The Measurement of the Wγ Cross Section at 8 TeV (PhD thesis defense)

Ekaterina Avdeeva

University of Nebraska - Lincoln

June, 2017

Introduction

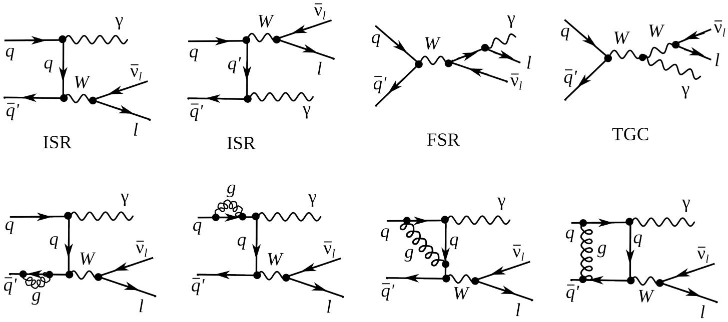


Figure: The Feynman diagrams. ISR(x2), FSR, and TGC.

Process signature:

- prompt, energetic, and isolated muon/electron
- prompt isolated photon
- significant missing energy due to neutrino

Goals:

- measure total and differential $\frac{d\sigma}{dp_T^{\gamma}}$ cross section

Motivation:

Theory

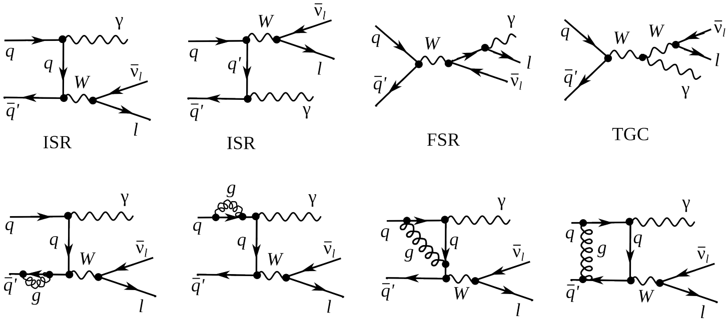


Figure: The Feynman diagrams. ISR(x2), FSR, and TGC.

Process signature:

- prompt, energetic, and isolated muon/electron
- prompt isolated photon
- significant missing energy due to neutrino

Goals:

- measure total and differential $\frac{d\sigma}{dp_T^\gamma}$ cross section

Motivation:

Experimental Setup

Phase Space Definition

Applied at generator-level to compute acceptance and MC-based cross section;
(tried to follow phase space definition of approved Zγ analysis [AN-2013-280] as close as possible)

- $p_T^{\gamma} > 15$ GeV
- $\Delta R(\gamma, lep) > 0.7$
- $|\eta^{\gamma}| < 2.5$, $|\eta^{lep}| < 2.5$
- $p_T^{lep} > 20$ GeV
- $Iso^{\gamma} < 5$ GeV (sum of all P_T within ΔR <0.3 based on MC-generated particles)
- for Zγ Check: M(lep,lep)>50 GeV
- for differential cross section, P_T^{γ} binning:
  15-20-25-30-35-45-55-65-75-85-95-120-500 GeV

Analysis Outline

Table: 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	$d\sigma/dP_T$	σ
select events	N^j_{sel}	N_{sel}
subtract background	$N^j_{sign} = N^j_{sel} - N^j_{bkg}$	$N_{sign} = N_{sel} -$
unfold	$N^i_{A\times\epsilon} = U_{ij} \cdot N^j_{sign}$	$-$
correct for the acceptance and efficiency	$N^i_{true} = \frac{N^i_{A\times\epsilon}}{(A\times\epsilon)^i}$	$N_{true} = \frac{N}{A}$
divide by luminosity and bin width	$\left(\frac{d\sigma}{dP^\gamma_T}\right)^i = \frac{N^i_{true}}{L\cdot(\Delta P^\gamma_T)^i}$	$\sigma = N_{true}$
estimate systematic uncertainties		

Data and MC 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 fb^{-1}. To select $W\gamma$ events, we use data collected by single muon and single electron triggers.

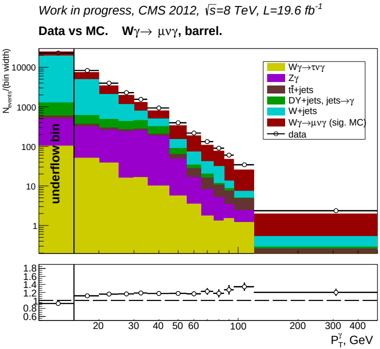
Table: Summary of simulated samples used in the measurement.

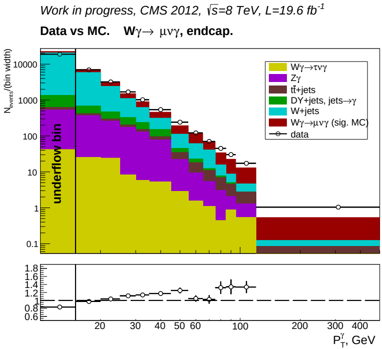
Process	Type	σ, pb
$W\gamma \rightarrow l\nu\gamma$	signal	554
W+jets$\rightarrow$ $l\nu$+jets	background	36257
DY+jets$\rightarrow$ ll+jets	background	3504
$t\bar{t}$+jets$\rightarrow$ 1l+X	background	99
$t\bar{t}$+jets$\rightarrow$ 2l+X	background	24
$Z\gamma \rightarrow ll\gamma$	background	172

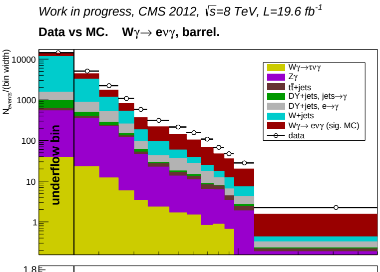
Cut-Based Event Selection

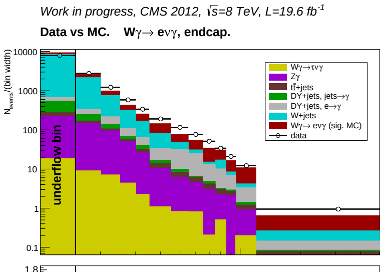
- - - - - - - **Muon Channel** - - - - - - -	- - - - - - **Electron Channel** - - - - - -
HLT_IsoMu24_eta2p1	HLT_Ele27_WP80
Photon selection:	
medium ID; $p_T > 15$; $\lvert\eta\rvert < 1.4442$ or $1.566 < \lvert\eta\rvert < 2.5$	
	ElectronConversionVeto$\rightarrow$PixelSeedVeto
Muon tight ID	Electron tight ID
$p_T^{\mu} > 25$ GeV; $\lvert\eta^{\mu}\rvert < 2.1$	$p_T^{ele} > 30$; $\lvert\eta^{ele}\rvert < 1.4442$ or $1.566 < \lvert\eta^{ele}\rvert < 2.5$
Second lepton veto:	
$p_T^{\mu 2} > 10$ GeV; $\lvert\eta^{\mu 2}\rvert < 2.4$	$p_T^{ele2} > 10$; $\lvert\eta^{ele2}\rvert < 1.4442$ or $1.566 < \lvert\eta^{ele2}\rvert < 2.5$; veto ID for ele2
$M_T^W > 40$ GeV	$M_T^W > 40$ GeV
	$110\ GeV > M_{e\gamma} > 70\ GeV$ excluded
$\Delta R(lep,\gamma) > 0.7$	

Data vs MC. P_T^γ



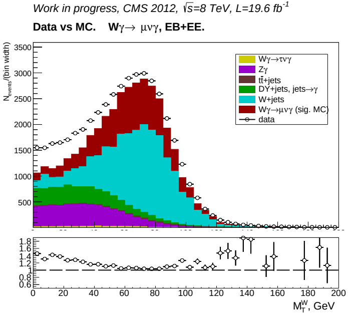


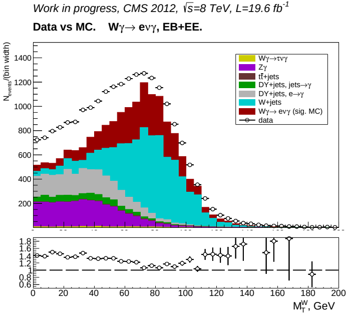




Data vs MC. M_T^W

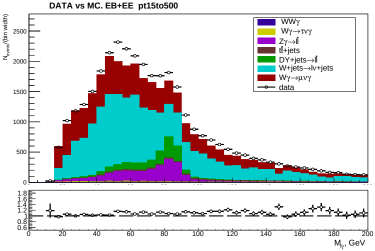
The selection requirement of $M_T^W > 40$ **GeV** is applied.

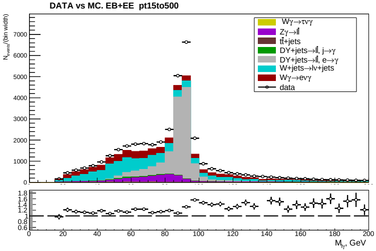




Data vs MC. $M_{l\gamma}$

The selection requirement of $M_{l,\gamma} < 70$ *or* $M_{l,\gamma} > 110$ **GeV** is applied in the **electron channel** only.

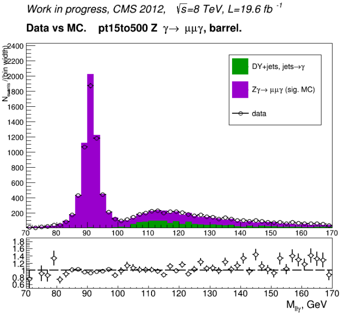


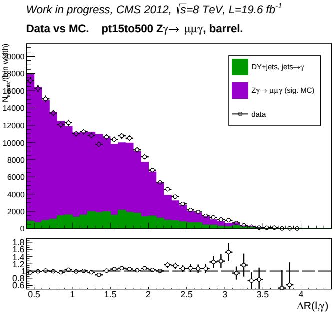


Jets$\to\gamma$ Background Subtraction. Method Description

- Fit function: $F(V_{fit}) = N_{true} \cdot T_{true}(V_{fit}) + N_{fake} \cdot T_{fake}(V_{fit})$
- $V_{fit} = I_{ch}^{\gamma}$ and $V_{fit} = \sigma_{i\eta i\eta}$ are considered
- T_{true}: Real-γ template from Z$\gamma \to \mu\mu\gamma$ FSR minus [small] fake contamination based on DYjets MC
- T_{fake}: Fake-γ template from Z$\gamma \to \mu\mu\gamma$ ISR minus [large] real contamination based on Zγ MC
- For I_{ch}, real-γ: use full data $p_T^{\gamma} > 15\ GeV$, fake-γ: use data only for certain p_T^{γ} to prepare template;
- For $\sigma_{i\eta i\eta}$, real-γ: merge data for $p_T^{\gamma} > 30\ GeV$ for preparing templates, fake-γ: merge data for $p_T^{\gamma} > 55\ GeV$;
- Same templates for the muon and the electron channels;
- Algorithm of template rebinning is applied if necessary;
- After fit is performed, extract $N_{real} = \epsilon_{real} \cdot N_{real}^{from-fit}$, $N_{fake} = \epsilon_{fake} \cdot N_{fake}^{from-fit}$, where $\epsilon_{real}(\epsilon_{fake})$ is efficiency of real-γ(fake-γ) events to pass the nominal cut of V_{fit} computed from the Zγ FSR sample (Zγ ISR data and MC samples).

$Z\gamma \rightarrow \bar{\mu}\mu\gamma$. FSR and ISR

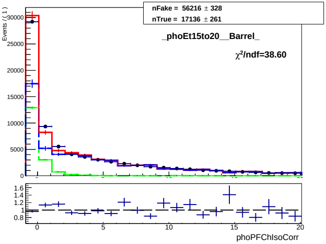


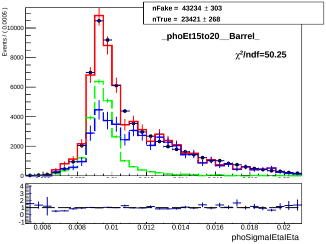


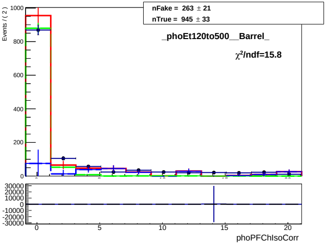
FSR selection: $M_{\mu\mu\gamma} < 101$ GeV and $\Delta R(\mu_1, \gamma) > 0.4$, nominal $Z\gamma$ selection otherwise

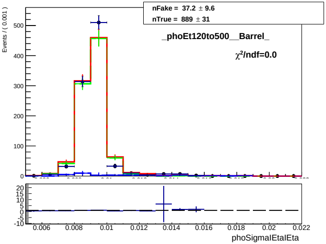
ISR selection: $M_{\mu\mu\gamma} > 101$ GeV and $\Delta R(\mu_1, \gamma) > 1.0$, nominal $Z\gamma$ selection otherwise

Jets $\rightarrow \gamma$ Background Subtraction. Fits Plots. Muon









More fit plots available in the backup and even more in the AN-15-055.

$e \rightarrow \gamma$ background estimation. Method Description and Fit Model

Method Description

- Get $N^{e\rightarrow\gamma}_{MC-Zpeak}$ (number of $e \rightarrow \gamma$ events under the Z-peak based on the MC prediction); done by counting
- Get $N^{e\rightarrow\gamma}_{data-Zpeak}$ (number of $e \rightarrow \gamma$ events under the Z-peak from data); done by fitting
- Get $N^{e\rightarrow\gamma}_{MC-nom}$ (number of $e \rightarrow \gamma$ events in the nominal range based on the MC prediction); done by counting
- Get $N^{e\rightarrow\gamma}_{data-nom}$ (number of $e \rightarrow \gamma$ events in the nominal range based on the MC predictionfrom data); done by scaling
  $N^{e\rightarrow\gamma}_{data-nom} = N^{e\rightarrow\gamma}_{MC-nom} \cdot N^{e\rightarrow\gamma}_{data-Zpeak} / N^{e\rightarrow\gamma}_{MC-Zpeak}$

Fit Model

$N_{sig} \cdot (RooNDKeysPdf\ x\ Gaussian) + N_{bkg} \cdot (RooCMSShapePdf)$

$e \rightarrow \gamma$ background estimation. Fit Binning and Details

- perform fits separately for each (pt,eta) bin
- pt binning is analysis binning
- eta binning is
  - 15-20-25-30-35-45-55 GeV
    - barrel: 0.00-0.10-0.50-1.00-1.44
    - endcap: 1.56-2.10-2.20-2.40-2.50
  - 55-65-75-85 GeV
    - barrel: 0.00-0.50-1.44
    - endcap: 1.56-2.20-2.50
  - 85-95-120-500 GeV
    - barrel: 0.00-1.44
    - endcap: 1.56-2.50
  - 10-15 GeV (underflow bin): MC-prediction is used without scaling
- to prepare yields, fit results in all fine barrel eta bins are summed up and fit results in all fine endcap eta bins are summed up
- templates for RooNDKeysPdf: $e \rightarrow \gamma$ portion of the DYjets MC separately for each pt-eta bin
- $e \rightarrow \gamma$ portion of the DYjets MC: a photon has a gen-level electron within dR=0.4

$M_{e,\gamma}$ Fit Plots. 15-20 GeV, barrel

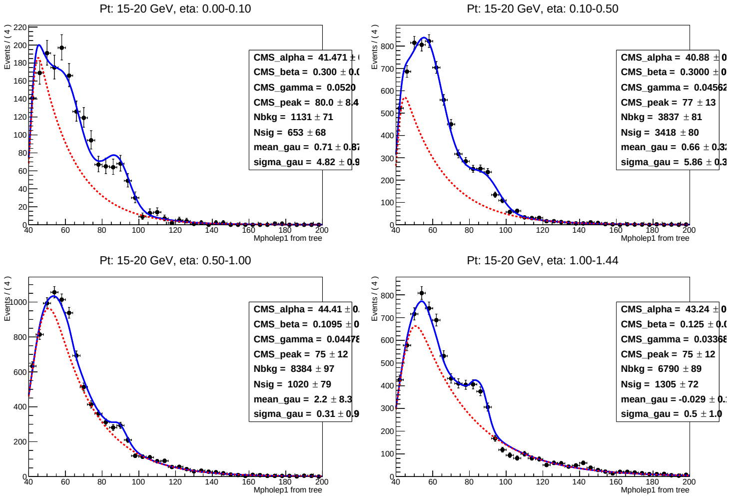


Figure: $M_{e,\gamma}$ fits, Wγ, electron channel, 15-20 GeV, barrel, 4 eta bins

$M_{e,\gamma}$ Fit Plots. 15-20 GeV, endcap

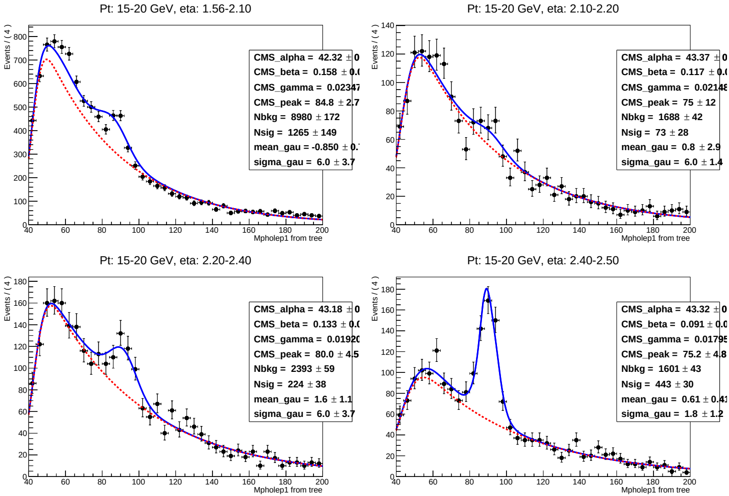


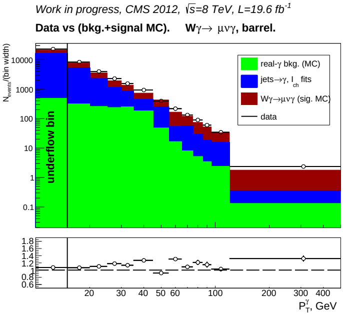
Figure: $M_{e,\gamma}$ fits, Wγ, electron channel, 15-20 GeV, endcap, 4 eta bins

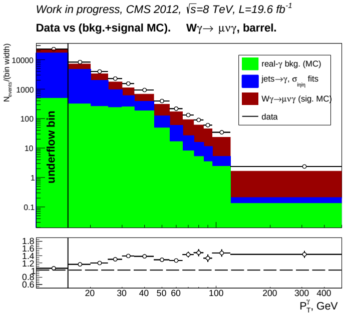
Other Backgrounds

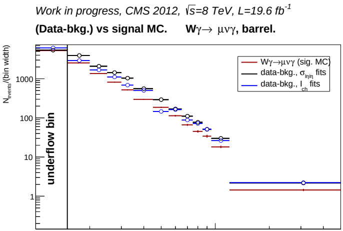
- true γ background. Main sources of true γ background are Zγ, W$\gamma \rightarrow \tau\nu\gamma$, WW$\gamma$. The MC-based estimation is used to subtract these backgrounds.
- $e \rightarrow \gamma$ background for muon channel. Sources of these backgrounds could be WW ($W \rightarrow \mu\nu_\mu + W \rightarrow e\nu_e$), WZ ($W \rightarrow \mu\nu_\mu + Z \rightarrow ee$ or $W \rightarrow e\nu_\mu + Z \rightarrow \mu\mu$) or ZZ ($Z \rightarrow \mu\mu + Z \rightarrow ee$). Negligible.
- *jets* $\rightarrow$ *lepton* + *true* γ (γ+jets and $\gamma\gamma$+jets events). Negligible.
- *jets* $\rightarrow$ *lepton* + *jets* $\rightarrow \gamma$. Negligible.

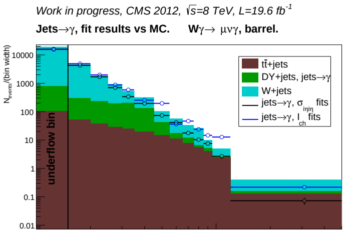
P_T^γ Spectrum before and after Background Subtraction. Muon Channel, Barrel

Top: data vs fake-γ background derived from the template method + real-γ background predicted by dedicated MC samples + signal MC, with I_{ch} and $\sigma_{i\eta i\eta}$ used as fit variables. Bottom: left: data yields after full background subtraction vs signal MC. I_{ch} vs $\sigma_{i\eta i\eta}$ fit results. Right: fake-γ data driven background prediction vs MC. Plotted with the stat error only.



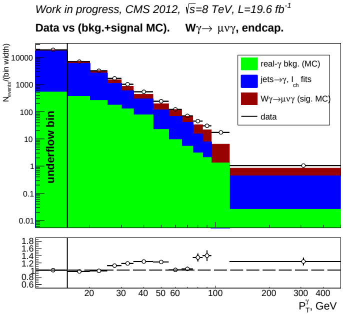


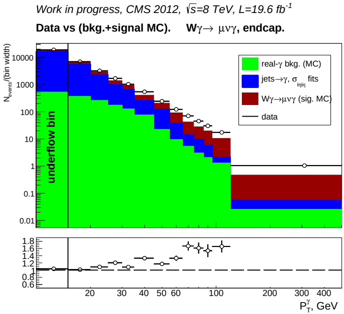


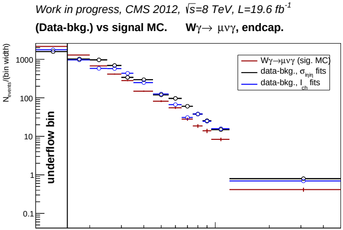


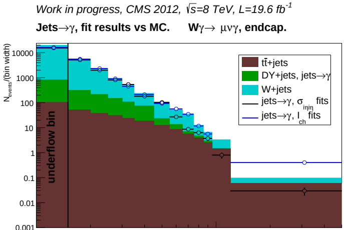
P_T^{γ} Spectrum before and after Background Subtraction. Muon Channel, Endcap

Top: data vs fake-γ background derived from the template method + real-γ background predicted by dedicated MC samples + signal MC, with I_{ch} and $\sigma_{i\eta i\eta}$ used as fit variables. Bottom: left: data yields after full background subtraction vs signal MC. I_{ch} vs $\sigma_{i\eta i\eta}$ fit results. Right: fake-γ data driven background prediction vs MC. Plotted with the stat error only.



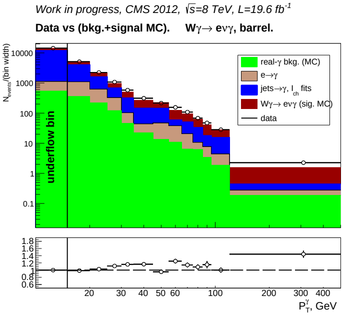


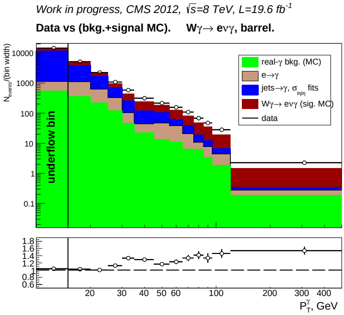


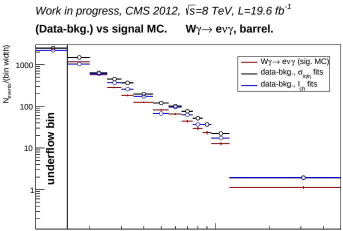


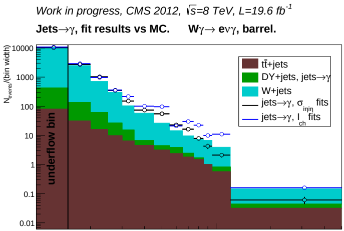
P_T^{γ} Spectrum before and after Background Subtraction. Electron Channel, Barrel

Top: data vs fake-γ background derived from the template method + real-γ background predicted by dedicated MC samples + signal MC, with I_{ch} and $\sigma_{i\eta i\eta}$ used as fit variables. Bottom: left: data yields after full background subtraction vs signal MC. I_{ch} vs $\sigma_{i\eta i\eta}$ fit results. Right: fake-γ data driven background prediction vs MC. Plotted with the stat error only.



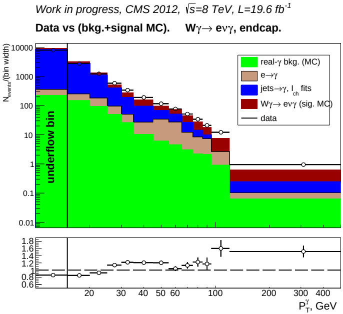


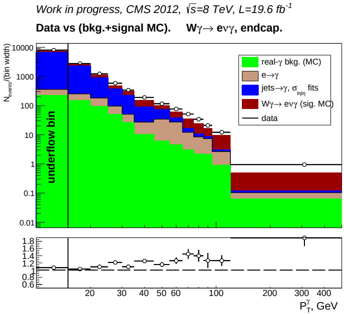


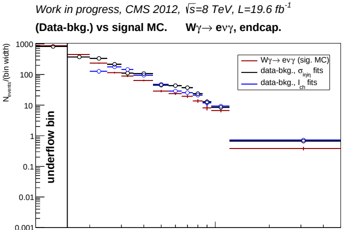


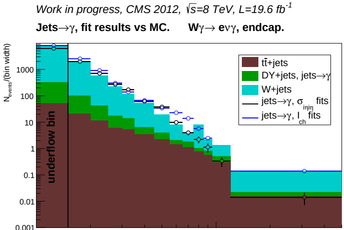
P_T^γ Spectrum before and after Background Subtraction. Electron Channel, Endcap

Top: data vs fake-γ background derived from the template method + real-γ background predicted by dedicated MC samples + signal MC, with I_{ch} and $\sigma_{i\eta i\eta}$ used as fit variables. Bottom: left: data yields after full background subtraction vs signal MC. I_{ch} vs $\sigma_{i\eta i\eta}$ fit results. Right: fake-γ data driven background prediction vs MC. Plotted with the stat error only.









Detector Resolution Unfolding. Muon channel

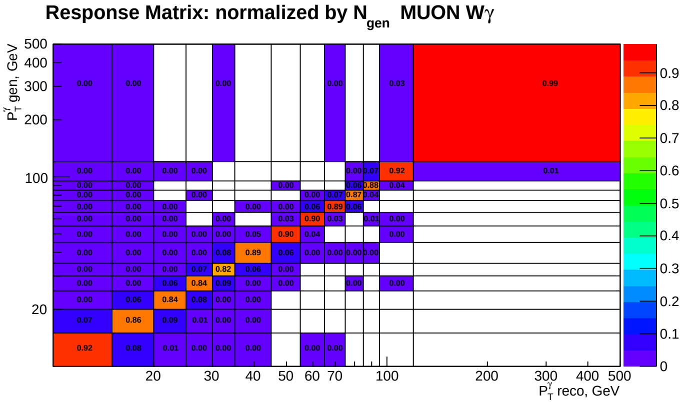
detector resolution unfolding: $N^i_{accXeff} = U_{ij} \cdot N^j_{bkg.subtr.}$

difference between $N^j_{bkg.subtr.}$ and $N^i_{accXeff}$ is P^γ_T binning only (reco vs true)

to scale results to the phase space (defined later) the bin-by-bin AccXEff correction is applied

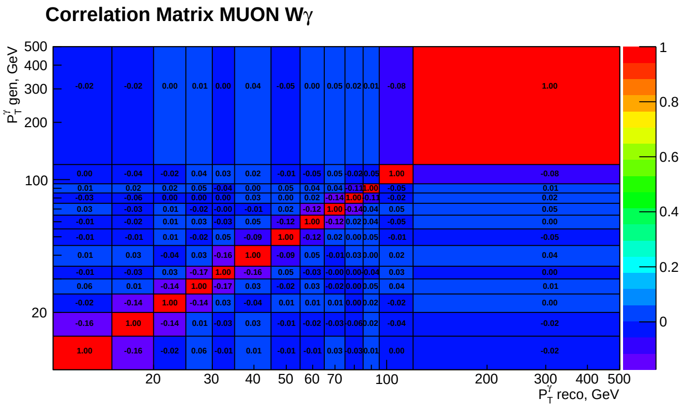
AccXEff correction: $N^i_{true} = \frac{N^i_{accXeff}}{\epsilon_i \cdot A_i}$

Unfolding constants are derived from signal MC sample ($W\gamma \to \mu\nu_\mu\gamma / W\gamma \to e\nu_e\gamma$) with the D'Agostini method using RooUnfold package



Correlation Matrices for Unfolded Yields. Muon channel

Correlation matrix on the unfolded yields (not on cross section), takes into account statistical error only (correlation matrices for the systematic errors are prepared separately).



Acc x Eff Correction. Method Description

- Computed as combined value bin-by-bin based on signal MC
- **Numerator:**
  - For total cross section: selected signal MC yields with PU weight applied
  - For differential cross section: selected signal MC yields with PU weight applied in p_T^γ GEN bins 15-20-25... Transition to p_T^γ RECO bins performed with unfolding
- **Denominator:**
  - For each event in unskimmed signal MC file, photon and lepton (two leptons for Zγ) which refer to Wγ (Zγ) and determine ΔR GEN and p_T^γ GEN
  - For total cross section: calculate number of events within phase space cut
  - For differential cross section: calculate number of events within phase space cut in 15-20-25... p_T^γ GEN bins

Systematic Uncertainties. List

- jets$\rightarrow \gamma$ background estimation
  - template statistics (randomize templates and take the RMS of the results as a error)
  - the whole difference $|N_{CHISO} - N_{SIHIH}|$
  - uncerntainty of Zγ MC normalization (4.6% as reported by the 8 TeV Zγ measurement); not part of weighted average
- M_T^W cut: vary Pt of all photons, electrons, jets available in the ggNtuple for each event as prescribed
- statistics of real-γ background MC samples; Wγ normalization (W$\gamma \rightarrow \tau\nu\gamma$) and Z$\gamma$ normalization
- statistics of $e \rightarrow \gamma$ enriched sample and all the other samples involved for the $e \rightarrow \gamma$ bck.estimation
- signal MC statistics for unfolding (randomize migration matrix, unfold for each, take RMS as an error)
- signal MC statistics for accXeff
- luminosity: 2.6%
- vary scale factors by their uncertainties (only check the variation effect on the accXeff and unfolding, not background estimation)
- vary pileup cross section by $\pm$5% (only check the effect on the accXeff and unfolding)

Systematic Uncertainties. Table. Muon channel

Table: Relative errors [%]. MUON WGamma

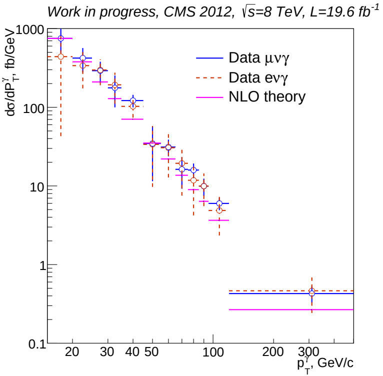
bin lims	err stat	syst $j \rightarrow \gamma$	Zg MC norm	real γ bkg	A.E. MCst	L	MET	PU	SF	unf MCst	syst tot
total	1	11	4	1	0	3	1	4	3	1	13
15-20	2	32	12	2	1	3	1	4	5	2	35
20-25	2	31	13	3	2	3	2	4	2	3	34
25-30	2	27	13	3	2	3	2	2	2	3	31
30-35	3	42	15	4	3	3	1	4	2	3	45
35-45	2	15	12	3	3	3	2	3	2	2	20
45-55	4	64	19	3	3	3	1	4	2	5	67
55-65	3	21	12	2	4	3	2	4	2	3	26
65-75	6	40	19	2	6	3	3	5	2	6	45
75-85	4	18	11	1	7	3	3	3	3	5	23
85-95	5	22	9	2	8	3	4	6	3	7	27
95-120	5	15	8	2	7	3	2	2	3	6	20
120-500	3	22	11	1	6	3	1	4	4	4	26

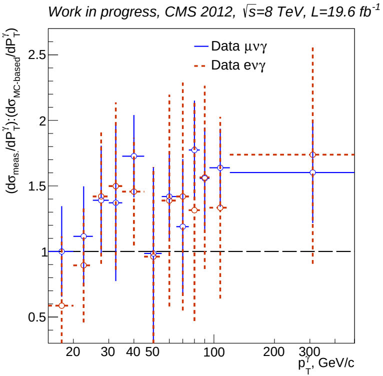
Systematic Uncertainties. Table. Electron channel

Table: Relative errors [%]. ELECTRON WGamma

bin lims	err stat	syst $j \rightarrow \gamma$	Zg MC norm	$e \rightarrow \gamma$ stat	real γ bkg	A.E. MCst	L	MET	PU	SF	unf MCst	syst tot
total	2	16	6	1	2	0	3	1	4	16	2	24
15-20	8	106	27	5	6	2	3	1	4	11	8	110
20-25	7	51	20	5	5	2	3	1	4	9	7	56
25-30	5	51	16	3	3	3	3	1	3	10	6	55
30-35	5	38	14	2	2	4	3	1	3	12	6	43
35-45	3	19	13	1	1	4	3	1	4	16	4	29
45-55	8	57	20	5	2	5	3	3	4	30	9	68
55-65	7	35	12	4	2	5	3	3	5	39	7	55
65-75	7	40	15	4	2	6	3	1	4	40	8	59
75-85	8	42	17	4	2	8	3	2	3	39	8	61
85-95	9	16	7	6	2	9	3	2	2	35	10	42
95-120	7	24	9	4	1	8	3	1	4	40	7	49
120-500	4	30	6	1	1	7	3	2	3	36	4	48

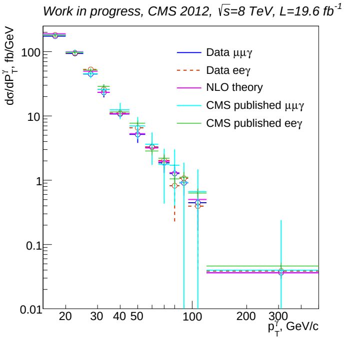
Differential Cross Section. Plots

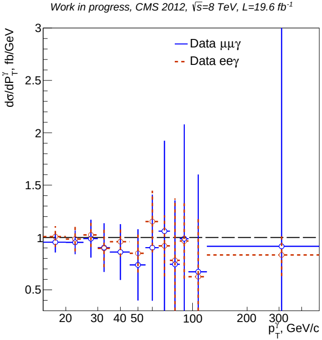




Total cross section ($P_T^{\gamma} > 15$ GeV):
MC-based: $\sigma = 9101$ fb
Measured, muon channel: $\sigma = 10949 \pm 91 \pm 1463$ fb
Measured, electron channel: $\sigma = 9146 \pm 185 \pm 2213$ fb

ZGamma check. Differential Cross Section. Plots





- Double-mu and double-ele datasets are used to perform the $Z\gamma$ check (the same samples as used to prepare templates for jets$\rightarrow \gamma$ background estimation)
- Total and differential cross section of $Z\gamma\gamma$ is measured and compared to the 8 TeV published CMS result
- The workflow for $Z\gamma$ is the same as $W\gamma$ except different selection criteria and also $Z\gamma$ has only jets$\rightarrow \gamma$ background
- For the muon channel, templates significantly overlay with the dataset, so, the result for the muon channel is a closure test rather that an actual measurement
- For the electron channel, templates and dataset are independent
- Intermediate plots are available in the AN (template fits, yields, data vs MC)
- The result agrees well with the 8 TeV published CMS result

Conclusions

- Cross section for muon and electron channels are computed;
- This is the first measurement of the differential Wγ cross section with CMS;
- Results agree with the MC prediction;
- Results between the two channels agree;
- Relative systematic uncertainties on total cross section are larger than those reported by the CMS 7 TeV measurement;
- Good agreement in the $Z\gamma$ check validates most parts of the analysis (those which are the same for the muon and electron channels).