Estimation of $ZZ \to ll\nu\nu$ background using $Z(\to ll) + \gamma$ data

Mangesh Sonawane
Supervisor: Dr. Beate Heinemann

June 2017 - May 2018

Contents

Abstract

In the search for Dark Matter at the LHC, events with large imbalance in transverse momentum are of interest. One such signature is $ll + E_T^{miss}$. *The dominant background contributing to the* $ll + E_T^{miss}$ *is* $ZZ \to ll\nu\nu$ *(≈60%). Currently, this background is determined using Monte Carlo simulation, with an uncertainty of* $\approx 10\%$. *The goal of this study is to establish a data driven method to estimate this background, and refine the uncertainty. However, the small branching ratio of* Z *decaying leptonically limits the precision to which we can estimate this directly from* ZZ *data using* $Z(\to ll) + \gamma$, *which is a pure signal and has a high* $BR * \sigma$. *In regions where* $p_T(\gamma) \gg M_Z$, *the two processes are kinematically similar. Defining a variable* R *as a function of transverse momentum:*

$$R(p_T) = \frac{\sigma_{ZZ}(p_T)}{\sigma_{Z\gamma}(p_T)}$$

we can use Monte Carlo to estimate the uncertainty on $R(p_T)$, *and use* R *with* $Z\gamma$ *data to obtain the contribution of* $ZZ \to ll\nu\nu$ *background.*

1 Introduction

Among the candidates for Dark Matter at the LHC are WIMPs (Weakly Interacting Massive Particles). The signature for WIMPs are events with large E_T^{miss}. One such signal we look at is $ll + E_T^{miss}$. For example, the production of Higgs in association with a Z, as shown in Fig.1, is one possible process giving the $ll + E_T^{miss}$ signature:

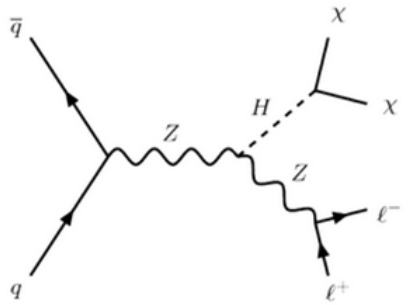


Figure 1: Feynman diagram showing the associated production of Higgs

WIMPs do not register in the detector, and thus result in a large missing transverse momentum (MET or E_T^{miss}).

Other processes that contribute to this signature are $ZZ \to ll\nu\nu$, $WZ \to lll\nu$,$WW \to l\nu l\nu$, Z+jets and W+jets. These processes contribute to the background. The dominant source of background is the $ZZ \to ll\nu\nu$ process, contributing $\approx 60\%$ of the background. Thus is it important to determine

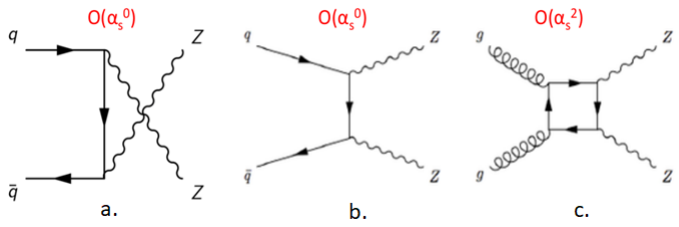


Figure 2: Feynman Diagram showing ZZ Production
a. & b. $q\bar{q} \to ZZ$ c. $gg \to ZZ$

this contribution to the background, along with the uncertainty associated with it. Currently, this is determined using Monte Carlo, with an uncertainty of $\approx 10\%$ [2].

The branching fraction of Z to any one flavor of lepton is $\approx 3.4\%$, and to neutrinos is $\approx 20\%$.

$$\begin{aligned} BR(Z \to ee \text{ or } \mu\mu) &\approx 6.8\% \\ BR(Z \to \nu\nu) &\approx 20\% \\ \text{Thus,} & \\ BR(ZZ \to ll\nu\nu) &= 2 * BR(Z \to ee \text{ or } \mu\mu) * BR(Z \to \nu\nu) \\ &= 2 * (0.068) * (0.2) \approx 3\% \end{aligned}$$

One method of estimating this contribution is to look at $ZZ \to llll$. However, this branching fraction is even lower, at $\approx 0.46\%$.

In similar vein to a earlier analysis that used γ+jets to calibrate Z+jets background [1], in the $p_T(\gamma) \gg M_Z$ region, the $Z\gamma \to ll\gamma$ process should be kinematically similar to $ZZ \to ll\nu\nu$ as the mass of the Z boson is negligible. Figures 2 and 3 show the leading order Feynman diagrams for the production of ZZ and $Z + \gamma$ respectively. The diagrams for $q\bar{q}$ and gg (a. b. and c.)are similar. In addition to having a higher $BR * \sigma$ as compared to $ZZ \to ll\nu\nu$, the $Z\gamma \to ll\gamma$ signal is also very pure. Thus, it should be possible to use $Z\gamma \to ll\gamma$ data to estimate $ZZ \to ll\nu\nu$ contribution to the background, and obtain a more accurate prediction.

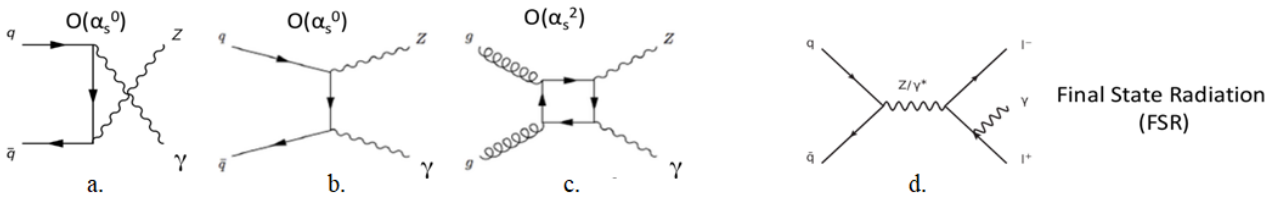


Figure 3: Feynman Diagram showing $Z + \gamma$ Production
a. & b. $q\bar{q} \to Z + \gamma$ c. $gg \to Z + \gamma$ d. Final State Radiation (FSR)

2 Approach

Following the method defined in the Ref [1], we define a variable $R(p_T)$ to be the ratio of the cross sections of $ZZ \to ll\nu\nu$ to $Z\gamma \to ll\gamma$ as a function of p_T.

$$R(p_T) = \frac{\sigma_{ZZ}(p_T)}{\sigma_{Z\gamma}(p_T)} \tag{1}$$

With the two processes being kinematically similar at high p_T, R depends on the coupling of the Z and γ to quarks. It should approach some value asymptotically.

The photon - quark and Z boson - quark couplings in the Standard Model are given by,

$$-ieQ_q\gamma^\mu \qquad \text{and} \qquad \frac{-ie}{2\sin\theta_W\cos\theta_W}\gamma^\mu(v_q - a_q\gamma_5) \tag{2}$$

respectively, where Q_q, v_q and a_q are respectively the electric, vector and axial neutral weak couplings of the quarks, and θ_W is the weak mixing angle. The cross sections are dependent on the matrix elements squared, which contain factors of Q_q^2 for γ, or $(v_q^2 + a_q^2)/4\sin^2\theta_W\cos^2\theta_W$ for Z. There is a contribution due to the Z mass which appears in the internal propagators and phase space integration. This contribution becomes less important in the $p_T(\gamma) \gg M_Z$ region.

Thus, in the high p_T region, the Z and γ cross sections would be in the ratio

$$R_q = \frac{v_q^2 + a_q^2}{4\sin^2\theta_W\cos^2\theta_W * Q_q^2}. \tag{3}$$

Considering the contributions from both u and d flavor quarks,

$$R = \frac{Z_u \langle u \rangle + Z_d \langle d \rangle}{\gamma_u \langle u \rangle + \gamma_d \langle d \rangle} \tag{4}$$

Substituting $\sin^2 \theta_W = 0.2315$, at moderate p_T values, $R \approx 1.4$[1].

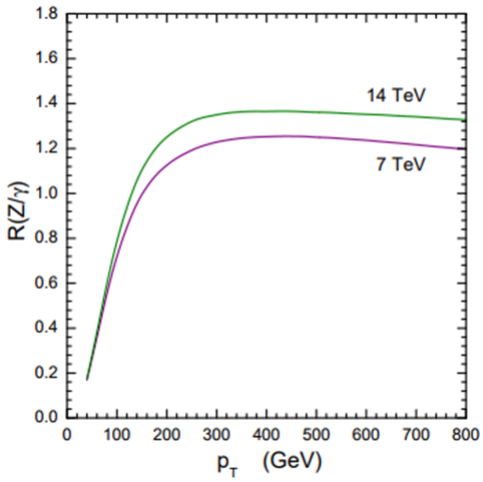


Figure 4: Ratio of the Z and γ p_T distributions [1]

This ratio R can be used as is for $ZZ \to ll\nu\nu$ and $Z\gamma \to ll\gamma$, as the contribution from the $Z \to ll$ is identically multiplied into the numerator as well as the denominator, and thus cancels out.

A Monte Carlo generator, MCFM v8.0 [3] at NLO, in this case, is used to generate cross sections of $ZZ \to ll\nu\nu$ and $Z\gamma \to ll\gamma$ processes, with a selection of generator level cuts. The samples are generated with cuts on $E^{miss}_{T,min}$ for the ZZ process $p_{T,min}(\gamma)$ for the $Z + \gamma$ process. A ratio of these cross sections is taken to obtain the R curve as a function of p_T. The uncertainty on R is calculated by varying several parameters at the generator level, such as the renormalization and factorization scales, the PDF sets used, photon fragmentation, etc. Effects of applying lepton cuts on the cross sections as well as the ratio, and the contributions of the $q\bar{q}$ and gg processes are also studied.

However, the MCFM generator only produces $Z \to ee$ instead of $Z \to ll$. Thus, this branching ratio needs to be accounted for to obtain the value of R.

$$R_{inc} = R * \frac{BR(Z \to ee)}{BR(Z \to ee) * BR(Z \to \nu\nu) * 2} \tag{5}$$

3 Generator Parameters

The samples are generated using MCFM v8.0 for the following data points[2]

$$\text{For } ZZ \to ee\nu\nu : E_T^{miss} > \{50, 75, 100, 125, 150, 200, 250, 300, 400, 500\} \text{ GeV}$$
$$\text{For } Z(\to ee) + \gamma : p_T(\gamma) > \{50, 75, 100, 125, 150, 200, 250, 300, 400, 500\} \text{ GeV}$$

The following generator level cuts are used for the first run of ZZ and $Z + \gamma$ processes[3]

[1] Equations (3) and (4), as well as the value of R are taken from Ref [1]

[2] MCFM does not generate $Z \to ll$ but $Z \to ee$. As electrons and muons have similar properties with the exception of mass, simply the branching fraction of $Z \to ee$ must be accounted for at a later stage.

[3] All lepton cuts are consistent with the ones used in the ATLAS Z+MET analysis

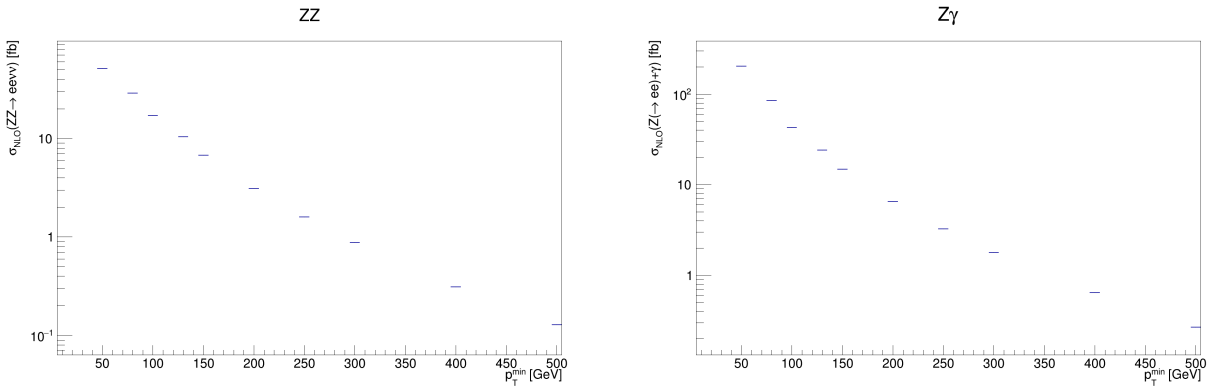
Cuts	$ZZ \to ee\nu\nu$	$Z(\to ee) + \gamma$
Process ID	87	300
M_{ee}	$81 < M_{ee} < 101$ GeV	$81 < M_{ee} < 101$ GeV
$M_{\nu\nu}$	$81 < M_{\nu\nu} < 101$ GeV	-
Order	NLO	NLO
PDFset	CT14.NN	CT14.NN
$p_T^{\text{lead}}(e)$	> 30 GeV	> 30 GeV
$\eta^{lead}(e)$	< 2.5	< 2.5
$p_T^{\text{sublead}}(e)$	> 20 GeV	> 20 GeV
$\eta^{sublead}(e)$	< 2.5	< 2.5
$\Delta R(\gamma, e)$	-	0.7
Renormalization scale	91.187 GeV	91.187 GeV (M_Z)
Factorization scale	91.187 GeV	91.187 GeV (M_Z)

Table 1: Parameters in input.DAT for MCFM

The constraint on M_{ee} in the case of $Z + \gamma$ suppresses the FSR process by ensuring that the lepton pair are from a Z decay only.

4 Results

Upon running the steering file with the parameters described above, the cross sections shown in Figure 5 are obtained. Throughout this analysis, this sample is the reference.



(a) $ZZ \to ee$ cross section

(b) $Z(\to ee) + \gamma$ cross section

Figure 5: Cross sections of ZZ and $Z + \gamma$ processes with the cuts as in Table 1. The Y axis is in $\log_{10}$ scale.

The resulting ratio is shown in Figure 6

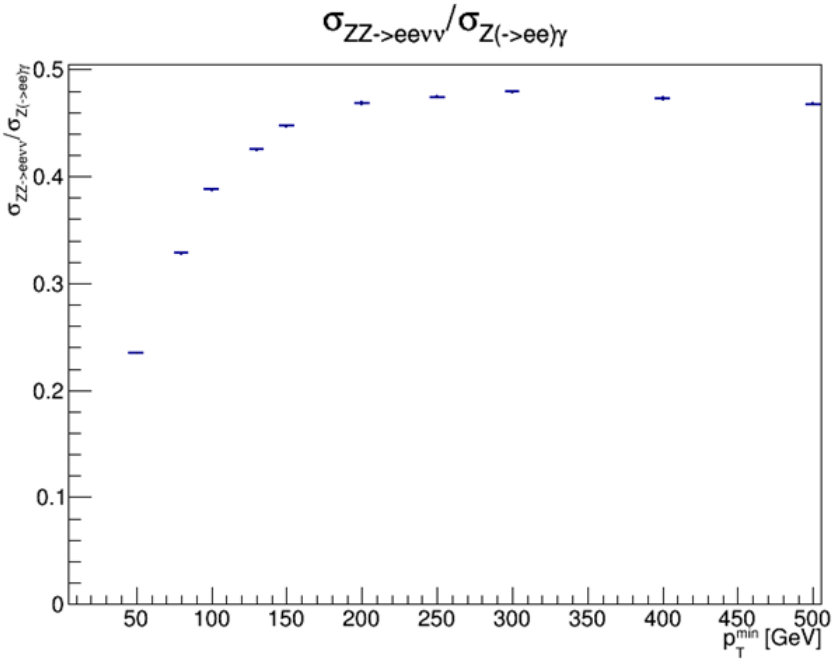


Figure 6: R curve as a function of p_T

The R value is observed to increase from ≈ 0.24 at 50 GeV to ≈ 0.47 at high p_T, where it is constant. When the branching ratio is accounted for as show in Equation 5, the resulting $R(p_T)$ curve is shown in Figure 7, increasing from ≈ 0.61 at 50 GeV to ≈ 1.2 at high p_T.

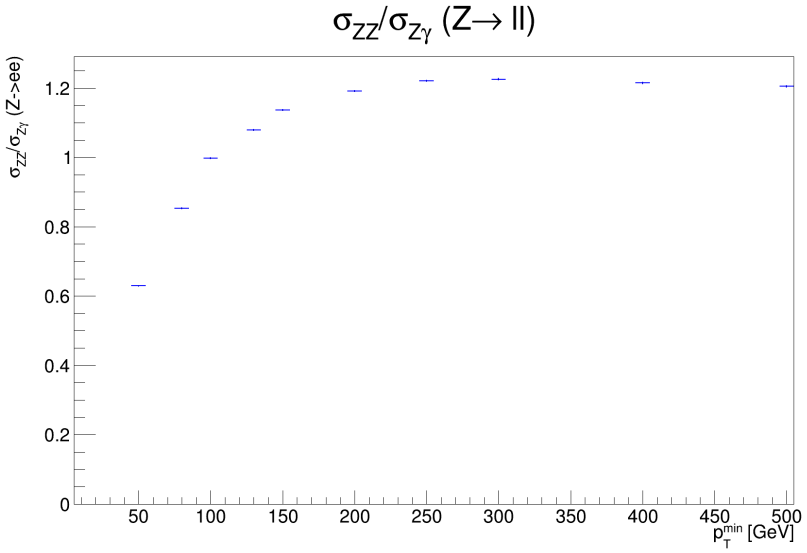


Figure 7: R curve as a function of p_T, accounting for the $Z \to ee$ and $Z \to \nu\nu$ branching ratios.

4.1 Lepton Cuts

To check the effects of lepton cuts on the ratio, samples with similar parameters as those in Table 1 are generated. However, we relax the cuts on leptons. Both the leading and subleading lepton should have $p_T > 5$ GeV, and $\eta < 10$. In the lower p_T regions, the cross section falls by nearly half in both processes. However, the ratio is not affected very much, as seen in Figure 8.

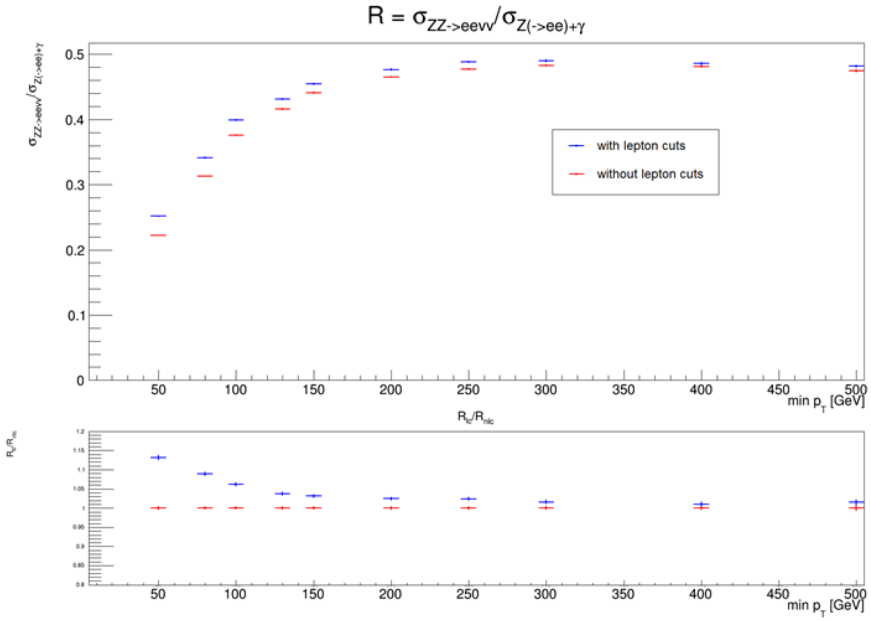


Figure 8: Comparison of reference R curve to R curve without lepton cuts

The R curves differ by $\approx 4\%$ at high p_T, and $\approx 7\%$ at 100 GeV.

4.2 Scale Variation

The Renormalization and Factorization scales are arbitrary parameters that address the UV and IR divergences respectively that arise while calculating cross sections. They are important when considering higher order effects in QCD. To obtain the uncertainties associated to these scales, the Renormalization (μ_R) and Factorization (μ_F) scales are each varied by a factor of 2 in either direction from the central value, $M_Z = 91.187$ GeV, to obtain the uncertainty. Figure 9.

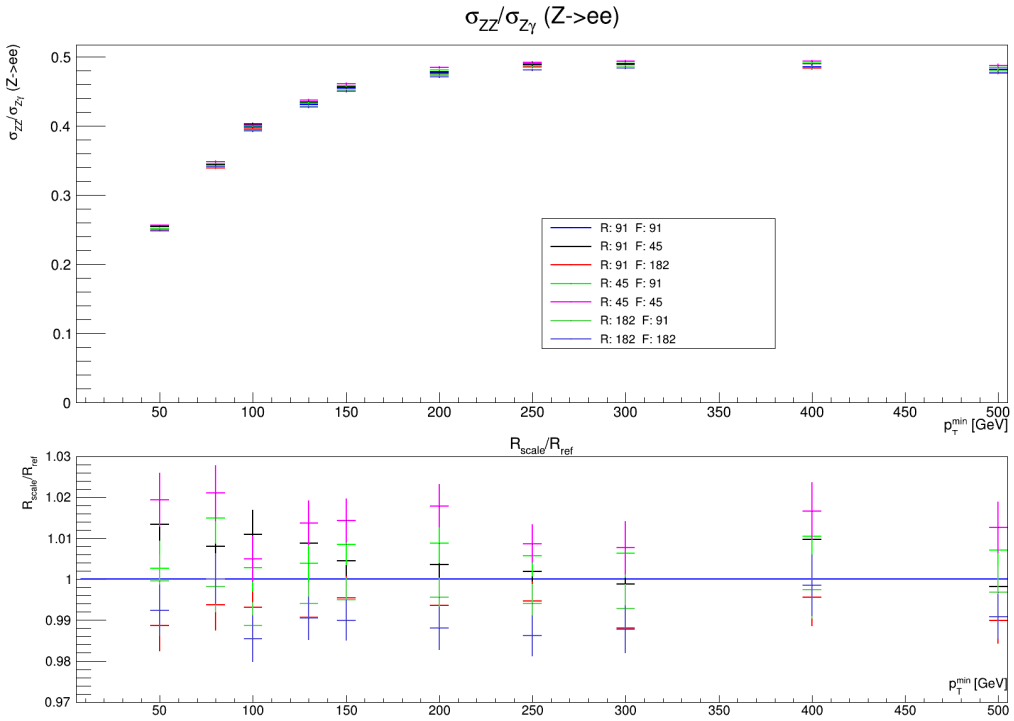
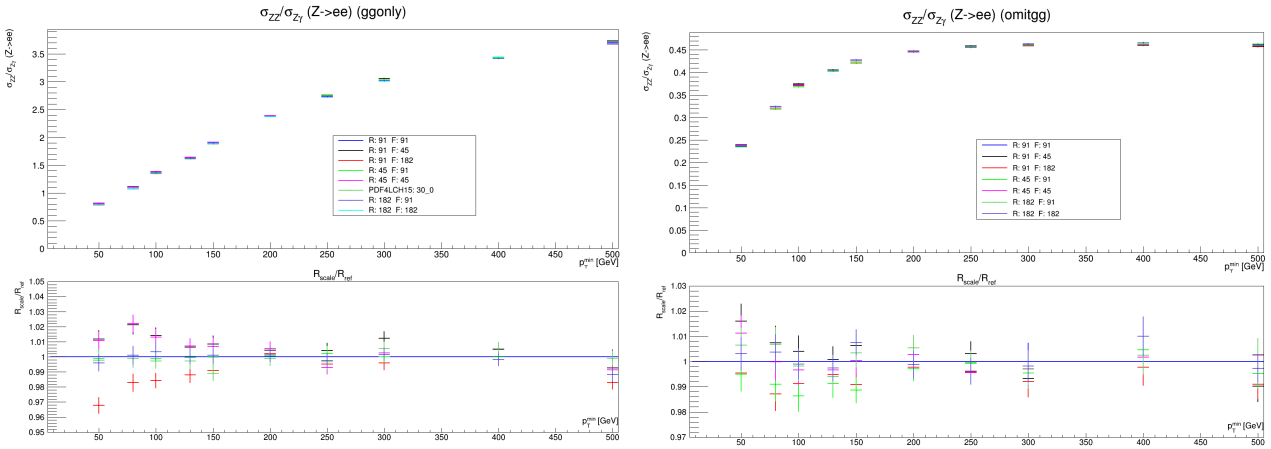


Figure 9: The ratio $R(p_T)$ for various choices for μ_R (R) and μ_F (F). The bottom panel shows the relative - with respect to the reference (R: 91, F: 91) for each scale. The uncertainties are statistical

The uncertainty due to the variation of scales around $R = 0.398$ is $\pm \approx 2\%$ for all p_T.

Looking at the the contribution of the gg subprocess separately from the $q\bar{q}$ and qg subprocesses, the result is shown in Figure 10.



(a) R_{gg} curve from gg subprocess only

(b) $R_{q\bar{q}/qg}$ curve from $q\bar{q}$ and qg subprocesses

Figure 10: The ratio $R(p_T)$ for various choices for μ_R (R) and μ_F (F) for the gg and qg+$q\bar{q}$ subprocesses separately. The bottom panel shows the relative difference with respect to the reference (R: 91, F: 91) for each scale. The uncertainties are statistical.

Gluon-gluon processes contribute to 8.6% of the total cross section for the ZZ process and 2.5% of the $Z + \gamma$ process. An uncertainty of $\pm \approx 2\%$ around $R_{gg} = 1.37$ at 100 GeV and is $< 4\%$ for all p_T. It remains to understand the shape and magnitude of the R curve for gg processes.

4.3 PDF variation

The PDF set used for reference is the `CT14.NN` PDF set. To study the variation due by varying PDFs, the PDF sets used are `PDF4LHC15`[5], constructed from the combination of `CT14,MMHT14` and `NNPDF3.0` PDF sets. These sets are provided by LHAPDF6[4]. `PDF4LHC15` gives access to different PDF groups. The group used here is `PDF4LHC15_nlo_30`, consisting of 30 PDF sets. While the most accurate uncertainties are given by `PDF4LHC15_nlo_100` sets, `PDF4LHC15_nlo_30` is used here for a faster, reasonably accurate estimate of the uncertainties.

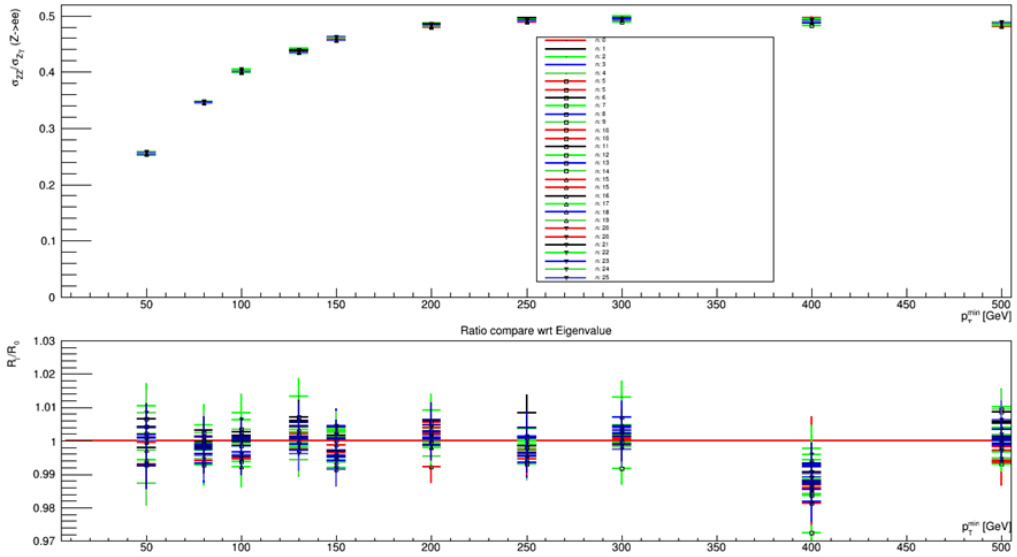


Figure 11: The ratio $R(p_T)$ for each of the 30 PDF sets in `PDF4LHC15_nlo_30`. The bottom plot shows the relative differences of sets 1-30, with respect to set 0 which is taken as the central value.

Fig.11 shows the comparison of the ratio curves $R(p_T)$ from the 30 member sets of `PDF4LHC15_nlo_30`. To measure the uncertainty due to these 30 sets, the relation as stated in Equation 20 in Ref [5] is used:

$$\delta^{PDF}\sigma = \sqrt{\sum_{k=1}^{N_{mem}} (\sigma^{(k)} - \sigma^{(0)})^2} \tag{6}$$

where N_{mem} is the number of member sets in the group, in this case, 30. The R curve obtained from the `PDF4LHC15_nlo_30` set is compared to the reference curve from `CT14.NN`, as shown in Figure 12:

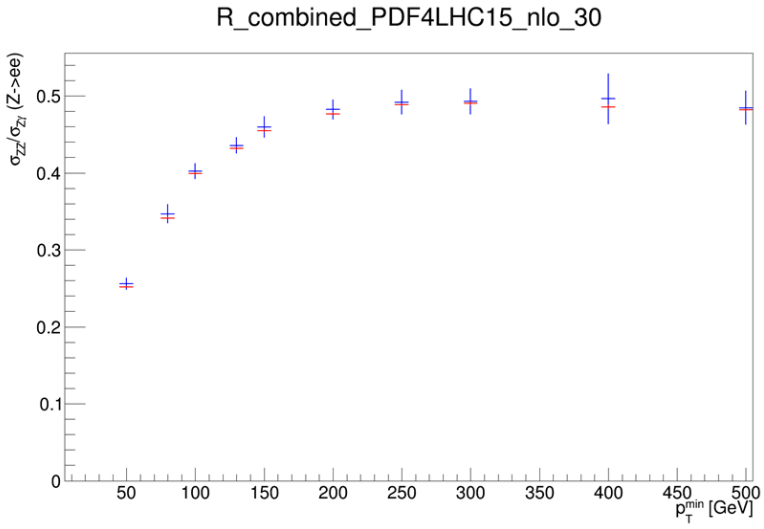


Figure 12: The ratio $R(p_T)$ of the `PDF4LHC15_nlo_30`, with combined uncertainties as given by Equation 6, to the reference constructed from the PDF set `CT14.NN`

Fig.12 shows a comparison between the central value of the sets in `PDF4LHC15_nlo_30` with the combined uncertainties, and the reference PDF set `CT14.NN`. The combined uncertainty around $R \approx 0.40$ is $\pm 2.55\%$ at 100 GeV. The PDF sets agree to within the uncertainty bounds. The contributions of the gg subprocess to the cross sections, and the R_{gg} curve are also studied, as shown in Figure 13.

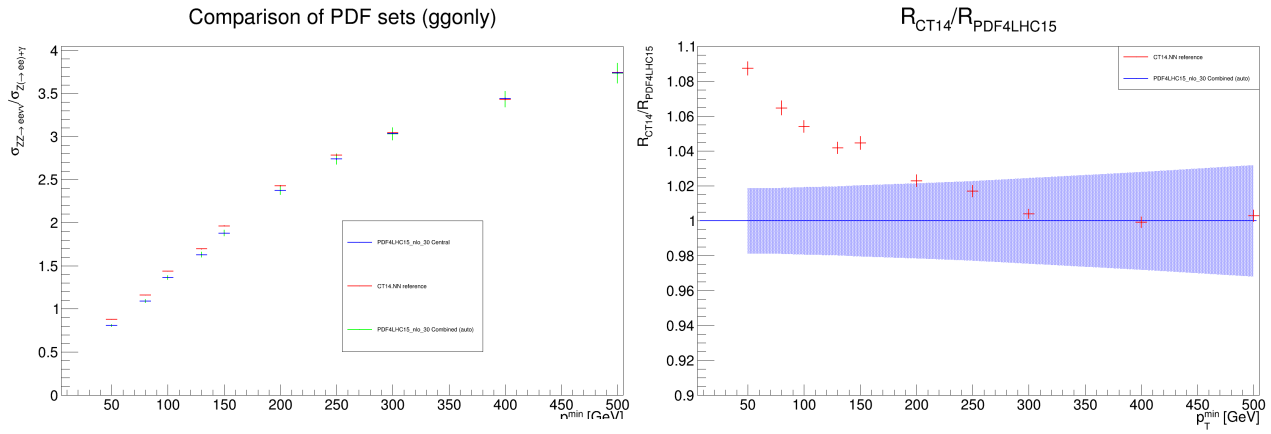


Figure 13: R_{gg} curve plotted from only the gg contribution to the cross sections of ZZ and $Z + \gamma$, using the combined uncertainties of `PDF4LHC15_nlo_30` sets. The figure on the right shows the ratio of the `CT14.NN` set to the `PDF4LHC15_nlo_30` set.

The gg contributions differ by a factor of 10. This curve appears to reach an a constant value at a higher p_T value than the ratio curve constructed from the total cross section. The gluon gluon process is of interest, thus it has also been compared to the reference `CT14.NN` set.

5 Conclusion

We propose a new method quantify the uncertainty from sources such as renormalization and factorization scales and different PDF distributions. From these, we observe that at high p_T, the value of R approaches 0.47. The uncertainty is quantified for $p_T > 100$ GeV slice to be $\approx 2\%$ from scale variation, and $\approx 2.55\%$ from PDF variation, around $R = 0.40$.

It remains to observe the effects of photon fragmentation, and to take a closer look at gg, $q\bar{q}$ and qg contributions to this ratio for `CT14.NN` as well as `PDF4LHC15_nlo_30` sets. In order to improve the uncertainty, we will run over the `PDF4LHC15_nlo_100` sets.

Acknowlegements

This work, part of the DESY Summer Programme 2017, was conducted under the patient supervision of Dr. Beate Heinemann. In addition to her guidance and advice, I had the help of Dr. Yee Chinn Yapp and Dr. Pieter Everaerts, who helped me work on streamlining the presentation of this work, and Dr. Sarah Heim, whose office is close enough to mine that I could bother her for the tiniest of details.

Bibliography

[1] *Using γ+ jets to calibrate the Standard Model $Z(\to \nu\nu)$+ jets background to new processes at the LHC*
S. Ask, M. A. Parker, T. Sandoval, M. E. Shea, W. J. Stirling
Cavendish Laboratory, University of Cambridge, CB3 0HE, UK; 2011
[arXiv:1107.2803]

[2] *Search for an invisibly decaying Higgs boson or dark matter candidates produced in association with a Z boson in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*
ATLAS Collaboration
arXiv:1708.09624

[3] *Monte Carlo for FeMtobarn processes (MCFM) v8.0 User Manual*
John Campbell, Keith Ellis, Walter Giele, Ciaran Williams
https://mcfm.fnal.gov/

[4] *LHAPDF6: parton density access in the LHC precision era*
Andy Buckley, James Ferrando, Stephen Lloyd, Karl Nordstrom, Ben Page, Martin Ruefenacht, Marek Schoenherr, Graeme Watt
arXiv:1412.7420

[5] *PDF4LHC recommendations for LHC Run II*
[arXiv:1510.03865]