

**Shining Light on Dark Matter,
One Photon at a Time**

by

Brandon Leigh Allen

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

Doctorate of Science in Physics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2019

© Massachusetts Institute of Technology 2019. All rights reserved.

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

Shining Light on Dark Matter, One Photon at a Time

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 Science 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

1	The Monophoton Analysis	9
1.1	Dataset	10
1.1.1	Trigger Efficiency	11
1.2	Event Selection	12
1.2.1	Signal Regions	14
1.2.2	Control Regions	14
1.2.3	Proxy Samples	15
1.2.4	Measurement Samples	16
1.3	Irreducible backgrounds	16
1.3.1	Simulation of $V+\gamma$ Processes	16
1.3.2	Data-driven Control Regions	20
1.4	Misidentified electrons	26
1.5	Misidentified hadrons	27
1.6	Beam halo	32
1.7	Spikes	33
1.8	Minor SM Backgrounds	35
1.9	Statistical Interpretation	35
1.10	Results	36
1.10.1	Pre-fit and post-fit distributions	36
1.10.2	Limits	40

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_T^{miss} , whether with real photons, other particles misidentified as photons, or unphysical photon signatures from various machine and detector processes.

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 $Z(\rightarrow \nu\bar{\nu})+\gamma$ and $W(\rightarrow \ell\nu)+\gamma$, 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 1.3.

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 mismea-

sured 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 contributions from these processes are described in detail in Sections 1.6 and 1.7, 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.

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 $E_T^\gamma > 175 \text{ GeV}$ and the integrated luminosity of the analyzed data sample is $(35.9 \pm 0.9) \text{ fb}^{-1}$ [?].

1.1.1 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 because photons leave no tracks and do not bend in the magnetic field.

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 without the pixel seed veto requirement and a muon object that passes the “tight” identification 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 $E_T > 175 \text{ GeV}$.

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 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.

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,

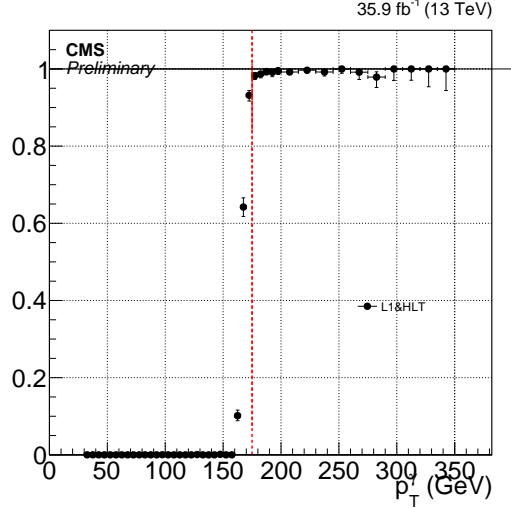


Figure 1-1: The efficiency turn-on of the HLT_Photon165_HE10 trigger for photons passing the candidate selection, measured using $\mu + e/\gamma$ events from the SingleMuon data set. Red vertical line corresponds to $E_T = 175$ GeV.

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.

1.2 Event Selection

From the recorded data, events are selected by requiring $E_T^{\text{miss}} > 170$ GeV and at least one photon with $E_T^\gamma > 175$ GeV in the fiducial region of the ECAL barrel ($|\eta| < 1.44$). Events with photons in the endcaps are not considered because the estimate of backgrounds due to beam halo and misidentified hadron are greatly complicated due to the x - y grid of the crystals in the endcaps.

Events with a high- p_T photon and large E_T^{miss} are subjected to further requirements to suppress SM background processes that feature a genuine high-energy photon, but not a significant amount of E_T^{miss} . One such SM process is γ +jets, where an apparently large E_T^{miss} is often the result of mismeasuring the energy of a jet. In contrast

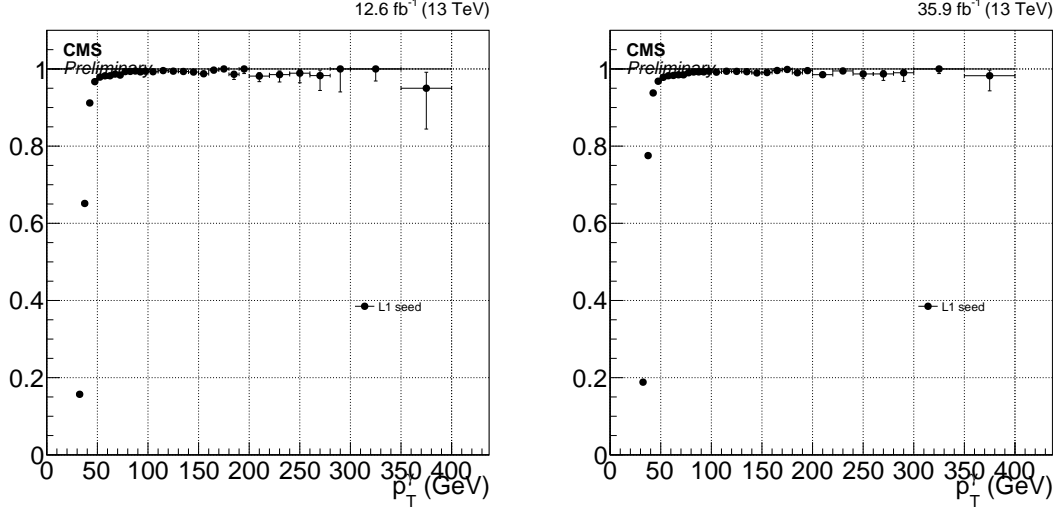


Figure 1-2: The efficiency of the L1 seed for the signal trigger in periods B and C (left) and the full data set (right). The drop in efficiency at high- E_T in the earlier period is fixed by the addition of **SingleJet** L1 seeds during the remainder of data-taking.

to signal-like processes, the E_T^{miss} is typically smaller than E_T^γ in these events, so requiring the ratio of E_T^γ to E_T^{miss} to be less than 1.4 rejects this background effectively with little effect on signal efficiency. Events are also rejected if the minimum opening angle between \vec{p}_T^{miss} and the directions of the four highest p_T jets, $\min\Delta\phi(\vec{p}_T^{\text{jet}}, \vec{p}_T^{\text{miss}})$, is less than 0.5. Only jets with $p_T > 30$ GeV and $|\eta| < 5$ are considered in the $\min\Delta\phi(\vec{p}_T^{\text{jet}}, \vec{p}_T^{\text{miss}})$ calculation. In the γ +jets process, rare pathological mismeasurements of E_T^γ also lead to large E_T^{miss} . For this reason, the candidate photon \vec{p}_T and \vec{p}_T^{miss} must be separated by more than 0.5 radians.

Variable	Selection	Motivation
E_T^γ	> 175 GeV	high- p_T photon passing trigger
$ \eta $	< 1.44	region with best background estimates
E_T^{miss}	> 170 GeV	characteristic signature of dark matter
$E_T^\gamma/E_T^{\text{miss}}$	< 1.4	reduce jet mismeasurement backgrounds
$\min\Delta\phi(\vec{p}_T^{\text{jet}}, \vec{p}_T^{\text{miss}})$	< 0.5	reduce jet mismeasurement backgrounds
$\Delta\phi(\vec{p}_T^\gamma, \vec{p}_T^{\text{miss}})$	> 0.5	reduce photon mismeasurement backgrounds

Table 1.1: Baseline selections for all events considered in the analysis.

The above selections, summarized in Table 1.1, constitute the baseline selections

common to all regions. To improve the purity of the signal region, we require a more stringent photon identification as well as additional object vetos. The contributions from the $Z(\rightarrow \nu\bar{\nu})+\gamma$ and $W(\rightarrow \ell\nu)+\gamma$ 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 while 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 various signal, control, and proxy regions used in the analysis are described in the following sections.

1.2.1 Signal Regions

The defining feature of the signal region is the application of both the e/γ and γ -specific portions of the photon ID, given in Tables ?? and ?? respectively. The former reduces the hadron misidentification rate with a collection of isolation and shower shape selections while the latter reduces the electron misidentification rate with the pixel seed veto and rejects non-collisions backgrounds with specifically tailored selections.

In the signal region, events are vetoed if they contain an electron or a muon with $p_T > 10$ 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_T^{miss} , and leptons such as $W(\rightarrow \ell\nu)+\gamma$, $t\bar{t}\gamma$, and $VV\gamma$.

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

1.2.2 Control Regions

The single-electron (single-muon) control region is defined by requiring exactly one tight electron (muon) with $p_T > 30$ GeV and $|\eta| < 2.5(2.4)$ in addition to re-

quiring the same photon ID as in the signal regions. To suppress the contributions from large- $E_{\text{T}}^{\text{miss}}$ processes other than $W(\rightarrow \ell\nu)+\gamma$, the transverse mass $m_{\text{T}} = \sqrt{2E_{\text{T}}^{\text{miss}}p_{\text{T}}^{\ell}[1 - \cos \Delta\phi(\vec{p}_{\text{T}}^{\text{miss}}, \vec{p}_{\text{T}}^{\ell})]}$ must be less than 160 GeV. Additionally, for the single-electron control region, $E_{\text{T}}^{\text{miss}}$ must be greater than 50 GeV to limit the contribution from the γ +jets process, where a jet is misidentified as an electron.

The dielectron (dimuon) control region is defined by exactly two electrons (muons) in addition to the photon, with $60 < m_{\ell\ell} < 120$ 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.

Finally, in the control regions, the recoil vector $\vec{U} = \vec{p}_{\text{T}}^{\text{miss}} + \sum_{\ell} \vec{p}_{\text{T}}^{\ell}$ serves as an analogue for the $\vec{p}_{\text{T}}^{\text{miss}}$ in the signal region. In the signal region, the $\vec{p}_{\text{T}}^{\text{miss}}$ is a proxy for the vector boson p_{T} while in the control regions, the recoil vector is used instead. Thus, the recoil \vec{U} must satisfy identical requirements to those for the $\vec{p}_{\text{T}}^{\text{miss}}$ in the signal region to keep the control region kinematics as similar as possible to the signal region kinematics.

1.2.3 Proxy Samples

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 $I_{\text{CH}} < 0.441$ 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.

1.2.4 Measurement Samples

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 $E_{\text{T}} > 175$ GeV and $|\eta| < 1.44$ plus at least one jet with $p_{\text{T}} > 100$ 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 exception of the following relaxed cuts: $\sigma_{i\eta i\eta} < 0.015$ and $I_{\text{CH}} < 11.0$ GeV. Additionally, we apply a $E_{\text{T}}^{\text{miss}} < 60$ 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. These are exactly the same as the hadron proxy samples, except that a high- p_{T} jet has replaced the high- $E_{\text{T}}^{\text{miss}}$, minimizing the kinematic differences between the two.

1.3 Irreducible backgrounds

1.3.1 Simulation of $V+\gamma$ Processes

The $Z(\rightarrow \nu\bar{\nu})+\gamma$ and $W(\rightarrow \ell\nu)+\gamma$ 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 $E_T^\gamma > 130$ GeV are employed for this purpose.

A study using a privately generated 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 1-3 and 1-4 show the comparisons of the private aMC@NLO samples and the MADGRAPH 5 samples used for the background estimation in the key kinematic distributions.

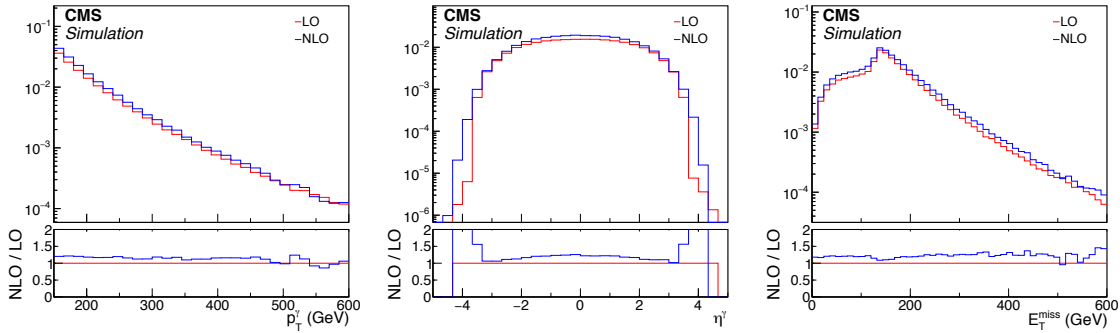


Figure 1-3: Distributions of E_T^γ (left), η^γ (middle), and p_T^Z (right) in $Z(\rightarrow \nu\bar{\nu})+\gamma$ process from the private aMC@NLO sample (blue) and the LO sample used for background prediction (red) along with the NLO / LO ratios.

To approximate the QCD higher-order effects, $Z(\rightarrow \nu\bar{\nu})+\gamma$ and $W(\rightarrow \ell\nu)+\gamma$ events are reweighted with E_T^γ by the factors given in Tab. 1.3. These factors are the ratios of QCD next-to-next-to leading order (NNLO) differential cross sections calculated by Grazzini et al. [?] 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 in the strict sense of the word. $V\gamma$ k-factors found in literature can be $\gg 1$ at high E_T^γ , if the denominator only accounts for the cross section of $q\bar{q} \rightarrow V\gamma$ process.)

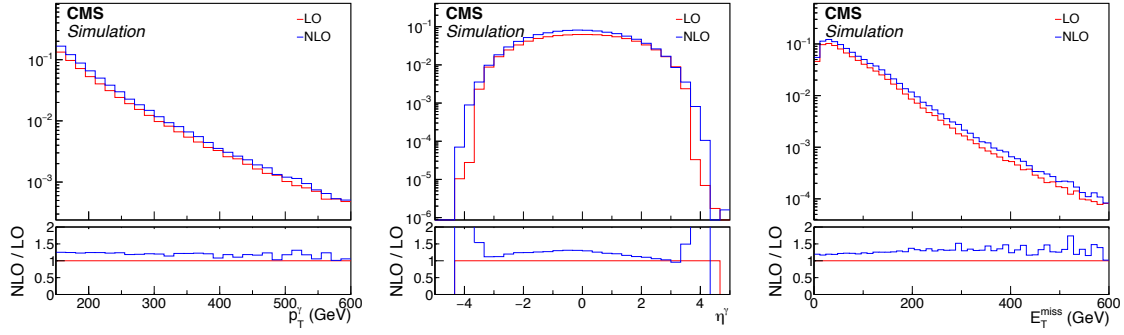


Figure 1-4: Distributions of E_T^γ (top left), η^γ (top right), and p_T^W (bottom left) in $W(\rightarrow \ell\nu)+\gamma$ process from the private aMC@NLO sample (blue) and the LO sample used for background prediction (red) along with the NLO / LO ratios.

E_T^γ range (GeV)	$Z(\rightarrow \nu\bar{\nu})+\gamma$	$W(\rightarrow \ell\nu)+\gamma$
[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, ∞)	1.15	1.15

Table 1.3: NNLO / LO correction factors for $Z(\rightarrow \nu\bar{\nu})+\gamma$ and $W(\rightarrow \ell\nu)+\gamma$ samples.

Additionally, 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 Sudakov suppression at high boson p_T and potentially the addition of photon-induced scattering processes [?, ?]. We apply the correction factors shown in Figure 1-5, which are combinations of Sudakov suppression factors and photon-induced enhancements, and are provided by the authors of Ref. [?] in addition to the NNLO QCD correction.

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

$$d\sigma^{\text{NNLO QCD+NLO EW}} = d\sigma^{\text{LO}} k^{\text{NNLO QCD}} (1 + \kappa^{\text{EW Sudakov}} + \kappa^{\text{EW } q\gamma}), \quad (1.1)$$

where $k^{\text{NNLO QCD}} = d\sigma^{\text{NNLO QCD}}/d\sigma^{\text{LO}}$, and the two κ terms are the Sudakov suppression and photon-induced enhancement components of the electroweak correction,

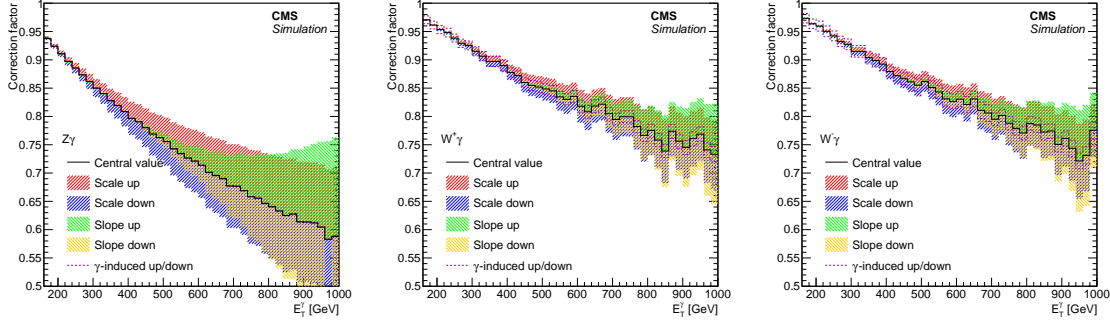


Figure 1-5: Electroweak NLO cross section corrections as a function of photon p_T for $Z(\rightarrow \nu\bar{\nu})+\gamma$ (left), $W^+ + \gamma$ (middle), and $W^- + \gamma$ (right) processes, overlaid with uncertainty bands. See text for descriptions of the individual components of the uncertainty. The uncertainty due to γ -induced production is negligible in $Z(\rightarrow \nu\bar{\nu})+\gamma$ production.

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. Further details on the measurement and values of various ρ are found in Chapter ??.

Four sources of systematic uncertainties considered for E_T^γ distribution ratios of the $V+\gamma$ 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 all considered as correlated between the E_T^γ bins.

The higher-order QCD renormalization and factorization scale uncertainties on the NNLO cross sections are assessed by varying the respective scales by factor 2 and 0.5 during the cross section computation. These uncertainties are between 7-8%, varying bin by bin, 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. We estimate the magnitude of the uncertainty on $\kappa^{\text{EW Sudakov}}$ and $\kappa^{\text{EW}q\gamma}$ to be $(\kappa^{\text{EW Sudakov}})^2$ and $\kappa^{\text{EW}q\gamma}$, 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 $\kappa^{\text{EW Sudakov}}$ is motivated by the fact that fully resummed leading-log Sudakov suppression is an exponential of $\kappa^{\text{EW Sudakov}}$.

For the Sudakov suppression, which is the dominant term in the electroweak correction, we further consider two types of systematic variations, inspired by ref. [?], which provides a prescription for electroweak correction uncertainties for $V + \text{jets}$ processes. 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 1-6). 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 $E_T^\gamma = 590 \text{ GeV}$.

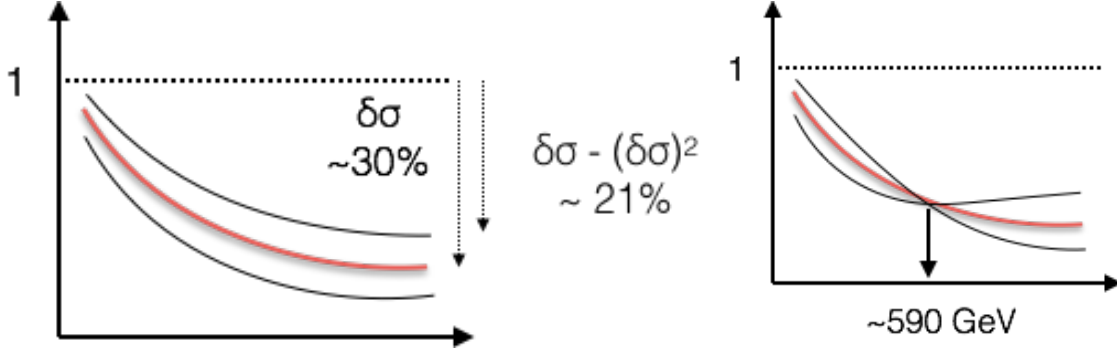


Figure 1-6: Electroweak correction variation scheme to cover the scale (left) and shape (right) uncertainties.

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 $Z(\rightarrow \nu\bar{\nu}) + \gamma$ and $W(\rightarrow \ell\nu) + \gamma$ processes, i.e., the variation of the ratio is bounded by the ratios of the upward and downward variations.

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.

1.3.2 Data-driven Control Regions

Contributions from the $Z(\rightarrow \nu\bar{\nu})+\gamma$ and $W(\rightarrow \ell\nu)+\gamma$ processes are estimated using observed data in four mutually exclusive single-electron, single-muon, dielectron, and dimuon control regions defined in Section 1.2. The ratios between the expected yields of these processes are constrained by MC simulations of $V+\gamma$ processes. This background estimation method exploits cancellation of some of the systematic uncertainties, both experimental and theoretical, in the ratios of the photon E_T^γ distributions of $V+\gamma$ processes, from here on referred to as “transfer factors”.

For example, in the transfer factor between the $Z(\rightarrow \nu\bar{\nu})+\gamma$ and $Z(\rightarrow \ell\bar{\ell})+\gamma$ processes, denoted $R_{\ell\ell\gamma}^{Z\gamma}$, the uncertainties due to photon energy calibration, jet energy resolution, and higher-order QCD effects are significantly reduced 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 and the statistical uncertainty due to the limited MC sample size. Figure 1-7 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, for which the numerator is the expected $Z(\rightarrow \nu\bar{\nu})+\gamma$ yield in the combined signal regions and the denominator is the expected $Z(\rightarrow \ell\bar{\ell})+\gamma$ yield in the relevant control region.

For increasing E_T^γ , the Z boson in a $Z(\rightarrow \ell\bar{\ell})+\gamma$ 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 $Z(\rightarrow \nu\bar{\nu})+\gamma$ 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

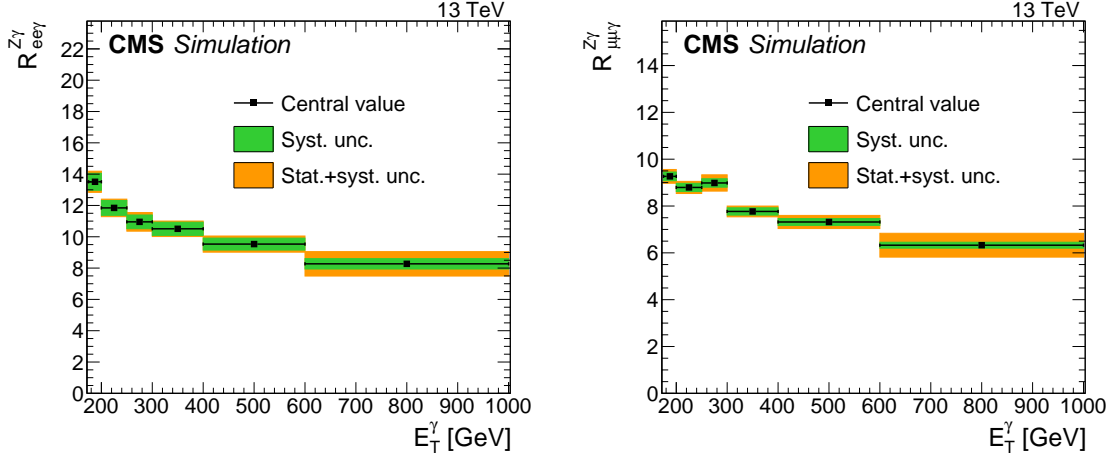


Figure 1-7: Transfer factors $R_{ee\gamma}^{Z\gamma}$ (left) and $R_{\mu\mu\gamma}^{Z\gamma}$ (right). The uncertainty bands in green (inner) and orange (outer) show the systematic uncertainty, and the combination of systematic and statistical uncertainty arising from limited MC sample size, respectively. The systematic uncertainties considered are the uncertainties in the data-to-simulation correction factors ρ for the lepton identification efficiencies.

control region in the i^{th} bin of the E_T^γ distribution can be expressed as

$$T_{\ell\ell\gamma,i} = \frac{N_i^{Z\gamma}}{R_{\ell\ell\gamma,i}^{Z\gamma}} + b_{\ell\ell\gamma,i}, \quad (1.2)$$

where $N^{Z\gamma}$ is the number of $Z(\rightarrow \nu\bar{\nu})+\gamma$ 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 $W(\rightarrow \ell\nu)+\gamma$ processes. 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 1.4 gives the breakdown of the $W(\rightarrow \ell\nu)+\gamma$ background passing the full

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

Table 1.4: 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 was roughly within acceptance ($|\eta| < 2.5$) but failed the lepton veto, and those where the lepton was out of acceptance ($|\eta| > 2.5$). For each W decay mode, the fraction out of total generated ($A \times \epsilon$) is shown.

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 constitutes approximately 60% of the $W(\rightarrow \ell\nu) + \gamma$ background. The remaining 40% of the $W(\rightarrow \ell\nu) + \gamma$ background comes from events where 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, which is lower for electrons than for muons, which translates to a larger background contribution from the electrons. Meanwhile, the E_T^{miss} requirement explains the behavior of the out-of-acceptance background ($|\eta| > 2.5$). A large fraction of electrons that are out of tracker acceptance are still captured by the calorimeters, while out-of-acceptance muons directly contribute to missing momentum leading to a larger background contribution from muons.

Figure 1-8 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 $W(\rightarrow \ell\nu) + \gamma$ yield in the combined signal regions, and the denominator is the estimated $W(\rightarrow \ell\nu) + \gamma$ 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.

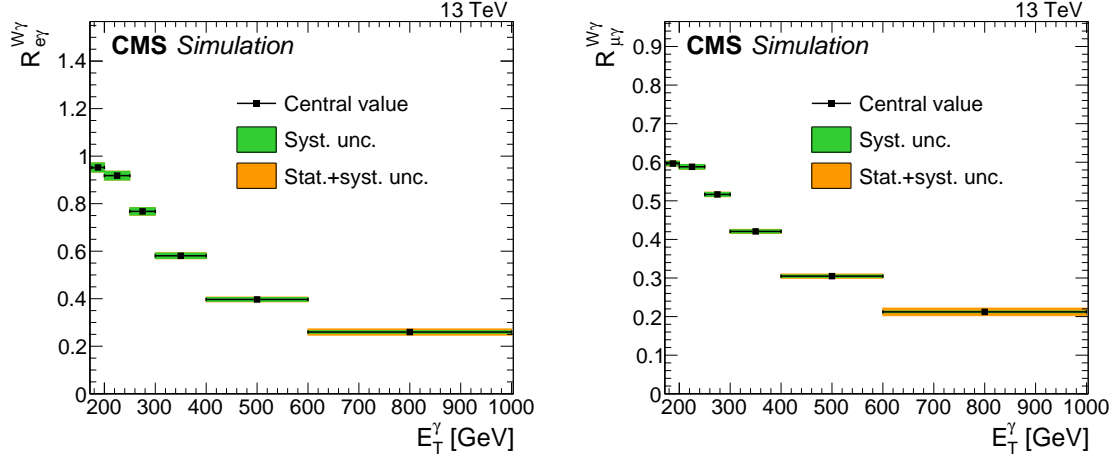


Figure 1-8: Transfer factors $R_{e\gamma}^{W\gamma}$ (left) and $R_{\mu\gamma}^{W\gamma}$ (right). The uncertainty bands in green (inner) and orange (outer) show the systematic uncertainty, and the combination of systematic and statistical uncertainty arising from limited MC sample size, respectively. The systematic uncertainties considered are the uncertainties in the data-to-simulation correction factors ρ for the lepton identification efficiencies.

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 connect the $Z(\rightarrow \nu\bar{\nu})+\gamma$ and $W(\rightarrow \ell\nu)+\gamma$ background yields in the signal regions, where the quantity $N^{W\gamma}$ is the number of $W(\rightarrow \ell\nu)+\gamma$ 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 1.3.1. Figure 1-9 shows the effect of each systematic uncertainty in $f_{W\gamma}^{Z\gamma}$ with respects to its nominal value for $Z(\rightarrow \nu\bar{\nu})+\gamma$ and $W(\rightarrow \ell\nu)+\gamma$ respectively.

The ratio $f_{W\gamma}^{Z\gamma}$ rises rather than falls with increasing E_T^γ because $W(\rightarrow \ell\nu)+\gamma$ events have a lower rather than higher signal region selection efficiency if the charged lepton falls within the tracker acceptance while the $Z(\rightarrow \nu\bar{\nu})+\gamma$ efficiency is indepen-

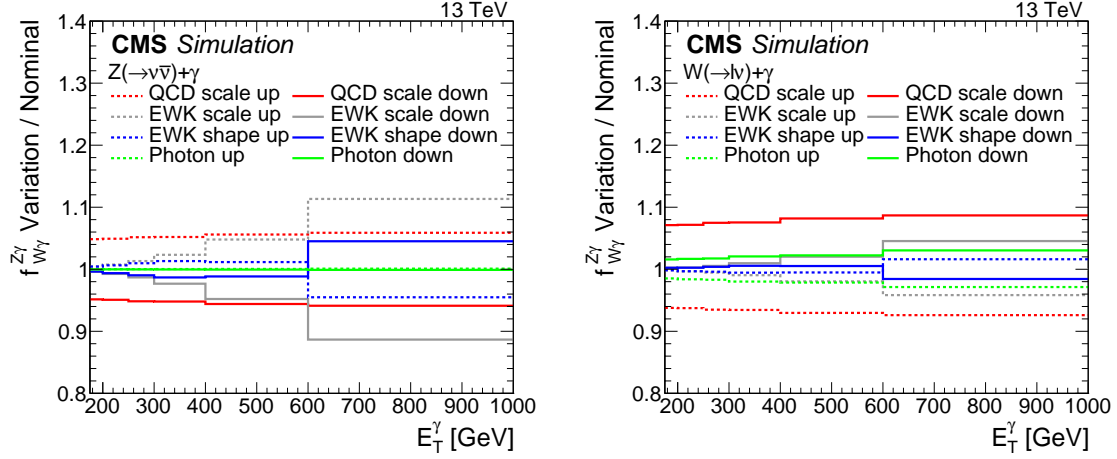


Figure 1-9: Systematic uncertainty in the transfer factors for $Z(\rightarrow \nu\bar{\nu})+\gamma$ (left) and $W(\rightarrow \ell\nu)+\gamma$ (right). The last bin includes all events with $E_T^\gamma > 1000$ GeV.

dent of E_T^γ . Figure 1-10 shows the transfer factor $f_{W\gamma}^{Z\gamma}$ between the $Z(\rightarrow \nu\bar{\nu})+\gamma$ and $W(\rightarrow \ell\nu)+\gamma$ 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 control region in the i^{th} bin of the E_T^γ distribution can be expressed as

$$T_{\ell\gamma,i} = \frac{N_i^{Z\gamma}}{R_{\ell\gamma,i}^{W\gamma} f_{W\gamma,i}^{Z\gamma}} + b_{\ell\gamma,i}, \quad (1.3)$$

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.

1.4 Misidentified electrons

An electron can be misidentified as a photon if the association of tracks or track seeds to the supercluster in ECAL 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_T^{miss} signature if the electron is misidentified.

The rate at which this misidentification occurs is proportional to the inefficiency $1 - \epsilon_e^{\text{track}}$ of the tracking, defined over the electrons passing the photon identification

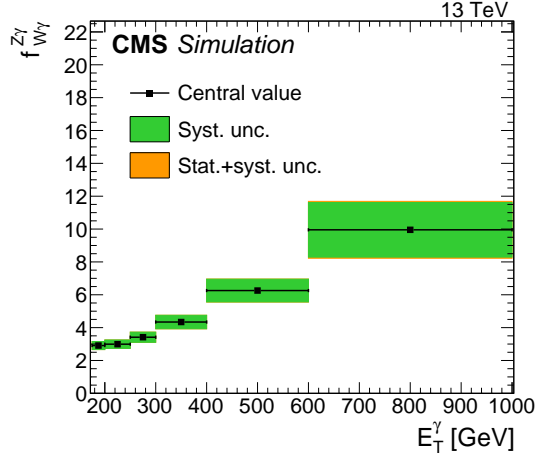


Figure 1-10: Transfer factor $f_{W\gamma}^{Z\gamma}$. The uncertainty bands in green (inner) and orange (outer) show the systematic uncertainty, and the combination of systematic and statistical uncertainty arising from limited MC sample size, respectively. The systematic uncertainties considered are the uncertainties from higher-order theoretical corrections.

criteria described in Sec. ?? except the electron veto. This partial identification is denoted as $e\gamma$ ID in the following. If one assumes that the kinematic and other critical properties of the electron plus E_T^{miss} events are unaffected by the electron misidentification, it is possible to model the electron misidentification background by taking a proxy sample with well-identified electrons and scaling this sample by $R_e = (1 - \epsilon_e^{\text{track}})/\epsilon_e^{\text{track}}$.

The “tag-and-probe” method described in Section ?? with appropriate changes is used to measure the efficiency corresponding to the factor R_e in data.

The first such change is that the sample is split into $e\gamma$ and ee categories depending on whether the probe passes or fails the electron veto requirement. Probes in both categories must also pass the $e\gamma$ ID. Denoting the area of the peak in each category $N_{e\gamma}$ and N_{ee} , respectively, the ratio $N_{e\gamma}/N_{ee}$ is equal to R_e up to minor systematic corrections.

The second such change is in the background model used in the TP fits. The backgrounds to the $e\gamma$ fit consist of processes with actual electron and photon in the final state, such as $W\gamma$ and $Z \rightarrow ee$ with a hard radiation off one of the electrons. Because of this, we scale the mass distribution of the $\mu + \gamma$ sample by the ratio of

electron-probe to muon-probe events taken from MC to account for the different rates of FSR and bremsstrahlung between muons and electrons. 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 1-11 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 background is derived.

Figure 1-12 shows the derived R_e factor as a function of E_T^γ . The electron proxy sample is reweighted by R_e depending on the p_T of the electron object.

1.5 Misidentified hadrons

The estimation of hadron misidentification background proceeds in multiple steps. First, we measure the fraction of hadronic objects within a pool of photon candidate objects in the EM object+jet measurement sample. This measurement is described in detail in Section ?? . Figure 1-13 and Table 1.5 show the final 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 1.5: 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 . Figure 1-14 shows the transfer factor R_h along with the various distributions used for its derivation.

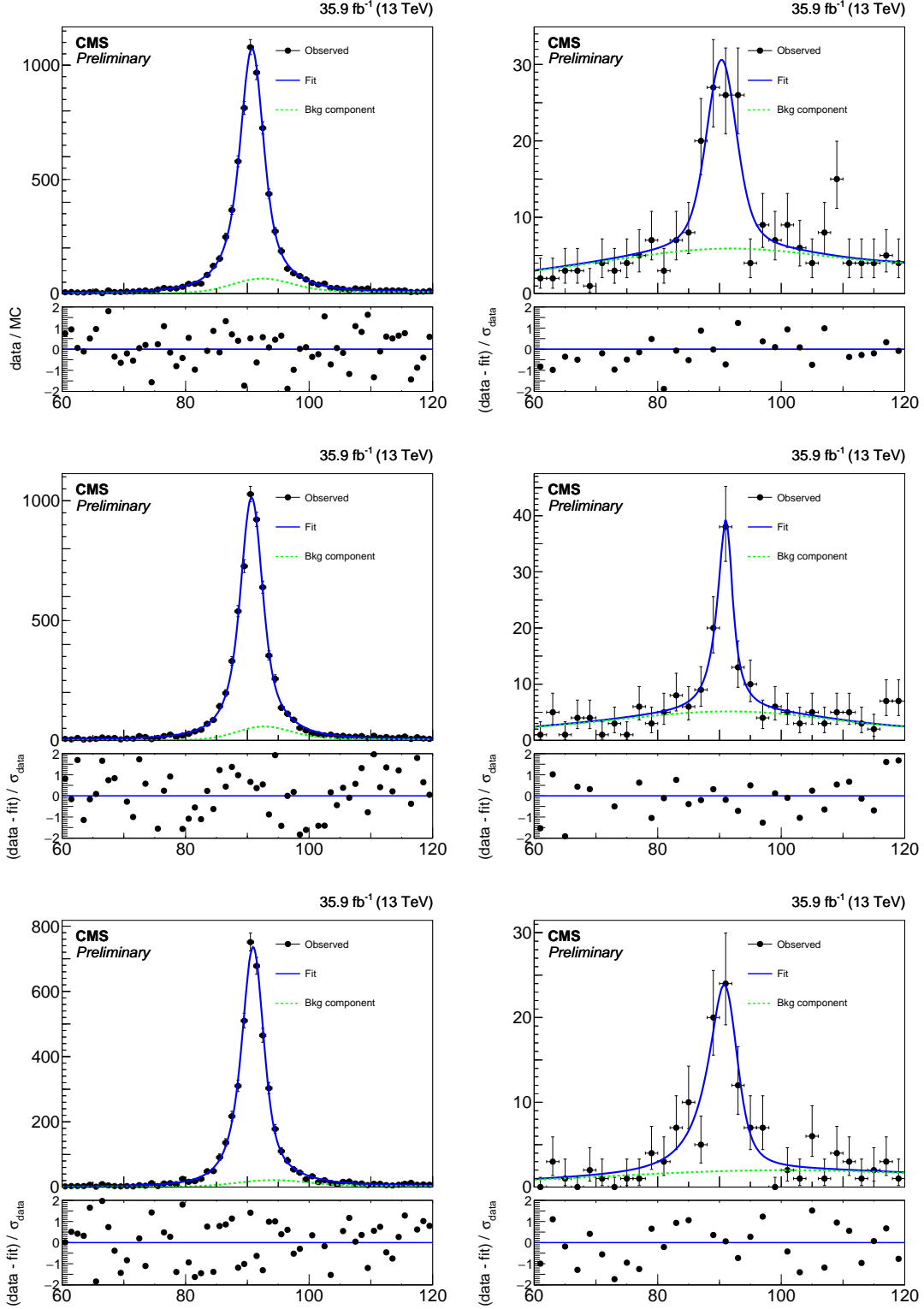


Figure 1-11: Fits to the mass distributions for ee (left) and $e\gamma$ (right) selections, in bins of probe p_T : $175 < p_T < 200$ GeV (top), $200 < p_T < 250$ GeV (middle), $p_T > 250$ GeV (bottom). The blue solid line represents the full fit model, and the green dashed line its background component.

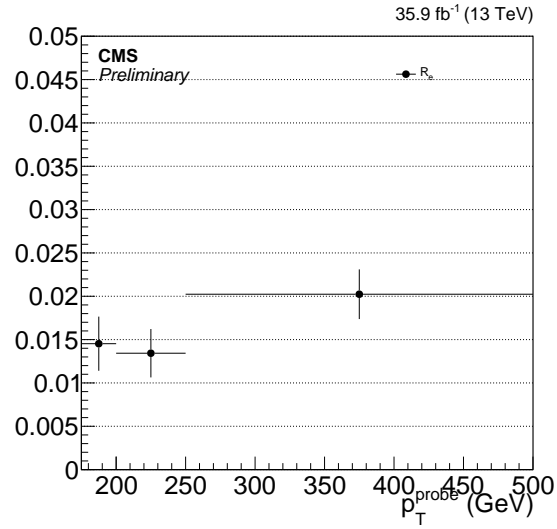


Figure 1-12: Electron to photon fake rate R_e .

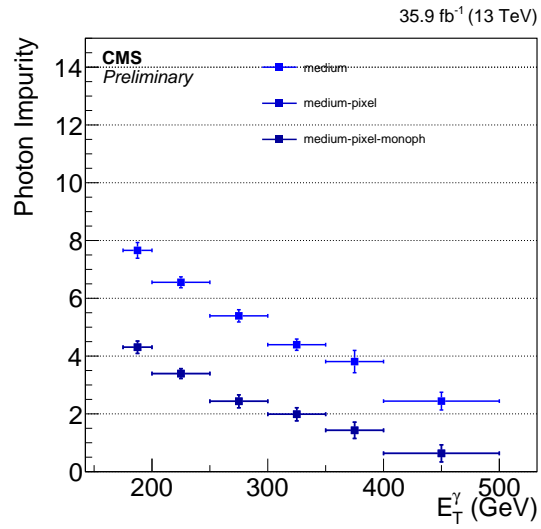


Figure 1-13: Impurities for photons as a function of p_T . The different bands show the effects of adding different stages of the full ID, starting with the baseline ID and isolation and successively adding the pixel seed veto.

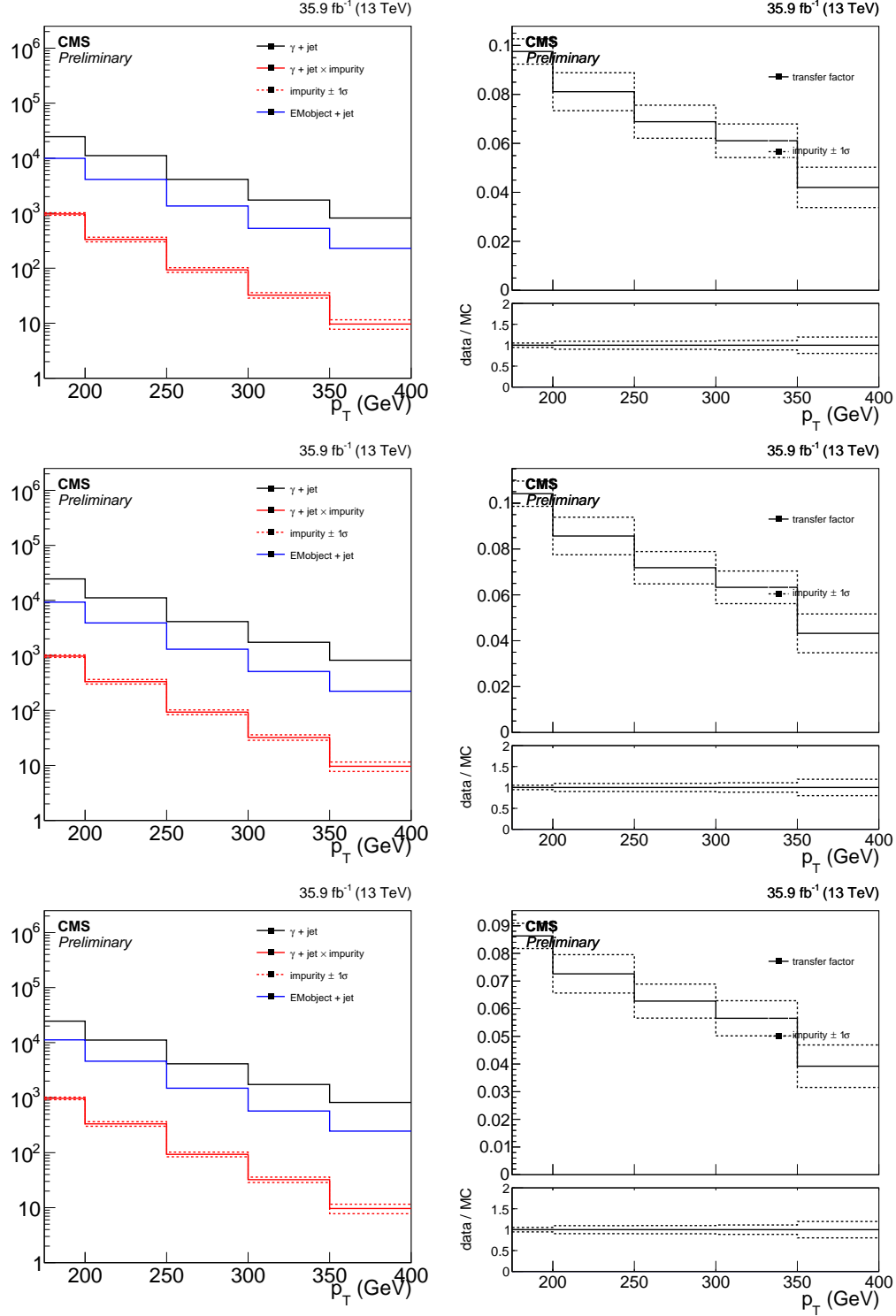


Figure 1-14: Left: The p_T distribution of the candidate photon object in the photon + jet control sample (black), the result of scaling it with the impurity (red), and the p_T distribution of the hadronic proxy object in the proxy + jet control sample (blue). Right: Hadronic transfer factor R_h , which is the ratio of the red and blue distributions in the left plot. Top: Nominal hadron proxy object. Middle: Tighter hadron proxy object. Bottom: Looser hadron proxy object.

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 arrive at an estimate of the misidentified hadron plus E_T^{miss} background of this analysis.

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 different distributions from the nominal, tight, and loose selections are shown in Figure 1-15. 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. Figure 1-16 shows the resulting shapes from moving the shapes generated by a one sigma shift in the purity.

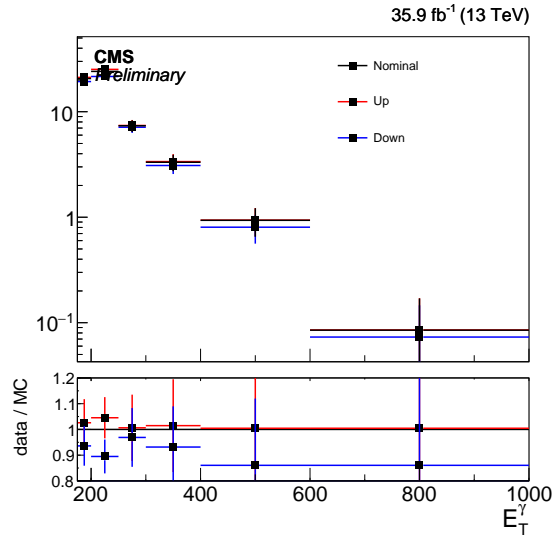


Figure 1-15: The p_T distribution of the estimated contribution from hadronic fakes in the signal region. The distribution labeled Up (Down) comes from the tighter (looser) selection. The systematic uncertainty resulting from this variation is around 5% at the low end of our p_T range and increases to 15% after $p_T > 400$ GeV.

1.6 Beam halo

Based on the beam halo features discussed in Section ??, a two-template fit to the ϕ' distribution of the photons in the candidate sample, where the templates are that of the halo shower and a uniform distribution, accurately estimates the amount of beam

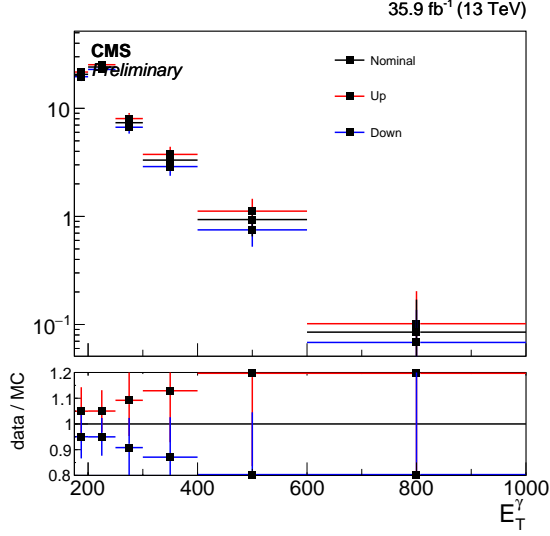


Figure 1-16: The p_T distribution of the estimated contribution from hadronic fakes in the signal region. The distribution labeled Up (Down) comes from varying the purity one sigma up (down). The systematic uncertainty resulting from this variation is around 5% at the low end of the p_T range and increases to 20% after $p_T > 400$ GeV.

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 $C_V = (\pi - 1)/\pi$, respectively. The corresponding fractions for beam halo events are determined by selecting a halo-enriched sample where the halo identification 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 i^{th} bin of the signal region $K \in \{H, V\}$. Using the notation introduced in Section 1.3, the total estimated background T_K in the two signal regions are

$$\begin{aligned} T_{K,i} &= C_K(N_i^{Z\gamma} + N_i^{W\gamma}) + hn_{K,i}^{\text{halo}} + C_K b_{K,i} \\ &= C_K(1 + f_{W\gamma i}^{Z\gamma - 1})N_i^{Z\gamma} + hn_{K,i}^{\text{halo}} + C_K b_{K,i}, \end{aligned} \quad (1.4)$$

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.

1.7 Spikes

Given the observations in Section ??, the time distribution of spike-like rec hits outside of the window $-15 < t < -10$ 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 Ref. [?]. 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 $-3 < t < 3$ ns.

Having established that there is no special population of ECAL spikes in the in-time region, we can estimate the number of ECAL spike events present in the signal candidate sample by an “ABCD” method, where

- A = Number of clusters with $\sigma_{i\eta i\eta}$ or $\sigma_{i\phi i\phi}$ less than 0.001 and seed time $-15 < t < -10$ 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 $-15 < t < -10$ 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,

and D is the estimated number of spike events in the signal region, obtained by

$$D = C \times \frac{B}{A}. \quad (1.5)$$

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

$$A = 4969$$

$$B = 1180$$

$$C = 54$$

$$\therefore D = 12.8 \pm 1.8(\text{stat.})$$

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 happens to be narrow, as observed in Fig. ???. 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 $-15 < t < -10$ 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 as the systematic uncertainty in the spike background estimate.

1.8 Minor SM Backgrounds

After the full selection described in Section 1.2, the SM γ +jets, $t\bar{t}\gamma$, $VV\gamma$, $Z(\rightarrow \ell\bar{\ell})+\gamma$, and $W \rightarrow \ell\nu$ processes are minor ($\sim 10\%$) 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 $E_{\text{T}}^{\text{miss}}$ in the signal region. However, the $E_{\text{T}}^{\text{miss}}$ is typically aligned with the photon or one of the jets in such cases, and therefore various selections on the kinematic relations between the $E_{\text{T}}^{\text{miss}}$, 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 1.6 and 1.7.

1.9 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. Uncertainties in various quantities are represented by nuisance parameters in the fit. Predictions for $Z(\rightarrow \nu\bar{\nu})+\gamma$, $W(\rightarrow \ell\nu)+\gamma$, 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 $Z(\rightarrow \nu\bar{\nu})+\gamma$ 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 $W(\rightarrow \ell\nu)+\gamma$ and $Z(\rightarrow \ell\bar{\ell})+\gamma$ samples in all regions are related to the yield of $Z(\rightarrow \nu\bar{\nu})+\gamma$ through the MC prediction through the transfer factors defined in Section 1.3. 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

$$\begin{aligned}
\mathcal{L} &= \prod_i \{ \mathcal{L}_{\text{signal}} \times \mathcal{L}_{\text{single-lepton}} \times \mathcal{L}_{\text{dilepton}} \} \times \mathcal{L}_{\text{nuisances}} \\
&= \prod_i \left\{ \prod_{K=H,V} \mathcal{P} \left(d_{K,i} \middle| T_{K,i}(\vec{\theta}) \right) \times \prod_{\ell=e,\mu} \mathcal{P} \left(d_{\ell\gamma,i} \middle| T_{\ell\gamma,i}(\vec{\theta}) \right) \times \prod_{\ell=e,\mu} \mathcal{P} \left(d_{\ell\ell\gamma,i} \middle| T_{\ell\ell\gamma,i}(\vec{\theta}) \right) \right\} \times \prod_j \mathcal{N}(\theta_j) \\
&= \prod_i \left\{ \prod_{K=H,V} \mathcal{P} \left(d_{K,i} \middle| \left(1 + f_{W\gamma,i}^{Z\gamma} \right)^{-1}(\vec{\theta}) C_K N_i^{Z\gamma} + h n_{K,i}^{\text{halo}}(\vec{\theta}) + C_K b_{K,i}(\vec{\theta}) \right) \right. \\
&\quad \times \prod_{\ell=e,\mu} \mathcal{P} \left(d_{\ell\gamma,i} \middle| \frac{N_i^{Z\gamma}}{R_{\ell\gamma,i}^{W\gamma}(\vec{\theta}) f_{W\gamma,i}^{Z\gamma}(\vec{\theta})} + b_{\ell\gamma,i}(\vec{\theta}) \right) \\
&\quad \times \prod_{\ell=e,\mu} \mathcal{P} \left(d_{\ell\ell\gamma,i} \middle| \frac{N_i^{Z\gamma}}{R_{\ell\ell\gamma,i}^{Z\gamma}(\vec{\theta})} + b_{\ell\ell\gamma,i}(\vec{\theta}) \right) \left. \right\} \times \prod_j \mathcal{N}(\theta_j), \tag{1.6}
\end{aligned}$$

following the notation introduced in Section 1.3, and where $\mathcal{P}(n|\lambda)$ is the Poisson probability of n for mean λ , \mathcal{N} denotes the unit normal distribution, and $d_{X,i}$ is the observed number of events in bin i of region X . Systematic uncertainties are treated as nuisance parameters in the fit and are represented by $\vec{\theta}$. Each quantity Q_j with a nominal value \bar{Q}_j and a standard deviation of the systematic uncertainty σ_j appears in the likelihood function as $\bar{Q}_j \exp(\sigma_j \theta_j)$.

1.10 Results

1.10.1 Pre-fit and post-fit distributions

Figure 1-17 shows the observed E_T^γ distributions in the four control regions compared with the results from simulations before and after performing the simultaneous fit across all the control samples and signal region, and assuming absence of any signal. Figure 1-18 shows the observed E_T^γ distributions in the horizontal and vertical signal regions compared with the results from simulations before and after performing a combined fit to the data in all the control samples and the signal region. The observed distributions are in agreement with the prediction from SM and noncollision backgrounds.

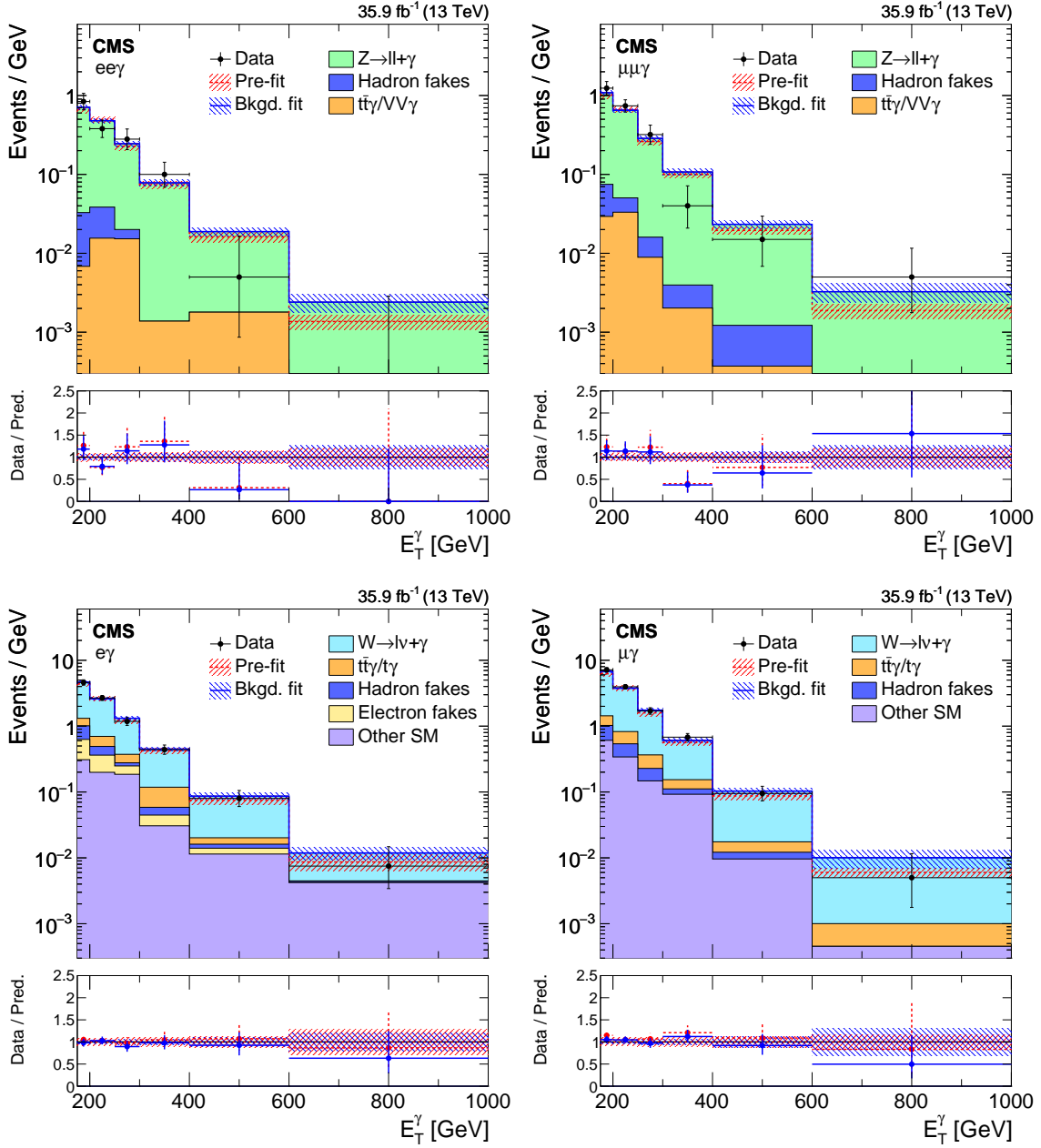


Figure 1-17: Comparison between data and MC simulation in the four control regions: $ee\gamma$ (upper left), $\mu\mu\gamma$ (upper right), $e\gamma$ (lower left), $\mu\gamma$ (lower right) before and after performing the simultaneous fit across all the control samples and signal region, and assuming absence of any signal. The last bin of the distribution includes all events with $E_T^\gamma > 1000$ 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.

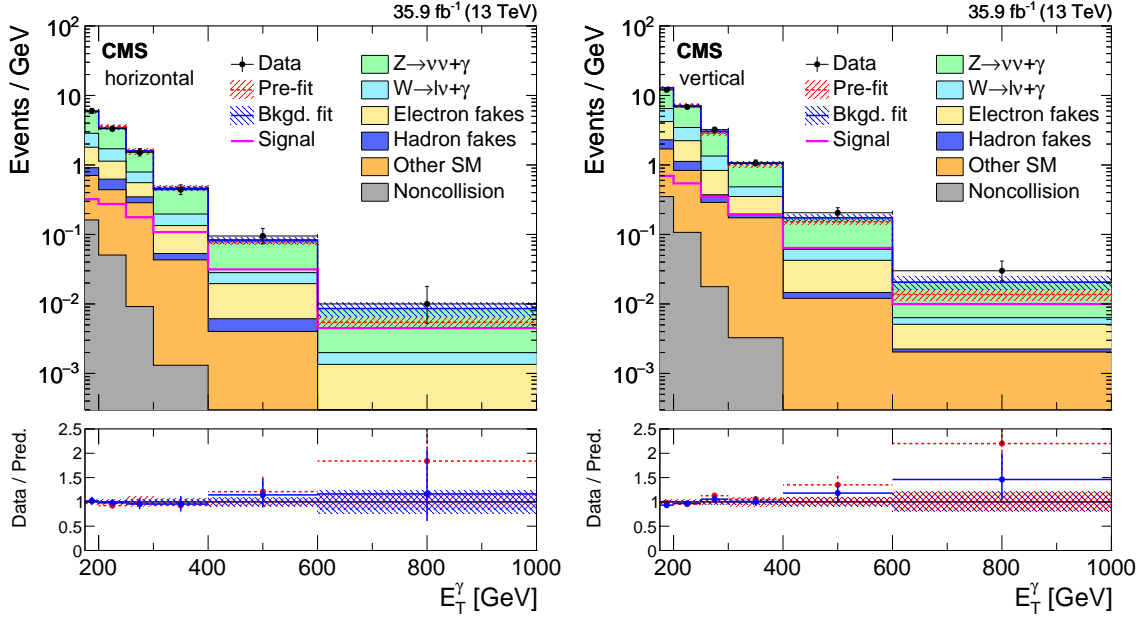


Figure 1-18: Observed E_T^γ distributions in the horizontal (left) and vertical (right) signal regions compared with the post-fit background expectations for various SM processes. The last bin of the distribution includes all events with $E_T^\gamma > 1000$ GeV. The expected background distributions 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 TeV vector mediator decaying to 1 GeV DM particles is overlaid.

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 1.6: Expected event yields in each E_T^γ bin for 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.

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 1.7: Expected event yields in each E_T^γ bin for 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.

The expected yields in each bin of E_T^γ for all backgrounds in the horizontal and vertical signal regions after performing a combined fit to data in all the control samples, excluding data in the signal regions, are given in Tables 1.6 and 1.7, respectively. The covariances between the predicted background yields across all the E_T^γ bins in the two signal regions are shown in Fig. 1-19. The expected yields together with the covariances can be used with the simplified likelihood approach detailed in Ref. [?] to reinterpret the results for models not studied in this thesis

1.10.2 Limits

No significant excess of events beyond the SM expectation is observed. 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 [?, ?, ?], using a test statistic based on the negative logarithm of the likelihood in Section 1.9.

Figure 1-20 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 $M_{\text{med}}-m_{\text{DM}}$ plane. The solid black (dashed red) curves are the observed (expected) contours of $\mu_{95} = 1$. The σ_{theory} hypothesis is

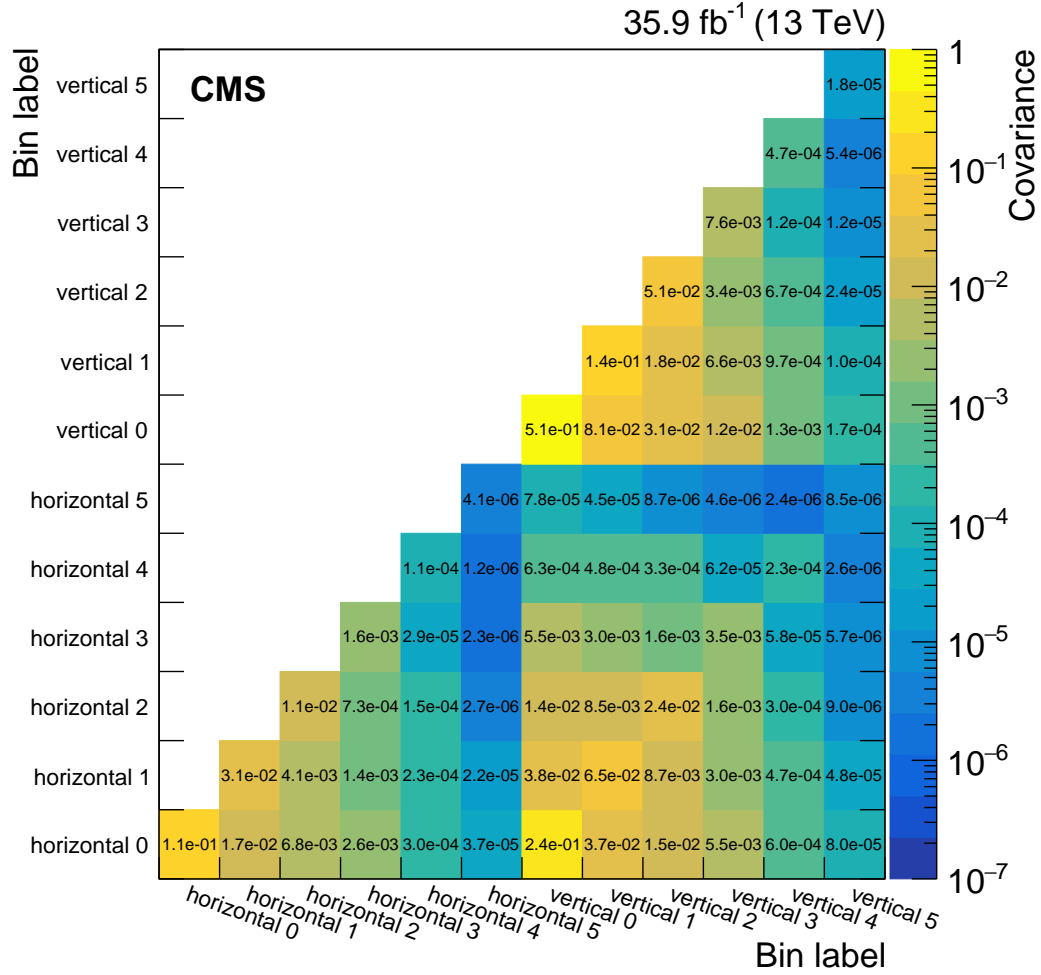


Figure 1-19: Covariances between the predicted background yields in all the E_T^γ bins of the horizontal and vertical signal regions. The bin labels specify which signal region the bin belongs to and what number bin it is for that region.

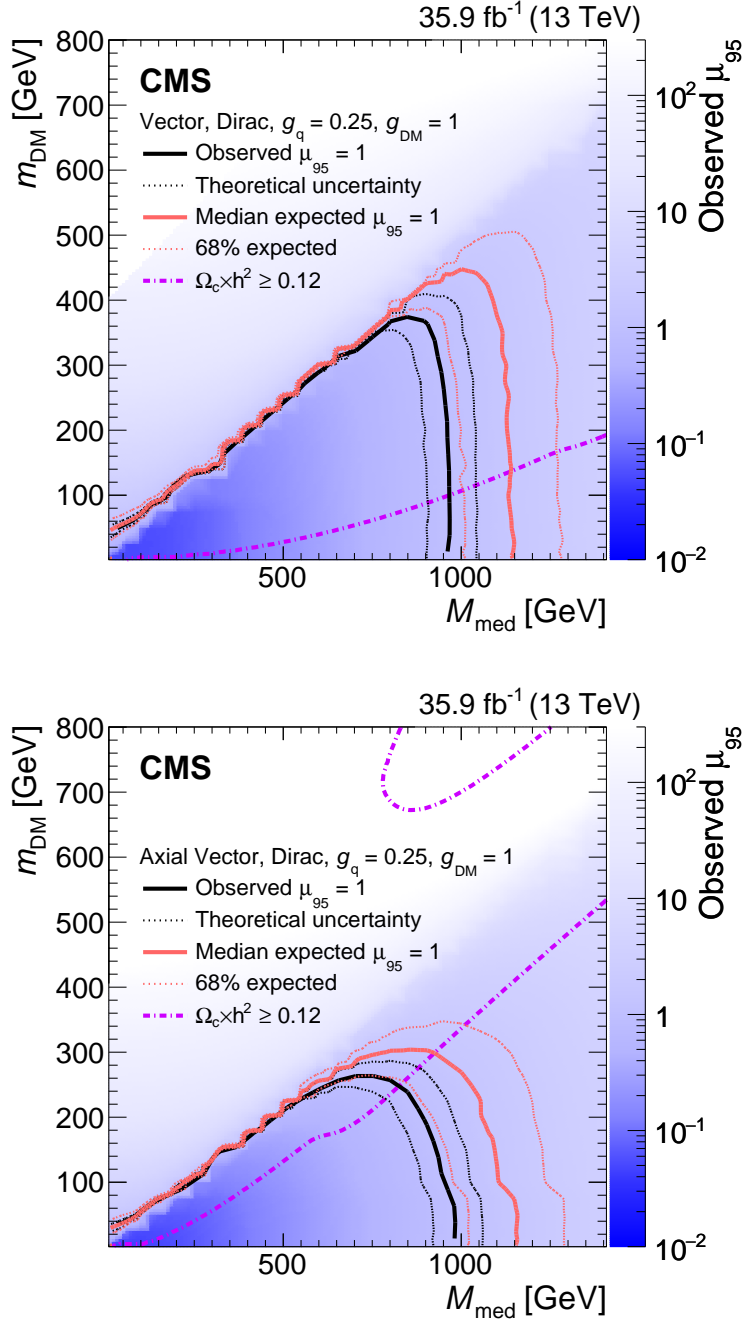


Figure 1-20: The ratio of 95% CL upper cross section limits to the theoretical cross section (μ_{95}), for DM simplified models with vector (top) and axial-vector (bottom) mediators, assuming $g_q = 0.25$ and $g_{DM} = 1$. Expected $\mu_{95} = 1$ contours are overlaid in red. The region under the observed 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.

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 GeV are excluded for values of m_{DM} less than 1 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 [?]. The expected DM abundance is estimated, separately for each model, using the thermal freeze-out mechanism implemented in the MADDM [?] framework and compared to the observed cold DM density $\Omega_c h^2 = 0.12$ [?], where Ω_c is the DM relic abundance and h is the dimensionless Hubble constant.

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 Ref. [?] and [?]. The exclusion contours for the vector and axial-vector models shown in Figure 1-20 are translated into the $\sigma_{\text{SI}}-m_{\text{DM}}$ and $\sigma_{\text{SD}}-m_{\text{DM}}$ planes shown in Figure 1-21. When compared to recent results by the CDMSLite [?], LUX [?], PandaX-II [?], XENON1T [?], and CRESST-II [?] 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 results by the PICO-60 [?], IceCube [?], PICASSO [?] and Super-Kamiokande [?] collaborations, the limits obtained from this search provide stronger constraints for DM masses less than 200 GeV for spin dependent models.

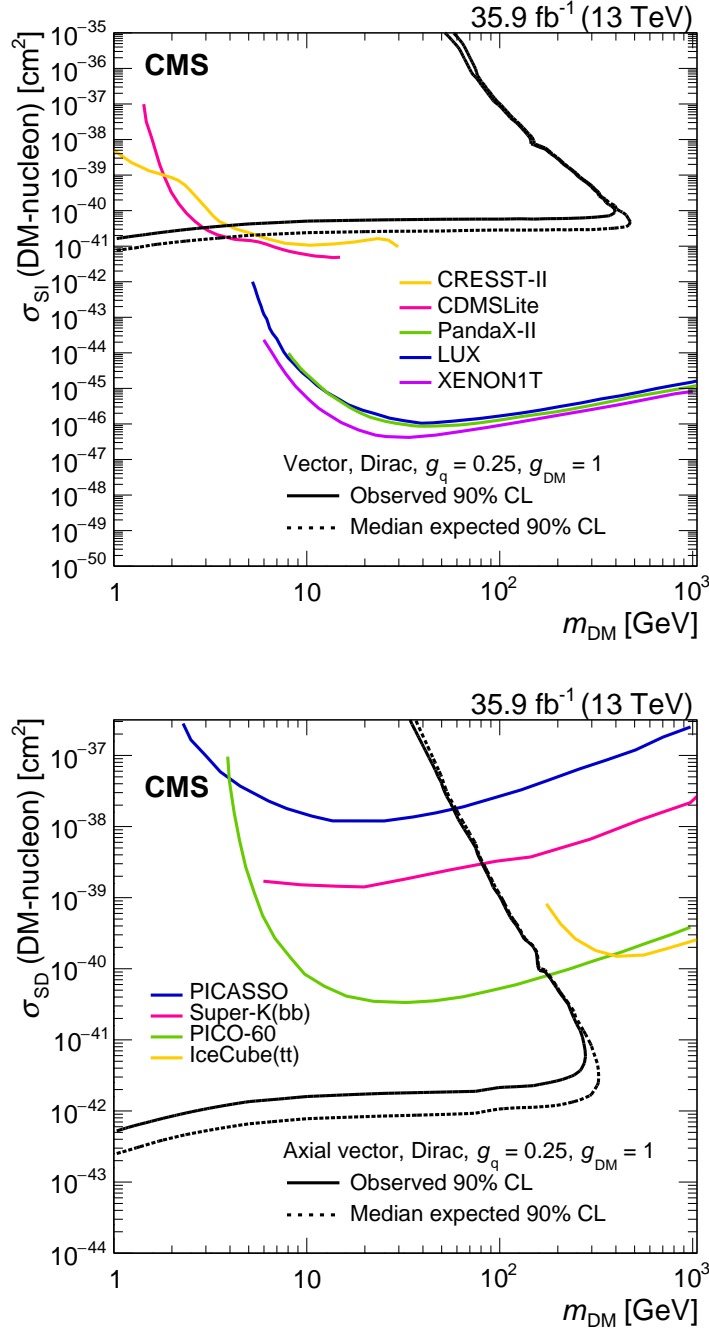


Figure 1-21: The 90% CL exclusion limits on the χ -nucleon spin-independent scattering cross sections involving the vector operator (top) and the χ -nucleon spin-dependent scattering cross sections involving the axial-vector operator (bottom) as a function of the m_{DM} . Simplified model DM parameters of $g_q = 0.25$ and $g_{\text{DM}} = 1$ 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.