



CERN-THESIS-2017-307
29/01/2018

A catalogue record is available from Utrecht University Library

ISBN: 978-90-393-6928-9

Copyright © 2018 by D.F.Lodato All rights reserved. No part of this book may be reproduced, stored in a database or retrieval system, or published, in any form or in any way, electronically, mechanically, by print, photoprint, microfilm or any other means without prior written permission of the author.

Typeset using L^AT_EX

Measurement of the prompt photon
production cross section in proton proton
collisions at $\sqrt{s} = 7$ TeV with ALICE

Meting van de productiedoorsnede van
prompte fotonen in botsingen van protonen
met $\sqrt{s} = 7$ TeV met ALICE

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op
gezag van de rector magnificus, prof. dr. G.J. van der Zwaan, ingevolge het
besluit van het college voor promoties in het openbaar te verdedigen op
maandag 29 januari 2018 des middags te 4.15 uur

door

DAVIDE FRANCESCO LODATO

geboren op 22 juli 1984 te Enna, Italië

Promotor: Prof.dr. T.Peitzmann

Copromotor: Dr.ir. M. v. Leeuwen

Vi Veri Veniversum Vivus Vici
The Vision and the Voice,
Aleister Crowley

Contents

1	Heavy Ion Physics	1
1.1	Quantum Chromo Dynamics	2
1.1.1	QCD Phase Diagram	2
1.2	QGP phenomenology and experimental observables	4
1.2.1	Nuclear modification factor	5
1.2.2	Direct photons	6
1.2.3	Azimuthal anisotropy of particle yields	7
1.3	Probing initial conditions with p-Pb collisions	9
2	Photons as a probe of nuclear matter	13
2.1	Photon production in hadron-hadron collisions	14
2.1.1	QCD in the low- x regime, limits of applicability	17
2.2	The Colour Glass Condensate effective theory	18
2.2.1	CGC and observables	21
2.3	Measuring direct photons: isolation techniques	26
3	The A Large Ion Collider Experiment	29
3.1	General overview	30
3.2	The Inner Tracking System	31
3.3	The Time Projection Chamber	32
3.4	The V0 detector	33
3.5	The Electro-Magnetic Calorimeter	35
3.5.1	Selecting rare events with EMCal: triggering	37
3.5.2	Clustering algorithm	38

4	Data analysis	41
4.1	Data sample, event selection	41
4.1.1	Event selection	42
4.1.2	EMCal data quality assurance studies	43
4.2	Charged Particle Veto	45
4.2.1	Multi-contribution clusters	48
4.3	Photon cluster identification	50
4.3.1	Transverse shower shape σ_{long}^2	51
4.3.2	Isolated photon clusters	54
4.4	Contamination estimation for isolated photon clusters	55
4.4.1	The ABCD method	56
5	Monte Carlo studies and corrections	59
5.1	Simulation samples	59
5.2	Validation of the Monte Carlo samples.	61
5.3	Validation of the analysis procedure	61
5.4	Reconstruction efficiency for isolated prompt photons	64
5.5	ABCD method closure test	66
5.5.1	ABCD method on MC background sample	69
5.6	A more realistic scenario:	
	the effect of signal	72
5.6.1	ABCD Method on MC-Mixed Sample:	
	the α factor	72
5.7	Non factorisation in MC and comparison to data	75
5.7.1	σ_{long}^2 distributions	76
5.7.2	Isolation activity distributions	77
5.8	Non-factorisation of σ_{long}^2 and $E_{\text{T}}^{\text{iso}}$	82
5.9	"Correcting" the MC	85
5.9.1	Smearing of σ_{long}^2 distributions	85
5.9.2	Same rejection in MC as in data	88
6	Corrected yield and systematic uncertainties	95
6.1	Corrected yield as a function of the transverse energy	96

6.2	Systematic uncertainties in data:	
	triggering	99
6.2.1	Out-of-time (bunch) triggered events	99
6.2.2	Efficiency and stability of the triggering algorithm . . .	101
6.2.3	In-bunch pile up events	103
6.3	Systematic uncertainties from MC:	
	the choice of the Mixing Ratio	104
6.4	Systematic uncertainties from data and MC	105
6.4.1	Residual hadronic contamination	105
6.4.2	Cluster definition via shower shape selection	109
6.5	Uncertainty from different isolation	
	probability in data and MC	117
6.5.1	Isolation probability and smearing	118
6.5.2	Anti isolation criterion	125
6.5.3	In-cone activity with charged particles	126
6.5.4	Determination of the central value	
	for the corrected yield	134
6.6	Total systematic uncertainty for the measurement	137
7	Results, outlook and conclusions	141
7.1	Integrated luminosity	142
7.2	Isolated photon cross section and	
	comparison with NLO pQCD calculation.	144
7.3	Isolated photons at forward rapidity	145
7.3.1	$2 \rightarrow 2$ process kinematics	148
7.3.2	Forward π^0 and prompt photon production at LHC. . .	150
7.3.3	Inclusive prompt photons vs. Drell-Yan process	151
7.3.4	Forward-mid γ -hadron azimuthal correlations	153
7.4	A F orward C alorimeter: FoCal	157
8	Conclusions and outlook	161
A	I solation activity and underlying event estimators.	165

B Studies on isolation probability in data and simulated sample.	173
Bibliography	189
Samenvatting	195
Acknowledgements	199
Curriculum Vitae	201

Chapter 1

Heavy Ion Physics

In this chapter I will present the theoretical framework in which the research activity presented in this thesis has been carried out.

After an introduction on the fundamental theory of strong interaction, *Quantum Chromo Dynamics* (QCD), and its unique features, the attention will be shifted to one of the open point of this theory: the existence, under certain specific conditions, of a new state of matter where strongly interacting particles, the quarks, are not confined in the form of hadrons but rather single objects, free to move in a medium permeated by the carriers of the strong force, the gluons. This new state of matter is called *Quark Gluon Plasma* (QGP), and it is expected to be produced in extreme conditions of temperature and density.

The way to verify the existence of such a new state of matter and to test its properties is to perform high-energy heavy ion collisions, and to study how the properties of the colliding matter change with respect to the simpler case of pp collisions. To do so, a number of experimental observables have been identified, measured and studied. In the end of the chapter some of these observables will be presented together with the insight that they provide us about the properties of the hadronic matter under such extreme temperature and density conditions.

1.1 Quantum Chromo Dynamics

Quantum Chromo Dynamics (QCD) is the theory of the interaction between quarks and gluons, the elementary constituents of matter within hadrons like protons and neutrons. QCD was formulated after the *Quantum Electrodynamics* (QED)¹ and adopted after evidence for quarks had been found. Both QED and QCD are quantum field theories with gauge interactions, although the features of QCD are quite different from the (more familiar) QED.

Indeed, while the QED coupling constant is a slowly increasing function of the energy scale, the QCD coupling constant α_S is a strongly decreasing function of the energy scale, as shown in Fig. 1.1. The higher the energy scale at which the QCD is studied, the lower the value of its coupling constant. On the contrary, for low values of the momentum transfer in a scattering process, the coupling constant α_S even diverges, indicating that perturbation theory is fundamentally not applicable in this regime. The momentum scale at which the value of α_S diverges is a fundamental scale of QCD and is referred to as Λ_{QCD} .

When performing high-energy nuclear collisions it is possible to concentrate a great amount of energy in a very small portion of space, where the nucleons melt into a soup of quarks and gluons (deconfinement). Macroscopically, this hot and dense medium will be characterised by a high particle density and high temperature; microscopically, the interactions between the quarks and gluons within the medium will be characterised by a low momentum transfer (low Q^2) which makes it impossible to study such collisions within the analytical framework of perturbation theory.

1.1.1 QCD Phase Diagram

Using Lattice QCD (LQCD) it is possible to model an ideal, static version of the hot and dense medium produced in high-energy heavy ion collisions and to predict the behaviour for some of its macroscopic characteristics. Figure

¹QED describes the interaction of electrically charged particles and it is the theory most tested and most precisely (up to few parts per billions) verified

1.1. Quantum Chromo Dynamics

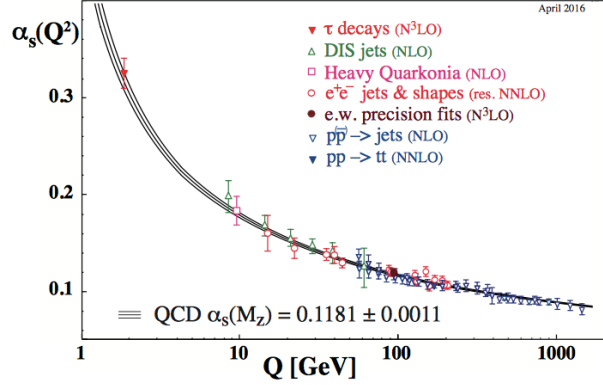


Figure 1.1: The dependence of the coupling α_s on the energy scale of the studied process. Figure taken from [1].

1.2 shows the dependence of the transition temperature of the system on its baryon chemical potential μ_B [2], related to the net surplus density of quarks over antiquarks (possibly bound within hadrons) in the system.

Increasing the temperature of the system leads to a phase transition of first order² from the hadron gas phase, in which quarks and gluons are confined within hadrons, to a new state of matter, the Quark Gluon Plasma. If the baryon chemical potential μ_B is approximately zero, the system will undergo through a phase transition at a certain finite temperature. Calculations have predicted a transition temperature for the system of ~ 170 MeV at $\mu_B = 0$. This value is very close to that of the QCD mass scale parameter Λ .

On the other hand, increasing the baryon chemical potential leads to a colour-superconducting phase, theorised to be the phase for hadronic matter in neutron stars.

Results from more recent calculations on the lattice also predict, for baryon chemical potentials smaller than a certain critical value $\mu_{B,c}$ and using physical quark masses, a rapid crossover instead of a first order transition.

At LHC, studying the properties of high-energy nuclear collisions by looking at the particles produced at mid-rapidity allows us to access the region of

²In Thermodynamics, the order of a phase transition is the order n of the derivative of the free energy $\frac{d^{(n)}F(T)}{dT^{(n)}}$ for which a discontinuity is found. $F(T)$ is the free energy of the system.

$\mu_B \approx 0$, with an estimated initial temperature of ~ 400 MeV [3]. Therefore a hot and dense medium is created in high-energy heavy ion collisions, and it is possible to characterise its properties.

Some of the properties of the QGP, together with the probes designed to test them, will be the subject of the next part of the chapter.

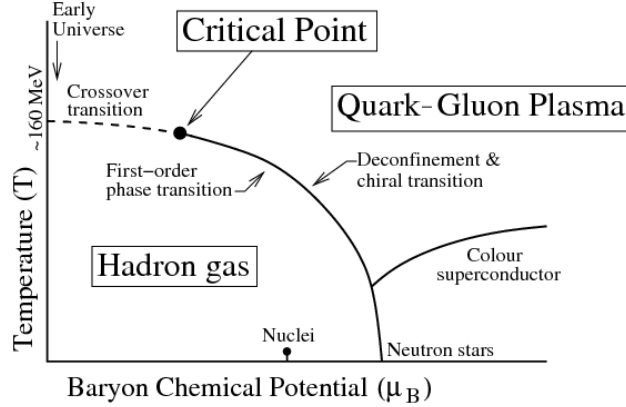


Figure 1.2: QCD Phase diagram. Taken from [4].

1.2 QGP phenomenology and experimental observables

In high-energy nuclear collisions, the energy density is large enough to create a Quark Gluon Plasma. After the collision, the quarks and gluons in the QGP interact among each other until local thermodynamic equilibrium is reached; afterwards, the hot and dense medium expands because of the large pressure gradient from the center of the collision to the surface. During this phase, the temperature of the system decreases significantly, eventually leading to a *freeze-out* of the system into hadrons. The expected lifetime of this new state of the hadronic matter is of the order of 10^{-23} seconds, which makes it impossible to measure its properties directly. Nonetheless, the production of such a hot and dense medium affects some of the experimental observables.

To quantify the changes on the observables due to the production of a hot and dense medium in AA collisions, the same observables are often measured

and studied also in pp (or pA) collisions, where the production of the QGP is not predicted.

In the next sections we will discuss some of the main results from heavy ion collisions obtained by the ALICE collaboration.

1.2.1 Nuclear modification factor

The dominant production mechanism for high- p_T hadrons, in hadron-hadron collisions, is the fragmentation of high- p_T quarks or gluons, called *partons*, produced in hard scatterings in the early stage of the collision. As such partons propagate through the hot and dense QCD medium they can lose energy, leading to changes in the production rate of hadrons at a given transverse momentum. The changes in the particle production rates at high p_T due to the presence of the hot and dense medium are often characterised using the nuclear modification factor

$$R_{AA} = \frac{dN_{AA}/dp_T}{\langle N_{\text{coll}} \rangle dN_{pp}/dp_T} \quad (1.1)$$

which is the ratio of the p_T -differential particle production yields in AA collisions (dN_{AA}/dp_T) to those in pp collisions (dN_{pp}/dp_T), scaled by the number of binary inelastic nucleon-nucleon collisions. The number of binary nucleon-nucleon collision is defined as

$$\langle N_{\text{coll}} \rangle = \frac{T_{AA}}{\sigma_{inel}^{pp}} = \frac{\int T_A(s) T_A(s-b) d^2s}{\sigma_{inel}^{pp}},$$

where T_{AA} is the nuclear overlap function and σ_{inel}^{pp} is the cross section for inelastic pp collisions [5].

Figure 1.3 shows the nuclear modification factor of pions, kaons, protons and unidentified charged hadrons, measured by the ALICE collaboration in Pb-Pb collisions at $\sqrt{s} = 2.76$ TeV. The different panels show results for different collision centralities, denoted percentiles of the total hadronic cross section. As one can see, the suppression of high- p_T charged particle production in AA collision increases with the centrality of the collision. For the most central collisions (0-5 %), the distributions show a characteristic shape with a minimum around 5 GeV/c, and an increasing trend at larger values of the transverse momentum.

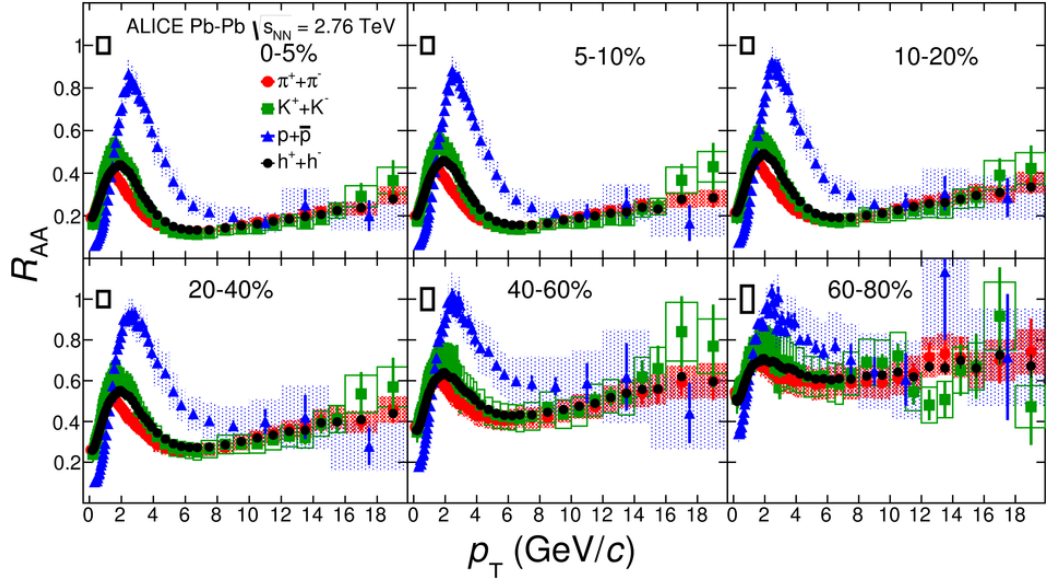


Figure 1.3: Nuclear modification factor R_{AA} of identified charged particles measured by the ALICE experiment in Pb-Pb collisions at $\sqrt{s} = 2.76$ TeV, for different centrality classes of the collision. Figure taken from [6].

1.2.2 Direct photons

One of the best tools to study QGP is the production of photons in high-energy heavy ion collisions. While photon production takes place throughout the whole collision process, it is possible to identify two different stages in which photons are produced, allowing us to extract information on the medium at different times of its evolution.

Photons produced directly in high-energy heavy ion collision can then be categorised as follow:

- "thermal" photons, produced *in a thermal medium* as a result of either quark Bremsstrahlung or emitted from processes like qg -Compton and $q\bar{q}$ annihilation, happening inside the medium. The spectrum of such photons carries information about the temperature, collective flow and space-time evolution of the plasma;
- "prompt" photons, mainly produced in hard-scatterings. These are produced early in the collision and experience the full evolution of the

1.2. QGP phenomenology and experimental observables

QGP. For sufficiently large values of the photon transverse momentum, the prompt photon production rate can be calculated in perturbative QCD, serving as a test for calculations of production rates in nuclear collisions.

Because photons interact only electromagnetically, they can escape the medium without interacting. Photons are also produced via other processes like those involving fragmentation of a quark or the decay of hadrons like π^0 , η and other resonances. Photons from both these processes form a background for the thermal and prompt photons and need to be properly taken into account.

Measurements of the transverse momentum distribution for direct photons performed at LHC within the ALICE experiment are presented in Fig. 1.4 [7]; the spectra can be fitted with an exponential function to obtain an effective slope parameter, which reflects the temperature of the medium. The obtained value is $T^\gamma = 304 \pm 11^{stat} \pm 40^{sys}$ MeV.

1.2.3 Azimuthal anisotropy of particle yields

In non-central heavy-ion collisions the overlap region between the two nuclei is initially asymmetric, almond-shaped, as illustrated by the sketch in Fig. 1.5. The reaction plane is defined as the plane containing both the beam axis and the impact parameter. As an approximation to this, the so-called event plane can be experimentally determined. In a hydrodynamical description of the final state of the collision, like that presented in section 1.2, the pressure gradient in the volume along the short axis (which lies in the reaction plane) is larger than the gradient along the larger axis. This generates an asymmetric expansion, with larger velocities developing along the short axis. As the system expands, the pressure gradient decreases and the initial spatial anisotropy of the particles translates into a momentum anisotropy. By measuring the particle production relative to the event plane, it is possible to determine the anisotropy of the particle production as a function of the transverse momentum.

Figure 1.6 shows the anisotropy, characterised by the second Fourier coefficient v_2 , for charged particles measured by the ALICE experiment in Pb-Pb

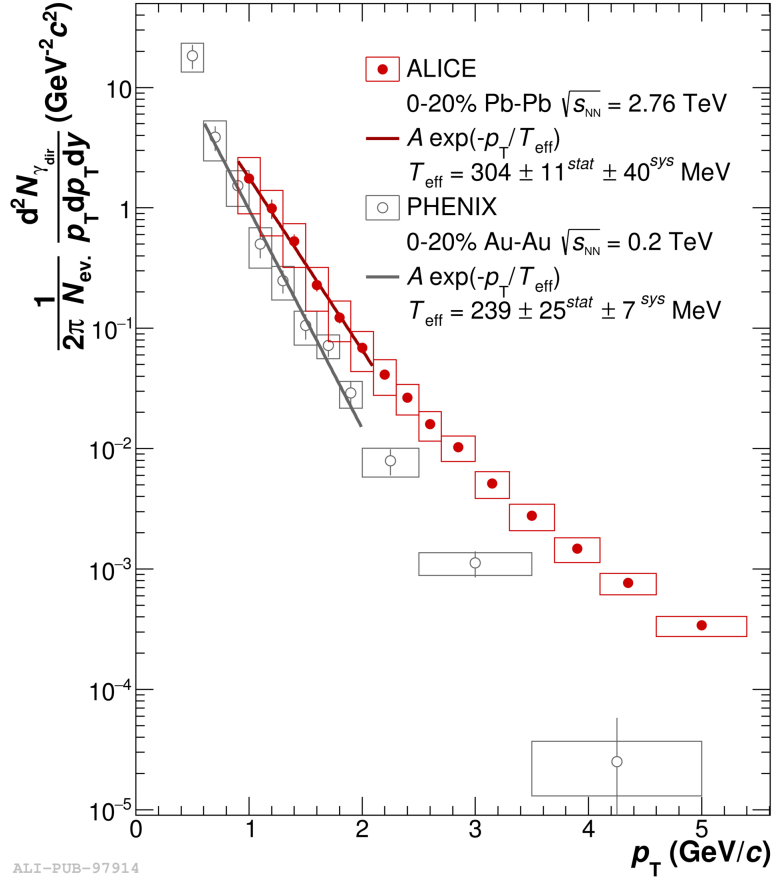


Figure 1.4: Direct photon spectrum, for low values of the reconstructed transverse momentum, as measured by the ALICE experiment in the 0-20 % most central collisions. The thermal contribution is estimated by means of a fit with an exponential function. For comparison, the result from the same measurement performed by the PHENIX collaboration is shown. Figure taken from [7].

collisions at $\sqrt{s_{NN}} = 2.76$ TeV for centrality class 20-30 %. A finite value of v_2 and a characteristic dependence on momentum is visible.

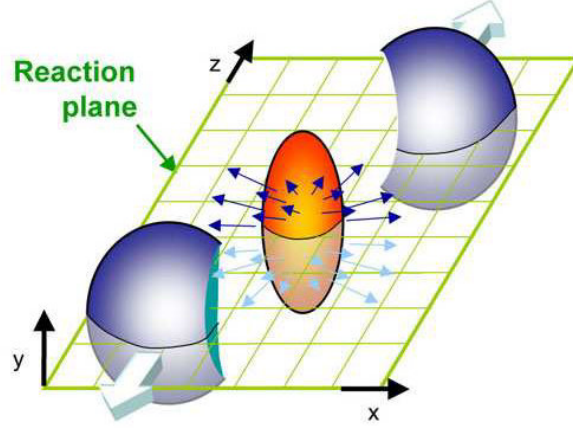
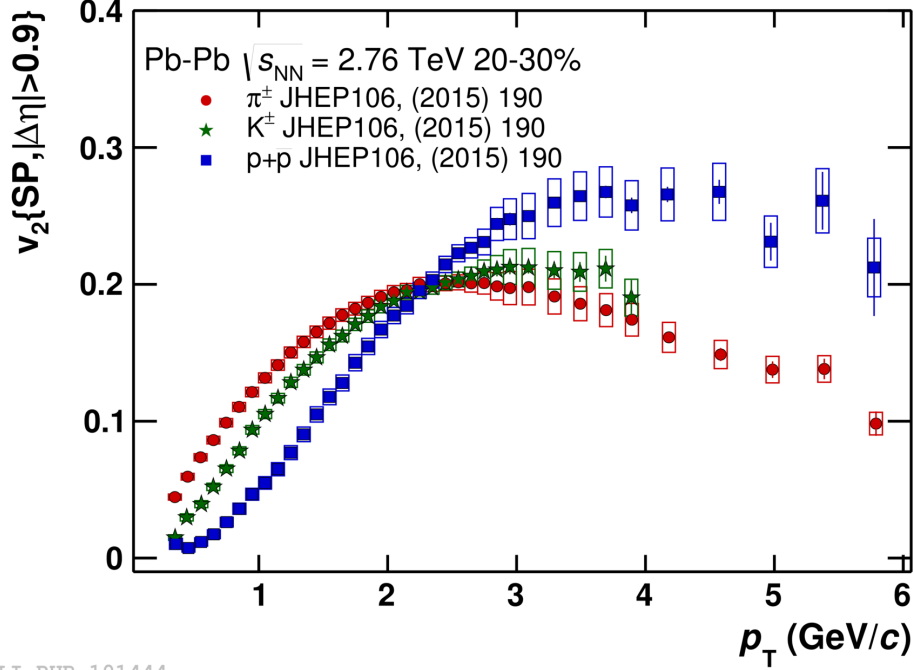


Figure 1.5: Representation of the dynamics for a semi-central nuclear collision and the effect of the pressure gradient on the particle spectra as a function of the transverse momentum.

1.3 Probing initial conditions with p-Pb collisions

When performing high-energy heavy ion collisions we deal with a very complicated system in which we attempt to disentangle the different contributions to the modifications observed in the experimental observables with respect to the case of pp collisions. The theoretical modelling of collisions involving nuclei assumes an incoherent superposition of nucleons in the nuclei with the addition of corrections to account for the presence of a nuclear medium. The scattering amplitude when colliding nuclei is then treated similarly to the pp case. These assumptions have deep implications on the interpretation of the experimental results and need to be carefully studied and verified. In order to verify whether our modelling of collisions between nuclei is correct, it is useful to perform the same measurements explained in the previous paragraph when colliding protons and nuclei.

In this case the (smaller) colliding system is expected to show some changes with respect to the pp case, due to the so-called cold nuclear matter effects. An example of cold nuclear matter effects is the *anti-shadowing* of the parton distribution functions, which means that the probability to find



ALI-PUB-101444

Figure 1.6: The v_2 coefficient as a function of the transverse momentum for some identified particles, measured by the ALICE collaboration in Pb-Pb collisions at $\sqrt{s} = 2.76$ TeV. The three distributions are taken from [8].

a parton in the proton that carries a momentum fraction x is larger when the nucleons are in a nuclear environment than in free nucleons. As a consequence, the probability of finding low- x partons in a nuclear environment is smaller than in free nucleons (*shadowing*).

Figure 1.7 shows the results of some of the studies performed by the ALICE collaboration on p-Pb collisions. On the left we see the nuclear modification factor R_{pA} , whose definition is similar to that given in 1.1, while on the right we present the distribution of charged particles in relative azimuthal angle $\Delta\varphi$ with respect to a trigger particle in the 0-20 most central p-Pb collisions, after subtracting the same quantity measured in 60-100 central collision. These last results have been fitted with a Fourier series and compared to the prediction obtained with the Hijing Monte Carlo generator [9]. A v_2 -like modulation is visible in the data.

1.3. Probing initial conditions with p-Pb collisions

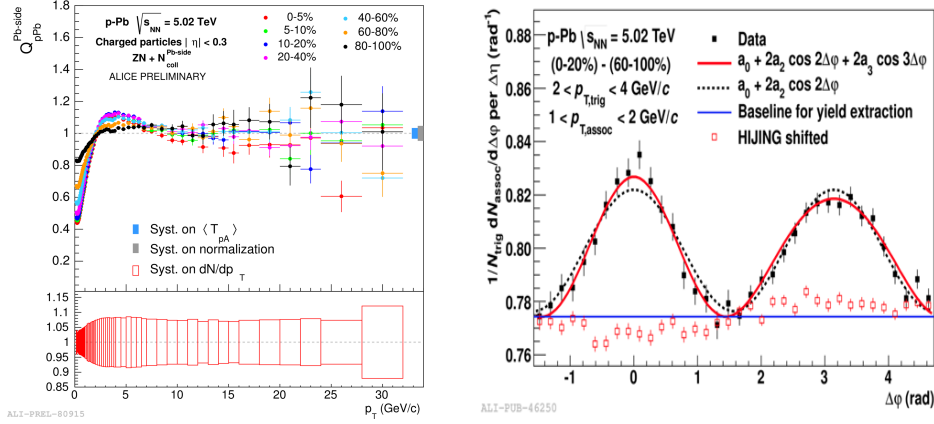


Figure 1.7: (left) Nuclear modification factor for charged particles in p-Pb collisions at $\sqrt{s} = 5.02$ TeV measured by the ALICE collaboration [10]. (right) Distribution of charged particles in relative azimuthal angle $\Delta\varphi$ with respect to a trigger particle in the 0-20 most central p-Pb collisions [11]. After subtracting the contribution given by peripheral collisions, the data points have been fitted to a development in Fourier series at the third order. For comparison, a prediction obtained by the Hijing Monte Carlo generator is shown.

Currently, observations of particle production in p-Pb collisions provide a somewhat puzzling picture: while the measured R_{p-Pb} is very close to 1, showing no or only small nuclear effects, the comparison of the associated yield for charged particles as a function of $\Delta\varphi$ shows a significant deviation from the expectation of a superposition of (mostly) independent nucleon-nucleon collisions.

While the observed v_2 -like effect in p-Pb collisions can be described by hydrodynamical calculations, like in Pb-Pb collisions, it is not clear whether the required density profiles can be generated in p-Pb collisions, and whether the system is large enough to reach the hydrodynamical stage. An alternative explanation is offered by colour-glass condensate models, in which the gluon densities are so large that multiple-gluon interactions are enhanced; quantum mechanical interference effects could produce a flow-like effect [12].

In the next chapter we will give a theoretical description of initial state effects, with emphasis on the Colour Glass Condensate, which is an effec-

tive field theory for the low- x gluon fields in hadrons. We will also discuss some experimental observables that are sensitive exclusively to initial state conditions.

Chapter 2

Photons as a probe of nuclear matter

In the previous chapter we have seen how a complete knowledge of the initial state of AA collisions is fundamental in order to identify final state effects like the production of the QGP in nucleus-nucleus. Proton-nucleus collisions provide a practical way to study initial state effect from the nucleus in a hadron collider, since no QGP formation is expected in such collisions.

Another approach is to use probes which are predicted to be unaffected by the final state modifications of the specific colliding system studied and compare the results of the measurements with the theoretical predictions. Direct photons are among the probes that fulfil this criterion. The qualifier *direct* indicates that these photon do not come from the decay of hadrons, such as π^0 , η , etc., but are directly produced in a hard scattering of initial state partons (quarks and/or gluons), probing the short-distance dynamics of the collision.

In this chapter we will discuss the production mechanisms through which photons are produced in hadronic collisions, the theoretical predictions computed via pQCD calculations and those given by other models, which assume different initial state conditions. We will conclude the chapter with a short review of the experimental techniques and challenges that detecting photons implies.

2.1 Photon production in hadron-hadron collisions

The processes for direct photon production in pp collisions are (at Leading Order in perturbation theory) the quark-gluon Compton scattering and the annihilation process. The Leading Order (LO) Feynman diagrams for such process are presented in Fig. 2.1.

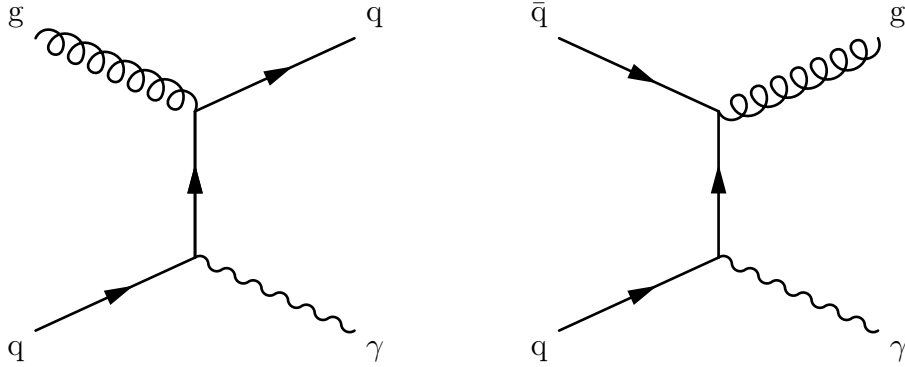


Figure 2.1: t-channel Feynman diagrams at leading order for the two main processes for direct photon production in hadronic collisions: (left) quark-gluon Compton scattering and (right) the annihilation process.

From these diagrams one can see that prompt photon production involves, through its main production channel¹, the interaction of a quark with a gluon. The cross section for such process is therefore sensitive to the gluon density inside the colliding hadrons.

The second consideration one can make is that, at LO, direct photon production processes (and the corresponding measurement) are ideally free from the uncertainties inherent, for example, in the fragmentation mechanisms of partons into hadrons typical of inclusive hadron production processes.

The differential cross section for direct photon production in collisions between two hadrons A and B can be expressed as:

¹Here we are restricting our study to the energy range $E_{\gamma prompt} \leq 100 \text{ GeV}/c$, which is relevant for this analysis

2.1. Photon production in hadron-hadron collisions

$$\frac{d^2\sigma^{A+B\rightarrow\gamma+X}}{dp_T d\eta} = \sum_{i,j=q,\bar{q},g} \int dx_1 dx_2 F_{i/A}(x_1, M) F_{j/B}(x_2, M) \frac{d^2\sigma^{i+j\rightarrow\gamma+X'}}{dp_T d\eta} \quad (2.1)$$

where:

- η is the pseudorapidity, useful in experimental particle physics to describe the emission angle of a particle relative to the beam axis;
- x_1 and x_2 are the fractions of hadron momenta carried by the colliding partons;
- $F_{i/A}(x_1, M)$ and $F_{j/B}(x_2, M)$ are the Parton Distribution Functions (PDFs) of the two colliding hadrons. In case one (or both) colliding hadron is a nucleus, the $F_{i/H}(x, M)$ are nuclear PDFs (nPDFs). The (n)PDFs represent the probability for a parton to be found within the hadron while carrying a momentum fraction x_i of the total hadron momentum;
- M represents the scale at which the cross-section is computed with a perturbative approach. Three scales are usually defined: μ_R, μ_f and μ_F , called respectively the renormalization, the factorisation and the fragmentation scale (relevant in Eq.2.2);
- $\frac{d^2\sigma^{i+j\rightarrow\gamma+X'}}{dp_T d\eta}$ is the differential cross section for direct photon production in parton-parton collision; this can be computed via pQCD calculations at a given order;

At colliders, Next to Leading Order (NLO) contributions become important and must be taken into account in the computation of the direct photon production cross section. The main NLO contributions to this process, reported in Fig. 2.2, yield corrections of $\mathcal{O}(\alpha\alpha_s^2)$ to the LO processes and introduce a dependence of the cross section on the fragmentation of partons into photons. It is then necessary to include the NLO contributions

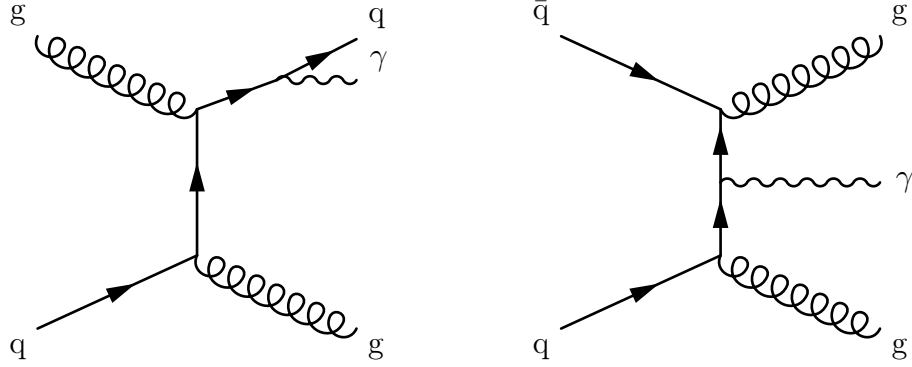


Figure 2.2: Examples of NLO contributions to the t-channel of direct photon production process in hadron hadron collisions for the qg-Compton scattering (left) and the annihilation process (right)

to the computation of the total cross section for direct photons; this can be expressed, similarly to Eq. 2.1, as

$$\frac{d^2\sigma^{A+B\rightarrow\gamma+q+X}}{dp_T d\eta} = \sum_{i,j=q,\bar{q},g} \int_0^1 \frac{dz}{z} dx_1 dx_2 F_{i/A}(x_1, M) F_{j/B}(x_2, M) \cdot D_q^\gamma(z, M_F) \frac{d^2\sigma^{i+j\rightarrow q+X'}}{dp_T d\eta} \quad (2.2)$$

with $\frac{d^2\sigma^{i+j\rightarrow q+X'}}{dp_T d\eta}$ being the cross section for the production of a quark q in a partonic collision; the probability of such quark to emit a photon via bremsstrahlung is then described by the Fragmentation Function (FF) $D_q^\gamma(z, M_F)$, where z is the fraction of the outgoing quark momentum taken by the fragmented photon.

Given the very limited amount of production channels, direct photons are one of the probes that help to characterise hadron-hadron collisions precisely. In fact, even when produced in nucleus-nucleus collision², direct photons leave the hot and dense medium unaffected, since they are colourless, allowing us to have direct knowledge of the initial state conditions of the collision.

We have seen how observables like the nuclear modification factor and the azimuthal anisotropy lead to opposite conclusions once used to discriminate final and initial state effects in collisions between protons and Pb nuclei; the

²this holds true also for pA and pp collisions

2.1. Photon production in hadron-hadron collisions

reason for this discrepancy is often associated with a different sensitivity of the two observables to the initial and final state conditions of the collisions.

We could use photons as a probe to disentangle between initial and final state effects. The study of direct photon production in pp collisions will give us the baseline behaviour for the production rate and can be a quantitative test of pQCD, while measuring the same quantity in pA collisions will allow us to study and quantify any deviation from the theoretical description of the cold nuclear matter structure at high energies. Finally the measurement of the production rate for direct photons in AA collisions will probe in addition the properties of the hot and dense medium produced.

2.1.1 QCD in the low- x regime, limits of applicability

One of the most powerful features of perturbation theory is that it allows to organise the calculations in terms with powers of the coupling constant. This provides a way to systematically improve the precision of calculations by including terms with higher powers of the coupling constant. For direct photons, calculations at NLO are available, and work is ongoing to explore the next order in perturbation theory.

When performing high energy hadronic collisions, the interaction probability for hard scatterings depends on the quarks and gluons PDFs (as we saw in Eq. 2.1 and 2.2), which are probability density distribution of x , and depend on the momentum Q^2 transferred in the hard scattering.

$$f_i = f_i(x, Q^2) \quad \text{for } i = q, \bar{q}, g. \quad (2.3)$$

Our knowledge of the PDFs is mostly derived from a number of *Deep Inelastic Scattering* (DIS) experiments, including the pioneering measurements at SLAC National Laboratory [13, 14] and the systematic measurements at HERA [15] and NMC [16], which cover a large kinematic range.

Since DIS are mediated via the exchange of a (virtual) photon (or a W/Z boson), the cross section measurements obtained at different interaction energies are (mainly) sensitive to the PDFs (and their evolution with Q^2) of charged partons within the proton, while the gluon PDF have been obtained indirectly considering the unitarity of the scattering process via summation

rules.

The evolution of the PDFs with the energy scale is described by the DGLAP (Dokshitzer, Gribov, Lipatov, Altarelli, Parisi) equations [17, 18], which allow, given a measurement of the structure functions of hadrons at a certain scale μ_0 , to know the parton density distributions at any $Q^2 \gg \Lambda_{\text{QCD}}^2$. For low values of Q^2 and x , the DGLAP approach leads to inaccurate results, because the terms of the (perturbatively expanded) PDFs are enhanced by powers of $\ln(1/x)$. In this regime, the evolution of the PDFs is described by the BFKL (Balitskii, Fadin, Kuraev and Lipatov) equations [19–21], which resum leading logarithmic terms of the form $(\alpha_S \ln(1/x))^n$ and, at NLO, $\alpha_S(\alpha_S \ln(1/x))^n$.

The power growth at NLO, which indicates a growth of the parton densities within the hadrons, is such that the cross section for a specific QCD process, computed for low values of x and Q^2 , would at some high energy violate the Froissart bound for which

$$\sigma_{tot} < const \cdot (\ln s)^2.$$

In order to have a consistent description of hadrons, non-linear effects, allowing for the gluon PDFs to *saturate* in the low x - low Q^2 regime, must be taken into account.

2.2 The Colour Glass Condensate effective theory

As we have seen in the previous section, the internal structure of a hadron depends on the scales at which the hadron is probed. When performing hadron-hadron collisions, a precise understanding of the partonic composition of the projectiles at all relevant scales is crucial. This is confirmed, as shown in section 1.3, by recent results from high multiplicity events in p-p and p-Pb collisions at LHC, where QGP-like signatures, usually expected for large colliding systems and high energy densities, have been observed despite the smaller system size.

2.2. The Colour Glass Condensate effective theory

As discussed in the previous section, both the DGLAP and BFKL equations do not include saturation effects that are expected to tame the large gluon densities at low x and low Q^2 .

The *Colour Glass Condensate* (CGC) is one of the most promising attempts to describe coherently the onset of gluon saturation effects in hadrons. The CGC is an effective field theory (EFT) for high-energy QCD scattering, based on the following three main ingredients:

- high gluon densities, as those measured already in DIS experiments at HERA, are treated as strong classical fields, allowing the calculation of hadronic wave functions at small- x through classical techniques; this is justified with the argument that the large occupation numbers of gluons with low x reduces quantum effects.
- quantum corrections are incorporated via non-linear renormalisation group equations, known as BK/JIMWLK³, which allow for a description of the saturation regime via a non-linear term, needed to recover unitarity of the theory;
- universality of the physics in the regime of saturated gluons and its independence of the details of the fragmentation region, which can already be understood with linear evolution.

The first ingredient is illustrated in Fig. 2.3 and can be summarised as follows: when performing high-energy collisions involving hadron(s), the energy scale at which the hadron is probed defines both the spatial and (consequently) the temporal resolution with which the hadron is probed. This means that processes like gluon emission/absorption from/by a parton within the hadron become "visible" to the probe with the increasing energy available for the colliding system. For even higher energies, the probe becomes sensitive also to gluon splitting processes, which lead to an increase of the gluon field within the probed hadron (BFKL regime). The gluon splitting process becomes more and more probable as the fraction x of momentum carried by the daughter gluon gets smaller and smaller, reaching a regime

³Balitsky, Kovchegov/Jalilian-Marian, Iancu, Weigert, Leonidov, Kovner

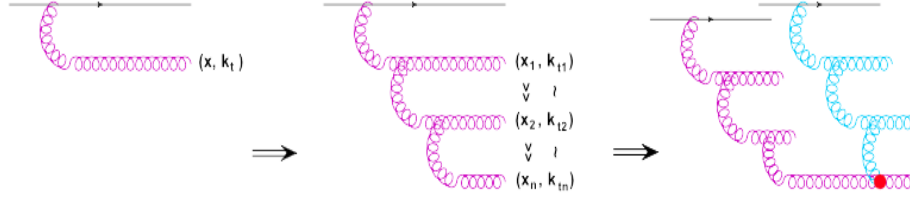


Figure 2.3: Schematic example of relevant diagrams for small- x evolution: (left) gluon emission, (middle) BFKL ladder and (right) fusion of two BFKL ladders in CGC non-linear evolution. Picture taken from [22].

(saturation regime) where different gluon cascades fuse, giving non-linear effects responsible for the saturation of the gluon PDFs.

In this regime, the quantum fluctuations are small compared to the strong gluon field given by the many low- x gluons; the strong gluon field can be described by means of a classical field emitted by valence quarks, the *sources* of the gluon field, which are described via a functional probability distribution $W_\Lambda[\rho]$, where Λ is a certain scale and ρ is the density of the sources inside the hadron. The many low- x gluons described by the classical field can themselves emit even softer gluons. These emission processes are the quantum corrections in the second point of the list above and are incorporated in the probability distribution $W_{\Lambda'}[\rho]$.

All these considerations, valid for hadrons, find an even stronger justification when considering gluon densities typical of heavy nuclei. In fact, for fixed, low-enough, values of x , the scale at which saturation is expected to become relevant in a nucleus is enhanced with respect to the nucleon case, by the *oomph* factor $\approx A^{\frac{1}{3}}$

$$\frac{xG_A(x, Q^2)}{\pi R_A^2} \approx A^{\frac{1}{3}} \frac{xG_N(x, Q^2)}{\pi R_N^2} \implies Q_{s,A}^2 \approx A^{\frac{1}{3}} Q_{s,N}^2. \quad (2.4)$$

where x is the momentum fraction of the parton within the hadron, G_A , G_N are, respectively, the gluon PDFs inside the nucleus and the nucleon; R_A and R_N are the respective radii; $Q_{s,A}^2$ and $Q_{s,N}^2$ are the saturation scales for nuclei and nucleons.

The above equation is obtained from geometrical considerations when calculating the density needed for gluons to overlap, and it expresses the

2.2. The Colour Glass Condensate effective theory

universality implied by the CGC EFT, linking the dynamics of the saturation regime only to the saturation scale with its particular dependence on energy and nuclear size. In the next section we will present a few experimental observables and compare them to the predictions obtained with CGC theory and CGC-based models.

2.2.1 CGC and observables

The CGC EFT was first introduced in 1993 by McLerran and Venugopalan, and has been acquiring more and more consideration within the scientific community since then. Nowadays it is possible to quantitatively predict the behaviour of many experimental observables for e-p and p(/A)-A collisions, relevant to understand the initial state conditions of hadrons in high energy collisions. A lot of work still needs to be done in order to establish a complete formalism and to *easily* produce analytical and/or numerical calculations in the mean field approximation of the CGC.

As already said, CGC effects are expected to be enhanced in p-A collision due to the larger gluon density per transverse area unit; in the following we will present some experimental results obtained by different experiments both at RHIC and at LHC, comparing these with the predictions produced within the CGC framework.

Fig. 2.4 shows the prediction for the number of charged particles as a function of the pseudorapidity in AA collisions at different collision energy and for different colliding systems. The results of the calculations are compared with the values measured by the ALICE collaboration at LHC (left) and by the PHOBOS and BRAHMS experiments at RHIC (right) [23]; the CGC model prediction and the measured values are in good agreement both at RHIC and LHC energies.

The multiplicity of particles produced in a hadronic collision is mainly related to soft processes like multi parton interactions and colour reconnection. These processes usually take place after the collision and therefore the distribution of particles multiplicity might not be the best tool to study initial state conditions of hadronic collisions. Computing the particle spectra as a function of the transverse momentum can refine the knowledge acquired from

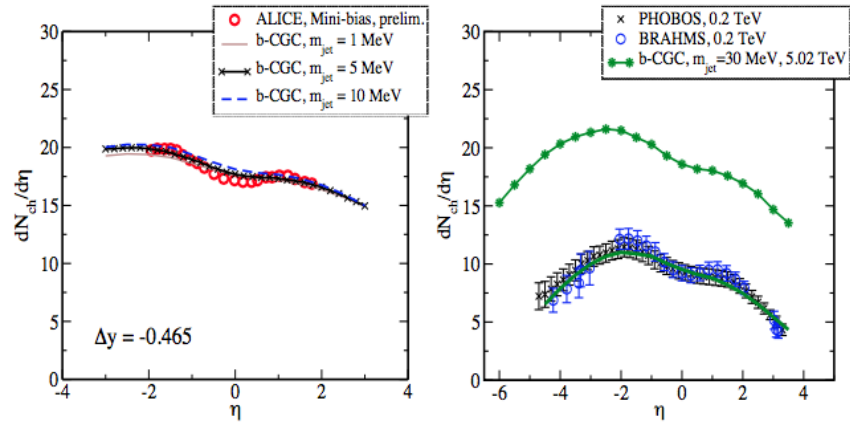


Figure 2.4: Pseudorapidity distributions of charged particles in pA collisions. On the left the results from p-Pb minimum bias collisions at $\sqrt{s_{NN}} = 5.02$ TeV obtained at LHC by the ALICE experiment, while the right panel shows the distributions measured by PHOBOS and BRAHMS collaborations at RHIC in dAu collisions at $\sqrt{s_{NN}} = 200$ GeV. The different curves correspond to CGC-based calculations with different settings. The lower green curve in the right plot indicates the prediction from b-CGC for collisions at $\sqrt{s_{NN}} = 200$ GeV. Figure taken from [23].

2.2. The Colour Glass Condensate effective theory

studying particles multiplicity distributions. To quantify the differences with respect to the pp case, we make use of the nuclear modification factor R_{pPb} , explained in the previous chapter. Figure 2.5 shows the modifications of the measured cross section for charged particle production in p-Pb collisions with respect to the same cross section measured in pp collision. The result obtained by the ALICE collaboration is compared with predictions from different models. As one can see, CGC-based calculations (upper panel, gray band) are consistent with the dependence of the nuclear modification factor on the transverse momentum of the charged particle observed in data.

Of course, in order to be more sensitive to the effects of saturations, it would be best to measure particles produced in hard scatterings between partons with low x . To select these cases this measurement should be repeated for particles produced at forward rapidity. A further improvement, as explained in section 2.1 is to select direct photons, which provide, as discussed in section 2.1, a more direct probe for the gluon density, since they emerge directly from the hard scattering process.

The measurement at high rapidity is particularly interesting since it explores the low- x regime, where saturation is expected to occur. Figure 2.6 shows the prediction for the R_{pPb}^γ for two different scenarios: in blue the prediction obtained from the JETPHOX Monte Carlo partonic event generator [25], [26] and [27], which includes nuclear shadowing effects in the EPS09 set of nuclear PDFs used; the orange band represents the prediction obtained within the CGC framework [28]. The latter predicts a much stronger suppression of the direct photon yield due to saturation effects in the gluon component of the wave function of the nucleus.

Further details and studies on the feasibility of such measurement will be presented in Chapter 7 as a part of the future prospects for the analysis presented in this thesis, while in the following of the chapter we will discuss the techniques to measure photons produced in high energy hadron collisions. In fact, even though the physics of the qg-Compton channel described in 2.1 is fairly simple, there are many processes involving the production of photons with similar signatures. To be able to discard these background processes a study of the surrounding hadronic environment, the very low cross section for such process, and the lack of redundancy in the detection of electromagnetic

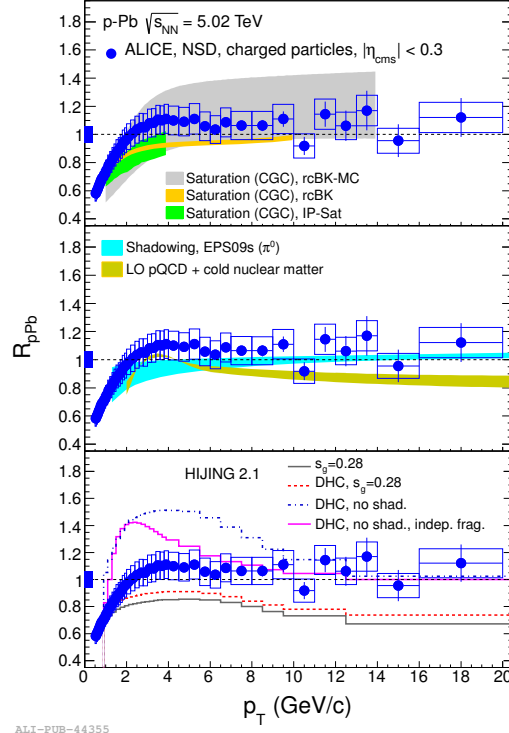


Figure 2.5: Transverse momentum dependence of the nuclear modification factor R_{pPb} of charged particles measured in p-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV [24]. The ALICE data in $|\eta_{cms}| < 0.3$ (symbols) are compared to model calculations (bands or lines, see text for details; for HIJING, DHC stands for decoherent hard collisions). The vertical bars (boxes) show the statistical (systematic) errors. The relative systematic uncertainty on the normalization is shown as a box around unity near $p_T = 0$.

2.2. The Colour Glass Condensate effective theory

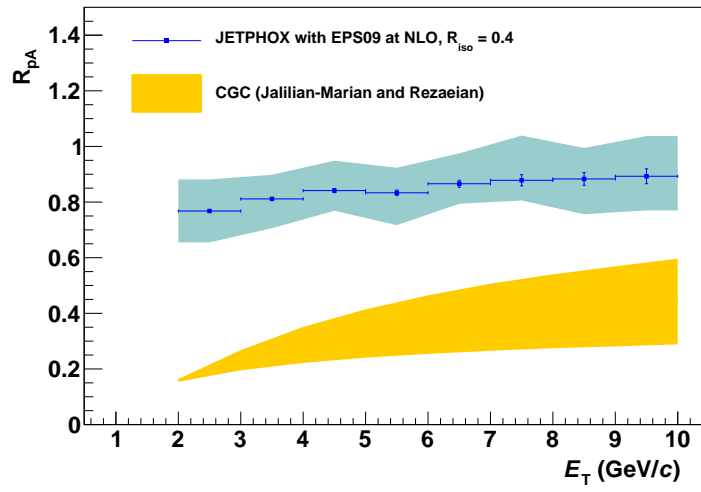


Figure 2.6: Predictions for the direct isolated photon nuclear modification factor R_{pPb} at $\eta = 4$ for p-Pb collisions at $\sqrt{s} = 8.8$ TeV, obtained using two scenarios for the gluon distributions in the nuclear environment. The blue points (with dark cyan band) have been computed by using the nPDF EPS09, which include shadowing effects, while the orange band is the prediction obtained from CGC-based calculation, taken from [28]. The same isolation criterion is used for both calculations.

neutral particles must be taken into account.

2.3 Measuring direct photons: isolation techniques

In Section 2.1 we have seen the main production channels for direct photons productions at hadron collider experiments. An experimental complication in measuring photons produced in pp collisions at energies typical of the LHC is the production of secondary photons coming from the decays of π^0 , η mesons, etc. These represent a source of background, which overwhelms the direct photon signal and greatly complicates the measurement of direct photon performed without any constrain on the environment surrounding the particle we want to study.

To reject the background from decay photons, measurements of direct photon production at collider experiments rely usually on the study of the hadronic activity in a cone of radius R in (pseudo) rapidity⁴ and azimuthal angle around the photon direction,

$$(y - y_\gamma)^2 + (\varphi - \varphi_\gamma)^2 \leq R^2, \quad (2.5)$$

where the y_γ (or η_γ) and φ_γ coordinates represents the direction of the candidate photon and the choice of the radius R is arbitrary and can change accordingly to the particular experiment specifications and design.

In addition to the rejection of the background of secondary photons, the isolation cut also affect the prompt-photon cross section itself; in particular the component of photons produced as a result of fragmentation processes is reduced. This is true for those cases in which the fraction $(1 - z)$ of the transverse momentum carried by the quark *after* having emitted a photon is high enough to not satisfy the isolation criterion adopted experimentally., discussed in section 2.1.

In the following of this thesis, the term "isolated photons" will be intended as synonym for direct prompt photons since, experimentally, the two

⁴The definition of rapidity and pseudorapidity coincide for a massless particle such as the photon.

2.3. Measuring direct photons: isolation techniques

definitions refer to the same observable.

From a theoretical point of view, the introduction of an isolation criterion brings a dependence of the cross section on the isolation parameters. This implies that the cross section might no longer be fully inclusive over the hadronic final state, with the consequence that the factorisation of the isolated photon cross section would not necessarily be valid. Still, it has been proven [25], that, with a suitable collinear and infrared-safe definition of the isolation criterion, factorisation of the isolated photon cross section is instead valid at any order of the QCD perturbation theory.

Throughout the years, physicist have formulated various isolation criteria, few of which are:

1. $\sum_i E_{T,i}^{had} \leq E_{T,max}$,
2. $\max E_{T,i}^{had} \leq E_{T,max}$;
3. $\sum_i E_{T,i}^{had} \leq \epsilon E_{T,\gamma}$;
4. $\sum_i E_{T,i} \theta(\delta - R_{i\gamma}) \leq \varkappa(\delta, E_\gamma)$ for all $\delta \leq R$;

where $E_{T,i}^{had}$ is the transverse energy of the i -th hadron, $E_{T,max}$ is a threshold value chosen by the experiment, ϵ is a constant factor chosen by the experiment, $\theta(\delta - R_{i\gamma})$ is a vanishing function for arguments smaller than zero and \varkappa is a function of the distance δ of both the i -th particle from the photon and the energy of the photon itself.

The different isolation criteria reported constrain, in some way, the kinematic of the fragmentation photon production processes; these result in the inclusion (or exclusion) of some of the contributions in the isolated photon cross section and the sensitivity of the measurement to them.

In this thesis we will adopt the first of the isolation criteria listed above, using values for $R = 0.4$. In order to choose the value for $E_{T,max}$ we will study the total hadronic transverse energy distribution, associated with the production of isolated photons, on Monte Carlo simulation. The results of this study will be presented in section 5.4.

Chapter 3

The A Large Ion Collider Experiment

In this chapter I will present an overview of the ALICE experiment and its detectors, focusing on those used for the analysis in this thesis.

ALICE (*A Large Ion Collider Experiment*) is one of the four big experiments installed along the ring of the *Large Hadron Collider* facility (LHC) at CERN in Geneva. It is a multipurpose experiment optimized for the study of heavy-ion collisions. The primary aim of the experiment is to study the behaviour of nuclear matter at high densities and temperatures, where deconfinement and chiral symmetry restoration are expected to happen. The new state of matter is known as the *Quark Gluon Plasma*, or QGP.

To understand the properties of the QGP and how much it differs from the ordinary cold hadronic matter, measurements of the same observables are performed in simpler colliding systems like proton-proton or proton-nucleus collision, which provide a baseline knowledge of the characteristics of those observables in cold nuclear matter.

3.1 General overview

ALICE consists of a number of different sub-detectors, most of which are mainly designed to identify charged particles and measure their momentum. The layout of the ALICE set-up is shown in Fig. 3.1. With the exception of the muon spectrometer, which is devoted to the study of quarkonium behaviour in dense matter, and covers the large rapidity range, most of the detectors in ALICE are placed in the central part, which covers $\pm 45^\circ$ ($|\eta| < 0.9$) over the full azimuth. These detectors are embedded within a large magnet with a solenoidal field of 0.5 T . Outside of the Inner Tracking System (ITS), there are: a cylindrical Time Projection Chamber (TPC), a large area array of Time-of-Flight (TOF) counters and a Transition Radiation Detector (TRD). In addition the Electromagnetic Calorimeter (EMCAL) allows also the detection and measurement of the energy for electrons, positrons and photons.

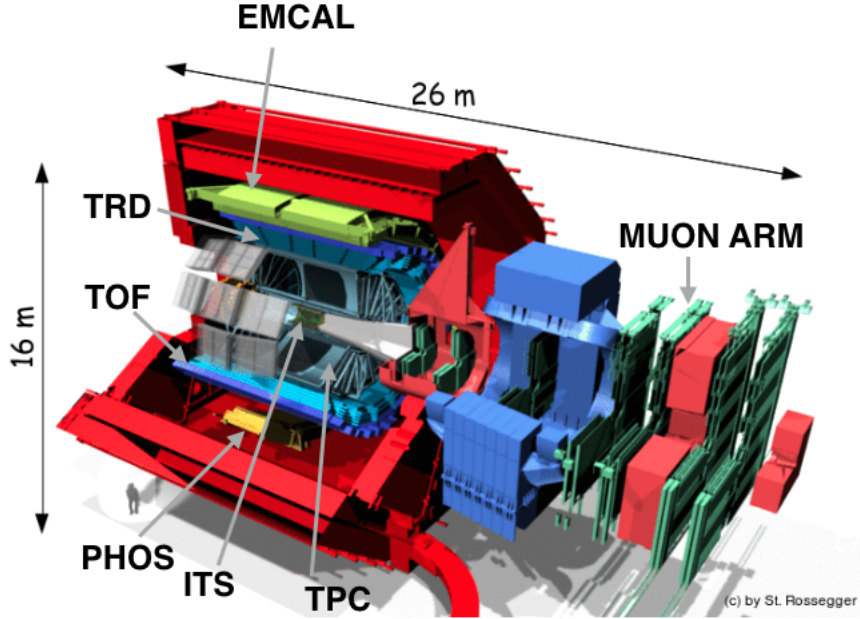


Figure 3.1: ALICE and its main detectors.

3.2 The Inner Tracking System

The ITS is the detector closest to the beam pipe. It consists of six cylindrical layers of silicon detectors, two layers of Silicon Pixel Detectors (SPD), two layers of Silicon Drift Detectors (SDD), and two layers of Silicon Strip Detectors (SSD). The main specifications and uses for each detector of the ITS can be found in Table 3.1 while a sketch of the design is shown in Fig. 3.2.

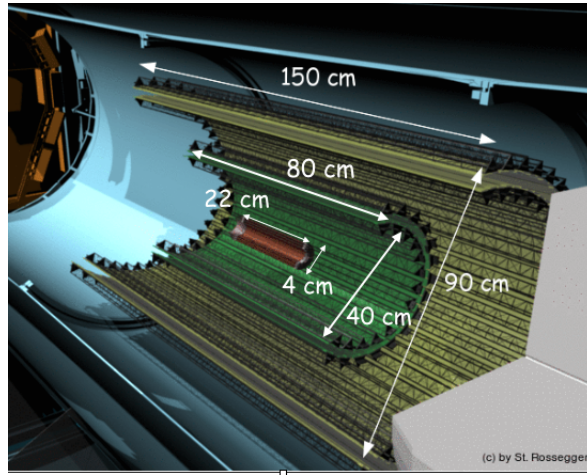


Figure 3.2: The Inner Tracking System of the ALICE Detector.

Silicon pixel detectors (SPD) are chosen for the innermost two layers in order to be able to achieve a good two-particle separation even in an environment with large particle density, typical of central heavy-ion collisions. Their granularity is such that the average occupancy is as low as 1.5% for the inner layer. Around the SPD, two layers of silicon drift detectors are installed, while the outer two layers are double-sided silicon micro-strip detectors. With the exception of the two innermost pixel layers, all layers identify particles via ionization energy loss (dE/dx) measurements in the non-relativistic ($1/\beta^2$) region. This gives the ITS a stand-alone capability as a low- p_T particle spectrometer.

The Inner Tracking System is also used to extend and complement the momentum measurement of the TPC and to improve the determination of

Parameter	Units	SPD	SDD	SSD
Cell Size	μm^2	50×300	150×300	95×40000
Spatial precision ($r\phi$)	μm	12	38	20
Spatial precision (z)	μm	70	28	830
Two Track resolution ($r\phi$)	μm	100	200	300
Two Track resolution (z)	μm	600	600	2400
Total number of readout channels	k	15729	133	2719

Table 3.1: ALICE ITS specifics.

the position of the extrapolated tracks, mainly used in secondary vertex reconstruction. The granularity (mainly) of the SPD is such that a displaced secondary vertex can be found with ALICE ITS down to an impact parameter of $\sim 100 \mu\text{m}$. The granularity of the SPD is optimized for a maximum track density of 8000 tracks per unit of rapidity and a spatial resolution of a few tens of μm [29].

3.3 The Time Projection Chamber

The heart of the ALICE experiment is the TPC. It is a 5.1 m long cylindrical gas detector with an outer radius of 2.5 m and an inner radius of 85 cm filled with a mixture of $\text{Ne}_2/\text{CO}_2/\text{N}_2$ (85.7%/9.5%/4.8%). A schematical view can be seen in Fig. 3.3.

An electric field is applied between the end-plates (anodes), which are divided into 18 trapezoidal sectors hosting the read-out chambers, and the central membrane of the TPC, which serves as a HV (100 kV) cathode. Charged particles that pass through the TPC ionize the gas, producing electron-ion pairs. The electrons subsequently drift towards the read out chambers. This design ensures the detection of charged particles and allows the reconstruction of their tracks through the sampling of up to 160 hits in the TPC. The high voltage gradient at which the field cage is operated results in a maximum drift time of about $90 \mu\text{s}$. To keep the occupancy low and ensure good

3.4. The V0 detector

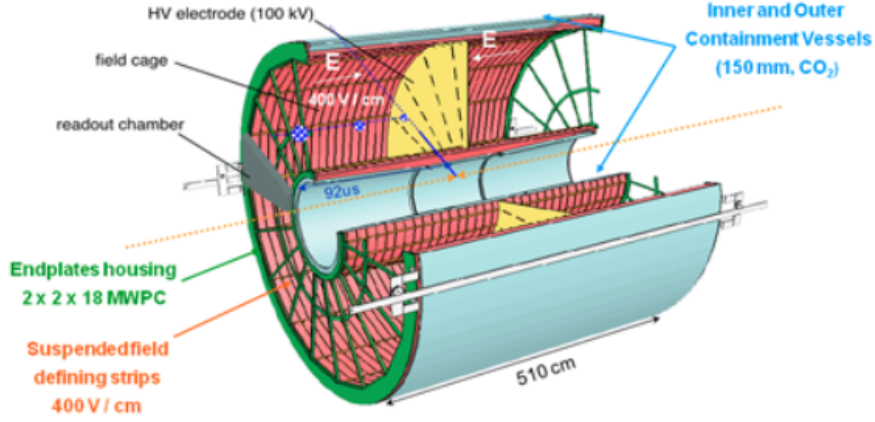


Figure 3.3: A schematic view of the ALICE TPC. The inner radius is 85 cm while the outer radius is 250 cm

dE/dx and position resolution, there are about $5.6 \cdot 10^5$ readout pads in the chambers. The TPC in combination with the ITS, provides track reconstruction with a momentum resolution of 10% for a transverse momentum up to 100 GeV/ c at 0.5 T magnetic field. A more detailed explanation of the ALICE TPC specifications and performance can be found in [30].

3.4 The V0 detector

The V0 detector is a small-angle detector consisting of two arrays of scintillator counters, named V0A and V0C, installed on both sides of the ALICE interaction point. The counters cover the pseudo-rapidity ranges $2.8 > \eta > 5.1$ (V0A) and $-3.7 < \eta < -1.7$ (V0C).

This detector has several functions, one of which is to provide the minimum-bias (MB) triggers for the experiment in all colliding systems.

The efficiency of tagging good pp collisions with respect to all other possible processes can be as high as 89 % when using V0A only, and down to 83 % when requiring a coincidence signal between the two V0 detectors. Although a lower tagging efficiency is found, the coincidence of both V0 detectors is more effective in suppressing beam background. Despite the presence of a significant background from secondaries produced in the vacuum chamber

elements or coming from beam-halo or halo-halo interactions, the relation between the number of impinging particles on the V0 arrays and the number of emitted primary particles remains close to linear. The multiplicity measurement in the V0 detectors serves as an indicator of the centrality of the collision. The time resolution (~ 1 ns) of the V0 detectors allows to distinguish collisions in subsequent bunch crossings. The layout of the two V0 detectors is shown in Fig. 3.4, while more specifications of the V0 detectors can be found in [31].

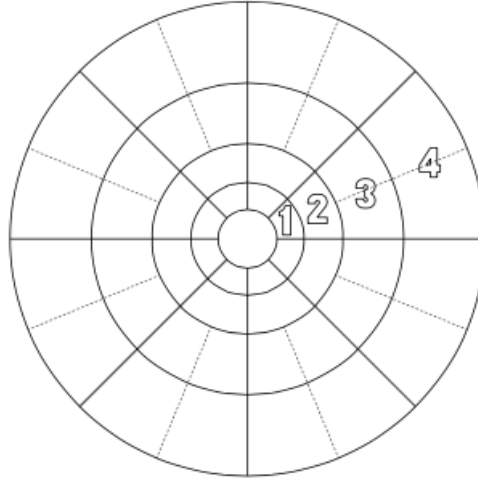


Figure 3.4: Segmentation of the V0A (full lines) and V0C (full and dotted lines) arrays.

Both detectors are placed as close as possible to the beam pipe. Each array consists of 32 counters distributed in 4 rings. Each of these rings covers $0.5 - 0.6$ unit of pseudo-rapidity and is divided in 8 sectors (45°) in azimuth. For the V0C array, elements of rings 3 and 4 are divided further into two identical sectors for optimised signal uniformity and a better time resolution. The sensitive material of the two arrays consists of scintillators read out with photomultipliers connected via optical fibers.

3.5 The Electro-Magnetic Calorimeter

The Electro-Magnetic Calorimeter (EMCal) was built to explore the physics of jet quenching in heavy-ion collisions, via the measurement of electrons, photons and the electromagnetic component of jets, at the LHC over a wide kinematic range. The EMCal is located right in front of the ALICE magnet coil at a radius of ~ 4.5 meters from the beam line. It has a coverage of 107° in azimuth and $|\eta| \leq 0.7$. The size of the EMCal is constrained by the available free space and the maximum weight that can be supported by the L3 magnet.

The EMCal is a layered Pb-scintillator sampling calorimeter. An illustration of the design of the EMCal detector is shown in Fig. 3.5.

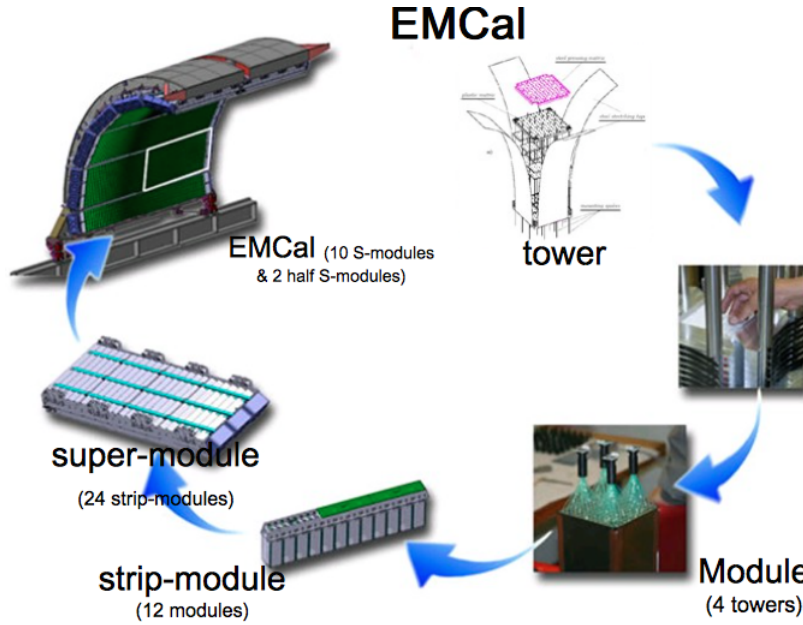


Figure 3.5: A schematic view of the ALICE EMCal composition and appearance in 2012, when the φ coverage was enlarged with the installation of two smaller supermodules covering 6° .

The smallest building block of the calorimeter are the towers; each tower is built with 76 alternating layers of 1.44 mm Pb and 1.76 mm scintillator. The front-face dimensions of the towers are $6 \times 6 \text{ cm}^2$, resulting in an individual

Tower Size at $\eta = 0$	$\sim 6.0 \times \sim 6.0 \times 24.6 \text{ cm}^3$
Tower Size	$\Delta\eta\Delta\phi = 0.0143 \times 0.0143$
Effective Radiation Length X_0	12.3 mm
Effective Molière Radius	3.20 cm
Sampling Fraction	10.5
Number of Towers (per Full SM)	1152

Table 3.2: ALICE EMCal and DCal specifics.

tower acceptance of $\Delta\phi\Delta\eta \sim 0.014 \times 0.014$ rad. Each tower has a thickness of 20.1 radiation lengths (X_0) and ~ 1.3 interaction lengths (λ) and an effective Molière radius of 3.20 cm. Groups of 2×2 towers form a module. Twelve of such modules are arranged to form a strip module. A full size super module is assembled from 24 strip modules, $12 \times 24 = 288$ modules, which span $\Delta\eta = 0.7$ and $\Delta\phi = 0.349$ rad (20°). At the time the data used in this analysis were taken (2011), the EMCal featured 10 full size SM.

A list of EMCal physical parameters can be found in Table 3.2.

- 1 The thickness of each tower allows for electrons and photons to produce
- 2 an electromagnetic shower of particles. Such a cascade is the product of
- 3 the interaction of high energy photons with the electromagnetic field of
- 4 the atoms in the Pb layers which results in the production of an electron
- 5 positron pair. Once produced, the electron and positron can emit photons
- 6 via bremsstrahlung. The pair production cross-section depends on the energy
- 7 as

$$\frac{d\sigma}{d\Omega} = \frac{Z^2}{137} \frac{r_0^2}{2\pi} m_0^2 c^4 F_{pair}^1$$

and vanishes at the critical value $E = 1.022$ MeV. In the above formula, Z is the atomic number of the material, r_0 is the electron radius, and m_0 is the electron mass.

When the energy of the particles in the shower reach the above mentioned

¹ F_{pair} is a function of energy, momentum and angle of both particles produced in the process

3.5. The Electro-Magnetic Calorimeter

value of 1.022 MeV, the number of particles in the shower stops increasing and the particles will lose energy via bremsstrahlung.

The total energy deposited in the scintillator layers by the particles of the shower is proportional to the energy of the impinging particle. Hadrons on the other hand rarely deposit a large amount of energy in EM calorimeters. The interaction of hadrons with matter is governed by the interaction length λ , which is a physical constant linked to the probability for hadrons to interact strongly with the nuclei of the detector. The interaction length is much larger than the radiation length, indicating a smaller interaction probability. Detailed GEANT3 simulations with the longitudinal segmentation described above have shown that the resolution of EMCal for electrons and photons depends on the energy of the particle as

$$\frac{\sigma_E}{E} = A + \frac{B}{\sqrt{E}} + \frac{C}{E}$$

with $A = (1.443 \pm 0.026)\%$, $B = (6.895 \pm 0.088)\%$ over the range $p_T = 5$ to 100 GeV/c, and C is a term due to electronic noise, hence independent of the energy. These values are in agreement with those obtained in an analysis of the data acquired during the 2009 test beams performed at the PS and SPS facilities [32, 33].

For each cell with a non-zero signal, the amplitude, the time of arrival and the index of the cell are stored and then used to reconstruct the original properties of the impinging particle by means of a clustering algorithm, which will be explained at the end of the chapter.

3.5.1 Selecting rare events with EMCal: triggering

The most important features of the EMCal are an efficient and fast trigger (Level 0/1) for high energy jets and the capability to measure the neutral portion of jet energy. The EMCal also enhances capabilities of ALICE to measure high p_T photons, neutral hadrons, and electrons.

Level 0 trigger

EMCal provides a trigger for rare events. In 2011, during the proton physics period, only the Level-0 (L0) trigger was available.

The L0 algorithm runs in so-called Trigger Region Units (TRUs), each of which covers one third of a super module. The signal amplitudes recorded by each tower of a module are summed inside the Front End Card and, together with the timing of the signal, are passed to the Level-0 triggering algorithm. Here the information of 2×2 modules is summed into patches of 4×4 towers and sent to the L0 firmware, together with the information from all other patches inside a TRU. The procedure just described is then repeated for all possible configurations of 2×2 neighbouring modules (sliding patch). The patches that recorded an energy deposit above a certain threshold (~ 5.5 GeV for 2011) are stored as L0 Information.

In 2012 the L1 trigger was implemented and used. The L0 trigger information is used as input for the L1 trigger, which allows to distinguish triggers from single electrons or photons (L1-EGA) and cases in which the trigger has been fired by jets (L1-EJE). Both L1-EGA and L1-EJE are based on a sliding patch procedure similar to that of L0. The difference is that the size of the L1 patch is 8×8 towers for the L1-EGA trigger and 32×32 towers for the L1-EJE allowing, in the second case, for patches to cross the boundaries of the EMCal supermodules.

3.5.2 Clustering algorithm

The calorimetric method for the identification of photons is based on the interaction of photons with the material through pair production and bremsstrahlung processes. Photons interacting with the EMCal material will produce more and more particles through these two interaction mechanisms and the end result usually is a structure called electromagnetic shower.

From now on the terms "electromagnetic shower", "cluster" and "energy deposited by a neutral particle" will be used as synonyms, (at least from an experimental point of view).

In order to reconstruct the total energy of the impinging particle, a clustering algorithm (also called *clusterizer V1*) is used; this is an attempt to reconstruct the showers produced by the interaction of particles with the EMCal (active) material and to extract information on such particles via the study of the transverse profile of the signal recorded by the cells of a cluster.

3.5. The Electro-Magnetic Calorimeter

The clusterizer first identifies all cells that measured an energy above a certain threshold (for this analysis the value of 300 MeV was used) and marks them as seeds. Clusters are formed by adding neighbouring cells to the seed as long as the energy they measure is above 100 MeV. This procedure is continued as long as there are valid cells found for the aggregation step. A cell can always belong to only one single cluster. The energy resolution obtainable with this clustering algorithm has been studied by analysing simulated events. Collisions between protons have been simulated with the Pythia generator [34]. The response of the passage of the collision products through the ALICE detectors has been simulated with GEANT3 [35].

After applying the clusterizer to the simulated response of the EMCal cells, and applying the energy calibration constants, we compare the reconstructed clusters energy with the energy of the particles that produced them.

In Fig. 3.6 shows the result of this study for two different clustering algorithm, the above-mentioned clusterizer V1 and *clusterizer V2*. The V2 clustering algorithm is similar to the clusterizer V1, but, before assigning a cell to a cluster, a check on the energy of the cell with respect to the energy of the seed of the cluster is performed: if the energy measured by the cell is lower than that measured by the seed, the cell is assigned to the cluster, otherwise the cell is not included in the cluster that is being built and the algorithm continues as long as valid adjacent cells are found. The excluded cell will afterwards act as seed for a new cluster, with the same aggregation procedure applied to cells not yet previously grouped into clusters.

A good energy resolution is found for clusters with reconstructed energy $E \geq 10$ GeV/ c , with the V2 algorithm showing slightly worse energy resolution.

The discrimination of clusters produced by single photon from those produced by π^0 and η meson decay photons is expected to be simpler, up to 20 GeV/ c , using the V1 algorithm, more suited to discriminate between single and multi-particle clusters. On the other hand the V2 algorithm is expected to perform better with measurement involving the reconstruction of the invariant mass of neutral pions.

For these reasons, we decided to use, in our analysis, the V1 algorithm instead of the clusterizer V2.

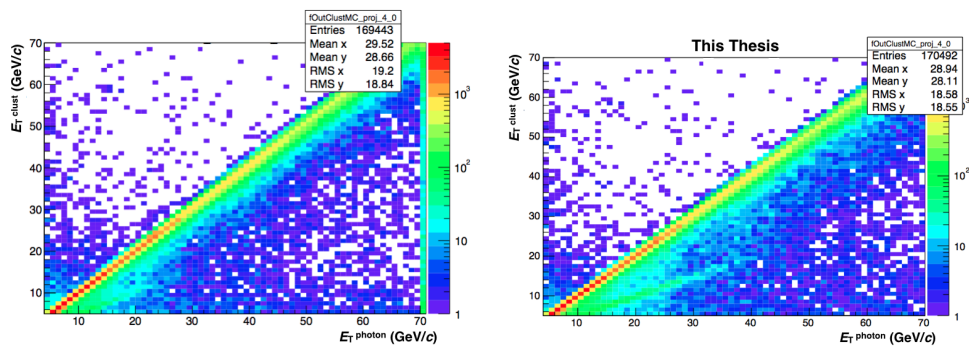


Figure 3.6: Cluster energy distribution as a function of the energy of the prompt photon that is associated with the cluster, reconstructed using the clusterizer V1 (left) and clusterizer V2 (right) algorithm.

Chapter 4

Data analysis

In this chapter the analysis method applied to the data sample to extract an uncorrected isolated photon spectrum will be explained. Corrections for purity and efficiency make use of Monte Carlo detector response simulations and will be discussed in the next chapter, and the result of these corrections on the measured isolated photon spectrum will be shown in Chapter 6.

This chapter first introduces the single event characterisation and the selection of good-quality events. Afterwards we will discuss the analysis method for the identification of the single photon signal from the ALICE-EMCal detector via the study of the transverse shape of the reconstructed clusters and finally we will show the results obtained after selecting isolated photons from decay photons by means of an isolation criterion, together with the estimation of the residual contamination in our sample.

4.1 Data sample, event selection

The results presented in this thesis have been obtained by analysing the data from pp collisions at $\sqrt{s} = 7$ TeV recorded in 2011 by the ALICE experiment. The trigger algorithm used to select events is the EMC-L0, which selects minimum bias events (showing a coincidence trigger on the forward scintillators V0A and V0C) with a large energy deposit in the EMCal. The trigger thresholds in this period were set to 4.5 (not used in this analysis) and 5.5 GeV/ c . The L0 trigger algorithm has been explained in Chapter 3. The

total number of events recorded with high energy threshold is approximately 10 millions. In order not to be biased by effects due to the onset of the triggering algorithm, the lowest transverse energy explored in this analysis is $E_T = 10 \text{ GeV}/c$.

4.1.1 Event selection

Among the events recorded, we accept only those for which the position of the reconstructed primary vertex, along the beam axis, is found to be less than 10 centimetres away from the nominal centre of the ALICE experiment [36]¹. The events that pass this selection are less affected by inefficiency of the tracking algorithm. Moreover the measurement of the transverse momentum of the particles will be more reliable. Figure 4.1 shows the distribution of the reconstructed z-coordinate of the primary vertex. After the selection, our event sample is reduced to 8.8 million events.

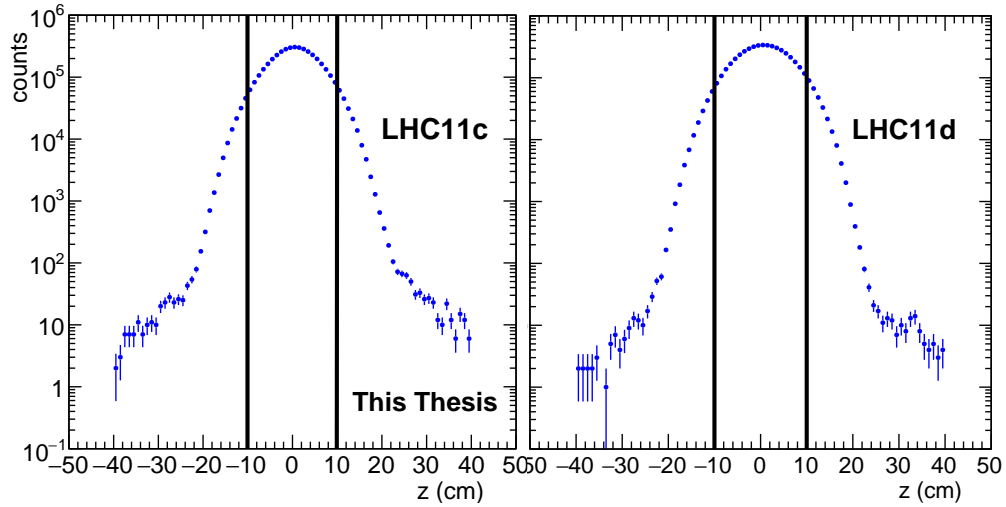


Figure 4.1: Distribution of the z-coordinate of the primary vertex, obtained from the tracking procedure. The accepted events have a reconstructed primary vertex within 10 cm from the center of the ALICE experiment (black vertical lines).

¹Justification for such selection can be found in Fig. 26 of the reference

4.1.2 EMCal data quality assurance studies

The EMCal provides information on both the energy deposited in each cell and the arrival time of the signal; while the former is used to measure the energy of the particle detected, the latter is used to make sure that the selected signals belong to the same bunch crossing for which the L0 trigger was fired.

In case the measured energy exceeds 300 MeV for a certain cell, a cluster is built from it, according to the algorithm explained in Section 3.5.2, and the time of arrival of the cell (called *seed cell*) signal is assigned to the cluster. In order to remove clusters produced by consecutive bunch crossings in pp collisions, we restricted our analysis to clusters with a production time in the time window $[-30, 30]$ ns around the triggered bunch crossing.

Fig. 4.2 shows the distribution of (relative) time for reconstructed clusters. The clusters associated with EMCal-L0 triggered events show a time distribution peaked around 0 ns, while clusters whose signal is recorded at an earlier or later time belong to an earlier bunch crossing or a subsequent one.

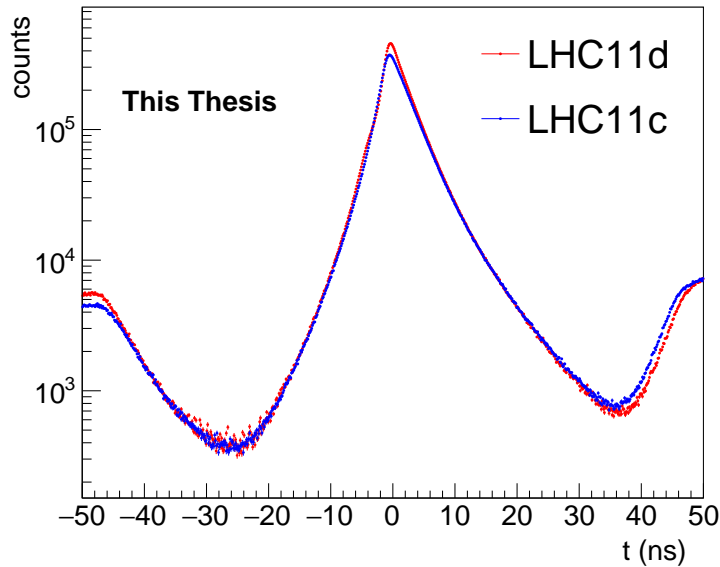


Figure 4.2: Distributions of time for clusters in both periods of data taking.

Cluster quality cuts

Fig. 4.3 shows the reconstructed position of clusters in the two periods of data taking analysed. The distribution for the LHC11c period shows more empty areas, where dead channels were masked, than the LHC11d period. For the LHC11c period, there were also two regions (in azure) with a reduced activity, which have been excluded from the analysis. To ensure a good energy resolution for the reconstructed clusters, a cut on the minimum distance (2 cells) between the reconstructed cluster centroid and any *bad channel* is applied ².

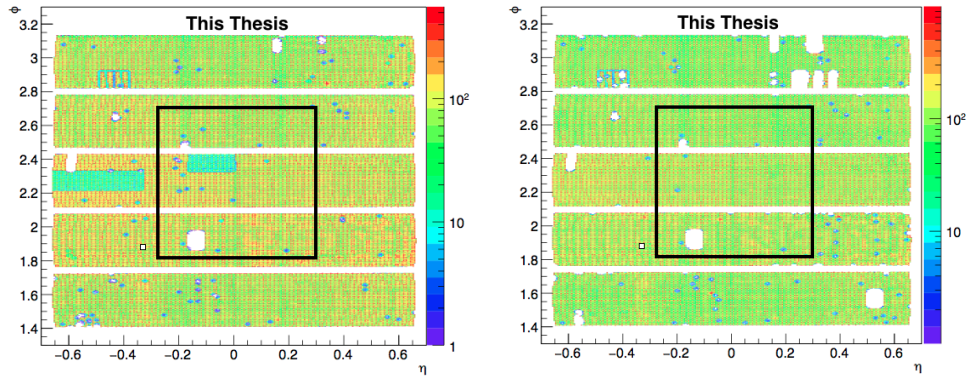


Figure 4.3: EMCal cluster distribution in the (η, ϕ) plane for the data collected during the LHC11c period (left) and LHC11d (right). The black borders indicate the $\eta - \phi$ acceptance for candidates in this analysis.

Moreover, all clusters made up of only one cell (minimum ionizing particles, MIPs) and all clusters with a reconstructed energy smaller than 500 MeV (for which we know EMCal doesn't show a good response) are cut out from our sample because the response of the EMCal cells is not good/reliable enough.

After selecting the clusters according to the selection criteria just explained, the topology of the cluster has been studied by means of the F_{cross} variable, defined as

²studies of this can be found in [32] and [37]

4.2. Charged Particle Veto

$$F_{cross} = 1 - \frac{E_{cross}}{E_{total}}$$

where E_{cross} is the sum of the energy of the cells that share a side with the seed cell, while $E_{cluster}$ is the cluster energy. If the calculated value for F_{cross} is higher than 0.97 the cluster is rejected. Figure 4.4 illustrates the topology of cells used to define the cut variable. Following the specifics reported in Section 3.5, we expect electromagnetic showers to have F_{cross} values smaller than 0.90.

In fact, it has been shown [38] that clusters with $F_{cross} > 0.97$ are mainly produced by neutral hadrons hitting directly the avalanche photodiode (APD), producing a great signal in a very limited portion of space (basically 1 cell).

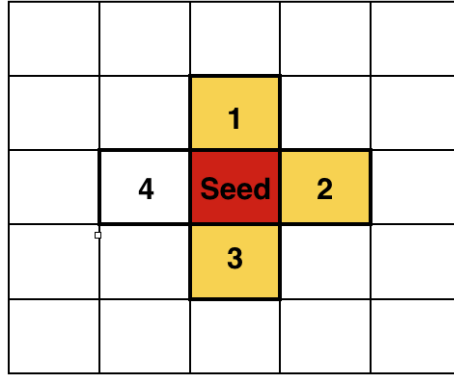


Figure 4.4: F cross cut topology. If the energy measured by the **seed** cell is more than 97 % that of the reconstructed cluster, this is rejected.

All three conditions also help to further reduce the contribution of neutral hadrons in our sample

4.2 Charged Particle Veto

The cuts described so far aim to exclude from the analysis both badly behaving cells, and clusters affected by them, and to eliminate possible sources of contaminations due to low-energy neutral hadrons.

Still our cluster sample is highly contaminated both with clusters produced by charged hadrons and electrons. These particles are also reconstructed by the charged particle tracking system of the ALICE experiment so we can extrapolate what their position would be at the EMCal surface. A cluster with a reconstructed centroid close enough to the extrapolated position of a track is then considered to be produced by a charged particle and excluded from our sample. This procedure is referred to as *Charged Particle Veto*. In order to establish the criterion for "closeness" between cluster and track we have studied the per-cluster distributions of the distance between track and cluster in pseudorapidity ($\Delta\eta_{C-T}$) and in azimuthal angle ($\Delta\varphi_{C-T}$) at a distance from the interaction point of 440 cm. Figures 4.5 and 4.6 show these distributions which both peak at zero and drop immediately for distances larger than 0.02 and 0.03 in η and φ respectively. Therefore clusters with a reconstructed centroid farther away than these values, in η and φ , from any track are considered neutral.

After vetoing all clusters which are likely to be produced by charged particles, the main contributors to our cluster sample are:

- prompt photons
- photons from fragmentation processes
- single decay photons from neutral hadrons
- decay photon pairs from neutral hadrons

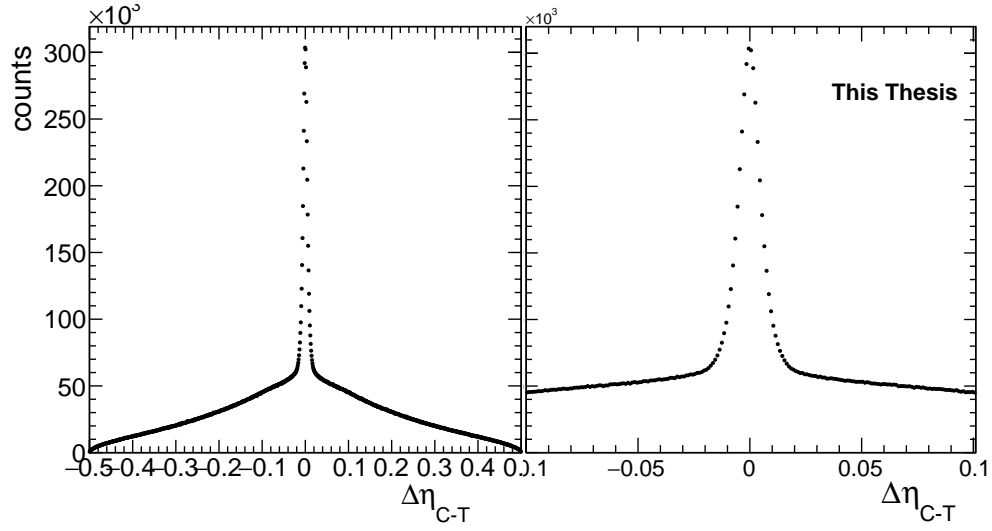


Figure 4.5: $\Delta\eta_{C-T}$ distributions for all reconstructed clusters and (right) a zoom on the correlation peak.

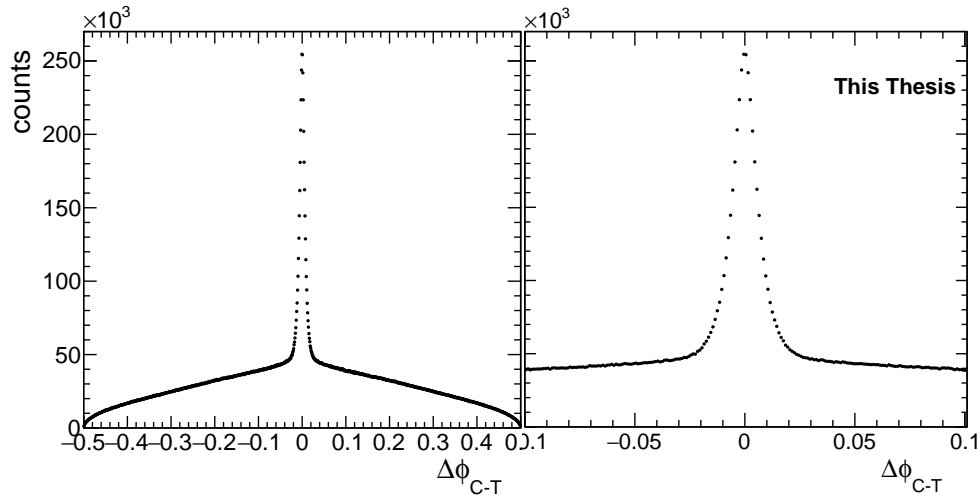


Figure 4.6: $\Delta\phi_{C-T}$ distributions for all reconstructed clusters and (right) a zoom on the correlation peak.

4.2.1 Multi-contribution clusters

To further select only clusters produced by direct and decay photons we need to find a variable sensitive to some characteristic of the cluster related to the number of particles that produced it.

While single-photon clusters are mainly expected to have only one maximum, clusters produced by a pair of decay photons from neutral mesons would typically have two local maxima. Clusters that show more than two local maxima will have contribution from more than two particles.

In order to compute the number of local maxima within a cluster, the energy measured by all cells within a cluster is weighted logarithmically with the distance of the cells from the seed (according to [39]) and then compared to that of the seed cell. If a cell within a cluster measures an energy above a certain threshold it will be tagged as a further maximum of the cluster. The logarithmic weighting is necessary to avoid that single particle clusters in which two neighbouring cells measure each a large fraction of the particle energy will be identified as clusters with two local maxima, complicating the clusters selection based on the number of local maxima. In fact, in case two seed cells are adjacent, the position of the local maximum of the cluster is computed by means of an average of the position of each cell center weighted with the energy measured by each cell.

Fig. 4.7 shows the number of local maxima found in the clusters transverse energy distribution as a function of their transverse momentum. Using the number of local maxima in a cluster and excluding all clusters with more than two local maxima, we reduce the contamination of our cluster sample due to neutral hadrons and other cases where multiple particles produce overlapping clusters.

Such a selection, however also excludes clusters produced by decay photons from neutral mesons in which one of the two photons converts in one of the detectors installed between the interaction point and the EMCal surface (for example the TPC or the TRD). If the conversion happens at less than 180 cm from the beam axis, a track can be properly reconstructed and the electrons are identified; the produced cluster will be matched with the electron track and excluded by the CPV algorithm; if the photon converts

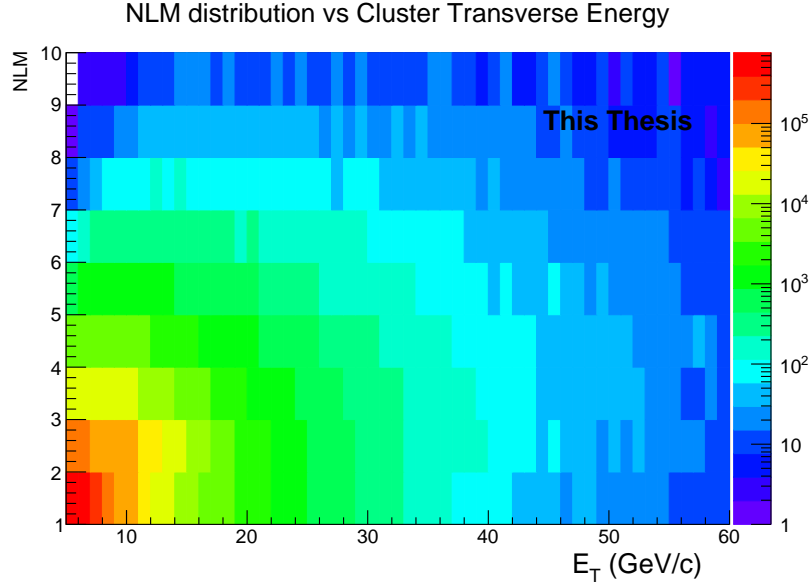


Figure 4.7: Number of maxima found in the transverse energy distribution of neutral clusters as a function of the transverse momentum of the cluster. The distribution refers to LHC11c period.

further away from the beam pipe, the electron tracks cannot be properly reconstructed and the cluster produced will be tagged as neutral. In case the opening angle of the decay partner of such photon is small enough, the resulting cluster will show three local maxima and the cluster will be disregarded.

Results in [40] have shown that the the majority of photon conversions happen in the ALICE-TRD, which is installed right in front of EMCal. Nonetheless photon conversions in the TRD do not add any complication, since the kinematics of the such process yields a collinear electron-positron pair producing, in most of the cases, one single cluster.

Fig. 4.8, shows the neutral clusters spectrum after applying the various cuts described up to now obtained from pp events from the 11c period.

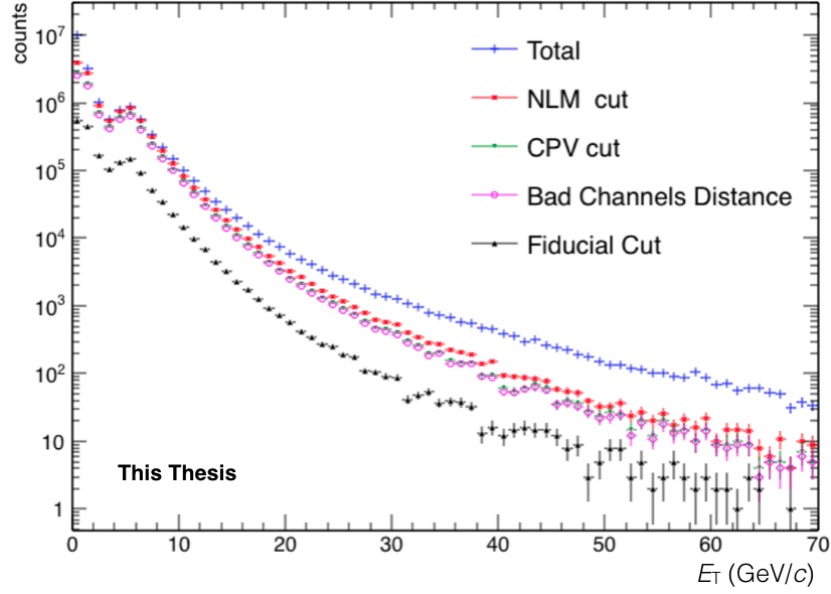


Figure 4.8: Transverse momentum distribution of clusters after each cut applied to obtain the neutral clusters p_T spectrum. These spectra are obtained from pp events from the 11c period.

4.3 Photon cluster identification

In this section the strategy to select photon clusters from our neutral cluster sample will be explained. Our particle of interest is not detected by any of the inner detectors in ALICE (ITS,TPC,TRD,ToF), while it deposits a large amount of its energy in the EMCal. Although neutral baryons like neutrons, Λ etc. and neutral mesons like K_0 are also not detected directly by the inner barrel detectors, due to the limited longitudinal thickness of EMCal, they will typically deposit only a fraction of their energy. Clusters produced by such particles will therefore fail one of the selection criteria explained in the previous section; our neutral cluster sample will constitute almost exclusively of clusters produced by either direct or decay photons.

4.3.1 Transverse shower shape σ_{long}^2

The procedure to reconstruct the particle energy used in the clusterizer V1 has been described in Chapter 3. The clusterizer V1 has been adopted for this analysis because of its discriminating power between one and two-particles clusters, which is crucial to identify clusters produced by both decay photons from neutral pions. The transverse energy distribution of the cluster is studied by means of *shower shape parameters*, which characterise the transverse extension of the cluster. These are defined as:

$$\sigma_{\text{long}}^2 = \frac{\sigma_{\varphi\varphi} + \sigma_{\eta\eta}}{2} + \sqrt{\left(\frac{\sigma_{\varphi\varphi} - \sigma_{\eta\eta}}{2}\right)^2 + (\sigma_{\varphi\eta})^2} \quad (4.1)$$

$$\sigma_{\text{short}}^2 = \frac{\sigma_{\varphi\varphi} + \sigma_{\eta\eta}}{2} - \sqrt{\left(\frac{\sigma_{\varphi\varphi} - \sigma_{\eta\eta}}{2}\right)^2 + (\sigma_{\varphi\eta})^2} \quad (4.2)$$

where $\sigma_{\varphi\varphi}$, $\sigma_{\eta\eta}$, $\sigma_{\varphi\eta}$ are the variances weighted by the cells energies, i.e.

$$\sigma_{\varphi\eta} = \frac{\sum_i w_i \varphi_i \eta_i}{\sum_i w_i} - \frac{(\sum_i w_i \varphi_i)(\sum_i w_i \eta_i)}{(\sum_i w_i)^2}.$$

and w_i are defined as:

$$w_i = 4.5 - \ln\left(\frac{E_i}{E_{\text{cluster}}}\right) \quad (4.3)$$

The σ_{long}^2 and the σ_{short}^2 parameters represent the dispersion of the cells signal distribution in the lateral direction, and can assume different values for clusters produced by one, two or more particles. The logarithmic weights in eq. 4.3 reduce biases in the determination of the cluster position and width due to the relatively coarse granularity of standard calorimeter readout and the non-gaussian distribution of the cluster transverse energy. The method was first introduced in [39], where a more detailed explanation can be found; in particular, the choice of the factors in Eq. 4.3 is based on the studies, performed on Test Beam data and simulated samples, reported in [41].

While the distance between the two decay photons at the EMCal surface depends only on the kinematics of the decay for neutral mesons like π^0 and η , the possibility to *separate* the two particles also depends on the granularity

of the detector and on the clustering procedure. This is illustrated for a few characteristic situations in Fig. 4.9.

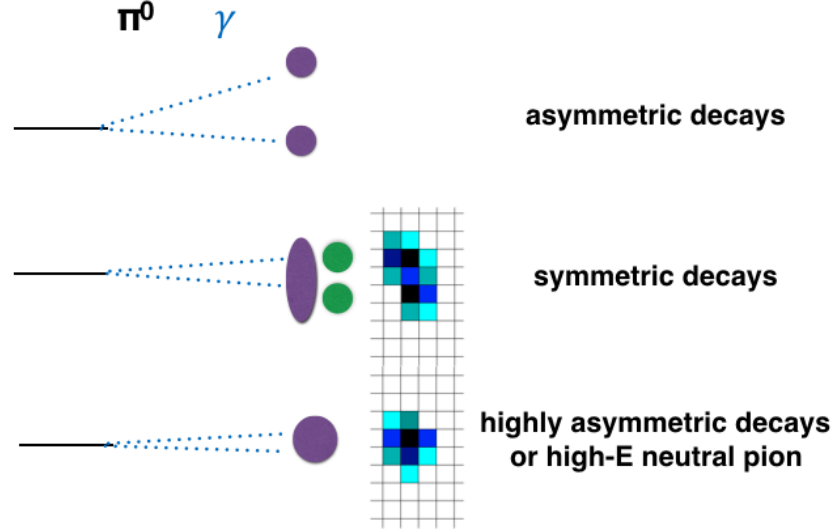


Figure 4.9: Illustration of the effect of both the opening angle and the granularity of the EMCal detector on the clustering procedure and the computed transverse shape for the clusters.

Fig. 4.10 shows the σ_{long}^2 distribution for neutral clusters and its evolution as a function of the transverse energy (E_T); a clear peak is visible around the value 0.27.

Clusters produced by a single particle are the main contributors to this peak, while the bump at larger values of σ_{long}^2 is due to two-particle clusters. As the transverse momentum increases, the broad peak from two-particle clusters moves in the distributions to lower values of σ_{long}^2 , disappearing below the single particle peak for very high E_T clusters.

An E_T -dependent cut on the σ_{long}^2 parameter is applied. The cut values used in this analysis as a function of E_T are presented in table 4.1. Clusters with a computed value of the σ_{long}^2 parameter smaller than those reported in the above-mentioned table will be addressed to as *narrow* clusters, while the rest will be called *wide* clusters.

The reason for the shift, for high energy clusters, of the σ_{long}^2 distribution towards lower values is mostly due to the decay kinematic of π^0 and η mesons:

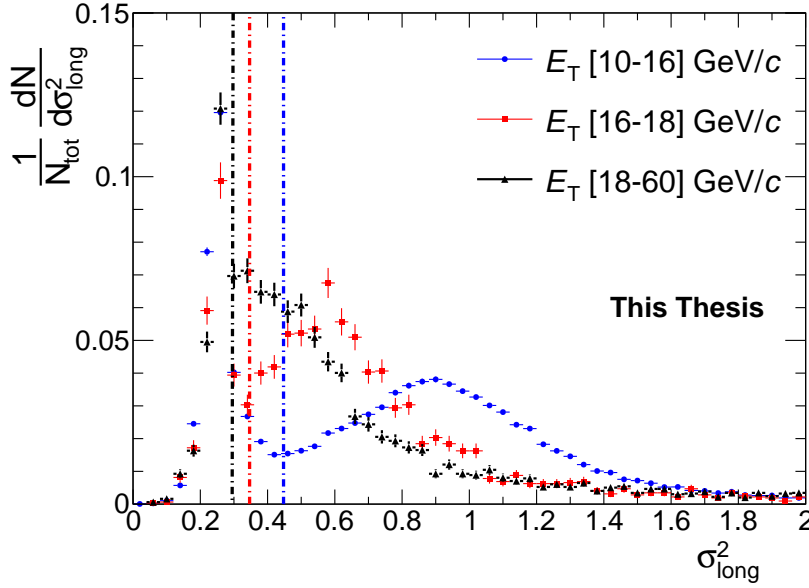


Figure 4.10: Probability distribution of the cluster transverse width σ_{long}^2 for neutral clusters in different bins of transverse energy, reconstructed in pp collisions at $\sqrt{s} = 7$ TeV. The coloured vertical bands correspond to the definition of narrow cluster for the different transverse energy ranges of the clusters.

E_T range (GeV/c)	σ_{long}^2 cut value
10 - 16	0.45
16 - 18	0.35
18 - 60	0.3

Table 4.1: σ_{long}^2 cut values as a function of the transverse momentum of the candidate cluster.

the higher the momentum of the meson, the smaller (on average) the opening angle between the two decay photons. In this case, the resulting cluster obtained via the clusterizer V1 algorithm will be a single, approximately rotationally symmetric cluster.

Moreover, the orientation of the two photons in the rest frame of the

mother particles is randomly distributed. This leads to a broad distribution in energy asymmetry of the two photons; when the decay is too asymmetric (one of the two photons gets the majority of the energy of the parent) or if one decay photon misses the EMCal, the transverse shower shape will be in the range of single photon clusters, drastically reducing the $\gamma - \pi^0$ discriminating power of the σ_{long}^2 variable for those specific cases.

In order to disentangle clusters produced by such configurations from clusters produced by direct photons we need an extra variable.

4.3.2 Isolated photon clusters

In Section 2.3 we have discussed the need for a selection of photon candidates based on the hadronic activity around the photon direction to reject photons produced in fragmentation processes; in the previous section we have shown that it is sometime impossible to distinguish between clusters produced by a direct photon and clusters produced by a decay photon (or, at high energy, even clusters from two decay photons). We expect photons produced via qg-Compton scattering to be surrounded by low hadronic activity (by definition), while both fragmentation photons and photons from neutral mesons decay are produced within a jet-like environment.

In this paragraph we will show how a selection on the total hadronic activity around a photon candidate reduces the contribution from decay and fragmentation photons in our neutral cluster sample. In the following of the thesis we will use isolation activity, in-cone activity and isolation energy as synonyms for the hadronic activity measured inside a circle of radius R around the neutral cluster candidate.

Isolation criterion

From an experimental point of view, any of the isolation criteria defined in Section 2.3 translates into the reconstruction of tracks from charged particles and the determination of their transverse momentum, and the measurement of neutral particle energy deposited in the EMCal. For each of the photon candidates the isolation activity is calculated by adding the transverse

4.4. Contamination estimation for isolated photon clusters

momenta of all charged particles and the transverse energy of all EMCal clusters, in a circular area with radius $R = 0.4$ in η and φ (called *isolation cone*), around the photon candidate cluster.

Using both charged and neutral component of the isolation activity it is possible to reject more efficiently the contribution to neutral clusters. However, in this cases there is also a correlation between the shower shape and the isolation energy, as the energy distribution in the vicinity of the photon candidate is used to determine both. Using only the charged component to measure the isolation activity reduces the correlation, at the expense of a worse resolution on the energy in the isolation cone.

The main result of this analysis has been obtained by considering the total isolation activity, while employing only the charged component will be used as a systematic check.

Fig. 4.11 shows the isolation activity distributions for narrow and wide clusters in data. By comparing the distributions on the left plot (relative to photon-like clusters) with those on the right (relative to π^0 -like clusters) a difference is found only in a few bins at small isolation energy values, where the isolated photon signal is expected.

4.4 Contamination estimation for isolated photon clusters

We have seen in sections 4.3.1 and 4.3.2, how both the transverse shower shape σ_{long}^2 and the isolation activity have some power to distinguish isolated photons from decay and fragmentation photons. Still, it is not possible to obtain a pure isolated photon clusters sample via separate selection on the transverse shower shape and on isolation energy distribution; in fact, due to probabilistic nature of the underlying processes, a significant fraction of the clusters from decay background is expected to be present in the signal area.

Since the two quantities are *ideally* expected to be uncorrelated we can use both variables and apply a *simultaneous single side band method* for the estimation of the residual background in our isolated photon cluster samples.

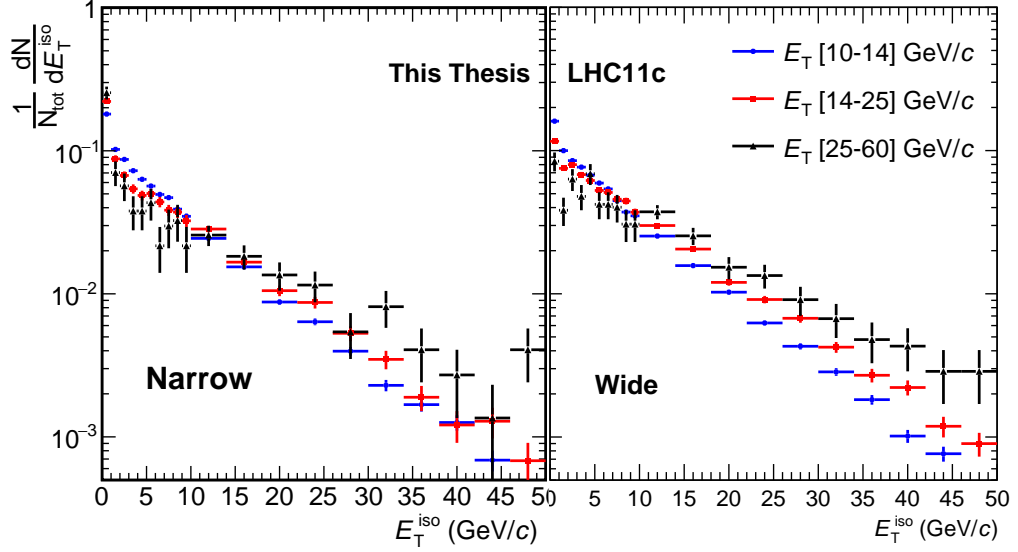


Figure 4.11: Total isolation activity probability distributions for narrow (left) and wide (right) clusters, in 3 different E_T intervals of cluster energy.

4.4.1 The ABCD method

The method mentioned above is commonly referred to as the ABCD method. It is a single side band method applied simultaneously to a pair of properly chosen discriminating variables $Var1$ and $Var2$. The combined use of the two variables defines the phase-space ($Var1 - Var2$ plane) of the system; a selection on these variables results in the division of the plane into four different areas called \mathcal{A} , \mathcal{B} , \mathcal{C} , \mathcal{D} (see Fig. 4.12). One of these region, in our case \mathcal{A} , is the *signal region*, while the remaining regions are called *background* or *control* regions. By applying this *data-driven* method it is possible to estimate, using the measured distributions of the two variable, the residual background in the signal region. In fact, the method uses simultaneously the sidebands (background dominated) in the two variables to determine the background in the signal region, extrapolating the information from the three *control zones*³.

³The underlying assumption in this procedure is that the background has monotonic behaviour with respect to the quantity taken into consideration.

4.4. Contamination estimation for isolated photon clusters

A pair of variables can be used to estimate the contamination via the ABCD method if they fulfil two fundamental requirements:

- both variables must have good discriminating power between signal and background;
- the two variables must be uncorrelated⁴ for the background.

If the last condition is fulfilled, given a background sample (in both variables) it is easy to see that **regardless** of the selection used, the background counts in the four regions satisfy the relation

$$\frac{N_{\mathcal{A}}}{N_{\mathcal{B}}} = \frac{N_{\mathcal{C}}}{N_{\mathcal{D}}} \quad (4.4)$$

This relation can then be used to estimate the number of background counts in the \mathcal{A} region ($N_{\mathcal{A}}$) from those in the control regions:

$$N_{\mathcal{A}}^{est} = N_{\mathcal{B}} \times \frac{N_{\mathcal{C}}}{N_{\mathcal{D}}} \quad (4.5)$$

As long as the contribution of signal counts to the background regions \mathcal{B} , \mathcal{C} and \mathcal{D} is small, the value $N_{\mathcal{A}}^{est}$ obtained using eq. 4.5 is the data-driven estimate of the background contamination in the signal region.

However, it is often impossible to find two totally uncorrelated variables showing a good or perfect separation between signal and background. For this reason a validation of the method and a closure test are performed on a Monte Carlo simulation to ensure a good signal to background separation and the validity of the assumptions.

Monte Carlo simulations can be used to choose the pairs of variables to use for the ABCD method and in defining the regions. Nonetheless, using data-driven methods drastically reduces the dependency of the estimate on theoretical models.

Fig. 4.12 illustrates an ideal case of ABCD plane in which a selection on two generic variables (1 and 2) can be made such that the signal (blue points) is almost exclusively present in the \mathcal{A} region, showing a high level of

⁴Meaning that the distribution of candidates in one cut variable Var1 is independent of the value of the other variable Var2: $P(Var2|Var1 = c)$ is independent of c

correlation between the two variables. On the other hand, the distribution of background counts, in red, fulfils the second requirement for the applicability of the ABCD method.

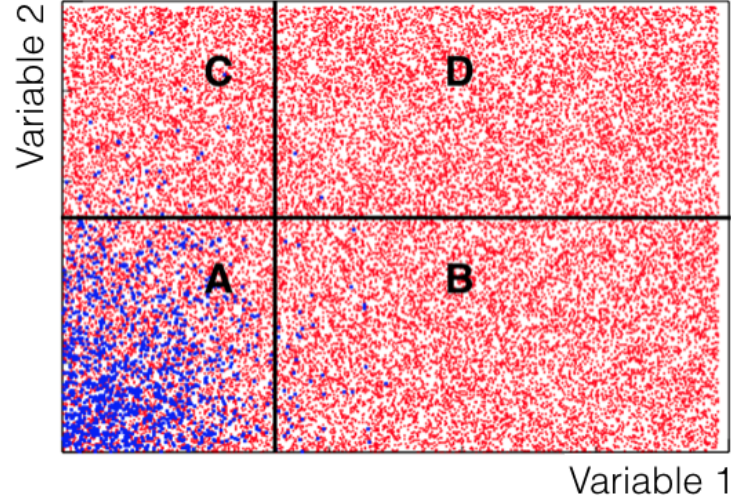


Figure 4.12: Ideal case where the ABCD method yields a correct estimation of the contamination in the signal region.

In the next chapter we will study the ABCD method in simulated events.

Chapter 5

Monte Carlo studies and corrections

In this chapter all checks performed and results obtained by analysing the Monte Carlo (MC) samples available will be presented.

At first we will give some of the properties of the samples used and the steps that lead to their validation; afterwards we will verify the cluster selection procedure applied to the data sample and described in the previous chapter.

We will also show the results on reconstruction and identification efficiency, as well as the results obtained by applying the ABCD method for the estimation of the contamination in the MC samples.

5.1 Simulation samples

To study the validity of the analysis procedure, the signal reconstruction efficiency and to verify the applicability of the ABCD method for the estimation of the contamination, different Monte Carlo (MC) productions, designed to best reproduce the experimental setup available during the period of data taking, have been used.

Events of pp collisions were simulated using an event generator and a detector simulation to produce simulated detector signals that can be analysed in the same way as real collision data. The pp collisions have been simulated

using Pythia 6.4, (Perugia 2011 tune). The detector condition, the triggering algorithm adopted in our period of data taking and the analysis strategy have been either reproduced or mimicked as accurately as possible.

For this analysis, we base our Monte Carlo simulation on PYTHIA minimum bias events, divided into various subsamples, each sampling a specific range in the underlying scale of the hard process. In order to mimic the experimental trigger conditions, the events were selected (before the detector simulation took place) to have at least a photon coming from a π^0 or η meson decay within the EMCAL $\eta - \varphi$ acceptance. The Monte Carlo samples obtained with such configuration are referred to as Jet-Jet (JJ) samples, and are meant to model the background component of the neutral cluster sample.

Two background simulation samples have been produced, both satisfying the condition that a photon from the decay of a neutral meson is emitted within the EMCAL acceptance. These two samples differ for the minimum energy (3.5 and 7 GeV/ c respectively) that the decay photon must have for the event to be reconstructed and stored. The reason for this is to ensure a uniform population of neutral clusters in the σ_{long}^2 parameter distribution, especially for low values of the reconstructed energy of the candidate clusters. In fact, because the measurement presented in this thesis starts at $E_T = 10$ GeV/ c , the candidates from the low-energy triggered simulation accepted in our neutral cluster sample will be produced by photons from highly asymmetric π^0 decays, while the high-energy triggered simulation will contribute to our neutral cluster sample with decay photons from symmetric or slightly asymmetric decay configurations.

In order to model the sample of signal events, i.e. events that contain a prompt photon, and to estimate the efficiency of our measurement procedure, two identical MC simulations have been produced, each anchored to one of the two semi-independent batches of data; in these MC signal samples pp events are forced to interact exclusively via quark-gluon Compton scattering or the annihilation process; the event is selected for reconstruction only if (one of) the photon(s) is produced inside the ALICE-EMCAL $\eta - \varphi$ acceptance. In the following we will refer to the Monte Carlo samples obtained with such configuration as Gamma-Jet (GJ) samples.

It is worth mentioning that those explained above are the only specific

5.2. Validation of the Monte Carlo samples.

settings with which all pp events have been simulated; for all physics processes, like initial and final state radiation, multi particle interactions etc., the default settings from Pythia have been used to better reproduce the conditions of the real hard scattering process.

Furthermore the experimental conditions of the ALICE detectors during the data taking are reproduced by means of a Geant3 model of all the ALICE detectors, not only those described in Ch. 3.

5.2 Validation of the Monte Carlo samples.

To verify that the response of the ALICE detectors to the simulated event is similar to that obtained during the data taking periods, a few basics tests have been performed. Among these, the number of clusters per event, the distribution of Number of Local Maxima (NLM) in a cluster, the charged particle transverse momentum spectrum and the neutral, and charged particle isolation activity are the interesting quantities to validate the simulations with respect to the response of EMCal.

Most of these tests have been performed only on the MC background samples, since these are the cases where the environment is the most challenging for the response of EMCal.

Figure 5.1 shows the $\eta - \varphi$ distribution of the reconstructed clusters in MC; a comparison of the plots in Fig. 5.1 with those in Fig. 4.3 shows no qualitative difference between Data and MC distributions; this confirms that the non functioning cells in data have been correctly excluded when producing the MC samples.

5.3 Validation of the analysis procedure

In this section we will present the results obtained from checking the various assumptions that go into the analysis procedure. The presence of pile-up events which give rise to early or late signals in the readout is not simulated; however, the time selection for the time of arrival of the cluster signal used in the analysis of the data sample is sufficient to reduce the effect

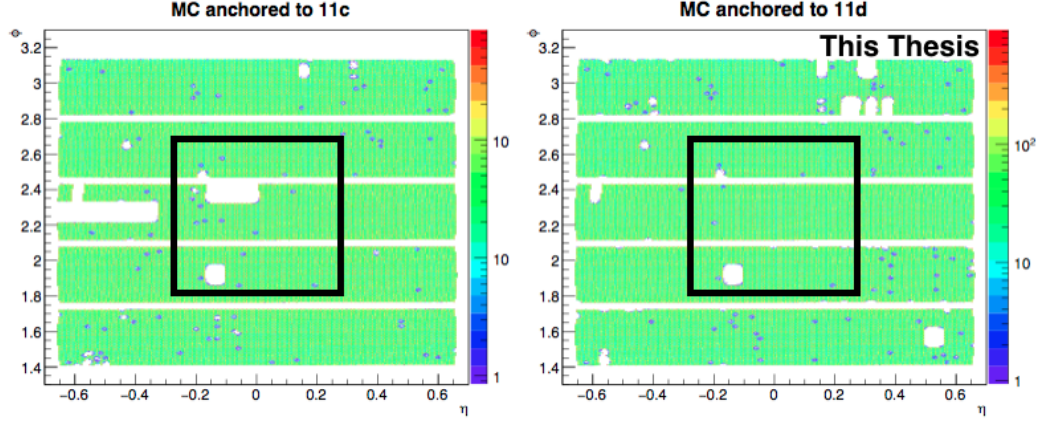


Figure 5.1: Distribution of clusters as a function of η and φ coordinates, computed from the MC samples. On the left the occupancy is obtained from the sample anchored to the LHC11c period, while on the right the same plot is shown for the MC sample anchored to LHC11d.

of out-of-time pile-up.

We will complete the study on pile-up rejection in the next chapter, as a part of the systematic uncertainty determination.

One crucial step in our analysis is the selection of neutral clusters, performed via the CPV algorithm.

Repeating the same procedure on the simulated background samples we find, Fig. 5.2, that the distributions of distance between clusters and the extrapolated position of tracks, at a distance $r = 440$ cm from the centre of the ALICE experiment, shows the same dependence on the distance as those obtained in data and presented in Figs. 4.5 and 4.6.

We apply the selection criteria that are used in the analysis of real data to the neutral cluster sample present in our MC simulations. We reject all clusters that have more than two local maxima; the effect of this cut as a function of the reconstructed transverse energy of the cluster is presented in Fig. 5.3 for the data and MC samples. As one can see, up to 20 GeV/ c the agreement on the fraction of clusters with the same number of local maxima is very good in MC and in data.

5.3. Validation of the analysis procedure

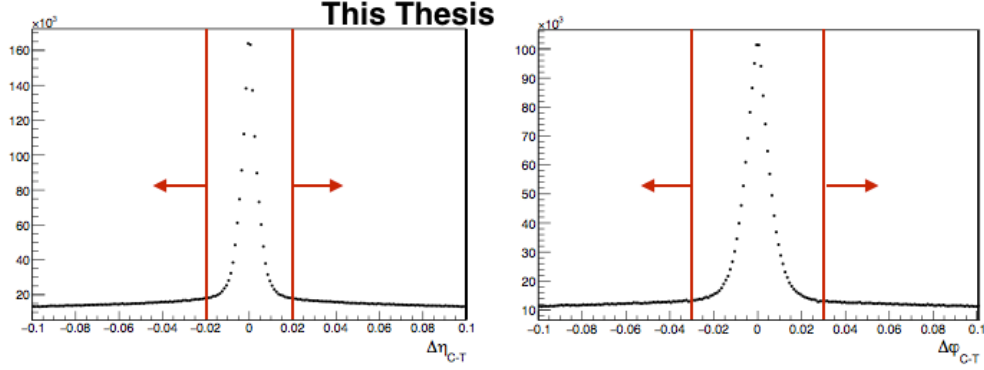


Figure 5.2: $\Delta\eta$ (left) and $\Delta\varphi$ (right) distributions between clusters and closest track as computed from the MC background sample. The distributions computed for data and MC sample drop for similar values of the cluster-track distance.

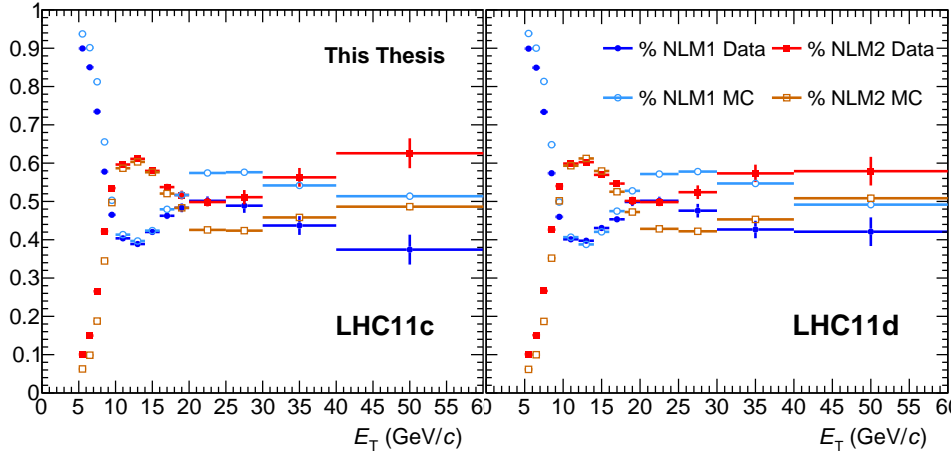


Figure 5.3: Percentage contribution of clusters with 1 (blue circles) or 2 (red squares) local maxima (NLM) in data and MC samples as a function of the reconstructed transverse energy for neutral clusters, for the two periods separately. Full markers represent results from data, while the open, light coloured markers represent the result obtained from analysing the MC sample.

5.4 Reconstruction efficiency for isolated prompt photons

After this initial selection for candidate clusters in our MC samples, we need to answer the questions left open in Section 4.3.1 and 4.3.2. We need to check if our assumptions about the presence of signal within certain ranges of the σ_{long}^2 and in-cone activity distributions are valid and how we can profit from this knowledge to estimate the contamination in our sample of narrow isolated clusters. For this we study the MC signal samples, separately for the two periods and independently from the simulated background samples.

The analysis on the MC signal samples has been carried out following the procedure explained in Section 5.2. Once a cluster produced by an isolated photon is reconstructed as neutral, its properties are stored and studied as a function of the transverse momentum. The transverse width of the reconstructed neutral clusters can allow us to identify single photons while the total isolation activity, computed as

$$E_{\text{T}}^{\text{iso}} = \left(\sum_{i=0}^{N_{\text{clust}}} E_{\text{T}}^{\text{i}} + \sum_{j=0}^{N_{\text{track}}} p_{\text{T}}^{\text{j}} \right) \quad (5.1)$$

in a cone of radius 0.4 around the cluster center allows to reject photons coming both from fragmentation processes and from neutral mesons decay. Eq. 5.1 is the experimental equivalent of the definition we decided to use in chapter 2 for isolated photons. Here we have explicitly split the different contributions to the total in-cone activity: the first term represents the pure neutral contribution, the second indicates the contribution due to well identified charged particles depositing only part of their energy into the electromagnetic calorimeter.

Fig. 5.4 shows the σ_{long}^2 and the $E_{\text{T}}^{\text{iso}}$ distributions for clusters produced by isolated prompt photons for different reconstructed photon energies. All distributions are normalised to their integral, so that it is possible to appreciate how much of the signal would be lost after defining a signal region for each of the two variables.

The cluster width distributions show a sharp peak at low values of σ_{long}^2 and a tail at high values, independent of the reconstructed photon energy.

5.4. Reconstruction efficiency for isolated prompt photons

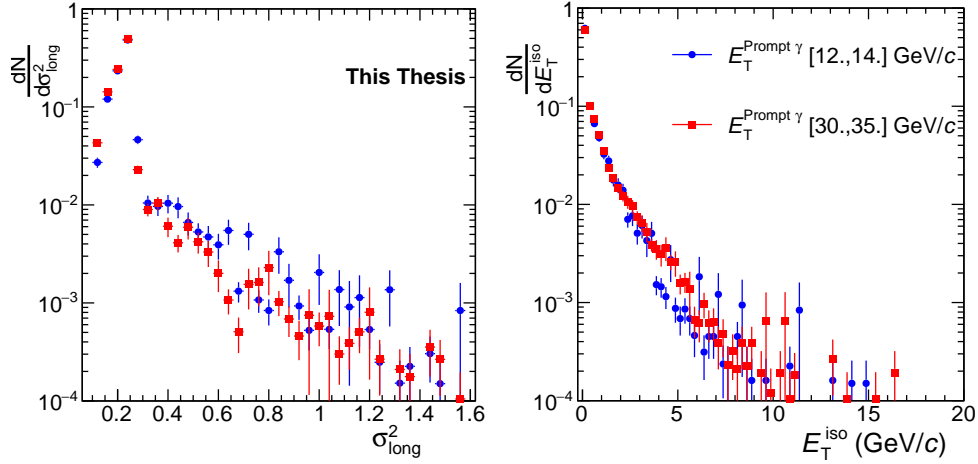


Figure 5.4: (left) Transverse width of isolated photons, studied via the distribution of σ_{long}^2 . (right) Distribution of the total isolation activity E_T^{iso} measured in a cone of radius 0.4 around the isolated photon direction. The distributions, shown for clusters with $12 < E_T < 14 \text{ GeV}/c$ (blue) and $30 < E_T < 35 \text{ GeV}/c$ (red), are from the gamma-jet simulation sample.

This tail is mainly due to the conversions of photons inside the TRD, placed in front of EMCal. When this happens, the signal in the EMCal material comes from the electrons produced in the conversion process.

The isolation activity distributions do not extend above 10 GeV/c , with increasing steepness below 3-4 GeV/c . Figure 5.5 shows the isolation efficiency as a function of the isolation cut value for direct photons with $12 < E_T < 14 \text{ GeV}/c$ and $30 < E_T < 35 \text{ GeV}/c$. The same study has been carried out for all E_T bins in the final spectrum.

We will therefore consider "isolated" all clusters with a total hadronic activity smaller than 2 GeV/c within a cone of radius 0.4 around their direction. This will ensure a high enough identification efficiency for isolated photons, while keeping both the signal leakage into the control region \mathcal{C} and the rejection of background (shown later in the chapter) at an acceptable level. The adopted selection for isolated clusters is also based on the measurement reported in [42], where the production of charged particles associated with the underlying event, performed by the ALICE collaboration, is estimated to contribute to the in-cone activity by $\sim 0.5 - 0.75 \text{ GeV}/c$.

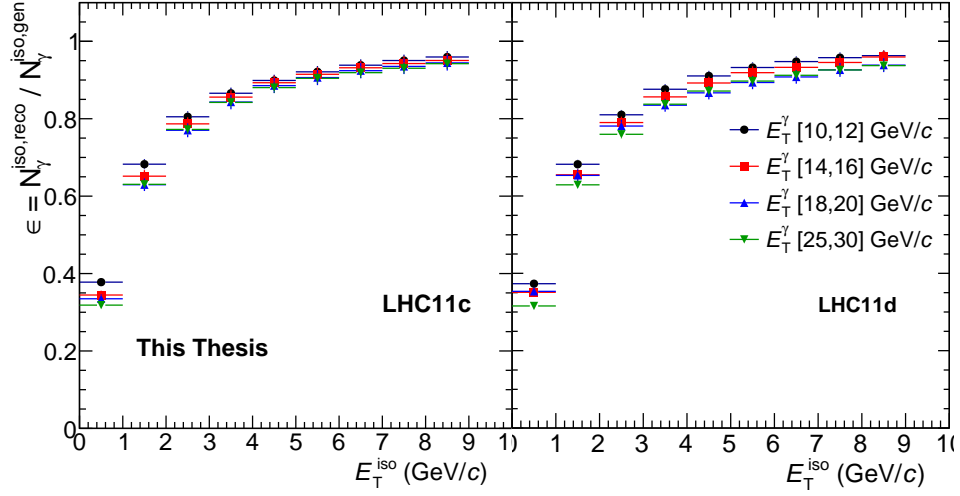


Figure 5.5: Identification efficiency for isolated photons as a function of the reconstructed transverse energy of the clusters. The distributions shown are computed for isolated prompt photons at different transverse energies for the MC-GJ sample anchored to LHC11c (left) and LHC11d (right) separately. The isolation activity distributions have been obtained by adding both the neutral and the charged component of the transverse momentum of particles within the isolation cone.

5.5 ABCD method closure test

In the previous section we have defined range of isolation activity for the signal and verified that our isolated photon sample is well contained in only one region of the $E_T^{\text{iso}} - \sigma_{\text{long}}^2$ plane. In this section we will study the feasibility of applying the ABCD method to estimate the contamination in the signal region.

To do so we need to understand if and where the transverse width distributions computed from our MC background sample are compatible with the same distributions computed in data. This will allow us to finalize the definition of all the regions used by the ABCD method to estimate the contamination in the signal region and to verify whether the in-cone activity and the σ_{long}^2 are indeed two uncorrelated observables for the background.

Fig. 5.6 shows the probability distributions of the transverse width of the

5.5. ABCD method closure test

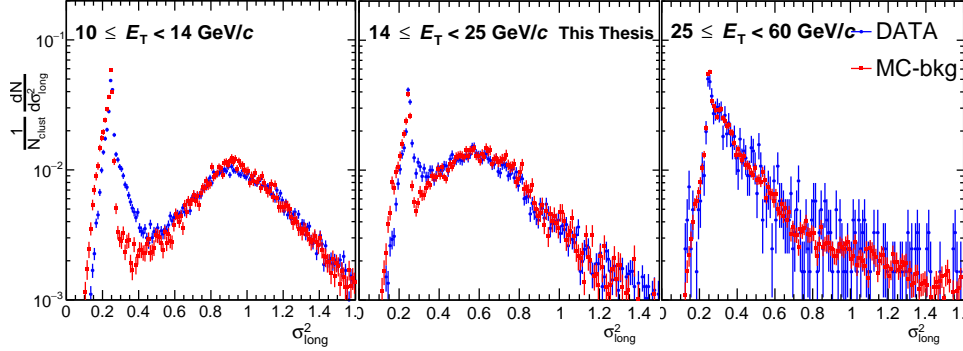


Figure 5.6: Transverse width distributions for candidate clusters computed from data (blue points) and from MC background sample (red points), for three different energies of the candidate clusters. The distributions are computed for the LHC11c period from data and the MC anchored to the corresponding data sample.

candidate clusters, computed for both our data and MC background samples. The two distributions agree quite well both for low and for high values of the σ_{long}^2 parameter while, especially for low values of the reconstructed energy of the candidate cluster, we observe a discrepancy between data and MC at intermediate values of the transverse width of the clusters.

For this reason, starting from the values reported in table 4.1 and considering the fact that our MC sample does not reproduce exactly the shape of the right side of the narrow cluster peak, we decided to introduce a gap of 0.1 between the end of the narrow cluster region and the beginning of the wide cluster region. Furthermore it has been decided to define an upper limit for the wide clusters control regions. Wide clusters will be used in this analysis to estimate the contamination in the signal region. They are defined as satisfying:

$$\sigma_{\text{long}}^2 \in [(\sigma_{\text{long}}^2)_{\text{Narrow}} + 0.1, (\sigma_{\text{long}}^2)_{\text{Narrow}} + 1.1],$$

refining slightly the definition of wide clusters given in Section 4.3.1. Table 5.1 reports the ranges used to define the wide cluster region, as a function of the reconstructed cluster transverse energy. The effect of the mismatch of the σ_{long}^2 distributions in data and simulated sample on the corrected yield will

be explored in the next chapter when determining the systematic uncertainty associated with the selection criterion on the cluster transverse width.

E_T range (GeV/ c)	Low- σ_{long}^2 limit	High- σ_{long}^2 limit
10-16	0.55	1.55
16-18	0.45	1.45
18-60	0.4	1.4

Table 5.1: Limits of the σ_{long}^2 parameter used for the definition of the wide cluster control regions \mathcal{B} and \mathcal{D} , as a function of the reconstructed transverse energy of the cluster.

We can now define completely the four regions of interest of the E_T^{iso} - σ_{long}^2 plane:

- \mathcal{A} Narrow isolated clusters region, where most of the signal is concentrated; this region is also populated by clusters produced by either single decay photons or pairs of decay photons from highly asymmetric neutral mesons decays;
- \mathcal{B} Wide isolated clusters region, mainly populated by clusters produced by pairs of photons from π^0 and η meson symmetric decays;
- \mathcal{C} Narrow non-isolated clusters region, the cluster properties in this region are similar to region \mathcal{A} , while the in-cone activity of the candidates exceeds the threshold value of 2 GeV/ c ;
- \mathcal{D} Wide non-isolated clusters region, same as \mathcal{B} region with in-cone activity above the threshold value as in \mathcal{C} .

The total counts in each region will be indicated with N_n^{iso} , N_w^{iso} , $\overline{N_n^{\text{iso}}}$ and $\overline{N_w^{\text{iso}}}$ for the \mathcal{A} , \mathcal{B} , \mathcal{C} and \mathcal{D} regions respectively, while the background counts in the same region will be indicated with the letter B instead of the letter N .

5.5. ABCD method closure test

Fig. 5.7 shows the distribution of candidate clusters in data within the $E_T^{iso} - \sigma_{\text{long}}^2$ plane and how this is divided into the four regions just defined. The distributions of the isolation activity for neutral clusters are quite broad, showing, in rare cases, that the measured in-cone activity around the cluster candidate is comparable to the transverse energy of the candidate itself; still it is possible to identify a peak at zero GeV/c extending up to 2-5 GeV/c , depending on the reconstructed energy of the candidate cluster.

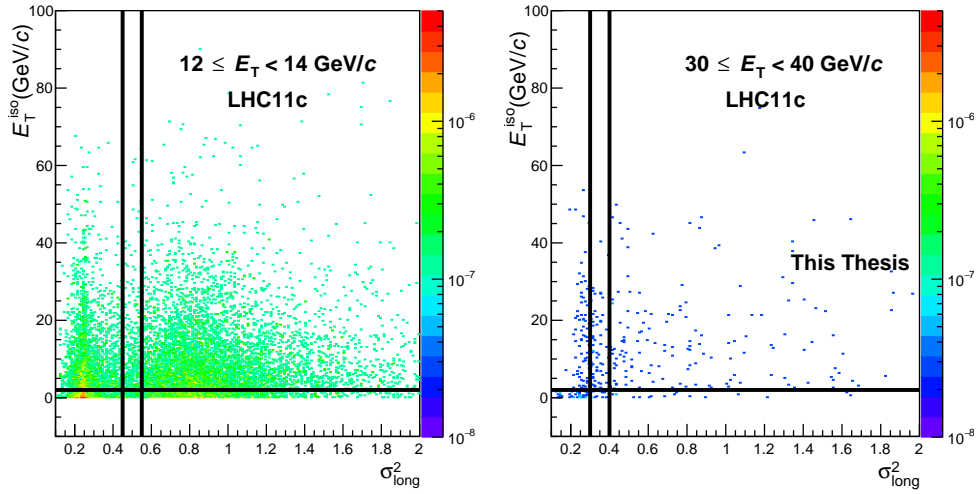


Figure 5.7: Distribution of candidate clusters as a function of σ_{long}^2 and E_T^{iso} , shown for reconstructed transverse energy of the candidate of $12 < E_T < 14 \text{ GeV}/c$ (left) and $30 < E_T < 40 \text{ GeV}/c$ (right).

In order to reduce the leakage of signal counts into the control region \mathcal{C} , one may want to increase the lower threshold in E_T^{iso} to stay safely away from the signal. This would, however, reduce the statistics in this region, resulting in an increase of the uncertainty of the estimated B_n^{iso} .

5.5.1 ABCD method on MC background sample

Fig. 5.8 shows the distribution of candidate clusters from the 2 JJ samples in the $E_T^{iso} - \sigma_{\text{long}}^2$ plane. We present this separately for the sample of events triggered by decay photons at 3.5 GeV/c (left panel) and for the sample of events triggered by decay photons at 7 GeV/c (right panel). We decided to

build our total sample of background events using the low energy triggered sample up to 16 GeV/ c , covering the remaining range of the p_T spectrum with the high energy triggered simulation sample.

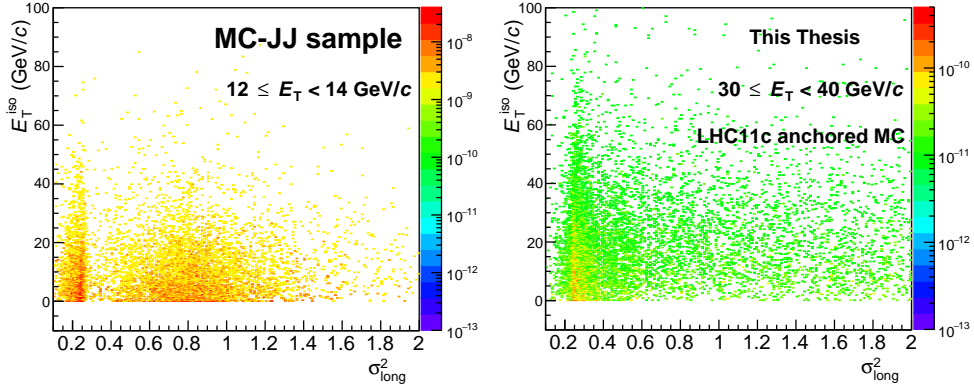


Figure 5.8: Cluster distribution as a function of σ_{long}^2 and E_T^{iso} computed from the background MC sample when selecting candidates in two different transverse energy ranges.

The α_{bkg} factor

To check the performance of the ABCD method on our MC background sample, the transverse momentum distribution of decay photon clusters in the signal region is compared with the estimate given by the ABCD method. Doing so we can check immediately if the total isolation activity and the transverse cluster shape are two uncorrelated observables in our background template. Figure 5.9 shows the result of this comparison. As one can see from the left plot there is a difference between the background estimated via the ABCD method in the signal region and the actual background simulated in the same region. The size of this difference is characterised by means of the factor α_{bkg} defined as

$$\alpha_{bkg} = \frac{B_n^{iso}}{B_n^{iso,est}} = B_n^{iso} \cdot \frac{B_w^{iso}}{B_w^{iso} \cdot B_n^{iso}} \quad (5.2)$$

The dependence of the α_{bkg} factor on the reconstructed transverse energy

5.5. ABCD method closure test

of the cluster is shown in the right plot of the same figure. At low cluster energies the ABCD method overestimates the background, while at high energies it underestimates the background in the signal region.

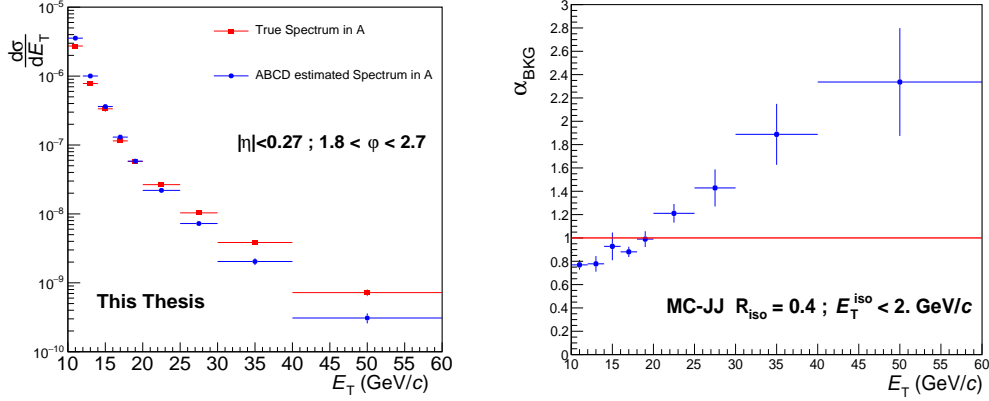


Figure 5.9: Transverse energy distribution of isolated neutral clusters from combined BKG sample (red) and ABCD estimated background spectrum on the same sample (blue)(left panel) and (right panel) ratio of the two spectra from the left panel. The deviation of this ratio from 1 demonstrates the non-factorisation of the σ_{long}^2 parameter and the isolation activity.

Before investigating this further, we are interested in studying how the ABCD method performs when applied to a more realistic sample in which signal and background are mixed.

In fact we know from Section 5.4 that given our isolation criterion, part of the signal leaks (mainly) into the \mathcal{C} region. This can modify the performance of the ABCD method in estimating the contamination in the signal region and for this reasons it is useful to have an estimate of how the contamination estimation changes in presence of signal counts in the control regions.

5.6 A more realistic scenario: the effect of signal

In order to study properly the ABCD method performance and applicability to the MC samples available for our analysis, we need to reproduce the situation in MC and study how the estimation of the contamination in the signal region changes. In the previous section we saw that the ABCD method applied to a pure background sample shows discrepancies between the true and the estimated contamination in the signal region. Still we also saw in Fig. 5.6 that, within the ranges used in the ABCD method, the distributions of the transverse width of the cluster in MC agree reasonably well in shape with the data, in the wide cluster region.

Figure 5.10 shows the distributions of signal (blue points) and background (red points) within the $E_T^{iso} - \sigma_{long}^2$ plane, for candidate clusters with reconstructed energy $18 \leq E_T < 20$ GeV/ c . While the background distribution seems to behave as sketched for the idealised case in Fig. 4.12, the signal is distributed differently, leaking into the \mathcal{C} and \mathcal{B} control regions. To investigate the leakage of signal into the wide clusters region we selected clusters from isolated prompt photon in this region and studied their composition; these clusters are associated to electrons produced via photon conversion in the TRD and TOF, placed in front of EMCal [40].

To produce a realistic sample of neutral clusters, where clusters from both decay and prompt photons are present, GJ and JJ samples are added together, each of them scaled by the σ_{inel} given by Pythia. The validity of this procedure will be demonstrated in the next paragraph, while the effects of choosing different mixture percentages will be studied in the next chapter as a part of the study of the systematic uncertainties related to the MC corrections.

5.6.1 ABCD Method on MC-Mixed Sample: the α factor

Results of the study performed on the MC sample obtained by mixing the signal and background samples are presented in Fig. 5.11, both in form

5.6. A more realistic scenario:
the effect of signal

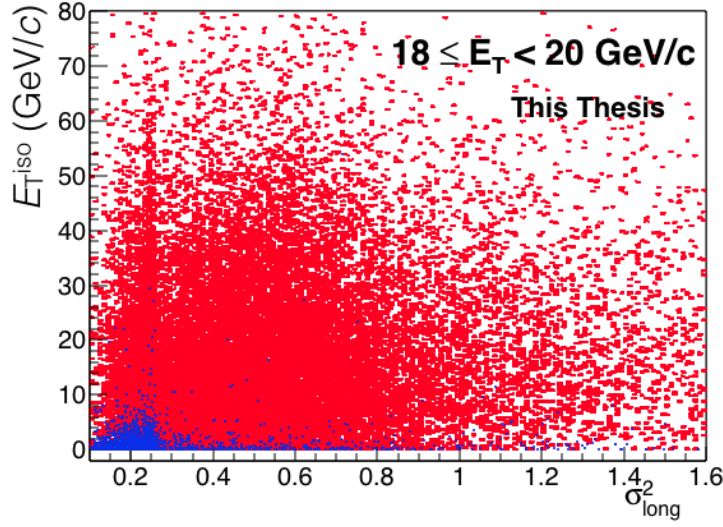


Figure 5.10: Cluster distribution as a function of σ_{long}^2 and $E_{\text{T}}^{\text{iso}}$ for the background sample (red) and for the signal sample (blue). One observes a leakage of the signal into the control regions \mathcal{B} and (mainly) \mathcal{C} .

of the estimated background spectrum and by means of the ratio defined in eq. 5.2, only slightly modified into

$$\alpha = \frac{B_n^{\text{iso}}}{B_n^{\text{iso}, \text{est}}} = B_n^{\text{iso}} \cdot \frac{N_w^{\text{iso}}}{N_w^{\text{iso}} \cdot N_n^{\text{iso}}} \quad (5.3)$$

in order to properly take into account the leakage of the signal into the control regions. By comparing Fig. 5.11 and Fig. 5.9, we can see that the presence of signal in the MC-mixed sample changes the estimation the contamination when using the ABCD method, slightly reducing the deviation of the α factor from the unity.

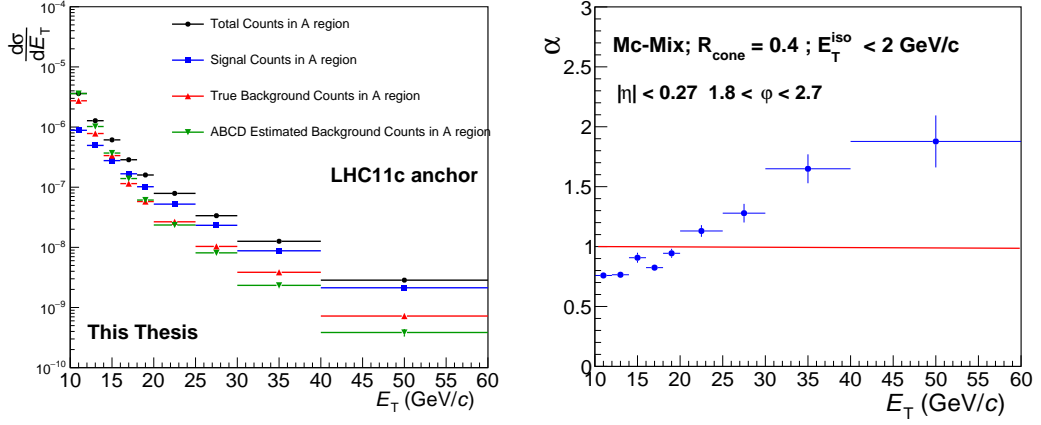


Figure 5.11: Transverse energy distribution of isolated neutral clusters for background (red), signal (blue) together with their sum (black) and the estimated background (green) spectrum in the \mathcal{A} region as a function of E_T (left panel). The right panel shows the ratio between the estimated background and the true background spectrum in the region of narrow, isolated clusters, taken from the right panel. The deviation of this ratio from 1 is smaller than that shown in Fig. 5.9 (b).

Fig. 5.12 shows the uncorrected ABCD-estimated contamination for both data periods and the MC-mixed samples anchored to them. The two estimates from data and MC agree reasonably well with each other. This is a necessary condition for the applicability of a further correction of the estimated contamination via the α factor.

Still, the necessary correction of the estimated contamination with the α factor is an *ad-hoc* correction which does not explain why the two theoretically unrelated parameters used in our method show some degree of correlation.

In the next sections we will present an extensive comparison between data and MC for the quantities of interest in our analysis, trying to shed some light on the causes of the non-factorisation between the transverse shower shape and the in-cone activity for the candidate neutral clusters.

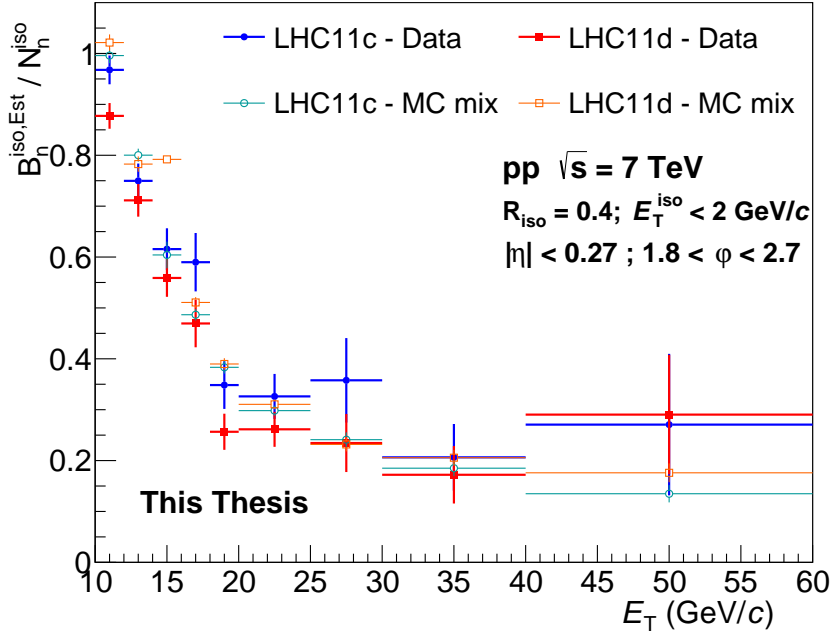


Figure 5.12: Uncorrected ABCD estimated contamination in the region of narrow, isolated cluster for both data periods (blue and red) and the respective estimates from MC sample (dark azure and orange) obtained when combining the signal and the background sample as explained in the text.

5.7 Non factorisation in MC and comparison to data

The conclusions from the studies presented in the previous sections can be summarised as follow:

- The simulated response of the detector partially fails to describe the transverse width of the single photon cluster distributions, showing a peak at low values of σ_{long}^2 which is more narrow than that observed in data.
- The observed deviation of the α factor from 1 is a consequence of the non-factorisation of the two parameters used in the ABCD method, which will therefore need an additional correction.

In this section we will investigate the possible causes that lead to an incorrect estimation of the background in the signal region when using the ABCD method.

5.7.1 σ_{long}^2 distributions

As we saw in Fig. 5.6 (a) and (b), the distribution of the σ_{long}^2 variable show a peak around 0.27 for both data and MC samples. Still the distributions do not agree well for values of the transverse width of clusters slightly above the value of the narrow cluster peak. A direct comparison of the distributions computed from data and simulation (computed by mixing the JJ and GJ events) samples is presented in Fig. 5.13, where we show the σ_{long}^2 probability distributions for clusters, separately for the cases of isolated clusters and for the non-isolated ones. The distributions computed from our simulated sample well reproduce the data distributions in the wide cluster region, while the peak in the narrow cluster region falls more steeply in the simulated sample than in data.

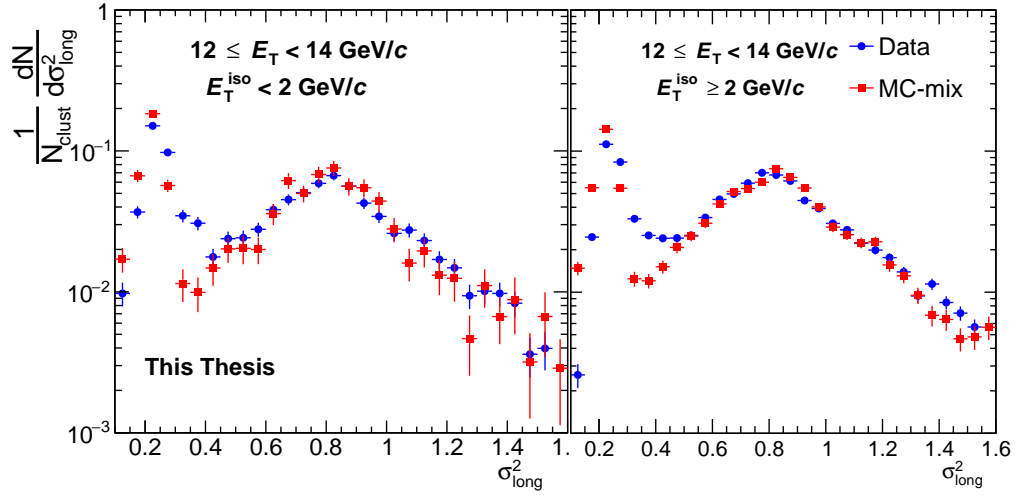


Figure 5.13: σ_{long}^2 probability distributions for isolated (right) and non isolated (left) clusters computed from the data and MC-mixed samples, according to the definition of isolated clusters $\Sigma(E_T + p_T) < 2 \text{ GeV}/c$.

In order to check the possible causes of this discrepancy, we have repeated

the analysis using simulated samples produced by varying, one at the time, various settings of the simulation. We have used a newer version of GEANT¹, for the modelling of the passage of particles through the ALICE subsystems placed in front of EMCal, to test eventual improvements in the description of the interaction of particles with the material; we have studied the σ_{long}^2 parameter distribution as a function of the pseudorapidity of the single isolated photon to study the dependence of the transverse cluster width distributions on the reconstructed position of the cluster within the EMCal supermodules. All these studies have been inconclusive on the cause of the observed difference.

During the final work on this thesis, a cross-talk effect in the read-out electronics was found that may explain the difference between the measured and simulated shower shape distributions. However, the effect is rather complex and has not been fully characterised yet, so for the purpose of this thesis we use an empirical course of action to improve the modelling of the transverse width of the clusters in simulated samples. This consists in convoluting the σ_{long}^2 distributions obtained from our MC sample with a distribution that models additional resolution effects that are apparently present in data, so that the narrow cluster distribution doesn't show the sharply falling peak observed. This procedure is called "smearing" and the corresponding results will be presented in paragraph 5.9.1.

5.7.2 Isolation activity distributions

We can also compare the hadronic activity inside the isolation cone for both our data and MC samples. Fig. 5.14 shows the total isolation activity probability distribution for narrow and wide clusters in data (left) and MC mixed sample (right); the narrow isolated cluster yield (first and second bin) is higher in data than in MC. On the other hand in Fig. 5.15 we compare the total isolation activity probability distribution for narrow (left) and wide (right) clusters via the ratio

$$R_{\text{D-MC}}^{\text{iso}} = \frac{(dN/dE_{\text{T}}^{\text{iso}})_{\text{Data}}}{(dN/dE_{\text{T}}^{\text{iso}})_{\text{MC}}}, \quad (5.4)$$

¹GEANT 4 instead of the original GEANT 3.

in both cases the distributions computed in data are steeper than the corresponding distributions in MC ($R_{D-MC}^{iso} < 1$ for $E_T^{iso} > 10$ GeV/c). The distributions and the ratios shown refer only to clusters with reconstructed transverse energy $E_T \in [12, 14]$ GeV/c, but the same behaviour is observed throughout the whole E_T^{clust} range investigated.

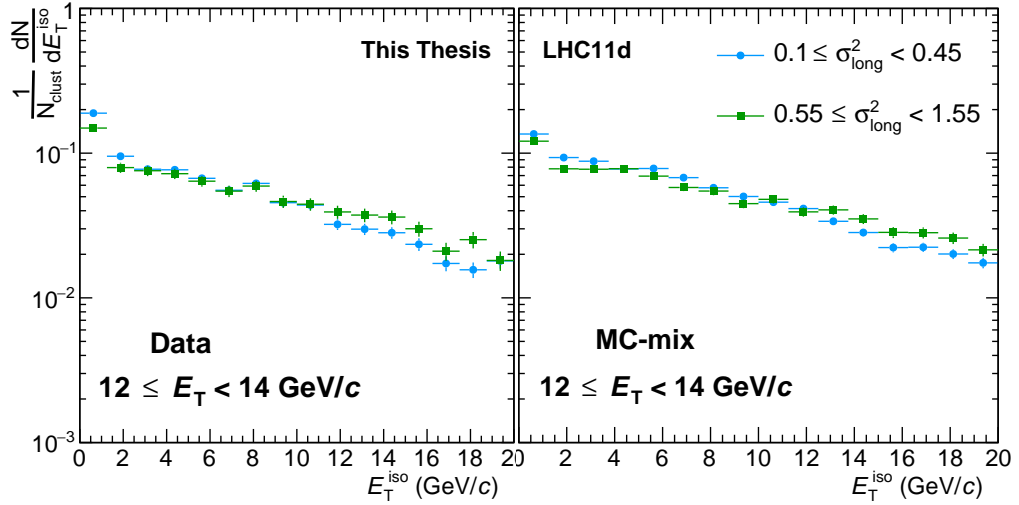


Figure 5.14: In-cone activity probability distributions for narrow and wide clusters computed separately from our data (left) and MC mixed (right) samples.

Ratio Isolated/Non-Isolated in data and MC

In order to illustrate how much the in-cone activity distributions computed in MC differ from those computed from our data sample, Fig. 5.16 shows the ratio between isolated and non isolated clusters as a function of the cluster transverse energy. By applying the same isolation cut value in MC-mixed sample as in data we find that the number of clusters tagged as isolated in MC is smaller ($\sim 30\%$) than in data for the LHC11c period, while for the LHC11d subperiod the discrepancy is $\sim 10\%$.

5.7. Non factorisation in MC and comparison to data

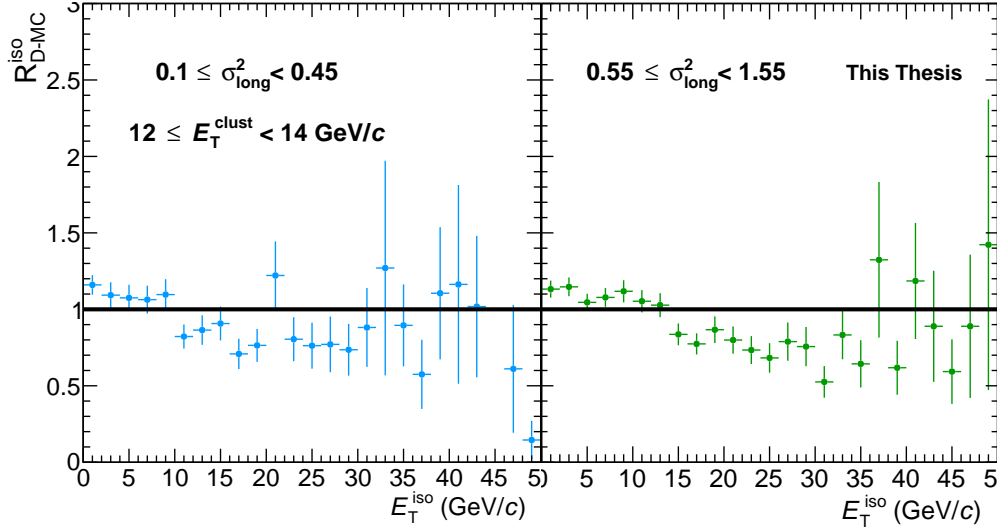


Figure 5.15: Ratio R_{D-MC}^{iso} as defined in Eq. 5.4 computed separately for narrow (left) and wide clusters (right). The ratio has been computed from the corresponding distributions in Fig. 5.14

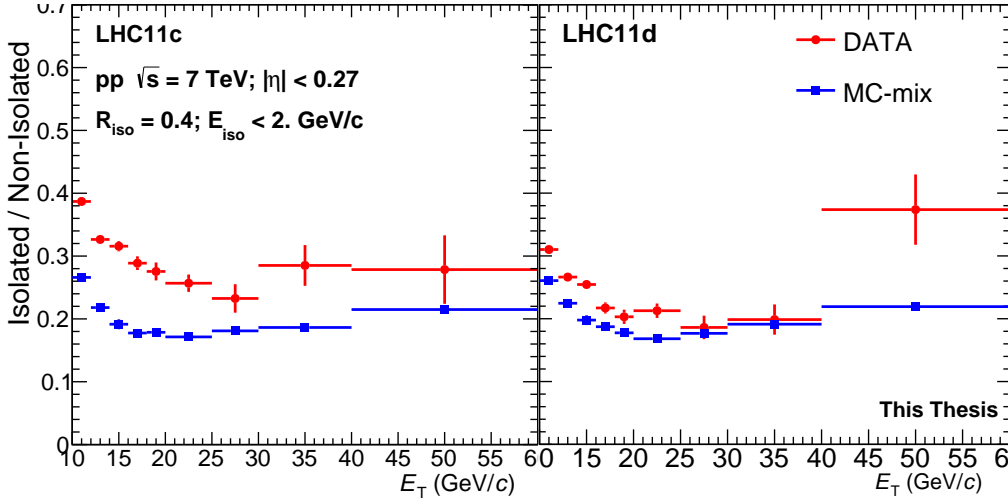


Figure 5.16: Dependence of the ratio $N_{\text{clust}}^{\text{iso}}/N_{\text{clust}}^{\text{iso}}$ on the reconstructed transverse energy of the clusters. The ratio has been computed for the data sample (red) and the MC mixed sample (blue). The ratio is reported separately for the 11c (left) and 11d (right) period.

Double Ratio

We can take a more detailed look at the discrepancies between data and MC by studying the ratio between isolated and non isolated clusters as a function of the σ_{long}^2 parameter. This is shown in Fig. 5.17 for the two reconstructed cluster energies for which the largest discrepancies between data and MC sample are found. We observe a larger number of isolated clusters in data than in our simulated sample; the difference of the ratio $N_{\text{clust}}^{\text{iso}}/N_{\text{clust}}^{\text{iso}}$ between data and MC-mixed sample is largest at low values of the σ_{long}^2 parameter, and it reduces in the wide cluster region.

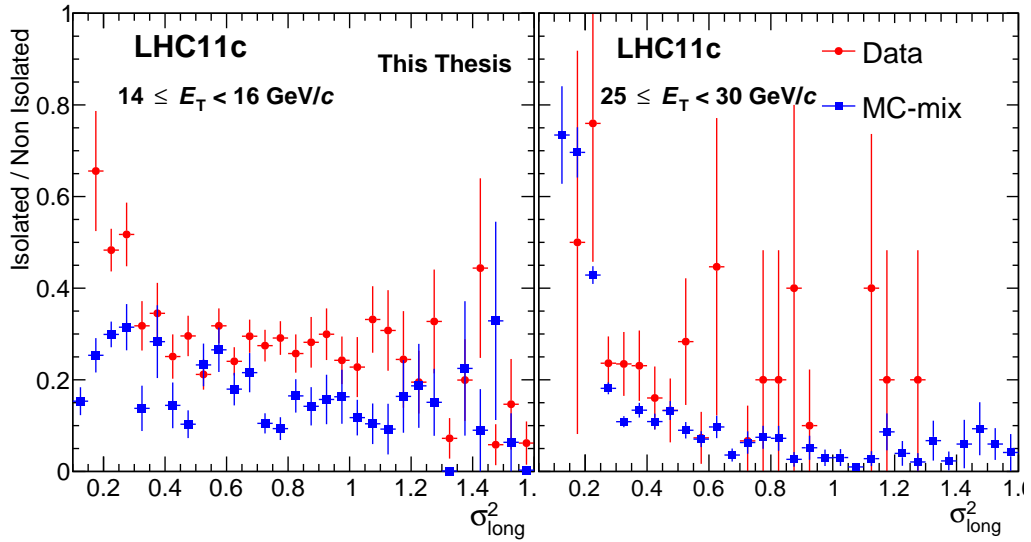


Figure 5.17: Dependence of the ratio $N_{\text{clust}}^{\text{iso}}/N_{\text{clust}}^{\text{iso}}$ on the σ_{long}^2 parameter for two reconstructed transverse energy of the clusters. The ratio has been computed for the data sample (red) and the MC sample (blue). Both plots refer to LHC11c data period.

To gauge directly the difference in the isolation activity distributions computed in data and MC-mixed sample, as a function of the cluster transverse width we can compute the double ratio

$$\frac{(N_{\text{clust}}^{\text{iso}}/N_{\text{clust}}^{\text{iso}})_{\text{Data}}}{(N_{\text{clust}}^{\text{iso}}/N_{\text{clust}}^{\text{iso}})_{\text{MC}}} \quad (5.5)$$

5.7. Non factorisation in MC and comparison to data

as a function of the σ_{long}^2 parameter and look for deviations of the double ratio from 1. Figure 5.18 shows the result of this study for one value of the cluster energy.

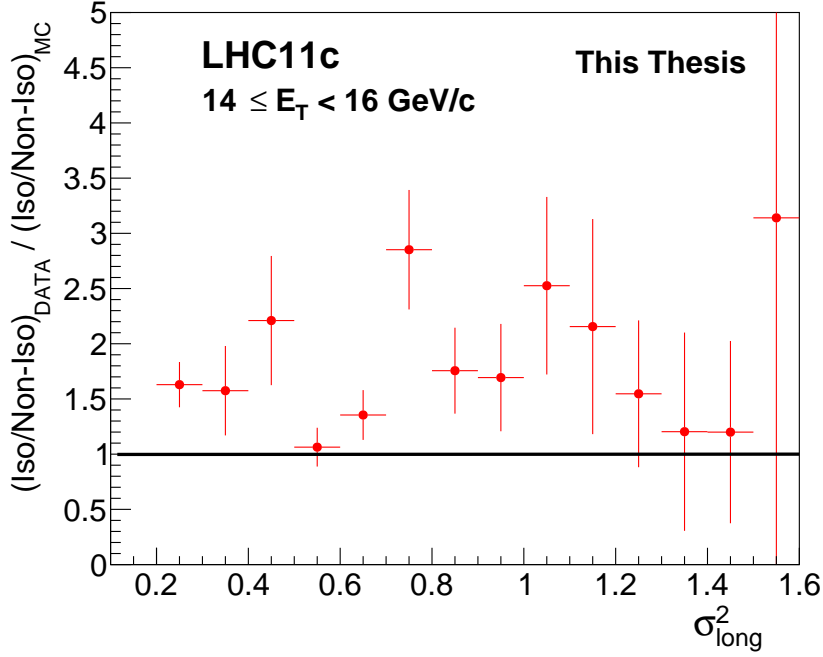


Figure 5.18: Double ratio as defined in Eq. 5.5 as a function of the σ_{long}^2 parameter for one reconstructed transverse energy of the clusters. Note that the first bin is off scale at a value of ≈ 8 .

As we can see, an offset is present between data and MC for all σ_{long}^2 values. The subsample of narrow clusters from data shows a far smaller in-cone activity than the respective MC subsample. We saw in Fig. 5.4 that the isolation probability distributions for isolated prompt photons are sharply peaked at zero and their 99th percentile is 7.5 GeV/ c , giving a limited contribution of signal counts to the N_n^{iso} sample². The cause for the discrepancy between data and MC could possibly be an inaccurate hadronization model used to produce our MC Jet-Jet samples or an inaccurate description of the

²Also the counts from GJ samples contributes to the σ_{long}^2 - E_T^{iso} weighted by the production cross section for isolated prompt photon, which is far smaller than that of JJ events

tracking efficiency, which in turn could result from a non optimal calibration of the TPC, affecting the simulated response of the detector in MC.

The situation is further complicated by a possible dependence of the double ratio shown in Fig. 5.18 on the σ_{long}^2 parameter, which would indicate that not only the probability of tagging a cluster as isolated is different in data and in the MC sample, but also that this difference in probability is changing for narrow and wide clusters, hinting at a non-factorisation of the transverse shower shape and the in-cone activity for candidate clusters.

In the following sections we will first explore the main reason for the observed mismatch between data and MC, and then we will sketch a method that will be used to make the MC adequate for the necessary corrections.

5.8 Non-factorisation of σ_{long}^2 and $E_{\text{T}}^{\text{iso}}$.

In order to proceed with the measurement of the isolated photon production cross-section we need to understand the causes of the non-factorisation, which is already present in a pure background sample.

One of the reasons is inherent to the method used for this measurement. The majority of the wide clusters in our sample is produced by both photons from neutral mesons; as a consequence, the reconstructed energy for such clusters will be (very close to) that of the mother particle. On the contrary, the narrow cluster subsample will consist of clusters produced by only one of the two decay photons³. In this case and for such values of transverse energy it is likely that the decay partner of the reconstructed candidate photon falls within the isolation cone of the candidate; as a result, the in-cone activity computed for such clusters can be larger than the in-cone activity measured for wide clusters of the same energy.

It has been shown in [43] that through an in-cone invariant mass analysis it is possible to partially correct for such a phenomenon; the correction brings a minor improvement in the non-factorisation, and will not be applied in this analysis.

³An exception to this is the case in which the detected photons produce one narrow clusters as the results of very asymmetric decays, for which one of the two photons carries the majority of the energy of the mother particle.

The studies on the EMCAL resolution are reported in [37]. These studies refer to single particle measurements (mainly electrons) performed during different test beams, where the experimental conditions are different compared to the physics measurements. We are now interested in the reconstruction performance for decay photons within a jet-like environment, as a function of the transverse energy of the neutral meson, for different decay kinematics. Figure 5.19 shows the distribution of clusters as a function of cluster E_{T} and transverse width for different transverse energy of the neutral pions reconstructed and identified within the JJ MC simulations used for this analysis.

As one can see, the narrow cluster region shows the expected behaviour: the measured cluster energy is smaller or equal (within the EMCal resolution) to that of the neutral pion, depending on the degree of asymmetry of its decay. On the other hand the wide cluster region shows an unexpected feature. In fact, in rare cases, the reconstructed cluster energy is twice as large as that of the original π^0 meson. This introduces an effect of **bin migration**, which means that, for a given reconstructed energy interval, some of the clusters population will consist of clusters produced by both decay photons from a π^0 (wide cluster region) and other particles not vetoed by the CPV algorithm. Together with this contribution of ill-defined clusters, the wide cluster region will also be populated by clusters produced by decay photons from symmetric decays of neutral pions with the same energy as that of the cluster; while the narrow cluster sample will instead consist of clusters produced by one decay photon from a higher energy neutral pion and clusters produced by both decay photons from highly asymmetric decays.

Our estimation of the contamination in the signal region by means of the ABCD method will hence be affected "by construction" by the different contributions to the cluster distribution, which may show different relations between shower shape and isolation energy. Still, the same feature is expected in our data sample, for which the clustering algorithm is the same as the one used when reconstructing the simulated EMCal response in MC.

In spite of the discrepancies observed between data and MC, many properties of the data are correctly described by the simulation, in particular the uncorrected contamination as shown in Fig. 5.12. On the basis of this the

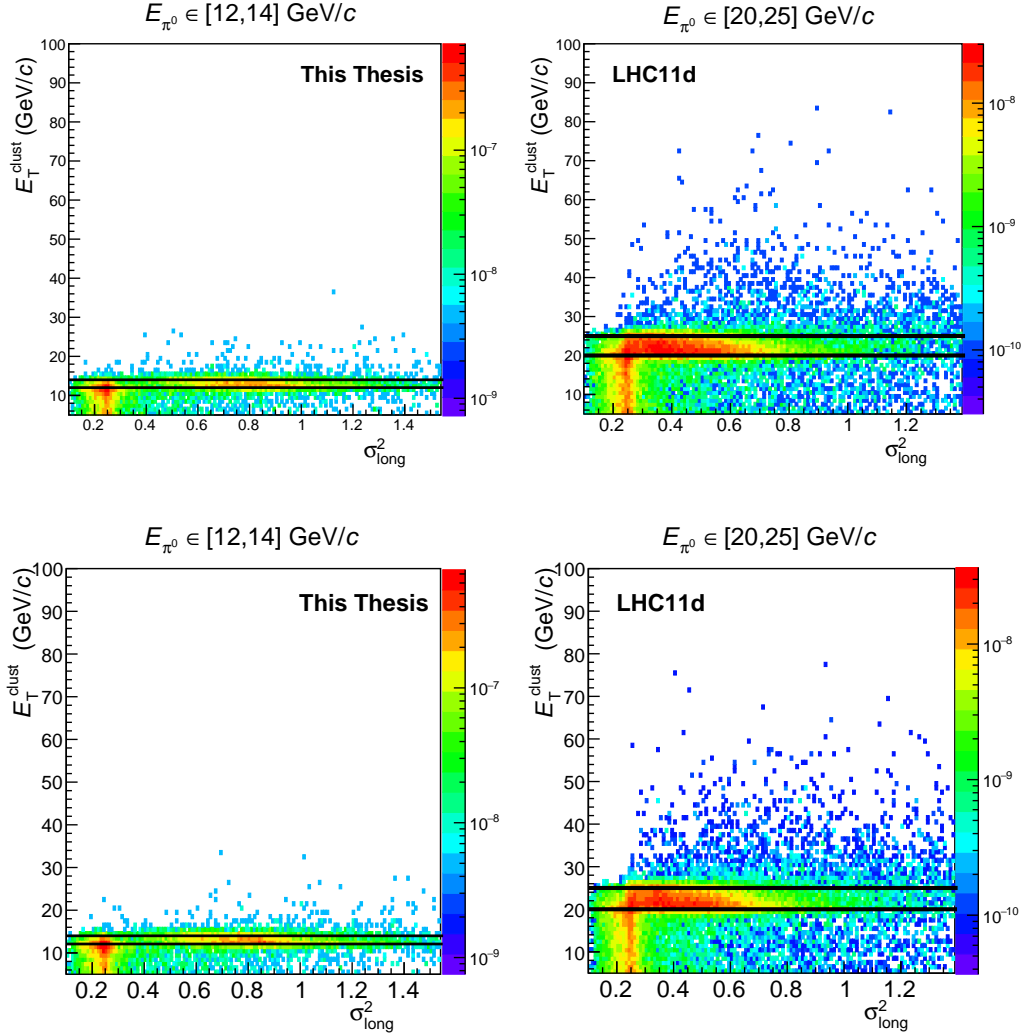


Figure 5.19: Reconstructed cluster energy as a function of the transverse shower shape parameter for two different bins of the neutral pion energy (black lines).

use of a simple calibration factor α described in Section 5.6.1 obtained from MC for the contamination is justified.

In the next chapter the use of the α factor to correct the estimated contamination in data will be proven further, showing its stability as a function of the signal region definition (both for the σ_{long}^2 and the E_T^{iso} distributions) and as a function of the weights applied to our signal and background sample

used while building the MC-mixed sample.

5.9 "Correcting" the MC

In the previous section we have shown that both the transverse shower profile of clusters and the in-cone activity computed in MC (mixed or not) samples differ substantially from the distributions found in data sample and presented in sections 4.3.1 and 4.3.2.

In the next paragraphs we will present the two methods adopted to correct for the discrepancy between data and MC samples.

5.9.1 Smearing of σ_{long}^2 distributions

To solve the observed discrepancy in the modelling of the transverse shower shape in our simulated sample compared to data, we have used a heuristic approach. The goal is to reproduce the tail, present only in data, at the right side of the peak of narrow clusters. The idea is to modify the σ_{long}^2 distribution to resemble more the one obtained from our data sample. As already said, this procedure consists in convoluting the σ_{long}^2 distributions with a function that models additional fluctuations that are apparently present in data.

In order to reproduce these resolution effects, we have tried convoluting the transverse width distributions obtained from the simulated sample with different functional forms; here we will present only the results obtained with the one that provides the best agreement with data.

We use a Monte Carlo convolution method, which consists in extracting a value in the range $[0, 2]$ according to a Landau probability density function:

$$p(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} e^{s \cdot \ln s + x \cdot s} ds \quad \text{with } s = \frac{x - M}{W}, \quad (5.6)$$

where M is the mean of the distribution and it is set to 0, while W is the width of the distribution and it is set to 0.005; we then add this quantity to the computed σ_{long}^2 for each cluster with 1 local maximum. This procedure is used in all MC samples and for all runs anchored to both data periods.

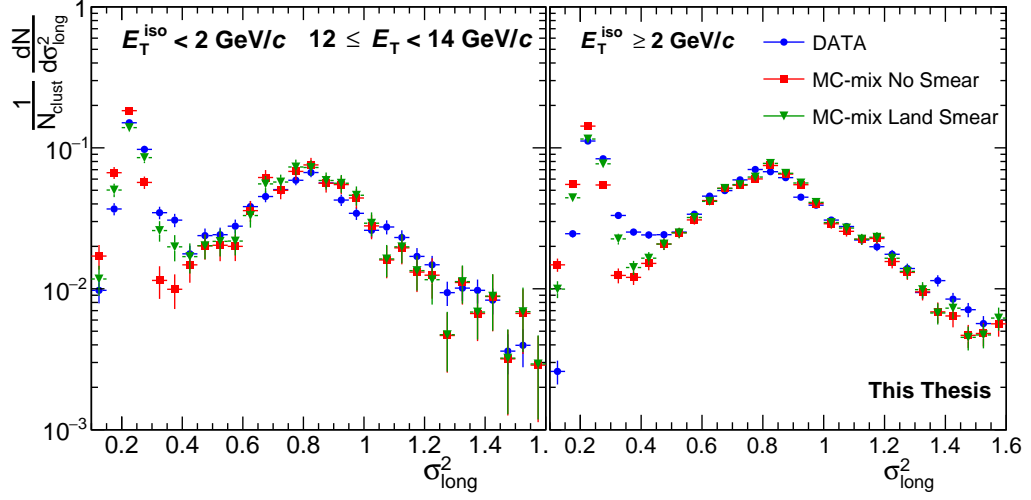


Figure 5.20: σ_{long}^2 probability distributions computed from the data sample (blue) and from the MC sample with (green) and without (red) smearing of the σ_{long}^2 parameter. The left panels show the distributions for isolated clusters while the right panels show the same distribution for clusters with a total activity in their isolation cone greater than 2 GeV/c.

Fig. 5.20 shows the effect of the smearing procedure on the σ_{long}^2 probability distributions computed from the MC-mixed sample and the comparison with the distributions computed from the data sample. A clear difference between the two MC distributions is visible, and the added smearing improves the agreement of the MC distributions compared to those from the data sample.

As a consequence of this smearing procedure, also quantities like efficiency and estimated contamination, which rely on the transverse shower shape discriminating power, might change.

Fig. 5.21 shows the reconstruction and identification efficiency, defined as

$$\epsilon = \frac{N_{\gamma}^{\text{iso, reco}}}{N_{\gamma}^{\text{iso, gen}}} \quad |\eta| \leq 0.27 \quad 1.8 \leq \varphi \leq 2.7 \quad (5.7)$$

as a function of the cluster transverse momentum for two different modelling of cluster transverse shower shape. The single photon (or narrow

5.9. "Correcting" the MC

cluster) definition used is the same reported in table 4.1 while the isolation criterion considered is $\Sigma(E_{T,i} + p_{T,i}) < 2 \text{ GeV}/c$ in a cone of radius 0.4, and applied both at reconstructed and generated level. By shifting counts from the peak region to higher values of σ_{long}^2 while keeping the same narrow cluster definition, we loose some identification power with respect to the non-smear scenario.

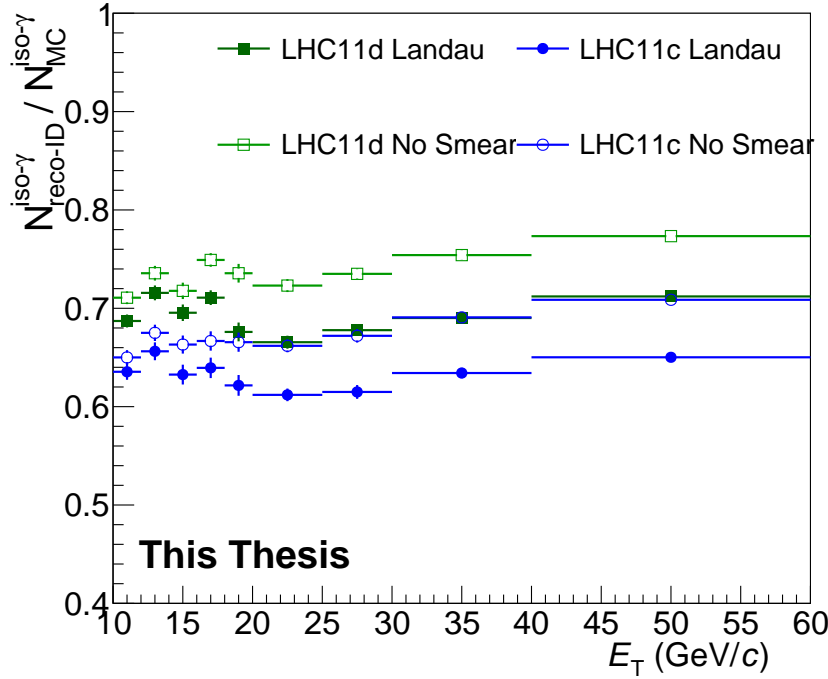


Figure 5.21: Reconstruction and identification efficiency for isolated photons as a function of the transverse momentum for two different modelling of the transverse shower shape. The spectrum of photon which verify the isolation criterion is then compared to the spectrum of the generated isolated photon (in the same $\eta - \varphi$ region).

As shown in paragraph 5.7.2 we can study the isolation probability as a function of the σ_{long}^2 parameter by means of the ratio $N_{\text{clust}}^{\text{iso}}/N_{\text{clust}}^{\text{iso}}_{\text{MC}}$; moreover we can estimate the discrepancy between data and MC via the double ratio and its dependence on the transverse shower shape. Figure 5.22 shows the double ratio defined in Eq. 5.5 computed after the smearing is applied to the σ_{long}^2 distributions. By comparing the plots in Figs. 5.18 and 5.22 we

can appreciate how the smearing procedure reduces both the offset initially present in the double ratio throughout the whole σ_{long}^2 range.

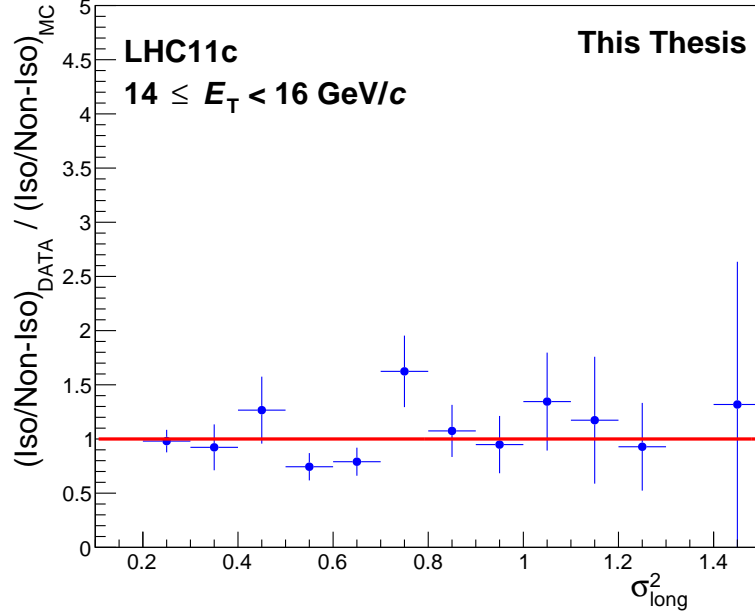


Figure 5.22: Double ratio, as defined in Eq. 5.5, as a function of the σ_{long}^2 parameter, computed after applying a Landau smearing to the σ_{long}^2 distributions in our MC samples.

We can now check how the estimation of the contamination in the signal region is affected by the smearing procedure. In Fig. 5.23 we show the results of the contamination estimated with the ABCD method with and without using a smeared σ_{long}^2 probability distributions, together with a comparison of the true contamination injected into the MC-mixed sample. Despite the changes in the identification efficiency and in the double ratio, the smearing has little effect on the α factor, thus helping only with a better modelling of the σ_{long}^2 distribution from MC, compared to those in data.

5.9.2 Same rejection in MC as in data

In the previous section we have studied the effect of an heuristic attempt to correct the σ_{long}^2 probability distributions computed from our MC sample.

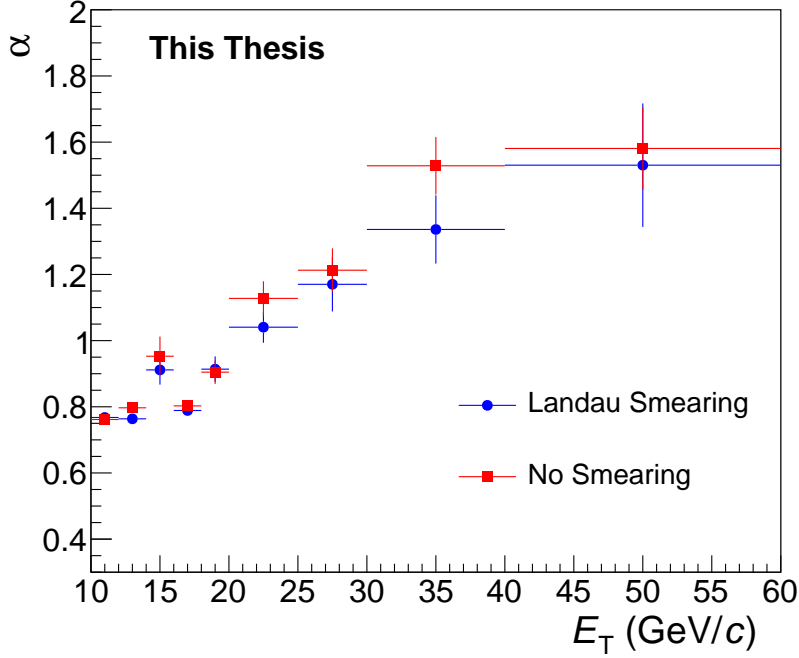


Figure 5.23: Calibration factor α computed after applying a Landau smearing to the σ_{long}^2 probability distributions in our MC samples anchored to LHC11c period.

In Fig. 5.5 and in Fig. 5.21 we have seen that by defining as **isolated** clusters with an in-cone activity smaller than 2 GeV/ c ensures an efficiency, defined in Eq. 5.7, of $\approx 65\%$ for narrow clusters, almost constant throughout the studied E_T range. Moreover Fig. 5.15 shows that the in-cone activity distributions differ between data and MC.

Although the origin of this mismatch is not completely understood, we can try to correct for the discrepancy found by slightly modifying our analysis strategy on the simulated sample in order to have the same amount (in percentage) of clusters failing the isolation criterion in data as in MC⁴. We refer to this procedure as **same rejection criterion**.

We first compute the isolation energy distributions of narrow and wide clusters in data; from these we calculate the percentage of cluster rejected as

⁴It is worth noting that given their shape, it is not possible to apply such algorithm to the σ_{long}^2 probability distributions without biasing the definition of wide (background) clusters, which is where the MC sample well models the distributions obtained in data.

non-isolated when a cut on the allowed total in-cone activity of 2 GeV/ c is applied, we then calculate which value of the allowed total in-cone activity gives the same percentage of non-isolated clusters in our MC mixed (GJ + JJ) sample. This procedure is performed differentially in E_T^{clust} ; in order to have an estimation of the percentage of rejected clusters in MC as precise as possible we vary the isolation criterion in steps of 0.25 GeV/ c .

Table 5.2 lists the values of the in-cone activity for which the same rejection is achieved, as a function of transverse energy of the cluster, while Figs. 5.24 and 5.25 show how the single and double ratio change after applying the Same Rejection algorithm to our analysis strategy. The obtained agreement between the isolation probability in data and MC is at the per mille level.

E_T range (GeV/ c)	E_T^{Iso} cut (GeV/ c)		N^{Iso} in MC (%)		Data – MC diff. ($\cdot 10^{-3}$)	
	LHC11c	LHC11d	LHC11c	LHC11d	LHC11c	LHC11d
10-12	3.00	2.50	71.58	75.79	5.6	5.6
12-14	3.00	2.50	76.12	78.47	7.3	4.0
14-16	3.50	2.75	76.26	79.80	0.8	3.5
16-18	3.25	2.50	78.23	81.61	3.7	4.2
18-20	3.50	2.25	77.96	83.71	0.9	0.1
20-25	3.25	2.75	79.97	82.36	1.5	2.1
25-30	3.00	2.50	81.19	83.04	1.9	1.3
30-40	3.75	2.25	78.49	82.63	3.5	2.6
40-60	4.00	5.25	77.22	72.68	1.9	0.5

Table 5.2: E_T^{iso} values, computed percentages of non-isolated clusters and difference Data-MC computed with the Same Rejection algorithm as a function of the transverse momentum of the candidate cluster. These values ensure the most similar amount of rejected clusters in MC when using a fixed value of 2 GeV/ c for the "anti-isolation" criterion in data.

5.9. "Correcting" the MC

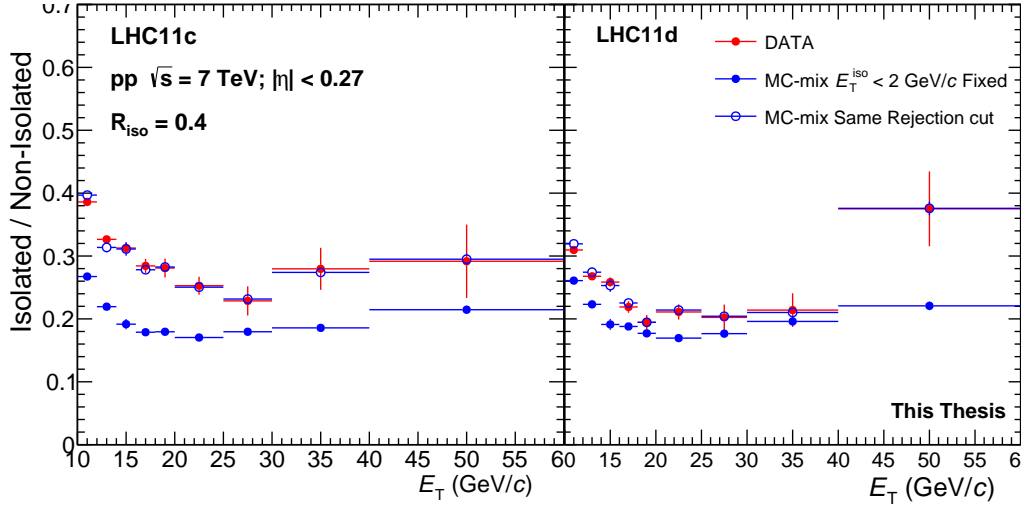


Figure 5.24: Dependence of the ratio $N_{\text{clus}}^{\text{iso}}/N_{\text{clus}}^{\text{iso}}_{\text{bar}}$ on the reconstructed transverse energy of the clusters. The ratio has been computed for MC mixed sample after applying the Same Rejection algorithm to the MC mixed sample (open blue circles). For comparison, the curves reported in Fig. 5.16 are also shown. The left panel refers to the LHC11c period while the right panel shows the results obtained from the LHC11d period.

The good agreement is also confirmed in the comparison in Fig. 5.24 where the ratio $N_{\text{iso}}/N_{\text{iso}}_{\text{bar}}$ is plotted for data and the MC sample when the same isolation criterion as in data is applied and with the isolation criterion computed with the same rejection algorithm. Still, we need to check what the effect of this procedure is on both the efficiency, since we are relaxing our definition of "isolated" photon, and on the non-factorisation of the two variables used with the ABCD method. Figs. 5.26 and 5.27 shows the results of these studies; the efficiency increases as a consequence of the relaxation of the isolation criterion, while the α factor, although somewhat modified, is still not compatible with one.

We showed that while both the smearing of the transverse shower shape distributions and the same rejection algorithm allows us to reach a reasonable level of agreement between the corresponding distributions in data and MC samples, neither of these methods solve completely the non factorisation between σ_{long}^2 and in-cone activity distributions.

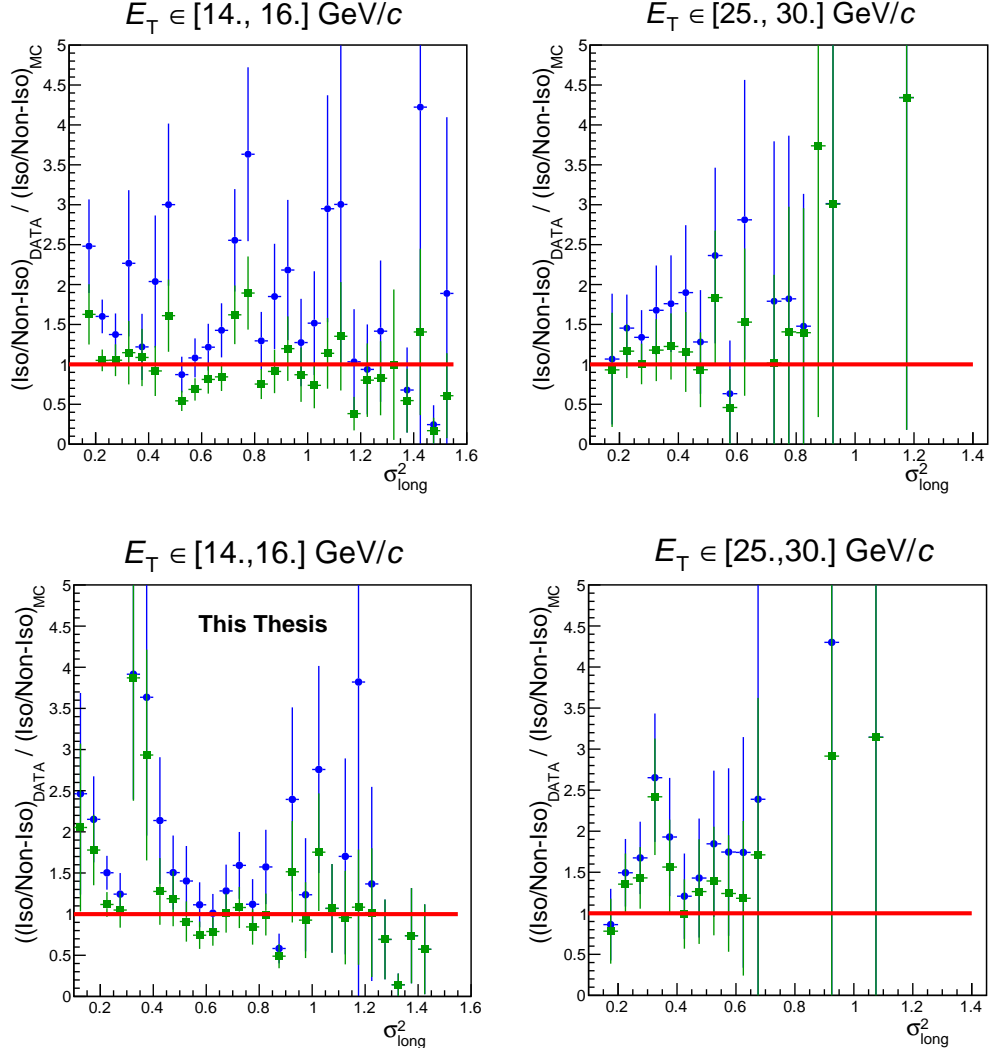


Figure 5.25: Double ratio, as defined in Eq. 5.5, as a function of the σ_{long}^2 parameter computed for a fixed value of the isolation criterion (blue points) and with the isolation criterion given by the same rejection algorithm applied on our MC mixed sample (green points). The double ratio is shown for two values of cluster transverse energy. The two panels show the results obtained from the LHC11c (left) and LHC11d (right) period separately.

Still, the same clustering algorithm has been used to reconstruct our data and MC samples, equally affecting both samples with the bin migration effect shown in Section 5.8. Proof of this is the uncorrected contamination, shown

5.9. "Correcting" the MC

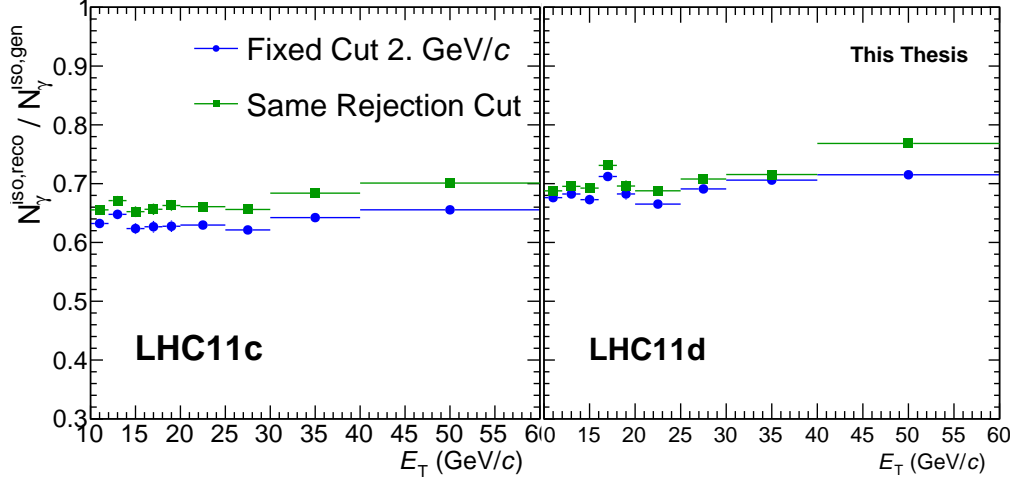


Figure 5.26: Reconstruction and identification efficiency after applying the same rejection algorithm to the E_T^{iso} probability distributions in our MC mixed sample.

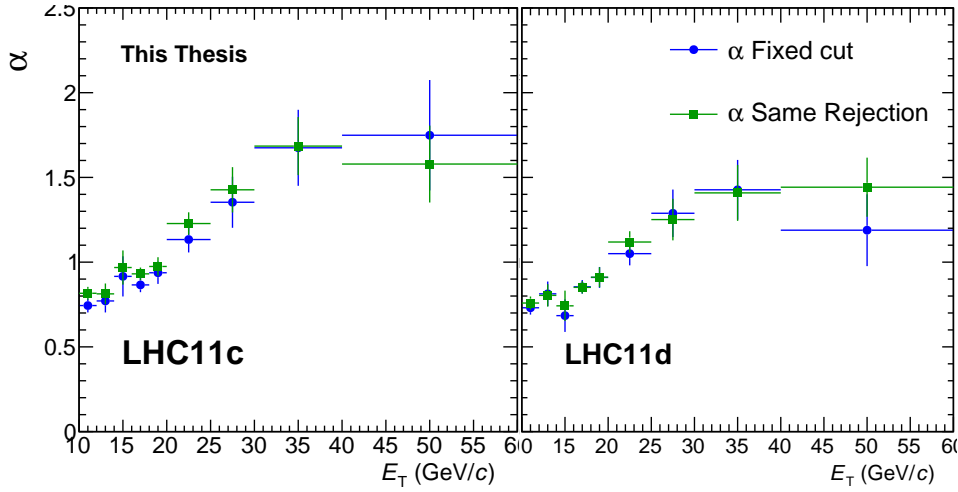


Figure 5.27: Calibration factor α computed after applying the same rejection algorithm to the E_T^{iso} probability distributions in our MC mixed sample.

in Fig. 5.12, obtained from our MC-mixed sample which is compatible, throughout the whole E_T range investigated, with that obtained from our

data sample.

Chapter 6

Corrected yield and systematic uncertainties

In the previous chapter we have studied the contamination estimation procedure by means of the ABCD method. A correction for the non-factorisation is needed. This correction is implemented via the α factor which is determined from simulated events. The background is first estimated by applying the ABCD method to the data, and this estimate is multiplied by α to correct for the fact that the shower shape and the in-cone activity show some degree of correlation.

We showed how the cross talk between the two variables is partially due to the choice of variables used for the estimate of the contamination via the ABCD method and partially a feature of the clustering algorithm. Since the same clustering algorithm and the same definition of in-cone activity are used in the data analysis, we concluded that a correction of the contamination estimate in data is indeed needed. In this chapter we will study the stability of the calibration factor α with the variation of all quantities related to the contamination estimate in MC, further justifying the use of this MC correction when analysing our data sample.

We also showed the differences, concerning the σ_{long}^2 and in-cone activity distributions, between our data and simulated event sample available; we presented the results of the smearing method, used to improve the modelling of the σ_{long}^2 distributions in simulated event samples, and the results

obtained by modifying the isolation criterion according to the same-rejection algorithm, which we used to reject non-isolated clusters from the mixed (GJ and JJ) sample of simulated event.

Both approaches seem to reduce the strength of non-factorisation of the two variables used in the ABCD method, but neither of them helps solving the problem entirely.

As discussed, the data-MC differences for the in-cone activity distributions could be due to the accuracy of the parton shower and hadronisation modelling in the simulation.

No explanation has been found yet for the data-MC difference for the σ_{long}^2 distributions, despite the same discrepancy has been observed in different data samples and the respective simulated event samples, independently of the collision system and the energy of the collision. For this reason, a modification of the MC modelling for transverse energy distribution of clusters in simulated events is needed. We will therefore use the results on efficiency and non-factorisation obtained with the smearing of the σ_{long}^2 distributions as described in Section 5.9.1.

6.1 Corrected yield as a function of the transverse energy

After applying all cuts explained in Section 4.3 and after applying the isolation criterion discussed in Section 5.4, we obtain the uncorrected density distribution of isolated narrow cluster transverse energy (also called *raw yield* or *raw spectrum* in the following of the chapter), shown in Fig. 6.1.

We first estimate the contamination C in our narrow isolated cluster sample via the ABCD method, as illustrated in Fig. 6.2; afterwards we correct this estimate with the factor α , see Fig. 5.23 (blue points), using the following relation:

$$C = \frac{B_n^{iso,est}}{N_n^{iso}} \rightarrow C = \alpha \cdot \frac{B_n^{iso,est}}{N_n^{iso}} = \frac{B_{n,MC}^{iso,true}}{B_{n,MC}^{iso,est}} \cdot \frac{B_n^{iso,est}}{N_n^{iso}} \quad (6.1)$$

By doing so, we can subtract the background contribution produced

6.1. Corrected yield as a function of the transverse energy

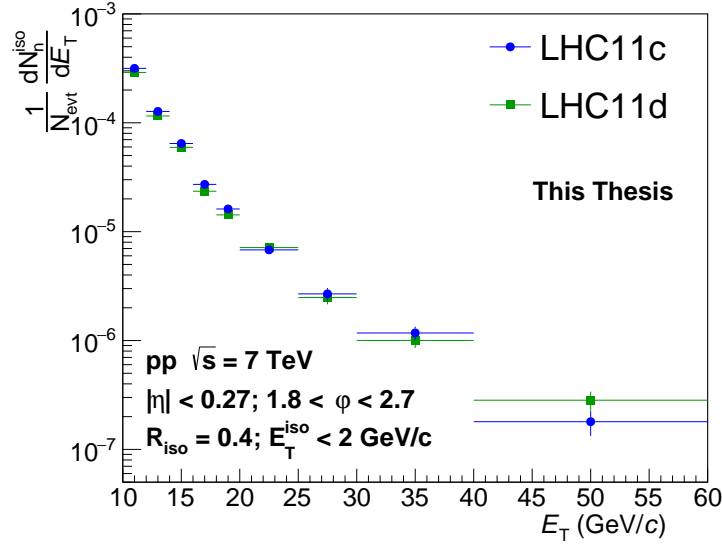


Figure 6.1: Raw spectrum for isolated narrow neutral clusters computed from the data collected in the LHC11c and LHC11d period.

mainly by neutral pion decay photons.

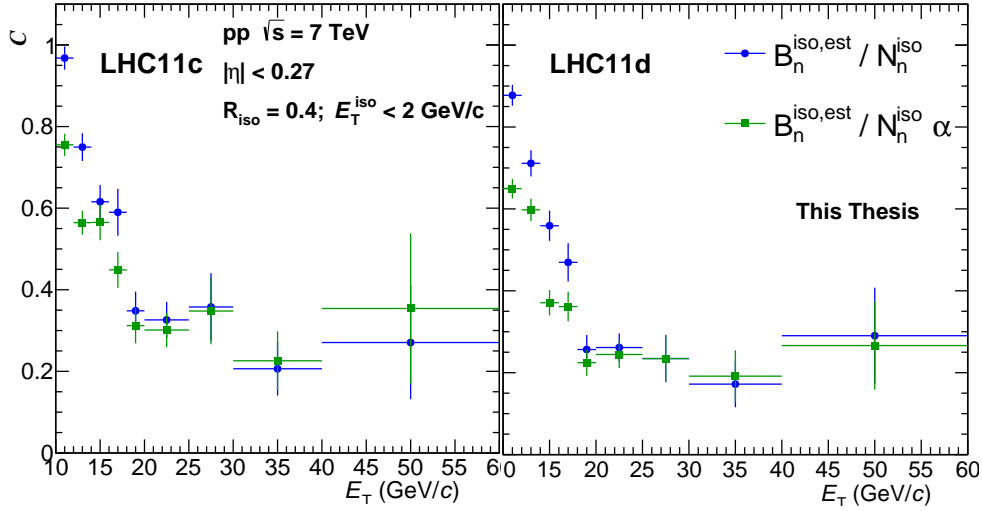


Figure 6.2: Uncorrected (blue) and corrected (green) contamination fraction in the signal region computed for both the LHC11c (left) and LHC11d (right) data samples separately.

After subtracting the estimated background, we need to correct the iso-

lated photon spectrum computed for each period for the efficiency of reconstruction and identification (studied in Section 5.4, Fig. 5.21) and for the trigger efficiency to obtain the fully corrected spectrum for isolated photon production, presented in Fig. 6.3. The lower panel shows a ratio of the corrected yield obtained from the LHC11d and LHC11c period. The two periods are compatible with each other within 2σ throughout the whole transverse energy range explored.

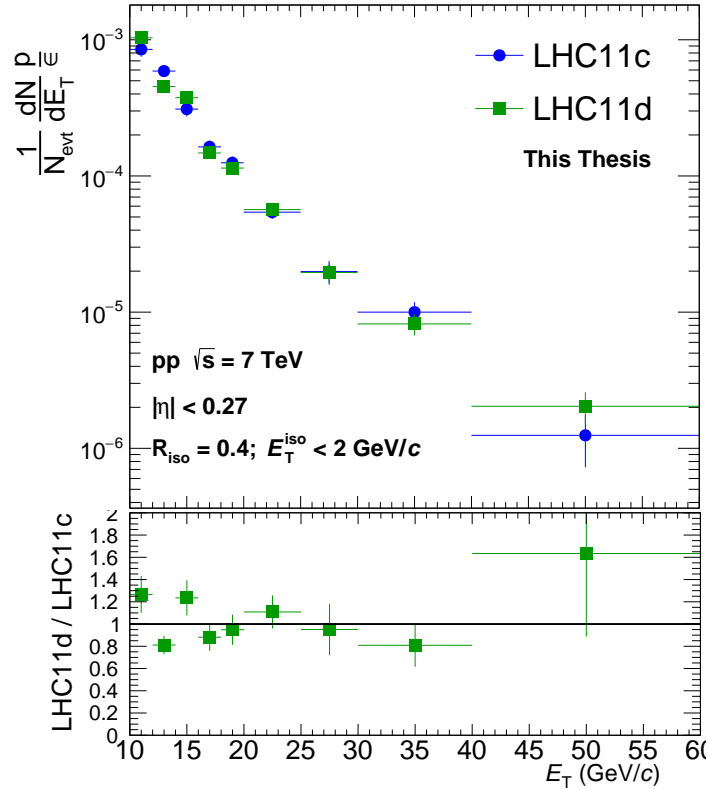


Figure 6.3: Isolated photon spectrum obtained from the two data samples available. The lower panel shows a comparison between the two spectra, which agree at least to a 95 % confidence level.

Before continuing with the computation of the cross-section of isolated photon production, we need to estimate the systematic uncertainties of our measurement. To explore the systematic uncertainties, we vary several cuts in the analysis and report the resulting changes in the corrected yield.

Some of the selection criteria might affect more than one quantity at the

6.2. Systematic uncertainties in data: triggering

same time. To disentangle the effect of using different selection criterion, the same procedure has been repeated varying, one at a time, each selection criteria applied, and the changes in the final result are compared to the corrected spectrum presented in Fig. 6.3.

The final result for the isolated photon production cross section will be given and discussed in the next chapter where we will include the total systematic uncertainty resulting from the studies presented in the following sections of this chapter.

6.2 Systematic uncertainties in data: triggering

In this section we will discuss the systematic uncertainties on the event selection, and in particular the triggering. While the use of a trigger as a tool for on-line selection of events is mandatory for our analysis, the periods analysed were the first where the EMCal-L0 trigger has been used, and the system was still being tuned. The problems in the trigger algorithm that have an effect on our analysis are: an imperfect synchronisation of the triggering system and instabilities of the triggering procedure during the data taking periods. In the next paragraphs we will study the bias introduced in our measurement. In particular effects like out-of-time and in-bunch pile up triggered events are not included in our simulated samples, where only one event is simulated at a time.

6.2.1 Out-of-time (bunch) triggered events

The clustering procedure assigns the arrival time of the seed cell signal to each cluster. The EMCal-L0 trigger algorithm used in this analysis requires a high energy deposition in EMCal and simultaneous hits on the V0 detector. In Section 4.1.2 we saw how selecting clusters with a time signal within 30 nanoseconds from the triggering time selects clusters from a single pp event.

During the acquisition of the analysed data the EMCal-L0 trigger signal was not fully synchronised, meaning that it could have arrived ~ 50 ns earlier than the actual bunch -crossing time. As a consequence, all reconstructed

clusters have been assigned a time shifted by ~ 50 ns later. The effect of this de-synchronisation is shown in Fig. 6.4, where we have now studied the time distribution already presented in Fig. 4.2, with a wider range, and as a function of the energy of the cluster. As one can see, a second peak is present in the time distribution. This peak is the result of the synchronisation problems just explained. The clusters belonging to the second peak have been excluded from this analysis, but it is important to estimate the effect of this loss due to bad triggering.

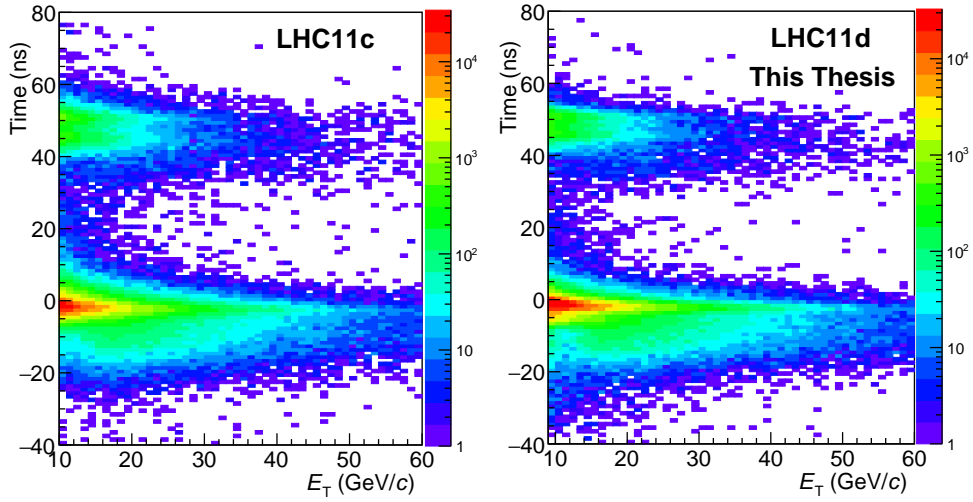


Figure 6.4: Cluster time distribution as a function of the cluster energy for the LHC11c (left) and LHC11d (right) periods of data taking.

In order to quantify the loss of clusters produced in events that fired the trigger but were assigned a later arrival time, we define the cluster rejection factor as

$$T = \frac{N_{clus}(-30, 30)}{N_{clus}(-30, 70)}. \quad (6.2)$$

The clusters used to compute the ratio in Eq. 6.2 satisfy only the requirement $\sigma_{\text{long}}^2 > 0.1$. Figure 6.5 shows the dependence of this ratio on the cluster energy for the two periods separately. We observe a decrease in the number of accepted clusters with respect to the total, with increasing energy. This dependence is a consequence of the trigger algorithm used at that time, which allowed for the electronic noise to be added while building the patches signal, therefore allowing for trigger signals to arrive earlier than the actual signal.

6.2. Systematic uncertainties in data: triggering

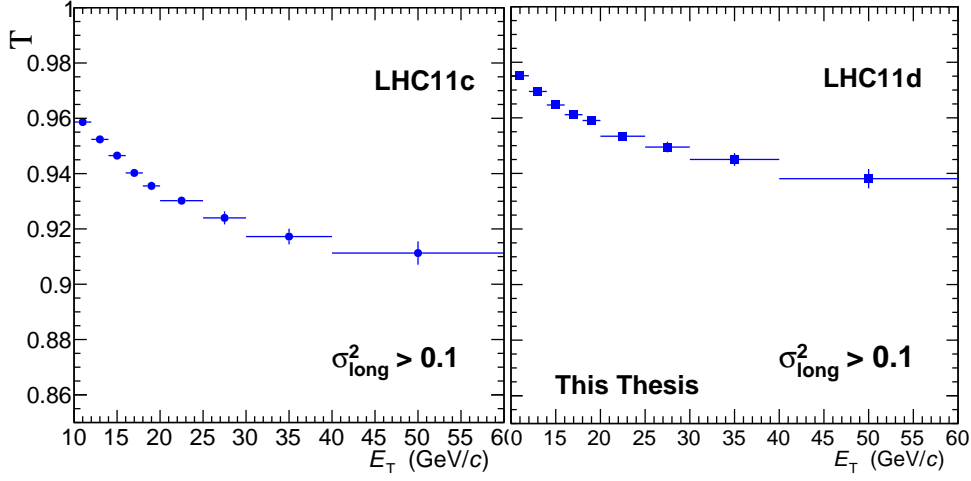


Figure 6.5: Dependence of the ratio in Eq. 6.2 on the reconstructed energy of the cluster.

We will take into account the modified trigger efficiency, quantified via the cluster rejection factor T , when computing the corrected yield from the neutral cluster raw spectrum in the following of this chapter.

6.2.2 Efficiency and stability of the triggering algorithm

As already said in Chapter 3, the results presented in this thesis are based in part on two EMC-L0 triggered data samples, with the trigger threshold set to 5.5 GeV/c. The EMC-L0 class of events is a subsample of the *minimum bias* class of events (named INT7), so we can rely on the latter to compute the efficiency of the EMC-L0 triggering algorithm. Counting the number of events which have fired the EMC-L0 trigger and the number of events belonging to the INT7 trigger class that present a high energy deposition in EMCal, we obtain the trigger efficiency, defined as:

$$\epsilon_{\text{trig}} = \frac{N_{\text{evt}, \text{EMC-L0} + \text{clust} > 10} \text{ GeV}}{N_{\text{evt}, \text{INT7} + \text{clust} > 10} \text{ GeV}} \quad (6.3)$$

We have found $\epsilon_{\text{trig}} = 0.90 \pm 0.06^{\text{stat}}$.

The trigger efficiency is constant as a function of the transverse energy. This was expected, since the range of transverse energy explored in this

analysis is well above the threshold set for the EMC-L0 algorithm. The efficiency is below 100 % because the sliding patch algorithm only considers towers within a TRU, as discussed in Section 3.5.1, which leads to reduced efficiency at the edges.

In order to further ensure that both our data samples have been recorded with similar detector conditions, we have studied the stability of the number of high energy clusters per event, corrected by the efficiency, as a function of the run number. This study is needed especially because the MC samples, used to compute both the α factor and the reconstruction and identification efficiency, have been produced mimicking the detector conditions of four runs only. This affects the efficiency estimate.

Figure 6.6 shows the percentage of high energy clusters per event for each run number (each with possible small changes in detector conditions). A fit with a constant function properly describes the data, with very few outliers due to statistical fluctuations. To estimate the effect of the instability of the trigger, the points have been fitted with a constant function and the relative uncertainty on the stability of the trigger efficiency is calculated from the RMS spread of the data points around the fit.

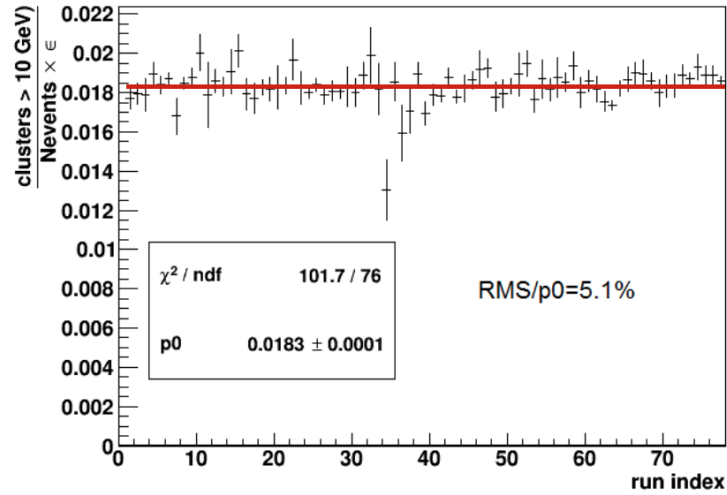


Figure 6.6: Number of high energy clusters per event, after correction for the computed reconstruction efficiency. Figure taken from [43].

6.2. Systematic uncertainties in data: triggering

6.2.3 In-bunch pile up events

Finally, the sample of recorded events which triggered the EMC-L0 trigger may contain events in which more than one inelastic proton-proton collision happens within the same bunch crossing (in-bunch pile up). Such events can bias both the determination of the production rate for neutral particles (and consequently candidate clusters) and the measurement of their in-cone activity.

The EMC-L0 trigger class considered in this analysis is rare enough that the bias in the determination of the production rate for neutral particle is about 0.03 %, and therefore negligible.

The bias in the measured isolation activity due to particles produced in a second inelastic collision in the same bunch crossing has been estimated by means of toy simulation. We have selected, in the same periods, the minimum bias trigger class, and stored the in-cone activity associated with a fake random trigger particle identifying this with the contribution from pile-up events to the in-cone activity of true neutral cluster candidates in collisions that fired the EMC-L0 trigger class. We then added artificially the isolation activity from fake trigger particles to the simulated event sample that models our background.

Despite allowing for 20 % of the events to be affected by in-bunch pile up, the percentage of isolated clusters obtained from the event sample with *embedded* pile up is 0.3 % higher than the percentage of non isolated clusters computed from pure background simulated sample.

The same study has been repeated on the simulated sample that models the production of prompt photons, with similar values for the difference of percentage of isolated clusters. Therefore, the bias on the measure of the in-cone activity of neutral particles due to in-bunch pile up is considered negligible.

6.3 Systematic uncertainties from MC: the choice of the Mixing Ratio

In this section we will focus on the systematic uncertainties exclusively related to the MC corrections applied in our analysis. The affected quantity is the calibration factor α , while the effect of different selection criteria on quantities like the efficiency will be studied in Section 6.4.

In Section 5.6 we have shown how the signal and background sample are used together to calculate the factor α , used to correct for the non-factorisation of the σ_{long}^2 and $E_{\text{T}}^{\text{iso}}$ variables. However, the size of the correction depends on the signal/background ratio in the simulated events, i.e. the relative contribution of both prompt photon and jet production to the total cross section.

In this section we will estimate the uncertainty on the calibration factor related to the signal fraction used to obtain the total MC mixed sample, giving us an insight on the effect of the leakage of the signal counts (mainly) into the \mathcal{C} region. Figure 6.7 shows the result of this study. Since the choice of the weight applied to the MC-GJ sample affects exclusively the computation of the α factor, it is possible to estimate the effect on the corrected spectrum directly from the variation of the calibration factor itself. The black points (used as reference in the lower panels) show the α factor computed from a mixture of signal and background sample that best reproduces the uncorrected contamination estimated in data, as shown in Fig. 5.12.

Despite varying the weights applied to the signal, the α factor remains stable at low values of the transverse energy of the cluster ($E_{\text{T}} < 16 \text{ GeV}/c$); at high values of the reconstructed transverse energy of the cluster there is a clear monotonic relation.

Therefore, the systematic uncertainty due to the choice of the weights applied to the signal sample is estimated to be of the order of 1 % for $E_{\text{T}} < 16 \text{ GeV}/c$ and increasing up to 6 %, for $40 \leq E_{\text{T}} < 60 \text{ GeV}/c$.

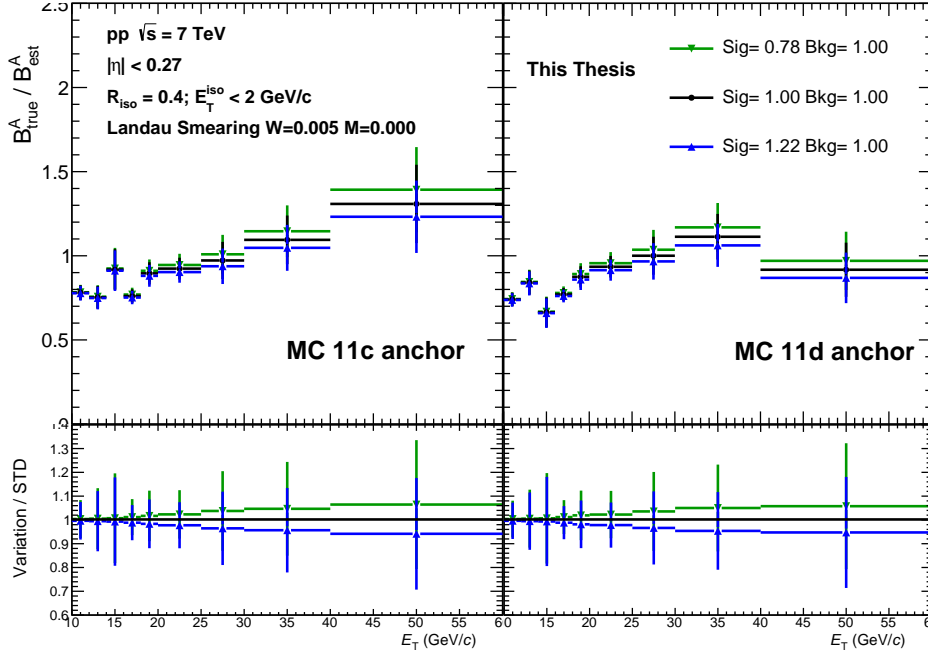


Figure 6.7: Variation of the calibration factor α for different choices of the signal/background ratio used for the MC analysis.

6.4 Systematic uncertainties from data and MC

In this section we will show the main sources of systematic uncertainties due to all selection criteria applied in the analysis. We will explore changes of the final results due to the variation of the track-cluster correlation distance allowed by the CPV algorithm, which will affect the neutral cluster sample used in the analysis, and the selection criterion for narrow and wide clusters, which will modify the raw yield and the estimate of the contamination in the \mathcal{A} region.

6.4.1 Residual hadronic contamination

In the analysis, a charged particle veto cut is applied to remove clusters that are produced by charged particles (electrons and hadrons). The non-perfect charged particle tracking efficiency might cause either fake matches, which would reduce the number of neutral clusters in our sample, or missed

matches, which would result in clusters produced by a photon and a charged particle, biasing the reconstructed energy of the photon. Furthermore, neutral hadrons might also be present in our neutral cluster sample.

In this section we will present the results of the studies performed to estimate the uncertainty related to the three effects mentioned.

Correction for random vetoes

In order to estimate how well our CPV algorithm performs in rejecting charged particles and how many photon cluster candidates are randomly rejected by the CPV we have studied the distance between good clusters¹ and the closest track as a function of the cluster transverse energy. This study is performed on all clusters with number of local maxima $NLM \leq 2$ and the results obtained from the LHC11d data sample are presented in Fig. 6.8 where we have plotted the distribution of the distance between each good-quality cluster and the track closest to it, for clusters with reconstructed transverse energy greater than 5 GeV/ c . A clear sharp peak near distance ΔR_{C-T} is found, and the selection is estimated to be more than 99 % efficient when the track is reconstructed.

Extrapolation of the distributions at distance $\Delta R_{C-T} > 0.01$ below the peak provide an estimate of the combinatorial rejection of neutral candidate clusters. All distributions in Figs. 6.8 (and Fig. 6.9 for MC) extend (on the x axis) up to more than three times the range shown here. The loss of candidates due to random vetoes is found to be of the order of 0.2 %, independent of the transverse energy of the cluster, and it is therefore considered negligible.

As a cross-check, we have repeated the same study on the MC samples and the corresponding distributions of distance, shown in Fig. 6.9, lead to conclusions similar to those obtained in data.

¹A cluster is considered *good* if it is made of more than two cells, and if it fulfils the topological cut presented in Section 4.1.2.

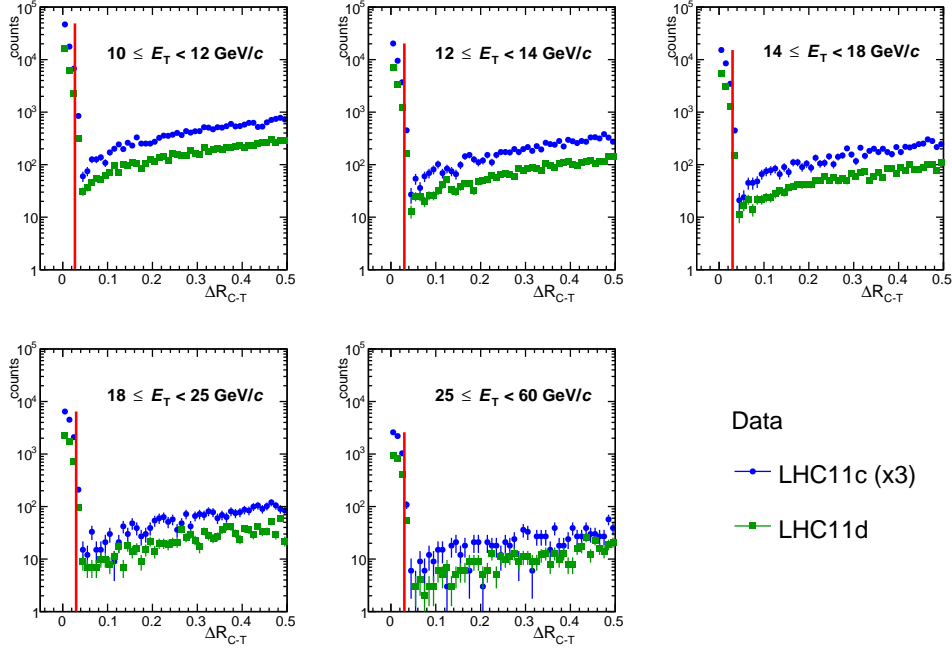


Figure 6.8: Distribution of the distance between the reconstructed position of accepted clusters and closest track, obtained in pp collisions L0-triggered data for both LHC11c (blue) and LHC11d (green) periods.

Studies on cluster-track distance cut

In the previous paragraph we have seen that the CPV algorithm used in this analysis ensures a highly pure candidate cluster sample. Still, both the raw yield and the estimated contamination (and consequently the corrected yield) might be sensitive to variations of the cluster-track correlation distance used to identify a cluster produced by a charged particle.

The variations used to study the systematic uncertainty of our veto method are summarised in Table 6.1, while Fig. 6.10 shows the effect of such variations of the raw yield in the \mathcal{A} region. As one can see, the variations of the correlation distance between cluster and track lead to a difference in the raw spectra up to 2 % and up to 1 % for the corrected yield, shown in Fig. 6.11.

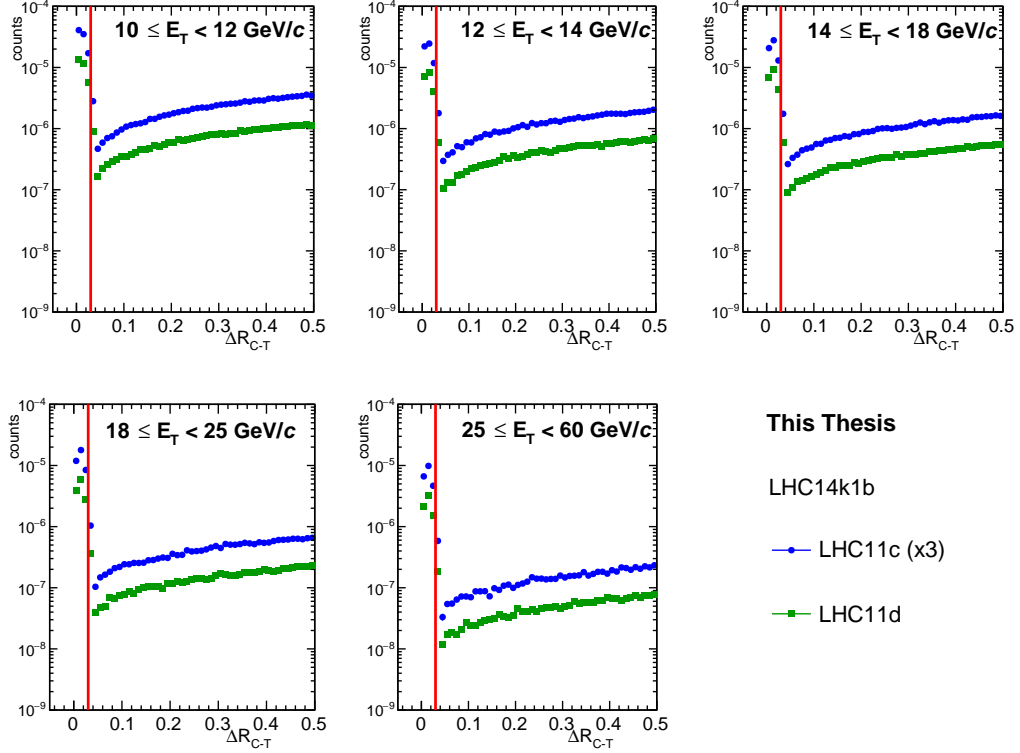


Figure 6.9: Distribution of the distance between the reconstructed position of accepted clusters and closest track, obtained from one of the MC-JJ samples available for this analysis, separated by anchor period.

Neutral hadron contamination

When measuring photons in a complicated environment such as that produced in high energy pp collisions, the determination of the neutral cluster yield depends on the veto procedure used to exclude charged particles. Moreover, particles like the K_L^0 meson, neutron or Λ baryon do not (always) decay within the sensitive volume of the ALICE experiment and cannot be detected by the tracking system because of their charge neutrality; as a consequence, any CPV algorithm will be "blind" to these particles that might pollute our sample. Still, because neutral hadrons do not interact with the electromagnetic field of the atoms in matter, but rather with the nuclei, the probability for such interaction is far smaller, such that these particles deposit very lit-

Correlation Distance	$\Delta\eta_{C-T}$	$\Delta\varphi_{C-T}$
Loose	0.015	0.025
Standard	0.02	0.03
Tight	0.025	0.035

Table 6.1: Variations of the CPV distances in η and φ coordinates used to select candidate clusters. The variations used are simultaneous in the two coordinates and are not changed as a function of the transverse momentum of the cluster.

the energy in the ALICE-EMCal. Figure 6.12 shows the relative abundances of particles that produce a cluster populating the \mathcal{A} region, computed from the simulated sample. As one can see, the residual contribution of neutral hadrons (green empty stars) is negligible.

6.4.2 Cluster definition via shower shape selection

In this section we will study how the cuts on the shower shape distributions affect the corrected yield. They will affect the estimation of the contamination in the \mathcal{A} region, the computation of the calibration factor and, in case of variation of the narrow cluster definition, the computation of the efficiency for reconstruction and identification. In the next paragraphs we will study effects on the different quantities separately, even though some of the parameters may be correlated.

We have repeated the analysis by keeping the lower limit of the narrow cluster region fixed while varying the upper limit of the narrow cluster region by 0.05 downward and upward. In addition, we have changed the lower limit of the wide cluster region by the same value 0.05; the upper limit of the σ_{long}^2 parameter for the wide cluster is changed by the same amount, so that the control regions \mathcal{B} and \mathcal{D} keep the same size in σ_{long}^2 .

Table 6.2 summarizes the modifications of the cut values used for the different transverse energy ranges.

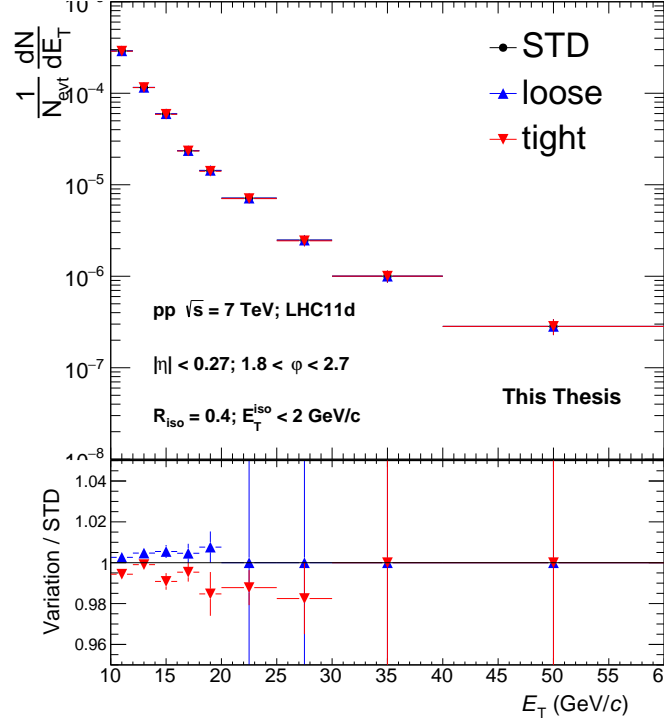


Figure 6.10: Raw yield as a function of the transverse energy of the cluster obtained using different CPV definitions. The plot refers to the results obtained from the LHC11d period. The lower panel shows the ratio of the spectra with a modified CPV distance cut to those obtained with the default cut.

Narrow clusters selection

Changing the limits of the narrow cluster region has a direct effect on both the efficiency and on the contamination in the isolated region; furthermore, since the definition of narrow cluster is common to both region \mathcal{A} and to the control region \mathcal{C} , changes in the narrow region definition will also affect the ABCD method estimate of the contamination in the isolated narrow cluster sample. Since changing the shower shape cut for narrow clusters affects the raw yield as well as the contamination and its estimate via the ABCD method, the effect is best quantified by comparing the corrected yield. The variations to the shower shape cuts for narrow and wide clusters are listed in

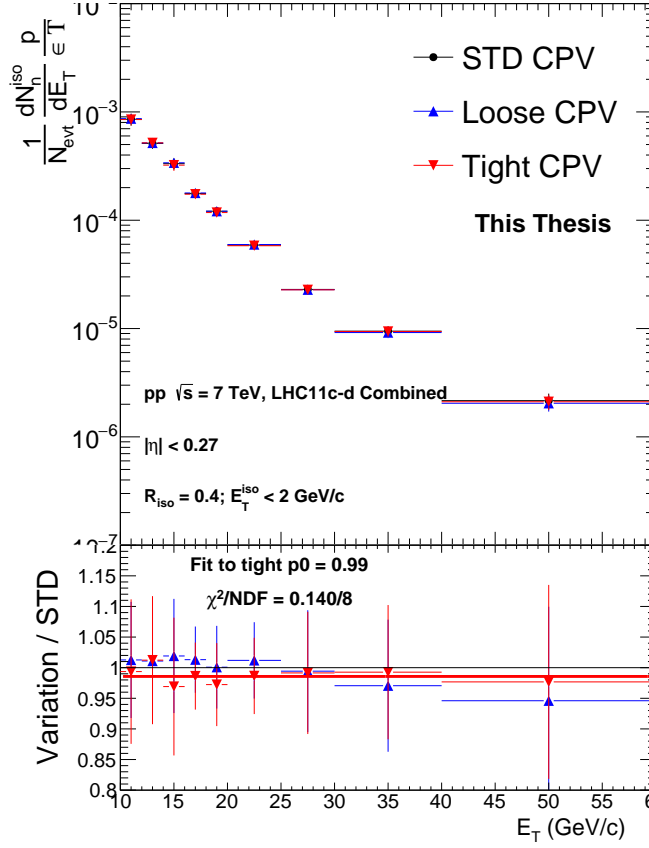


Figure 6.11: Corrected isolated photon yield as a function of the transverse energy, obtained using using different CPV definitions. The plot refers to the results obtained from the whole data sample analysed. The lower panel shows the ratios of the spectra with a modified CPV distance cut to those obtained with the default cut, together with an estimate of the larger deviation, from the two variations tried, by means of a fit to the points with a constant function.

table 6.2.

Figure 6.13 shows the effect that changing, separately, the selection criterion for narrow and wide clusters has on the raw yield, while the changes of the contamination corrected for the α factor (Fig. 6.14), according to Eq. 6.1 are presented in Fig. 6.15. The correction factor α is quite stable with respect to changes in the definition of narrow cluster region, even though the

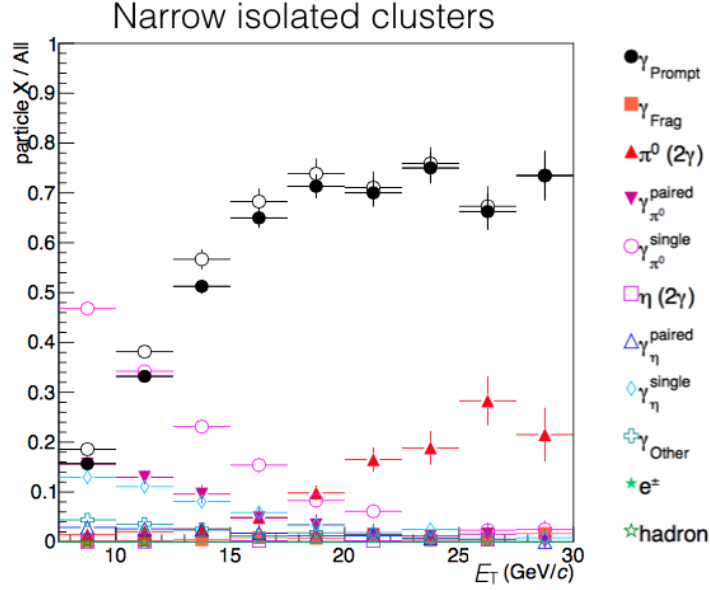


Figure 6.12: Percentages of particles producing a narrow cluster with $E_T^{iso} < 2$ GeV/ c . The contribution of neutral hadrons to the total sample of neutral clusters is negligible. Figure taken from [44].

E_T^{clust} (GeV/ c)	Upper limit narrow region (σ_{long}^2)			Lower limit wide region (σ_{long}^2)		
Variation	Low	STD	High	Low	STD	High
10-16	0.4	0.45	0.5	0.5	0.55	0.6
16-18	0.3	0.35	0.4	0.4	0.45	0.5
18-60	0.27	0.3	0.35	0.35	0.4	0.45

Table 6.2: Variations of the σ_{long}^2 values used to study how the definition of the narrow and of the wide cluster region affect the corrected yield for isolated photons.

cut modifications lead to changes of the raw yield of 15 %.

The stability of the α factor with respect to the choice of the weights applied to the MC-JJ and MC-GJ sample has been shown in Fig. 6.7, while

6.4. Systematic uncertainties from data and MC

Fig. 6.14 proves the stability of the same quantity with respect to the choice of the signal range for the σ_{long}^2 parameter. We conclude that the α factor is stable under sizeable changes of the analysis and that a correction strategy that the ABCD method on data and a correction factor from simulated events is likely to be more accurate than e.g. using a background estimation that is directly taken from simulated events.

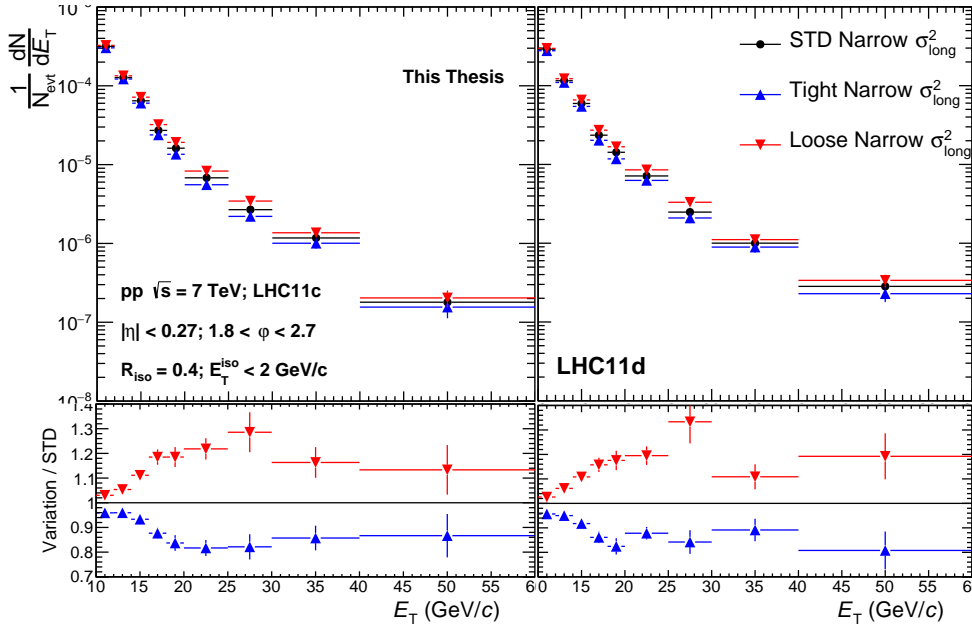


Figure 6.13: Raw yield of isolated narrow neutral clusters for different selections of narrow cluster, for the two data subsamples analysed. The variations used are reported in table 6.2. The lower panels show a comparison of the yields computed by varying the selection with respect to the yield obtained with standard selection.

Figure 6.16 shows the variation of the reconstruction and identification efficiency for both periods separately. We observe an increase of the efficiency with the width of the selected narrow cluster region, as expected.

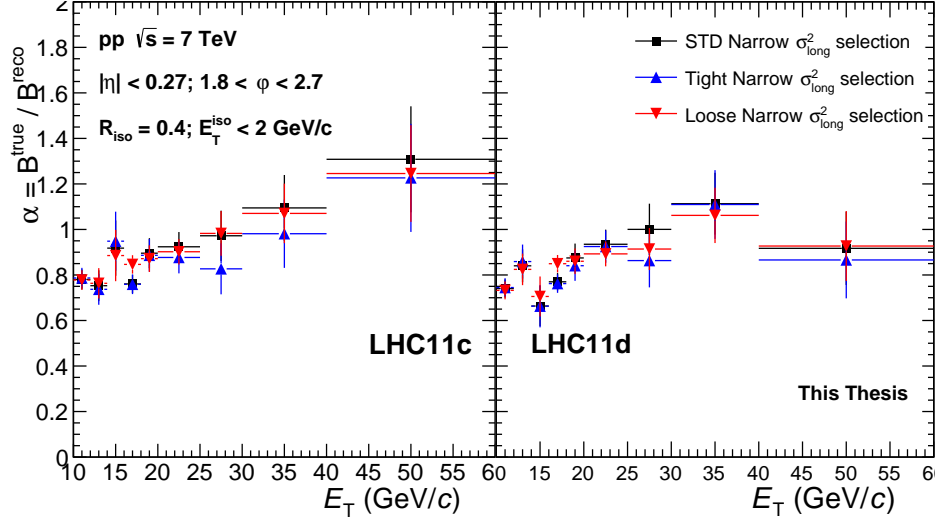


Figure 6.14: Contamination correction factor α , as a function of the transverse momentum, using various definitions for narrow clusters in the MC mixed samples.

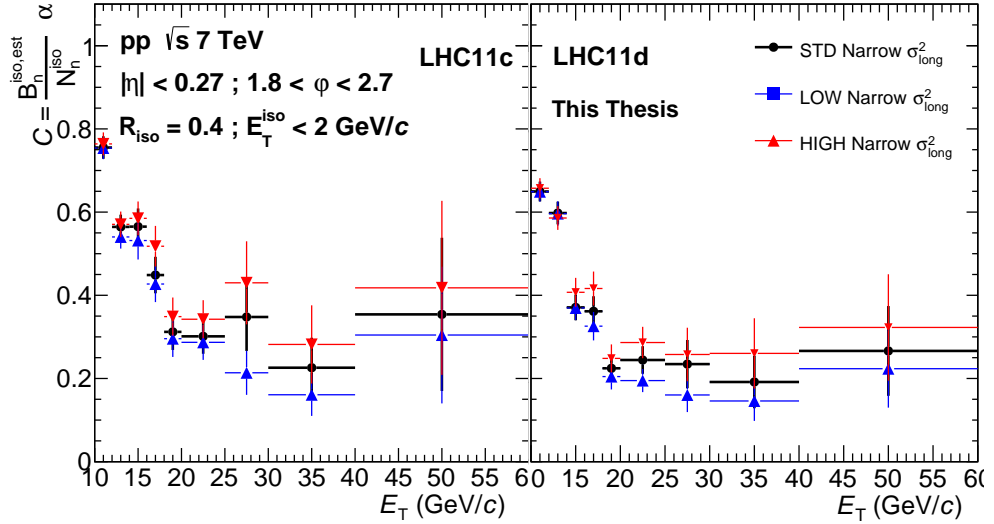


Figure 6.15: Contamination, defined in Eq. 6.1, as a function of the transverse energy of the cluster, for different definitions of narrow cluster. The different panels show the results obtained in the two periods separately.

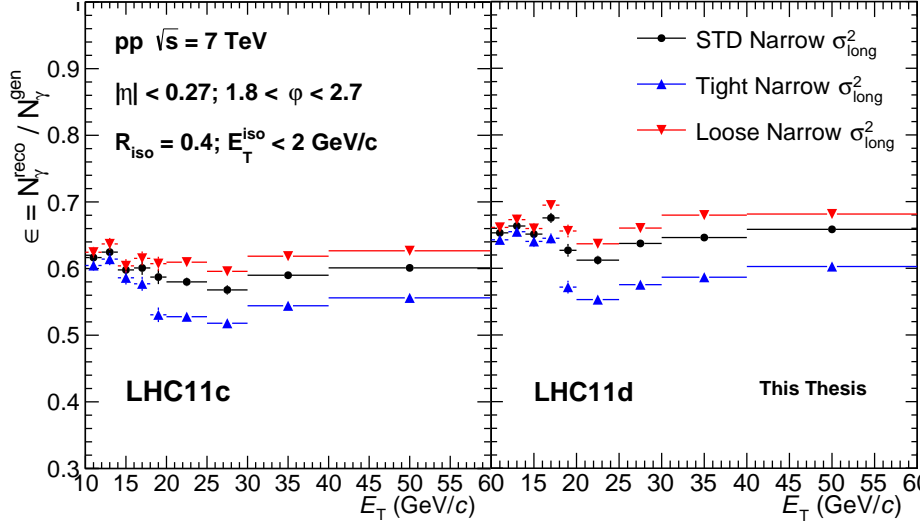


Figure 6.16: Reconstruction and identification efficiency, as a function of the reconstructed energy of the cluster, for different definitions of narrow cluster, computed from the MC signal samples anchored to the different periods.

Wide clusters selection

In this paragraph we will present the effect of changing the selection for wide clusters; we are interested in the modification of the estimated contamination in the \mathcal{A} region when varying the selection of wide clusters; the leakage of signal counts in the \mathcal{B} region, the efficiency in tagging clusters from decay photons and, consequently, the non-factorisation of the σ_{long}^2 and E_T^{iso} parameters will be the quantities affected by the different selection. These effects are explored by changing the gap between the narrow and wide region of the $E_T^{\text{iso}} - \sigma_{\text{long}}^2$ plane. The selection of narrow clusters remains unchanged, while varying simultaneously both limits of the range for wide clusters. Such variations affect the estimated contamination in data and the computation of the α factor from the MC-mixed samples, while leaving the efficiency unmodified.

Figure 6.17 shows the effect on the fully corrected yield of isolated prompt photon due to varying separately the narrow and the wide cluster region definition. Modifying the definition of narrow clusters has direct consequences, as we saw in the previous paragraph, both on the efficiency and on the con-

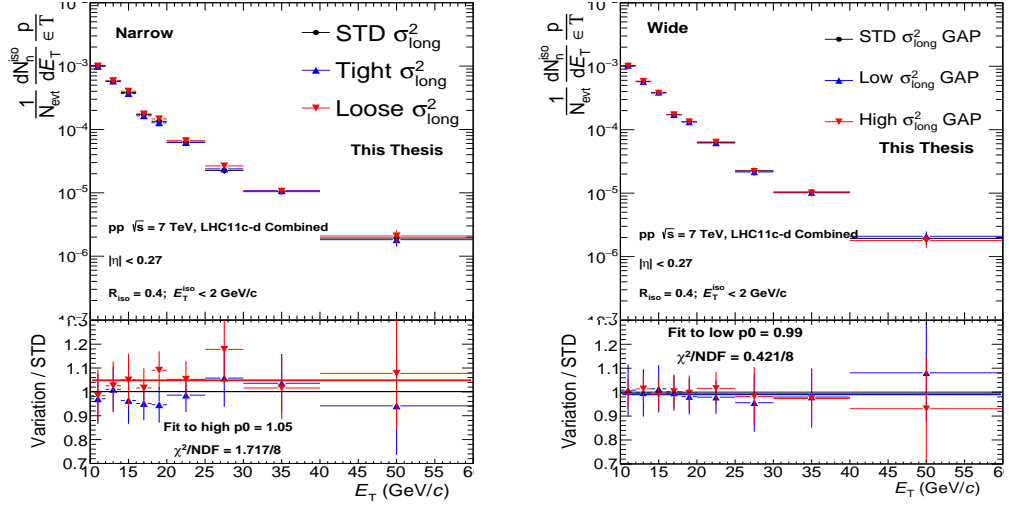


Figure 6.17: Isolated photon spectrum for different values of the right hand limit of the narrow cluster range (left) and for different values of the left hand limit of the wide cluster range (right). The upper panels show the corrected yields, while the lower panels shows the ratios of the resulting spectra with varied cuts to the standard analysis settings (black points), all reported in table 6.2. The horizontal lines are constants fit to the ratios.

tamination of the isolated cluster sample, resulting in greater changes on the corrected yield compared to the (almost negligible) changes observed when the wide cluster region definition is modified. In order to quantify the discrepancy between the various corrected yields the ratios have been fit with a constant (E_T - independent) and the results of this procedure are shown in the lower panels for the largest deviation, with respect to the standard selection.

6.5 Uncertainty from different isolation probability in data and MC

In Section 5.9.1 we have discussed the modelling of the transverse shower shape distributions in MC and the difference with respect to those computed from the data sample, showing how smearing the distributions of the σ_{long}^2 parameter obtained in the simulated sample helps reducing the discrepancy in the isolation probability for wide clusters in data and MC (Figs. 5.18 and 5.22). The data-MC agreement was quantified by a double ratio of isolated to non-isolated clusters in data and MC, Eq. 5.5, and we observed that using smeared distributions of the transverse width parameter in MC helps reducing the offset of the double ratio with respect to unity, especially for wide clusters. In this section we will illustrate the effect of choosing different functions for the additional smearing of the σ_{long}^2 parameter on the correction factor α , on the efficiency, and on the isolated photon spectrum.

Moreover, to improve the agreement of the isolation probability computed in data and MC mixed sample, we can modify the $E_{\text{T}}^{\text{iso}}$ selection for the background regions \mathcal{C} and \mathcal{D} (both in data and simulated samples)

Finally we will study the effect on the corrected yield obtained when using only reconstructed charged particles in the measurement of the isolation activity. As discussed in Section 4.3.2, this is expected to improve the factorisation of the background distribution in σ_{long}^2 and $p_{\text{T}}^{\text{iso}}$.

All these modifications to the analysis strategy will modify the double ratio, leading to a different estimation of the contribution of the background in the narrow cluster region. In fact, it is found that the double ratio depends on the σ_{long}^2 parameter for most of the explored variations. An additional correction to Eq. 6.1 is introduced to take this dependence into account.

In the next sections, results of these studies will be presented and the uncertainty on the isolated photon yield due to the different isolation probability in data and MC determined.

6.5.1 Isolation probability and smearing

To study the changes on the double ratio with respect to the smearing function used to improve the data-MC agreement, we have repeated the analysis on the MC samples with varying the functional form used to smear the MC σ_{long}^2 distribution:

1. Landau smearing with Width = 0.005 and Mean = 0.000;
2. Landau smearing with Width = 0.010 and Mean = 0.000;

The first smearing function is the default, which has already been discussed in Section 5.9.1, and the effect of its use in the computation of the corrected yield shown in Section 6.1, while the second option is used as a variation of the smearing procedure. Other functional forms have been tried to smear the distributions of the σ_{long}^2 parameter obtained from the MC samples (Gaussian function and Landau shifted function), but none of these provided a satisfactory description of the σ_{long}^2 distributions obtained from data.

In the following, we will refer to the first smearing as *standard* (STD), while the second will be referred to as *wide* (WIDE) smearing.

Figs. 6.18 and 6.19 show respectively the α factor and the efficiency computed when using different modelling of the transverse shower shape distribution in MC. We observe that, at low E_T , the α factor is independent of the smearing used, while using a WIDE smearing results in an α factor slightly closer to unity for high values of the energy of the cluster. Applying the WIDE smearing to the σ_{long}^2 distributions obtained from the simulated samples also reduces the efficiency for reconstruction and identification by $\sim 2\%$ at low E_T and up to 6% for high values of the reconstructed energy of the cluster.

The double ratio defined in Eq. 5.5 also changes as a function of the smearing used. The effects of the different smearing on the double ratio are reported in Fig. 6.20, for the LHC11c period, as a function of the reconstructed cluster energy. The blue (STD) and red (WIDE smearing) distributions are compatible with each other, especially for high values of the cluster transverse width.

6.5. Uncertainty from different isolation probability in data and MC

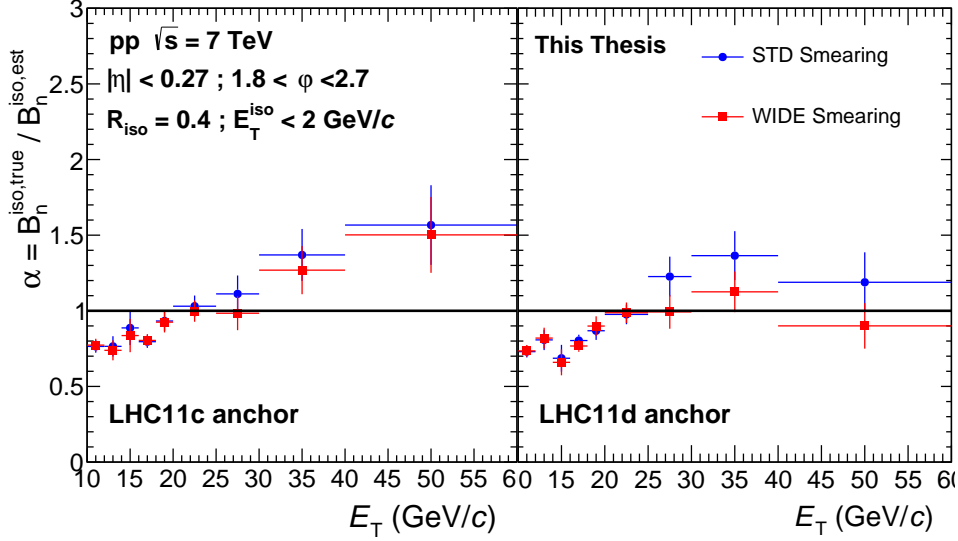


Figure 6.18: Dependence of the α factor on the reconstructed transverse energy of the cluster for two smearing procedures. The two panels show the results obtained from simulated samples anchored to the different periods of data taking analysed.

We observe that, at all cluster energies, the double ratio, for high values of the σ_{long}^2 parameter, can be modelled with a linear function.

It is worth remembering that while the cluster population in this region consists almost entirely of clusters produced by decay photons from neutral mesons (background clusters), the narrow cluster region is populated by clusters from both isolated and decay photons. We attempt to estimate the contribution of background clusters in the narrow cluster region by a fit to the points of the double ratio in the wide cluster region and extrapolating the resulting functional form to the region of narrow clusters.

Since the choice of the function used to model the dependence of the double ratio on the σ_{long}^2 parameter in the wide cluster region will affect the estimation of the background cluster in the narrow region, the fitting procedure has been repeated with various functions. We consider a polynomial function of the first order and an exponentially falling function, with the constraint that the slope must not be lower than -1, provide the best description of the trend of the double ratio in the wide cluster region.

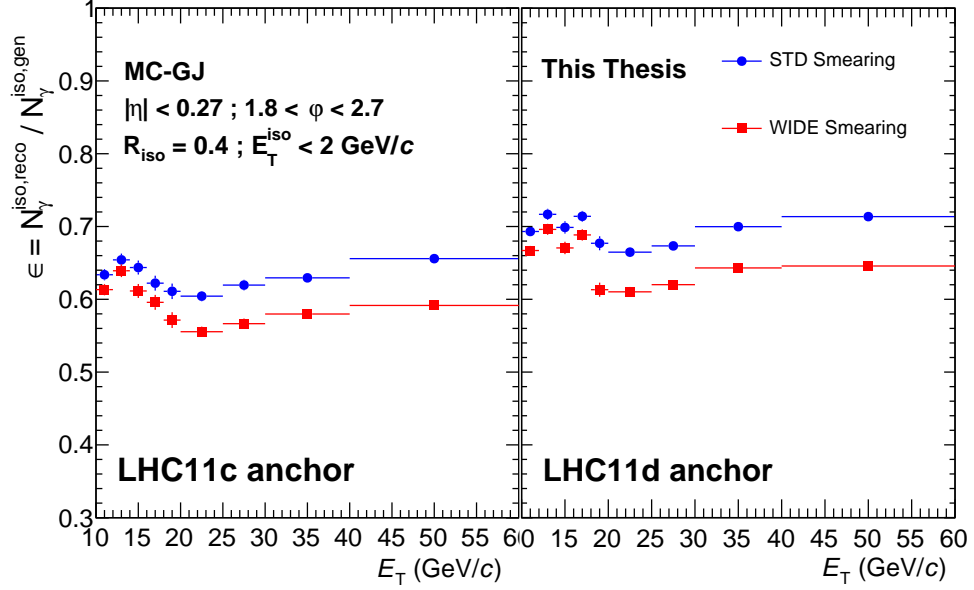


Figure 6.19: Reconstruction and identification efficiency for isolated photons, computed for the LHC11c (left) and LHC11d (right) period separately. The efficiency decreases with the width of the Landau function used to smear the distributions of the σ_{long}^2 parameter obtained from the MC signal sample.

Assuming a linear dependence of the double ratio on the σ_{long}^2 parameter yields a correction of the α factor:

$$\alpha \rightarrow \alpha' = \alpha \cdot \frac{p_0 + p_1 \cdot x_n}{p_0 + p_1 \cdot x_w}, \quad (6.4)$$

where p_0 and p_1 are respectively the constant and the slope obtained from the fit and x_n and x_w are the barycenters of the σ_{long}^2 distributions from data in the narrow and wide cluster region respectively.

Similarly, if we assume an exponential dependence ($f(x) = q_0 \cdot e^{q_1 \cdot x} + q_2$) of the double ratio, the correction of the α factor is:

$$\alpha \rightarrow \alpha' = \alpha \cdot \frac{q_0 \cdot e^{q_1 \cdot x_n} + q_2}{q_0 \cdot e^{q_1 \cdot x_w} + q_2}, \quad (6.5)$$

with definitions for q_i and x_j analogous to those reported for Eq. 6.4.

Figs. 6.21 and 6.22 show examples of the results of this study for the sub

6.5. Uncertainty from different isolation probability in data and MC

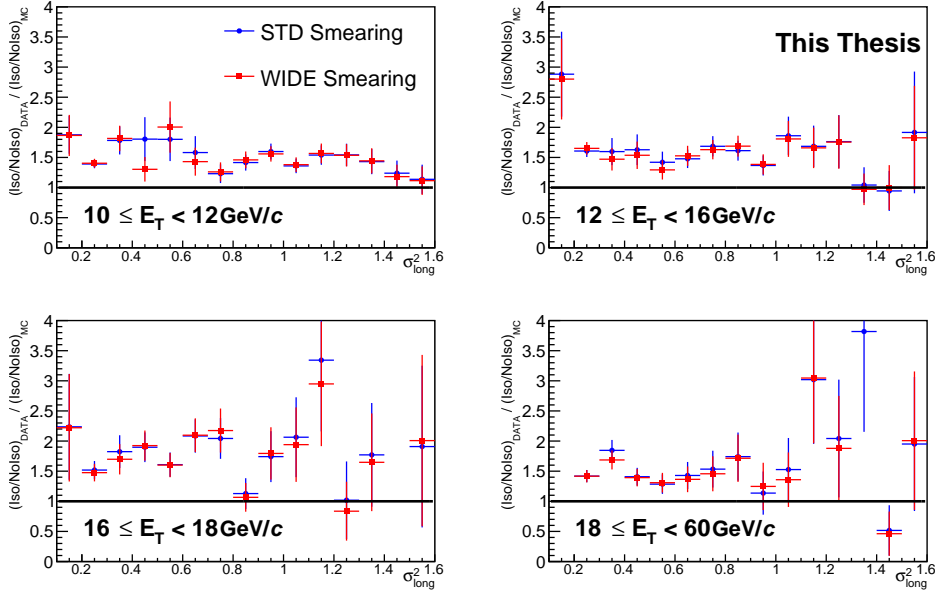


Figure 6.20: Double ratio, defined in Eq. 5.5 computed with different smearing of the σ_{long}^2 distributions in MC mixed sample anchored to LHC11c.

period LHC11c (linear assumption) and LHC11d (exponential dependence) respectively, where the first collection of plots is computed with the STD smearing while the second collection is computed with the WIDE smearing. A complete collection of all results produced by this study can be found in Appendix B. In case of the exponential dependence of the double ratio on the transverse width of the cluster, all fits to the double ratios have been performed by fixing the q_1 parameter to the value of -0.3.

Figure 6.23 shows the alpha factor, as a function of the transverse energy of the cluster, when different functions are used to fit the double ratio in the wide cluster region and to estimate the contribution of the background in the narrow region.

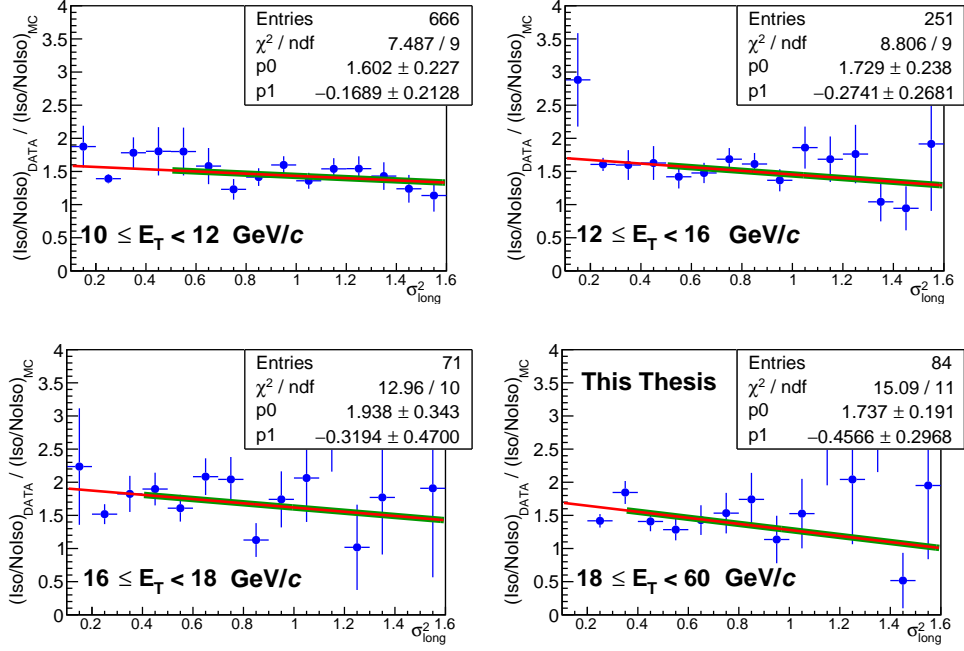


Figure 6.21: Double ratio, as defined in Eq. 5.5, as a function of the σ_{long}^2 parameter for four different ranges of the reconstructed transverse energy of clusters. The results shown refer to the LHC11c sub period, with a standard smearing of the σ_{long}^2 parameter. The green lines represent the result of the fit to the points performed with a polynomial function of the first order, while the red lines show the estimation of the background in the narrow cluster region via extrapolation of the resulting fitting function.

6.5. Uncertainty from different isolation probability in data and MC

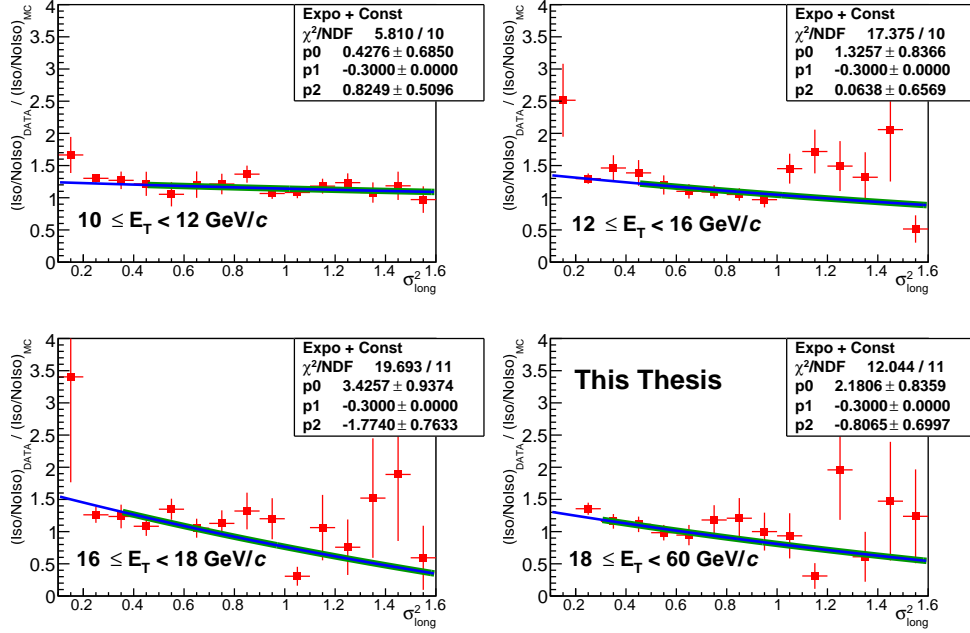


Figure 6.22: Double ratio, as defined in Eq. 5.5, as a function of the σ_{long}^2 parameter for four different ranges of the reconstructed transverse energy of clusters. The results shown refer to the LHC11d sub period, with a wide smearing of the σ_{long}^2 parameter. The green lines represent the result of the fit to the points performed with an exponentially falling function, while the blue lines show the estimation of the background in the narrow cluster region via extrapolation of the resulting fitting function.

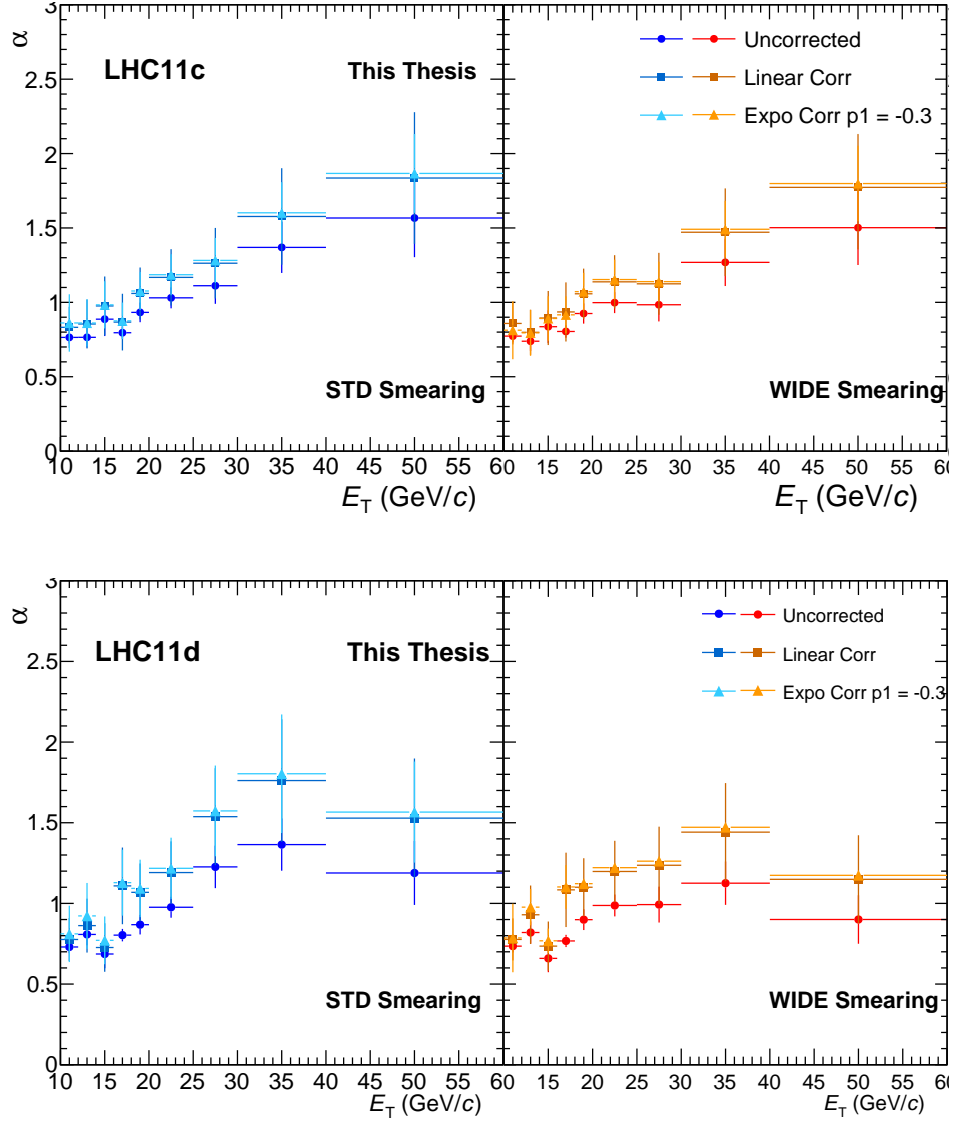


Figure 6.23: Correction factor α , defined accordingly to Eqs. 5.3, 6.4 and 6.5, as a function of the transverse energy of the cluster. The results shown refer to the LHC11c (top panels) and LHC11d (bottom panels) periods. The left column shows the modification of the alpha factor when the standard smearing is applied, while the right column refers to the MC modelling of the σ_{long}^2 parameter with a WIDE smearing.

6.5.2 Anti isolation criterion

To quantify the effect that the leakage of prompt photon clusters in the control region \mathcal{C} has on the double ratio in Eq. 5.5, we can modify the definition of non-isolated clusters, similarly to what we did for the definition of wide clusters in case of the σ_{long}^2 distributions. Separating the value of total in-cone activity allowed for isolated clusters and the minimum value of the same quantity for a cluster to be considered non-isolated, we expect to control the leakage of the signal into the \mathcal{C} region, therefore checking the role of the signal in the non-factorisation of the σ_{long}^2 and $E_{\text{T}}^{\text{iso}}$ variables. For this test, the isolation energy threshold for non-isolated clusters was changed from $E_{\text{T}}^{\text{iso}} > 2 \text{ GeV}/c$ to $E_{\text{T}}^{\text{iso}} > 4 \text{ GeV}/c$, while the selection of isolated clusters was kept at $E_{\text{T}}^{\text{iso}} < 2 \text{ GeV}/c$. As a consequence, the ratio between isolated and non isolated clusters in data and MC changes, leading to modifications of the double ratio defined in Eq. 5.5. The resulting double ratio is shown in Fig. 6.24, and compared to the standard isolation selection. We observe that, in case of the modified anti-isolation criterion, the offset of the double ratio with respect to one is smaller than the offset obtained with the Standard analysis procedure, indicating that the isolation probability for wide cluster in MC is now more similar to that computed in data.

By changing the anti-isolation criterion we change the definition of the \mathcal{C} and \mathcal{D} regions, affecting the estimation of the contamination in the \mathcal{A} region via the ABCD method. Moreover, also our estimate of the contribution of background clusters to the narrow cluster region will change. To quantify these effects, the same fitting procedure explained in the previous section is repeated. The effect of this estimation on the α factor is presented in Fig. 6.25, separately for the LHC11c (left) and LHC11d (right) period, for different anti-isolation criteria used and with different background cluster estimation in the narrow region. At low E_{T} we observe small changes in the computed alpha factor, while at high values of the reconstructed energy of the cluster we observe an offset of the alpha factor computed when a tighter anti-isolation criterion is used; this offset increases even more when the contribution of background clusters in the narrow cluster region is obtained by extrapolating the result of the fit with a linear or exponential function as in

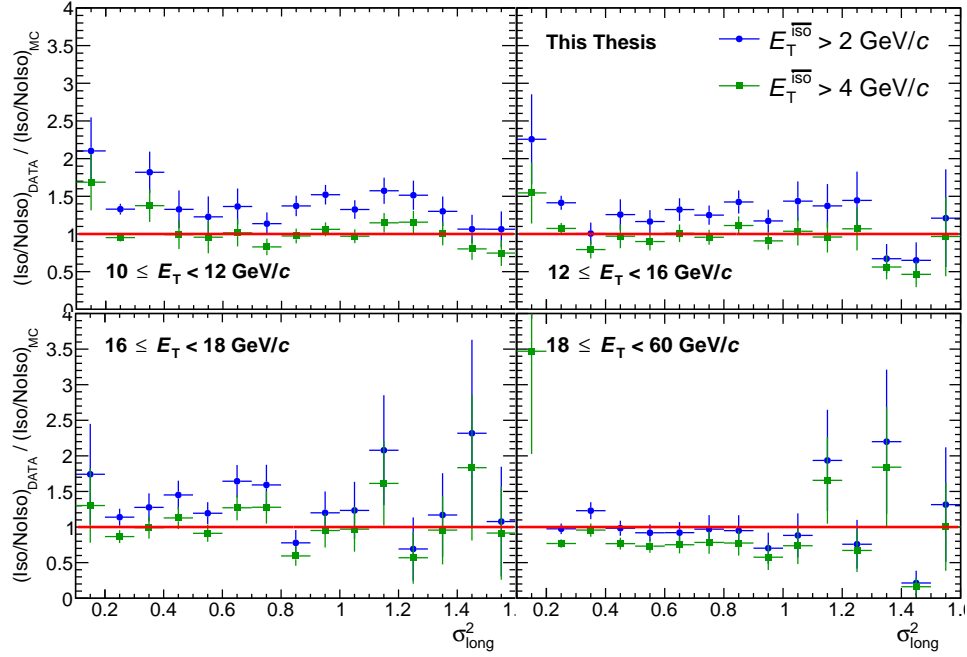


Figure 6.24: Double ratio as a function of the σ_{long}^2 parameter for two different selections of non isolated clusters and for different values of the cluster transverse energy. The results shown refer to the LHC11c period.

Eqs. 6.4 and 6.5.

6.5.3 In-cone activity with charged particles

In the previous chapter we have studied the differences between data and MC samples in both the σ_{long}^2 and the in-cone activity distributions. As argued in Section 5.8, the main cause of non-factorisation of the two variables used in the ABCD method is the fact that EMCAL is used both to identify the neutral candidate clusters and to measure (part of) the in-cone activity. Narrow clusters are likely single decay photons from a higher energy neutral pion and their decay partner will fall into the isolation cone of the candidate clusters, increasing their isolation activity; wide clusters may rather be produced by pairs of decay photons: the reconstructed cluster energy will likely be closer to the neutral pion energy and the measured isolation activity for such clusters

6.5. Uncertainty from different isolation probability in data and MC

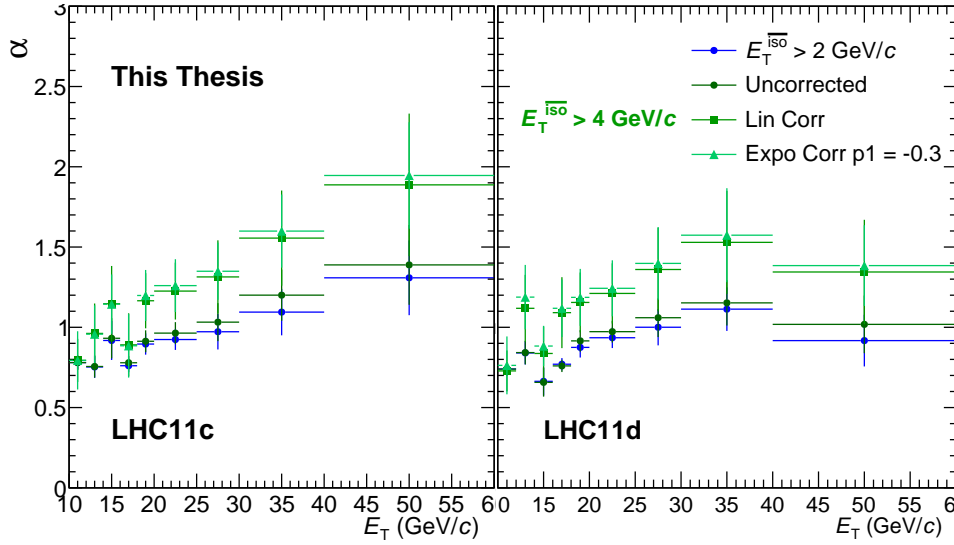


Figure 6.25: Correction factor α computed, according to Eq. 5.3, 6.4 and 6.5, for the LHC11c (left) and LHC11d (right) period separately. The different corrections are presented with different markers in shades of green. A comparison with the uncorrected case in which the standard definition of non-isolated clusters is used (blue points) is also shown.

will not include the decay partner contribution.

To reduce this correlation effect between the two parameters, we have repeated the analysis estimating the isolation activity of each candidate by the sum of the measured transverse momentum of charged particles only, indicated with p_T^{iso} and referred to, in the following, also as *charged* component isolation (or in-cone) activity.

The selection of candidates, the measurement of their energy and the computation of the σ_{long}^2 parameter will be carried out using EMCAL, while the measurement of the in-cone activity will be carried out exclusively from the response of the ALICE tracking system.

Using only the charged component of the isolation energy reduces the selectivity of the isolation criterion. In Section 5.9.2 we presented the same-rejection algorithm used to compute which value of the total in-cone activity obtained in simulated samples should be used to obtain the same percentage of clusters tagged as non-isolated. By applying the same rejection algorithm

to the *total* and *charged* in-cone activity distributions in data, we can obtain a cut value for the p_T^{iso} distributions for which the rejection of non isolated clusters is similar to the case of total in-cone activity with anti-isolation criterion $E_T^{iso} > 2 \text{ GeV}/c$.

The result of such study tells is, that one would have to use as isolation criterion the condition:

$$p_T^{iso} = \sum_{i=0}^{N_{track}} p_T^i < 0.25 \text{ GeV}/c. \quad (6.6)$$

However, when using such a low threshold value, we may become sensitive to soft particle production in the underlying event, for which modelling uncertainties are large. In addition, considering that the reconstruction and identification for charged particles with the full ALICE tracking system is possible for $p_T \geq 150 \text{ MeV}/c$, such a low isolation threshold will already reject candidates which have one low momentum particle in the cone.

Therefore we decided to use a higher threshold of

$$p_T^{iso} = \sum_{i=0}^{N_{track}} p_T^i < 1.25 \text{ GeV}/c. \quad (6.7)$$

as a default for the analysis with charged particle isolation. This value is motivated by the consideration that, because of isospin symmetry², about 2/3 of the total momentum in the event is carried by charged particles, making the theoretical estimate

$$p_T^{iso} \approx 2/3 \cdot E_T^{iso} \sim 1.33 \text{ GeV}/c. \quad (6.8)$$

The theoretical values reported in Eqs. 6.7 and 6.8 differ by a small amount, still this can lead to sizeable difference in the raw yields. To quantify this effect, we have repeated the study setting the threshold for isolated clusters to $p_T^{iso} < 1.50 \text{ GeV}/c$. The raw yield obtained with this isolation criterion were consistent, within a few percent, with those obtained with the isolation

²According to the isospin symmetry, all charge state of the pions are equally populated so that 2/3 of them are charged. As pions are the most abundant particles, we take this value for pions as the general scaling factors for all charged particles.

6.5. Uncertainty from different isolation probability in data and MC

criterion from Eq. 6.7, therefore we concluded that the small difference between the values reported in 6.7 and 6.8 does not affect the determination of the corrected yield. Figure 6.26 shows the isolated narrow cluster spectrum computed with two isolation criteria for the charged component of the isolation activity. As a comparison, the raw yield computed by using the total in-cone activity is also shown. The lower panels present the ratios of the spectra, obtained with the charged component, to the standard result. A large difference (up to $\sim 60\%$) is seen, at low values of the reconstructed cluster energy, for the raw yield in both LHC data taking period analysed.

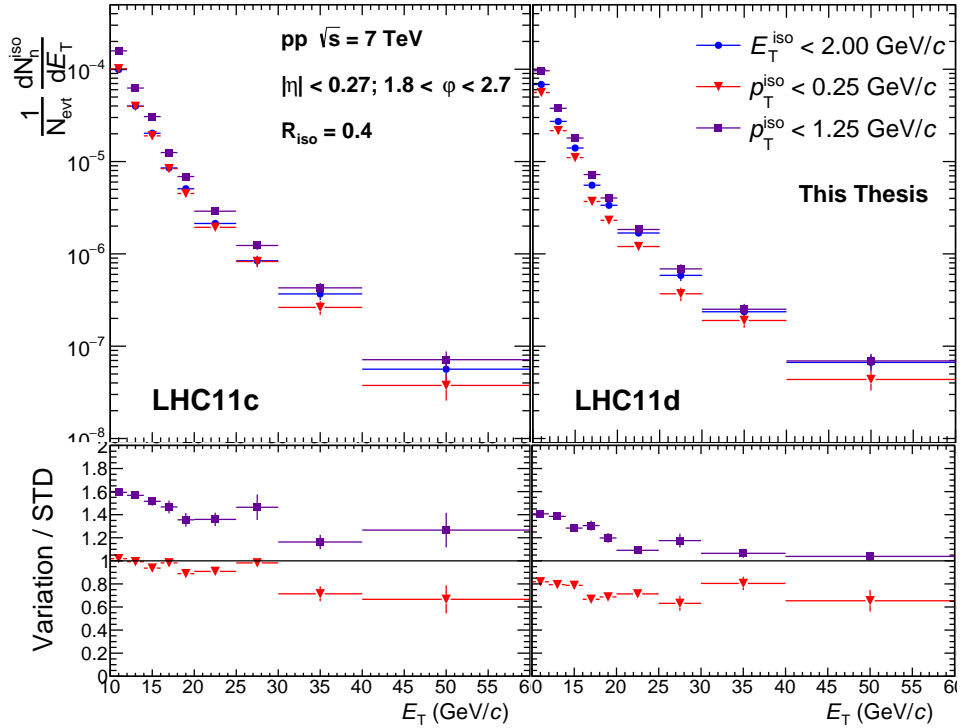


Figure 6.26: Raw yield computed with two different estimators of the in-cone activity and for two different isolation criteria in case the isolation activity is computed exclusively with charged particles. The lower panels show the ratios between the yield computed using charged particles and the yield computed with the total isolation activity.

Figure 6.27 shows the different components of the in-cone activity measured in data. The neutral component is well reproduced in our MC sample,

while the MC distributions of the in-cone activity of charged particles deviate from the spectra obtained from the data sample. The disagreement between the distributions from data and simulations is larger for the charged particle (p_T^{iso}) in-cone activity distributions than for the total isolation energy (E_T^{iso}) distributions, and the effect is overall larger for the LHC11c data taking period than for the LHC11d one.

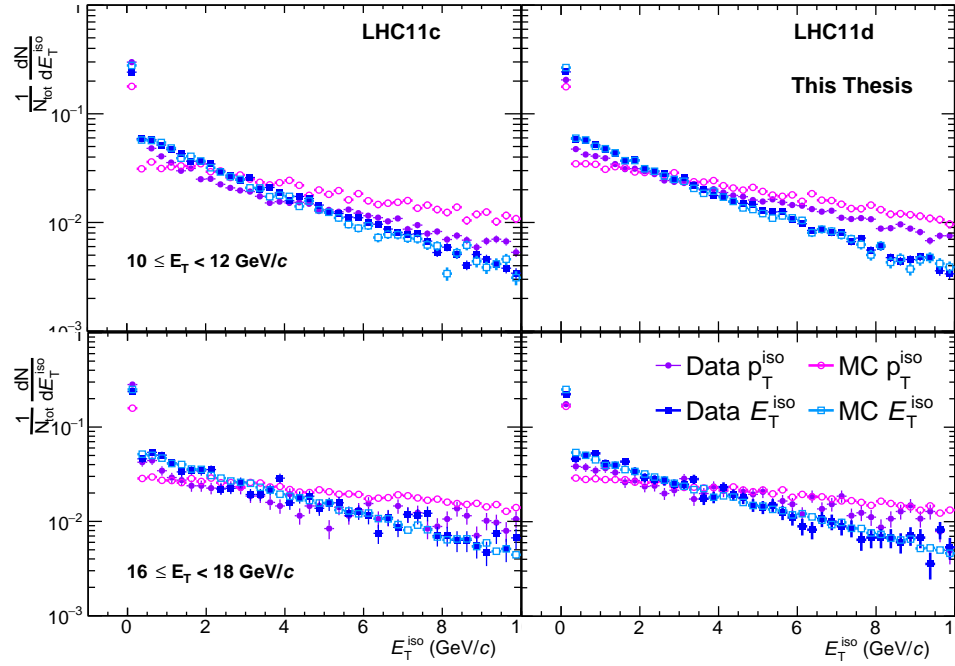


Figure 6.27: Different components of the in-cone activity probability distribution for LHC11c (left) and LHC11d (right) period. In dark colours the distributions obtained from the data sample, while light colours represent the distributions from the respective MC mixed samples. The neutral component (shades of red) of the isolation activity is well modelled in MC, while the charged component shows a different slope with respect to the distributions computed in data.

The mismatch between data and simulation can have its origin in the event generator, which may not fully describe the isolation energy distribution and/or in the detector response simulation. The larger difference between the track-based in-cone activity distributions in the LHC11c period is due to imperfect calibration of the TPC, which leads to a poor description

6.5. Uncertainty from different isolation probability in data and MC

of the tracking performance by the detector simulation.

Because of the small cross section for isolated photon production in pp collision, we expect the mismatch of the in-cone activities in data and MC to be related (almost) exclusively to the MC-JJ sample, while the change of the cut value for the allowed charged in-cone activity does not bias the estimation of the efficiency of our measurement significantly.

In order to check that the reasons for the non-factorisation of the E_T^{iso} and the σ_{long}^2 parameter is indeed the use of ECal to measure both quantities, we have computed the α factor when using the ABCD method on the $p_T^{iso} - \sigma_{long}^2$ plane, and compared the results to those obtained with the standard analysis.

Figure 6.28 shows the result of such comparison. We observe a sizeable change of the dependence of the alpha factor on the cluster transverse energy.

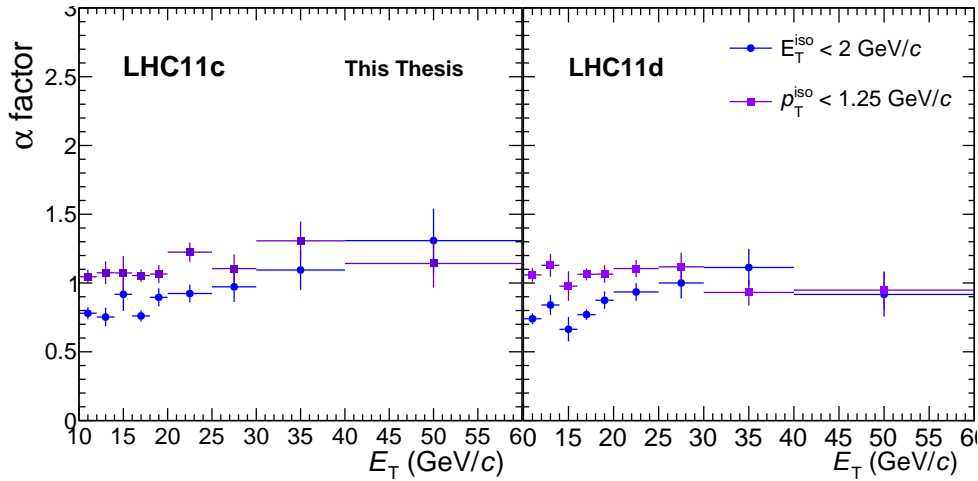


Figure 6.28: Correction factor α computed, according to Eq. 5.3, for the LHC11c (left) and LHC11d (right) period separately, when using two different estimator of the in-cone activity.

The test of a different estimator for the in-cone activity was justified by the non-factorisation of the E_T^{iso} and σ_{long}^2 parameters. As expected, the α correction factor is closer to 1 when using only charged particles for the estimation of the in-cone activity, indicating a better factorisation of the background in the $p_T^{iso} - \sigma_{long}^2$ variables. However, the α factor does not quite reach unity, which shows that there is still a remaining non-factorising

contribution. This is probably caused by the fact that narrow clusters are produced by photons from a neutral pion with energy higher than the cluster energy, decaying with a wide opening angle, while the reconstructed energy for wide clusters is close to the energy of the π^0 . For the same cluster energy, narrow clusters are then originating from π^0 of higher energy than wide clusters. This might imply that the energy of the jet, where the π^0 is found, would also be higher. The average activity in the jet cone would then likely be different for the two cases.

Since non-factorising effects can not be fully avoided, we will continue to use the contamination correction factor α also for the analysis with p_T^{iso} as estimator of the isolation activity.

Figure 6.29 shows the estimated background percentage in the \mathcal{A} region, corrected for the α factor (left), and the reconstruction and identification efficiency (right). Both quantities are shown for the two data taking periods separately.

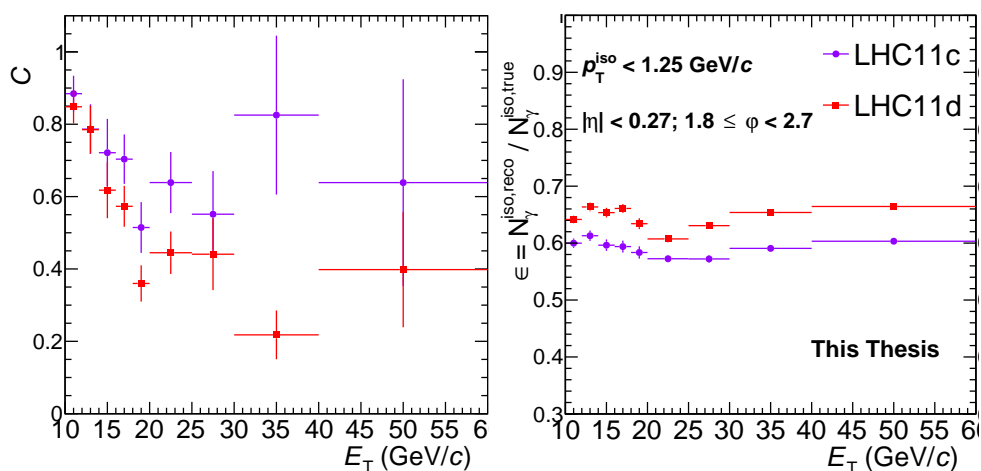


Figure 6.29: Contamination (left), as defined in Eq. 6.1 and reconstruction efficiency (right) computed when estimating the in-cone activity around the candidate direction using exclusively charged particles. Both quantities shown are computed considering the isolation criterion in Eq. 6.7.

Still, similarly to what has been done in the previous two sections, we need to quantify the difference in the isolation probability for narrow and

6.5. Uncertainty from different isolation probability in data and MC

wide clusters in data and MC and, in case a non constant dependence of the double ratio on the σ_{long}^2 parameter is found, to determine a correction for the α factor, according to Eq. 6.4 and Eq. 6.5.

Figure 6.30 shows the modifications to the α factor, as a function of the transverse energy of the cluster, for different modelling of the background cluster contribution to the narrow cluster region. While for the LHC11c period the α factor shows small differences whether we compute it via Eqs. 5.3, 6.4 or 6.5, the variations in the α computed from the simulation anchored to the LHC11d period are larger when extrapolating the background cluster contribution to the narrow cluster region.

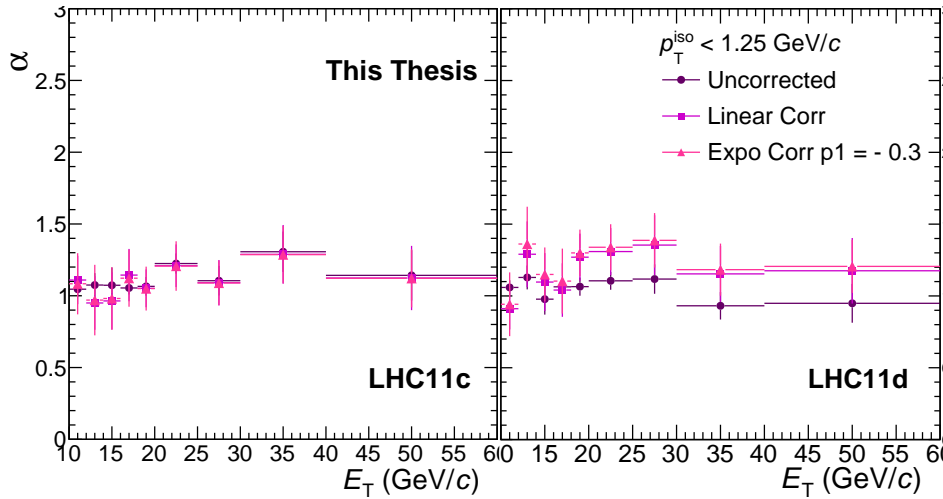


Figure 6.30: Contamination correction factor α , defined according to Eqs. 5.3, 6.4 and 6.5, as a function of the transverse energy of the cluster. The left plot refers to the MC sample anchored to the LHC11c period, while the right plot shows the results obtained from the simulations anchored to the LHC11d period.

We observe that all fits to the double ratios show a value for the χ^2 , divided by the number of degrees of freedom for the fits, very close to 1. We have also performed fits to all double ratios, computed with the different variations, by varying the value of the slope for the exponential function used. Among all possible values, we tried values of the slope -0.5, -1, -2, and -3. For all these cases we observed that the fitting algorithm was compensating for the different value of slope by modifying (sometimes enormously) the value of the other two parameters, without changing the resulting fit function. For this reason, it has been decided not to include the estimates of the background to the narrow cluster region obtained when setting the slope of the exponential fit function to -0.5, -1, -2, -3.

6.5.4 Determination of the central value for the corrected yield

In the previous sections we have presented the different strategies used to study the possible bias due to different isolation probabilities in data and simulated samples, for narrow and wide clusters. We have seen how, for all four variations (2 different smearing, change of the anti-isolation criterion and change of the estimator for the in-cone activity) of the analysis strategy, the main differences for the correction of the α factor are found between the different assumptions on the dependence of the double ratio on σ_{long}^2 , while the different values for the slope of the exponential function yield essentially the same correction of the alpha factor.

To obtain the isolated photon yield corrected for the different isolation probability in data and MC we computed, for each of the variations discussed in the previous sections, the raw yield and corrected this for the reconstruction efficiency. We then subtracted the corrected contamination defined in Eq. 6.1 with the α factor computed according to Eqs. 5.3, 6.4 or 6.5.

Afterwards we computed, as a function of the cluster transverse energy, the mean value

$$N_{\gamma}^{\text{iso,Mean}}(E_T) = \frac{\sum_{i=1}^4 (N_{\gamma}^{\text{iso,Uncorr}} + N_{\gamma}^{\text{iso,Lin}} + N_{\gamma}^{\text{iso,Exp}})}{12}, \quad (6.9)$$

where i represents the variation to the analysis strategy (2 smearings,

6.5. Uncertainty from different isolation probability in data and MC

anti-isolation criterion and in-cone activity estimator) and the standard deviation:

$$\sigma(E_T) = \sqrt{\frac{\sum_{j=1}^{12} (N_{\gamma}^{iso,j}(E_T) - N_{\gamma}^{iso,Mean}(E_T))^2}{12}} \quad (6.10)$$

The first will be our best estimate for the measurement presented in this thesis, while the latter will be the estimate of the systematic uncertainty due to the different isolation probability for narrow and wide clusters in data and MC.

Figure 6.31 shows the result of this study, the four upper panels show the corrected yield for the four variations of the analysis strategy obtained with the various corrections for the different isolation probability in data and MC. As a comparison, mean isolated photon spectrum is also shown. The lower panels show the ratio of all variations with respect to the computed mean value, while the standard deviation is shown in light grey boxes.

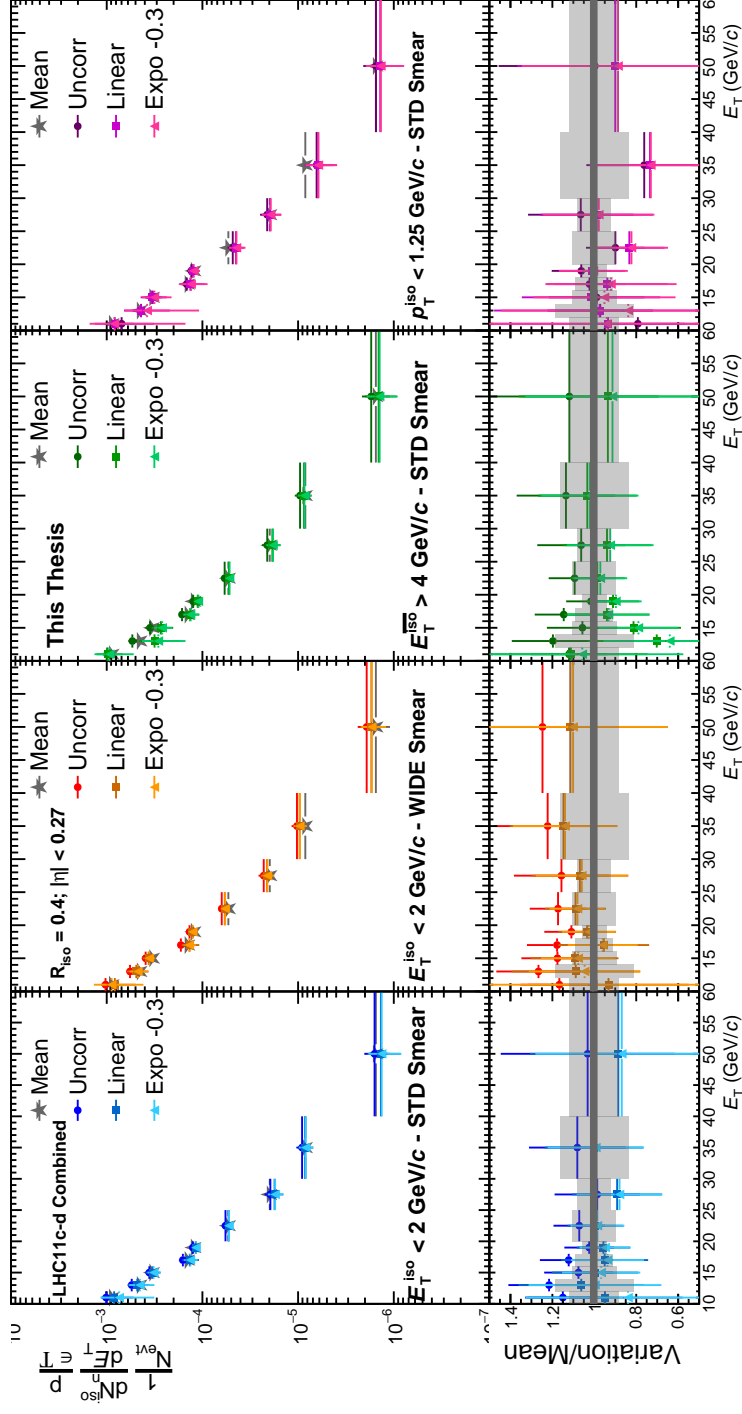


Figure 6.31: Isolated photon yield, as a function of the transverse energy of the cluster. Each variations of the analysis strategy is shown in a different column and includes the results obtained from all corrections of the α factor from different modelling of the double ratio (upper panels). The mean corrected yield (grey stars) is also shown, for comparison. The lower panels show the ratio of the different corrected yields with respect to the computed mean spectrum, while the standard deviation, as a function of the transverse energy, is shown in light grey boxes.

6.6. Total systematic uncertainty for the measurement

E_T (GeV/ c)	Error (%)
10 - 12	16
12 - 14	13
14 - 16	9.6
16 - 18	8.2
18 - 20	5.5
20 - 60	15

Table 6.3: Systematic uncertainty, as a function of the transverse energy of the cluster, relative to the different isolation probability for narrow and wide clusters in data and simulated samples.

In order to quantify the systematic uncertainty associated with the modelling of the isolation energy, it has been decided to not use the bin-by-bin value obtained for the standard deviation as defined in Eq. 6.10, but to use a systematic uncertainty that changes smoothly with the reconstructed E_T of the cluster.

The reasons for this decision is a significant contribution, already at $E_T = 20$ GeV/ c , of the background from merged decay photon clusters to the narrow cluster distribution; moreover, starting at the same energy, the standard deviation is mainly driven by the variations to the isolated photon yield obtained with a different estimator of the in-cone activity (rightmost plots in Fig. 6.31). The values chosen for the systematic uncertainty, as a function of the transverse energy of the cluster, are reported in table 6.3.

6.6 Total systematic uncertainty for the measurement

In Sections 6.2, 6.3 and 6.4 we discussed the systematic uncertainties related to triggering, the choice of the mixing ratio for simulated signal and background sample, and the various selections for candidate cluster used in

our analysis strategy. All the variation discussed are independent of the isolation probability discrepancy between data and MC. The total systematic uncertainty for our measurement of isolated photon production in pp collisions at $\sqrt{s} = 7$ TeV is calculated by adding the contributions from all the variations to the analysis strategy discussed in this chapter. Table 6.4 shows a summary of all systematic uncertainties, including the computation of the total systematic uncertainty introduced in our measurement by the analysis strategy adopted. The total systematic uncertainty has been obtained by summing quadratically the single contributions via

$$\frac{\Delta N_{\gamma}^{iso}(E_T)}{N_{\gamma}^{iso}} = \sqrt{\sum_k \left(\frac{\Delta N_{\gamma}^{iso,k}}{N_{\gamma}^{iso,k}} \right)^2} = \sqrt{\sum_k \varepsilon^2}, \quad (6.11)$$

with k being the systematic uncertainty source number in table 6.4

The main source of uncertainty in our analysis is associated with a non-perfect modelling, in the simulated sample, of both the σ_{long}^2 and E_T^{iso} distributions.

The causes of the non-perfect agreement between the σ_{long}^2 distributions computed in the simulated and data samples has been discussed in Sections 5.9.1 and 6.5.1.

The mismatch in the modelling, in simulated sample, of the in-cone activity from charged particles obtained in data has been discussed in Section 6.5.3, while the neutral component of the isolation activity for candidate clusters is well reproduced in the simulated sample, as shown in Fig. 6.27. Because investigations on the cause for this mismatch have been inconclusive, we decided to include the result of the analysis performed with a different estimator for the in-cone activity in the computation of the systematic uncertainty related to different isolation probability for narrow and wide clusters in data and simulated sample.

6.6. Total systematic uncertainty for the measurement

Variation number (k)	Source of systematic uncertainty	Error ε (%)
1	Trigger efficiency	7
2	Run by run trigger stability	5.1
3	Mixing Ratio	1 - 6
4	Charged Particle Veto	1
5	Narrow cluster definition	5
6	Wide cluster definition	1
7	Data-MC difference in the isolation probability for narrow and wide clusters	5.5 - 16
Total Error		14 - 22

Table 6.4: Single contributions to the total systematic uncertainty estimated for the measurement of isolated photon production cross section in pp collisions at $\sqrt{s} = 7$ TeV, performed with the ALICE experiment. In case the uncertainty is not constant throughout the whole transverse energy range investigated, the minimum and maximum values are given.

Chapter 7

Results, outlook and conclusions

In Section 6.1 we have computed the transverse momentum distribution of the corrected yield for isolated photons in pp collisions. We have also presented the result of the studies on the systematic uncertainties for the analysis method used.

In order to compute the differential cross section for isolated photon production

$$\frac{d^2\sigma_{\gamma^{iso}}}{dE_T d\eta} = \frac{1}{\mathcal{L} \cdot \epsilon_{trig}} \frac{d^2N_{\gamma^{iso}}}{dE_T d\eta} \quad (7.1)$$

we need to determine the integrated luminosity \mathcal{L} corresponding to the analysed data sample, while the estimate for the trigger efficiency ϵ_{trig} has been presented in Section 6.2.2.

In this chapter we will present the result of these studies and compute the isolated photon cross section in pp collisions at $\sqrt{s} = 7$ TeV, measured at mid-rapidity with the ALICE experiment. A comparison between our measurement and the state-of-the-art NLO pQCD calculations will also be shown.

We will conclude the chapter presenting the results of studies on the x_2 sensitivity for the measurement of isolated photon production at forward rapidity, performed on Monte Carlo simulations produced with the Pythia event generator.

7.1 Integrated luminosity

The integrated luminosity \mathcal{L} for our data sample triggered with EMCAL with the L0 trigger algorithm can be expressed as:

$$\mathcal{L} = \frac{N_{\text{evt,EMC-L0}} \cdot P_{\text{EMC-L0}}}{\sigma_{\text{EMC-L0}}} \quad (7.2)$$

where $P_{\text{EMC-L0}}$ is the average number of interactions that satisfy the EMC-L0 trigger condition in bunch crossings that have at least one interaction that satisfies the same trigger condition. Because of the extremely low rate with which the EMC-L0 trigger class is fired, we can safely assume $P_{\text{EMC-L0}} = 1$.

Since the EMC-L0 trigger class is a subsample of the Minimum Bias (MB) trigger class, we can express the $\sigma_{\text{EMC-L0}}$ in terms of the measured cross section for the INT7 (V0AND) trigger class σ_{INT7} which, for the 2011 period of data taking, is $\sigma_{\text{INT7}}^{2011} = 53.7 \pm 1.9^{\text{syst}}$ mb [45]. We get:

$$\sigma_{\text{EMC-L0}} = \frac{\sigma_{\text{INT7}} \cdot R_{\text{EMC-L0}}}{P_{\text{INT7}}} \quad (7.3)$$

where we used P_{INT7} , which is analogous to $P_{\text{EMC-L0}}$, and $R_{\text{EMC-L0}}$, called *trigger rejection factor* (TRF), which indicates the rate at which the EMC-L0 trigger class is fired with respect to the INT7 trigger class.

Substituting Eq. 7.3 into Eq. 7.2 we get:

$$\mathcal{L} = \frac{P_{\text{INT7}} \cdot N_{\text{evt,EMC-L0}}}{\sigma_{\text{INT7}} \cdot R_{\text{EMC-L0}}}. \quad (7.4)$$

To obtain the value for P_{INT7} , we recall that the probability for 2 protons from two bunches travelling in opposite directions, to have an inelastic collision is very low and independent of all other potential collisions. The probability of having n collisions in one bunch crossing follows a Poisson distribution:

$$P(X = n) = e^{-\mu_{\text{INT7}}} \cdot \frac{\mu_{\text{INT7}}^n}{n!} \quad (7.5)$$

With μ_{INT7} the expectation value for the average number of pp collisions per bunch crossing.

7.1. Integrated luminosity

We note that the INT7 trigger fires when at least one interaction takes place in a bunch crossing. The probability for this is:

$$P(X \geq 1) = P(\mathbb{N}) - P(X = 0) = 1 - e^{-\mu_{\text{INT7}}}. \quad (7.6)$$

To determine μ_{INT7} , we relate the probability $P(X \geq 1)$ to the trigger scaler for the INT7 class of triggers:

$$R_{L0b, \text{INT7}} = P(X \geq 1) \cdot nBC_{\text{orbit}} \cdot \nu_{\text{LHC}}, \quad (7.7)$$

where nBC_{orbit} is the number of unmasked bunches per orbit and ν_{LHC} is the LHC frequency, which for the periods analysed was 11245 Hz.

Combining Eqs. 7.6 and 7.7 we get:

$$\mu_{\text{INT7}} = -\ln \left(1 - \frac{R_{L0b, \text{INT7}}}{nBC_{\text{orb}} \cdot \nu_{\text{LHC}}} \right) \approx 0.3. \quad (7.8)$$

The term P_{INT7} in Eq. 7.4 represents the number of interaction per bunch crossing with at least one collision:

$$P_{\text{INT7}} = \frac{\sum_{n=1}^{\infty} n \cdot P(X = n)}{\sum_{n=1}^{\infty} P(X = n)} = \frac{\mu_{\text{INT7}}}{P(X \geq 1)} = \frac{\mu_{\text{INT7}}}{1 - e^{-\mu_{\text{INT7}}}}. \quad (7.9)$$

Using the value for μ_{INT7} from Eq. 7.8 we can compute P_{INT7} . Figure 7.1 shows $R_{\text{EMC-L0}}/P_{\text{INT7}}$ for different runs from both periods analysed. The weighted average over both periods is $3.4 \pm 0.2^{\text{stat}} \cdot 10^{-4}$.

Using the values for $\sigma_{\text{INT7}}^{2011}$, $R_{\text{EMC-L0}}$, P_{INT7} and the number of events satisfying the event selection criteria explained in Section 4.1.1, which totals to $N_{\text{evt}} \approx 8.8$ M, we can compute the value for the integrated luminosity for the two periods combined:

$$\mathcal{L} = 473 \pm 22^{\text{stat}} \pm 17^{\text{syst}} \text{ nb}^{-1}. \quad (7.10)$$

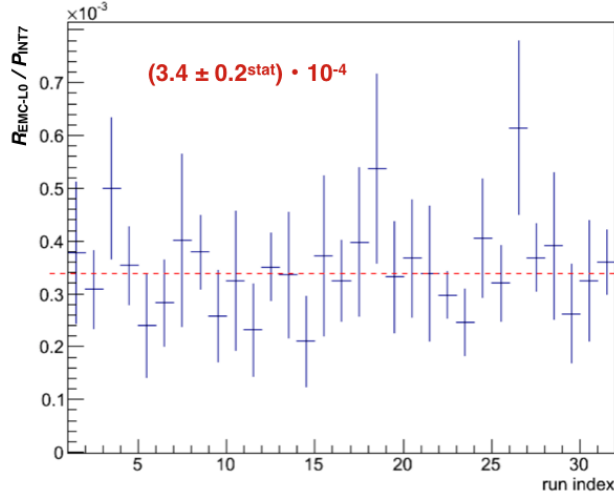


Figure 7.1: Value of the ratio $R_{\text{EMC-L0}}/P_{\text{INT7}}$ for different run numbers from both periods analysed. The red line is a fit to the points with a constant function. Figure taken from [43].

7.2 Isolated photon cross section and comparison with NLO pQCD calculation.

We now have all the ingredients needed to compute the cross section for isolated photon production in pp collision at \sqrt{s} 7 TeV. Figure 7.2 shows the results obtained when using the selection criterion for isolated photons in Eq.5.1:

$$E_{\text{T}}^{\text{iso}} = \left(\sum_{i=0}^{N_{\text{clust}}} E_{\text{T}}^i + \sum_{j=0}^{N_{\text{track}}} p_{\text{T}}^j \right) < 2 \text{ GeV}/c$$

The central values for the measured cross section have been computed from the mean corrected yield presented in Fig. 6.31 (grey stars), with the systematic uncertainties reported in table 6.4. The theoretical prediction of the cross section for isolated photon production is also shown for comparison. The NLO pQCD calculation is obtained by simulating pp collisions at $\sqrt{s} = 7$ TeV with the JETPHOX Monte Carlo partonic event generator. The PDF chosen for these calculations are the CTEQ6.6M [46], with the BFGII [47] parton-to-photon fragmentation functions. Further details on these calcula-

7.3. Isolated photons at forward rapidity

tions, as well as details on the determination of the uncertainties in these calculations, can be found in [43].

Since the theoretical calculation is performed at the partonic level, a correction factor

$$\bar{C}(E_T) \simeq \epsilon_{\text{iso}} = \frac{N_{\gamma}^{\text{iso},\text{had}}(E_T)}{N_{\gamma}^{\text{iso},\text{part}}(E_T)}. \quad (7.11)$$

is needed to compare the calculation with the measured cross section. The superscripts *had* and *part* in Eq. 7.11 stand for *hadron* and *parton*, while $N_{\gamma}^{\text{iso},x}$ indicates the number of isolated photons, as a function of the transverse energy, respectively after hadronisation and at parton level. The value found is $\bar{C} = 0.88 \pm 0.01$ [43], independent of the cluster transverse energy and takes into account the effect of both hadronisation and underlying event on the theoretical isolated direct photon cross section prediction. The bottom panel of Fig. 7.2 shows the ratio Data / Theory for the isolated prompt photon production cross section. We can see that the measured cross section agrees within 1.5σ with the JETPHOX prediction throughout the whole transverse energy range investigated.

Isolated photon measurements are also available from ATLAS [48] and CMS [49] in the same system. These measurements use slightly different experimental cuts, so may not be directly comparable. We can, however, judge their respective agreement with JETPHOX calculations using the corresponding cuts. These comparisons show very similar behaviour of the different data sets when compared to the theoretical calculations. The measurement presented in Fig. 7.2 extends, towards lower values, the transverse energy range explored at LHC.

7.3 Isolated photons at forward rapidity

In Chapter 2 we have discussed the production mechanisms for direct photons in hadronic collisions, showing (Fig. 2.6) the different predictions for the direct photon production rate at forward rapidity in p-Pb collisions obtained assuming the validity of collinear pQCD (+ shadowing) and the CGC, respectively.

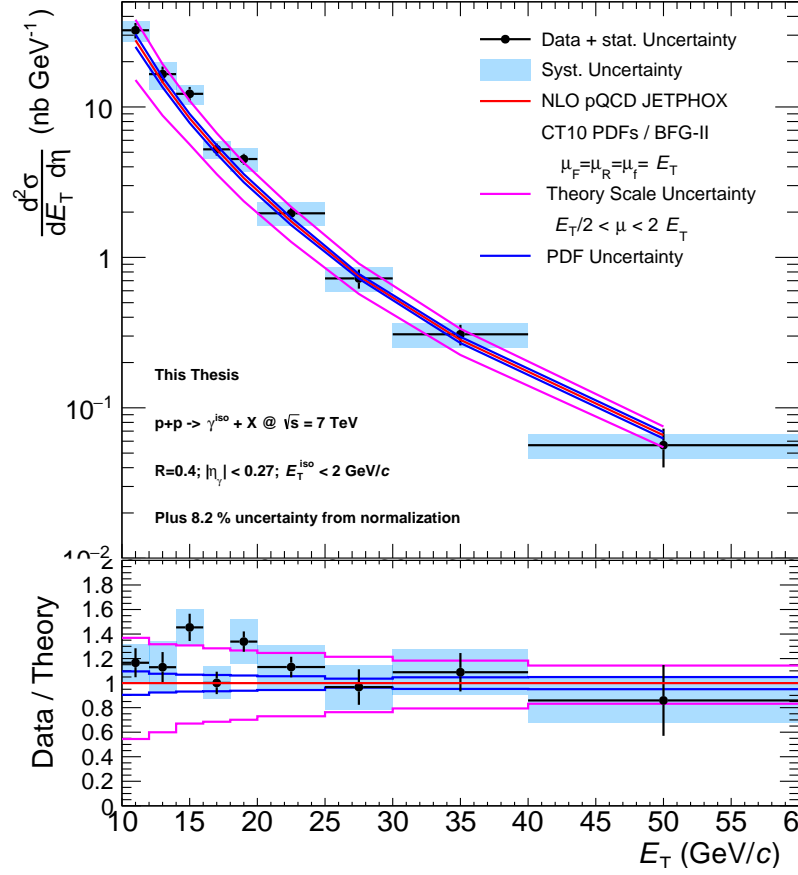


Figure 7.2: Isolated photon cross section measured by the ALICE experiment in pp collision at $\sqrt{s} = 7$ TeV and NLO pQCD calculation obtained with JETPHOX MC generator(upper panel). The lower panel shows the ratio data over theory.

7.3. Isolated photons at forward rapidity

The measurement of the nuclear modification factor (R_{pPb}^{iso}) in p-Pb collisions is an extremely sensitive experimental observable which would allow us to discriminate between the two theoretical scenarios. In general, inclusive measurements for single particle cross section at pseudorapidities $|\eta| \geq 1.0$ have already provided a first hint of saturation effects at RHIC energies with the STAR [50] and BRAHMS [51] experiments. These collaborations measured the production rates for neutral pions (STAR) and charged hadrons (BRAHMS) in the pseudorapidity range $3 \leq \eta \leq 4$ and up to $\eta = 3.2$, respectively, in d-Au collisions with a collision energy of 200 GeV per nucleon pair. It should be noted, however, that such measurements have an inherent uncertainty due to the hadronisation phase, which changes the relation between the final state observables and the initial state characteristics of the hard scattering. On the other hand, photons do not undergo hadronisation and preserve information about the hard scattering.

To further study the hard scattering, and to better distinguish the pQCD + shadowing and the CGC scenarios, one can perform angular correlation measurements, with a (direct) photon as trigger particle.

In the collinear pQCD formalism, the hard scattering (at LO) is a $2 \rightarrow 2$ process, while in the CGC framework this is modified, for high parton density of the projectile, with a $2 \rightarrow \text{many}$ or $2 \rightarrow 1$ process becoming relevant as the saturation regime sets in.

Saturation of the gluon PDF should start to become evident at RHIC energies for values of $x \sim 10^{-4}$ at $Q^2 \approx 1 - 2 \text{ GeV}^2$ [52], so probing the gluon PDF at $x_2 \leq 10^{-4}$ means being able to tell if the saturation regime has already set in and, if so, to study its properties and evolution in the explored range.

To study the feasibility of measurements of single particle production rates and azimuthal particle correlation in pp collisions, we have analysed the results of simulated pp collisions, using the Pythia (version 6.4.4) event generator, which includes only LO diagrams for the isolated photon production mechanism. In the next sections we will present the results of these studies.

7.3.1 $2 \rightarrow 2$ process kinematics

Before illustrating the analysis steps and the results of this study, it is good to recall a few relations between the final state particle properties and the initial state ones. This way we will already have the pQCD prediction for our sensitivity to the momentum fraction x carried by the gluon from the nucleus in one of the two processes involved.

In the initial state, two partons with 4-momentum components $(x_1, 0, 0, x_1)\frac{\sqrt{s}}{2}$ and $(x_2, 0, 0, -x_2)\frac{\sqrt{s}}{2}$ ¹ collide. Since in general $x_1 \neq x_2$, the centre of momentum (CM) of the two colliding partons is not motionless, $p_{CM}^{i,Lab} \neq 0$, where the superscript *Lab* means in the reference frame of the two colliding protons. This translates to a non-zero rapidity for the CM of the two partons. The expression for the CM pseudorapidity is

$$\eta_{CM}^i = \frac{1}{2} \ln \frac{1 + \beta_{CM}}{1 - \beta_{CM}} \quad (7.12)$$

By substituting the expression $\beta = \frac{|p|}{E} = \frac{x_1 - x_2}{x_1 + x_2}$, we obtain

$$\eta_{CM}^i = \frac{1}{2} \log \frac{x_1}{x_2} \quad (7.13)$$

Since the two partons propagate with equal and opposite pseudorapidity η^* in the CM frame, we conclude that η^* is half the pseudorapidity difference between the two outgoing partons (identified with 3 and 4) in the lab frame:

$$\eta_3^* = -\eta_4^* = \eta^{*,fin} = \frac{\eta_3 - \eta_4}{2} \quad (7.14)$$

The pseudorapidity η_{CM}^f of the CM in the laboratory frame is

$$\eta_{CM}^f = \frac{\eta_3 + \eta_4}{2}. \quad (7.15)$$

Using momentum conservation, we obtain:

$$\eta^i = \eta^f \Rightarrow \frac{x_1}{x_2} = e^{(\eta_3 + \eta_4)} \quad (7.16)$$

¹Here we use the Minkowski metric with signature (1,-1,-1,-1) and we assume all partons are massless.

7.3. Isolated photons at forward rapidity

The collision energy of the two incoming partons is

$$\sqrt{\hat{s}} = \sqrt{x_1 x_2 S} \quad (7.17)$$

while the energy of the two partons in the centre of momentum frame is:

$$E^* = p_T \cosh \eta^*. \quad (7.18)$$

By substituting the value of $\eta^{*,fin}$ from Eq. 7.14 in 7.18 and combining this with Eq. 7.17 we obtain:

$$\sqrt{x_1 x_2 S} = 2p_T \cosh \left(\frac{\eta_3 - \eta_4}{2} \right). \quad (7.19)$$

The equations 7.16 and 7.19 give us the relations between the fractions of momentum of the two colliding partons and the rapidities of the two outgoing partons.

$$\begin{cases} \eta^i = \eta^f \Rightarrow \frac{x_1}{x_2} = e^{(\eta_3 + \eta_4)} \\ \sqrt{x_1 x_2 S} = 2p_T \cosh \left(\frac{\eta_3 - \eta_4}{2} \right). \end{cases} \quad (7.20)$$

By solving this simple second order system of equations for x_1 and x_2 separately, and defining $x_T = 2p_T/\sqrt{s}$ we have:

$$\begin{cases} x_1 = \frac{p_T}{\sqrt{s}}(e^{\eta_3} + e^{\eta_4}) = \frac{x_T}{2}(e^{\eta_3} + e^{\eta_4}) \\ x_2 = \frac{p_T}{\sqrt{s}}(e^{-\eta_3} + e^{-\eta_4}) = \frac{x_T}{2}(e^{-\eta_3} + e^{-\eta_4}) \end{cases} \quad (7.21)$$

By selecting a pair of back-to-back jets we can calculate the initial state momentum fractions from the final state rapidities of the outgoing partons and by selecting the rapidity range we automatically restrict our study to the desired x range. Moreover, substituting one of the two jets with a trigger photon we also improve our resolution in x_2 .

The above calculations clearly state that the higher the rapidity considered, the lower the gluon x_2 values explored, allowing to study the dependence of the production rate of direct photons on the momentum fraction x_2 .

The reason to choose direct photons as a trigger particle is to preserve the kinematic relation between the colliding partons and the scattering products, without complications introduced by the hadronisation process.

In Ch. 2 we have discussed the two main processes involved in the production of prompt photons in hadron-hadron collisions. A complication to the simple L.O. scenario is given by quark-to-photon fragmentation and (gluon) Bremsstrahlung processes, discussed at the end of Section 2.1.

7.3.2 Forward π^0 and prompt photon production at LHC.

In the previous section we have pointed to studies performed by the STAR collaboration at RHIC on inclusive π^0 production rate measurement and π^0 -charged hadrons forward-mid back-to-back production. Such measurements hinted to the onset of the saturation regime for values of $x \sim 10^{-3} - 10^{-2}$. To explore the x_2 range that is probed by a similar measurement at the LHC, we have studied pp collisions at $\sqrt{s} = 8.8$ TeV simulated with Pythia. To estimate the range in momentum fraction x_2 that can be probed, we have forced the colliding parton to interact either via qg-Compton scattering or the annihilation process

The event sample used to study the production of neutral pions at forward rapidity consist of ten million Minimum Bias pp collisions at $\sqrt{s} = 8.8$ TeV, where, in each event, only the highest p_T neutral pion, satisfying the condition $p_T^{\pi^0} \geq 1.5$ GeV/c and produced in the η range $[4., 5.]$, is selected; at the same time, the distribution of the x_2 value given by Pythia for all events fulfilling these requirements, is stored.

The same study has been performed on single direct photon measurements. After simulating one million p-p collisions with only the annihilation and qg-Compton process activated (MSEL=10), we stored the x_2 values given by Pythia for each event that presented a direct photon produced in the same η range as before and with a $p_T \geq 4$ GeV/c.

Figure 7.3 shows the x_2 distributions, given by Pythia, obtained from both simulated samples. Both distributions have been normalised for the cross section of the involved process(es) as given by Pythia and for the number of simulated events. It is clear that even without placing constraints on the

7.3. Isolated photons at forward rapidity

presence of an azimuthally correlated particle, and so without a complete constrain of the kinematic of the hard scattering, the resolving power in x_2 when detecting prompt photons is far better than the case of single pion detection: the x_2 mean value is roughly 10^3 times smaller and the range explored is restricted to very small values when compared to the single pion detection.

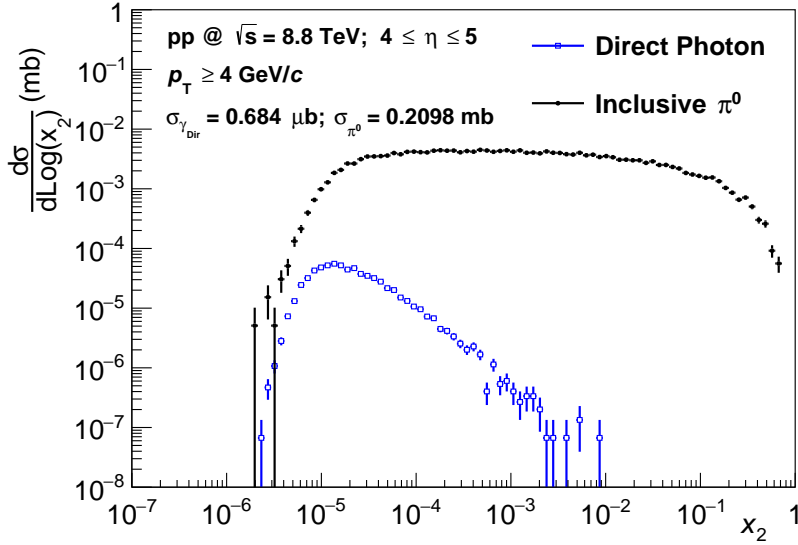


Figure 7.3: Cross section, as a function of the x_2 variable, for inclusive prompt photon (blue markers) and inclusive π^0 production (black points) at pseudo-rapidity $4 \leq \eta \leq 5$. The distributions shown refer to simulated pp collisions at $\sqrt{s} = 8.8$ TeV. The x_2 range explored in case of prompt photons is far narrower and limited only to low values.

7.3.3 Inclusive prompt photons vs. Drell-Yan process

Another process that might be sensitive to saturation effects is the Drell-Yan lepton-pair production. The LHCb collaboration [53], has shown preliminary results of the measurement of lepton pairs from Drell-Yan processes at forward rapidity, theoretically probing x_2 values as low as 10^{-5} . To study the feasibility of such a measurement at LHC, we have simulated ten million

pp collisions at $\sqrt{s} = 8.8$ TeV, where only the Drell-Yan production process was allowed. In order to make a comparison with the capabilities of the ALICE(/LHCb) experiment to measure leptons produced in such processes, we have restricted our analysis only to Drell-Yan pairs produced at pseudorapidity $4 \leq \eta \leq 5$.

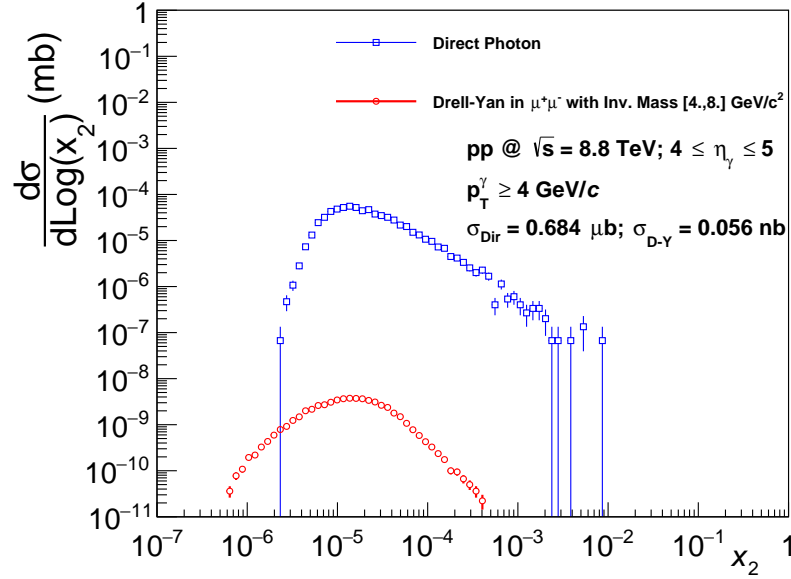


Figure 7.4: Cross section, as a function of the x_2 variable, for Drell-Yan lepton pair production at $4 \leq \eta \leq 5$ (red points) in pp collisions at $\sqrt{s} = 8.8$ TeV. For comparison, the cross section for isolated photon production from Fig. 7.3 (blue points) is also shown.

Only Drell-Yan virtual photons decaying into muon pairs with an invariant mass of $4 \leq M \leq 9$ GeV/ c^2 have been considered. This selection criterion was used in order to better compare these calculations with realistic data analysis, where one would like to study such process in a mass range where the contribution to the invariant mass spectrum from J/Ψ and Υ resonances is negligible. No constraints on the rapidity of the two muons have been applied.

Figure 7.4 shows the comparison of the x_2 distributions obtained by analysing both the sample of pp collisions producing a direct photon and

the sample where Drell-Yan process was forced to be simulated. In this case we need to scale all distributions for the probability of the only decay channel considered in $q\bar{q} \rightarrow \gamma^* \rightarrow \mu^+\mu^-$ to happen. This is taken to be the same as the branching ratio of the decay $Z^0 \rightarrow \mu^+\mu^-$, so $BR_{\gamma^* \rightarrow \mu^+\mu^-} = 3.37\%$.

Although measuring Drell-Yan pairs produced at large rapidity is quite promising from the point of view of the sensitivity to the x_2 of the process, the cross section for this process is very small and when compared with the prompt photon cross section and considering the expected integrated luminosity at LHC, it is clear that studies of this process will be even more complicated, because of the limited sample size, than direct photon measurements at the ALICE experiment.

7.3.4 Forward-mid γ -hadron azimuthal correlations

As argued in Section 7.3.1, measuring back-to-back azimuthally correlated particles with respect to the trigger photon provides complete information on the kinematic of the hard scattering, and thus accesses information on the initial state of the collision. Due to the fragmentation and hadronisation processes, more than one particle (jets) originate from the outgoing parton.

In our study, we restricted the correlation to two particles, assuming that the rapidity of a single particle is a sufficiently accurate measure of the parton rapidity. The trigger particle (π^0) must fulfil the kinematic requirements presented in Section 7.3.2, while the kinematic cuts applied for the associated particles are: $4 \leq \eta^{assoc} \leq 5$, $1.5 \text{ GeV}/c \leq p_T^{assoc} \leq p_T^{trig}$, $\Delta\varphi_i = \varphi_{trig} - \varphi_i \in [8/9\pi, 10/9\pi]$. If, within an event, more than one neutral pion satisfies the selection criteria for trigger (or associated) particle, only the highest p_T particle is taken into account.

The values of x_2 of the partons involved in the collision, given by Pythia, are stored and compared to the x_2 values computed via the second relation in Eq. 7.21.

Figure 7.5 shows the comparison of the x_2 distributions for events with two correlated π^0 in the two cases.

The mean value of the peak of the *computed* distribution is shifted towards lower values of x_2 , as expected since the p_T of the hadron is always

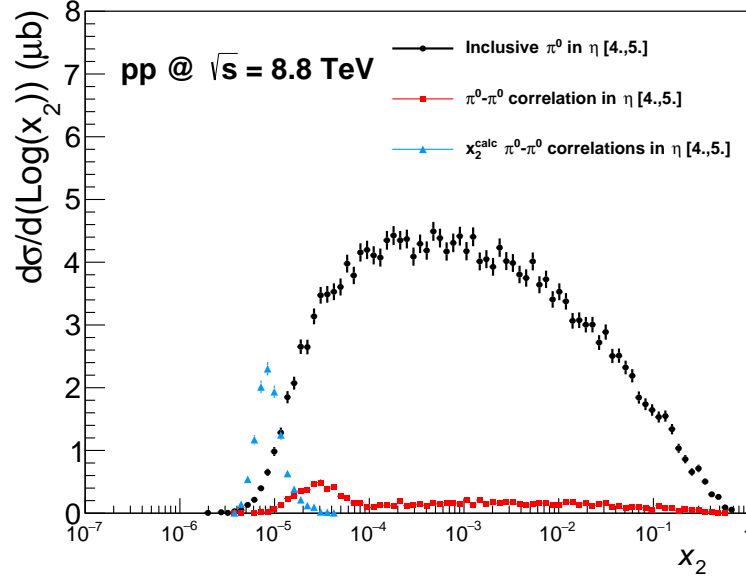


Figure 7.5: Distributions of the x_2 variable for the measurement of inclusive π^0 production at forward rapidity (black markers), production of back-to-back (in azimuth) π^0 pairs (red markers), and (for the latter case) comparison of the dependence of the cross section on the x_2 variable, as calculated according to second relation in Eq. 7.21 (azure triangles).

smaller than (or at best equal to) the p_T of the parton. Moreover the true x_2 distribution taken from Pythia presents a quite long tail towards higher values of x_2 . This makes it almost impossible to achieve a well defined x_2 sensitivity with such studies.

As a next step, we have selected direct photons, fulfilling the same requirements presented in Section 7.3.2, as trigger particle, and studied the azimuthal distribution of neutral pions with respect to it. Figure 7.6 shows the results of this study. Both the calculated and the Pythia x_2 distributions are shown. It is clear that the x_2 distribution for $\gamma - \pi^0$ angular correlation allows for more precise determination of the x_2 variable. Moreover, the computed distribution is almost identical to the true distribution obtained from Pythia.

To investigate a broader range of the x_2 distribution, we have selected

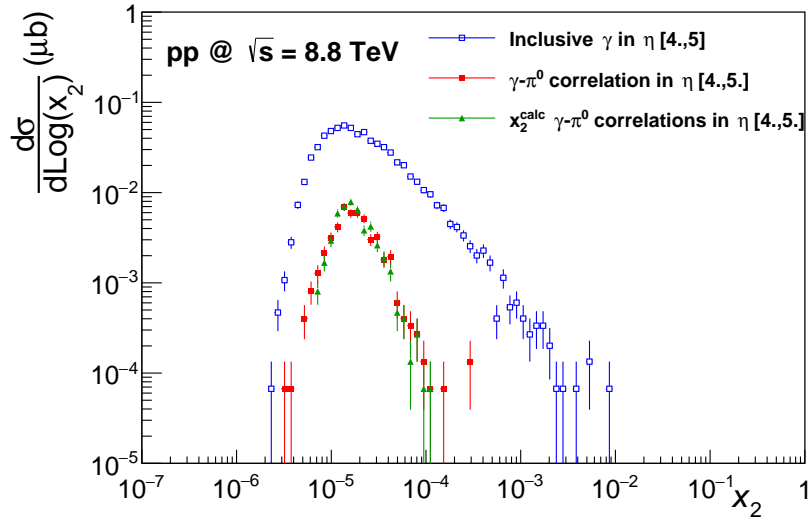


Figure 7.6: x_2 distributions for the measurement of inclusive prompt photon production at forward rapidity (blue markers), $\gamma - \pi^0$ production back-to-back in the azimuthal angle (red markers), and (for the latter case) comparison of the dependence of the cross section on the x_2 variable, as calculated according to second relation in Eq. 7.21 (green triangles).

photons produced at $4 \leq \eta_\gamma \leq 5$, and paired them with π^0 produced back-to-back in the azimuthal angle, as a function of the rapidity of the associated π^0 .

Figure 7.7 (left) shows the result of this study. On the left we observe the expected correlation in the upper half of the figure, but in addition, there is a band at $x_2 \sim 10^{-5}$, with no dependence on the rapidity of the π^0 . The main reason for this structure in the Pythia model is that partons occasionally undergo multiple parton interactions (MPI). The right panel of Fig. 7.7 shows the result of the same analysis when multiple interactions are switched off in Pythia. The low-density structure is much reduced, with some remaining events outside the expected correlation band due to initial and final state shower processes and 3 or 4-jets events. There is, however, no experimental possibility to suppress MPI, so this serves only to better understand the underlying processes, and, in the following, we will not make use of these results.

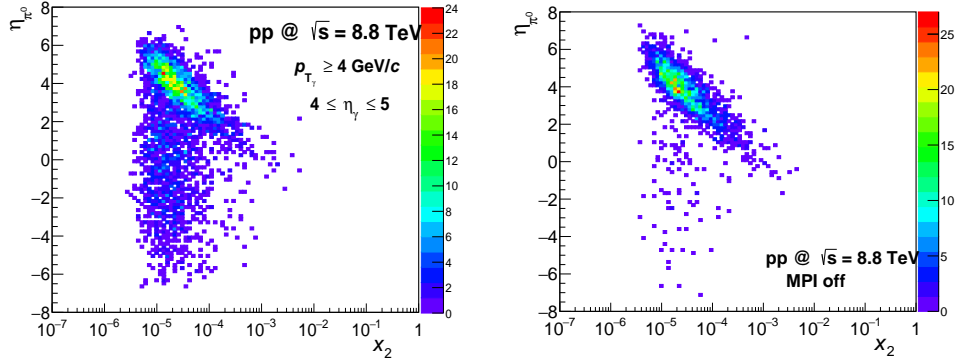


Figure 7.7: Distribution, in the $\eta_{\pi^0} - x_2$ plane, of events where a prompt photon is produced at pseudorapidity $4 \leq \eta \leq 5$ in association with a neutral pion with $1.5 \text{ GeV}/c \leq p_T^{\pi^0} \leq p_T^\gamma$ produced back-to-back, in the transverse plane, with respect to the direction of the prompt photon. The left panel refers to simulated events where all physical processes known in a pp collision are allowed to happen, while the right panel shows the same distribution of events when the probability for multi-parton interactions is set to zero.

By projecting the distributions in Fig. 7.7 (left panel) on the x axis,

7.4. A Forward Calorimeter: FoCal

we can study the x_2 sensitivity when measuring direct photons produced at forward rapidity associated with neutral pions produced at different rapidities and back-to-back in the transverse plane. The expected relation between x_2 and η_{π^0} is seen, especially if the associated particle is also produced at forward rapidity. Still, the presence of a tail for all remaining distributions indicates that $\gamma - \pi^0$ azimuthal coincidence measurement performed with forward emitted isolated photon as trigger and mid-rapidity emitted pions as associate particles is not very useful to constrain the x_2 of the measurement.

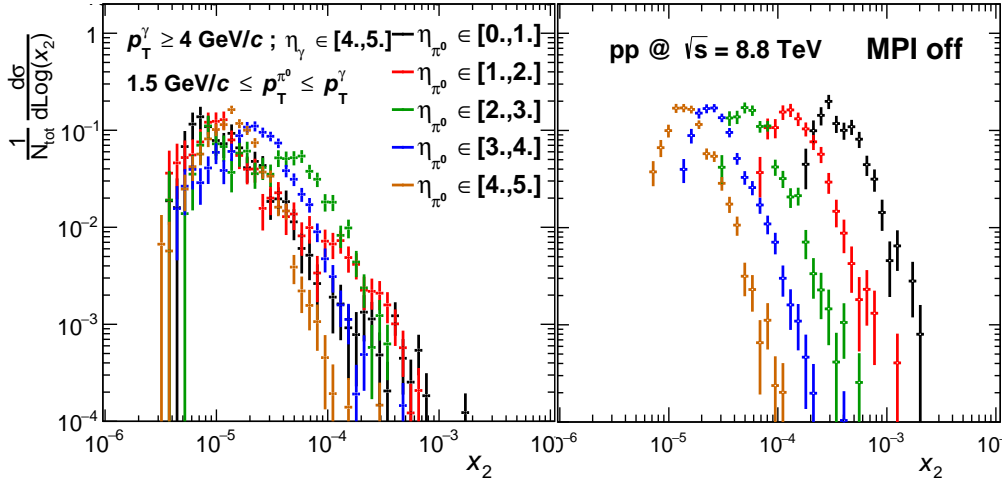


Figure 7.8: Cross section, as a function of the x_2 variable, for different selections of the pseudorapidity of the neutral pion associated to a prompt photon emitted at $4 \leq \eta \leq 5$. The distributions are obtained by projecting on the x-axis the distributions shown on the left panel of Fig. 7.7. The right panel shows the same projections from the right panel of Fig. 7.7.

7.4 A Forward Calorimeter: FoCal

In the previous section we have presented studies on x_2 sensitivity from different measurements involving the production of photons at large rapidity. The main experimental complication in measuring such photons is the discrimination of clusters from single photons from those produced by photons

from π^0 decay with small opening angle. One way to solve this problem is to change the design of calorimeters, improving the transverse granularity. This is the main motivation for the proposed high-resolution forward calorimeter (FoCal) as an upgrade of the ALICE experiment.

The proposed FoCal consists of one Si-W sampling electromagnetic calorimeter (FoCal-E) and a hadronic one (FoCal-H) to be placed 7 meters away from the centre of the ALICE experiment. The proposed design of the EM calorimeter features a structure of silicon pad layers, with two high-granularity layers of silicon pixels (with $30\ \mu\text{m}$ granularity) inserted at 5 and 10 radiation lengths for $\gamma - \pi^0$ discrimination. Monolithic Active Pixel Sensors (MAPS) are a promising technology for this application, since they can be produced relatively cheaply, using standard CMOS technology. MAPS technology is also used for the ITS upgrade detector.

Studies performed on a prototype detector entirely made of MAPS and obtained by analysing the data collected in test beams at DESY, PS and SPS facilities have shown [54] that a two shower separation capability down to few mm, with energy resolution only slightly worse than those observed in analog calorimeters based on Si-W technology can be achieved, as shown in Fig. 7.9.

Therefore single photon measurement can be performed down to $p_T^\gamma \sim 4\ \text{GeV}/c$. At lower p_T , the reconstruction efficiency for π^0 becomes smaller than 0.9 because the two photon separation for pairs of photons from highly asymmetric configurations for the π^0 decay is limited by statistical fluctuations in the shower development. Considering also that the signal to background ratio γ/π^0 is about 0.01, a background rejection efficiency lower than 90 % leads to a very large uncertainty in the measurement of prompt photons produced at large rapidity.

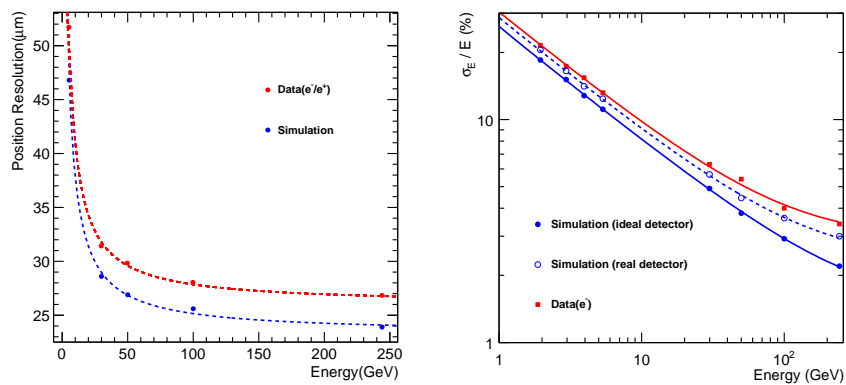


Figure 7.9: Position (left) and energy resolution (right) as a function of the measured energy of the particle. These results have been obtained by analysing the data collected at DESY and SPS facilities with the FoCal prototype detector. The computed quantities are compared to the results from simulations with different configurations of the modelled detector. Figures taken from [54].

Chapter 8

Conclusions and outlook

High energy heavy ion collisions allow us to study a new state of hadronic matter in which the quarks and gluons are not confined within hadrons, but instead form a plasma, called QGP. The interactions in this hot and dense medium are characterised by low momentum transfer, making it impossible to study such collisions within the theoretical framework of perturbation theory. Instead, the hot and dense medium produced in collisions involving nuclei is characterised via the modifications of observables, directly related to macroscopic characteristics of the system, with respect to the simpler, and better understood, case of collisions between protons.

This characterisation of the hot and dense medium produced in high energy heavy ion collisions is based on the modelling the nucleus as an incoherent superposition of nucleons within the nuclei, with the addition of corrections that account for the presence of a nuclear medium (nPDFs).

It is possible to validate our modelling of the nucleus by performing collisions between protons and nuclei and by measuring the same observables and their modifications with respect to the behaviour predicted by the perturbation theory, which also relies on the knowledge derived from a number of *Deep Inelastic Scattering* (DIS) experiments, including the pioneering measurements at SLAC National Laboratory and the systematic measurements at HERA.

Still, since DIS are mediated via the exchange of a (virtual) photon (or a W/Z boson), the cross section measurements are (mainly) sensitive to the

PDFs (and their evolution with Q^2) of charged partons within the proton, while the gluon PDF have been obtained from the evolution of the PDFs with Q^2 and from momentum sum rules.

The evolution of the PDFs with the energy scale is described by the DGLAP equations in a wide range of Q^2 and x , except for low values of these quantities, where the DGLAP approach leads to inaccurate results. In this regime, the evolution of the PDFs is described by the BFKL equations. Both with the DGLAP and with the BFKL equations we observe a growth of the gluon parton density functions, which, in order to have a consistent description of hadrons, must be balanced by non-linear effects allowing for the gluon PDFs to *saturate* in the low x - low Q^2 regime.

One of the proposed description of the gluon fields in the saturation regime is the *Colour Glass Condensate* (CGC), an effective field theory based on the assumption that, in the low- x low- Q^2 regime, the gluon field in hadrons can be described with a mean field approximation, and hence treated as a classical, coherent field.

Different experimental observables can help test the validity of the CGC EFT; among these, the production of direct photons in hadronic collisions at forward rapidity is one of the cleanest and, at the same time, most challenging observables.

In fact prompt photons are produced, in hadronic collisions, mainly via quark gluon-Compton scattering, which, at Leading Order (LO), allows to directly relate the energy and momentum of the final state photon to those of the initial state gluon.

The experimental challenges in measuring photons at hadron colliders comes mainly from the huge background of decay photons from neutral hadrons, especially neutral pions, for which the production cross section overwhelms that of direct photons by several order of magnitudes. Another complication comes from the fact that at collision energies as high as those available at LHC, NLO contributions to the direct photon production become sizeable. For these reasons, one relies on isolation techniques to discriminate the true direct (isolated) photon signal from the irreducible background.

Presently, ALICE can only perform measurement of isolated photon at mid-rapidity, which is the main topic of this thesis. The presented results

refer to the data collected in 2011, where the EMC-L0 trigger algorithm was used to reduce the amount of non-interesting events stored.

In Ch. 4 we have discussed the analysis strategy that allowed us, starting from the response of the single EMCal modules, to obtain a sample of clusters using the V1 clustering algorithm (3.5.2). After rejecting clusters based on simple QA requirements and after rejecting clusters produced by charged particles via a CPV procedure, we are left with a neutral cluster sample, which is made of direct, fragmentation and decay photons. For each of these clusters, we have measured the total hadronic activity in a cone of radius 0.4, which allows us to select isolated clusters, while a selection on the transverse width of the cluster is used to disentangle clusters from single photons and clusters produced by pairs of photons from decay of neutral pions in symmetric configurations. Using these two quantities simultaneously it is possible to define a region of narrow (one particle) - isolated ($E_T < 2$ GeV/ c) clusters. The estimation of the contamination in this sample is done via the ABCD method, a single side band method applied simultaneously on the two discriminating variables.

In order to validate the analysis procedure used in data, the same analysis has been carried out on a combination of different MC simulations used to model the signal and the background in pp collisions at $\sqrt{s} = 7$ TeV. This strategy allows also for a refinement of the selection criteria and for tests of the applicability of the ABCD method, which yields a correction factor for the estimated contamination in data to account for non-factorising effects between the two variables. The correction for the reconstruction and identification efficiency of isolated photons is computed on the simulated sample that models our signal. These corrections are presented in Ch. 5.

The study of the systematic uncertainties in our analysis procedure has been carried out and presented in Ch. 6, while the computation of the cross section for prompt photon production at mid-rapidity is presented in Ch. 7. A comparison with the state-of-the-art NLO pQCD calculation performed with JETPHOX is also presented. The measured cross section is in agreement with the theoretical prediction within uncertainties.

The presented measurement will serve as a baseline for similar studies in different colliding systems and at different rapidities.

While the current measurement is sensitive to parton distributions at moderate values of x , possible future measurements at forward rapidity allows to access the low- x regime, where the onset of saturation phenomena must occur. In Ch. 7 we also presented studies of the sensitivity to the kinematic variable x of the measurement of isolated photon production at large rapidity in hadron-hadron collisions. These studies prove that by performing gamma-hadron azimuthal correlations it is possible to study, at LHC, the low- Q^2 low- x regime, where saturation is predicted to become relevant.

Despite the outstanding sensitivity to low values of the kinematic variable x , measurements of photon produced at large rapidity present an experimental complication regarding the identification of two-photon showers from neutral pion decay. To overcome this, a new design for an electromagnetic calorimeter is needed, which we briefly describe as an outlook.

The solution proposed as an upgrade for ALICE is a Si-W sampling calorimeter, based on silicon pad sensors, with the addition of two high granularity layers of silicon pixels inserted at 5 and 10 radiation lengths for $\gamma - \pi^0$ discrimination. The proposed technology for this application is the Monolithic Active Pixel Sensors (MAPS). Results from the analysis of data collected at the DESY, PS and SPS facilities with a prototype detector entirely made of MAPS have demonstrated the capability to separate two particle with a separation distance down to a few mm and an energy resolution only slightly worse than those observed in analog calorimeters based on Si-W technology.

Appendix A

Isolation activity and underlying event estimators.

Estimators for the hadronic activity

In this section we will list the possible experimental strategies to adopt in order to estimate the isolation activity inside a cone of radius R around the direction of the candidate neutral cluster. In fact, while from a theoretical point of view the estimation of the isolation activity reduces to a simple sum of the transverse momentum of the particles produced within the isolation cone, a measurement of the isolation activity can be done according to different strategies.

Similar arguments apply also to the estimation of the underlying event, which becomes really important when performing measurements in collisions involving high energy nuclei. In fact, in case of heavy ion collisions, the probability for soft processes increase, greatly biasing, without a detailed study, the outcome of the measurement of the specific observable related to the hard scattering.

EMCal cells

One of the most immediate course of action to estimate the activity around the photon candidate and the underlying event associated with the hard scattering is to consider the response of the single elements of the de-

tector used to identify candidates and to measure their energy. Because the measurement presented in this analysis is performed via the calorimetric method, this translates in a sum of the energy measured by the single cells of EMCal.

This strategy, which relies exclusively on the response of the detector, can be pursued only when a hadronic calorimeter is also used. In fact, the response of a typical electromagnetic calorimeter to charged hadrons (like π^\pm) quickly saturates for energy of the particle above few hundreds MeV, therefore making the estimation of the total hadronic activity unreliable without the use of an hadronic calorimeter, designed specifically to measure the energy of hadrons (although some limitations in the energy resolution have still to be overcome nowadays).

EMCal clusters

A second strategy to estimate the neutral hadronic activity associated with the hard scattering is to group the response of the calorimeter in clusters; by clustering the response of the single modules of the detector, a more direct correspondence between the energy of the cluster and that of the electron (or photon) that released its energy into the active material of the calorimeter can be found.

Still, the estimation of the hadronic activity produced by charged particles will be biased by the same limitations, assuming same experimental conditions, discussed in the previous section; moreover the clustering procedure might be biased by the inclusion, in the resulting cluster, of the energy produced by the interaction of charged hadrons produced close to a neutral particle.

In order to reduce the bias in the determination of the energy of neutral particles due to these phenomena, a two-way correction is used in the analysis. On the one hand, clusters matched to a track, reconstructed with the tracking system and extrapolated to half of the EMCal depth, have been excluded from the computation of the neutral hadronic activity and the transverse momentum of the matching track is added to the hadronic activity from charged particles instead. On the other hand, if a track is reconstructed close

to a cluster, without exactly matching it, the maximum between the track transverse momentum and the cluster energy is calculated and added either to the neutral or to the charged hadronic activity.

Tracks

A third way, used in the analysis as a systematic check related to the isolation probability for narrow and wide clusters, is to consider only charged particles in the estimation of the hadronic activity of a hard scattering. The advantage of this strategy is that it relies on the response of multiple detectors (ITS, TPC, TOF, TRD), therefore obtaining larger precision for the estimation of the hadronic activity of charged particles. The disadvantage of this strategy is an almost entire lack of knowledge of the neutral hadronic activity which, on the one hand, helps in partially recovering the factorisation of the two variables used in the ABCD method, and, on the other hand, does not take into account the hadronisation processes that lead to the production of neutral particles.

Isolation Methods

In section 2.3 we have listed the different isolation criteria developed through the years, without going into much details for all the criteria not used in the analysis. The criterion adopted in the analysis ensures a smaller contamination in the narrow neutral cluster sample for high energies of the candidate clusters. This feature is mandatory to be able to compare our measurement with those carried out by the ATLAS and CMS collaborations.

In the following of this section we will explain quickly the isolation criteria listed in the main text but not used in the analysis, explaining the main advantages and disadvantages for each of them.

Threshold proportional Trigger cluster E_T

A "democratic" method to identify isolated neutral cluster is to use an isolation criterion with a threshold for isolated clusters proportional to the

transverse energy E_T of the candidate cluster. The use of this method results in an almost absent dependence of the raw yield of isolated narrow clusters on the transverse energy, compensated, as presented in Fig.5.5 by an increasing identification efficiency for isolated photon with the reconstructed energy of the cluster, at the cost of increasing the contamination in the isolated narrow neutral cluster region which, for high values of the reconstructed energy of the cluster, is already underestimated with the ABCD method applied to the σ_{long}^2 and E_T^{iso} variables.

Frixione's Method

A relatively new method to identify isolated photon is the Frixione's method, named after the designer of such method, Stefano Frixione. This method can be seen as a refinement of the method discussed in the previous section, where the determination of the threshold for the isolation activity allowed within a cone depends not only on the energy of the candidate cluster, but also on the radius of the isolation cone.

In particular, given a maximum radius for the isolation cone R_{max} , the isolation activity must satisfy the criterion

$$E_T^{\text{iso}}(R_i) \leq \kappa(R_i) E_T^{\text{iso}}(R_{\text{max}}, E_T^{\text{clust}})$$

for all $R_i < R_{\text{max}}$, for the candidate to be identified as isolated. Such a condition on the hadronic activity inside the isolation cone(s) has two effects: on the one hand it ensures the infrared and collinear safety of the computed (and measured) isolated photon yield at all orders of the perturbative expansion; on the other hand, the condition on the hadronic activity relies heavily on the knowledge of the parton-to-photon fragmentation and hadronisation mechanisms.

Therefore, the isolated photon yield measured with the use of the Frixione's method and obtained with the help of simulated collisions might be biased by the lack of a deep understanding of the hadronisation mechanisms, while performing the same measurement with a data-driven strategy might require very large statistics just to tune the optimal values for κ on a control sample.

Estimators for the Underlying Event

When performing hadron-hadron collisions the particles produced by the hard scattering are not the only ones detected. In fact, because hadrons consist of quarks and gluons interacting with each other, a collision between hadrons is usually characterised both by a hard scattering between partons and soft processes like multi-parton interactions (and many others) which usually lead to the production of low momentum particles. Observables related to the hard scattering are linked, via energy-momentum conservation laws to the state of hadrons before the collision, allowing to gain knowledge on the strong interaction between partons in the hadrons. On the other particles produced in soft processes are emitted isotropically, affecting the measurement of the observables from hard scatterings.

Therefore, a detailed study of the production of particles from soft processes, called Underlying Event (UE), is needed to estimate the bias introduced in our measurement. This usually translates in a measurement of the activity in a portion of the phase space where the production of particles from hard scatterings is excluded (for the particular event). In order not to be biased by fluctuations, one tends to maximize the portion of phase space used to determine the hadronic activity related to soft processes which is subtracted from the hadronic activity associated with the hard scattering.

In the following we will present the most popular methods used to estimate the contribution to the hadronic activity due to soft processes.

φ Band

After identifying the candidate cluster and its direction, and after defining the isolation cone, the estimation of the UE can be done by summing the hadronic activity in a band with same dimension as the isolation cone in η , $\eta_{\text{gamma}} - R_{\text{iso-cone}} \leq \eta_{\text{Band}} \leq \eta_{\text{gamma}} + R_{\text{iso-cone}}$, and extending on φ for all the length of the detector used: $\varphi_{\gamma} + R_{\text{iso-cone}} \leq \varphi_{\text{Band}} \leq \varphi_{\gamma} + \pi - R_{\text{iso-cone}}$.

Fig. A.1, left panel, illustrates the concept of a φ -band to estimate the UE. This method has two disadvantages: the limited portion of phase space usable to estimate the UE (with possible biases due to fluctuations) and the sensitivity to azimuthal correlation of particles produced in the hard

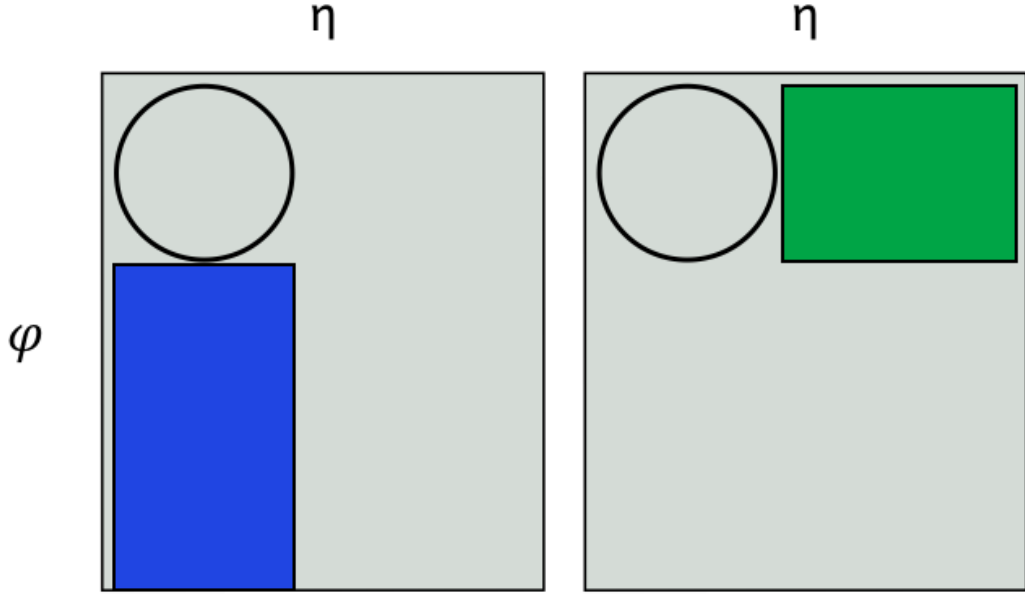


Figure A.1: Illustration of the φ -Band (left) and η -Band (right) methods for the estimation of the underlying event.

scattering.

η Band

The concept of an η band is similar to the concept of φ -band. The band extends in φ along the range $[\varphi_\gamma - R_{iso-cone}; \varphi_\gamma + R_{iso-cone}]$ while it extends in the η direction for all the length of the detector used. While using the η -Band method to estimate the UE solves the bias from the azimuthal correlation of particles from hard scatterings, usually the phase space available for the η -Band is even smaller than that of the φ -Band. Therefore, this method will, in principle, be more affected by fluctuations of the underlying event.

Perpendicular Cones

A different way to study the production of particles from soft processes is to measure the hadronic activity in cones perpendicular to the direction of the candidate. This technique has been first used in 2 particles correlation (or jets) analysis. In a $2 \rightarrow 2$ process the final state is usually characterised

by two high p_T , back-to-back in azimuth, jets, so the "safest" portion of the phase space to estimate the UE is a cone whose centre is orthogonal in φ to the direction of the jets (Fig. A.2, left panel). In our case, one of the two jets is substituted with the candidate photon with the isolation cone around it.

Although the radius of the cones might be different from that of the isolation cone of the candidate cluster, this method might be biased by the event-by-event fluctuations of the underlying event.

To overcome such bias, in the next section we present a further method for the estimate of the UE. This method combines ideas from the previous three methods and has been adopted in all recent analysis involving the reconstruction of jets and the measurement of observables related to jets production.

Full TPC

We have seen how the previous methods are, more or less, affected by event-by-event fluctuations of the underlying event and/or correlations from particles produced in the hard scattering. With the Full-TPC method, all the available phase space, except that associated with the production of particles from the hard scattering, is used to estimate the UE. Fig.A.2, right panel, illustrates the portion of the phase space associated to the underlying event (red) and that associated with the trigger particle (white); the blue band along η in the opposite direction in φ with respect to the trigger particle is taken out to take into account the longitudinal momentum unbalance typical of hadron-hadron collisions.

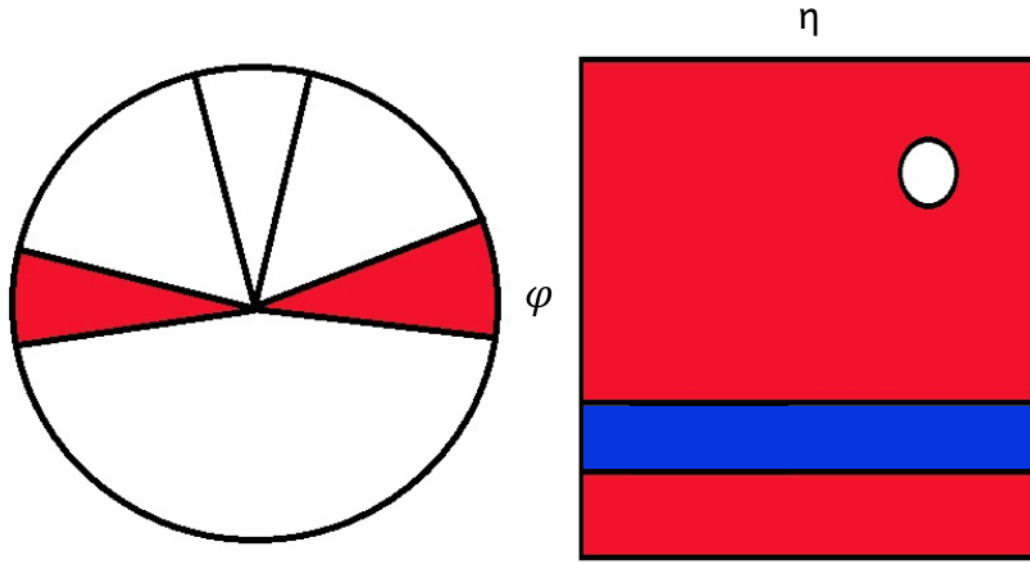


Figure A.2: Illustration of the perpendicular cones (left) and Full-TPC (right) methods for the estimation of the underlying event.

Appendix B

Studies on isolation probability in data and simulated sample.

In this Appendix we will include all plots that have lead to the results presented in Section 6.5.

As already explained in the main body of the thesis, to estimate the background contribution in the narrow cluster region we computed the double ratio in Eq. 5.5,

$$\frac{(N_{\text{clust}}^{\text{iso}}/\overline{N_{\text{clust}}^{\text{iso}}})_{\text{Data}}}{(N_{\text{clust}}^{\text{iso}}/\overline{N_{\text{clust}}^{\text{iso}}})_{\text{MC}}} \quad (\text{B.1})$$

as a function of the transverse width of the cluster (quantified via the σ_{long}^2 parameter) for all variations presented in 6.5 and separately for the two periods analysed. We then performed fits to the points assuming a linear dependence of the double ratio on the σ_{long}^2 variable first, and then assuming an exponential dependence. To model the dependence of the isolation probability for background cluster in data and MC, only the wide cluster region (defined in Section 5.5), which includes the \mathcal{B} and \mathcal{D} regions, are fitted with the above-mentioned functions.

To estimate the contribution of background clusters in the narrow cluster region we have extrapolated the results of the fit at low values of the σ_{long}^2 parameter.

The effect of this procedure on the factor α , used to correct the contamination estimated in our data samples via the ABCD method, have already

been presented in Eqs. 6.4 and 6.5. To obtain the modification to the α factor shown in Figs.6.23, 6.25, and 6.30 we have performed fits to the points of the various double ratio with a linear function and an exponential with slope fixed to -0.3.

Support plots for Section 6.5.1

Figure 6.21 shows the result, for the LHC11c period, of the fit to the points of the double ratio, for different transverse energies of the cluster, with a linear function; the σ_{long}^2 distributions used to compute the double ratio have been obtained according to the smearing procedure explained in Section 5.9.1 (STD Smearing).

Figure 6.22 shows instead the result for the LHC11d period, obtained with a different smearing function (WIDE smearing) and assuming an exponential dependence of the double ratio on the transverse width of the clusters.

In this section we present the results of the remaining 6 variations used to estimate the different isolation probability for narrow and wide background clusters in data and MC.

Results for the LHC11c period

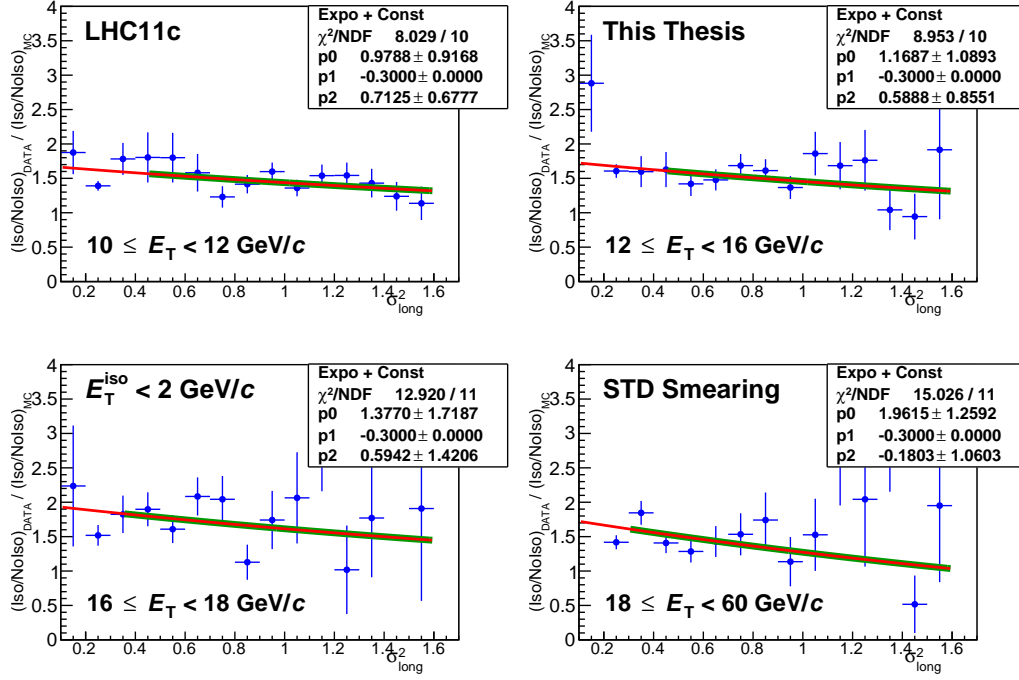


Figure B.1: Double ratio, as defined in Eq. 5.5, as a function of the σ_{long}^2 parameter for four different ranges of the reconstructed transverse energy of clusters. The results shown refer to the LHC11c sub period, with the STD smearing of the σ_{long}^2 parameter. The green lines represent the result of the fit to the points performed with an exponentially falling function, while the blue lines show the estimation of the background in the narrow cluster region via extrapolation of the resulting fitting function.

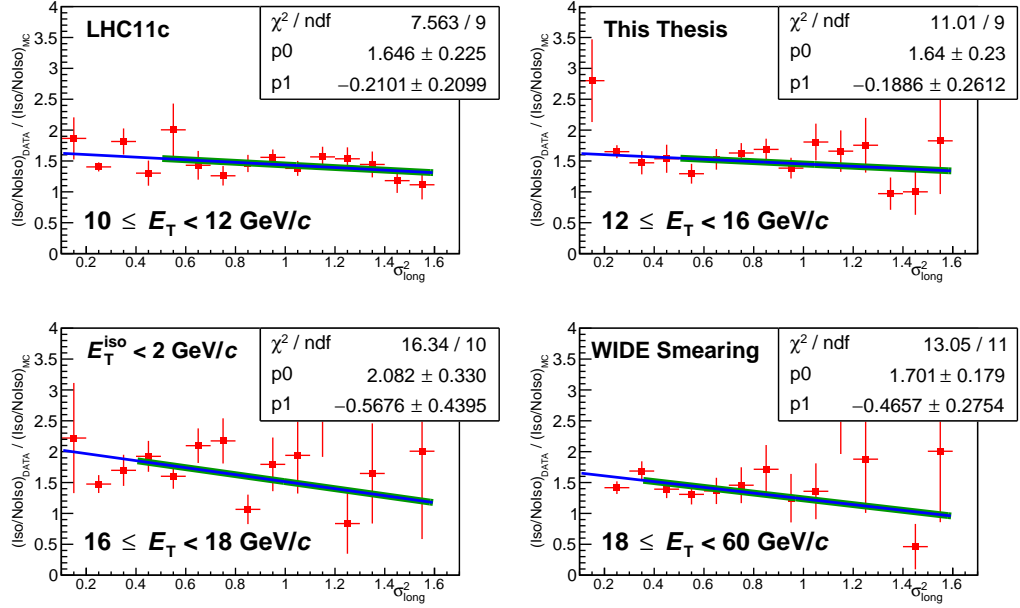
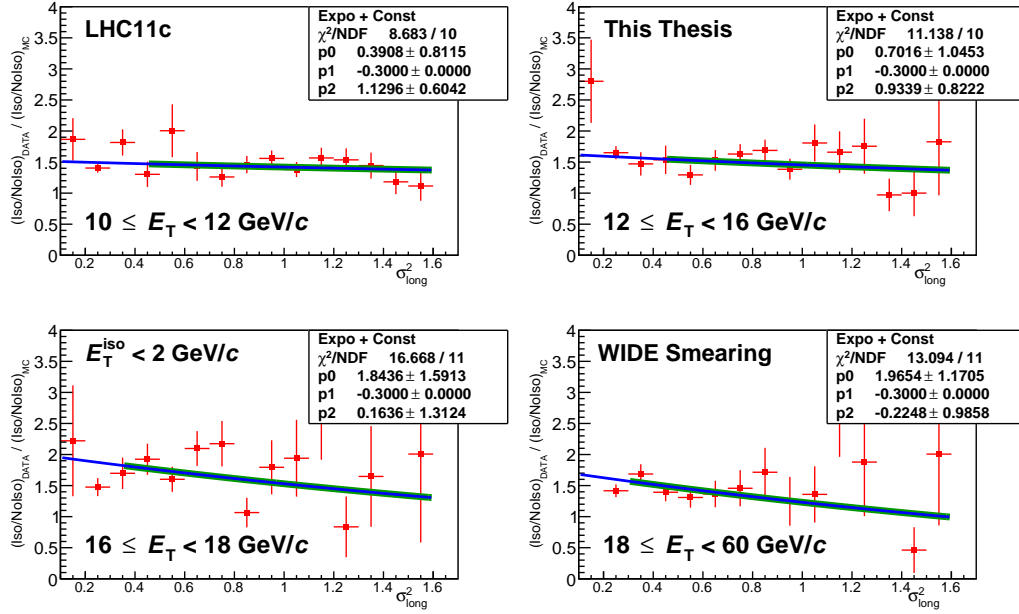


Figure B.2: Double ratio, as defined in Eq. 5.5, as a function of the σ_{long}^2 parameter for four different ranges of the reconstructed transverse energy of clusters. The results shown refer to the LHC11c sub period, with the WIDE smearing of the σ_{long}^2 parameter. The green lines represent the result of the fit to the points performed with a polynomial function of the first order, while the red lines show the estimation of the background in the narrow cluster region via extrapolation of the resulting fitting function.



Results for the LHC11d period

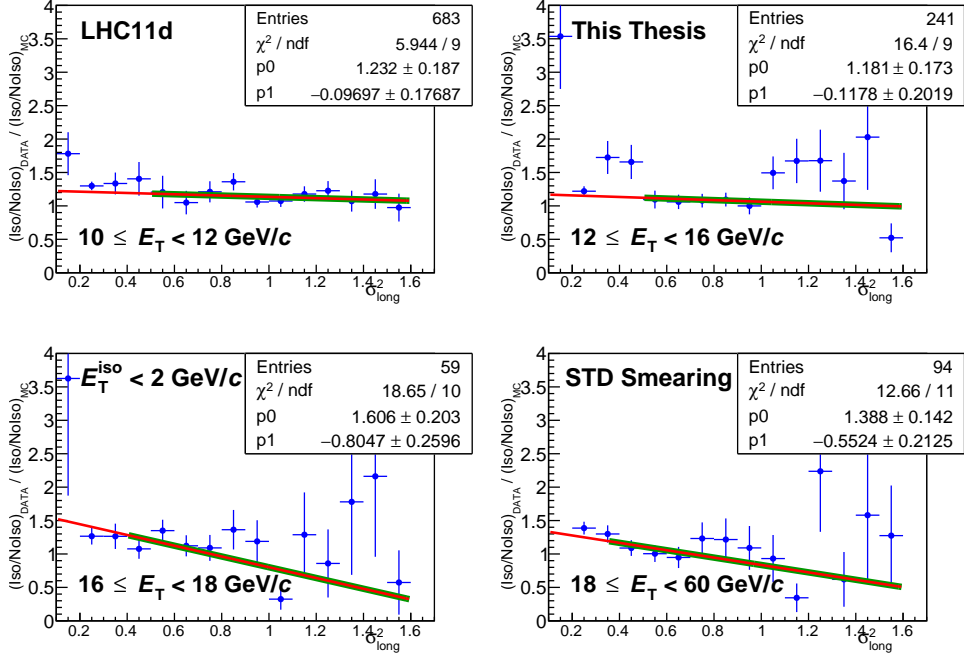


Figure B.4: Double ratio, as defined in Eq. 5.5, as a function of the σ_{long}^2 parameter for four different ranges of the reconstructed transverse energy of clusters. The results shown refer to the LHC11d sub period, with the STD smearing of the σ_{long}^2 parameter. The green lines represent the result of the fit to the points performed with a polynomial function of the first order, while the red lines show the estimation of the background in the narrow cluster region via extrapolation of the resulting fitting function.

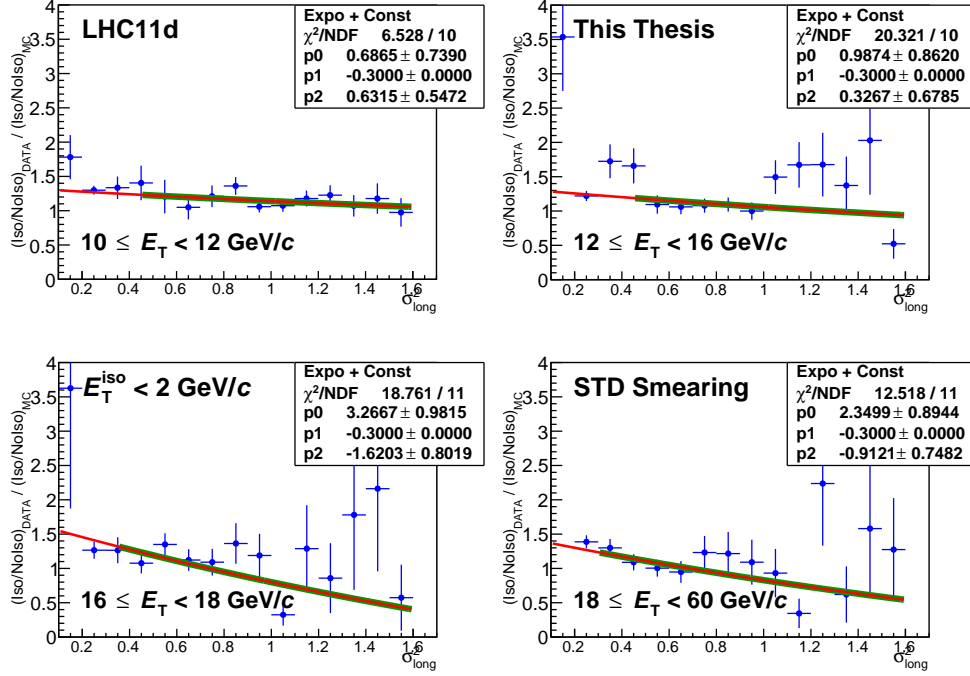


Figure B.5: Double ratio, as defined in Eq. 5.5, as a function of the σ_{long}^2 parameter for four different ranges of the reconstructed transverse energy of clusters. The results shown refer to the LHC11d sub period, with the STD smearing of the σ_{long}^2 parameter. The green lines represent the result of the fit to the points performed with an exponentially falling function, while the blue lines show the estimation of the background in the narrow cluster region via extrapolation of the resulting fitting function.

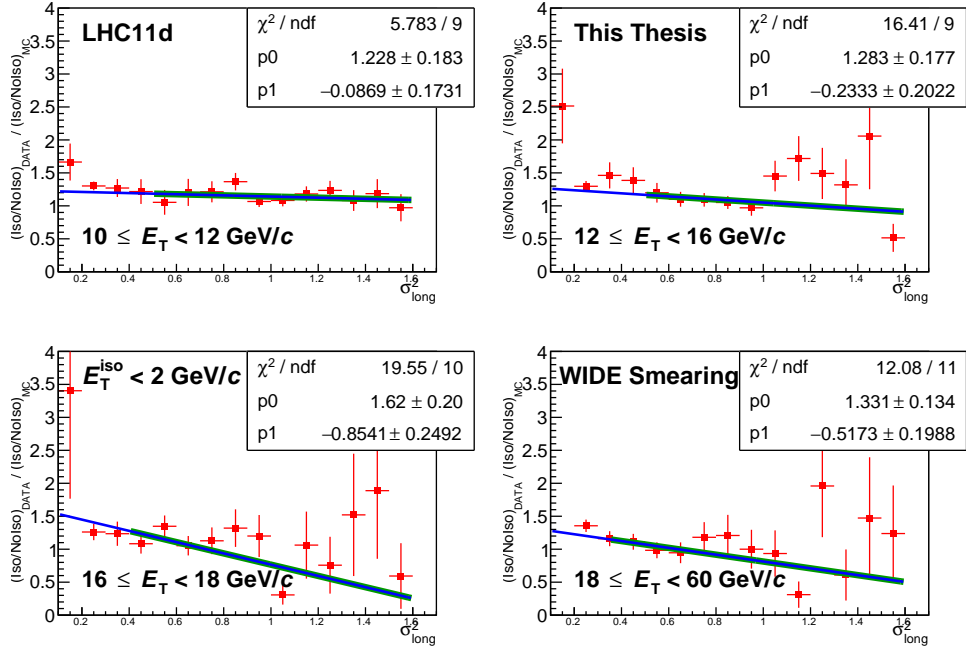


Figure B.6: Double ratio, as defined in Eq. 5.5, as a function of the σ_{long}^2 parameter for four different ranges of the reconstructed transverse energy of clusters. The results shown refer to the LHC11d sub period, with the WIDE smearing of the σ_{long}^2 parameter. The green lines represent the result of the fit to the points performed with a polynomial function of the first order, while the red lines show the estimation of the background in the narrow cluster region via extrapolation of the resulting fitting function.

Support plots for Section 6.5.2

In Section 6.5.2 we have seen (Fig. 6.24) that modifying the selection criterion for non isolated clusters helps reducing the offset of the double ratio in Eq. B.1 with respect to one. This translates in a reduction of the discrepancy in the isolation probability for narrow and wide clusters in data and MC. In this section we will present, separately per period, the studies that allowed us to reach the correction for the α factor shown in Fig. 6.25.

Results for the LHC11c period

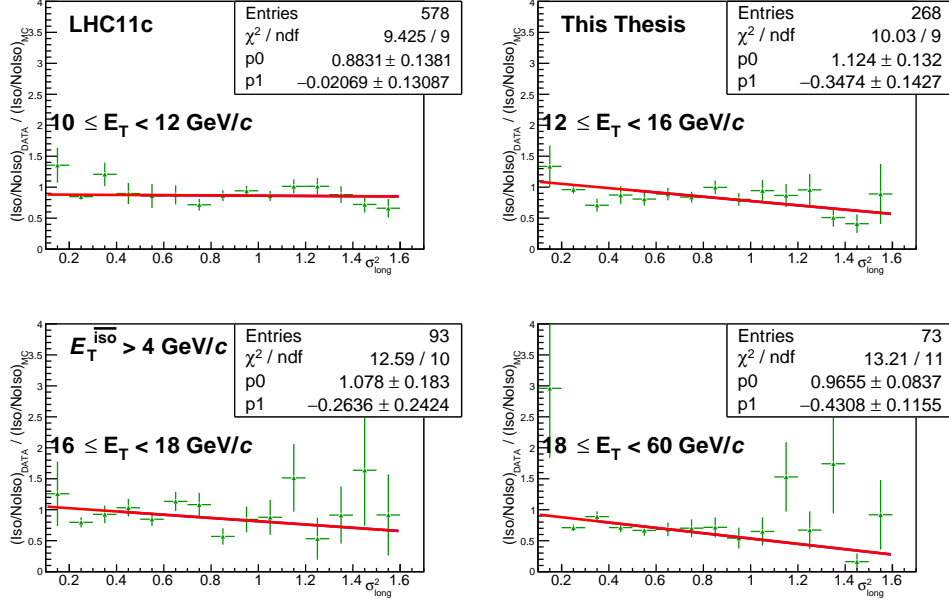


Figure B.7: Double ratio, as defined in Eq. 5.5, as a function of the σ_{long}^2 parameter for four different ranges of the reconstructed transverse energy of clusters. The results shown refer to the LHC11c sub period, with the STD smearing of the σ_{long}^2 parameter. The green lines represent the result of the fit to the points performed with a polynomial function of the first order, while the red lines show the estimation of the background in the narrow cluster region via extrapolation of the resulting fitting function.

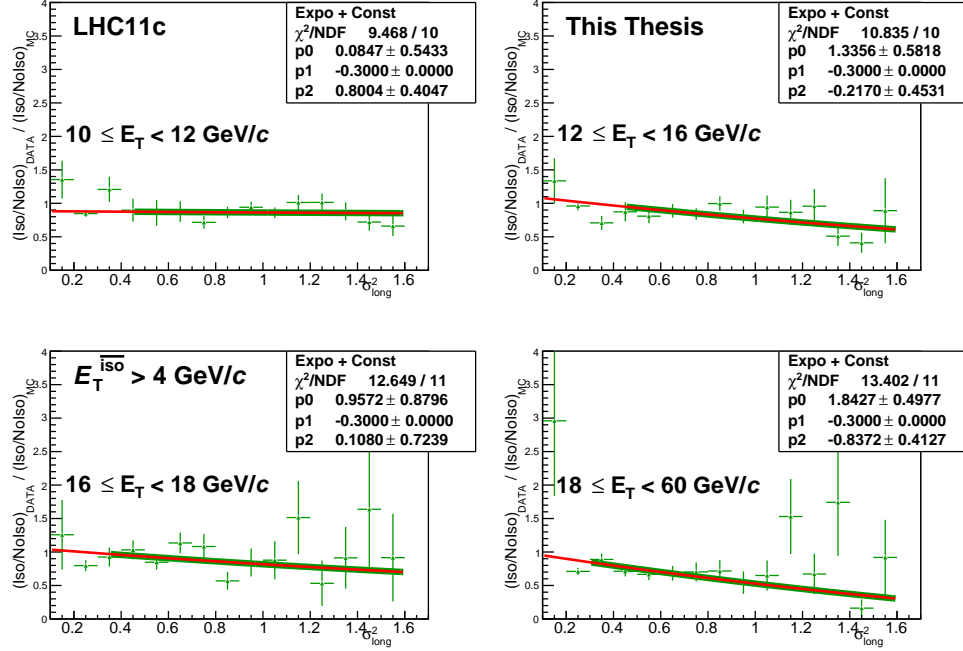


Figure B.8: Double ratio, as defined in Eq. 5.5, as a function of the σ_{long}^2 parameter for four different ranges of the reconstructed transverse energy of clusters. The results shown refer to the LHC11c sub period, with the STD smearing of the σ_{long}^2 parameter. The green lines represent the result of the fit to the points performed with an exponentially falling function, while the blue lines show the estimation of the background in the narrow cluster region via extrapolation of the resulting fitting function.

Results for the LHC11d period

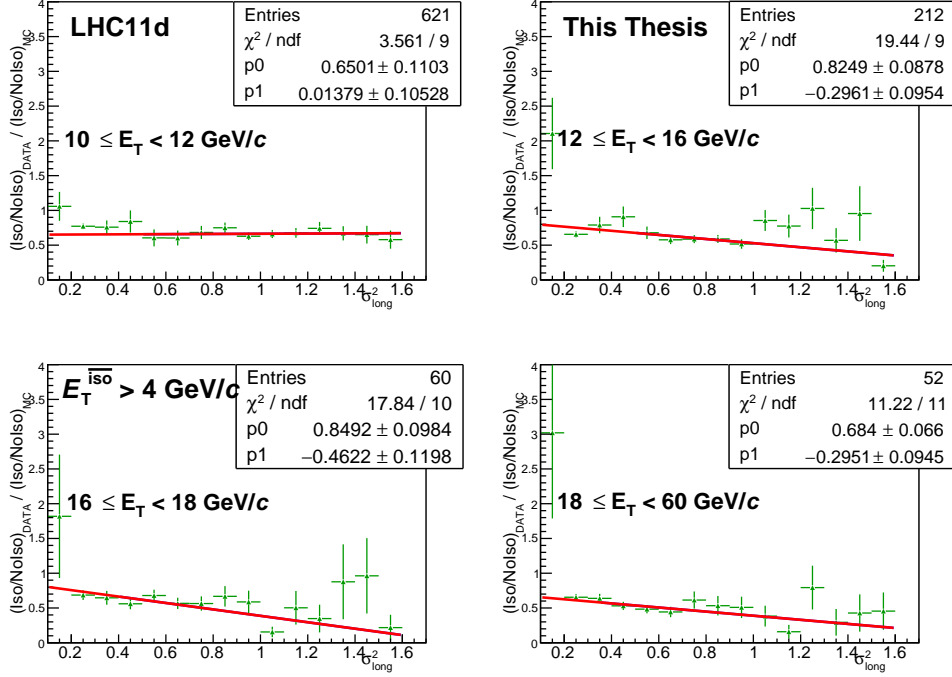


Figure B.9: Double ratio, as defined in Eq. 5.5, as a function of the σ_{long}^2 parameter for four different ranges of the reconstructed transverse energy of clusters. The results shown refer to the LHC11d sub period, with the STD smearing of the σ_{long}^2 parameter. The green lines represent the result of the fit to the points performed with a polynomial function of the first order, while the red lines show the estimation of the background in the narrow cluster region via extrapolation of the resulting fitting function.

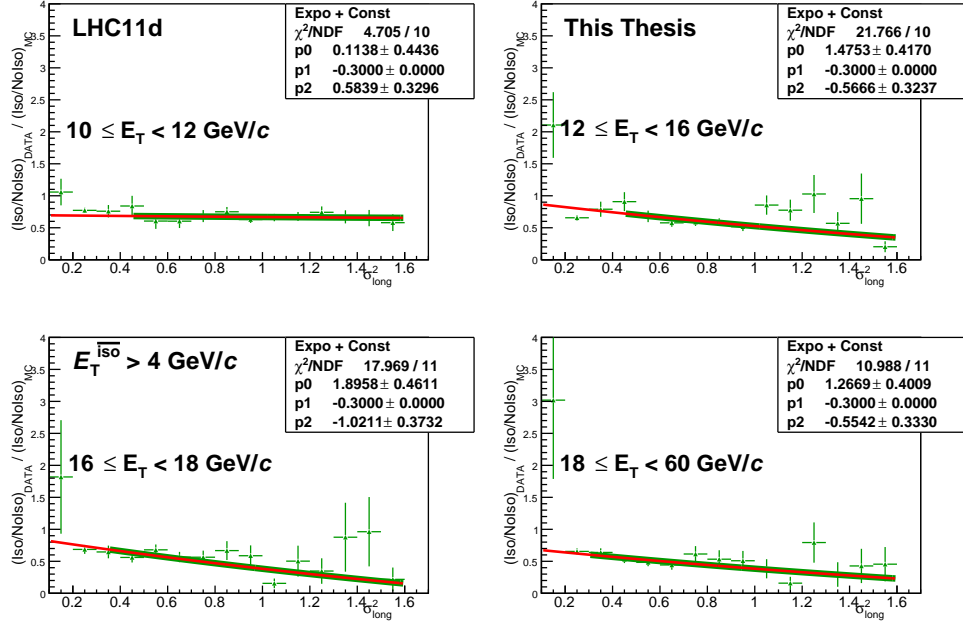


Figure B.10: Double ratio, as defined in Eq. 5.5, as a function of the σ_{long}^2 parameter for four different ranges of the reconstructed transverse energy of clusters. The results shown refer to the LHC11d sub period, with the STD smearing of the σ_{long}^2 parameter. The green lines represent the result of the fit to the points performed with an exponentially falling function, while the blue lines show the estimation of the background in the narrow cluster region via extrapolation of the resulting fitting function.

Support plots for Section 6.5.3

In Section 6.5.3 we have discussed extensively the differences between the in-cone activity from charged particles measured in data and in the simulated samples. We have also presented a small digression on the criterion adopted in this analysis to select isolated neutral clusters when considering the isolation activity due to charged particles only. In this section we will present the results, separately per period, of the studies that lead to the computation of the calibration factor α presented in Fig.6.30.

Results for the LHC11c period

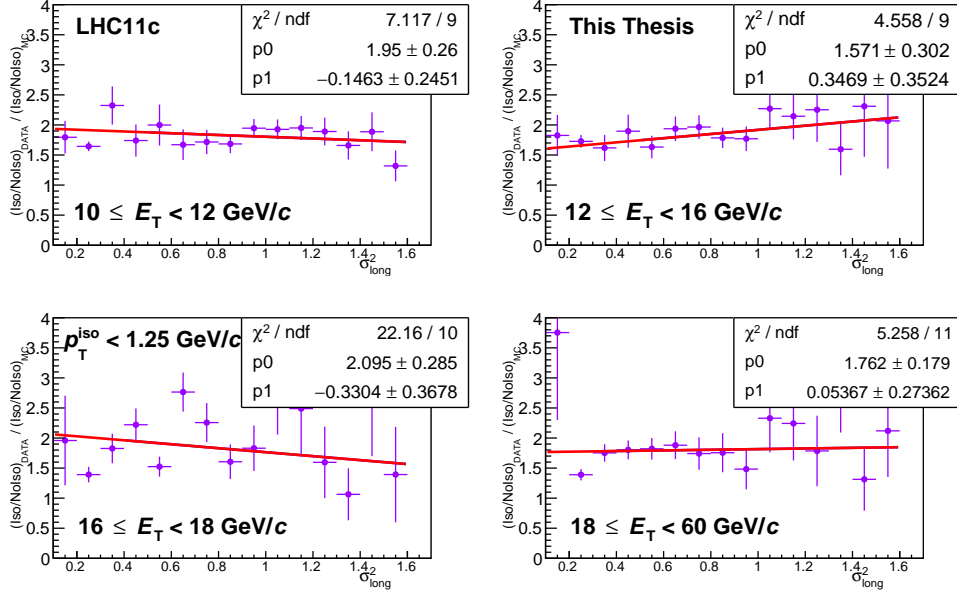


Figure B.11: Double ratio, as defined in Eq. 5.5, as a function of the σ_{long}^2 parameter for four different ranges of the reconstructed transverse energy of clusters. The results shown refer to the LHC11c sub period, with the STD smearing of the σ_{long}^2 parameter. The green lines represent the result of the fit to the points performed with a polynomial function of the first order, while the red lines show the estimation of the background in the narrow cluster region via extrapolation of the resulting fitting function.

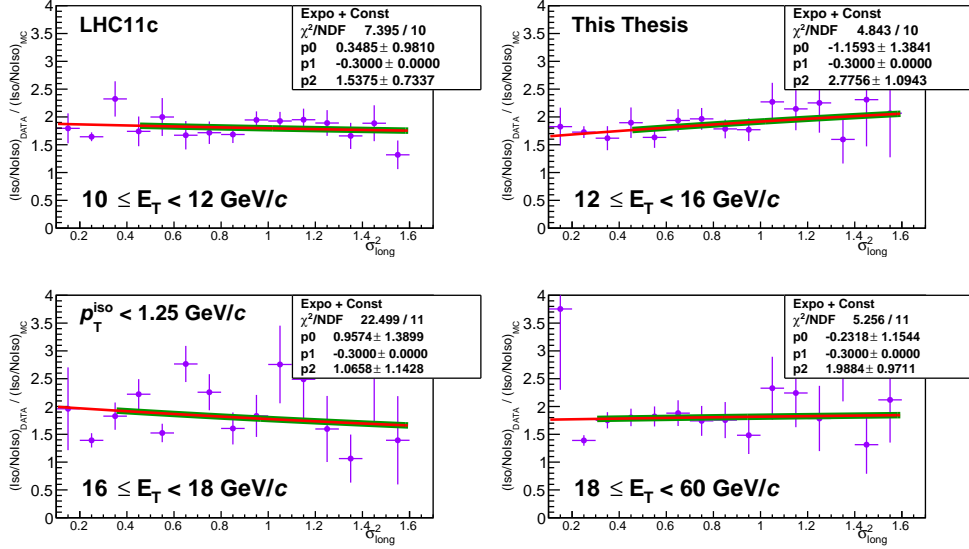


Figure B.12: Double ratio, as defined in Eq. 5.5, as a function of the σ_{long}^2 parameter for four different ranges of the reconstructed transverse energy of clusters. The results shown refer to the LHC11c sub period, with the STD smearing of the σ_{long}^2 parameter. The green lines represent the result of the fit to the points performed with an exponentially falling function, while the blue lines show the estimation of the background in the narrow cluster region via extrapolation of the resulting fitting function.

Results for the LHC11d period

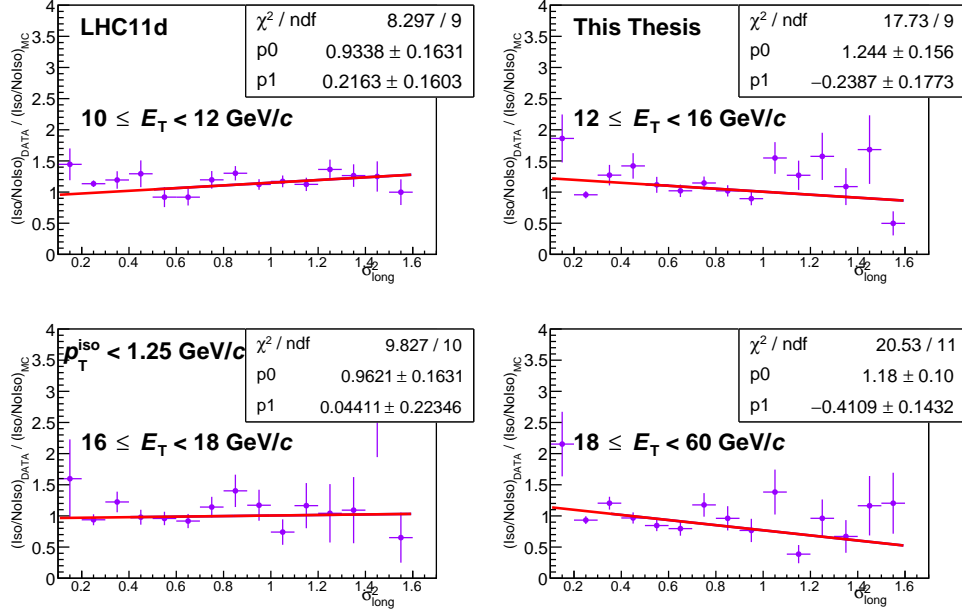


Figure B.13: Double ratio, as defined in Eq. 5.5, as a function of the σ_{long}^2 parameter for four different ranges of the reconstructed transverse energy of clusters. The results shown refer to the LHC11d sub period, with the STD smearing of the σ_{long}^2 parameter. The green lines represent the result of the fit to the points performed with a polynomial function of the first order, while the red lines show the estimation of the background in the narrow cluster region via extrapolation of the resulting fitting function.

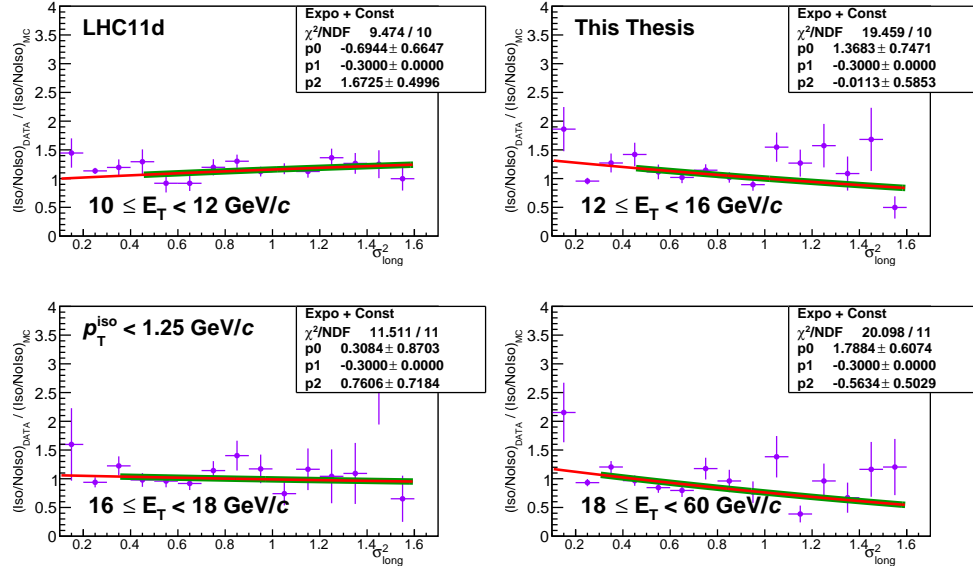


Figure B.14: Double ratio, as defined in Eq. 5.5, as a function of the σ_{long}^2 parameter for four different ranges of the reconstructed transverse energy of clusters. The results shown refer to the LHC11d sub period, with the STD smearing of the σ_{long}^2 parameter. The green lines represent the result of the fit to the points performed with an exponentially falling function, while the blue lines show the estimation of the background in the narrow cluster region via extrapolation of the resulting fitting function.

Bibliography

- [1] C. Patrignani et al. Review of Particle Physics. *Chin. Phys.*, C40(10):100001, 2016.
- [2] N. Cabibbo and G. Parisi. Exponential Hadronic Spectrum and Quark Liberation. *Phys. Lett.*, 59B:67–69, 1975.
- [3] Karen M. Burke et al. Extracting the jet transport coefficient from jet quenching in high-energy heavy-ion collisions. *Phys. Rev.*, C90(1):014909, 2014.
- [4] Rajeev S. Bhalerao. Relativistic heavy-ion collisions. In *Proceedings, 1st Asia-Europe-Pacific School of High-Energy Physics (AEPSHEP): Fukuoka, Japan, October 14-27, 2012*, pages 219–239, 2014.
- [5] A. Bialas, M. Bleszynski, and W. Czyz. Multiplicity Distributions in Nucleus-Nucleus Collisions at High-Energies. *Nucl. Phys.*, B111:461–476, 1976.
- [6] Jaroslav Adam et al. Centrality dependence of the nuclear modification factor of charged pions, kaons, and protons in Pb-Pb collisions at $\sqrt{s_{\text{NN}}} = 2.76$ TeV. *Phys. Rev.*, C93(3):034913, 2016.
- [7] Davide Francesco Lodato. Direct photon yield in pp and in Pb-Pb collisions measured with the ALICE experiment. In *8th International Conference on Hard and Electromagnetic Probes of High-energy Nuclear Collisions: Hard Probes 2016 (HP2016) Wuhan, Hubei, China, September 23-27, 2016*, 2017.

- [8] Betty Bezverkhny Abelev et al. Elliptic flow of identified hadrons in Pb-Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV. *JHEP*, 06:190, 2015.
- [9] Miklos Gyulassy and Xin-Nian Wang. HIJING 1.0: A Monte Carlo program for parton and particle production in high-energy hadronic and nuclear collisions. *Comput. Phys. Commun.*, 83:307, 1994.
- [10] Jaroslav Adam et al. Centrality dependence of particle production in p-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. *Phys. Rev.*, C91(6):064905, 2015.
- [11] Betty Abelev et al. Long-range angular correlations on the near and away side in p -Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. *Phys. Lett.*, B719:29–41, 2013.
- [12] Sören Schlichting and Prithwish Tribedy. Collectivity in Small Collision Systems: An Initial-State Perspective. *Adv. High Energy Phys.*, 2016:8460349, 2016.
- [13] G. Miller, E. D. Bloom, G. Buschhorn, D. H. Coward, H. DeStaebler, J. Drees, C. L. Jordan, L. W. Mo, R. E. Taylor, J. I. Friedman, G. C. Hartmann, H. W. Kendall, and R. Verdier. Inelastic electron-proton scattering at large momentum transfers and the inelastic structure functions of the proton. *Phys. Rev. D*, 5:528–544, Feb 1972.
- [14] E. M. Riordan. The Discovery of quarks. *Science*, 256:1287–1293, 1992.
- [15] F. D. Aaron et al. Combined Measurement and QCD Analysis of the Inclusive e^+p Scattering Cross Sections at HERA. *JHEP*, 01:109, 2010.
- [16] P. Amaudruz et al. Proton and deuteron f_2 structure functions in deep inelastic muon scattering. *Phys. Lett.*, B295:159–168, 1992.
- [17] Guido Altarelli and G. Parisi. Asymptotic Freedom in Parton Language. *Nucl. Phys.*, B126:298–318, 1977.
- [18] Yuri L. Dokshitzer. Calculation of the Structure Functions for Deep Inelastic Scattering and e^+e^- Annihilation by Perturbation Theory in Quantum Chromodynamics. *Sov. Phys. JETP*, 46:641–653, 1977. [*Zh. Eksp. Teor. Fiz.*73,1216(1977)].

- [19] L. N. Lipatov. Reggeization of the Vector Meson and the Vacuum Singularity in Nonabelian Gauge Theories. *Sov. J. Nucl. Phys.*, 23:338–345, 1976. [*Yad. Fiz.*23,642(1976)].
- [20] E. A. Kuraev, L. N. Lipatov, and Victor S. Fadin. The Pomeranchuk Singularity in Nonabelian Gauge Theories. *Sov. Phys. JETP*, 45:199–204, 1977. [*Zh. Eksp. Teor. Fiz.*72,377(1977)].
- [21] I. I. Balitsky and L. N. Lipatov. The Pomeranchuk Singularity in Quantum Chromodynamics. *Sov. J. Nucl. Phys.*, 28:822–829, 1978. [*Yad. Fiz.*28,1597(1978)].
- [22] Javier L. Albacete and Cyrille Marquet. Gluon saturation and initial conditions for relativistic heavy ion collisions. *Prog. Part. Nucl. Phys.*, 76:1–42, 2014.
- [23] Amir H. Rezaeian. CGC predictions for p+A collisions at the LHC and signature of QCD saturation. *Phys. Lett.*, B718:1058–1069, 2013.
- [24] Betty Abelev et al. Transverse momentum distribution and nuclear modification factor of charged particles in p -Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. *Phys. Rev. Lett.*, 110(8):082302, 2013.
- [25] S. Catani, M. Fontannaz, J. P. Guillet, and E. Pilon. Cross-section of isolated prompt photons in hadron hadron collisions. *JHEP*, 05:028, 2002.
- [26] Patrick Aurenche, Michel Fontannaz, Jean-Philippe Guillet, Eric Pilon, and Monique Werlen. A New critical study of photon production in hadronic collisions. *Phys. Rev.*, D73:094007, 2006.
- [27] Z. Belghobsi, M. Fontannaz, J. Ph. Guillet, G. Heinrich, E. Pilon, and M. Werlen. Photon - Jet Correlations and Constraints on Fragmentation Functions. *Phys. Rev.*, D79:114024, 2009.
- [28] Jamal Jalilian-Marian and Amir H. Rezaeian. Prompt photon production and photon-hadron correlations at RHIC and the LHC from the Color Glass Condensate. *Phys. Rev.*, D86:034016, 2012.

- [29] G. Dellacasa et al. ALICE technical design report of the inner tracking system (ITS). 1999.
- [30] J. Alme et al. The ALICE TPC, a large 3-dimensional tracking device with fast readout for ultra-high multiplicity events. *Nucl. Instrum. Meth.*, A622:316–367, 2010.
- [31] P Cortese et al. ALICE technical design report on forward detectors: FMD, T0 and V0. 2004.
- [32] P. Cortese et al. ALICE electromagnetic calorimeter technical design report. 2008.
- [33] J. Allen et al. Performance of prototypes for the ALICE electromagnetic calorimeter. *Nucl. Instrum. Meth.*, A615:6–13, 2010.
- [34] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. PYTHIA 6.4 Physics and Manual. *JHEP*, 05:026, 2006.
- [35] R. Brun, F. Bruyant, M. Maire, A. C. McPherson, and P. Zancarini. GEANT3. 1987.
- [36] Betty Bezverkhny Abelev et al. Performance of the ALICE Experiment at the CERN LHC. *Int. J. Mod. Phys.*, A29:1430044, 2014.
- [37] U. Abeysekara et al. ALICE EMCal Physics Performance Report. 2010.
- [38] M.Cosentino. An investigation on exotic clusters, 2012. ALICE EMCal weekly meeting.
- [39] T. C. Awes, F. E. Obenshain, F. Plasil, S. Saini, S. P. Sorensen, and G. R. Young. A Simple method of shower localization and identification in laterally segmented calorimeters. *Nucl. Instrum. Meth.*, A311:130–138, 1992.
- [40] G.Conesa Balbastre. Emcal photon clusters originating from conversions in simulation, 2015. ALICE PWG-GA General Meetings.
- [41] G.Conesa Balbastre. Clusterization and shower shape, update, 2011. ALICE PWG-GA General Meetings.

- [42] Betty Abelev et al. Underlying Event measurements in pp collisions at $\sqrt{s} = 0.9$ and 7 TeV with the ALICE experiment at the LHC. *JHEP*, 07:116, 2012.
- [43] M. Germain A. Mas et al. Isolated direct photon cross section measurement in pp collisions at $\sqrt{s} = 7$ TeV with the ALICE detector at LHC, 2016. ALICE Internal Analysis Note, ANA-1530.
- [44] A.Vauthier. Isolated photons purity estimation, 2016. ALICE Internal Analysis Note, ANA-3275.
- [45] Betty Bezverkhny Abelev et al. Measurement of quarkonium production at forward rapidity in pp collisions at $\sqrt{s} = 7$ TeV. *Eur. Phys. J.*, C74(8):2974, 2014.
- [46] Pavel M. Nadolsky, Hung-Liang Lai, Qing-Hong Cao, Joey Huston, Jon Pumplin, Daniel Stump, Wu-Ki Tung, and C. P. Yuan. Implications of CTEQ global analysis for collider observables. *Phys. Rev.*, D78:013004, 2008.
- [47] L. Bourhis, M. Fontannaz, and J. P. Guillet. Quarks and gluon fragmentation functions into photons. *Eur. Phys. J.*, C2:529–537, 1998.
- [48] G. Aad et al. Measurement of the inclusive isolated prompt photon cross section in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *Phys. Rev.*, D83:052005, 2011.
- [49] Serguei Chatrchyan et al. Measurement of the Differential Cross Section for Isolated Prompt Photon Production in pp Collisions at 7 TeV. *Phys. Rev.*, D84:052011, 2011.
- [50] J. Adams et al. Forward neutral pion production in $p+p$ and $d+Au$ collisions at $s(NN)^{1/2} = 200$ -GeV. *Phys. Rev. Lett.*, 97:152302, 2006.
- [51] I. Arsene et al. Quark gluon plasma and color glass condensate at RHIC? The Perspective from the BRAHMS experiment. *Nucl. Phys.*, A757:1–27, 2005.

- [52] Adrian Dumitru, Arata Hayashigaki, and Jamal Jalilian-Marian. The Color glass condensate and hadron production in the forward region. *Nucl. Phys.*, A765:464–482, 2006.
- [53] Inclusive low mass Drell-Yan production in the forward region at $\sqrt{s} = 7$ TeV. 2012.
- [54] A.P. de Haas et al. The FoCal prototype - an extremely fine-grained electromagnetic calorimeter using CMOS pixel sensors. *Submitted to JINST*, 2017.

Samenvatting

Hoog-energetische zware-ionenbotsingen stellen ons in staat een nieuwe toestand van de hadronische materie te bestuderen waarin de quarks en gluonen niet gebonden (confined) zijn in hadronen, maar een quark-gluonplasma (QGP) vormen, waarin ze vrij kunnen bewegen. De interacties in deze materietoestand met hoge temperatuur en dichtheid worden gekenmerkt door een lage impulsoverdracht, waardoor het onmogelijk is om dergelijke botsingen binnen het theoretische kader van de storingsrekening te beschrijven. Bij het onderzoek aan het QGP ligt daarom de nadruk op het vergelijken van verschijnselen in botsingen met atoomkernen met waarneming in botsingen van protonen, waarin geen QGP gevormd wordt.

Het startpunt van dergelijke vergelijkingen is het idee dat in eerste benadering de zware-ionenbotsingen gemodelleerd kunnen worden als een incoherente superpositie van botsingen van nucleonen, met de toevoeging van correcties die rekening houden met de aanwezigheid van een nucleair medium, zogenaamde nucleaire parton-dichtheidsverdelingen (nPDF's). We kunnen deze aanpak en de bijbehorende aannamen testen door metingen in botsingen van protonen met loodkernen te vergelijken met voorspellingen van storingsrekening, waarbij ook de kennis over de protonstructuur die is verkregen door een aantal *Deep Inelastic Scattering* (DIS) -experimenten, waaronder de baanbrekende metingen bij SLAC National Laboratory en de systematische metingen bij HERA, van belang is.

Omdat DIS echter wordt gemedieerd via de uitwisseling van een (virtueel) foton (of een W / Z -deeltje), zijn de metingen van de werkzame doorsnede (zogenaamde structuurfuncties) vooral gevoelig voor de PDF's (en hun evolutie met Q^2) van geladen partonen in het proton, terwijl de gluon PDF is verkregen uit de evolutie van de PDF's met Q^2 en van impuls-somregels.

De evolutie van de PDF's met de energieschaal wordt beschreven door de DGLAP-vergelijkingen. De DGLAP-aanpak resulteert in een exponentiële groei van de gluon-PDF in het regime van lage- x en lage Q^2 . Om een consistente beschrijving van hadronen te verkrijgen, worden daarom ook niet-lineaire effecten verwacht, waardoor de gluondichtheid *verzadigd*.

Een van de voorgestelde beschrijving van de gluonvelden in het verzadigingsregime is het *Color Glass Condensate* (CGC), een effectieve veldentheorie gebaseerd op de aanname dat bij lage x en lage Q^2 het gluonenveld in hadronen kan worden beschreven met een *mean field*-benadering, en dus behandeld kan worden als een klassiek veld.

Er zijn verschillende experimentele 'signaturen' voorspeld van de CGC-gedrag. Een van de theoretisch meest eenduidige metingen in die categorie is de productie van directe fotonen in de voorwaartste richting in hadronische botsingen. De zogenaamde prompte fotonen worden in hadronbotsingen, voornamelijk geproduceerd via quark-gluon-Compton-verstrooiing. In dit proces is het mogelijk om uit de energie en impuls van de geproduceerde deeltjes, de kinematica van het inkomende gluon en quark te bepalen.

Met de huidige detectoren, kan ALICE alleen van prompte fotonen meten die geproduceerd worden onder bijna loodrechte hoek met de bundel (mid-rapidity). Het onderwerp van dit proefschrift is de meting van zogenaamde foton met elektromagnetische calorimeter van ALICE, waarbij fotonen geselecteerd worden die geïsoleerd zijn, dwz. waarbij geen of weinig andere deeltjes in buurt geproduceerd worden.

De gepresenteerde resultaten hebben betrekking op de gegevens die in 2011 zijn verzameld, waarbij het EMC-L0-trigger algoritme werd gebruikt om de hoeveelheid niet-interessante opgeslagen gebeurtenissen te verminderen. Het onderzochte transversale energiegebied is $10 \leq p_{rmT}^\gamma < 60$ GeV/ c . Hiermee wordt het impulsbereik voor dergelijke metingen bij de Large Hadron Collider naar lagere waarden uitgebreid. De schatting van de achtergrond in de meting is uitgevoerd via de ABCD-methode, een *single side band*-methode die gelijktijdig wordt toegepast op twee variabelen die signaal en achtergrond kunnen onderscheiden.

Een onderzoek naar de systematische onzekerheden in onze meting wordt ook gepresenteerd. De gemeten productiedoorsnede voor fotonen is, bin-

nen de onzekerheden, in overeenstemming met de beschikbare NLO-pQCD-berekening die met de JETPHOX MC-generator is geproduceerd.

Onze meting dient als startpunt voor vergelijkbare studies in pPb en PbPb-botsingen en, door dezelfde meting uit te voeren voor fotonen in de voorwaartse richting, voor het bestuderen van de afhankelijkheid van de productiedoorsnede voor prompte fotonen van de impulsfractie x_2 van het gluon. Dit zal helpen bij het begrijpen en uiteindelijk bestuderen van niet-lineaire effecten zoals verzadiging, wat nodig is om een consistente beschrijving van hadronen te verkrijgen.

Acknowledgements

Although I like to think of myself as someone that hardly misses the possibility to thank anyone who help (or tries to, in his/her own way) me, I believe some acknowledgements on the completion of this thesis are in order.

First, I would like to thank my promoter Thomas, for giving me the opportunity both to fulfil a dream and for enriching me by sharing many times his vast experience.

Second, I will never be able to thank enough my supervisor Marco: the constant help, feedback, useful discussions, teachings and guidance (even when many people would not have bothered) have been (and still continue to be) invaluable to me.

Then, my parents (Enzo and Marinella) and my siblings (Ivano -the inspiration- and Andrea -the artist-): thanks for letting me be, helping me be here, and for loving me and allowing me to bother you, unconditionally.

Since almost two years, I feel blessed to have met Natália, my girlfriend, who has been supporting me during many problematic situations and who is shedding light on my life with her love. *Obrigado, tesouro meu. Desculpe se as vezes eu sou um pouco chato..*

I would also like to thank Stefano and Gaetano, Gigi, Dario, Salvo, Adriano, and Elisabetta: lifelong bros, who I missed every day of my life outside Italy; and then Andrea, Sandro, Jacopo, Laura DiG. and Maria (my -forever-housemates), Alberto, Redmer, Drazen, Val, G.G.N.A. (I know he prefers it this way), Barbara, Cristina, Henrique, Deepa, Antonio (x2), Morena, Claudia (x2), Sophie, Laura and Tania, Darius, Annelies, Martijn and Alexandrou. Thanks for making me feel home many and many times, regardless of where I was, to the point that I now have homesickness when I think of the Netherlands.

Curriculum Vitae

PhD Candidate (Subatomic Physics) 2012-2018

Utrecht University, the Netherlands

Thesis Title:

"Measurement of the prompt photon production
cross section in proton proton collisions
at $\sqrt{s} = 7$ TeV with ALICE"

M.Sc. Physics 2006-2011

University of Catania and INFN-LNS, Italy

Thesis Title:

"Neutral hadron reconstruction
with the LHCf experiment"

B.Sc. Physics 2003-2006

University of Catania and INFN, Italy

Thesis Title:

"Study of the response of SiC detectors
to α particles and light ions"

