Searching for Dark Matter with the CMS Detector in proton-proton collisions containing a single high-p_{T} photon and large $E_{\mathrm{T}}^{\mathrm{miss}}$

by

Brandon Leigh Allen

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

Doctorate of Philosophy in Physics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2019

Author . . . .
Department of Physics
May 18, 2019

Certified by . . . .
Christoph E.M. Paus
Professor
Thesis Supervisor

Accepted by . . . .
Nergis Mavalvala
Associate Department Head for Education

Searching for Dark Matter with the CMS Detector in proton-proton collisions containing a single high-p_T photon and large $E_\mathrm{T}^\mathrm{miss}$

by

Brandon Leigh Allen

Submitted to the Department of Physics
on May 18, 2019, in partial fulfillment of the
requirements for the degree of
Doctorate of Philosophy in Physics

Abstract

A search is conducted for new physics in final states containing a photon and missing transverse momentum in proton-proton collisions at $\sqrt{s} = 13$ TeV. The data collected by the CMS experiment at the CERN LHC correspond to an integrated luminosity of 35.9 inverse femtobarns. No deviations from the predictions of the standard model are observed. The results are interpreted in the context of dark matter production and limits on new physics parameters are calculated at 95% confidence level. For the two simplified dark matter production models considered, the observed (expected) lower limits on the mediator masses are both $950\,(1150)$ GeV for 1 GeV dark matter mass.

Thesis Supervisor: Christoph E.M. Paus
Title: Professor

Acknowledgments

This is the acknowledgements section. You should replace this with your own acknowledgements.

Contents

Chapter 1

The Monophoton Analysis

In this chapter, we discuss the search for dark matter produced in association with a single high-p_T photon. Our benchmark signal models are the vector and axial dark matter mediators discussed previously. However, many Standard Model (SM) processes are also capable of producing events with a single high-p_T photon and large $E_\mathrm{T}^\mathrm{miss}$, whether with real photons, other particles misidentified as photons, or unphysical photon signatures from various machine and detector processes. [However, many doesn't read well. I think this first paragraph deserves an actual introduction to the analysis and not an immediate dive into the details] [You immediately start talkinga bout the backgrounds. Would be better to put some diagrams and discussion about the final state. This you can tie in with the above comment.]

The most significant are the irreducible backgrounds, where the underlying physics process produces the exact same signature as the signal with only real and properly identified physics objects. In this case, the irreducible backgrounds are the associated production of a high-energy γ with either a Z boson that subsequently decays to a pair of neutrinos or a W boson that decays to a charged lepton and a neutrino, with the charged lepton outside of the detector acceptance. These two processes are denoted as $Z(\to \nu\overline{\nu})+\gamma$ and $W(\to \ell\nu)+\gamma$, respectively, and together they account for approximately 70% of the SM background, with 50% from the former and 20% from the latter. The modeling of the irreducible backgrounds is explained in Section 1.6.

Additional backgrounds arise from events where the candidate photon object is a

misidentified electron (Section 1.4) or an electromagnetic shower caused by hadrons (Section 1.5). The background events from electron misidentification are mostly W boson production ($W \rightarrow e\nu$), whereas those from hadron misidentification are due to multiple sources such as $Z(\rightarrow \nu\nu) + \text{jets}$ and QCD multijets with grossly mismeasured jet energy. Misidentification itself is rare, but because these processes have high cross sections, the amount of background is substantial, approximately 15% and 5% respectively. Since object misidentification rates depend on subtle details of the detector, the MC simulation often fails to accurately describe them. Therefore, the contributions from these background processes are estimated by employing data-driven techniques, where the misidentification rates are measured in data and applied to proxy samples with well-identified electrons or hadrons.

Finally, apparent large energy deposits in ECAL from non-collision processes mimic $\gamma + E_{\text{T}}^{\text{miss}}$ events and therefore need to be controlled. Known sources of such background include bremsstrahlung of beam halo or cosmic ray muons and anomalous ECAL energy deposits resulting from the interaction of particles in the ECAL photodetectors referred to as "ECAL spikes". These methods used to estimate the contributions from these processes are described in detail in Sections 1.7 and 1.8, respectively.

The estimates of the contributions from $Z(\rightarrow \nu\bar{\nu}) + \gamma$, $W(\rightarrow \ell\nu) + \gamma$, and beam halo processes are allowed to float in the fits to data performed to extract the potential signal contribution and set limits on new physics models. Meanwhile, for all other background processes, the yields in the fits are fixed to the estimates from data-driven methods or MC cross section calculation. [The above paragraph doesn't fit. You bring up fit out of nowhere and floating is jargon. I would just remove this and outline this when you do the signal extraction]

1.1   Dataset

The data sample was collected with a single-photon trigger: `HLT_Photon165_HE10`. This high-level trigger algorithm is relatively simple, only requiring at least one photon

candidate with $E_{\text{T}} > 165\,\text{GeV}$ reconstructed with a corresponding L1 seed. The photon candidate must also have $H/E < 0.1$ to discriminate against jets, where H/E is the ratio of HCAL to ECAL energy deposits in the central calorimeter tower corresponding to the candidate. The photon energy reconstructed at the HLT is less precise relative to that derived later in the offline reconstruction. Therefore, the online thresholds in the trigger on both H/E and E_{T}^{γ} are less restrictive than their offline counterparts.

The trigger efficiency is measured to be about 98% for events passing the analysis selection with $E_{\text{T}}^{\gamma} > 175\,\text{GeV}$ and the integrated luminosity of the analyzed data sample is $(35.9 \pm 0.9)\,\text{fb}^{-1}$ [1].

1.1.1 Trigger Efficiency

We measure the trigger efficiency by looking for trigger objects that match the candidate photon object in an appropriate data set because the trigger decisions are based on the existence of a single photon object in the event. A trigger object is the four-momenta of an object reconstructed at the trigger level that is used for making trigger decisions. A trigger object is matched to the candidate when their angular separation $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ is less than a certain threshold. For the photon candidate object, a line that connects the detector origin and the cluster position was used to define its direction.

The trigger efficiency measurement is performed on the SingleMuon data set, exploiting events mostly from leptonic $t\bar{t}$ ($e\mu$) topology. Events with a candidate-quality photon object without the pixel seed veto requirement (e.g. a physical electron) and a muon object that passes the "tight" identifictaion requirement and matches the trigger object of the `HLT_IsoMu24` or `HLT_IsoTkMu24` triggers are used. The matching rate of the photon object and the trigger object is the trigger efficiency. Figure 1-1 shows the L1+HLT combined efficiency as a function of the photon E_{T}. It can be seen that the trigger is fully efficient for $E_{\text{T}} > 175\,\text{GeV}$.

For the first period of data taking, the `HLT_Photon165_HE10` trigger was seeded only by an isolated e/γ L1 trigger. This L1 seed becomes inefficient at high E_{T} due
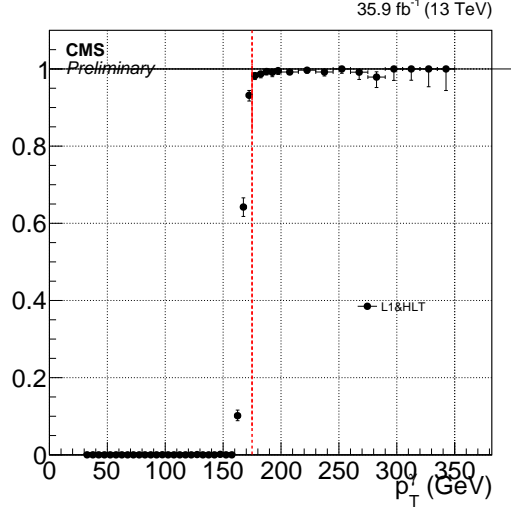
Figure 1-1: The efficiency turn-on of the `HLT_Photon165_HE10` trigger for photons passing the candidate selection, measured using $\mu + e/\gamma$ events from the SingleMuon data set. Red vertical line corresponds to $E_\mathrm{T} = 175\,\mathrm{GeV}$.

to a misconfiguration in the H/E computation algorithm as indicated by the drop in efficiency at high-E_T shown in the left side of Figure 1-2. To mitigate the effect, in the later periods, the trigger was seeded by the logical **OR** of `SingleEG40` and `SingleJet` L1 triggers, combining multiple with various p_T thresholds. [multiple is not specific enough, we need to know all the triggers so we can reproduce it]

Even with this addition, the measured trigger efficiency is not 100% at the plateau, but it is flat with respect to E_T as shown on the right of Figure 1-2. In principle, the efficiency should be applied to all simulation-based background estimates whose normalization is fixed by theoretical calculation of the cross section. However, the only simulation-based background processes with absolute normalization are those that contribute at $\mathcal{O}(1)\%$, with large systematic uncertainties. Therefore we deem the slight discrepancy of the trigger efficiency from unity as irrelevant.

1.1.2 Pileup Reweighting

The distribution of the number of pileup interactions inserted into MC events differ from the true pileup distribution, estimated from the measurement of instantaneous luminosity, beam intensity of each proton bunch, and the total cross section of proton
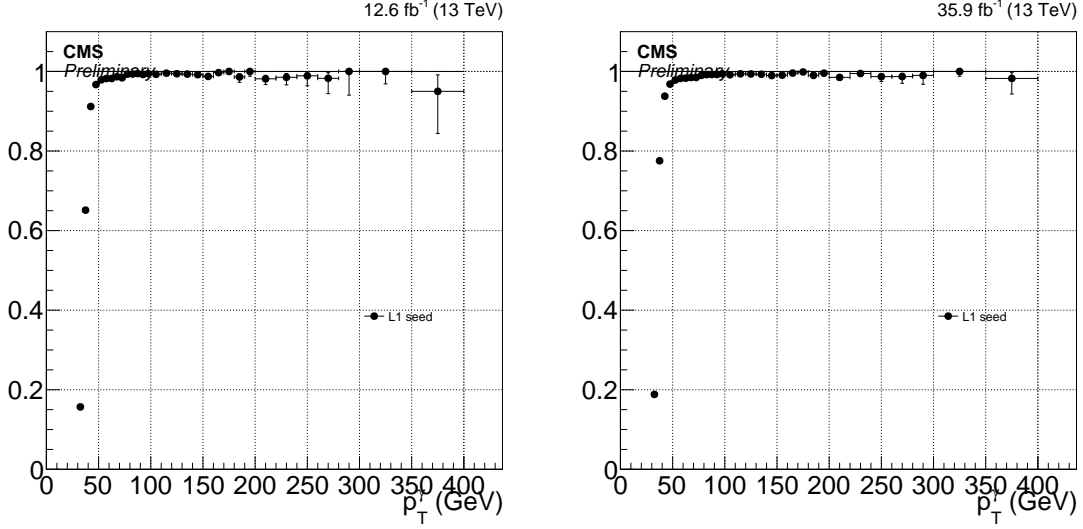
Figure 1-2: The efficiency of the L1 seed for the signal trigger in periods B and C (left) and the full data set (right). The drop in efficiency at high-E_{T} in the earlier period is fixed by the addition of `SingleJet` L1 seeds during the remainder of data-taking.

inelastic scattering $(69.2\ \mathrm{mb}^{-1})$. [we should know htis calculation each proton bunch multiplied by the total cross...]

Figure 1-3 shows the pileup distributions in data and MC and the ratio of the two. Each simulated MC event has its weight multiplied by the value of the ratio evaluated at the number of true pileup interaction injected into the event in order to reproduce the observed pileup distribution in the MC samples.

1.2 Event Selection

From the recorded data, events are selected by requiring at least one photon with $E_{\mathrm{T}}^{\gamma} > 175\,\mathrm{GeV}$ in the fiducial region of the ECAL barrel ($|\eta| < 1.44$). Events with photons in the endcaps are not considered because the estimation of backgrounds from beam halo and misidentified hadrons is greatly complicated because the crystals in the ECAL endcaps do not form a grid in η and ϕ.

To reduce the hadron misidentification rate, we require the leading photon to pass the the e/γ ID, a collection of isolation and shower shape selections given in
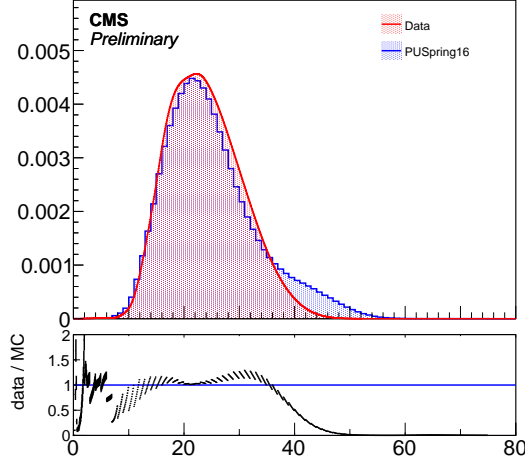
Figure 1-3: The pileup distributions in data and MC. The ratio of the two is used to reweight the MC samples.

Table **??**. Furthermore, to reduce the electron misidentification rate and reject non-collision backgrounds, we require the leading photon to pass the γ-specific ID, which consists of the pixel seed veto and additional selections given Table **??**.

Events with a high-p_T photon are subjected to further requirements to suppress SM background processes that feature a genuine high-energy photon, but not a significant amount of $E_\mathrm{T}^\mathrm{miss}$. As shown on the left of Figure 1-4, the main SM process is γ+jets, mismeasuring the energy of a jet results in a sizeable amount of artificial $E_\mathrm{T}^\mathrm{miss}$. To reject the vast majority of the γ+jets background, we require $E_\mathrm{T}^\mathrm{miss} > 170\,\mathrm{GeV}$. Looking at the right of Figure 1-4, we see that the $E_\mathrm{T}^\mathrm{miss}$ is typically smaller than E_T^γ for the remaining γ+jets events while the $E_\mathrm{T}^\mathrm{miss}$ and E_T^γ are approximately equal for signal-like processes. Thus, we require the ratio of E_T^γ to $E_\mathrm{T}^\mathrm{miss}$ to be less than 1.4 to reject this remainder of this background effectively with little effect on signal efficiency.

To protect against artificial $E_\mathrm{T}^\mathrm{miss}$ from jet mismeasurement, events are rejected if the minimum opening angle between $\vec{p}_\mathrm{T}^\mathrm{miss}$ and the directions of the four highest p_T jets, $\min\Delta\phi(\vec{p}_\mathrm{T}^\mathrm{jet}, \vec{p}_\mathrm{T}^\mathrm{miss})$, is less than 0.5; only jets with $p_\mathrm{T} > 30\,\mathrm{GeV}$ and $|\eta| < 5$ are considered in the $\min\Delta\phi(\vec{p}_\mathrm{T}^\mathrm{jet}, \vec{p}_\mathrm{T}^\mathrm{miss})$ calculation. Rare pathological mismeasurements of E_T^γ also lead to large $E_\mathrm{T}^\mathrm{miss}$. For this reason, the candidate photon $\vec{p}_\mathrm{T}$ and $\vec{p}_\mathrm{T}^\mathrm{miss}$
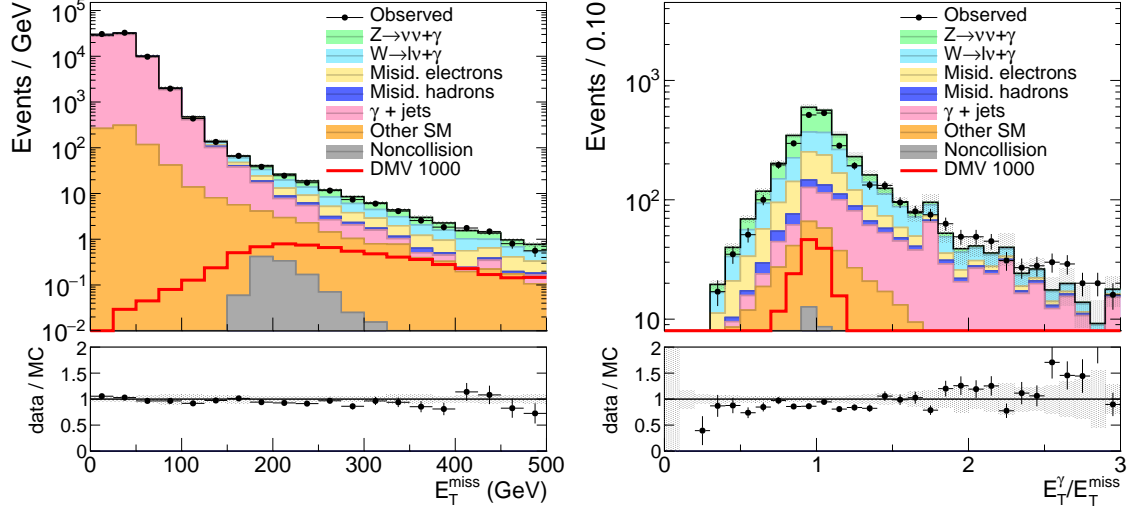
14

Figure 1-4: Left: The E_T^{miss} distribution in the signal region before applying any selection on E_T^{miss}. The γ+jets background dominates below $150\,\text{GeV}$, while the DM signal and irreducible backgrounds dominate in the tails. Right: The ratio of E_T^γ to E_T^{miss} after requiring $E_T^{miss} > 170\,\text{GeV}$. Signal-like events are clustered near 1 while γ+jets events dominate at values above 1.4.
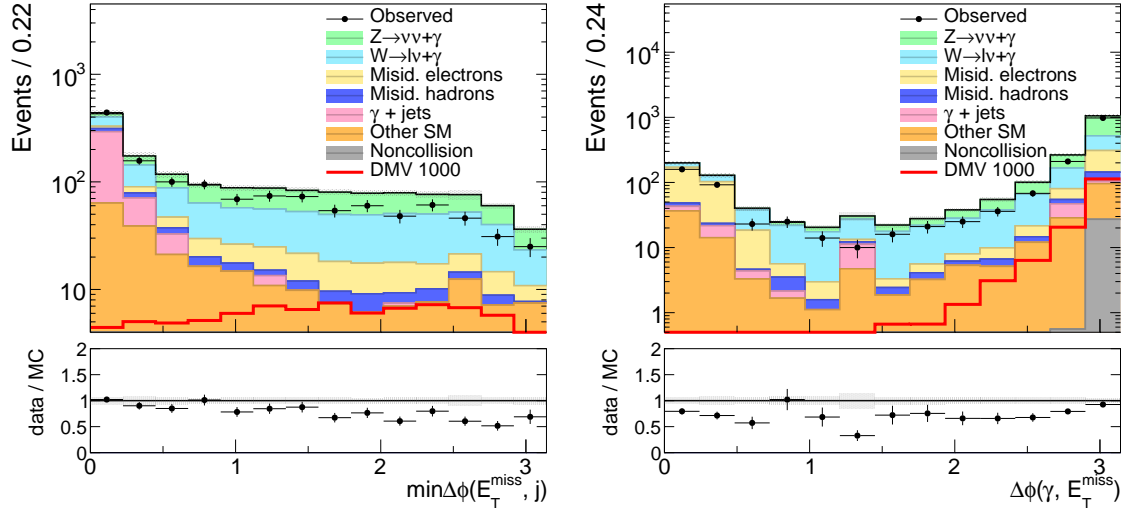

Figure 1-5: Signal region N-1 plots for the $\min\Delta\phi(\vec{p}_T^{\,\text{jet}}, \vec{p}_T^{\,\text{miss}})$ (left), and $\Delta\phi(\gamma, p_T^{miss})$ (right) selections. The requirement that the angular separations exceed 0.5 radians was chosen to reduce the rate of events with artificial E_T^{miss} from energy mismeasurement of jets and photons, respectively.

must be separated by more than 0.5 radians. Figure 1-5 shows the N-1 plots in the signal region for these two angular selections.

Additionally, events are vetoed if they contain an electron or a muon with $p_\mathrm{T} > 10\,\mathrm{GeV}$ that is separated from the photon by $\Delta R > 0.5$. This lepton veto rejects SM processes that produce a high-p_T photon, $E_\mathrm{T}^\mathrm{miss}$, and leptons such as $W(\to \ell\nu)+\gamma$, $t\bar{t}\gamma$, and $VV\gamma$.

Variable	Selection	Motivation
E_T^γ	$> 175\,\mathrm{GeV}$	high-p_T photon passing trigger
$|\eta|$	< 1.44	region with best background estimates
e/γ ID	Pass	reject hadronic background
γ-specific ID	Pass	reject electron and non-collision backgrounds
$E_\mathrm{T}^\mathrm{miss}$	$> 170\,\mathrm{GeV}$	characteristic signature of dark matter
$E_\mathrm{T}^\gamma / E_\mathrm{T}^\mathrm{miss}$	< 1.4	reduce jet mismeasurement backgrounds
$\min\Delta\phi(\vec{p}_\mathrm{T}^\mathrm{jet}, \vec{p}_\mathrm{T}^\mathrm{miss})$	> 0.5	reduce jet mismeasurement backgrounds
$\Delta\phi(\gamma, p_\mathrm{T}^\mathrm{miss})$	> 0.5	reduce photon mismeasurement backgrounds
Lepton Veto	Pass	reject backgrounds with real leptons

Table 1.1: Selections for the signal region.

The above selections, summarized in Table 1.1, constitute the signal selection for the analysis.

The contributions from the $Z(\to \nu\bar{\nu})+\gamma$ and $W(\to \ell\nu)+\gamma$ processes to the signal region are modeled by fitting to observed data in control regions where one or two leptons (electrons or muons) are identified in addition to the photon candidate. Meanwhile, the contributions from misidentified electrons and hadrons are modeled by proxy regions where some of the selections in the photon identification have been inverted. The additional requirements for the signal and various control regions used in the analysis are described in the following sections.

1.2.1 Signal Regions

Furthermore, to constrain the beam halo normalization, the signal region is split into two parts according to the variable ϕ' introduced in Equation 1.9. The region defined by $|\phi'| < 0.5$ is called the horizontal region, its complement $0.5 < |\phi'| < \pi/2$ is called the vertical region, and the two together are referred to as the combined signal

regions. [I think its fine to discuss the splitting of the signal region into categories later. Here, it just makes it hard to follow]

1.2.2 Proxy Samples

[Again this could go into a big table] To estimate the background due to misidentified electrons, an electron proxy sample is used. This proxy sample is obtained by identical event selection as that of the signal region but with the pixel-seed veto inverted on the photon candidate object. Such a photon candidate is referred to as a electron proxy object. This yields a sample of events with similar kinematics to the signal region and well-identified electron candidates, differing only from the misidentified electron events in that a pixel hit was associated with the photon object. Thus, these exact events are used to estimate the misidentified electron background after scaling them by the electron-to-photon misidentification rate.

To estimate the background due to misidentified hadrons, a hadron proxy sample is used. This proxy sample is obtained by identical event selection as that of the signal region but where the photon candidate passes the e/γ and γ-specific IDs with except for at least one of the following cuts: $\sigma_{i\eta i\eta} < 0.01022$ and $I_{\mathrm{CH}} < 0.441\,\mathrm{GeV}$. Such a photon candidate is referred to as a hadron proxy object. This yields a sample of events with similar kinematics to the signal region and well-identified proxies for misidentified hadrons. Thus, these exact events are used to estimate the misidentified hadron background after scaling them by the hadron-to-photon misidentification rate.

Additional tight and loose hadron proxy objects and samples are made by tightening and loosening the constant term in the I_{NH} and I_γ requirements on the proxy object. The specific values for each proxy object are shown in Table 1.2.

	I_{NH} (GeV)	I_γ (GeV)
Nominal	2.792	2.176
Loose	10.910	3.630
Tight	0.264	2.362

Table 1.2: Constant terms in the I_{NH} and I_γ selections for the hadron proxy objects.

1.2.3 Measurement Samples

[here again, this section is not very clear]

To measure the photon purity and part of the photon efficiency, an EM object+jet measurement sample is formed by requiring an EM object with $E_T > 175\,\text{GeV}$ and $|\eta| < 1.44$ plus at least one jet with $p_T > 100\,\text{GeV}$ and $|\eta| < 2.5$ which passes the loose jet ID. An EM object is a photon candidate that passes the e/γ ID with the execption of the following relaxed cuts: $\sigma_{i\eta i\eta} < 0.015$ and $I_{\text{CH}} < 11.0\,\text{GeV}$. Additionally, we apply an $E_T^{\text{miss}} < 60\,\text{GeV}$ cut to make this region orthogonal to the signal region.

To measure the hadron misidentification rate, a hadron proxy+jet measurement sample is formed by replacing the the EM object in the EM object+jet sample with a hadron proxy object, one for each type of hadron proxy. The hadron proxy+jet samples have the exact same selection as the hadron proxy samples, except that a high-p_T jet is required instead of high-E_T^{miss}, minimizing the kinematic differences between the two.

1.3 Efficiencies and Scale Factors

[I think it might be better to put this before the previous section or at least merge them. The prevous section is out of place.] While we try to model the CMS detector as accurately as possible with our MC simulations, there are still differences between the behavior of physics objects within the simulations and those from data taken with the detector. Most importantly, this results in different efficiencies for photons and leptons in data and MC, which we must measure. To improve our MC, we reweight it by the ratio of the efficiency in data to that in MC, known as the scale factor.

When measuring the scale factor for photons, we factorize the photon ID into the e/γ portion and the γ-specific portion. The e/γ portion of the ID consists of a collection of isolation and shower shape selections designed to reduce the hadron misidentification rate. We measure the efficiency of the e/γ portion using the "tag-and-probe" (TP) method with $Z \to ee$ events as these variables have similar efficiencies for physical electrons and photons. The γ-specific portion of the ID consists of the pixel seed

veto and non-collision rejection cuts. We measure the efficiency of γ-specific portion on a sample of physical photons in the EM object+jet measurement sample using a $\sigma_{i\eta i\eta}$ template fit method.

We perform both efficiency estimates as a function of p_{T} with the binning [175,200], [200,250], [250,300], [300,350], [350,400] and [400,∞). This binning was chosen based on the number of available events in data for the failing probes fit in the TP method and the background template for the $\sigma_{i\eta i\eta}$ fits, as these samples are the smallest and drive the uncertainty of the methods.

1.3.1 e/γ ID Efficiency

The efficiency corresponding to the e/γ part of the photon ID is estimated by exploiting Z boson decays into pairs of electrons and positrons. Using the TP method, a high-quality electron object (tag) is identified in a single photon data sample, and the accompanying electron is sought for in the pool of elecromagnetic objects (probes) in the event. The area under the peak in the mass distribution of the tag-probe system around the Z boson mass (between 81 GeV and 101 GeV) is then measured once applying the $e\gamma$ ID requirements on the probe and once inverting all of the requirements simultaneously. Denoting the two areas under the peaks in the passing and failing samples N_{pass} and N_{fail}, respectively, the resulting effiency $\epsilon_{e/\gamma}$ is given by

$$\epsilon_{e/\gamma} = \frac{N_{\mathrm{pass}}}{N_{\mathrm{pass}} + N_{\mathrm{fail}}}. \tag{1.1}$$

The TP measurement is performed on a subset of the single photon triggered events where there is an electron object (tag) passing the "tight" identification criteria in addition to the triggering photon (probe). All possible tag-probe combinations are considered; if the tag object can also serve as a probe and the probe object as a tag, which is a common occurrence in the case when the probe is electron-like (passes the $e\gamma$ ID), then the two combinations are considered independently to avoid the bias caused by preferring to use one object over another as the probe.

The tag-probe mass distributions are then fit to extract N_{pass} and N_{fail}. The fit

model is composed of two templates, where one template describes a pure $Z \rightarrow ee$ line shape and the other describes the background contributions. The main backgrounds to the fits are those with real electrons such as W+jets, diboson, and $t\bar{t}$ productions. Minor contributions from processes that do not involve true electrons, such as diphoton production with a strongly asymmetric conversion on one of the photons and misidentification of a QCD jet as an electron, are predicted to be negligible from MC studies.

The $Z \rightarrow ee$ template shape is an analytic convolution of the Breit-Wigner distribution and the Crystal Ball function. The mass and width parameters of the Breit-Wigner distribution are fixed to PDG values while the Crystal Ball parameters are allowed to float in the fit. We are able to use the analytic Breit-Wigner distribution instead of a template taken from MC because at this high probe p_{T} scale the selected events are mostly of the Z+jets topology with a boosted Z boson. This makes the selection rather inclusive in terms of the tag-probe invariant mass and ensures that the Breit-Wigner distribution accurately models the mass distribution even though the tag and probe are under kinematically exclusive selections.

The background template is taken from events collected by the single photon trigger where an additional muon object is present, making use of the fact that the most of the background processes in both fits are symmetric in lepton flavor. In order to mitigate statistical fluctuations in the background sample, the actual template is constructed by a Gaussian kernel estimation of the mass distribution of this muon-probe sample.

The floating parameters of the fits are therefore the normalizations of the $Z \rightarrow ee$ and background templates and the Crystal Ball smearing parameters. Selected example fits are shown in Figure 1-6.

The statistical uncertainty of the fits is estimated by generating toy data from the nominal fit result with the same number of entries as the fit target distribution. The mass distribution of the toy data is then fit with the same model with the parameters floating. This procedure is repeated 100 times to obtain a distribution of the $Z \rightarrow ee$ event yields, and its standard deviation is taken as the statistical uncertainty of the
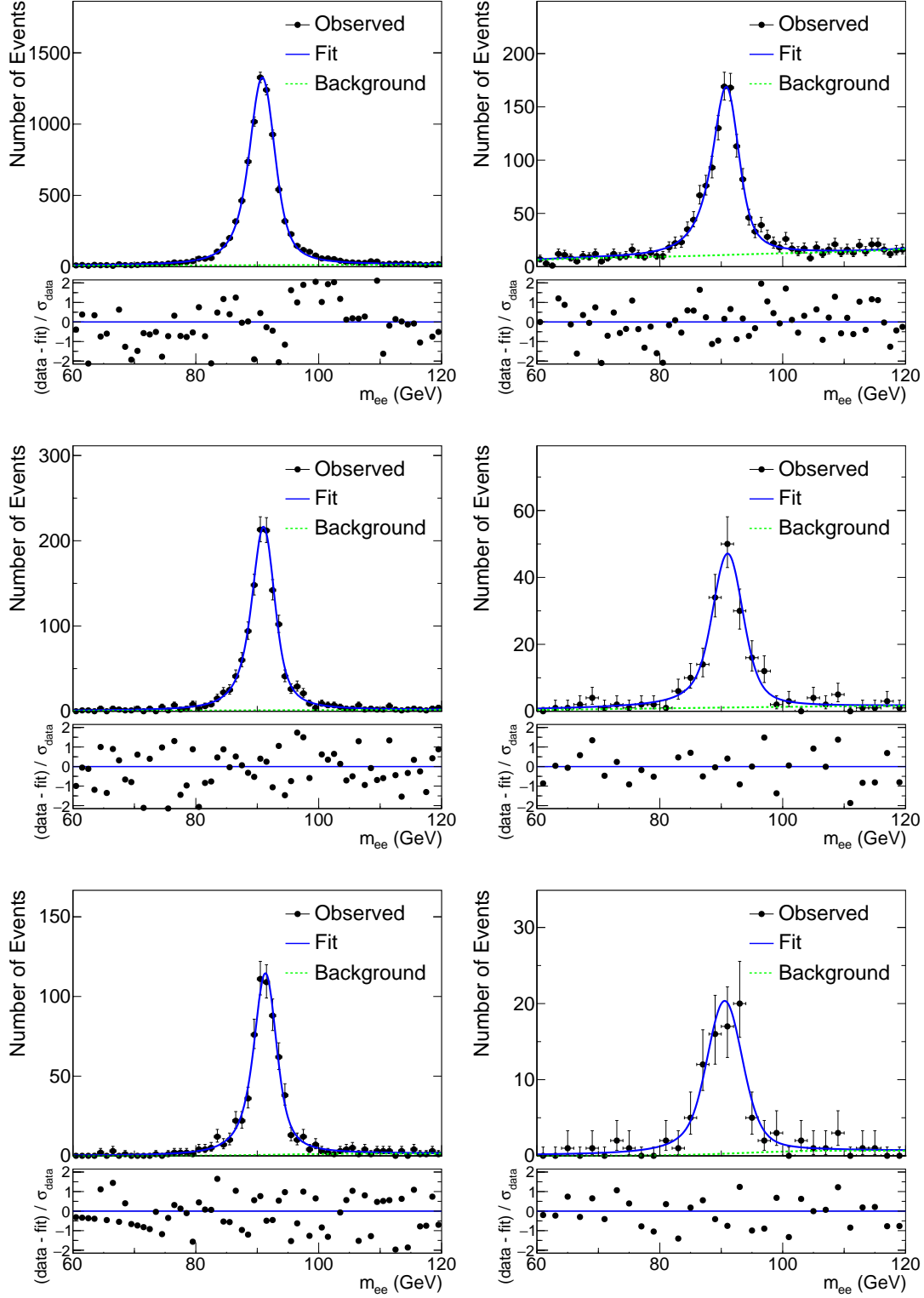
20

Figure 1-6: Fits to the mass distributions for pass (left) and fail (right) selections, in bins of probe p_{T}: $175 < p_{\mathrm{T}} < 200\,\mathrm{GeV}$ (top), $300 < p_{\mathrm{T}} < 350\,\mathrm{GeV}$ (middle), $p_{\mathrm{T}} > 400\,\mathrm{GeV}$ (bottom). The blue solid line represents the full fit model, and the green dashed line its background component.

fit. The relative statistical uncertainty on the yields is at most 10% for the fits in the high-p_T bins and decreases to the sub-percent level for the lowest p_T bins.

To estimate the effect of potential mismodeling in the fits, alternative fits varying the background and signal templates are performed first. In the alternative-background fit, a simple linear function is tested. In the alternative-signal fit, no Crystal Ball convolution is performed to the signal template and the mass and width of the Breit-Wigner function are allowed to vary. Resulting best-fit distributions of these alternative models are then used to generate a large number of toy distributions, which are fit by the nominal model. The average shift of the fit result from the nominal value is then taken as the uncertainty. The relative uncertainty on the efficiency varies from 2 to 4% depending on the probe p_T bin.

The MC efficiency is taken from counting the number truth-matched electrons passing and failing the e/γ part of the ID from a $Z \to ee$ sample. Additionally, the MC efficiency is computed using the same procedure as in data as a cross-check. The efficiencies obtained from these two methods are consistent within their uncertainties.

$p_\mathrm{T}^\mathrm{probe}$ (GeV)	MC Fit	Truth
(175, 200)	1.014 ± 0.008	1.009 ± 0.016
(200, 250)	1.003 ± 0.008	0.999 ± 0.014
(250, 300)	1.014 ± 0.010	1.016 ± 0.019
(300, 350)	1.002 ± 0.014	0.997 ± 0.022
(350, 400)	0.986 ± 0.012	0.987 ± 0.022
(400, 6500)	0.988 ± 0.011	0.999 ± 0.016

Table 1.3: e/γ scale factors as a function of photon p_T.

The data efficiencies, MC efficiencies, and resulting scale factors as a function of p_T are shown in Figure 1-7. The scalefactors are consistent with unity within the uncertainties. The numerical values are given in Table 1.3. We use the bin by bin scale factor corresponding to the truth values in the analysis.

1.3.2 γ-specific ID Efficiency

To measure the efficiency of the γ-specific component of the photon ID, we use a $\sigma_{i\eta i\eta}$ template fit to extract the number of true photons from a pool of photon objects
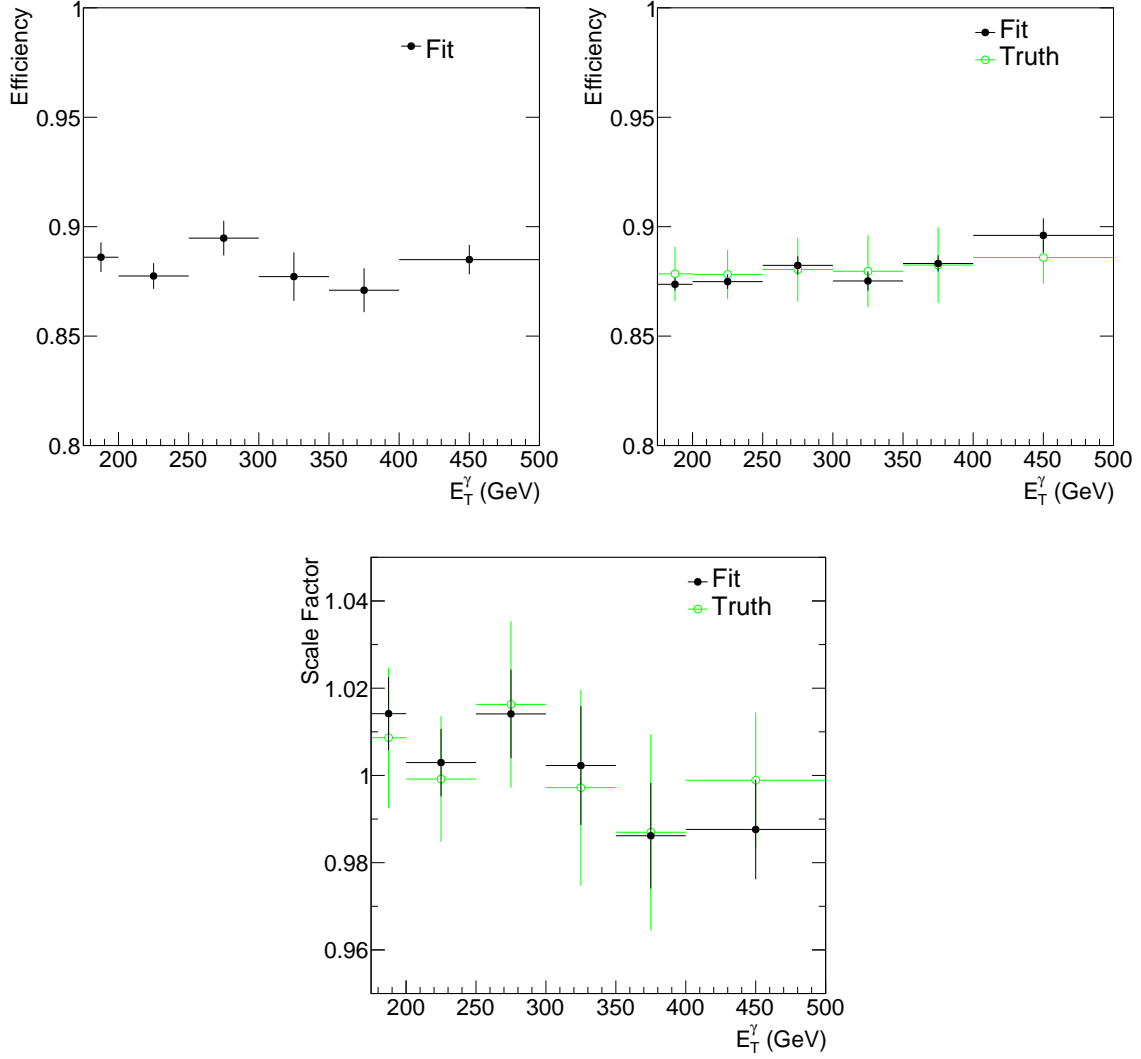
Figure 1-7: The e/γ component of the photon identification efficiency for data (top-left) and MC (top-right) and corresponding scale factor (bottom) as a function of photon p_{T}. Values derived from the fit methodology are shown in black and values derived from truth counting in MC are shown in light green.

passing the e/γ ID.

The measurement is performed using the EM object+jet measurement sample. We fit the $\sigma_{i\eta i\eta}$ distribution of the EM object with a template describing the $\sigma_{i\eta i\eta}$ shape of true photons and another describing the hadronic background. The real photon template is taken from γ+jets MC requiring the photon to pass the e/γ ID except for the $\sigma_{i\eta i\eta}$ requirement. The fake photon template is taken from the same data control sample, requiring $5\,\mathrm{GeV} < I_{\mathrm{CH}} < 7\,\mathrm{GeV}$. The number of true photons in the target sample is the integral of the post-fit real photon template below $\sigma_{i\eta i\eta} = 0.0104$, the same value as in our photon ID.

The fit is performed once for all EM objects and then once for EM objects passing the γ-specific ID criteria. The ratio of the numbers of true photons obtained in the two fits is the efficiency.

The $\sigma_{i\eta i\eta}$ template fit method in its simplest form fits the observed distribution with the following fit function:

$$P(f; \sigma_{i\eta i\eta}) = f \cdot h_s(\sigma_{i\eta i\eta}) + (1 - f) \times h_b(\sigma_{i\eta i\eta}), \qquad (1.2)$$

where h_s is the signal template, h_b is the background template, and f is the fraction of true photons in the target sample. Both the target template and the fit function are normalized to unity, removing the number of photon candidates in the target sample N as a fit parameter and leaving f as the only free parameter.
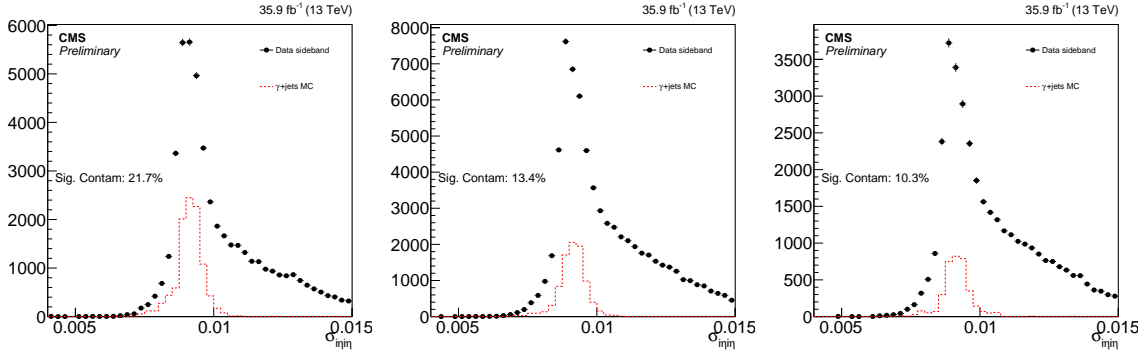

Figure 1-8:   Signal contamination in the [3.5,5.0] (left), [5.0,7.5] (middle), and [7.5,9.0] (right) isolation sidebands.

However, the hadronic background template, taken from the data control sample, has contributions from real photons with a I_{CH} value exceeding the ID requirements. The amount of this "photon contamination" depends on the sideband choice, but is non-zero even for a sideband with very large I_{CH}. As described below, we perform additional fits with the background templates from alternative sidebands $3.5\,\text{GeV} < I_{\text{CH}} < 5\,\text{GeV}$ ("near") and $7.5\,\text{GeV} < I_{\text{CH}} < 9\,\text{GeV}$ ("far") to assess the systematic uncertainty. The photon contamination of the nominal and far sideband is 10-15%, and in the near sideband, it can go up to approximately 20% (see Figure 1-8).

To remove the photon contamination from the background templates, we start with the true photon shape in the sideband $h_{s'}$, which differs from the signal template h_s in the I_{CH} selection applied to the photons. Then, we create a new background template h_b^{sub} from the original background template h_b by subtracting $h_{s'}$. After normalization to unity, we obtain the expression

$$h_b^{\text{sub}}(\sigma_{i\eta i\eta}) = \frac{h_b(\sigma_{i\eta i\eta}) - S'/B \cdot h_{s'}(\sigma_{i\eta i\eta})}{1 - S'/B}, \tag{1.3}$$

where B is the number of photon candidates in the sideband and S' is the number of true photons in the sideband.

To determine S', we start with the number of true photons in the target sample, $f \cdot N$. We then scale this by the ratio of the relative fractions of true MC photons in the I_{CH} sideband r_{sb} and in the signal region r_{sig}, giving us the expression

$$S' = f \cdot \frac{r_{\text{sb}}}{r_{\text{sig}}} \cdot N. \tag{1.4}$$

Going back to our original fit function and replacing h_b with h_b^{sub} gives us

$$P(f; \sigma_{i\eta i\eta}) = f \cdot h_s(\sigma_{i\eta i\eta}) + (1 - f) \times \frac{h_b(\sigma_{i\eta i\eta}) - S'(f)/B \cdot h_{s'}(\sigma_{i\eta i\eta})}{1 - S'(f)/B}, \tag{1.5}$$

which converges to the original fit function if $S' = 0$, i.e., if there is no photon contamination in the sideband. Note that f is still the only free parameter for this new function as S' only depends on f and $r_{\text{sb}}/r_{\text{sig}}$ is set constant in the fit (see

discussion of systematics for more detail).

There are four main sources of systematic uncertainty for this measurement. The first comes from the sideband choice, as the relative rates of different types of fake photons varies with I_{CH}. The second comes from the true photon I_{CH} shape, as this is used to determine the normalization of true photons in the sideband. Currently, this shape is taken from MC and thus there is the potential to mismodel the effects of the underlying event and pile-up. The third comes from the true photon $\sigma_{i\eta i\eta}$ distribution. As we take this from MC as well, we can mismodel the signal template shape. Finally, at high p_{T}, we suffer from low yields in our I_{CH} sidebands, which leads fluctutations that negatively influence the fit.

The uncertainty due to sideband choice is the larger of the differences of the purities measured using the near and far sidebands versus the nominal sideband. Figure 1-9 shows fits using the three sidebands for the [175,200] p_{T} bin on the left and for the [400,∞) p_{T} bin on the right.

To measure the uncertainty due to the I_{CH} shape, we look at the I_{CH} for electrons in $Z \to ee$ events in both data and MC. Using these distributions, we obtain a data/MC scale factor which we apply to the MC true photon I_{CH} distribution to obtain a scaled MC distribution. Then, we recount the photons using this new distribution and take the difference in the values obtained using the raw MC and scaled MC distributions as a systematic uncertainty.

To measure the uncertainty due to the signal template $\sigma_{i\eta i\eta}$ shape, we look at the $\sigma_{i\eta i\eta}$ distributions for electrons from $Z \to ee$ events in both data and MC. Using these distributions, we obtain a data/MC scale factor which we apply to the MC true photon $\sigma_{i\eta i\eta}$ distribution to obtain a scaled MC distribution. Then, we recount the photons using this new distribution and take the difference in the values obtained using the raw MC and scaled MC distributions as a systematic uncertainty.

To estimate the uncertainty due to statistical fluctuations in our background templates, we generate toys from the background template from data. We then repeat the fit with each of these toys and plot the distribution of the difference between the purity value obtained from the toy templates versus the nominal template. We take
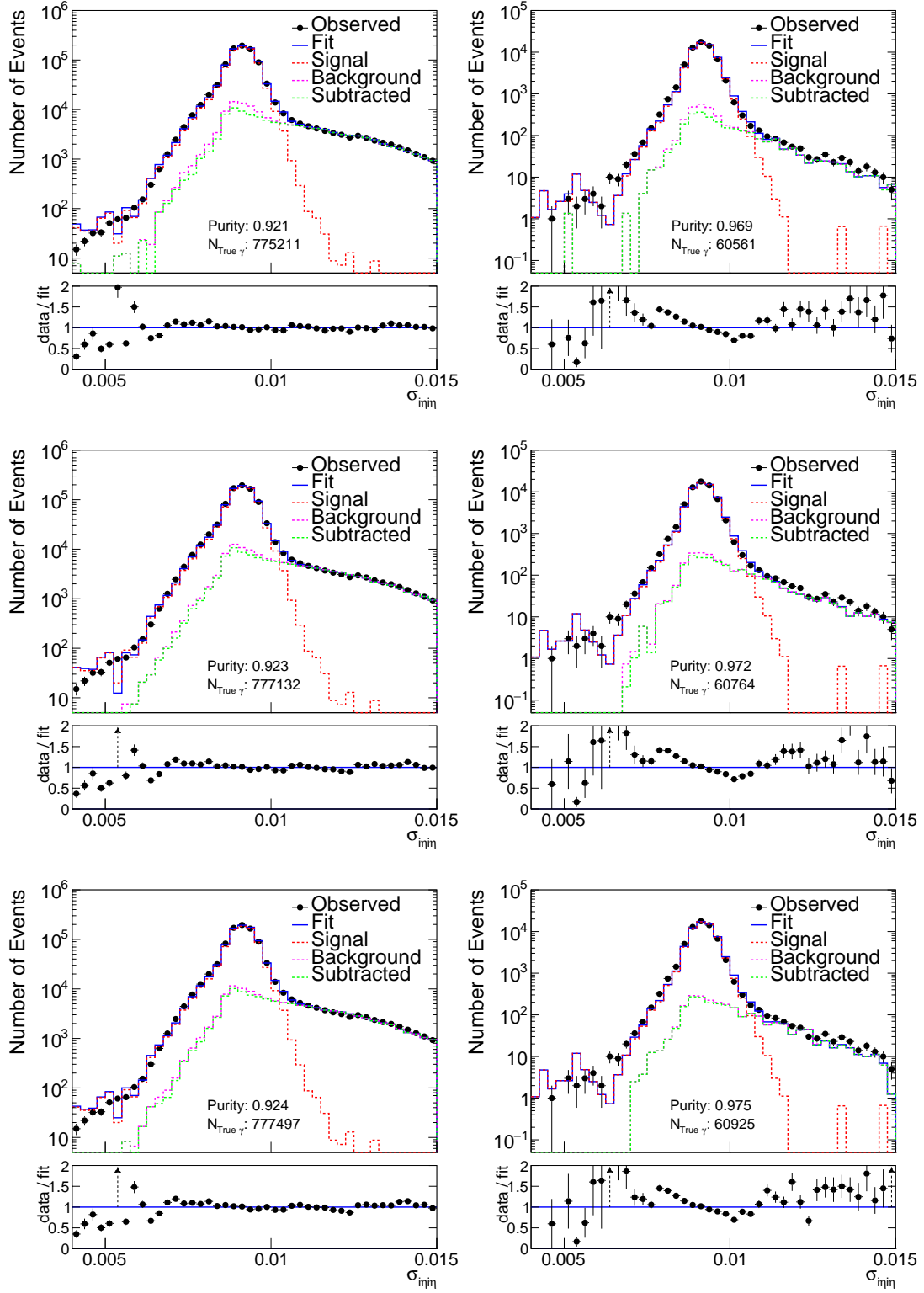
Figure 1-9: Fits to the $\sigma_{i\eta i\eta}$ distributions for the [175, 200] (left) and [400,∞) (right) p_{T} bins using the [3.5,5.0] (top), [5.0,7.5] (middle), and [7.5,9.0] (bottom) isolation sidebands. The blue solid line represents the full fit model, the red dashed line its signal component, and the green dashed line its background component.

the standard deviation of this distribution as a systematic uncertainty.

p_T Range (GeV)	Sources of Systematic Uncertainty			
	Sideband	I_CH Shape	Signal Shape	Bgkd. Stats
(175, 200)	0.09	0.18	0.05	0.04
(200, 250)	0.01	0.16	0.06	0.03
(250, 300)	0.14	0.16	0.06	0.05
(300, 350)	0.12	0.16	0.07	0.08
(350, 400)	0.23	0.11	0.05	0.10
(400, ∞)	0.27	0.09	0.05	0.05

Table 1.4: Relative uncertainties on the estimated number of true photons in the numerator sample.

The values obtained for each systematic uncertainty on the true photon count of the numerator (denominator) are shown in Table 1.4 (Table 1.5)in bins of p_T. In the efficiency, each uncertainty source is considered as fully correlated.

p_T Range (GeV)	Sources of Systematic Uncertainty			
	Sideband	I_CH Shape	Signal Shape	Bgkd. Stats
(175, 200)	0.21	0.15	0.06	0.04
(200, 250)	0.09	0.14	0.07	0.04
(250, 300)	0.12	0.14	0.07	0.06
(300, 350)	0.07	0.13	0.08	0.09
(350, 400)	0.33	0.12	0.07	0.13
(400, ∞)	0.25	0.10	0.12	0.09

Table 1.5: Relative uncertainties on the estimated number of true photons in the denominator sample.

The MC efficiency of the γ-specific ID is determined by counting the number of truth-matched photons passing the e/γ part of the ID and the full ID. However, there is a complication, the γ+jets region in data has approximately 5% contamination from electrons before applying the pixel veto, as shown in Figure 1-10. To replicate this effect in the MC, we combine appropriately cross-section weighted γ+jets, W+jets, and $t\bar{t}$ samples and truth match to both electrons and photons. Additionally, to account for the NLO cross-section ratio uncertainties with respect to γ+jets at this p_T range, we apply a 14% uncertainty on the W+jets and $t\bar{t}$ yields, where the specific value comes from the uncertainty on the γ+jets to W+jets ratio in the monojet

analysis [2]. This uncertainty is uncorrelated between the numerator and denominator as a negligible amount of electron events survive the pixel veto.
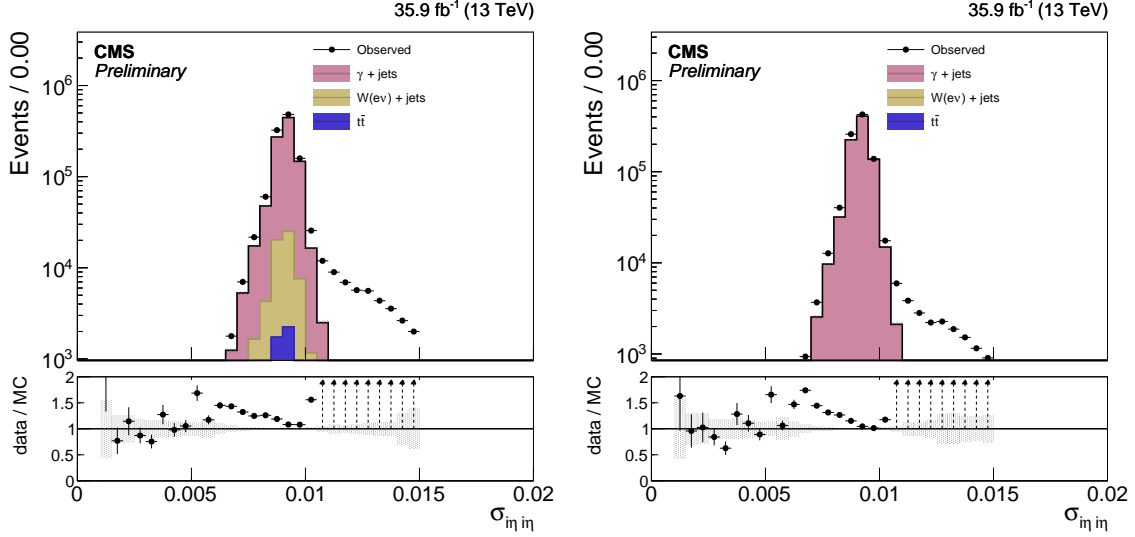

Figure 1-10: Electron contamination in γ+jets region before (left) and after (right) applying the pixel seed veto.

The data efficiency, MC efficiency, and the scale factor for the γ-specific ID as a function of p_T are shown in Figure 1-11. As there is no significant trend in the scale factor as a function of p_T, we apply a flat scale factor of 0.9840 ± 0.0090 for all of the MC-based background and signal models in the analysis.

1.3.3 Lepton Veto Efficiency

The lepton veto requirement in the signal region has a non-unity efficiency over events that do not have genuine leptons, because particles such as pions and protons can mimic leptons to become "fake leptons" and cause the event to be rejected. To measure the possible difference between data and MC of this lepton veto efficiency, we compare dimuon events in data and MC. In a high-purity $Z \to \mu\mu$ sample with the dimuon mass close to M_Z, events with a genuine third lepton is negligibly rare, and therefore the efficiency loss from rejecting events with a third lepton is dominantly due to fake leptons.

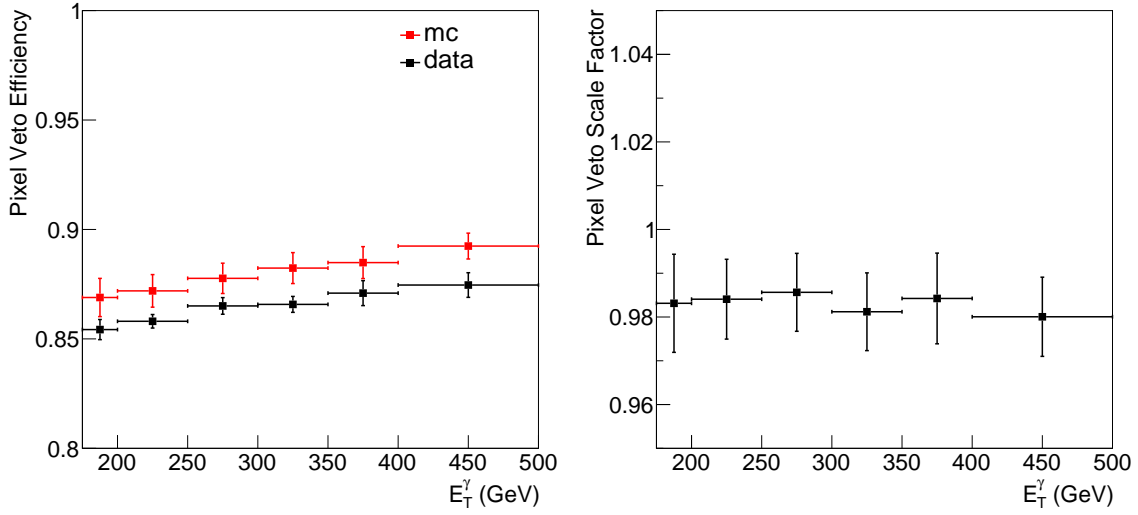For this measurement, collision events are taken from the SingleMuon data set

Figure 1-11: Photon pixel veto efficiencies (left) and corresponding scale factor (right) as a function of photon p_{T}.

and the MC events from a mixture of Drell-Yan, $t\bar{t}$, WW, WZ, and ZZ samples. We require two muons passing the "tight" identification working point defined with the mass between 61 and 121 GeV. The failing category consists of events containing an additional electron or muon object that passes the loose selection criteria while the passing category consists of those without an additional lepton. The efficiency is inspected as a function of number of vertices, number of jets, and H_{T} in the event, and in all cases data and MC are consistent as shown in Figure 1-12.

It should be noted, however, that the absolute lepton veto efficiency in MC dimuon sample is significantly different from that of the $Z(\to \nu\bar{\nu})+\gamma$ sample, which more closely features the properties of the signal candidate sample. The full difference in the efficiencies between the dimuon and $Z(\to \nu\bar{\nu})+\gamma$ samples is taken as the systematic uncertainty in the lepton veto scale factor, which is therefore 1.00 ± 0.02.

Additionally, a small fraction of events with real leptons pass the lepton veto due to the leptons failing the loose ID requirements. This effect is most relevant for $W(\to \ell\nu)+\gamma$ events in the signal region and for $Z(\to \ell\bar{\ell})+\gamma$ events in the single lepton control regions. We compute a scale factor $\mathrm{SF}_{\mathrm{veto}} = (1.-\epsilon_{\mathrm{data}})/(1.-\epsilon_{\mathrm{MC}})$ using the data and MC efficiencies for the loose lepton IDs with a flat 1% uncertainty for the efficiencies. All scale factors are consistent with unity within the uncertainties.
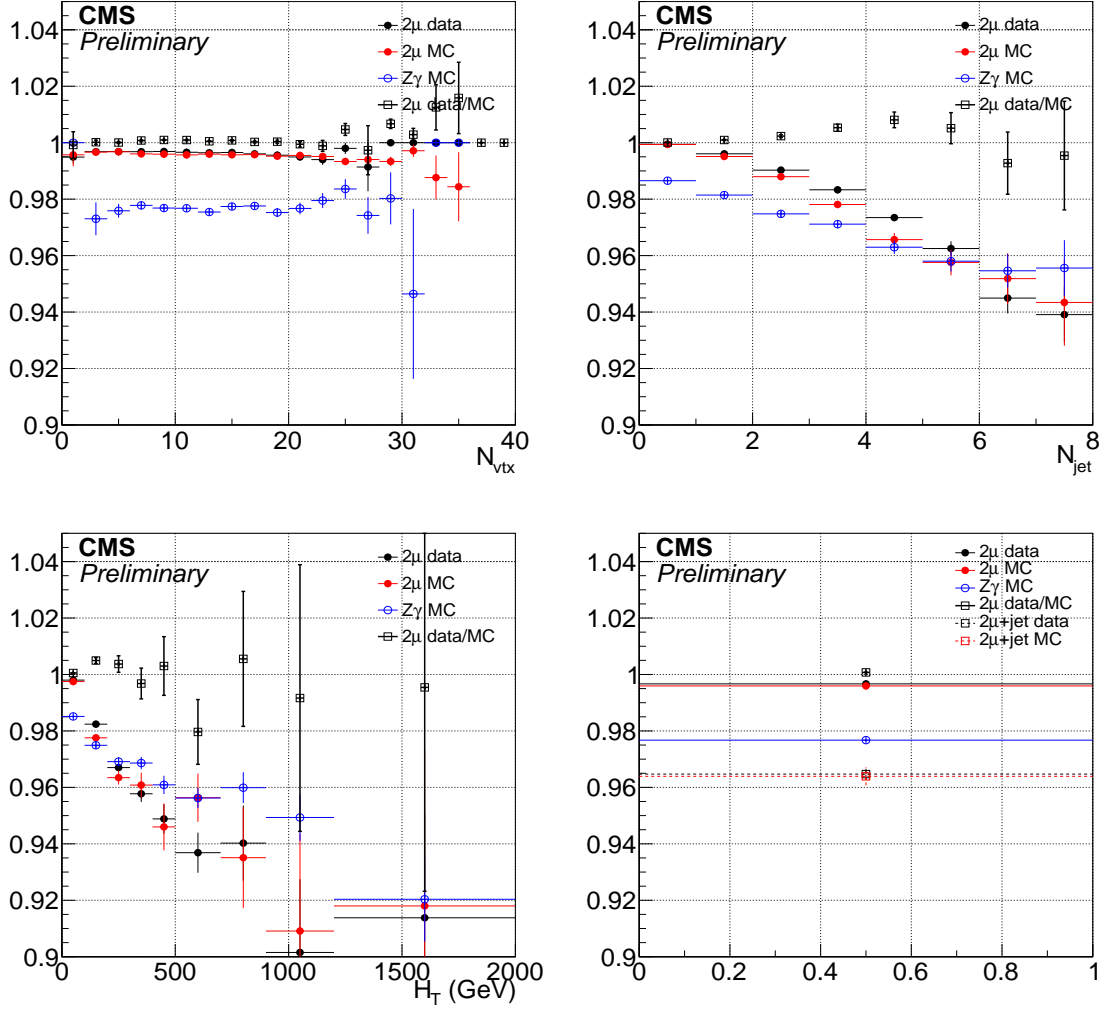
30

Figure 1-12: Lepton veto efficiencies and data/MC scale factors as functions of N_{vtx}, N_{jet}, and H_{T}, and the corresponding inclusive values. While dimuon and $Z(\to \nu\bar{\nu})+\gamma$ samples have significantly different efficiencies, data and MC agree well within dimuon samples, giving scale factors consisten with 1 almost everywhere. This is true even when additionally requiring a high-p_{T} jet in the event, as seen in the inclusive efficiency plot. Thus, the difference between $Z(\to \nu\bar{\nu})+\gamma$ and dimuon efficiencies itself is taken as the uncertainty.

This veto scale factor is applied to MC events with a reconstructed lepton that fails the loose ID. If there are multiple such leptons in an event, we apply the scale factor only for the hardest muon and electron. After applying the scale factors, the final MC yields for $W(\to \ell\nu)+\gamma$ in the signal region and $Z(\to \bar{\ell}\ell)+\gamma$ in the single lepton control regions change by less than 0.5%.

1.4   Misidentified electrons

An electron can be misidentified as a photon if the association of tracks or track seeds to the ECAL supercluster fails in the reconstruction step. The production of a single W boson decaying to an electron and a neutrino is a high-rate process, and it mimicks the photon plus $E_{\mathrm{T}}^{\mathrm{miss}}$ signature if the electron is misidentified.

The rate at which this misidentification occurs is $R_e = (1 - \epsilon_e^{\mathrm{track}})/\epsilon_e^{\mathrm{track}}$, where $\epsilon_e^{\mathrm{track}}$ is the tracking efficiency of electrons passing the photon identification criteria except for the electron veto. We measure the factor R_e in data using the TP method described in Section 1.3.1 with changed definitions for passing and failing probes and an adjustment to the background model. The ee category contains passing probes with a pixel seed while the $e\gamma$ category contains failing probes without a pixel seed. Probes in both categories must pass the remainder of the e/γ and γ-specific IDs. Denoting the area of the peak in each category N_{ee} and $N_{e\gamma}$, respectively, the ratio $N_{e\gamma}/N_{ee}$ is equal to R_e up to minor systematic corrections.

Additionally, the backgrounds to the $e\gamma$ fit consist of processes with an electron and an actual photon in the final state, such as $W\gamma$ and $Z \to ee$ with a hard radiation off one of the electrons. To account for the higher rate of bremsstrahlung experienced by electrons than by muons, we scale the mass distribution of the $\mu + \gamma$ sample by the ratio of electron-probe to muon-probe events taken from MC. As an alternative template to assess the systematic effect introduced by the choice of the background template, the unscaled mass distribution is also tested.

Figure 1-13 shows the six fits performed on ee and $e\gamma$ in bins of probe p_{T}, from which the R_e factor used for the estimation of the electron misidentification back-
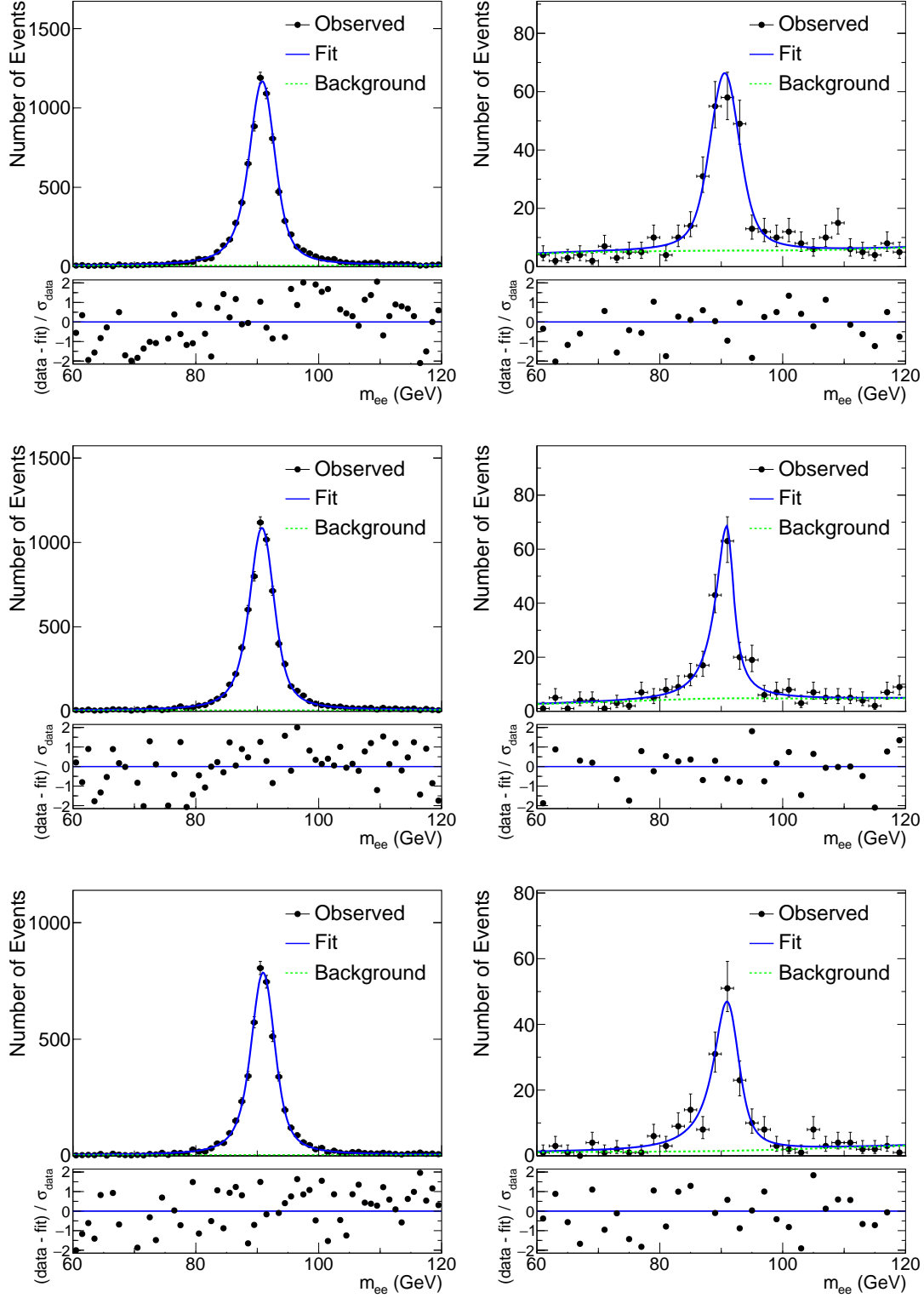
Figure 1-13: Fits to the mass distributions for ee (left) and $e\gamma$ (right) selections, in bins of probe p_{T}: $175 < p_{\mathrm{T}} < 200\,\mathrm{GeV}$ (top), $200 < p_{\mathrm{T}} < 250\,\mathrm{GeV}$ (middle), $p_{\mathrm{T}} > 250\,\mathrm{GeV}$ (bottom). The blue solid line represents the full fit model, and the green dashed line its background component.

Figure 1-14: Electron to photon fake rate R_e.
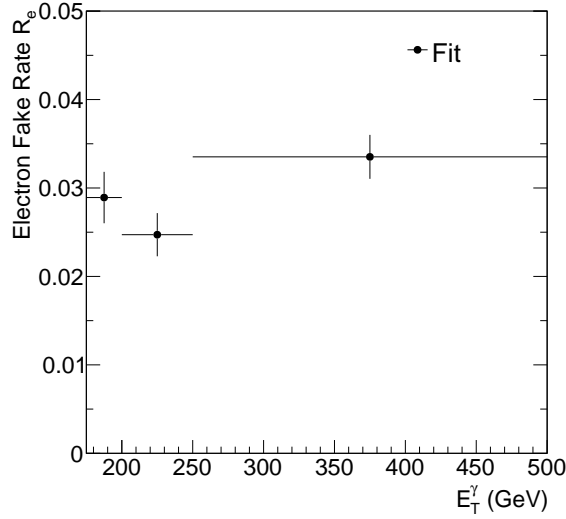
ground is derived. Figure 1-14 shows the derived R_e factor as a function of E_T^γ.

After making the reasonable assumption that the kinematic and other critical properties of the electron plus E_T^{miss} events are unaffected by the electron misidentification, we model the electron misidentification background by taking the electron proxy sample and reweighting each event by R_e depending on the p_T of the electron proxy object.

1.5 Misidentified hadrons

A hadron can be misidentified as a photon if fragmentation processes results in mainly neutral hadrons that subsequently decay to collimated pairs of photons. The production of Z+jets where the Z boson decays to neutrinos is a high-rate process, and it mimicks the photon plus E_T^{miss} signature if the hadrons from the jet are misidentified.

Without the presence of additional charged tracks or neutral hadron energy deposits, the only way to distinquish these EM-like hadrons from real photons is through the shower shape. Thus, we measure the fraction of hadronic objects within a pool of photon candidate objects in the EM object+jet measurement sample using the $\sigma_{i\eta i\eta}$ template fit method from Section 1.3.2. Figure 1-15 and Table 1.6 show the final
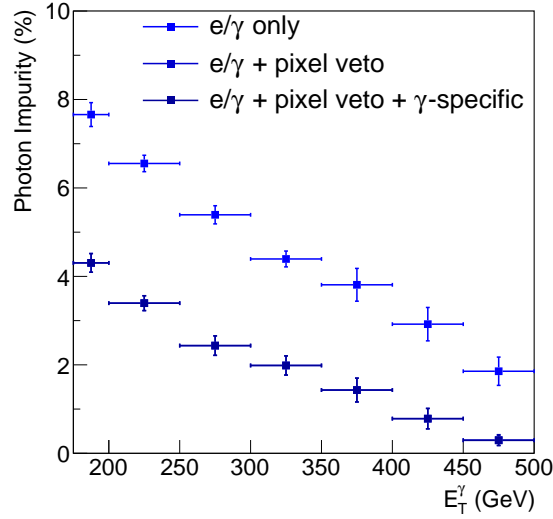
Figure 1-15: The percent impurity for photons as a function of p_{T}. The different bands show the effects of adding different stages of the full ID, starting with the e/γ portion of the ID and successively adding the pixel seed veto followed by the rest of the γ-specific portion of the ID. These last two curves overlap, as the non-collision rejection cuts do not effect the rate at which hadrons are misidentified as photons.

impurity and associated uncertainties as a function of p_{T}.

p_{T} (GeV)	Nominal (%)	Sources of Systematic Uncertainty			
		Sideband	CH Iso Shape	Signal Shape	Bgkd. Stats
(175, 200)	4.31 ± 0.21	0.09	0.18	0.05	0.04
(200, 250)	3.39 ± 0.17	0.01	0.16	0.06	0.03
(250, 300)	2.44 ± 0.22	0.14	0.16	0.06	0.05
(300, 350)	1.99 ± 0.23	0.12	0.16	0.07	0.08
(350, 400)	1.43 ± 0.28	0.23	0.11	0.05	0.10
(400, ∞)	0.63 ± 0.30	0.27	0.09	0.05	0.05

Table 1.6: Impurities for photons as a function of p_{T}.

The hadronic transfer factor R_h measures the rate at which hadronic proxy objects result in hadrons that are misidentified as candidate photons. The factor R_h is obtained by dividing the estimated number of misidentified hadrons in the EM object+jet measurement sample by the number of events in the hadron proxy+jet measurement sample as a function of p_{T}. [This section and honstly the last few would benefit from a table or even a Venn diagram showing the regions and the rates] Figure 1-16 shows the transfer factor R_h along with the various distributions used for

its derivation.

Under the assumption that the R_h stays constant regardless of whether the event has a high-p_T jet or E_T^{miss}, the hadron proxy sample is weighted by R_h to determine the number of events due to misidentified hadrons in the signal region.

To estimate the uncertainty on this background, we repeat the above method using additional proxy and measurement samples with tighter and looser definitions of the hadron proxy object. The tight and loose shapes are taken as the one sigma band around the nominal estimate. Additionally, there is an uncertainty coming from the estimation of the photon purity, with values given in Table 1.6.

1.6 Irreducible backgrounds

1.6.1 Simulation of V+γ Processes

The $Z(\to \nu\overline{\nu})+\gamma$ and $W(\to \ell\nu)+\gamma$ background contributions are modeled using MC simulations. Samples generated at the leading order (LO) in QCD by MADGRAPH 5 with up to two additional partons and a generator-level requirement of $E_T^\gamma > 130\,\text{GeV}$ are employed for this purpose.

A study using an aMC@NLO sample with high E_T^γ threshold confirms that the predicted kinematic distributions would not change drastically by using the NLO sample. Figures 1-17 and 1-18 show the comparisons of the aMC@NLO samples[1] and the MADGRAPH 5 samples used for the background estimation in the key kinematic distributions.

To approximate the QCD higher-order effects, $Z(\to \nu\overline{\nu})+\gamma$ and $W(\to \ell\nu)+\gamma$ events are reweighted with E_T^γ by the factors given in Tab. 1.7. These factors are the ratios of QCD next-to-next-to leading order (NNLO) differential cross sections calculated by Grazzini et al. [3] to the LO cross sections given in the centrally produced samples. Note that the denominator cross section includes contributions from processes with up to two additional partons, and is therefore not a LO cross section
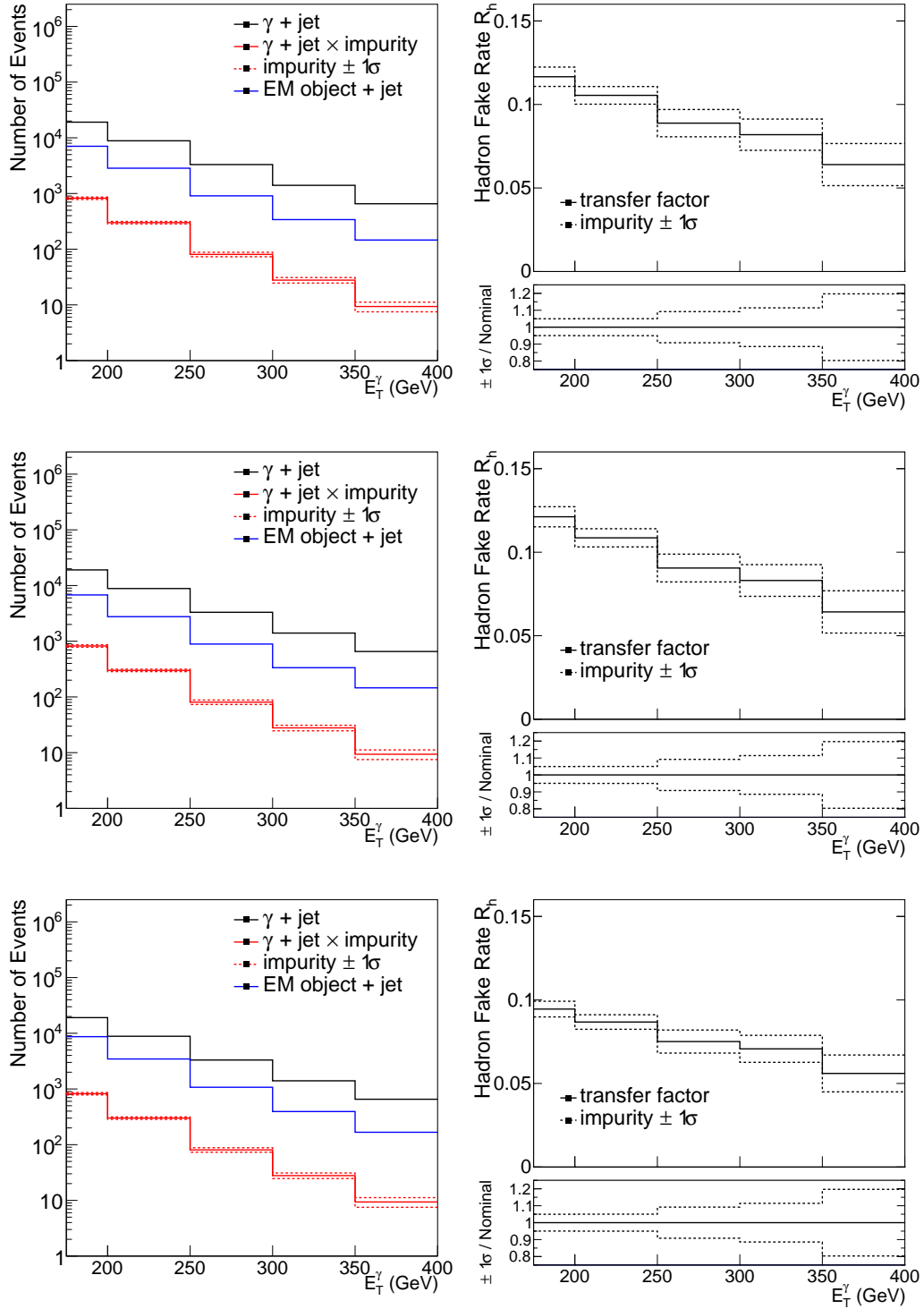
---

[1]These samples were privately produced.

Figure 1-16: Left: The p_T distribution of the candidate photon object in the γ+jets control sample (black), the result of scaling it with the impurity (red), and the p_T distribution of the EM object in the EM object + jet control sample (blue). Right: Hadronic transfer factor R_h, which is the ratio of the red and blue distributions in the left plot. The dashed bands indicate the uncertainty on R_h due to the impurity measurement. Top: Nominal hadron proxy object. Middle: Tighter hadron proxy object. Bottom: Looser hadron proxy object.
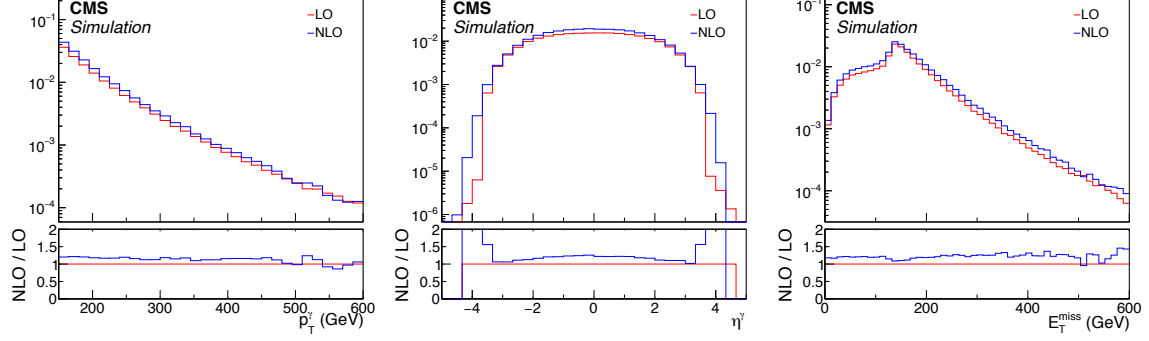
37

Figure 1-17: Distributions of E_T^γ (left), η^γ (middle), and p_T^Z (right) in $Z(\to \nu\bar{\nu})+\gamma$ process from the private aMC@NLO sample (blue) and the LO sample used for background prediction (red) along with the NLO / LO ratios.
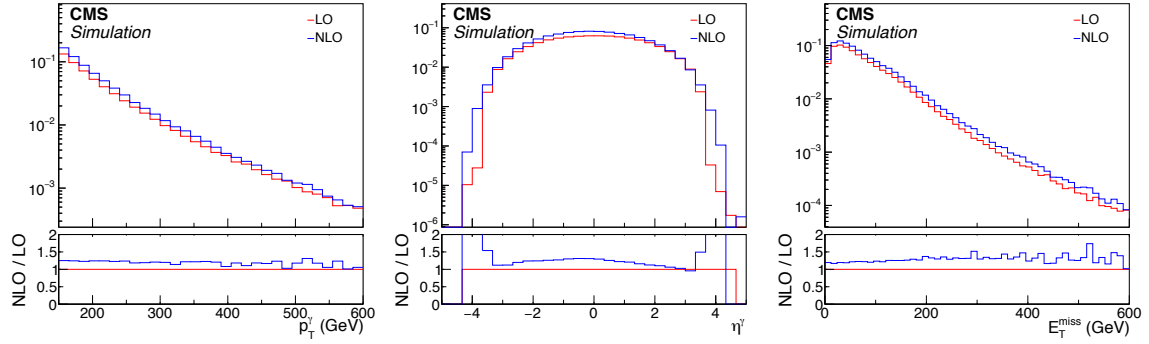

Figure 1-18: Distributions of E_T^γ (top left), η^γ (top right), and p_T^W (bottom left) in $W(\to \ell\nu)+\gamma$ process from the private aMC@NLO sample (blue) and the LO sample used for background prediction (red) along with the NLO / LO ratios.

in the strict sense of the word. $V\gamma$ k-factors found in literature are $\gg 1$ at high E_T^γ, if the denominator only accounts for the cross section of $q\bar{q} \to V\gamma$ process.

E_T^γ range (GeV)	$Z(\to \nu\bar{\nu})+\gamma$	$W(\to \ell\nu)+\gamma$
$[175, 190]$	1.44	1.40
$[190, 250]$	1.41	1.37
$[250, 400]$	1.35	1.31
$[400, 700]$	1.29	1.26
$[700, \infty)$	1.15	1.15

Table 1.7: Correction factors $k_{\mathrm{QCD}}^{\mathrm{NNLO}}$ for $Z(\to \nu\bar{\nu})+\gamma$ and $W(\to \ell\nu)+\gamma$ samples.

Higher-order electroweak correction factors are also applied as a function of E_T^γ. Out of various electroweak higher-order effects, ones that can give sizeable ($\gg \mathcal{O}(\alpha)$) corrections to the cross section are the Sudakov suppression at high boson p_T and potentially the addition of photon-induced scattering processes [4, 5]. We apply the correction factors shown in Figure 1-19, which are the combinations of Sudakov suppression factors and photon-induced enhancements, and are provided by the authors of Reference [5] in addition to the NNLO QCD correction.
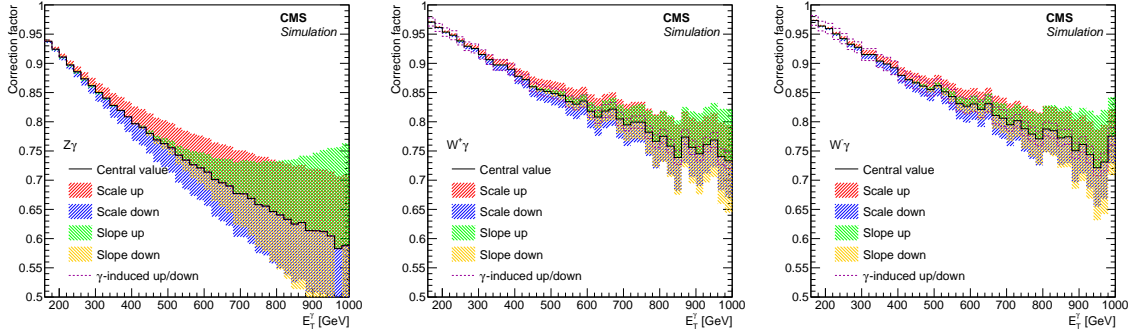

Figure 1-19: Electroweak NLO cross section corrections as a function of photon p_T for $Z(\to \nu\bar{\nu})+\gamma$ (left), $W^+ + \gamma$ (middle), and $W^- + \gamma$ (right) processes, overlaid with uncertainty bands. The scale and slope bands show the behavior of the two Sudakov uncertainties described in the text. The uncertainty due to γ-induced production is negligible in $Z(\to \nu\bar{\nu})+\gamma$ production.

The differential cross section after the full higher-order corrections is

$$d\sigma_{\mathrm{QCD}}^{\mathrm{LO}} \cdot k_{\mathrm{QCD}}^{\mathrm{NNLO}} \cdot \left(1 + k_{\mathrm{EW}}^{\mathrm{Sudakov}} + k_{\mathrm{EW}}^{q\gamma}\right), \tag{1.6}$$

where $k_{\mathrm{QCD}}^{\mathrm{NNLO}} = d\sigma_{\mathrm{QCD}}^{\mathrm{NNLO}}/d\sigma_{\mathrm{QCD}}^{\mathrm{LO}}$, and the two k_{EW} terms are the Sudakov suppression and photon-induced enhancement components of the electroweak correction, respectively.

Furthermore, subtle differences between simulation and observation in the reconstruction and identification efficiencies for various particle candidates are accounted for with the set of selection efficiency correction factors ρ. The value of an individual ρ typically lies within a few percent of unity.

Four sources of systematic uncertainties considered for E_{T}^{γ} distribution ratios of the V+γ processes are higher-order QCD corrections, higher-order EWK corrections, choice of PDF set, and data-to-simulation correction factors ρ. The four uncertainties are evaluated for each E_{T}^{γ} bin and fully correlated between the different bins..

The higher-order QCD renormalization and factorization scale uncertainties on the NNLO cross sections are assessed by varying the respective scales by factors 2 and 0.5 during the cross section computation. The uncertainties vary between 7-8% across the bins and are considered uncorrelated in the ratio between the $Z(\rightarrow \nu\bar{\nu})+\gamma$ and $W(\rightarrow \ell\nu)+\gamma$ processes.

Theoretical uncertainties on the electroweak corrections are not well understood to date for $V\gamma$ processes. Our treatement of them is informed by discussions with the authors of Reference [6], which provides a prescription for electroweak correction uncertainties for V+jets processes. We estimate the magnitude of the uncertainty on $k_{\mathrm{EW}}^{\mathrm{Sudakov}}$ and $k_{\mathrm{EW}}^{q\gamma}$ to be $\left(k_{\mathrm{EW}}^{\mathrm{Sudakov}}\right)^2$ and $k_{\mathrm{EW}}^{q\gamma}$, i.e., square of the correction for Sudakov suppression and the 100% of the correction itself for the photon-induced enhancement. The choice of using the square of $k_{\mathrm{EW}}^{\mathrm{Sudakov}}$ is motivated by the fact that fully resummed leading-log Sudakov suppression is an exponential of $k_{\mathrm{EW}}^{\mathrm{Sudakov}}$.

For the Sudakov suppression, which is the dominant term in the electroweak correction, we further consider two types of systematic variations. In this paper, electroweak correction as a function of the boson p_{T} is varied in overall scale and in slope. The slope variation is realized by selecting a point in the boson p_{T} spectrum and letting the shift in correction cross over at the point (see Figure 1-20). Following this prescription, we let the Sudakov suppression vary in overall scale and in slope,

where we choose our crossover point for the slope variation to be at $E_T^\gamma = 600\,\text{GeV}$, to prevent the low E_T^γ bins from overconstraining the EWK uncertainties in the last bin.
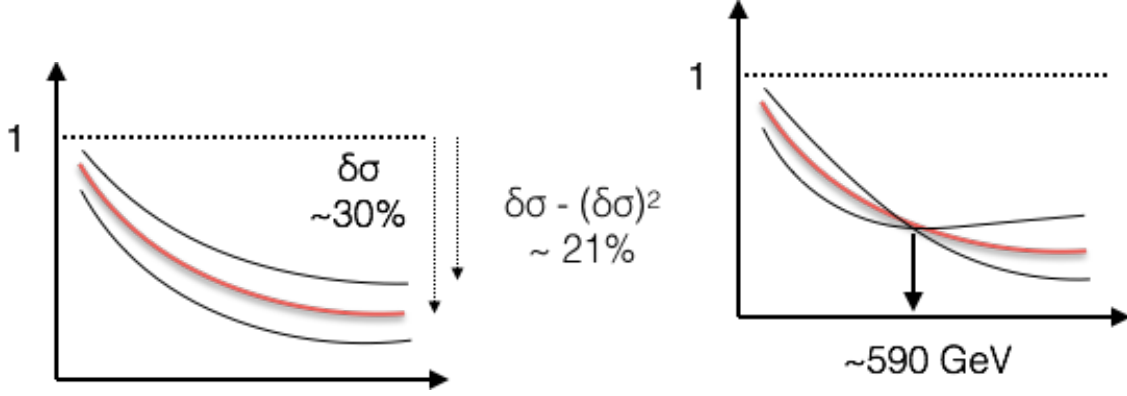

Figure 1-20: Electroweak correction variation scheme to cover the scale (left) and shape (right) uncertainties.

The PDF uncertainty is evaluated by varying the weight of each event using the weights provided in the NNPDF set, and taking the standard deviation of the resulting E_T^γ distributions. This uncertainty is considered fully correlated in the ratio between the $Z(\to \nu\bar{\nu})+\gamma$ and $W(\to \ell\nu)+\gamma$ processes, i.e., the variation of the ratio is bounded by the ratios formed by the simultaneous upward and downward variations of the numerator and denominator.

Finally, data-to-simulation correction factors ρ for the lepton identification efficiencies have associated uncertainties that do not cancel when taking ratios between regions defined by different lepton selection requirements. The lepton efficiencies are measured using the "tag-and-probe" method where the tag object is an electron (muon) object passing the tight ID and matched to a SingleElectron (SingleMuon) trigger and the probe object is a PF electron (muon) without any ID applied. The passing (failing) categories are defined by events with probes passing (failing) the ID definition in question. The electron (muon) scale factors ρ are approximately unity with a flat 2% (1%) systematic uncertainty.

1.6.2 Lepton Control Regions

Contributions from the $Z(\to \nu\bar{\nu})+\gamma$ and $W(\to \ell\nu)+\gamma$ processes are estimated using observed data in four mutually exclusive single-electron, single-muon, dielectron, and dimuon control regions.

In all of these regions, the recoil vector $\vec{\mathbf{U}} = \vec{p}_{\mathrm{T}}^{\mathrm{miss}} + \sum_{\ell} \vec{p}_{\mathrm{T}}^{\ell}$ serves as an analogue for the $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ in the signal region. In the signal region, the $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ is a proxy for the vector boson p_{T} while in the control regions, the recoil vector is used instead. Thus, the recoil $\vec{\mathbf{U}}$ must satisty identical requirements to those for the $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ in the signal region to keep the control region kinematics as similar as possible to the signal region kinematics. This forms the baseline selection for the lepton control regions, as shown in Table 1.8. The lepton veto is understood to apply for any leptons beyond those specifically required for each region.

Variable	Selection	Motivation
E_{T}^{γ}	$> 175\,\mathrm{GeV}$	high-p_{T} photon passing trigger
$|\eta|$	< 1.44	region with best background estimates
e/γ ID	Pass	reject hadronic background
γ-specific ID	Pass	reject electron and non-collision backgrounds
$\mathbf{U}$	$> 170\,\mathrm{GeV}$	characteristic signature of dark matter
$E_{\mathrm{T}}^{\gamma}/\mathbf{U}$	< 1.4	reduce jet mismeasurement backgrounds
$\min\Delta\phi(\vec{p}_{\mathrm{T}}^{\mathrm{jet}}, \vec{\mathbf{U}})$	> 0.5	reduce jet mismeasurement backgrounds
$\Delta\phi(\gamma, \vec{\mathbf{U}})$	> 0.5	reduce photon mismeasurement backgrounds
Lepton Veto	Pass	reject backgrounds with additional real leptons

Table 1.8: Baseline selections for the lepton control regions. The lepton veto is understood to apply for any leptons beyond those specifically required for each region.

The single-electron (single-muon) control region is defined by requiring exactly one tight electron (muon). To suppress the contributions from large-$E_{\mathrm{T}}^{\mathrm{miss}}$ processes other than $W(\to \ell\nu)+\gamma$, the transverse mass $m_{\mathrm{T}} = \sqrt{2E_{\mathrm{T}}^{\mathrm{miss}}p_{\mathrm{T}}^{\ell}[1 - \cos\Delta\phi(\vec{p}_{\mathrm{T}}^{\mathrm{miss}}, \vec{p}_{\mathrm{T}}^{\ell})]}$ must be less than $160\,\mathrm{GeV}$. Additionally, for the single-electron control region, to limit the contribution from the γ+jets process where a jet is misidentified as an electron, the $E_{\mathrm{T}}^{\mathrm{miss}}$ must be greater than $50\,\mathrm{GeV}$. Tables 1.9 and 1.10 summarize the additional selection for the single-electron and single-muon control regions, respectively.

The dielectron (dimuon) control region is defined by requiring exactly two elec-

Variable	Selection	Motivation
Leading e	Tight ID	well-identified electron
m_T	$< 160\,\mathrm{GeV}$	suppress other large-$E_\mathrm{T}^\mathrm{miss}$ processes
$E_\mathrm{T}^\mathrm{miss}$	$> 50\,\mathrm{GeV}$	reject jets misidentified as electrons

Table 1.9: Additional selections for the single-electron control region.

Variable	Selection	Motivation
Leading μ	Tight ID	well-identified muon
m_T	$< 160\,\mathrm{GeV}$	suppress other large-$E_\mathrm{T}^\mathrm{miss}$ processes

Table 1.10: Additional selections for the single-muon control region.

trons (muons), with $60 < m_{\ell\ell} < 120\,\mathrm{GeV}$, where $m_{\ell\ell}$ is the mass of the dilepton system. The leading lepton must pass the tight ID requirements, while the trailing lepton only needs to pass the loose ID requirements. Tables 1.11 and 1.12 summarize the additional selection for the dielectron and dimuon control regions, respectively.

Variable	Selection	Motivation
Leading e	Tight ID	well-identified electron
Trailing e	Loose ID	well-identified electron
m_{ee}	$> 60\,\mathrm{GeV}, < 120\,\mathrm{GeV}$	originate from Z-boson

Table 1.11: Additional selections for the dielectron control region.

1.6.3 Data-driven Estimation

The various ratios between the expected $Z(\to \nu\bar{\nu})+\gamma$ yield in the combined signal regions and the expected $W(\to \ell\nu)+\gamma$ and $Z(\to \ell\bar{\ell})+\gamma$ yields in the control regions are constrained by MC simulations. When a ratio is calculated using MC samples as a function of E_T^γ, it is referred to as the "transfer factor" between the two processes. This background estimation method exploits the cancellation of some of the systematic uncertainties, both experimental and theoretical, in the transfer factors between the different V+γ processes.

For the transfer factor $R_{\ell\ell\gamma}^{Z\gamma}$, the numerator is the expected $Z(\to \nu\bar{\nu})+\gamma$ yield in the combined signal regions and the denominator is the expected $Z(\to \ell\bar{\ell})+\gamma$ yield in the relevant dilepton control region. The uncertainties due to photon energy calibration, jet energy resolution, and higher-order QCD effects are significantly reduced on $R_{\ell\ell\gamma}^{Z\gamma}$

Variable	Selection	Motivation
Leading μ	Tight ID	well-identified muon
Trailing μ	Loose ID	well-identified muon
$m_{\mu\mu}$	$> 60\,\mathrm{GeV}, < 120\,\mathrm{GeV}$	originate from Z-boson

Table 1.12: Additional selections for the dimuon control region.

compared to when such effects are considered for individual processes. The only uncertainties in the transfer factor $R_{\ell\ell\gamma}^{Z\gamma}$ that do not largely cancel are those on lepton identification efficiency, the statistical uncertainty due to the limited MC sample size, and a minor uncertainty due to the different acceptances between the $Z(\to \nu\bar{\nu})+\gamma$ and $Z(\to \ell\bar{\ell})+\gamma$ processes. Figure 1-21 shows the transfer factor $R_{ee\gamma}^{Z\gamma}$ ($R_{\mu\mu\gamma}^{Z\gamma}$) between the dielectron (dimuon) control region and the combined signal regions.

[mention that the postfit uncertainties are governed by the statistical uncertainty of the signal region.]
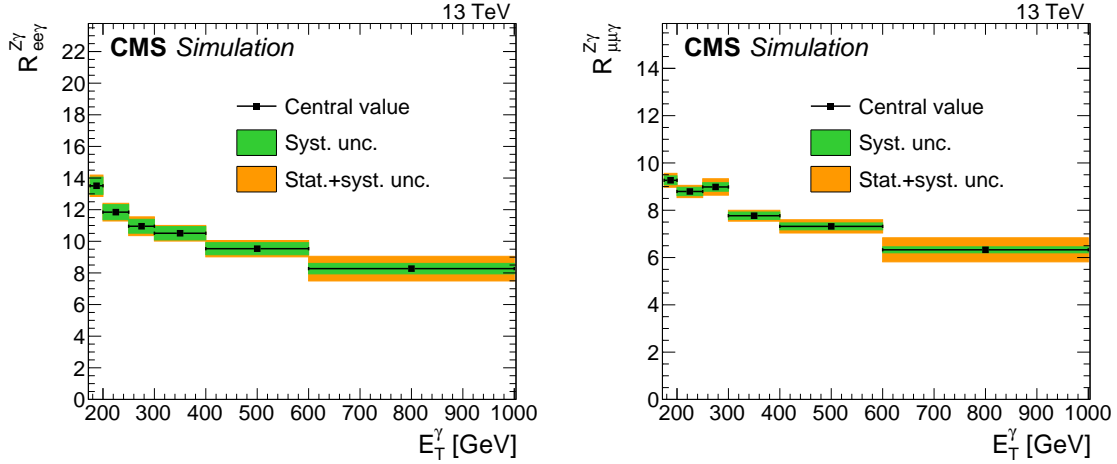

Figure 1-21: Transfer factors $R_{ee\gamma}^{Z\gamma}$ (left) and $R_{\mu\mu\gamma}^{Z\gamma}$ (right). The numerator is the expected $Z(\to \nu\bar{\nu})+\gamma$ yield in the combined signal regions and the denominator is the expected $Z(\to \ell\bar{\ell})+\gamma$ yield in the dielectron (left) or dimuon (right) control region. The uncertainty bands in green (inner) and orange (outer) show the systematic uncertainty, and the combination of systematic and statistical uncertainty arising from limited MC sample size, respectively. The systematic uncertainties considered are the uncertainties in the data-to-simulation correction factors ρ for the lepton identification efficiencies.

For increasing E_{T}^{γ}, the Z boson in a $Z(\to \ell\bar{\ell})+\gamma$ event tends to emerge with lower rapidity, and hence so do its decay products. As a consequence, the charged

leptons are more likely to fall within the inner tracker acceptance, which increases the dilepton control region selection efficiency of these events. In contrast, the signal region selection efficiency of $Z(\to \nu\bar{\nu})+\gamma$ events is unaffected by the rapidity of the final state neutrinos, as long as the observed E_T^miss has the appropriate magnitude and azimuthal direction. This causes the distinctive drop in the ratio $R_{\ell\ell\gamma}^{Z\gamma}$ with increasing E_T^γ.

Using the transfer factor $R_{\ell\ell\gamma}^{Z\gamma}$, the total estimated event yield $T_{\ell\ell\gamma}$ in each dilepton control region in the ith bin of the E_T^γ distribution is given by

$$T_{\ell\ell\gamma,i} = \frac{N_i^{Z\gamma}}{R_{\ell\ell\gamma,i}^{Z\gamma}} + b_{\ell\ell\gamma,i}, \tag{1.7}$$

where $N^{Z\gamma}$ is the predicted number of $Z(\to \nu\bar{\nu})+\gamma$ events in the combined signal regions and $b_{\ell\ell\gamma}$ is the predicted contribution from other background sources in the dilepton control region, namely $t\bar{t}\gamma$, $VV\gamma$, and misidentified hadrons. The subscript i indicates that the quantities are evaluated in bin i of the E_T^γ distribution.

Similar considerations apply to events arising from the $W(\to \ell\nu)+\gamma$ process. A large fraction of such events are rejected by the electron and muon vetoes in the signal region selection and end up in the control regions instead. However, hadronic tau events and events where the leptons are out of acceptance or fail to be reconstructed will remain in the signal region, on top of the vetoes having imperfect efficiencies. In the ratio of these two classes of events, denoted $R_{\ell\gamma}^{W\gamma}$, the only uncertainties that remain non-negligible are those associated with the lepton identification efficiency and the MC statistical uncertainty.

Table 1.13 gives the breakdown of the $W(\to \ell\nu)+\gamma$ background passing the full event selection for the signal region, categorized by the lepton flavor and, for the case of electrons and muons, the lepton pseudorapidity at the parton-level. From this breakdown, one sees that events where the W boson decays to a τ and a neutrino constitute approximately 60% of the $W(\to \ell\nu)+\gamma$ background. The remaining 40% of the $W(\to \ell\nu)+\gamma$ background comes from events where the W boson decays to a μ or e and a neutrino. Events containing an electron are more likely to be within the

detector acceptance, while those with a muon are more likely to be out of acceptance. For the in-acceptance background ($|\eta| < 2.5$), the identification efficiency is lower for electrons than for muons, which translates to a larger background contribution from the electrons. The requirement for large $E_{\mathrm{T}}^{\mathrm{miss}}$ removes events with out-of-acceptance electrons because the energy from these electrons is captured by the calorimeters and retains events with out-of-acceptance muons because they contribute directly to the missing momentum. The overall result is a larger background contribution from events with out-of-acceptance muons than from events with out-of-acceptance electrons. [This discussion is lacking. You should probably discuss how hte MET calcualtion changes for muon in and out of acceptance when compared to electrons]

Subprocess	$A \times \epsilon \times 10^3$	
$W \to e\nu + \gamma$	1.68	
$|\eta^e| < 2.5$		1.35
$|\eta^e| > 2.5$		0.32
$W \to \mu\nu + \gamma$	1.83	
$|\eta^\mu| < 2.5$		0.74
$|\eta^\mu| > 2.5$		1.08
$W \to \tau\nu + \gamma$	5.03	

Table 1.13: The breakdown of simulated $W+\gamma$ events passing the full event selection. Events are categorized in the W decay mode. Events with $e\nu$ and $\mu\nu$ final states are further divided into those where the lepton is roughly within acceptance ($|\eta| < 2.5$) but failed the lepton veto, and those where the lepton is out of acceptance ($|\eta| > 2.5$). For each W decay mode, the fraction out of total generated ($A \times \epsilon$) is shown.

Figure 1-22 shows the transfer factor $R_{e\gamma}^{W\gamma}$ ($R_{\mu\gamma}^{W\gamma}$) between the single-electron (single-muon) control region and the combined signal regions, for which the numerator is the estimated $W(\to \ell\nu)+\gamma$ yield in the combined signal regions, and the denominator is the estimated $W(\to \ell\nu)+\gamma$ yield in the relevant control region. The ratio $R_{\ell\gamma}^{W\gamma}$ decreases with increasing E_{T}^γ in a similar manner to $R_{\ell\ell\gamma}^{Z\gamma}$. The underlying logic is the same; e.g., that the signal region selection efficiency is unaffected by E_{T}^γ while the control region acceptances increase with increasing E_{T}^γ due to increased lepton efficiency resulting from lower W rapidity.

Finally, to benefit further from the larger statistical power that the single-lepton control samples provides, an additional transfer factor $f_{W\gamma}^{Z\gamma} = N^{Z\gamma}/N^{W\gamma}$ is defined to
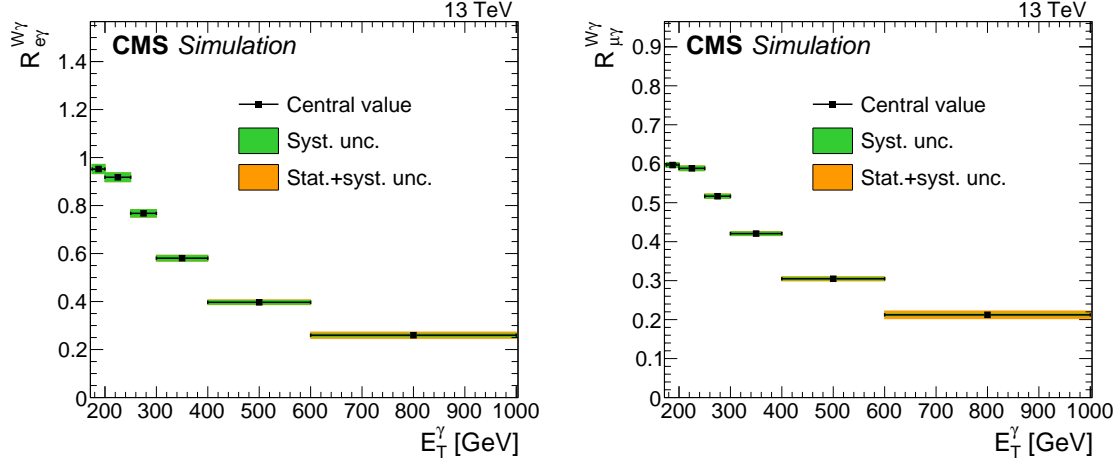
Figure 1-22: Transfer factors $R_{e\gamma}^{W\gamma}$ (left) and $R_{\mu\gamma}^{W\gamma}$ (right). The numerator is the expected $W(\to \ell\nu)+\gamma$ yield in the combined signal regions and the denominator is the expected $W(\to \ell\nu)+\gamma$ yield in the mono-electron (left) or mono-muon (right) control region. The uncertainty bands in green (inner) and orange (outer) show the systematic uncertainty, and the combination of systematic and statistical uncertainty arising from limited MC sample size, respectively. The systematic uncertainties considered are the uncertainties in the data-to-simulation correction factors ρ for the lepton identification efficiencies.

connect the $Z(\to \nu\bar{\nu})+\gamma$ and $W(\to \ell\nu)+\gamma$ background yields in the signal regions, where the quantity $N^{W\gamma}$ is the number of $W(\to \ell\nu)+\gamma$ events in the combined signal regions. When calculating the ratio $f_{W\gamma}^{Z\gamma}$, all experimental uncertainties associated with the data-to-simulation correction factors ρ cancel since both processes result in very similar event configurations. The main uncertainties in $f_{W\gamma}^{Z\gamma}$ are those from higher-order theoretical corrections, discussed in Section 1.6.1. Figure 1-23 shows the effect of each systematic uncertainty in $f_{W\gamma}^{Z\gamma}$ with respects to its nominal value for $Z(\to \nu\bar{\nu})+\gamma$ and $W(\to \ell\nu)+\gamma$, respectively.

The ratio $f_{W\gamma}^{Z\gamma}$ rises rather than falls with increasing E_{T}^{γ} because $W(\to \ell\nu)+\gamma$ events have a lower signal region selection efficiency when the charged lepton falls within the tracker acceptance due to the lepton veto while the $Z(\to \nu\bar{\nu})+\gamma$ efficiency is independent of E_{T}^{γ}. Figure 1-24 shows the transfer factor $f_{W\gamma}^{Z\gamma}$ between the $Z(\to \nu\bar{\nu})+\gamma$ and $W(\to \ell\nu)+\gamma$ processes in the combined signal region.

Using $R_{\ell\gamma}^{W\gamma}$ and $f_{W\gamma}^{Z\gamma}$, the total estimated event yield $T_{\ell\gamma}$ in each single-lepton
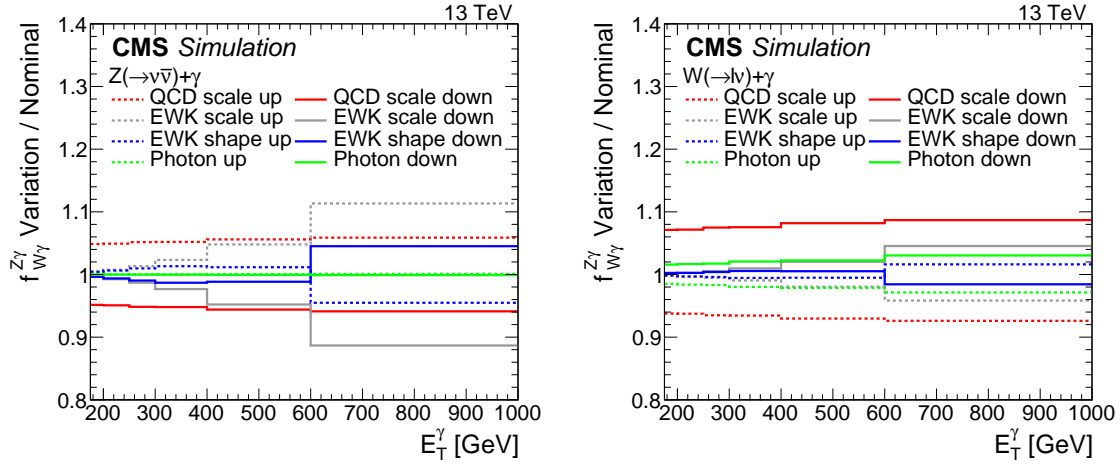
Figure 1-23: Systematic uncertainties on the transfer factor $f_{W\gamma}^{Z\gamma}$ due to higher-order corrections on the $Z(\to \nu\bar{\nu})+\gamma$ process (left) and the $W(\to \ell\nu)+\gamma$ process (right). The last bin includes all events with $E_T^\gamma > 1000\,\mathrm{GeV}$.
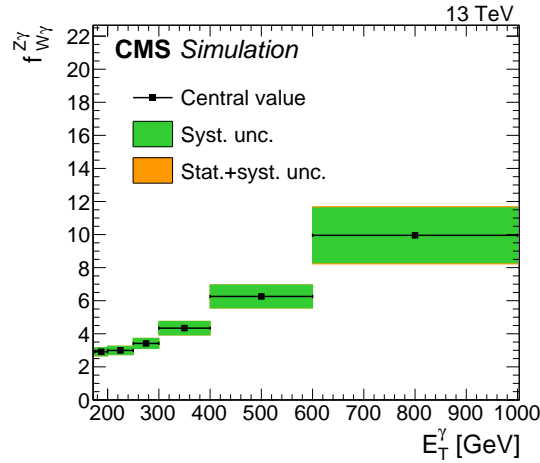

Figure 1-24: Transfer factor $f_{W\gamma}^{Z\gamma}$ in the combined signal regions. The numerator is the expected $Z(\to \nu\bar{\nu})+\gamma$ yield and the denominator is the expected $W(\to \ell\nu)+\gamma$ yield. The uncertainty bands in green (inner) and orange (outer) show the systematic uncertainty, and the combination of systematic and statistical uncertainty arising from limited MC sample size, respectively. The systematic uncertainties considered are the uncertainties from higher-order theoretical corrections.

48

control region in the ith bin of the E_T^γ distribution is given by

$$T_{\ell\gamma,i} = \frac{N_i^{Z\gamma}}{R_{\ell\gamma,i}^{W\gamma} f_{W\gamma,i}^{Z\gamma}} + b_{\ell\gamma,i}, \qquad (1.8)$$

where $b_{\ell\gamma}$ is the predicted contribution from other background sources in the single-lepton regions, namely misidentified electrons and hadrons and other minor SM processes.

1.7   Beam halo

[I have to say I don't remember all these control regions well and the fact that the definitions are somewhere else really confuses me. I really think it would be good to put a reminder or table or discussion in each section.] Bremsstrahlung photons emitted by beam halo muons in the ECAL volume generate a physical EM shower in the ECAL crystals [7]. Large energy deposits are rare, but the rate of beam halo penetration during the 2016 run was substantial. The characteristic features of a shower caused by a halo particle include coincident hits in the barrel muon system and a "trail" of low-energy clusters in ECAL along the particle trajectory. The beam halo MET filter described in Section ?? exploits the former, while the E_MIP variable described in Section ?? captures the latter.

Because beam halo particles are produced through complex LHC machine effects, the observed distribution of the halo showers is not symmetric in the azimuthal angle in the detector coordinates. The right side of Figure 1-25 is a ϕ^γ distribution of the halo showers obtained from the SinglePhoton dataset, requiring $E_\text{T}^\text{miss} > 140\,\text{GeV}$. Halo showers are defined as those that fail the MIP-tagging and those in events tagged by the CSC beam halo tagger. On the other hand, reconstructed showers from all other sources are symmetric in ϕ^γ as shown on the left side of Figure 1-25.

To ensure that the the distribution of Fig. 1-25 is a valid template for halo showers and invariant under photon selection requirements, the ϕ^γ distribution is folded such that the two peaks of the halo showers coincide. To match the positions of the peaks
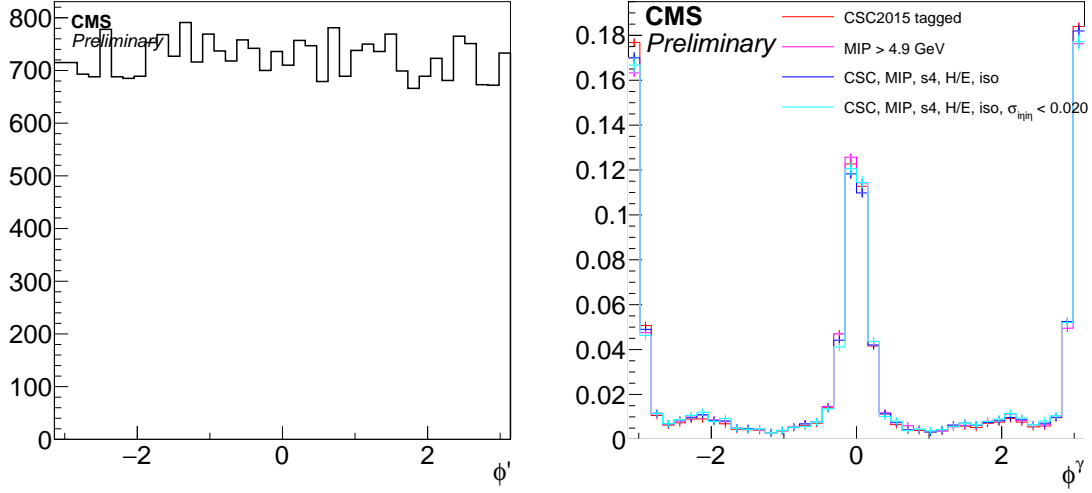
Figure 1-25: Left: The ϕ^γ distribution from $Z(\to \nu\bar{\nu}) + \gamma$ MC simulation. Right: The ϕ^γ distribution of the halo-like showers, tagged in multiple ways. Histograms are normalized to unity. The cyan histogram is the ϕ^γ distribution after applying photon identification selections except for the shower shape. It can be seen that the ϕ^γ distribution is highly stable against the listed identification selections.

in the halo template, the distribution is shifted by 0.005 and then folded along 0. The new angular variable ϕ'

$$\phi' := \left| \left[[\phi^\gamma + 0.005]_{-\pi}^{\pi} - \frac{\pi}{2} \right]_{-\pi}^{\pi} \right| - \frac{\pi}{2}, \tag{1.9}$$

where $[\cdot]_\pi^\pi$ signifies casting the content into range $[-\pi, \pi]$, exhibits a unimodal distribution for the halo template, as shown in Fig. **??**.

The contribution of real photons into the halo control sample is negligible. This is confirmed from the $\sigma_{i\eta i\eta}$ distribution of the halo control sample and the correlation between $\sigma_{i\eta i\eta}$ and E_{MIP} in a MC true-photon sample. The $\sigma_{i\eta i\eta}$ distribution of the halo control sample features a small peak at $\sigma_{i\eta i\eta} \sim 0.01$, which can be attributed to contributions from true photons, as the photon $\sigma_{i\eta i\eta}$ distribution overlaid in Figure 1-26 suggests. However, the contribution of true photons diminishes rapidly with increasing $\sigma_{i\eta i\eta}$. Additionally, Figure 1-27 illustrates that the shape of the true-photon $\sigma_{i\eta i\eta}$ does not change significantly with respect to E_{MIP}. From these two observations, we can see that there are is only a negligible number of true photons in the halo control
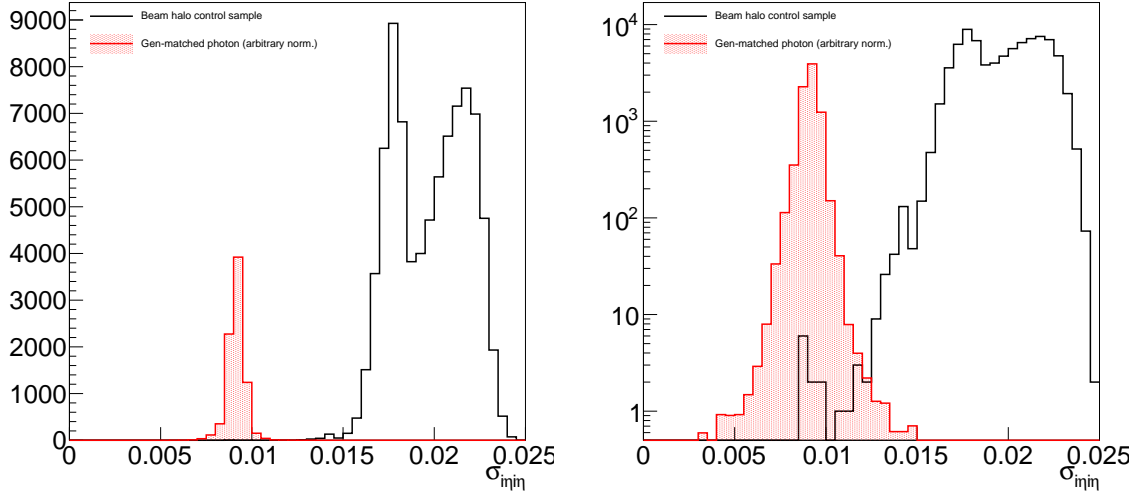
Figure 1-26: The $\sigma_{i\eta i\eta}$distribution of the beam halo control sample and a reference distribution from truth-matched MC photons. Left: linear scale, Right: log scale. There is a small peak at $\sigma_{i\eta i\eta} \sim 0.01$ in the beam halo control sample, which is not visible in linear-scale.

sample.

While the peaking behavior is a robust feature of the halo showers, their rate is not easily predictable. Therefore, a two-template fit to the ϕ' distribution of the photons in the candidate sample, where the templates are a uniform distribution (Figure 1-25 left) and that of the halo shower (Figure **??**), accurately estimates the amount of beam halo background present in the signal region. For this analysis, the splitting of the signal region functions in a similar manner, enabling us to determine the beam halo contribution during the signal extraction procedure.

In the horizontal (H) and vertical (V) signal regions, collision processes occupy the relative fractions of phase space $C_H = 1/\pi$ and $C_V = (\pi - 1)/\pi$ corresponding to $|\phi'| < 0.5$ and $0.5 < |\phi'| < \pi/2$, respesctively. The corresponding fractions for beam halo events are determined by selecting a halo-enriched sample where the halo veto is inverted. Thus, a fit of the two signal regions provides an estimate of the overall normalization of the beam halo background, denoted h.

The E_T^γ dependence of the halo background is encoded in $n_{K,i}^{\text{halo}}$, the unit-normalized beam halo prediction in the ith bin of the signal region $K \in \{H, V\}$. Using the
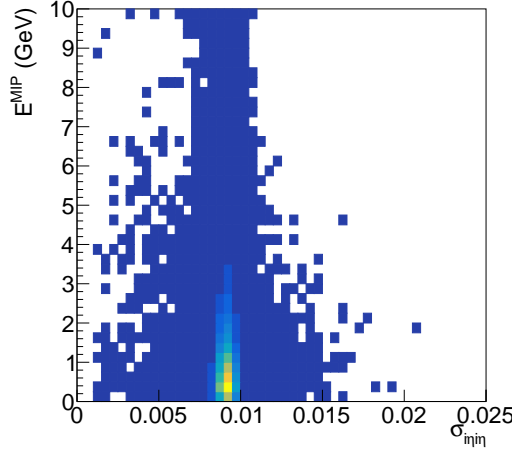
Figure 1-27: Correlation between $\sigma_{i\eta i\eta}$ and E_{MIP} in truth-matched MC photons.

notation introduced in Section 1.6, the total estimated background T_K in the two signal regions are

$$
\begin{aligned}
T_{K,i} &= C_K \left(N_i^{Z\gamma} + N_i^{W\gamma} + b_{K,i} \right) + h n_{K,i}^{\mathrm{halo}} \\
&= C_K \left(\left[1 + \left(f_{W\gamma,i}^{Z\gamma} \right)^{-1} \right] N_i^{Z\gamma} + b_{K,i} \right) + h n_{K,i}^{\mathrm{halo}},
\end{aligned}
\tag{1.10}
$$

where $b_{K,i}$ is the total contribution to bin i of region K from electron and hadron misidentification, ECAL spikes, and other minor SM background processes.

1.8 ECAL spikes

Noise in the photodetector or the detector electronics can result in spurious photon signals. Most of the time, such spurious signal is filtered out by multiple layers of protection, starting from the so-called "spike killer" algorithm at the Level 1 trigger [8]. Nevertheless, in rare cases, noise in a single ECAL channel coincides with pileup or other energy deposit in the nearby crystals and appear as a high-energy photon cluster.

The origin of ECAL spikes is believed to be interactions of neutrons and other hadronic particles (collectively called neutral hadrons hereafter) with the photocathode material of the ECAL avalanche photo diodes (APD) [9]. Nuclear fission at the
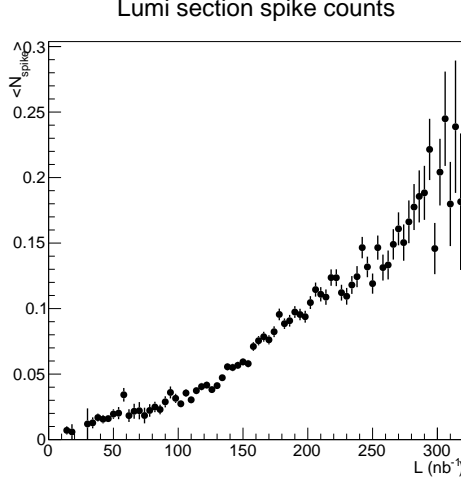
52

Figure 1-28: Average number of spike clusters in a luminosity section, identified by $\sigma_{i\eta i\eta} < 0.001$ and $E > 50\,\mathrm{GeV}$, in muon-triggered events, versus integrated luminosity of the luminosity section.

APD surface then causes a large electron avalanche, which is mistaken as a large photon yield scintillation in the ECAL crystal. Evidences supporting this hypothesis is documented in Reference [8]. In Figure 1-28, scaling of the rate of spikes with the instantaneous luminosity is confirmed, up to much higher luminosity values than was observed at the time when Reference [8] was written.

A known feature of such spurious photon clusters is that the recorded pulse shape of the seed crystal, i.e., the channel with the noise, is not what is expected from a real electromagnetic shower in ECAL. In particular, this translates to a distinctive early rec hit time distribution, since the rec hit time is extracted from a fit to the pulse shape assuming a normal pulse.

In the normal CMS data reconstruction, rec hits that are tagged as spike-like are ignored in clustering. Rec hits are tagged as spikes if there is very little energy deposit recorded in the surrounding crystals, or if the reconstructed time is out of an allowed window. Identical algorithms are employed in the HLT and offline reconstructions.

To study an unbiased spike sample, ECAL DIGI samples stored in the SingleMuon AOD datasets are reconstructed into ECAL clusters with no spike cleaning applied. DIGIs associated with the standard and "uncleaned" photon objects are stored in AOD, and ones for the uncleaned photons is rich in spike-like hits. Figure 1-29 shows
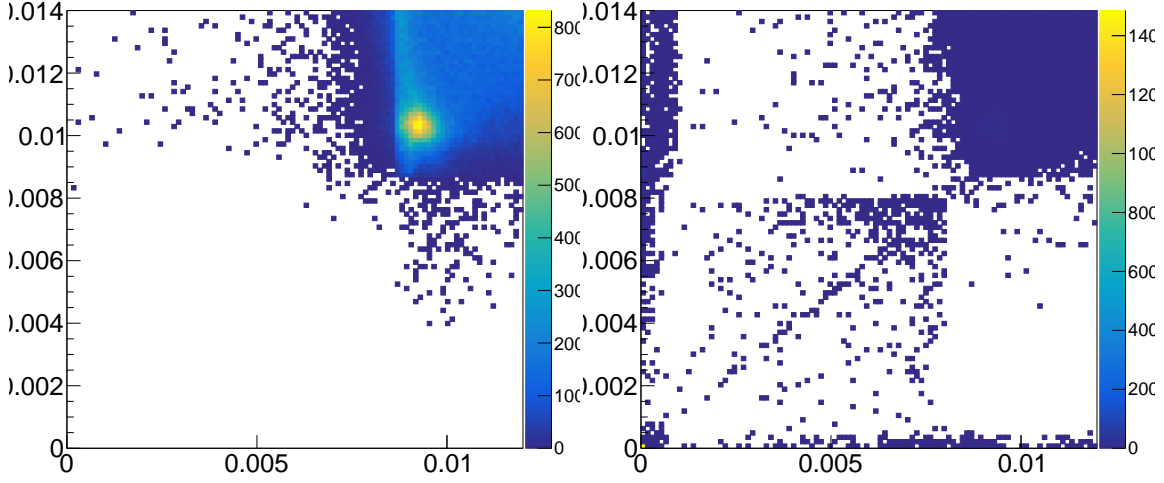
Figure 1-29: Two-dimensional distributions in $\sigma_{i\phi i\phi}$ and $\sigma_{i\eta i\eta}$ of ECAL clusters in the standard reconstruction (left) and the special reconstruction with no spike cleaning (right).

how narrow clusters are cleaned away in the normal reconstruction.

Figure 1-30 shows the spacial and temporal distributions of the rec hits seeding narrow ($\sigma_{i\eta i\eta} < 0.001$) clusters. The spacial distribution appears mostly random, indicating that there is no single source of spike-like rec hits. The two highest peaks in the time distribution at $t \sim -15$ ns and $t \sim 10$ ns are characteristic of pulse shapes, which rise faster than the pulse from the normal scintillation. The second peak is understood to come from the next bunch crossing.

The small peak at $t \sim 0$ in the time distribution of Fig. 1-30 is due to actual "physical" clusters that happened to have a very narrow cluster shape. By processing the γ+jets MC simulation events through this special reconstruction, we see that about 0.5% of ECAL clusters from prompt photons have $\sigma_{i\eta i\eta} < 0.001$ as shown in Figure 1-31.

To understand the time distribution, one can investigate the original DIGI samples from which rec hits are made. At each event readout, a single ECAL channel outputs 10 ADC signals corresponding to a sampling of the analog pulse output of multi-gain preamplifier (MGPA) in range $t_0 - 125$ ns $< t < t_0 + 100$ ns, where t_0 is the time of the triggering bunch crossing. These 10 signal points can be described well by the
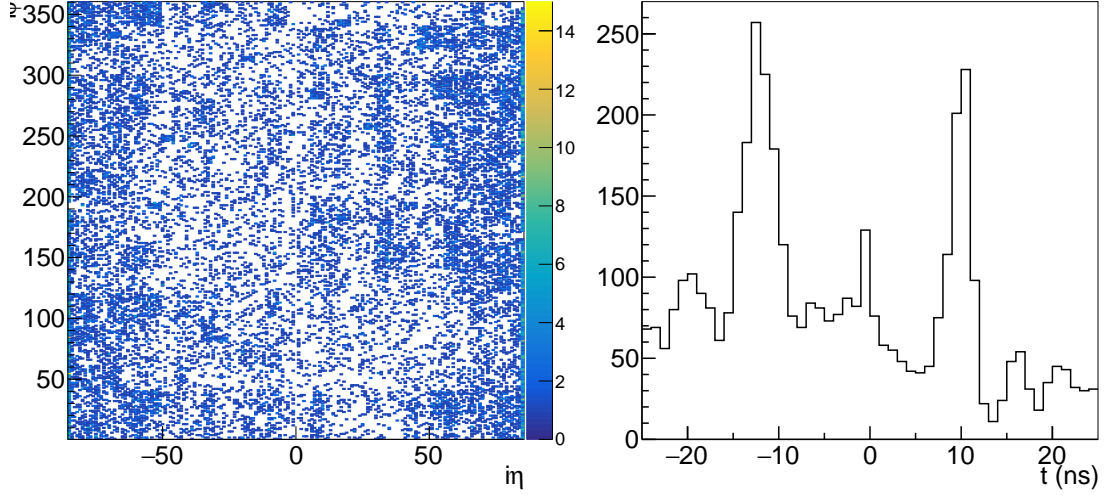
Figure 1-30: η–ϕ and time distributions of seed hits of narrow ($\sigma_{i\eta i\eta} < 0.001$) clusters.
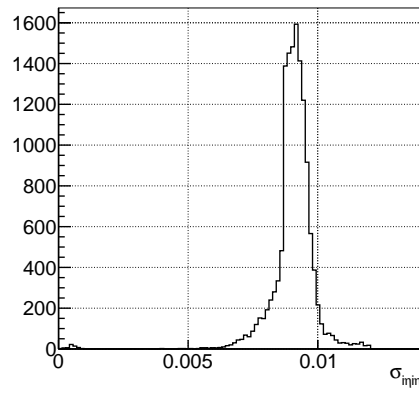

Figure 1-31: $\sigma_{i\eta i\eta}$ distribution of uncleaned clusters from γ+jets MC simulation.
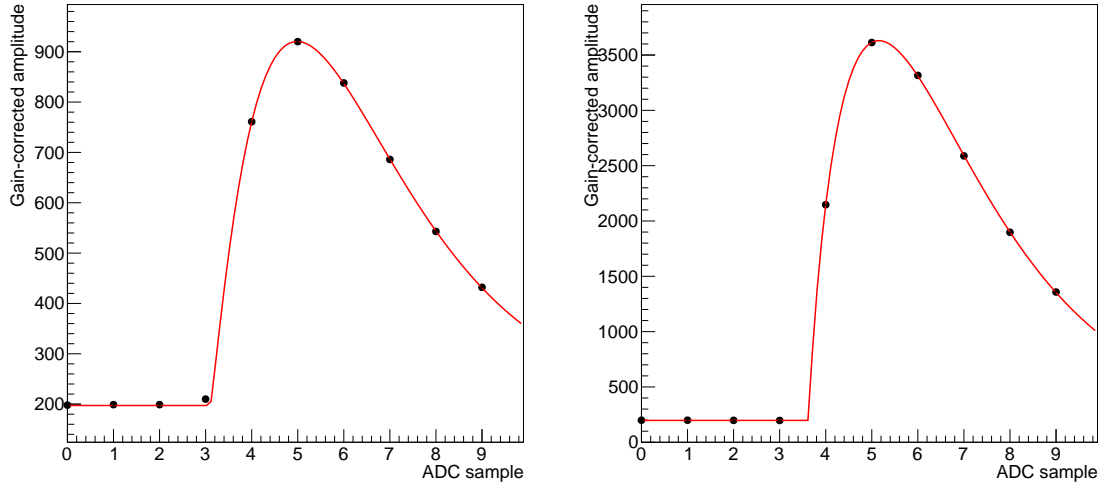
Figure 1-32:   Example ECAL DIGIs and corresponding pulse shapes reconstructed through χ^2 fits of Equation 1.11, for normal (left) and spike-like (right) hits.

formula

$$f(t) = A \left(1 - \frac{t - \tau}{\alpha \beta} \right)^{\alpha} \exp\left(-\frac{t - \tau}{\beta} \right).$$ (1.11)
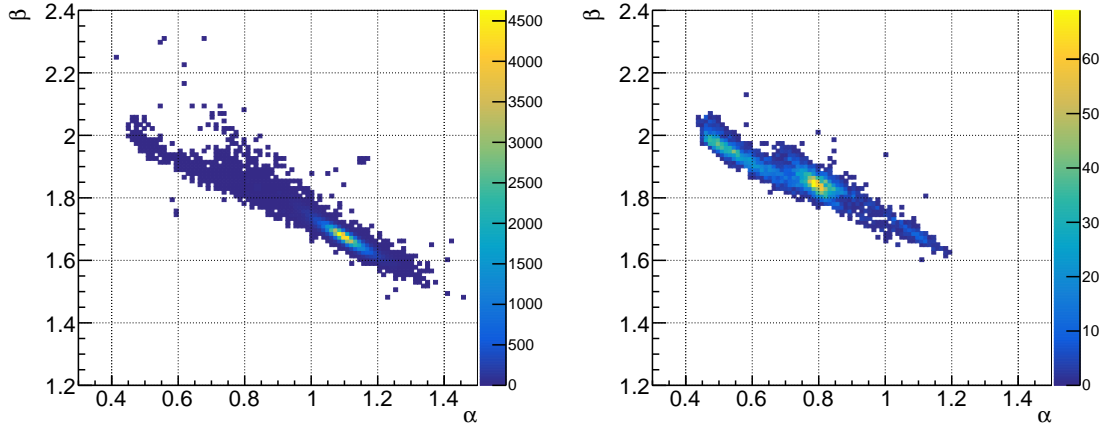

Figure 1-33: α–β distributions of the seed hits of physical wide clusters (left) and spike-like clusters (right).

In the formula, parameters A and τ correspond to the pulse amplitude and peak time, whereas α and β control the shape of the pulse. Figure 1-32 illustrates various observed pulse shapes fit with the above formula with all parameters floating. A χ^2

fit is employed using the average noise amplitude of each MGPA channel as the errors on the data points. The noise is measured in ECAL calibration cycles in the inter-fill period and is recorded in the conditions database.
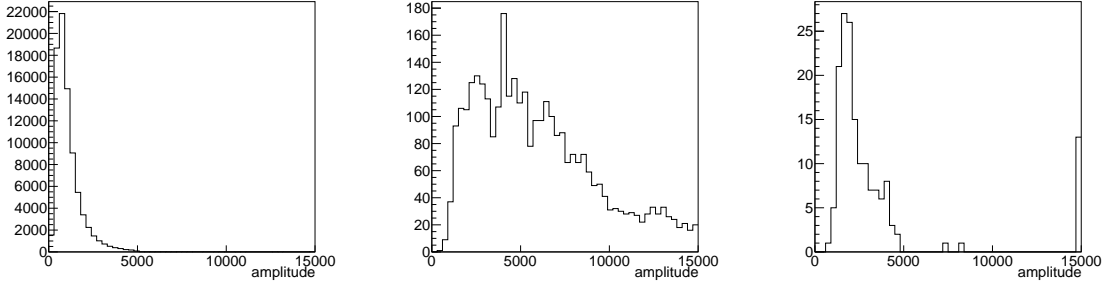

Figure 1-34: Seed crystal pulse amplitude distributions of physical wide clusters (left), narrow clusters with $\alpha < 0.9$ (center), and narrow clusters with $\alpha > 0.9$ (right).

In the α–β parameter space, seed rec hits of wide clusters concentrate around $(\alpha, \beta) \sim (1.1, 1.7)$, while spike-like hits populate the region $\alpha < 0.9$ as shown in Figure 1-33. In fact, the pulse amplitude distribution of narrow-cluster seeds with $\alpha > 0.9$ is unlike that of the narrow-cluster seeds with $\alpha < 0.9$, and resembles the amplitude distribution of wide-cluster seeds shown in Figure 1-34. This suggests that the population $\alpha > 0.9$ correspond to the clusters of physical, prompt photons. It then follows that spike hits can be regarded to exclusively have sharp pulse shapes. [I am not sure you can definitively conlclude this, but all it does is just put a bound. So the method is fine, but I think you should be more exacting on the language.]

Given these observations, the time distribution of spike-like rec hits outside of the window $-15\,\mathrm{ns} < t < -10\,\mathrm{ns}$ (and the equivalent with one-bunch-crossing shift) is understood to be due to delayed interactions of neutral hadrons with the APDs, as documented also in Reference [8]. In other words, ECAL spike clusters which survive the time cleaning cut of the standard reconstruction are a part of a broad tail of a distribution, and there is no evidence of spike signals that specifically populate the "in-time" region $-3\,\mathrm{ns} < t < 3\,\mathrm{ns}$.

Having established that there is no special population of ECAL spikes in the in-time region, we can estimate the number D of ECAL spike events present in the signal

candidate sample to be

$$D = C \times \frac{B}{A}, \tag{1.12}$$

where

- A = Number of clusters with $\sigma_{i\eta i\eta}$ or $\sigma_{i\phi i\phi}$ less than 0.001 and seed time $-15\,\mathrm{ns} < t < -10\,\mathrm{ns}$, counted in the special-reconstruction sample,

- B = Number of clusters with both $\sigma_{i\eta i\eta}$ and $\sigma_{i\phi i\phi}$ greather than 0.001 and seed time $-15\,\mathrm{ns} < t < -10\,\mathrm{ns}$, counted in the special-reconstruction sample,

- C = Number of clusters with $\sigma_{i\eta i\eta}$ or $\sigma_{i\phi i\phi}$ less than 0.001 but an in-time seed, counted in the standard-reconstruction sample passing all other signal event selection.

The special-reconstruction samples for A and B are from the SinglePhoton datasets, with only the timing cleaning removed from the offline reconstruction. In this way, the selection bias over spikes from the L1T, HLT, and offline reconstruction is equally applied to samples A, B, and C. Plugging in the observed numbers, we have

$$A = 4969$$
$$B = 1180$$
$$C = 54$$
$$\therefore D = 12.8 \pm 1.8 (\mathrm{stat.})$$

There are, however, at least two reasons to believe that this method overestimates the number of spike events in the signal region. One is that the population C contains some physical, prompt photon clusters that just happen to be narrow, as observed in Fig. 1-31. Another reason is that there is likely a correlation between the cluster width and the seed time such that the ratio of true D to C is smaller than B/A. This statement is based on the standard hypothesis that the wide-cluster spike is an ECAL spike embedded in a physical EM shower cluster. Under this model, spikes in wide clusters are mainly caused by prompt neutral hadrons in a jet, which implies

that they strongly prefer seed time $-15\,\mathrm{ns} < t < -10\,\mathrm{ns}$. Given that this is a minor background with a relatively large uncertainty, as described below, we will still use this estimate as the nominal value of predicted spike contribution in the signal region.

The uncertainty in the estimate of D is evaluated by two modifications to A, B, and C. First, the three values are recomputed with using $\sigma_{i\eta i\eta} < 0.001$ as the only definition of narrow cluster. This results in a minor change of the value of D of 12.1 ± 1.7. Next, A and B are computed using a lower-p_T SinglePhoton sample, requiring triggers `Photon135_PFMET100` or `Photon120_R9Id90_HE10_IsoM` to have fired, instead of the signal trigger. The second modification gives $D = 9.1 \pm 1.3$. We then take twice the discrepancy between the nominal D and the D value from the second modification to obtain a 33% systematic uncertainty on the spike background estimate.

1.9   Minor SM Backgrounds

After the full selection described in Section 1.2, the SM γ+jets, $t\bar{t}\gamma$, $VV\gamma$, $Z(\to \ell\bar{\ell})+\gamma$, and $W \to \ell\nu$ processes are minor ($\sim10\%$) background processes in the signal region. These processes, collectively denoted as minor SM backgrounds, contribute in the signal region when the jet energy is severely mismeasured or the leptons fail to be reconstructed resulting in large $E_\mathrm{T}^\mathrm{miss}$ in the signal region. However, the $E_\mathrm{T}^\mathrm{miss}$ is typically aligned with the photon or one of the jets in such cases, and therefore various selections on the kinematic relations between the $E_\mathrm{T}^\mathrm{miss}$, photons, and jets are used to reduce these backgrounds to a manageable rate. The estimates for all five processes are taken from MADGRAPH5_aMC@NLO simulations at LO in QCD and are listed in Tables 1.14 and 1.15.

1.10   Statistical Interpretation

The potential signal contribution is extracted from the data via simultaneous fits to the E_T^γ distributions in the signal and control regions defined in Section 1.2. Un-

certainties in various quantities are represented by nuisance parameters in the fit. Predictions for $Z(\to \nu\bar{\nu})+\gamma$, $W(\to \ell\nu)+\gamma$, and the beam halo backgrounds are varied in the fit. Beam halo is not a major background, but the extraction of its rate requires a fit to the observed distributions in the signal region.

Free parameters of the fit are the yield of $Z(\to \nu\bar{\nu})+\gamma$ background in each bin of the signal regions ($N_i^{Z\gamma}$) and the overall normalization of the beam halo background (h). Bin-by-bin yields of $W(\to \ell\nu)+\gamma$ and $Z(\to \ell\bar{\ell})+\gamma$ samples in all regions are related to the yield of $Z(\to \nu\bar{\nu})+\gamma$ through the MC prediction through the transfer factors defined in Section 1.6. The transfer factors are allowed to shift within the aforementioned theoretical and experimental uncertainties. The background-only likelihood that is maximized in the fit is

$$
\begin{aligned}
\mathcal{L} &= \prod_i \left\{ \mathcal{L}_{\text{signal}} \times \mathcal{L}_{\text{single-lepton}} \times \mathcal{L}_{\text{dilepton}} \right\} \times \mathcal{L}_{\text{nuisances}} \\
&= \prod_i \left\{ \prod_{K=H,V} \mathcal{P}\left(d_{K,i} \Big| T_{K,i}(\vec{\theta})\right) \times \prod_{\ell=e,\mu} \mathcal{P}\left(d_{\ell\gamma,i} \Big| T_{\ell\gamma,i}(\vec{\theta})\right) \times \prod_{\ell=e,\mu} \mathcal{P}\left(d_{\ell\ell\gamma,i} \Big| T_{\ell\ell\gamma,i}(\vec{\theta})\right) \right\} \times \prod_j \mathcal{N}(\theta_j) \\
&= \prod_i \left\{
\begin{aligned}
&\prod_{K=H,V} \mathcal{P}\left(d_{K,i} \Big| C_K \left(\left[1 + \left(f_{W\gamma,i}^{Z\gamma}(\vec{\theta})\right)^{-1}\right] N_i^{Z\gamma} + b_{K,i}(\vec{\theta})\right) + h n_{K,i}^{\text{halo}}(\vec{\theta})\right) \\
&\qquad \times \prod_{\ell=e,\mu} \mathcal{P}\left(d_{\ell\gamma,i} \Big| \frac{N_i^{Z\gamma}}{R_{\ell\gamma,i}^{W\gamma}(\vec{\theta}) f_{W\gamma,i}^{Z\gamma}(\vec{\theta})} + b_{\ell\gamma,i}(\vec{\theta})\right) \\
&\qquad \times \prod_{\ell=e,\mu} \mathcal{P}\left(d_{\ell\ell\gamma,i} \Big| \frac{N_i^{Z\gamma}}{R_{\ell\ell\gamma,i}^{Z\gamma}(\vec{\theta})} + b_{\ell\ell\gamma,i}(\vec{\theta})\right)
\end{aligned}
\right\} \times \prod_j \mathcal{N}(\theta_j),
\end{aligned}
$$

$$(1.13)$$

following the notation introduced in Section 1.6, and where $\mathcal{P}(n|\lambda)$ is the Poisson probability of n for mean λ, $\mathcal{N}$ denotes the unit normal distribution, and $d_{K,i}$ is the observed number of events in bin i of region K. Systematic uncertainties on R_e, R_h, $k_{\text{QCD}}^{\text{NNLO}}$, $k_{\text{EW}}^{\text{Sudakov}}$, $k_{\text{EW}}^{q\gamma}$, and the various scale factors ρ are treated as nuisance parameters in the fit and are represented by $\vec{\theta}$. Quantities such as $R_{\ell\ell\gamma}^{Z\gamma}$, $R_{\ell\gamma}^{W\gamma}$, $f_{W\gamma}^{Z\gamma}$, n_K^{halo}, and b_K appear in the likelihood function as $Q(\vec{\theta}) = \overline{Q} \sum_j \exp(\sigma_j \theta_j)$, where $\overline{Q}$ is the nominal value of the quantity and σ_j is standard deviation of each systematic uncertainty θ_j on the quantity.

1.11 Results

Using the likelihood from the previous section, we perform a maximum likelihood fit to the observed E_T^γ distributions in all the control samples and signal region simultaneously. The minimization of the negative log likelihood is performed using the Minuit2 algorithm through an interface provided in the RooStats package [10].
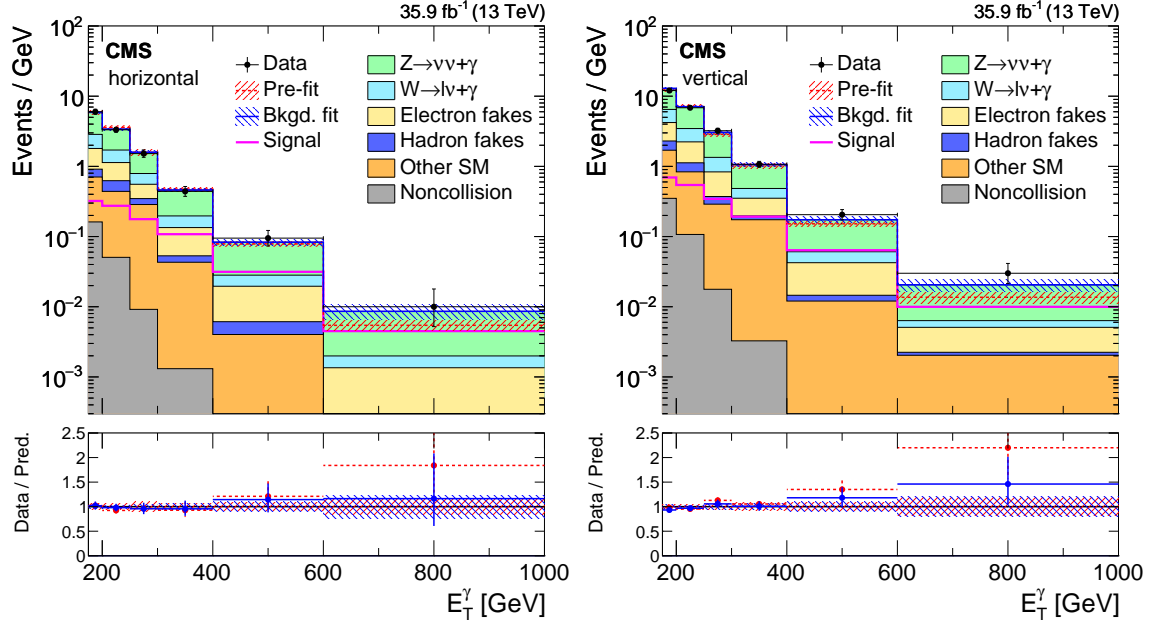

Figure 1-35: Observed E_T^γ distributions in the horizontal (left) and vertical (right) signal regions compared with the post-fit background expectations for various SM processes. The last bin of the distribution includes all events with $E_T^\gamma > 1000\,\mathrm{GeV}$. The expected background distributions are evaluated after performing a combined fit to the data in all the control samples and the signal region. The ratios of data with the pre-fit background prediction (red dashed) and post-fit background prediction (blue solid) are shown in the lower panels. The bands in the lower panels show the post-fit uncertainty after combining all the systematic uncertainties. The expected signal distribution from a $1\,\mathrm{TeV}$ vector mediator decaying to $1\,\mathrm{GeV}$ DM particles is overlaid.

The comparison between the observed distributions and the results from simulations before and after performing the simultaneous fit are shown in Figures 1-35 and 1-36 for the signal and control regions, respectively. The observed distributions are in agreement with the prediction from SM and non-collision backgrounds and no significant excess of events beyond the SM expectation is observed.
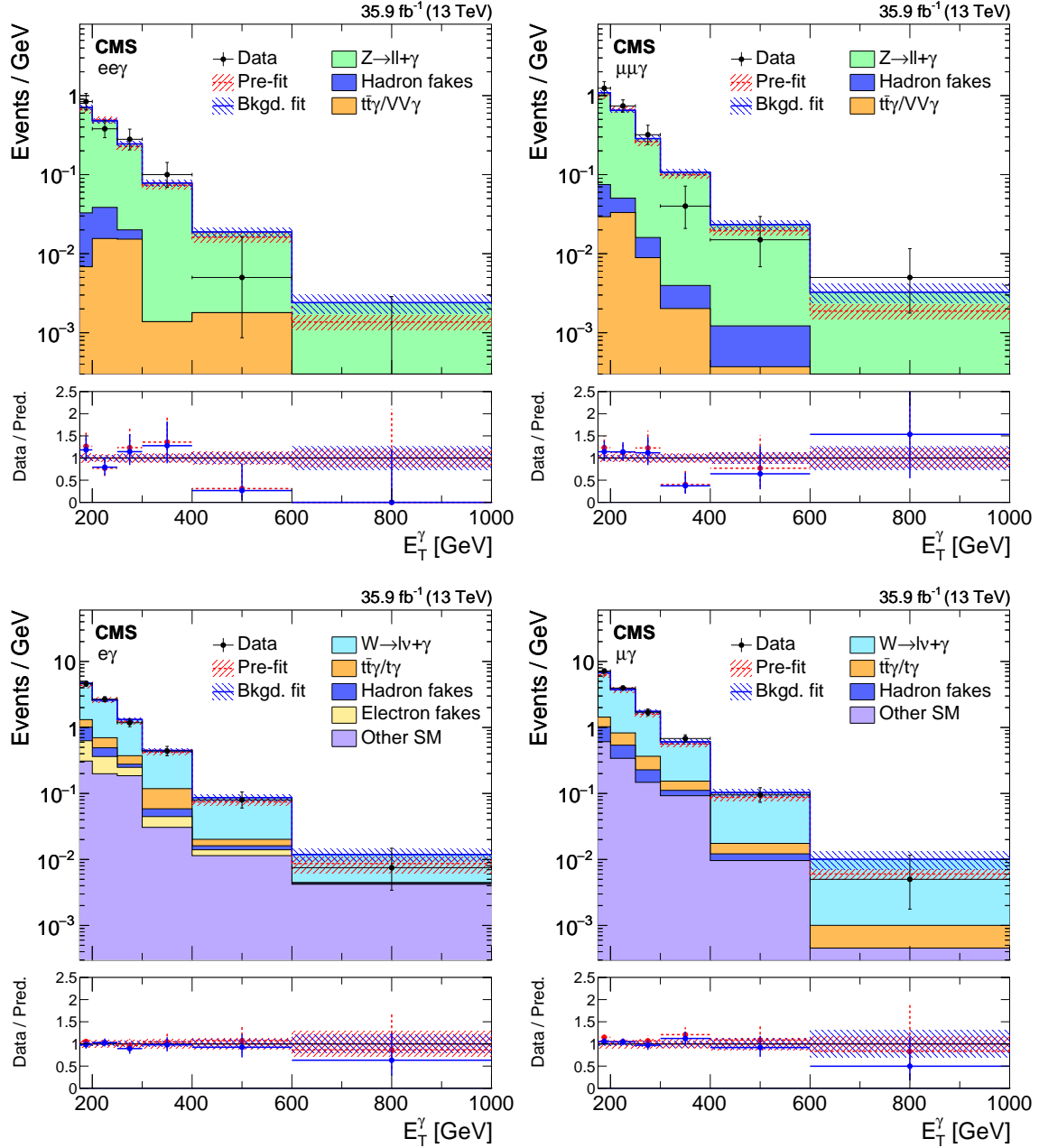
61

Figure 1-36: Comparison between data and MC simulation in the four control regions: $ee\gamma$ (upper left), $\mu\mu\gamma$ (upper right), $e\gamma$ (lower left), $\mu\gamma$ (lower right) before and after performing the simultaneous fit across all the control samples and signal region, and assuming absence of any signal. The last bin of the distribution includes all events with $E_T^\gamma > 1000\,\mathrm{GeV}$. The ratios of data with the pre-fit background prediction (red dashed) and post-fit background prediction (blue solid) are shown in the lower panels. The bands in the lower panels show the post-fit uncertainty after combining all the systematic uncertainties.

1.11.1 Limits

[Can you pick a few mass points and show the expected and and observed and how it fits in the band. I don't see a discussion ofwhether there is an excess or not.] Upper limits are determined for the production cross section of the new-physics processes mentioned in Section **??**. For each model, a 95% confidence level (CL) upper limit is obtained utilizing the asymptotic CL_s criterion [11–13], using a test statistic based on the negative logarithm of the likelihood in Section 1.10.
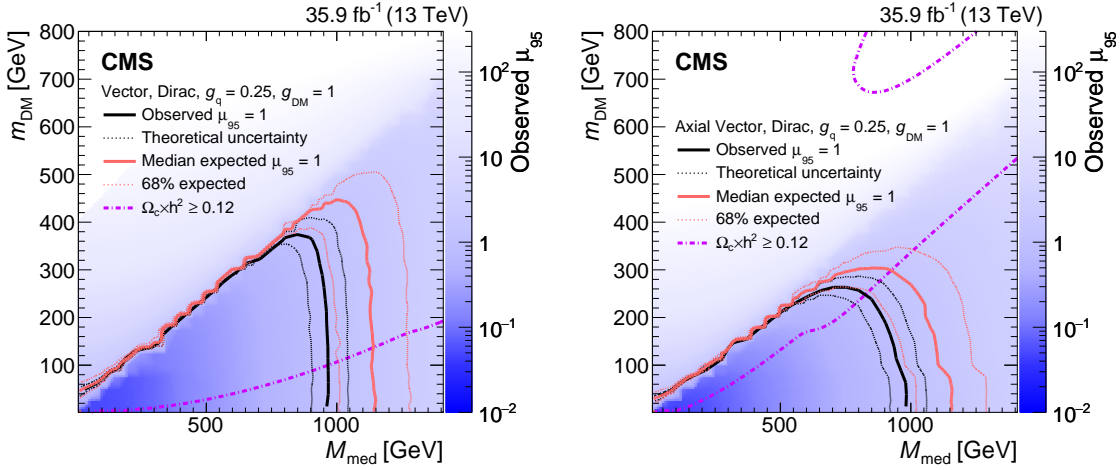

Figure 1-37: The ratio of 95% CL upper cross section limits to the theoretical cross section (μ_{95}), for DM simplified models with vector (left) and axial-vector (right) mediators, assuming $g_q = 0.25$ and $g_{DM} = 1$. The expected $\mu_{95} = 1$ contours are overlaid in red. The region under the observed contour is excluded. For DM simplified model parameters in the region below the lower violet dot–dash contour, and also above the corresponding upper contour in the right hand plot, cosmological DM abundance exceeds the density observed by the Planck satellite experiment.

Figure 1-37 shows the 95% CL upper cross section limits with respect to the corresponding theoretical cross section ($\mu_{95} = \sigma_{95\%}/\sigma_{\text{theory}}$) for the vector and axial-vector mediator scenarios, in the M_{med}–m_{DM} plane. The solid black (dashed red) curves are the observed (expected) contours of $\mu_{95} = 1$. The σ_{theory} hypothesis is excluded at 95% CL or above in the region with $\mu_{95} < 1$. The uncertainty in the expected upper limit includes the experimental uncertainties. For the simplified DM

LO models considered, mediator masses up to $950\,\text{GeV}$ are excluded for values of m_DM less than $1\,\text{GeV}$.

The results for vector and axial-vector mediators are compared to constraints from the observed cosmological relic density of DM as determined from measurements of the cosmic microwave background by the Planck satellite experiment [14]. The expected DM abundance is estimated, separately for each model, using the thermal freeze-out mechanism implemented in the MADDM [15] framework and compared to the observed cold DM density $\Omega_c h^2 = 0.12$ [14], where Ω_c is the DM relic abundance and h is the dimensionless Hubble constant.

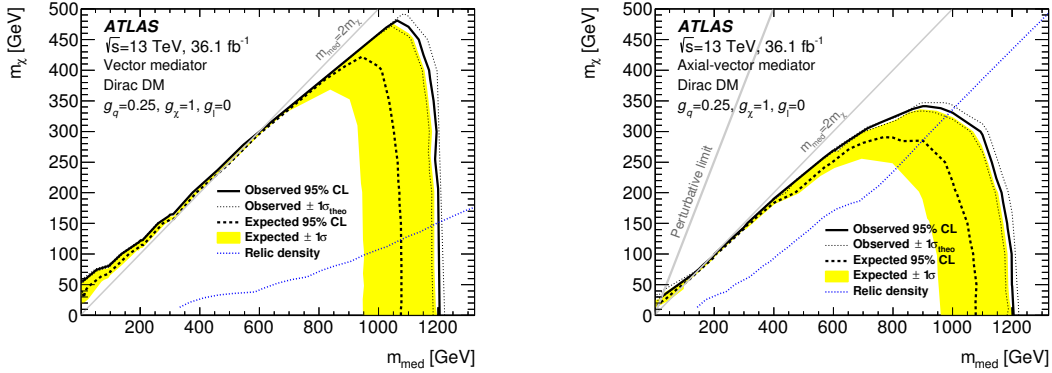## 1.11.2 Comparison to Results from Other Experiments

Figure 1-38: The equivalent of Figure 1-37 with results from ATLAS. The expected $\mu_{95} = 1$ contours are overlaid in yellow. The region under the observed contour is excluded. Reprinted from Reference [16]

The results from a similar analysis by ATLAS [16] are shown in Figure 1-38, with mediator masses up to $1200\,\text{GeV}$ are excluded for values of m_DM less than $1\,\text{GeV}$. These observed results are stronger than those shown in this thesis, even though the expected limits are worse, and the following differences between the two analyses explain why. The event selection differs by a requirement of $E_\text{T}^\gamma > 150\,\text{GeV}$ facilitated by a lower trigger threshold and a veto for events with more than one jet. In terms of analysis strategy, the ATLAS analysis does not have the single electron control region or the transfer factor $f_{W\gamma}^{Z\gamma}$ linking the $Z(\to \nu\bar{\nu})+\gamma$ and $W(\to \ell\nu)+\gamma$ background

yields but does have a low E_T^{miss} control region to estimate the γ+jets background. Additionally, the signal extraction fit uses only three bins of E_T^{γ} with lower boundaries of 150 GeV, 225 GeV, and 300 GeV. Finally, the ATLAS result has a slight pre-fit deficit of observed events, while the result shown here have a slight pre-fit excess.
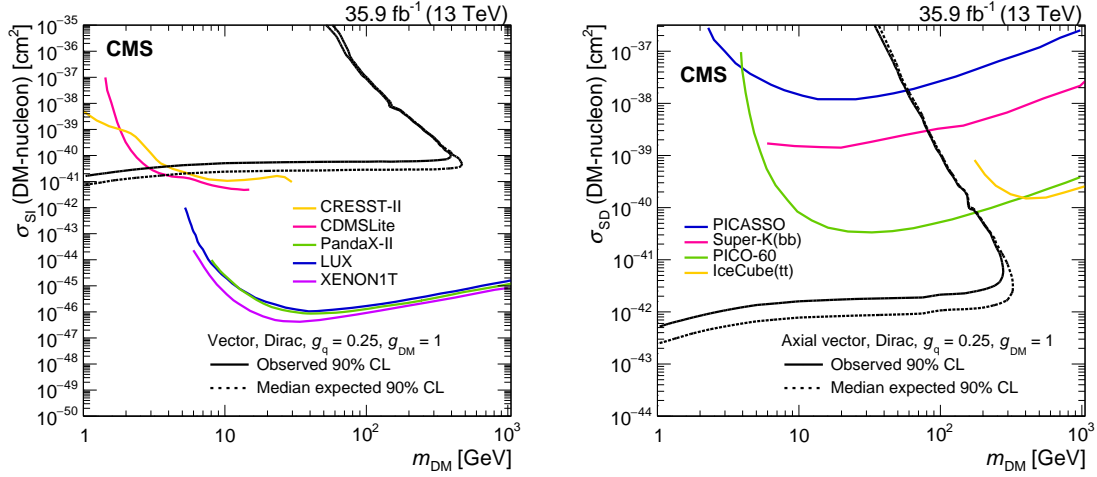

Figure 1-39: The 90% CL exclusion limits on the χ–nucleon spin-independent scattering cross sections involving the vector operator (top) and the χ–nucleon spin-dependent scattering cross sections involving the axial-vector operator (bottom) as a function of the m_{DM}. Simplified model DM parameters of $g_q = 0.25$ and $g_{DM} = 1$ are assumed. The region to the upper left of the contour is excluded. On the plots, the median expected 90% CL curve overlaps the observed 90% CL curve. Also shown are corresponding exclusion contours, where regions above the curves are excluded, from the recent results by the direct and indirect detection experiments listed in the text.

To enable a direct comparison with results from direct and indirect detection experiments, the 95% CL limits on the mediator mass for the vector and axial-vector models are translated to 90% CL limits on the spin-independent and spin-dependent DM–nucleon scattering cross sections, σ_{SI} and σ_{SD} respectively, following the prescriptions given in Reference [17] and [18]. The exclusion contours for the vector and axial-vector models shown in Figure 1-37 are translated into the σ_{SI}–m_{DM} and σ_{SD}–m_{DM} planes shown in Figure 1-39. When compared to recent results by the CDMSLite [19], LUX [20], PandaX-II [21], XENON1T [22], and CRESST-II [23] collaborations, the limits obtained from this search provide stronger constraints for DM masses less than 2 GeV for spin independent models. When compared to recent re-

sults by the PICO-60 [24], IceCube [25], PICASSO [26] and Super-Kamiokande [27] collaborations, the limits obtained from this search provide stronger constraints for DM masses less than $200\,\mathrm{GeV}$ for spin dependent models.

1.11.3   Interpretation of Additional Models

In addition to the background-only fit to all of the signal and control regions, we performed a simultaneous maximum likelihood fit to the observed E_T^γ distributions in the control regions only. The results of this fit enable the interpretation of new physics models not studied in this thesis with the simplified likelihood approach detailed in Reference [28]. The predicted background yields in each bin of E_T^γ in the horizontal and vertical signal regions after performing the control region only fit are given in Tables 1.14 and 1.15, respectively. The covariances between the predicted background yields across all the E_T^γ bins in the two signal regions are shown in Fig. 1-40.

E_T^γ [GeV]	[175, 200]	[200, 250]	[250, 300]	[300, 400]	[400, 600]	[600, 1000]
$Z\gamma$	81.2 ± 8.0	88.2 ± 8.4	38.8 ± 4.8	26.8 ± 3.7	8.8 ± 1.9	1.4 ± 0.7
$W\gamma$	27.9 ± 3.7	29.9 ± 3.9	11.4 ± 1.7	6.3 ± 1.2	1.4 ± 0.4	0.1 ± 0.1
Misid. electrons	22.5 ± 2.7	25.7 ± 2.7	10.5 ± 1.0	8.2 ± 0.7	2.7 ± 0.2	0.5 ± 0.0
Misid. hadrons	5.2 ± 2.2	9.3 ± 1.8	3.1 ± 0.7	1.0 ± 0.3	0.4 ± 0.1	0.0 ± 0.0
Other SM	13.6 ± 2.0	19.6 ± 1.3	13.9 ± 0.4	4.2 ± 0.2	0.8 ± 0.0	0.1 ± 0.0
ECAL spikes	4.3 ± 1.3	2.7 ± 0.8	0.5 ± 0.1	0.1 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Total prediction	154.6 ± 8.3	175.4 ± 8.8	78.2 ± 5.3	46.6 ± 4.0	14.1 ± 2.1	2.1 ± 0.8
Observed	150 ± 12	166 ± 13	76.0 ± 8.7	44.0 ± 6.6	19.0 ± 4.4	4.0 ± 2.0

Table 1.14: Expected event yields in each E_T^γ bin for the various background processes in **the horizontal signal region**. The background yields and the corresponding uncertainties are obtained after performing a combined fit to data in all the control samples, excluding data in the signal region. The observed event yields in the horizontal signal region are also reported.
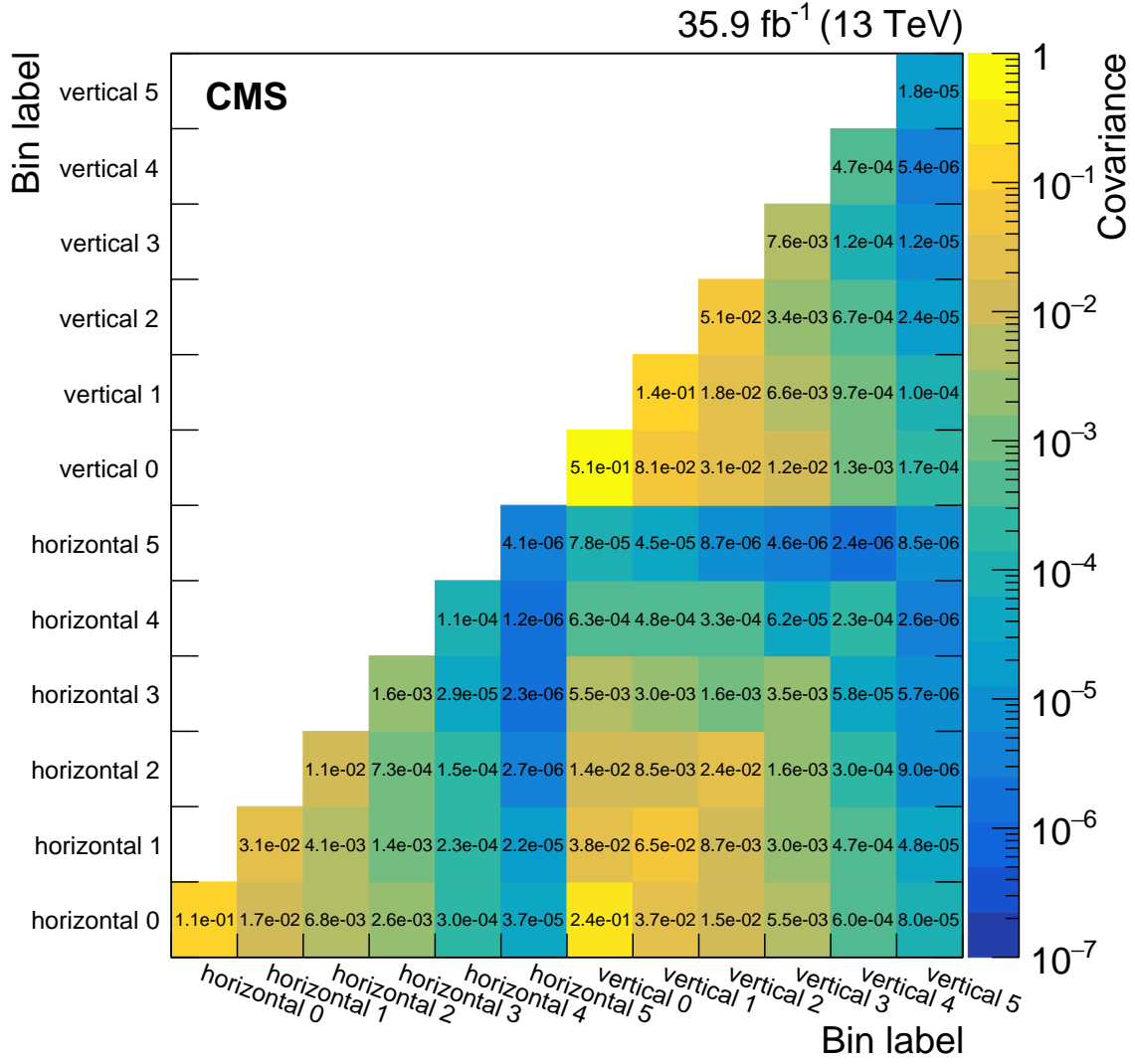
.

Figure 1-40: Covariances between the predicted background yields in all the E_T^γ bins of the horizontal and vertical signal regions. The bin labels specify which signal region the bin belongs to and what number bin it is for that region.

E_{T}^{γ} [GeV]	[175, 200]	[200, 250]	[250, 300]	[300, 400]	[400, 600]	[600, 1000]
$Z\gamma$	172 ± 17	190 ± 18	83 ± 10	58.6 ± 7.9	18.0 ± 3.9	3.1 ± 1.6
$W\gamma$	59.9 ± 7.8	63.6 ± 7.8	24.6 ± 3.5	13.4 ± 2.4	3.0 ± 0.8	0.3 ± 0.2
Misid. electrons	48.4 ± 5.6	56.2 ± 5.1	23.4 ± 1.8	15.7 ± 1.4	5.6 ± 0.4	1.2 ± 0.1
Misid. hadrons	15.1 ± 4.4	14.5 ± 3.1	4.2 ± 0.8	2.3 ± 0.8	0.5 ± 0.1	0.1 ± 0.1
Other SM	33.8 ± 4.1	36.6 ± 2.7	13.6 ± 0.5	17.1 ± 0.6	2.4 ± 0.1	0.8 ± 0.0
ECAL spikes	9.3 ± 2.8	5.7 ± 1.7	0.9 ± 0.3	0.3 ± 0.1	0.0 ± 0.0	0.0 ± 0.0
Total prediction	339 ± 18	366 ± 19	150 ± 11	107.5 ± 8.7	29.6 ± 4.3	5.4 ± 1.7
Observed	301 ± 17	342 ± 19	161 ± 13	107 ± 10	41.0 ± 6.4	12.0 ± 3.5

Table 1.15: Expected event yields in each E_{T}^{γ} bin for the various background processes in **the vertical signal region**. The background yields and the corresponding uncertainties are obtained after performing a combined fit to data in all the control samples, excluding data in the signal regions. The observed event yields in the vertical signal region are also reported.

Bibliography

[1] *CMS luminosity measurements for the 2016 data taking period*, CMS Physics Analysis Summary CMS-PAS-LUM-17-001 (CERN, 2017).

[2] A. M. Sirunyan et al., "Search for new physics in final states with an energetic jet or a hadronically decaying w or z boson and transverse momentum imbalance at $\sqrt{s}$= 13 TeV", Phys. Rev. D **97**, 092005 (2018).

[3] G. Bozzi, S. Catani, G. Ferrera, D. de Florian, and M. Grazzini, "Production of Drell-Yan lepton pairs in hadron collisions: Transverse-momentum resummation at next-to-next-to-leading logarithmic accuracy", Phys. Lett. **B696**, 207–213 (2011).

[4] A. Denner, S. Dittmaier, M. Hecht, and C. Pasold, "NLO QCD and electroweak corrections to W+γ production with leptonic W-boson decays", JHEP **04**, 018 (2015).

[5] A. Denner, S. Dittmaier, M. Hecht, and C. Pasold, "NLO QCD and electroweak corrections to Z+γ production with leptonic Z-boson decays", JHEP **02**, 057 (2016).

[6] J. M. Lindert et al., "Precise predictions for $V+$ jets dark matter backgrounds", Eur. Phys. J. **C77**, 829 (2017).

[7] M. J. Ambrose et al., "A novel beam halo monitor for the CMS experiment at the LHC", JINST **10**, P11011 (2015).

[8] D. Petyt et al., *Characterization and treatment of anomalous signals in the cms electromagnetic calorimeter*, CMS Note 2010/357 (2010).

[9] D. A. Petyt et al., "Mitigation of anomalous APD signals in the CMS electromagnetic calorimeter", J. Phys.: Conf. Series **404**, 012043 (2012).

[10]K. Belasco et al., "The RooStats Project", PoS **ACAT2010**, Comments: 11 pages, 3 figures, ACAT2010 Conference Proceedings, 057 (2010).

[11]T. Junk, "Confidence level computation for combining searches with small statistics", Nucl. Instrum. Meth. A **434**, 435 (1999).

[12]A. L. Read, "Presentation of search results: the cl_s technique", J. Phys. G **28**, 2693 (2002).

[13]G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics", Eur. Phys. J. C **71**, [Erratum: 10.1140/epjc/s10052-013-2501-z], 1554 (2011).

[14]P. A. R. Ade et al., "Planck 2015 results. XIII. Cosmological parameters", Astron. Astrophys. **594**, A13 (2016).

[15]M. Backović, K. Kong, and M. McCaskey, "MadDM v.1.0: computation of dark matter relic abundance using MadGraph5", Phys. Dark Univ. **5-6**, 18 (2014).

[16]M. Aaboud et al., "Search for dark matter at $\sqrt{s} = 13$ tev in final states containing an energetic photon and large missing transverse momentum with the atlas detector", The European Physical Journal C **77**, 393 (2017).

[17]G. Busoni et al., "Recommendations on presenting LHC searches for missing transverse energy signals using simplified s-channel models of dark matter", 2016.

[18]D. Abercrombie et al., "Dark Matter Benchmark Models for Early LHC Run-2 Searches: Report of the ATLAS/CMS Dark Matter Forum", edited by A. Boveia, C. Doglioni, S. Lowette, S. Malik, and S. Mrenna (2015).

[19]R. Agnese et al., "New results from the search for low-mass weakly interacting massive particles with the cdms low ionization threshold experiment", Phys. Rev. Lett. **116**, 071301 (2016).

[20]D. S. Akerib et al., "Results from a search for dark matter in the complete lux exposure", Phys. Rev. Lett. **118**, 021303 (2017).

[21]X. Cui et al., "Dark matter results from 54-ton-day exposure of PandaX-II experiment", Phys. Rev. Lett. **119**, 181302 (2017).

[22] E. Aprile et al., "Dark matter search results from a one Tonne × Year exposure of XENON1T", 2018.

[23] G. Angloher et al., "Results on light dark matter particles with a low-threshold CRESST-II detector", Eur. Phys. J. C **76**, 25 (2016).

[24] C. Amole et al., "Dark matter search results from the PICO−60 C3f8 bubble chamber", Phys. Rev. Lett. **118**, 251301 (2017).

[25] M. Aartsen et al., "Improved limits on dark matter annihilation in the sun with the 79-string IceCube detector and implications for supersymmetry", Journal of Cosmology and Astroparticle Physics **2016**, 022–022 (2016).

[26] E. Behnke et al., "Final results of the PICASSO dark matter search experiment", Astropart. Phys. **90**, 85 (2017).

[27] K. Choi et al., "Search for neutrinos from annihilation of captured low-mass dark matter particles in the Sun by Super-Kamiokande", Phys. Rev. Lett. **114**, 141301 (2015).

[28] The CMS Collaboration, *Simplified likelihood for the re-interpretation of public CMS results*, CMS Note CMS-NOTE-2017-001 (2017).