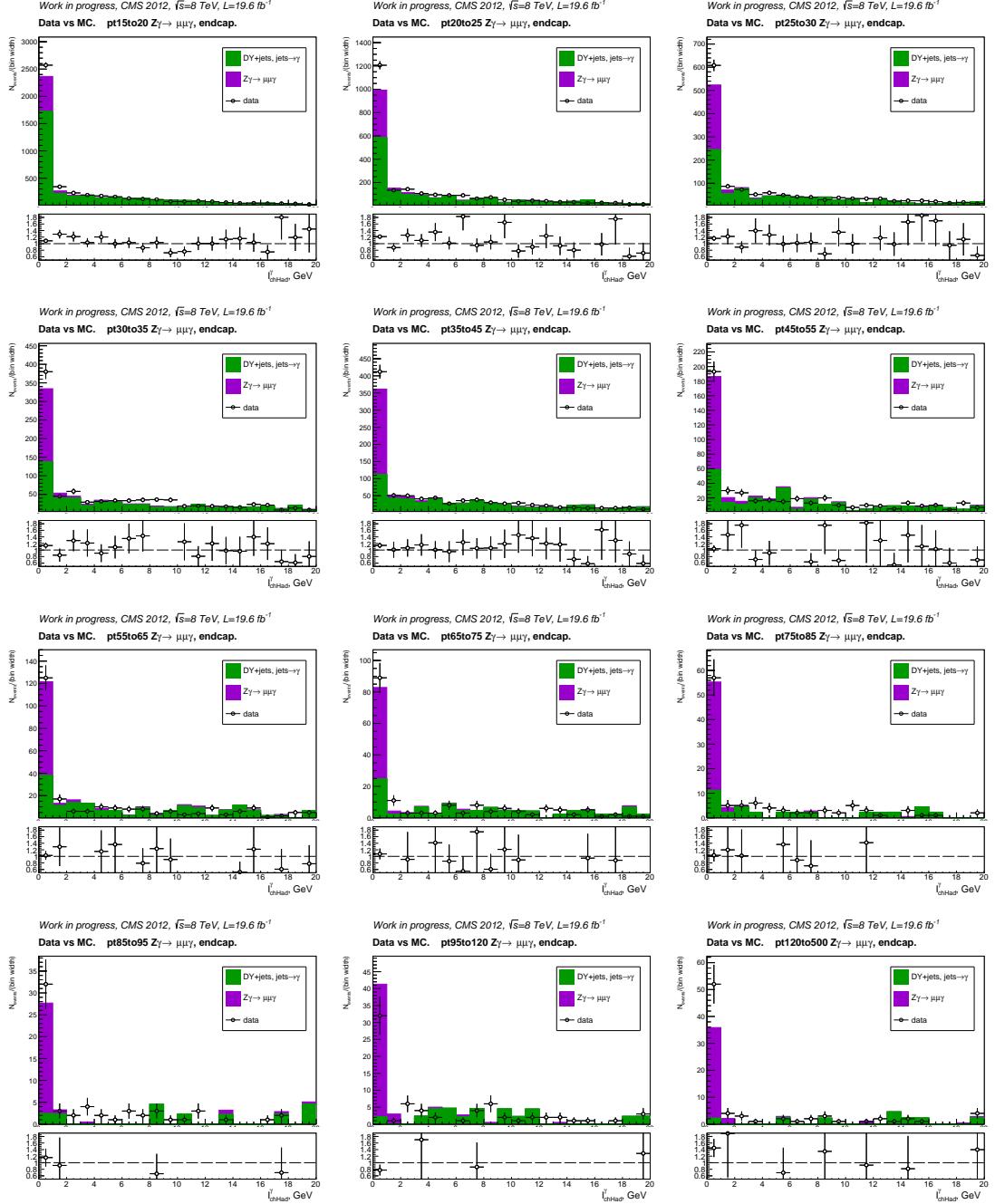


Figure B.7:  $Z\gamma$ -selected ISR events, data vs MC. Distributions of  $l_{\text{chHad}}^\gamma$  used for preparing fake- $\gamma$  templates. Real- $\gamma$  contribution to ISR region is subtracted based on  $Z\gamma$  signal MC prediction to prepare fake- $\gamma$  templates. Ranges of  $P_T^\gamma$  are shown in the plot titles and cover the total range of  $15 \text{ GeV} < P_T^\gamma < 500 \text{ GeV}$ . EE photons.

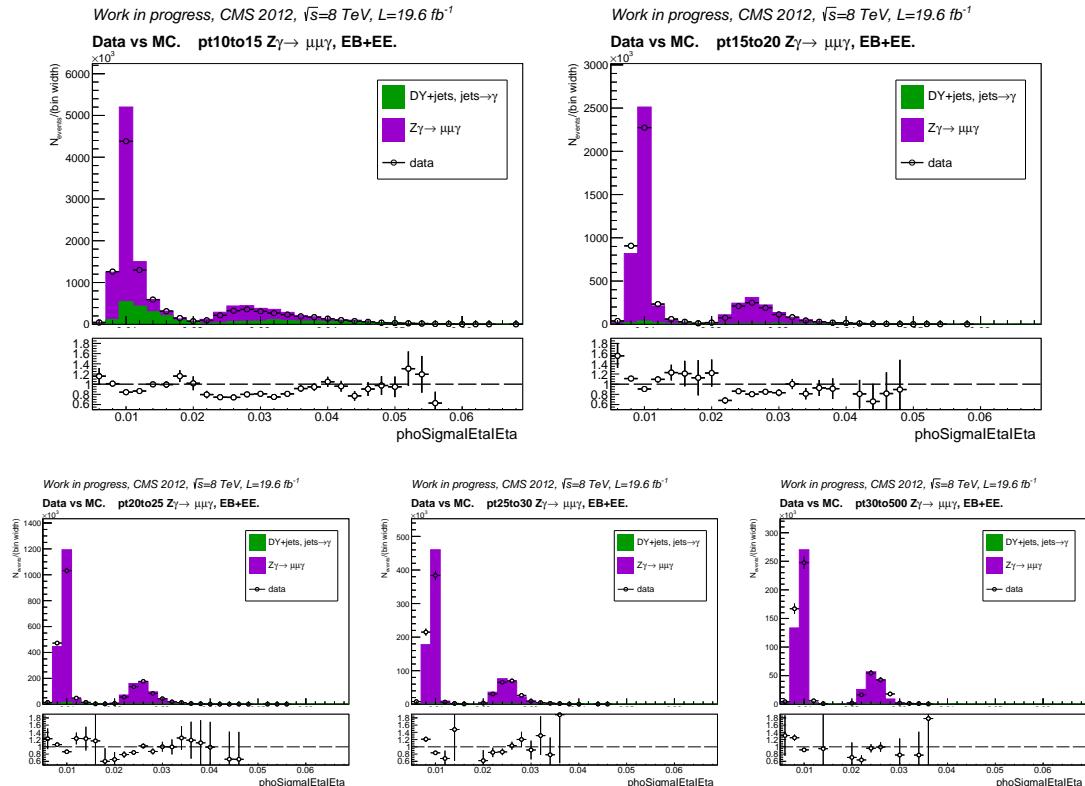


Figure B.8:  $Z\gamma$ -selected FSR events, data vs MC. Distributions of  $\sigma_{i\eta i\eta}$  are used for preparing real- $\gamma$  templates. Fake- $\gamma$  contribution to FSR region is subtracted based on DY+jets MC prediction to prepare real- $\gamma$  templates. The templates are prepared separately for barrel and endcap photons.

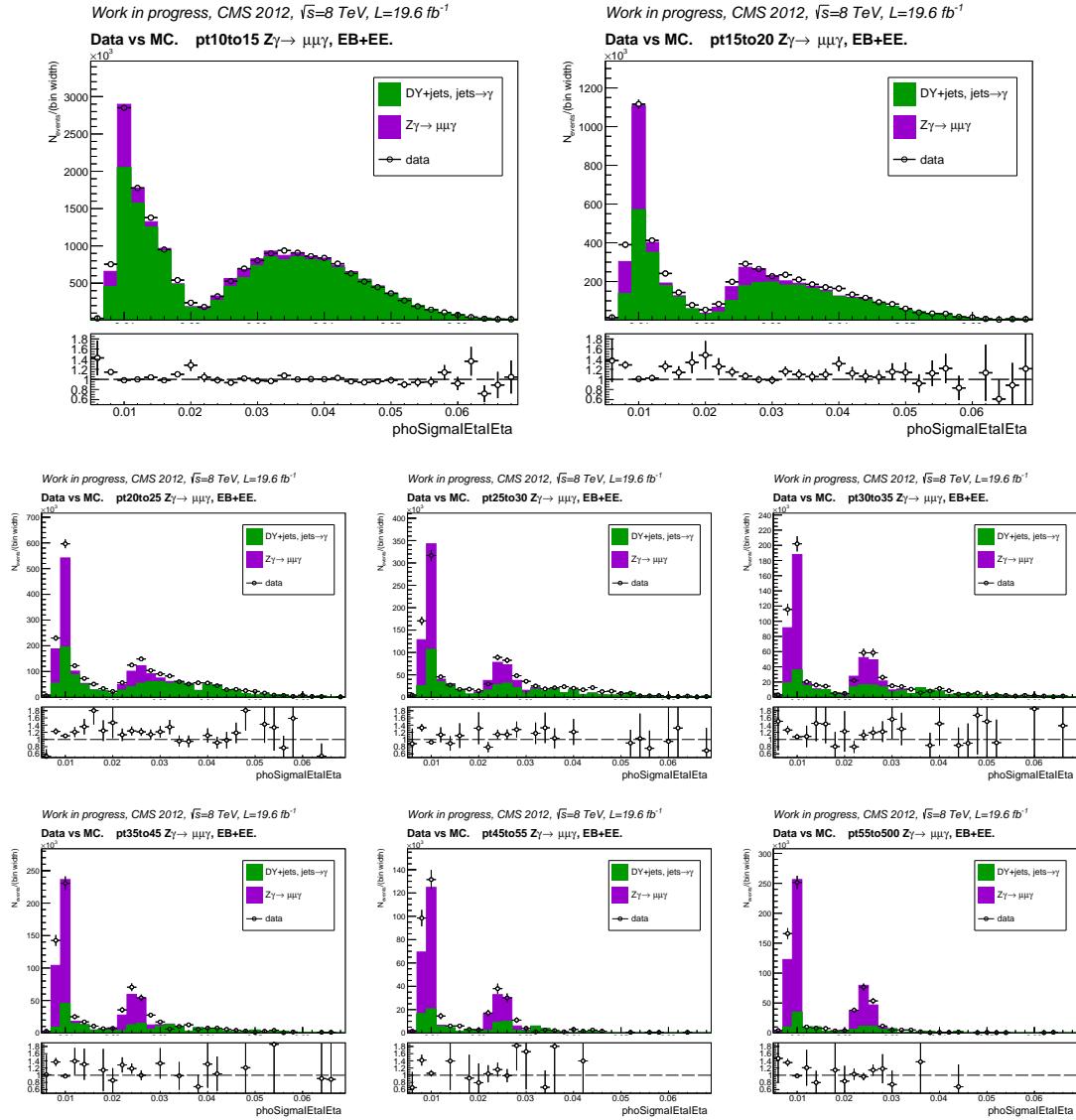


Figure B.9:  $Z\gamma$ -selected ISR events, data vs MC. Distributions of  $\sigma_{i\eta i\eta}$  are used for preparing real- $\gamma$  templates. Fake- $\gamma$  contribution to ISR region is subtracted based on DY+jets MC prediction to prepare real- $\gamma$  templates. The templates are prepared separately for barrel and endcap photons.

## Appendix C

### Studies of $E_T^{miss}$ Dependence of Templates

Figure C.1 shows the  $I_{ch}^\gamma$  and  $\sigma_{i\eta i\eta}^\gamma$  fake- $\gamma$  templates. Black histograms represent nominal  $Z\gamma$  ISR selection while red histograms have additional requirements of  $E_T^{miss} > 10$  GeV and  $\Delta\phi(E_T^{miss}, \gamma) < 0.5$ . These histograms do not show template dependence of  $E_T^{miss}$ . Fig. C.2-C.5 show  $I_{ch}^\gamma$  and  $\sigma_{i\eta i\eta}^\gamma$  distributions of  $W\gamma$  and  $W+\text{jets}$  MC with different  $M_T^W$  requirements.

The  $\sigma_{i\eta i\eta}^\gamma$  distributions of  $W+\text{jets}$  for different  $M_T^W$  requirements are different which indicates  $\sigma_{i\eta i\eta}^\gamma$  fake template dependence of  $M_T^W$ . We use  $Z\gamma$ -selected sample to prepare fake- $\gamma$  and real- $\gamma$   $\sigma_{i\eta i\eta}^\gamma$  templates, and we cannot apply the same  $M_T^W$  requirement on the sample as we apply in  $W\gamma$  selection. That causes a systematic uncertainty due to mismodeling of the  $\sigma_{i\eta i\eta}^\gamma$  shape. This uncertainty is covered of by comparison between fit results of  $\sigma_{i\eta i\eta}^\gamma$  and  $I_{ch}^\gamma$  templates.

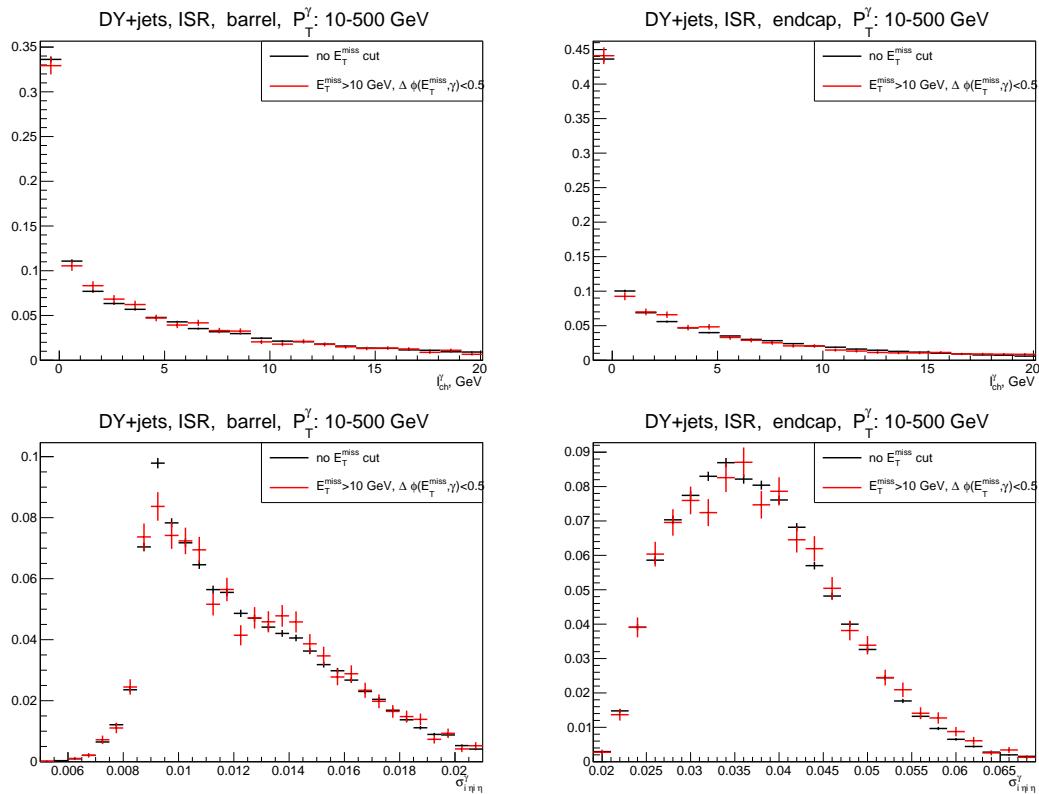


Figure C.1:  $Z\gamma$  ISR-selected data (includes DY+jets and  $Z\gamma$  events),  $I_{ch}^\gamma$  (top) and  $\sigma_{i\eta i\eta}^\gamma$  (bottom) fake- $\gamma$  templates with and without  $E_T^{\text{miss}}$  requirement.

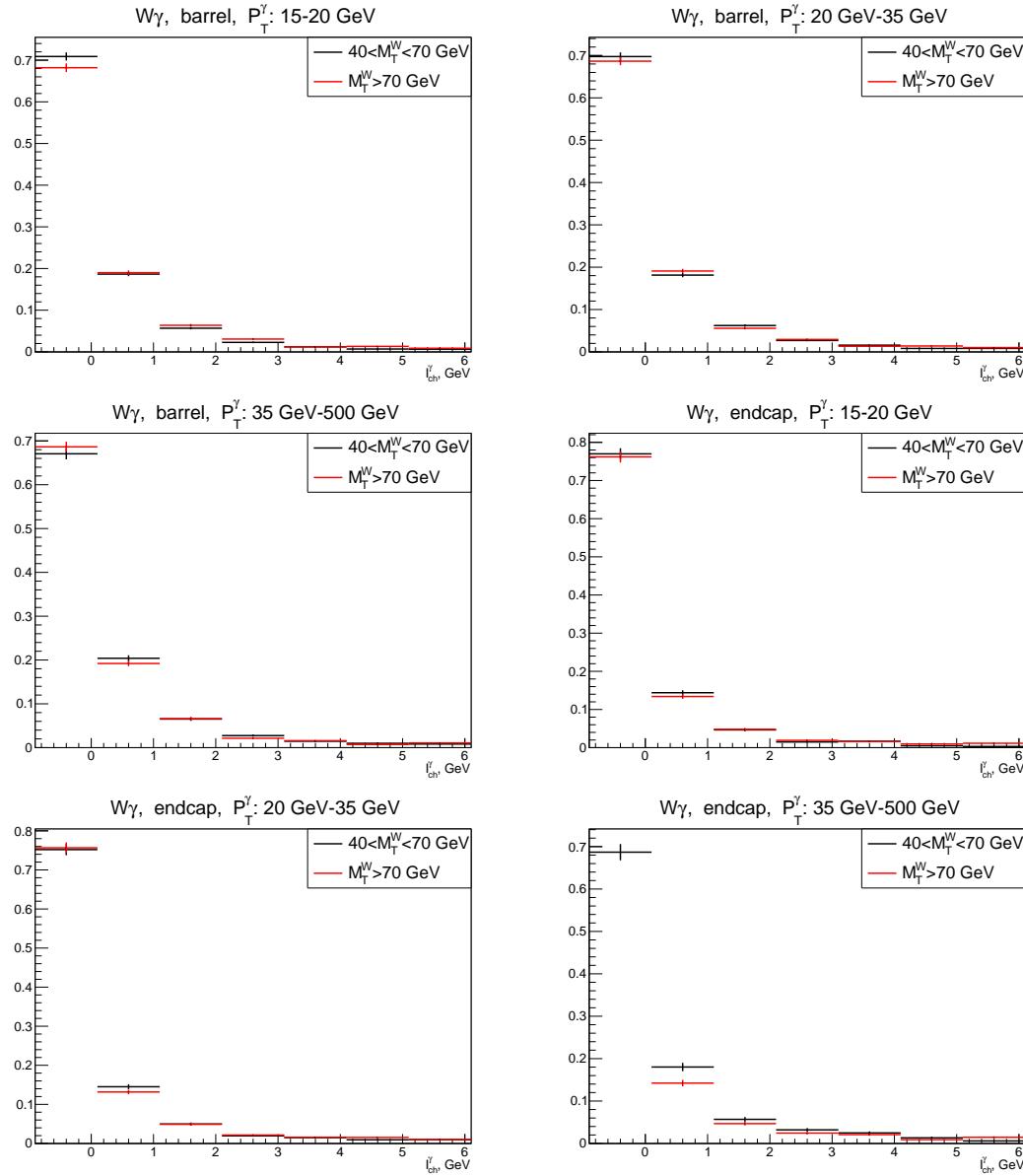


Figure C.2:  $W\gamma$  MC, muon channel,  $I_{ch}^{\gamma}$  templates.

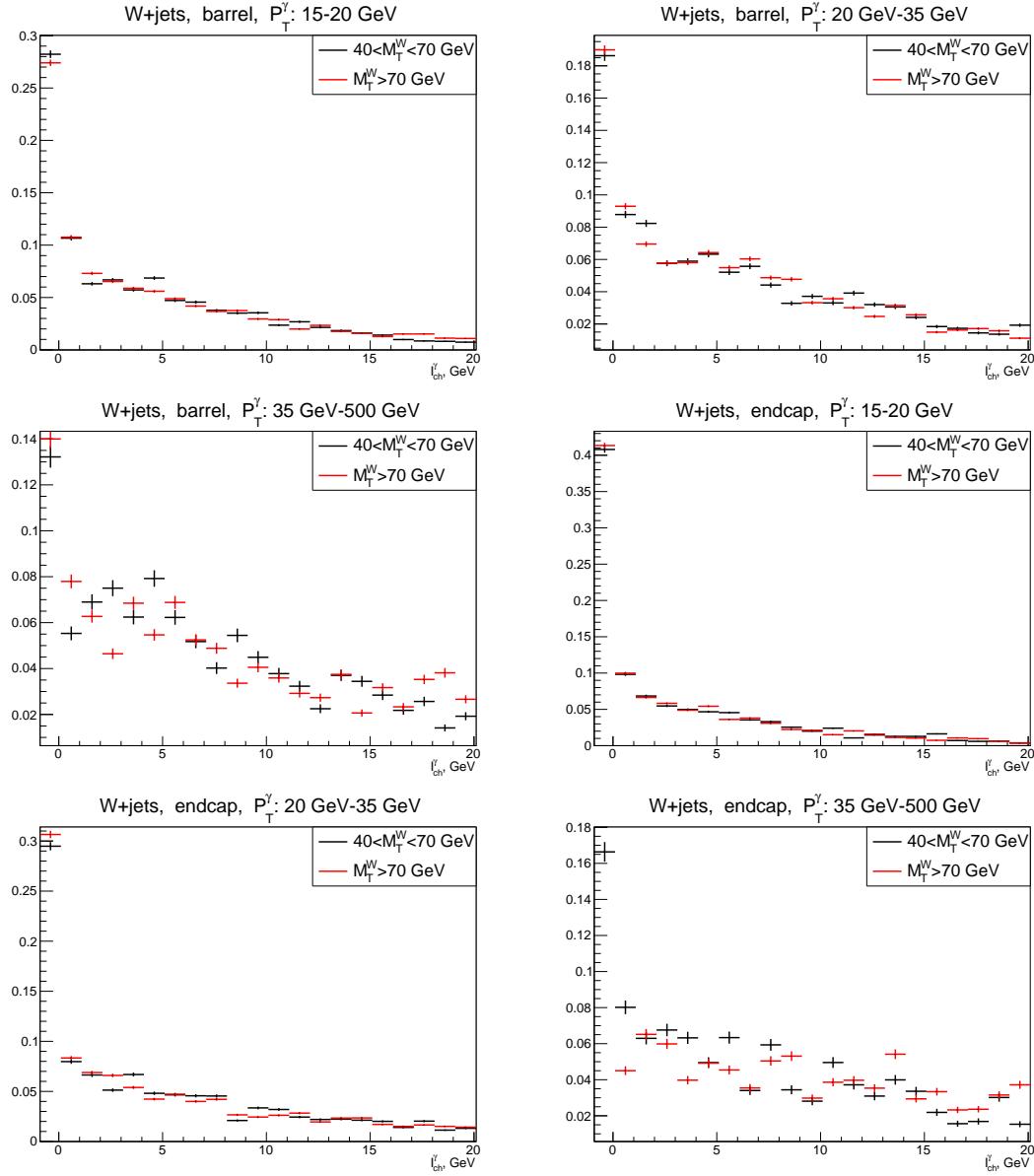


Figure C.3:  $W+jets$  MC, muon channel,  $I_{ch}^{\gamma}$  templates.

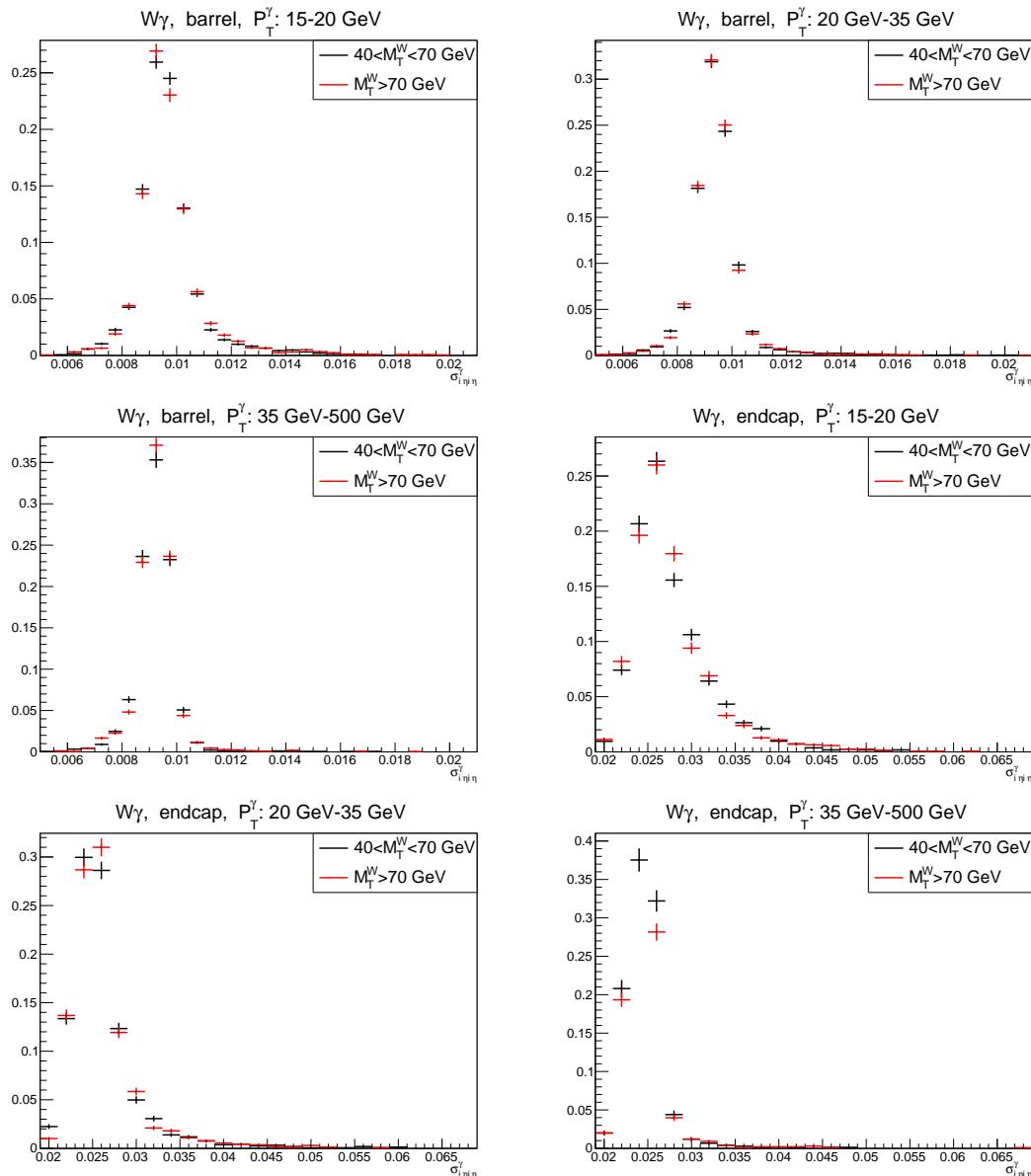


Figure C.4:  $W\gamma$  MC, muon channel,  $\sigma_{inj}^\gamma$  templates.

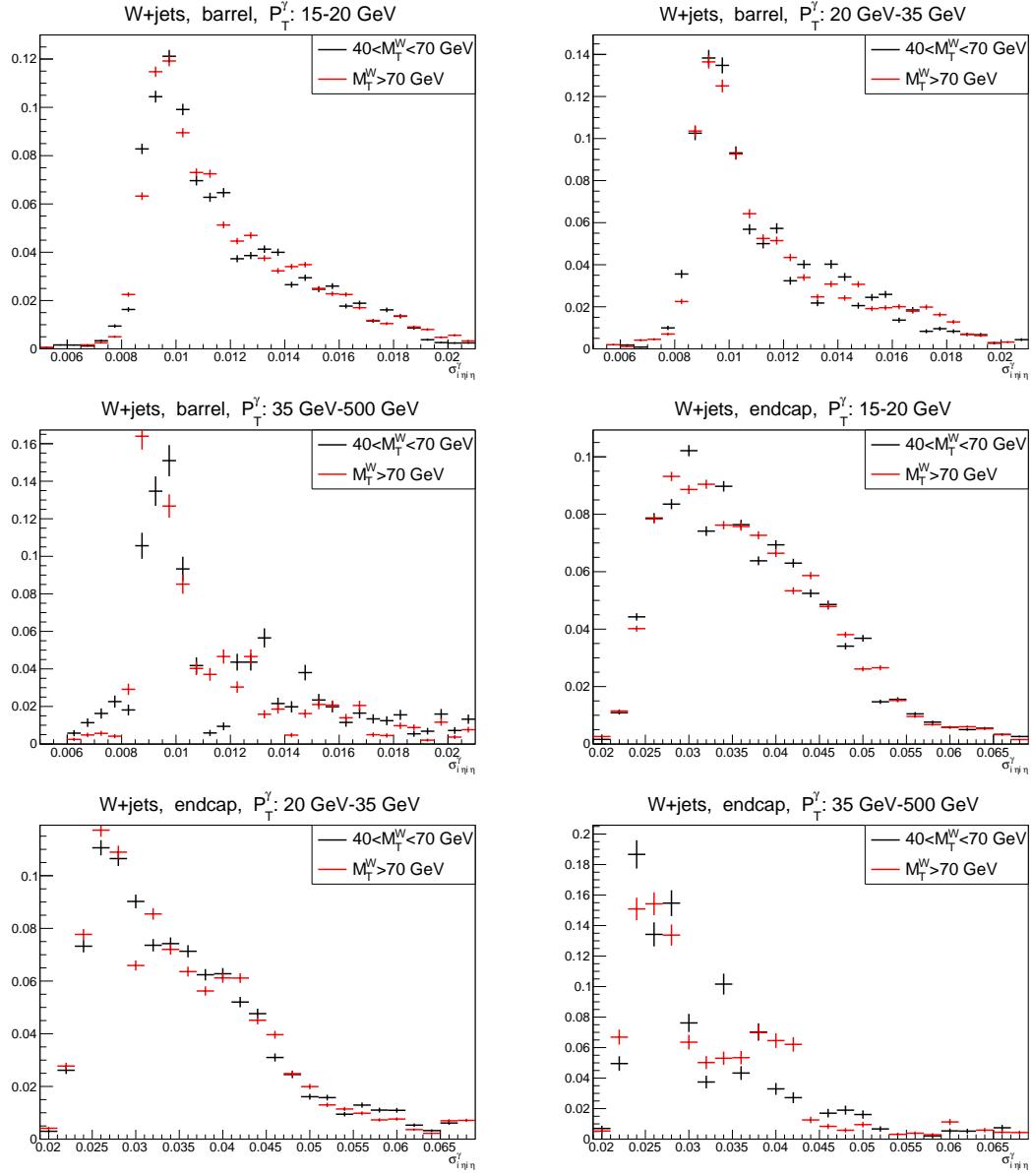


Figure C.5:  $W + \text{jets}$  MC, muon channel,  $\sigma_{i\eta i\eta}$  templates.

## Appendix D

### Template Fit Plots, $W\gamma$ , Data

This appendix contains fit results for jets $\rightarrow\gamma$  background estimation. On any plot, black histogram is data, green is a real- $\gamma$  template, blue is a fake- $\gamma$  template, and red is the fit function. These fits are part of the procedure of jets $\rightarrow\gamma$  background estimation which is described in Ch. 5.4.1.

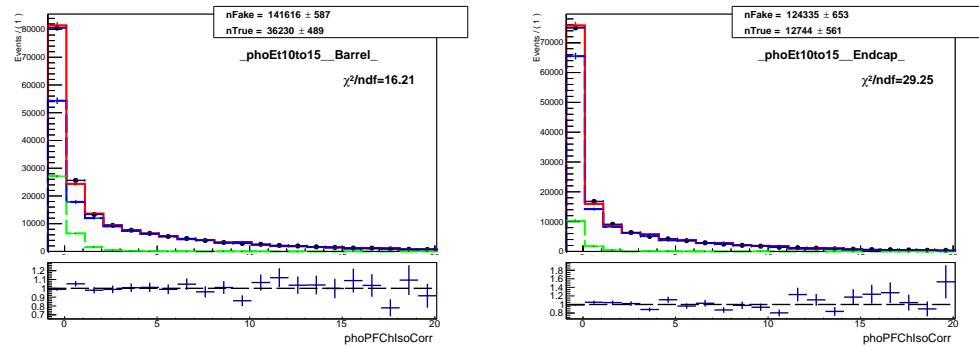


Figure D.1: Fits of  $I_{ch}^\gamma$  templates,  $W\gamma$ , muon channel, underflow bin (10 – 15 GeV).

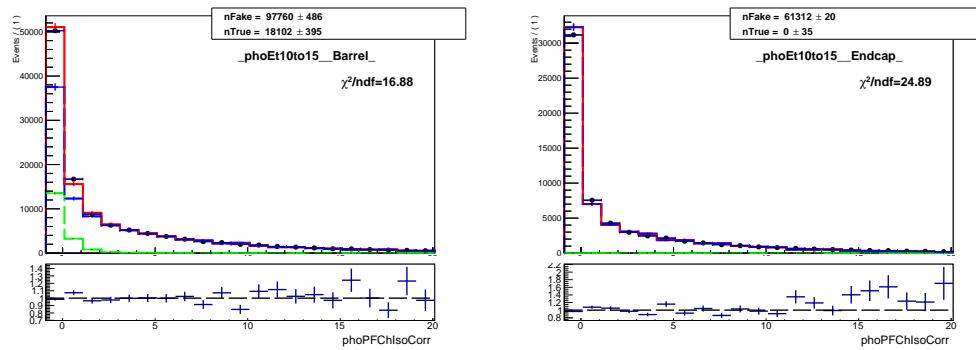


Figure D.2: Fits of  $I_{ch}^{\gamma}$  templates,  $W\gamma$ , electron channel, underflow bin (10 – 15 GeV).

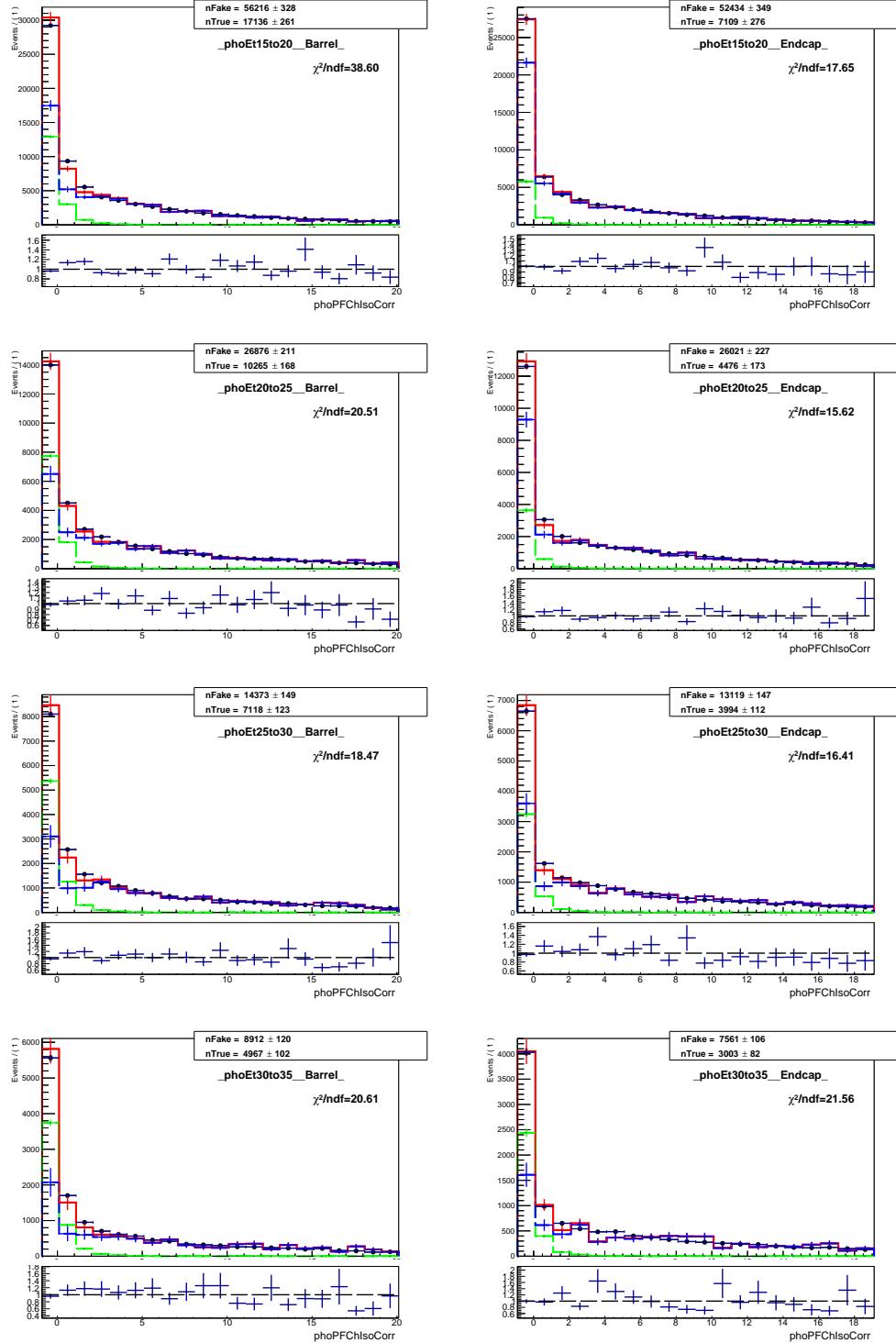


Figure D.3: Fits of  $I_{ch}^{\gamma}$  templates,  $W\gamma$ , muon channel.

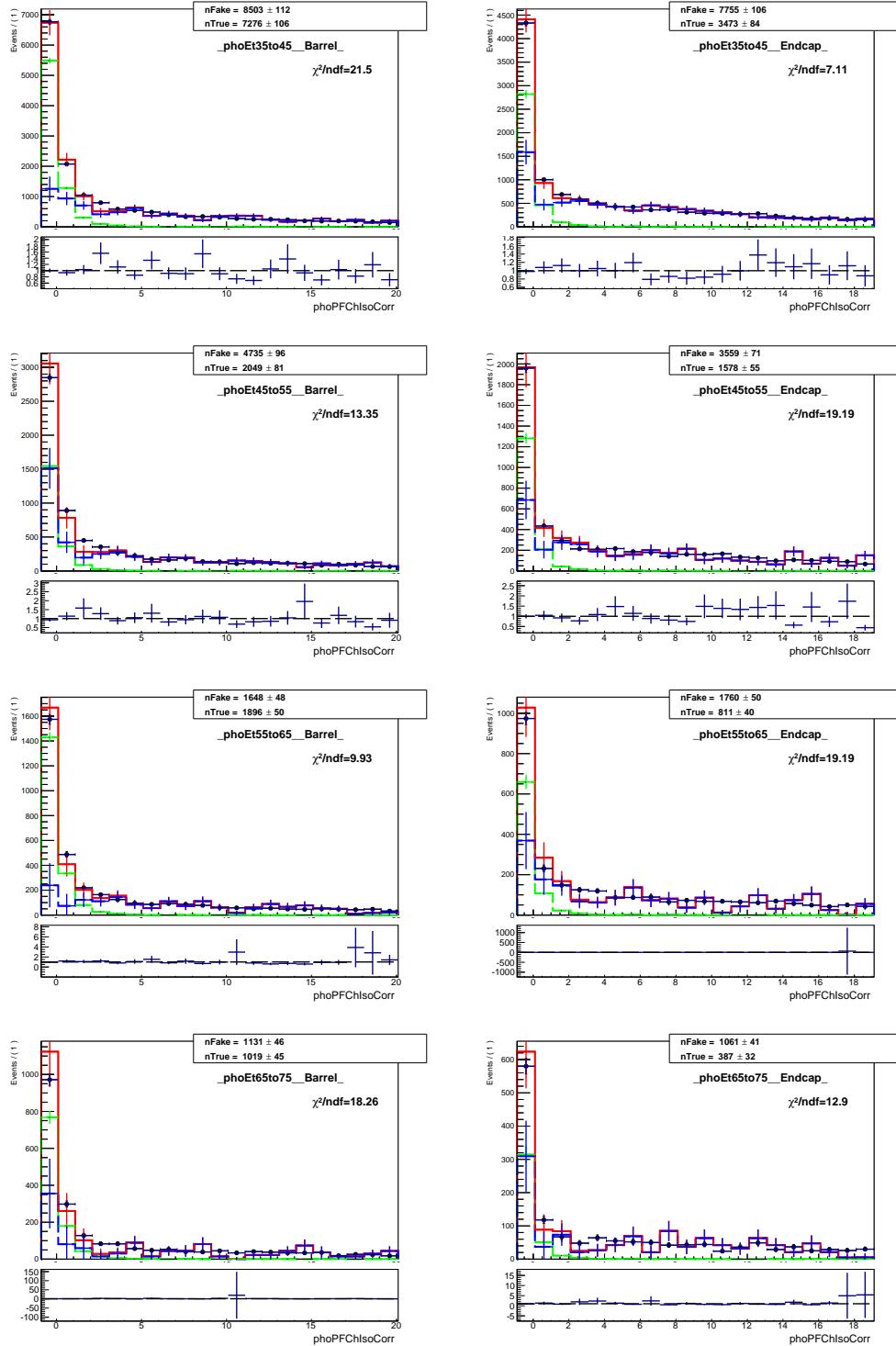


Figure D.4: Fits of  $I_{ch}^{\gamma}$  templates,  $W\gamma$ , muon channel.

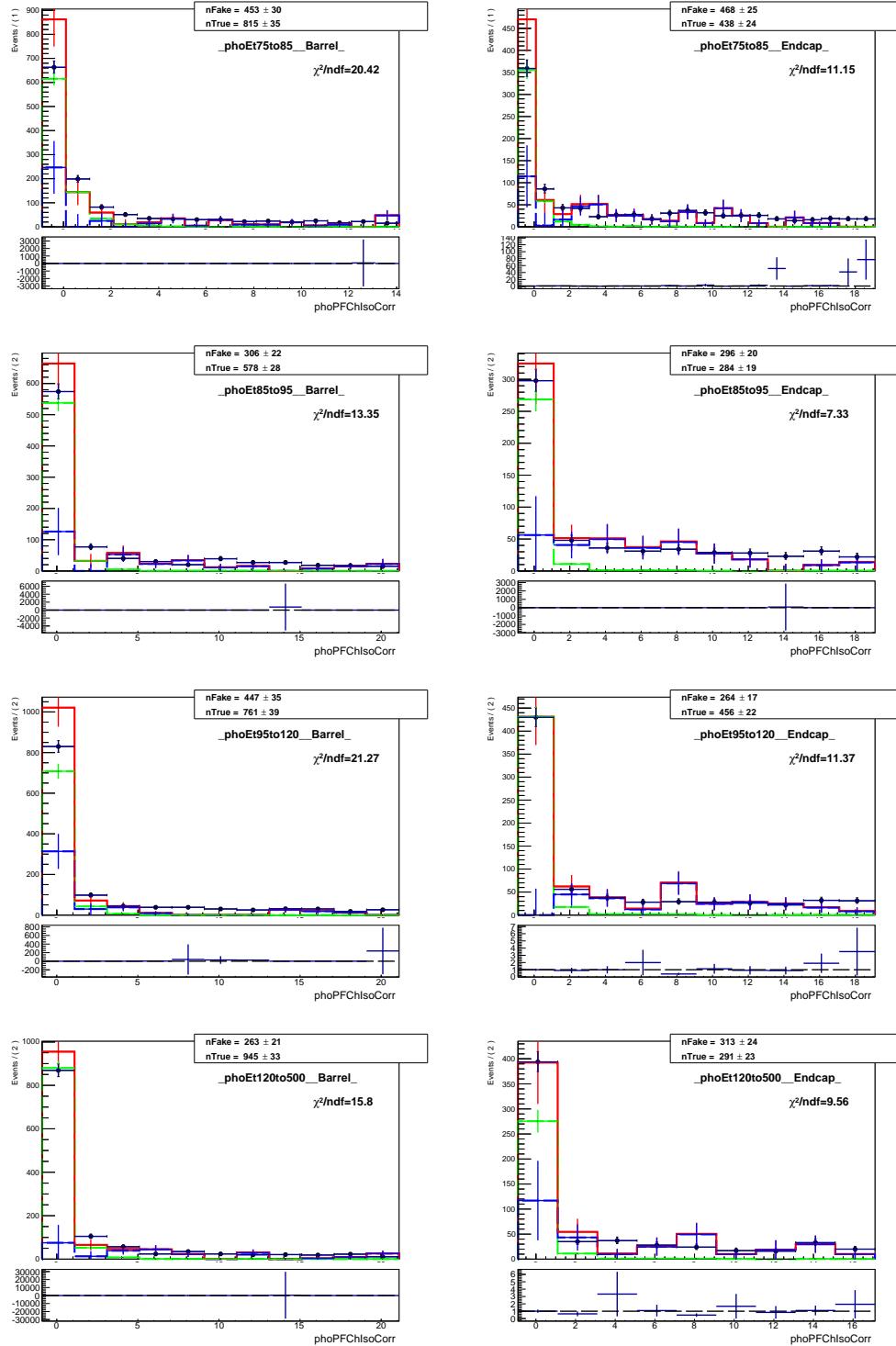


Figure D.5: Fits of  $I_{ch}^{\gamma}$  templates,  $W\gamma$ , muon channel.

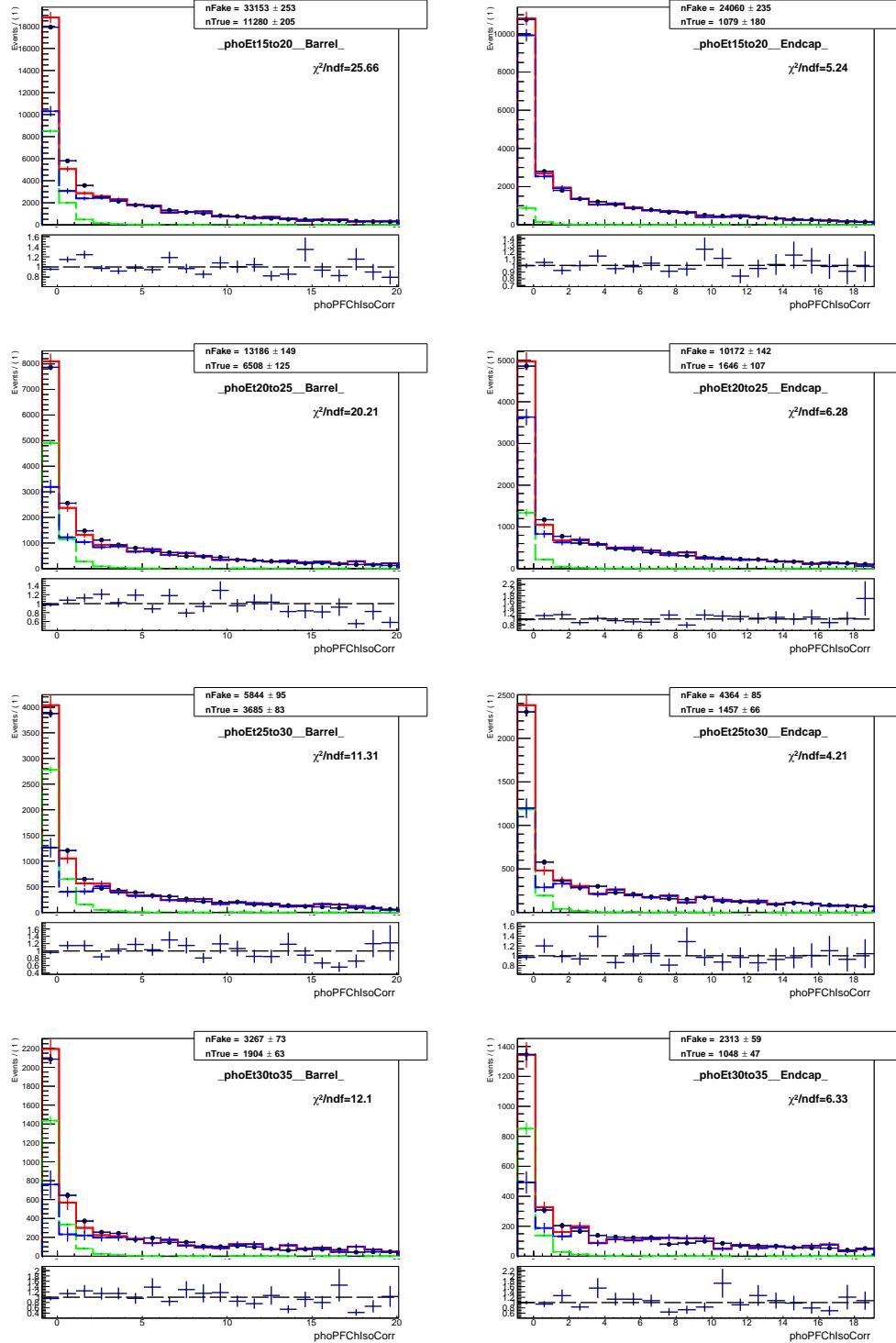


Figure D.6: Fits of  $I_{ch}^{\gamma}$  templates,  $W\gamma$ , electron channel.

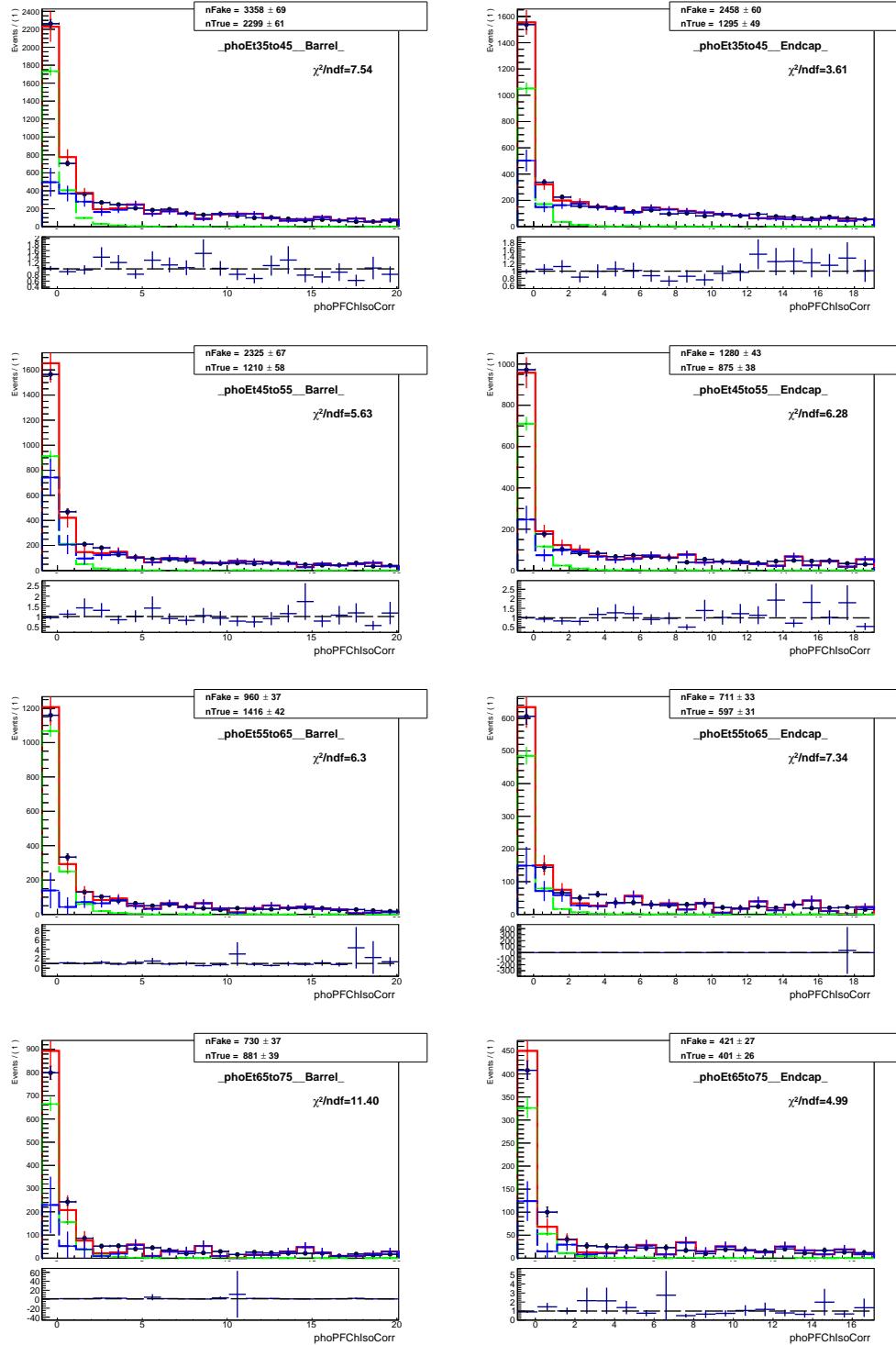


Figure D.7: Fits of  $I_{ch}^{\gamma}$  templates,  $W\gamma$ , electron channel.

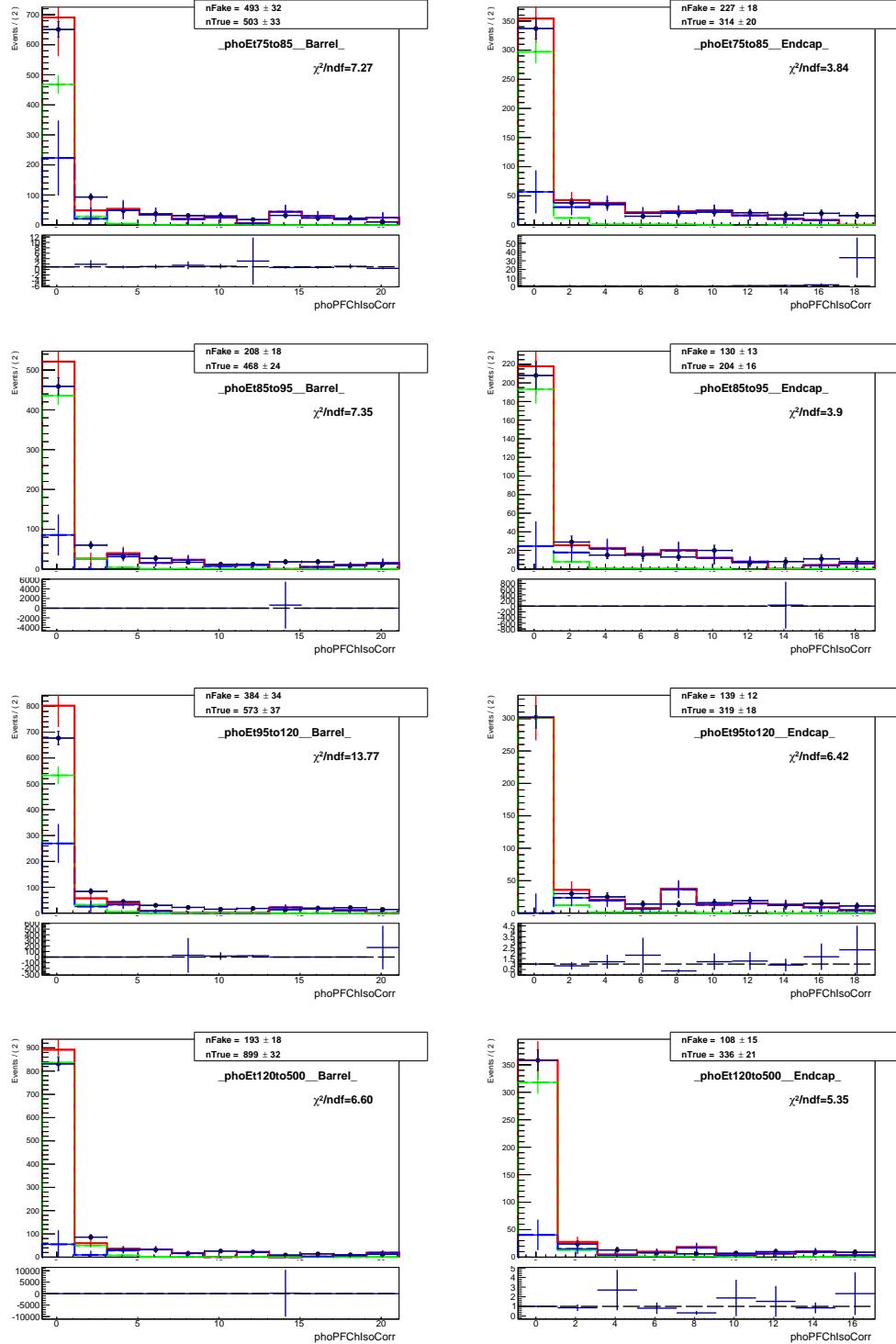


Figure D.8: Fits of  $I_{ch}^{\gamma}$  templates,  $W\gamma$ , electron channel.

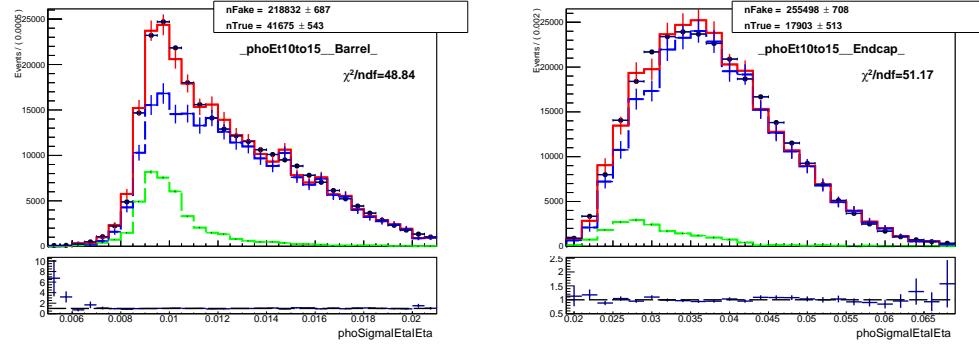


Figure D.9: Fits of  $\sigma_{i\eta i\eta}$  templates,  $W\gamma$ , muon channel, underflow bin ( $10 - 15$  GeV).

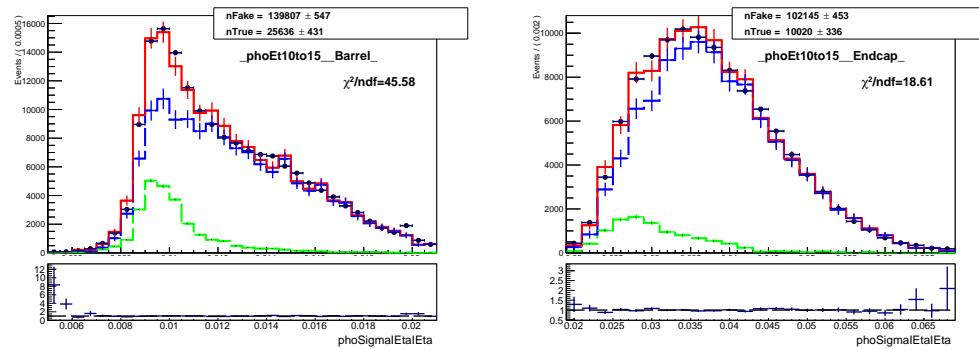


Figure D.10: Fits of  $\sigma_{i\eta i\eta}$  templates,  $W\gamma$ , electron channel, underflow bin ( $10 - 15$  GeV).

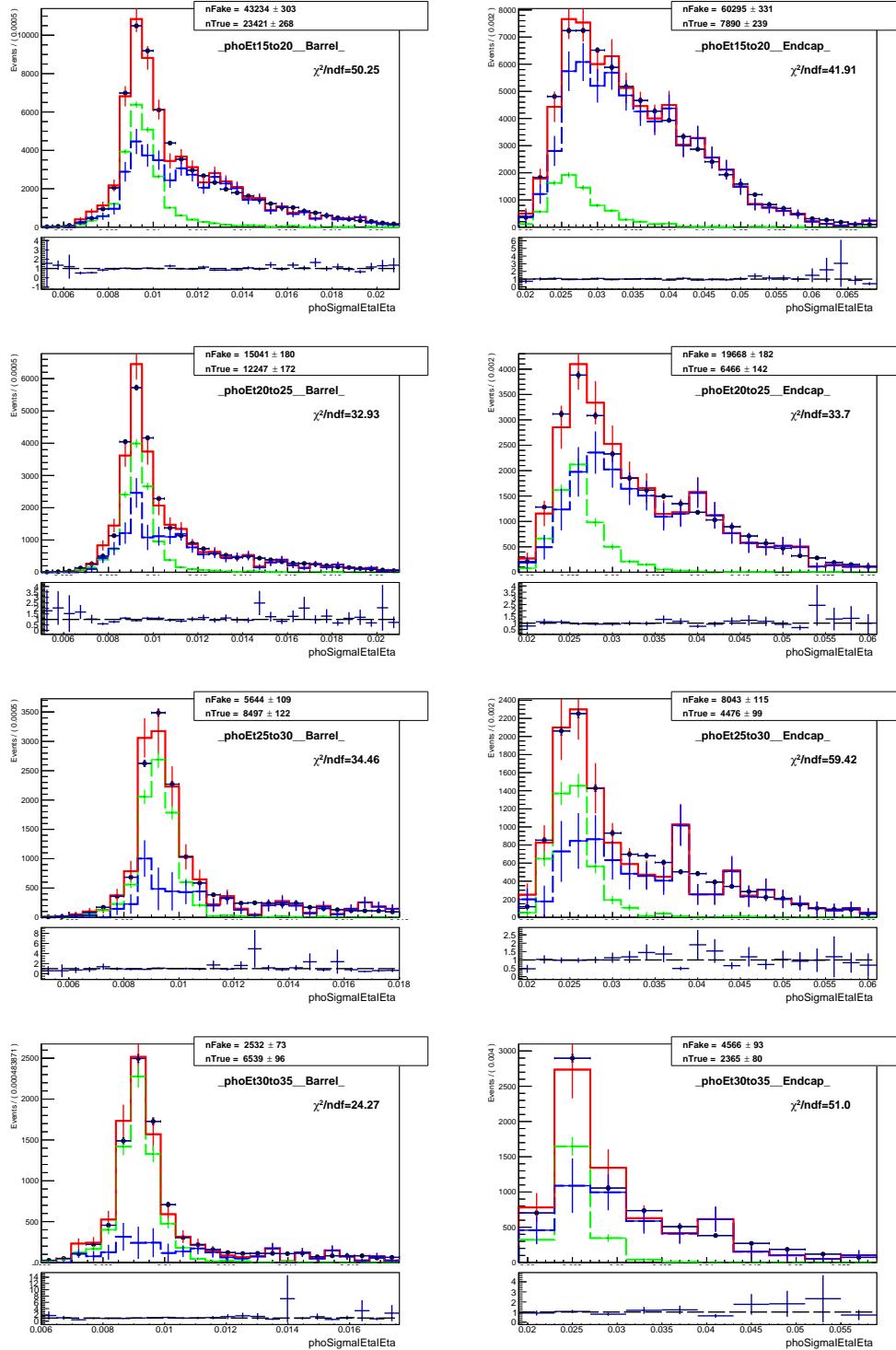


Figure D.11: Fits of  $\sigma_{i\eta i\eta}$  templates,  $W\gamma$ , muon channel.

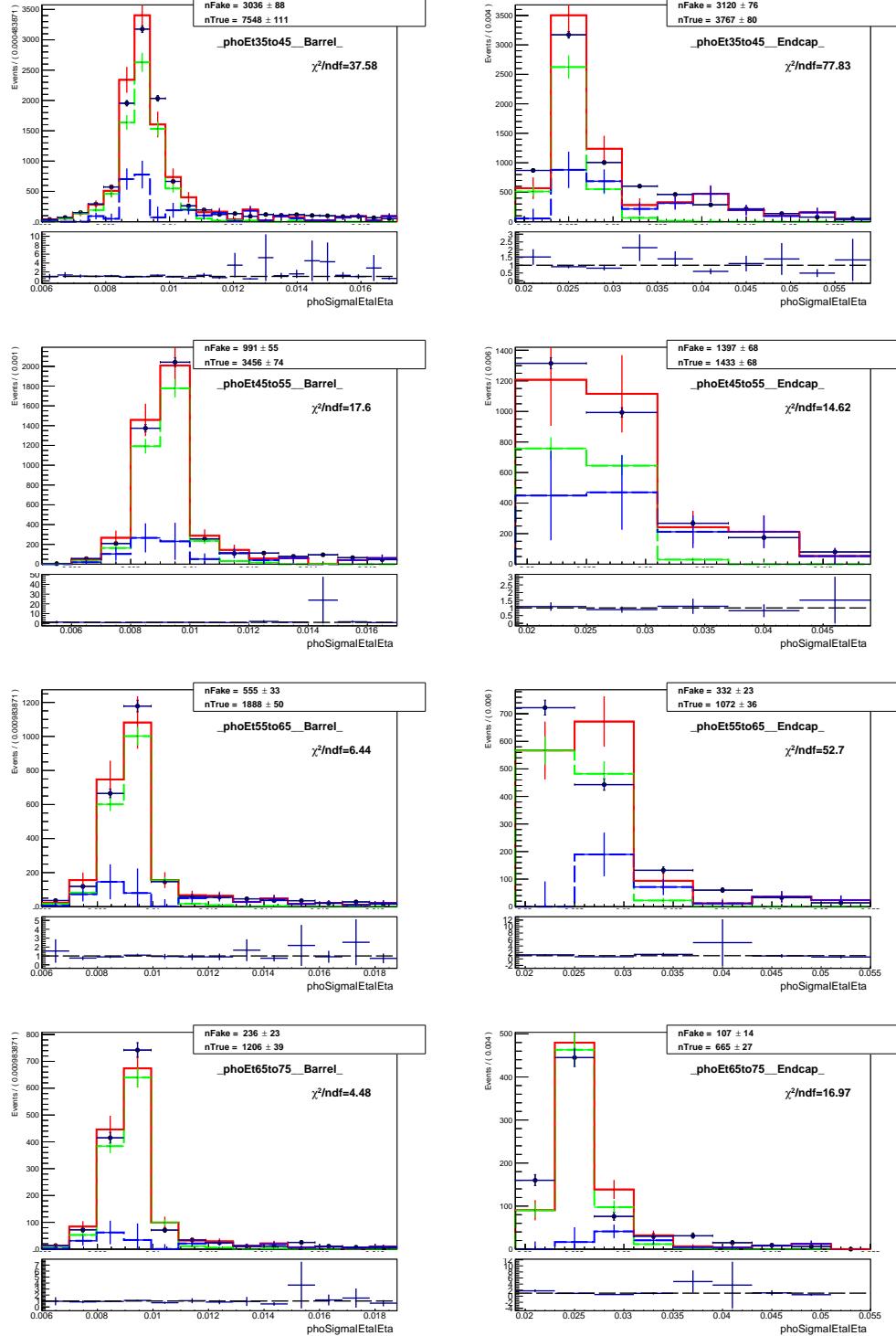


Figure D.12: Fits of  $\sigma_{i\eta i\eta}$  templates,  $W\gamma$ , muon channel.

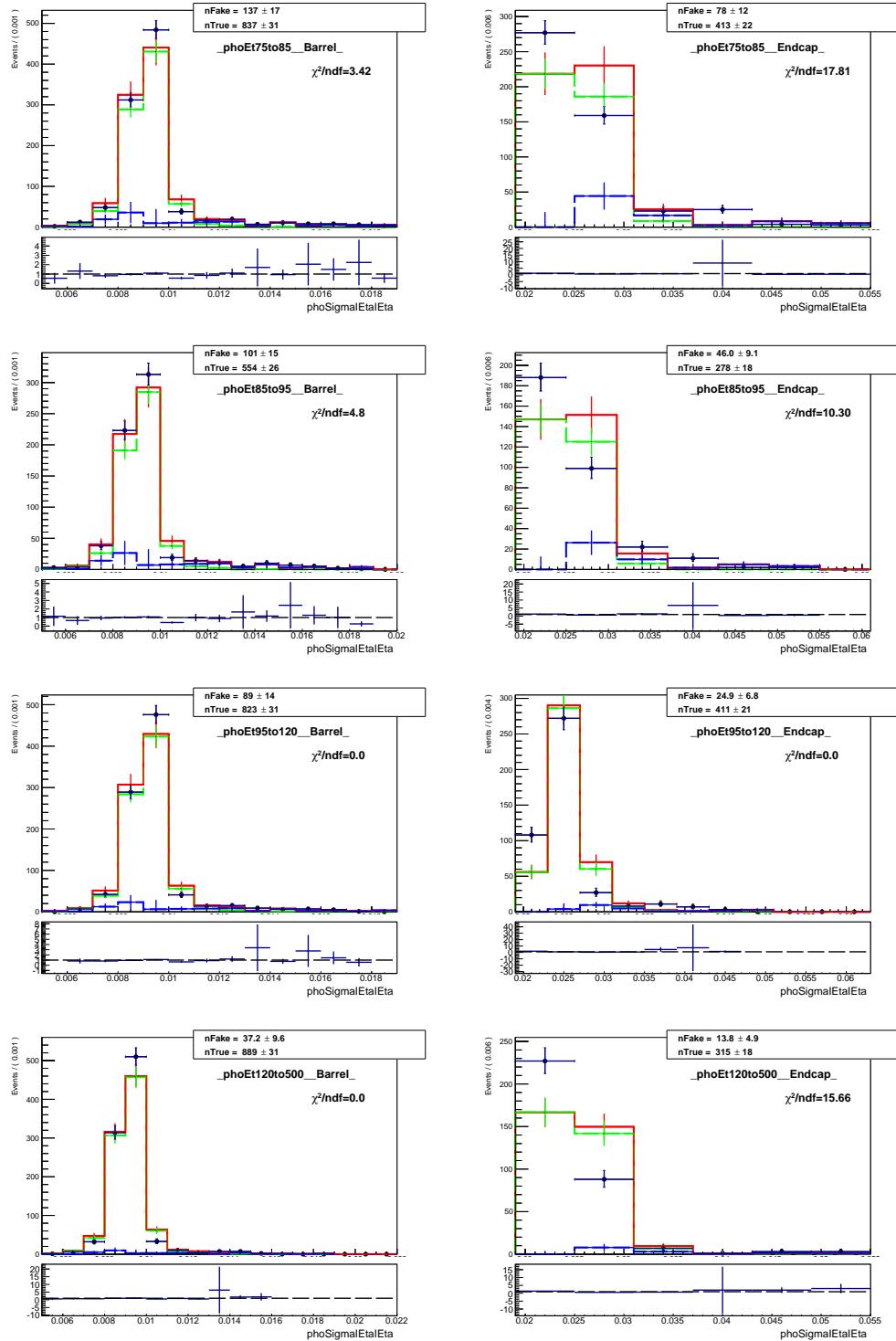


Figure D.13: Fits of  $\sigma_{ijkl}$  templates,  $W\gamma$ , muon channel.

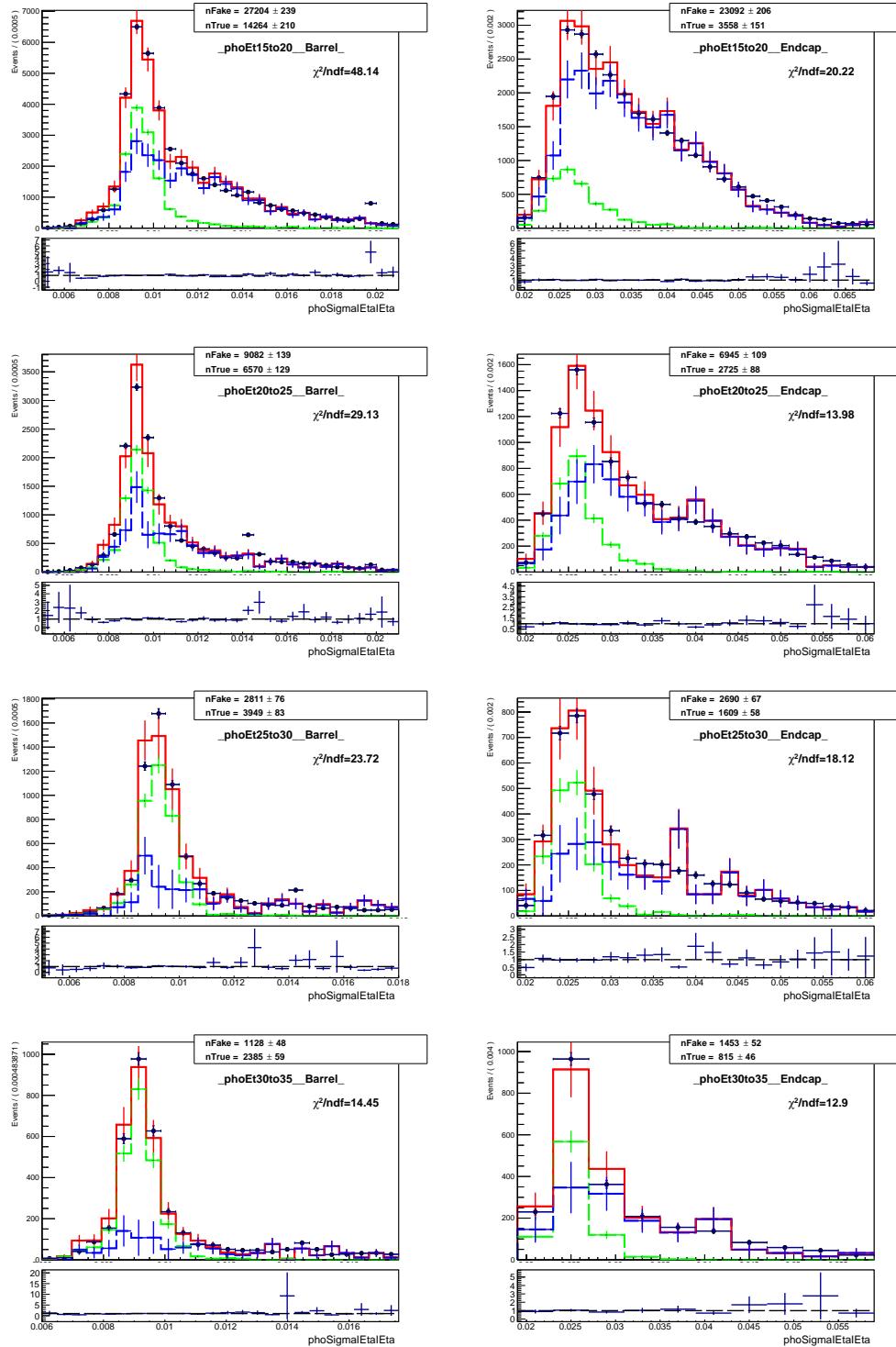


Figure D.14: Fits of  $\sigma_{i\eta i\eta}$  templates,  $W\gamma$ , electron channel.

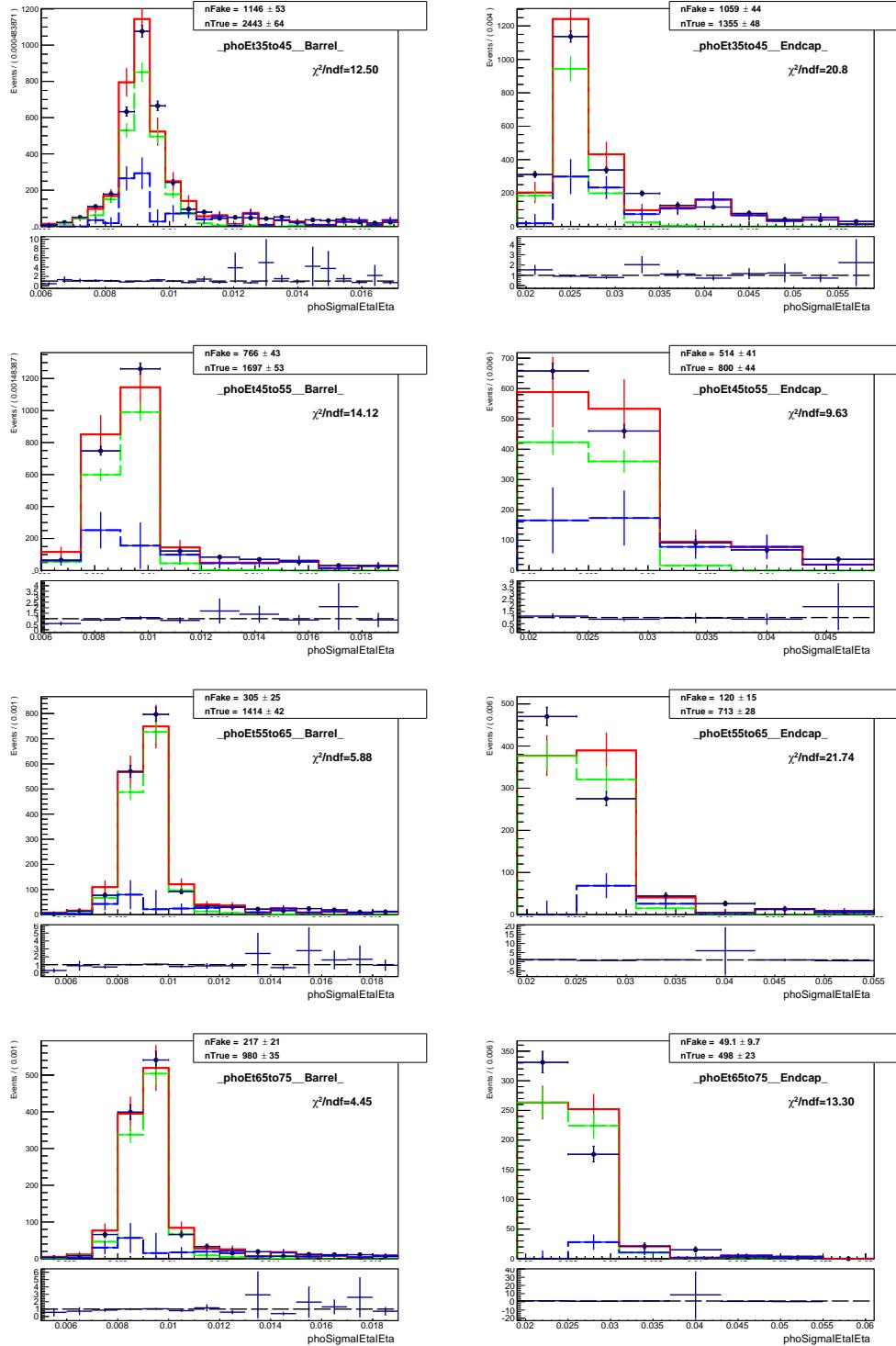


Figure D.15: Fits of  $\sigma_{i\eta i\eta}$  templates,  $W\gamma$ , electron channel.

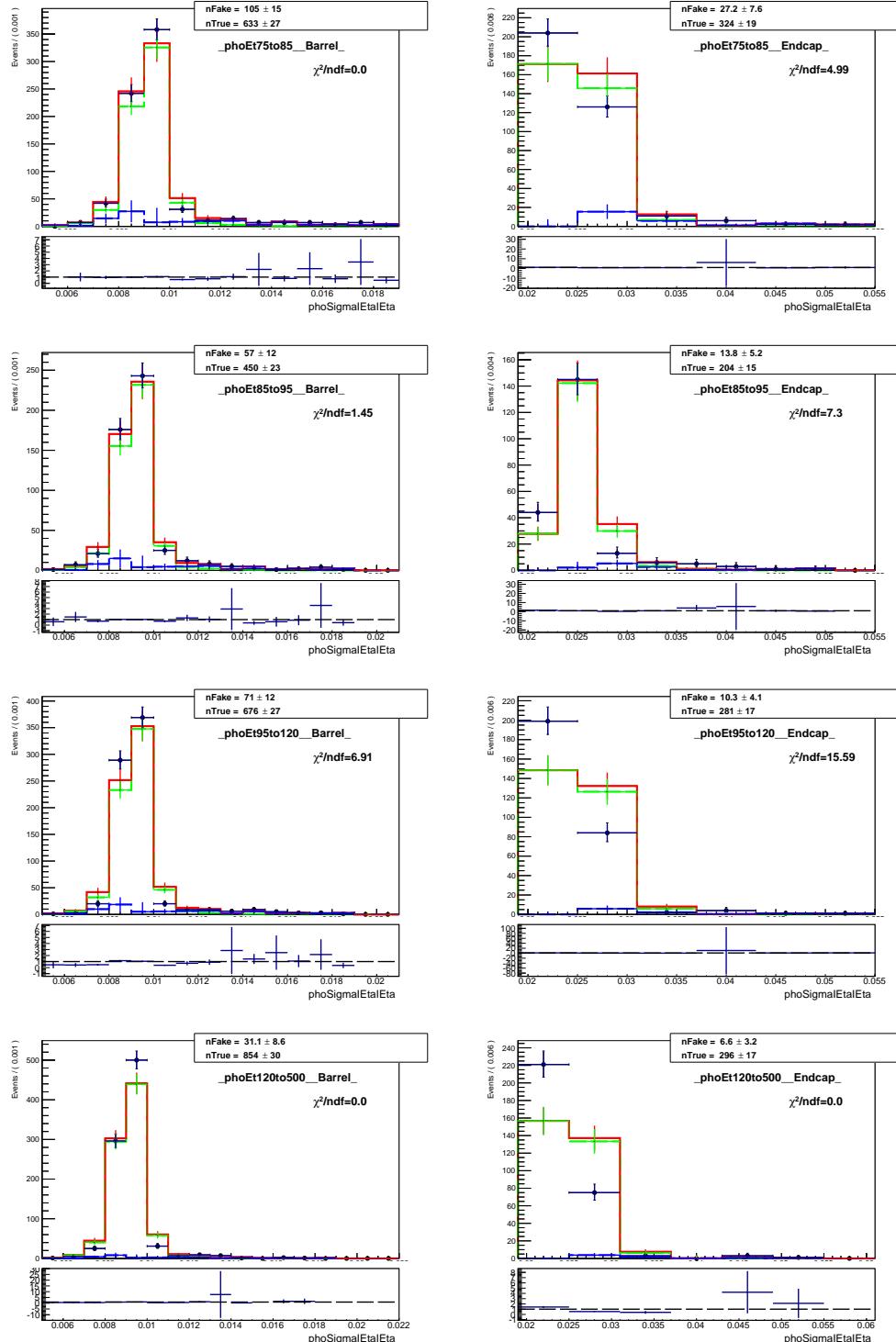


Figure D.16: Fits of  $\sigma_{i\eta i\eta}$  templates,  $W\gamma$ , electron channel.

## Appendix E

### MC Closure Check

Fit results of jets $\rightarrow \gamma$  background estimation on pseudodata prepared as  $W+\text{jets}$  and  $W\gamma$  samples appropriately weighted and mixed together. Templates are prepared from  $W\gamma$  and  $W+\text{jets}$  MC samples, same as those that constitute pseudodata. On all the plots, the black histogram is the pseudodata, green is a real- $\gamma$  template, blue is a fake- $\gamma$  template, and red is a fit function. Cyan is a histogram of  $W+\text{jets}$  MC sample in a given channel.

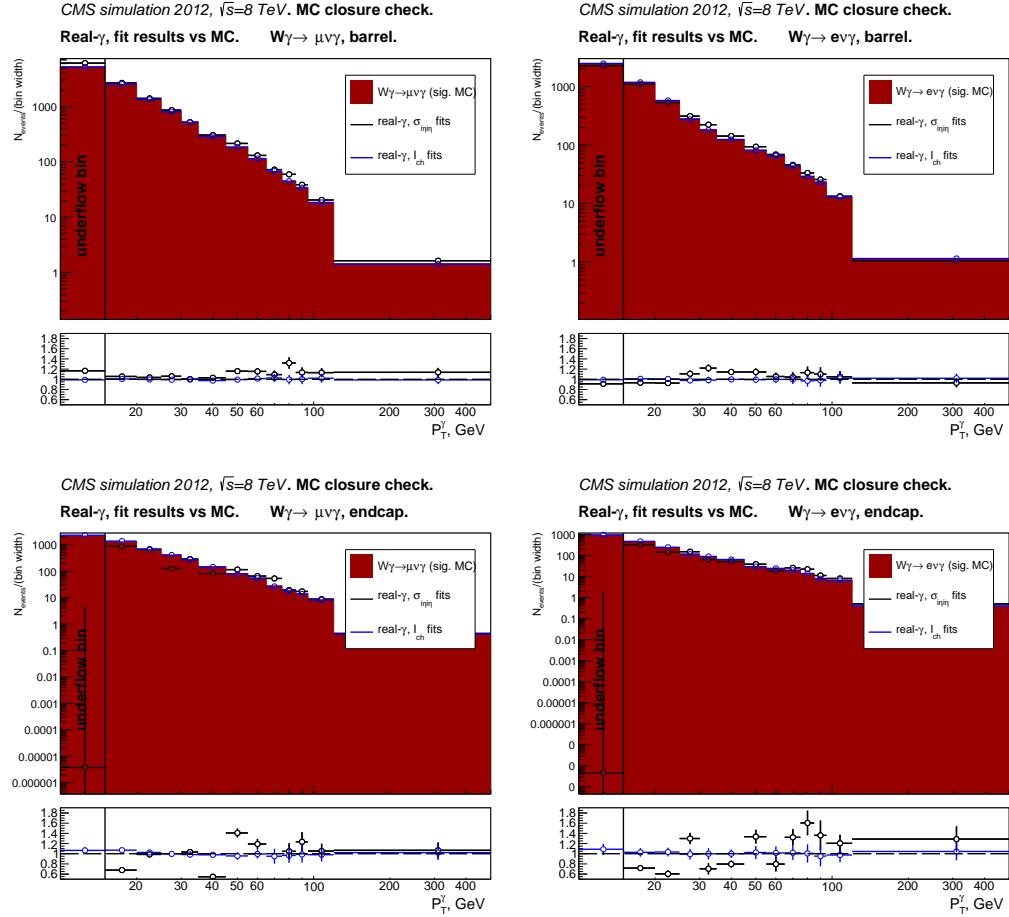


Figure E.1: Real- $\gamma$  yields derived from fits of pseudodata superimposed with  $W\gamma$  MC. Left column - muon channel, right column - electron channel. Top to bottom: barrel and endcap photons.

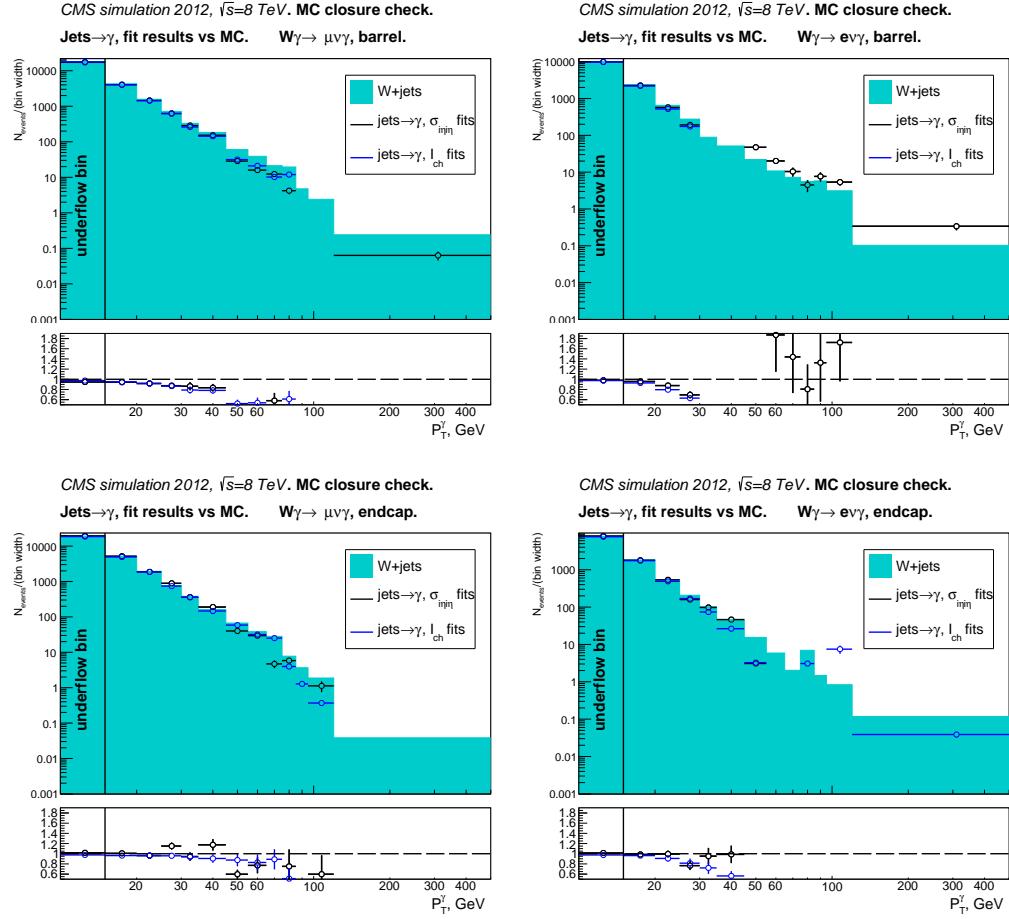


Figure E.2:  $W+jets$  yields derived from fits of pseudodata superimposed with  $W+jets$  MC. Left column - muon channel, right column - electron. Top to bottom: barrel and endcap photons.

To perform the MC closure check, the pseudodata samples are prepared to mimic each of the data samples needed for the fit. This includes:

- $W\gamma$ -selected samples to be used for fit (full analysis selection without  $I_{ch}$  or  $\sigma_{i\eta i\eta}$  requirements), those samples are prepared separately for the muon and electron channels;
- $W\gamma$  in full selection to be plotted into the [data vs bkg+signal] plots, separately for the two channels;
- $Z\gamma \rightarrow \mu\mu\gamma$  FSR selected sample for the real- $\gamma$  templates;
- $Z\gamma \rightarrow \mu\mu\gamma + (\text{DY+jets})$  ISR selected sample for the fake- $\gamma$  templates.

To prepare the pseudodata  $W\gamma$ -selected samples, the  $W\gamma$ ,  $W+\text{jets}$ ,  $Z\gamma$ ,  $\text{DY}+\text{jets}$ ,  $t\bar{t}\gamma$ ,  $t\bar{t}+\text{jets}$ , and  $WW\gamma$  MC samples are merged with the  $W\gamma$  selection applied. To prepare the pseudodata  $Z\gamma$ -selected samples that are used for the template extraction, the  $Z\gamma$  and  $\text{DY}+\text{jets}$  MC samples are merged with the appropriate  $Z\gamma$  selections applied. Luminosity normalizations, PU and scale factor weights are applied on all MC samples. Then the pseudodata are treated as if they were data to perform the fits, prepare the plots and subtract the background. All the MC samples used in the nominal  $W\gamma$  measurement are used for this MC closure check the same way.

Figures E.3-E.6 show the selection and background estimation results in data (left column) and pseudodata (right column). Top plots are data and pseudodata superimposed with MC. Because the pseudodata was prepared as MC samples merged together, there is an exact match in right-top plots by construction. Middle and bottom plots are results of the background estimation in data and pseudodata. Plots show data (pseudodata) vs background estimates and signal MC where

jets  $\rightarrow \gamma$  background is estimated from fits of  $I_{ch}^{\gamma}$  (middle) and  $\sigma_{i\eta i\eta}^{\gamma}$  (bottom) templates. Plots of fits themselves on the pseudodata are available in App. ??.

The agreement in pseudodata vs background plus signal plots is significantly better than in data plots. Thus, the poor agreement in data can be partially explained by wrong normalizations of the signal MC and other real- $\gamma$  MC samples. The disagreement between  $I_{ch}^{\gamma}$  and  $\sigma_{i\eta i\eta}^{\gamma}$  fit results in pseudodata is smaller than in data however is still significant in many  $P_T^{\gamma}$  bins which indicates systematic bias of the fit procedure.

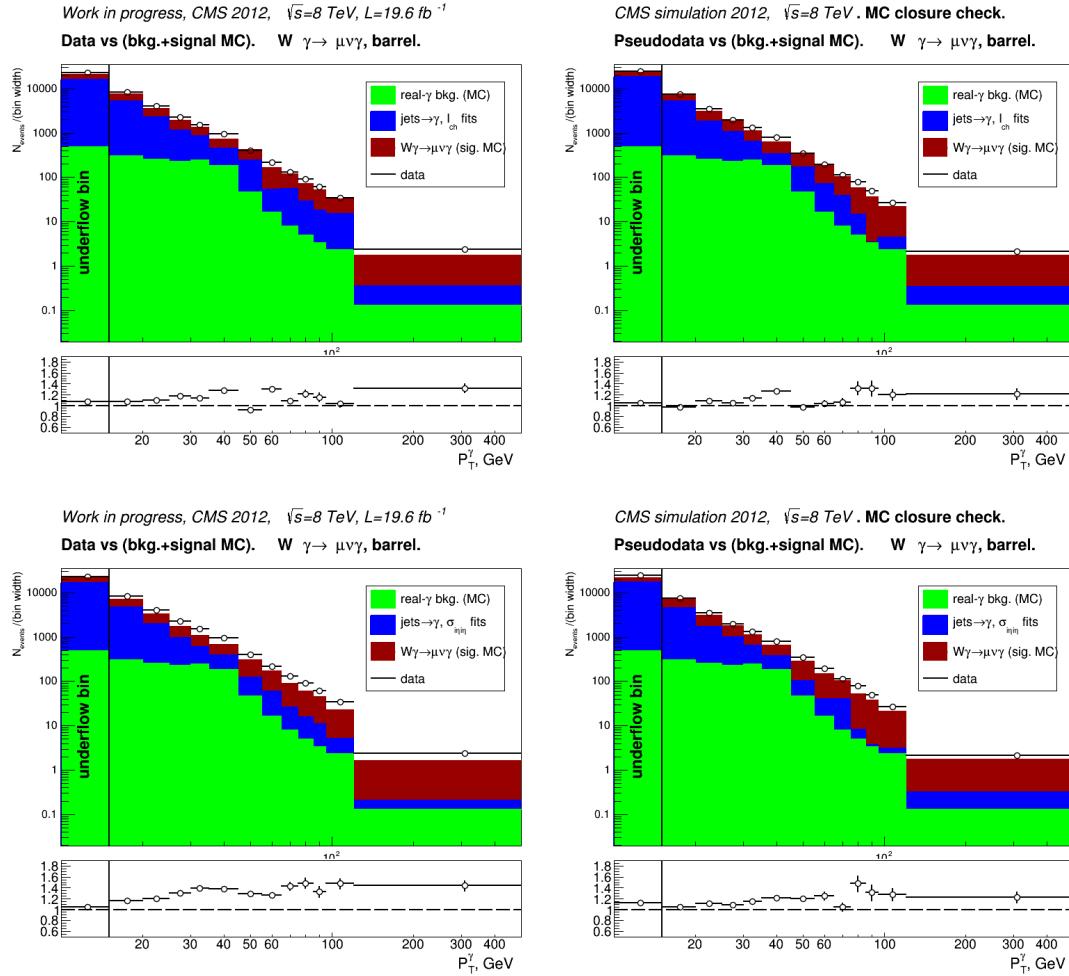


Figure E.3: Data (left) and pseudodata (right) vs background estimates and signal MC in bins of  $P_T^\gamma$ . Jets  $\rightarrow \gamma$  background estimated from fits of  $I_{ch}^\gamma$  (top) and  $\sigma_{i\eta i\eta}^\gamma$  (bottom). Muon channel. Barrel photons.

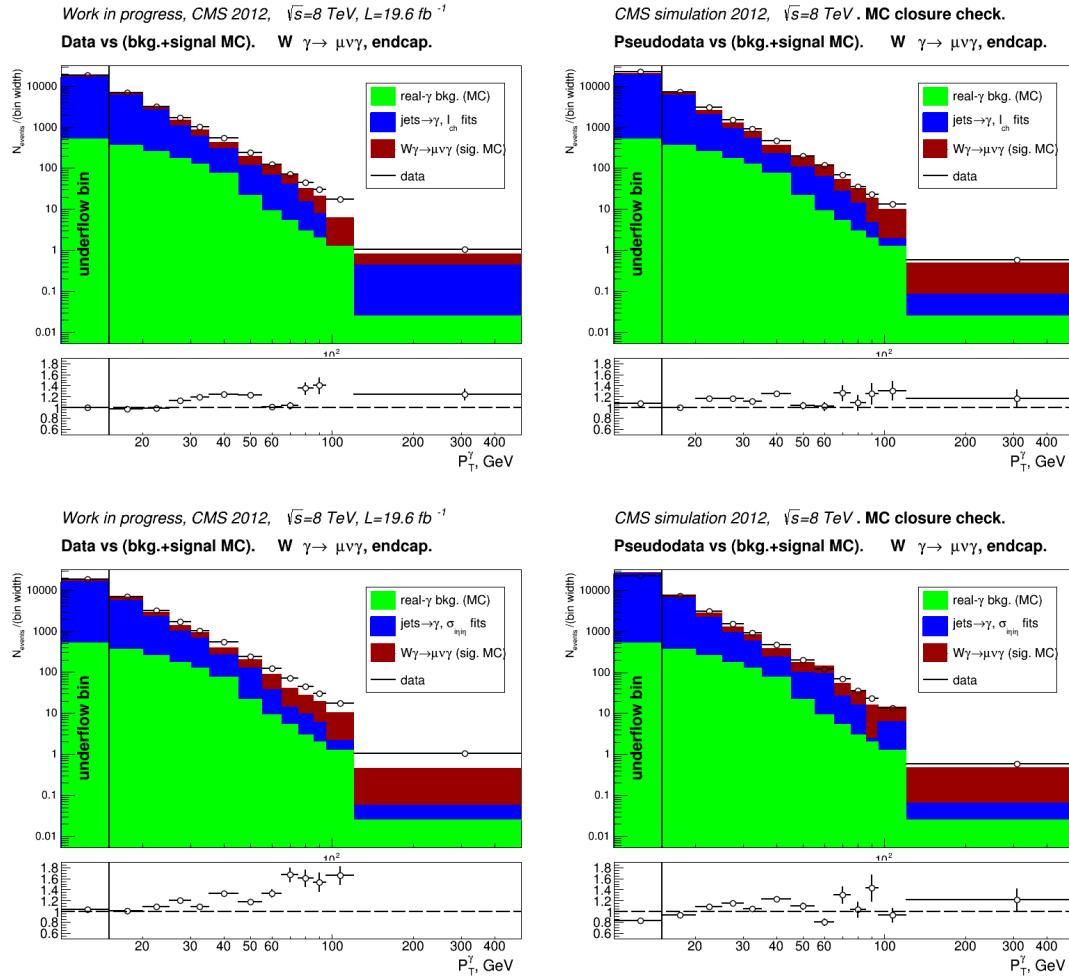


Figure E.4: Data (left) and pseudodata (right) vs background estimates and signal MC in bins of  $P_T^\gamma$ . Jets $\rightarrow\gamma$  background estimated from fits of  $I_{ch}^\gamma$  (middle) and  $\sigma_{i\eta i\eta}^\gamma$  (bottom). Muon channel. Endcap photons.

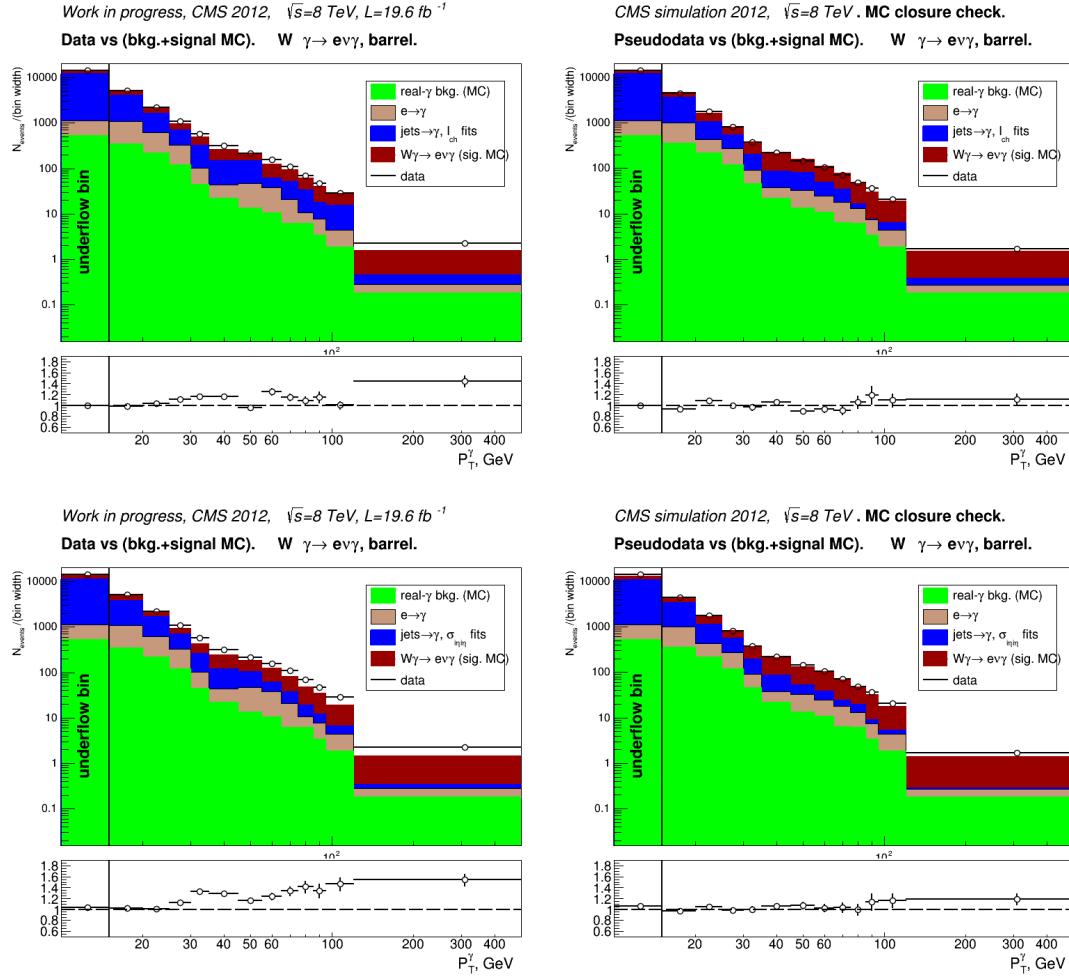


Figure E.5: Data (left) and pseudodata (right) vs background estimates and signal MC in bins of  $P_T^\gamma$ . Jets $\rightarrow\gamma$  background estimated from fits of  $I_{ch}^\gamma$  (middle) and  $\sigma_{i\eta i\eta}^\gamma$  (bottom). Electron channel. Barrel photons.

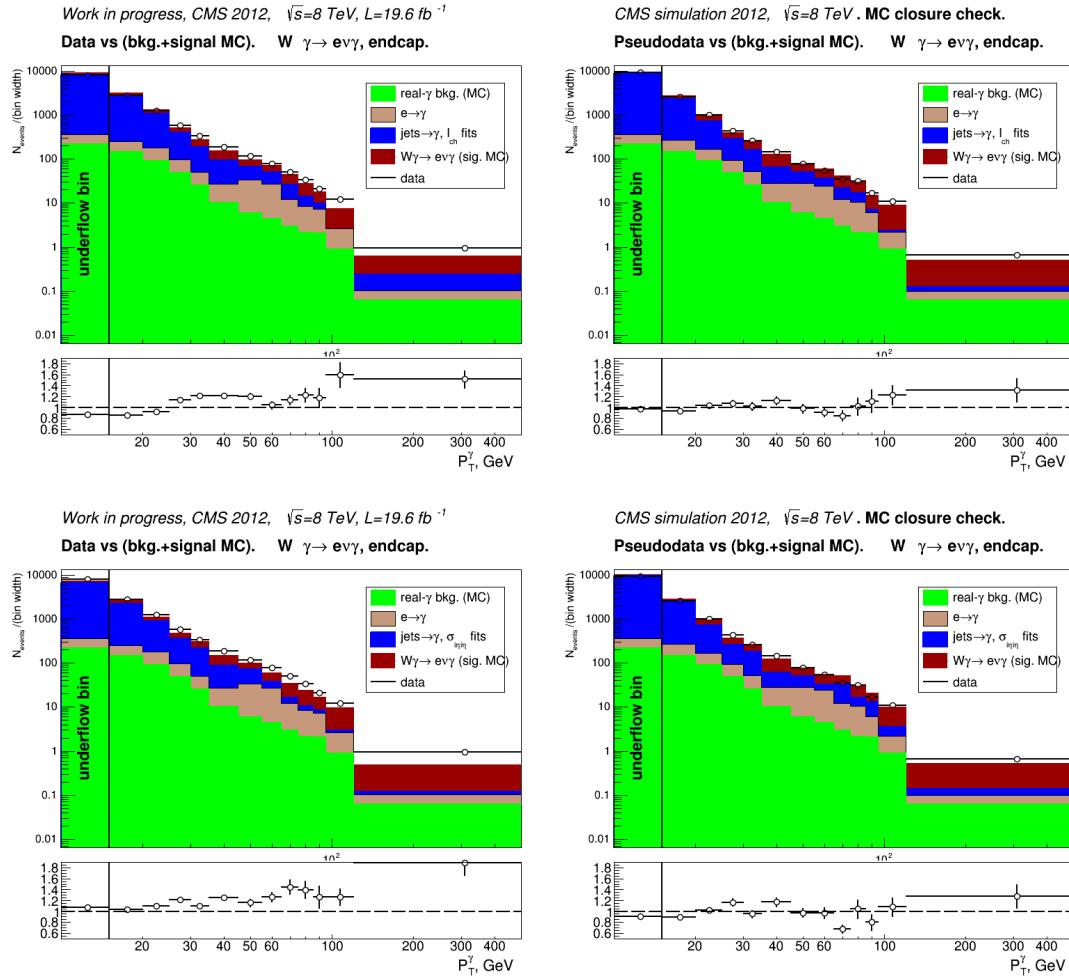


Figure E.6: Data (left) and pseudodata (right) vs background estimates and signal MC in bins of  $P_T^\gamma$ . Jets  $\rightarrow \gamma$  background estimated from fits of  $I_{ch}^\gamma$  (middle) and  $\sigma_{ijinj}^\gamma$  (bottom). Electron channel. Endcap photons.

## Appendix F

### Fit Plots of $M_{e\gamma}$

Fit results of electron-photon invariant mass  $M_{e\gamma}$  for the  $e \rightarrow \gamma$  data-driven estimation in the electron channel. The procedure of the background estimation is described in Ch. 5.4.2.

The number of  $e\gamma$  events in data under the Z-peak  $N_{MC-Zpeak}^{e\rightarrow\gamma}$  is extracted from the fit of the model:

$$F_{fit}^{e\rightarrow\gamma} = N_{e\rightarrow\gamma} \cdot (RooNDKeysPdf * Gaussian) + N_{else} \cdot (RooCMSShapePdf). \quad (F.1)$$

The function *RooNDKeysPdf* is part of the *RooFit* package [49] and the *RooCMSShapePdf* was developed specifically for CMS [50].

$F_{fit}^{e\rightarrow\gamma}$  has eight fit parameters. “Nsig” and “Nbkg” is the plots are  $N_{e\rightarrow\gamma}$  and  $N_{else}$ , respectively, “mean\_gau” and “sigma\_gau” are parameters of the Gaussian distribution, “CMS\_alpha” and “CMS\_beta” are parameters of the exponential component of the *RooCMSShapePdf*, and “CMS\_gamma” and “CMS\_peak” are parameters of the turn over component of the *RooCMSShapePdf*.

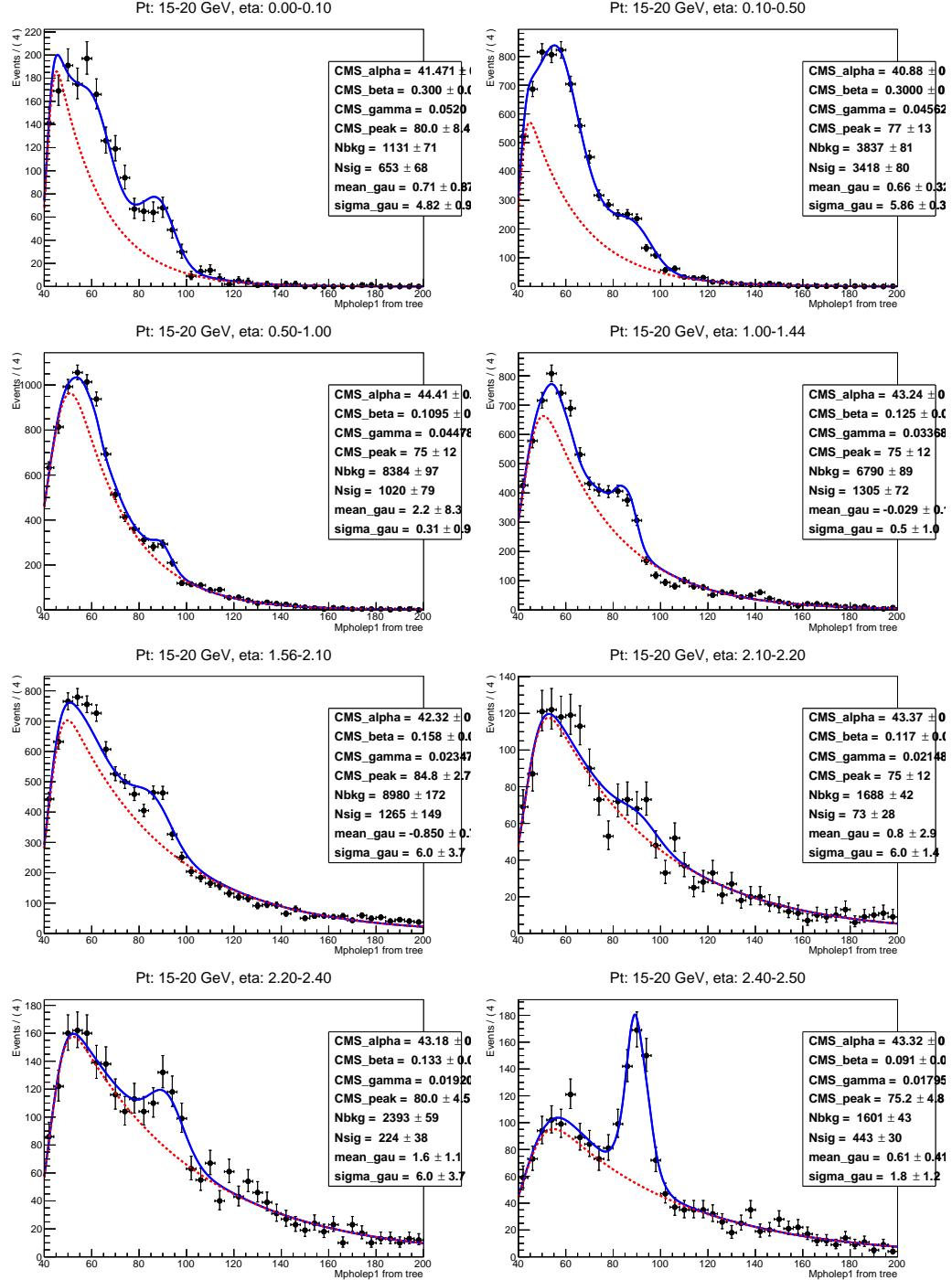


Figure F.1:  $M_{e\gamma}$  fits,  $W\gamma$ , electron channel, 15-20 GeV, 8  $\eta^\gamma$  bins.

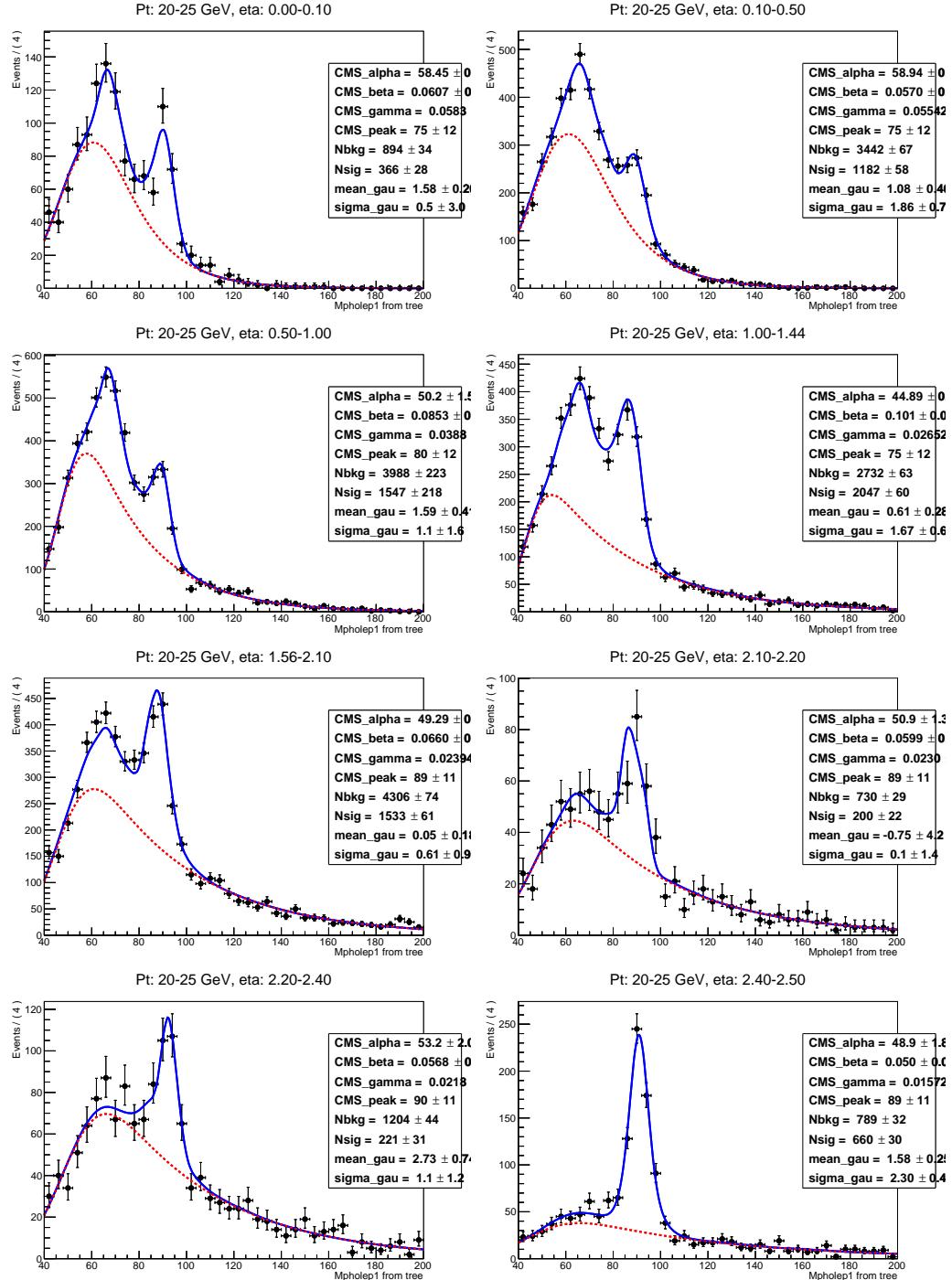


Figure F.2:  $M_{e\gamma}$  fits,  $W\gamma$ , electron channel, 20-25 GeV, 8  $\eta^\gamma$  bins.

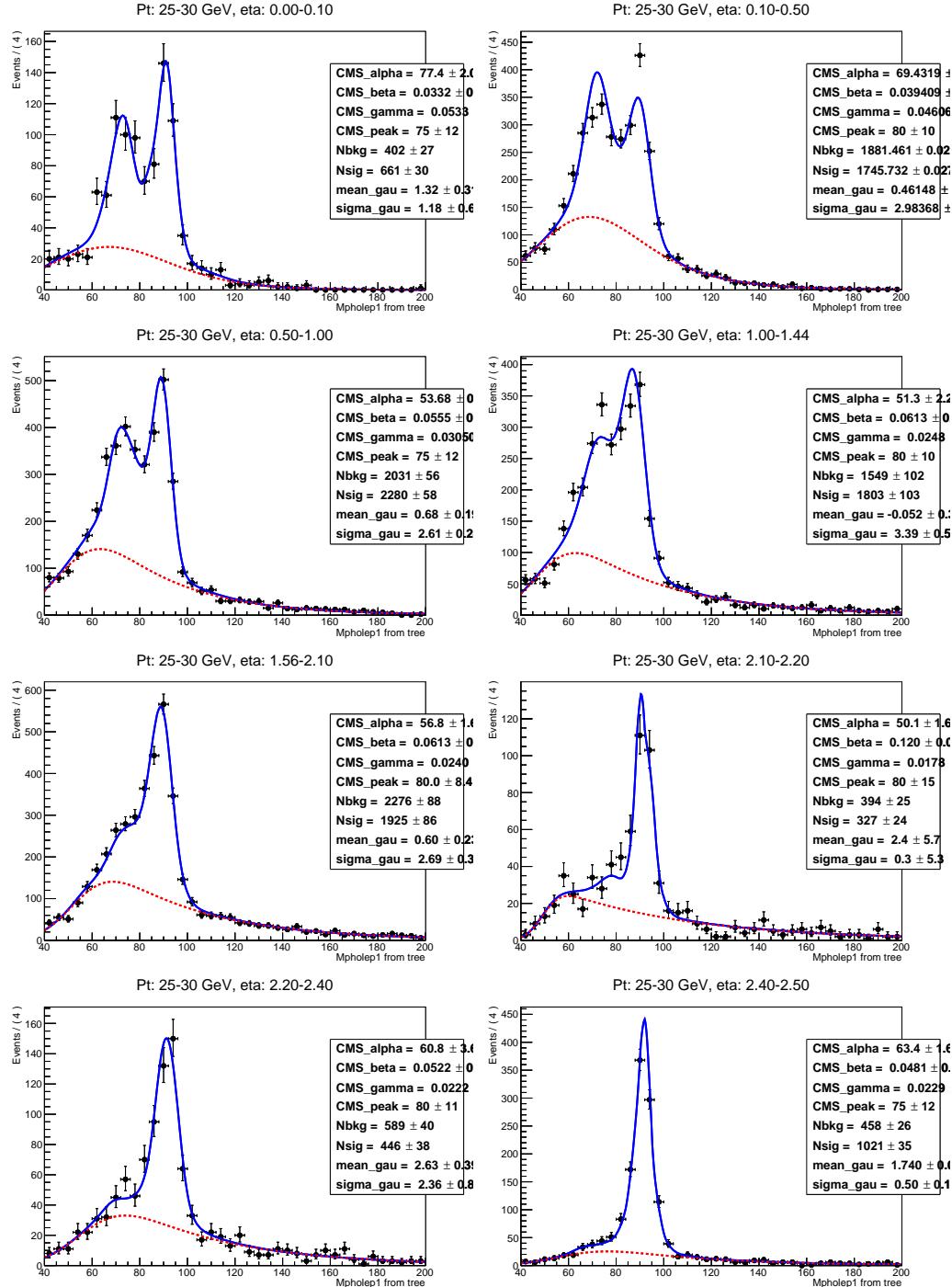


Figure F.3:  $M_{e\gamma}$  fits,  $W\gamma$ , electron channel, 25-30 GeV, 8  $\eta^\gamma$  bins.

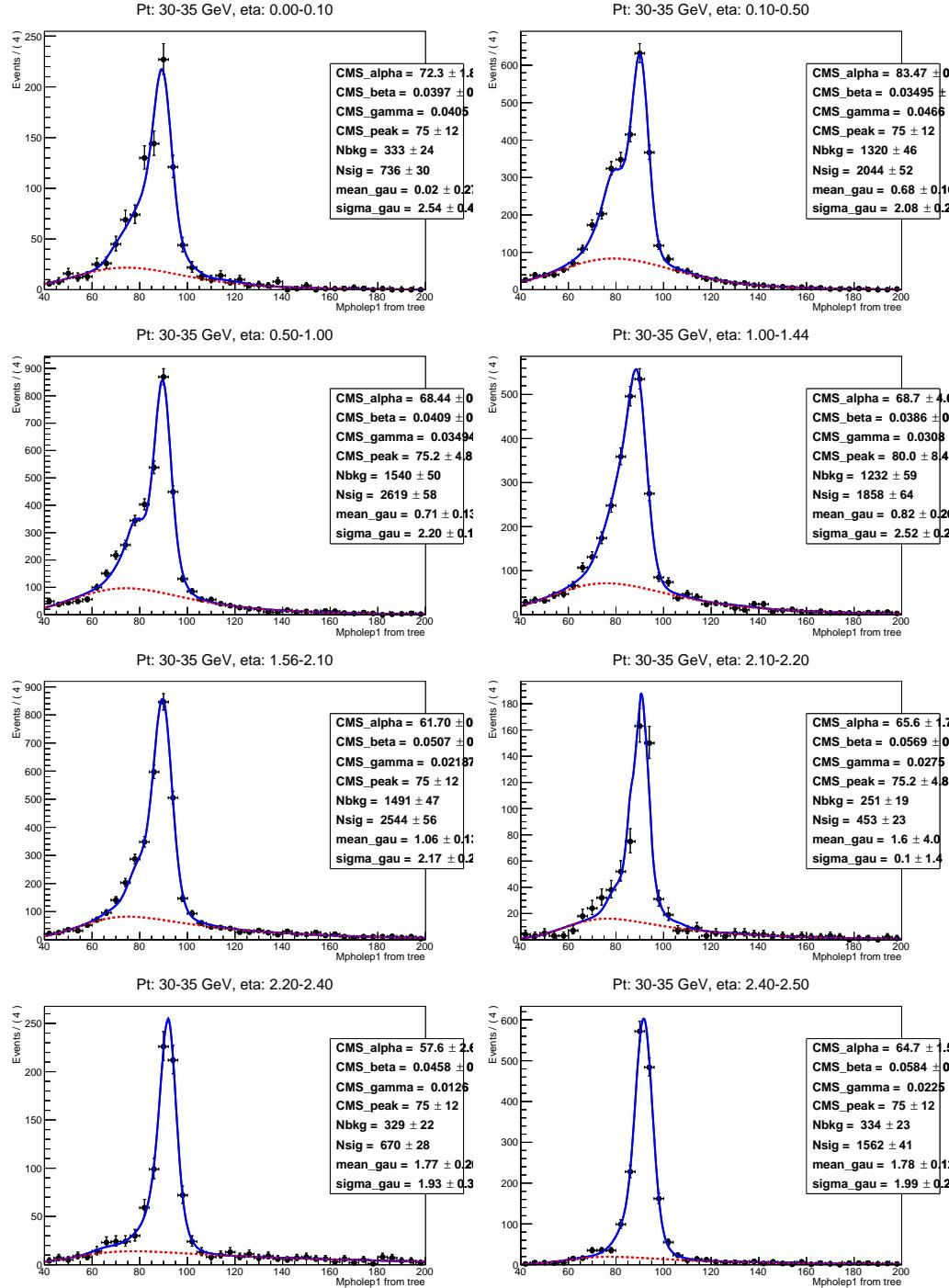


Figure F.4:  $M_{e\gamma}$  fits,  $W\gamma$ , electron channel, 30-35 GeV, 8  $\eta^\gamma$  bins.

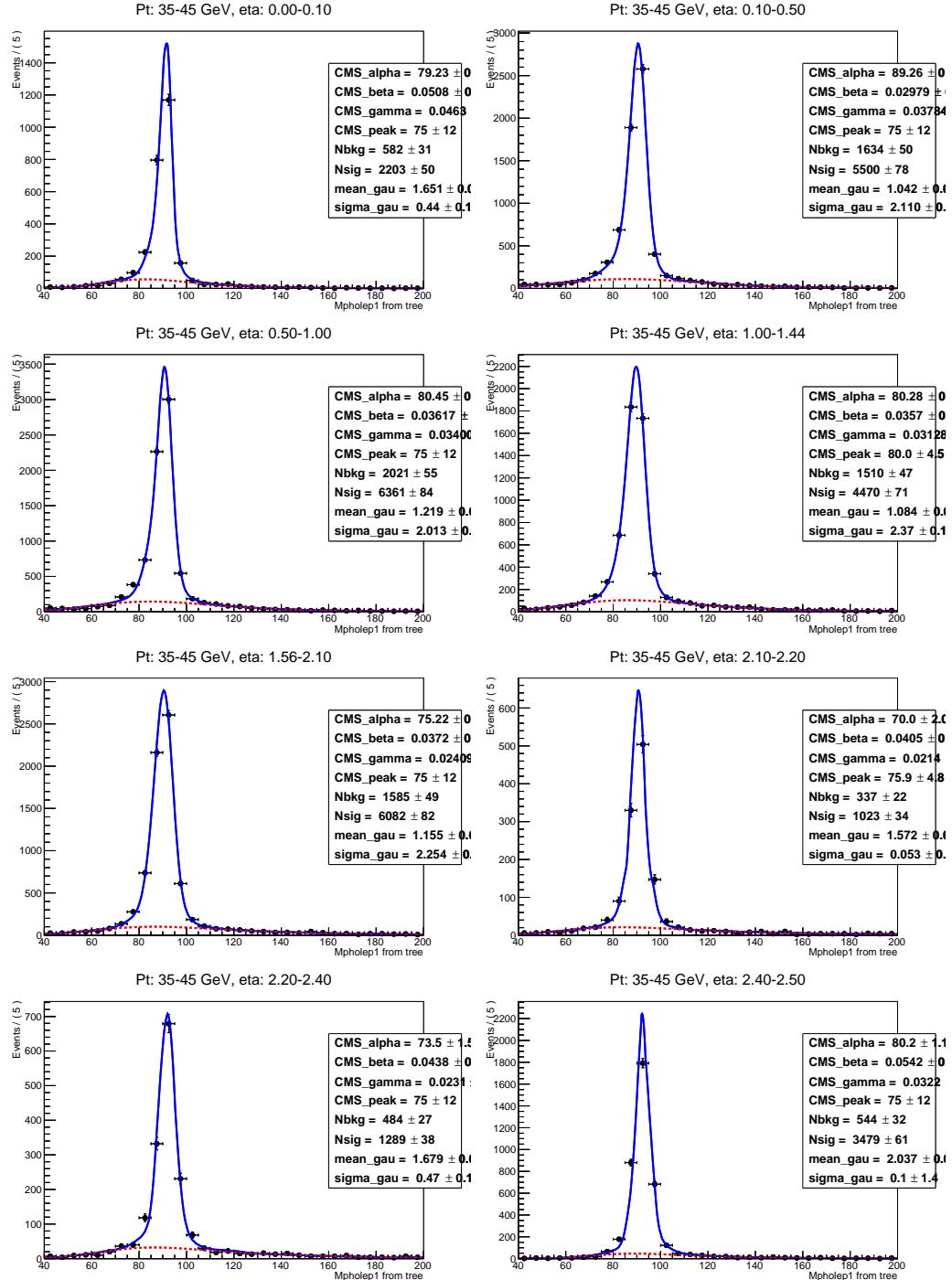


Figure F.5:  $M_{e\gamma}$  fits,  $W\gamma$ , electron channel, 35-45 GeV, 8  $\eta^\gamma$  bins.

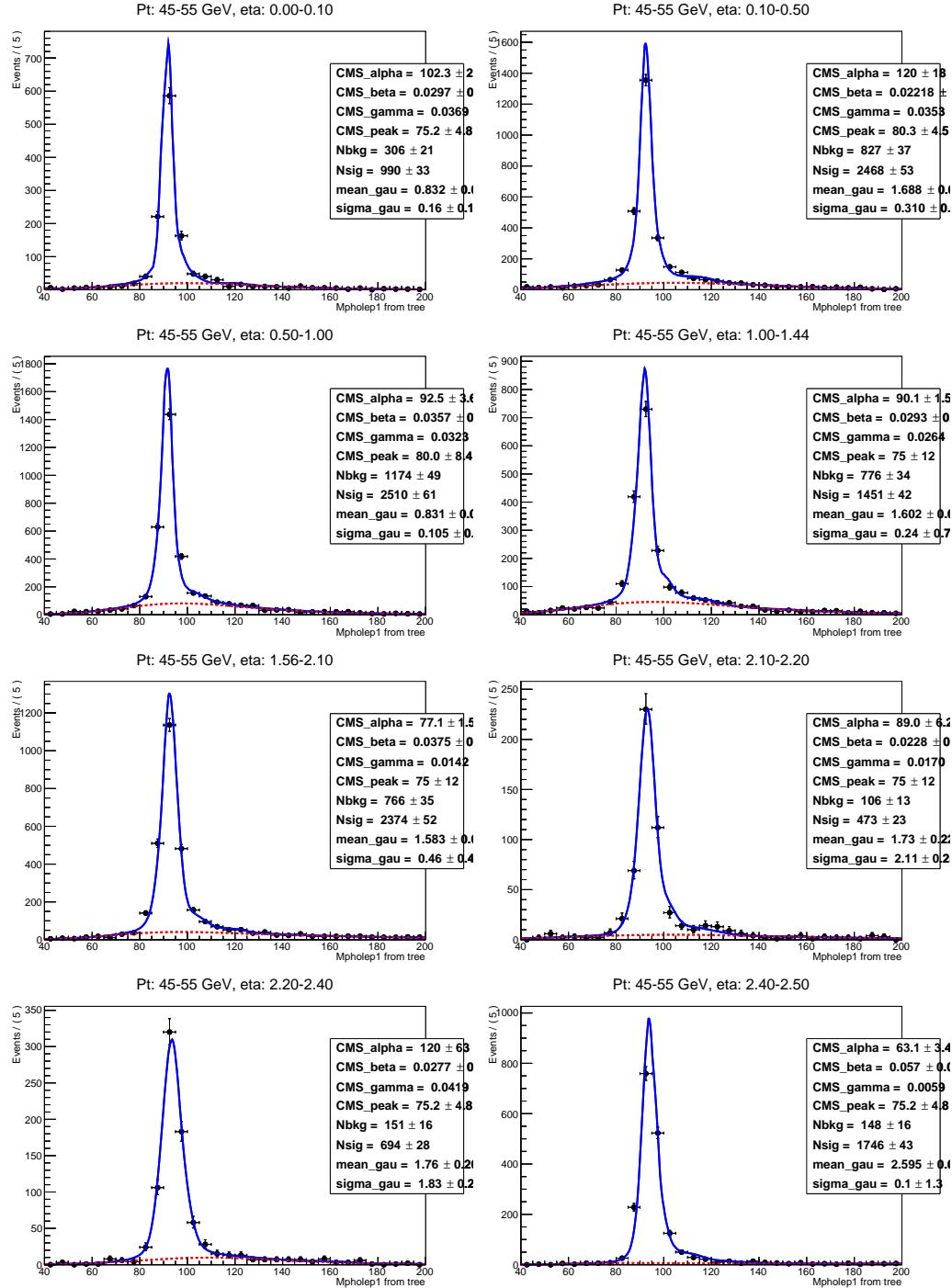


Figure F.6:  $M_{e\gamma}$  fits,  $W\gamma$ , electron channel, 45-55 GeV, 8  $\eta^\gamma$  bins.

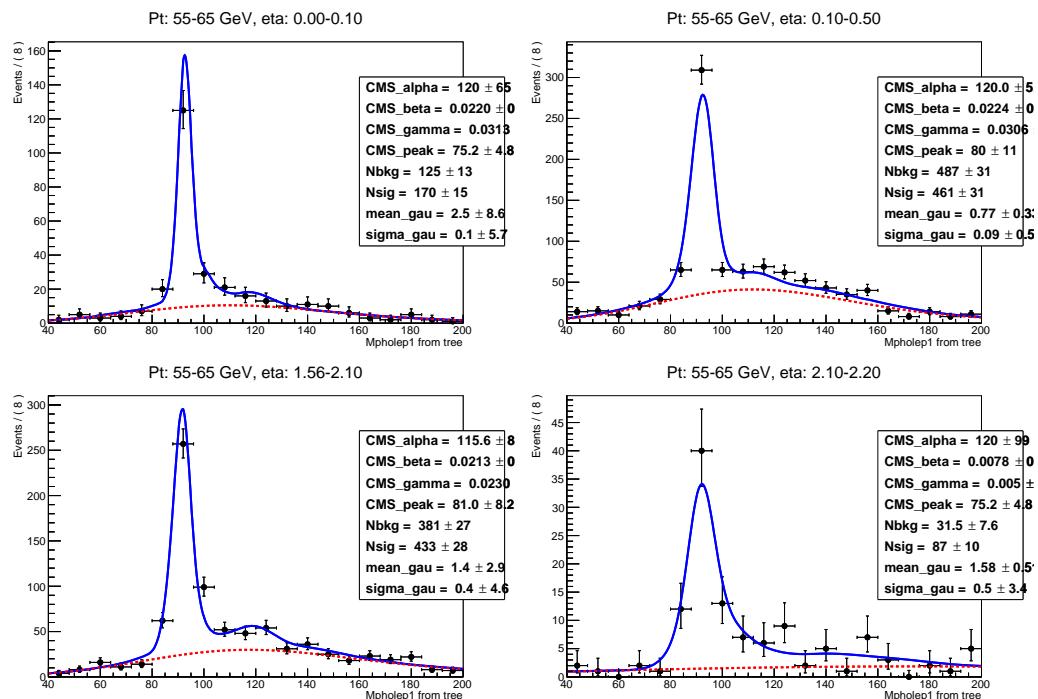


Figure F.7:  $M_{e\gamma}$  fits,  $W\gamma$ , electron channel, 55-65 GeV, 4  $\eta^\gamma$  bins.

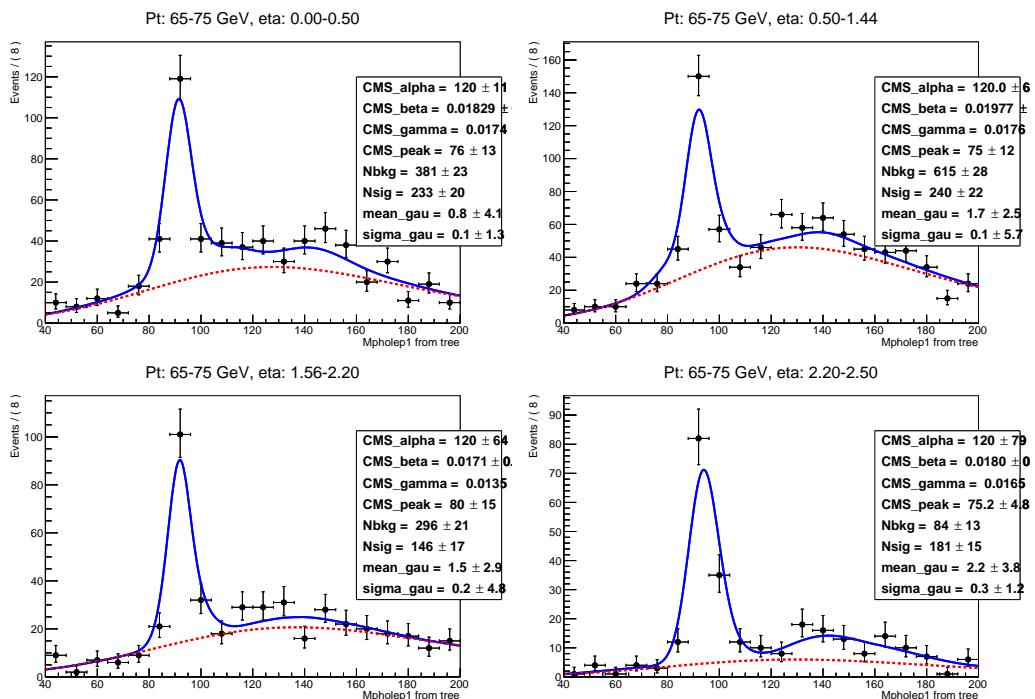


Figure F.8:  $M_{e\gamma}$  fits,  $W\gamma$ , electron channel, 65-75 GeV, 4  $\eta^\gamma$  bins.

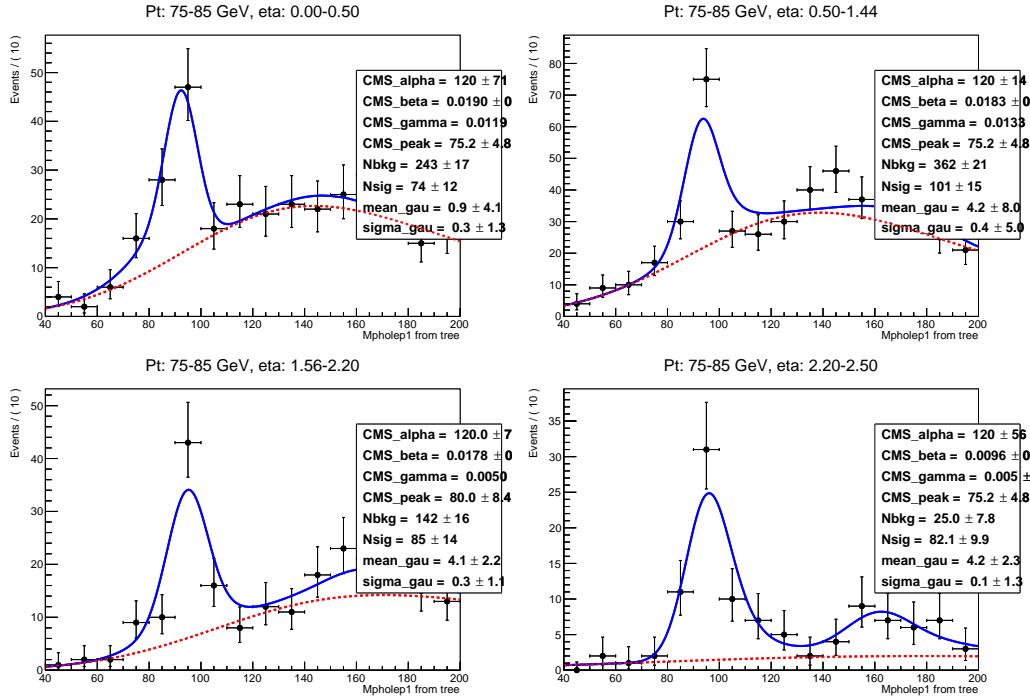


Figure F.9:  $M_{e\gamma}$  fits,  $W\gamma$ , electron channel, 75-85 GeV, 4  $\eta^\gamma$  bins.

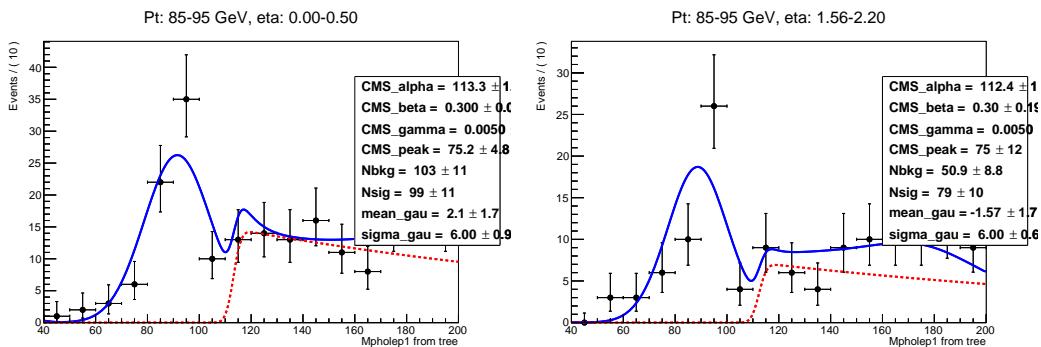


Figure F.10:  $M_{e\gamma}$  fits,  $W\gamma$ , electron channel, 85-95 GeV, 2  $\eta^\gamma$  bins.

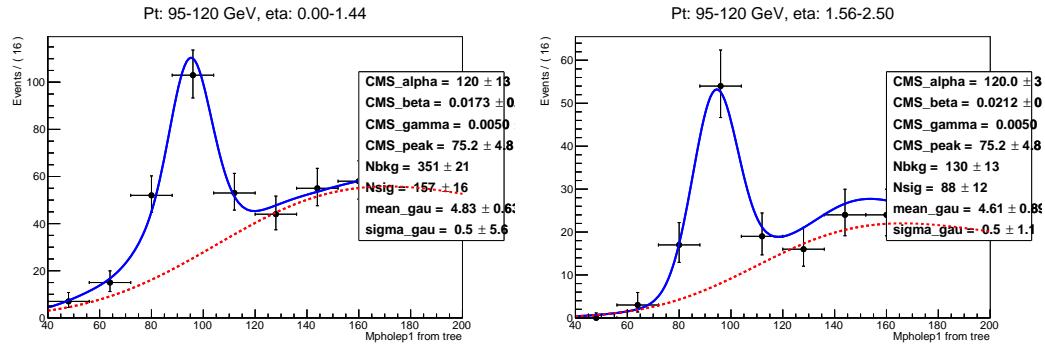


Figure F.11:  $M_{e\gamma}$  fits,  $W\gamma$ , electron channel, 95-120 GeV, 2  $\eta^\gamma$  bins.

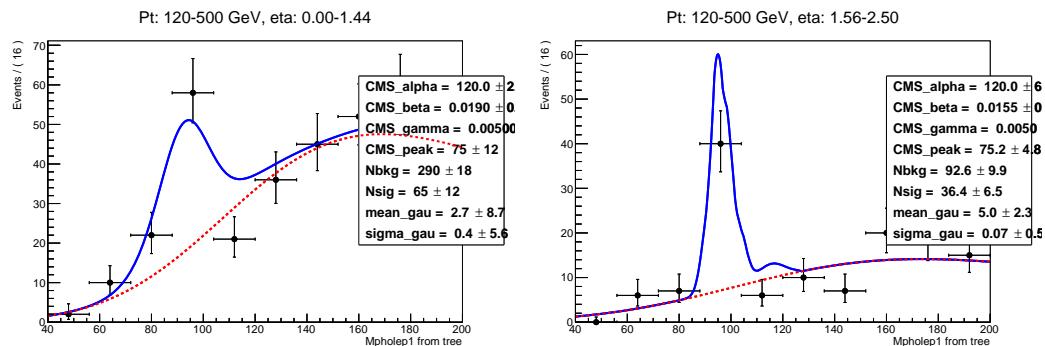


Figure F.12:  $M_{e\gamma}$  fits,  $W\gamma$ , electron channel, 120-500 GeV, 2  $\eta^\gamma$  bins.

## Appendix G

### Fit Plots of $M_{e\gamma}$ without Requirement on $M_T^W$

Fit results of electron-photon invariant mass  $M_{e\gamma}$  without  $M_T^W$  requirement applied for the estimation of the systematic uncertainty for the procedure of the  $e \rightarrow \gamma$  background estimation in the electron channel. The procedure of the background estimation is described in Ch. 5.4.2, the procedure of the estimation of the systematic uncertainty is described in Ch. 5.7.2.

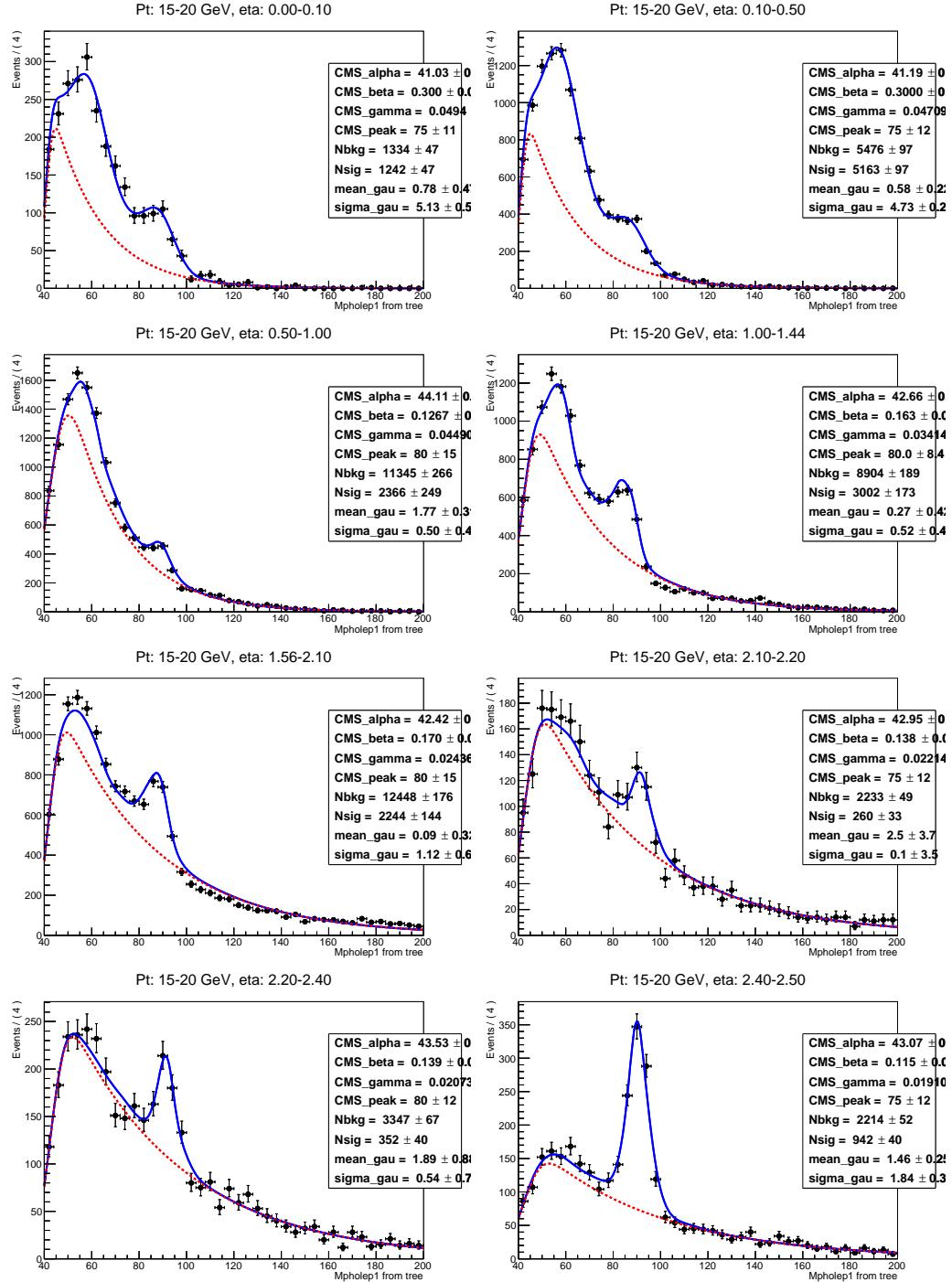


Figure G.1:  $M_{e\gamma}$  fits,  $W\gamma$ , electron channel, underflow bin (15-20 GeV), 8  $\eta^\gamma$  bins.

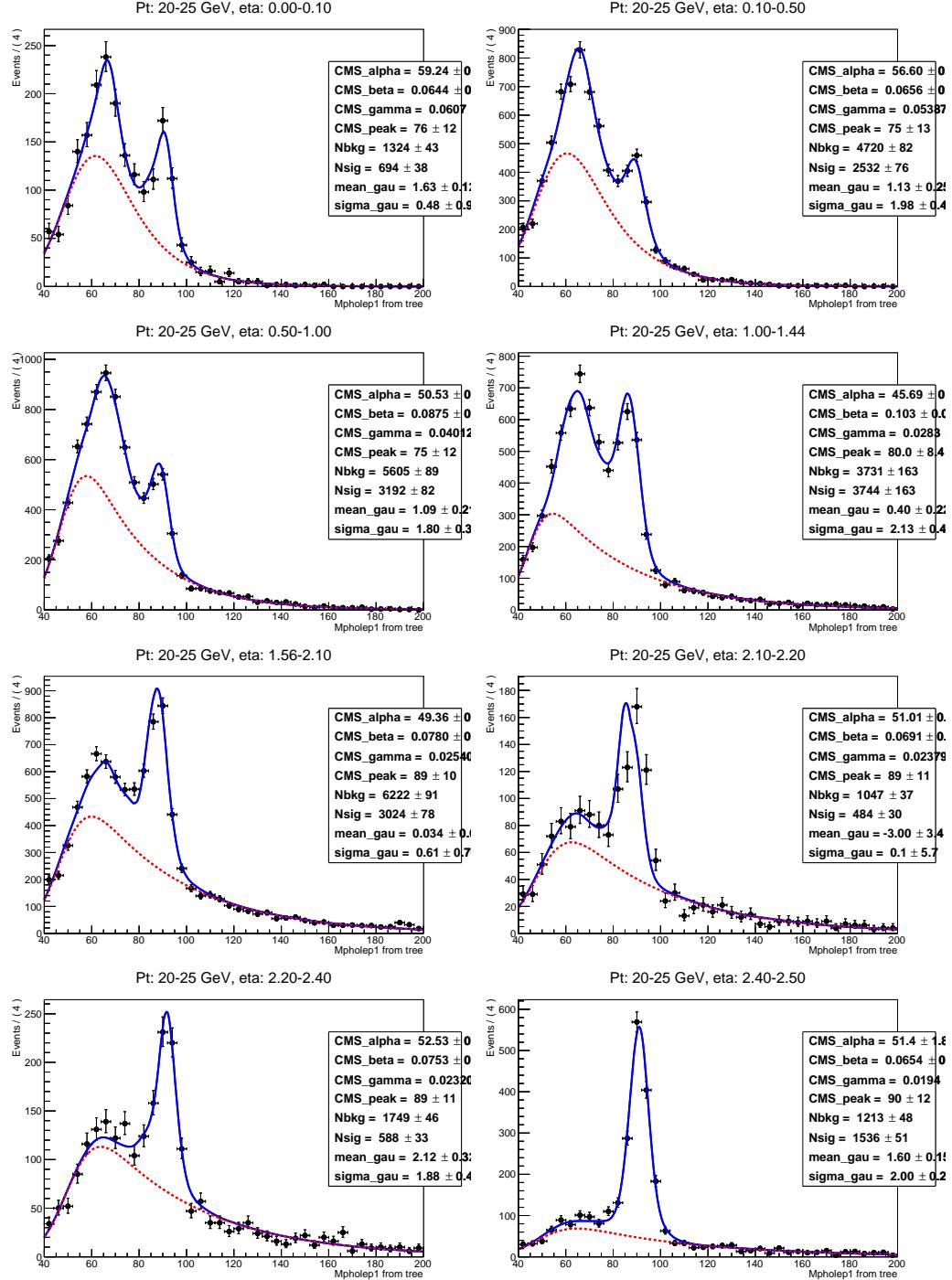


Figure G.2:  $M_{e\gamma}$  fits,  $W\gamma$ , electron channel, underflow bin (20-25 GeV), 8  $\eta^\gamma$  bins.

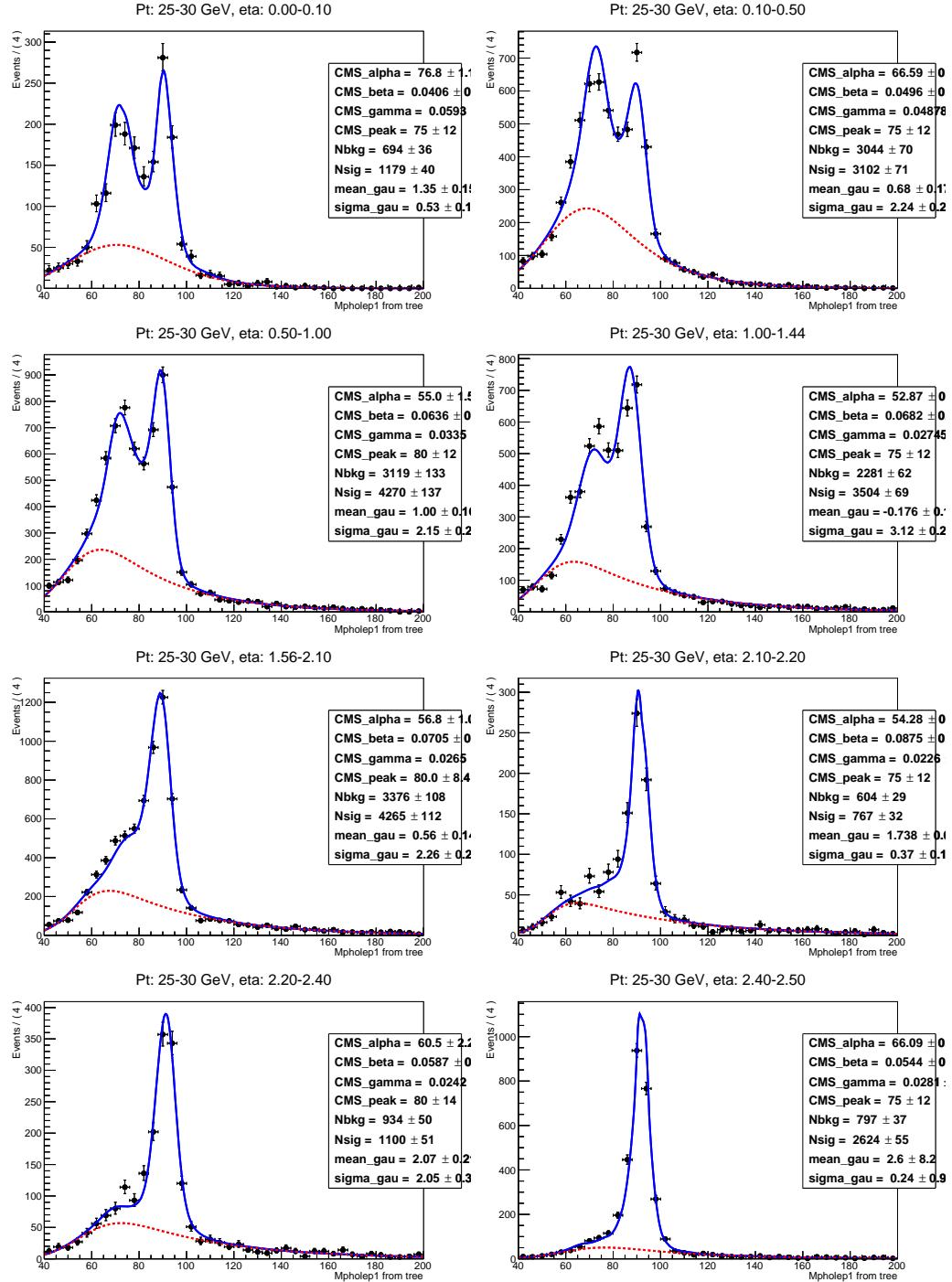


Figure G.3:  $M_{e\gamma}$  fits,  $W\gamma$ , electron channel, underflow bin (25-30 GeV), 8  $\eta^\gamma$  bins.

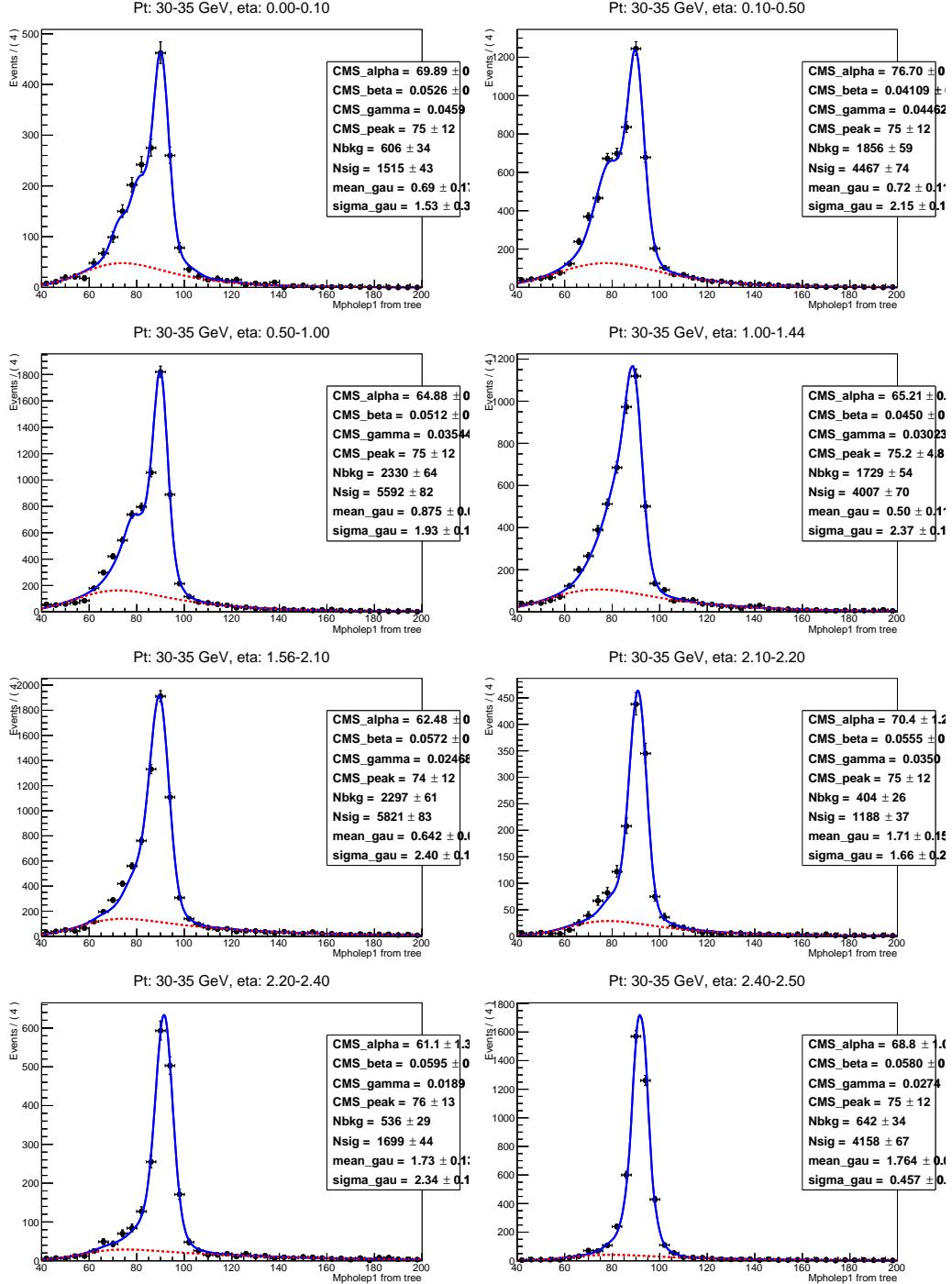


Figure G.4:  $M_{e\gamma}$  fits,  $W\gamma$ , electron channel, underflow bin (30-35 GeV), 8  $\eta^\gamma$  bins.

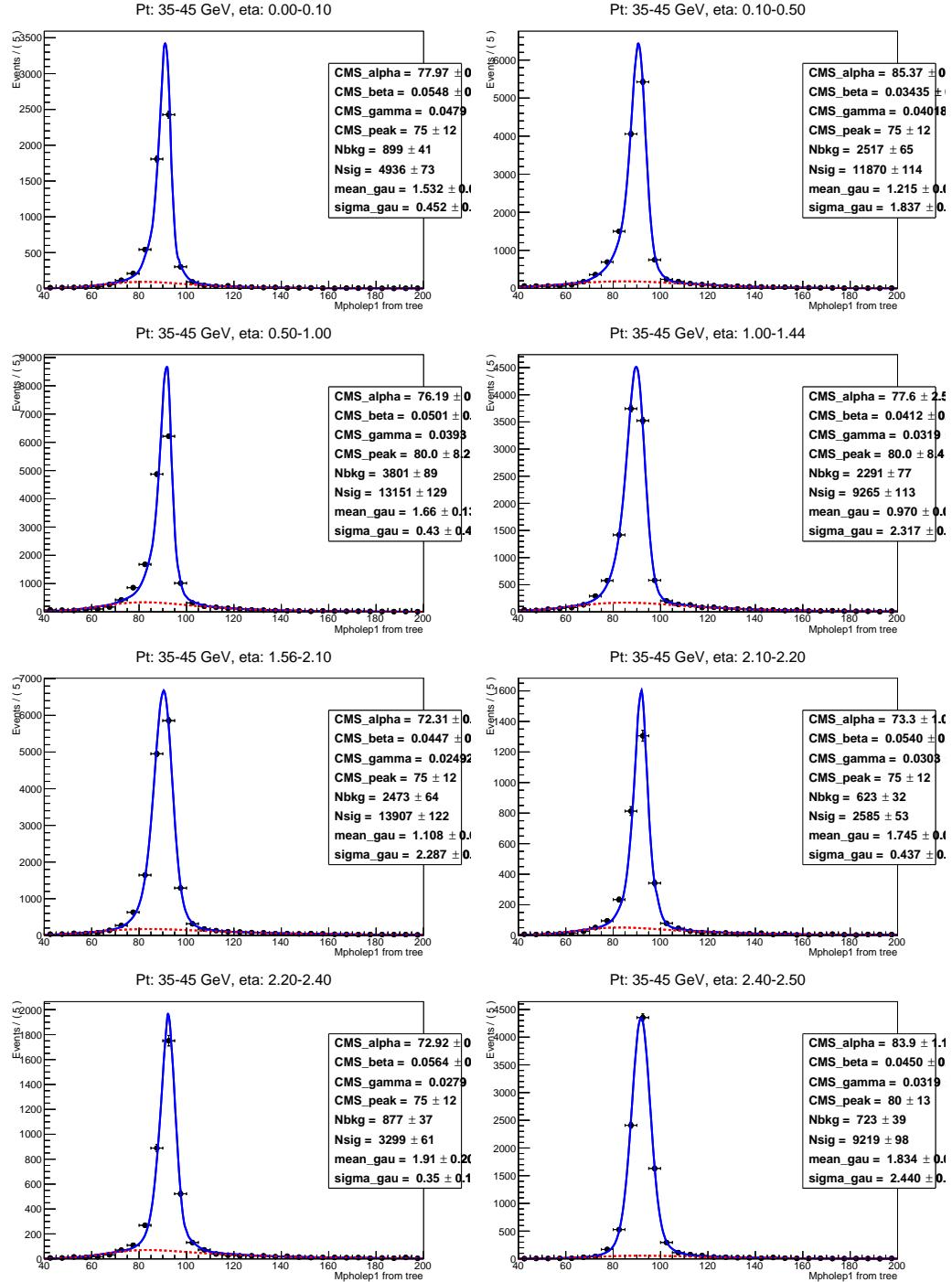


Figure G.5:  $M_{e\gamma}$  fits,  $W\gamma$ , electron channel, underflow bin (35-45 GeV), 8  $\eta^\gamma$  bins.

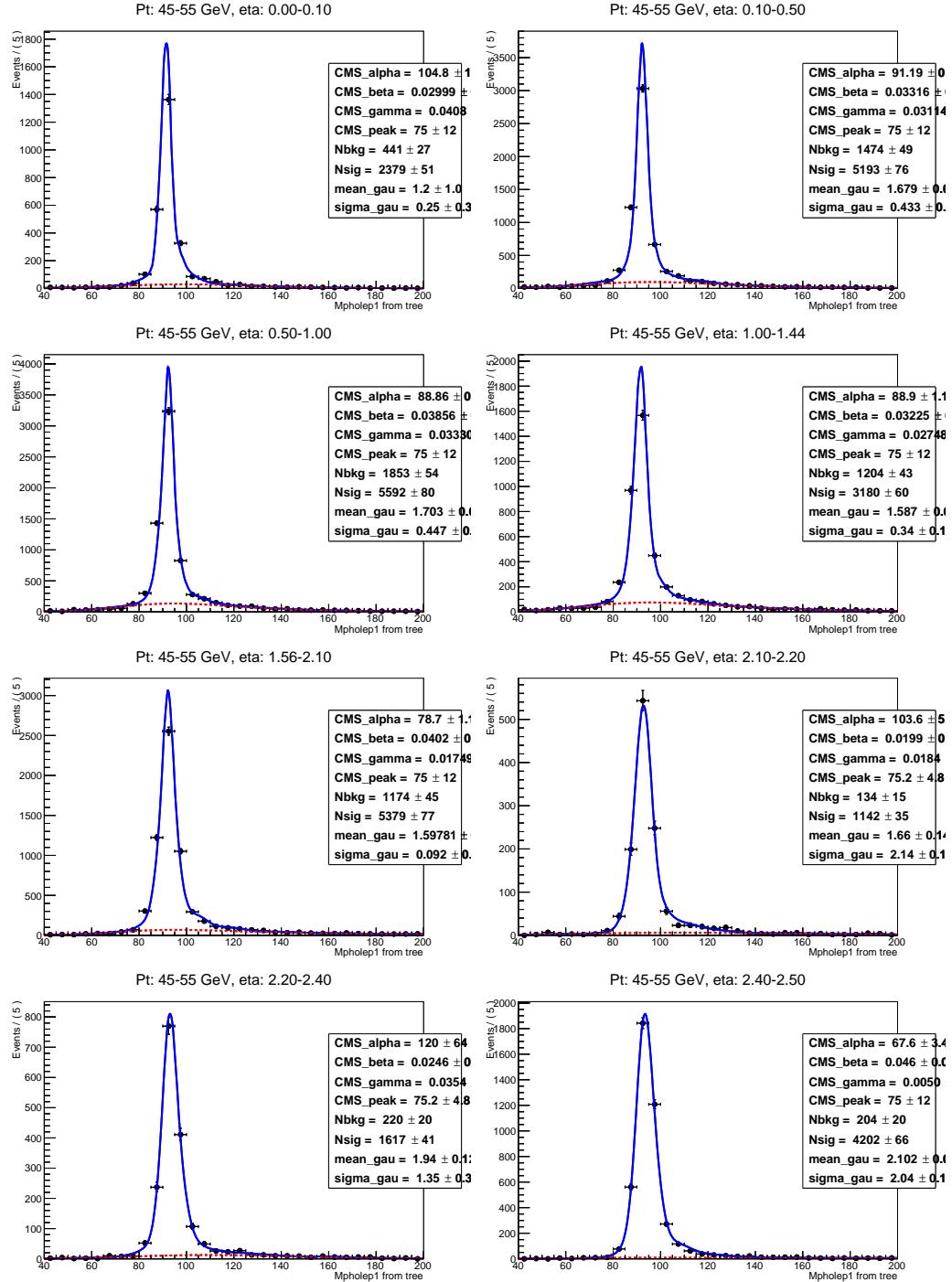


Figure G.6:  $M_{e\gamma}$  fits,  $W\gamma$ , electron channel, underflow bin (45-55 GeV), 8  $\eta^\gamma$  bins.

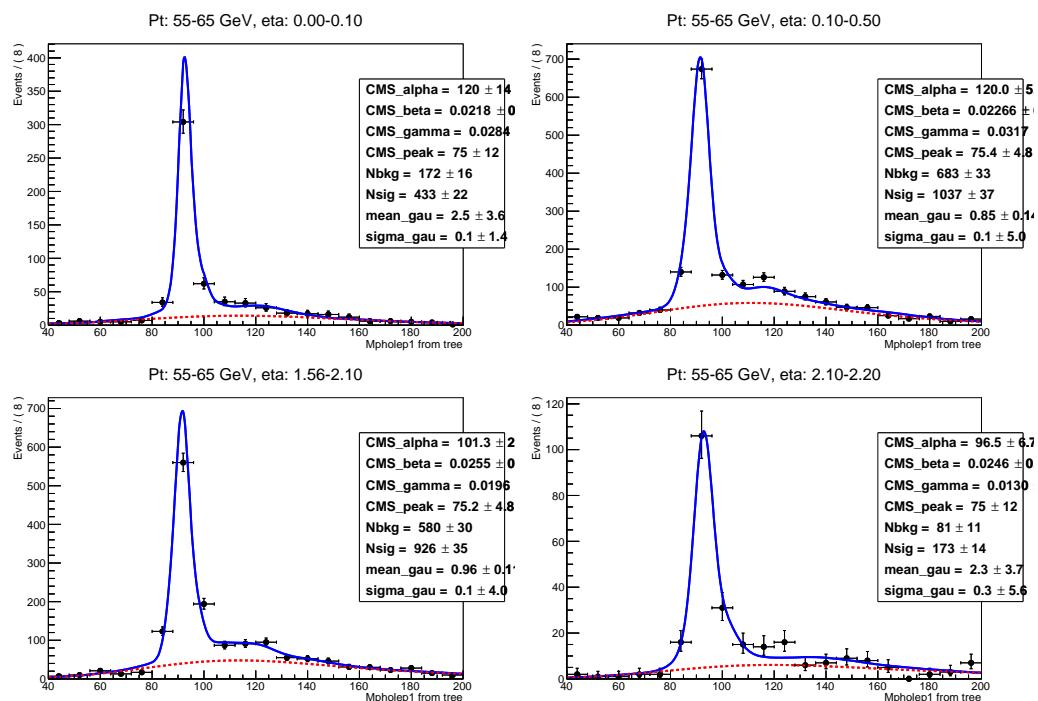


Figure G.7:  $M_{e\gamma}$  fits,  $W\gamma$ , electron channel, underflow bin (55-65 GeV), 4  $\eta^\gamma$  bins.

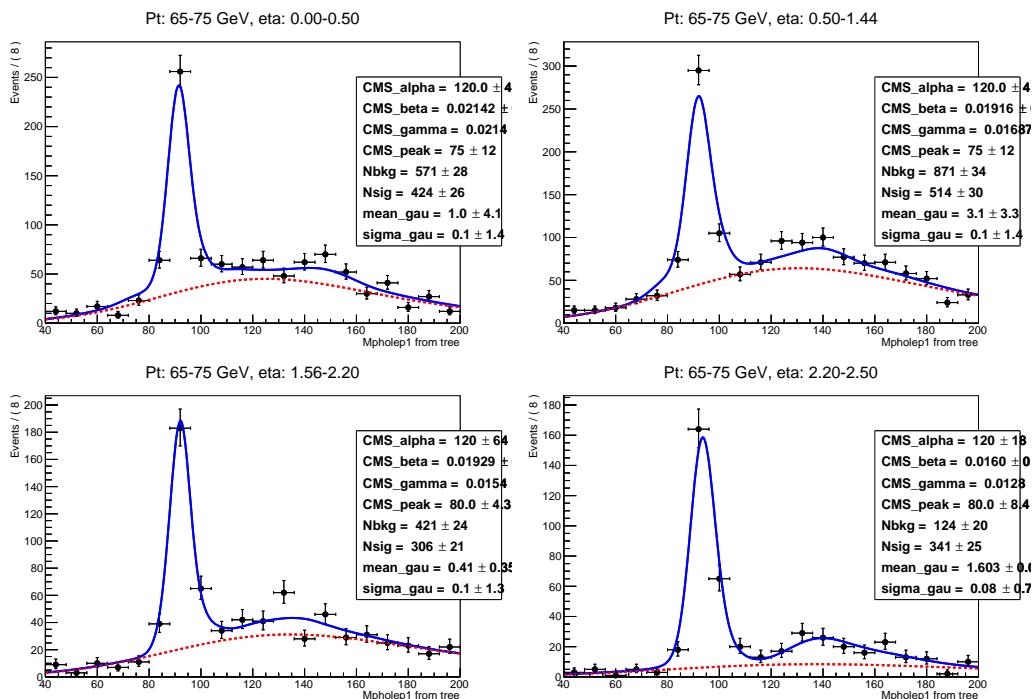


Figure G.8:  $M_{e\gamma}$  fits,  $W\gamma$ , electron channel, underflow bin (65-75 GeV), 4  $\eta^\gamma$  bins.

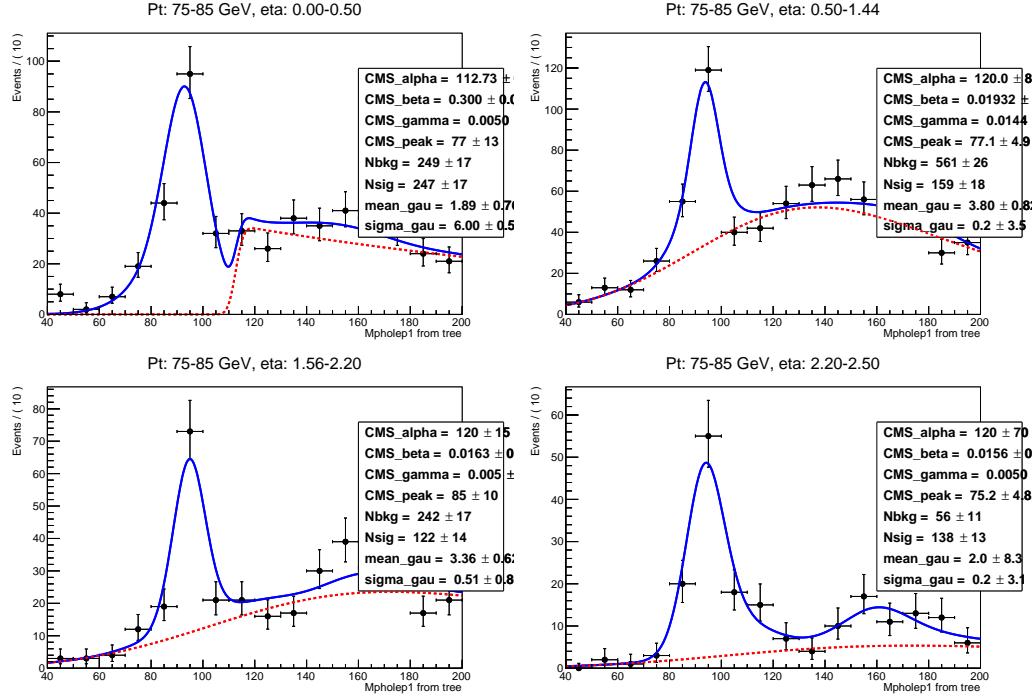


Figure G.9:  $M_{e\gamma}$  fits,  $W\gamma$ , electron channel, underflow bin (75-85 GeV), 4  $\eta^\gamma$  bins.

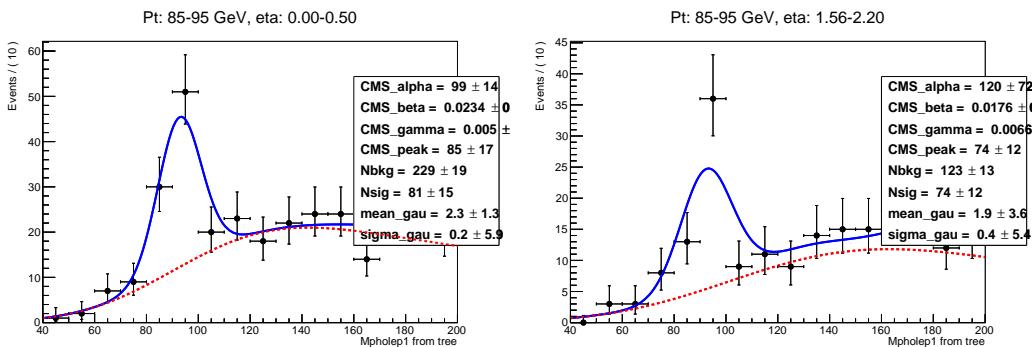


Figure G.10:  $M_{e\gamma}$  fits,  $W\gamma$ , electron channel, underflow bin (85-95 GeV), 2  $\eta^\gamma$  bins.

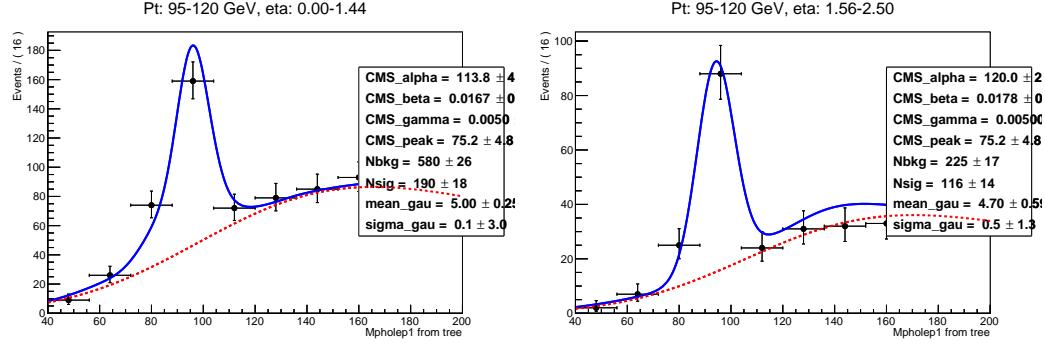


Figure G.11:  $M_{e\gamma}$  fits,  $W\gamma$ , electron channel, underflow bin (95-120 GeV), 2  $\eta^\gamma$  bins.

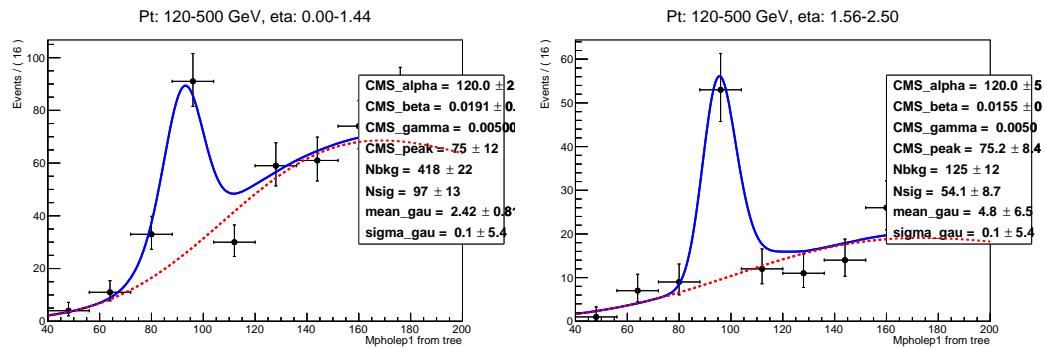


Figure G.12:  $M_{e\gamma}$  fits,  $W\gamma$ , electron channel, underflow bin (120-500 GeV), 2  $\eta^\gamma$  bins.

# Appendix H

## Tables for $e \rightarrow \gamma$ Background Estimation

This appendix presents results of  $e \rightarrow \gamma$  background estimation. Tab. H.1-H.2 show results of  $e \rightarrow \gamma$  background estimation when fits are performed on data with all selection criteria applied except Z-mass window requirement. These results are used for the background subtraction. Tables H.3-H.4 show results of  $e \rightarrow \gamma$  background estimation when fits are performed on data without  $M_T^W$  requirement. These results are used for the estimation of the systematic uncertainties.

In each table, the first column is a  $P_T^\gamma$  bin, the second column is yields of weighted DY+jets MC in conditions of full nominal selection, the third column is yields of  $e \rightarrow \gamma$ -enriched dataset with or without  $M_T^W$  requirement. These yields are extracted from fit. The fourth column is yields of weighted DY+jets MC in conditions of  $e \rightarrow \gamma$ -enriched selection with or without  $M_T^W$  requirement, consistently with the dataset. The fifth column is the scale which is computed as the yield in the third column divided over the yield in the fourth column. The sixth column is the estimated  $e \rightarrow \gamma$  background in the nominally selected dataset.

The value is computed as the yield in the second column multiplied by the scale. The values in the sixth column as used for the background subtraction (Tab. H.1-H.2) or estimation for the systematic uncertainty (Tab. H.3-H.4). The seventh column is yields of the weighted signal MC ( $W\gamma \rightarrow e\nu\gamma$ ) in the nominally selected conditions, it is quoted for comparison purposes, to estimate how significant is  $e \rightarrow \gamma$  background compared to the signal.

Table H.1:  $e \rightarrow \gamma$  background. Barrel. Fits with  $M_T^W$  requirement applied

bin lims	DY+jets nom. sel.	Data $e \rightarrow \gamma$ enr.	DY+jets $e \rightarrow \gamma$ enr.	scale	$e \rightarrow \gamma$ yield	SigMC ( $W\gamma \rightarrow e\nu\gamma$ )
15-20	1917±63	6395±149	3300±83	1.94±0.0667	3715±177	5857±110
20-25	1175±49	5141±235	2987±79	1.72±0.091	2023±137	2868±77
25-30	543±33	6489±122	3418±84	1.9±0.0591	1030±71	1411±54
30-35	166±18	7257±105	4215±94	1.72±0.0461	286±33	915±43
35-45	134±16	18534±144	11597±158	1.6±0.0251	215±27	1247±51
45-55	186±20	7417±97	4134±94	1.79±0.0473	335±37	820±41
55-65	130±16	1426±48	685±38	2.08±0.136	272±39	654±37
65-75	86±13	473±29	286±24	1.65±0.177	143±27	440±30
75-85	42±9	174±19	165±19	1.05±0.168	45±12	295±25
85-95	20±6	140±14	66±12	2.1±0.445	42±16	234±22
95-120	38±9	156±16	94±14	1.65±0.307	63±19	318±26
120-500	36±9	64±11	67±12	0.957±0.246	34±12	429±30

Table H.2:  $e \rightarrow \gamma$  background. Endcap. Fits with  $M_T^W$  requirement applied

bin lims	DY+jets nom. sel.	Data $e \rightarrow \gamma$ enr.	DY+jets $e \rightarrow \gamma$ enr.	scale	$e \rightarrow \gamma$ yield	SigMC ( $W\gamma \rightarrow e\nu\gamma$ )
15-20	458±31	2004±159	1805±61	1.11±0.096	508±55	2253±68
20-25	402±29	2613±77	2432±72	1.07±0.0451	432±36	1177±49
25-30	216±21	3719±102	3527±85	1.05±0.0388	228±23	574±34
30-35	123±16	5228±78	5374±109	0.973±0.0247	120±16	445±31
35-45	173±19	11873±114	12355±164	0.961±0.0158	166±18	638±37
45-55	223±21	5286±75	4212±94	1.25±0.0334	280±28	287±24
55-65	182±19	1010±38	813±41	1.24±0.0787	226±28	237±22
65-75	82±13	327±22	299±25	1.09±0.121	89±17	194±21
75-85	68±13	167±17	184±21	0.907±0.141	61±15	137±18
85-95	40±10	107±22	82±14	1.29±0.35	52±19	81±14
95-120	48±11	88±11	97±15	0.901±0.188	43±13	166±20
120-500	22±7	36±6	54±11	0.662±0.184	15±6	145±18

Table H.3:  $e \rightarrow \gamma$  background. Barrel. Fits without  $M_T^W$  requirement applied

bin lims	DY+jets nom. sel.	Data $e \rightarrow \gamma$ enr.	DY+jets $e \rightarrow \gamma$ enr.	scale	$e \rightarrow \gamma$ yield	SigMC ( $W\gamma \rightarrow e\nu\gamma$ )
15-20	1917±63	11771±321	7491±125	1.57±0.0503	3012±138	5857±110
20-25	1175±49	10162±201	6933±120	1.47±0.0387	1722±85	2868±77
25-30	543±33	12055±173	7526±125	1.6±0.0353	869±57	1411±54
30-35	166±18	15580±138	9753±144	1.6±0.0275	265±30	915±43
35-45	134±16	39220±218	27310±242	1.44±0.015	193±24	1247±51
45-55	186±20	16343±135	10411±149	1.57±0.0261	293±31	820±41
55-65	130±16	3256±65	1722±60	1.89±0.0765	247±33	654±37
65-75	86±13	938±39	600±36	1.56±0.115	135±23	440±30
75-85	42±9	405±24	274±24	1.48±0.162	63±16	295±25
85-95	20±6	156±19	125±16	1.25±0.226	25±9	234±22
95-120	38±9	189±18	155±18	1.22±0.188	46±13	318±26
120-500	36±9	96±13	89±14	1.08±0.226	38±12	429±30

Table H.4:  $e \rightarrow \gamma$  background. Endcap. Fits without  $M_T^W$  requirement applied

bin lims	DY+jets nom. sel.	Data $e \rightarrow \gamma$ enr.	DY+jets $e \rightarrow \gamma$ enr.	scale	$e \rightarrow \gamma$ yield	SigMC ( $W\gamma \rightarrow e\nu\gamma$ )
15-20	458±31	3798±158	4004±92	0.948±0.0452	434±35	2253±68
20-25	402±29	5631±103	5586±109	1.01±0.027	405±31	1177±49
25-30	216±21	8755±138	8528±132	1.03±0.0228	222±22	574±34
30-35	123±16	12865±120	13762±175	0.935±0.0148	115±15	445±31
35-45	173±19	29009±176	29847±254	0.972±0.0102	168±19	638±37
45-55	223±21	12339±114	10099±145	1.22±0.021	273±27	287±24
55-65	182±19	2012±50	1700±59	1.18±0.0511	215±25	237±22
65-75	82±13	646±32	606±36	1.07±0.0842	87±16	194±21
75-85	68±13	260±19	316±28	0.823±0.0957	56±12	137±18
85-95	40±10	139±15	148±19	0.944±0.16	38±11	81±14
95-120	48±11	115±13	185±21	0.626±0.104	30±8	166±20
120-500	22±7	54±8	85±14	0.632±0.149	14±5	145±18

## Appendix I

### Data vs MC Plots of $M_{e\gamma}$

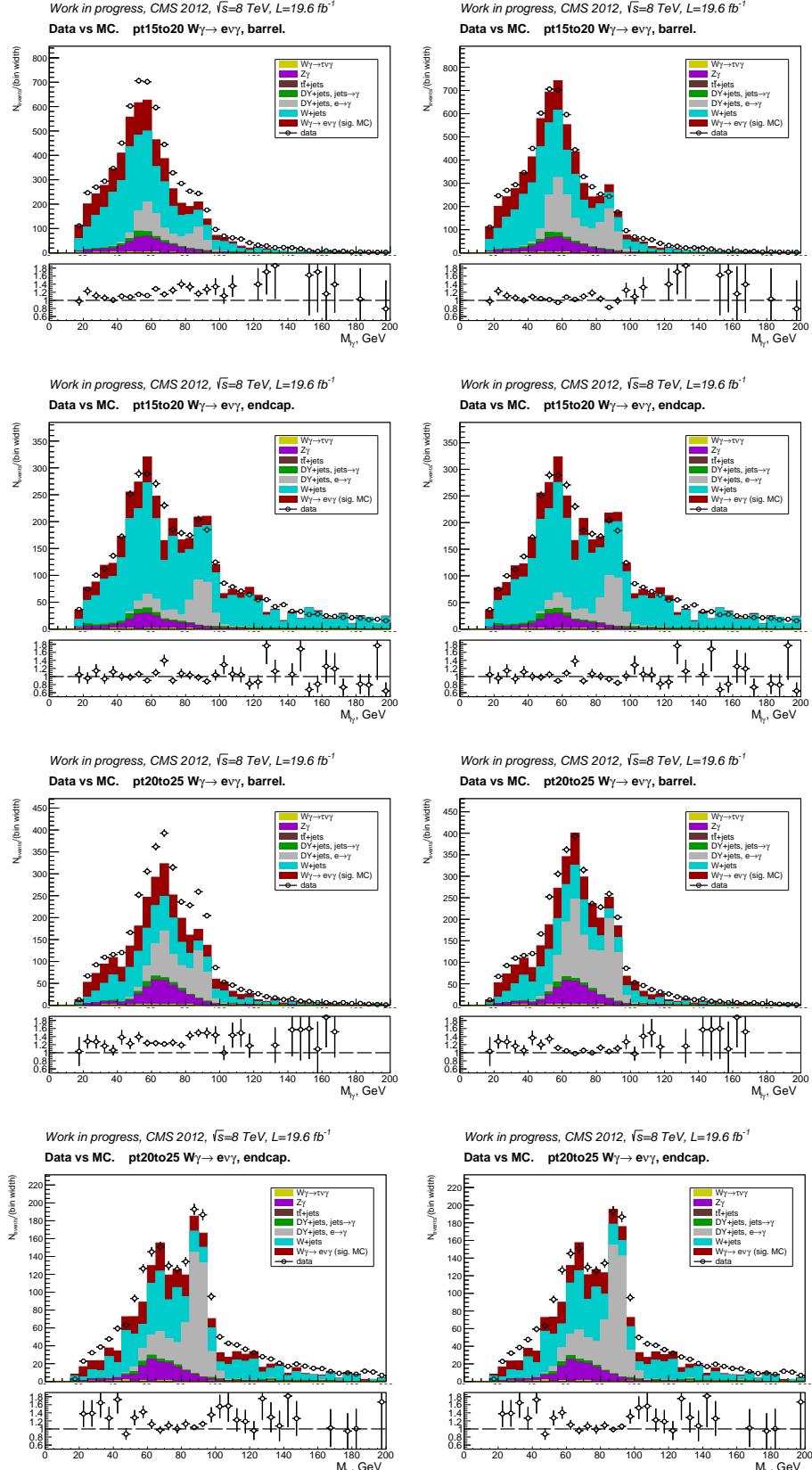


Figure I.1:  $M_{e\gamma}$  distribution, data vs MC. Bins 15 – 20 – 25 GeV. Left: all MC samples are normalized to luminosity of data, PU weight and SFs, right: DY+jets( $e \rightarrow \gamma$ ) also normalized to  $e \rightarrow \gamma$  background estimates.

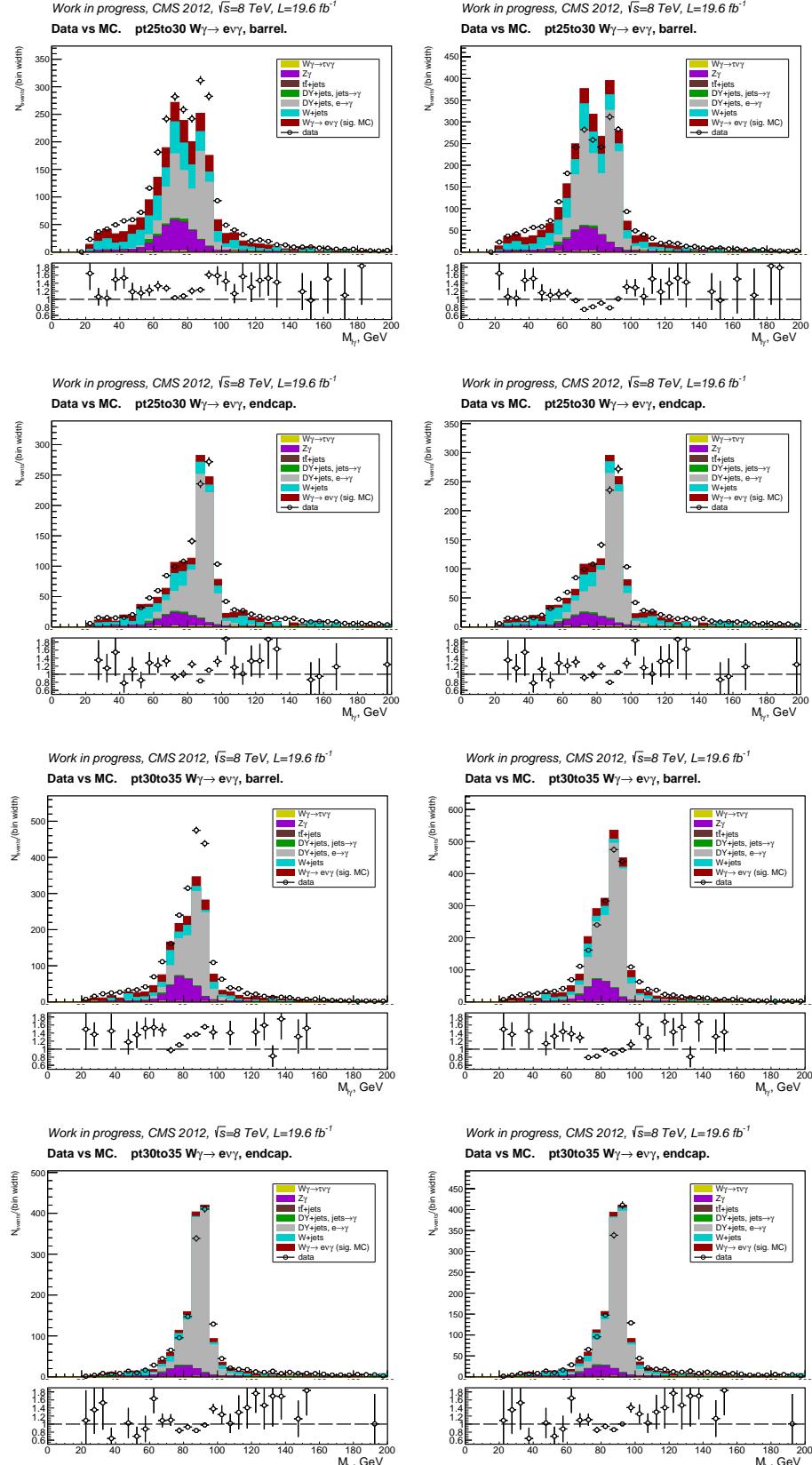


Figure I.2:  $M_{e\gamma}$  distribution, data vs MC. Bins 25 – 30 – 35 GeV. Left: all MC samples are normalized to luminosity of data, PU weight and SFs, right: DY+jets( $e \rightarrow \gamma$ ) also normalized to  $e \rightarrow \gamma$  background estimates.

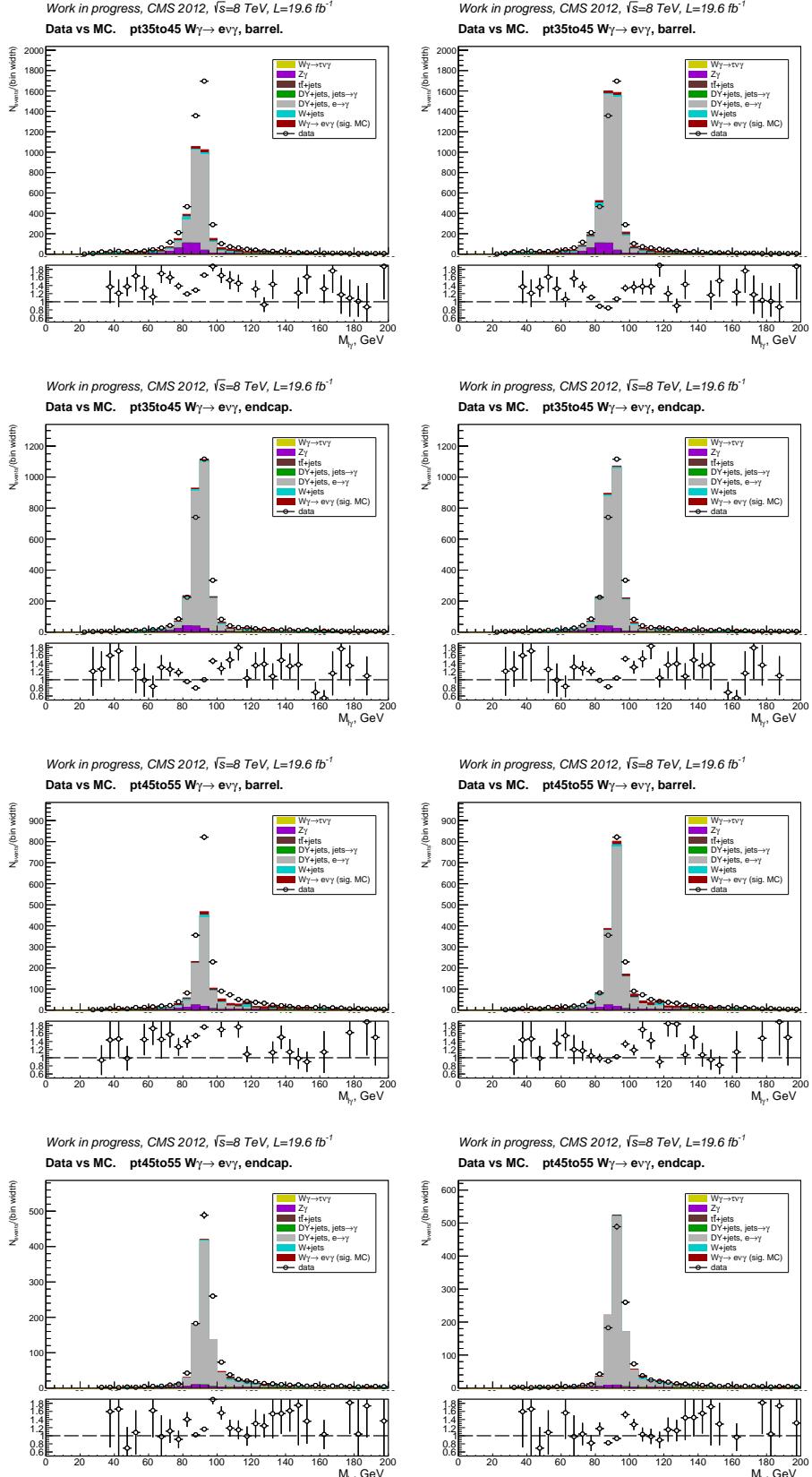


Figure I.3:  $M_{e\gamma}$  distribution, data vs MC. Bins 35 – 45 – 55 GeV. Left: all MC samples are normalized to luminosity of data, PU weight and SFs, right: DY+jets( $e \rightarrow \gamma$ ) also normalized to  $e \rightarrow \gamma$  background estimates.

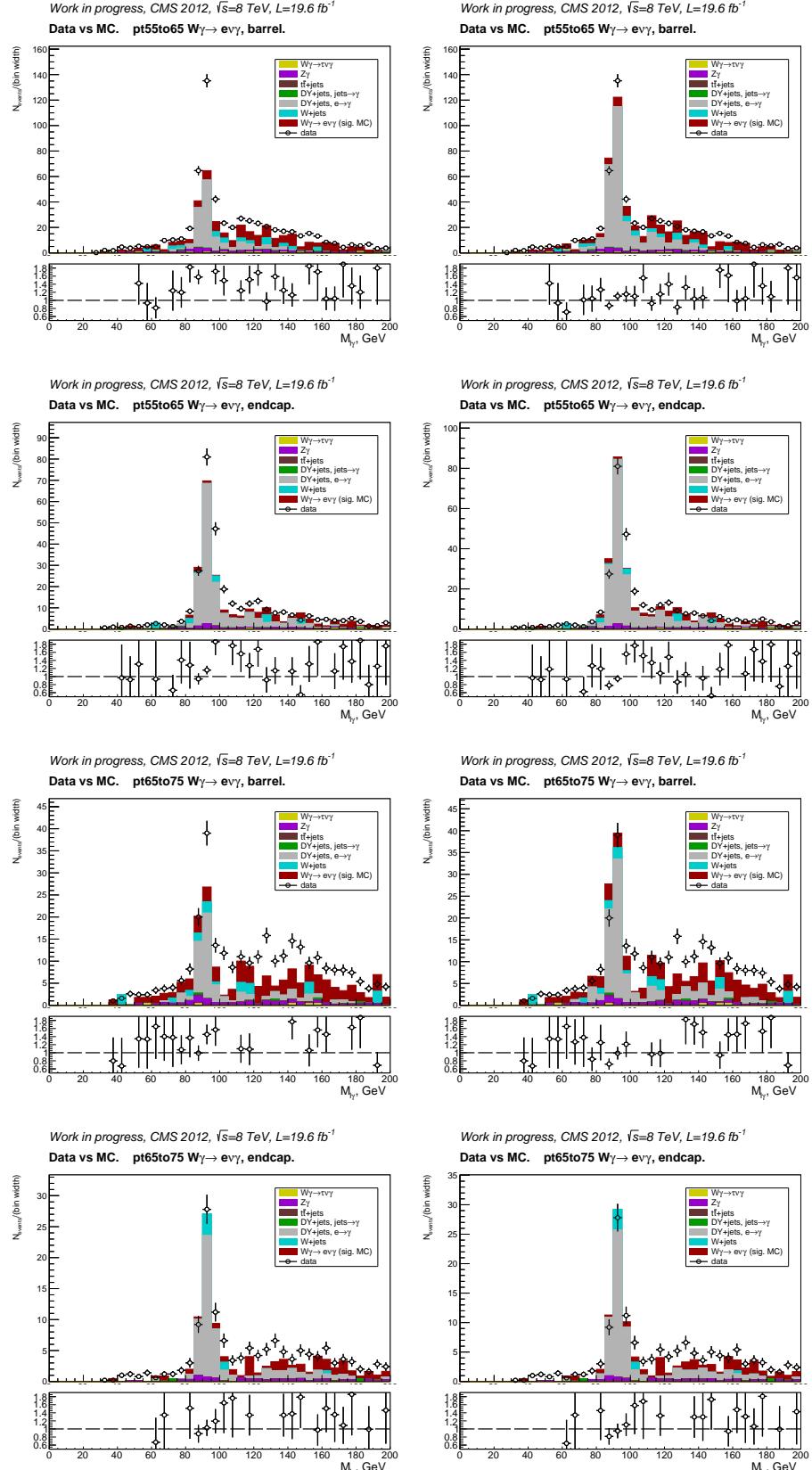


Figure I.4:  $M_{e\gamma}$  distribution, data vs MC. Bins 55 – 65 – 75 GeV. Left: all MC samples are normalized to luminosity of data, PU weight and SFs, right: DY+jets( $e \rightarrow \gamma$ ) also normalized to  $e \rightarrow \gamma$  background estimates.

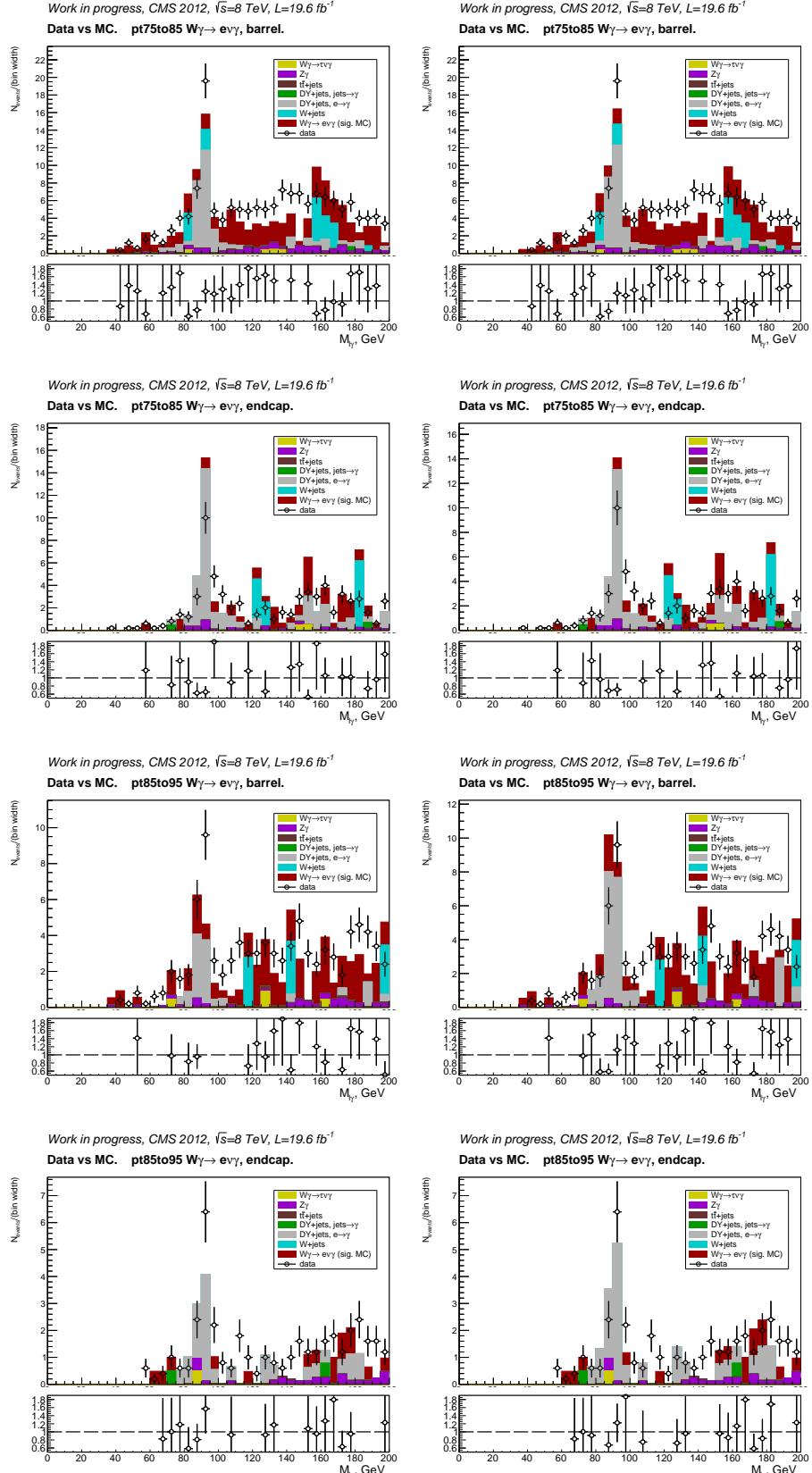


Figure I.5:  $M_{e\gamma}$  distribution, data vs MC. Bins 75 – 85 – 95 GeV. Left: all MC samples are normalized to luminosity of data, PU weight and SFs, right: DY+jets( $e \rightarrow \gamma$ ) also normalized to  $e \rightarrow \gamma$  background estimates.

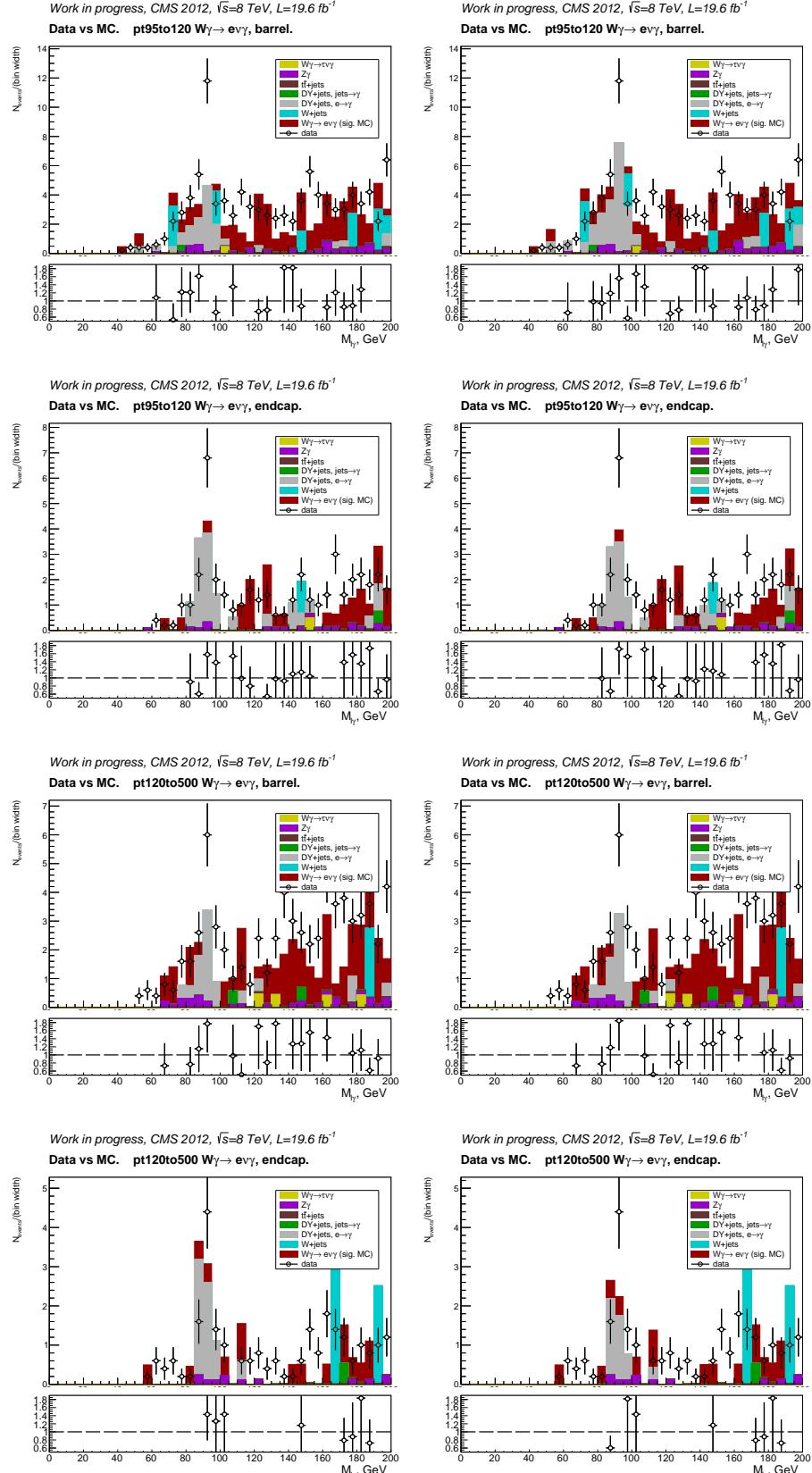


Figure I.6:  $M_{e\gamma}$  distribution, data vs MC. Bins 95 – 120 – 500 GeV. Left: all MC samples are normalized to luminosity of data, PU weight and SFs, right: DY+jets( $e \rightarrow \gamma$ ) also normalized to  $e \rightarrow \gamma$  background estimates.

## **Appendix J**

# **Correlation Matrices for Different Sources of the Systematic Uncertainties**

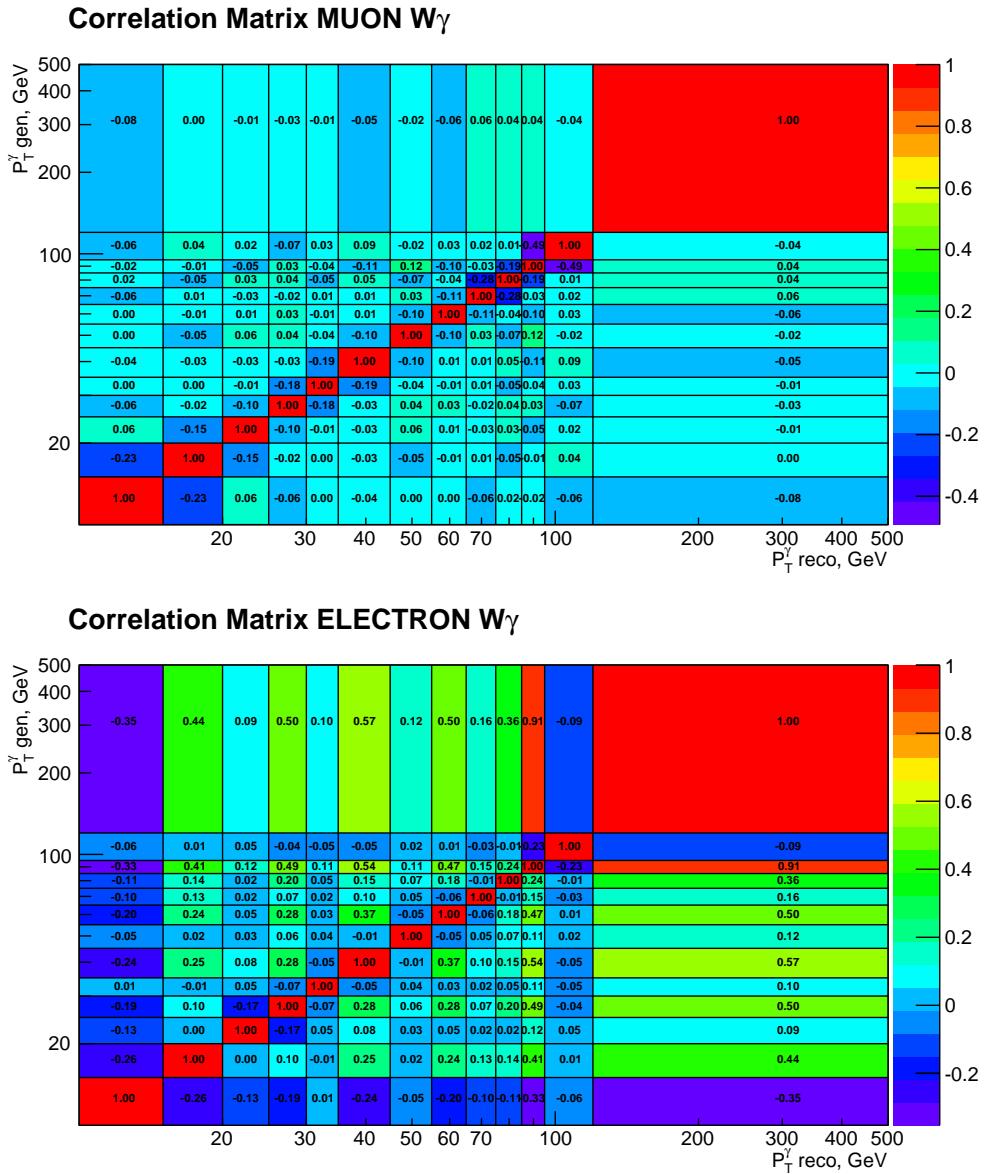


Figure J.1: Correlation Matrices for systematic error due to the difference between  $I_{ch}^\gamma$  and  $\sigma_{in in}^\gamma$  fit results.

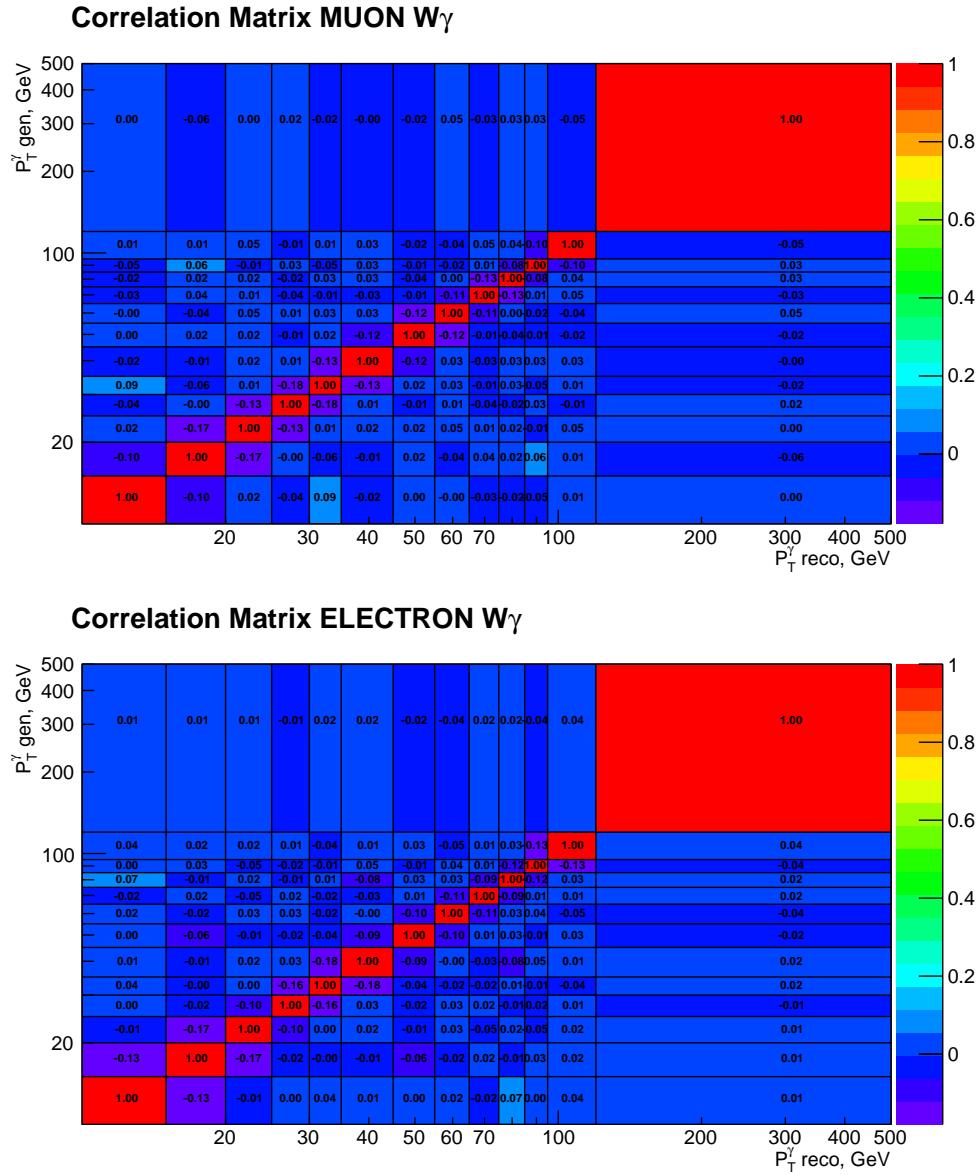


Figure J.2: Correlation Matrices for systematic error due to uncertainty on the  $Z\gamma$  MC sample normalization.

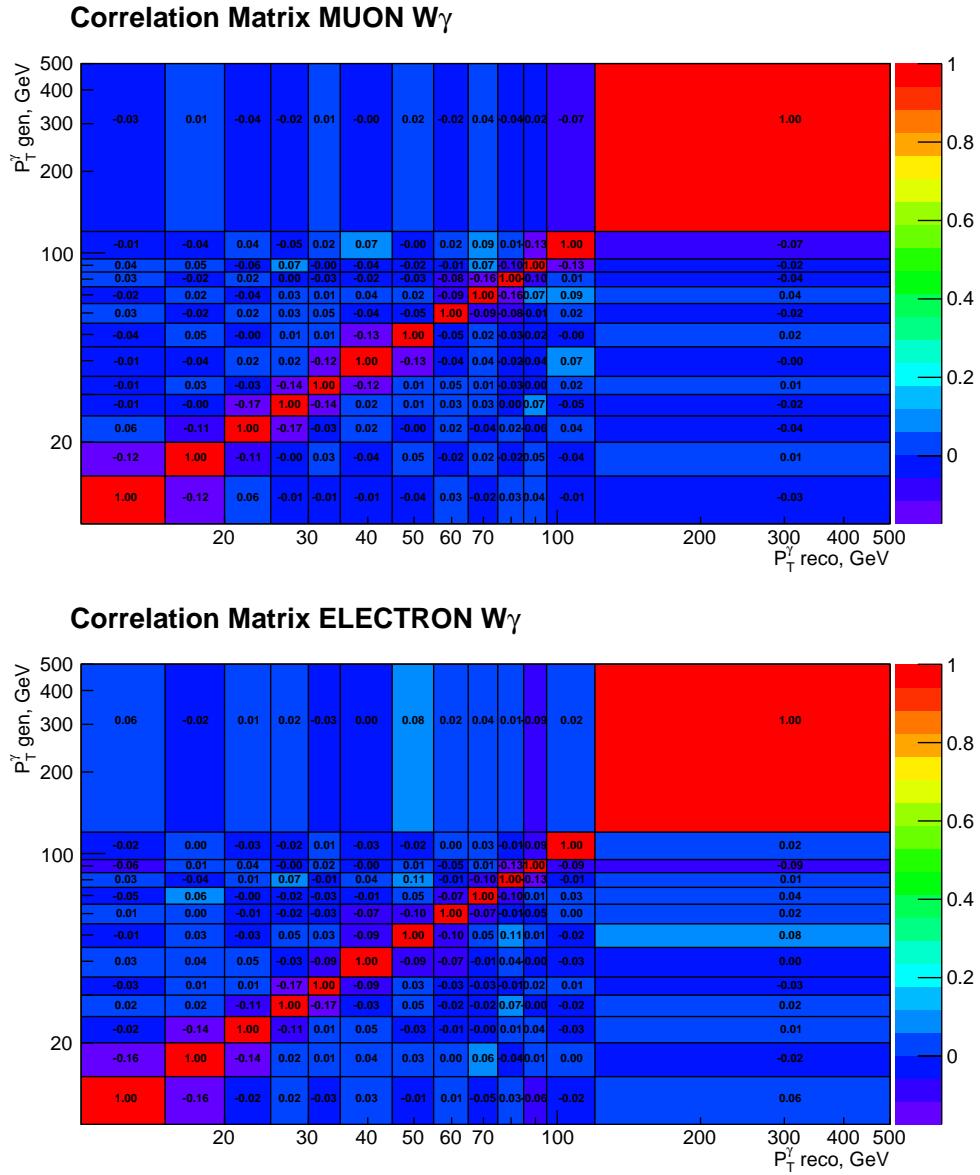


Figure J.3: Correlation Matrices for systematic error due to the template statistical power.

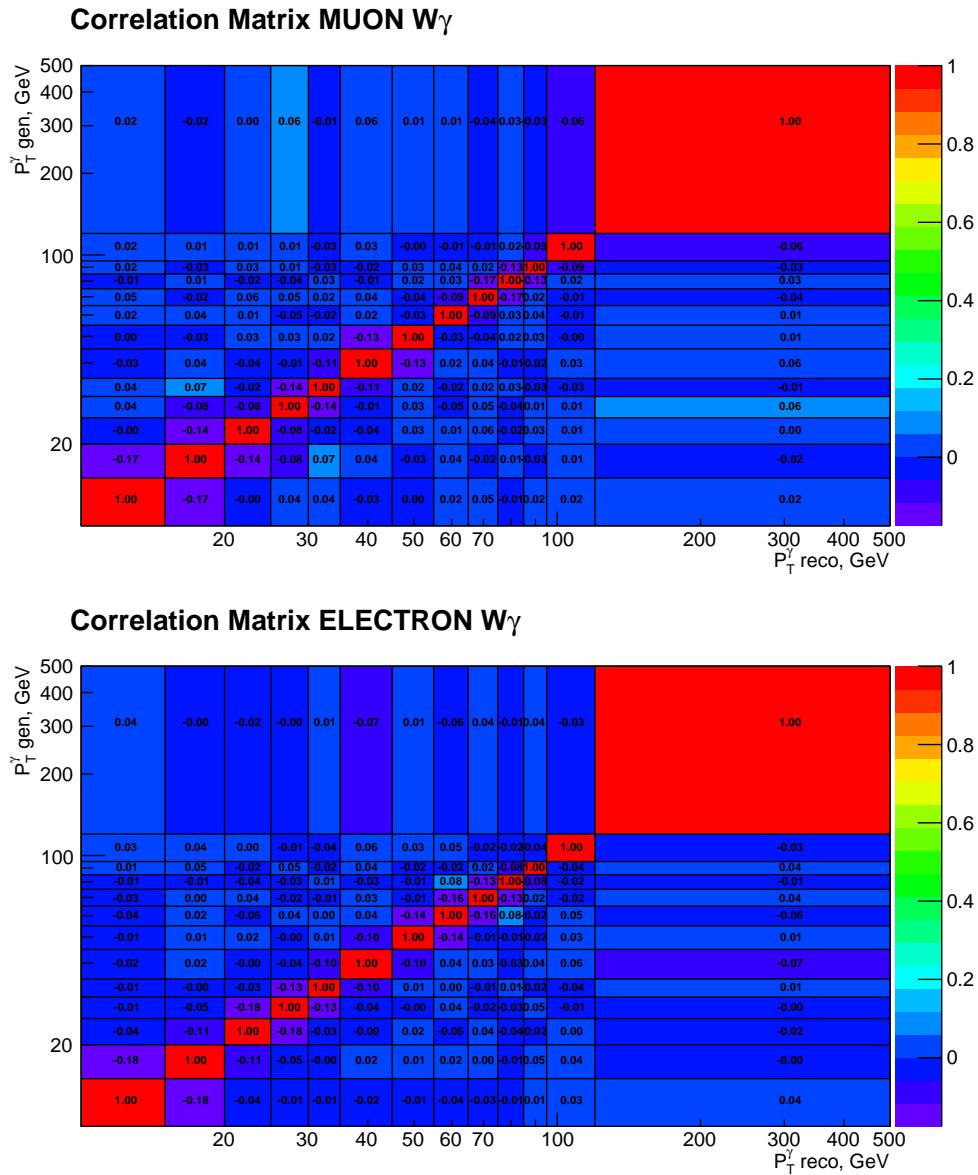


Figure J.4: Correlation Matrices for systematic error due to real- $\gamma$  background subtraction.

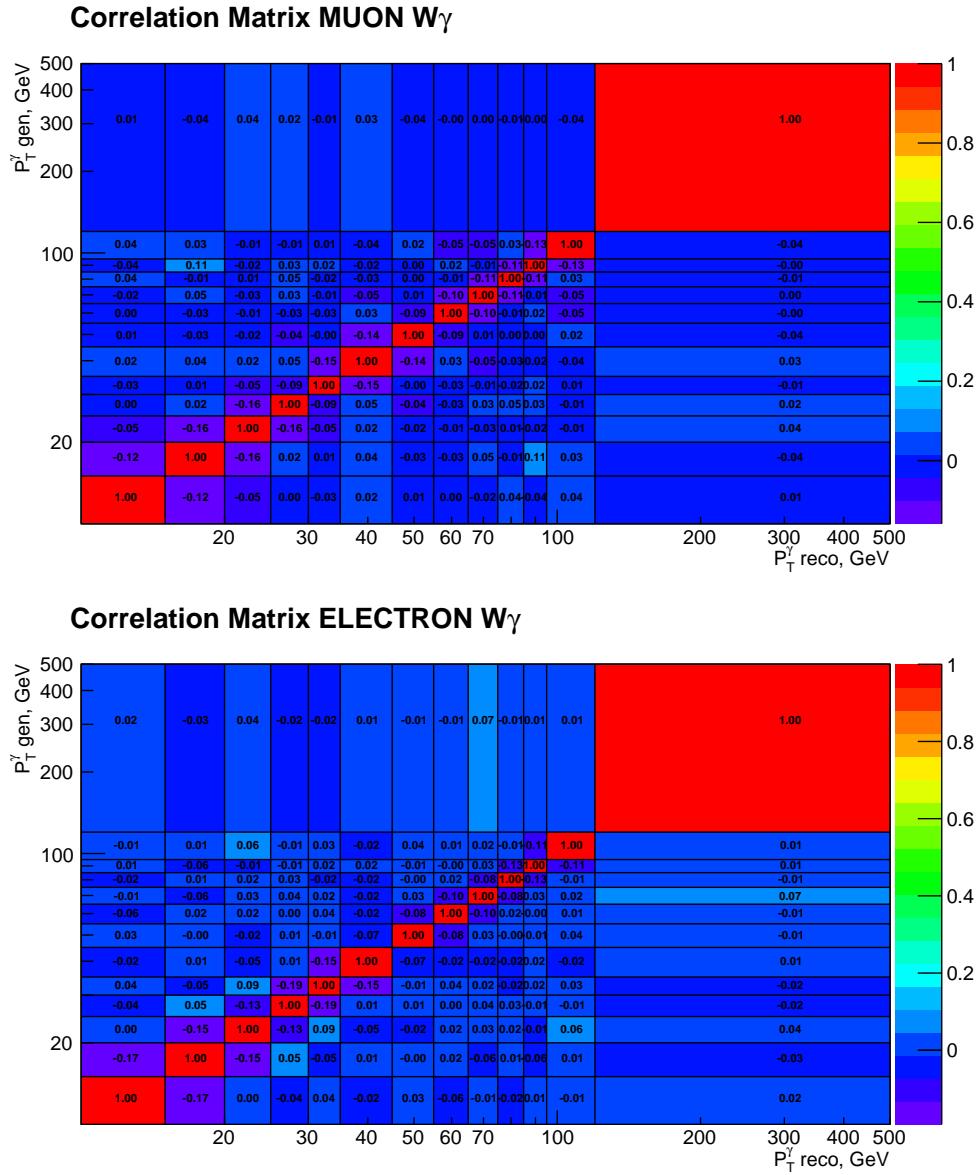


Figure J.5: Correlation Matrices for systematic error due to signal MC statistics for unfolding.

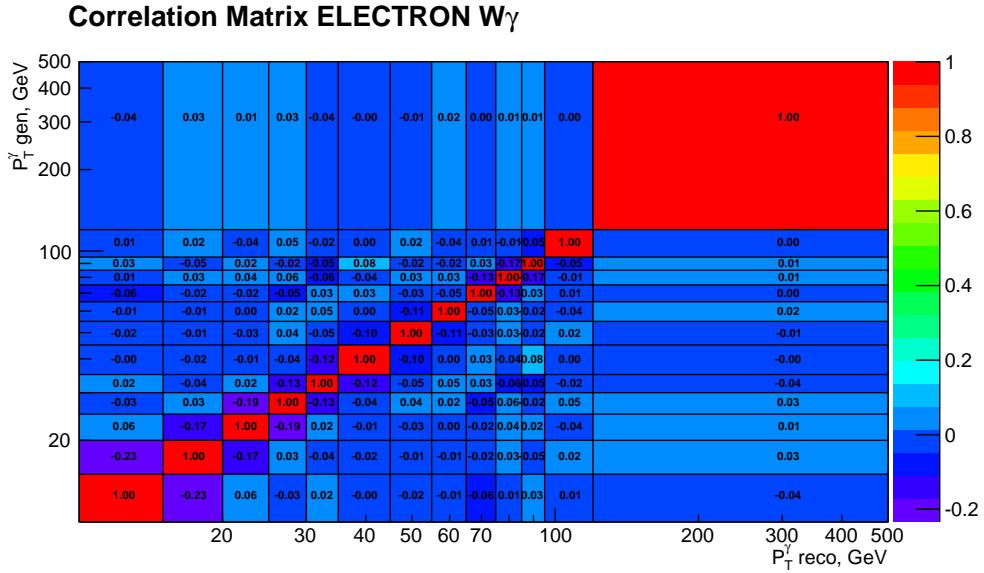


Figure J.6: Correlation Matrix for systematic error due to statistics of different samples for  $e \rightarrow \gamma$  background estimation.

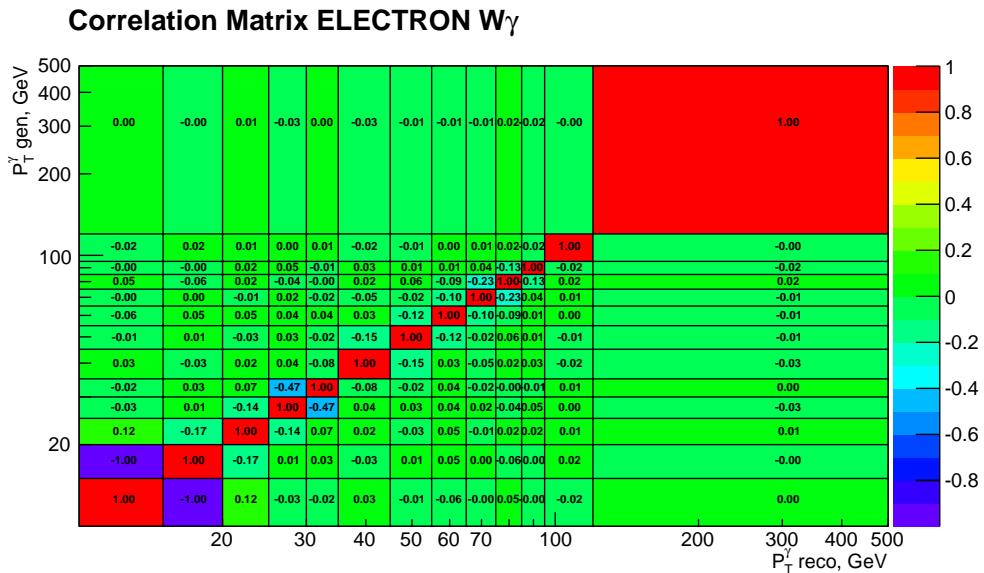


Figure J.7: Correlation Matrix for systematic error due to fit bias for  $e \rightarrow \gamma$  background estimation.

# Appendix K

## Code and Software

The CMS software (CMSSW) [51] is the tool developed to process all CMS responses, reconstruct particles and prepare data in a condition convenient for the final physics measurements. CMSSW is mostly written on C++ and python programming languages. It has hundreds of contributors that use GITHUB [52] to share their work. All CMS physics measurements use CMSSW.

The procedure of the tracker alignment and validation described in Ch. 4 is also a part of the CMSSW although the Millepede-II algorithm itself is implemented in the external software tool.

The samples for the physics measurements are stored in a format of ROOT trees. The ROOT tree contains multiple parameters for each entry and allows easy access to all parameters. These properties makes it convenient to use ROOT trees for particle physics measurements where, usually, one entry corresponds to one event. The ROOT trees provided by reconstruction algorithms of CMSSW are referred as “tuples”. Tuples are further processed by different large physics subgroups that prepare “ntuples”. Ntuples store only information that is necessary for a specific class of measurements and arrange it in a more convenient way for

this specific class of measurements.

The author of this dissertation used ntuples prepared by Central Taiwan University and Kansas State University groups mostly for various diboson and triboson measurements. The code of the program that prepares the ntuples is available at [53].

The code for the CMS  $W\gamma$  measurement at  $\sqrt{s} = 8$  TeV was written by the author of this dissertation using C++ language, ROOT and RooFit [49] packages, *RooUnfold* [54] class for the detector resolution unfolding, and *RooCMSShapePdf* [50] for  $e \rightarrow \gamma$  background estimation. Auxiliary shell scripts are used to run the chain of C++ programs corresponding to separate physics measurement steps. The code is available at [52].

Several cross check were performed with other collaborators to make sure the code is free of major errors. Especially the event selection and background estimation for the electron channel is fully implemented by both Kansas State University group and the author of this dissertation in separate frameworks. These procedures are carefully cross checked between two developers.

# Appendix L

## $Z\gamma$ Check

For the  $Z\gamma$  check, the same procedures we use for the  $W\gamma$  measurement, we also use for the  $Z\gamma$  measurement except those that are not applicable. Figure L.1 shows the data vs MC distribution of  $Z\gamma$ -selected samples. The selected sample mostly consists of  $Z\gamma$  signal event and DY+jets background. DY+jets background is a source of jets $\rightarrow\gamma$  background and is estimated the same way as it is done for our nominal  $W\gamma$  measurement.

The templates are derived from  $Z\gamma \rightarrow \mu\mu\gamma$  sample, therefore, the  $Z\gamma$  check in the muon channel is not a valid physics measurement but a closure check because the templates for the jets $\rightarrow\gamma$  background estimation procedure are largely overlap with the fitted data. At the same time, the  $Z\gamma$  check in the electron channel is a valid physics measurement. Fit results on data and pseudodata (MC mixtures) show good agreement for both channels (Fig. L.2-L.5). The fit plots themselves are available in App. ??-??.

Major systematic uncertainties are estimated the same way as it is done for  $W\gamma$  measurement and are listed in Tab. L.1-L.2. Measured cross section values compared to the MC-based cross section are listed in the Tab. L.3-L.4. Figure L.6

shows an agreement between muon and electron channels, agreement with the MC-based cross section and with the published  $Z\gamma$  measurement with CMS at  $\sqrt{s} = 8$  TeV [46].

The good agreement between  $Z\gamma$  cross section of this measurement and the published one validates steps of the  $W\gamma$  measurement that are the same between  $Z\gamma$  and  $W\gamma$  measurements. The list of the same steps include muon, electron, and photon selection, jets  $\rightarrow \gamma$  background estimation, detector resolution unfolding, acceptance and efficiency corrections.

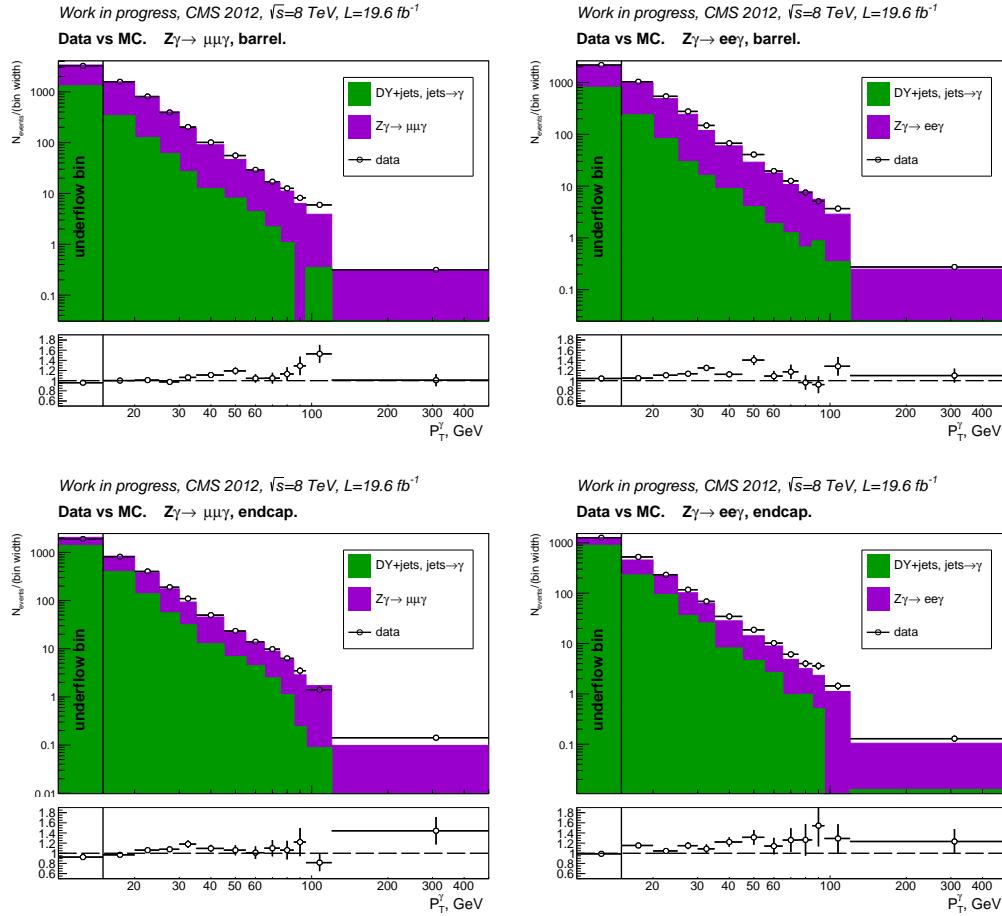


Figure L.1: Data vs MC plots. Left column - muon channel, right column - electron channel. Top to bottom: barrel and endcap photons

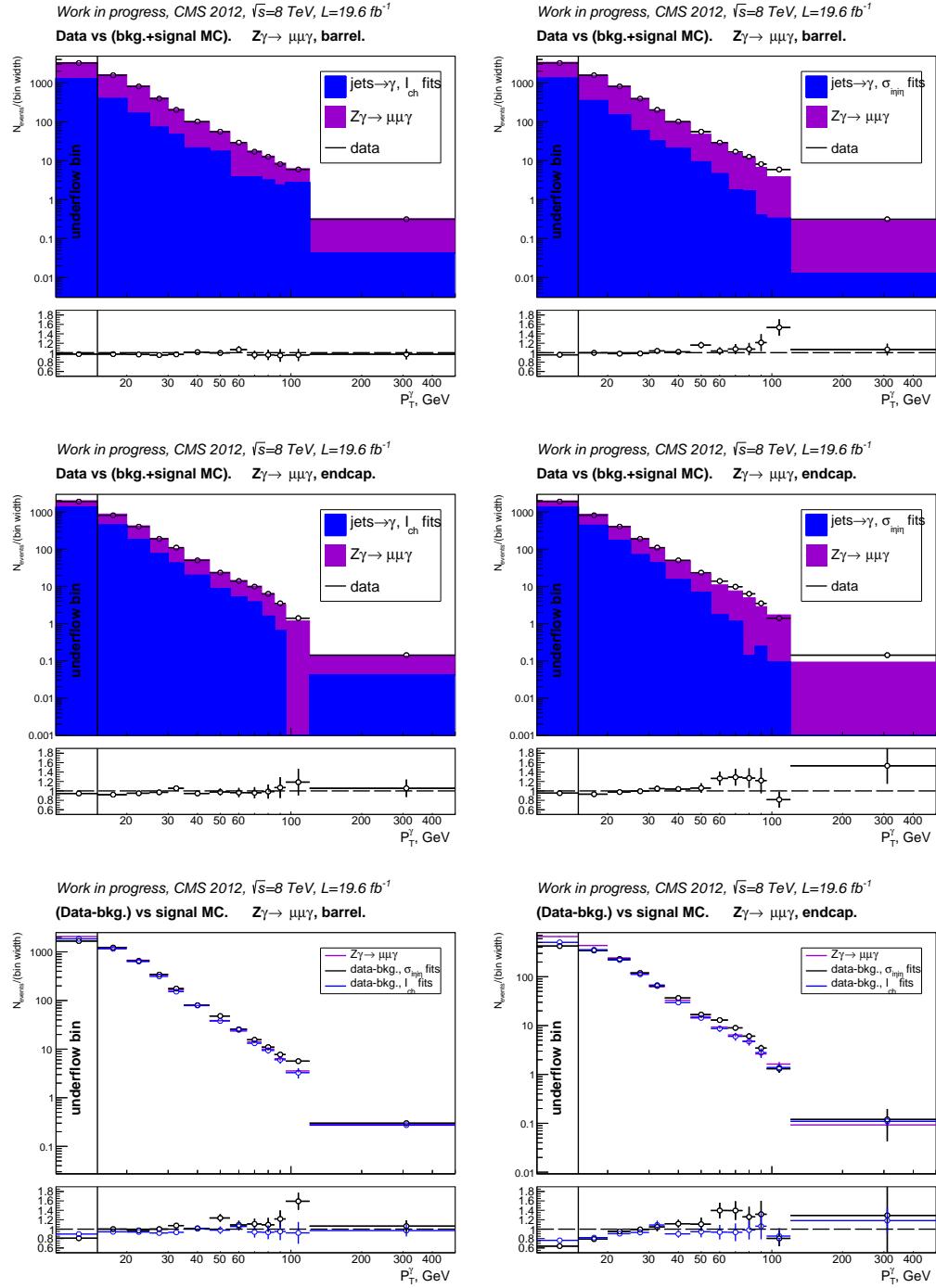


Figure L.2:  $Z\gamma$  check. Muon channel. Top and middle: data vs fake- $\gamma$  background derived from the template method + real- $\gamma$  background predicted by dedicated MC samples + signal MC, with  $I_{ch}$  and  $\sigma_{i\eta}$  used as fit variables. Bottom: data yields after full background subtraction vs signal MC.  $I_{ch}$  vs  $\sigma_{i\eta}$  fit results.

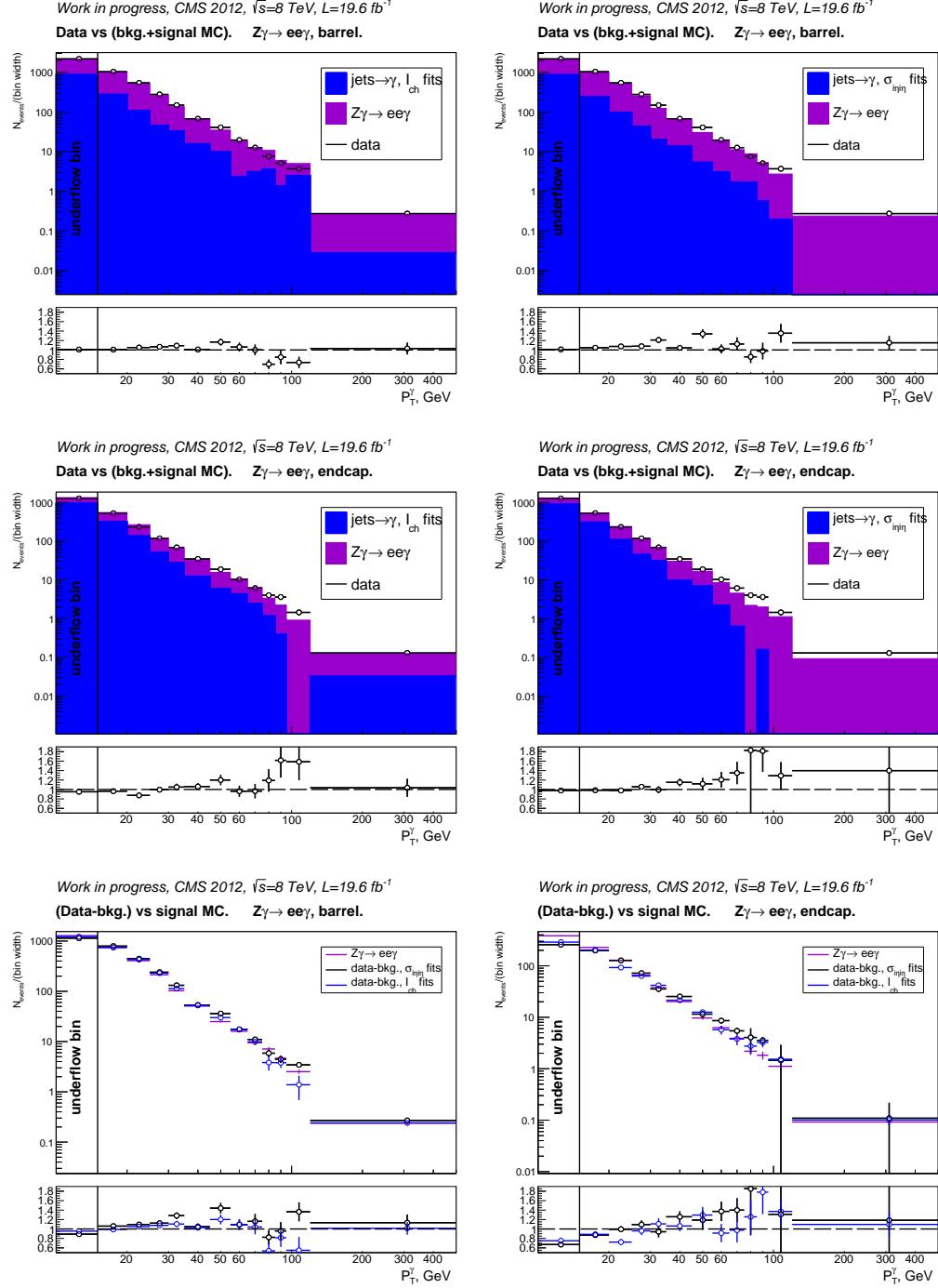


Figure L.3:  $Z\gamma$  check. Electron channel. Top and middle: data vs fake- $\gamma$  background derived from the template method + real- $\gamma$  background predicted by dedicated MC samples + signal MC, with  $I_{ch}$  and  $\sigma_{i\eta}$  used as fit variables. Bottom: data yields after full background subtraction vs signal MC.  $I_{ch}$  vs  $\sigma_{i\eta}$  fit results.

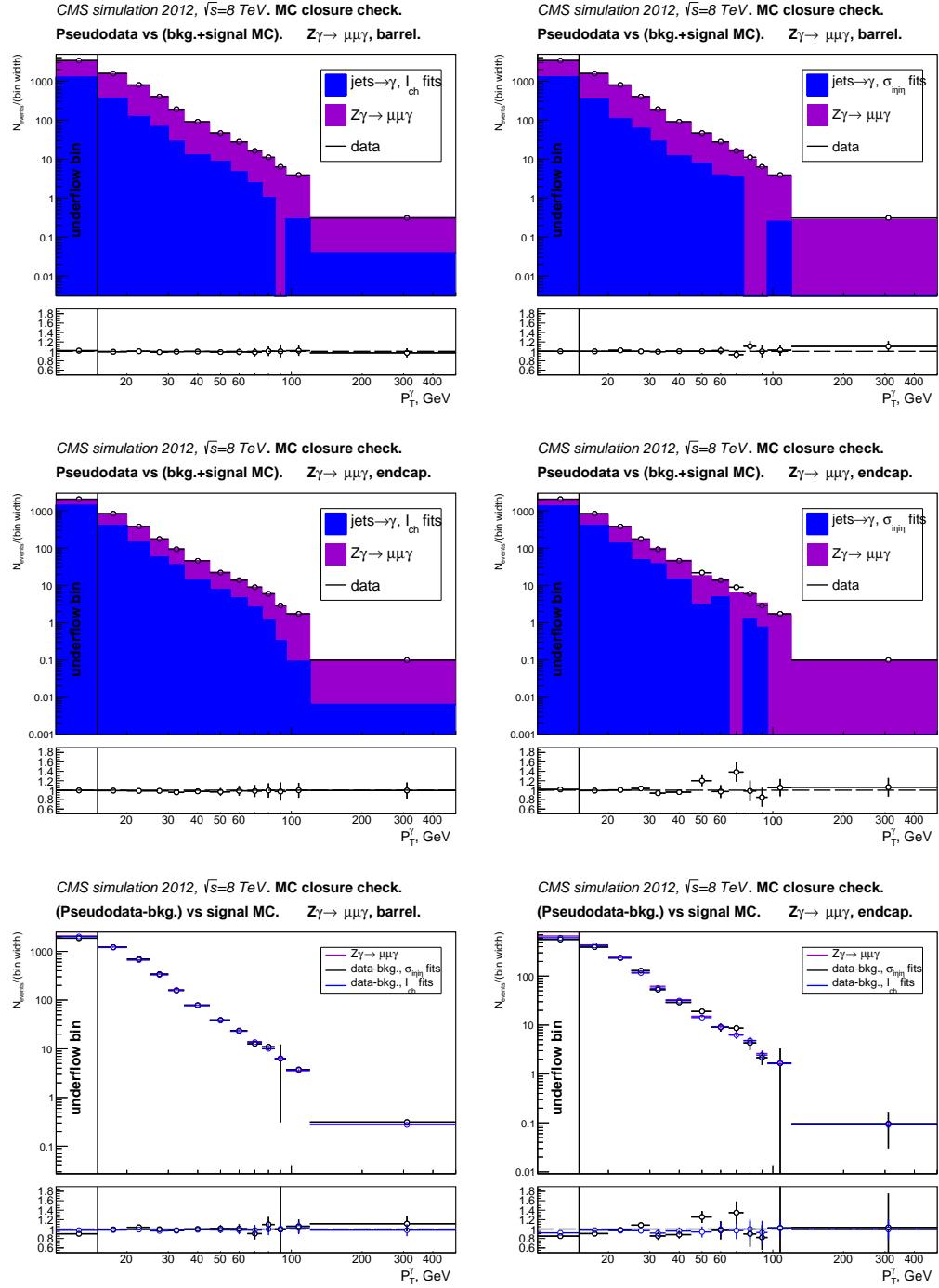


Figure L.4:  $Z\gamma$  MC closure check. Muon channel. Top and middle: pseudodata vs fake- $\gamma$  background derived from the template method + real- $\gamma$  background predicted by dedicated MC samples + signal MC, with  $I_{ch}$  and  $\sigma_{inj}$  used as fit variables. Bottom: pseudodata yields after full background subtraction vs signal MC.  $I_{ch}$  vs  $\sigma_{inj}$  fit results.

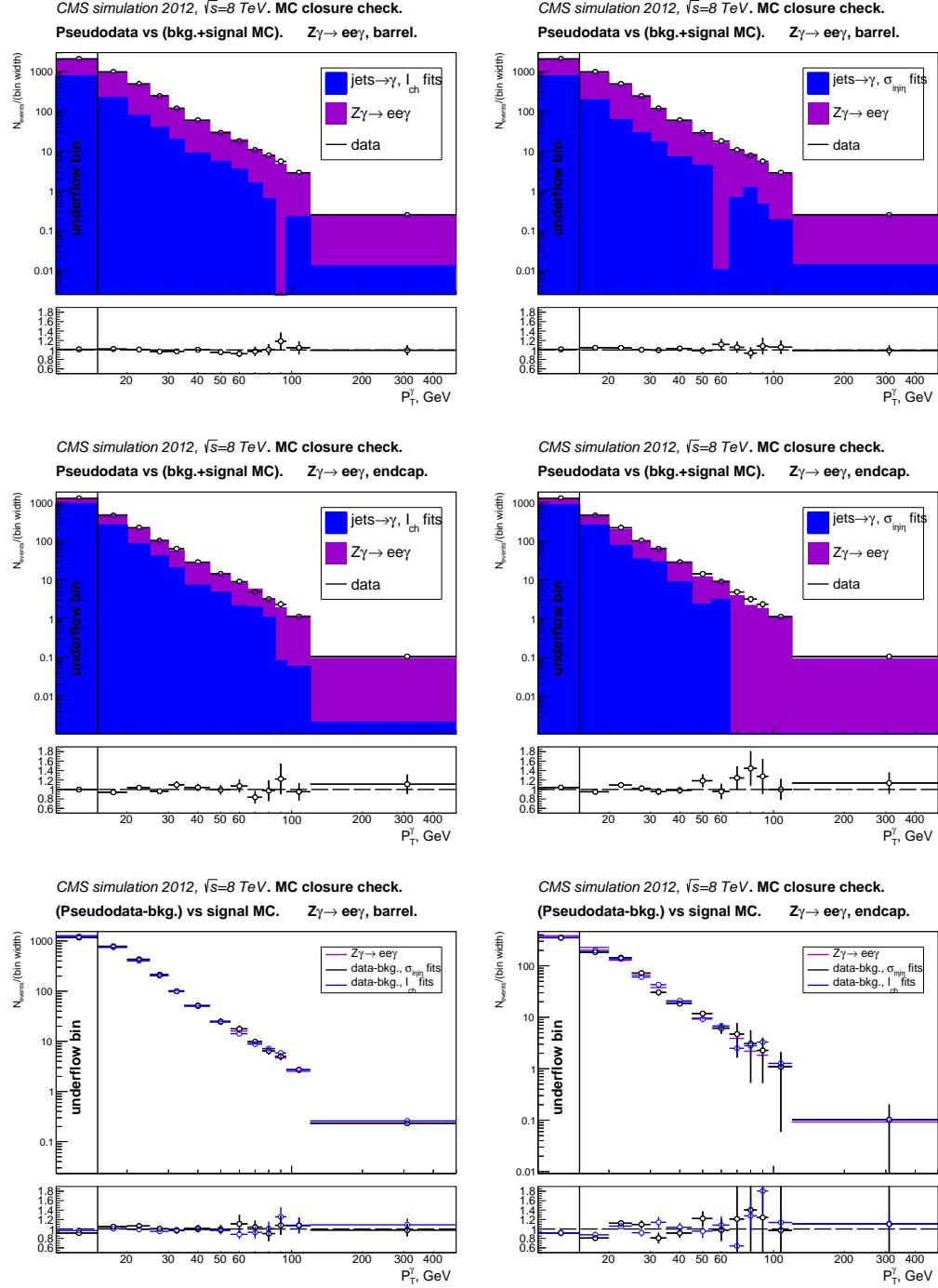


Figure L.5:  $Z\gamma$  MC closure check. Electron channel. Top and middle: pseudodata vs fake- $\gamma$  background derived from the template method + real- $\gamma$  background predicted by dedicated MC samples + signal MC, with  $I_{ch}$  and  $\sigma_{inj}$  used as fit variables. Bottom: pseudodata yields after full background subtraction vs signal MC.  $I_{ch}$  vs  $\sigma_{inj}$  fit results.

Table L.1: Relative errors [%]. MUON  $Z\gamma$ 

bin lims	err stat	syst $ N_{Ich} - N_{sigi\eta} $	$Z\gamma$ MC norm	$A \times \epsilon$ MC stat	syst lumi	unf MC stat	syst total	syst + stat total
total	1	1	1	0	3	1	3	3
15-20	2	2	2	1	3	2	4	5
20-25	2	2	3	1	3	2	5	5
25-30	3	3	4	2	3	3	7	8
30-35	4	6	5	3	3	5	10	10
35-45	4	3	6	3	3	4	9	9
45-55	6	8	8	4	3	6	14	15
55-65	7	5	7	5	3	7	13	14
65-75	9	7	8	6	3	9	16	18
75-85	10	8	6	7	3	10	16	19
85-95	12	8	8	9	3	12	19	23
95-120	11	10	6	8	3	11	18	21
120-500	8	5	9	7	3	9	16	18

Table L.2: Relative errors [%]. ELECTRON  $Z\gamma$ 

bin lims	err stat	syst $ N_{Ich} - N_{sigi\eta} $	$Z\gamma$ MC norm	$A \times \epsilon$ MC stat	syst lumi	unf MC stat	syst total	syst + stat total
total	1	1	1	0	3	1	3	4
15-20	2	3	3	1	3	2	5	6
20-25	3	2	3	1	3	3	5	6
25-30	4	3	4	2	3	4	7	8
30-35	5	4	5	3	3	6	10	11
35-45	5	4	6	3	3	5	10	11
45-55	6	6	6	4	3	7	11	13
55-65	9	7	8	5	3	9	15	17
65-75	10	8	8	7	3	11	18	20
75-85	14	11	12	9	3	16	25	28
85-95	15	9	6	10	3	17	23	28
95-120	10	5	6	9	3	11	16	19
120-500	9	3	7	8	3	10	15	17

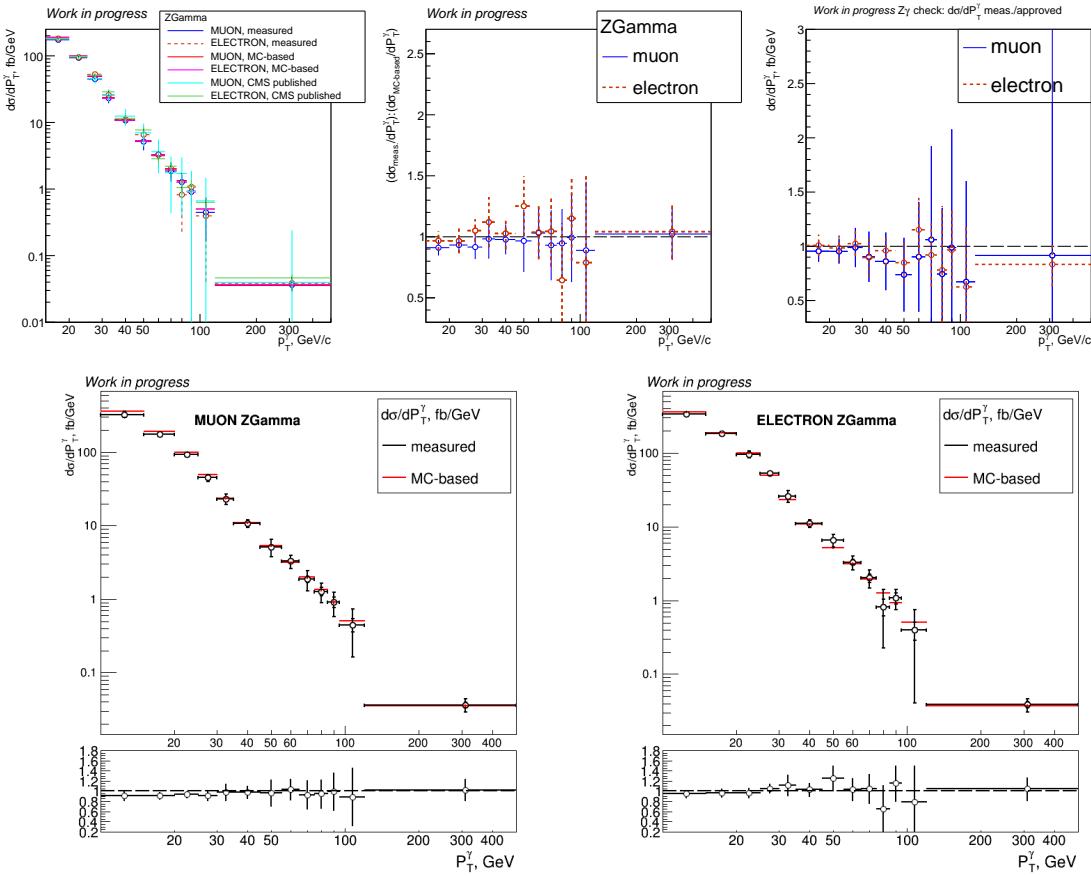


Figure L.6:  $Z\gamma$  differential cross section. Top, left to right: the  $Z\gamma$  differential cross section; the ratio of measured over the MC-based  $Z\gamma$  differential cross section; the ratio of the measured over published  $Z\gamma$  differential cross section. Bottom: the  $Z\gamma$  measured differential cross section overlaid with the MC-based cross section, the muon channel (left) and the electron channel (right).

Table L.3: Cross section and errors. MUON  $Z\gamma$ 

bin lims	$d\sigma/dP_T$ MC based	$d\sigma/dP_T$ meas.
total	2085	$1985 \pm 18 \pm 62$
15-20	191	$178 \pm 3 \pm 8$
20-25	100	$98 \pm 2 \pm 5$
25-30	49	$43 \pm 1 \pm 3$
30-35	24	$23 \pm 1 \pm 2$
35-45	11	$11 \pm 0 \pm 1$
45-55	5.3	$5.2 \pm 0.3 \pm 0.7$
55-65	3.2	$3.4 \pm 0.2 \pm 0.4$
65-75	2.0	$2.1 \pm 0.2 \pm 0.3$
75-85	1.3	$1.5 \pm 0.1 \pm 0.2$
85-95	0.9	$1.1 \pm 0.1 \pm 0.2$
95-120	0.50	$0.56 \pm 0.06 \pm 0.10$
120-500	0.036	$0.040 \pm 0.003 \pm 0.006$

Table L.4: Cross section and errors. ELECTRON  $Z\gamma$ 

bin lims	$d\sigma/dP_T$ MC based	$d\sigma/dP_T$ meas.
total	2061	$2091 \pm 24 \pm 70$
15-20	188	$180 \pm 4 \pm 9$
20-25	99	$101 \pm 3 \pm 6$
25-30	50	$51 \pm 2 \pm 4$
30-35	23	$25 \pm 1 \pm 2$
35-45	11	$12 \pm 1 \pm 1$
45-55	5.2	$6.5 \pm 0.4 \pm 0.7$
55-65	3.2	$3.3 \pm 0.3 \pm 0.5$
65-75	1.9	$2.3 \pm 0.2 \pm 0.4$
75-85	1.3	$1.2 \pm 0.2 \pm 0.3$
85-95	0.9	$1.1 \pm 0.2 \pm 0.3$
95-120	0.50	$0.66 \pm 0.07 \pm 0.11$
120-500	0.037	$0.042 \pm 0.004 \pm 0.006$

# Bibliography

- [1] Particle data group. Available at: [pdg.lbl.gov](http://pdg.lbl.gov).
- [2] PhD thesis. Senka's thesis, CMS Wg- $\mu$ nug, 7 TeV.
- [3] Available at: <https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsSMPaTGC#>
- [4] 7TeV Wg CMS paper.
- [5] 7TeV Wg ATLAS paper.
- [6] website: <http://cds.cern.ch/record/1621583/files/>.
- [7] website: <https://muon.files.wordpress.com/2014/01/cms-sigma-vs-e.png>.
- [8] Available at: <https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults>.
- [9] Available at: <https://inspirehep.net/record/837874/plots>.
- [10] F. Meier Aeschbacher. *Measurement of the  $\Lambda_b^0$  baryon lifetime using the Compact Muon Solenoid experiment at Cern's Large Hadron Collider*. PhD thesis, ETH Zurich, 2013.
- [11] CMS collaboration. Alignment of the cms tracker with large hadron collider data.

- [12] CERN brochure, <http://cds.cern.ch/record/1165534/files/CERN-Brochure-2009-003-Eng.pdf>.
- [13] CMS Collaboration. Observation of a new boson at a mass of 125 gev with cms experiment at the lhc. *Phys. Lett. B*, 716(1):30–61, 2012. Available at: <https://arxiv.org/abs/1207.7235>.
- [14] 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(1):1–29, 2012. Available at: <https://arxiv.org/abs/1207.7214>.
- [15] Nasa website. Available at: <https://science.nasa.gov/astrophysics/focus-areas/what-is-dark-energy>.
- [16] A. Pich. *The Standard Model of Electroweak Interactions*, 2007. Updated version of the lectures given at the 2006 European School of High Energy Physics (Aronsborg, Sweden, 18 June - 1 July 2006) and at the 4th CERN - CLAF School of High Energy Physics (Vina del Mar, Chile, 18 February - 3 March 2007). Available at: <https://arxiv.org/abs/0705.4264>.
- [17] D. Griffiths. *Introduction to Elementary Particles*. WILEY-VCH, 2008.
- [18] Available at: [https://www-do.fnal.gov/results/publications\\_talks/thesis/snyder/html/nod.html](https://www-do.fnal.gov/results/publications_talks/thesis/snyder/html/nod.html)
- [19] Martin Halzen. *Quarks and leptons*.
- [20] Hard scattering. Available at: <http://iopscience.iop.org/article/10.1088/0034-4885/70/1/R02>.
- [21] LO (1969) theory paper.
- [22] NNLO theory paper.

- [23] aTGC paper, LEP parametrization.
- [24] Do Combination of  $W\gamma$ , WW, WZ and VW using 8.6 fb-1 of 2 TeV  $p\bar{p}$  collisions  
Phys.Lett. B718 (2012) 451-459.
- [25] LEP Combination of WW and single W using 0.7 fb-1 per experiment of  $e^+e^-$   
collisions at WW pair production energies arXiv:1302.3415.
- [26] ATLAS WW using 20.3 fb-1 of 8 TeV pp collisions Submitted to JHEP.
- [27] CMS WW using 4.9 fb-1 of 7 TeV pp collisions Eur. Phys. J. C 73 (2013) 2610.
- [28] CMS WW using 19.4 fb-1 of 8 TeV pp collisions Submitted to EPJC.
- [29] ATLAS VW ( $V=W,Z \rightarrow jj$ ) using 4.6 fb-1 of 7 TeV pp collisions JHEP 01 (2015)  
049.
- [30] CMS VW ( $V=W,Z \rightarrow jj$ ) using 5.0 fb-1 of 7 TeV pp collisions Eur.Phys.J. C73  
(2013) 2283.
- [31] MCFM generator.
- [32] LHC TDR.
- [33] Available at: <http://home.cern/topics/large-hadron-collider>.
- [34] Available at: <https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults>.
- [35] CMS TDR.
- [36] <https://cds.cern.ch/record/345832/files/lhc-project-report-165.pdf>.
- [37] CMS collaboration. Description and performance of track and  
primary-vertex reconstruction with the cms tracker. Available at:  
<https://arxiv.org/abs/1405.6569>.

- [38] CMS collaboration. Energy calibration and resolution of the cms electromagnetic calorimeter in pp collisions at  $\text{sqrt}(s) = 7 \text{ tev}$ . Available at: <https://arxiv.org/abs/1306.2016>.
- [39] CMS collaboration. The performance of the cms muon detector in protonproton collisions at  $s=7 \text{ tev}$  at the lhc. Available at: <https://arxiv.org/pdf/1306.6905.pdf>.
- [40] CMS collaboration. The cms trigger system. Available at: <https://arxiv.org/pdf/1609.02366.pdf>.
- [41] CMS collaboration. <https://inspirehep.net/record/925379/files/PFT-09-001-pas.pdf>.
- [42] Millepede ii algorithm.
- [43] Hip algorithm.
- [44] Cms tracker detector performance results 2015: Alignment. Available at: <https://twiki.cern.ch/twiki/bin/view/CMSPublic/TkAlignmentPerformance2015>.
- [45] MadGraph generator.
- [46] CMS collaboration. Zg, 8 tev.
- [47] CMS collaboration. Wgg, 8 tev.
- [48] D'agostini unfolding.
- [49] Rooftit.
- [50] Roocmsshapepdf.
- [51] Cms software. Available at: <http://cms-sw.github.io/index.html>.

[52] Github.

[53] ggntuplizer.

[54] Roounfold.

[55] UNKNOWN [TEMPORARY]. Source of sm figure [temporary]. Available at:  
<http://www.isgtw.org/spotlight/go-particle-quest-first-cern-hackfest>.

[56] Schroeder Peskin. *Introduction to Quantum Field Theory*.

[57] L. Gray. *A measurement of the  $Z\gamma$  Cross Section and Limits on anomalous Gauge Couplings at  $\sqrt{s} = 7$  TeV Using CMS*. PhD thesis, University of Wisconsin - Madison, 8 2012.

[58] Wikipedia article. cross section (physics). Available at:  
[https://en.wikipedia.org/wiki/Cross\\_section\\_%28physics%29](https://en.wikipedia.org/wiki/Cross_section_%28physics%29).

[59] NLO theory paper.

[60] CMS 8 TeV WW lnulnu.

[61] CMS 7 TeV Zg nunug.

[62] website: [https://inspirehep.net/record/1235541/files/LHC\\_all.png](https://inspirehep.net/record/1235541/files/LHC_all.png).

[63] Available at: <http://cds.cern.ch/journal/CERNBulletin/2014/24/News%20Articles/170660>

[64] F. Meier Aeschbacher. Hunting down suspicious results: recent improvements in the alignment of the cms tracker detector, 2010. Available as an internal CMS presentation at:  
<https://indico.cern.ch/event/243940/contributions/1562955/attachments/416194/578159/>

[65] Root.

[66] mycode.