

UNIVERSITÀ DEGLI STUDI DI MILANO

FACOLTÀ DI SCIENZE E TECNOLOGIE

Corso di Laurea Triennale in Fisica

TESI DI LAUREA TRIENNALE

**Measurement of the jets faking photons background
for the Dark Matter search in the Mono-Photon
channel with the ATLAS detector**

Codice PACS: 14.80.-j

Relatore interno:

Prof. Leonardo Carminati

Tesi di laurea di:

Gianluca Fugante

Matricola: 867571

Correlatori:

Dott.sa Silvia Resconi

Dott. Davide Pietro Mungo

Anno Accademico 2018/2019

A chi ha creduto in me

Contents

Introduction	1
1 Dark Matter	3
1.1 Evidences of the existence of Dark Matter	4
1.1.1 The rotation curves of galaxies	4
1.1.2 Gravitational lensing	4
1.1.3 The Cosmic Microwave Background	5
1.2 Dark Matter candidates	5
1.3 Detection strategies	6
1.3.1 Direct searches	6
1.3.2 Indirect searches	7
1.3.3 Searches at colliders	7
2 LHC and ATLAS	9
2.1 The Large Hadron Collider	9
2.1.1 The accelerator structure	9
2.1.2 The CERN accelerator complex	10
2.2 The ATLAS experiment	11
2.2.1 Inner Detector	12
2.2.2 Calorimeters	13
2.2.3 Muon Spectrometer	15
2.2.4 Trigger System	15
3 Photons and E_T^{miss} reconstruction in the ATLAS detector	17
3.1 Photons	17
3.1.1 Energy cluster in the EM calorimeter	17
3.1.2 Reconstruction	18
3.1.3 Identification	19
3.1.4 Isolation	21
3.2 Missing transverse momentum	22
4 The Mono-Photon analysis	23
4.1 Events selection	23
4.1.1 Pre-selection	23
4.1.2 Selection in the Signal Region	24
4.2 Background estimation	25
4.2.1 Control Regions	25
4.2.2 Simultaneous fitting technique	26
4.3 Results	27

5 Jets faking photons	31
5.1 Basic method	31
5.1.1 Correlation in the background	33
5.1.2 Signal leakage	33
5.1.3 Signal yields and purity	34
5.2 Validation test	35
5.3 Coefficients from MC	36
5.3.1 Signal leakage coefficients	36
5.3.2 Correlation factor	38
5.3.3 R prime	39
5.3.4 Analysis	40
5.4 Results	40
Conclusions	43

Introduction

Despite the astonishing success of the Standard Model of particle physics (SM) in describing high energy physics results, culminated with the Higgs Boson discovery in 2012, it is widely accepted that it is not a complete theory. For example there are several cosmological and astrophysical observations indicating that SM can account for only $\sim 5\%$ of the mass and energy in the universe. In particular $\sim 27\%$ of the Universe mass-energy is of completely unknown nature: this is known as Dark Matter (DM).

One of the most interesting DM candidate is a weakly interacting massive particle (WIMP). If DM interacts with SM particles, WIMPs can be produced in proton-proton collisions at colliders like LHC and, as they don't interact with the detector, they can be measured only as missing transverse momentum (E_T^{miss}). To tag those events a detectable physics object, such as single photon or jet, as well as Z, W or Higgs boson, is required to be produced in association with DM particles. Such searches with the LHC detectors are known as Mono-X searches: first results on Run 1 measurements and preliminary analysis of Run 2 data, taken during 2015 and 2016, have shown no deviations from SM predictions. However, the increased amount of data collected by ATLAS detector in 2015-2018 will permit to enhance their discovery or exclusion potential, thus extending the reach of DM searches.

The *Mono-Photon* analysis fits in the Mono-X framework, searching for events where the final state carries a Mono-Photon signature. The $E_T^{miss} + \gamma$ final state is relatively clean, only few SM processes show such signature in their final state. The dominant backgrounds consist in processes with a Z or W boson produced in association with a photon (mainly $Z(\rightarrow \nu\nu) + \gamma$) and are estimated by normalizing the Monte Carlo predictions for those backgrounds with a simultaneous fitting technique, which is based on control regions enriched by a specific process. Other backgrounds, like $W/Z + \text{jet}$, top and diboson, in which electrons or jets are reconstructed as photons, are estimated with data-driven techniques since MC simulations don't describe accurately fake photons. In particular the jets faking photons background estimation for the ATLAS Mono-Photon search is the subject of the work presented in this thesis.

Photons in ATLAS are reconstructed from clusters in the electromagnetic calorimeter and tracks in the inner detector. Photons candidates are then required to fulfill a set of identification criteria on shower shapes variables in the calorimeter and finally an isolation selection is applied to define the photon sample. The estimation of the jets faking photons background is done using an ABCD method, in which the candidate photons are divided in four regions according to their isolation and identification (tightness), assuming that true photons are fully contained in the tight-isolated region, while the other regions contain only background photons (fake photons), and that the isolation and tightness variables are uncorrelated. The model has to be corrected to take into account signal leakage, from the signal region to each of the other three regions, and the correlation within all the four regions: this is done by means of MC simulations. To account for these effects three signal leakage coefficient (c_1, c_2, c_3) and one correlation factor (R) are computed using different MC samples: $W + \gamma$, $Z + \gamma$, $Z(\nu\nu) +$

γ for signal photons, and $W + \text{jets}$, $Z + \text{jets}$ for background photons. Coefficients of W and Z bosons are pretty different, especially for $c1$ (signal leakage from the tight-isolated region to the tight-not isolated region): this difference is not fully understood and has been included as a systematic uncertainty. To evaluate the compatibility between simulations and real data an R' (R prime) coefficient has been computed, which is simply a correlation factor in a completely not-isolated region where the signal contamination is expected to be negligible: results for simulations and real data are compatible within one sigma, we can then assume that also MCs predictions in the signal region can be trusted.

With this method it has been possible to estimate purities for 2015-2016 data in each region of the analysis: a purity going from 83% to 100% in the various regions has been measured, with a statistical uncertainty going from 0.3% to 14% and a systematic error between 0.6% and 29%, mainly due to systematics on the tightness. Results are in good agreement with those of the published analysis on the same dataset. The new analysis on Run 2 dataset is still ongoing with a factor of 4 more data that will produce much more precise evaluation of the number of signal Mono-Photon events. The methodologies developed in this work are ready to be applied to the entire Run 2 dataset as soon as the analysis will be unblinded. Results are expected to be published in January 2020.

Chapter 1

Dark Matter

The Standard Model of particle physics (SM) is currently the best model to describe how the fundamental constituents of the Universe interact at a microscopic scale. The SM classifies all known elementary particles, shown in Fig. 1.1, and describes three of the four known fundamental forces: the electromagnetic, weak, and strong interactions, but it does not include the gravitational force. According to this theory the building blocks of all the known matter are fermions, $\frac{1}{2}$ -spin particles divided in quarks and leptons. The left-handed components of both leptons and quarks are grouped in three doublets, while the right handed components are singlets. Every fermion has its own antiparticle, i.e. a particle with the same mass but opposite quantum numbers. The interactions between these elementary particles are mediated by integer- spin particles known as *gauge bosons*: the strong interaction is mediated by *gluons*, the electromagnetic interaction by *photons*, while the weak interaction is mediated by W^+ , W^- and Z^0 bosons. The last particle described by the SM is the Higgs boson, discovered at LHC in 2012 [1] [2], which was introduced to explain the mass of the massive gauge bosons and other elementary particles.

Despite the impressive agreement between experimental results and SM prediction, the SM doesn't account for the gravitational force, moreover some experimental results, such as neutrino oscillations, have come to challenge the SM theory. The most important deviation from this model is certainly Dark Matter (DM): cosmological observations show that SM can account for only 5% of the energy and mass in the Universe, the remaining part is made of Dark Matter (27%) and Dark Energy (68%).

Standard Model of Elementary Particles				
three generations of matter (fermions)		interactions / force carriers (bosons)		
QUARKS	I	II	III	
mass charge spin	$=2.2 \text{ MeV}/c^2$ $\frac{2}{3}$ $\frac{1}{2}$ up	$=1.28 \text{ GeV}/c^2$ $\frac{2}{3}$ $\frac{1}{2}$ charm	$=173.1 \text{ GeV}/c^2$ $\frac{2}{3}$ $\frac{1}{2}$ top	0 0 1 gluon
	$=4.7 \text{ MeV}/c^2$ $\frac{1}{3}$ $\frac{1}{2}$ down	$=98 \text{ MeV}/c^2$ $\frac{1}{3}$ $\frac{1}{2}$ strange	$=148 \text{ GeV}/c^2$ $\frac{1}{3}$ $\frac{1}{2}$ bottom	0 0 1 photon
LEPTONS	$=0.511 \text{ MeV}/c^2$ -1 $\frac{1}{2}$ electron	$=105.66 \text{ MeV}/c^2$ -1 $\frac{1}{2}$ muon	$=1.7768 \text{ GeV}/c^2$ -1 $\frac{1}{2}$ tau	$=91.19 \text{ GeV}/c^2$ 0 1 Z boson
	$=2.2 \text{ eV}/c^2$ $\frac{1}{2}$ $\frac{1}{2}$ electron neutrino	$<1.7 \text{ MeV}/c^2$ 0 $\frac{1}{2}$ muon neutrino	$=15.5 \text{ MeV}/c^2$ 0 $\frac{1}{2}$ tau neutrino	$=80.39 \text{ GeV}/c^2$ ± 1 1 W boson
SCALAR BOSONS				
GAUGE BOSONS VECTOR BOSONS				

Figure 1.1: Table of particles included in the Standard Model.

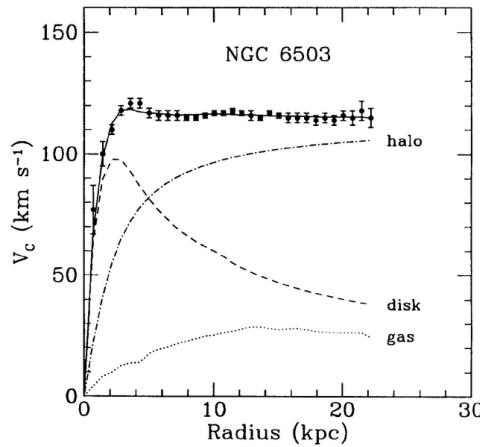


Figure 1.2: Rotation curve of galaxy NGC 6503 [3]. It is evident that the halo contribution is necessary to justify experimental observations.

It is thus natural to wonder if the SM can be considered as the low-energy limit of a more fundamental theory, which would extend the theory to include a description of gravity and of the nature of Dark Matter. Dark Matter searches are of crucial importance to test the extensions of the SM contemplating particles which could constitute Dark Matter and to understand the nature of Dark Matter itself.

1.1 Evidences of the existence of Dark Matter

As discussed previously, the existence of Dark Matter is indicated by several cosmological and astrophysical observations. In the following sections three of the main experimental evidences will be presented.

1.1.1 The rotation curves of galaxies

The rotation curves of galaxies - i.e. the circular velocity of the stars and gas in a galaxy, as function of their distance from the galactic center - provide the first and most convincing evidence of the existence of Dark Matter. It is possible in fact to infer the mass distribution of a galaxy from its rotation curve. Classical mechanics predicts a decrease of circular velocity with the increase of the distance R from the center of galaxies, following the power law $v \sim \frac{1}{\sqrt{R}}$, known as *Keplerian decrease*. Experimental data, however, show a flat dependence of the rotation velocity up to three times the galaxy's optical radius. This discrepancy between theory and observations can be explained by assuming that a hidden halo (the Dark Matter halo) is present in the outer region of galaxy. The presence of the hidden mass of the halo can explain the observed flat rotation curves, as it is clearly visible from Fig. 1.2. Historically this is the most relevant topic in convincing the scientific community of the existence of Dark Matter.

1.1.2 Gravitational lensing

The theory of General Relativity states that a massive body bends space-time around it. Photons don't have mass and can't interact with other particles or bodies by the gravitational force, but if an intense gravitational field is present the geometry of space-time is modified,

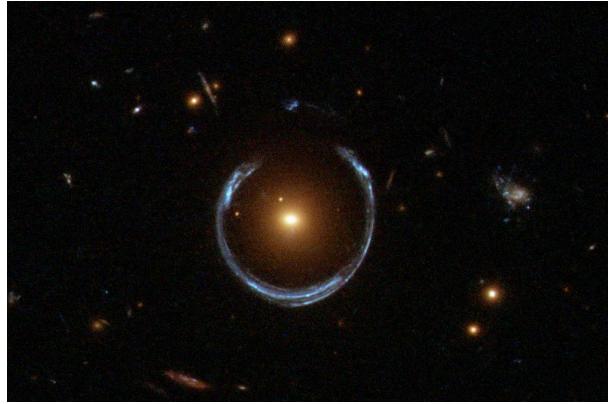


Figure 1.3: Example of gravitational lensing. The source of light is a blue galaxy in the background, while the object deviating the light is the yellow galaxy (LRG 3-757). The effect of gravitational lensing is the blue ring around the yellow light spot. Image Credit: ESA/Hubble and NASA.

bending photons trajectories even if they don't have mass. So if a large mass distribution is present between a source of light and the observer, the light signal looks distorted: this phenomenon is known as *gravitational lensing*. An example of this effect is shown in Fig. 1.3.

This is not just a spectacular effect, but can also indirectly detect Dark Matter. The curvature of light can indeed be used to infer the mass of the object interposed between the source of light and the observer, allowing to measure discrepancies between the value of the mass computed taking into account the effect of gravitational lensing and the mass derived from the visible matter of the object. This once again leads to the hypothesis of the presence of Dark Matter.

1.1.3 The Cosmic Microwave Background

The last important evidence of DM existence comes from the *Cosmic Microwave Background* (CMB), whose study can also indicate the total amount of Dark Matter present in the Universe.

The CMB, represented in Fig. 1.4, is an electromagnetic radiation which fills the whole Universe with a black-body spectrum at a temperature of 2.73 K. It is supposed to be originated during the *epoch of recombination*, approximately 379000 years after the Big Bang, when the Universe became transparent to photons (before this epoch, matter and light were strongly coupled and photons were not free to propagate). The CMB is known to be isotropic, but from an accurate analysis of the CMB anisotropies it is possible to evaluate the composition of today's Universe. The latest and most precise measurement of the CMB was carried out by the PLANCK experiment (ESA): PLANCK's measurements [6] allowed to determine that today's Universe is made up by 5% of ordinary (fermionic) matter, 27% of Dark Matter and 68% of Dark Energy.

1.2 Dark Matter candidates

In this section the most important DM candidates will be discussed. However it is necessary to point out that the particle physics approach is not the only possibility. An alternative approach is to modify Newtonian dynamics and the laws of General Relativity to explain

observations without calling into question the existence of hidden massive particles [7]. These theories are still controversial as they don't really explain some important observations [8].

Despite having collected compelling evidence of the existence of Dark Matter, its nature is yet to be understood. The particle physics approach is based on the assumption that Dark Matter is made up by particles belonging to the family of WIMPs (*Weakly Interacting Massive Particles*), that are stable particles with mass lying between 10 GeV and 1 TeV. This hypothesis is supported by cosmological theories: DM has to be made of particles with a lifetime comparable to the age of the Universe and interacts with SM particles with a cross section around the weak scale. Alternative proposals, as WISPs (*Weakly Interacting Sub-eV Particles*), are summarized in [4], [5]. The WIMPs hypothesis has been for decades the most credited explanation of the nature of Dark Matter for two main reasons:

- The abundance of Dark Matter is consistent with a Dark Matter particle of mass in the WIMP mass range. These particles are expected to interact with ordinary matter by means of the weak force, as neutrinos do. They would also have the same abundance of DM, compatible with the weak cross section. This is known as the *WIMP miracle*, which assumes that Dark Matter was created thermally in the early stages of the Universe and can now only annihilate;
- WIMPs are predicted by a number of consistent theories extending the SM, such as *Supersymmetry* and *Extra Dimensions* and other BSM theories.

These considerations make WIMPs great candidates for DM particles and, if this is true, their observation would be possible in three ways: direct searches, indirect searches and searches at colliders.

1.3 Detection strategies

1.3.1 Direct searches

Direct searches of WIMPs is based on the assumption that a constant flux of them may pass through the Earth and weakly interact with ordinary matter. It would be thus possible to measure the recoil energy of nucleons weakly interacting with these particles. The experimental setup should then include a huge detector which has to be optimized to minimize the background contamination: the most important backgrounds usually come from cosmic rays, natural radioactivity and the intrinsic radioactivity of the detector. The WIMP interaction with the detector's nuclei can have different features: it could be either elastic or inelastic,

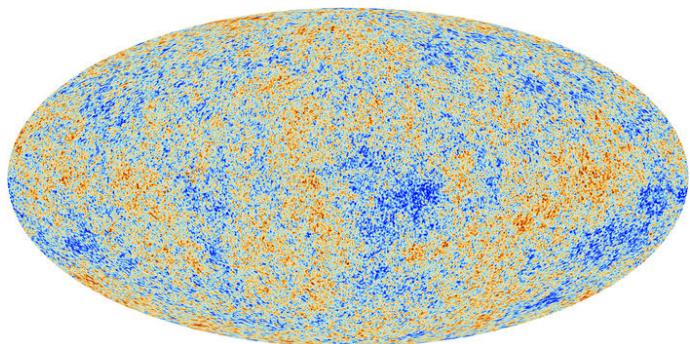


Figure 1.4: Cosmic Microwave Background measured by the PLANCK experiment [6].

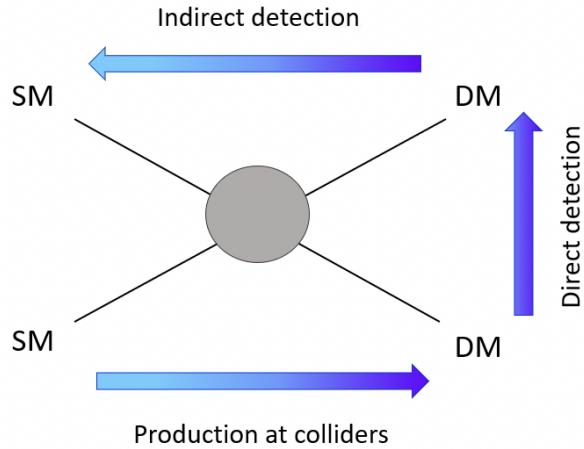


Figure 1.5: Diagram representing the DM detection strategies.

spin-dependent or spin-independent. Up to now, experiments of direct searches just came to set upper and lower limits of the WIMPs mass, without revealing their nature.

1.3.2 Indirect searches

The second method of detection is the indirect method: the idea is to detect the electromagnetic radiation or SM particles produced in the annihilation of WIMPs far in the Universe, that occurs with a high probability in regions of space with high density of DM particles, like galaxies centers. Antiparticles, neutrinos and gamma rays are the most useful annihilation products to be detected. For this reason the indirect strategy focuses on finding an excess in the flux of the annihilation products. Identifying the possible source of the flux and observe whether it could be a region containing a large density of Dark Matter or not helps to discriminate between background and DM decay products.

1.3.3 Searches at colliders

Assuming that WIMPs interact with SM particles they can also be produced in pp collisions in colliders like LHC. Moreover, as said before, WIMPs have to be stable particles and, as they only interact weakly with ordinary matter, they escape the detector. However, the presence of a WIMP would leave a particular signature inside the detector: the total momentum in the transverse plane (i.e. the vectorial sum of the transverse momenta p_T of the particles produced in a pp collision) is expected to amount to zero, but if a particle escapes the detector, it leaves a transverse momentum imbalance from which is possible to infer the production of a WIMP.

Searching for DM in colliders, two approaches are possible: the first one consists in looking for all decay channels expected in a determined theoretical framework, as Supersymmetry, which incorporates Dark Matter candidates, the second one consists in searching for a particular Dark Matter signature. Considering a DM particle manifesting itself as a missing transverse momentum and produced in association with a detectable SM object: the two decay products would recoil against each other, leaving in the detector an event with a large missing transverse momentum and only one SM particle. The presence of a SM particle is necessary to tag the event. The SM object, which could be a jet, a photon, a vector boson (W or Z) or a Higgs boson, defines the name of these signatures, known as *mono-X signatures*.

[9] [10] [11] [12]. These searches look for an excess of events with the mono-X signature with respect to SM processes with the same signature.

Chapter 2

LHC and ATLAS

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [13] is the largest and most powerful particle accelerator in the world. It is located in an underground circular tunnel at CERN, Geneva, where it can run both proton-proton (pp) and heavy ions collisions. Since it was built it reached the maximum center-of-mass energy of $\sqrt{s} = 13$ TeV, which record an instantaneous luminosity of $2 \cdot 10^{34}$ $\text{cm}^{-2} \text{s}^{-1}$ for pp collisions. The data used in this thesis have been collected by the ATLAS detector at LHC between 2015 and 2018, with an integrated luminosity of $\mathcal{L} = 140 \text{ fb}^{-1}$.

2.1.1 The accelerator structure

LHC is a 27 km long circular accelerator, placed about 100 m underground. It is not perfectly round, but consists of eight 2.45 km long arcs and eight 545 m long straight sections.

Like any other particle accelerator, the LHC has three main components: the beam pipes, the accelerating structures, and the magnet system. Inside its two beam pipes, each 6.3 cm in diameter, proton or heavy ion beams travel in opposite directions in an ultra-high vacuum of 10^{-13} bar. This low pressure is necessary to minimize the number of collisions with residual gas molecules and the subsequent loss of the accelerated particles. The second part of an accelerator consists of its accelerating structures. Before protons are introduced into the two LHC beam pipes, they are accelerated in smaller accelerators to about 6% of their final energy. Inside the LHC, the particles acquire their final energy from eight accelerating cavities. Beams enter the LHC at 99.9997828% of the speed of light and after acceleration they reach 99.999991%, rising the protons energy from 938 MeV to 7 TeV. To give them a circular trajectory, the pipes are surrounded by a large magnet system that deflects the protons' path, the larger the energy of a particle becomes, the stronger the magnets need to be, and this is the most challenging limitation for every particle accelerator. The beam momentum can indeed be expressed as:

$$p = 0.3RB \quad (2.1)$$

where R is the average radius of the accelerator and B the magnetic field at the center of the cavity.

The most critical part of the magnet system consists of 1232 dipole magnets, each with a length of about 16 m and a weight of 35 t, which create a maximum magnetic field of 8.33 T. Each magnet contains two coils, each surrounding one of the two beam pipes. The current runs through the coils to create two opposite magnetic fields, one in each beam pipe. This is how two particles of the same charge can follow the same track in opposite directions.

In addition to the dipole magnets, there are quadrupole magnets for focusing the beams, and thousands of additional smaller sextupole and octupole magnets for correcting the beam size and position. All magnet coils and the accelerator cavities are built from NbTi, that become superconducting at very low temperatures. To reach their maximum performance, the magnets need to be at 1.9 K.

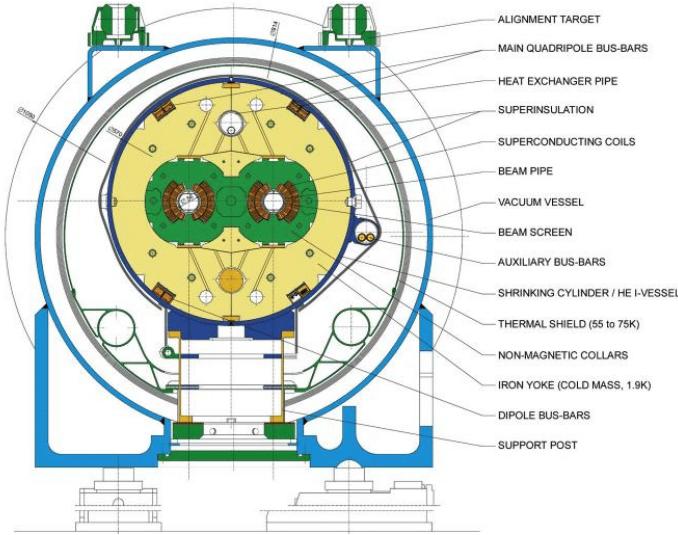


Figure 2.1: Cross section of the LHC dipole magnets.

Inside the cavities the protons are collimated in beams, each composed of more than 2800 bunches of $1.5 \cdot 10^{11}$. Being the protons bunched together, the collisions take place at discrete time intervals, so that the number of events detected in a fixed time interval is always proportional to the luminosity \mathcal{L}^1 , defined as:

$$\mathcal{L} = \frac{n_{beam} N_{bunch}^2 f_{rev} \gamma}{4\pi \epsilon_n \beta^*} \quad (2.2)$$

where n_{beam} is the number of bunches per beam, N_{bunch} is the number of protons per bunch, f_{rev} is the revolution frequency of a bunch around the ring, γ is the relativistic factor, ϵ_n is related to the r.m.s. of the protons' position and momentum distributions along the ring, and β^* reflects the transverse size of the particle beam, which is $16 \mu\text{m}$ at the interaction point.

2.1.2 The CERN accelerator complex

Protons are delivered to the LHC by a chain of pre-accelerators. Free protons are obtained applying an electric field to a source of simple hydrogen, they are then injected in the linear LINAC2 accelerator and accelerated to an energy of 50 MeV first. The Synchrotron Booster bring them to an energy of 1.4 GeV, then the Proton Synchrotron (PS) accelerates the protons to an energy of 25 GeV and the protons finally achieve an energy of 450 GeV in the Super Proton Synchrotron (SPS). At this energy they are injected into the LHC, where the RF (Radiofrequency) cavities accelerate each proton beam to a maximum energy of 6.5 TeV. A scheme of CERN's complete accelerator facility is shown in Fig. 2.2.

¹The instantaneous luminosity is the ratio of the number of events detected at a certain time to the cross section of the interaction involved in the events.

2.2 The ATLAS experiment

ATLAS (A Toroidal LHC ApparatuS) [14] is one of the general purpose detectors at LHC and the largest particle detector ever built for a collider. It is a cylindrical detector located in a cavern at Point 1, it is 44 m long with a diameter of 25 m and weights 7000 tonnes. It has a forward–backward symmetric cylindrical geometry and near 4π coverage in solid angle. ATLAS is composed by three main layers: the **Inner Detector**, the **Calorimeters** and the **Muon Spectrometer**. The three sub-detectors are complementary: the Inner Detector tracks charged particles, the Calorimeters measure the energy of photons and electrons or hadrons, and the muon system provides tracking measurements of highly penetrating muons. A solenoid magnet surrounds the Inner Detector, allowing measurements of the momentum of charged particles, and the toroid magnet plays the same role for the Muon Spectrometer. A representation of the ATLAS setup is shown in Fig. 2.3.

In the detector a cartesian coordinate system is defined: the z axis is aligned to the beam direction with the origin in the nominal collision point, while in the transverse plane the x axis points to the center of LHC and the y axis points upwards. The pseudorapidity, which is relativistic invariant, is defined in terms of the polar angle θ (in spherical coordinates) as:

$$\eta = -\ln \tan \theta/2 \quad (2.3)$$

The angular difference between two points in the detector is then:

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \quad (2.4)$$

Polar coordinates (r, ϕ) are used in the transverse plane, where the vector sum of the momenta of all the produced particles has to be zero, while this is not true for z components because the momenta of the interactions partons is unknown. These quantities are also known as transverse quantities (p_T, E_T^{miss}) and are of fundamental importance for this work.

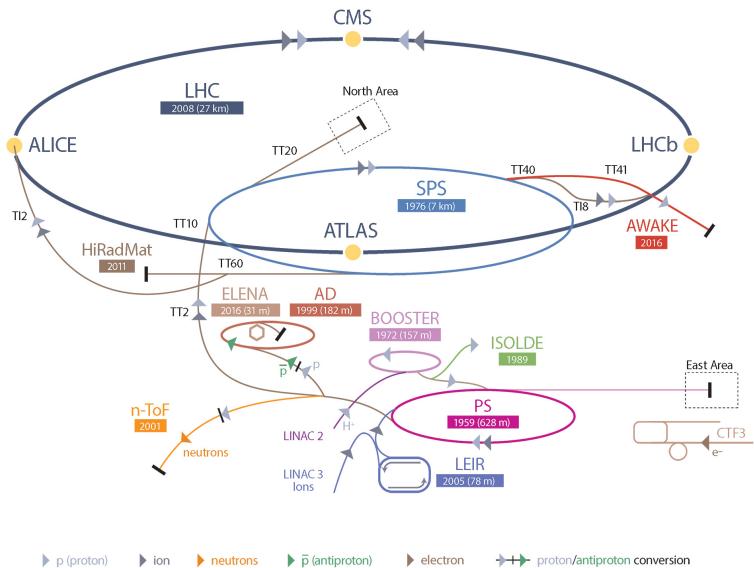


Figure 2.2: CERN’s accelerator complex. In figure are represented the complete pre-accelerators chain and the main experiments.

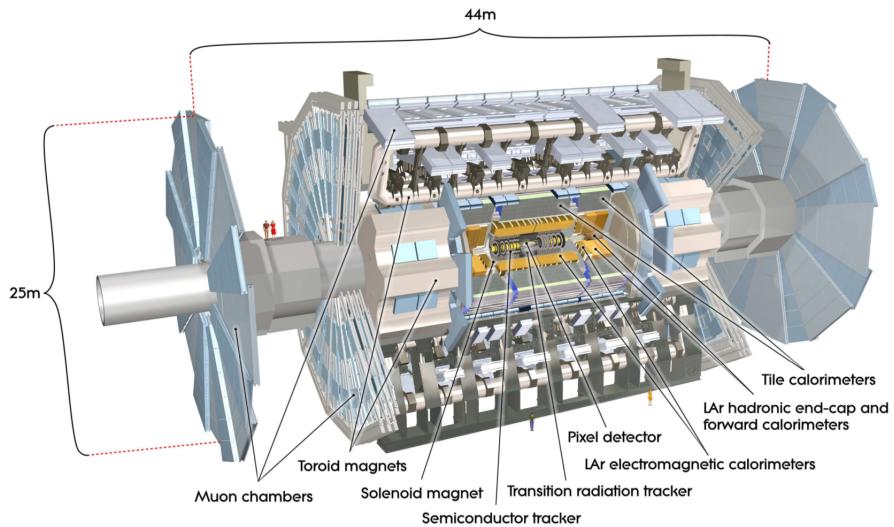


Figure 2.3: Representation of the main components of the ATLAS detector

2.2.1 Inner Detector

The inner-most layer of the ATLAS experiment is the Inner Detector (ID), a 6.2 m long cylindrical detector of radius 2.1 m, surrounded by a 2 T magnetic field generated by the central solenoid. The ID is responsible for tracking charged particles and measuring their energy and momentum, as well as finding the primary and secondary vertex. Because of the large number of particles that emerge from the pp collisions (approximately 1000 particles every 25 ns), this detector needs a fine granularity to achieve high-precision measurements.

A representation of the ID layout is shown in Fig. 2.4. It is composed by a barrel and two end-caps and encloses three independent sub-detectors: the **Pixel Detector**, the **Semiconductor Tracker (SCT)** and the **Transition Radiation Tracker (TRT)**.

The precision tracking detectors (pixels and SCT) cover the region $|\eta| < 2.5$. In the barrel region, they are arranged on four concentric cylinders around the beam axis while in the end-cap regions they are located on disks perpendicular to the beam axis. The highest granularity is achieved around the vertex region using silicon pixel detectors. The pixel layers are segmented in $R - \phi$ and z . All pixel sensors are identical and have a minimum pixel size in $R - \phi \times z$ of $50 \times 400 \mu\text{m}^2$ for an intrinsic single hit resolution of $12 \times 72 \mu\text{m}^2$. For the SCT, four cylinders made of two layers of sensors glued back to back with a 40 mrad stereo angle are used. The mean pitch of the strips is approximately $80 \mu\text{m}$, for an intrinsic hit resolution of $16 \mu\text{m}$ on $560 \mu\text{m}$ the $R - \phi$ out direction.

A large number of hits (typically 36 per track) is provided by the 4 mm diameter straw tubes of the TRT. The TRT only provides $R - \phi$ information, with a single hit resolution of $130 \mu\text{m}$. In the barrel region, the straws are parallel to the beam axis and are 144 cm long, while in the end-cap region, the 37 cm long straws are arranged radially in wheels. The combination of precision trackers at small radii with the TRT at a larger radius gives very robust pattern recognition and high precision. The straw hits at the outer radius contribute significantly to the momentum measurement, since the lower precision per point compared to the silicon is compensated by the large number of measurements and longer measured track length. The space between the straw tubes is filled with polymer fibers to create transition radiation (TR). The TR photons are also measured in the straw, giving a substantial discriminating power between electrons and pions.

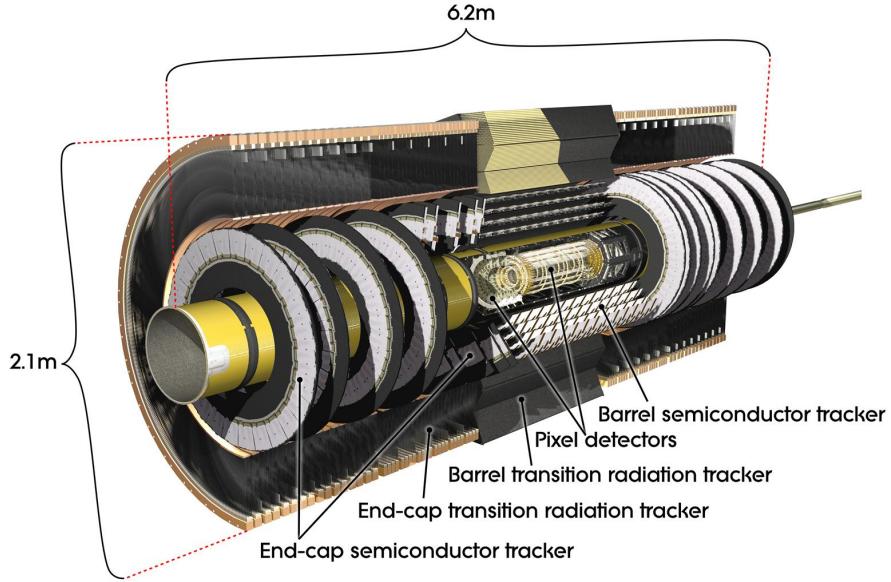


Figure 2.4: Representation of the ATLAS Inner Detector.

2.2.2 Calorimeters

ATLAS's calorimeters measure the energy carried by all particles, both charged and neutral, except for the highly penetrating muons. This energy is converted in detectable signals (scintillation light or electric signals) by means of the materials that constitute the calorimeters. These calorimeters are *sampling calorimeters*, made of alternating layers of *absorbing material* and *active material*. An incident particle interacting with the absorbing material produces a shower of secondary particles with progressively degraded energy, while the active material collects the deposited energy providing a measurable signal. To estimate correctly the energy of a particle a calorimeter needs to fully contain its shower.

The ATLAS's calorimeters system, shown in Fig. 2.5, is composed by the **Electromagnetic Calorimeter (EM)**, the **Hadronic Calorimeter (Had)**, and the **Forward Calorimeter (FCal)**.

The EM calorimeter is divided into a barrel part ($|\eta| < 1.475$) and two end-cap components ($1.375 < |\eta| < 3.2$), while the regions where the barrel meets the end-caps ($1.375 < |\eta| < 1.475$) are called *crack regions* and are not used for measurements. The EM calorimeter is a sampling calorimeter where the active medium is liquid Argon (LAr), alternated with lead plates as absorbing materials. The nominal EM calorimeter resolution, for electrons and photons, is:

$$\frac{\sigma_E}{E} \approx \frac{10\% - 17\%}{\sqrt{E}} \oplus 0.7\% \quad (2.5)$$

As Fig. 2.6 shows, the EM is segmented in 190000 cells along η and ϕ and is composed of four longitudinal layers to record the longitudinal development of the electromagnetic shower, where the first one has the finest granularity in η needed to discriminate prompt photons and two almost collinear photons as decay products of a π^0 .

The Hadronic Calorimeter covers a pseudorapidity of $|\eta| < 3.2$ and is divided in three sub-calorimeters: the Tile Calorimeter in the barrel and the Hadronic End-cap Calorimeter in the two end-caps. They are sampling calorimeters, but the first one uses steel as an absorber and

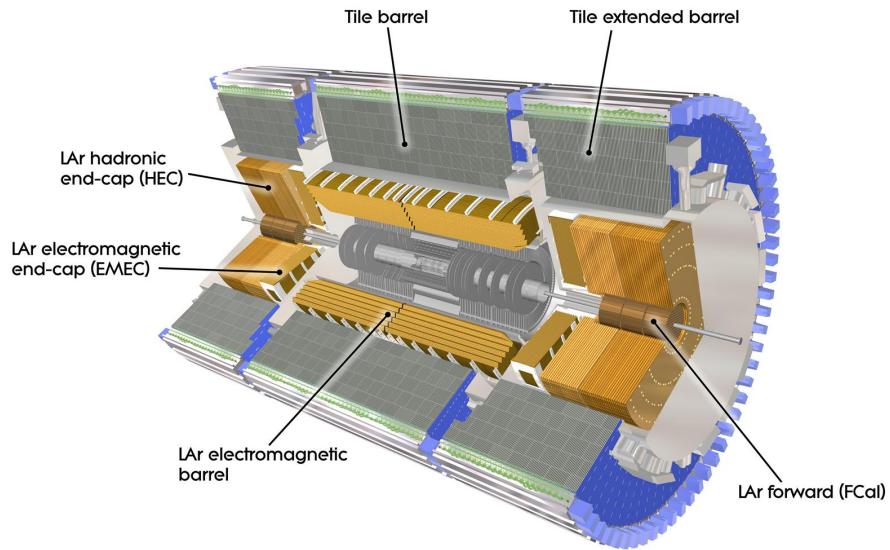


Figure 2.5: Representation of the ATLAS calorimeters system.

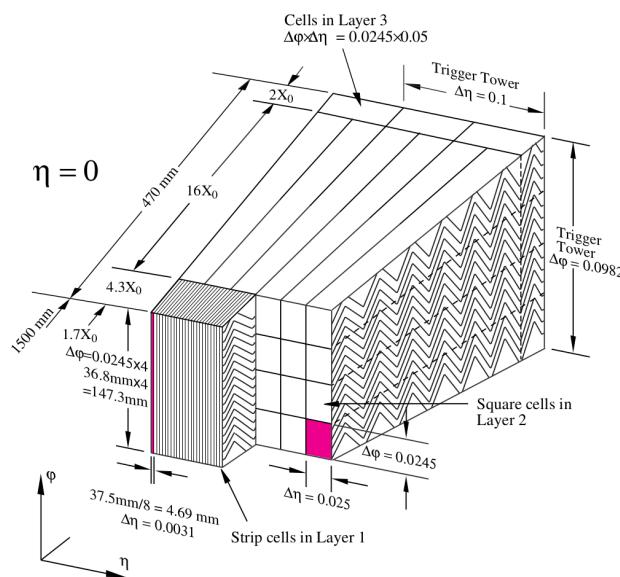


Figure 2.6: Scheme of the Electromagnetic Calorimeter segmentation in η and ϕ

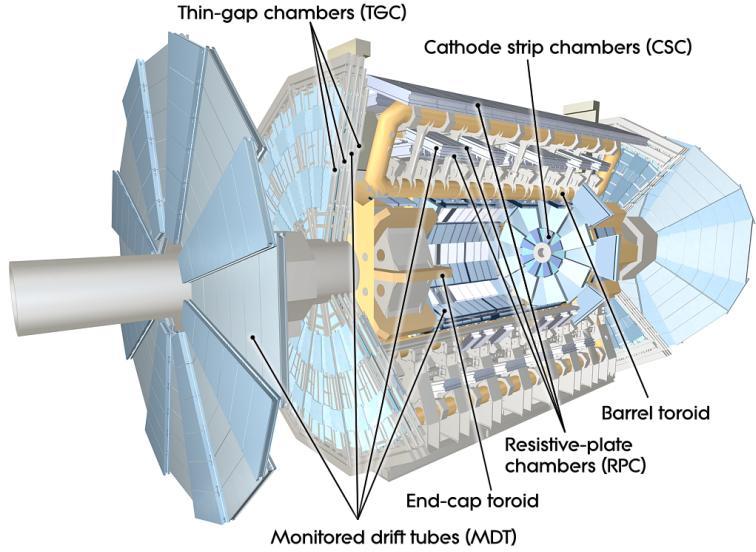


Figure 2.7: Representation of the ATLAS Muon Spectrometer.

scintillating tiles as an active material, while the second ones are liquid-Argon calorimeters, using copper or tungsten as absorbing materials. The nominal energy resolution for hadronic jets is:

$$\frac{\sigma_E}{E} \approx \frac{50\%}{\sqrt{E}} \oplus 3.0\% \quad (2.6)$$

The Forward Calorimeter covers the pseudorapidity range $|\eta| < 4.9$. It consists of three layers in the end-caps, that use liquid Argon as an active material and copper or tungsten as absorbing material. The first module delivers measurements of the energy of electromagnetic showers, while the last two modules mainly record the energy of hadronic showers. In the FCal the nominal energy resolution for hadronic jets is:

$$\frac{\sigma_E}{E} \approx \frac{100\%}{\sqrt{E}} \oplus 1.0\% \quad (2.7)$$

2.2.3 Muon Spectrometer

Muons are highly penetrating particles that can escape the calorimeters, and this is why ATLAS has a dedicated Muon Spectrometer in the outer layer. This detector covers the pseudorapidity range $|\eta| < 2.7$ and a magnetic field, delivered by dedicated toroid magnets, bends muons trajectories, allowing high precision measurements of the muons' momentum. Measurements are provided by muon chambers of two different types: Monitored Drift Tubes (MDTs) in the barrel and Cathode Strip Chambers (CSCs) in the end-cap region, while fast triggering on muon tracks is provided by the Resistive Plate Chambers (RPCs) and the Thin Gap Chambers (TGCs), respectively in the barrel and in the end-caps.

2.2.4 Trigger System

As LHC produces a number of collisions much greater than the storable quantity of data, a system to preselect potentially interesting events is needed: this is the trigger system. The trigger system has two distinct levels: L1 and the high level trigger. Each trigger level refines

the decisions made at the previous level and, where necessary, applies additional selection criteria. The trigger system has to decrease the event rate from the nominal bunch crossing rate of 40 MHz to a rate of about 1 kHz, that is the maximum rate at which data can be processed.

The L1 (Level 1) trigger is an hardware trigger that uses coarse-granularity informations from the muon chambers, calorimeters and forward detector. It searches for high transverse-momentum muons, electrons, photons, jets, and τ leptons decaying into hadrons, as well as large missing and total transverse energy. In each event, the L1 trigger also defines one or more Regions-of-Interest (RoI), regions within the detector where the selection process has identified interesting features. Events passing the L1 trigger selection are analyzed by the next stages of the detector-specific electronics and subsequently passed to the data acquisition. The High Level Trigger (HLT) is a software-based trigger. It uses all the information available within the RoIs, identified by the L1 trigger, exploiting the full detector granularity and offline-like algorithms. The HLT reduces the output rate to approximately 1 kHz with a nominal processing time of about 4 s.

Chapter 3

Photons and E_T^{miss} reconstruction in the ATLAS detector

3.1 Photons

Photons reconstruction, as well as electrons reconstruction, is based on the tracks reconstructed in the Inner Detector and on the clusters of energy deposits in the electromagnetic calorimeter. The reconstruction is designed to separate electrons, unconverted photons and photon conversions to electron-positron pairs, these objects have indeed very similar behaviours inside the detector. For a complete description of the reconstruction process see [16] and [17].

3.1.1 Energy cluster in the EM calorimeter

The photon reconstruction process follows a *topo-cluster* approach. Topo-clusters are formed following signal-significance patterns, in which the variable that determines the seeding and the growth of a topo-cluster is the cell significance, ζ_{cell}^{EM} , which is defined as:

$$\zeta_{cell}^{EM} = \left| \frac{E_{cell}^{EM}}{\sigma_{noise}^{EM}} \right| \quad (3.1)$$

where E_{cell}^{EM} is the absolute cell energy at EM scale and σ_{noise}^{EM} is the expected cell noise. The process begins identifying cells with a $\zeta_{cell}^{EM} \geq 4$, then around these seeds-clusters, neighboring cells with cell $\zeta_{cell}^{EM} \geq 2$ are collected. If two clusters contain the same cell with absolute energy 2σ above the noise, they are merged in a unique one. Finally a set of neighboring cells with $\zeta_{cell}^{EM} \geq 0$ is added to the topo-cluster. Hadronic calorimeter cells, and not only cells from LAr calorimeter, are initially included in the cluster, but, after a selection designed to identify the clusters that are originated from electromagnetic showers, hadronic calorimeter cells are removed from the selected clusters.

A candidate electron seed topo-cluster is required to have a $E_T > 1$ GeV and has to be matched with a track with at least four hits in the silicon tracker. For photons, since no track match is required, seed topo-clusters with transverse energy $E_T > 1.5$ GeV are used.

Once a seed topo-cluster has been found, the algorithm starts looking for satellite clusters among all the remaining topo-clusters with lower E_T in a window of $\Delta\eta \times \Delta\phi = 0.075 \times 0.125$ around the seed cluster barycentre. When the seed cluster and satellite clusters are selected, a *supercluster* is formed combining all the cells associated to each cluster. Finally the superclusters are restricted to 3 (5) cells in the barrel (endcap) along the η direction.

3.1.2 Reconstruction

Once super-clusters are reconstructed, the reconstruction algorithm searches for tracks with a direction compatible with the cluster barycenter in the EM calorimeter. If a matching track is found, the candidate is classified as an electron, otherwise it is classified as an unconverted photon. If there is a conversion vertex (secondary vertex) it can be either a **double-track vertex** or a **single-track vertex**. Double-track vertex candidates are reconstructed from pairs of oppositely charged tracks in the ID that are likely to be electrons or positrons. The requirement to be an electron changes whether the considered track has hits on the silicon detector or not. Tracks originating from a photon conversion are parallel at the place of conversion, therefore track pairs need to fulfill a series of geometric criteria that change whether both tracks, none or only one track have hits in the silicon detector. After this geometric selection the retained tracks are fitted in order to retrieve the final double-track vertex candidate: only fitted vertices satisfying "quality" criteria are retained. The reconstruction efficiency of double-track vertex candidates drops for conversions taking place in the outer layers of the ID: the TRT has a poorer resolution compared to the silicon detector, therefore softer tracks are more easily missed and a pair can be reconstructed as a single track if the two tracks are very close to each other. These two issues are relevant for asymmetric photon conversions, where one of the electrons carries away most of the energy, and for high-energy photons. To face this efficiency drop, it has been proven to be necessary to include the possibility of reconstructing a single-track conversion vertex.

The final step consists in matching a conversion vertex, if found, to energy clusters in the EM calorimeter. The conversion vertex candidate is extrapolated to the second layer of the EM calorimeter, and its extrapolated angular position ($\eta_{vertex}, \phi_{vertex}$) is compared with the barycenter ($\eta_{cluster}, \phi_{cluster}$) of the energy cluster.

- A conversion vertex built from tracks with hits on the silicon detector is considered to be matched to an energy cluster if their distance in both the η and ϕ directions is smaller than 0.05. If the conversion vertex is a single-track vertex, the distance in the ϕ coordinate increases to 0.1 in the direction of the bending of the track;
- A conversion vertex built from tracks with hits only in the TRT is required to have $\Delta\phi < 0.02$ (0.03) in the direction of (opposite to) the bending of the tracks and $\Delta\eta < 0.35$ (0.2) in the barrel (end-caps) of the TRT. $\Delta\phi$ and $\Delta\eta$ are the angular distances between the extrapolated position of the reconstructed tracks to the second layer of the EM calorimeter and the position of the energy cluster.

After matching conversion vertices to energy cluster in the EM calorimeter, the reconstruction of a candidate particle as an electron or as a photon is finally performed requiring the following criteria:

- Unconverted photon: energy cluster with no matching tracks nor conversion vertex;
- Converted photon: cluster with a matching track with less than four hits in the silicon detector. If a track associated with an energy cluster has no hits in the pixel detector and is part of a double-track vertex located in the silicon detector, the candidate particle is reconstructed as a photon;
- Electron: cluster with a matching track with at least two hits in the pixel detector, four hits in the silicon detector and no matching conversion vertices. If the track is part of a conversion vertex, the particle is reconstructed as an electron either if the vertex

is not a double-track vertex located in the silicon detector or if only one of the tracks associated with the vertex has hits in the innermost layer of the pixel detector;

- A candidate particle is reconstructed both as an electron and as a photon if it does not fulfill any of the requirements listed above or if the ratio of the energy of the cluster to the momentum of the matching track is above 10 or if p_T from the track is smaller than 2 GeV.

3.1.3 Identification

At this point a list of possible photon candidates is created. The next step is to discriminate *prompt photons*, i.e. photons not originating from the decays of hadrons, and *background photons*. This process is carried out by means of identification variables [17] based on the transverse and longitudinal profiles of the energy deposits in the EM calorimeter: these profiles are known as *shower shapes*. Distinctive features of prompt photons are a narrower energy deposit in the EM calorimeter and a smaller leakage to the hadronic calorimeter, compared with background photons originating from hadronic jets. In addition, isolated $\pi^0 \rightarrow \gamma\gamma$ decays result into two separate local energy maxima in the first calorimeter's layer, whose fine segmentation is useful to distinguish between these photon pairs originating from π^0 decays and prompt photons. Selection criteria depend on the p_T and η of the candidate photon to account for differences of the shower shapes in different regions of the detector and with the transverse energy.

Type	Name	Description	Loose	Tight
Acceptance	-	$ \eta < 2.37$, with $1.37 < \eta < 1.52$ excluded	✓	✓
Hadronic leakage	R_{had1}	Ratio of E_T in the first layer of the hadronic calorimeter to the E_T of the energy cluster in the EM calorimeter (used for a pseudorapidity range of $ \eta < 0.8 \wedge \eta > 1.37$)	✓	✓
	R_{had}	Ratio of E_T in the hadronic calorimeter to the E_T of the energy cluster in the EM calorimeter (used for a pseudorapidity range of $0.8 < \eta < 1.37$)	✓	✓
L2 of the EM calorimeter	R_η	Ratio of the energy in 3×7 cells to the energy in 7×7	✓	✓
	w_{η^2}	Lateral shower width, $\sqrt{(\sum E_i(i - i_{max})^2) / (\sum E_i)}$, where i runs over all strips in a window of 20×2 strips and i_{max} is the index of the highest-energy strip measured in L1	✓	✓
	R_ϕ	Ratio of the energy in 3×3 cells to the energy in 3×7		✓

L1 of the EM calorimeter	w_{s^3}	Lateral shower width, $\sqrt{(\sum E_i(i - i_{max})^2)/(\sum E_i)}$, where i runs over all strips in a window of 3×2 strips and i_{max} is the index of the highest-energy strip calculated from three strips around the strip with maximum energy deposit		✓
	w_{stot}	Total lateral shower width, $\sqrt{(\sum E_i(i - i_{max})^2)/(\sum E_i)}$, where i runs over all strips in a window of 20×2 strips and i_{max} is the index of the highest-energy strip measured in L1		✓
	f_{side}	Energy outside the core of the three central strips but within seven strips divided by energy within the three central strips		✓
	ΔE_s	Difference between the energy associated with the second maximum in L1 and the energy reconstructed in the strip with the minimum value found between the first and second maxima		✓
	E_{ratio}	Ratio of the energy difference between the maximum energy deposit and the energy deposit in the secondary maximum in the cluster to the sum of these energies		✓
	f_1	Ratio of the energy in L1 to the total energy of the EM cluster		✓

Table 3.1: List and description of the variables used by the identification algorithm. Variables employed to define the Loose and Tight identification working points are highlighted.

Photons candidates can be divided in *tight* and *loose* photons (the two selections are highlighted in Table 3.1): loose photons are those which pass the selections on hadronic leakage and on the middle strips, while tight photons pass all identification criteria. Tight and loose selections will be largely used in the last chapter, in addition three more selections will be used. These selections are built from the tight one releasing different criteria, for each selection the variables that won't be checked are:

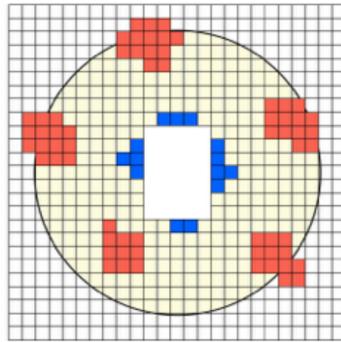


Figure 3.1: TopoEtCone40: isolation cone (yellow), topological clusters (red) and p_T leakage (blue). The central rectangle is removed.

- Tight-3: f_{side} , ΔE , w_{s3} ;
- Tight-4: f_{side} , ΔE , w_{s3} , E_{ratio} ;
- Tight-5: f_{side} , ΔE , w_{s3} , E_{ratio} , w_{stot} .

where each selection excludes tight photons. These variables are chosen in order to minimize the correlation between tightness and isolation.

3.1.4 Isolation

Requiring isolation criteria for the prompt photon candidates allows a further rejection of background photons from hadronic leakage. Photon isolation criteria are based on two different isolation variables: calorimeter isolation and track isolation.

The calorimetric isolation variable E_T^{iso} is defined as the sum of transverse energies in topological clusters in the EM calorimeter and in the hadronic calorimeter within a cone of radius $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.4$ (or 0.3 or 0.2) around the photon candidate. The contribution of the energy of the candidate photon itself is excluded from the sum subtracting a rectangle of 5×7 cells from the center of the cone, as shown in Fig. 3.1. Also corrections for photon energy leakage in the isolation cone are applied. The result is stored in the *TopoEtCone40* (or 30 or 20) variable. The ATLAS collaboration consider isolated (in the calorimeters) those photons which pass the p_T dependent criterium:

$$TopoEtCone40 < 0.022p_T + 2.45 \text{ GeV} \quad (3.2)$$

The track isolation, instead, is defined by means of the *ptcone20*, *ptcone30* and *ptcone40* variables. Inside the Inner Detector the p_T of all tracks greater than 1 GeV are summed up in a cone of radius $\Delta R = 0.2$ (or 0.3 or 0.4) are summed up and stored in the respective variable. The ATLAS collaboration consider isolated (in the ID) those photons which pass the p_T dependent cut:

$$\frac{ptcone20}{p_T} < 0.05 \quad (3.3)$$

A candidate photon is considered isolated if it passes both isolation cuts.

3.2 Missing transverse momentum

The missing transverse momentum (E_T^{miss}) is a fundamental variable exploited to identify final states with particles that only weakly interact with the detector material, such as the neutrinos or other particles predicted by new physics theories, like WIMPs.

A precise measurement of the E_T^{miss} is crucial as it gives important information about particles invisible to the detector, but it is also of difficult reconstruction as it needs all detected particles properly reconstructed and identified. This variable is also very sensible to electronic noise, underlying events, pile-up and cosmic rays background.

To reconstruct the E_T^{miss} an algorithm sums up all contributions from different objects. This summation is done in a specific sequence: electrons, photons, tau, jets and muons. A *Soft Term* is also included to account for energy clusters not topologically connected to any of the reconstructed objects. In particular the Soft Term is one of the most delicate to be computed as it has to account for low p_T objects, electronic noise, pile-up and cosmic rays background.

The x (y) components of the E_T^{miss} are then computed as follow:

$$E_{x(y)}^{miss} = E_{x(y)}^{miss,e} + E_{x(y)}^{miss,\gamma} + E_{x(y)}^{miss,\tau} + E_{x(y)}^{miss,jets} + E_{x(y)}^{miss,mu} + E_{x(y)}^{miss,SoftTerm} \quad (3.4)$$

where $E_{x(y)}^{miss,k}$ is the negative sum of the momenta of all reconstructed objects of type k projected onto the $x(y)$ direction. Finally the E_T^{miss} is:

$$E_T^{miss} = \sqrt{(E_x^{miss})^2 + (E_y^{miss})^2} \quad (3.5)$$

A detailed description of the E_T^{miss} variable can be found in [18] and [19].

In this analysis the E_T^{miss} will be largely used, but some modifications are required. In the next chapter a Signal Region (SR) and four Control Regions (CR) will be defined, but each of these regions needs a specific E_T^{miss} definition. In the SR the current definition will be used, while e.g. in the 2eCR the electron term will be ignored, and similarly for the 1muCR, 2muCR and gammajetCR. The modified E_T^{miss} definition is necessary to its spectrum in the SRs and CRs are comparable.

Chapter 4

The Mono-Photon analysis

Dark Matter searches at colliders are based on one simple idea: if DM particles interact with SM particles they can be produced in pp collisions and, as they don't interact with the detector, they can be measured as missing transverse momentum (E_T^{miss}). The E_T^{miss} , however, needs to be associated with at least one detectable physics object in order to be triggered. Events with E_T^{miss} plus a single object in their final state are of particular importance: searches that investigate these events are known as Mono-X searches, where X stands for the single detectable object. The ATLAS Mono-Photon analysis fits in the Mono-X framework searching for DM in the final state containing an energetic photon and large E_T^{miss} .

The ATLAS' Mono-Photon analysis is characterized by a relatively clean final state which can be mimicked by only a few SM processes. An excess in the selected events with respect to SM predictions carrying the same signature would constitute an evidence of physics beyond the SM. The observation of a significant discrepancy with the background due to SM processes could be an indication of production of $DM + \gamma$ events.

The $E_T^{miss} + \gamma$ signature is predicted by multiple theories of physics beyond the SM (BSM), but the Mono-Photon analysis focuses only on a Simplified DM model [20] [21] with five free parameters, i.e. the mass of the DM particle m_χ , the mass of the mediator through which χ interacts with quarks m_{med} , the width of the mediator Γ_{med} , and the couplings of the mediator to quarks (gq) and to the DM particle ($g\chi$).

In this chapter an overview of the Mono-Photon analysis will be presented, focusing on features interesting for the jets faking photons background estimation. For a complete description see [9].

4.1 Events selection

4.1.1 Pre-selection

The pre-selection criteria are common to many ATLAS physics analysis: they are designed to discriminate interesting events from instrumental and experimental noise. For each event the pre-selection requirements are:

- Data quality: the event must be in the good run list (GRL)²;
- Trigger: the event must have passed the `HLT_g140_loose` trigger requirement, asking for a loose photon with $p_T > 140$ GeV;

²Events are listed in the GRL if they were collected when LHC and relevant ATLAS sub-detectors worked at their best functionality.

- Good vertex: a primary vertex must be reconstructed with at least two associated good-quality tracks with $p_T > 400$ MeV and $|\eta| < 2.5$;
- Jet cleaning: reject events with any *LooseBad* jet, overlapping with neither leptons nor photons, with calibrated $p_T > 20$ GeV. This selection is meant to reject cosmic rays, beam induced background and noise burst in the calorimeters.

4.1.2 Selection in the Signal Region

The Signal Region (SR) is designed to maximise the expected signal significance and is defined as follow:

- $E_T^{miss} > 150$ GeV;
- At least one loose photon with $p_T > 150$ GeV within the range $|\eta| < 2.37$ excluding the calorimeter crack region $1.37 < |\eta| < 1.52$;
- E_T^{miss} significance³ > 8.5 GeV $^{\frac{1}{2}}$;
- The leading photon must be tight, as explained in Sec. 3.1.3;
- The leading photon must be isolated: $\text{TopoEtcone40} < 0.022p_T + 2.45$ GeV and $\text{ptcone20}/p_T < 0.05$, as explained in Sec. 3.1.4;
- The leading photon must not overlap with E_T^{miss} : $\Delta\phi(\gamma, E_T^{miss}) > 0.4$;
- Photon pointing: the z coordinate pointed by the photon with respect to the identified primary vertex must be less than 250 mm;
- Jet veto: events with at most one jet are retained. If there is a jet, it must have $\Delta\phi(jet, E_T^{miss}) > 0.4$;
- Lepton veto: events with any electron or muon are discarded.

It is also possible to define different SRs with increasing E_T^{miss} threshold. In particular, inclusive and exclusive SRs have been defined for different E_T^{miss} bins. The three inclusive SRs (ISRs) are defined in Table 4.1:

ISR1	ISR2	ISR3
$E_T^{miss} > 150$ GeV	$E_T^{miss} > 225$ GeV	$E_T^{miss} > 300$ GeV

Table 4.1: E_T^{miss} thresholds for inclusive SRs.

while the two exclusive SRs (ESRs) are presented in Table 4.2:

³The E_T^{miss} significance, defined as $E_T^{miss}/\sigma(E_T^{miss})$, gives the significance of E_T^{miss} with respect to its resolution ($\sigma(E_T^{miss})$) and is a powerful quantity for the separation of events with true E_T^{miss} from events where the E_T^{miss} is fake.

ESR1	ESR2
$150 \text{ GeV} < E_T^{\text{miss}} < 225 \text{ GeV}$	$225 \text{ GeV} < E_T^{\text{miss}} < 300 \text{ GeV}$

Table 4.2: E_T^{miss} thresholds for exclusive SRs.

4.2 Background estimation

The SM background in the signal region is composed by different sources, the following SM processes may lead to a final state with a high p_T photon and large E_T^{miss} :

- $\gamma + Z(\nu\nu)$: this is the only irreducible SM background;
- $\gamma + W(e\nu)$: the e is not reconstructed or is reconstructed as a γ ;
- $\gamma + W(\mu\nu)$: the μ is not reconstructed;
- $\gamma + W(\tau\nu)$: the τ can either decay leptonically and be missed or reconstructed as a jet;
- $\gamma + Z(l\bar{l})$: both taus, electrons or muons are missed;
- jet + $Z(\nu\nu)$: the jet fakes a γ ;
- jet + $W(e\nu)$: the e or a jet fakes a γ ;
- jet + $W(\mu\nu)$ and jet + $W(\tau\nu)$: the μ or τ is not reconstructed or the τ is reconstructed as a jet and a jet fakes a γ ;
- $t\bar{t}$, single- t and diboson: similar to the $W + \text{jet}$ background;
- QCD and $\gamma + \text{jet}$: a high fake E_T^{miss} is produced by a miscalibration or misreconstruction of a jet;
- $\gamma + t\bar{t}$: leptons from W 's are not reconstructed (minor contribution).

For the dominant background $\gamma + Z(\nu\nu)$ and secondary contributions, like $\gamma + W(l\nu)$, $\gamma + Z(l\bar{l})$ and $\gamma + \text{jet}$, control regions are built reverting one or more cuts used to define the SRs, allowing one of these processes to become dominant. The control regions are fitted simultaneously using rescaled MC predictions match the data. The $W/Z + \text{jet}$, top and diboson contributions are estimated with data driven techniques based on electron or jet photon-fake studies, while beam-induced background is highly suppressed by applying the jet cleaning and the photon pointing criteria.

4.2.1 Control Regions

Four Control Regions (CRs), for each of the SRs, are defined to constrain the normalization of $\gamma + W$, $\gamma + Z$ and $\gamma + \text{jet}$ backgrounds.

- **Single-Muon Control Region** (1muCR): the same selection of the SR is applied, except for the muon veto: it is required that exactly one muon is present in the event. This control region is used to extract the normalization of the $W(l\nu) + \gamma$ background in the SR;

- **Two-Muon (Two-Electron) Control Region** (2muCR and 2eCR): the same selection of the SR is applied, except for the lepton veto: it is required that exactly two muons (electrons) are present in the event and no pre-selected electron (muon). These control regions are used to constrain the normalization of both the $Z(\nu\nu) + \gamma$ and the $Z(ll) + \gamma$ in the SR;
- **Photon-Jet Control Region** (gammajetCR): the same selection of the SR is applied except for a lower E_T^{miss} range: $85 \text{ GeV} < E_T^{miss} < 110 \text{ GeV}$ to enrich this region of $\gamma + \text{jet}$ background. The request $\Delta\phi(\gamma, E_T^{miss}) < 3.0$ is applied to reduce possible signal contamination. This specific CR has been defined in Run-2 to constrain the normalization of $\gamma + \text{jet}$ background. Due to the different E_T^{miss} range, the gammajetCR is only one (and not one for each of the SRs).

To make sure that the E_T^{miss} spectrum in all the SRs is similar to the one in the CRs, electrons and muons are treated as invisible particles in the corresponding CRs, as explained in Sec. 3.2. Moreover the E_T^{miss} significance and photon pointing cuts are removed.

4.2.2 Simultaneous fitting technique

To finally estimate the background a *simultaneous fitting technique* is used. This technique enables the signal and background yields estimation in both CRs and SRs, exploiting the data constrain in all these regions at the same time. Using a simultaneous fit technique allows a combination of multiple CRs, where the predicted event yield in each region R (either the signal region or a control region) is described as a random variable which follows a Poisson distribution:

$$N_R^{obs} \propto Pois \left(N_R^{obs} | \mathcal{L} \times \sigma \times (A \times \epsilon)_R \right. \\ \left. + k_{Z\gamma} \times N_R(Z(\nu\nu) + \gamma) \right. \\ \left. + k_{W\gamma} \times N_R(W(l\nu) + \gamma) \right. \\ \left. + k_{Z\gamma} \times N_R(Z(ll) + \gamma) \right. \\ \left. + k_{\gamma+jet} \times N_R(\gamma + jet) \right. \\ \left. + N_R^{e-to-\gamma} + N_R^{jet-to-\gamma} \right) \quad (4.1)$$

where:

- N_R^{obs} is the signal yield in the region R , \mathcal{L} the luminosity, σ the signal cross section, A the acceptance and ϵ the reconstruction efficiency;
- $N_R(Z(\nu\nu) + \gamma)$, $N_R(W(l\nu) + \gamma)$, $N_R(Z(ll) + \gamma)$ and $N_R(\gamma + jet)$ are the expected real- γ background yields in the region R , provided by MC simulations. Each yield is multiplied by their scale factor k ;
- $N_R^{e-to-\gamma}$ and $N_R^{jet-to-\gamma}$ are the two additional background sources from electrons faking photons and from jets faking photons.

The free parameters of the fit are the *signal strength* μ and the three background normalization parameters $k_{Z\gamma}$, $k_{W\gamma}$ and $k_{\gamma+jet}$.

The MC event yields depend from the nuisance parameters θ_i , corresponding in this case to the uncertainties on predictions on each background yield in every region: the presence of the nuisance parameters allows the expected background yields to fluctuate from the nominal MC prediction, but the gaussian constraints applied to the nuisance parameters disfavour the simultaneous fit to assign to θ_i a value far from the best expected value.

At first stage fit is performed sing only CRs (*background-only fit*), two different fit configurations are used: the background-only inclusive fit, which determines the normalisations for $W + \gamma$, $Z + \gamma$ and $\gamma + \text{jets}$ backgrounds for each inclusive SR independently and the background-only multiple-bin fit, which determines the normalisations for the three exclusive SRs simultaneously. In the first case, four CRs corresponding to a given SR are used to obtain the normalisations, while in the second one, all ten CRs associated with the three exclusive SRs (ESR1, ESR2 and ISR3) are used.

4.3 Results

In this section the main results of the previous Mono-Photon analysis [9], which refers to Run 2 data taken during 2015 and 2016 will be presented.

Table 4.3 shows the fitted values of the normalisation factors for $W + \gamma$, $Z + \gamma$ and $\gamma + \text{jets}$ backgrounds from the background fits:

Signal region	E_T^{miss} [GeV]	$k_{W\gamma}$	$k_{Z\gamma}$	$k_{\gamma+\text{jets}}$	$k'_{W\gamma}$	$k'_{Z\gamma}$	$k'_{\gamma+\text{jets}}$
SRI1	> 150	1.05 ± 0.09	1.10 ± 0.09	1.07 ± 0.25			
SRI2	> 225	1.04 ± 0.11	1.14 ± 0.13	1.06 ± 0.25			
SRI3	> 300	1.04 ± 0.15	1.27 ± 0.23	1.06 ± 0.24	1.03 ± 0.14	1.27 ± 0.23	
SRE1	150–225				1.06 ± 0.10	1.10 ± 0.10	1.07 ± 0.25
SRE2	225–300				1.02 ± 0.12	1.09 ± 0.14	

Table 4.3: Normalisation factors k obtained from a background-only inclusive-SR fit performed in each inclusive SR (the first three columns) and scale factors k' obtained from a background-only multiple-bin fit performed simultaneously in the three regions ESR1, ESR2 and ISR3 (the last three columns), where $k'_{\gamma+\text{jet}}$ applies to all exclusive signal regions. The errors shown include both the statistical and systematic uncertainties.

As can be noticed the normalization factors are well compatible with 1.

Table 4.4 shows the observed number of events and the total SM background prediction after the background-only fit in all SRs, as obtained from the simultaneous fit on data taken during 2015-16 and corresponding to an integrated luminosity of 36.1 fb^{-1} . The event yields observed in data are consistent within uncertainties with the predicted SM background event yields in all inclusive SRs.

Since no deviations from SM predictions have been observed, these results are therefore interpreted in terms of exclusion limits [22] in models that would produce an excess of $E_T^{\text{miss}} + \gamma$ events. Secondly exclusion plots are produced for the simplified models discussed in the introduction of this chapter. Fig. 4.1 shows for example the observed and expected contours corresponding to a 95% CL exclusion as a function of m_{med} and m_χ for simplified DM models with axial-vector or vector mediators and different g coupling values (reported in figure). The region of the plane under the limit curves is excluded and the region not allowed due to perturbative unitarity violation is to the left of the Perturbative limit line.

The analysis was firstly able to put upper limits on the fiducial cross section ($\sigma \times A$): the observed (expected) upper limits on the fiducial cross section for the production of events

	SRI1	SRI2	SRI3	SRE1	SRE2
Observed events	2400	729	236	1671	493
Fitted Background	2600 \pm 160	765 \pm 59	273 \pm 37	1900 \pm 140	501 \pm 44
$Z(\rightarrow vv)\gamma$	1600 \pm 110	543 \pm 54	210 \pm 35	1078 \pm 89	342 \pm 41
$W(\rightarrow \ell\nu)\gamma$	390 \pm 24	109 \pm 9	33 \pm 4	282 \pm 22	75 \pm 8
$Z(\rightarrow \ell\ell)\gamma$	35 \pm 3	7.8 \pm 0.8	2.2 \pm 0.4	27 \pm 3	5.7 \pm 0.7
$\gamma + \text{jets}$	248 \pm 80	22 \pm 7	5.2 \pm 1.0	225 \pm 80	17 \pm 6
Fake photons from electrons	199 \pm 40	47 \pm 11	13 \pm 3	152 \pm 28	34 \pm 8
Fake photons from jets	152 \pm 22	37 \pm 15	9.7 $^{+10}_{-9.7}$	115 \pm 24	27 \pm 9
Observed events in 1muCR	1083	343	116	740	227
Observed events in 2muCR	254	86	27	168	59
Observed events in 2eleCR	181	59	21	122	38
Observed events in PhJetCR	5064	5064	5064	5064	5064

Table 4.4: Observed event yields in 36.1 fb^{-1} of data compared to expected yields from SM backgrounds in all signal regions, as predicted from the simultaneous fit to their respective CRs. The uncertainty includes both the statistical and systematic uncertainties.

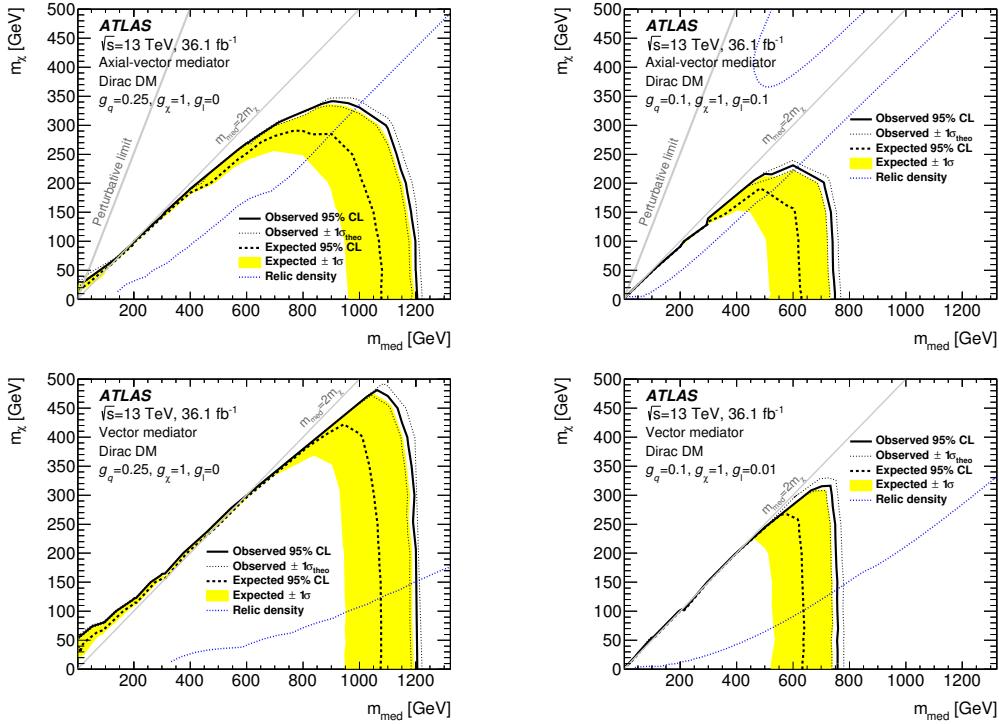


Figure 4.1: The observed and expected 95% CL exclusion contours for a simplified model of DM production involving different mediators and Dirac DM couplings. The region on the left is excluded by the perturbative limit and the relic density curve is also shown: at higher mediator masses, the DM would be overabundant, while at lower values, it would be underabundant.

with a photon and large missing transverse momentum are 7.0 and 2.3 fb (10.6 and 3.0 fb) at 95% CL for E_T^{miss} thresholds of 150 GeV and 300 GeV respectively.

For the simplified DM model considered, the search excludes axial-vector and vector mediators with masses below 750–1200 GeV for χ masses below 230–480 GeV at 95% CL, depending on the couplings chosen.

Chapter 5

Jets faking photons

The jets faking photons background represents approximately the 8% of the total background of the Mono-Photon analysis and it is due to $Z + \text{jet}$ or $W + \text{jet}$ events in which the jet is mistakenly reconstructed and identified as a photon. The classification of the reconstructed object into a photon or a jet is performed by an algorithm based on the detector inputs. This algorithm necessarily introduces an efficiency for true photons and a jet contamination in a selected sample of photons. The simulations do not describe accurately the performance of this algorithm, so the $\text{jet} \rightarrow \gamma$ fraction is estimated with a purely data-driven technique. This is done applying the *Two Dimensional Sideband Method*, which was used in 2015 and 2016 analysis too. In this chapter the method, the validation and the results of the estimation of the jets faking photons background in all the regions of the 2019 ATLAS' Mono-Photon analysis using the full 2015-2018 Run 2 statistics will be presented.

5.1 Basic method

The *Two Dimensional Sideband Method*, also known as the ABCD method, is an almost data-driven technique employed to determine a background contamination in a given signal region. This method relies on counting photon candidates in four regions of a two-dimensional plane defined by an isolation variable and an identification variable (tightness). The identification variable is represented by the value of $\text{TopoEtCone40} - 0.022p_T - 2.45$ GeV, described in Sec. 3.1.4: the candidate photon is considered isolated if it is lower than 0 GeV, not-isolated if it is greater than 3 GeV. Note that the gap of 3 GeV is defined to minimize signal leakage in the CRs. On the other hand the identification variable is divided in two bins corresponding to the Tight and Tight-4 selections described in Sec. 3.1.3. Tight-3 and Tight-5 selections will be used to estimate systematic uncertainties.

The plane is so divided into four regions as shown in Fig. 5.1, where the signal region corresponds to the Tight-Isolated region and the three CRs are assumed to be populated only by background photons. In each region the number of candidate photons is:

- N^A : number of Tight - Isolated candidates;
- N^B : number of Tight - Not-isolated candidates;
- M^A : number of Tight-4 - Isolated candidates;
- M^B : number of Tight-4 - Not-isolated candidates.

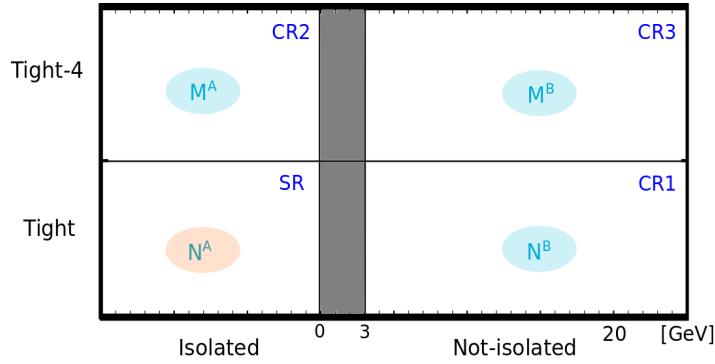


Figure 5.1: Scheme of the regions defined by the Two Dimensional Sideband Method.

For real data the numbers of signal and background photon candidates in the SR are defined as N_{sig}^A and N_{bkg}^A respectively, while for MC simulation as N_{sigMC}^A and N_{bkgMC}^A (and similarly in the CRs).

The method relies on two simplifying hypotheses:

1. the correlation between tightness and isolation is negligible for the background;
2. in the three CRs the number of signal photon candidates is negligible compared to the number of fake candidates:

$$\begin{aligned} N_{bkg}^B &\gg N_{sig}^B \\ M_{bkg}^A &\gg M_{sig}^A \\ M_{bkg}^B &\gg M_{sig}^B \end{aligned} \quad (5.1)$$

The first assumption leads to:

$$\frac{N_{bkg}^A}{N_{bkg}^B} = \frac{M_{bkg}^A}{M_{bkg}^B} \quad (5.2)$$

while the second one implies that:

$$\begin{aligned} N^B &= N_{bkg}^B \\ M^A &= M_{bkg}^A \\ M^B &= M_{bkg}^B \end{aligned} \quad (5.3)$$

Therefore, combining the two hypotheses:

$$N_{bkg}^A = N^B \frac{M^A}{M^B} \quad (5.4)$$

that leads directly to:

$$N_{sig}^A = N^A - N_{bkg}^A = N^A - N^B \frac{M^A}{M^B} \quad (5.5)$$

$$P := \frac{N_{sig}^A}{N^A} = 1 - \frac{N^B}{N^A} \frac{M^A}{M^B} \quad (5.6)$$

where the purity P of the sample is defined as the ratio of the number of signal photons and the total number of candidates in the SR.

A purely data-driven formula for the purity is then obtained even if it requires that both the two hypotheses are satisfied, although this is not always the case.

5.1.1 Correlation in the background

By means of MC simulations it is possible to take into account non-negligible correlations in the background 2D distributions. Assuming that MC and real data are in good agreement for background events, the number of candidates in a given region can be approximated with the corresponding MC prediction. This is done in the following way:

$$\begin{aligned}
N_{sig}^A &= N^A - N_{bkg}^A = N^A - N_{bkg}^A \frac{N_{bkg}^B \times M_{bkg}^A / M_{bkg}^B}{N_{bkg}^B \times M_{bkg}^A / M_{bkg}^B} \\
&= N^A - \left(N_{bkg}^B \frac{M_{bkg}^A}{M_{bkg}^B} \right) \left(\frac{N_{bkg}^A}{N_{bkg}^B} \frac{M_{bkg}^B}{M_{bkg}^A} \right) \\
&\approx N^A - \left(N_{bkg}^B \frac{M_{bkg}^A}{M_{bkg}^B} \right) \left(\frac{N_{bkgMC}^A}{N_{bkgMC}^B} \frac{M_{bkgMC}^B}{M_{bkgMC}^A} \right) \\
&\approx N^A - \left(N^B \frac{M^A}{M^B} \right) \left(\frac{N_{bkgMC}^A}{N_{bkgMC}^B} \frac{M_{bkgMC}^B}{M_{bkgMC}^A} \right) \\
&= N^A - \left(N^B \frac{M^A}{M^B} \right) R_{MC}
\end{aligned} \tag{5.7}$$

that implicitly defines the *correlation factor* R_{MC} which can be estimated by means of pure background MC simulations. Its statistical uncertainty is obtained by the error propagation:

$$\sigma_{R_{MC}} = R_{MC} \sqrt{\frac{1}{N_{bkgMC}^A} + \frac{1}{N_{bkgMC}^B} + \frac{1}{M_{bkgMC}^A} + \frac{1}{M_{bkgMC}^B}} \tag{5.8}$$

which depends only on the available MC statistic.

5.1.2 Signal leakage

MC simulations can also estimate the signal leakage from the SR to each one of the CRs, releasing thus the second hypothesis. As for the correlation factor, real photons are approximated with MC simulations, assuming non negligible signal contamination in the CRs, the number of events in these regions can be written as:

$$\begin{aligned}
N^B &= N_{bkg}^B + N_{sig}^B = N_{bkg}^B + N_{sig}^A \frac{N_{sig}^B}{N_{sig}^A} \\
M^A &= M_{bkg}^A + M_{sig}^A = M_{bkg}^A + N_{sig}^A \frac{M_{sig}^A}{N_{sig}^A} \\
M^B &= M_{bkg}^B + M_{sig}^B = M_{bkg}^B + N_{sig}^A \frac{M_{sig}^B}{N_{sig}^A}
\end{aligned} \tag{5.9}$$

Three *signal leakage coefficients* are then defined as:

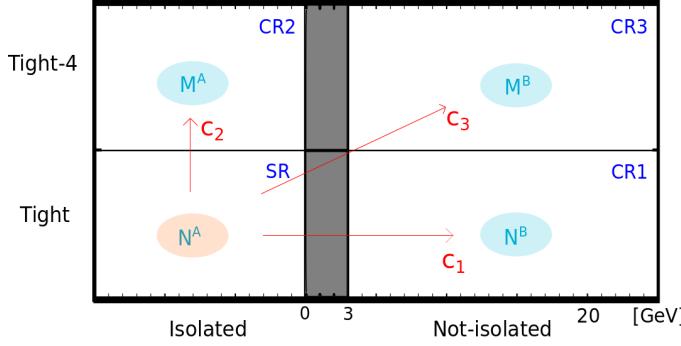


Figure 5.2: Representation of the signal leakage coefficients.

$$\begin{aligned}
 \frac{N_{sig}^B}{N_{sig}^A} &:= c_1 \approx \frac{N_{sigMC}^B}{N_{sigMC}^A} \\
 \frac{M_{sig}^A}{N_{sig}^A} &:= c_2 \approx \frac{M_{sigMC}^A}{N_{sigMC}^A} \\
 \frac{M_{sig}^B}{N_{sig}^A} &:= c_3 \approx \frac{M_{sigMC}^B}{N_{sigMC}^A}
 \end{aligned} \tag{5.10}$$

A schematic representation of the meaning of signal leakage coefficients is reported in Fig. 5.2.

Applying these coefficients to Eq. 5.5 leads to:

$$N_{sig}^A = N^A - N_{bkg}^B \frac{M_{bkg}^A}{M_{bkg}^B} = N^A - (N^B - N_{sig}^A c_1) \frac{M^A - N_{sig}^A c_2}{M^B - N_{sig}^A c_3} \tag{5.11}$$

and resolving for N_{sig}^A gives the number of signal photons in the SR corrected for signal leakage in the CRs:

$$N_{sig}^A = \frac{(M^B + N^A c_3 - N^B c_2 - M^A c_1) \left(-1 + \sqrt{1 + \frac{4(c_1 c_2 - c_3)(N^A M^B - N^B M^A)}{(M^B + N^A c_3 - N^B c_2 - M^A c_1)^2}} \right)}{2(c_1 c_2 - c_3)} \tag{5.12}$$

5.1.3 Signal yields and purity

It is now possible to account for both signal leakage and correlation in the background at the same time. Starting from Eq. 5.7:

$$N_{sig}^A \approx N^A - \left(N_{bkg}^B \frac{M_{bkg}^A}{M_{bkg}^B} \right) \left(\frac{N_{bkgMC}^A}{N_{bkgMC}^B} \frac{M_{bkgMC}^B}{M_{bkgMC}^A} \right) \tag{5.13}$$

replacing N_{bkg}^B with $N^B - N_{sig}^A c_1$ and so on:

$$N_{sig}^A \approx N^A - \left((N^B - N_{sig}^A c_1) \frac{M^A - N_{sig}^A c_2}{M^B - N_{sig}^A c_3} \right) R_{MC} \tag{5.14}$$

some simple algebra gives N_{sig}^A and dividing by N^A one finds a formula for the purity which accounts for both corrections:

$$P = \frac{(M^B + N^A c_3 - N^B c_2 R_{MC} - M^A c_1 R_{MC})}{2N^A(c_1 c_2 R_{MC} - c_3)} \cdot \left(-1 + \sqrt{1 + \frac{4(c_1 c_2 R_{MC} - c_3)(N^A M^B - N^B M^A R_{MC})}{(M^B + N^A c_3 - N^B c_2 R_{MC} - M^A c_1 R_{MC})^2}} \right) \quad (5.15)$$

The derivatives needed for errors propagation are computed with Mathematica [23], a modern technical computing system that enables, among other things, to compute derivatives of an analytical expression of a function.

The systematic uncertainties are estimated varying different assumptions of the method. For example the not-tight control region is moved from Tight-4 to Tight-3 and Tight-5 and two new purities are obtained. The maximum difference between these purities with the nominal one is quoted as systematic uncertainty from the not-tight selection. The same will be done moving the isolation gap from 3 GeV to 2 GeV and 4 GeV, and similarly for the systematic uncertainties on the four coefficients. The total systematic uncertainty is the square sum of all systematic uncertainties.

5.2 Validation test

To evaluate the consistency of the method, it has been tested on a mixed MC sample of $W(\mu\nu) + \gamma$ and $W + \text{jets}$ events with known purity. The samples are all MC generated with Sherpa [24] at NLO⁴.

The correlation factor and the signal leakage coefficients have been computed on the mixed sample matching the photon candidates with true-level⁵ jets (for R_{MC}) and with true-level photons (for c_1, c_2 and c_3). A particular focus is given to the SR - ISR1 as it is the most populated region.

Results in the SR - ISR1 for the coefficients, with statistical uncertainties, are reported in Table 5.1:

R_{MC}	c_1	c_2	c_3
2.76 ± 0.49	$7.08 \pm 0.80 \%$	$4.50 \pm 0.63 \%$	$0.49 \pm 0.21 \%$

Table 5.1: Coefficients used in the validation test.

The results for the purities with propagated and systematic uncertainties respectively are reported in Table 5.2:

Purity	
True	84.27 %
Calculated	$86.41 \pm 5.87 \pm 7.26 \%$

Table 5.2: Table of purities calculated to validate the ABCD method.

⁴Events are typically weighted by the process cross section. MC samples at NLO (Next to Leading Order) are characterized by weights assigned to each event to properly reproduce the normalization and kinematics of the process.

⁵MC simulations are obtained in three steps: at first the particles are generated, then they are simulated to interact with the detector and in the end they are reconstructed with the same procedure used for real data. In this process the generation stage is also known as *true-level*.

As can be noticed the calculated purities are compatible within the errors with the expected results.

5.3 Coefficients from MC

Once validated the method on a simplified MC sample it is possible to start analyzing different MCs (all NLO Sherpa samples) to get the coefficients needed by the method and assess the systematic uncertainties.

At first very strange results have been observed due to large negative weights, in few events especially in the CRs of the Mono-Photon analysis. These anomalous weights led to very high unnatural correlation factors.

This difficulty has been managed in the following way:

- weights with an absolute value > 100 are rescaled to 1;
- events with $p_T < 140$ GeV at truth-level are excluded: the analysis selects candidate photons with $p_T > 150$ GeV, but some events can be reconstructed with higher p_T , with respect to true-level p_T , possibly gaining a very high cross section and therefore a high weight;
- regions where a certain sample ($W/Z + \gamma/jets$) is dominant have been merged to increase statistics:
 - $W + \gamma/jets \rightarrow SR + 1muCR$;
 - $Z + \gamma/jets \rightarrow SR + 2muCR + 2eCR$;
 - the gammajetCR has been treated separately as it has a different E_T^{miss} cut.
- the track isolation ($ptcone20/p_T$) has been released in the CRs of the method:
 - not-isolated events are now those which fail the calorimetric isolation OR the track isolation;
 - a gap of 0.05 on the track isolation variable has been excluded to prevent signal leakage.

Upper limits on track and calorimetric isolation have been set, respectively to 1 and 140 GeV, to exclude pathological events.

5.3.1 Signal leakage coefficients

Signal leakage coefficients have been calculated using three signal samples of $Z(l\bar{l}) + \gamma$, $Z(\nu\nu) + \gamma$ and $W + \gamma$. Results are reported in Fig. 5.3, 5.4 and 5.5.

A systematic difference between c_1 coefficients calculated from W and Z bosons samples is evident in the first plot. Fig. 5.6 shows the normalized calorimetric isolation profiles in the ISR1, where a difference in the tails of the distributions is clearly visible, that results in a systematic difference in the coefficients that accounts for signal leakage from isolated to not-isolated regions. Since the origin of this effect is not fully understood a systematic uncertainty is assigned to c_1 coefficients to cover the differences observed between different samples. Also c_3 is thus affected by this systematic, even if the results are compatible.

It can be noticed also a systematic increase of c_1 and c_3 with the increase of the E_T^{miss} threshold. This behaviour is once again explained by the calorimetric isolation profiles, in

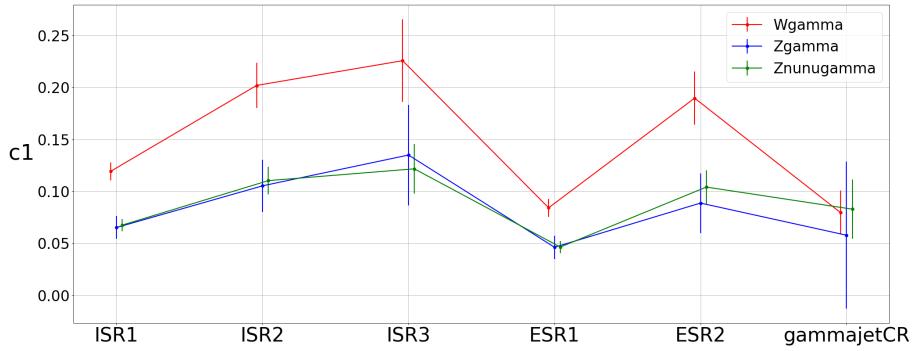


Figure 5.3: Representation of the signal leakage coefficients c_1 , accounting for signal leakage in the tight - not-isolated region.

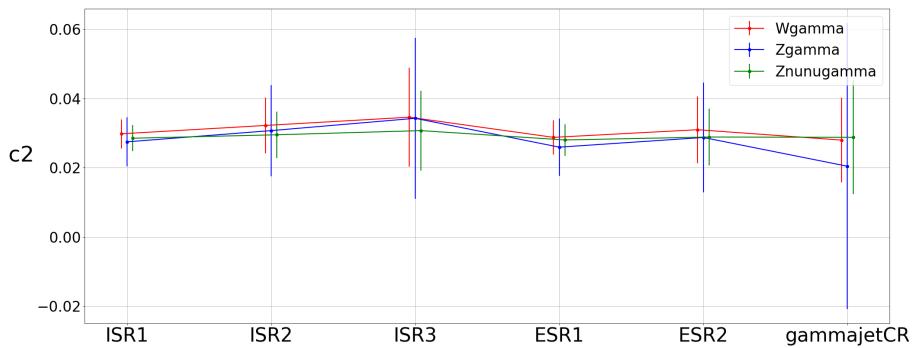


Figure 5.4: Representation of the signal leakage coefficients c_2 , accounting for signal leakage in the not-tight - isolated region.

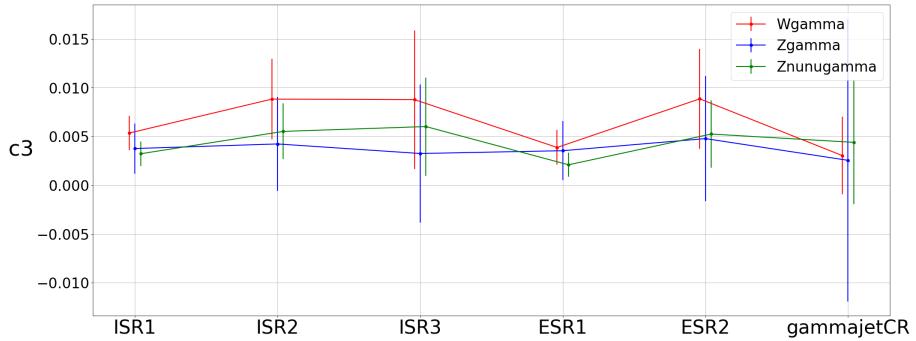


Figure 5.5: Representation of the signal leakage coefficients c_3 , accounting for signal leakage in the not-tight - not-isolated region.

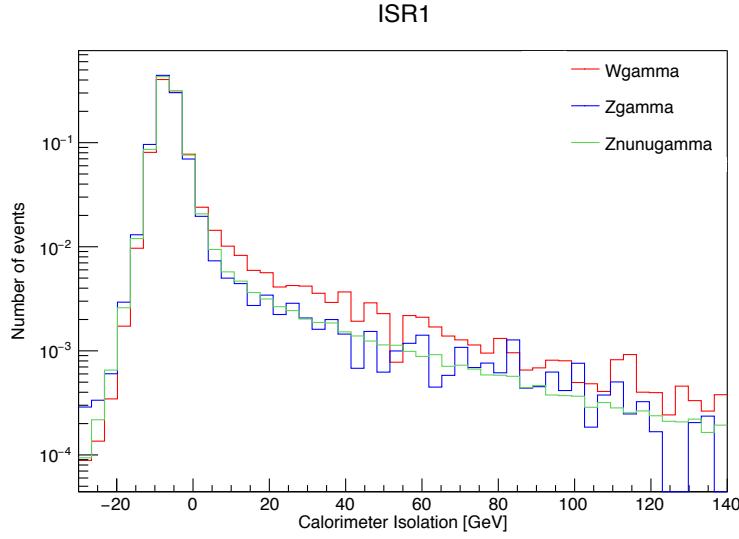


Figure 5.6: Normalized calorimetric isolation profiles in the ISR1 of the three MC samples used to compute the signal leakage coefficients.

particular in the tails. Fig. 5.7 shows the calorimetric isolation profiles of the $Z(\nu\nu) + \gamma$ sample in the inclusive signal regions: as can be seen the tails move upward with the increase of the E_T^{miss} threshold, resulting in increasing coefficients values.

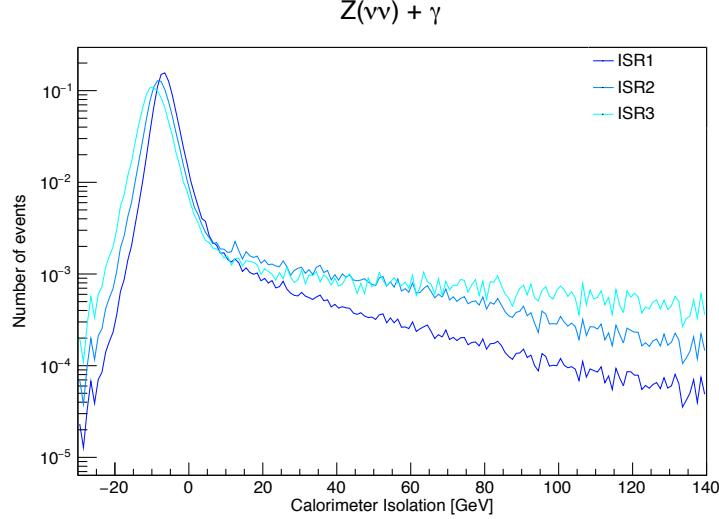


Figure 5.7: Normalized calorimetric isolation profiles, in logarithmic scale, of the $Z(\nu\nu) + \gamma$ sample in the ISR1, ISR2 and ISR3.

5.3.2 Correlation factor

The correlation factors have been calculated separately on two background samples of $Z +$ jets and $W +$ jets. Results are reported in Fig. 5.8.

As can be seen in the plot, the results are compatible within the uncertainties and don't show any evident systematics.

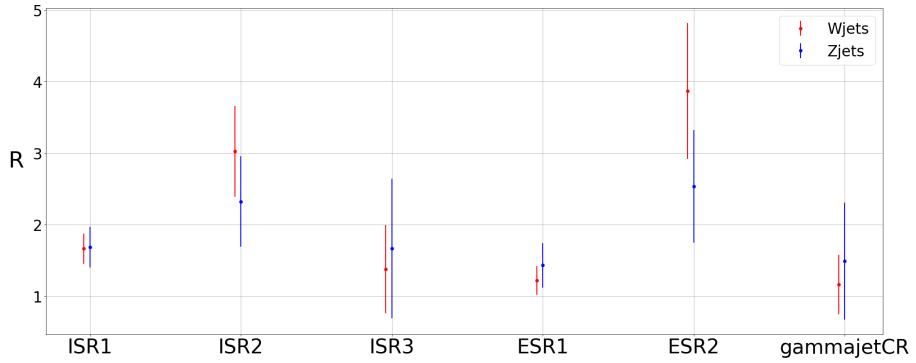


Figure 5.8: Representation of the results obtained for the correlation factor using two samples of $Z + \text{jets}$ and $W + \text{jets}$.

5.3.3 R prime

To evaluate the accuracy of the MC description of the correlation factor a special correlation factor (R_{prime}) in a completely not-isolated region has been computed: in this way it is possible to calculate it for both real data and MC without signal contamination problems. The MC used are the same as for the usual correlation factor, while data are the 2015-16 data. The populations in the not-isolated regions are divided by rectangular selections as Fig. 5.9 shows. The cuts on the track isolation and calorimetric isolation are chosen in order to maintain a good population in all these regions and are set to 0.2 and 50 GeV respectively, so that:

- N_{prime}^A : events with $(3 \text{ GeV} < \text{calo isolation} < 50 \text{ GeV} \text{ and } \text{track isolation} < 0.2)$ or $(\text{calo isolation} < 3 \text{ GeV} \text{ and } 0.1 < \text{track isolation} < 0.2)$
- N_{prime}^B : events with $(50 \text{ GeV} < \text{calo isolation} < 100 \text{ GeV} \text{ and } \text{track isolation} < 1.0)$ or $(\text{calo isolation} < 50 \text{ GeV} \text{ and } 0.2 < \text{track isolation} < 1.0)$

and similarly for M_{prime}^A and M_{prime}^B

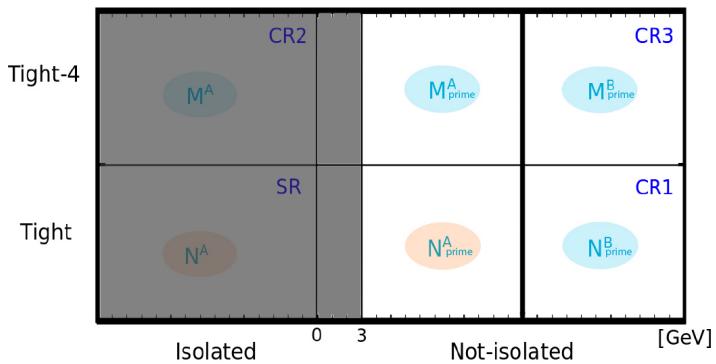


Figure 5.9: Scheme of the new regions defined to compute R_{prime} .

Results are reported in Fig. 5.10. The MC predictions for R_{prime} and are compatible within uncertainties with those extracted from 2015-16 data, except in the gammajetCR where the statistics is very poor. Both data and MCs exhibits a R_{prime} correlation factor close to 1.

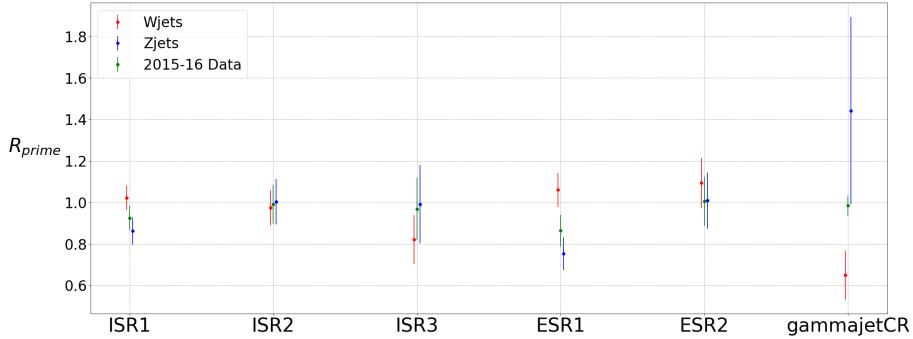


Figure 5.10: Representation of the results obtained for R_{prime} using two samples of $Z + jets$ and $W + jets$.

5.3.4 Analysis

For all the coefficients (c_1 , c_2 , c_3 and R_{MC}) a weighted mean over different MC samples is computed with its statistical uncertainties for the five SRs and the gammajetCR. Systematic uncertainties on c_1 and c_3 has been assigned as the RMS of the values from different MC. These uncertainties are propagated to the purity as the difference between the values obtained using c_x varied up and down by the size of its uncertainty.

The 5+1 coefficients are applied to the 20+1 region of the analysis: each value is used in the corresponding SR and in all the Mono-Photon CRs (for example the results in the ISR1 will be used in the SR - ISR1, 1muCR - ISR1, 2muCR - ISR1 and 2eCR - ISR1) except for the coefficients in the gammajetCR that are treated only in this particular CR.

5.4 Results

It is now possible to compute the purities on 2015-16 data and compare them with the previous analysis results.

Statistical uncertainties are computed with the propagation of errors only from the real data populations, while the errors propagated from the coefficients are quoted as systematic errors as they are determined by the limited statistics of the MCs.

Systematic uncertainties on the purities, on the other hand, are obtained as discussed in Sec. 5.1.3 and 5.3.4. Having released the track isolation in CRs of the method, the gap between the isolated and not-isolated regions needs to be redefined. The now rectangular gap is varied respectively on the track isolation and calorimetric isolation as $(0.05, 3 \text{ GeV}) \rightarrow (0.01, 2 \text{ GeV})$ or $(0.05, 3 \text{ GeV}) \rightarrow (0.10, 4 \text{ GeV})$. The different systematic errors are finally summed in quadrature with the statistical errors propagated from the coefficients, giving the final systematic error on the purities. Finally the total error is the sum in quadrature of the total systematic and statistical errors.

Fig. 5.11 reports the results of fake photons numbers, both for the previous analysis and the published analysis [9] on 2015-16 data, calculated as:

$$N_{FakePhotons} = (1 - P)N^A \quad (5.16)$$

where the errors are obtained simply by multiplying the error on P by N^A . The two analysis results are compatible within uncertainties in each region of the analysis, although in the new analysis the systematic uncertainties are larger.

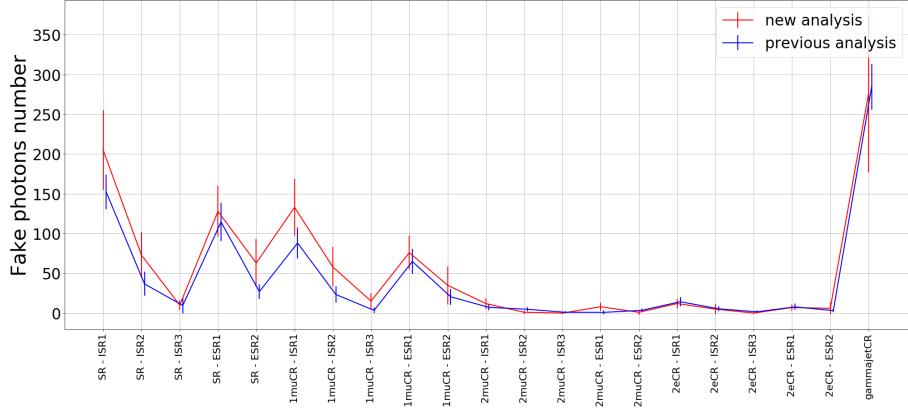


Figure 5.11: Comparison between results of the previous analysis (blue) and of the new analysis (red) on data taken during 2015-16.

Table 5.3 shows the details of the results obtained by the ongoing analysis on 2015-16 data. The results for the fake photons numbers in each region are dominated by the systematic error coming from the tightness control regions and by the statistical uncertainty on R_{MC} , while the uncertainties on the other coefficients don't really affect the total error.

regions	mean	stat.	tightness syst.	isolation syst.	c1 stat.	c2 stat.	c3 stat.	R stat.	c1 syst.	c3 syst.	total syst.	total error
SR - ISR1	205	21	36	3	3	8	0	22	13	0	45	50
SR - ISR2	73	15	19	2	2	6	0	13	6	0	25	29
SR - ISR3	10	4	2	0	1	2	0	4	1	0	5	6
SR - ESR1	128	15	20	5	2	6	0	17	7	0	29	32
SR - ESR2	63	16	20	3	2	6	0	14	5	0	26	30
1muCR - ISR1	133	15	28	5	1	3	0	14	4	0	32	36
1muCR - ISR2	58	12	19	3	1	2	0	10	2	0	22	25
1muCR - ISR3	15	5	6	1	0	1	0	6	0	0	8	10
1muCR - ESR1	76	10	14	2	1	2	0	10	2	0	18	21
1muCR - ESR2	35	9	21	1	1	2	0	7	2	0	22	24
2muCR - ISR1	12	4	5	0	0	1	0	1	1	0	6	7
2muCR - ISR2	1	2	1	0	0	1	0	0	0	0	1	2
2muCR - ISR3	-0	0	0	0	0	0	0	0	0	0	0	1
2muCR - ESR1	8	3	3	1	0	0	0	1	1	0	4	5
2muCR - ESR2	1	3	1	0	0	1	0	0	0	0	1	3
2eCR - ISR1	12	5	3	1	0	0	0	1	1	0	4	6
2eCR - ISR2	5	4	4	0	0	0	0	1	0	0	4	6
2eCR - ISR3	0	1	0	0	0	0	0	0	0	0	0	1
2eCR - ESR1	7	3	3	0	0	0	0	1	1	0	3	4
2eCR - ESR2	6	5	5	1	0	0	0	1	0	0	5	7
gammajetCR	276	18	26	20	8	32	1	85	5	0	97	99

Table 5.3: Results for fake photon numbers of the current analysis of 2015-16 data with statistical (stat.) and systematic (syst.) errors.

Conclusions

In this thesis work an estimation of the background yields from jets faking photons for the Mono-Photon analysis was performed.

The jets faking photons background estimation has been performed using the ABCD method, a data-driven technique that allows to calculate the background contamination in a given signal region. The method needs some MC inputs, which are calculated using simulated samples of W or Z bosons decays associated to a γ or a jet in their final state.

Evaluating the number of fakes photons in every region of the analysis turned out to be complicated by the poor statistics in MC samples in specific regions of the analysis. This problem has been solved by releasing the track isolation in the control regions of the method and merging the populations where a certain sample is dominant. It is important to point out that these two modifications are both necessary to be able to calculate all the coefficients in the kinematic regions and in the gammajetCR. All the coefficients are in good agreement across different samples, with the exception of c_1 coefficients (signal leakage from the tight-isolated to the tight-not-isolated region) that shows a systematic difference between W and Z samples, which has been accounted for as a systematic uncertainty. This behaviour is not completely understood, but comes from the difference of the tails in the calorimetric isolation profiles.

The number of jets misreconstructed as photons has been calculated in every region of the ATLAS Mono-Photon analysis using the unblinded 2015-16 data. A signal purity going from 83% to 100% depending on the analysis region has been measured. For these results the errors vary from 0.6% to 29% and are completely dominated by the systematics coming from the tightness and from the statistical error on the correlation factor. The errors on the other coefficients are much smaller than these two and don't really affect the total error. The results reported here are compatible with the ones published in [9], although with larger uncertainties.

The current analysis is still ongoing and preliminary results of the jets faking photons contributions in the analysis CRs using the full Run 2 dataset have been produced. The publication of results from the whole Mono-Photon analysis are expected in the beginning of 2020.

Bibliography

- [1] The ATLAS collaboration. *Observation of a New Particle in the Search for the Standard Model Higgs Boson with the ATLAS Detector at the LHC*. CERN, Geneva, Phys. Lett. B 716 (2012) 1-29.
- [2] The CMS collaboration. *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*. CERN, Geneva, Phys. Lett. B 716 (2012) 30.
- [3] Doroshkevich, A., Lukash, Vladimir, Mikheeva, Elena, *A solution to the problems of cusps and rotation curves in dark matter halos in the cosmological standard model*. Phys. Usp. 55 3-17 (2012).
- [4] Gianfranco Bertone, Dan Hooper, and Joseph Silk, *Particle Dark Matter: Evidence, Candidates and Constraints*. Phys. Rept. 405:279-390, 2005.
- [5] Andreas Ringwald, *Exploring the role of axions and other WISPs in the dark universe*. 2012.
- [6] Planck collaboration: P. A. R. Ade et al, *Planck 2013 results. I. Overview of products and scientific results*.
- [7] Mordehai Milgrom, *MOND theory*. Canadian Journal of Physics, 2015.
- [8] Douglas Clowe, Marusa Bradac, Anthony H. Gonzalez, Maxim Markevitch, Scott W. Randall, Christine Jones, and Dennis Zaritsky, *A direct empirical proof of the existence of dark matter*. Astrophys. J. 648:L109-L113, 2006.
- [9] The ATLAS collaboration, *Search for dark matter at $\sqrt{s} = 13$ TeV in final states containing an energetic photon and large missing transverse momentum with the ATLAS detector*. CERN, Geneva, Eur. Phys. J. C 77 (2017) 393.
- [10] The ATLAS collaboration, *Search for dark matter in events with a hadronically decaying vector boson and missing transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*. ATLAS, JHEP 10 (2018) 180.
- [11] The ATLAS collaboration, *Search for large missing transverse momentum in association with one top-quark in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*. ATLAS, JHEP 05 (2019) 41.
- [12] The CMS collaboration, *Search for dark matter particles produced in association with a Higgs boson in proton-proton collisions at $\sqrt{s} = 13$ TeV*. CMS, Aug. 2019.
- [13] Fabiola Gianotti, *Collider Physics: LHC*
- [14] The ATLAS collaboration. *The ATLAS Experiment at the CERN Large Hadron Collider*. ATLAS, JINST 3 (2008).

- [15] Stefano Manzoni, *Physics with photons with the ATLAS Run 2 data: calibration and identification, measurement of the Higgs boson mass and search for supersymmetry in di-photon final state*. Milano, 2017.
- [16] *Expected photon performance in the ATLAS experiment*. Tech. Rep. ATL-PHYS-PUB-2011-007, CERN, Geneva, Apr 2011.
- [17] The ATLAS collaboration, *Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run-1 data*. ATLAS, Eur. Phys. J. C 76 (2016) 666.
- [18] The ATLAS Collaboration, *Performance of missing transverse momentum reconstruction for the ATLAS detector in the first proton-proton collisions at $\sqrt{s} = 13$ TeV*. ATLAS, Eur. Phys. J. C 78 (2018) 903.
- [19] M. G. Ratti, *Analysis of Final States with Large Missing Transverse Momentum and a High-Momentum Photon for the Search of Dark Matter with the ATLAS Detector at the LHC*. Milano, 2014
- [20] J. Abdallah et al., *Simplified Models for Dark Matter Searches at the LHC*. Phys. Dark Univ. 9-10 (2015) 8-23.
- [21] D. Abercrombie et al., *Dark Matter Benchmark Models for Early LHC Run-2 Searches: Report of the ATLAS/CMS Dark Matter Forum*. 2015.
- [22] Cowan, G., Cranmer, K., Gross E., Vitells O., *Asymptotic formulae for likelihood-based tests of new physics*. Eur. Phys. J. C71:1554, 2011.
- [23] <https://www.wolfram.com/mathematica/>
- [24] T. Gleisberg, S. Hoeche, F. Krauss, M. Schoenherr, S. Schumann, F. Siegert, J. Winter, *Event generation with SHERPA 1.1*. JHEP 0902:007, 2009.