Di-photon cross section measurement and Higgs sensitivity study in the two-photon final state with the ATLAS detector

Li Yuan

THESE de DOCTORAT
de L'UNIVERSITE PIERRE ET MARIE CURIE

Spécialité physique des particules

La physique de la particule à la matière condensée (ED389)

Mesure de la section efficace de production de paires de photons et étude de sensibilité de recherche du Higgs dans le canal $H \to \gamma\gamma$ avec le détecteur ATLAS

présentée par

Li YUAN

pour obtenir

Le GRADE de DOCTEUR EN SCIENCES
DE L'UNIVERSITE PIERRE ET MARIE CURIE

Thèse soutenue le 22 Juin 2011 devant la Commission d'examen

M.	Jacques CHAUVEAU	
M.	Yuanning GAO	Rapporteur
M.	Shan JIN	
M^{me}	Lydia ROOS	Directrice de thèse
M.	Guillaume UNAL	
M.	Patrice VERDIER	Rapporteur

Résumé

Ce travail de thèse a été effectué auprès du détecteur ATLAS. Trois méthodes indépendantes sont proposées pour mesurer l'efficacité du déclenchement basé sur la détection de photons. Elles sont d'abord évaluées à l'aide de simulations Monte-Carlo, puis appliquées sur les données enregistrèes en 2010. Les deux méthodes utilisant des lots de photons montrent des résultats cohérents au niveau de quelques pourcents. Pour la méthode basée sur un lot d'électrons, la taille du lot est trop faible pour tirer des conclusions. Une mesure de la section efficace de production QCD de di-photons est effectuée sur les données 2010, correspondant à une luminosité de 37.2±1.2 pb^{-1}. Elle se base sur une méthode d'ajustement bi-dimensionnel pour extraire le signal. Les sections efficaces en fonction des observables $M_{\gamma\gamma}$, $p_{T\gamma\gamma}$ et $\Delta\phi_{\gamma\gamma}$ sont comparées avec les prédictions des générateurs DIPHOX et RESBOS. Un bon accord est trouvé pour la variable $M_{\gamma\gamma}$, tandis que des écarts sont observés dans les distributions $p_{T\gamma\gamma}$ et $\Delta\phi_{\gamma\gamma}$. Dans l'étude du canal $H \to \gamma\gamma$ basée sur échantillon simulé, une détérioration de 4% dans la limite d'exclusion est observée après incorporation de l'incertitude systématique découlant de la résolution sur la masse invariante de la paire de photons. Dans l'analyse des données réelles, la méthode d'ajustement 2D est également appliquée pour décomposer le bruit de fond dans la gamme de masse invariante [100, 150] GeV. La sensibilité de la recherche du boson de Higgs dans ce canal est alors réévaluée. Dans la gamme de masse [110-140] GeV, une production de Higgs de 3.2 à 4.2 fois le Modèle standard devrait être exclue avec $1fb^{-1}$ de données à $\sqrt{s} = 7$ TeV.

Mots-clés : QCD, section efficace, di-photon, efficacité de déclenchement, Higgs, sensibilité.

Abstract

This work is done with the ATLAS collaboration. Three independent methods are proposed to measure the photon trigger efficiency. They are first evaluated using Monte Carlo simulation and then applied on 2010 data. The two photon-based methods show consistent results, with efficiency discrepancy at a few % level. For the method based on electron sample, the statistics is too low to draw conclusion. A detailed QCD di-photon cross-section measurement is performed on a data sample corresponding to a luminosity of 37.2 ±1.2 pb^{-1}, in which a 2D fit method is introduced to extract the signal yields. The differential distributions of the observables $M_{\gamma\gamma}$, $p_{T\gamma\gamma}$ and $\Delta\phi_{\gamma\gamma}$ are derived and compared with the predictions from the DIPHOX and RESBOS generators. A good agreement is found for the $M_{\gamma\gamma}$ distribution, whereas discrepancies are observed in the $p_{T\gamma\gamma}$ and $\Delta\phi_{\gamma\gamma}$ distributions. In the study of the $H \to \gamma\gamma$ channel based on a simulated sample, a deterioration of 4% in the exclusion limit is observed after incorporating the systematic uncertainty arising from the resolution of the di-photon invariant mass. In the real data analysis, the 2D fit method is also applied to decompose the backgrounds in the di-photon candidate invariant mass range [100, 150] GeV. The sensitivity of the Higgs boson search in the two photon final state is then reappraised. 3.2−4.2 times the SM predicted cross-section in the 110-140 GeV mass range is expected to be excluded with $1fb^{-1}$ of data at $\sqrt{s} = 7$ TeV.

Keywords : QCD, cross section, di-photon, trigger efficiency, Higgs, sensitivity.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 The Standard Model of particle physics

One of the most successful theoretical models in particle physics is the so-called Standard Model (abbreviated as SM), in which, all matter is built from three generations of quarks and leptons [1].

The leptons carry integer electric charge. In each generation, a charged lepton is paired with a neutral lepton, called neutrino. The 2^{nd} and 3^{rd} generation of charged leptons are the muon and the tau, which are heavy versions of the electron. They are unstable and decay spontaneously to electrons, neutrinos and other particles.

The quarks carry fractional charges of $+2/3e$ or $-1/3e$. They are also grouped into three generations. They are named: up and down, charm and strange, top and bottom. There is a further degree of freedom for each quark, which is called color. In SM, the color charge of each quark can have three possible values, say red, blue or green. In this case, the total charge of quarks is then $3{\times}3 \times (2/3e + 1/3e) = 3e$, and the total charge of leptons is $3 \times (-1 + 0)e = -3e$. Therefore the total charge of quarks and leptons is 0. This is the actual condition that the SM should be free of so-called anomalies and is a renormalizable field theory.

There are four types of fundamental interactions in nature: the strong, the electromagnetic, the weak and the gravitational interactions. Each one has its corresponding force carriers. The detailed information on quarks, leptons and force carriers, including the symbol, the electric charge, the number of color charges and the mass, is summarized in Figure 1.1.

Strong interaction is the force that holds quarks together to form the protons, the neutrons and other particles, and also binds neutrons and protons within nuclei. This force is mediated by a massless particle, the gluon.

Electromagnetic interaction is the force between charged particles arising from their electric and magnetic fields. It is mediated by photon exchange.

Weak interaction is responsible for some particle decay, nuclear β decay, and neutrino absorption and emission. This force is mediated by the heavy W and Z bosons.

Gravitational interaction acts between all types of particles. It is by far the weakest of all the fundamental interactions. It is supposedly mediated by the exchange of a spin 2 boson, the graviton.

The description of fundamental interactions by four independent and unrelated interaction fields is a rather unsatisfactory scheme. It was speculated that the different inter-

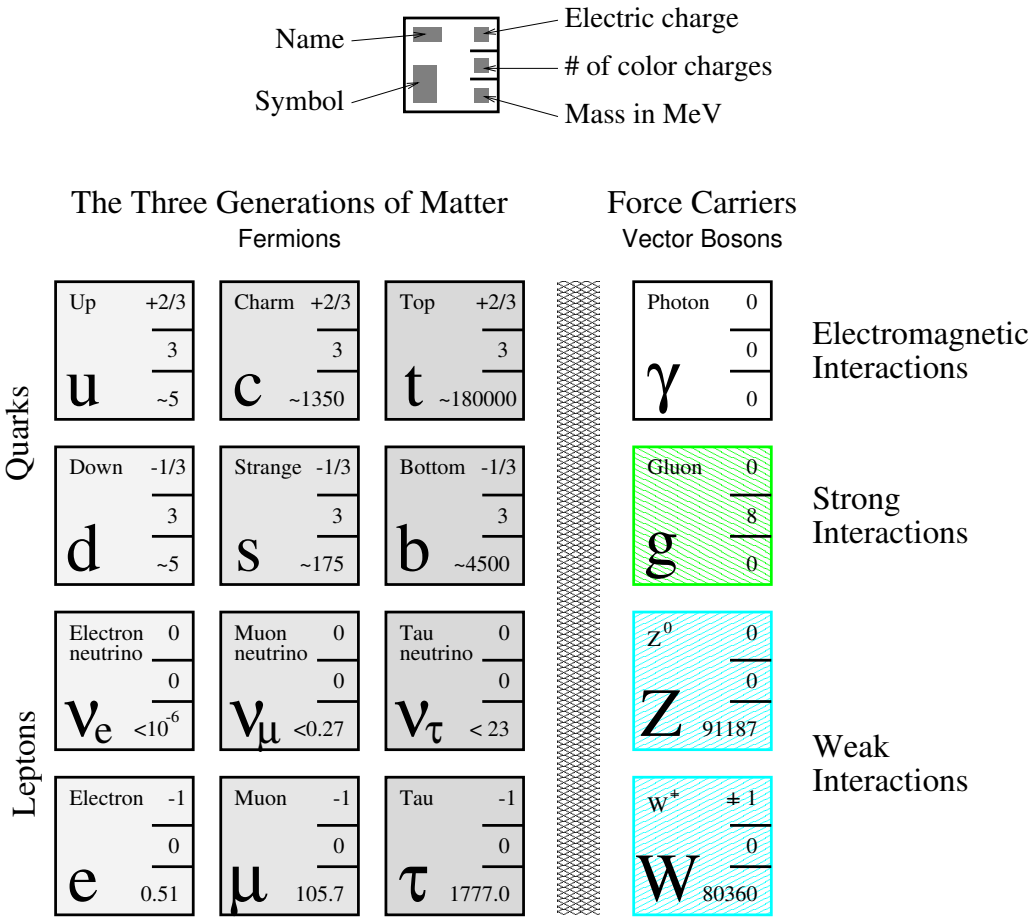


Figure 1.1: Quarks, leptons and interaction force carriers in SM.

actions are different aspects of a single, unified field. Enormous efforts were made towards the unification of the forces, especially during the 1960s [2]. It was partially achieved by the Standard Model, which provides a unified framework to describe the electromagnetic, weak and strong interactions based on a combination of local gauge symmetry groups: $SU(2)_L \otimes U(1)_Y \otimes SU(3)_C$.

The electroweak theory, proposed by Glashow, Salam and Weinberg [3] [4] to describe the electromagnetic and weak interactions between quarks and leptons, is based on the gauge symmetry group $SU(2)_L \otimes U(1)_Y$. It postulates four massless mediating bosons, arranged in a triplet and a singlet as members of multiplets of "weak isospin" I and "weak hypercharge" Y. Three of the bosons, are the components of an $I = 1$ triplet of the non-abelian group $SU(2)$, while the forth, is an isoscalar ($I = 0$) belonging to the abelian $U(1)$ group of weak hypercharge.

The quantum chromodynamic theory (QCD) [5] [6] [7] [8], based on the gauge symmetry group $SU(3)_C$, describes the strong interactions between the colored quarks. The bosons called gluons mediate the quark-quark interactions, each carries a color and an anticolor and belongs to an octet of $SU(3)_C$ group.

1.1.1 Gauge invariance in Quantum Electrodynamics

Quantum Electrodynamics (QED), the quantum field theory that describes the electromagnetic interactions, is characterized by two crucial properties, renormalizability and gauge invariance. The renormalization is the procedure that absorbs the divergent terms from the self-interaction contribution. The bare mass or charge m_0 or e_0 (unmeasurable value in the absence of the self-energy) is then multiplied by a dimensionless term, and becomes a physical value m or e that can be determined from experiment. The gauge invariance implies that the Lagrangian is invariant under the gauge transformation. The free Lagrangian that describes a dirac fermion is of the form:

$$\mathcal{L}_0 = i\overline{\phi(x)}\gamma^\mu\partial_\mu\phi(x) - m\overline{\phi(x)}\phi(x) \tag{1.1}$$

where $\phi(x)$ is a Dirac spinsor and γ^μ the Dirac matrices. The Lagrangian is invariant under $U(1)$ transformation,

$$\phi(x) \rightarrow \phi'(x) = exp[ie\theta]\phi(x) \tag{1.2}$$

where e is the charge, θ is an arbitrary real constant. If θ depends on space-time coordinate, the free Lagrangian is no longer invariant, as indicated:

$$\partial_\mu\phi(x) \rightarrow exp[ie\theta](\partial_\mu + ie\partial_\mu\theta)\phi(x) \tag{1.3}$$

In order to restore the gauge invariance, a new spin 1 field $A_\mu(x)$ is introduced, which transforms as:

$$A_\mu(x) \rightarrow A'_\mu(x) = A_\mu(x) + \frac{1}{e}\partial_\mu\theta \tag{1.4}$$

A covariant derivative is defined as:

$$D_\mu = \partial_\mu - ieA_\mu \tag{1.5}$$

It has the same transforming form as the field itself:

$$D_\mu\phi(x) \rightarrow (D_\mu\phi)'(x) = exp[ie\theta]D_\mu\phi(x) \tag{1.6}$$

The Lagrangian then becomes:

$$\mathcal{L} = i\overline{\phi(x)}\gamma^\mu D_\mu \phi(x) - m\overline{\phi(x)}\phi(x) = \mathcal{L}_0 + eA_\mu(x)\overline{\phi(x)}\gamma^\mu\phi(x) \tag{1.7}$$

A gauge-invariant kinetic term is also needed to make $A_\mu(x)$ a true propagating field:

$$\mathcal{L}_{kin} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{1.8}$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the usual electromagnetic field strength. A mass term is forbidden since it violates the gauge invariance. Therefore the photon is predicted to be massless.

1.1.2 Electroweak theory

The electroweak theory is based on $SU(2)_L \otimes U(1)_Y$, where $SU(2)_L$ is associated to the weak isospin I and $U(1)_Y$ on the weak hypercharge Y, which are related to the electric charge Q by: $Q = I_3 + Y/2$, where I_3 is the third component of the weak isospin. It is constructed based on the same principle of gauge invariance as described for the QED Lagrangian in the previous section. There is one gauge field associated to $U(1)_Y$, B_μ, and three associated to $SU(2)_L$, W^i_μ. B_μ couples to both the left- and right-handed components of the fermion fields, ϕ_L and ϕ_R, while the W^i_μ gauge fields only couple to the left-handed components. The gauge invariant Lagrangian of the electroweak theory can be written as:

$$\mathcal{L}_{EW} = i\overline{\phi}_L\gamma^\mu D_{\mu L}\phi_L + i\overline{\phi}_R\gamma^\mu D_{\mu R}\phi_R - \frac{1}{4}W^i_{\mu\nu}W_i^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \tag{1.9}$$

where the first two terms describe the interactions between fermions and the gauge fields, the last two terms are the gauge field terms. $D_{\mu L}$ and $D_{\mu R}$ are the covariant derivatives, which ensure the gauge invariance of the Lagrangian. They are defined as:

$$D_{\mu L}\phi_L = (\partial_\mu + ig\frac{\sigma_i}{2}W^i_\mu + ig'\frac{Y_L}{2}B_\mu)\phi_L \tag{1.10}$$

$$D_{\mu R}\phi_R = (\partial_\mu + ig'\frac{Y_R}{2}B_\mu)\phi_R \tag{1.11}$$

where g and g' are the $SU(2)_L$ and $U(1)_Y$ coupling constants, σ_i are the Pauli matrices, Y_L and Y_R are the weak hypercharge for the left- and right-handed components of the fermion fields respectively. The field strengths are defined by:

$$W^i_{\mu\nu} = \partial_\mu W^i_\nu - \partial_\nu W^i_\mu - g\epsilon^{ijk}W^j_\mu W^k_\nu \tag{1.12}$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \tag{1.13}$$

Here, ϵ^{ijk} is the structure constant of $SU(2)_L$. Linear combinations of the four gauge fields B_μ and W^i_μ represent the observable weak gauge bosons γ, Z and $W^\pm$:

$$W^\pm_\mu = \frac{W^1_\mu \mp W^2_\mu}{\sqrt{2}} \tag{1.14}$$

$$A_\mu = B_\mu \cos\theta_W + W^3_\mu \sin\theta_W \tag{1.15}$$

$$Z_\mu = -B_\mu \sin\theta_W + W^3_\mu \cos\theta_W \tag{1.16}$$

The Weinberg angle θ_W is defined as:

$$\cos\theta_W = \frac{g'}{\sqrt{g^2+g'^2}} \quad and \quad \sin\theta_W = \frac{g}{\sqrt{g^2+g'^2}} \tag{1.17}$$

1.1.3 Introduction to QCD

Quantum Chromodynamics (QCD), the gauge field theory that describes the strong interactions, is the $SU(3)_C$ component of the $SU(2)_L \otimes U(1)_Y \otimes SU(3)_C$ SM gauge group. Strong interactions are characterized by three basic properties: asymptotic freedom, confinement and dynamical chiral symmetry breaking.

The free QCD Lagrangian is described as:

$$\mathcal{L}_0 = \sum_f \overline{q}_f (i\gamma^\mu \partial_\mu - m_f) q_f, \tag{1.18}$$

where q_f is a vector notation for the quark field in the color space : $q_f = (q_f^1, q_f^2, q_f^3)$, and the index f denotes the quark flavour. This Lagrangian is invariant under arbitrary global $SU(3)_C$ transformations in color space. The transformation matrix can be written in the form:

$$U = exp[-ig_s \frac{\lambda^a}{2} \theta_a], \tag{1.19}$$

where λ^a $(a = 1, 2, \ldots, 8)$ denote the generators of the fundamental representation of the $SU(3)_C$ group, and θ_a are arbitrary parameters. The matrices λ^a are traceless and satisfy the commutation relations:

$$[\lambda^a, \lambda^b] = 2if^{abc}\lambda^c, \tag{1.20}$$

with f^{abc} the $SU(3)_C$ structure constants, which are real and totally antisymmetric.

If we require the Lagrangian to be also invariant under local $SU(3)_C$ transformations, $\theta_a = \theta_a(x)$, we need to change the quark derivatives by covariant objects. By including the 8 independent gauge parameters and 8 different gauge bosons (gluons) $G_a^\mu(x)$, the covariant derivative is then of the form:

$$D^\mu = [\partial^\mu - ig_s \frac{\lambda^a}{2} G_a^\mu(x)] = [\partial^\mu - ig_s G^\mu(x)] \tag{1.21}$$

Here the compact matrix notation is introduced:

$$[G^\mu(x)]_{\alpha\beta} = (\frac{\lambda^a}{2})_{\alpha\beta} G_a^\mu(x). \tag{1.22}$$

To build a gauge-invariant kinetic term for the gluon fields, the corresponding field strengths should be introduced:

$$\begin{aligned} G^{\mu\nu}(x) &= \frac{i}{g_s}[D^\mu, D^\nu] = \partial^\mu G^\nu - \partial^\nu G^\mu - ig_s[G^\mu, G^\nu] = \frac{\lambda^a}{2} G_a^{\mu\nu}(x), \\ G_a^{\mu\nu}(x) &= \partial^\mu G_a^\nu - \partial^\nu G_a^\mu + g_s f^{abc} G_b^\mu G_c^\nu. \end{aligned} \tag{1.23}$$

Under a gauge transformation,

$$G^{\mu\nu} \to (G^{\mu\nu})' = U G^{\mu\nu} U^\dagger.$$

Taking the proper normalization for the gluon kinetic term, the Lagrangian is then of the form:

$$\begin{aligned} \mathcal{L}_{QCD} &= -\frac{1}{4}(\partial^\mu G_a^\nu - \partial^\nu G_a^\mu)(\partial_\mu G_\nu^a - \partial_\nu G_\mu^a) + \sum_f \overline{q}_f^\alpha (i\gamma^\mu \partial_\mu - m_f) q_f^\alpha \\ &\quad + g_s G_a^\mu \sum_f \overline{q}_f^\alpha \gamma_\mu \left(\frac{\lambda^a}{2}\right)_{\alpha\beta} q_f^\beta \\ &\quad - \frac{g_s}{2} f^{abc}(\partial^\mu G_a^\nu - \partial^\nu G_a^\mu) G_\mu^b G_\nu^c - \frac{g_s^2}{4} f^{abc} f_{ade} G_b^\mu G_c^\nu G_\mu^d G_\nu^e. \end{aligned} \tag{1.24}$$

In the first line, one can find the correct kinetic terms for the different fields and also the terms for the corresponding propagators. In the second line, the color interaction between quarks and gluons is given, it involves the $SU(3)_C$ matrices λ^a. Finally, owing to the non-abelian character of the colour group, the $G_a^{\mu\nu}G_{\mu\nu}^a$ term generates the cubic and quartic gluon self-interactions shown in the last line; the strength of these interactions is given by the same coupling g_s which appears in the fermionic piece of the Lagrangian. The Lagrangian 1.24 looks actually very simple because of its colour-symmetry properties. All interactions here are given in terms of a single universal coupling g_s, which is called the *strong coupling constant*. The existence of self-interactions among the gauge fields, which actually introduce the loop contribution in the Feynman diagrams, can explain properties like asymptotic freedom and confinement.

The computation of perturbative corrections to the tree-level results involves divergent loop integrals. In order to get finite results with physical meaning, renormalization is performed to reabsorb the divergences through a redefinition of the original field and couplings. After the redefinition, the coupling, denoted $\alpha_s(Q^2)$, becomes:

$$\alpha_s(Q^2) = \frac{\alpha_s(Q_0^2)}{1 - \frac{\beta_1 \alpha_s(Q_0^2)}{2\pi} \ln(Q^2/Q_0^2)} \tag{1.25}$$

where Q^2 is the energy scale, β_1 is the first coefficient of the β-function[1] [8] [9]. From this equation, one can see the coupling depends on the energy scale Q^2. One can easily get

$$\lim_{Q^2 \to \infty} \alpha_s(Q^2) = 0.$$

The coupling decreasing at short distance indicates QCD indeed has the property of asymptotic freedom.

On the other hand, at large distances, the strength of the coupling increases: this is the confinement property. However, this indeed can not be described by perturbative QCD anymore. There are certainly other tools, such as QCD sum rules and lattice calculations, to investigate particular aspects of non-perturbative physics.

1.1.4 The Higgs mechanism

The description of electroweak and strong interactions in SM is based on gauge field theory, in which the quantum field must preserve gauge invariance.

In the Lagrangian of Equation 1.7, there is no term corresponding to the square of field $A_\mu(x)$: it implies that the associated boson, the photon, is massless. On the contrary, we know from the experimental results that the vector bosons of the weak interaction, the

[1]This function is used to denote the scale dependence of $\alpha_s(Q^2)$, as indicated:

$$\mu \frac{d\alpha}{d\mu} = \alpha\beta(\alpha) \tag{1.26}$$

$$\beta(\alpha) = \beta_1 \frac{\alpha}{\pi} + \beta_2 (\frac{\alpha}{\pi})^2 + ... \tag{1.27}$$

The renormalization of the QCD coupling includes the contribution from gluon self-interactions. The calculation of the relevant one-loop diagrams gives:

$$\beta_1 = \frac{2N_f - 11N_C}{6} \tag{1.28}$$

where N_f, N_C denote the number of flavors, color charges for quark respectively. Thus, $\beta_1 < 0$ if $N_f \leq 16$.

W and Z, are massive, and with a mass about 80−90 GeV. However, if a mass item like $\frac{1}{2}M_A^2 A_\mu A^\nu$ is inserted in the Lagrangian, the invariance under the gauge transformation is violated. The Higgs mechanism was introduced to address this problem.

In the Higgs mechanism, a doublet of complex scalar fields is introduced:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \tag{1.29}$$

here ϕ^0 and ϕ^+ denote a neutral and a charged field respectively. The scaler field ϕ is integrated in the Lagrangian of the electroweak theory in the form as:

$$\mathcal{L}_H = (D^\mu \phi)^\dagger (D_\mu \phi) - V(\phi), \tag{1.30}$$

where $V(\phi)$ is a potential term with the form:

$$V(\phi) = \frac{1}{2}\mu^2\phi^2 + \frac{1}{4}\lambda\phi^4, \tag{1.31}$$

and the covariant derivative is of the form:

$$D_\mu \phi = (\partial_\mu + ig\frac{\sigma_i}{2}W_\mu^i + ig'\frac{Y_\phi}{2}B_\mu)\phi, \quad Y_\phi = 1. \tag{1.32}$$

The Lagrangian 1.30 is invariant under $SU(2)_L \otimes U(1)_Y$ transformation. The minimum value of $V(\phi)$ occurs at $\phi = \phi_0$, when $\frac{\partial V}{\partial \phi} = 0$ or

$$\phi(\mu^2 + \lambda\phi^2) = 0 \tag{1.33}$$

If $\mu^2 > 0$, the situation for a massive particle, then $\phi = \phi_0$ when $\phi = 0$; this is the normal situation for the lowest energy vacuum state with $V = 0$. However, if $\mu^2 < 0$, then

$$\phi = \phi_0 \quad when \quad \phi = \pm\sqrt{\frac{-\mu^2}{\lambda}} \tag{1.34}$$

Here the lowest energy state has ϕ finite, with $V = -\frac{\mu^4}{4\lambda}$, so that V is everywhere a non-zero constant. Figure 1.2 shows V as a function of ϕ, both for $\mu^2 > 0$ and $\mu^2 < 0$. In either case a symmetric curve result, but for $\mu^2 < 0$ there are two minima, $\phi_0 = +\sqrt{\frac{-\mu^2}{\lambda}}$ and $-\sqrt{\frac{-\mu^2}{\lambda}}$.

The absolute value of the field at the minima of the potential is the so called *vacuum expected value* of the field, $|\phi_0| = \sqrt{\frac{-\mu^2}{\lambda}} = \upsilon/\sqrt{2}$. The reference ground state for the local gauge transformation is chosen as:

$$\phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \upsilon \end{pmatrix}, \quad (\phi_1 = \phi_2 = \phi_4 = 0,\ \phi_3 = \upsilon) \tag{1.35}$$

It breaks the $SU(2)_L$ symmetry while the Lagrangian remains invariant under $SU(2)_L \otimes U(1)_Y$ transformations. This behavior is called *spontaneous symmetry breaking.*

The excitation of the scalar field can be written as:

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \upsilon + H \end{pmatrix} \tag{1.36}$$

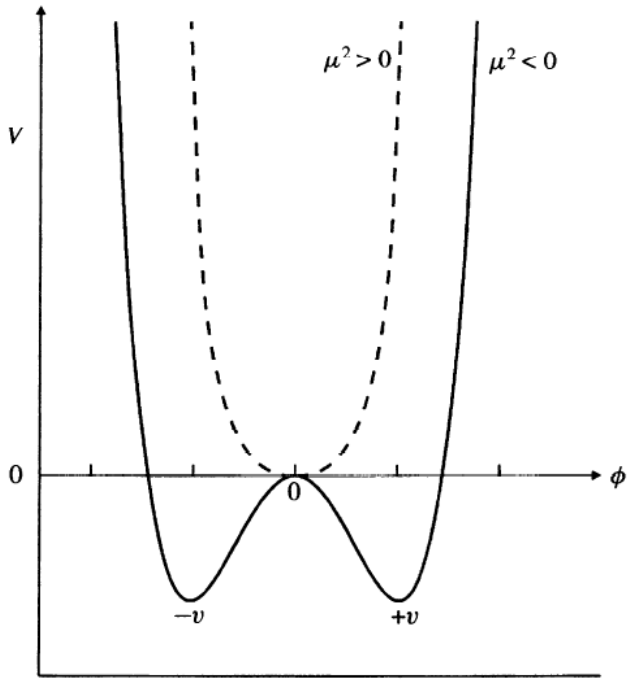


Figure 1.2: The potential V as a function of a one-dimensional scaler field ϕ for the two cases $\mu^2 > 0$ and $\mu^2 < 0$.

with H denoting the scalar Higgs field which represents the physical Higgs boson. Inserting the Equations 1.15− 1.16 and 1.36 into the Lagrangian 1.30, the following result is obtained:

$$\mathcal{L}_H = \frac{1}{2}\partial^\mu H \partial_\mu H + \frac{g^2 v^2}{8}(W_\mu^+ W^{+\mu} + W_\mu^- W^{-\mu}) + \frac{g^2 v^2}{8\cos^2\theta_W} Z_\mu Z^\mu - V(\phi), \tag{1.37}$$

One can get the mass terms for Z and $W^\pm$ bosons given by:

$$m_W = \cos\theta_W m_Z = \frac{gv}{2} \tag{1.38}$$

and the photon remains massless, $m_A = 0$.

The Higgs mass is given by the relation

$$m_H = \sqrt{-2\mu^2} = \sqrt{2\lambda v^2}.$$

λ reminds a free parameter of the potential, thus the Higgs mass is a free parameter of the theory.

Fermion masses are generated by introducing additional terms in the Lagrangian describing the Yukawa coupling of the fermions to the scalar field ϕ. The Yukawa coupling terms are of the form:

$$\mathcal{L}_{Yukawa} = g_\psi(\overline{\psi_L}\phi\psi_R + \overline{\psi_R}\phi^\dagger\psi_L) \tag{1.39}$$

where g_ψ denotes the Yukawa coupling constant of the fermion. The local gauge transformation H introduces two terms per fermion in the Lagrangian, which is of the form:

$$\frac{g_L}{\sqrt{2}} v(\overline{\psi_L}\psi_R + \overline{\psi_R}\psi_L) + \frac{g_L}{\sqrt{2}}(\overline{\psi_L}\psi_R + \overline{\psi_R}\psi_L)H \tag{1.40}$$

The mass of the fermions are then given by the Yukawa couplings: $m_L = \frac{g_L v}{\sqrt{2}}$. This is for the lower member of the left-handed doublets, which works for the leptons when assuming the neutrinos are massless. For the upper member of quark doublets, the Higgs field would need to have a non-null upper component.

1.1.5 Theoretical constraints on the Higgs mass

The Higgs mass is a free parameter of the SM. There are theoretical constraints which can be derived from assumptions on the energy range in which the SM is valid before perturbation theory breaks down and new physics should emerge. These include constraints from unitarity in scattering amplitudes, perturbativity of the Higgs self-coupling, stability of the electroweak vacuum and fine-tuning.

- **Perturbative unitarity**

  The elastic scattering diagrams of WW are shown in Figure 1.3. If one includes only the first two types of contribution (a) and (b) in the perturbative approximation, the cross-section will increase with the scattering energy, therefore the unitarity for energy above 1.2 TeV will be violated. The Higgs mechanism introduces the third diagram (c). For a certain Higgs mass, it balances the contributions that increase with the scattering energy. Taking into account the fact that WW can be coupled with the other channels (ZZ, HH, ZH, W^+H and W^+Z), the unitarity constraint yields an upper limit on the Higgs mass [12]: $m_H \lesssim 710$ GeV.

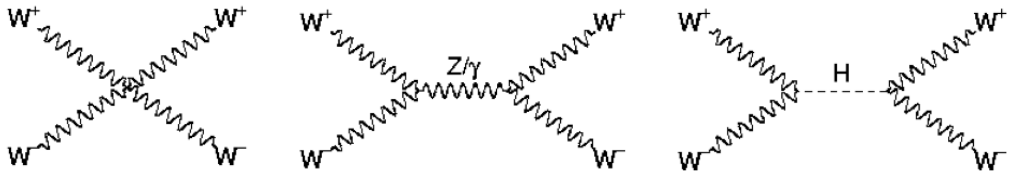


Figure 1.3: Some Feynman diagrams for W^+W^- scattering.

- **Triviality and stability bounds**

  As discussed in Section 1.1.3, the couplings which appear in the SM Lagrangian, depend on the considered energy. This is also the case for the quartic Higgs coupling which will be monotonically increasing with the energy scale $|Q|$. The variation of the quartic Higgs coupling with the energy scale Q is described by the Renormalization Group Equation (RGE) [13] [14] [15],

$$\frac{d}{dQ^2}\lambda(Q^2) = \frac{3}{4\pi^2}\lambda^2(Q^2) + higher\ orders \tag{1.41}$$

  The solution of this equation, choosing the natural reference energy point to be the electroweak symmetry breaking scale, $Q_0 = \nu$, reads at one-loop

$$\lambda(Q^2) = \lambda(\nu^2)\left[1 - \frac{3}{4\pi^2}\lambda(\nu^2)\log\frac{Q^2}{\nu^2}\right] \tag{1.42}$$

In order to get finite physical results where self-coupling λ remains finite, an energy cut-off Λ_C is set at Higgs boson mass itself, $\Lambda_C = M_H$. From simulations of gauge theories on the lattice, where the non-perturbative effects are properly taken into account, one can obtain a rigorous bound $M_H < 640$ GeV.

If including the additional contributions from fermions and gauge bosons besides the Higgs boson, the coupling of λ could not be too small, otherwise it will lead to a scalar potential $V(Q^2) < V(\nu)$, where the vacuum is not stable anymore since it has no minimum. A requirement on the coupling of λ puts a strong constraint on the Higgs boson mass, which depends on the value of the cutoff Λ_C. For the stability bound, one simply requires that the coupling λ remains positive at the cutoff scale, $\lambda(\Lambda_C) > 0$. Including the theoretical uncertainties by a variation of the cut-off Λ_C using the matching conditions for the top quark and Higgs boson masses, and the experimental errors mainly on $\alpha_s = 0.118 \pm 0.002$ and $m_t = 175 \pm 6$ GeV, one obtains the constraints from the stability (lower band) and triviality (upper band) shown in Figure 1.4. It shows the allowed range of M_H as a function of the scale of New Physics Λ_C (between the bands).

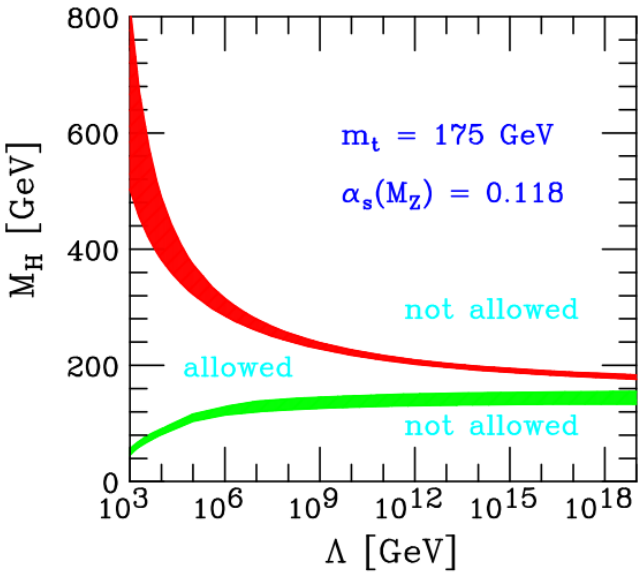


Figure 1.4: The triviality (upper) bound and the vacuum stability bound (lower) bound on the Higgs mass as a function of the New Physics or cut-off scale Λ_C. The colored/shaded bands illustrate the impact of various uncertainties. See Reference [16].

- **The fine-tuning constraint**

  The last theoretical constraint comes from the fine-tuning problem originating from the radiative corrections to the Higgs boson mass. The Feynman diagrams for one-loop radiative corrections to the SM Higgs boson mass are shown in Figure 1.5. Cutting off the loop integral momenta at a scale Λ, and keeping only the dominant contribution in this scale, one obtains

$$M_H^2 = (M_H^0)^2 + \frac{3\Lambda^2}{8\pi^2\nu^2}[M_H^2 + 2M_W^2 + M_Z^2 - 4m_t^2] \tag{1.43}$$

where M_H^0 is the bare mass contained in the unrenormalized Lagrangian and where we retained only the contribution of the top heavy quark for the fermion loops. If the cutoff Λ is very large, for instance of the order of the Grand Unification scale $\sim 10^{16}$ GeV, one needs a very fine arrangement of 16 digits between the bare Higgs mass and the radiative corrections to have a physical Higgs boson mass in the range of the electroweak symmetry breaking scale, $M_H \sim 100$ GeV to 1 TeV, as is required for the consistency of the SM. This is the naturalness of fine-tuning problem. One can obtain a very useful information by considering the fine-tuning problem in the SM at scales of a few tens of TeV. In the vicinity of these scales, a Higgs boson with a mass $M_H \sim$ 200 GeV can still allow for an acceptable amount of fine-tuning.

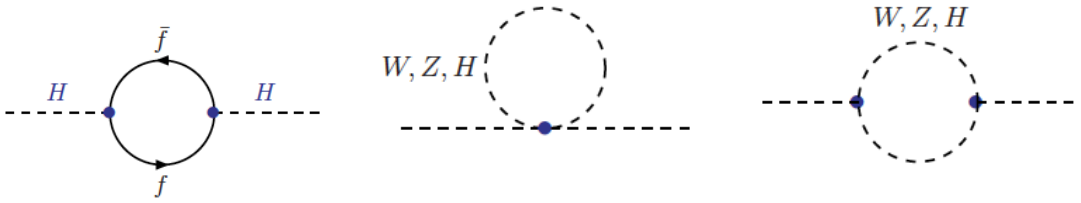


Figure 1.5: Feynman diagrams for the one-loop radiative corrections to the SM Higgs boson mass. The loop involves Higgs boson, massive gauge boson and fermion.

1.2 Review of Higgs searches

1.2.1 Direct searches

The four experiments (ALEPH, DELPHI, L3 and OPAL) at Large Electron-Positron collider (LEP) have collected a total of 2461 pb^{-1} of e^+e^- collision data at centre-of-mass energies between 189 and 209 GeV. At LEP, the SM Higgs boson was expected to be produced mainly in association with the Z boson through the Higgs-strahlung process $e^+e^- \to HZ$, and decay mainly into $b\bar{b}$ quark pairs. Thus the main final states that contributed to the direct searches are: the four-jet final state $(H \to b\bar{b})(Z \to b\bar{b})$, the missing energy final state $(H \to b\bar{b})(Z \to \nu\overline{\nu})$, the leptonic final state $(H \to b\bar{b})(Z \to e^+e^-$ or $\mu^+\mu^-)$ and the τ lepton final states $(H \to b\bar{b})(Z \to \tau^+\tau^-)$ and $(H \to \tau^+\tau^-)(Z \to q\bar{q})$. The combination of the results from the four experiments has set a lower bound of 114.4 GeV (Figure 1.6) on the mass of the SM Higgs boson at 95% confidence level [18]. At the mass of 115 GeV, the ALEPH experiment [17] has reported a 2.8σ excess of SM Higgs boson production. However, it was not confirmed by the other LEP experiments.

At Tevatron, the two experiments (CDF and DØ) also perform the direct search for SM Higgs boson with $p\overline{p}$ collision data at centre-of-mass energy of 1.96 TeV. During summer 2010, the Tevatron experiments have presented a combined exclusion limit based on up to 5.9 fb^{-1} for CDF and up to 6.7 fb^{-1} for DØ. The result is shown in Figure 1.7. A SM Higgs in the mass range 158< m_H <175 GeV is excluded. An updated result was shown by the Tevatron experiments at the winter 2011 conferences [28]. The analysis, in the 130 to 200 GeV mass range, leads to a new exclusion region: 158< m_H <173 GeV.

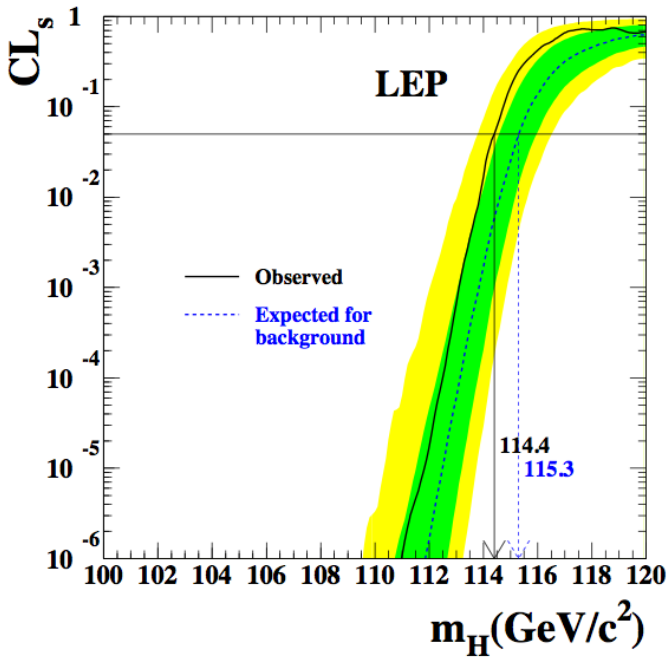


Figure 1.6: The exclusion limit for SM Higgs at LEP. The ratio of $CL_s = CL_{s+b}/CL_b$ for the signal and background hypothesis is shown. These quantities are detailed in Chapter 6. Solid line: observation; dashed line: median background expectation. The green and yellow shaded bands around the median expected line correspond to the 68% and 95% probability bands. The intersection of the horizontal line for $CL_s = 0.05$ with the observed curve is used to define the 95% confidence level lower bound on the mass of the Standard Model Higgs boson [18].

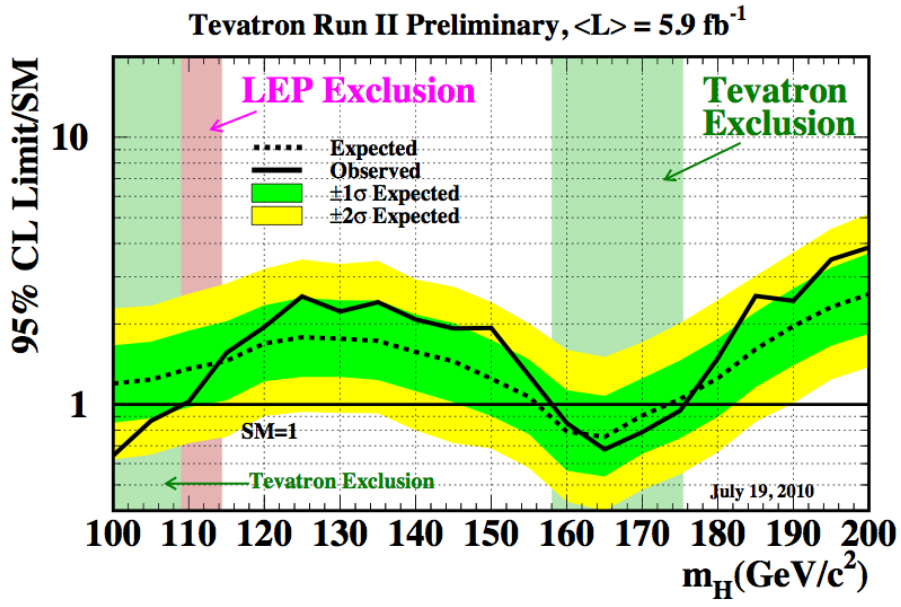


Figure 1.7: The exclusion limits for SM Higgs at Tevatron. Observed and expected (median, for the background-only hypothesis) upper limits are shown on the ratios to SM cross-section at 95% of confidence level, as a function of the Higgs boson mass. The bands indicate the 68% and 95% probability regions where the limits can fluctuate, in the absence of signal [27].

1.2.2 Indirect constraints

In SM, the EW sector is basically described by three parameters: the W and Z masses and the electromagnetic constant coupling α_{em}. Since the Higgs boson contributes to the radiative corrections, precision measurements of electroweak observables can be used to put indirect limits on its mass.

In the high-Q^2 domain, many SM parameters can be measured, i.e. the W boson mass m_W, width Γ_W and the top mass m_t. The combined result from Tevatron [19] [20] and LEP gives $m_W = 80.399 \pm 0.023$ GeV, $\Gamma_W = 2.085 \pm 0.042$ GeV [26]. For the mass of the top quark, the result from Tevatron is $m_t = 173.3 \pm 1.1$ GeV [21].

From the low-Q^2 interactions, the results of atomic parity violation in caesium [22], the electroweak mixing angle measurement in Moller scattering [23], and neutrino-nucleon neutral-to-charged current cross-section ratios performed by the NuTeV collaboration [24], are used in the constraint of the SM.

The SM predictions are calculated with ZFITTER based on a χ^2 minimization technique, as a function of five SM parameters: $\Delta\alpha^{(5)}_{had}(m_Z^2)$, $\alpha_S(m_Z^2)$, m_Z, m_t and $\log_{10}(m_H/GeV)$ [25]. m_Z denotes the mass of the Z boson and $\Delta\alpha^{(5)}_{had}(m_Z^2)$ is the five-quark hadronic contribution to the running QED coupling constant. Those five parameters are varied simultaneously in the fits. Indirect constraints on some parameters, i.e. m_W, can thus be obtained. They are found to be in good agreement with the direct measurements.

The most stringent constraints on m_H are obtained when all high-Q^2 measurements are used in the fit. In Figure 1.8, from Ref. [26], the observed value of $\Delta\chi^2 = \chi^2 - \chi^2_{min}$ for a fit including all high-Q^2 results is plotted as a function of m_H. The 95% one-sided confidence level (C.L.) upper limit on m_H (taking the theory-uncertainty band into account) is 158

GeV. The 95% C.L. lower limit of 114.4 GeV obtained from direct searches at LEP-II and the region between 158 GeV and 175 GeV excluded by the Tevatron experiments are not used in the determination of this limit.

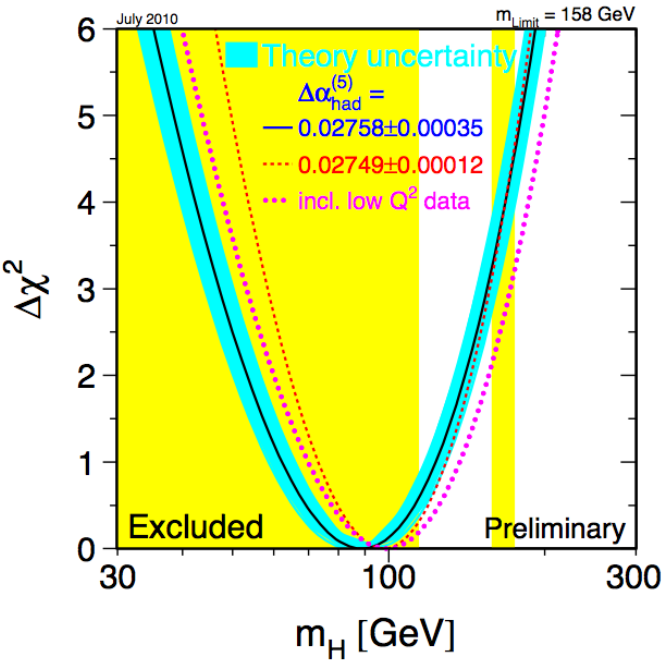


Figure 1.8: $\Delta\chi^2 = \chi^2 - \chi^2_{min}$ vs. m_H curve. The line is the results from the fit using all high-Q^2 data; the band represents an estimate of the theoretical error. The yellow band show the exclusion limit from the direct searches at LEP and Tevatron. The dotted curve corresponds to a fit including also the low-Q^2 data [25].

Using most of the up-to-date experimental results as the inputs, the Gfitter group has performed a global electroweak fit of the SM. The complete fit which includes the results from direct searches for the Higgs boson at LEP and Tevatron (up to the summer 2010), gives the constrained result: $m_H = 120.6^{+17.9}_{-5.2}$ GeV. Figure 1.9 shows the $\Delta\chi^2$ as a function of m_H distribution obtained with the complete fit [29].

1.3 Review of di-photon cross-section measurements

One aspect of this thesis is the measurement of the di-photon production cross-section. The di-photon production is very interesting on its own right. The direct measurement of the transverse momentum of the $\gamma\gamma$ system ($p_{T_{\gamma\gamma}}$), especially in the low $p_{T_{\gamma\gamma}}$ region, is sensitive to the initial state soft gluon radiation, whereas the high $p_{T_{\gamma\gamma}}$ region is sensitive to the parton fragmentation into photons. The fragmentation contribution can also be tested through the distribution of the azimuthal angle between the two photons, $\Delta\phi_{\gamma\gamma}$. QCD di-photon production is also the main background of many interesting physics processes, and first of all, of the SM Higgs decay in the $\gamma\gamma$ final state. Other processes include new physics signatures involving $\gamma\gamma$ final state, such as extra dimensions and cascade decay of heavy new particles.

Measurements involving the di-photon final state have been previously carried out at

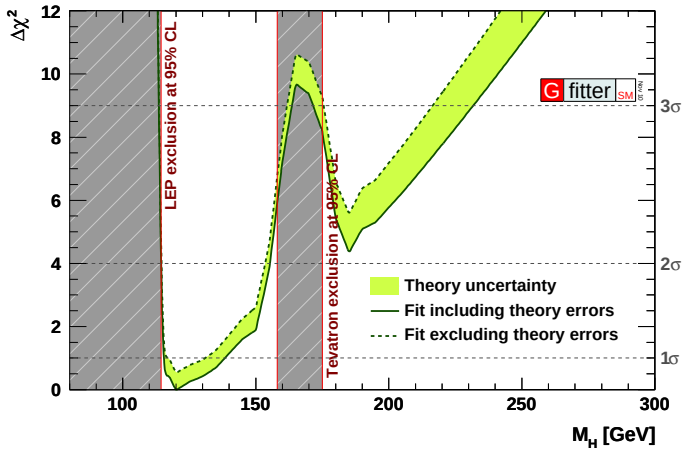


Figure 1.9: $\Delta\chi^2$ as a function of m_H obtained by Gfitter group. The solid (dashed) line corresponds to the results when including (ignoring) the theoretical errors.

fixed-target [30] [31] and collider experiments [32] [33] [34] [35]. Figure 1.10 shows the differential cross-sections measured by the CDF collaboration with 5.4 fb^{-1} of data at a centre-of-mass energy of 1.96 TeV. The data are compared with three pertinent model calculations, at leading order (PYTHIA [36]), fixed next-to-leading order (DIPHOX [83]), and p_T resummed matched to next-to-leading order (RESBOS [80]) in the strong coupling at the hard scattering. However, none of the models adequately describes all aspects of the data. One can also find the results obtained by DØ collaboration in Figure 1.11. The conclusion from those plots is the same. The study of the di-photon production at LHC offers a complementary check of the models since the relative fraction of the various subprocesses is different than at the Tevatron. Moreover, the very high statistics expected at LHC will allow to do precision tests.

1.4 Motivation and structure for this thesis

As discussed in the previous sections, both theoretical constraints and experimental searches favor a SM Higgs boson with a low mass. The Higgs decay channel with two photons in the final state is one of the most promising channels for Higgs search at LHC in the low mass region. The LHC machine has been operating with a centre-of-mass energy at 7 TeV in 2010, and the detectors recorded high quality data. The main topic of this thesis is the di-photon cross-section measurement based on the ATLAS data recorded in 2010. A study of the background to the Higgs decay channel in two photons is also presented, as well as the prospects for excluding a low mass Higgs.

The next chapters of this thesis are organized as follows:

- **Chapter 2**: a general description of the LHC machine and the ATLAS detector is given. The designed parameters and performance for each subdetector of ATLAS are presented. The data taking status in 2010 running is also included.
- **Chapter 3**: this chapter presents the algorithms used in the ATLAS experiment for the photon reconstruction, calibration and identification, which are fundamental ingredients for the physics analyses in the following chapters.

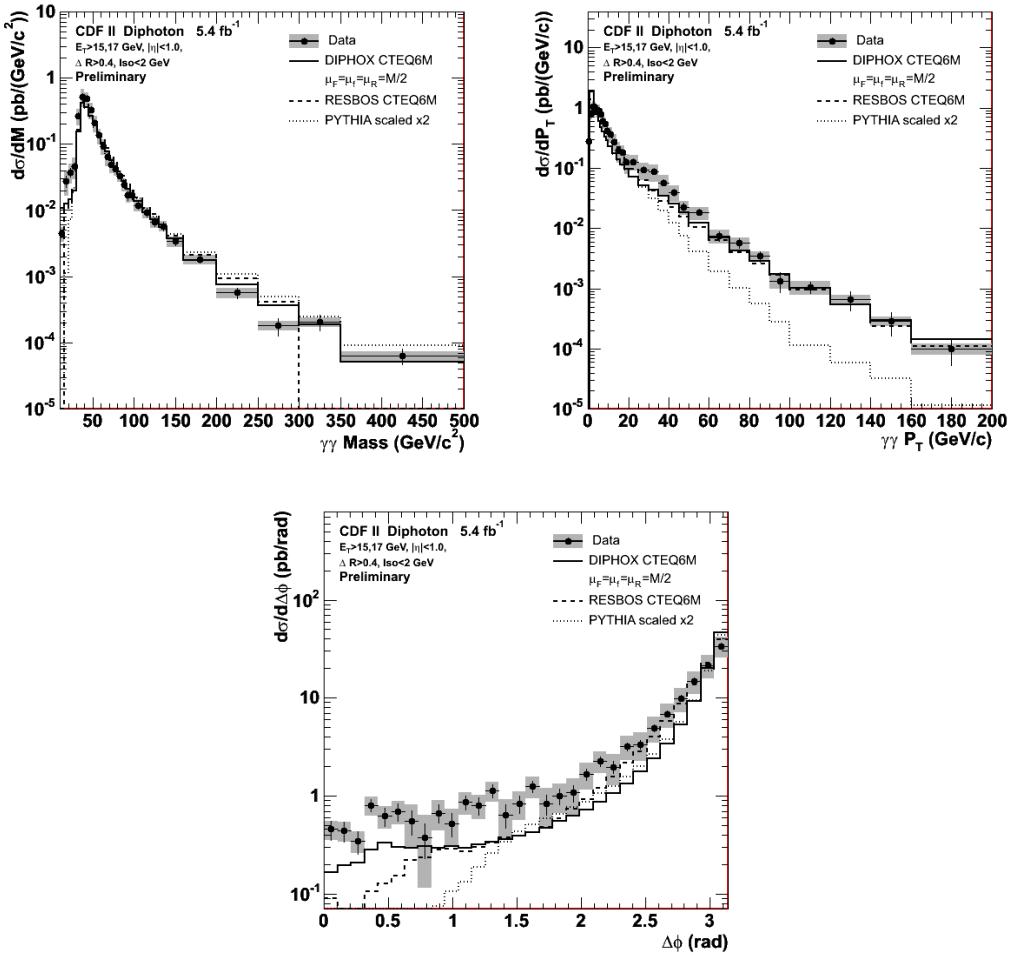


Figure 1.10: The measured differential cross-section as functions of $M_{\gamma\gamma}$ (top left), $p^T_{\gamma\gamma}$ (top right), $\Delta\phi_{\gamma\gamma}$ (bottom), from the CDF experiment compared with three theoretical models [34].

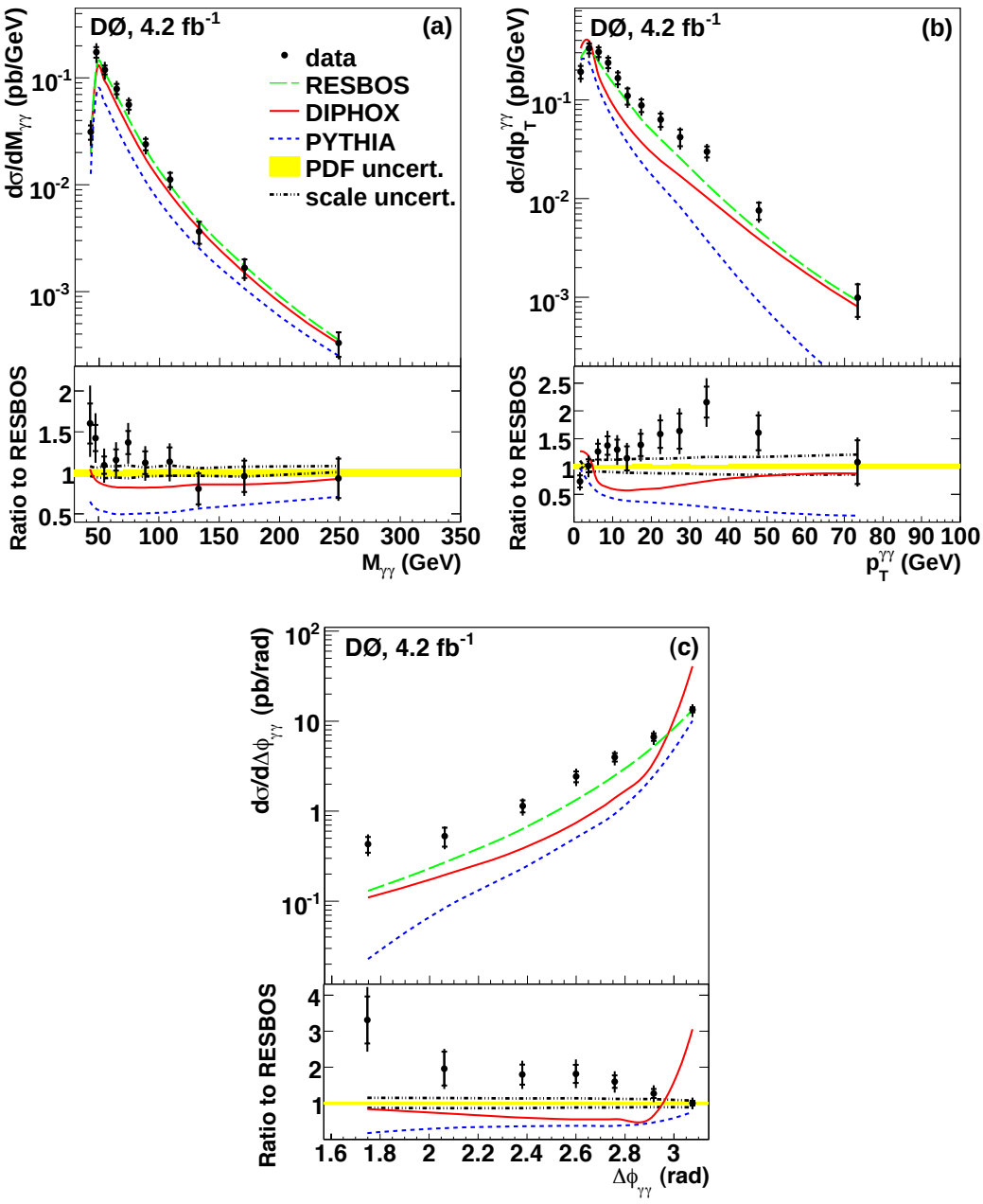


Figure 1.11: The measured differential cross-section as functions of $M_{\gamma\gamma}$ (top left), $p^T_{\gamma\gamma}$ (top right), $\Delta\phi_{\gamma\gamma}$ (bottom), in DØ experiment compared with three theoretical models. In the bottom plots, the ratio of differential cross-sections between data and RESBOS are displayed as black points with uncertainties. The solid (dashed) line shows the ratio of the predictions from DIPHOX (PYTHIA) to those from RESBOS [35].

- **Chapter 4**: the methods developed to measure the photon trigger efficiency are discussed in details. The results from the Monte Carlo study and from the 2010 data are shown.
- **Chapter 5**: this chapter presents the measurement of the di-photon cross-section. A two-dimensional fit method to extract the di-photon yield is described. The measured differential cross-sections are shown and compared to theoretical predictions.
- **Chapter 6**: $H \rightarrow \gamma\gamma$ analysis. A background study and the exclusion prospect of a SM Higgs in the $H \rightarrow \gamma\gamma$ channel are discussed.
- **Chapter 7**: Conclusion.

Chapter 2

Accelerator and detector

2.1 Introduction to LHC

The Large Hadron Collider (LHC) [37] is a proton-proton collider located at CERN, about 20 kilometers from Geneva centre. It re-uses the 27 kilometers circumference underground tunnel (Figure 2.1) [44] that was built for the previous big accelerator, the Large Electron-Positron collider (LEP). As for accelerators like the LHC, the most important parameters are the beam energy and the number of interesting collisions, since the production rate of a particular process varies with these two quantities (Figure 2.2). The main constraints for the energy of the LHC that can be reached are from the magnets, cavities, the size of the tunnel and other essential elements of the machine. With all the constraint, the energy per proton beam for the LHC could be up to 7 TeV. The number of interesting collisions is quantified by the luminosity, defined as:

$$\mathcal{L} = \frac{N_b^2 n_b f_{rev} \gamma_r}{4\pi \varepsilon_n \beta_*} F \tag{2.1}$$

Here, N_b is the number of particles per bunch, n_b is the number of bunches per beam, f_{rev} is the revolution frequency, γ_r is the relativistic gamma factor, ε_n is the normalized transverse beam emittance, β_* is the beta function at the collision point, and F is the geometric luminosity reduction factor due to the crossing angle at the interaction point. The LHC is designed to collide proton beams at a centre-of-mass energy of 14 TeV and an instantaneous luminosity of $10^{34} cm^{-2} s^{-1}$, which will extend the frontiers of particle physics. The nominal number of protons per bunch is 1.15×10^{11}. Beam crossing are 25 ns apart, thus bunches of protons will collide 40 million times per second. The LHC has also the capacity to collide heavy ions, in particular lead nuclei, at 2.8 TeV per nucleon, at a design luminosity of $10^{27} cm^{-2} s^{-1}$.

The LHC uses the most advanced superconducting magnet and accelerator technologies ever employed. As for a particle-particle collider, there should be two rings with counter-rotating beams, unlike particle-antiparticle colliders that can have both beams sharing the same phase space in a single ring. Due to the hard limit on space, it led to the "two-in-one" superconducting magnet design. This design accommodates the windings for the two beam channels in a common cold mass and cryostat, with magnetic flux circulating in the opposite sense through the two channels. In order to have proton beams at 7 TeV energy, LHC relies on superconducting magnets that are expected to operate at a field above 8 T. At the experiments of Tevatron at FNAL, HERA at DESY and RHIC at BNL, the

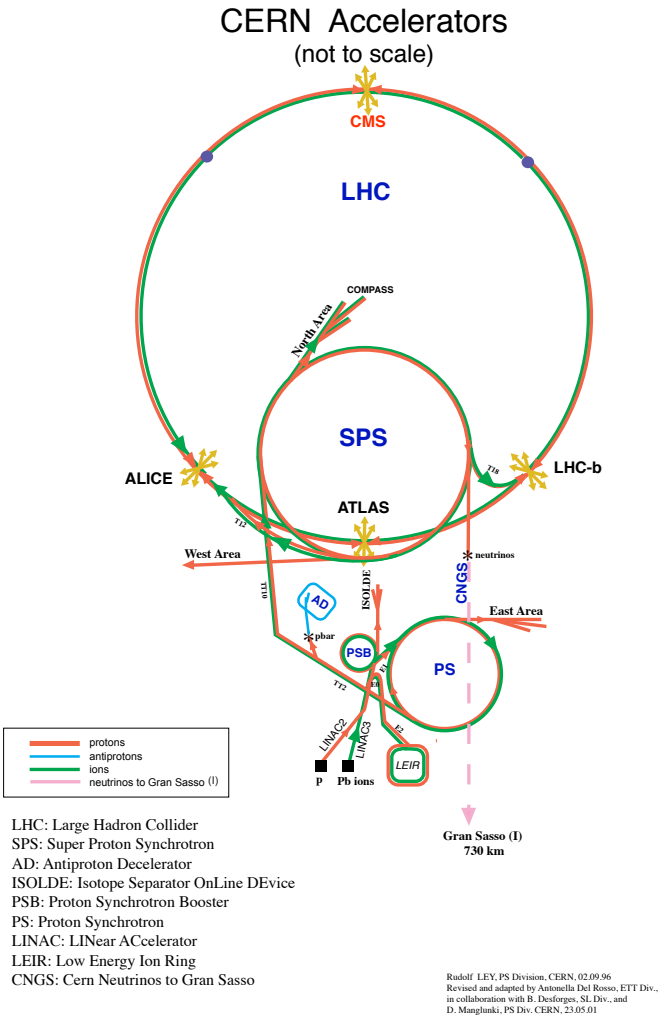


Figure 2.1: Global view of CERN accelerator complex.

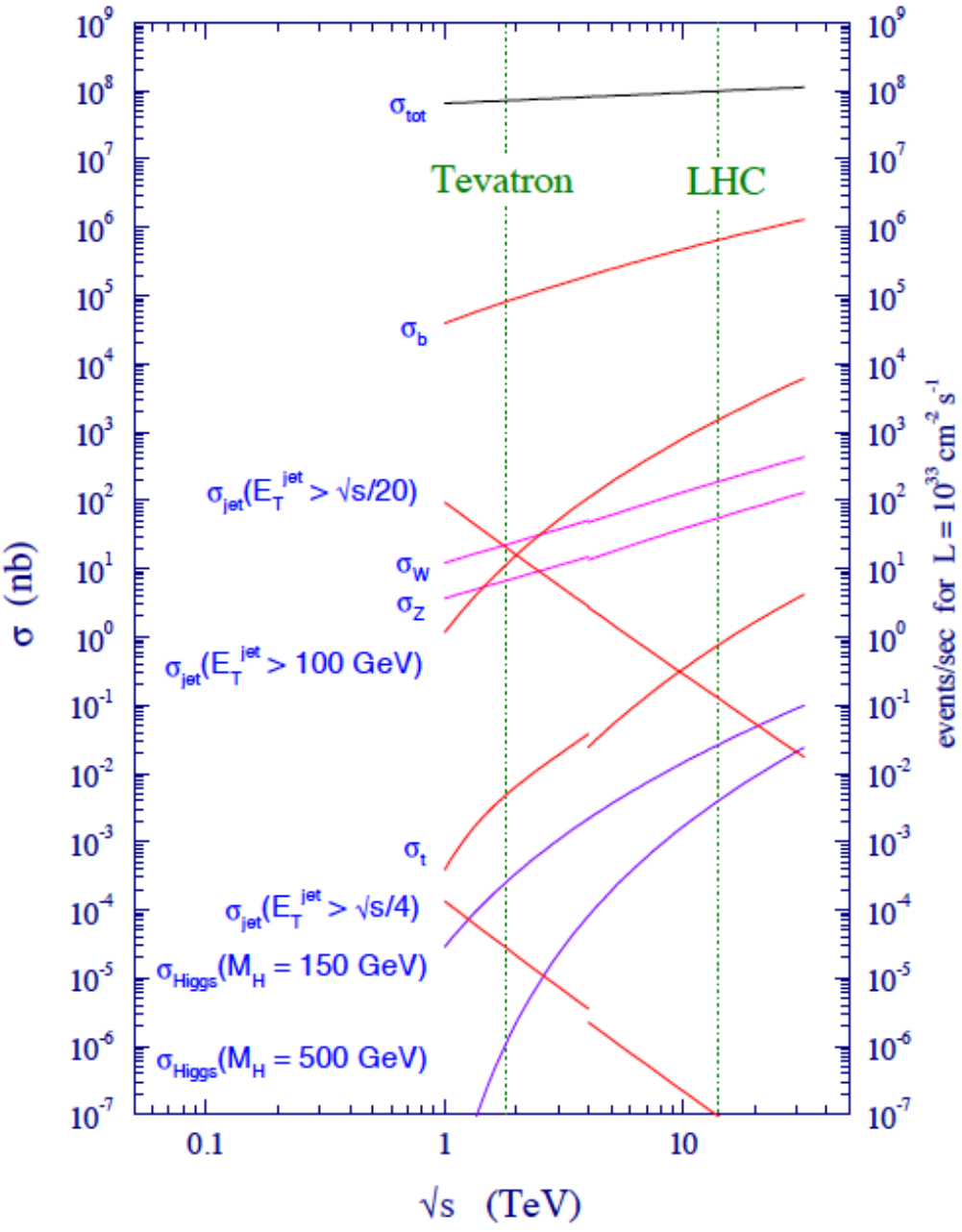


Figure 2.2: The cross section of the physics processes as a function of the centre of mass energy at hadron collider.

superconducting magnets are classical niobium-titanium magnets, cooled by supercritical helium at temperature of 4.2 K, which can reach a magnetic field of 5 T. The LHC magnet system also use niobium-titanium magnets, but cooled by superfluid helium at temperature below 2 K, which allows the magnetic field to be above 8 T.

There are six experiments installed at the LHC: ALICE [38], ATLAS [39], CMS [40], LHCb [41], LHCf [42] and TOTEM [43]. The first four detectors are the main detectors installed in four huge underground caverns built around the four collision points of the LHC beams.

- ATLAS and CMS are two general purpose detectors designed to cover the widest possible range of physics at the LHC, from the search for Higgs boson to supersymmetry (SUSY) and extra dimensions.
- ALICE is a detector specialized in analyzing lead-ion collisions. It will study the properties of quark-gluon plasma.
- LHCb is designed mainly for the study of B-physics. A specialized study of the slight asymmetry between matter and antimatter present in interactions of B-physics, may lead to the discovery of new physics.
- LHCf is a small experiment that is constructed to measure neutral particles produced very close to the direction of the beams at the LHC.
- TOTEM is also a small experiment dedicated to the measurement of the total proton-proton cross section with a luminosity-independent method. It is installed near the interaction point used by CMS.

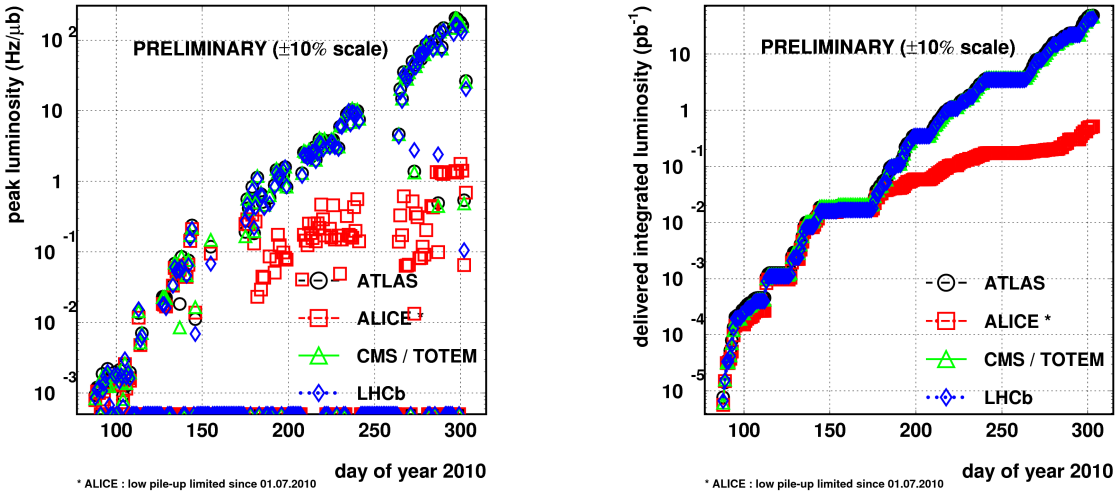
In December 2009, the first proton-proton collisions at the centre-of-mass energy of 900 GeV were delivered by the LHC. There were 5×10^{10} protons per bunch, two bunches per beam and the luminosity was around $10^{26}\ cm^{-2}s^{-1}$ to $10^{27}\ cm^{-2}s^{-1}$. About 1.5 million events in total were collected by all the experiments. At the end of March 2010, first collisions at the centre-of-mass energy of 7 TeV were recorded at the LHC. The number of protons per bunch then progressively increased from 5×10^{10} up to 1.2×10^{11}, which is surprisingly above the design parameter for beam-beam effects. Beam crossings were 150 ns apart. The proton-proton collisions continued until the end of October 2010, and were followed by a period of heavy ion Pb-Pb collisions for about four weeks afterwards. In summary, the LHC has run with stable beams for 1074 hours. The highest instantaneous luminosity reached in 2010 is $2 \times 10^{32} cm^{-2}s^{-1}$, as shown in Figure 2.3(a). The delivered integrated luminosity is about 48 pb^{-1} (Figure 2.3(b)) for proton-proton collisions, and $9.7\mu b^{-1}$ for Pb-Pb collisions.

2.2 The ATLAS Detector

The ATLAS (**A** **T**oroidal **L**HC **A**pparatu**S**), one of the two general-purpose detectors at the LHC, is designed to accommodate the wide spectrum of possible physics signatures.

2.2.1 Coordinate system and nomenclature

The coordinate system and nomenclature used to describe the ATLAS detector are briefly summarized here since they are used repeatedly throughout this thesis. The nominal interaction point is defined as the origin of the coordinate system, while the beam direction



(a) (b)

Figure 2.3: Peak (a) and integrated (b) luminosity for each LHC experiment in 2010 collisions.

defines the z-axis and the x-y plane is transverse to the beam direction. The positive x-axis is defined as pointing from the interaction point to the centre of the LHC ring and the positive y-axis is defined as pointing upwards. The side-A of the detector is defined as that with positive z and side-C is that with negative z. The azimuthal angle ϕ is measured around the beam axis, and the polar angle θ is the angle from the beam axis. The pseudorapidity is defined as:

$$\eta = -\frac{1}{2}\ln\tan(\theta/2)$$

In the case of massive objects such as jets, the rapidity

$$y = \frac{1}{2}\ln(\frac{E+p_z}{E-p_z})$$

is used instead, as the pseudorapidity can not approximate rapidity when the object mass is not negligible with respect to its momentum. The transverse momentum p_T, the transverse energy E_T, and the missing transverse energy $E_{T_{miss}}$ are defined in the $x-y$ plane. The distance ΔR in the pseudorapidity-azimuthal angle space is defined as:

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$$

2.2.2 Physics requirement and performance goals

The main goals of the ATLAS detector are to perform precision tests of QCD and electroweak interactions and search for Higgs boson and new physics. In particular, the search for the SM Higgs boson has been used as a benchmark to the performance of important sub-systems of ATLAS. The production and decay mechanisms of the Higgs boson depends on its mass. More information can be found in Section 6.1.1 and 6.1.2. At low masses ($m_H < 2m_Z$), the natural width is only a few MeV, thus the observed width is mostly determined by the instrumental resolution. The predominant decay mode into hadrons is

difficult to detect due to high QCD backgrounds: the two-photon final state is thus an important discovery channel. The reconstruction and identification of photons is critical issue in this channel. For masses above 130 GeV, the decay channels, $H \rightarrow ZZ^{\star}$ (with each Z decaying to a pair of oppositely charged leptons) and $H \rightarrow WW$ (with each W decaying into charged lepton and neutrino), are the experimentally cleanest channel to detect and study the properties of the Higgs boson. Therefore, identification of electrons and muons is very important. It requires good performance of the tracking system, calorimeter and muon spectrometer. For masses around and over 600 GeV, the dominate decay mode is the diboson final state. WW and ZZ decays into jets or involving neutrinos are needed to extract a signal. The tagging of forward jets from the vector boson fusion production is also important for searching for the Higgs boson.

The context of high luminosity, high beam energy and high background production at the LHC imposes extraordinary difficulties on the detector. The benchmark physics goals can be interpreted as a set of general requirements for ATLAS:

- The detectors require fast, radiation-hard electronics and sensor elements. In addition, high granularity is needed to handle the particle fluxes and to reduce the influence of overlapping events.

- Large acceptance in pseudorapidity with almost full azimuthal angle coverage is required.

- Good charged-particle momentum resolution and reconstruction efficiency in the inner tracker are essential. For offline tagging of τ-leptons and b-jets, vertex detectors close to the interaction region are required to observe secondary vertices.

- Very good electromagnetic (EM) calorimetry for electron and photon identification and measurements, complemented by full-coverage hadronic calorimetry for accurate jet and missing transverse energy measurements, are important requirements, as these measurements form the basis of many of the physics studies.

- Good muon identification and momentum resolution over a wide range of momenta and the ability to determine unambiguously the charge of high p_T muons are fundamental requirements.

- Highly efficient triggering on low transverse momentum objects with sufficient background rejection, is a prerequisite to achieve an acceptable trigger rate for most physics processes of interest.

The main performance goals for ATLAS detector are listed in Table 2.1.

2.2.3 Overview of the ATLAS detector

The ATLAS detector is mainly composed of six sub-systems: the magnet system, the inner detector, the calorimeter, the muon spectrometer, the trigger system and the data acquisition system. The overall ATLAS detector layout is shown in Figure 2.4. It is 25 m in height and 44 m in length. The overall weight of the detector is approximately 7000 tons. It is nominally forward-backward symmetric with respect to the interaction point.

The magnet configuration comprises a thin superconducting solenoid surrounding the inner-detector cavity, and three large superconducting toroids (one barrel and two end-caps)

Detector Component	Required Resolution	η coverage	
		measurement	Trigger
Tracking	$\sigma_{p_T}/p_T = 0.05\% p_T \oplus 1\%$	± 2.5	
EM calorimetry	$\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$	± 3.2	± 2.5
Hadronic calorimetry (jets)			
barrel and end-cap	$\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$	± 3.2	± 3.2
forward	$\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$	$3.1 < \lvert\eta\rvert < 4.9$	$3.1 < \lvert\eta\rvert < 4.9$
Muon spectrometer	$\sigma_{p_T}/p_T = 10\%$ at $p_T = 1$ TeV	± 2.7	± 2.4

Table 2.1: General performance goals of the ATLAS detector. Note that, for high-p_T muons, the muon-spectrometer performance is independent of the inner-detector system. The units for E and p_T are in GeV.

arranged with an eight-fold azimuthal symmetry around the calorimeters. This fundamental choice has driven the design of the rest of the detector.

The inner detector is immersed in the 2 T solenoidal field to measure the transverse momentum of the charged particles.

The high granularity liquid-argon (LAr) electromagnetic sampling calorimeters is installed on top of the solenoidal magnet system, and is used to measure the energy of electrons and photons. Outside the electromagnetic calorimeter, is the hadronic calorimeter, used to measure the energy of jets and hadrons.

The calorimeters are surrounded by the muon spectrometer. The air-core toroid system, with a long barrel and two inserted end-cap magnets, generates strong bending power in a large volume within a light and open structure. Excellent muon momentum resolution is achieved with three layers of high precision tracking chambers.

2.2.4 Magnet system

ATLAS features a unique hybrid system of four large superconducting magnets. This magnetic system is 22 m in diameter and 26 m in length, with a stored energy of 1.6 GJ.

The superconducting magnet system provides the magnetic field over a volume of approximately 12000 m^3 (defined as the region in which the field exceeds 50 mT). The spatial arrangement of the coil windings is shown in Figure 2.5.

The four large superconducting magnets are:

- **Central solenoid** [45]: It is placed outside the inner detector and before the electromagnetic calorimeter. The solenoid is aligned on the beam axis and provides a 2 T axial magnetic field for the inner detector. In order to achieve the desired calorimeter performance, the material thickness in front of the calorimeter needs to be minimized. The solenoid is designed to be as thin as possible. The layout was optimized so that the solenoid assembly contributes a total of ~ 0.66 radiation lengths at normal incidence. The solenoid and the electromagnetic calorimeter share one common vacuum vessel, therefore eliminating two vacuum walls.

- **Barrel toroid** [46]: It consists of eight coils assembled radially and symmetrically around the beam axis. The coils are encased in individual racetrack-shaped, stainless-steel vacuum vessels. The coils assembly are supported by eight inner and eight outer

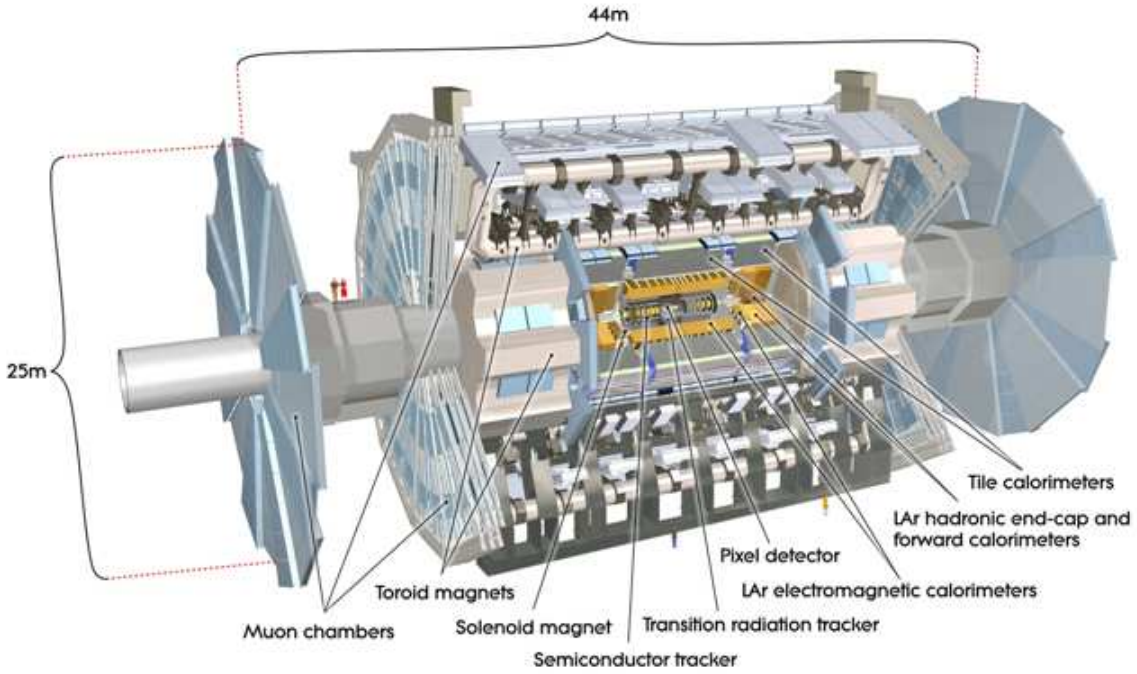


Figure 2.4: Global view of the ATLAS detector.

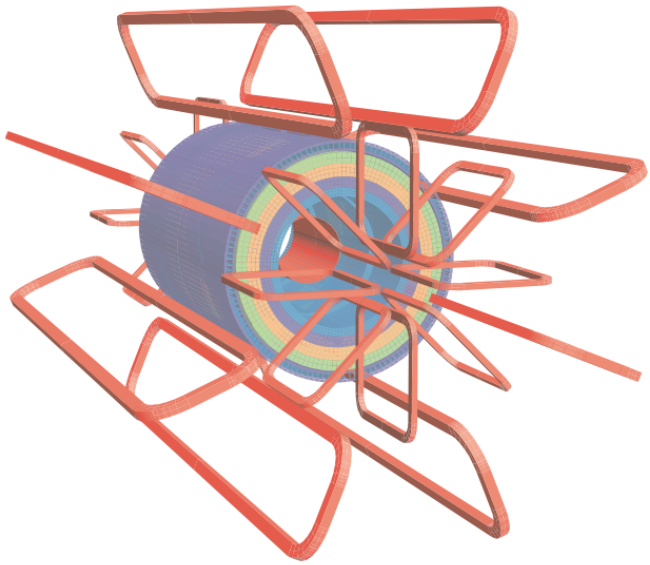


Figure 2.5: Geometry of magnet windings and tile calorimeter steel. The barrel toroid and two end-cap toroids are visible. The tile calorimeter is indicated in the figure by the multicolor structure. There are four layers with different magnetic properties together with an outside return yoke. The solenoid (the innermost red part) is located inside the calorimeter.

rings of struts. Services are brought to the coils through a cryogenic ring linking the eight cryostats to a separate service cryostat, which provides connections to the power supply, the helium refrigerator, the vacuum systems and the control system. The whole barrel toroid system is 25.3 m in length, with inner and outer diameters of 9.4 m and 20.1 m respectively.

- **End-cap toroid** [47]: There are two end-cap toroids inserted in the barrel toroid at each end and line up with the solenoid. Each end-cap toroid consists of eight racetrack, double-pancake coils in an aluminium alloy housing. The end-cap toroid coil system is rotated by 22.5° with respect to the barrel toroid coil system in order to provide radial overlap and to optimise the bending power at the interface between the two coil systems. The peak magnetic fields on the superconductors in end-cap toroid is 4.1 T.

2.2.5 Inner detector

The ATLAS Inner Detector [48] [49], abbreviated as ID, is placed right beyond the beam pipe. It is a tracking detector designed to provide excellent momentum resolution and both primary and secondary vertex measurements for charged tracks above a given p_T threshold. The lowest p_T is 0.5 GeV, but it had been lowered to 0.1 GeV in some early data analysis on minimum bias events. The pseudorapidity range is $|\eta| < 2.5$. It also provides information for charged particle identification, especially electrons and muons.

The ID is contained within a cylindrical envelope of length ±3512 mm and of radius 1150 mm, within the solenoidal magnetic field. A general view of the ID layout is given in Figure 2.6. It consists of three independent but complementary sub-detectors. Each sub-detector consists of a barrel part and two end-cap parts, one at each side of the barrel. Figures 2.7 and 2.8 show the sensors and structural elements traversed by 10 GeV tracks in the barrel and end-cap regions respectively.

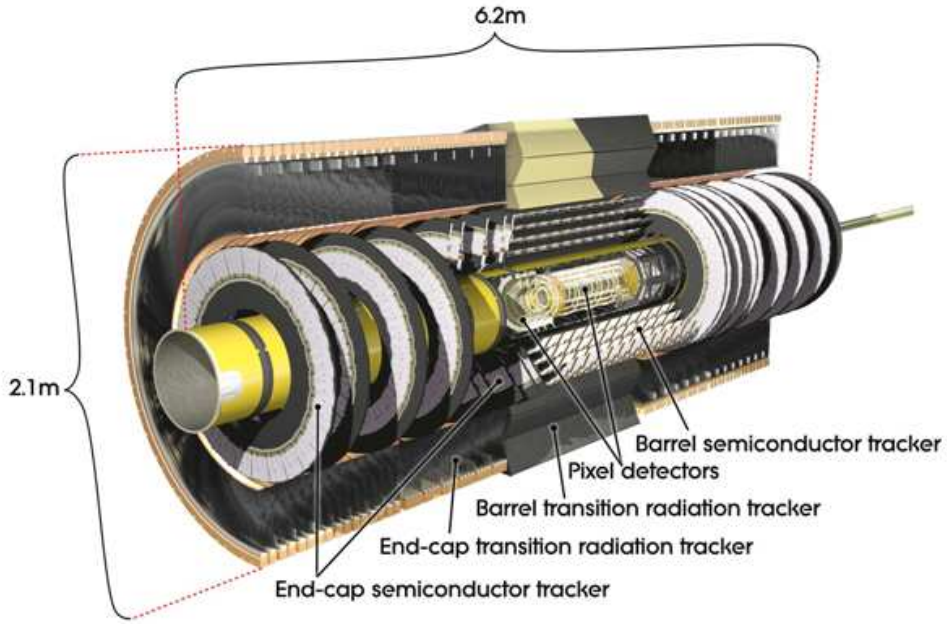


Figure 2.6: Over view of ATLAS inner detector.

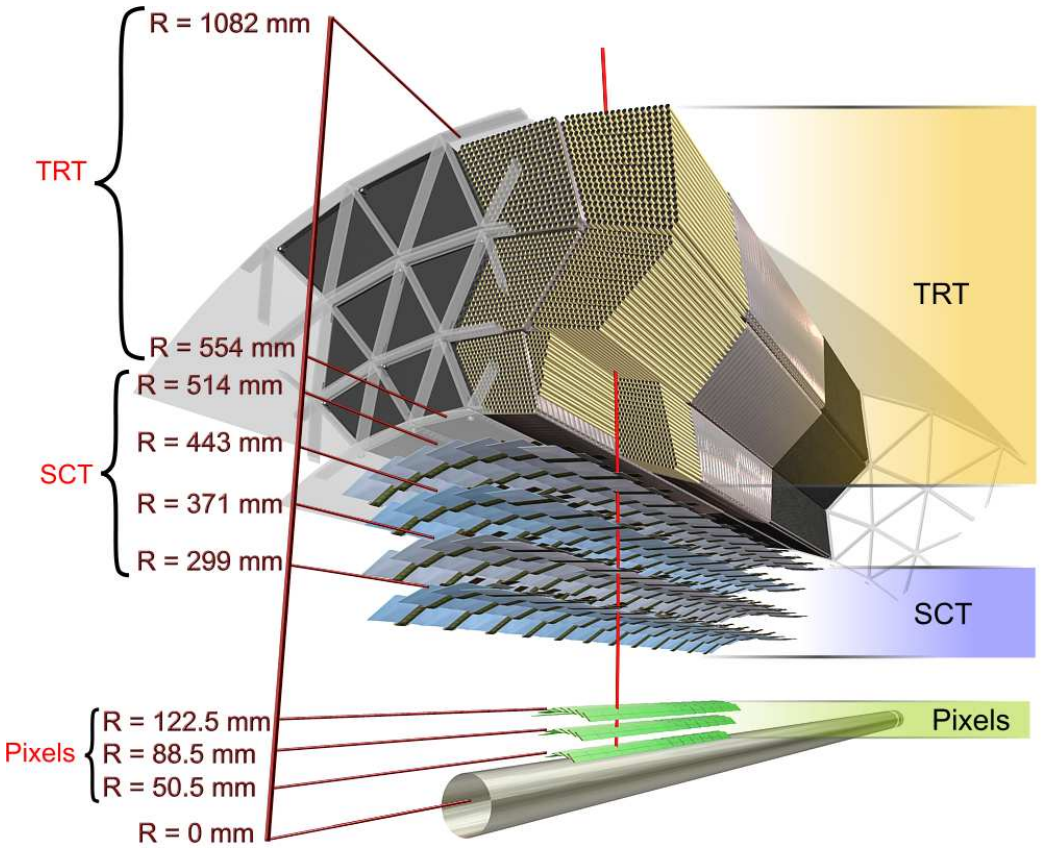


Figure 2.7: Drawing of the sensors and structural elements traversed by a charged track of 10 GeV p_T in the barrel inner detector ($\eta = 0.3$).

The dimensions of each sub-detector are listed in Table 2.2 and shown in Figure 2.9. At inner radii, high-resolution pattern recognition capabilities are available using discrete space-points from silicon pixel layers (Pixels) [50] and stereo pairs of silicon microstrip (SCT) layers. At larger radii, the transition radiation tracker (TRT) comprises many layers of gaseous straw tube elements interleaved with transition radiation material. With an average of 36 hits per track, it provides continuous tracking to enhance the pattern recognition and improve the momentum resolution over $|\eta| < 2.0$ and electron identification complementary to that of the calorimeter over a wide range of energies.

- **Pixels**: The pixel detector in ATLAS consists of three barrel layers (at average radii of ∼4 cm, 10 cm, and 13 cm) and two end-caps (each with three disk layers between radii of 11 and 20 cm). The layers are segmented in $R-\phi$ and z. All the layers contains a total of 140 million detector elements, each element with a size of $\Delta_{R-\phi} \times \Delta_z = 50 \times 400\ \mu m^2$. It corresponds to a resolution of 10 μm in $R-\phi$ plane and 115 μm in the z direction. The system is designed to be highly modular, containing approximately 1500 barrel modules and 700 disk modules. Each module is 62.4 mm long and 21.4 mm wide, serving 24 × 160 detector elements read out by 16 chips. The main quantitative goals were to reach a single hit resolution of 10 and 100 μm in transverse and longitudinal dimensions respectively, a transverse impact parameter resolution better than 15 μm, and a longitudinal primary vertex resolution better than 1 mm.

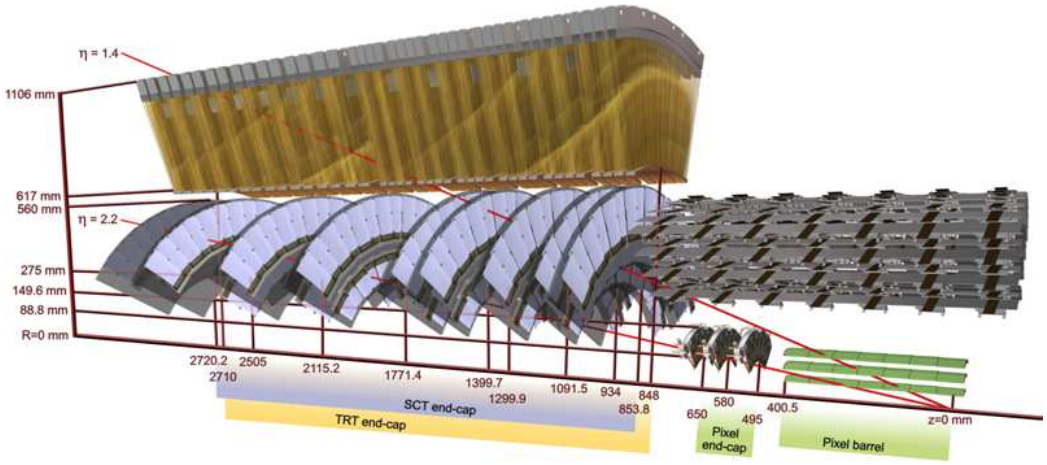


Figure 2.8: Drawing of the sensors and structural elements traversed by two charged tracks of 10 GeV p_T in the end-cap inner detector ($\eta = 1.4$ and 2.2).

Item		**Radial extension(mm)**	**Length(mm)**
Overall ID envelope		$0 < R < 1150$	$0 < \|z\| < 3512$
Beam-pipe		$29 < R < 36$	
Pixel	Overall envelope	$45.5 < R < 242$	$0 < \|z\| < 3092$
3 cylindrical layers	Sensitive barrel	$50.5 < R < 122.5$	$0 < \|z\| < 400.5$
2 x 3 disks	Sensitive end-cap	$88.8 < R < 149.6$	$495 < \|z\| < 650$
SCT	Overall envelope	$255 < R < 549(barrel)$	$0 < \|z\| < 805$
		$251 < R < 610(endcap)$	$810 < \|z\| < 2797$
4 cylindrical layers	Sensitive barrel	$299 < R < 514$	$0 < \|z\| < 749$
2 x 9 disks	Sensitive end-cap	$275 < R < 560$	$839 < \|z\| < 2735$
TRT	Overall envelope	$554 < R < 1082(barrel)$	$0 < \|z\| < 780$
		$617 < R < 1106(endcap)$	$827 < \|z\| < 2744$
73 straw planes	Sensitive barrel	$563 < R < 1066$	$0 < \|z\| < 712$
160 straw planes	Sensitive end-cap	$644 < R < 1004$	$848 < \|z\| < 2710$

Table 2.2: Main parameters of the ID

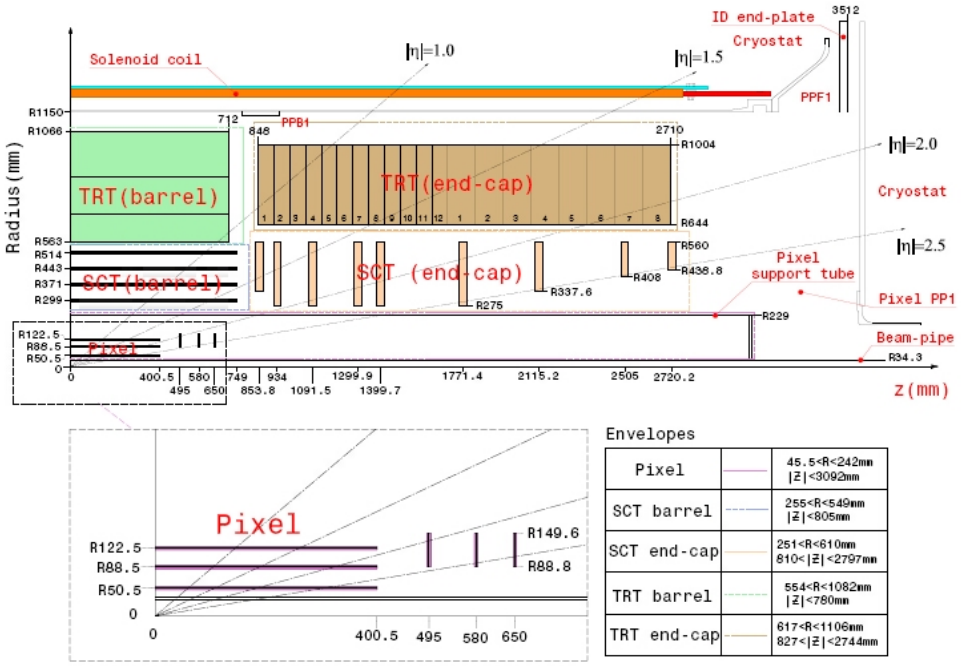


Figure 2.9: Plan view of a quarter-section of the ATLAS inner detector showing each of the major detector elements with its active dimensions and envelopes. The labels PP1, PPB1 and PPF1 indicate the patch-panels for the ID services.

- **SCT**: The SCT system is designed to provide eight precision measurements per track in the intermediate radial range, contributing to the measurement of momentum, impact parameter and vertex position, as well as providing good pattern recognition by the use of high granularity. The system is composed of four barrel layers [51] with radii of 30.0, 37.3, 44.7 and 52.0 cm, and two end-caps [52], each with nine wheels. The detector is also segmented into many elements in $R-\phi$ and z. Each element has a size of $6.36 \times 6.40\ cm^2$ with 768 readout strips of 80 μm pitch. The end-cap wheels use tapered strips. The strips are arranged radially, with $\sim$ 12 cm in length (at the outer radii) or $6-7$ cm in length (at the innermost radii). This size is optimized to cover the acceptance range $|\eta| < 2.5$. The SCT detector covers a total area of 63 m^2 and contains 6.2 million readout channels. The nominal spatial resolution is 16 μm in $R-\phi$ plane and 580 μm in the z direction. Tracks can be distinguished if separated by more than 200 μm.

- **TRT**: This system is based on the use of straw detectors, which can operate at the very high rates at the LHC by virtue of their small diameter and the isolation of the sense wires within individual gas volumes. The TRT contains up to 73 layers of straws interleaved with fibres (barrel) [53] and 160 straw planes interleaved with foils (end-cap) [54], which provide transition radiation for electron identification. All charged tracks with $p_T > 0.5$ GeV and $|\eta| < 2.0$ traverse at least 36 straws, except in the barrel-end-cap transition region ($0.8 < |\eta| < 1.0$), where this number decreases to a minimum of 22 crossed straws. Typically, seven to ten high-threshold hits from transition radiation are expected for electrons with energies above 2 GeV. The large number of straws per track guarantees a combined measurement accuracy of better

than 50 μm at the LHC design luminosity, averaged over all straws and including a systematic error of $\sim$ 30 μm from alignment. The TRT is able to discriminate electrons from other charged particles using transition radiation (TR). It consists of soft x-ray photons emitted by charged particles transversing the boundary between materials with different dielectric constants. The number of TR photons produced depends on the γ, where $\gamma = E/mc^2$, of the particle and the number of boundaries crossed by the particle.

2.2.6 Calorimeter

The ATLAS calorimeters [55] [56] comprise an electromagnetic calorimeter part and a hadronic calorimeter part. Each part consists of a number of sampling detectors with full ϕ-symmetry and coverage around the beam axis. The electromagnetic calorimeter is placed beyond the inner detector, and the hadronic calorimeter is outside the electromagnetic calorimeter. The overall system is shown in Figure 2.10, and its main parameters are given in Table 2.3. The calorimeters closest to the beam-line are housed in three cryostats, one barrel and two end-caps. The barrel cryostat contains the electromagnetic barrel calorimeter, whereas the two end-cap cryostats each contain an electromagnetic end-cap calorimeter (EMEC), a hadronic end-cap calorimeter (HEC), located behind the EMEC, and a forward calorimeter (FCal) to cover the region closest to the beam. All these calorimeters use liquid argon as the active detector medium; liquid argon has been chosen for its intrinsic linear behavior, its stability of response over time and its intrinsic radiation-hardness. The hadronic barrel calorimeter, normally called tile calorimeter, is placed outside the electromagnetic barrel calorimeter. It uses scintillator as the active medium.

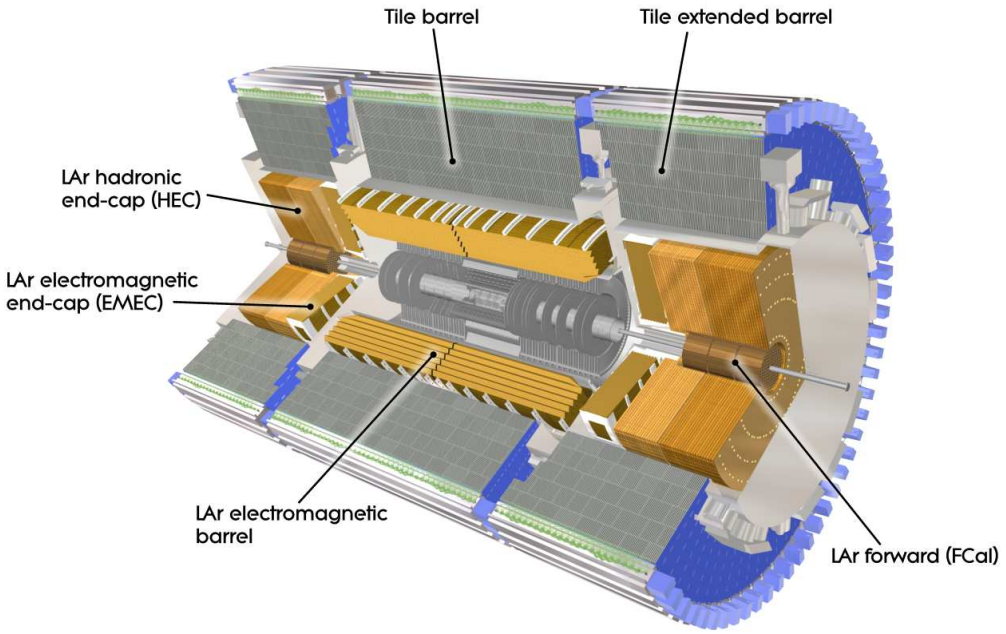


Figure 2.10: Over view of ATLAS calorimeter.

- **Electromagnetic calorimeter**

	Barrel		End-cap	
EM calorimeter				
Number of layers and $\|\eta\|$ coverage				
presampler	1	$\|\eta\| < 1.52$	1	$1.5 < \|\eta\| < 1.8$
Calorimeter	3	$\|\eta\| < 1.35$	2	$1.375 < \|\eta\| < 1.5$
	2	$1.35 < \|\eta\| < 1.475$	3	$1.5 < \|\eta\| < 2.5$
			2	$2.5 < \|\eta\| < 3.2$
Granularity $\Delta\eta \times \Delta\phi versus \|\eta\|$				
presampler	0.025×0.1	$\|\eta\| < 1.52$	0.025×0.1	$1.5 < \|\eta\| < 1.8$
Calorimeter 1st layer	$0.025/8 \times 0.1$	$\|\eta\| < 1.40$	0.050×0.1	$1.375 < \|\eta\| < 1.425$
	0.025×0.025	$1.40 < \|\eta\| < 1.475$	0.025×0.1	$1.425 < \|\eta\| < 1.5$
			$0.025/8 \times 0.1$	$1.5 < \|\eta\| < 1.8$
			$0.025/6 \times 0.1$	$1.8 < \|\eta\| < 2.0$
			$0.025/4 \times 0.1$	$2.0 < \|\eta\| < 2.4$
			0.025×0.1	$2.4 < \|\eta\| < 2.5$
			0.1×0.1	$2.5 < \|\eta\| < 3.2$
Calorimeter 2nd layer	0.025×0.025	$\|\eta\| < 1.40$	0.05×0.025	$1.375 < \|\eta\| < 1.425$
	0.075×0.025	$1.40 < \|\eta\| < 1.475$	0.025×0.025	$1.425 < \|\eta\| < 2.5$
			0.1×0.1	$2.5 < \|\eta\| < 3.2$
Calorimeter 3rd layer	0.05×0.025	$\|\eta\| < 1.35$	0.05×0.025	$1.5 < \|\eta\| < 2.5$
Number of readout channels				
Presampler	7808		1536(both sides)	
Calorimeter	101760		62208(both sides)	
LAr hadronic end-cap				
$\|\eta\| coverage$			$1.5 < \|\eta\| < 3.2$	
Number of layers			4	
Granularity $\Delta\eta \times \Delta\phi$			0.1×0.1	$1.5 < \|\eta\| < 2.5$
			0.2×0.2	$2.5 < \|\eta\| < 3.2$
Readout channels			5632(both sides)	
LAr forward calorimeter				
$\|\eta\| coverage$			$3.1 < \|\eta\| < 4.9$	
Number of layers			3	
Granularity $\Delta x \times \Delta y(cm)$			$FCal1 : 3.0 \times 2.6$	$3.15 < \|\eta\| < 4.30$
			$FCal1 :\sim 4x finer$	$3.10 < \|\eta\| < 3.15$
				$4.30 < \|\eta\| < 4.83$
			$FCal2 : 3.3 \times 4.2$	$3.24 < \|\eta\| < 4.50$
			$FCal2 :\sim 4x finer$	$3.20 < \|\eta\| < 3.24$
				$4.50 < \|\eta\| < 4.81$
			$FCal3 : 3.3 \times 4.2$	$3.32 < \|\eta\| < 4.60$
			$FCal3 :\sim 4x finer$	$3.29 < \|\eta\| < 3.32$
				$4.60 < \|\eta\| < 4.75$
Readout channels			3524(both sides)	
Scintillator tile calorimeter				
	Barrel		Extended barrel	
$\|\eta\| coverage$	$\|\eta\| < 1.0$		$0.8 < \|\eta\| < 1.7$	
Number of layers	3		3	
Granularity $\Delta\eta \times \Delta\phi$	0.1×0.1		0.1×0.1	
last layer	0.2×0.1		0.2×0.1	
Readout channels	5760		4092(both sides)	

Table 2.3: Main parameters of the calorimeter system.

As already mentioned in Section 2.2.2, the search for the SM Higgs boson has been used as a benchmark to the performance of important sub-systems of ATLAS. In particular, the electromagnetic calorimeter has been optimized for the decay channels of $H \to \gamma\gamma$ and $H \to ZZ^{\star}$. The most crucial discriminant for $H \to \gamma\gamma$ channel is the two photon invariant mass. It is computed as:

$$m_{\gamma\gamma} = \sqrt{2E_{T,1}E_{T,2}(\cosh(\eta_1 - \eta_2) - \cos(\phi_1 - \phi_2))} \tag{2.2}$$

The following requirements on the calorimeter play an important role in this channel, especially for the two photon invariant mass reconstruction:

- **Acceptance coverage**: The branching ratio of $H \to \gamma\gamma$ is quite small, at the level of $\mathcal{O}(10^{-3})$. Therefore the largest possible acceptance is needed to get the best sensitivity.
- **Energy resolution**: An excellent energy resolution and linearity is needed for the invariant mass reconstruction. A stochastic term of $10\%/\sqrt{E}$ and a constant term of 0.7% is expected to contribute to the energy resolution. A linearity response better than 0.5% (1%) up to 300 GeV (3 TeV) is required.
- **Position resolution**: It is related to the direction (η and ϕ) reconstruction of the photons. Accurate reconstruction of the direction is needed for the two photons in the invariant mass reconstruction as indicated in Equation 2.2.
- **Photon identification**: The QCD background (especially the di-jet background) is much larger than the signal production for $H \to \gamma\gamma$, thus excellent γ/jet separation is needed to reject the background. A refined granularity of the calorimeter is really important in the γ/jet separation.

The precision electromagnetic calorimeters are lead-liquid argon detectors with accordion shape absorbers and electrodes. The accordion geometry provides naturally a full coverage in ϕ without any cracks, and a fast extraction of the signal at the rear or at the front of the electrodes. In the barrel [57], the accordion waves are axial and run in ϕ, and the folding angles of the waves vary with radius to keep the liquid-argon gap constant. In the end-caps, the waves are parallel to the radial direction and run axially. Since the liquid-argon gap increases with radius in the end-caps [58], the wave amplitude and the folding angle of the absorbers and electrodes vary with radius.

The calorimeters have several active layers in depth, three in the precision-measurement region ($0 < |\eta| < 2.5$) and two in the higher-η region ($2.5< |\eta| < 3.2$) and in the overlap region between the barrel and the end-cap. In the precision measurement region, an accurate position measurement is obtained by finely segmenting the first layer in η. The η-direction of photons is determined by the position of the photon cluster in the first and the second layers.

The barrel electromagnetic calorimeter is made of two half-barrels, one covers the region with $z > 0$ ($0 < \eta < 1.475$) and the other covers the region with $z < 0$ ($-1.475 < \eta < 0$). Each half-barrel is 3.2 m in length, 1.4 m for inner radii and 2.0 m for outer radii, and 57 tons in weight. As mentioned before, there are three layers in the longitudinal direction, each layer with different granularity. For electrons and photons, most of their energy deposits in the middle layer. The granularity is $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ for each cell in the middle layer. The first layer (also

called strip layer) is finely segmented along η, as shown in Figure 2.11, so that there are eight strips in front of a middle cell. The back layer collects only the tail of the electromagnetic shower, thus it is less segmented in η. In the region $(0 < |\eta| < 1.8)$, there is a presampler placed before the first layer. It is used to measure the energy lost in front of the electromagnetic calorimeter.

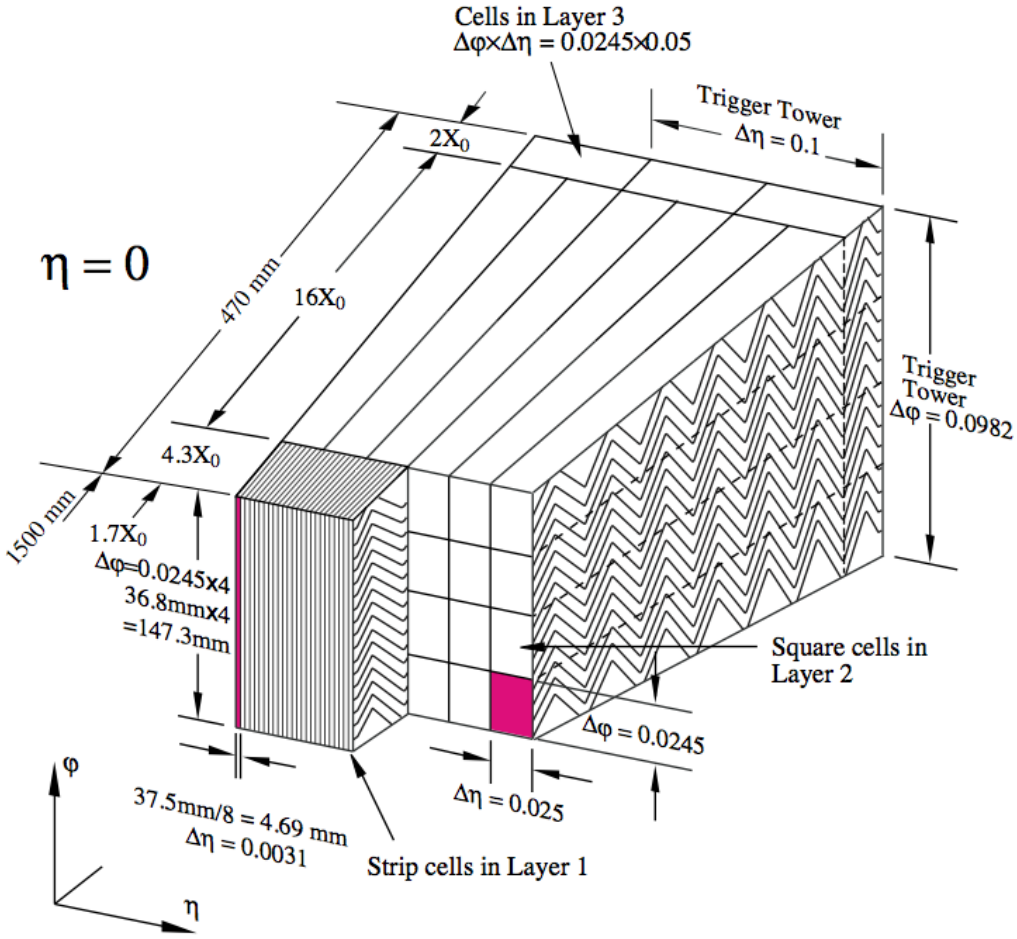


Figure 2.11: The granularity in η and ϕ of the cells of each of the three layers for electromagnetic calorimeter in the barrel. The granularity for trigger towers is also shown.

The end-cap electromagnetic calorimeter consists of two wheels, one on each side of the electromagnetic barrel. Each wheel is 63 cm in thickness, 27 tons in weight and with 330 mm for inner radii 2098 mm for outer radii. It covers the region $1.375 < |\eta| < 3.2$. There is relatively more material in the transition region between barrel and the end-cap calorimeters. The presampler is used in the region of $1.5 < |\eta| < 1.8$ to improve the energy measurement. The end-cap electromagnetic calorimeter in the region of $1.5 < |\eta| < 2.5$ is divided into three longitudinal layers as the barrel electromagnetic calorimeter. The granularity for the middle layer cells is the same as the one in the barrel: $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$. The first layer is finely segmented in η, while the back layer has a twice coarser granularity in η.

- **Tile calorimeter**

The tile calorimeter [59] is located behind the electromagnetic calorimeter. It uses steel as the absorber and scintillator as the active medium. It consists of a central barrel and two extended barrels. The central barrel is 5.8 m in length and 7.4λ (interaction lengths) in radial depth, covering the region $|\eta| < 1.0$. Each extended barrel is 2.6 m in length, 2.28 m for inner radii and 4.25 m for outer radii, covering the region of $0.8 < |\eta| < 1.7$. The choice of this technology provides maximum radial depth for the least cost for ATLAS. The tile calorimeter is divided into three longitudinal layers. The granularity is $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ for the cells in the first two layers and $\Delta\eta \times \Delta\phi = 0.1 \times 0.2$ for the cells in the third layer.

- **Hadronic endcap calorimeter**

  The hadronic calorimetry is extended to larger pseudorapidities by the hadronic end-cap calorimeter (HEC) [60], a copper/liquid-argon detector, which covers the region of $1.5 < |\eta| < 3.2$. The HEC consists of two wheels in each end-cap cryostat: a front wheel (HEC1) and a rear wheel (HEC2), each wheel containing two longitudinal sections. The wheels are cylindrical with an outer radii of 2030 mm. Each of the four HEC wheels is constructed of 32 identical wedge-shaped modules.

- **Forward calorimeter**

  The forward calorimeters (FCal) are located in the same cryostats as the end-cap calorimeters and cover the region of $3.1 < |\eta| < 4.9$. The close vicinity and coupling between these systems result in a quite hermetic design, which minimizes energy losses in cracks between the calorimeter systems and also limits the backgrounds which reach the muon system. As the FCal modules are located at high η, at a distance of approximately 4.7 m from the interaction point, they are exposed to high particle fluxes. Each FCal is split into three 45 cm deep modules: one electromagnetic module (FCal1) and two hadronic modules (FCal2 and FCal3), as illustrated in Figure 2.12.

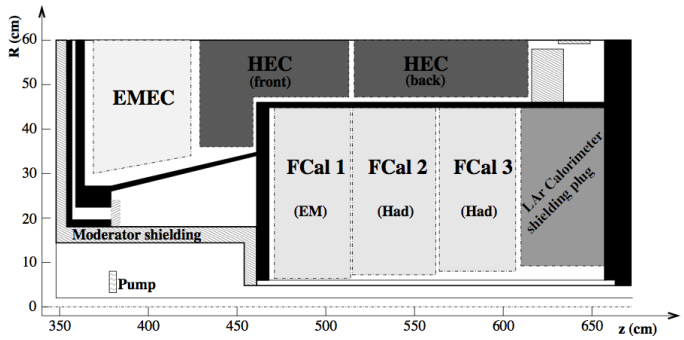


Figure 2.12: Schematic diagram showing the three FCal modules located in the end-cap cryostat. The material in front of the FCal and the shielding plug behind it are also shown. The black regions are structural parts of the cryostat. The diagram has a larger vertical scale for clarity.

- **Material before and in the calorimeters**

  Cumulative amount of material, in units of interaction length, as a function of $|\eta|$ is shown in Figure 2.13 in front of the electromagnetic calorimeters, in the electromagnetic calorimeters themselves, in each hadronic layer, and the total amount at the end

of the active calorimetry. Also shown for completeness is the total amount of material in front of the first active layer of the muon spectrometer (up to $|\eta| < 3.0$).

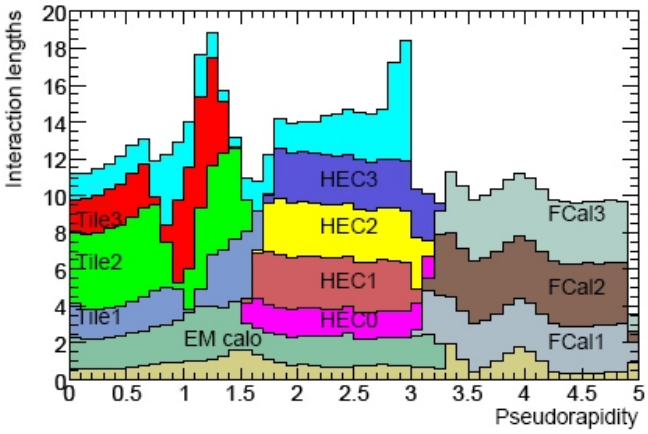


Figure 2.13: Over view of material distribution in front of and in the electromagnetic and hadronic calorimeters as a function of $|\eta|$.

2.2.7 Muon spectrometer

The muon spectrometer [61, 62] forms the outermost part of the ATLAS detector. It is designed to provide an accurate measurement of the muons momentum in a large dynamic range. It is also designed to trigger on muons in the region $|\eta|$ <2.4. The layout of the muon spectrometer is shown in Figure 2.14. A transverse momentum resolution of approximately 10% for 1 TeV tracks is expected, which translates into a sagitta along the z (beam) axis of about 500 μm, to be measured with a resolution of ≤ 50 μm. Muon momenta down to a few GeV (~ 3 GeV, due to energy loss in the calorimeters) may be measured by the spectrometer alone. Even at the high end of the accessible range (~ 3 TeV), the standalone measurements still provide adequate momentum resolution and excellent charge identification.

Precision-tracking chambers in the barrel region are located between and on the eight coils of the superconducting barrel toroid magnet, while the end-cap chambers are in front and behind the two end-cap toroid magnets. The ϕ symmetry of the toroids is reflected in the symmetric structure of the muon chamber system, consisting of eight octants. Each octant is subdivided in the azimuthal direction in two sectors with slightly different lateral extensions, a large and a small sector, leading to a region of overlap in ϕ. This overlap of the chamber boundaries minimises gaps in detector coverage and also allows for the relative alignment of adjacent sectors using tracks recorded by both a large and a small chamber.

The chambers in the barrel are arranged in three concentric cylindrical shells around the beam axis at radii of approximately 5 m, 7.5 m, and 10 m. In the two end-cap regions, muon chambers form large wheels, perpendicular to the z-axis and located at distances of $|z| \approx 7.4$ m, 10.8 m, 14 m, and 21.5 m from the interaction point.

The precision momentum measurement is performed by the Monitored Drift Tube chambers (MDTs). They cover the pseudorapidity range $|\eta| < 2.7$ (except in the innermost end-cap layer where their coverage is limited to $|\eta|$ <2.0). These chambers consist of three to eight layers of drift tubes, operated at an absolute pressure of 3 bar, which achieve an

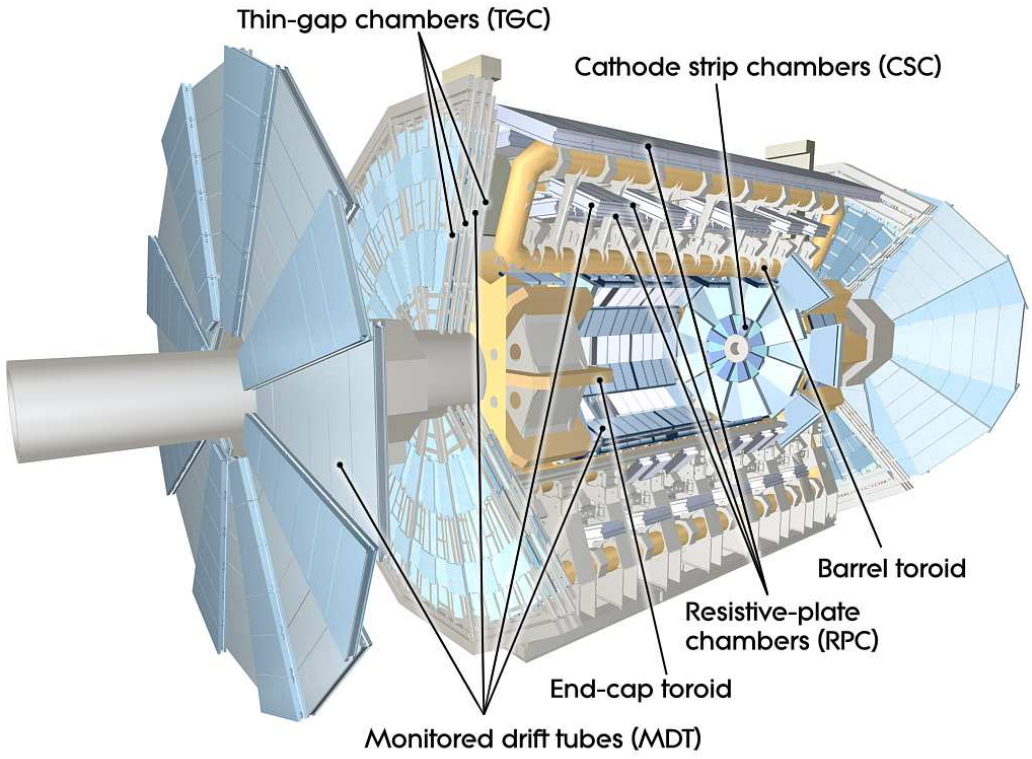


Figure 2.14: Over view of the muon system.

average resolution of 80 μm per tube, or about 35 μm per chamber.

In the forward region ($2 < |\eta| < 2.7$), Cathode-Strip Chambers (CSC) are used in the innermost tracking layer due to their higher rate capability and time resolution. The CSCs are multiwire proportional chambers with cathode planes segmented into strips in orthogonal directions. This allows both coordinates to be measured from the induced-charge distribution. The resolution of a chamber is 40 μm in the bending plane and about 5 mm in the transverse plane. The difference in resolution between the bending and non-bending planes is due to the different readout pitch, and to the fact that the azimuthal readout runs parallel to the anode wires.

An essential design criterion of the muon system was the capability to trigger on muon tracks. The precision-tracking chambers have therefore been complemented by a system of fast trigger chambers capable of delivering track information within a few tens of nanoseconds after the passage of the particle. In the barrel region ($|\eta| < 1.05$), Resistive Plate Chambers (RPC) were selected for this purpose, while in the end-cap ($1.05 < |\eta| < 2.4$) Thin Gap Chambers (TGC) were chosen.

2.2.8 Trigger and data acquisition system

In the LHC environment, a very powerful and efficient trigger system is indispensable for the experiments. The trigger system [63] in the ATLAS experiment includes three levels: Level 1, Level 2 and the event filter level. One can see the overview of the ATLAS trigger system in Figure 2.15. The L2 and event filter together form the High-Level Trigger (HLT) [65]. Each trigger level refines the decisions made at the previous level and, where necessary,

applies additional selection criteria.

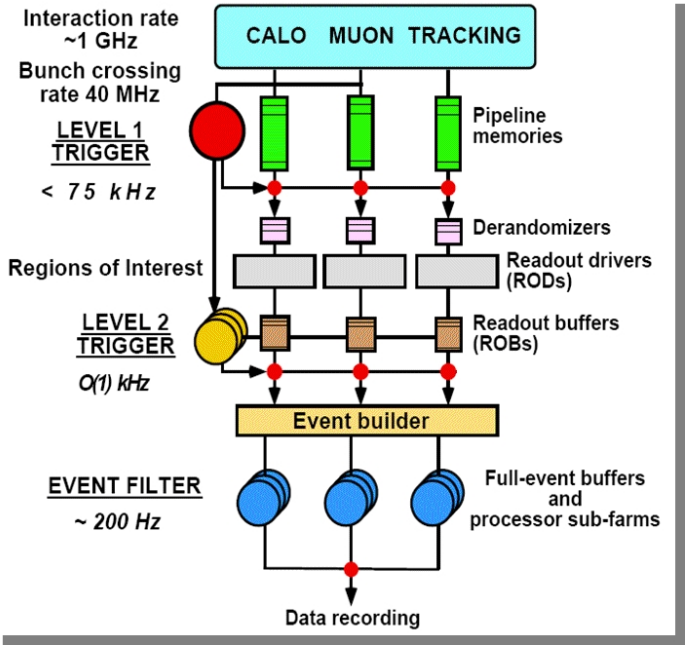


Figure 2.15: Over view of the trigger system.

The L1 trigger [64] searches for high transverse-momentum muons, electrons, photons, jets, and τ-leptons decaying into hadrons, as well as large missing and total transverse energy. High transverse-momentum muons are identified using trigger chambers in RPC and TGC. Calorimeter selections are based on reduced-granularity information from all the calorimeters. Results from the L1 muon and calorimeter triggers are processed by the central trigger processor, which implements a trigger "menu" made up of combinations of trigger selections. In each event, the L1 trigger defines one or more Regions-of-Interest (RoIs), in which interesting features have been identified and the criteria passed is associated. This information is subsequently used by the high-level trigger. After L1, the collision rate of 40 MHz will be reduced to 75 kHz (or up to 100 kHz).

The L2 selection is seeded by the RoI provided by the L1 trigger. The RoIs are the regions where the L1 trigger has identified possible trigger objects within the event. L2 selections use, at full granularity and precision, all the available detector data within the RoIs (approximately 2% of the total event data). The L2 menus are designed to reduce the trigger rate to approximately 3.5 kHz, with an average event processing time of approximately 40 ms.

The final stage of the event selection is carried out by the event filter, which uses offline analysis procedures on fully-built events to further select events. This level reduces the event rate to roughly 200 Hz. Its selections are implemented with an average event processing time of about four seconds.

The data acquisition system (DAQ) first receives and buffers the event data from the detector-specific readout electronics at the L1 trigger rate. Then it transmits to the L2 trigger any data (typically the data of RoI's) requested by the trigger. The events passing the L2 selection will be moved to the event filter, and perform the event-building. The events passing the event filter level selection, are finally moved to permanent event storage.

In addition to controlling movement of data down the trigger selection chain, the data acquisition system also provides for the configuration, control and monitoring of the ATLAS detector during data-taking.

2.3 ATLAS data taking status and performance

In December 2009, the ATLAS detector recorded data from a first series of LHC runs for proton-proton collisions at $\sqrt{s} = 0.9$ TeV and 2.36 TeV, with integrated luminosities of about 9 μb^{-1} and 0.7 μb^{-1} respectively. The ATLAS operating procedure in 2009 maintained the calorimeters and TRT in standard operating conditions, but the silicon trackers and muon chambers were at a reduced or "standby" voltage until after stable beams were declared by the LHC.

In 2010, the LHC delivered an integrated luminosity of about 48 pb^{-1} for proton-proton collisions at $\sqrt{s} = 7$ TeV, 94% of which were recorded by ATLAS with the full detector at nominal high voltage, as illustrated in Figure 2.16. As already mentioned in Section 2.1, the highest instantaneous luminosity reached in 2010 is $2 \times 10^{32} cm^{-2} s^{-1}$. More than 97% of the each subdetector was operational throughout 2010, and some of the non-operational channels were repaired during the winter shutdown. For each of the subsystems, more than 90% of the recorded data are considered of good quality for physics analysis.

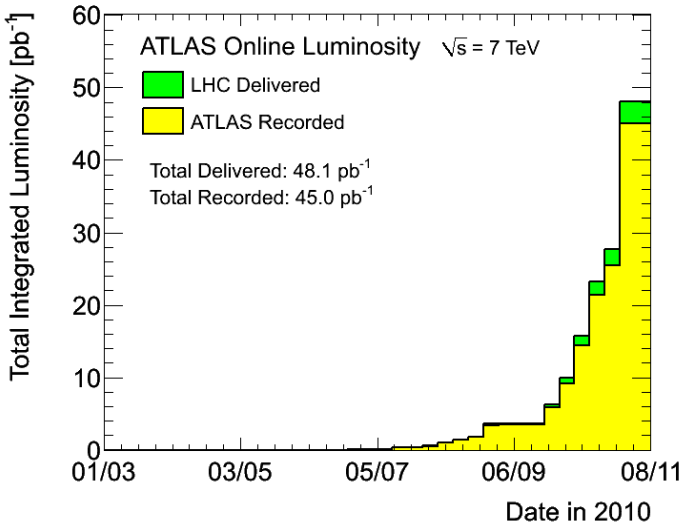


Figure 2.16: Integrated luminosity versus day delivered to (green) and recorded by ATLAS (yellow) in pp collisions at $\sqrt{s} = 7$ TeV.

The LHC integrated luminosity is measured in real time using different algorithms based on the counting rates of dedicated Cherenkov detectors (LUCID). In addition, other techniques with the ATLAS detector (timing distribution in the trigger scintillators, timing in the forward calorimeters and particle vertices reconstructed online with the tracking detectors) are also used to measure the luminosity. The different methods agree within about 5%. The absolute luminosity is determined, using beam separation scans for calibration, with an uncertainty of 11% dominated by the systematic uncertainty on the LHC beam current measurement.

2.3.1 Performance of the Inner Detector

- **Pixels**: The fraction of operating modules amounts to 97.4% and is found to be stable along the data taking period. Non-operating modules are due to various service disconnections (high and low voltage) or non-configurable modules. The hit efficiency for a track in the Pixel detector is about 99%. Figure 2.17 shows the primary vertex position in the transverse (x, y) plane for an LHC fill. The measured dispersions in both directions (108 and 96 μm respectively) agree with the 100 μm predicted from the beam instrumentation. As will be discussed in Chapter 6, reconstructed individual primary vertices help to improve the resolution on physical quantities, which use the angles between particles, like invariant masses. An alignment is performed and the spatial resolution is investigated by looking at unbiased residuals, which are the differences between the measured position of the cluster and the predicted position from the trajectory in the tracker. In the barrel region, unbiased residual widths of 19 and 105 μm have been obtained for transverse and longitudinal dimensions respectively, which are in agreement with the desired resolution.

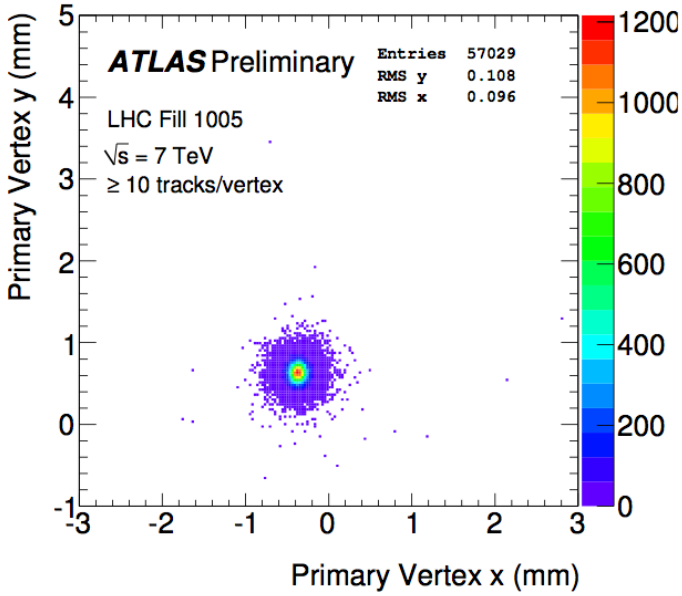
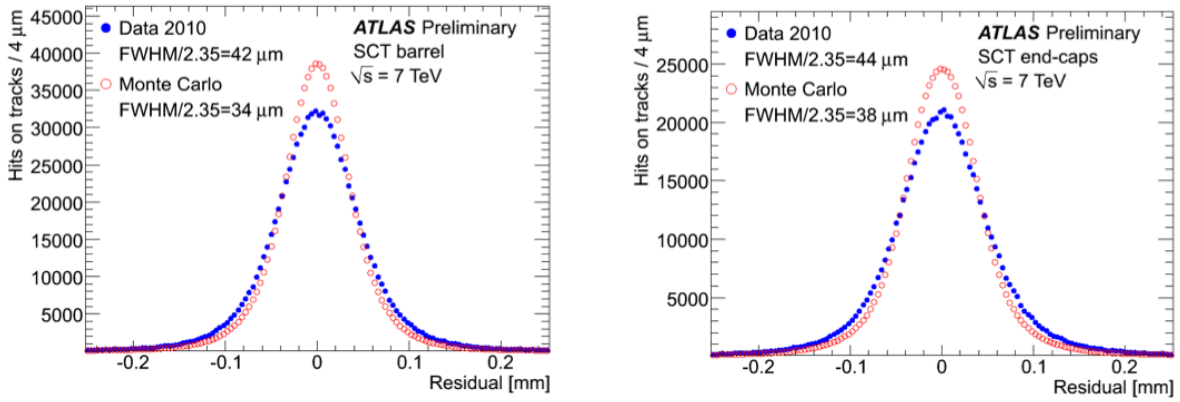


Figure 2.17: Primary vertices position in the transverse plane for an LHC fill.

- **SCT**: There are 30 modules (out of 4088) excluded of the configuration of the SCT in the 2010 data taking. 99.3% of the SCT detector is fully functional. The SCT has on average an intrinsic hit efficiency of 99.8%±0.1%. The alignment for 2010 is performed using the data-sets collected in 2009 collision at $\sqrt{s}$ = 0.9 TeV, The alignment is already very close to the ideal one as demonstrated in Figure 2.18, where the x track-residual distributions from track extrapolation are shown for 7 TeV collision data compared to MC simulation with nominal alignment. The plots show evidence that the SCT is well aligned and its geometry approximates closely the perfect, MC-simulated geometry. This is really important for performing accurate tracking and invariant mass reconstructions.

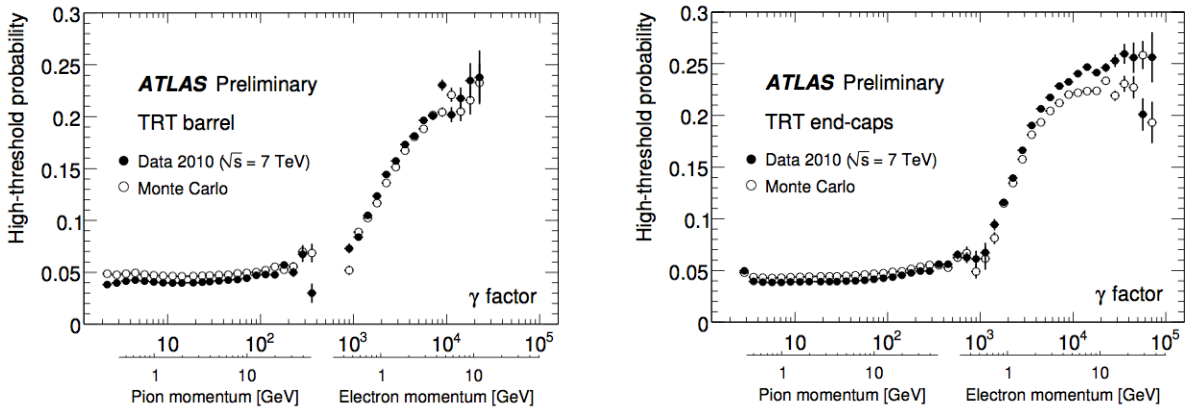
- **TRT**: The hit efficiency for TRT is about 94% if the track is within 1 mm from the



(a) (b)

Figure 2.18: Track x-residual distribution for SCT with perfect MC-simulated geometry (red circle) and for the current alignment with 7 TeV collision data (blue dot). (a) barrel (b) end-cap. [66].

straw centre. Figure 2.19 shows the probability of a TRT high-threshold (HT) hit as a function of $\gamma = E/mc^2$. The results from the barrel are in agreement with those from MC simulation, whereas in the endcap the simulated sample shows slightly higher HT probability for hadrons and lower HT probability for electron candidates, which indicates that improvements are needed for the simulation of the TR in the endcap. The spatial residual in TRT is approaching the design resolution of 130 μm.



(a) (b)

Figure 2.19: The probability of a TRT high-threshold (HT) hit as a function of the Lorentz factor, $\gamma = E/m$, for TRT. (a) barrel (b) end-cap. [67]

2.3.2 Performance of the calorimeter

During the data-taking between April and October 2010, some optical links in the calorimeter readout system were non-functional. The inoperable optical links increased over time.

As of November 2010, 2% of the readout channels were inoperable. During the shutdown from the end of 2010 and the beginning of 2011, those channels were repaired. In the last few weeks of 2010 data-taking, which corresponds to 55% of the total integrated luminosity recorded by ATLAS, about 30 out of 1524 optical links were affected. The (η, ϕ) maps of the dead readout optical links for the presampler, strip and middle layers in EM calorimeter in this period are shown in Figure 2.20. Those maps are called Object Quality maps used to constraint the quality of the photon and electron reconstruction. If cells of the photon or electron object are inside the inoperable region, the object is rejected. The distribution of the energy deposition in the calorimeter cells is well described by simulated minimum bias events for both electromagnetic and hadronic calorimeters in all η regions. The resolution of the energy measurement is controlled using randomly triggered events. An example of the energy distribution in the Tile calorimeter cells is shown in Figure 2.21. The energy associated to a track was determined using topology clusters with barycenter at $\Delta R < 0.2$ from the track. Shower contamination from photons and neutral hadrons was subtracted using late showering charged hadrons as control sample. Data and MC agree within 5%. The calibration of the LAr calorimeters shows good stability. The calibration constants are monitored over long time-periods. The stability of the pedestal is below 0.03 ACD counts, which is approximately 2 MeV, and the stability of the gain below 0.1% over all calorimeters. With 2010 collision data, an alignment of a few hundred picoseconds per front-end board (128 readout channels) was achieved, the goal being < 100 ps per readout channel for all LAr calorimeters.

2.3.3 Performance of the muon spectrometer

The p_T resolution in the muon spectrometer was measured using cosmic ray data taken in 2009 by splitting tracks crossing the centre of the detector into two "collision-like" tracks and comparing their measured p_T's. An example of the resolution is shown in Figure 2.22 obtained from one reconstruction algorithm in the sectors of the MDT located on the coils of the toroid. The data points are fitted to an empirical parameterization including energy loss fluctuation, multiple scattering and spectrometer resolution. The results obtained reach the designed goals.

2.3.4 Trigger and event selection

During the ATLAS startup phase, the focus of the trigger selection strategy was to commission the trigger and to ensure that well known Standard Model processes are observed. With peak luminosity less than few times of 10^{27} $cm^{-2}s^{-1}$, the minimum bias L1 trigger operated using hits in scintillator counters (MBTS), and HLT operated in pass-through mode. At peak luminosity around 10^{29} $cm^{-2}s^{-1}$, HLT chains have been activated to cope with increasing rate while running with low L1 thresholds. The trigger rates were kept about constant while the luminosity quickly increased by prescaling low threshold trigger items. An example of the L1 calorimetric trigger efficiency is shown in Figure 2.23. Good agreement between the collision data and MC-simulated sample is observed.

2.3.5 Material mapping in the ID

A very precise mapping of the material in the ID was obtained from the rate and vertex position of secondary hadronic interactions [70] or photon conversions [71] and from the energy flow [72] in the calorimeter. The three methods give results compatible with each

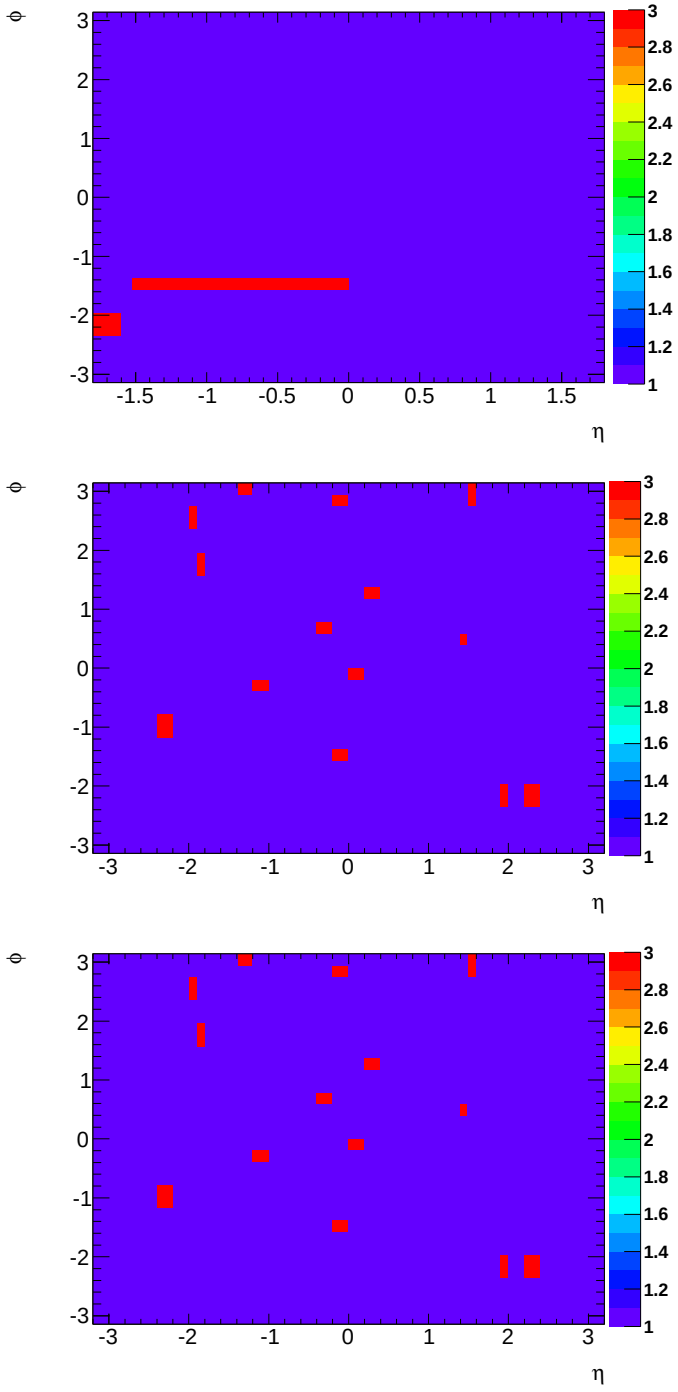


Figure 2.20: Maps of the dead readout optical links of the Liquid Argon Calorimeter. Top: presampler. Middle: strip layer. Bottom: middle layer. The red regions are not functional.

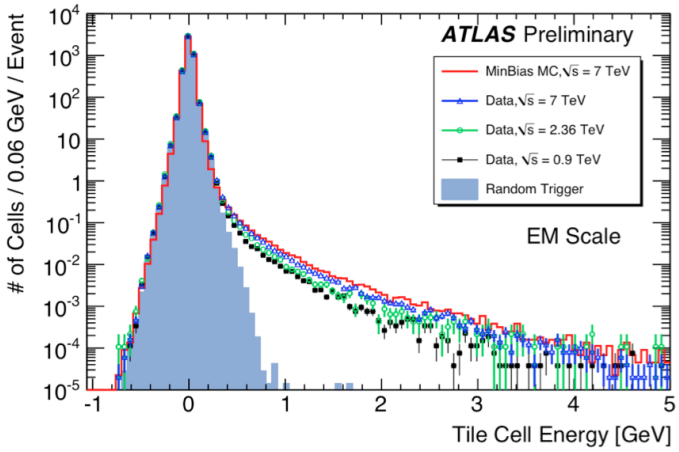


Figure 2.21: Energy deposition in the Tile calorimeter cells for collision data at different centre-of-mass energies, randomly triggered events and MC simulated minimum bias events. [68].

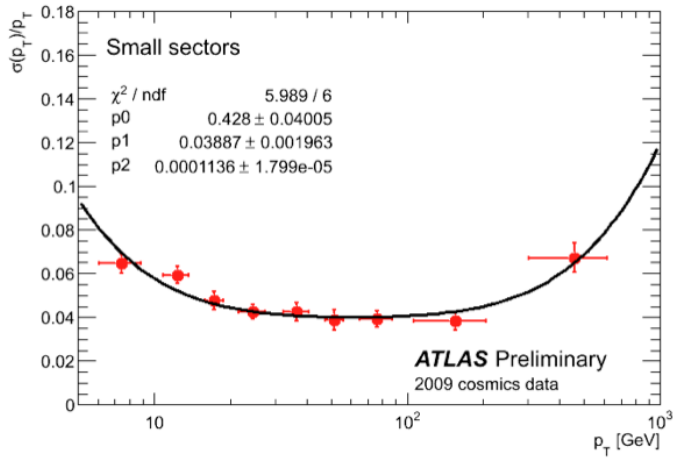


Figure 2.22: Transverse momentum resolution versus transverse momentum obtained from cosmic ray data (red dot), fitted with an empirical function (black line). [68].

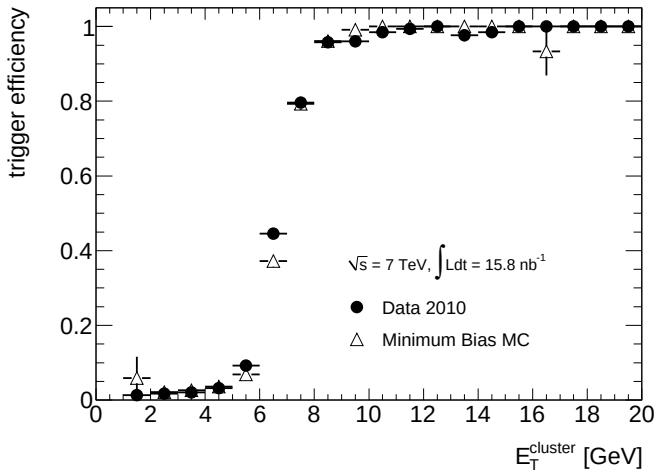
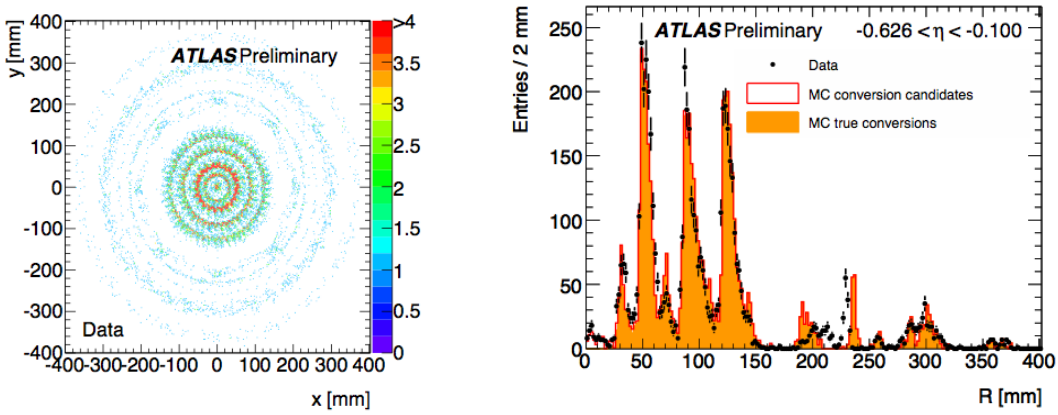


Figure 2.23: The efficiency for a L1 calorimetric trigger item as a function of the transverse energy of the reconstructed photons. The transverse energy threshold of this trigger item is 5 GeV. Result for data is obtained with collision data at $\sqrt{s} = 7$ TeV. [69].

other and show some discrepancy from the MC description of the detector. The maps obtained from data are used to improve the simulation.

The distribution of photon conversion vertices can be used to map the distribution of material in the ID. After applying some geometric selection criteria and the requirement on the fit quality of the conversion vertex, the position of selected photon conversion vertices are displayed in Figure 2.24. In this Figure, the beam pipe (R = 34.3 mm), the three barrel Pixel layers (R = 50.5, 88.5, 122.5 mm) and the first two SCT barrel layers (R = 299, 371 mm), together with the Pixel Support Tube (R = 229 mm) and various other support structures are clearly seen. In the xy projection, the cooling pipes on the Pixel detector modules and the overlap regions in the first SCT layer are visible. As shown in Figure 2.24 (b) clear shift in the simulated radial positions is observed for the Pixel Support Tube and global Pixel Support structure (around R = 200 mm), while the overall amount of material seems to be in good agreement.



(a) (b)

Figure 2.24: Distribution of reconstructed photon conversion vertices in the xy projection within $|\eta| < 1$ (a), and the radial distribution of conversion vertices in $-0.626 < \eta < -0.1$ (b). [73].

Chapter 3

Photon reconstruction, calibration and identification

3.1 Photon reconstruction

In the ATLAS experiment, rectangular clusters are first reconstructed from electromagnetic calorimeter (ECAL). The calorimeter is divided into a grid of 200×256 CaloTowers in the $\eta - \phi$ space within the range of $|\eta| < 2.5$. All layers of the ECAL are considered. Each CaloTower has a size of $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$. A rectangular precluster is first formed if a local maximum in a window of 5×5 CaloTowers above a fixed threshold of 2.5 GeV is found. Cells are then grouped into a rectangular window of $N_{\eta}^{cluster} \times N_{\phi}^{cluster}$ centered on a layer-dependent seed position, following the sequence of middle layer, strip layer, presampler and the back layer. In the middle layer, the precluster barycenter position is used as the seed position. The size of $N_{\eta}^{cluster} \times N_{\phi}^{cluster}$ depends on the particle type. More detailed information about the calorimeter clustering algorithm can be found in [86]. Clusters with a small energy leakage[1] (less than 2%) into the hadronic calorimeter (HCAL) are considered as electrons and photons.

Due to the presence of material upstream of the ECAL, photons can undergo a conversion. To classify a cluster as an electron, a converted photon or an unconverted photon, a two-step algorithm is designed. The first step is the track matching. If there is no track matching the cluster, it is characterized as unconverted photon. If there is at least one track matching the cluster, then all the reconstructed conversion vertex candidates[2] found by the Inner Detector (ID) reconstruction algorithm are checked.

The best matched track to the electromagnetic cluster is normally characterized as the one with an impact point[3] closest to the cluster itself. This track is checked against the track(s) originating from the best conversion vertex candidate matched to the same cluster.

[1] It is defined as the ratio of the transverse energy measured in the hadronic calorimeter over the transverse energy measured in ECAL. The transverse energy is calculated by: $E_T = E/\cosh(\eta)$, where E is the measured energy, η is computed from the position of the barycenter of the cluster.

[2] The conversion vertex candidates are reconstructed by a constrained vertex fit of two electron-like tracks [75]. There are also conversion vertex candidates that with only one single electron-like track associated, which are called single-track conversion vertex candidates. The associated single electron-like track must not be originating from the primary vertex, and the conversion vertex candidate is assigned at the first hit of the track.

[3] The impact point is the hit point in the calorimeter extrapolated from the trajectory of a track in the Inner Detector.

If the track coincides with a track coming from the conversion vertex, then this cluster is treated as a converted photon. If the coinciding track has a hit in the innermost Pixel layer (B-layer), and the other track of a double-track conversion vertex candidate lacks a B-layer hit, this cluster will not be considered as a converted photon. If the track does not coincide with any of the tracks assigned to the conversion vertex candidate, then it is considered as an electron, unless the track transverse momentum p_T is smaller than that of the candidate converted photon p_T.

If the cluster has a best matched track that is with only Transition Radiation Tracker hits (TRT-only track) and with $p_T > 2$ GeV and $E/p < 10$ (here, E is the cluster energy, p is the track momentum), it is then considered as a converted photon, regardless of whether a conversion vertex candidate has been matched to its electromagnetic cluster or not. Otherwise if TRT-only tracks have $p_T < 2$ GeV, then the cluster is automatically considered to be unconverted photon candidates.

Finally, the four momentum of the converted photon candidates is computed using in addition the available tracking information. Although the energy itself comes from the electromagnetic cluster, the η and ϕ directions are taken from the corresponding track parameters at the reconstructed conversion vertex. In the case of TRT-only tracks the η direction is picked up from the cluster η-pointing, which is determined by the barycenters of the middle layer and the strip layer.

3.2 Photon calibration

The number of ECAL cells used for the photon energy measurement depends on the photon classification.[4] In the barrel, a rectangular window of 3×5 second layer cell size around the cluster barycenter is used for unconverted photon and 3×7 used for converted photon. In the endcap, a rectangular window of 5×5 is used for all the photons. Two algorithms are designed to determine the particle energy from the energies measured in the presampler and the three calorimeter layers: the "calibration hits" (CH) and the "longitudinal weights" (LW) method [76]. The calibration coefficients are extracted from Monte Carlo simulations. They are updated each time when a more refined simulation of the calorimeter is available. In the past, the two algorithms have been validated on test beam data. Both methods are available in the ATLAS reconstruction program and the CH method is the default one. The CH method is more sophisticated. It takes full consideration of the η, energy and radiation length dependence for the various contributions which enter in the computation of the photon energy. The LW method derives the coefficients with a global χ^2 optimization over a wide energy range and is not strictly related to the different contributions. In this section, only the CH method is detailed.

This method determines the energy of photons by summing up four different contributions: the energy deposited in front of the calorimeter (including the energy between the cold calorimeter wall and the strip layer), the energy deposited inside the cluster, outside the cluster (lateral leakage) and the energy deposited beyond the calorimeter (longitudinal leakage). The four terms are parameterized as a function of the measured energies in the presampler (where present) and in the three longitudinal layers. The coefficients are also computed as a function of the pseudorapidity value which is the center of a middle cell and

[4] Although photon calibration is described in this section, the same procedure is applied to the calibration of the clusters classified as electrons, leading to the extraction of specific calibration coefficients.

stored in a database. The energy can be expressed with the following formula:

$$\begin{aligned} E_\gamma &= a(E_{tot}^{Acc},\eta) + b(E_{tot}^{Acc},\eta)\cdot E_{ps}^{clLAr} + c(E_{tot}^{Acc},\eta)\cdot (E_{ps}^{clLAr})^2 \\ &+ \frac{s_{cl}^{Acc}(X,\eta)}{f_{out}(X,\eta)}\cdot (\sum_{i=1,3} E_i^{clLAr})\cdot (1+f_{leak}(X,\eta))\cdot (F(\eta,\phi)) \end{aligned} \tag{3.1}$$

Where:

- E_γ is the photon energy.
- $a(E_{tot}^{Acc},\eta)$, $b(E_{tot}^{Acc},\eta)$ and $c(E_{tot}^{Acc},\eta)$ are parameters determined as a function of the energy deposited in the accordion (E_{tot}^{Acc}) and $|\eta|$. The coefficients a and b are commonly named *offset* and *slope*. For pseudorapidity values corresponding to the barrel calorimeter ($|\eta| \leq 1.475$) the parameterization is limited to the first two terms (c=0).
- E_{ps}^{clLAr} is the energy measured in the presampler cluster corrected for the fraction deposited in the passive materials.
- X is the longitudinal barycenter of the shower (Shower Depth) defined as:

$$X = \frac{\sum_{i=0,3} E_i^{clLAr} X_i}{\sum_{i=0,3} E_i^{clLAr}} \tag{3.2}$$

  Here, E_i^{clLAr} are the energies deposited in the cluster in the active medium of the presampler and the three compartments of the calorimeter (strip, middle, back), and X_i is the depth, expressed in radiation length, of the longitudinal center of each compartment computed from the centre of ATLAS. X_i are functions of $|\eta|$.
- $s_{cl}^{Acc}(X,\eta)$ is a correction factor of the accordion sampling fraction in the cluster. It is parameterized as a function of X and $|\eta|$.
- $f_{out}(X,\eta)$ is the correction for the energy deposited in the calorimeter outside the cluster (lateral leakage). It is parameterized as a function of X and $|\eta|$.
- $f_{leak}(X,\eta)$ is the longitudinal leakage correction. It is parameterized as a function of $|\eta|$ and X.
- $F(\eta,\phi)$ is the energy correction depending on the impact point inside a cell (energy modulation).

In the region $|\eta| \geq 1.8$, not instrumented with the presampler, the energy deposited in front of the calorimeter is parameterized as a function of the shower longitudinal barycenter as shown in Equation 3.2 computed with the information given by the three calorimeter layers only.

The performance studies of single calorimeter modules using the test beam data and fully simulated Monte Carlo events indicate that a photon energy resolution can be described by the following formula:

$$\frac{\sigma}{E} = C \oplus \frac{s}{\sqrt{E}} \tag{3.3}$$

where E is the photon energy and σ is the photon energy resolution. s is a sampling term resulting from the sampling fluctuations. C is an overall constant term which consists

of local constant term, short-range constant term and long-range constant term. Local constant term is mainly due to the imperfect corrections for the lateral and longitudinal leakage and for the ϕ-modulation. The short-range constant term includes the contributions from mechanics, such as absorber and liquid gap tolerances, and from calibration, such as amplitude accuracy, readout stability and difference between calibration and physics signals. The long-range constant term include the effects from time dependence of charge measured in liquid, temperature gradients, high voltage variations and materials. The designed goal for the overall constant term in ATLAS is at the level of 0.7% or better over the full range of ECAL devoted to precision physics.

3.3 Photon identification

The largest background contamination for the true photons that comes from the hard scattering or parton fragmentation are jets, especially the jets that fragment into a leading π^0 or η mesons. The photons from π^0 or η decays are typically less isolated due to the hadronic activity around, and give larger shower in the ECAL due to the presence of two photons.

In order to separate true photons from fake photons resulting from jets, several discriminating shower shape variables are defined using mostly the information of electromagnetic calorimeter.

- Hadronic leakage:
  - R_{had_1}, R_{had}: ratio of E_T in the first layer of the hadronic calorimeter to the E_T of the EM cluster. In the pseudorapidity range $0.8 < |\eta| < 1.37$ which is not covered by the first hadronic layer, the total hadronic E_T to the EM E_T ratio is used.

$$R_{\mathrm{had_1}} = \frac{E_T^{\mathrm{had_1}}}{E_T} \tag{3.4}$$

$$R_{\mathrm{had}} = \frac{E_T^{\mathrm{had}}}{E_T} \tag{3.5}$$

- Middle layer of ECAL:
  - R_η: ratio in η of cell energies in 3×7 versus 7×7 cells,

$$R_\eta = \frac{E_{3\times7}^{S2}}{E_{7\times7}^{S2}} \tag{3.6}$$

  - R_ϕ: ratio in ϕ of cell energies in 3×3 versus 3×7 cells,

$$R_\phi = \frac{E_{3\times3}^{S2}}{E_{3\times7}^{S2}} \tag{3.7}$$

  - ω_2: lateral width of the shower,

$$\omega_2 = \sqrt{\frac{\sum E_i \eta_i^2}{\sum E_i} - \left(\frac{\sum E_i \eta_i}{\sum E_i}\right)^2} \tag{3.8}$$

- Strip layer of ECAL:
  - $\omega_{s\,3}$: shower width for three strip cells around the strip cell with the highest energy deposited,
$$\omega_{s\,3} = \sqrt{\frac{\sum E_i (i - i_{\max})^2}{\sum E_i}} \tag{3.9}$$
  - $\omega_{s\,\mathrm{tot}}$: total lateral shower width,
$$\omega_{s\,\mathrm{tot}} = \sqrt{\frac{\sum E_i (i - i_{\max})^2}{\sum E_i}} \tag{3.10}$$
  - F_{side}: ratio of energy outside a core of 3 central strip cells but within 7 strip cells, over the energy in the core of 3 central strip cells,
$$F_{\mathrm{side}} = \frac{E(\pm 3) - E(\pm 1)}{E(\pm 1)} \tag{3.11}$$
  - ΔE: difference between the energy of the strip cell with the second highest energy deposited and the energy of the strip cell with the smallest energy deposited between the two leading strip cells,
$$\Delta E = E^{S1}_{2^{\mathrm{nd}}\max} - E^{S1}_{\mathrm{min}} \tag{3.12}$$
  - E_{ratio}: ratio of the energy difference associated with the highest and second highest energy deposited over the sum of these energies.
$$E_{\mathrm{ratio}} = \frac{E^{S1}_{1^{\mathrm{st}}\,\max} - E^{S1}_{2^{\mathrm{nd}}\,\max}}{E^{S1}_{1^{\mathrm{st}}\,\max} + E^{S1}_{2^{\mathrm{nd}}\,\max}} \tag{3.13}$$
  - $R_{max\,1}$: energy in the strip cell with the second highest energy deposited, rescaled by a monotonic function of the total transverse energy E_T. This variable was not used anymore starting from the year of 2009 in ATLAS. Because the variable of E_{ratio} is correlated with $R_{max\,1}$, and shows larger rejection power.
$$R_{\mathrm{max\,1}} = \frac{E^{S1}_{2^{\mathrm{nd}}\,\max}}{1000 + 0.009\,E_T} \tag{3.14}$$

The cut on the hadronic leakage can greatly reduce the rate of fake photons coming from jets, which give relatively large energy deposition in the HCAL. The cuts on the variables based on the middle layer of the ECAL can help to reject the fake photons with wider showers. The variables defined from the strip layer of ECAL, make use of the refined segmentation of this layer and help to further reduce the fake photons originating from single π^0 or η mesons which decay in $\gamma\gamma$.

A comparison on the rejection power between converted and unconverted photon is performed for all the discriminating variables except the variable $R_{max\,1}$. This study is based on the Monte Carlo simulation samples listed in Table 3.1. A filter is applied at particle level to the generated sample to minimize the simulation and reconstruction time. The filters used in the production of those samples are:

Dataset	Process	σ [pb]	Filter	Filter Eff. [%]	Events	Luminosity $[pb^{-1}]$
106384	$H \to \gamma\gamma$	3.527E-2	SymPhotonFilter	78	99873	3.6303E+6
105802	JF17 (mostly jj)	1.461E+9	JetFilter	7.06	9849209	0.095

Table 3.1: MC samples used for the study on the shower shape variables.

- SymPhotonFilter: at least 2 photons with $p_T > 20$ GeV, $|\eta| < 2.7$
- JetFilter: at least 1 jet with $p_T > 17$ GeV, $|\eta| < 2.7$

In Figure 3.1, 3.2 and 3.3, the left column is the discriminating variable distribution for true and fake unconverted photons, the middle column is for converted photons and the right column is the background selection efficiency at the optimized cut as a function of the target signal efficiency. Here, true photons are the photons from the Higgs decay, and fake photons are the other photon candidates. Six of the nine discriminating variables show higher rejection power for the converted photons, two of them show higher rejection power for the unconverted photons, while the other one is similar for both unconverted and converted photons. The main difference between converted and unconverted signal photons lies in the shower width along the ϕ direction. Consequently, the shower shape variables that are not sensitive to the shower along the ϕ direction, show similar distribution for converted and unconverted signal photons. It is indeed the case for the variables R_{had_1}, R_η and $\omega_{s\,\mathrm{tot}}$. On the contrary, for the background fake photons, the dominant contribution comes from the parton fragmenting into a leading π^0 or η. Most of π^0 or η decay into two photons immediately. If the decay is with small angle and at least one photon is converted, this fake photon is mostly reconstructed as a converted photon and the shower is wide due to the several electrons or photons involved, thus making the discrimination from signal photon easier. On the other hand, if the tracks associated to the reconstructed converted background photon are not electron tracks (which means not from the photon conversion), it is a sign for a jet with more track-related activity around. Both cases will make the distribution of R_{had_1}, R_η and $\omega_{s\,\mathrm{tot}}$ for the converted candidates to be more background-like. Thus, converted photons are easily discriminated with those variables. The first case also makes the variable of E_{ratio} and ΔE more discriminating for converted photons than unconverted photons. Finally for the variables of R_ϕ and $\omega_{s\,3}$, which are sensitive to the photon conversion, the unconverted photons show better discrimination.

Figure 3.4 shows an example of the behavior of the mean of each calorimetric discriminating variables as a function of the pseudorapidity $|\eta|$, for both true and fake photons with $E_T > 20$ GeV before any selection, using also the Monte Carlo simulation samples listed in Table 3.1.

The clear dependence on pseudorapidity $|\eta|$ reflects the distribution of the material upstream before the calorimeter. Let us take the variable F_{side} as an example. Since there is more material in the region $0.8 < |\eta| < 1.6$, the photon shower starts earlier than in other regions. The shower is then slightly wider, thus F_{side} is larger. One can also see that the profiles for converted photon and unconverted photon are different due to the fact that converted photons with electron pairs are bent in the solenoid magnetic field and give wider showers especially in the ϕ direction. Consequently, the cut on each variable must be optimized in seven $|\eta|$ bins, and separately for converted and unconverted photons.

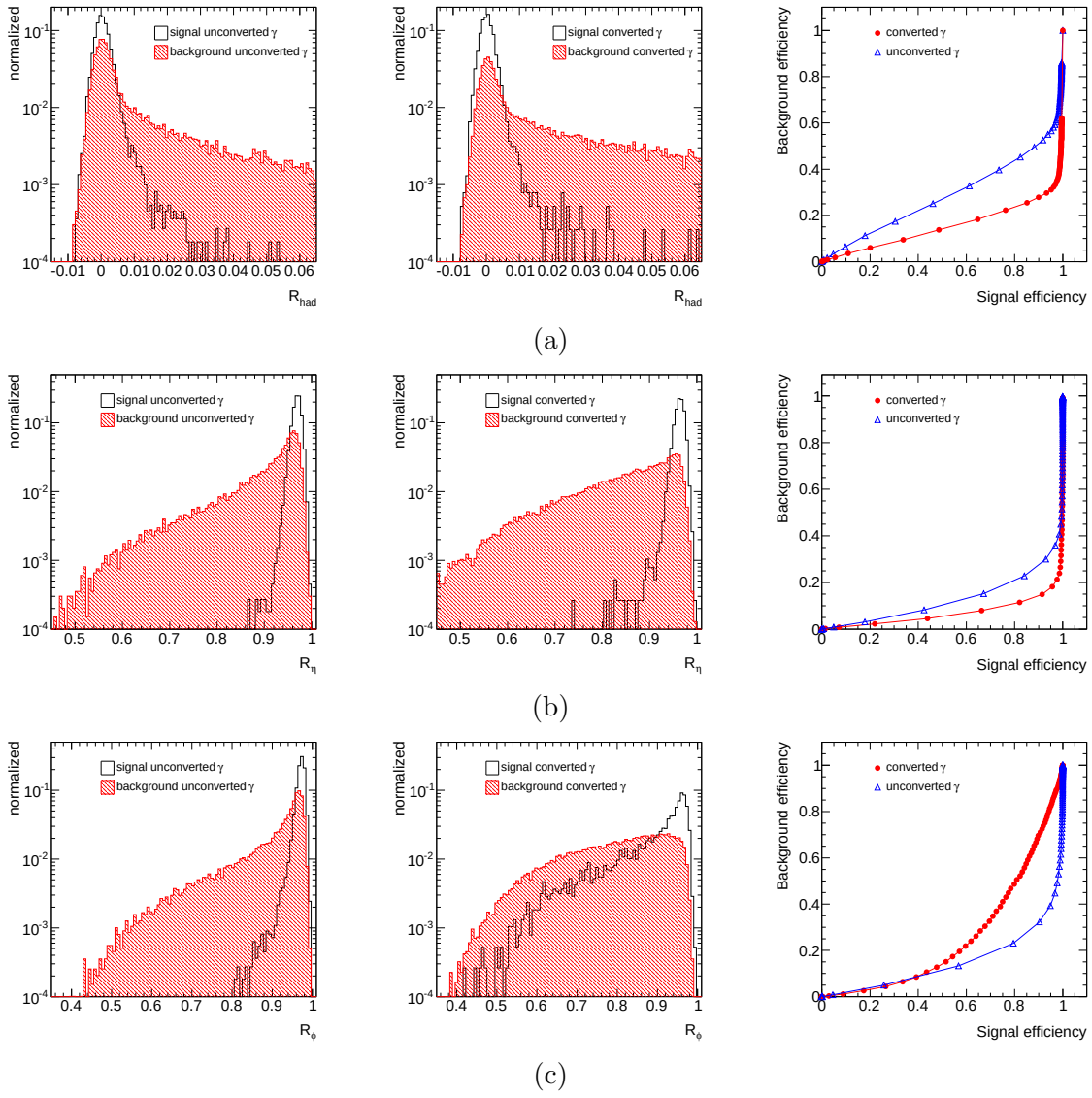


Figure 3.1: Probability distribution function of (a) R_{had} (b) R_η (c) R_ϕ for true and fake photons in the region $|\eta| < 0.6$ and with 25 GeV $< p_T <$ 40 GeV before any selection separately plotted for converted (left column) and unconverted (middle column) photons. Background selection efficiency comparison for converted and unconverted photons at the optimized cut as a function of the target signal efficiency is shown in the right column.

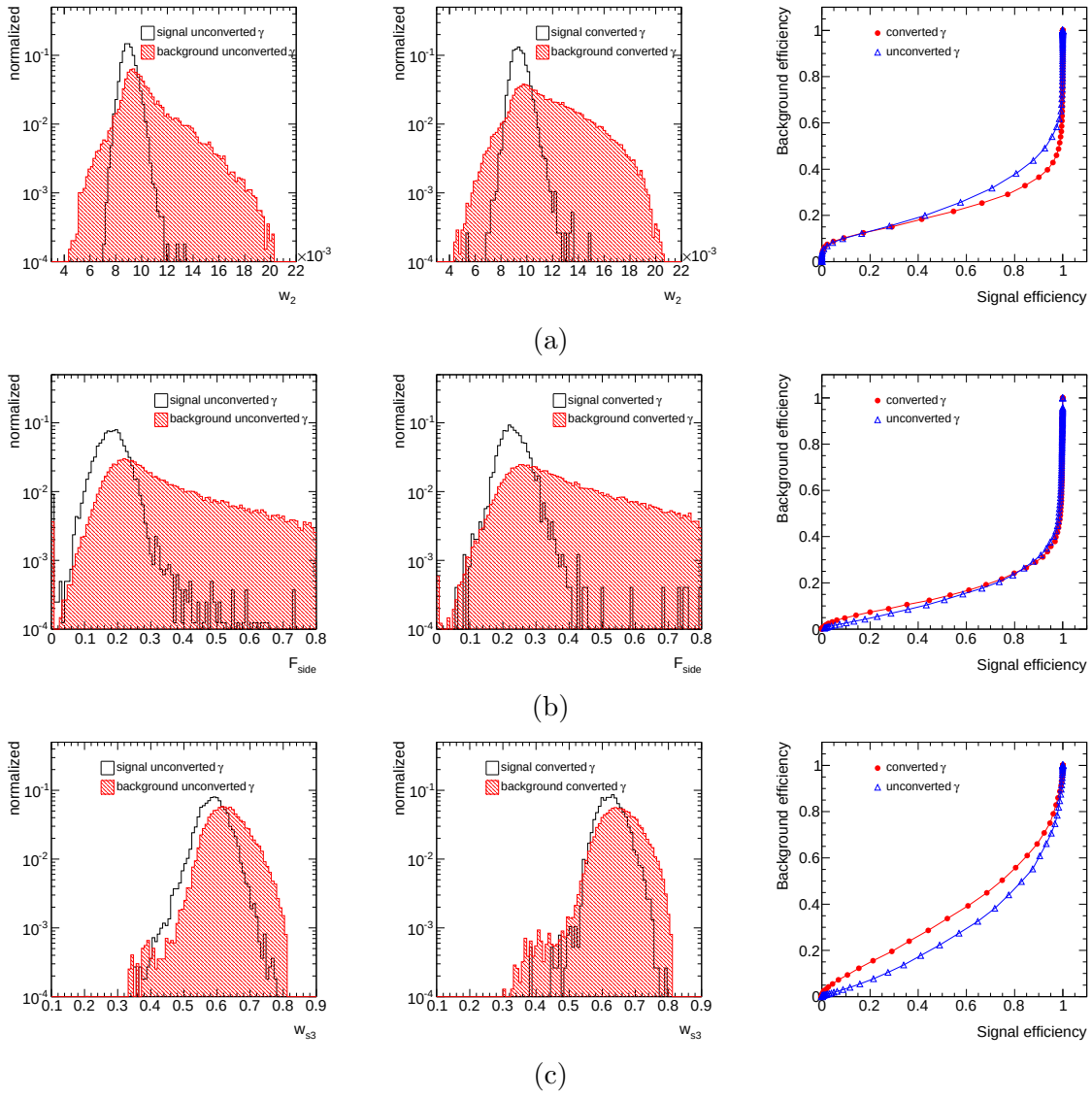


Figure 3.2: Probability distribution function of (a) w_2 (b) F_{side} (c) $w_{s,3}$ for true and fake photons in the region $|\eta| < 0.6$ and with 25 GeV $< p_T <$ 40 GeV before any selection separately plotted for converted (left column) and unconverted (middle column) photons. Background selection efficiency comparison for converted and unconverted photons at the optimized cut as a function of the target signal efficiency is shown in the right column.

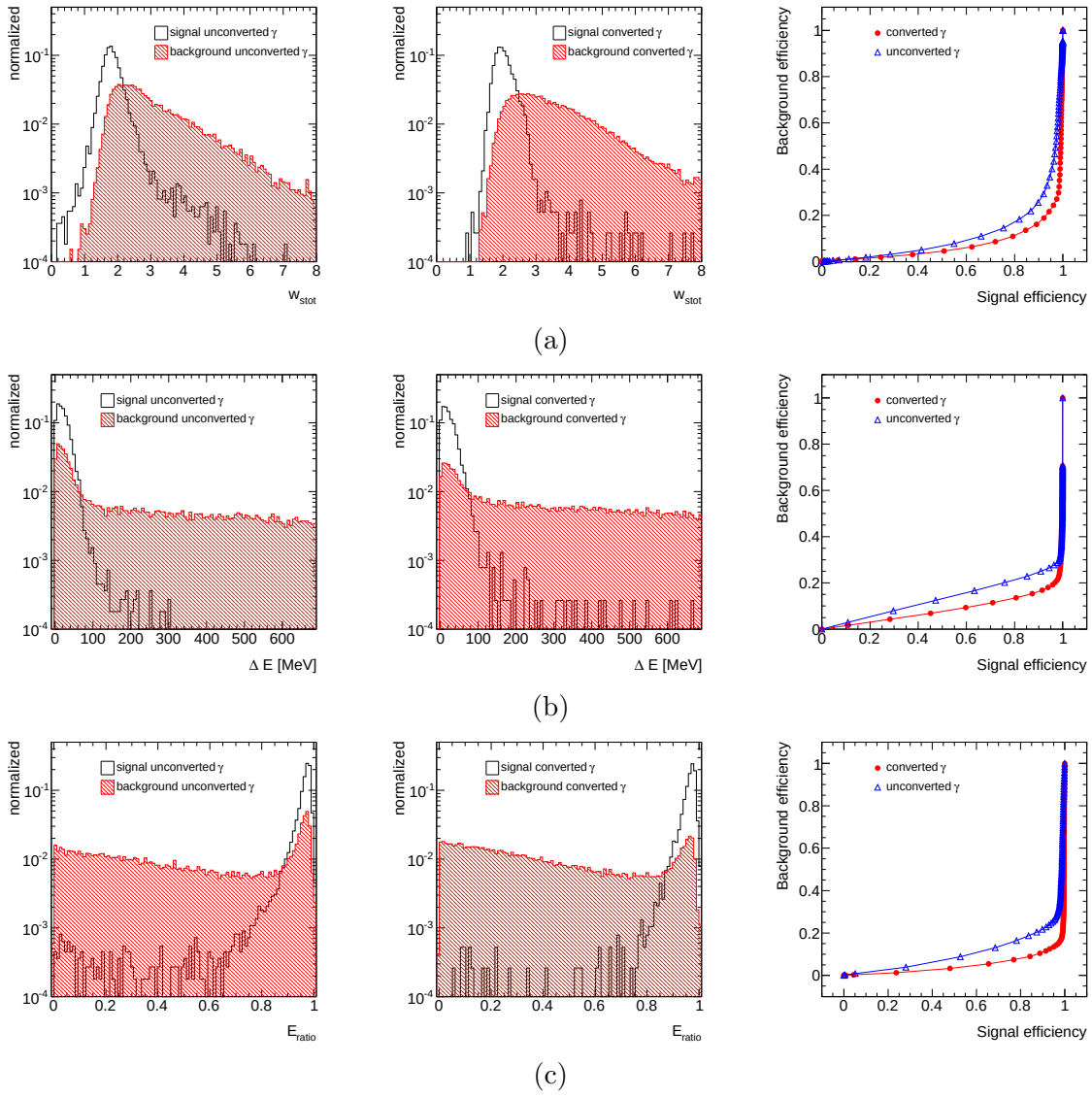


Figure 3.3: Probability distribution function of (a) $w_{s,tot}$ (b) ΔE (c) E_{ratio} for true and fake photons in the region $|\eta| < 0.6$ and with 25 GeV $< p_T <$ 40 GeV before any selection separately plotted for converted (left column) and unconverted (middle column) photons. Background selection efficiency comparison for converted and unconverted photons at the optimized cut as a function of the target signal efficiency is shown in the right column.

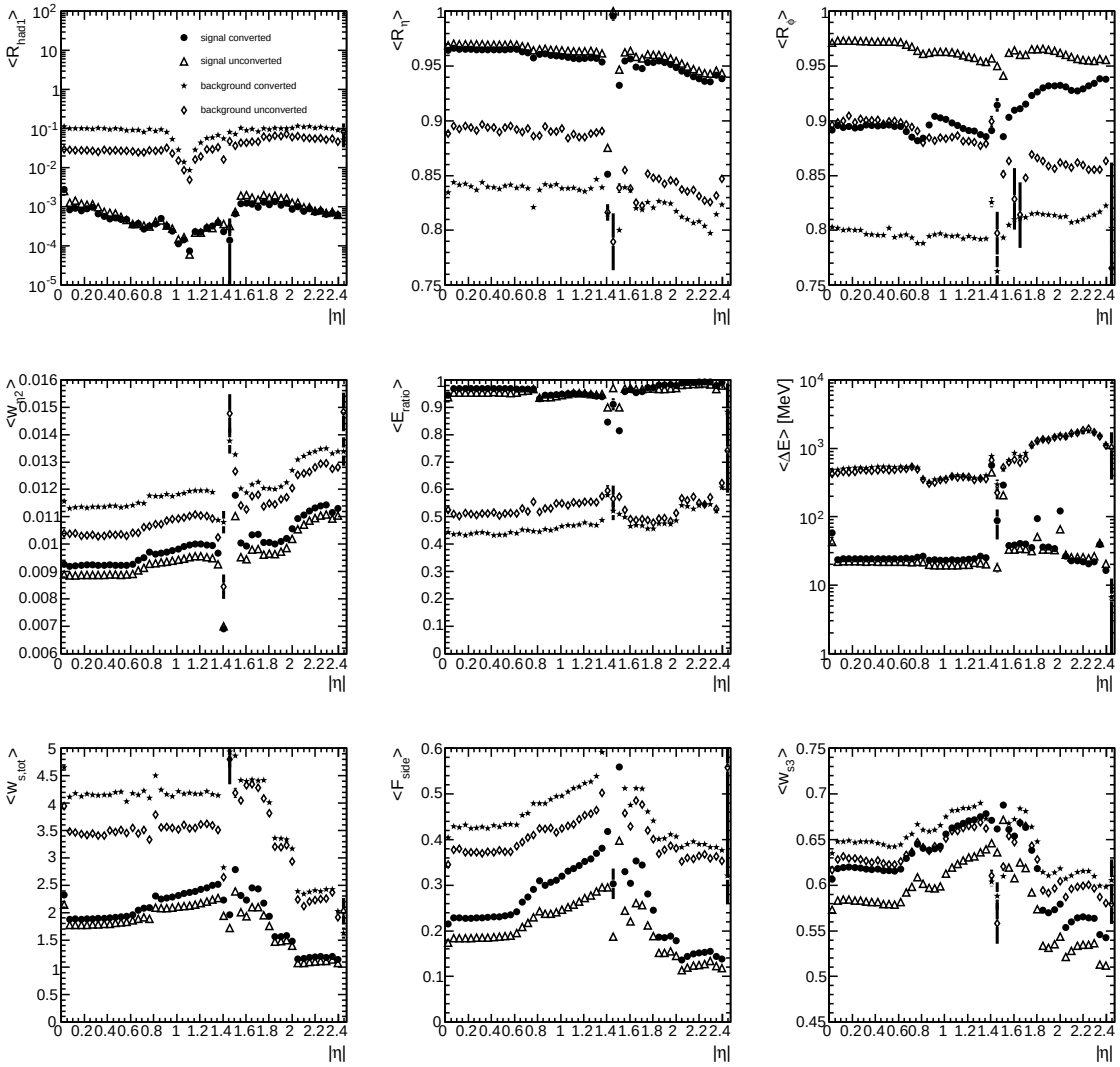


Figure 3.4: Distribution of the means of each calorimetric discriminating shower shape variables as a function of the pseudorapidity $|\eta|$ for true and fake photons with $E_T > 20$ GeV before any selection. Both true and fake photons are separated in converted and unconverted candidates. See reference [77].

Set	1	2	3	4
Label	**PhotonTight**	**PhotonTight**	**PhotonTight**	**RobustTight**
Time	Before mid 2009	mid 2009 $\sim$ Feb. 2010	Feb. 2010 $\sim$ mid 2010	from mid 2010 onwards
Tuning dependence	E_T and η	E_T and η	η	η
Separate conv/unconv	No	Yes	Yes	Yes
		Variables used		
R_{had_1}, R_{had}	$\surd$	$\surd$	$\surd$	$\surd$
R_η	$\surd$	$\surd$	$\surd$	$\surd$
R_ϕ	$\surd$	$\surd$	$\surd$	$\surd$
w_2	$\surd$	$\surd$	$\surd$	$\surd$
F_{side}	$\surd$	$\surd$	$\surd$	$\surd$
$w_{s\,3}$	$\surd$	$\surd$	$\surd$	$\surd$
$w_{s\,\mathrm{tot}}$	$\surd$	$\surd$	$\surd$	$\surd$
ΔE	$\surd$	$\surd$	$\surd$	$\surd$
E_{ratio}			$\surd$	$\surd$
$R_{max\,1}$	$\surd$	$\surd$		

Table 3.2: Evolution for the tight photon identification in ATLAS.

In ATLAS, the photons are selected by cutting on nine shower shape variables. The tight selection used in this work has changed with time. One can define three different periods (Table 3.2). In the oldest period (Set 1), the cut values were optimized in transverse energy E_T and $|\eta|$ bins using Monte Carlo simulation samples. An additional refinement was then done by taking into account the difference between converted and unconverted photons (Set 2). At the beginning of 2010, cuts were simplified to be tuned only on η bins and the $R_{max\,1}$ variable was replaced by E_{ratio} which was found to be more powerful(Set 3). For the data analysis, the selection was simplified and only the $|\eta|$ dependence was used. In addition, cut values were changed from the original Monte Carlo optimization in order to take into account the discrepancy observed between real data and the simulation samples. For the simulation samples used in the photon trigger efficiency study presented in the next chapter, the first set of cuts was applied. For the Higgs search with the di-photon final state based on Monte Carlo samples, as described in Chaper 6, the second set of cuts was used, whereas for the real data analysis, documented in the following three Chapters, Set 4 was used.

In the ATLAS reconstruction program, each of the shower shape variables is associated to a bit number, as shown in Table 3.3. After calling the function isEM() for a given photon candidate, a bitmask is returned, indicating the passed or failed status of the variables. If the photon candidate fails the cut on a given variable, the corresponding bit is set to 1, otherwise 0. Thus, if one requires the photon candidate to pass the selection of all nine variables in the **RobustTight** as well as within the η acceptance, the requirement will be of the form (isEM bitmask & 0x1efc01) == 0.

As an example, the cuts for the **RobustTight** selection (Set 3) are listed in Table 3.4. As already mentioned before, this set of cuts takes into account the differences in the shower shapes observed between the real data and simulated samples. In addition, a **Loose** selection is defined. It is based on the cuts only on the hadronic and middler layer variables: R_{had}, R_η and ω_2. The cuts, as shown in Table 3.5, are looser than the **RobustTight** ones. This selection is used as a preselection in several analyses, i.e. the di-photon analysis detailed in Chapter 5. Besides, it is also used as a benchmark for the trigger selection, which is defined to be slightly looser than the **Loose** one. In practice, the trigger selection on the calorimetric information for the 2010 data is identical for photon and electron triggers. Because the electron selection makes additional use of tracking, the cuts on the calorimetric shower shaper variables can be looser than for the photon. Therefore, the trigger criteria are fixed according to the electron loose selection.

bit	
0	ClusterEtaRange
1	ClusterEtaRange_Loose
2	ClusterHadronicLeakage_Loose
3	ClusterMiddleEnergy_Loose
4	ClusterMiddleEratio37_Loose
5	ClusterMiddleEratio33_Loose
6	ClusterMiddleWidth_Loose
7	
8	
9	
10	ClusterHadronicLeakage
11	ClusterMiddleEnergy
12	ClusterMiddleEratio37
13	ClusterMiddleEratio33
14	ClusterMiddleWidth
15	ClusterStripsEratio
16	ClusterStripsDeltaEmax2
17	ClusterStripsDeltaE
18	ClusterStripsWtot
19	ClusterStripsFracm
20	ClusterStripsWeta1c
21	ClusterStripsDEmaxs1
22	TrackMatchBoverP
23	AmbiguityResolution
24	
25	
26	
27	
28	
29	Isolation
30	ClusterIsolation
31	TrackIsolation

Table 3.3: Photon bits definition of IsEM.

$\|\eta\|$	0-0.6	0.6-0.8	0.8-1.15	1.15-1.37	1.37-1.52	1.52-1.81	1.81-2.01	2.01-2.37	2.37-2.47
Unconverted photon candidates									
R_{had_1}, R_{had}	0.0089	0.007	0.006	0.008	–	0.019	0.019	0.0137	–
R_η	0.950784	0.9398	0.9418	0.9458	–	0.932066	0.928	0.924	–
R_ϕ	0.954	0.95	0.59	0.82	–	0.93	0.947	0.935	–
w_2	0.0107194	0.011459	0.010759	0.011359	–	0.0114125	0.0110	0.0125	–
F_{side}	0.284	0.36	0.36	0.514	–	0.67	0.211	0.181	–
$w_{s\,3}$	0.66	0.69	0.697	0.81	–	0.73	0.631	0.58	–
$w_{s\,\mathrm{tot}}$	2.95	4.4	3.26	3.4	–	3.8	2.4	1.64	–
ΔE	92	92	99	111	–	92	110	148	–
E_{ratio}	0.63	0.84	0.823	0.887	–	0.88	0.71	0.78	–
Converted photon candidates									
R_{had_1}, R_{had}	0.00748	0.007	0.00489	0.008	–	0.0149	0.016	0.011	–
R_η	0.940784	0.9268	0.9298	0.9308	–	0.918066	0.924	0.913	–
R_ϕ	0.4	0.426	0.493	0.437	–	0.535	0.479	0.692	–
w_2	0.0116194	0.011359	0.012859	0.012659	–	0.0138125	0.012	0.0129	–
F_{side}	0.32	0.428	0.483	0.51	–	0.508	0.252	0.215	–
$w_{s\,3}$	0.697	0.709	0.749	0.78	–	0.773	0.652	0.614	–
$w_{s\,\mathrm{tot}}$	2.8	2.95	2.89	3.14	–	3.7	2.0	1.48	–
ΔE	200	200	122	86	–	123	80	132	–
E_{ratio}	0.908	0.911	0.808	0.803	–	0.67	0.915	0.962	–

Table 3.4: Values of the photon tight selection cuts for the different discriminating variables in the different $|\eta|$ regions, for unconverted and converted candidates. R_{had} is used for $0.8 \leq |\eta| < 1.37$, R_{had_1} elsewhere.

$\|\eta\|$	0-0.6	0.6-0.8	0.8-1.15	1.15-1.37	1.37-1.52	1.52-1.81	1.81-2.01	2.01-2.37	2.37-2.47
R_{had_1}, R_{had}	0.010	0.010	0.010	0.008	0.010	0.025	0.019	0.014	0.019
R_η	0.927	0.912	0.921	0.916	0.750	0.906	0.920	0.908	0.915
w_2	0.012	0.012	0.013	0.013	0.025	0.015	0.013	0.013	0.012

Table 3.5: Values of the photon loose selection cuts for the different discriminating variables in the different $|\eta|$ regions.

Chapter 4

Photon Trigger Efficiency

4.1 Introduction

Precise knowledge of the photon trigger efficiency is required in photon-related physics analyses. For instance, it is an important ingredient in the prompt-photon cross-section measurement performed with the first ATLAS data [90] and, in the di-photon analysis and the Higgs search presented in this thesis. New Physics searches involving photons in the final state may also need this measurement.

Unlike electrons and muons, for which high-statistics pure control samples can be obtained from lepton-antilepton decays of narrow resonances such as the Z, Υ or J/ψ, there is no easy way to select a high-statistics, pure photon control sample with the first ATLAS data. In fact, the "standard" sources of photon control samples are typically di-photon decays of resonances such as π^0 or η, or radiative decays of Z bosons ($Z \to ee\gamma$): the former decays do not help in determining photon efficiencies in the p_T region of interest for physics analyses cited before ($p_T \gtrsim 10 - 20$ GeV), while the latter ones have a very small cross-section and are thus limited by statistics and not suitable for photon efficiency determination with the early (few pb^{-1}) ATLAS dataset.

The ATLAS photon triggers are described in Section 4.2. The selection criteria for the trigger items relevant for this study is presented in Section 4.3. Section 4.4 presents how the efficiency, relative to the selection presented in Chapter 3, can be evaluated by three data-driven methods, namely the *tag&probe*, the *bootstrap*, and the *electron-to-photon extrapolation* methods. A Monte Carlo study of the g20_loose single photon trigger efficiency determination is presented. The goal of this study is to investigate in details the feasibility of using these data-driven methods to extract and cross-check the trigger efficiency for photons, and eventually demonstrate the ability of reducing the systematic uncertainty on this quantity. In this perspective the limits and complementarity of the first two methods are studied and compared to the third one. Possible biases on the determination of the relative trigger efficiency are estimated with respect to the true photon trigger efficiency. It should be noted that data-driven methods can be used for the estimation of the photon trigger efficiency only relative to a given photon selection, which is needed in order to increase the purity of the data sample from which to start. In the following this will also be called *marginal* or *relative* trigger efficiency.

All of the three methods discussed in this chapter are limited - in a way which depends on the method chosen - by the size (which affects the statistical uncertainty) and the photon purity (which affects the systematic uncertainty) of the selected control sample. In

particular the two methods relying directly on photon candidates, the tag&probe and the bootstrap methods, are dominated by jets mimicking photons. If the hypothesis that a jet passing tight photon selection criteria has the same trigger efficiency as a true prompt photon was correct, these two methods would give direct access to the photon trigger efficiency. Since this assumption is not properly satisfied, the residual biases are treated as systematics. Finally, having data-driven methods to estimate the photon trigger efficiency is an important tool also for diagnostic and monitoring of the electron-photon triggers performances.

Finally, Section 4.6 presents the results after applying these methods on real data. The good agreement among the three methods ensures the reliability of the results.

4.2 Photon triggers in ATLAS

In ATLAS, electrons and photons are reconstructed by the trigger system in the range $|\eta|$ <2.5. The selection variables for photons at each of the trigger levels are summarized in the following [78].

- **level-1 (L1) selection**

  The whole spatial space ($|\eta| < 2.5$ and $|\phi| < \pi$) is segmented into trigger towers. Each trigger tower has a size of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$. All the cells' energy within the trigger tower are summed over all the layers of the electromagnetic and hadronic calorimeter. If the energy of four central trigger towers in a sliding window of 4×4 trigger towers is the local maximum with respect to its nearest overlapping neighbours, this 4×4 window is considered as a candidate. In the 2×2 tower core of this window, there are four combinations of two neighboring towers. If the energy in at least one of the combinations passes the electromagnetic cluster threshold (the threshold depends on the trigger item. e.g. for L1_EM5, it is 5 GeV), it is then considered to contain an electron or photon candidate. Figure 4.1 shows the L1 trigger tower scheme.

  Isolation requirements can also be defined based on this scheme. The requirements to be imposed on the isolation energy are as follows:

  - The total transverse energy in the 12 towers of the electromagnetic calorimeter surrounding the 2×2 tower core is less than the electromagnetic isolation threshold.
  - The total transverse energy in the 4 towers of the hadronic calorimeter behind the 2×2 tower core of the electromagnetic calorimeter is less than the hadronic core threshold.
  - The total transverse energy in the 12 towers surrounding the 2×2 tower core of the hadronic calorimeter is less than the hadronic isolation threshold.

  The thresholds for the isolation requirement also depends on the trigger item. For a trigger item like L1_EM14I, the thresholds are 4, 3, 2 GeV, respectively. During the data-taking in 2009−2010, isolation requirements were not imposed. It may be implemented in later 2011 in order to control the bandwidth as the instantaneous luminosity increases.

- **level-2 (L2) selection**

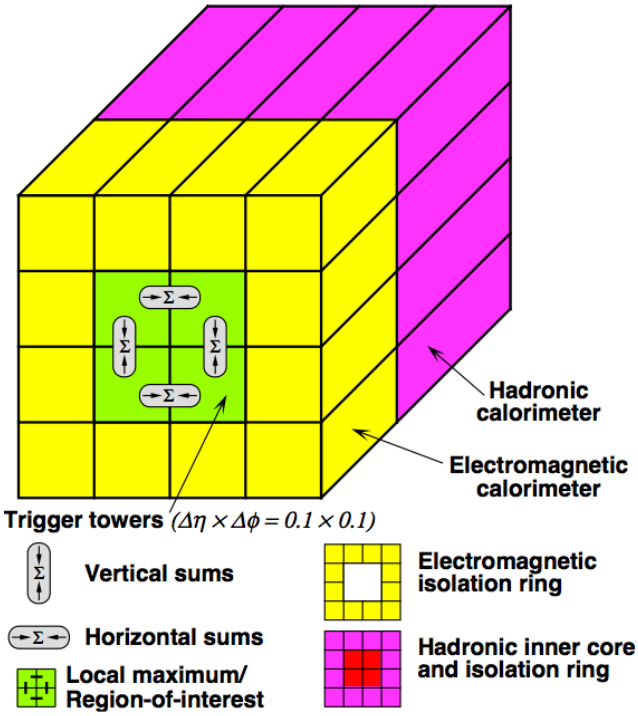


Figure 4.1: L1 trigger tower scheme. The trigger towers used to determine the energy of the electromagnetic cluster as well as isolation are shown.

The L2 calorimeter reconstruction is seeded by the L1 Region of Interest (RoI), which is the sliding 4×4 window defined above. Only the sample of the detector data with RoI passing the L1 requirement is passed to the L2. The L2 cluster building algorithm scans the cells in the second layer of the electromagnetic calorimeter and searches for the cell with the highest energy. A cluster with size of $\Delta\eta \times \Delta\phi = 0.075 \times 0.175$ is built around this most energetic cell. The shower shape variables are calculated using the information from the first and second layers of the electromagnetic calorimeter, as well as the first layer of the hadronic calorimeter. Together with the transverse energies in both calorimeters, they are used by the L2 selection. Requirements are imposed on the following quantities:

- The transverse energy of the electromagnetic cluster (required to be greater than a given threshold).
- The transverse energy in the first layer of the hadronic calorimeter (required to be below a given threshold).
- The ratio of energy in the first layer of the electromagnetic calorimeter over the total energy: f_1.
- The ratio of the energy in the 3×7 cell window of the second layer of electromagnetic calorimeter over the 7×7 cell window. It is required to be larger than a given value.
- The ratio of the difference of the most energetic cell energy and the second most energetic cell's energy over the sum of these two energies. It is required to be larger than a given value for an electron or photon candidate.

	L1		L2		EF	
	L1 item	selection	L2 item	selection	EF item	selection
g20_loose	L1_EM18	$E_T > 18$ GeV	L2_g20_loose	$E_T > 19$ GeV, $E_{T,\mathrm{had}} < 3GeV$, $R_\eta > 0.89$, $f_1 > 0.005$, $E_{\mathrm{ratio}} > 0.6$	EF_g20_loose	$E_T > 20$ GeV, cuts on $E_{T,\mathrm{had}}/E_T$, R_η, R_ϕ, ω_2, f_1, $\omega_{s\,\mathrm{tot}}$, F_{side}, ΔE, $\omega_{s\,3}$
g10HLTPass	L1_EM 7	$E_T > 7$ GeV	L2_g10HLTPass	no requirement	EF_g10HLTPass	no requirement

Table 4.1: Trigger items used in the photon trigger efficiency measurement based on Monte Carlo simulations in this chapter. The relevant selection criteria are also included. Cuts on the shower shape variables at EF for g20_loose are η dependent.

- **event-filter level (EF) selection**

  At EF trigger level, offline reconstruction and tools are used as much as possible. However, due to the limited time for trigger selection, EF reconstruction only accesses the detector data which has passed L2 selection. The EF cluster building is seeded on the RoI given by L2. In the photon trigger selection, only the calorimeter information is used. Beside the requirement on the transverse energy E_T, most of the shower shape variables as in the offline reconstruction are calculated: R_{had}, R_η, ω_2, R_ϕ, $\omega_{s\,3}$, $\omega_{s\,\mathrm{tot}}$, F_{side}, ΔE and E_{ratio}. The selection criteria on those variables depends on each specific trigger item.

The main physics analyses involving photons in the final state (direct photon productions, $H \to \gamma\gamma$, ...) are based on trigger paths requiring one or two trigger photons with a transverse energy greater than ∼20 GeV. Specific energy thresholds, trigger prescales and the sets of cuts on trigger quantities at L1, L2 and Event Filter depends on luminosity and trigger rate conditions and can change accordingly. For an instantaneous luminosity of 10^{31} $cm^{-2}s^{-1}$, the reference physics trigger (i.e. not prescaled) for photon analysis in data taking is designed to be the so called g20_loose single photon trigger stream at EF. The Monte Carlo studies reported in this chapter aim to the measurement of the g20_loose relative trigger efficiency for photons to be used in physics analysis at nominal luminosity. The methods described in this chapter work equally well for higher threshold triggers. Although the EF marginal efficiency was also investigated, only L1 and L2 results are shown, because of a wrong implementation of the EF selection criteria in the ATLAS software release 14 available at the time of this study, which was inconsistent with the tight selection criteria. There is no such problem in the real data analysis, thus results for all the three levels are presented. Another important point is that both the trigger cuts and the photon tight selection have changed since the ATLAS software releases 14 (used in the Monte Carlo analysis) and 16 (used in the data analysis), as described in Chapter 3.

4.3 Triggers items and the corresponding selection criteria

The selection criteria for g20_loose in the Monte Carlo simulation samples is listed in Table 4.1. Another trigger item relevant for these studies is the support trigger called g10HLTPass which provides an un-biased sample in the transverse energy domain around the g20_loose trigger turn-on. The details of the trigger selection in the Monte Carlo simulation for this item is also included in Table 4.1. Those cuts were the default ones in the ATLAS software release 14, used to produce the Monte Carlo simulation samples relevant to this chapter.

	L1		L2		EF	
	L1 item	selection	L2 item	selection	EF item	selection
g20_loose	L1_EM14	$E_T > 14$ GeV	L2_g20_loose	$E_T > 19$ GeV, $E_{T,\mathrm{had}}/E_T$, R_η, f_1	EF_g20_loose	$E_T > 20$ GeV, cuts on $E_{T,\mathrm{had}}/E_T$, R_η, ω_2
g15_loose	L1_EM10	$E_T > 10$ GeV	L2_g15_loose	$E_T > 14$ GeV, $E_{T,\mathrm{had}}/E_T$, R_η, f_1	EF_g15_loose	$E_T > 15$ GeV, cuts on $E_{T,\mathrm{had}}/E_T$, R_η, ω_2
g10_loose	L1_EM5	$E_T > 5$ GeV	L2_g10_loose	$E_T > 9$ GeV, $E_{T,\mathrm{had}}/E_T$, R_η, f_1	EF_g10_loose	$E_T > 10$ GeV, cuts on $E_{T,\mathrm{had}}/E_T$, R_η, ω_2
e3_NoCut	L1_EM2	$E_T > 2$ GeV	L2_e3_NoCut	prescaled	EF_e3_NoCut	prescaled
e10_NoCut	L1_EM5	$E_T > 5$ GeV	L2_e10_NoCut	prescaled	EF_e10_NoCut	prescaled
e15_medium	L1_EM10	$E_T > 10$ GeV	L2_e15_medium	$E_T > 14$ GeV, $E_{T,\mathrm{had}}/E_T$, R_η, track $p_T > 2$ GeV $\Delta\eta(clus, track) < 0.2$	EF_e15_medium	$E_T > 15$ GeV, $\Delta\eta(clus, track)$, impact parameter a_0, number of SCT hits, number of Pixel hits, $E_{T,\mathrm{had}}/E_T$, E_{ratio}, $\omega_{s\,\mathrm{tot}}$, R_η, ω_2
e10_medium	L1_EM5	$E_T > 5$ GeV	L2_e10_medium	$E_T > 9$ GeV, $E_{T,\mathrm{had}}/E_T$, R_η, track $p_T > 2$ GeV $\Delta\eta(clus, track) < 0.2$	EF_e15_medium	$E_T > 10$ GeV, $\Delta\eta(clus, track)$, impact parameter a_0, number of SCT hits, number of Pixel hits, $E_{T,\mathrm{had}}/E_T$, E_{ratio}, $\omega_{s\,\mathrm{tot}}$, R_η, ω_2

Table 4.2: Trigger items used in the photon trigger efficiency measurement with the real data. The relevant selection criteria are also included. Cuts on the shower shape variables at L2 and EF are η dependent.

In the real data taking, the trigger menu evolves together with the luminosity increase. The trigger selections for the various trigger items are adjusted in order to record the interesting data with as high as possible efficiency and within the data acquisition bandwidth limit. Table 4.2 summarizes the trigger items used for the trigger efficiency measurement in real data. The corresponding selection criteria are also included. The trigger item g10HLTPass defined in the Monte Carlo samples is not in the menu for the real data, a similar trigger item called e10_NoCut is used instead.

4.4 Photon trigger efficiency measurement: description of the methods

In any of the three methods presented in the following, the same approach is used: a sample of photon candidates (*probe*) is first selected according to the criteria defined in Chapter 3. The relative trigger efficiencies for L1, L2 and EF are then evaluated by measuring the fraction of probe photon candidates that are matched[1] to a trigger object respectively at L1, L2 and EF. The way the probe photons are selected depends on the method and so do the purity and sample composition. The L1, L2 and EF definitions are nested, i.e. an object that passes a higher-level trigger passes also all lower-level trigger requirements. Applying the photon selection criteria has a two-fold purpose:

- increase the purity $P = \frac{N_{\mathrm{true}}}{N_{\mathrm{true}}+N_{\mathrm{fake}}}$ of the selected sample of probes, which contains N_{true} true photons and N_{fake} fake photons;

[1]The matching criterion used in this study is: $\Delta R = \sqrt{(\eta_{trigger} - \eta_{cluster})^2 + (\phi_{trigger} - \phi_{cluster})^2} < 0.15$.

- reduce the difference in marginal trigger efficiency for true ($\varepsilon_{\text{true}}$) and fake ($\varepsilon_{\text{fake}}$) photons.

The reason is that the marginal photon-trigger efficiency is measured by counting the number $N_{\text{true}}\varepsilon_{\text{true}} + N_{\text{fake}}\varepsilon_{\text{fake}}$ of probes that pass the trigger and dividing it by the total number $N_{\text{true}} + N_{\text{fake}}$ of probes. Thus, our estimate of the efficiency is:

$$\varepsilon = \frac{N_{\text{true}}\varepsilon_{\text{true}} + N_{\text{fake}}\varepsilon_{\text{fake}}}{N_{\text{true}} + N_{\text{fake}}} = \varepsilon_{\text{true}}P + \varepsilon_{\text{fake}}(1-P) = \varepsilon_{\text{true}} - (\varepsilon_{\text{true}} - \varepsilon_{\text{fake}})(1-P) \quad (4.1)$$

Therefore, the measurement bias is proportional to the difference between the true and fake photon efficiencies and to $(1-P)$. It decreases with the tightness of the selection criteria both because the purity is higher and because the fake photons "look more like" true photons thus reducing the difference $(\varepsilon_{true} - \varepsilon_{fake})$.

In the following, three data driven methods are described.

4.4.1 The *tag* & *probe* method

The tag&probe method ideally consists in selecting a clean sample (or a sample with known purity) of di-photon events using one photon as a *tag* and then measure the efficiency of interest using the second photon as a probe.

The method presented here is actually a "pseudo" tag&probe method, since the purity of the sample is not measured (unlike for instance the case for electrons from $Z \to ee$ where the purity of the sample is inferred from the electron pair invariant mass distribution sidebands). The idea here is to use a single-photon-trigger sample with at least two selected photons in each event, one of which is matched to a trigger object (tag), and use the other good photons as probes (see Figure 4.2).

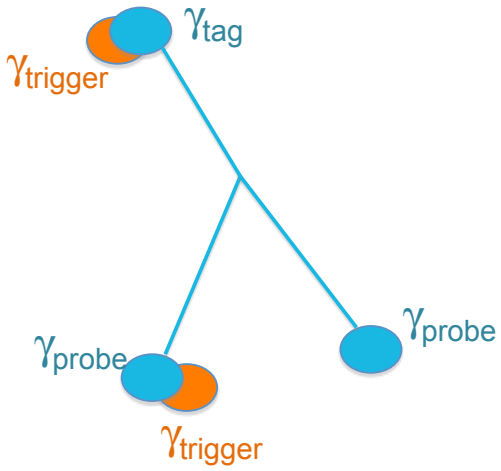


Figure 4.2: Illustration of the tag-and-probe principle.

In practice, the following steps are performed:

1. select events that pass the single photon trigger under study;
2. require one selected photon to be matched to one trigger object. This photon is considered to be the tag photon;

3. all the other selected photons are regarded as probe photons;
4. compute the efficiency as the ratio between the number of probe photons matched to a trigger object and the total number of probe photons.

If at least one probe photon is matched to a trigger object, the tag photon will also be included in the probe photon control sample (i.e. swap the role of the tag and probe photons).

4.4.2 The *bootstrap* method

In the bootstrap method, events are required to contain at least one selected photon matched to a trigger object satisfying looser criteria, usually a lower transverse energy threshold, than those of the trigger line under study in order to enhance the purity of the photon control sample. In the following, this lower energy threshold is called the base trigger. The method proceeds as follows:

1. select the events with at least one photon candidate;
2. require the leading p_T photon candidate to be matched to a trigger object that passes the requirement of the base trigger.
3. count the number of photons passing 1. and 2. that are matched to a trigger object that pass the requirements of the trigger under study;
4. measure the efficiency of the trigger under study as the ratio between the photon candidates passing 1. 2. and 3. and those passing only 1. and 2.

4.4.3 The *electron-to-photon extrapolation* method

For electrons, the tag&probe method applied to the $Z \rightarrow e^+e^-$ process can be used to select a relatively pure electron sample. Since photon and electron interactions in the electromagnetic calorimeter are similar, and that they both deposit almost completely their energy in the calorimeter, the photon trigger efficiency can be measured by assuming that it is similar to the electron trigger efficiency, once the tight photon selection criteria are applied to the electron candidates. Therefore, the standard tag&probe method based on a constraint on the Z mass is used to select the electron sample, to which we apply the tight photon cuts. The photon-like electrons selected in this way are then used to measure for the trigger efficiency under study. The procedure is the following:

1. select electron-positron pairs with invariant mass M_{ee} satisfying $|M_{ee} - 91.2 \text{ GeV}| < 10$ GeV;
2. require one e (either the electron or the positron) to pass the tight electron identification criteria and the other e to pass the loose electron identification criteria; the former is the tag, the second is the probe;
3. require the probe to pass the photon selection;
4. require the tag electron to be matched to a trigger photon;
5. determine the (marginal) efficiency of the trigger by counting the fraction of probe candidates that are matched to a trigger object which passes the requirements of the trigger under study.

4.5 Monte Carlo study

4.5.1 Reconstructed photon selection criteria

In the photon-related analysis, a certain photon identification criterion is needed. As mentioned before, the photon trigger efficiency discussed in this chapter is a marginal trigger efficiency, which is relative to the photons passing the selection. Besides the identification criterion, isolation requirement is also imposed. The selection criteria includes:

- Tight photon identification using the first set of cuts defined in Section 3.3.
- Crack region rejection: $(|\eta| < 1.37)$ || $(1.52 < |\eta| < 2.37)$;
- Track isolation: sum of the transverse momenta (p_T) of all tracks with $p_T > 1$ GeV passing a certain quality selection in the $0.1 < \Delta R < 0.3$ range to be lower than 4 GeV. The lower ΔR range limit is set to preserve converted photon tracks.
- Overlap removal: if two photon candidates are found within a cone $\Delta R < 0.1$, that with lower p_T is removed.

Similar but slightly different photon selection is applied on the real data, which is the **RobustTight** cuts defined in Section 3.3.

4.5.2 Monte Carlo samples

All the Monte Carlo samples used in this analysis were generated with PYTHIA 6.4 by the ATLAS MC production and simulation group. For this photon trigger efficiency study, di-jets (jj), photon-jet (γj), di-photon ($\gamma\gamma$) samples as well as an inclusive $Z \to e^+e^-$ sample are used. They were simulated and processed by the ATLAS software, ATHENA [79], release 14. The dataset ID, process, cross section, event-filter (with efficiency), number of reconstructed events and equivalent luminosity of these samples are summarized in Table 4.3.

The JF17 sample is a mixture of jj, γj and γ^*, W, Z^0 events; all non di-jet events in JF17 are removed at analysis level in order to avoid overlaps with the exclusive γj sample. In the following therefore, unless explicitly stated otherwise, di-jet events refer to the JF17 sample after non-di-jet-events removal, while γj events refer to the exclusive photon-jet MC sample.

Dataset	Process	σ [pb]	Filter	Filter Eff. [%]	Events used	Luminosity [pb^{-1}]
106379	γj	2.890E+5	AsymJetFilter	3.58	2907432	281.0
108086	γj	2.885E+5	SymJetFilter	7.06	3469364	170.3
108087	γj	2.885E+5	PhotonFilter	43.04	980337	7.9
105802	JF17 (mostly jj)	1.461E+9	JetFilter	7.06	9688178	0.09
105964	$\gamma\gamma$	934.449	PhotonFilter	12.81	99939	834.9
106050	$Z \to e^+e^-$	1143.96	LeptonFilter	96.0	734461	668.8

Table 4.3: MC samples used in this study

In order to have a high generation efficiency, the samples were generated with a cut in the $\hat{p}_T$ of the hard scattering at parton level, corresponding to PYTHIA parameter `CKIN(3)`, which is required to be:

- $> 15\ GeV$ for the jj and γj samples
- $> 10\ GeV$ for the $\gamma\gamma$ sample
- for $Z \to e^+e^-$, this cut is replaced by the requirement that the e^+e^- invariant mass be greater than 60 GeV.

On top of this, to minimize simulation and reconstruction time, a filter is applied at particle level (after the showering) to the generated sample, requiring the presence in the event of either one or more photons with a certain transverse energy (*PhotonFilter*), either one or more leptons (*LeptonFilter*), or jets (*JetFilter*). In the latter case, a jet is defined as an energy flow of stable particles through a region (square or cone) in a (η, ϕ) grid, mimicking a jet reconstruction algorithm. A cut in the transverse energy of such a particle-level "jet" is then applied in the filter. The filters used in the production of the aforementioned samples are:

- AsymJetFilter: at least 1 jet with $p_T > 30\ GeV$, $|\eta| < 2.7$ and 1 jet with $p_T > 20\ GeV$, $|\eta| < 2.7$
- SymJetFilter: at least 2 jets with $p_T > 17\ GeV$, $|\eta| < 2.7$
- PhotonFilter: at least 1 or 2 photons (1 for γj, 2 for $\gamma\gamma$) with $p_T > 15\ GeV$, $|\eta| < 2.7$
- JetFilter: at least 1 jet with $p_T > 17\ GeV$, $|\eta| < 2.7$
- LeptonFilter: at least 1 lepton with $|\eta| < 2.8$

The samples listed in Table 4.3 have been produced in ATLAS especially for direct-photon cross-section or $H \to \gamma\gamma$ analyses. Therefore the generator cuts and filter thresholds are rather high with respect to the needs of this study, which focuses on the efficiency behavior close to trigger turn-on at around 20 GeV. As a consequence, in principle, the exclusive photon-jet MC sample that is best suited for this analysis would be the one with dataset ID 108087, which has a single photon filter at 15 GeV, being the most unbiased one at low E_T. The other two photon-jet MC samples have the filter requirement of two jets, which at some extent are more biased. However, the statistics for the sample with dataset ID 108087 is rather limited and when the tag&probe selection is applied, the remaining probe sample is very small as it will be shown in the next section. Therefore, for the study of the γ-jet contribution to the tag&probe method, the sample with dataset ID 108086 is chosen as the best compromise between available statistics and unbiasedness. For the other physics contributions to the analysis (di-jets, di-photons, and $Z \to ee$) the samples used are the ones listed in Table 4.3. The biases on the measured trigger efficiencies and on sample composition, due to the generator filters and cuts will be addressed as a systematics of the method as detailed in Section 4.5.4.

4.5.3 Results of the tag&probe method

As described in Section 4.4.1, this method requires at least two photon candidates in each selected event, one of the two photons (tag) is matched with the g20_loose trigger object

requirement, the other photon candidate is the probe. The main issue for the extraction of the photon relative trigger efficiency from this data-driven method resides in the selected sample composition in a two-fold way: in terms of true photons or jets faking a photon and in terms of different contributions coming from the relevant physics processes involved. As already stated in Section 4.4, any data-driven method attempting to measure the photon efficiency will be unbiased in the case this quantity is exactly the same for true and fake photons. The tight photon selection with respect to which the trigger efficiency is evaluated makes the true and fake photons have similar efficiencies. The observable efficiency with data-driven method will be an average of the two, weighted by the sample composition. It is then crucial to evaluate these discrepancies and their dependance on the sample composition to validate the tag&probe method and the achievable uncertainty on the trigger efficiency. In this section, the trigger efficiencies ϵ_{L1} and ϵ_{L2} are extracted from the different Monte Carlo samples for true photons and compared to the corresponding efficiencies for fakes and the differences evaluated. The main issues on the physics process composition are addressed. Finally, the results and the main systematics of the tag&probe method are discussed.

- **True and fake photons trigger efficiency turn-on curves**

  A comparison is presented between true and fake photon efficiency for probes selected in di-jets (jj), photon-jet (γj+$j\gamma$) and di-photon ($\gamma\gamma$) events, after applying the tag&probe selection as described in Section 4.4. Figure 4.3 shows the L1 (left) and L2 (right) trigger efficiency as a function of transverse momentum for true and fake photons. Here, "true" photon means a reconstructed photon candidate that is truth-matched with a photon originating from the hard-process photon or from parton bremsstrahlung, whereas any other candidate is considered as "fake" photon.

  The only statistically significant comparison between true and fake photons can be done for the case of γj sample (Figure 4.3 second row), where the fake p_T trigger turn-on is steeper for true photons than for fakes. Also the L2 efficiency plateau for fake photons is slightly lower than for true ones. In the case of the di-photon sample (Figure 4.3 first row), the fake fraction is very small as one can see from the large statistical fluctuations. For the di-jet sample (Figure 4.3 third row), both true and fake photon turn-on curves suffer from fluctuations due to the low statistics of the selected probe sample. In fact, even though the di-jet sample in itself is large, of the order of 10^7 events, given the typical jet rejection rate of the offline photon selection which is $O(10^3)$, the requirement of two reconstructed photons gives typically a reduction factor of $O(10^5)$. The remaining $O(10^2)$ events with two photon candidates are not enough to determine with high precision the trigger turn-on curve. The consequences of this issue are detailed below. Parameterization for di-jet sample is proposed.

- **Tag & probe sample composition**

  Using the physics samples listed in Table 4.3, their relative process cross-sections and the corresponding luminosities, we evaluated the tag&probe sample composition in terms of expected number of probes as a function of p_T, rescaled to an ATLAS integrated luminosity of 20 pb^{-1}. The results of this evaluation are shown in Table 4.4 together with the statistical uncertainties.

  The cross-section of di-jets events is so large ($O(10^9)$ pb) that, even though all the events from the official production have been used, the equivalent integrated luminosity corresponds to 0.094 pb^{-1}. As already mentioned, very few probe photons from

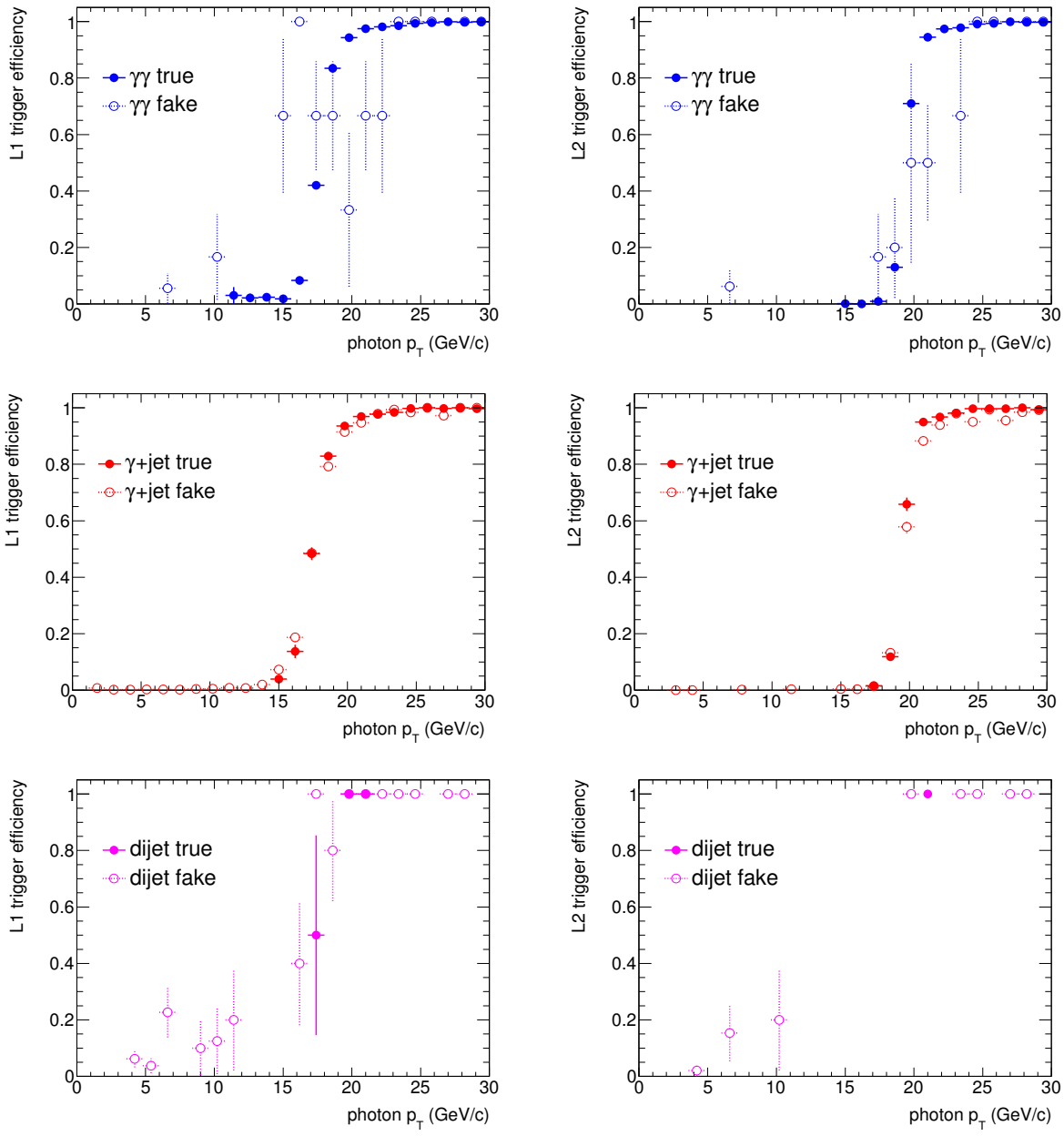


Figure 4.3: g20_loose efficiency as a function of the transverse momentum for true and fake photons in the $\gamma\gamma$, γj, jj simulated samples. Top: $\gamma\gamma$; Middle: γj; Bottom: jj. Left column (a): L1 trigger efficiency. Right-column (b): L2 trigger efficiency.

Sample	Photon p_T [GeV/c]		
	< 14	$14 - 20$	> 20
$\gamma\gamma$	14 ± 1	341 ± 3	909 ± 5
γj	1797 ± 15	577 ± 8	1420 ± 13
jj	49362 ± 3241	4043 ± 927	4255 ± 952

Table 4.4: Expected sample composition in terms of the number of probe photons as a function of p_T

di-jet events survive after the two-photons tag&probe selection. However, when all the samples are scaled to the same luminosity, di-jet events dominates over photon-jets and di-photon ones in the composition of the tag&probe sample, as can be seen in Table 4.4. The estimate of the di-jet contribution is then affected by a large statistical uncertainty due to the large luminosity rescaling factor ($20/0.094 = 212.8$). Thus it is very important to find another way to estimate accurately the contribution from the di-jet sample.

- **parameterization of the di-jet sample**

  To overcome the di-jet sample statistic issue, an approach to "factorize" the two-photon tag&probe selection is proposed. This approach starts by selecting events with at least one jet in the truth-jets collection with $E_T > 14$ GeV and then consider all the jet pairs which are likely to fake a photon pair. It is required that the leading p_T jet of the selected pair is matched to a reconstructed photon candidate. The photon fake rate for the leading selected jet is evaluated. Then one releases the photon matching requirement on the first jet of the pair and loops over all the possible jets matching a reconstructed photon in the event, considering all the possible combinations. From this second step, one can extract the photon fake rate for the subleading jet. The di-photon fake rate is then obtained by multiplying the one for the leading jet and that for the subleading one. The total number of expected di-photon candidates after the tag&probe selection is then obtained by summing up all the possible independent jet pairs combinations. This procedure has the benefit of keeping a decent sample statistics (only one jet rejection suppression factor) while not disregarding the di-jet process event topology features (dependence of the di-photon fake rate from jet multiplicity in the event). The underlying hypothesis here is that the photon fake rates of the two jets of the selected pairs are uncorrelated and can then be simply multiplied to obtain a joint di-photon rate. This hypothesis as we will see in the following is verified at least in the photon p_T range relevant for this analysis. Possible disagreements are treated as systematics and detailed in the corresponding systematics section.

  The p_T and η distributions (denoted as h_{probe} in the following) for fake photon probe candidates from di-jet events entering the tag&probe selection are estimated in the following way. Events where the tag is faked by the true jet that has the highest p_T (*leading* jet) and the probe matches with a true subleading jet are first considered. R_L is defined as the ratio of the number of leading jets matching a tag photon (passing the tight photon selection and matching a g20_loose object) over the number of total leading jets, and h_{SL} is the p_T distribution of the photon candidates matching with

subleading jets. The p_T distribution of those probes is then proportional to $R_L \times h_{SL}$. To handle the case in which the probe matches the true leading jet and the tag if faked by the subleading jet, R_{SL} and h_L are defined similarly. The total p_T distribution is then of the form $R_L \times h_{SL} + R_{SL} \times h_L$. However, the contribution from other jets beyond the leading and the subleading ones have to be taken into account. Furthermore, the dependence of R_L, R_{SL}, h_L, h_{SL} as a function of η and p_T must be parameterized. The resulting formula for h_{probe} can be expressed in the following way:

$$\begin{aligned} h_{probe} &= F_{scale} \times \sum_{\substack{a=1,5\\b=1,6}} \left\{ R_L(a,b) \left[\sum_{\substack{i=1,5\\j=1,6}} h_{other}(i,j) f_2(i,j) \mid_{L(a,b)} \right] \right\} \\ &+ F_{scale} \times \sum_{\substack{a=1,5\\b=1,6}} \left\{ \left[\sum_{\substack{i=1,5\\j=1,6}} R_{SL}(i,j) f_1(i,j) \mid_{L(a,b)} \right] h_L(a,b) \right\} \end{aligned} \quad (4.2)$$

where

- $|\eta|$ bins: $a, i = \{0-0.8,\ 0.8-1.37,\ 1.52-1.8,\ 1.8-2.0,\ 2.0-2.37\}$;
- p_T bins [GeV]: $b, j = \{14-20,\ 20-30,\ 30-40,\ 40-50,\ 50-80,\ >80\}$;
- F_{scale} is a luminosity scaling factor: for an equivalent luminosity of 20 pb^{-1}, the factor is $20/0.094 = 212.8$;
- $R_L(a,b)$ is the ratio of the number of leading jets in bin (a,b) matched to a photon candidate and to a g20_loose trigger object over the number of total leading jets in bin (a,b);
- $h_{other}(i,j)$ is the p_T distribution of the photon candidates matching with any other jets except the leading jets in bin (a,b);
- $f_2(i,j) \mid_{L(a,b)}$ is the fraction of the number of other jets (except the leading jets) in bin (i,j) over the total number of the other jets when the leading jet lies in bin (a,b).
- $R_{SL}(i,j)$ is the ratio of the number of subleading jets in bin (i,j) matched to a photon candidate and to a g20_loose trigger object over the number of total subleading jets in bin (i,j);
- $f_1(i,j) \mid_{L(a,b)}$ is the fraction of the number of subleading jets in bin (i,j) over the number of the total subleading jets when the leading jet lies in bin (a,b);
- $h_L(a,b)$ is the p_T distribution of the photon candidates matching with leading jets in bin (a,b);

After applying this parameterization on the di-jet sample to emulate the tag&probe selection, a much smoother probe photon distribution is obtained and the estimated statistics with 20 pb^{-1} is in good agreement with the standard method (requiring at least two photon candidates in each event). The estimated number of probes with this parameterization and a comparison to the standard method are given in Table 4.5 in coarse p_T bins.

Table 4.5: Probe photon sample composition as a function of p_T

Sample	Photon p_T $[GeV]$		
	< 14	$14 - 20$	> 20
jj (standard)	49362 ± 3241	4043 ± 927	4255 ± 952
jj (parameterised)	22560 ± 371	4638 ± 118	3160 ± 205

The expected number of probes obtained with the standard method and with the parameterization method are consistent within errors above 14 GeV, while they are inconsistent below 14 GeV. This is mainly due to the selection of jets used in the parameterization procedure, which considers only jets with $E_T > 14$ GeV. This choice was motivated by the need of keeping the number of jet pairs combinations as low as possible for the calculation of Equation 4.2 while trying not to bias the photon p_T region under study which is well above that E_T threshold.

Another interesting quantity that can be used to compare the two estimates of the di-jet contribution to the tag&probe method is the p_T spectrum of the selected probes, which is shown in Figure 4.4. Again, a large discrepancy can be seen in the low p_T region, while a satisfactory agreement, at this stage, is observed around and above the trigger turn-on region.

- **Improved tag&probe sample composition**

  Once the parameterization procedure for the di-jet sample contribution is assessed, a new estimate of the tag&probe sample composition is performed by combining the three physics processes according to their cross-sections. A combined data sample, meant to be the best estimation of the real data composition, is then obtained. Figure 4.5 shows the predicted probe sample composition in terms of the different physics processes as a function of the probe photon p_T.

  Figure 4.6 shows the composition of the probe photon sample in terms of the physics processes and of true and fake photons. The different colors correspond to $\gamma\gamma$, γj, jj relative fractions respectively and the dashed (solid) areas correspond to the fractions of true (fake) photons. It is consequently possible to have an estimate of the probe photon purity, defined as the fraction of probes matched to a true photon over the total number of probes. In Figure 4.7, the purity is shown as a function of the photon p_T for L1 and L2 probes. In the turn-on region, the purity is predicted to be in the range between 20% and 50% while it increases to around 80% at higher photon transverse momenta. From the information contained in Figure 4.6 one can notice that especially around the trigger turn-on (i.e. around 20 GeV/c) a sizable fraction of true photons is predicted to come from bremsstrahlung photons in the di-jet sample. In the higher p_T regime the hard scattering production of prompt photons from γ-jet and di-photon samples are taking over.

  In this study, the purity is estimated from the Monte Carlo prediction. This prediction, in the low p_T region, is very sensitive to the showering model that is used. Since the purity knowledge affects the observable efficiency, this uncertainty must be taken into account when computing the systematic uncertainty of the tag&probe method.

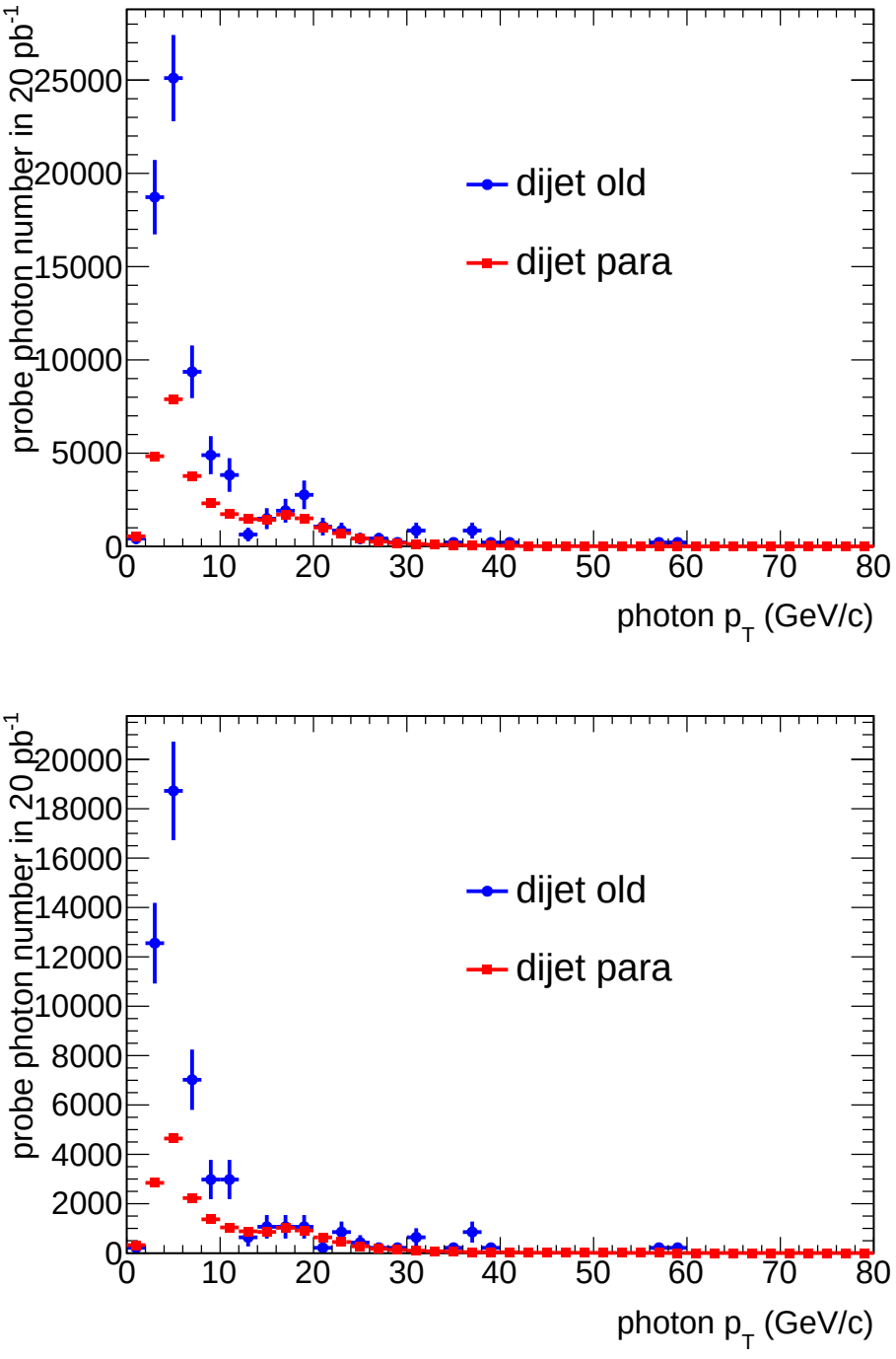


Figure 4.4: p_T spectrum of L1 (top) and L2 (bottom) probe photons in the di-jet sample from direct tag&probe selection (blue circles, **old**) and from the parameterization (red squares, **para**).

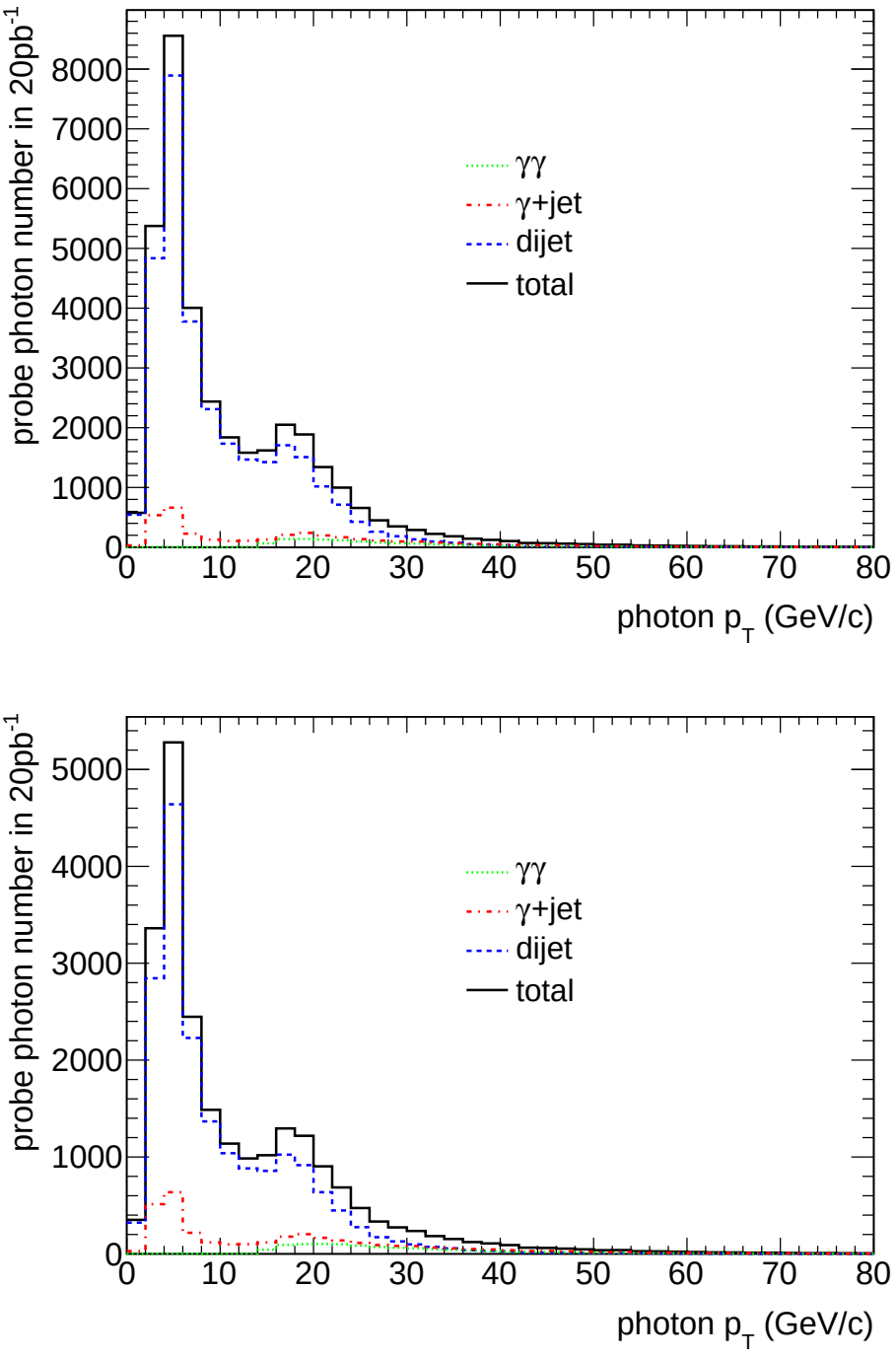


Figure 4.5: Expected p_T spectrum of L1 (top) and L2 (bottom) probe photons in the tag&probe method, for photon candidates from $\gamma\gamma$, γj and jj events. The number of probes is estimated for an integrated luminosity of 20 pb^{-1}.

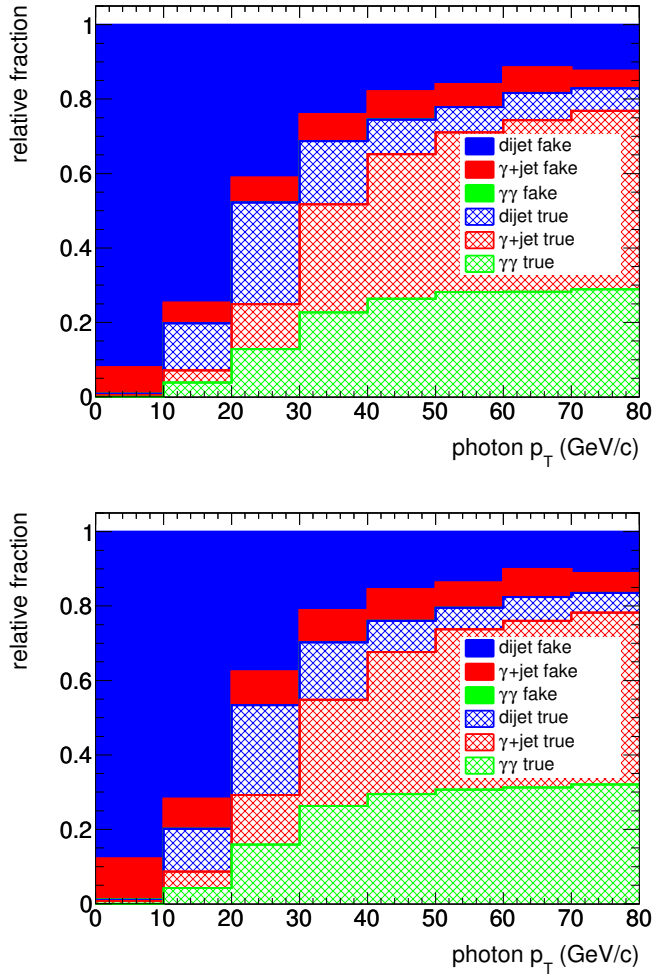


Figure 4.6: Sample composition with $\gamma\gamma$, γj, jj selected by tag&probe method at: L1 (top) and L2 (bottom)

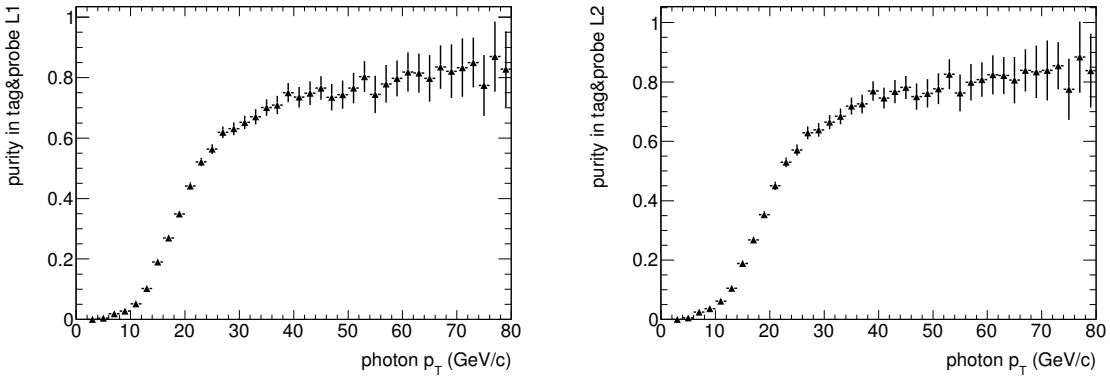


Figure 4.7: Purity of the probe photon sample, selected with the tag&probe method, as a function of p_T, at L1 (left) and L2 (right).

- **Results on trigger efficiency**

  Figure 4.8 shows the efficiency as a function of the transverse momentum of the probe at L1 and L2 levels. The $\gamma\gamma$, γj, jj samples are scaled to 20 pb^{-1} and combined to obtain the baseline data sample in which, for jj sample, the parameterization described previously has been applied. Even though the purity in the turn-on region is low, since the efficiency of true photon probes is close to the efficiency of fakes, the measured overall efficiency (black solid circles in Figure 4.8) is close to the true photon probe efficiency (red triangles in Figure 4.8). The detailed results for trigger efficiency and purity in several p_T bins are summarized in Tables 4.6 for L1 and L2 trigger respectively.

4.5.4 Tag&probe systematics

In the perspective of using the tag&probe method to determine the relative trigger efficiency for photons, differences between the efficiency measured on the probe sample and the true photon efficiency as predicted by Monte Carlo are treated as systematics. In this way, in principle, one tends to overestimate the systematics with respect to the choice of directly correcting for the bias using the Monte Carlo or other independent information from the data. Given that the size of those biases is comparable to or of the same order of the expected statistical error with 20 pb^{-1} of integrated luminosity, we chose to quote the full discrepancy as systematics on the trigger efficiency.

The biases on the determination of the efficiency can be divided into two classes, coming from:

- the knowledge of the sample purity
- the intrinsic biases in the method

The systematic uncertainty coming from the purity of the probe sample is directly connected to the uncertainty on the relative sample composition in terms of $\gamma\gamma$, γj, jj samples which have different purities. The overall observable efficiency as a function of the purity P is $\varepsilon(P) = \varepsilon_{\text{true}} \times P + \varepsilon_{\text{fake}} \times (1-P)$; the associated uncertainty is $\sigma_\varepsilon = \frac{1}{2}[\varepsilon(P+\sigma_P) - \varepsilon(P-\sigma_P)]$, with σ_P the uncertainty on the purity itself. In the limit of exactly equal efficiencies between true and fake photons ($\varepsilon_{\text{true}} = \varepsilon_{\text{fake}}$) the efficiency does not depend on the purity and hence the knowledge of the sample composition is irrelevant. As can be seen in Figure 4.8, $\varepsilon_{\text{true}}$ and $\varepsilon_{\text{fake}}$ (red triangles and green squares respectively) are predicted to be very close but with some significant differences. This will be reflected in a systematic uncertainty on the efficiency coming from a mild dependence of the observable efficiency on the purity. Two main sources of purity systematic uncertainties are considered:

1. the generator level filter biases on the γj sample
2. the parameterization of the di-jet sample.

- **Purity-related systematic uncertainty from the generator level biases**

  The generator level biases can be estimated by comparing the three γj samples shown in Figure 4.9. Tables 4.7 and 4.8 give the expected number of probes with an integrated luminosity of 20 pb^{-1} as estimated from the different γj samples. The sample with ID=108087 was generated with a single photon filter at 15 GeV and hence is in

Quantity	photon p_T [GeV/c]							
	$14-16$	$16-18$	$18-20$	$20-24$	$24-30$	$30-50$	$50-70$	>70
L1								
ϵ(true γ) [%]	2.48 ± 0.72	33.90 ± 1.61	86.78 ± 1.06	97.21 ± 0.49	99.42 ± 0.26	99.94 ± 0.08	99.86 ± 0.26	99.87 ± 0.34
ϵ(fake γ) [%]	5.49 ± 0.60	34.92 ± 1.18	84.39 ± 0.98	96.32 ± 0.54	99.21 ± 0.37	99.95 ± 0.12	100.00 ± 0.00	100.00 ± 0.00
ϵ(total) [%]	4.92 ± 0.50	34.65 ± 0.97	85.23 ± 0.74	96.74 ± 0.37	99.34 ± 0.21	99.95 ± 0.06	99.89 ± 0.21	99.89 ± 0.29
purity [%]	19.01 ± 0.84	26.91 ± 0.83	34.90 ± 0.94	47.57 ± 1.03	59.67 ± 1.29	70.25 ± 1.26	79.09 ± 2.59	85.06 ± 3.12
L2								
ϵ(true γ)[%]	0.01 ± 0.01	0.78 ± 0.36	32.31 ± 1.75	95.15 ± 0.78	99.27 ± 0.34	99.88 ± 0.13	99.77 ± 0.35	99.75 ± 0.49
ϵ(fake γ)[%]	0.15 ± 0.12	0.80 ± 0.27	26.85 ± 1.49	90.22 ± 1.05	97.85 ± 0.71	98.80 ± 0.62	99.17 ± 1.37	100.00 ± 0.00
ϵ(Total)[%]	0.12 ± 0.10	0.79 ± 0.22	28.78 ± 1.15	92.61 ± 0.66	98.71 ± 0.35	99.57 ± 0.19	99.66 ± 0.39	99.79 ± 0.42
purity [%]	18.84 ± 1.02	26.77 ± 1.01	35.29 ± 1.14	48.46 ± 1.26	60.59 ± 1.50	71.91 ± 1.35	80.48 ± 2.64	85.93 ± 3.17

Table 4.6: L1, L2 trigger efficiency for true, fake and total photon candidates, and purity of the selected control samples, as a function of p_T, obtained with the tag&probe method, together with their statistical uncertainties.

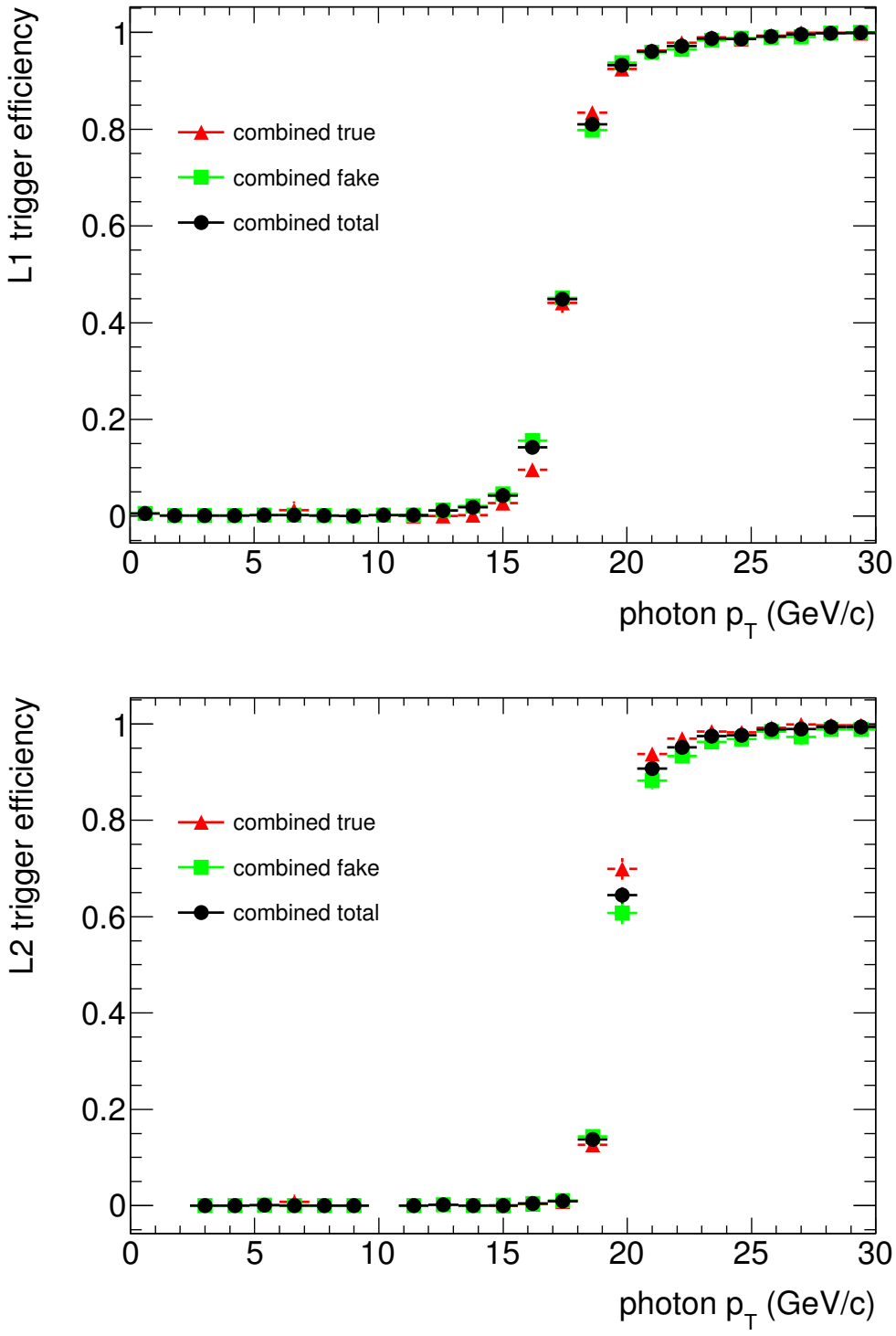


Figure 4.8: g20_loose efficiency as a function of the probe transverse momentum for the combined $\gamma\gamma$, γj, jj data sample, for true photons (red triangles), fake ones (green squares) and all photon candidates (black circles). Top: L1 trigger efficiency. Bottom: L2 trigger efficiency.

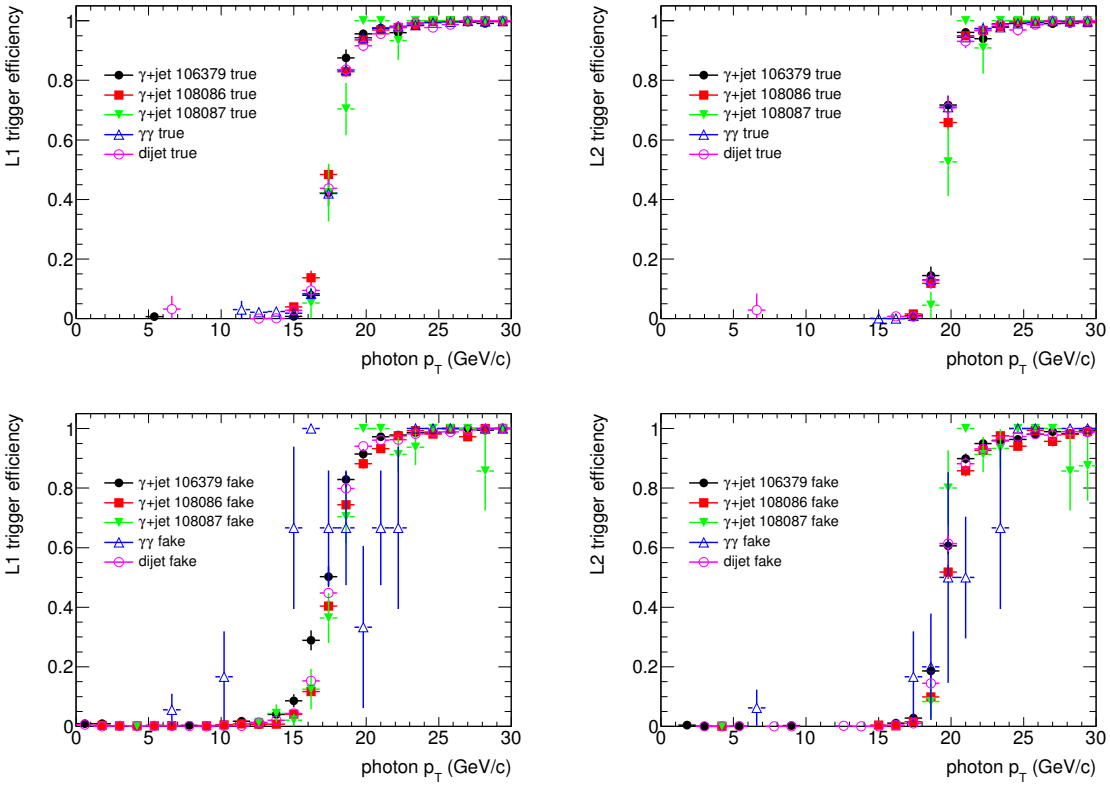


Figure 4.9: *g20_loose* efficiency as a function of p_T for true photons (top plots) and fake photons (bottom plots) compared for the: $\gamma\gamma$, di-jet and the three γj samples. For L1 (left) and L2 (right).

principle the less biased in the low momentum region; as already stated, we did not chose it as the reference γj sample only because of its limited statistics. The sample with dataset ID=106379, compared to the one with ID=108087, underestimates the γj contribution: this is a consequence of the two-asymmetric-jet generator-level filter that requires two jets with p_T greater than 20 and 30 GeV, thus affecting the estimate at lower momenta. Since the baseline sample (dataset ID=108086) is generated with a symmetric-jet filter requiring at least two jets with p_T >17 GeV, it probably underestimates the γj contribution, especially for the p_T range $14 - 20$ GeV.

As a practical example of how the efficiency systematic uncertainty related to the generator-level filter bias is computed, let us consider the p_T region $14 - 20$ GeV. If the number of probes in γj events is scaled from 577 (estimated with the dataset with ID=108086) to 925 (estimated with the dataset with ID=108087), the fraction of γj in the combined tag&probe sample changes from 8% to 13% and the overall purity increases by 1%. The effect on the efficiency is found to be negligible (below 0.01%). Note that for this estimation, only the rescaling is applied: the efficiency in the γj component itself is kept unchanged.

γj dataset ID	photon p_T [GeV/c]		
	< 14	$14 - 20$	> 20
106379	1003 ± 8	341 ± 3	959 ± 8
108086	1797 ± 15	577 ± 8	1420 ± 13
108087	12044 ± 203	925 ± 56	1473 ± 71

Table 4.7: Estimated number of photon probes obtained in γj events with the tag&probe method as a function of p_T at L1.

γj dataset ID	photon p_T [GeV/c]		
	< 14	$14 - 20$	> 20
106379	999 ± 8	124 ± 3	919 ± 8
108086	1712 ± 14	499 ± 8	1214 ± 12
108087	10068 ± 185	788 ± 52	1240 ± 65

Table 4.8: Estimated number of photon probes obtained in γj events with the tag&probe method as a function of p_T at L2.

- **Purity-related systematic uncertainty from the parameterization of the di-jet sample**

  The purity of the probe sample is affected by a large systematic uncertainty dominated by the estimation of the di-jet contribution. A comparison between the purity estimates obtained with the parameterization and the standard method indicates that the typical variations in the turn-on region are in the range between 10 and 20%. This effect dominates the purity-related systematics. In order to estimate this uncertainty, the di-jet parameterization probe p_T spectrum (Figure 4.5) is rescaled to the one predicted by applying the standard two-photon selection on the di-jet sample, for each p_T bin, allowing in this way for maximum sample composition shifts. A rather

conservative choice to define the systematics on the purity from this effect, is made:

$$\sigma_P = |P_{para} - P_{standard}|/2$$

and the corresponding efficiency difference $\Delta\epsilon$ are calculated. Values for the systematic error on the efficiency around 0.5% or below are found in the turn-on region. The full systematic error from the purity uncertainty is detailed in Tables 4.9 and 4.10 where the two contributions from the generator-level filter bias and the di-jet purity are summed in quadrature.

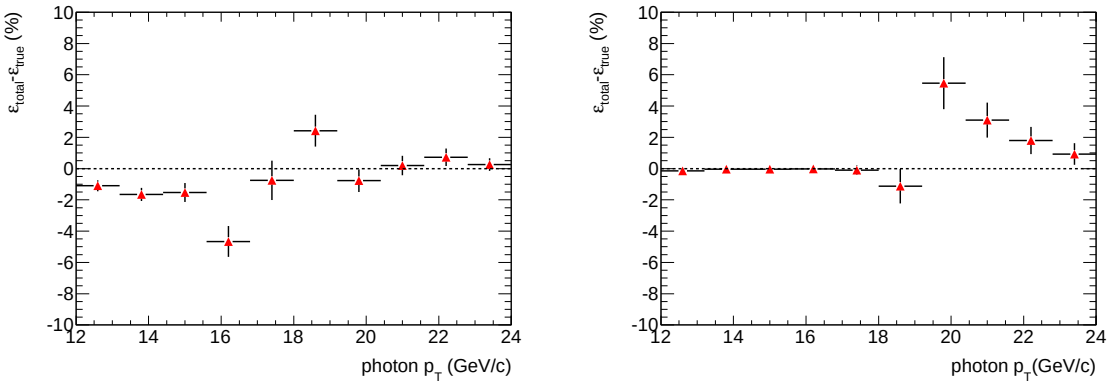


Figure 4.10: Efficiency residuals (in %) between the efficiency measured with the tag&probe method and the true photon efficiency as a function of p_T; L1 (left) and L2 (right).

- **Method systematic uncertainty**

  All the remaining systematic effects intrinsically related to the tag&probe method are estimated from the residuals between the measured efficiency with respect to the true photon efficiency. The results are shown in Figure 4.10 (for L1 and L2), which shows the differences between the true and measured efficiencies displayed in Figure 4.8. The full residual for each probe photon p_T bin is reported in Tables 4.9 and 4.10 in the "Method systematics" row. This choice is over-conservative with respect to the potentialities of tag&probe method. It is essentially motivated by the comparison with the statistical uncertainty predicted for 20 pb^{-1}, in the turn-on region. For higher luminosities it will certainly be worthwhile to shrink the systematics by correcting for known effects.

- **Total systematic uncertainty**

  Tables 4.9 and 4.10 summarize the various contributions to the systematic uncertainties discussed in the previous sections. The total systematic is well controlled within 1% for the plateau region ($p_T > 24$ GeV). It is relatively larger in the trigger turn-on region ($14 < p_T < 20$ GeV). As shown above, it is introduced by the efficiency discrepancy between the true and fake photons.

4.5.5 Results of bootstrap method

Results on efficiency and purity

As detailed in Section 4.4.2, this method requires that at least one photon candidate matches the looser trigger criteria named g10HLTPass. With this requirement, the selected sample is

Uncertainty	photon p_T [GeV/c]					
	$14-16$	$16-18$	$18-20$	$20-24$	$24-30$	>30
Statistical uncertainty [%]	0.50	0.97	0.74	0.37	0.21	0.06
Method systematics [%]	2.44	0.75	1.56	0.46	0.09	0.01
Purity systematics [%]	0.52	0.20	0.15	0.19	0.11	0.03
Total uncertainty [%]	2.54	1.24	1.73	0.62	0.26	0.07

Table 4.9: L1 trigger efficiency uncertainty from the tag&probe method

Uncertainty	photon p_T [GeV/c]					
	$14-16$	$16-18$	$18-20$	$20-24$	$24-30$	>30
Statistical uncertainty [%]	0.10	0.22	1.15	0.66	0.35	0.16
Method systematics [%]	0.11	0.02	3.53	2.54	0.56	0.24
Purity systematics [%]	0.02	0.00	0.30	1.25	0.25	0.01
Total uncertainty [%]	0.15	0.22	3.72	2.91	0.70	0.29

Table 4.10: L2 trigger efficiency uncertainty from the tag&probe method

three orders of magnitude higher than the tag&probe sample. However the estimated photon purity in the turn-on region is only between 6% and 18% as can be seen in Figure 4.11. The largest contribution to the probe sample is from di-jet events as shown in Table 4.11.

Process	photon p_T [GeV/c]		
	<14	$14-20$	>20
$\gamma\gamma$	10 ± 1	630 ± 7	1286 ± 25
γj	3397 ± 195	275755 ± 1296	746727 ± 6056
jj	4398936 ± 69245	3444894 ± 46793	2591702 ± 91250

Table 4.11: Expected number of candidates, in p_T bins, from each physics process, selected with the bootstrap method.

Note that the prescale factor of g10HLTPass trigger is not taken into account for the rate estimates. The typical expected prescale factor is in the range $10-100$. The γj sample used in this study is the dataset with ID=108087, with a single photon filter at 15 GeV, which has relatively high efficiency to the less stringent requirements of the bootstrap selection thus allowing to exploit this unbiased sample in this case. Table 4.12 show the efficiencies and purities obtained with the bootstrap method in several p_T bins, for L1 and L2 respectively. The true γ in the first row of both tables is the reconstructed photon candidate matching a hard process photon or a photon from parton bremsstrahlung in the $\gamma\gamma$, γj, jj combined sample. The topology for the sample selected with the bootstrap method is a single-photon-oriented sample, whereas the sample selected by tag&probe is a di-photon one. It is expected that true photon samples with different topologies have different marginal trigger efficiencies due to the different selection requirements sampling different detector regions. This is what one can observe when comparing the true γ efficiency in Table 4.6 with the true γ efficiency in Table 4.12. As shown in Table 4.12, even though

the purity of the sample selected with the bootstrap method is low especially for the turn-on region, the efficiency for true γ is close to that for fake ones so that the final efficiency measured with the bootstrap method is close to the true one. The results of the bootstrap method determination of the relative trigger efficiency are plotted in Figure 4.12.

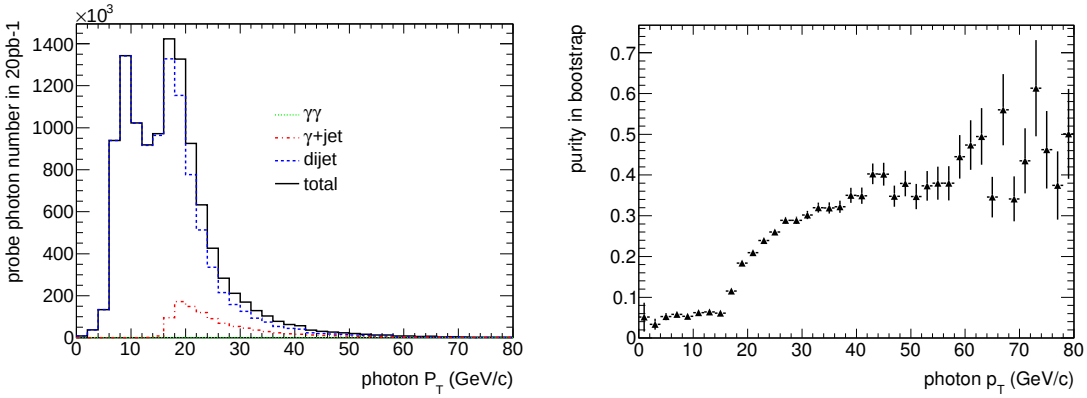


Figure 4.11: Bootstrap sample composition (left) and photon purity (right) as a function of p_T.

4.5.6 Bootstrap systematics

The systematic uncertainties for the bootstrap method are evaluated following the over-conservative procedure used in the tag&probe case. The full difference between the measured efficiency and the efficiency predicted for true photons is taken as systematic error on the efficiency itself. The values of this estimate are reported in Tables 4.13 and 4.14 together with the predicted statistical uncertainty, both scaled to an integrated luminosity of 20 pb^{-1}. These residuals are also plotted in Figure 4.13.

In the case of the bootstrap results, the purity systematics, in the way it was investigated in the tag&probe method, are neglected for two reasons: the fact that the unbiased γ-jet was used in this case and the striking dominance of the di-jet contribution in the sample composition which did not need a parameterization procedure. This two arguments lead to consider purity systematics negligible with respect to what is estimated in tag&probe and the uncertainty on the fake/true photon composition within the dominant di-jet sample is then over-covered by the "method systematics". It is mandatory to rescale this uncertainty according to the g10HLTPass trigger actual prescale factor in the data when it is fixed. It is worthwhile noticing that the bootstrap systematics is limited and lower than or of the same order of the tag&probe one and that even after a harsh prescale, this single photon method is competitive in terms of statistics with the di-photon tag&probe.

4.5.7 Comparison between the tag&probe and bootstrap samples

The photon samples selected in the tag&probe and bootstrap methods are different. As shown in the top plot of Figure 4.14, the p_T spectrum in the bootstrap method is harder than the one of the probes selected with tag&probe. As an example, in the bootstrap method, there is no peak around 5 GeV like in the tag&probe method because the g10PassHLT trigger criteria have been applied to all photon candidates while the probes in tag&probe are not required to be matched with a trigger object. The right plot of Figure 4.14 shows

Quantity	photon p_T [GeV/c]							
	$14-16$	$16-18$	$18-20$	$20-24$	$24-30$	$30-50$	$50-70$	>70
L1								
ϵ(true γ) [%]	6.42 ± 1.29	40.46 ± 1.13	85.42 ± 0.62	97.58 ± 0.03	99.53 ± 0.01	99.70 ± 0.01	99.91 ± 0.01	99.83 ± 0.03
ϵ(bootstrap) [%]	6.27 ± 0.36	37.71 ± 0.57	85.67 ± 0.42	97.11 ± 0.01	99.53 ± 0.01	99.97 ± 0.002	100.00 ± 0.00	99.99 ± 0.004
purity [%]	6.12 ± 0.33	11.53 ± 0.27	18.39 ± 0.32	22.18 ± 0.03	27.55 ± 0.05	33.15 ± 0.06	39.50 ± 0.15	49.16 ± 0.25
L2								
ϵ(true γ) [%]	0.75 ± 0.50	0.83 ± 0.13	30.14 ± 0.71	94.12 ± 0.04	99.20 ± 0.02	99.58 ± 0.01	99.80 ± 0.02	99.78 ± 0.03
ϵ(bootstrap) [%]	0.14 ± 0.05	1.06 ± 0.12	28.69 ± 0.53	92.09 ± 0.02	98.63 ± 0.01	99.37 ± 0.01	99.53 ± 0.02	99.97 ± 0.01
purity [%]	6.12 ± 0.33	11.53 ± 0.27	18.39 ± 0.32	22.18 ± 0.03	27.55 ± 0.05	33.15 ± 0.06	39.50 ± 0.15	49.16 ± 0.25

Table 4.12: L1, L2 trigger efficiency for signal photons in the combined $\gamma\gamma$, γj, jj sample and for the control sample selected with the bootstrap method, and purity of the selected control samples, as a function of p_T.

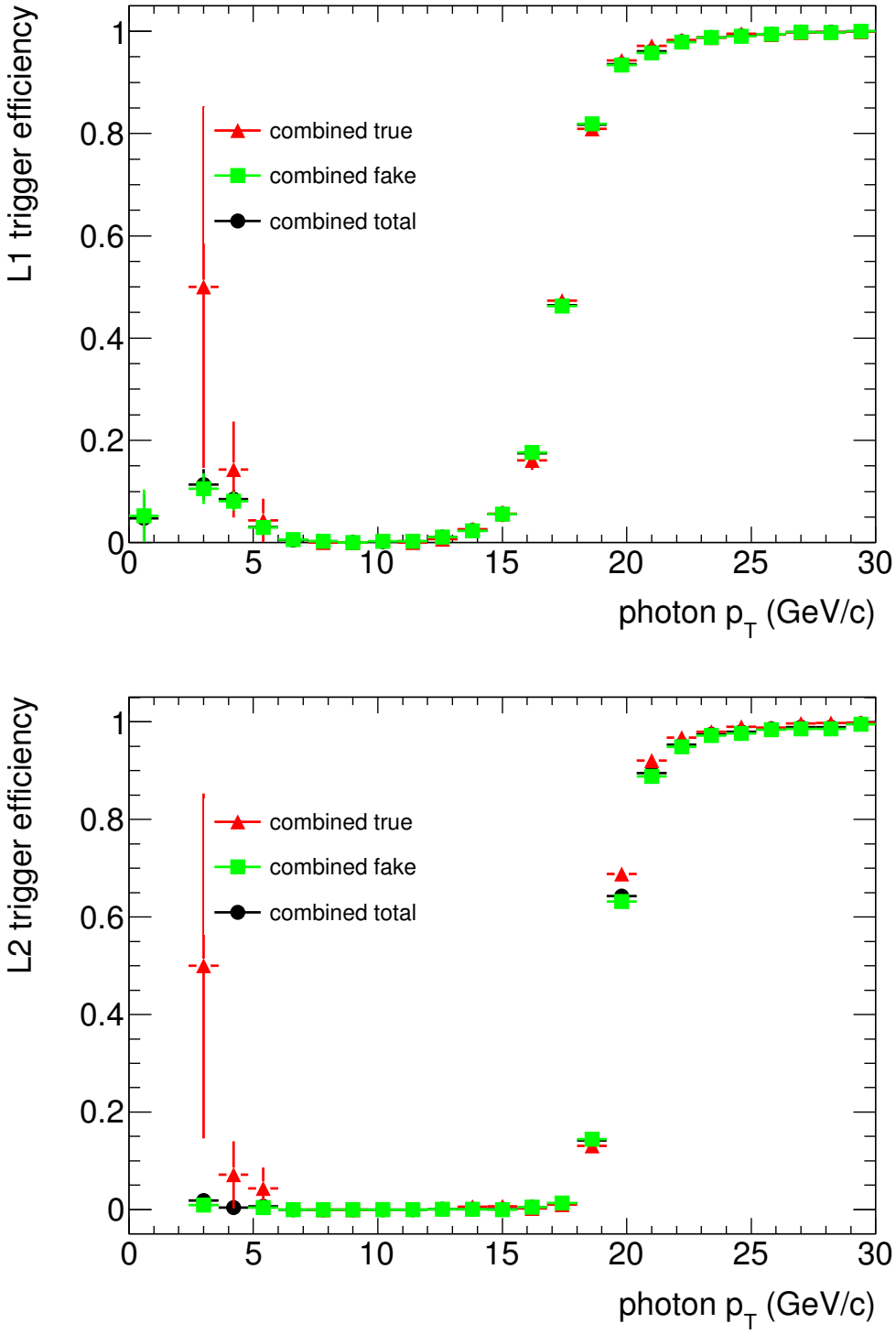


Figure 4.12: *g20_loose* trigger efficiency as a function photon p_T for the combined $\gamma\gamma$, γj, jj data sample with the bootstrap method. Red triangles: true photon efficiency; green squares: fake photon efficiency; black circles: combined (measured) efficiency. Top: L1. Bottom: L2.

Uncertainty	photon p_T [GeV/c]					
	$14-16$	$16-18$	$18-20$	$20-24$	$24-30$	>30
Statistical uncertainty [%]	0.36	0.57	0.42	0.01	0.01	0.002
Method uncertainty [%]	0.15	2.75	0.24	0.47	0.003	0.24
Total [%]	0.38	2.81	0.48	0.47	0.01	0.24

Table 4.13: L1 trigger efficiency uncertainty from the bootstrap method

Uncertainty	photon p_T [GeV/c]					
	$14-16$	$16-18$	$18-20$	$20-24$	$24-30$	>30
Statistical uncertainty [%]	0.05	0.12	0.53	0.02	0.01	0.01
Method uncertainty [%]	0.62	0.22	1.45	2.04	0.57	0.21
Total [%]	0.62	0.25	1.54	2.04	0.57	0.21

Table 4.14: L2 trigger efficiency uncertainty from the bootstrap method

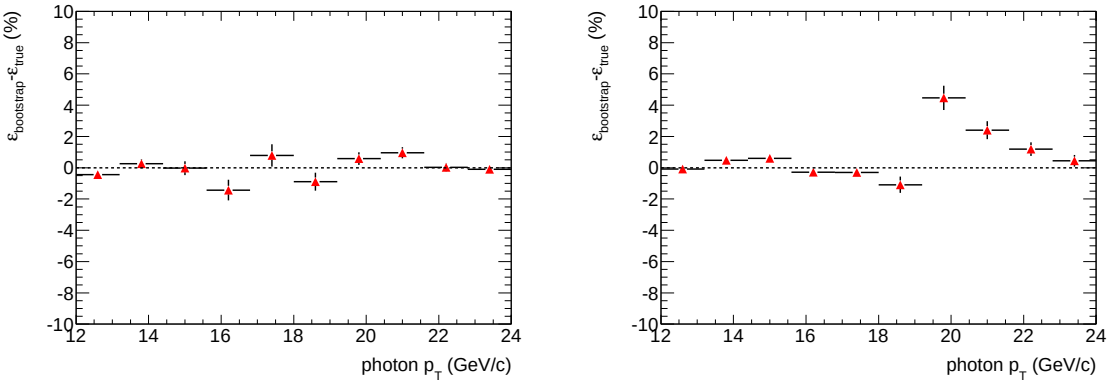


Figure 4.13: Efficiency residuals (in %) between the measurable efficiency with the bootstrap method and the true photon efficiency as a function of p_T; L1 (left) and L2 (right).

a more centrally distributed pseudo-rapidity for the di-photon tag&probe selection due to a bias toward higher event total transverse energy with respect to single photon selection performed in the bootstrap case. Another example is that the tag&probe sample is more enriched in $\gamma\gamma$ events due to the di-photon selection and as a consequence shows a higher photon purity. With more data, one could envisage to re-weight the distributions of photons selected in the tag&probe procedure with the bootstrap one, in order to unfold for possible efficiency differences due to for instance the different pseudo-rapidity spectra.

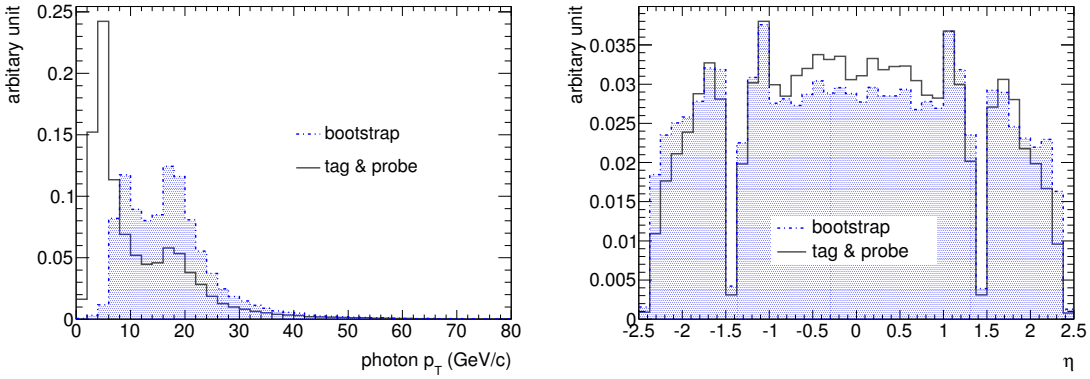


Figure 4.14: Normalized p_T (left) and η (right) distribution in the samples selected by the tag&probe (hollow histogram) and bootstrap (dashed histogram) methods.

4.5.8 The electron to photon extrapolation method

This method takes advantage of lepton-antilepton decays of the narrow Z resonance. Using the selection criteria listed in Section 4.4.3, a pure electron/positron control sample is obtained with a background fraction below 1%. The main issue for this method is the recalibration of the electron/positron showers: electrons, converted photons and unconverted photons cluster energies are measured using different calibration constants as discussed in Section 3.2. Since the photon trigger efficiency is measured as a function of p_T, the transverse momentum p_T of the probe need to be corrected. Figure 4.15 shows the trigger efficiency as a function of the electron/positron p_T (before recalibration). A comparison to the superimposed signal photon efficiency curve (from a combined $\gamma\gamma$, γj and jj sample) shows a large discrepancy between the two curves.

Electron or positron clusters can be re-reconstructed using either converted or unconverted photon calibration constants, depending on the chosen hypothesis. Figure 4.16 shows the measured efficiency as a function of p_T after recalibration with converted (left) and unconverted (right) photon calibration constants. The agreement with the signal γ curve is strongly improved.

4.5.9 Electron extrapolation systematics

The systematic uncertainty for the electron extrapolation method is again estimated by taking the full difference between measured efficiency and the efficiency predicted for true photons. The values of this estimate are reported in Tables 4.15 and 4.16 together with the predicted statistical uncertainty, both are scaled to an integrated luminosity of 20 pb^{-1}. These residuals are also plotted in Figure 4.17. There are two important issues that were

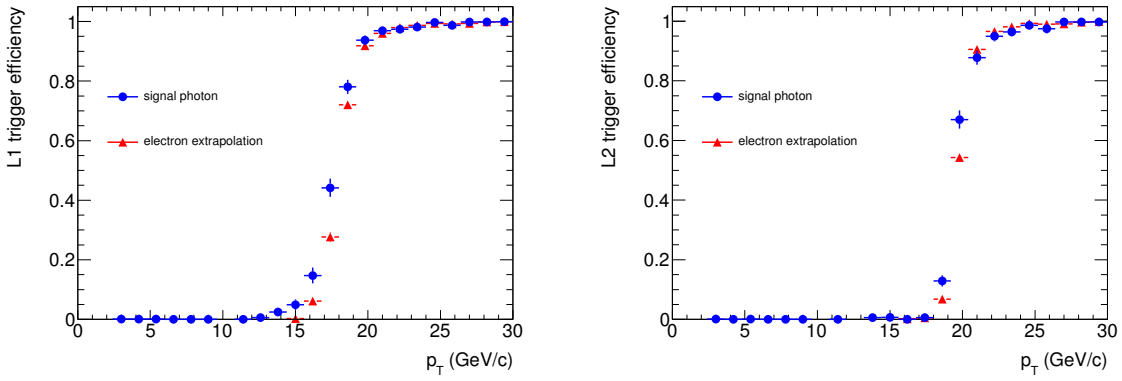


Figure 4.15: *g20_loose* efficiency as a function of the electron transverse momentum p_T, superimposed with signal photon efficiency from $\gamma\gamma$, γj, jj combined sample at L1 (left) and L2 (right).

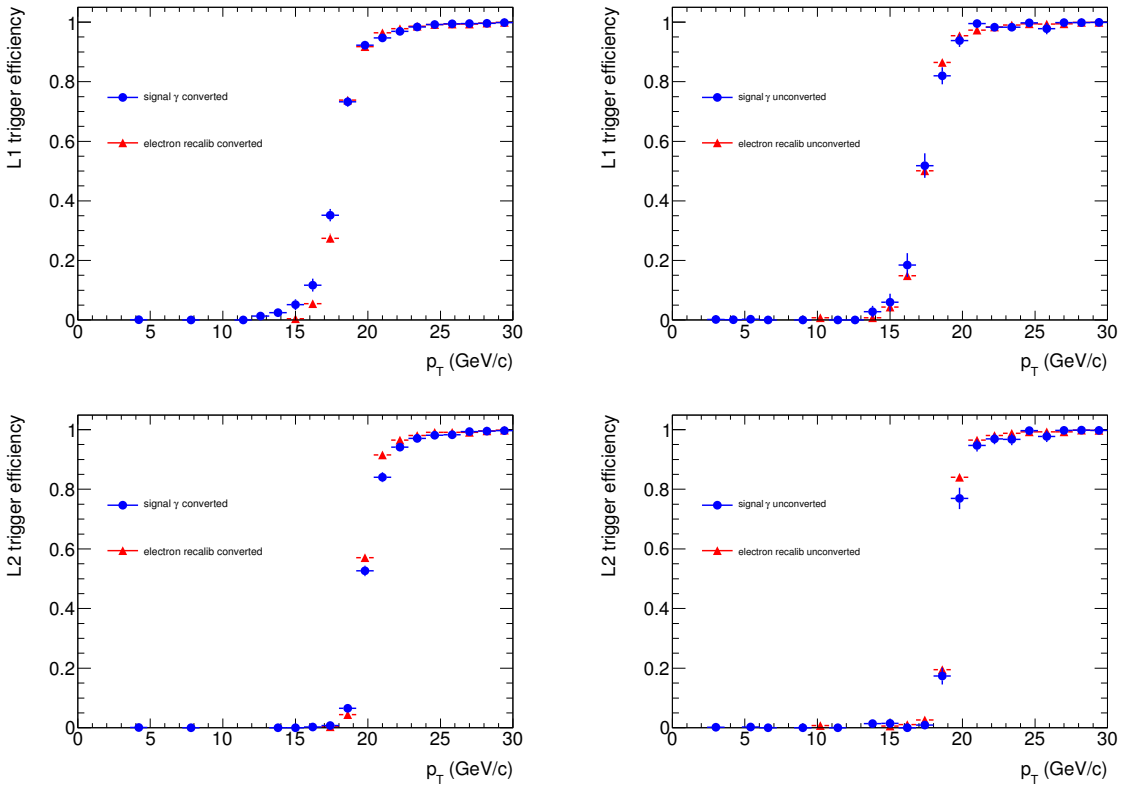


Figure 4.16: g20_loose efficiency as a function of the transverse momentum with electron extrapolation method, superimposed signal photon efficiency from $\gamma\gamma$, γj, jj combined sample; Top: L1, Bottom: L2. Electron recalibrated as converted photon and compared with signal converted photon (left) and electron recalibrated as unconverted photon and compared with signal unconverted photon (right).

not considered in this evaluation and that can potentially reduce the observed discrepancies: the shower shapes corrections and the kinematic of the Z sample to be considered while porting the efficiency extracted from the Z electrons to the di-photon sample.

Uncertainty	Photon p_T [GeV/c]					
	$14-16$	$16-18$	$18-20$	$20-24$	$24-30$	>30
Statistical [%]	1.84	3.27	1.69	0.26	0.06	0.00
Method systematics [%]	3.88	8.67	1.30	0.44	0.14	0.14
Total uncertainty [%]	4.30	9.26	2.13	0.51	0.15	0.14

Table 4.15: Uncertainties (in absolute %) on the L1 trigger efficiency from the electron-extrapolation method with 20 pb^{-1}

Uncertainty	Photon p_T [GeV/c]					
	$14-16$	$16-18$	$18-20$	$20-24$	$24-30$	>30
Statistical [%]	0.50	0.82	2.39	0.35	0.06	0.00
Method systematics [%]	0.58	0.57	5.82	2.58	0.31	0.14
Total [%]	0.76	1.00	6.30	2.60	0.32	0.14

Table 4.16: Uncertainties (in absolute %) on the L2 trigger efficiency from the electron-extrapolation method with 20 pb^{-1}

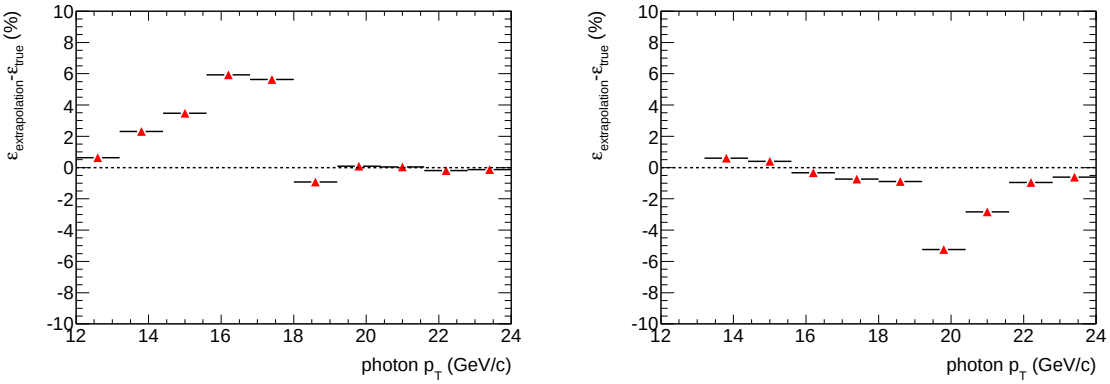


Figure 4.17: Residuals (in %) between the efficiency measured with the electron-extrapolation method and the true photon efficiency as a function of p_T: L1 (left) and L2 (right).

4.5.10 Comparison of the three data-driven methods

From the results shown in the previous sections, it can be noticed that the three methods show an overall good agreement with the true photon efficiency, at least with the data statistics expected with an integrated luminosity of 20 pb^{-1}. Nevertheless some discrepancies are observed, in particular in the case of the electron-extrapolation method. A review of the

results, together with the plans for the application of these methods to the early ATLAS data, are given in the following.

- **Trigger efficiency determination biases**

  In Figure 4.18, the L1 and L2 efficiency turn-on curves as measured with the three data-driven methods, together with the true photon efficiency, are compared. In Tables 4.17 and 4.18, the trigger efficiency differences between the three methods are given. They are compared with the total uncertainty of the tag&probe which is taken here as a reference.

Measured quantity	Photon p_T [GeV/c]					
	$14-16$	$16-18$	$18-20$	$20-24$	$24-30$	>30
$\epsilon_{tag\&probe} - \epsilon_{bootstrap}$ [%]	1.35	3.06	0.44	0.36	0.19	0.04
$\epsilon_{tag\&probe} - \epsilon_{extrapolation}$ [%]	2.38	2.85	1.50	1.28	0.33	0.06
Total uncertainty in tag&probe [%]	2.54	1.24	1.73	0.62	0.26	0.07

Table 4.17: Absolute L1 trigger efficiency differences (in %) between the three methods.

Measured quantity	Photon p_T [GeV/c]					
	$14-16$	$16-18$	$18-20$	$20-24$	$24-30$	>30
$\epsilon_{tag\&probe} - \epsilon_{bootstrap}$ [%]	0.02	0.27	0.29	0.70	1.01	0.49
$\epsilon_{tag\&probe} - \epsilon_{extrapolation}$ [%]	0.06	0.62	7.57	5.31	1.90	0.85
Total uncertainty in tag&probe [%]	0.15	0.22	3.72	2.91	0.70	0.29

Table 4.18: Absolute L2 trigger efficiency differences (in %) between the three methods.

For the signal photon curve, photon candidates are required to be matched to true photons originating from the hard process or parton bremsstrahlung in the di-jet, photon-jet and di-photon combined sample. In general, across the whole p_T range and especially in the high p_T region, the difference between the three methods is well below the separation power given by the data statistics. In the turn-on region around the trigger threshold, some discrepancies are observed. In particular, in the steep rise region, the tag&probe and the bootstrap methods are in good mutual agreement and are close to the true photon efficiency. As reported in Sections 4.5.4 and 4.5.6, the systematic uncertainties of the two methods are evaluated to be of the same order of the statistical uncertainty and their biases are statistically not significant. For the electron-extrapolation method, the measured efficiency appears somewhat more biased with respect to the expected true photon efficiency on one side and with respect to the other two methods on the other side. This is reflected in Tables 4.17 and 4.18, where the differences of the efficiencies obtained with the tag&probe and the bootstrap methods are within the tag&probe total uncertainty while the difference between the efficiencies obtained with the tag&probe and the electron-extrapolation methods is higher in the turn-on region and also above 20 GeV.

- **Data samples statistics and trigger efficiency uncertainty**

  In Table 4.19, the expected number of photon candidates for 20 pb^{-1} for each method is reported. It appears that, in the turn-on region, the electron extrapolation method starts to be a valuable method for estimating the relative trigger efficiency for photons with much higher integrated luminosities than the reference one of 20 pb^{-1} used in this study. The predicted statistics for this luminosity is in fact very poor compared to those expected for the tag&probe and the bootstrap methods in this region, while it is comparable to the tag&probe method at higher transverse momenta.

Method	Photon p_T [GeV/c]		
	< 14	$14 - 20$	> 20
tag&probe	24372 ± 387	5556 ± 129	5489 ± 223
bootstrap	$(44.0 \pm 0.3)10^5$	$(37.2 \pm 0.5)10^5$	$(33.4 \pm 1.0)10^5$
extrapolation	32 ± 7	136 ± 21	6302 ± 93

Table 4.19: Estimated number of photon candidates expected for each method as a function of p_T with an integrated luminosity of 20 pb^{-1}.

Method	Photon p_T [GeV/c]					
	$14 - 16$	$16 - 18$	$18 - 20$	$20 - 24$	$24 - 30$	> 30
tag&probe	2.54	1.24	1.73	0.62	0.26	0.07
bootstrap	0.38	2.81	0.48	0.47	0.01	0.24
extrapolation	4.30	9.26	2.13	0.51	0.15	0.14

Table 4.20: Total uncertainty on the L1 photon-trigger efficiency (in absolute %) for the three methods with an integrated luminosity of 20 pb^{-1}.

Method	Photon p_T [GeV/c]					
	$14 - 16$	$16 - 18$	$18 - 20$	$20 - 24$	$24 - 30$	> 30
tag&probe	0.15	0.22	3.72	2.91	0.70	0.29
bootstrap	0.62	0.25	1.54	2.04	0.57	0.21
extrapolation	0.76	1.00	6.30	2.60	0.32	0.14

Table 4.21: Total uncertainty on the L2 photon-trigger efficiency (in absolute %) for the three methods with an integrated luminosity of 20 pb^{-1}.

In Tables 4.20 and 4.21, the total uncertainties summing up the statistical and systematic uncertainties for each of the three data-driven methods are given in the different p_T bins, for L1 and L2 photon-trigger efficiencies respectively. The total uncertainties, at least in the high p_T region above 20 GeV, are dominated by the method systematics and suffer from large bin to bin fluctuations. Considering the precision on the uncertainties quoted in Tables 4.20 and 4.21, one can say that the tag&probe and bootstrap methods give good results in the turn-on region, while in the plateau region (above 20 GeV), the three methods give similar results.

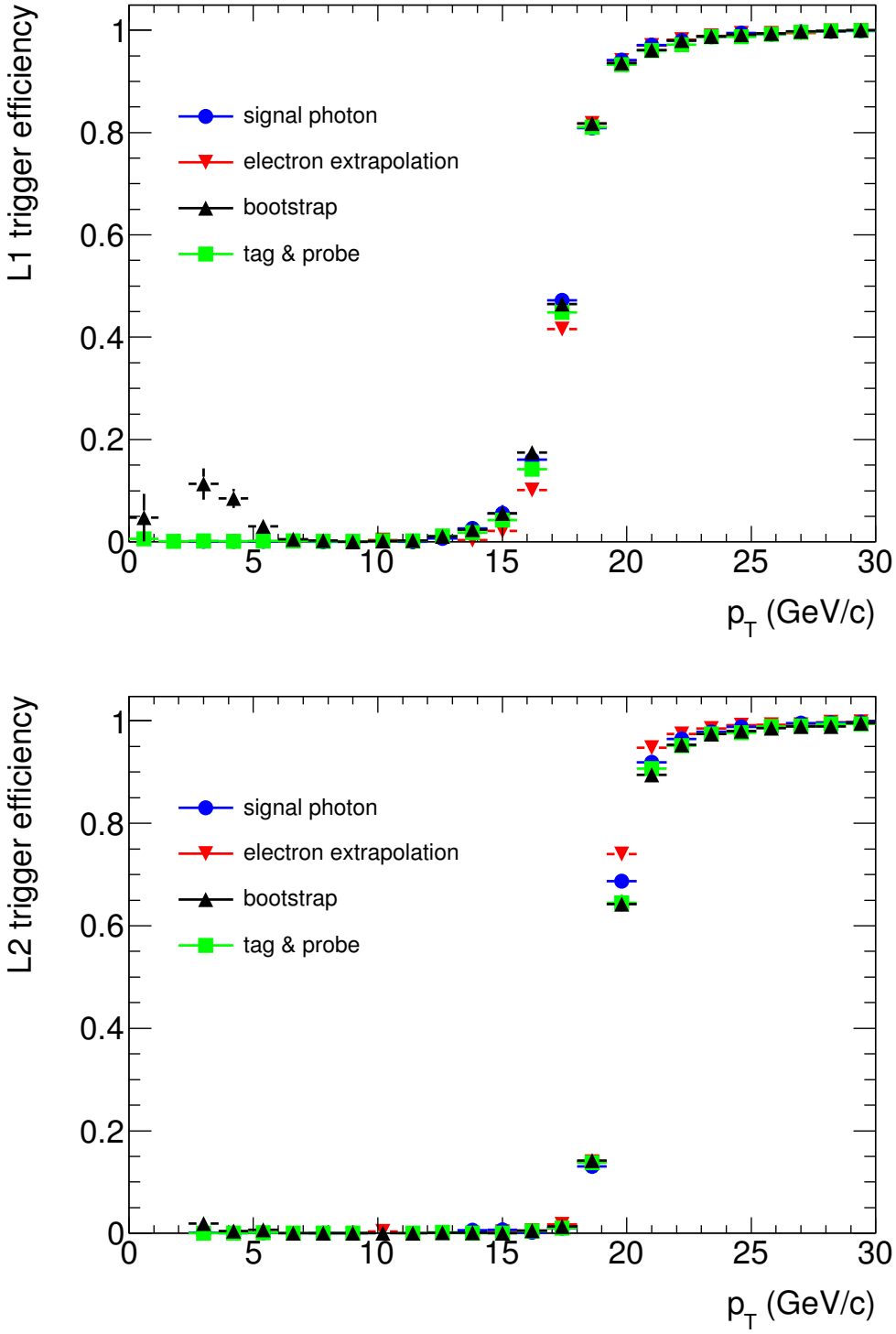


Figure 4.18: Comparison of the trigger marginal efficiency for signal photons (blue circles) and for photon candidates selected with the tag& probe (green squares), the bootstrap (black triangles) and the electron extrapolation (red triangles) methods. Top: L1. Bottom: L2.

4.5.11 Conclusion

A feasibility study of three data-driven methods for the extraction of the relative photon-trigger efficiency was performed at simulation level. The accuracy that can be achieved by each of the three methods has been predicted for an integrated luminosity of 20 pb^{-1}. A comparison between the three methods and with respect to the true photon efficiency, as expected from Monte Carlo, has been presented. The two direct-photon-based methods (tag&probe and bootstrap) agree with each other within the estimated uncertainties while the extrapolation to photons of the electron sample selected in the $Z \rightarrow e^+e^-$ study shows larger biases in particular in the trigger turn-on region around the trigger p_T threshold at 20 GeV. In the high p_T region ($p_T > 25$ GeV) all of the three methods agree with each other, essentially because in this p_T region the photon trigger efficiency is expected to be almost 100% for both true and fake photons. The total uncertainties of these data-driven methods for the p_T region above turn-on, which is relevant for the main photon-based ATLAS analyses, have been estimated in this work to be well below 1% absolute, already with an integrated luminosity of 20 pb^{-1}. In the lower p_T region around the trigger turn-on, an uncertainty of a few % is expected in the data-driven determination of the photon-trigger efficiency thus allowing a fair check of the trigger performances.

4.6 Efficiency measurement on real data

As stressed before, the trigger efficiency measured here is a marginal trigger efficiency. Photon selection needs to be applied first. In real data, the tight photon selection criteria includes:

- Photons are required to pass the **RobustTight** selection defined in Section 3.3.
- Crack region rejection: $(|\eta| < 1.37)$ || $(1.52 < |\eta| < 2.37)$;
- Calorimetric isolation, as defined in Section 5.3.1: $E_T^{iso} < 3$ GeV.

4.6.1 g20_loose efficiency measurement

Even though around 45 pb^{-1} of data were recorded by ATLAS during the 2010 data taking period, only a small subset of it can be used for all the three methods described above. For the tag&probe method, g20_loose is used for the event preselection, since at least the tag photon is required to pass g20_loose trigger. However, due to the fast evolution of the instantaneous luminosity during the data-taking, g20_loose was highly prescaled. For the bootstrap method, e10_NoCut is used as the lower threshold unbiased calorimetric trigger, which is actually a L1 calorimetric trigger L1_EM5 also highly prescaled. This trigger is used as the event preselection for the bootstrap method. In addition, g20_loose is required to be unprescaled. Otherwise the events that labelled as "failed" actually due to the prescaling may enter the probe sample, which introduces the efficiency loss. Consequently, the corresponding luminosity for the selected events is even smaller than for the tag&probe method. As for the electron-to-photon extrapolation method, e10_medium (or e15_medium when e10_medium starts to be prescaled) is used for the electron sample selection. Besides, g20_loose is also required to be unprescaled. All the relevant selection in each trigger item is summarized in Table 4.2. Table 4.22 summarizes the corresponding luminosity after the event preselection on the trigger for each method.

For more information, Table 4.23 shows the list of photon related triggers which are not prescaled over all the periods of 2010 data taking.

method	trigger requirement	Luminosity $[nb^{-1}]$	N_{probe}
tag&probe	g20_loose	7023.93	3792 (L1) 3697 (L2) 3666 (EF)
bootstrap	e10_NoCut and g20_loose unprescaled	141.188	23766
extrapolation	e10_medium (or e15_medium) and g20_loose unprescaled	3520.7	1162 (unconv) 1986 (conv)

Table 4.22: Corresponding luminosity and the statistics for the probe sample after the event preselection on the trigger for the three methods.

trigger items	L1 item	L2 item	EF item
2g15_loose	L1_2EM10	L2_2g15_loose	EF_2g15_loose
g20_loose_xe20_noMu	L1_EM14_XE10	L2_g20_loose_xe20_noMu	EF_g20_loose_xe20_noMu
g20_loose_xe30_noMu	L1_EM14_XE15	L2_g20_loose_xe30_noMu	EF_g20_loose_xe30_noMu
g40_loose	L1_EM14	L2_g40_loose	EF_g40_loose
g50_loose	L1_EM14	L2_g50_loose	EF_g50_loose
g40_tight	L1_EM14	L2_g40_tight	EF_g40_tight

Table 4.23: Unprescaled photon triggers in 2010 data-taking.

The p_T spectrum of the photon candidates selected to measure the photon trigger efficiency is shown in Figure 4.19. The corresponding luminosity for the bootstrap method after the preselection on the trigger is small, the selection of the probe is based on the requirement of only one tight and isolated photon is required for each event. Therefore, more photon candidates can be selected with this method. On the contrary, for the tag&probe method, the corresponding luminosity after the preselection is higher. However, since at least two tight and isolated photon candidates are required for each event, the final statistics is low. The electron extrapolation method has the lowest statistics. It is limited by the relatively small production cross-section for the $Z \rightarrow ee$ events compared with the QCD processes.

Figure 4.20 shows the measured g20_loose photon trigger efficiency for the tight and isolated photons as a function of the transverse momentum, using 2010 data. The results from all three methods are overlaid, only statistical uncertainty is shown. The turn-on curves, show that the g20_loose trigger works efficiently for tight and isolated photons. The photon candidate, with transverse momentum above 22 GeV are already in the plateau region. However, for the electron-to-photon extrapolation method, very low statistics is obtained in the photon p_T region below 20 GeV, thus giving no insight to the turn-on region. As expected from the Monte Carlo studies, tag&probe and bootstrap methods give consistent results within the statistical uncertainty, which validate the robustness for each of the methods.

4.6.2 g10_loose efficiency for inclusive photon cross-section measurement

Based on the early data recorded in 2010, a first inclusive photon cross-section measurement was performed [90]. In this measurement, the photon candidates with transverse energy above 15 GeV are considered, with g10_loose as the trigger selection. The g10_loose trigger efficiency is measured with the bootstrap method. This is the first time that this method has been applied in a photon-related physics measurement using the real data recorded by ATLAS experiment.

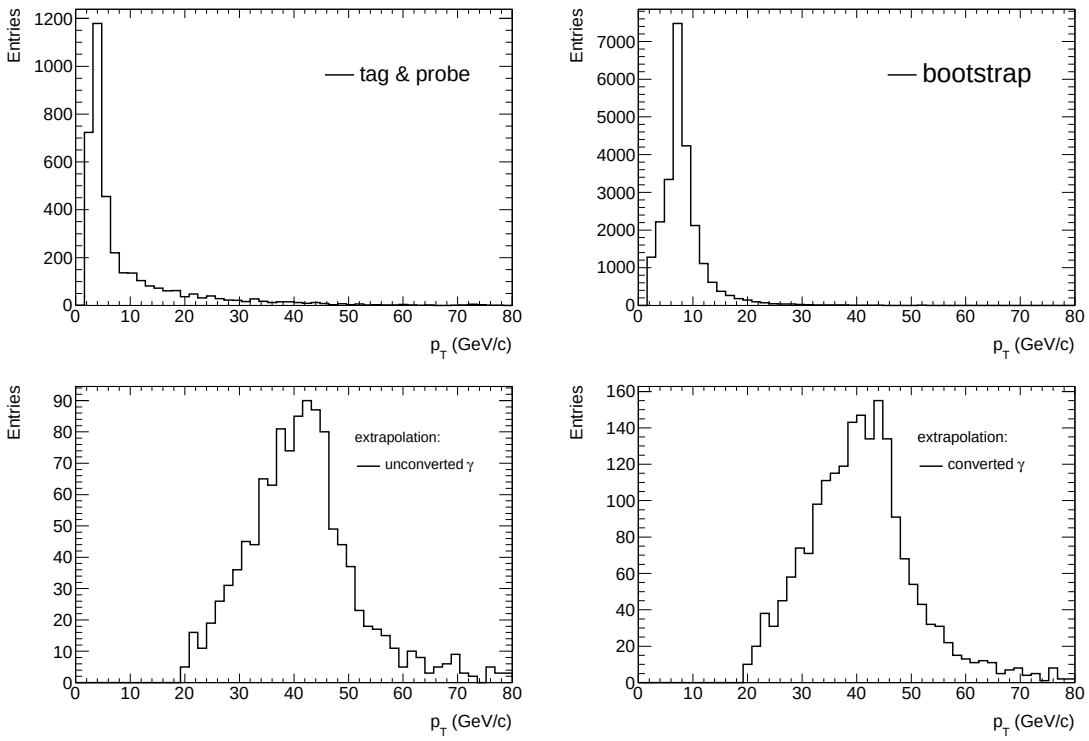


Figure 4.19: p_T spectrum for the photon candidates selected in each method. Top left: tag&probe, Top right: bootstrap, Bottom left: extrapolation for unconverted photon, Bottom right: extrapolation for converted photon.

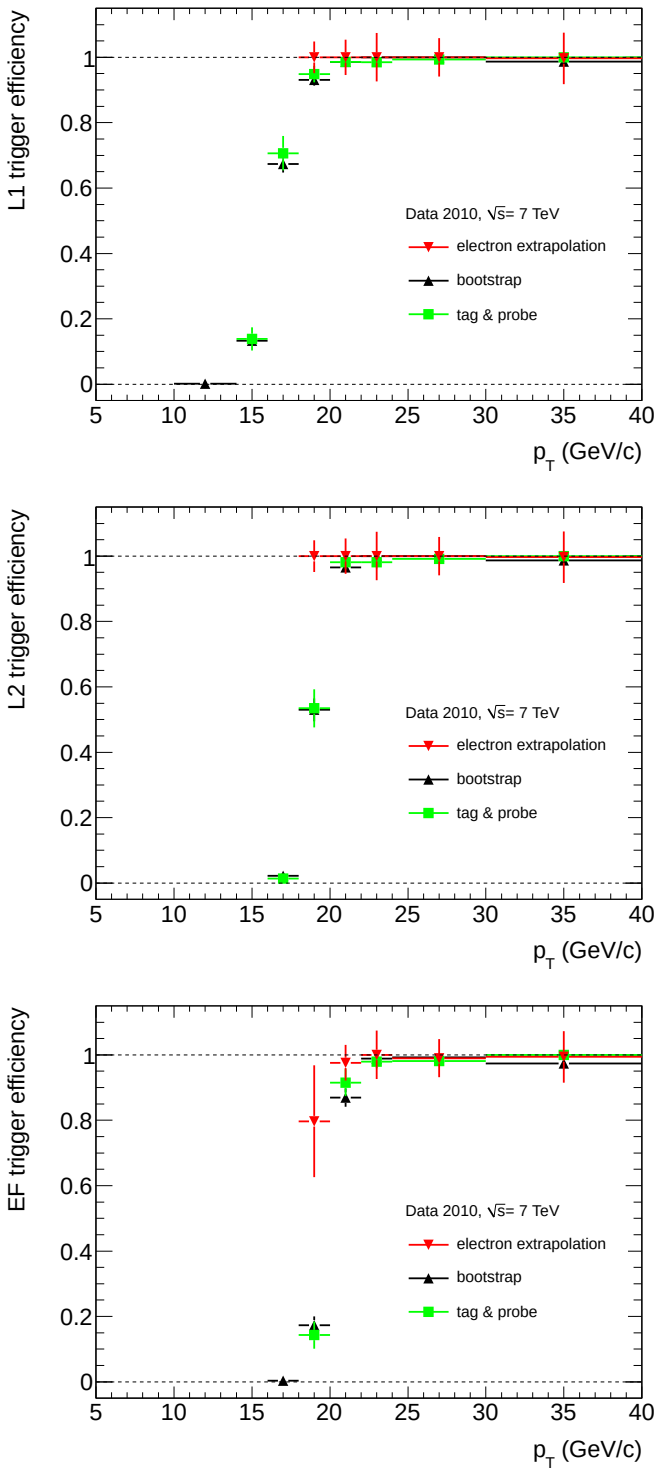


Figure 4.20: g20_loose trigger efficiency measured on 2010 data for tight and isolated photons as a function of the transverse momentum for the three levels of g20_loose. Top: L1; Middle: L2; Bottom: Event Filter level.

In practice, it is performed with two steps. For the first step, a lower threshold L1 calorimeter trigger (L1_EM2) is checked with a sample of minimum-bias triggers.[2] The measured efficiency of this trigger is 100% for all photon candidates passing the tight identification and isolation requirement with transverse energy $E_T > 15$ GeV. In the second step, the g10_loose efficiency is checked with events passing the e3_NoCut (which is actually L1_EM2 as shown in Table 4.2). The trigger efficiency for the photon candidates passing the tight identification and isolation requirement with transverse energy $E_T > 15$ GeV can then be determined by the second step, which is measured to be $\varepsilon^{trig} = (99.5\pm0.2)\%$. The g10_loose efficiencies for each trigger level in the different η ranges are shown in Table 4.24, which is almost constant over all η regions. The efficiency for g10_loose is also checked using Monte Carlo samples. It is found to be 99.75±0.07% which is consistent with the result measured in real data. The absolute difference of the trigger efficiency for a pure signal sample and for a pure background sample is found to be smaller than 0.5%. A comparison on the trigger efficiency curve between the real data and the simulated background sample can be found in Figure 4.21. Consistent results is observed in the comparison.

	$\|\eta\|$ ranges				
	$0-0.6$	$0.6-1.37$	$1.52-1.81$	$1.81-2.37$	all
L1	100.0±0.0	99.92±0.08	99.62±0.27	99.84±0.16	99.87±0.06
L2	100.0±0.0	99.83±0.12	99.62±0.27	99.68±0.22	99.81±0.08
EF	99.88±0.12	99.32±0.24	99.42±0.33	99.52±0.27	99.53±0.15

Table 4.24: The measured g10_loose efficiency at each trigger level for reconstructed photon candidates passing the tight identification and isolation requirement and with transverse energy $E_T > 15$ GeV.

[2]The information is provided by the Minimum Bias Trigger Scintillators (MBTS). They are installed in the region of $2.09 < |\eta| < 3.84$ and $z = 3.6$ m between Inner Detector and EM endcap calorimeter. The events fulfilling the criteria for good collision candidates are considered to pass the triggers.

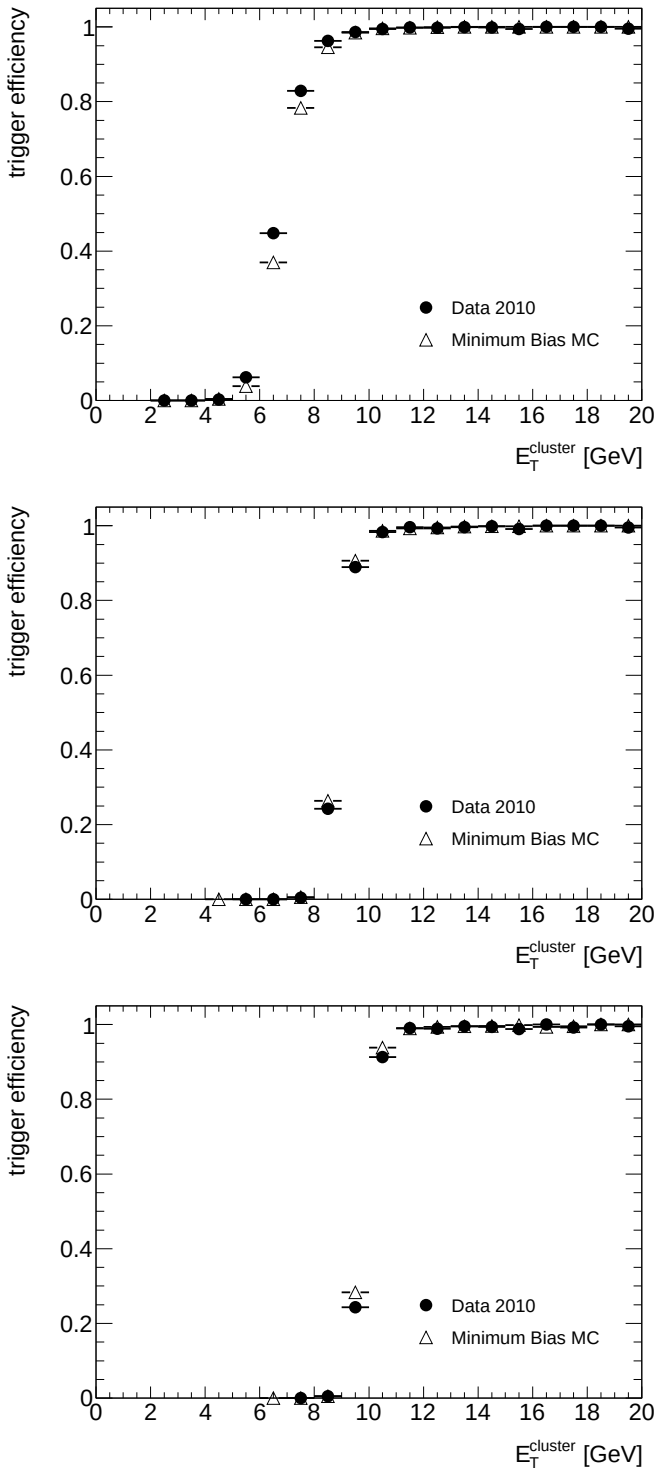


Figure 4.21: g10_loose trigger efficiency measured on 2010 data for tight and isolated photons as a function of the transverse momentum for the three levels of g10_loose. Top: L1; Middle: L2; Bottom: Event Filter level.

Chapter 5

Isolated di-photon cross-section measurement

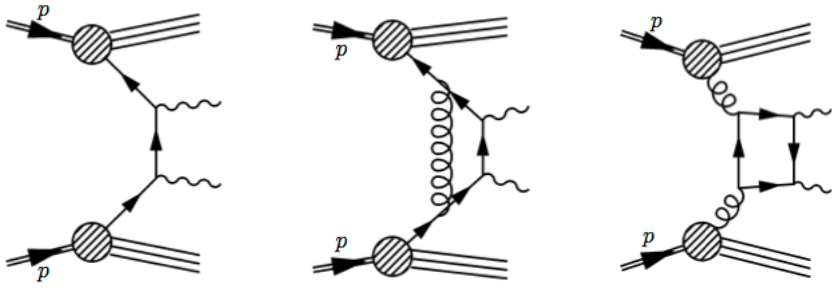
In this chapter, a measurement of the isolated di-photon cross-section in proton-proton collisions at 7 TeV is presented. Results are compared with the Standard Model prediction.

5.1 Di-photon production and background processes

5.1.1 signal processes

In the Stand Model, there are mainly three categories of processes that contribute to the di-photon production [80, 81]: two direct photons, single bremsstrahlung/fragmentation and double bremsstr- ahlung/fragmentation.

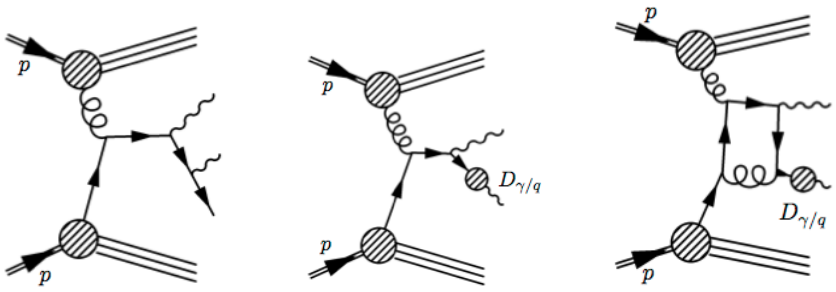
The two-direct-photon process occurs mainly through quark-antiquark annihilation. The first diagram in Figure 5.1 shows the leading order contribution from the Born process ($q\overline{q} \rightarrow \gamma\gamma$) and one example of next-to-leading order (NLO) contribution is shown in the second diagram. The third diagram is the so-called box process where gluon-gluon scattering leads to two direct photons via a quark loop. The Born process is of the order of $\mathcal{O}(\alpha_{em}^2)$ while the box process is suppressed by an extra $\mathcal{O}(\alpha_s^2)$ factor. However, it is enhanced by the high gluon density of the proton-proton collisions at LHC. Therefore, these two contributions are comparable.



(a) (b) (c)

Figure 5.1: Examples of diagrams for the di-photon production in which both photons are direct. See reference [81].

The single bremsstrahlung/fragmentation process, gq (or $\overline{q}$)$\rightarrow \gamma\gamma q$ (or $\overline{q}$), corresponds to the production of one direct photon and one photon from the bremsstrahlung or fragmentation of a hard parton produced in the short distance subprocess. Due to the collinearity with the parton, the bremsstrahlung/fragmentation photon is expected to be not well isolated from the experimental point of view. Figure 5.2(a) shows an example of the single bremsstrahlung process. It is of order of $\mathcal{O}(\alpha_s\alpha_{em}^2)$ and therefore can be considered as NLO correction to the Born process but is also favored by the high gluon density of the LHC. The process of single fragmentation is shown in Figure 5.2(b). An example of NLO correction to this process is shown in Figure 5.2(c). Divergences arise in the limit where the photon is collinear to the parton. Quark or gluon fragmentation functions, $D_{\gamma/q\ or\ g}(z, M_f^2)$, are therefore introduced to overcome this technical difficulty. They describe the showering of a parton during which a collinear photon is emitted, absorbing the divergences and some higher order corrections. z is the fraction of the parton momentum transferred to the photon. M_f sets the scale of the corrections absorbed in the fragmentation function, while the remaining corrections must be calculated. It is not a physical parameter. The separation between the bremsstrahlung emission and the fragmentation is arbitrary and only the sum of the two processes has a physical meaning. For convenience, the bremsstrahlung emission and the fragmentation are collectively called fragmentation in the following.



(a) (b) (c)

Figure 5.2: Examples of diagrams for the di-photon production where one photon is produced via fragmentation.

The process of double fragmentation, with both photons resulting from the fragmentation, can also make a significant contribution. Figure 5.3(a) shows an example of the leading order contribution for double fragmentation. Figure 5.3(b), (c) and (d) are examples of NLO corrections.

For the cross-section measurement in this chapter, the photons are required to be isolated, so as to suppress the background (described later in Section 5.1.2), but also to reduce the fragmentation part which is less precisely known.

The real data are compared with the following computations:

- DIPHOX [83]: it is fixed-order parton level generator. All the processes discussed above, including the two direct photons, single fragmentation and double fragmentation, are computed at NLO. There is only one exception for the box process, which is at LO. Thanks to the NLO calculation of the single and double fragmentation, DIPHOX is expected to give a rather accurate description of the azimuthal distance between the two photons ($\Delta\phi_{\gamma\gamma}$) at low values, and of the transverse momentum of the photon pair ($p_{T\gamma\gamma}$) in the large $p_{T\gamma\gamma}$region.

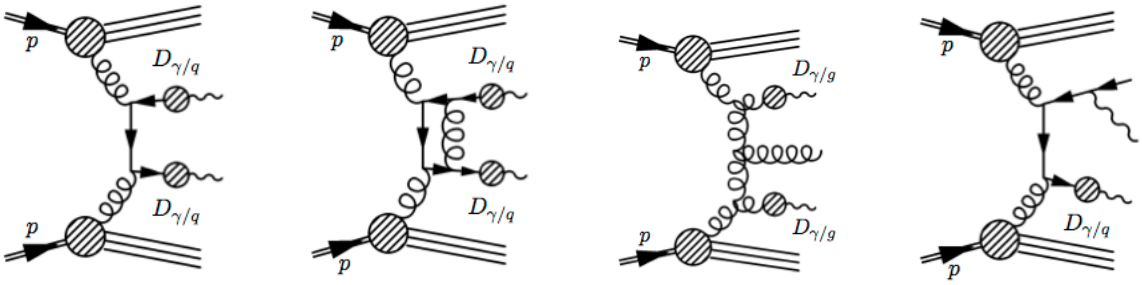


Figure 5.3: Examples of diagrams for the di-photon production where both photons are produced at fragmentation.

- RESBOS [80]: it is also a parton level generator, at NLO for the total cross-section. It includes the two direct photons process, such as the Born and box processes. It also includes the single fragmentation processes, which is only computed at LO. Besides, the next-to-leading logarithms (NLL) resummation formalism which absorbs the initial-state singularities due to the emission of soft or collinear gluons, is applied. This resummation formalism helps to give an accurate description of the lower part of the $p_{T_{\gamma\gamma}}$ spectrum.

5.1.2 Background process

The backgrounds for both the di-photon analysis and the Higgs search ($H \to \gamma\gamma$) mainly come from the production of a photon and a jet or the production of multi-jets, namely photon-jet and di-jet events in which jet(s) are mis-identified as photon(s). After photon identification and isolation, these backgrounds can be greatly reduced. The remaining jets mis-identified as photons are mainly jets that fragment into a leading π^0 or η meson.

Isolated electrons may be wrongly identified as photons since they are similar from the calorimetric point of view. Thus, the physics processes to ee final states (Drell-Yan, $Z \to ee$, $WW \to e\nu e\nu$) and to $e\gamma$ final states ($\gamma W \to \gamma e\nu$, $\gamma Z \to \gamma ee$) can contribute to the di-photon and $H \to \gamma\gamma$ background.

5.2 Data and Monte Carlo samples

In this analysis, the data collected from the $\sqrt{s} = 7$ TeV proton-proton collisions in 2010, periods E to I, are used. It corresponds to an integrated luminosity of $37.2 \pm 1.3\ pb^{-1}$. The data periods A to D are not used because the di-photon triggers in these periods were not enabled. The corresponding luminosity for each period from E to I requiring 2g15_loose trigger can be found in Table 5.1.

The Monte Carlo (MC) samples used in this analysis are summarized in Table 5.2, which includes the dataset ID, process type, cross-section, event-filter, the corresponding filter efficiency, the number of reconstructed events and the equivalent luminosity $\mathcal{L}$ for each sample. All these samples were generated with PYTHIA 6.4 by the ATLAS MC production group. The label convention for γj is different from the one in Chapter 4. Here, γj events are photon-jet events with the reconstructed leading p_T photon candidate being a

period	Run range	Integrated luminosity $[nb^{-1}]$
E	160387 - 161948	985
F	162347 - 162882	1749
G	165591 - 166383	6686
H	166466 - 166964	7063
I	167575 - 167844	20765

Table 5.1: The collected integrated luminosity for each period of data taking.

true photon, and the subleading p_T photon candidate being a jet, and vice-versa for $j\gamma$. In Chapter 4, γj meant for all the photon-jet events. The detailed information for the filters applied in these samples is listed below:

- SymPhotonFilter: at least 2 photons with $p_T > 14$ GeV ($p_T > 10$ GeV for the $\gamma\gamma$ sample with Dataset ID 105961), $|\eta| < 2.7$
- PhotonFilter: at least 1 photon with $p_T > 17$ GeV, $|\eta| < 2.7$
- SymJetFilter: at least 2 jets with $p_T > 10$ GeV, $|\eta| < 2.7$

The bunch-train pileup in Table 5.2 is setup as: double trains with 225 ns separation, which is equivalent to the duration of 9 bunches crossing. Within each train, 8 filled bunches are included with 150 ns bunch separation. The average number of minimum-bias interactions is 2.2. For systematics studies, several samples are also generated with additional material in the Inner Detector and the calorimeter and the events are then reconstructed with the nominal geometry. The extra material is as follows:

- 5% of the whole Inner Detector
- 20% of the pixel services
- 20% of the SCT services
- 15% of radiation length material at end of SCT/TRT endcaps
- 15% of radiation length material at the Inner Detector endplate
- 20% of radiation length for the region of $1.5 < |\eta| < 4.5$ in the calorimeter

5.3 Photon and event selection

For reconstructed photon candidates, the **RobustTight** (**T**) criterion as defined in Section 3.3, is applied to reject large contamination from jets.

5.3.1 Photon isolation

A calorimetric isolation is applied besides the cuts on the discriminating shower shape variables in order to further reject the jet background. The calorimetric isolation E_T^{iso} is defined as the sum of the transverse energy of the cells in ECAL and HCAL inside a

Dataset ID	Process	N_{evts}	σ $[nb]$	Filter	Filter Eff	Lumin $[nb^{-1}]$
Bunch-train pileup, nominal geometry						
105961	$\gamma\gamma$	499938	1.503	SymPhotonFilter	0.22483	1.479E+6
119065	$\gamma\gamma$+(γj+$j\gamma$)	49995	445.87	SymPhotonFilter	0.000511	2.194E+5
108100	γj+$j\gamma$	4696242	1438.4	SymJetFilter	0.050991	6.403E+4
115802	$\gamma\gamma$+(γj+$j\gamma$)+jj	996859	1147100	PhotonFilter	0.00018122	4.795E+3
106043	$W \rightarrow e\nu$	6989096	8.9374	-	1	7.820E+5
106046	$Z \rightarrow ee$	4995076	0.85521	-	1	5.841E+6
Bunch-train pileup, distorted geometry (extra material)						
119065	$\gamma\gamma$+(γj+$j\gamma$)	49994	445.87	SymPhotonFilter	0.000511	2.194E+5
106043	$W \rightarrow e\nu$	998879	8.9374	-	1	1.118E+5
106046	$Z \rightarrow ee$	998876	0.85521	-	1	1.168E+6
No pileup						
119065	$\gamma\gamma$+(γj+$j\gamma$)	49995	445.87	SymPhotonFilter	0.000511	2.194E+5
108100	γj+$j\gamma$	498919	1438.4	SymJetFilter	0.050991	6.802E+3
106043	$W \rightarrow e\nu$	699916	8.9374	-	1	7.831E+4
106046	$Z \rightarrow ee$	499938	0.85521	-	1	5.846E+5

Table 5.2: The size of the Monte Carlo samples used in this study and their corresponding integrated luminosities. The cross-section and filter efficiency for each sample are obtained from the information given by the AMI database [84]. The $\gamma\gamma$+(γj+$j\gamma$) sample is a combined sample with di-photon and photon-jet events. γj events are photon-jet events with the reconstructed leading p_T photon candidate being a true photon, and subleading p_T photon candidate being a jet, and vice-versa for $j\gamma$.

cone of $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.4$ around the photon. Cells from the photon shower itself are excluded from the calculation. The leakage from the photon outside the cluster is parameterized as a function of the photon transverse energy based on simulated single photon samples. This leakage is then subtracted from the isolation transverse energy. In this way, the isolation transverse energy becomes independent of the photon transverse energy. Besides, the contribution from the underlying event is also subtracted [92]. Low energy jets are separately used in each of the two pseudorapidity regions ($|\eta| < 1.5$ and $1.5 < |\eta| < 3.0$), to compute an ambient transverse energy density. This transverse energy density is multiplied by the area of the isolation cone, and then subtracted from the isolation transverse energy. Since the ambient transverse energy density is not equally distributed, there might be not much energy inside the cone. Thus, this subtraction may lead to a negative isolation transverse energy. With this correction, the distribution of the isolation transverse energy is centered at zero for photons from the hard-scattering.

The isolation criterion for photon candidate selection is $E_T^{iso} < 3$ GeV. This criterion is chosen for an optimal ratio of signal to background, which keeps a reasonable signal photon efficiency and good background rejection. The overall efficiency for signal photons within the acceptance passing the **T** selection is estimated to be around 90%.

5.3.2 Other photon identification criteria

The isolation transverse energy E_T^{iso} exhibits non-negligible correlations with some shower shape variables, especially those measured in the middle layer. This fact would be an issue, because in data-driven methods used to subtract the background as described in the later sections, the isolation transverse energy of the jets is estimated by requiring the photon candidates to fail the **RobustTight** selection. The variables measured in the strip layer, are defined in a relatively small region at the core of the cluster that is not included in the isolation calculation. Thus they show small correlation with the isolation, especially the variables $\omega_{s\,3}$ and F_{side}. Another two variables (ΔE and E_{ratio}) are found to be relatively uncorrelated with isolation. The variable $\omega_{s\,\mathrm{tot}}$ is defined over a wider range, and is found to be more strongly correlated with isolation. For this reason, a modified loose selection **Loose'** ($\mathbf{L}'$) is defined, in which all the tight cuts on hadronic and Middle layer variables are applied, whereas for the strip layer variables only the cut on $\omega_{s\,\mathrm{tot}}$ is implemented, which means omitting four variables ($\omega_{s\,3}$, F_{side}, ΔE, E_{ratio})[1]. The objects passing the **Loose'** selection but failing **RobustTight** selection are called "non-**RobustTight**" ($\tilde{\mathbf{T}}$), and form a background control sample. Monte Carlo studies show that for the jet background the E_T^{iso} distributions in the **T** and $\tilde{\mathbf{T}}$ samples are fairly similar. All the identification criteria used in this thesis are summarized in Table 5.3.

The definition of the **Loose'** selection contains a sort of arbitrariness in the choice of which strips cuts are being omitted. For systematic studies, the definition can be modified, by omitting only two variables ($\omega_{s\,3}$, F_{side}) or all the five ($\omega_{s\,3}$, F_{side}, ΔE, E_{ratio}, $\omega_{s\,\mathrm{tot}}$)[2]. Omitting only two variables can reduce the correlation with isolation but leads to very limited statistics. Omitting five variables will increase the correlation, but gains about 80% more statistics for the $\tilde{\mathbf{T}}$ background control sample compared with the nominal **Loose'** definition. Therefore the systematic is evaluated only from the five-variable configuration.

[1] Technically, a bit mask 0x45fc01 is applied to the isEM bits.
[2] The bit masks are respectively 0x67fc01 and 0x41fc01.

criterion	**Loose** (**L**)	**Loose'** (**L**$'$)	**RobustTight** (**T**)	non-**RobustTight** ($\tilde{\mathbf{T}}$)
R_{had}	◯	√	√	√
R_{η}	◯	√	√	√
ω_2	◯	√	√	√
R_{ϕ}		√	√	√
$\omega_{s\,\mathrm{tot}}$		√	√	√
$\omega_{s\,3}$			√	⊗
ΔE			√	⊗
F_{side}			√	⊗
E_{ratio}			√	⊗

Table 5.3: The corresponding variables and the cut level involved in each of the identification criteria used in this thesis. The label "◯" means that the variable needs to pass the loose requirement, while the label "√" means it must pass the tight requirement. For the non-**RobustTight** criterion, at least one of the variables labelled with "⊗" must fail the tight requirement.

5.3.3 Event selection

As mentioned in Section 5.2, data from periods E to I are used to perform this study. The event preselection is organized as follows:

- *Skimming*: events must contain at least two photon candidates passing **Loose** selection, and with a reconstructed transverse energy $E_T > 16$ GeV, measured by the EM calorimeter. This requirement is overridden by the subsequent ones, and is listed only for completeness — its purpose is actually to reduce significantly the amount of data without biasing the analysis.

- *Good Run List*: events must satisfy the standard ATLAS Good Run List[3] for electron and photon studies. It requires good conditions for the beams and several detectors (trackers, calorimeters, muon spectrometer and magnet system).

- *Trigger*: events must be selected by the 2g15_loose trigger.

- *Primary vertex*: each event must contain at least one reconstructed primary vertex[4] having at least three charged tracks associated.

- *Two photons in acceptance*: each event must have at least two photons with transverse energy $E_T > 16$ GeV, a pseudo-rapidity η_{S2} measured by the middle layer of ECAL in the fiducial region $|\eta_{\mathrm{S2}}| < 2.37$, with the exclusion of the gap $1.37 < |\eta_{\mathrm{S2}}| < 1.52$ between the barrel and the endcap. Moreover, Object Quality maps are used to

[3] data10_7TeV.pro04.merged_LBSUMM_eg_standard_7TeV.xml

[4] The primary vertex is defined as the primary interaction point, which is reconstructed through a fitting method [87] on tracks. Reconstructed tracks passing the preselection and originating from the interaction region are first selected. The z coordinates of the selected tracks are computed at the perigee with respect to the beam spot center. A global maximum of z coordinate distribution is then picked up and considered as a vertex seed. The adaptive vertex fitting algorithm [88], in which the seed position and the tracks are the input, is used to compute the vertex position. Tracks with $\chi^2 > 49$ are considered to be incompatible with the vertex, and used to seed a new vertex. This procedure is repeated until no new vertex can be found or no unassociated tracks are left.

identify and reject any selected photon cluster affected by some detector problems such as dead and non nominal HV regions, dead Front End Boards (including dead optical links OTX) and isolated dead or high noise channels. A criterion based on the *jet cleaning* is also applied. This criterion is applied to reject bad quality jets in the electromagnetic calorimeter [85]. These jets are affected by some noise bursts affecting ECAL. They are characterized by cells with bad calorimeter reconstruction "quality"[5] and with large reconstructed energy of which a large fraction is only in ECAL. In practice, if the reconstructed jet associated to the photon has an electromagnetic energy fraction $f_{EM} > 0.95$ and a fraction of LAr bad cells $f_{quality} > 0.8$, this photon will be rejected.

- *Two loose photons*: the photons described in the previous step must pass the **Loose** selection.

- *Choice of the photon pair*: the best photon pair is then selected by choosing the two photon candidates with highest E_T among those fulfilling the acceptances and **Loose** requirements described above. Since the isolation transverse energy $E_T^{\rm iso}$ is measured inside a cone $R < 0.4$, it is important that a candidate of the pair does not enter the cone of the other one. Therefore a separation criterion:

$$\Delta R_{\gamma\gamma} = \sqrt{(\eta_1^{\rm clus} - \eta_2^{\rm clus})^2 + (\phi_1^{\rm clus} - \phi_2^{\rm clus})^2} > 0.4 \tag{5.1}$$

is imposed between the two clusters. If this requirement is not fulfilled, the full event is discarded. A requirement $\Delta R_{\gamma\gamma} > 0.8$ may appear to be preferable, because it would guarantee no overlap between the two isolation cones. However, as shown in Figure 5.4, the isolation transverse energy distributions appear to be similar for $\Delta R_{\gamma\gamma} > 0.8$ and $0.4 < \Delta R_{\gamma\gamma} < 0.8$, both for **RobustTight** and non-**RobustTight** samples, while it is completely different for $\Delta R_{\gamma\gamma} < 0.4$. Therefore the choice has been done in order to maximize the acceptance.

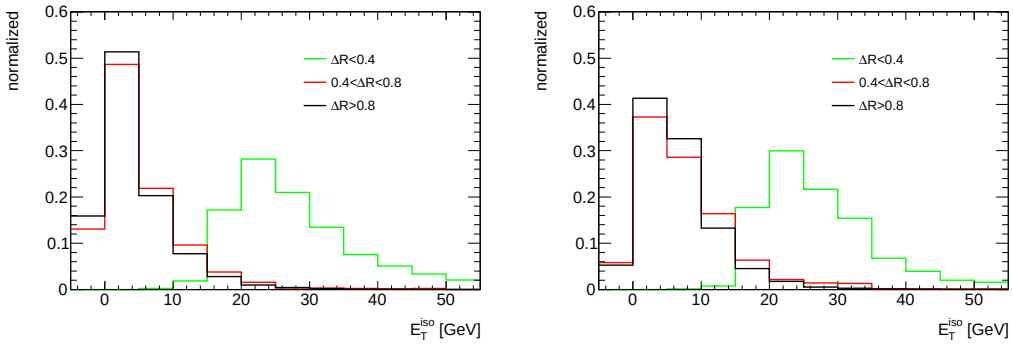


Figure 5.4: Isolation transverse energy distributions for **RobustTight** (left) and non-**RobustTight** (right) candidates, for different $\Delta R_{\gamma\gamma}$ separations: $\Delta R_{\gamma\gamma} > 0.8$ (black), $0.4 < \Delta R_{\gamma\gamma} < 0.8$ (red) and $\Delta R_{\gamma\gamma} < 0.4$ (green).

[5]The quality is determined by the difference between the measured pulse and the reference pulse that is used for the cell energy reconstruction. If a large difference is observed, the cell is considered to be a bad quality cell.

After the event preselection, if both photon candidates satisfy the **RobustTight** identification cuts (Section 3.3) and fulfill the calorimetric isolation requirement $E_T^{\text{iso}} < 3$ GeV (Section 5.3.1), the event is classified as a di-photon signal candidate. Otherwise, if both photon candidates satisfy the **Loose'** selection (Section 5.3.2), the event is used to populate the distributions in the background control regions. The choice of the best photon pair based on E_T-ranking is done before the **RobustTight** selection, in order to have similar kinematic spectra for the signal candidates and the control regions. A summary of event numbers after each step of selection can be found in Table 5.4. There are 5365 photon pairs fulfilling the **RobustTight** selection, of which 2022 are also isolated, as indicated in the last two columns "**TT**" and "**TITI**" of Table 5.4.

Selection	skim	GRL	trigger	vertex	2γ	**LL**	ΔR	**TT**	**TITI**
Event number	266098	227387	113749	113725	66459	65442	63673	5365	2022

Table 5.4: Number of events after different requirements.

5.4 Extraction of the di-photon signal

An isolation template fit was used in ATLAS to measure the inclusive isolated photon cross-section [90]. In this method, signal and background isolation templates are separately derived using data-driven methods. A total template summing signal and background templates together is used to perform a binned maximum likelihood fit to the isolation transverse energy distribution of the tightly identified photon candidates (passing **RobustTight** selection).

In the di-photon analysis presented here, this technique is extrapolated to be a two-dimensional fit.

5.4.1 Extraction of the yields

The signal ($\gamma\gamma$) and background yields are extracted from an extended maximum likelihood fit to the two-dimensional distribution of the isolation transverse energies $E_{\text{T},1}^{\text{iso}}$ and $E_{\text{T},2}^{\text{iso}}$ of the leading and the subleading photon candidates. The fit is performed to the **TT** sample, *i.e.* both photons are required to pass the **RobustTight** criteria. The total 2D isolation transverse energy distribution is modeled with the sum of four components: one signal ($\gamma\gamma$) and three background (γj[6], $j\gamma$[7], jj) categories, weighted by their corresponding yields:

$$W_{\text{tot}}^{\text{TT}} F_{\text{tot}}^{\text{TT}} = W_{\gamma\gamma}^{\text{TT}} F_{\gamma\gamma}^{\text{TT}} + W_{\gamma j}^{\text{TT}} F_{\gamma j}^{\text{TT}} + W_{j\gamma}^{\text{TT}} F_{j\gamma}^{\text{TT}} + W_{jj}^{\text{TT}} F_{jj}^{\text{TT}}, \tag{5.2}$$

where $F_{\text{tot}}^{\text{TT}}$ represents the 2D probability density function of the two isolation transverse energies for all the events with both photon candidates passing the **RobustTight** criteria, and F_i^{TT} represents the 2D PDF for the i-th event category. The fit is performed in a large isolation transverse energy range ($-5 < E_T^{\text{iso}} < 25$ GeV) for both photons, but the

[6]Photon-jet events for which the leading p_T photon candidate is a true photon and subleading p_T candidate is a jet.

[7]Photon-jet events for which the leading p_T photon candidate is a jet and subleading p_T candidate is a true photon.

fit results are expressed in terms of the yields W^{TITI} inside the signal region, where both photon candidates are isolated ($-5 < E_T^{\mathrm{iso}} < 3$ GeV).

If the correlations between the leading p_T and subleading p_T photon candidates are negligible for a certain event category, the corresponding 2D PDF can be factorized as the product of two one-dimensional PDFs. The linear correlations expected from the simulation for the four event categories are summarized in Table 5.5. The two-dimensional isolation transverse energy distributions for di-photon candidates selected on Monte Carlo samples are shown in Figure 5.5 for the four categories. The correlations between the isolation transverse energies of the two candidates are found to be negligible for $\gamma\gamma$, γj and $j\gamma$ events, and the corresponding PDFs are factorized as the product of two one-dimensional PDFs.

On the other hand, for the jj background, there are too few events to draw a conclusion from the simulation sample. However, the correlation can be studied on real data. The background isolation transverse energy distribution is derived from candidates (the "probes") in the $\tilde{\mathbf{T}}$ region with different requirements on the other photon candidate in the event (the "tag"), including: no requirement, $\mathbf{T}$, $\tilde{\mathbf{T}}$, isolated (E_T^{iso} <3 GeV) and non-isolated (E_T^{iso} >3 GeV). The distributions are shown in Figure 5.6. If the tag (either photon or jet) is isolated, the probe background tends to be more isolated, and vice-versa. One can conclude that the correlation between the isolation transverse energies of the two photon candidates is large in the jj event category. The 2D jj PDF is therefore modeled with a two-dimensional distribution that takes into account the observed correlations.

An additional complication arises from the fact that leading and subleading true photons have slightly different isolation PDFs, as described in Section 5.4.2, and the same holds for fake photon candidates. The final fit becomes:

$$W_{\mathrm{tot}}^{\mathrm{TT}} F_{\mathrm{tot}}^{\mathrm{TT}} = W_{\gamma\gamma}^{\mathrm{TT}} F_{\gamma,1}^{T} F_{\gamma,2}^{T} + W_{\gamma j}^{\mathrm{TT}} F_{\gamma,1}^{T} F_{j,2}^{T} + W_{j\gamma}^{\mathrm{TT}} F_{j,1}^{T} F_{\gamma,2}^{T} + W_{jj}^{\mathrm{TT}} F_{jj}^{\mathrm{TT}} \tag{5.3}$$

where $F_{\gamma,1}^{T}(E_{\mathrm{T},1}^{\mathrm{iso}})$ and $F_{\gamma,2}^{T}(E_{\mathrm{T},2}^{\mathrm{iso}})$ are the one-dimensional leading and subleading photon isolation PDFs, $F_{j,1}^{T}(E_{\mathrm{T},1}^{\mathrm{iso}})$ and $F_{j,2}^{T}(E_{\mathrm{T},2}^{\mathrm{iso}})$ are the one-dimensional leading and subleading jet isolation PDFs, and $F_{jj}^{\mathrm{TT}}(E_{\mathrm{T},1}^{\mathrm{iso}}, E_{\mathrm{T},2}^{\mathrm{iso}})$ is the two-dimensional jj isolation PDF.

With this fit function, one can directly extract the number of signal ($W_{\gamma\gamma}^{\mathrm{TT}}$) and background events ($W_{\gamma j}^{\mathrm{TT}}, W_{j\gamma}^{\mathrm{TT}}, W_{jj}^{\mathrm{TT}}$) in the **TT** sample as well as – exploiting the knowledge of the partial integrals of the various PDFs in the isolation signal region – the events number $W_{\gamma\gamma}^{\mathrm{TITI}}$, $W_{\gamma j}^{\mathrm{TITI}}$, $W_{j\gamma}^{\mathrm{TITI}}$, W_{jj}^{TITI} below the isolation cut (**TITI** sample).

final states	entries in MC after selection	linear correlation
$\gamma\gamma$	40413	$(5.5 \pm 0.5)\%$
γj	2240	$(5.8 \pm 2.1)\%$
$j\gamma$	524	$(-1.7 \pm 4.4)\%$
jj	6	$(46 \pm 33)\%$

Table 5.5: Linear correlation coefficient between the isolation transverse energies of the two photon candidates from the simulation, in different truth-level final states.

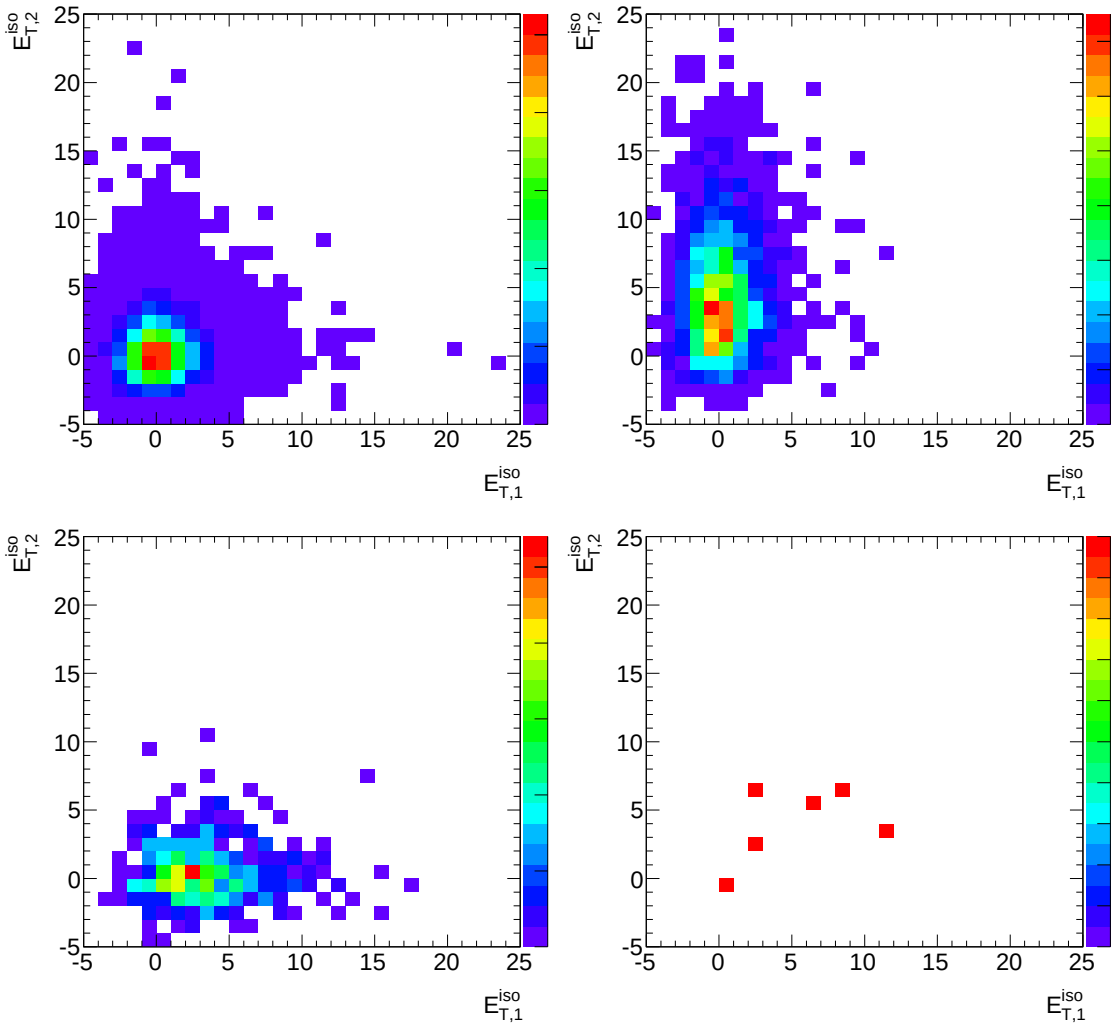


Figure 5.5: Two-dimensional distributions of the two photon isolation transverse energies, as expected from simulation. Top left: $\gamma\gamma$. Top right: γj. Bottom left: $j\gamma$. Bottom right: jj.

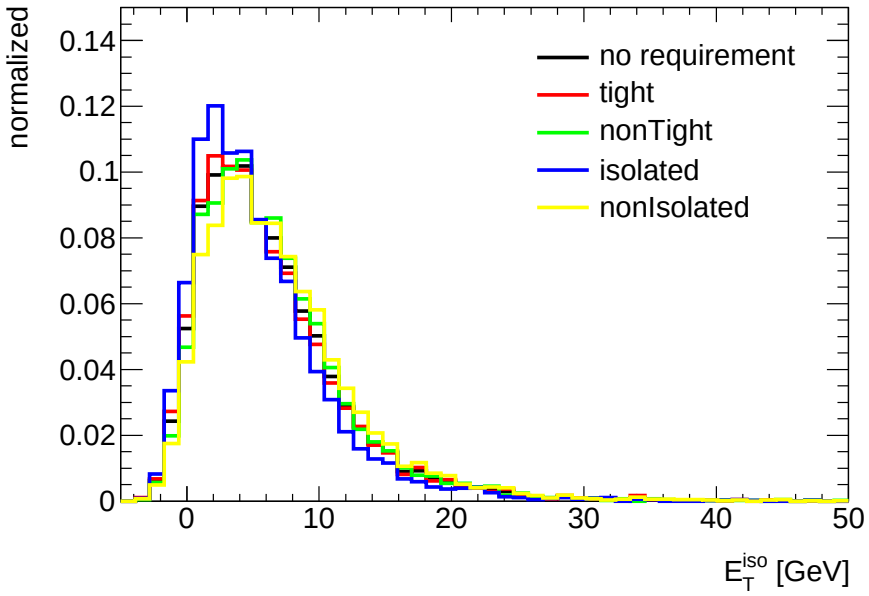


Figure 5.6: The background isolation transverse energy distribution for the probe under the different conditions on the tag. Black: no requirement, red: **T**, green: $\tilde{\mathbf{T}}$, blue: isolated, yellow: non-isolated.

5.4.2 Signal isolation transverse energy one-dimensional PDF

Electrons are used to estimate the signal photon isolation transverse energy distribution in real data. Differences between electrons and signal photons mainly come from two sources: the electron isolation transverse energy is expected to be slightly larger because of the material upstream of the calorimeter, and the photon isolation transverse energy spectrum exhibits a larger tail because of the contribution of the bremsstrahlung and the fragmentation photons. In the following, these differences are estimated from Monte Carlo and a correction is applied to the isolation transverse energy distribution of electrons in real data.

A relatively pure electron sample can be selected from $W \to e\nu$ and $Z \to ee$ decays. The energy taken away by the neutrino is not detectable. Thus a variable called missing transverse energy is calculated for selecting the events with neutrino involved. The missing transverse energy E_T^{miss} is reconstructed with only the calorimeter. It is computed by:

$$E_x^{\text{miss}} = -\sum_{i=1}^{N_{cell}} E_i \sin\theta_i \cos\phi_i \tag{5.4}$$

$$E_y^{\text{miss}} = -\sum_{i=1}^{N_{cell}} E_i \sin\theta_i \sin\phi_i \tag{5.5}$$

$$E_T^{\text{miss}} = \sqrt{(E_x^{\text{miss}})^2 + (E_y^{\text{miss}})^2} \tag{5.6}$$

Here, E_i is the cell energy, θ_i is the cell polar angle and ϕ_i is the cell azimuthal angle. Cells belonging to the reconstructed clusters in the ECAL and HCAL above the noise are all summed up.

A summary of the requirements to select the electron samples from $W \to e\nu$ and $Z \to ee$ events is as follows:

- For $W \to e\nu$ events:
  - an electron with a cluster $E_T > 20$ GeV passing **RobusterTight**[8] identification level;
  - missing transverse energy $E_T^{\text{miss}} > 25$ GeV;
  - transverse mass of the $e - E_T^{\text{miss}}$ system, $m_{\text{T}} > 40$ GeV;
  - azimuthal angle between missing transverse energy and nearest jet (with energy > 10 GeV): $\Delta\phi > 2.5$;
- For $Z \to ee$ events:
  - an electron and a positron candidate, both with cluster $E_T > 20$ GeV, passing **RobusterTight** identification level;
  - invariant mass of the electron-positron pair within the range: $66.118 < m_{ee} < 116.118$ GeV;

Figure 5.7 shows the isolation distribution comparison between data and MC with the electron samples selected. The isolation energy from data tends to be slightly larger than the simulated sample. One possible explanation for this discrepancy is the underestimation of the contribution from the underlying events in real data. Or on the contrary, the contribution from the pile-up effect might be lower estimated in the simulated sample.

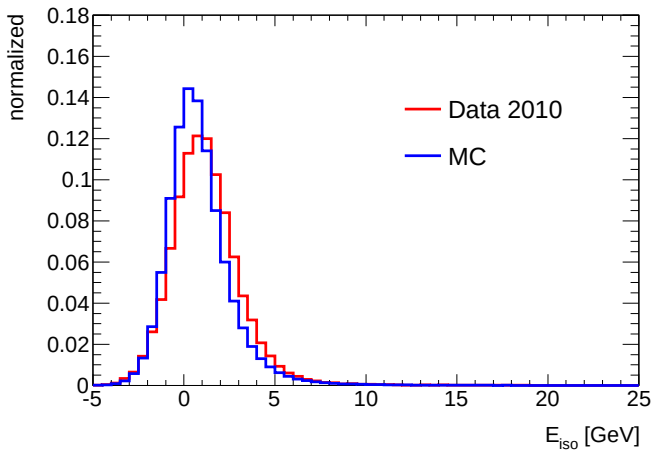


Figure 5.7: Isolation transverse energy distributions of electrons from $W \to e\nu$ and $Z \to ee$ decays. Red line: from data, blue line: from simulated sample.

[8]This criterion is optimized for electron selection in ATLAS. Many variables are cut on, including the shower shape variables (R_{had}, R_η, ω_2, $\omega_{s\,\text{tot}}$ and E_{ratio}) track quality variables (number of hits in the pixel and silicon trackers, the transverse impact parameter), a cluster-track matching variable ($\Delta\eta$), the ratio of cluster energy to track momentum, the number of hits in the TRT, and the ratio of high-threshold to the total number of hits in the TRT. In addition, Electrons are vetoed from the conversion.

The electron isolation transverse energy distribution in data is fitted with a Crystal Ball function [89]:

$$F(t) = \begin{cases} \exp\left(-t^2/2\right) , & \text{for } t > -\alpha \\ (n/|\alpha|)^n \cdot \exp\left(-|\alpha|^2/2\right) \cdot (n/|\alpha| - |\alpha| - t)^{-n} , & \text{otherwise} \end{cases} \quad (5.7)$$

where $t = (E_T^{\text{iso}} - \mu)/\sigma$, and μ, σ, α and n are the free parameters of the fit. The fitted distribution is shown in Figure 5.8.

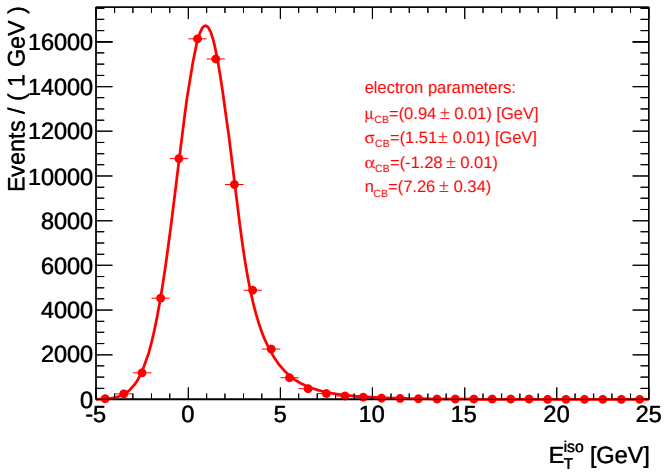


Figure 5.8: Isolation transverse energy distribution of electron candidates in data, selected from $W \to e\nu$ and $Z \to ee$ decays. The distribution is fitted with a Crystal Ball function.

To extract the electron-to-photon correction, simulated samples listed in Table 5.2 are used: the $\gamma\gamma + (\gamma j + j\gamma)$ as well as a luminosity-weighted sample combining $W \to e\nu$ and $Z \to ee$ events. Samples with bunch-train pile-up are used. The fitted isolation transverse energy distributions are shown in Figure 5.9 for electrons and in Figure 5.10 for leading and subleading photons separately. The fitted parameters are summarized in Table 5.6, together with the resulting parameter corrections to be applied to the electron distribution shape. The different corrections between leading and subleading photons indicates that there is a longer tail for subleading photons isolation transverse energy distribution. This can be explained partly by the contribution from the bremsstrahlung and one fragmentation processes. In those processes, a parton recoils the other direct photon in the final state of the hard process. The photon from the fragmentation tends to have smaller transverse energy than the direct photon, and has larger isolation energy since it is close to or inside the parton initiated jet. The estimated photon isolation transverse energy shape parameters after electron-to-photon correction are summarized in Table 5.7.

5.4.3 Background isolation transverse energy one-dimensional PDF

As already mentioned in Section 5.3.2, the objects passing **Loose'** selection but failing **RobustTight** (**T**) selection are called "non-**RobustTight**" ($\tilde{\mathbf{T}}$), and form a background control sample. The background isolation transverse energy obtained in this background control sample is used to model the isolation transverse energy distribution in the signal (**T**) region. The distribution for the leading photon candidate is obtained requiring the

parameters	μ	σ	α	n
electrons	0.274 ± 0.001	1.477 ± 0.001	-1.064 ± 0.002	8.237 ± 0.072
Leading photons	-0.03 ± 0.02	1.45 ± 0.01	-1.23 ± 0.03	9.90 ± 1.68
Subleading photons	0.09 ± 0.02	1.46 ± 0.01	-1.14 ± 0.03	3.17 ± 0.17
correction	$\Delta\mu$	$\Delta\sigma$	$\Delta\alpha$	Δn
Leading photon	-0.30	-0.03	-0.17	$+1.66$
Subleading photon	-0.18	-0.02	-0.08	-5.07

Table 5.6: Result of the Crystal Ball fits to the electron and photon isolation transverse energy distributions in Monte Carlo samples. The last two lines are the correction extracted from the Monte Carlo fits to be applied to the electron distribution shape from data.

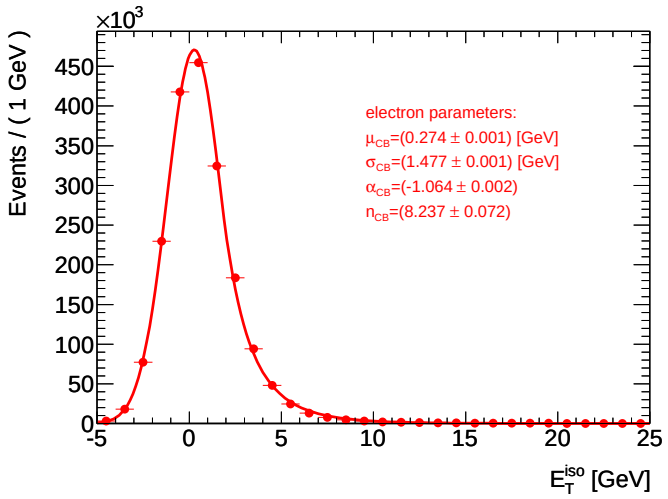


Figure 5.9: Isolation transverse energy distribution of simulated electrons from $W \to e\nu$ and $Z \to ee$ decays. The distribution is fitted with a Crystal Ball function.

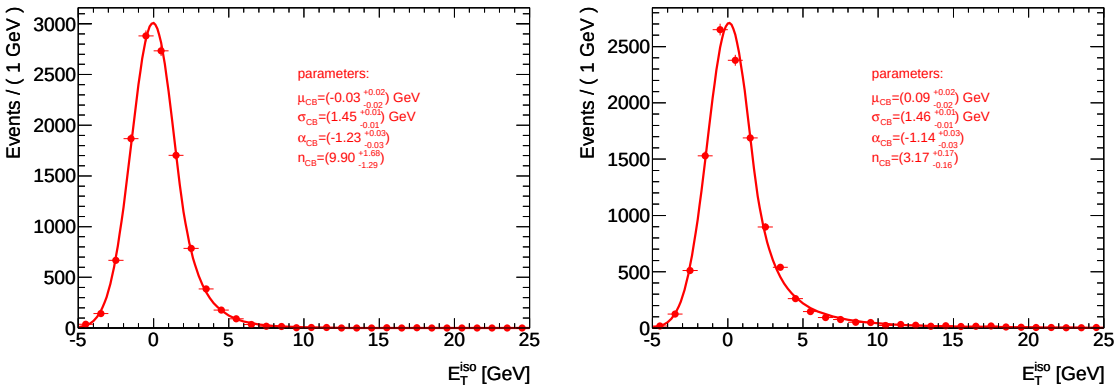


Figure 5.10: Isolation transverse energy distribution of simulated photon from $\gamma\gamma$+(γj+$j\gamma$) sample. The distribution is fitted with a Crystal Ball function. Left: leading, Right: subleading

parameters	μ	σ	α	n
Leading photon	0.64	1.48	−1.45	8.92
Subleading photon	0.76	1.49	−1.36	2.19

Table 5.7: Crystal Ball parameters of the photon isolation transverse energy shape estimated from electrons in data after electron-to-photon Monte Carlo corrections.

leading photon candidate to be $\tilde{\mathbf{T}}$ and the subleading photon to be $\mathbf{T}$. Similar but reverse conditions are required to get the isolation transverse energy distribution for the subleading photon candidate. The distributions are then fitted with the sum of a Novosibirsk function and a Gaussian with the same peak. The Novosibirsk function has the functional form:

$$P(x) = e^{-0.5^{(\ln q_y)^2/\Lambda^2+\Lambda^2}} \tag{5.8}$$

$$q_y = 1 + \Lambda(x - x_0)/\sigma \times \frac{\sinh(\Lambda\sqrt{\ln 4})}{\Lambda\sqrt{\ln 4}} \tag{5.9}$$

Here x_0 is the peak position, σ is the width of the peak, and Λ is a parameter describing the tail of the distribution. The leading and subleading background isolation transverse energy PDFs are shown in Figure 5.11.

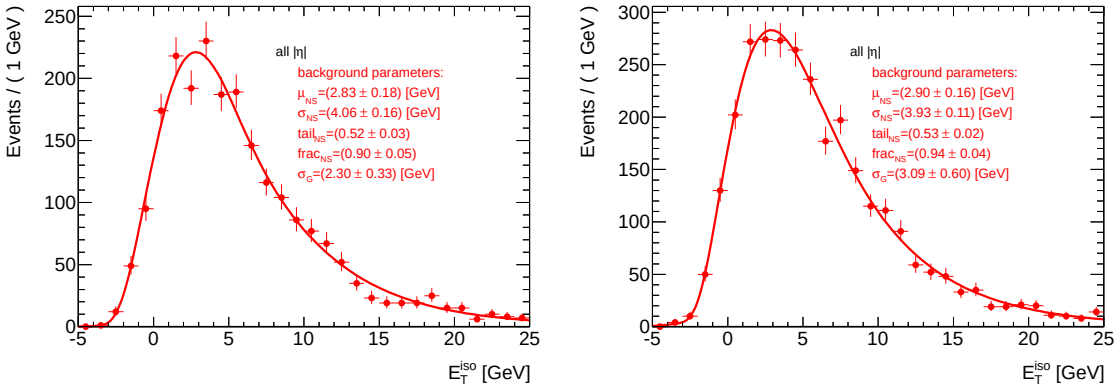


Figure 5.11: One-dimensional jet isolation transverse energy PDFs. Left: leading-p_T candidate PDF. Right: subleading-p_T candidate PDF.

5.4.4 Two-dimensional jj isolation transverse energy PDF

The two-dimensional isolation transverse energy PDF for jj events is derived from a data control sample obtained after applying all the di-photon analysis selection criteria with the exception that both photon candidates are required to be $\tilde{\mathbf{T}}$. The two-dimensional histogram is smoothed using an adaptive kernel estimation technique, and the corresponding "2D keys" PDF is used in the final fit. The adaptive kernel estimation technique actually makes use of a bunch of Gaussian distributions to describe data distribution. The mean value and the width for each Gaussian distribution are determined by the local distribution density of data. More information about this technique can be found in [91]. The two-dimensional isolation transverse energy distribution in selected jj events in the $\tilde{\mathbf{T}}$ sample and the corresponding smoothed PDF are shown in Figure 5.12.

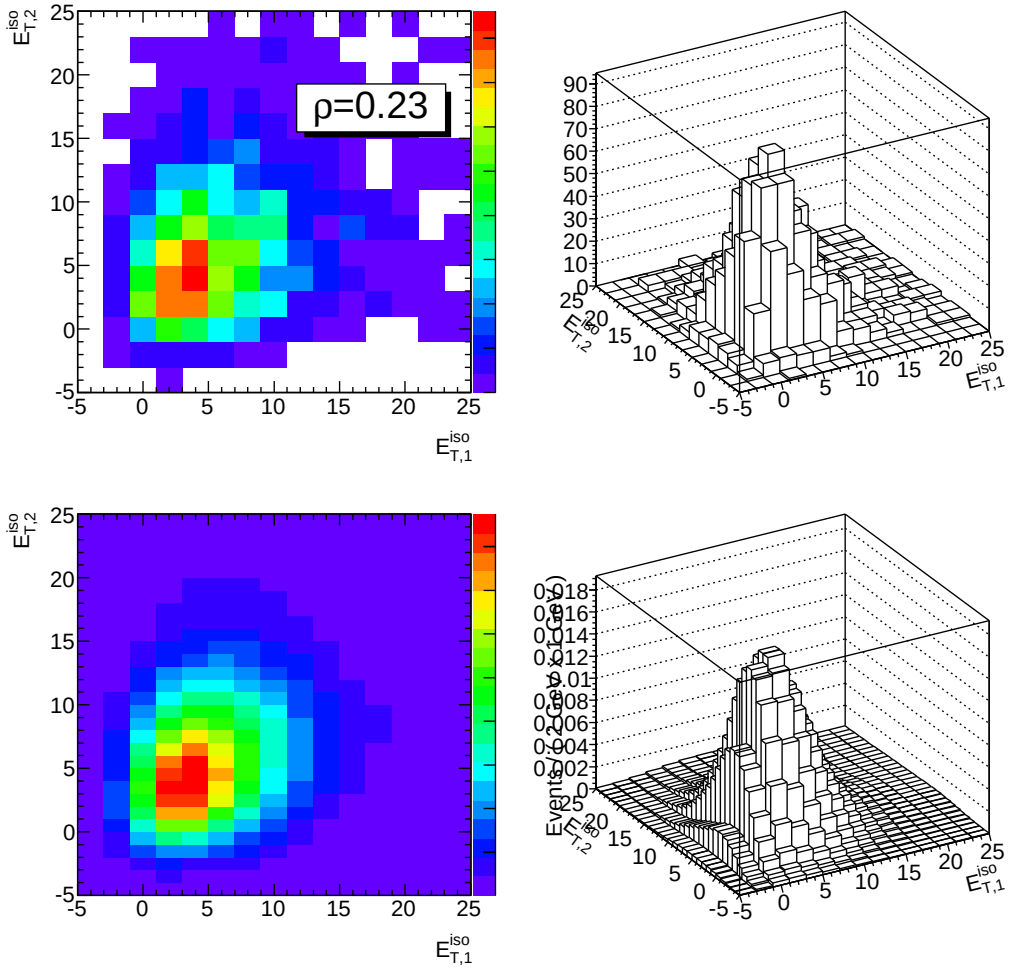


Figure 5.12: Two-dimensional jj isolation transverse energy PDF. Top: data histogram. Bottom: continuous interpolation.

5.4.5 Tests on Monte Carlo

In order to check that the fit is not biased, a test is performed on a simulated sample of $\gamma\gamma$ and $\gamma j + j\gamma$ events as listed in Table 5.2. The same selection as in Section 5.3.3 is applied. Due to the statistics limitation of the simulated samples, the jj component is not included (there are almost no jj candidates in MC after the **TT** cuts). The first test is to fit the selected sample using as signal and background PDFs the isolation transverse energy distributions from true photons and true jets after **T** cuts. Then, to make the fit more realistic and closer to what will be done on data, it is repeated using the true signal isolation transverse energy PDF for photons while using for jets, the background isolation transverse energy PDF obtained from true jets in the $\tilde{\mathbf{T}}$ control region. The results from these two steps are summarized in Table 5.8 together with the true yields for comparison. The results from the fit for the signal yield in both cases are in agreement, within the statistical uncertainties, with the true yields.

	True yields	Fit results using true γ and **T** true j PDFs	Fit results using the true γ and $\tilde{\mathbf{T}}$ true j PDFs	relative difference
$W_{\gamma\gamma}$	2505	2537 ± 72	2587 ± 72	2%
$W_{\gamma j}$	827	803 ± 26	736 ± 24	8%
$W_{j\gamma}$	253	231 ± 23	272 ± 27	18%

Table 5.8: Results of 2D fit on a luminosity-weighted sample of $\gamma\gamma$ and γj+$j\gamma$ Pythia simulated samples, using either true photon and jet PDFs (for photon and jets passing the **RobustTight** identification criteria), or the true photon PDF and the true jet PDF determined from photon candidates failing the **RobustTight** identification criteria. The true yields are given in the left column. The rightmost column is the absolute value of the difference between the two fit results (3rd and 4th columns) divided by the fit result in the third column.

5.4.6 Results on real data

The results from the 2D fit to the **TT** sample, expressed in terms of the yields inside the **TITI** region, are summarized in Table 5.9. A significant di-photon signal in the **TITI** region is observed as far as the statistical errors are concerned. The dominant background contribution comes from the γj events where the leading photon candidate is a photon and the subleading candidate comes from a jet. The isolation transverse energy distributions of the two photon candidates in the **TT** region are shown in Figure 5.13, together with the projections of the fit.

event category	yields in **TT**	yields in **TITI**
$W_{\gamma\gamma}$	1770 ± 69	1353 ± 53
$W_{\gamma j}$	1240 ± 67	348 ± 19
$W_{j\gamma}$	622 ± 52	168 ± 14
W_{jj}	1643 ± 64	174 ± 7

Table 5.9: Results of the 2D fit. Event yields are given in the **TT** and the **TITI** regions.

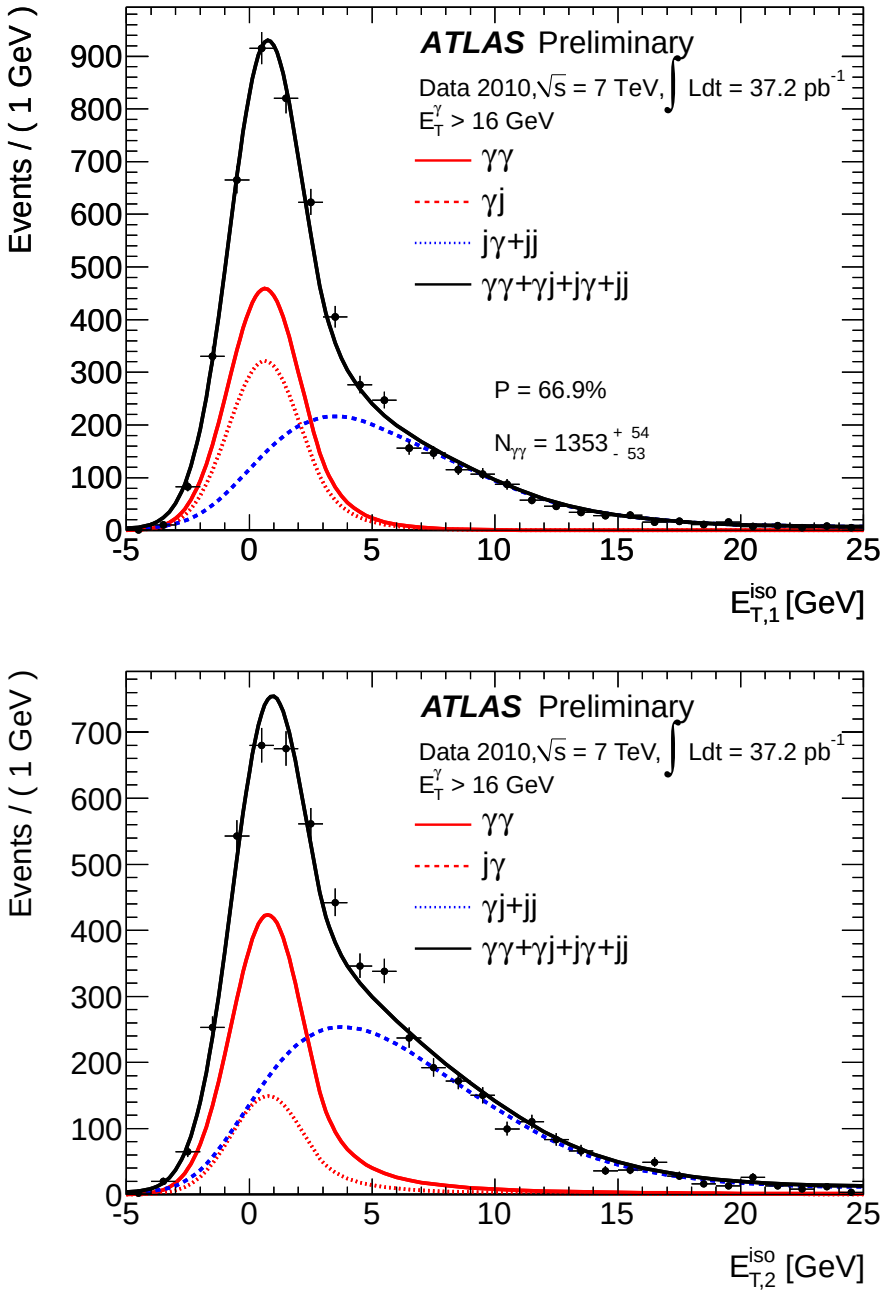


Figure 5.13: Projections of the 2D template fit on the two photon candidates' isolation transverse energies. Top: leading photon candidates, Bottom: subleading photon candidates. An analytic parameterization of the γ and j (except for jj) isolation PDFs has been used. For the jj background a smooth interpolation (with the kernel technique) of the 2D binned template has been used.

5.4.7 Systematic uncertainties

The main systematic uncertainties originate from the imperfect knowledge of the signal and background templates. They arise from both the limited statistics of the control samples used to extract the photon and jet PDFs (which translate into a statistical uncertainty on the PDF parameters), and from the assumptions that the isolation distributions of the control samples provide a good representation of the isolation distributions in the data sample being fitted.

The systematic uncertainties arising from the limited knowledge of the photon and jet PDFs parameters are estimated in the following way:

- using the central values and the covariance matrices of the PDFs parameters as returned from the fits to the isolation distributions of the control samples (with a Crystal Ball lineshape for the signal and with a Novosibirsk lineshape for the single-jet background), 1000 alternative sets of parameters are randomly generated;
- each time a new set of parameters is generated, the fit to the data is repeated, using this set of parameters as the nominal one;
- the RMS of the distribution of the fitted yield in each category i is taken as the systematic uncertainty on the yield W_i.

The procedure is applied separately to the photon and to the jet PDFs.

For the jj PDF a similar study is carried out in the following way. The jj PDF is derived by requiring both photon candidates to fail the **RobustTight** identification criteria and to pass the non-**RobustTight** ones, and a continuous 2D PDF is obtained by smoothing the data histogram using a 2D kernel estimation technique. Again, the limited size of the data control sample may translate into a poor model of the true distribution. We use therefore the nominal jj distribution to produce 200 alternative samples with similar size, which are then used to construct the 2D isolation histogram and then a smoothed PDF. The latter is then used in the fit, replacing the nominal one. The RMS of the distribution of the fitted yield in each category i is taken as the systematic uncertainty on the yield W_i.

As described in Section 5.4.2 the signal template is obtained from the prompt electron (from W/Z) isolation transverse energy distribution after correcting it for the difference observed in simulation between electrons from W/Z and photons in $\gamma\gamma$ events. The photon isolation transverse energy distribution in the simulation, and therefore also the values of the corrections, depend on the relative fraction of direct and fragmentation photons in the selected sample, the latter having typically a larger tail at high isolation values. A first source of error on this relative fraction comes from the MC samples composition. The MC used to compute the PDF parameters only includes $\gamma\gamma$ and γj+$j\gamma$ samples thus underestimating the fraction of fragmentation photons by not taking into account the jj contribution. A possible way to account for this systematics is to determine the PDF parameters for a MC sample including jj events. However, the equivalent luminosity of the sample including jj events is samll, thus the fitted parameters given in Table 5.10 exhibit large statistical errors.

A second limitation of this estimation of the uncertainty from the relative amount of fragmentation photons is that it relies on the prediction of PYTHIA. Therefore, a more conservative approach is adopted.

For this second method, one considers that the uncertainty on the fragmentation to direct photon ratio should affect mostly the subleading photon since the leading one is

parameters	μ	σ	α	n
leading photon PDF				
$\gamma\gamma$ + (γj+$j\gamma$) sample	-0.03 ± 0.02	1.45 ± 0.01	-1.23 ± 0.03	9.93 ± 1.52
$\gamma\gamma$ + (γj+$j\gamma$) + jj sample	0.18 ± 0.13	1.62 ± 0.10	-1.52 ± 0.23	1.70 ± 0.62
subleading photon PDF				
$\gamma\gamma$ + (γj+$j\gamma$) sample	0.10 ± 0.02	1.45 ± 0.01	-1.14 ± 0.03	3.13 ± 0.16
$\gamma\gamma$ + (γj+$j\gamma$) + jj sample	-0.18 ± 0.16	1.31 ± 0.12	-0.71 ± 0.14	3.89 ± 1.61

Table 5.10: Comparison between the parameters from the Crystal Ball fits to the $\gamma\gamma$ + (γj+$j\gamma$) sample and $\gamma\gamma$ + (γj+$j\gamma$) + jj sample, the latter containing relatively more photons from bremsstrahlung and fragmentation.

likely to be a direct photon. Two extreme cases are thus considered where the values 100% (all subleading photons originate from the direct process) and 0% (all subleading photons originate from fragmentation) are assumed. The isolation transverse energy distributions of direct photons and of fragmentation photons are fitted with Crystal Ball function, whose parameters are summarized in Table 5.11, and two corresponding sets of electron-to-photon correction factors are determined. The fit is repeated on the data using, as subleading photon PDF, the electron isolation energy PDF measured in data after applying either the correction factors obtained from the direct photon MC or from the fragmentation photon MC (bottom part of Table 5.11). The leading photon PDF is kept unchanged.

parameters	μ	σ	α	n
direct photons	0.05 ± 0.02	1.43 ± 0.02	-1.11 ± 0.03	19.2 ± 9.7
bremsstrahlung photons	0.21 ± 0.04	1.57 ± 0.03	-0.92 ± 0.04	2.44 ± 0.21
subleading photon PDF				
all subleading are direct	0.72	1.46	-1.33	18.3
all subleading are brem	0.93	1.60	-1.14	1.46

Table 5.11: Parameters from the Crystal Ball fits to the simulated direct and fragmentation photon isolation transverse energy distributions for subleading photons. Resulting subleading photon PDFs in the two extreme hypotheses: all subleading photons are direct and all subleading photons are fragmentation photons.

A similar procedure is used to estimate the uncertainty on the estimated yields due to the limited knowledge of detector geometry and material in the simulation. The electron-to-photon correction is evaluated from electron and $\gamma\gamma + (\gamma j + j\gamma)$ samples simulated with additional material in the Inner Detector and calorimeter as mentioned in Section 5.2. The signal PDFs for leading and subleading photons are derived by applying this correction to the electron sample from data (Table 5.12) and the fit of the event yields is repeated.

As described in Section 5.4.4 the background template is obtained from the $\tilde{\mathbf{T}}$ data control sample. This may be contaminated by the presence ("leakage") of some residual photon signal failing the identification criteria. A modification on the fit is performed to exclude the signal leakage from the background template. The leakage ratio is derived from simulated samples. The difference between the yields obtained with the nominal configuration (signal leakage neglected) and with the signal leakage corrected fit is taken as

parameters	μ	σ	α	n
electrons	0.322 ± 0.003	1.483 ± 0.003	-1.037 ± 0.005	9.12 ± 0.20
leading photons	0.01 ± 0.02	1.44 ± 0.01	-1.21 ± 0.03	8.92 ± 1.38
subleading photons	0.12 ± 0.02	1.46 ± 0.02	-1.14 ± 0.03	3.22 ± 0.19
signal photon PDFs with additional material				
leading photons	0.63	1.47	-1.45	7.06
subleading photons	0.74	1.49	-1.38	1.36

Table 5.12: Parameters from the Crystal Ball fits to the electron, leading, and subleading photon isolation distribution in Monte Carlo samples simulated with additionnal material in the Inner Detector. Resulting leading and subleading photon PDFs.

a systematics. Also, non-negligible correlations between the isolation and the identification variables may have the consequence that the isolation transverse energy distribution of $\tilde{\mathbf{T}}$ fake photons is different from that of fake photons passing the **T** identification criteria. As described in Section 5.4.5, the fit on Monte Carlo using a background isolation transverse energy determined from $\tilde{\mathbf{T}}$ true jets is able to correctly estimate the signal and background yields within 1σ of the true value. A systematic uncertainty is conservatively assigned by taking the ratio between the yields fitted with **T** true jet isolation and the yields fitted with the $\tilde{\mathbf{T}}$ true jets. The impact of choosing a larger set of variables to define the $\tilde{\mathbf{T}}$ region is also evaluated. A background isolation transverse energy PDF is thus extracted from a new $\tilde{\mathbf{T}}$ sample with five strip variables reversed. The 2D fit to the data is then performed with this new background PDF, and the difference with the nominal result is taken as a systematic uncertainty.

The electron-to-photon correction is estimated using a di-photon Monte Carlo sample and requiring that both the leading and the subleading candidates match a true photon. As a consequence the fit method relies on the assumption that the isolation transverse energy PDF of the leading (resp. subleading) photons in $\gamma\gamma$ events is similar to the isolation transverse energy PDF of the leading (resp. subleading) photons in γj (resp. $j\gamma$) events. To determine the systematics linked to this effect, the isolation transverse energy PDF of leading photons in γj events and subleading photons in $j\gamma$ is computed, meaning the other photon candidate is required to be a fake. The fit is performed using those PDFs for the γj and $j\gamma$ components. The results are very similar to the nominal case. The PDFs used for this check being known to a much worse precision for statistical reasons, it is preferable to keep the fit PDF modeling as described in Section 5.4.4 and consider this check as a systematics.

A summary of all the systematics discussed above can be found in Table 5.13. The dominant systematic uncertainty comes from the definition of the background control sample. It is evaluated by omitting all the five strip variables compared with the nominal results where all the five strip variables except $\omega_{s\,\mathrm{tot}}$ are omitted. Thus, this systematic uncertainty reflects the correlations between the isolation transverse energy and the shower shape variable $\omega_{s\,\mathrm{tot}}$. The subleading systematic uncertainty is from the relative fraction of the direct and fragmentation photons. The reason is that the tail of the signal distribution depends strongly on this fraction. A direct measurement of the signal photon isolation transverse energy distribution on the data would overcome this limitation. A method proposed in ATLAS is to subtract from the **T** sample distribution the background distribution extracted

from the $\tilde{\mathbf{T}}$ control region. The normalization is derived from the $E_T^{iso} > 7$ GeV where both samples are largely dominated by fake photon candidates. The other sources of systematic uncertainties give small contributions, such as the signal and background isolation transverse energy PDF parameters, the signal leakage into the $\tilde{\mathbf{T}}$ background control samples and the effect of the material on the signal.

	$W_{\gamma\gamma}^{\mathrm{TITI}}$	$W_{\gamma j}^{\mathrm{TITI}}$	$W_{j\gamma}^{\mathrm{TITI}}$	W_{jj}^{TITI}
signal PDF parameters	± 9	± 9	± 2	± 1
jet PDF parameters	± 16	± 13	± 7	± 0
jj PDF	± 14	± 7	± 6	± 4
direct/brem photon fraction	± 115	± 140	± 10	± 16
effect of material on signal	$+21$	-40	-1	-4
signal leakage	$+16$	-12	-5	0
jet PDF from $\tilde{\mathbf{T}}$ cands	± 27	± 28	± 30	0
bkg control sample	$+179$	-84	-12	-59
γj and $j\gamma$ PDF	± 5	± 12	± 13	± 7
Total	$^{+217}_{-120}$	$^{+144}_{-172}$	$^{+35}_{-38}$	$^{+18}_{-62}$

Table 5.13: Systematic uncertainties on the measured yields.

5.4.8 Differential spectra

The differential spectra $dN_{\gamma\gamma}/dX$, $X = M_{\gamma\gamma}$, $p_{\mathrm{T},\gamma\gamma}$ and $\Delta\phi_{\gamma\gamma}$, are obtained by performing the 2D fit in separate bins of the quantity X and by dividing the (**RobustTight**, isolated) di-photon yield in that bin by the bin width ΔX. The background PDFs (both the single jet PDF for the γj and $j\gamma$ components and the di-jet PDF) are determined from data, after reversing the **RobustTight** photon identification criteria, in the particular X bin under study (*i.e.* the quantity X is required to be in the bin under study for the di-photon candidates that fail the isEM cuts and are used to parameterize the background isolation distribution).

A summary of the bin by bin yield, statistical and systematic errors for the $M_{\gamma\gamma}$, $p_{\mathrm{T},\gamma\gamma}$ and $\Delta\phi_{\gamma\gamma}$ differential spectra is given in Tables 5.14- 5.16. The systematics considered are the same as in the inclusive analysis.

Figures 5.14-5.16 show the $M_{\gamma\gamma}$, $p_{\mathrm{T},\gamma\gamma}$ and $\Delta\phi_{\gamma\gamma}$ distributions. The N_{jj}, $N_{\gamma j+j\gamma}$, and $N_{\gamma\gamma}$ contributions are stacked and respectively dark blue, light blue and white. The statistical error bars are black. The total error computed as a quadratic combination of all systematics and statistical errors are red. Both the yields and the errors (Tables 5.14-5.16) are normalized to the bin width.

5.5 Other methods

Beside the 2D fit technique, two methods were developed in parallel to extract the di-photon yields.

	$W_{\gamma\gamma}^{\mathrm{TITI}}$	$W_{\gamma j}^{\mathrm{TITI}}$	$W_{j\gamma}^{\mathrm{TITI}}$	W_{jj}^{TITI}
$0 < M_{\gamma\gamma}(GeV) < 30$	$96 \pm 14^{+17}_{-12}$	$10 \pm 4^{+11}_{-12}$	$13 \pm 4^{+5}_{-6}$	$24 \pm 3^{+3}_{-6}$
$30 < M_{\gamma\gamma}(GeV) < 40$	$176 \pm 21^{+53}_{-14}$	$35 \pm 8^{+23}_{-33}$	$26 \pm 7^{+9}_{-16}$	$25 \pm 4^{+6}_{-8}$
$40 < M_{\gamma\gamma}(GeV) < 50$	$305 \pm 27^{+97}_{-26}$	$76 \pm 11^{+39}_{-59}$	$31 \pm 8^{+11}_{-21}$	$55 \pm 5^{+7}_{-15}$
$50 < M_{\gamma\gamma}(GeV) < 60$	$230 \pm 24^{+74}_{-21}$	$76 \pm 10^{+27}_{-41}$	$31 \pm 6^{+13}_{-18}$	$40 \pm 4^{+7}_{-13}$
$60 < M_{\gamma\gamma}(GeV) < 70$	$133 \pm 18^{+47}_{-15}$	$58 \pm 8^{+18}_{-26}$	$17 \pm 5^{+6}_{-9}$	$31 \pm 3^{+5}_{-11}$
$70 < M_{\gamma\gamma}(GeV) < 80$	$75 \pm 14^{+22}_{-7}$	$38 \pm 6^{+10}_{-15}$	$15 \pm 4^{+4}_{-5}$	$18 \pm 3^{+3}_{-6}$
$80 < M_{\gamma\gamma}(GeV) < 100$	$154 \pm 19^{+37}_{-17}$	$65 \pm 7^{+19}_{-26}$	$21 \pm 4^{+7}_{-9}$	$24 \pm 3^{+5}_{-8}$
$100 < M_{\gamma\gamma}(GeV) < 150$	$82 \pm 14^{+14}_{-7}$	$39 \pm 5^{+9}_{-9}$	$20 \pm 4^{+5}_{-6}$	$21 \pm 3^{+4}_{-7}$
$150 < M_{\gamma\gamma}(GeV) < 200$	$26 \pm 7^{+2}_{-2}$	$4 \pm 2^{+2}_{-2}$	$5 \pm 2^{+1}_{-1}$	$5 \pm 1^{+1}_{-1}$

Table 5.14: Measured yields for the $M_{\gamma\gamma}$ differential spectrum with their statistical and their asymmetrical systematic errors respectively

	$W_{\gamma\gamma}^{\mathrm{TITI}}$	$W_{\gamma j}^{\mathrm{TITI}}$	$W_{j\gamma}^{\mathrm{TITI}}$	W_{jj}^{TITI}
$0 < p_{\mathrm{T},\gamma\gamma}(GeV) < 10$	$573 \pm 34^{+147}_{-31}$	$77 \pm 13^{+59}_{-80}$	$39 \pm 9^{+11}_{-24}$	$89 \pm 5^{+7}_{-29}$
$10 < p_{\mathrm{T},\gamma\gamma}(GeV) < 20$	$311 \pm 28^{+114}_{-18}$	$133 \pm 12^{+34}_{-68}$	$36 \pm 7^{+9}_{-18}$	$59 \pm 4^{+7}_{-19}$
$20 < p_{\mathrm{T},\gamma\gamma}(GeV) < 30$	$168 \pm 20^{+27}_{-16}$	$60 \pm 7^{+18}_{-20}$	$31 \pm 5^{+7}_{-8}$	$29 \pm 3^{+5}_{-11}$
$30 < p_{\mathrm{T},\gamma\gamma}(GeV) < 40$	$101 \pm 17^{+23}_{-12}$	$48 \pm 7^{+14}_{-18}$	$27 \pm 5^{+6}_{-8}$	$15 \pm 2^{+4}_{-7}$
$40 < p_{\mathrm{T},\gamma\gamma}(GeV) < 50$	$66 \pm 13^{+8}_{-8}$	$30 \pm 4^{+8}_{-9}$	$16 \pm 4^{+3}_{-3}$	$9 \pm 2^{+2}_{-4}$
$50 < p_{\mathrm{T},\gamma\gamma}(GeV) < 60$	$52 \pm 10^{+12}_{-6}$	$15 \pm 3^{+6}_{-6}$	$8 \pm 2^{+2}_{-2}$	$6 \pm 1^{+1}_{-3}$
$60 < p_{\mathrm{T},\gamma\gamma}(GeV) < 80$	$32 \pm 9^{+5}_{-6}$	$17 \pm 3^{+5}_{-4}$	$6 \pm 2^{+2}_{-1}$	$3 \pm 1^{+1}_{-1}$
$80 < p_{\mathrm{T},\gamma\gamma}(GeV) < 100$	$12 \pm 5^{+4}_{-3}$	$3 \pm 2^{+3}_{-3}$	$5 \pm 2^{+2}_{-2}$	$1 \pm 1^{+1}_{-1}$
$100 < p_{\mathrm{T},\gamma\gamma}(GeV) < 150$	$1 \pm 6^{+4}_{-2}$	$6 \pm 2^{+2}_{-2}$	$4 \pm 1^{+2}_{-2}$	$0 \pm 0^{+1}_{-1}$
$150 < p_{\mathrm{T},\gamma\gamma}(GeV) < 200$	$1 \pm 1^{+0}_{-0}$	$1 \pm 1^{+0}_{-0}$	$1 \pm 0^{+0}_{-0}$	$0 \pm 0^{+0}_{-0}$

Table 5.15: Measured yields for the $p_{\mathrm{T},\gamma\gamma}$ differential spectrum with their statistical and asymmetrical systematic errors respectively

	$W_{\gamma\gamma}^{\mathrm{TITI}}$	$W_{\gamma j}^{\mathrm{TITI}}$	$W_{j\gamma}^{\mathrm{TITI}}$	W_{jj}^{TITI}
$0. < \Delta\phi_{\gamma\gamma} < 1$	$97 \pm 15^{+19}_{-17}$	$19 \pm 5^{+12}_{-13}$	$10 \pm 4^{+5}_{-2}$	$37 \pm 4^{+2}_{-19}$
$1. < \Delta\phi_{\gamma\gamma} < 2.$	$115 \pm 20^{+61}_{-18}$	$57 \pm 8^{+21}_{-34}$	$57 \pm 7^{+12}_{-28}$	$11 \pm 3^{+6}_{-7}$
$2. < \Delta\phi_{\gamma\gamma} < 2.5$	$160 \pm 21^{+55}_{-21}$	$54 \pm 8^{+24}_{-34}$	$35 \pm 7^{+8}_{-15}$	$29 \pm 4^{+5}_{-13}$
$2.5 < \Delta\phi_{\gamma\gamma} < 2.8$	$228 \pm 24^{+70}_{-20}$	$66 \pm 9^{+28}_{-41}$	$33 \pm 7^{+9}_{-16}$	$46 \pm 4^{+7}_{-19}$
$2.8 < \Delta\phi_{\gamma\gamma} < 3.0$	$318 \pm 27^{+84}_{-25}$	$98 \pm 11^{+36}_{-54}$	$26 \pm 6^{+8}_{-13}$	$74 \pm 5^{+7}_{-30}$
$3.0 < \Delta\phi_{\gamma\gamma} < 3.1416$	$365 \pm 26^{+80}_{-32}$	$108 \pm 10^{+38}_{-48}$	$22 \pm 5^{+6}_{-10}$	$72 \pm 5^{+7}_{-30}$

Table 5.16: Measured yields for the $\Delta\phi_{\gamma\gamma}$ differential spectrum with their statistical and asymmetrical systematic errors respectively

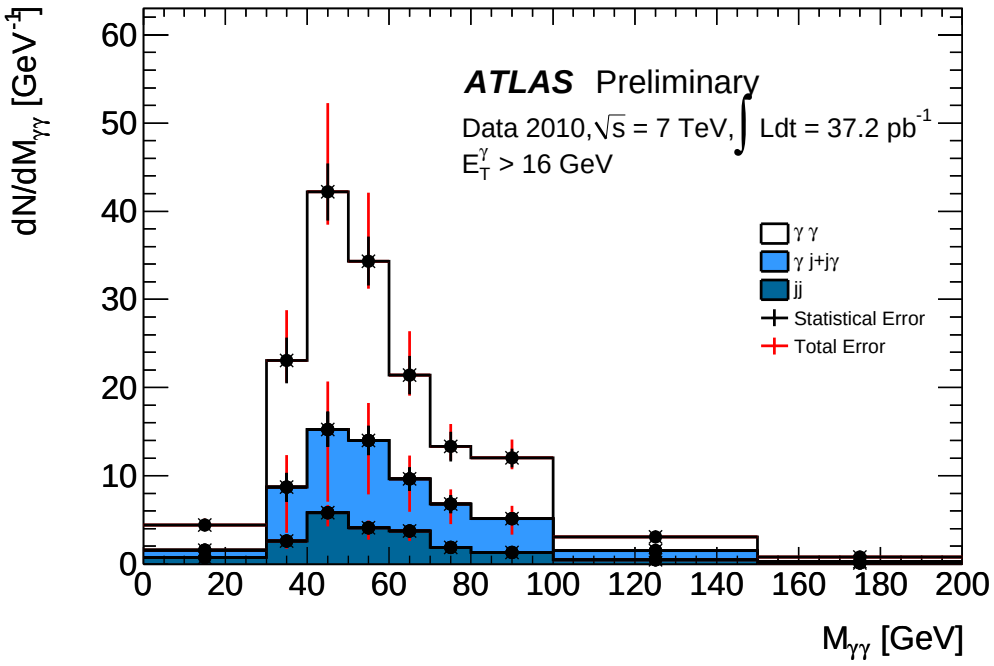


Figure 5.14: $M_{\gamma\gamma}$ differential spectrum for the N_{jj} (dark blue), $N_{\gamma j+j\gamma}$ (light blue), and $N_{\gamma\gamma}$ (white) contributions. Both statistical (black line) and total error (red line) are represented. Yields and errors are normalized to the bin width.

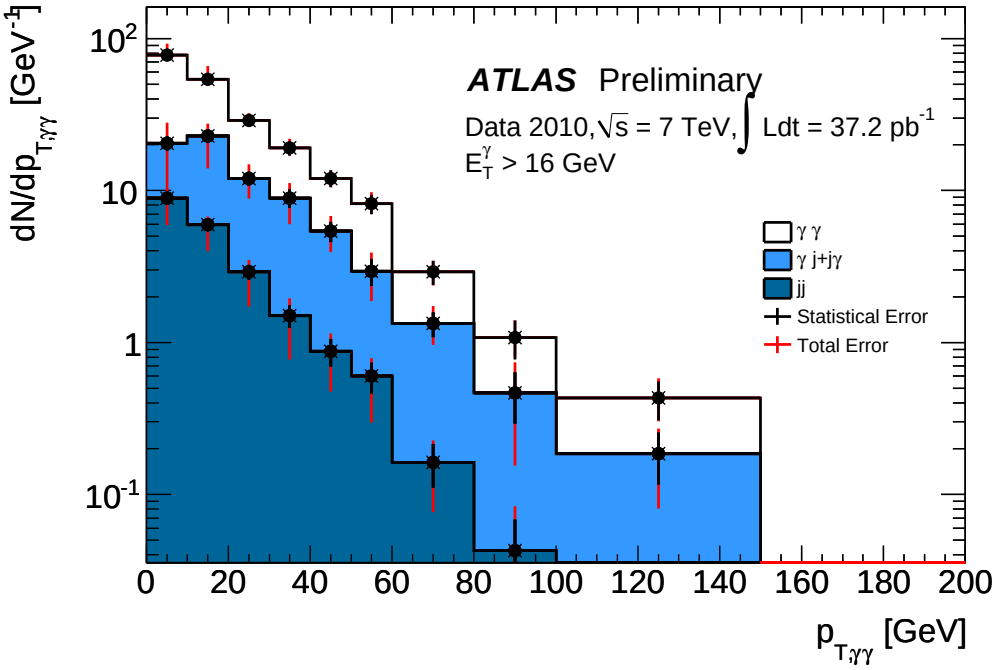


Figure 5.15: $p_{\mathrm{T},\gamma\gamma}$ differential spectrum for the N_{jj} (dark blue), $N_{\gamma j+j\gamma}$ (light blue), and $N_{\gamma\gamma}$ (white) contributions. Both statistical (black line) and total error (red line) are represented. Yields and errors are normalized to the bin width.

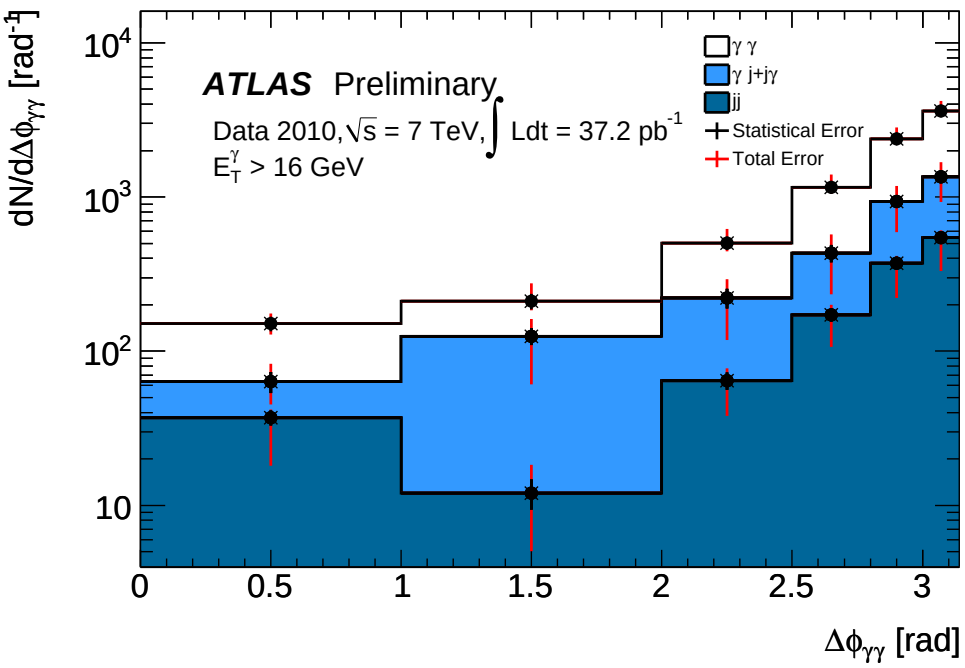


Figure 5.16: $\Delta\phi_{\gamma\gamma}$ differential spectrum for the N_{jj} (dark blue), $N_{\gamma j+j\gamma}$ (light blue), and $N_{\gamma\gamma}$ (white) contributions. Both statistical (black line) and total error (red line) are represented. Yields and errors are normalized to the bin width.

5.5.1 2×2D sideband

The 2-D sideband method has been introduced in ATLAS to measure the inclusive isolated photon cross-section [90]. It is based on the data-driven background measurement in a signal region (*e.g.* where photon candidates are well identified and isolated) by extrapolating the background from control regions located in the sidebands of identification and isolation variables forming a 2D-plane.

To extract the di-photon signal yields, this technique needs to be applied twice. The first step is applied on the leading p_T photon candidates. The 2D-plane for leading p_T photon candidates is illustrated on the top of Figure 5.17. There are three background control regions (M^A, M^B and N^B) and one signal photon enriched region (N^A). Assuming no correlation between the identification variables and the isolation transverse energy, and assuming no signal leakage into the background control regions, the number of background events in the signal enriched region can be derived as:

$$N^A_{\text{bkg}} = N^B_{\text{bkg}} \frac{M^A_{\text{bkg}}}{M^B_{\text{bkg}}}. \tag{5.10}$$

The number of signal events in region N^A reads:

$$\begin{aligned} N^A_{\text{sig}} &= N^A - N^B_{\text{bkg}} \frac{M^A_{\text{bkg}}}{M^B_{\text{bkg}}} \\ &= N^A - N^B \frac{M^A}{M^B} \end{aligned} \tag{5.11}$$

However, the signal leakage and the correlation between the identification variables and the isolation are not negligible, and need to be taken into account. Three parameters are

introduced for the signal leakage into the three background control regions: c_1, c_2 and c_3. c_1 indicates the leakage in the **T** and non-isolated region. $1+c_1$ is actually the inverse of the signal photon isolation efficiency. It can be extracted either with the signal photon isolation PDF derived from the electron sample or with the signal photon isolation distribution by subtracting the background from data as discussed later in Section 5.5.2. c_2 is actually connected with the tight identification efficiency, which is derived with Monte Carlo photon sample but the shower shape variables are corrected to match the distribution observed on data. c_3 is the product of the two parameters assuming negligible correlation between isolation and the tight selection. Another parameter θ is defined as:

$$\theta = \frac{M^B_{\rm bkg}}{M^A_{\rm bkg}} \frac{N^A_{\rm bkg}}{N^B_{\rm bkg}}. \tag{5.12}$$

It quantifies the correlation between the identification variables and isolation transverse energy. After including the four parameters mentioned above, the formula to extract the signal events in the signal enriched region becomes:

$$N^A_{\rm sig} = N^A - \left[(N^B - c_1 N^A_{\rm sig}) \frac{M^A - c_2 N^A_{\rm sig}}{M^B - c_3 N^A_{\rm sig}} \right] \theta. \tag{5.13}$$

It is a quadratic equation with one variable, which has one physical solution.

The same procedure is then applied to the subleading p_T photon candidates requiring that the leading photon is tightly identified and isolated. Afterwards, the signal yield $N'^A_{\rm sig}$ for subleading p_T photon in N'^A region is derived. The 2D-plane for subleading p_T photon candidates is illustrated on the bottom of Figure 5.17.

The two signal yields $N^A_{\rm sig}$ and $N'^A_{\rm sig}$ are used to extract the number of di-photon events for which both photon candidates are tightly identified and isolated. One additional constrain is needed, which is the fraction α of tightly identified and isolated di-photon candidates in the photon-jet sample where the leading p_T photon candidate is a true jet:

$$W^{\rm TITI}_{j\gamma} = \alpha (W^{\rm TITI}_{j\gamma} + W^{\rm TITI}_{\gamma j}). \tag{5.14}$$

Indeed, since the photon-jet contamination in the signal region is expected to be small, one does not expect a very large systematic uncertainty coming from this parameter on the final result.

Using this constraint, the solution for the yields reads:

$$\begin{aligned}
W^{\rm TITI}_{\gamma\gamma} &= \frac{\epsilon' \left(\alpha f' N^A_{\rm sig} + (\alpha - 1) N'^A_{\rm sig} \right)}{(\alpha - 1)\epsilon' + \alpha f'}, \\
W^{\rm TITI}_{\gamma j} &= \frac{(\alpha - 1) f' \left(\epsilon' N^A_{\rm sig} - N'^A_{\rm sig} \right)}{(\alpha - 1)\epsilon' + \alpha f'}, \\
W^{\rm TITI}_{j\gamma} &= \frac{\alpha f' \left(-\epsilon' N^A_{\rm sig} + N'^A_{\rm sig} \right)}{(\alpha - 1)\epsilon' + \alpha f'}, \\
W^{\rm TITI}_{jj} &= \frac{f' \left(\alpha \epsilon' f' N^A + (\alpha - 1)\epsilon'^2 \left(N^A - N^A_{\rm sig} \right) - \alpha f' N'^A_{\rm sig} \right)}{\epsilon' \left((\alpha - 1)\epsilon' + \alpha f' \right)}.
\end{aligned} \tag{5.15}$$

Where ϵ' is the efficiency of a subleading p_T true photon to pass the cuts when the leading p_T candidates has passed the cuts. It can be written as a combination of the signal

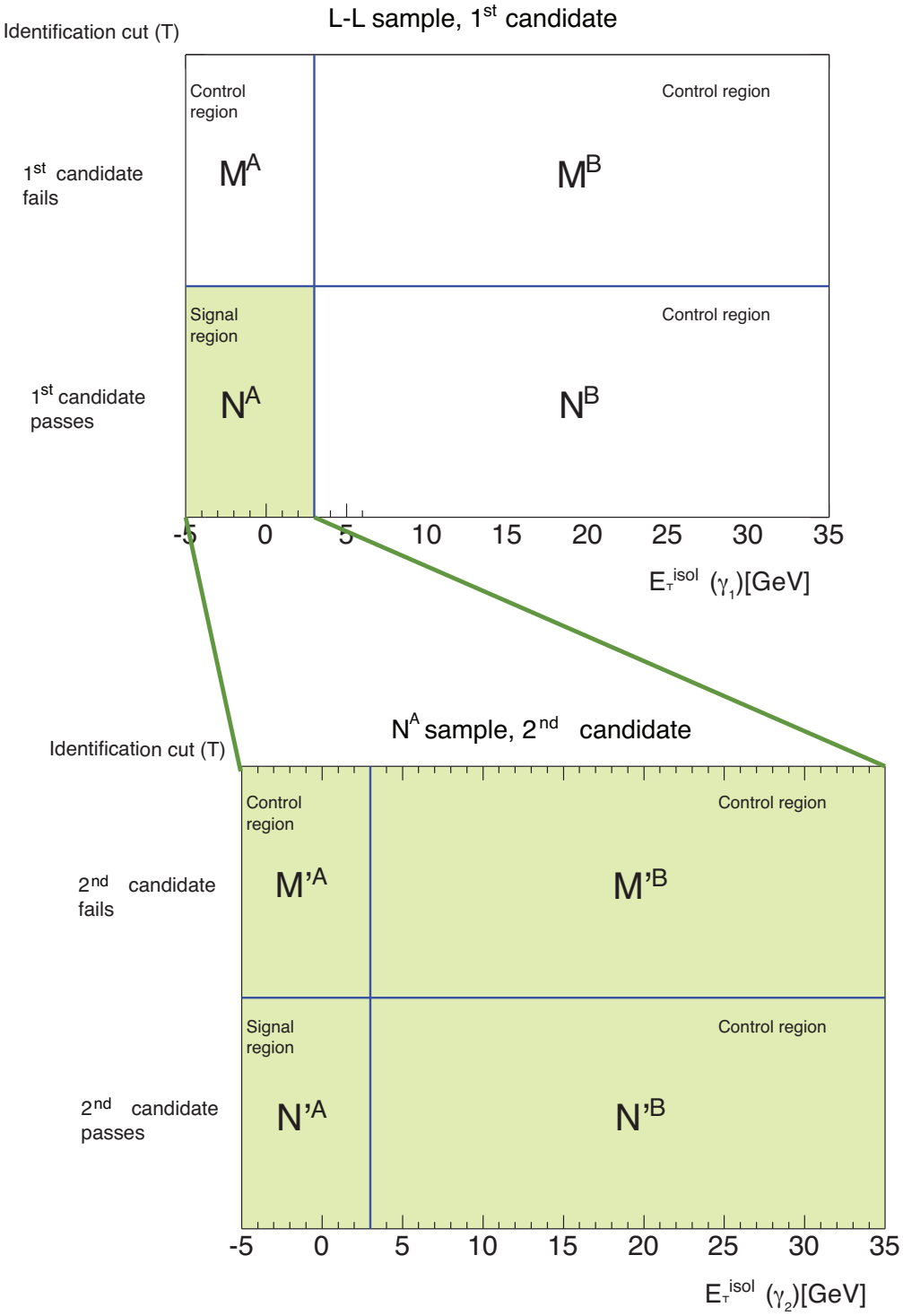


Figure 5.17: Schematic representation of the two-dimensional sideband method. The top plane is formed with the isolation (x-axis) and **Robust Tight** identification (y-axis) criteria for the leading photon, for all pairs of candidates passing the **Loose'** identification criterion. The three background regions are M^A, M^B, $N^B$and the signal region is N^A. Events from the N^Aregion are then passed to the bottom plane that works similarly than the first plane.

leakage coefficients which are obtained by data-driven method. f' is the fake rate of a subleading p_T jet to pass the cuts when the leading p_T candidate has passed the cut. It can be interpreted as:

$$f' = \frac{N'^A - N'^A_{\rm sig}}{N^A - \frac{1}{\epsilon'} N'^A_{\rm sig}}, \tag{5.16}$$

The correlation coefficients θ (Equation 5.12) and the α parameter (Equation 5.14) are not measured on data. In the results given below, θ is set to 1, assuming no correlation, whereas α is taken from Monte Carlo. Both are varied to estimate the associated systematic uncertainty.

Additional systematics come from the definition of the **Loose'** selection and from the signal leakage into the background control region.

5.5.2 4×4 matrix

This method has been developed in the past by CDF [34] and DØ [35] collaborations. The core component in this method is the 4×4 efficiency matrix. In ATLAS, the criterion used to define the efficiency matrix is the calorimetric isolation as mentioned in Section 5.3.1.

The analysis is performed on the **TT** region. The calorimetric isolation requirement $E_T^{\rm iso} < 3$ GeV is then used to classify the events into four categories, depending on whether the leading p_T or subleading p_T photon candidate passes (labelled as P) or fails (labelled as F) the requirement. $N_{\rm PP}, N_{\rm PF}, N_{\rm FP}, N_{\rm FF}$[9] are the number of events in each category.

The number of events for each component $W^{\rm TT}_{\gamma\gamma}$, $W^{\rm TT}_{\gamma j}$, $W^{\rm TT}_{j\gamma}$, $W^{\rm TT}_{jj}$can then be inferred, by solving the linear system of four equations:

$$\begin{pmatrix} N_{\rm PP} \\ N_{\rm PF} \\ N_{\rm FP} \\ N_{\rm FF} \end{pmatrix} = {\rm E} \begin{pmatrix} W^{\rm TT}_{\gamma\gamma} \\ W^{\rm TT}_{\gamma j} \\ W^{\rm TT}_{j\gamma} \\ W^{\rm TT}_{jj} \end{pmatrix} \tag{5.17}$$

Here E is the 4×4 matrix, whose coefficients are combination of the photon efficiency and the jet fake rate after the calorimetric isolation requirement with respect to the **RobustTight** selection. Both the photon efficiency and jet fake rate can be derived by data-driven method. The jet fake rate is measured directly from the isolation transverse energy distribution of the $\tilde{\mathbf{T}}$ background control sample, assuming, as in the 2D fit method, that it is similar to the jet distribution in the **T** region. The **T** sample is composed of signal photon and background jet. The background jet isolation transverse energy distribution in the **T** region can be subtracted using the $\tilde{\mathbf{T}}$ distribution. The normalization factor is taken form the $E_T^{\rm iso} > 7$ GeV region, where both distributions are largely dominated by the background. After the subtraction of the background jet component, one is left with the distribution of pure signal photons, from which the signal photon efficiency can be extracted.

The matrix E is further refined to take into account the correlation between the two photon candidates:

$${\rm E}_{ab(a=1,4;b=1,2)} = \begin{pmatrix} \frac{1}{2}\left[\epsilon_1\epsilon_2^{\hat{\rm P}} + \epsilon_1^{\hat{\rm P}}\epsilon_2\right] & \frac{1}{2}\left[\epsilon_1 f_2^{\hat{\rm P}} + \epsilon_1^{\hat{\rm P}} f_2\right] \\ \frac{1}{2}\left[\epsilon_1\left(1-\epsilon_2^{\hat{\rm P}}\right) + \epsilon_1^{\hat{\rm F}}(1-\epsilon_2)\right] & \frac{1}{2}\left[\epsilon_1\left(1-f_2^{\hat{\rm P}}\right) + \epsilon_1^{\hat{\rm F}}(1-f_2)\right] \\ \frac{1}{2}\left[(1-\epsilon_1)\epsilon_2^{\hat{\rm F}} + \left(1-\epsilon_1^{\hat{\rm P}}\right)\epsilon_2\right] & \frac{1}{2}\left[(1-\epsilon_1) f_2^{\hat{\rm F}} + \left(1-\epsilon_1^{\hat{\rm P}}\right) f_2\right] \\ \frac{1}{2}\left[(1-\epsilon_1)\left(1-\epsilon_2^{\hat{\rm F}}\right) + \left(1-\epsilon_1^{\hat{\rm F}}\right)(1-\epsilon_2)\right] & \frac{1}{2}\left[(1-\epsilon_1)\left(1-f_2^{\hat{\rm F}}\right) + \left(1-\epsilon_1^{\hat{\rm F}}\right)(1-f_2)\right] \end{pmatrix} \tag{5.18}$$

[9] The 1st index is meant to denote the leading p_T photon candidates, whereas the 2nd index is for the subleading p_T photon candidates.

$$\mathrm{E}_{ab(a=1,4;b=3,4)} = \begin{pmatrix} \frac{1}{2}\left[f_1\epsilon_2^{\hat{\mathrm{P}}} + f_1^{\hat{\mathrm{P}}}\epsilon_2\right] & \frac{1}{2}\left[f_1 f_2^{\hat{\mathrm{P}}} + f_1^{\hat{\mathrm{P}}} f_2\right] \\ \frac{1}{2}\left[f_1\left(1-\epsilon_2^{\hat{\mathrm{P}}}\right) + f_1^{\hat{\mathrm{F}}}\left(1-\epsilon_2\right)\right] & \frac{1}{2}\left[f_1\left(1-f_2^{\hat{\mathrm{P}}}\right) + f_1^{\hat{\mathrm{F}}}\left(1-f_2\right)\right] \\ \frac{1}{2}\left[\left(1-f_1\right)\epsilon_2^{\hat{\mathrm{F}}} + \left(1-f_1^{\hat{\mathrm{P}}}\right)\epsilon_2\right] & \frac{1}{2}\left[\left(1-f_1\right) f_2^{\hat{\mathrm{F}}} + \left(1-f_1^{\hat{\mathrm{P}}}\right) f_2\right] \\ \frac{1}{2}\left[\left(1-f_1\right)\left(1-\epsilon_2^{\hat{\mathrm{F}}}\right) + \left(1-f_1^{\hat{\mathrm{F}}}\right)\left(1-\epsilon_2\right)\right] & \frac{1}{2}\left[\left(1-f_1\right)\left(1-f_2^{\hat{\mathrm{F}}}\right) + \left(1-f_1^{\hat{\mathrm{F}}}\right)\left(1-f_2\right)\right] \end{pmatrix} \tag{5.19}$$

The efficiencies ϵ_i and the fake rates f_i are computed separately for the leading ($i = 1$) and sub-leading ($i = 2$) candidate. The P and F upperscripts indicate whether the other candiate passes or fails the isolation cut. In practice, the efficiencies and fake rates are computed as a function of the pseudorapidity. A vector ($N_{\mathrm{PP}}, N_{\mathrm{PF}}, N_{\mathrm{FP}}, N_{\mathrm{FF}}$) is associated to each event with $N_{XX} = 0$ or 1, depending on which of the four categories the event belongs to. ($W_{\gamma\gamma}^{\mathrm{TT}}$, $W_{\gamma j}^{\mathrm{TT}}$, $W_{j\gamma}^{\mathrm{TT}}$, W_{jj}^{TT}) is then a vector of weights giving the probabilities of a given event to belong to each of the four components. The total yields are obtained by summing the weights of all events, or events in a given $M_{\gamma\gamma}$, $p_{T\gamma\gamma}$or $\Delta\phi_{\gamma\gamma}$ bin.

The main systematics come from the statistical uncertainties on the efficiencies and fake rates, as well as the definition of the non-isolated and $\tilde{\mathbf{T}}$ region.

For this method, the main systematic uncertainties includes:

- Uncertainties in the coefficients that enter the matrix.
- Definition of **Loose'** and definition of non-isolated region.

5.5.3 Comparison of the three methods

The way to subtract the background isolation transverse energy distribution in the **T** sample for all the three methods is principally the same. It is based on the assumption that the background jet isolation transverse energy distribution in **T** can be illustrated with the one in $\tilde{\mathbf{T}}$. For the signal photon isolation distribution extraction, the 4×4 matrix method directly makes use of the signal photon candidates distribution after the subtraction of the background distribution, whereas the 2D fit method extrapolates from the electron distribution. The three methods give compatible results. The overall signal and background yields in the **TITI** region are summarized in Table 5.5.3.

event category	2D fit	2×2D sideband	4×4 matrix
$W_{\gamma\gamma}^{\mathrm{TITI}}$	1353 ± 53	1325 ± 75	1239 ± 60
$W_{\gamma j}^{\mathrm{TITI}}$	348 ± 19	420 ± 36	419 ± 23
$W_{j\gamma}^{\mathrm{TITI}}$	168 ± 14	136 ± 12	170 ± 21
W_{jj}^{TITI}	174 ± 7	142 ± 10	193 ± 9

Table 5.17: Signal and background yields together with the statistical uncertainties in the **TITI** sample obtained from the 2D fit, 2×2D sideband and 4×4 matrix method.

As discussed in the previous sections, the definition of $\tilde{\mathbf{T}}$ is a common systematic uncertainty. It is also the dominant uncertainty for all three methods. Table 5.5.3 summarizes the

	2D fit	2×2D sideband	4×4 matrix
$W_{\gamma\gamma}^{\text{TITI}}$	$1353 \pm 53(\text{stat})^{+217}_{-120}(\text{syst})$	$1325 \pm 75(\text{stat})^{+116}_{-146}(\text{syst})$	$1239 \pm 60(\text{stat})^{+186}_{-163}(\text{syst})$

Table 5.18: Signal yields in the **TITI** sample, the statistical and systematic uncertainties obtained from the 2D fit, 2×2D sideband and 4×4 matrix method.

di-photon yields from each of the three methods including their statistical and systematic uncertainties.

The comparison on the differential yields is displayed in Figure 5.18, for the three observables $M_{\gamma\gamma}$, $p_{T\gamma\gamma}$, $\Delta\phi_{\gamma\gamma}$. Good agreement is observed.

The three methods have in common the usage of shower shape variables and calorimetric isolation energy, as variables to discriminate the photon signal from the jet background, and the data-driven evaluation of the jet background properties, relying on the "non-**RobustTight**" control sample defined in 5.3.2 and 5.4.3. For this reason, they cannot be considered as completely independent, and therefore their outcomes cannot be combined.

5.6 Differential cross-section

The goal of this study is to measure the di-photon production differential cross-sections. The observables considered here are $M_{\gamma\gamma}$, $p_{T\gamma\gamma}$ and $\Delta\phi_{\gamma\gamma}$. The formula to compute the differential cross-section can be illustrated as:

$$\frac{d\sigma}{dX} = \frac{1}{\mathcal{L}} \frac{dn(X^{\text{true}})}{dX^{\text{true}}} \tag{5.20}$$

Here X means for any observable as $M_{\gamma\gamma}$, $p_{T\gamma\gamma}$ or $\Delta\phi_{\gamma\gamma}$. The differential yield dn of true di-photon events is needed, as a function of the true quantity X^{true}. The true di-photon events are defined as the events with two photons in the final state that can be either direct or fragmentation photon. Both photons must satisfy the following requirements:

- $p_T > 16$ GeV and $|\eta| < 2.37$, with the exclusion of the region $1.37 < |\eta| < 1.52$.
- $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} > 0.4$.
- a partonic isolation $E_T^{iso(part)} < 4$ GeV in a cone $R < 0.4$ around each photon.[10]

The yields extracted in Section 5.4.8 are reconstructed yields, they need to be interpreted to the yields of the true quantity. In order to connect the yields of the true quantity with the reconstructed yields, the following formula is defined:

$$N_i^{\mathbf{TITI},trigger} = \underbrace{\frac{N_i^{\mathbf{TITI},trigger}}{N_i^{\mathbf{TITI}}}}_{\epsilon_i^{trigger}} \underbrace{\frac{N_i^{\mathbf{TITI}}}{N_i^{\mathbf{II}}}}_{\epsilon_i^{\mathbf{TT}}} \sum_\alpha \left(\underbrace{\text{Prob}(X_i^{\text{rec}}|X_\alpha^{\text{true}})}_{M_{i\alpha}} \underbrace{\frac{N_\alpha^{\mathbf{II}}}{n_\alpha}}_{\epsilon_\alpha^{rec}} n_\alpha \right) \tag{5.21}$$

where:

[10] In RESBOS and DIPHOX, an upper limit is imposed on a transverse momentum carried by the remnants of the fragmenting parton when it is inside the isolation cone (e.g $E_{T_{max}} < 15$ GeV). In other generators such as PYTHIA, there is no such variable. Thus, the transverse momenta of all true particles in the cone are summed up, excluding the contribution from neutrinos and muons, and removing the underlying event based on the energy density measured from soft truth jets.

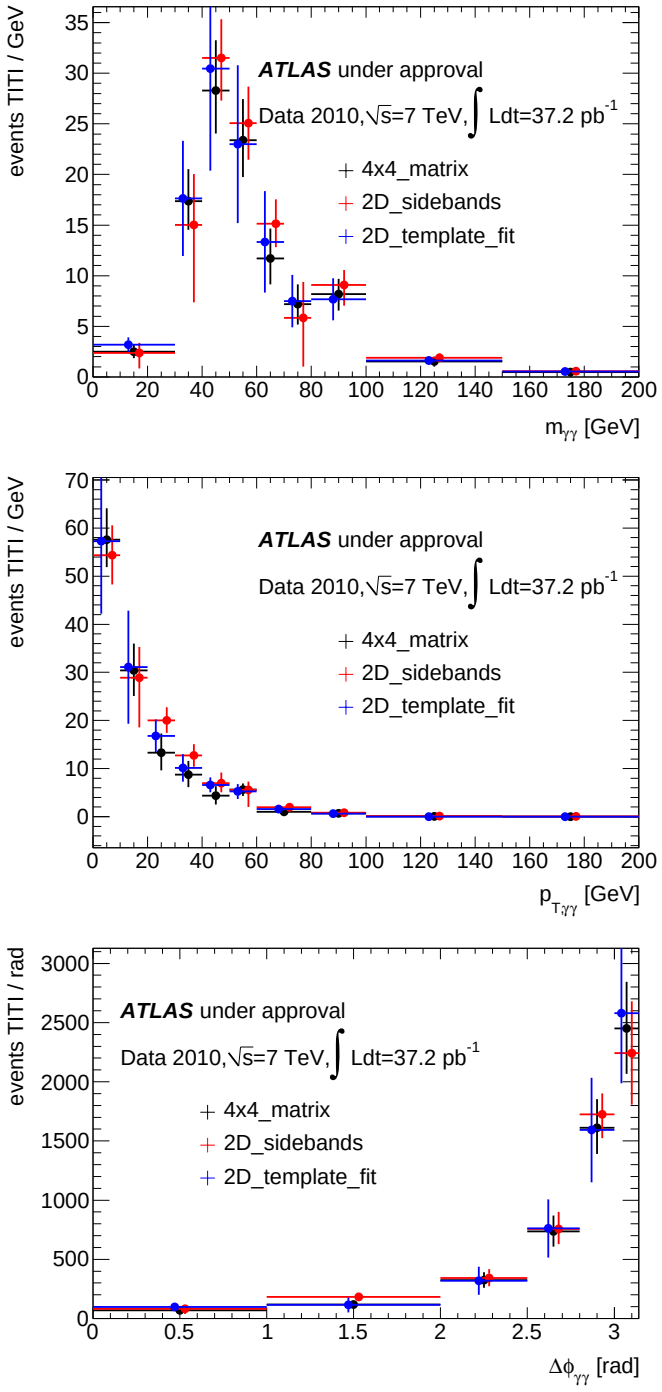


Figure 5.18: Differential yields in the **TITI** sample, as a function of the three observables $M_{\gamma\gamma}$, $p_{T\gamma\gamma}$, $\Delta\phi_{\gamma\gamma}$, obtained with the three methods. The vertical error bars display the total errors, accounting for both the statistical uncertainties and the systematic effects. The bars are artificially shifted horizontally, to better display the three outcomes.

- n_α is the number of di-photon events produced in α-th bin of the variable X^{true}, that satisfy the requirements defined above, including all kinematic, separation and isolation requirements.
- ϵ_α^{rec} is the reconstruction efficiency in α-th bin: it accounts for all p_T^{clus}, $|\eta_{\text{S2}}|$, ΔR cuts, OQ map, and calorimetric isolation.
- $M_{i\alpha}$ is the migration matrix, i.e. probability that a true value X_α^{true} be measured as X_i^{rec}.
- $\epsilon_i^{\mathbf{TT}}$ is the tight photon selection (**RobustTight**) efficiency of the photon-pair, in i-th bin, evaluated for photons passing the calorimetric isolation cut.
- $\epsilon_i^{trigger}$ is the trigger efficiency, for photon-pairs passing the calorimetric isolation and the tight photon selection cuts.
- $N_i^{\mathbf{TITI},trigger}$ is the pure reconstructed di-photon yields in i-th bin.

Equation 5.21 can be reversed to obtain:

$$n_\alpha = \frac{1}{\epsilon_\alpha^{rec}} \sum_i \left(U_{\alpha i} \frac{N_i^{\mathbf{TITI},trigger}}{\epsilon_i^{trigger} \epsilon_i^{\mathbf{TT}}} \right) \tag{5.22}$$

where $U_{\alpha i}$ is the "unfolding" matrix — naively $\mathbf{U} = \mathbf{M}^{-1}$. The differential signal yields $W_{\gamma\gamma}^{\text{TITI}}$ extracted with the methods as discussed in previous sections still contains the electron background that fakes photons. To proceed on, the background from electrons needs to be extracted from $W_{\gamma\gamma}^{\text{TITI}}$to get the pure di-photon yields $N_i^{\mathbf{TITI},trigger}$. Besides, the other elements such as ϵ_α^{rec}, $U_{\alpha i}$, $\epsilon_i^{\mathbf{TT}}$ and $\epsilon_i^{trigger}$ also need to be derived. All of those elements have been evaluated within the ATLAS di-photon working group.

5.6.1 Extract the background from electrons

As mentioned in Section 5.1.2, the physics processes to ee final states (Drell-Yan, $Z \to ee$, $WW \to e\nu e\nu$) and to $e\gamma$ final states ($\gamma W \to \gamma e\nu$, $\gamma Z \to \gamma ee$) can contribute to the di-photon events, due to the fact that isolated electrons are misidentified as isolated photons. Those events need to be subtracted from the $W_{\gamma\gamma}^{\text{TITI}}$ yields in order to obtain the pure di-photon events. In the two-body final state with electrons or photons, the following matrix equation can be built up:

$$\begin{pmatrix} N_{\gamma\gamma} \\ N_{\gamma e} \\ N_{ee} \end{pmatrix} = \begin{pmatrix} 1 & f_{e\to\gamma} & (f_{e\to\gamma})^2 \\ 2f_{\gamma\to e} & (1 + f_{e\to\gamma} f_{\gamma\to e}) & 2f_{e\to\gamma} \\ (f_{\gamma\to e})^2 & f_{\gamma\to e} & 1 \end{pmatrix} \times \begin{pmatrix} N_{\gamma\gamma}^{corr} \\ N_{\gamma e}^{corr} \\ N_{ee}^{corr} \end{pmatrix} \tag{5.23}$$

Here, N_{xy} are the number of final states experimentally classified as xy that pass the **TITI** selection. Therefore $N_{\gamma\gamma} = W_{\gamma\gamma}^{\text{TITI}}$, as defined in Section 5.4.1. N_{xy}^{corr} are the number of final state that correctly classified as the true state passing the **TITI** selection. $f_{e\to\gamma}$ and $f_{\gamma\to e}$ are defined as the ratio between objects passing the tight identification and isolation selection that are wrongly classified and those correctly classified. They can be expressed as follows:

$$f_{e\to\gamma} = \frac{N_{e\to\gamma}}{N_{e\to e}} \tag{5.24}$$

$$f_{\gamma\to e} = \frac{N_{\gamma\to e}}{N_{\gamma\to\gamma}} \tag{5.25}$$

Here $N_{e\to\gamma}$ is the number of true electrons reconstructed as photons and passing the photon **RobustTight** and isolation selection. $N_{e\to e}$ is the number of true electrons that correctly reconstructed as electrons and passing the electron **RobusterTight** and isolation selection. Similar definition applies for $N_{\gamma\to e}$ and $N_{\gamma\to\gamma}$. Equation 5.23 can be inverted as:

$$\begin{pmatrix} N_{\gamma\gamma}^{corr} \\ N_{\gamma e}^{corr} \\ N_{ee}^{corr} \end{pmatrix} = \frac{1}{(1-f_{e\to\gamma}f_{\gamma\to e})^2} \begin{pmatrix} 1 & -f_{e\to\gamma} & (f_{e\to\gamma})^2 \\ -2f_{\gamma\to e} & (1+f_{e\to\gamma}f_{\gamma\to e}) & -2f_{e\to\gamma} \\ (f_{\gamma\to e})^2 & -f_{\gamma\to e} & 1 \end{pmatrix} \times \begin{pmatrix} N_{\gamma\gamma} \\ N_{\gamma e} \\ N_{ee} \end{pmatrix} \tag{5.26}$$

With the inputs of $f_{e\to\gamma}$ and $f_{\gamma\to e}$, one can extract the number of correctly identified $\gamma\gamma$ events after subtraction of the $e\gamma$ and ee background:

$$N_{\gamma\gamma}^{corr} = \frac{N_{\gamma\gamma} - \left[f_{e\to\gamma}N_{\gamma e} - (f_{e\to\gamma})^2 N_{ee}\right]}{(1-f_{e\to\gamma}f_{\gamma\to e})^2} \tag{5.27}$$

The parameter $f_{e\to\gamma}$ is derived using reconstructed ee, γe and $\gamma\gamma$ final states falling in a ± 5 GeV window around the Z mass. Since the true $\gamma\gamma$ and γe final states in that window are negligible, one can derive $f_{e\to\gamma}$ from Equation 5.23 as:

$$f_{e\to\gamma} = \frac{N_{\gamma e}^{Z\ mass}}{2N_{ee}^{Z\ mass}} \tag{5.28}$$

where $N_{\gamma e}^{Z\ mass}$ and $N_{ee}^{Z\ mass}$ are the numbers of reconstructed γe and ee final states in the Z mass window. Table 5.19 shows the values of $f_{e\to\gamma}$ measured with the data driven technique and MC truth on the simulation samples, as well as the value measured from real data. The good agreement between the data driven value and the true value validates the method. However, the values differ among the two different simulation samples and the real data.

sample	data driven $f_{e\to\gamma}$	MC truth $f_{e\to\gamma}$
JF17 MC	0.180 ± 0.041	0.183
$Z \to ee$ MC	0.132 ± 0.0004	0.129
data	0.11206 ± 0.0046	

Table 5.19: $f_{e\to\gamma}$ value estimated with the data driven method compared to the one coming from the MC truth.

The parameter $f_{\gamma\to e}$ can in principle also be obtained with data-driven method using the radiation decay of Z into $ee\gamma$. However, due to the low statistics, it is impossible to derive this parameter in this way. Thus, it is obtained from the Monte Carlo simulations.

For each observable of $M_{\gamma\gamma}$, $p_{T\gamma\gamma}$ and $\Delta\phi_{\gamma\gamma}$, three histograms with the reconstructed final states of ee, γe and $\gamma\gamma$ passing the **TITI** selection are first obtained. For each bin, with the entries of $N_{\gamma\gamma}$, $N_{\gamma e}$ and N_{ee}, together with $f_{e\to\gamma}$ and $f_{\gamma\to e}$, the number of correctly identified true di-photon yield is then derived by the Equation 5.27.

Figure 5.19 shows the distribution of the electron background as a function of the three observables $M_{\gamma\gamma}$, $p_{T\gamma\gamma}$ and $\Delta\phi_{\gamma\gamma}$. The impurities are defined as follows. For the total electron background:

$$I_e = 1 - \frac{N_{\gamma\gamma}^{corr}}{N_{\gamma\gamma}} \tag{5.29}$$

And for the γe and ee background:

$$I_{\gamma e} = \frac{f_{\gamma\to e} N_{\gamma e}^{corr}}{N_{\gamma\gamma}} \tag{5.30}$$

$$I_{ee} = \frac{f_{\gamma\to e}^2 N_{ee}^{corr}}{N_{\gamma\gamma}} \tag{5.31}$$

The total impurity from γe and ee events is on average smaller than 5%. The only exception with a high impurity is in the invariant mass bin of 80∼100 GeV, which is expected from the Z mass peak.

5.6.2 Trigger efficiency

The efficiency $\epsilon_i^{trigger}$ is measured with real data. In this analysis, the 2g15_loose trigger is applied.

The object-level trigger efficiency is defined as the fraction of tightly identified and isolated photon candidates associated with a trigger object (*i.e.* passing trigger requirements). It is estimated by the bootstrap method, in which the measurement is done on a sample selected with a lower threshold, unbiased trigger (here, L1_EM5). The efficiencies are calculated for photon candidates[11] with transverse energy above 16 GeV.

A comparison between data and simulation for the object-level trigger efficiencies (L1_EM5 and g15_loose) as a function of the transverse energy is shown in Figure 5.20.

For this analysis, the 2g15_loose event-level trigger efficiency needs to be determined. An event-level efficiency is defined as the fraction of di-photon signal candidate events[12] that pass the trigger. Due to the high prescale factor applied to L1_EM5, the event-level efficiency for 2g15_loose has to be estimated in an alternative way to the bootstrap method. The event-level L1_EM5 efficiency ϵ_0 is first measured on a sample of minimum-bias triggers and found to be $100^{+0.00}_{-0.06}$%, for events with at least one photon candidate with transverse energy above 16 GeV. The object-level g15_loose efficiency ϵ_1 for photon candidates with transverse energy above 16 GeV is then measured on an unbiased sample triggered by L1_EM5 and found to be $99.14^{+0.31}_{-0.40}$%. Di-photon events with the subleading photon matched to a g15_loose trigger object are then selected[13]. The fraction ϵ_2 of events from this sample passing 2g15_loose is found to be $99.35^{+0.47}_{-0.96}$%. The final event-level trigger efficiency for 2g15_loose is eventually determined as $\epsilon_0 \times \epsilon_1 \times \epsilon_2$, and amounts to:

$$\epsilon^{trigger} = 98.50^{+0.56}_{-1.04} \pm 1\% \tag{5.32}$$

The systematic uncertainty of 1% is propagated from the electron trigger efficiency that is measured with tag&probe method with W or Z events in the egamma trigger group. Checks with Monte Carlo photon samples give consistent result.

[11] The efficiencies quoted in this section are relative to the photon **RobustTight** identification and calorimetric isolation selection.

[12] The di-photon signal candidate events are the events passing the **TITI** selection.

[13] The subleading photon is preferred to the leading one for its p_T spectrum similarity to the g15_loose objects found in the L1_EM5-triggered sample.

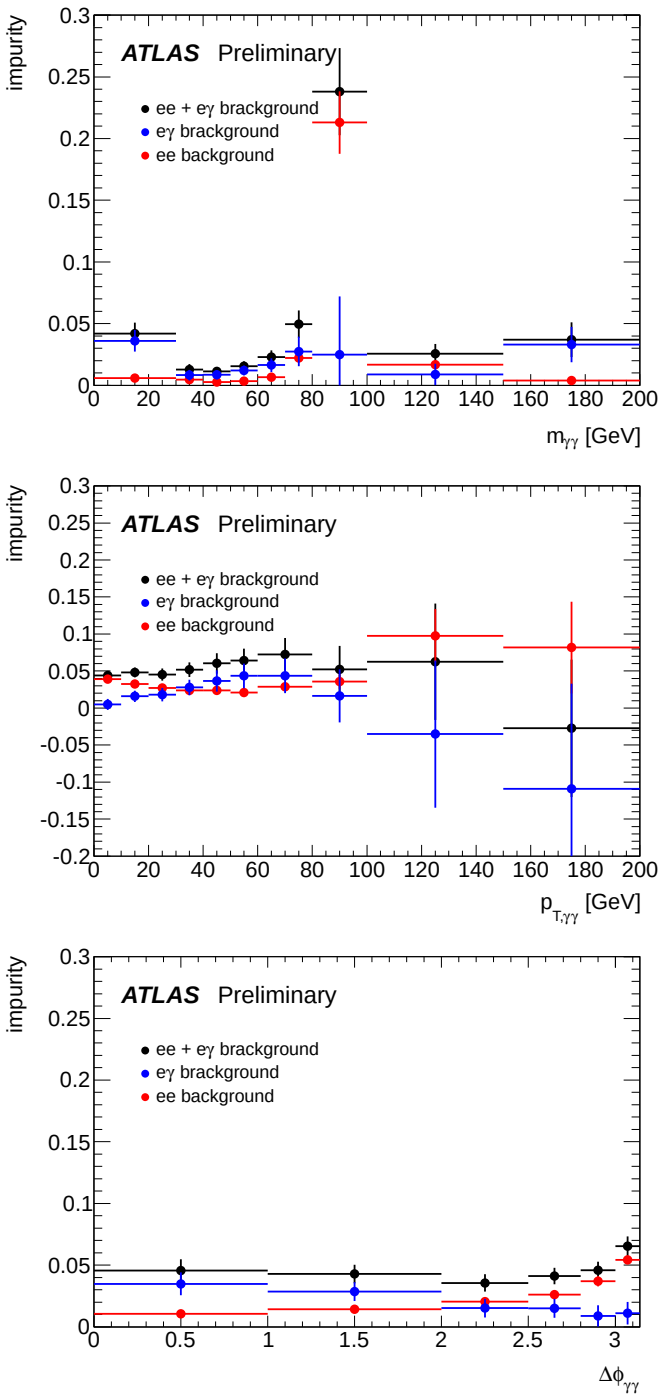


Figure 5.19: The impurity from γe and ee sources as well as the total impurity as a function of the observables $M_{\gamma\gamma}$, $p_{T\gamma\gamma}$ and $\Delta\phi_{\gamma\gamma}$.

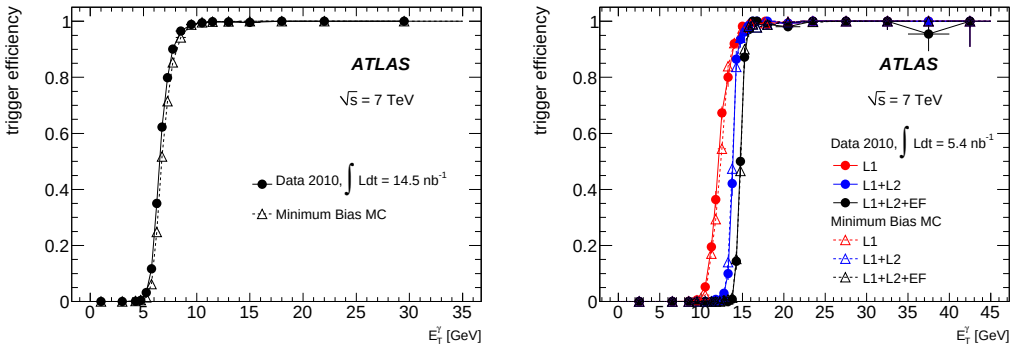


Figure 5.20: Trigger efficiency with respect to tightly identified and isolated photon candidates as a function of the transverse energy measured in data (circles) and simulated signal and background events (triangles) for L1_EM5 and g15_loose. The luminosities quoted on the figures correspond to the recorded luminosities in the samples used to compute the efficiencies.

5.6.3 Identification efficiency

For a given bin X_i^{rec}, the number of events passing **TT** and **TITI** selections are respectively expressed as:

$$\begin{aligned} N_i^{\mathbf{TT}} &= \sum_{k \in X_i^{\mathrm{rec}}} W_{\gamma\gamma}^{\mathrm{TT}(k)} \\ N_i^{\mathbf{TITI}} &= \sum_{k \in X_i^{\mathrm{rec}}} \left(\epsilon_1^{\mathbf{I}|\mathbf{T}^{(k)}} \epsilon_2^{\mathbf{I}|\mathbf{T}^{(k)}} W_{\gamma\gamma}^{\mathrm{TT}(k)} \right) \end{aligned} \tag{5.33}$$

where $\epsilon_1^{\mathbf{I}|\mathbf{T}^{(k)}}, \epsilon_2^{\mathbf{I}|\mathbf{T}^{(k)}}$ are the efficiencies of the isolation cut, applied to the leading and subleading photons, that pass the **RobustTight** selection. These efficiencies can be obtained using the signal photon isolation transverse energy PDFs discussed in Section 5.4.2.

The yield of di-photon events that pass the isolation selection can be written as:

$$N_i^{\mathbf{II}} = \sum_{k \in X_i^{\mathrm{rec}}} \left(\frac{\epsilon_1^{\mathbf{I}|\mathbf{T}^{(k)}} \epsilon_2^{\mathbf{I}|\mathbf{T}^{(k)}} W_{\gamma\gamma}^{\mathrm{TT}(k)}}{\epsilon_1^{\mathbf{T}|\mathbf{I}^{(k)}} \epsilon_2^{\mathbf{T}|\mathbf{I}^{(k)}}} \right) \tag{5.34}$$

The event identification efficiency can then be derived by taking the ratio of Equations 5.33 and 5.34.

$$\epsilon_i^{\mathbf{TT}} = \frac{N_i^{\mathbf{TITI}}}{N_i^{\mathbf{II}}} = \frac{\sum_{k \in X_i^{\mathrm{rec}}} \left(\epsilon_1^{\mathbf{I}|\mathbf{T}^{(k)}} \epsilon_2^{\mathbf{I}|\mathbf{T}^{(k)}} W_{\gamma\gamma}^{\mathrm{TT}(k)} \right)}{\sum_{k \in X_i^{\mathrm{rec}}} \left(\frac{\epsilon_1^{\mathbf{I}|\mathbf{T}^{(k)}} \epsilon_2^{\mathbf{I}|\mathbf{T}^{(k)}} W_{\gamma\gamma}^{\mathrm{TT}(k)}}{\epsilon_1^{\mathbf{T}|\mathbf{I}^{(k)}} \epsilon_2^{\mathbf{T}|\mathbf{I}^{(k)}}} \right)}$$

At this stage, the unknown ingredient to compute ϵ_i^{TT} is the photon identification efficiency $\epsilon^{\mathbf{T}|\mathbf{I}}$ for photons passing the isolation selection. It is calculated using Monte Carlo samples, after correction of the shower shape variables used in the **RobustTight** selection

for the difference observed between data and simulation samples. The correction factors are obtained comparing the shower shape distributions in data and simulation after applying the **RobustTight** selection criteria. They are separately obtained for unconverted and converted candidates, for the different η and E_T bins. The main systematic uncertainties for the photon identification efficiency include:

- The variation of the efficiency in the different η and E_T bins. This uncertainty varies from 1% at high E_T to 5% at low E_T.
- The choice of the photon candidate sample to derive the correction factors. The nominal factors are obtained by comparing the data and the simulation with the candidates passing the **RobustTight** selection.
- The knowledge of the material for the detector in the simulation.
- The relative fraction of direct and fragmentation photons.
- The classification between converted and unconverted photons.

The total systematic uncertainties for the photon identification efficiency is evaluated to be below 2% for small $|\eta|$ and no more than 3% for large $|\eta|$. The overall systematic uncertainty for the event identification efficiency ϵ_i^{TT} is found to be dominated by the single photon identification efficiency. Figure 5.21 displays the event identification efficiency as a function of the observables $M_{\gamma\gamma}$, $p_{T\gamma\gamma}$ and $\Delta\phi_{\gamma\gamma}$. The overall efficiency is on average at the level of 60%.

5.6.4 Unfolding matrix

The unfolding matrix $U_{\alpha i}$ is used to convert the experimentally reconstructed distribution N_i^{II} to true distribution N_α^{II}, given the fact that event in X_α^{true} bin may migrate to X_i^{rec} bin due to experimental effects (e.g the energy scale and resolution). It is estimated from the $\gamma\gamma + (\gamma j + j\gamma)$ sample as listed in Table 5.2. In addition, the events that pass the reconstruction-level kinematic cuts but fail the truth-level kinematic cuts are not considered as signal events. Those events, called "below-threshold" events, need to be removed from the signal yields. The fraction I_i^{BT} of those events is also evaluated from Monte Carlo for each bin of the observable X_i^{rec}. With the fraction I_i^{BT}, the "below-threshold" events are removed from the signal yields, the unfolding matrix is then applied to obtain the true yield in the following way:

$$N_\alpha^{\mathbf{II}} = \sum_i U_{\alpha i} N_i^{\mathbf{II}} (1 - I_i^{BT}) \tag{5.35}$$

In practice, the iterative Bayesian unfolding algorithm [93] is used to perform the unfolding process. This algorithm is implemented in the RooUnfold package.

5.6.5 Reconstruction efficiency

The last ingredient needed to compute the cross-section in Equation 5.22 is the reconstruction efficiency ϵ_α^{rec} as defined in Equation 5.21. It is actually the fraction of true signal events satisfying the requirements listed at the beginning of Section 5.6, i.e that pass the event preselection as mentioned in Section 5.3.3 together with the calorimetric isolation.

This efficiency is derived using Monte Carlo samples, such as the $\gamma\gamma + (\gamma j + j\gamma)$ sample in Table 5.2. Before evaluating the efficiency, two issues need to be addressed. The first one

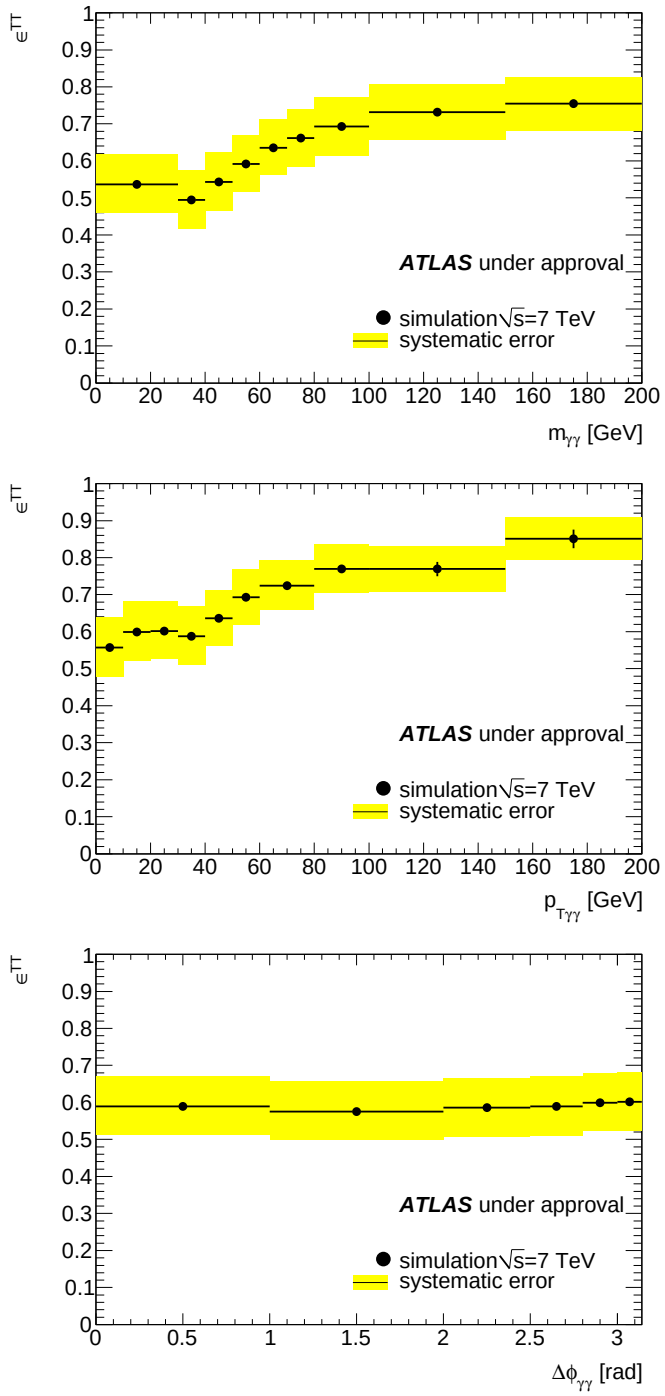


Figure 5.21: Event identification efficiency as a function of the observables $M_{\gamma\gamma}$, $p_{T\gamma\gamma}$ and $\Delta\phi_{\gamma\gamma}$. The yellow bands show the systematic error, whereas the statistical error is displayed by the verticla black lines, and is almost negligible.

is that the di-photon pair is chosen after applying the **Loose** photon selection, which introduces a **Loose** identification efficiency. The potential bias from this selection being applied before or after the calorimetric isolation and choice of the photon pair needs to be checked. The yields following the selection sequence of **Loose** selection, isolation requirement and the choice of photon pair, are compared with the yields following the selection sequence of isolation requirement, choice of photon pair and **Loose** selection. Results are found to be in good agreement in all the bins of the observables, with relative deviations below 1%. The second issue is the difference of the calorimetric isolation distributions between data and simulations, which implies that a correction to the isolation profile in Monte Carlo sample is necessary. A simple linear transformation is introduced:

$$E_T^{\mathrm{iso}^{\mathrm{data}}} = \alpha + \beta \cdot E_T^{\mathrm{iso}^{\mathrm{MC}}} \tag{5.36}$$

The α and β parameters for the leading and the subleading photons are shown in Table 5.20 using the $\gamma\gamma + (\gamma j + j\gamma)$ sample. After applying this transformation, the Monte Carlo calorimetric isolation distribution shows good agreement with the data. Besides, the difference between data and Monte Carlo may also result from the physical event itself, such as the kinematics and the fragmentation. Thus, the particle-level isolation $E_T^{\mathrm{iso(part)}^{\mathrm{MC}}}$ should be also corrected. A linear transformation is also introduced to model the relationship of particle-level isolation $E_T^{\mathrm{iso(part)}}$ and experimentally calorimetric isolation E_T^{iso}. It is illustrated as:

$$E_T^{\mathrm{iso}} = m \cdot E_T^{\mathrm{iso(part)}} + q \tag{5.37}$$

The m and q parameters from $\gamma\gamma + (\gamma j + j\gamma)$ sample are summarized in Table 5.21. They are assumed to work for data as well. Consequently, the corrected particle-level isolation becomes:

$$E_T^{\mathrm{iso(part)}^{\mathrm{data}}} = \frac{\alpha + (\beta - 1)q}{m} + \beta \cdot E_T^{\mathrm{iso(part)}^{\mathrm{MC}}} \tag{5.38}$$

	α [MeV]	β
leading γ	445	1.24
sub-leading γ	428	1.24

Table 5.20: Correction factors to the calorimetric isolation in Monte Carlo, according to the equation $E_T^{\mathrm{iso}^{\mathrm{data}}} = \alpha + \beta \cdot E_T^{\mathrm{iso}^{\mathrm{MC}}}$.

	m	q [MeV]
$\gamma\gamma + (\gamma j + j\gamma)$ sample	0.53	415

Table 5.21: Correction factors for describing the relationship between the particle-level isolation and the experimentally calorimetric isolation.

For a nominal reconstruction efficiency, only the transformation of Equation 5.36 is applied on the experimentally calorimetric isolation using $\gamma\gamma + (\gamma j + j\gamma)$ sample, which is generated with a nominal description of the detector. The particle-level isolation is not changed. The efficiency is on average around 55%. The main efficiency loss results from the

acceptance selection (-18%) and the calorimetric isolation requirement (-20%). The main systematic uncertainties relevant for this efficiency includes: the energy resolution, the energy scale, the particle-level isolation modeling, the knowledge of the detector geometry and the generator effect.

5.6.6 Final result

The differential cross-sections measured with the yields extracted with the 2D fit method, after electron background subtraction, efficiency corrections and unfolding as described above, are shown in Figure 5.22–**??**. They are compared with the predictions of RESBOS and DIPHOX for di-photon events satisfying the kinematic and isolation requirements given in Section 5.6. The theoretical uncertainties take into account the uncertainties on the scales (renormalization, factorization and fragmentation) which are fixed for each event to $M_{\gamma\gamma}$ for the nominal result, and on the parton density function.[14]

Observable $M_{\gamma\gamma}$ shows a good agreement between data and both DIPHOX and RESBOS, except in the very low $M_{\gamma\gamma}$ bin. Given the present level of uncertainty, a reasonable agreement is also found for the $p_{T\gamma\gamma}$ differential cross-section. However, the experimental measurement is above the theoretical predictions in the intermediate [30-80] GeV region. The trend is slightly stronger for RESBOS. The present uncertainty level does not allow to compare the data agreement with RESBOS and DIPHOX in the low $p_{T\gamma\gamma}$ region where the RESBOS prediction is expected to be more reliable. Finally, a discrepancy is observed in the full range for the $\Delta\phi_{\gamma\gamma}$ differential cross-section where the predicted shapes significantly differ from the experimental result. This was also observed at Tevatron both by D∅ and CDF.

In order to get a better understanding of the observed discrepancies, more precise measurements are needed. In 2011, one expects around 20 times more data than in 2010, thus allowing to greatly reduce the statistical uncertainties. Most of the systematic uncertainties have a statistical origin: for example, one can expect that the uncertainty from the definition of the non-**RobustTight** sample will be reduced with more data. Furthermore, with increased statistics, double differential cross-sections can be measured and will certainly help to understand the phenomenology of di-photon production.

[14] At the time of this thesis, the full estimation of the uncertainties is done for DIPHOX and the result is used for both the DIPHOX and RESBOS predictions.

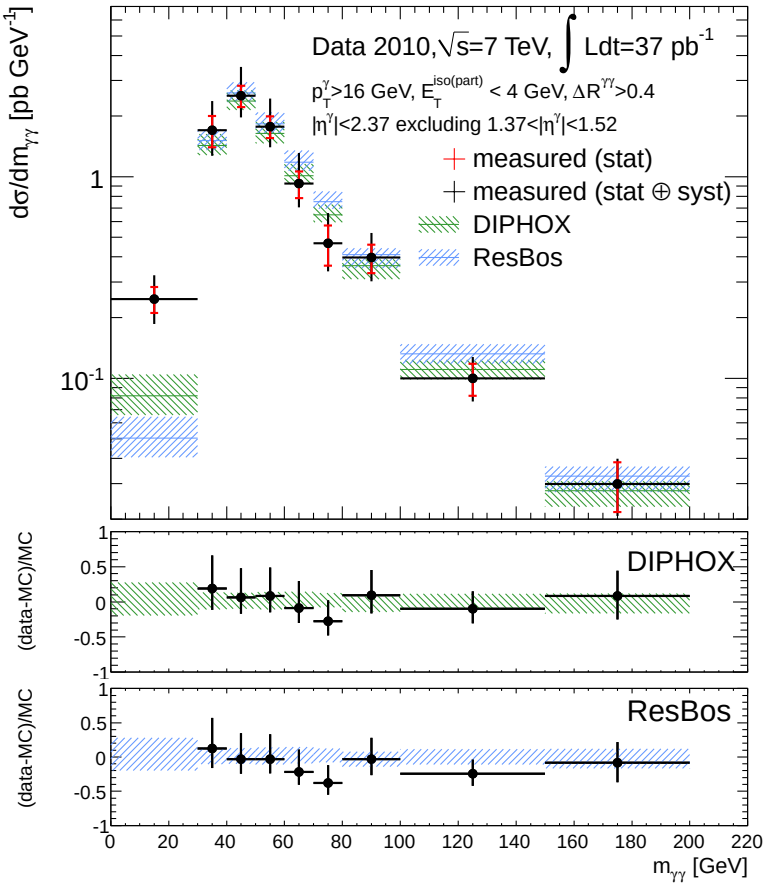


Figure 5.22: Di-photon production differential cross-section as a function of $M_{\gamma\gamma}$, compared with DIPHOX and RESBOS predictions.

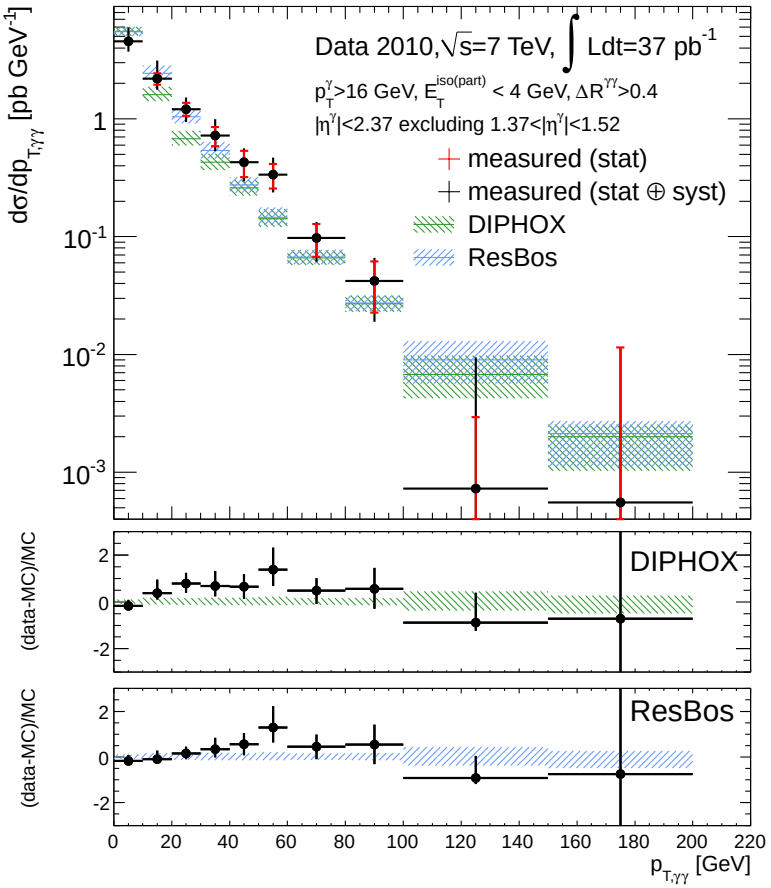


Figure 5.23: Di-photon production differential cross-section as a function of $p_{T\gamma\gamma}$, compared with DIPHOX and RESBOS predictions.

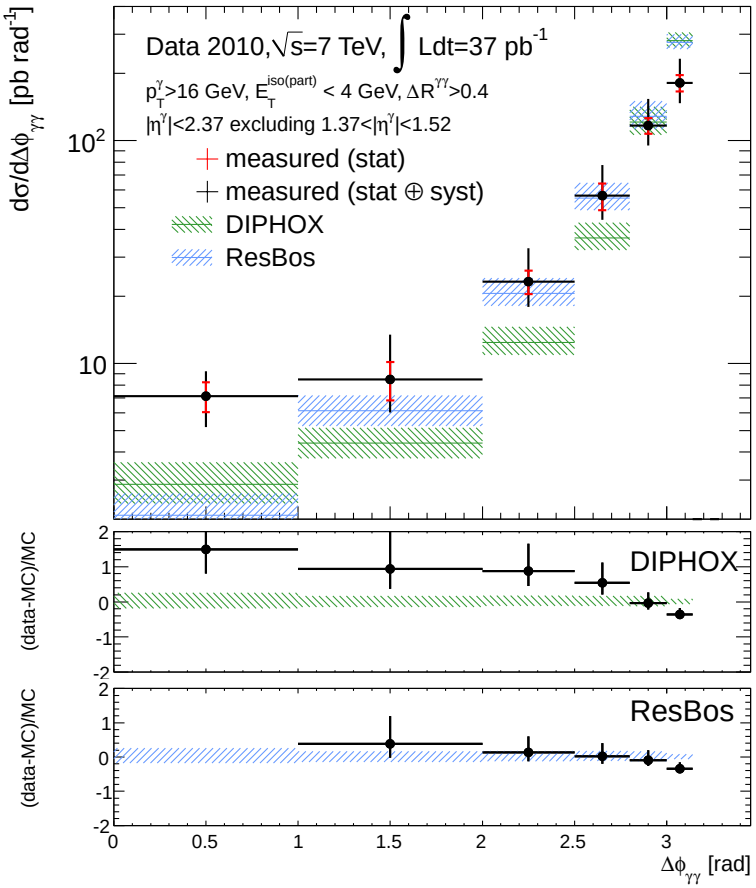


Figure 5.24: Di-photon production differential cross-section as a function of $\Delta\phi_{\gamma\gamma}$, compared with DIPHOX and RESBOS predictions.

Chapter 6

$H \rightarrow \gamma\gamma$ analysis

The search for the SM Higgs boson with the di-photon final state has been a benchmark study for the design of the ATLAS detector, and as such, was studied for a very long time. An estimation of the discovery potential at the center-of-mass energy of 14 TeV was reported in 1999 [94]. A statistical significance of 3.9 (6.5) σ was expected with an integrated luminosity of 30 (100) fb^{-1} for a Higgs mass of 120 GeV . Those performances were established with an inclusive analysis and based on signal and background processes calculated at LO.

In 2007, the discovery potential was re-assessed, also at the center-of-mass energy of 14 TeV [95]. The signal and background processes were calculated at NLO. A refined study was performed splitting the sample into different event categories according to the number of jets associated, the pseudorapidity of the photon candidates and their converted or unconverted nature. A multivariate maximum likelihood method was applied using the invariant mass and the transverse momentum of the two photons, as well as photon decay angle in the rest frame of the Higgs candidate, as discriminants. The statistical significance for a Higgs boson signal with a mass of 120 GeV was evaluated to be 3.2$\sim$4.1 σ with a luminosity of 10 fb^{-1} .

With the decision to run the LHC at a reduced center-of-mass energy, the sensitivity had to be evaluated again. This chapter presents the inclusive $H \rightarrow \gamma\gamma$ analysis on Monte Carlo simulation at center-of-mass energy of 7 TeV as well as the study of the 2010 data, corresponding to an integrated luminosity of 37 pb^{-1}.

6.1 Signal and background processes

6.1.1 Higgs production

In the SM, the Higgs boson couples preferentially to heavy particles, including the vector bosons W and Z, the top quark and the bottom quark in a lesser extent. Thus, there are mainly four production modes for the SM Higgs boson in hadron collisions. One can find in Figure 6.1 the corresponding Feynman diagrams for these four modes.

- The dominant process is the gluon-gluon fusion, $gg \rightarrow H$, where the Higgs boson couples indirectly to gluons via a triangular loop of quarks, mainly the top.

- The subleading production process is the vector boson fusion, $qq \rightarrow V^{\star}V^{\star} \rightarrow qq + H$. In this process, the quarks interact via a W/Z boson, and there is no color exchange

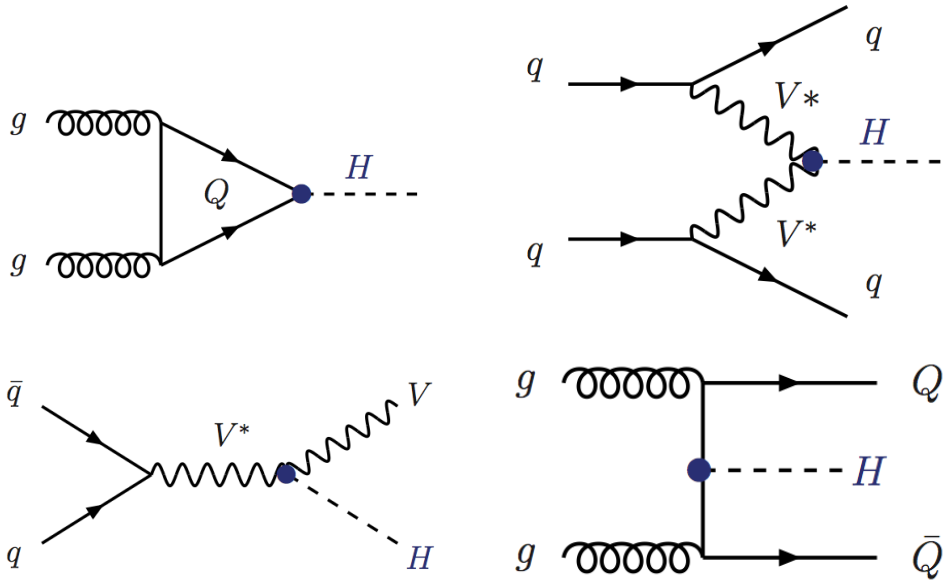


Figure 6.1: The main SM Higgs boson production modes in hadron collisions.

between the initial-state quarks. The two final quarks fragment into two forward jets, and no QCD activity is expected between the two forward jets except the Higgs decay products. This particular topology can be used to suppress the background in an exclusive analysis.

- The associated production with W/Z mode, $q\bar{q} \to V + H$, is called Higgs-strahlung.[1] Once the Higgs is discovered, this mode is interesting to study the Higgs coupling to the vector bosons.

- The associated production with heavy quarks ($gg, q\bar{q} \to Q\overline{Q} + H$), mostly $t\bar{t}H$, is an important process to measure the Yukawa coupling between the Higgs and the top quark.

Figure 6.2 shows the cross-section of Higgs production at the LHC for the different modes as a function of the Higgs mass, at a centre-of-mass energy of 7 TeV. Due to the large gluon luminosity at the LHC, the gluon-gluon fusion mode dominates over the whole range of the Higgs mass. The bump at $M_H \sim 350$ GeV near the $t\bar{t}$ threshold comes from the enhancement of the imaginary part of the amplitude due to the top loop. As labeled in the plot, the cross-section includes the QCD and EW corrections, and all the modes are computed at NNLO except the $t\bar{t}H$ which is computed only at NLO.

6.1.2 Higgs decay

Figure 6.3 shows the branching ratios for the main Higgs decay modes as a function of the Higgs mass.

In the low mass region (normally meant for the range [100, 140] GeV), the dominant decay channel is $H \to b\bar{b}$, with a branching ratio around 70% at the Higgs mass of 120

[1] By analogy to bremsstrahlung.

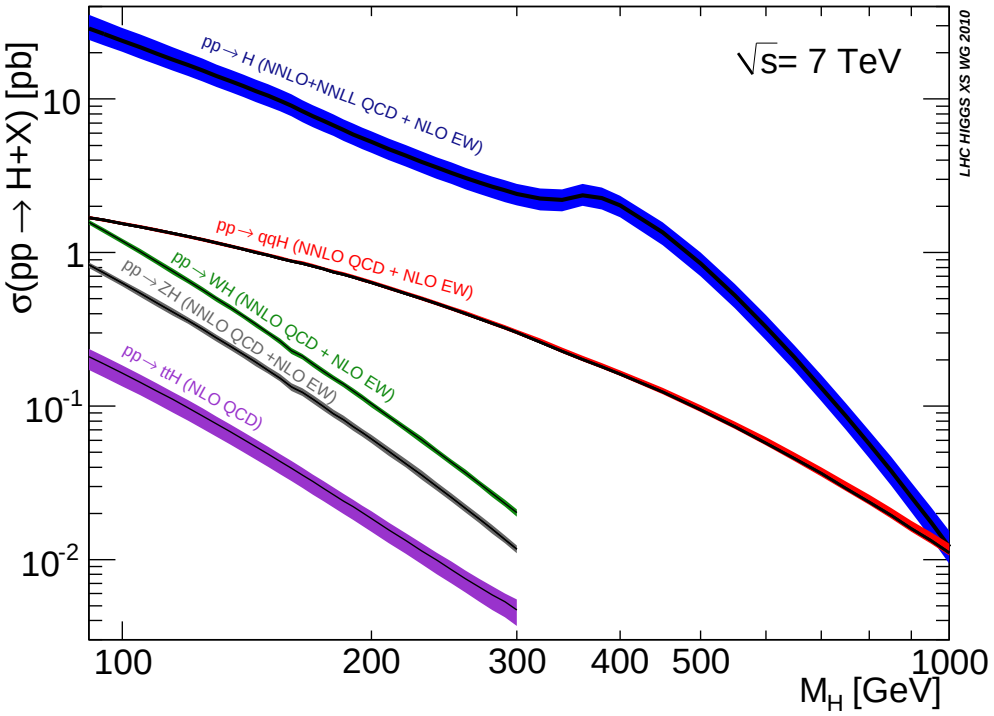


Figure 6.2: Cross-section of the Higgs production modes at LHC (7 TeV) as a function of its mass from Reference [96].

GeV. However, the sensitivity of Higgs search with this channel is small. The main reason is that the QCD background is large. In addition, it is difficult to perform an accurate reconstruction of the b-jets, and therefore, to reconstruct the invariant mass peak with the resolution needed to discriminate the signal against the QCD background. Several algorithms are developed in ATLAS to tag the b-jets, making use of their specific topology such as long distance vertex and high multiplicity.

The $H \to gg$, $c\bar{c}$ channels are also predicted to have a high branching ratio at low mass. However, it is impossible to discriminate the gluon or charm jets against the QCD backgrounds. Therefore, these modes are hopeless.

The $H \to \tau\tau$ channel has a moderate branching ratio, about 8% at the Higgs mass of 120 GeV, and does not suffer from huge QCD background. However, there is also a large background contamination from $Z \to \tau\tau$ process, which is also difficult to reject.

Compared with the channels mentioned above, the branching ratio of the $H \to \gamma\gamma$ channel is small, at the level of 2×10^{-3}, firstly due to the indirect Higgs coupling to the photon, mainly through W boson and charged fermions. The partial decay width for $H \to \gamma\gamma$ has the following form [98]:

$$\Gamma(H \to \gamma\gamma) = \frac{G_\mu \alpha^2 M_H^3}{128\sqrt{2}\pi^3} |\sum_f N_c Q_f^2 A_{1/2}^H(\tau_f) + A_1^H(\tau_W)|^2 \tag{6.1}$$

where $A_{1/2}^H(\tau_f)$ is the form factor associated to the fermion loops and $A_1^H(\tau_W)$ is the one associated to the W boson loops. $\tau_i = M_H^2/4M_i^2$ with $i = f, W$ are defined by the corresponding masses of heavy particles in the loop. The $Hf\overline{f}$ coupling is proportional to the

mass of the fermion m_f, thus the contribution of light fermions is negligible. Therefore, the Higgs boson decay into two photons is mainly through top quark and W boson loops (Figure 6.4), which interfere destructively, leading to a small branching ratio. However, the two-photon final state still offers several advantages for Higgs discovery: it is a simple signature, the invariant mass of the two photons can be reconstructed with high accuracy, and additional observables like the Higgs candidates transverse momentum or $\cos\theta^\star$ (See Section 6.2) can be fully reconstructed and used to discriminate the signal from the background. This channel also shows a lot of interest after the Higgs boson discovery. Since it gives access to all production modes, the relative production rates can be tested and compared with the SM prediction. This channel can also offer an accurate measurement of the Higgs mass, and the study of $\cos\theta^\star$ will allow for the determination of the spin of the Higgs boson.

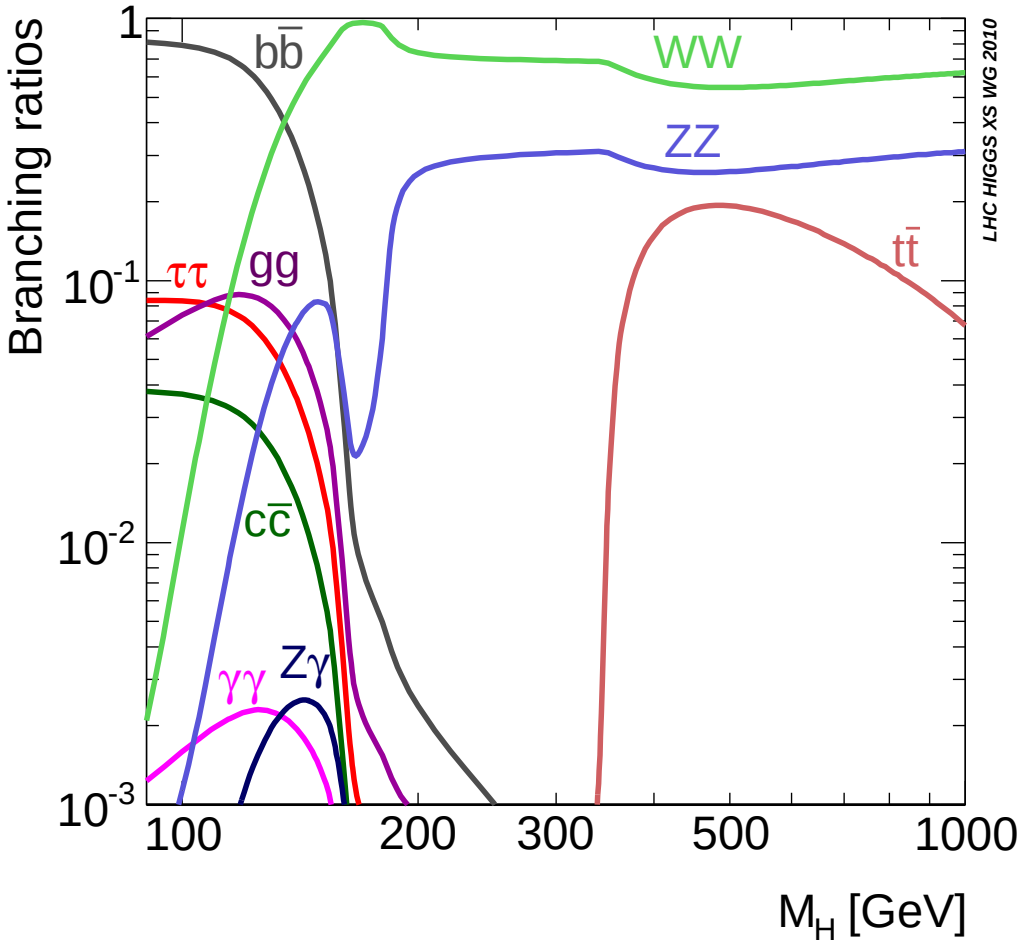


Figure 6.3: Branching ratios for the Higgs decay modes as a function of the Higgs mass from Reference [96].

6.1.3 Background processes

- **Irreducible background**

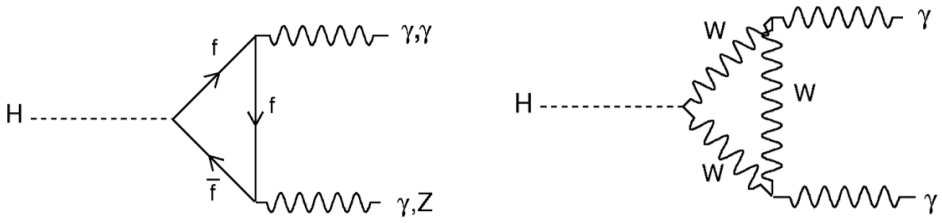


Figure 6.4: Feynman diagrams for $H \to \gamma\gamma$ or $Z\gamma$.

The QCD di-photon production discussed in Section 5.1.1 is regarded here as an irreducible background, since it has the same signature as $H \to \gamma\gamma$. The two-photon invariant mass is the main discriminating variable between the two processes. Photon isolation can mostly help to reduce the contribution with one or two photons from fragmentation.

- **Reducible background**

  The jet background for the QCD di-photon production is also a background for $H \to \gamma\gamma$. As discussed in Section 5.1.2, it is due to photon-jet and di-jet events and comes from the fact that the jet(s) are mis-identified as photon(s). The cross-section for the photon-jet and di-jet events is significantly large compared with the $H \to \gamma\gamma$ signal process and also much larger than the irreducible background. However, with photon identification and isolation, these backgrounds can be greatly reduced. Finally, as for the di-photon study of Chapter 5, the processes to ee final states (Drell-Yan, $Z \to ee$, $WW \to e\nu e\nu$) and to $e\gamma$ final states ($\gamma W \to \gamma e\nu$, $\gamma Z \to \gamma ee$) can contribute to the background because electrons are mis-identified as photons.

6.2 Discriminating variables

In this thesis, only the inclusive analysis for $H \to \gamma\gamma$ is presented. The discriminating variables are based on the di-photon system, and take advantage of the Higgs resonance and its related kinematic behavior. The variables that can be considered are the following:

- $M_{\gamma\gamma}$: the photon pair invariant mass. Photon pairs that come from the Higgs decay should give a mass peak above the continuum exponential QCD background. The distribution is almost independent on the prediction models. Thus this variable is supposed to be the most robust discriminant.

- $p_{T\gamma\gamma}$: the transverse momentum of the photon pair. The distribution for this variable is largely affected by the parton density function and also the model to simulate the parton radiation.

- $\cos\theta^\star$: the cosine of the angle between one photon of the pair and a predefined axis in the photon pair rest frame. Two definitions are commonly used in ATLAS. The first one is the boost axis (BA) of the di-photon system. The second one is defined in the so-called Collins-Soper (CS) frame, which is based on the initial-state hadron momenta ($\overrightarrow{P_1}$ and $\overrightarrow{P_2}$) as the axis which bisects the momenta ($\overrightarrow{P_1}$ and -$\overrightarrow{P_2}$), as shown

in Figure 6.5. This variable has only little discriminating power, but the signal distribution relies only on the hypothesis that Higgs is a spin-0 particle.

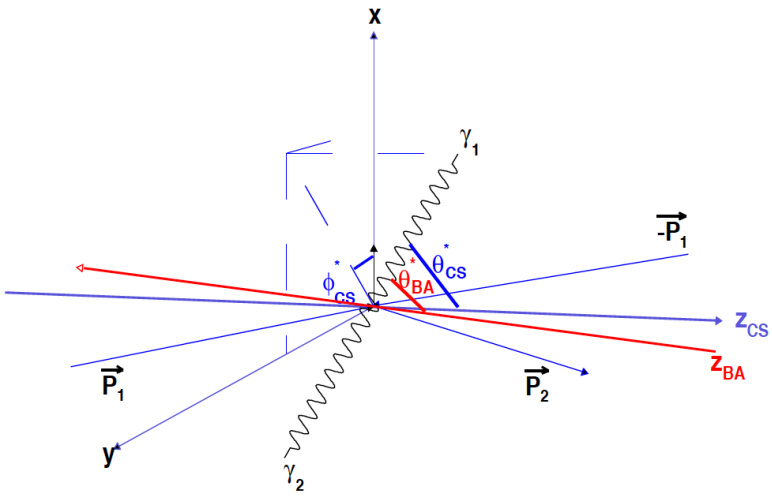


Figure 6.5: Two definitions for the angle between the two photons. z_{CS} bisects the protons direction, while z_{BS} bisects the direction of the two-photon system. $P_{\gamma 1 \gamma 2}$ and P_{p1p2} correspond respectively to the plan of the two photons in their rest frame and of the two protons, which are the CS and boosted axis (BA) frame. $\theta^{\star}_{BA}$ corresponds to BA the polar angle in the BA frame. $\phi^{\star}_{CS}$ corresponds to the azimuthal angle in the CS frame.

6.3 Principles of an early $H \rightarrow \gamma\gamma$ analysis

In order to perform a robust $H \rightarrow \gamma\gamma$ inclusive analysis, only the di-photon invariant mass is used as the discriminating variable. In 2009−2010, such a robust inclusive analysis was performed using MC simulated events at center-of-mass energy of 10 TeV before the decision to run at 7 TeV. Results were then extrapolated to 7 TeV in order to assess the ATLAS sensitivity with 1 fb^{-1}, which was the expected integrated luminosity at the end of 2011. Almost the same analysis was repeated on 2010 data. In this section, the principles of the analysis are presented.

6.3.1 Photon selection

The Higgs search on 2010 data use the same acceptance and identification criteria as in the QCD di-photon analysis described in Section 5.3.3, i.e. $|\eta|$ and **RobustTight** selection. Before the data taking, the standard photon identification was the so-called **PhotonTight** selection as described in Table 3.2. For the analysis based on 10 TeV Monte Carlo samples, the second set of cuts in Table 3.2 was used.

In addition to the photon identification, an isolation requirement is applied to the photon candidates in order to further reduce the photon-jet and di-jet background. In the 10 TeV MC analysis, a track-based isolation is used. As defined in Section 4.4, the transverse momenta of tracks with $p_T > 1$ GeV inside a $0.1 < \Delta R < 0.3$ ring around the photon direction are summed up. The criteria $\sum_{Trks} p_T < 4$ GeV is then applied, which keeps the

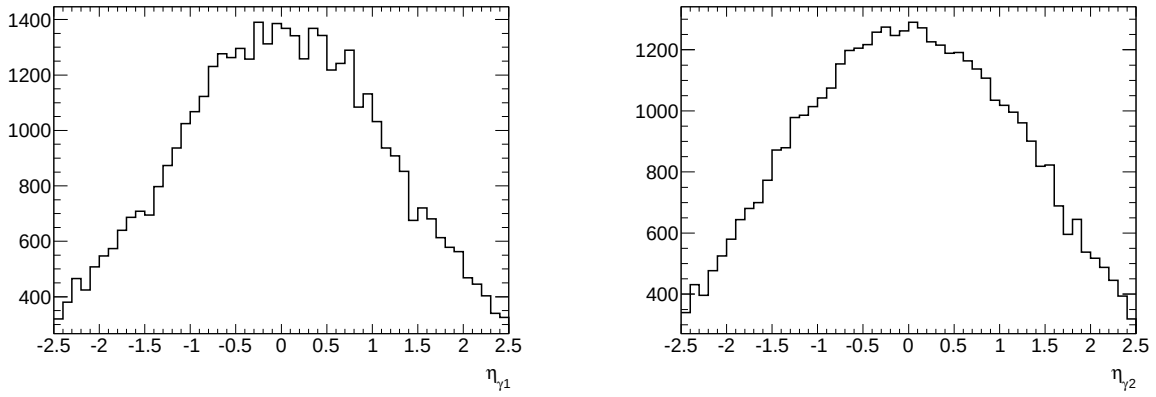
true photon efficiency at 99.2% and rejects fake photons by a factor of 1.6. In the real data analysis, a calorimetric isolation as defined in Section 5.3.1 is used instead. The criteria of $E_T^{iso} < 3$ GeV is applied. The main advantage to use this calorimetric isolation in the real data analysis is the smaller discrepancy observed between data and MC, which leads to reduced systematics.

6.3.2 Event selection

At the event level, three requirements are imposed:

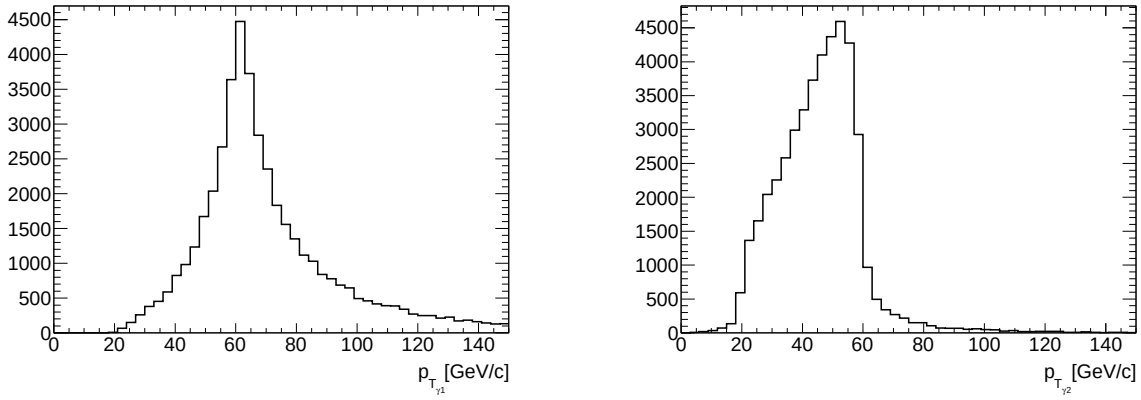
- There must be at least two photons that pass the photon identification and isolation requirement within the acceptance.
- Kinematic cut: the leading p_T photon is required to be $p_T^{\gamma 1} > 40$ GeV and the subleading p_T photon is required to be $p_T^{\gamma 2} > 25$ GeV.
- Trigger selection: the trigger item 2g20_loose is used in the 10 TeV MC analysis. This trigger item chains for each level: L1_2EM18, L2_2g20_loose and EF_2g20_loose. It selects events with at least two photons having transverse energy above 20 GeV at the EF level. In the real data analysis, two trigger items are applied depending on the data periods. For data periods A to E3, which corresponds to a luminosity of 0.8 pb^{-1}, the lowest unprescaled L1 calorimeter trigger L1_EM14 is used. For the latest data periods (from E4 onwards) corresponding to 35.8pb^{-1}, 2g15_loose is used.

  The η and transverse momentum (p_T) distribution for the reconstructed leading p_T and subleading p_T photon that come from Higgs decay are shown in Figure 6.6 and 6.7. The invariant mass and transverse momentum of the photon pair (which are the reconstructed mass and transverse momentum of Higgs) and $\Delta\phi$ between the two photons are also computed, and shown in Figure 6.8. The distributions are obtained with a simulation sample generated with PYTHIA, in which Higgs mass is at 120 GeV. A SymPhotonFilter (requiring at least two photons with $p_T > 20$ GeV and in $|\eta| < 2.7$) is applied when generating the sample.



(a) (b)

Figure 6.6: The η distribution for the leading (a) and subleading p_T photon (b) that come from Higgs decay. The PYTHIA simulated sample for Higgs mass at 120 GeV is used.



(a) (b)

Figure 6.7: The transverse momentum (p_T) distribution for the leading (a) and subleading p_T photon (b) that come from Higgs decay. The PYTHIA simulated sample for Higgs mass at 120 GeV is used.

In addition, for the real data analysis, there are three more requirements:

- Events must satisfy the egamma Good Run List, which is defined in Section 5.3.3, requiring the detector operated with a good data quality.
- A primary vertex consistent with the average beam spot position and with at least three associated tracks is required.
- As mentioned in Section 5.3.3, Object Quality maps are used to make both photons be outside the calorimeter regions affected by readout (dead OTX) and high voltage problems. The criterion of *jet cleaning* is also applied to reject spurious signal from ECAL coherent noise burst.

6.3.3 Primary vertex reconstruction

Pairs of photons passing the whole selection described above, one with transverse momentum $p_T > 40$ GeV and the other with $p_T > 25$ GeV, are considered as Higgs boson candidates. If more than two photons in the event pass the selection, only the pair formed by the two with highest p_T is considered. The invariant mass $M_{\gamma\gamma}$ of the two photon candidates is reconstructed using their kinematic informations, i.e. the polar angles θ_1, θ_2 expressed as pseudorapidities η_1, η_2, the azimuthal angles ϕ_1, ϕ_2 and the transverse energies E_T^1, E_T^2:

$$M_{\gamma\gamma} = \sqrt{2E_T^1 E_T^2[\cosh(\eta_1 - \eta_2) - \cos(\phi_1 - \phi_2)]} \tag{6.2}$$

The directions η_1 and η_2 of the photons depend on the longitudinal position z of their origin. In order to get the best possible performances, different methods are used to compute z [100], depending on the nature of the photon. For unconverted photons and for converted photons whose tracks are not reconstructed with silicon hits in the inner detector, the best method, so-called calorimeter pointing method, is to draw a straight line (in the $z-R$ plane) passing through the barycenters of the first and the second samplings of the electromagnetic calorimeter. The line gives the direction of the photon, and the intersection of the line with

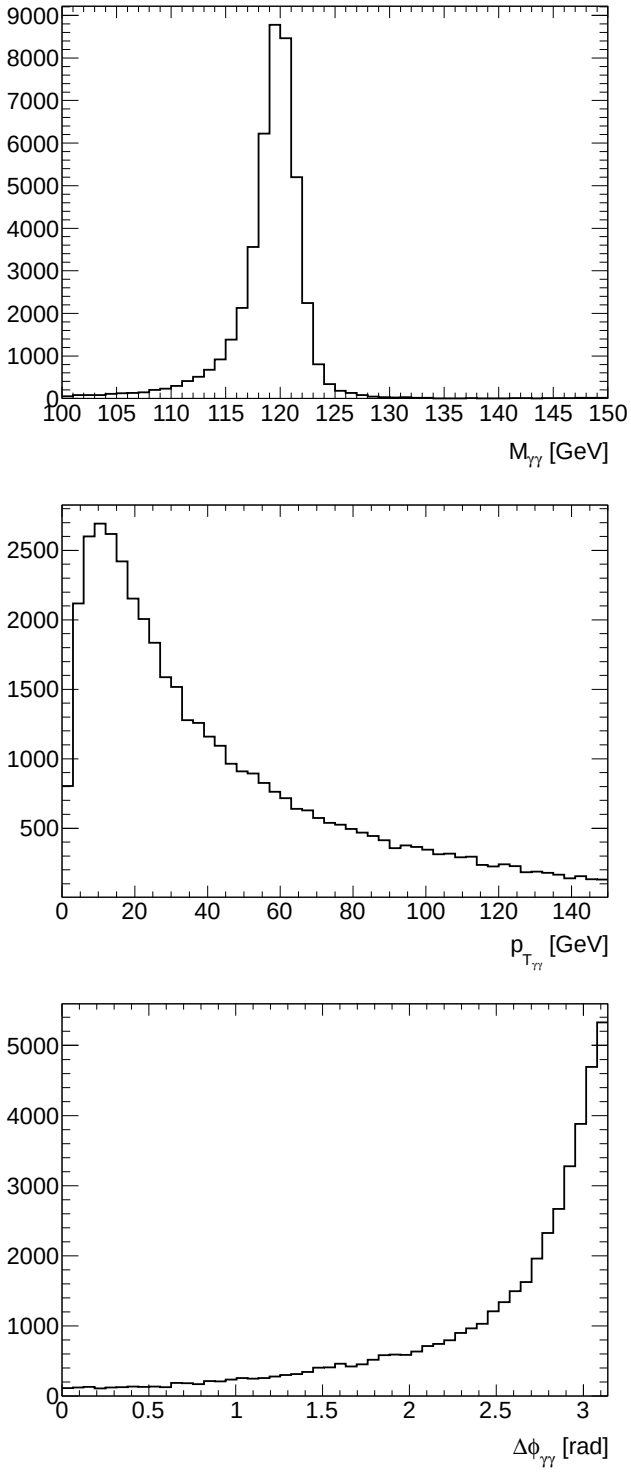


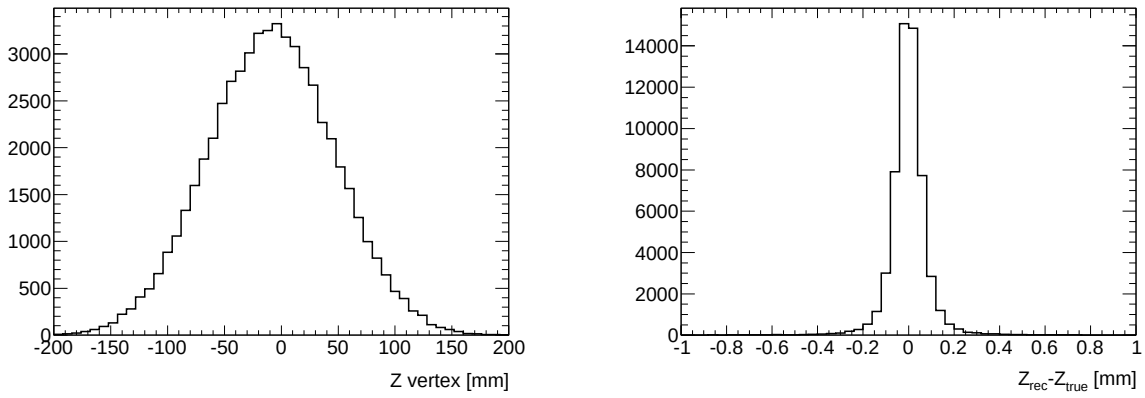
Figure 6.8: The distribution of the invariant mass (top), transverse momentum (middle) of the photon pair and $\Delta\phi$ (bottom) between the two photons where both photons come from the Higgs decay. No requirement on the photon p_T and η is applied yet. The distribution is obtained using PYTHIA simulated sample for Higgs mass at 120 GeV.

the beam axis yields the vertex of the photon. For the converted photons for which tracks have at least one silicon hit in the inner detector, the best method is to use the straight line passing through the first sampling of the electromagnetic calorimeter and the conversion vertex. The z determined with any of these procedures is called $z_{calo+ID}$ in the following. The computation of $z_{calo+ID}$ and its error provides the likelihood $\mathcal{L}_{calo+ID}(z)$:

$$\mathcal{L}_{calo+ID}(z) = e^{-\frac{z^2}{2\sigma_{IR}^2}} \times e^{-\frac{(z-z_{calo+ID})^2}{2\sigma_{z_{calo+ID}}^2}} \times \frac{1}{e^{-\frac{z^2}{2\sigma_{IR}^2}}} \ . \tag{6.3}$$

where σ_{IR} is the interaction region size. In this product the first term is the probability of having a hard process vertex z considering its mean value $z = 0$ and its uncertainty σ_{IR}; the second factor is the probability of $z_{calo+ID}$ with its uncertainty $\sigma_{z_{calo+ID}}$; the third factor is the inverse of the probability of having a Minimum Bias vertex z considering its mean value $z = 0$ and its uncertainty σ_{IR}.

Figure 6.9 (a) shows the z vertex distribution from the reconstruction. Figure 6.9 (b) shows the discrepancy between reconstructed and true position. The later distribution exhibits a gaussian core with a width around 0.15 mm.



(a) (b)

Figure 6.9: The distribution of the reconstructed primary vertex position (a). The discrepancy between reconstructed and true position (b). The PYTHIA simulated sample for Higgs mass at 120 GeV is used.

6.3.4 Extraction of the exclusion limit

The profile likelihood method is used to set the exclusion limit. Only the di-photon invariant mass is used as the discriminant, thus the likelihood model is of the form:

$$L = \mu N_S P_S(M_{\gamma\gamma}) + N_B P_B(M_{\gamma\gamma}, \xi) \tag{6.4}$$

where μ is the signal strength parameter defining the hypothesis, μ=0 corresponding to absence of a signal, μ=1 the signal rate expected from SM; N_S is the total number of signal events expected from SM; N_B is the overall number of background events, and ξ is the exponential slope of background shape: they are the only two nuisance parameters;

P_S, P_B are PDFs respectively for the signal and the background di-photon invariant mass distribution. The distribution for signal is modeled by the sum of a Crystal Ball function which describes the core of the distribution and a Gaussian function which describes the outlying events. Figure 6.10 shows an example with the fit for the signal shape. For background it is modeled by an exponential function of slope ξ.

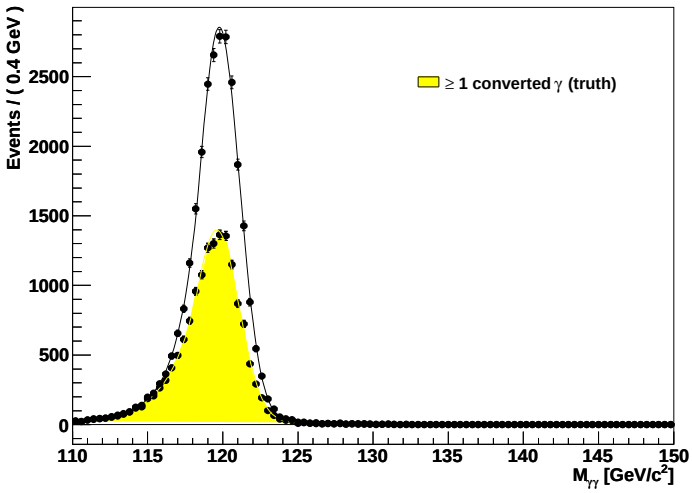


Figure 6.10: The di-photon invariant mass distribution after the selection using PYTHIA simulated sample for Higgs mass at 120 GeV. The distribution for events with at least one converted photon is overlaid (yellow histogram). The distribution is fitted with a sum of Crystal Ball function and a Gaussian function.

To test a hypothesized value of μ, the profile likelihood is constructed as:

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\theta}})}{L(\hat{\mu}, \hat{\theta})} \tag{6.5}$$

Here $\hat{\hat{\theta}}$ is the value of θ maximizing L for a fixed μ. It is the conditional maximum-likelihood estimator (MLE) of θ. $\hat{\theta}$ and $\hat{\mu}$ are the values that maximize L when both θ and μ are left floating, thus $\hat{\theta}$ and $\hat{\mu}$ are the MLEs in this case. From this definition, one can see that $0 \leq \lambda \leq 1$, with $\lambda(\mu) = 1$ implying good agreement between the data and the hypothesized value of μ. To evaluate the significance of the hypothesis, the test statistics is defined as:

$$q_\mu = \begin{cases} -2\ln\lambda(\mu), & \hat{\mu} \leq \mu \\ 0, & \hat{\mu} > \mu \end{cases} \tag{6.6}$$

Two distributions of q_μ are generated, one in the case of background only and one for background plus signal. These distributions are obtained by generating a large number of pseudo-experiments, or toy Monte Carlo simulation. In order to test a certain μ hypothesis, μ_x, each pseudo-experiment is generated with $\mu_x N_S$ number of signal events, and N_B of background events. The combined sample is first fitted with both μ and θ to be floated to obtain the denominator of Equation 6.5. Then it is fitted with a fixed value of μ_x to get the numerator of Equation 6.5. q_μ value for μ_x hypothesis for the background plus signal is then derived. The same procedure is repeated with the sample of only background events to get the q_μ value for the case of background only. With many pseudo-experiments, the

two distributions of q_μ for background and background plus signal can be derived. The latter distribution follows a $1/2\delta$ function at 0 plus a $1/2\chi^2$ function. The confidence level at which a certain hypothesis is excluded (CL_{s+b}) is given by $CL_{s+b} = 1 - p_\mu$, where the p-value p_μ is calculated by

$$p_\mu = \int_{q_{\mu,obs}}^{\infty} f(q_\mu|\mu) dq_\mu$$

In a sensitivity study, $q_{\mu,obs}$ is the median value from the background only q_μ distribution, and $f(q_\mu|\mu)$ is the background plus signal distribution function. The $\pm 1\sigma$ or $\pm 2\sigma$ uncertainty on the excluded value is derived by varying the integral range of the background plus signal distribution. This procedure is repeated for different signal hypotheses in order to find the one corresponding to an exclusion at 95% confidence level (C.L.).

For the extraction of the exclusion limit from a real data sample, $q_{\mu,obs}$ is the value of the test statistics that is actually measured on this sample.

There is another definition for the exclusion limit, which is calculated from:

$$1 - CL_s = 1 - \frac{1 - CL_{s+b}}{1 - CL_b} \tag{6.7}$$

here the p-value calculated on the background only distribution 1-CL_b is used as normalization factor for the p-value 1-CL_{s+b} calculated from the background plus signal distribution. This definition will not exclude the background only hypothesis at 2σ, and gives more conservative results. It is frequently used in Tevatron experiments.

6.4 Assessment of the exclusion limit from Monte Carlo simulation

In the 10 TeV MC analysis, the signal and background yields after all the selection defined above are evaluated for a luminosity of 200 pb^{-1}. The yields are then extrapolated to 1 fb^{-1} at $\sqrt{s} = 7$ TeV. A summary of the analysis is given here, with a focus on a study which incorporates systematic uncertainties into the exclusion limit.

6.4.1 Event-level trigger efficiency measurement

The photon trigger efficiency discussed in Chapter 4 is a marginal trigger efficiency, which is with respect to photon identification and isolation selection. The event-level trigger efficiency discussed in this section is also a marginal trigger efficiency: it is evaluated for events passing selection criteria (1) and (2) defined in Section 6.3.2. In order to estimate the efficiency both for signal and background, $H \to \gamma\gamma$, QCD $\gamma\gamma$ and photon-jet samples are used. Because of the low statistics obtained after the selection, the efficiency can not be evaluated for the di-jet background. The efficiencies are simply estimated as the ratio of Higgs candidate events (true or not) passing the trigger requirements: none of the data-driven methods developed in Chapter 4 is used. The results are shown in Table 6.1. This trigger item is efficient in selecting the Higgs candidate events, both for events coming from the signal $H \to \gamma\gamma$ sample or from the QCD background sample ($\gamma\gamma$ and photon-jet). It is expected, since the selection criteria (1) and (2) are tighter than the trigger level selection on the shower shape, and since the photon candidates passing the kinematic requirements ($p_T > 25$ or 40 GeV) are already in the plateau region of the trigger.

Sample	Efficiency (%)		
	L1_2EM18	L2_2g20_loose	EF_2g20_loose
$H \to \gamma\gamma$	99.98 ± 0.01	99.97 ± 0.01	99.86 ± 0.01
$\gamma\gamma$	99.96 ± 0.03	99.96 ± 0.03	99.83 ± 0.06
$\gamma j + j\gamma$	99.84 ± 0.07	99.68 ± 0.09	99.19 ± 0.15

Table 6.1: Efficiency of the 2g20_loose trigger item with respect to the Higgs selection, measured on $H \to \gamma\gamma$, $\gamma\gamma$, $\gamma j + j\gamma$ MC events.

6.4.2 Signal and background estimation

The analysis is focused on the invariant mass window of $100 < M_{\gamma\gamma} < 150$ GeV, in which the tails of the signal distribution are well included.

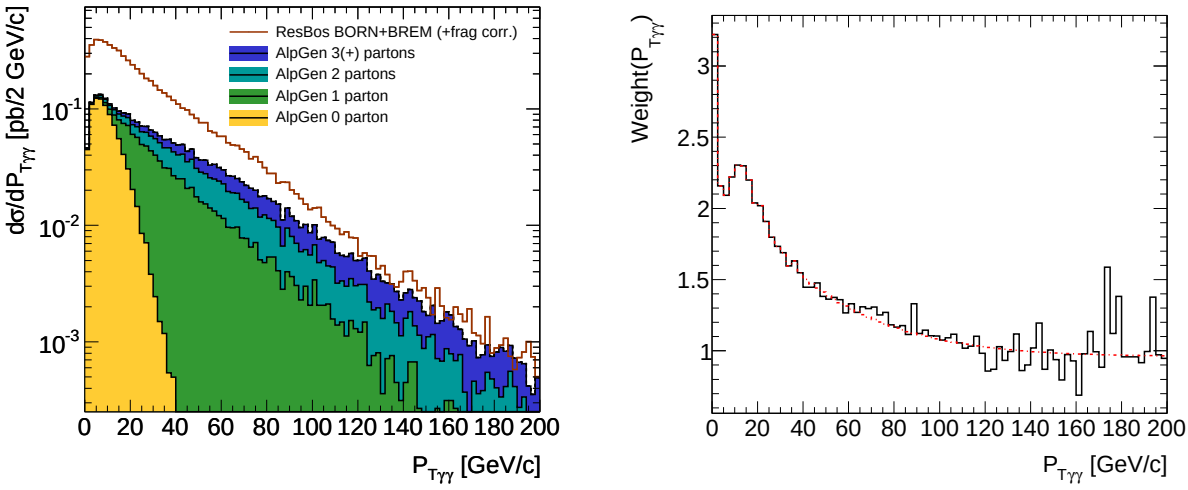
- **signal**: the samples of signal events, with different production modes, are fully simulated with the PYTHIA generator. The samples are normalized to an equivalent integrated luminosity of 200 pb^{-1}, according to the PYTHIA cross-sections. The number of events in the combined sample is then scaled by the ratio of the cross-section predicted in Reference [96], at NLO, over the cross-section predicted by the PYTHIA generator.

- **background**: for the irreducible background, both PYTHIA and ALPGEN samples are used. ALPGEN [99] is used for the generation of the Born and bremsstrahlung contributions, with the emission of up to 3 jets. The di-photon invariant mass distribution is obtained by an event-by-event reweighting of the ALPGEN fully simulated samples based on the transverse momentum of the di-photon pair, using RESBOS spectrum, as shown in Figure 6.11. RESBOS cross-section is firstly rescaled to account for the fragmentation treatment performed in DIPHOX. The Box process cross-section and distributions are obtained with a PYTHIA fully simulated sample, rescaled by the RESBOS NLO prediction. Finally, all of them are normalized to the expected integrated luminosity. For the reducible background, it is impossible to obtain the di-photon invariant mass distribution with decent statistics applying the selection on the fully simulated samples. Thus a parameterization method is used to predict the distribution. The jet rejection factor is defined as:

$$R = \frac{N_{jets}}{N_{fake\,\gamma}} \tag{6.8}$$

  where N_{jets} is the total number of jets produced in the acceptance region, and $N_{fake\,\gamma}$ is the number of jets that pass the photon identification and isolation criteria. This jet rejection factor is parameterized as a function of the transverse momentum of the quark-initiated jets and gluon-initiated jets respectively. Besides, the transverse momentum of the jets faking photons is also parameterized in the p_T bins of the initial jets. This parameterization is then applied on the generation level samples. A good agreement is observed in the di-photon invariant mass shapes from the photon-jet background between the one obtained with this parameterization procedure and the one from fully simulated sample, as shown in Figure 6.12. This validates the procedure which is then used to predict the $M_{\gamma\gamma}$ spectrum for jj events. In addition, the distribution from the photon-jet background is scaled by a k-factor, which is derived

from the ratio of the prediction from JETPHOX (NLO) and PYTHIA (LO). There is no higher order correction on the di-jet contribution, since it only contributes to the total background of about 3%.



(a) (b)

Figure 6.11: The transverse momentum ($p_{T_{\gamma\gamma}}$) of the photon pair for ALPGEN and RESBOS (a). The distribution for RESBOS has been corrected to account for the fragmentation treatment in DIPHOX. Disctribution of the weight used to reweight ALPGEN as a function of $p_{T_{\gamma\gamma}}$ (b).

6.4.3 Extrapolation to 7 TeV

The extrapolation of the results from 10 TeV to 7 TeV is performed as follows:

- The number of signal events is first evaluated using the 7 TeV PYTHIA Monte Carlo samples. Yields are then scaled by the ratio of the predicted NLO cross section [96] at $\sqrt{s} = 7$ TeV over the one predicted by PYTHIA at LO. The shape of di-photon invariant mass distribution is assumed to be the same. Table 6.2 shows the reconstruction efficiencies and expected numbers of signal events with 1 fb^{-1}. For processes with no available simulated samples, the expected numbers are estimated by using numbers of the VBF process[2].

- The numbers of background events for each components is scaled by the ratio of the PYTHIA predicted LO cross-section at 7 TeV over the one at 10 TeV.

This procedure is based on the assumption that the jet rejections are the same at the centre-of-mass energy of 7 TeV and 10 TeV. After this extrapolation, about 13 signal events

[2]The expected number of events, $N_A(mass)$, is calculated by

$$\frac{\sigma_{A,mass}}{\sigma_{A,120}} \times \frac{\sigma_{VBF,120}}{\sigma_{VBF,mass}} \frac{N_{VBF,mass}}{N_{VBF,120}} \times N_{A,120}. \quad (6.9)$$

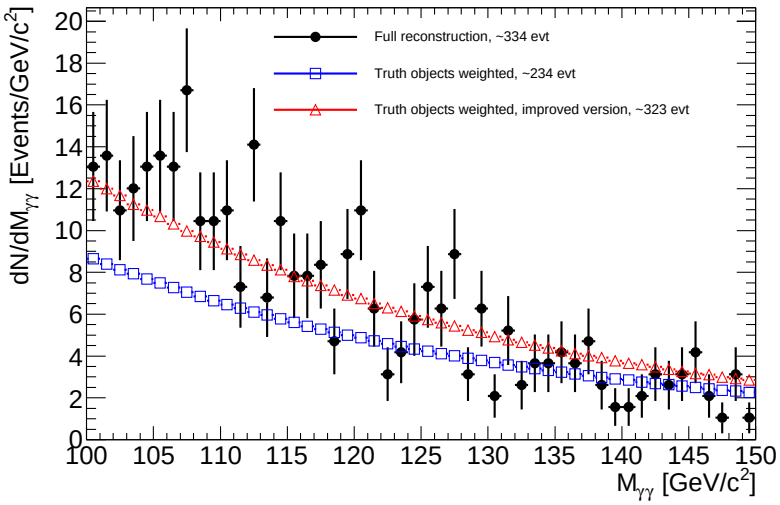


Figure 6.12: Di-photon invariant mass distribution obtained with the fully simulated sample by directly applying the selection (black dot), with the fast simulated sample by applying only the parameterization of the jet rejection factor (blue square) or together with the parameterization of the jet fake photon p_T (red triangles).

production	m_H = 110 GeV		m_H = 115 GeV		m_H = 120 GeV		m_H = 130 GeV		m_H = 140 GeV	
process	Eff [%]	N_{evts}	Eff [%]	N_{evts}	Eff [%]	N_{evts}	Eff [%]	N_{evts}	Eff [%]	N_{evts}
ggH	41	10.37	42	10.56	44	10.81	47	9.96	49	7.66
VBF	53	1.11	54	1.20	55	1.22	57	1.19	58	0.99
$W(\rightarrow jj)H$	-	0.44	-	0.43	50	0.40	-	0.33	-	0.23
$W(\rightarrow l\nu)H$	-	0.23	-	0.22	54	0.21	-	0.17	-	0.12
$Z(\rightarrow jj)H$	-	0.24	-	0.22	54	0.21	-	0.17	-	0.12
$Z(\rightarrow \nu\bar{\nu})H$	-	0.07	-	0.07	55	0.07	-	0.06	-	0.04
$Z(\rightarrow ll)H$	-	0.04	-	0.04	54	0.03	-	0.03	-	0.02
ttH	-	0.10	-	0.09	46	0.09	-	0.08	-	0.06
Total		12.60		12.85		13.05		12.00		9.25

Table 6.2: Reconstruction efficiencies with 7 TeV simulated samples and expected number of signal events for an integrated luminosity 1 fb^{-1}.

for a Higgs mass of 120 GeV and 8490 background events are expected in the invariant mass window of $100 < M_{\gamma\gamma} < 150$ GeV at 7 TeV for a luminosity of 1 fb^{-1}.

Table 6.3 summarizes the number of events for the signal (all channels combined for m_H = 120 GeV) and the main backgrounds for a luminosity of 1 fb^{-1} at $\sqrt{s} = 7$ TeV.

Figure 6.13 shows the confidence level obtained by the CL_{s+b} method described above, as a function of the hypothesis μ. The junction point between the red horizontal line and the black line is the central value for the exclusion limit at 95%, while the junction points with the yellow and green lines are respectively deviation of 1σ and 2σ. Exercise is done for each of the five mass points (110, 115, 120, 130, 140 GeV). Figure 6.14 displays the exclusion limits at 95% using CL_{s+b} method for the five Higgs mass points with a luminosity of 1 fb^{-1}.

6.4.4 Incorporating systematics uncertainties

Systematic uncertainties arise from both signal and background. In the likelihood function, there are two parameters related to the background, N_B and ξ, which are evaluated by

Sample	N_{evts}
signal	13
$\gamma\gamma$ background	5540
photon-jet background	2500
di-jet background	360
Drell-Yan	90

Table 6.3: Expected number of events obtained for the signal of $m_H = 120$ GeV and the main backgrounds in the invariant mass range $100 < M_{\gamma\gamma} < 150$ GeV after all the selections with a luminosity of 1 fb^{-1} at 7 TeV.

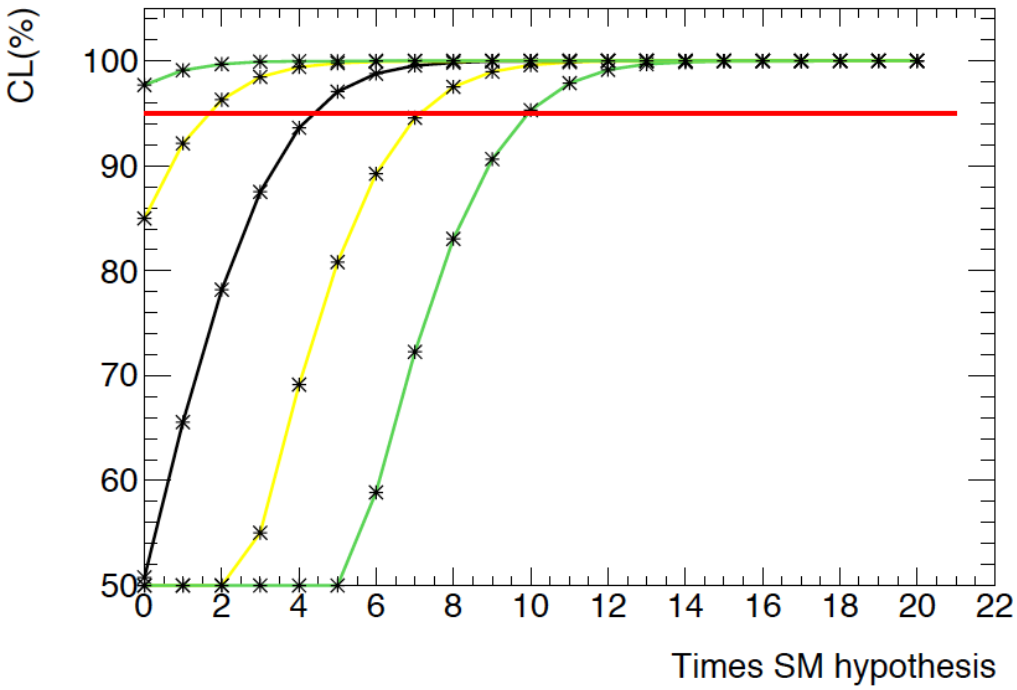


Figure 6.13: Estimated Confidence Levels as a function of the hypothesis on the number of SM signal cross-section for $m_H = 120$ GeV using the CL_{s+b} method at 7 TeV with a luminosity of 1 fb^{-1}.

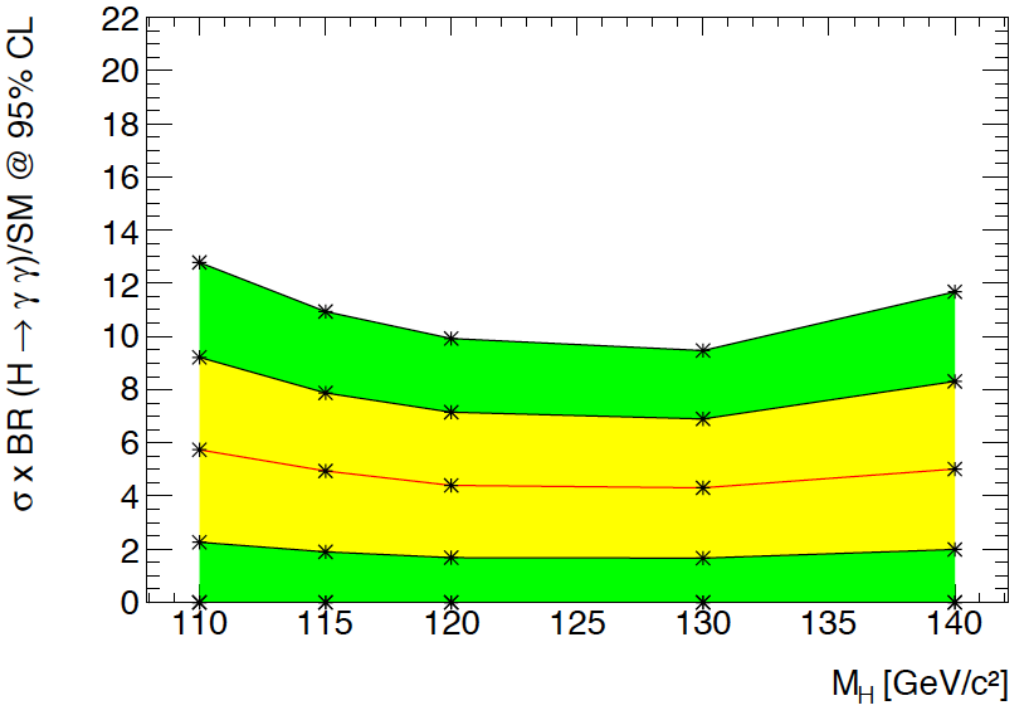


Figure 6.14: Expected exclusion limit of the cross section at 95% CL using the CL_{s+b} method at 7 TeV with a luminosity of 1 fb^{-1}.

the fit. Consequently, the systematic uncertainties which are related to these nuisance parameters are automatically taken into account. However, it is not the case for the signal. Systematics to be considered are the following:

- The width of the Crystal Ball function. It is related to the energy resolution of the detector. As described in [101], if the long range constant term for the energy of the photon varies by 0.5% (1%), the width $\sigma_{\gamma\gamma}$ is increased by 0.8% (13%) in comparison to the resolution with the default long range constant calibration value of 0.7%.

- The uncertainty on the integrated luminosity. Since the overall normalization N_B of the background is directly measured on the data, this uncertainty affects only the signal. For early data analysis, it is at the level of 11%.

- The photon reconstruction efficiency. It is considered only for signal as well since the background normalization N_B is measured directly on data. It is varied by 1% in this analysis.

- The cross-section. Higgs boson production cross-sections have been studied extensively by the LHC Higgs cross-section working group and the results are compiled in [96]. The theoretical uncertainties is estimated to be of 15% in the m_H range relevant for this analysis.

These systematic uncertainties can be incorporated into the exclusion limit with the Cousins and Highland method [102]. The idea is to smear the parameters mentioned above with a Gaussian distribution at the Toy Monte Carlo generator level. The systematic

uncertainty introduced by the $\sigma_{\gamma\gamma}$ of the Crystal-Ball function is evaluated. The default value for $\sigma_{\gamma\gamma}$ of the Crystal Ball function is 1.31 GeV, which is best resolution obtained on the Monte Carlo signal sample study. It is considered to be the central value for the Gaussian smearing. And the σ for the smearing is set to be 0.3 GeV in order to cover the deviation when varying the constant term. Besides, two other central values of 1.41GeV and 1.51 GeV are also studied. For each pseudo experiment, a random value for $\sigma_{\gamma\gamma}$ of the Crystal Ball function is produced by a Gaussian generator, with $\sigma = 0.3$ GeV around the central value. This random value of $\sigma_{\gamma\gamma}$ is then used to generate the signal sample. The value for $\sigma_{\gamma\gamma}$ in the fit is always set to be the central value for the background plus signal hypothesis. In this way, one can derive the q_μ distribution after the smearing for the background plus signal hypothesis, as shown in Figure 6.15.

One can easily see that the test statistics q_μ distribution does not agree with $1/2\chi^2$ distribution after the smearing. Therefore, one need to use the new q_μ distribution to calculate the p-value. In Table 6.4, the exclusion limit results before and after the smearing are shown. The exclusion limit increase by about 4% when incorporating the systematic uncertainty of the mass resolution.

$\sigma_{\gamma\gamma}$ [GeV]	after smearing	no smearing
1.31	4.60	4.41
1.41	4.75	4.60
1.51	4.91	4.75

Table 6.4: Expected exclusion limit of the SM at 95% CL using the CL_{s+b} method at 7 TeV with a luminosity of 1 fb^{-1} with and without the smearing the $\sigma_{\gamma\gamma}$ of the Crystal Ball function.

6.5 Sensitivity on 2010 data

As mentioned before, a robust analysis with only di-photon invariant mass as the discriminant has been performed using 2010 data. In this analysis [103], the event-level trigger efficiency is evaluated using data-driven method and a first measurement of the background on data is presented. The expected exclusion sensitivity with 1 fb^{-1} is also estimated for a Higgs mass range of 110−140 GeV.

6.5.1 Trigger efficiency measurement

As mentioned in Section 6.3.2, two trigger items are used in this analysis: L1_EM14 and 2g15_loose. This section presents the measurement of their efficiency with respect to the Higgs selection.

The measurement of the event-level trigger efficiency for L1_EM14 is done on a sample of events passing the lower threshold trigger L1_EM2. Although this trigger is prescaled, the sample statistics is sufficient. It is found to be fully efficient for tightly identified isolated photons with $p_T > 40$ GeV (or $p_T > 25$ GeV). L1_EM14 is then checked on the Higgs candidates events passing L1_EM2.

The measurement of the efficiency of 2g15_loose follows the same method as discussed in Section 5.6.2, the main difference being the analysis selection. The object-level g15_loose

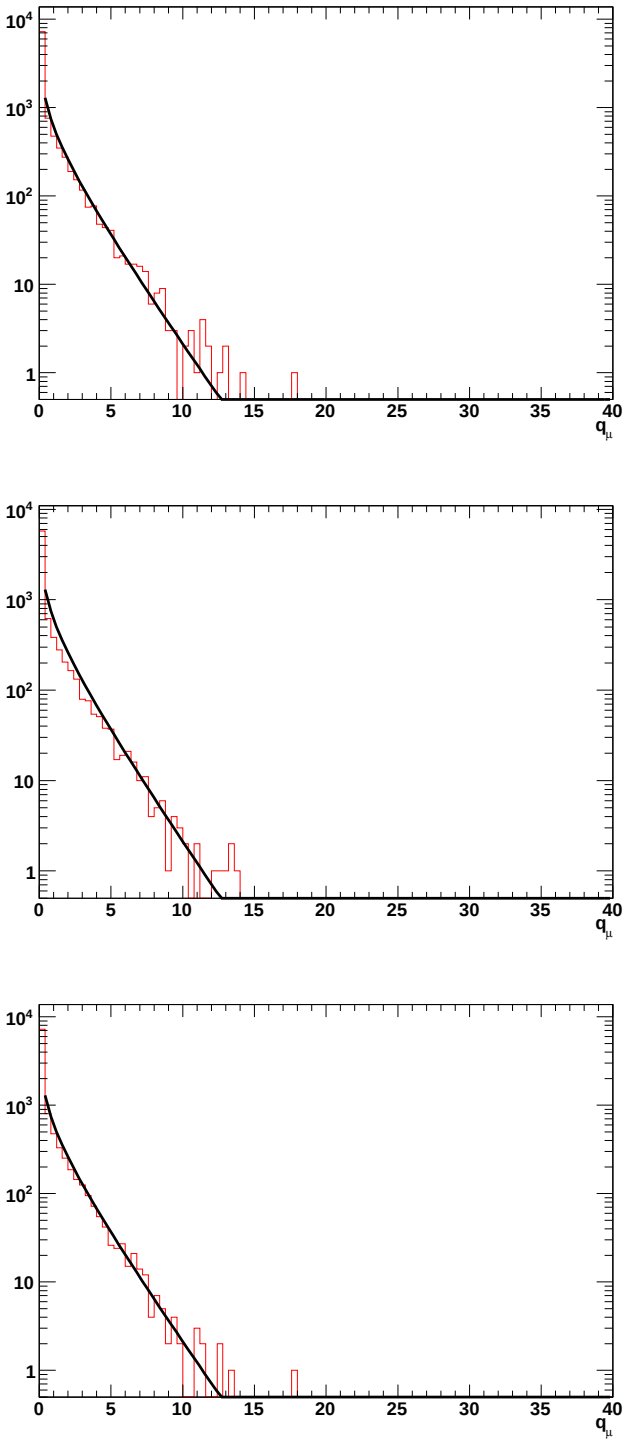


Figure 6.15: The q_μ distribution (red line)after a Gaussian smearing of σ of the Crystal-Ball function, at three central value. Top: 1.31 GeV, Middle: 1.41 GeV, Bottom: 1.51 GeV. The hypothesis is set to be $\mu = 10$. The $1/2\chi^2$ distribution is also overlapped in the plot for a comparison.

efficiency ϵ_1 for photon candidates with transverse energy above 40 GeV is measured by the bootstrap method, and it is found to be $100.0^{+0.0}_{-1.9}\%$). Higgs candidate events with the leading photon passing the g15_loose trigger are selected. The fraction ϵ_2 of events from this sample passing 2g15_loose is found to be $100.0^{+0.0}_{-4.7}\%$. The final event-level trigger efficiency for 2g15 _loose is determined as $\epsilon_1 \times \epsilon_2$.

The event-level trigger efficiencies for L1_EM14 and 2g15_loose are summarized in Table 6.5; only statistical uncertainties are quoted in the Table, to which a $\pm 1\%$ systematic uncertainty is to be added.

trigger item	L1_EM14	2g15_loose
efficiency [%]	$100.0^{+0.0}_{-4.5}$	$100.0^{+0.0}_{-5.0}$

Table 6.5: event level trigger efficiency with respect to the Higgs selection

6.5.2 Analysis on real data: background decomposition using the diphoton analysis technique

The 2D fit method introduced in Section 5.4 can be extrapolated to the Higgs analysis. The same procedure is applied, the only difference being the kinematic cuts:

- Derive the signal photon isolation PDFs from the electron isolation transverse energy distribution in data, separately for leading p_T and subleading p_T photon candidates.
- Derive the fake photon isolation PDFs by reversing a subset of the identification cuts, separately for leading p_T and subleading p_T photon candidates.
- Derive the jj 2D isolation PDF by reversing the identification cuts simultaneously for both photon candidates.
- Sum up the four 2D isolation PDFs: $\gamma\gamma$, γj, $j\gamma$ and jj, and fit to the 2D isolation distribution for the two photons candidates, when both pass the **RobustTight** photon identification.

Before the fit on the data, a validation of the fit method is performed by evaluating the results using the "cocktail" sample of $\gamma\gamma$ and $\gamma j + j\gamma$ simulated events which correspond to an equivalent luminosity of around 110 pb^{-1}. No single jj simulated event is obtained after all the selection cuts. Results after the fit are summarized in Table 6.6. A reasonable agreement is obtained between the fitted number and the true number of events with no significant bias. An equivalent test using a larger sample of inclusive $\gamma\gamma$ candidates has already been shown in Section 5.4.5. An additional systematic uncertainty is assigned to the result of the 2D fit in order to account for the potential bias in the number of fitted events.

83 events are obtained in the invariant mass region [100, 150] GeV after the selection described in Section 6.3.2. Table 6.7 summarizes the decomposition results obtained with the 2D fit method. The 83 Higgs candidate events are dominated by the QCD $\gamma\gamma$ background. Figure 6.16 shows the isolation profile for the leading and the subleading photon candidates. As a check, the number of events in each component is evaluated for different values of the isolation cut, the same cut being applied to the two photon candidates. The

	$N_{\gamma\gamma}$	$N_{\gamma j}$	$N_{j\gamma}$	N_{jj}
True number in Sig. Region	270	30	9	0
Fitted number in Sig. Region	265.1 ± 18.6	40.4 ± 3.5	11.5 ± 2.7	0.1 ± 0.4

Table 6.6: Results of the 2D fit method, applied to a cocktail of simulated $\gamma\gamma$, $\gamma j + j\gamma$ events.

results are shown in Figure 6.17. There is a small increase of the di-photon event yields while the photon-jet and di-jet event yields increase rapidly as the isolation cut is looser. It is the expected pattern that the gain for the photon efficiency is small while the fake rejection decreases rapidly with a looser isolation requirement.

The systematic uncertainties are evaluated in a similar way to what is described in Section 5.4.7. The main three sources of systematics: the definition of non-RobustTight background control region, the relative fraction of direct and fragmentation photons and the correlation between the isolation and the identification variables, are evaluated in this analysis. Results are shown in Table 6.8.

event category	yields
$N_{\gamma\gamma}$	$67.4 \pm 9.6 \pm 8.2$
$N_{\gamma j}$	$10.0 \pm 2.2 \pm 1.5$
$N_{j\gamma}$	$3.4 \pm 1.6 \pm 2.5$
N_{jj}	$1.7 \pm 0.6 \pm 0.1$

Table 6.7: Results from the 2D fit method in the $H \to \gamma\gamma$ sample.

Source	$\Delta N_{\gamma\gamma}$	$\Delta N_{\gamma j}$	$\Delta N_{j\gamma}$
Fake photon PDF	± 8.1	± 0.1	± 2.4
photon PDF	± 0.1	± 0.2	± 0.1
Bias on MC	± 1.2	± 1.4	± 0.9

Table 6.8: Dominant systematic uncertainties on the number of events, evaluated using the 2D fit method.

Similarly, the 2×2D sideband method can also be adapted to this analysis framework. The results are given in Table 6.9. They are consistent with 2D fit results.

event category	2×2D sideband
$N_{\gamma\gamma}$	$68.6 \pm 11.0 \pm 3.4$
$N_{\gamma j}$	$8.6 \pm 4.7 \pm 2.3$
$N_{j\gamma}$	$2.0 \pm 0.9 \pm 3.4$
N_{jj}	$1.2 \pm 0.6 \pm 0.5$

Table 6.9: Results from the 2×2D sideband in the $H \to \gamma\gamma$ sample.

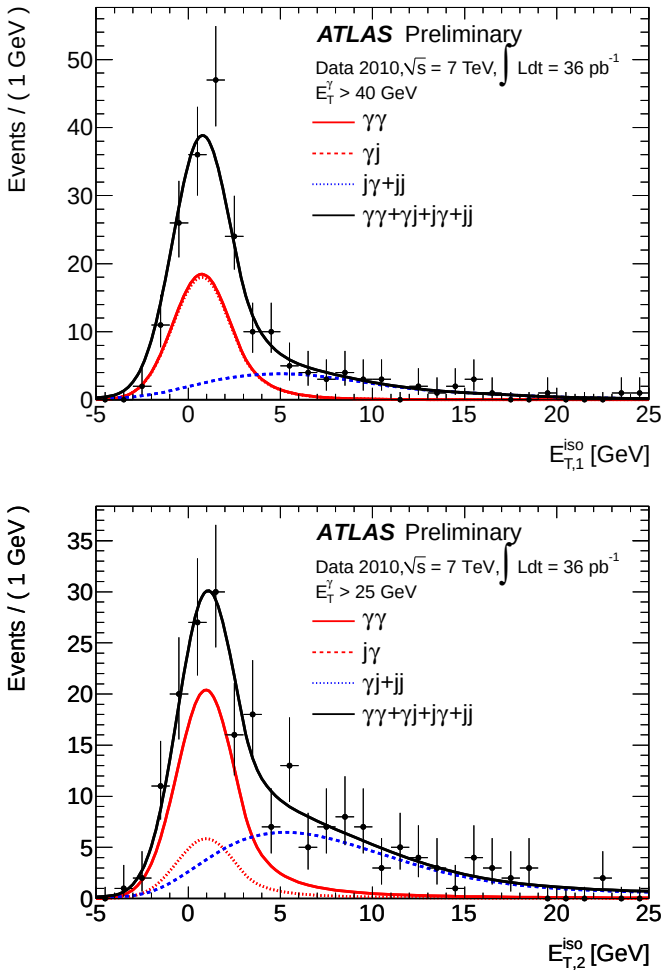


Figure 6.16: Distributions of the isolation variable for the leading (top) and subleading (bottom) photon candidates, for the $H \to \gamma\gamma$ sample. The overlaid curves are the result of the 2D template fit for each of the components.

6.5.3 Comparison with the Monte Carlo prediction

If the numbers of di-photon, photon-jet and di-jet events reported in Section 6.4.3 are normalized to the integrated luminosity of the full 2010 data sample, 203, 92 and 13 events are expected respectively. The yields actually measured on the data are significantly lower. However, there are many differences between the Monte Carlo and the 2010 data analyses, which were studied in details within ATLAS. It is found that the main effect comes from the change of the isolation criterion from the track-based isolation ($E_{\mathrm{T}}^{\mathrm{track}} < 4$ GeV) to the calorimetric isolation requirement ($E_{\mathrm{T}}^{\mathrm{iso}} < 3$ GeV). The latter is found to be more effective in rejecting the background, reducing significantly the photon-jet and di-jet contributions. The prediction of the irreducible background must be corrected by many effects, none of them playing a dominant role. They include the change of isolation criterion already mentioned, for a given isolation requirement the difference of efficiency between Monte Carlo and data, the reduction of the acceptance in the data from the OQ maps. Some theoretical factors

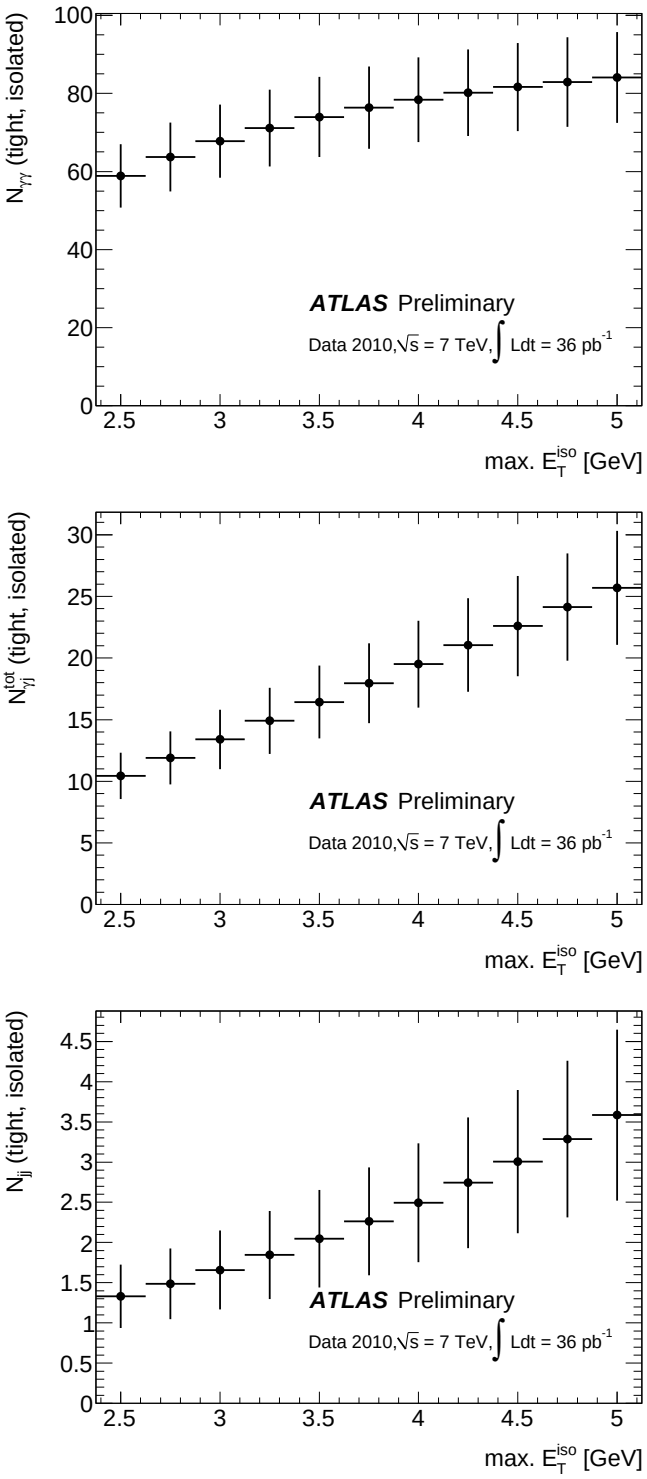


Figure 6.17: The number of diphoton (top), photon-jet (middle) and dijet (bottom) events, as a function of the cut on the isolation transverse energy.

also impact the total predicted number of $\gamma\gamma$ events. One major effect comes from using the actual parton density function for 7 TeV centre-of-mass energy collisions, whereas in the 7 TeV Monte Carlo study presented in Section 6.4.3, the $\gamma\gamma$ contribution was just rescaled from the 10 TeV Monte Carlo.

After these corrections are applied, the predictions for the main background from Monte Carlo simulation are summarized in Table 6.10. The number of $\gamma\gamma$ events agrees with the estimated ones shown in Sections 6.5.2 within the theoretical systematic uncertainties of $\pm 27\%$ [105]. Figure 6.18 shows the di-photon invariant mass spectrum for the 2010 data sample, and the comparison of the prediction. The blue curves represent the background composition estimated by the 2D fit method. They are overlaid in a cumulative way: jj (dotted), $\gamma j + j\gamma$ (dashed) and $\gamma\gamma$ (solid). The exponential coefficient for each component is derived from the MC 10 TeV analysis. The sum of $\gamma\gamma$ NLO prediction and the estimated yields for $\gamma j + j\gamma + jj$ from 2D fit method is also shown (the light blue line). The background from Drell-Yan process is estimated with a standalone technique, and is found to contribute around 3 events.

event category	yields
$N_{\gamma\gamma}$	86 ± 23
$N_{\gamma j} + N_{j\gamma}$	27 ± 14
N_{jj}	1 ± 1

Table 6.10: The expected number of events after the experimental and theoretical corrections for Monte Carlo simulations.

6.5.4 Projected sensitivity

With 37 pb^{-1}, the expected limit of the exclusion sensitivity is around 20 to 30 times the cross-section predicted for the SM Higgs boson. This expected limit is already comparable with recent results from Tevatron in the $H \to \gamma\gamma$ channel [106] [107].

This observed inclusive distribution of di-photon events is used to estimate the sensitivity for the $H \to \gamma\gamma$ channel with 1 fb^{-1}. In the 2011 data-taking period, the detector acceptance is expected to increase after the repair of faulty optical links in the Liquid Argon calorimeter (LAr) readout system. The average interactions per bunch crossing is expected to increase to around 5, which will reduce the signal efficiency and increase the width of the invariant mass. The event rates in the 2010 data sample are then rescaled to an integrated luminosity of 1 fb^{-1}, taking into account all the relevant effects mentioned before. The projected sensitivity is then evaluated based on the rescaled sample. Results are shown in Table 6.11. A Higgs boson produced with 3.2$\sim$4.2 times the SM predicted cross-section in the 110$-$140 GeV mass range can be excluded with 1 fb^{-1} of data at $\sqrt{s} = 7$ TeV. Figure 6.19 shows the central value of the exclusion limit as a function of the Higgs mass, together with $\pm 1\sigma$ and $\pm 2\sigma$ deviation. The sensitivity is found to degrade by $+0.5$ due to the deterioration of photon energy resolution as shown in the dashed line.

6.6 Conclusion and prospects

A dedicated study of the event-level trigger efficiency was performed both on MC simulations and on 2010 data. Results show that the trigger items used in the study are very efficient

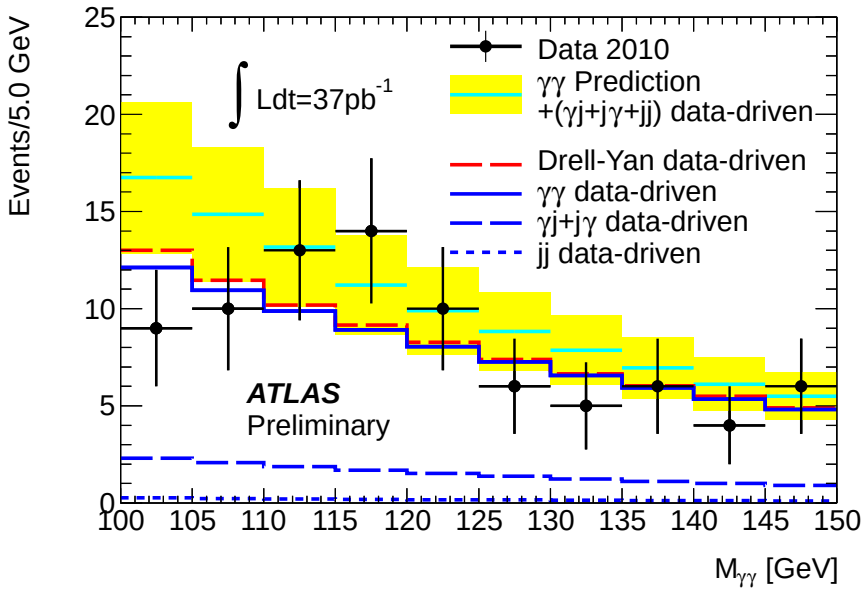


Figure 6.18: The di-photon invariant mass the 2010 data sample after the Higgs selection. The estimated background composition from 2D fit method is overlaid. The blue curves represent the cumulative jj (dotted), γj+$j\gamma$ (dashed) and $\gamma\gamma$ (solid) components. The red curve corresponds to the Drell-Yan contribution. The yellow band is the prediction for the sum of the reducible and the irreducible background components, where the reducible background is normalized to the result of the 2D fit method, and the irreducible component is normalized to the di-photon NLO prediction.

Higgs boson mass [GeV]	110	115	120	130	140
Upper bound [× SM]	3.7	3.5	3.2	3.3	4.2

Table 6.11: The upper bound of the exclusion at 95% CL, in units of the Standard Model Higgs boson cross-section, as a function of the Higgs boson mass found by projecting the current data to 1 fb^{-1}.

in selecting the signal events. This trigger efficiency evaluation helped the physics group to understand the trigger performance, and to give feedback on physics need to the trigger group for the future data taking. In the 2010 data taking, the 2g15_loose was used as the primary di-photon trigger. At an instantaneous luminosity of around 10^{32} $cm^{-1}s^{-1}$, the corresponding bandwidth was about 1 Hz. Since this trigger item works efficiently for the di-photon and $H \rightarrow \gamma\gamma$ analyses, it is expected to be used as a primary trigger as much as possible, hopefully until the instantaneous luminosity reaches 6×10^{32} $cm^{-1}s^{-1}$. Afterwards, 2g20_loose will become the primary di-photon trigger in order to balance the physics need and the bandwidth limit.

A study on incorporating the systematic uncertainties into the exclusion limit using the Cousins and Highland prescription was presented. A deterioration of 4% of the exclusion limit is observed when taking into account the systematic uncertainty arising from the resolution of di-photon invariant mass. Similar studies can be performed for the systematic

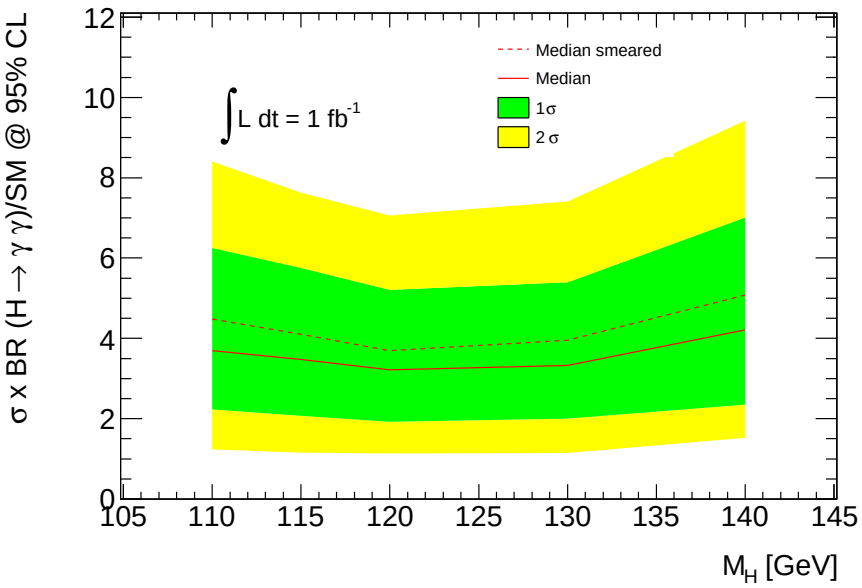


Figure 6.19: The estimated number of excluded Standard Model signal cross section at 95% CL as a function of the Higgs boson mass by projecting the current data to 1 fb^{-1}. The green (respectively yellow) bands correspond to the expected exclusion in the case of a $\pm 1\sigma$ (respectively $\pm 2\sigma$) fluctuation of the background only test statistic. The dashed curve corresponds to the exclusion after degrading the energy resolution of the photons from signal. The deterioration of the energy resolution is based on the comparison of the Z$\rightarrow e^+e^-$ sample between 37 pb^{-1} data and simulation may be pessimistic for 1 fb^{-1} data.

uncertainties from the luminosity, photon efficiency and signal cross-section.

In the 2010 data sample, 83 events passing all selection cuts in the invariant mass region $100 < M_{\gamma\gamma} < 150$ GeV were observed. The 2D fit method was applied to extract the background composition in terms of $\gamma\gamma$, γj, $j\gamma$ and jj events. The results are consistent with another data-driven method within 5%. The overall background and each of its contributions measured in data are compatible with the predictions. These results are used to reappraise the sensitivity of the ATLAS search for the Higgs boson in the two photon channel. With a simple inclusive search for a narrow resonant production of two photons and $1fb^{-1}$ of data at $\sqrt{s} = 7$ TeV, ATLAS expects to exclude a Higgs boson produced with 3.2$\sim$4.2 times the SM predicted cross-section in the 110-140 GeV mass range.

Chapter 7

Conclusion

Three independent although correlated topics have been presented in this thesis. The three analyses were performed using the data recorded by the ATLAS experiment in 2010. They are all related to the di-photon system.

- A dedicated study of the photon trigger efficiency measurement was presented. Three independent data-driven methods were proposed and tested on Monte Carlo samples, namely the *tag&probe*, *bootstrap* and *electron-to-photon-extrapolation* methods. A detailed comparison of the three methods with respects to the true photon efficiency was performed. The two methods based on photon measurements (tag&probe and bootstrap) agree with each other within the estimated uncertainties, while the extrapolation to photons of the electron sample selected from $Z \rightarrow ee$ decays shows larger biases, in particular in the trigger turn-on region around the p_T threshold at 20 GeV. In the high p_T region ($p_T >$ 25 GeV), the three methods agree, essentially because this is the plateau region where the efficiency is close to 100% for both true and fake photons. The total uncertainty from any of these data-driven measurements performed on 20 pb^{-1} of data was estimated to be well below 1%, in the p_T region above the threshold, which is relevant for the main photon-based analyses. In the lower p_T region, around the trigger turn-on, an uncertainty of a few % is expected from the Monte Carlo study. These three methods were applied on the real data. As expected, the two photon-based methods show consistent results, with efficiency discrepancy at a few % level, well within the statistical uncertainty. For the extrapolation method, the available statistics in the p_T turn-on region was too low to perform a meaningful comparison.

- The measurement of the QCD isolated di-photon cross-section measurement was performed. The interest of this measurement is two-fold: it is a probe of QCD, and it is the main background for the SM Higgs search in the two-photon final state. One of the crucial points in this measurement is the extraction of the di-photon signal yields. A two-dimensional isolation fit method was developed. The total 2D isolation distribution from a preselected sample was modeled with the sum of four components: one signal ($\gamma\gamma$) and three background (γj, $j\gamma$, jj) categories, weighted by their corresponding yields. The jet probability density functions entering the fit were extracted from data. The photon PDFs were derived from the electron distribution in data corrected by Monte Carlo. An integrated di-photon signal yield of $1353 \pm 53^{+217}_{-128}$ was measured. This technique was applied in intervals of the di-photon observables: the

mass ($M_{\gamma\gamma}$) and the transverse momentum ($p_{T\gamma\gamma}$) of the di-photon system, and the azimuthal angle ($\Delta\phi_{\gamma\gamma}$) between the two photons. The measurement of the event-level trigger efficiency for di-photon candidate events was performed. It makes use of the methods that are developed to measure the object-level efficiency. The technique adopted for the measurement of the di-photon trigger efficiency is unbiased with respect to the correlation between the two photon candidates. After taking into account the photon identification efficiency, the unfolding factor and the reconstruction efficiency, the differential cross-sections as a function of $M_{\gamma\gamma}$, $p_{T\gamma\gamma}$ and $\Delta\phi_{\gamma\gamma}$ are derived and compared to the NLO prediction from generators (DIPHOX and RESBOS). A good agreement is found for the $M_{\gamma\gamma}$ differential cross-section, especially in the comparison with RESBOS. It is not the case for the differential cross-sections as a function of $p_{T\gamma\gamma}$ and $\Delta\phi_{\gamma\gamma}$ which are sensitive to the parton-to-photon fragmentation and collinear emission of soft gluons in the initial state: DIPHOX and RESBOS both fail to describe the experimental data, although this must be confirmed by more precise measurements.

- The last part of the thesis is the $H \rightarrow \gamma\gamma$ analysis. The measurement of the trigger efficiency done for the di-photon analysis was repeated for the triggers used in this analysis. A study on incorporating the systematic uncertainties into the exclusion limit using Cousins and Highland perscription is presented. A deterioration of 4% in the exclusion limit is observed when taking into account the systematic uncertainty from the resolution of di-photon invariant mass. Similar studies can be performed for the systematic uncertainties arising from the luminosity, the photon efficiency and the signal cross-section. The 2D isolation fit method developed for the di-photon analysis was applied to analyze the various background components in the di-photon invariant mass region $100 < M_{\gamma\gamma} < 150$ GeV. The irreducible two-photon background was measured to be 80% of the total. These results contributed to the understanding of the 2010 data sample, which was used to estimate the sensitivity of the Higgs search in the two-photon channel. With a simple inclusive search on $1fb^{-1}$ of data at $\sqrt{s}$ = 7 TeV, ATLAS expects to exclude a Higgs boson produced with 3.2$\sim$4.2 times the SM predicted cross-section in the 110-140 GeV mass range.

The trigger is evolving due to the rapid increase of the LHC luminosity. At present, the main di-photon trigger is similar to the one that was studied in this thesis, except the transverse energy threshold is higher (20 GeV). In the future, one might need to consider trigger with tighter identification cuts, or asymmetric transverse energy trigger (g20_g30_loose for example), or including isolation requirements. In the latter case, isolation-based analyses as those presented here must be redefined. With the accumulation of new data, the di-photon cross-section measurement becomes even more interesting. Already in 2011, it will be possible to greatly reduce the uncertainties and provide rather accurate tests of the models, and even double-differential cross-sections. The most exciting study is of course the search for the Higgs. Given the very good performances of the LHC, one can expect to reach the sensitivity for a SM Higgs with the two-photon channel alone, in the low mass range. This requires to go beyond the inclusive approach presented here, by including more discriminant variables or candidate categories. By combining more decay channels, prospective studies show that ATLAS will be able to exclude the Higgs boson in the 115-500 GeV mass range with 4 fb^{-1} of data [108]. Such an integrated luminosity is indeed reachable by the end of the present LHC run. As far as an evidence for Higgs is concerned, the region below 150

GeV is the most difficult one. The $H \to \gamma\gamma$ study will be therefore a crucial analysis, if the Higgs is there.

Bibliography

[1] Donald H. Perkins, "Introduction to High Energy Physics", Cambridge University press 4th Edition.

[2] S. Weinberg, "The Making of the Standard Model", hep-ph/0401010. Jan 2004.

[3] S. Glashow, Nucl. Phys. **22**, 579(1961).

[4] S. Weinberg, Phys. Rev. Lett. **19**, 1264(1967).

[5] M. Gell-Mann, Phys. Lett. **8**, 214(1964).

[6] G. Zweig, CERN-Report 8182/TH40 (1964).

[7] H. Fritzsch, M. Gell-Mann and H. Leutwyler, Phys. Lett. B **47**, 365(1973).

[8] D. Gross and F. Wilczek, Phys. Rev. Lett. **30**, 1343(1973).

[9] H. D. Politzer, Phys. Rev. Lett. **30**, 1346(1973).

[10] S. Coleman and D. J. Gross, Phys. Rev. Lett. **31**, 851(1973).

[11] J. Goldstone, Nuov. Cim **19**, 154(1961).

[12] B. W. Lee, C. Quigg and H. B. Thacker, Phys. Rev. D **16**, 1519(1977).

[13] T. Cheng, E. Eichten and L. Li, Phys. Rev. D **9**, 2259(1974).

[14] B. Pendleton and G. Ross, Phys. Lett. B **98**, 291(1981).

[15] C. Hill, Phys. Rev. D **24**, 691(1981).

[16] T. Hambye and K. Riesselmann, "Matching conditions and Higgs mass upper bounds revisited", Phys. Rev. D **55**, 7255 (1997).

[17] ALEPH Collaboration, "Final results of the searches for neutral Higgs bosons in e^+e^- collisions at $\sqrt{s}$ up to 209GeV", Phys. Lett. B **526**, 191 (2002).

[18] G. Abbiendi et al, "Search for the Standard Model Higgs Boson at LEP", hep-ex/0306033, Jun 2003.

[19] The Tevatron Electroweak Working Group for the CDF and the DØ Collaborations, "Updated Combination of CDF and DØ results for the Mass of the W Boson", Eprint arXiv:0908.1374 [hep-ex], 2009.

[20] The Tevatron Electroweak Working Group for the CDF and the DØ Collaborations, "Combination of CDF and DØ Results on the Width of the W boson", Eprint arXiv:1003.2826 [hep-ex], 2010.

[21] The Tevatron Electroweak Working Group for the CDF and the DØ Collaborations, "Combination of CDF and DØ Results on the Mass of the Top Quark", Eprint arXiv:1007:3178 [hep-ex], 2010.

[22] C. S. Wood et al, Science **275**, 1759(1997).

[23] SLAC E158 Collaboration, P. Anthony et al, Phys. Rev. Lett. **92**, 181602(2004).

[24] NuTeV Collaboration, G. P. Zeller et al, Phys. Rev. Lett. **95**, 2002, 091802, erratum: 90(2003) 239902

[25] The LEP Collaborations, the LEP Electroweak Working Group and the SLD Electroweak and Heavy Flavour Groups, "Precision Electroweak Measurements and Constraints on the Standard Model", Phys. Rept. **427**, 257(2006).

[26] The LEP collaborations, the LEP Electroweak Working Group, the Tevatron Electroweak Working Group and the SLD electroweak and heavy flavour groups, "Precision Electroweak Measurements and Constraints on the Standard Model", arXiv:1012.2367v2, Jan 2011.

[27] The CDF, DØ collaborations and the Tevatron New Physics, Higgs Working Group, "Combined CDF and DØ Upper Limits on Standard Model Higgs Boson Production with up to 6.7 fb^{-1} of Data", arXiv:1007.4587v1, Jul 2010.

[28] The CDF, DØ collaborations, the TEVNPHWG Working Group, "Combined CDF and DØ Upper Limits on Standard Model Higgs Boson Production with up to 8.2 fb^{-1} of Data", arXiv:1103.3233v2, Mar 2011.

[29] Gfitter group, "A Generic Fitter Project for HEP Model Testing", "http://gfitter.desy.de/".

[30] E. Bonvin, et al., WA70 Collaboration, Phys. Lett. B **236**, 523(1990).

[31] M. Begel, Ph. D. Thesis, University of Rochester, FERMILAB-THESIS-1999-05, 1999.

[32] C. Albajar, et al., UA1 Collaboration, Phys. Lett. B **209**, 385(1988).

[33] J. Alitti, et al., UA2 Collaboration, Phys. Lett. B **288**, 386(1992).

[34] CDF Collaboration, "Measurement of the Cross Section for Prompt Diphoton Production with 5.4 fb^{-1} of CDF Run II Data", http://www-cdf.fnal.gov/physics/new/qcd/diphoXsec_2010/public_diphoton.html, 2010.

[35] V. M. Abazov, et al., D0 Collaboration, Phys. Lett. B **690**, 108(2010).

[36] T. Sjostrand, S. Mrenna and P. Skands, "PYTHIA 6.4 Physics and Manual", arXiv:hep-ph/0603175v2.

[37] E. Lyndon (ed.) and B. Philip (ed.)", "LHC Machine", J. Instrum. **3**, S08001 (2008).

[38] ALICE Collaboration, "The ALICE experiment at the CERN LHC", J. Instrum. **3**, S08002 (2008).

[39] ATLAS Collaboration, "The ATLAS Experiment at the CERN Large Hadron Collider", J. Instrum. **3**, S08003 (2008).

[40] CMS Collaboration, "The CMS experiment at the CERN LHC", J. Instrum. **3**, S08004 (2008).

[41] LHCb Collaboration, "The LHCb Detector at the LHC", J. Instrum. **3**, S08005 (2008).

[42] LHCf Collaboration, "The LHCf detector at the CERN Large Hadron Collider", J. Instrum. **3**, S08006 (2008).

[43] TOTEM Collaboration, "The TOTEM experiment at the CERN Large Hadron Collider", J. Instrum. **3**, S08007 (2008).

[44] R. LEY PS Division.", CERN 02.09.96 revised and updated by A. Del Russo, HIT Div, in collaboration with H. Desforges and D. Manglanki, CERN, 23.05.01.

[45] A. Yamamoto et al, "The ATLAS central solenoid", Nucl. Instrum. Meth. A **584**, 53 (2008).

[46] A. Foussat et al, "Assembly concept and technology of the ATLAS barrel toroid", IEEE T. Appl. Supercond **16**, 565 (2006).

[47] D. E. Baynham et al, "Engineering status of the end cap toroid magnets for the ATLAS experiment at LHC", IEEE T. Appl. Supercond **10**, 357 (2000).

[48] ATLAS Collaboration, "ATLAS inner detector : Technical Design Report, 1", ATLAS-TDR-004,CERN-LHCC-97-016, CERN, Geneva, 1997.

[49] ATLAS Collaboration, "ATLAS inner detector : Technical Design Report, 2", ATLAS-TDR-005,CERN-LHCC-97-017, CERN, Geneva, 1997.

[50] ATLAS Collaboration, "ATLAS pixel detector : Technical Design Report", ATLAS-TDR-011,CERN-LHCC-98-013, CERN, Geneva, 1998.

[51] A. Abdesselam et al, "The barrel modules of the ATLAS semiconductor tracker", Nucl. Instrum. Meth. A **575** 642 (2006).

[52] A. Abdesselam et al, "The ATLAS semiconductor tracker end-cap module", Nucl. Instrum. Meth. A **575** 353 (2007).

[53] E. Abat et al, "The ATLAS TRT Barrel Detector", JINST **3**, P02014 (2008).

[54] E. Abat et al, "The ATLAS TRT End-cap Detectors", JINST **3**, P10003 (2008).

[55] ATLAS Collaboration, "ATLAS liquid-argon calorimeter : Technical Design Report", ATLAS-TDR-002,CERN-LHCC-96-041, CERN, Geneva, 1996.

[56] ATLAS Collaboration, "ATLAS tile calorimeter : Technical Design Report", ATLAS-TDR-003,CERN-LHCC-96-042, CERN, Geneva, 1996.

[57] B. Aubert et al, "Construction, assembly and tests of the ATLAS electromagnetic barrel calorimeter", Nucl. Instrum. Meth. A **558**, 388 (2006).

[58] M. Aleksa et al, "Construction, assembly and tests of the ATLAS electromagnetic end-cap calorimeter", JINST **3**, P06002 (2008).

[59] J. Abdallah et al, "Design, construction and installation of the ATLAS hadronic barrel scintillator-tile calorimeter", ATL-TILECAL-PUB-2008-001, CERN, Geneva, 2007.

[60] D. M. Gingrich et al, "Construction, assembly and testing of the ATLAS hadronic end-cap calorimeter", JINST **2**, P05005 (2007).

[61] ATLAS Collaboration, "ATLAS muon spectrometer : Technical Design Report", ATLAS-TDR-010,CERN-LHCC-97-022, CERN, Geneva, 1997.

[62] M. Deile et al, "Performance of the ATLAS precision muon chambers under LHC operating conditions", Nucl. Instrum. Meth. A **518**, 65 (2004).

[63] ATLAS Collaboration, "ATLAS Trigger Performance : Status Report", CERN-LHCC-98-015, CERN, Geneva, 1998.

[64] ATLAS Collaboration, "ATLAS level-1 trigger : Technical Design Report", ATLAS-TDR-012,CERN-LHCC-98-014, CERN, Geneva, 1998.

[65] ATLAS Collaboration, "ATLAS DAQ, EF, LVL2 and DCS : Technical Progress Report", CERN-LHCC-98-016, CERN, Geneva, 1998.

[66] P. Vankov, "ATLAS Silicon Microstrip Tracker Operation and Performance", ATL-INDET-PROC-2010-041, CERN, Geneva, 2010.

[67] J. D. Degenhardt, "Performance of the ATLAS Transition Radiation Tracker with Cosmic Rays and First High Energy Collisions at the LHC", ATL-INDET-PROC-2010-040, CERN, Geneva, 2010.

[68] L. Masetti, "ATLAS: Status Report and Performance for Physics Objects", ATL-PHYS-PROC-2011-007, CERN, Geneva, 2011.

[69] ATLAS Collaboration, "Evidence for prompt photon production in pp collisions at $\sqrt{s}$ = 7 TeV with the ATLAS detector", ATLAS-CONF-2010-077, CERN, Geneva, 2010.

[70] ATLAS Collaboration, "Mapping the material in the ATLAS Inner Detector using secondary hadronic interactions in 7 TeV collisions", ATLAS-CONF-2010-058, CERN, Geneva, 2010.

[71] ATLAS Collaboration, "Photon Conversions at $\sqrt{s}$ = 900 GeV measured with the ATLAS Detector", ATLAS-CONF-2010-007, CERN, Geneva, 2010.

[72] ATLAS Collaboration, "Probing the response of the ATLAS electromagnetic calorimeter and material upstream with energy flow from $\sqrt{s}$ = 7 TeV minimum bias events", ATLAS-CONF-2010-037, CERN, Geneva, 2010.

[73] K. Tackmann, "ATLAS Inner Detector material studies", ATL-PHYS-PROC-2010-051, CERN, Geneva, 2010.

[74] W. Lampl et al, "Calorimeter Clustering Algorithms : Description and Performance", Technical Report ATL-LARG-PUB-2008-002, ATL-COM-LARG-2008-003, CERN, Geneva, Apr 2008.

[75] ATLAS Collaboration, "Reconstruction of Photon Conversions, in Expected Performance of the ATLAS Experiment: Detector, Trigger and Physics", CERN-OPEN-2008-020, arXiv:0901.0512, CERN, Geneva, 2008.

[76] D. Banfi et al., "The measurements of the energy of electrons and photons in the ATLAS electromagnetic calorimeters", ATL-PHYS-INT-2010-038, CERN, Geneva, 2010.

[77] ATLAS Collaboration, "Expected photon performance in the ATLAS experiment", ATL-PHYS-PUB-2011-007, CERN, Geneva, 2011.

[78] ATLAS Collaboration, "Performance of the Electron and Photon Trigger in p-p Collisions at a centre of mass energy of 900 GeV", ATLAS-CONF-2010-022, CERN, Geneva, 2010.

[79] D. Malon et al., "Grid-Enabled Data Access in the ATLAS Athena Framework", Proceedings of International Conference on Computing in High Energy and Nuclear Physics, 2001.

[80] S. Mrenna C. Balazs, E. L. Berger and C. -P. Yuan, "Photon pair production with soft gluon resummation in hadronic interactions", Phys. Rev. D, 57:6934, (1998).

[81] P. Nadolsky C. Balazs, E. L. Berger and C. -P. Yuan, "Photon pair production with soft gluon resummation in hadronic interactions", Phys. Rev. D, 76:013009, (2007).

[82] A. Djouadi, D. Graudenz, M. Spira and P. Zerwas, Nucl. Phys. B **453**, 17 (1995).

[83] T. Binoth, J.-Ph. Guillet, E. Pilon and M. Werlen, Eur. Phys. J. C **16**, 311 (2000).

[84] ATLAS Collaboration, "AMI database for ATLAS datasets", "http://ami.in2p3.fr:8080/AMI/servlet/net.hep.atlas.Database.Bookkeeping.AMI.Servlet.Command?linkId=512".

[85] ATLAS Collaboration, "Data-Quality Requirements and Event Cleaning for Jets and Missing Transverse Energy Reconstruction with the ATLAS Detector in Proton-Proton Collisions at a Center-of-Mass Energy of $\sqrt{s}$ = 7 TeV", ATLAS-CONF-2010-038, CERN, Geneva, June 2010.

[86] W. Lampl **et al**, "Calorimeter Clustering Algorithms: Description and Performance", Technical Report ATL-LARG-PUB-2008-002, ATL-COM-LARG-2008-003, CERN, Geneva, Apr 2008.

[87] ATLAS Collaboration, "Performance of primary vertex reconstruction in proton-proton collisions at $\sqrt{s} = 7$ TeV in the ATLAS experiment", ATLAS-CONF-2010-069, CERN, Geneva, Jul 2010.

[88] R. Fruhwirth, W. Waltenberger and P. Vanlaer", "Adaptive vertex fitting", J. Phys. G **34**,(2007).

[89] J. E. Gaiser, "Charmonium Spectroscopy from Radiative Decays of the J/Psi and Psi-Prime", Ph. D. Thesis, SLAC-R-255, 178 (1980).

[90] ATLAS Collaboration, "Measurement of the inclusive isolated photon cross section in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector", Phys. Rev. D **83**, 052005 (2011).

[91] K. S. Cranmer, "Kernel Estimation in High-Energy Physics", Comput. Phys. Commun. **136**, 198-207 (2001).

[92] M. Cacciari, G. P. Salam and G. Soyez, "The Catchment Area of Jets", JHEP **04**, 042 (2008).

[93] G. D' Agostini, "A Multidimensional unfolding method based on Bayes' theorem", Nucl. Instr. Method. A **362**, 487 (1995).

[94] ATLAS Collaboration, "ATLAS detector and physics performance: Technical Design Report, 2", ATLAS-TDR-015,CERN-LHCC-99-015, CERN, Geneva, 1999.

[95] ATLAS Collaboration, "Prospects for the Discovery of the Standard Model Higgs Boson Using the H to gamma gamma Decay", ATL-PHYS-PUB-2009-053,ATL-COMPHYS-2009-204, CERN, Geneva, 2009.

[96] LHC Higgs Cross Section Working Group and Dittmaier, S. and Mariotti, C. and Passarino, G. and Tanaka (Eds.), R.", "Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables", CERN-2011-002, CERN, Geneva, 2011, arXiv:hep-ph1101.0593.

[97] N. Andari et al., "Higgs Production Cross Sections and Decay Branching Ratios", ATL-COM-PHYS-2010-046, CERN, Geneva, Jan 2010.

[98] A. Djouadi, "The Anatomy of Electro-Weak Symmetry Breaking. I: The Higgs boson in the Standard Model", hep-ph/0503172, May 2005.

[99] M. L. Mangano et al., "ALPGEN, a generator for hard multiparton processes in hadronic collisions", JHEP **07**, 001 (2003).

[100] M. Escalier et al., "Reconstruction of the z vertex and direction of the photon", ATL-PHYS-INT-2010-013, CERN, Geneva, Jan 2010.

[101] R. Ishmukhametov et al., "The effect of detector non-uniformity on the exclusion limit calculation for the Higgs particle in the $H \to \gamma\gamma$ channel", ATL-PHYS-INT-2010-055, CERN, Geneva, Jun 2010.

[102] R. D. Cousins and V. L. Highland, "Incorporating systematic uncertainties into an upper limit", Nucl. Instrum. Meth. A **320**, 331-335 (1992).

[103] HSG1 working group, "Measurement of the backgrounds to the $H \to \gamma\gamma$ search and reappraisal of its sensitivity with 37 pb^{-1} of data recorded by the ATLAS detector", ATLAS-COM-CONF-2011-006, CERN, Geneva, Jan 2011.

[104] ATLAS Collaboration, "ATLAS Sensitivity Prospects for Higgs Boson Production at the LHC Running at 7 TeV", ATL-PHYS-PUB-2010-009, CERN, Geneva, July 2010.

[105] Bernat, P. and Fayard, L. and Kado, M. and Polci, F., "Irreducible background and signal studies in the $H \to \gamma\gamma$ analysis", ATL-COM-PHYS-2009-253, CERN, Geneva, 2009.

[106] Kilminster, Ben, "Higgs boson searches at the Tevatron", arXiv:hep-ex1012.2172.

[107] Peters, Krisztian, "Search for the Higgs boson in the gamma gamma final state at the Tevatron", arXiv:hep-ex1009.0859.

[108] ATLAS Collaboration, "Further investigations of ATLAS Sensitivity to Higgs Boson Production in different assumed LHC scenarios", ATL-PHYS-PUB-2011-001, CERN, Geneva, Jan 2011.

Acknowledgements

Finalizing this document, plenty of past scenes come into my mind. Those scenes make me recall the support and help from my supervisors, my friends and my family. At this stage, I would like to express my deep acknowledgements to all of you.

I would like to first express my gratitude to my Chinese supervisor Shan Jin. It is him leading me into the research field of experimental particle physics. It is also him suggesting me to join the ATLAS experiment, to work in the frontier of the physics analysis in the hadron colliding. I could not make this work without his support. I adore his critical thinking, rigorous scholarship style and meticulous research spirit. His passion for the physics and outstanding research work set an excellent example for me.

Special acknowledgement should be delivered to my French supervisor Lydia Roos. It is her leading me into the world of the ATLAS experiment. It is her following every step and every progress of my work. It is her correcting my wrong attitude towards the research work. It is her setting the self-confidence for my own. It is her training me in searching for the right way to solve a problem. It is her persuading me to persevere in a work and reach the end. Most of the work presented in this document can not be fulfilled without her participation. Her fast reaction to the physics problem, her comprehensive analysis, her passion, insistence and patience in the work impress me. I am very glad that I have done my Ph.D work with her.

I would like to express many thanks to my colleagues in LPNHE in Paris. Special thanks to Giovanni Marchiori. I am really impressed by his excellent computing skills and rather efficient work. Each time I went to him for a computing or physics problem, he can always help me to find a good solution or offer some wonderful ideas. I appreciate a lot the fruitful discussion and the pleasant jointly work with him. Thanks to Sandro De Cecco for the fruitful and patient discussion on the photon trigger efficiency measurement. Without his nice work and many help, I could not finish that work. Many thanks to other people in the Higgs group in LPNHE, Bertrand Laforge, Jose Ocariz, Giovanni Calderini, Sandrine Laplace, Yuji Enari, Heberth Torres, Trinh Nguyet, Olivier Davignon, Camilla Rangel, Irena Nikolic and Marine Kuna for the active discussion and good suggestions for my work during the group meetings. Thanks to Federic Derue for his great expertise in the egamma software. Thanks to Dikai Li for the helpful discussion about the technic problems encountered during the analysis work, and also the great help in the everyday life. I would like to thank Marumi Kado from LAL in Orsay for his listening and encouragement for my various work. Thanks to Marc Escalier for offering the help in checking my document and good suggestions.

I would like to thank my colleagues in IHEP in Beijing as well. Yanping Huang, her great passion and insistence in the research work sets a great example for me. Yu Bai, his excellent understanding of the physics impress me a lot. Xifeng Ruan, his intelligent and industrious work also impresses me. Thanks to Zhichao Zhan, Hongbang Liu, Junping

Tian, Hongwei Liu, Ying Yang, Liwen Yao for setting up a great working climate in the office, for sharing the experience and the interest in the work and in everyday life.

I would like to express my deep gratitude to my parents. They always stand in my back to support me. I could not achieve without their understanding and support.

Finally, I wish to thank all those people I have mentioned or may forget to mention. Thank you all who have contributed to this achievement.