UNIVERSITÉ PARIS DIDEROT

THÈSE

en vue d'obtenir le grade de

DOCTEUR ÈS SCIENCES

présentée par

Laurent Millischer

Mesure de la section efficace de production de paires de photons isolés avec le détecteur CMS

devant être soutenue le 30 septembre 2011 devant la commission d'examen composée de

Philippe Schwemling	président
Marc Besançon	co-directeur de thèse
Federico Ferri	co-directeur de thèse
Éric Kajfasz	rapporteur
Tommaso Tabarelli de Fatis	rapporteur
Wolfgang Adam	examinateur
Éric Pilon	examinateur

Measurement of the Inclusive Production Cross Section of Prompt Photon Pairs with the CMS Detector at the LHC

Laurent Millischer

October 24, 2011

Contents

Introduction

The predictions made by the Standard Model of particle physics have been tested experimentally to a considerable degree of precision. The quest is now on, at high energy particle colliders, for the only remaining piece within the framework of the Standard Model to be discovered, namely the Higgs boson, which is responsible for the electroweak symmetry breaking.

The asymptotic freedom of Quantum Chromodynamics (QCD), the theory of strong interactions in the Standard Model, makes QCD predictions possible in perturbation theory. The comparatively high value of the strong coupling constant however is responsible for the slow apparent convergence of the perturbative series. The computation of several orders of the series is therefore required for QCD predictions to be accurate and comparable to measurements performed at particle colliders. Two orders are state-of-the-art and three in final preparation.

The mass of the Standard Model Higgs boson is known from direct searches performed at the Large Electron-Positron Collider (LEP) to be greater than 114.4 GeV/c^2 and, as of August 2011, to lie below 150 GeV/c^2. This upper limit is determined on one hand by direct searches at the Large Hadron Collider (LHC) and the Tevatron and on the other by precision fits of electroweak data. In the lower part of this mass window, searches of the Higgs boson decaying to two photons are among the the most sensitive.

Studies of isolated photon pairs at the LHC allow to test QCD in its perturbative regime at energy scales never probed before. Moreover, the measurement of the cross section of diphotons is crucial, as photon pairs produced in hard QCD interactions form the largest irreducible background to the search for a low mass Higgs boson that decays to diphotons. Even though this background can be estimated from data, a solid understanding of its kinematic properties and reliable theoretical predictions give additional handles to enhance the sensitivity of Higgs boson searches in the diphoton channel.

This thesis presents a comprehensive study of the production cross section of pairs of prompt isolated photons. An overview of the diphoton production in the Standard Model in chapter 1 is followed by a description of the LHC machine and the CMS detector, with emphasis on the electromagnetic calorimeter, in chapter 2. The reconstruction of photons and the measurement of the photon energy and momentum are described in chapter 3.

A new method of determining the number of prompt isolated photons in a sample of photon candidates, based on a data-driven fit on an electromagnetic isolation variable, is described and validated in chapter 4 before being applied in chapter 5 to the measurement of the differential cross section of single isolated photons. Finally, chapter 6 presents the first measurement of the differential cross section of isolated photon pairs performed with the CMS detector at a centre-of-mass energy of 7 TeV.

Chapter 1

Photon Pairs in the Standard Model

1.1 The Standard Model

1.1.1 Introduction

The Standard Model of particle physics is a theory describing all of the known elementary particles and their interplay via the strong, the weak and the electromagnetic forces. The effects of gravity, the fourth of the fundamental interactions, are negligible with respect to the other three and are not accounted for in the Standard Model.

Fermions

Matter is made of fermions that are classified into two categories, leptons and quarks, depending on the way they interact: all fermions couple to the weak and electromagnetic forces, yet only the quarks interact via the strong force.

Leptons are split into three generations, each one comprising a massive particle of negative unitary electric charge and a lighter, neutral one, the associated neutrino. The electron and the electronic neutrino for instance make up the first generation as can be seen in table 1.1.

Generation	Symbol	Mass	Charge	Symbol	Mass	Charge
electron	e	0.5110 MeV	-1	ν_e	< 2 eV	0
muon	μ	105.7 MeV	-1	ν_μ	< 2 eV	0
tau	τ	1777 MeV	-1	ν_τ	< 2 eV	0

Table 1.1: Leptons [1]

Quarks are also split in three generations of two particles. Although the number of lepton-generations of the above kind is found experimentally [2] to be set, the standard model does not exclude the existence of further quark generations. The particles in the left-hand column of table 1.2 have an electric charge of $\frac{2}{3}$, whereas the charge of the right-hand ones is $-\frac{1}{3}$. Thus, the proton being made of two up quarks and a down quark (`uud`) has a positive unit charge, whilst the neutron (`udd`) is neutral.

Name	Symbol	Mass	Charge	Name	Symbol	Mass	Charge
up	u	$1.5 - 3.3$ MeV/c^2	2/3	down	d	$3.5 - 6.0$ MeV/c^2	$-1/3$
charm	c	$1.27\ ^{+0.07}_{-0.11}$ GeV/c^2	2/3	strange	s	$104\ ^{+26}_{-34}$ MeV/c^2	$-1/3$
top	t	171.2 ± 2.1 GeV/c^2	2/3	bottom	b	$4.20\ ^{+0.17}_{-0.07}$ GeV/c^2	$-1/3$

Table 1.2: Quarks [3]

Interactions

In the Standard Model, the dynamics and interactions of particles are determined by the Lagrangian density function, or *Lagrangian*, similar to that of classical mechanics in that the equations of motion of a system derive from it via the action principle. It states that a system evolving from one state to another does so by a path corresponding to a stationary action, defined as the spatial integral of the Lagrangian.

Particles in the Standard Model are described as quantised excitations of fields. The dynamics of fermions of mass m, excitations of the field $\psi(x)$, are determined by the Dirac Lagrangian [4]. Fermion-fermion interactions are introduced in the theory via the gauge-invariance condition, that is requiring the Lagrangian to be invariant under a symmetry group called gauge group of which examples are given in table 1.3. This invariance condition gives rise to new fields $A_\mu(x)$ [5] coupling to the fermion fields $\psi(x)$ via a scalar coupling constant g in terms of the Lagrangian.

These terms illustrate the meaning of the word interaction. Indeed, as the integral of the Lagrangian is expressed in units of energy, terms involving the fermion and the gauge boson fields show that energy exchange is possible among the fields $\psi(x)$ and $A_\mu(x)$, in other words that fermions can interact via gauge bosons. The coupling constant g then determines the strength of the interaction.

Symmetries in the Standard Model

In the context of the Standard Model, the three fundamental interactions correspond to invariances under the symmetry groups $SU(3)$ for the strong force and $SU(2) \times U(1)$ for both the weak and the electromagnetic force as presented in table 1.3.

Symmetry	Group	Fields	Dim.	Gauge bosons
Strong	$SU(3)$	g_a^μ a=1..8	8	8 gluons
Electroweak	$SU(2) \times U(1)$	K_a^μ a=1..3 and L^μ	4	2 W, 1 Z et 1 photon

Table 1.3: Gauge groups of the Standard Model. The dimension of the gauge group equals the number of gauge bosons: eight in the case of the strong, four in the case of the electroweak symmetries.

In the previous paragraph it was stated that there are particles originating from the gauge invariance. These particles are excitations of the gauge fields $A_\mu(x)$ and thus are called gauge bosons, summarised in table 1.4.

Interaction	Gauge boson	Multiplicity	Symbol	Mass	Charge
Strong	gluons	8	g	0	0
Weak	W boson	2	$W^{\pm}$	80.398 ± 0.023 GeV/c^2	± 1
	Z boson	1	Z^0	91.1876 ± 0.0021 GeV/c^2	0
Electromagnetic	photon	1	γ	0	0

Table 1.4: Vector bosons [6]. The weak and electromagnetic interactions are grouped together because they correspond to the same gauge group, as will be shown in section 1.1.2.

Particles and Fields

Not all gauge groups act on all fields. The gauge group of the strong interaction acts only on quark and on gluons. The group $SU(2)$ acts on $SU(2)$ doublets via the matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.1}$$

Among the fermions, only left-handed fields are doublets and the group is therefore noted $SU(2)_L$. In the classical version of the Standard Model, neutrinos are considered as being massless and are left-handed. The $U(1)$ symmetry group acts only on fields with non-zero *hypercharge* Y, and is therefore noted $U(1)_Y$.

With only the fields $\psi(x)$ and $A_\mu(x)$, it is not possible to give mass to the fields, by including mass terms of the following kind

$$-m^2 A_\mu(x) A^\mu(x) \tag{1.2}$$

to the Lagrangian, that are not gauge invariant. Yet the experimental data presented in tables 1.1, 1.2 and 1.4 shows that fermions and some gauge bosons have masses.

1.1.2 Electroweak Symmetry Breaking

The Higgs Field

Gluons are found to be massless and the Lagrangian is symmetric under $SU(3)$. The $W^{\pm}$ and Z^0 gauge bosons however, associated with the $SU(2) \times U(1)$ symmetry, are massive, therefore the Standard Model Lagrangian cannot be symmetric under this particular gauge group and the symmetry is said to be broken.

The Higgs mechanism allows to write mass terms for the W and Z bosons and give a longitudinal polarisation state. Several ways of breaking the $SU(2) \times U(1)$ symmetry are possible. In the Standard Model, the electroweak symmetry is broken spontaneously via the introduction of an iso-doublet of scalar fields having a non-vanishing vacuum expectation value, the signature of which is a spin 0 particle, the famous Higgs Boson [7].

Higgs Mass

As of August 2011, the scalar Higgs boson predicted by the Standard Model in the context of the spontaneous symmetry breaking has not been discovered. Higgs boson searches are among the priorities of the Large Hadron Collider described in section 2.1.

From the expressions of the vector boson masses given in equation (**??**) and the experimental values of table 1.4 one gathers that

$$\frac{g'}{\sqrt{g^2 + g'^2}} \approx 0.23 \tag{1.3}$$

The coupling constants g and g' can be determined experimentally and allow to compute the Higgs field vacuum expectation value $v \approx 250$ GeV. The only unknown parameter left is the Higgs boson mass m_H. Its value is not predicted by the model model and has to be determined experimentally. Before any results from the LEP, the Tevatron and the LHC experiments, a region from 15 MeV upwards was considered [8] but as of 2011, a lower bound [9] at 114.4 GeV/c^2 and further exclusion regions [10, 11] were established. From fits to electroweak precision data, such as one can deduce an upper bound to the Higgs mass that goes as far down as 153 GeV/c^2 [12].

The Higgs boson mass m_H is submitted to corrections coming from loop diagrams [13] as shown in figure 1.1.

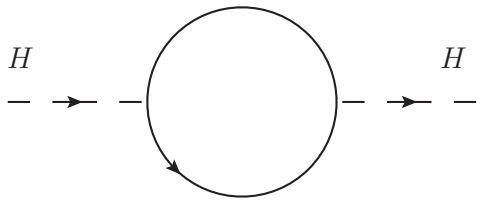


Figure 1.1: Generic loop correction to the Higgs boson mass

These correction terms to the Higgs boson mass are be proportional to Λ^2 where Λ is the scale up to which the Standard Model is valid [14]. Assuming it breaks down at the Planck mass, that is the scale where the gravitational interaction becomes comparable to the strong, weak and electromagnetic interactions, $\Lambda = M_{\text{Planck}} = O(10^{19})$ GeV, the terms multiplying the correction terms would have to cancel out to 10^{-34} in order for the Higgs mass to stay below 1 TeV, which is what one expects. This appears to be highly unlikely and is known as the hierarchy problem.

Fermion Masses

Finally, the introduction of the Higgs field also allows to account for fermion masses, by adding mass terms to the Lagrangian such as

$$-m_e^2 \bar{e}e \tag{1.4}$$

where $m_e^2 = \lambda_e v$ and λ_e is the *Yukawa* coupling, determining the strength with which the Higgs field couples to the fermion.

1.1.3 Fields and Couplings

The Standard Model Lagrangian defines the dynamics and interactions of the above mentioned particles. The cross section of a given process, defined by an initial and a final state, can be computed via graphical representations of the process, the Feynman diagrams [15].

From sections 1.1.1 and 1.1.2 one gathers:

- the gluon, the gauge boson of the strong interaction, couples to gluons and quarks,
- the photon, the gauge boson of the remaining electroweak symmetry $U(1)_Q$, couples to all electrically charged particles, that is particles with $Q \neq 0$, and

- the Higgs boson couples to all massive particles: quarks, leptons and the $W^{\pm}$ and Z^0 vector bosons.

Running couplings

In a similar way the mass of a particle receives corrections from loop diagrams, as the one shown on figure 1.1, an interaction vertex does too. Computing the higher order corrections to the vertices leads to divergences. In order to remove the diverging terms in the obtained expression, one introduces a mass scale μ_R [16] called the *renormalisation scale*.

When taking into account the corrections, one obtains a new expression of the coupling constant, depending on the renormalisation scale, that is the mass scale of the considered physical process. The dependence of the new coupling on μ_R is quantified by the *beta-function*

$$\beta(\alpha) = \frac{\partial \alpha}{\partial \ln \mu_R} \tag{1.5}$$

Taking into account only the first set of corrections, that is the corrections that involve only one loop or two additional vertices, one obtains the following expressions of the beta functions for the photon-fermion interactions

$$\beta(\alpha) = \frac{2\alpha^2}{3\pi} \tag{1.6}$$

and the quark-gluon interaction

$$\beta(\alpha_s) = \left(\frac{2N_f}{3} - 11\right)\frac{\alpha_s^2}{2\pi} \ . \tag{1.7}$$

where $N_f < 6$ is the number of quarks flavours contributing to the correction loop. We note the negative sign of the above quark-gluon beta function. Both the electromagnetic and the strong coupling constants have been measured experimentally at certain mass scales and examples of those measurements are:

$$\alpha(m_e) = 7.2973 \times 10^{-3} \qquad \text{from [17]} \tag{1.8}$$

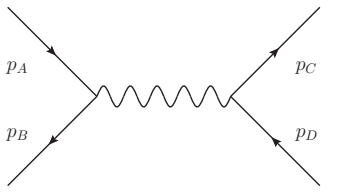
$$\alpha_s(m_{Z^0}) = 0.1184 \pm 0.0007 \qquad \text{from [18]} \tag{1.9}$$

Quantum Chromodynamics

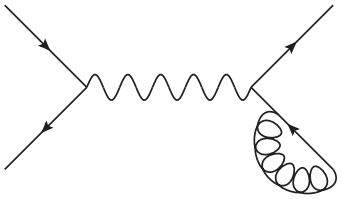
The condition $\alpha \ll 1$ is important in order to compute cross sections in *perturbation theory*, that is by successive refinement of the computation, starting with the simplest amplitude and adding *higher order* terms, with greater powers of α, corresponding to more complex Feynman diagrams. An example of such a higher order diagram, a correction to the one shown in figure 1.2(a) can be seen on figure 1.2(b). As opposed to *quantum electrodynamics*, the theory describing the interaction of photons and fermions, for which perturbation theory works well, *quantum chromodynamics* (QCD), the theory describing the interactions of quarks and gluons, is subject to two caveats.

The first resides in the rapid increase of α_s at low mass scales, due to the negative sign of the one-loop QCD beta function (1.7), making perturbative QCD (pQCD) predictions impossible for $\alpha_s \gtrsim 1$. The decrease of α_s at high mass scales however, called *asymptotic freedom* [19], allows for cross section predictions to be made at mass scales probed by the LHC.

The second caveat lies in the comparatively high values of α_s even at those mass scales (1.9) leading to significant contributions of orders beyond the lowest one in pQCD to the predicted cross sections.



(a) Tree-level



(b) One-loop correction

Figure 1.2: Generic $AB \to CD$ scattering process shown at tree-level (a). A one-loop QCD correction diagram is shown on figure (b).

1.2 Photon Pair Production

In this section, we describe the production mechanisms of photon pairs, called *diphotons*, within the Standard Model. We shed light on the insights gained by a study of diphoton production, give the outline of the cross-section prediction with the tools of pQCD and list the sources of uncertainties of this prediction.

1.2.1 Diphoton Production Processes

From the couplings described in section 1.1.3 one can derive the processes leading to the production of photon pairs in the parton-parton interaction, as opposed to photon pairs produced in the decay of secondary particles such as the π^0 or the η mesons. In the framework of pQCD, we show the leading-order diagram of each process, corresponding to the amplitude with lowest order in α_s and examples of higher-order corrections.

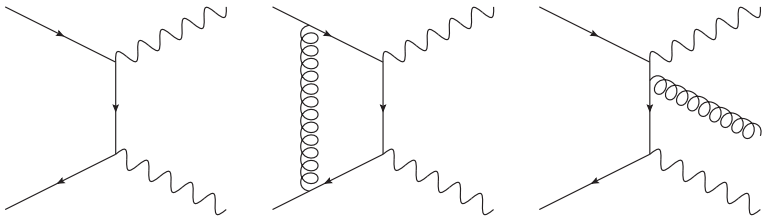
In chapter 6 we measure the *inclusive* diphoton production cross section, that is the production rate of events with at least one photon pair plus possibly other particles. Diphotons produced along other objects, such as jets, will contribute to the inclusive cross section. Two sets of higher-order corrections to the leading-order diagrams have therefore to be taken into account, those resulting in the emission of further particles, called *real corrections*, and those preserving the final state, called *virtual corrections*.

Two kinds of processes contribute to the diphoton production cross section.

Direct photons

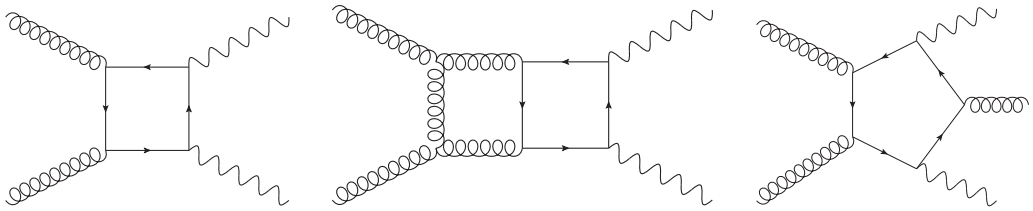
Direct photons are produced in the hard interaction at high energy scales [20], an example of which is the simplest photon pair production diagram shown on figure 1.3. It involves the annihilation of a quark-antiquark pair and is called the *Born* process.

On top of this direct process, a photon pair can be produced from gluons via a quark loop, as shown on figure 1.4. Although the leading order box diagram involves two more vertices than the leading order Born diagram and is therefore expected to contribute less to the diphoton production cross section, the momentum distribution between the constituents of the photon, described in detail in section 1.2.2, increases the contribution of the box diagram to similar importance than the Born diagram [21].



(a) Leading order (b) Virtual correction (c) Real correction

Figure 1.3: The Born process: (a) leading order Feynman diagram and examples of (b) a virtual and (c) a real correction

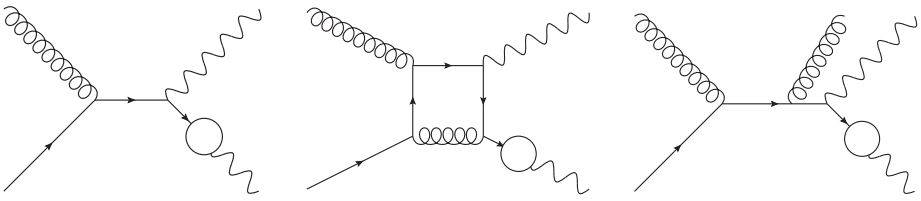


(a) Leading order (b) Virtual correction (c) Real correction

Figure 1.4: The box process: (a) leading order Feynman diagram and examples of (b) a virtual and (c) a real correction

Fragmentation photons

Further contributions to the diphoton production cross section come from the collinear fragmentation off a quark or a gluon which was produced in the hard interaction. The fragmentation occurs at a scale of the order of the typical hadronic mass [20]. As explained in section 1.2.2, the divergences appearing for collinear parton-photon final states are factorised and absorbed into so-called *fragmentation functions*.



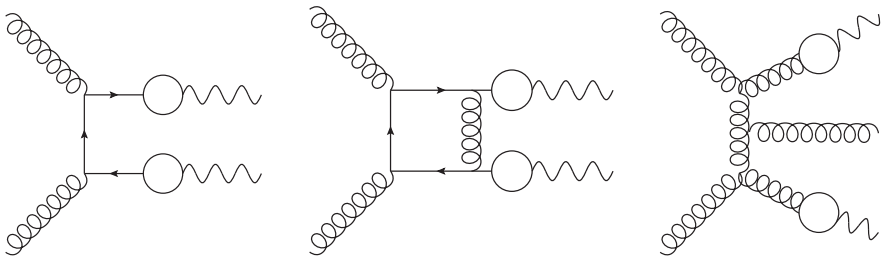
(a) Leading order (b) Virtual correction (c) Real correction

Figure 1.5: Diphoton production diagrams with one photon produced in collinear fragmentation off a parton. The (a) leading order diagram is shown along with examples of (b) virtual and (c) real correction to the first order in α_s. The white circle represents the fragmentation process.

Examples of diphoton production diagrams where either one or both of the two photons in the pair are produced via fragmentation are shown on figures 1.5 and 1.6 respectively.

Higgs boson decay photons

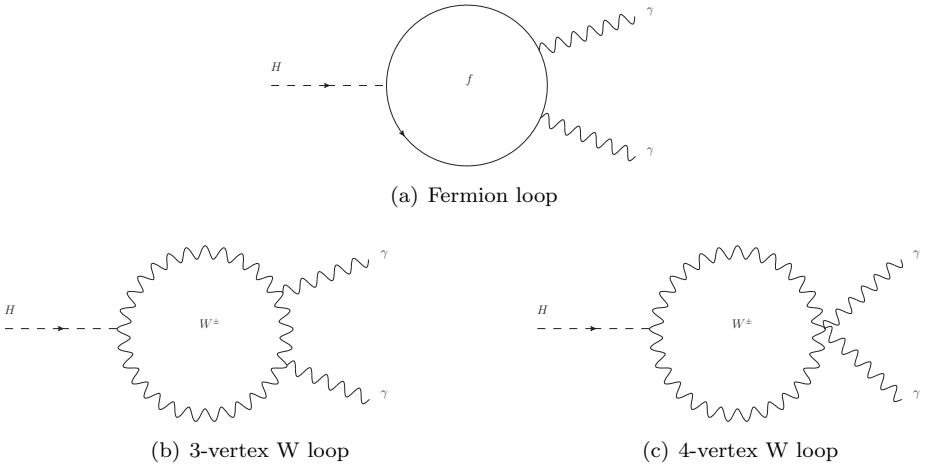
Finally, within the Standard Model as described in sections 1.1.1 and 1.1.2, the decay of the Higgs boson can produce photon pairs. As the Higgs is neutral and the photon massless, there is no possible coupling



(a) Leading order (b) Virtual correction (c) Real correction

Figure 1.6: Diphoton production diagrams with both photons produced in collinear fragmentation off a parton. The (a) leading order diagram is shown along with examples of (b) virtual and (c) real correction to the first order in α_s. The white circle represents the fragmentation process.

of the scalar to the electromagnetic vector boson. Based on the possible vertices given at the beginning of section 1.1.3, one can draw the leading order diagrams contributing to the diphoton decay channel of the Higgs boson.



(a) Fermion loop

(b) 3-vertex W loop (c) 4-vertex W loop

Figure 1.7: Leading order diagrams of the diphoton decay of the Higgs boson, which can be mediated by (a) a fermion or (b and c) a W loop.

The decay is mediated either by a fermion, figure 1.7(a), or a W loop, the latter decaying to two photons via a three-point vertex, figure 1.7(b), or a four-point vertex, figure 1.7(c). The leading-order decay rate of the $H \to \gamma\gamma$ channel can be found in chapter 2 of reference [22] as a function of the usual constants and the Higgs mass.

The production rate of the Higgs boson and its decay rates in other channels (lepton pair, quark pair, vector boson pair, loop mediated gluon pair) also depend on the Higgs boson mass, and so does the $H \to \gamma\gamma$ cross section. It appears in grey, labelled $\gamma\gamma$, on figure 1.8 at a centre-of-mass energy of 7 TeV, corresponding to the conditions of the 2010 data taking at the LHC. The $H \to \gamma\gamma$ cross section is highest at low values of the Higgs mass and always smaller than 0.05 pb.

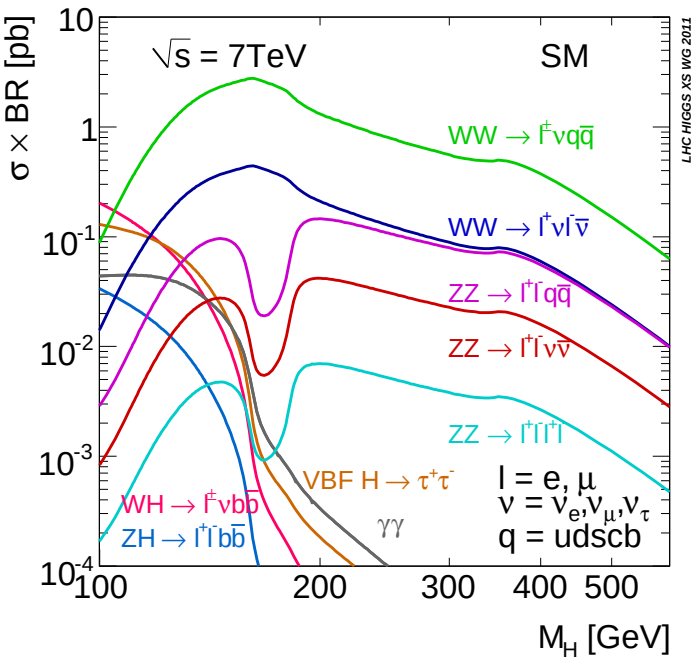


Figure 1.8: Higgs boson decay cross sections from [23] at a centre-of-mass energy of 7 TeV, corresponding to the conditions of the 2010 data taking at the LHC. The $H \to \gamma\gamma$ cross section is highest at low values of the Higgs mass and always smaller than 0.05 pb.

Isolated diphotons

Decays of the Higgs boson in the diphoton channel provide a very clear signature: two *isolated* photons, that is photons not produced along hadronic energy. Studying isolated diphotons at the LHC is a way to search for the Higgs boson at small masses but it also allows to probe the processes discussed in this section, giving valuable experimental input to validate the pQCD cross section predictions.

In diphoton measurements at the LHC, as shown in chapter 4, non-isolated photons, for instance decay products from π^0 mesons, whose production cross section is eight orders of magnitude larger than those of direct photons [24], contaminate the photon sample selected to perform the cross section measurement. In order to reduce the contamination of those photons, an *isolation criterion* is imposed, requiring the hadronic energy deposited in a cone in the rapidity-azimuthal plane around the photon to be smaller than a given threshold value.

On top of reducing the contribution of decay photons, this criterion also rejects fragmentation photons which are very likely to be accompanied by hadrons. Therefore, to compare the pQCD predictions with the measurement, a similar isolation criterion has to be applied to the calculations. It has to be noted that the transverse energy carried by partons, that carried by the hadronic particles resulting from parton hadronisation and that measured by the detector can differ. These differences will be dealt with in section 6.5.

A different definition of the photon isolation as a series of cones, was suggested in [20], and allows for a infra-red-safe treatment of the isolation in the theoretical computation.

1.2.2 Diphoton Production Cross Section

In this section we present elements of the isolated diphoton production cross section computation.

Parton distribution functions

The cross section measured at the LHC does not correspond to the hard parton-parton interaction cross section of processes such as $q\bar{q} \to \gamma\gamma$ or $gg \to \gamma\gamma$ but to the the proton-proton interaction cross section $pp \to \gamma\gamma$. The energy of a proton is divided among its constituents, the partons, namely three valance quarks and a sea of gluons and quark-antiquark pairs, this distribution is quantified by *Parton Distribution Functions* (PDF).

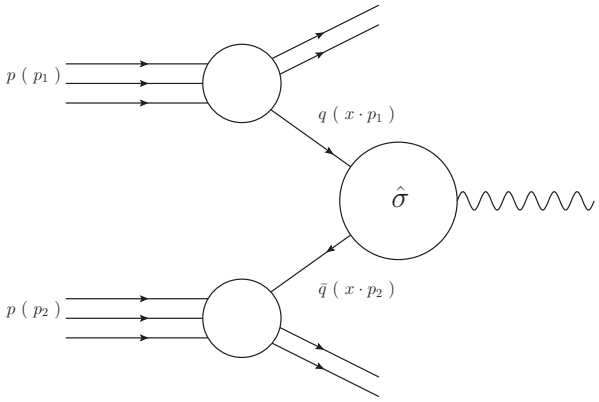


Figure 1.9: Kinematics for a generic cross section calculation at the LHC, used in equation (1.10).

For a generic hard process such as the one shown on figure 1.9, the proton-proton interaction cross section σ is related to the hard parton-parton interaction cross section $\hat{\sigma}$ via

$$\sigma = \int dx_1 \; f_{q/p}(x_1, \mu_\times^2) \int dx_2 \; f_{q/\bar{p}}(x_2, \mu_\times^2) \; \hat{\sigma}(x_1 p_1, x_2 p_2, \mu_\times^2) \tag{1.10}$$

where $f_{q/p}(x_1, \mu^2)$ is the PDF for a given quark in the proton to carry the fraction x_1 of the proton energy at a mass scale $\mu_\times$, called *factorisation scale*, at which the non-perturbative dynamics of partons within the proton factorises off the perturbative hard interaction cross section.

Parton distribution functions are extracted via a global fit on several experimental results sensitive to PDFs and their evolution with $\mu_\times^2$ is given by the DGLAP equations [26]. The PDFs represented on figure 1.10 at an energy scale $\mu_\times^2 = 400$ GeV^2 relevant for the LHC, show the predominance of the valence up-quark and the down-quark over the quarks in the sea as well as the high gluon density at small energy fraction.

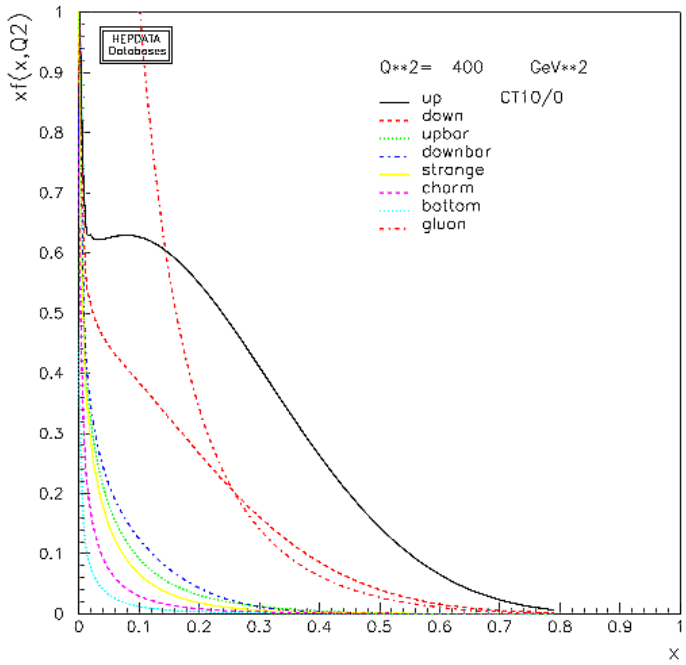


Figure 1.10: Representation of the parton distribution functions for quarks, antiquarks and gluons in a proton, at an energy scale of 20 GeV, simulated with [25]. It shows the predominance of the valence up-quark and the down-quark at high energy fraction as well as the high gluon density at small energy fraction.

Fragmentation functions

When computing the contributions of photons from the collinear fragmentation off partons to the diphoton production cross section, as explained in [24], final state collinear singularities arise in the calculation when a gluon or a quark experiences a series of successive splitting leading to a photon fragmentation.

It is possible to factorise these singularities and absorb them into quark and gluon fragmentation functions $D_{\gamma/qg}(z, \mu_F)$ defined at an arbitrary fragmentation scale μ_F and for a fraction z of the parton energy taken by the fragmenting photon. It is then possible [27] to write the contribution from fragmentation photons to the diphoton production cross section as

$$\hat{\sigma}(\text{frag}) = \sum_{i,j,k\,=\,q,\bar{q},g} \int dx_1\; dx_2\; \frac{dz}{z^2}\; f_{i/p}(x_1, \mu_\times)\; f_{j/p}(x_2, \mu_\times)\; D_{\gamma/k}(z, \mu_F) \left(\frac{\alpha_s(\mu_R)}{2\pi}\right)^2\; \hat{\sigma}_{ij}^{\;k} \qquad (1.11)$$

where i and j are the two interacting partons and $f_{ij/p}(x_2, \mu_\times)$ their parton distribution function within the proton, depending on the factorisation scale $\mu_\times$, and $\hat{\sigma}_{ij}^{\;k}$ is the production cross section of the parton k from the incoming partons i and j. The fragmentation function $D_{\gamma/k}(z, \mu_F)$ of a photon off the parton k depends on the fraction z and the framentation scale μ_F and α_s on the renormalisation scale μ_R.

It has to be noted that the distinction of diphotons into a one-fragmentation, a two-fragmentation and a direct contribution is somewhat ambiguous as only the sum of the three has no dependence on μ_F if computed to all orders and thus is the relevant physical observable [24].

Parton generator codes

Fixed order parton generator codes, such as `DIPHOX` [24] and `gamma2MC` [28], compute the cross section including all contributing diagrams up to a given order in α and α_s. Parton generator codes are not actual event generators and cannot be interfaced with a Monte-Carlo description of a detector: `DIPHOX` for instance produces some events with negative weight and keeps only the longitudinal kinematical information of the residues of the fragmentation processes.

The theoretical prediction compared to the inclusive diphoton production cross section measurement of chapter 6 was performed using `DIPHOX` to compute the contributions from quark-annihilation (figure 1.3) and fragmentation (figures 1.5 and 1.6) up to and including the order $\alpha_s\alpha^2$, accounting for the behaviour of the photon fragmentation function as α/α_s for $\mu_F \gg 1$ GeV.

As the box contribution is only described at leading order in `DIPHOX`, the gluon fusion diagram and its corrections (figure 1.4) were computed by `gamma2mc` up to and including the order $\alpha_s^3\alpha^2$. Its contributions are of comparable importance due to the very high gluon densities at the LHC, as shown on figure 1.10 despite the fact that they exceed the other diagrams by two orders of α_s.

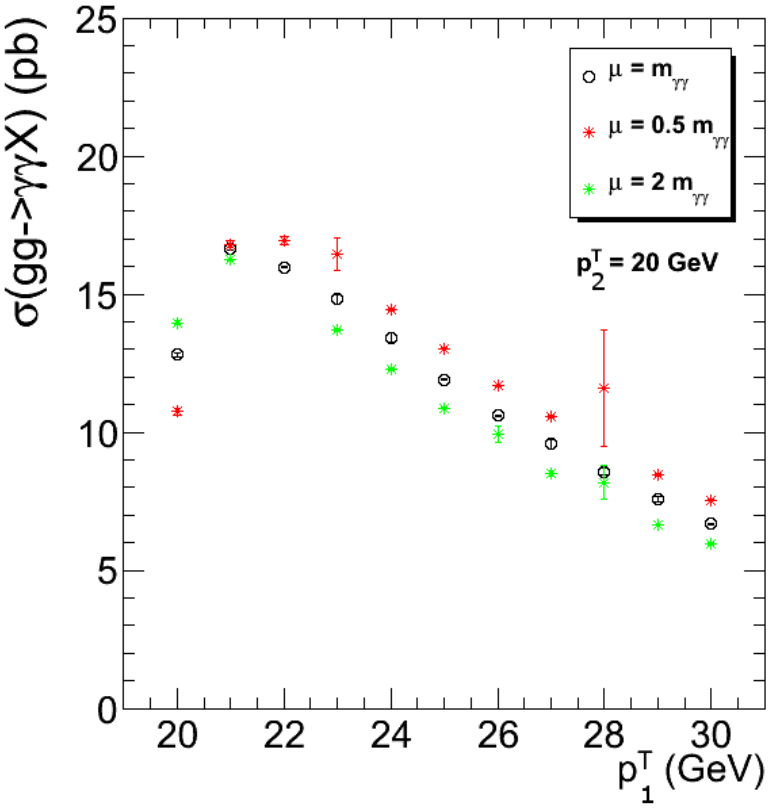


Figure 1.11: NLO contribution to the gluon fusion box diagram, figure 6.1(b), as predicted by gamma2MC versus $p_{T,1}$, the selection threshold on the leading photons, where the selection threshold on the trailing photon has been set to 20 GeV. One would expect the cross section to decrease with increasing $p_{T,1}$, which it does not for it being smaller than 23 GeV. The kinematic acceptance thresholds of table 6.2 have been chosen as a consequence of this observation. The large error bars on the predictions at 23 and 28 GeV arise as a consequence of the weights which are very high for some events.

When computing the NLO prediction to the gluon fusion box diagram with equal selection thresholds

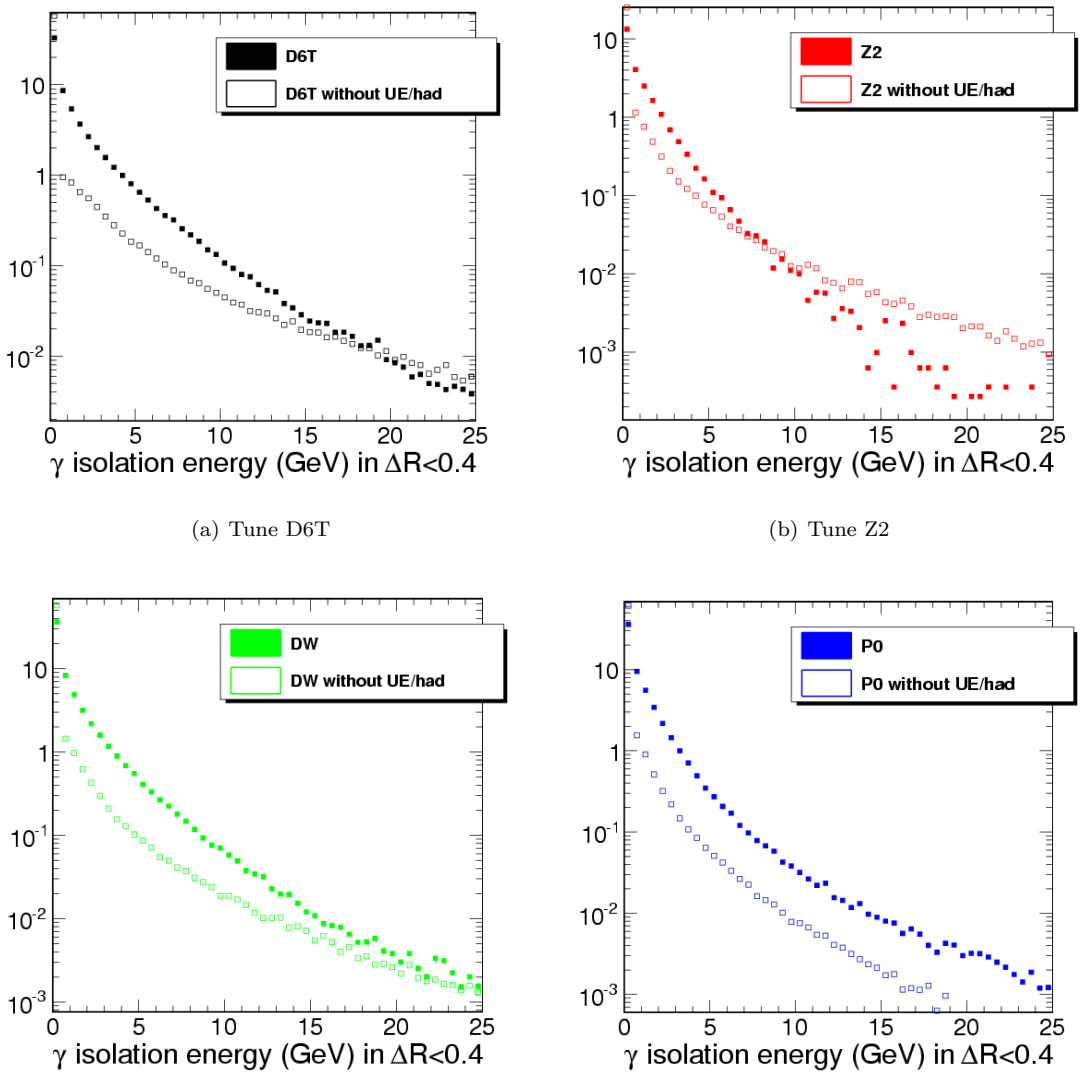
on the transverse energies of both photons of the photon pair, virtual corrections by single soft gluon radiations rarely pass the selection requirements inducing an infra-red sensitivity on the plotted differential cross section. Figure 1.11 illustrates this effect: the cross section computed with `gamma2mc` for the gluon fusion process with a set threshold of 20 GeV on the transverse energy of the sub-leading photons is expected to decrease when the threshold on the leading photon transverse energy $p_{T,1}$ increases; it does so only for $p_{T,1} > 23$ GeV. It is therefore judicious to choose asymmetric transverse energy thresholds to obtain solid predictions.

An alternative to fixed-order calculations are calculations using soft gluon resummation techniques, developed for the computation of vector boson cross sections [29] and later applied to diphoton production cross sections [30]. They are particularly appropriate in the calculation of the differential cross section as a function of the transverse momentum $p_{T,\gamma\gamma}$ of the diphoton pair where at small values of $p_{T,\gamma\gamma}$ the photon pair is accompanied by soft and collinear gluon radiation. Cross section predictions including soft gluon resummation are performed by `RESBOS` [31].

Photon isolation

As stated before, an isolation criterion, requiring the total hadronic transverse energy deposited in a cone of $R < 0.4$ around each photon to be less than 5 GeV, is included in the parton generator calculations to describe the isolation criterion imposed in the diphoton cross section measurement. Contributions to this hadronic energy can come from residues of the parton fragmentation, from a third parton in the final state or from both. The requirement imposed in the measurement however concerns the isolation computed from the transverse energy of particles resulting from parton hadronisation or being produced in underlying events (UE), hadronisation products from partons that have not taken part in the hard interaction.

In order to correct the predicted cross section for the absence of underlying event and hadronisation in the parton generator codes, the `PYTHIA` [32] parton shower generator is used in [33] to obtain the distributions, shown on figure 1.12, of the aforementioned isolation variable and deduce the efficiency on the isolation criterion with and without UE and hadronisation. A correction factor of 0.953 ± 0.003 is applied to the prediction.



(a) Tune D6T

(b) Tune Z2

(c) Tune DW

(d) Tune P0

Figure 1.12: The predictions of various PYTHIA tunes are compared independently for various isolation-energy distributions after switching on and off the underlying event and hadronization mechanisms, with numbers of events on the y-axis, from [33].

1.2.3 Diphoton Cross Section Uncertainties

When computing the production cross section of isolated photon pairs to a given order in α and α_s, limited knowledge of some of the ingredients of the computation leads to uncertainties on the predicted cross section.

PDF and α_s uncertainties

Several sets of parton distribution function are available, they differ in several respects, by

- the choice of the functional parametrisation of the PDFs for the different partons,
- the fit procedure chosen to perform the global fit, that is the choice of the floated parameters, the type of minimisation and the imposed constraints,
- the choice of the experimental results used in the global fit.

The PDF4LHC [34] working group recommends to combine three sets of PDFs: CTEQ [35], MSTW08 [36] and NNPDF2.1 [37] to determine the uncertainties on a cross section prediction computed at next-to-leading order and to combine this uncertainty in quadrature with that on α_s.

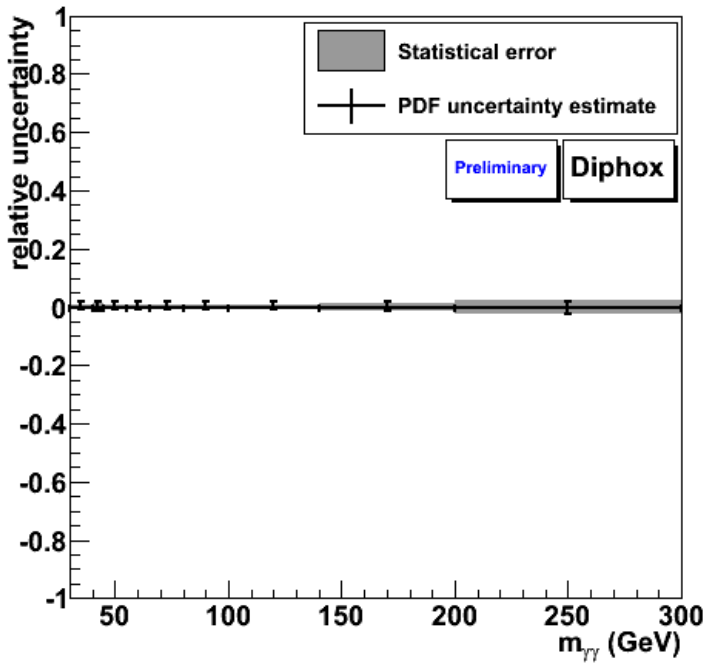


Figure 1.13: PDF uncertainties on the prediction of the invariant mass spectrum, with acceptance and binning of chapter 6, determined with NNPDF2.1.

Following the PDF4LHC working group recommendations, the PDF uncertainties are estimated by constructing the envelope provided by the central values and of the different PDF sets (CTEQ, MSTW08 and NNPDF2.1) and their errors obtained from varying α_s and the PDFs available in each set. Figures 1.13 and 1.14 show the PDF uncertainties on the prediction of the invariant mass spectrum, determined with NNPDF2.1 and CTEQ respectively and figure 1.15 the above mentioned envelope.

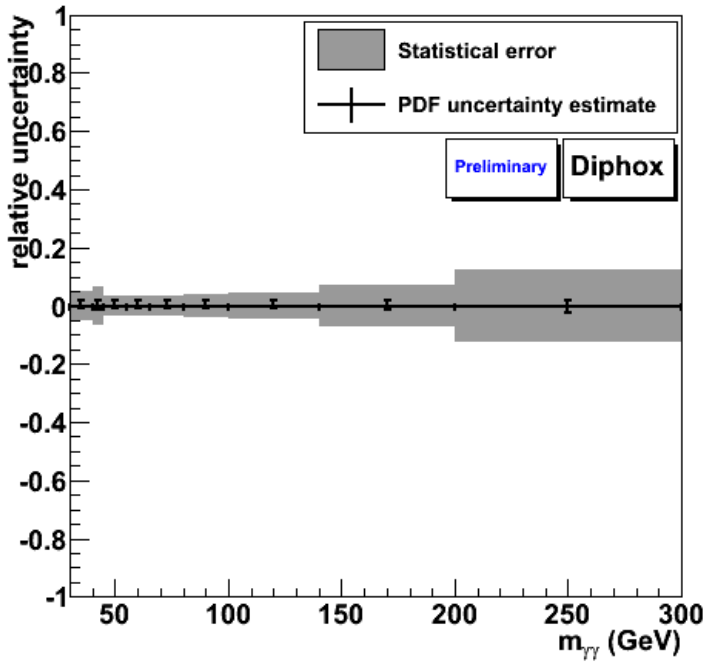


Figure 1.14: PDF uncertainties on the prediction of the invariant mass spectrum, with acceptance and binning of chapter 6, determined with CTEQ

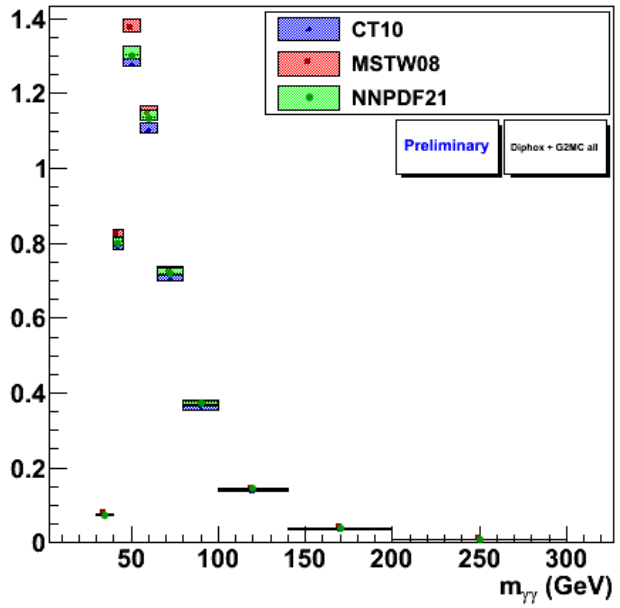


Figure 1.15: Comparison of the differential diphoton production cross-section (in pb) as a function of the diphoton invariant mass, with acceptance and binning of chapter 6, obtained with DIPHOX using the CTEQ, MSTW08 and NNPDF2.1 PDFs.

Scale uncertainties

The calculations presented in section 1.2.2 depend on three arbitrarily chosen mass scales

- μ_R the renormalisation scale, determining the value of the running strong coupling α_s
- $\mu_\times$ the factorisation scale, determining the decoupling of the hard interaction cross section and the parton distribution functions,
- μ_F the fragmentation scale, determining the decoupling of the hard parton production cross section and the fragmentation functions.

As these scales are chosen arbitrarily, any dependence of the calculated cross section on μ_R, $\mu_\times$ and μ_F has to be treated as an uncertainty of the prediction. When computing higher orders in perturbation theory, the dependence of the cross section on the scales decreases. This is shown on figure 1.16 where the scale-related uncertainty on the predicted differential inclusive Z boson production cross section versus the Z boson rapidity decreases from 30 % to 9 % and less than 1 % when moving from a leading order prediction to a prediction at next-to-next-to-leading-order (NNLO).

By convention [24], after having set a scale μ at a value relevant for the considered process, $\mu = M$, one varies the scale in the range $M/2 < \mu < 2M$ to determine the uncertainty of the predicted quantity on the scale. The differential cross section versus the invariant diphoton mass predicted in [24] at scales relevant to Higgs boson searches in the diphoton channel changes by 5 % when the scales are thus varied.

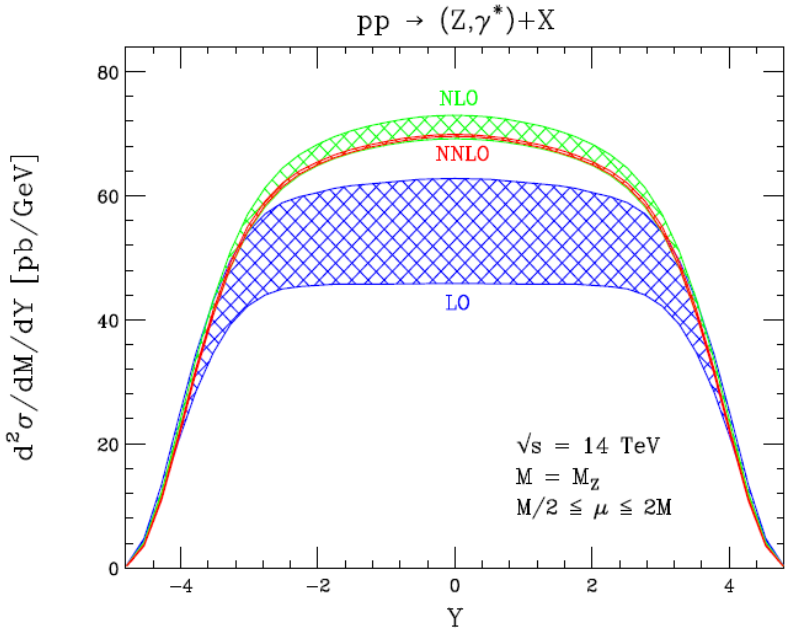


Figure 1.16: Predicted differential inclusive Z boson production cross section versus the Z boson rapidity. The bands show the uncertainty linked to the variation of the renormalisation and factorisation scales μ_R and $\mu_\times$. This uncertainty decreases with added orders in perturbation theory, from [38].

Chapter 2

Experimental Setup

2.1 The Accelerator LHC – Large Hadron Collider

The Large Hadron Collider is the last of a series of accelerators built by the European Organisation of Nuclear Research (CERN) in Geneva and, as of 2011, the highest energy particle collider in the world. It was designed to extend the efforts of the Tevatron accelerator at Fermilab National Accelerator Laboratory (FNAL) in Chicago in terms of tests of the Standard Model and probes of physics lying beyond it. With its nominal centre-of-mass energy of 14 TeV, the LHC will probe energy ranges so far unexplored: it will discover or exclude the Standard Model Higgs boson, search for supersymmetric particles and new massive gauge bosons.

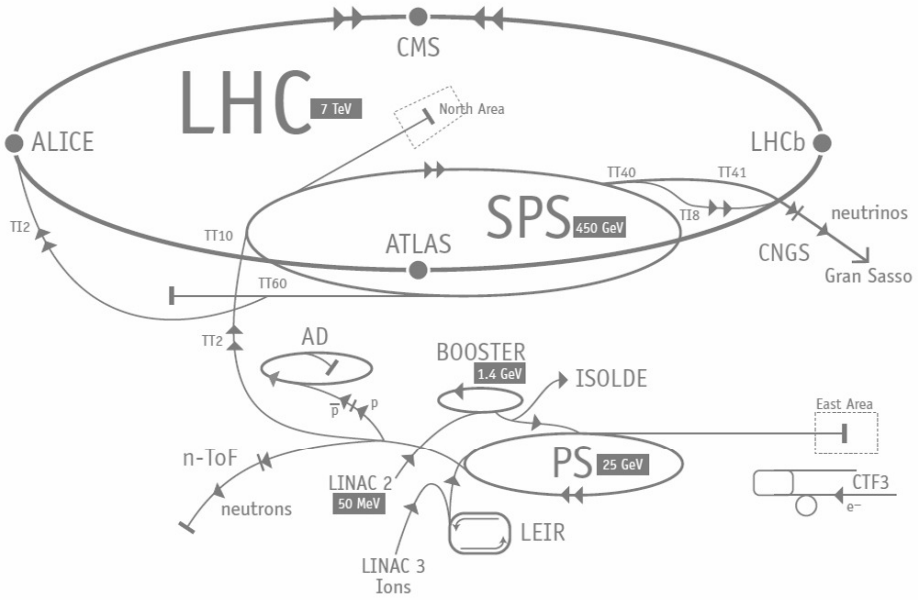


Figure 2.1: Schematic representation of the LHC accelerating structure. From the LINAC 2, trough the Booster, the PS, the SPS, the protons are accelerated from an energy of 50 MeV to an energy of 450 GeV before being injected in the LHC, where they are accelerated to reach 3.5 TeV.

2.1.1 Modus operandi

The LHC project was approved in 1994, civil engineering upgrading the existing infrastructure from the LEP collider began in 1998 and installation in 2007, after LEP was fully dismantled. Although being a project of the European Organisation for Nuclear Research holding twenty European member-states, contributions were made by many non-European countries such as Canada, India and the USA.

Figure 2.1 shows the overall structure of the accelerating structure leading to the Large Hadron Collider: two proton beams are circulated in the ring and made to collide at a nominal energy of 7 TeV per proton in the points labelled LHCb, ATLAS, ALICE and CMS, that correspond to the four major experiments on the LHC. As of 2011 protons are accelerated to 3.5 TeV and the design energy of 7 TeV will be reached in a second phase.

In order to reach the collision energy, the acceleration is performed in five steps. Starting from a bottle of hydrogen serving as a source, protons are passed on to the Linear Accelerator (LINAC 2), accelerating them from 90 keV to 50 MeV. The PS Booster, made of four stacked rings, accelerates the protons to 1.4 GeV, before transferring them to the Proton Synchrotron (PS) that splits the bunches and brings them to an energy of 25 GeV. Finally, the Super Proton Synchrotron (SPS), the last injector before the LHC, 7 km in length, accelerates the protons to 450 GeV. The LHC can serve as a proton-proton and as a heavy-ion collider, in the latter mode the acceleration chain is somewhat different, starting from LINAC 3 via the Low Energy Ion Ring (LEIR) and leading to the PS.

The LHC itself is made of different components serving different purposes. The radiofrequency (RF) cavities have a longitudinal component in phase with the beam, thus providing the acceleration up to the nominal 7 TeV per proton. Dipoles provide the vertical magnetic bending field to keep the protons on track, quadrupoles inserted after every third dipole handle the focussing of the beam and sextupoles deal with its chromatic aberrations, the fluctuations of the proton energy around the nominal value.

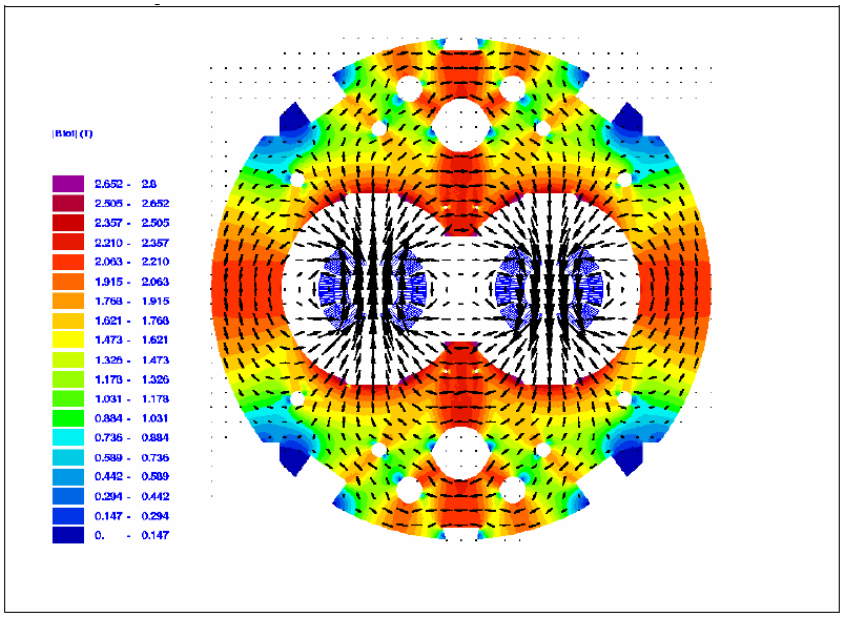


Figure 2.2: Magnetic map in a transverse cut of an LHC dipole, published in the LHC design report [39].

The LHC being a same-charge particle collider, it requires an inverted bending field for the two contrarily rotating beams. Figure 2.2 shows the design chosen to provide this field geometry. For proton-antiproton accelerators, like the Tevatron at FNAL, it is possible to use a single beam pipe to

perform the acceleration. However, the production and storage of anti-protons is challenging and leads to a serious limitation on the circulating beam intensities.

2.1.2 Centre-of-Mass Energy

The centre-of-mass energy, noted $\sqrt{s}$, the square root of the s-channel Mandelstam variable, is defined as the sum of the energies of two colliding protons. In particle colliders, it corresponds to the energy of a two-proton system in its centre of mass, hence its name. At the LHC its nominal value is 14 TeV, yet the centre-of-mass energy of the 2010 data taking was 7 TeV. The equation of motion of a centripetally accelerated proton circulating in an ideal ring yields the dependence of its momentum p on the intensity of the vertical bending field B and the bending radius R given in equation (2.1). With the parameter R being set by the already existing LEP tunnel to 2.8 km, the centre-of-mass energy depends only on the intensity B of the magnetic bending, 4.2 T in the 2010 run.

$$p\ [\mathrm{GeV/c}]\ = 0.3 \times B\ [\mathrm{T}]\ \times R\ [\mathrm{m}]\ . \tag{2.1}$$

The momentum of a proton is distributed among its constituent partons: three valance quarks and a sea of gluons and quark-antiquark pairs. The probability for a parton to carry a fraction $x \in [0, 1]$ of the proton momentum is given by the Parton Distribution Function introduced in chapter 1. Therefore, the interaction energy at parton level $\sqrt{\hat{s}}$ verifies

$$\sqrt{\hat{s}} \leq \sqrt{s} \tag{2.2}$$

reaching $\sqrt{s}$ for both partons carrying a momentum fraction $x = 1$. The centre-of-mass energy thus describes the energy range probed by an experiment. At its nominal configuration, the LHC will explore an energy range seven times larger than the Tevatron, the centre-of-mass energy of which is 1.96 TeV.

2.1.3 Luminosity

The luminosity, denoted $\mathcal{L}$ is another key parameter describing the operation of an accelerator. It is a measure of the interaction rate determining the number of events N occuring for a given process of cross section σ via the proportionality relation given in (2.3).

$$N = \int dt \mathcal{L}(t) \times \sigma \tag{2.3}$$

The integral of the instantaneous luminosity over the considered time range measures the total amount of accumulated data and, having the inverse dimension of the cross-section, is usually measured in pb^{-1}.

For two bunches of protons colliding head on with an interaction cross-section σ, the instantaneous luminosity is proportional to both the numbers of protons in each bunch (N_p), to the number of bunches crossing per second ($N_b \times f_{\mathrm{rev}}$ where N_b is the number of bunches in the ring and f_{rev} the revolution frequency), the interaction rate being given by the fraction σ/A, where A is the transverse section of the bunches.

$$\mathcal{L} = \frac{N_p^2 N_b f_{\mathrm{rev}}}{A}, \tag{2.4}$$

Assuming transverse Gaussian particle density distributions, the area A can be estimated by $4\pi\sigma_x\sigma_y$. In an equivalent formalism [40] A is expressed as $\epsilon_n\beta^*/\gamma$ where γ is the Lorentz factor, ϵ_n the normalised transverse emittance and β^* the betatron function with design values of 3.75 μm and 0.55 m respectively. If the beams collide with a non-zero crossing angle ϕ, the area A has to be divided by a geometric factor F [41].

Thus, equation (2.5) gives the expression of the instantaneous luminosity as a function of the caracteristic accelerator variables,

$$\mathcal{L} = \frac{N_p^2 N_b f_{\text{rev}} F}{4\pi\sigma_x\sigma_y} \ , \tag{2.5}$$

and table 2.1 gives the values of the above variables, determining the instantaneous luminosity, for both the nominal scenario and the the 2010 data-taking.

Variable		Unit	Nominal	2010 run
Nb. protons per bunch	N_p		1.15×10^{11}	$\sim 1.15 \times 10^{11}$
Nb. bunches in ring	N_b		2808	$\lesssim 450$
Revolution frequency	f_{rev}	[Hz]	11245	11245
Geometric factor	F		0.836	0.836
x-width	σ_x	[cm]	1.67×10^{-3}	$\sim 5.6 \times 10^{-3}$
y-width	σ_y	[cm]	1.67×10^{-3}	$\sim 5.6 \times 10^{-3}$
Instantaneous luminosity	$\mathcal{L}$	[$\text{cm}^{-2}\ \text{s}^{-1}$]	1×10^{34}	$\lesssim 2 \times 10^{32}$

Table 2.1: Caracteristic variables of the LHC determining the instantaneous luminosity. The nominal [40] values and the values for the 2010 data-taking are given.

The measurement of the luminosity in CMS exploits the relation given in equation (2.3) by measuring the rate of a process, the cross section of which is well known, such as the proton-proton total cross section [40]. The instantaneous luminosity is monitored online, using the hadronic forward (HF) calorimeters, described in section 2.2.2, and exploiting the linear dependence of the total transverse energy deposited in the HF and the number of interactions. The results thus obtained are cross checked offline.

The HF method yields the relative variation of the luminosity. To access the absolute figures of the luminosity, a calibration is required which is either performed by using simulated events [42] or dedicated Van der Meer scans [43] that consist of measuring the transverse beam widths σ_x and σ_y appearing in equation (2.5) by means of fast beam current transformers.

The instantaneous and integrated luminosities over the data-taking period of 2010 are shown on figures 2.4 and 2.3 respectively. The systematic uncertainties on the luminosity measurement come from beam-halo backgrounds and uncertainties on the beam current measurement. The final uncertainty on the integrated luminosity is 4 % [44].

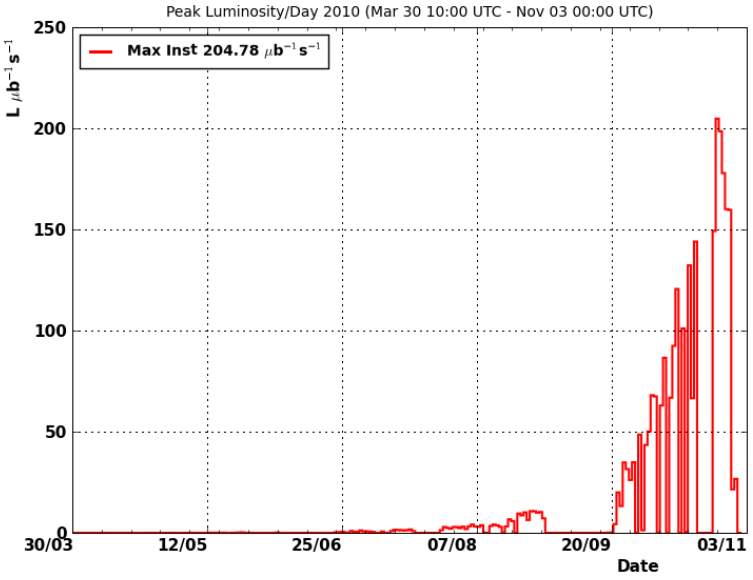


Figure 2.3: Evolution of the instantaneous luminosity during the 2010 run, reaching $2 \times 10^{32}\text{cm}^2\text{s}^{-1}$ towards the end of the run, from [45]. The characteristic parameters of the LHC yielding this instantaneous luminosity are given in table 2.1.

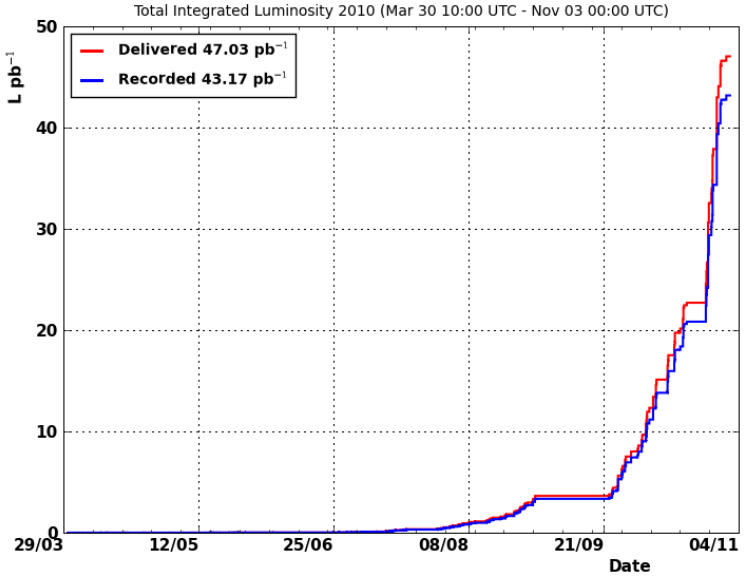


Figure 2.4: Evolution of the integrated luminosity during the 2010 run, from [45]. About 43 pb^{-1} were recorded by CMS.

2.2 The detector CMS – Compact Muon Sollenoid

The CMS experiment, shown on figure 2.5, is a general purpose detector and one of the four major experiments at the LHC, along with ATLAS, the other general purpose detector, ALICE, specialised in heavy-ion collisions and LHCb, focussing on b physics.

2.2.1 Introduction

The general architecture of the CMS detector is cylindrical and follows that of its eponymous superconducting solenoid, about 13 m in length and 6 m in diameter, which lies at the heart of its design. It generates within its bore a constant magnetic field of 3.8 T along the axis of the beamline in order to achieve good momentum resolution. All specialised subdetectors and appearing on the transverse cut of the detector shown on figure 2.7, follow the CMS architecture with a cylindrical central barrel part and disk-shaped endcap parts aimed at achieving a maximal hermeticity.

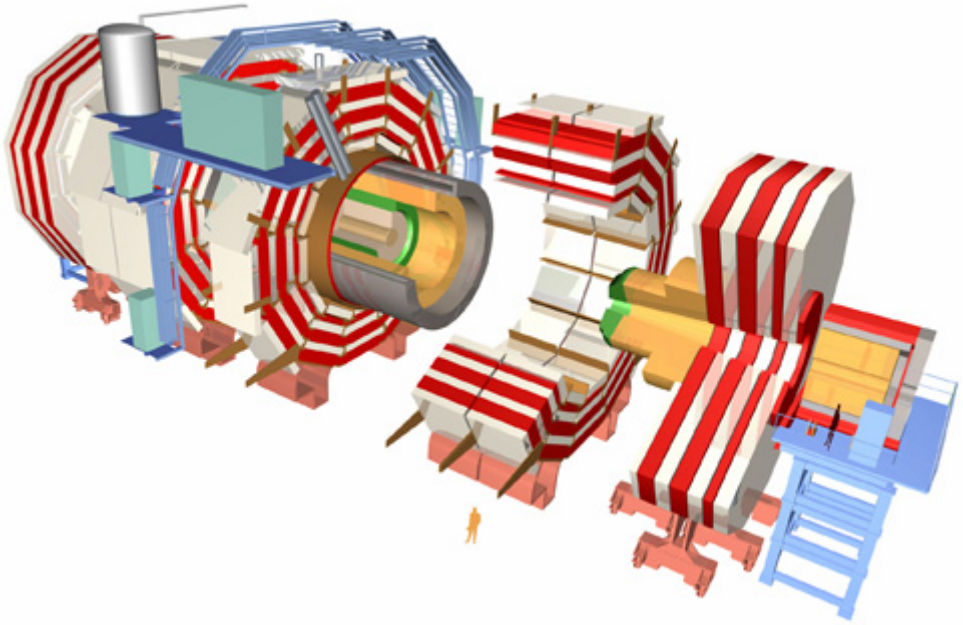


Figure 2.5: General view of CMS. The detector has a cylindrical structure around the axis of the beam pipe. The return yoke of the magnet and the muon system are shown in alternating layers of red and white respectively, the superconducting solenoid generating the magnetic field in grey, the hadronic (`HCAL`) and electromagnetic (`ECAL`) calorimeters in yellow and green, and closest to the beamline, the tracker in fawn.

For the CMS reference frame, one adopts the following coordinates: from the origin, the centre of the detector, the x-axis points horizontally towards the centre of the LHC, the y-axis upwards and the z-axis along the beam line, eastwards. Two angular coordinates are commonly used: the azimuthal angle φ measured from the x-axis in the x-y plane and the pseudorapidity η defined via the polar angle θ measured from the z-axis as follows

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right). \tag{2.6}$$

Photon plus jet events act as an major background in the measurement of the diphoton production cross section and all the subdetectors playing major a role in this measurement are required to detect and identify the particles produced in such an event. An event display from a photon plus jet event simulated with PYTHIA is shown on figure 2.6, where the photon and the jet appear on the right and left hand side of the display respectively.

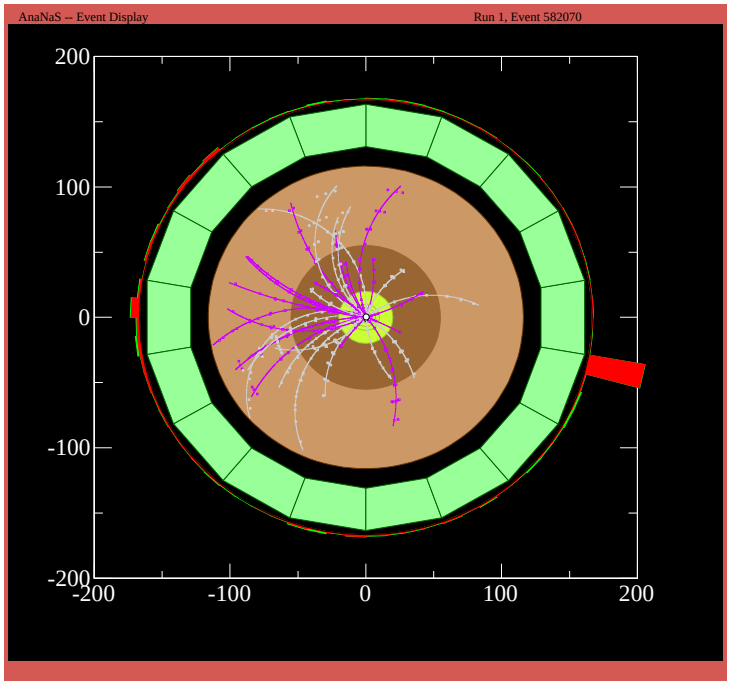


Figure 2.6: Event display of a photon plus jet event simulated with PYTHIA. Deposits in ECAL appear in red, the deposits of the photon can be clearly seen on the right hand side. Tracks are shown in grey and purple and deposits in HCAL in green. The jet appears on the left hand side of the event display.

2.2.2 Subsystems

Figure 2.7 shows the five kinds of stable particles detected by CMS: muons, electrons, photons, stable hadrons, charged and neutral. Among those, the particles produced in a photon plus jet event, photons, electrons and hadrons, are reconstructed essentially by three subdetectors: the tracking system, the electromagnetic calorimeter (ECAL) and the hadron calorimeter (HCAL). Among stable Standard Model particles, neutrinos are the only ones to go undetected. Their signature therefore is missing energy.

The Tracking System The tracking system, described in detail in reference [46], is designed to detect the tracks of charged particles and to measure their transverse momentum from their bending angle. It lies within a radius ρ of 110 cm around the beam axis. Its cylindrical barrel and ring-shaped endcap parts cover a pseudorapidity range of $|\eta| < 2.5$. As the particle flux decreases with increasing ρ [47], the tracking system consists of three layers, the granularity decreases with increasing ρ.

- up to $\rho = 10$ cm the pixel detector, made of three layers and 66 million pixels of 100×150 μm^2 is used for the precise vertex localisation.

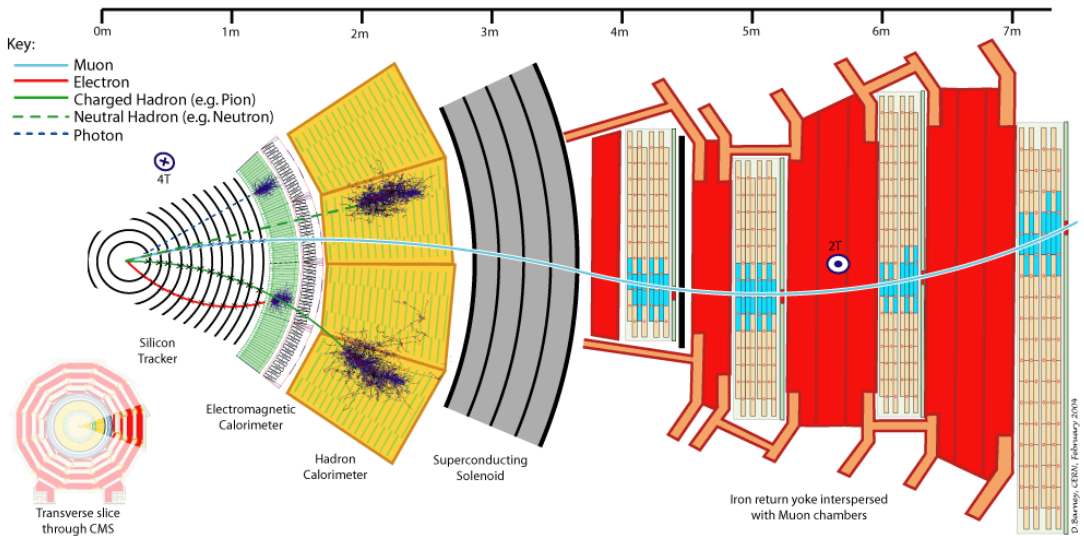


Figure 2.7: A slice of the transverse cut of the CMS detector, showing the trajectory of a **photon** (dashed blue) interacting in `ECAL` only, a **neutral hadron** (dashed green) depositing its energy in `HCAL`, a **muon** (continuous teal), interacting in the tracker and the muon system, a **charged hadron** (continuous green) interacting with the tracker before depositing its energy in `HCAL` and an **electron** (continuous red) interacting in the tracker and `ECAL`.

- For 20 cm $< \rho <$ 55 cm the inner silicon tracker with strips of 10 cm $\times$ 80 μm and
- for 55 cm $< \rho <$ 110 cm the outer silicon tracker, with slightly bigger strips, completes the system.

The p_T resolution of the tracking system worsens with $|\eta|$ and p_T and for a muon of 100 GeV at η=0 is smaller than 2% as shown on figure 2.8. The resolution for the longitudinal and transverse impact parameters δ_Z and δ_T worsens with $|\eta|$ and improves with p_T, for a 100 GeV muon at η=0 it is { 50 μm, 10 μm}. The track reconstruction efficiency improves with p_T and varies with η. Figure 2.9 shows the dependence of the track reconstruction efficiency for simulated charged hadrons on the pseudorapidity.

Electromagnetic Calorimeter The electromagnetic calorimeter (`ECAL`) detects photons and electrons and measures their energy and momentum. Made of crystals of lead tungstate, its pseudorapidity coverage reaches $|\eta| < 3$. Playing a crucial part in the measurement of the photon pair production cross section, it is described at length in section 2.3.

In a photon plus jet event, the photon leg deposits its energy in `ECAL` and the electromagnetic part of the jet leg will also be measured by this subdetector.

Hadron Calorimeter The hadron calorimeter, described in detail in reference [49], is designed to detect stable hadrons, charged and neutral, and to measure their energy and position. It follows the cylindrical geometry of CMS, thereby maximising the hermeticity for an optimal measurement of the missing energy.

Its structure is of alternating absorbing brass or steel and scintillating plastic material. The pseudorapidity coverage is performed by a barrel ($|\eta|$ <1.4) and two endcaps (1.5< $|\eta|$ <3.0) and completed by an outer part ($|\eta|$ <1.26) lying beyond the magnet coil, made of scintillators and pushing the depth of the hadron calorimeter in the central pseudorapidity region to over 10 radiation lengths. The scintillator layers consist of tiles of $\Delta\eta \times \Delta\varphi = 0.087 \times 0.087$ in size in the barrel, and increasing size (up to

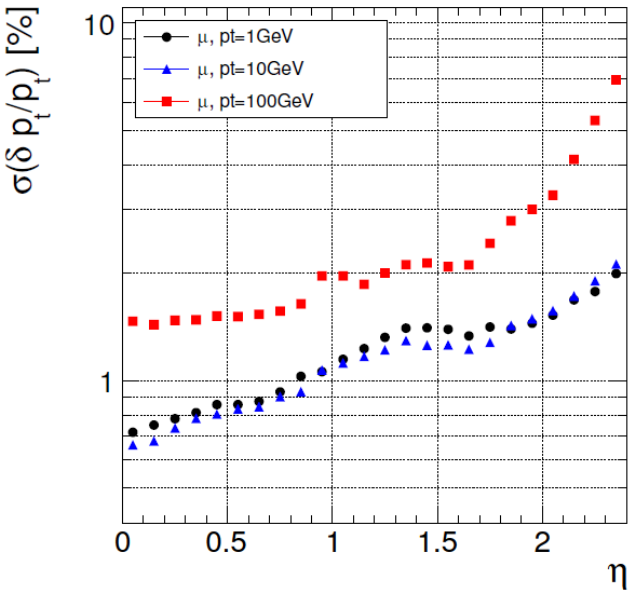


Figure 2.8: Resolution of the transverse momentum of tracks of simulated muons, from [48]

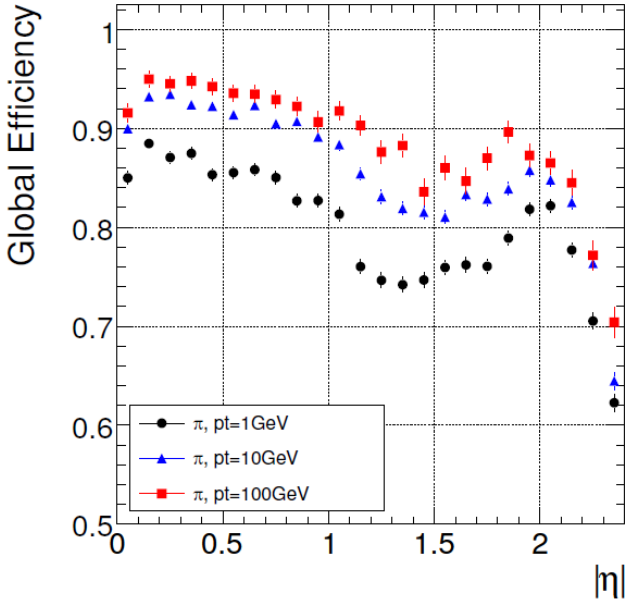


Figure 2.9: Global efficiency of the tracker for an example of charged hadrons: $\pi^{\pm}$, defined as the probability to reconstruct the track for a particle with $p_T > 0.9$ GeV/c and a production vertex inside the beam pipe. The efficiency, computed with simulated events, dips at the connection of the half cylinders making up the tracker barrel ($|\eta| = 0$) and in the barrel-endcap transition region ($|\eta| \simeq 1.4$), and drops when approaching the end of the pseudorapidity coverage ($|\eta| \simeq 2.1$).

$\Delta\eta \times \Delta\varphi = 0.350 \times 0.174$) in the endcap. The tiles of different layers are ordered projectively on top of each other, thus forming HCAL *towers*, corresponding to matrices of 5×5 crystals of the ECAL.

The forward hadron calorimeter (HF), covering a pseudorapidity range of $2.9 < |\eta| < 5.0$, measures the energetic forward objects, increases the hermeticity for a better measurement of the missing energy and is used in the measurement of the instantaneous luminosity as explained in section 2.1.3. It is situated 11 meters from the interaction point

Muon system The muon system measuring direction and energy of the produced muons. Located in the red iron magnet return yokes of the CMS detector, it is split in a barrel and two endcaps. Its p_T resolution deteriorates with increasing $|\eta|$ and muon p_T and is about 1% at null $|\eta|$ for a muon of 50 GeV.

2.2.3 Data Acquisition Flow

The CMS Trigger and Data Acquisition System (TriDAS) is designed to select *events* for archiving and later analysis based on the information gathered by the CMS subdetectors, delivered at a bunch crossing frequency of 40 MHz. It is described in detail in references [50] and [51].

Trigger Starting from a nominal interaction rate of 40 MHz, the event rate has to be drastically reduced to a rate of about 100 Hz before being saved and analysed. This reduction being too important to be performed by one system while keeping a high efficiency for those physical phenomena that are of interest for the analysis, it is split in two stages.

- The level one trigger (L1) has to reduce the data rate by more than two orders of magnitude to a design value of 100 kHz. Implemented entirely in the hardware, the L1 trigger is based on so called *trigger primitives*, objects reconstructed with low granularity in the ECAL, the HCAL and the muon system. In the ECAL, the L1 is triggered in the presence of trigger primitives above certain E_T threshold. The L1 trigger decision is reached in 3.2 μs while the data is buffered.
- The high level trigger (HLT) reduces the dataflow by another three orders of magnitude to 100 Hz. As it has to sustain an average event input rate of 100 kHz, the strong timing constraints put on the selection algorithm are met thanks to partial event reconstruction. The HLT is implemented in software and reaches its decision by taking into account information about tracks and particle isolation. After this step, the event – about 1.5 MB for a proton-proton collision – can be saved.

Data flow After the L1 acceptance of an event, the information of the 75 million channels is bundled by 600 Front End Driver boards (FEDs) and handed to the event builder for a coherent assembly of the full event. After acceptance by the HLT, the event is stored to be reconstruced offline, its raw data being built to *physics objects*, such as muons, jets, photons and electrons. This is performed in two steps.

- Local reconstruction within every subdetector. The tracking system agregates hits in the pixel and silicon detectors to track-objects, the muon system builds muon candidate objects from hits in the muon chambers. As explained in detail in chapter 3, hits in the ECAL are aggregated to superclusters.
- Global reconstruction combining the information from local reconstruction. The track and muon objects are combined into global muon objects. In a similar way, track and supercluster objects are matched to form electrons.

A gamma plus jet event sees its jet leg reconstructed with a combination of information from the tracking system and the calorimeters. The photon leg energy and impact position in ECAL are measured in the electromagnetic calorimeter.

2.3 The Subdetector ECAL – Electromagnetic Calorimeter

The electromagnetic calorimeter (ECAL) is located between the tracking system and the hadron calorimeter, its aim is to measure the energy and momentum of photons and electrons with high accuracy. A complete description of its original design, arrangement and performance can be found in reference [52] and its performance with the first collision events in [53].

2.3.1 Architecture

The ECAL, a longitudinal cut of which is shown on figure 2.10, consists of a barrel part (EB) the inner surface of which is at a radius of $\rho = 129$ cm, covering a pseudorapidity range of $|\eta| < 1.479$, two endcap parts (EE) at a distance of $z = 314$ cm of the interaction point, covering a range of $1.479 < |\eta| < 3.0$ and a preshower device (ES) in front of EE, covering the range $1.653 < |\eta| < 2.6$.

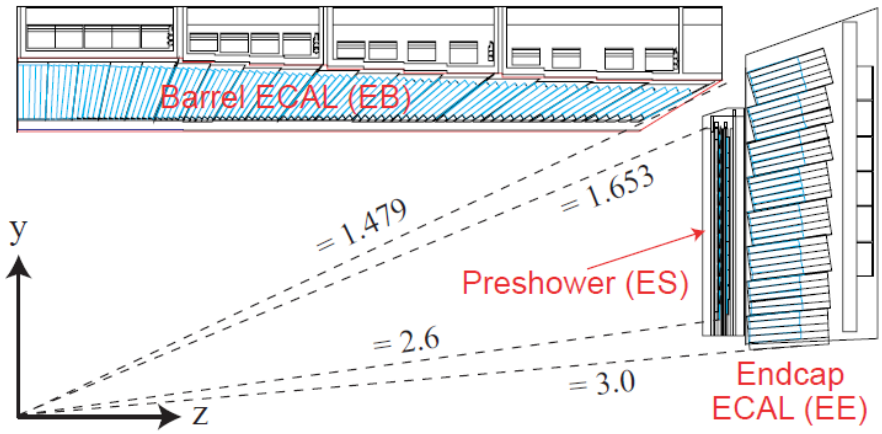


Figure 2.10: Partial view of the longitudinal cut of ECAL, showing the barrel, endcap and preshower.

The EB is made of 61200 crystals, arranged in 36 supermodules separated by 6 mm [52], each covering half of the barrel length. Their crystals are tilted by 3° in η and ϕ with respect to the direction of the interaction point in order to prevent a particle escaping through a gap between two crystals. The barrel crystals cover a surface of $\Delta\eta \times \Delta\varphi = 0.0174 \times 0.0174$ each and have a length of 23 cm. The scintillation light of the barrel crystals is collected by silicon avalanche photodiodes (APD), having an active area of 5×5 mm^2, two of which are glued to the rear side of each crystal.

Each EE disk is made of 7324 crystals, arranged mainly in 5×5 matrices called supercrystals, in a $x - y$ pattern instead of the $\eta - \varphi$ pattern in EB. Their front faces have a surface of 28.6×28.6 mm^2 and their length is 22 cm. Due to increased radiation, vacuum phototriodes (VPT) were used as photodetectors in EE, one of which is glued to each endcap crystal.

The ES is a sampling calorimeter, consisting of two alternating layers of absorbing elements, lead plates that initiate electromagnetic showers, and active elements, silicon strip detectors, measuring the deposited energy.

The radiation length X_0 of a material is defined via equation (2.7) showing the energy of an electron as a function of its penetration depth x into the material. Its Molière radius R_M is the radius of a cylinder whose axis is parallel to the direction of an incoming particle and that contains 90 % of the particle's electromagnetic shower energy.

$$E(x) = E(0) \times e^{-x/X_0} \tag{2.7}$$

These two factors are essential in determining the position and energy resolution of a calorimeter. The ECAL crystals are made of lead-tungstate ($PbWO_4$) and have small radiation length ($X_0 = 0.89$ cm) and Molière radius ($R_M = 2.2$ cm). These characteristics allow for a compact electromagnetic calorimeter with high granularity. Moreover, the crystals respond quickly as 80 % of the light is emitted within 25 ns. EB crystals have a depths of 25.8 X_0, EE crystals of 24.7 X_0.

For the smooth operation of ECAL, the stability of some parameters is monitored.

- Temperature is kept stable at 18 ± 0.05 °C in order to preserve the energy resolution, as the light output of the crystal and the APD gain depend on the temperature, -2.5 % per degree.
- Low and high voltage supplying the photodetectors and the readout cards respectively is kept stable. The operating voltage for the APDs lies between 340 and 430 V depending on the considered photodetector. The APD gain increases by 3.1 % per Volt.

On top of these, the optical transmission of the ECAL crystals is affected by irradiation. Indeed, the high levels of particle fluxes, via ionization radiation, induce the creation of light absorbing colour centres in the crystals. On figure 2.11 one sees the simulation of the transparency variations of a barrel crystal at $|\eta| = 0.92$. The periods of exponential decrease (as between hours 0 and 10) correspond to radiation exposure due to collisions, whereas the periods of transparency increase (as between hours 10 and 12) correspond to the resorption of the colour centres in the absence of incoming radiation.

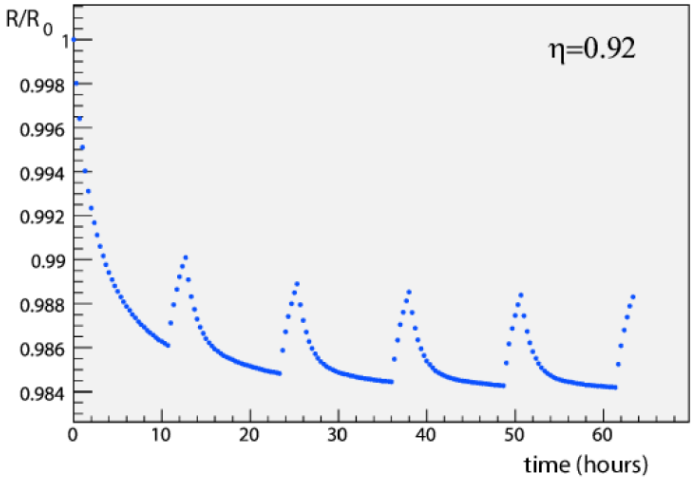


Figure 2.11: Simulation of crystal transparancy evolution from [54]. The periods of exponential decrease (as between hours 0 and 10) correspond to radiation exposure due to collisions, whereas the periods of transparency increase (as between hours 10 and 12) correspond to the resorption of the colour centres in the absence of incoming radiation.

The transparency of every crystal is monitored by a laser monitoring system, described in detail in [55]. It uses two laser sources, a blue one with a wave-length of $\lambda = 440$ nm very close to the scintillation emission peak and a red one at $\lambda = 796$ nm hardly affected by the optical transmission variations and used to monitor the response change due to other sources than the crystal transparency.

The measurement of the APD or VPT response to the laser light on one hand and the response of a PN diode to the direct exposure to laser light on the other, allows the computation of the transparency

to laser light $R(t)$ defined as follows:

$$R(t) = APD(t)/PN(t). \tag{2.8}$$

This ratio can be linked to the the variation in response to scintillation light $S(t)$ by the power law

$$\left(\frac{S(t)}{S(t_0)}\right) = \left(\frac{R(t)}{R(t_0)}\right)^{\alpha} \tag{2.9}$$

where the coefficient α takes on two values depending on the type of crystal considered. In order to insure an ECAL energy measurement independent of the radiation load received by the crystals, the output of each photodetector is corrected by the adequate factor S/S_0.

2.3.2 Data acquisition

Out of the approximate 1.5 MB per event, 100 kB are dedicated to ECAL data. If all crystals were to be read out and saved, ECAL data alone would amount to about 2.0 MB. To achieve the data reduction, a *zero suppression* is applied and only crystals with significant energy deposits $|E| > 80$ (130) MeV are read out in the barrel (endcaps), with the exception of those in regions of high interest as defined by the selective readout processor (SRP), described in detail in [56].

The SRP decision is based on transverse energy deposited in *towers* of 5×5 crystals in the barrel and of a more complex structure in the endcaps. Depending on the total transverse energy E_T^{tower} deposited in the tower, the latter is classified in one of three categories, as shown schematically on figure 2.12:

- if $E_T^{tower} > 2$ GeV (drawn in black on figure 2.12) all the crystals of the tower and the neighbouring towers ($9 \times 25 = 255$ crystals in the barrel) are read out without the application of the zero suppression,
- if $1 < E_T^{tower} < 2$ GeV (drawn in dark grey on figure 2.12) all the crystals of the tower itself are red out without the application of the zero suppression, and
- if $E_T^{tower} < 1$ GeV (drawn in light grey on figure 2.12) zero suppression is applied in the tower, that is only crystals with $E_T \gtrsim 100$ MeV are saved.

By the application of the SRP, the average number of ECAL crystals read out in a collision event of the 2010 data-taking, is about 3000.

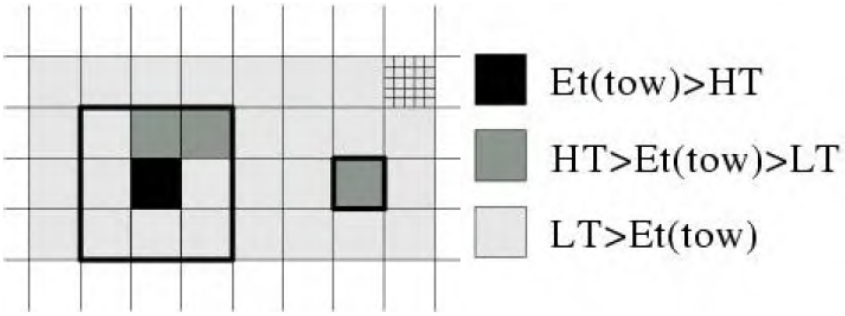


Figure 2.12: Schematic view of the SRP in the barrel. Three regions of interest are defined, depending on the total transverse energy E_T^{tower} deposited in a tower. If $E_T^{tower} > 2$ GeV (black) all the crystals of the tower and the neighbouring towers are read out without the application of the zero suppression, if $1 < E_T^{tower} < 2$ GeV (dark grey) all the crystals of the tower itself are red out without the application of the zero suppression, and if $E_T^{tower} < 1$ GeV (light grey), zero suppression is applied in the tower, that is only crystals with $|E| > 80$ (130) MeV are recorded in the barrel (endcaps).

Chapter 3

Photon Reconstruction

Photons produced in CMS deposit their energy in the crystals of the electromagnetic calorimeter (ECAL) by an electromagnetic shower. The photon reconstruction algorithm, described in [57], reconstructs the photon energy and momentum from the ECAL crystals.

The reconstruction of the crystal energy in ECAL is described in section 3.1, the *clustering* of the crystals corresponding to local energy deposits in section 3.2 and the measurement of the photon energy and position in section 3.3 and 3.4 respectively.

3.1 Energy Reconstruction in the ECAL Crystals

3.1.1 Energy measurement

The photodetectors measure the scintillation light yield of the electromagnetic shower in each crystal, their output is amplified and digitised by the front end electronics. Three amplification gains (factors of 12, 6 and 1) are applied depending on the energy deposited in the crystal, in order to insure the highest resolution and prevent saturation. A channel saturates at approximately 1.7 TeV in the barrel and 3.0 TeV in the endcaps.

For every event the APD or VPT output is sampled ten times with a sampling rate of 40 MHz, thereby constituting the raw data used for the crystal energy measurement. The digitisation is synchronised to the LHC timing such that the maximum of the signal pulse induced by the collision of a given bunch crossing occurs at the sixth sample, as shown on figure 3.1.

The amplitude $\mathcal{A}$ of the pulse in each crystal is computed as a linear combination of the digitised samples as shown in equation (3.1), where w_i are the weights and S_i the digitised values in analog-to-digital converter (ADC) counts.

$$\mathcal{A} = \sum_{i=1}^{10} w_i \times S_i \tag{3.1}$$

The measured output of the photodetector is submitted to fluctuations due to electronic noise, which has an amplitude corresponding to about 35 MeV in the barrel for an amplification gain of 12. The amplitude reconstruction with the weight method of equation 3.1 is equivalent to a χ^2 fit of the amplitude and the baseline [59].

3.1.2 Intercalibration

The crystal amplitude in ADC counts is converted into values of energy via a global energy scale coefficient that will be described in section 3.3.3 and the intercalibration coefficients c_i, known for each

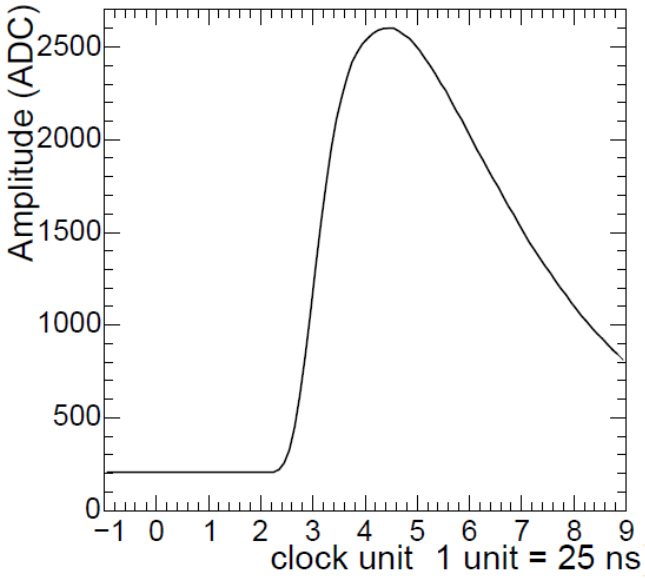


Figure 3.1: Signal pulse in a crystal in response to an electron beam of 120 GeV, from [58]. The digitisation is synchronised to the LHC timing such that the maximum of the signal pulse induced by the collision of a given bunch crossing occurs at the sixth sample.

crystal, compensating the light yield variations per deposited energy from one crystal to the other. The intercalibration of the crystals is required to obtain an optimal energy resolution of the detector.

The intercalibration coefficients were first determined for some supermodules in laboratory measurements [60] and with test-beam data [61]. Subsequently, they were determined for the whole calorimeter by insuring a φ-symmetry of measured minimum bias deposits in ECAL [62] and using the mass peaks of π^0 and η mesons decaying to photon pairs and Z bosons decaying to electron-positron pairs [63]. Preshower strips were calibrated with cosmic rays and subsequently with the tracks of charged particles. The calibration of ECAL with the first data is described in [64].

3.1.3 ECAL Anomalous Signals

Isolated high-energy deposits in ECAL channels have been observed during proton-proton collisions, at all centre-of-mass energies of the 2010 data taking, 900 GeV, 2.36 TeV and 7 TeV [65]. They arise from direct ionisation of the ECAL photodetectors in the barrel, the avalanche photodiodes (APD) by highly ionizing particles produced in the collisions. Those deposits have an estimated rate of about 1 in 1000 minimum bias events [65] which scales with the centre-of-mass energy consistently with the measured increase of charged particle multiplicity [66].

As these *anomalous signals* often correspond to single crystal deposits and interact directly with the APD without previous shower development in the crystal, they can be flagged and rejected based on their shape and their timing as described in [67].

As those isolated signals are made of energy deposited in a single crystal, the topological variable defined as the ratio of the crystal energy $E1$ and the energy sum of the four adjacent crystals $E4$ allows discrimination between photons and photons candidates from *anomalous signals*. The distribution of the variable $(1 - E4/E1)$ is shown on figure 3.2 for the highest energy crystal in each event with a deposit of $E_T > 3$ GeV for data and simulated photon events that do not include a simulation of the anomalous signals. The peak at $(1 - E4/E1) \sim 1$ appearing in the data corresponds to the single-crystal isolated

high energy deposits.

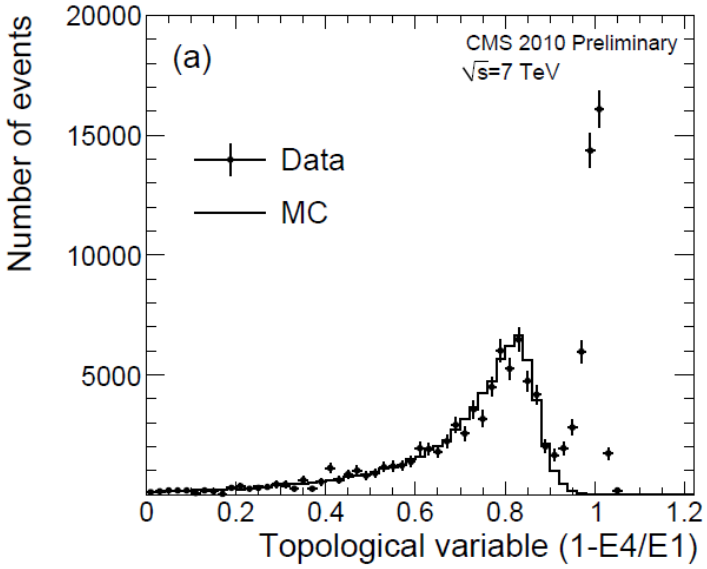


Figure 3.2: Distribution of the variable $(1 - E4/E1)$ for the highest energy crystal in each event with a deposit of $E_T > 3$ GeV for data and simulated photon events that do not include a simulation of the anomalous signals. The peak at $(1 - E4/E1) \sim 1$ appearing in the data corresponds to the single-crystal isolated high energy deposits, from [65].

A selection requirement on this variable $(1 - E4/E1) < 0.95$ has a rejection power that increases with the transverse energy threshold of a given photon sample, reaching 99.2 % for $E_T > 10$ GeV.

Further rejection of the anomalous signals can be achieved by investigating the pulse shapes of the crystals. For collision photons, the measured pulse shape is a convolution of the emission time for scintillating light in crystal and the electronics response time whereas for anomalous signals, crossing the crystal without electromagnetic shower development and interacting directly in the APD, only the response time of the electronics is taken into account.

Figure 3.3 shows the distribution of the reconstructed time, corresponding to the maximum of the signal pulse for the highest energy crystal in each event with a deposit of $E_T > 3$ GeV before and after the selection on the variable $(1 - E4/E1) < 0.95$.

The reconstructed time of collision photons peaks at zero, that of anomalous signals at -10 ns with a timing distribution up to +60 ns due to anomalous signals produced by low momentum neutrons. By requiring the reconstructed time to lie within 5 standard deviations σ_t of 0 (the dependence of σ_t on the reconstructed amplitude is given in [65]), one reduces the contamination of anomalous signals in a sample of photons with $E_T > 25$ GeV to less than 0.2 % [68].

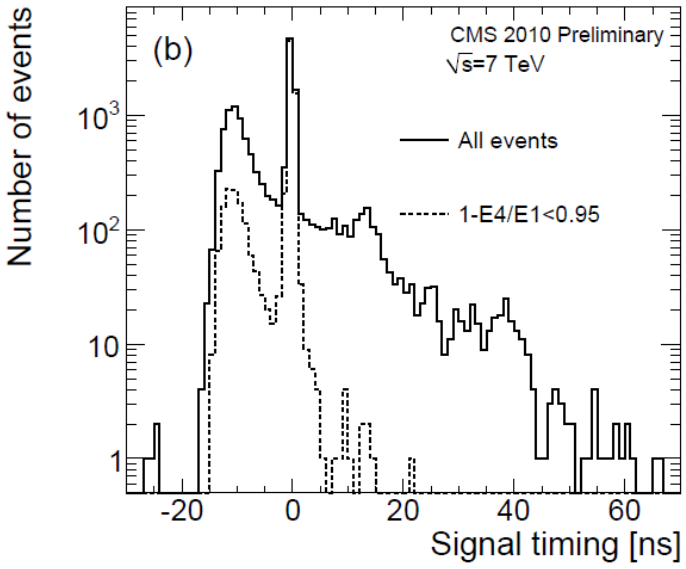


Figure 3.3: Distribution of the reconstructed time, corresponding to the maximum of the signal pulse for the highest energy crystal en each event with a deposit of $E_T > 3$ GeV before and after the selection on the variable $(1 - E4/E1) < 0.95$. The reconstructed time of collision photons peaks at zero, that of anomalous signals at -10 ns with a timing distribution up to +60 ns due to anomalous signals produced by low momentum neutrons, from [65].

3.2 Clustering

Non-converted photons induce a shower in the electromagnetic calorimeter, depositing approximately 94 % of their energy in a matrix of 3×3 and 97 % in a matrix of 5×5 crystals centred around the photon impact position in ECAL [69].

As shown on figure 3.4, before reaching the calorimeter, the trajectory of a photon with $|\eta| < 2.5$ crosses on average 1 radiation length of detector material. This results in an approximate conversion rate of 60 % for a collision photon. For a converted photon, the produced electron-positron pair is separated by the magnetic field along φ and the energy deposits of the photons are spread in that same direction. Two electrons of $p_T = 10$ GeV/c are separated by 0.074 rad, or about 4 barrel crystals, two electrons of $p_T = 30$ GeV/c by 0.025 rad, or about 1.5 barrel crystals.

Because of the different geometry of the ECAL barrel and the endcaps, two different *clustering algorithms* are applied in each. They aggregate crystals in local deposits, called *basic clusters* which are then gathered into *superclusters* that are extended along φ. Along with the calibration, these clustering methods, described in detail in references [69] and [51], insure that the energy of the obtained supercluster yields the best estimation of that of the incoming photon.

3.2.1 ECAL Barrel

The *Hybrid* clustering algorithm is applied in the ECAL barrel. The basic clusters, each built from strips of 1×5 crystals around a local maximum crystal with a transverse energy above the seeding threshold of 1 GeV, are assembled along φ to a supercluster of 5 crystals in η. The algorithm stops if it reaches 17 added strips in φ on each side of the seed crystal, as on figure 3.5.

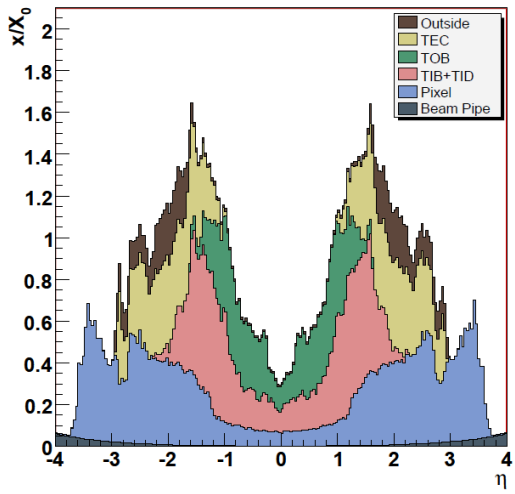


Figure 3.4: Detector material crossed by a photon before reaching ECAL, in radiation lengths, versus the pseudorapidity. TEC is the tracker endcap, TOB the tracker outer barrel, TIB the tracker inner barrel, TID the tracker inner disks. From [70].

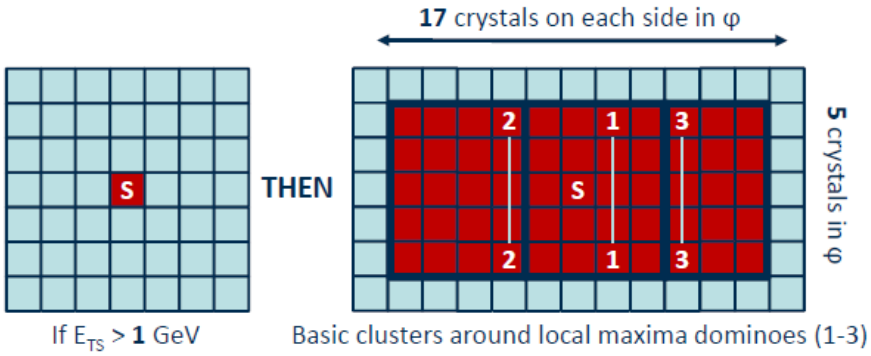


Figure 3.5: Hybrid Superclustering algorithm applied in the Barrel. The seed crystal is labelled **S** and has a transverse energy greater than 1 GeV. Each of the crystal sets labelled **1**, **2** and **3** are the building blocks of the *Hybrid* basic clusters.

3.2.2 ECAL Endcaps and Preshower

The *Multi5x5* clustering algorithm is applied in the ECAL endcaps. Basic clusters shown in figure 3.6 are 5×5 crystal matrices built around seeding crystals having $E_T > 0.18$ GeV. One crystal is assigned to one and one only basic cluster but crystals on the edge of a 5×5 matrix are allowed to seed other clusters [71]. Basic clusters are gathered together in a supercluster, around basic clusters with $E_T > 1.0$ if their pseudorapidity and azimuthal distances respectively satisfy

$$\Delta\eta < 0.14 \quad \text{and} \quad \Delta\varphi < 0.40. \tag{3.2}$$

A preshower cluster is built in front of each basic cluster of the endcap supercluster in a search area centred on the point determined by extrapolating the basic cluster position to the preshower plane in

the direction of the nominal vertex position [51]. The energy of the preshower clusters is then added to that of the endcap clusters.

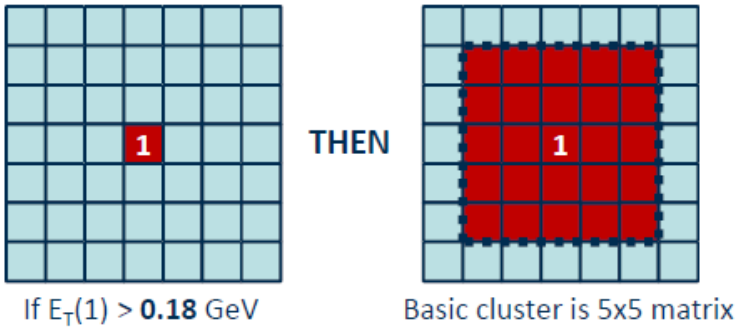


Figure 3.6: Multi5 $\times$ 5 clustering algorithm applied in the endcaps. The superclusters are made from 5×5 crystal matrix basic clusters built around seed crystals (labelled **1**) with $E_T^{seed} > 0.18$ GeV and not being further apart than $\Delta\eta < 0.14$ and $\Delta\varphi < 0.40$.

In every collision event, these clustering algorithms yield a set of superclusters. A photon candidate is defined from each of these superclusters. Let us note here that electrons, like photons induce a shower in the crystals, depositing their energy in the calorimeter. These deposits are clustered by the above mentioned algorithms, thus reconstructing all electrons as photons. It will be shown in chapter 4 how electrons misidentified as photons are rejected by applying a veto on tracks matched with the supercluster.

3.3 Photon Energy Reconstruction

In this section we detail the steps leading to the measurement of the photon momentum and energy and discuss the resolution of the CMS `ECAL` on the latter.

3.3.1 Energy Measurement

As stated in section 3.2, the energy of non-converted photons is well contained in a fixed matrix of 5×5 crystals and the clustering algorithms are aimed at collecting the energy of converted photons spread along φ. The variable $R9$, the ratio of the energy of the 3×3 crystal matrix around the seed crystal and the energy of the supercluster,

$$R9 = \frac{E_{3\times3}}{E_{SC}} \tag{3.3}$$

measures the spread of the energy deposited in a supercluster, tending towards 1 for a very collimated deposit and towards 0 in the case of widespread deposits, as is the case for converted photons. The photon energy is measured in the following way:

- if $R9 > 0.94$ (0.95) in the barrel (endcaps), the photon energy is taken to be the energy sum of all crystals contained in a 5×5 matrix around the crystal of highest energy,
- if $R9 < 0.94$ (0.95) in the barrel (endcaps), the photon energy is taken to be the energy of its supercluster.

3.3.2 Energy Resolution

The ECAL energy resolution can be parametrised [72] as a quadratic sum of the following terms

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2 \tag{3.4}$$

where S is a stochastic term, N the noise and C the constant term. The term C is the limitation on the energy resolution for high energy photons, it is determined in test beams to be $0.26 \pm 0.04\%$ as shown on figure 3.7.

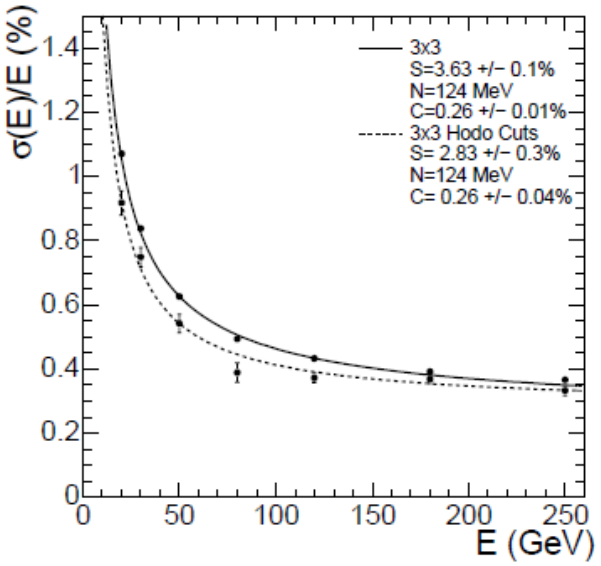


Figure 3.7: Energy resolution for electrons in one supermodule of ECAL in the test beam [72]. The energy is reconstructed with a 3×3 crystal matrix around the impact crystal. The resolution for electrons with $E > 50$ GeV is smaller than 0.6 %.

Figure 3.7 shows the resolution of electrons in the test beam experiments versus their energy. A fit following the parametrisation of equation (3.4) is superimposed. In the test beam experiments, the incident electron energy is known, the electron energy is reconstructed with a 3×3 crystal matrix around the impact crystal and resolution is extracted from a Gaussian fit on the reconstructed energy distribution [72]. The electrons were emitted without magnetic field and without detector material in front of ECAL, therefore they have not emitted bremsstrahlung and these results can be compared to the ones of non-converted photons. The ECAL energy resolution for electrons with an energy above 50 GeV is smaller than 0.6 %.

Figure 3.8 shows the distribution of the ratio of reconstructed and generated energy of photons in the ECAL barrel with $E > 50$ GeV and $R_9 > 0.94$ simulated with PYTHIA. To extract the ECAL energy resolution for those photons, the distribution is fitted with a crystal-ball function, defined in equation (5.7). The fitted mean is compatible with 1 and the fitted resolution is $\sigma_{\text{fit}}/E = 1.1$ % where the statistical uncertainty on the fit is negligible.

The difference between the resolution measured in the test beam and the one obtained with simulated collision events is linked to the presence in simulation of material between the interaction point and the calorimeter resulting in photon conversions and conversion-product bremsstrahlung on one hand and to the strong magnetic field which requires the use of clustering algorithms more complex than the fixed

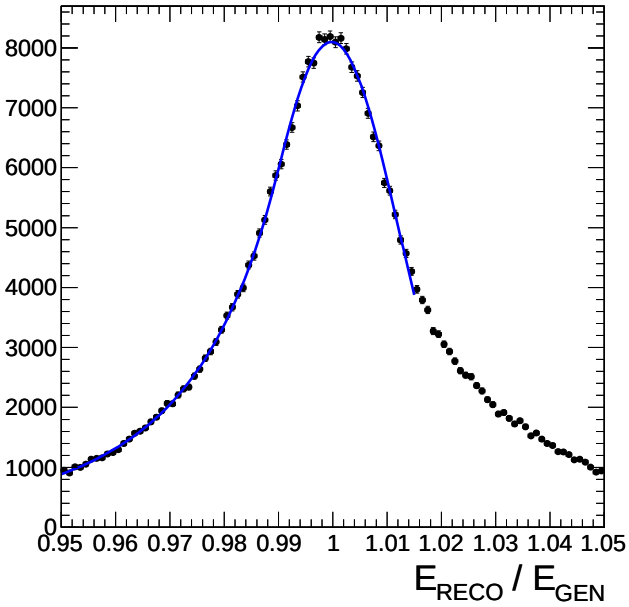


Figure 3.8: Distribution of the ratio of reconstructed and generated energy of photons with $E > 50$ GeV and $R_9 > 0.94$ simulated with PYTHIA. The distribution is fitted with a crystal-ball function, defined in equation (5.7), to extract the ECAL energy resolution. The fitted mean is compatible with 1 and the fitted resolution is found to be $\sigma_{\text{fit}}/E = 1.1$ % with negligible statistical fit uncertainty.

3×3 crystal matrix on the other hand. Furthermore, a more precise intercalibration was applied in the test beam.

3.3.3 Energy Scale

The energy of a photon is computed from the amplitudes measured in the crystals belonging to its cluster via the global energy scale coefficient G as follows [73]

$$E_\gamma = F_\gamma \cdot G \cdot \sum_{\text{crystals}} \mathcal{A}_i c_i \tag{3.5}$$

where F_γ is a correction factor, described in section 3.3.4 and accounting for energy loss due to bremsstrahlung and variations of the shower containment. The sum runs over all the crystals in the photon cluster and $\mathcal{A}_i$ is the reconstructed amplitude in ADC counts in each crystal.

Two global energy scale factors are used, one in the barrel and one in the endcaps [73]. Both were measured in the test beam campaign by equalising the energy sum of a 5×5 matrix centred around the crystal impacted by the electrons to the known electron energy and subsequently corrected for effects of the magnetic field on the shower containment and photodetector gain and for the small temperature difference between the test beam and the LHC environments.

Once the ECAL installed, the global energy scales were determined by tuning the invariant mass peaks of known particles, the Z boson decaying to two electrons for instance [73]. The uncertainty on G is 0.5 % in the barrel and 1.4 % in the endcaps.

3.3.4 Cluster Energy Corrections

Even though the summed energy of all crystals in a supercluster is close to that of the incoming photon, it has to be corrected for the effects of the photon interactions with the detector material and the containment of the electromagnetic shower. The corrections of three effects were tuned on simulation and are applied to the supercluster energy.

- Due to the offpointing crystals in the ECAL barrel the lateral energy leakage of the electromagnetic shower depends on the pseudorapidity η. This dependence is corrected for superclusters in the barrel. As most crystals in the endcaps point towards the centre of the detector or are tilted at a very small angle [74] this correction is not required in the endcaps.
- Photon conversions and subsequent bremsstrahlung by the conversion products spread the photon energy along φ. Energy losses due to these interactions in the detector material are expressed as a function of the ratio of the supercluster sizes σ_φ and σ_η along the φ and η directions, where the size σ_φ is defined in equation 3.6 and σ_η has a similar definition.
- The bremsstrahlung and conversion processes depend on the photon transverse energy and the amount of material the photon crosses before reaching ECAL. This dependence is expressed as a function of the variables E_T^{SC} and η.

$$\sigma_\varphi = \sum_{\text{crystals}} \sqrt{\frac{E_i}{E_{\text{SC}}}(\varphi_i - \varphi_{\text{SC}})} \tag{3.6}$$

In simulation, these corrections amount for about 1 % of a the uncorrected supercluster energy [57].

3.4 Photon Position Reconstruction

3.4.1 Cluster Position

Knowing the photon energy, only the photon direction has to be determined in order to access its momentum. First, the position of a basic cluster $\vec{x}_{\text{bc}}$ is determined by a weighted mean of the positions of its crystals $\vec{x}_i$ [69], using the logarithms of the cluster crystals as weights [75], as follows

$$\vec{x}_{\text{bc}} = \frac{\sum_i w_i \cdot \vec{x}_i}{\sum_i w_i} \quad \text{where} \quad w_i = \max\left(0, w_0 + \log\frac{E_i}{E_{\text{bc}}}\right). \tag{3.7}$$

where the sums run over all crystals in the basic cluster and E_i is the crystal energy. The constant w_0 is 4.7 after empirical studies [76] thus determining that crystals in the cluster containing more than 90 % of the cluster energy contribute to the position measurement.

The position of a supercluster is then calculated by computing the energy weighted mean of its basic clusters and the impact position of a photon in ECAL is taken to be the position of its supercluster.

3.4.2 Vertex Association

The photon being neutral, it does not leave a track in the pixel and the tracker detector and we cannot determine the position of its collision vertex. However, we can infer the position of its vertex from the tracks of the underlying event, potentially of jets produced in association with the photon and, in the case of a converted photon, the reconstructed tracks of the electron-positron pair.

The primary vertex reconstruction algorithm described in [77] reconstructs vertices from hits in the pixel detector associated to tracks. In the case of several reconstructed primary vertices, the photon vertex is chosen to be the one with the highest $\sum p_T^2$ of the reconstructed tracks pointing to it [78].

The probability to associate the correct vertex to a reconstructed photon with this method in the case of a Higgs boson of mass $m_H = 115$ GeV/c^2 is estimated in [79] to be 80 % at low luminosity and 55 % at nominal luminosity with 17.3 pilup events per beam crossing.

Chapter 4

Isolated Prompt Photons

Photons issuing from the hard parton-parton interaction, called *prompt photons*, play an important role in the understanding of quantum chromodynamics (QCD) and physics searches. They constitute clear signatures of new processes such as the decay of a light Standard Model Higgs boson, extra-dimensional gravitons [80] and some supersymmetric states, the neutralino decaying to a photon and a gravitino [81] for instance. These final state photons are not produced along other particles and are therefore not accompanied by electromagnetic or hadronic energy. Properly reconstructing and identifying those *isolated prompt photons* is particularly important for those searches and extensive studies on that matter have been performed at CMS [82].

In this chapter we will discuss their identification by briefly reviewing in section 4.1 the different particles that are falsely identified as isolated prompt photons and how to reject them and describing in detail in section 4.2 a statistical method to accurately estimate the number of events of a process with one ore several prompt isolated photons in the final state. This method will be applied to the measurement of the cross sections of single isolated prompt photons in chapter 5 and pairs of isolated prompt photons in chapter 6.

4.1 Prompt Photon Identification

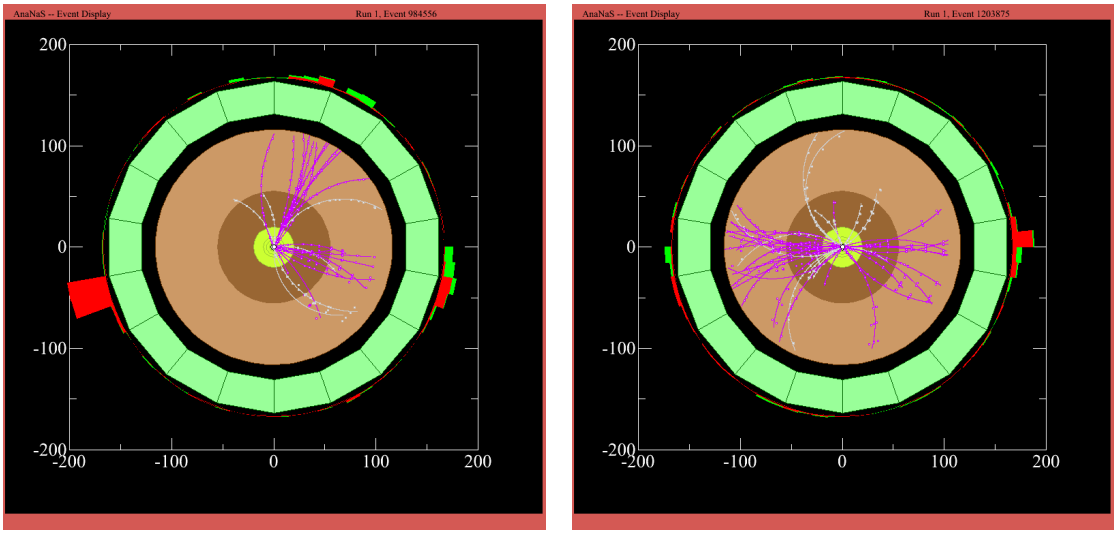
4.1.1 Non-prompt photons

In this section we will discuss the two main sources of background for studies of prompt isolated photons: both photons from jets and electrons leave deposits in the ECAL that can be falsely identified as single prompt isolated photons.

Neutral mesons such as the π^0 and η mesons are produced in jets and decay to a pair of photons, their $\gamma\gamma$ branching ratios are 99 % and 39 % respectively [83]. If the meson carries a substantial part of the jet energy, the two photons will be collimated and have close impact points in ECAL resulting in a single isolated supercluster being reconstructed by the reconstruction algorithms. The event displays of figure 4.1 show a prompt photon and a photon candidate from meson decay.

Even though this misidentification rarely happens to a jet, less than 3 % of simulated jets from partons with transverse energy greater than 30 GeV are reconstructed as photons, the production cross section for jets is three orders of magnitude higher than for prompt photons thus making non-prompt photons from jets the most important source of falsely identified isolated prompt photons.

Electrons just like photons deposit their energy in the crystals of ECAL. In the absence of magnetic field, electrons and photons have a similar calorimetric signature, in the magnetic field of CMS electrons



(a) Prompt photon

(b) Non-prompt photon

Figure 4.1: Event displays from PYTHIA simulation, showing a prompt photon in the left hand part of figure (a) and a non-prompt photon reconstructed from neutral meson decay products in the right-hand part of figure (b).

are bent away from the bremsstrahlung photons that are emitted tangentially to their trajectory, thus leaving wider superclusters along the φ axis. Photons on the other hand can convert in the material before reaching ECAL and the resulting electron-positron pair will also be separated by the field along the φ axis.

4.1.2 Identification variables

Two kinds of variables characterising the photon candidate signature will be introduced in this section, both will be used to distinguish isolated prompt photons from non-prompt photons, for instance the ones reconstructed from neutron meson decay products. The first kind of variables concerns the shape of the photon candidate energy deposits in the calorimeters. Throughout this section we will use two samples of photon candidates: from photon plus jet events with $\hat{p}_T > 30$ GeV/c and from jet events with $\hat{p}_T > 30$ GeV/c, both simulated with PYTHIA. Where $\hat{p}_T$ is defined as the transverse momentum of the hardest parton in the final state.

The variable $\sigma_{i\eta i\eta}$ defined in equation (4.1) quantifies the spread along η of the energy deposits in the central 5×5 matrix of the supercluster.

$$\sigma^2_{i\eta i\eta} = \frac{\sum_i^{5\times5} w_i \left(i\eta_i - i\eta_{seed}\right)^2}{\sum_i^{5\times5} w_i} \quad \text{where} \quad w_i = \max\left(0\ ;\ 4.7 + ln\frac{E_i}{E_{5\times5}}\right) \tag{4.1}$$

where i runs over the crystals of the supercluster that belong toa 5×5 matrix. The energy deposits of an isolated prompt photon can be extended along the φ axis if it is converted to an e^+e^- pair, but are always collimated along η. On the other hand, because of the small separation in η between the two photons from the neutral meson photons that contribute to a non-prompt photon supercluster, the energy deposits will be spread along η leading to higher values of $\sigma_{i\eta i\eta}$. In the barrel, 96.9 ± 0.1 % of simulated prompt photons and only 38.0 ± 0.5 % of simulated non-prompt photons with $E_T > 30$ GeV

have $\sigma_{i\eta i\eta} < 0.01$. The normalised $\sigma_{i\eta i\eta}$ distributions of isolated prompt photons and photons from neutral meson decays are compared on figure 4.2.

The variable H/E defined in equation (4.2) is the ratio of the energy deposited in HCAL behind the photon supercluster and the photon energy as measured by ECAL.

$$\texttt{H/E} = \sum_{i \in C} H(\text{tower})^i / E_\gamma \tag{4.2}$$
where C is the set of all towers in a disk
of radius $R = 0.15$

Non-prompt photons are produced in jets along particles depositing energy in HCAL and take on much higher values of H/E than isolated prompt photons. In the barrel, 99.2 ± 0.1 % of simulated prompt photons and only 73.0 ± 0.4 % of simulated non-prompt photons with $E_T > 30$ GeV have H/E < 0.05. Figure 4.3 shows their normalised H/E distributions.

The other kind of variables concerns the surroundings of the photon, determining whether energy is measured around it in one of the following subdetectors: ECAL, HCAL or the tracker. Non-prompt photons are produced in jets and are therefore surrounded by significantly more charged tracks and hadronic energy.

Non-prompt photons are produced in jets along particles depositing energy in HCAL. Therefore prompt isolated photons take on much lower values of the variable hcalIso, defined in (4.3) as the sum of the transverse energy deposited in HCAL around the photon direction, than non-prompt photons.

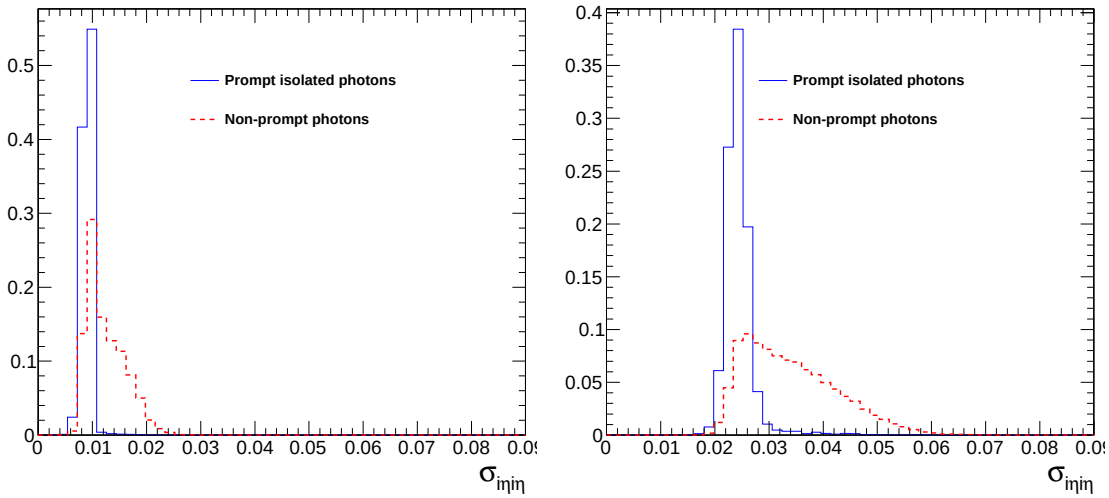
$$\texttt{hcalIso} = \sum_{i \in C} H_T^i(\text{tower}) \tag{4.3}$$
where C is the set of all towers in a ring
with inner radius $R_{in} = 0.15$
and outer radius $R_{out} = 0.4$

In the barrel, 98.9 ± 0.1 % of simulated prompt photons and only 65.5 ± 0.5 % of simulated non-prompt photons with $E_T > 30$ GeV have hcalIso < 2 GeV. The two normalised hcalIso distributions are compared on figure 4.4.

Charged particles produced in jets along non-prompt photons are detected by the tracker. The variable trackIso is defined in equation (4.4) as the sum of the transverse momenta of charged particles coming from the event's primary vertex in a hollow cone around the photon direction. The requirement on particles coming from the primary vertex is imposed in order not to include conversion electrons and positrons in trackIso.

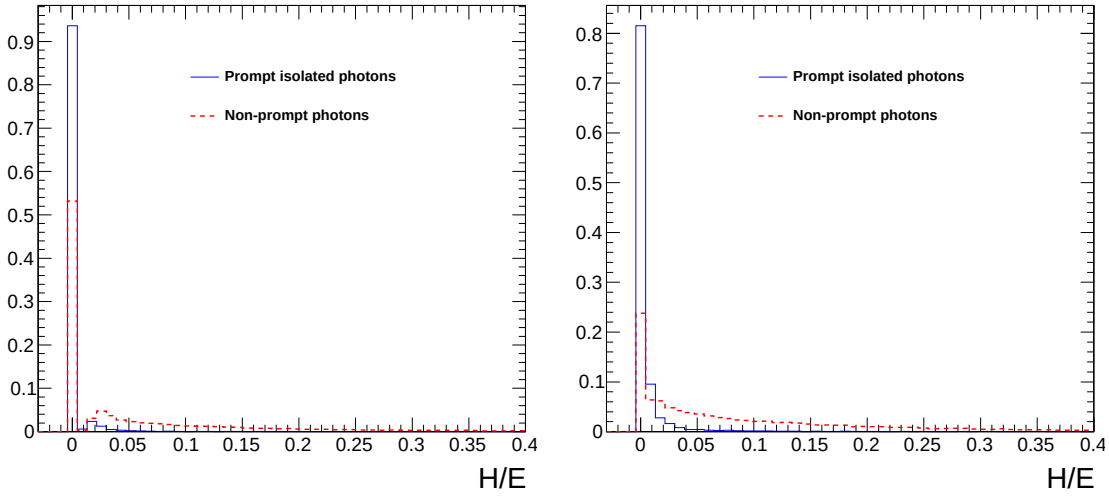
$$\texttt{trackIso} = \sum_{i \in C} p_T^i(\text{track}) \tag{4.4}$$
where C is the set of all tracks in a hollow cone
with inner radius $R_{in} = 0.15$
and outer radius $R_{out} = 0.4$
coming from the event's primary vertex

Prompt isolated photons are not produced alongside tracks and take on much lower values of the trackIso than non-prompt photons. In the barrel, 94.4 ± 0.2 % of simulated prompt photons and only



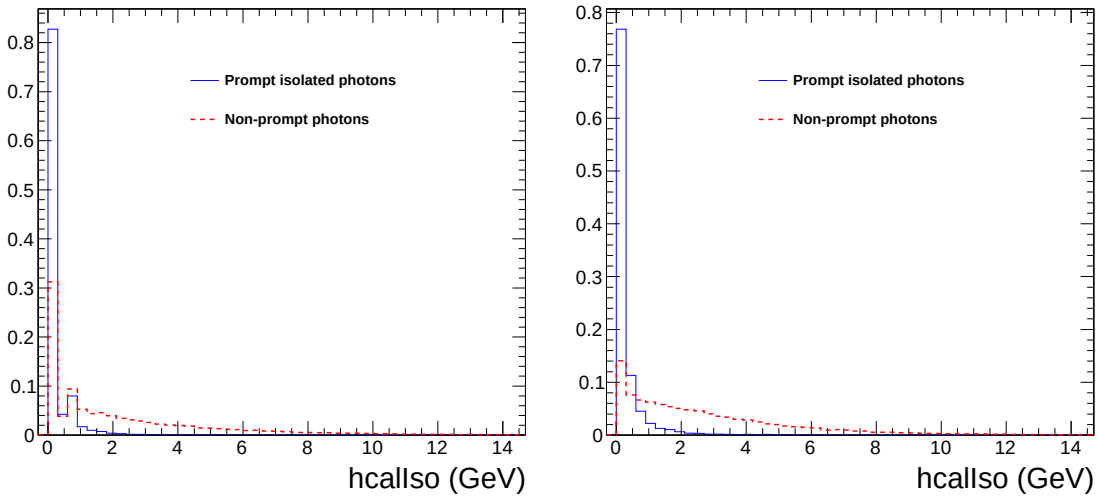
(a) Barrel (b) Endcaps

Figure 4.2: Normalised distributions of the $\sigma_{i\eta i\eta}$ variable of prompt isolated photons and non-prompt photons in single photon events with $E_T > 30$ GeV simulated with PYTHIA.



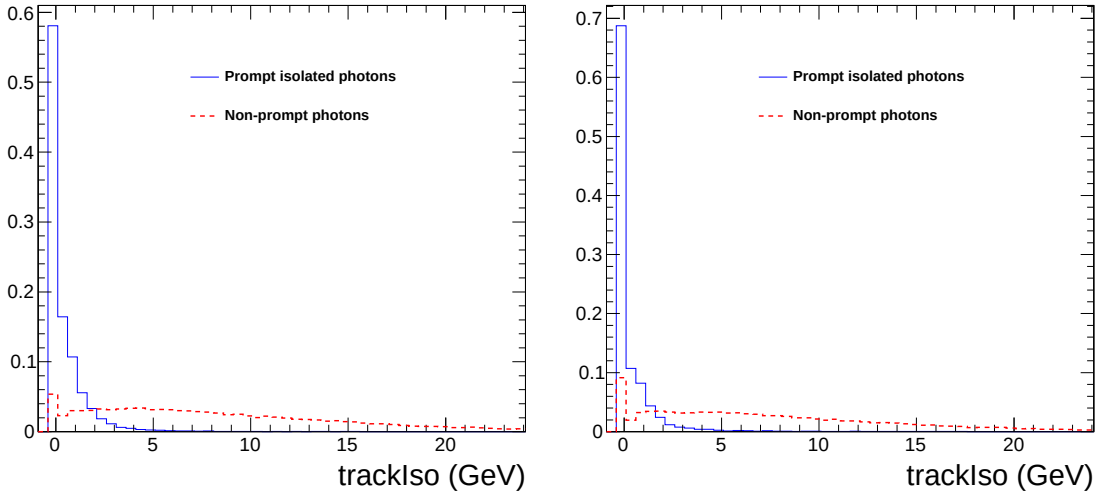
(a) Barrel (b) Endcaps

Figure 4.3: Normalised distributions of the H/E variable of prompt isolated photons and non-prompt photons in single photon events with $E_T > 30$ GeV simulated with PYTHIA.



(a) Barrel (b) Endcaps

Figure 4.4: Normalised distributions of the `hcalIso` variable of prompt isolated photons and non-prompt photons in single photon events with $E_T > 30$ GeV simulated with `PYTHIA`.



(a) Barrel (b) Endcaps

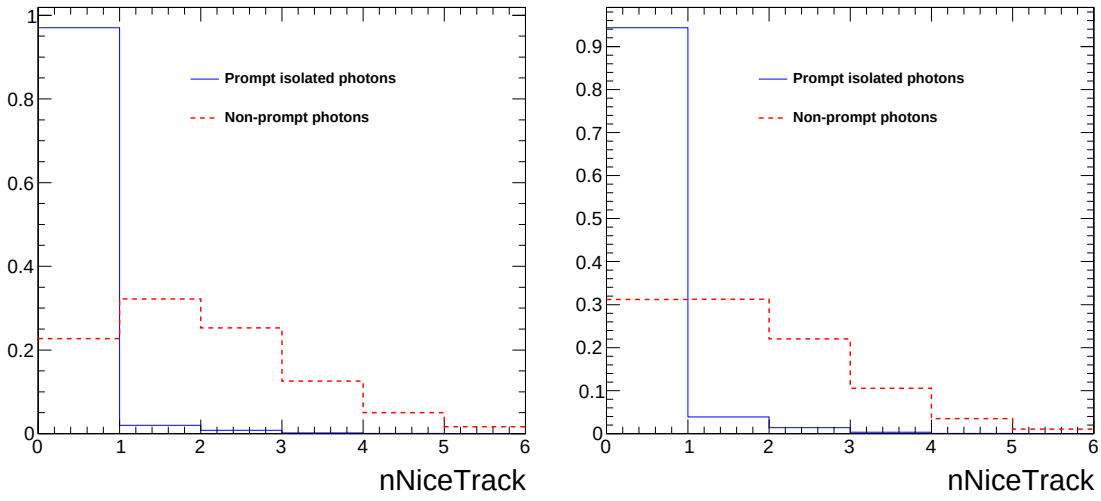
Figure 4.5: Normalised distributions of the `trackIso` variable of prompt isolated photons and non-prompt photons in single photon events with $E_T > 30$ GeV simulated with `PYTHIA`.

19.5 ± 0.4 % of simulated non-prompt photons with $E_T > 30$ GeV have `trackIso` < 2 GeV. The two normalised `trackIso` distributions are compared on figure 4.5.

As a further variable to distinguish prompt and non-prompt photons, we define a track as an *impinging track* if it has a momentum greater than 3 GeV, originates at the event's primary vertex and hits `ECAL` within $R < 0.4$ of a photon's impact position, the definition criteria are detailed in table 4.1. Non-prompt photons being produced in jets are much likelier to have a *impinging track* in their vicinity. In the barrel, 97.6 ± 0.1 % of simulated prompt photons and only 27.0 ± 0.4 % of simulated non-prompt photons with $E_T > 30$ GeV have no *impinging track* in their vicinity. This can be seen on the normalised distributions contrasted on figure 4.6. Adding to the photon selection a veto on the number of *impinging tracks* increases the rejection of non-prompt photons but also allows an essential step in the statistical estimation of those photons as described in section 4.2.4: the selection of a pure sample of non-prompt photons by selecting photon candidates with one *impinging track*.

Definition of an *impinging track*	
impact point	track and photon impact points on `ECAL` to be closer than $R = 0.4$
pixel seed	track required to have hits in the first two layers of the pixel detector
momentum	track momentum as measured by tracker to be greater than 3 GeV
impact parameter	track required to approach the event's primary vertex by less than 2 mm along the z-axis and 1 mm in the transverse plane

Table 4.1: Requirements for a track to be defined as a *impinging track*



(a) Barrel (b) Endcaps

Figure 4.6: Normalised distributions of the number of *impinging tracks* of prompt isolated photons and non-prompt photons in single photon events with $E_T > 30$ GeV simulated with `PYTHIA`.

Finally, the electromagnetic isolation variable `ecalIso` measures the transverse energy deposited in `ECAL` around the photon supercluster. It is defined in (4.5) as the sum the transverse energies of all crystals with $E_T > 300$ MeV in a ring around the photon impact position of which a strip along the φ axis has been removed. Figure 4.7 shows the geometrical definition of this variable in the (η, φ) plane. The strip along φ is designed to exclude the deposits from the electron-positron pair in the case the photon converted in the material in front of `ECAL`. The threshold of 300 MeV has been chosen in order not to take into account any electronic noise and most minimum ionising particles depositing an energy of less than 300 MeV [84] in the the electromagnetic calorimeter.

$$\begin{aligned} \texttt{ecalIso} = \sum_{i \in C} E_T^i (&\text{crystal} \notin \text{supercluster and } E_T^i > 300 \text{ MeV}) \\ &\text{where } C \text{ is a ring with inner radius } R_{in} = 3.5 \text{ crystals} \\ &\text{and outer radius } R_{out} = 0.4 \text{ of which a strip along} \\ &\text{the } \varphi \text{ axis of width } \Delta\eta = 5 \text{ crystals has been removed.} \end{aligned} \tag{4.5}$$

The two photons from neutral meson decays that make up a non-prompt photon's supercluster can be separated in both η and φ directions thereby depositing energy in the isolation surface (blue on the plot), particularly if one or both photons are converted, which is likely to happen with a probability of 70 %. Furthermore, non-prompt photons are produced in jets along with other particles that deposit energy in the `ECAL`. Therefore isolated prompt photons take on much lower values of `ecalIso` than non-prompt photons. In the barrel, 99.4 ± 0.1 % of simulated prompt photons and only 63.2 ± 0.5 % of simulated non-prompt photons with $E_T > 30$ GeV have `ecalIso` < 2 GeV. In the endcaps, 99.7 ± 0.1 % of simulated prompt photons and only 89.2 ± 0.3 % of simulated non-prompt photons with $E_T > 30$ GeV have `ecalIso` < 2 GeV. This appears on the distributions of figure 4.8.

Electrons interact with the pixel and tracker detectors and leave a track that points to the deposits in `ECAL`. We define the variable `pixelMatch` to be `true` if the photon supercluster is matched with track hits in the layers of the pixel detector or the tracker endcap. For this, the supercluster position is backpropagated along a helix trajectory to the innermost part of the detector. Both positive and negative charge hypotheses for the track are considered [85]. By imposing a veto on the variable `pixelMatch` and requiring no *impinging track* to accompany the photon candidate, one rejects 98 % of all events in a sample of electrons from $Z \to ee$ decays simulated with `PYTHIA`.

4.2 Statistical Estimation of the Prompt Photon Yield

4.2.1 Introduction

A sample of photon candidates obtained by applying stringent selection requirements on the aforementioned variables still contains a majority of non-prompt photons: the proportion of non-prompt photons in a sample of `PYTHIA`-simulated photons with transverse energy greater than 30 GeV decreases from 91 % to only 73 % after applying the selection requirements given in table 4.2.
In order to measure the number of prompt photons in a selected sample, called the prompt photon *yield*, we have to use a method based on a statistical treatment of the whole sample of events.

On top of the requirements to reject `ECAL` anomalous signals as described in chapter 3 a selection on $\sigma_{i\eta i\eta} > 0.001$ is applied to reject those rare events where two adjacent anomalous signals are falsely identified as a photon.

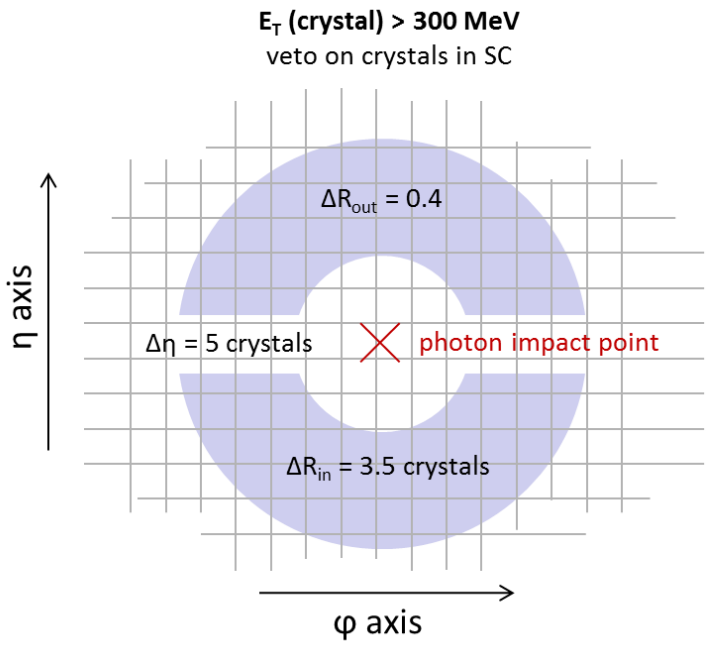


Figure 4.7: Definition of the `ecalIso` variable as the sum the transverse energies of all crystals with $E_T > 300$ MeV in a ring around the photon impact position. The strip along φ is designed to exclude the deposits from the electron-positron pair in the case the photon converted in the detector material before reaching `ECAL`.

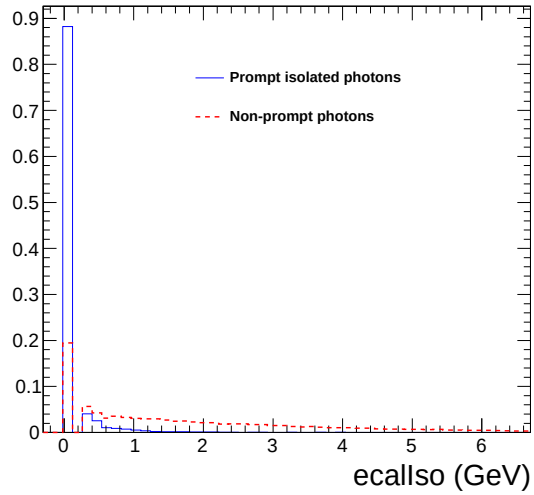
Variable	Requirement
`pixelMatch`	`veto`
`H/E`	< 0.05
$\sigma_{i\eta i\eta}$	< 0.011 (0.03)
	> 0.001
`hcalIso`	< 2 (4) GeV
`trackIso`	< 2 (4) GeV
`ecalIso`	< 2 GeV
n(*impinging tracks*)	$= 0$

Table 4.2: Photon selection requirements. Selection values in the endcaps are presented in parenthesis if different from the barrel. The requirement on $\sigma_{i\eta i\eta} > 0.001$ rejects adjacent `ECAL` anomalous signals.

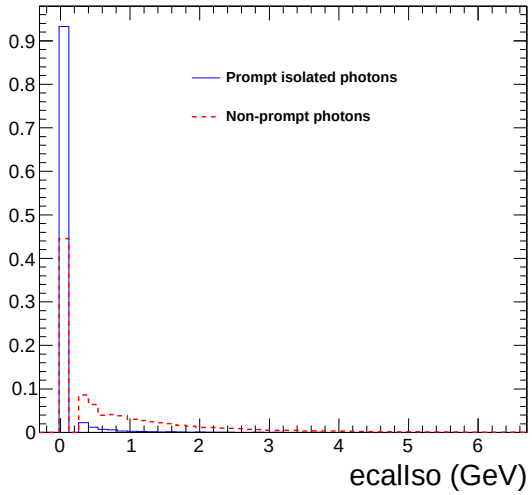
4.2.2 Maximum likelihood fit

Figure 4.8 shows that the shapes of the `ecalIso` distributions of prompt and non-prompt photons differ greatly, non-prompt photons taking on far higher values. Given a sample of prompt and non-prompt photons, *signal* and *background* of this study, and the knowledge of their respective `ecalIso` distributions, one can compute the proportion of signal events by fitting the `ecalIso` distribution of the entire sample with a linear combination of the signal and background distributions.

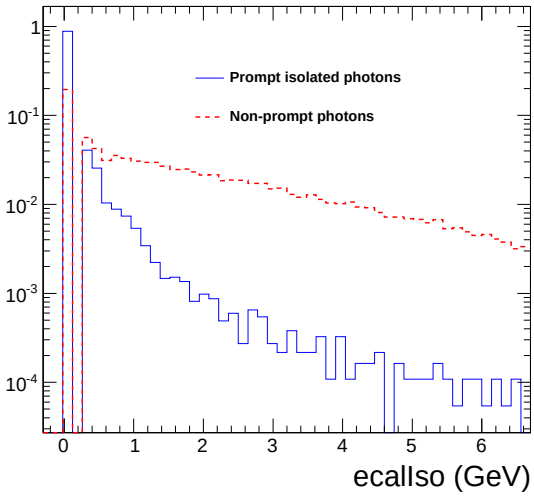
As we cannot obtain pure samples of either isolated prompt photons nor non-prompt photons passing



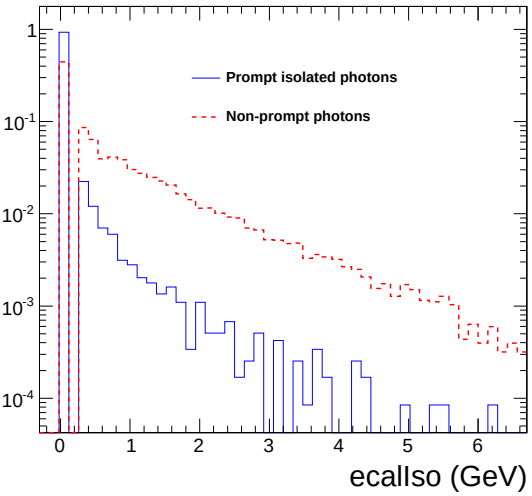
(a) Barrel – linear scale



(b) Endcaps – linear scale



(c) Barrel – logarithmic scale



(d) Endcaps – logarithmic scale

Figure 4.8: Normalised distributions of the `ecalIso` variable of prompt isolated photons and non-prompt photons in single photon events with $E_T > 30$ GeV simulated with `PYTHIA`. Only crystals with a transverse energy greater than 300 MeV are counted in the `ECAL` isolation sum, thus a gap appears below 300 MeV in the `ecalIso` distribution.

all selection criteria, and therefore cannot access the true signal and background `ecalIso` distributions, we are required to use distributions of `ecalIso`, reproducing the true distributions, to extract the amount of signal in the sample. Sections 4.2.3 and 4.2.4 describe the methods used to obtain these distributions from specially selected control samples. Throughout this chapter all selection requirements of table 4.2 except the one on `ecalIso` will be applied.

The `ecalIso` fit of the signal and background distributions to a data sample containing N events is performed via a maximum likelihood method [86]. The signal and background `ecalIso` distributions are used as *probability density functions* (pdf) called f_S and f_B respectively. The likelihood function $\mathcal{L}$, depending on r, the proportion of signal events to be determined by the fit, is expressed as the product of the probabilities $\mathcal{P}(i)$ of each event i in the sample, given the `ecalIso` value $\mathcal{I}_i$ of its photon candidate and given the signal and background probability density functions of the variable `ecalIso`.

$$\mathcal{L}(r) = \prod_i^N \mathcal{P}(i \in S) \cdot \mathcal{P}(\mathcal{I}_i \mid S) + \mathcal{P}(i \in B) \cdot \mathcal{P}(\mathcal{I}_i \mid B) \tag{4.6}$$

$$= \prod_i^N r \cdot f_S(\mathcal{I}_i) + (1-r) \cdot f_B(\mathcal{I}_i) \tag{4.7}$$

Where in equation (4.6) the factors $\mathcal{P}(i \in S)$ and $\mathcal{P}(i \in B)$ represent the probability for the event i to be signal and background respectively, $\mathcal{I}_i$ the value of `ecalIso` of the photon in event i and $\mathcal{P}(\mathcal{I}_i \mid S)$ represents the probability for a photon to have `ecalIso` $= \mathcal{I}_i$ given that it is a prompt photon.

If one considers N, the number of events in the sample, to be the realisation of a Poisson variable with mean Φ, the likelihood function becomes :

$$\mathcal{L}(r, \Phi) \;=\; \frac{e^{-\Phi}\Phi^N}{N!} \prod_i^N r \cdot f_S(\mathcal{I}_i) + (1-r) \cdot f_B(\mathcal{I}_i). \tag{4.8}$$

By defining $N_S = r \cdot \Phi$ and $N_B = (1-r) \cdot \Phi$ one obtains the expression of the *extended likelihood* [87] function (4.9) used throughout this chapter and the next. It depends on the number of signal and background events determined by the fits, called signal and background *yields* respectively.

$$\mathcal{L}(N_S, N_B) = \frac{e^{-(N_S+N_B)}}{N!} \prod_i^N N_S \cdot f_S(\mathcal{I}_i) + N_B \cdot f_B(\mathcal{I}_i) \tag{4.9}$$

By maximising $\mathcal{L}$ with respect to N_S and N_B, one obtains an estimate of the signal yield in the sample.

Studies on the agreement between the data and the simulation, like the ones performed in [57], show discrepancies between the simulated and observed events of up to 30 % in the most populated bins of the studied electromagnetic isolation distribution and a bias towards lower values of the electromagnetic isolation in the simulated events. This can be seen on figure 4.9. In order to avoid uncertainties and biases due to an inaccurate description of the data by the simulation, we will extract the signal and background `ecalIso` distributions from the data itself.

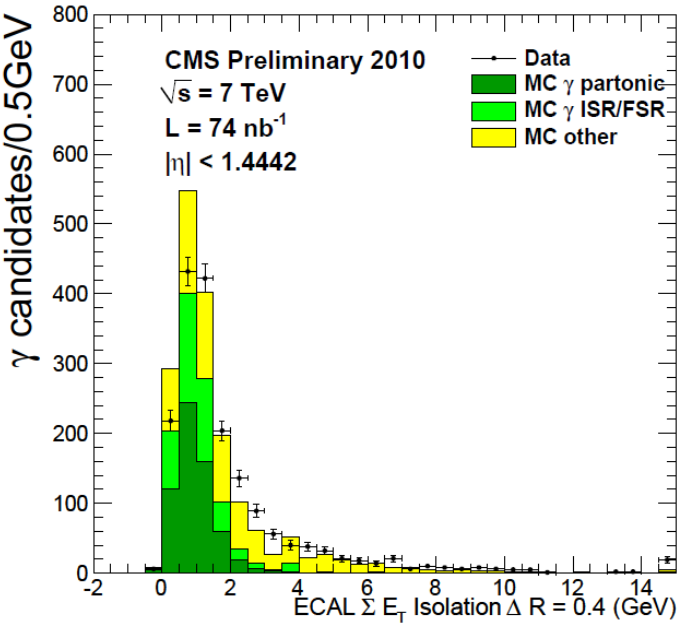


Figure 4.9: Comparison of the distributions of an `ECAL` isolation variable between data and simulation, published in [57]. The origin of the simulated photons is indicated, direct photons, labelled "MC γ partonic", are shown in dark green, fragmentation photons, labelled "MC γ ISR/FSR", are shown in light green and non-prompt photons, labelled "MC other" are shown in yellow. The simulated distribution is scaled to the number of collected photons. The depicted `ECAL` isolation variable has the same geometrical definition as `ecalIso` but the threshold on crystals is $|E| < 80$ MeV.

4.2.3 Signal pdf extraction

Deposits from the photon itself are excluded from the `ecalIso` sum as shown on figure 4.7. Crystals belonging to the supercluster are not considered, deposits of electrons and positrons from conversions are excluded by the strip along the φ axis and the electronic noise, described in chapter 2 to have an amplitude of about 35 MeV in the barrel, is negligible compared to the 300 MeV threshold on crystals to be counted in the sum. Hence the contributions to `ecalIso` are twofold:

- the **underlying event** – jets from partons that have not taken part in the hard interaction
- and **pileup** – further colliding protons from the same bunch.

On average these contributions do not depend on φ and have a slight dependence on η. By computing `ecalIso` in each event around a point chosen at the photon pseudorapidity and randomly in an azimuthal region of $\pi/2$ around the axis perpendicular to its impact point, one collects deposits due to the underlying event and pile-up. In order to insure that these two contributions are the only ones, we require that no reconstructed candidates, such as jets, electrons, photons and muons hit the isolation region. The `ecalIso` distribution thus obtained is a good estimator of the signal distribution.

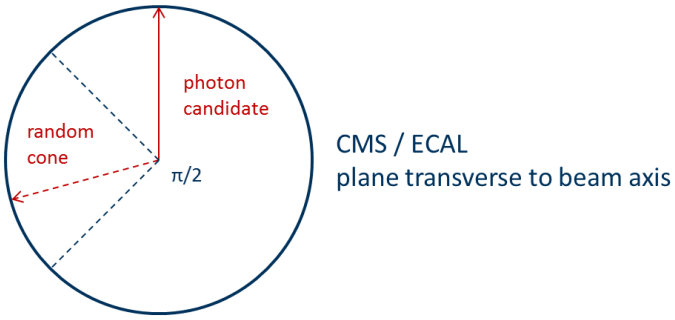


Figure 4.10: The *random cone* method. The `ecalIso` variable is measured in a *random cone* at the same pseudorapidity as the photon candidate and in an azimuthal region of $\pi/2$ around the axis perpendicular to the photon's impact point.

The azimuthal regions at the opposite of the photon candidate has been excluded in order to avoid potential back-to-back particles produced in the same interaction, the region at $\Delta\varphi = 0$ in order to avoid taking into account deposits from the photon itself. This technique, called *random cone* method, is sketched on figure 4.10 and its validity will be proven in the following paragraphs.

Figure 4.11 shows the `ecalIso` distributions obtained when applying the *random cone* technique to a sample of prompt photon and a sample of non-prompt photons, both simulated with `PYTHIA` using the same underlying event tune Z2 [32]. The agreement between the two distributions is good, showing that the kind of events on which the *random cone* method is applied, does not affect the extracted `ecalIso` distribution. Thus it is possible to extract the signal pdf on a sample of photon candidates, independently of the proportion of prompt photons it contains.

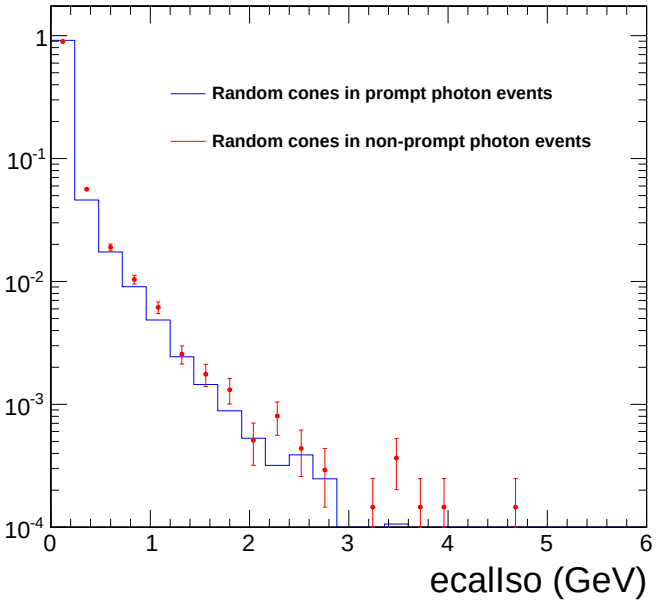


Figure 4.11: `ecalIso` distributions obtained when applying the *random cone* technique to a sample of prompt photon and a sample of non-prompt photons, both simulated with `PYTHIA` using the same underlying event tune Z2 [32]. The agreement between the two distribution is good.

Validation on simulation

On a sample of simulated photons, one can compare the `ecalIso` distributions of the random cones and the photons themselves. This is done in figure 4.12 for the `ECAL` barrel and endcaps, for diphoton events simulated with `MADGRAPH` [88].

The agreement between the `ecalIso` distribution of photons and random cones is very good in the most populated bins up to 4 GeV in the barrel and about 2 GeV in the endcaps. Beyond these values photon energy leakage into the isolation zone is responsible for the discrepancy.

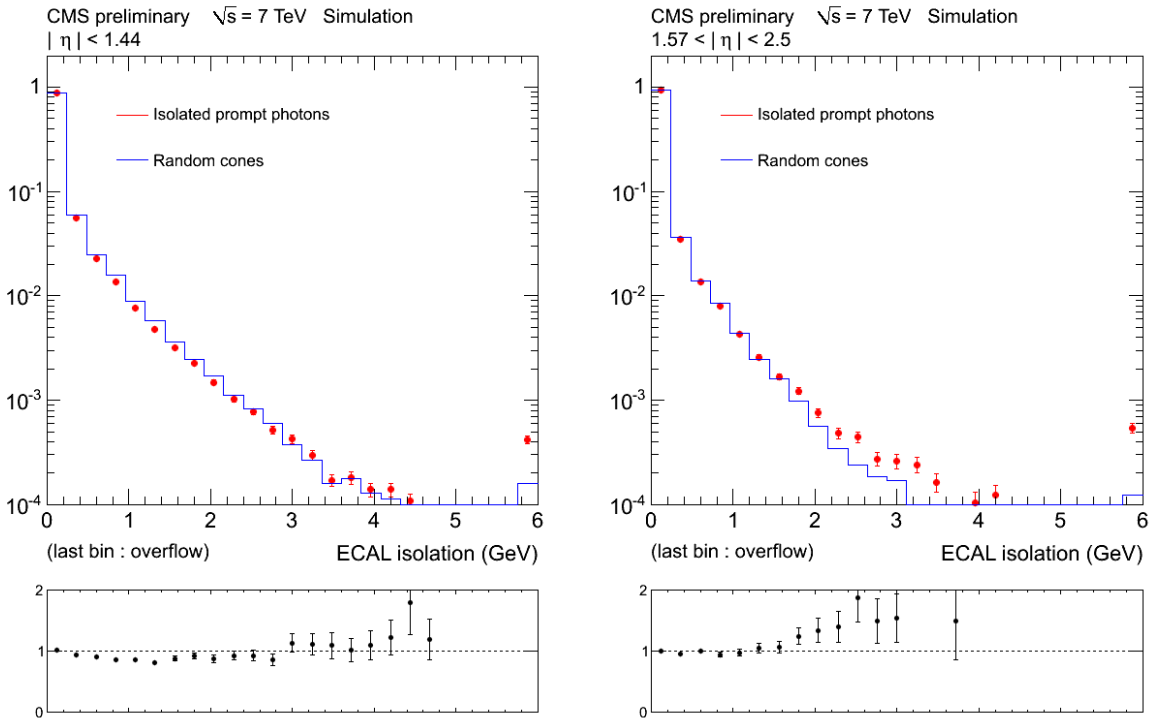
Validation on data electrons

The signature in `ECAL` of electrons and positrons that have emitted a small fraction of their energy by bremsstrahlung is similar to that of prompt isolated photons. Electrons and positrons from Z and W decays are isolated and can thus be used to extract an `ecalIso` distribution reproducing the signal pdf and can serve as a cross check on data of the distribution extracted with *random cones*.

In order to keep only electrons and positrons with collimated deposits in `ECAL`, we add a requirement on the variable `fBrem` < 0.2, defined in equation (4.10) from the initial $p_{T,i}(e)$ and final $p_{T,f}(e)$ momenta of the electron or positron tracks as measured by the tracker, selecting only electrons with low bremsstrahlung losses.

$$\texttt{fBrem} = \frac{p_{T,i}(e) - p_{T,f}(e)}{p_{T,i}(e)} \tag{4.10}$$

In order to obtain a pure sample of electrons from $Z \to e^+e^-$ decays, we selected events with a pair of isolated electrons imposing the strict selection requirements shown in table 4.3 and an invariant



(a) Barrel (b) Endcaps

Figure 4.12: Comparison of the normalised `ecalIso` distributions of prompt photons and random cones in prompt diphoton events simulated with `MADGRAPH`. The bottom plots show the bin-by-bin ratio of the prompt photon and the random cone distributions. The last bin contains the overflow.

mass around the mass of the Z boson ($m \in [70, 110]$ GeV/c^2) on photons with transverse energy above 20 GeV. The second hardest electron has to pass an additional requirement on `ecalIso` to further reject background. The harder electron is used for the pdf extraction and is required to be *low-radiating*, that is to have `fBrem` < 0.2.

Variable	Selection on e_1	Selection on e_2
$\sigma_{i\eta i\eta}$	< 0.011 (0.03)	< 0.011 (0.03)
`H/E`	< 0.05	< 0.05
`hcalIso`	< 2 (4) GeV	< 2 (4) GeV
`trackIso`	< 2 (4) GeV	< 2 (4) GeV
`ecalIso`	-	< 2 GeV
`pixelMatch`	`required`	`required`
`fBrem`	< 0.2	-

Table 4.3: Selection requirements for a sample of low-radiating isolated electrons from $Z \to e^+e^-$ decays. The selection values in the endcaps are indicated in parenthesis if different from the barrel.

The fraction of background events in the selected sample is estimated from simulation to be smaller than 1 %. The `ecalIso` distributions of random cones and electrons from Z boson decays are compared on data events on figure 4.14 and on simulated events on figure 4.13. The agreement between the `ecalIso` distributions of electrons and random cones in the same events is very good in the most populated bins up to `ecalIso` = 1 GeV.

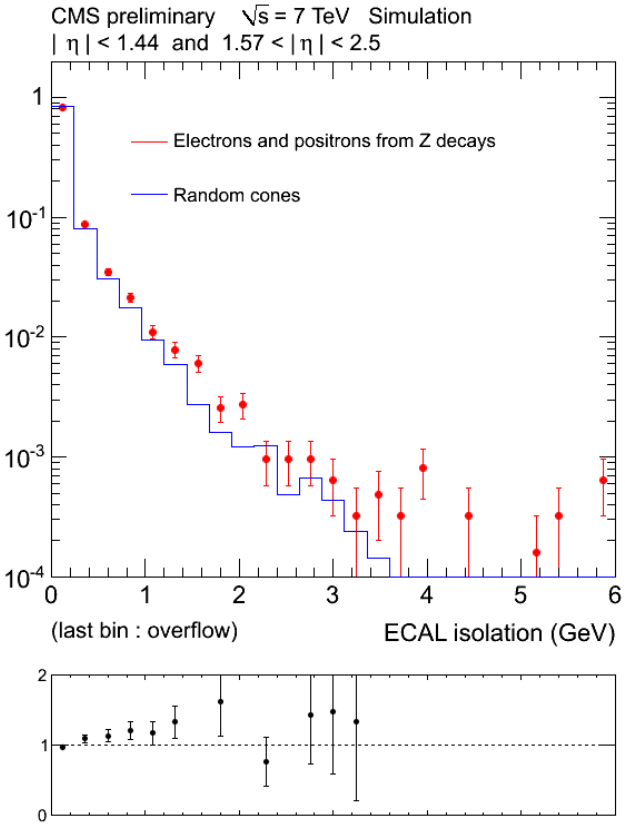


Figure 4.13: Comparison of the normalised distributions of `ecalIso` of low-radiating electrons from $Z \to e^+e^-$ events simulated with `PYTHIA`. The bottom plots show the bin-by-bin ratio of the electron and the random cone distributions. The agreement is fair up to about 3 GeV. The last bin contains the overflow.

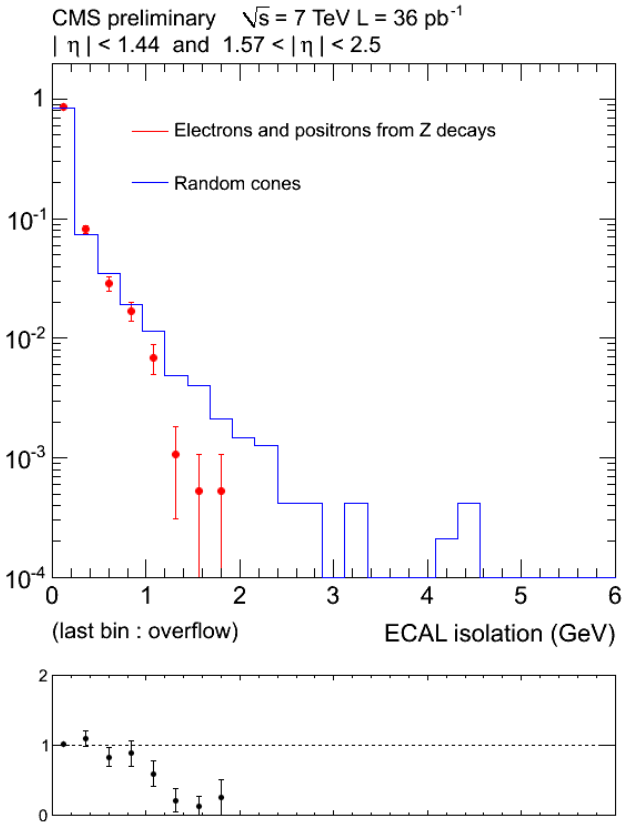


Figure 4.14: Comparison of the normalised distributions of `ecalIso` of low-radiating electrons from $Z \to ee$ decays and random cones in data. The bottom plots show the bin-by-bin ratio of the electron and the random cone distributions. The agreement is good in the most populated bins up to about 1 GeV.

The signal pdf extraction validation is also performed on data electrons from $W \to e\nu$ decays. The production cross section of W bosons being greater than the production cross section of Z bosons by an order of magnitude [89], a larger number electrons is available for this validation. Unlike in the case of the Z boson decay paragraph, it is not possible to apply a selection requirement on the reconstructed mass of the W boson, given that we can only access the transverse momentum of the neutrino. Therefore one cannot obtain a pure signal sample by applying simple selections to reject the background. The `ecalIso` distribution of electrons from W boson decay is extracted using a statistically optimal background subtraction method, the `sPlot` technique [90].

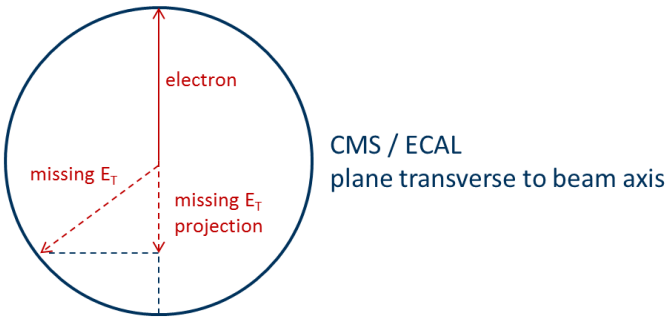
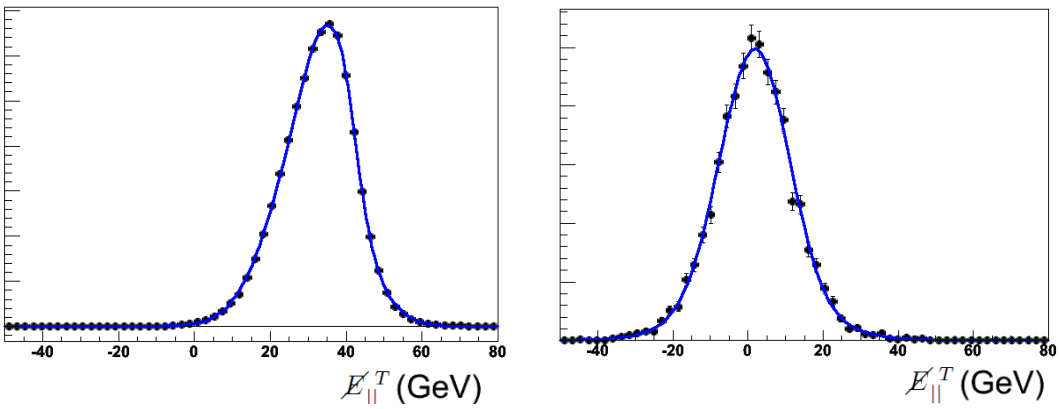


Figure 4.15: Definition of $\not{E}_{||}^{T}$, the projection of the missing transverse energy on the electron axis

This technique uses a maximum likelihood fit on a *discrimination variable*, which is chosen to be $\not{E}_{||}^{T}$, the projection of the missing transverse energy in an event on the axis of the electron as defined in figure 4.15. Knowing the probability density functions (pdf) of the signal (electrons from $W \to e\nu$) and of several background processes ($Z \to ee$, $W \to \tau\nu$, misidentified electrons from γ-jet and QCD events), one can determine the probability for an event to be an electron from a W boson decay. When weighting each event with the signal-probability and plotting the distribution of `ecalIso`, one obtains an estimation of the signal distribution in the case of an unbiased fit with realistic pdfs [90].



(a) $W \to e\nu$ (b) γ-Jet and QCD

Figure 4.16: Probability density functions of the signal $W \to e\nu$ (a) and the combined γ-Jet and QCD background (b) for the fit on the variable $\not{E}_{||}^{T}$.

Events with a low-radiating electron were selected, vetoes on a second electron and jets have been

added to reduce the backgrounds from $Z \to ee$, γ-jet and QCD events.

The $\not\!\!E_{||}^{T}$ pdfs for the signal and for the background processes $Z \to ee$, $W \to \tau\nu$ are extracted from simulation using the `PYTHIA` event generator. The parametrisation of the $\not\!\!E_{||}^{T}$ pdfs for the remaining background processes, γ-jet and QCD events, are determined from simulation using the `PYTHIA` event generator and the parameters are floated in the fit, as they are not well reproduced by the simulation. The pdfs for $W \to e\nu$ and the γ-Jet and QCD background are shown on figure 4.16. The cross section ratios between the signal and the processes $Z \to ee$ and $W \to \tau\nu$ are taken from the Particle Data Group [91], combined with the efficiencies for each process, obtained from simulation and fixed in the fit.

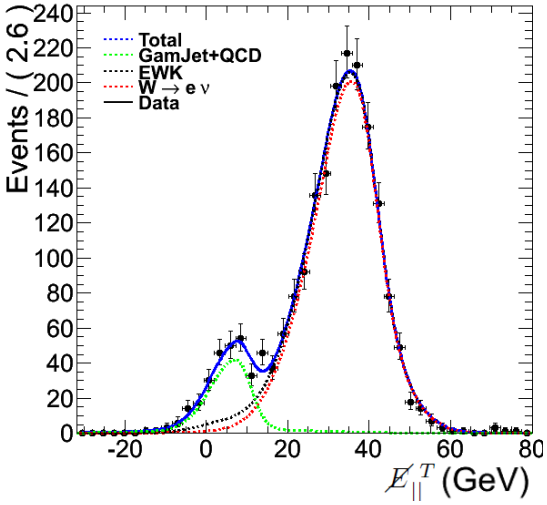


Figure 4.17: Result of the fit on $\not\!\!E_{||}^{T}$. The fitted contribution of the signal (red) and the background (green) sum up to the total (blue) contribution. All three are superimposed on the data sample distribution (black points).

Figure 4.17 shows the $\not\!\!E_{||}^{T}$ distribution from data with the superimposed result of the fit showing a very good matching of the pdfs to the data. The results of the fit are reported in table 4.4.

Species	Number of events fitted
$W \to e\nu$ (signal)	1728 ± 44
$Z \to ee$	set to 0.5% of $N(W \to e\nu)$
$W \to \tau\nu$	set to 4.5% of $N(W \to e\nu)$
γ-jet & QCD	228 ± 20

Table 4.4: Number of events obtained for each species from the fit on $\not\!\!E_{||}^{T}$ for the sPlot technique. The ratios of the $Z \to ee$ and $W \to \tau\nu$ have been determined from simulation and are fixed parameters in the fit.

Figure 4.18 shows the distribution of the variable `ecalIso` for photon-like electrons from $W \to e\nu$ decays obtained with the `sPlot` technique from the fit on $\not\!\!E_{||}^{T}$ compared to the `ecalIso` distribution of random cones. The observed agreement is very good on the entire spectrum.

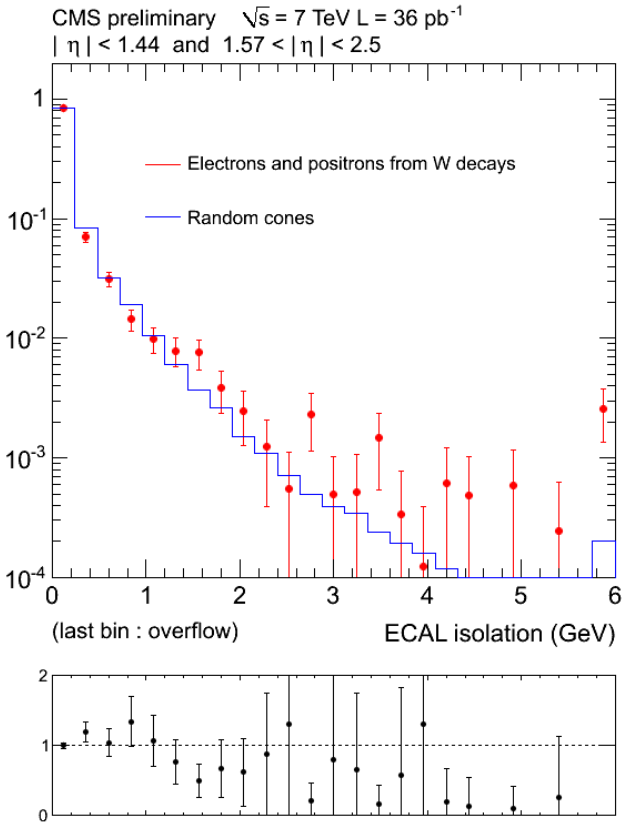


Figure 4.18: Comparison of the normalised distributions of `ecalIso` of electrons from $W \to e\nu$ decays and random cones in data in the `ECAL` barrel. The agreement is excellent in the first bin and very good in the most populated bins up to `ecalIso` = 2.5 GeV. The last bin contains the overflow.

Conclusion

The *random cone* method is a powerful technique to extract the signal pdf for the discrimination fit on `ecalIso`. This method has been validated with simulated diphoton events and with isolated data electrons from the decays of W and Z bosons.

4.2.4 Background pdf extraction

Non-prompt photons that pass the event selections are similar to signal photons in many respects and it is therefore rather challenging to obtain a pure sample of such events for the extraction of the `ecalIso` background pdf.

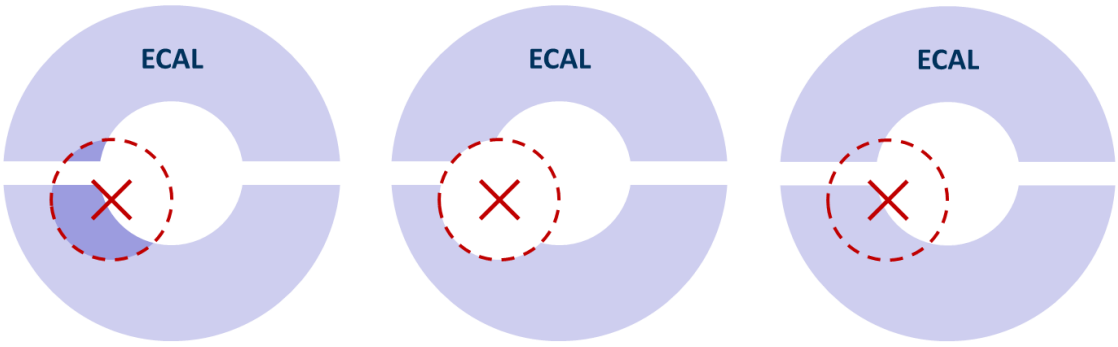
Previous measurements [92] have shown that the extraction of the background pdf was a crucial step in the analysis and a potential source of significant systematic uncertainties. In [92], a fit on the variable $\sigma_{i\eta i\eta}$ is used to perform the statistical discrimination, and the background distribution is obtained by requiring `trackIso` to be greater than 2 GeV, in order to obtain a pure sample of otherwise signal-like background events. However, the correlation between the variables used for the fit and the reverse selection induces a bias in the extracted pdf which is hard to correct and results in significant systematic uncertainties of about 15 %.

We propose a new approach to extract the background pdf. With the method we developed we aim to significantly reduce the systematic uncertainties. This method, which we call the *impinging track* method, identifies non-prompt photons based on the presence of one and one only single *impinging track*. The background `ecalIso` distribution is obtained by subsequently removing the `ECAL` deposits related to the *impinging track* in order not to bias the `ECAL` isolation. The method proceeds in three steps, schematically shown on figure 4.19.

1. Selecting photons passing all selection criteria given in table 5.1 but changing the veto on *impinging tracks* given in table 4.2 and requiring exactly one *impinging track* to accompany the photon as shown on figure 4.19(a), one obtains a pure sample of events with a non-prompt photon. Studies on simulation show that it contains less than 0.1 % of prompt photons.

2. Requiring one *impinging track* induces a bias in the `ecalIso` distribution towards higher values, as the particles leaving the *impinging track* in the tracker – mainly charged pions and kaons in simulated events – are produced in jets along further particles: charged kaons decaying to charged and neutral pions for example. To compensate this effect, all `ECAL` deposits within a radius $R_c = 0.05$ around the *impinging track*'s impact point are removed as shown on figure 4.19(b).

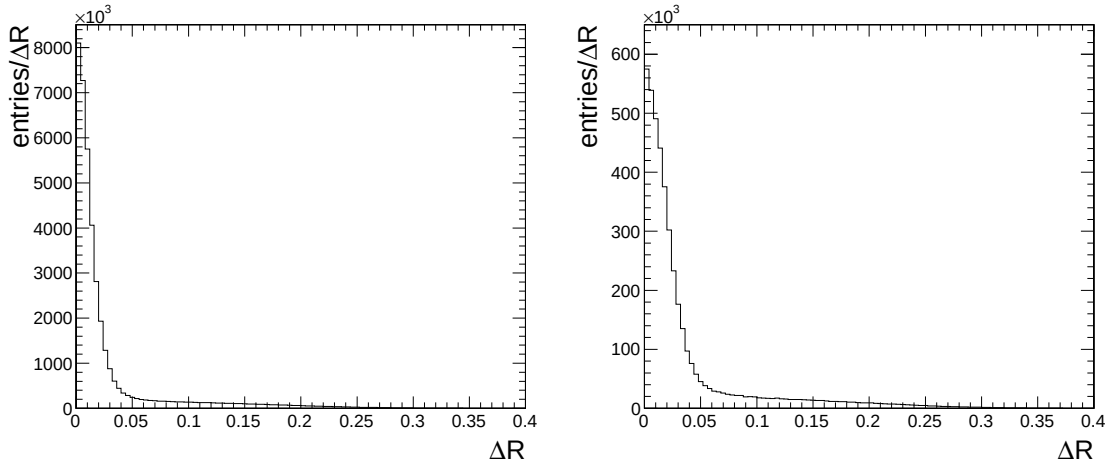
3. In order to compensate for a potential removal of deposits unrelated to the *impinging track*, the value of `ecalIso` computed in the reduced isolation region is normalised to the total isolation area as shown on figure 4.19(c).



(a) Step 1: selection (b) Step 2: removal (c) Step 3: normalisation

Figure 4.19: Schematic view of the *impinging track* method

The number of crystals contributing to the `ecalIso` sum per unit of R, the distance between the crystal and the *impinging track* impact point in the (η, φ) plane, is shown in figure 4.20 for photons selected in the data. The excess observed below $R = 0.05$ is due to the energy deposited in the isolation region by particles produced along the *impinging track*. The radius chosen for the removal of this energy in the step shown on figure 4.19(b) is chosen accordingly.



(a) Barrel

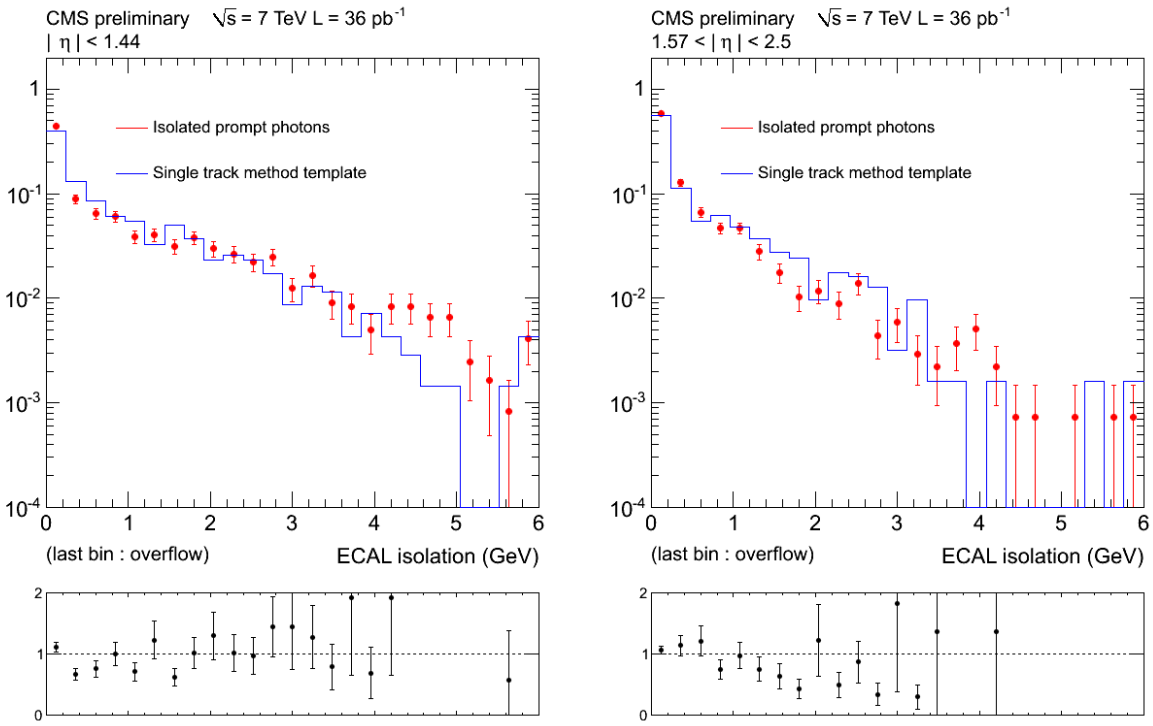
(b) Endcaps

Figure 4.20: Density of crystals with deposits of $E_T > 300$ MeV around the impact point of a *impinging track* versus R, the distance between the crystal and the *impinging track* impact point in the (η, φ) plane, in a sample of photon candidates selected in data.

Extraction validation on simulated events

A comparison of the `ecalIso` distribution of the signal-like non-prompt photons and the pdf obtained via the *impinging track* method is shown on figure 4.21 for the `ECAL` barrel and endcaps. The overall agreement is good.

Further validation of the background pdf extraction method is performed on data.



(a) Barrel (b) Endcaps

Figure 4.21: Comparison of the normalised distribution of `ecalIso` of non-prompt photons and the pdf extracted with the *impinging track* method from `PYTHIA`-simulated jet events from partons with $E_T > 30$ GeV. The bottom plots show the bin-by-bin ratio of the non-prompt photon and the extracted pdf distributions. The distributions agree within statistical uncertainties. The last bin contains the overflow.

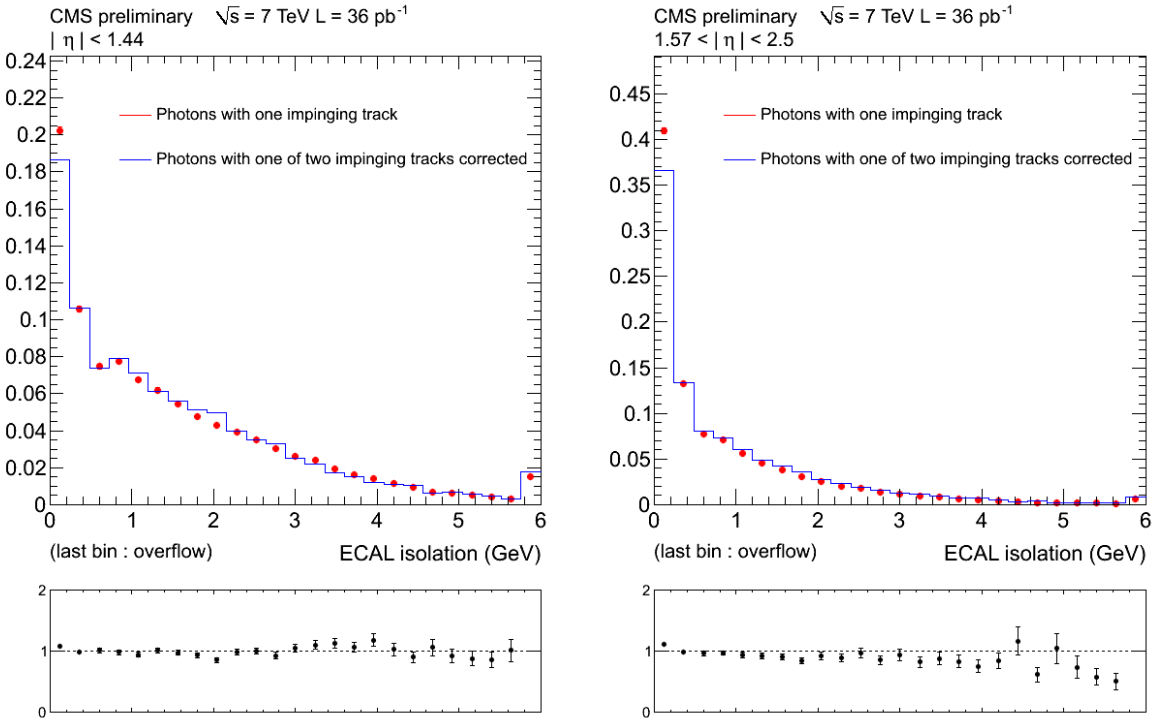
Extraction validation on data

In order to validate the *impinging track* method on data, one compares the following:

- the `ecalIso` distribution of photons with one and one only *impinging track*, that is only step 1 of the *impinging track method* applied and
- the `ecalIso` distribution of photons with exactly two *impinging tracks*, the impact points of the *impinging tracks* separated by more than $2 \times R_c = 0.1$, one of which has its `ECAL` deposits removed via steps 2 and 3 of the *impinging track* method.

The `ecalIso` distributions of those two samples are compared on figure 4.22

These distributions show an excellent agreement demonstrating the effectiveness of the `ECAL` energy deposit removal principle and indicating the strength of the *impinging track* method.



(a) Barrel

(b) Endcaps

Figure 4.22: Comparison of the normalised `ecalIso` distribution of photon candidates produced along one *impinging track* and photon candidates produced along two separated *impinging tracks*, the `ECAL` deposits of one of which have been removed, on data events. The bottom plots show the bin-by-bin ratio of the distributions with one and two *impinging tracks* respectively. The distributions agree very well. The last bin contains the overflow.

4.3 Conclusion

We have investigated in this chapter the different backgrounds to an isolated prompt photon signature and how to perform their rejection. Electrons can be easily discarded but selection requirements are not sufficient to eliminate all non-prompt photons.

Therefore, a statistical background estimation method based on a fit of an electromagnetic isolation variable `ecalIso` was discussed. The entirely data-driven extraction of the signal and background pdfs using the *random cone* technique and the *impinging track* method respectively was validated on simulated events and on data.

The *impinging track* method used for the extraction of the `ecalIso` distribution for non-prompt photons, could be applied to extract the distributions of other variables. Indeed, requiring a photon candidate to be produced along one *impinging track* allows to select a pure sample of non-prompt photons and the bias introduced by this requirement compensated by a subtraction on an event-to-event basis.

For instance, the $\sigma_{i\eta i\eta}$ distribution of non-prompt photons could be extracted by selecting photon candidates accompanied by *impinging tracks*, the `ECAL` impact position of which are required to lie far enough away from the photon supercluster to not impact its lateral shower shape. In this case, an event-by-event subtraction would not be required.

Chapter 5

Inclusive Photon Production Cross Section Measurement

5.1 Introduction

Measuring the isolated prompt photon production cross section allows to test computations of perturbative quantum chromodynamics (pQCD) and yield important informations on the parton distribution functions (PDF) used in these computations. It also yields information about the major irreducible in searches for a light Higgs boson decaying to two photons.

In this chapter we measure the the inclusive photon production cross section in order to validate the method introduced throughout chapter 4 that was developed for the measurement of the production cross section of pairs of isolated prompt photons. The measured differential transverse energy cross section spectrum is compared to a previous measurement performed by the CMS collaboration [68] the results of which are reproduced by the new method. The main systematic uncertainty is assessed and could be significantly improved with respect to that of the previous measurement.

5.1.1 Measurement scope

The cross section measurement carried out in [68] as a function of the transverse photon energy is performed in four bins of photon pseudorapidity. The comparison is performed in two pseudorapidity bins $|\eta| \in [0, 1.44]$ and $|\eta| \in [1.57, 2.5]$ corresponding to photons reconstructed in the `ECAL` barrel and endcaps respectively and in eight transverse energy intervals

$$E_T \in [50, 55]\ [55, 60]\ [60, 65]\ [65, 70]\ [70, 80]\ [80, 100]\ [100, 120]\ [120, 200]\ \text{(in GeV)}. \tag{5.1}$$

Two methods are used in [68] to extract the number of prompt isolated photons: one based on a *combined isolation* variable, defined as the sum of the isolation in `ECAL`, in `HCAL` and in the tracker; the other exploits converted photons to increase the precision of the measurement at low transverse photon momenta. In this chapter we will compare the statistical discrimination method based on `ecalIso` with the *combined isolation* method of [68]. This method measures the E_T spectrum of photons with $|\eta| > 1.57$ starting at 50 GeV, hence the lower bound shown in equation (5.1).

It follows from equation (2.3) that the inclusive differential cross section in a bin is given by the ratio of the number of isolated prompt photons produced in that bin and the product of the bin size and the integrated luminosity. The number of isolated prompt photons produced in a bin can be accessed by dividing the resolution corrected prompt photon yield as given by a background rejection fit on a sample of photon candidates by the selection efficiency of that sample.

The cross section can be expressed as follows:

$$\frac{d\sigma}{dE_T\ |d\eta|}(\gamma + X) = \frac{\mathcal{U}(N^p)}{\epsilon \cdot \Delta E_T \cdot \Delta|\eta| \cdot \mathcal{L}} \quad , \tag{5.2}$$

where $\mathcal{U}(N^p)$ is the yield of isolated prompt photons corrected for the detector resolution, ϵ is the efficiency of the event selection detailed in section 5.1.2, ΔE_T and $\Delta|\eta|$ the widths of the considered bin and $\mathcal{L}$ the integrated luminosity of the sample used for the measurement.

5.1.2 Selection

The trigger selection in this study keeps events that contain a photon with transverse energy greater than 40 GeV and satisfying a loose isolation requirement. The latter includes conditions on the variables `trackIso`, the sum of the transverse momenta of all tracks in a hollow cone around the photon, on `hcalIso`, the sum of the transverse energy deposited in the `HCAL` towers in a hollow cone around the photon impact position, both defined in section 4.1.2, and on a new electromagnetic isolation variable `trigIso`, the ratio of the sum of transverse energies of `ECAL` crystals with $|E| > 80$ MeV in a surface defined by outer and inner radii around the photon impact position on `ECAL` of 3.5 crystals and 0.3 radians respectively and a strip along the φ axis of 5 crystals in width on one hand and the transverse energy of the photon on the other, as follows:

$$\texttt{trigIso} = \sum_{i \in C} E_T^i(\text{crystal} \notin \text{supercluster, and } |E^i| > 80 \text{ MeV}) \ / \ E_T^\gamma \tag{5.3}$$
$$\text{where } C \text{ is a ring with inner radius } R_{in} = 3.5 \text{ crystals}$$
$$\text{and outer radius } R_{out} = 0.3 \text{ and a strip along}$$
$$\text{the } \varphi \text{ axis of } \Delta\eta = 5 \text{ crystals}$$

In each event, the photon candidate with the highest transverse energy is required to pass stringent isolation and shower shape selection criteria, as detailed in table 5.1, designed to reject background from electrons, photon pairs from neutral meson decays and non-isolated photons in jets. A shower-shape selection criterion, $\sigma_{i\eta i\eta} > 0.001$, rejects fake photons reconstructed from anomalous signals in two adjacent crystals of `ECAL`. It is applied on top of the selections detailed in section 3.1.3. The selection on `trigIso` is tighter than the one required by the loose isolation of the trigger path, it is applied to insure a fully efficient trigger selection.

We note that the variables `trigIso` and `ecalIso` have similar definitions, both measure the electromagnetic energy deposited around a photon candidate with differences in the crystal energy threshold, in the geometric definition of the isolation surface and in the normalisation to the transverse energy of the photon candidate. A selection on $\texttt{trigIso} < 0.2$ is equivalent to a selection on $\sum_{i \in C(\texttt{trig})} E_T^i > 10$ GeV. As can be seen on figures 5.3 and 5.4, loosely selecting on `trigIso` still allows for the use of the statistical signal estimation method described in chapter 4. Indeed the distributions of `ecalIso` in the data still reach up further than 6 GeV and allow for a differentiation of signal and background.

Finally, we define the efficiency of the trigger selection defined at the beginning of this paragraph as the probability for a photon passing the selection requirements to fire the trigger. Given the selections of table 5.1, this efficiency is determined from photon events simulated with `PYTHIA` to be greater than 99.9 %.

Variable	Requirement in barrel (endcaps)
E_T	> 50 GeV
`pixelMatch`	`veto`
`H/E`	< 0.05
$\sigma_{i\eta i\eta}$	$< 0.011(0.03)$
	> 0.001
`hcalIso`	$< 2(4)$ GeV
`trackIso`	$< 2(4)$ GeV
`trigIso`	< 0.2
n(*impinging tracks*)	$= 0$

Table 5.1: Photon selection requirements. Selection values in the endcaps are presented in parenthesis if different from the barrel. The selection criterion $\sigma_{i\eta i\eta} > 0.001$ rejects fake photons reconstructed from anomalous signals in two adjacent crystals of `ECAL`. It is applied on top of the selections detailed in section 3.1.3.

5.2 Prompt Photon Yield Extraction

5.2.1 Probability density functions

In every bin of the observable E_T the sample of selected photon candidates contains signal events, isolated prompt photons, and background events such as non-prompt photons from neutral meson decays and non-isolated photons in jets. We perform a maximum likelihood fit on the variable `ecalIso` using as signal and background probability density functions (pdf) the distributions extracted on events of that bin as described in section 4.2 in order to extract the isolated prompt photon yield.

- The **signal pdf** is extracted with the *random cone* technique.
- The **background pdf** with the *impinging track* method.

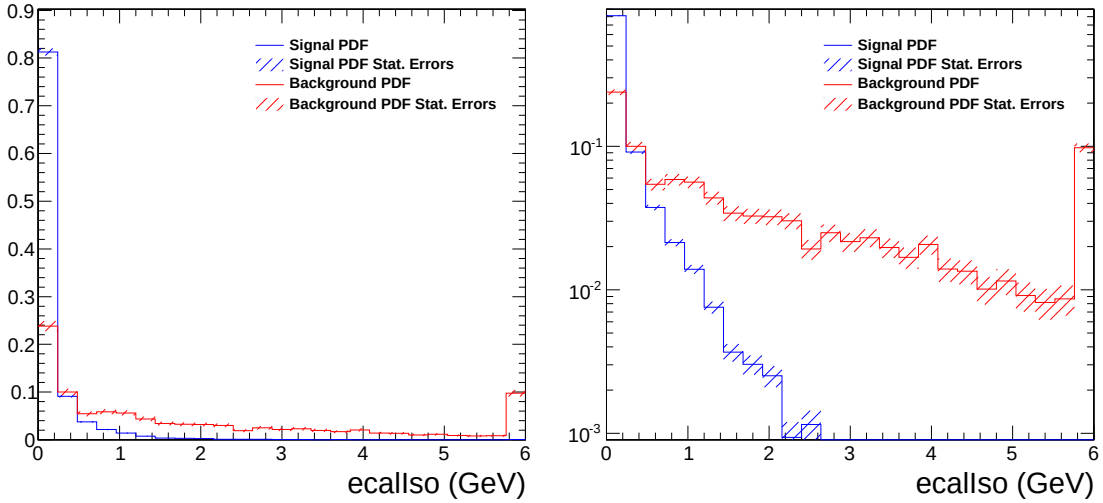
Signal and background pdfs used for the fit are shown for two bins of E_T and $|\eta|$, a bin with few events in the barrel and a bin in the endcap with many, on figures 5.1 and 5.2.

5.2.2 Raw signal yields

Using the pdfs extracted with the methods introduced in section 5.2.1, a fit on `ecalIso` is performed in each E_T bin and for each $|\eta|$ region and the number of prompt isolated photons extracted. These numbers need to be corrected to take into account the detector resolution and are therefore called *raw yields*. This correction is described in section 5.3.1.

Figures 5.3 and 5.4 show the fit for photons with $E_T \in [80, 100]$ GeV in the barrel and photons with $E_T \in [50, 55]$ GeV in the endcaps respectively. Given that the only free parameters of the fit are the signal and the background yields and there is no ad-hoc parameter allowing the shapes of the pdfs to accommodate to the data, the excellent agreement observed indicates that the pdfs describe the data very well.

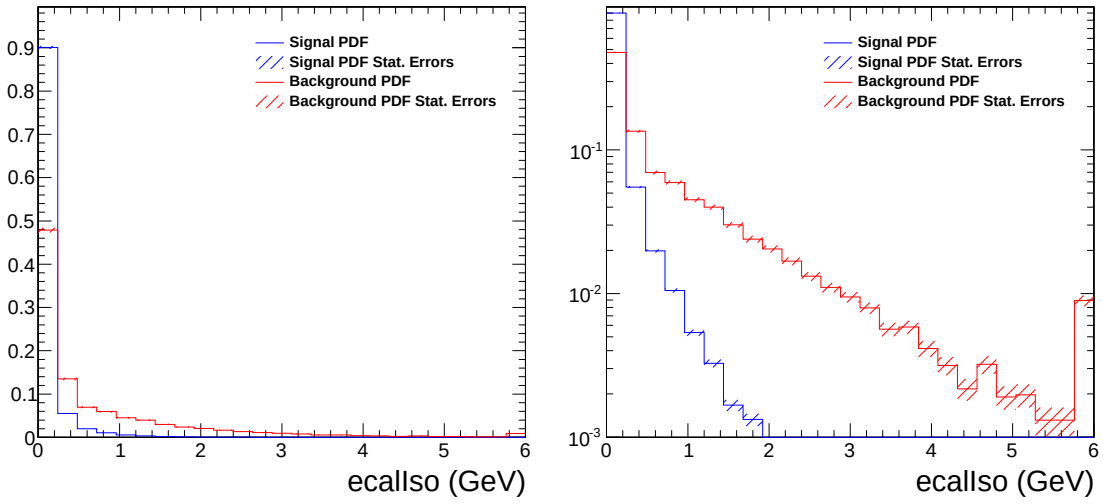
The same procedure is applied in all bins of the E_T spectrum in the barrel and the endcap and the obtained normalised raw signal yield, N^p / ΔE_T using the notations of equation (5.2), is plotted on figure 5.5.



(a) Linear scale

(b) Logarithmic scale

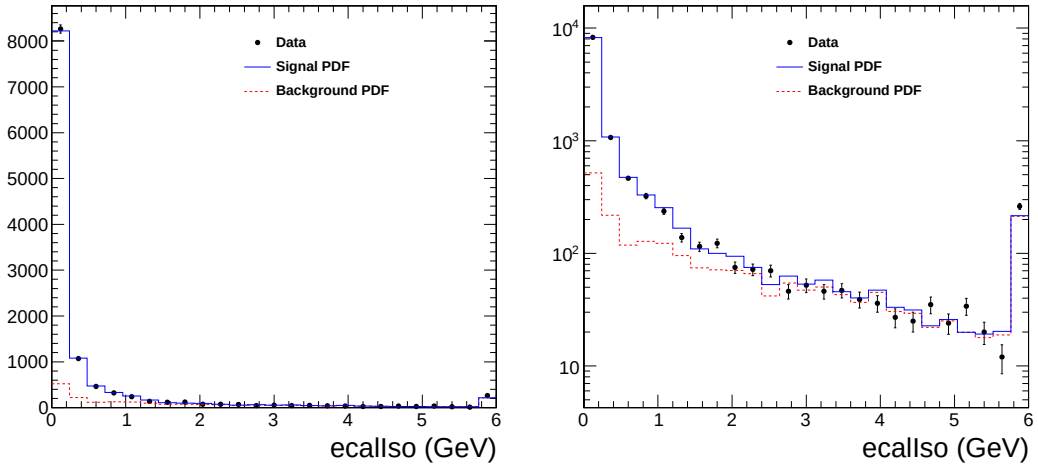
Figure 5.1: Signal and background pdfs in the region $E_T \in [80, 100]$ GeV for photons with superclusters in the barrel, that is having $|\eta| \in [0, 1.44]$. The last bin contains the overflow.



(a) Linear scale

(b) Logarithmic scale

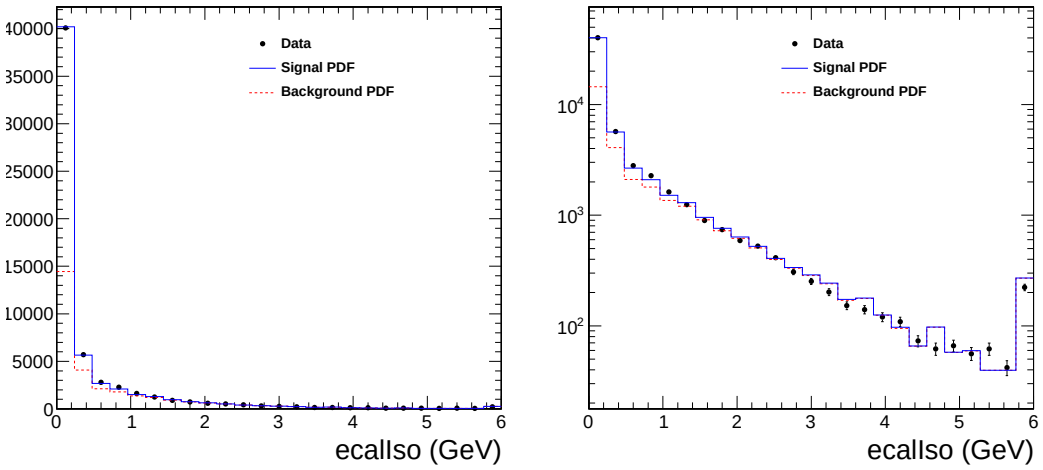
Figure 5.2: Signal and background pdfs in the region $E_T \in [50, 55]$ GeV for photons with supercluster in the endcaps, that is having $|\eta| \in [1.57, 2.5]$. The last bin contains the overflow.



(a) Linear scale

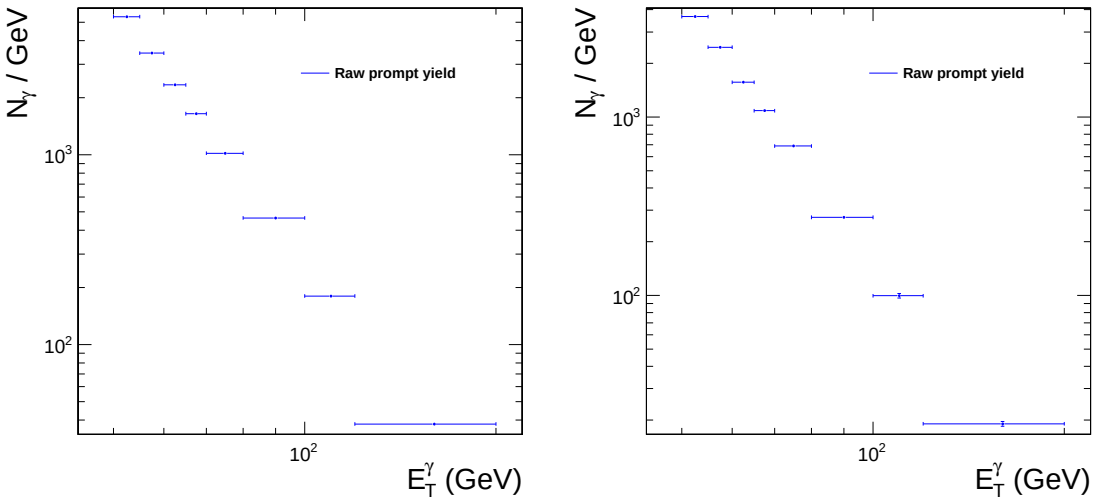
(b) Logarithmic scale

Figure 5.3: Result of the fit on `ecalIso` in the bin corresponding to the range $E_T \in [80, 100]$ GeV in the barrel. The black points represent the data, the continuous blue line the sum of the signal and the background pdfs and the dashed red line the background component. The last bin contains the overflow.



(a) Linear scale (b) Logarithmic scale

Figure 5.4: Result of the fit on `ecalIso` in bin corresponding to the range $E_T \in [50, 55]$ GeV in the endcaps. The black points represent the data, the continuous blue line the sum of the signal and the background pdfs and the dashed red line the background component. The last bin contains the overflow.



(a) Barrel (b) Endcap

Figure 5.5: Normalised raw signal yield spectra in the barrel (a) and the endcaps (b).

5.3 Cross Section Measurement

Two further steps have to be taken to access the differential cross section: the treatment of the detector resolution and the computation of the selection efficiencies.

5.3.1 Unfolding

The raw spectrum of figure 5.5 is the combination of the physical E_T spectrum and the CMS detector resolution. This combination is represented [93] by the integral

$$f(E_T^{meas}) = \int dE_T \; \mathcal{R}(E_T^{meas}, E_T) \; g(E_T) \tag{5.4}$$

where $f(E_T^{meas})$ and $g(E_T)$ are the measured and physical E_T spectra respectively and $\mathcal{R}(E_T^{meas}, E_T)$ the detector resolution function. To access the physical spectrum, the raw spectrum has to be unfolded. This is performed by determining a detector response matrix from simulated photon events.

Indeed, if one supposes the resolution function to be constant in each bin, one can express the content of a bin of the reconstructed spectrum as a linear combination of the bin contents of the physical spectrum and the coefficients of a square *response matrix*.

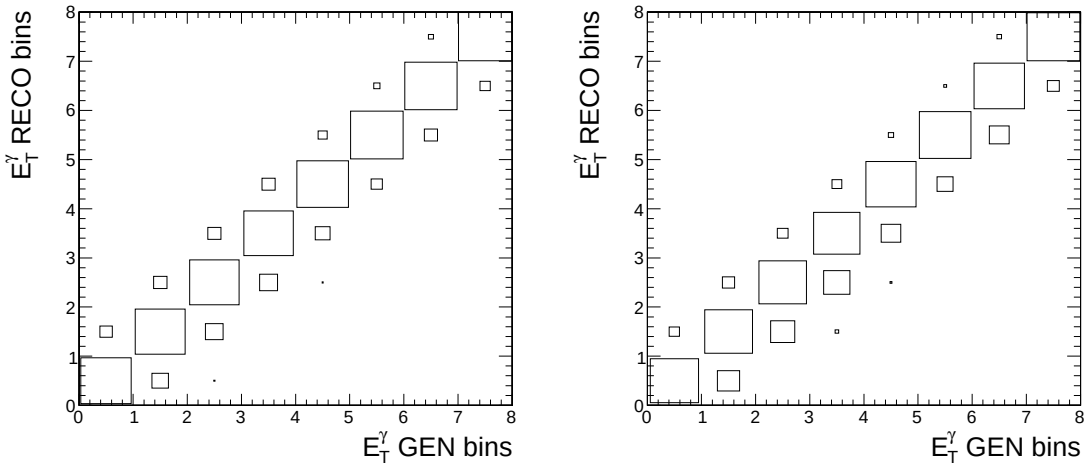
The square matrix $\mathcal{R}$ maps the true value of the simulated photon transverse energy to its measured value, the element (i, j) quantifying the probability for a photon with true E_T in bin j to have a reconstructed E_T in bin i. Off-diagonal cells therefore quantify the bin-to-bin migration and an ideal detector would be characterised by a Dirac-type resolution function and therefore by a diagonal matrix. Under these assumptions, it is possible to obtain the binned physical spectrum $\underline{G} = \{G_j\}$ by applying the inverted response matrix to the binned measured spectrum $\underline{F} = \{F_i\}$.

$$\underline{g} = \mathcal{R}^{-1}\underline{f} \tag{5.5}$$

Given the excellent performance of the CMS electromagnetic calorimeter, which plays the central part in the measurement of the photon transverse energy, the resolution is small with respect to the bin sizes and the response matrix shown on figure 5.6 is close to unity. The unfolding of the detector resolution therefore has a small impact on the measured yields.

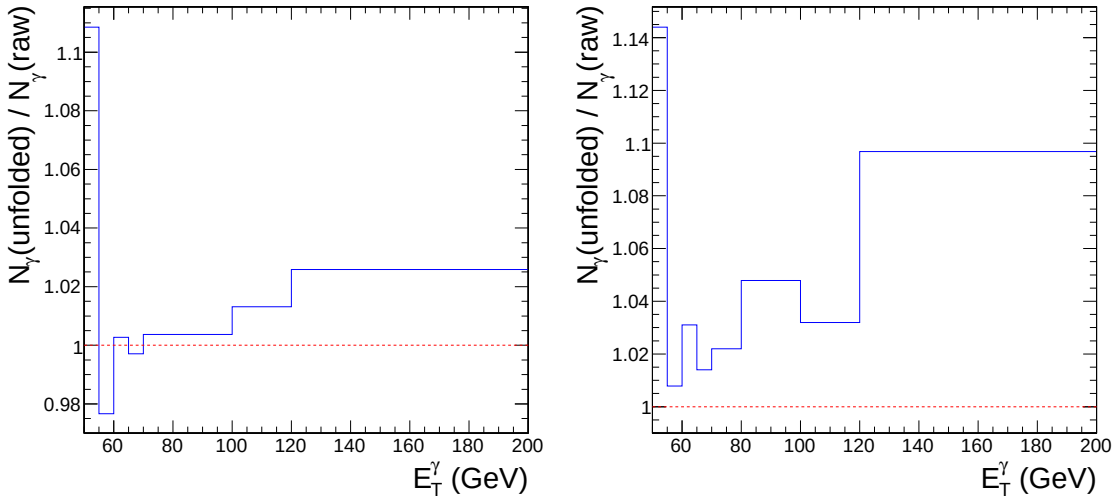
By inverting the response matrix determined from photon events simulated with `PYTHIA` and multiplying it with the vector of raw yields, one obtains the vector of unfolded yields. The ratio of unfolded and raw yields, is plotted on figure 5.7.

As can be seen on figure 5.8, the distribution of the ratio of reconstructed and generated transverse energies of photons simulated with `PYTHIA` with E_T greater than 50 GeV though peaking at 1 has a tail towards lower values of reconstructed photon energies. Indeed, about 5 % of photons with $E_T > 50\ GeV$ have a reconstructed energy more than 10 % lower than their generated energy. The impact of this underestimation is a migration of events to lower bins which is corrected by the unfolding procedure. This appears on figure 5.7, showing the ratio of unfolded and raw yields in each bin, where bins of higher transverse energy tend to be lifted with respect to bins of lower transverse energy. Photons produced with transverse energy close to the selection threshold of 50 GeV have a certain probability not to be reconstructed above this threshold. The unfolding procedure corrects for this effect by increasing the first bin of the spectrum.



(a) Barrel (b) Endcap

Figure 5.6: Detector response matrices (a) in the barrel and (b) in the endcaps, determined with photon events simulated with PYTHIA. The excellent performance of CMS ECAL playing the central part in the measurement of the photon E_T leads to nearly-unitary response matrices and a small impact of the unfolding on the measured yields.



(a) Barrel (b) Endcap

Figure 5.7: Ratio of unfolded and raw yields showing the impact of the unfolding on the measured spectrum. The tail towards low values of the ratio of reconstructed and generated transverse momenta, shown on figure 5.8 is compensated by lifting bins of higher with respect to lower transverse energy. The first bin is lifted to correct for photons produced above and reconstructed below the selection threshold of 50 GeV.

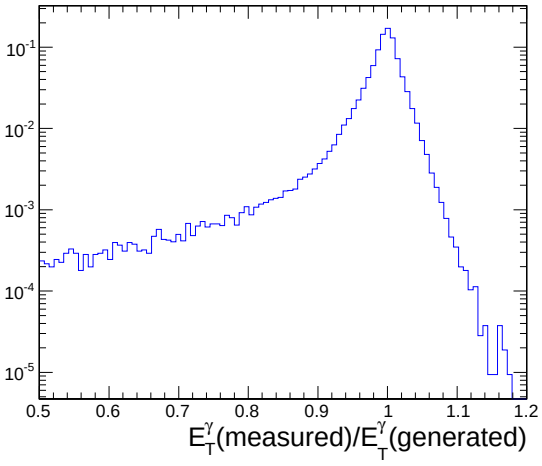


Figure 5.8: Photon transverse energy resolution from photons with $E_T > 50$ GeV simulated with PYTHIA

5.3.2 Selection efficiencies

The efficiency ϵ used in equation (5.2) is the probability for a photon entering the acceptance to be included in the sample used for the measurement. The efficiency ϵ can be split in three parts: the trigger efficiency, the reconstruction efficiency and the identification efficiency. Their computation is detailed in this section.

First, as stated in section 5.1.2, the trigger efficiency is defined as the probability of an event passing the selection requirements to pass the high level trigger requirements used for this study. The trigger efficiency is determined on PYTHIA-simulated photon events to be greater than 99.9%, with a negligible statistical uncertainty.

Second, the photon reconstruction efficiency, defined as the probability for a photon with $E_T > 50$ GeV produced in the pseudorapidity acceptance to be reconstructed in the detector, has been determined from PYTHIA-simulated photon events to be 97.4 % in the barrel and 96.7 % in the endcaps, with a negligible statistical uncertainty.

Third and last, the photon identification efficiency is defined as the probability for a reconstructed prompt photon entering the acceptance to satisfy the selection criteria summarized in table 5.1. All the efficiencies of the selection requirements can be extracted from data, using techniques that will be detailed in this section, with the exception of the pixelMatch veto.

The photon identification efficiency is therefore extracted from PYTHIA-simulated photon events and then corrected by the ratios of data and simulation for those efficiencies that can be extracted from data, as shown in equation (5.9)

The tag and probe technique

The efficiencies of the selection requirements on trigIso, hcalIso, trackIso, H/E and $\sigma_{i\eta i\eta}$ can be computed from data using the so called *tag and probe* technique. Based on the identical signature of electrons and photons in the calorimeter, this method computes the selection efficiencies on electrons from $Z \to ee$ events. These events are tagged using a very well reconstructed electron satisfying stringent electron identification selections, passed by 80 % of simulated electrons, and the probe as any electron

of more than 20 GeV. The tag and probe system is required to have an invariant mass lying between 60 and 120 GeV.

Two invariant mass distributions are fitted to extract the fraction of signal events. On one hand the distribution of all events with at least one tag and one probe, on the other hand the distribution of events with the probe passing the isolation and shower-shape selections described in table 5.1. For the fit, a convolution of a Breit-Wigner funtion, the parameters of which

$$m_Z = 91.1876 \pm 0.0021 \text{ GeV}/c^2 \qquad \Gamma_Z = 2.4952 \pm 0.0023 \text{ GeV}/c^2 \tag{5.6}$$

were taken from the Particle Data Group [91], and an ad-hoc crystal-ball function defined in equation (5.7) with four free parameters is used to model the signal shape and and exponential function to model the background.

$$r(x) = \begin{cases} e^{-(x-\mu)^2/2\sigma^2}/\sqrt{2\pi\sigma^2} & x > -a\cdot\sigma+\mu \\ A(B-(x-\mu)/\sigma)^{-n}/\sqrt{2\pi\sigma^2} & x \leq -a\cdot\sigma+\mu \end{cases} \tag{5.7}$$

where

$$A = \left(\frac{n}{|a|}\right)^n \mathrm{e}^{-a^2/2} \quad \text{and} \quad B = \frac{n}{|a|} - |a| \tag{5.8}$$

The ratio of the fitted yields $\mathrm{N}_{ee}(\text{pass})/\mathrm{N}_{ee}(\text{all})$ gives the efficiency of the considered selection. Figures 5.9 and 5.10 show the fit for the combined isolation and shower-shape selections in the barrel and endcaps respectively.

The random cone method

To determine the efficiency of the veto on the number of *impinging tracks*, we use the number of *impinging tracks* obtained in random cones in data, assuming that the density of *impinging tracks* close to photon candidates, coming from pileup and underlying event interactions, is the same as in random points of the detector. This assumption is verified on single photon events simulated with `PYTHIA`.

Computing the efficiency

The photon identification efficiency labelled $\epsilon_{\text{ ID}}$ is computed from `PYTHIA`-simulated photon events and then corrected by the ratios of data and simulation efficiencies, for those efficiencies that can be extracted from data:

$$\epsilon_{\text{ ID}} = \epsilon_{\text{ ID}}[\text{SIM}] \times \frac{\epsilon_{\text{tag\&probe}}[\text{DATA}]}{\epsilon_{\text{tag\&probe}}[\text{SIM}]} \times \frac{\epsilon_{\text{random cones}}[\text{DATA}]}{\epsilon_{\text{random cones}}[\text{SIM}]} \tag{5.9}$$

Figure 5.11 shows that the identification efficiency computed from `PYTHIA`-simulated photon events is independent of the photon transverse energy.

As stated at the beginning of this section, the overall photon efficiency is computed as the product of the three efficiencies: the trigger efficiency $\epsilon_{\text{ TRIG}}$, the reconstruction efficiency $\epsilon_{\text{ RECO}}$ and the identification efficiency $\epsilon_{\text{ ID}}$:

$$\epsilon = \epsilon_{\text{ TRIG}} \times \epsilon_{\text{ RECO}} \times \epsilon_{\text{ ID}} \tag{5.10}$$

Finally, we can summarise the results in table 5.2 and give the total single photon identification efficiencies in the barrel and the endcaps. The uncertainties on the identification efficiency are a quadratic sum of the statistical uncertainties of the simulated identification efficiency, the agreement of the efficiency extracted with the tag and probe method on data and simulation and the uncertainty determined from studies on $Z \to \mu\mu\gamma$ events [94] on the reproduction of the `pixelMatch` efficiency by simulation.

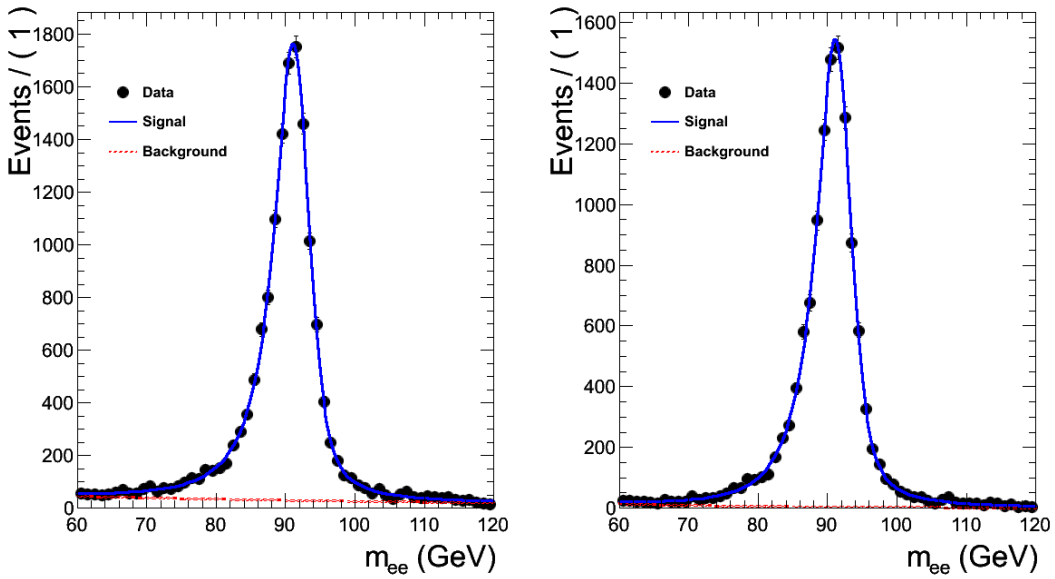


Figure 5.9: Extraction of the efficiencies of the selection requirements on isolation and shower-shape variables with the tag and probe method in the barrel. The fit on the invariant mass of the tag and the probe is shown for all probes on the left and for those probes that pass the identification requirements on the right.

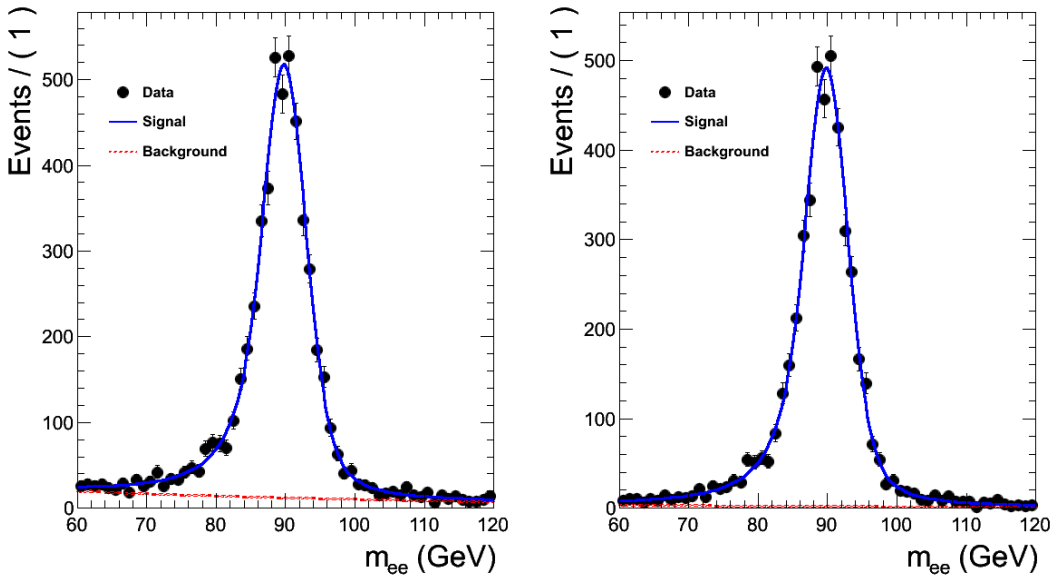


Figure 5.10: Extraction of the efficiencies of the selection requirements on isolation and shower-shape variables with the tag and probe method in the endcap. The fit on the invariant mass of the tag and the probe is shown for all probes on the left and for those probes that pass the identification requirements on the right.

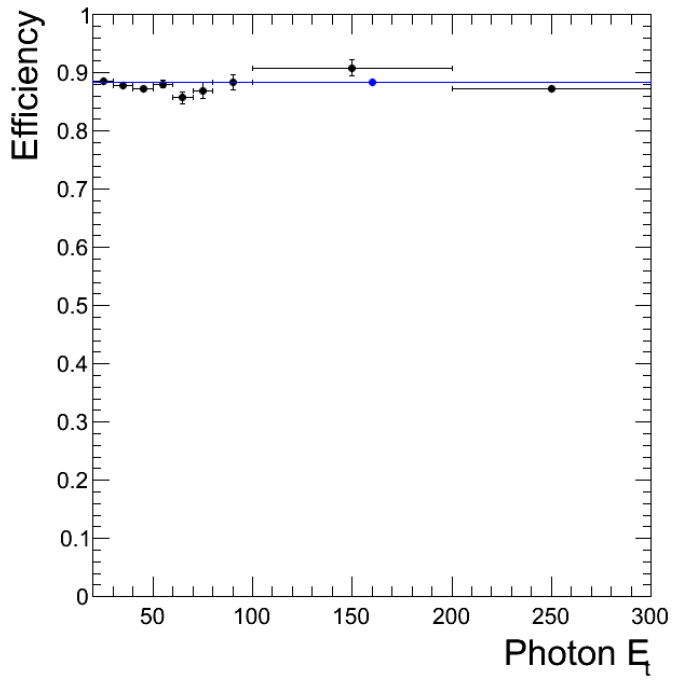


Figure 5.11: The photon identification efficiency computed from PYTHIA-simulated photon events is independent of the photon transverse energy.

	Barrel	Endcaps
ϵ TRIG	99.9 ± 0.0 %	99.9 ± 0.0 %
ϵ RECO	97.4 ± 0.0 %	96.7 ± 0.0 %
ϵ ID	87.0 ± 1.3 %	85.8 ± 4.2 %
Total	84.7 ± 1.3 %	82.9 ± 4.0 %

Table 5.2: Efficiencies [%] of the photon identification

5.4 Comparison of two methods

In this concluding section we contrast the results obtained in this chapter with the statistical discrimination method based on `ecalIso` with the ones obtained with the *combined isolation* method in [68]. After having computed the main systematic uncertainty of the `ecalIso` method in section 5.4.1, we will compare it to the one of the *combined isolation* method in section 5.4.3 and contrast the measured differential cross sections in section 5.4.2.

5.4.1 Uncertainties on the pdf extraction

Given the shape of the pdfs used in the fit, as shown on figure 5.3 for instance, with the probability density peaking at `ecalIso` = 0, the discrimination power resides mainly in the size-difference of the first bins of the signal and the background distributions. The uncertainty linked to the knowledge of the pdfs is therefore dominated by the uncertainty of those first bins. We denote s and b the fraction of the distribution contained in the first bin for the signal and background respectively, that is the probability for a signal or a background event to have `ecalIso`=0. Variations in the shape of the tails are assumed to be a second order effect as will be shown in chapter 6.

To estimate the systematic uncertainties linked to the knowledge of the pdfs, we vary s and b by the amounts δs and δb respectively. Those variations reflect our knowledge of the signal and background pdfs, they arise from the pdf extraction methods presented in section 4.2.3 and 4.2.4:

- The signal probability of the first bin, s, is varied by the difference in first bin probability of the photon `ecalIso` and the *random cone* distributions in simulated events, as shown on figure 4.12.
- The background probability of the first bin, b, is varied by the difference in first bin probability of the `ecalIso` distributions of data photons with one *impinging track* and data photons with two *impinging tracks* one of which had its deposits corrected in data, as shown on figure 4.22.

Those variations are given in table 5.3

Category	δs	δb
Barrel	0.01	0.03
Endcaps	0.01	0.05

Table 5.3: Variations of the first bin of the pdfs used for the determination of the systematic uncertainties on the pdf knowledge, as $s \pm \delta s$ and $b \pm \delta b$

One can make a simple calculation to estimate the uncertainty due to the knowledge of the pdf. s and b are the probabilities for a signal and a background event respectively to be in the first bin of the `ecalIso` distribution and δs and δb the uncertainties on these quantities. Let a be the percentage of events in the first bin of the data `ecalIso` distribution, and N_S the number of signal events, N_B the number of background events and $N = N_S + N_B$ the total numbers of events. We can write

$$sN_S + bN_B = aN \ . \tag{5.11}$$

As $N_S + N_B = N$, we can write the number of signal events as

$$N_S = \frac{a-b}{s-b}N \tag{5.12}$$

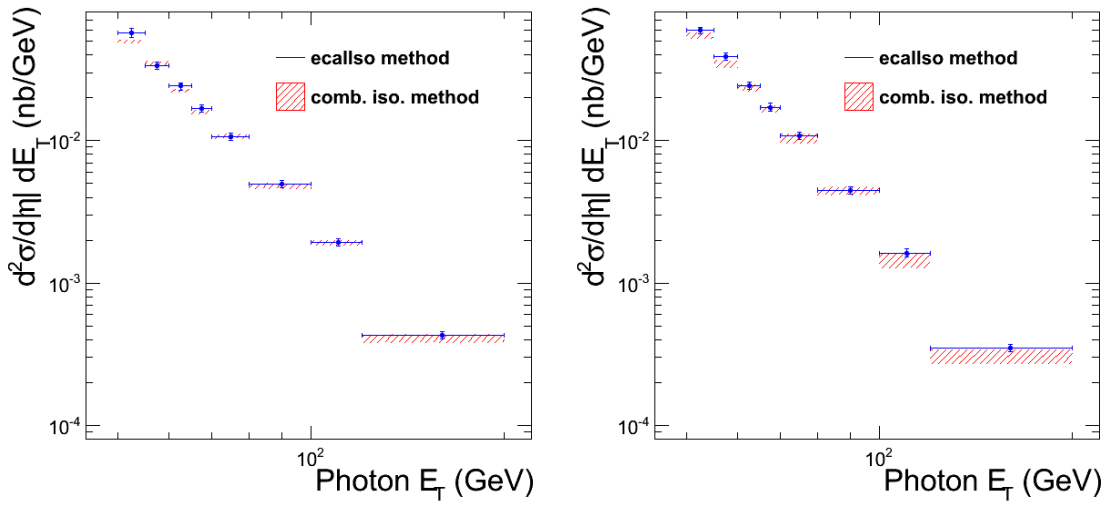
By propagating the errors on s and b to N_S, we obtain the relative signal yield uncertainty

$$\frac{\delta N_S}{N_S} = \frac{\delta s}{s-b} \oplus \frac{\delta b\ |a-s|}{(s-b)\ |a-b|}\ . \tag{5.13}$$

Using the result of equation (5.13) along with the values of table 5.3, we can compute the systematic uncertainties linked to the pdf knowledge in all bins. Thus, in the bin corresponding to $E_T \in [50, 55]$ GeV in the barrel, $s = 0.81$ and $b = 0.27$ with $a = 0.62$ thus yielding an uncertainty of $\delta N_S/N_S = 0.035$. We note that the systematic uncertainty is dominated by the uncertainty on the background ($\delta N_S/N_S$ (b) $= 0.030$). The values of that uncertainty will be given in section 5.4.3 and contrasted with the ones obtained with the *combined isolation* method [68].

5.4.2 Comparing differential E_T spectra

Finally, figure 5.12 shows the comparison of the differential cross section spectra obtained with the `ecalIso` and the *combined isolation* methods.



(a) Barrel (b) Endcaps

Figure 5.12: Comparison of the differential photon production cross sections measured in [68], labelled *comb. iso. method* and with the `ecalIso` method, labelled *ecalIso method*.

The spectrum measured with the `ecalIso` method and the spectrum measured with the *combined isolation* method agree within systematic uncertainties.

5.4.3 Comparing uncertainties

To extract the prompt photon yields, a signal and background component fit on a discrimination variable is also applied in [68]. The pdf extractions however are different:

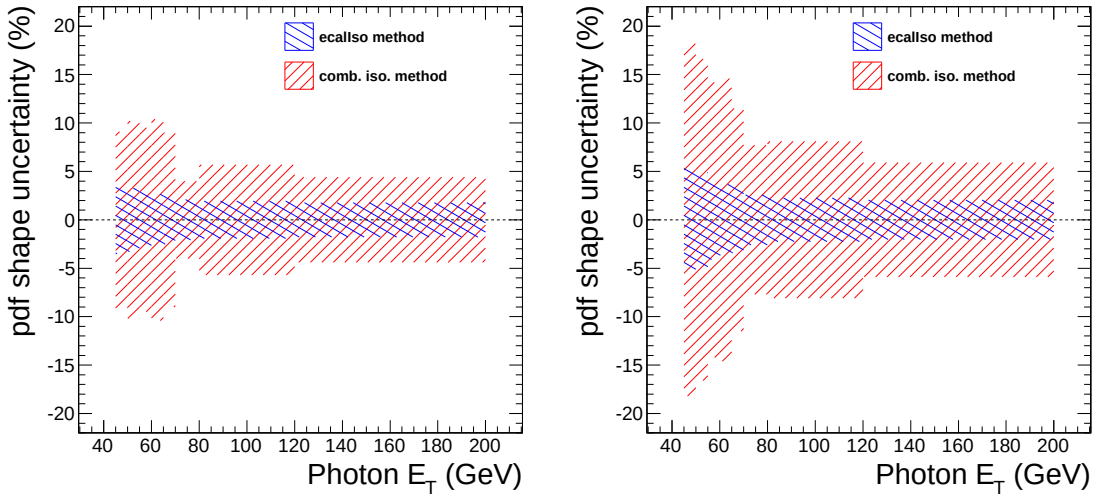
- the **signal pdf** is extracted from simulation and corrected for the differences observed in Z decays between the data and the simulation,
- the **background pdf** is extracted from a background enriched sample obtained from by inverting the requirement on $\sigma_{i\eta i\eta}$, used in the event selection.

The systematic uncertainty linked to the knowledge of the pdfs is then asserted by varying the correction applied to the signal pdf and changing the background pdf by the differences obtained when comparing data and simulation.

Figure 5.13 shows the comparison of these systematic uncertainties with the ones of the method described in section 5.4.1. The uncertainties of the data-driven `ecalIso` method are significantly smaller over the entire E_T spectrum. Equation (5.13) shows that the contribution of the background pdf knowledge is proportional to

$$\left|\frac{a-s}{a-b}\right| = \left|\frac{1}{p} - 1\right| \tag{5.14}$$

where the purity $p = N_S/N$. As the purity increases with increasing E_T and the background term is dominating, the systematic uncertainty decreases with E_T as shown on figure 5.13.



(a) Barrel (b) Endcaps

Figure 5.13: Comparison of systematic uncertainty linked to the knowledge of the signal and background pdfs in [68], labelled *comb. iso. method* and with the computation described in section 5.4.1, labelled *ecalIso method*.

5.4.4 Conclusion

In this chapter, we have measured the differential inclusive single photon cross section versus the photon transverse energy, using the variable `ecalIso` for the discrimination of prompt isolated photons and non-prompt photons produced in jets. Both the signal and the background pdf for the discrimination fit on `ecalIso` are extracted with data-driven methods, thereby allowing a good control of the main systematic uncertainty. The measured spectrum agrees within systematic uncertainties with that measured by CMS. These results validate the prompt photon extraction procedure described in chapter 4 that will be applied to the measurement of the prompt isolated photon pair production cross section in chapter 6.

Chapter 6

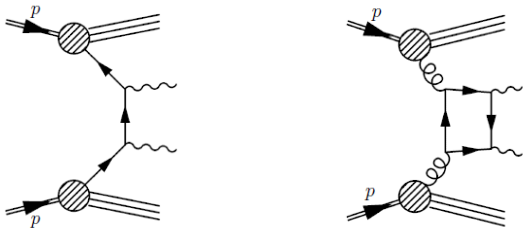
Inclusive Photon Pair Production Cross Section Measurement

6.1 Introduction

6.1.1 Scope of the measurement

As shown in chapter 1, pairs of isolated prompt photons are a very clean signature of a light Standard Model Higgs boson. They are also a signature of extra-dimensional gravitons [80] and some supersymmetric states, as the neutralino decaying to a photon and a gravitino [81]. For those searches of rare processes or new physics, pairs of prompt isolated photons, henceforth called *prompt isolated diphotons*, from QCD processes constitute a major background. Even though in most cases this background can be estimated from the data directly, a good understanding of its kinematic properties and solid theoretical predictions allow for a strengthening of searches for rare processes.

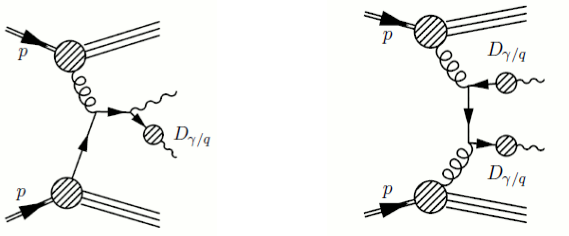
In this chapter we present the measurement of the integrated and differential cross sections of isolated prompt diphoton production. This provides an accurate description of the reducible and irreducible background to the Higgs signal in the diphoton decay channel on one hand and experimental input for existing perturbative QCD parton generator codes on the other. The diphoton production cross section has been previously measured in proton-antiproton collisions at Tevatron by the CDF [95, 96] and D0 [97] collaborations and at 7 TeV in proton-proton collisions by the ATLAS collaboration [98]. This is the first measurement performed with the CMS detector.



(a) Born (b) Box

Figure 6.1: Leading order direct prompt diphoton production processes

Two kinds of production processes contribute to the prompt diphoton cross section. *Direct* photons are produced in the hard interaction as shown on figure 6.1 and *fragmentation* photons fragment off a



(a) One fragmentation (b) Two fragmentations

Figure 6.2: Leading order prompt diphoton production processes from fragmentation

quark produced in the hard interaction and are also counted as isolated prompt photon. Fragmentation processes are depicted on figure 6.2. Even though the gluon-gluon fusion or so-called *box* process of figure 6.1(b) is of higher order of α_s in the perturbation theory than the quark-antiquark annihilation or so-called *Born* process of figure 6.1(a), the proton parton distribution functions (PDF) at the LHC (figure 1.10) are such that the gluon density fosters the *box* to a level comparable with the *Born* process.

6.1.2 Outline of the measurement

The measurement is performed using the full dataset recorded by CMS in 2010, corresponding to an integrated luminosity of 36 ± 1.4 pb^{-1}. Three trigger paths were used to select the dataset, their definition and the integrated luminosity collected by each of them are given in table 6.1. The cross section is measured in the kinematic acceptance described in table 6.2, all quoted cross sections are therefore fiducial.

Name	Description	Collected
`HLT1`	two photons with $E_T > 15$ GeV	2.9 pb^{-1}
`HLT2`	two photons with $E_T > 17$ GeV	15 pb^{-1}
`HLT3`	one photon, one isolated supercluster both with $E_T > 17$ GeV	18 pb^{-1}

Table 6.1: High Level Trigger (HLT) requirements for the dataset used for the measurement. The isolation criteria applied to the supercluster for `HLT3` are less stringent than those used for event selection as given in table 6.5.

The transverse energy thresholds of the photons were chosen to be greater than those of the trigger paths of table 6.1, to ensure full trigger efficiency. Furthermore, asymmetric thresholds were chosen to allow for the comparison of the measured cross section with a solid theoretical prediction, as explained in section 1.2.2.

The requirements on the reconstructed photon transverse energy were chosen to be tighter than the trigger requirements by a few GeV in order to insure a fully efficient trigger. The gap in the pseudorapidity acceptance is chosen to avoid the transition region between the `ECAL` barrel and endcaps where the photon reconstruction performs worse than in the rest of the detector. Finally, the variable `ecalIso` measuring the electromagnetic energy deposited around a photon inside a radius $R_{out} = 0.4$ as depicted on figure 4.7 is used in this measurement for the statistical background subtraction. In order to avoid the contamination of the one photon's `ECAL` isolation surface by the supercluster of the other photon, a separation requirement in the (η, φ) plane $R > 0.45$ is added to the acceptance, where R is

	Variable	Acceptance
E_T (γ_1)	leading photon transverse energy	> 23 GeV
E_T (γ_2)	trailing photon transverse energy	> 20 GeV
$\|\eta(\gamma)\|$	photon pseudorapidity	$\in [0.00, 1.44] \cup [1.57, 2.5]$
R	(η, φ) distance between photons	> 0.45

Table 6.2: Kinematic acceptance for the diphoton cross section measurement. We call *leading* and *trailing* the photons with the highest and second highest transverse energies.

defined as

$$R(\gamma_1, \gamma_2) = \sqrt{(\eta_1 - \eta_2)^2 + (\varphi_1 - \varphi_2)^2} \tag{6.1}$$

Symbol	Definition	Equ.
$m_{\gamma\gamma}$	Invariant mass of the photon pair	(6.2)
$p_{T,\gamma\gamma}$	Transverse momentum of the photon pair	(6.3)
$\Delta\varphi_{\gamma\gamma}$	Difference in azimuthal angle between the two photons	(6.4)
$\cos\,\theta^*$	Cosine of the scattering angle in the Collins-Soper frame	(6.5)

Table 6.3: Differential cross section spectra. The mathematical definition of each variable is given in the equation listed in the third column.

The differential cross section is measured as a function of four variables shown in table 6.3 chosen either because of their sensitivity to the QCD processes, as will be described in section 6.5.3 for the variable $\Delta\varphi_{\gamma\gamma}$, or because of their discriminating power in a diphoton channel Higgs boson search. The invariant mass of the two photons $m_{\gamma\gamma}$ is defined in equation (6.2) as a function of the photon energies and θ_{12}, the angle between the two photons. The transverse momentum $p_{T,\gamma\gamma}$ of the diphoton system is expressed in equation (6.3) as a function of the two photon transverse momenta and the difference in azimuthal angle between the two photons which in turn is defined in equation (6.4) to be comprised in $[0, \pi]$. Finally the cosine of the scattering angle in the Collins-Soper frame $\cos\,\theta^*$ is expressed as a function of the difference of the photon rapidities or pseudorapidities in equation (6.5).

$$m^2_{\gamma\gamma} = 2 \cdot E_1 \cdot E_2 \cdot (1 - \cos\,\theta_{12}) \tag{6.2}$$

$$p^2_{T,\gamma\gamma} = p^2_{T,\gamma_1} + p^2_{T,\gamma_2} + 2 \cdot p_{T,\gamma_1} \cdot p_{T,\gamma_2} \cdot \cos\Delta\varphi_{\gamma\gamma} \tag{6.3}$$

$$\Delta\varphi_{\gamma\gamma} = \varphi_1 - \varphi_2 \in [0, \pi] \tag{6.4}$$

$$\cos\,\theta^* = \left| \tanh\left(\frac{y_1 - y_2}{2}\right) \right| = \left| \tanh\left(\frac{\eta_1 - \eta_2}{2}\right) \right| \tag{6.5}$$

In order to equalise the statistical uncertainties on the measured spectra, a small number of unequally sized bins, presented in table 6.4, were chosen to measure the spectra of the differential cross sections.

Symbol	Unit	Number	Bins edges
$m_{\gamma\gamma}$	GeV	10	[0,30] [30,40] [40,45] [45-55] [55,65] [65,80] [80,100] [100,140] [140,200] [200,300]
$p_{T,\gamma\gamma}$	GeV	10	[0,4] [4,6] [6,8] [8,12] [12,18] [18,30] [30,40] [40,50] [50,80] [80,180]
$\Delta\varphi_{\gamma\gamma}$	π	9	[0.0,0.2] [0.2,0.4] [0.4,0.6] [0.6,0.8] [0.8,0.88] [0.88,0.92] [0.92,0.95] [0.95,0.98] [0.98,1.00]
$\cos\theta^*$		5	[0,0.2] [0.2,0.4] [0.4,0.6] [0.6,0.8] [0.8,1.0]

Table 6.4: Binning of the differential cross section spectra

Events for the measurement entering the acceptance defined in table 6.2 were selected using the background rejection variables described in section 4.1.2 and detailed in table 6.5.

Selection	on γ_1 and γ_2
`trigIso`	< 0.3
`H/E`	< 0.05
$\sigma_{i\eta i\eta}$	< 0.011 (0.03)
`hcalIso`	< 2 (4) GeV
`trackIso`	< 2 (4) GeV
`pixelMatch`	`veto`
n(*impinging tracks*)	$= 0$

Table 6.5: Event selection for the sample used for the cross section measurement. The variable `trigIso` is defined in equation (5.3) and is used to ensure that the selections on reconstructed quantities are tighter than the trigger requirements.

The isolation requirement on the HLT3 trigger path includes a selection requirement on the variable `trigIso`, defined in equation (5.3) and a selection requirement on `trigIso` is included in the event selection in order not to impact the trigger efficiency. As shown in section 5.1.2, even though the variables `trigIso` and `ecalIso` have similar definitions, a selection on `trigIso` < 0.2 is equivalent to a selection on $\sum_{i\in C(\texttt{trig})} E_T^i < 10$ GeV and so has a negligible impact on the signal distribution. Figures 5.3 and 5.4 show that loosely selecting on `trigIso` still allows for the use of the statistical signal and background discrimination based on `ecalIso` as the distributions of `ecalIso` in the data reach up further than 6 GeV and allow for a differentiation of signal and background.

Section 6.1.3 will shed light on the method applied to perform the measurement.

6.1.3 Computing the cross section

As explained in section 4.2, a sample of photon candidates selected by requirements on the background rejection variables of section 4.1.2 still contains a proportion of non-prompt photons. Thus, after applying the requirements of table 6.5, each photon of the selected diphoton candidate can either be an isolated prompt photon or a non-prompt photon leading to the signal and background definitions of table 6.6.

Nature	γ_1	γ_2
Signal	Prompt	Prompt
Background	Prompt	Non-prompt
	Non-prompt	Prompt
	Non-prompt	Non-prompt
	Electron	Electron

Table 6.6: Definition of signal and background in the diphoton cross section measurement. As electrons can be misidentified as photons, di-electron Drell-Yan events are one of the background species.

As explained in section 4.1.1, electrons can be misidentified as photons. Di-electron Drell-Yan events contaminate the selected sample. They are not rejected by the fitting method described in section 6.2.1 because their electron clusters are isolated in `ECAL`, but will be corrected as explained in section 6.3.2.

The photons are not ordered according to any criteria, transverse energy or pseudorapidity for instance. Throughout the rest of this chapter events with two non-prompt photons will be denoted *nn*, events with one prompt and one non-prompt photon *pn* and events with two prompt photons *pp*. Given this notation, the integrated cross section that we want to measure is given by

$$\sigma(\gamma\gamma + X) = \frac{\mathcal{U}(N^{pp})}{\epsilon \cdot \mathcal{L}} \tag{6.6}$$

where $\mathcal{U}(N^{pp})$ is the number of prompt diphotons corrected for potential fit biases and detector resolution effects, N^{pp} is henceforth called *prompt diphoton yield*, ϵ the diphoton selection efficiency as defined and computed in section 6.3.4 and $\mathcal{L}$ the integrated luminosity of the selected sample. The differential cross section in a given bin of the observable Ω is given by

$$\frac{d\sigma}{d\Omega}(\gamma\gamma + X) = \frac{\mathcal{U}(N^{pp})}{\epsilon \cdot \Delta\Omega \cdot \mathcal{L}} \tag{6.7}$$

where $\Delta\Omega$ is the width of the considered bin as given in table 6.4.

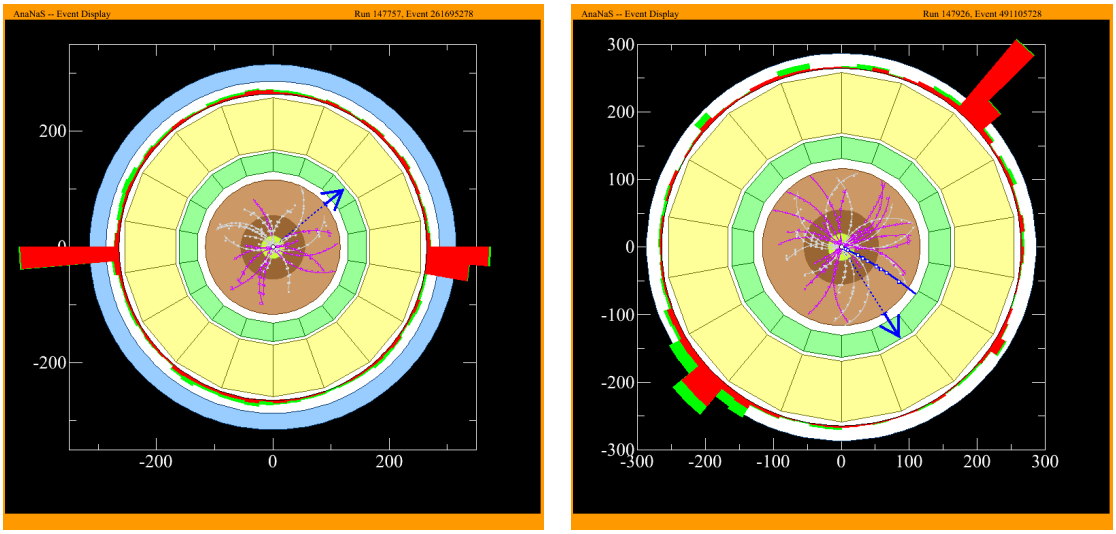
Section 6.2 discusses how the prompt diphoton yield N^{pp} is determined.

6.2 Prompt Diphoton Yields

6.2.1 Signal extraction fit

The statistical method used for the discrimination of prompt isolated photons and non-prompt photons from jets based on the electromagnetic isolation variable `ecalIso` introduced in section 4.1.2 is applied here to the measurement of the prompt isolated diphoton cross section. It exploits the fact that prompt photons are likely to have small values of `ecalIso` as opposed to non-prompt photons. Figure 6.3 shows event displays of diphotons likely to be made of two prompt and two non-prompt photons.

As there are now three species of events, the signal *pp* and two background species *pn* and *nn*, the likelihood function $\mathcal{L}$ given in equation (6.8) and used in the maximum likelihood fit is expressed as a



(a) Candidates likely to be prompt photons (b) Candidates likely to be non-prompt photons

Figure 6.3: Event displays from 2010 data taken by CMS, showing two photon pair candidates. The diphoton candidate shown on (a) has $\mathcal{I}_1 = \mathcal{I}_2 = 0$ GeV and is therefore likely to contain two prompt photons whereas the diphoton candidate shown on (b) has $\mathcal{I}_1$ and $\mathcal{I}_2$ greater than 4 GeV, therefore likely to contain two non-prompt photons.

function of the three yields and probability density functions (pdf). The two-dimensional pdfs depend on $\mathcal{I}_1$ and $\mathcal{I}_2$, the `ecalIso` values of the two photons.

$$\mathcal{L}(N_{pp}, N_{pn}, N_{nn}) = \frac{e^{-(N_{pp}+N_{pn}+Nnn)}}{N!} \prod_i^N [\ N_{pp} \cdot f_{pp}(\mathcal{I}_1^i, \mathcal{I}_2^i) + \ N_{pn} \cdot f_{pn}(\mathcal{I}_1^i, \mathcal{I}_2^i) + \ N_{nn} \cdot f_{nn}(\mathcal{I}_1^i, \mathcal{I}_2^i)\ ] \tag{6.8}$$

It has been checked that $\mathcal{I}_1$ and $\mathcal{I}_2$ are independent in all of the three species. Figure 6.4 shows the two-dimensional distribution of $\mathcal{I}_1$ versus $\mathcal{I}_2$ for background enriched data events, where one of the photons is required to have at least one *impinging track*.
We therefore express the common probability of $\mathcal{I}_1$ and $\mathcal{I}_2$ as the product of the probability of each

$$\mathcal{P}(\mathcal{I}_1 \texttt{ and } \mathcal{I}_2) = \mathcal{P}(\mathcal{I}_1) \cdot \mathcal{P}(\mathcal{I}_2) \tag{6.9}$$

In the context of the likelihood function of equation (6.8) this allows us to express the two-dimensional pdfs as a function of the one-dimensional pdfs as defined in chapter 4.

$$f_{pp}(\mathcal{I}_1, \mathcal{I}_2) = f_p(\mathcal{I}_1) \times f_p(\mathcal{I}_2) \tag{6.10}$$
$$f_{pn}(\mathcal{I}_1, \mathcal{I}_2) = f_p(\mathcal{I}_1) \times f_n(\mathcal{I}_2) + f_n(\mathcal{I}_1) \times f_p(\mathcal{I}_2) \tag{6.11}$$
$$f_{nn}(\mathcal{I}_1, \mathcal{I}_2) = f_n(\mathcal{I}_1) \times f_n(\mathcal{I}_2) \tag{6.12}$$

The one-dimensional `ecalIso` pdfs of equations (6.10) to (6.12) are

- $f_p(\mathcal{I})$ – the pdf for prompt photons, extracted with the *random cone* technique
- $f_n(\mathcal{I})$ – the pdf for non-prompt photons, extracted using the *impinging track method*.

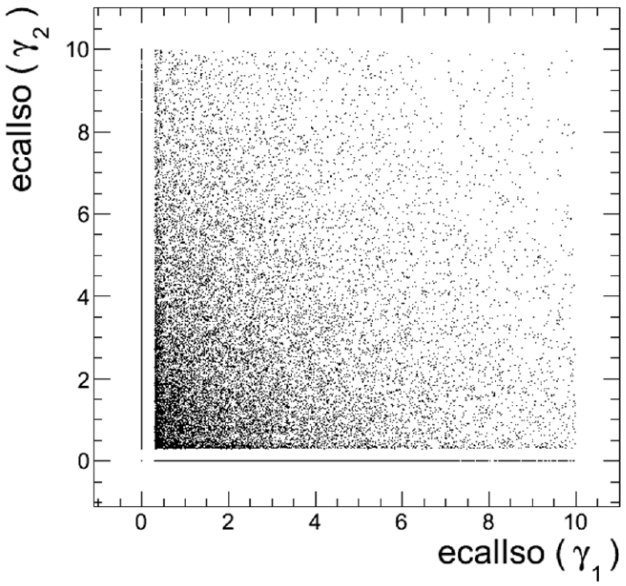
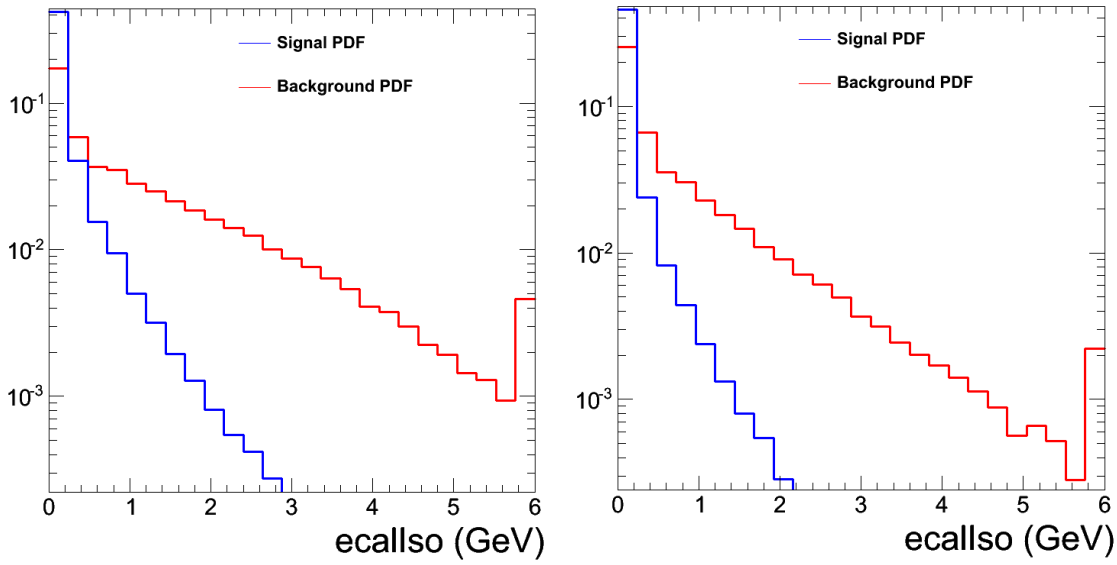


Figure 6.4: Two-dimensional distribution of `ecalIso`(γ_1) versus `ecalIso`(γ_2) for a background-enriched data sample, where one of both photons is required to have at least one *impinging track*. The correlation coefficient is 9.3 %.



(a) Barrel (b) Endcaps

Figure 6.5: Shapes of the pdfs extracted using all single photon candidates. Section 6.2.2 describes how these pdfs are used for the fit in each bin of the differential measurement. The last bin contains the overflow.

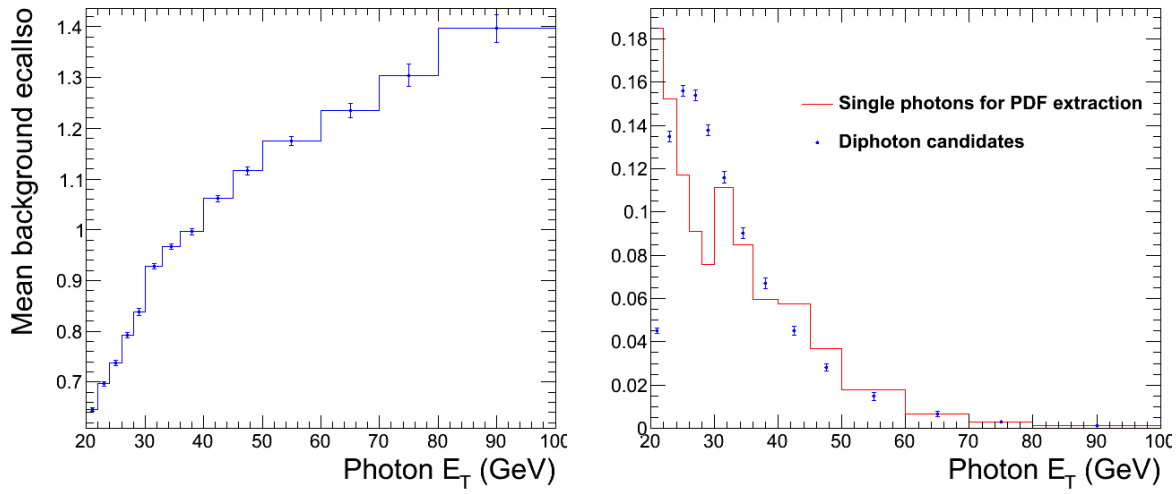
Figures 6.5(a) and 6.5(b) show the `ecalIso` distributions obtained when applying the signal and back-

ground pdf extraction techniques to all single photon candidates in the barrel and the endcaps respectively. We split the sample of diphotons passing the event selection requirements into three categories: events with both photons in the barrel, events with both photons in the endcaps and events with one photon in the barrel and one photon in the endcaps. Throughout this section and the next, the measurement in the first category will be highlighted and the results in all categories will be given in section 6.5.

6.2.2 Extraction of the pdfs

It has to be noted that the shape of the signal and background pdfs depends on the properties of the single photon events used for their extraction.

As explained in section 4.1.2 and shown on figure 4.8, the variable `ecalIso` discriminates between prompt and non-prompt photons because of the leakage of the non-prompt photon energy into the isolation surface. Thus for non-prompt photons the average `ecalIso` value obtained with the *impinging track* method increases with the the transverse energy of the photon. The `ecalIso` value extracted at $\Delta\varphi = \pi/2$ of the photon with the *random cone* method does not depend on its transverse energy. The average energy deposited in `ECAL` increases with the pileup of the event and therefore both the signal and the background average `ecalIso` values increase with the number of reconstructed vertices in the event. Finally, both average `ecalIso` values decrease with increasing pseudorapidity. These dependences appear on figures 6.6(a), 6.7 and 6.8.

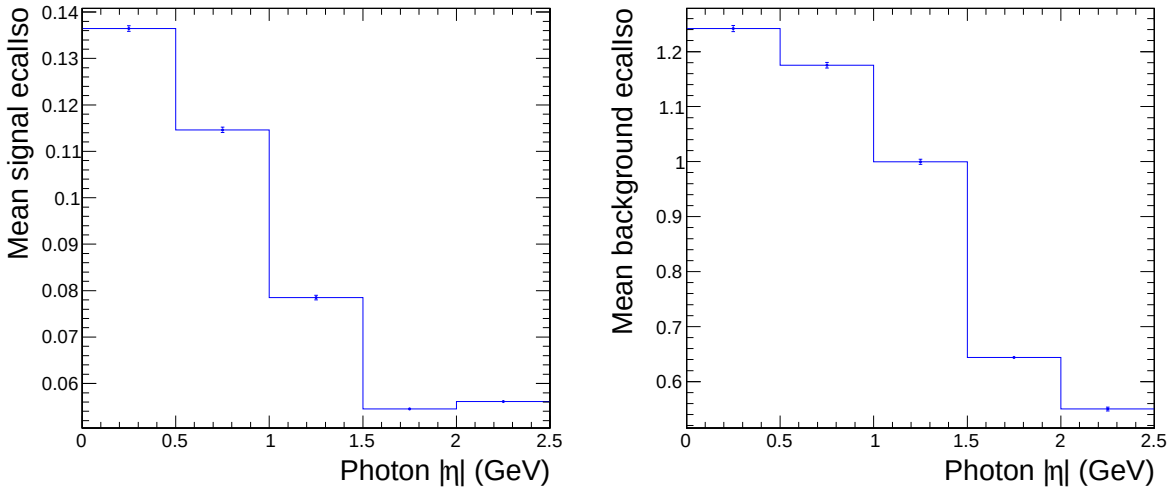


(a) `ecalIso` average vs. E_T

(b) E_T comparison

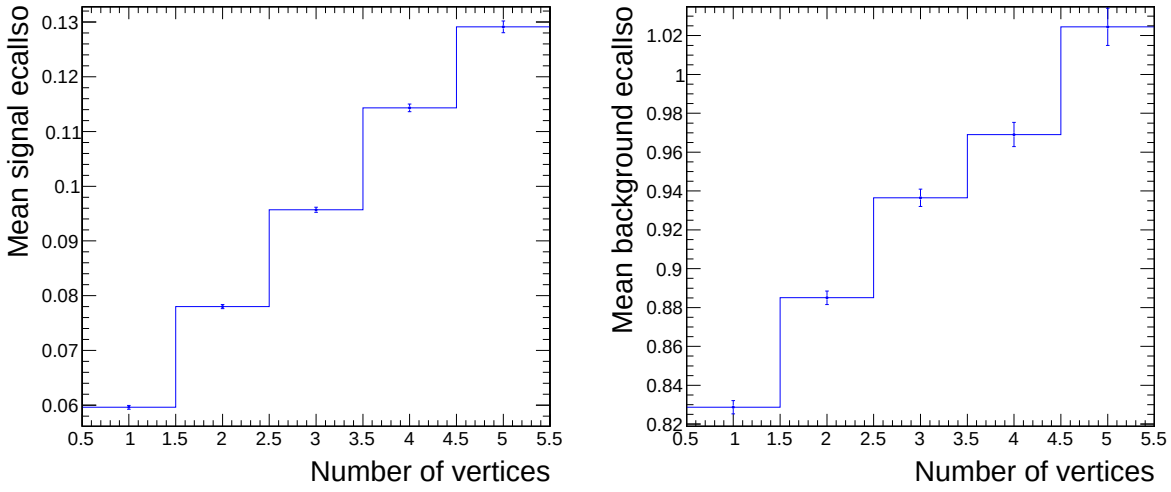
Figure 6.6: Illustration of the dependence of the average `ecalIso` on the photon E_T. Plot (a) shows the increase of the average `ecalIso` value in the background pdf extracted with the *impinging track* method versus the photon E_T and plot (b) the different E_T distribution between diphoton candidates and single photons used for the pdf extraction. We note that the E_T spectrum of single photons is shaped by the trigger prescales.

The single photon sample on which the pdfs are extracted and the diphoton sample used for the measurement are selected with different trigger paths, the prescale factors of which have changed over the period of data-taking. This induces different E_T, as shown on figure 6.6 (b), and N_V distributions. We note that the E_T spectrum of single photons on figure 6.6 (b) is shaped by the trigger prescales: the single photon HLT path with a 20 GeV E_T threshold was prescaled by factors of 5 and 26 compared with the single photon HLT paths with E_T threshold at 30 and 50 GeV respectively. Selecting on the



(a) Signal (b) Background

Figure 6.7: Average `ecalIso` value versus photon pseudorapidity. Plot (a) shows the average `ecalIso` value of the signal pdf extracted with the *random cone* technique and plot (b) that of the background pdf extracted with the *impinging track* method versus the photon $|\eta|$.



(a) Signal (b) Background

Figure 6.8: Average `ecalIso` value versus the number of reconstructed vertices. Plot (a) shows the average `ecalIso` value of the signal pdf extracted with the *random cone* technique and plot (b) that of the background pdf extracted with the *impinging track* method versus N_V.

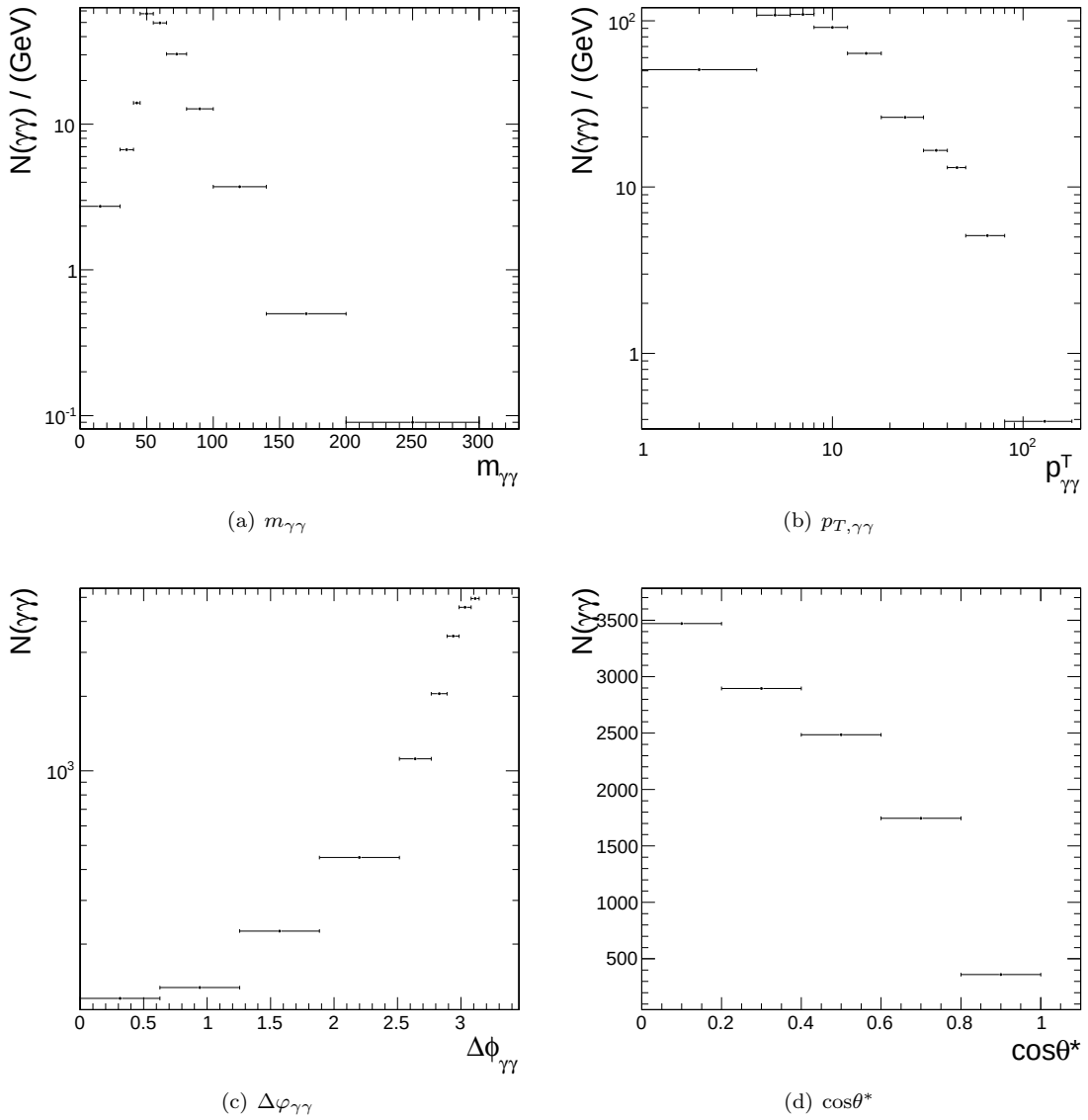
diphoton observables of table 6.3 can further bias the E_T and $|\eta|$ spectra. Indeed, equation (6.2) shows that bins of high $m_{\gamma\gamma}$ contain photons with high transverse energy and equation (6.5) demonstrates that bins of high $\cos\theta^*$ contain photons with high pseudorapidities. This results in significant differences in the E_T, $|\eta|$ and N_V distributions of the single photons used for the pdf extraction and the diphoton candidates.

Therefore, in every bin of the differential measurement, we extract a particular signal and background pdf by assigning a weight to the single photon events in order to match their E_T, $|\eta|$ and N_V distributions with the ones of a large diphoton sample. Candidates of this sample were selected with all the requirements of table 6.5 but accepting photon candidates with up to one *impinging track*. In the case of figure 6.6 (b) a single photon event with $E_T = 20$ GeV would be assigned a low weight of about 0.2 and a single photon event with $E_T = 30$ GeV a high weight of about 1.8. Thus a different set of signal and background pdfs is used in each bin of the observables.

6.2.3 Signal yields

By performing the maximum likelihood fit described in section 6.2.1 on the `ecalIso` distributions of the selected diphoton candidates, of which four spectra are shown on figure 6.9, we extract the prompt diphoton yields in every bin. These yields from the fit need to be corrected before being used for the cross section calculation, as explained in sections 6.3.1 and 6.3.3 and are therefore called *raw yields*.

The projections on $\mathcal{I}_1$ and $\mathcal{I}_2$ of the `ecalIso` fit in two bins of the differential spectra, the bin of the $\Delta\varphi_{\gamma\gamma}$ spectrum corresponding to $\Delta\varphi_{\gamma\gamma} \in [0.92\pi, 0.95\pi]$ and the bin of the $m_{\gamma\gamma}$ spectrum corresponding to $m_{\gamma\gamma} \in [55, 56]$ GeV/c^2, are shown on plots 6.10 and 6.11. The proportion of prompt diphotons in every bin of the four differential distributions is shown on figure 6.12 and the fit result, the raw yield spectra, on figure 6.13.



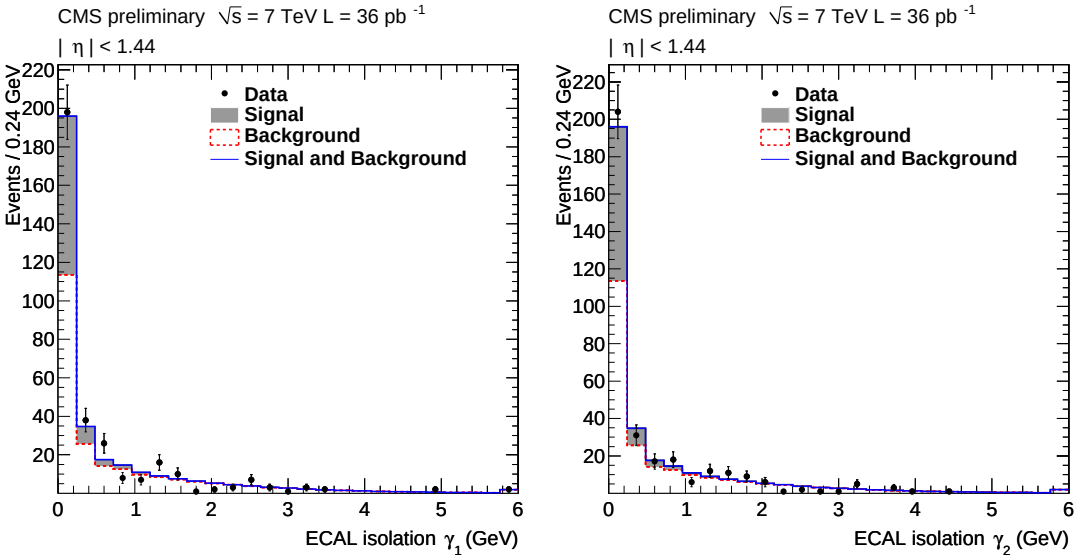
(a) $m_{\gamma\gamma}$

(b) $p_{T,\gamma\gamma}$

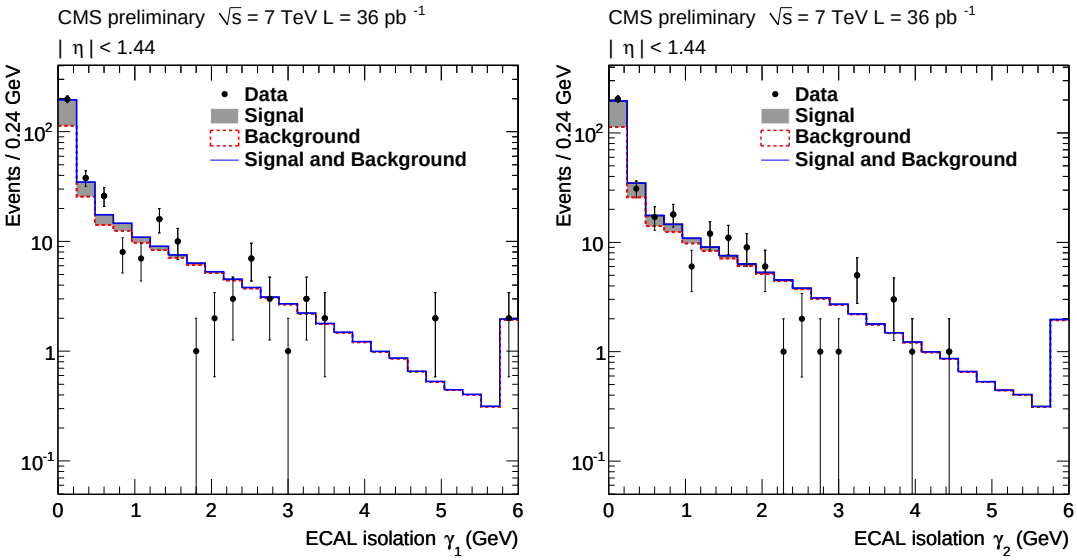
(c) $\Delta\varphi_{\gamma\gamma}$

(d) $\cos\theta^*$

Figure 6.9: Plots showing the number of selected diphoton candidates with both photons in the barrel on which the `ecalIso` fit is performed, in each bin of the differential distributions.

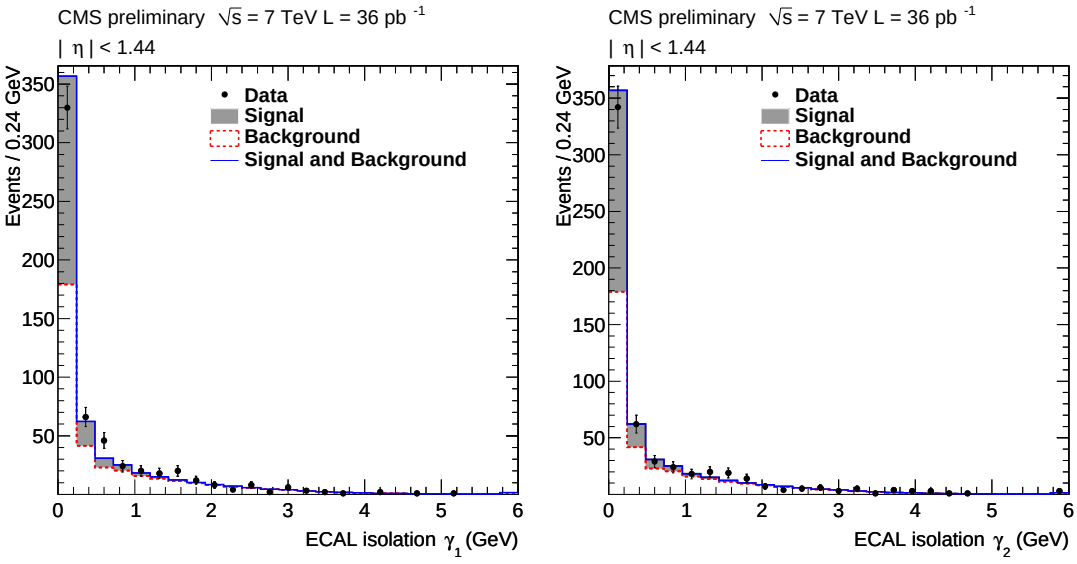


(a) Linear scale

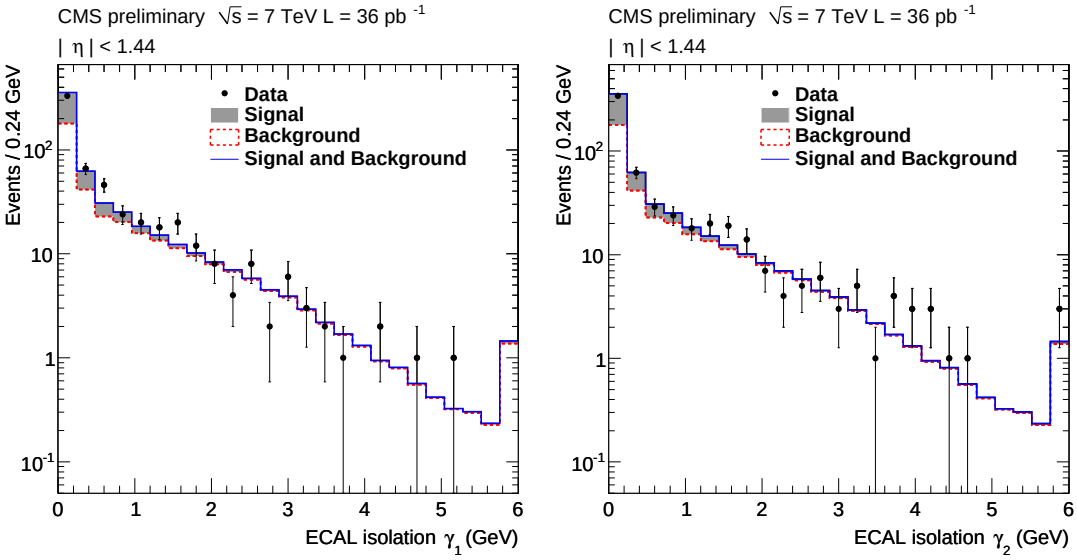


(b) Logarithmic scale

Figure 6.10: Fit result in the bin of the $\Delta\varphi_{\gamma\gamma}$ spectrum corresponding to $\Delta\varphi_{\gamma\gamma} \in [0.89, 0.92]\pi$, for photons in the barrel. The data with superimposed signal and background (blue) and background only (dashed red) pdfs are shown in (a) linear and (b) logarithmic scales. The last bin contains the overflow.

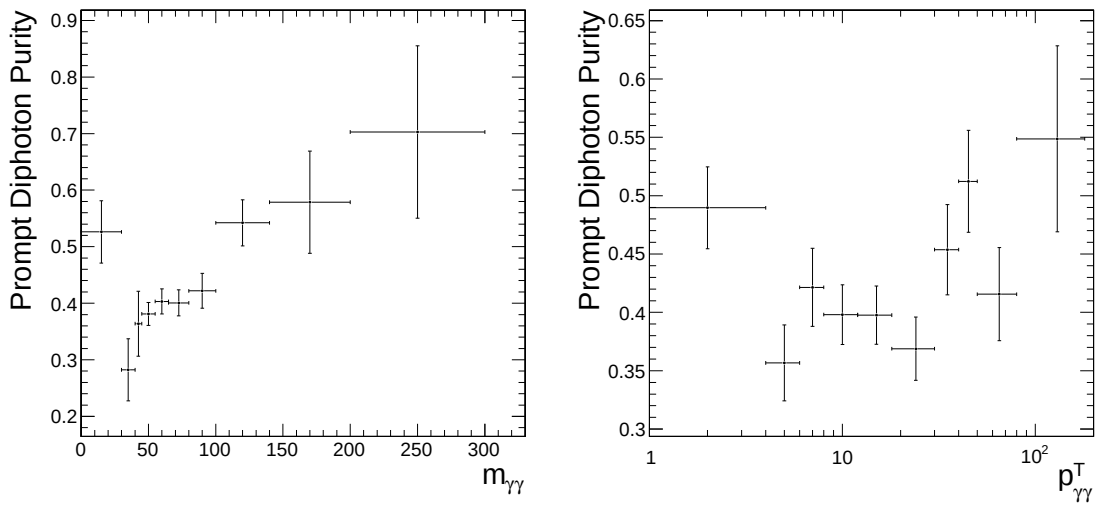


(a) Linear scale

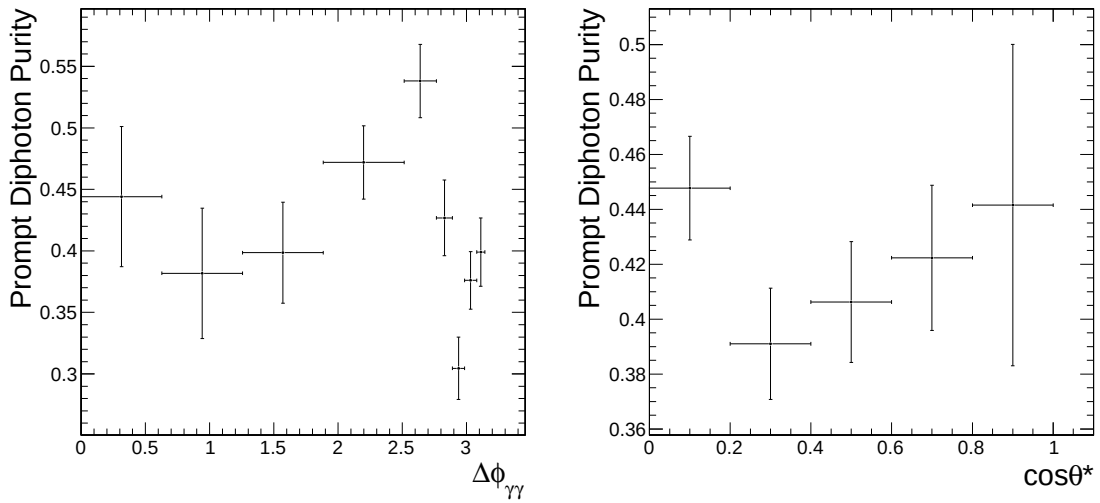


(b) Logarithmic scale

Figure 6.11: Fit result in the bin of the $m_{\gamma\gamma}$ spectrum corresponding to $m_{\gamma\gamma} \in [45, 55]$ GeV/c^2, for photons in the barrel. The data with superimposed signal and background (blue) and background only (dashed red) *pdfs* are shown in (a) linear and (b) logarithmic scales. The last bin contains the overflow.

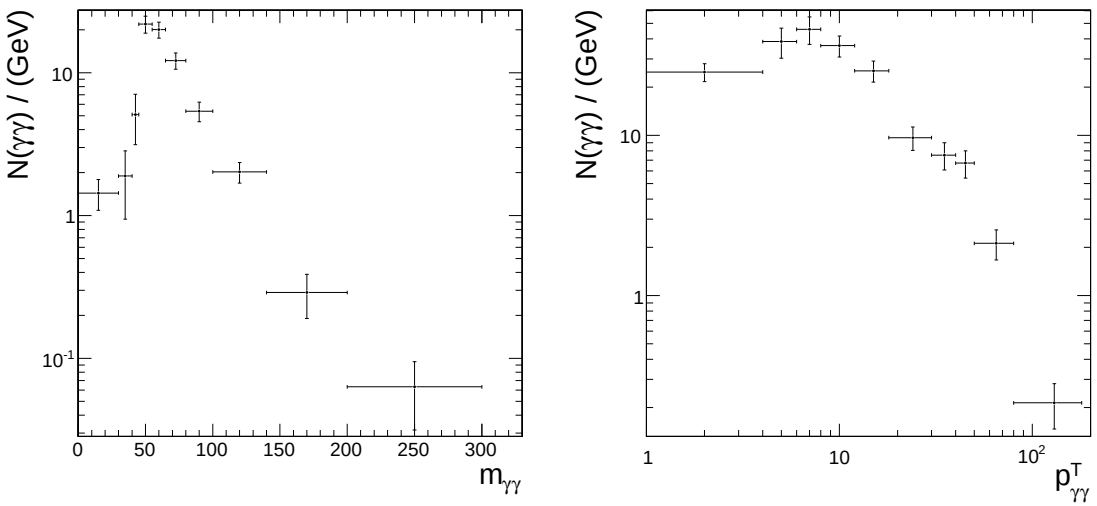


(a) $m_{\gamma\gamma}$ (b) $p_{T,\gamma\gamma}$

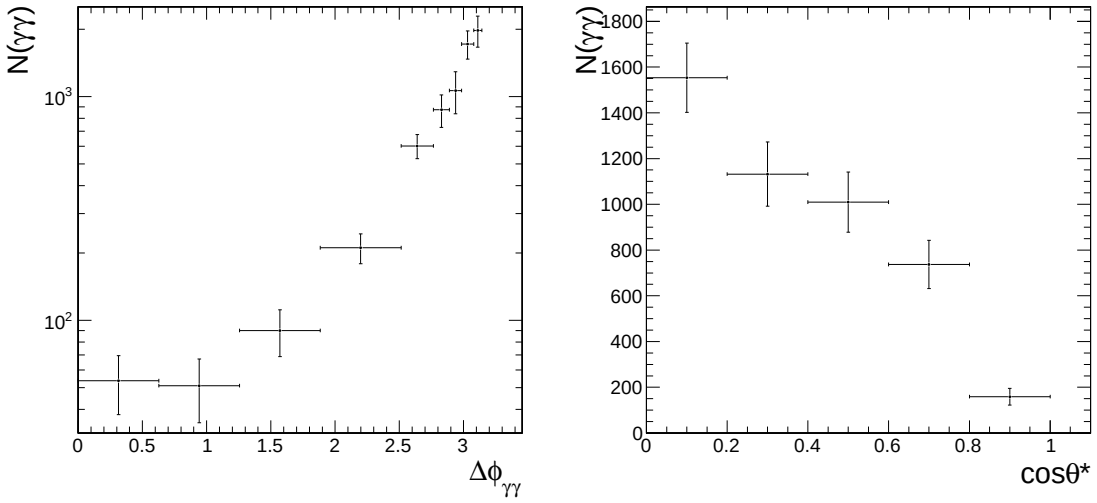


(c) $\Delta\varphi_{\gamma\gamma}$ (d) $\cos\theta^*$

Figure 6.12: Plots showing the prompt diphoton purity as given by the `ecalIso` fit, in each bin of the differential distributions.



(a) $m_{\gamma\gamma}$ (b) $p_{T,\gamma\gamma}$



(c) $\Delta\varphi_{\gamma\gamma}$ (d) $\cos\theta^*$

Figure 6.13: Plots showing the prompt diphoton raw yield as given by the `ecalIso` fit, in each bin of the differential distributions.

6.3 Cross Section Measurement

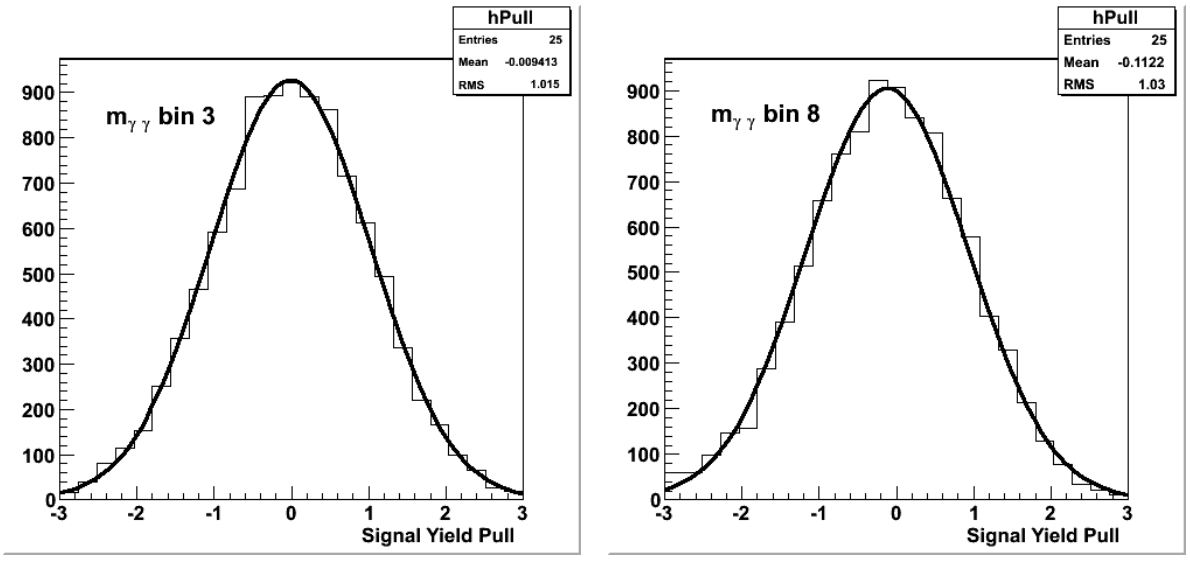
6.3.1 Bias correction

Given that for a small sample the diphoton yield extracted with the maximum likelihood fit is a biased estimator of the signal yield [86], this bias is estimated by performing toy Monte Carlo experiments and correct the raw yields by the estimated amount.

Three sets of toy Monte Carlo events are generated (one for each species pp, pn and nn) with the `ecalIso` pdfs defined in equations (6.10) to (6.12). The generation is performed using Gaussian distribution with mean values set to the yields extracted from the fit, N_{pp}, N_{pn} and N_{nn}, and widths set to the fit errors, δN_{pp}, δN_{pn} and δN_{nn}. A maximum likelihood fit with the same `ecalIso` pdfs is then performed on the sample thus obtained. The variable X_{pull} defined as

$$X_{\text{pull}} = \frac{N_{pp}^{\text{toy fit}} - N_{pp}}{\delta N_{pp}} \tag{6.13}$$

follows a standard normal distribution for an unbiased fit. The toy Monte Carlo experiment is repeated 10000 times and the distribution of X_{pull} fitted with a Gaussian function to extract its width σ_{pull} and mean value μ_{pull}. Figure 6.14 shows two examples of such a fitted pull distribution in bins of the $m_{\gamma\gamma}$ spectrum.



(a) Bin: 45 GeV/c^2 < $m_{\gamma\gamma}$ < 55 GeV/c^2

(b) Bin: 100 GeV/c^2 < $m_{\gamma\gamma}$ < 140 GeV/c^2

Figure 6.14: Fit pull distributions for two bins of the $m_{\gamma\gamma}$ spectrum with both photons in the barrel. Plot (a): in this bin a high number of candidates (576) leads to a small bias (-0.9 % of the statistical error). Plot (b): in this bin a small number of candidates (30) leads to a higher bias (-11 % of the statistical error).

In each bin of the differential measurement the raw prompt diphoton yield N_{pp} is corrected by the computed bias, $\mu_{\text{pull}} \times \delta N_{pp}$. Figure 6.15 shows that the impact of the bias correction on the four differential spectra goes from 0 to 11 % for photons in the barrel. It demonstrates that the bias of the diphoton yield extracted with the maximum likelihood fit increases with decreasing number of fitted events.

Finally, the fits on the integrated samples of diphotons are hardly biased at all and μ_{pull} amounts to less than 1 % of the statistical error.

6.3.2 Electron contamination

Electrons are rejected by the veto on the `pixelMatch` variable as well as by the veto on the `impinging tracks`. Even though the rejection power of these requirements on electron channel Drell-Yan events is is very high, 99.9 %, the Drell-Yan cross section is higher than the diphoton cross section by two orders of magnitude and this leads to a non-negligible contamination of the diphoton sample with dielectron events in the bin around the Z boson peak. From section 4.2.3 we gather that the `ecalIso` distribution of isolated photons and isolated electrons is similar and that dielectron events passing the diphoton selection requirements will be identified as prompt-prompt events by the maximum likelihood fit.

Using the next-to-leading simulation of the Drell-Yan process performed with `POWHEG` [99] that was shown in [89] to be in good agreement with the measurement of the cross section of $Z \to ee$, we compute the rate of Drell-Yan events passing the diphoton selection requirements. By scaling it to the integrated luminosity of the sample used for this measurement we obtain the contamination rate in each bin as shown on figure 6.16.

This contamination, which is negligible in most bins with the notable exception of the bin of the $m_{\gamma\gamma}$ spectrum corresponding to the Z boson peak, is subtracted bin-by-bin. The combined uncertainty on the cross section measured in [89] and the rejection efficiency of the selection requirements lead to an uncertainty of 5 % on the number of dielectron contaminating the diphoton sample. Given that the contamination is small, this leads to an uncertainty of $\delta N/N < 0.25\%$ on the the diphoton yield which is negligible compared to other systematic uncertainties, which are of the order of 8 % as shown in section 6.4. The impact of that correction on the integrated cross section is smaller than one percent.

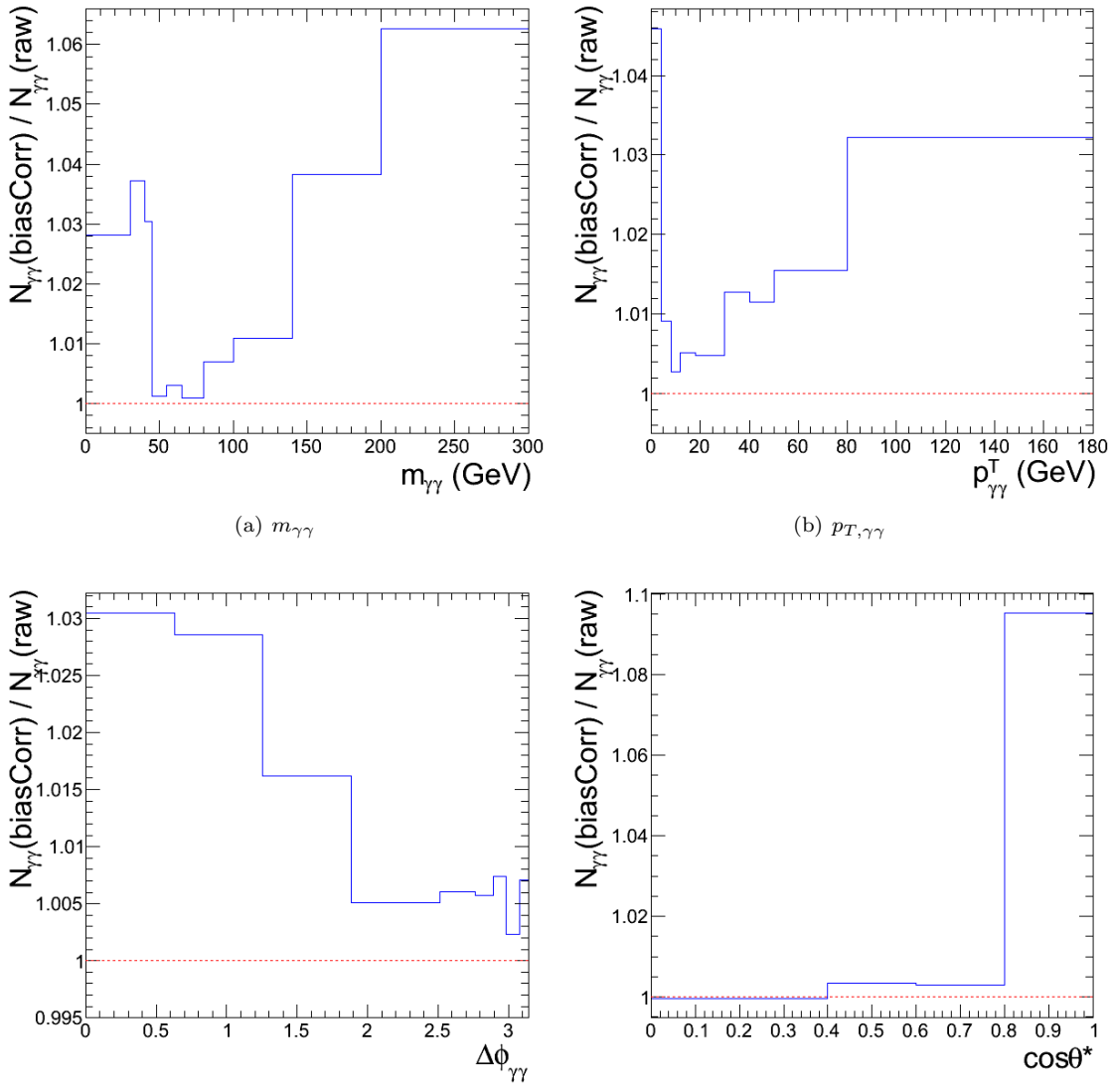
6.3.3 Unfolding

As shown in section 5.3.1, the four measured differential distributions are the combination of the physical spectra and the CMS detector resolution. If $\underline{F}^{\Omega} = \{F_i\}$ is the binned measured and $\underline{G}^{\Omega} = \{G_i\}$ the binned physical spectrum of the variable $\Omega \in \{m_{\gamma\gamma}, p_{T,\gamma\gamma}, \Delta\varphi_{\gamma\gamma}, \cos\theta*\}$ they are related by the *response matrix* $\mathcal{R}^{\Omega}$ via

$$\underline{F}^{\Omega} = \mathcal{R}^{\Omega}\underline{G}^{\Omega} \tag{6.14}$$

The response matrices given in figures 6.17 and 6.18 are determined from a sample of prompt diphoton events simulated with `PYTHIA`. The excellent performance of the CMS `ECAL` in terms of energy and position resolution, leads to nearly diagonal matrices which, as is shown in [100], can be simply inverted to solve equation 6.14 without the need for an additional term to insure the smoothness of the obtained physical spectrum.

The bin-per-bin impact of the unfolding is shown on figure 6.19, it is negligible for the angular variables $\Delta\varphi_{\gamma\gamma}$ and $\cos\theta^*$. The effect on the energy-related spectra $m_{\gamma\gamma}$ and $p_{T,\gamma\gamma}$ is greater, particularly in the steep slopes as for instance the one at $m_{\gamma\gamma} = 40$ GeV/c^2 due to the kinematic selections.



(a) $m_{\gamma\gamma}$

(b) $p_{T,\gamma\gamma}$

(c) $\Delta\varphi_{\gamma\gamma}$

(d) $\cos\theta^*$

Figure 6.15: Distribution of the ratio of prompt diphoton yields before and after bias correction. Contrasting these distributions with the ones shown on figure 6.9 demonstrates the dependence of the inherent fit bias on the number of candidates the fit is performed on.

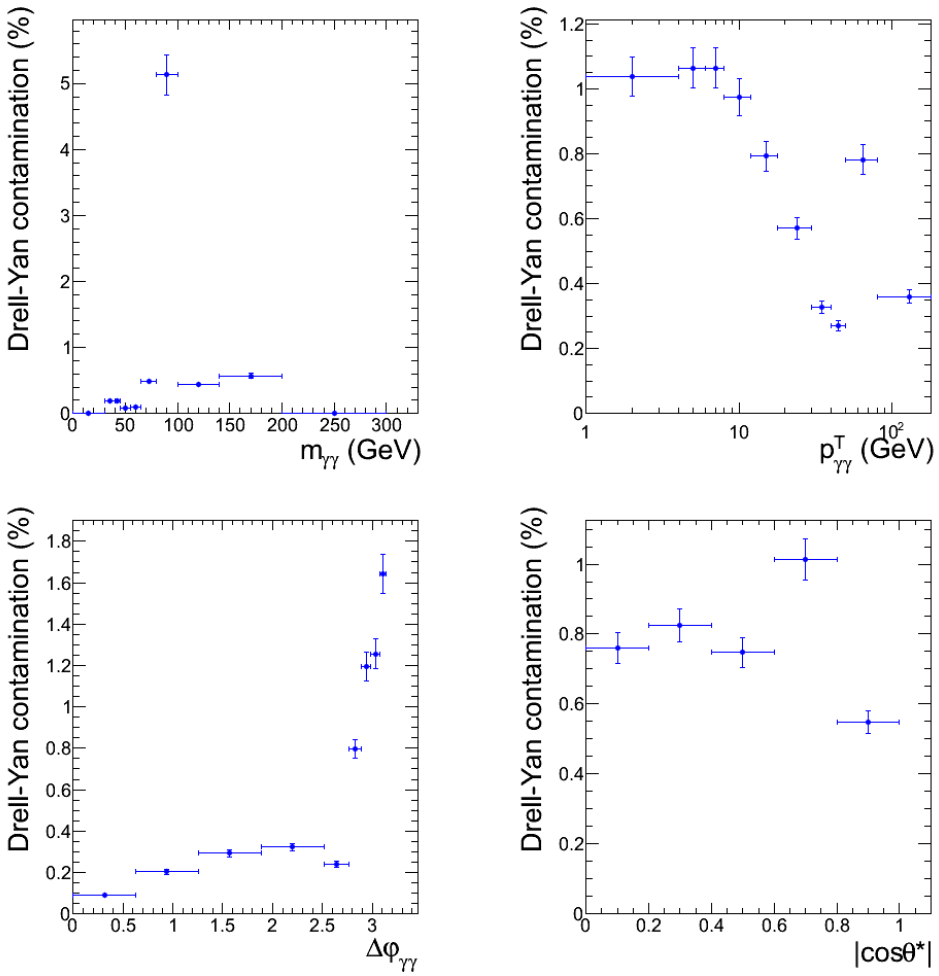
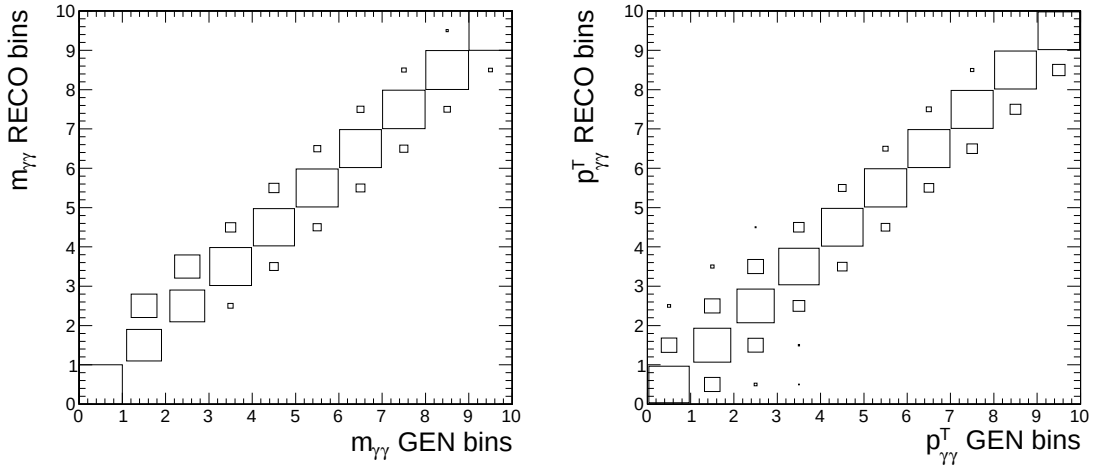
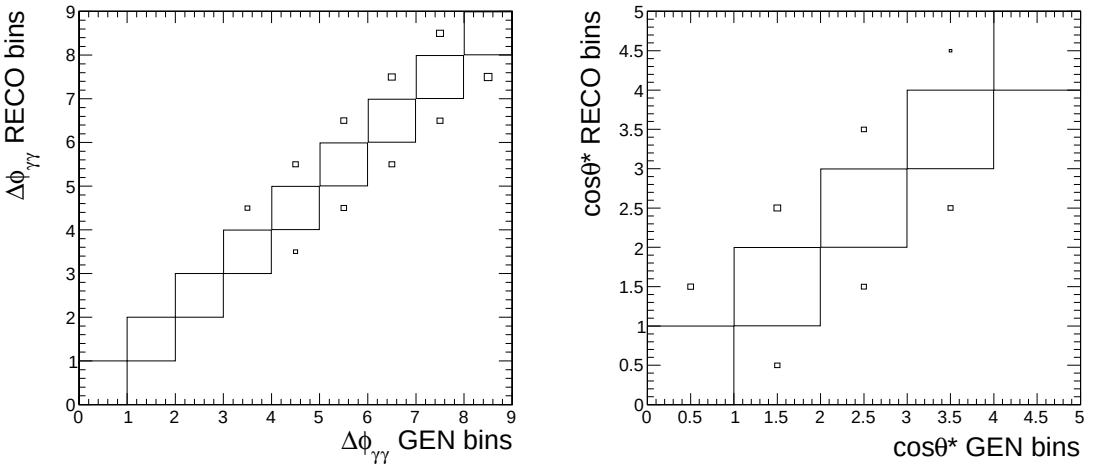


Figure 6.16: Bin-by-bin contamination of the prompt diphoton yield with electron channel Drell-Yan events for photons in the barrel. The e^+e^- contribution to most bins is negligible with the notable exception of the bin around 90 GeV/c^2 in the invariant mass spectrum, corresponding to the Z boson peak.



(a) $m_{\gamma\gamma}$ (b) $p_{T,\gamma\gamma}$



(c) $\Delta\varphi_{\gamma\gamma}$ (d) $\cos\theta^*$

Figure 6.17: Detector response matrices for the four differential spectra, determined from prompt diphoton events simulated with PYTHIA. The excellent detector energy and position resolution of the CMS ECAL leads to nearly diagonal matrices.

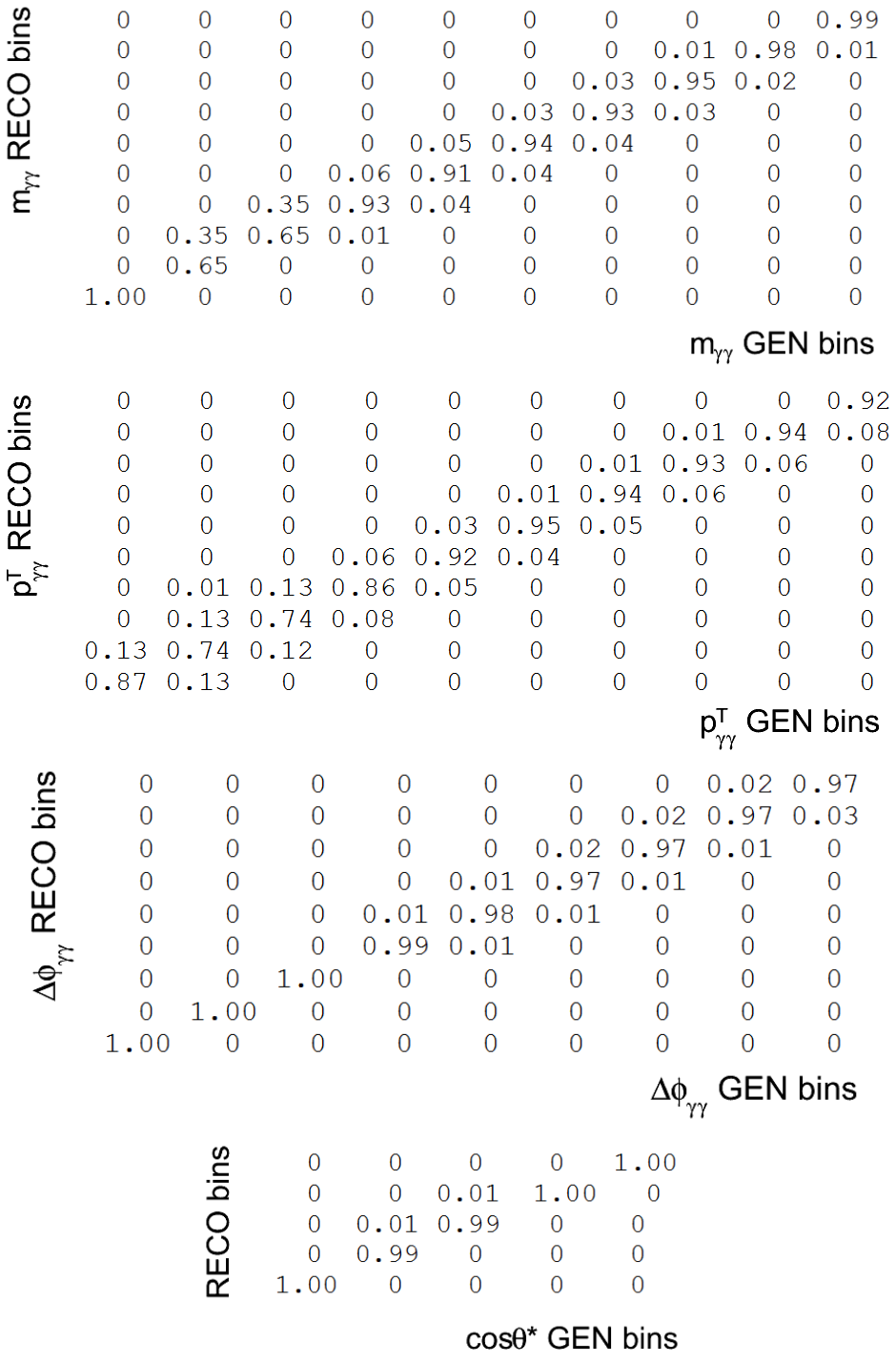
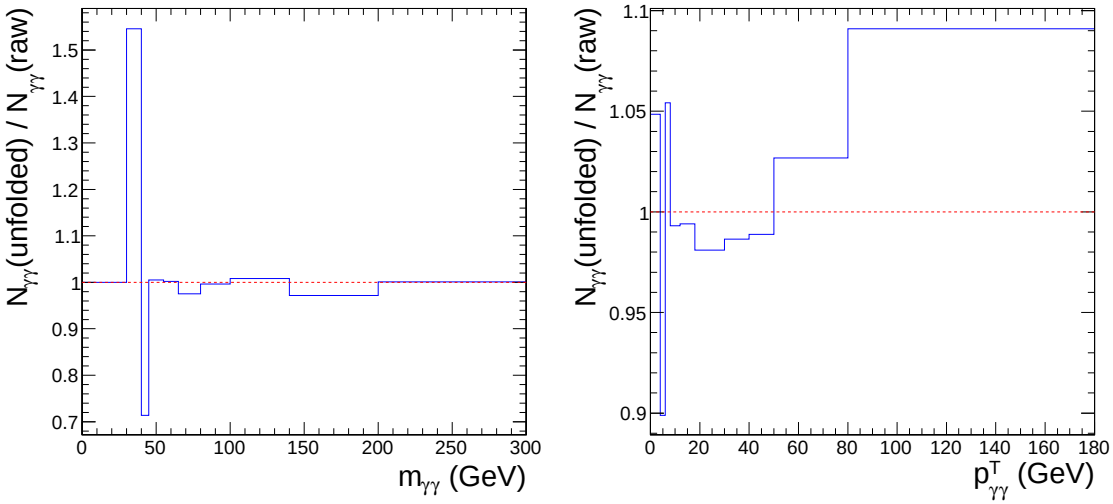
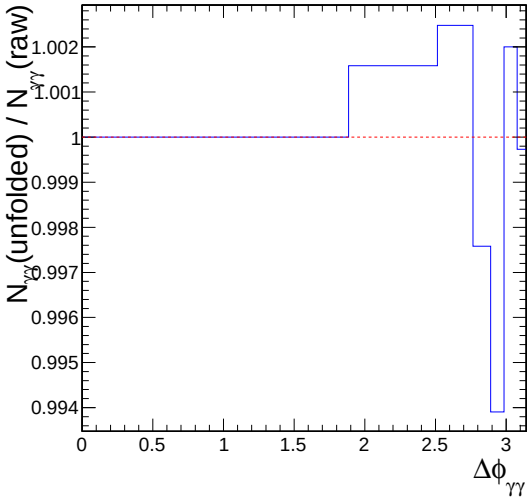


Figure 6.18: Detector response matrices for the four differential spectra, determined from prompt diphoton events simulated with PYTHIA. The excellent detector energy and position resolution of the CMS ECAL leads to nearly diagonal matrices.

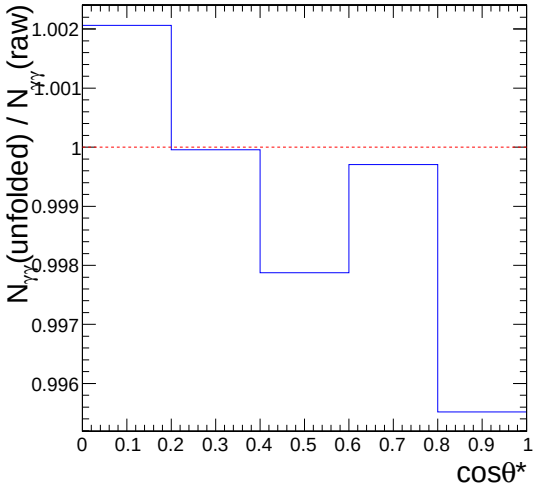


(a) $m_{\gamma\gamma}$

(b) $p_{T,\gamma\gamma}$



(c) $\Delta\varphi_{\gamma\gamma}$



(d) $\cos\theta^*$

Figure 6.19: Distributions of the yield ratio before and after unfolding. The unfolding has a negligible impact on the angular variables $\Delta\varphi_{\gamma\gamma}$ and $\cos\theta^*$ and acts mainly on the energy-related spectra $m_{\gamma\gamma}$ and $p_{T,\gamma\gamma}$.

6.3.4 Efficiencies

To measure the production cross section of prompt diphotons one has to know the efficiency with which one selects the final sample in order to obtain the number of produced diphotons. Three correction factors are to be computed: the reconstruction correction factor, the identification efficiency and the trigger efficiency.

Reconstruction correction factor

The reconstruction correction factor ϵ_{RECO} is related to the probability for a prompt diphoton event to be reconstructed by CMS. It has to be extracted with prompt diphoton events simulated with `PYTHIA`. The kinematic acceptance $\mathcal{A}$ of the measurement being defined by

$$\mathcal{A} \quad \begin{cases} E_1^T > 23 \text{ GeV } \texttt{and } E_2^T > 20 \text{ GeV} \\ |\eta_{1,2}| < 1.44 \texttt{ or } 1.57 < |\eta_{1,2}| < 2.5, \\ R(\gamma_1, \gamma_2) > 0.45 \end{cases} \tag{6.15}$$

in the bin i of the Ω spectrum, $\Omega \in \{m_{\gamma\gamma}, p_{T,\gamma\gamma}, \Delta\varphi_{\gamma\gamma}, \cos\theta*\}$, $\epsilon_{RECO}(\Omega_i)$ is given by the ratio of events in the reconstructed and generated acceptance. Be $\mathcal{A}_{RECO}$ ($\mathcal{A}_{GEN}$) the boolean variable that is `true` if the reconstructed (generated) quantities of an event satisfy equation (6.15) and $\cancel{\mathcal{A}_{RECO}}$ ($\cancel{\mathcal{A}_{GEN}}$) its conjugate. Then

$$\epsilon_{RECO}(\Omega_i) = \frac{n\left(\mathcal{A}_{RECO} \;\&\; \Omega_{GEN} \in i\right)}{n\left(\mathcal{A}_{GEN} \;\&\; \Omega_{GEN} \in i\right)} \; . \tag{6.16}$$

We note that the reconstruction correction factor can be greater than one. Indeed, omitting the condition $\Omega_{GEN} \in i$, equation (6.16) can be expressed as

$$\epsilon_{RECO}(\Omega_i) = \frac{n\left(\mathcal{A}_{RECO} \;\&\; \cancel{\mathcal{A}_{GEN}}\right) + n\left(\mathcal{A}_{RECO} \;\&\; \mathcal{A}_{GEN}\;\right)}{n\left(\cancel{\mathcal{A}_{RECO}} \;\&\; \mathcal{A}_{GEN}\;\right) + n\left(\mathcal{A}_{RECO} \;\&\; \mathcal{A}_{GEN}\right)} \; . \tag{6.17}$$

If the number of events generated outside the acceptance and reconstructed within is greater than the number of events generated in the acceptance and reconstructed outside of it

$$n\left(\mathcal{A}_{RECO} \;\&\; \cancel{\mathcal{A}_{GEN}}\right) > n\left(\cancel{\mathcal{A}_{RECO}} \;\&\; \mathcal{A}_{GEN}\right) \tag{6.18}$$

then the reconstruction correction factor is greater than one. Such an example is the first bin of the $m_{\gamma\gamma}$ distribution as can be seen on figure 6.21(a).

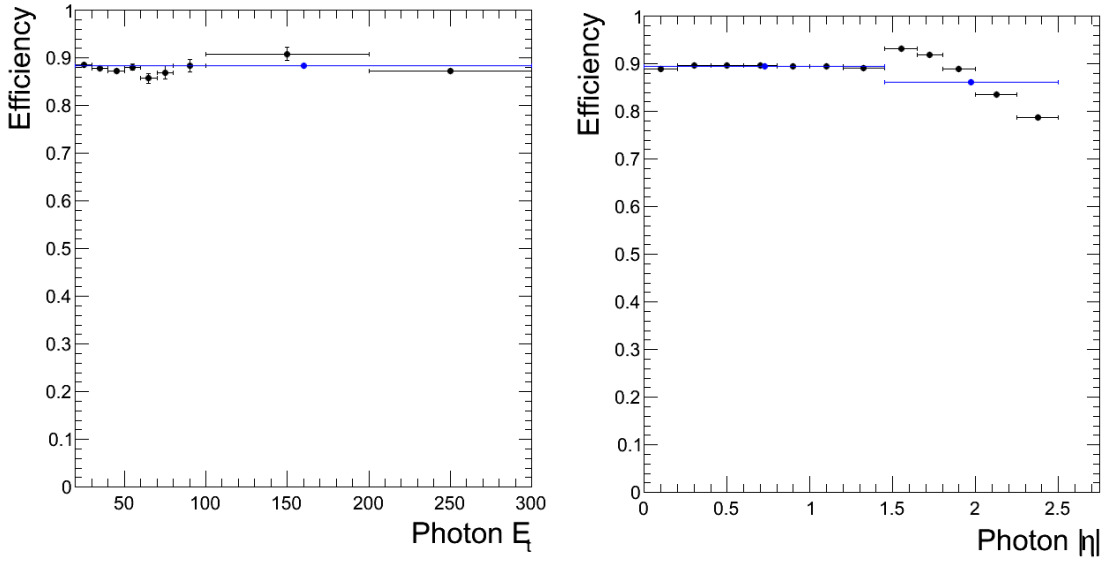
Identification efficiency

The photon identification efficiency is defined as the probability for a reconstructed prompt photon satisfying the selection requirements on E_T and $|\eta|$ to also satisfy the identification selection criteria described in section 6.1.2 and summarized in table 6.5. Studies with `PYTHIA`-simulated prompt diphoton events presented in figure 6.20 show that it is independent of the photon transverse energy but presents a dependence on the photon pseudorapidity in the endcaps.

Combined reconstruction and identification correction factor

As shown in section 5.3.2, all identification efficiencies with the exception of the requirement on the variable `pixelMatch` can be extraced from data: the efficiency of the requirements on `trigIso`, `hcalIso`, `trackIso`, $\sigma_{i\eta i\eta}$ and `H/E` can be extracted from data via the *tag and probe* method and the efficiency of the requirement on the number of *impinging tracks* with *random cones*.

The combined reconstruction and identification correction factor will thus be extracted in every bin Ω_i of the differential measurement from simulated events, $\epsilon_{\text{RECO \& ID}}[\text{SIM}](\Omega_i)$, and corrected by the



(a) Versus E_T

(b) Versus $|\eta|$

Figure 6.20: The identification efficiency determined on PYTHIA-simulated photon events is independent of the photon transverse energy but presents a dependence on the photon pseudorapidity in the endcaps.

ratios of the efficiencies determined via *tag and probe* and *random cones* in the data and in simulation. Thus, the dependencies of the reconstruction and identification efficiencies on $|\eta_\gamma|$ are taken into account and potential inaccuracies of the simulation are compensated by the terms shown in equation (6.21).

$$\epsilon_{\text{RECO \& ID}}(\Omega_i) = \epsilon_{\text{RECO \& ID}}[\text{SIM}](\Omega_i) \left(\frac{\epsilon_{\text{t\&p}}[\text{DATA}]}{\epsilon_{\text{t\&p}}[\text{SIM}]}\right)^2 \left(\frac{\epsilon_{\text{rndCone}}[\text{DATA}]}{\epsilon_{\text{rndCone}}[\text{SIM}]}\right)^2 \tag{6.19}$$

where the first factor, the reconstruction and identification correction factor in simulation, is defined as the ratio of all events falling within the reconstructed acceptance and passing the reconstructed identification requirements on one hand and all events falling within the generated acceptance, among events the generated diphoton observable Ω of which falls in the bin i:

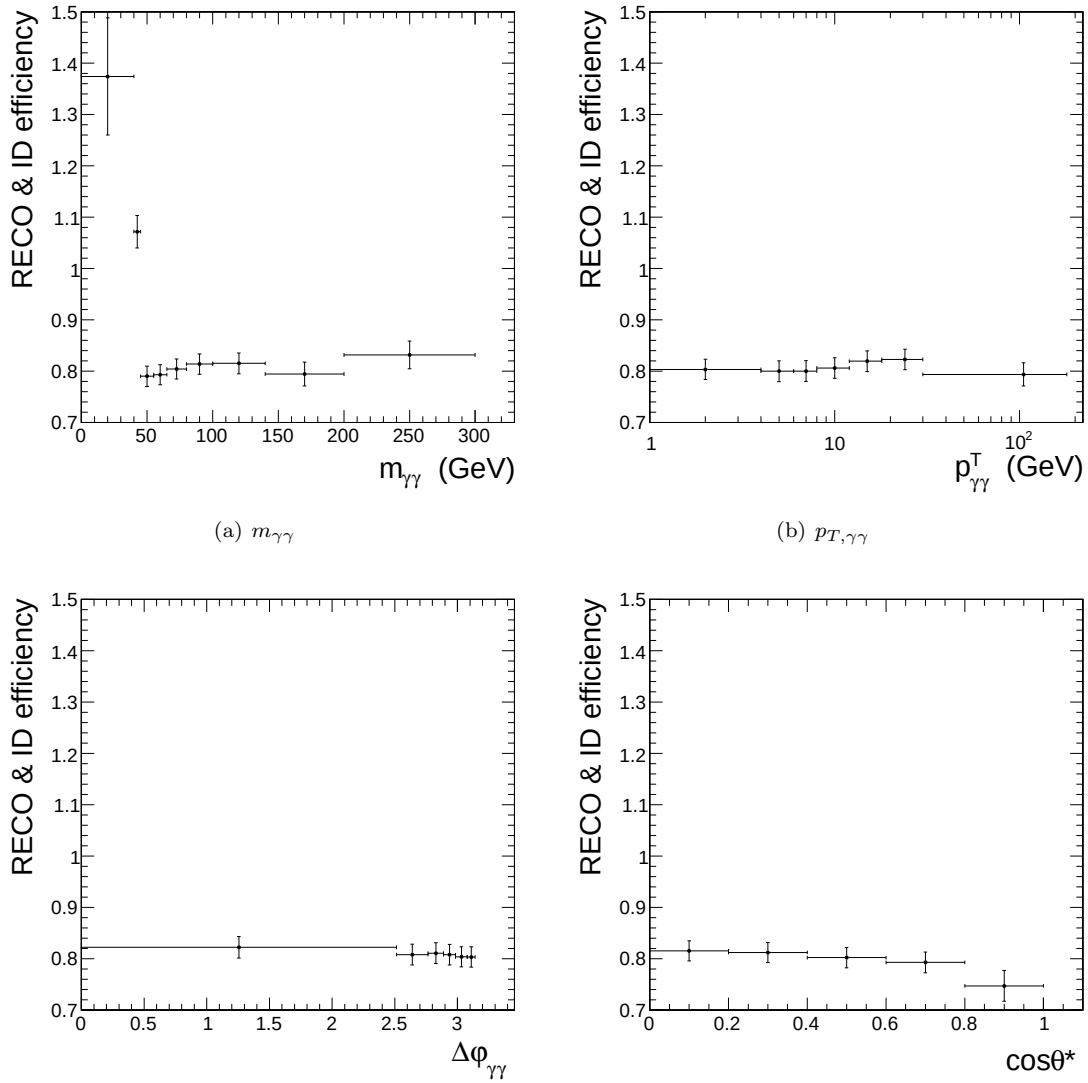
$$\epsilon_{\text{RECO \& ID}}[\text{SIM}](\Omega_i) = \frac{n\left(\mathcal{A}_{RECO} \;\&\; \Omega_{GEN} \in i \;\&\; \mathcal{C}_{RECO}\right)}{n\left(\mathcal{A}_{GEN} \;\&\; \Omega_{GEN} \in i\right)}. \tag{6.20}$$

where $\mathcal{C}_{RECO}$ denotes the identification selection requirements. Finally the bin-per-bin results of the computed efficiency are shown in figure 6.21 and the overall ratios are given in equation (6.21).

$$\left(\frac{\epsilon_{\text{t\&p}}[\text{DATA}]}{\epsilon_{\text{t\&p}}[\text{SIM}]}\right)^2 = 0.947 \pm 0.008 \qquad \left(\frac{\epsilon_{\text{rndCone}}[\text{DATA}]}{\epsilon_{\text{rndCone}}[\text{SIM}]}\right)^2 = 0.999 \pm 0.001 \tag{6.21}$$

For the $p_{T,\gamma\gamma}$ and $\Delta\varphi_{\gamma\gamma}$ spectra the correction factor does not depend on the considered bin. The first bins of the invariant mass spectrum have efficiencies greater than unity, they are mostly populated by photons close to the E_T selection threshold and have their reconstruction correction factor increased by photons reconstructed in the acceptance without being generated in it.

Finally, the integrated combined reconstruction and identification efficiency for photons with $|\eta| <$



(a) $m_{\gamma\gamma}$

(b) $p_{T,\gamma\gamma}$

(c) $\Delta\varphi_{\gamma\gamma}$

(d) $\cos\theta^*$

Figure 6.21: Combined reconstruction and identification efficiencies for photons in the barrel. The errors shown refer to the systematic uncertainties on the efficiency determination that is addressed in section 6.4.3. The first bins of the invariant mass spectrum (a) are mostly populated by photons close to the E_T selection threshold and have their reconstruction correction factor increased by photons generated outside the acceptance but reconstructed inside it.

1.44 and photons with $|\eta| < 1.44$ or $1.57 < |\eta| < 2.5$ are the following:

$$\epsilon^{\text{EB}}_{\text{RECO \& ID}} = 80.8 \pm 1.9 \ \% \tag{6.22}$$

$$\epsilon^{\text{ALL}}_{\text{RECO \& ID}} = 76.2 \pm 3.3 \ \% \tag{6.23}$$

We note that the correction factor defined in equation (6.20) also corrects for the discrepancy between the isolation measured at reconstruction level and the isolation at particle level as defined in section 1.2.2.

Trigger efficiency

The trigger efficiency is defined as the probability of an event passing all the selection requirements summarised in table 6.5 to fire on of the three high level trigger paths used for this analysis and listed in table 6.1.

The trigger inefficiency computed on simulated diphoton events generated with `PYTHIA`, with emulated HLT, is found to be smaller than $7 \cdot 10^{-4}$ for diphoton events in the kinematic region considered in this measurement. The trigger efficiency is taken to be 100 %.

6.4 Systematic Uncertainties

When refering to equations (6.6) and (6.7) we have computed the factors to access the integrated and differential cross sections. Before presenting them we review the different sources of uncertainties affecting the measurement of those cross sections.

6.4.1 Energy scale

A source of systematic uncertainty is the ECAL energy scale which is known on the 2010 dataset with a precision of 0.5 % in the barrel and 1.4 % in the endcaps [73]. The resulting uncertainty on the photon energy has two separate effects on the measurement uncertainty. It impacts the transverse energy acceptance and the spectra of the observables computed from the photon energy, that is $m_{\gamma\gamma}$ and $p_{T,\gamma\gamma}$ as shown in equations (6.2) to (6.5).

Uncertainty on acceptance

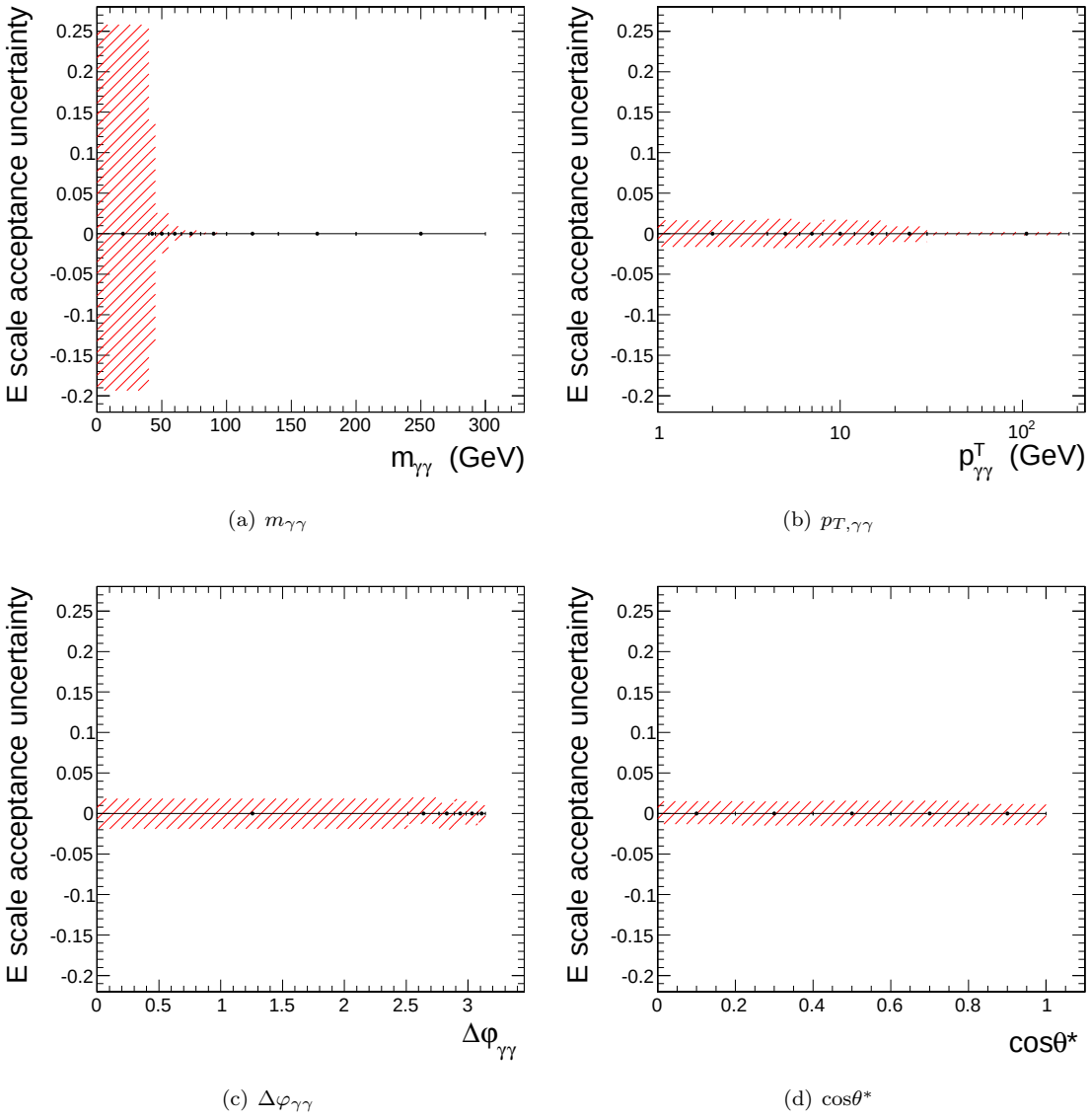
The energy scale uncertainty on the kinematic acceptance is determined via its impact on the reconstruction correction factor, from prompt diphoton events simulated with PYTHIA. It is taken to be the variation of the number of selected diphotons in simulation observed when varying the threshold on the reconstructed transverse energy by the uncertainty on the ECAL energy scale, that is 0.5 % in the barrel and 1.4 % in the endcaps.

The impact on the reconstruction energy is shown on figure 6.22. For the angular spectra $\Delta\varphi_{\gamma\gamma}$ and $\cos\theta^*$ and the major part of the $p_{T,\gamma\gamma}$ spectrum this uncertainty is constant at about 1.5 %. Above 60 GeV/c^2 in the $m_{\gamma\gamma}$ spectrum and 40 GeV/c in the $p_{T,\gamma\gamma}$ spectrum photons have transverse energies far greater than the 23,20 GeV selection thresholds and are therefore not affected by the uncertainty on the transverse energy acceptance. The region below 40 GeV/c^2 in the $m_{\gamma\gamma}$ spectrum is only populated by photons close to the 23,20 GeV selection threshold and therefore has high uncertainties of up to 26 %.

Bin-to-bin migration

The uncertainty on the ECAL energy scale affects the computed values of $m_{\gamma\gamma}$ and $p_{T,\gamma\gamma}$ and leads to bin-to-bin migrations of some events. In order to assess the uncertainty related to those migrations, the measured diphoton mass and transverse momentum spectra are fitted with an ad-hoc function. The difference of the integral of this function in the nominal bins and the bins varied by the uncertainty on the ECAL energy scale is taken to be the systematic uncertainty in every bin.

Figure 6.23 shows the bin-per-bin relative systematic uncertainty for the two variables. The highest impact is observed where the distributions show steep slopes as for example $m_{\gamma\gamma}$ around 40 GeV/c^2. As the energy measurement does not impact the measured angular observables η and φ, the angular spectra $\Delta\varphi_{\gamma\gamma}$ and $\cos\theta^*$ are not affected by this ECAL energy scale uncertainty.



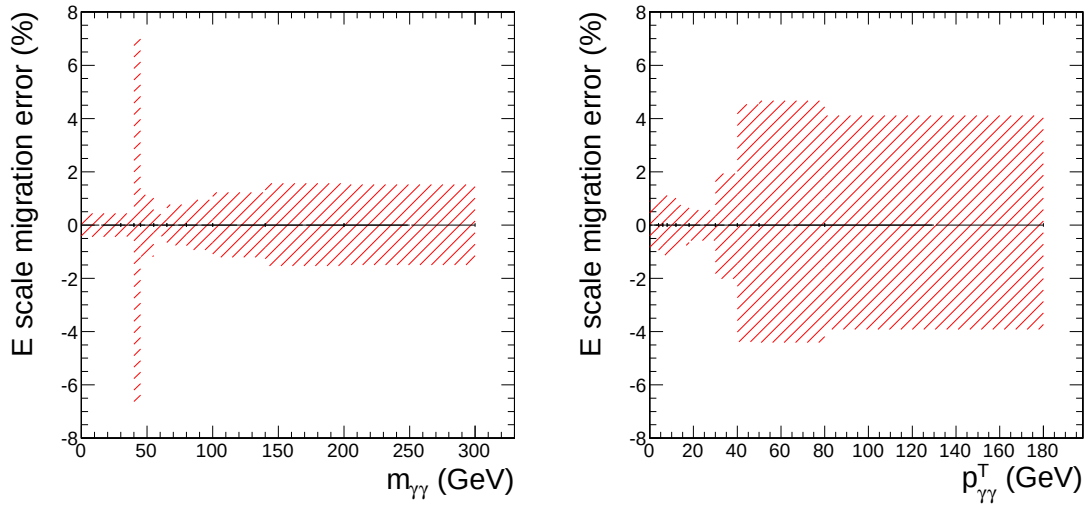
(a) $m_{\gamma\gamma}$

(b) $p_{T,\gamma\gamma}$

(c) $\Delta\varphi_{\gamma\gamma}$

(d) $\cos\theta^*$

Figure 6.22: ECAL energy scale systematic uncertainty on the kinematic acceptance for photons in the barrel. The region below 40 GeV/c^2 in the $m_{\gamma\gamma}$ spectrum is only populated by photons close to the 23,20 GeV selection threshold and therefore has high uncertainties, the regions above 60 GeV/c^2 in the $m_{\gamma\gamma}$ spectrum and above 40 GeV/c in the $p_{T,\gamma\gamma}$ spectrum on the other hand are only populated by photons with E_T far greater than these thresholds and are not affected by this uncertainty. For the remaining phase space the uncertainty is constant at about 1.5 %.



(a) $m_{\gamma\gamma}$ (b) $p_{T,\gamma\gamma}$

Figure 6.23: Relative systematic uncertainty on bin-to-bin migration in the $m_{\gamma\gamma}$ or $p_{T,\gamma\gamma}$ spectra. The highest impact is observed in the regions where the distributions show steep slopes as for example $m_{\gamma\gamma}$ around 40 GeV/c^2. As the energy measurement does not impact the measured angular observables η and φ, the angular spectra $\Delta\varphi_{\gamma\gamma}$ and $\cos\theta^*$ are not affected by this ECAL energy scale uncertainty.

6.4.2 Extraction of the pdfs

An imperfect knowledge of the signal and background pdfs used in the maximum likelihood fit induces a systematic uncertainty on the fit result. In this section we estimate this uncertainty. There are two contributions to the imperfect knowledge of the pdfs,

- the uncertainties of the **pdf extraction methods** described in chapter 4, namely the *random cone* method for the signal and the *impinging track method* for the background pdf and
- the uncertainties related to the **weighting of the single photon events** described in section 6.2.2 to match the diphoton E_T, $|\eta|$ and N_V distributions.

As explained in section 5.4.1, the discrimination power of maximum likelihood fit resides mainly in the first bin of the pdfs. To compute the uncertainty on the pdfs, we will therefore quantify the uncertainty on the first bins.

Concerning the *random cone* method, we determine the biggest discrepancy between the contents of the first bin of the following pairs of

- the *random cone* method applied on MADGRAPH-simulated [88] prompt diphoton events and the ecalIso distribution of those simulated diphoton events, compared on figure 4.12
- the *random cone* method applied on data events and the ecalIso distribution of low-radiating electrons as defined in section 4.2.3 from $Z \to e^+e^-$ events, compared on figure 4.14
- the *random cone* method applied on data events and the ecalIso distribution of low-radiating electrons from $W \to e\nu$ events, compared on figure 4.18.

The discrepancy of the first bins is highest in the comparison of the extracted pdf to data electrons from Z boson decay and amounts to 1 %.

Concerning the *impinging track* method, we determine the difference between the contents of the first bin of the pdf extracted on single photon data events with two *impinging tracks* compared to the actual `ecalIso` distribution of single photon data events with one *impinging track*, as compared on figure 4.22 in chapter 4. The observed difference is 0.03 in the barrel and 0.05 in the endcaps. The comparison is performed on `PYTHIA`-simulated events, shown on figure 4.21, and found to be compatible within statistical uncertainties.

On top of the discrepancies due to the pdf extraction methods, the weighting procedure described in section 6.2.2 is performed in two different ways.

- First, the single photons used for the pdf extraction are weighted in order to match their E_T, $|\eta|$ and N_V distributions to the ones of diphoton candidates in data, on which the requirement on the number of *impinging tracks* was relaxed.
- Second, the single photons used for the pdf extraction are weighted in order to match their E_T, $|\eta|$ and N_V distributions to the ones of `PYTHIA`-simulated prompt diphotons.

The signal and background pdfs obtained with the two sets of weights are compared and the disagreement between the first bins, of 0.5 % for both signal and background, added in quadrature to the previous differences.

By varying the first bin of the signal and the background pdfs by the aforementioned amounts, summarised in table 6.7, towards higher and lower values, one obtains several sets of modified pdfs.

Type	Signal	Background
Barrel	0.012	0.012
Endcaps	0.03	0.05

Table 6.7: Variation of the first bin

The systematic uncertainties are determined via toy Monte Carlo experiments, similarly to the ones described in section 6.3.1. Every set of modified pdfs is used to generate events which are then fitted with the nominal pdfs. This experiment is repeated 10000 times. The pull distribution thus obtained are fitted with Gaussian functions and the minimum and maximum mean values μ_{pull} are taken to be the positive and negative systematic uncertainties, after the removal of the inherent fit bias as determined in section 6.3.1.

6.4.3 Efficiency extraction

There are six contributions to the systematic uncertainty of the selection efficiency extraction described in section 6.3.4, which are summed in quadrature.

- The statistical uncertainty on the combined reconstruction and identification efficiency extracted in every bin from `PYTHIA`-simulated prompt photon events, varying from 0.2 to 11 %,
- the uncertainty on the identification requirement on the variables `pixelMatch` determined from studies on $Z \to \mu\mu\gamma$ events [94] to be 0.5 % in the barrel and 4 % in the endcaps,
- the statistical uncertainty from the *tag and probe* fit, 0.3 % as shown in section 5.3.2,
- the difference between the efficiency determined with the *tag and probe* method on `PYTHIA`-simulated $Z \to ee$ events and the actual value of that efficiency, 1.2 %,

- the difference between the efficiency on the requirement on the number of *impinging tracks* determined with *random cones* on PYTHIA-simulated prompt diphoton events and the actual value, 0.1 %,
- the systematic uncertainty associated to the trigger efficiency extraction is negligible.

The combined systematic uncertainties for the differential measurement are shown on figure 6.21. The uncertainties on the integrated efficiencies are 1.9 % for photons with $0.00 < |\eta| < 1.44$ in the barrel and 3.3 % for photons with $0.00 < |\eta| < 1.44$ or $1.57 < |\eta| < 2.5$.

6.4.4 Further uncertainties

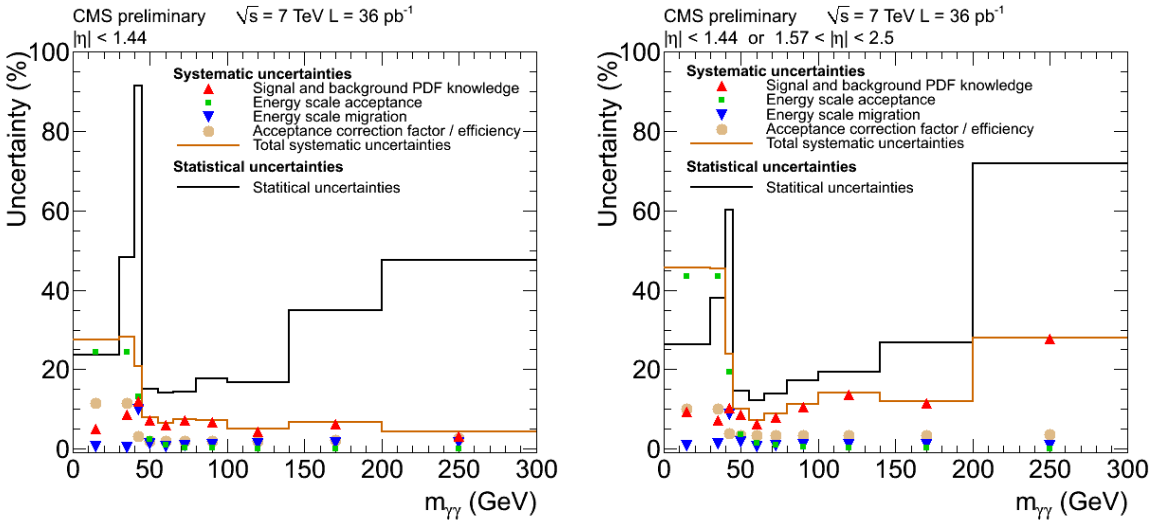
The following uncertainties are computed to be negligible compared to the aforementioned ones:

- the *random cone* method is applied on PYTHIA-simulated prompt diphoton and non-prompt diphoton events to determine the impact of the **purity of the diphoton candidate sample** on the signal pdf shape and none found,
- the error on the **bias correction** is small as 10000 toy Monte Carlo experiments were run,
- given the small **contamination with Drell-Yan events** the error on the removal of this contamination has no impact on the overall systematic uncertainty.

The uncertainty on the integrated luminosity used in equations (6.6) and (6.7) has been taken from [44] to be 4 %. Finally, the robustness of the fit is checked by varying the binning of the pdfs from the actual value of 25 bins to 50 and 100 bins, which does not affect the measured yields.

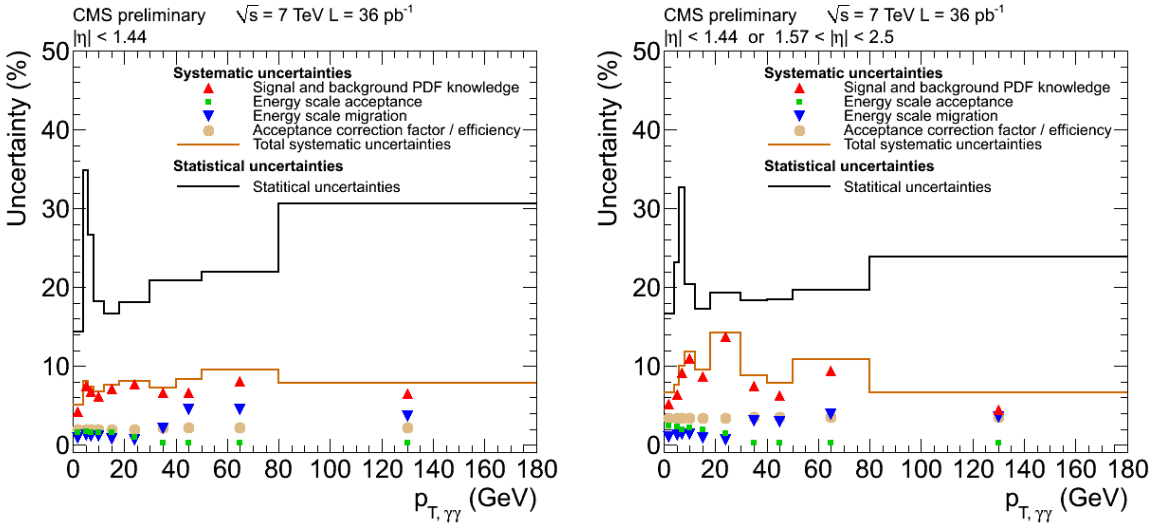
6.4.5 Summary

The systematic uncertainties presented in sections 6.4.1 to 6.4.4 are summarised in table 6.8 and the total uncertainties are computed. Plots 6.24 to 6.27 show the contribution of the different sources of systematic uncertainties to the total systematic uncertainty in each bin and contrast them with the statistical uncertainty.



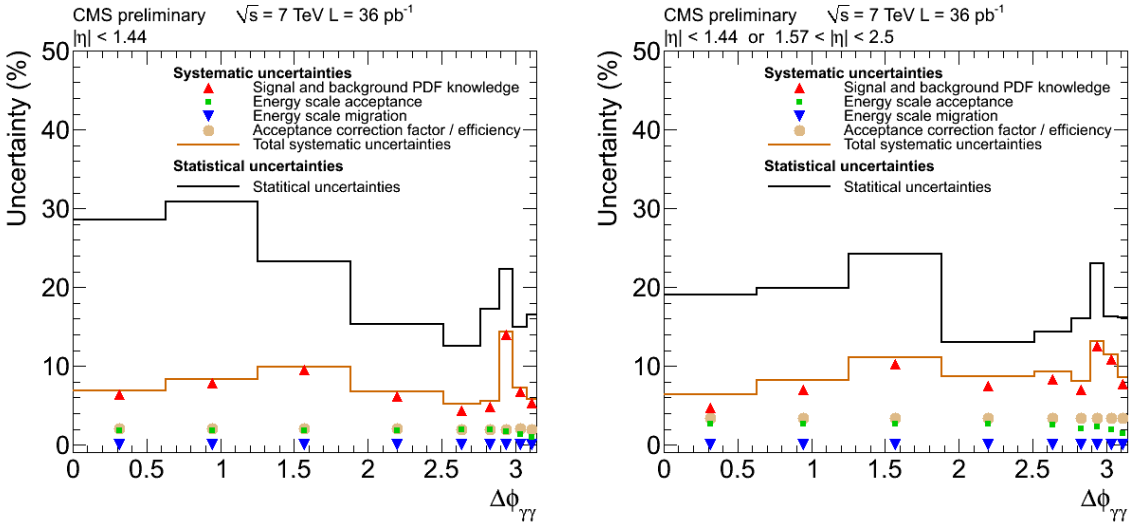
(a) $0 < |\eta| < 1.44$ (b) $0 < |\eta| < 1.44$ or $1.57 < |\eta| < 2.5$

Figure 6.24: Statistical and systematic uncertainties on the diphoton invariant mass spectrum in for photons in $0 < |\eta| < 1.44$ (a) and $0 < |\eta| < 1.44$ or $1.57 < |\eta| < 2.5$ (b).



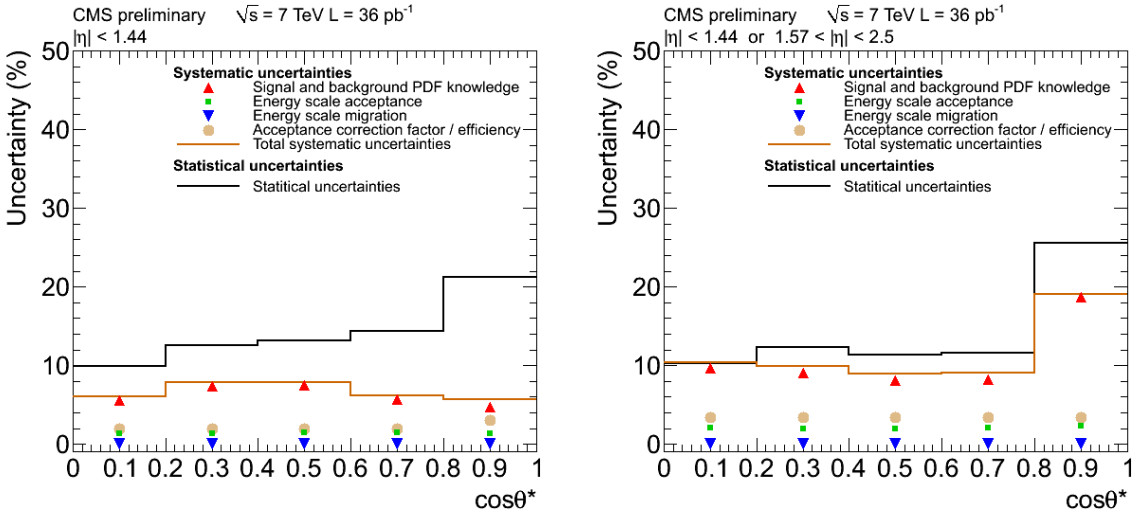
(a) $0 < |\eta| < 1.44$ (b) $0 < |\eta| < 1.44$ or $1.57 < |\eta| < 2.5$

Figure 6.25: Statistical and systematic uncertainties on the diphoton transverse momentum spectrum in for photons in $0 < |\eta| < 1.44$ (a) and $0 < |\eta| < 1.44$ or $1.57 < |\eta| < 2.5$ (b).



(a) $0 < |\eta| < 1.44$ (b) $0 < |\eta| < 1.44$ or $1.57 < |\eta| < 2.5$

Figure 6.26: Statistical and systematic uncertainties on the $\Delta\varphi$ spectrum in for photons in $0 < |\eta| < 1.44$ (a) and $0 < |\eta| < 1.44$ or $1.57 < |\eta| < 2.5$ (b).



(a) $0 < |\eta| < 1.44$ (b) $0 < |\eta| < 1.44$ or $1.57 < |\eta| < 2.5$

Figure 6.27: Statistical and systematic uncertainties on the $\cos\theta^*$ spectrum in for photons in $0 < |\eta| < 1.44$ (a) and $0 < |\eta| < 1.44$ or $1.57 < |\eta| < 2.5$ (b).

		$\|\eta\| < 1.44$		$\|\eta\| < 1.44$ or $1.57 < \|\eta\| < 2.5$	
Source	Section	Integrated	Differential	Integrated	Differential
Energy scale on acceptance	6.4.1	1.4 %	$\sim$ 1.5 %	2.0 %	$\sim$ 2 %
Energy scale on migration (a)	6.4.1	-	$\sim$ 1 %	-	$\sim$ 1.5 %
Signal and background pdf knowledge	6.4.2	6.3 %	$\sim$ 7 %	7.9 %	$\sim$ 9 %
Efficiency correction factor	6.4.3	1.9 %	$\sim$ 2 %	3.3 %	$\sim$ 3.5 %
Luminosity	6.4.4	4.0 %	4.0 %	4.0 %	4.0 %
Total (c)		7.9 %	$\sim$ 8 %	9.7 %	$\sim$ 11 %

Table 6.8: Summary of the non-negligible systematic uncertainties. The uncertainties on the ECAL energy scale bin-to-bin migration (a) are given for the $m_{\gamma\gamma}$ and $p_{T,\gamma\gamma}$ spectra. The total systematic uncertainties (b) are given including the luminosity uncertainty.

6.5 Results

In this section we recall the elements of the next-to-leading (NLO) order cross section prediction and compare to the measured integrated and differential prompt diphoton cross sections.

6.5.1 Theoretical cross section prediction

A combination of the parton generator codes DIPHOX [24] and gamma2MC [28] were used to compute the theoretical prediction, the former for the quark annihilation (figure 6.1(b)) and fragmentation contributions (figures 6.2(a) and 6.2(b)), the latter for the gluon fusion contribution shown on figure 6.1(b). All contributions were taken into account to NLO order accuracy, that is up to and including the order $\alpha_s\alpha^2$ for the diagrams computed by DIPHOX, accounting for the behaviour of the photon fragmentation function as α/α_s. The gluon fusion diagram, of order $\alpha_s^2\alpha^2$, and its corrections up to and including the order $\alpha_s^3\alpha^2$ were computed by the gamma2mc parton generator code.

Even though the gluon fusion diagrams exceed the other diagrams by two orders of α_s, their contributions are of comparable importance due to the very high gluon densities at the LHC, as shown on figure 1.10. This is the reason for their being taken into account.

In the computation, a selection criterion on the parton level isolation is imposed, requiring the total hadronic transverse energy deposited in a cone $\Delta R < 0.4$ around each photon to be less than 5 GeV. Contributions to this hadronic energy can come from residues of the parton fragmentation, from a third parton in the final state or from both. In order to correct the predicted cross section for the absence of underlying event (UE) and hadronisation in the parton generator codes, the PYTHIA [32] parton shower generator is used to determine the efficiency on the isolation criterion with and without UE and hadronisation. A correction factor of 0.953 ± 0.003 is applied to the prediction.

The calculation is performed with three different sets of parton distribution functions (PDF) CT10 [35], MSTW08 [36] and NNPDF2.1 [37]. The PDF4LHC recommendations [34] were followed to determine the uncertainties due to the PDFs and α_s. The PDFs are varied within the envelope of maximum differences of the three PDFs including their errors, the error on α_s is added in quadrature. The error within a given set of PDFs is estimated with NNPDF2.1 and DIPHOX to be about 2%. For the two other sets of PDFs, the uncertainty is also compatible with 2% within the statistical errors. Thus, the PDF uncertainty is dominated by the bin-to-bin differences between the three PDFs. The uncertainty linked

to the theoretical scales used in the computation, the fragmentation, renormalisation and factorisation scales, is estimated by taking the initial and final scales among three values $m_{\gamma\gamma}/2$, $m_{\gamma\gamma}$ and $2m_{\gamma\gamma}$. All combinations for which the scales differ at most by a factor of two were included. The total uncertainty is taken to be the maximum and minimum cross section values in each bin.

Finally, the integrated cross section predictions of this computation are :

$$\begin{array}{rrrrrrrr} \sigma(\gamma\gamma)|_{|\eta|<2.5} = & 52.7 & ^{+5.8}_{-4.2} & \text{(scale)} & \pm 2.0 & \text{(pdf)} & \text{pb} \\ \sigma(\gamma\gamma)|_{|\eta|<1.44} = & 27.3 & ^{+3.0}_{-2.2} & \text{(scale)} & \pm 1.1 & \text{(pdf)} & \text{pb.} \end{array} \tag{6.24}$$

The differential cross sections are shown on plots 6.28 to 6.35 compared with the measured cross section values.

6.5.2 Integrated cross section

The measured integrated fiducial cross sections are :

$$\begin{array}{rrrrrrrrr} \sigma(\gamma\gamma)|_{|\eta|<2.5} = & 62.4 & \pm 3.6 & \text{(stat)} & ^{+5.3}_{-5.8} & \text{(syst)} & \pm 2.5 & \text{(lumi)} & \text{pb} \\ \sigma(\gamma\gamma)|_{|\eta|<1.44} = & 31.0 & \pm 1.8 & \text{(stat)} & ^{+2.0}_{-2.1} & \text{(syst)} & \pm 1.2 & \text{(lumi)} & \text{pb.} \end{array}$$

The theoretical predictions are at 1.0 σ and 0.8 σ from the measured results for photons in the barrel and the whole electromagnetic calorimeter respectively. Theoretical computations at NNLO [101] are expected to have smaller scale uncertainties.

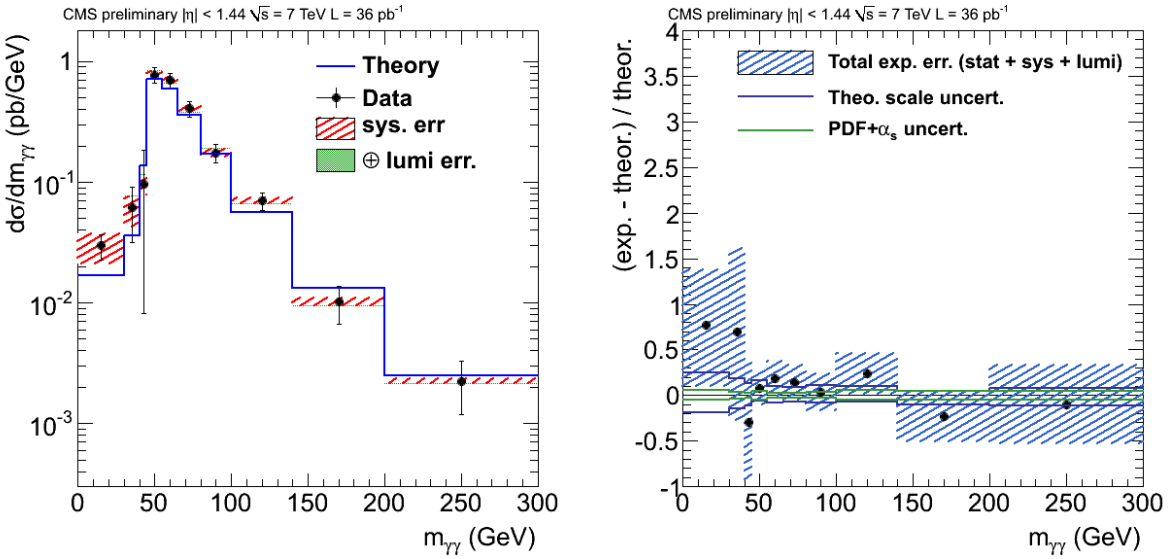
6.5.3 Differential cross sections

The measured fiducial differential cross sections are given in tables 6.9 to 6.12. On plots 6.28 to 6.35 the measured and predicted spectra are compared, plots (a), and the agreement of the data and the theory is quantified on plots (b).

$d\sigma/dm_{\gamma\gamma}$ [fb/(GeV/c^2)]

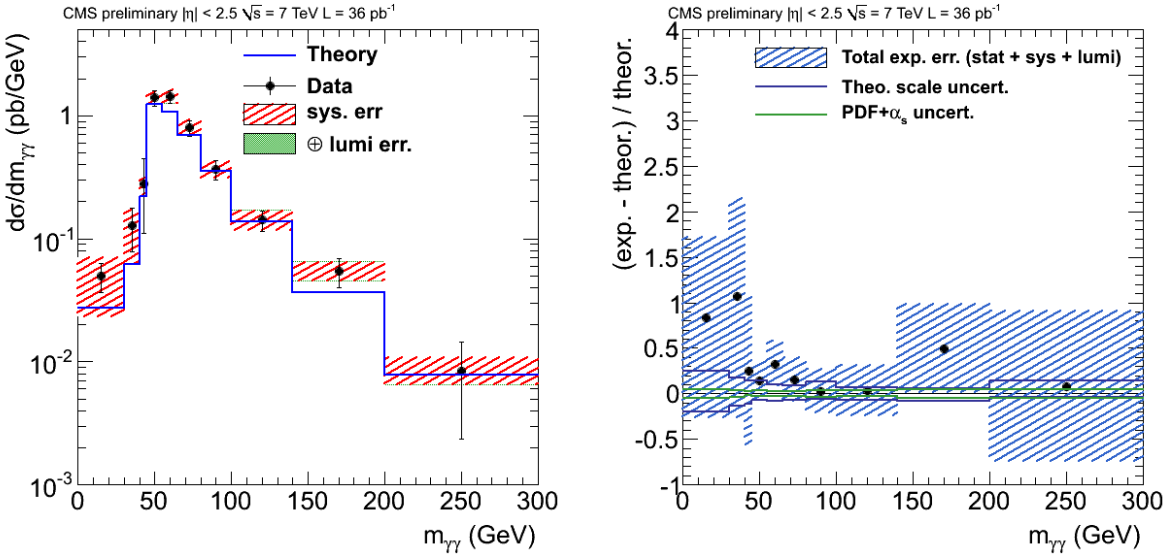
Bin [GeV]	Acceptance							
$m_{\gamma\gamma}$	$\|\eta\| < 1.44$				$\|\eta\| < 1.44$ or $1.57 < \|\eta\| < 2.5$			
		stat.	sys.			stat.	sys.	
$0-30$	29.9	± 7.1	$+6.9$	-8.6	50	± 13	$+14$	-24
$30-40$	61	± 30	$+15$	-18	127	± 49	$+35$	-61
$40-45$	97	± 88	$+20$	-20	280	± 170	$+65$	-67
$45-55$	770	± 120	$+62$	-54	1400	± 200	$+140$	-120
$55-65$	705	± 100	$+46$	-39	1430	± 180	$+100$	-93
$65-80$	408	± 59	$+30$	-31	800	± 110	$+70$	-65
$80-100$	175	± 31	$+13$	-12	365	± 63	$+41$	-37
$100-140$	70	± 12	$+3.5$	-3.4	142	± 28	$+20$	-18
$140-200$	10.2	± 3.5	$+0.69$	-0.64	54	± 15	$+6.5$	-5.9
$200-300$	2.2	± 1.1	$+0.098$	-0.087	8.4	± 6	$+2.3$	-1.9

Table 6.9: Measured differential cross section as a function of the variable $m_{\gamma\gamma}$ with statistical (stat.) and systematic uncertainties (sys.).



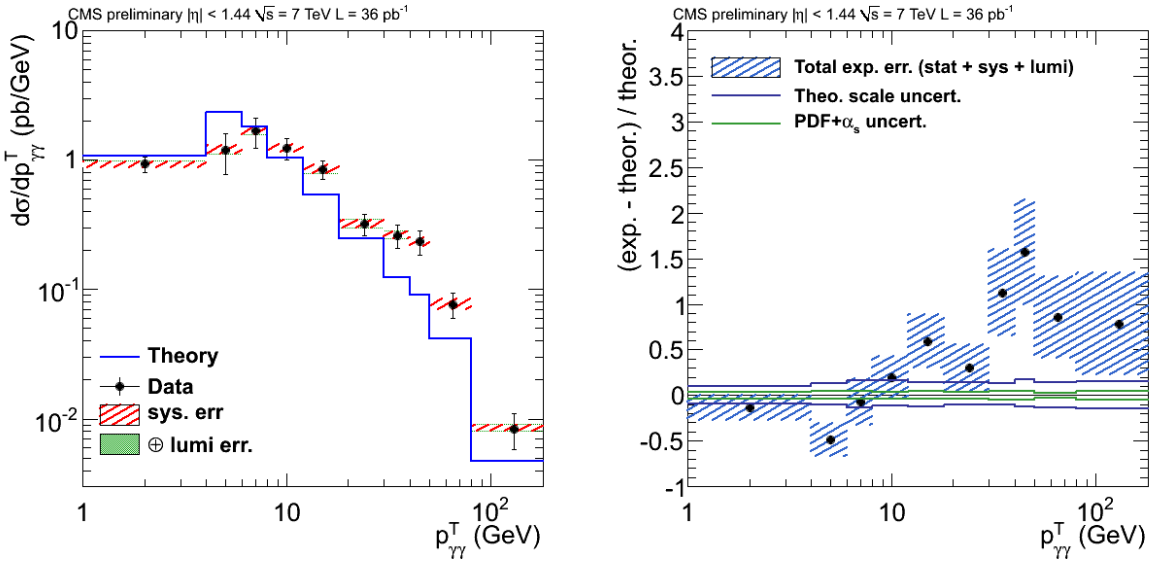
(a) Barrel (b) Endcaps

Figure 6.28: Measured diphoton production cross section as a function of the invariant mass of the photon pair (a) and bin-by-bin comparison with the theory (b) for photons within the pseudorapidity region $|\eta| < 1.44$. The total systematic uncertainties are represented by the shaded area, the different contributions are added in quadrature sequentially.



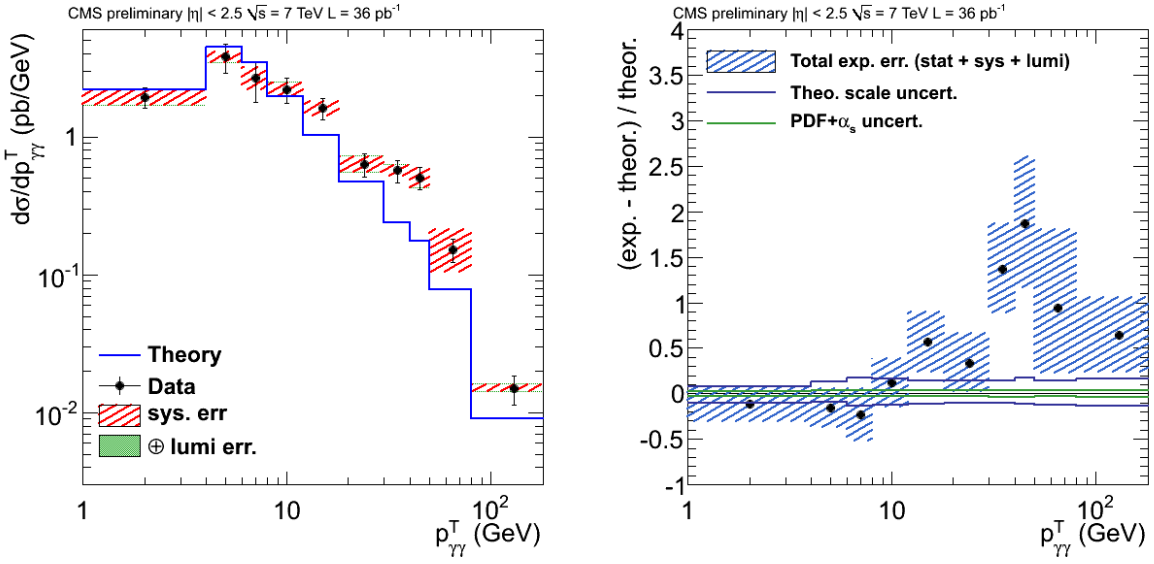
(a) Barrel (b) Endcaps

Figure 6.29: Measured diphoton production cross section as a function of the invariant mass of the photon pair (a) and bin-by-bin comparison with the theory (b) for photons within the pseudorapidity region $|\eta| < 1.44$ or $1.57 < |\eta| < 2.5$. The total systematic uncertainties are represented by the shaded area, the different contributions are added in quadrature sequentially.



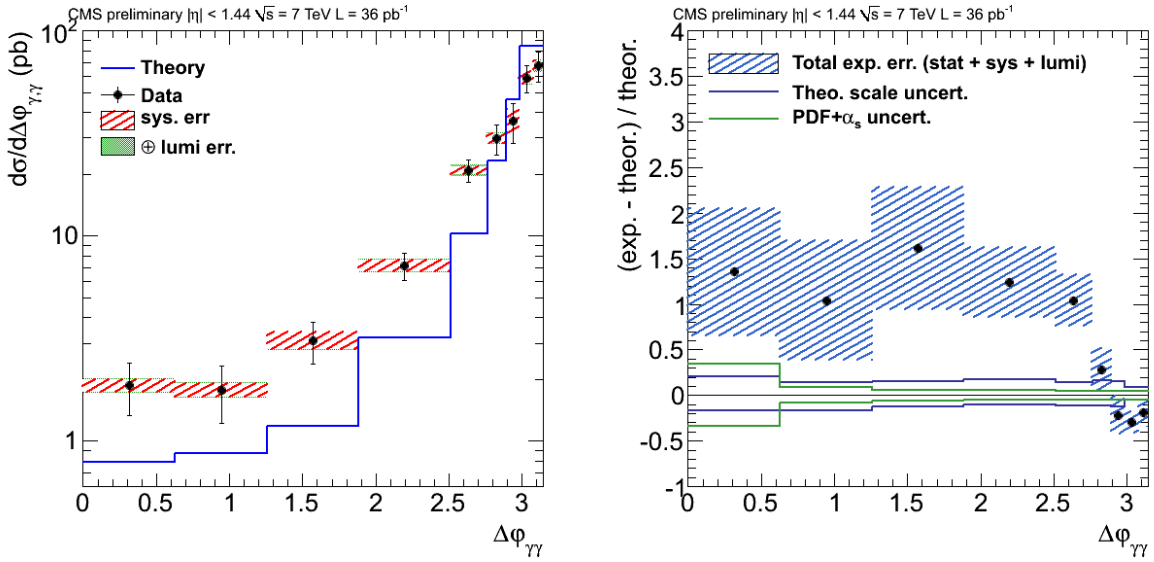
(a) Barrel (b) Endcaps

Figure 6.30: Measured diphoton production cross section as a function of the transverse momentum of the photon pair (a) and bin-by-bin comparison with the theory (b) for photons within the pseudorapidity region $|\eta| < 1.44$. The total systematic uncertainties are represented by the shaded area, the different contributions are added in quadrature sequentially.



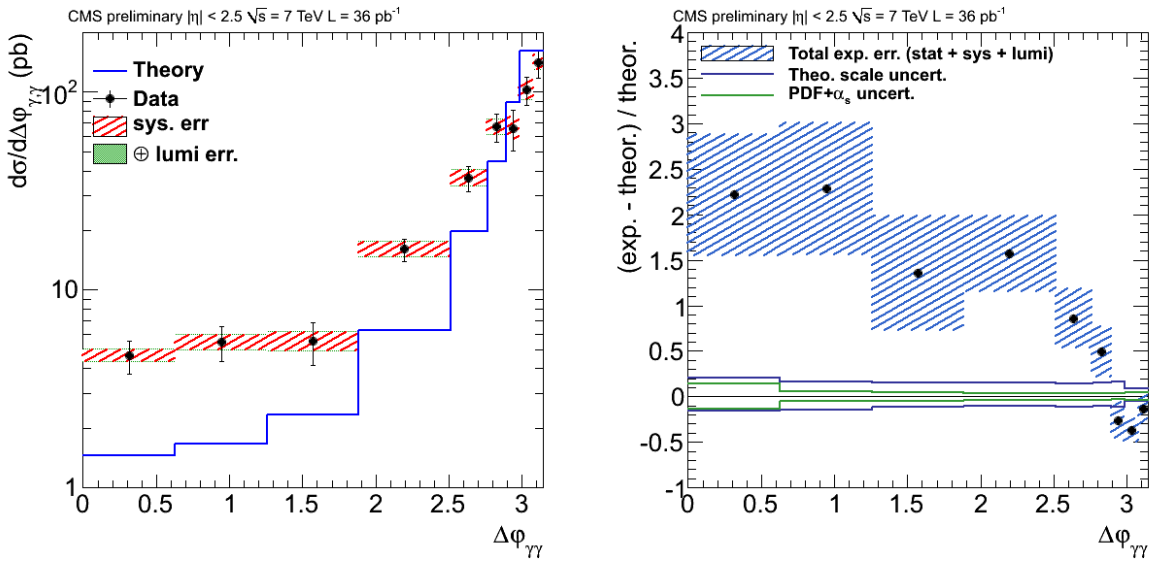
(a) Barrel (b) Endcaps

Figure 6.31: Measured diphoton production cross section as a function of the transverse momentum of the photon pair (a) and bin-by-bin comparison with the theory (b) for photons within the pseudorapidity region $|\eta| < 1.44$ or $1.57 < |\eta| < 2.5$. The total systematic uncertainties are represented by the shaded area, the different contributions are added in quadrature sequentially.



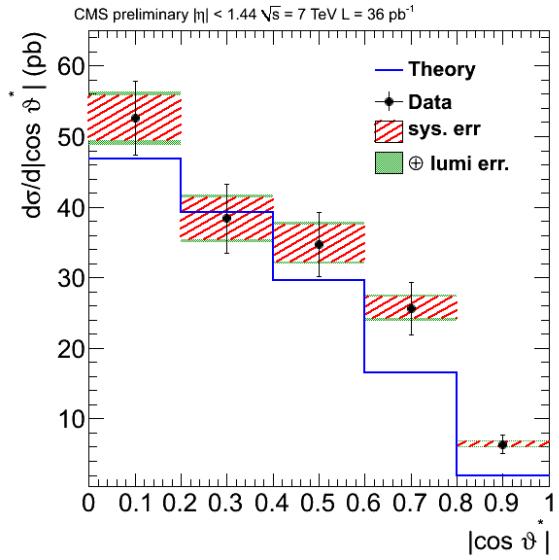
(a) Barrel (b) Endcaps

Figure 6.32: Measured diphoton production cross section as a function of the azimuthal angle between the two photons (a) and bin-by-bin comparison with the theory (b) for photons within the pseudorapidity region $|\eta| < 1.44$. The total systematic uncertainties are represented by the shaded area, the different contributions are added in quadrature sequentially.

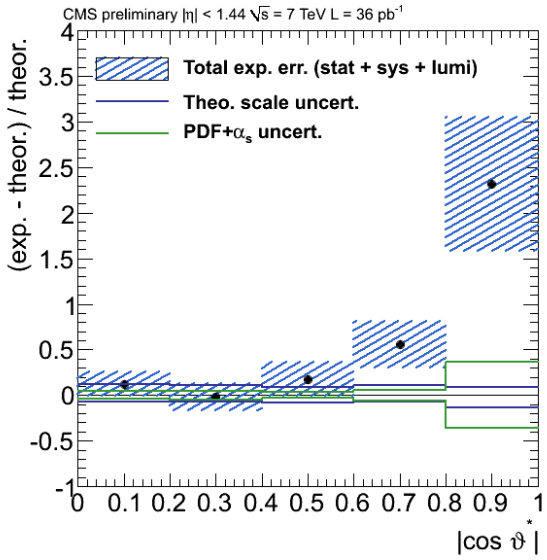


(a) Barrel (b) Endcaps

Figure 6.33: Measured diphoton production cross section as a function of the azimuthal angle between the two photons (a) and bin-by-bin comparison with the theory (b) for photons within the pseudorapidity region $|\eta| < 1.44$ or $1.57 < |\eta| < 2.5$. The total systematic uncertainties are represented by the shaded area, the different contributions are added in quadrature sequentially.

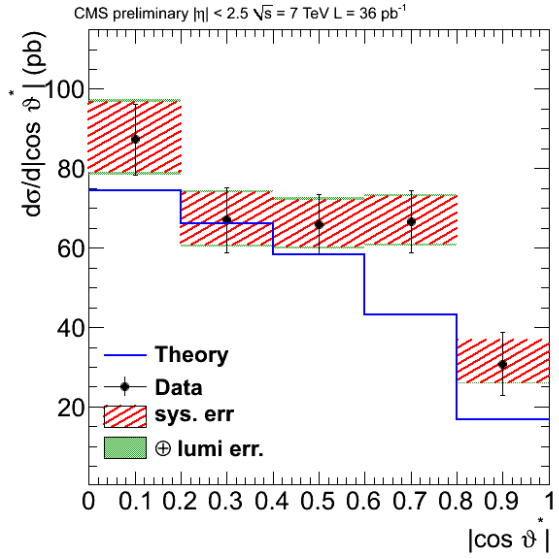


(a) Barrel

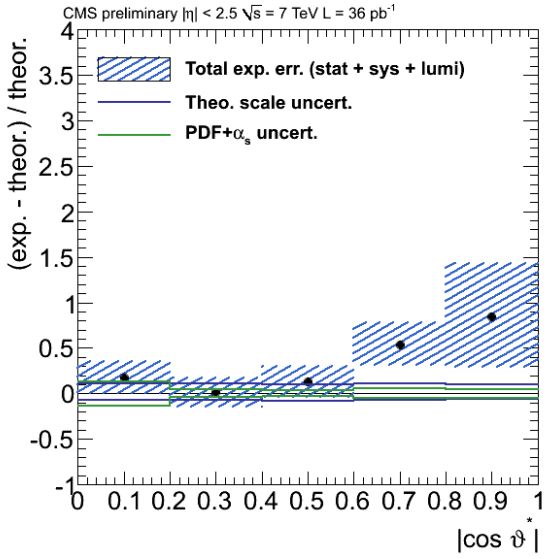


(b) Endcaps

Figure 6.34: Measured diphoton production cross section as a function of the absolute value of the cosine scattering angle in the Collins-Soper frame (a) and bin-by-bin comparison with the theory (b) for photons within the pseudorapidity region $|\eta| < 1.44$. The total systematic uncertainties are represented by the shaded area, the different contributions are added in quadrature sequentially.



(a) Barrel



(b) Endcaps

Figure 6.35: Measured diphoton production cross section as a function of the absolute value of the cosine scattering angle in the Collins-Soper frame (a) and bin-by-bin comparison with the theory (b) for photons within the pseudorapidity region $|\eta| < 1.44$ or $1.57 < |\eta| < 2.5$. The total systematic uncertainties are represented by the shaded area, the different contributions are added in quadrature sequentially.

$d\sigma/dp_{T\gamma\gamma}$ [fb/(GeV/c)]

Bin [GeV]	Acceptance							
$p_{T,\gamma\gamma}$	$\|\eta\| < 1.44$				$\|\eta\| < 1.44$ or $1.57 < \|\eta\| < 2.5$			
		stat.	sys.			stat.	sys.	
$0-4$	930	±130	$+44$	-47	1940	±320	$+120$	-130
$4-6$	1200	±420	$+97$	-85	3800	±880	$+270$	-290
$6-8$	1680	±450	$+120$	-110	2660	±870	$+270$	-240
$8-12$	1240	±220	$+82$	-76	2210	±450	$+260$	-220
$12-18$	850	±140	$+65$	-62	1610	±280	$+150$	-150
$18-30$	320	±58	$+26$	-22	630	±120	$+89$	-76
$30-40$	262	±55	$+19$	-17	570	±100	$+50$	-44
$40-50$	234	±49	$+20$	-19	507	±93	$+40$	-36
$50-80$	77	±17	$+7.3$	-6.7	153	±30	$+16$	-16
$80-180$	8.4	±2.6	$+0.66$	-0.52	15.0	±3.6	$+1.0$	-0.86

Table 6.10: Measured differential cross section as a function of the variable $p_{T,\gamma\gamma}$ with statistical (stat.) and systematic uncertainties (sys.).

$d\sigma/d\Delta\varphi_{\gamma\gamma}$ [pb]

Bin	Acceptance							
$\cos\theta^*$	$\|\eta\| < 1.44$				$\|\eta\| < 1.44$ or $1.57 < \|\eta\| < 2.5$			
		stat.	sys.			stat.	sys.	
$0-\pi/5$	1.87	±0.53	$+0.13$	-0.13	4.65	±0.89	$+0.29$	-0.30
$\pi/5-2\pi/5$	1.77	±0.55	$+0.15$	-0.14	5.5	±1.1	$+0.45$	-0.45
$2\pi/5-3\pi/5$	3.09	±0.72	$+0.31$	-0.29	5.5	±1.3	$+0.61$	-0.54
$3\pi/5-4\pi/5$	7.2	±1.1	$+0.49$	-0.44	16.1	±2.1	$+1.4$	-1.2
$4\pi/5-0.88\pi$	20.8	±2.6	$+1.0$	-0.96	36.7	±5.3	$+3.4$	-3.0
$0.88\pi-0.92\pi$	29.8	±5.1	$+1.7$	-1.5	67	±11	$+5.4$	-5.0
$0.92\pi-0.95\pi$	36.2	±8.1	$+5.1$	-4.7	66	±15	$+8.6$	-7.6
$0.95\pi-0.98\pi$	58.8	±8.8	$+4.2$	-3.8	103	±17	$+12$	-11
$0.98\pi-\pi$	68	±11	$+3.9$	-3.8	141	±23	$+12$	-11

Table 6.11: Measured differential cross section as a function of the variable $\Delta\varphi$ with statistical (stat.) and systematic uncertainties (sys.).

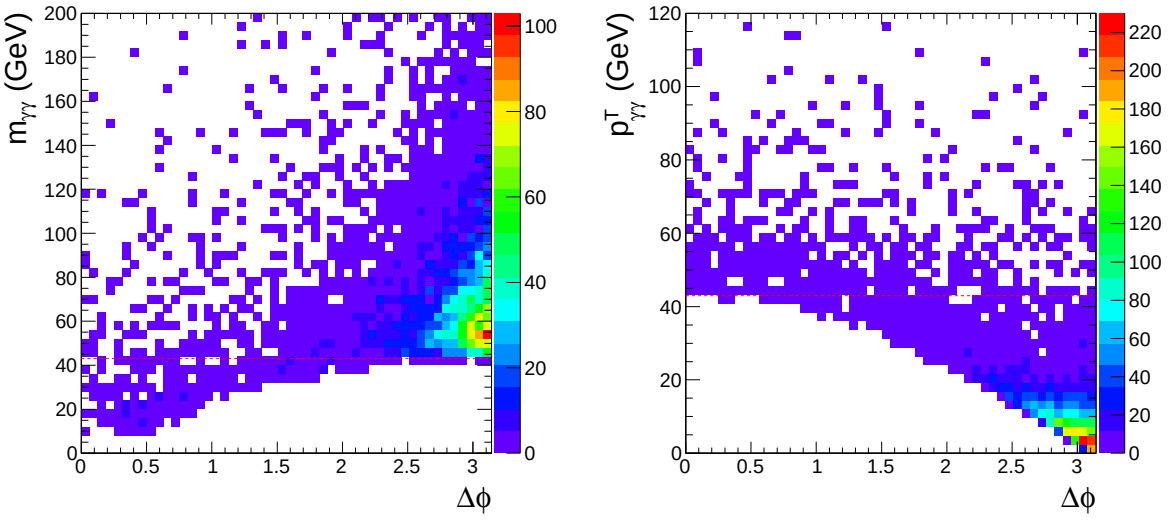
$d\sigma/d\cos\theta^*$ [pb]

Bin	Acceptance							
$\cos\theta^*$	$\|\eta\| < 1.44$				$\|\eta\| < 1.44$ or $1.57 < \|\eta\| < 2.5$			
		stat.	sys.			stat.	sys.	
$0 - 0.2$	52.6	± 5.2	$+3.1$	-3.2	87.3	± 9	$+9.1$	-7.9
$0.2 - 0.4$	38.4	± 4.9	$+3.0$	-3.0	67.0	± 8.2	$+6.6$	-6.0
$0.4 - 0.6$	34.8	± 4.6	$+2.7$	-2.5	66.0	± 7.5	$+5.9$	-5.3
$0.6 - 0.8$	25.6	± 3.7	$+1.6$	-1.5	66.7	± 7.7	$+6.1$	-5.3
$0.8 - 1$	6.4	± 1.4	$+0.34$	-0.36	30.8	± 7.9	$+5.9$	-4.7

Table 6.12: Measured differential cross section as a function of the variable $\cos\theta^*$ with statistical (stat.) and systematic uncertainties (sys.).

As can be seen on the plots 6.33 (b) and 6.32 (b) of the previous section, contrasting the theory and the measurement, in the region $\Delta\varphi_{\gamma\gamma} \lesssim 0.9\ \pi$ the theoretical cross section is underestimated. This region receives contributions only from processes beyond the lowest order diagrams shown on figure 6.1, for which the two photons are back-to-back, that is have $\Delta\varphi_{\gamma\gamma} = \pi$, because of momentum conservation. From figure 6.36(a) and equation (6.2), where $\cos\theta_{12}$ is the angle between the two photons, and knowing the selection requirements of $p_T > 20$, 23 GeV, we gather that the region $m_{\gamma\gamma} < 43\ GeV$ is mostly populated with events having $\Delta\varphi_{\gamma\gamma} \lesssim\ 0.9\ \pi$. Figure 6.36(b) and equation (6.3) show that the same is true for the region $p_{T,\gamma\gamma} > 43\ GeV$.

As shown on figures 6.28 and 6.29 on one hand and 6.30 and 6.31 on the other, the theoretical calculation underestimates the measurement at low invariant masses and in the "shoulder" above $p_{T,\gamma\gamma} \approx 40\ GeV$ as it is the case in the $\Delta\varphi_{\gamma\gamma}$ spectrum.



(a) $m_{\gamma\gamma}$ (b) $p_{T,\gamma\gamma}$

Figure 6.36: Diphoton candidate invariant mass and transverse momentum versus the difference in azimuthal angles $\Delta\varphi_{\gamma\gamma}$. The regions $m_{\gamma\gamma} < 43\ GeV$ and $p_{T,\gamma\gamma} > 43\ GeV$ are mainly populated by events with $\Delta\varphi_{\gamma\gamma} \lesssim\ 0.9\ \pi$

Analogous disagreements were observed in the diphoton production in proton-antiproton collisions [96] and [97] as well as in proton-proton collisions [98]. The most recent plots from CDF are shown on figures 6.38 to 6.40 and those from ATLAS on figures 6.41 and 6.42. This disagreement provides input for the calculation of processes not covered by current theoretical predictions, such as 2→3 processes at one loop and 2→4 processes at tree level.

In [102] the discrepancy in the $p_{T,\gamma\gamma}$ and $\Delta\varphi_{\gamma\gamma}$ spectra is studied. The fragmentation contribution to the cross section, on figure 6.2, is collinearly enhanced when the two photons come close to each other, an example of such a configuration is shown on figure 6.37. The theoretical description in the regions corresponding to $\Delta\varphi_{\gamma\gamma} \lesssim\ 0.9\ \pi$ is to *effective lowest order* and hence underestimated.

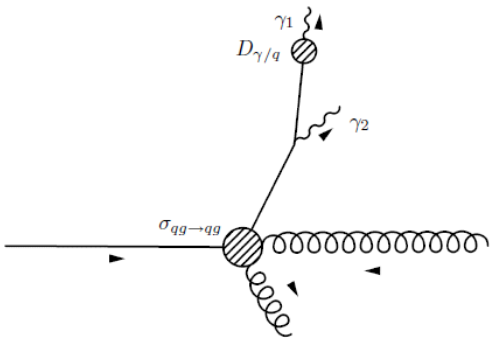


Figure 6.37: Kinematical configuration for which the fragmentation contribution is collinearly enhanced when the two photons become close to each other, published in [102]

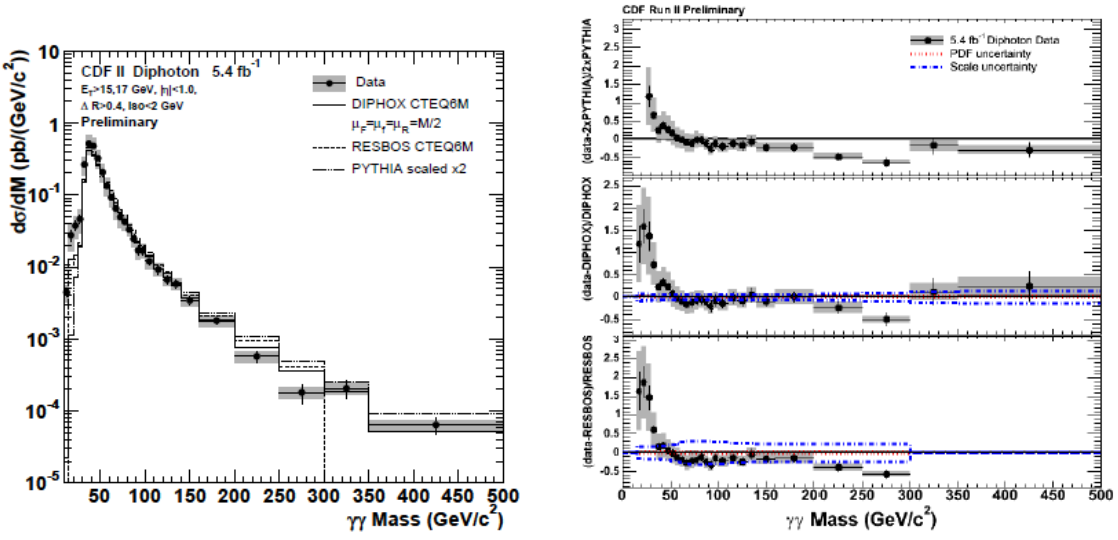


Figure 6.38: Diphoton production cross section as a function of the invariant mass of the photon pair and bin-by-bin comparison with a theoretical prediction as measured by CDF [96].

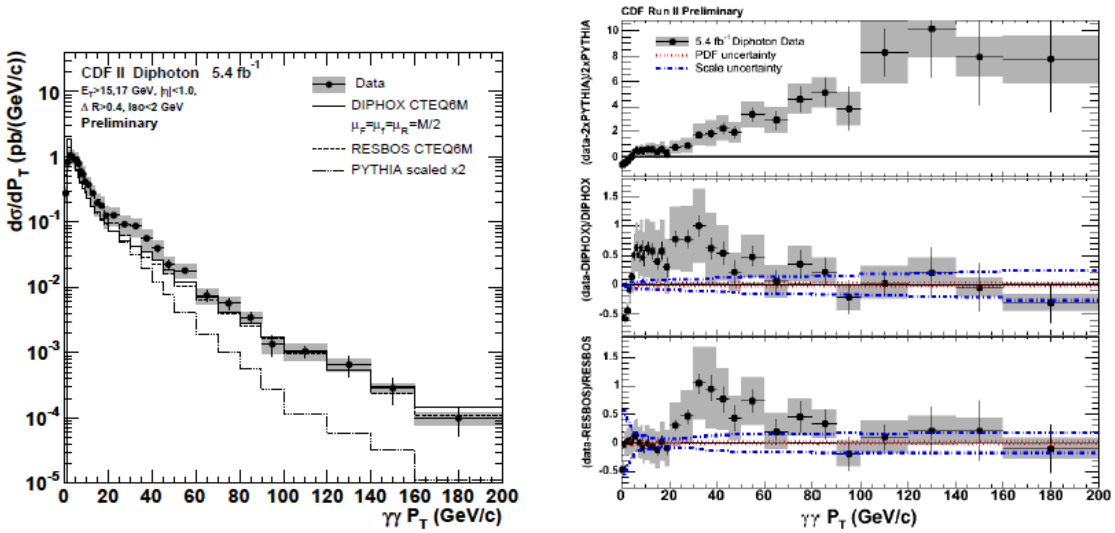


Figure 6.39: Diphoton production cross section as a function of the transverse momentum of the photon pair and bin-by-bin comparison with a theoretical prediction as measured by CDF [96].

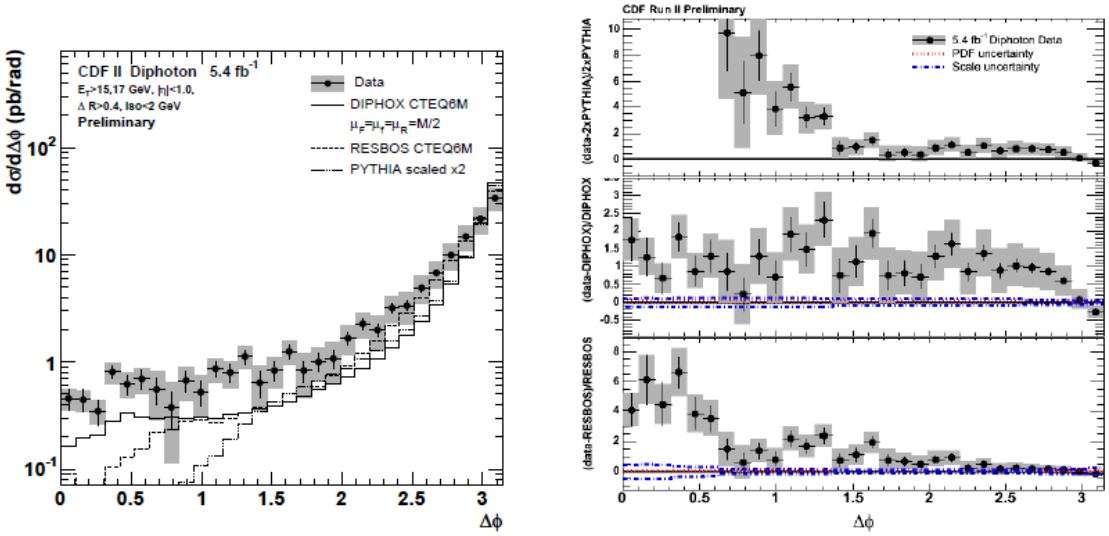
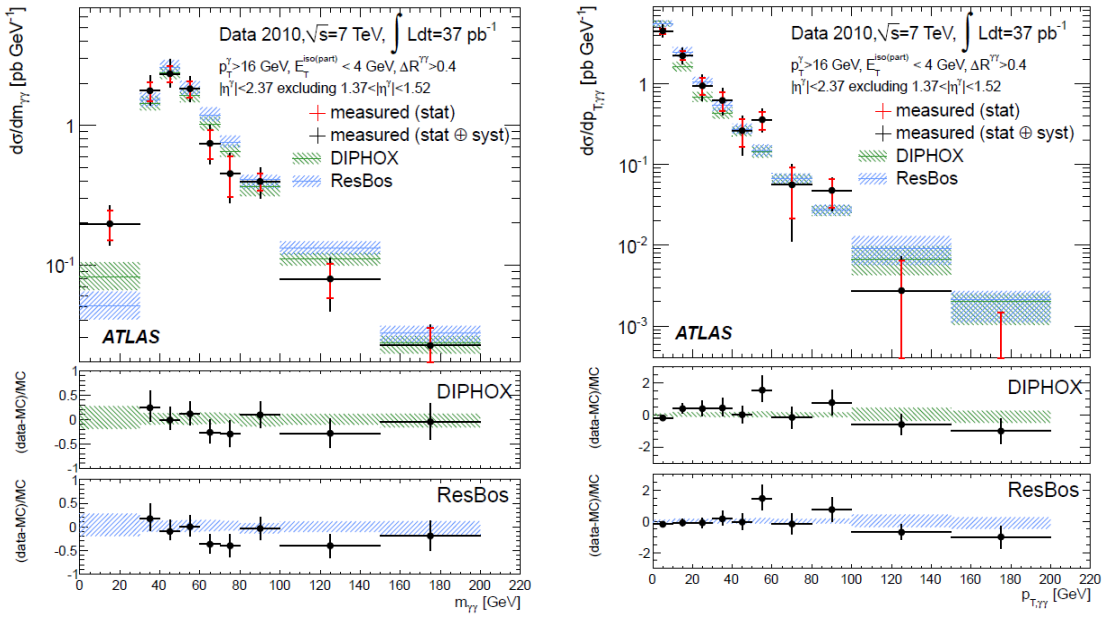


Figure 6.40: Diphoton production cross section as a function of the azimuthal angle between the two photons and bin-by-bin comparison with a theoretical prediction as measured by CDF [96].



(a) $m_{\gamma\gamma}$ (b) $p_{T,\,\gamma\gamma}$

Figure 6.41: Diphoton production cross section as a function of the invariant mass (a) and the transverse momentum (b) of the photon pair and bin-by-bin comparison with a theoretical prediction as measured by ATLAS [98].

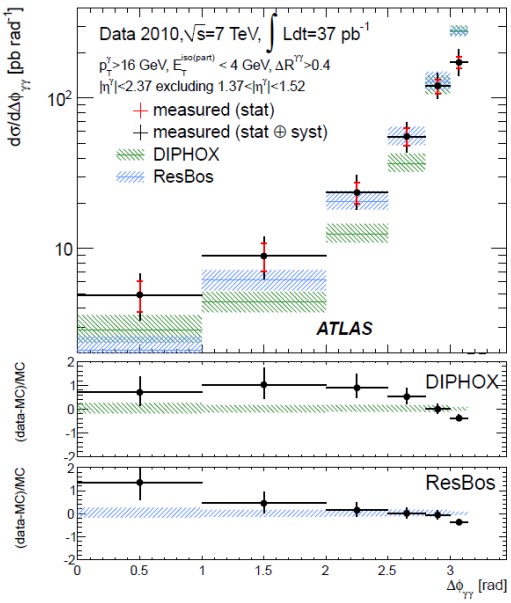


Figure 6.42: Diphoton production cross section as a function of the azimuthal angle between the two photons and bin-by-bin comparison with a theoretical prediction as measured by ATLAS [98].

Conclusion

In this thesis, the comprehensive measurement of the inclusive production cross section of isolated prompt photon pairs with the CMS detector was presented. This result represents valuable input for the validation of perturbative quantum chromodynamics (pQCD) calculations and it sheds light on the important irreducible background to light Higgs boson searches in the diphoton channel.

First, we developed a method to perform the discrimination of prompt isolated photons and non-prompt photons from jets, which constitute the main background of prompt photons analyses. As it is not possible to reject all this background by applying selection requirements on discriminating variables, a statistical method was chosen to perform the discrimination. It is based on a maximum likelihood fit on an electromagnetic isolation variable and was used to extract the number of prompt isolated photons from a sample of photon candidates.

Both the signal and the background probability density functions (pdf) used in the fit were extracted from data. The signal distribution of the electromagnetic isolation variable, that is the distribution of prompt isolated photons, was obtained by measuring the electromagnetic energy deposited in cones chosen randomly in each event. The background distribution, that of non-prompt photons from jets, was extracted from a control sample, selected by requiring a track to hit the electromagnetic calorimeter close to the photon impact position and by subsequently removing the deposits left by that track. Both extraction methods were validated on simulated as well as on data events.

The new discrimination method, different with respect to previous methods in that it does not rely on simulation for the extraction of the pdfs of the discrimination fit, was then applied to the measurement of the differential inclusive production cross section of single photons versus the photon transverse energy. The spectrum measured by CMS with a different method was reproduced for photons in the barrel and the endcap. The systematic uncertainty on the extraction of the pdfs was computed and found to be smaller than the ones of that measurement.

Finally, the discrimination method was used to measure the inclusive production cross section of isolated prompt photon pairs. The differential cross section was measured as a function of four variables chosen either because of their sensitivity to the underlying QCD processes or because of their ability to describe the large irreducible background to searches of a light Higgs boson decaying to two photons: the invariant mass of the two photons $m_{\gamma\gamma}$, the transverse momentum $p_{T,\gamma\gamma}$ of the diphoton system, the azimuthal angle between the two photons $\Delta\varphi_{\gamma\gamma}$ and the cosine of the scattering angle in the Collins-Soper frame $\cos\theta^*$.

The extraction of the number of prompt isolated photon pairs from a sample of diphoton candidates was performed with a two-dimensional maximum likelihood fit, where the pdfs were built from those extracted on single photons. The systematic uncertainties on the measurement have been estimated, they are dominated by the uncertainties on pdf extraction.

A state-of-the-art fixed order theoretical prediction, computing all contributing processes to next-to-leading order (NLO), was compared to the measurement. The predicted integrated cross section

was found to agree with the measurement and disagreements between the predicted and the measured spectra were observed in the region of phase space where the angle between the two photons of the pair is small. Indeed, as photons are back-to-back in the leading-order calculation, this phase-space region is populated only by contributions from NLO diagrams and hence underestimated by the theoretical prediction. Similar disagreements have been observed by recent measurements performed by the CDF and the ATLAS collaborations.

This measurement was performed with 36 pb^{-1} of data. As of July 2011, more than 1.6 fb^{-1} have been collected by CMS. Performing the same measurement on that extended dataset would allow two major improvements. On one hand, given the higher numbers of collected diphoton candidates, a sample of diphoton candidates could be used for the extraction of the two-dimensional pdfs instead of a sample of single photon candidates. On the other, thanks to the greater collected data sample, the differential spectra could be measured using much smaller bins.

This finer binning would be of especially high interest in the region of $m_{\gamma\gamma}$, where the sensitivity to the $H \to \gamma\gamma$ decay is highest, that is between 110 and 130 GeV. In the context of searches of a light Higgs boson, the method for statistical discrimination of prompt isolated photons and non-prompt photons from jets developed in this thesis, could be used for an in depth study of the reducible and irreducible backgrounds.

As the instantaneous luminosity of the 2011 data-taking increased drastically compared to the 2010 data-taking, one expects an impact of the higher number of pileup events on the discrimination power of the electromagnetic isolation variable used for the fit and on the pdf extraction methods. These impacts would have to be investigated.

Furthermore, in order to better understand the discrepancies observed in the regions where the fixed-order NLO prediction does not reproduce the measurement, it would be interesting to compare the differential cross section spectra with a theoretical prediction performing soft gluon resummation, as done for instance by `RESBOS` [31]. Finally, a fixed-order next-to-next-to-leading order (NNLO) theoretical prediction is expected to be published soon [103] and will shed further light on the aforementioned discrepancies.

Bibliography

[1] C. Amsler, "Particle physics booklet," *Physics Letters B*, vol. 667, no. 1, pp. 13–20, 2008.

[2] (Aleph Collaboration), "A precise determination of the number of families with light neutrinos and of the Z boson partial widths," *Physics Letters B*, vol. 235, no. 3-4, pp. 399–411, 1990.

[3] C. Amsler, "Particle physics booklet," *Physics Letters B*, vol. 667, no. 1, pp. 21–22, 2008.

[4] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory; 1995 ed.*, pp. 15–19. Boulder, CO: Westview, 1995.

[5] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory; 1995 ed.*, pp. 133–169. Boulder, CO: Westview, 1995.

[6] C. Amsler, "Particle physics booklet," *Physics Letters B*, vol. 667, no. 1, pp. 8–12, 2008.

[7] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory; 1995 ed.*, pp. 506–512. Boulder, CO: Westview, 1995.

[8] J. Ellis, M. Gaillard, and D. Nanopoulos, "A phenomenological profile of the higgs boson," *Nuclear Physics B*, vol. 106, pp. 292–340, 1976.

[9] R. Barate *et al.*, "Search for the standard model Higgs boson at LEP," *Phys. Lett.*, vol. B565, pp. 61–75, 2003.

[10] T. Aaltonen *et al.*, "Combination of Tevatron searches for the standard model Higgs boson in the W+W- decay mode," 2010.

[11] (The CMS Collaboration), "SM Higgs Combination," *CMS PAS HIG-11-011*, 2010.

[12] A. Hoecker, "Status of the global electroweak fit of the Standard Model," 2009.

[13] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory; 1995 ed.*, pp. 175–208. Boulder, CO: Westview, 1995.

[14] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory; 1995 ed.*, pp. 786–789. Boulder, CO: Westview, 1995.

[15] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory; 1995 ed.*, pp. 102–108. Boulder, CO: Westview, 1995.

[16] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory; 1995 ed.*, p. 378. Boulder, CO: Westview, 1995.

[17] D. Hanneke, S. Fogwell, and G. Gabrielse, "New Measurement of the Electron Magnetic Moment and the Fine Structure Constant," *Phys.Rev.Lett.*, vol. 100, p. 120801, 2008.

[18] S. Bethke, "The 2009 World Average of α_s," *Eur.Phys.J.*, vol. C64, pp. 689–703, 2009.

[19] D. J. Gross and F. Wilczek, "Ultraviolet behavior of non-abelian gauge theories," *Phys. Rev. Lett.*, vol. 30, pp. 1343–1346, Jun 1973.

[20] S. Frixione, "Isolated photons in perturbative QCD," *Phys.Lett.*, vol. B429, pp. 369–374, 1998.

[21] B. Bailey and J. Owens, "An Order alpha-s two photon background study for the intermediate mass Higgs boson," *Phys.Rev.*, vol. D47, pp. 2735–2738, 1993.

[22] J. Gunion, *The Higgs Hunter's Guide*. Boulder, Colorado, USA: Westview Press, 2000.

[23] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino, and R. Tanaka (Eds.), "Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables," *CERN-2011-002*, CERN, Geneva, 2011.

[24] T. Binoth, J. Guillet, E. Pilon, and M. Werlen, "A full next-to-leading order study of direct photon pair production in hadronic collisions," *The European Physical Journal C - Particles and Fields*, vol. 16, pp. 311–330, 2000. 10.1007/s100520050024.

[25] "The Durham HepData Project," 2011.

[26] G. Altarelli and G. Parisi, "Asymptotic Freedom in Parton Language," *Nucl. Phys.*, vol. B126, p. 298, 1977.

[27] P. Aurenche, M. Fontannaz, J.-P. Guillet, E. Pilon, and M. Werlen, "A New critical study of photon production in hadronic collisions," *Phys.Rev.*, vol. D73, p. 094007, 2006.

[28] Z. Bern, L. Dixon, and C. Schmidt, "Isolating a light higgs boson from the diphoton background at the cern lhc," *Phys. Rev. D*, vol. 66, p. 074018, Oct 2002.

[29] G. Altarelli, R. Ellis, M. Greco, and G. Martinelli, "Vector Boson Production at Colliders: A Theoretical Reappraisal," *Nucl.Phys.*, vol. B246, p. 12, 1984.

[30] P. Chiappetta, R. Fergani, and J. Guillet, "Double prompt photon production from hadronic collisions," *Phys.Lett.*, vol. B348, pp. 646–656, 1995.

[31] C. Balazs, P. M. Nadolsky, C. Schmidt, and C. Yuan, "Diphoton background to Higgs boson production at the LHC with soft gluon effects," *Phys.Lett.*, vol. B489, pp. 157–162, 2000.

[32] T. Sjöstrand, S. Mrenna, and P. Skands, "Pythia 6.4 physics and manual," *Journal of High Energy Physics*, vol. 2006, no. 05, p. 026, 2006.

[33] S. Ahuja, O. Bondu, N. Chanon, G. Chen, B. Choudhary, M. Déjardin, D. d'Enterria, J. Fan, F. Ferri, S. Gascon-Shotkin, P. Gras, J. Malcles, L. Millischer, J. Tao, and H. Xiao, "Theoretical predictions for the inclusive photon pair production cross section in pp collisions at $\sqrt{s} = 7$ tev and correction of non-perturbative effects," *CMS Analysis Note AN-11-038*, 2011.

[34] M. Botje *et al.*, "The PDF4LHCWorking Group Interim Recommendations," *arXiv:1101.0538*.

[35] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, J. Pumplin, and C.-P. Yuan, "New parton distributions for collider physics," *Phys. Rev. D*, vol. 82, p. 074024, Oct 2010.

[36] A. Martin, W. Stirling, R. Thorne, and G. Watt, "Uncertainties on $alpha_s$ in global pdf analyses and implications for predicted hadronic cross sections," *The European Physical Journal C - Particles and Fields*, vol. 64, pp. 653–680, 2009. 10.1140/epjc/s10052-009-1164-2.

[37] R. D. Ball, V. Bertone, F. Cerutti, L. D. Debbio, S. Forte, A. Guffanti, J. I. Latorre, J. Rojo, and M. Ubiali, "Impact of heavy quark masses on parton distributions and lhc phenomenology," *Nuclear Physics B*, vol. 849, no. 2, pp. 296 – 363, 2011.

[38] C. Anastasiou, L. J. Dixon, K. Melnikov, and F. Petriello, "High precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at NNLO," *Phys.Rev.*, vol. D69, p. 094008, 2004.

[39] O. S. Brüning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock, *LHC Design Report*. Geneva: CERN, 2004.

[40] D. Acosta *et al.*, *CMS physics: Technical Design Report*, vol. I: Detector Performance and Software of *Technical Design Report CMS*, p. 315. Geneva: CERN, 2006.

[41] W. Wittmer, *Luminosity Tuning at the Large Hadron Collider*. PhD thesis, Technische Universität Graz, 2006.

[42] (The CMS Collaboration), "Measurement of CMS Luminosity," *CMS PAS EWK-10-004*, 2010.

[43] S. van der Meer, "Calibration of the effective beam height in the ISR," Tech. Rep. CERN-ISR-PO-68-31. ISR-PO-68-31, CERN, Geneva, 1968.

[44] "Measurement of CMS Luminosity," *CMS-DP-2011-002.*, 2011.

[45] (The CMS Collaboration), "CMS Luminosity - Public Results," 2011.

[46] (The CMS Collaboration), "The tracker project technical design report," *CERN/LHCC 98-6*, 1998.

[47] D. Acosta *et al.*, *CMS physics: Technical Design Report*, vol. I: Detector Performance and Software of *Technical Design Report CMS*, p. 25. Geneva: CERN, 2006.

[48] D. Acosta *et al.*, *CMS physics: Technical Design Report*, vol. I: Detector Performance and Software of *Technical Design Report CMS*, p. 315. Geneva: CERN, 2006.

[49] (The CMS Collaboration), "HCAL Technical Design Report," *CERN/LHCC 97-31*, 1997.

[50] (The CMS Collaboration), "The tridas project technical design report, volume 1: The trigger systems," *CERN/LHCC*, vol. 2002-26, 2002.

[51] (The CMS Collaboration), "The TriDAS Project Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger," *CERN/LHCC*, vol. 2002-26, 2002.

[52] (The CMS Collaboration), "CMS ECAL Technical Design Report," *CERN/LHCC 97-33*, 1997.

[53] G. Franzoni, "Performance of CMS ECAL with first LHC data," *Nucl.Instrum.Meth.*, vol. A628, pp. 90–94, 2011.

[54] D. Acosta *et al.*, *CMS physics: Technical Design Report*, vol. I: Detector Performance and Software of *Technical Design Report CMS*, p. 15. Geneva: CERN, 2006.

[55] M. Anfreville et al. , "Laser monitoring system for the CMS lead tungstate crystal calorimeter," *Nucl. Instr. and Meth.*, vol. A594, 2008.

[56] N. Almeida and others, "Data filtering in the readout of the CMS Electromagnetic Calorimeter," *JINST*, vol. 3, pp. 2–11, 2008.

[57] (The CMS Collaboration), "Photon Reconstruction and Identification at $\sqrt{s} = 7$ TeV," *CMS PAS EGM-10-005*, 2010.

[58] D. Acosta *et al.*, *CMS physics: Technical Design Report*, vol. I: Detector Performance and Software of *Technical Design Report CMS*, p. 164. Geneva: CERN, 2006.

[59] P. Paganini and I. van Vulpen, "Pulse amplitude reconstruction in the CMS ECAL using the weights method," *CMS Note 2004/025*, 2004.

[60] L. M. Barone, F. Cavallari, S. Costantini, I. Dafinei, M. Diemoz, E. Longo, P. Meridiani, G. Organtini, and R. Paramatti, "Improvement on PbWO4 Crystal Intercalibration Precision from Light Yield Measurements at the INFN-ENEA Regional Center," *CMS Rapid Note RN 2004/003*, 2004.

[61] F. Cavallari, G. Franzoni, A. Ghezzi, P. Govoni, A. Meregaglia, and M. Obertino, "Relative Light Yield comparison between laboratory and testbeam data for CMS ECAL PbWO4 crystals," *CMS Rapid Note RN 2004/002*, 2004.

[62] D. Futyan and C. Seez, "Intercalibration of the CMS electromagnetic calorimeter crystals in phi using symmetry of energy deposition," *J.Phys.G*, vol. G29, pp. 1299–1326, 2003.

[63] P. Meridiani and R. Paramatti, "On the use of $Z \to ee$ events for ECAL calibration," *CMS Note 2006/039*, 2006.

[64] A. Ghezzi, "Calibration of the CMS electromagnetic calorimeter at the LHC start up," *Nucl.Instrum.Meth.*, vol. A617, pp. 88–89, 2010.

[65] (The CMS Collaboration), "Electromagnetic calorimeter commissioning and first results with 7 TeV data," *CMS NOTE EGM-10-002*, 2010.

[66] (The CMS Collaboration), "Transverse Momentum and Pseudorapidity Distributions of Charged Hadrons in pp Collisions at $\sqrt{s} = 7$ TeV," *Phys. Rev. Lett.*, vol. 105, p. 022002, Jul 2010.

[67] (The CMS Collaboration), "Electromagnetic calorimeter commissioning and first results with 7 TeV data," *CMS Note 2010/012*, 2010.

[68] (The CMS Collaboration), "Measurement of the Differential Isolated Prompt Photon Production Cross Section in pp Collisions at $\sqrt{s} = 7$ TeV," *CMS PAS QCD-10-037*, 2010.

[69] D. Acosta *et al.*, *CMS physics: Technical Design Report*, vol. I: Detector Performance and Software of *Technical Design Report CMS*, p. 365. Geneva: CERN, 2006.

[70] R. Ranieri, "The tracker geometry validation procedure," *CMS Note*, 2007.

[71] N. Adam, D. Bandurin, D. Ellman, V. Halyo, A. Hunt, and J. Werner, "Electron Reconstruction at Low pT," *CMS Analysis Note AN 2009/074*, 2009.

[72] D. Acosta *et al.*, *CMS physics: Technical Design Report*, vol. I: Detector Performance and Software of *Technical Design Report CMS*, p. 15. Geneva: CERN, 2006.

[73] R. P. S. Nourbakhsh and L. Soffi, "Energy scale of cms electromagnetic calorimeter with 2010 data," *CMS Analysis Note AN-11-038*, 2011.

[74] M. Anderson, A. Askew, A. Barfuss, D. Evans, F. Ferri, K. Kaadze, Y. Maravin, P. Meridiani, and C. Seez, "Review of clustering algorithms and energy corrections in ECAL," *CMS Internal Note IN 2010/008*, 2010.

[75] T. Awes, F. Obenshain, F. Plasil, S. Saini, S. Sorensen, and G. Young, "A simple method of shower localization and identification in laterally segmented calorimeters," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 311, no. 1-2, pp. 130 – 138, 1992.

[76] E. Meschi, T. Monteiro, C. Seez, and P. Vikas, "Electron Reconstruction in the CMS Electromagnetic Calorimeter," *CMS Internal Note IN 2010/008*, 2010.

[77] S. Cucciarelli and M. Konecki, "Track Parameter Evaluation and Primary Vertex Finding with the Pixel Detector," *CMS Note 2003/026*, 2003.

[78] D. Acosta *et al.*, *CMS physics: Technical Design Report*, vol. I: Detector Performance and Software of *Technical Design Report CMS*, p. 260. Geneva: CERN, 2006.

[79] S. Cucciarelli, "The Performance of the CMS Pixel Detector and the Primary Vertex Finding," *CMS Conference Report CR 2003/057*, 2003.

[80] (The CDF Collaboration), "Search for High-Mass Diphoton States and Limits on Randall-Sundrum Gravitons at CDF," *Phys. Rev. Lett.*, vol. 99, p. 171801, Oct 2007.

[81] (The CDF Collaboration), "Search for supersymmetry with gauge-mediated breaking in diphoton events with missing transverse energy at cdf ii," *Phys. Rev. Lett.*, vol. 104, p. 011801, Jan 2010.

[82] (The CMS Collaboration), "Isolated Photon Reconstruction and Identification at $\sqrt{s} =$ 7 TeV," *CMS NOTE EGM-10-006*, 2010.

[83] K. Nakamura (Particle Data Group), "Particle physics booklet," *JP G*, vol. 37, no. 075021, 2010.

[84] (The CMS Collaboration), "Measurement of the Muon Stopping Power in Lead Tungstate," *JINST*, vol. 5, p. P03007, 2010.

[85] W. Adam, S. Baffioni, F. Beaudette, D. Benedetti, C. Broutin, D. Chamont, C. Charlot, E. DiMarco, D. Futyan, S. H. D. Lelas, A. Martelli, P. Meridiani, M. Pioppi, I. Puljak, D. Sabes, R. Salerno, M. Sani, C. Seez, Y. Sirois, P. Vanlaer, and D. Wardrope, "Electron Reconstruction in CMS," *CMS Analysis Note AN-09-164*, 2009.

[86] F. James, *Statistical Methods in Experimental Physics*. Elsevier Science Ltd, 2006.

[87] R. Barlow, "Extended maximum likelihood," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 297, no. 3, pp. 496 – 506, 1990.

[88] J. Alwall, P. Demin, S. de Visscher, R. Frederix, M. Herquet, *et al.*, "MadGraph/MadEvent v4: The New Web Generation," *JHEP*, vol. 0709, p. 028, 2007.

[89] (The CMS Collaboration), "Measurements of inclusive W and Z cross sections in pp collisions at $\sqrt{s} = 7$ TeV," *Journal of High Energy Physics*, vol. 2011, pp. 1–40, 2011.

[90] M. Pivk and F. R. Le Diberder, "sPlot: a statistical tool to unfold data distributions," *Nucl. Instrum. Meth.*, vol. A555, pp. 356–369, 2005.

[91] K. Nakamura (Particle Data Group), "Particle physics booklet," *JP G*, vol. 37, no. 075021, 2010.

[92] (The CMS Collaboration), "Measurement of the Isolated Prompt Photon Production Cross Section in pp Collisions at $\sqrt{s} =$ 7 TeV," *Phys. Rev. Lett.*, vol. 106, Feb 2011.

[93] V. Blobel, "An unfolding method for high energy physics experiments," 2002.

[94] A. Bornheim, M. Gataullin, , and J. Veverka, "Photon Pixel Match Veto Efficiency Measurement with $Z \to \mu\mu\gamma$ Events Selected at 7 TeV," *CMS Analysis Note AN-11-089*, 2011.

[95] (The CDF Collaboration), "Measurement of the Cross Section for Prompt Diphoton Production in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV," *Phys. Rev. Lett.*, vol. 95, p. 022003, Jul 2005.

[96] (The CDF Collaboration), "Measurement of the Cross Section for Prompt Diphoton Production with 5.4 fb^{-1} of CDF Run II Data," *CDF Note 10160*, May 2010.

[97] (The D0 Collaboration), "Measurement of direct photon pair production cross sections in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV," *Physics Letters B*, vol. 690, no. 2, pp. 108 – 117, 2010.

[98] (The ATLAS Collaboration), "Measurement of the isolated di-photon cross-section in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector," *Phys. Rev. D*, July 2011.

[99] S. Alioli, P. Nason, C. Oleari, and E. D. Re, "NLO vector-boson production matched with shower in POWHEG," *JHEP*, vol. 07, p. 060, 2008.

[100] G. Cowan, "A survey of unfolding methods for particle physics," *Advanced Statistical Techniques in Particle Physics: Proceedings. Edited by M.R. Whalley and L. Lyons. Durham, UK, Inst. For Particle Physics Phenomenology, (IPPP-02-39)*, pp. 248–257, 2002.

[101] C. Anastasiou, E. Glover, and M. Tejeda-Yeomans, "Two loop QED and QCD corrections to massless fermion boson scattering," *Nucl.Phys.*, vol. B629, pp. 255–289, 2002.

[102] T. Binoth, J. P. Guillet, E. Pilon, and M. Werlen, "Beyond leading order effects in photon pair production at the Fermilab Tevatron," *Phys. Rev. D*, vol. 63, p. 114016, May 2001.

[103] M. Grazzini, "Higgs Production: Theory Overview," *Higgs Hunting*, 2011.