# Estimating $ZZ \rightarrow ll\nu\nu$ background in the $ll + E_T^{miss}$ final state using $Z\gamma \rightarrow ll\gamma$ data

A Thesis

submitted to
Indian Institue of Science Education and Research, Pune
in partial fulfillment of the requirements for the
BS-MS Dual Degree Programme

by

Mangesh Sonawane
Registration Number: 20121083

Indian Institute of Science Education and Research, Pune
Dr. Homi Bhabha Road,
Pashan, Pune 411008, INDIA

June 2017 - April 2018

Conducted at : DESY
Notkestraße 85,
22607, Hamburg
Germany

Supervisor: Dr. Beate Heinemann

# Certificate

This is to certify that this dissertation, entitled "Estimating $ZZ \to ll\nu\nu$ background in the $ll + E_T^{miss}$ final status using $Z\gamma \to ll\gamma$ data", submitted towards the partial fulfilment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research (IISER), Pune, represents the work carried out by Mangesh Sonawane at the Deutsches Elektronen-Synchrotron (DESY), Hamburg, under the supervision of Dr. Beate Heinemann, Professor of Experimental Particle Physics at the Institute of Physics, University of Freiburg, during the academic year 2017-2018.

Mangesh Sonawane | Dr. Beate Heinemann

Committee:
Dr. Beate Heinemann
Dr. Seema Sharma

I dedicate this thesis to my parents, Avinash and Ranjana Sonawane, my mentors, Dr. Sourabh Dube and Dr. Seema Sharma, and to my friends and colleagues and IISER, without whose timely advice and support this thesis would not have been made possible.

# Declaration

I hereby declare that the matter containined within the thesis entitled "Estimating $ZZ \rightarrow ll\nu\nu$ background in the $ll + E_T^{miss}$ final status using $Z\gamma \rightarrow ll\gamma$ data", contains the results of the work carried out by me at the Deutsches Elektronen-Synchrotron (DESY) Hamburg, under the supervision of Dr. Beate Heinemann, and the same has not been submitted elsewhere for any other degree.

Mangesh Sonawane　　　　　　　　　　Dr. Beate Heinemann

Committee:
Dr. Beate Heinemann
Dr. Seema Sharma

# Acknowledgements

I would like to express my deepest gratitude for Dr. Beate Heinemann for her guidance and patient mentoring. It's not just technical skills that I have acquired under her supervision, but also an understanding of how a physicist approaches the subject and tackles the inevitable problems that surface.

$\langle$ placeholder $\rangle$

# Abstract


In the search for Dark Matter (DM) at the LHC, SM particles are produced in association with DM particles, which are invisible as they don't interact with the detector. Thus events with large imbalance in transverse momentum are of interest. One such signature is $ll + E_T^{miss}$. The dominant background contributing to the search for DM in the $ll + E_T^{miss}$ is $ZZ \rightarrow ll\nu\nu$. Currently, this background is determined using Monte Carlo simulation, with an uncertainty of $\approx 10\%$ [1]. The goal of this study is to establish a data driven method to estimate this background, and reduce the uncertainty. Using $Z\gamma \rightarrow ll\gamma$ , which is a process with low backgrounds and has a high $BR{*}\sigma$, it is possible to estimate the $ZZ \rightarrow ll\nu\nu$ contribution. In regions where $p_T(\gamma) \gg M_Z$, the two processes are kinematically similar. They have the same production mechanisms, but differ due to the photon and Z boson couplings to the quarks being different, as well as the difference in mass (photons are massless, while Z bosons are massive). Introducing a transfer factor $R$ as the ratio $\sigma(ZZ)/\sigma(Z\gamma)$ which is determined from simulation, the contribution of $ZZ \rightarrow ll\nu\nu$ to the background can be estimated from $Z\gamma \rightarrow ll\gamma$ data. The uncertainty on the prediction of $R$ due to theoretical aspects is estimated in this work.

# Contents

# Chapter 1

# Introduction

# Chapter 2

# The Large Hadron Collider

# Chapter 3

# Analysis

## 3.1 Motivation

# Chapter 4

# Theoretical uncertainties on cross sections and ratio R

MCFM is a program that calculates cross sections for femtobarn-level processes at LO or NLO. In this study, MCFM v8.0 [4] is used to produce cross sections of $ZZ \to ll\nu\nu$ and $Z\gamma \to ll\gamma$ processes at NLO, with a selection of generator level cuts. The samples are generated with cuts on $E_T^{miss} = p_T(Z \to \nu\nu)$ for the $ZZ$ process and $E_T^{miss} = p_T(\gamma)$ for the $Z + \gamma$ process. A ratio of these cross sections is taken to obtain the $R$ distribution 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 on the ratio are studied. The contributions of the $q\bar{q}$ and $gg$ processes are estimated separately.

MCFM does not produce $Z \to ll$ 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 to obtain the inclusive value of $R$.

$$R_{inc} = R * \frac{BR(Z \to ee)}{BR(Z \to ee) * BR(Z \to \nu\nu) * 2} \tag{4.1}$$

## 4.1 Generator Parameters

The samples are generated using MCFM v8.0 for the following data points:

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

Table 4.1 lists the generator level settings used for the $ZZ$ and $Z + \gamma$ processes. All lepton cuts are consistent with the ones used in the ATLAS Z+$E_T^{miss}$ analysis.

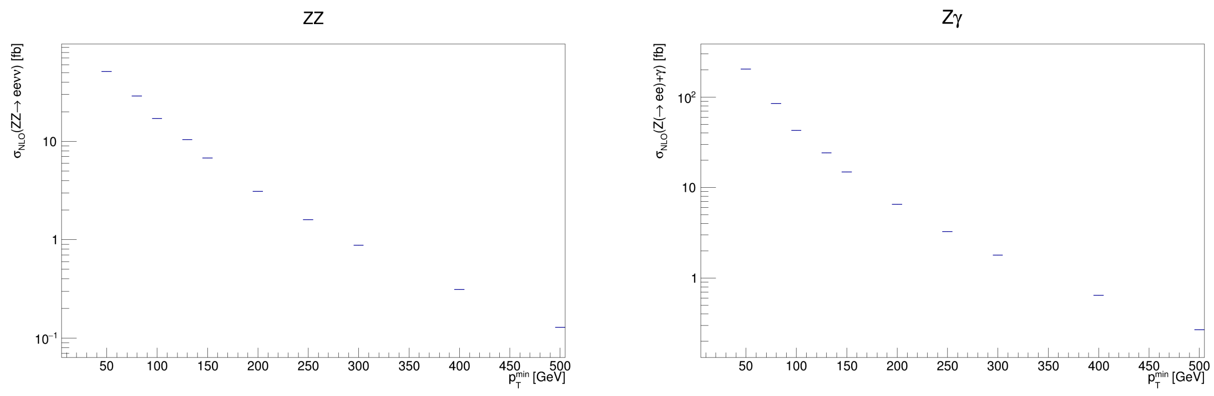
| **Cuts** | $ZZ \to ee\nu\nu$ | $Z(\to ee) + \gamma$ |
|---|---|---|
| Process ID | 87 | 300 |
| $M_{ee}$ | $81 < M_{ee} < 101$ GeV | $81 < M_{ee} < 101$ GeV |
| $M_{\nu\nu}$ | - GeV | - |
| Order | NLO | NLO |
| PDF set | CT14 | CT14 |
| $p_T^{\text{lead}}(e)$ | $> 30$ GeV | $> 30$ GeV |
| $\|\eta^{lead}(e)\|$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

Table 4.1: Settings 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.

## 4.2 Results

Using the settings listed in Table 4.1, the cross sections shown in Figure 4.1 are obtained. Throughout this analysis, this sample is the reference.




(a) (b)

Figure 4.1: Cross sections of $ZZ \to ee\nu\nu$ (left) and $Z\gamma \to ee\gamma$ (right) processes with the cuts as in Table 1. The vertical axis is in $\log_{10}$ scale. The leptonically decaying $Z$ boson decays to an $e^+e^-$ pair. There is no flavor constraint on the neutrinos.

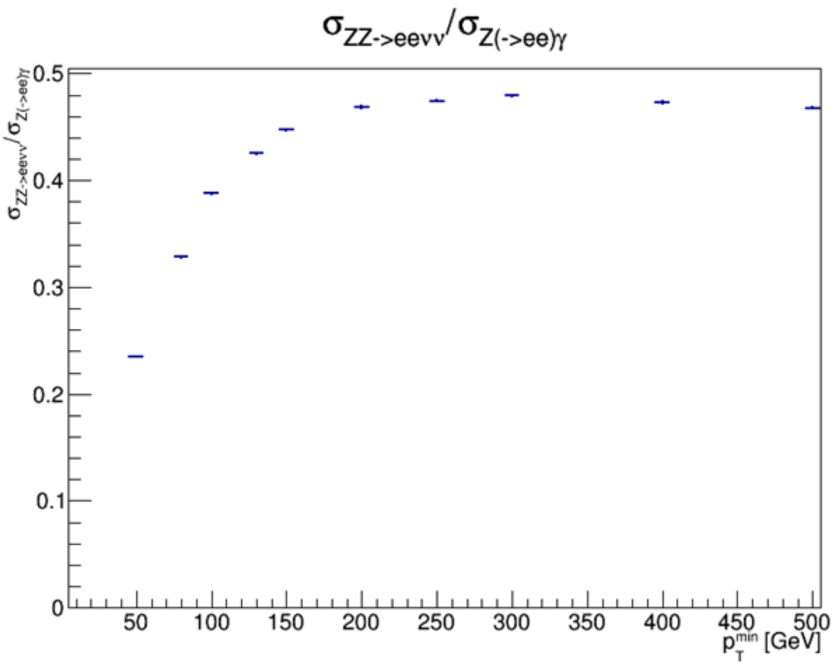



Figure 4.2: The transfer factor $R$ as a function of $p_T$, taken as a ratio of plots 4.1a and 4.1b. The leptonically decaying $Z$ boson decays to an $e^+e^-$ pair.

The ratio $R = \sigma(ZZ \to ee\nu\nu)/\sigma(Z\gamma \to ee\gamma)$ is shown in Figure 4.2. The $R$ value is observed to increase from $\approx 0.24$ at 50 GeV to $\approx 0.47$ at high $p_T$, where it is constant. When the branching ratio of $Z$ boson decaying selectively to $e^+e^-$, or to $\nu\nu$, is accounted for as shown in Equation 4.1, the resulting ratio $R(p_T)$ is shown in Figure 4.3, which shows the ratio of $\sigma(ZZ)$ to $\sigma(Z\gamma)$, i.e. if the $Z$ bosons do not decay further. The value of $R$ is observed to increase from $\approx 0.61$ at 50 GeV to $\approx 1.2$

at high $p_T$, in reasonable agreement with the simple approximate calculation presented in Section 2 of $R \approx 1.28$.

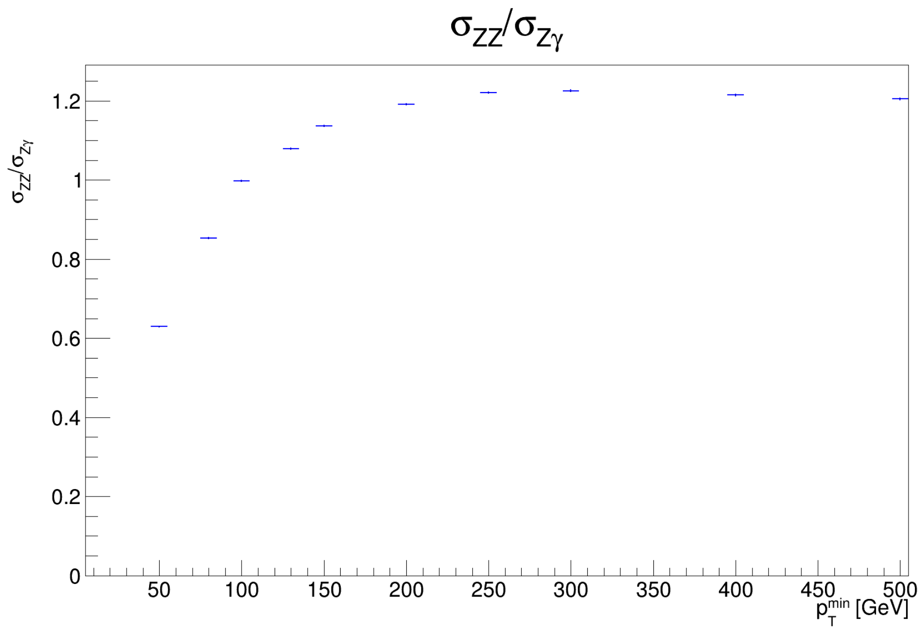
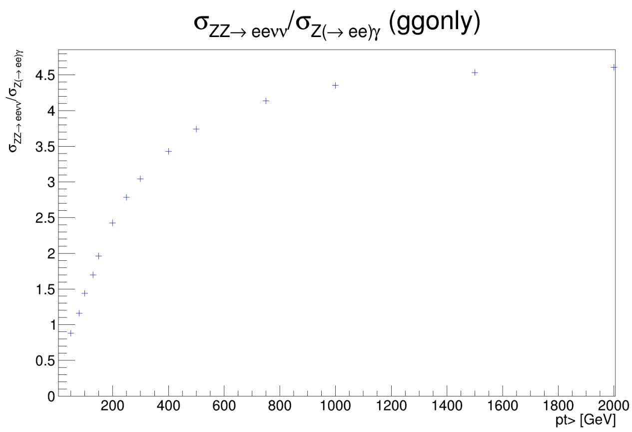
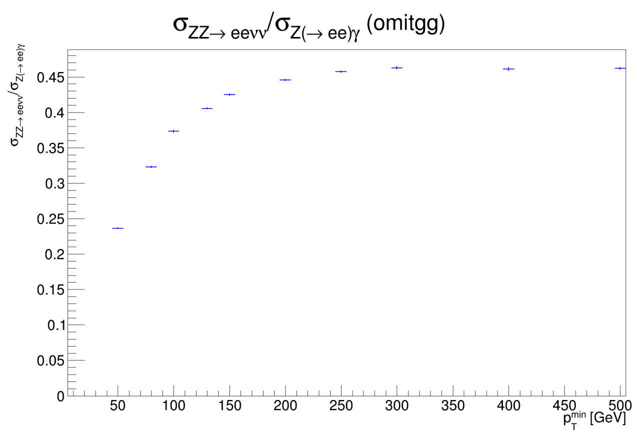



Figure 4.3: The transfer factor $R$ as a function of $p_T$, adjusted for the $Z \to ee$ and $Z \to \nu\nu$ branching ratios. This shows the $R = \sigma(ZZ)/\sigma(Z\gamma)$, where the $Z$ bosons do not decay.

Gluon-gluon processes contribute to 8.6% of the total cross section for the $ZZ$ process and 2.5% of the $Z + \gamma$ process. Figure 4.4 shows the transfer factor $R$ obtained from the $gg$ process, as well as $R$ from the $q\bar{q}$ and $qg$ processes.




(a)




(b)

Figure 4.4: $R(p_T)$ from the contributing $gg$ processes (a), and $qg$ and $q\bar{q}$ processes together (b). The $Z$ bosons decay further to $e^+e^-$ for the leptonic $Z$ boson, or $\nu\nu$ for the invisibly decaying $Z$ boson.

The $R_{gg}$ distribution is observed to approach an asymptotic value at a much higher $p_T = 1.5$ TeV. The shape and scale of the $R_{gg}$ distribution (Figure 4.4a) remain to be understood, as they differ from Figure 4.2.

### 4.2.1 Effect of Lepton Cuts

To check the effects of lepton cuts on the ratio, samples with the same parameters as those in Table 4.1 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. The ratio is affected by up to 15% as seen in Figure 4.5, and therefore for all following studies the lepton cuts are applied as they emulate the experimental cuts needed in the analysis.

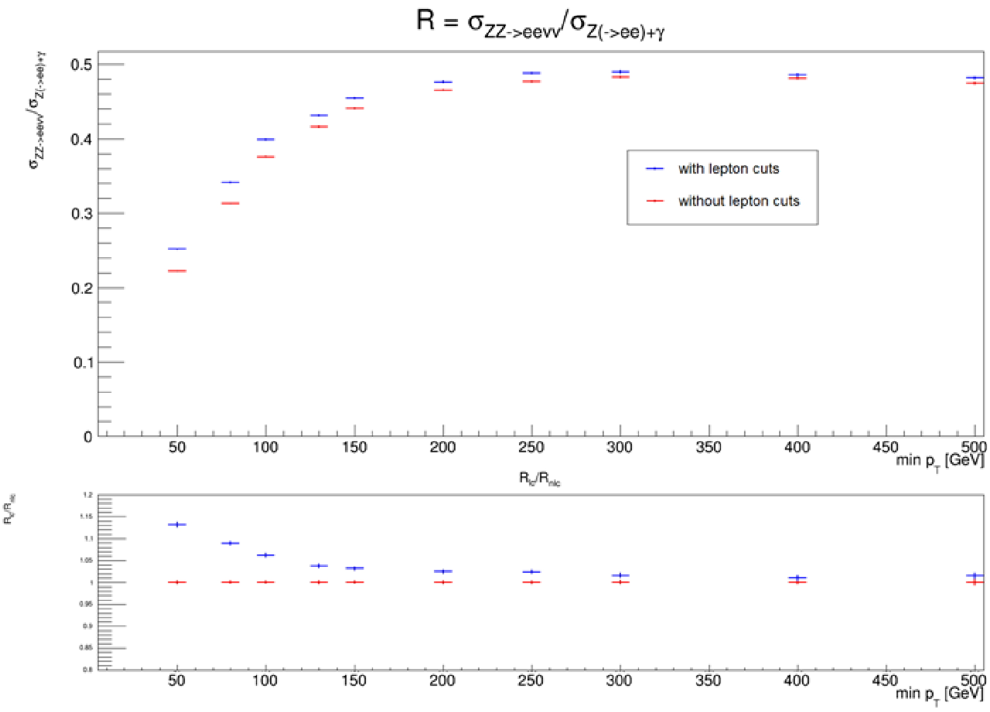



Figure 4.5: Comparison of reference the $R$ distribution to the $R$ distribution without lepton cuts

### 4.2.2 Uncertainty from Scale Variation

In higher order QCD calculations, perturbative corrections may be added to the vertices or propagators in a Feynman diagram. Physically, these corrections occur at very small time scales. These loop integrals that correspond to these corrections diverge.
The higher the order, the more difficult the calculation is. It is possible to introduce an arbitrary cut-off scale $\mu$ such that up to a given order, the effect of these corrections can be absorbed into the strong coupling constant $\alpha_s(\mu)$.
Two kinds of divergences are encountered: infrared divergences, and ultraviolet divergences. Infrared divergences occur for an on-shell internal propagator, and ultraviolet divergences are logarithmic divergences that occur as the integration variable approaches $\infty$. They correspond to physics at long and short distances[1], respectively. The infrared divergences are addressed by the inclusion of the factorization scale $\mu_F$, while the ultraviolet divergences are addressed by the inclusion of the renormalization scale $\mu_R$. These parameters are arbitrary, and are set by hand. These are then varied between $\frac{1}{2}\mu < \mu < 2\mu$ to obtain an indication of the dependence of the matrix element on the scales, and thus, the uncertainty around the chosen scale.

To address scale uncertainties in this study, the prescription used in [5] is followed. The central scale, $\mu_0$ is chosen to be *something* for both $ZZ \to ll\nu\nu$ and $Z\gamma \to ll\gamma$ samples, and seven-point variations are applied, i.e.

$$\frac{\mu_i}{\mu_0} = (1,1),(1,2),(2,1),(2,2),(0.5,1),(1,0.5),(0.5,0.5) \tag{4.2}$$

where $i = 0,...,6$. The central cross section value is taken to be the mean of the maximum and minimum cross sections resulting from this variation, and the uncertainty to be the half the difference

[1]Long distances are those where soft interactions take place, away from the hard parton-parton interaction. Short distances are those where the hard parton parton interactions occur.

between the same.

$$\sigma^{(V)}_{NLO} = \frac{1}{2}\left[\sigma^{(V,max)}_{NLO} + \sigma^{(V,min)}_{NLO}\right] \tag{4.3}$$

$$\delta\sigma^{(V)}_{NLO} = \frac{1}{2}\left[\sigma^{(V,max)}_{NLO} - \sigma^{(V,min)}_{NLO}\right] \tag{4.4}$$

where

$$\begin{aligned}\sigma^{(V,max)}_{NLO} &= max\left\{\sigma^{(V)}_{NLO}(p_T(V),\mu_i)|0 \leq i \leq 6\right\}\\ \sigma^{(V,min)}_{NLO} &= min\left\{\sigma^{(V)}_{NLO}(p_T(V),\mu_i)|0 \leq i \leq 6\right\}\end{aligned} \tag{4.5}$$

and $V = Z \to \nu\nu$ for $ZZ \to ll\nu\nu$ , or $V = \gamma$ for $Z\gamma \to ll\gamma$ . This uncertainty is propagated to $R$.

To estimate the degree of correlation between the processes, the process dependent part of the cross sections may be used. Since the study is conducted at NLO, the highest available term in the perturbative expansion is considered to define a K-factor.

$$\Delta K^{(V)}_{NLO} = \sigma^{(V)}_{NLO}(p_T)/\sigma^{(V)}_{LO}(p_T) \tag{4.6}$$

To estimate the unknown process dependent correlation effects, the difference between the QCD K-factors of the $ZZ \to ll\nu\nu$ and $Z\gamma \to ll\gamma$ processes is taken.

$$\delta^{(2)}\sigma_{NLO} = \Delta K^{(\gamma)}_{NLO}(p_T) - K^{(Z)}_{NLO}(p_T) \tag{4.7}$$

### 4.2.3 Uncertainty from PDF variation

Parton Distribution Functions (PDFs) characterize the fraction of proton momentum carried by partons as probability distributions. PDF sets are collections of PDFs that model parton momenta as accurately as possible. The PDF set used for reference is the `CT14`[6] PDF set. The uncertainty on the PDFs is studied by using the 30 variations provided by the `PDF4LHC15` set[7], constructed from the combination of `CT14,MMHT14`[8] and `NNPDF3.0`[9] PDF sets. These sets are provided by LHAPDF6[10]. `PDF4LHC15` provides a set of variations that include those determined by different groups (MSTW, CTEQ and NNPDF). The set used here is `PDF4LHC15_nlo_30`, consisting of 30 members. While the most accurate uncertainties are given by `PDF4LHC15_nlo_100` set, `PDF4LHC15_nlo_30` is used here for a faster, reasonably accurate estimate of the uncertainties.

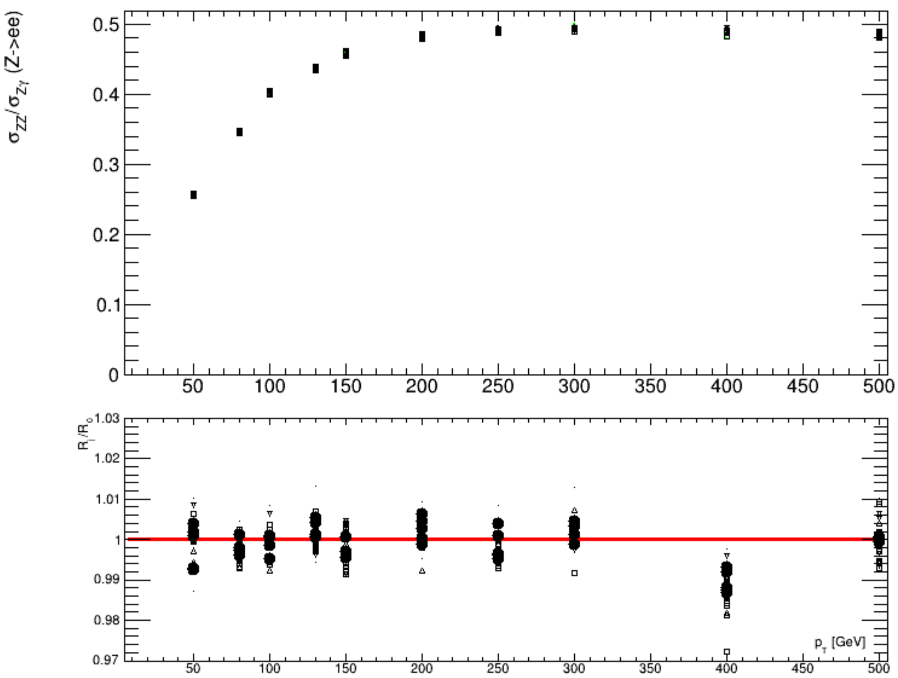



Figure 4.6: The ratio $R(p_T)$ for each of the 30 PDF sets in `PDF4LHC15_nlo_30`. The bottom plot shows the relative differences of sets 1-30, with respect to set 0 which is taken as the central value.

Fig.4.6 shows the comparison of the ratio $R(p_T)$ from the 30 member sets of `PDF4LHC15_nlo_30`. To measure the uncertainty due to these 30 sets, analogous to Equation 20 in Ref [7], Equation 4.8 is used:

$$\delta^{PDF} R = \sqrt{\sum_{k=1}^{N_{mem}} (R^{(k)} - R^{(0)})^2} \tag{4.8}$$

where $N_{mem}$ is the number of member sets in the group, in this case, 30. The $R$ distribution obtained from the `PDF4LHC15_nlo_30` set is compared to the reference distributions from `CT14`, as shown in Figure 4.7:

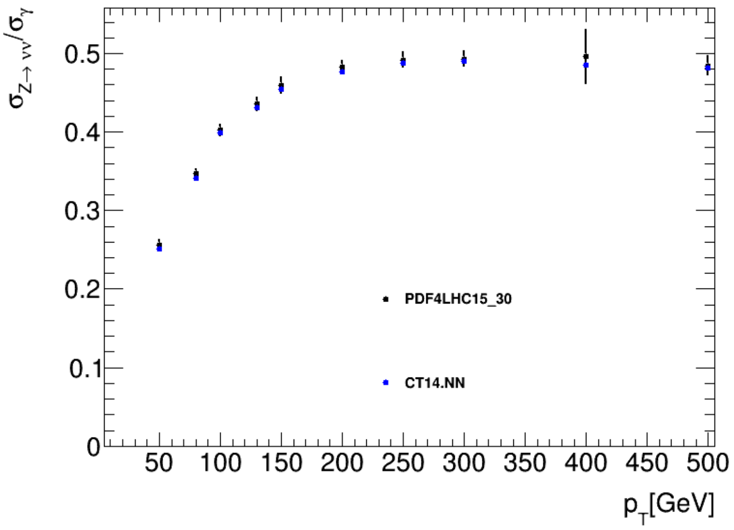



Figure 4.7: The ratio $R(p_T)$ calculated using the PDF sets in `PDF4LHC15_nlo_30` with combined uncertainties as given by Equation 4.8 (blue), compared to the reference constructed from the PDF set `CT14` (red).

Figure 4.7 shows a comparison between the central value of the sets in `PDF4LHC15_nlo_30` with the combined uncertainties, and the reference PDF set `CT14`. The combined uncertainty around $R \approx 0.40$ is $\pm 2.00\%$ at 100 GeV. The $R$ distributions drawn from the two PDF sets agree to within the uncertainty bounds.

### 4.2.4 Uncertainty from Photon Fragmentation

The $Z\gamma \to ll\gamma$ process may contain photons that arise from the hadron showers. It is therefore important to isolate the prompt photon from hadronic activity. This reduces unwanted background from pion decays, or fragmentation processes.

Experimentally, photon isolation is implemented with the following cuts:

$$\sum_{\in R_0} E_T(\text{had}) < \epsilon_h p_T^\gamma \qquad \text{or} \qquad \sum_{\in R_0} E_T(\text{had}) < E_T^{max} \tag{4.9}$$

limiting the transverse hadronic energy $E_T(had)$ in a cone of size $R_0 = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ around the photon, to some fraction of the photon $p_T$, or some fixed small cut-off.

The smooth cone isolation method of Frixione [11] is an alternative isolation procedure, which simplifies calculations by avoiding fragmentation contribuitions. The following isolation prescription is applied to the photon:

$$\sum_{R_{j\gamma} \in R_0} E_T(\text{had}) < \epsilon_h p_T^\gamma \left(\frac{1 - \cos R_{j\gamma}}{1 - \cos R_0}\right)^n . \tag{4.10}$$

where $R_{j\gamma}$ is the separation of the photon and the $j^{th}$ hadron. This requirement constrains the sum of hadronic energy inside a cone of radius $R_{j\gamma}$, for all separations $R_{j\gamma}$ less than a chosen cone

size $R_0$. This prescription allows soft radiation inside the photon cone, but collinear singularities are removed. The smooth cone isolation is infrared finite, thus fragmentation contributions do not need to be included.
The relative isolation, given by Equation 4.2.4 is used in experimental analyses, while smooth isolation is difficult to implement experimentally. However, comparing both methods gives us an estimate of the uncertainty due to the modelling of photon fragmentation.

In this analysis, $R_0$ is chosen to be 0.4 to agree with the experimental definition. The central value is chosen to be from the sample using smooth cone isolation (Frixione) with $\epsilon_h = 0.075$ and $n = 1$. These parameters are varied within a reasonable range to assess the uncertainty as shown in Figure 4.8.

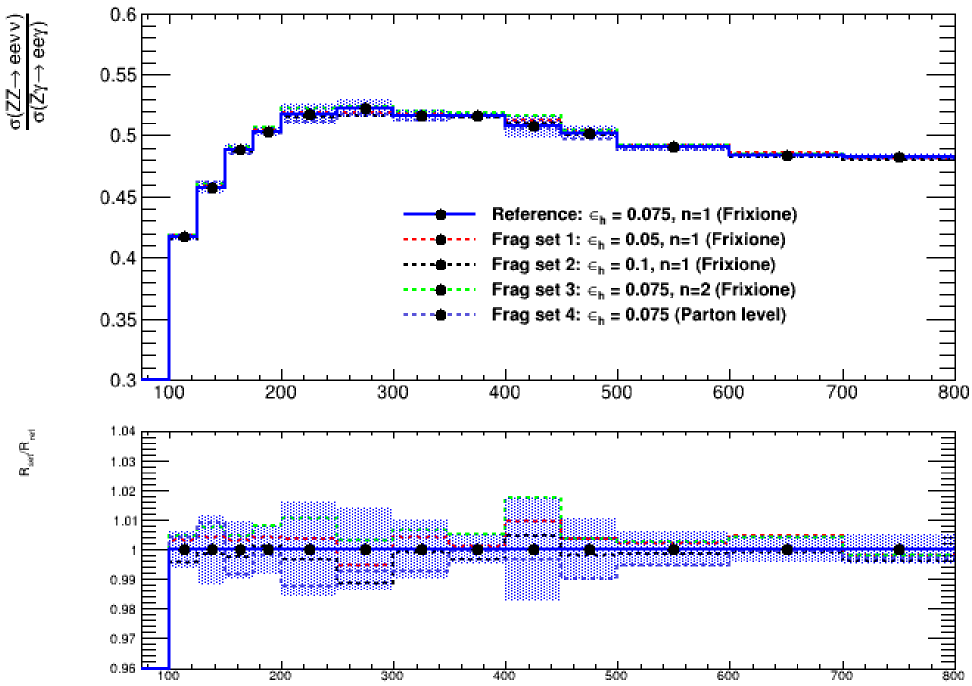



Figure 4.8: $R$ distribution as a function of $p_T$, showing the uncertainty due to variation of photon isolation parameters $\epsilon_h$ and $n$ in the smooth cone isolation procedure (Frixione), and $\epsilon_h$ in the photon isolation procedure. The lower panel shows the relative deviation of the varied sets from the central value, as well as the uncertainty band.

The uncertainty is calculated from the four sets listed in Figure 4.8:

$$\begin{aligned} \delta R_i &= |R_i - R_{ref}| \qquad\qquad i \in (1,2,3,4) \\ \delta R &= \sqrt{\max_{i=1,2,3} (\delta R_i)^2 + (\delta R_4)^2} \end{aligned} \tag{4.11}$$

as the effects assessed by changing the isolation definition in set 4, and varying the parameters in sets 1-3 are different.
The uncertainty is $< 2\%$ over the whole range, which has been extended up till 800 GeV.

# Bibliography


[1] *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

[2] *Using $\gamma+$ 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]

[3] *2017 Review of Particle Physics - Particle Listings*
**C. Patrignani *et al.* (Particle Data Group)**
Chin. Phys. C, 40, 100001 (2016)

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

[5] *Precise predictions for V +jets dark matter backgrounds*
**J.M. Lindert, S. Pozzorini, et al.**
arXiv:1705.04664

[6] *New parton distribution functions from a global analysis of quantum chromodynamics*
**Sayipjamal Dulat, Tie Jiun Hou, Jun Gao, Marco Guzzi, Joey Huston, P. Nadolsky, Jon Pumplin, Carl Schmidt, Daniel Stump, C. P. Yuan**
arXiv:1506.07443

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

[8] *Parton distributions in the LHC era: MMHT 2014 PDFs*
**L. A. Harland-Lang, A. D. Martin, P. Motylinski, R. S. Thorne**
arXiv:1412.3989

[9] *Parton distributions for the LHC Run II*
**The NNPDF Collaboration: Richard D. Ball, Valerio Bertone, Stefano Carrazza, Christopher S. Deans, Luigi Del Debbio, Stefano Forte, Alberto Guffanti, Nathan P. Hartland, Jose I. Latorre, Juan Rojo, Maria Ubiali**
arXiv:1410.8849

[10] *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

[11] *Isolated photons in perturbative QCD*
**S. Frixione**
Phys. Lett.B429(1998)369–374, hep-ph/9801442