Photon production and its objective of study have already been discussed in the chapter on “The Theoretical Perspective”. Prompt photon production plays an important role in probing the perturbative QCD. With the measurement of the differential cross section of inclusive isolated photons in pp collisions at the LHC energies, an unexplored region of photon production can thus be scrutinized. By probing the proton at a much deeper level one can get a more deeper insight into the structure of the proton and in a way yield more precise knowledge of its parton density functions. This chapter presents a measurement of the double-differential cross section of inclusive isolated photons in pp collisions at $\sqrt{s} = 7$ $TeV$. The amount of analysed data corresponds to 35.9 ± 1.4 $pb^{-1}$ of certified integrated luminosity recorded by the CMS detector upto November 2010. Isolated prompt photons were measured in wider ranges of transverse energies, $P_T = 25 - 400$ $GeV$, into pseudorapidity $|\eta| < 2.5$ and corresponding to the kinematic range of 0.007 < $x_T$ < 0.114, where $x_T = 2P_T/\sqrt{s}$. This measurement [185] extends the previous inclusive photon production measurement from CMS with 2.9 $pb^{-1}$ of recorded data at $\sqrt{s} = 7$ $TeV$ [83], which has already been summarized in Section 1.5. The term “inclusive” was used in regard to that no other additional physics objects were
explicitly required for the analysis and all the photons that satisfy the selection criteria were measured. The measurement was performed in fifteen bins of photon $P_T$ and in four bins of photon pseudorapidity $|\eta^\gamma|$. Prompt photons were separated from neutral meson background on a statistical basis, exploiting the distributions of specific variables defining the photon in the detector. The upcoming sections in this chapter will describe in detail about the datasets used, the event and object selections, the signal extraction procedure, the acceptance and efficiency calculation, and the unfolding mechanism. The unfolded cross section results were then compared with pQCD NLO calculations.

5.1 Data and Monte Carlo Samples

As mentioned, the full data sample employed for this analysis equals $35.9 \pm 1.4 \text{ pb}^{-1}$. Table 5.1 lists the processed datasets reconstructed in CMSSW 3.9.7 version of the CMS reconstruction software which were used for this analysis. The table also mentions the relevant run-ranges used.

<table>
<thead>
<tr>
<th>Run range</th>
<th>Data-set Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>136035-144114</td>
<td>/EG/Run2010A-Dec22ReReco_v1/AOD</td>
</tr>
<tr>
<td>146428-149294</td>
<td>/Photon/Run2010B-Dec22ReReco_v1/AOD</td>
</tr>
</tbody>
</table>

Table 5.1: The data-sets used for the analysis.

Due to increase in instantaneous luminosity at the LHC with time during the 2010 run period, the lowest unprescaled photon trigger $P_T$ threshold used in the HLT also changed with time. Thus, a run-range dependent trigger selection was used such that the lowest unprescaled HLT path was selected at every luminosity variation. This procedure was used in order to simplify the analysis and the corresponding loss in luminosity was of the order of only a few percent. For each bin of photon $P_T$ analysed, an OR of the available unprescaled trigger paths below the particular values of photon $P_T$ in that bin were used. Table 5.2 summarizes the unprescaled HLT trigger paths used for each photon $P_T$ bin and the corresponding integrated luminosity in that bin.
Consider a particular $P_T$ bin, 35 - 55 GeV, the HLT_Photon20_Cleaned_L1R trigger was used on the runs till the point a prescale factor is applied on the HLT path. The runs after that were picked out using the next lowest $P_T$ unprescaled path upto the time even that becomes prescaled. For this bin 35 - 55 GeV, an OR of the triggers used amounts to a total of $8.23 \text{ pb}^{-1}$ of data for photons. One can infer from the table that the integrated luminosity varies between $2.4 \text{ pb}^{-1}$ to $35.9 \text{ pb}^{-1}$. Also, it can be seen that the HLT threshold is at least 5 GeV less than photon candidate $P_T$, signifying that the HLT turns on fully or becomes 100% efficient around that value.

Monte Carlo simulated events, needed to model the data at various points in the analysis, were generated using PYTHIA 6.4 generator [152] using the Z2 tune [186]. The details of the respective studies will be discussed in the subsequent sections of this chapter. The necessary simulation steps described in Section 3.1 were considered for the generation. The list of PYTHIA generated MC samples used in the analysis is given in Table 5.3. The first table in Table 5.3 gives the list of all signal samples used which correspond to prompt photon processes. The second table in Table 5.3 on the other hand shows the list of QCD di-jet background samples. The last table summarizes the $Z \rightarrow ee$ samples used for various studies in the analysis. Along with assuming start-up conditions for calibration and alignment and no pileup added, the processed samples were reconstructed with CMS reconstruction software version CMSSW_3_8_6.
### Signal $\gamma + \text{jet}$ samples

<table>
<thead>
<tr>
<th>MC Sample $p_T$ bin (GeV/c)</th>
<th>$\sigma$ (pb)</th>
<th>N Events</th>
<th>$\int L dt$ (pb$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>15 - 30</td>
<td>$1.7 \times 10^3$</td>
<td>825840</td>
<td>$4.8 \times 10^6$</td>
</tr>
<tr>
<td>30 - 50</td>
<td>$1.7 \times 10^4$</td>
<td>925480</td>
<td>$5.5 \times 10^3$</td>
</tr>
<tr>
<td>50 - 80</td>
<td>$2.7 \times 10^3$</td>
<td>1024608</td>
<td>$3.8 \times 10^2$</td>
</tr>
<tr>
<td>80 - 120</td>
<td>$4.5 \times 10^2$</td>
<td>998215</td>
<td>$2.2 \times 10^3$</td>
</tr>
<tr>
<td>120 - 170</td>
<td>$8.4 \times 10^1$</td>
<td>1023361</td>
<td>$1.2 \times 10^4$</td>
</tr>
<tr>
<td>170 - 300</td>
<td>$2.3 \times 10^1$</td>
<td>1100000</td>
<td>$4.9 \times 10^4$</td>
</tr>
<tr>
<td>300 - 470</td>
<td>$1.5 \times 10^0$</td>
<td>1068904</td>
<td>$7.2 \times 10^5$</td>
</tr>
<tr>
<td>470 - 800</td>
<td>$1.3 \times 10^{-1}$</td>
<td>1083499</td>
<td>$8.2 \times 10^6$</td>
</tr>
</tbody>
</table>

### Background QCD di-jet samples

<table>
<thead>
<tr>
<th>MC Sample $p_T$ bin (GeV/c)</th>
<th>$\sigma$ (pb)</th>
<th>N Events</th>
<th>$\int L dt$ (pb$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>15 - 30</td>
<td>$8.2 \times 10^8$</td>
<td>4154640</td>
<td>$5.1 \times 10^{-2}$</td>
</tr>
<tr>
<td>30 - 50</td>
<td>$5.3 \times 10^7$</td>
<td>2814660</td>
<td>$5.3 \times 10^{-2}$</td>
</tr>
<tr>
<td>50 - 80</td>
<td>$6.4 \times 10^6$</td>
<td>2341546</td>
<td>$3.7 \times 10^{-1}$</td>
</tr>
<tr>
<td>80 - 120</td>
<td>$7.8 \times 10^5$</td>
<td>2408299</td>
<td>$3.1 \times 10^0$</td>
</tr>
<tr>
<td>120 - 170</td>
<td>$1.2 \times 10^5$</td>
<td>3035200</td>
<td>$2.6 \times 10^1$</td>
</tr>
<tr>
<td>170 - 300</td>
<td>$2.4 \times 10^4$</td>
<td>3210080</td>
<td>$1.3 \times 10^2$</td>
</tr>
<tr>
<td>300 - 470</td>
<td>$1.2 \times 10^3$</td>
<td>2149440</td>
<td>$1.8 \times 10^3$</td>
</tr>
<tr>
<td>470 - 600</td>
<td>$7.0 \times 10^1$</td>
<td>1999732</td>
<td>$2.9 \times 10^4$</td>
</tr>
<tr>
<td>600 - 800</td>
<td>$1.6 \times 10^1$</td>
<td>1979055</td>
<td>$1.3 \times 10^5$</td>
</tr>
</tbody>
</table>

### Other samples studied

<table>
<thead>
<tr>
<th>MC Sample $p_T$ bin (GeV/c)</th>
<th>$\sigma$ (pb)</th>
<th>N Events</th>
<th>$\int L dt$ (pb$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Z\text{Jet} \rightarrow ee$ (0 - 15)</td>
<td>$4.28 \times 10^4$</td>
<td>2.0 $\times 10^5$</td>
<td>$4.68 \times 10^1$</td>
</tr>
<tr>
<td>$Z\text{Jet} \rightarrow ee$ (15 - 20)</td>
<td>$1.45 \times 10^2$</td>
<td>2.0 $\times 10^5$</td>
<td>$1.38 \times 10^3$</td>
</tr>
<tr>
<td>$Z\text{Jet} \rightarrow ee$ (20 - 30)</td>
<td>$1.31 \times 10^2$</td>
<td>1.5 $\times 10^5$</td>
<td>$1.15 \times 10^3$</td>
</tr>
<tr>
<td>$Z\text{Jet} \rightarrow ee$ (30 - 50)</td>
<td>$8.40 \times 10^1$</td>
<td>1.5 $\times 10^5$</td>
<td>$1.79 \times 10^3$</td>
</tr>
<tr>
<td>$Z\text{Jet} \rightarrow ee$ (50 - 80)</td>
<td>$3.23 \times 10^1$</td>
<td>1.0 $\times 10^5$</td>
<td>$3.10 \times 10^3$</td>
</tr>
<tr>
<td>$Z\text{Jet} \rightarrow ee$ (80 - 120)</td>
<td>$9.99 \times 10^0$</td>
<td>1.0 $\times 10^5$</td>
<td>$1.00 \times 10^4$</td>
</tr>
<tr>
<td>$Z\text{Jet} \rightarrow ee$ (120 - 170)</td>
<td>$2.74 \times 10^0$</td>
<td>1.0 $\times 10^5$</td>
<td>$3.66 \times 10^4$</td>
</tr>
<tr>
<td>$Z\text{Jet} \rightarrow ee$ (170 - 230)</td>
<td>$7.22 \times 10^{-1}$</td>
<td>1.0 $\times 10^5$</td>
<td>$1.39 \times 10^5$</td>
</tr>
<tr>
<td>$Z\text{Jet} \rightarrow ee$ (230 - 300)</td>
<td>$1.94 \times 10^{-1}$</td>
<td>1.0 $\times 10^5$</td>
<td>$5.16 \times 10^5$</td>
</tr>
<tr>
<td>$Z\text{Jet} \rightarrow ee$ (300+)</td>
<td>$7.58 \times 10^{-1}$</td>
<td>1.0 $\times 10^5$</td>
<td>$1.32 \times 10^5$</td>
</tr>
</tbody>
</table>

Table 5.3: Monte Carlo samples used in the analysis.
For computational reasons, the MC samples were generated by splitting the process in several bins of $\hat{p}_T$ (the transverse momenta exchanged in the partonic interactions) and thus resulting in higher statistics for the corresponding region in phase space. For utilizing the complete phase space of the process as given by the MC simulation, the individual samples of each $\hat{p}_T$ bin for a particular process were combined by providing appropriate weights according to the corresponding sample cross section. Any overlap in between the $\hat{p}_T$ bins were removed.

The signal definition for events selected from simulation sample was given by combining direct and fragmentation photons and imposing a requirement on the generator-level isolation (sum of transverse energy of hadronic activity around the photon) to be less than 5 GeV. Rest of the photons not passing this requirement were considered as background. The cone size used to calculate the energy sum was chosen to match the one used for measuring the isolation at reconstruction level, $R = 0.4$. The generator level isolation values are typically much less than 5 GeV so this cut is not sensitive to uncertainties in the generator modeling of isolated photon processes.

5.2 Event Selection

Events were shortlisted for the analysis using dedicated single photon high level trigger paths, which have been extensively discussed in Sections 2.3 and 3.2.1.1. The summary of trigger paths used in this analysis has already been mentioned in Table 5.2. As explained in Section 3.2.1 a trigger path, for example "HLT_PhotonX_Cleaned_L1R", was formed from a sequence of modules. Seeded by an L1 trigger object of 8 GeV, the next sequences implement the high level trigger algorithms. For the concerned trigger path here, the energy deposits formed using the reconstruction algorithm at the HLT were supposed to have a minimum transverse energy threshold of $X$ GeV. The trigger path names shown in Table 5.2 depict what thresholds were used during the different run periods. Also, due to the anomalous signals seen in single crystals in the ECAL barrel, a swiss cross kind
of rejection criterion was applied on the superclusters at the HLT level (Section 3.2.1.1). Thus, the word “Cleaned” has been added to the HLT path name.

The events passing the trigger sequences were then subjected to additional set of selections to suppress the non-collision background. These consist of requirements on good vertex and track quality. Events with at least one good quality primary vertex were considered. The vertex was required to be within 24 cm of the nominal center of the detector along the beam axis (3σ of the luminous region) and have at least four degrees of freedom in the fit [187]. Further, for the events with more than ten reconstructed tracks, at least 25% of them were required to be of good quality [187].

5.3 Isolated Photons Selection

Once events with good vertex and tracks were selected, requirements were applied to select isolated photon candidates from the selection. Each event was required to have at least one isolated photon candidate within the detector acceptance with pseudorapidity of the supercluster $|\eta_{SC}| < 2.5$ and the transverse momenta of the candidate $P_T > 25$ GeV.

The fake signals coming from the anomalous interactions in the ECAL were already removed using a topological selection at the trigger level. The selection was enforced at the analysis level too. Additionally, the candidates originating in anomalous interactions were further suppressed offline by requiring a topological selection criteria based on the shower shape and the timing of the most energetic crystal in the supercluster to be consistent with collision events. The measured timing distribution of the normal electromagnetic shower signal is the convolution of time from light emission of the lead tungstate crystals and the response time of the electronics. For spikes, however, only the electronics time contributes to the measured timing. The measured hit time of the most energetic crystal (seed crystal) was required to be consistent with the collision timing within 5 standard deviations. Furthermore, the candidates were required to satisfy these conditions, $\sigma_{inj} >$
0.001 and $\sigma_{\phi} > 0.001$ in the barrel region. The contamination of spikes after selections was estimated to be less than 0.2\%. The other class of anomalous events which were observed in the endcaps region, were removed using the condition on shower shape, $\sigma_{\eta} > 0.001$. There were a total of 46 such events in the whole 2010 data set.

Further, high purity samples were selected requiring at least one good photon object based on some pre-selection criteria formed out of a number of identification variables already described in Section 4.1.2. The pre-selection helped to reject significant contribution of photons coming from background. However, a considerable amount of contamination coming from neutral mesons may still remain in the sample. For extracting the signal yield from the high purity sample thus selected, a template based method was employed. Signal and background templates were built based on a particular photon variable not used at the pre-selection level and which displayed different distribution patterns for the signal and background events. For the inclusive photon measurement, two separate templates were built based on totally different photon variables: a) *Conversion identification*, which used $E_T/p_T$ as the template variable ($E_T/p_T$ - the ratio of the transverse energy measured in the electromagnetic calorimeter ($E_T$) to the transverse momentum measured in the tracker ($p_T$) for converted photons) and b) *Isolation*, which used $ISO$ as the template variable ($ISO = ISO_{\text{Track}} + ISO_{\text{ECAL}} + ISO_{\text{HCAL}}$ - the combined isolation sum measured in the tracker, electromagnetic and hadronic calorimeters). The following sections of this chapter will discuss about the isolation method only. Table 5.4 summarizes the pre-selection and the side-band selection criteria for the isolation template method. As explained the pre-selection was applied to enhance the signal to background ratio in the fitting region, the side-band selection on the other hand was used to obtain background-enriched control samples which help in determining the parameter values for background probability density functions (bPDF).

**Binning of variables for measurement:** The analysis measures the cross section of inclusive photon production in various bins of photon pseudorapidity and photon transverse momenta. In terms of pseudorapidity, the analysis was performed in four bins: $|\eta^\gamma|$
Table 5.4: Photon pre-selection and side-band criteria for the isolation method.

<table>
<thead>
<tr>
<th>Photon Variable</th>
<th>Signal Region</th>
<th>Side-band Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pixel seed veto</td>
<td>True</td>
<td>True</td>
</tr>
<tr>
<td>H/E</td>
<td>&lt; 0.05</td>
<td>&lt; 0.05</td>
</tr>
<tr>
<td>barrel: $\sigma_{i\eta}$</td>
<td>0.010</td>
<td>0.0110 - 0.0115</td>
</tr>
</tbody>
</table>
| endcaps: $\sigma_{i\eta}$ | 0.028 | > 0.038

< 0.9, 0.9 < $|\eta|$ < 1.4442, 1.566 < $|\eta|$ < 2.1 and 2.1 < $|\eta|$ < 2.5. A narrow transition region 1.4442 < $|\eta|$ < 1.566 in pseudorapidity was excluded from the measurement due to low photon reconstruction efficiency in that region owing to lack of detector material there. The barrel region was sub-divided into two regions, $|\eta|$ < 0.9 and 0.9 < $|\eta|$ < 1.4442, because of the fact that the material budget in the two regions varied with respect to each other. The tracker material in front of ECAL region corresponding to $|\eta|$ < 0.9 was relatively much less as compared to 0.9 < $|\eta|$ < 1.4442. The fact can be well understood from Figure 2.11. The division of $|\eta|$ into two barrel bins can in fact also provide an additional piece of information in terms of studying the shape of the cross sections versus the photon pseudorapidity. The endcaps region was also sub-divided to add more information for understanding the cross section dependence on the photon pseudorapidity. The cross section measurement of the photon upto $|\eta|$ < 2.5 can be said to be an improvement over the earlier measurements provided by the CDF and DØ experiments at Fermilab, where the cross section was measured upto $|\eta|$ < 1.0 only. The measurement was also split in fifteen photon transverse momenta ($P_T$) bins given as following:


Larger statistic samples at lower values of $P_T$ motivates the $P_T$ binning to be much finer there. One can thus extract a finer detail picture of the cross section in that region of phase space. However, the trigger selection may result in a varied statistical sample for each $P_T$ bin. Equally sized bins of 5 GeV and a growing bin-size after 70 GeV ensured a reasonable compromise between the expected statistical error in each bin.
5.4 Signal Extraction

The signal yield \(N^\gamma\) from the high purity data formed after the pre-selection was determined using signal and background templates built on a property of the photon (not used at the pre-selection level) and whose pattern differed considerably among signal and background events. The purity value of the sample was actually extracted from the fit performed on the distribution of the respective photon variable obtained from data. The fit was performed using a function formed using the combined sum of the signal and the background component distribution functions of the respective variable. The shapes of the component distributions were taken from simulation events and were also validated by methods based on data. This section will basically describe the procedure used for obtaining signal yield from the isolation method.

5.4.1 Isolation Method

The isolation method utilizes the difference in shape of the isolation energy sum around the photon candidate from the tracker and calorimeters, \(ISO = Iso_{Track} + Iso_{ECAL} + Iso_{HCAL}\), for the signal extraction procedure. The good separation power of the isolation sum helps in modeling well the signal and background probability density functions in the fitting region. For signal photon, only particles from underlying event, pile-up, and detector noise may contribute to the isolation sum, \(ISO\). For this reason the signal \(ISO\) distribution falls off at around 5 GeV. While for photons coming from background, particles produced together with neutral mesons (\(\pi^0\) and \(\eta\)) and their decay products contribute to the isolation energy sum and lead to a widened \(ISO\) distribution.

The signal and background component distributions were parametrized with analytic functional forms, mainly driven by the \(ISO\) distributions observed in the simulation and the control sample from data. For every \(P_T\) bin, the signal component shape \((S)\) was taken from Monte Carlo simulation and corrected for differences between data and MC, while the background component shape \((B)\) was derived from data. The signal fraction
was estimated from a extended maximum likelihood fit to the ISO distribution in data.

**Signal Component:** The signal component shape functional form was derived from the convolution of the lifetime function (exponential function) with a Gaussian distribution, as given below:

\[
S(x) = \frac{1}{p_0} e^{\left(\frac{x^2}{p_0^2}\right) - \left(\frac{x - p_1}{p_0}\right)} \times \left[1 - F_{\text{req}}\left(\frac{p_2}{p_0} - \frac{x - p_1}{p_2}\right)\right] 
\]

(5.1)

**Background Component:** The background component shape was derived from an inverted ARGUS function, as given below:

\[
B(x) = \left[1 - e^{p_3(x - p_4)}\right] \times \left[1 - p_5(x - p_4)\right]^{p_6}
\]

(5.2)

Here, \(x\) stands for the ISO variable. These functional forms, \(S(x)\) and \(B(x)\) have both been tested and validated with simulation and data signal or background control sample. The \(S(x)\) describes well the ISO distributions for both electrons in data and MC and signal photons in MC, while the \(B(x)\) describes well the ISO distributions for both side-band data and background photons in MC.

The templates used to initialize and constrain the functions were obtained as following,

- The signal template was built from the signal events from \(\gamma + \text{jet} \) MC. The events from signal MC were scaled using correction factors obtained to account for any differences observed between \(Z^0 \rightarrow e^+e^-\) data and MC events. To derive these factors, low energy bremsstrahlung electrons were selected from \(Z^0 \rightarrow e^+e^-\) decays as described in Ref [188]. Bremsstrahlung activity is evaluated from the relative difference between the momentum measured at the last point on the electron track \((p_{\text{out}})\) and the momentum measured at the origin \((p_{\text{in}})\). Low bremsstrahlung electrons were thus selected based on this ratio, given as \((p_{\text{in}} - p_{\text{out}})/p_{\text{in}} < 0.15).\n
- The background template was extracted from data using the side-band selections given in Table 5.4 and thereafter applying corrections accounting for the difference
between the MC background in signal region and the MC side-band.

**Fitting:** For each photon $P_T$ bin, data were fit using the functional form: $f(\text{ISO}) = N_s S(\text{ISO}) + N_b B(\text{ISO})$, where $N_s$ and $N_b$ were the estimated number of signal and background events in the $P_T$ bin. The fitting was performed using an unbinned, extended maximum likelihood technique, by minimizing the function $-\ln L$ given as,

$$-\ln L = (N_s + N_b) - \sum_{i=1}^{N} \ln(N_s P_s^i + N_b P_b^i)$$  \hspace{1cm} (5.3)

where, $P_s$ and $P_b$ are the signal and background probability density functions evaluated with the ISO of the photon candidate $i$, $N$ is the number of measured events, $N_s$ and $N_b$ are the number of signal and background events. The fit to the measured ISO distribution was performed for the region defined as, $-1 < x < 20 \text{ GeV}$.

The signal and background probability density functions were obtained after normalizing the integrals of $S(x)$ and $B(x)$ to unity within the fit range, respectively:

$$P_s^i = \frac{1}{\int_{-1}^{20} S(x)dx} S(x^i) \quad \text{and} \quad P_b^i = \frac{1}{\int_{-1}^{20} B(x)dx} B(x^i)$$  \hspace{1cm} (5.4)

The fitting procedure was performed in two steps. For the first step, the signal and background templates were fitted with the functions $S(x)$ and $B(x)$ respectively. After that in the second step, the ISO distribution in data was fitted such that the various fitting parameters, used in equations 5.1 and 5.2, were either determined from the fit to the data directly or were fixed or constrained while performing the fit. The shape parameters for signal ($p_1$ and $p_2$) and for background ($p_5$ and $p_6$) were free parameters without any constraint (together with $N_s$ and $N_b$) and were obtained from fit to the data directly. While $p_0$ - the signal exponential tail and $p_3$ - the background turn-on power were constrained to be within the uncertainty of the template fit. The parameter $p_4$ which
is the background starting point, was fixed.

Figure 5.1 shows the results of likelihood fit for barrel and endcaps photon data in one bin of photon $P_T$ analyzed. Table 5.5 lists the fitted yields, for all $P_T$ bins, with statistical uncertainties and Figure 5.2 shows the purity as a function of $P_T$ and $|\eta|$ bins. For the last $P_T$ bin 300 - 400 GeV/c, the fitting was not performed due to small number of events. The purity in this case was assumed to be 100% and the yields after pre-selection were simply reported.

### 5.5 Cross Section Measurement

The differential inclusive photon production cross section measured can be described according to the following expression:
\[
\frac{d^2\sigma}{dP_T d\eta} = N_s(\Delta P_T) \frac{U}{\Delta P_T \Delta \eta . L(\Delta P_T) \varepsilon(\Delta P_T)}
\]  

(5.5)

where, \(N_s(\Delta P_T)\) is the number of prompt isolated photons (signal yield measured from signal extraction method) in a particular \(\Delta P_T\) bin, \(\varepsilon(\Delta P_T)\) is the selection efficiency in that \(\Delta P_T\) bin, \(U\) denotes the bin-by-bin unfolding correction accounting for the smearing in the reconstructed energy, position and isolation quantities, and \(L(\Delta P_T)\) is the integrated luminosity corresponding to the analysed dataset.

The efficiency and the unfolding corrections were obtained from data whenever possible. The details of the procedure of extraction of the two quantities is summarized in the following sub-sections.

**5.5.1 Efficiency Estimation**

The selection efficiency for the photons can be factorized into contributions coming from various inputs to this analysis. The following expression takes into account all the con-
tributions required for evaluating the selection efficiency:

\[
\varepsilon = \varepsilon_{TRIG} \times \varepsilon_{RECO} \times \varepsilon_{ID1} \times \varepsilon_{ID2}
\]  

(5.6)

In order to measure user-defined object efficiencies, a data-driven technique called “Tag and Probe” was developed. The technique employs di-object resonances observed in data. As photons and electrons leave similar kind of energy deposits in the electromagnetic calorimeter, the “Tag and Probe” technique uses \( Z \rightarrow ee \) events to calculate efficiencies for analysis measuring photons. The tag is an electron which is well reconstructed by a stringent electron selection criteria. The probe is an electron supercluster (SC) which
passes the photon selection criteria and depends on the type of efficiency to be estimated. The tag and probe electrons thus selected are then required to have a reconstructed invariant mass compatible with the Z-mass. The probes are then examined depending on certain selection criteria defined by the efficiency which is being estimated. The efficiency is then evaluated either by the simple counting method (counting the number of pass and fail probes) or by the fitting method where fits are performed to the invariant mass distribution of the electrons in the pass and fail samples. Lineshapes of passing and failing probe (with tag object) are fit separately with a signal + background model. The efficiency is computed using the ratio of the signal yields in these two lineshapes. This technique was first described for CMS in [189]. \[ε_{\text{TRIG}}\] and \[ε_{\text{ID1}}\] were computed with the help of the “Tag and Probe” technique.

The text given below will be expanding on the definitions and methods used for estimation of \[ε_{\text{TRIG}}\] and \[ε_{\text{RECO}}\]. While the text following it will be giving details on the estimation of \[ε_{\text{ID1}}\] and \[ε_{\text{ID2}}\], using the relevant definitions of each, for the isolation method.

**Trigger Efficiency (\[ε_{\text{TRIG}}\]):** For determining the efficiency of single photon HLT paths (mentioned in Table 5.2), data-driven methods like “Tag and Probe” were used. Tags were chosen as reconstructed electrons with \[P_T > 20 \text{ GeV}\] and lying within the detector acceptance (within \[|\eta| < 1.4442\] and \[1.566 < |\eta| < 2.5\]) excluding the transition region. The tag electrons were further required to pass a stringent electron selection criteria [189]. \[ε_{\text{TRIG}}\], as measured in \(Z \rightarrow ee\) events, was found to have a plateau at 99.8 ± 0.1 % in the barrel region and 99.0 ± 0.7 % in the endcaps.

Probes were photons with \(P_T > 20 \text{ GeV}\) and lying within the detector acceptance. Probe photons were further selected with the set of photon pre-selection requirements as given in Table 5.4. An invariant mass cut, requiring \(60 < M_{e\gamma} < 120 \text{ GeV}\) (mass window of the Z boson), was imposed on the tag-probe pair. This requirement ensured a high-purity sample of tag-probe pairs. For the “Tag and Probe” technique, the efficiency of an HLT path was defined as the ratio of the number of probes passing an HLT path (having an HLT object match) to the total number of probes (including both failing and passing an HLT path). Following this procedure \[ε_{\text{TRIG}}\], as measured in \(Z \rightarrow ee\) events, was found to have a plateau at 99.8 ± 0.1 % in the barrel region and 99.0 ± 0.7 % in the endcaps.
Reconstruction Efficiency ($\varepsilon_{\text{RECO}}$): The reconstruction efficiency ($\varepsilon_{\text{RECO}}$) was defined as the ratio of the number of true prompt photons that were reconstructed within detector acceptance to the number of true prompt photons that were generated with true $P_T$ and have a generator-level isolation less than 5 GeV. The value of $\varepsilon_{\text{RECO}}$ was found to be 99.8% for all $P_T$ and $|\eta|$ bins as determined from simulated photon signal events.

Photon pre-selection ($\varepsilon_{\text{ID1}}$): The efficiency due to the requirements of the photon pre-selection ($\varepsilon_{\text{ID1}}$) as given in Table 5.4, can simply be defined as the ratio of the number of reconstructed, true prompt photons satisfying the pre-selection requirements (with additional requirement of generator-level isolation < 5 GeV) to the total number of true prompt photons that were reconstructed. Initially, the value of $\varepsilon_{\text{ID1}}$ was determined from simulated photon signal events and then compared with results from data derived using the “Tag and Probe” technique employing electrons from $Z \rightarrow e^+e^-$ decays. The level of agreement between the two was quite good and the small differences, of the order of a percent, were taken as a data-to-simulation scaling factor. The simulated isolation efficiency was later multiplied by this data-to-simulation scaling factor to obtain $\varepsilon_{\text{ID1}}$ for efficiency equation. The simulation predicts a few percent difference in the efficiency $\varepsilon_{\text{ID1}}$ between photons and electrons; half of this difference was taken as a systematic uncertainty. In addition, the scaling factor was measured in various time periods that correspond to different average numbers of pile-up events due to multiple pp interactions in the same bunch crossing. The envelope of the full variation, approximately 3 - 5% depending on the selection criteria and photon $P_T$, was taken as the systematic uncertainty. The scaling factor obtained using “Tag and Probe” method varies from 0.971 ± 0.073 to 0.955 ± 0.032 for the barrel and from 0.998 ± 0.056 to 0.990 ± 0.056 for the endcaps as $P_T$ increases from 20 GeV to 45 GeV.

Pixel Veto ($\varepsilon_{\text{ID2}}$): The symbol $\varepsilon_{\text{id2=pe}}$ represents the efficiency of the pixel veto requirement. The value of $\varepsilon_{\text{pe}}$ was determined from simulated photon signal events first and then multiplied by a data-to-simulation scaling factor. The data-to-simulation scal-
ing factor of efficiency of the pixel veto requirement was estimated with the photons from the final-state radiation of muons in $Z^0$ decays. The scaling factor was measured to be $0.996 \pm 0.013$ for the barrel and $0.959 \pm 0.062$ for the endcaps.

Figure 5.3: Measured signal efficiency $\varepsilon$ in the four $|\eta|$ regions for the isolation selection criteria. Data-to-simulation scaling factors have been applied. The error bars include the systematic uncertainties.

Figure 5.3 shows the total selection efficiency for the isolation method, after taking into account all the scaling factors, as a function of photon $P_T$ in the four regions. Uncertainties on the $\varepsilon_{ID1}$, $\varepsilon_{ID2}$, and trigger efficiency $\varepsilon_{TRIG}$ were included as sources of systematic uncertainty on the final cross section measurement shown later.

5.5.2 Unfolding Detector Effects

Once the fitted signal yield has been obtained, a bin-by-bin unfolding correction factor $U$ was applied to the yield. The factor takes into account any mismeasured photon candidate $P_T$ contributions. The contributions to the mismeasured photon candidate $P_T$ come from various detector effects such as resolution, calibration or any imperfections
in the reconstruction algorithm in the detector. So the cross section in the various $P_T$ bins needed to be corrected to obtain the true values. For binned data, the distorted distribution can be linked to the true distribution by the following expression:

\[
\hat{R}T = M
\]  

(5.7)

where, $T$ is a vector of true values of an observable, while $M$ is a vector of measured values of that observable and $\hat{R}$ is a matrix forming mapping between true and measured values and is known as the response matrix. The response matrix covers the systematic shifts as well as bin migrations of the true values and is usually built by using the simulation of the detector. To recover the true distribution, the response matrix is inverted and applied on the measured observable $M$:

\[
T = \hat{R}^{-1}M
\]  

(5.8)

Since the response matrix was derived from Monte Carlo simulation, it is subject to statistical fluctuations. Several methods have been developed to minimize the effect of statistical fluctuations. The unfolding procedure for the analysis concerned in this chapter uses bin-by-bin method. For the analysis, the corrections were obtained from simulation for each $P_T - \eta$ bin, by taking the ratio of the generator to the reconstruction level photon $P_T$ spectrum. Direct photons, simulated by PYTHIA as described in Section 3.1, were used to derive the correction. The same set of selection criteria as used in this analysis were applied on photons considered to obtain the unfolding corrections.

### 5.6 Systematic Uncertainties

All experimental measurements are prone to uncertainties arising from random or systematic uncertainty sources. While random errors are statistical fluctuations, about the mean value in either direction, due to the precision limitations of the measurement de-
Systematic uncertainties are, on the contrary, inaccuracies that are consistently reproduced in the same direction. They can show up due to uncertainties associated with the nature of the measurement apparatus (calibration of the measurement device, acceptance of the detector, ...), assumptions made by the experimenter, or due to parameters of the model used to make inferences based on the observed data that themselves are not precisely known. A goal in any experiment is to reduce the magnitude of systematic errors below the size of the random errors. The text given below will describe and/or mention all the sources of uncertainties affecting the inclusive photon cross section which were investigated, with respect to the isolation method.

- An overall uncertainty of 4% was assigned on the integrated luminosity.

- The uncertainty on ECAL energy scale was estimated from $Z^0$ mass peak positions and was found to be 0.6% for the barrel and 1.5% for the endcaps region. This results in a 4% shift effect in the photon cross section.

- The value of uncertainty on the trigger efficiency was limited by the available size of the $Z^0 \rightarrow ee$ samples.

- Uncertainties from the signal extraction procedure were procured by performing pseudo-experiments. The signal and background distribution was varied in the generated pseudo-experiments according to the uncertainty on the shape parameters. The result of each pseudo-experiment was then fitted using the original fit model. The variation of the fitted yield was assigned as the systematic uncertainty.

  - **Shape Parameters:** The mis-modeling for the input signal template arises due to the imperfect representation of the fragmentation photons inside the PYTHIA6 generator, from pile up and from detector noise effects. The signal exponential tail, $p_0$, which is constrained in the fit was varied up to ±30% to get an account of the imperfect modelling of pile up events. In order to obtain the effect of mis-modeling of the non direct photons inside PYTHIA, the contribution was removed from the events and then the results compared with the
central values. Small differences were observed in the isolation distributions for the direct and nondirect photons. The uncertainty on the background shape parameters, $p_3$ and $p_4$, was influenced mainly by the size of background enriched samples that were selected within the side-band region and that were used to derive the data-to-simulation scaling factors. The differences observed between the constrained values obtained by applying and not applying the scaling factors was conservatively included as a systematic uncertainty.

— **Fitting bias:** To evaluate any possible fitting biases, tests were performed by generating pseudo-experiments for each bin and then fitting with the same fitter settings as used to fit the real data. For each experiment, $N_s$, $N_b$ and the signal and background shape parameters were floated within the statistical error extracted from the fit. The uncertainty due to imperfect fitting was obtained from these pseudo-experiments by taking the difference between the observed signal yields and the yields expected under the fitting model.

- Uncertainties from the efficiency estimation procedure: The uncertainty on the pre-selection criteria ($\varepsilon_{ID1}$) and on the $\varepsilon_{ID2}$ for the isolation method were dominated by the limited number of $Z^0 \rightarrow ee$ and $Z^0 \rightarrow \mu\mu\gamma$ events available, which were used to evaluate such kind of efficiencies. The total statistical and systematic uncertainty on these efficiencies was propagated to the final cross section result as a systematic uncertainty on each efficiency. Additionally the pile-up conditions, the background estimate, and the difference between photons and electrons observed in simulation were also observed to contribute systematic uncertainties to the result.

- The electron background from $Z \rightarrow ee$ decays was estimated from the product of the integrated luminosity, the production cross section measured in Ref. [190], and the efficiency from simulated $Z$ events multiplied by a data-to-simulation efficiency scaling factor. The contribution of electron background from $W \rightarrow ev$ decays and Drell-Yan processes was also estimated following a similar procedure. The total
contamination of the sample due to mis-identification of electrons from Z and W events was found to be less than 1% (negligible) and was taken as a systematic uncertainty.

- For the unfolding procedure, the difference in the correction factors obtained by using the $P_T$ spectrum of direct photons in PYTHIA and as against using those obtained from the NLO pQCD predictions were taken as a systematic uncertainty. The corresponding change on the cross section was estimated and can be seen in Table 5.6.

| Source                        | $|\eta^\gamma| < 0.9$ | $0.9 < |\eta^\gamma| < 1.44$ | $1.57 < |\eta^\gamma| < 2.1$ | $2.1 < |\eta^\gamma| < 2.5$ |
|-------------------------------|-------------------------|-------------------------|-------------------------|-------------------------|
| Luminosity                    | 4.0                     | 4.0                     | 4.0                     | 4.0                     |
| Energy scale                  | 4.0                     | 4.0                     | 4.0                     | 4.0                     |
| Trigger efficiency            | 0.1                     | 0.1                     | 0.7                     | 0.7                     |

<table>
<thead>
<tr>
<th>Isolation method</th>
<th>Efficiency</th>
<th>$P_T$ for $P_T = 300 - 400 \text{ GeV}$</th>
<th>Signal/background shape</th>
<th>Electron background</th>
<th>Unfolding correction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3.6 - 7.6</td>
<td>1.8 - 13</td>
<td>1.8 - 13</td>
<td>&lt; 0.1</td>
<td>2.0</td>
</tr>
<tr>
<td>Fit bias</td>
<td>0.1 - 2.9</td>
<td>0.1 - 2.8</td>
<td>4.9 - 16</td>
<td>&lt; 0.1</td>
<td>2.0</td>
</tr>
<tr>
<td>$N^\gamma$ for $P_T = 300 - 400 \text{ GeV}$</td>
<td>4.5</td>
<td>8.3</td>
<td>10</td>
<td>&lt; 0.1</td>
<td>2.0</td>
</tr>
<tr>
<td>Electron background</td>
<td>&lt; 0.1</td>
<td>&lt; 0.1</td>
<td>4.9 - 16</td>
<td>&lt; 0.1</td>
<td>2.0</td>
</tr>
<tr>
<td>Unfolding correction</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
</tr>
<tr>
<td>Total</td>
<td>3.8 - 18</td>
<td>3.9 - 35</td>
<td>8.7 - 18</td>
<td>10 - 23</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.6: Systematic uncertainties (in percent) for each source in the four $|\eta^\gamma|$ regions. The ranges, when quoted, indicate the variation over photon $P_T$.

All the systematic uncertainties have been summarized in Table 5.6. One can also see the $P_T$ dependence of each of the systematic uncertainty listed in the table in Figure 5.4. A strong $P_T$ dependence can be seen due to the uncertainty contribution associated with the signal/background shapes and this dependence was basically governed by two competing effects. Firstly, purity (uncertainty) increases (decreases) with an increase in the $P_T$ and thus leads to a reduction on the sensitivity on precise knowledge of the background template shape. Also, the template statistics decreases with increasing uncertainty and the difference between the simulated and observed ISO distributions increases with $P_T$. Due to the improvements in the LHC instantaneous luminosity and the change in the pile
up event quantity between the data analyzed for $P_T < 55 \text{ GeV}$ and above it, a transition in the systematic uncertainty pattern can be observed at 55 GeV. Additionally, due to the lack of high $P_T$ candidates that satisfy side-band selection criteria, a background-enriched sample with photon $P_T = 80 - 100 \text{ GeV}$ was used to derive the data to simulation scaling factors for all the $P_T$ bins above 80 GeV. This effect can be observed in the figures, and a discontinuity can be seen (in 3 out of the 4 $|\eta|$ regions) at around 100 GeV. Due to the large difference between the simulated and observed isolation distributions, the systematic uncertainty on the cross section for low-$P_T$ photons in the outer barrel was larger than in the central barrel.

![Systematic uncertainties](image)

**Figure 5.4:** Relative systematic uncertainties on the photon cross section measured with the photon isolation method in the four $|\eta|$ regions.

### 5.7 Results and Comparison to Theory

The results from both the conversion and the isolation methods were compared (wherever both were available) and then combined by weighting the result from each method by the corresponding uncertainty in each $P_T - \eta$ bin. The procedure is well described in
Ref. [191]. Overall, a good agreement between the results measured from the two methods was observed. The weights were obtained by inverting the covariance error matrix between the two methods. The elements of the covariance matrix were defined according to the expression: $C_{ij} = \rho_{ij} \sigma_i \sigma_j$, where $\sigma_i$ and $\sigma_j$ define the total uncertainties from the two methods respectively and $\rho_{ij}$ defines the correlation coefficient. $\rho_{ij}$ is determined according to the correlation between the sources of uncertainty from the two methods. At low values of $P_T$ the conversion method has smaller uncertainty than the isolation method, and thereby it receives a higher weight in the combination at low $P_T$. However, the situation gets reversed at higher values of $P_T$, where the isolation method receives higher weight than the conversion method.

The cross section for each $P_T - \eta$ bin were calculated according to the weighted sum from the conversion and the isolation methods and using equation 6.2. The corresponding measured values for the signal yield, efficiency, bin by bin unfolding correction and the luminosity for each $P_T - \eta$ bin, were fed into the equation to evaluate the cross section.

Figure 5.5 shows the measured cross section results from data, for each $P_T - \eta$ bin, along with the theoretical predictions obtained from the NLO pQCD calculations using JETPHOX 1.3.0 [78] [79].

In order to obtain the JETPHOX predictions, the following input settings were used. Calculations in JETPHOX were performed using the CT10 PDFs [192] and the Bourhis, Fontannaz, Guillet (BFG) set II of fragmentation functions [193]. The hadronic energy surrounding the photon in this case was required to be below 5 GeV within a cone of size $R = 0.4$ at the parton level. The generator-level isolation criterion used to extract the isolated cross section was chosen in order to match the analysis level criteria as closely as possible. The renormalization, factorization, and fragmentation scales ($\mu_R, \mu_F$ and $\mu_f$) were all set to the $P_T$ of the photon.

Systematic uncertainties due to the various choices for the parameters used in the theory calculations were derived as per the following procedure:

- The effect of the choice of the theory scales set in the calculations were estimated by
Figure 5.5: Measured isolated prompt photon differential cross sections (markers) as a function of transverse energy in the four pseudorapidity regions and the predictions from JETPHOX 1.3.0 (histograms). The error bars are the quadratic sums of statistical and systematic uncertainties on the measurements. The cross sections are scaled by the factors shown in the legend for easier viewing.

varying the three scales independently between $P_T/2$ and $2P_T$, such that the ratio of one scale to the other is at most two. The cross section predictions change from $\pm 22\%$ to $\pm 7\%$ with increasing $P_T$.

- The uncertainty on the predictions due to the choice of the PDFs was determined from the $52 + 1$ CT10 PDF sets using the Hessian method [194] [195]. The $\alpha_S$ uncertainty was estimated by measuring the difference between the CT10 PDF when $\alpha_S$ was set to 0.118 and two CT10as sets when $\alpha_S$ was set to 0.118 $\pm$ 0.001. The combined PDF and $\alpha_S$ uncertainties (added in quadrature) were found to be within 2.5 - 8.0%, 1.6 - 8.2%, 2.4 - 8.5% and 1.7 - 11% in the four $|\eta^\gamma|$ regions.

- The use of the BFG set I of fragmentation functions instead of the BFG set II yielded negligible differences in the predictions.
Figure 5.6: Ratios of the measured isolated prompt photon differential cross section to the NLO pQCD predictions from JETPHOX 1.3.0. The vertical error bars show the statistical uncertainties, while the shaded areas show the statistical and systematic uncertainties added in quadrature. The 4% luminosity uncertainty on the data was not included. The two curves show the uncertainties on the theoretical predictions due to their dependence on the $\mu_R$, $\mu_F$ and $\mu_f$ scales, and on the variation of CT10 $\alpha_S$ and PDFs. A correction to account for extra activity ($\tilde{C} = 0.975 \pm 0.006$) was applied to the theoretical predictions.

To account for the contributions from the underlying event and parton-to-hadron fragmentation, the theoretical predictions from JETPHOX were multiplied with an additional correction factor $C$. The factor was obtained by taking the ratio between the isolated fraction of the total prompt photon cross section at the hadron level and the same fraction obtained after turning off both multiple-parton interactions (MPI) and hadronization, using PYTHIA generated events. $C$ was measured to be $0.975 \pm 0.006$ from an average calculated over four different sets of parameters (Z2 [186], D6T, DWT and Perugia-0 [196]).
used in PYTHIA. The uncertainty on C was obtained by the root mean square of the results obtained with the different PYTHIA parameter sets. The correction led to a reduction in the predicted cross section values, owing to the presence of extra activity inside the isolation cone and thus resulting in some photons failing the isolation requirements.

The NLO pQCD predictions seem to agree well within uncertainties with the measured cross section from data, as shown in Figure 5.5. A more clear picture can be drawn by reproducing the same plot in terms of the ratio of data to theory shown in Figure 5.6, which also shows the uncertainty bands from both data and theory. The results also lead to an observation that for the lower $P_T$ photons in the regions $|\eta^\gamma| < 0.9$, $0.9 < |\eta^\gamma| < 1.44$, and $1.57 < |\eta^\gamma| < 2.1$, the theoretical calculated values tend to give higher cross section as compared to the cross section measured in data.