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

Mangesh Sonawane
Supervisor: Dr. Beate Heinemann

July 18 - September 7 2017

Contents

Abstract

In the search for Dark Matter at the LHC, we look for events with large imbalance in transverse momentum. 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* $\approx 10\%$. *We aim 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 data. Using* $Z(\to ll) + \gamma$, *which is a pure signal and has a high* $BR * \sigma$. *In regions where* $p_T(\gamma) >> M_Z$, *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 fine tune the uncertainty on this curve, and use R *on* $Z\gamma$ *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: WIMPs do not register in the detector,

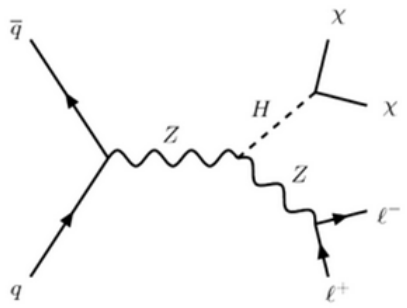


Figure 1: Associated production of Higgs

and thus result in a large missing transverse momentum (MET or E_T^{miss}).

Other processes that contribute to this signal 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 in the signal. The dominant source of background is the $ZZ \to ll\nu\nu$ process, contributing to $\approx 60\%$ of the background. Thus is it important

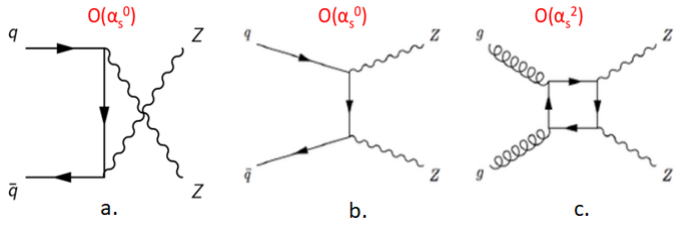


Figure 2: ZZ Production
a. & b. $q\bar{q} \to ZZ$ c. $gg \to ZZ$

to determine with precision the contribution to the background, along with the uncertainty associated with it. Currently, this is determined using Monte Carlo simulation, up to an uncertainty of $\pm 10\%$.

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

$$Z \to ee/\mu\mu \qquad \approx 6.8\%$$
$$Z \to \nu\nu \qquad \approx 20\%$$

Therefore,

$$ZZ \to ll\nu\nu \qquad \approx 3\%$$

It is difficult to determine with precision this contribution due to the small branching fraction of the $ZZ \to ll\nu\nu$. We could estimate this contribution by looking 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 [**?**], in the $p_T(\gamma) >> M_Z$ region, the $Z(\to ll) + \gamma$ process should be kinematically similar to $ZZ \to ll\nu\nu$. In addition to having a higher $BR * \sigma$ as compared to $ZZ \to ll\nu\nu$, the $Z(\to ll) + \gamma$ signal is also very pure. Thus, it should be possible to use $Z(\to ll) + \gamma$ data to estimate $ZZ \to ll\nu\nu$ contribution to the background, and obtain a more accurate prediction.

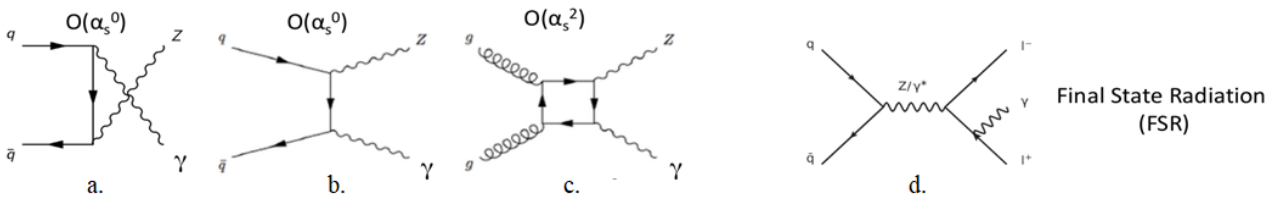


Figure 3: $Z + \gamma$ Production
a. & b. $q\bar{q} \to Z + \gamma$ c. $gg \to Z + \gamma$ d. Final State Radiation (FSR)

2 Approach

Following the method defined in the analysis paper[**?**], we define a variable $R(p_T)$ to be the ratio of the cross sections of $ZZ \to ll\nu\nu$ to $Z(\to ll) + \gamma$ 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. We would expect to see it 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 $p_T(\gamma) >> M_Z$ 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 $\sin^2\theta_W = 0.2315$, we obtain at moderate p_T values, $R \approx 1.4$[1].

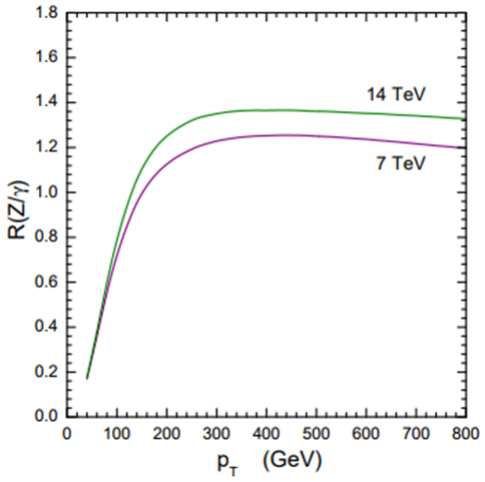


Figure 4: Ratio of the Z and γ p_T distributions [**?**]

We use a Monte Carlo generator, MCFM v8.0 [**?**] at NLO, in this case, to generate cross sections of $ZZ \to ll\nu\nu$ and $Z(\to ll) + \gamma$ 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. We then take a ratio of these cross sections to obtain the R curve as a function of p_T. We calculate the uncertainty on R by varying several parameters at the generator level, such as the renormalization and factorization scales, the PDF sets used, photon fragmentation, etc. We also aim to understand the 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.

However, the MCFM generator only produces $Z \to ee$ instead of $Z \to ll$. Thus, we account for this branching ration 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]

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

We use the following generator level cuts 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 [**?**]

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

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

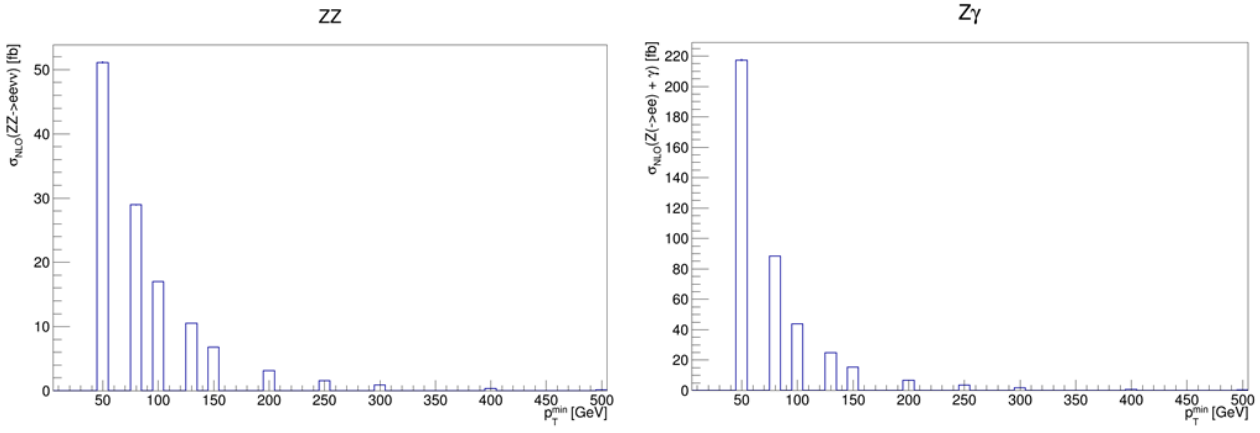
Cuts	$ZZ \to ee\nu\nu$	$Z(\to ee) + \gamma$
Process ID	87	300
M_{ee}	$81 < M_{ee} < 101$ GeV	$81 < M_{ee} < 101$ GeV
$M_{\nu\nu}$	$81 < M_{\nu\nu} < 101$ GeV	-
Order	NLO	NLO
PDFset	CT14.NN	CT14.NN
$p_T^{\text{lead}}(e)$	> 30 GeV	> 30 GeV
$\eta^{lead}(e)$	< 2.5	< 2.5
$p_T^{\text{sublead}}(e)$	> 20 GeV	> 20 GeV
$\eta^{sublead}(e)$	< 2.5	< 2.5
$p_T^{min}(\gamma)$	40 GeV	-
$\Delta R(\gamma, e)$	-	0.7
Renormalization scale	91.187 GeV	91.187 GeV
Factorization scale	91.187 GeV	91.187 GeV

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, we obtain the following cross sections. 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.

Taking the ratio:

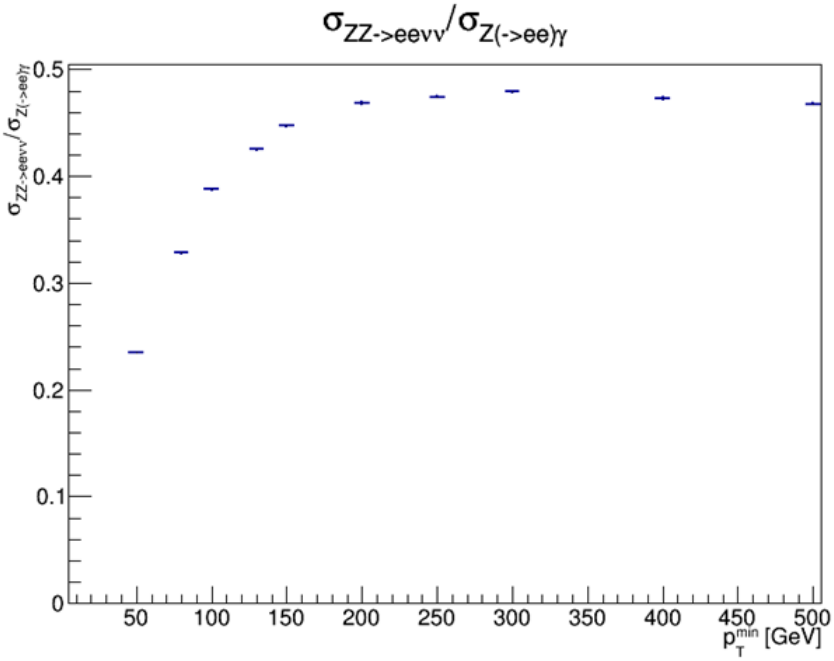


Figure 6: R curve as a function of p_T

We see that the R curve does seem to approach a constant value ≈ 0.47

4.1 Scale Variation

The Renormalization and Factorization scales are arbitrary parameters, important when considering higher order effects in QCD. Therefore, we vary the Renormalization and Factorization scales each by a factor of 2 in various combinations, to obtain the uncertainty due to these.

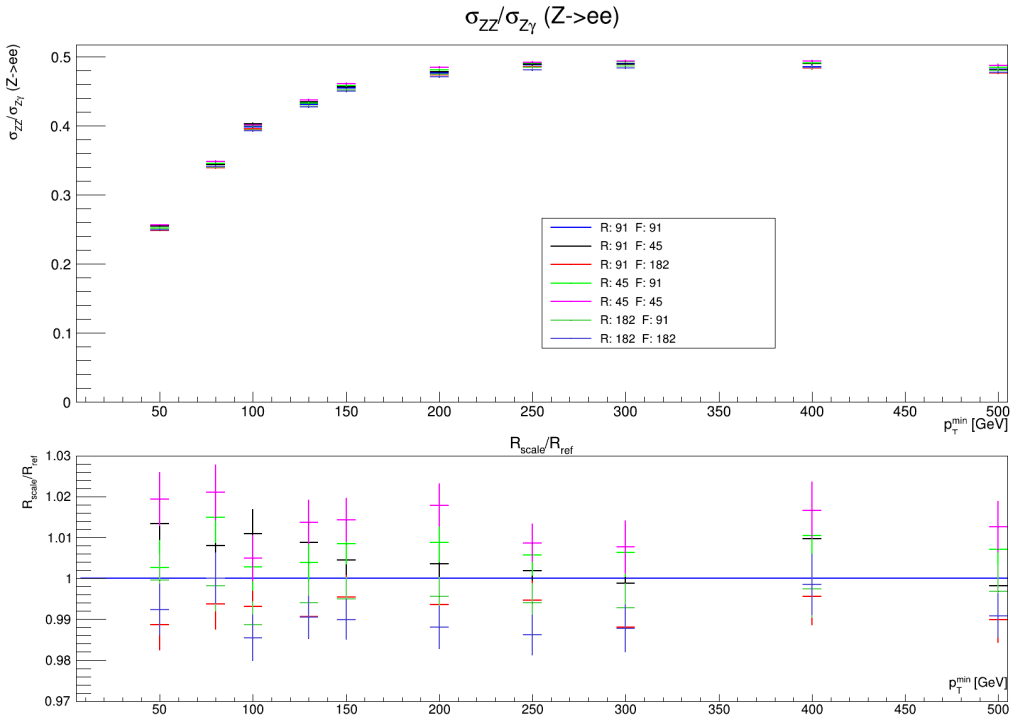


Figure 7: Varying the Renormalization (R:) and Factorization (F:) scales by a factor 2 in different combinations

The uncertainty due to the variation of scales around $R \approx 0.40$ is $\pm 2\%$ at 100 GeV.
Looking at the the contribution of the gg subprocess separately from the $q\bar{q}$ and qg subprocesses:

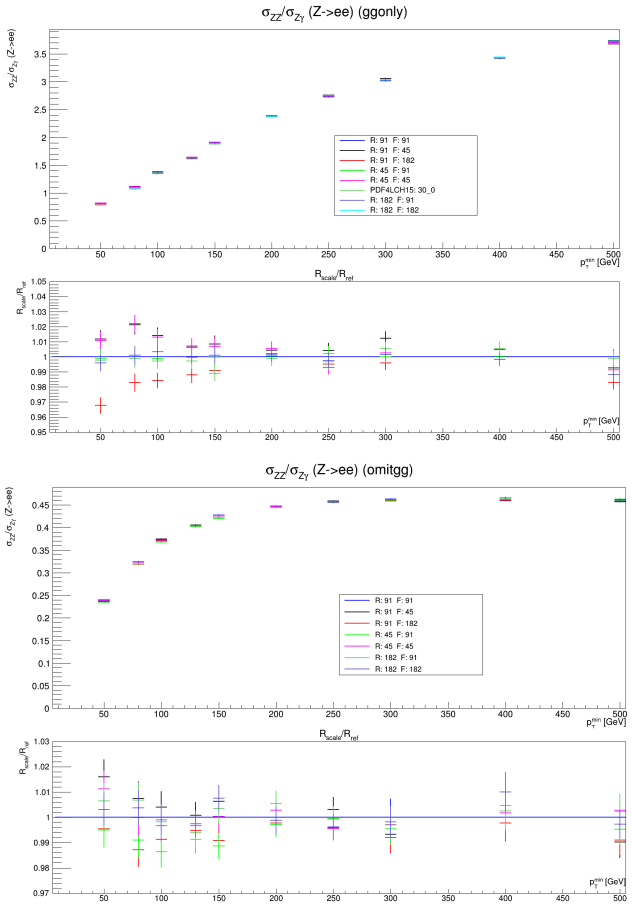


Figure 8: Varying the Renormalization (R:) and Factorization (F:) scales by a factor 2 for the gg subprocesses only

Gluon-gluon processes contribute to 8.6% of the total cross section for the ZZ process and 2.5% of the $Z + \gamma$ process. We observe an uncertainty of $\pm 2\%$ around $R_{gg} \approx 1.4$ at 100 GeV. It remains to understand the shape and magnitude of the R curve for gg processes. An additional data set in this plot is the cross sections using the `PDF4LHC15_nlo_30` PDFset. This is covered in the next heading.

4.2 PDF variation

The PDF set used for reference is the `CT14.NN` PDF set. To study the variation due by varying PDFs, we have used the `PDF4LHC15`[**?**], which is constructed from the combination of `CT14,MMHT14` and `NNPDF3.0` PDF sets. These sets are provided by LHAPDF6[**?**]. `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, we have used `PDF4LHC15_nlo_30` as we required a faster, reasonable accurate estimate of the uncertainties.

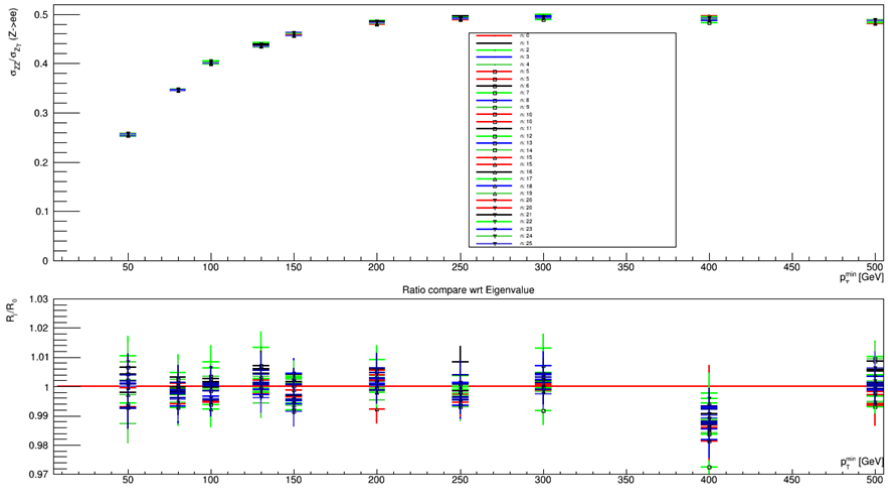


Figure 9: Comparison of the 30 PDF sets in `PDF4LHC15_nlo_30`

To measure the uncertainty due to these 30 sets, we use the relation as stated in Eq. (20) in Ref [**?**]:

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

where N_{mem} is the number of member sets in the group, in this case, 30. We compare the R curve obtained from the `PDF4LHC15_nlo_30` set to the reference curve from `CT14.NN`:

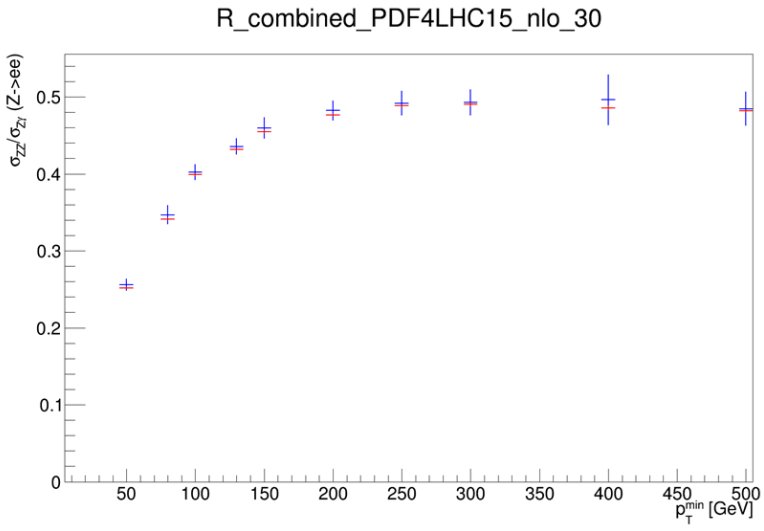


Figure 10: Comparison of the `PDF4LHC15_nlo_30`, with combined uncertainties, to the reference `CT14.NN`

The combined uncertainty around $R \approx 0.40$ is $\pm 2.55\%$ at 100 GeV. The PDF sets agree to within the uncertainty bounds. We also look at the contributions of the gg subprocess to the cross sections, and the R_{gg} curve.

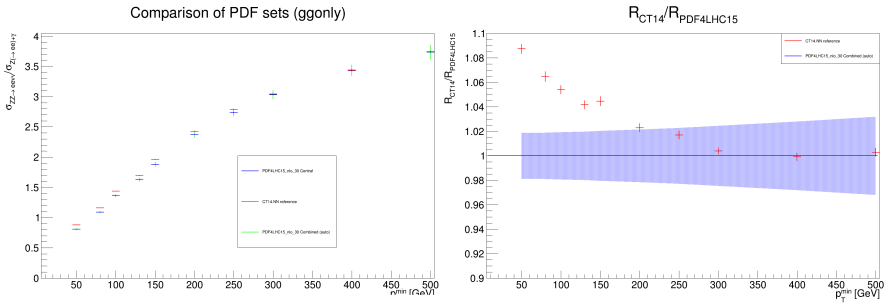


Figure 11: R_{gg} curve plotted from only the gg contribution to the cross sections of ZZ and $Z + \gamma$, using the combined uncertainties of `PDF4LHC15_nlo_30` sets

It is immediately apparent that the gg contributions differs by a factor of 10. The curve reaches an asymptotically flat value at ≈ 2 TeV and beyond. The gluon gluon process is of interest, thus it has also been compared to the reference `CT14.NN` set.

4.3 Lepton Cuts

To check the effects of lepton cuts on the ratio, we generate samples with similar parameters as those in Table 1. 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.

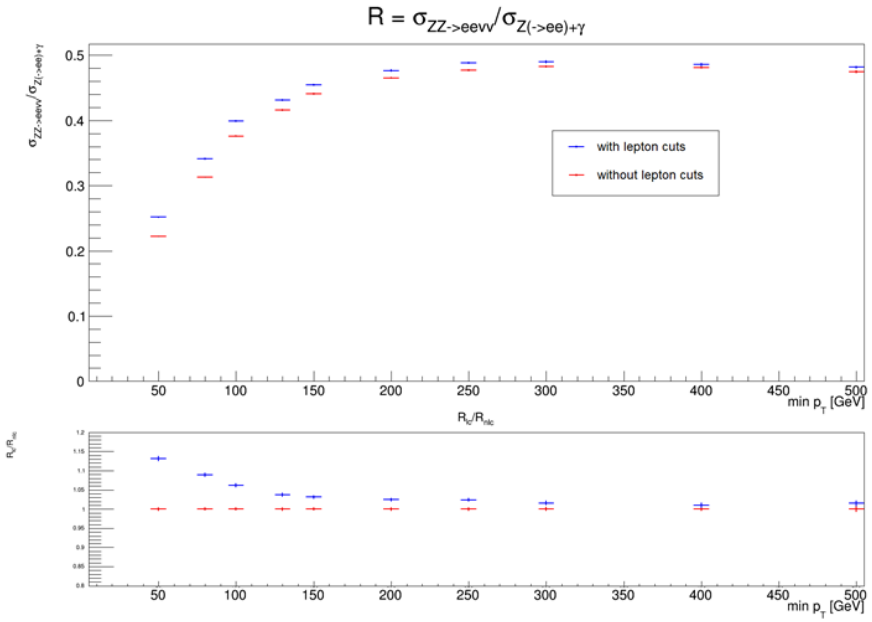


Figure 12: Comparison of reference R curve to R curve without lepton cuts

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

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

[3] *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

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