Measurements of the inclusive isolated-photon and photon-plus-jet production in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector

Submitted by Daniel Camarero Muñoz in partial fulfilment of the requirements for the degree of Doctor in Physics

Supervisors: Claudia Glasman Kuguel and Juan Terrón Cuadrado

Univesidad Autónoma de Madrid

Faculty of Sciences

Department of Theoretical Physics



Universidad Autónoma de Madrid

Madrid, September 2021

CERN-THESIS-2021-240
15/09/2021



El sol brillaba con plenitud y fuerza. Mientras regresaba a casa, de pronto observé mi sombra ante mí, del mismo modo que veía la sombra de la otra guerra detrás de la actual. Durante todo ese tiempo, aquella sombra ya no se apartó de mí; se cernía sobre mis pensamientos noche y día; quizá su oscuro contorno se proyecta también sobre muchas páginas de este libro. Pero toda sombra es, al fin y al cabo, hija de la luz y sólo quien ha conocido la claridad y las tinieblas, la guerra y la paz, el ascenso y la caída, sólo éste ha vivido de verdad.

Stefan Zweig, *El mundo de ayer*

A mi familia y amigos

Agradecimientos

La finalización de este trabajo concluye una etapa de más de cuatro años que he vivido con gran intensidad. Un periodo de tiempo tan largo está necesariamente acompañado de innumerables lecciones y altibajos, que junto a las personas adecuadas se puede convertir en una experiencia de gran valor. Esto ha sido posible gracias a un cierto número de personas que han influido de manera más o menos directa en mí y en el desarrollo de este trabajo.

En el año 2015 conocí a Claudia, que ha sido mi tutora durante los periodos de máster y de doctorado. También conocí a Juan, mi codirector durante el periodo doctorado, aunque prefiero llamarlo tutor, ya que creo que esta palabra le hace más justicia. Durante aquel año y el posterior desarrollé mi trabajo de máster con Claudia, que tuvo una paciencia y dedicación fuera de lo normal, gracias a las cuales pude aprender mucho en muy poco tiempo. En aquel momento no pensaba en la posibilidad de trabajar con ellos en una etapa de doctorado, esto es algo que se produjo hasta cierto punto inesperadamente, y que no dudé ni un momento en perseguir. Durante este periodo de mi formación, de 2017 hasta el día de hoy, he aprendido gran cantidad de cosas, mi manera de pensar, capacidad de entendimiento y forma de trabajar se han modelado porque he tenido principalmente dos referencias claras sobre cómo desempeñar un trabajo de calidad sin perder de vista el punto de vista humano, que considero inseparable del profesional. Cuando he necesitado ayuda o consejo de cualquier tipo, lo he tenido a mi disposición, y siempre he estado acompañado en los momentos de mayor incertidumbre. En este último año y medio en el que todos hemos sufrido la pandemia y trabajar con la misma eficacia que antes ha sido más complicado, con ellos ha resultado fácil, hemos perseverado juntos, y para mí han sido un ejemplo de firmeza y compromiso. Por todos estos motivos admiro a Claudia y a Juan, y aunque el grado de éxito o fracaso en la consecución de un proyecto como este depende en mayor medida del desempeño de uno mismo, el mío no es ni será posible de entender ahora o en el futuro sin su enorme influencia, y es algo por lo que siempre les estaré agradecido.

Al comienzo de esta etapa, ante el aluvión inicial de nueva información, procedimientos y compromisos formales, mi compañera Ana me ayudó y me dio consejo con una predisposición total. Gracias a Ana pude progresar rápido al principio de mi trabajo en la cualificación de ATLAS, cuando parecía más abrumador, y también gracias a ella pude ser independiente para realizarlo al cabo de unos meses. Más allá de aquello, al trabajar

en el mismo grupo, antes de que Ana finalizase su doctorado, tuve tiempo suficiente para observar y tratar de aprender de su manera de trabajar y afrontar los desafíos. Creo que fue una estudiante extraordinaria y que en la misma medida lo sigue siendo ahora como investigadora independiente.

Durante estos últimos dos años también he podido trabajar con Josu, que para mí ha sido un ejemplo en todo, desde su dedicación y su capacidad para definir límites hasta su extraordinaria lucidez, que más de una vez me ha arrojado luz cuando la necesitaba. Siempre que le he pedido ayuda, la he recibido, y le estoy agradecido. Más o menos a la vez, aunque de manera menos estrecha, he trabajado con Javier, con quien a partir de mis estancias en Ginebra he coincidido más y de quien también he podido aprender tanto de física como de la complejidad y la realidad del mundo de la investigación, o de la pequeña parte de ese mundo con la que estamos relacionados. Tanto Javier como Josu son dos investigadores independientes que ya llevan cierto tiempo haciendo su camino, y si en el futuro siguiese sus pasos, espero convertirme en algo parecido a lo que ellos representan, con el mismo arrojo pero también su pureza y humildad. La compañía de Manuel también es algo que debo agradecer. Como estudiantes que prácticamente hemos desarrollado este periodo en paralelo y dado que ambos hemos realizado un tipo de trabajo similar, compartir nuestros problemas e inquietudes de cualquier tipo ha sido algo sencillo y tranquilizador. También he podido aprender mucho de él, de nuestras conversaciones de física, de historia, de filosofía y de lo humano. La pandemia nos ha arrebatado uno o dos años de convivencia, pero siempre guardaré con cariño en mi memoria los cinco meses que hemos compartido y los aproximadamente dos años que hemos coincidido en el CERN. A Juan Carlos, cuya amistad ha sido una suerte, le agradezco su ayuda y lo bien que se ha portado conmigo. Siempre ha sido un compañero de tenis ejemplar, salvo en la digestión de la derrota. Junto con Javier, Josu y Manuel, hemos pasado muy buenos momentos juntos, algunos de los mejores que guardo de estos últimos años. En Heidelberg, en 2018, conocí a mi amigo Javier, a quién también quiero agradecer su compañía y conversación, que a menudo ha sido una suerte de bálsamo. Desde entonces hemos mantenido la amistad y siempre es una alegría poder coincidir él y con Davide, al que también guardo aprecio. Durante estos años en el CERN, además de las personas que ya he mencionado, he conocido a otras muchas que me han acompañado día a día, como a Mercedes, Guillermo, Pilar, Carlos, Miguel, Emma, Santiago y Neus, a las que también quiero que agradecer su compañía.

Del principio de mi periodo de doctorado en la UAM, quiero darle las gracias a Isabel, mi compañera de despacho preferida. Le estoy agradecido por su cercanía, sensibilidad y su calidad como persona, que me hicieron mucho más llevadero el inicio de este camino. También la compañía de Raúl, Isa, Marina, Doris y Ania fueron importantes, ya que entre todos llegamos a formar un grupo de amigos, y por ello les estoy muy agradecido. Todos ellos han tenido éxito en su camino y les deseo lo mejor para el futuro.

Quiero agradecerle a mi familia, su comprensión, cariño y apoyo incondicional; ellos me han permitido desarrollarme con total libertad y nunca me ha faltado nada cuando lo he necesitado. A Andrea quiero agradecerle su compañía y apoyo durante este periodo de

tiempo. A mis amigos de toda la vida, desde el colegio a la universidad y de otros lugares, que por fortuna son muchos y representan una fuente de alegría, sosiego y aprendizaje inagotables, gracias por vuestro cariño. A Rossana, gracias por acompañarme en el camino.

Por último, el haber llegado hasta aquí también se lo debo a todos los profesores que me han influido en Valladolid, desde la escuela a la universidad, algunos de los cuales han sido extraordinarios. También, de mi periodo como estudiante de máster en Madrid quiero agradecerles su trabajo a los profesores que me han influido. En mi etapa de doctorado en el CERN, he trabajado con varios estudiantes, investigadores y profesores de los que he podido aprender mucho y quiero agradecérselo. En cada una de las instituciones que he mencionado anteriormente, hay un gran número de personas trabajando en su administración y en el funcionamiento de sus servicios, a las que quiero agradecerles el haber recibido siempre un trato y atención excelentes.

Abstract

This thesis presents various measurements of the differential cross section for isolated-photon production inclusively and in association with one jet in proton-proton collisions at a centre-of-mass energy of 13 TeV with the ATLAS detector at the LHC. These measurements are sensitive to the parton distribution functions in the proton. In addition, a study of the *in situ* calibration of the jet energy scale and resolution of jets using photon-plus-jet events in proton-proton collisions at a centre-of-mass energy of 13 TeV is also included.

The inclusive isolated-photon differential cross section was measured using a sample of events corresponding to an integrated luminosity of $36.1\,\text{fb}^{-1}$ collected during the 2015 and 2016 data-taking periods. The measurement was performed in four different regions of the pseudorapidity of the photon as a function of the photon transverse energy. A later measurement was performed using the full integrated luminosity of the ATLAS Run II data of $139\,\text{fb}^{-1}$, increasing the regions of pseudorapidity of the photon from four to six and measuring the cross section for two different photon isolation cone sizes. In both measurements, theoretical predictions are found to provide a good description of the data within the experimental and theoretical uncertainties. The production of isolated-prompt photons associated with one jet was studied using $138.4\,\text{fb}^{-1}$ of the ATLAS Run II dataset. The differential cross section was measured as a function of the transverse energy of the photon for different angular configurations of the photon and the jet. The next-to-leading order perturbative QCD calculations describe adequately the data.

The jet energy scale and resolution were measured using in situ techniques in photon-plus-jet events in which the photon and the leading jet transverse momenta are balanced. The jet energy scale (resolution) measurement presented in this dissertation was performed using a sample of events corresponding to an integrated luminosity of $80\,\text{fb}^{-1}$ ($58.5\,\text{fb}^{-1}$) collected during the $2015-2017$ (2018) data-taking periods. The in situ γ+jet corrections to the jet energy scale are at the level of 2% in most of the measured range of the photon transverse momentum, and the uncertainty on the jet energy scale is of the order of 1% for the different measured configurations. Resolutions at the order of $17\%-24\%$ ($4\%-5\%$) with experimental uncertainties of $\approx 25\%-45\%$ ($\approx 35\%$) are obtained at photon transverse momenta of $25-45$ GeV (≈ 1.2 TeV) for the two measured regions of jet rapidity and jet collections.

Resumen

Esta tesis presenta varias mediciones de la sección eficaz diferencial de la producción de fotones aislados inclusivos y en asociación con un jet en colisiones protón-protón a una energía del centro de masas de 13 TeV con el detector ATLAS en el LHC. Estas mediciones son sensibles a las funciones de distribución de los partones en el protón. Además, también se incluye un estudio de la calibración *in situ* de la escala de la energía y la resolución de los jets usando sucesos con un fotón más un jet en colisiones protón-protón a una energía del centro de masas de 13 TeV.

La sección eficaz diferencial de fotones aislados inclusivos se midió usando una muestra de sucesos correspondientes a una luminosidad integrada de $36.1\,\text{fb}^{-1}$ recogida durante los periodos de toma de datos de 2015 y 2016. La medición se realizó en cuatro regiones diferentes de la pseudorapidez del fotón en función de su energía transversa. Posteriormente, se realizó una medición empleando la luminosidad integrada total de $139\,\text{fb}^{-1}$ de los datos del Run II de ATLAS, incrementando las regiones de pseudorapidez del fotón de cuatro a seis y midiendo la sección eficaz para dos tamaños del cono de aislamiento del fotón. Para ambas mediciones, las predicciones teóricas proporcionan una buena descripción de los datos dentro de las incertidumbres experimentales y teóricas. La producción de fotones *prompt* aislados asociados con un jet fue estudiada usando $138.4\,\text{fb}^{-1}$ del conjunto de datos del Run II de ATLAS. La sección eficaz diferencial fue medida en función de la energía transversa del fotón para diferentes configuraciones angulares del fotón y del jet. Los cálculos perturbativos de QCD al orden siguiente al principal describen los datos adecuadamente.

La escala de la energía y la resolución de los jets fueron medidas usando técnicas in situ en sucesos fotón más jet en los que los momentos transversos del fotón y del jet principal están balanceados. La medición de la escala (resolución) de la energía presentada en esta disertación se realizó usando una muestra de sucesos correspondiente a una luminosidad integrada de $80\,\text{fb}^{-1}$ ($58.5\,\text{fb}^{-1}$) recogida durante los periodos de toma de datos de $2015-2017$ (2018). Las correcciones in situ γ+jet se encuentran al nivel de 2% en la mayor parte del rango medido del momento transverso del fotón, y la incertidumbre en la escala de la energía es del orden de 1% para las diferentes configuraciones medidas. Se obtienen resoluciones del orden de $17\% - 24\%$ ($4\% - 5\%$) con incertidumbres experimentales de $\approx 25\% - 45\%$ ($\approx 35\%$) a momentos transversales del fotón de $25 - 45$ GeV (≈ 1.2 TeV) para las dos regiones de rapidez y colecciones de jets medidas.

Contents

1 Introduction **18**

2 Experimental set-up **28**

2.1 The LHC accelerator 28

2.1.1 Main machine layout and performance 29

2.1.2 Timeline of the LHC and ATLAS data-taking 32

2.2 The ATLAS detector 36

2.2.1 Physics requirements and detector overview 37

2.2.2 Tracking 39

2.2.3 Calorimetry 42

2.2.4 The muon spectrometer 47

2.2.5 Forward detectors 50

2.2.6 Trigger and data-acquisition system 50

3 Theoretical framework **54**

3.1 The Standard Model 54

3.1.1 Mathematical formulation of the Standard Model 56

3.2 Quantum Chromodynamics 58

3.2.1 The Renormalisation Group Equations 59

3.2.2 The running coupling constant 60

3.2.3 The β function 61

3.2.4 Asymptotic freedom and confinement 62

3.2.5 Parton model, structure and parton distribution functions 63

3.2.6 The QCD improved parton model 69

3.3 Jet algorithms 71

3.3.1 Properties and comparisons between algorithms . . . 73
3.4 Theory of prompt-photon production . . . 76
3.4.1 Photon isolation . . . 79
3.4.2 Fixed-order pQCD calculations . . . 81
3.4.3 Recent progress on Standard Model photon physics . . . 82

4 Monte Carlo event generators **86**
4.1 Monte Carlo simulations . . . 86
4.1.1 The complex skeleton of high-energy processes . . . 87
4.1.2 Event generators . . . 88
4.2 Physics overview . . . 89
4.2.1 Hard processes and parton distributions . . . 90
4.2.2 Initial- and final-state radiation . . . 91
4.2.3 Combining matrix elements and parton showers . . . 93
4.2.4 Multiple parton interactions . . . 94
4.2.5 Hadronisation . . . 95
4.2.6 Simulation of detector effects . . . 100
4.3 Monte Carlo samples . . . 103

5 Photon reconstruction, identification and calibration **108**
5.1 Trigger reconstruction and identification of electrons and photons . . . 108
5.1.1 Photon and electron triggers at L1 . . . 109
5.1.2 HLT photon reconstruction and identification for pp data-taking . . . 109
5.2 Photon reconstruction . . . 111
5.2.1 Topocluster reconstruction . . . 112
5.2.2 Track reconstruction, track-cluster matching and photon conversion reconstruction . . . 113
5.2.3 Supercluster reconstruction . . . 113
5.2.4 Photons for analysis . . . 114
5.2.5 Photon reconstruction performance . . . 115
5.3 Photon calibration . . . 115
5.3.1 Electron and photon energy estimate and expected resolution from the simulation . . . 116

5.3.2 Corrections applied to the data . . . 117
5.3.3 Data/MC energy scale and resolution measurements with $Z \to ee$ decays . . . 120
5.3.4 Systematic uncertainties in the energy scale and resolution . . . 121
5.3.5 Cross-checks with $J/\Psi \to ee$ and $Z \to ll\gamma$ . . . 122
5.4 Photon identification . . . 123
5.4.1 Optimisation of the photon identification . . . 125
5.5 Photon isolation . . . 126
5.6 Additional corrections to simulated events . . . 128

6 Jet reconstruction and calibration **130**
6.1 Jet reconstruction . . . 130
6.2 Jet energy scale calibration . . . 131
6.2.1 Simulation-based jet calibrations . . . 132
6.2.2 In situ jet calibrations . . . 136
6.2.3 Systematic uncertainties . . . 145
6.3 Jet energy resolution . . . 146
6.3.1 Resolution measurement using dijet events . . . 147
6.3.2 Noise measurement using random cones . . . 148
6.3.3 Combination of *in situ* JER measurements . . . 152

7 Measurement of the inclusive isolated-photon cross section at $\sqrt{s} = 13$ TeV using $36\,\mathrm{fb}^{-1}$ of ATLAS data **153**
7.1 Data selection . . . 153
7.2 Monte Carlo simulations . . . 162
7.3 Background evaluation and signal extraction . . . 163
7.3.1 Background from electrons faking photons . . . 168
7.4 Control plots . . . 170
7.5 Selection efficiency and purity . . . 172
7.6 Cross-section measurement procedure . . . 177
7.7 Systematic uncertainties . . . 181
7.7.1 Signal modelling . . . 181
7.7.2 Background subtraction . . . 183

7.7.3 Photon reconstruction . . . 192
7.7.4 Unfolding procedure . . . 195
7.7.5 Running conditions . . . 198
7.7.6 Photon energy scale and resolution . . . 200
7.7.7 Total systematic uncertainty . . . 204
7.8 Theory predictions . . . 207
7.8.1 Hadronisation and UE corrections to the NLO QCD calculations . . . 207
7.8.2 Theoretical uncertainties in the NLO pQCD calculations . . . 209
7.8.3 Theoretical uncertainties in the NNLO pQCD calculations . . . 210
7.9 Results . . . 215

8 Measurement of the cross section for inclusive isolated photon production in *pp* collisions at $\sqrt{s} = 13$ TeV using $139\,\text{fb}^{-1}$ of ATLAS data 225
8.1 Data selection . . . 225
8.2 Monte Carlo simulations . . . 231
8.3 Background evaluation and signal extraction . . . 235
8.3.1 Background from electrons faking photons . . . 240
8.4 Control plots . . . 241
8.5 Selection efficiency and purity . . . 244
8.6 Cross-section measurement procedure . . . 247
8.7 Systematic uncertainties . . . 251
8.7.1 Signal modelling . . . 251
8.7.2 Background subtraction . . . 253
8.7.3 Photon reconstruction . . . 260
8.7.4 Unfolding procedure . . . 264
8.7.5 Running conditions . . . 267
8.7.6 Photon energy scale and resolution . . . 269
8.7.7 Total systematic uncertainty . . . 273
8.8 Theory predictions . . . 276
8.8.1 Hadronisation and UE corrections to the NLO QCD calculations . . . 276
8.8.2 Theoretical uncertainties in the NLO pQCD calculations . . . 276
8.9 Results . . . 283

9 In situ jet energy scale and resolution calibration using γ+jet events 297
9.1 In situ JES calibration using $\boldsymbol{\gamma}$+jet events . . . 297
9.1.1 The MPF and direct balance methods . . . 297
9.1.2 Data selection . . . 299
9.1.3 Monte Carlo simulations . . . 300
9.1.4 Fitting procedure . . . 301
9.1.5 Signal purity estimation . . . 302
9.1.6 Response distributions . . . 304
9.1.7 Systematic uncertainties . . . 308
9.1.8 Closure of the in situ calibration . . . 311
9.1.9 Results . . . 313
9.2 In situ JER calibration using $\boldsymbol{\gamma}$+jet events . . . 313
9.2.1 Strategy . . . 313
9.2.2 Data selection . . . 314
9.2.3 Monte Carlo simulations . . . 314
9.2.4 The direct balance method applied to the JER . . . 314
9.2.5 Measurement of the JER with $\mathbf{58.5\,fb^{-1}}$ of ATLAS data . . . 315

10 Measurement of the isolated photon plus jet cross section in pp collisions at $\sqrt{s}$ = 13 TeV using $138\,\mathrm{fb}^{-1}$ with the ATLAS detector 324
10.1 Data selection . . . 324
10.2 Monte Carlo simulations . . . 333
10.3 Background evaluation and signal extraction . . . 333
10.3.1 Background from electrons faking photons . . . 337
10.4 Control plots . . . 339
10.5 Selection efficiency and purity . . . 341
10.6 Cross-section measurement procedure . . . 344
10.7 Systematic uncertainties . . . 348
10.7.1 Signal modelling . . . 348
10.7.2 Background subtraction . . . 350
10.7.3 Photon reconstruction . . . 357
10.7.4 Unfolding procedure . . . 361

10.7.5 Running conditions . . . 364
10.7.6 Photon energy scale and resolution . . . 366
10.7.7 Jet energy scale and resolution . . . 370
10.7.8 Total systematic uncertainty . . . 374
10.8 Theoretical predictions . . . 377
10.8.1 Theoretical uncertainties in the NLO pQCD calculations . . . 377
10.9 Results . . . 380

11 Summary and conclusions **385**

Bibliography **406**

A Converted photon fractions **407**

B Ambient transverse energy density **409**

C Data-driven corrections to $E_{\mathrm{T}}^{\mathrm{iso}}$ **411**

D Cross-checks using reweighted MC for the inclusive photon analysis using $36\,\mathrm{fb}^{-1}$ **424**

E Cross-check using Bayesian unfolding for the inclusive photon analysis using $36\,\mathrm{fb}^{-1}$ **430**

F Trigger efficiencies for the inclusive photon analysis using $139\,\mathrm{fb}^{-1}$ **436**

G Cross-checks using reweighted MC for the inclusive photon analysis using $139\,\mathrm{fb}^{-1}$ **440**

H Fits to the MPF and DB distributions **446**

I Trigger efficiencies for the photon plus jet analysis using $138\,\mathrm{fb}^{-1}$ **451**

J Fraction of brem component in PYTHIA **455**

K Cross-checks using reweighted MC for the photon plus jet analysis using $138\,\mathrm{fb}^{-1}$ **458**

Chapter 1

Introduction

The Large Hadron Collider (LHC) [1], built by the European Organization for Nuclear Research (CERN), has extended the frontiers of particle physics over the years with its unprecedented high energy and luminosity. With this rich physics potential, the LHC has provided more precise measurements of the Standard Model (SM) parameters and sensitive searches of new phenomena. The discovery of the Higgs boson in 2012 [2, 3] by ATLAS (A Toroidal LHC Apparatus) [4] and CMS (Compact Muon Solenoid) [5], represents one of the most important milestones in the history of particle physics, and one of the reasons why the LHC was built. Moreover, the high luminosity and large cross-sections at the LHC allow further high precision tests of Quantum Chromodynamics (QCD), electroweak interactions, and flavour physics, and at the same time, the investigation of numerous models of new physics.

Prompt photons with large transverse momenta constitute colourless probes of the hard interaction and their production in proton-proton (pp) collisions, $pp \to \gamma + X$, provides a testing ground for perturbative QCD (pQCD) at large hard-scattering scales (Q^2) and over a wide range of the incoming parton momentum fraction (x). Prompt photons are defined as those that are not secondaries from hadron decays, and their production is understood to proceed via two processes: the one in which the photon arises directly from the hard interaction (the direct-photon process), and the one in which the photon is emitted in the fragmentation of a high transverse momentum (p_{T}) parton (the fragmentation-photon process) [6, 7].

In pp collisions, due to the abundance of photons from neutral hadron decays and the contribution from the fragmentation process, prompt-photon production is studied by requiring the photons to be isolated. Measurements of inclusive isolated-photon production in pp collisions at centre-of-mass energies ($\sqrt{s}$) of 7, 8 and 13 TeV have been published by the ATLAS [8–10] and CMS [11, 12] collaborations.

At the LHC, the dominant contribution to prompt-photon production arises from the $qg \to q\gamma$ process. As a consequence, the production is sensitive to the gluon density in the proton [13–15] even at the leading order (LO) in pQCD. The measurement of the

production cross section of inclusive isolated-photons in different ranges of the photon transverse energy (E_{T}^{γ}) and pseudorapidity (η^{γ}) provides detailed information on the parton distribution functions (PDF) in different regions of Q^2 and x, which at LO is equal to $2\,\frac{E_{\mathrm{T}}^{\gamma}}{\sqrt{s}}\,e^{\pm\eta^{\gamma}}$. A recent study [16] has included the ATLAS measurement of inclusive-photon production at 8 TeV [9] in a next-to-next-to-leading-order (NNLO) QCD fit to obtain the PDF of the proton; figure 1.1 shows the kinematic coverage of the ATLAS 8 TeV data in the (x, Q^2) plane. In ref. [16] it is concluded that the inclusion of the ATLAS data to the NNPDF3.1 NNLO global analysis does not create any tension with the other datasets included in the fit; moreover, the constraints on the gluon PDF from the inclusion of the ATLAS data are consistent with those of the rest of the datasets in the NNPDF3.1 analysis. Figure 1.2 shows the moderate reduction of the gluon PDF uncertainty in the region $10^{-3} \lesssim x \lesssim 0.4$, which is consistent with the kinematic coverage of the ATLAS 8 TeV data shown in figure 1.1.

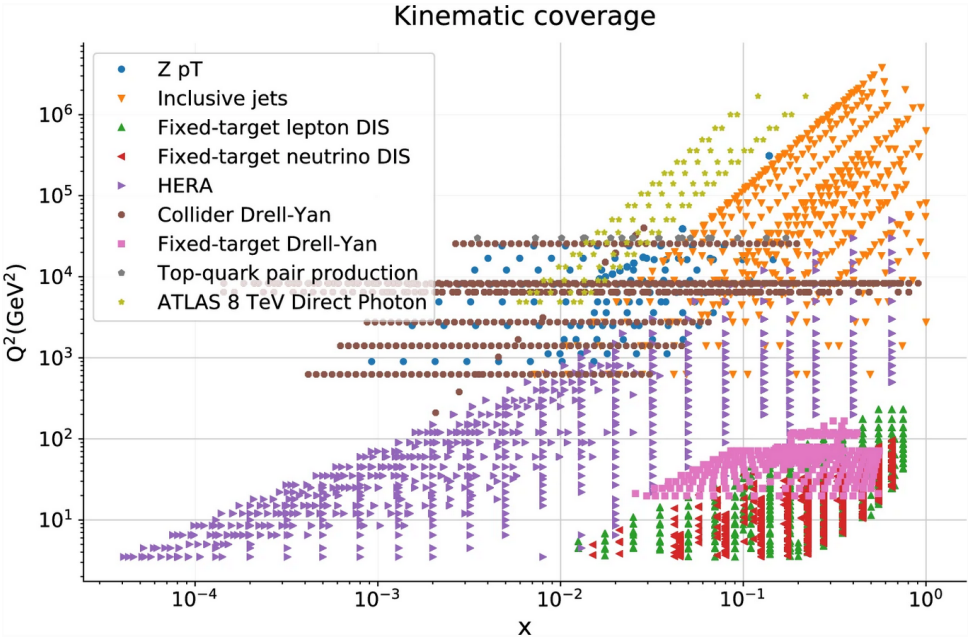


Figure 1.1: The coverage in the (x, Q^2) kinematic plane of the ATLAS 8 TeV photon measurement [9] (using LO kinematics), compared to the dataset used in the global NNPDF3.1 fit. From ref. [16].

Measurements of inclusive prompt-photon production also provide benchmarks to use in investigating novel approaches to parton radiation [17], next-to-leading-order (NLO) QCD corrections with matched parton showers [18], the relevance of threshold logarithms in QCD and electroweak corrections [19] and NNLO QCD corrections [20–22]. More specifically, the production of prompt photons, which is less sensitive to hadronisation effects than that of jets, can be used to pursue an alternative QCD description based on the k_{t} factorisation approach combining off-shell amplitudes and transverse-momentum-dependent parton densities [17]. The prompt-photon data allow the investigation of

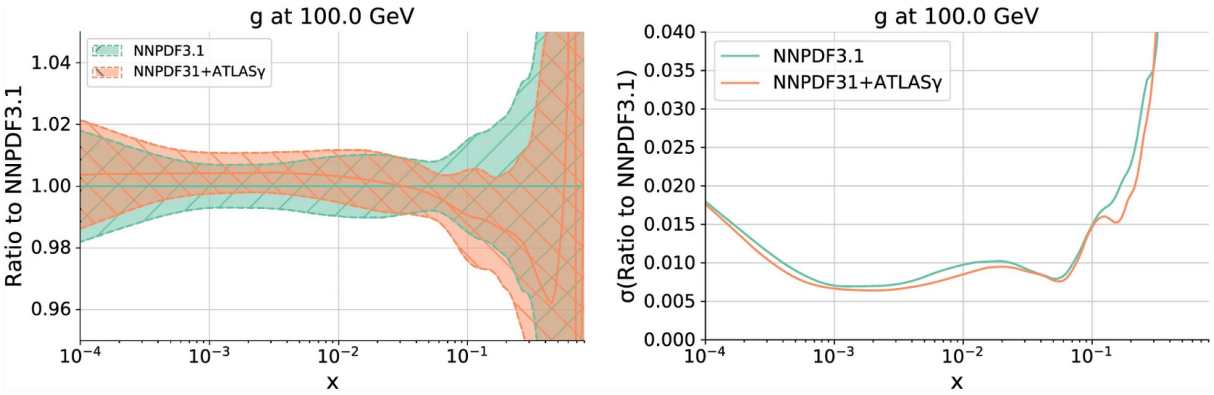


Figure 1.2: The comparison of the gluon PDF at $Q = 100$ GeV between the NNPDF3.1 and NNPDF3.1 + ATLASγ fits, normalised to the central value of the former is shown on the left. The corresponding relative one-sigma PDF uncertainties in each case are shown on the right. From ref. [16].

the fragmentation contribution, which can be done either via fragmentation functions or through the inclusion of parton showers [18]. Electroweak corrections are likely to play an important role at the TeV scale, thereby, measurements at photon transverse energies in that region would help to unveil such phenomena [19]. The prompt-photon measurements at the LHC are characterised by small uncertainties that demand precise theoretical predictions, e.g. those from NNLO pQCD calculations [20–22], to fully exploit these data. Measurements involving isolated photons have also been used to constrain the contributions from new light scalar particles decaying into a photon pair [23].

Jets are a prevalent feature of the final state in high-energy pp interactions at the LHC. They are made of collimated showers of hadrons and represent important elements in many SM measurements and in searches for new phenomena. Jets are reconstructed using a clustering algorithm run on a set of input four-vectors, typically obtained from topologically associated energy deposits, charged-particle tracks, or simulated particles. Measurements of the energy scale and resolution of jets, as well as their associated systematic uncertainties, are therefore essential both for precision measurements and for sensitive searches for new physics. The calibrations account for features of the detector, the jet reconstruction algorithm, jet fragmentation, and the busy data-taking environment resulting from multiple pp interactions, referred to as pile-up. Finally, in situ techniques are used to measure the difference in jet response and resolution between data and Monte Carlo (MC), and residual corrections are applied to jets in data only.

The production of prompt photon plus jet events, $pp \to \gamma + \text{jet} + X$, also represents an important test of pQCD predictions. In addition, the study of the angular correlations between the photon and the jet can be used to constrain the photon fragmentation functions [7]. Photon-plus-jet events can also be used to calibrate the jet energy scale and its resolution by profiting from momentum conservation in the transverse plane and the accurate energy measurement of the photon in the electromagnetic (EM) calori-

meter [24–26]. Moreover, precise knowledge of the photon-plus-jet production processes is useful to validate the MC generators used for background studies in searches for physics beyond the SM which involve photons, such as the search for new phenomena in final states with a photon and a jet [27, 28]. Measurements of isolated-photon plus jet production in pp collisions at $\sqrt{s} = 7, 8$ and 13 TeV have been published by the ATLAS [29–33] and CMS [12, 34–36] collaborations.

The measurement of the differential cross section, $\mathrm{d}\sigma/\mathrm{d}E_{\mathrm{T}}^{\gamma}$, in three jet rapidity intervals and for two different angular configurations of the photon-plus-jet system, namely the same-sign ($\eta^{\gamma} \cdot y^{\mathrm{jet}} \geq 0$) and the opposite-sign ($\eta^{\gamma} \cdot y^{\mathrm{jet}} < 0$) configurations, gives access to the PDF in different regions of Q^2 and x, which in the leading-order approximation is given by $\frac{E_{\mathrm{T}}^{\gamma}}{\sqrt{s}}\left(e^{\pm\eta^{\gamma}} + e^{\pm y^{\mathrm{jet}}}\right)$. Therefore, the measurement of the differential cross section in these configurations provides the possibility to identify regions of phase space which are specially sensitive to the proton PDF. Furthermore, the separation of the phase space of the measurement into these different angular configurations allows the comparison between the data and theoretical predictions in regions where the relative contribution of the fragmentation component to the total cross section is different.

The structure of this thesis is organised as follows. Chapter 2 describes the LHC accelerator with an emphasis on the ATLAS detector. Chapter 3 describes the fundamentals of the theoretical framework underlying the work presented in this thesis; a description of jet algorithms and an overview of the theory of prompt-photon production are also presented. Chapter 4 describes the basic characteristics of MC generators and the samples used for this thesis specifically. Chapter 5 describes the reconstruction, calibration and identification of photons in ATLAS. Chapter 6 describes the reconstruction of jets and their calibration in ATLAS. The next chapters of this dissertation present various measurements of the production of isolated prompt photons inclusively and in association with one jet in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector at the LHC. The first of these studies is the measurement of the inclusive isolated-photon cross section using $36.1\,\mathrm{fb}^{-1}$ of ATLAS data, which was published in ref. [37], and is presented in chapter 7. An extension of the previous measurement using the full ATLAS Run II integrated luminosity was performed subsequently, increasing the number of measured regions of the photon pseudorapidity from four to six, and measuring the differential cross sections in E_{T}^{γ} for two different photon isolation cone sizes. This measurement is currently under review in the ATLAS collaboration and is presented in chapter 8. A study of the calibration of the jet energy scale (JES) and resolution (JER) using in situ techniques in final states in which a jet recoils against a well-measured photon is also presented in chapter 9. The JES (JER) calibration presented in this thesis was performed using $80\,\mathrm{fb}^{-1}$ ($58.5\,\mathrm{fb}^{-1}$) of the ATLAS $\sqrt{s} = 13$ TeV data. The measurement of the JES, recently published in ref. [26], represents an improvement in terms of systematic and statistical uncertainties thanks to the increase of integrated luminosity and the improvements from the collaboration recommendations with respect to the previous calibration at $\sqrt{s} = 13$ TeV [25]. Finally, in chapter 10, the measurement of the cross section of isolated-photon plus jet production using $138.4\,\mathrm{fb}^{-1}$ is documented in detail. This measurement is ongoing. The

first of these measurements, included in chapter 7, represented and improvement in terms of statistical and systematic uncertainties, extending the reach in E_{T}^{γ} with respect to the previous measurement at $\sqrt{s} = 13$ TeV. The two latter measurements presented in chapters 8 and 10, respectively, benefit from the inclusion of the full Run II integrated luminosity, providing higher values in E_{T}^{γ} with smaller statistical uncertainties and also benefits from reduced systematic uncertainties thanks to the continuous improvement in terms of precision of the ATLAS combined performance groups which provide recommendations to evaluate key systematic uncertainties for these analyses. Therefore, these measurements will be part of the legacy of the ATLAS collaboration in its exploration of the 13 TeV energy frontier.

Introducción en castellano

El Gran Colisionador de Hadrones (LHC) [1], construido por la Organización Europea para la Investigación Nuclear (CERN), ha extendido las fronteras de la física de partículas a lo largo de los años con su alta energía y luminosidad sin precedentes. Con este rico potencial físico, el LHC ha proporcionado mediciones más precisas de los parámetros del Modelo Estándar (SM) y búsquedas sensibles de nuevos fenómenos. El descubrimiento del bosón de Higgs en 2012 [2,3] por ATLAS [4] y CMS [5], representa uno de los hitos más importantes en la historia de la física de partículas, y una de las razones por las cuales se construyó el LHC. Además, la alta luminosidad y grandes secciones eficaces alcanzadas en el LHC permiten más pruebas de alta precisión de la Cromodinámica Cuántica (QCD), interacciones electrodébiles, y física del sabor, y al mismo tiempo, la investigación de numerosos modelos de nueva física.

Los fotones prompt con momentos transversales grandes constituyen sondas incoloras de la interacción dura, y su producción en colisiones protón-protón (pp), $pp \to \gamma + X$, proporciona un campo de pruebas para QCD perturbativa (pQCD) a grandes escalas de la interacción dura (Q^2) y en un rango amplio de la fracción del momento del partón entrante (x). Los fotones prompt se definen como aquellos que no son secundarios de las desintegraciones de hadrones, y se entiende que su producción proviene de dos procesos: aquel en el que el fotón surge directamente de la interacción dura (el proceso de fotones directos), y aquel en el que el fotón es emitido en la fragmentación de un partón de momento transversal (p_{T}) alto (el proceso de fotones de fragmentación).

En colisiones pp, debido a la abundancia de fotones que provienen de desintegraciones de hadrones y a la contribución del proceso de fragmentación, la producción de fotones prompt se estudia requiriendo que los fotones estén aislados. Mediciones de la producción inclusiva de fotones aislados en colisiones pp a energías del centro de masas ($\sqrt{s}$) de $7, 8$ y 13 TeV han sido publicadas por las colaboraciones ATLAS [8–10] y CMS [11, 12].

En el LHC, la contribución dominante a la producción de fotones prompt surge del proceso $qg \to q\gamma$. Como consecuencia, la producción es sensible a la densidad de gluones en el protón [13–15] incluso al orden principal (LO) en pQCD. La medición de la sección eficaz de producción de fotones aislados inclusivos en diferentes rangos de la energía transversal (E_{T}^{γ}) y pseudorapidez del fotón (η^{γ}) proporciona información detallada sobre las funciones de distribución de partones (PDF) en diferentes regiones de Q^2 y x, que al LO es igual a $2\,\frac{E_{\mathrm{T}}^{\gamma}}{\sqrt{s}}\,e^{\pm\eta^{\gamma}}$. Un estudio reciente [16] ha incluido la medición de ATLAS de

la producción de fotones inclusivos aislados a 8 TeV [9] en un ajuste de QCD del orden siguiente al siguiente al principal (NNLO) para obtener las PDF del protón; la figura 1.3 muestra la cobertura cinemática de los datos de ATLAS a 8 TeV en el plano (x, Q^2). En la ref. [16] se concluye que la inclusión de los datos de ATLAS en el análisis global NNPDF3.1 NNLO no crea ninguna tensión con los otros conjuntos de datos incluidos en el ajuste; además, las restricciones en la PDF del gluón obtenidas de la inclusión de los datos de ATLAS son consistentes con las obtenidas del resto de los conjuntos de datos en el análisis NNPDF3.1. La figura 1.4 muestra una reducción moderada de la incertidumbre de la PDF del gluón en la región $10^{-3} \lesssim x \lesssim 0.4$, que es consistente con la cobertura cinemática proporcionada por los datos de ATLAS a 8 TeV representados en la figura 1.3.

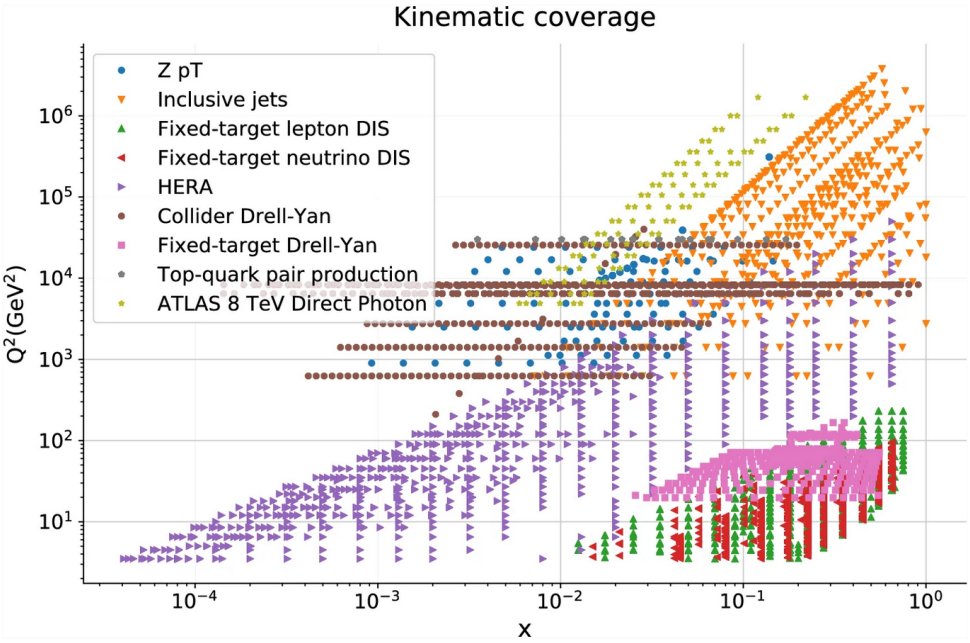


Figure 1.3: La cobertura en el plano cinemático (x, Q^2) de la medición de fotones de ATLAS a 8 TeV [9] (usando cinemática a LO), comparada con el conjunto de datos usado en el ajuste global NNPDF3.1. De la ref. [16].

Las mediciones inclusivas de la producción de fotones prompt también proporcionan puntos de referencia para investigar nuevos enfoques de la radiación de partones [17], correcciones al orden siguiente al principal (NLO) en QCD con duchas partónicas emparejadas [18], la relevancia de los logaritmos de umbral en correcciones de QCD y electrodébiles [19] y correcciones de QCD a NNLO [20–22]. Más específicamente, la producción de fotones prompt, que es menos sensible a los efectos de hadronización que la producción de *jets*, puede ser empleada para buscar una descripción alternativa de QCD basada en el enfoque de factorización k_t combinando amplitudes fuera de la capa de masas y densidades de partones dependientes del momento transversal del partón [17]. Los datos de fotones prompt permiten la investigación de la contribución de fragmentación, que puede realizarse mediante la inclusión de funciones de fragmentación o a través de la inclusión de

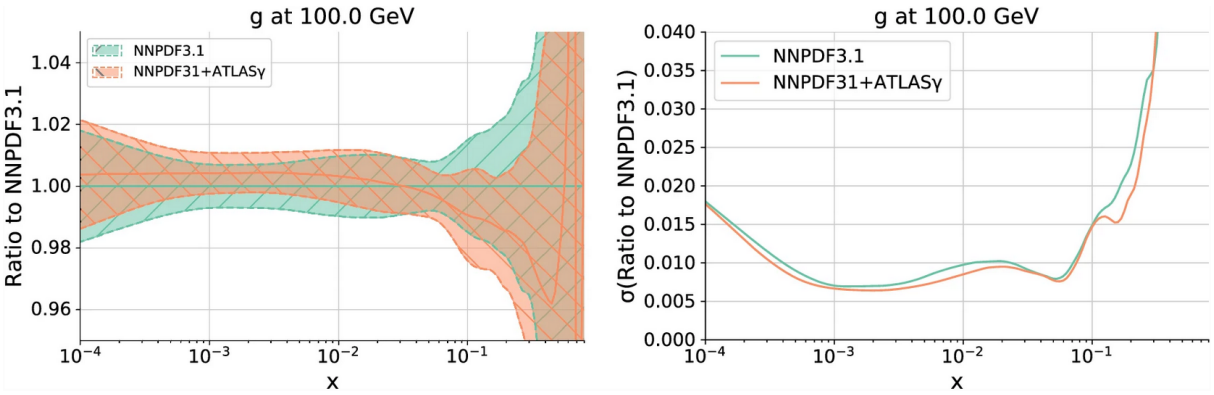


Figure 1.4: A la izquierda se muestra la comparación de la PDF del gluón a $Q = 100$ GeV entre los ajustes NNPDF3.1 y NNPDF3.1 + ATLASγ, normalizados al valor central del primero. A la derecha se muestra la correspondiente incertidumbre relativa de un sigma de las PDF en casa caso. De la ref. [16].

duchas de partones [18]. Es probable que las correcciones electrodébiles jueguen un papel importante en la escala de los TeV, por lo que las mediciones a energías transversales del fotón en esa región ayudarían a descubrir esa clase de fenómenos [19]. Las mediciones de fotones prompt en el LHC están caracterizadas por tener incertidumbres pequeñas que demandan cálculos teóricos precisos, e.g. cálculos a NNLO en pQCD [20–22], para poder aprovechar al máximo estos datos. Las mediciones que involucran fotones aislados también han sido usadas para restringir las contribuciones de nuevas partículas ligeras que se desintegran en un par de fotones [23].

Los jets son una característica predominante del estado final en las interacciones de pp de alta energía en el LHC. Están constituidos por duchas colimadas de hadrones y representan elementos importantes en muchas mediciones del SM y en búsquedas de nuevos fenómenos. Los jets se reconstruyen usando algoritmos de agrupamiento ejecutados sobre un conjunto de cuadrivectores de entrada, que típicamente se obtienen a partir de depósitos de energía asociados topológicamente, trazas de partículas cargadas, o partículas simuladas. Las mediciones de la escala y la resolución de la energía de los jets, así como de sus incertidumbres asociadas, son, por lo tanto, esenciales para las mediciones de precisión y para búsquedas sensibles a nueva física. Las calibraciones tienen en cuenta las características del detector, el algoritmo de reconstrucción de los jets, la fragmentación de los jets, y el entorno concurrido de toma de datos que resulta de las interacciones múltiples de pp, denominado como *pile-up*. Finalmente, se emplean técnicas *in situ* para medir la diferencia en la respuesta y la resolución de los jets entre los datos y el Monte Carlo (MC), y se aplican correcciones residuales solamente a los jets en los datos.

La producción de sucesos con un fotón prompt más un jet, $pp \to \gamma + \text{jet} + X$, también representa un test importante de las predicciones de pQCD. Además, el estudio de las correlaciones angulares entre el fotón y el jet se puede usar para restringir las funciones de fragmentación del fotón [7]. Sucesos con un fotón más un jet se pueden emplear para

calibrar la escala y la resolución de la energía de los jets aprovechando la conservación del momento en el plano transversal y la medición precisa de la energía del fotón en el calorímetro electromagnético (EM) [24–26]. Además, un conocimiento preciso de los procesos de producción de fotón más jet es útil para validar los generadores de MC usados para estudios de los *backgrounds* en búsquedas de física más allá del SM que involucran fotones, como por ejemplo para la búsqueda de nuevos fenómenos en estados finales con un fotón y un jet [27, 28]. Mediciones de la producción de un fotón aislado más uno o más jets en colisiones pp a $\sqrt{s} = 7, 8$ y 13 TeV han sido publicadas por las colaboraciones ATLAS [29–33] y CMS [12, 34–36].

La medición de la sección eficaz diferencial, $\mathrm{d}\sigma/\mathrm{d}E_{\mathrm{T}}^{\gamma}$, en tres intervalos de rapidez del jet y para dos configuraciones angulares diferentes del sistema fotón más jet, es decir, configuraciones de mismo signo ($\eta^{\gamma} \cdot y^{\mathrm{jet}} \geq 0$) y signo opuesto ($\eta^{\gamma} \cdot y^{\mathrm{jet}} < 0$), proporciona acceso a las PDF en diferentes regiones de Q^2 y x, que en la aproximación del orden principal viene dada por $\frac{E_{\mathrm{T}}^{\gamma}}{\sqrt{s}}\left(e^{\pm\eta^{\gamma}} + e^{\pm y^{\mathrm{jet}}}\right)$. Por lo tanto, la medición de la sección eficaz diferencial en estas configuraciones proporciona la posibilidad de identificar regiones del espacio de fases que son especialmente sensibles a las PDF del protón. Además, la separación del espacio de fases de la medición en estas configuraciones angulares diferentes permite la comparación entre los datos y las predicciones teóricas en regiones donde la contribución relativa de la componente de fragmentación a la sección eficaz total es diferente.

La estructura de la tesis está organizada como sigue. El capítulo 2 describe el acelerador LHC haciéndo énfasis en el detector ATLAS. El capítulo 3 describe los fundamentos del marco teórico subyacente al trabajo presentado en esta tesis; también se presenta una descripción de los algoritmos de los jets y una vista general de la teoría de producción de fotones prompt. El capítulo 4 describe las características básicas de los generadores de MC y las muestras empleadas específicamente para esta tesis. El capítulo 5 describe la reconstrucción, calibración e identificación de fotones en ATLAS. El capítulo 6 describe la reconstrucción de los jets y su calibración en ATLAS. Los siguientes capítulos de esta disertación presentan varias mediciones de la producción inclusiva y en asociación con un jet de fotones prompt aislados en colisiones de pp a $\sqrt{s} = 13$ TeV con el detector ATLAS en el LHC. El primero de estos estudios es la medición de la sección eficaz inclusiva de fotones aislados usando $36.1\,\mathrm{fb}^{-1}$ de los datos de ATLAS, que fue publicada en la ref. [37], y se presenta en el capítulo 7. Una extensión de la anterior medición empleando toda la luminosidad integrada de ATLAS durante el *Run II* fue realizada con posterioridad, incrementando el número de regiones de pseudorapidez del fotón medidas de cuatro a seis, y midiendo la sección eficaz diferencial en E_{T}^{γ} para dos tamaños del cono de aislamiento del fotón diferentes. Esta medición se encuentra actualmente bajo revisión por la colaboración ATLAS y se presenta en el capítulo 8. Un estudio de la calibración de la escala de la energía de los jets (JES) y de su resolución (JER) usando técnicas in situ en estados finales en los que un jet recula contra un fotón bien medido se presenta en el capítulo 9. La calibración de la JES (JER) presentada en esta tesis fue realizada usando $80\,\mathrm{fb}^{-1}$ ($58.5\,\mathrm{fb}^{-1}$) de los datos de ATLAS a $\sqrt{s} = 13$ TeV. La medición de la

JES, recientemente publicada en la ref. [26], representa una mejora en términos de las incertidumbres sistemáticas y estadística gracias al incremento de la luminosidad integrada y a las mejoras de las recomendaciones de la colaboración con respecto a la anterior calibración a $\sqrt{s} = 13$ TeV [25]. Finalmente, en el capítulo 10, la medición de la sección eficaz de producción de un fotón aislado más un jet empleando $138.4\,\text{fb}^{-1}$ se documenta en detalle. Esta medición está en curso. La primera de estas mediciones, incluida en el capítulo 7, representó una mejora en términos de incertidumbres estadística y sistemática, extendiendo el alcance en E_{T}^{γ} con respecto a la medición anterior a $\sqrt{s} = 13$ TeV. Las dos últimas mediciones presentadas en los capítulos 8 y 10 respectivamente, se benefician de la inclusión de toda la luminosidad integrada del Run II, proporcionando valores superiores en E_{T}^{γ} con incertidumbres estadísticas inferiores y también beneficios debido a las incertidumbres sistemáticas reducidas gracias a la continua mejora en términos de precisión de los grupos de rendimiento combinado de ATLAS que proporcionan las recomendaciones para evaluar las incertidumbres sistemáticas clave para estos análisis. Por lo tanto, estas mediciones formarán parte del legado de la colaboración ATLAS en su exploración de la frontera de energía de los 13 TeV.

Chapter 2

Experimental set-up

In this chapter, the most relevant characteristics of the LHC accelerator [1] are introduced, as well as a description of the ATLAS detector [4]. An overview of the trigger system based on ref. [38, 39], is also provided at the end of this chapter.

2.1 The LHC accelerator

The LHC is a proton-proton accelerator designed to run at high energies and luminosities. It was built by CERN between 1998 and 2008, and consists of a two-ring superconducting-hadron accelerator and collider installed in the existing 26.7 km tunnel that was constructed between 1984 and 1989 for the Large Electron-Positron Collider (LEP) [40]. The tunnel has eight straight sections and eight arcs and lies between 45 and 170 m below the surface on a plane inclined at 1.4% sloping towards the *Léman* lake.

The LHC design depends on some basic principles linked with the latest technology. Being a particle-particle collider, there are two rings with counter-rotating beams, unlike particle-antiparticle colliders that can have both beams sharing the same phase space in a single ring. The tunnel geometry, originally designed for the LEP machine, has eight crossing points, flanked by long straight sections for radio frequency (RF) cavities that compensated the high synchrotron radiation losses. A proton machine such as the LHC does not have the same synchrotron radiation problem and could have longer arcs and shorter straight sections for the same circumference. Of the eight possible interaction regions, only four were equipped for the LHC [41].

The size of the internal diameter of the tunnel in the arcs makes extremely difficult to install two completely separated proton rings. This hard limit led to the adoption of the twin-bore magnet design. Beams are kept on their circular paths by using 1232 dipole magnets, while an additional 392 quadrupole magnets are used to keep the beams focused. The magnet system makes use of a technology based on NbTi Rutherford cables, and cools the magnets to a temperature below 2 K using superfluid helium, operating at

fields above 8 T. The external view of one of the dipoles is shown in figure 2.1.

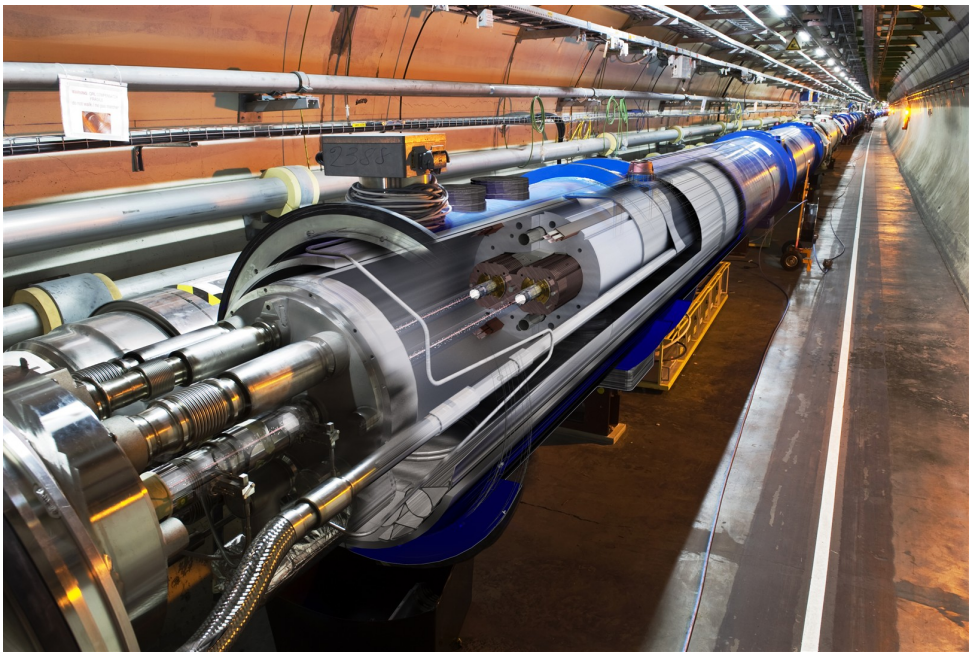


Figure 2.1: 3D view photo of the LHC Machine. From ref. [42].

There are two transfer tunnels, each of approximately 2.5 km in length, connecting the LHC to the CERN accelerator complex that acts as an injector. A description is provided in the following lines for the sequence of accelerators that operated during the Run II in the LHC. The proton source is a simple bottle of hydrogen gas. An electric field is used to strip hydrogen atoms of their electrons to yield protons. First, the Linear Particle Accelerator (Linac 2), accelerates the protons to the energy of 50 MeV. The beam is then injected into the Proton Synchrotron Booster (PSB), which accelerates the protons to 1.4 GeV, followed by the Proton Synchrotron (PS), which pushes the beam to 25 GeV. Protons are then sent to the Super Proton Synchrotron (SPS) where they are accelerated to 450 GeV. The protons are finally transferred to the two beam pipes of the LHC. It takes 4 minutes and 20 seconds to fill each LHC ring, and 20 minutes for the protons to reach their maximum energy of 6.5 TeV. Beams circulate for many hours inside the LHC beam pipes under normal operating conditions, until they are brought into collision inside four detectors: ALICE (A Large Ion Collider Experiment) [43], ATLAS [4], CMS [5] and LHCb (Large Hadron Collider beauty) [44], where the total energy at the collision point is equal to 13 TeV. Figure 2.2 shows the layout of the accelerator chain.

2.1.1 Main machine layout and performance

Performance goals

The main goal of the LHC is to reveal physics beyond the Standard Model (SM) with

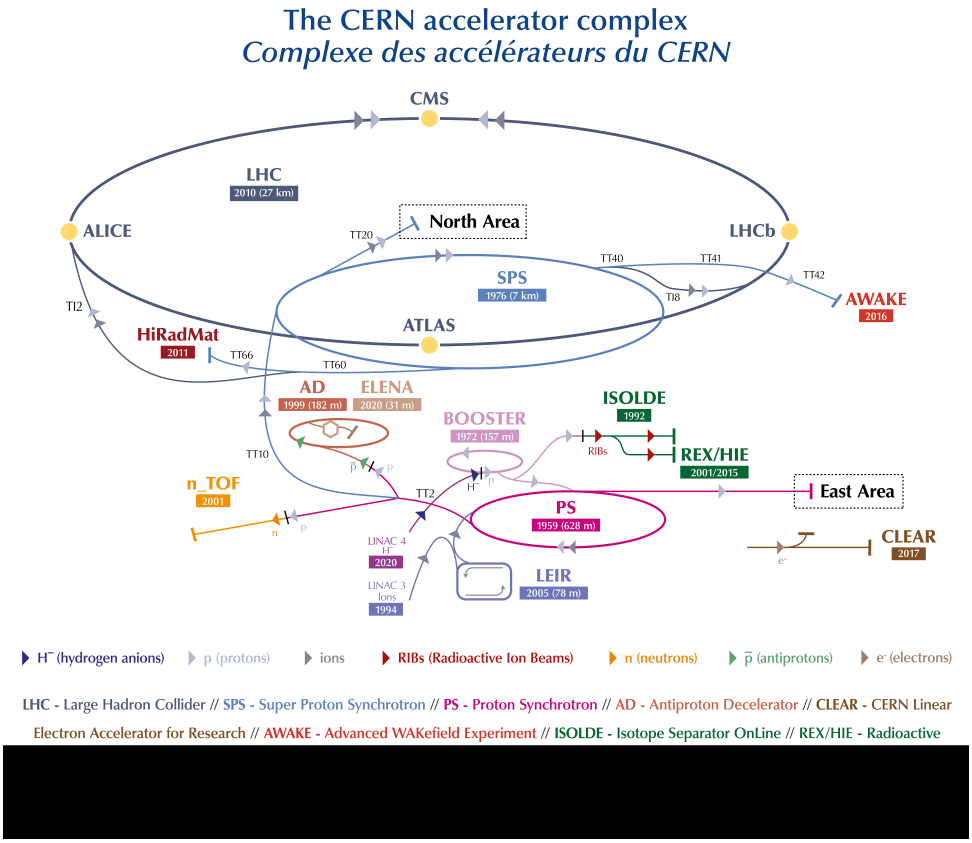


Figure 2.2: Layout of the CERN accelerator complex. From ref. [45].

centre-of-mass collision energies ($\sqrt{s}$) of up to 14 TeV. The luminosity of an accelerator is one of the most important parameters, being a measurement of the number of collisions per unit of area and time. The number of events per second generated in the LHC collisions is given by

$$N_{\text{event}} = \mathcal{L}\sigma_{\text{event}} ,$$

where σ_{event} is the cross section for the event under study and $\mathcal{L}$ is the machine luminosity. The machine luminosity depends only on the beam parameters and can be written for a Gaussian beam distribution as

$$\mathcal{L} = \frac{N_{\text{b}}^2 n_{\text{b}} f_{\text{rev}} \gamma_{\text{r}}}{4\pi\varepsilon_{\text{n}}\beta^*} F ,$$

where N_{b} is the number of particles per bunch, n_{b} the number of bunches per beam, f_{rev} the revolution frequency, γ_{r} the relativistic gamma factor, ε_{n} the normalised transverse beam emittance, β^* the beta function at the collision point, and F the geometric luminosity reduction factor due to the crossing angle at the interaction point (IP)

$$F = \left(1 + \left(\frac{\theta_c \sigma_z}{2\sigma^*}\right)^2\right)^{-1/2} .$$

where θ_c is the full crossing angle at the IP, σ_z the root-mean-squared (RMS) bunch length, and σ^* the transverse RMS beam size at the IP. The above expression assumes round beams, with $\sigma_z \ll \beta$, and with equal beam parameters for both beams. The exploration of rare events in the LHC collisions therefore requires both high beam energies and high beam intensities.

Typical values of the previous parameters are: $N_b \sim 10^{11}$, $f_{rev} \sim 11$ kHz, $\beta^* \sim 0.55$ m and $F = 0.85$. The operating parameters have been optimised to provide a higher luminosity during the different years of data-taking. The peak instantaneous luminosity for proton operation surpassed the design value for the LHC of $1 \times 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ in 2016.

The integral of the delivered luminosity over time is called integrated luminosity. It is a measurement of the collected data size, and can be written as

$$L = \int \mathcal{L}\,\mathrm{d}t\,.$$

The LHC has two high luminosity experiments, ATLAS and CMS, for proton operation. There are also two low luminosity experiments: LHCb, for B-physics, and TOTEM (TOTal Elastic and diffractive cross-section Measurements) [46] for the detection of protons from elastic scattering at small angles. In addition to proton beams, the LHC also operates with heavy ions, and ALICE is a dedicated LHC ion experiment.

Pile-up performance limitation

The performance of the detector can be severely affected by the high number of inelastic *pp* interactions in the same bunch crossing. Pile-up corresponds to additional information recorded by the detector during the hard-scattered interaction of interest. Two categories of pile-up can be defined, namely the out-of-time pile-up, due to interactions from neighbouring bunch crossings, and the in-time pile-up, resulting from additional interactions in the same bunch crossing. The amount of pile-up has an impact in the reconstruction of objects in the detector, such as tracks or vertex reconstruction [47]. The number of *pp* interactions per bunch crossing follows a Poisson distribution with mean value μ. During a fill, μ decreases with decreasing beam intensity and increasing emittance, and μ also varies between bunches. The number of interactions averaged over all bunch crossings and averaged over the data analysed is referred to as $\langle\mu\rangle$. The average number of interactions per bunch crossing $\langle\mu\rangle$ increased from 9.1 in 2011 to 20.7 in 2012 [48].

In data, μ can be calculated using

$$\mu = \frac{\mathcal{L}\sigma_{inel}}{n_b f_{rev}} ,$$

where σ_{inel} is the total inelastic cross section (80 mb at $\sqrt{s} = 13$ TeV) [49] and the rest of the parameters are defined in section 2.1.1. The uncertainty on μ depends on the uncertainties on the luminosity and the total inelastic cross section. The luminosity measurement is performed with dedicated detectors and calibrated using special LHC fills.

The delivered luminosity (see *Performance goals* in section 2.1.1) with 2011 and 2012 data as a function of $\langle\mu\rangle$ is shown in figure 2.3, while for 2015, 2016, 2017, 2018, and the full Run II data it is shown in figure 2.4. The mean number of interactions per crossing corresponds to the mean of the Poisson distribution of the number of interactions per bunch crossing calculated for each bunch. In 2011, $\langle\mu\rangle = 9.1$; in 2012, $\langle\mu\rangle = 20.7$; in 2015, $\langle\mu\rangle = 13.4$; in 2016, $\langle\mu\rangle = 25.1$; in 2017, $\langle\mu\rangle = 37.8$; in 2018, $\langle\mu\rangle = 36.1$, and for the full Run II, $\langle\mu\rangle = 33.7$.

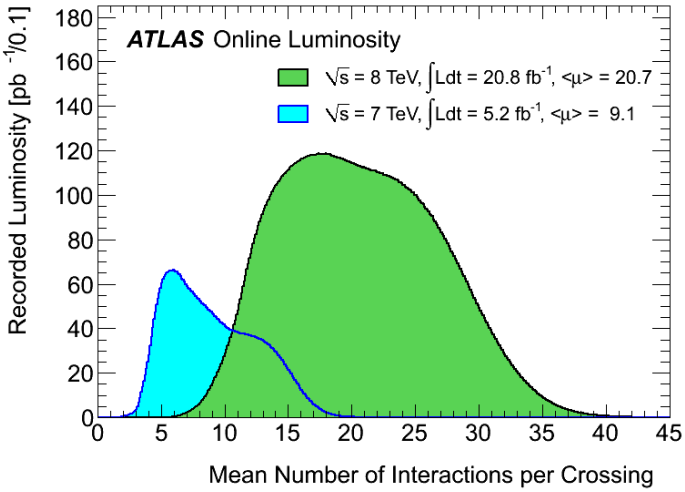


Figure 2.3: The luminosity-weighted distribution of the mean number of interactions per crossing for the 2011 and 2012 data. This shows the full 2011 and 2012 pp runs. The integrated luminosities and $\langle\mu\rangle$ values are given in the figure. From ref. [50].

2.1.2 Timeline of the LHC and ATLAS data-taking

The construction of the LHC was finalised in 2008. It was expected to operate at $\sqrt{s} =$ 10 TeV by the end of 2008. Unfortunately, on September 19$^{\text{th}}$ 2008, a magnet quench occurred in about 100 bending magnets in sectors 3 and 4, where an electrical connection fault between two magnets led to a loss of approximately 6 tonnes of liquid helium, breaking the vacuum conditions of the beam pipe [52]. This event shaped the evolution of the LHC data-taking during the following years. To run with full safety, the LHC operated up to $\sqrt{s} = 8$ TeV until the first Long Shutdown (LS1) in 2013 [53]. A timeline of the LHC and ATLAS data-taking is summarised in the following lines:

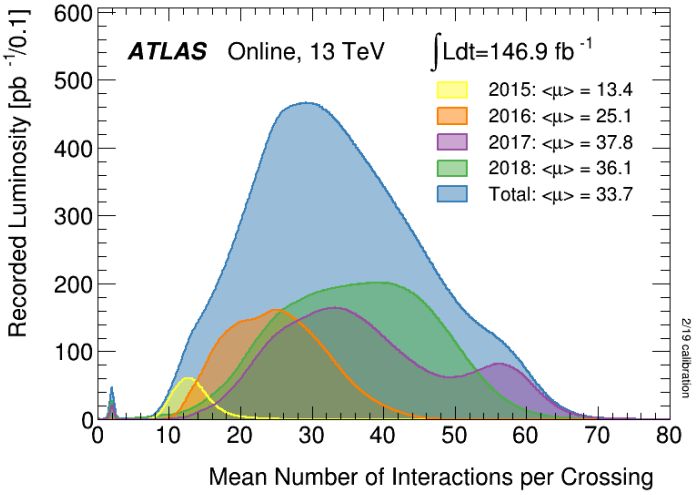


Figure 2.4: The luminosity-weighted distribution of the mean number of interactions per crossing for the $2015 - 2018$ pp collision data at $\sqrt{s} = 13$ TeV. All data recorded by ATLAS during stable beams is shown, and the integrated luminosities and $\langle\mu\rangle$ values are given in the figure. From ref. [51].

– During 2009, the first LHC beams circulated at $\sqrt{s} = 900$ GeV, which allowed ATLAS to obtain approximately $9\,\mu\text{b}^{-1}$ enabling important calibration work to be carried out. On December 16^{th} 2009, the LHC ended its first full period of operation. Collisions at 2.36 TeV set a new world record.

– After a short technical stop, beams circulated again on February 28^{th} 2010. On March 19^{th}, two 3.5 TeV proton beams successfully circulated in the LHC for the first time. This represented an important step on the way to the start of the LHC research programme. The number of bunches was increased up to 368, with 1.0×10^{11} protons per bunch, reaching a maximum peak luminosity of $2.1 \times 10^{32}\,\text{cm}^{-2}\text{s}^{-1}$, as shown in figure 2.5. A total integrated luminosity of $45\,\text{pb}^{-1}$ was recorded by ATLAS.

– During 2011, the $\sqrt{s}$ was maintained at 7 TeV while the number of bunches was increased to 1380 and the bunch intensity to 1.45×10^{11} protons per bunch, reaching a maximum peak luminosity of $3.6 \times 10^{33}\,\text{cm}^{-2}\text{s}^{-1}$, as shown in figure 2.5. A total integrated luminosity of $5.08\,\text{fb}^{-1}$ was recorded by ATLAS.

– During 2012, the $\sqrt{s}$ was increased up to 8 TeV. The number of bunches was reduced to 1376 and the intensity increased up to 1.6×10^{11} protons per bunch, reaching a maximum peak luminosity of $7.7 \times 10^{33}\,\text{cm}^{-2}\text{s}^{-1}$, as shown in figure 2.5. A total integrated luminosity of $21.3\,\text{fb}^{-1}$ was recorded by ATLAS.

– Following the LS1 in 2013 and 2014 during which the LHC and, in particular, the superconducting magnet interconnections were consolidated, the machine was re-

commissioned at an energy of 6.5 TeV per beam in 2015, signalling the start of LHC Run II. The LHC Run II covered the years from 2015 until the second Long Shutdown (LS2) that started in December 2018 [54].

- During the summer of 2015, the first collisions at $\sqrt{s} = 13$ TeV were delivered by the LHC and recorded by ATLAS. This period was dedicated to establishing operation at 6.5 TeV per beam and with standard 25 ns bunch spacing to prepare for substantial luminosity production during the years running up to the LS2. The number of bunches per beam was increased to 2244, with 1.15×10^{11} protons per bunch, and a maximum peak luminosity of $5 \times 10^{33}\,\mathrm{cm^{-2}s^{-1}}$ was reached, as shown in figure 2.6(a). A total integrated luminosity of $3.9\,\mathrm{fb^{-1}}$ was recorded by ATLAS.

- The first year with substantial luminosity production was 2016. On June 26$^{\mathrm{th}}$, after careful optimisation of the machine settings and beam brightness in the injectors, the LHC attained its design luminosity of $1 \times 10^{34}\,\mathrm{cm^{-2}s^{-1}}$ with 2040 bunches per beam and 1.2×10^{11} protons per bunch. The preparation of the injector chain for the successful production of high brightness 25 ns beams resulted in a lower transverse emittance in combination with a reduction of the half crossing angle which gave rise to a further gradual increase of the peak luminosity as can be observed in figure 2.6(b), with a record peak luminosity of $1.38 \times 10^{34}\,\mathrm{cm^{-2}s^{-1}}$ and 2076 bunches per beam. A total integrated luminosity of $35.6\,\mathrm{fb^{-1}}$ was recorded by ATLAS.

- In 2017, after a short technical stop, with a magnet replacement in sector 1-2, the first beam was injected on April 29$^{\mathrm{th}}$. The initial set of 2556 number of bunches was reduced to 1916 to mitigate the observed abnormal background radiation and sudden beam losses. The intensity of the beam was maintained at 1.2×10^{11} protons per bunch. The beam parameters were further optimised[1] allowing efficient luminosity production, resulting in a new record of $2.09 \times 10^{34}\,\mathrm{cm^{-2}s^{-1}}$ for the peak luminosity as shown in figure 2.6(c). The pile-up in the ATLAS and CMS experiments was too high, consequently, the instantaneous luminosity was levelled to $1.5 \times 10^{34}\,\mathrm{cm^{-2}s^{-1}}$. A total integrated luminosity of $46.9\,\mathrm{fb^{-1}}$ was recorded by ATLAS.

- In 2018, the number of bunches was increased to 2556 with much lower beam losses observed than in 2017. Stable beams with 2556 bunches were reached on May 5$^{\mathrm{th}}$, with an intensity of 1.2×10^{11} protons per bunch. An approximately constant peak luminosity of $1.9 \times 10^{34}\,\mathrm{cm^{-2}s^{-1}}$ was achieved for most of the data-taking period, as can be observed in figure 2.6(d). The 2018 proton run ended with a record total integrated luminosity of $60.6\,\mathrm{fb^{-1}}$ recorded by ATLAS.

The cumulative total integrated luminosity versus time for the full Run I (Run II) is shown in figure 2.7 (2.8). The luminosity delivered to ATLAS in the different data-taking periods from 2011 to 2018 can be compared in figure 2.9. The data used for the

[1]For the 8b4e-BCS beam type, the number of bunches was further reduced to 1868, increasing the intensity to 1.25×10^{11} protons per bunch.

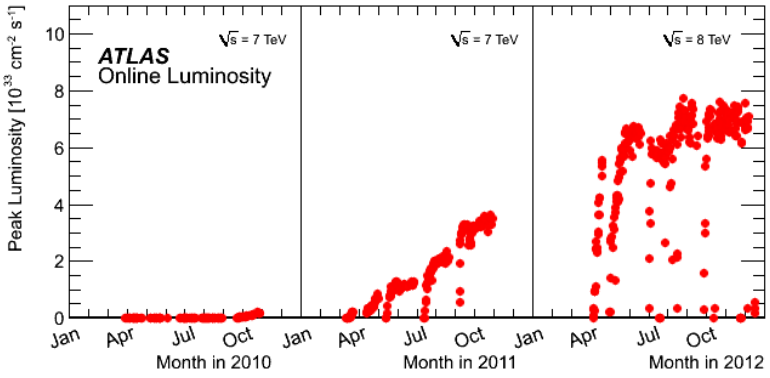


Figure 2.5: The peak instantaneous luminosity delivered to ATLAS per day versus time during the *pp* runs of 2010, 2011 and 2012. From ref. [50].

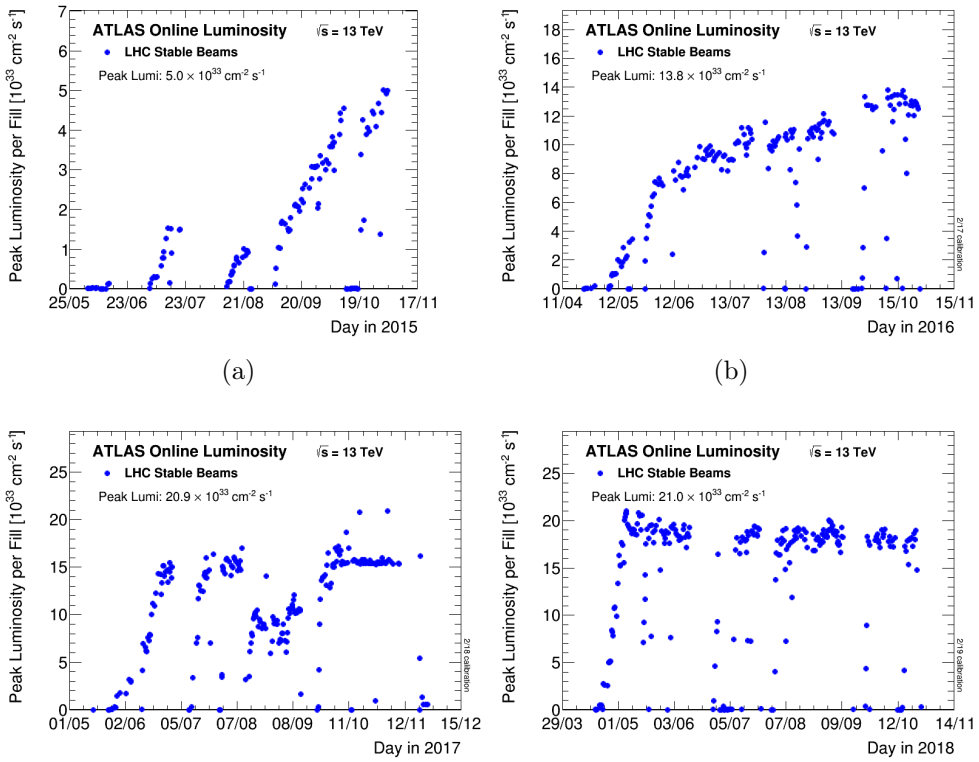


Figure 2.6: The peak instantaneous luminosity delivered to ATLAS during stable beams for *pp* collisions at $\sqrt{s} = 13$ TeV is shown for each LHC fill as a function of time in 2015 (a), 2016 (b), 2017 (c) and 2018 (d). From ref. [51].

results presented in this thesis were recorded during the 2015 and 2016 or the full Run II data-taking periods.

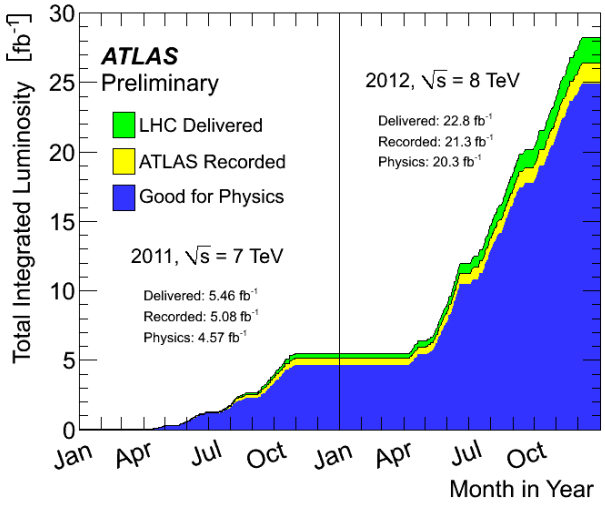


Figure 2.7: Cumulative luminosity versus time delivered to ATLAS (green), recorded by ATLAS (yellow), and certified to be good quality data (blue) during stable beams and for pp collisions at $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV in 2011 and 2012. From ref. [50].

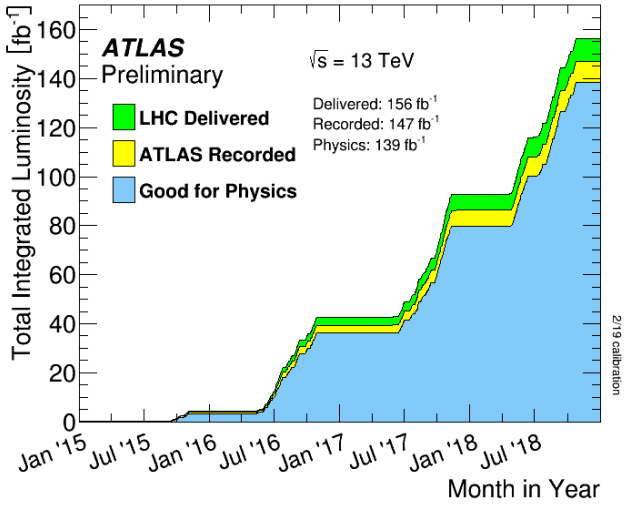


Figure 2.8: Cumulative luminosity versus time delivered to ATLAS (green), recorded by ATLAS (yellow), and certified to be good quality data (blue) during stable beams for pp collisions at $\sqrt{s} = 13$ TeV from 2015 to 2018. From ref. [51].

2.2 The ATLAS detector

The LHC at CERN promises a major step forward in the understanding of the fundamental nature of matter. The ATLAS experiment is a general-purpose detector for the LHC [4, 55], whose design was guided by the need to accommodate the wide spectrum of possible physics signatures. ATLAS was built for probing pp and AA collisions, with

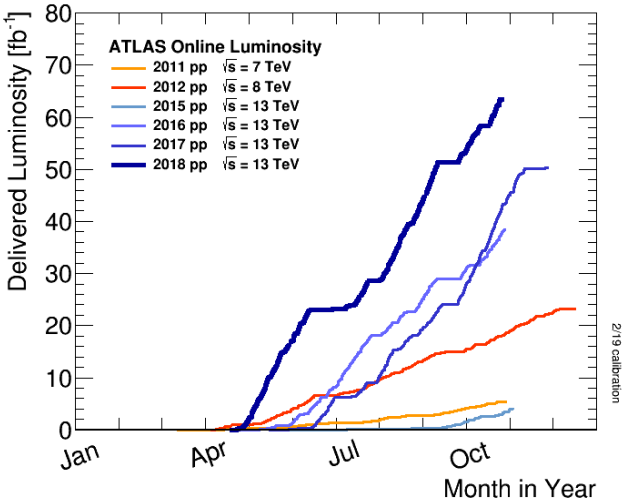


Figure 2.9: Cumulative luminosity versus day delivered to ATLAS during stable beams and for high-energy pp collisions from 2011 to 2018. From ref. [51].

the major benefit of exploring the TeV mass scale where ground-breaking discoveries are expected. The investigation of the electroweak symmetry breaking and, linked to this, the search for the Higgs boson, discovered in 2012 [2,3], as well as the search for Physics beyond the SM are on the focus. The high luminosity and increased cross sections at the LHC enable further high precision tests of QCD, electroweak interactions and flavour physics. This detector represents the work of a large collaboration of several thousands physicists, engineers, technicians, and students over a period of fifteen years of dedicated design, development, fabrication, and installation.

2.2.1 Physics requirements and detector overview

The coordinate system and nomenclature used to describe the ATLAS detector and the particles emerging from the pp collisions are briefly described in the following lines. The nominal IP is defined as the origin of the coordinate system, while the beam direction defines the z-axis and the x-y plane is transverse to the beam direction. The positive x-axis is defined as pointing from the IP to the centre of the LHC ring and the positive y-axis is defined as pointing upwards. The side-A of the detector is defined as that with positive z and side-C is that with negative z. The azimuthal angle ϕ is measured around the beam axis, and the polar angle θ is the angle from the beam axis. The pseudorapidity is defined as $\eta = -\ln[\tan(\theta/2)]$ (in the case of massive objects such as jets, the rapidity $y = 1/2 \ln[(E + p_z)/(E - p_z)]$ is used). The transverse momentum p_T, the transverse energy E_T and the missing transverse energy E_T^{miss} are defined in the x-y plane. The distance ΔR in the pseudorapidity-azimuthal angle space is defined as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

The large LHC luminosity and resulting interaction rate are needed because of the small cross sections expected for many processes of interest. However, with an inelastic pp cross section of 80 mb at $\sqrt{s} = 13$ TeV [49], the LHC faces a serious experimental difficulty as it implies that every candidate event for new physics is, on the average, accompanied by ≈ 30 inelastic events per bunch crossing separated only by 25 ns, as observed with the full Run II dataset. The nature of pp collisions introduces another difficulty, since QCD jet production cross sections dominate over many rare processes, requiring the identification of various experimental signatures such as $E_{\mathrm{T}}^{\mathrm{miss}}$ or secondary vertices.

These benchmark physics goals can be turned into a set of general requirements for the LHC detectors:

- Due to the experimental conditions at the LHC, the detector requires fast, radiation-hard electronics and sensor elements. In addition, high detector granularity is needed to handle the particle fluxes and to reduce the influence of overlapping events.
- Large acceptance in pseudorapidity with almost full azimuthal angle coverage is required.
- Good charged-particle momentum resolution and reconstruction efficiency in the inner tracker are essential. For offline tagging of τ-leptons and b-jets, vertex detectors close to the interaction region are required to observe secondary vertices.
- Very good electromagnetic calorimetry for electron and photon identification and measurements, complemented by full-coverage hadronic calorimetry for accurate jet and $E_{\mathrm{T}}^{\mathrm{miss}}$ measurements are important requirements.
- Good muon identification and momentum resolution over a wide range of momenta and the ability to determine unambiguously the charge of high-p_{T} muons are fundamental requirements.
- Highly efficient triggering on low-p_{T} objects with sufficient background rejection is a prerequisite to achieve an acceptable trigger rate for most physics processes of interest.

The overall ATLAS detector layout is shown in figure 2.10 and its main performance goals are listed in table 2.1. The ATLAS detector has a forward-backward symmetry with respect to the IP. The magnet configuration comprises a thin superconducting solenoid surrounding the inner-detector cavity, and three large superconducting toroids (one barrel and two end-caps) are arranged with an eight-fold azimuthal symmetry around the calorimeters. These fundamental choices drove the design of the rest of the detector.

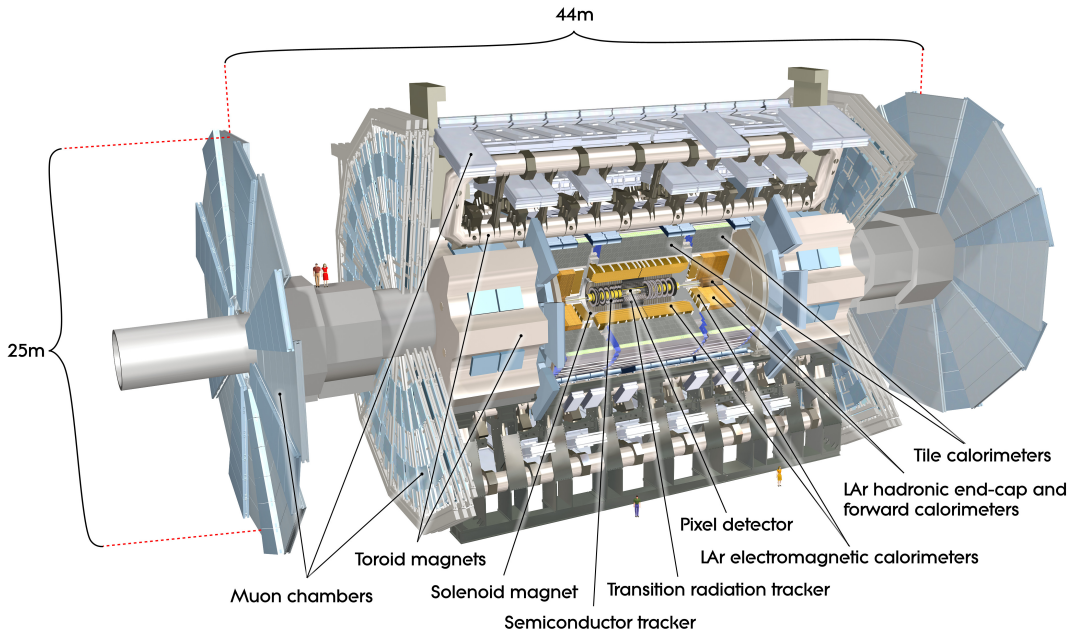


Figure 2.10: Cut-away view of the ATLAS detector. The dimensions of the detector are 25 m in height and 44 m in length. The overall weight of the detector is approximately 7000 tonnes. From ref. [56].

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

Table 2.1: General performance goals of the ATLAS detector. The units for E and p_T are in GeV.

2.2.2 Tracking

The ATLAS Inner Detector (ID) is the closest sub-detector to the beams interaction. It is designed to provide hermetic and robust pattern recognition, excellent momentum resolution, electron identification and both primary and secondary vertex measurements for charged tracks above a given p_T threshold. The ID consists of three independent but complementary sub-detectors. The use in conjunction of high-resolution Semi-Conductor pixel and silicon microstrip Trackers (SCT) in the inner part of the detector and straw tubes of the Transition Radiation Tracker (TRT) offer these features.

The layout of the ID is illustrated in figure 2.11. It is immersed in a 2 T magnetic

field generated by the central solenoid allowing the measurement of charged particles p_{T} using the curvature produced by the magnetic field. The solenoid is a single layer of Al stabilised Nb/Ti/Cu conductor cooled by liquid Helium at a temperature of 4.5 K, which extends over a length of 5.3 m with a diameter of 2.5 m between the ID and the calorimeter. The precision-tracking detectors (pixels and SCT) cover the region $|\eta| < 2.5$. In the barrel region, they are arranged on concentric cylinders around the beam axis, while in the end-cap regions, they are located on disks perpendicular to the beam axis. The highest granularity is achieved around the vertex region using silicon pixel detectors. Figure 2.12 shows the sensors and structural elements in the ID barrel region, and table 2.2 summarises the main detector characteristics.

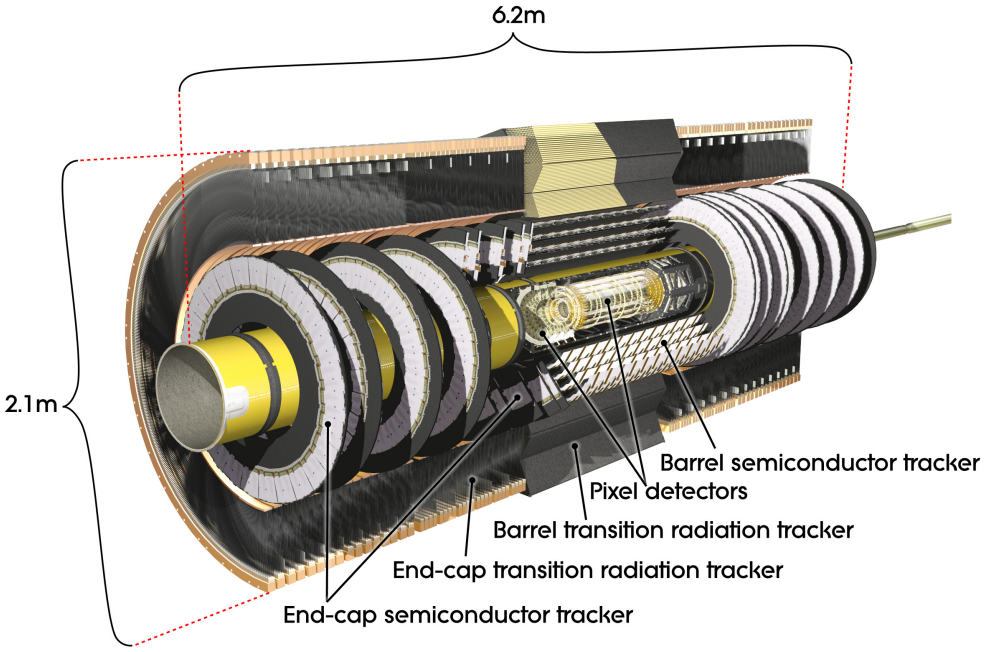


Figure 2.11: Cut-away view of the ATLAS inner detector. From ref. [56].

The pixel detector

The pixel detector is designed for precise tracking measurement of charged particles with 40 MHz bunch crossing identification and for vertices reconstruction. It consists of three cylindrical barrel layers in the central region $|\eta| < 1.37$, and three disks on both sides for the forward measurement. The pixel layers are segmented in $R - \phi$ and z, with typically three pixel layers crossed by each track. The three barrel layers are at a radius of 50.5 mm, 88.5 mm and 122.5 mm from the beam pipe. Each layer is composed of hundreds of modules consisting of 8 double pixel sensors directly coupled to the 16 readout. The pixel sensors are identical and have a minimum size in $R-\phi \times z$ of 50×400 μm^2, with a thickness of 250 μm. The intrinsic accuracies in the barrel are 10 μm ($R - \phi$) and 115 μm (z), and in the disks are 10 μm ($R - \phi$) and 115 μm (R). The pixel detector has approximately 80.4 million readout channels. During the LS1 a new pixel detector layer, the Insertable

B-Layer (IBL) [57], was added to the pixel detector between the new and smaller beam pipe and the previous innermost pixel layer, at 32 mm from the IP. Among the benefits of the introduction of this new layer, there are an increase in tracking robustness against failure of pixel modules, higher tagging efficiency of secondary vertices of hadrons with a bottom quark (b-tagging), and better vertexing performance and tracking precision due to the closer location to the IP.

SCT

For the SCT, eight strip layers (four space points) are crossed by each track. In the barrel region, this detector uses small-angle (40 mrad) stereo strips to measure both coordinates, with one set of strips in each layer parallel to the beam direction, measuring $R-\phi$. They consist of two 6.4 cm long daisy-chained sensors with a strip pitch of 80 μm. In the end-cap region, the detectors have a set of strips running radially and a set of stereo strips at an angle of 40 mrad. The mean pitch of the strips is also approximately 80 μm. The intrinsic accuracies per module in the barrel are 17 μm ($R-\phi$) and 580 μm (z) and in the disks are 17 μm ($R-\phi$) and 580 μm (R). The total number of readout channels in the SCT is approximately 6.3 million.

TRT

A large number of hits (typically 36 per track) is provided by the 4 mm diameter straw tubes of the TRT, which enables track-following up to $\eta = 2.0$. The TRT only provides $R-\phi$ information, for which it has an intrinsic accuracy of 130 μm per straw. In the barrel region, the straws are parallel to the beam axis and are 144 cm long, with their wires divided into two halves, approximately at $\eta = 0$. In the end-cap region, the 37 cm long straws are arranged radially in wheels. The total number of TRT readout channels is approximately 351000.

The combination of precision trackers at small radii with TRT at a larger radius gives very robust pattern recognition and high precision in both $R-\phi$ and z coordinates. The straw hits at the outer radius contribute significantly to the momentum measurement, since the lower precision per point compared to the silicon is compensated by the large number of measurements and longer measured track length.

The ID system provides tracking measurements in a range matched by the precision measurements of the EM calorimeter. The electron identification capabilities are enhanced by the detection of transition-radiation photons in the straw tubes. The semiconductor trackers also allow impact parameter measurements and vertexing for heavy-flavour and τ-lepton tagging. The secondary vertex measurement performance is enhanced by the innermost layer of the pixel detector.

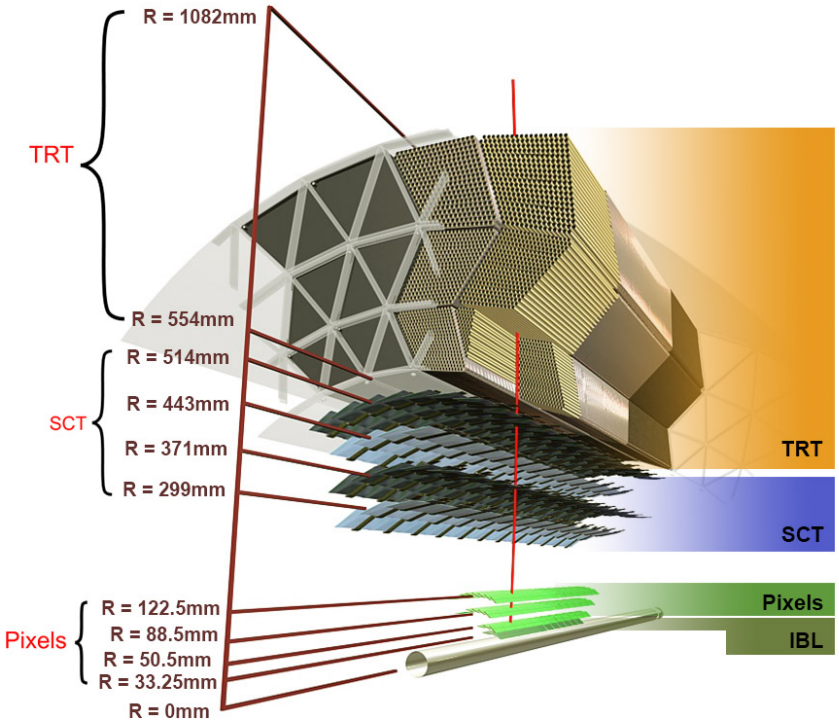


Figure 2.12: A 3D visualization of the structure of the barrel of the ID. The beam pipe, the IBL, the Pixel layers, the four cylindrical layers of the SCT and the three layers of the TRT barrel modules consisting of 72 straw layers are shown. From ref. [58].

Subdetector	Element size	Intrinsic resolution [μm]	Barrel layer radii [mm]	Disk layer $\lvert z\rvert$ [mm]
IBL	$50\mu\text{m} \times 250\mu\text{m}$	10×60	33.25	
Pixel	$50\mu\text{m} \times 400\mu\text{m}$	10×115	50.5, 88.5, 122.5	495, 580, 650
SCT	$80\mu\text{m}$	17	299, 371, 443, 514	from 839 to 2735
TRT	4 mm	130	from 554 to 1082	from 848 to 2710

Table 2.2: Summary of the main characteristics of the ID subdetectors. The intrinsic resolution of the IBL and the Pixel sensors are reported along $r-\phi$ and z, while for SCT and TRT only the resolution along $r-\phi$ is given. For SCT and TRT the element size refers to the spacing of the read-out strips and the diameter of the straw tube, respectively.

2.2.3 Calorimetry

The ATLAS calorimeters are designed to measure the energy that a particle loses as it passes through the detector. They consist of a number of sampling detectors with full ϕ-symmetry and coverage around the beam axis, alternating layers of an absorber, a dense material with which the incident particle interacts, and an active medium that provides a detectable signal.

A view of the sampling calorimeters is presented in figure 2.13. The pseudorapidity

coverage, granularity and segmentation in depth of the calorimeters are summarised in table 2.3. These calorimeters cover the range $|\eta| < 4.9$, using different techniques suited to the widely varying requirements of the physics processes of interest and of the radiation environment over this large η-range. Over the η region matched to the ID, the fine granularity of the EM calorimeter is ideally suited for precision measurements of electrons and photons. The coarser granularity of the rest of the calorimeter is sufficient to satisfy the physics requirements for jet reconstruction and $E_{\mathrm{T}}^{\mathrm{miss}}$ measurements.

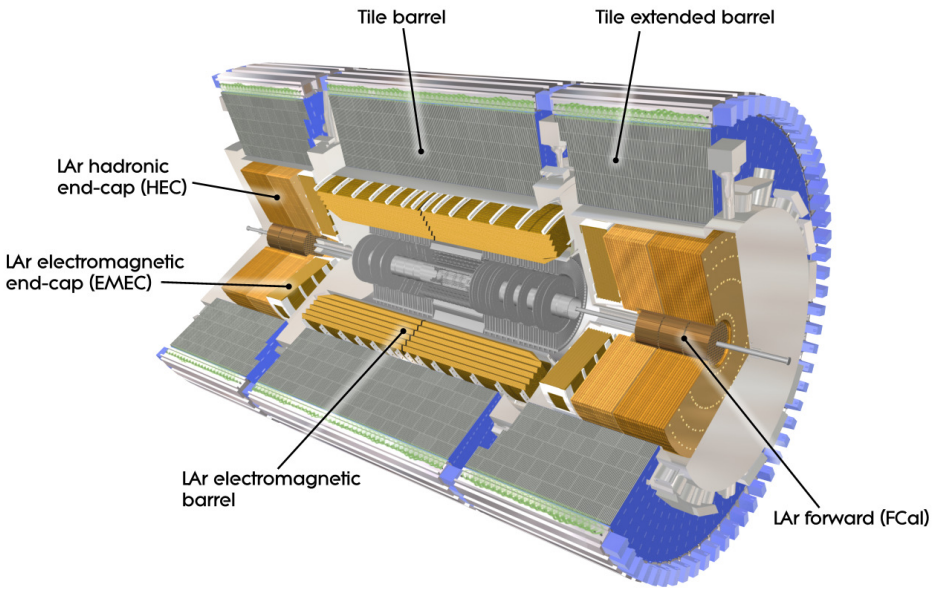


Figure 2.13: Cut-away view of the ATLAS calorimeter system. From ref. [56].

Calorimeters must provide an adequate containment for EM and hadronic showers and must also limit punch-through into the muon system. The total thickness of the EM calorimeter is greater than 22 radiation lengths[2] (X_0) in the barrel and $24X_0$ in the end-caps. The approximate 9.7 interaction lengths[3] (λ) of active calorimeter in the barrel

[2]The radiation length of a material is both (a) the mean distance over which a high-energy electron loses all but $1/e$ of its energy by bremsstrahlung, and (b) 7/9 of the mean free path for pair production by a high-energy photon. It is also the appropriate scale length for describing high-energy EM cascades. It can be approximated by the following formula

$$X_0 = \frac{716.405\,A}{Z(Z+1)\ln\left(\frac{287}{\sqrt{Z}}\right)}\,,$$

where Z is the atomic number and A is the mass number of the nucleus.

[3]The nuclear interaction length is the distance in a given substance after which an incident number of hadrons diminishes by a factor $1/e$ by nuclear interaction,

$$N(x) = N_0 \exp(-x/\lambda)\,,$$

where $N(x)$ is the number of non-interacting particles after a distance x, and N_0 is the number of

(10λ in the end-caps) are adequate to provide good resolution for high-energy jets. The total thickness of 11λ at $\eta = 0$ is sufficient to reduce the punch-through well below the irreducible level of prompt or decay muons. Together with the large η-coverage, this thickness ensures a good $E_{\mathrm{T}}^{\mathrm{miss}}$ measurement.

The calorimeters closest to the beam-line are housed in three cryostats, one barrel and two end-caps. The barrel cryostat contains the EM barrel calorimeter, whereas the two end-cap cryostats each contain an EM End-Cap calorimeter (EMEC), a Hadronic End-Cap calorimeter (HEC), located behind the EMEC, and a Forward Calorimeter (FCal) to cover the region closest to the beam. Electromagnetic calorimetry is provided by the barrel and end-cap high-granularity LAr calorimeters within the region $0 < |\eta| < 3.2$. The calorimeters have several active layers in depth, three in the precision-measurement region ($0 < |\eta| < 2.5$) and two in the higher-η region ($2.5 < |\eta| < 3.2$) and in the overlap between the barrel and the end-cap. The calorimeter system also has EM coverage at higher η ($3.1 < |\eta| < 4.9$) provided by the FCal. Furthermore, in the region ($0 < |\eta| < 1.8$), the EM calorimeters are complemented by presamplers, an instrumental argon layer, which provides a measurement of the energy lost in front of the EM calorimeters. Hadronic calorimetry is provided by a steel/scintillator-tile calorimeter composed of three parts, one central barrel and two extended barrels that cover the range $|\eta| < 1.7$, two copper/LAr detectors which cover the region $1.5 < |\eta| < 3.2$ and the FCal, a copper-tungsten/LAr detector. The hadronic calorimetry thus reaches one of its main design goals, namely coverage up to $|\eta| = 4.9$.

2.2.3.1 LAr electromagnetic calorimeter

The EM calorimeter is divided into a barrel part ($|\eta| < 1.475$) and two end-cap components ($1.375 < |\eta| < 3.2$), each housed in their own cryostat. Each end-cap wheel is subdivided into 8 wedge-shaped modules. Eight out of the 16 modules of the EM end-cap calorimeter were assembled at UAM[4]. The position of the central solenoid in front of the EM calorimeter demands optimisation of the material to achieve the desired calorimeter performance. As a consequence, the central solenoid and the LAr calorimeter share a common vacuum vessel, thereby eliminating two vacuum walls. The barrel calorimeter consists of two identical half-barrels, separated by a small gap (4 mm) at $z = 0$. Each end-cap calorimeter is mechanically divided into two coaxial wheels: an outer wheel covering the region $1.375 < |\eta| < 2.5$, and an inner wheel covering the region $2.5 < |\eta| < 3.2$. The EM calorimeter is a lead-LAr detector with accordion-shaped kapton electrodes and lead absorber plates over its full coverage. The accordion geometry provides complete ϕ symmetry without azimuthal cracks. The lead thickness in the absorber plates has been optimised as a function of η in terms of EM calorimeter performance in energy resolution.

The EM calorimeter is composed of three sampling layers for $|\eta| < 2.5$. The first layer has a thickness of about 4.4 X_0. In the ranges $|\eta| < 1.4$ and $1.5 < |\eta| < 1.8$, it is

incoming particles.

[4]Universidad Autónoma de Madrid.

LAr EM calorimeter				
	Barrel		**End-cap**	
Number of layers and $\|\eta\|$ coverage				
Presampler	1	$\|\eta\| < 1.52$	1	$1.5 < \|\eta\| < 1.8$
Calorimeter	3	$\|\eta\| < 1.35$	2	$1.375 < \|\eta\| < 1.5$
	2	$1.35 < \|\eta\| < 1.475$	3	$1.5 < \|\eta\| < 2.5$
			2	$2.5 < \|\eta\| < 3.2$
Granularity $\Delta\eta \times \Delta\phi$ versus $\|\eta\|$				
Presampler	0.025×0.1	$\|\eta\| < 1.52$	0.025×0.1	$1.5 < \|\eta\| < 1.8$
Calorimeter 1st layer	$0.025/8 \times 0.1$	$\|\eta\| < 1.40$	0.050×0.1	$1.375 < \|\eta\| < 1.425$
	0.025×0.025	$1.40 < \|\eta\| < 1.475$	0.025×0.1	$1.425 < \|\eta\| < 1.5$
			$0.025/8 \times 0.1$	$1.5 < \|\eta\| < 1.8$
			$0.025/6 \times 0.1$	$1.8 < \|\eta\| < 2.0$
			$0.025/4 \times 0.1$	$2.0 < \|\eta\| < 2.4$
			0.025×0.1	$2.4 < \|\eta\| < 2.5$
			0.1×0.1	$2.5 < \|\eta\| < 3.2$
Calorimeter 2nd layer	0.025×0.025	$\|\eta\| < 1.40$	0.050×0.025	$1.375 < \|\eta\| < 1.425$
	0.075×0.025	$1.40 < \|\eta\| < 1.475$	0.025×0.025	$1.425 < \|\eta\| < 2.5$
			0.1×0.1	$2.5 < \|\eta\| < 3.2$
Calorimeter 3rd layer	0.050×0.025	$\|\eta\| < 1.35$	0.050×0.025	$1.5 < \|\eta\| < 2.5$
Number of readout channels				
Presampler	7808		1536 (both sides)	
Calorimeter	101760		62208 (both sides)	
Scintillator tile calorimeter				
	Barrel		**Extended barrel**	
$\|\eta\|$ coverage	$\|\eta\| < 1.0$		$0.8 < \|\eta\| < 1.7$	
Number of layers	3		3	
Granularity $\Delta\eta \times \Delta\phi$	0.1×0.1		0.1×0.1	
Last layer	0.2×0.1		0.2×0.1	
Readout channels	5760		4092 (both sides)	
LAr hadronic end-cap calorimeter				
			End-cap	
$\|\eta\|$ coverage			$1.5 < \|\eta\| < 3.2$	
Number of layers			4	
Granularity $\Delta\eta \times \Delta\phi$			0.1×0.1	$1.5 < \|\eta\| < 2.5$
			0.2×0.2	$2.5 < \|\eta\| < 3.2$
Readout channels			5632 (both sides)	
LAr forward calorimeter				
			Forward	
$\|\eta\|$ coverage			$3.1 < \|\eta\| < 4.9$	
Number of layers			3	
Granularity $\Delta x \times \Delta y$ (cm)			FCal1: 3.0×2.6	$3.15 < \|\eta\| < 4.30$
			FCal1: $\sim$ four times finer	$3.10 < \|\eta\| < 3.15$
				$4.30 < \|\eta\| < 4.83$
			FCal2: 3.3×4.2	$3.24 < \|\eta\| < 4.50$
			FCal2: $\sim$ four times finer	$3.20 < \|\eta\| < 3.24$
				$4.50 < \|\eta\| < 4.81$
			FCal3: 5.4×4.7	$3.32 < \|\eta\| < 4.60$
			FCal3: $\sim$ four times finer	$3.29 < \|\eta\| < 3.32$
				$4.60 < \|\eta\| < 4.75$
Readout channels			3524 (both sides)	

Table 2.3: Main parameters of the calorimeter system.

segmented into high-granularity strips in the η direction, with a cell size of $0.025/8 \times 0.1$ in the $\eta - \phi$ plane. For $1.4 < |\eta| < 1.5$ and $1.8 < |\eta| < 2.5$, the η-segmentation of the first layer is coarser. The fine granularity in the η of the strips is sufficient to provide discrimination between single photon showers and two overlapping showers coming from the decays of neutral hadrons. The second layer has a thickness of about 17 X_0. In the ranges $|\eta| < 1.4$ and $1.425 < |\eta| < 2.5$ it is segmented in the η direction with a cell size of

0.025×0.025 in the $\eta - \phi$ plane. For $1.4 < |\eta| < 1.425$, the η-segmentation is coarser. It collects most of the energy deposited in the calorimeter by photon and electron showers. The third layer has a thickness of about 2 X_0 and a granularity of 0.050×0.025 in the $\eta - \phi$ plane. It is used to correct for leakage of high-energy showers beyond the EM calorimeter. For the end-cap inner wheel ($2.5 < |\eta| < 3.2$), the calorimeter is segmented in two sections in depth and has a coarser granularity (0.1×0.1) than for the rest of the acceptance. In the region of $|\eta| < 1.8$, a presampler detector is used to correct for the energy lost by electrons and photons upstream of the calorimeter. The presampler consists of an active LAr layer of thickness 1.1 cm (0.5 cm) in the barrel (end-cap) region with a granularity of 0.025×0.1 in the $\eta - \phi$ plane.

2.2.3.2 Hadronic calorimeters

Tile calorimeter

The Tile calorimeter is a sampling calorimeter using steel as the absorber and scintillating tiles as the active material. It is placed directly outside the EM calorimeter envelope. Its barrel covers the region $|\eta| < 1.0$, and its two extended barrels the range $0.8 < |\eta| < 1.7$. The barrel and extended barrels are divided azimuthally into 64 modules. Radially, the tile calorimeter extends from an inner radius of 2.28 m to an outer radius of 4.25 m. It is segmented in depth in three layers, approximately 1.5, 4.1 and 1.8 λ thick for the barrel and 1.5, 2.6 and 3.3 λ for the extended barrel. The total detector thickness at the outer edge of the tile-instrumented region is 9.7 λ at $\eta = 0$. Two sides of the scintillating tiles are read out by wavelength shifting fibres into two separate photomultiplier tubes. In η, the readout cells built by grouping fibres into the photomultipliers are pseudo-projective towards the interaction region.

LAr hadronic end-cap calorimeter

The HEC consists of two independent wheels per end-cap, located directly behind the EMEC and sharing the same LAr cryostats. To reduce the drop in material density at the transition between the end-cap and the forward calorimeters (around $|\eta| = 3.1$), the HEC extends out to $|\eta| = 3.2$, thereby overlapping with the FCal. Similarly, the HEC η range also slightly overlaps that of the tile calorimeter ($|\eta| < 1.7$) by extending to $|\eta| = 1.5$. Each wheel is built from 32 identical wedge-shaped modules assembled with fixtures at the periphery and at the central bore. Each wheel is divided into two segments in depth, for a total of four layers per end-cap. The wheels closest to the IP are built from 25 mm parallel copper plates, while those further away use 50 mm copper plates (for all wheels the first plate is half-thickness). The outer radius of the copper plates is 2.03 m, while the inner radius is 0.475 m (except in the overlap region with the FCal where the radius becomes 0.372 m). The copper plates are interleaved with 8.5 mm LAr gaps, providing the active medium for this sampling calorimeter.

LAr forward calorimeter

The FCal is integrated into the end-cap cryostats, as this provides clear benefits in terms of uniformity of the calorimetric coverage as well as reduced radiation background levels in the muon spectrometer. To reduce the amount of neutron albedo in the ID cavity, the front face of the FCal is recessed by about 1.2 m with respect to the EM calorimeter front face. This severely limits the depth of the calorimeter and therefore calls for a high-density design. The FCal covers the region $3.1 < |\eta| < 4.9$ and is approximately 10 λ deep, with three modules in each end-cap: the first, made of copper, is optimised for EM measurements, while the other two, made of tungsten, measure predominantly the energy of hadronic interactions. Each module consists of a metal matrix, with regularly spaced longitudinal channels filled with the electrode structure consisting of concentric rods and tubes parallel to the beam axis. The LAr in the gap between the rod and the tube is the sensitive medium. This geometry allows for excellent control of the gaps, which are as small as 0.25 mm in the first section to avoid problems due to ion buildup.

2.2.4 The muon spectrometer

The muon spectrometer forms the outer part of the ATLAS detector and is designed to detect charged particles exiting the barrel and end-cap calorimeters and to measure their momentum in the pseudorapidity range $|\eta| < 2.7$. It is also designed to trigger on these particles in the region $|\eta| < 2.4$. It is based on the magnetic deflection of muon tracks in the large superconducting air-core toroid magnets, instrumented with separate trigger and high-precision tracking chambers. Over the range $|\eta| < 1.4$, magnetic bending is provided by the large barrel toroid. For $1.6 < |\eta| < 2.7$, muon tracks are bent by two smaller end-cap magnets inserted into both ends of the barrel toroid. Over $1.4 < |\eta| < 1.6$, usually referred to as the transition region, magnetic deflection is provided by a combination of barrel and end-cap fields. This magnet configuration provides a field which is mostly orthogonal to the muon trajectories, while minimising the degradation of resolution due to multiple scattering. The anticipated high level of particle flux has had a major impact on the choice and design of the spectrometer instrumentation, affecting performance parameters such as rate capability, granularity, ageing properties, and radiation hardness.

The conceptual layout of the muon spectrometer is shown in figure 2.14, and the main parameters of the muon chambers are listed in table 2.4.

In the barrel region, tracks are measured in chambers arranged in three cylindrical layers around the beam axis; in the transition and end-cap regions, the chambers are installed in planes perpendicular to the beam, also in three layers.

2.2.4.1 The toroid magnets

A system of three large air-core toroids generates the magnetic field for the muon spectrometer. The two end-cap toroids are inserted in the barrel toroid at each end and line up with the central solenoid. Each of the three toroids consists of eight coils assembled

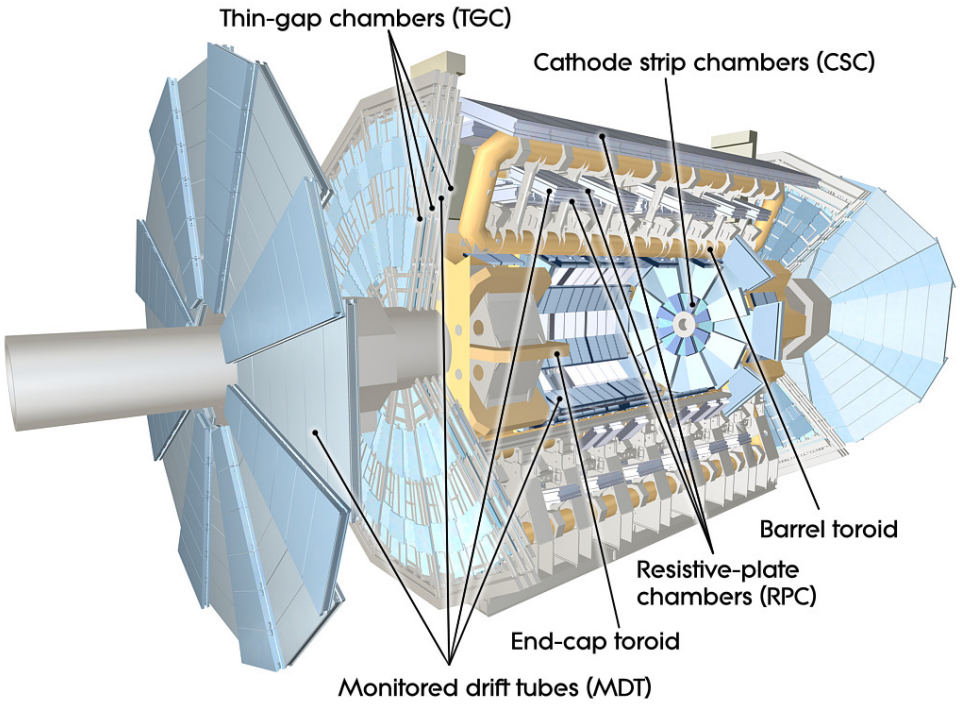


Figure 2.14: Cut-away view of the ATLAS muon system. From ref. [56].

radially and symmetrically around the beam axis.

The performance in terms of bending power is characterised by the field integral $\int Bdl$, where B is the field component normal to the muon direction and the integral is computed along an infinite-momentum muon trajectory, between the innermost and outermost muon-chamber planes. The barrel toroid provides 1.5 to 5.5 Tm of bending power in the pseudorapidity range $|\eta| < 1.4$, and the end-cap toroids approximately 1 to 7.5 Tm in the region $1.6 < |\eta| < 2.7$. The bending power is lower in the transition regions where the two magnets overlap ($1.4 < |\eta| < 1.6$).

2.2.4.2 Muon chamber types

Over most of the η-range, a precision measurement of the track coordinates in the principal bending direction of the magnetic field is provided by the Monitored Drift Tubes (MDT's). The mechanical isolation in the drift tubes of each sense wire from its neighbours guarantees a robust and reliable operation. At large pseudorapidities, Cathode Strip Chambers (CSC which are multiwire proportional chambers with cathodes segmented into strips) with higher granularity are used in the innermost plane over $2.0 < |\eta| < 2.7$, to withstand the demanding rate and background conditions. The stringent requirements on the relative alignment of the muon chamber layers are met by the combination of precision mechanical-assembly techniques and optical alignment systems both within and between

Monitored drift tubes	**MDT**
Coverage	$\|\eta\| < 2.7$ (innermost layer: $\|\eta\| < 2.0$)
Number of chambers	1088 (1150)
Number of channels	339000 (354000)
Function	Precision tracking
Cathode drift tubes	**CSC**
Coverage	$2.0 < \|\eta\| < 2.7$
Number of chambers	32
Number of channels	31000
Function	Precision tracking
Resistive plate chambers	**RPC**
Coverage	$\|\eta\| < 1.05$
Number of chambers	544 (606)
Number of channels	359000 (373000)
Function	Triggering, second coordinate
Thin gap chambers	**TGC**
Coverage	$1.05 < \|\eta\| < 2.7$ (2.4 for triggering)
Number of chambers	3588
Number of channels	318000
Function	Triggering, second coordinate

Table 2.4: Main parameters of the muon spectrometer. Numbers in brackets for the MDT and the RPC refer to the final configuration of the detector in 2009.

muon chambers.

The trigger system covers the pseudorapidity range $|\eta| < 2.4$. Resistive Plate Chambers (RPC) are used in the barrel and Thin Gap Chambers (TGC) in the end-cap regions. The trigger chambers for the muon spectrometer serve a threefold purpose: provide bunch-crossing identification, provide well-defined p_{T} thresholds, and measure the muon coordinate in the direction orthogonal to that determined by the precision-tracking chambers.

2.2.4.3 Muon chamber alignment and B-field reconstruction

The overall performance over the large areas involved, particularly at the highest momenta, depends on the alignment of the muon chambers with respect to each other and with respect to the overall detector.

The accuracy of the stand-alone muon momentum measurement necessitates a precision of 30 μm on the relative alignment of chambers both within each projective tower and between consecutive layers in immediately adjacent towers. The internal deformations and relative positions of the MDT chambers are monitored by approximately 12000

precision-mounted alignment sensors, all based on the optical monitoring of deviations from straight lines. Because of geometrical constraints, the reconstruction and/or monitoring of the chamber positions rely on somewhat different strategies and sensor types in the end-cap and barrel regions, respectively. The accuracy required for the relative positioning of non-adjacent towers to obtain adequate mass resolution for multi-muon final states, lies in the few millimetre range. This initial positioning accuracy is approximately established during the installation of the chambers.

For magnetic field reconstruction, the goal is to determine the bending power along the muon trajectory to a few parts in a thousand. The field is continuously monitored by a total of approximately 1800 Hall sensors distributed throughout the spectrometer volume. Their readings are compared with magnetic-field simulations and used for reconstructing the position of the toroid coils in space, as well as to account for magnetic perturbations induced by the tile calorimeter and other nearby metallic structures.

2.2.5 Forward detectors

In addition to the main ATLAS detector systems, three smaller sets of detectors were built to provide good coverage in the very forward region [59]. The main function of two of the systems is to determine the luminosity delivered to ATLAS.

Ordered according to their distance from the IP, as shown in figure 2.15, the first system is a Cerenkov detector called LUCID (LUminosity measurement using Cerenkov Integrating Detector). LUCID detects inelastic pp scatterings in the forward direction. It is the main relative luminosity monitor in ATLAS, and is located at a distance of ± 17 m from the IP, near the TAS (Target Absorber Secondaries) collimator. The second system is the ZCD (Zero-Degree Calorimeter), located at a distance of ± 140 m from the IP. This corresponds to the location where the LHC beam-pipe is divided into two separate pipes. The ZDC is embedded in the TAN (Target Absorber Neutral), located between the beam-pipes just after the split. The ZDC primary purpose is to detect forward neutrons in heavy-ion collisions and its modules consist of layers of alternating quartz rods and tungsten plates which measure neutral particles at $|\eta| > 8.2$. The most remote system is the absolute luminosity detector ALFA (Absolute Luminosity For ATLAS). ALFA consists of scintillating-fibre trackers located inside Roman pots at a distance of approximately ± 240 m from the IP which are designed to approach as close as 1 mm to the beam.

2.2.6 Trigger and data-acquisition system

The trigger system is an essential component of any collider experiment as it is responsible for deciding whether to keep an event from a given bunch-crossing interaction for later study. During the Run I of the LHC, the trigger system of the ATLAS experiment operated efficiently at instantaneous luminosities of up to 8×10^{33} cm^{2}s^{-1} and primarily

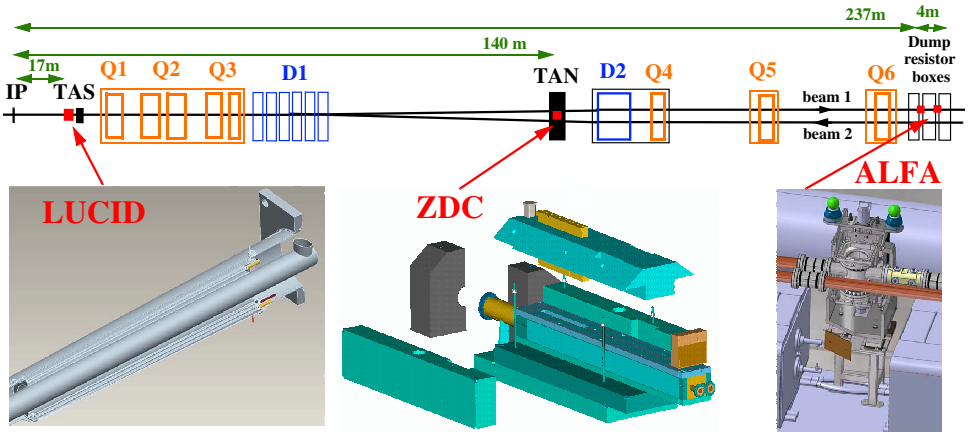


Figure 2.15: Placement of the forward detectors along the beam-line around the ATLAS IP. From ref. [56].

at $\sqrt{s}$ of 7 TeV and 8 TeV. In Run II, the increased $\sqrt{s}$ of 13 TeV, higher luminosity and increased pile-up meant that, without upgrades of the trigger system, the trigger rates would have exceeded the maximum allowed rates when running with the trigger thresholds needed to satisfy the physics programme of the experiment. For this reason, the LS1 between LHC Run I and Run II operations was used to improve the trigger system with almost no component left untouched [38, 60].

The Trigger and Data Acquisition (TDAQ) system shown in figure 2.16 consists of a hardware-based first-level trigger (L1) and a software-based High-Level Trigger (HLT). The L1 trigger decision is formed by the Central Trigger Processor (CTP), which receives inputs from the L1 Calorimeter (L1Calo) and L1 Muon (L1Muon) triggers as well as several other subsystems such as the Minimum Bias Trigger Scintillators (MBTS), the LUCID Cerenkov counter and the ZDC. The CTP is also responsible for applying preventive dead-time. It limits the minimum time between two consecutive L1 accepts to avoid overlapping readout windows, and restricts the number of L1 accepts allowed in a given number of bunch-crossings to avoid front-end buffers from overflowing.

The L1 trigger, implemented in custom hardware, utilises coarser-granularity signals from the calorimeter and the muon chambers to reduce the event rate from the 40 MHz bunch crossing rate to below 100 kHz; it has 2.5 μs to decide which events to keep to satisfy this factor 400 reduction. L1 also defines Regions-of-Interest (RoIs) which have calorimeter clusters with high transverse energy or muon tracks in the muon chambers.

Events accepted by L1 are processed by the HLT, based on algorithms implemented in software which must further reduce the number of events recorded to disk to an average of 1 kHz within a few seconds. The HLT uses fine-granularity calorimeter information, precision measurements from the muon spectrometer and tracking information from the ID, which are not available at L1. HLT reconstruction can be executed either within the RoIs identified at L1 or for the full detector (full-scan). The selection of particle

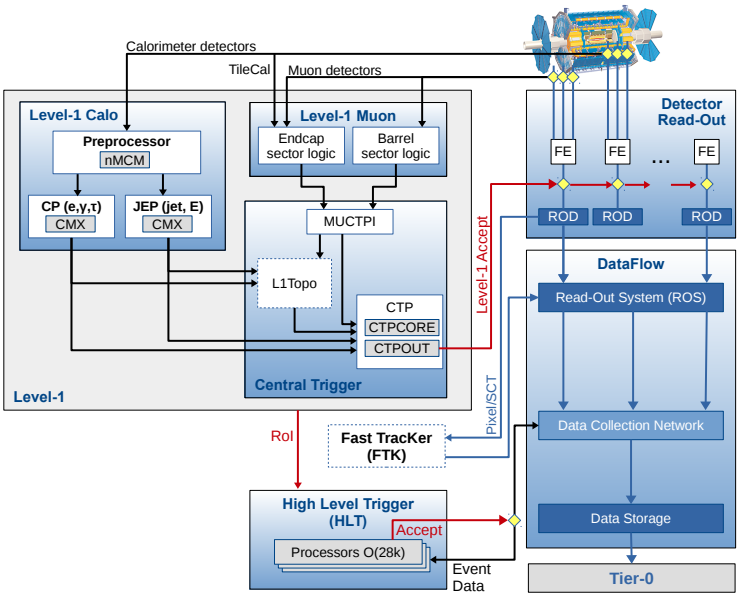


Figure 2.16: The ATLAS TDAQ system in Run II with emphasis on the components relevant for triggering. From ref. [38].

candidates by the HLT is performed at each step, so that if it fails at a certain step, subsequent steps are not executed. This is essential to reduce the time needed by the HLT to reconstruct the event and make a decision.

A sequence of L1 and HLT trigger algorithms is called a *trigger* and is meant to identify one or more particles of a given type and a given threshold of transverse energy or momentum. For example, electron and photon triggers select events with one or more electrons or photons in the detector. The configuration of the trigger is controlled by the *trigger menu*, which defines a full list of the L1 and HLT triggers and their configurations. Menu composition and trigger thresholds are optimised for the LHC running conditions (beam type, luminosity, etc.) to fit within the event acceptance rate and the bandwidth constraints of the data acquisition system of the ATLAS detector as well as the offline storage constraints.

In addition to the triggers described above, there are *rerun* triggers which never accept an event on their own, but are configured to run only on the events accepted by other triggers, and their decision is recorded for offline use. This information is used for studies of the trigger efficiency, which is calculated separately for each object (leg) of the multi-object triggers.

Trigger thresholds and identification criteria have to be modified sometimes to maintain a stable output rate. To ensure an optimal trigger menu within the rate constraints of a given LHC luminosity, prescale factors can be applied to both the L1 and HLT triggers independently and configured during data-taking. They allow the experiment to either disable triggers completely or to set the fraction of events that may be accepted.

Chapter 3

Theoretical framework

In this chapter, the most relevant aspects of the theoretical framework underlying the work presented in this thesis are reviewed. The structure and content of the more fundamental sections involving the SM and QCD are to a great extent based on ref. [61]. A brief description of the jet algorithms is provided using ref. [62, 63], with the focus on those algorithms used for the results presented in this thesis. Finally, an overview of the theory of prompt-photon production is given based on the most recent studies available on the literature, from both the experimental and theoretical sides.

3.1 The Standard Model

The Standard Model of particle physics is the mathematical theory that describes all the known elementary particles and their interactions. Developed in the early 1970s, it has successfully explained almost all experimental results and precisely predicted a wide variety of phenomena. Over time and through many experiments, the SM has become established as a well-tested physics theory.

There are four fundamental forces in nature: the strong, the weak, the electromagnetic, and the gravitational force. They work over different ranges and have different strengths. Gravity is the weakest but it has an infinite range. The electromagnetic force has infinite range but it is many times stronger than gravity. The weak and strong forces are effective only over a very short range and dominate only at the level of subatomic particles. The weak force is much stronger than gravity, but it is the weakest of the other three. Three of these fundamental forces result from the exchange of mediator particles (bosons). Particles of matter transfer discrete amounts of energy by exchanging bosons with each other. The SM includes the electromagnetic, weak and strong forces and all their bosons, and explains how these forces act on all matter particles.

All matter is made out of fermions, which are particles of half-integer spin. The interactions among fermions are interpreted via the exchange of the aforementioned bosons,

which are particles of integer spin. Fermions are subdivided into two kinds of elementary particles, leptons and quarks. There are six leptons, classified according to their charge (Q), electron number (L_e), muon number (L_μ), and tau number (L_τ). They fall naturally into the three families (or generations) listed in table 3.1.

Lepton	Q	L_e	L_μ	L_τ
First generation				
e	-1	1	0	0
ν_e	0	1	0	0
Second generation				
μ	-1	0	1	0
ν_μ	0	0	1	0
Third generation				
τ	-1	0	0	1
ν_τ	0	0	0	1

Table 3.1: Lepton classification.

There are also six antileptons, with all the signs reversed, increasing the total number of leptons in the SM up to 12. The electron, the muon and the tau (and their antiparticles) all have an electric charge and a sizeable mass, whereas the neutrinos are electrically neutral and have very small mass. Since neutrinos are neutral, the question whether each neutrino is its own antiparticle (Majorana particles) or not (Dirac particles) remains unanswered.

Similarly, there are six *flavours* of quarks, which are classified according to charge, strangeness (S), charm (C), beauty (B), and truth (T). The quarks also fall into three generations that are shown in table 3.2. In the same way as for leptons, all signs would be reversed for the corresponding antiquarks. Quarks also come in three different *colours*, giving 36 of them in all, and only mix in such a way as to form colourless objects.

Quark	Q	S	C	B	T
First generation					
d	$-1/3$	0	0	0	0
u	$2/3$	0	0	0	0
Second generation					
s	$-1/3$	-1	0	0	0
c	$2/3$	0	1	0	0
Third generation					
b	$-1/3$	0	0	-1	0
t	$2/3$	0	0	0	1

Table 3.2: Quark classification.

The lightest and most stable particles make up the first generation, whereas the heavier

and less-stable particles belong to the second and third generations. All stable matter in the universe is made from particles that belong to the first generation; any heavier particles quickly decay to more stable ones.

Each of the three forces unified in the SM is described by a Quantum Field Theory (QFT), corresponding to the exchange of a spin-1 mediator, known as a gauge boson. The strong interaction is responsible for binding the quarks and it is mediated by the gluon gauge bosons, which are massless, electrically neutral but carry colour charge. The massless photons, mediate the electromagnetic interaction between charged particles. Finally, the gauge bosons mediating the weak force are the charged $W^{\pm}$ and the neutral Z bosons. These last two bosons carry mass and make the weak interaction short-ranged. On the other hand, the photon is massless, and, as a consequence, the interaction has infinite range. An exception occurs with the gluons and the strong force, despite being massless, *colour confinement* and the *asymptotic freedom* of the theory explain why the interaction becomes strong at low energies, leading to the confinement of quarks and gluons within composite hadrons.

In principle, the particles included in the SM are massless, unlike the particles observed in nature. The basic equations of the unified theory of the weak force and electromagnetism describe correctly the force-carrying particles, namely the photon, and the $W^{\pm}$ and Z bosons, except for a major problem. All these particles emerge without a mass. A mechanism called ElectroWeak Symmetry Breaking (EWSB) was introduced with which the $W^{\pm}$ and Z bosons could acquire mass. In the SM, the Higgs scalar field induces a spontaneous breaking of the electroweak symmetry when it acquires a non-vanishing vacuum expectation value. This mechanism is commonly referred to as the Brout-Englert-Higgs mechanism [64–67]. The Higgs field provides mass to fermions of the SM via Yukawa interactions. The Brout-Englert-Higgs mechanism postulates the existence of one scalar particle, the Higgs boson, which was discovered during 2012 by ATLAS and CMS collaborations at the LHC with a mass around 125 GeV [2, 3].

A schema of the particles that are included in the SM and their most important properties is shown in figure 3.1.

3.1.1 Mathematical formulation of the Standard Model

In the SM all of the fundamental interactions derive from a single general principle, the requirement of local gauge invariance. The symmetry group of the SM is

$$\mathrm{SU(3)_C} \times \mathrm{SU(2)_L} \times \mathrm{U(1)_{Y_W}} \,.$$

The group $\mathrm{SU(3)_C}$ describes the strong interactions, also known as Quantum Chromodynamics, and the index C refers to the colour charge, which is conserved under the SU(3) symmetry. The conservation of colour charge implies the existence of eight gauge bosons that are gluons. Under $\mathrm{SU(3)_C}$, quarks are colour triplets while leptons are colour

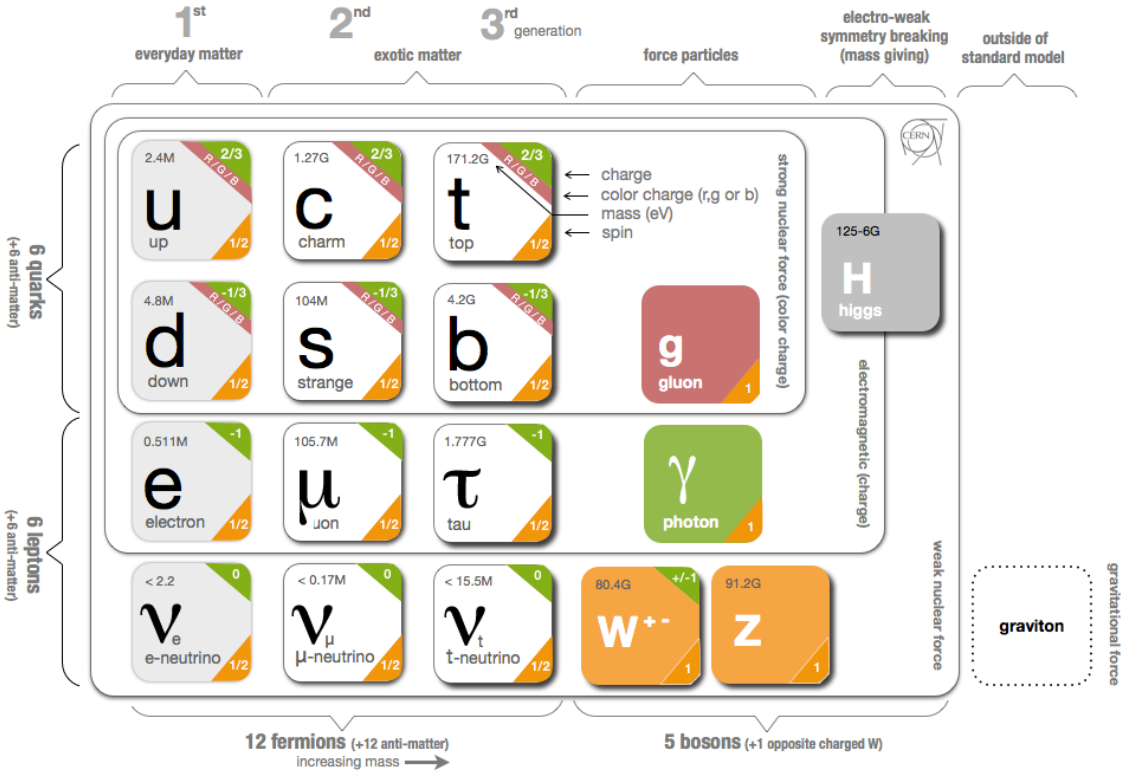


Figure 3.1: The particle content of the SM. All fermions participate in the weak interaction, but only quarks feel the strong force. Both quarks and charged leptons interact via the electromagnetic force. Neutrinos, being neutral and colourless, only interact via the weak force. Although it has not been discovered yet, the graviton should be the corresponding force-carrying particle of gravity. From ref. [68].

singlets; therefore, quarks carry a colour index ranging between one and three, whereas leptons do not take part in strong interactions.

$\mathrm{SU(2)_L} \times \mathrm{U(1)_{Y_W}}$ describes the *non-Abelian* gauge group of the electroweak interactions, described by the Glashow-Weinberg-Salam (GWS) model [69–71]. The index L refers to the left chiral nature of the SU(2) coupling and Y_W to the weak hypercharge that is carried by the gauge field B_μ and is related to the electric charge Q and the weak isospin T_3 by

$$Y_\mathrm{W} = 2\,(Q - T_3)\,.$$

The gauge bosons of the group $\mathrm{SU(2)_L} \times \mathrm{U(1)_{Y_W}}$ are γ, Z, and $W^\pm$. The group symmetry of the electromagnetic interactions, $\mathrm{U(1)_{EM}}$ appears in the SM as a subgroup of $\mathrm{SU(2)_L} \times$

$U(1)_{Y_W}$ and it is in this sense that the weak and electromagnetic interactions are said to be unified.

The SM Lagrangian can be separated into two terms; the first one describing the electroweak interactions and the second accounting for the QCD interactions

$$\mathcal{L}_{SM} = \mathcal{L}_{EW} + \mathcal{L}_{QCD} .$$

3.2 Quantum Chromodynamics

One of the most fundamental principles of QCD is that hadronic matter is made of quarks. The idea of quarks arose from the need to have a physical manifestation for the group SU(3) of flavour ($SU(3)_F$) observed in the spectrum of the lowest-mass mesons and baryons. The properties of the six known quarks are shown in table 3.3. The observed baryons are interpreted as three-quark states. The quark constituents of the baryons are forced to have half-integral spin to account for the spins of the low-mass baryons. Quarks in the spin-$\frac{3}{2}$ baryons are then in a symmetrical state of space, spin and $SU(3)_F$ degrees of freedom. However, the requirements of Fermi-Dirac statistics imply the total antisymmetry of the wave function. The solution to this problem was the introduction of the colour degree of freedom: a colour index a with three possible values ($a = 1, 2, 3$) carried by each quark. The baryon wave functions are totally antisymmetric in this new index.

Quark	Mass	Charge (Q)	Weak isospin T_3	Weak hypercharge (Y)
d	$4.67^{+0.48}_{-0.17}$ MeV	$-1/3$	$-1/2$	$1/3$
u	$2.16^{+0.49}_{-0.26}$ MeV	$2/3$	$+1/2$	$1/3$
s	93^{+11}_{-5} MeV	$-1/3$	$-1/2$	$1/3$
c	$1.27^{+0.02}_{-0.02}$ GeV	$2/3$	$+1/2$	$1/3$
b	$4.18^{+0.03}_{-0.02}$ GeV	$-1/3$	$-1/2$	$1/3$
t	$172.76^{+0.30}_{-0.30}$ GeV	$2/3$	$+1/2$	$1/3$

Table 3.3: The properties of the six quarks. The mass values included correspond to the Particle Data Group (PDG) averages reported in ref. [72]. The weak isospin describes how the quarks transform under $SU(2)_L$. The weak hypercharge describes the transformations under $U(1)_{Y_W}$. Under the unitary irreducible representation of $SU(2)_L \times U(1)_{Y_W}$, the electric charge describes quarks transformations.

The colour hypothesis has to be supplemented by a requirement that only colour singlet states can exist in nature. If the colour group of transformations is SU(3), the

quarks q_{a} transform according to the fundamental representation, and the antiquarks $\overline{q}^{\mathrm{a}}$, according to the complex conjugate representation, then, the basic colour singlet states are precisely mesons $q_{\mathrm{a}}\overline{q}^{\mathrm{a}}$ and the baryons $\epsilon^{\mathrm{abc}}q_{\mathrm{a}}q_{\mathrm{b}}q_{\mathrm{c}}$, where ϵ^{abc} is the totally antisymmetric tensor.

The Feynman rules required for a perturbative analysis of QCD can be derived from a Lagrangian[1] given by

$$\mathcal{L}_{\mathrm{QCD}} = -\frac{1}{4}F^{\mathrm{A}}_{\alpha\beta}F^{\alpha\beta}_{\mathrm{A}} + \sum_{\text{flavours}} \overline{q}_{\mathrm{a}}\left(i\gamma_\mu D^\mu - m\right)_{\mathrm{ab}} q_{\mathrm{b}}\,. \tag{3.1}$$

These terms describe the interaction of spin-$\frac{1}{2}$ quarks of mass m and massless spin-1 gluons. To simplify the equations, the spinor indices of γ^μ and q_{a} are suppressed. The metric is given by $g^{\alpha\beta} = \mathrm{diag}(1,-1,-1,-1)$, $\hbar = c = 1$ and the γ matrices satisfy the anticommutation relations $\{\gamma^\mu,\gamma^\nu\} = 2g^{\mu\nu}$.

The field strength tensor $F^{\mathrm{A}}_{\alpha\beta}$ is derived from the gluon field A^{A}_{α}

$$F^{\mathrm{A}}_{\alpha\beta} = \left[\partial_\alpha A^{\mathrm{A}}_\beta - \partial_\beta A^{\mathrm{A}}_\alpha - g_s f^{\mathrm{ABC}}A^{\mathrm{B}}_\alpha A^{\mathrm{C}}_\beta\right]\,, \tag{3.2}$$

where the indices A, B and C run over the eight colour degrees of freedom of the gluon field. The third non-Abelian term on the right hand side of equation 3.2 gives rise to triplet and quartic gluon self-interactions and ultimately to the property of asymptotic freedom.

The explicit sum in equation 3.1 runs over the n_f different quark flavours, g_{S} in equation 3.2 is the strong coupling constant, and f^{ABC} are the structure constants of $\mathrm{SU}(3)_{\mathrm{C}}$. The quark fields q_{a} are in the triplet representation of the colour group ($a = 1,2,3$), and D^μ is the covariant derivative, which acts on triplet and octet fields and takes the form

$$(D_\alpha)_{\mathrm{ab}} = \partial_\alpha\delta_{\mathrm{ab}} + ig_{\mathrm{S}}(t^{\mathrm{C}}A^{\mathrm{C}}_\alpha)_{\mathrm{ab}},\;\; (D_\alpha)_{\mathrm{AB}} = \partial_\alpha\delta_{\mathrm{AB}} + ig_{\mathrm{S}}(T^{\mathrm{C}}A^{\mathrm{C}}_\alpha)_{\mathrm{AB}}\,,$$

where t and T are matrices in the fundamental and adjoint representations of SU(3), respectively.

3.2.1 The Renormalisation Group Equations

The evaluation of Feynman diagrams that contain loops can lead to divergent integrals over loop momenta. A regularisation procedure is adopted to make finite the divergent expressions. This procedure introduces additional parameters and makes possible to re-express the divergences of perturbation theory in a well-defined way. The regularised

[1]The description provided of QCD omits the so-called *Gauge-fixing* and *Faddeev-Popov* terms introduced to properly quantise the QCD Lagrangian.

divergences of perturbation theory are then removed by absorbing them into definitions of physical quantities through a renormalisation procedure. This renormalisation is done by some specified (but arbitrary) prescription that introduces a new dimensional scale μ_{R}. Renormalised quantities in the theory, such as the basic vertex coupling strength g_{S}, depend explicitly on μ_{R}. Different renormalisation prescriptions with different μ_{R} must all lead to the same observable amplitudes. The transformations of renormalised operators as μ_{R} is changed form a Lie group, first recognised by Stueckelberg and Peterman, and named by them the *Renormalisation Group*. The equations that express the invariance of the physics under changes of the parameter μ_{R} are known as the Renormalisation Group Equations (RGE).

3.2.2 The running coupling constant

In order to introduce the concept of running coupling, a dimensionless physical observable R which depends on a single energy scale Q is considered as an example. The scale Q is assumed to be much larger than all other dimensionful parameters, such as quark masses, that are therefore set to zero.

When R is calculated as a perturbation series in the coupling $\alpha_{\mathrm{S}} = g_{\mathrm{S}}/4\pi$ (defined in analogy with the fine structure constant of Quantum Electrodynamics (QED)), the perturbation series requires renormalisation to remove *ultraviolet* (UV) divergences. Since the renormalisation procedure introduces a second mass scale μ_{R} (the point at which the subtraction which removes the UV divergences is performed), R depends in general on the ratio Q^2/μ_{R}^2 and the renormalised coupling α_{S} depends on the choice made for the subtraction point.

The Lagrangian of QCD makes no mention of the scale μ_{R} required to define the theory at the quantum level. Since R is dimensionless, it can only depend on the ratio Q^2/μ_{R}^2 and the renormalised coupling α_{S}. The μ_{R} independence of R may be expressed by

$$\mu_{\mathrm{R}}^2 \frac{d}{d\mu_{\mathrm{R}}^2} R\left(Q^2/\mu_{\mathrm{R}}^2, \alpha_{\mathrm{S}}\right) \equiv \left[\mu_{\mathrm{R}}^2 \frac{\partial}{\partial \mu_{\mathrm{R}}^2} + \mu_{\mathrm{R}}^2 \frac{\partial \alpha_{\mathrm{S}}}{\partial \mu_{\mathrm{R}}^2} \frac{\partial}{\partial \alpha_{\mathrm{S}}}\right] R = 0\,.$$

This equation can be rewritten in a more compact form introducing $t = \ln\left(\frac{Q^2}{\mu_{\mathrm{R}}^2}\right)$ and $\beta\left(\alpha_{\mathrm{S}}\right) = \mu_{\mathrm{R}}^2 \frac{\partial \alpha_{\mathrm{S}}}{\partial \mu_{\mathrm{R}}^2}$,

$$\left[-\frac{\partial}{\partial t} + \beta\left(\alpha_{\mathrm{S}}\right) \frac{\partial}{\partial \alpha_{\mathrm{S}}}\right] R\left(e^{\mathrm{t}}, \alpha_{\mathrm{S}}\right) = 0\,. \tag{3.3}$$

This first order partial differential equation is solved by implicitly defining a new function, the *running coupling* $\alpha_{\mathrm{S}}\left(Q^2\right)$,

$$t = \int_{\alpha_S}^{\alpha_S(Q^2)} \frac{\mathrm{d}x}{\beta(x)}, \quad \alpha_S\left(\mu_R^2\right) \equiv \alpha_S \,. \tag{3.4}$$

By differentiating equation 3.4,

$$\frac{\partial \alpha_S(Q^2)}{\partial t} = \beta\left(\alpha_S\left(Q^2\right)\right), \quad \frac{\partial \alpha_S(Q^2)}{\partial \alpha_S} = \frac{\beta\left(\alpha_S(Q^2)\right)}{\beta(\alpha_S)},$$

so that $R\left(1, \alpha_S\left(Q^2\right)\right)$ is a solution of equation 3.3. All the scale dependence in R enters through the running of the coupling constant $\alpha_S\left(Q^2\right)$. The knowledge of the quantity $R\left(1, \alpha_S\left(Q^2\right)\right)$, calculated in fixed-order perturbative theory allows to predict the variation of R with Q if equation 3.4 is solved. QCD is an asymptotically-free theory, this is discussed in section 3.2.4. As a consequence, $\alpha_S\left(Q^2\right)$ becomes smaller as the scale Q increases, and for sufficiently large Q, equation 3.4 can always be solved using perturbation theory.

3.2.3 The β function

The running of the coupling constant α_S is determined by the renormalisation group equation

$$Q^2 \frac{\partial \alpha_S}{\partial Q^2} = \beta(\alpha_S) \,. \tag{3.5}$$

In QCD, the β function has the perturbative expansion

$$\beta(\alpha_S) = -b\alpha_S^2\left(1 + b'\alpha_S + b''\alpha_S^2 + O\left(\alpha_S^3\right)\right), \tag{3.6}$$

where

$$b = \frac{(33 - 2n_f)}{12\pi},$$

$$b' = \frac{(153 - 19n_f)}{2\pi(33 - 2n_f)},$$

$$b'' = \frac{(77139 - 15099n_f + 325n_f^2)}{288\pi^2(33 - 2n_f)},$$

and n_f is the number of quark flavours contributing in the loops[2]. The β function coefficients are extracted from the higher-order corrections to the *bare* vertices of the theory.

[2]The three-loop coefficient b'' depends on the scheme used for the renormalisation of the theory. The result given here is for the $\overline{MS}$ scheme.

As a consequence of the non-Abelian character of QCD interactions, the b coefficients in QCD and QED have the opposite sign for $n_f < 16$.

From equations 3.5 and 3.6,

$$Q^2 \frac{\partial \alpha_S(Q^2)}{\partial Q^2} = -b\alpha_S^2(Q^2)\left[1 + b'\alpha_S(Q^2) + O\left(\alpha_S^2(Q^2)\right)\right], \tag{3.7}$$

if both $\alpha_S(\mu_R^2)$ and $\alpha_S(Q^2)$ are in the perturbative region, the right-hand-side of equation 3.7 can be truncated neglecting the b' and higher coefficients. The solution is

$$\alpha_S(Q^2) = \frac{\alpha_S(\mu_R^2)}{1 + \alpha_S(\mu_R^2)\, bt}, \tag{3.8}$$

which gives the relation between $\alpha_S(Q^2)$ and $\alpha_S(\mu_R^2)$, being both in the perturbative region. When t becomes very large, the running coupling $\alpha_S(Q^2)$ decreases to zero. This property is called asymptotic freedom. The approach to zero is rather slow, since α_S decreases like an inverse power of $\ln(Q^2)$. The sign of b is crucial. With the opposite sign of b, the coupling would increase with Q, as it does in QED.

3.2.4 Asymptotic freedom and confinement

Confinement explains why the quark and gluon degrees of freedom are not observed as states which propagate over macroscopic distances. Asymptotic freedom explains why the methods of perturbation theory are useful at high energy. These are two limiting cases of the theory, that appear at low and high energies, respectively.

The coupling constant is said to run, being large at low energy and becoming smaller at high energy, where t becomes very small and large in equation 3.8, respectively. This behaviour is illustrated in figure 3.2.

When the coupling constant increases, in the low-energy regime, QCD is strongly interacting, leading to the confinement of quarks and gluons. Confinement implies that neither quarks nor gluons can appear in isolation, they can only exist within colourless composite hadrons. On the other hand, at high energies, QCD interacts weakly, allowing perturbation theory calculations at large Q. The world average value of $\alpha_S(M_Z^2)$ published in the latest PDG review [72] is $\alpha_S(M_Z^2) = 0.1179 \pm 0.0010$, as a result of a combination of a wide range of measurements from different reactions.

Perturbative QCD (pQCD) tells us how the coupling constant varies with the scale, not the absolute value itself, which must be obtained from experiment. The value of the coupling constant can be chosen at a convenient reference scale, as M_Z, which is large enough to be in the perturbative regime. Setting $\mu_R = M_Z$, the value of the coupling can be deduced at any other large scale Q.

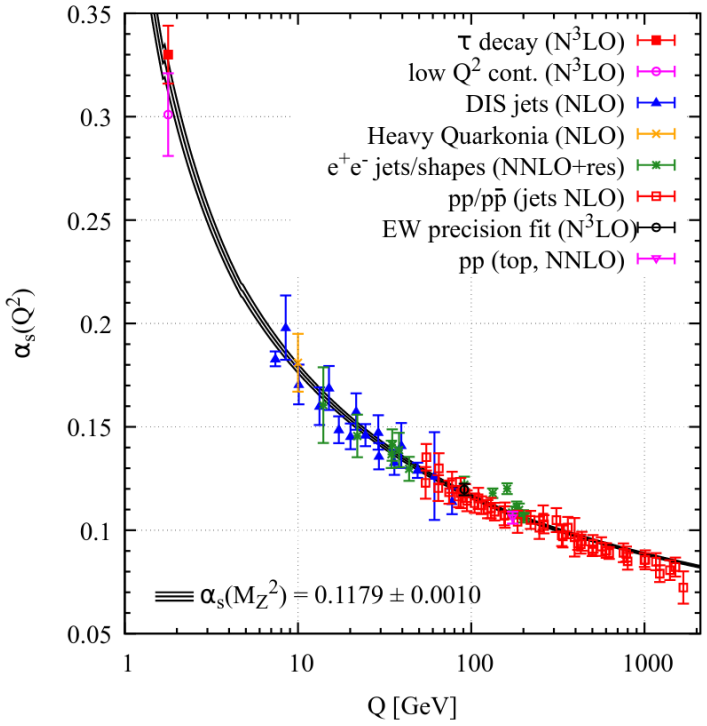


Figure 3.2: Summary of measurements of α_S as a function of the energy scale Q. From ref. [72].

An alternative approach, adopted historically is to introduce a dimensionful parameter Λ directly into the definition of $\alpha_S(Q^2)$, called the Landau Pole. The Λ parameter represents the scale at which the coupling would diverge, if extrapolated outside the perturbative domain. It indicates the order of magnitude of the scale at which $\alpha_S(Q^2)$ becomes strong. The value of Λ depends on the renormalisation scheme used, and is ~ 200 MeV for the $\overline{\text{MS}}$ scheme. The running coupling $\alpha_S(Q^2)$ becomes large for scales comparable with the masses of light hadrons, which could be an indication that the confinement of quarks and gluons inside hadrons is actually a consequence of the growth of the coupling at low scales.

In terms of Λ, the asymptotic solution for α_S can be written in the leading order (LO) approximation as

$$\alpha_S\left(Q^2\right) = \frac{1}{b \ln\left(Q^2/\Lambda^2\right)}.$$

3.2.5 Parton model, structure and parton distribution functions

The parton model was devised before the development of QCD. The basic assumption of the model is that the interactions of hadrons are due to interactions of partons; in certain

types of reactions, the structure of the hadrons may be described by an instantaneous distribution of partons.

High-energy lepton-nucleon scattering plays a key role in determining the partonic structure of the proton. The process $lN \to l'X$ is illustrated in figure 3.3. The circle represents the internal structure of the nucleon, which can be expressed in terms of structure functions. The quantities k^μ and k'^μ are the four-momenta of the incoming and outgoing leptons, respectively. The momentum of the target nucleon (assumed hereafter to be a proton) is represented by p^μ. The exchanged particle is a γ, $W^\pm$, or Z; it transfers four-momentum $q^\mu = k^\mu - k'^\mu$ to the proton. Then, the standard deep inelastic variables are defined by

$$
\begin{aligned}
Q^2 &= -q^2 \,, \\
M^2 &= p^2 \,, \\
\nu &= p \cdot q = M\left(E' - E\right) \,, \\
x &= \frac{Q^2}{2\nu} = \frac{Q^2}{2M\left(E - E'\right)} \,, \\
y &= \frac{q \cdot p}{k \cdot p} = 1 - E'/E \,,
\end{aligned}
$$

where the energy variables refer to the target rest frame and M is the proton mass.

The cross section for this process, considering a simple model where a photon scatters off a point-like quark that carries a fraction ϵ of its momentum, i.e. $p_{\mathrm{q}}^\mu = \epsilon p^\mu$, can be written as[3]

$$
\frac{d^2\sigma}{dx\, dQ^2} = \frac{4\pi\alpha^2}{Q^4}\left[\left[1 + (1-y^2)\right] F_1 + \frac{(1-y)}{x}\left(F_2 - 2xF_1\right)\right] \,, \tag{3.9}
$$

where $F_{\mathrm{i}}\left(x, Q^2\right)$ are structure functions that parameterise the structure of the hadron. The *Bjorken limit* is defined as $Q^2 \to \infty$ with x fixed. In this limit, the structure functions are observed to obey a scaling law, depending only on the variable x ($F_{\mathrm{i}}\left(x, Q^2\right) \to F_{\mathrm{i}}\left(x\right)$). Bjorken scaling implies that the virtual photon scatters off pointlike constituents, otherwise, the structure functions would depend on the ratio Q/Q_0, with $1/Q_0$ a length characterizing the size of the constituents.

Comparing equation 3.9 with the double differential cross sections for the quark scattering process $eq \to eq$, the mass-shell constraint for the outgoing quark implies $x = \epsilon$, and the *Callan-Gross relation* between the structure functions is obtained[4]

[3] The equation is written in the *infinite momentum frame*, in which the hadron is moving very fast, $p^\mu \approx (P, 0, 0, P)$, with $P \gg M$.

[4] This result is a direct consequence of the spin-$\frac{1}{2}$ property of quarks.

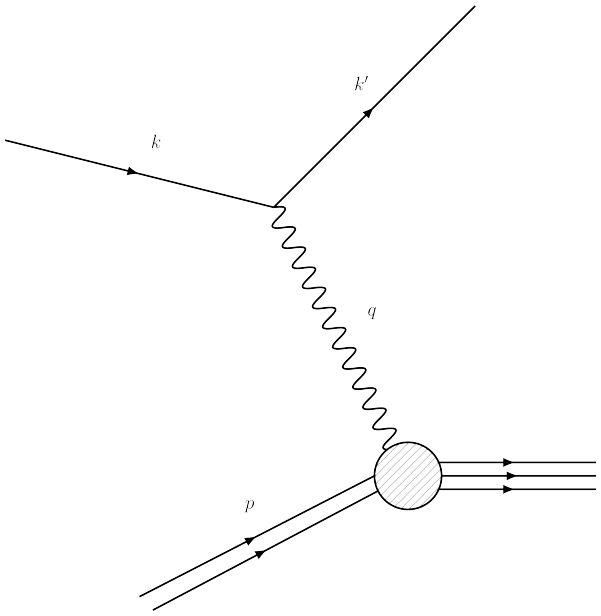


Figure 3.3: Deep inelastic charged lepton-proton scattering.

$$F_2 = xe_{\mathrm{q}}^2 \delta\left(x-\epsilon\right) = 2xF_1\,,$$

which suggests that the structure function $F_2\left(x\right)$ probes a quark constituent with momentum fraction $\epsilon = x$. The structure functions depend on x, suggesting that the quark constituents carry a range of momentum fractions. $q\left(\epsilon\right)d\epsilon$ represents the probability that a quark q carries a momentum fraction between ϵ and $\epsilon + d\epsilon$, where $0 < \epsilon < 1$. The virtual photon scatters incoherently off the quark constituents. The proton structure functions are obtained by weighting the quark structure functions with the probability distribution $q\left(\epsilon\right)$,

$$F_2\left(x\right) = 2xF_1\left(x\right) = \sum_{q,\overline{q}} \int_0^1 \mathrm{d}\epsilon\, q\left(\epsilon\right) xe_{\mathrm{q}}^2 \delta\left(x-\epsilon\right) = \sum_{q,\overline{q}} e_{\mathrm{q}}^2 xq\left(x\right)\,,$$

and these relations can be inverted to give the individual quark distributions in terms of the measured structure functions. The proton consists of three valence quarks (uud) that carry its electric charge and baryon quantum numbers and an infinite *sea* of light $q\overline{q}$ pairs. When the proton is probed at a momentum scale Q, the sea contains all quark flavours with mass $m_{\mathrm{q}} \ll Q$. At a scale of $O\left(1\text{ GeV}\right)$, assuming the sea is symmetric in the quark flavours,

$$u\left(x\right) = u_{\mathrm{V}}\left(x\right) + S\left(x\right)\,,$$

$$d\left(x\right) = d_{\mathrm{V}}\left(x\right) + S\left(x\right)\,,$$

$$S(x) = \overline{u}(x) = \overline{d}(x) = \overline{s}(x) = s(x) \ ,$$

and the sum rules

$$\int_0^1 \mathrm{d}x \, u_{\mathrm{V}}(x) = 2, \quad \int_0^1 \mathrm{d}x \, d_{\mathrm{V}}(x) = 1 \, .$$

The experimental result

$$\sum_{\mathrm{q}} \int_0^1 \mathrm{d}x \, x \left[q(x) + \overline{q}(x) \right] = \int_0^1 \mathrm{d}x \, x \left[u_{\mathrm{V}}(x) + d_{\mathrm{V}}(x) + 6S(x) \right] \approx 0.5 \, ,$$

indicates that the quarks only carry about the 50% of the proton momentum, the rest can be attributed to the gluon constituents. A good knowledge of the PDF is crucial to make predictions for the SM and beyond the SM processes at hadron colliders. The PDF are non-perturbative functions that have to be determined experimentally, and several collaborations around the world are dedicated to provide them and their uncertainties. The PDF obtained by different collaborations (such as NNPDF, MMHT, CTEQ or MSTW) may differ on the choices of the datasets used, the treatment of data errors, the order of perturbation theory or the parameterisation at the initial scale (Q_0). Figure 3.4 shows a typical set of Parton Distribution Functions (PDF) extracted from fits to data in the next-to-next-to-leading-order (NNLO) NNPDF3.0 (MMHT2014) global analysis [73] ([74]) at scales $\mu^2 = 10$ GeV^2 and $\mu^2 = 10^4$ GeV^2.

The behaviour of the structure functions predicted in the Bjorken limit is related to the assumption that the transverse momentum of the partons in the infinite momentum frame of the proton is small. However, in QCD, this assumption is violated since hard gluons can be radiated by quarks, leading to logarithmic scaling violations, that are particularly large at small x. This feature is illustrated in figure 3.5 for the proton structure function F_2^{p} at two different Q^2 values. The radiation of gluons produces the evolution of the structure functions. As Q^2 increases, more gluons are radiated. This process leads both to the softening of the initial quark momentum distributions and to the growth of the gluon density and the $q\overline{q}$ sea density as x decreases.

In QCD, the processes are described in terms of scale-dependent PDF typically represented by $f_{\mathrm{i}}(x, \mu_{\mathrm{F}}^2)$, where $i = q$ or g, and, μ_{F} is the factorisation scale, typically set to represent the scale of the probe Q. For parton distributions, x always refers to the nucleon momentum fraction of the parton. For $Q^2 \gg M^2$, the structure functions are of the form

$$F_{\mathrm{a}} = \sum_{\mathrm{i}} C_{\mathrm{a}}^{\mathrm{i}} \otimes f_{\mathrm{i}} + \mathcal{O}\left(M^2/Q^2\right) \ ,$$

where $\otimes$ denotes the convolutional integral

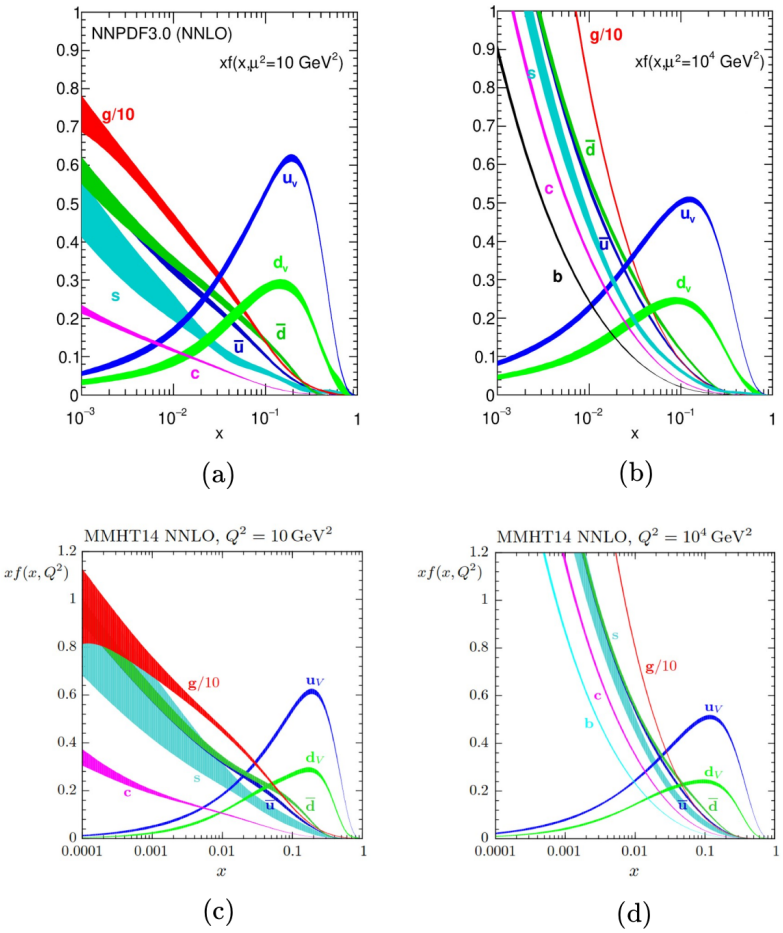


Figure 3.4: The bands display x times the parton distributions $f(x)$ (where $f = u_{\mathrm{V}}, d_{\mathrm{V}}, \overline{u}, \overline{d}, s \simeq \overline{s}, c = \overline{c}, b = \overline{b}, g$) obtained in the NNLO NNPDF3.0 global analysis (a,b) [73] at scales $\mu^2 = 10$ GeV^2 (left) and $\mu^2 = 10^4$ GeV^2 (right), with $\alpha_{\mathrm{S}}(M_{\mathrm{Z}}^2) = 0.118$. The analogous results obtained in the NNLO MMHT2014 analysis obtained in ref. [74] are shown (c,d).

$$C \otimes f = \int_{\mathrm{x}}^{1} \frac{\mathrm{d}y}{y} C(y)\, f\left(\frac{x}{y}\right) ,$$

where the coefficient functions $C_{\mathrm{a}}^{\mathrm{i}}$ are given as a power series in α_{S}. The PDF f_{i} corresponds, at a given x, to the density of parton i in the proton integrated over transverse momentum up to μ_{F}. Its evolution in μ_{F} is described in QCD by a DGLAP (Dokshitzer–Gribov–Lipatov–Altarelli–Parisi) equation [76–79], which has the form

$$\frac{\partial f_{\mathrm{i}}(x, \mu_{\mathrm{F}}^2)}{\partial \ln \mu_{\mathrm{F}}^2} = \frac{\alpha_{\mathrm{S}}(\mu_{\mathrm{F}}^2)}{2\pi} \sum_{\mathrm{j}} \int_{\mathrm{x}}^{1} \frac{\mathrm{d}\epsilon}{\epsilon} P_{\mathrm{ij}}\left(\frac{x}{\epsilon}, \alpha_{\mathrm{S}}(\mu_{\mathrm{F}}^2)\right) f_{\mathrm{j}}(\epsilon, \mu_{\mathrm{F}}^2) , \tag{3.10}$$

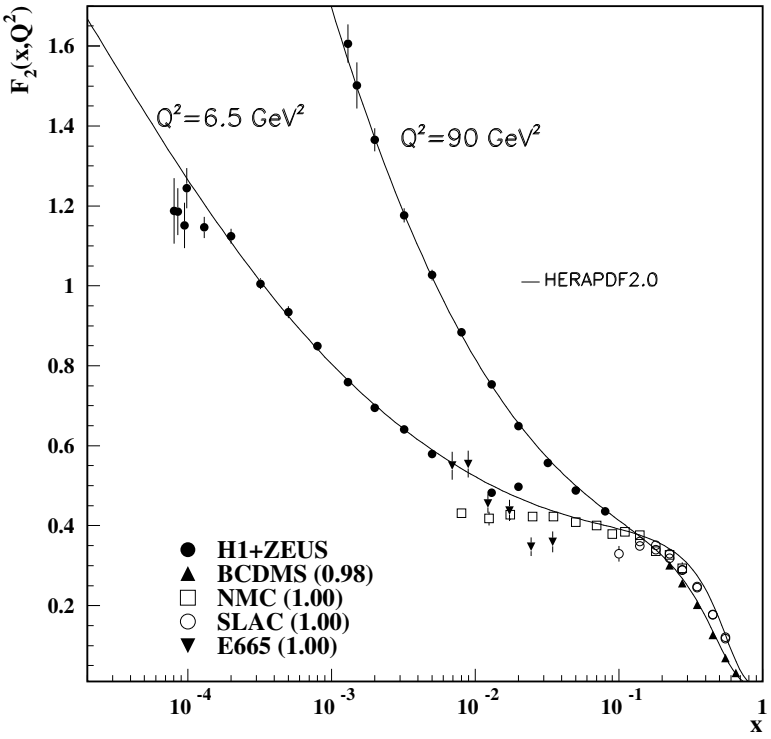


Figure 3.5: The proton structure function F_2^{p} given at two Q^2 values (6.5 GeV^2 and 90 GeV^2), which exhibit scaling at the *pivot* point $x \sim 0.14$. The curves were obtained using the PDF from the HERAPDF analysis [75]. From ref. [72].

where the P_{ij}, which describe the parton splitting $j \to i$, are also given as a power series in α_{S}. Equation 3.10 is the analogue of the β function equation 3.5 describing the variation of α_{S} with μ_{R}, and is one of the most important equations of pQCD. Although the evolution of the PDF can be predicted from a particular scale μ with equation 3.10, they can not be predicted a priori at any particular μ_0. Thus, they must be measured at a starting point μ_0 before the predictions of QCD can be compared to the data at other scales.

The P splitting functions describe the probability of a given parton splitting into two others. The LO Altarelli-Parisi splitting functions are

$$P_{\mathrm{qq}} = \frac{4}{3}\left[\frac{1+x^2}{(1-x)}\right]_+ = \frac{4}{3}\left[\frac{1+x^2}{(1-x)_+}\right] + 2\delta\left(1-x\right) ,$$

$$P_{\mathrm{qg}} = \frac{1}{2}\left[x^2 + (1-x)^2\right] ,$$

$$P_{\mathrm{gq}} = \frac{4}{3}\left[\frac{1+(1-x)^2}{x}\right] ,$$

$$P_{\mathrm{gg}} = 6\left[\frac{1-x}{x} + x(1-x) + \frac{x}{(1-x)_+}\right] + \left[\frac{11}{2} - \frac{n_{\mathrm{f}}}{3}\right]\delta(1-x) ,$$

where n_{f} is the number of quark active flavours and the notation $[F(x)]_+$ defines a distribution such that for any sufficiently regular test function, $f(x)$,

$$\int_0^1 \mathrm{d}x\, f(x)\, [F(x)]_+ = \int_0^1 \mathrm{d}x\, (f(x) - f(1))\, F(x) .$$

In general, the splitting functions can be expressed as a power series in α_{S}. The series contains terms proportional to $\ln \mu_{\mathrm{F}}^2$, $\ln(1/x)$ and $\ln(1-x)$. The splitting functions are known up to NNLO. The precision of the experimental data demands that at least next-to-leading-order (NLO), and preferably NNLO, DGLAP evolution is used in comparisons between QCD theory and experiment. Beyond the LO, it is necessary to specify and use consistently, both a renormalisation and a factorisation scheme. The renormalisation scheme used almost universally is the modified minimal subtraction ($\overline{\mathrm{MS}}$) scheme [80]. The most popular choice for the factorisation scheme is also the $\overline{\mathrm{MS}}$ [81].

As for the renormalisation of the coupling constant, $f_{\mathrm{i}}(x)$ can be regarded as an unmeasurable, bare distribution. The collinear singularities are absorbed into this bare distribution at a factorisation scale μ_{F}, which plays a similar role as the renormalisation scale μ_{R}. The renormalised distribution $f_{\mathrm{i}}(x, \mu_{\mathrm{F}}^2)$ can not be calculated from first principles in perturbation theory, since it receives contributions from the long-distance (non-perturbative) part of the strong interaction. It can be determined from structure function data at any particular scale. The ability to factorise the long- and short-distance contributions to the structure functions in this way is a fundamental property of the theory [82].

3.2.6 The QCD improved parton model

The high-energy interactions of hadrons are described by the QCD improved parton model. In this model, a hard scattering (i.e. high Q^2) process between two hadrons is the result of an interaction between the quarks and gluons that constitute the incoming hadrons. The factorisation property of the theory [82] allows to derive predictions for the cross sections by separating (factorising) the long-distance (soft) component from the short-distance (hard) perturbative process.

The cross section of a hard scattering process initiated by two hadrons with four-momenta P_1 and P_2 represented in figure 3.6 can be expressed by the a convolution of the universal parton density functions and the short-distance partonic cross section

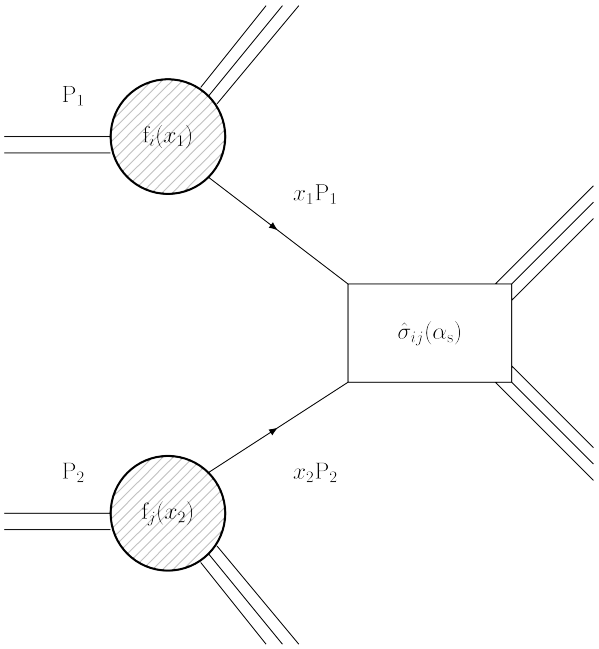


Figure 3.6: The parton model description of a hard scattering process.

$$\sigma\left(P_1, P_2\right) = \sum_{\mathrm{i,j}} \int \mathrm{d}x_1 \mathrm{d}x_2 f_{\mathrm{i}}\left(x_1, \mu_{\mathrm{F}}^2\right) f_{\mathrm{j}}\left(x_2, \mu_{\mathrm{F}}^2\right) \hat{\sigma}_{\mathrm{ij}}\left(p_1, p_2, \alpha_{\mathrm{S}}\left(\mu_{\mathrm{R}}^2\right), \frac{Q^2}{\mu_{\mathrm{F}}^2}, \frac{Q^2}{\mu_{\mathrm{R}}^2}\right) .$$

The momenta of the partons which participate in the hard interaction are $p_1 = x_1 P_1$ and $p_2 = x_2 P_2$. The characteristic scale of the hard scattering is denoted by Q. This scale can be the mass of a weak boson or a heavy quark, or the transverse momentum of a jet or a photon. The functions $f_{\mathrm{i}}\left(x, \mu_{\mathrm{F}}^2\right)$ are the QCD parton distribution functions, defined at a factorisation scale μ_{F} for a parton species i as a function of the momentum fraction x_{i}. The short-distance cross section for the scattering of partons of types i and j is denoted by $\hat{\sigma}_{\mathrm{ij}}$. Since the coupling is small at high energy, the short-distance cross section can be calculated as a perturbation series in α_{S}. The $(n+k)$th-order approximation of the short-distance cross section is given by

$$\hat{\sigma} = \alpha_{\mathrm{S}}^{\mathrm{k}} \sum_{\mathrm{m}=0}^{\mathrm{n}} c^{(\mathrm{m})} \alpha_{\mathrm{S}}^{\mathrm{m}} ,$$

where the $c^{(\mathrm{m})}$ are functions of the kinematic variables and the factorisation scale, and different hard processes contribute with different leading powers of k.

In the LO approximation ($n = 0$), the short-distance cross section is identical to the normal parton scattering cross section. At higher orders, the short-distance cross section is derived from the parton scattering cross section by removing long-distance pieces that are factorised into the parton distribution functions. This is necessary because

the perturbative calculation of the cross section contains contributions from interactions that occur long before the hard scattering, and these pieces can be factored out and absorbed into the description of the incoming hadrons. The remaining contributions to the cross section involve only high-momentum transfers, and are insensitive to the physics of low-momentum scales. This factorisation property of the cross section can be proved to all orders in perturbation theory [83]. Factorisation is a fundamental property of the theory that turns QCD into a reliable calculational tool with controllable approximations, extending its power with respect to the original parton model of Feynman [84].

The factorisation scale μ_F is an arbitrary parameter that can be thought as the scale that separates long- and short-distance physics. Thus, a parton emitted with a small transverse momentum, less than the scale μ_F, is considered part of the hadron structure and is absorbed into the parton distribution. A parton emitted with a transverse momentum larger than μ_F is considered as part of the short-distance cross section. The scale μ_F should be chosen to be of the order of the hard scale Q which characterises the parton-parton interaction. The more terms included in the perturbative expansion, the weaker the dependence on μ_F will be. The same statement is true for μ_R.

As the scales are varied, the perturbative coefficients at higher order change in such a way that the cross section to all orders is independent of the scales

$$\frac{\partial\sigma}{\partial\mu_F} = \frac{\partial\sigma}{\partial\mu_R} = 0\,.$$

The simplifying assumption $\mu = \mu_R = \mu_F$ is often made, and the standard choice of scales is $\mu = Q$, the hard scattering scale.

3.3 Jet algorithms

The structure of hadronic final states in high-energy collisions is described in terms of jets, that can be defined intuitively as a large amount of hadronic energy in a small angular region. Jet algorithms are defined in order to specify a jet configuration unambiguously, both on the theory and experimental sides.

Different classes of jet algorithms have been developed that can be classified in cone- and cluster-type algorithms. Cone algorithms, such as the Seedless Infrared-Safe Cone jet algorithm (SISCone) [85], are defined by maximizing the amount of energy which can be covered by cones of fixed size, while in clustering algorithms, such as the k_t [62], Cambridge/Aachen [86] or anti-k_t [63], particles are assigned to jets iteratively according to whether a given energy-angle resolution variable y_{ij} exceeds a fixed resolution parameter y_{cut}. Cluster algorithms have been shown to be more appropriate for the study of the internal structure of jets and they do not suffer from ambiguities related to the treatment of overlapping cones in multi-jet events [62], since these ambiguities which can be present

in cone algorithms are avoided by the unique assignment of every particle to exactly one hard final-state jet or beam jet.

In the case of hadron collisions, an adequate jet definition must fulfil the requirements of being:

1. Simple to use in experimental analyses.
2. Simple to use in theory calculations.
3. Infrared and collinear safe.
4. Subject to small hadronisation corrections.
5. Able to factorise initial-state collinear singularities into universal distributions.
6. Not strongly affected by contamination from hadron remnants and the soft underlying event.

Requirements 1 and 2 are self-evident, and related with the idea that the algorithm has to be independent of the type of input provided, reconstructing the same topology at parton, particle or detector level. Condition 3 implies that the final jet configuration cannot be affected by the presence of extra soft particles radiated from the primary partons (infrared safety) or by particles radiated with small angles with respect to the parent parton (collinear safety). Conditions 3 and 4 are guaranteed in clustering algorithms by means of a proper choice of the resolution variable and recombination scheme. The jet definition for hadron collisions has to fulfil the extra requirements 5 and 6. The initial-state collinear singularities present in the QCD calculation at parton level are removed so that the hard scattering part is calculable in perturbation theory. This is achieved by factorising these singular contributions in a universal way and associating them with hadron structure functions. The fulfilment of condition 6 is required afterwards by optimizing the jet definition.

The requirements 1-6 are implemented in a jet-clustering algorithm by defining:

1. A test variable (energy-angle resolution) d_{ij}.
2. A recombination procedure.
3. A hard scattering scale.

The test variable is needed to decide whether or not two hadrons h_{i}, h_{j} belong to the same cluster, while the recombination procedure tells us how the resolution variables are related to those of the hadrons belonging to them. The introduction of a hard-scattering scale is necessary in the case of hadron collisions.

The k_t, Cambridge/Aachen and anti-k_t are among the most widely-used sequential recombination algorithms for analysing data from hadronic collisions. An extension relative to the k_t and Cambridge/Aachen algorithm definitions encompasses the anti-k_t case as well. In the first step the distances d_{ij} between entities (particles, pseudojets) i and j and d_{iB} between entity i and the beam B are introduced. Then, the inclusive clustering proceeds by identifying the smallest of the distances between d_{ij} and d_{iB}; if it is a d_{ij}, the entities i and j are recombined, while if it is a d_{iB}, the algorithm proceeds calling i a jet and removing it from the list of entities. The distances are recalculated and the procedure repeated until no entities are left.

The extended definition of the distance measures

$$d_{ij} = \min\left(k_{ti}^{2p}, k_{tj}^{2p}\right) \frac{\Delta_{ij}^2}{R^2}\,,$$

$$d_{iB} = k_{ti}^{2p}\,, \tag{3.11}$$

where $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$, and k_{ti}, y_i and ϕ_i are the transverse momentum, rapidity and azimuth of particle i, respectively. R is the usual radius parameter, and p is added to govern the relative power of the energy versus geometrical (Δ_{ij}) scales.

For $p = 1$ the inclusive k_t algorithm is recovered. For values $p > 0$ the behaviour of the jet algorithm with respect to soft radiation is similar to that observed for the k_t algorithm, since the ordering between particles for finite Δ_{ij} is maintained for all positive values of p. The case $p = 0$ corresponds to the inclusive Cambridge/Aachen algorithm which uses angular ordering independently of the transverse momentum of the entities. Negative values of p show a similar behaviour with respect to soft radiation. The case $p = -1$ corresponds to the anti-k_t algorithm.

The analyses presented in this thesis made use of the anti-k_t and k_t jet algorithms. A brief description comparing their properties is provided in the following section.

3.3.1 Properties and comparisons between algorithms

The functionality of the anti-k_t algorithm can be understood by considering an event with a few well-separated hard particles with transverse momenta k_{t1}, k_{t2}, ... and many soft particles. The distance $d_{1i} = \min(1/k_{t1}^2, 1/k_{ti}^2)\,\Delta_{1i}^2/R^2$ between the hard particle 1 and the soft particle i is exclusively determined by the transverse momentum of the hard particle and the Δ_{1i} separation. Note that the distance d_{ij} between two similarly separated soft particles will instead be much larger. A consequence of this property is that soft particles will tend to cluster with hard ones long before they cluster among themselves. If a hard particle has no hard neighbours within a distance $2R$, it will accumulate all the soft particles around within a cone of radius R, resulting in a perfectly conical jet.

When an additional hard particle 2 is present at an angular distance $R < \Delta_{12} < 2R$, there will be two jets. It is not possible, for both of them, to be perfectly conical. If

$k_{t1} \gg k_{t2}$, then, jet 1 will be conical and jet 2 will be partially conical, missing the part overlapping with jet 1. When $k_{t1} = k_{t2}$, neither jet will be conical and the overlapping part will simply be divided by a straight line equally between the two. For a general situation with $k_{t1} \sim k_{t2}$, both jets will be clipped, with the boundary b between them defined by $\Delta_{1b}/k_{t1} = \Delta_{2b}/k_{t2}$

In the case where $\Delta_{12} < R$, the hard particles will cluster to form a single jet. If $k_{t1} \gg k_{t2}$, then, it will be a conical jet centred on particle 1. For $k_{t1} \sim k_{t2}$, the shape will be more complex, being the union of cones (radius $< R$) around each hard particle plus a cone (of radius R) centred on the final jet.

Thanks to these properties, soft particles do not modify the shape of the jet, while hard particles do. In this sense, the jet boundary in this algorithm is said to be resilient with respect to soft radiation, but flexible with respect to hard radiation. A priori, it is not clear whether it is better to have regular (soft-resilient) or less regular (soft-adaptable) jets. Regularity implies a certain rigidity in the ability of the jet algorithm to adapt a jet to the successive branching nature of QCD radiation. On the other hand, knowledge of the typical shape of jets is often quoted as facilitating experimental calibration of jets. Soft-resilience can also simplify certain theory calculations, as well as eliminate parts of the momentum-resolution loss caused by *underlying event* (UE) and pile-up contamination.

3.3.1.1 Comparison between algorithms

The behaviour of different jet algorithms is illustrated in figure 3.7. The authors considered a parton-level together with $\sim 10^4$ random soft *ghost* particles (as in ref. [87]) and then clustered them with four different algorithms. Because of the infrared safety property, the addition of ghosts does not change the momenta of the final jets. The use of ghost particles is motivated by the fact that the number of ghosts contained in a jet provides a measure of its *active area*[5]. The region within which the random ghosts are clustered into each partonic jet is represented. For the k_t and Cambridge/Aachen algorithms the jet areas depend somewhat on the specific set of ghosts and that jagged borders of the jets are a consequence of the randomness of the ghosts. For SISCone, single-particle jets are seen to be regular, while composite jets have more varied shapes. Finally, with the anti-k_t algorithm, the hard jets are circular, with radius R, while softer jets have more complex shapes.

[5]The concept of *jet area* was introduced to study the way UE and pile-up can affect a jet [87]. A good measure of the jet susceptibility to contamination by the UE and pile-up is given by the extent of the region over which it is possible to capture UE and pile-up particles, i.e. the *catchment area* of the jet. Two definitions were introduced in ref. [87] of the definition of the area of a jet, the passive and active areas, which can be understood as a measure of the susceptibility of the jet to contamination from an UE with a pointlike or uniform and diffuse structure, respectively.

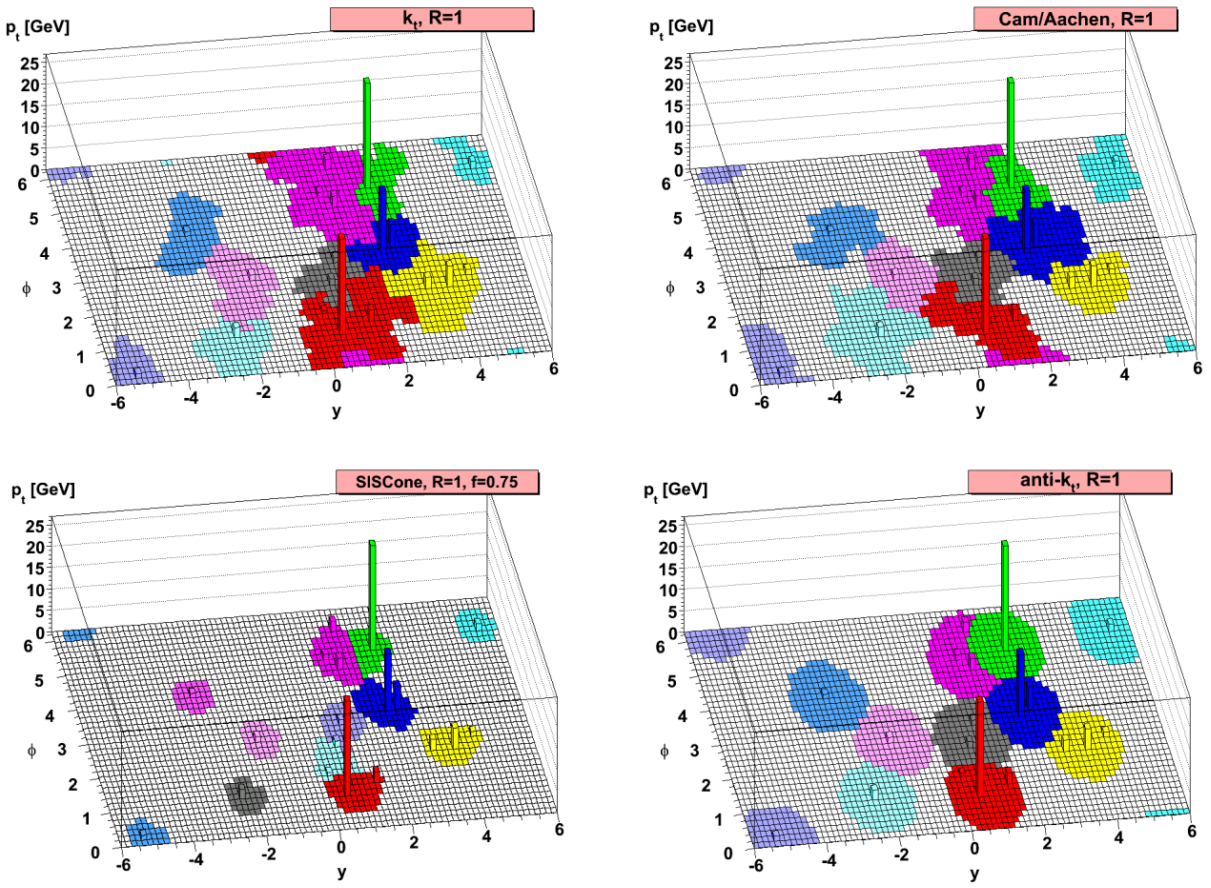


Figure 3.7: A sample parton-level event, together with many random soft ghosts, clustered with four different algorithms, illustrating the active catchment areas of the resulting hard jets. From ref. [63].

3.3.1.2 Jet resolution variable and recombination schemes

The resolution variables d_{ij} and d_{iB} shown in equation 3.11 are chosen as a particular extrapolation of a more fundamental definition of the jet resolution variables. This extrapolation is Lorentz invariant under longitudinal boosts. The main constraints in choosing an extrapolation are that d_{ij} and d_{iB} have to be increasing functions of the angles and that their functional dependence is responsible for the detailed form of the factorisation of structure functions and hard cross-section contributions.

Introducing longitudinal-boost-invariant variables is not enough to guarantee the same property for the full clustering procedure, which is influenced by the recombination scheme considered as well. Distinct recombination prescriptions differ among themselves in the way in which the resolution variables are defined for the pseudoparticles formed after merging two into a single cluster. The simplest recombination prescription, called the E-scheme, is used in the analyses presented in this thesis. It amounts to computing the

resolution variable for a pseudoparticle (p_{i}, p_{j}) as if it was a particle with four-momentum $p_{\mathrm{ij}} = p_{\mathrm{i}} + p_{\mathrm{j}}$. Other recombination prescriptions can be considered and may be more suitable to implement in computer codes for experimental analyses.

3.4 Theory of prompt-photon production

High transverse momentum (prompt) photons constitute colourless probes of the hard interaction and their production in proton-proton collisions, $pp \to \gamma + X$, provides a testing ground for QCD. Prompt photons are defined as those that are not secondaries from hadron decays. As it was previously introduced in chapter 1, prompt-photon production is understood to proceed via two processes: the direct-photon process, in which the photon arises directly from the hard interaction, and the fragmentation-photon process, in which the photon is emitted in the fragmentation of a high transverse momentum parton [6, 7]. From a topological point of view, when a direct photon is produced, it is most likely that it will be separated from the hadronic activity, whereas a photon produced from a fragmentation process, is most probably accompanied by hadrons. An exception occurs when the fragmentation photon carries away most of the momentum of the fragmenting parton. These fragmentation configurations are rare and less typical, and they are precisely those which are not suppressed by the isolation criterion [88].

At LO in perturbation theory, there are two subprocesses: (a) the annihilation process $q\overline{q} \to \gamma g$, and (b) the Compton process $qg \to \gamma q$, shown in figure 3.8. At medium and large x, there is a natural hierarchy of parton distributions in the proton, $q \gg g \gg \overline{q}$, while at small x, $g \gg q, \overline{q}$. As a consequence, in proton-proton collisions, the qg Compton process dominates over essentially all the p_{T} range. This makes direct photon production particularly useful for constraining the gluon distribution.

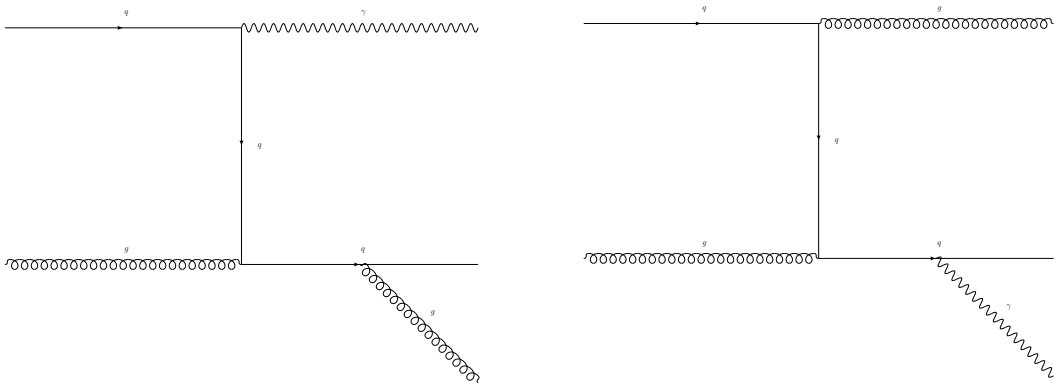


(a) (b)

Figure 3.8: Feynman diagrams for the LO direct-photon processes (a) $q\overline{q} \to \gamma g$ and (b) $qg \to \gamma q$.

First-order perturbative corrections to this processes are represented in figure 3.9. In figure 3.9(a), there is a collinear singularity when the momenta of the final-state quark and gluon are parallel. This divergence cancels when real and virtual gluon contributions are

summed, and the net effect is a finite $\mathcal{O}(\alpha_S)$ correction to the LO process. On the other hand, in the diagram of figure 3.9(b) there is another collinear singularity, this time, when the photon and quark momenta are parallel. This singularity, however, does not cancel, but has to be absorbed into a photon fragmentation function $D_q^\gamma(z, \mu_f^2)$ that represents the probability of finding a photon carrying longitudinal momentum fraction z in a quark jet at scale μ_f. This fragmentation function is not calculable in perturbation theory, and obeys a DGLAP evolution equation similar to that for the hadron fragmentation functions. The contribution to the cross section from figure 3.9(b) contains a piece of the form



(a) (b)

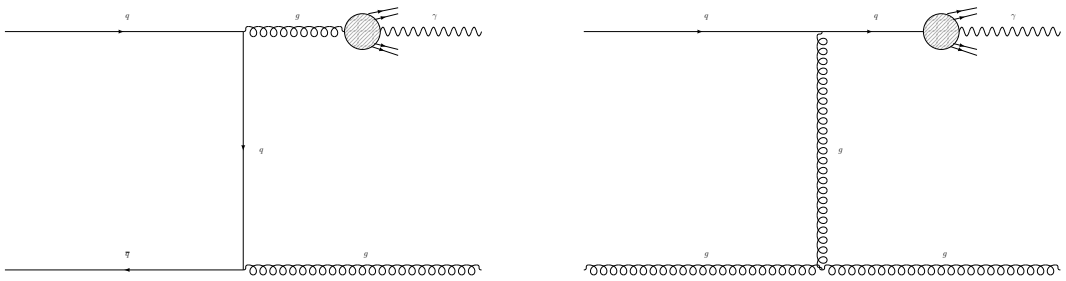
Figure 3.9: Feynman diagrams that contribute to the direct photon cross section at NLO.

$$\hat{\sigma}(qg \to gq) \otimes D_q^\gamma\left(z, \mu_f^2\right) . \tag{3.12}$$

The photon-fragmentation contribution appears when a final-state quark-photon collinear singularity occurs in the calculation of the contribution from subprocesses such as $qg \to gq\gamma$. At higher orders, multiple final-state collinear singularities appear in any subprocess where a high-p_T parton undergoes a cascade of successive collinear splittings ending up with a quark-photon splitting. These singularities are factorised to all orders in α_S according to the factorisation theorem [82], and are absorbed into quark and gluon fragmentation functions of the photon, $D_q^\gamma(z, \mu_f^2)$ and $D_g^\gamma(z, \mu_f^2)$, respectively.

The photon fragmentation function increases uniformly with the scale over the whole z range, i.e. $D_k^\gamma(z, \mu_f^2) \sim d^\gamma(z) \ln \mu^2$ as $\mu^2 \to \infty$. When the p_T is large with respect to ~ 1 GeV, the $\ln p_T^2$ growth of the fragmentation function in equation 3.12 compensates one of the $\alpha_S(p_T^2)$ couplings in the subprocess cross section, and the contribution is effectively of order $\alpha_S(p_T^2)\,\alpha_{EM}$, i.e. the same as the LO contribution. Feynman diagrams corresponding to the LO fragmentation component are shown in figure 3.10.

The inclusive differential cross section in E_T^γ for the production of a non-isolated photon is given by the sum of the fragmentation and direct contributions, and can be written as



(a) (b)

Figure 3.10: Feynman diagrams for the LO fragmentation-photon processes (a) $q\overline{q} \to \gamma(g)g$ and (b) $qg \to \gamma(q)g$.

$$\frac{d\sigma}{dE_{\mathrm{T}}^{\gamma}} = \frac{d\sigma_{\mathrm{dir}}}{dE_{\mathrm{T}}^{\gamma}} + \frac{d\sigma_{\mathrm{frag}}}{dE_{\mathrm{T}}^{\gamma}} = \sum_{\mathrm{a,b=q,\overline{q},g}} \int \mathrm{d}x_{\mathrm{a}}\mathrm{d}x_{\mathrm{b}}\, f_{\mathrm{a}}\left(x_{\mathrm{a}}, \mu_{\mathrm{F}}^2\right) f_{\mathrm{b}}\left(x_{\mathrm{b}}, \mu_{\mathrm{F}}^2\right) \times \tag{3.13}$$

$$\left[\mathrm{d}\hat{\sigma}_{\mathrm{ab}}^{\gamma}\left(p_{\gamma}, x_{\mathrm{a}}, x_{\mathrm{b}}, \mu_{\mathrm{R}}, \mu_{\mathrm{F}}, \mu_{\mathrm{f}}\right) + \sum_{\mathrm{c=q,\overline{q},g}} \int_{z_{\min}}^{1} \frac{\mathrm{d}z}{z^2} \mathrm{d}\hat{\sigma}_{\mathrm{ab}}^{\mathrm{c}}\left(p_{\gamma}, x_{\mathrm{a}}, x_{\mathrm{b}}, z, \mu_{\mathrm{R}}, \mu_{\mathrm{F}}, \mu_{\mathrm{f}}\right) D_{\mathrm{c}}^{\gamma}(z, \mu_{\mathrm{f}}^2) \right],$$

where $D_{\mathrm{c}}^{\gamma}(z, \mu_{\mathrm{f}}^2)$ is the fragmentation function of a parton c to a photon carrying momentum fraction z, $f_{\mathrm{a}}\left(x_{\mathrm{a}}, \mu_{\mathrm{F}}^2\right)$ is the PDF of a parton a, μ_{R} and μ_{F} are the standard renormalisation and factorisation scales, and μ_{f} is the fragmentation scale. Corrections to the direct component of the partonic cross section $\hat{\sigma}_{\mathrm{ab}}^{\gamma}$ are known up to the NNLO in pQCD, while the fragmentation component $\hat{\sigma}_{\mathrm{ab}}^{\mathrm{c}}$ is only known at NLO.

At LO, the theory calculations for the direct and fragmentation processes converge separately, and can be considered independently. However, this distinction has no physical meaning beyond the LO, since both kind of processes need to be considered at the same time to cancel the final-state infrared and collinear singularities. Therefore, beyond the LO, both direct and fragmentation processes cannot be considered separately. From a theoretical point of view, the distinction is defined by an arbitrary choice. It follows from the necessity of factorising the final-state collinear singularities and absorbing them into the fragmentation functions. This factorisation requires the introduction of an arbitrary fragmentation scale μ_{f}, which is a non-physical parameter. More generally, it relies on the arbitrary choice of the factorisation scheme, which defines the finite part of the higher-order corrections that is absorbed in the fragmentation functions together with the singularities; the remaining finite part is then included in the higher-order contributions to the partonic cross sections. The dependence on this arbitrariness, and in particular, on μ_{f}, cancels only in the sum of the direct and fragmentation contributions, so only this sum is a physical observable.

3.4.1 Photon isolation

In order to distinguish prompt photons from those produced through hadron decays, experimental analyses apply isolation criteria to the measured photons. These criteria consist commonly of applying a restriction to the amount of hadronic transverse energy in the vicinity of a photon candidate, becoming part of the photon definition. A perfect isolated photon would not be surrounded by hadronic energy in a certain region typically defined by a cone in pseudorapidity η and azimuthal angle ϕ. However, this definition is not possible from an experimental point of view and is not theoretically well defined. The reason behind is that in order to construct an observable infrared safe under QCD corrections, the emission of soft partons must be allowed everywhere in the final-state phase space. Therefore, the deposition of a certain amount of parton energy close to the photon is necessary to allow the cancellation of infrared divergences between soft real radiation and virtual corrections.

The isolation criterion used by collider experiments is schematically as follows. A photon is said to be isolated if in a cone of radius R in the (η,ϕ)-plane around the photon direction, the amount of deposited hadronic transverse energy $E_{\mathrm{T}}^{\mathrm{iso}}(R)$ is smaller than a threshold value $E_{\mathrm{T}}^{\mathrm{max}}$, that can be defined as a constant value or as a function of the transverse energy of the photon candidate E_{T}^{γ}

$$E_{\mathrm{T}}^{\mathrm{iso}}(R) < E_{\mathrm{T}}^{\mathrm{max}} \,. \tag{3.14}$$

The application of an isolation criterion implies that the cross section is no longer fully inclusive over the hadronic final state and, hence, that the factorised equation 3.13 is not necessarily valid. To preserve factorisation, the isolation requirement must fulfil the following conditions:

- Infrared safety: the cross section must be insensitive to the momenta of arbitrarily soft particles.
- Collinear safety: when some final-state particles are produced collinearly, the cross section must depend on their total momentum rather than on the momentum of each of them.
- Final-state collinear factorisability: guaranteeing that all the long-distance phenomena related to the low-momentum fragmentation of the photon may be absorbed and factorised in the universal fragmentation function $D_{\mathrm{a}}^{\gamma}(z, \mu_{\mathrm{f}})$.
- Initial-state collinear factorisability: in hadron-hadron collisions, the cross section is affected by additional long-distance phenomena related to the non-perturbative binding of the colliding partons into the incoming hadrons. At the parton level, these phenomena lead to initial-state singularities that have to be absorbed and factorised in the non-perturbative parton distributions of the colliding hadrons.

This property guarantees that the isolation criterion does not spoil the factorisation of the initial-state collinear singularities.

A proof for the preservation of the factorisation property was provided in ref. [88], and the main conclusion on the generalised isolation criterion in equation 3.14 is that QCD factorisation is valid at any order in perturbation theory.

3.4.1.1 Photon isolation prescriptions

The isolation criterion defined in equation 3.14 is commonly referred to as the *fixed-cone isolation* prescription. For the results presented in this thesis, a linear dependence of $E_{\mathrm{T}}^{\mathrm{max}}$ on E_{T}^{γ} is chosen

$$E_{\mathrm{T}}^{\mathrm{max}} = \epsilon E_{\mathrm{T}}^{\gamma} + E_{\mathrm{T}}^{\mathrm{thres}} .$$

This prescription can be implemented in experimental analyses in a standard manner. However, when implemented in a theory calculation, it introduces a sensitivity to the photon fragmentation process.

An alternative is the implementation of the *dynamical-cone isolation* prescription proposed by Frixione [89], where, starting from a cone with radius R_{d}, smaller concentric sub-cones with $r_{\mathrm{d}} \leq R_{\mathrm{d}}$ are considered. The allowed hadronic energy $E_{\mathrm{T}}^{\mathrm{iso}}(r_{\mathrm{d}})$ decreases with decreasing r_{d}, and reaches zero for $r_{\mathrm{d}} = 0$. Hence, the photon candidate is only accepted if the admitted hadronic energy $E_{\mathrm{T}}^{\mathrm{iso}}(r_{\mathrm{d}})$ for any sub-cone radius r_{d} is below a certain threshold

$$E_{\mathrm{T}}^{\mathrm{max}}(r_{\mathrm{d}}) = \epsilon_{\mathrm{d}} E_{\mathrm{T}}^{\gamma} \left(\frac{1 - \cos r_{\mathrm{d}}}{1 - \cos R_{\mathrm{d}}} \right)^{\mathrm{n}} \text{ for all } r_{\mathrm{d}} < R_{\mathrm{d}} ,$$

where ϵ_{d} and n are free parameters. With this implementation, exactly collinear hard radiation in the photon direction is forbidden, while soft radiation is admitted over the whole phase space. The most important consequence is that the theory prediction is independent of the photon fragmentation process, which leads to a considerable simplification of the evaluation of higher order corrections. However, because of finite detector resolution, the dynamical-cone isolation can only be approximated in experimental analyses.

Due to the lower computational complexity associated with the dynamical-cone isolation procedure, calculations of NNLO QCD corrections to direct photon production have been performed using this prescription only. To compare these theory calculations with measurements, based all on a fixed-cone isolation prescription, the free parameters are adjusted to mimic the effect of the fixed-cone isolation. The present discrepancy between the isolation procedures used in experimental measurements and theory calculations is nevertheless unsatisfactory, and prevents quantitative statements on the impact of varying isolation parameters. Fully consistent NNLO predictions with a fixed-cone isolation

still require the computation of fragmentation contributions to this order, but an improvement over the present predictions is possible by means of the implementation of a *hybrid-cone isolation*, which can be formulated as a combination of a dynamical isolation with small value of R_{d} with a fixed-cone isolation with larger value of R, such that $R^2 \gg R_{\mathrm{d}}^2$. The dynamical-cone isolation is applied first, such that events that are very close to the collinear divergence are vetoed, and the dependence on the fragmentation process is eliminated in an infrared-safe manner. After that, the fixed-cone isolation is applied to all events that passed the dynamical-isolation requirement. The hybrid isolation procedure only discards some events within a small inner fraction of the cone area, potentially introducing an unknown $R-$independent shift of the total amount of $E_{\mathrm{T}}^{\mathrm{iso}}(R)$. The hybrid-cone prescription has been used by the ATLAS collaboration [32,33,37] to compare the measured cross sections with the NLO pQCD predictions from the SHERPA [90, 91] MC generator. In addition, this prescription has also been adopted in ref. [21] for the recent NNLO pQCD theory calculation in the NNLOJET framework.

3.4.2 Fixed-order pQCD calculations

Due to the complexity of multi-loop calculations, perturbative calculations are only known at fixed-order, especially for the high-energy processes involving hadrons in which the renormalisation and factorisation effects are entangled with each other. The impossibility of accounting for all the orders in a pQCD calculation introduces a dependence of the fixed-order predictions for a physical observable on the renormalisation and factorisation scales chosen. Hence, the fixed-order predictions for observables in pQCD are usually assumed to suffer from an uncertainty in fixing both scales. This ambiguity in making fixed-order predictions occurs because one usually assumes an arbitrary scale, typically representing the momentum flow of the process which is assumed to be the effective virtuality of the strong interaction in that process, together with an arbitrary range to ascertain its uncertainty. This *ad hoc* assignment of the scales, however, causes the coefficients of the QCD running coupling at each perturbative order to depend strongly on the choice of the scale and scheme used.

3.4.2.1 JETPHOX NLO pQCD calculations

For the three analyses presented in this thesis, NLO fixed-order pQCD predictions were computed using the program JETPHOX 1.3.1_2 [88,92]. JETPHOX provides NLO pQCD calculations of the direct and fragmentation contributions to the prompt-photon cross section. The number of massless quark flavours is set to five. The renormalisation, factorisation and fragmentation scales are chosen to be $\mu_{\mathrm{R}} = \mu_{\mathrm{F}} = \mu_{\mathrm{f}} = E_{\mathrm{T}}^{\gamma}$. Variations of these scales are considered to get an estimate of the uncertainties due to missing higher-order terms in the perturbative expansion (see sections 7.8.2, 8.8.2, 10.8.1).

The nominal prediction is obtained using the MMHT2014 [74] PDF set and the BFG set II of parton-to-photon fragmentation functions [93], both of which are determined

at NLO. The strong coupling constant is set to $\alpha_S(M_Z^2) = 0.120$, since it is the value assumed in the MMHT2014 PDF fit. For the EM coupling, the low-energy limit of $\alpha_{EM} = 1/137.036$ is used. A parton-level fixed-cone isolation criterion is used, requiring the total transverse energy from the partons inside a cone of size $\Delta R = 0.4$ around the photon direction to be below the same $E_{T,cut}^{iso}$ applied at particle and detector level, defined as $E_T^{iso} < 4.2 \times 10^{-3} \times E_T^{\gamma} + 4.8$ GeV.

Predictions based on other PDF sets, namely CT14 [94], ABMP16 [95], HERAPDF2.0 [75] and NNPDF3.0 [73], are also computed in the inclusive isolated-photon analysis with 36 fb^{-1} of ATLAS data; in each case, $\alpha_S(M_Z^2)$ is set to the value assumed in the PDF fit. For the inclusive isolated-photon and photon-plus-jet analyses with the full Run II data, predictions based on the more recent PDF sets CT18 [96], and NNPDF3.1 [97] were also calculated.

3.4.2.2 NNLOJET NNLO pQCD calculations

The NNLO pQCD prediction is calculated [21] in the NNLOJET framework. The NNLO corrections include three types of parton-level contributions, namely the two-loop corrections to photon-plus-one-parton production, the virtual corrections to photon-plus-two-parton production and the tree-level photon-plus-three-parton production. Only direct-photon processes are included; fragmentation processes are circumvented by the application of the Frixione's criterion at small ΔR using parameters $R_d = 0.1$, $n = 2$, and $\epsilon_d = 0.1$. The photon transverse energy is used as the baseline choice for the renormalisation and factorisation scales. The NNPDF3.1 [97] PDF set at NNLO is used. The EM coupling is taken in the G_μ-scheme [98] as $\alpha_{EM}^{G_\mu} = 1/132.232$. Photon isolation is implemented following the hybrid-cone isolation prescription introduced in section 3.4.1.1. For the NNLOJET prediction, the fixed-cone isolation at $\Delta R = 0.4$ is applied using the same requirement as for the particle and detector levels and introduced in 3.4.2.1. The prediction at NLO QCD in the NNLOJET framework is also calculated to illustrate the improvements achieved including the NNLO QCD corrections. The NLO pQCD prediction of NNLOJET differs from that of JETPHOX in the fragmentation contribution, the photon isolation prescription, the value of α_{EM} and the PDF set.

3.4.3 Recent progress on Standard Model photon physics

Since the LHC started its operation, various measurements of inclusive isolated-photon production in pp collisions at $\sqrt{s} = 7$, 8 and 13 TeV have been performed by the ATLAS [8–10, 37] and CMS [11, 12] collaborations. In addition, measurements of the production of isolated photons in association with jets in pp collisions at $\sqrt{s} = 7$, 8 and 13 TeV are available from ATLAS [29–33] and CMS [12, 34–36] collaborations.

The prompt-photon measurements at the LHC are characterised by small uncertainties that demand precise theory predictions to fully exploit the data. Until recently, the

highest-fixed-order pQCD calculation available to be compared with many of the available measurements was at NLO accuracy. In 2017, the first NNLO pQCD calculation [20] based on the N-jettiness slicing procedure and computed with the MCFM program was published. Comparisons between the MCFM predictions at NNLO accuracy and the ATLAS 8 TeV measurement [9] were provided in a limited phase-space region, finding a significant improvement compared to previous calculations. In a more recent publication [16], updated NNLO theory predictions with the MCFM program were compared with the ATLAS 8 TeV and 13 TeV [10] measurements in the four rapidity regions in which the inclusive isolated-photon cross sections were measured. The inclusion of NNLO QCD and leading-logarithm (LL) electroweak corrections was found to provide a good quantitative agreement with both measurements, except for the most forward rapidity region in the former measurement. In addition, this study included the ATLAS measurement at 8 TeV in a NNLO QCD fit to obtain the parameterisations of the proton. It was observed that the inclusion of the ATLAS data led to a reduction in the gluon density uncertainties [16].

Since many years, there is a persistent discrepancy between the isolation prescriptions used in experimental measurements and theory calculations. All the measurements are performed using a fixed-cone isolation prescription, while theory calculations which do not incorporate the fragmentation component of the cross section are usually performed using a dynamical-cone isolation. This is the case for the MCFM predictions published in ref. [16, 20], for which Frixione's isolation is applied with parameters $R_{\mathrm{d}} = 0.4$, $n = 2$, and $\epsilon_{\mathrm{d}} = 0.1$. With the improvements of NNLO pQCD calculations and the increasingly more precise measurements performed at the LHC, the impact of this mismatch becomes more important. The publication of the first NNLO pQCD predictions using a hybrid-cone isolation [21] represented a step forward towards bringing the theory predictions closer in line with experiment. In these theory predictions, the calculation of the NNLO QCD corrections uses the antenna subtraction method and is performed in the NNLOJET framework.

For the NNLOJET theory calculations, as described in section 3.4.2.2, the choice of Frixione's parameters is $R_{\mathrm{d}} = 0.1$, $n = 2$, and $\epsilon_{\mathrm{d}} = 0.1$. The inner cone radius is set to $R_{\mathrm{d}} = 0.1$, while the outer one is $\Delta R = 0.4$, to guarantee that the condition $R^2 \gg R_{\mathrm{d}}^2$ is fulfilled. This condition guarantees that the hybrid procedure exactly reproduces the behaviour of the fixed-cone isolation used in experiment, and in particular, the R-dependence of the cross section. In addition, as R_{d} approaches to zero, the cross section diverges as $\sigma \sim \log\left(\frac{1}{R_{\mathrm{d}}}\right)$, following from the factorisation of the cross section in the photon-quark collinear limit at NLO [21]. In the limit $R_{\mathrm{d}} \to 0$, the fragmentation contribution to the cross section has to be included to compensate the divergence with a negative and divergent mass-factorisation counter term from the quark-to-photon fragmentation function. In the study performed in ref. [21], it is concluded that there exists a value of R_{d} for which a calculation based on the hybrid procedure should produce exactly the same results as a calculation including the fragmentation contribution to the cross section and using a fixed-cone prescription, since the vetoed real radiation cross section inside R_{d} and the negative fragmentation counter-term exactly compensate each other.

From the comparisons performed in ref. [21], the exact compensation is thought to take place at around $R_{\mathrm{d}} = 0.1$ or above. The NNLOJET NNLO pQCD theory predictions in ref. [21], were facilitated by the authors and could be used by the ATLAS collaboration in ref. [37] to provide comparisons with the measured cross sections. The NNLO pQCD calculations shown uncertainties due to the scale variations reduced by a factor in the range $2 - 20$ relative to those of a NLO pQCD calculation. Overall, the NNLO pQCD predictions were found to give an excellent description of the data, providing a precise test of the theory at $\mathcal{O}\left(\alpha_{\mathrm{EM}}\alpha_{\mathrm{S}}^{3}\right)$ in the range of photon transverse energies from 125 GeV up to and beyond 1 TeV. The increase of integrated luminosity with respect to the previous measurement at 13 TeV [10] guarantees its potential to further constrain the PDF, particularly the gluon density in the proton, within a global NNLO QCD fit. In addition, the study in ref. [21] provided comparisons of the NNLOJET NNLO pQCD calculations with the measurement of isolated-photon-plus-jet cross sections at 13 TeV [32], showing that the NNLO corrections can induce substantial changes in shape that often could not be resolved in NLO calculations due to the much larger scale uncertainties. Another relevant consequence previously mentioned of the use of the hybrid-cone isolation prescription is that the exact dependence of the cross section on the fixed cone R can be predicted, as long as the condition $R^2 \gg R_{\mathrm{d}}^2$ is fulfiled. More stringent tests of pQCD predictions can be performed in the future, once measurements for different cone sizes are available.

The default settings of perturbative calculations obtained with the MCFM program in ref. [20] and those of NNLOJET [21] differ in several aspects. The MCFM (NNLOJET) theory prediction uses the CT14 (NNPDF3.1) PDF set. The MCFM prediction implements a dynamical-cone isolation prescription, while the NNLOJET uses the hybrid procedure. Different choices for the value of α_{EM} were taken, the MCFM used $\alpha_{\mathrm{EM}} = 1/127.9$ while the NNLOJET predictions are obtained in the G_{μ}-scheme with $\alpha_{\mathrm{EM}}^{G_{\mu}} = 1/132.232$. To provide a numerical comparison between both NNLO pQCD calculations in ref. [21], the NNLOJET calculations were repeated with the CT14 PDF set, using a dynamical-cone isolation with the same Frixione's parameters choice and applying a constant rescaling factor to account for the difference in the EM constant values. The two theory predictions rely on independent implementations of the NNLO matrix elements and use different methods for the extraction and cancellations of infrared singularities among the different subprocesses. Mutual agreement between calculations within MC uncertainties is reported in ref. [21].

Electroweak corrections to the cross sections are not included in the calculations of JETPHOX or NNLOJET introduced in section 3.4.2. These corrections are likely to play an important role at the TeV scale, therefore, representing a theoretical aspect of increasing importance for the present and future photon measurements at the LHC. Electroweak Sudakov corrections for single photon production are available [99]. However, these constitute only part of the full electroweak corrections since real corrections are not included. The full electroweak corrections and the real corrections partially cancel out when the presence of W or Z bosons in the final state is not vetoed. The measurements of inclusive isolated-photon or isolated photon-plus-jet cross sections presented in this thesis

do not preclude the presence of massive gauge bosons in addition to the high-transverse momentum photon or photon-plus-jet.

With the improvements of the recent years from the theory side, predictions are on a solid basis with much smaller uncertainties that are competitive with those of experimental nature. The hybrid procedure, still contains an intrinsic theoretical ambiguity from the removal of the collinear contributions inside the narrow inner cone that can induce a small shift on the cross sections. Future progress from the theory side will require the calculation of the fragmentation component at NNLO to reduce the mismatch between experiment and theory for the isolation procedure. Once this is achieved, theory predictions at NNLO will be able to apply the same isolation criteria as used in the experiments. On the other hand, the explosion of the field of jet substructure has provided useful tools to explore soft and collinear dynamics within jets, together with an increasing interest on adapting these tools to handle photons and study photon isolation. In particular, a recent proposal [100] introduced soft-drop isolation, a new substructure-based photon-isolation technique with a solid theoretical basis, guaranteeing infrared and collinear safety and equivalent at leading (non-trivial) order to the definition of Frixione's isolation provided in section 3.4.1.1. Soft-drop isolation is a democratic technique, based on clustering algorithms that make it well suited to identify direct-photons in jet-rich environments. In addition, this technique could provide a new way to handle detector granularity, making possible its implementation both at the experimental and theoretical levels.

Chapter 4

Monte Carlo event generators

In this chapter, MC programs and their applications in particle physics are introduced. A brief description of some of the different stages that intervene in the generation of a hadron-collider event is provided. The elaboration of the different sections is based on the wide and detailed descriptions provided in ref. [90, 101], while more recent updates introduced in ref. [91, 102] are also included.

4.1 Monte Carlo simulations

The Monte Carlo technique is a way of calculating difficult integrals that may be hard to solve by ordinary numerical interpolation methods. High-energy collisions between elementary particles normally produce complex final states, which are populated by many hadrons, leptons, photons and neutrinos. The relation between the final states and the underlying physics description is not simple due to the lack of understanding of the physics and the fact that any analytical approach is not feasible due to the large particle multiplicities. An additional difficulty is related to the need to simulate complicated geometrical factors that represent detectors, a routine situation for experimenters.

Monte Carlo methods allow the generation of complete events. The complexity of the generation of such events is handled by subdividing the problem into a set of separate steps whose solution is simpler. The main aspects of the events are simulated, such as the hard-process selection, initial-state radiation (ISR) and final-state radiation (FSR), beam remnants, fragmentation, decays, etc. Therefore, the generated events can be compared with those experimentally observed. The MC programs can be used to extract physics from comparisons with existing data or to study physics at future experiments.

4.1.1 The complex skeleton of high-energy processes

High-energy processes have, to first approximation, a simple structure at the level of interactions between fundamental particles, i.e. quarks, leptons and gauge bosons that represent the basic skeleton for each process. Corrections to this picture can be subdivided into three main classes:

1. Bremsstrahlung-type modifications: produced by the emission of additional final-state particles by branchings such as $e \to e\gamma$ or $q \to qg$. The latter are due to the largeness of the strong coupling constant α_{S}, and because of the presence of the triple gluon vertex that make QCD emission of quarks and gluons especially prolific. Photon emissions may give sizeable effects in e^+e^- and ep processes. These emissions are commonly referred to as the *parton showers*, wherein a single initial parton may give rise to a whole bunch of partons in the final state. The bulk of the bremsstrahlung corrections are universal and only depend on a few key numbers, such as the momentum transfer scale of the process.

2. Higher-order corrections: there are higher-order corrections which involve a combination of loop graphs and the soft parts of the bremsstrahlung graphs mentioned in point 1. A combination between these two is needed to cancel some divergences. For a complete description, it is not possible to consider bremsstrahlung separately. For high-precision studies, it is crucial to take into account results of loop calculations, but usually effects are minor for the qualitative aspects of high-energy processes.

3. Quark and gluon confinement: the perturbative language used to describe the short-distance interactions of quarks, leptons and gauge bosons used in points 1 and 2 is sufficient for the case of leptons and colourless bosons, but that is not the case for quarks and gluons. The description of quarks and gluons must be complemented with the structure of the incoming hadrons and a picture of the hadronisation process, wherein the coloured partons are transformed into jets of colourless hadrons, photons and leptons. The hadronisation process can be further subdivided into fragmentation and decays. Fragmentation describes the way the creation of new quark-antiquark pairs can break up a high-mass system into lower-mass ones, ultimately hadrons.

In this way, the original skeleton process becomes much more complex, with hundreds of final state particles. However, the original physics has not disappeared, but the skeleton process has been dressed up and is no longer directly visible. A direct comparison between theory and experiment is complicated, and still requires a simulation of detector effects to generate events as detailed as could be observed by a perfect detector.

4.1.2 Event generators

The objective of an event generator is precisely to use computers to perform the simulation of detector effects. This is not done in one step, but rather by subdividing the full problem into a number of components that can be handled reasonably accurately. The general idea is that the hard process is used as input to generate bremsstrahlung corrections, the result of this exercise is then hadronised in a final step. Most of the steps involve the branching of one object into two, or a very small number, with the daughters free to branch in their turn. Hence, a lot of of book-keeping work is necessary, which can be left to computers given its repetitive nature.

The output of an event generator should be in the form of *events*, with the same average behaviour and fluctuations as real data. There is, however, an underlying distinction in the nature of fluctuations. In data, fluctuations arise from the quantum mechanics of the underlying theory, whereas MC techniques select all relevant variables according to the desired probability distributions, ensuring (quasi-)randomness in the final events. Some loss of information is entailed, since quantum mechanics is based on amplitudes, not probabilities.

An event generator can be used in many different ways. Some of the main applications are:

1. As a help to give physicists a feeling for the kind of events that can be expected to find, and at what rates.
2. As a key tool for the design of a new detector, so that detector performance is optimised for the study of promising physics scenarios.
3. As a tool for planning analysis strategies that should be used on real data, for instance, when optimising signal-to-background conditions.
4. As a method for the estimation of detector acceptance corrections that have to be applied to data to extract the *true* physics signal.
5. As a useful framework within which it is possible to interpret the observed phenomena in terms of a more fundamental underlying theory.

In MC programs, the role of the event production is taken by the event generators described in this section. The behaviour of the detectors is simulated in programs such as GEANT4 [103]. This latter activity is called event simulation or detector simulation, and the output of this simulation has exactly the same format as the real data recorded by the detector, and can therefore be put through the same event reconstruction and physics analysis chain.

4.2 Physics overview

Event generators contain simulations of many physics aspects to provide detailed descriptions of high-energy events. A description of the different steps involved in the event generation and simulation of each of the physics characteristics is provided trying to follow a time ordering:

1. Initially two beam particles come towards each other. In a hadron-hadron collision, each particle is characterised by a set of parton distributions which defines the partonic substructure in terms of flavour composition and energy sharing.

2. One shower initiator parton from each of the beams starts off a sequence of branchings such as $q \to qg$, which builds up an initial-state shower.

3. From each of the two showers, one incoming parton enters the hard process, where certain number of outgoing partons are produced, usually two. The nature of this process determines the main characteristics of the event.

4. The hard process may produce a set of short-lived resonances, such as $Z^0/W^\pm$ gauge bosons, whose decay to partons has to be considered in close association with the hard process itself.

5. The outcoming partons may branch, in the same way as the incoming partons, to build up final-state showers.

6. In addition to the hard process, further hard or semihard interactions may occur between the other partons of two incoming hadrons. These additional interactions between remnants of the incoming hadrons constitute the UE.

7. When a shower initiator is taken out of a beam particle, a beam remnant is left behind. This remnant may have an internal structure, and a net colour charge that relates it to the rest of the final state.

8. The QCD confinement mechanism ensures that the outgoing quarks and gluons are not observable, but instead fragment to colourless hadrons.

9. Many of the produced hadrons are unstable and decay further.

Figure 4.1 represents a hadron-collider event, where a $t\bar{t}H$ final state is produced and evolves by including effects of QCD bremsstrahlung in the initial and final state, the UE, hadronisation and, finally, the decays of unstable hadrons into stable ones. These different aspects are described in more detail in the following sections, pointing to the differences between the MC generators taken as a reference, PYTHIA [102] and SHERPA [90, 91], also used in the analyses presented in this thesis.

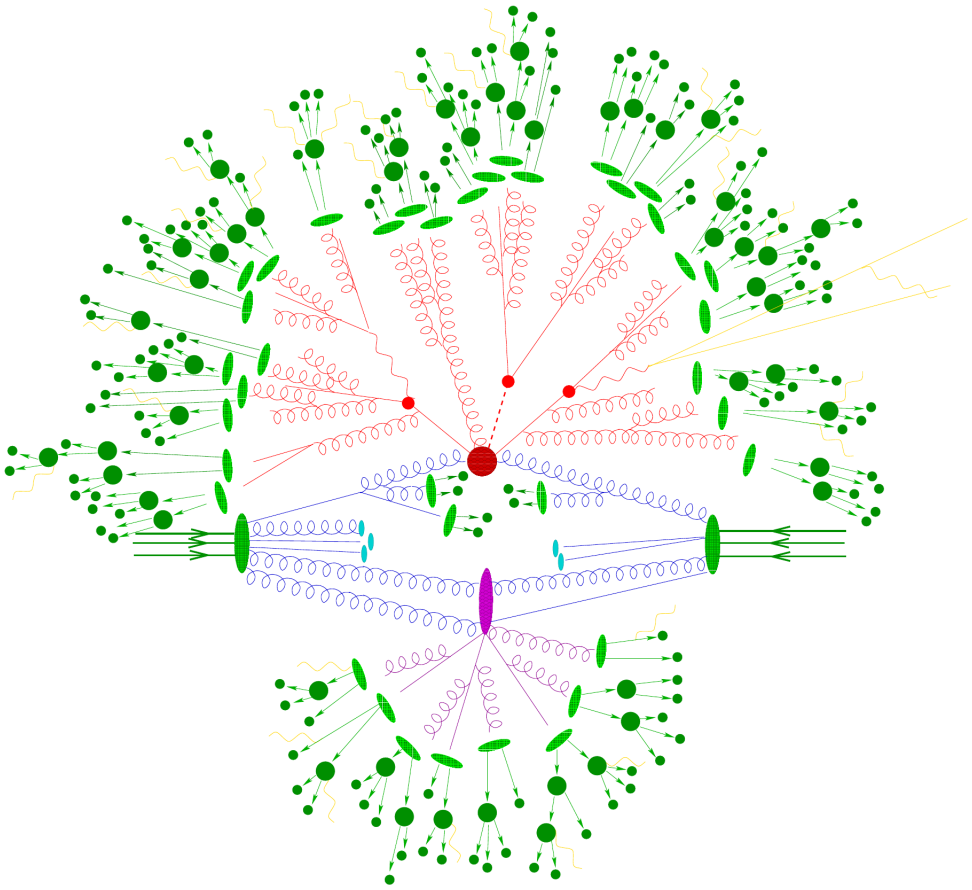


Figure 4.1: Pictorial representation of a $t\bar{t}H$ event as produced by an event generator. The hard interaction (big red blob) is followed by the decay of both top quarks and the Higgs boson (small red blobs). Additional hard QCD radiation is produced (red) and a secondary interaction takes place (purple blob) before the final-state partons hadronise (light green blobs) and hadrons decay (dark green blobs). Photon radiation occurs at any stage (yellow). From ref. [90].

4.2.1 Hard processes and parton distributions

The hard-process generation can be classified according to the number of final-state objects i.e.: $2 \to 1$ processes, $2 \to 2$ ones, $2 \to 3$ ones, etc. This type of classification is convenient from a programming point of view, since the amount of final-state objects is directly related with the complexity of the phase space and the whole generation procedure [101]. Another classification could be according to the physics scenario. In the case of PYTHIA and SHERPA MC generators, the former classification has been used for the analyses presented in this thesis:

– PYTHIA is optimised for $2 \to 1$ and $2 \to 2$ processes, and was used for the generation of prompt-photon and dijet samples with LO matrix elements.

– SHERPA simulates hard production processes with many-particle final states. In what concerns the study of prompt-photon production, SHERPA can provide matrix elements for photon-plus-jet final states with up to three additional partons ($2 \to n$ processes with n from 2 to 5).

The proton structure was parameterised by the LO NNPDF2.3 [104] PDF set for PYTHIA and by the NLO CT10 [105] or NNLO NNPDF3.0 [73] PDF sets for SHERPA, depending on the generator version.

4.2.2 Initial- and final-state radiation

In processes containing charged or coloured objects, or both of them, in the initial or final states, photons and/or gluons can be radiated and generate large corrections to the overall topology of events. The starting $2 \to 2$ process topology can be changed to a $2 \to 3$, $2 \to 4$, and so on, final-state topologies by the inclusion of such corrections. This kind of hard emissions are increasingly important in determining the event structure due to the increase of the available energies.

Two traditional approaches exist to the modelling of perturbative corrections [90]. One is the matrix-element method, in which Feynman diagrams are calculated order by order. This is, in principle, the exact approach, and takes into account exact kinematics and the full interference and helicity structure. However, the calculations become increasingly difficult in higher orders, in particular for the loop graphs. The second possible approach is the parton-shower one, where an arbitrary number of branchings of one parton into two (or more) may be combined, to yield a description of multijet events, with no explicit upper limit on the number of partons involved. This approach relies on approximations of the matrix-element expressions.

The two approaches are complementary, and both are widely used. For simplicity and flexibility reasons, the parton-shower option is often the first choice. However, with the improvement of calculational techniques and the use of faster computers, Born-level calculations have been made to higher orders and have seen increasing use. The most desirable goal would be to have an approach which combined the best aspects of both kinds of calculations.

4.2.2.1 Parton showers

The separation of radiation into initial- and final-state showers is arbitrary, but certainly convenient. In general, the assignment of radiation to a given mother parton is a good approximation for an emission close to the direction of motion of that parton, but not for a wide-angle emission in between two jets, where interference terms are expected to be important [101].

In both initial- and final-state showers, the structure is given in terms of branchings $a \to bc$. Each of these processes is characterised by a splitting kernel $P_{\mathrm{a}\to\mathrm{bc}}(z)$. The branching rate is proportional to the integral $\int P_{\mathrm{a}\to\mathrm{bc}}(z)\,\mathrm{d}z$. Where z represents the energy sharing, with daughter b taking a fraction z and daughter c the remaining $1-z$ of the mother energy. Once formed, daughters b and c may branch subsequently, and so on.

Each parton is characterised by some virtuality scale Q^2, which gives an approximate sense of time ordering to the cascade. Different definitions of Q^2 are possible. In the case of PYTHIA, two distinct alternatives are implemented. In the initial-state shower, Q^2 values are gradually increasing as the hard scattering is approached, while Q^2 is decreasing in the final-state showers. The shower evolution is cut off at some lower scale Q_0, typically 1 GeV for QCD branchings. On the other hand, from above, a maximum scale Q_{max} is introduced, where the showers are matched to the hard interaction itself.

Final-state showers

Final-state showers are time-like, i.e. partons have $m^2 = E^2 - p^2 \geq 0$. As mentioned before, the choice of the evolution variable Q^2 of the cascade is not unique. In recent versions of PYTHIA, a p_{T}-ordered shower algorithm, with $Q^2 = p_{\mathrm{T}}^2 = z\,(1-z)\,m^2$ is implemented. Regardless of the exact definition of the ordering variable, the general strategy is the same: starting from some maximum scale Q^2_{max}, an original parton is evolved downwards in Q^2 until a branching occurs. The selected Q^2 value defines the p_{T} of the branching parton, for this particular choice of ordering variable. The z value obtained from the splitting kernel represents the parton energy division between the daughters. These daughters may evolve downwards, in this case with maximum virtuality already defined by the previous branching, and so on down to the Q_0 cut-off.

Initial-state showers

In contrast to final-state showers, initial-state ones are space-like, hence, in the sequence of branchings $a \to bc$ that lead up from the shower initiator to the hard interaction, particles a and b have $m^2 = E^2 - p^2 < 0$. Particle c, which does not participate in the hard scattering, may be on the mass shell, or have a time-like virtuality. In the latter case, a time-like shower will evolve off it. To first approximation, the evolution of the space-like main branch is characterised by the evolution variable Q^2, which is required to be strictly increasing along the shower, i.e. $Q^2_{\mathrm{b}} > Q^2_{\mathrm{a}}$.

The ISR and FSR shift around the kinematics of the original hard interaction. In deep inelastic scattering, this means that the x and Q^2 values that can be derived from the momentum of the scattered lepton do not automatically agree with the values originally picked. Effects of any original kinematics selection cuts are therefore smeared out, an unfortunate side-effect of the parton-shower approach.

The differences in the implementation of the parton showering approach between the main MC generators used in this thesis are summarised below:

– In PYTHIA, the initial- and final-state algorithms are based on a p_{T}-ordered evolution that uses $Q^2 = p_{\mathrm{T}}^2 = z\,(1-z)\,m^2$ as the final and $Q^2 = p_{\mathrm{T}}^2 = (-\,(1-z)\,m^2)$ as the initial evolution variable. The p_{T}-ordered shower leads to the correct behaviour of coherence effects[1], and no additional requirements are needed. The initial-state evolution, the multiple interactions and the final-state evolution are interleaved. In this context, some of that FSR gets to be associated with dipoles stretched between a final-state parton and the *hole* left by an initial-state one, which therefore now can take a recoil.

– SHERPA uses a shower generator based on virtuality ordered DGLAP parton evolution with superimposed angular-ordering constraints. The evolution variable is defined as $Q^2 = p^2 - m^2$. Colour coherence during evolution is taken into account by an explicit angular veto, which means that a branching is rejected if the opening angle of the emission is larger than the one of the previous branching.

4.2.3 Combining matrix elements and parton showers

The precise theoretical description of multijet final states is one of the most important challenges for simulating events at modern high-energy collider experiments [90]. To obtain reliable estimates of the SM backgrounds to experimental analyses, an accurate approach to this problem is needed by means of the best tools available. The underlying hard process is therefore accounted for by full matrix-element calculations, while the subsequent evolution and the conversion of hard partons into jets is modelled by QCD parton cascades and phenomenological hadronisation models.

The development of an event is determined by several scales which make difficult a separation to unambiguously disentangle the components that belong to the hard process and those that belong to the evolution of the hard parton. Given an n jet event of well separated partons, the jet structure is not changed when a further collinear or soft parton is emitted. However, an additional hard, large-angle parton emission will add an extra jet changing the n to an $n+1$ jet final state. Therefore, a merging scheme for matrix-element calculations and parton showers has to be defined, that determines on an event-by-event basis which of the two possibilities has to be followed. The primary goal is to avoid double counting by preventing events to appear multiple times, once for each possibility.

There are mainly two solutions for this problem that have been proposed and implemented in the past few years: the MLM scheme [106], which relies on a geometric analysis of the unconstrained radiation pattern in terms of cone jets, and the CKKW scheme or merging algorithm [107, 108], which employs an analytical reweighting of the matrix elements supplemented by a constrained parton-shower evolution. A common characteristic to all schemes is that sequences of tree-level multileg matrix elements with increasing

[1] In QCD showers, corrections to the LL picture, so-called coherence effects, lead to an ordering of subsequent emissions in terms of decreasing angles.

final-state multiplicity are merged with parton showers to yield a fully inclusive sample with no double counting and correct at LL accuracy.

Matching and merging NLO matrix elements and parton showers

Matching matrix elements and parton showers resolves the double-counting of the NLO corrections in the matrix-element calculation with the first parton-shower emission [91]. For the matching of NLO matrix elements, a variant of the MC@NLO [109] is used in SHERPA. The underlying basic idea is the realisation that parton showers organise the radiation pattern, and in particular the first emission, by identifying and factorising the singular soft and collinear limits of the emission matrix elements. On the other hand, multijet merging allows the combination of final states of increasing matrix-element-parton multiplicity, evolved by a parton shower, into an inclusive description. This enables prediction for higher jet rates at NLO or LO accuracy, depending on the order of the underlying matrix-element calculation, up to a certain maximum matrix-element-parton multiplicity. Higher jet numbers are accounted for by the parton shower off the highest-multiplicity matrix element.

The multijet-merging approach uses the notion of jets to classify emissions as either jet production or jet evolution, and to additively combine towers of exact matrix elements with increasing jet multiplicities into one inclusive sample. The separation scale of both regimes is defined as Q_{cut}, and emissions with $Q \geq Q_{\text{cut}}$ get accounted for by exact matrix elements, while radiation with $Q < Q_{\text{cut}}$ is described by the parton shower instead. The SHERPA merging algorithm for tree-level matrix elements, called MEPS@LO has been promoted to include matrix elements at NLO accuracy in QCD, called MEPS@NLO, and implemented in SHERPA [110].

4.2.4 Multiple parton interactions

The hardest partonic interaction in hadronic collisions may be accompanied by softer ones, all of them falling in the realm of pQCD. Experimental evidence for this assumption came from studies carried out by the CDF collaboration at the Fermilab Tevatron during Run I and Run II [111, 112]. It was shown that a correct description at the hadron level of particle multiplicities and jet activities can only be achieved by MC event generators which incorporate a model for multiple parton scatterings [90].

The basic underlying idea of this model is to postulate the probability distribution for the possible multiple scatterings that may occur. According to the model of ref. [113], this is given as a function of the non-diffractive cross section σ_{ND}, the hard perturbative cross section σ_{hard}, and the outgoing transverse momentum p_{T} in the scattering by

$$p\left(p_{\text{T,b}}\right) = f_{\text{c}} f\left(b\right) \frac{1}{\sigma_{\text{ND}}} \frac{\mathrm{d}\sigma_{\text{hard}}}{\mathrm{d}p_{\text{T}}}\,, \tag{4.1}$$

where the prefactors f_{c} (normalisation) and $f\left(b\right)$ (proton shape function) incorporate an

additional impact-parameter dependence of the distribution on an event-by-event basis. In the collinear factorisation approach, a phase-space cut on the hard matrix elements, e.g. a minimum jet resolution cut $Q_{\mathrm{cut}}^{\mathrm{resol}}$, has to be introduced to obtain a well-defined differential cross section in perturbation theory. Then, $\mathrm{d}\sigma_{\mathrm{hard}}/\mathrm{d}p_{\mathrm{T}}$ receives its dominant contributions from $2 \to 2$ processes and, therefore, the definition of p_{T} in equation 4.1 is unambiguous [90].

The multiple-interactions models used in PYTHIA and SHERPA, generate multiple scatterings according to the ideas from ref. [90] summarised above. However, each generator approach introduces differences in the relation of secondary interactions with parton showers and the hard scattering that deviate from the original formalism.

4.2.5 Hadronisation

Once the parton shower has concluded, there is a configuration of coloured partons at some low scale of the order of a few GeV in transverse momentum. In this low-energy regime, the coloured partons are transformed into colourless hadrons, a process called either hadronisation or fragmentation. Hadronisation is performed by phenomenological models that are usually based on some qualitative ideas on how the parton-to-hadron transition proceeds, like, e.g., local parton-hadron duality or infrared safety, defining the coarse properties of the model [90]. However, many of the important subtle details, which are often related to how flavour is created and distributed in the procedure are essentially undefined and subject to phenomenological parameters that must be fixed from extensive comparisons with data.

4.2.5.1 Beam remnants

In high-energy collisions with hadronic initial states, there is the problem that one or more final-state particles will be colour-connected to the hadron remnants. Then, the breakup of the initial hadron into partons has to be modelled adequately. This can be done in a way that only a minimal set of particles (quarks and diquarks) is produced to reconstruct the constituent flavour configuration of the initial hadron. After the colour connections are established, the complete set of final-state partons can be hadronised affecting energy flows and other similar properties of the event at hadron level.

The colour distribution in the hadron remnants cannot be extracted from first principles. Therefore, MC generators employ different phenomenological models.

4.2.5.2 Hadronisation models

As previously mentioned, the fragmentation process has yet to be understood from first principles, starting from the QCD Lagrangian. A number of different phenomenological models has been developed. Three main schools are usually distinguished [101], string

fragmentation (SF), independent fragmentation (IF) and cluster fragmentation (CF), but many variants and combinations are also used. The aim for each of these models is to provide internal consistency, a good representation of existing data, and a predictive power for properties not yet studied or results at higher energies.

In the following, the string- and cluster-fragmentation models are introduced as implemented in PYTHIA and SHERPA generators, respectively. In both models, the parton configurations coming from the parton showers, UE and beam remnants are cast into the form of colour-connected singlets, which will decay non-perturbatively [91]. These decays proceed by *popping* flavour/anti-flavour pairs and inserting them into the singlet structure, which then decays into more singlets with reduced masses. The only flavours[2] which are allowed to be produced in this way are the light u, d, and s quarks and, possibly diquarks[3] made from them. The differences between these models are in the logic in which the non-perturbative flavour production is implemented. In the string model, the singlets form coloured lines (strings) of the type $qgg...gg\overline{q}$, which decay from their ends into a hadron and a *shorter* string. The flavour necessary to form the hadron is compensated by the anti-flavour of the new string end. In contrast, in the cluster model, gluons decay non-perturbatively to form colour-neutral quark/anti-quark or quark/diquark clusters. Clusters are interpreted as massive hadron resonances and undergo binary decays, until the clusters formed are light enough to be hadrons.

String-fragmentation model

The string fragmentation model or *Lund model* [114] is the default for PYTHIA [101]. All current models are of a probabilistic and iterative nature. Hence, the fragmentation process is described in terms of simple underlying branchings, of the type jet $\rightarrow$ hadron + remainder-jet, string $\rightarrow$ hadron + remainder-string, and so on. The probabilistic rules for the production of new flavours and for the sharing of energy and momentum between the products are given at each branching.

Given an example system of a colour-singlet $q\overline{q}$ 2-jet event, as produced in e^+e^- annihilation. Here lattice QCD studies support a linear confinement picture (in the absence of dynamical quarks), i.e. the energy stored in the colour dipole field between a charge and an anticharge increases linearly with the separation between the charges, if the short-distance Coulomb term is neglected [101].

The assumption of linear confinement provides the starting point for the string model. As the q and $\overline{q}$ partons move apart from their common production vertex, the physical picture is similar to a colour flux tube[4] that is stretched between the q and the $\overline{q}$. The assumption of an uniform tube along its length leads to a confinement picture with linearly rising potential.

[2] Quarks and diquarks are commonly called *flavours*.

[3] Diquarks are hypothetical bound states of two quarks or two anti-quarks, forming a colour sextet or anti-sextet, which in the large-N_c limit is re-interpreted as a colour anti-triplet or triplet [91]. They carry baryonic quantum numbers, being baryons in this picture bound states of a quark and a diquark.

[4] The transverse dimensions of the tube are of typical hadronic sizes, roughly 1 fm.

As the q and $\overline{q}$ separate from each other, the potential energy stored in the string increases, and the string may break producing a new $q'\overline{q}'$ pair, so that the system splits into two colour-singlet systems, $q\overline{q}'$ and $q'\overline{q}$. If the invariant mass of any of these string pieces is large enough, further breaks may occur. In the Lund string model, the string breakup processes are assumed to proceed until only on-mass-shell hadrons remain, each hadron corresponding to a small piece of string with a quark in one end and an antiquark in the other.

The procedure through which quark-antiquark pairs $q'\overline{q}'$ which lead to string breakups are generated in the Lund model is based on the idea of quantum mechanical tunnelling, which leads to a flavour-independent Gaussian spectrum for the p_{T} of $q'\overline{q}'$ pairs. The total p_{T} of a hadron is made up out of the p_{T} contributions from the quark and antiquark that together form the hadron. Contributions of very soft perturbative gluon emissions may also effectively be included in this description.

A schematic picture of the production of a multihadronic final state in e^+e^- annihilation according to the string model is shown in figure 4.2.

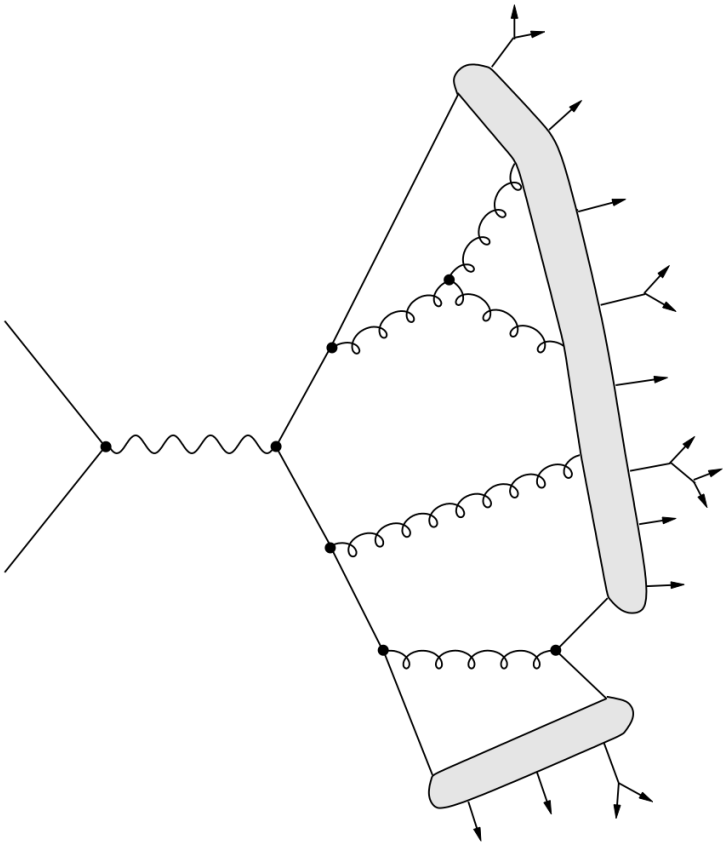


Figure 4.2: Parton shower with string hadronisation model for $e^+e^- \to$ hadrons. From ref. [61].

Cluster-fragmentation model

Since version 1.1, SHERPA employs its own module AHADIC, which implements a cluster-hadronisation model as described in ref. [115, 116]. The basic assumption underlying these models is local parton-hadron duality, which implies that quantum numbers on the hadron level follow very closely the flow of quantum numbers on the parton level. In this framework, the mass spectrum of the emerging colour-neutral clusters is dominated by typical hadron masses or masses slightly above [90]. These clusters are typically thought as some kind of *hadron matter*, carrying the flavour and momentum quantum numbers of hadrons. For this reason, light clusters are translated into hadrons, or, when they are too heavy, they are treated as unknown heavy resonances that decay further into lighter ones.

In AHADIC, the hadronisation of quarks and gluons starts by boosting and scaling the four-momenta of all partons such that all of them reside on constituent-mass shells, whereas gluons remain massless. In addition, diquarks states emerge, which effectively carry baryonic quantum numbers.

Since hadrons are made of flavour constituents only, in cluster models the gluons must decay. In AHADIC this is facilitated in a dipole framework, and there will be a colour-connected spectator to compensate the recoil when the massless gluon decays into massive flavours. In these non-perturbative forced decays, the transverse momentum measure p_{T} is bounded from above by a parameter $p_{\mathrm{T}}^{\mathrm{max}}$ which is typically of the order of the parton-shower cut-off scale. As a result of these decays, quark/anti-quark and diquark/anti-diquark pairs are produced.

For the production of flavours, the available phase space leads to kinematic enhancements of lighter-flavour pairs depending on the mass of the dipole. It is supplemented by constant weight factors, $\mathcal{P}_{\mathrm{q}}$, also used in cluster decays to hadrons. After all gluons have decayed, those flavours which are colour-connected are then merged into colour-neutral clusters. Diquarks are also allowed as cluster constituents, so that baryonic quantum numbers can be produced at every stage of the hadronisation process.

At every step, all clusters are checked whether they fall into a regime where the mass of the cluster is of the order of the mass of the heaviest hadron with a flavour wave component identical to the flavour content of the cluster or a regime where the mass of the cluster is lighter than the summed mass of the two heaviest hadrons it can decay into by adding a flavour-antiflavour pair to its flavour content, or both. If a cluster is outside both regimes, it will decay further into two clusters; whereas if it falls into one or both transition regimes, the probabilities for the allowed transitions into one or two clusters are calculated.

Figure 4.3 shows the cluster hadronisation of the same parton shower as in figure 4.2.

4.2.5.3 Decays of unstable particles

A large number of the particles produced in the fragmentation process are unstable and may decay into stable or almost stable ones. Hence, the particles included must have

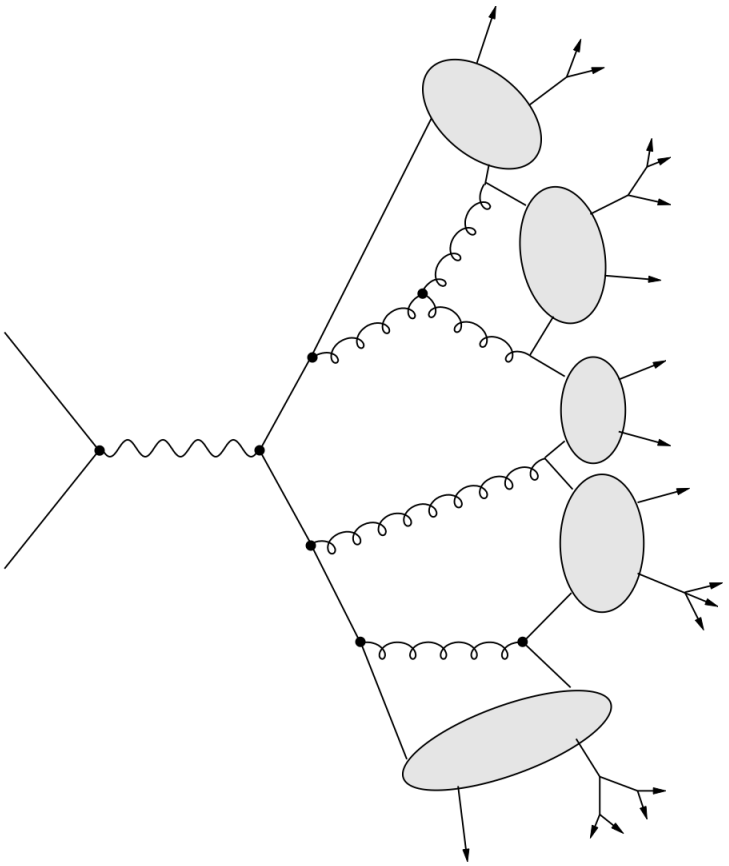


Figure 4.3: Parton shower with cluster hadronisation model for $e^+e^- \rightarrow$ hadrons. From ref. [61].

their proper mass distributions and decay properties. However, while a lot of experimental information is available, there is still very much which is missing. In the case of charm mesons, to obtain a complete and consistent set of decay channels, it is necessary to combine measurements of exclusive branching ratios with inclusive multiplicity distributions [101]. For bottom mesons, less information is known, and for some B baryons, a rather simple phase-space type of generator has been used for hadronic decays.

It is normally assumed that decay products are distributed according to phase space, with no dependence on the dynamics in their relative distribution [101]. Nevertheless, additional assumptions are needed in many cases, e.g. for semileptonic decays of charm and bottom hadrons one needs to include the proper weak matrix elements. Moreover, particles can also be produced polarised such that their decay products are non-isotropically distributed.

This way of treating decays makes use of a set of tables where branching ratios and decay modes are stored. It encompasses leptons and all hadrons made from d, u, s, c and b quarks. The decay products can be hadrons, leptons, photons or other heavy states allowed to decay to partonic states.

4.2.6 Simulation of detector effects

As previously mentioned, the behaviour of detectors is simulated with toolkits such as GEANT4. Another important feature that needs to be accounted is the simulation of pile-up. Both topics are experiment dependent, and in the following, they are addressed in the context of the LHC and the ATLAS Experiment in particular.

4.2.6.1 GEANT4

The standard simulation strategy of ATLAS is based on GEANT4. This toolkit provides a diverse and wide set of software components that can be used for a variety of settings. From simple studies of basic phenomena and geometries to full-scale detector simulations for experiments at the LHC and other facilities. Details on the internal design and functioning are described in this section.

Design overview

GEANT4 is driven by the software needs of modern experiments. Typical components of a software system such as the event generation, detector simulation, reconstruction and analysis can be used separately or in combination. This toolkit was designed as the basis for the simulation. Thus it was required to have well-defined interfaces to the other components and to provide parts to be used by them. Additional requirements are that it is modular and flexible, and that the implementation of physics is transparent and open to user validation. The modular and hierarchical structure for the toolkit is represented in figure 4.4, where subdomains are linked in an unidirectional flow of dependencies. The key parts of the simulation of the passage of particles through matter are: geometry and materials; particle interaction in matter; tracking management; digitisation and hit management; event and track management; visualisation and visualisation framework; user interface. These domains led to the creation of class categories with coherent interfaces.

The categories (packages) shown at the bottom of the diagram in figure 4.4 are used by higher categories and provide the foundation of the toolkit. These include the category *global* covering the system of units, constants, numerics and random number handling; the categories *materials*, *particles*, *graphical representations* and *geometry* including the volumes for detector description and the navigation in the geometry model; and *intercoms* which provides a way of interacting with GEANT4 through the user interface and also a way of communicating between modules.

Above these reside categories required to describe the tracking of particles and the physical processes they undergo. The *track* category contains classes for tracks and steps, used by *processes*, which contains implementations of models of physical interactions. One of such processes, *transportation*, handles the transport of particles in the geometry model. All these processes may be invoked by the *tracking* category, which manages the contributions to the evolution of the state of a track and undertakes to provide information in sensitive volumes for hits and digitisation.

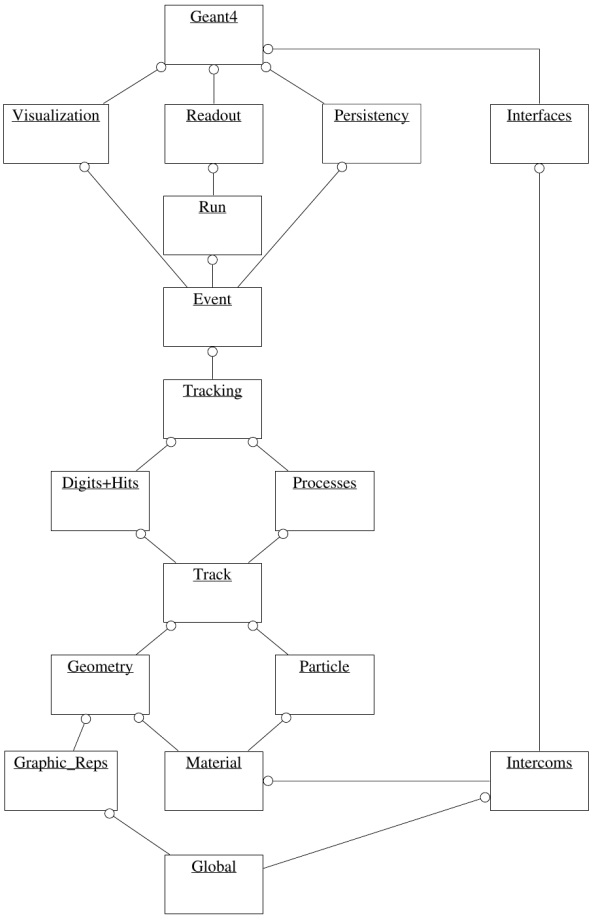


Figure 4.4: The Top Level Category Diagram of the GEANT4 toolkit. The open circle on the joining line represents an using relationship; the category at the circle end uses the adjoined category. From ref. [103].

Over these the *event* category manages events in terms of their tracks and *run* manages collections of events that share a common beam and detector implementation. The *readout* category allows the handling of pile-up.

Finally, capabilities that use all of the above and connect to facilities outside the toolkit are provided by the *visualisation*, *persistency* and *interface* categories.

Openness and transparency

One of the most important design goals of GEANT4 was to make the design and implementation of the physics open and transparent. The use of object-oriented technology allows to establish a clear and customisable correspondence between particles and processes and offers a choice of models for each process. The result is a highly granular implementation, each component of which can be inspected at the source-code level.

GEANT4 has been, and continues to be developed and maintained by a number of

cooperating institutes, HEP experiments and universities that form an international collaboration. It is built on their experience on the field of MC simulation of physics detectors and physical processes.

4.2.6.2 Simulation of pile-up in the ATLAS Experiment

The high luminosity of the LHC results in a significant background to interesting events coming from the high number of inelastic pp interactions in the same and neighbouring bunch crossings. As described in section 2.1.1, due to the increase of pp collisions per bunch crossing during the Run II, understanding and modelling this background is critical for performing analyses in ATLAS. There are five components of the pile-up background [117]:

1. In-time pile-up: additional pp collisions occurring in the same bunch-crossing as the collision of interest.
2. Out-of-time pile-up: additional pp collisions occurring in bunch-crossings just before and after the collision of interest.
3. Cavern background: the gas of neutrons and photons inundating the cavern during a typical run of the LHC. These mostly contribute with random hits in the muon system.
4. Beam halo events: the effect of protons from a bunch scraping against an up-stream collimator. The scraping results in sprays of muons running approximately parallel to the beam-line.
5. Beam gas events: collisions between the proton bunch and residual gas inside the beam-pipe. They generally occur off-center in the detector.

There are two basic approaches to model these backgrounds. In the first, each of the previous components is simulated individually. Once this is done, they are added on top of signal events prior to the conversion of energy deposits to detector signals in an step which is known as *digitisation*. Digitisation CPU requirements are demanding, and increase proportionally to the number of soft collisions. The *overlay* technique is an alternative method which attempts to simulate pile-up using specially-collected data or simulated events [118]. Compared to the current digitisation, simulated pile-up events are pre-mixed in an independent step and hard-scatter events are then overlaid on the merged background. The CPU requirements and the stress on storage are much lower compared to those of the digitisation method. This alternative technique is promising and will be especially useful for the future LHC upgrade scenarios including as many as 200 pp collisions per bunch-crossing.

4.3 Monte Carlo samples

For the analyses presented in this thesis, samples of MC events were generated to study the characteristics of signal events. The MC samples were also used to determine the correction factors necessary to obtain the measured cross sections and the predicted cross sections at particle level. In addition, MC samples were used to estimate hadronisation and UE corrections for the NLO and NNLO pQCD calculations.

Simulated signal samples

Samples of prompt-photon events were generated using the programs PYTHIA 8.186 [102] and SHERPA 2.1.1 [90]. In both programs, the event generation was performed using tree-level matrix elements, with the inclusion of initial- and final-state parton showers. Fragmentation into hadrons is described by the Lund string model [114] in the case of PYTHIA, and by a modified version of the cluster model [116] in the case of SHERPA. The proton structure was parameterised by the LO NNPDF2.3 [104] PDF set for PYTHIA and by the NLO CT10 [105] PDF set for SHERPA. All samples include a simulation of the UE with parameter values set according to the ATLAS 2014 tune[5] series (A14 tune) for PYTHIA [119] or to the tune developed by the authors of SHERPA for use in conjunction with the NLO CT10 PDF set. The PYTHIA simulation of the signal includes LO photon-plus-jet events from direct processes (the hard subprocess $qg \to q\gamma$ and $q\overline{q} \to g\gamma$, called the *hard* component) and photon bremsstrahlung in LO QCD dijet events (called the *bremsstrahlung* component). The bremsstrahlung component is modelled by final-state QED radiation arising from calculations to all $2 \to 2$ QCD processes. The SHERPA samples were generated with LO matrix elements for photon-plus-jet final states with up to three additional partons ($2 \to n$ processes with n from 2 to 5); the matrix elements were merged with the SHERPA parton shower using the ME+PS@LO prescription [120]. The bremsstrahlung component is accounted for in SHERPA through the matrix-element $2 \to n$ processes with $n \geq 3$. In the generation of the SHERPA samples, a requirement on the photon isolation at the matrix-element level is imposed using Frixione's criterion defined in ref. [89] and introduced in section 3.4.1.1. This criterion, requires the total transverse energy inside a cone of size r_{d} in the $\eta - \phi$ plane around the generated final-state photon, excluding the photon itself, to be below a certain threshold $E_{\mathrm{T}}^{\mathrm{max}}(r_{\mathrm{d}})$. The parameters for the threshold were chosen to be $R_{\mathrm{d}} = 0.3$, $n = 2$, and $\epsilon_{\mathrm{d}} = 0.025$. As described in section 3.4.1.1, this criterion is applied to avoid divergences in the matrix elements when the photon is collinear with a parton.

All the samples of generated events were passed through the GEANT4-based ATLAS detector- and trigger-simulation programs [121]. These samples were reconstructed and analysed by the same program chain as the data. The samples used for the inclusive isolated-photon analysis with $36\,\mathrm{fb}^{-1}$ are listed in table 4.1. For the inclusive isolated-photon and photon-plus-jet analyses with the full Run II dataset, the MC samples were generated in different *campaigns* to match better the running conditions for each of the

[5] A tune is a set of parton-shower-generator parameter values.

different data-taking periods, namely, $2015-2016$, 2017, and 2018. Three campaigns (a, d and e) were used, and the mean values for the cross sections and filter efficiencies and the total number of generated events are listed in table 4.2.

Generator	E_{T}^{γ} **range** [GeV]	**Number of events**	**Cross section [nb]**	**Filter efficiency**	**Filter type**
PYTHIA	$70-140$	2743800	1.0621E+05	3.9049E-05	
	$140-280$	996400	6.7046E+03	4.8927E-05	
	$280-500$	967199	3.4426E+02	5.6152E-05	
	$500-800$	994200	2.3729E+01	5.2447E-05	
	$800-1000$	92000	2.2869E+00	3.4994E-05	
	$1000-1500$	99500	7.0203E-01	3.6900E-05	
	$1500-2000$	98300	3.9345E-02	3.6688E-05	
	$2000-2500$	94700	3.3277E-03	3.3896E-05	
	$2500-3000$	98300	3.2495E-04	2.9921E-05	
	$3000-$ inf	97700	3.1480E-05	2.8355E-05	
SHERPA 2.1.1	$70-140$	1999200	3.1290E+00	3.9960E-01	CVetoBVeto
	$70-140$	1993400	3.1329E+00	4.8199E-01	CFilterBVeto
	$70-140$	499400	3.1352E+00	1.1728E-01	BFilter
	$140-280$	9966800	2.4741E-01	4.1049E-01	CVetoBVeto
	$140-280$	1993200	2.4739E-01	5.0284E-01	CFilterBVeto
	$140-280$	997600	2.4937E-01	1.2874E-01	BFilter
	$280-500$	997200	1.3648E-02	4.1834E-01	CVetoBVeto
	$280-500$	500000	1.3617E-02	4.7349E-01	CFilterBVeto
	$280-500$	99800	1.3874E-02	1.4069E-01	BFilter
	$500-1000$	200000	9.2334E-04	3.8226E-01	CVetoBVeto
	$500-1000$	100000	9.2185E-04	4.7148E-01	CFilterBVeto
	$500-1000$	20000	9.3819E-04	1.4811E-01	BFilter
	$1000-2000$	100000	1.8432E-05	3.7036E-01	CVetoBVeto
	$1000-2000$	50000	1.8388E-05	4.6691E-01	CFilterBVeto
	$1000-2000$	10000	1.9046E-05	1.5751E-01	BFilter
	$2000-4000$	101000	7.9682E-08	3.6729E-01	CVetoBVeto
	$2000-4000$	50500	8.0331E-08	4.6642E-01	CFilterBVeto
	$2000-4000$	10000	8.2035E-08	1.6920E-01	BFilter

Table 4.1: Information relative to the MC samples used for the generation of signal events in the inclusive isolated-photon analysis at $\sqrt{s} = 13$ TeV using $36\,\mathrm{fb}^{-1}$ of ATLAS data.

MC Samples for theoretical predictions

The program SHERPA 2.2.2 was used in order to compute NLO pQCD predictions. This version of SHERPA consistently combines parton-level calculations of $\gamma+(1,2)$-jet events at NLO and $\gamma+(3,4)$-jet events at LO in pQCD [91,122,123] supplemented with a parton shower [124] while avoiding double-counting effects [110]. The samples of events generated

Generator	E_{T}^{γ} **range** [GeV]	**Number of events**	**Cross section [nb]**	**Filter efficiency**	**Filter type**
PYTHIA	140 − 280	12594050	6.7046E+03	4.9477E-05	
	280 − 500	12528500	3.4426E+02	5.5860E-05	
	500 − 800	12481300	2.3729E+01	5.2918E-05	
	800 − 1000	1551700	2.2869E+00	3.5012E-05	
	1000 − 1500	1521600	7.0203E-01	3.7058E-05	
	1500 − 2000	1561000	3.9345E-02	3.7124E-05	
	2000 − 2500	365700	3.3277E-03	3.4292E-05	
	2500 − 3000	367200	3.2495E-04	3.0272E-05	
	3000 − inf	355500	3.1480E-05	2.8649E-05	
SHERPA 2.1.1	140 − 280	39309000	2.4730E-01	3.9258E-01	CVetoBVeto
	140 − 280	23673000	2.4729E-01	4.7798E-01	CFilterBVeto
	140 − 280	11779000	2.4929E-01	1.2902E-01	BFilter
	280 − 500	14940000	1.3635E-02	3.8650E-01	CVetoBVeto
	280 − 500	5971000	1.3639E-02	4.7393E-01	CFilterBVeto
	280 − 500	1168000	1.3872E-02	1.4016E-01	BFilter
	500 − 1000	3146000	9.2510E-04	3.8004E-01	CVetoBVeto
	500 − 1000	2367000	9.2404E-04	4.7146E-01	CFilterBVeto
	500 − 1000	479000	9.4462E-04	1.4921E-01	BFilter
	1000 − 2000	1965000	1.8485E-05	3.7085E-01	CVetoBVeto
	1000 − 2000	990000	1.8471E-05	4.7111E-01	CFilterBVeto
	1000 − 2000	280000	1.8963E-05	1.6142E-01	BFilter
	2000 − 4000	399000	7.9543E-08	3.6837E-01	CVetoBVeto
	2000 − 4000	599000	7.9553E-08	4.6613E-01	CFilterBVeto
	2000 − 4000	50000	8.2641E-08	1.6751E-01	BFilter
	4000 − inf	12000	2.5115E-12	4.2234E-01	CVetoBVeto
	4000 − inf	12000	2.5136E-12	4.2551E-01	CFilterBVeto
	4000 − inf	12000	2.6003E-12	1.5193E-01	BFilter

Table 4.2: Information relative to the MC samples used for the generation of signal events in the inclusive isolated-photon and photon-plus-jet analyses at $\sqrt{s} = 13$ TeV using the full Run II ATLAS luminosity. The mean values of the cross sections and filter efficiencies of the three campaigns (a, d and e) and the total number of generated events are shown.

with SHERPA 2.2.2 differ from those from SHERPA 2.1.1 described previously. The former include the higher-order virtual corrections for $\gamma + (1, 2)$-jet processes. Photon isolation at matrix-element level is applied by using Frixione's criterion with $R_{\mathrm{d}} = 0.1$, $n = 2$, and $\epsilon_{\mathrm{d}} = 0.1$. The parameters are chosen to be as loose as possible to minimise a bias from the application of the photon isolation requirement at particle level. Since Frixione's isolation criterion requires the isolation energy upper limit to be exactly zero at $r_{\mathrm{d}} = 0$, the criterion cannot be strictly looser than any non-zero photon isolation requirement at particle or detector level for all $r_{\mathrm{d}} < R_{\mathrm{d}}$. Dynamic factorisation and renormalisation

scales are adopted (E_{T}^{γ}) as well as a dynamical merging scale with $\overline{Q}_{\mathrm{cut}} = 20$ GeV [125]. The strong coupling constant is set to $\alpha_{\mathrm{S}}(M_{\mathrm{Z}}^2) = 0.118$. For the EM coupling, the low-energy limit of 1/137.036 is used. Fragmentation into hadrons and simulation of the UE are performed using the same models as for the SHERPA 2.1.1. The NNLO NNPDF3.0 PDF set [73] is used in conjunction with the corresponding SHERPA tuning.

The samples are listed in table 4.3 for the inclusive isolated-photon analysis with $36\,\mathrm{fb}^{-1}$. For the analyses based on the full Run II data, the samples are listed in table 4.4.

Generator	E_{T}^{γ} **range** [GeV]	**Number of events**	**Cross section [nb]**	**Filter efficiency**
SHERPA 2.2.2	70 – 140	9879000	4.5259E+00	1.0000E+00
	140 – 280	9862600	3.7595E-01	1.0000E+00
	280 – 500	2970000	2.1867E-02	1.0000E+00
	500 – 1000	989400	1.4638E-03	1.0000E+00
	1000 – inf	491580	2.9729E-05	1.0000E+00

Table 4.3: Information relative to the MC samples used for the calculation of theory predictions for the inclusive isolated-photon analysis at $\sqrt{s} = 13$ TeV using $36\,\mathrm{fb}^{-1}$ of ATLAS data.

Generator	E_{T}^{γ} **range** [GeV]	**Number of events**	**Cross section [nb]**	**Filter efficiency**
SHERPA 2.2.2	140 – 280	42916800	3.7606E-01	1.0000E+00
	280 – 500	12957500	2.1860E-02	1.0000E+00
	500 – 1000	8297100	1.4629E-03	1.0000E+00
	1000 – inf	6080720	2.9865E-05	1.0000E+00

Table 4.4: Information relative to the MC samples used for the calculation of theory predictions for the inclusive isolated-photon and photon-plus-jet analyses at $\sqrt{s} = 13$ TeV using the full Run II ATLAS luminosity. The mean values of the cross sections and filter efficiencies of the three campaigns (a, d and e) and the total number of generated events are shown.

Simulated background samples

The main background to isolated-photon events arises from jets misidentified as photons. This background is subtracted using a data-driven technique, and no MC sample is used to simulate this background. The background from electrons or positrons misidentified as photons is evaluated using MC samples generated with the program SHERPA 2.2.1 [91, 110, 122–124, 126]. The $pp \to Z/\gamma^* \to e^+e^- + X$ and $pp \to W \to e\nu + X$ processes were generated with matrix elements calculated for up to two additional partons at NLO and up to four partons at LO. The NNLO NNPDF3.0 PDF set [73] was used in conjunction with a dedicated tune developed by the SHERPA authors.

Simulation of pile-up

Pile-up from additional pp collisions in the same and neighbouring bunch crossings was simulated by overlapping each MC event with a variable number of simulated inelastic pp collisions generated using PYTHIA 8.186 with the ATLAS set of tuned parameters for minimum-bias events (A2 tune) [127] and the MSTW2008LO PDF set [128]. The MC events are weighted (*pile-up reweighting*) to reproduce the distribution of the average number of interactions per bunch crossing observed in the data.

Chapter 5

Photon reconstruction, identification and calibration

In this chapter, the most important aspects of the trigger reconstruction and identification of photons and electrons based on ref. [39] are introduced. The latest photon and electron offline reconstruction methods are described, as well as the most relevant characteristics of the photon calibration and identification procedures [129–131]. A brief description of the isolation requirement from an experimental point of view is provided at the end of the chapter, together with some additional requirements that are applied to simulated events at analysis level.

5.1 Trigger reconstruction and identification of electrons and photons

Photon and electron reconstruction at the HLT stage is performed on each EM RoI provided by the L1, which satisfies E_{T} and isolation requirements as specified by the trigger menu. It proceeds in a series of sequential steps. In the HLT, fast algorithms are executed first, allowing precision algorithms (which closely reproduce the offline algorithms and require more CPU time) to run at a reduced rate later in the trigger sequence.

Fast algorithms are executed using calorimeter and ID information within the RoI to perform the initial selection and identification of the photon and electron candidates and achieve early background rejection.

If a particle candidate satisfies the criteria defined for the fast selection, the precision algorithms are executed in the HLT, where access to detector information outside the RoI is possible. These precision online algorithms are similar to their offline counterparts, with the following exceptions: the bremsstrahlung-aware re-fit for electron tracks and electron

and photon dynamic, variable-size topoclusters[1] [132] are not used online; photon candidates are identified using only the calorimeter information online; the online algorithms use $\langle\mu\rangle$ to assess pile-up, while the number of primary vertices (N_{PV}) is used offline. In addition to the above, some cell-energy-level corrections are not available online, such as the correction for transient changes in LAr high-voltage, or differ in implementation, such as the bunch crossing position-dependent pile-up correction.

The offline electron and photon reconstruction is described in detail in section 5.2.

5.1.1 Photon and electron triggers at L1

The L1Calo trigger for photons and electrons uses calorimeter information in the central ($|\eta| < 2.5$) region to build an EM RoI. A sliding-window algorithm is used, with a window consisting of 4×4 trigger towers with granularity 0.1×0.1 in η and ϕ, longitudinally segmented into EM and hadronic towers. Within the window, the algorithm uses the maximum E_{T} from the four possible pairs of nearest-neighbour EM towers in a 2×2 central region; this is used for EM transverse energy reconstruction. The energy of the trigger towers is calibrated at the EM energy scale. This EM scale is not the same as the one used in the offline reconstruction, which can lead to trigger inefficiencies relative to offline reconstruction. The schematic view of the trigger towers used as input to the L1Calo trigger algorithms is shown in figure 5.1.

A nominal transverse energy threshold is applied (e.g. $E_{\mathrm{T}} > 22$ GeV, denoted by a trigger name EM22). The threshold can also be η-dependent, due to the energy scale depending on η, within the granularity of 0.1. Typical variations of the threshold are -2 to $+3$ GeV relative to the nominal value. Optionally, a selection to reject hadronic activity can be applied: candidate electrons and photons are rejected if the sum of transverse energies in hadronic towers matched to the 2×2 central region is at least 1 GeV and exceeds $E_{\mathrm{T}}/23 - 0.2$ GeV. Finally, an EM isolation requirement can also be applied: candidate photons and electrons are rejected if the sum of transverse energies in the 12 towers surrounding the 2×2 central region in the EM layer is at least 2 GeV and exceeds $E_{\mathrm{T}}/8 - 1.8$ GeV. No requirements based on hadronic activity or EM isolation are applied above 50 GeV of E_{T} reconstructed at L1.

5.1.2 HLT photon reconstruction and identification for *pp* data-taking

The HLT fast algorithm reconstructs clusters from the calorimeter cells within the EM RoIs identified by L1. To minimise the HLT latency, the fast algorithm uses only the second layer of the EM calorimeter to find the cell with the largest deposited transverse

[1] Topoclusters are dynamic, variable-size clusters of energy deposits measured in topologically connected EM and hadronic calorimeter cells.

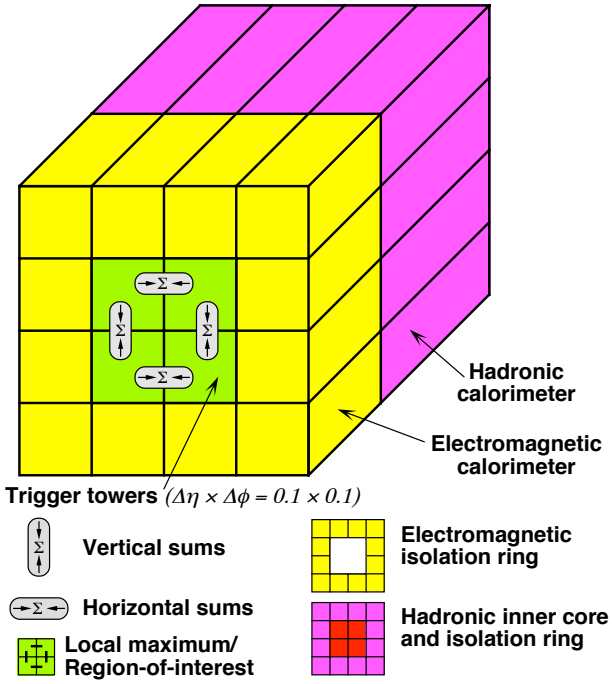


Figure 5.1: Schematic view of the trigger towers used as input to the L1Calo trigger algorithms. From ref. [38].

energy in the RoI. This cell is referred to as the *pre-seed*. Nine possible 3×7 windows ($\Delta\eta \times \Delta\phi = 0.075 \times 0.175$) around the *pre-seed* cell are checked to ensure that the local maximum, the cluster seed, is found. The final cluster position is obtained by calculating the energy-weighted average cell positions inside a 3×7 window centred on the cluster seed. To compute the accumulated energy in all EM calorimeter layers, a cluster size of 3×7 is used in the barrel and a cluster size of 5×5 in the end-caps. Several corrections, based on the offline reconstruction algorithms, are used at the fast algorithm step to improve the resolution of the cluster position and energy. Only selections on the cluster E_{T} and shower shape parameters[2] R_{had}, R_{η} and E_{ratio}, which have a good discrimination power between background and signal, are applied.

In the precision step, offline-like algorithms are used for the reconstruction of calorimeter quantities. After retrieving the cell information from regions slightly larger than the RoI, the precision HLT reconstruction uses the offline sliding-window algorithm to construct clusters [133]. The energy of the clusters is calibrated using a multivariate technique such that the response of the calorimeter layers is corrected in data and simulation [130]. The online photon identification relies on the same cluster shower shapes that are used in the offline algorithms (see table table 5.1), and three identification Working Points (WP) are defined: *loose*, *medium* and *tight*. The *medium* is used only in the HLT.

[2]The different shower shapes used for identification criteria are described in table 5.1.

An optional requirement on calorimeter-only isolation in photon triggers uses topoclusters, similar to the offline isolation calculation [131]. Full-scan topocluster reconstruction is needed to compute the energy density of the event on-the-fly in the HLT; this is then used to subtract ambient noise in the isolation cone. An isolation cone of size $\Delta R = 0.2\,(0.4)$ around the photon candidate is used for the very loose (tight) isolation requirement, denoted by *icalovloose* (*icalovtight*). If the ratio of the transverse energy in the topoclusters to the transverse energy of the photon candidate is less than 10% (3%, with an energy offset of 2.45 GeV), then the photon is considered isolated for *icalovloose* (*icalovtight*) by the HLT. These isolation criteria are over 98% efficient for offline photons satisfying *tight* isolation. The full-scan topocluster reconstruction is executed only once per event at the end of the trigger sequence (as it is very CPU intense) and is common to all isolated triggers and all trigger signatures.

5.2 Photon reconstruction

Since the analyses presented in this thesis were performed in different periods of time, slightly different reconstruction procedures were applied. The inclusive photon analysis using $36\,\text{fb}^{-1}$ (see chapter 7) and the in situ jet energy scale (JES) calibration using γ+jet events (see chapter 9) were performed using the reconstruction methods reported in ref. [129]. Those analyses developed at a latter stage made use of the updated reconstruction methods published in ref. [131], namely the inclusive photon analysis using $139\,\text{fb}^{-1}$ (see chapter 8) and the photon plus jet analysis using $139\,\text{fb}^{-1}$ (see chapter 10). The sliding-window algorithm used in ATLAS for the reconstruction of fixed-size clusters of calorimeter cells [129] was replaced for the updated reconstruction methods [131], improving the offline reconstruction of electrons and photons to use dynamic, variable-size clusters, called superclusters. While the use of fixed-size clusters provides a linear energy response and good stability as a function of pile-up, dynamic clusters can change in size as needed to recover energy from bremsstrahlung photons or from electrons from photon conversions. Changes to the photon conversion reconstruction were also applied to improve the reconstruction efficiency of double-track Si conversions and to reduce the fraction of unconverted photons mistakenly reconstructed as single- or double-track TRT conversions (see section 5.2.5).

An electron is defined as an object consisting of a cluster built from energy deposits in the calorimeter (supercluster) and a matched track (or tracks). A converted photon is a cluster matched to a conversion vertex (or vertices), and an unconverted photon, is a cluster matched to neither an electron track nor a conversion vertex. The fraction of converted photons observed in the inclusive-isolated photon analysis using the Run II dataset is shown in appendix A.

For the reconstruction of electrons and photons with $|\eta| < 2.5$, the algorithm first prepares the tracks and clusters it will use. Clusters of energy deposits measured in topologically connected EM and hadronic calorimeter cells are selected. These clusters are

matched to ID tracks, which are re-fitted accounting for bremsstrahlung. The algorithm also builds conversion vertices and matches them to the selected topoclusters. The electron and photon superclusters can then be built separately using the matched clusters as input. After applying initial-position corrections and energy calibrations to the resulting superclusters, the supercluster-building algorithm matches tracks to the electron superclusters and conversion vertices to the photon superclusters. Then, the final objects to be used for analyses are built, their energies are calibrated and discriminating variables used to separate electrons or photons from background are added.

5.2.1 Topocluster reconstruction

In the first step of the topocluster reconstruction algorithm, protoclusters in the EM and hadronic calorimeters are formed using a set of noise thresholds in which the cell which initiates the cluster is required to have a certain significance $|\zeta_{\text{cell}}^{\text{EM}}|$, where

$$\zeta_{\text{cell}}^{\text{EM}} = \frac{E_{\text{cell}}^{\text{EM}}}{\sigma_{\text{noise,cell}}^{\text{EM}}},$$

and $E_{\text{cell}}^{\text{EM}}$ is the cell energy at the EM scale[3] and $\sigma_{\text{noise,cell}}^{\text{EM}}$ is the expected cell noise. The expected cell noise includes the known electronic noise and an estimate of pile-up noise corresponding to the average instantaneous luminosity expected for Run II. The formed protoclusters collect neighbouring cells with significance $|\zeta_{\text{cell}}^{\text{EM}}| \geq 2$, and these cells become seeds in the next iteration, collecting each of their neighbours in the protocluster. In the presence of negative-energy cells induced by calorimeter noise, the algorithm uses $|\zeta_{\text{cell}}^{\text{EM}}|$ instead of $\zeta_{\text{cell}}^{\text{EM}}$ to avoid biasing the cluster energy upwards. Protoclusters with two or more local maxima[4] are split into separate clusters.

The electron and photon reconstruction starts with topoclusters but only uses the energy from the cells in the EM calorimeter, except in the transition region of $1.37 < |\eta| < 1.63$, where the energy measured in the presampler and the tile calorimeter is also used. The fraction of the EM energy over the total cluster energy is defined as f_{EM}. Only clusters with EM energy greater than 400 MeV are considered, and a preselection requirement of $f_{\text{EM}} > 0.5$ was chosen for the initial topoclusters. These clusters are referred to as EM topoclusters in the following.

[3] The EM scale is the basic signal scale accounting correctly for the energy deposited in the calorimeter by electromagnetic showers.

[4] A cell is considered a local maximum when it has $E_{\text{cell}}^{\text{EM}} > 500$ MeV, at least four neighbours, and when none of them has a larger signal.

5.2.2 Track reconstruction, track-cluster matching and photon conversion reconstruction

Tracks reconstructed in the ID are loosely matched to the fixed-size clusters. These seed clusters are required to pass loose shower shape requirements in hadronic leakage and energy distribution in η in order to be used to create ROIs, within which standard pattern recognition is first performed. If the standard pattern recognition [134] fails for a silicon track seed (a set of silicon detector hits used to start a track) within a ROI, a modified pattern recognition based on a Kalman filter formalism [135] is used. Then, track candidates are fitted with the global χ^2 fitter [136], allowing for additional energy loss when the standard track fit fails. In addition, tracks with silicon hits loosely matched to fixed-size clusters are re-fitted using a Gaussian Sum Filter (GSF) algorithm [137]. The loosely matched, re-fitted tracks are then matched to the EM topoclusters, extrapolating the track from the perigee to the second layer of the calorimeter. Either the measured track momentum or the rescaled magnitude of the momentum are used to match the cluster energy [131].

When multiple tracks are matched to a cluster, those with hits in the pixel detector are preferred, then tracks with hits in the SCT but not in the pixel detector. Within each category, tracks with a better ΔR match to the cluster in the second layer of the calorimeter are preferred unless the differences are less than 0.01. The tracks are then extrapolated through the calorimeter, first with the track momentum rescaled to the cluster energy and successively without rescaling. If the ΔR differences between both extrapolations are small, the track with more pixel hits is preferred. Finally, the highest-ranked track is used to define the reconstructed electron properties.

Tracks loosely matched to fixed-size clusters serve as input to the reconstruction of the conversion vertex. Tracks with silicon hits (Si tracks) and tracks reconstructed only in the TRT (TRT tracks) are used. Two-track conversion vertices are reconstructed from two opposite-charge tracks forming a vertex consistent with that of a massless particle, while single-track vertices are essentially tracks without hits in the innermost sensitive layers. To increase the converted-photon purity, the tracks used to build conversion vertices are required to have a high probability to be electron tracks.

The conversion vertices are then matched to the EM topoclusters. If there are multiple conversion vertices matched to a cluster, double-track conversions with two Si tracks are preferred over other double-track conversions, followed by single-track conversions. In addition, the vertices with smallest conversion radius are preferred.

5.2.3 Supercluster reconstruction

The reconstruction of photon superclusters proceeds in two stages: in the first stage, EM topoclusters are tested for use as seed cluster candidates, which form the basis of superclusters; in the second, those EM topoclusters near the seed candidates are iden-

tified as satellite cluster candidates. Satellite clusters may emerge from bremsstrahlung radiation or topocluster splitting, and they are added to the seed candidates to form the final superclusters when they fulfil the necessary conditions.

The initial list of EM topoclusters is sorted according to decreasing E_{T} (with an exception with clusters in the transition region of $1.37 < |\eta| < 1.63$). The clusters are tested one by one for use as seed clusters. For photon reconstruction, a cluster must have $E_{\mathrm{T}} > 1.5$ GeV to be classified as seed, with no requirement made on any track or conversion vertex matching. Then, the algorithm classifies a cluster as satellite if it falls within a window of $\Delta\eta \times \Delta\phi = 0.075 \times 0.125$ around the seed cluster barycentre. These cases tend to represent EM showers originating from the same initial photon. For photons with conversion vertices made up only of tracks containing silicon hits, a cluster is added as a satellite if its best-matched track belongs to the conversion vertex matched to the seed cluster. The previous steps rely on tracking information to discriminate between distant radiative photons or conversion electrons from pile-up noise or other unrelated clusters.

The seed clusters together with their associate satellite clusters are called superclusters. In the last step of the supercluster-building algorithm, calorimeter cells are assigned to a given supercluster. Only cells from the presampler and the first three LAr calorimeter layers are considered, with an exception in the transition region of $1.4 < |\eta| < 1.6$, where the energy measured in the tile calorimeter is also added. The sensitivity of the superclusters to pile-up noise is limited restricting the size of each constituent topocluster to a maximal width of 0.075 or 0.125 in the η direction in the barrel or end-cap region, respectively. No restriction is applied in the ϕ-direction.

5.2.4 Photons for analysis

An initial energy calibration and position correction is applied to the electron and photon superclusters once they are built. Tracks are matched to electron supercluster and conversion vertices to photon superclusters. This matching is performed in the same way as the matching to EM topoclusters was performed, but using superclusters instead. Since electron and photon superclusters are built independently, a given seed cluster can produce both of them. In those cases, the procedure depicted in figure 5.2 is followed, with the purpose of classifying the ambiguous cases as photons or electrons. Otherwise, both an electron and a photon object are created.

Once the final track and conversion vertices matching is performed, the energies of electrons and photons are recalibrated, following the procedure described in ref. [130]. Subsequently, shower shape and other discriminant variables [129] are calculated for electron and photon identification.

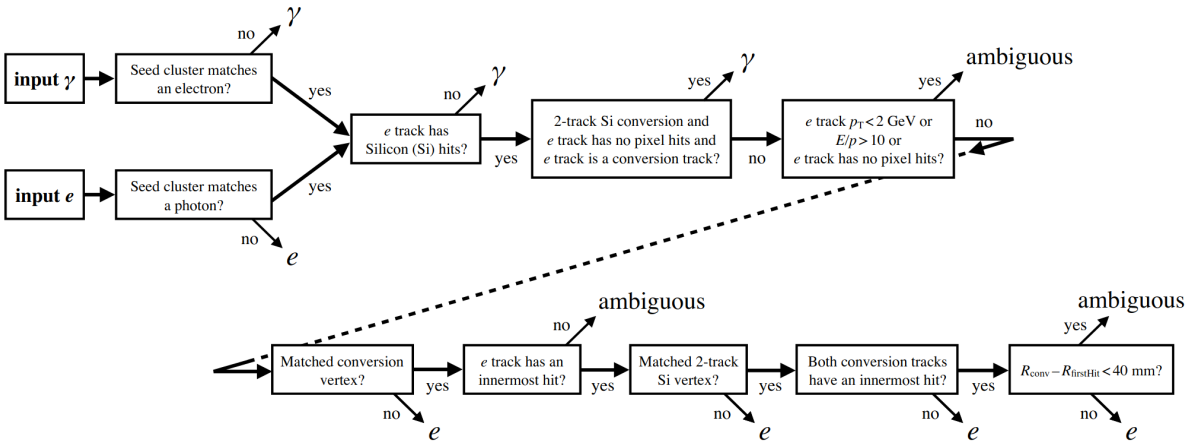


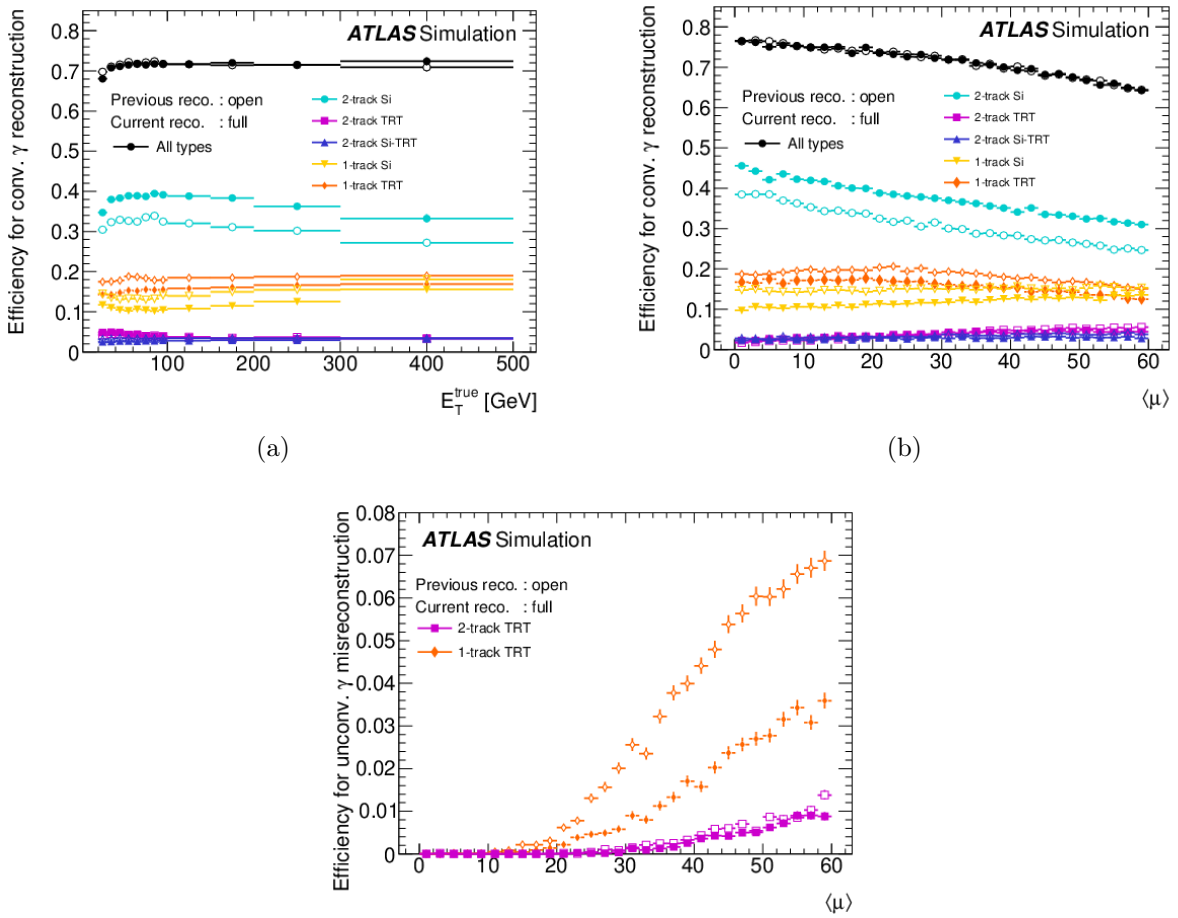
Figure 5.2: Flowchart showing the logic of the ambiguity resolution for particles initially reconstructed both as electrons and photons. An *innermost hit* is a hit in the functioning pixel nearest to the beam-line along the track trajectory, E/p is the ratio of the supercluster energy to the measured momentum of the matched track, R_{conv} is the radial position of the conversion vertex, and $R_{\text{firstHist}}$ is the smallest radial position of a hit in the track or tracks that make a conversion vertex. From ref. [131].

5.2.5 Photon reconstruction performance

The reconstruction efficiencies for converted photons as a function of the particle level E_{T} are shown in figure 5.3. The efficiencies obtained with the latest version of the reconstruction software described in this section (full markers) are compared with those obtained with the previous version (open markers) described in ref. [129]. An improvement in the reconstruction efficiency for double-track Si conversions and the corresponding reduction of single-track Si conversions is visible in figures 5.3(a) and 5.3(b). A slight reduction in double- and single-track TRT conversion efficiencies is also visible, and related with the significant reduction of the probability for true unconverted photons to be reconstructed as TRT conversions, as can be observed in figure 5.3(c).

5.3 Photon calibration

The calibration of the energy response of electrons and photons from the energy of a cluster of cells in the EM calorimeter is divided in various steps that are described in detail in the following sections. As explained at the beginning of section 5.2, for the analyses presented in this thesis which made use of the photon reconstruction described in ref. [129], photons were calibrated with the methods reported in ref. [130]. On the other hand, for those analyses which used the updated reconstruction methods published in ref. [131], the photon calibration procedures were updated as well (see section 5 in ref. [131]). The updated calibration methods reported in ref. [131] include an optimised measurement of



(a) (b)

(c)

Figure 5.3: On the top, the converted photon reconstruction efficiency and contributions of the different conversion types as a function of $E_{\mathrm{T}}^{\mathrm{true}}$ (a) and $\langle\mu\rangle$ (b). On the bottom, the probability of an unconverted photon to be mistakenly reconstructed as a converted photon as a function of $\langle\mu\rangle$. From ref. [131].

the energy resolution of the photon using a multivariate regression algorithm based on the properties of the shower development in the EM calorimeter (see section 5.3.1). In addition, the absolute energy scale using $Z \to ee$ decays was updated (see section 5.3.3), together with systematic uncertainties related to pile-up and material effects.

5.3.1 Electron and photon energy estimate and expected resolution from the simulation

The energy of electrons and photons is computed from the energy of the reconstructed cluster, and a correction is derived to account for the energy lost in the material upstream

of the calorimeter, for the energy deposited in the cells neighbouring the cluster in η and ϕ, and for the energy lost beyond the LAr calorimeter. The properties of the shower development are used to compute a single correction for all these effects using multivariate regression algorithms tuned on simulated single particles separately for converted and unconverted photons. The training of the algorithm, based on *boosted decision trees*, is done in intervals of $|\eta|$ and E_{T}. The procedure covers the whole EM calorimeter up to $|\eta| = 2.5$, including the transition region between the barrel and the end-cap. The variables considered in the regression are detailed in ref. [130]. The same algorithm is applied to data and simulation. As a result of this step, the energy resolution is optimised and the impact of the material in front of the calorimeter is minimised.

5.3.2 Corrections applied to the data

Corrections are applied to the data to account for residual differences between data and simulation. They include the intercalibration of the different calorimeter layers, corrections for energy shits induced by pile-up and corrections applied to improve the uniformity of the energy response. Since the absolute energy is set with $Z \to ee$ decays, only the relative calibration of the energy scales of the first two layers and the presampler is needed. In addition, given the small fraction of energy deposited in the third layer of the calorimeter, no dedicated corrections for its intercalibration are derived.

Intercalibration of the first and second calorimeter layers

To study the relative calibration of the first and second calorimeter layers, muon energy deposits, which are insensitive to the amount of passive material in front of the EM calorimeter, are used. This relative calibration is derived by comparing the energy deposits in data with simulation predictions. The analysis uses $Z \to \mu\mu$ decays, requiring $p_{\mathrm{T}}^{\mu} > 27$ GeV. The calorimeter cells crossed by the muon tracks are determined by extrapolating the track to each layer of the calorimeter, taking into account the geometry of the calorimeter, the misalignment between the ID and the calorimeter, and the magnetic field encountered by the muon along this path.

The observed muon energy in each of the two layers is described by the convolution of a Landau distribution, representing the energy deposit, and a noise distribution. The Most Probable Value (MPV) of the deposited muon energy is extracted using a fit of the convolution function to the observed muon energy distribution (*fit method*). An alternative method where the mean is computed over a restricted window to minimise the sensitivity to the tails of the distribution is also used (*truncated-mean method*). The same procedure is applied to data and MC and the relative calibration of the two layers is computed as $\alpha_{1/2} = \left(\langle E_1\rangle^{\mathrm{data}}/\langle E_1\rangle^{\mathrm{MC}}\right) / \left(\langle E_2\rangle^{\mathrm{data}}/\langle E_2\rangle^{\mathrm{MC}}\right)$, where $\langle E_1\rangle$ ($\langle E_2\rangle$) denotes the MPV in the first (second) layer.

The residual pile-up dependences of the extracted MPV values are further reduced for both methods performing the analysis as a function of $\langle\mu\rangle$. Once this correction is

applied, the final result is given by the average of the two signal extraction methods, the fit and truncated mean. In addition, several effects are investigated to estimate the uncertainty in the $\alpha_{1/2}$ value measured with muons from the average of the results of the two methods mentioned above. Figure 5.4 shows the results for $\alpha_{1/2}$ and the comparison of the two methods.

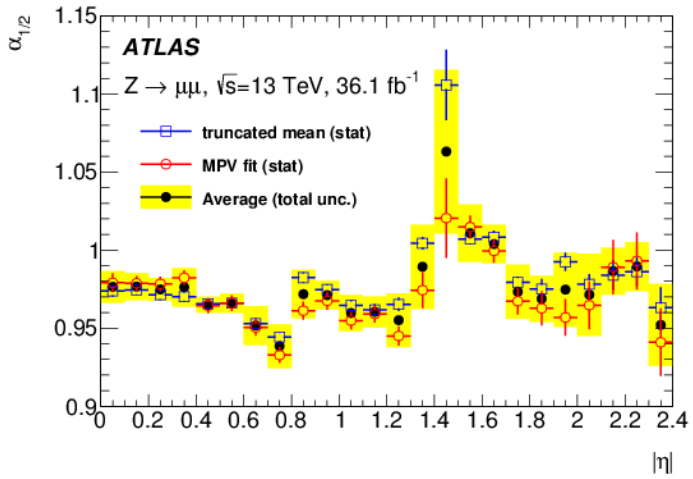


Figure 5.4: Ratio $\alpha_{1/2} = \left(\langle E_1\rangle^{\text{data}}/\langle E_1\rangle^{\text{MC}}\right) / \left(\langle E_2\rangle^{\text{data}}/\langle E_2\rangle^{\text{MC}}\right)$ as a function of $|\eta|$, as obtained from the study of the muon energy deposits in the first two layers of the calorimeters. The results from the two methods are shown with their statistical uncertainties. The final average measurement is shown with its total uncertainty including the statistical and systematic uncertainties. From ref. [130].

The relative calibration of the first two layers of the EM calorimeter can also be probed using $Z \to ee$ decays by investigating the variation of the mean of the dielectron invariant mass as a function of the ratio of the energies of the electron or positron candidates in the first two layers. Good agreement with respect to the nominal result obtained with muons from $Z \to \mu\mu$ decays was obtained except in the $|\eta|$ range 1.2 to 1.8, as reported in ref. [130]. The reason for this discrepancy is due to the large sensitivity of the results of the method based on $Z \to ee$ to the interval used to compute the average invariant mass [130].

Presampler energy calibration

The presampler energy scale α_{PS} is determined from the ratio of the presampler energies in data and MC. The measured energy from electrons from Z decays in the presampler is sensitive to α_{PS} and to the material in front of the presampler. To be sensitive only to α_{PS}, the correlation between the shower development and the material in front of the presampler is studied. An additional correction is applied accounting for the impact on the energies deposited in the first two layers $E_{1/2}$ due to any mismodelling of the material between the presampler and the calorimeter. The presampler energy scale is extracted as

$$\alpha_{\mathrm{PS}} = \frac{E_0^{\mathrm{data}}(\eta)}{E_0^{\mathrm{MC}}(\eta)} \times \frac{1}{1 + A(\eta)\left(\frac{E_{1/2}^{\mathrm{data}}(\eta)}{E_{1/2}^{\mathrm{MC}}(\eta) b_{1/2}(\eta)} - 1\right)},$$

where $E_0^{\mathrm{data}}(\eta)$ and $E_0^{\mathrm{MC}}(\eta)$ are the average energies deposited in the presampler by electrons from Z decays in data and MC. The factor $b_{1/2}(\eta)$ is the ratio of $E_{1/2}$ in data and MC for unconverted photons with small energy deposit in the presampler, and it accounts for the mismodelling of the material between the presampler and the calorimeter. $E_{1/2}^{\mathrm{data}}(\eta)$ and $E_{1/2}^{\mathrm{MC}}(\eta)$ are the average values of $E_{1/2}$ in data and MC extracted for electrons from Z decays. Once the correction with $b_{1/2}(\eta)$ is applied, this ratio is proportional to the amount of material in front of the presampler. $A(\eta)$ is estimated with MC and represents the correlation between the changes in $E_{1/2}$ and E_0 when varying the amount of material in front of the presampler.

This measurement is performed in intervals of $|\eta|$ and excluding the transition region between calorimeters $1.37 < |\eta| < 1.52$. Uncertainties in the measurements of α_{PS} include the statistical uncertainties of the various input quantities in the data and simulation. Figure 5.5 shows the result for α_{PS} as a function of $|\eta|$.

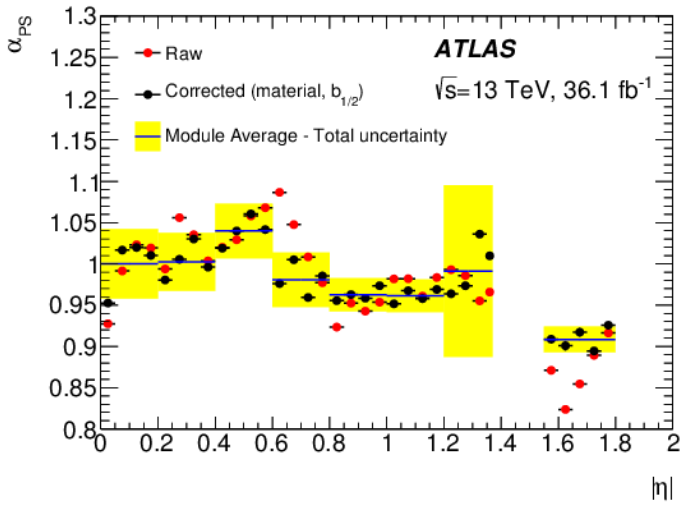


Figure 5.5: Measurement of the presampler energy scale ratio between data and simulation. The red dots show the measurement before the material and $b_{1/2}(\eta)$ corrections are applied. The black dots represent the measurement once both corrections are applied. The values averaged per presampler module in $|\eta|$ are shown together with the total uncertainties, represented by shaded areas. From ref. [130].

Pile-up energy shifts

After bipolar shaping, the average energy induced by pile-up interactions should be zero in an ideal situation of bunch trains with an infinite number of bunches and with the same luminosity in each pair of colliding bunches. However, in practice, bunch-to-bunch

luminosity variations and the finite bunch-train length can create significant energy shifts which depend on the position inside the bunch train and on the luminosity. To mitigate this effect on the estimation of the cell energies, the average expected pile-up energy shift is subtracted cell-by-cell. This average is estimated as a function of the bunch position inside the LHC ring, taking into account the instantaneous luminosity per bunch, the expected pulse shape as a function of time, the optimal filtering coefficients used to estimate the amplitude of the signal and a normalisation factor derived from data with single colliding bunches.

To further reduce the impact of pile-up-induced energy shifts for EM clusters, an additional correction is applied separately to each calorimeter layer as a function of $\langle\mu\rangle$ and η.

Improvements in the uniformity of the energy response

After all the corrections previously described are applied to the electron and photon candidates in the data separately for each calorimeter layer, the energy is computed using a regression algorithm. Corrections for variations in the energy response as a function of the impact point of the shower in the calorimeter affecting only the data are applied to the energy of the electron and photon candidates. Two effects are corrected, accounting for the energy loss between the barrel and calorimeter modules and for the effect of high-voltage inhomogeneities in the calorimeter.

5.3.3 Data/MC energy scale and resolution measurements with $Z \to ee$ decays

The difference in energy scale between data and MC is defined as α_{i}, where i corresponds to different regions in η. On the other hand, the mismodelling of the energy resolution is parameterised as an η-dependent constant term c_{i}. The energy scale correction is applied to the data, while the resolution correction is applied to the simulation using

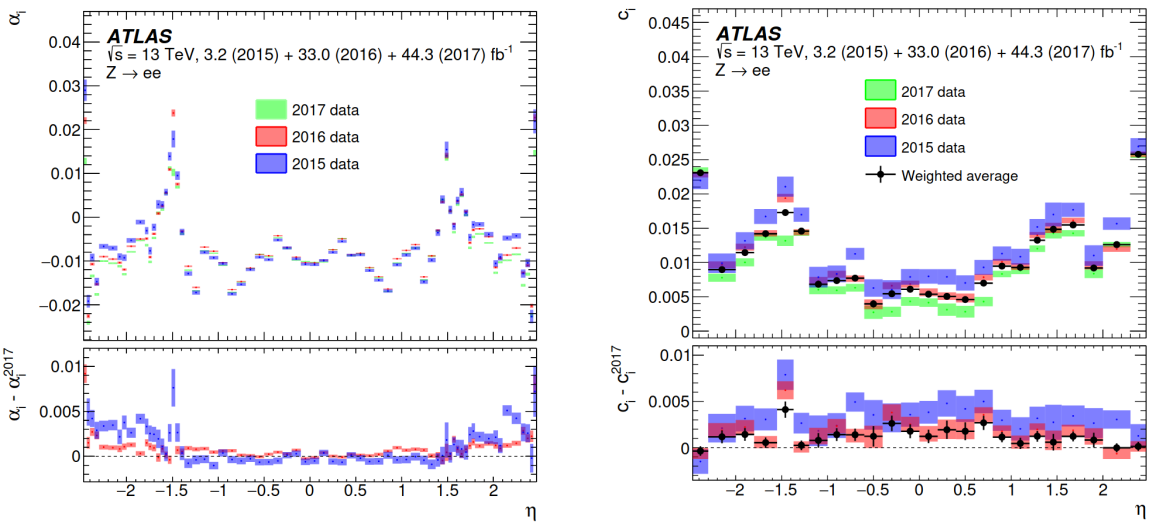
$$E^{\mathrm{data,corr}} = E^{\mathrm{data}} / \left(1 + \alpha_{\mathrm{i}}\right) \,, \; \left(\frac{\sigma_{\mathrm{E}}}{E}\right)^{\mathrm{MC,corr}} = \left(\frac{\sigma_{\mathrm{E}}}{E}\right)^{\mathrm{MC}} \oplus c_{\mathrm{i}} \,,$$

where the symbol $\oplus$ denotes a sum in quadrature.

For samples of $Z \to ee$ decays, where the electrons are reconstructed in η regions i and j, the effect of the energy scale correction on the dielectron invariant mass is given to first order by $m_{\mathrm{ij}}^{\mathrm{data,corr}} = m_{\mathrm{ij}}^{\mathrm{data}} / \left(1 + \alpha_{\mathrm{ij}}\right)$, with $\alpha_{\mathrm{ij}} = \left(\alpha_{\mathrm{i}} + \alpha_{\mathrm{j}}\right)/2$. Similarly, the difference in the MC mass resolution is given by $\left(\sigma_{\mathrm{m}}/m\right)_{\mathrm{ij}}^{\mathrm{MC,corr}} = \left(\sigma_{\mathrm{m}}/m\right)_{\mathrm{ij}}^{\mathrm{MC}} \oplus c_{\mathrm{ij}}$, with $c_{\mathrm{ij}} = \left(c_{\mathrm{i}} \oplus c_{\mathrm{j}}\right)/2$. Hence, the values of α_{ij} and c_{ij} are determined via an optimisation of the agreement between the invariant mass distributions in data and MC, separately for each (i, j) region, and finally, the α_{i} and c_{i} parameters are extracted from a simultaneous fit of all the regions.

Two methods are used to extract the parameters, and their difference is taken as a systematic uncertainty. In the first method, the best estimates of α_{ij} and c_{ij} are obtained by minimising the χ^2 of the difference between data and MC. The second method is used as a cross-check, where the sum of three Gaussian functions is fitted to the data and MC invariant mass distributions in each (i, j) region. The parameters α_i and c_i are extracted from the differences, between data and MC, of the means and widths of the fitted distributions.

Results of α_i and c_i derived in fine η intervals and separately for 2015, 2016 and 2017 are shown in figure 5.6. The differences in α_i for the different years are mainly due to two effects, namely the variations of the LAr temperature, and the increase of the instantaneous luminosity [131]. For the constant term corrections c_i, a dependence on the pile-up level is observed through the different values obtained for 2015 to 2017 data [131].



(a) (b)

Figure 5.6: The energy scale factors α_i (a) and the additional constant term c_i (b) as a function of η. The shaded areas correspond to the statistical uncertainties. The bottom panels show the differences between α_i (a) and c_i (b) measured in a given data-taking period and the measurements obtained using 2017 data. From ref. [131].

5.3.4 Systematic uncertainties in the energy scale and resolution

There are several systematic uncertainties that impact the measurement of the energy of electrons and photons (converted or unconverted) in a way that depends on their transverse energy or pseudorapidity. After the Z-based calibration, the energy scale and its uncertainty are fixed for electrons with transverse energy close to the average of those produced in Z decays. The relative uncertainty for a given electron or photon with transverse energy E_T and pseudorapidity η can be written as

$$\delta E_{\mathrm{i}}^{\mathrm{e},\gamma}\left(E_{\mathrm{T}},\eta\right) = \Delta E_{\mathrm{i}}^{\mathrm{e},\gamma}\left(E_{\mathrm{T}},\eta\right) - \Delta E_{\mathrm{i}}^{\mathrm{e}}\left(\langle E_{\mathrm{T}}^{\mathrm{e}(Z\to ee)}\rangle,\eta\right) ,$$

where $\Delta E_{\mathrm{i}}^{\mathrm{e},\gamma}\left(E_{\mathrm{T}},\eta\right)$ is, for a given uncertainty variation i, its relative impact on the energy as a function of η and E_{T} before the application of the Z-based calibration, and $\langle E_{\mathrm{T}}^{\mathrm{e}(Z\to ee)}\rangle \sim 40$ GeV is the average transverse energy for electrons produced in Z boson decays. The Z-based calibration absorbs the effect for electrons with $E_{\mathrm{T}} = \langle E_{\mathrm{T}}^{\mathrm{e}(Z\to ee)}\rangle$ and leaves the residual uncertainty $\delta E_{\mathrm{i}}^{\mathrm{e},\gamma}\left(E_{\mathrm{T}},\eta\right)$.

For a given uncertainty variation i, $\delta E_{\mathrm{i}}^{\mathrm{e},\gamma}\left(E_{\mathrm{T}},\eta\right)$ can change sign as a function of E_{T}. Moreover, for most of the considered uncertainty variations, their impact on the photon energy is computed separately for converted and unconverted categories. Converted photons have a shower development more similar to that of electrons and, as a consequence, usually smaller energy scale systematic uncertainties than unconverted photons.

The different sources of systematic uncertainties affecting the energy scale of electrons and photons are: the energy shift induced by pile-up, the impact of the layer calibration, the impact of the E_4 scintillator calibration, those due to the material in front of the calorimeter, the non-linearity of the cell energy measurement, the modelling of the lateral shower shape and the modelling of the photon reconstruction classification.

Energy resolution uncertainties

A formalism similar to that for the energy scale uncertainty can be used to describe the resolution uncertainties. Defining $\Delta\sigma_{\mathrm{i}}^{e,\gamma}\left(E_{\mathrm{T}},\eta\right)$ as the uncertainty in the relative energy resolution for a given particle from a given uncertainty variation i, the residual uncertainty after the adjustment of the resolution based on the Z decays can be written as

$$\delta\Sigma_{\mathrm{i}}^{\mathrm{e},\gamma}\left(E_{\mathrm{T}},\eta\right) = \Delta\Sigma_{\mathrm{i}}^{\mathrm{e},\gamma}\left(E_{\mathrm{T}},\eta\right) - \Delta\Sigma_{\mathrm{i}}\left(\langle E_{\mathrm{T}}^{\mathrm{e}(Z\to ee)}\rangle,\eta\right) ,$$

where Σ denotes the square of the relative energy resolution σ.

The different contributions to the energy resolution are: the shower and sampling fluctuations in the calorimeter, the fluctuations in energy loss upstream of the calorimeter, the effect of electronics and pile-up noise and the impact of residual non-uniformities affecting the measurement of the energy in the data.

5.3.5 Cross-checks with $J/\Psi \to ee$ and $Z \to ll\gamma$

$J/\Psi \to ee$ decays

The energy scale of low-energy electrons (average transverse energies around 10 GeV) is probed using $J/\Psi \to ee$ events. The known mass of the J/Ψ resonance provides a completely independent check of the energy calibration for low-energy electrons. The full

calibration procedure described in the previous sections is applied, including the energy scale derived from $Z \to ee$ events. In this way, the difference between data and simulation for $J/\Psi \to ee$ events is quantified using residual energy scale differences, $\Delta\alpha$, extracted from the peak position of the reconstructed invariant mass. If the energy calibration is correct, the $\Delta\alpha$ differences should be consistent with zero within the combined uncertainties of the $J/\Psi \to ee$ measurement and the systematic uncertainty of the energy calibration.

$Z \to ll\gamma$ decays

The energy scale correction of photons is assumed to be the same as the one extracted from $Z \to ee$ decays previously described, within uncertainties. In this case, the radiative decays of the Z boson can be used to check the energy scale of photons in the low-energy region in particular. Converted and unconverted photons are studied separately and the electron and muon channels are treated independently and then combined. All the corrections previously described are applied to electrons and photons, and residual energy scale factors for photons are derived by comparing the data with MC. The residual photon energy scale difference is parameterised as an additional correction to the photon energy $\Delta\alpha$, similarly as it is done in the $J/\Psi \to ee$ study. The mass distribution of the $ll\gamma$ system in the simulation is modified by applying the correction $\Delta\alpha$ to the photon energy, and the value $\Delta\alpha$ that minimises the χ^2 comparison between data and simulation is computed. If the energy calibration is correct, the value of $\Delta\alpha$ should be consistent with zero within the uncertainties in the measurement and in the photon energy scale.

5.4 Photon identification

Photon identification criteria are designed to efficiently select isolated-prompt photons and reject the main background coming from hadronic jets. The photon identification is constructed from a *cut-based* selection using the shower shape variables described in table 5.1. Those variables which use the first layer of the EM calorimeter play a particularly important role in rejecting π^0 decays into two highly-collimated photons.

The primary identification selection is labelled as tight, and there are two less restrictive selections called medium and loose. As introduced in section 5.1.2, the medium working point is used only in the HLT. The loose identification criteria remained unchanged since the beginning of Run II, and represents the main selection used in the triggering of photon and diphoton events in 2015 and 2016. It uses the R_{had}, R_{had_1}, R_η, and w_{η_2} shower shape variables. The medium selection adds a loose cut on E_{ratio}, and became the main trigger selection for diphotons in the beginning of 2017 to maintain an acceptable trigger rate. Since the reconstruction of photons in the ATLAS trigger system does not differentiate between converted and unconverted photons, the loose and medium criteria are the same for both photon categories. Because the shower shapes vary due to the geometry of the calorimeter, the three cut-based selections are optimised separately

Category	**Description**	**Name**	**Loose**	**Tight**
Acceptance	$\lvert\eta\rvert < 2.37$, with $1.37 \leq \lvert\eta\rvert < 1.52$ excluded	–	✓	✓
Hadronic leakage	Ratio of E_{T} in the first layer of the hadronic calorimeter to E_{T} of the EM cluster (used over the ranges $\lvert\eta\rvert < 0.8$ and $\lvert\eta\rvert > 1.37$).	$R_{\mathrm{had_1}}$	✓	✓
	Ratio of E_{T} in the hadronic calorimeter to E_{T} of the EM cluster (used over the range $0.8 < \lvert\eta\rvert < 1.37$).	R_{had}	✓	✓
EM second layer	Ratio of the energy in 3×7 $\eta \times \phi$ cells over the energy in 7×7 cells centred around the most energetic cell.	R_{η}	✓	✓
	Lateral shower width, $\sqrt{\frac{(\sum E_{\mathrm{i}}\eta_{\mathrm{i}}^2)}{(\sum E_{\mathrm{i}})} - \left(\frac{(\sum E_{\mathrm{i}}\eta_{\mathrm{i}})}{(\sum E_{\mathrm{i}})}\right)^2}$, where E_{i} is the energy and η_{i} is the pseudorapidity of cell i and the sum is calculated within a window of 3×5 cells.	w_{η_2}	✓	✓
	Ratio of the energy in 3×3 $\eta \times \phi$ cells over the energy of 3×7 cells centred around the most energetic cell.	R_{ϕ}		✓
EM first layer	Lateral shower width, $\sqrt{\frac{(\sum E_{\mathrm{i}}(i-i_{\mathrm{max}})^2)}{\sum E_{\mathrm{i}}}}$, where i runs over all cells in a window of 3 cells around the highest-energy cell.	$w_{s\,3}$		✓
	Total lateral shower width, $\sqrt{\frac{(\sum E_{\mathrm{i}}(i-i_{\mathrm{max}})^2)}{\sum E_{\mathrm{i}}}}$, where i runs over all cells in a window of $\Delta\eta \sim 0.0625$ and i_{max} is the index of the highest-energy cell.	$w_{s\,\mathrm{tot}}$		✓
	Energy fraction outside core of three central cells, within seven cells.	f_{side}		✓
	Difference between the energy of the cell associated with the second maximum, and the energy reconstructed in the cell with the smallest value found between the first and second maxima.	ΔE_s		✓
	Ratio of the energy difference between the maximum energy deposit and the energy deposit in a secondary maximum in the cluster to the sum of these energies.	E_{ratio}		✓
	Ratio of the energy in the first layer of the EM calorimeter to the total energy of the EM cluster.	f_1		✓

Table 5.1: Discriminating variables used for loose and tight photon identification.

in bins of $|\eta|$. The tight identification is also optimised in separate bins of E_{T} for its application in analyses [131] (see section 5.4.1).

The selection based on R_{had} and second layer requirements aims to reject most of the jet background. The selection on the first layer variables is optimised to reduce

further the contribution from jets with one or more hard π^0 decaying to photons and carrying most of the jet energy, which would pass the R_{had} and second layer requirements. The requirements on the first layer variables are different for converted and unconverted photon candidates due to the opening angle of the e^+e^- conversion pair, which is amplified by the magnetic field, and from the additional interaction of the conversion pair with the material upstream of the calorimeters. The first layer requirements also depend on the $|\eta^\gamma|$ interval to account for different shower developments and the varying amount of material upstream of the calorimeter.

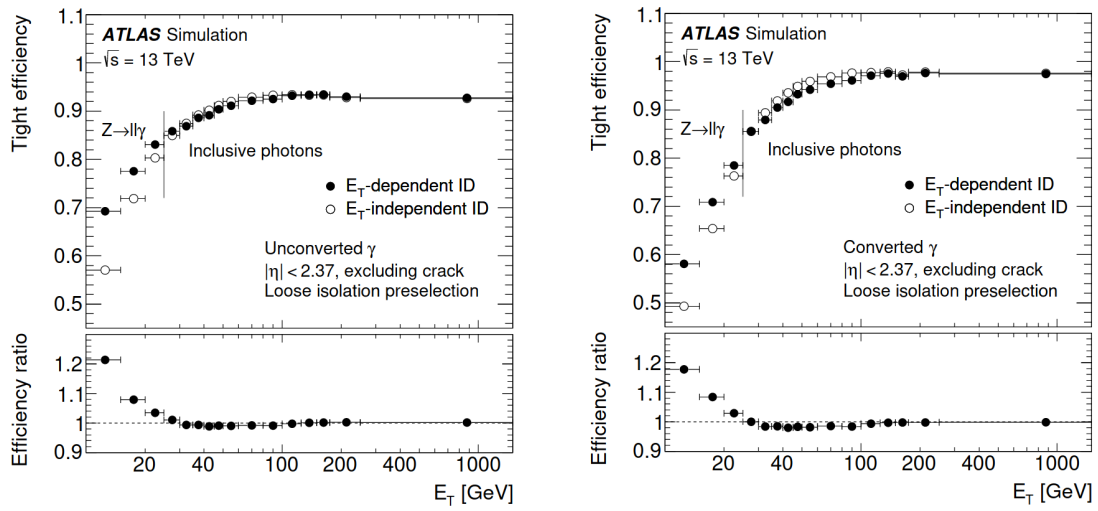
The loose cut uses four of the shower-shape variables defined in table 5.1, namely R_{had}, R_{had_1}, R_η, and w_{η_2}, while the tight cut is based on eleven discriminating variables: R_{had}, R_{had_1}, R_η, w_{η_2}, R_ϕ, $w_{s\,3}$, $w_{s\,\text{tot}}$, f_{side}, ΔE_s, E_{ratio}, and f_1. An additional selection criteria, called the *loose′* selection, is defined and used offline for the different analyses presented in this thesis. Different loose′ menus are used, but every variation is defined relaxing a certain number of cuts from the tight criteria. The nominal loose′ (*LoosePrime4*) criteria contains all tight cuts minus the $w_{s\,3}$, f_{side}, ΔE_s, and E_{ratio}, that are relaxed. For systematic studies, the loose′ definition was modified removing the tight cut on the variable $w_{s\,\text{tot}}$, to obtain a looser criteria (*LoosePrime5*) or including the tight cut on E_{ratio} (*LoosePrime3*) or both E_{ratio} and ΔE_s (*LoosePrime2*) for the two tighter criteria definitions. Non-tight photon candidates are defined as those photons which satisfy the loose′ criteria, but fail at least one of the tight requirements. The non-tight photon candidates are used for the data-driven background subtraction procedure described in section 7.3.

5.4.1 Optimisation of the photon identification

The tight identification optimisation procedure reported in ref. [131] was updated with respect its previous version [129]. As explained at the beginning of sections 5.2 and 5.3, the analyses presented in this thesis that made use of the full Run II dataset were performed using the latest photon reconstruction, calibration and identification methods reported in ref. [131], while the inclusive photon analysis using $36\,\text{fb}^{-1}$ and the in situ JES calibration using γ+jet events used the reconstruction, calibration and identification methods published in ref. [129, 130].

The tight identification was optimised in ref. [131] using a Toolkit for Multivariate Analysis (TMVA). MC simulations that provide prompt photons and representative backgrounds at different transverse momenta were used. For photons with $10 < E_{\text{T}}^\gamma < 25$ GeV, a $Z \to ll\gamma$ MC sample was used as signal. The corresponding background sample was obtained from data consisting of Z+jet events collected using a similar selection, but with relaxed requirements on the dilepton and dilepton+photon invariant masses. Above $E_{\text{T}}^\gamma = 25$ GeV, an inclusive-photon production MC sample was compared with a dijet background MC sample enriched in high-E_{T} energy deposits using a generator-level filter. No isolation selection was applied to the training samples and the shower shape variables were corrected to match those observed in data.

The result of the tight identification optimisation in terms of the efficiencies as a function of E_{T}^{γ} for the signal MC training samples is shown in figure 5.7 for both unconverted and converted categories. The optimised selection, labelled as $E_{\mathrm{T}} - dependent$, is compared with a reference selection which uses criteria that do not change with E_{T}^{γ} ($E_{\mathrm{T}} - independent$). This new, $E_{\mathrm{T}} - dependent$ tight identification allows the efficiencies of low- and high-E_{T} photon regions to be tuned separately; it is tuned to give a $\sim 20\%$ higher efficiency at low E_{T}^{γ}, and an improved background rejection at high E_{T}^{γ}.



(a) (b)

Figure 5.7: Efficiencies of the tight photon identification for unconverted (left) and converted (right) signal photons as a function of photon E_{T}. The signal events were taken from the sample of $Z \to ll\gamma$ photons with $E_{\mathrm{T}}^{\gamma} < 25$ GeV, and from inclusive-photon production for $E_{\mathrm{T}}^{\gamma} > 25$ GeV. In each case, the $E_{\mathrm{T}} - dependent$ and $E_{\mathrm{T}} - independent$ selections are compared. In both figures, the bottom panel shows the ratio between the $E_{\mathrm{T}} - dependent$ and the $E_{\mathrm{T}} - independent$ identification efficiencies. From ref. [131].

To asses the performance of the $E_{\mathrm{T}} - dependent$ tight photon identification on data, three photon efficiency measurements were performed using distinct data samples. The first uses an inclusive-photon production data selection, the second uses photons radiated from leptons in $Z \to ll\gamma$ decays, and the third uses electrons from $Z \to ee$ decays, with a method that transforms the electron shower shapes to resemble the photon shower shapes. These efficiency measurements are described in detail in ref. [129].

5.5 Photon isolation

The activity near a photon or a lepton can be quantified from the energy deposits in the calorimeters or from the tracks of nearby charged particles. These two possibilities are used to construct isolation variables.

The raw calorimeter isolation variable ($E_{\mathrm{T,raw}}^{\mathrm{iso}}$) is built summing the transverse energy of the positive-energy topoclusters whose barycentre falls within a cone centred around the electron or photon cluster barycentre. The topoclusters are calibrated to the EM scale. This raw calorimeter variable includes the EM particle energy ($E_{\mathrm{T,core}}$), which is subtracted by removing the energy of the EM calorimeter cells contained in a $\Delta\eta \times \Delta\phi = 5 \times 7$ (in EM-middle-layer units) rectangular cluster around the barycentre of the EM particle cluster. An advantage of this method is that a stable subtraction for real or fake/non-prompt objects for any transverse momentum and pile-up is performed. However, not all the EM particle energy is subtracted, and an additional leakage correction is needed. The leakage correction is parameterised as a function of E_{T} and $|\eta|$, using MC samples of single electrons or photons without pile-up. In addition, a correction for pile-up and UE contributions to the isolation cone is also estimated. This correction comes from the so-called *jet-area* method [138], where low-energy jets are used to compute an ambient transverse energy density in an event-by-event basis (see appendix B), which is then multiplied by the area of the isolation cone and subtracted from the transverse isolation energy. The addition of this correction improves the matching between the isolation definitions between measurements and theory, since fixed-order pQCD calculations do not include pile-up or UE effects.

The fully corrected calorimeter isolation variable is computed as

$$E_{\mathrm{T}}^{\mathrm{coneXX}} = E_{\mathrm{T,raw}}^{\mathrm{isoXX}} - E_{\mathrm{T,core}} - E_{\mathrm{T,leakage}}\left(E_{\mathrm{T}}, \eta, \Delta R\right) - E_{\mathrm{T,pile-up}}\left(\eta, \Delta R\right) \ ,$$

where XX refers to the size of the employed cone, and $\Delta R = \mathrm{XX}/100$.

After the application of the previous corrections, it is observed at analysis level that the MC simulation of the $E_{\mathrm{T}}^{\mathrm{iso}}$ distribution does not describe the position of the peak in the data, with certain dependence on the $|\eta^{\gamma}|$ region. Data-driven corrections to the simulated $E_{\mathrm{T}}^{\mathrm{iso}}$ variable were determined using a Crystal-Ball distribution so that the MC plus the background (as estimated using the non-tight photons in the data) distribution describes better the data. These data-driven corrections to $E_{\mathrm{T}}^{\mathrm{iso}}$ are applied to MC simulated events with the subsequent improvement in terms of the description of the $E_{\mathrm{T}}^{\mathrm{iso}}$ distribution (see appendix C).

In addition to the calorimeter isolation described above, the track isolation variable ($p_{\mathrm{T}}^{\mathrm{cone20}}$) is computed by summing the p_{T} of the selected tracks within a cone centred around the electron track or the photon cluster direction. The tracks matched to the electron or converted photon are excluded. For photon isolation, the tracks considered are required to have $p_{\mathrm{T}} > 1$ GeV and $|\eta| < 2.5$, at least seven silicon (Pixel + SCT) hits, at most one shared hit[5], at most two silicon holes (i.e. missing hits in the Pixel and SCT detectors), and at most one pixel hole.

[5]Shared hits are defined as $n_{\mathrm{Pixel}}^{\mathrm{sh}} + n_{\mathrm{SCT}}^{\mathrm{sh}}/2$, where $n_{\mathrm{Pixel}}^{\mathrm{sh}}$ and $n_{\mathrm{SCT}}^{\mathrm{sh}}$ are the numbers of hits assigned to several tracks in the Pixel and SCT detectors.

5.6 Additional corrections to simulated events

Additional corrections are applied to the simulated events to match the overall event conditions of the data sample and to account for known differences between data and simulation. These corrections are:

- Pile-up corrections: the MC events are reweighted such that the distribution of $\langle\mu\rangle$ matches the one in the data. The $\langle\mu\rangle$ distribution in data is scaled down applying a general factor of 1/1.09 for the analyses using MC15 samples, while it is recommended to use a factor equal to 1/1.03 for the latest MC16 (see sections 7.2 and 8.2, respectively). This scaling[6] improves the agreement between data and MC for the distribution of N_{PV} [139]. The scale factor applied to the data is measured with an uncertainty; to estimate the systematic effect on the cross section due to this uncertainty, two variations (up and down) of the scale factor applied to the data are used to reweight the MC.

- Corrections to the photon identification:

  1. The distributions of the photon transverse shower shapes in MC simulation do not perfectly describe those observed in data [129]. Small systematic differences in their average values are observed. Fudge factors are applied to the photon shower-shape variables in the signal MC simulated events to correct for the differences observed between data and MC simulation in the photon-cluster discriminating variables. Examples of the simulated shower-shape distributions before and after the corrections, for converted and unconverted photon candidates originating from Z boson radiative decays, are shown in figure 5.8.
  2. Scale factors are applied to the photons in MC to account for the differences observed in the tight identification efficiency between data and MC simulation [129, 131].

[6] The reason why the data are scaled down instead of scaling up the MC is purely technical. Due to the discrete nature of the values of μ used in MC, it is more practical to scale the μ in data by the inverse scale factor.

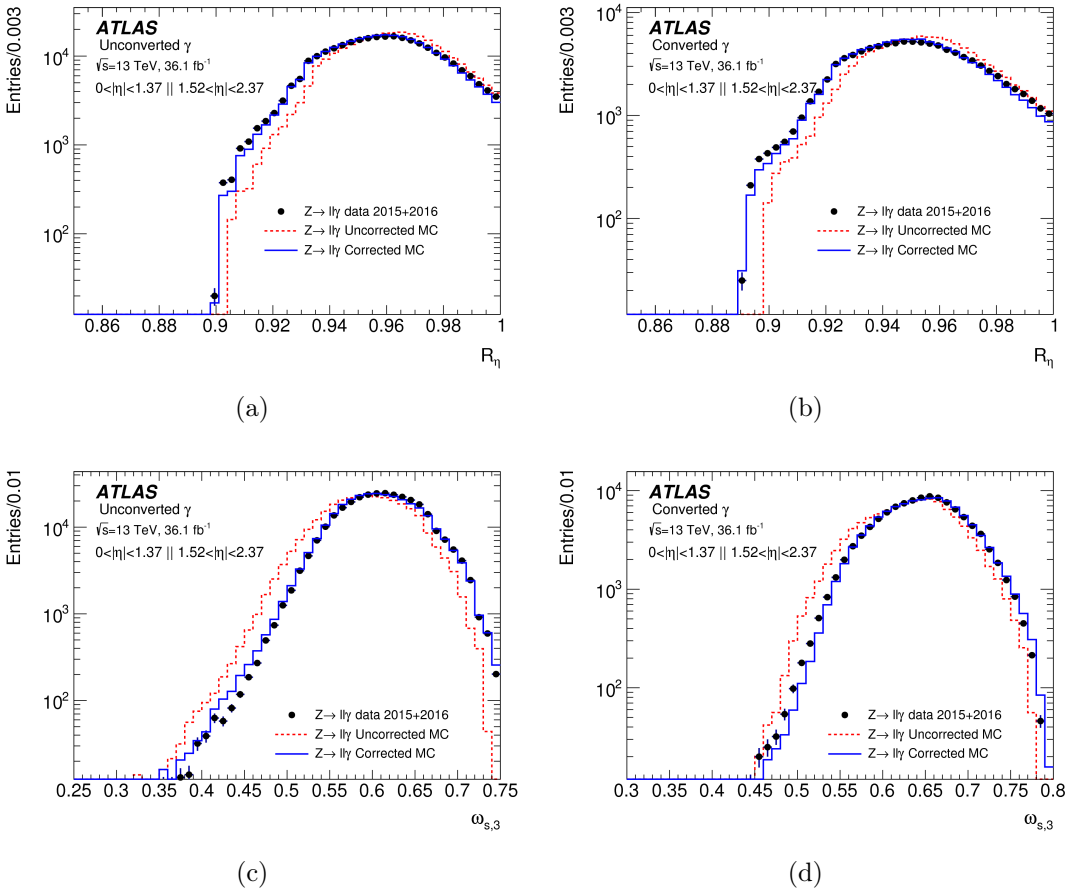


Figure 5.8: Distributions of calorimetric discriminating variables R_η and $w_{s\,3}$ for unconverted (left) and converted (right) photon candidates with $E_{\mathrm{T}}^{\gamma} \in [10, 50]$ GeV and $|\eta^{\gamma}| < 2.37$ (excluding $1.37 < |\eta^{\gamma}| < 1.52$) selected from $l^{+}l^{-}\gamma$ events (black dots). Distributions for true photons from simulated $Z \to ll\gamma$ events are also shown for the uncorrected simulation (dashed lines) and the simulation corrected by the average shift between data and simulation distributions (solid line). The definition of both discriminating variables is provided in table 5.1. From ref. [129].

Chapter 6

Jet reconstruction and calibration

In this chapter, the most important aspects of jet reconstruction are described, with the focus on the primary jet definition used in ATLAS physics results [26]. In addition, jet energy scale and resolution calibration procedures are described in detail, summarising the different stages and characteristics while providing a description of the most recent techniques employed in the ATLAS collaboration [26].

6.1 Jet reconstruction

The primary jet definition used in the studies presented in this thesis uses the anti-k_t [63] algorithm with a radius parameter $R = 0.4$ as implemented in the FASTJET 3.2.2 [140, 141] software package. This jet definition is the most widely used across the ATLAS collaboration. As described in section 3.3, the algorithm used to reconstruct jets can take as input different types of four-vector objects, such as stable particles defined by MC generators, charged-particle tracks, calorimeter energy deposits, or combinations of the latter two, as it is done in the particle-flow (PFlow) reconstruction technique [142].

Calorimeter cells used for jet reconstruction are first clustered into three-dimensional, massless topoclusters using a nearest-neighbour algorithm [132]. Cells are added to a topocluster if the ratio of the cell energy to the expected cell noise is above a certain threshold, which is defined to control the growth of each topocluster, as described in section 5.2.1. The resulting energy of the topocluster is defined at the EM scale. For jet reconstruction, only positive-energy topoclusters are used as inputs. Jets that are produced in the hard-scattering process are expected to originate from the primary vertex, which is defined as the reconstructed vertex with at least two associated tracks and the largest sum of squared track momentum. Therefore, an event-by-event correction to account for the position of the primary vertex in each event, referred to as the *origin correction*, is applied to every topocluster, based on its depth within the calorimeter and pseudorapidity [26]. Jets which are reconstructed using only calorimeter-based energy

information use the origin-corrected EM scale topoclusters, and are referred to as *EMtopo jets*. The origin correction was applied to the jet four-momentum rather than to its constituents in previous approaches [24, 25]; except for this update, the EMtopo was the primary jet definition used in ATLAS physics prior to the end of Run II.

With the purpose of improving hadronic final-state measurements by making a more complete use of the information from both the tracking and calorimeter systems, the use of jets reconstructed with the PFlow technique, referred to as *PFlow jets*, has been promoted in ATLAS by the end of Run II. The PFlow algorithm based on ref. [142] combines measurements from both the tracker and the calorimeter to form the input signals for jet reconstruction, which are intended to approximate individual particles. The energy deposited in the calorimeter by charged particles is subtracted from the observed topoclusters and replaced by the momenta of tracks that are matched to those topoclusters. The resulting PFlow jets, exhibit improved energy and angular resolution, reconstruction efficiency and pile-up stability compared to calorimeter jets [142]. The updates to the PFlow algorithm since its description provided in ref. [142] are described in ref. [26].

For the reconstruction of PFlow objects and calibration of both EMtopo and PFlow jets collections, tracks are reconstructed within the full acceptance of the inner detector ($|\eta| < 2.5$), required to have a $p_{\mathrm{T}} > 500$ MeV, and satisfy quality criteria based on the number of hits in the ID subdetectors. To suppress the effects of pile-up, tracks must be associated with the primary vertex. Tracks are matched to jets using ghost association [138], a procedure that treats them as four-vectors of infinitesimal magnitude during the jet reconstruction and assigns them to the jet with which they are clustered.

Particle-level jets are compared with PFlow and EMtopo jets to determine the energy scale and resolution. The particle-level jets used for calibration purposes are reconstructed using stable final-state particles and exclude muons, neutrinos, and particles from pile-up interactions, and they are selected with the same requirements on the transverse momentum and pseudorapidity as EMtopo and PFlow jets, $p_{\mathrm{T}} > 7$ GeV and $|\eta| < 4.5$. In addition, particle-level jets used in the calibration are geometrically matched to reconstructed EMtopo or PFlow jets with an angular distance $\Delta R < 0.3$.

6.2 Jet energy scale calibration

The jet energy scale calibration aims to restore the jet energy to that of jets reconstructed at the particle level. The global picture is illustrated in figure 6.1. The four-momentum of the jets is corrected in all the stages, where the jet p_{T}, energy, and mass are restored.

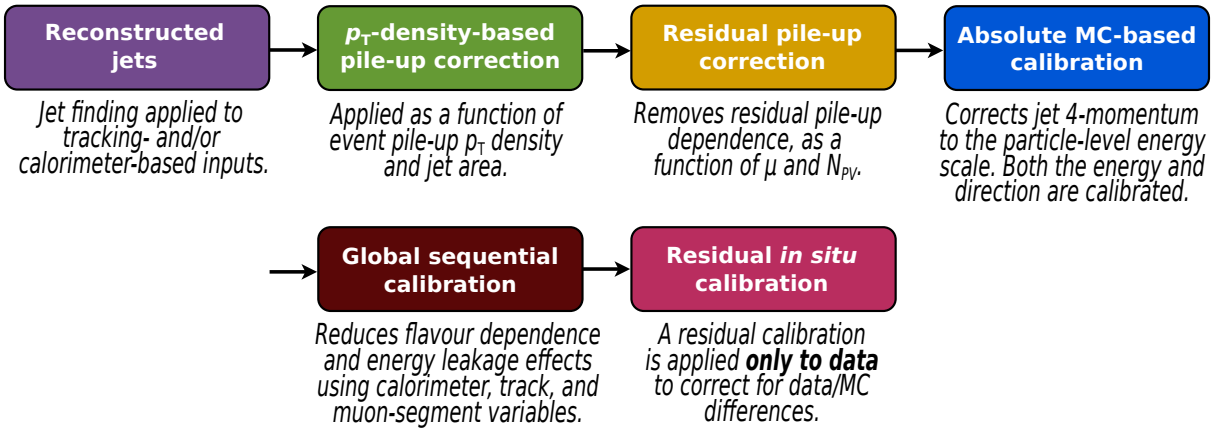


Figure 6.1: Stages of the jet energy scale calibration chain. Each one is applied to the four-momentum of the jet.

After the jet reconstruction step, pile-up corrections are applied, removing the excess energy due to additional pp interactions within the same or nearby bunch crossings. The corrections are divided into two components: one based on the jet area and transverse momentum density of the event, and a residual correction derived from MC simulation and parameterised as a function of $\langle\mu\rangle$ and N_{PV}. The *absolute* JES calibration corrects the jet energy and direction to that of particle-level jets. Furthermore, the Global Sequential Calibration (GSC) improves the jet p_{T} resolution and associated uncertainties by removing the dependence of the reconstructed jet response on observables constructed using information from tracking, calorimeter, and muon chamber detector systems. These corrections are described in section 6.2.1. Finally, a residual in situ calibration is applied to correct for the remaining differences between data and MC simulation. It is derived using well-measured reference objects such as photons, Z bosons, and calibrated jets. The in situ calibration methods are described in section 6.2.2.

6.2.1 Simulation-based jet calibrations

The MC-based calibrations are described below.

Pile-up corrections

A jet p_{T}-density-based subtraction of the per-event pile-up contribution to the jet p_{T} is performed in the first step. The jet area, A, is a measure of the susceptibility of the jet to pile-up, and is calculated by determining the relative number of ghost associated particles with a jet after jet clustering. The pile-up contribution is then estimated from the median p_{T} density, ρ, of the jets in the $y-\phi$ plane, $\langle p_{\mathrm{T}}/A\rangle$. The calculation of ρ is performed using the k_{t} jets with radius $R = 0.4$ from positive-energy topoclusters and $|\eta| < 2$.

The k_t algorithm is chosen for its sensitivity to soft radiation, and the median of the p_T density is used for ρ to reduce the bias from hard-scatter jets which populate the high-p_T tails of the distribution. An example MC distribution of ρ for representative values of N_{PV} is shown in figure 6.2. The ratio of the ρ-subtracted jet p_T to the uncorrected jet p_T is applied as a scale factor to the jet four-momentum, therefore not affecting the jet direction.

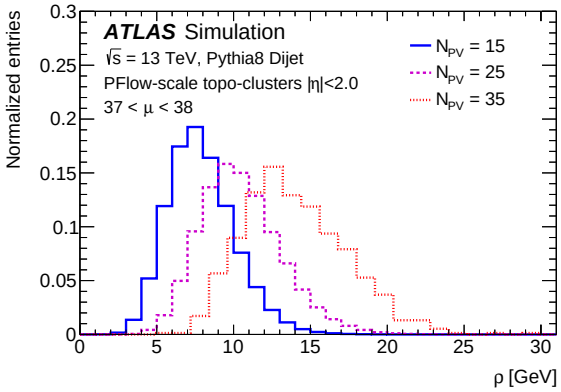


Figure 6.2: The per-event median p_T density, ρ, at $N_{PV} = 15$ (solid), $N_{PV} = 25$ (long dashed), and $N_{PV} = 35$ (short dashed) for $37 < \mu < 38$ obtained with MC simulation.

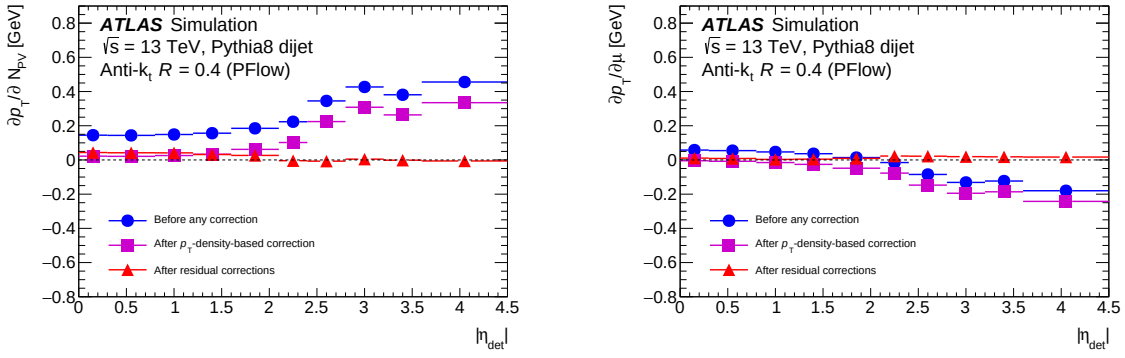
Since the ρ calculation is derived from the central, lower-occupancy regions of the calorimeter, it does not describe the pile-up sensitivity in the forward calorimeter region or in the higher-occupancy core of high-p_T jets. As a consequence, some dependence of the anti-k_t jet p_T on the pile-up activity remains, and a residual correction is needed. The residual pile-up correction is defined as the difference in p_T between the reconstructed and particle-level jets, and it is derived as a function of $\langle\mu\rangle$ and N_{PV}, providing sensitivity to in-time and out-of-time pile-up, respectively. The jet p_T after both the p_T-density-based and residual corrections is given by

$$p_T^{cor} = p_T^{reco} - \rho \times A - \alpha \times (N_{PV} - 1) - \beta \times \mu\,,$$

where p_T^{reco} refers to the p_T of the reconstructed jet before any pile-up correction is applied. The reconstructed and particle-level jets are geometrically matched with an angular distance $\Delta R < 0.3$. Independent linear fits are used to derive the α and β coefficients in bins of p_T^{true} and $|\eta_{det}|$, where p_T^{true} is the p_T of the particle-level jet that matches the reconstructed jet and $|\eta_{det}|$ is the jet η pointing to the geometric centre of the detector. In each $|\eta_{det}|$ bin, the fitted values of both coefficients at $p_T^{true} = 25$ GeV are taken as their nominal values, reflecting their behaviour in the p_T region where pile-up effects are most relevant.

The dependences of the p_T-density-based and residual corrections on N_{PV} and $\langle\mu\rangle$ as a function of $|\eta_{det}|$ for PFlow jets are shown in figure 6.3. Four systematic uncertainties

are introduced to account for MC-mismodelling of N_{PV}, $\langle\mu\rangle$, the ρ topology, and the p_{T} dependence of the residual pile-up corrections.



(a) (b)

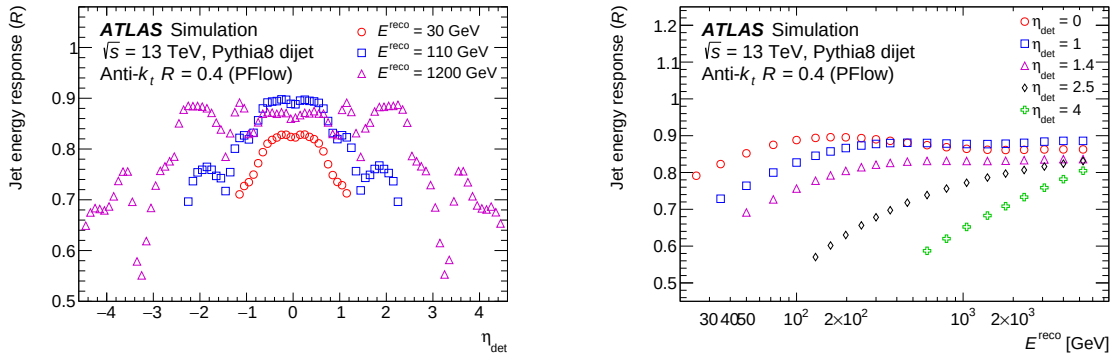
Figure 6.3: Dependence of PFlow jet p_{T} (a) in-time pile-up and (b) out-of-time pile-up as a function of $|\eta_{\mathrm{det}}|$ for $p_{\mathrm{T}}^{\mathrm{true}} = 25$ GeV.

Jet energy scale and η calibration

The absolute jet energy scale and η calibrations correct the reconstructed jet four-momentum to the particle-level energy scale, accounting for non-compensating calorimeter response, energy losses in dead material, out-of-cone (OOC) effects and biases in the jet η reconstruction caused by the transition between different calorimeter technologies and changes in calorimeter granularity. The calibration is derived for $R = 0.4$ anti-k_{t} jets in dijet events after the application of the pile-up corrections (see *Pile-up corrections* in section 6.2.1).

The average jet energy response, $\mathcal{R}$, defined as the mean of a Gaussian fit to the core of the $E^{\mathrm{reco}}/E^{\mathrm{true}}$ distribution, is measured in bins of E^{true} and η_{det}, where E^{reco} (E^{true}) is the energy of the reconstructed (particle-level) jet. The average response is parameterised as a function of E^{reco} using a numerical inversion procedure, as described in ref. [143]. The response for PFlow jets as a function of E^{reco} (η_{det}) for representative η_{det} (E^{reco}) bins is shown in figure 6.4.

A bias in the reconstructed jet η, defined as a significant deviation from zero in the signed difference between the reconstructed and particle-level jet η is observed and shown in figure 6.5 as a function of $|\eta_{\mathrm{det}}|$ for PFlow jets. It is largest in jets that encompass two calorimeter regions with more different energy responses due to the changes in calorimeter geometry and technology. Therefore, a second correction is derived as the difference between the reconstructed and particle-level η parameterised as a function of E^{true} and η_{det} to remove such a bias. A numerical inversion procedure is used again to derive the correction in E^{reco} bins. This calibration only alters the jet p_{T} and η, not the full four-momentum. EMtopo and PFlow jets calibrated with the full jet energy scale and η correction are considered to be at the EM+JES and PFlow+JES scale, respectively.



(a) (b)

Figure 6.4: The average energy response as a function of reconstructed jet η_{det} (a) and E^{reco} (b).

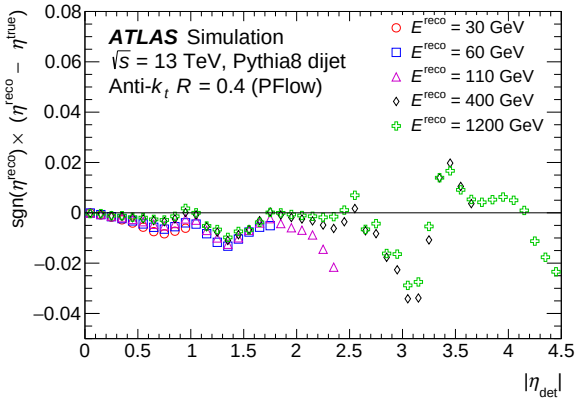


Figure 6.5: The signed difference between the reconstructed and particle-level jet η.

Global sequential calibration

Even after the application of the previous jet calibrations (from now on referred to as MCJES), for a given $(p_{\text{T}}^{\text{true}}, \eta_{\text{det}})$ bin, the response depends on the flavour and energy distribution of the jet constituents and the fluctuations of the jet development in the calorimeter. The average particle composition and shower shape of a jet varies between initiating particles, and more notably depending on whether the jet is quark or gluon-initiated. In the former case, the jet will often include hadrons with a higher fraction on the jet p_{T} that penetrate further into the calorimeter, while a gluon-initiated jet will typically contain more particles of softer p_{T}, leading to a lower calorimeter response and a wider transverse profile. The GSC is a series of multiplicative corrections that reduce the effects from these fluctuations and improve the jet resolution without changing the average jet energy response. The jet resolution, $\sigma_{\mathcal{R}}$, is defined as the standard deviation of a Gaussian fit to the jet p_{T} response distribution, which in turn is defined as the ratio of $p_{\text{T}}^{\text{reco}}$ to $p_{\text{T}}^{\text{true}}$.

The GSC is based on global jet observables such as the longitudinal structure of the energy depositions within calorimeters, tracking information associated with the jet, and information related to the muon chambers behind a jet. For this calibration, reconstructed jets are geometrically matched to particle-level jets and the same numerical inversion procedure as for the MCJES calibrations is used. Six observables are identified that improve the resolution of the JES through the GSC, providing an independent four-momentum correction as a function of $p_{\mathrm{T}}^{\mathrm{true}}$ and η_{det} for each observable. Each of the corrections is applied independently and sequentially to the jet four-momentum for jets with $|\eta| < 3.5$. The six stages of the GSC account for the dependence of the jet response on:

- f_{charged}, the fraction of the jet p_{T} measured from ghost-associated tracks with $p_{\mathrm{T}} > 500$ MeV ($|\eta_{\mathrm{det}}| < 2.5$).
- f_{Tile0}, the fraction of the jet energy measured in the first layer of the hadronic Tile calorimeter ($|\eta_{\mathrm{det}}| < 1.7$).
- f_{LAr3}, the fraction of the jet energy measured in the third layer of the EM LAr calorimeter ($|\eta_{\mathrm{det}}| < 3.5$).
- n_{trk}, the number of tracks with $p_{\mathrm{T}} > 1$ GeV ghost-associated with the jet ($|\eta_{\mathrm{det}}| < 2.5$).
- w_{trk}, also known as track-width, the average p_{T}-weighted transverse distance in the $\eta - \phi$ plane between the jet axis and all tracks of $p_{\mathrm{T}} > 1$ GeV ghost-associated with the jet ($|\eta_{\mathrm{det}}| < 2.5$).
- n_{segments}, also known as punch-through, the number of muon track segments ghost-associated with the jet ($|\eta_{\mathrm{det}}| < 2.7$).

The f_{charged} correction is only applied to PFlow jets. The n_{segments} correction reduces the tails of the response distribution caused by high-p_{T} jets that are not fully contained in the calorimeter. All corrections are derived as a function of the jet p_{T}, except the punch-through, which is derived as a function of the jet energy since this effect is more correlated with the energy escaping the calorimeters. The dependence of the jet response on each observable is reduced to less than 2% after the full GSC is applied.

The fractional jet resolution, defined as $\sigma_{\mathcal{R}}/\mathcal{R}$, is used to determine the size of fluctuations in the jet energy reconstruction and is shown for PFlow jets with $0.2 < |\eta_{\mathrm{det}}| < 0.3$ in MC simulation in figure 6.6. The fractional jet resolution improves as more corrections are applied, and the jet response dependence on the jet flavour is reduced.

6.2.2 In situ jet calibrations

After the application of the MCJES and GSC (from now on referred to as MCJES+GSC) calibrations, there are still differences between the jet response in data and simulation.

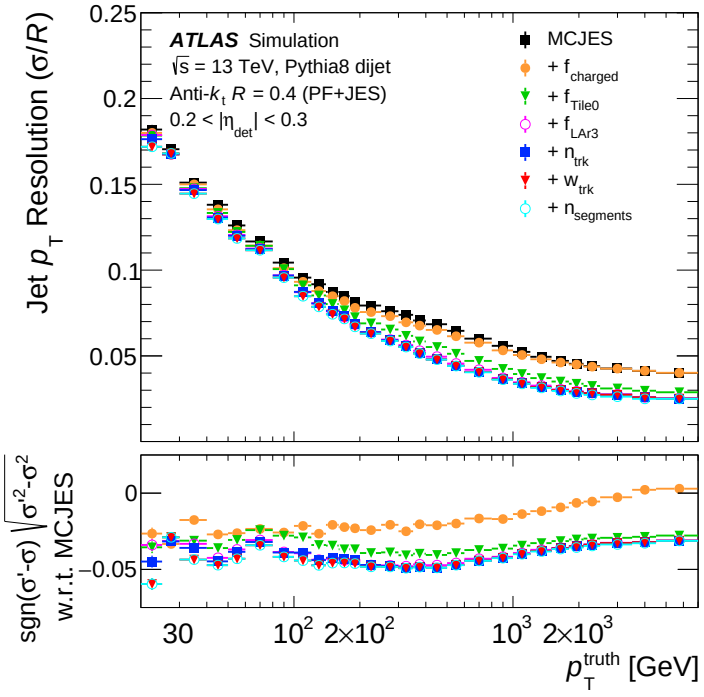


Figure 6.6: Resolution of jets at the PFlow+JES scale with $0.2 < |\eta_{\text{det}}| < 0.3$ measured in dijet MC simulation after each stage of the GSC. All jet flavours, including b-jets, are considered. The lower panel shows the difference in quadrature between the resolution before any GSC correction is applied (σ) and after the corresponding GSC step is applied (σ').

These discrepancies are caused by the imperfect simulation of both the detector materials and the physics processes involved. In order to account for these differences, a residual in situ calibration is derived, measuring the jet response in data and simulation separately and using the ratio as an additional correction to the data.

The jet response is calculated by balancing the p_{T} of a jet against that of a well-calibrated reference object or system. Then, the response $\mathcal{R}_{\text{insitu}}$ is defined as the average ratio of the jet p_{T} to the reference object p_{T} in bins of the reference object p_{T}. The average is taken from the peak location found by fitting the response distribution with a Gaussian function. The double ratio from the response in data and simulation

$$c = \frac{\mathcal{R}^{\text{data}}_{\text{in situ}}}{\mathcal{R}^{\text{MC}}_{\text{in situ}}},$$

is transformed via numerical inversion from a function of the reference object p_{T} to a function of the jet p_{T} (and jet η where applicable). The double ratio is more robust to secondary effects as long as they are well modelled in the simulation, and therefore constitutes a reliable measure of the jet energy scale difference between data and MC.

The in situ calibrations are divided in three stages. The first is the η intercalibration,

which corrects the energy scale of forward ($0.8 \leq |\eta| < 4.5$) jets to match that of central jets ($|\eta| < 0.8$) balancing the p_{T} in dijet events. In the second stage, the Z+jet and γ+jet analyses are performed, where the balance between the hadronic recoil in an event against the p_{T} of a calibrated Z or γ is used. Finally, the multijet balance (MJB) analysis uses a system of calibrated low-p_{T} jets to calibrate a single high-p_{T} jet. The Z/γ+jet and MJB analyses are derived only for central jets, but are also applicable to forward jets due to the effect of the η intercalibration. The three measurements are translated to a function of the jet p_{T} before being combined to provide a single smooth calibration applicable across the full momentum range.

Since the three in situ analyses are performed sequentially, systematic uncertainties are propagated from each to the next. Within each analysis, systematic uncertainties arise from three sources: the modelling of the physics processes in simulation, uncertainties in the measurement of the reference object and uncertainties in the expected p_{T} balance due to the event topology.

Relative calibration measurement in η using dijet events

The η intercalibration produces a correction which is applied to forward jets to bring them to the same energy scale as central jets, which are considered to be well-calibrated at the MCJES+GSC stage. Events with exactly two jets in different η regions of the detector are used, and there is no requirement for any of the jets to be in the central region; instead, all regions are calibrated relative to one another.

Dijet topologies are selected in which the two leading jets are back-to-back in ϕ and with no substantial contamination from a third jet. The jet p_{T} balance is quantified by the asymmetry, which is used as a metric for the response difference between the two detector regions

$$\mathcal{A} = \frac{p_{\mathrm{T}}^{\mathrm{left}} - p_{\mathrm{T}}^{\mathrm{right}}}{p_{\mathrm{T}}^{\mathrm{avg}}} ,$$

where $p_{\mathrm{T}}^{\mathrm{avg}} = (p_{\mathrm{T}}^{\mathrm{left}} + p_{\mathrm{T}}^{\mathrm{right}})/2$. The response ratio $\mathcal{R}$ of the two jets defines the calibration factor c for each jet

$$\mathcal{R} = \frac{c^{\mathrm{left}}}{c^{\mathrm{right}}} = \frac{2 + \langle \mathcal{A} \rangle}{2 - \langle \mathcal{A} \rangle} \simeq \frac{p_{\mathrm{T}}^{\mathrm{left}}}{p_{\mathrm{T}}^{\mathrm{right}}} .$$

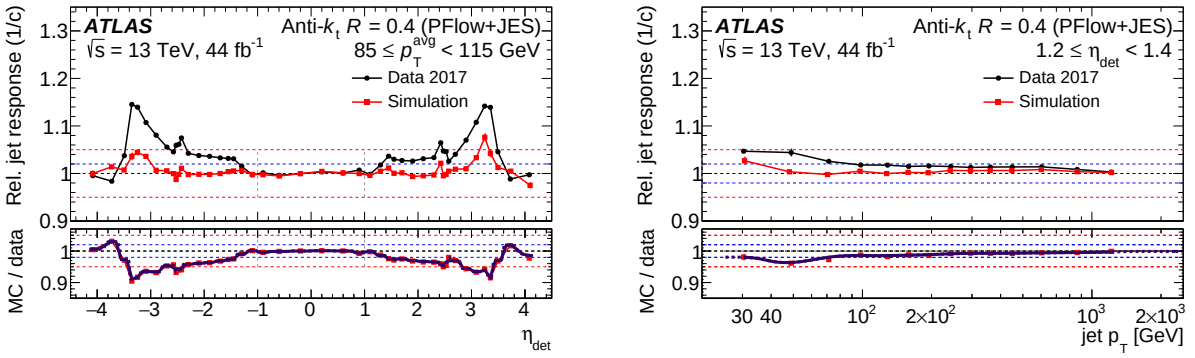
The average response ratio, $\langle \mathcal{R}_{\mathrm{ijx}} \rangle$, is measured in each bin i of η^{left}, j of η^{right}, and x of $p_{\mathrm{T}}^{\mathrm{avg}}$. The η variables are in detector coordinates rather than jet coordinates (η_{det}), since the properties of interest correlate to specific regions of detector hardware. The following function relates the correction factors and responses in each of the N bins

$$S\left(c_{1\mathrm{x}}, ..., c_{\mathrm{Nx}}\right) = \sum_{j=2}^{N} \sum_{i=1}^{j-1} \left(\frac{1}{\Delta \langle \mathcal{R}_{\mathrm{ijx}} \rangle} \left(c_{\mathrm{ix}} \langle \mathcal{R}_{\mathrm{ijx}} \rangle - c_{\mathrm{jx}} \right) \right)^2 + X\left(c_{\mathrm{ix}}\right) ,$$

where $\Delta\langle\mathcal{R}_{\mathrm{ijx}}\rangle$ is the statistical uncertainty in each bin. The function X (c_{ix}) quadratically imposes a penalty on correction factors deviation from 1. The correction factors to be used in the calibration are obtained by minimising this function.

The events used in this analysis may pass either a central jet trigger or a forward one, or both. The jets are required to satisfy the Jet Vertex Tagger (JVT) criteria to minimise contributions from pile-up and must pass basic cleaning requirements [144,145]. Each selected event must have two jets with $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 4.5$. To ensure a clean topology, events are further required to have no third jet with significant p_{T} via: $p_{\mathrm{T}}^{\mathrm{third}}/p_{\mathrm{T}}^{\mathrm{avg}} < 0.25$ and the two leading jets are required to be back-to-back in azimuthal plane such that $\Delta\phi > 2.5$ rad.

The calibration is derived in bins of $p_{\mathrm{T}}^{\mathrm{avg}}$ and η_{det} to guarantee both sufficient statistics in sparse regions and resolution of narrow detector features. Figure 6.7 shows the measured response in data and MC simulation for the 2017 dataset as a function of η_{det} for a particular $p_{\mathrm{T}}^{\mathrm{avg}}$ bin range 6.7(a), and as a function of $p_{\mathrm{T}}^{\mathrm{avg}}$ for a particular η_{det} range 6.7(b). It is observed that the simulation approximately reproduces the η_{det}-dependent features observed in data, although the response in data is consistently higher than observed in simulation.



(a) (b)

Figure 6.7: Relative response of jets calibrated with PFlow+JES in data (black circles) and MC (red squares). The response is shown as a function of η_{det} for jets with $85 \leq p_{\mathrm{T}}^{\mathrm{avg}} < 115$ GeV (a) and as a function of $p_{\mathrm{T}}^{\mathrm{avg}}$ for jets with $1.2 \leq \eta_{\mathrm{det}} < 1.4$ (b). The lower panel shows the response ratio of simulation to data (red squares) and the smoothed in situ calibration factor derived from the ratio (solid curve), which is used to perform the calibration correction. The dashed lines provide reference points for the viewer.

Calibration measurement using Z+jet and γ+jet events

The next stage of the in situ calibration corrects the differences between data and MC simulation using the momentum balance between the measured hadronic activity in the event and the p_{T} of a well-calibrated photon or Z boson.

The Z/γ+jet analyses rely on the energy scale of the photon or the electrons and

muons from the Z decay being well measured. The three objects are cleanly measured in the ATLAS detector and the uncertainties in their energy scales are small [131, 146]. The response is obtained separately in $Z \to e^+e^-$ and $Z \to \mu^+\mu^-$ events, and the three measurements are combined at a later stage. The Z+jet measurement is limited at moderate to high p_{T} by low statistics, and thus covers a range in p_{T} from 17 GeV to 1 TeV. The γ+jet measurement benefits from much higher statistics and extends to 1.2 TeV with small loss in sensitivity. However, this measurement is limited at low p_{T} due to the trigger prescales and the prevalence of soft jets misidentified as photons and so begins at 25 GeV.

The Missing-E_{T} Projection Fraction (MPF) technique is used in both the Z/γ+jet analyses, where the reference object p_{T} is balanced against the full hadronic recoil in an event. Therefore, the calorimeter response to hadronic showers is obtained directly. This technique is robust against UE and pile-up effects, since each of them cancels out directionally on average over a large number of events, and is not strongly affected by jet definitions since these become important only in the application of the calibration. Taking $p_{\mathrm{T}}^{\mathrm{recoil}}$ as the total transverse momentum of the hadronic activity in a clean Z/γ+jet event and $p_{\mathrm{T}}^{\mathrm{ref}}$ as the transverse momentum of the photon or the Z boson, the conservation of transverse momentum implies that

$$\vec{p}_{\mathrm{T,true}}^{\,\mathrm{ref}} + \vec{p}_{\mathrm{T,true}}^{\,\mathrm{recoil}} = 0\,. \tag{6.1}$$

The balance could be altered by the presence of ISR or FSR; this additional radiation is suppressed with cuts on the difference in the azimuthal angle, $\Delta\phi$, between the jet and the reconstructed photon or Z boson. If the calorimeter response to the hadronic activity in the event is r_{MPF} and the response for the calibrated reference object is 1, assuming that the missing energy in the event is due to the low response to the hadronic recoil, then, at detector level equation 6.1 transforms into

$$\vec{p}_{\mathrm{T}}^{\,\mathrm{ref}} + r_{\mathrm{MPF}}\,\vec{p}_{\mathrm{T}}^{\,\mathrm{recoil}} = -\vec{E}_{\mathrm{T}}^{\,\mathrm{miss}}\,. \tag{6.2}$$

Taking the projection in the direction of the reference object ($\hat{n}_{\mathrm{ref}}$) in equation 6.2, the response to the hadronic recoil is then seen to depend only on the missing energy in the event and the momentum of the reference object. The MPF response $\mathcal{R}_{\mathrm{MPF}}$ is defined by measuring r_{MPF} in bins of the reference object p_{T}, and is defined as follows

$$\mathcal{R}_{\mathrm{MPF}} = \left\langle 1 + \frac{\hat{n}_{\mathrm{ref}}\vec{E}_{\mathrm{T}}^{\,\mathrm{miss}}}{p_{\mathrm{T}}^{\mathrm{ref}}} \right\rangle , \tag{6.3}$$

where average response is extracted from the mean of the fit to the r_{MPF} distribution using a Gaussian distribution, in each corresponding $p_{\mathrm{T}}^{\mathrm{ref}}$ bin, and the response is mapped from reference to jet p_{T} by finding the average jet p_{T} in the events in each bin before the application of the MJB step.

The missing energy in each event is reconstructed from the calorimeter topoclusters in the case of the EMTopo jet calibration, while it is reconstructed from particle-flow objects in the case of the PFlow jet calibration, ensuring that the energy scale is consistent. The $Z \to ee$ ($Z \to \mu\mu$) events are required to pass a dielectron (dimuon) trigger with $p_{\mathrm{T}}^{e_1,e_2} > 15$ GeV ($p_{\mathrm{T}}^{\mu_1,\mu_2} > 14$ GeV). The electrons and muons that enter the analysis were required to have $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.47$ excluding the region $1.37 < |\eta| < 1.52$ for electrons, and $|\eta| < 2.4$ in the case of muons. In addition, all the Z+jet events are selected with a requirement on the electron or muon pair invariant mass to be close to the Z boson mass: $66 < m_{\mathrm{ee}/\mu\mu} < 116$ GeV. On the other hand, γ+jet events are selected using single-photon triggers with the lowest threshold corresponding to 15 GeV. Offline photons were required to have $E_{\mathrm{T}} > 25$ GeV and $|\eta^{\gamma}| < 1.37$, and satisfy tight identification and isolation criteria (see section 9.1.2).

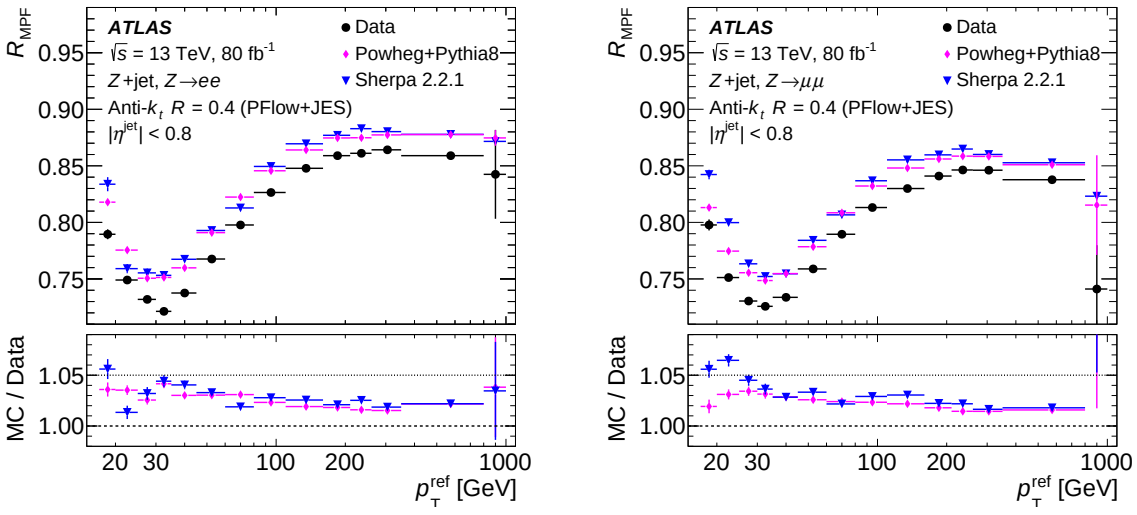
All jets within $\Delta R = 0.2$ of a photon or $\Delta R = 0.35$ of a lepton are removed. Jets must satisfy basic cleaning requirements and pass the JVT selection to suppress pile-up. Selected events were required to have one jet with $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 0.8$. Both Z+jet and γ+jet analyses have further selection requirements on the jets and event topology to suppress pile-up and ISR and FSR (see section 9.1.2).

The MPF responses calculated for data and two different MC generators in Z+jet and γ+jet events are shown in figures 6.8 and 6.9, respectively. The lower panels show the ratio of MC to data, and results using POWHEG+PYTHIA8 (Z+jet) and PYTHIA8 (γ+jet) constitute the nominal calibration.

Two small correction factors are derived in simulation and use the particle-level calorimeter response, defined as the ratio of measured energy in the calorimeter deposited by particles that belong to a particle-level jet to the total energy of the particle-level jet. The *topology correction* accounts for the differences in calorimeter response for sparse energy depositions versus those in dense cores of jets, and is estimated by taking the average of the ratio of $\mathcal{R}_{\mathrm{MPF}}$ to the particle-level calorimeter response in each p_{T} bin. The *showering correction* accounts for the flow of particles entering or exiting across the boundaries of the jet definition, and is calculated from the ratio of the particle-level calorimeter response to the measured response of the reconstructed jet, which depends on the jet algorithm and radius. The global correction factor accounting for both corrections was found to be less than 2% for jets of $p_{\mathrm{T}} < 50$ GeV and negligible above that. This correction factor would be applied identically to $\mathcal{R}_{\mathrm{MPF}}$ in both data and MC, and the effect would cancel out on the calibration double ratio. Therefore, only the uncertainty in the derivation of the correction factor is relevant, and it is estimated from the comparison of two different hadronic physics models (FTFP BERT [147] and QGSPBIC [148]) in the simulation of the particle/detector interactions.

An alternative approach for both Z/γ+jet analyses explicitly balancing the reference p_{T} against the p_{T} of a reconstructed jet is used to cross-check the jet energy scale calibration (see section 9.1.1).

High-p_{T} jet calibration using multijet balance



(a) (b)

Figure 6.8: Average PFlow jet response as a function of the reference p_{T} for Z+jet events where the Z boson decays into electrons (a) and into muons (b) calculated using the MPF technique. The two Z boson decay channels are combined at a later stage. The data are represented with black dots while the prediction of POWHEG+PYTHIA8 (SHERPA 2.2.1) is represented with pink diamonds (blue triangles). The ratio of MC to data for both generators is shown in the bottom panel and defines the in situ correction to be applied. The dotted lines at 1 and 1.05 serve as a reference.

In the final stage of the in situ calibration, a correction for jets with p_{T} above the range of the Z/γ+jet analyses is derived with the MJB technique. Events with a single high-p_{T} jet balanced against a system of low-p_{T} jets (the recoil system) are selected. The jets which form the recoil system are selected to ensure they are well calibrated using a combination of the Z/γ+jet results, while the leading jet is left at the scale of the η intercalibration. Therefore, the response of the system is defined as

$$\mathcal{R}_{\mathrm{MJB}} = \left\langle \frac{p_{\mathrm{T}}^{\mathrm{lead}}}{p_{\mathrm{T}}^{\mathrm{ref}}} \right\rangle ,$$

where $p_{\mathrm{T}}^{\mathrm{ref}}$ represents the vector sum of all jets in the recoil system. The response is again measured in bins of $p_{\mathrm{T}}^{\mathrm{ref}}$ and the correction is then mapped to the uncalibrated leading jet by finding the average $p_{\mathrm{T}}^{\mathrm{lead}}$ of the events in each bin.

Since the MJB analysis can only include events where all jets of the recoil system are well-calibrated, those events with very high-p_{T} leading jets are often excluded as their second and third leading jets can have a p_{T} outside the range of calibration covered by the Z/γ+jet analyses. This problem is addressed in the MJB analysis by proceeding in two iterations. In the first iteration, a combination of the Z/γ+jet results is used to calibrate

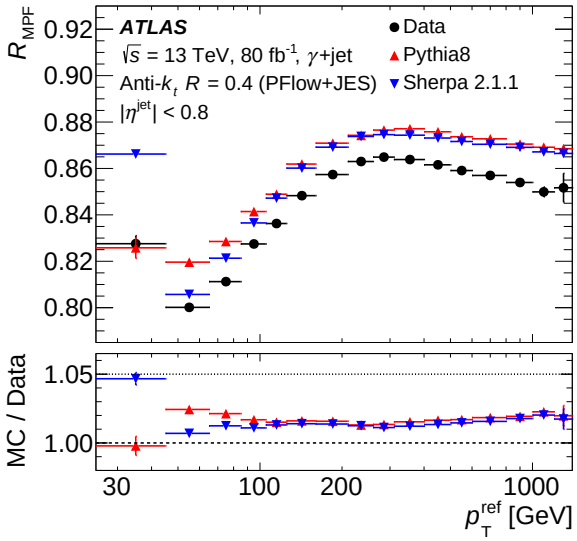


Figure 6.9: Average PFlow jet response as a function of the reference p_{T} for γ+jet events calculated using the MPF technique. The data are represented with black dots while the prediction of PYTHIA 8 (SHERPA 2.1.1) is represented with red (blue) triangles. The ratio of MC to data for both generators is shown in the bottom panel and defines the in situ correction to be applied. The dotted lines at 1 and 1.05 serve as a reference.

the jets of the recoil system, so that only events with subleading jets of $p_{\mathrm{T}} < 1.2$ TeV are included. In the second iteration, events with subleading jets up to $p_{\mathrm{T}} = 1.8$ TeV are included and calibrated using the MJB results from the first iteration. The range of the calibration is therefore extended to $p_{\mathrm{T}}^{\mathrm{lead}} = 2.4$ TeV.

Events are selected using a variety of single-jet triggers with each corresponding to a unique range of $p_{\mathrm{T}}^{\mathrm{lead}}$. At least three jets with $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.8$ are required. To suppress dijet topologies the subleading jet must not have a momentum above 0.8 $p_{\mathrm{T}}^{\mathrm{lead}}$. Jets are required to pass JVT selections to limit the effects of pile-up and the isolation of the leading jet from the recoil system is ensured by requiring that the difference in the azimuthal angle, $\Delta\phi$, between them is at least 0.3 rad.

The MJB response in data and four different MC generators is shown in figure 6.10. The lower panel shows the ratio of the response of each MC sample to data. The ratio measured using the SHERPA sample represents the nominal correction.

In situ combination

A data-to-MC response ratio of the form

$$\left\langle p_{\mathrm{T}}^{\mathrm{jet}}/p_{\mathrm{T}}^{\mathrm{ref}} \right\rangle_{\mathrm{data}} / \left\langle p_{\mathrm{T}}^{\mathrm{jet}}/p_{\mathrm{T}}^{\mathrm{ref}} \right\rangle_{\mathrm{MC}}$$

is measured in each of the four *absolute* in situ measurements of $Z(\to ee)$+jet, $Z(\to \mu\mu)$+jet, γ+jet, and the MJB. The different measurements are combined to produce a

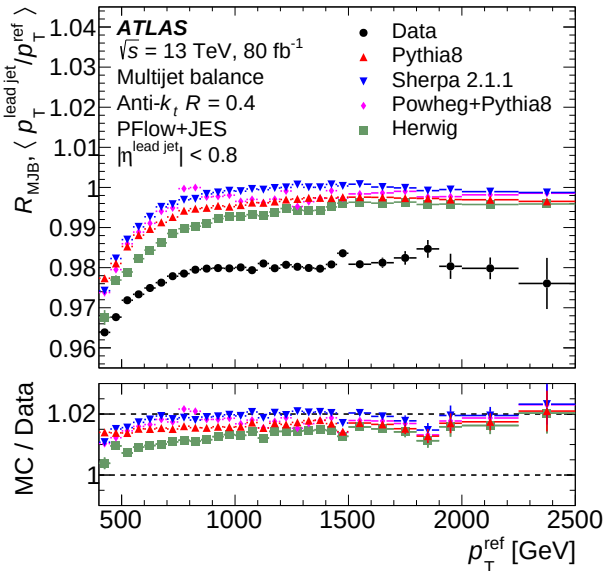


Figure 6.10: Average PFlow jet response as a function of the reference p_{T}. The response is shown for data and for simulation using four different MC generators, and the MC-to-data response ratios are shown in the bottom panel. The dotted lines at 1 and 1.02 serve as a reference.

single calibration covering the full range of jet p_{T} from 17 GeV to 2.4 TeV. Given the overlap of the measurements in various p_{T} ranges, the procedure must account for their relative statistical power and the tension between them in the same p_{T} range.

A detailed description of the combination procedure is provided in ref. [149]. Each of the absolute in situ measurements is converted from a parameterisation in terms of $p_{\mathrm{T}}^{\mathrm{ref}}$ into $p_{\mathrm{T}}^{\mathrm{jet}}$ and divided into finer bins of 1 GeV using second-order polynomial splines. A χ^2 minimisation is performed in each bin, taking as input the measurements available in that p_{T} range and their uncertainties, to determine a weight for each measurement in each bin. Therefore, the measurement with the smallest statistical and systematic uncertainties dominates the response estimation in a particular bin. The weights used for each input measurement are shown in figure 6.11. The final calibration curve is determined by smoothing the outputs from the minimisation with a Gaussian kernel.

The final in situ combination as a function of the jet p_{T} is shown in figure 6.12. The inverse of the solid curve ($R_{\mathrm{MC}}/R_{\mathrm{data}}$) is taken as the scaling factor and applied to data. The combined measurement for PFlow+JES jets is compared with each of the four individual measurements. The total size of the correction is approximately 3% at low p_{T} and decreases to around 2% for jets above 200 GeV.

Each uncertainty component from the in situ analyses is individually propagated through the combination procedure. First, the relevant measured response is varied by $1\,\sigma$ in the uncertainty component within its standard binning. After that, the finer rebinning, χ^2 minimisation, and combination procedure are repeated using instead the weights

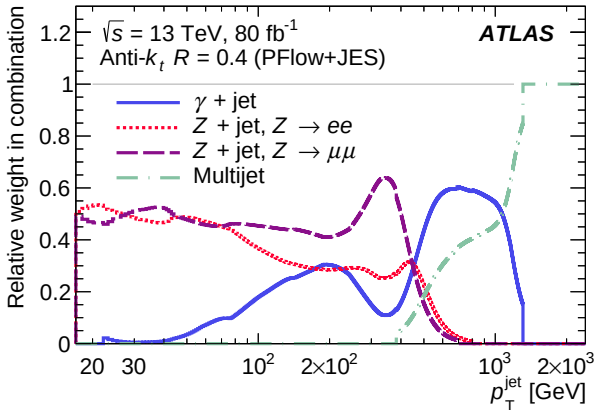


Figure 6.11: The weight assigned to different techniques in the combination of in situ measurements of the PFlow+JES jet response in data and simulation as a function of the jet p_{T}. For each p_{T} bin, the weights of the Z+jet, γ+jet and MJB methods are shown.

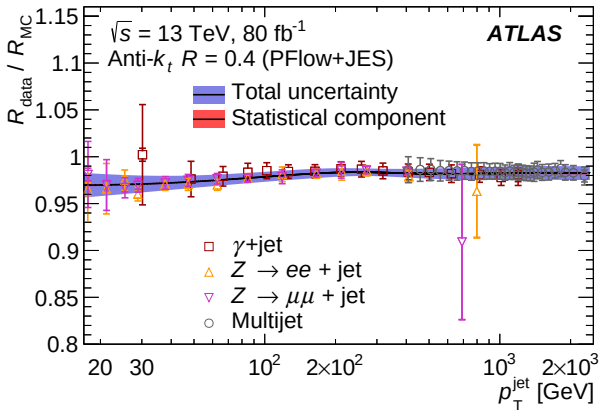


Figure 6.12: Ratio of the PFlow+JES jet response in data to that in the nominal MC event generators as a function of the jet p_{T} for the Z+jet, γ+jet and MJB in situ calibrations. The inner (outer) error bars represent the statistical (total) uncertainty of each of the measurements. The final correction and its statistical and total uncertainty bands are also shown.

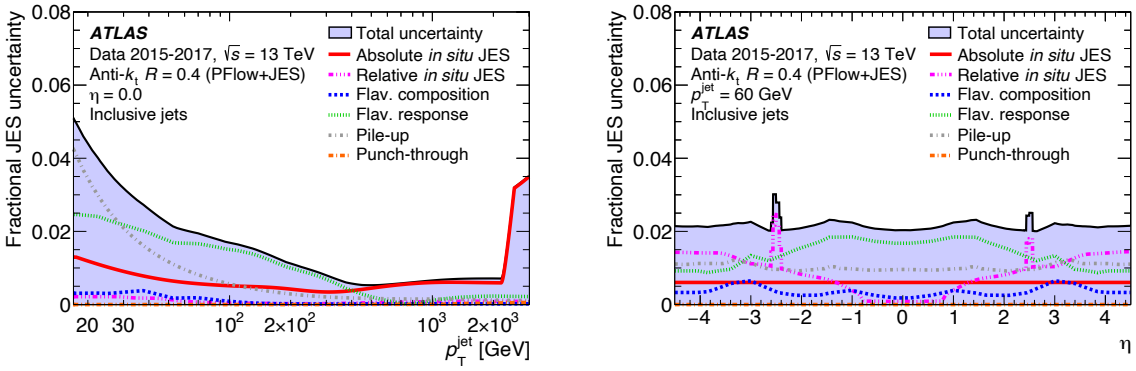
determined by the nominal result. The difference between the combined calibration curve with the systematically shifted input and the nominal calibration curve is taken as $1\,\sigma$ in the varied uncertainty.

6.2.3 Systematic uncertainties

The full uncertainty in the jet energy scale consists of 125 individual terms which are derived from the in situ measurements, pile-up effects, flavour dependence and other effects that are summarised in table 2 of ref. [26]. The majority of the components stem from

the in situ measurements, covering the effects of analysis selection cuts, event topology dependence, MC mismodelling, statistical limitations and the uncertainties associated with the calibration of electrons, muons and photons.

The total jet energy scale uncertainty is shown in figure 6.13(a) as a function of the jet p_{T} for fixed $\eta^{\mathrm{jet}} = 0$ and in figure 6.13(b) as a function of the jet η for fixed $p_{\mathrm{T}}^{\mathrm{jet}} = 60$ GeV. The uncertainties in the η intercalibration analysis are labelled as *relative in situ JES*. Uncertainties in all other in situ measurements are combined into the *absolute in situ JES* term, including also the single-particle uncertainty derived from studies of the response to individual hadrons and used to cover the region beyond 2.4 TeV.



(a) (b)

Figure 6.13: Fractional jet energy scale systematic uncertainty components for anti-k_{t} PFlow $R = 0.4$ jets as a function of the jet p_{T} at $\eta^{\mathrm{jet}} = 0$ (a) and as a function of the jet η at $p_{\mathrm{T}} = 60$ GeV. The total uncertainty, determined as the quadrature sum of all components, is shown as a filled region topped by a solid black line.

6.3 Jet energy resolution

Precise knowledge of the jet energy resolution (JER) is important for detailed measurements of SM jet production, measurements of the properties of the SM particles that decay into jets (e.g. W/Z bosons and top quarks) and searches for physics beyond the SM involving jets. The dependence of the relative JER on the jet p_{T} may be parameterised using a functional form expected for calorimeter-based resolutions, with three independent contributions, namely the noise (N), stochastic (S) and constant (C) terms [26]:

$$\frac{\sigma\left(p_{\mathrm{T}}\right)}{p_{\mathrm{T}}} = \frac{N}{p_{\mathrm{T}}} \oplus \frac{S}{\sqrt{p_{\mathrm{T}}}} \oplus C\,. \tag{6.4}$$

The noise term is due to pile-up and to the contribution of electronic noise to the signal measured by the detector front-end electronics. The noise term is expected to be significant in the low-p_{T} region, below ≈ 30 GeV. Statistical fluctuations in the amount of

deposited energy are captured by the stochastic term, which represents the limiting term in the resolution up to several hundred GeV in jet p_{T}. Finally, the constant term corresponds to fluctuations that are a constant fraction of the jet p_{T}, such as energy depositions in passive material, the starting point of the hadron showers, and non-uniformities of response across the calorimeter. The constant term is expected to dominate the high-p_{T} region, above ≈ 400 GeV.

To measure the JER, a precise measurement of the jet p_{T} is needed. Therefore, jets must either recoil against a well-measured reference object or be balanced against another jet in a well-defined dijet system. The latter approach has been followed by ATLAS in the most recent JER calibration [26].

6.3.1 Resolution measurement using dijet events

The JER measurements using dijet events for the balance method rely on the approximate scalar balance between the transverse momenta of the two leading jets. The deviations from the exact balance are measured with the asymmetry

$$\mathcal{A} = \frac{p_{\mathrm{T}}^{\mathrm{probe}} - p_{\mathrm{T}}^{\mathrm{ref}}}{p_{\mathrm{T}}^{\mathrm{avg}}}, \tag{6.5}$$

and are due to the experimental resolution, the presence of additional radiation in the event, and biases due to the event selection. In equation 6.5, $p_{\mathrm{T}}^{\mathrm{ref}}$ is the p_{T} of the reference jet required to be located in a well-calibrated region of the detector $0.2 \leq |\eta_{\mathrm{det}}^{\mathrm{ref}}| < 0.7$. The probe jet, however, can have $|\eta_{\mathrm{det}}^{\mathrm{probe}}| < 4.5$. The probe jet is the jet for which the resolution is to be measured and $p_{\mathrm{T}}^{\mathrm{avg}} = (p_{\mathrm{T}}^{\mathrm{probe}} + p_{\mathrm{T}}^{\mathrm{ref}})/2$. The standard deviation of $\mathcal{A}$ for a particular $\left(p_{\mathrm{T}}^{\mathrm{avg}}, \eta_{\mathrm{det}}^{\mathrm{probe}}\right)$ bin is denoted by $\sigma_{\mathcal{A}}$, and in the case of a measurement of the probe jet asymmetry, may be expressed as

$$\sigma_{\mathcal{A}}^{\mathrm{probe}} = \frac{\sigma_{p_{\mathrm{T}}}^{\mathrm{probe}} \oplus \sigma_{p_{\mathrm{T}}}^{\mathrm{ref}}}{\langle p_{\mathrm{T}}^{\mathrm{avg}} \rangle} = \left\langle \frac{\sigma_{p_{\mathrm{T}}}}{p_{\mathrm{T}}} \right\rangle^{\mathrm{probe}} \oplus \left\langle \frac{\sigma_{p_{\mathrm{T}}}}{p_{\mathrm{T}}} \right\rangle^{\mathrm{ref}},$$

where $\sigma_{p_{\mathrm{T}}}^{\mathrm{probe}}$ and $\sigma_{p_{\mathrm{T}}}^{\mathrm{ref}}$ are the standard deviations of $p_{\mathrm{T}}^{\mathrm{probe}}$ and $p_{\mathrm{T}}^{\mathrm{ref}}$, respectively. For calibrated jets, $\langle p_{\mathrm{T}}^{\mathrm{probe}} \rangle = \langle p_{\mathrm{T}}^{\mathrm{ref}} \rangle = \langle p_{\mathrm{T}}^{\mathrm{avg}} \rangle$ in the reference region. The reference jet relative resolution $\langle \sigma_{p_{\mathrm{T}}}/p_{\mathrm{T}} \rangle^{\mathrm{ref}}$, must therefore be subtracted from the measured asymmetry distribution to extract the energy resolution for the probe jet as

$$\left\langle \frac{\sigma_{p_{\mathrm{T}}}}{p_{\mathrm{T}}} \right\rangle^{\mathrm{probe}} = \sigma_{\mathcal{A}}^{\mathrm{probe}} \ominus \left\langle \frac{\sigma_{p_{\mathrm{T}}}}{p_{\mathrm{T}}} \right\rangle^{\mathrm{ref}}, \tag{6.6}$$

which is valid in the probe region as well, up to a correction factor accounting for the potential overall imbalance between the reference and the probe jets in that region. In

addition, the p_{T} balance of the measured jets is affected on an event-by-event basis by physics effects such as additional radiation, non-perturbative processes including hadronisation and multiparton interactions that may lead to particle losses and additions in the measured jets. Therefore, the measured dijet imbalance asymmetry distribution represents a convolution of the intrinsic detector resolution and the particle-level balance affected by the previous effects. The determination of $\sigma_{\mathcal{A}}^{\mathrm{probe}}$ may account for such effects by subtracting the particle-level quantity from the measured one

$$\left(\sigma_{\mathcal{A}}^{\mathrm{probe}}\right)_{\mathrm{det}} = \left(\sigma_{\mathcal{A}}^{\mathrm{probe}}\right)_{\mathrm{meas}} \ominus \sigma_{\mathcal{A}}^{\mathrm{truth}} . \tag{6.7}$$

An iterative fitting procedure is followed to extract the impact of these effects and to isolate the intrinsic detector resolution, $\left(\sigma_{\mathcal{A}}^{\mathrm{probe}}\right)_{\mathrm{det}}$, by assuming a Gaussian convolution of detector effects with the particle-level balance. Firstly, the asymmetry distribution measured at particle level in MC is fitted with an adhoc function $\mathcal{A}^{\mathrm{truth}}$ based on exponential curves. Secondly, the measured asymmetry distribution is fitted by the function

$$\mathcal{A}^{\mathrm{meas}} = \mathcal{A}^{\mathrm{truth}} \otimes \mathcal{R}\left(\mu_{\mathcal{A}}^{\mathrm{det}}, \sigma_{\mathcal{A}}^{\mathrm{det}}\right) ,$$

where $\mathcal{A}^{\mathrm{truth}}$ is taken from the particle-level fit and $\mathcal{R}\left(\mu_{\mathcal{A}}^{\mathrm{det}}, \sigma_{\mathcal{A}}^{\mathrm{det}}\right)$ is a Gaussian distribution with width $\sigma_{\mathcal{A}}^{\mathrm{det}}$ representing the detector resolution for the probe jet and offset $\mu_{\mathcal{A}}^{\mathrm{det}}$ accounting for any residual non-closure in the JES calibration.

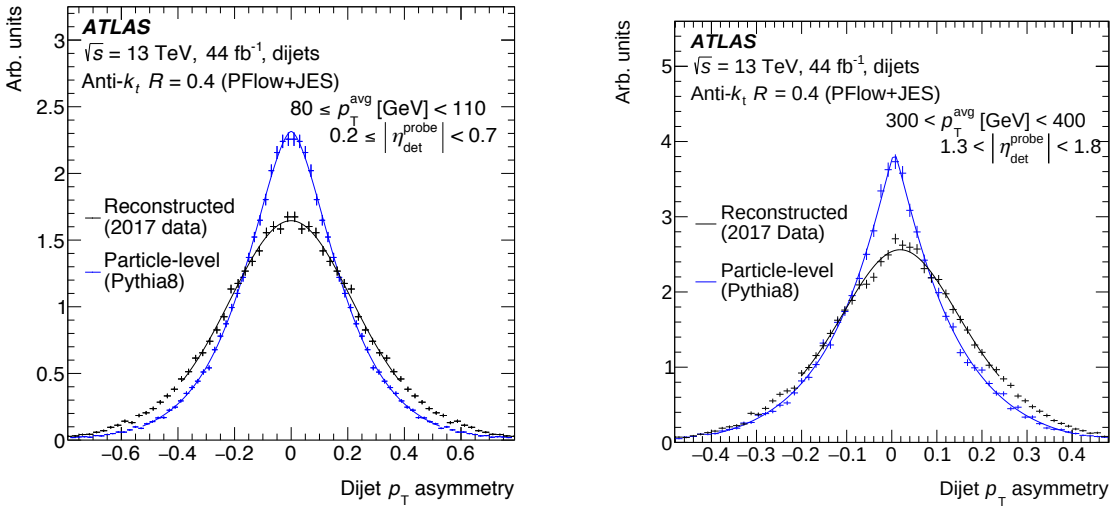
The final jet energy resolution measurement $\langle \sigma_{p_{\mathrm{T}}}/p_{\mathrm{T}} \rangle$ is calculated according to equation 6.6 after first correcting the asymmetry width $\sigma_{\mathcal{A}}$ according to equation 6.7.

The data used for the dijet balance measurement are collected using specific combinations of central and forward jet triggers for each of the $p_{\mathrm{T}}^{\mathrm{avg}}$ ranges of the measurement. Jets must also pass the JVT criteria. Topology cuts are applied to select well-defined dijet production processes with minimal contributions due to additional radiation or higher-order processes.

Example asymmetry distributions are shown in figure 6.14 in two representative bins of $p_{\mathrm{T}}^{\mathrm{avg}}$ and η_{det}. An iterative Gaussian fit to the core of the asymmetry is used to extract the JER. The result of the measurement of the relative JER and its systematic uncertainties is shown in figure 6.15 for a single narrow range of η_{det} and as a function of $p_{\mathrm{T}}^{\mathrm{avg}}$. The systematic uncertainties are dominated by imprecise knowledge of the scale of the jets at low p_{T} and the non-closure of the dijet balance method.

6.3.2 Noise measurement using random cones

A direct estimation of the noise term of equation 6.4 is obtained by measuring the fluctuations in the energy deposits due to pile-up using data samples that are collected y random unbiased triggers. These measurements are performed using the *random cones*



(a) (b)

Figure 6.14: Asymmetry distribution measured in data and particle-level MC for PFlow jets in two example p_{T} and η ranges. Error bars represent the statistical uncertainty. Jets with $80 \leq p_{\mathrm{T}}^{\mathrm{avg}} < 110$ GeV are shown in the range $0.2 \leq |\eta_{\mathrm{det}}^{\mathrm{probe}}| < 0.7$ (a), where the distributions are symmetric by construction. Jets with $300 < p_{\mathrm{T}}^{\mathrm{avg}} < 400$ GeV and in a range $1.3 < |\eta_{\mathrm{det}}| < 1.8$ where they can be asymmetric, are shown in (b).

method in which the energy deposits in the calorimeter are summed at the energy scale of the constituents in circular areas analogous to the jet area for anti-k_{t} $R = 0.4$ jets. This approach is selected due to its sensitivity to any non-Gaussian behaviour of the noise contributions to JER. The random fluctuations of the deposited energy $\Delta p_{\mathrm{T}}^{\mathrm{RC}}$ are measured taking the difference of two random cone sums, $p_{\mathrm{T}}^{\mathrm{c1}}$ and $p_{\mathrm{T}}^{\mathrm{c2}}$, at random ϕ values and within opposite $\pm\Delta\eta$ regions. A multiple selection of non-overlapping cones is performed within each event to maximise the statistical power. The random cone balance is given by

$$\Delta p_{\mathrm{T}}^{\mathrm{RC}} = p_{\mathrm{T}}^{\mathrm{c1}} - p_{\mathrm{T}}^{\mathrm{c2}} ,$$

and the estimation of the pile-up noise is determined by the central 68% confidence interval of the distribution of $\Delta p_{\mathrm{T}}^{\mathrm{RC}}$, σ_{RC}, sampled over many events as a function of both η and pile-up levels, as indexed by the number of interactions per bunch crossing, μ. The noise term due to pile-up is determined as

$$N^{\mathrm{PU}} = \frac{\sigma_{\mathrm{RC}}}{2\sqrt{2}} , \tag{6.8}$$

where the width of the distribution is divided by 2 to obtain the half-width of the distribution, and by $\sqrt{2}$ to obtain the fluctuations corresponding to just a single random cone.

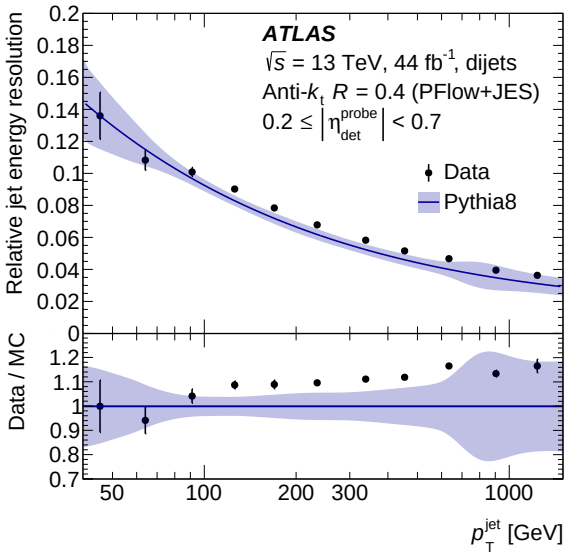
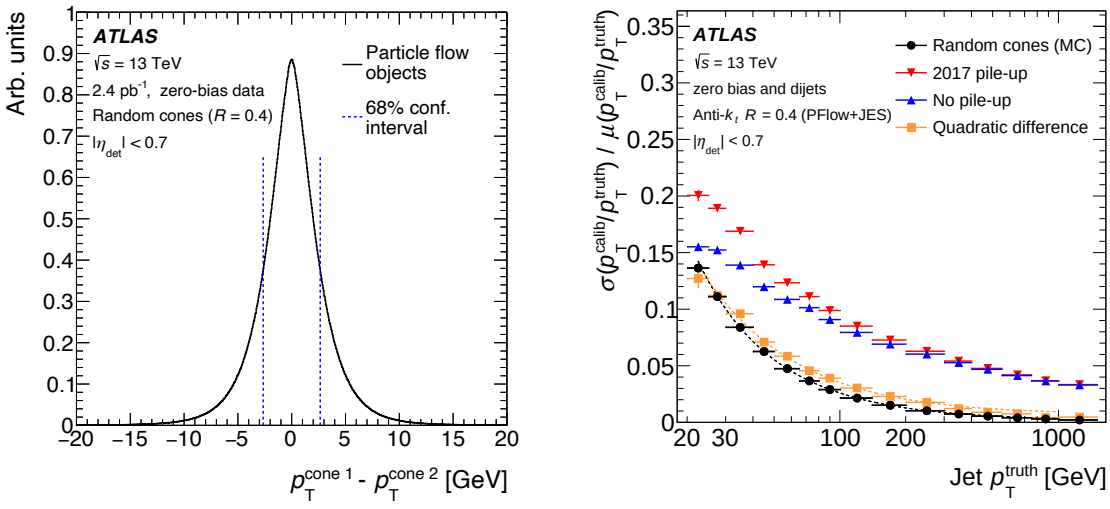


Figure 6.15: Relative jet energy resolution as a function of p_{T} for PFlow jets in the central region of the detector, measured using the dijet balance method. The resolution in data is shown in black points with error bars representing the statistical uncertainties, and the detector level MC resolution is shown by the blue curve with total systematic uncertainty given by the blue band.

The distribution of $\Delta p_{\mathrm{T}}^{\mathrm{RC}}$ is shown in figure 6.16(a). The energy scale of the noise term estimated in equation 6.8 is the constituent energy scale and cannot be compared with that of the jets measured in section 6.3.1. A conversion factor is required to perform that comparison, and the nominal JES calibration factor is used to perform this conversion to the appropriate energy scale.

A closure test of the random cone measurements is performed by comparing the in situ measurement of the calibrated N^{PU} with the expectation from MC simulation. The contribution to the JER due to pile-up noise in MC is estimated determining the JER with simulated events both with and without pile-up. A subtraction in quadrature is performed between the extracted resolutions. The two JER determinations in MC with and without pile-up are shown in figure 6.16(b), and their quadratic difference is compared with the in situ measurement from the random cones method. The random cone measurement is fitted with N/p_{T}, while the quadratic difference is fitted with $N/p_{\mathrm{T}} \oplus S/\sqrt{p_{\mathrm{T}}}$ to account for the non-negligible stochastic contributions.

The total noise contribution to the JER includes electronic noise in addition to the pile-up. Since the random cones are not sensitive to electronic noise due to the topocluster formation, a fit is performed to the JER measured in a dedicated MC sample with $\mu = 0$ and the electronic noise term is extracted as $N^{\mu=0}$. The total noise term used in the JER combination is therefore taken as $N = N^{\mathrm{PU}} \oplus N^{\mu=0}$, and it is shown as a function of $|\eta|$ in figure 6.17 along with its systematic uncertainties.



(a) (b)

Figure 6.16: The difference in the random cone sums, $\Delta p_{\mathrm{T}}^{\mathrm{RC}}$, measured in the central region ($|\eta_{\mathrm{det}}| < 0.7$) in randomly triggered data using PFlow objects (a). Comparison between the pile-up noise term N^{PU} determined using the random cones method (black solid circles) and the expectation from the MC simulation (orange squares) as extracted from the difference in quadrature of MC simulation with (red downward triangles) and without (blue upward triangles) pile-up. Results are shown at the PFlow+JES energy scale for jets in the central region of the detector ($|\eta_{\mathrm{det}}| < 0.7$).

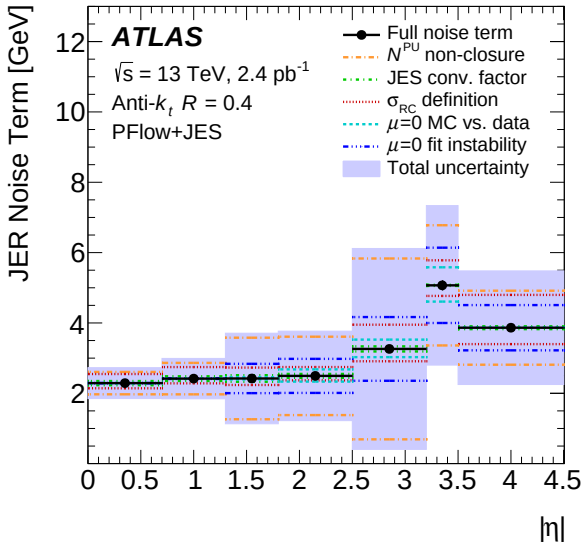
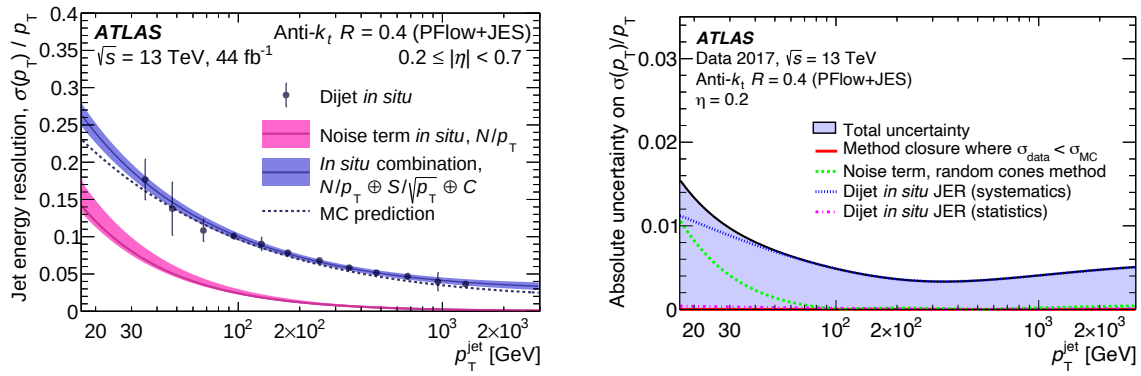


Figure 6.17: Noise term due to pile-up estimated using the random cones method and its uncertainty as a function of $|\eta|$.

6.3.3 Combination of *in situ* JER measurements

A combined measurement of the JER is obtained performing a fit to the dijet balance measurement using a constraint on the noise term (N) derived from the random cones measurement and $\mu = 0$ simulation sample. The implementation of the statistical combination is performed in a similar manner as the JES combination described in section 6.2.2. The functional form of equation 6.4 is used to interpolate across $p_{\mathrm{T}}^{\mathrm{jet}}$. A fit to the dijet measurement data is performed, fixing the noise term to the value measured by the random cones analysis.

The resulting combined measurement of the JER for PFlow+JES jets is shown in figure 6.18(a). The dijet measurement is also shown along with the total in situ combination, while the constraint on the noise term derived from the random cones is demonstrated by plotting N/p_{T} and its uncertainties for illustrative purposes. Figure 6.18(b) shows the absolute uncertainties on the combined JER measurement.



(a) (b)

Figure 6.18: The relative jet energy resolution as a function of p_{T} for fully calibrated PFlow+JES jets (a). The expectation from MC is compared with the relative resolution as evaluated in data through the combination of the dijet balance and random cones techniques. The absolute uncertainty on the relative jet energy resolution as a function of the jet p_{T} (b). Uncertainties from the two in situ measurements and from the data/MC difference are shown separately.

Chapter 7

Measurement of the inclusive isolated-photon cross section at $\sqrt{s} = 13$ TeV using $36\,\text{fb}^{-1}$ of ATLAS data

This chapter presents the measurement of the isolated-prompt photon production in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector at the LHC using an integrated luminosity of $36.1\,\text{fb}^{-1}$, which was published in ref. [37]. The differential cross section as a function of the photon transverse energy was measured in different regions of the photon pseudorapidity for $E_{\text{T}}^{\gamma} > 125$ GeV and $|\eta^{\gamma}| < 2.37$. The differential cross section as a function of the absolute value of the photon pseudorapidity was also measured in different regions of photon transverse energy. NLO pQCD calculations from JETPHOX and SHERPA as well as NNLO pQCD calculations from NNLOJET were compared with the measurement, using several parameterisations of the PDF. The predictions were found to provide a good description of the data within the experimental and theoretical uncertainties.

7.1 Data selection

The data used in this analysis were collected with the ATLAS detector during the pp collision running periods of 2015 and 2016, when the LHC operated at a centre-of-mass energy of $\sqrt{s} = 13$ TeV. Only events taken during stable beam conditions and satisfying detector- and data-quality requirements, which include the calorimeters and inner tracking detectors being in nominal operation, are considered. In the collected data sample, the average number of pp interactions per bunch crossing is ≈ 24, while the total integrated luminosity is $36.1 \pm 0.8\,\text{fb}^{-1}$, out of which $3.22 \pm 0.07\,\text{fb}^{-1}$ corresponds to the 2015 running period. The uncertainty in the combined $2015-2016$ integrated luminosity is 2.1% [150],

obtained using the LUCID-2 detector [151] for the primary luminosity measurements.

Trigger requirements

The data sample used consists of events triggered by a single-photon high-level trigger with a nominal transverse energy threshold of 120 GeV (140 GeV) for the 2015 (2015 and 2016) data-taking period and loose photon identification requirements[1] [39, 131]. The single-photon trigger with the threshold of 120 GeV was kept unprescaled during 2015 whereas it had to be prescaled during 2016 due to the increase in instantaneous luminosity. The single-photon trigger with the threshold of 140 GeV was the lowest-threshold unprescaled photon trigger during the 2016 data-taking period. The selection criteria applied by the trigger on the shower-shape variables computed from the energy profiles of the showers in the calorimeters are looser than the photon identification criteria applied in the offline analysis and allow a plateau of constant efficiency close to 100%, for true prompt photons of $E_{\mathrm{T}}^{\gamma} > 125$ GeV and $|\eta^{\gamma}| < 2.37$, as shown at the end of this section.

Event selection

The sample of isolated-photon events was selected offline using the following criteria:

- The events were required to have at least one reconstructed primary vertex, with at least two associated tracks of $p_{\mathrm{T}} > 400$ MeV, consistent with the average beam-spot position.
- Events with problems associated to noise burst and data-integrity errors in the LAr calorimeter were rejected. This requirement was applied only to data.
- Events for which a part of the detector information was missing were removed from the sample. This requirement was applied only to data.
- Events were rejected if the tile calorimeter output was corrupted. This requirement was applied only to data.
- Events with SCT problems were rejected. This requirement was applied only to data.

Photon selection

Photon candidates were reconstructed from clusters of energy deposited in the EM calorimeter and classified as unconverted photon candidates or converted photon candidates following the criteria described in ref. [129]. The main background in the prompt-photon production measurement comes from high-energy π^0 or η mesons which can be misidentified as a photon because they decay into an almost collinear photon pair. Such energetic

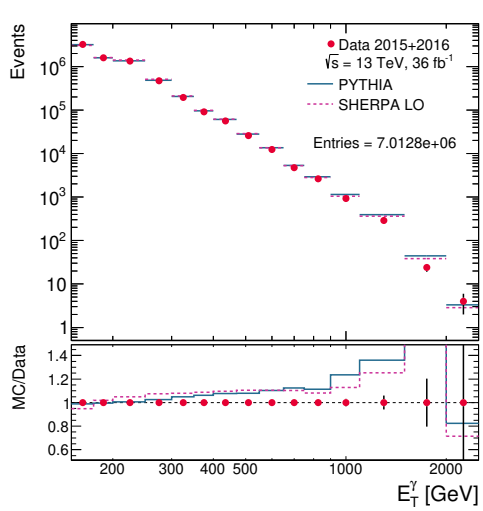
[1] The discriminating variables used for loose and tight photon identification can be found in table 5.1.

π^0 or η mesons are produced copiously inside jets. This background is reduced by the photon identification criteria and by requiring the photon candidate to be isolated.

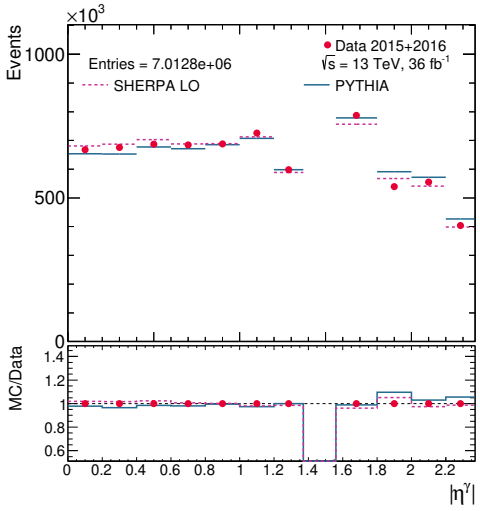
The photon-candidate selection criteria proceed as follows:

- The starting point of the selection are photons reconstructed from EM clusters and using tracking information provided by the ID [129]. Both converted and unconverted photons were kept. Photons reconstructed near regions of the calorimeter affected by read-out or high-voltage failures were rejected. The photon-candidate energy was calibrated with the procedures described in ref. [130].
- Events in which the photon with the highest transverse energy (leading photon) had $E_{\mathrm{T}}^{\gamma} > 125$ GeV and $|\eta^{\gamma}| < 2.37$ were selected. Photons in the region $1.37 < |\eta^{\gamma}| < 1.56$ were rejected. Photon candidates with $125 < E_{\mathrm{T}}^{\gamma} < 150$ GeV ($E_{\mathrm{T}}^{\gamma} > 150$) GeV were selected from the 2015 (2015 and 2016) data-taking period.
- In events with multiple candidates satisfying these requirements, the candidate with highest transverse energy (leading photon) was retained for further study.
- The candidates were required to pass the loose′ (LoosePrime4) identification criteria defined in section 5.4.
- The leading photon was required to pass the tight identification criteria defined in section 5.4.
- The isolation transverse energy, $E_{\mathrm{T}}^{\mathrm{iso}}$, of the leading photon, computed from topo-clusters inside a cone of size $\Delta R = 0.4$ around the photon candidate and corrected for photon energy leakage and ambient transverse energy (as described in section 5.5), was required to be smaller than $E_{\mathrm{T,cut}}^{\mathrm{iso}} = 4.2 \times 10^{-3} \times E_{\mathrm{T}}^{\gamma} + 4.8$ GeV, which amounts to a cut on $E_{\mathrm{T}}^{\mathrm{iso}}$ of 5.3 (15.3) GeV for $E_{\mathrm{T}}^{\gamma} = 125$ (2500) GeV.
- Additional corrections were applied to the simulated events to match the overall event conditions of the data sample and to account for known differences between data and simulation, as described in section 5.6.

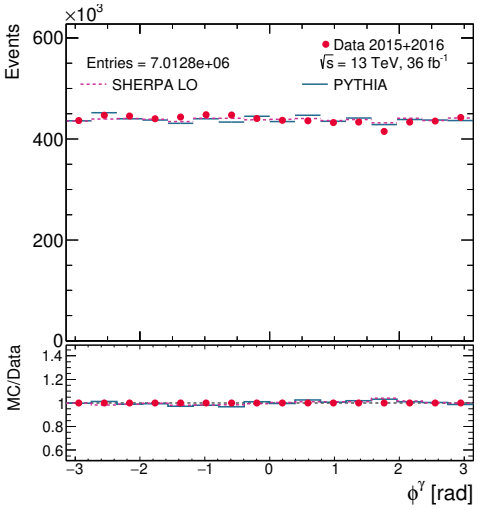
The total number of selected events with a photon candidate with $E_{\mathrm{T}}^{\gamma} > 150$ GeV ($125 < E_{\mathrm{T}}^{\gamma} < 150$ GeV) is 7 012 797 (650 410). Figure 7.1 shows the E_{T}^{γ}, $|\eta^{\gamma}|$ and ϕ^{γ} data and MC distributions. The kinematic requirements and the number of selected events in data in each $|\eta^{\gamma}|$ region are summarised in table 7.1. Each region in $|\eta^{\gamma}|$ is divided into 16 bins of E_{T}^{γ} starting at $E_{\mathrm{T}}^{\gamma} = 125$ GeV and ending at $E_{\mathrm{T}}^{\gamma} = 2500$ GeV. The binning is given by the following array of values of E_{T}^{γ} (in units of GeV): 125, 150, 175, 200, 250, 300, 350, 400, 470, 550, 650, 750, 900, 1100, 1500, 2000, and 2500. Some of the high-E_{T}^{γ} bins are not measured depending on the $|\eta^{\gamma}|$ region.



(a)



(b)



(c)

Figure 7.1: The measured E_{T}^{γ} (a), $|\eta^{\gamma}|$ (b) and ϕ^{γ} (c) distributions for events with $E_{\mathrm{T}}^{\gamma} >$ 150 GeV (dots). For comparison, the MC simulations of the signal from PYTHIA (solid histograms) and SHERPA (dashed histograms) are also included. The MC simulations are normalised to the number of data events. The lower panel of each figure shows the ratio of the MC distributions to the data.

	Phase space region			
Requirement on E_{T}^{γ}	$E_{\mathrm{T}}^{\gamma} > 125$ GeV			
Isolation requirement	$E_{\mathrm{T}}^{\mathrm{iso}} < 4.2 \times 10^{-3} \times E_{\mathrm{T}}^{\gamma} + 4.8$ GeV			
Requirement on $\lvert\eta^{\gamma}\rvert$	$\lvert\eta^{\gamma}\rvert < 0.6$	$0.6 < \lvert\eta^{\gamma}\rvert < 1.37$	$1.56 < \lvert\eta^{\gamma}\rvert < 1.81$	$1.81 < \lvert\eta^{\gamma}\rvert < 2.37$
Number of events with $125 < E_{\mathrm{T}}^{\gamma} < 150$ GeV	182 754	248 538	74 405	144 713
Number of events with $E_{\mathrm{T}}^{\gamma} > 150$ GeV	2 030 144	2 696 077	814 623	1 471 953

Table 7.1: Definition of the phase-space region for the measurement and predictions. The (next-to) last row indicates the number of data events selected in each $|\eta^{\gamma}|$ region for $E_{\mathrm{T}}^{\gamma} > 150$ GeV ($125 < E_{\mathrm{T}}^{\gamma} < 150$ GeV). For $125 < E_{\mathrm{T}}^{\gamma} < 150$ GeV only data from the 2015 running period are used, while for $E_{\mathrm{T}}^{\gamma} > 150$ GeV data from the 2015 and 2016 running periods are used.

Trigger efficiency, uncertainty and correction

A correction to account for the trigger inefficiency at turn-on was applied to the data. The correction for $E_{\mathrm{T}}^{\gamma} > 150$ GeV was evaluated using the HLT_g120_loose lower-threshold trigger, which was unprescaled (prescaled) during the 2015 (2016) running period. This correction was also evaluated for the $125 < E_{\mathrm{T}}^{\gamma} < 150$ GeV bin, using the HLT_g100_loose trigger, which was prescaled during the 2015 period. For $E_{\mathrm{T}}^{\gamma} > 150$ GeV, the correction factor is

$$C_{\mathrm{trigger}} = \frac{N_{\mathrm{HLT_g120_loose}}}{N_{\mathrm{HLT_g120_loose} \otimes \mathrm{HLT_g140_loose}}},$$

where $N_{\mathrm{HLT_g120_loose}}$ is the number of data events that pass all the selection criteria listed previously in this section and fulfil the conditions of the trigger HLT_g120_loose; $N_{\mathrm{HLT_g120_loose} \otimes \mathrm{HLT_g140_loose}}$ is the number of data events that pass all the selection criteria listed previously and fulfil the conditions of both triggers HLT_g120_loose and HLT_g140_loose. Figures 7.2 to 7.5 show the estimated correction for the turn-on as a function of E_{T}^{γ} in the different $|\eta^{\gamma}|$ regions for tight isolated, tight non-isolated, non-tight isolated, and non-tight non-isolated photons (these categories are defined in section 7.3). The factor C_{trigger} was applied to the data in each E_{T}^{γ} bin of each $|\eta^{\gamma}|$ region, according to the (non-)tight and (non-)isolated regions, to correct the data for the turn-on. For the $125 < E_{\mathrm{T}}^{\gamma} < 150$ GeV bin, the correction factor C_{trigger} was derived following the same procedure but using the trigger HLT_g100_loose (HLT_g120_loose) instead of the HLT_g120_loose (HLT_g140_loose).

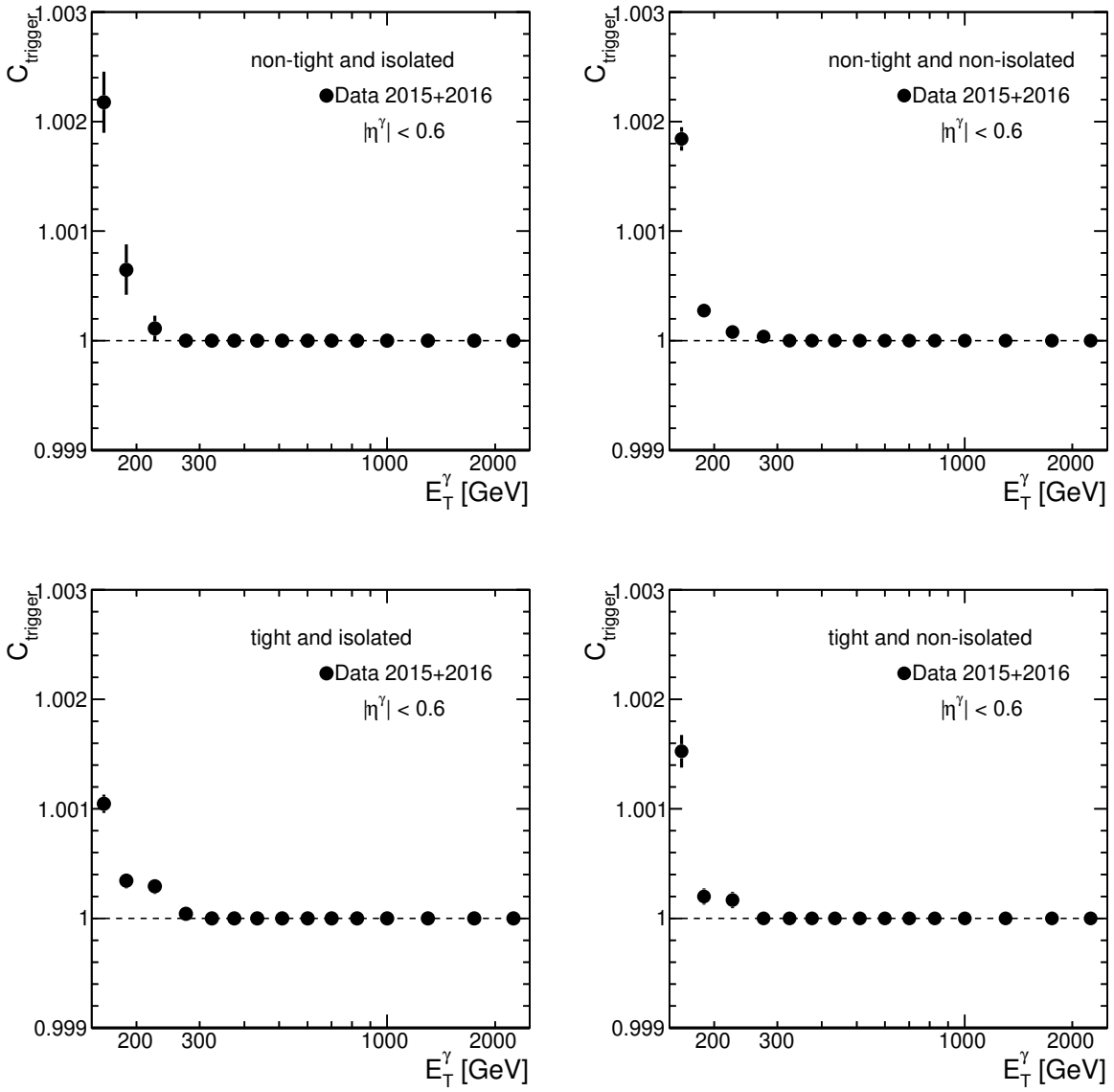


Figure 7.2: Correction for trigger inefficiency in data for $|\eta^\gamma| < 0.6$ as a function of E_{T}^γ for non-tight and isolated (control region C), non-tight and non-isolated (control region D), tight and isolated (signal region A), and tight and non-isolated (control region B) photons.

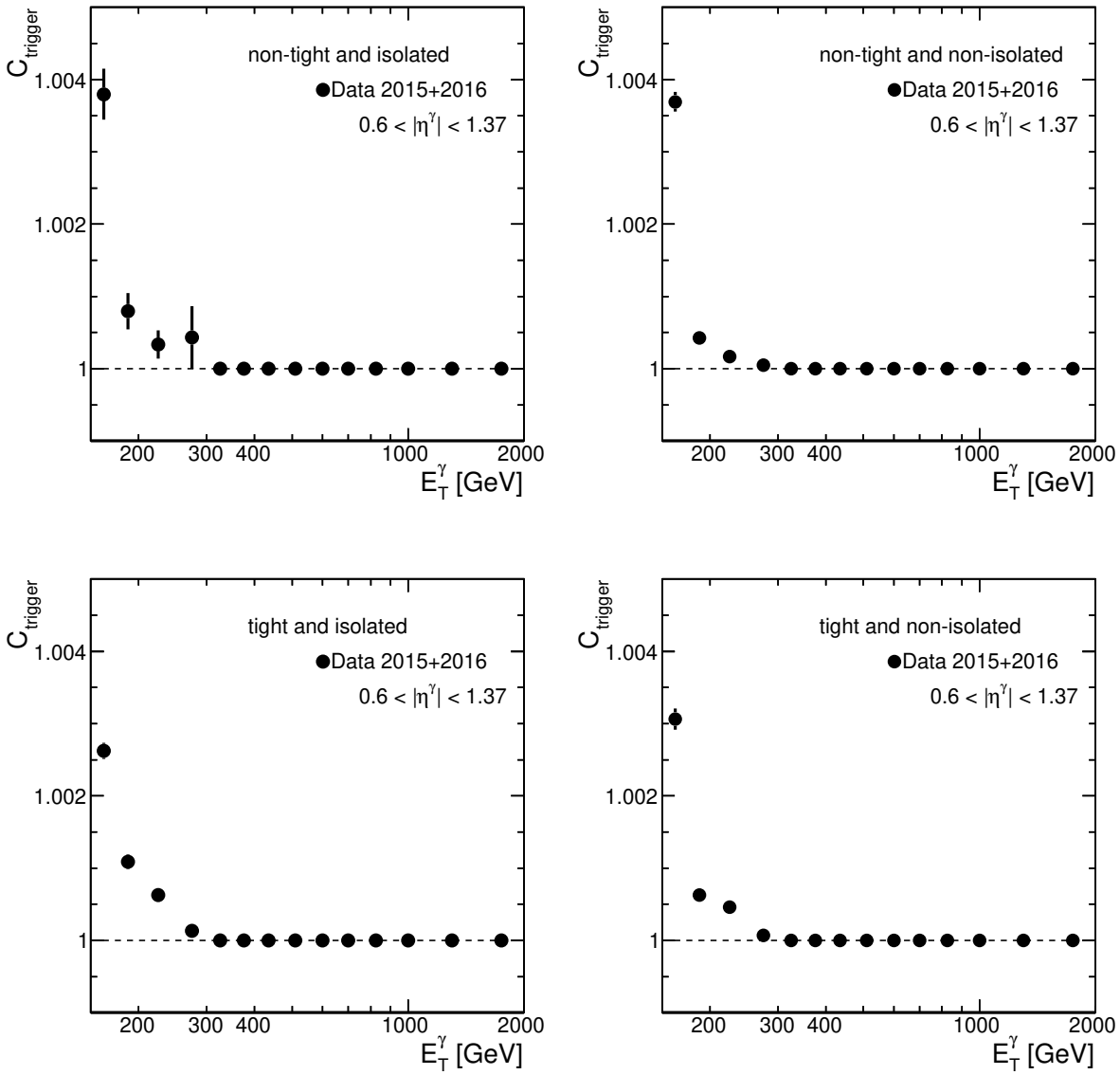


Figure 7.3: Correction for trigger inefficiency in data for $0.6 < |\eta^{\gamma}| < 1.37$ as a function of E_{T}^{γ} for non-tight and isolated (control region C), non-tight and non-isolated (control region D), tight and isolated (signal region A), and tight and non-isolated (control region B) photons.

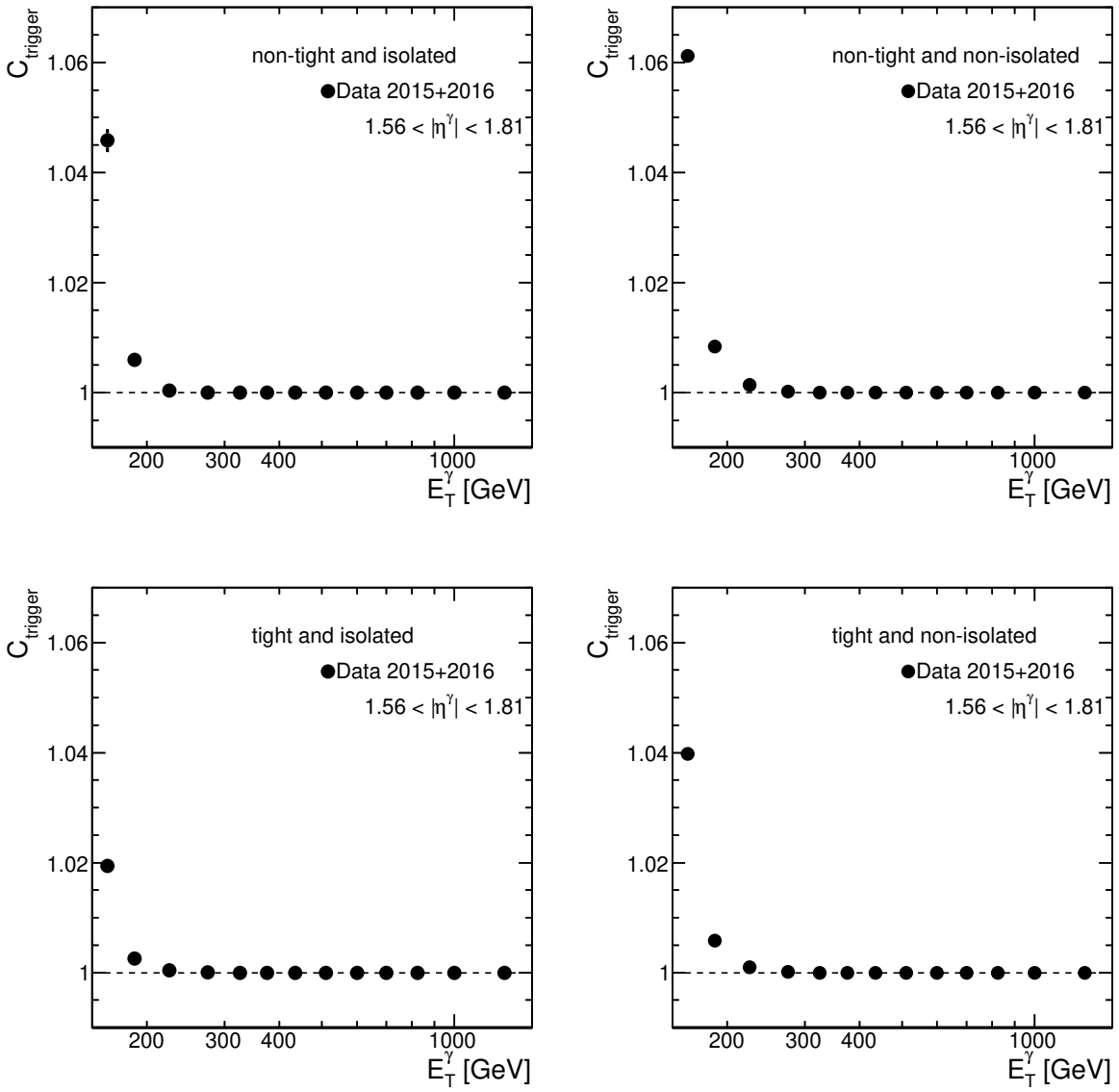


Figure 7.4: Correction for trigger inefficiency in data for $1.56 < |\eta^{\gamma}| < 1.81$ as a function of E_{T}^{γ} for non-tight and isolated (control region C), non-tight and non-isolated (control region D), tight and isolated (signal region A), and tight and non-isolated (control region B) photons.

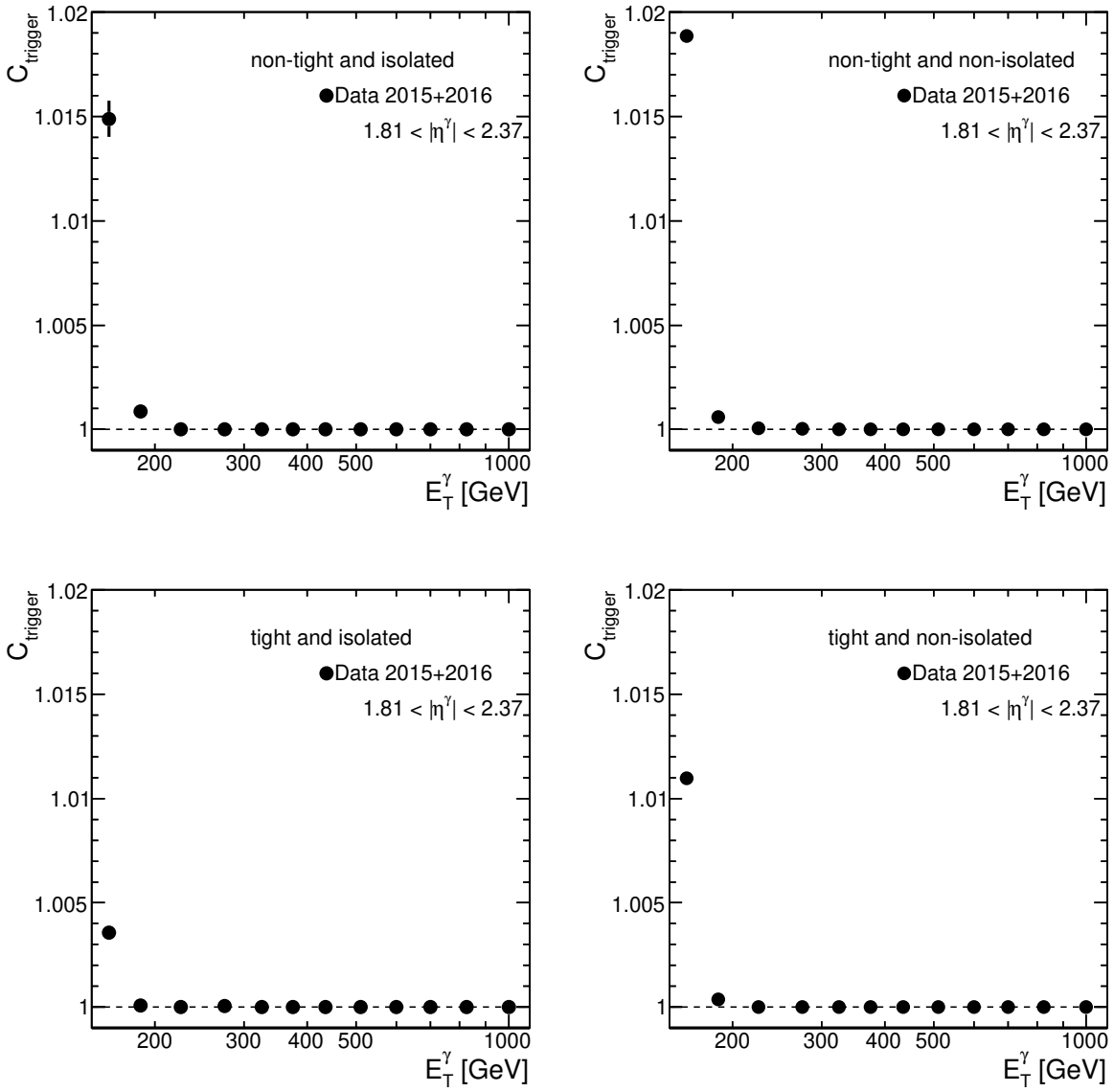


Figure 7.5: Correction for trigger inefficiency in data for $1.81 < |\eta^{\gamma}| < 2.37$ as a function of E_{T}^{γ} for non-tight and isolated (control region C), non-tight and non-isolated (control region D), tight and isolated (signal region A), and tight and non-isolated (control region B) photons.

After applying these corrections to the data, the efficiencies of the HLT_g120_loose and HLT_g100_loose triggers were evaluated with respect to a L1 trigger, as explained in ref. [152]. The efficiency of the HLT_g120_loose trigger for data is shown in figure 7.6. The uncertainty in the trigger efficiency was estimated as the sum in quadrature of the trigger inefficiency and the statistical uncertainty. The values propagated into the cross section are shown in table 7.2.

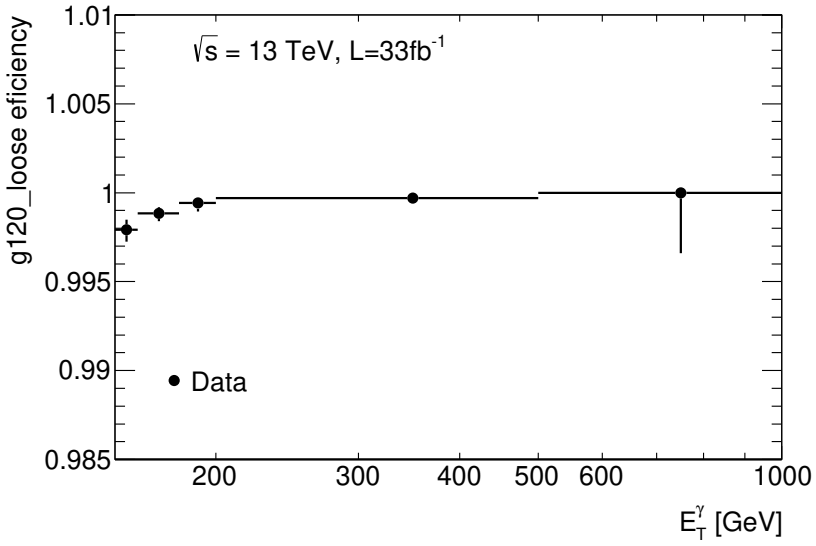


Figure 7.6: Efficiency of the HLT_g120_loose trigger as a function of E_{T}^{γ} for data.

E_{T}^{γ} region (GeV)	Relative uncertainty (%)
125 – 150	0.27
150 – 175	0.22
175 – 200	0.08
200 – 470	0.04
470 – 2500	0.40

Table 7.2: Estimated trigger uncertainty for each E_{T}^{γ} region

7.2 Monte Carlo simulations

Samples of MC events were generated to study the characteristics of the signal events, to determine the correction factors necessary to obtain the measured cross section and to compute the predicted cross sections at particle level. In addition, MC samples were also used to estimate hadronisation and UE corrections for the NLO and NNLO pQCD calculations. The different samples used for this analysis are described in detail in section 4.3.

As described in section 5.6, the MC events were reweighted such that the distribution of $\langle\mu\rangle$ matches the one in the data. Figure 7.7 shows the distribution of $\langle\mu\rangle$ for the data after applying a scale factor of 1/1.09 (see section 5.6), and for the PYTHIA simulation (before and after the MC is reweighted).

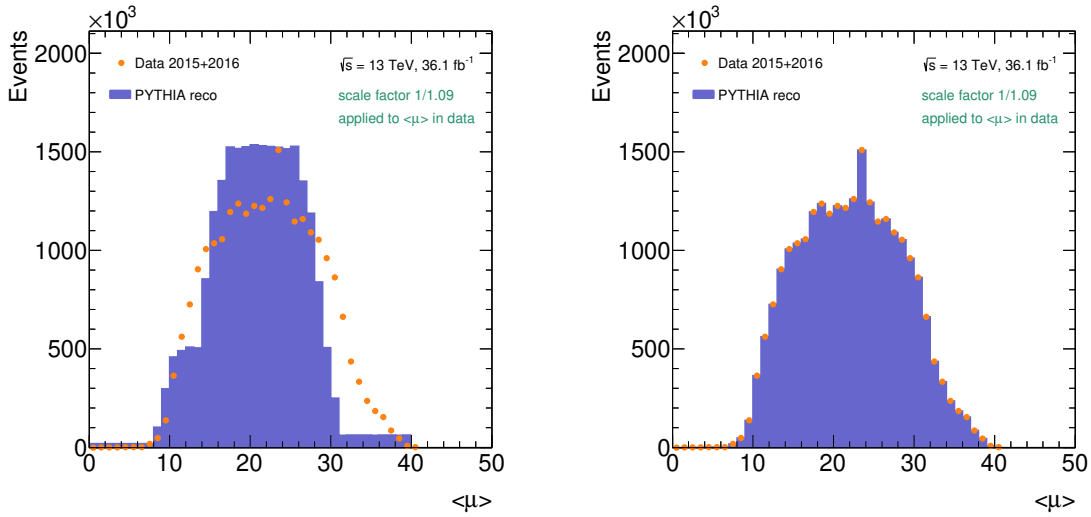
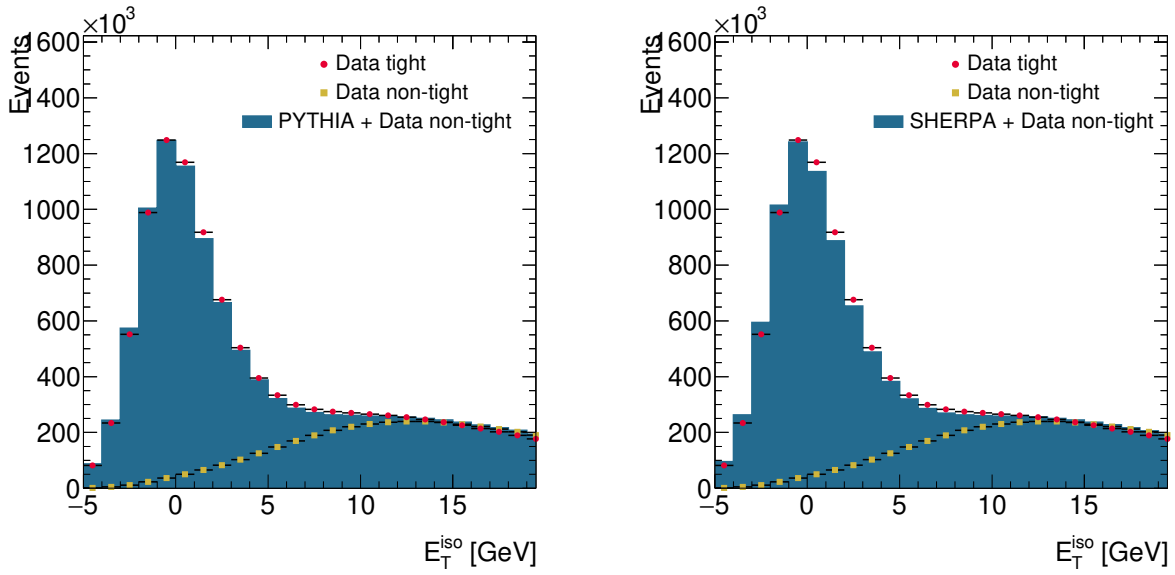


Figure 7.7: Distribution of the average number of interactions per bunch crossing in data (dots) and PYTHIA MC (histogram) before (left) and after (right) reweighting.

7.3 Background evaluation and signal extraction

A non-negligible amount of background remains in the selected sample even after the application of the tight identification and isolation requirements described in sections 5.4 and 5.5, respectively. This background originates predominantly from multijet processes in which a jet is misidentified as a photon; this jet usually contains a light neutral meson, mainly a π^0, which decays into two collimated photons which carry most of the energy of the jet.

The isolation profile of a signal-enriched sample was extracted from the data. For this study, a sample was obtained applying the selection criteria described in section 7.1, except for the tight identification and isolation requirements, and for photons with $E_{\mathrm{T}}^{\gamma} > 150$ GeV. Two subsamples were selected by applying the tight identification criteria: the subsample of candidates fulfiling the requirements (tight subsample) and the subsample of candidates that pass the loose′ criteria but fail the tight requirements (non-tight subsample). The non-tight subsample is expected to be enriched in background candidates. The $E_{\mathrm{T}}^{\mathrm{iso}}$ distributions for these subsamples are shown in figure 7.8. The $E_{\mathrm{T}}^{\mathrm{iso}}$ distribution of the non-tight candidates was scaled so that the integral for $E_{\mathrm{T}}^{\mathrm{iso}} > 10$ GeV, where the contribution from the signal is expected to be small, matched that of the tight candidates. The isolation profile is compared to the MC simulations of PYTHIA 7.8(a) and SHERPA 7.8(b). It is observed that the MC simulations exhibit a peak which reproduces that in the data; this is a consequence of applying the data-driven corrections on $E_{\mathrm{T}}^{\mathrm{iso}}$ to the MC simulations as discussed in section 5.5, otherwise the peak would be somewhat shifted towards lower $E_{\mathrm{T}}^{\mathrm{iso}}$ values with respect to the data (see appendix C).



(a) (b)

Figure 7.8: The measured $E_{\mathrm{T}}^{\mathrm{iso}}$ distribution before the isolation requirement and after applying the tight identification requirement (red dots) and for those events which fail the tight identification (golden squares). The non-tight distribution was normalised so that the integral of the tight and non-tight distributions for $E_{\mathrm{T}}^{\mathrm{iso}} > 10$ GeV coincides. For comparison, the MC simulations of the signal from PYTHIA (a) and SHERPA (b) (solid histogram) are also included. The non-tight contributions were added to the MC distributions.

The estimation of the background from the signal MC samples is not reliable. Therefore, a background subtraction method was devised which does not rely on MC samples of background and uses signal-suppressed control regions. The same *2D-sideband* subtraction method applied in ref. [9, 10] was used. The method is briefly explained below.

The background contamination was estimated and then subtracted using a counting technique based on the observed number of events in control regions of a two-dimensional plane. This plane was defined by using the photon identification (γ_{ID}) and the $E_{\mathrm{T}}^{\mathrm{iso}}$ variables, as shown in figure 7.9. These two variables were chosen because they are expected to be uncorrelated for the background events. Therefore, four regions were defined in this plane:

– Signal region A: contains tight and isolated ($E_{\mathrm{T}}^{\mathrm{iso}} < 4.2 \times 10^{-3} \times E_{\mathrm{T}}^{\gamma} + 4.8$ GeV) photon candidates.

– Control region B: contains tight and non-isolated ($E_{\mathrm{T}}^{\mathrm{iso}} > 4.2 \times 10^{-3} \times E_{\mathrm{T}}^{\gamma} + (4.8 + 2)$ GeV) photon candidates. An upper cut on $E_{\mathrm{T}}^{\mathrm{iso}}$ of 50 GeV is also imposed.

– Control region C: contains non-tight and isolated ($E_{\mathrm{T}}^{\mathrm{iso}} < 4.2 \times 10^{-3} \times E_{\mathrm{T}}^{\gamma} + 4.8$ GeV)

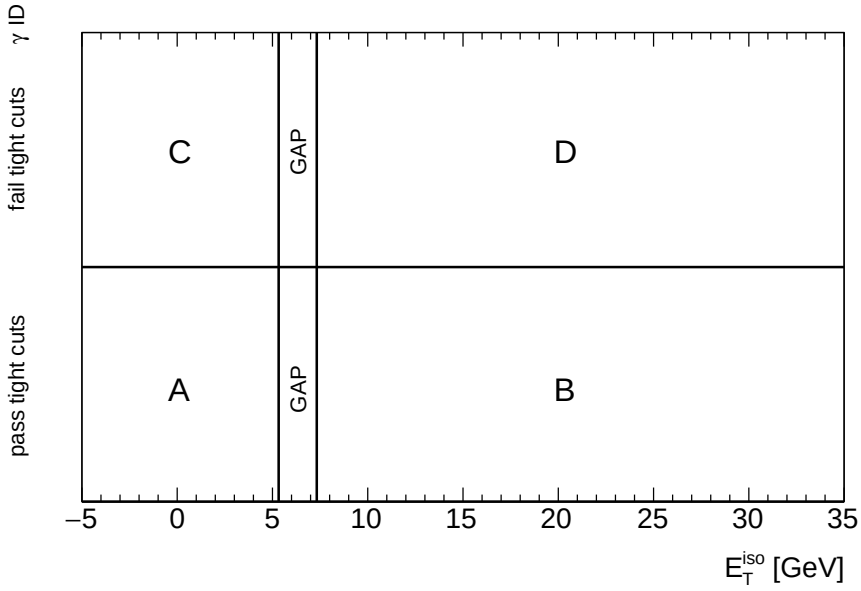


Figure 7.9: Illustration of the two-dimensional plane of the γ_{ID} vs $E_{\text{T}}^{\text{iso}}$ used to estimate the background yield in the signal region, A, from the observed yields in three control regions, B, C and D. The vertical lines correspond to the requirements on $E_{\text{T}}^{\text{iso}}$ for $E_{\text{T}}^{\gamma} = 125$ GeV (line on the left to define regions A and C and line on the right to define regions B and D).

photon candidates.

– Control region D: contains non-tight and non-isolated ($E_{\text{T}}^{\text{iso}} > 4.2 \times 10^{-3} \times E_{\text{T}}^{\gamma} + (4.8 + 2)$ GeV) photon candidates. An upper cut on $E_{\text{T}}^{\text{iso}}$ of 50 GeV is also imposed.

Regions B and D are defined with a *gap* of 2 GeV in $E_{\text{T}}^{\text{iso}}$ from regions A and C to have well separated background-control and signal regions and minimise migrations across the borders. The gap is chosen large enough in comparison to any difference between data and simulations, but still providing sufficient statistics in the control regions to perform the data-driven subtraction. Other choices for the size of this gap are taken as systematic uncertainties (see section 7.7.2).

Therefore, the number of signal events in the signal region A is given by

$$N_{\text{A}}^{\text{sig}} = N_{\text{A}} - R^{\text{bg}} \left(N_{\text{B}} - \epsilon_{\text{B}} N_{\text{A}}^{\text{sig}} \right) \frac{\left(N_{\text{C}} - \epsilon_{\text{C}} N_{\text{A}}^{\text{sig}} \right)}{\left(N_{\text{D}} - \epsilon_{\text{D}} N_{\text{A}}^{\text{sig}} \right)}, \tag{7.1}$$

where $N_{\text{A}}^{\text{sig}}$ is the expected number of signal events, N_{K} with $K = A, B, C, D$ is the number of observed events in each region and

$$R^{\mathrm{bg}} = \frac{N_{\mathrm{A}}^{\mathrm{bg}} N_{\mathrm{D}}^{\mathrm{bg}}}{N_{\mathrm{B}}^{\mathrm{bg}} N_{\mathrm{C}}^{\mathrm{bg}}}\,, \tag{7.2}$$

was taken as $R^{\mathrm{bg}} = 1$ for the nominal results; $N_{\mathrm{K}}^{\mathrm{bg}}$ with $K = A$, B, C, D is the number of background events in each region. The quantity R^{bg} is a measure of the correlation between the γ_{ID} and $E_{\mathrm{T}}^{\mathrm{iso}}$ variables in background events. Deviations of R^{bg} with respect to unity were taken as systematic uncertainties (see section 7.7.2). Equation 7.1 takes into account the expected number of signal events in the three background control regions via the signal leakage fractions, ϵ_{K} with $K = B$, C, D.

The signal leakage fractions were extracted from the MC simulations of the signal and are shown in figure 7.10 for PYTHIA and SHERPA. The fractions ϵ_{C}, which represent the signal leakage into the non-tight and isolated control region, are approximately constant in all η^{γ} regions and around $0.02-0.05$, being very similar for PYTHIA and SHERPA. The fractions ϵ_{B}, which represent the signal leakage into the tight and non-isolated background control region, are approximately constant for all $|\eta^{\gamma}|$ regions and take values of ≈ 0.1 for PYTHIA and ≈ 0.05 for SHERPA. The difference between PYTHIA and SHERPA is due to the different treatment of the bremsstrahlung (hereafter referred to as brem) component in both MC generators. In the case of PYTHIA, photons can be radiated in the parton shower without a restriction on the opening angle with respect to the parent parton and, as a result, photons can be emitted very close to the parton direction; however, in SHERPA, photons are not emitted in the parton shower, and the brem component is simulated through matrix elements of $2 \to N$ processes, with $N \geq 3$. In this case, the collinear singularity is circumvented by restricting the emission with the implementation of the Frixione requirement; as a result, photons are not emitted close to the parent parton. The fractions ϵ_{D}, which represent the signal leakage into the non-tight and non-isolated control region, increase as a function of E_{T}^{γ}, and take very different values for PYTHIA and SHERPA. This control region is the most affected by the different treatment of the brem component in both models.

It was observed that the application of an upper cut on $E_{\mathrm{T}}^{\mathrm{iso}}$ of 50 GeV is safe to avoid regions in which the tail of the MC distributions overestimates the data. In this way, the effect on the signal leakage fractions of the contributions at large values of $E_{\mathrm{T}}^{\mathrm{iso}}$ is suppressed and, thereby, it makes the results of the subtraction using the 2D-sideband method more robust than by not using an upper cut on $E_{\mathrm{T}}^{\mathrm{iso}}$ for the background control regions B and D. A systematic uncertainty was estimated by not including this upper cut on $E_{\mathrm{T}}^{\mathrm{iso}}$ for the definition of the control regions (see section 7.7.2).

The signal yield was determined from the observed number of events in the data in the four regions of the γ_{ID} vs $E_{\mathrm{T}}^{\mathrm{iso}}$ plane and the signal leakage fractions obtained from the simulated events using equation 7.1. The signal purities, computed as $P = N_{\mathrm{A}}^{\mathrm{sig}}/N_{\mathrm{A}}$ using the signal leakage fractions from PYTHIA and SHERPA are shown in figure 7.11. For some of the points at high E_{T}^{γ} in which the purity is above unity or consistent with unity but with large statistical uncertainty, the purity was set to unity. The signal purity is $\geq 90\%$ and the estimations using either PYTHIA or SHERPA to compute the signal leakage

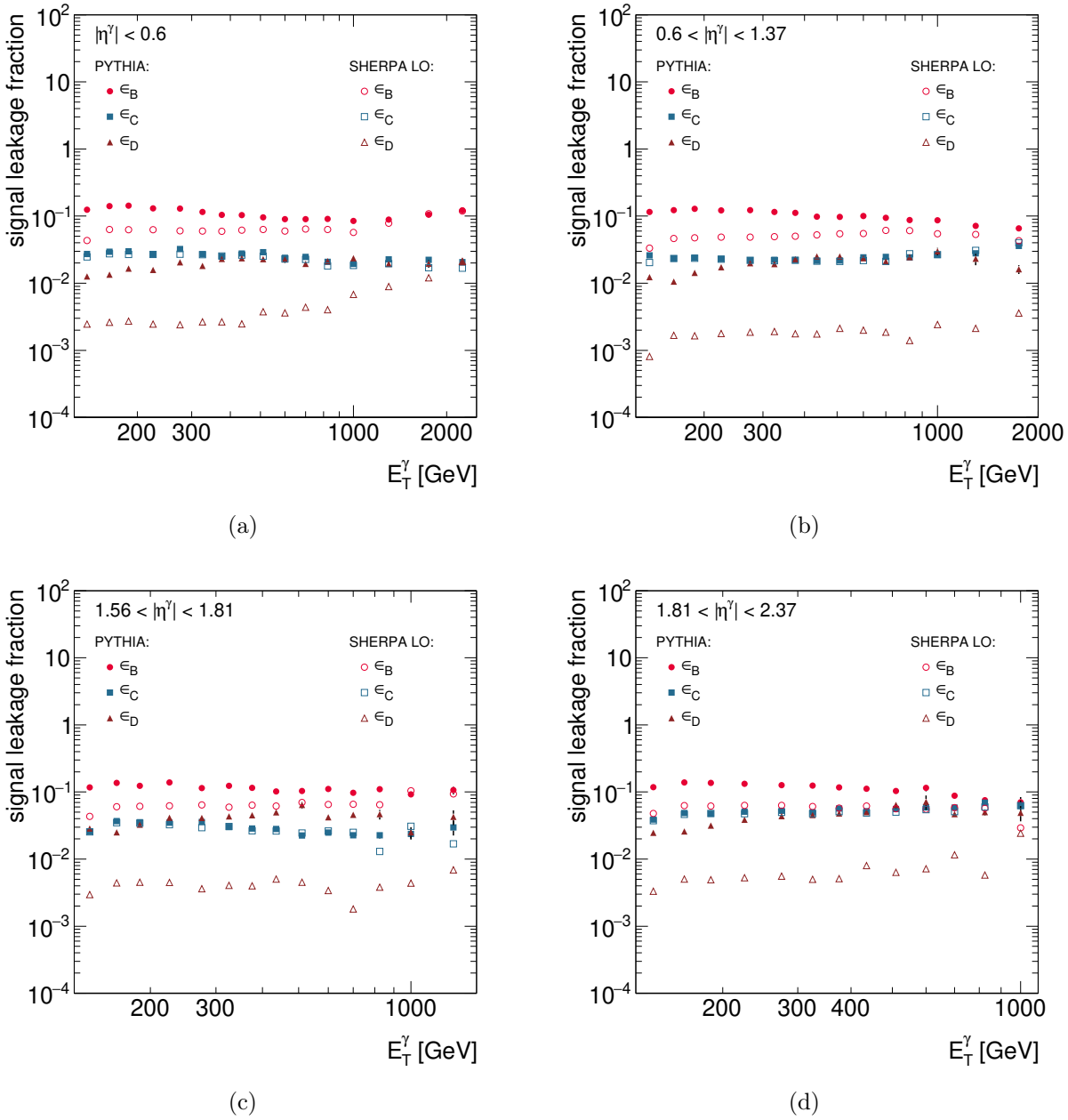


Figure 7.10: Signal leakage fractions from PYTHIA (solid markers) and SHERPA (open markers) for the control regions B (dots), C (squares) and D (triangles) as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$.

fractions show very similar values. The nominal signal yields were extracted using the signal leakage fractions from PYTHIA; while the estimation using signal leakage fractions from SHERPA was taken as a systematic uncertainty on the purity (see section 7.7.1).

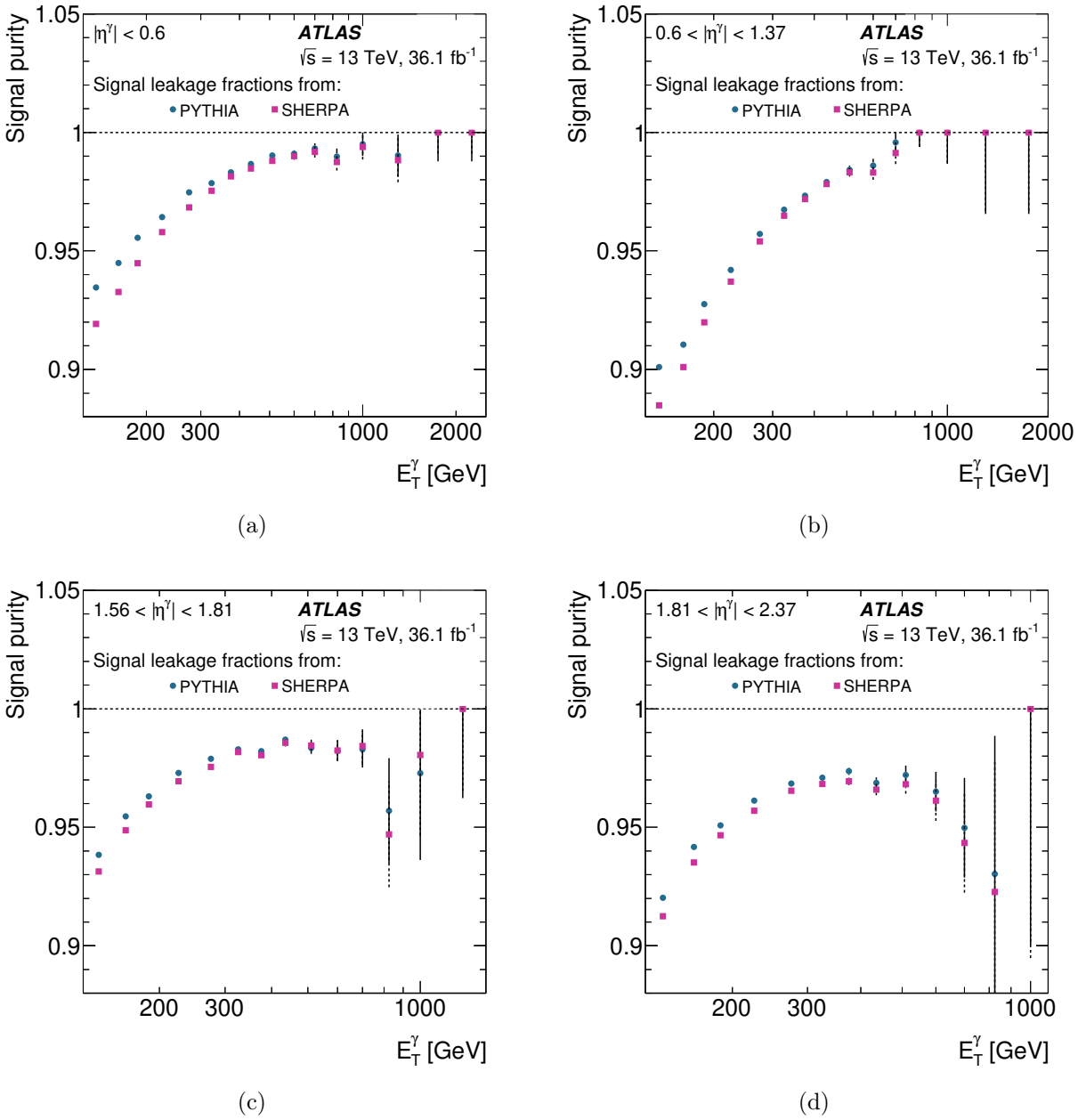


Figure 7.11: Estimated signal purities in data using signal leakage fractions from PYTHIA (dots) and SHERPA (squares) as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$.

7.3.1 Background from electrons faking photons

Electrons or positrons can be misidentified as photons and represent an additional source of background. This background is largely suppressed by the photon selection. The residual background contribution is evaluated using SHERPA 2.2.1 MC simulations of

W+jets and Z+jets as described in section 4.3.

The fraction of W+jets and Z+jets background to inclusive photon production as a function of E_{T}^{γ} for different regions of $|\eta^{\gamma}|$ was estimated with a particle-level matching to $e^{\pm}$ using an angular distance requirement $\Delta R < 0.2$. The fraction for W+jets (Z+jets) is shown in figure 7.12 (7.13) for the subsample of photons with $E_{\mathrm{T}}^{\gamma} > 150$ GeV.

The background from electrons faking photons from W+jets (Z+jets) was found to be approximately smaller than: 0.22% (0.07%) for $|\eta^{\gamma}| < 0.6$; 0.31% (0.10%) for $0.6 < |\eta^{\gamma}| < 1.37$; 0.50% (0.18%) for $1.56 < |\eta^{\gamma}| < 1.81$; and 0.69% (0.25%) for $1.81 < |\eta^{\gamma}| < 2.37$. Since the background was found to be subpercent in the phase-space regions of the analysis, no subtraction was performed and a systematic uncertainty of the size of the evaluated background was assigned (see section 7.7.2).

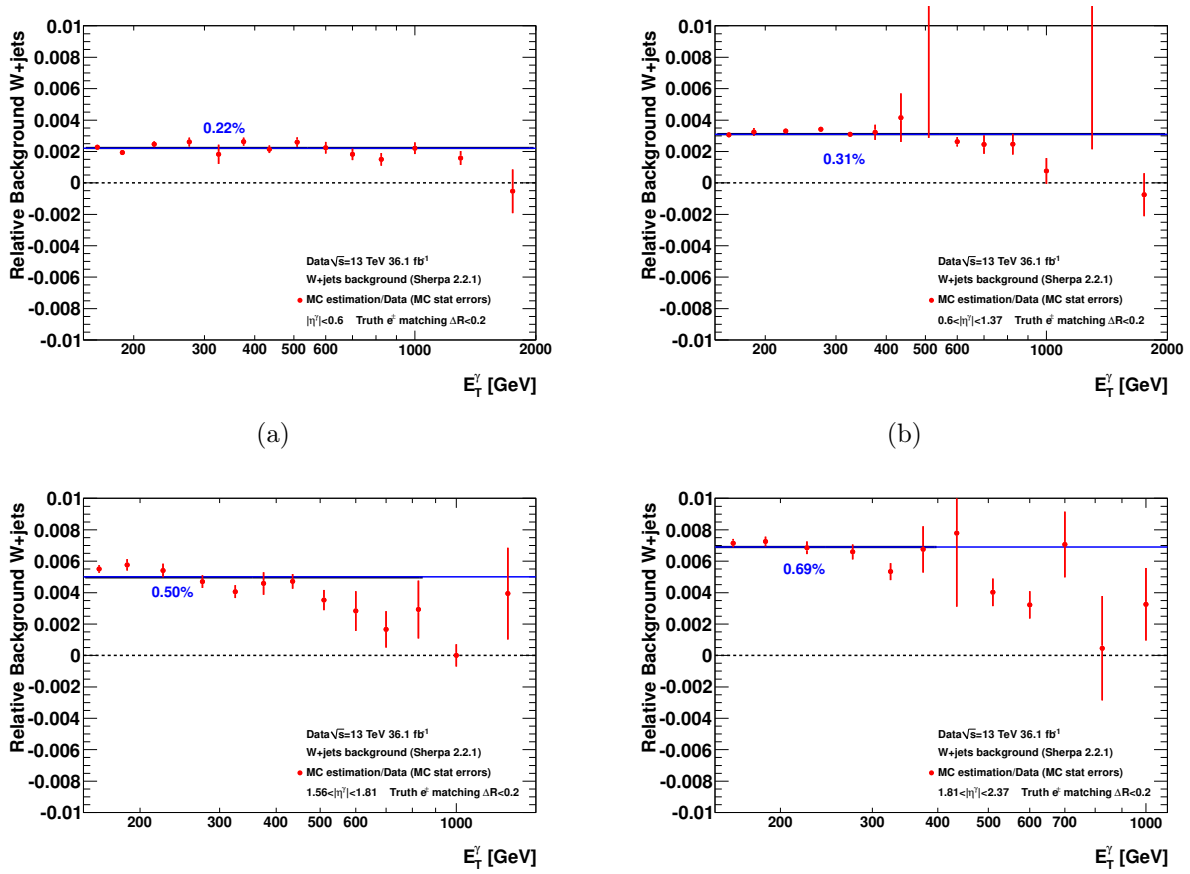


Figure 7.12: Fraction of W+jets background to inclusive photon production as a function of E_{T}^{γ} for different regions in $|\eta^{\gamma}|$ using the samples of SHERPA 2.2.1. The error bars represent the statistical uncertainties of the MC samples. The black line represents the result of a fit to a constant.

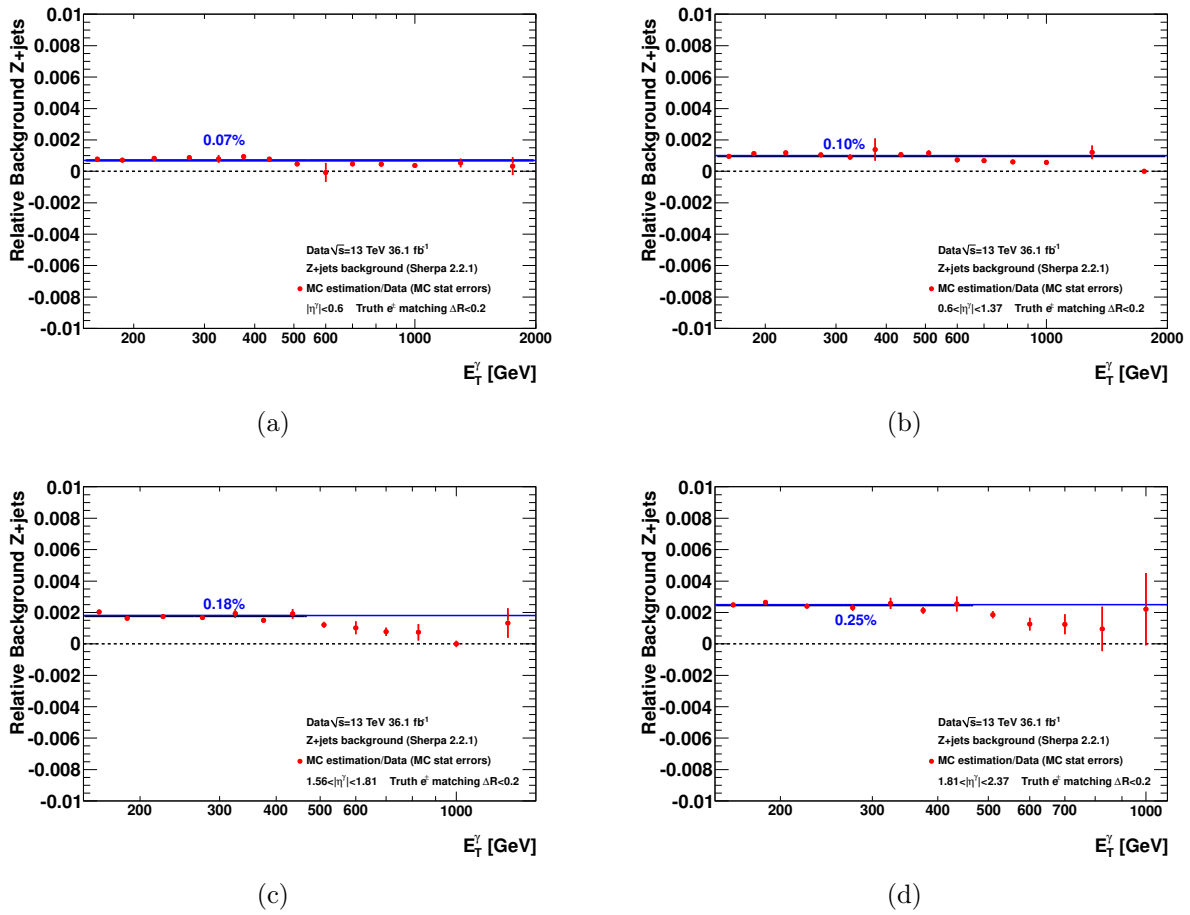


Figure 7.13: Fraction of Z+jets background to inclusive photon production as a function of E_{T}^{γ} for different regions in $|\eta^{\gamma}|$ using the samples of SHERPA 2.2.1. The error bars represent the statistical uncertainties of the MC samples. The black line represents the result of a fit to a constant.

7.4 Control plots

The estimated signal yield using the signal leakage fractions from PYTHIA is shown in figure 7.14 as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. The measured distributions decrease with increasing E_{T}^{γ} by approximately six orders of magnitude within the measured range. For comparison, the detector level predictions of PYTHIA and SHERPA MC generators are also included in figure 7.14. Both PYTHIA and SHERPA simulations provide a reasonable description of the data, except at high E_{T}^{γ}.

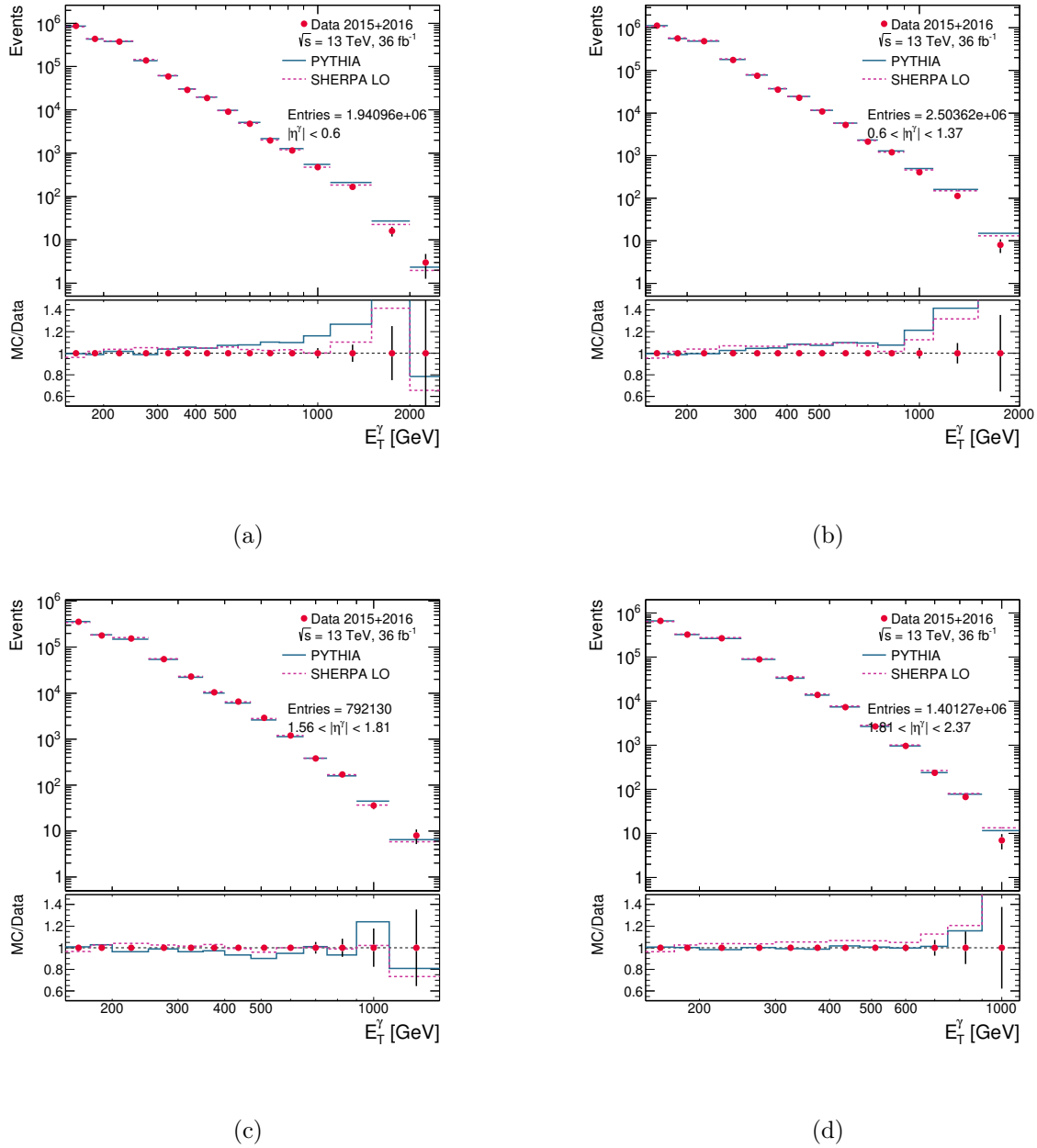


Figure 7.14: The extracted signal yield in data (dots) using signal leakage fractions from PYTHIA as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. For comparison, detector level simulations from PYTHIA (solid histograms) and SHERPA (dashed histograms) are also included. The MC distributions are normalised to the number of data events. The ratio of the MC distributions to the data is shown in the lower panel of the figures. The error bars display the statistical uncertainty of the data and the uncertainty in the signal leakage fractions added in quadrature. Results are shown for the subsample of $E_{\mathrm{T}}^{\gamma} > 150$ GeV.

7.5 Selection efficiency and purity

The quality of the reconstruction and the selection efficiency and purity were studied using the samples of MC events from PYTHIA and SHERPA. The MC events were required to fulfil the same kinematic and isolation requirements both at the detector and particle levels. In addition, the particle- and detector-level photons were required to be matched using the angular requirement $\Delta R < 0.2$ to remove fake photons.

Reconstruction quality

To assess the quality of the reconstruction of E_{T}^{γ} in the different $|\eta^{\gamma}|$ regions, the observables at detector and particle levels were compared in an event-by-event basis. Figures 7.15 and 7.16 show the correlation between the detector and particle levels for E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$ for the samples of PYTHIA and SHERPA, respectively. A good reconstruction quality is observed.

Selection efficiency and purity

The integrated selection efficiency was computed as

$$\varepsilon = \frac{N^{\mathrm{reco,part}}}{N^{\mathrm{part}}}\,, \tag{7.3}$$

where $N^{\mathrm{reco,part}}$ is the number of MC events which pass all the selection requirements both at the detector and particle levels, and N^{part} is the number of MC events that pass the selection requirements at the particle level. The integrated selection efficiency was found to be $83.26 \pm 0.05\%$ ($83.65 \pm 0.08\%$) from the PYTHIA (SHERPA) samples.

The bin-to-bin selection efficiency was computed as

$$\varepsilon(i) = \frac{N^{\mathrm{reco,part}}(i)}{N^{\mathrm{part}}(i)}\,, \tag{7.4}$$

where $N^{\mathrm{reco,part}}(i)$ is the number of MC events which pass all the selection requirements both at the detector and particle levels and are generated and reconstructed in bin i, and $N^{\mathrm{part}}(i)$ is the number of MC events that pass the selection requirements at the particle level and are located in bin i. Figure 7.17 shows the bin-to-bin selection efficiencies as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$ for PYTHIA and SHERPA. The bin-to-bin selection efficiency is typically $70\% - 80\%$ and is very similar for PYTHIA and SHERPA.

The integrated selection purity was computed as

$$\mathcal{P} = \frac{N^{\mathrm{reco,part}}}{N^{\mathrm{reco}}}\,, \tag{7.5}$$

where N^{reco} is the number of MC events that pass the selection requirements at the detector level. The integrated selection purity was found to be $92.86 \pm 0.05\%$ ($93.85 \pm 0.07\%$) from the PYTHIA (SHERPA) samples.

The bin-to-bin selection purity was computed as

$$\mathcal{P}(i) = \frac{N^{\text{reco,part}}(i)}{N^{\text{reco}}(i)}\,, \tag{7.6}$$

where $N^{\text{reco}}(i)$ is the number of MC events which pass the selection requirements at the detector level and are located in bin i. Figure 7.18 shows the bin-to-bin selection purities as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$ for PYTHIA and SHERPA. The bin-to-bin selection purity is above 80% and is very similar for PYTHIA and SHERPA.

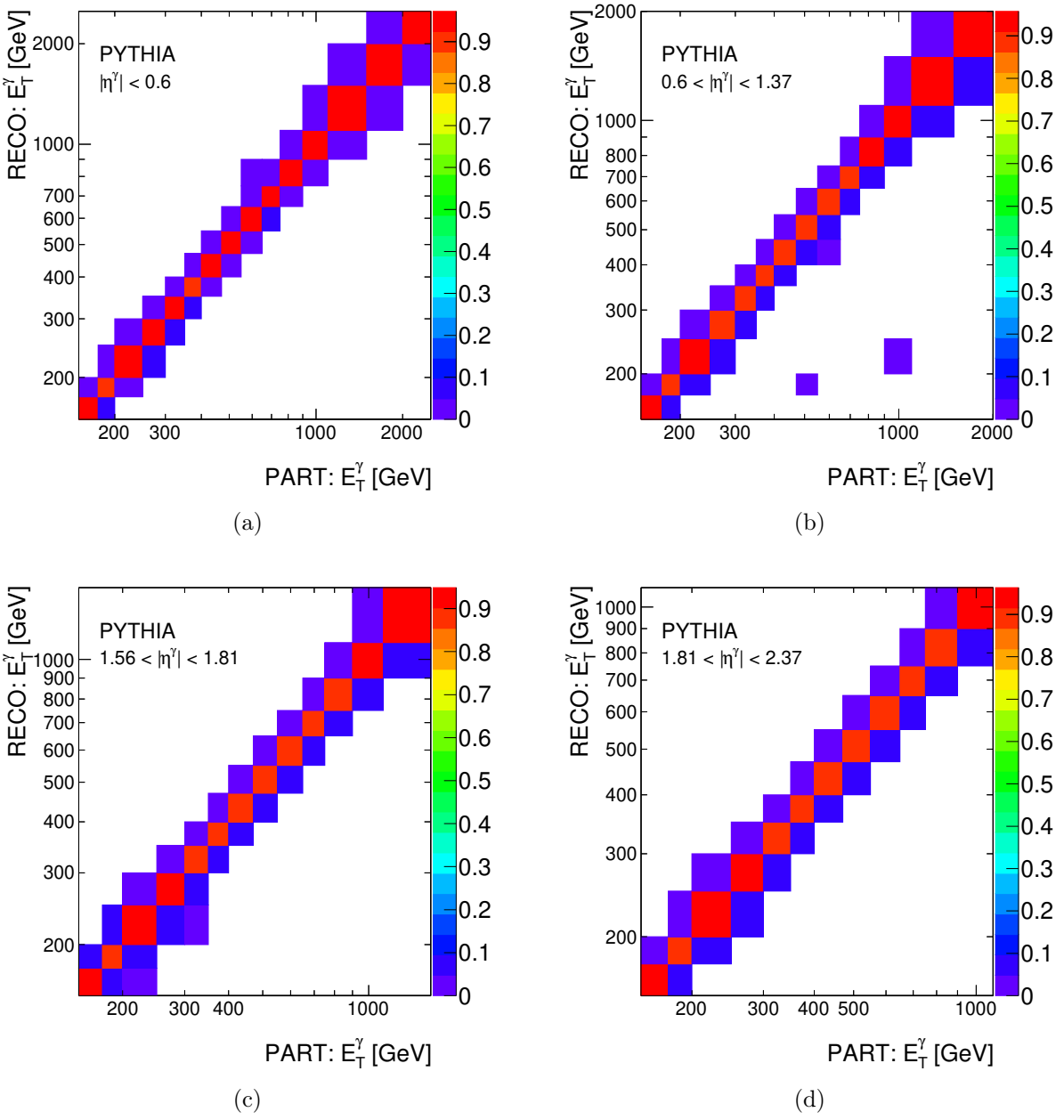


Figure 7.15: The $(E_{\mathrm{T}}^{\gamma})_{\mathrm{reco}}$ vs $(E_{\mathrm{T}}^{\gamma})_{\mathrm{part}}$ correlation for the PYTHIA MC samples in the different regions of $|\eta^{\gamma}|$. Normalisation is column wise. Results are shown for the subsample of $E_{\mathrm{T}}^{\gamma} > 150$ GeV.

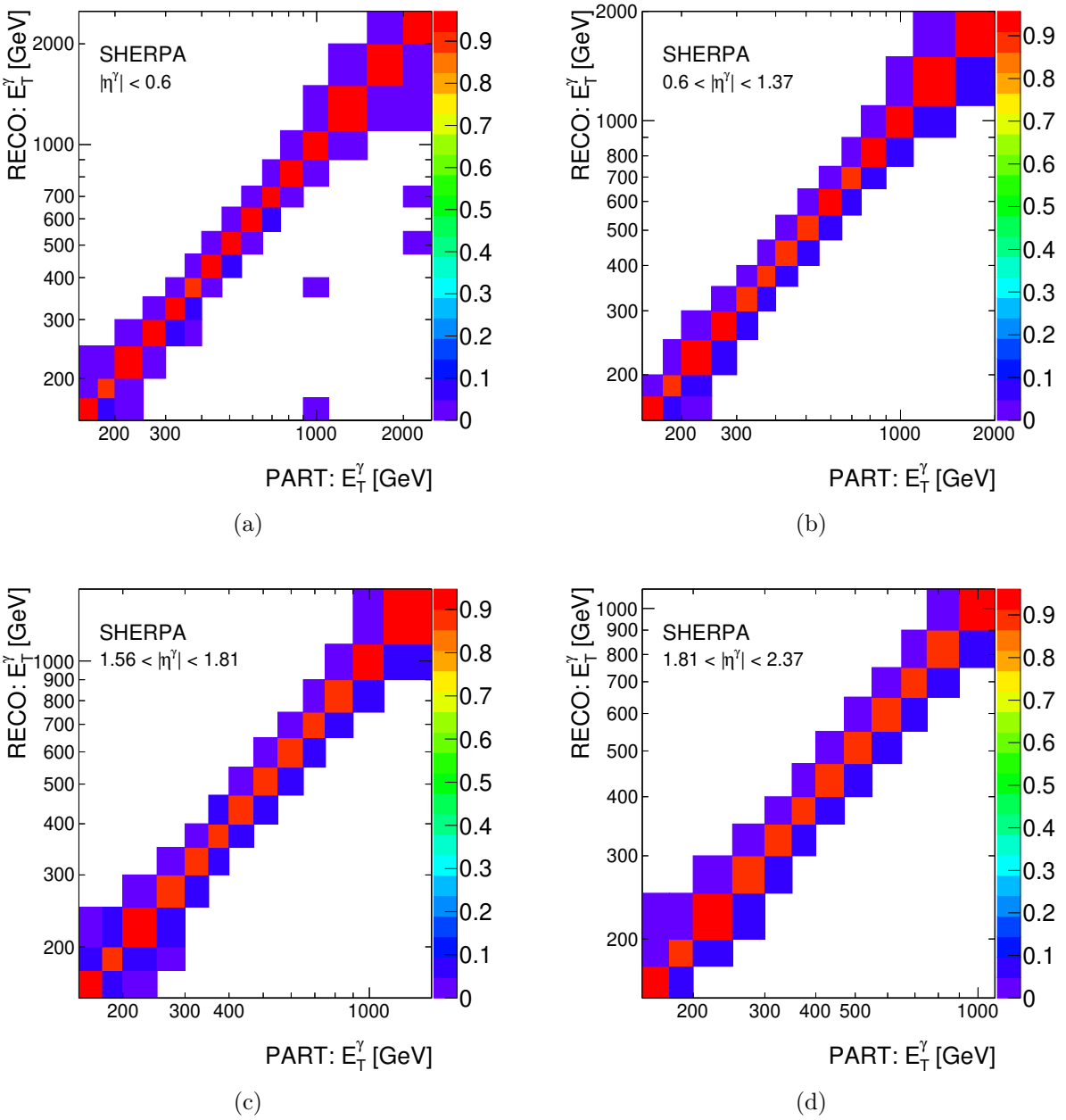


Figure 7.16: The $(E_{\mathrm{T}}^{\gamma})_{\mathrm{reco}}$ vs $(E_{\mathrm{T}}^{\gamma})_{\mathrm{part}}$ correlation for the SHERPA MC samples in the different regions of $|\eta^{\gamma}|$. Normalisation is column wise. Results are shown for the subsample of $E_{\mathrm{T}}^{\gamma} > 150$ GeV.

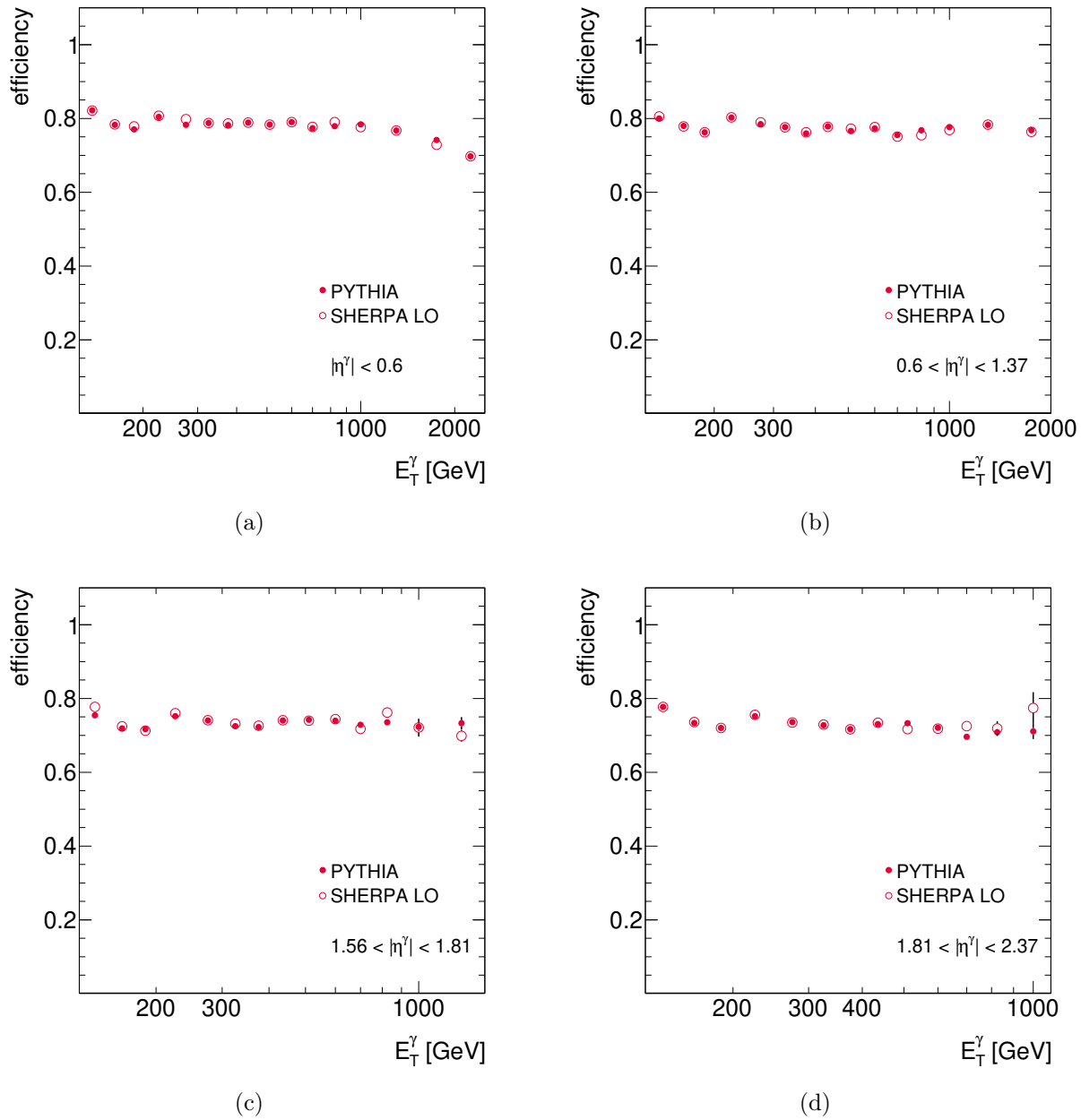


Figure 7.17: Selection efficiencies from PYTHIA (solid circles) and SHERPA (open circles) as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$.

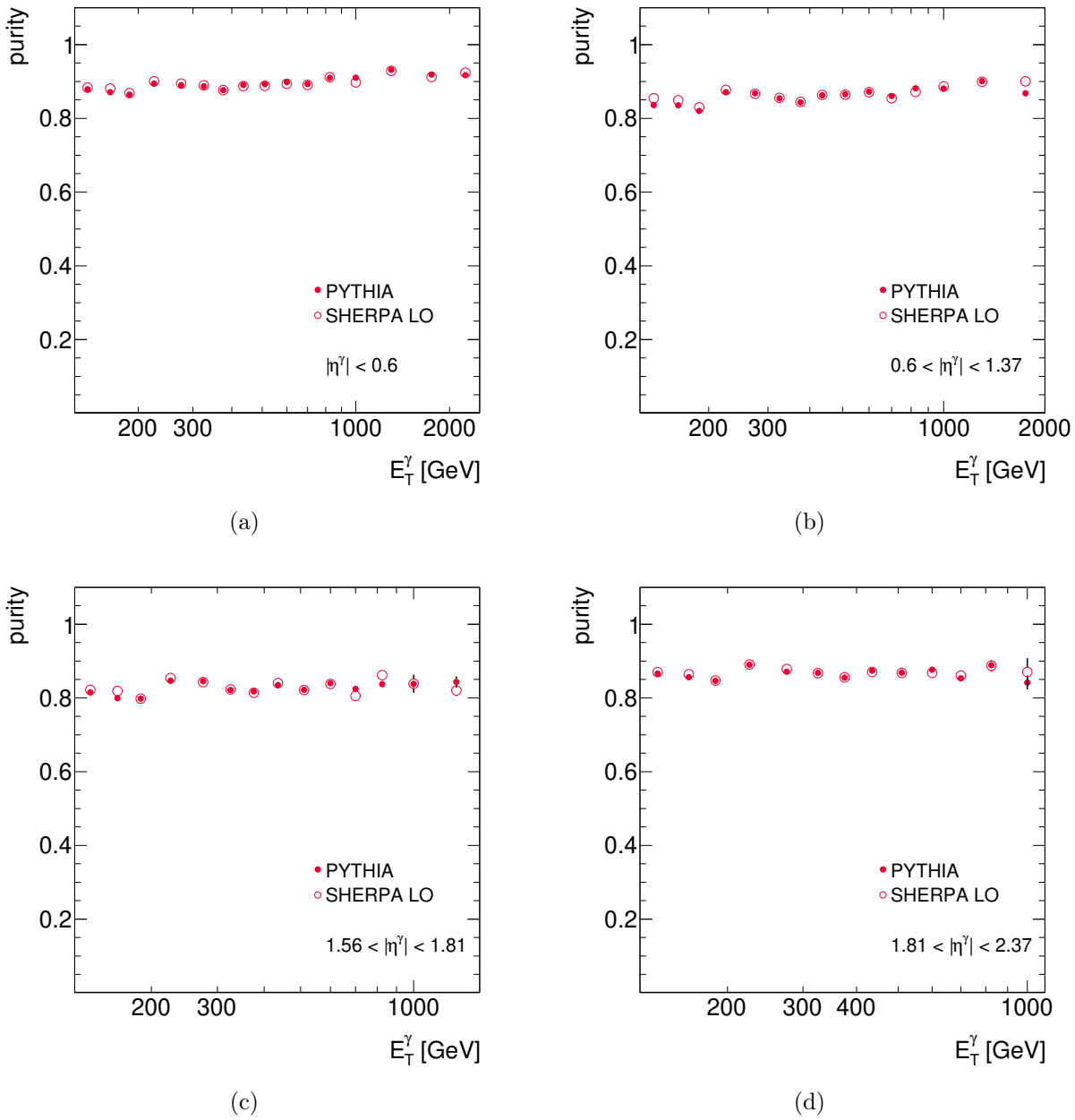


Figure 7.18: Selection purities from PYTHIA (solid circles) and SHERPA (open circles) as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$.

7.6 Cross-section measurement procedure

The isolated-photon cross section was measured for photons with $E_{\mathrm{T}}^{\gamma} > 125$ GeV and $|\eta^{\gamma}| < 2.37$ (excluding the region $1.37 < |\eta^{\gamma}| < 1.56$). Isolation was ensured by requiring

$E_{\mathrm{T}}^{\mathrm{iso}} < 4.2 \times 10^{-3} \times E_{\mathrm{T}}^{\gamma} + 4.8$ GeV. The previous cuts define the phase space of the measurement. The particle-level isolation requirement on the photon is built summing the transverse energy of all stable particles[2], except for muons and neutrinos, in a cone of size $\Delta R = 0.4$ around the photon direction after the contribution from the UE is subtracted; the same subtraction procedure used on data is applied at the MC particle level (see section 5.5); the particles associated with the overlaid pp collisions were not considered. The data were unfolded to the particle level to this region of phase space, as explained in the following.

The isolated-photon differential cross section was measured as a function of E_{T}^{γ} in four different regions of $|\eta^{\gamma}|$, namely, $|\eta^{\gamma}| < 0.6$, $0.6 < |\eta^{\gamma}| < 1.37$, $1.56 < |\eta^{\gamma}| < 1.81$ and $1.81 < |\eta^{\gamma}| < 2.37$.

The data distributions after background subtraction (see section 7.3) were corrected to the particle level using bin-by-bin correction factors determined using the MC samples. These correction factors take into account the efficiency of the selection criteria and the purity and efficiency of the photon reconstruction. For this approach to be valid, the uncorrected distributions of the data must be adequately described by the MC simulations at the detector level. This condition was satisfied by both PYTHIA and SHERPA MC samples (see figure 7.14). Therefore, the data distributions were corrected to the particle level via the formula

$$\frac{\mathrm{d}\sigma}{\mathrm{d}E_{\mathrm{T}}^{\gamma}}(i) = \frac{N_{\mathrm{A}}^{\mathrm{sig}}(i)\, C^{\mathrm{MC}}(i)}{\mathcal{L}\, \Delta E_{\mathrm{T}}^{\gamma}(i)} \,,$$

where $(\mathrm{d}\sigma/\mathrm{d}E_{\mathrm{T}}^{\gamma})\,(i)$ is the differential cross section as a function of E_{T}^{γ}, $N_{\mathrm{A}}^{\mathrm{sig}}(i)$ is the number of background-subtracted data events in bin i, $C^{\mathrm{MC}}(i)$ is the correction factor in bin i, $\mathcal{L}$ is the integrated luminosity and $\Delta E_{\mathrm{T}}^{\gamma}(i)$ is the width of bin i. The correction factors were computed as

$$C^{\mathrm{MC}}(i) = \frac{N_{\mathrm{part}}^{\mathrm{MC}}(i)}{N_{\mathrm{reco}}^{\mathrm{MC}}(i)} \,.$$

The correction factor $C^{\mathrm{MC}}(i)$ is determined for each bin i and is computed using the MC samples of events, where $N_{\mathrm{part}}^{\mathrm{MC}}(i)$ is the number of events which fulfil the kinematic constraints of the phase-space region at the particle level, i.e. the leading photon at particle level has to satisfy the requirement on E_{T}^{γ} (in particular, that of the given bin), the requirement on $|\eta^{\gamma}|$ (for the given region), and the constraint on the photon isolation at particle level, and $N_{\mathrm{reco}}^{\mathrm{MC}}(i)$ is the number of events which fulfil all the selection criteria at detector level, i.e. the event and the photon selection (e.g. the requirements on E_{T}^{γ}, $|\eta^{\gamma}|$, and isolation are applied to the reconstructed quantities). The quantity $N_{\mathrm{reco}}^{\mathrm{MC}}(i)$ can be rewritten in the following way

[2] All particles with a decay length of $c\tau > 10$ mm are considered as stable.

$$N_{\text{reco}}^{\text{MC}}(i) = N_{\text{reco/part-inside}}^{\text{MC}}(i) + N_{\text{reco/part-outside}}^{\text{MC}}(i)\,,$$

where $N_{\text{reco/part-inside}}^{\text{MC}}(i)$ is the number of events which, in addition to satisfying the requirements at the detector level, also fulfil the kinematic constraints at the particle level (the leading photon at particle level fulfils the conditions of the phase-space region); these are the events for which the photon is generated and reconstructed in the same bin i. The ratio of this quantity to $N_{\text{part}}^{\text{MC}}(i)$ is an estimation of the efficiency. $N_{\text{reco/part-outside}}^{\text{MC}}(i)$ is the number of events which satisfy the requirements at the detector level, but from the particle-level point of view, at least one of the kinematic constraints on the leading photon is not fulfilled. These are events which are generated in other bins of E_{T}^{γ}, or in other $|\eta^{\gamma}|$ regions and that migrate into the given bin i. There is also a contribution of events which migrate across the cut on the isolation value. By using this correction factor, $C^{\text{MC}}(i)$, both the efficiency of reconstructing events in the given bin as well as taking into account the migrations, are included in the unfolding.

The nominal cross section was measured using the correction factors from PYTHIA, and the deviations in the results obtained by using SHERPA to correct the data were taken as a systematic uncertainty on the effect of the QCD-cascade and hadronisation models in the corrections (see section 7.7.4). The correction factors are shown in figure 7.19, and were found to vary around $1.1 - 1.3$, being very similar for PYTHIA and SHERPA.

A cross-check of the bin-by-bin unfolding method was performed by reweighting the MC distributions of PYTHIA at detector and particle levels to the data. A further cross-check of the unfolding was performed by adjusting the fraction of the brem and hard components in PYTHIA. The details of both cross-checks are described in appendix D.

The cross section was also obtained using an iterative Bayesian unfolding method [153] and compared with the nominal result. The two results were found to be consistent with each other independently of whether the PYTHIA or SHERPA MC simulations were used for the unfolding. The cross-check is described in detail in appendix E.

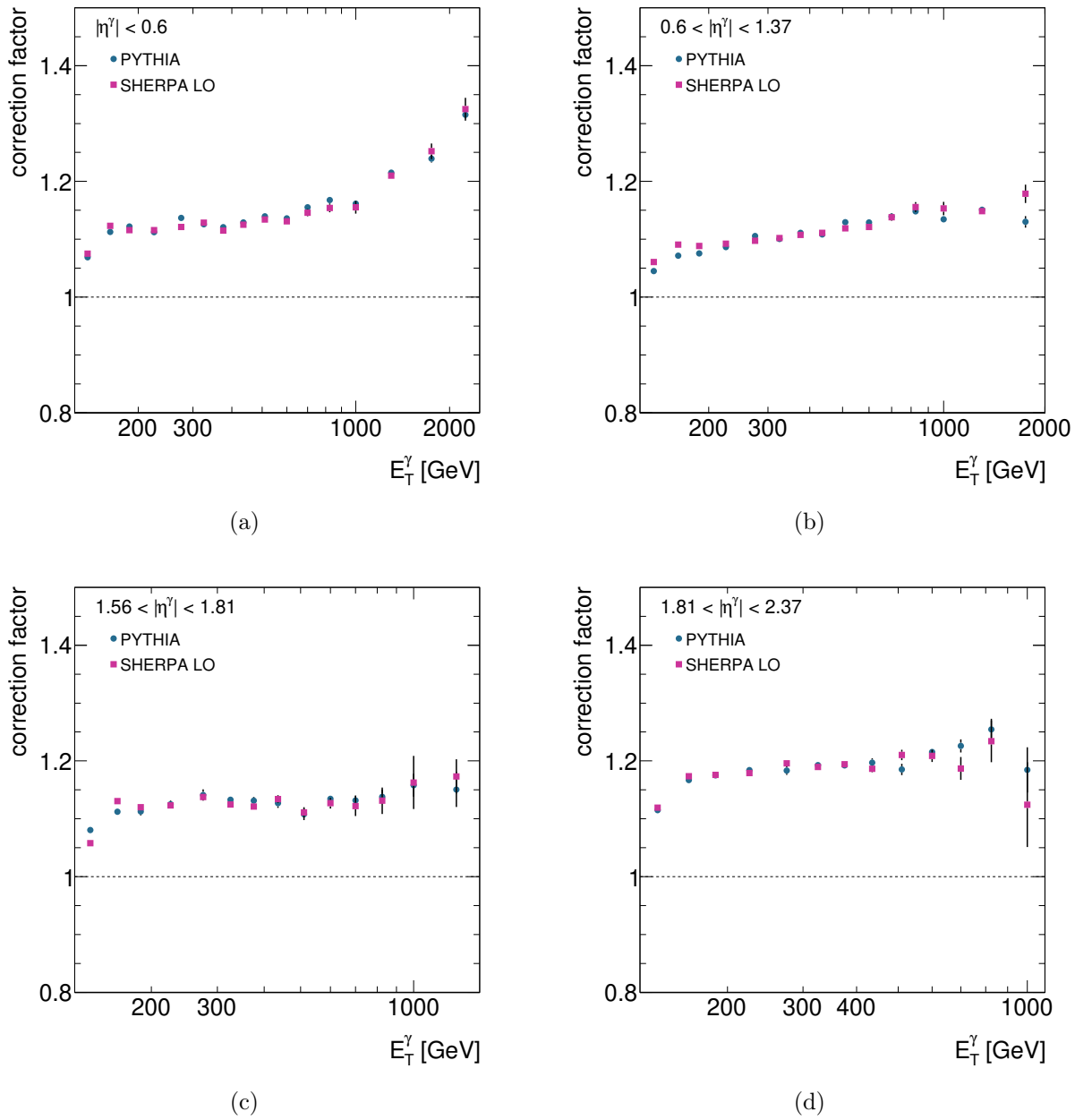


Figure 7.19: Correction factors from PYTHIA (dots) and SHERPA (squares) as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$.

7.7 Systematic uncertainties

The sources of systematic uncertainty that affect the measurement are discussed below. These sources include the signal modelling, the background subtraction, the photon identification and isolation, the unfolding procedure, the modelling of pile-up, the trigger efficiency, the luminosity measurement and the photon energy scale and resolution. For some of the systematic uncertainties, the bootstrap technique [154] was used to evaluate the data statistical influence on the uncertainties; the results were then used in a fit to smooth the systematic uncertainties.

7.7.1 Signal modelling

The nominal signal leakage fractions used in the extraction of the signal yield were evaluated with the PYTHIA simulation. The effect of the implementation of the matrix-element calculation in the generator as well as the models used for the parton shower and hadronisation were evaluated by comparing the nominal results with those obtained using the SHERPA simulation for the determination of the signal leakage fractions (see section 7.3). The differences were taken as a systematic uncertainty. The resulting uncertainty in the measured cross section is shown in figure 7.20 and is always less than 2% and typically less than 1%.

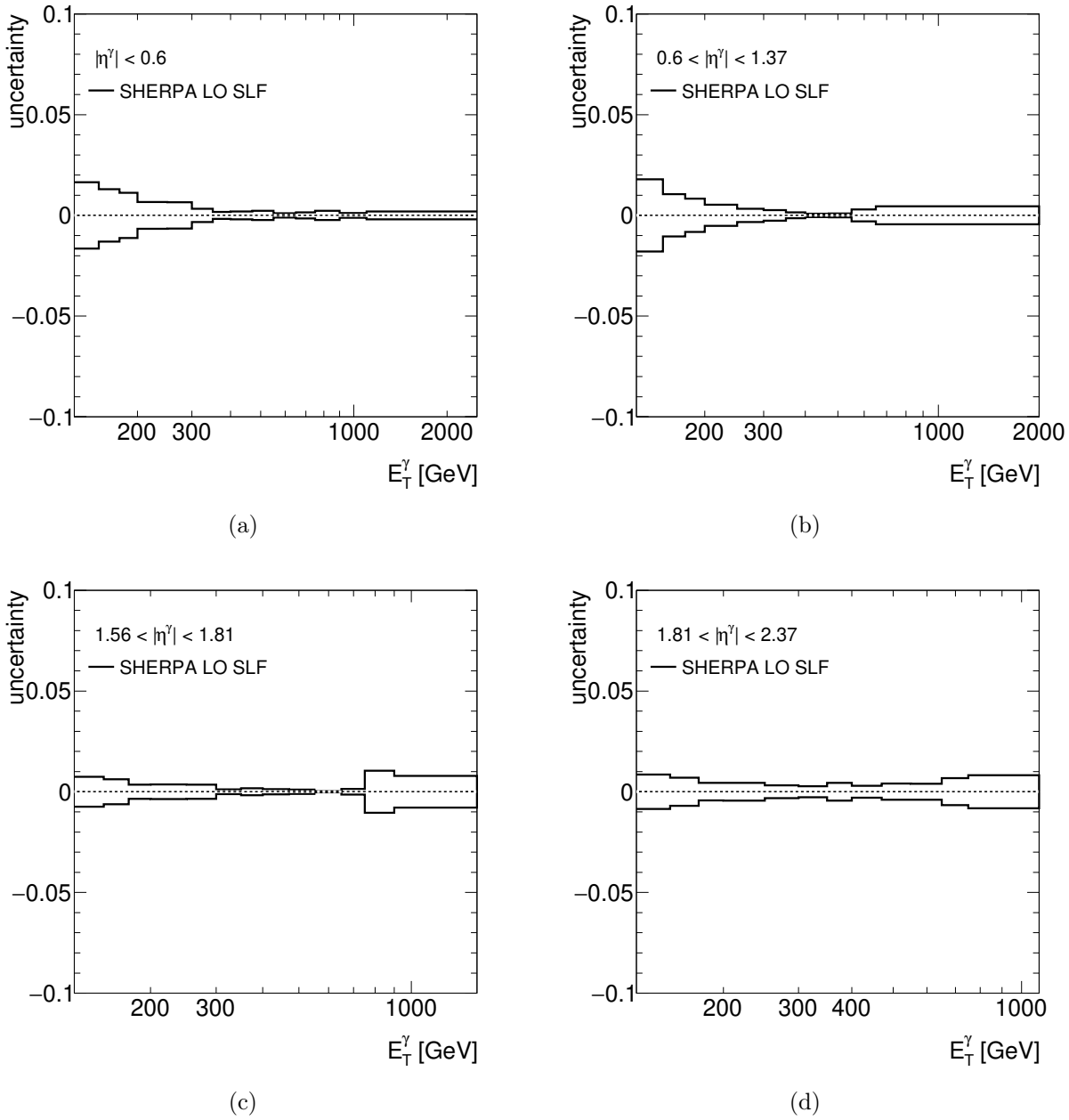


Figure 7.20: Systematic uncertainty in the measured cross section due to the QCD-cascade and hadronisation models as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. The uncertainty was symmetrised. Statistical fluctuations in the estimation of this systematic uncertainty were taken care of by rebinning, merging those bins affected by relatively large statistical fluctuations with neighbouring bins.

7.7.2 Background subtraction

Choice of background control regions

The data-driven background subtraction depends on the choice of background control regions. This dependence is studied by varying the lower limit on $E_{\mathrm{T}}^{\mathrm{iso}}$ for regions B and D by ± 1 GeV, removing the upper limit on $E_{\mathrm{T}}^{\mathrm{iso}}$ for regions B and D, as well as changing the choice of inverted photon identification variables. For each variation, the differences between the nominal signal yield and those extracted using the modified background control regions are taken as systematic uncertainties. The resulting uncertainties in the measured cross section due to the first two variations are less than 0.2% and are shown in figures 7.21 and 7.22, respectively. For the third variation, the analysis is repeated increasing (LoosePrime5) or decreasing (LoosePrime3 or LoosePrime2) the number of the shower shape variables computed from the energy deposits in the first layer of the EM calorimeter that are removed from the list of tight requirements used in the classification of non-tight photon candidates (see section 5.4). The final uncertainty was estimated as the envelope of these variations, leading to an uncertainty in the cross section that is less than 2% and typically less than 1%. Figure 7.23 shows the resulting uncertainty in the measured cross section.

Identification and isolation correlation in the background

The isolation and identification photon variables used to define the 2D-sideband method to subtract the background (see section 7.3) were assumed to be uncorrelated for background events ($R^{\mathrm{bg}} = 1$ in equation 7.2). Any correlation between these variables would affect the estimation of the purity of the signal and would lead to systematic effects in the background-subtraction procedure. A data-driven method used in ref. [9, 10] was devised to determine the maximum possible deviations of R^{bg} from the previous assumption. The dependence of the signal yield on this assumption was investigated in validation regions, which were defined within the non-isolated background control regions B and D. Region B was subdivided into two regions: region B' of tight photon candidates with $E_{\mathrm{T,cut}}^{\mathrm{iso}} + 2$ GeV $< E_{\mathrm{T}}^{\mathrm{iso}} < E_{\mathrm{T,cut}}^{\mathrm{iso}} + 10$ GeV and region B'' of tight photon candidates with $E_{\mathrm{T}}^{\mathrm{iso}} > E_{\mathrm{T,cut}}^{\mathrm{iso}} + 10$ GeV. Likewise, region D is subdivided into two regions, D' and D'', using the same separation in $E_{\mathrm{T}}^{\mathrm{iso}}$ as above. The four validation regions B', B'', D' and D'' were used to extract values of R^{bg} from the data after accounting for the signal leakage fractions in those regions using either PYTHIA or SHERPA MC simulations. The dependence on the signal leakage is investigated by increasing the lower limits on $E_{\mathrm{T}}^{\mathrm{iso}}$ for the validation regions, $E_{\mathrm{T,cut}}^{\mathrm{iso}} + 2$ GeV ($E_{\mathrm{T,cut}}^{\mathrm{iso}} + 10$ GeV), each time by 1 GeV up to $E_{\mathrm{T,cut}}^{\mathrm{iso}} + 7$ GeV ($E_{\mathrm{T,cut}}^{\mathrm{iso}} + 15$ GeV) for regions B' and D' (B'' and D''), keeping the width in $E_{\mathrm{T}}^{\mathrm{iso}}$ fixed to 8 GeV for the regions B' and D'. The maximum deviations from unity in the validation regions are shown in figure 7.24 (7.25) for PYTHIA (SHERPA), and vary in the range $15\% - 40\%$ depending on E_{T}^{γ} and $|\eta^{\gamma}|$. The differences between the nominal signal yield and those obtained using the maximum deviations of R^{bg} from unity observed in the validation regions were taken as systematic uncertainties. The resulting uncertainty

in the cross section is less than 2.5% and typically less than 1%. Figure 7.26 shows the uncertainty in the measured cross section.

Background from electrons faking photons

As discussed in section 7.3.1, the background contribution from electrons or positrons misidentified as photons is subpercent and a systematic uncertainty is included by assigning the full size of this background (adding the W+jets and Z+jets contributions linearly, depending on the E_{T}^{γ} and $|\eta^{\gamma}|$ region). Figure 7.27 shows the resulting uncertainty in the measured cross section.

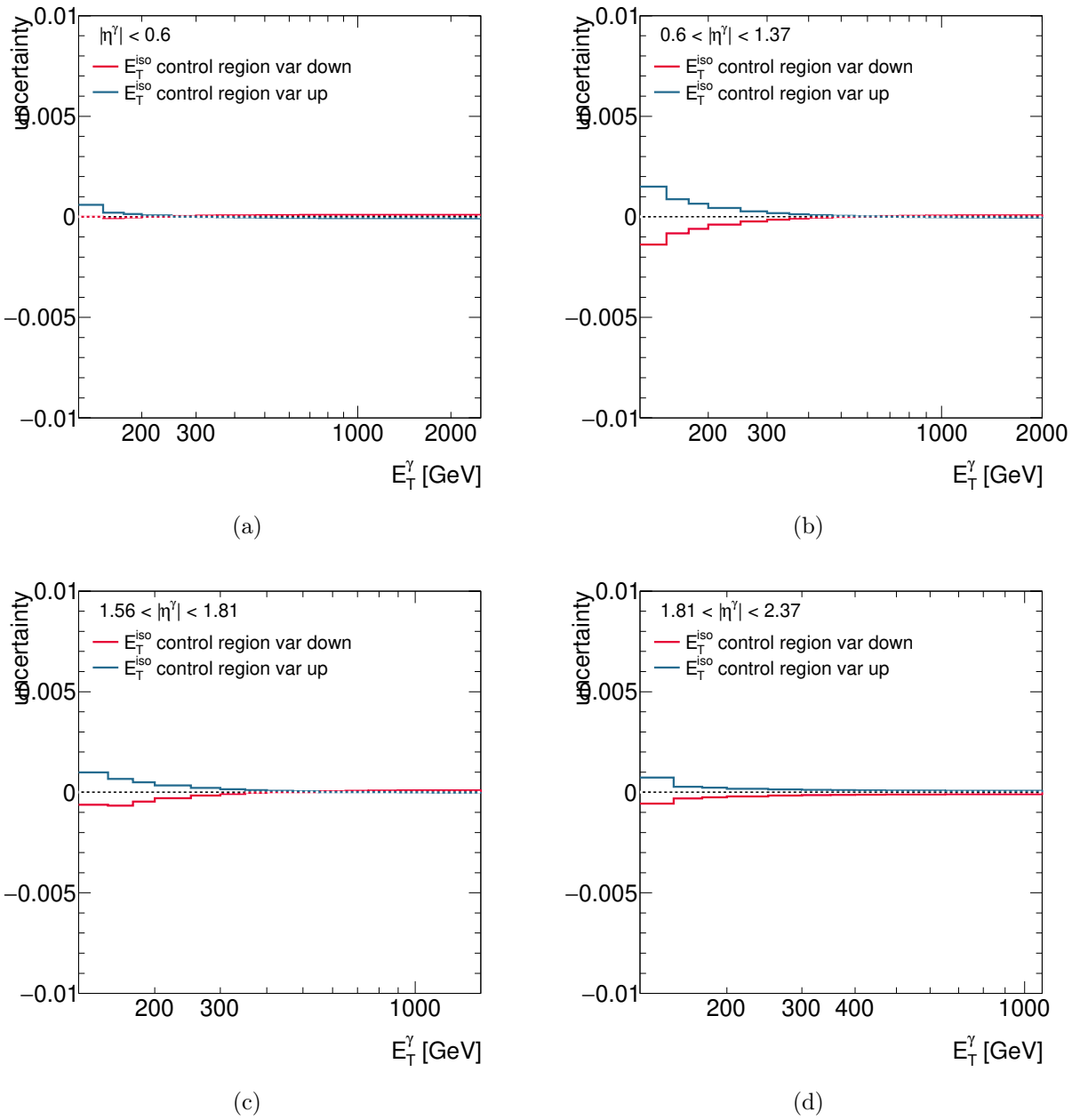


Figure 7.21: Systematic uncertainty in the measured cross section due to the choice of background control regions (different $E_{\mathrm{T}}^{\mathrm{iso}}$ requirements) as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. Statistical fluctuations in the estimation of this systematic uncertainty were taken care of by using a bootstrap technique.

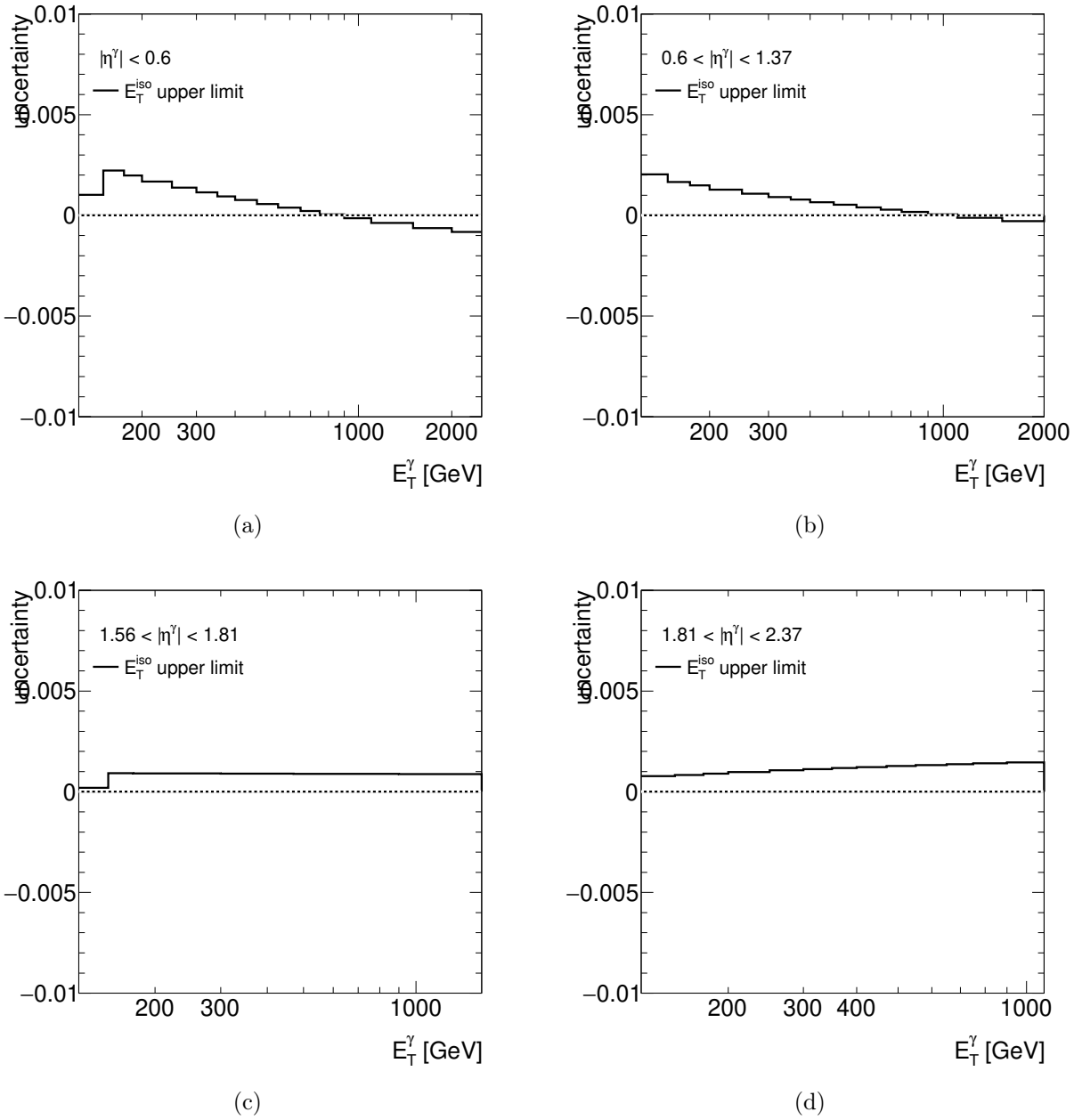


Figure 7.22: Systematic uncertainty in the measured cross section due to the choice of background control regions (removal of the upper limit on $E_{\mathrm{T}}^{\mathrm{iso}}$ for regions B and D) as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. Statistical fluctuations in the estimation of this systematic uncertainty were taken care of by using a bootstrap technique.

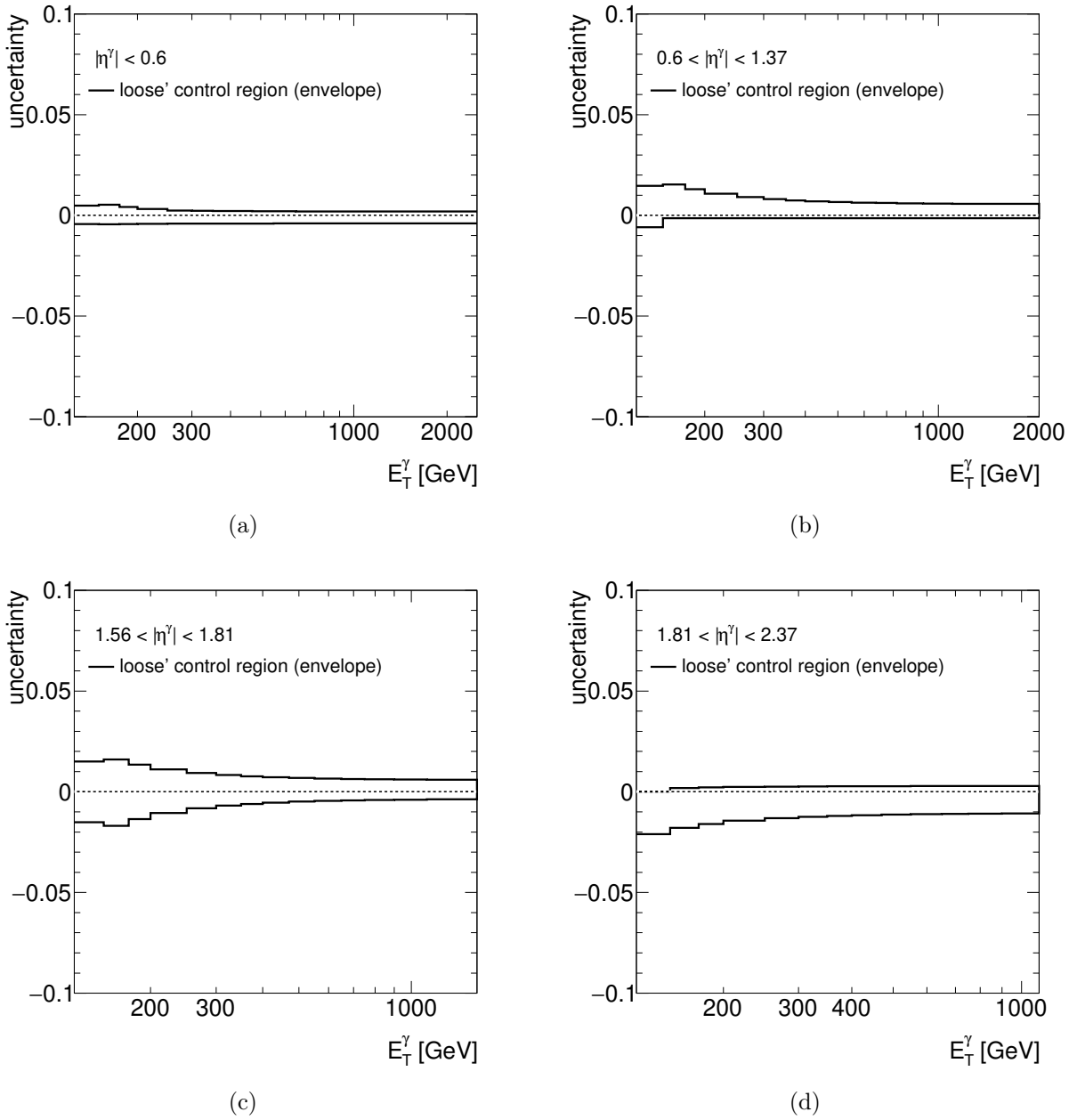


Figure 7.23: Systematic uncertainty in the measured cross section due to the choice of background control regions (different loose′ criteria) as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. Statistical fluctuations in the estimation of this systematic uncertainty were taken care of by using a bootstrap technique.

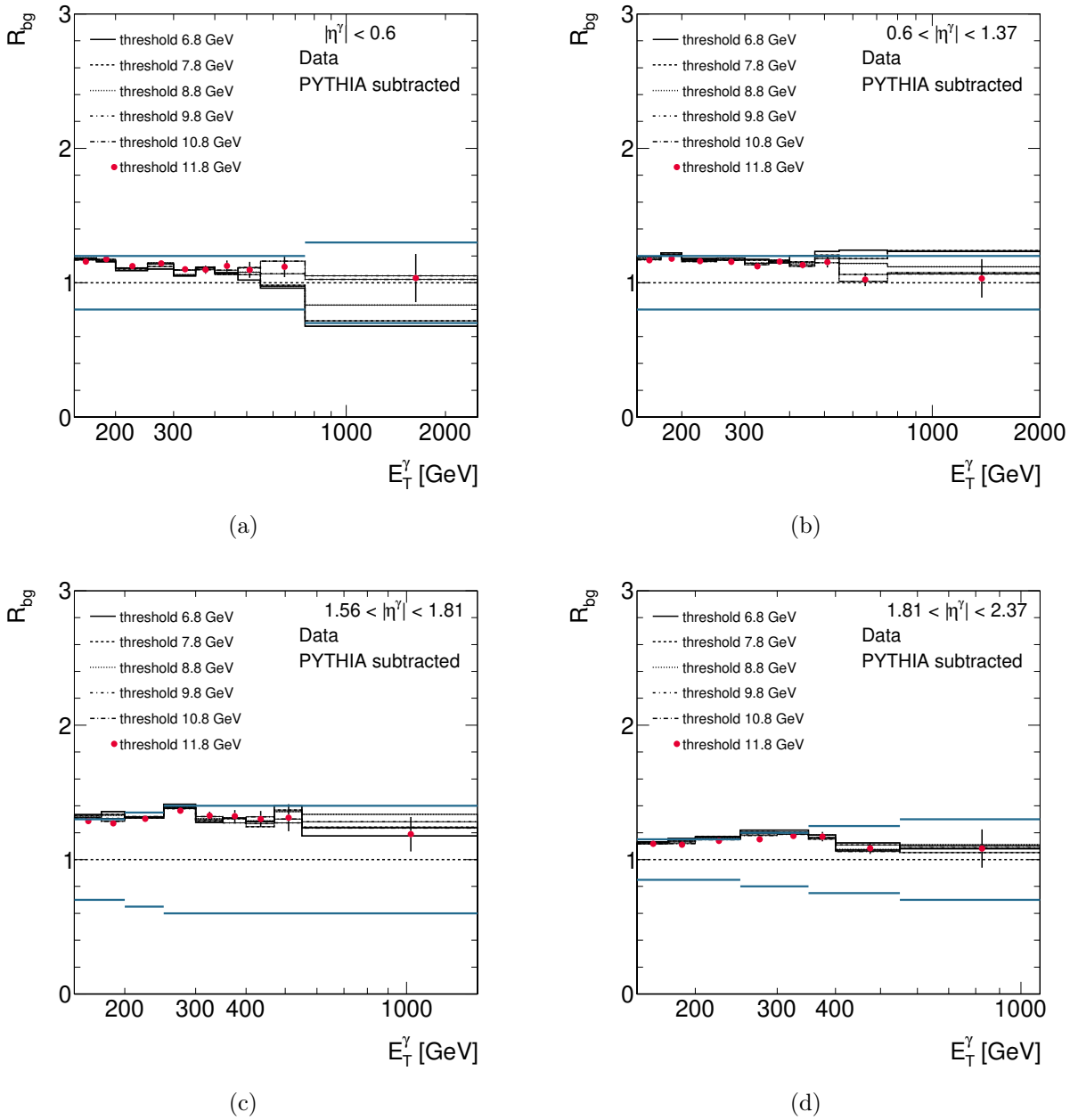


Figure 7.24: The estimated R^{bg} as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$ and for different definitions of the control regions. The signal leakage contributions in the control regions were removed using the PYTHIA samples. The range in R^{bg} which was taken as the uncertainty is indicated by the solid blue lines. Results are shown for the subsample of $E_{\text{T}}^{\gamma} > 150$ GeV.

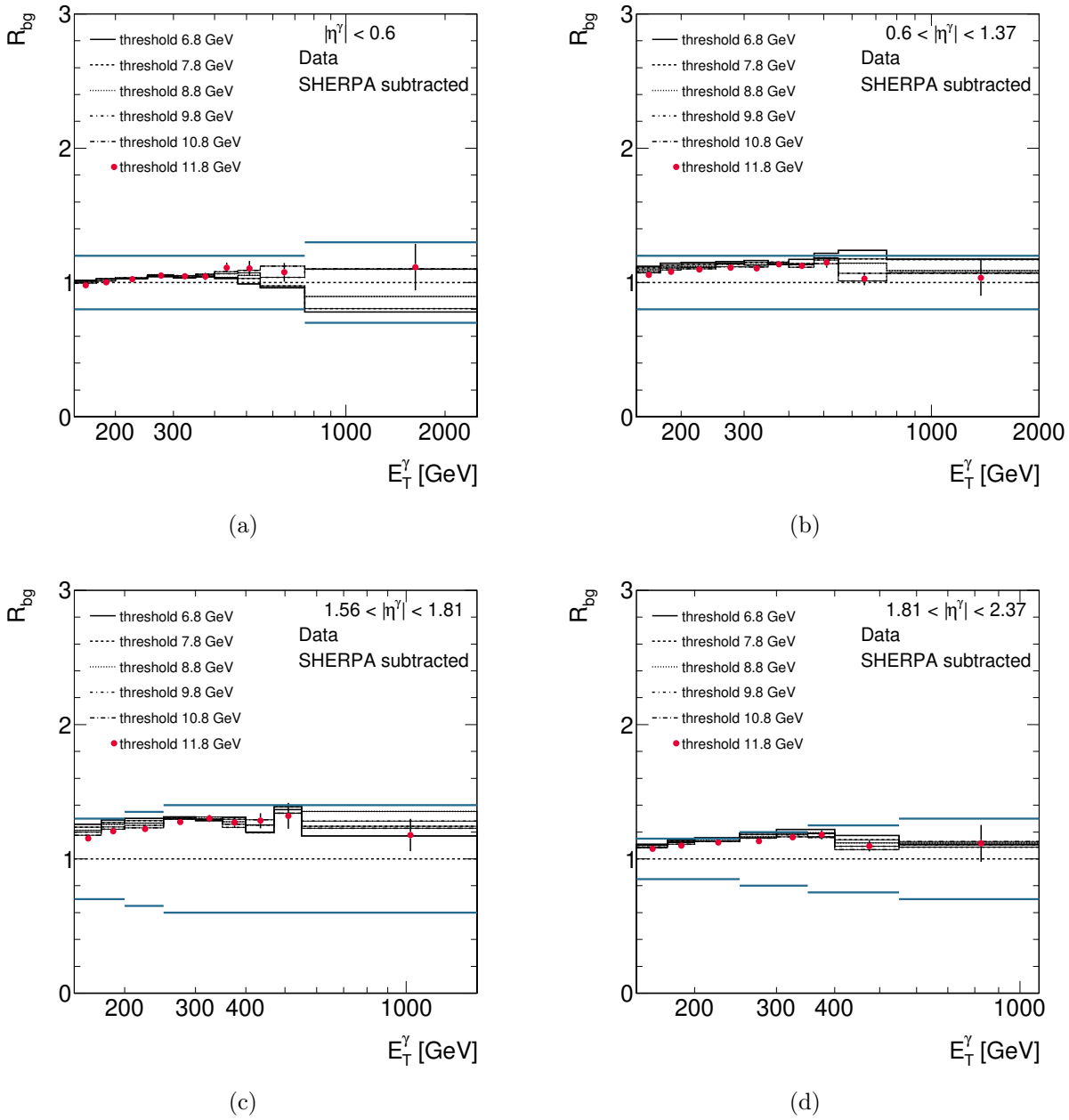


Figure 7.25: The estimated R^{bg} as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$ and for different definitions of the control regions. The signal leakage contributions in the control regions were removed using the SHERPA samples. The range in R^{bg} which was taken as the uncertainty is indicated by the solid blue lines. Results are shown for the subsample of $E_{\text{T}}^{\gamma} > 150$ GeV.

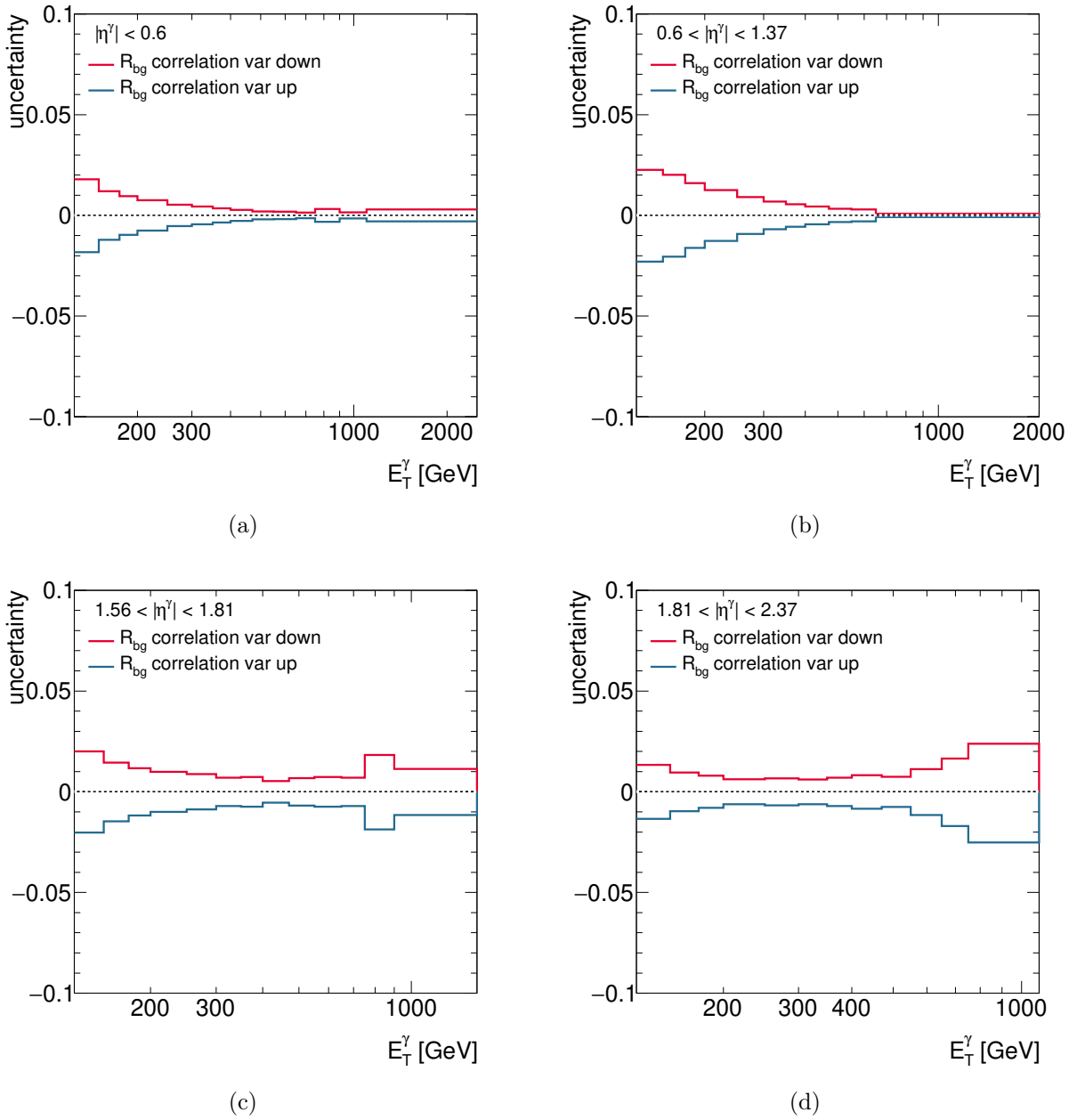


Figure 7.26: Systematic uncertainty in the measured cross section due to the correlation of the identification and isolation variables in the background as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$.

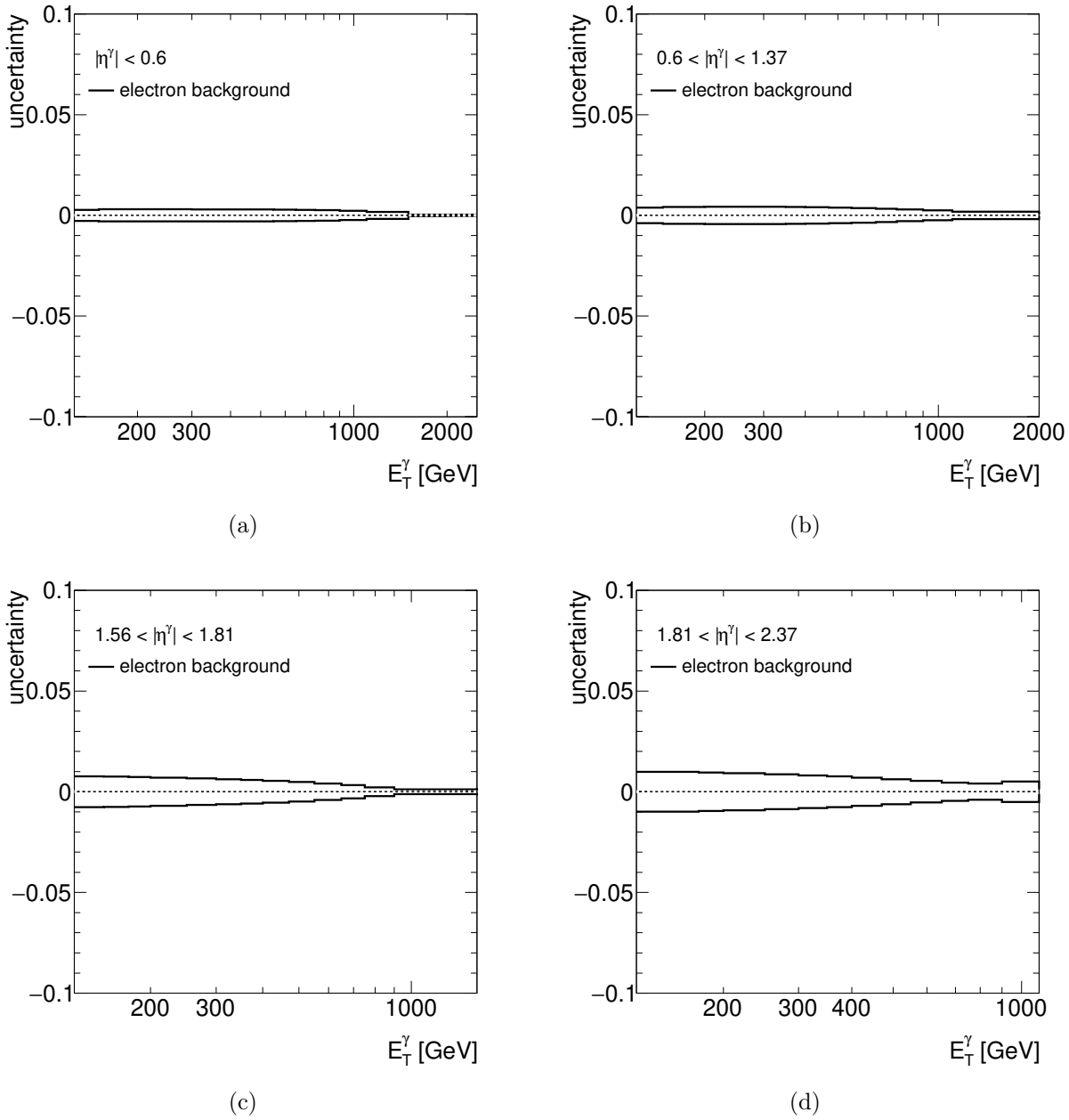


Figure 7.27: Systematic uncertainty in the measured cross section due to the electron background as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$.

7.7.3 Photon reconstruction

Photon-identification efficiency

The uncertainty in the efficiency scale factors, which were applied to simulated events to match the tight photon identification efficiency measured in data [129], is propagated to the measured cross section. The resulting uncertainty in the cross section is shown in figure 7.28, and is in the range $1\% - 3\%$.

$E_{\mathrm{T}}^{\mathrm{iso}}$ modelling

The uncertainty in the cross section due to the modelling of $E_{\mathrm{T}}^{\mathrm{iso}}$ in simulated events is evaluated by comparing the nominal results with those obtained using MC samples of events in which the data-driven correction to $E_{\mathrm{T}}^{\mathrm{iso}}$ is not applied (see section 5.5). The resulting uncertainty in the measured cross section is less than 2% except for $E_{\mathrm{T}}^{\gamma} > 500$ GeV in the regions $0.6 < |\eta^{\gamma}| < 1.37$ and $1.56 < |\eta^{\gamma}| < 1.81$, where it increases to 3%. Figure 7.29 shows the resulting uncertainty in the measured cross section.

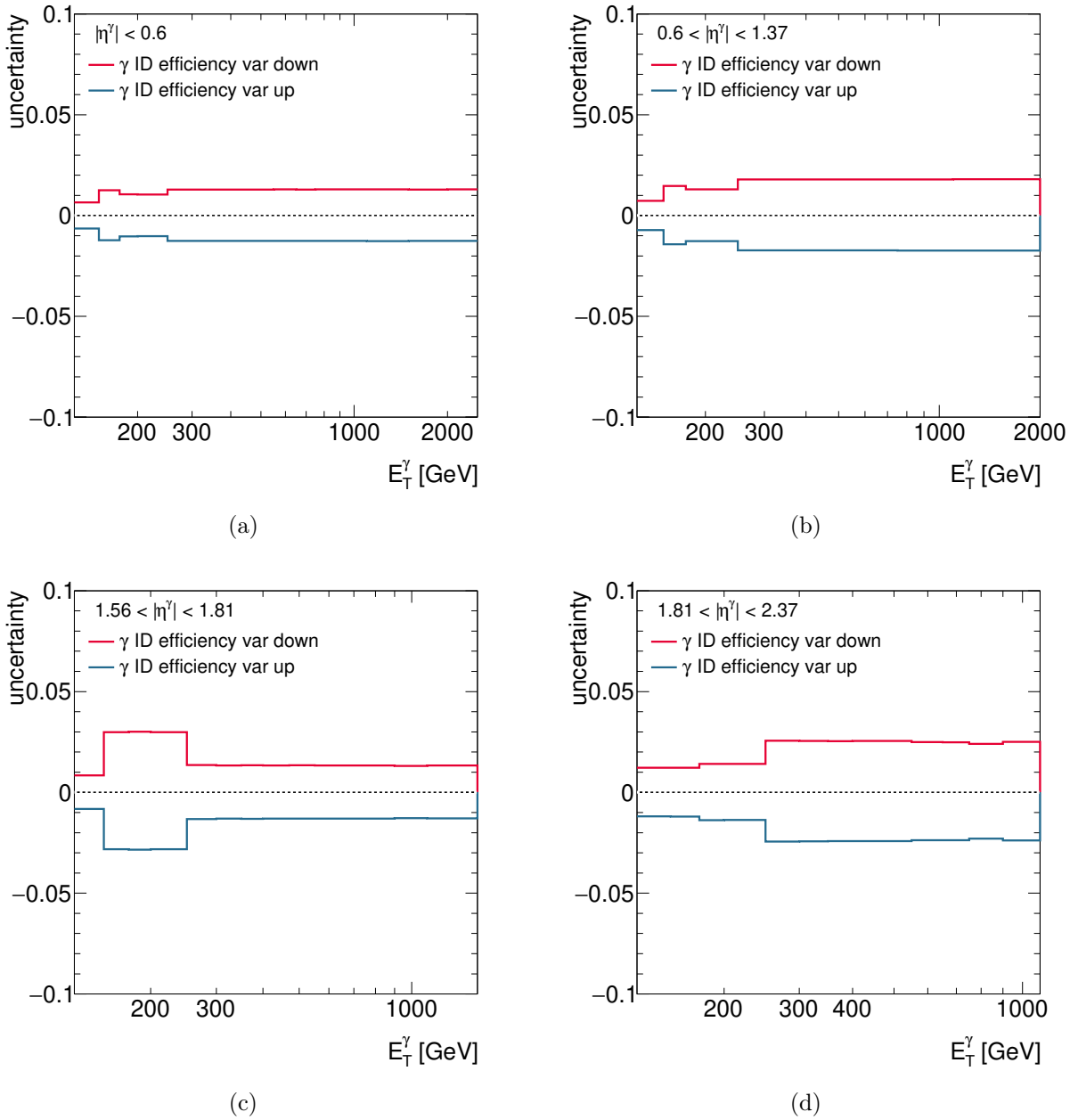


Figure 7.28: Systematic uncertainty in the measured cross section due to the uncertainty in the photon identification efficiency as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$.

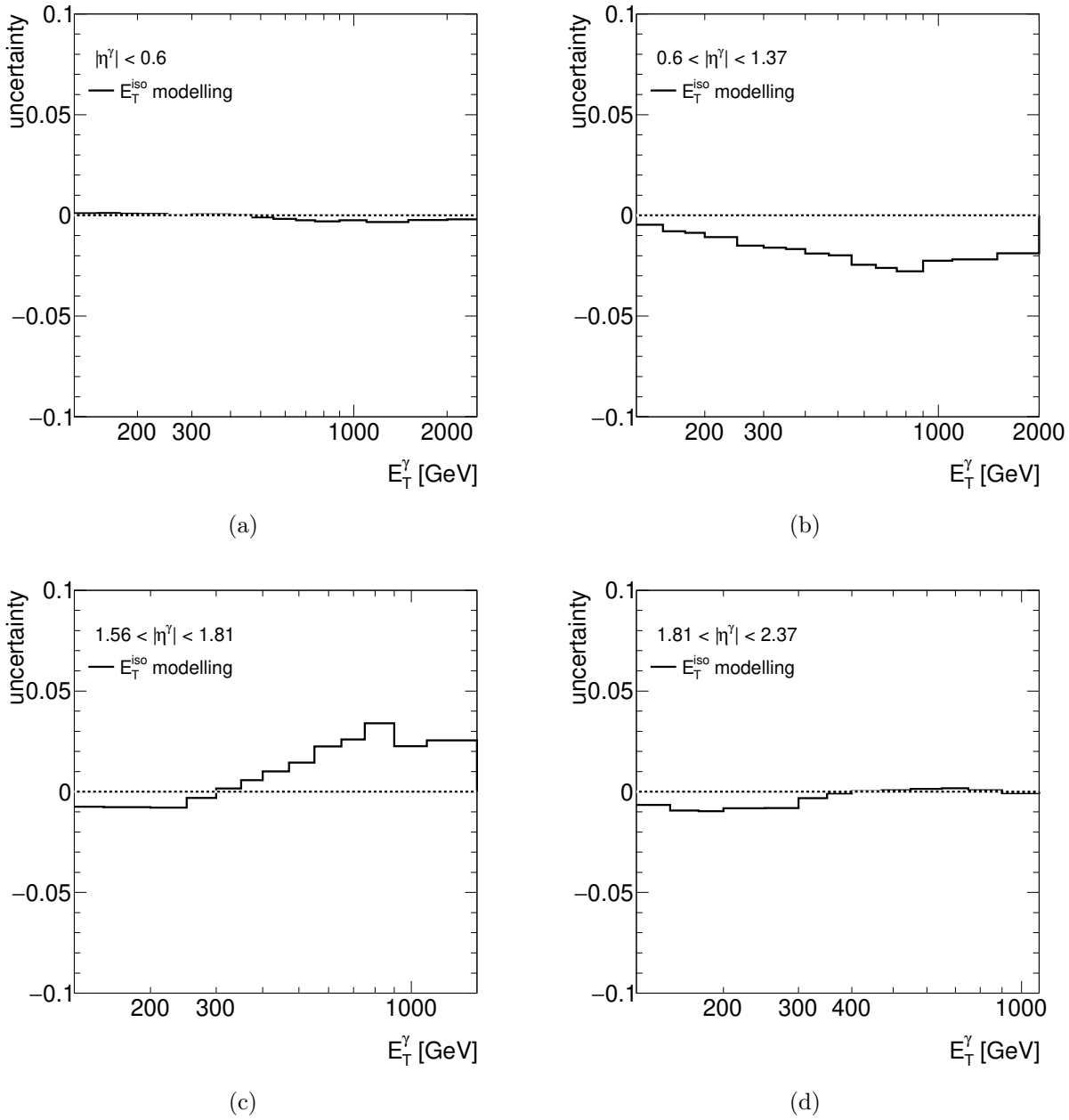


Figure 7.29: Systematic uncertainty in the measured cross section due to the effect of the isolation correction as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$.

7.7.4 Unfolding procedure

QCD-cascade and hadronisation model dependence

The nominal unfolding correction factors are computed using the PYTHIA simulation. As discussed in section 7.6, the effects on the unfolding correction factors due to the matrix elements and the parton shower and hadronisation models employed in the generators are investigated. The difference in the cross section between the nominal result and that obtained using the SHERPA simulation for the determination of the unfolding correction factors is taken as a systematic uncertainty. The uncertainty in the measured cross section is less than 2% and typically less than 1%. Figure 7.30 shows the resulting uncertainty in the measured cross section.

MC-sample statistics

The uncertainty in the cross section due to the statistical uncertainty of the MC samples is propagated into the measured cross section and typically amounts to less than 0.5%. Figure 7.31 shows the resulting uncertainty in the measured cross section.

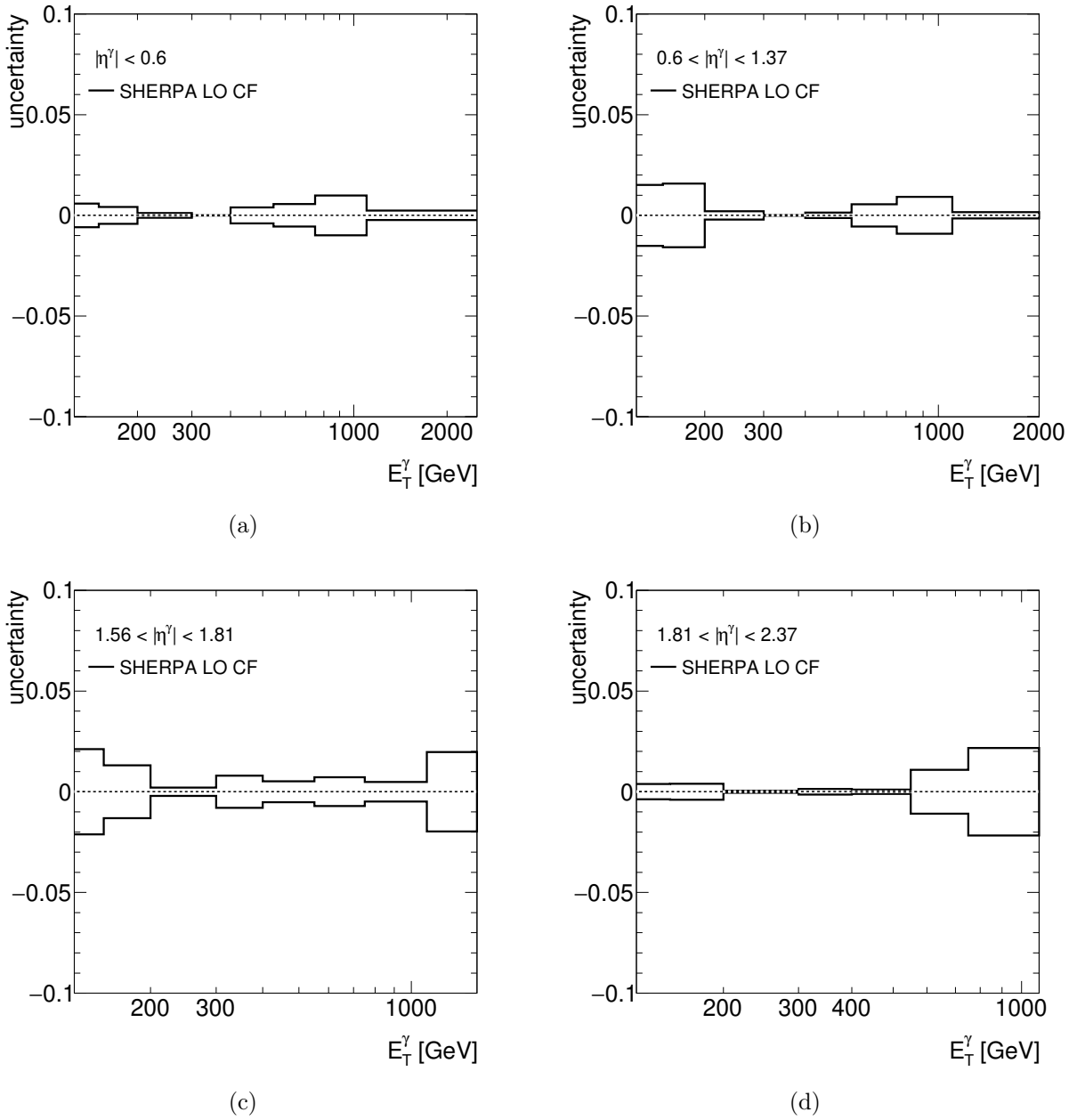


Figure 7.30: Systematic uncertainty in the measured cross section due to the QCD-cascade and hadronisation models as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. Statistical uncertainties in the estimation of this systematic uncertainty were taken care of by rebinning, merging those bins affected by relatively large statistical fluctuations with neighbouring bins.

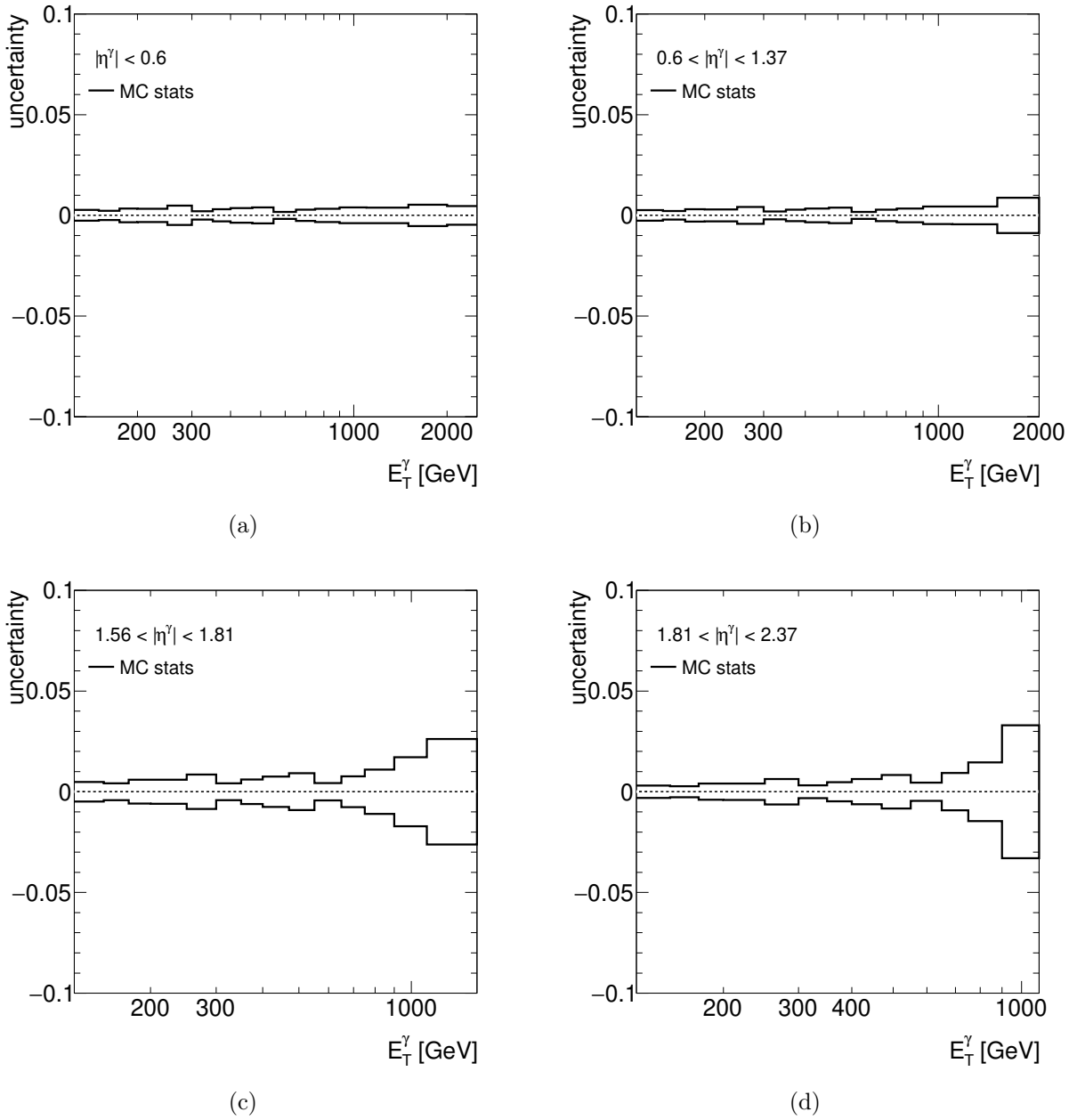


Figure 7.31: Statistical uncertainty of the MC samples as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$.

7.7.5 Running conditions

Pile-up

A variation in the pile-up reweighting of simulated events is included to cover the uncertainty in the ratio of the predicted to measured inelastic cross sections [49]. To match the in-time and out-of-time pile-up conditions in the data, the distribution of $\langle\mu\rangle$ in simulated events was reweighted to that of the data after applying a factor 1/1.09 (see section 5.6). The uncertainty was estimated by changing the factor applied to the data to 1/1.0 or 1/1.18. The resulting uncertainty in the measured cross section is shown in figure 7.32, and is less than 1.5%.

Trigger efficiency

The uncertainty in the trigger efficiency is estimated using the same methodology as in ref. [38], and is propagated into the measured cross section. The uncertainty in the measured cross section is at most 0.4% (see table 7.2).

Integrated luminosity

The uncertainty in the integrated luminosity is 2.1%. It is fully correlated between all bins with $E_{\mathrm{T}}^{\gamma} > 150$ GeV of the cross section. For the bin with $125 < E_{\mathrm{T}}^{\gamma} < 150$ GeV only data from the 2015 running period is used and the uncertainty in the luminosity measurement is not fully correlated with the uncertainty in the combined integrated luminosity corresponding to the 2015 and 2016 running periods. However, the impact of the uncorrelated part can be safely neglected.

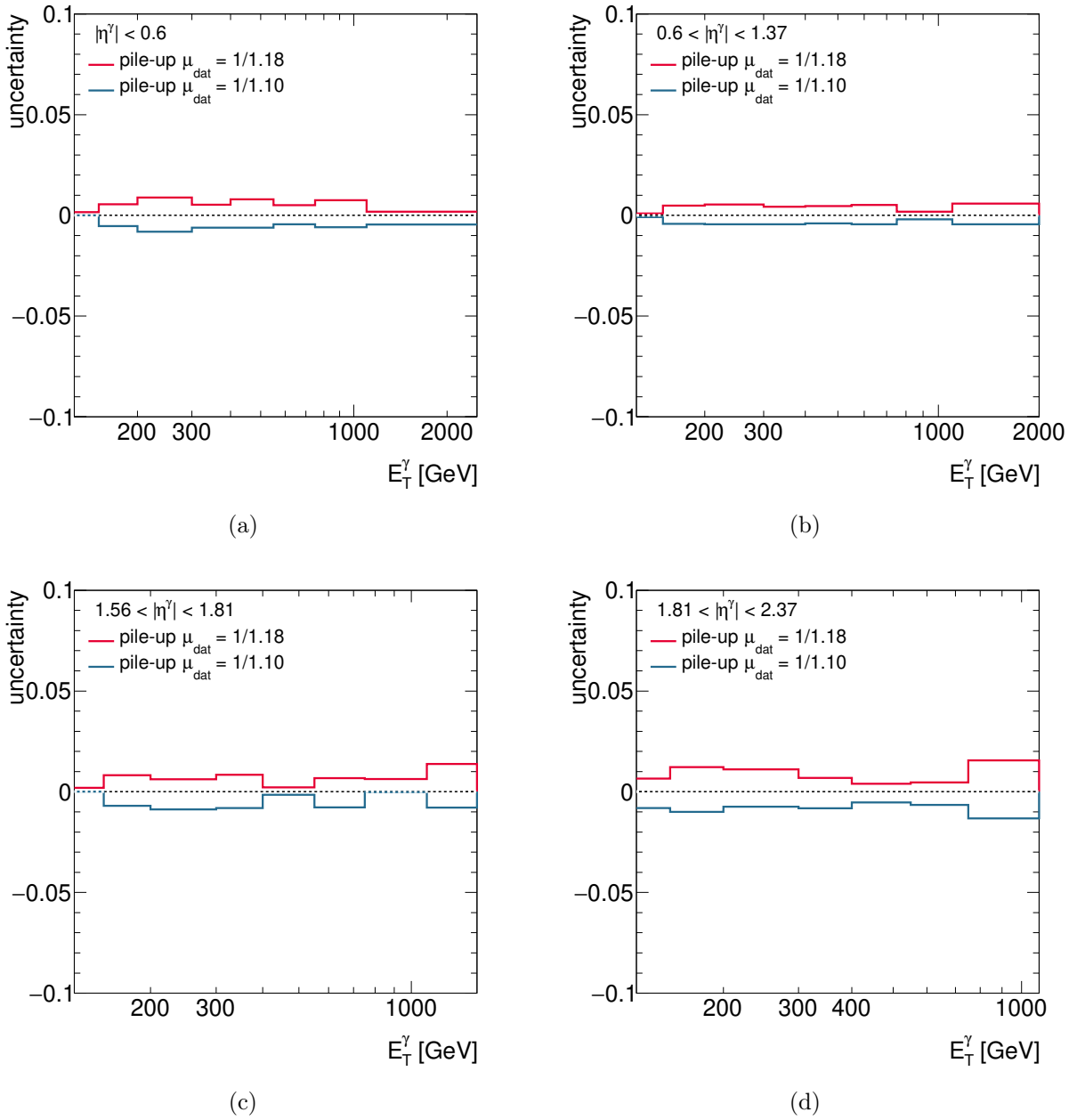


Figure 7.32: Systematic uncertainty in the measured cross section due to the pile-up reweighting as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. Statistical fluctuations in the estimation of this systematic uncertainty were taken care of by rebinning, merging those bins affected by relatively large statistical fluctuations with neighbouring bins.

7.7.6 Photon energy scale and resolution

A detailed analysis of the sources of uncertainty in the energy scale and resolution for photons was made with Run I data [155]. The same model was implemented and updated for Run II data [130]. The sources of uncertainty reported in ref. [130] include: the uncertainty in the overall energy scale adjustment using $Z \to e^+e^-$ decays; the uncertainty in the non-linearity of the energy measurement at the cell level; the uncertainty in the relative calibration of the different calorimeter layers; the uncertainty in the amount of material in front of the calorimeter; the uncertainty in the modelling of the reconstruction of photon conversions; the uncertainty in the modelling of the lateral shower shape; the uncertainty in the modelling of the sampling term[3]; the uncertainty in the measurement of the constant term in Z-boson decays. The sources of uncertainty are modelled using independent components to account for their η dependence. A total of 76 individual components influencing the energy scale and resolution of the photon are identified to asses the overall uncertainty in the energy measurement. All the uncertainty components are propagated separately through the analysis to keep track of the information about the correlations between different bins. The systematic uncertainty in the cross section is evaluated by varying each individual source of uncertainty separately by $\pm 1\sigma$ in the MC simulations and then adding the uncertainty contributions in quadrature. The resulting uncertainties in the measured cross section are typically less than 0.5% for the energy resolution and in the range $1\% - 16\%$ for the energy scale. The resulting uncertainty in the measured cross section is shown in figure 7.33 (7.34) for the energy resolution (scale). The three main contributions to the photon energy scale uncertainty in each $|\eta^\gamma|$ region are also shown in Figure 7.35.

[3]The relative energy resolution is parameterised as $\sigma(E)/E = a/\sqrt{E} \oplus b/E \oplus c$, where a is the sampling term, b is the noise term and c is the constant term.

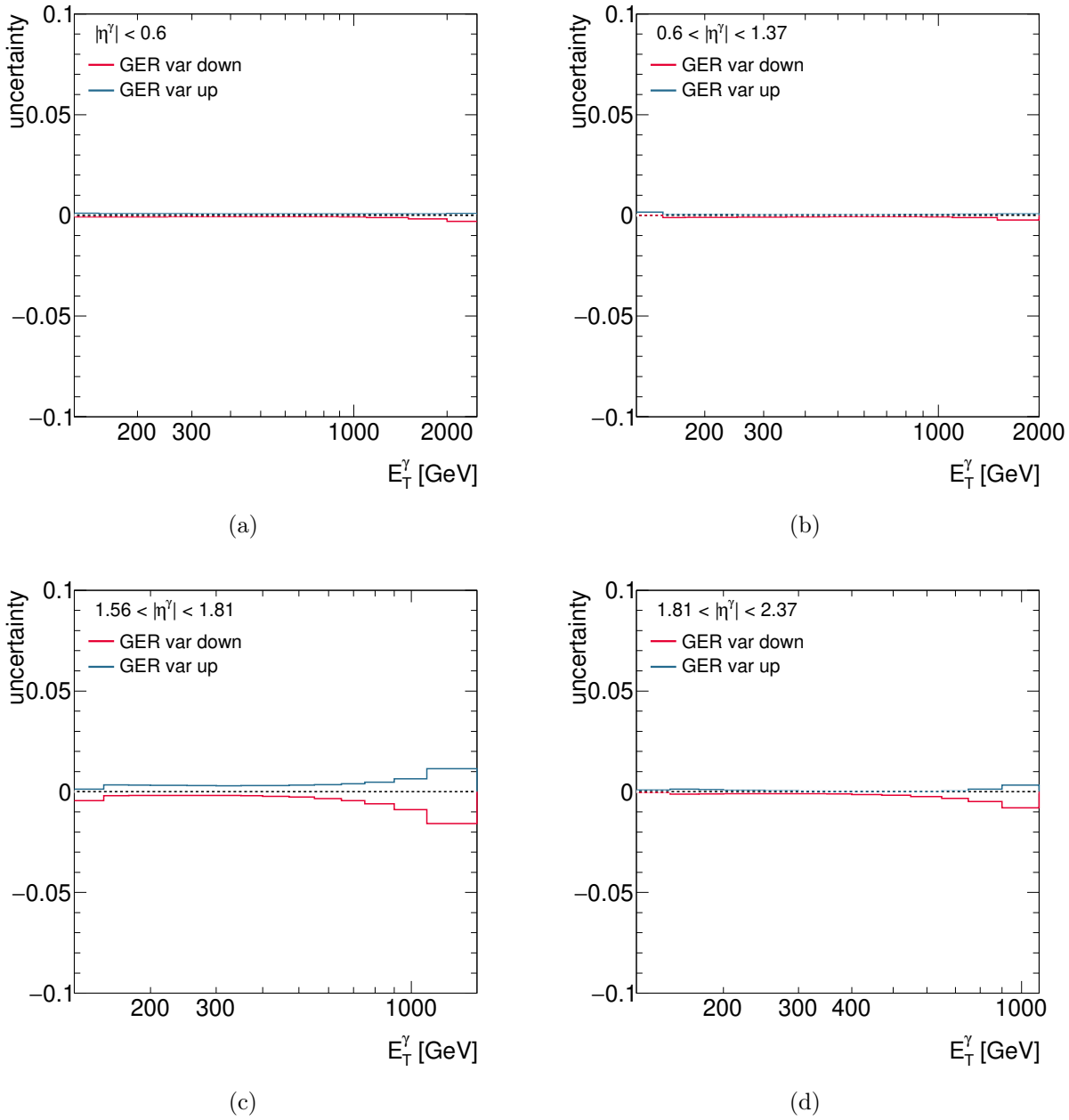


Figure 7.33: Systematic uncertainty in the measured cross section due to the photon energy resolution uncertainty as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. Statistical fluctuations in the estimation of this systematic uncertainty were taken care of by using a bootstrap technique on each individual component.

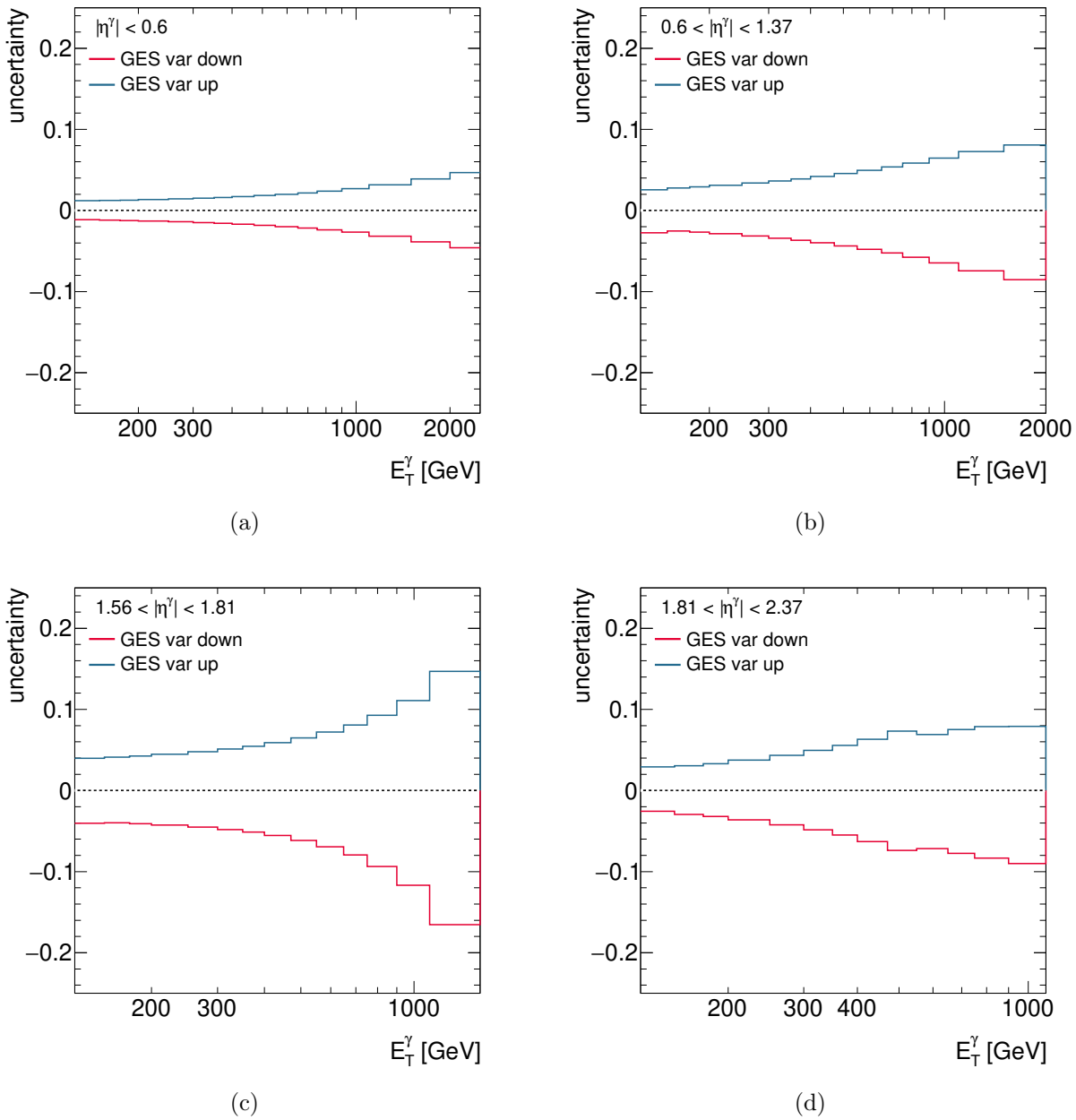


Figure 7.34: Systematic uncertainty in the measured cross section due to the photon energy scale uncertainty as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. Statistical fluctuations in the estimation of this systematic uncertainty were taken care of by using a bootstrap technique on each individual component.

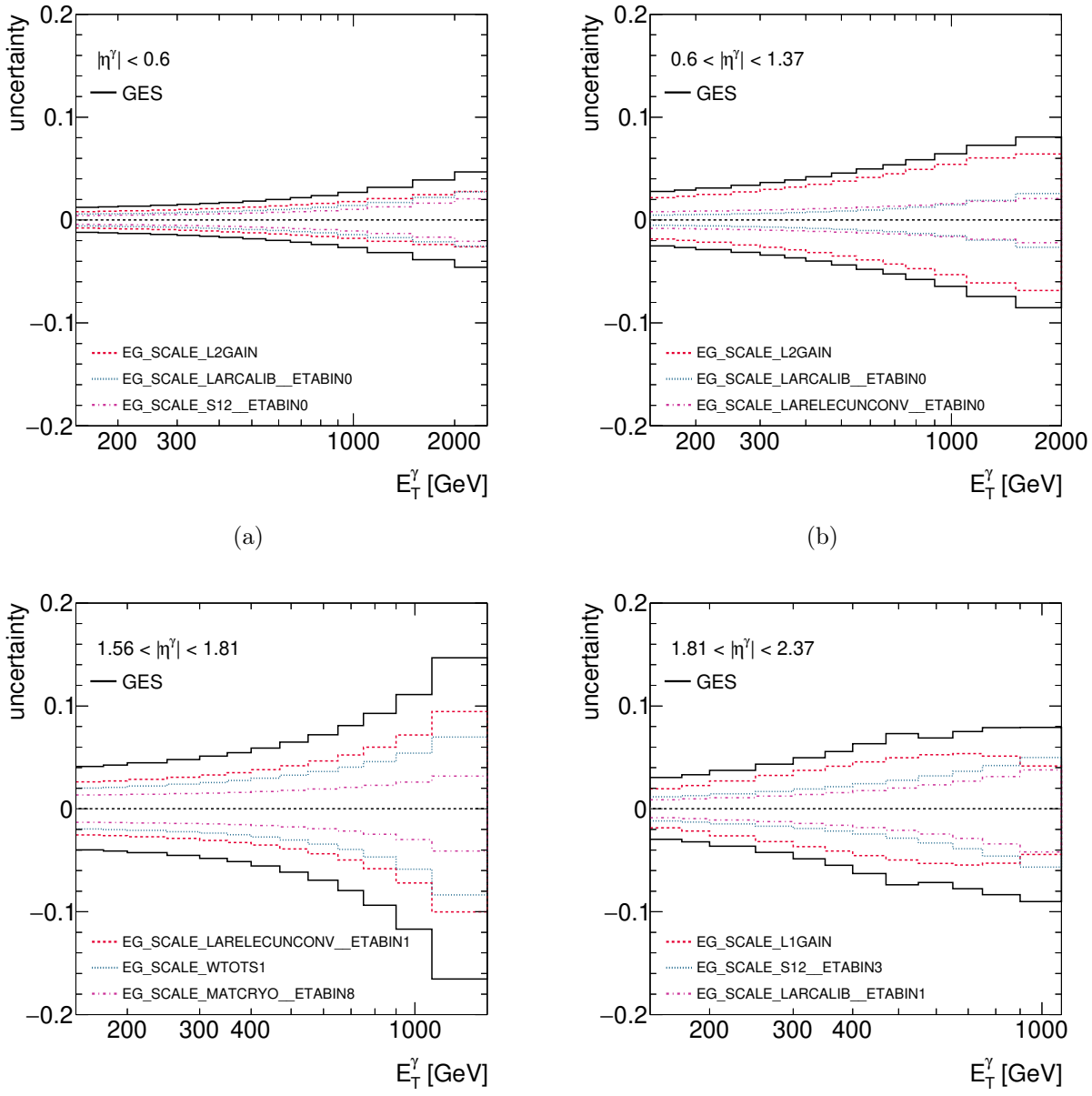


Figure 7.35: Systematic uncertainty in the measured cross section due to the photon energy scale uncertainty as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. The total uncertainty is represented with the black line; the three main contributions in each $|\eta^{\gamma}|$ region are also shown. Results are shown for the subsample of $E_{\mathrm{T}}^{\gamma} > 150$ GeV.

7.7.7 Total systematic uncertainty

The total systematic uncertainty is calculated by adding in quadrature the uncertainties listed above. The total systematic is in the range $3\% - 17\%$, depending on E_{T}^{γ} and the $|\eta^{\gamma}|$ region. The dominant sources of uncertainty arise from the photon energy scale, the photon identification efficiency and the integrated luminosity. Figure 7.36 shows the total systematic uncertainty for each measured cross section together with the dominant components. The systematic uncertainty dominates the total experimental uncertainty for $E_{\mathrm{T}}^{\gamma} \lesssim 1$ TeV and $|\eta^{\gamma}| < 1.37$, whereas for higher E_{T}^{γ} values, the statistical uncertainty of the data limits the precision of the measurement, as can be seen in figure 7.37. This represents an improvement relative to the previous measurement [10], where the region dominated by the systematic uncertainty was $E_{\mathrm{T}}^{\gamma} \lesssim 600$ GeV.

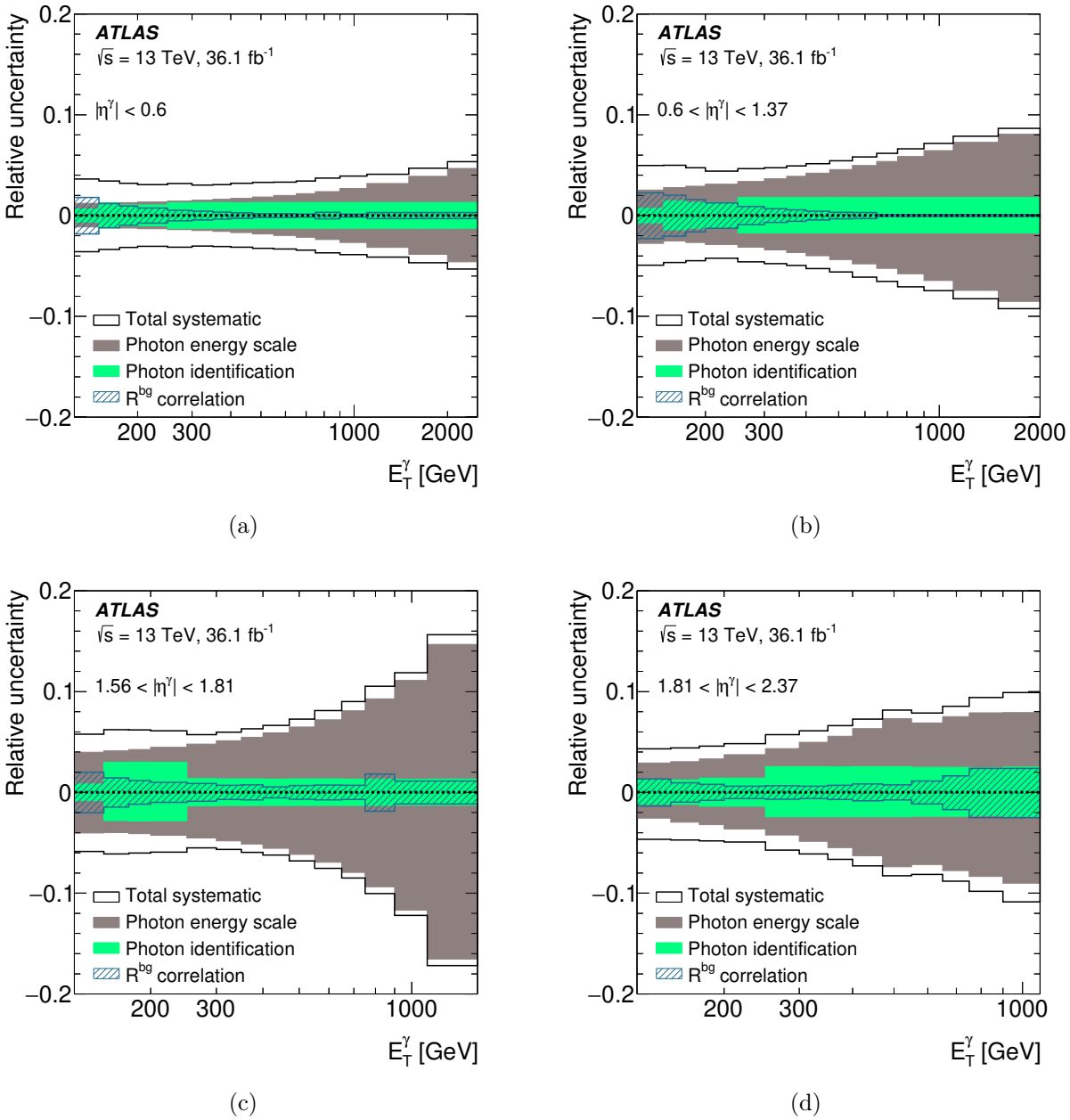


Figure 7.36: The relative total systematic uncertainty in the cross section (white areas) as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. The relative uncertainty due to the photon energy scale (grey areas), the relative uncertainty due to the photon identification efficiency (green areas) and the relative uncertainty due to R^{bg} (blue hatched areas) are also shown.

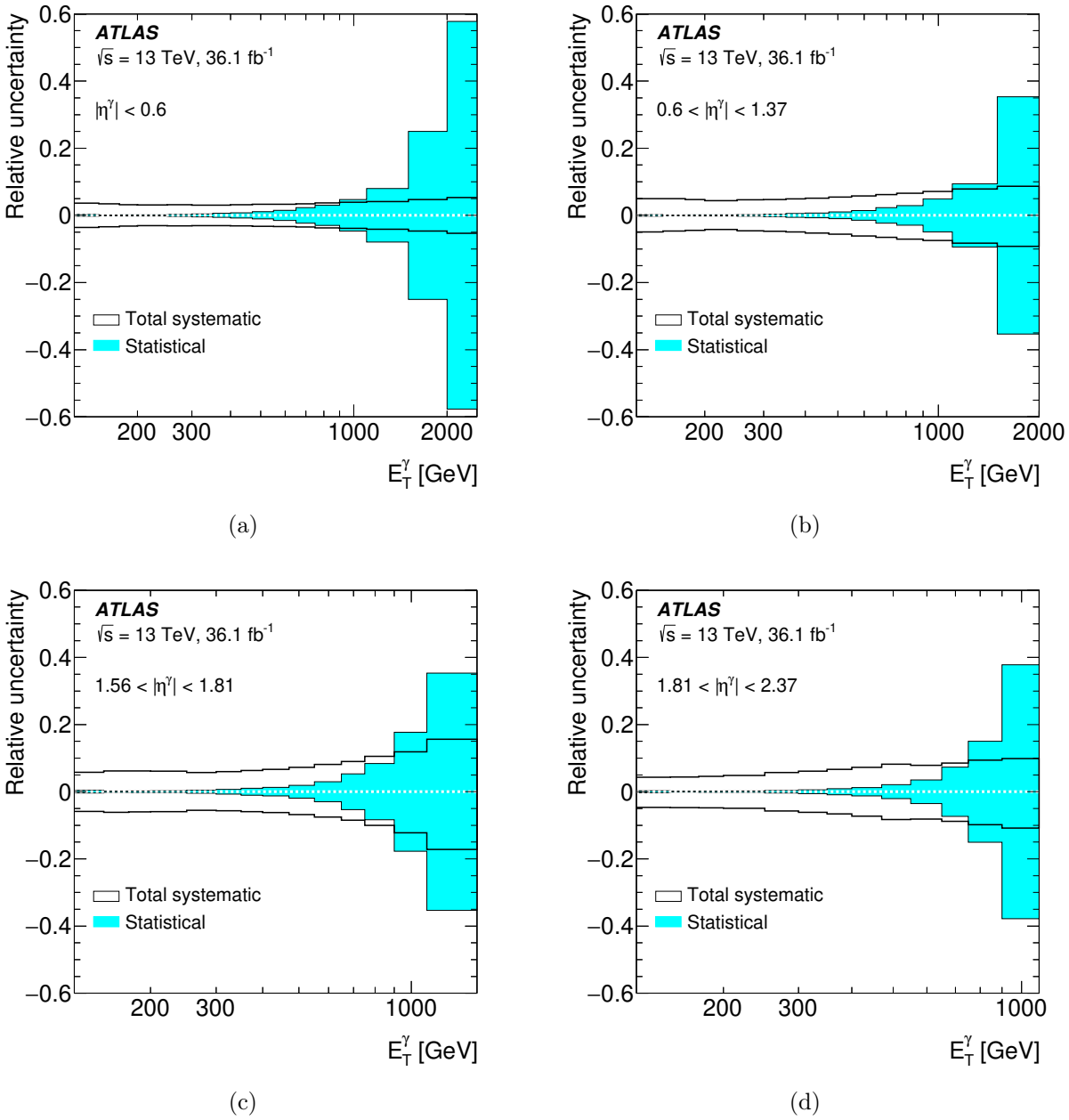


Figure 7.37: The relative total systematic uncertainty in the cross section (white areas) as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. The relative statistical uncertainty from the data is also shown (cyan areas).

7.8 Theory predictions

The NLO pQCD predictions were calculated using two programs, namely JETPHOX 1.3.1_2, which is described in section 3.4.2.1, and SHERPA 2.2.2, which was introduced in section 4.3. In addition, the NNLO pQCD prediction was calculated with NNLOJET, and is described in section 3.4.2.2.

There are several differences between the calculations using JETPHOX, SHERPA 2.2.2 and NNLOJET: the calculations of NNLOJET include NNLO QCD corrections; the calculations using SHERPA 2.2.2 include higher-order corrections as well as parton showers; the application of the Frixione's criterion (NNLOJET and SHERPA 2.2.2) at matrix-element level allows the fragmentation contribution to be ignored; and the prediction for the cross section using SHERPA 2.2.2 is at particle level and includes UE effects. A compilation of the major features of the three different approaches is shown in table 7.3.

Program	Order in α_S	Fragmentation	Parton shower	Isolation cone	α_{EM}	Particle level
JETPHOX	NLO	Yes	No	Fixed	1/137.036	No
SHERPA 2.2.2	NLO for γ+ (1,2)-jet LO for γ+ (3,4)-jet	No	Yes	Hybrid	1/137.036	Yes
NNLOJET	NNLO	No	No	Hybrid	1/132.232	No

Table 7.3: Major features of the three predictions used for inclusive isolated-photon production.

Electroweak corrections are not included in the calculations of JETPHOX, NNLOJET or SHERPA 2.2.2. As mentioned in section 3.4.3, the corrections are likely to play an important role at the TeV scale.

7.8.1 Hadronisation and UE corrections to the NLO QCD calculations

Since the NLO pQCD prediction from JETPHOX is at parton level while the measurement is at the particle level, correction factors for the non-perturbative (NP) effects of hadronisation and UE are needed for the prediction to match the data. The data are corrected for pile-up and UE effects and the distributions are unfolded to a phase-space definition in which the requirement on $E_{\mathrm{T}}^{\mathrm{iso}}$ at particle level is applied after the subtraction of the UE. Therefore, corrections for NP effects are expected to be close to unity. They are evaluated by computing the ratio of the particle-level cross section for a PYTHIA sample with UE effects to the parton-level cross section without UE effects (see section 4.3). Figure 7.38 shows the hadronisation and UE correction factors. No significant dependence on η^{γ} or the MC tuning[4] used was observed for any of the correction factors. The

[4]The parton-to-hadron corrections were also estimated using PYTHIA 8.186 with the LO NNPDF2.3 A14 tune as well as two variations around it, namely Var1Up and Var1Down. The corrections thus

correction factors were found to be consistent with unity within ±1%; thus, no correction was applied to the NLO pQCD prediction of JETPHOX and an uncertainty of 1% was assigned.

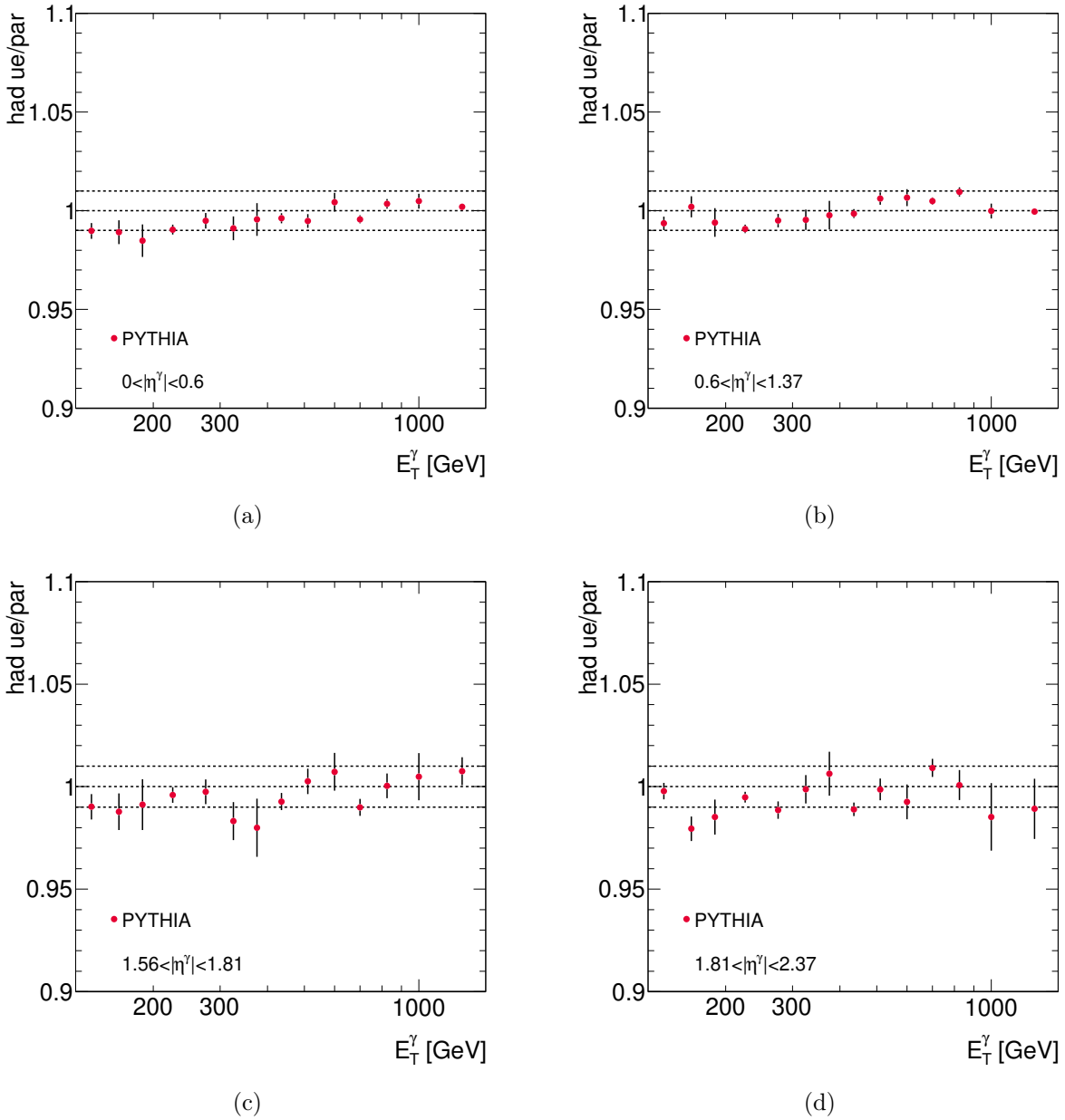


Figure 7.38: Hadronisation and UE correction factors from PYTHIA (dots) as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$.

obtained were consistent with the results based on PYTHIA 8.165 with the AU2 CTEQ6L1 tune within ±1%.

7.8.2 Theoretical uncertainties in the NLO pQCD calculations

The major source of uncertainty affecting the predictions is due to terms beyond NLO in QCD. This uncertainty is evaluated by repeating the calculations using values of μ_{R}, μ_{F}, and μ_{f} scaled by the factors 0.5 and 2. The three scales are either varied simultaneously or independently; in addition, configurations in which one scale is fixed and the other two are varied simultaneously are also considered. In all cases, the condition $0.5 \leq \mu_{\mathrm{A}}/\mu_{\mathrm{B}} \leq 2$ is imposed, where $\mathrm{A}, \mathrm{B} = \mathrm{R}, \mathrm{F}, \mathrm{f}$. The final uncertainty is then computed as the largest deviation from the nominal value among the 14 possible variations; this is done separately for the positive and negative contributions, so the uncertainty is not necessarily symmetric. In the case of SHERPA 2.2.2, which does not include the fragmentation contribution, μ_{R} and μ_{F} are varied as above and the largest deviation from the nominal prediction among the six possible variations is taken as the uncertainty. The resulting uncertainty in the prediction is in the range $10\% - 15\%$ ($20\% - 30\%$) for JETPHOX (SHERPA 2.2.2). The fact that the uncertainties are larger for SHERPA 2.2.2 than for JETPHOX is attributed to multijet configurations which are accounted for only in the case of SHERPA 2.2.2 [37].

The uncertainty in the prediction of JETPHOX due to the uncertainty in the proton PDF was estimated by repeating the calculations using the 50 sets from the MMHT2014 error analysis and applying the Hessian method [156]. In the case of SHERPA 2.2.2, it is estimated using the 100 replicas from the NNPDF3.0 analysis. The resulting uncertainty in the prediction is in the range $1\% - 6\%$ for JETPHOX and $1\% - 9\%$ for SHERPA 2.2.2, depending on E_{T}^{γ} and the $|\eta^{\gamma}|$ region.

The uncertainty in the prediction of JETPHOX (SHERPA 2.2.2) due to the uncertainty in α_{S} is evaluated by repeating the calculations using two additional sets of proton PDF from the MMHT2014 (NNPDF3.0) analysis for which different values of $\alpha_{\mathrm{S}}(M_{\mathrm{Z}}^2)$ are assumed in the fits, namely 0.118 (0.117) and 0.122 (0.119); in this way, the correlation between α_{S} and the PDF is preserved. The uncertainties in JETPHOX associated to α_{S} are evaluated as the differences between the varied and the nominal cross sections scaled by a factor 1.5/2. This scaling factor accounts for the difference in $\alpha_{\mathrm{S}}(M_{\mathrm{Z}}^2)$ between the nominal and the $\alpha_{\mathrm{S}}(M_{\mathrm{Z}}^2)$-varied calculations (0.002) and the uncertainty in $\alpha_{\mathrm{S}}(M_{\mathrm{Z}}^2)$ (0.0015) recommended by the authors of MMHT2014 [74]. The uncertainties in SHERPA 2.2.2 due to α_{S} are evaluated as the differences between the varied and the nominal cross sections scaled by a factor 1.5 to match the prescription used for MMHT2014. The resulting uncertainty in the prediction of JETPHOX (SHERPA 2.2.2) is less than 3% (6% except for the highest measured E_{T}^{γ} point in the region $1.81 < |\eta^{\gamma}| < 2.37$, where it increases to 12%).

In addition, an uncertainty of 1% was included due to the hadronisation corrections in JETPHOX, as mentioned in section 7.8.1. The SHERPA 2.2.2 prediction is based on particle-level observables after applying the requirements listed in table 7.1 and, therefore, can be compared directly with the measurement. Although no tunes are available to assess the impact of hadronisation and UE in the SHERPA 2.2.2 prediction, it is expected

that the associated uncertainty is of a similar size as that evaluated using PYTHIA, and, therefore, can be neglected in comparison to the other uncertainties.

The total theoretical uncertainty is obtained by adding in quadrature the individual uncertainties listed above (see figures 7.39 and 7.40). The dominant theoretical uncertainty for the predictions using JETPHOX or SHERPA 2.2.2 is that arising from the variation of the scales. The uncertainty arising from the PDF ($\alpha_S(M_Z^2)$) is the second dominant uncertainty for the prediction of JETPHOX (SHERPA 2.2.2). A prediction from JETPHOX using as scales $\mu_R = \mu_F = \mu_f = E_T^\gamma/2$ was also calculated together with the associated theoretical uncertainties.

7.8.3 Theoretical uncertainties in the NNLO pQCD calculations

The uncertainty in the NNLO pQCD prediction due to higher-order terms is evaluated by varying the renormalisation and factorisation scales. The resulting uncertainties are compared with those of the NLO pQCD prediction from NNLOJET in figure 7.41. The uncertainty in the NNLO pQCD prediction arising from the scale variations is in the range $0.6\% - 5\%$ and are smaller than those in the NLO pQCD prediction by a factor in the range $2 - 20$, depending on E_T^γ and $|\eta^\gamma|$. The uncertainties due to the PDF and α_S are not available, but they are expected to be similar to, or smaller than, those affecting the JETPHOX prediction. For the comparisons with the data in section 7.9, the total uncertainty associated with the NNLOJET prediction was taken as the sum in quadrature of the scale uncertainty, the statistical uncertainty of the calculation, the uncertainty of 1% assigned to the non-perturbative corrections (see section 7.8.1) and, as an approximation, the uncertainties due to the PDF and α_S as estimated from JETPHOX at NLO (see figure 7.42).

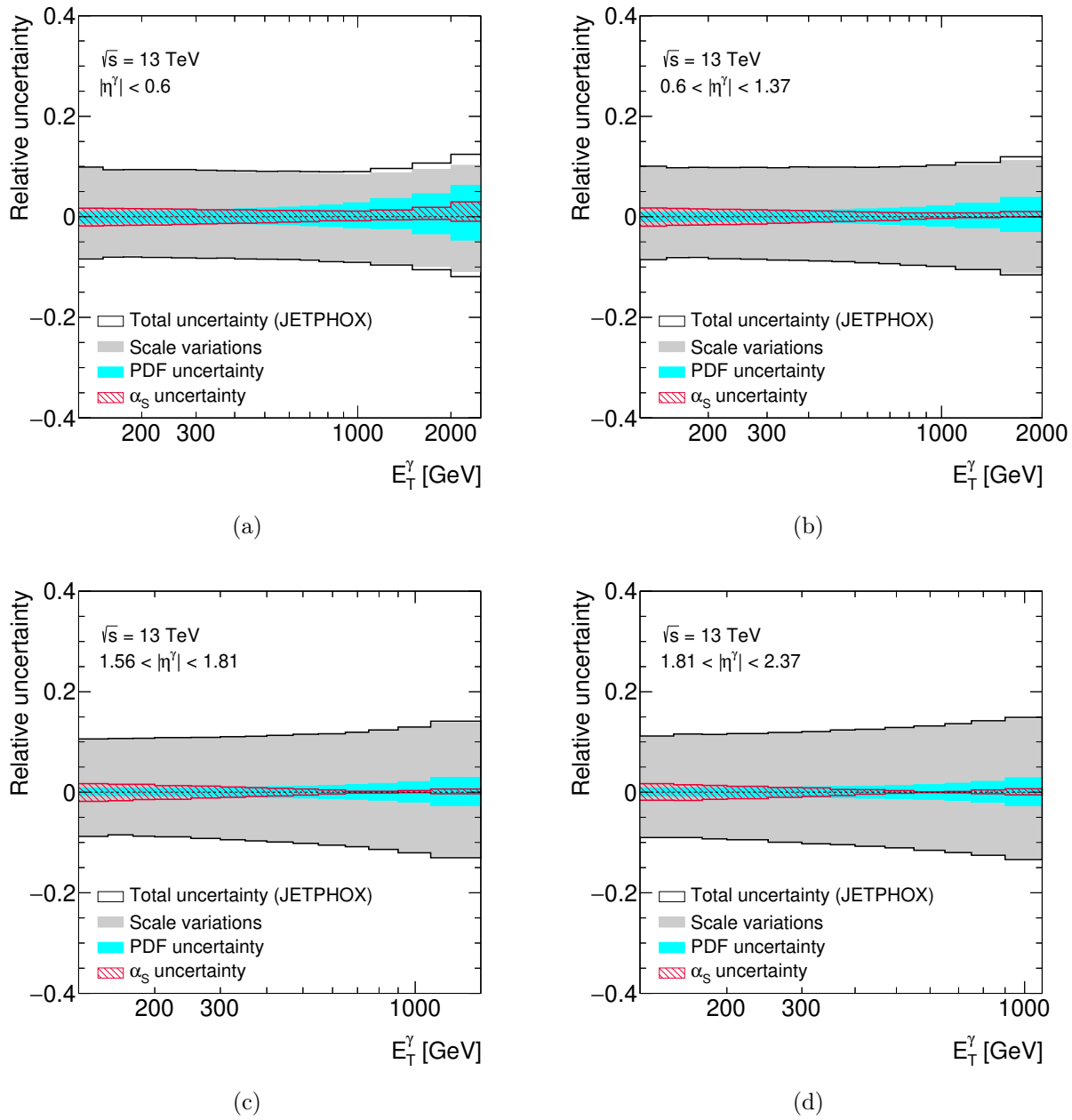


Figure 7.39: The relative total theoretical uncertainty in the cross-section prediction of JETPHOX as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. The uncertainty from the scale variations, the uncertainty from the PDF and the uncertainty from the value of α_{S} are also included.

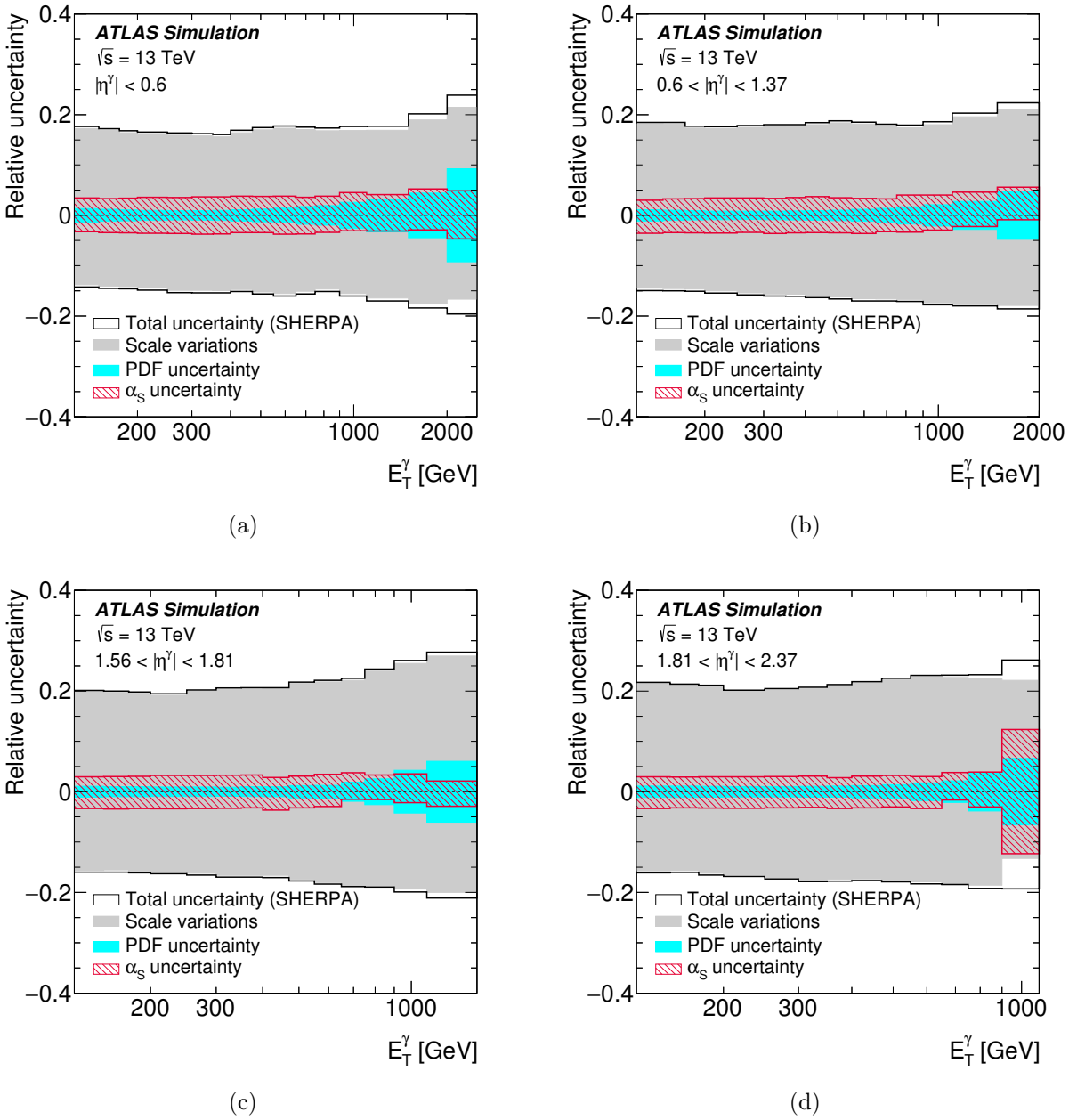


Figure 7.40: The relative total theoretical uncertainty in the cross-section prediction of SHERPA 2.2.2 as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. The uncertainty from the scale variations, the uncertainty from the PDF and the uncertainty from the value of α_{S} are also included.

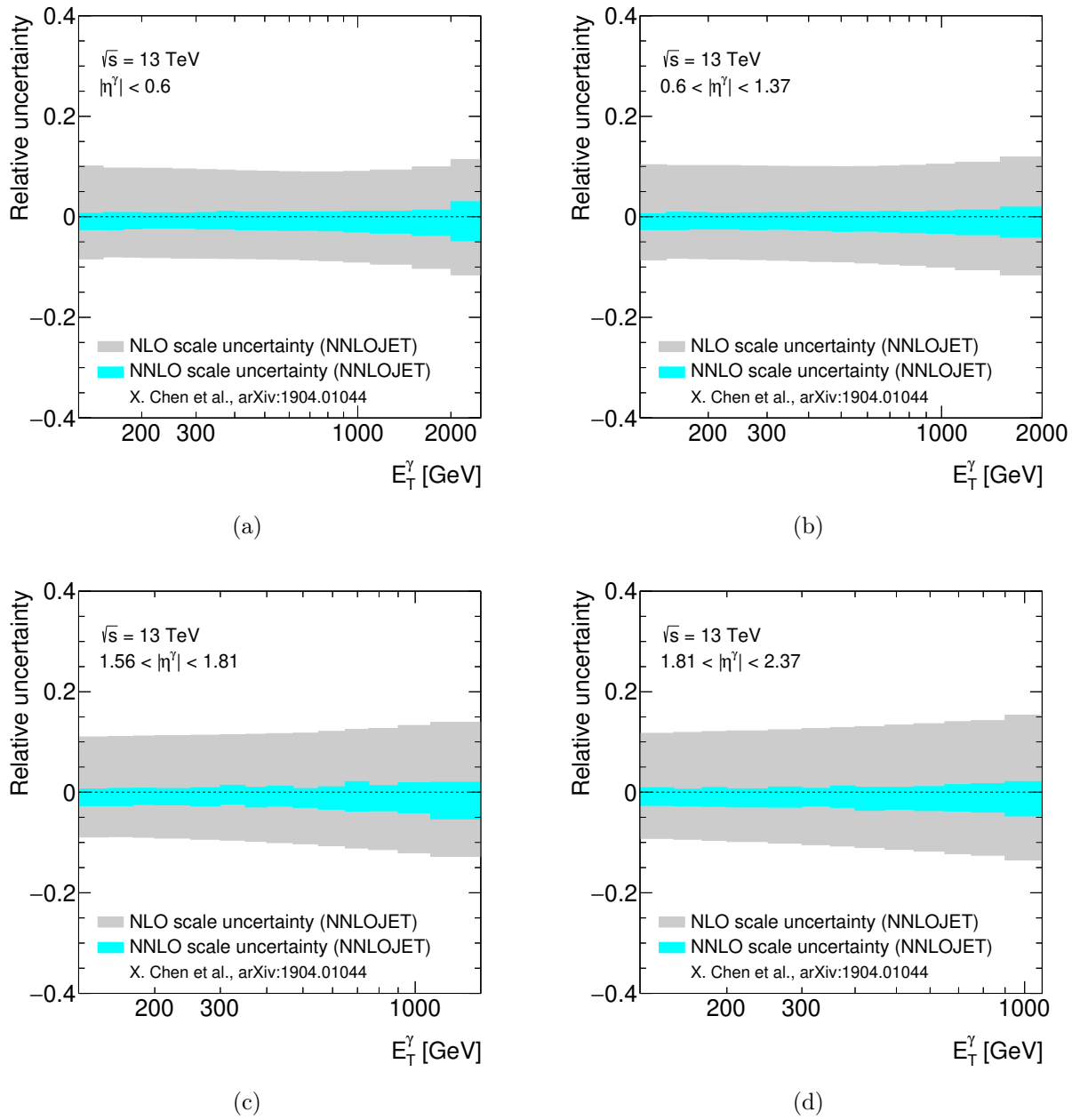


Figure 7.41: The relative theoretical uncertainty in the cross-section prediction of NNLOJET arising from the scale variations as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$: for NLO and NNLO QCD predictions.

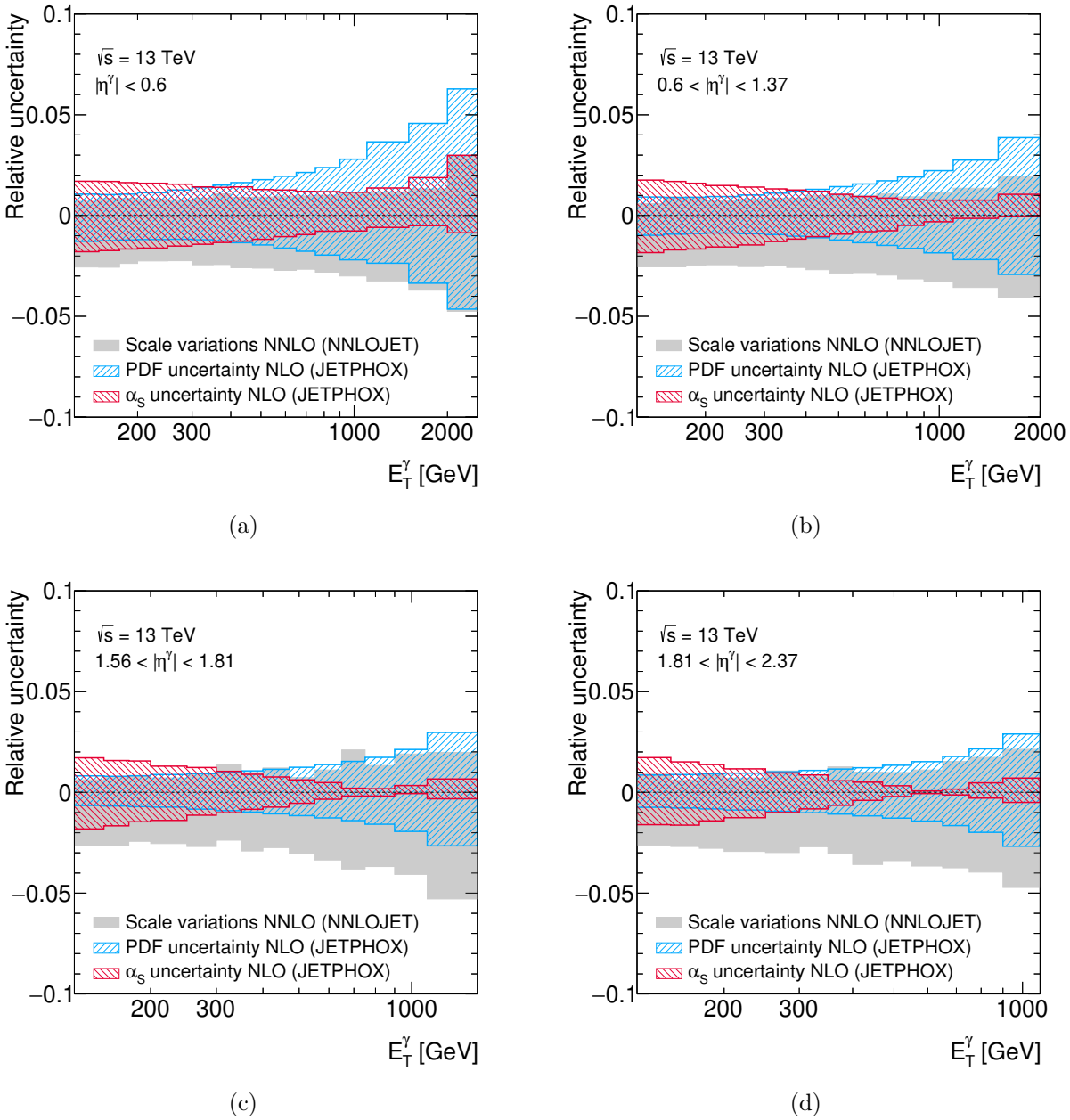


Figure 7.42: The relative theoretical uncertainty in the NNLO pQCD prediction from NNLOJET due to the scale variations (grey areas) and the relative theoretical uncertainty in the NLO QCD prediction from JETPHOX due to the uncertainties in the PDF (blue hatched areas) and α_S (red hatched areas).

7.9 Results

The measurement of the inclusive isolated-photon differential cross section as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$ is shown in figure 7.43. In the region $|\eta^{\gamma}| < 0.6$ ($0.6 < |\eta^{\gamma}| < 1.37$) the cross section decreases by seven orders of magnitude in the measured range $125 < E_{\mathrm{T}}^{\gamma} < 2500$ GeV ($125 < E_{\mathrm{T}}^{\gamma} < 2000$ GeV). In the forward regions $1.56 < |\eta^{\gamma}| < 1.81$ and $1.81 < |\eta^{\gamma}| < 2.37$, the measured ranges are $125 < E_{\mathrm{T}}^{\gamma} < 1500$ GeV and $125 < E_{\mathrm{T}}^{\gamma} < 1100$ GeV, respectively, and the cross section spans approximately six orders of magnitude.

The predictions of the leading-logarithm parton-shower MC models of PYTHIA and SHERPA are compared to the measurement in figure 7.43. These predictions were normalised to the measured integrated cross section in each $|\eta^{\gamma}|$ region. The difference in normalisation between the data and PYTHIA (SHERPA) is $\approx 10\% - 20\%$ ($\approx 40\%$). The calculations of PYTHIA give a good description of the shape of the data for $E_{\mathrm{T}}^{\gamma} \lesssim 500$ GeV for $|\eta^{\gamma}| < 1.37$; the description of PYTHIA is good in the whole measured range for $1.56 < |\eta^{\gamma}| < 2.37$. The predictions of SHERPA also give a good description of the shape of the data in the same phase-space regions.

The NLO pQCD predictions of JETPHOX and SHERPA 2.2.2 are compared with the measurement in figures 7.44 and 7.45. Figure 7.44 shows the inclusive isolated-photon differential cross section as a function E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. Figure 7.45 shows the inclusive isolated-photon double-differential cross section $\mathrm{d}^2\sigma/\mathrm{d}|\eta^{\gamma}|\mathrm{d}E_{\mathrm{T}}^{\gamma}$ as a function of $|\eta^{\gamma}|$ in different regions of E_{T}^{γ}. At high E_{T}^{γ} the decrease of the double-differential cross section with increasing $|\eta^{\gamma}|$ is more prominent. Whereas the measured double-differential cross section decreases by 12% from the most central point to the most forward point for the lowest E_{T}^{γ} range, the decrease is 98% for the highest E_{T}^{γ} range. The nominal prediction of SHERPA 2.2.2 is above that of JETPHOX; this is attributed to the fact that the former includes contributions from parton-showers, virtual corrections for γ+2-jet and higher-order tree-level matrix elements for the processes $2 \to n$ with $n = 4$ and 5, which are not present in the prediction of JETPHOX [37]. Both types of predictions adequately reproduce the dependence of the single and double-differential cross section on E_{T}^{γ} and $|\eta^{\gamma}|$ as observed in the data.

Although the predictions are overall consistent with the measurement within uncertainties, a detailed comparison is presented in terms of the ratios of theory predictions to data. The ratios of the predictions of JETPHOX with $\mu_{\mathrm{R}} = \mu_{\mathrm{F}} = \mu_{\mathrm{f}} = E_{\mathrm{T}}^{\gamma}$ using different PDF sets to the measured cross section as a function of E_{T}^{γ} are shown in figures 7.46 and 7.47; the full theoretical uncertainty, including the variations of the scales, is shown for the prediction based on the MMHT2014 PDF set. The predictions based on the MMHT2014, CT14 and NNPDF3.0 PDF sets are similar and closest to the data for $|\eta^{\gamma}| < 1.37$ for most of the range in E_{T}^{γ}. For $1.56 < |\eta^{\gamma}| < 2.37$, the prediction based on the HERAPDF2.0 PDF set is closest to the data. The prediction based on the ABMP16 PDF set is further away from the data than that based on the MMHT2014 PDF set for $E_{\mathrm{T}}^{\gamma} < 1$ TeV in the region $|\eta^{\gamma}| < 1.37$. The prediction of JETPHOX with

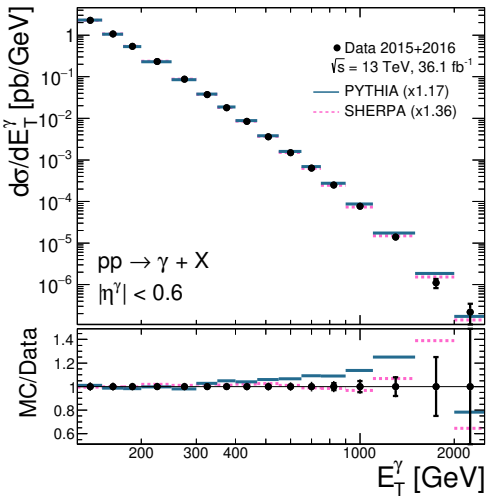
$\mu_R = \mu_F = \mu_f = E_T^\gamma$ exhibits a tendency to underestimate the data over most of the measured range in E_T^γ. However, the prediction of JETPHOX with a different choice for the scales, namely $\mu_R = \mu_F = \mu_f = E_T^\gamma/2$, provides an improved description of the normalisation of the data, as can be seen in figure 7.48; the full theoretical uncertainty, including the variations of the scales with respect to $E_T^\gamma/2$, is shown for the prediction based on the MMHT2014 PDF set. The ratio of the prediction of SHERPA 2.2.2 to the measured cross section as a function of E_T^γ is also shown in figure 7.48. Good agreement is observed between this prediction and the data distribution. In general, the NLO pQCD predictions are consistent with the measurement within the experimental and theoretical uncertainties.

The NNLO pQCD prediction from NNLOJET, which represents a significant step forward in precision, is compared with the data in figures 7.44, 7.45, 7.49 and 7.50. The inclusion of the NNLO QCD corrections increase the normalisation of the prediction and significantly reduces the uncertainty due to the scale variations, which is dominant for NLO pQCD calculations. With the described choice of the input parameters, the NNLO pQCD prediction provides an excellent description of the data except in the region $1.56 < |\eta^\gamma| < 1.81$, where there is a tendency in the theory to underestimate the data. The improvements in the description of the data by the NNLO pQCD prediction relative to that of NLO pQCD are shown in figures 7.49 and 7.50.

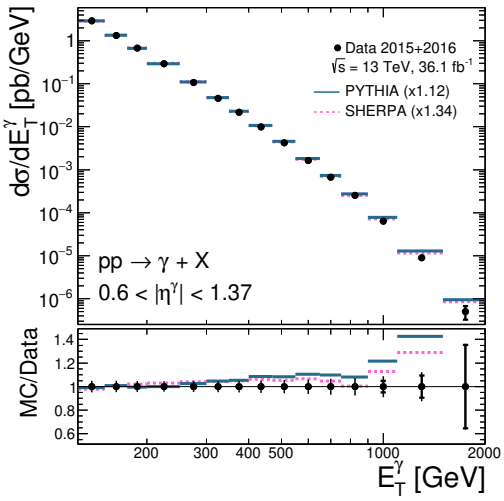
The integrated cross section for isolated-photon production was measured by integrating the differential cross section in the different regions of $|\eta^\gamma|$. The measured integrated cross section in the phase-space region given by $E_T^\gamma > 125$ GeV and $|\eta^\gamma| < 2.37$ (excluding the region of $1.37 < |\eta^\gamma| < 1.56$) and with isolation requirement $E_T^{\text{iso}} < 4.2 \times 10^{-3} \times E_T^\gamma + 4.8$ GeV is

$$\sigma_{\text{meas}} = 406.70^{+13.03}_{-13.01}\ (\text{syst.}) \pm 8.54\ (\text{lumi.})\ \text{pb}\,,$$

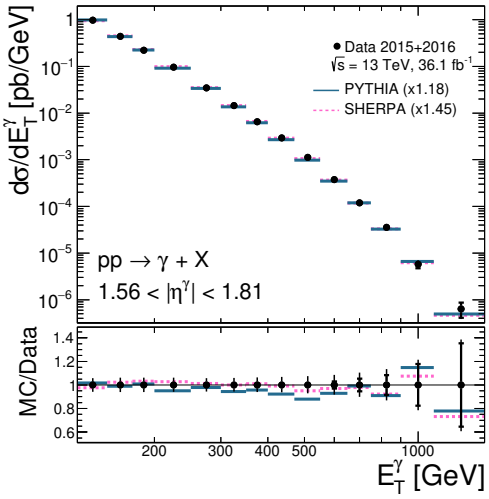
which is consistent with the previous published measurement and has a similar precision ($\sigma_{\text{meas}} = 399^{+13}_{-13}$ (syst.) $\pm$ 8 (lumi.) pb [10]).



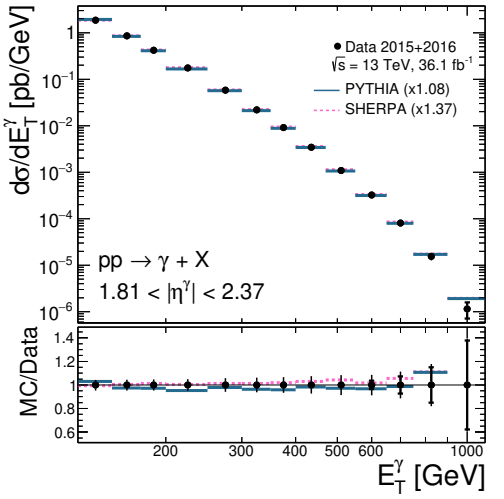
(a)



(b)



(c)



(d)

Figure 7.43: Measured cross section for isolated-photon production (dots) as a function of E_{T}^{γ} in $|\eta^{\gamma}| < 0.6$ (a) $0.6 < |\eta^{\gamma}| < 1.37$ (b) $1.56 < |\eta^{\gamma}| < 1.81$ (c) $1.81 < |\eta^{\gamma}| < 2.37$ (d). The predictions from PYTHIA (solid lines) and SHERPA (dashed lines) are also shown; these predictions are normalised to the measured integrated cross section in each region of $|\eta^{\gamma}|$ using the values indicated in parentheses. The bottom part of each figure shows the ratio of the MC predictions to the measured cross section. The inner (outer) error bars represent the statistical uncertainties (the statistical and systematic uncertainties added in quadrature). For most of the points, the inner error bars are smaller than the marker size and, thus, not visible.

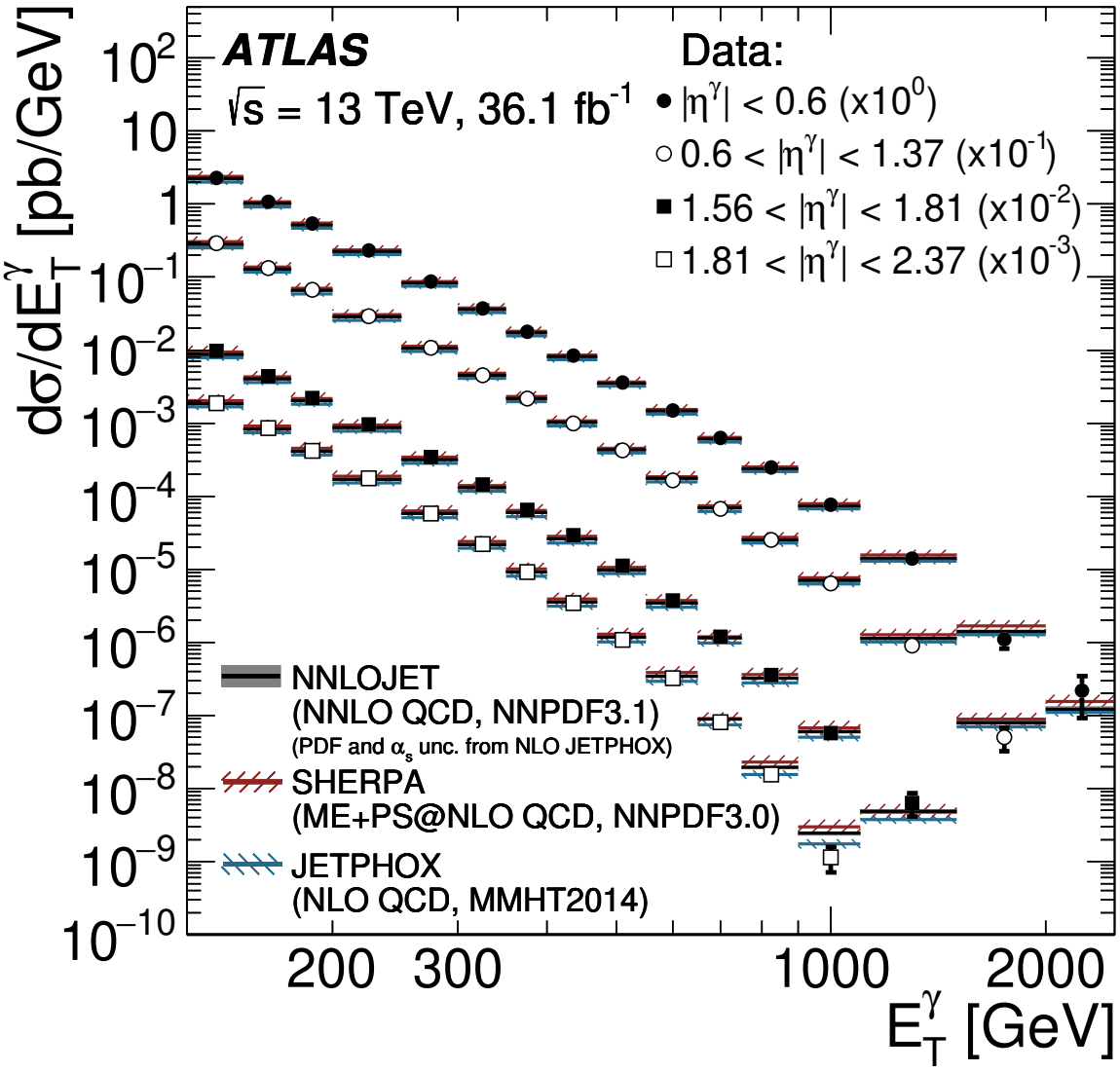


Figure 7.44: The measured differential cross section for isolated-photon production as a function of E_{T}^{γ} in $|\eta^{\gamma}| < 0.6$ (black dots), $0.6 < |\eta^{\gamma}| < 1.37$ (open circles), $1.56 < |\eta^{\gamma}| < 1.81$ (black squares), $1.81 < |\eta^{\gamma}| < 2.37$ (open squares). The NLO pQCD prediction from JETPHOX, the ME+PS@NLO pQCD prediction from SHERPA 2.2.2 and the NNLO pQCD prediction from NNLOJET are also shown. The measurement and the predictions are normalised by the factors shown in parentheses to aid visibility. The error bars represent the data statistical uncertainties and systematic uncertainties added in quadrature. For most of the points, the error bars are smaller than the marker size and, thus, not visible. The bands represent the theoretical uncertainty associated with the predictions; in the case of NNLOJET, the uncertainties due to the PDF and α_{S} are estimated at NLO with JETPHOX.

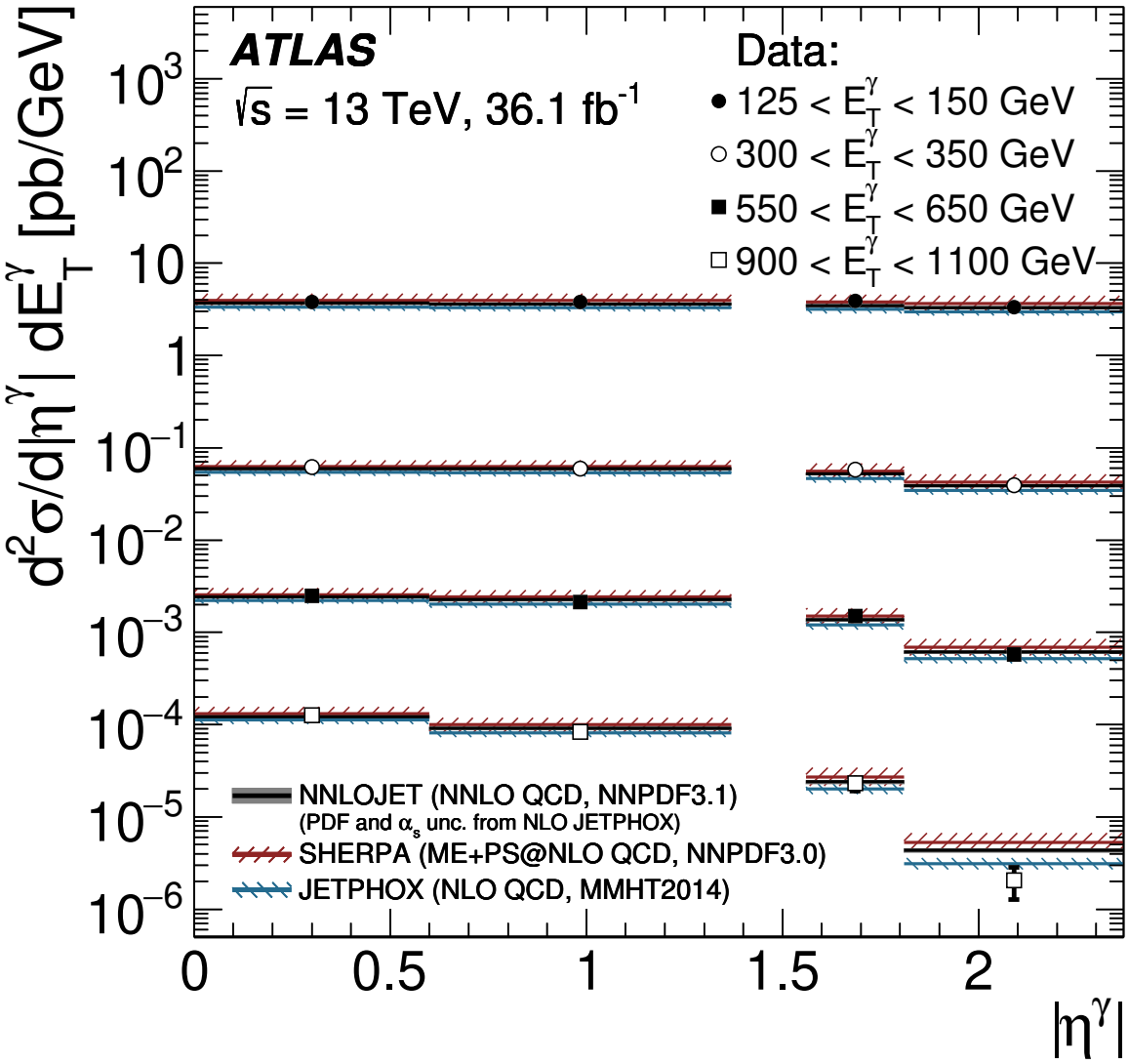


Figure 7.45: The measured double-differential cross section for isolated-photon production as a function of $|\eta^\gamma|$ in $125 < E_{\mathrm{T}}^\gamma < 150$ GeV (black dots), $300 < E_{\mathrm{T}}^\gamma < 350$ GeV (open circles), $550 < E_{\mathrm{T}}^\gamma < 650$ GeV (black squares), and $900 < E_{\mathrm{T}}^\gamma < 1100$ GeV (open squares). The NLO pQCD prediction from JETPHOX, the ME+PS@NLO pQCD prediction from SHERPA 2.2.2 and the NNLO pQCD prediction from NNLOJET are also shown. The error bars represent the data statistical uncertainties and systematic uncertainties added in quadrature. For most of the points, the error bars are smaller than the marker size and, thus, not visible. The bands represent the theoretical uncertainty associated with the predictions; in the case of NNLOJET, the uncertainties due to the PDF and α_{S} are estimated at NLO with JETPHOX.

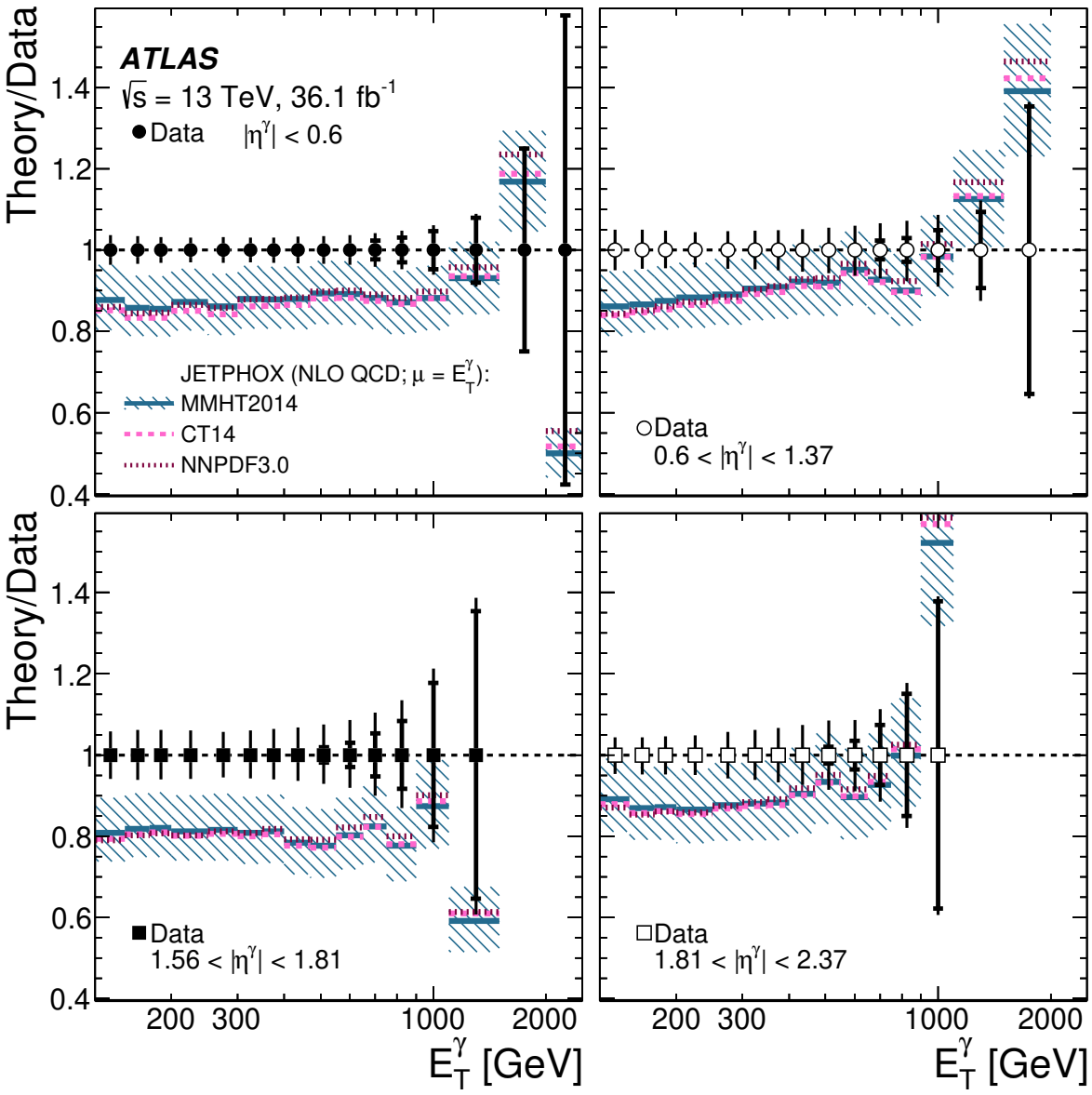


Figure 7.46: The ratio of the NLO pQCD prediction of JETPHOX with $\mu_{\mathrm{R}} = \mu_{\mathrm{F}} = \mu_{\mathrm{f}} = E_{\mathrm{T}}^{\gamma}$ using the MMHT2014 PDF set to the measured differential cross section for isolated-photon production (solid lines) as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. The symbols for the data are centred at unity and the inner (outer) error bars represent the relative data statistical uncertainties (statistical and systematic uncertainties added in quadrature). The hatched bands represent the full theoretical uncertainty associated to the prediction based on the MMHT2014 PDF set, including the variations of the scales. For comparison, the predictions using the CT14 (dashed lines) and NNPDF3.0 (dotted lines) PDF sets are also included.

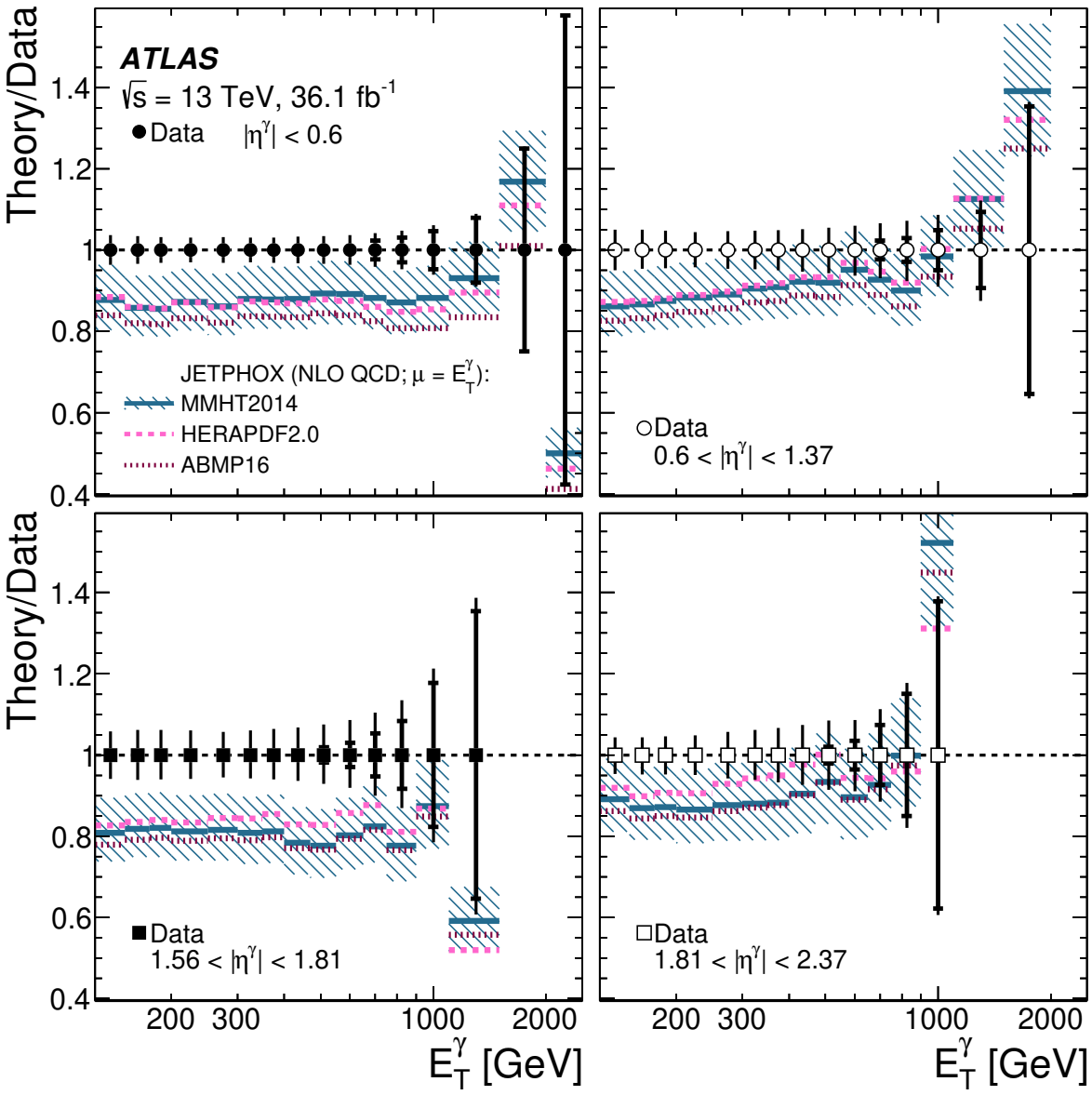


Figure 7.47: The ratio of the NLO pQCD prediction of JETPHOX with $\mu_{\mathrm{R}} = \mu_{\mathrm{F}} = \mu_{\mathrm{f}} = E_{\mathrm{T}}^{\gamma}$ using the MMHT2014 PDF set to the measured differential cross section for isolated-photon production (solid lines) as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. The symbols for the data are centred at unity and the inner (outer) error bars represent the relative data statistical uncertainties (statistical and systematic uncertainties added in quadrature). The hatched bands represent the full theoretical uncertainty associated to the prediction based on the MMHT2014 PDF set, including the variations of the scales. For comparison, the predictions using the HERAPDF2.0 (dashed lines) and ABMP16 (dotted lines) PDF sets are also included.

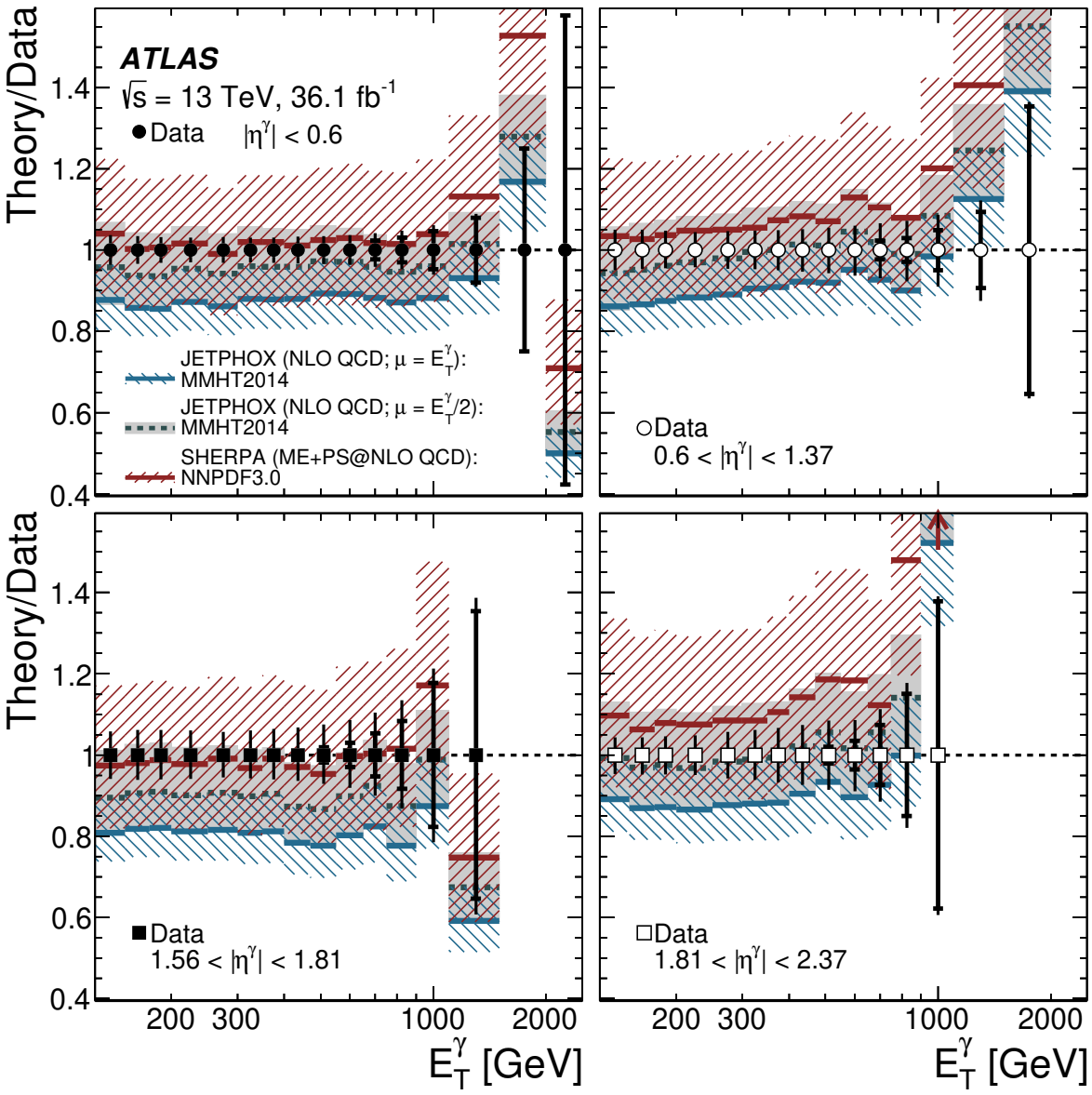


Figure 7.48: The ratio of the NLO pQCD prediction of JETPHOX with $\mu_{\mathrm{R}} = \mu_{\mathrm{F}} = \mu_{\mathrm{f}} = E_{\mathrm{T}}^{\gamma}$, the ratio of the NLO pQCD prediction of JETPHOX with $\mu_{\mathrm{R}} = \mu_{\mathrm{F}} = \mu_{\mathrm{f}} = E_{\mathrm{T}}^{\gamma}/2$, and the ratio of the ME+PS@NLO pQCD prediction of SHERPA 2.2.2 to the measured differential cross section for isolated-photon production as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. The symbols for the data are centred at unity and the inner (outer) error bars represent the relative data statistical uncertainties (statistical and systematic uncertainties added in quadrature). The bands represent the full theoretical uncertainty, which includes the variations of the scales.

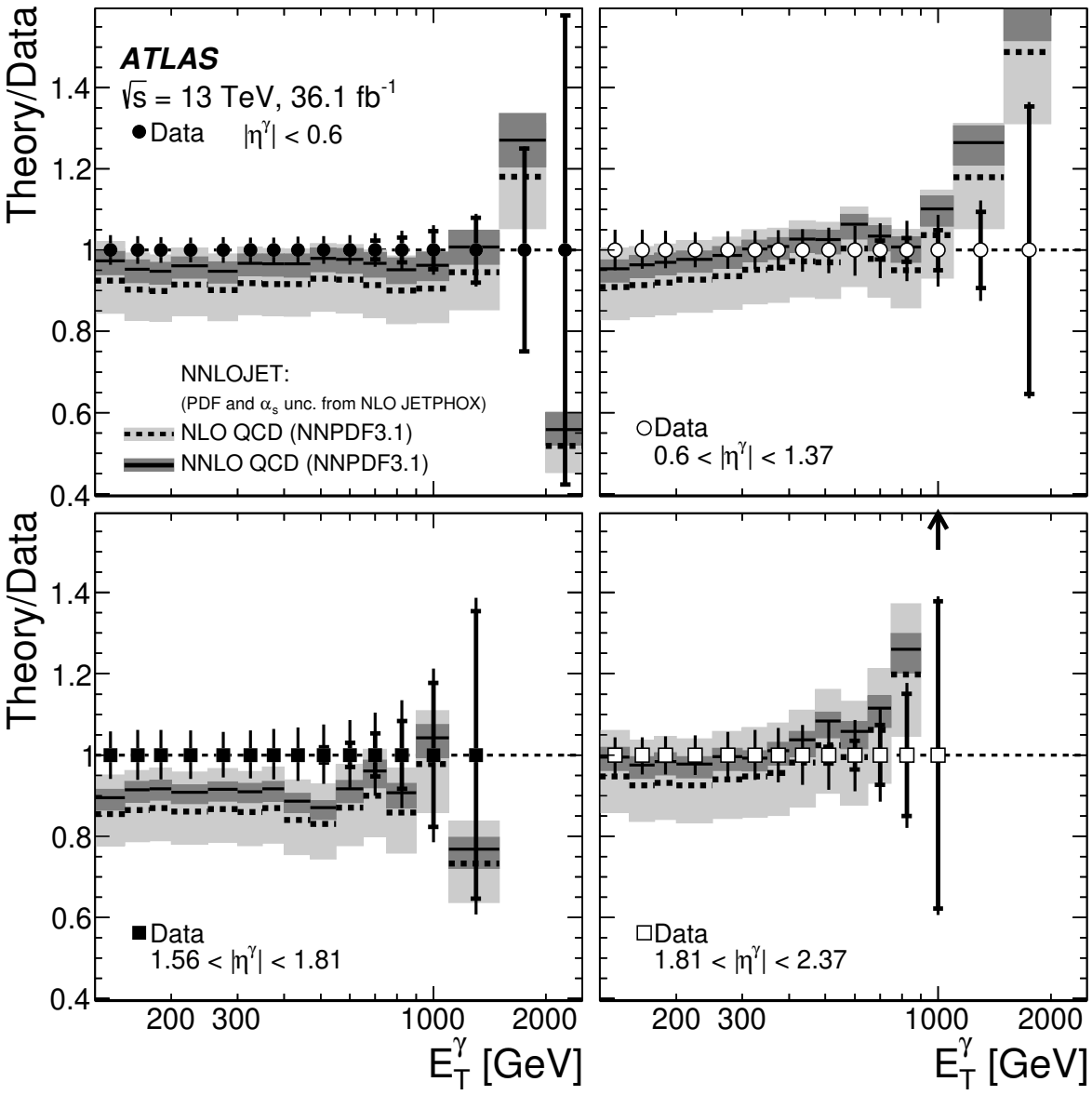


Figure 7.49: The ratio of the NNLO (NLO) pQCD prediction of NNLOJET using the NNPDF3.1 PDF set to the measured differential cross section for isolated-photon production as a function of E_{T}^{γ} in different regions of $|\eta^{\gamma}|$ is shown as a solid (dashed) line. The symbols for the data are centred at unity and the inner (outer) error bars represent the relative data statistical uncertainties (statistical and systematic uncertainties added in quadrature). The shaded bands represent the theoretical uncertainties, which include those due to the PDF and α_{S} as estimated at NLO with JETPHOX.

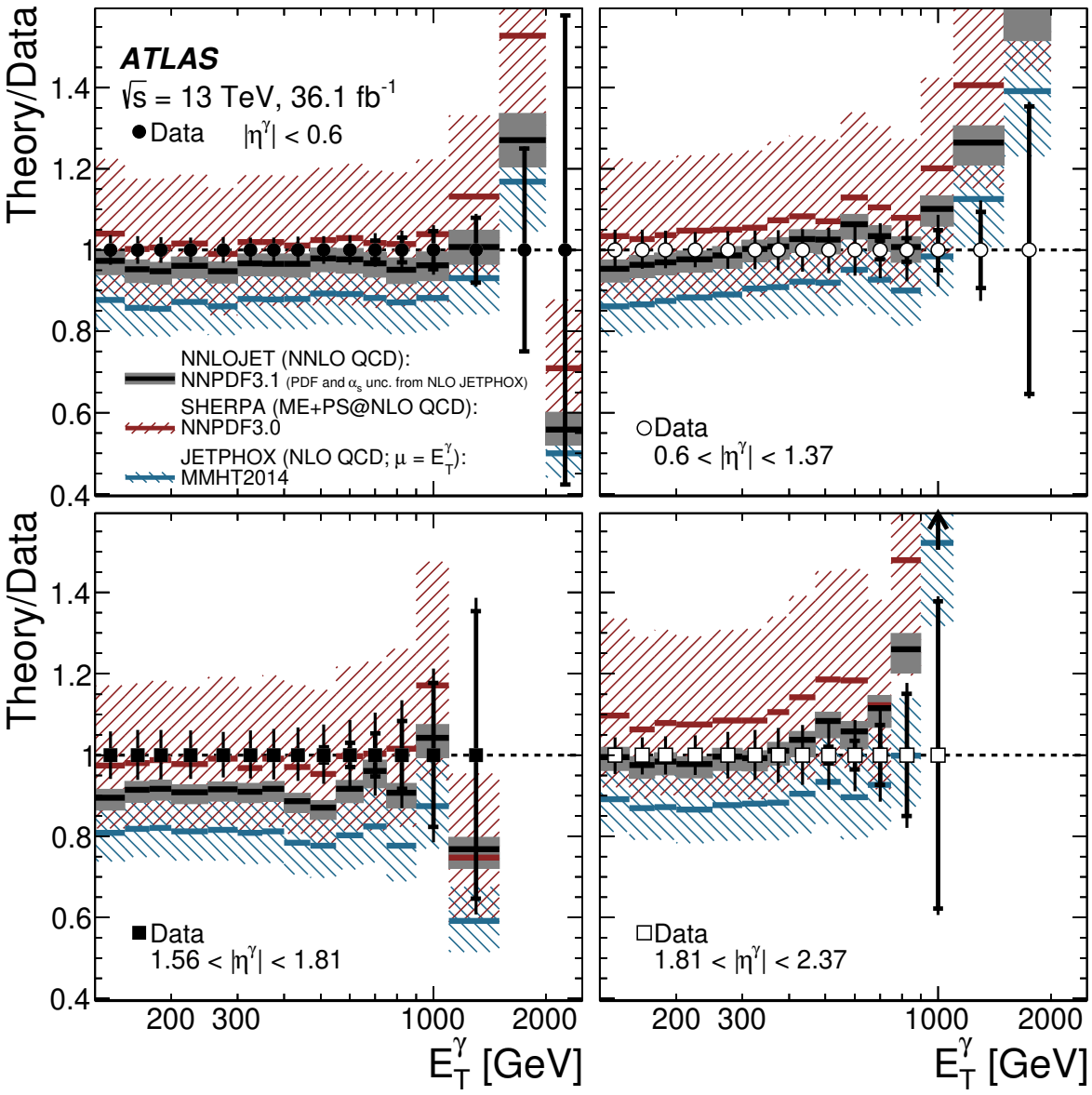


Figure 7.50: The ratio of the NLO pQCD prediction of JETPHOX with $\mu_{\mathrm{R}} = \mu_{\mathrm{F}} = \mu_{\mathrm{f}} = E_{\mathrm{T}}^{\gamma}$, the ratio of the ME+PS@NLO pQCD prediction of SHERPA 2.2.2 and the ratio of the NNLO pQCD prediction of NNLOJET to the measured differential cross section for isolated-photon production as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. The symbols for the data are centred at unity and the inner (outer) error bars represent the relative data statistical uncertainties (statistical and systematic uncertainties added in quadrature). The bands represent the theoretical uncertainty; in the case of NNLOJET, the uncertainties due to the PDF and α_{S} are estimated at NLO with JETPHOX.

Chapter 8

Measurement of the cross section for inclusive isolated photon production in pp collisions at $\sqrt{s} = 13$ TeV using $139\,\text{fb}^{-1}$ of ATLAS data

In this chapter, the measurement of the isolated-photon differential cross section in pp collisions at $\sqrt{s} = 13$ TeV using $139\,\text{fb}^{-1}$ of ATLAS data is presented. This measurement is currently under review in the ATLAS collaboration. The differential cross section is measured as a function of the photon transverse energy in different regions of the photon pseudorapidity. In addition, the dependence of the production cross section on the photon isolation cone size is measured. NLO pQCD calculations from JETPHOX and SHERPA are compared with the measurement, using several parameterisations of the proton parton distribution functions. The predictions provide a good description of the data within the experimental and theoretical uncertainties. This measurement constitutes an improvement with respect to those published earlier, measuring the cross section for different isolation-cone radii and in more $|\eta^{\gamma}|$ regions. The new measurement benefits from reduced systematic uncertainties.

8.1 Data selection

The data used in this analysis were collected with the ATLAS detector during the proton-proton collision running periods 2015 to 2018, when the LHC operated at a centre-of-mass energy of $\sqrt{s} = 13$ TeV. Only events taken during stable beam conditions and satisfying detector- and data-quality requirements, which include the calorimeters and inner tracking detectors being in nominal operation, are considered. In the collected data sample, the average number of pp interactions per bunch crossing is ~ 34, while the total integrated luminosity is $139.0 \pm 2.4\,\text{fb}^{-1}$. The uncertainty in the combined $2015 - 2018$

luminosity is 1.7% [150], obtained using the LUCID-2 detector [151] for the primary luminosity measurements.

Trigger requirements

The data sample used consists of events triggered by a single-photon high-level trigger with a nominal transverse energy threshold of 140 GeV for the 2015 − 2018 data-taking period and loose photon identification requirements[1] [39, 131]. The single-photon trigger with the threshold of 140 GeV was the lowest-threshold unprescaled photon trigger during the 2016 − 2018 data-taking period. For the 2015 period, the lowest-threshold unprescaled trigger had a threshold of 120 GeV. The selection criteria applied by the trigger on the shower-shape variables computed from the energy profiles of the showers in the calorimeters are looser than the photon identification criteria applied in the offline analysis and allow a plateau of constant efficiency close to 100%, for true prompt photons of $E_{\mathrm{T}}^{\gamma} > 250$ GeV and $|\eta^{\gamma}| < 2.37$, as shown at the end of this section.

Event selection

The sample of isolated-photon events was selected offline using similar criteria to those summarised in section 7.1. The requirements on the reconstructed primary vertex were changed during Run II [157]. First, the efficiency for pile-up vertices was slightly improved by eliminating an occasional failure mode in which tracks from a secondary vertex could generate a spurious primary vertex. Second, the p_{T} threshold for storing reconstructed tracks was raised from 400 MeV to 500 MeV to reduce the average event size on disk. For consistency and reproducibility, the p_{T} cut used in vertex reconstruction was identically tightened.

Photon selection

Photon candidates were reconstructed (see section 5.2) and calibrated (see section 5.3) using updated methods with respect to those used in the previous inclusive photon measurement [37], based on ref. [129] and [130], respectively. The main changes in the reconstruction come from the clustering procedure while from the calibration side, similar procedures to those applied in ref. [130] but updated for the new reconstruction were used. The main background in the prompt-photon production measurement comes from high-energy π^0 or η mesons which can be misidentified as a photon because they decay into an almost collinear photon pair. Such energetic π^0 or η mesons are produced copiously inside jets. This background is reduced by the photon identification criteria and by requiring the photon candidate to be isolated.

The photon-candidate selection criteria proceeds in a similar way as described in section 7.1, with the following modifications:

– Events in which the leading photon had transverse energy $E_{\mathrm{T}}^{\gamma} > 250$ GeV and $|\eta^{\gamma}| < 2.37$ were selected. Photons in the region $1.37 < |\eta^{\gamma}| < 1.56$ were rejected.

[1] The discriminating variables used for loose and tight photon identification are described in table 5.1.

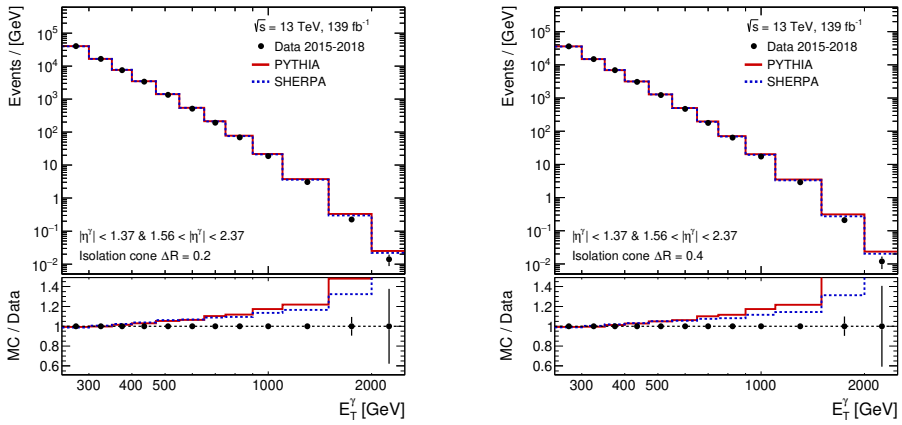
– The isolation transverse energy, $E_{\mathrm{T}}^{\mathrm{iso}}$, of the leading photon, computed from topo-clusters and corrected for photon energy leakage and ambient transverse energy (as described in section 5.5), was required to be lower than $4.2 \times 10^{-3} \times E_{\mathrm{T}}^{\gamma} + 4.8$ GeV, which amounts to a cut on $E_{\mathrm{T}}^{\mathrm{iso}}$ of 5.9 (15.3) GeV for $E_{\mathrm{T}}^{\gamma} = 250$ (2500) GeV. Two different samples were selected, one with isolation cone radius $R = 0.4$ and another with $R = 0.2$ (called $R = 0.4$ and $R = 0.2$ samples, respectively).

The total number of selected events with a photon candidate with $E_{\mathrm{T}}^{\gamma} > 250$ GeV is 3 652 433 (3 289 941) for the $R = 0.2$ ($R = 0.4$) sample. Figure 8.1 shows the E_{T}^{γ}, $|\eta^{\gamma}|$ and ϕ^{γ} data and MC distributions. The kinematic requirements and the number of selected events in data in each $|\eta^{\gamma}|$ region are summarised in table 8.1. Each sample was separated in six regions in $|\eta^{\gamma}|$ to perform the measurement of the differential cross section individually in each region. Each region in $|\eta^{\gamma}|$ is divided into 12 bins of E_{T}^{γ} starting at $E_{\mathrm{T}}^{\gamma} = 250$ GeV and ending at $E_{\mathrm{T}}^{\gamma} = 2500$ GeV. The binning is given by the following array of values of E_{T}^{γ} (in units of GeV): 250, 300, 350, 400, 470, 550, 650, 750, 900, 1100, 1500, 2000, and 2500. Some of the high-E_{T}^{γ} bins are not measured depending on the $|\eta^{\gamma}|$ region.

	Phase space region	
Requirement on E_{T}^{γ}	$E_{\mathrm{T}}^{\gamma} > 250$ GeV	
Isolation requirement	$E_{\mathrm{T}}^{\mathrm{iso}} < 4.2 \times 10^{-3} \times E_{\mathrm{T}}^{\gamma} + 4.8$ GeV	
Sample	$R = 0.2$	$R = 0.4$
Region $\lvert\eta^{\gamma}\rvert < 0.6$	1 128 799	1 015 083
Region $0.6 < \lvert\eta^{\gamma}\rvert < 0.8$	382 153	346 514
Region $0.8 < \lvert\eta^{\gamma}\rvert < 1.37$	1 073 194	991 171
Region $1.56 < \lvert\eta^{\gamma}\rvert < 1.81$	439 290	382 732
Region $1.81 < \lvert\eta^{\gamma}\rvert < 2.01$	257 478	226 895
Region $2.01 < \lvert\eta^{\gamma}\rvert < 2.37$	371 519	327 546

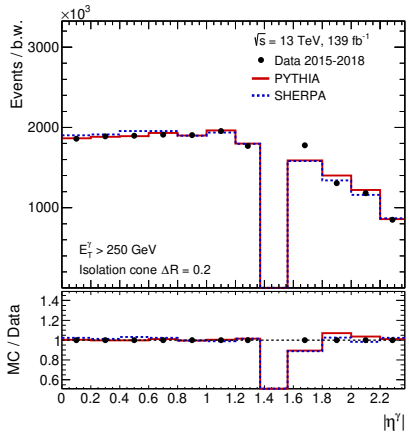
Table 8.1: Definition of the phase-space region for the measurement and predictions. The number of data events selected in each $|\eta^{\gamma}|$ region is shown for the two samples selected according to the value of the isolation cone radius parameter.

The average number of interactions per bunch crossing for the different data-taking periods is 23.5, 37.6 and 36.7 for the data collected with the HLT_g140_loose trigger in the 2015 − 2016, 2017, and 2018 years, respectively. The differences in the pile-up conditions between data and MC simulation were accounted for separately for each of the periods before their combination.

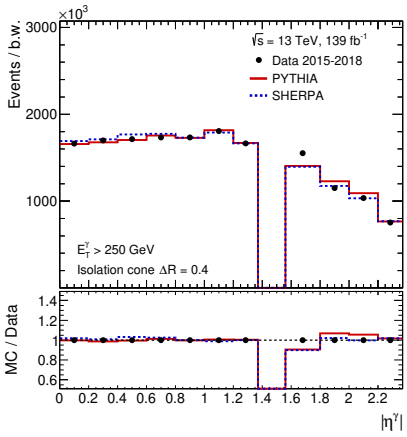


(a)

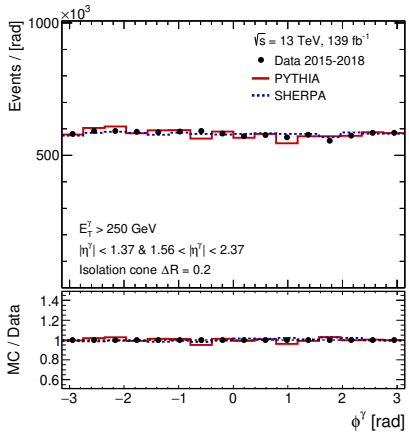
(b)



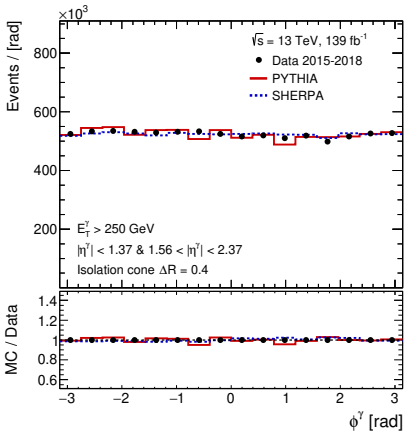
(c)



(d)



(e)



(f)

Figure 8.1: The measured E_{T}^{γ} (a,b), $|\eta^{\gamma}|$ (c,d) and ϕ^{γ} (e,f) distributions for events with $E_{\mathrm{T}}^{\gamma} > 250$ GeV (dots). The distributions for the sample with $R = 0.2$ are shown in (a,c,e), while those corresponding to a $R = 0.4$ are shown in (b,d,f). For comparison, the MC simulations of the signal from PYTHIA (solid histograms) and SHERPA (dashed histograms) are also included. The MC simulations are normalised to the number of data events. The lower panel of each figure shows the ratio of the MC distributions to the data.

Trigger efficiency, uncertainty and correction

The efficiency of the HLT_g140_loose trigger used in this analysis for photons with $E_{\mathrm{T}}^{\gamma} > 250$ GeV was evaluated using a lower-threshold trigger, the HLT_g120_loose, which was unprescaled (prescaled) during the 2015 (2016 − 2018) running period. The efficiency was estimated as

$$\epsilon_{\mathrm{trigger}} = \frac{N_{\mathrm{HLT_g120_loose}\otimes\mathrm{HLT_g140_loose}}}{N_{\mathrm{HLT_g120_loose}}},$$

where $N_{\mathrm{HLT_g120_loose}}$ is the number of data events that pass all the selection criteria listed previously in this section and fulfil the conditions of the trigger HLT_g120_loose; $N_{\mathrm{HLT_g120_loose}\otimes\mathrm{HLT_g140_loose}}$ is the number of data events that pass all the selection criteria listed previously and fulfil the conditions of both triggers HLT_g120_loose and HLT_g140_loose. Figure 8.2 shows the estimated trigger efficiencies as a function of E_{T}^{γ} in the different $|\eta^{\gamma}|$ regions for tight isolated photons. The trigger efficiencies for tight non-isolated, non-tight isolated, and non-tight non-isolated photons are shown in appendix F (these categories were defined in section 7.3).

The trigger efficiency was found to be very close to unity for all the $|\eta^{\gamma}|$ regions along the full E_{T}^{γ} spectrum and in the different tight isolated, tight non-isolated, non-tight isolated, and non-tight non-isolated regions. In addition, the trigger efficiency was estimated for each data-taking period separately, and no significant differences were observed. Therefore, no correction was applied to the data.

Following a similar strategy to that described in section 7.1, the efficiency of the HLT_g140_loose trigger was evaluated with respect to a L1 trigger, as explained in ref. [158]. The efficiency of the HLT_g140_loose trigger for data is shown in figure 8.3 separately for each of the 2015, 2016, 2017 and 2018 data-taking periods. Given the smallness of the L1 trigger efficiency, no correction was applied to the data. An uncertainty was considered on the measured cross sections which was estimated as the sum in quadrature of the trigger inefficiency and its uncertainty. The relative uncertainty, as shown in table 8.2 for two regions of E_{T}^{γ}, was propagated into the cross section for each running period separately.

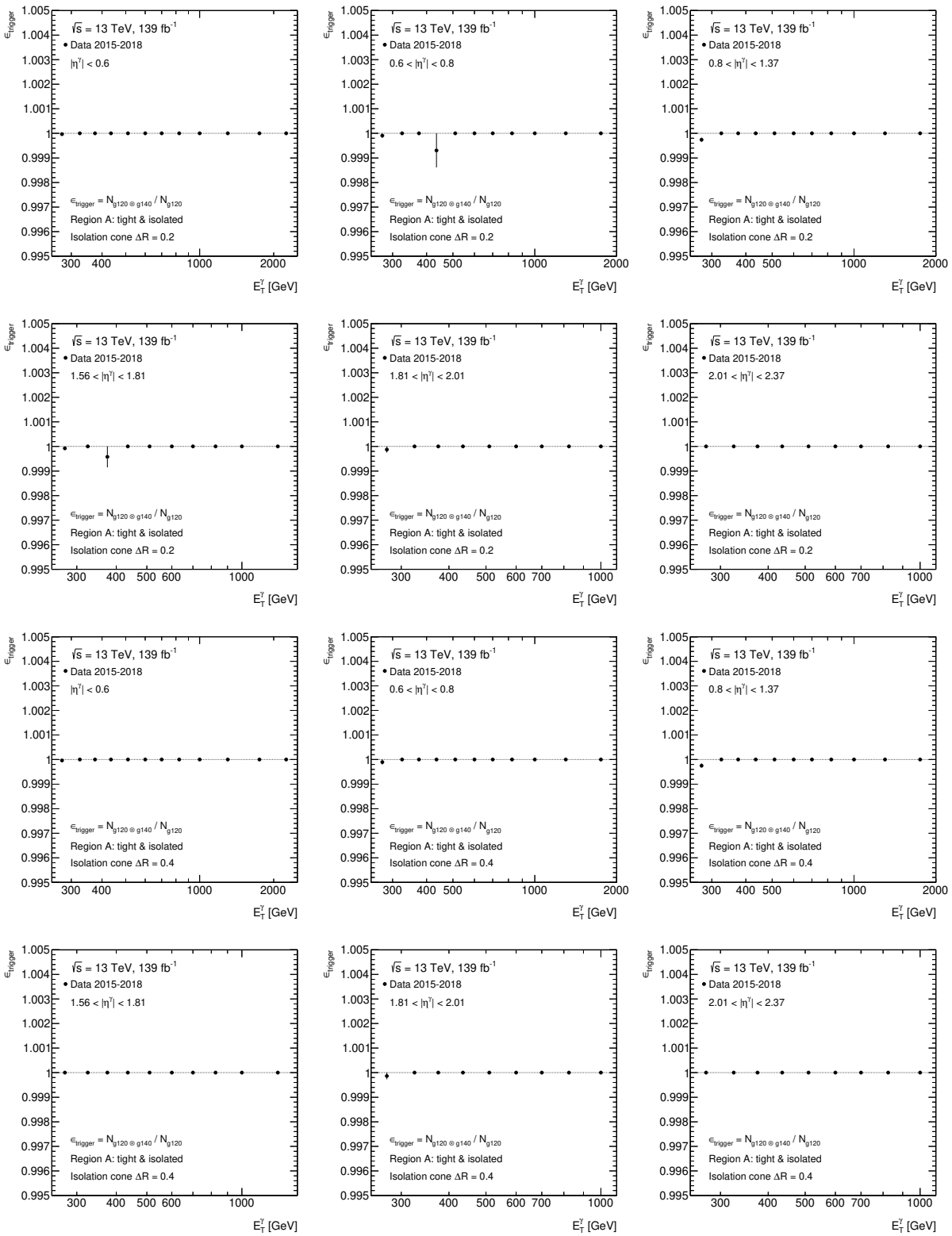


Figure 8.2: Trigger efficiency in data as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$ for tight isolated photons in the $R = 0.2$ (top) and $R = 0.4$ (bottom) samples.

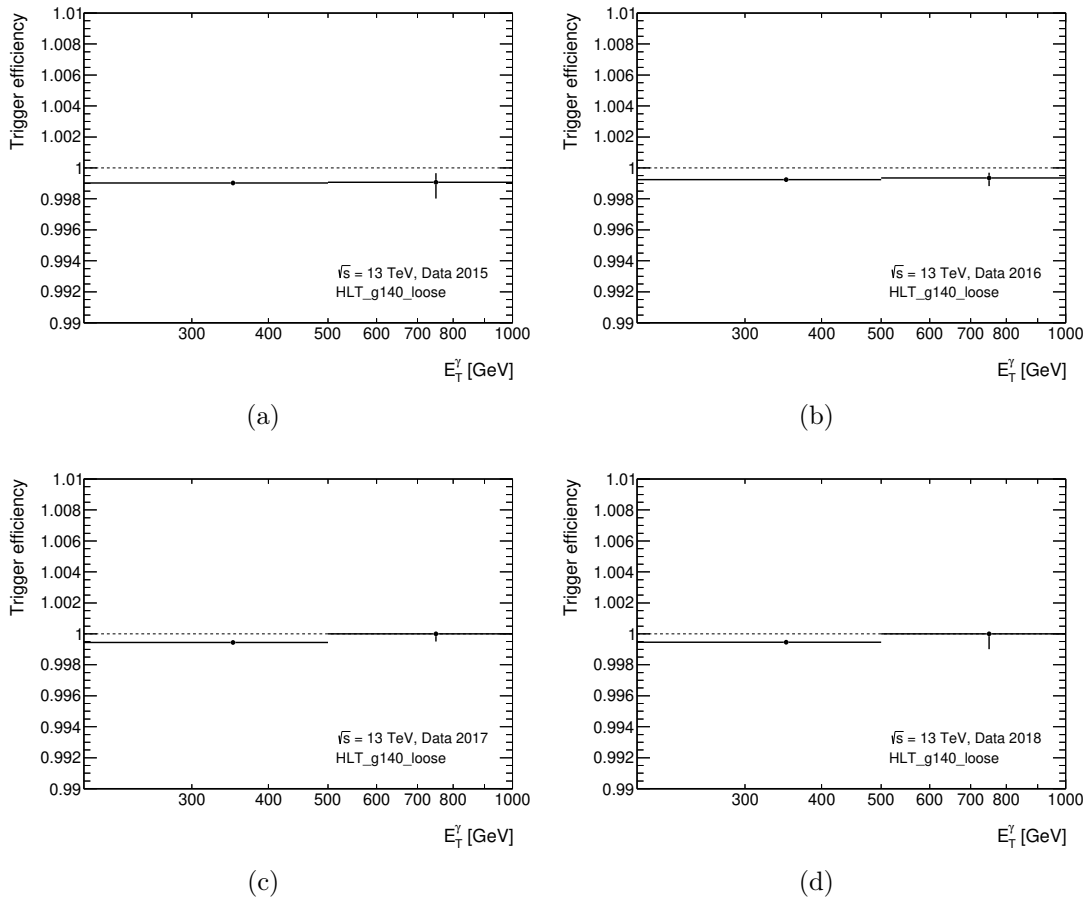


Figure 8.3: Efficiency of the HLT_g140_loose trigger as a function of E_{T}^{γ} for 2015 (a), 2016 (b), 2017 (c), and 2018 (d) data.

	Relative uncertainty (%)			
E_{T}^{γ} region (GeV)	2015	2016	2017	2018
$250 - 470$	0.10	0.08	0.06	0.06
$470 - 2500$	0.15	0.09	0.05	0.10

Table 8.2: Relative uncertainty in % for each running period on the trigger efficiency of the HLT_g140_loose with respect to a L1 trigger.

8.2 Monte Carlo simulations

Samples of MC events were generated to study the characteristics of the signal events, to perform the unfolding and to compute the predicted cross sections at particle level. In addition, MC samples were also used to estimate hadronisation and UE corrections for

the NLO pQCD calculations. The different samples used for this analysis are described in detail in section 4.3.

As indicated in section 4.3, three campaigns of MC simulation were produced to match the different data-taking periods along the Run II. As an example, the MC campaign *a* (*mca* sample) was designed to match the 2015 and 2016 data conditions, $\langle\mu\rangle$-profile, trigger menu and data conditions as they occurred during 2015 and 2016. For the 2017 and 2018 data-taking periods the MC campaigns *d* (*mcd* sample) and *e* (*mce* sample) were produced, respectively.

As described in section 5.6, the MC events were reweighted such that the distribution of $\langle\mu\rangle$ matches the one in the data. Figure 8.4 shows the distribution of $\langle\mu\rangle$, in each running period separately, for the data after applying a scale factor of 1/1.03 (see section 5.6), and for PYTHIA and SHERPA simulations (before and after the MC is reweighted). As expected, the reweighting factors for both MC simulations are very similar. Figure 8.5 shows the data and MC distributions for the full Run II period, after reweighting the MC simulations.

In addition to the corrections described in section 5.6, a correction to the photon reconstruction efficiency was derived using radiative Z events. The correction was derived independently for the 2015 and 2016, 2017, and 2018 data-taking periods. Scale factors were applied to account for the difference between data and MC simulation. The derived scale factors are shown in figure 8.6 for different $|\eta^{\gamma}|$ regions.

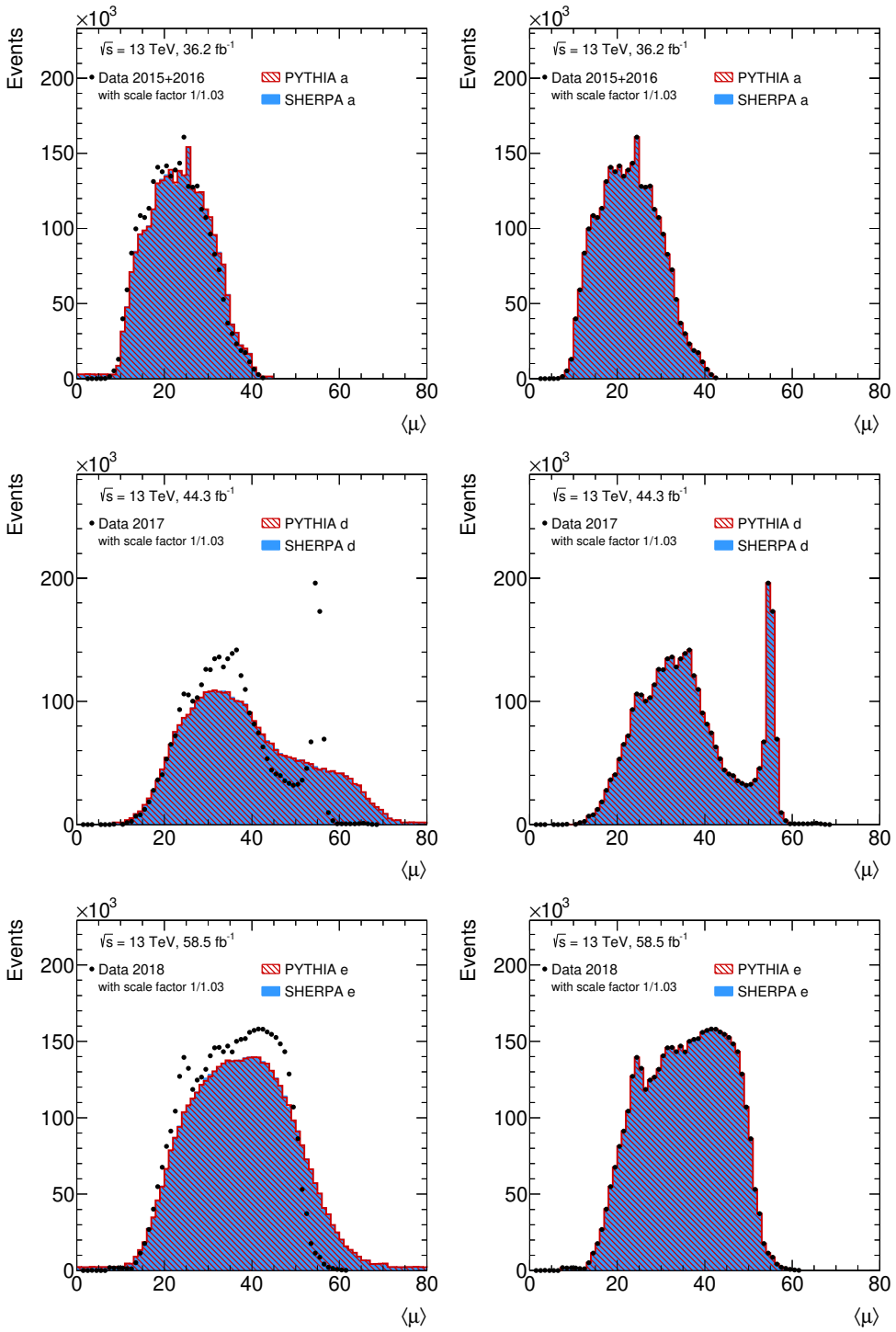


Figure 8.4: Distribution of the average number of interactions per bunch crossing in data (dots) and PYTHIA and SHERPA MC (histograms) before (left) and after (right) reweighting for each running period separately.

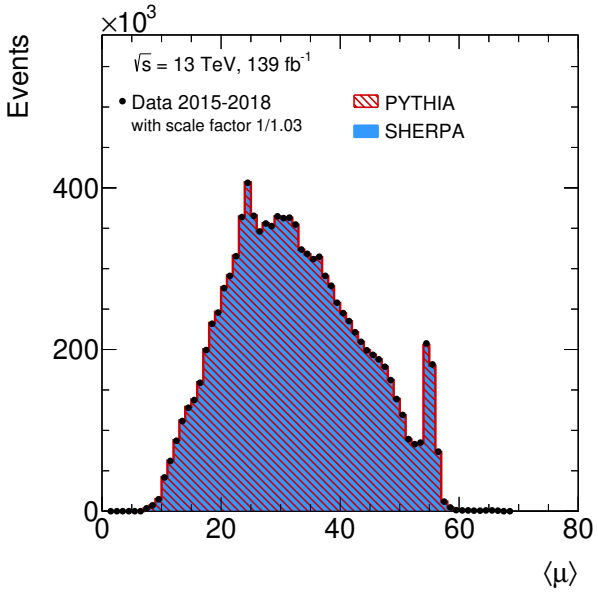


Figure 8.5: Distribution of the average number of interactions per bunch crossing in data (dots) and PYTHIA and SHERPA MC (histograms) after reweighting for the full Run II period (the MC events were reweighted such that the distribution of $\langle\mu\rangle$ matches the one in the data for each running period separately).

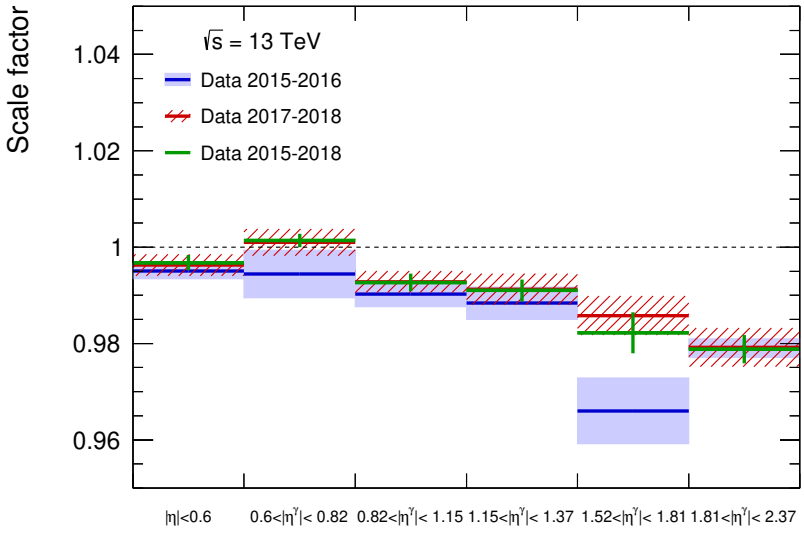


Figure 8.6: Reconstruction-efficiency scale factors in different $|\eta^{\gamma}|$ regions for each running period. The total uncertainty of the scale factors is represented with the error bars (bands) for the 2015 − 2018 (2015 − 2016 and 2017 − 2018) estimations.

8.3 Background evaluation and signal extraction

As discussed in section 7.3, there is a non-negligible amount of background remaining in the selected data sample even after the application of the tight identification and isolation requirements described in sections 5.4 and 5.5, respectively. The background from multijet processes in which a jet is misidentified as a photon is estimated via the same data-driven method applied in the previous inclusive photon measurements [9,10,37] and described in detail in section 7.3.

The isolation profile of a signal-enriched sample was extracted from the data. The sample used for this study was obtained after the application of the selection criteria described in section 8.1, except for the tight identification and the isolation requirements. Following the same methodology which is described in section 7.3, for the two samples using different isolation cone radius parameter, two subsamples were selected by applying the tight identification criteria: the tight subsample and the non-tight subsample, enriched in background candidates. The $E_{\mathrm{T}}^{\mathrm{iso}}$ distributions for these subsamples are shown in figure 8.7. The $E_{\mathrm{T}}^{\mathrm{iso}}$ distribution of the non-tight candidates was scaled so that the integral for $E_{\mathrm{T}}^{\mathrm{iso}} > 10$ GeV, where the contribution from the signal is expected to be small, matched that of the tight candidates. A good description of the data is obtained, showing similar results either using PYTHIA or SHERPA simulations to compare with the data. Differences between the results obtained for the sample with $R = 0.2$ and $R = 0.4$ are observed: the signal peak around zero is wider and the tail at high values of $E_{\mathrm{T}}^{\mathrm{iso}}$ is more populated for the $R = 0.4$ sample. In addition, it is observed that the MC simulations exhibit a peak which reproduces that in the data after the application of the data-driven corrections on $E_{\mathrm{T}}^{\mathrm{iso}}$ to the MC simulations; this feature was previously discussed in section 7.3.

As aforementioned, the same 2D-sideband method as in the previous publications was used. The control regions in the two-dimensional plane (see figure 7.9) were defined in the same way. The descriptions provided in section 7.3 for the methodology of the background extraction are also valid for this analysis. The following paragraphs provide a description of the results obtained for the signal leakage fractions and purities for this analysis.

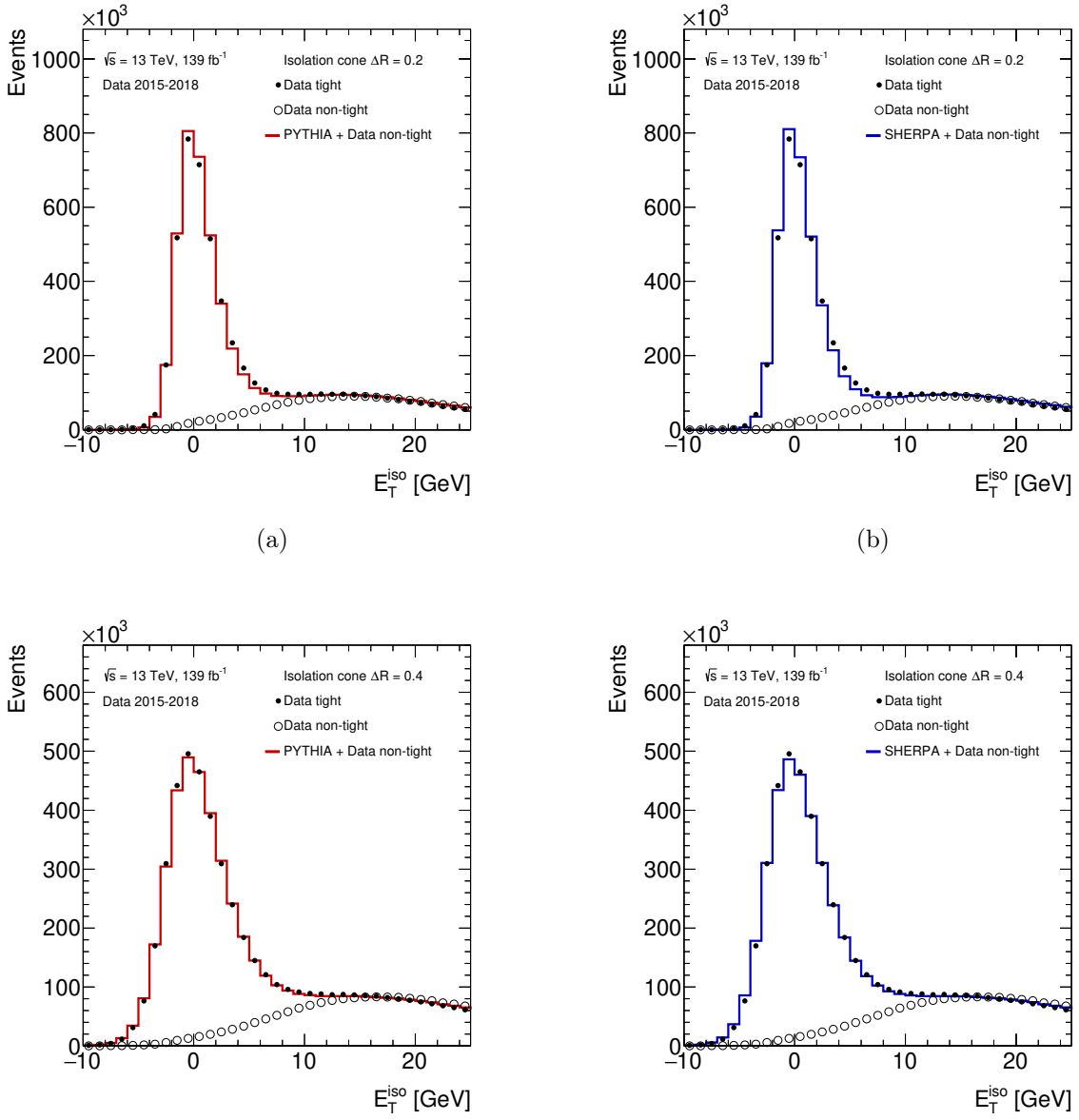


Figure 8.7: The measured $E_{\mathrm{T}}^{\mathrm{iso}}$ distribution before the isolation requirement and after applying the tight identification requirement (dots) and for those events which fail the tight identification (open circles). Distributions for the $R = 0.2$ ($R = 0.4$) sample are shown at the top (bottom). The non-tight distribution was normalised so that the integral of the tight and non-tight distributions for $E_{\mathrm{T}}^{\mathrm{iso}} > 10$ GeV coincides. For comparison, the MC simulations of the signal from PYTHIA (a,c) and SHERPA (b,d) (histograms) are also included. The non-tight contributions were added to the MC distributions.

The signal leakage fractions were extracted from the MC simulations of the signal and are shown in figure 8.8 for PYTHIA and SHERPA for both $R = 0.2$ and $R = 0.4$ samples. The fractions ϵ_{C}, which represent the signal leakage into the non-tight and isolated control region, are approximately constant in all $|\eta^{\gamma}|$ regions and around $0.02 - 0.05$, being very similar for PYTHIA and SHERPA and for both $R = 0.2$ and $R = 0.4$ samples. The fractions ϵ_{B}, which represent the signal leakage into the tight and non-isolated background control region, are approximately constant for all $|\eta^{\gamma}|$ regions, except for SHERPA in the $R = 0.2$ sample, where they increase as a function of E_{T}^{γ} and differ more compared to those of PYTHIA. The fractions take values of ≈ 0.1 (≈ 0.05) for PYTHIA and ≈ 0.05 (around $0.01 - 0.05$) for SHERPA for the $R = 0.4$ ($R = 0.2$) sample. As discussed in section 7.3, the differences between PYTHIA and SHERPA are attributed to the different treatment of the brem component in both MC generators. The fractions ϵ_{D}, which represent the signal leakage into the non-tight and non-isolated control region, increase as a function of E_{T}^{γ}, and take very different values for PYTHIA and SHERPA. The fractions are similar for PYTHIA in both $R = 0.2$ and $R = 0.4$ samples, while large differences are observed in the case of SHERPA. This control region is the most affected by the different treatment of the brem component in both models.

The signal yield was determined from the observed number of events in the data in the four regions of the γ_{ID} vs $E_{\mathrm{T}}^{\mathrm{iso}}$ plane and the signal leakage fractions obtained from the simulated events using equation 7.1. The signal purities, computed as $P = N_{\mathrm{A}}^{\mathrm{sig}}/N_{\mathrm{A}}$ using the signal leakage fractions from PYTHIA and SHERPA are shown in figure 8.9. For some of the points at high E_{T}^{γ} in which the purity is above unity, the purity was set to unity. The signal purity is $\geq 90\%$ and the estimations using either PYTHIA or SHERPA to compute the signal leakage fractions show very similar values. It is observed that the signal purities in the $R = 0.4$ sample are higher than for the $R = 0.2$ sample. The nominal signal yields were extracted using the signal leakage fractions from SHERPA; while the estimation using signal leakage fractions from PYTHIA was taken as a systematic uncertainty on the purity (see section 8.7.1).

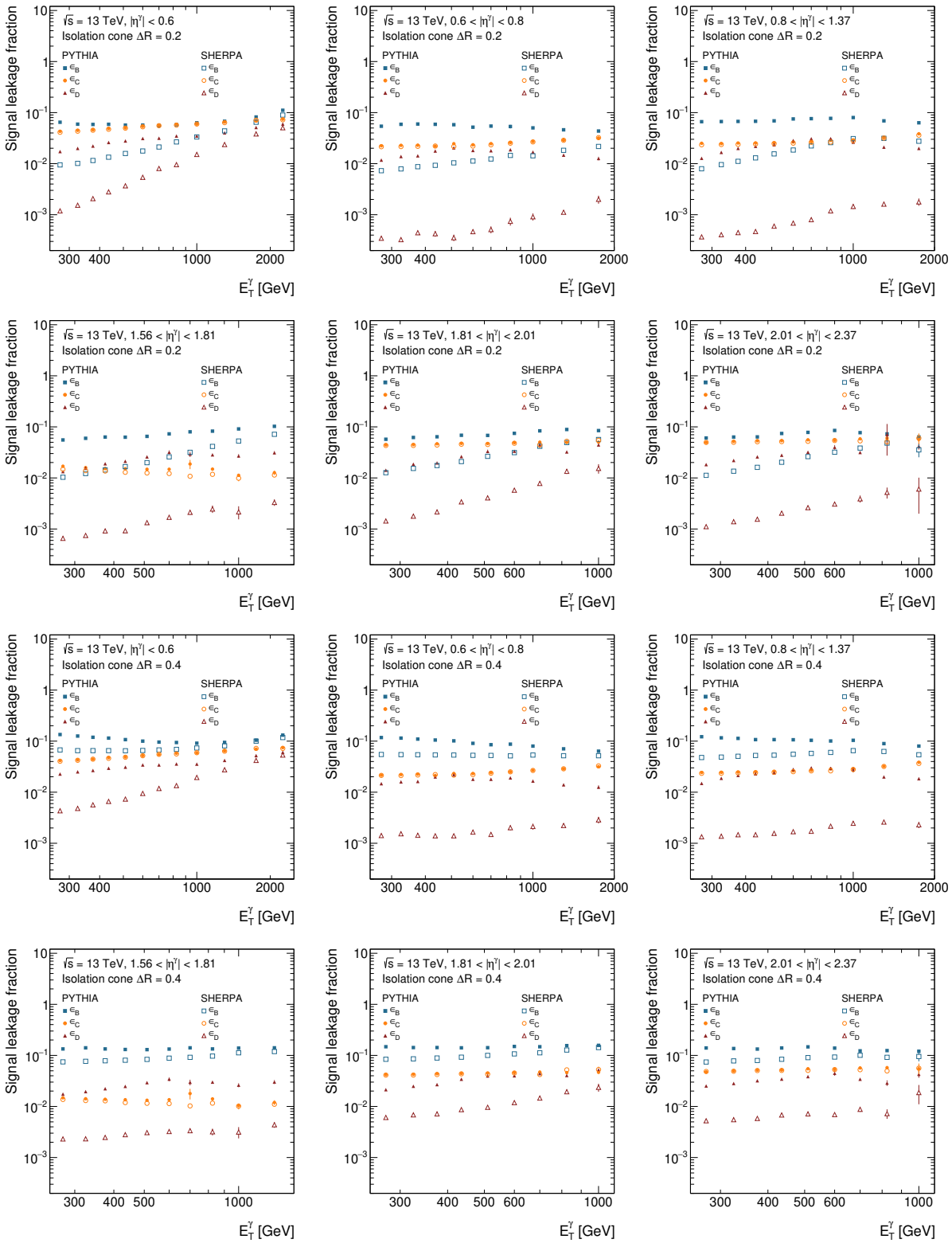


Figure 8.8: Signal leakage fractions from PYTHIA (solid markers) and SHERPA (open markers) for the control regions B (squares), C (circles) and D (triangles) as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$ for the $R = 0.2$ (top) and $R = 0.4$ (bottom) samples.

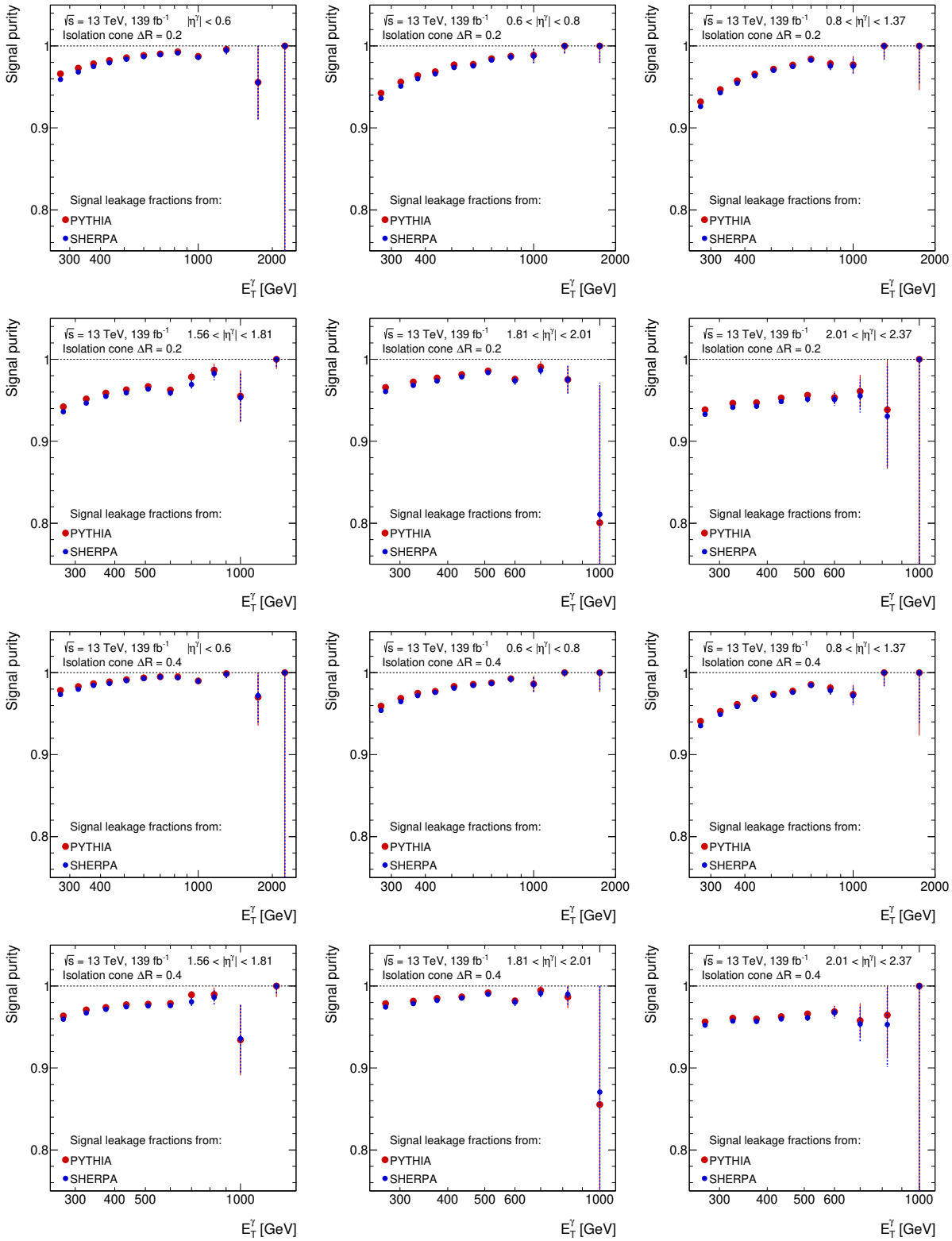


Figure 8.9: Estimated signal purities in data using signal leakage fractions from PYTHIA (red dots) and SHERPA (blue dots) as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$ for the $R = 0.2$ (top) and $R = 0.4$ (bottom) samples.

8.3.1 Background from electrons faking photons

The background contribution of electrons or positrons misidentified as photons was estimated applying the same methodology which is described in section 7.3.1, using SHERPA 2.2.1 MC simulations of W+jets and Z+jets to estimate the residual background. Details of the SHERPA samples are given in section 4.3.

The fraction of W+jets and Z+jets background to inclusive photon production as a function of E_{T}^{γ} for different regions of $|\eta^{\gamma}|$ was estimated with a particle-level matching to $e^{\pm}$ using an angular distance requirement $\Delta R < 0.2$. The total fraction of electron background is calculated adding linearly the contributions of W+jets and Z+jets, and is shown in figure 8.10 (8.11) for the sample with isolation cone radius $R = 0.2$ ($R = 0.4$).

The background from electrons faking photons was found to be approximately smaller than: 0.4% for $|\eta^{\gamma}| < 0.6$; 0.5% for $0.6 < |\eta^{\gamma}| < 0.8$; 0.6% for $0.8 < |\eta^{\gamma}| < 1.37$; 0.9% for $1.56 < |\eta^{\gamma}| < 1.81$; 1.2% for $1.81 < |\eta^{\gamma}| < 2.01$; and 1.3% for $2.01 < |\eta^{\gamma}| < 2.37$. A systematic uncertainty of the size of the evaluated background was assigned (see section 8.7.2).

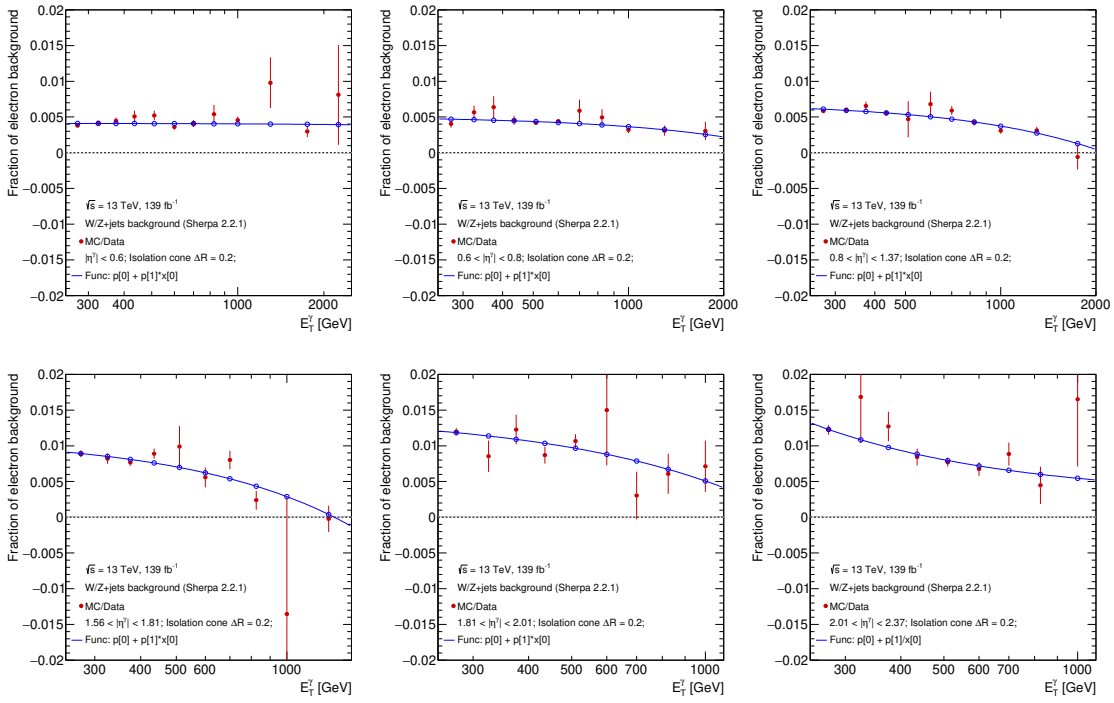


Figure 8.10: Fraction of electron background to inclusive photon production as a function of E_{T}^{γ} for different regions in $|\eta^{\gamma}|$ using the samples of SHERPA 2.2.1 (red dots). The error bars represent the statistical uncertainties of the ratios. The blue line represents the result of a fit to a smooth function. Results are shown for the sample with isolation cone radius $R = 0.2$.

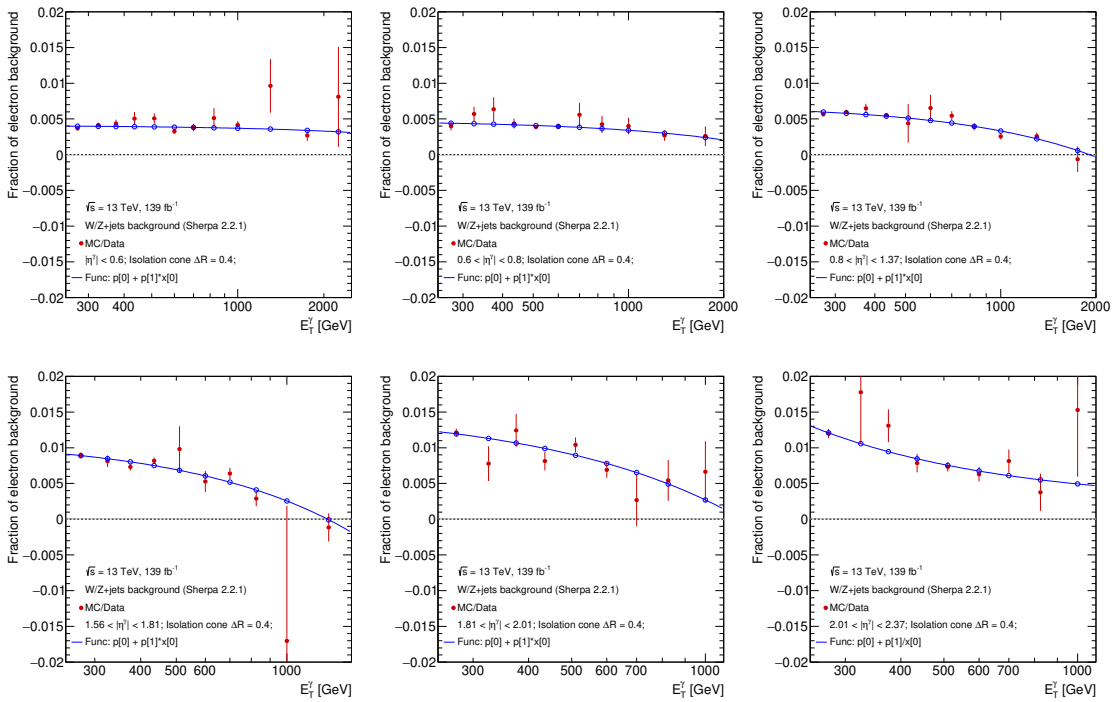


Figure 8.11: Fraction of electron background to inclusive photon production as a function of E_{T}^{γ} for different regions in $|\eta^{\gamma}|$ using the samples of SHERPA 2.2.1 (red dots). The error bars represent the statistical uncertainties of the ratio. The blue line represents the result of a fit to a smooth function. Results are shown for the sample with isolation cone radius $R = 0.4$.

8.4 Control plots

The estimated signal yields using the signal leakage fractions from SHERPA are shown in figures 8.12 and 8.13 as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$ for the $R = 0.2$ and $R = 0.4$ samples, respectively. The signal yields obtained using signal leakage fractions from PYTHIA are very similar, as evidenced by the similar signal purity (see figure 8.9). The measured distributions decrease with increasing E_{T}^{γ} by approximately five orders of magnitude within the measured range. The signal yield in the $R = 0.2$ sample is larger than for $R = 0.4$. For comparison, the simulations of PYTHIA and SHERPA are also included in the figures, and both PYTHIA and SHERPA provide a reasonable description of the data in all $|\eta^{\gamma}|$ regions, except at high E_{T}^{γ}.

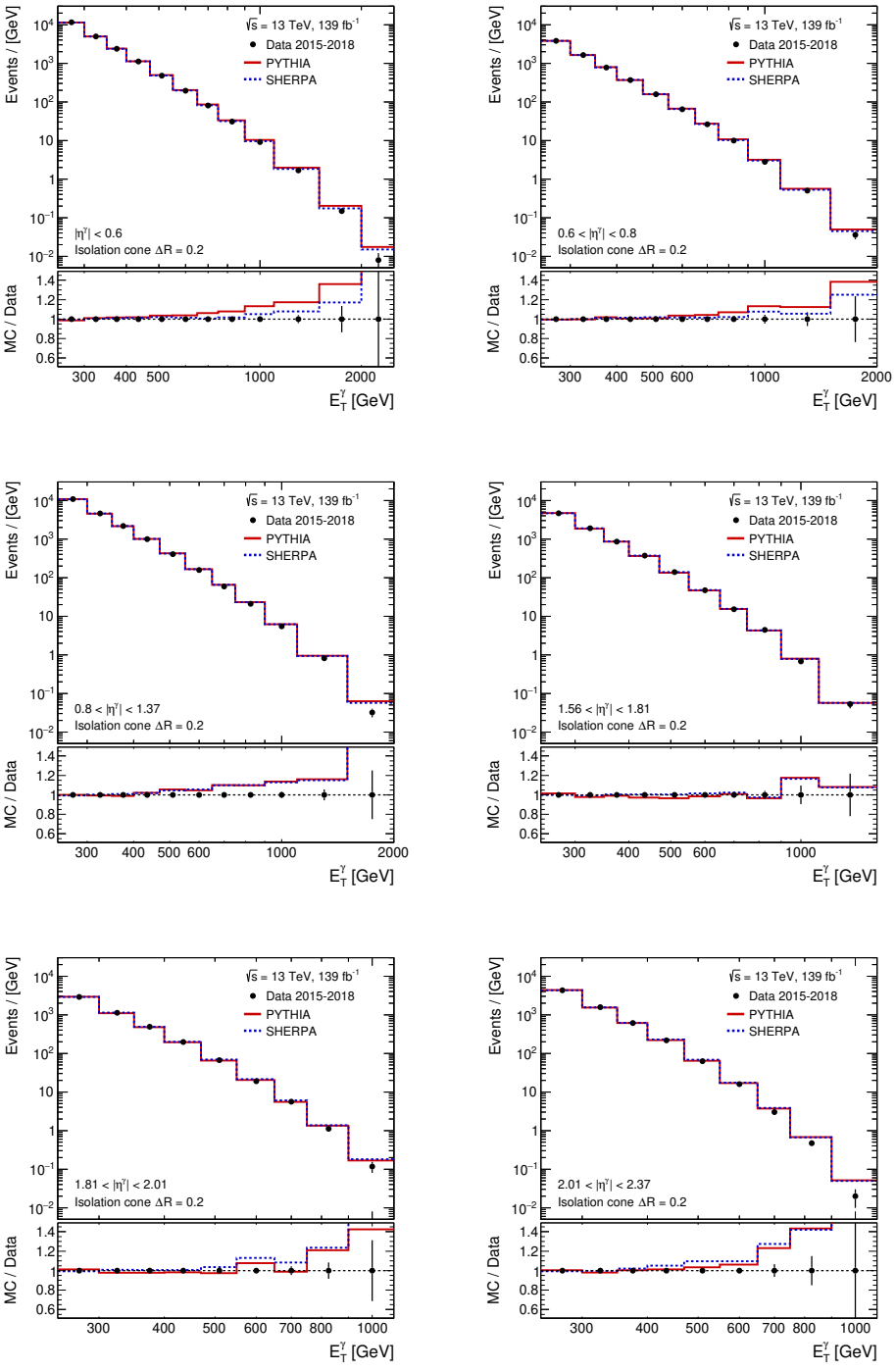


Figure 8.12: The extracted signal yield in data (dots) using signal leakage fractions from SHERPA as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$ for the $R = 0.2$ sample. For comparison, detector level simulations from PYTHIA (solid histograms) and SHERPA (dashed histograms) are also included. The MC distributions are normalised to the number of data events. The ratio of the MC distributions to the data is shown in the lower panel of the figures. The error bars display the statistical uncertainty of the data.

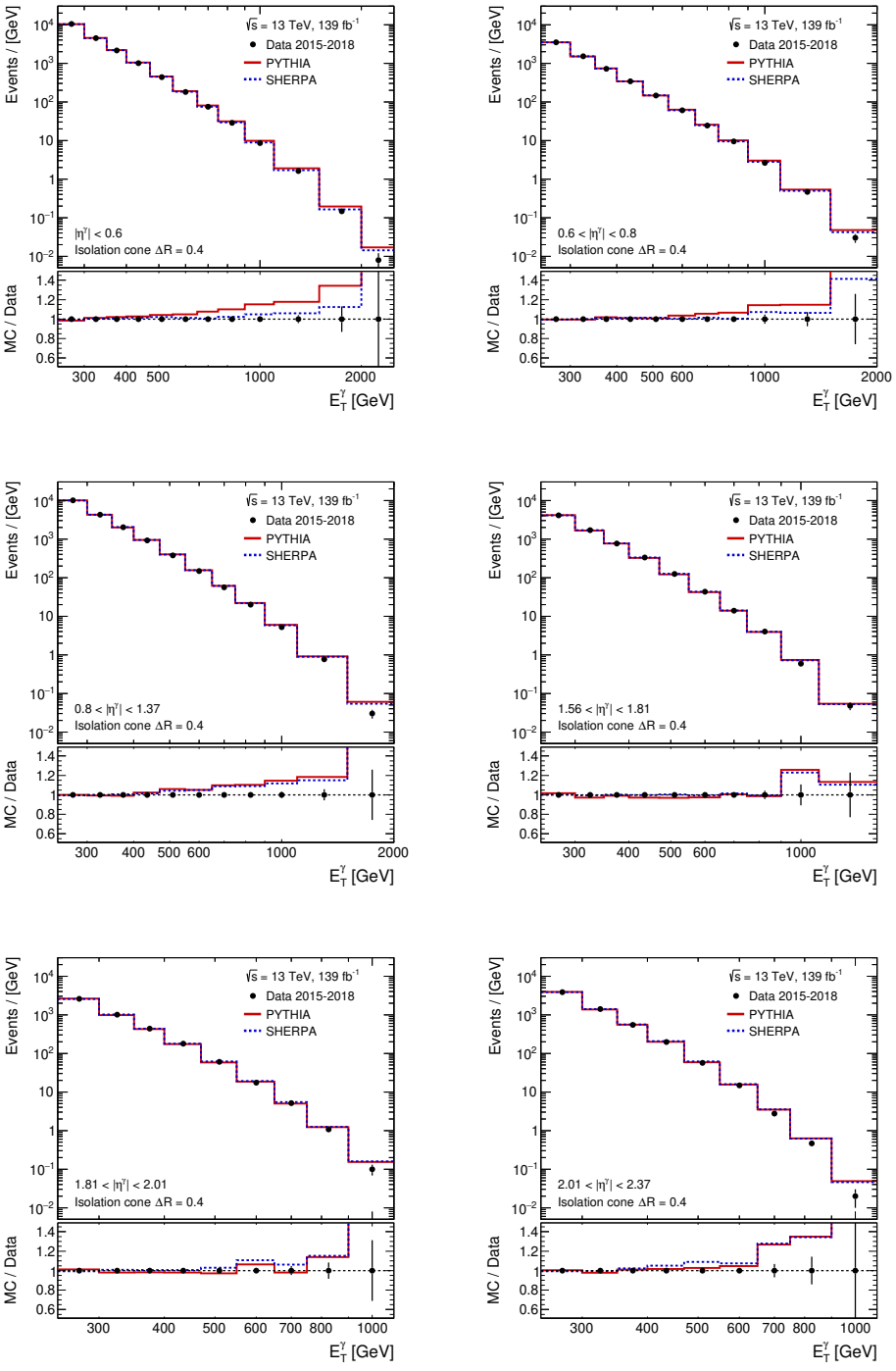


Figure 8.13: The extracted signal yield in data (dots) using signal leakage fractions from SHERPA as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$ for the $R = 0.4$ sample. For comparison, detector level simulations from PYTHIA (solid histograms) and SHERPA (dashed histograms) are also included. The MC distributions are normalised to the number of data events. The ratio of the MC distributions to the data is shown in the lower panel of the figures. The error bars display the statistical uncertainty of the data.

8.5 Selection efficiency and purity

The quality of the reconstruction and the selection efficiency and purity were studied using the samples of MC events from PYTHIA and SHERPA. The MC events were required to fulfil the same kinematic and isolation requirements both at the detector and particle levels. In addition, the particle- and detector-level photons were required to be matched using the angular requirement $\Delta R < 0.2$ in order to remove fake photons.

As previously defined in section 7.5, the integrated selection efficiency is computed in the same way, using equation 7.3. The integrated selection efficiency was found to be $(85.22 \pm 0.03)\%$ for PYTHIA and $(85.71 \pm 0.01)\%$ for SHERPA for the $R = 0.2$ sample, while for $R = 0.4$, it was $(81.64 \pm 0.03)\%$ for PYTHIA and $(81.78 \pm 0.01)\%$ for SHERPA.

The bin-to-bin selection efficiency was computed using equation 7.4. Figure 8.14 shows the bin-to-bin selection efficiencies as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$ for PYTHIA and SHERPA and the two $R = 0.2$ and $R = 0.4$ samples. The bin-to-bin selection efficiency is between $65\% - 85\%$ and is very similar for PYTHIA and SHERPA, being slightly higher for $R = 0.2$ than for the $R = 0.4$ sample.

The integrated selection purity is computed in the same way as in section 8.5, using equation 7.5. The integrated selection purity was found to be $(95.42 \pm 0.03)\%$ for PYTHIA and $(96.51 \pm 0.01)\%$ for SHERPA for the $R = 0.2$ sample, while for $R = 0.4$, it was $(94.06 \pm 0.03)\%$ for PYTHIA and $(94.16 \pm 0.01)\%$ for SHERPA.

The bin-to-bin selection purity was computed using equation 7.6. Figure 8.15 shows the bin-to-bin selection purities as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$ for PYTHIA and SHERPA and the two $R = 0.2$ and $R = 0.4$ samples. The bin-to-bin selection purity is above 85% and is very similar for PYTHIA and SHERPA, being slightly higher for $R = 0.2$ than for the $R = 0.4$ sample.

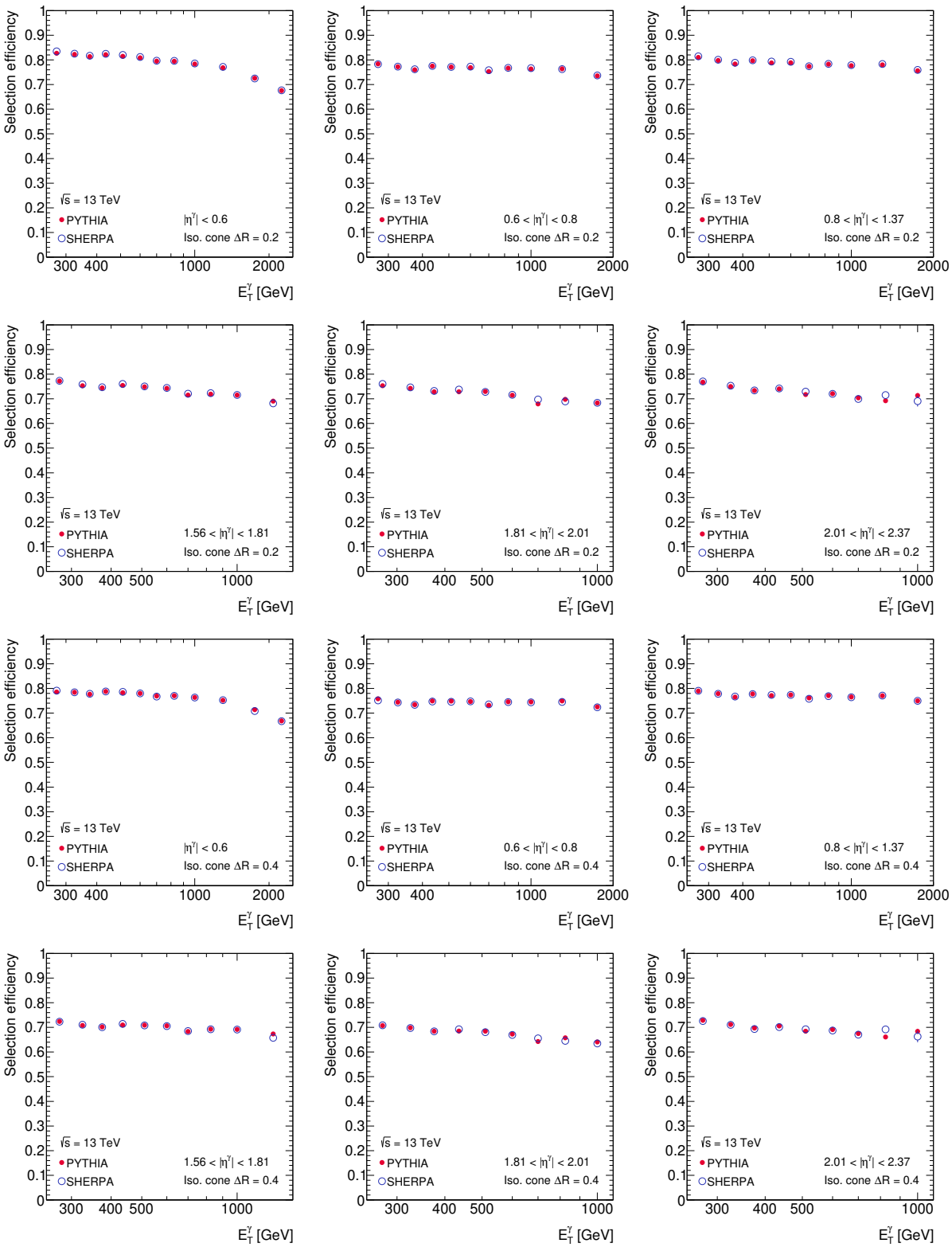


Figure 8.14: Selection efficiencies from PYTHIA (solid circles) and SHERPA (open circles) as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. Results for the $R = 0.2$ ($R = 0.4$) sample are shown at the top (bottom).

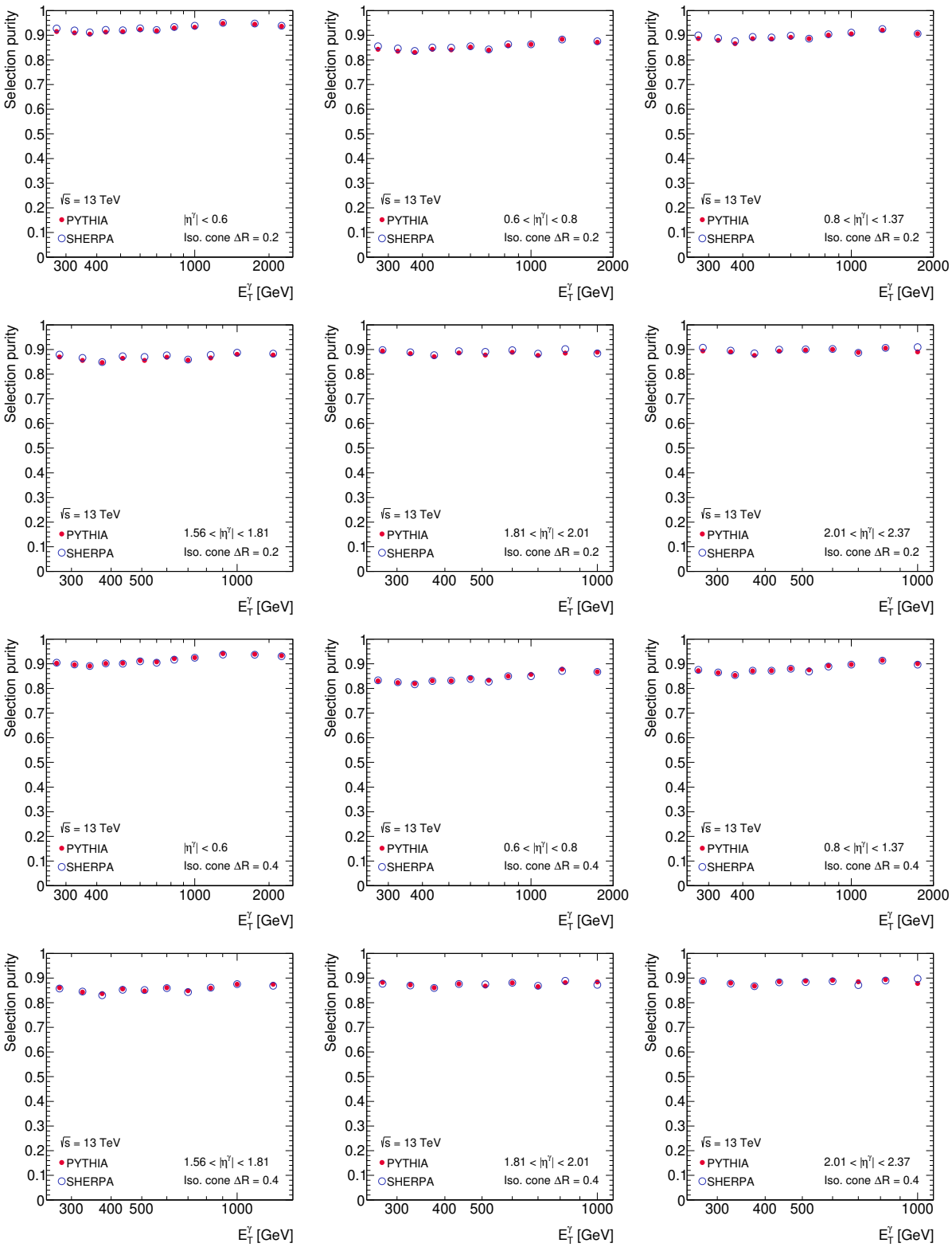


Figure 8.15: Selection purities from PYTHIA (solid circles) and SHERPA (open circles) as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. Results for the $R = 0.2$ ($R = 0.4$) sample are shown at the top (bottom).

8.6 Cross-section measurement procedure

The isolated-photon cross section was measured for photons with $E_{\mathrm{T}}^{\gamma} > 250$ GeV and $|\eta^{\gamma}| < 2.37$ (excluding the region $1.37 < |\eta^{\gamma}| < 1.56$). Isolation was ensured by requiring $E_{\mathrm{T}}^{\mathrm{iso}} < 4.2 \times 10^{-3} \times E_{\mathrm{T}}^{\gamma} + 4.8$ GeV. The previous cuts define the phase space of the measurement. The particle-level isolation requirement on the photon is built summing the transverse energy of all stable particles, except for muons and neutrinos, in a cone of size $\Delta R = 0.2$ or $\Delta R = 0.4$ around the photon direction after the contribution from the UE is subtracted; the same subtraction procedure used on data is applied at the MC particle level (see section 5.5); the particles associated with the overlaid *pp* collisions were not considered. The data were unfolded to the particle level to this region of phase space, as explained in the following.

The isolated-photon differential cross section was measured as a function of E_{T}^{γ} in six different regions of $|\eta^{\gamma}|$, namely, $|\eta^{\gamma}| < 0.6$, $0.6 < |\eta^{\gamma}| < 0.8$, $0.8 < |\eta^{\gamma}| < 1.37$, $1.56 < |\eta^{\gamma}| < 1.81$, $1.81 < |\eta^{\gamma}| < 2.01$ and $2.01 < |\eta^{\gamma}| < 2.37$ for the two isolation cone radii $R = 0.2$ and $R = 0.4$ separately.

The data distributions after background subtraction (see section 8.3) were unfolded to the particle level using the iterative application of the Bayes' theorem to obtain the measured differential cross section. The Bayesian unfolding method [153] as implemented in RooUnfold [159] was used. A detailed description of the method is provided in appendix E. In this method, the repeated application of Bayes' theorem is used to invert the response matrix, and regularisation is achieved by stopping the iterations. The results were found to be fairly insensitive to the number of iterations; two iterations were used in this analysis.

The unfolding matrices used in the Bayes' method are shown in figures 8.16 and 8.17 for PYTHIA and SHERPA, respectively. The reconstructed photons which are not matched to a truth photon (*reco unmatched*) as well as reconstruction inefficiencies (truth photons which are not matched to a reconstructed photon; *truth unmatched*) are taken into account in the Bayesian method. The unmatched fractions are shown in figure 8.18 for both $R = 0.2$ and $R = 0.4$ samples.

The nominal cross section was measured using the unfolding matrices from SHERPA, and the differences in the cross section obtained by using PYTHIA to unfold the data were taken as a systematic uncertainty of the effect of the QCD-cascade and hadronisation models in the corrections (see section 8.7.4).

A cross-check of the unfolding method was performed by reweighting the MC distributions of SHERPA at detector and particle levels to the data. A further cross-check of the unfolding was performed by adjusting the fraction of the brem and hard components in PYTHIA. The details of both cross-checks are described in appendix G.

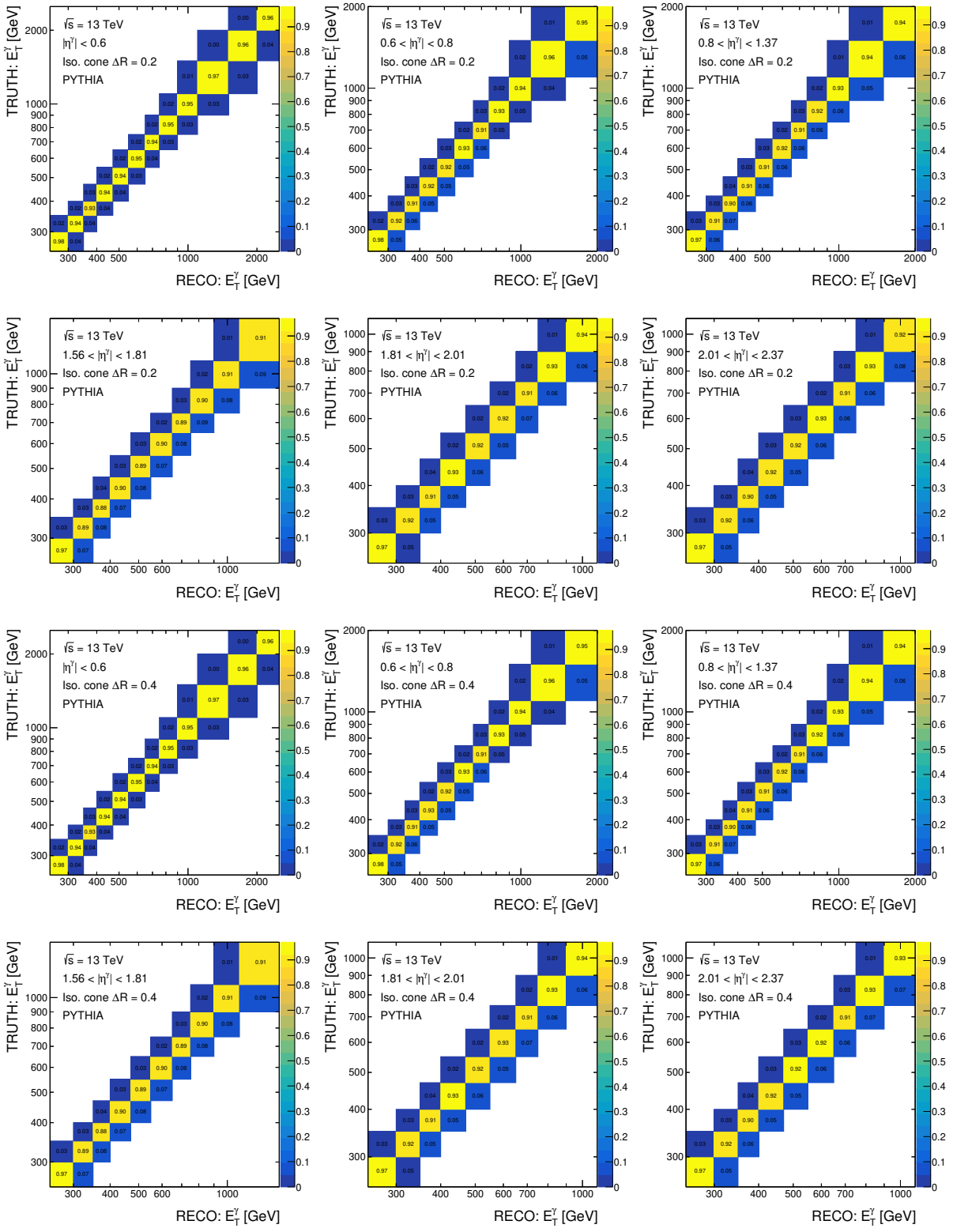


Figure 8.16: Unfolding matrices obtained with PYTHIA MC in the different regions of $|\eta^\gamma|$. Normalisation is column wise. Results for the $R = 0.2$ ($R = 0.4$) sample are shown at the top (bottom).

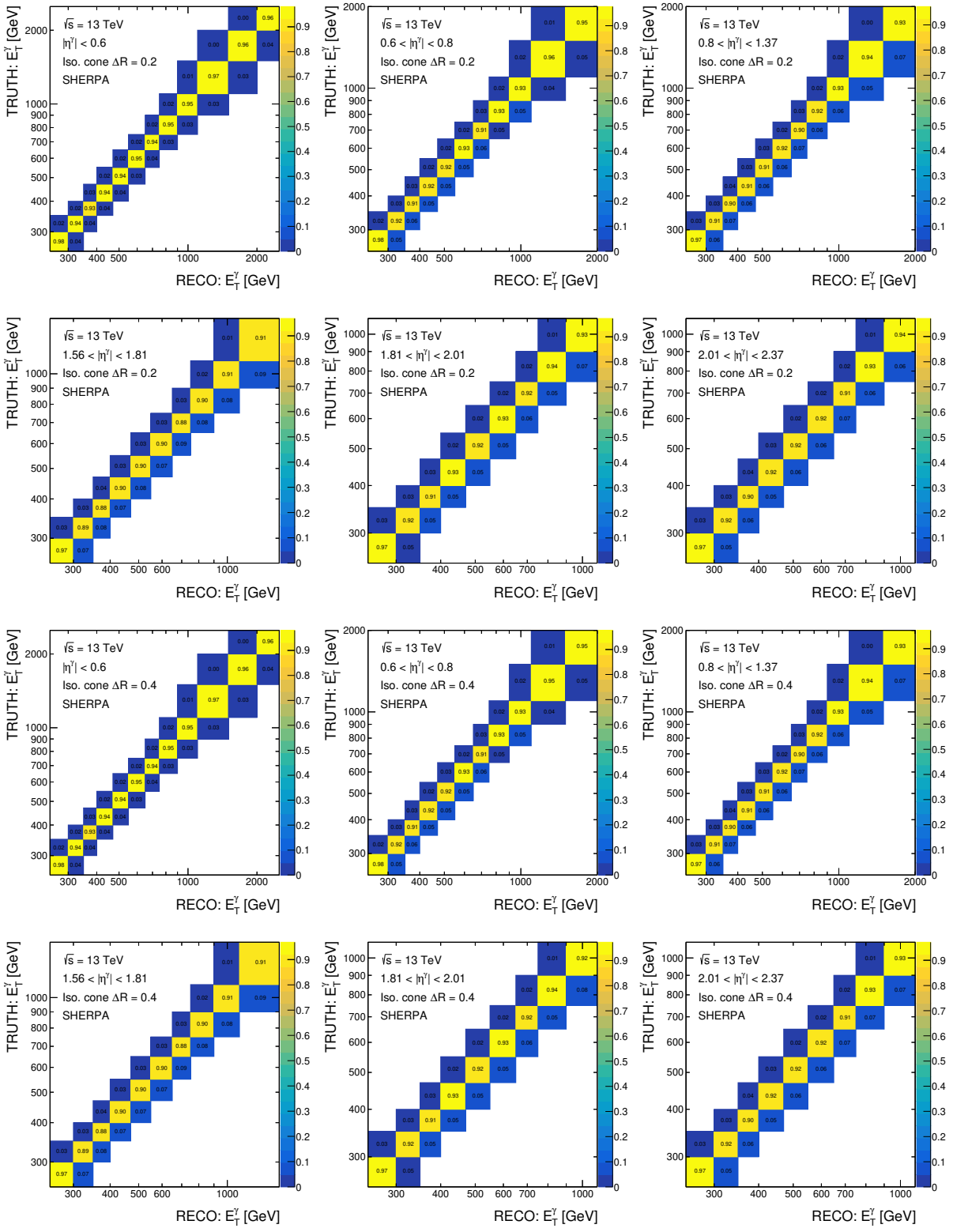


Figure 8.17: Unfolding matrices obtained with SHERPA MC in the different regions of $|\eta^\gamma|$. Normalisation is column wise. Results for the $R = 0.2$ ($R = 0.4$) sample are shown at the top (bottom).

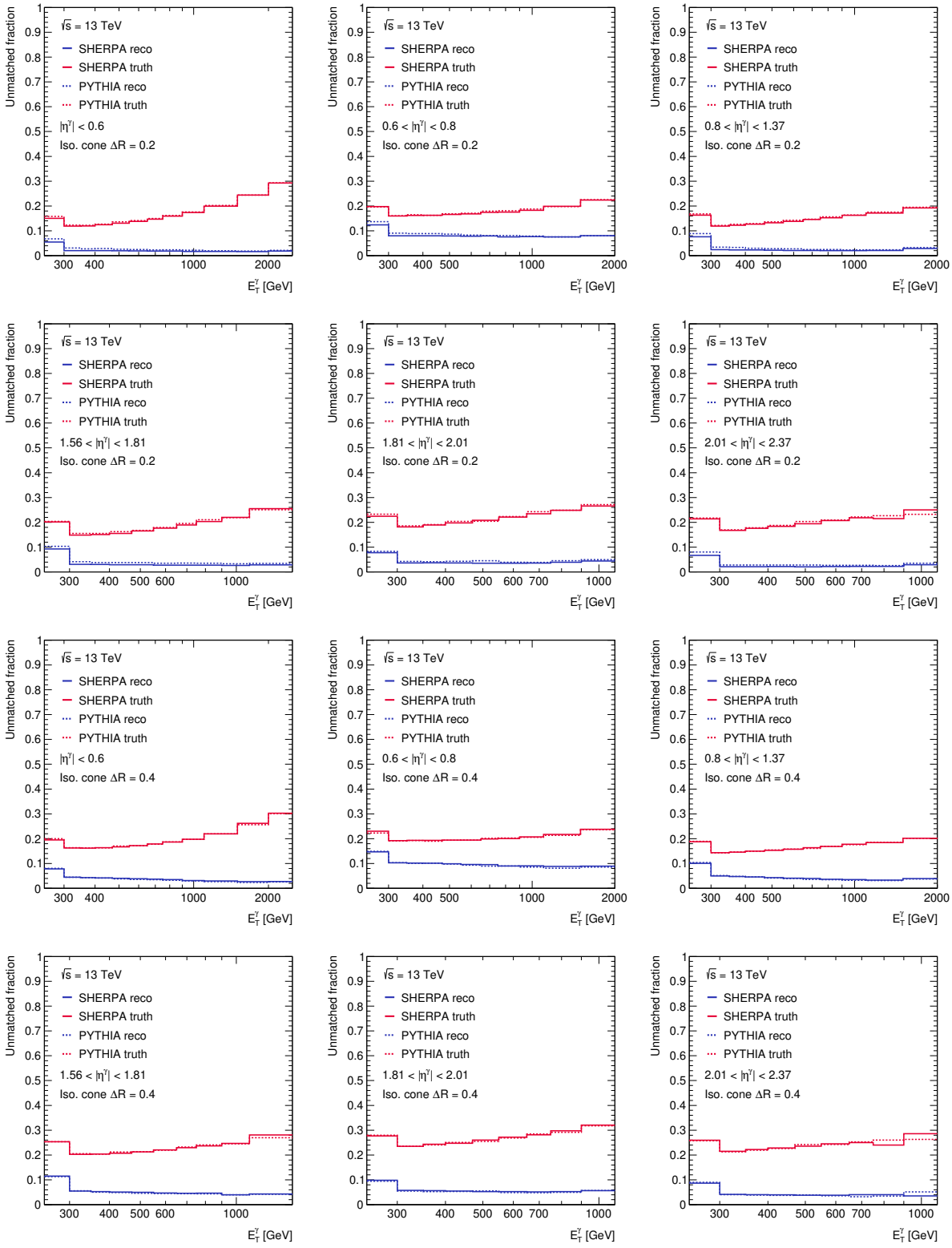


Figure 8.18: The reco (truth) unmatched fractions represented as blue (red) lines as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. SHERPA (PYTHIA) results are represented with solid (dashed) lines. Results for the $R = 0.2$ ($R = 0.4$) sample are shown at the top (bottom).

8.7 Systematic uncertainties

The sources of systematic uncertainty that affect the measurement are discussed below. These sources include the signal modelling, the background subtraction, the photon reconstruction, the unfolding procedure, the running conditions and the photon energy calibration. For some of the systematic uncertainties, the bootstrap technique [154] was used to evaluate the data statistical influence on the uncertainties; the results were then used in a fit to remove the statistical fluctuations in the systematic uncertainties.

8.7.1 Signal modelling

The nominal signal leakage fractions used in the extraction of the signal yield were evaluated with the SHERPA simulation. The effect of the implementation of the matrix-element calculation in the generator as well as the models used for the parton shower and hadronisation were evaluated by comparing the nominal results with those obtained using the PYTHIA simulation for the determination of the signal leakage fractions (see section 8.3). The differences were taken as a systematic uncertainty. The resulting uncertainty in the measured cross section is shown in figure 8.19 and is always less than 1% and typically less than 0.5%. A small increase of the uncertainty is observed for $R = 0.2$ compared to the $R = 0.4$ sample.

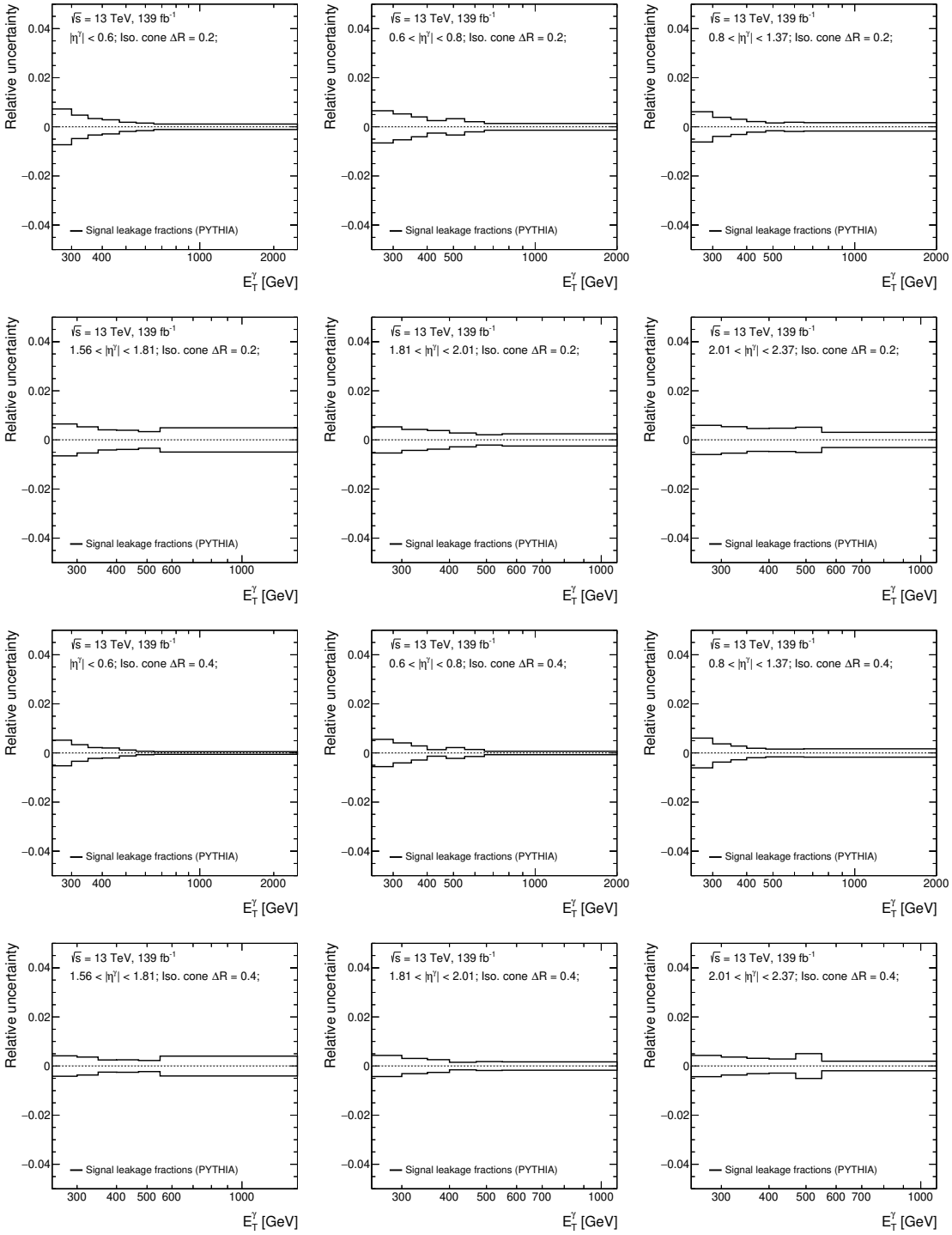


Figure 8.19: Systematic uncertainty in the measured cross section due to the QCD-cascade and hadronisation models as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. The uncertainty was symmetrised. Statistical fluctuations in the estimation of this systematic uncertainty were taken care of by rebinning, merging those bins affected by relatively large statistical fluctuations with neighbouring bins. Results for the $R = 0.2$ ($R = 0.4$) sample are shown at the top (bottom).

8.7.2 Background subtraction

Choice of background control regions

The data-driven background subtraction depends on the choice of background control regions. This dependence is studied following the same methodology described in section 7.7.2, by varying the lower limit on $E_{\mathrm{T}}^{\mathrm{iso}}$ for regions B and D by ± 1 GeV, removing the upper limit on $E_{\mathrm{T}}^{\mathrm{iso}}$ for regions B and D, as well as changing the choice of inverted photon identification variables. The resulting uncertainties in the measured cross section due to the first two variations are less than 0.3% and are shown in figures 8.20 and 8.21, respectively. Small differences are observed for the two different radii. The uncertainty in the cross section from the third variation is shown in figure 8.22 and is less than 1.5% and typically less than 1%. A small increase of the uncertainty is observed for $R = 0.2$ compared to the $R = 0.4$ sample.

Identification and isolation correlation in the background

As explained in section 7.3, the isolation and identification photon variables used to define the 2D-sideband method to subtract the background were assumed to be uncorrelated for background events ($R^{\mathrm{bg}} = 1$ in equation 7.2). The same data-driven method described in section 7.7.2 was used to determine the maximum possible deviations of R^{bg} from this assumption. The maximum deviations from unity in the validation regions are shown in figure 8.23 for PYTHIA and SHERPA, and vary in the range $10\% - 25\%$ depending on E_{T}^{γ} and $|\eta^{\gamma}|$. The uncertainty in the cross section was estimated in the same way as described in section 7.7.2, and is shown in figure 8.24. This uncertainty is less than 2% and typically less than 1%. A small increase of the uncertainty is observed for $R = 0.2$ compared to the $R = 0.4$ sample.

Background from electrons faking photons

As discussed in section 8.3.1, the background contribution from electrons or positrons misidentified as photons is subpercent and a systematic uncertainty is included by assigning the full size of this background from the results of the fits shown in figures 8.10 and 8.11 accounting for the dependence on E_{T}^{γ} and the $|\eta^{\gamma}|$ region. Figure 8.25 shows the resulting uncertainty in the measured cross section. No difference is observed between the uncertainties for the $R = 0.2$ and $R = 0.4$ samples.

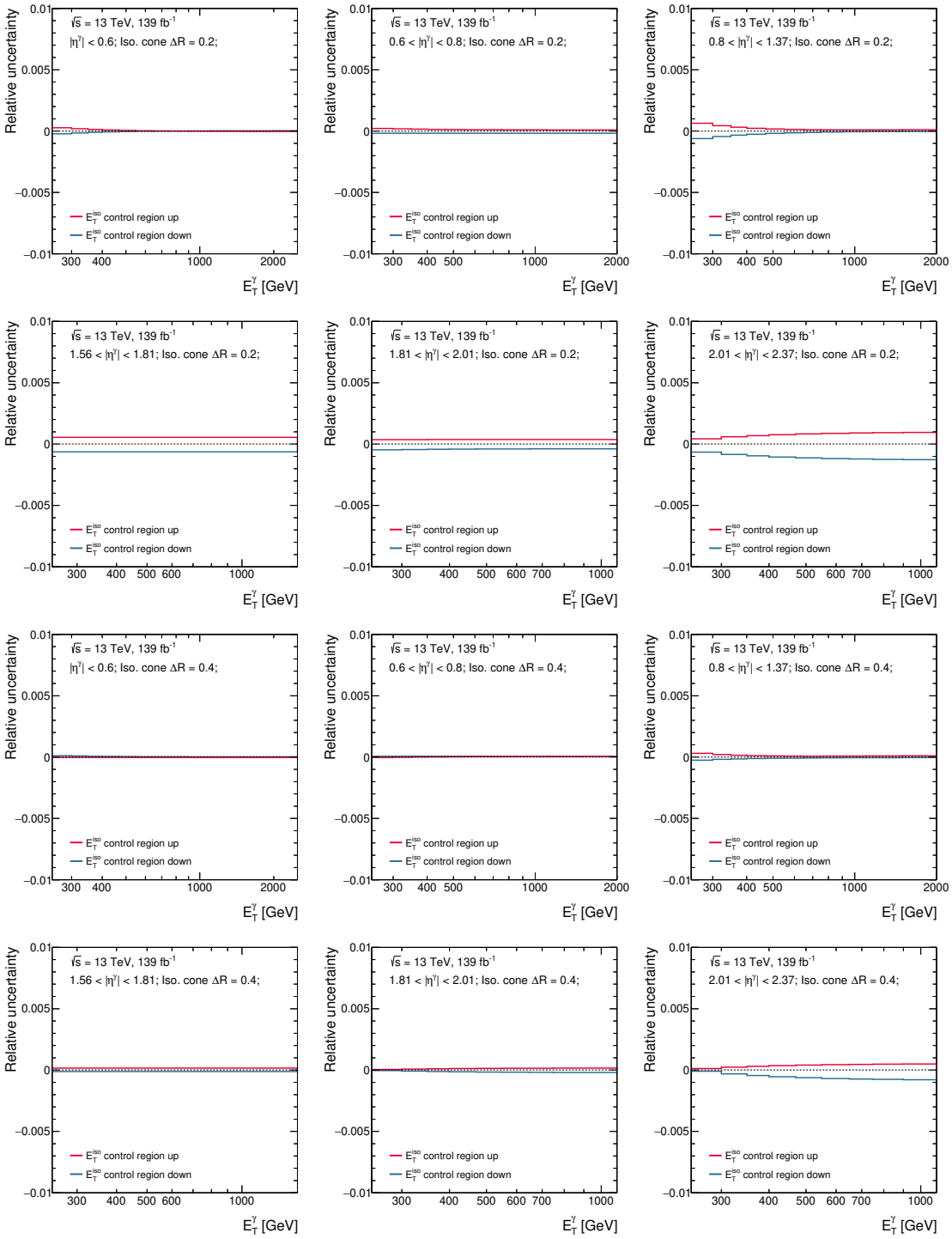


Figure 8.20: Systematic uncertainty in the measured cross section due to the choice of background control regions (different $E_{\mathrm{T}}^{\mathrm{iso}}$ requirements) as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. Statistical fluctuations in the estimation of this systematic uncertainty were taken care of by using a bootstrap technique. Results for the $R = 0.2$ ($R = 0.4$) sample are shown at the top (bottom).

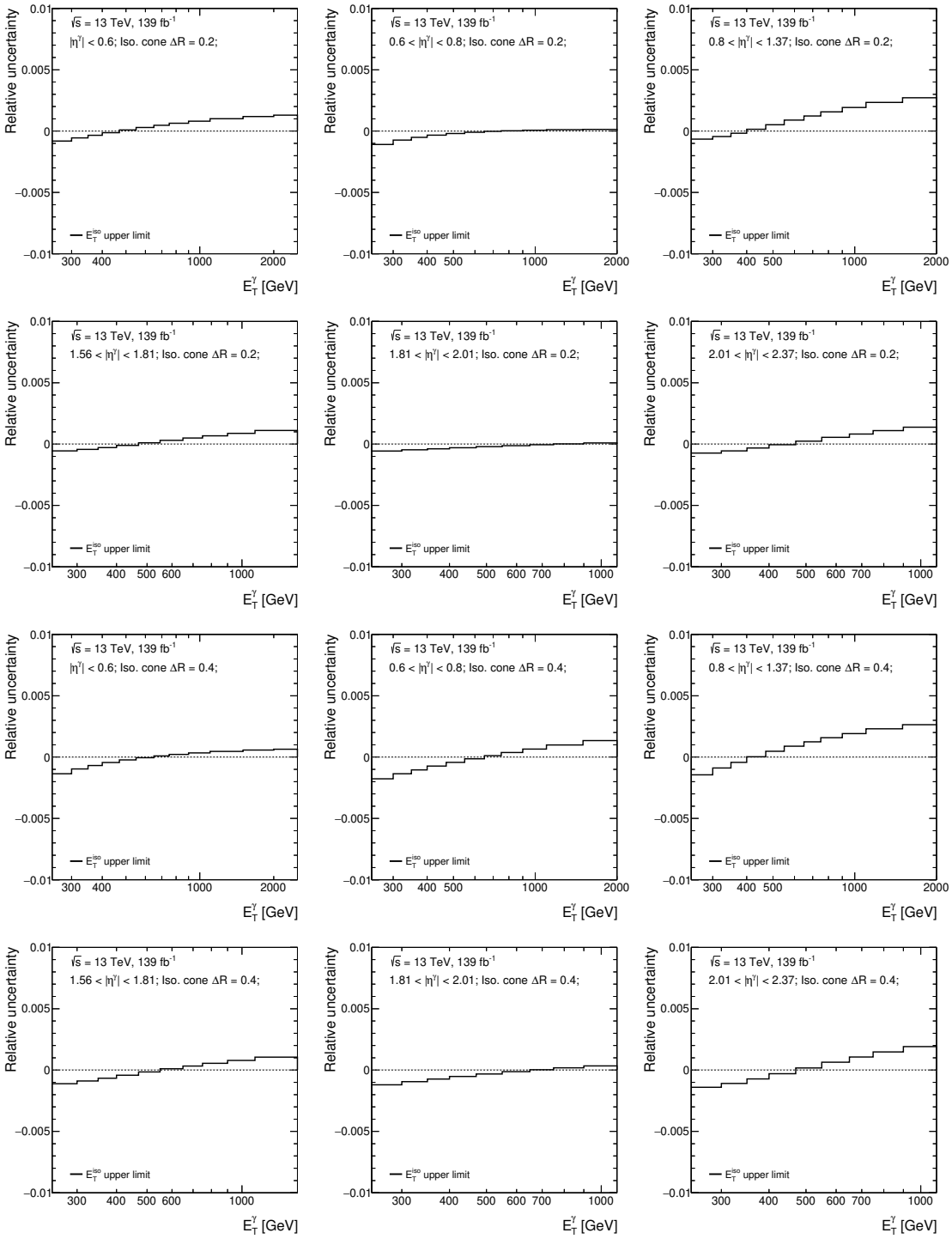


Figure 8.21: Systematic uncertainty in the measured cross section due to the choice of background control regions (removal of the upper limit on $E_{\mathrm{T}}^{\mathrm{iso}}$ for regions B and D) as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. Statistical fluctuations in the estimation of this systematic uncertainty were taken care of by using a bootstrap technique. Results for the $R = 0.2$ ($R = 0.4$) sample are shown at the top (bottom).

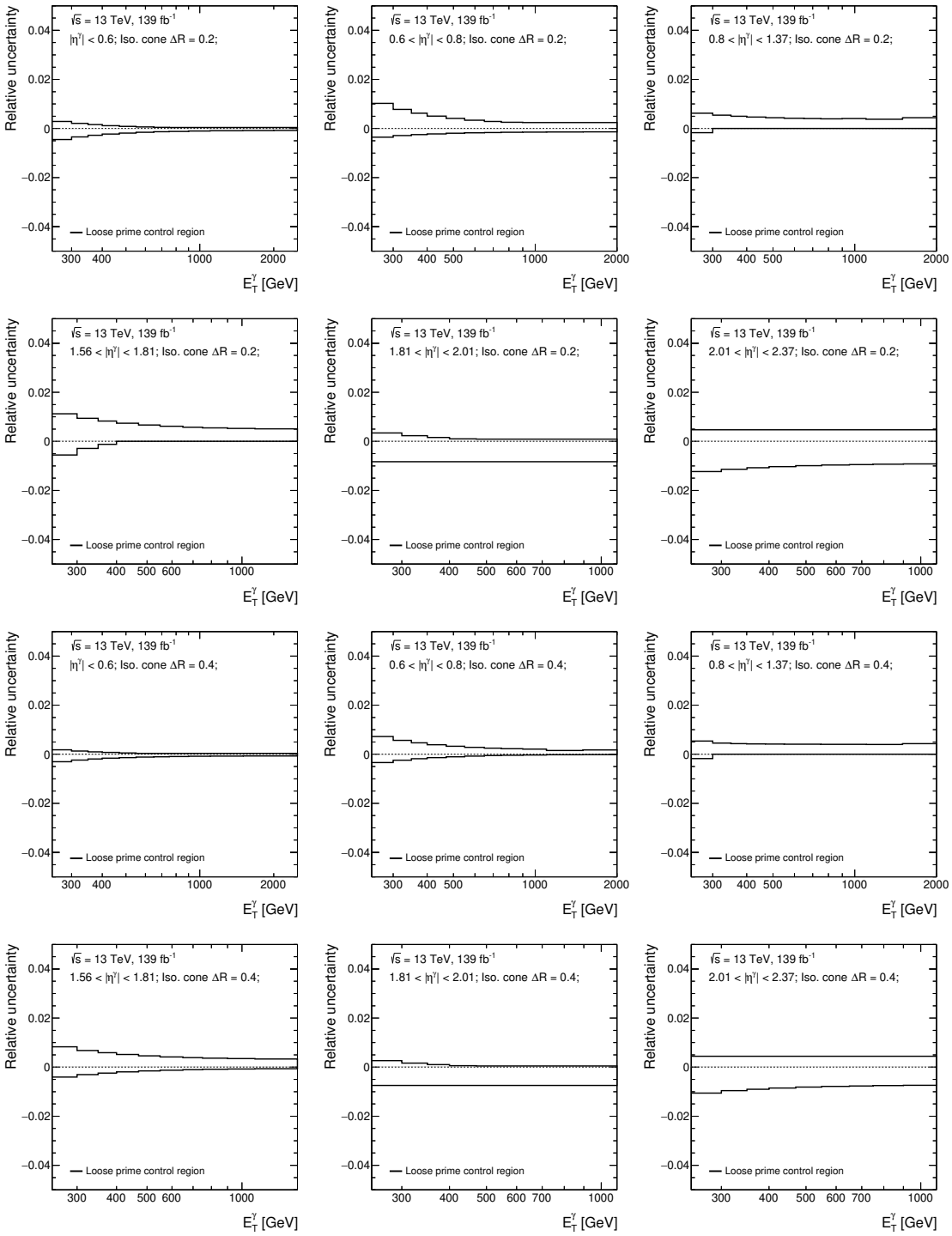


Figure 8.22: Systematic uncertainty in the measured cross section due to the choice of background control regions (different loose′ criteria) as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. Statistical fluctuations in the estimation of this systematic uncertainty were taken care of by using a bootstrap technique. Results for the $R = 0.2$ ($R = 0.4$) sample are shown at the top (bottom).

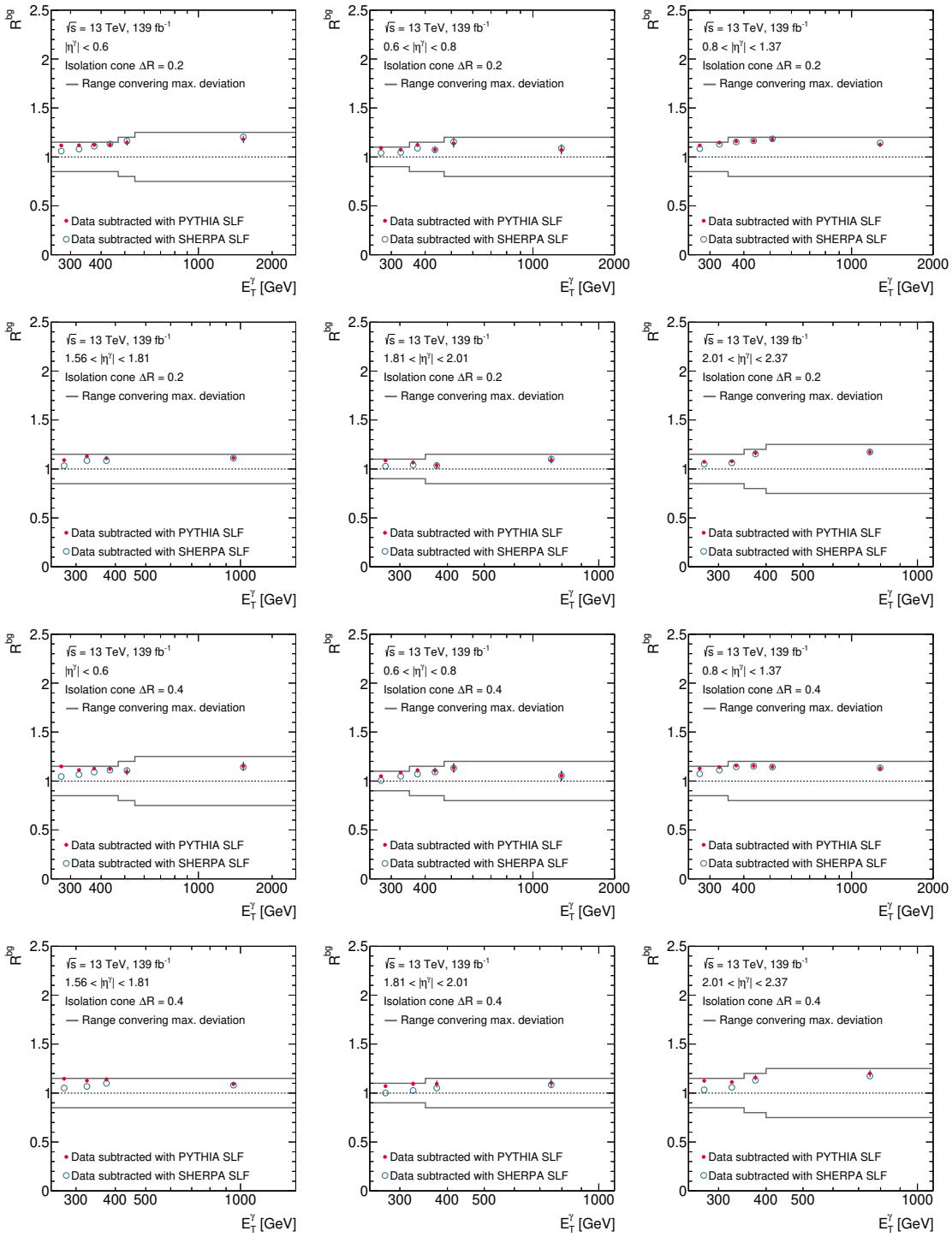


Figure 8.23: The estimated R^{bg} as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. The signal leakage contributions in the control regions were removed using either PYTHIA (red dots) or SHERPA (blue circles) samples. The range in R^{bg} which was taken as the uncertainty is indicated by the solid grey lines. Results for the $R = 0.2$ ($R = 0.4$) sample are shown at the top (bottom).

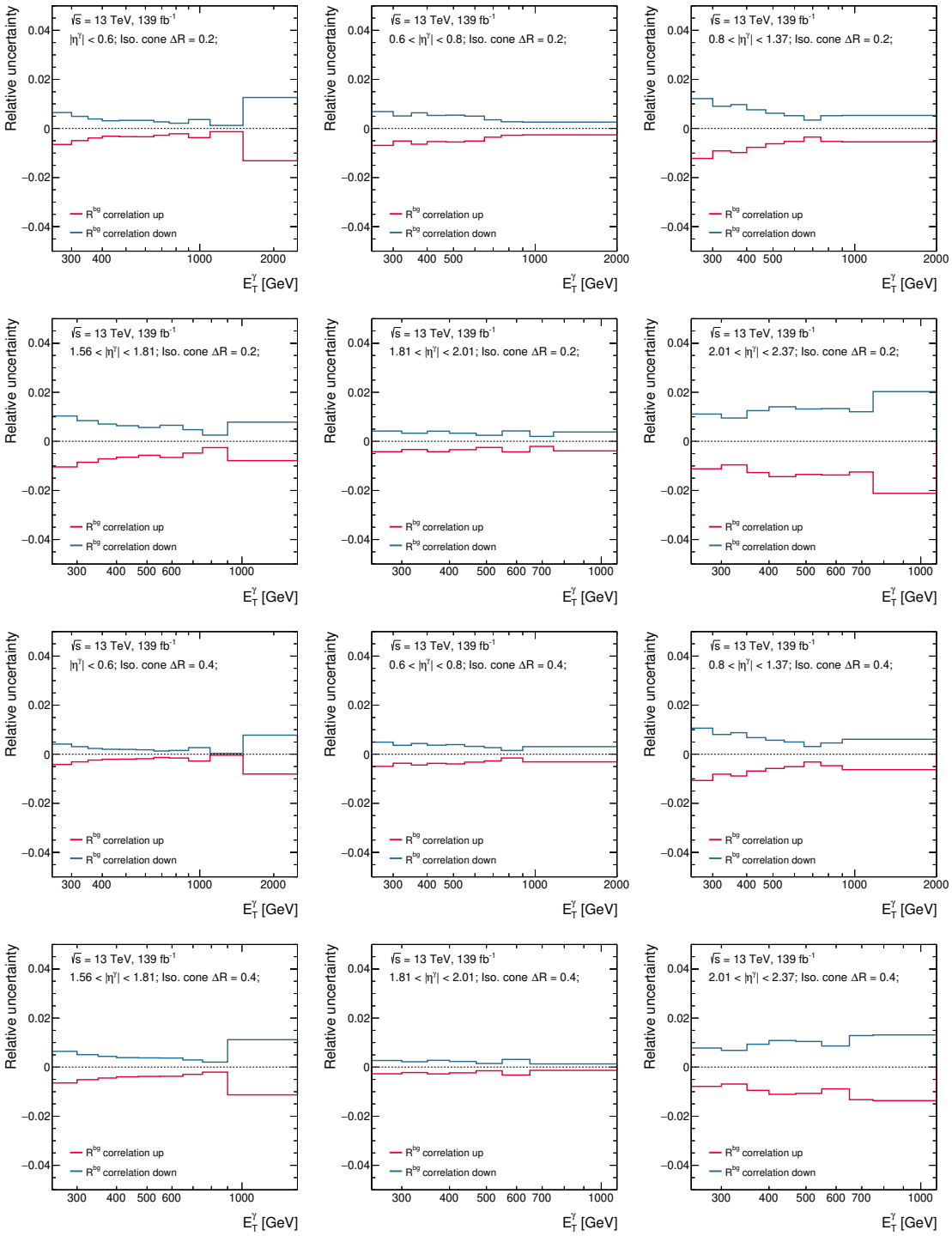


Figure 8.24: Systematic uncertainty in the measured cross section due to the correlation of the identification and isolation variables in the background as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. Results for the $R = 0.2$ ($R = 0.4$) sample are shown at the top (bottom).

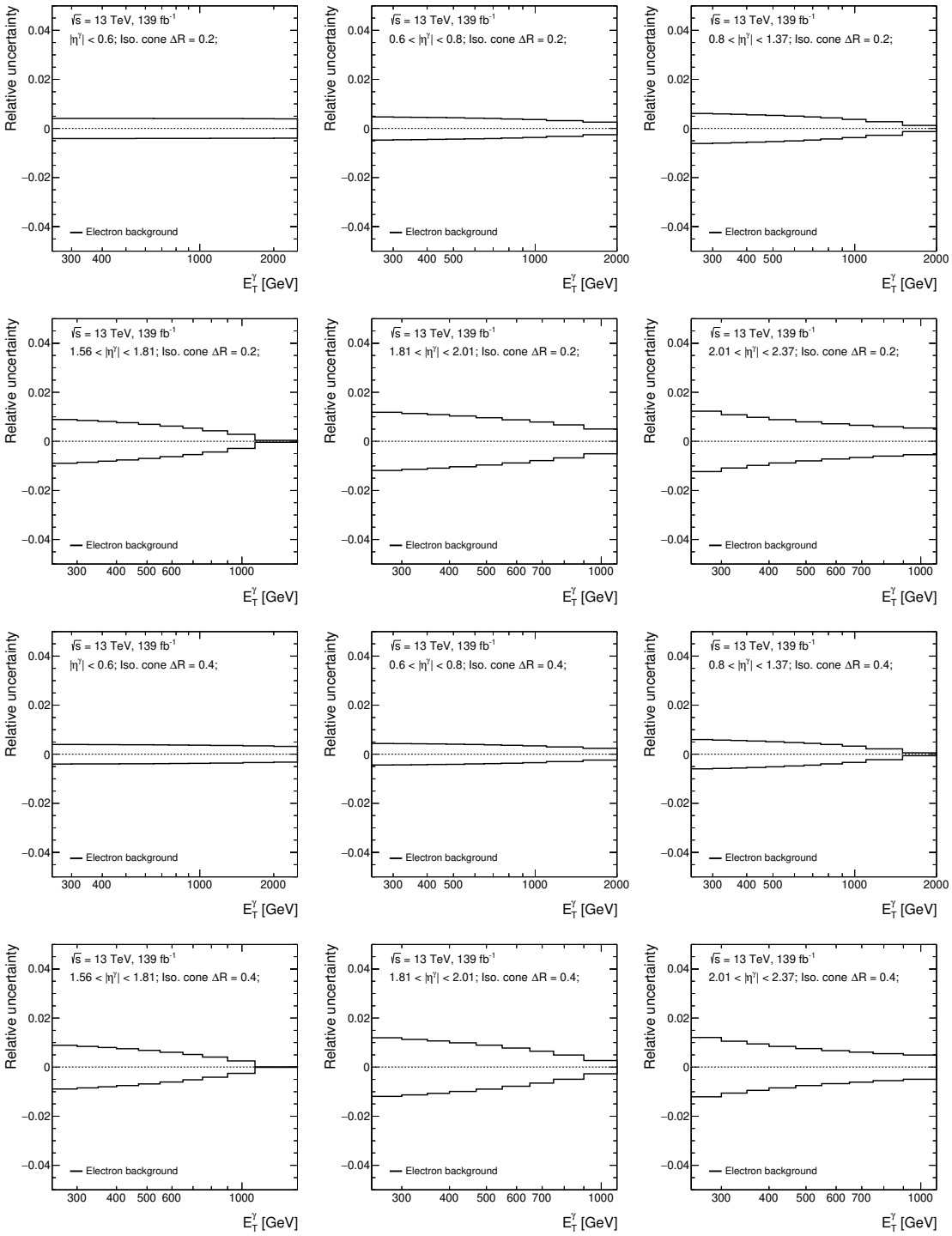


Figure 8.25: Systematic uncertainty in the measured cross section due to the electron background as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. Results for the $R = 0.2$ ($R = 0.4$) sample are shown at the top (bottom).

8.7.3 Photon reconstruction

Photon-reconstruction efficiency

The uncertainty on the photon-reconstruction efficiency was estimated by propagating the uncertainty in the scale factors which were applied to the MC events (see section 8.2). The resulting uncertainty in the measured cross section is in the range $0.2\% - 0.5\%$ and is shown in figure 8.26. No difference is observed between the uncertainties for the $R = 0.2$ and $R = 0.4$ samples.

Photon-identification efficiency

The uncertainty in the efficiency scale factors, which were applied to simulated events to match the tight photon identification efficiency measured in data [131], is propagated to the measured cross section. The resulting uncertainty in the cross section is shown in figure 8.27, and is in the range $0.5\% - 0.75\%$. No difference is observed between the uncertainties for the $R = 0.2$ and $R = 0.4$ samples.

$E_{\mathrm{T}}^{\mathrm{iso}}$ modelling

The uncertainty in the cross section due to the modelling of $E_{\mathrm{T}}^{\mathrm{iso}}$ in simulated events was estimated in the same way described in 7.7.3. The resulting uncertainty in the measured cross section is less than 2% except for $E_{\mathrm{T}}^{\gamma} > 500$ GeV in the region $0.8 < |\eta^{\gamma}| < 1.37$ ($1.56 < |\eta^{\gamma}| < 1.81$), where it increases up to 3.5% (5%). Figure 8.28 shows the resulting uncertainty in the measured cross section. A small increase of the uncertainty is observed for $R = 0.4$ compared to the $R = 0.2$ sample.

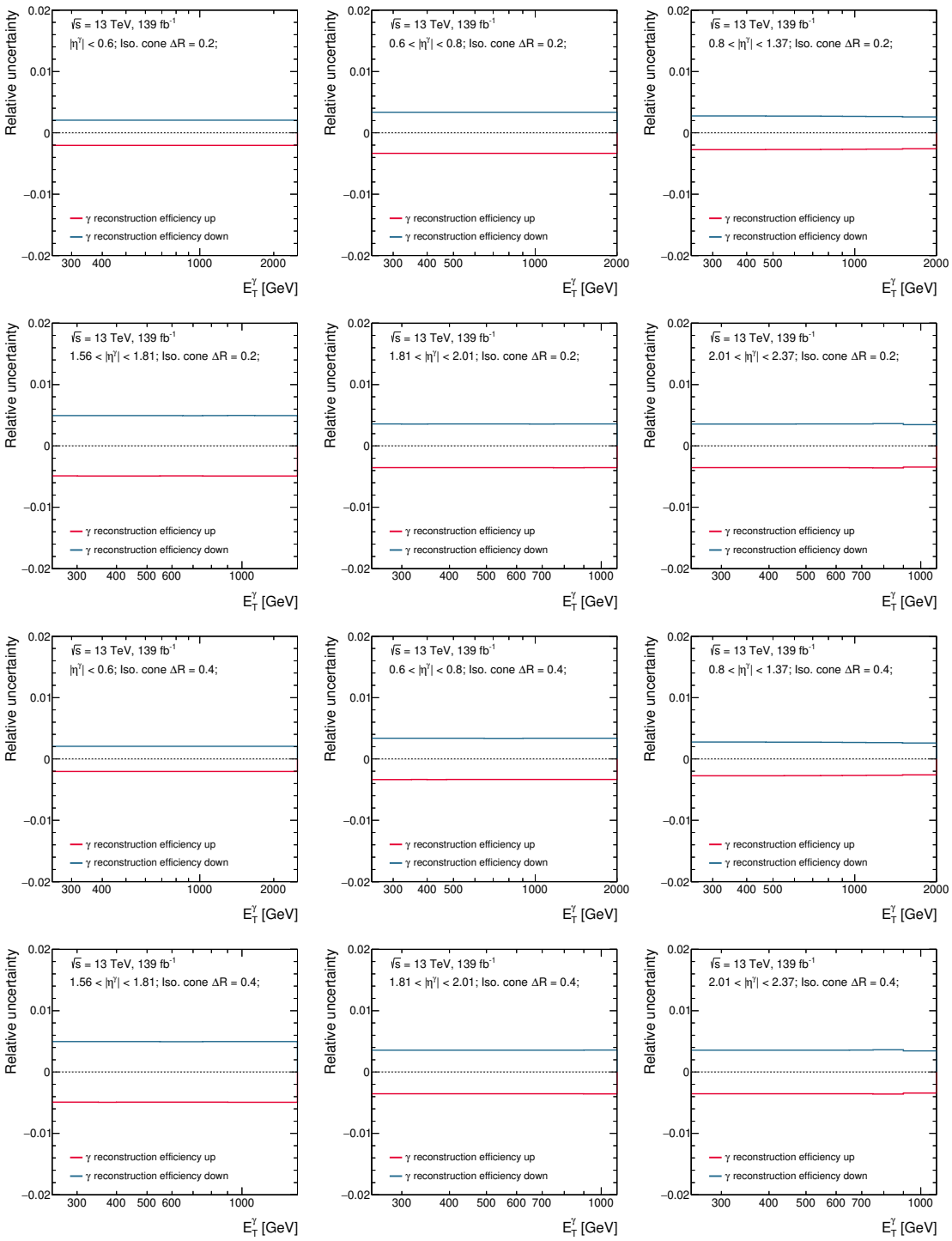


Figure 8.26: Systematic uncertainty in the measured cross section due to the uncertainty in the photon reconstruction efficiency as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. Results for the $R = 0.2$ ($R = 0.4$) sample are shown at the top (bottom).

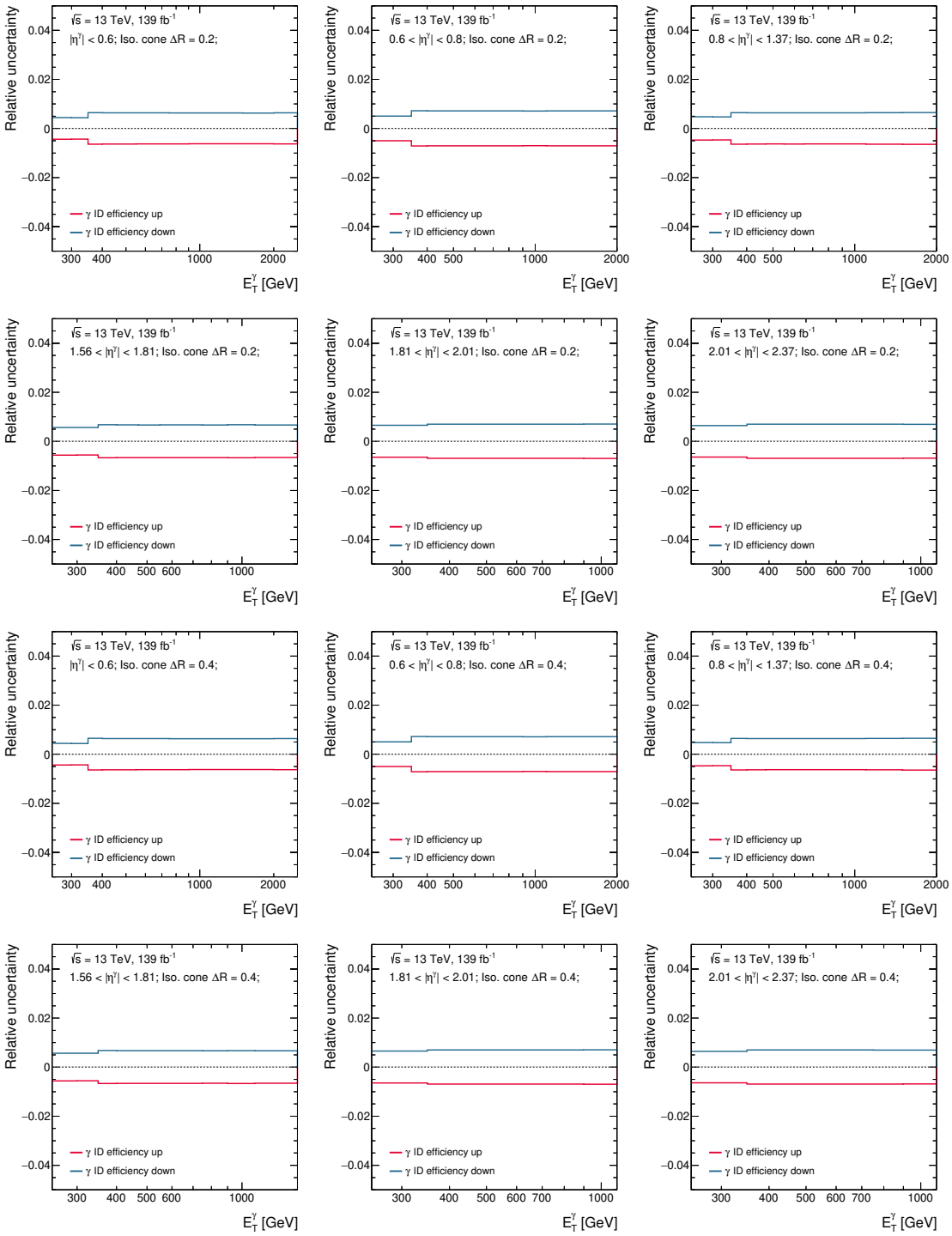


Figure 8.27: Systematic uncertainty in the measured cross section due to the uncertainty in the photon identification efficiency as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. Results for the $R = 0.2$ ($R = 0.4$) sample are shown at the top (bottom).

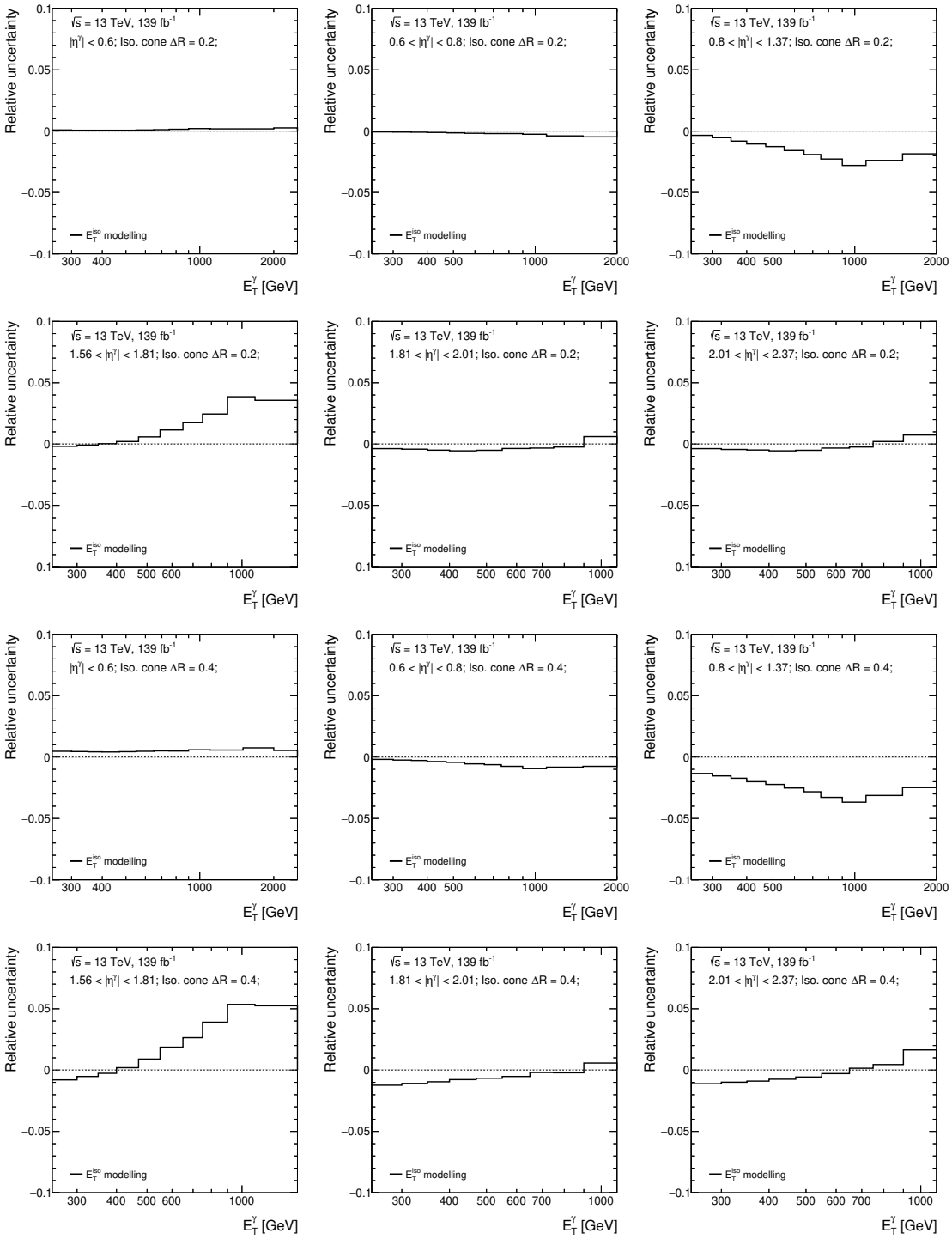


Figure 8.28: Systematic uncertainty in the measured cross section due to the effect of the isolation correction as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. Results for the $R = 0.2$ ($R = 0.4$) sample are shown at the top (bottom).

8.7.4 Unfolding procedure

QCD-cascade and hadronisation model dependence

The nominal unfolding correction matrices were computed using the SHERPA MC. As discussed in section 8.6, the effects on the unfolding procedure due to the matrix elements and the parton shower and hadronisation models employed in the generators are investigated. The difference in the cross section between the nominal result and that obtained using the PYTHIA simulation for the unfolding is taken as a systematic uncertainty. The uncertainty in the measured cross section is less than 2% and typically less than 1%. Figure 8.29 shows the resulting uncertainty in the measured cross section. Some differences are observed between the uncertainties obtained for the $R = 0.2$ and $R = 0.4$ samples.

Unfolding closure

A systematic uncertainty due to the unfolding closure (see appendix G) was included. Figure G.2 shows the resulting uncertainty on the measured cross section. The uncertainty is typically smaller than 0.1%, except in the tails of the most forward $|\eta^\gamma|$ region, where it increases up to 0.3%.

MC-sample statistics

The uncertainty in the cross section due to the statistical uncertainty of the MC samples is propagated into the measured cross section and typically amounts to less than 0.5%. Figure 8.30 shows the resulting uncertainty in the measured cross section. No difference is observed between the uncertainties for the $R = 0.2$ and $R = 0.4$ samples.

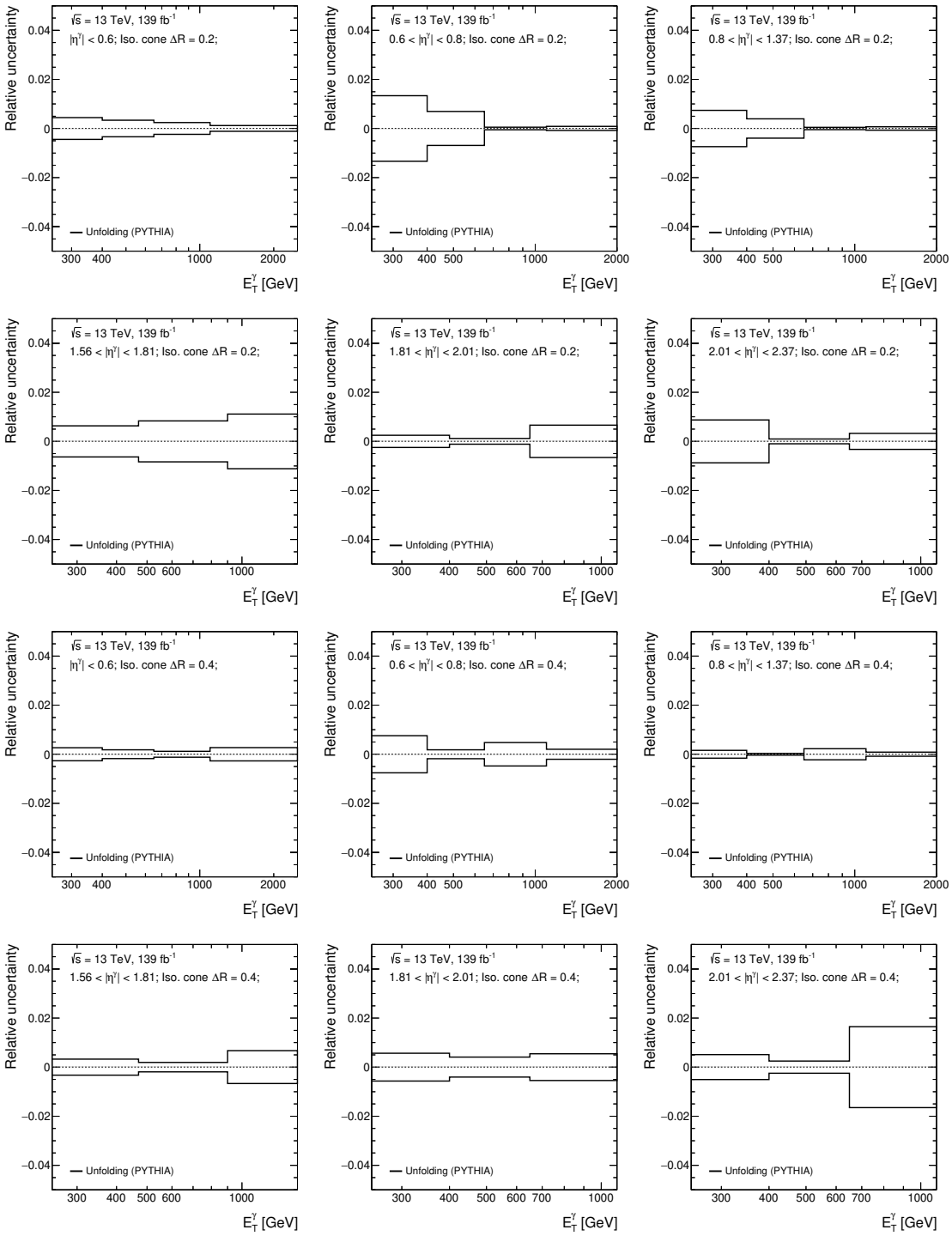


Figure 8.29: Systematic uncertainty in the measured cross section due to the QCD-cascade and hadronisation models as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. Statistical uncertainties in the estimation of this systematic uncertainty were taken care of by rebinning, merging those bins affected by relatively large statistical fluctuations with neighbouring bins. Results for the $R = 0.2$ ($R = 0.4$) sample are shown at the top (bottom).

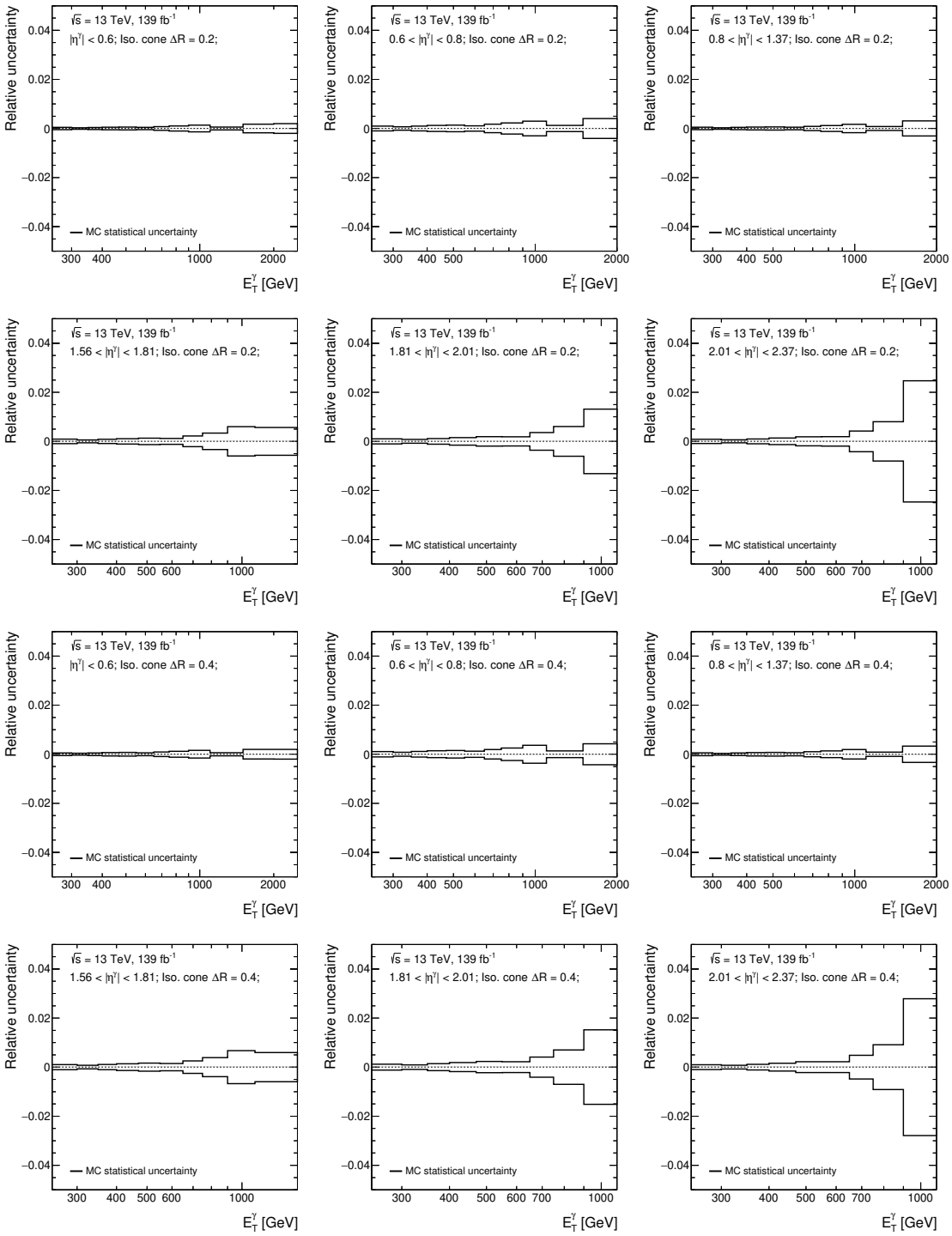


Figure 8.30: Statistical uncertainty of the MC samples as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. Results for the $R = 0.2$ ($R = 0.4$) sample are shown at the top (bottom).

8.7.5 Running conditions

Pile-up

A variation in the pile-up reweighting of simulated events is included to cover the uncertainty in the ratio of the predicted to measured inelastic cross sections [49]. To match the in-time and out-of-time pile-up conditions in the data, the distribution of $\langle\mu\rangle$ in simulated events was reweighted to that of the data after applying a factor 1/1.03 (see section 5.6). The uncertainty was estimated by changing the factor applied to the data to 1/0.94 or 1/1.12. The resulting uncertainty in the measured cross section is shown in figure 8.31, and is less than 1.5%. The uncertainty for $R = 0.2$ is substantially smaller than that for the $R = 0.4$ sample.

Trigger efficiency

The uncertainty in the trigger efficiency is estimated by propagating into the cross section the relative uncertainties shown in table 8.2 for each running period separately. The uncertainty in the measured cross section is less than 0.1%.

Integrated luminosity

The uncertainty in the integrated luminosity is 1.7%. It is fully correlated between all bins of the measured cross section.

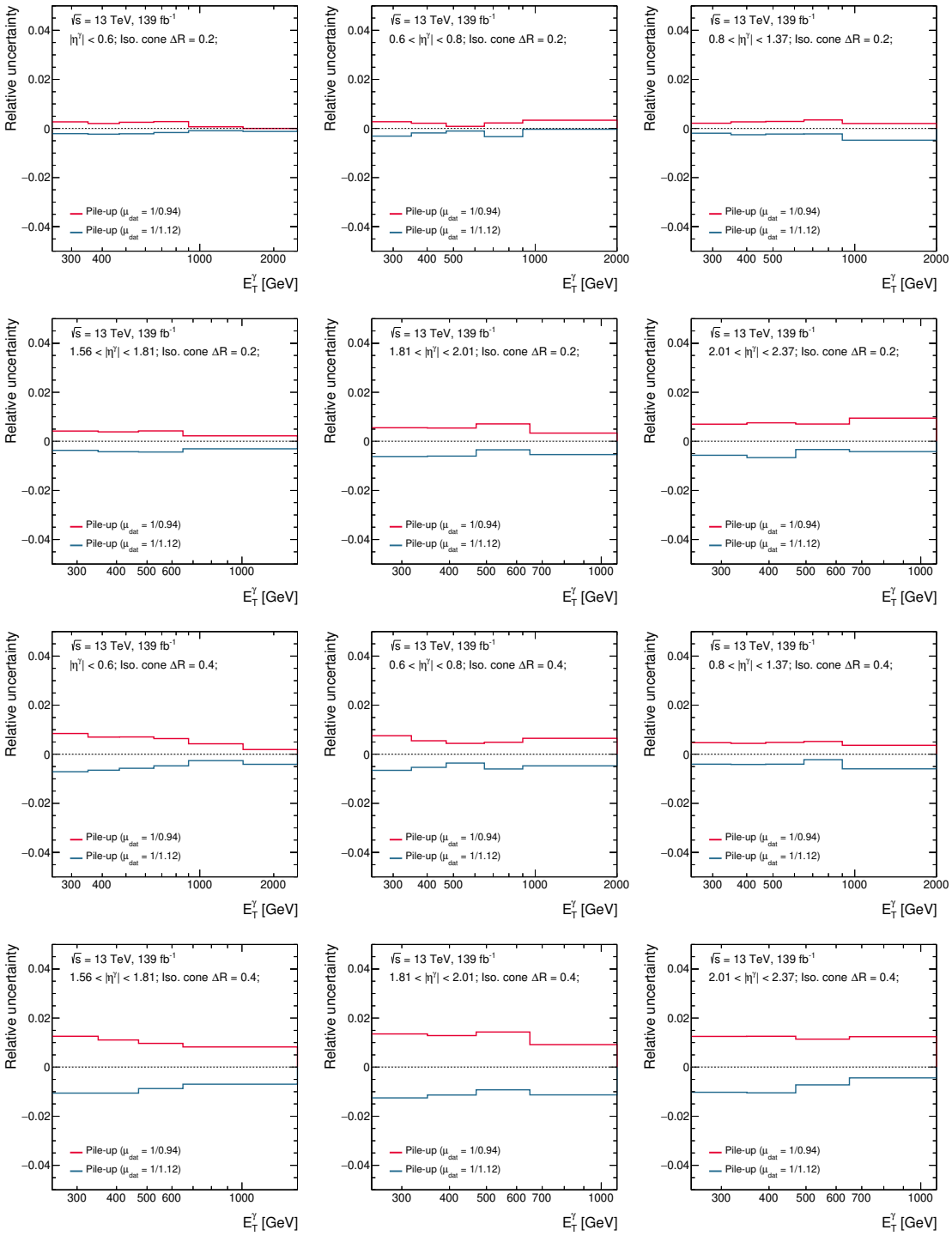


Figure 8.31: Systematic uncertainty in the measured cross section due to the pile-up reweighting as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. Statistical fluctuations in the estimation of this systematic uncertainty were taken care of by rebinning, merging those bins affected by relatively large statistical fluctuations with neighbouring bins. Results for the $R = 0.2$ ($R = 0.4$) sample are shown at the top (bottom).

8.7.6 Photon energy scale and resolution

As mentioned in section 7.7.6, an updated version of the model used to study the sources of uncertainty in the energy scale and resolution for photons with Run I data [155] was implemented for Run II in ref. [130]. A further update of the models implemented for Run II was published recently [131]. In this update, the amount of passive material located between the IP and the EM calorimeter was measured using the ratio of the energies deposited by electrons from Z-boson decays in the first and second layers of the EM calorimeter ($E_{1/2}$). The sensitivity of the calibrated energy to the detector material was re-evaluated to reflect the changes in the reconstruction introduced in ref. [131]. The systematic uncertainty due to the innermost pixel detector layer and the services of the pixel detector were also updated with respect to ref. [130], using a more accurate description of the systems in the simulation [131]. A total of 79 individual components influencing the energy scale and resolution of the photon are identified to asses the overall uncertainty in the energy measurement. The same procedure described in section 7.7.6 was followed to evaluate the systematic uncertainties in the measured cross sections. The resulting uncertainties are typically less than 0.5% for the energy resolution and in the range $2\% - 20\%$ for the energy scale. The uncertainty in the measured cross section is shown in figure 8.32 (8.33) for the energy resolution (scale). No difference is observed between the uncertainties for the $R = 0.2$ and $R = 0.4$ samples. Figure 8.34 shows the three main contributions to the photon energy scale in each $|\eta^{\gamma}|$ region.

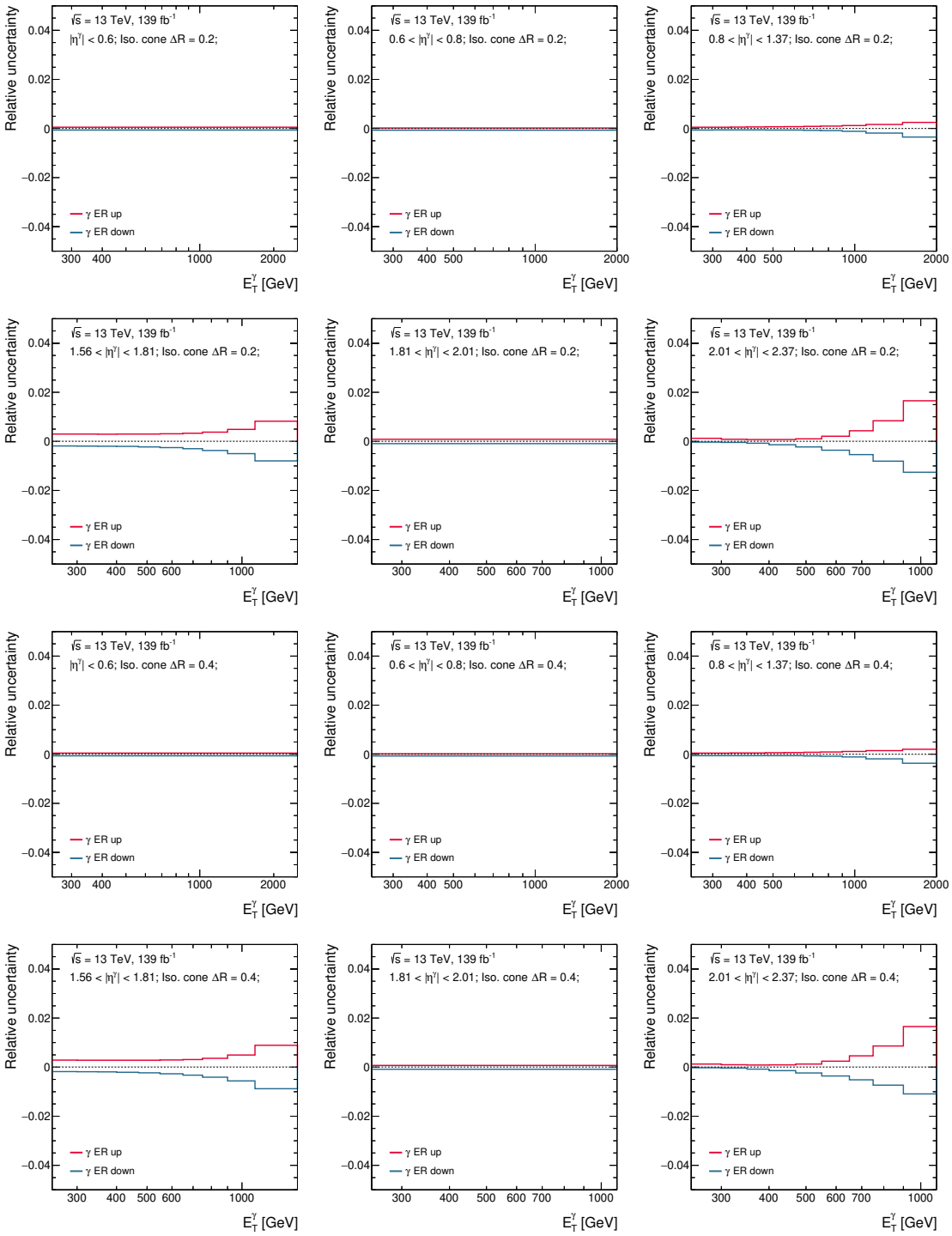


Figure 8.32: Systematic uncertainty in the measured cross section due to the photon energy resolution uncertainty as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. Statistical fluctuations in the estimation of this systematic uncertainty were taken care of by using a bootstrap technique on each individual component. Results for the $R = 0.2$ ($R = 0.4$) sample are shown at the top (bottom).

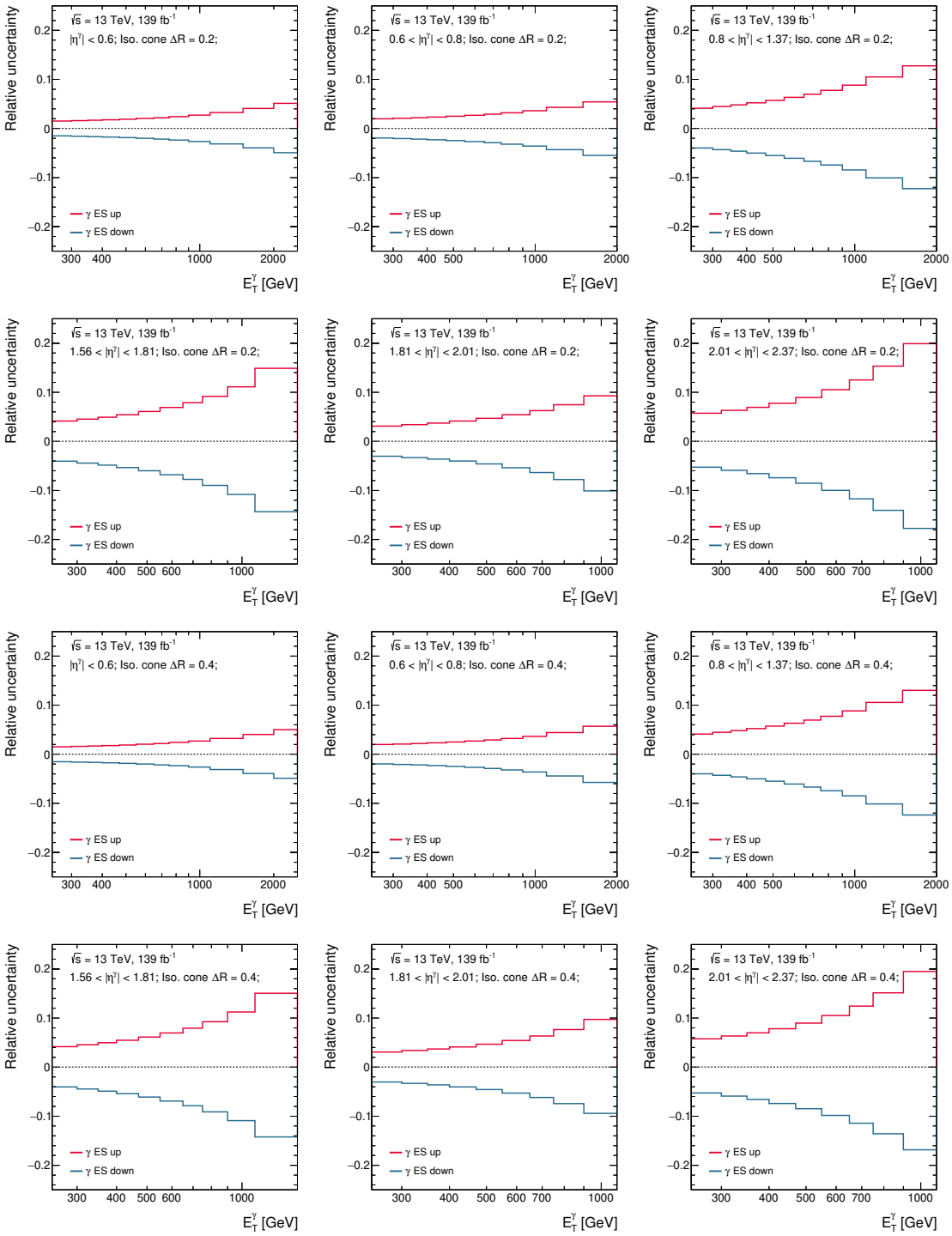


Figure 8.33: Systematic uncertainty in the measured cross section due to the photon energy scale uncertainty as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. Statistical fluctuations in the estimation of this systematic uncertainty were taken care of by using a bootstrap technique on each individual component. Results for the $R = 0.2$ ($R = 0.4$) sample are shown at the top (bottom).

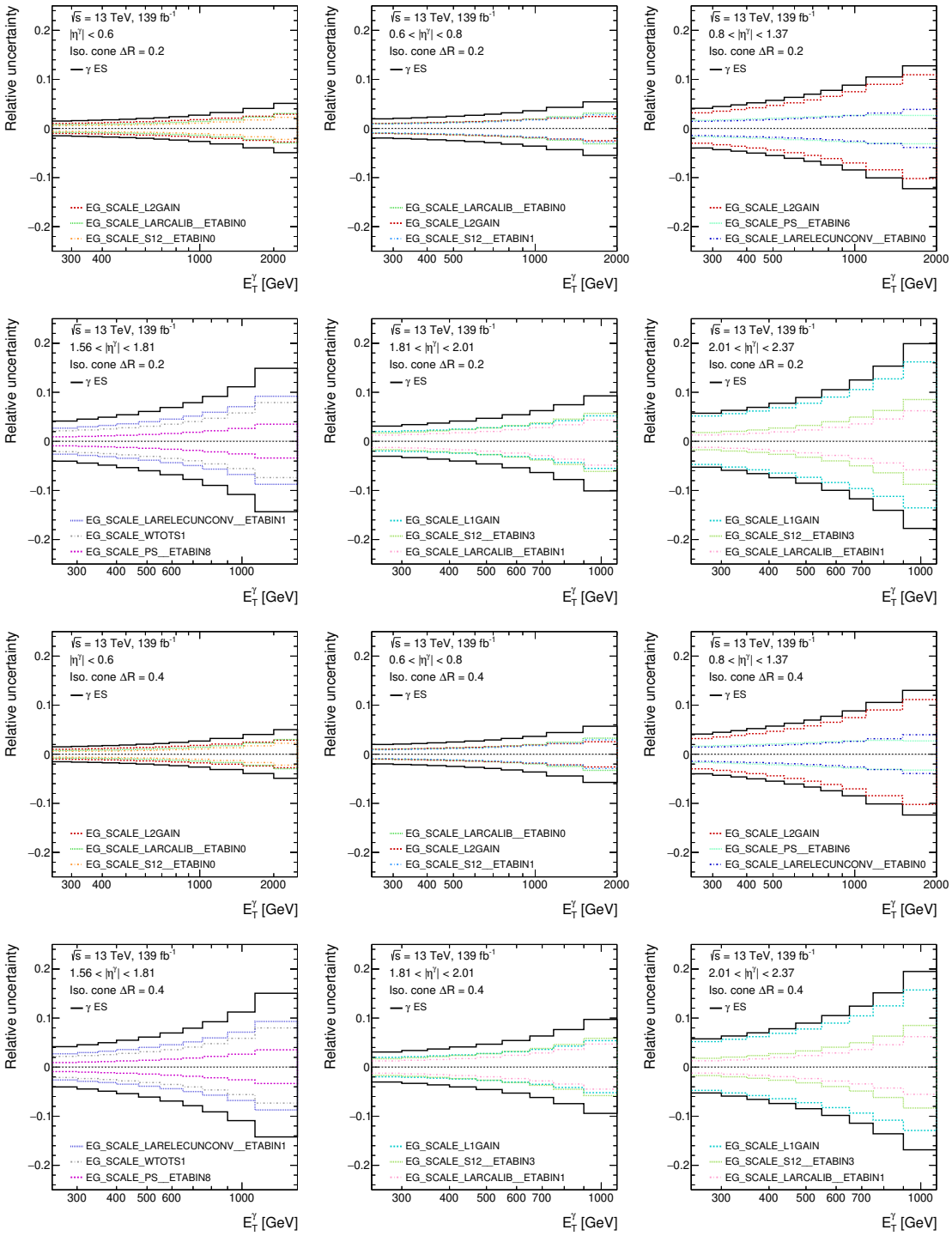


Figure 8.34: Systematic uncertainty in the measured cross section due to the photon energy scale uncertainty as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. The total uncertainty is represented with the black line; the three main contributions in each $|\eta^{\gamma}|$ region are also shown. Results for the $R = 0.2$ ($R = 0.4$) sample are shown at the top (bottom).

8.7.7 Total systematic uncertainty

The total systematic uncertainty is calculated by adding in quadrature the uncertainties listed above. The total systematic is in the range $3\% - 20\%$, depending on E_{T}^{γ} and the $|\eta^{\gamma}|$ region. The dominant sources of uncertainty arise from the photon energy scale, the integrated luminosity and the R^{bg} correlation (pile-up) for $R = 0.2$ ($R = 0.4$). Figure 8.35 shows the total systematic uncertainty for each measured cross section together with the dominant components. The systematic uncertainty dominates the total experimental uncertainty for $E_{\mathrm{T}}^{\gamma} \lesssim 1.5$ TeV for $|\eta^{\gamma}| < 0.6$ and $0.8 < |\eta^{\gamma}| < 1.37$, $E_{\mathrm{T}}^{\gamma} \lesssim 1.1$ TeV for $0.6 < |\eta^{\gamma}| < 0.8$ and $1.56 < |\eta^{\gamma}| < 1.81$, and $E_{\mathrm{T}}^{\gamma} \lesssim 900$ GeV for $1.81 < |\eta^{\gamma}| < 2.37$, whereas for higher E_{T}^{γ} values, the statistical uncertainty of the data limits the precision of the measurement, as can be seen in figure 8.36. This represents an improvement relative to previous measurements [37], where the region dominated by the systematic uncertainty was $E_{\mathrm{T}}^{\gamma} \lesssim 1.1$ TeV for $|\eta^{\gamma}| < 1.37$ and $E_{\mathrm{T}}^{\gamma} \lesssim 900$ GeV ($E_{\mathrm{T}}^{\gamma} \lesssim 750$ GeV) for $1.56 < |\eta^{\gamma}| < 1.81$ ($1.81 < |\eta^{\gamma}| < 2.37$).

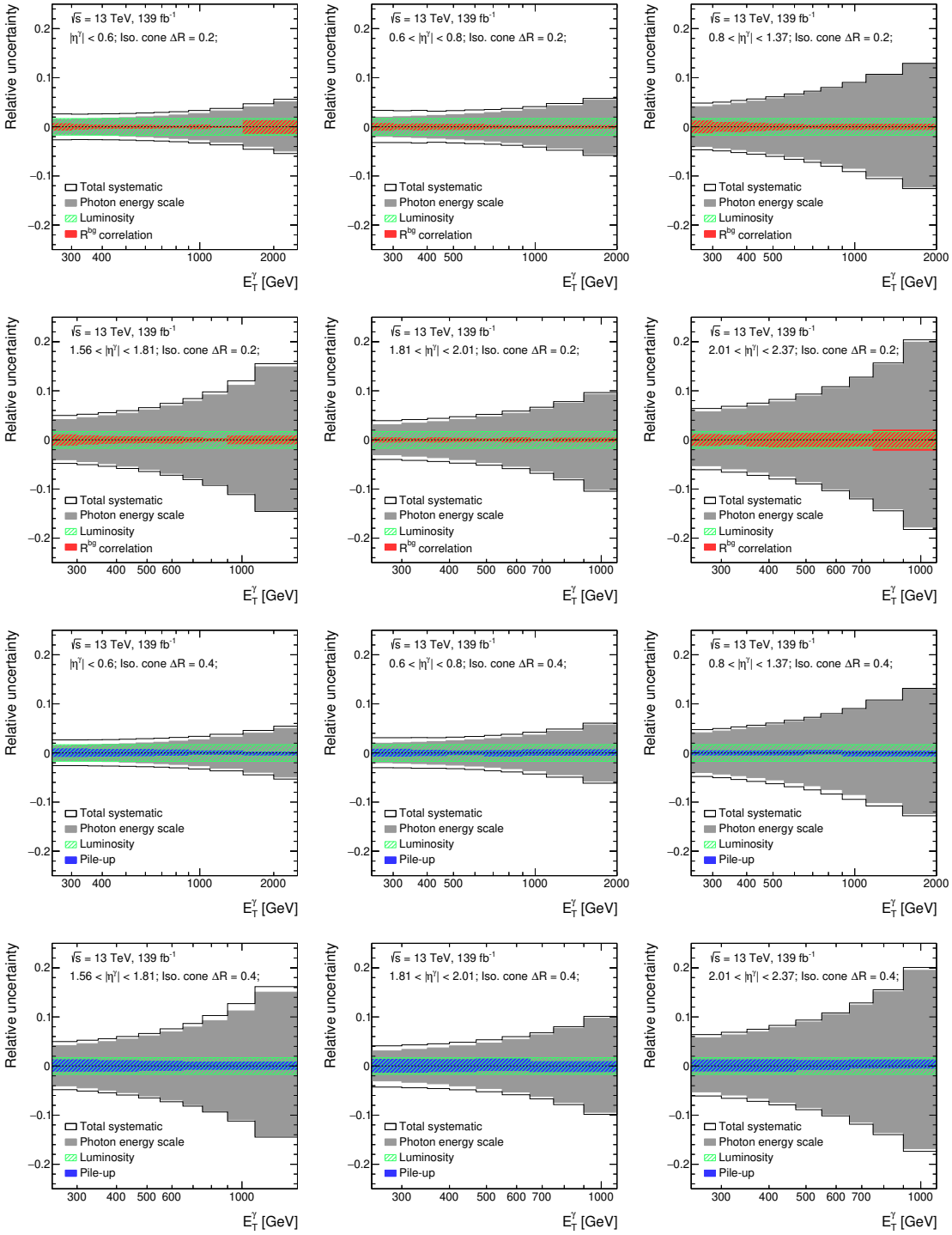


Figure 8.35: The relative total systematic uncertainty in the cross section (white areas) as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. The relative uncertainty due to the photon energy scale (grey areas), the relative uncertainty due to that on the integrated luminosity (green hatched areas), and the relative uncertainty due to the R^{bg} correlation (red areas) or pile-up (blue areas) for $R = 0.2$ or $R = 0.4$, respectively, are also shown. Results for the $R = 0.2$ ($R = 0.4$) sample are shown at the top (bottom).

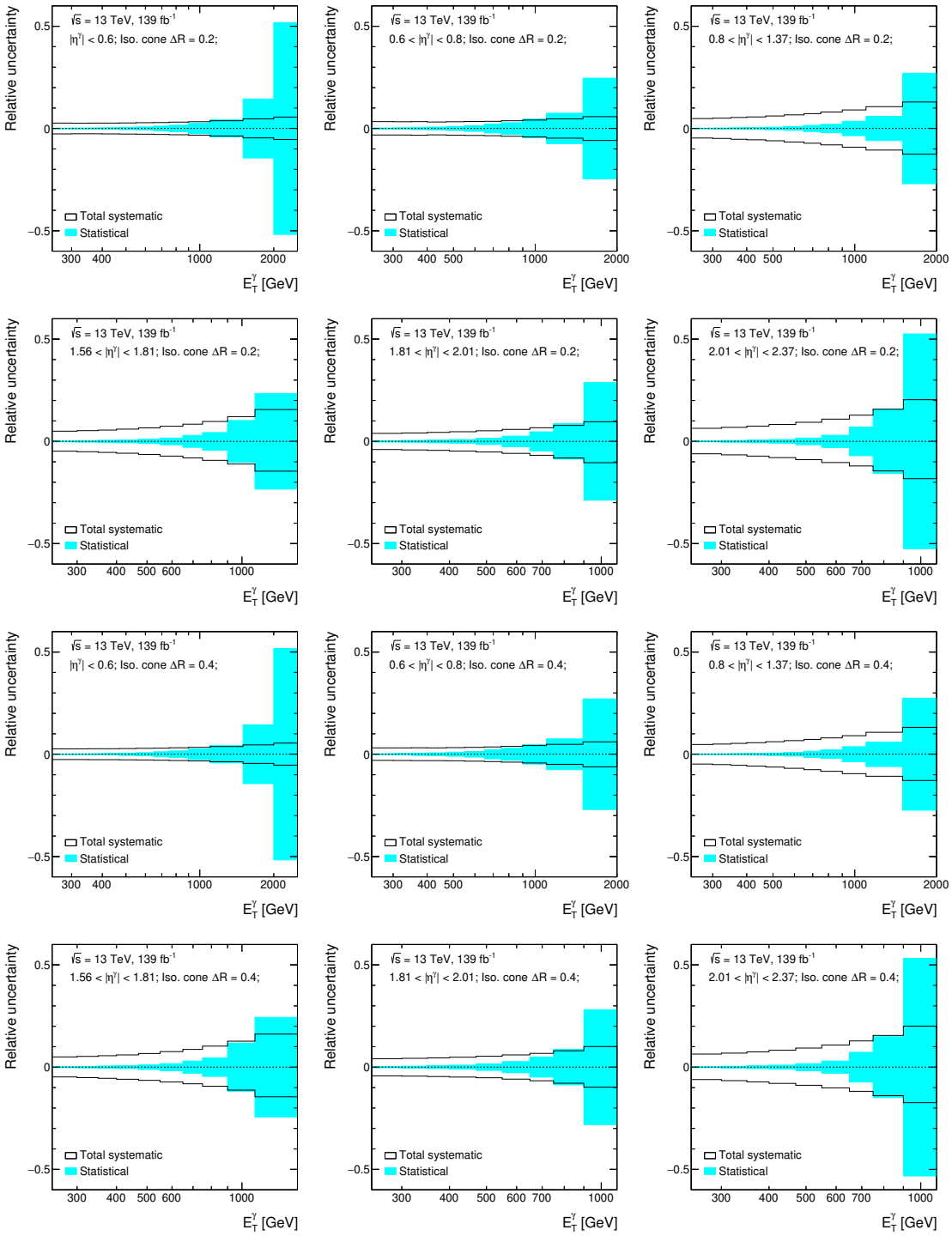


Figure 8.36: The relative total systematic uncertainty in the cross section (white areas) as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. The relative statistical uncertainty from the data is also shown (cyan areas). Results for the $R = 0.2$ ($R = 0.4$) sample are shown at the top (bottom).

8.8 Theory predictions

The NLO pQCD predictions were calculated using two programs, namely JETPHOX 1.3.1_2, which is described in section 3.4.2.1, and SHERPA 2.2.2, which is introduced in section 4.3. The main properties of the calculations using JETPHOX and SHERPA 2.2.2 were summarised in table 7.3.

8.8.1 Hadronisation and UE corrections to the NLO QCD calculations

Since the NLO pQCD prediction from JETPHOX is at parton level while the measurement is at the particle level, correction factors for the NP effects of hadronisation and UE are needed for the prediction to match the data. The calculation of the corrections is currently being performed using a similar approach as described in section 7.8.1, by computing the ratio of the particle-level cross section for a PYTHIA sample with UE effects to the parton-level cross section without UE effects. As no change with respect to the previous analysis is expected and the actual corrections are yet not available, a 1% uncertainty, as estimated in the previous analysis (see section 7.8.1), has been included.

8.8.2 Theoretical uncertainties in the NLO pQCD calculations

As has been shown in section 7.8.2 for the previous inclusive photon analysis, the major source of uncertainty affecting the predictions is due to terms beyond NLO in QCD. This uncertainty was evaluated following the same procedure which is described in section 7.8.2 for both JETPHOX and SHERPA 2.2.2 NLO pQCD calculations. The resulting uncertainty in the prediction is in the range $10\% - 19\%$ for JETPHOX and $20\% - 24\%$ SHERPA 2.2.2 (except for the highest measured E_{T}^{γ} point in the region $2.01 < |\eta^{\gamma}| < 2.37$, where it increases up to $\approx 31\%$). The fact that the uncertainties are larger for SHERPA 2.2.2 than for JETPHOX is attributed to multijet configurations which are accounted for only in the case of SHERPA 2.2.2 [37].

The uncertainty in the predictions of JETPHOX and SHERPA 2.2.2 due to the uncertainty in the proton PDF was estimated following the same methodology described in section 7.8.2. The resulting uncertainty in the prediction is in the range $1\% - 6\%$ for JETPHOX and $1\% - 8\%$ for SHERPA 2.2.2 (except for the highest measured E_{T}^{γ} point in the region $2.01 < |\eta^{\gamma}| < 2.37$, where it increases up to $\approx 15\%$), depending on E_{T}^{γ} and the $|\eta^{\gamma}|$ region.

The uncertainty in the prediction of JETPHOX (SHERPA 2.2.2) due to the uncertainty in α_{S} is evaluated by repeating the calculations using two additional sets of proton PDF from the MMHT2014 (NNPDF3.0) analysis for which different values of $\alpha_{\mathrm{S}}\,(M_{\mathrm{Z}}^2)$ are assumed in the fits, namely 0.118 (0.117) and 0.122 (0.119); then, the same approach

described in section 7.8.2 is followed. The resulting uncertainty in the prediction of JETPHOX (SHERPA 2.2.2) is less than 3% (7%).

In addition, an uncertainty of 1% was included due to the hadronisation corrections in JETPHOX, as mentioned in section 8.8.1. The total theoretical uncertainty was obtained by adding in quadrature the individual uncertainties listed above (see figures 8.37, 8.38, and 8.39). The dominant theoretical uncertainty for the predictions using JETPHOX or SHERPA 2.2.2 is that arising from the variation of the scales. The uncertainty arising from the PDF ($\alpha_{\mathrm{S}}(M_{\mathrm{Z}}^2)$) is the second dominant uncertainty for the prediction of JETPHOX (SHERPA 2.2.2). The comparisons between $R = 0.2$ and $R = 0.4$ show that the uncertainty due to the terms beyond NLO is slightly larger for the former, while the differences in the uncertainties due to the proton PDF and due to the uncertainty in α_{S} are negligible.

A prediction from JETPHOX using as scales $\mu_{\mathrm{R}} = \mu_{\mathrm{F}} = \mu_{\mathrm{f}} = E_{\mathrm{T}}^{\gamma}/2$ (see figure 8.38) was also made and the associated theoretical uncertainties are calculated using the same procedure as above, including variations of the scales by the factors 0.5 and 2 with respect to $E_{\mathrm{T}}^{\gamma}/2$. The comparison of the uncertainty due to the scale variations for this prediction and the one from JETPHOX using $\mu_{\mathrm{R}} = \mu_{\mathrm{F}} = \mu_{\mathrm{f}} = E_{\mathrm{T}}^{\gamma}$ shows small differences in the highest-E_{T}^{γ} bins. The differences observed in the uncertainties due to the proton PDF and due to the uncertainty in α_{S} are negligible.

The sensitivity of the observables to the proton PDFs was investigated by comparing the prediction of JETPHOX based on MMHT2014 with alternative predictions based on the CT14, CT18, NNPDF3.1, and HERAPDF2.0 PDF sets. Figures 8.40 and 8.41 show the relative differences between the calculations of JETPHOX using the different PDF sets for $\mu_{\mathrm{R}} = \mu_{\mathrm{F}} = \mu_{\mathrm{f}} = E_{\mathrm{T}}^{\gamma}$ and $\mu_{\mathrm{R}} = \mu_{\mathrm{F}} = \mu_{\mathrm{f}} = E_{\mathrm{T}}^{\gamma}/2$, respectively. In addition, the predictions for both $R = 0.2$ and $R = 0.4$ are included.

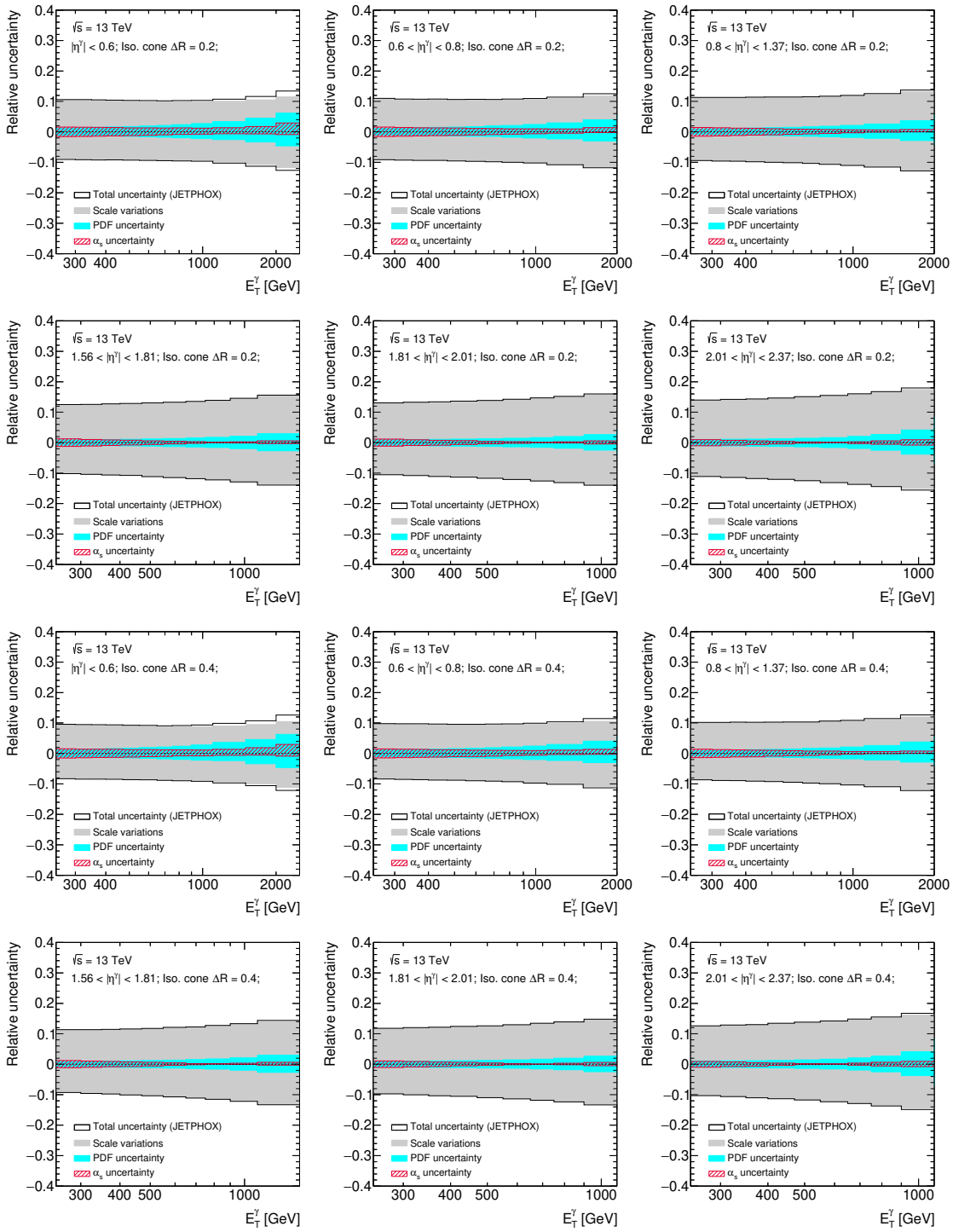


Figure 8.37: The relative total theoretical uncertainty in the cross-section prediction of JETPHOX using $\mu_{\mathrm{R}} = \mu_{\mathrm{F}} = \mu_{\mathrm{f}} = E_{\mathrm{T}}^{\gamma}$ as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. The uncertainty from the scale variations, the uncertainty from the PDF and the uncertainty from the value of α_{S} are also included. Results for the $R = 0.2$ ($R = 0.4$) sample are shown at the top (bottom).

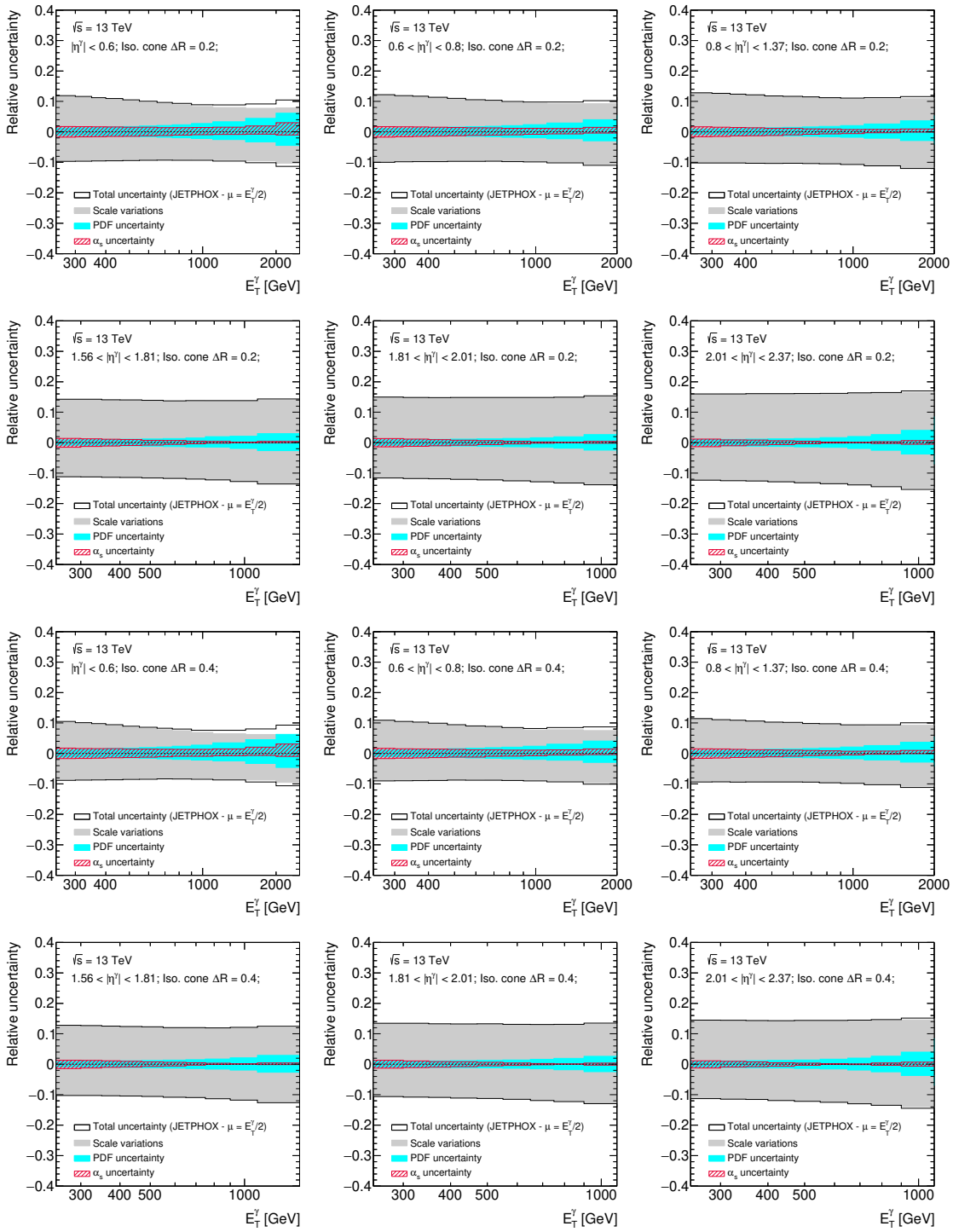


Figure 8.38: The relative total theoretical uncertainty in the cross-section prediction of JETPHOX using $\mu_{\mathrm{R}} = \mu_{\mathrm{F}} = \mu_{\mathrm{f}} = E_{\mathrm{T}}^{\gamma}/2$ as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. The uncertainty from the scale variations, the uncertainty from the PDF and the uncertainty from the value of α_{S} are also included. Results for the $R = 0.2$ ($R = 0.4$) sample are shown at the top (bottom).

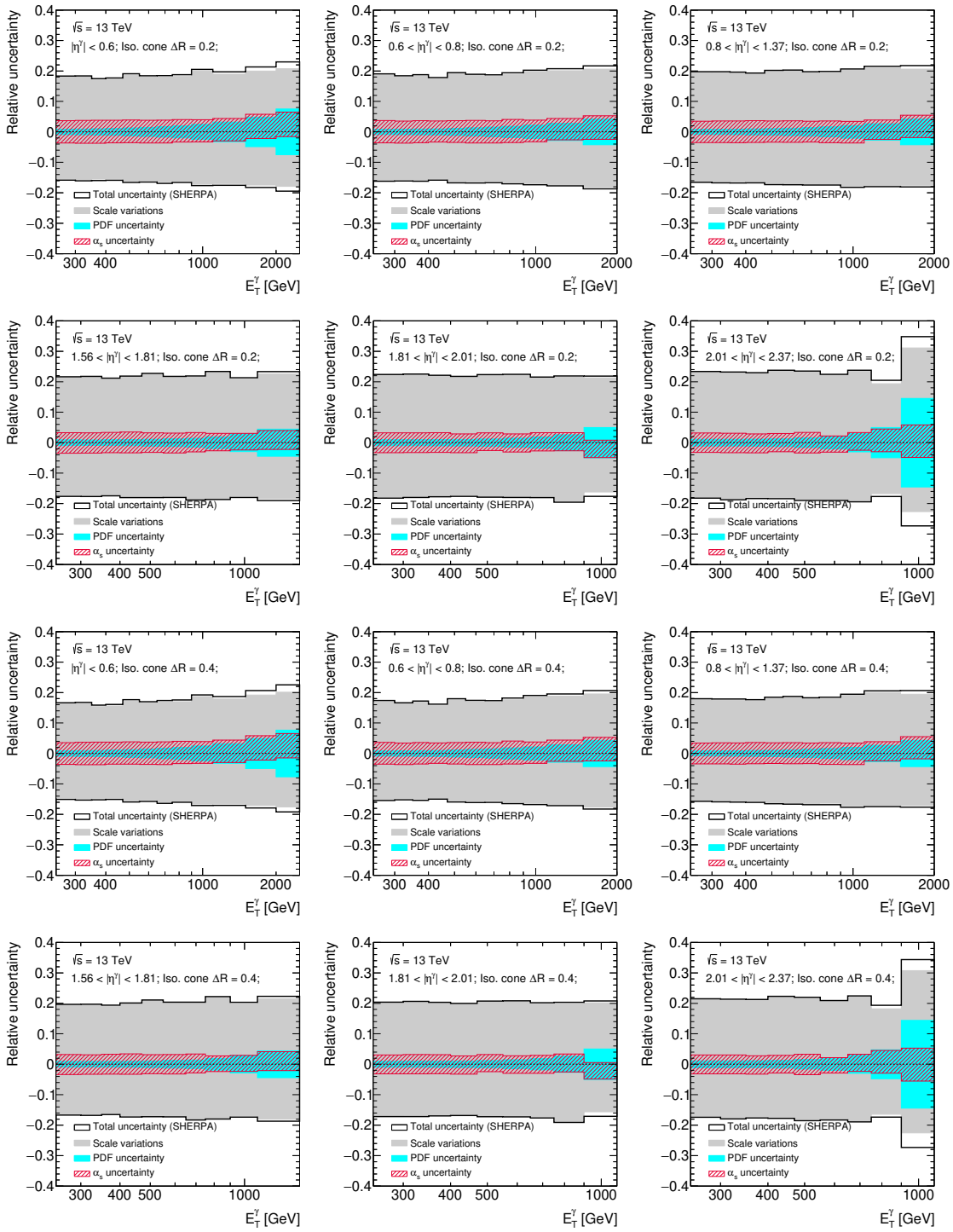


Figure 8.39: The relative total theoretical uncertainty in the cross-section prediction of SHERPA 2.2.2 using $\mu_{\mathrm{R}} = \mu_{\mathrm{F}} = \mu_{\mathrm{f}} = E_{\mathrm{T}}^{\gamma}$ as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. The uncertainty from the scale variations, the uncertainty from the PDF and the uncertainty from the value of α_{S} are also included. Results for the $R = 0.2$ ($R = 0.4$) sample are shown at the top (bottom).

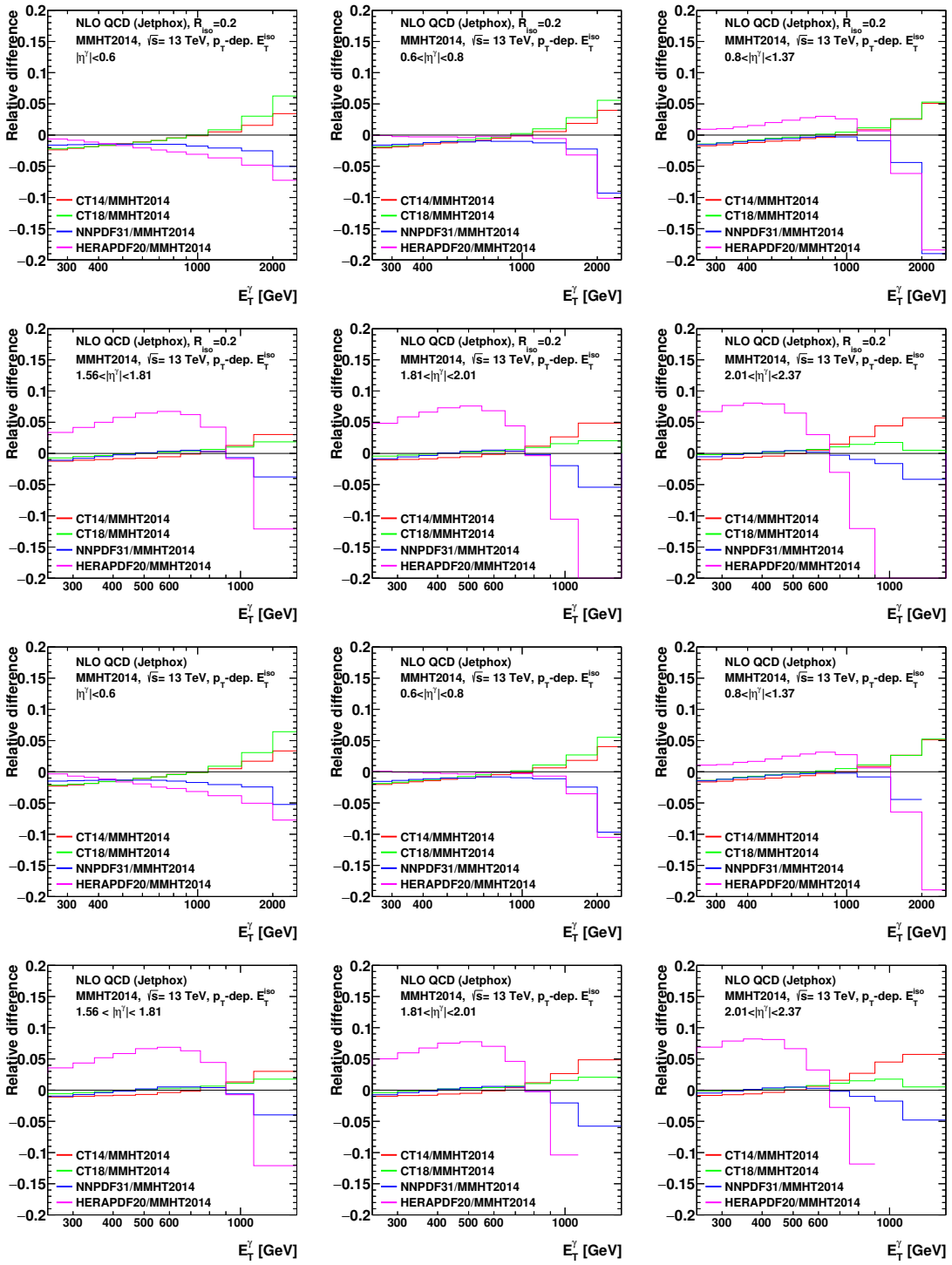


Figure 8.40: Relative difference between the JETPHOX predictions based on the CT14 (red lines), CT18 (green lines), NNPDF3.1 (blue lines), or HERAPDF2.0 (magenta lines) PDF sets and those based on the MMHT2014 PDF using $\mu_{\mathrm{R}} = \mu_{\mathrm{F}} = \mu_{\mathrm{f}} = E_{\mathrm{T}}^{\gamma}$ as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. Results for the $R = 0.2$ ($R = 0.4$) sample are shown at the top (bottom).

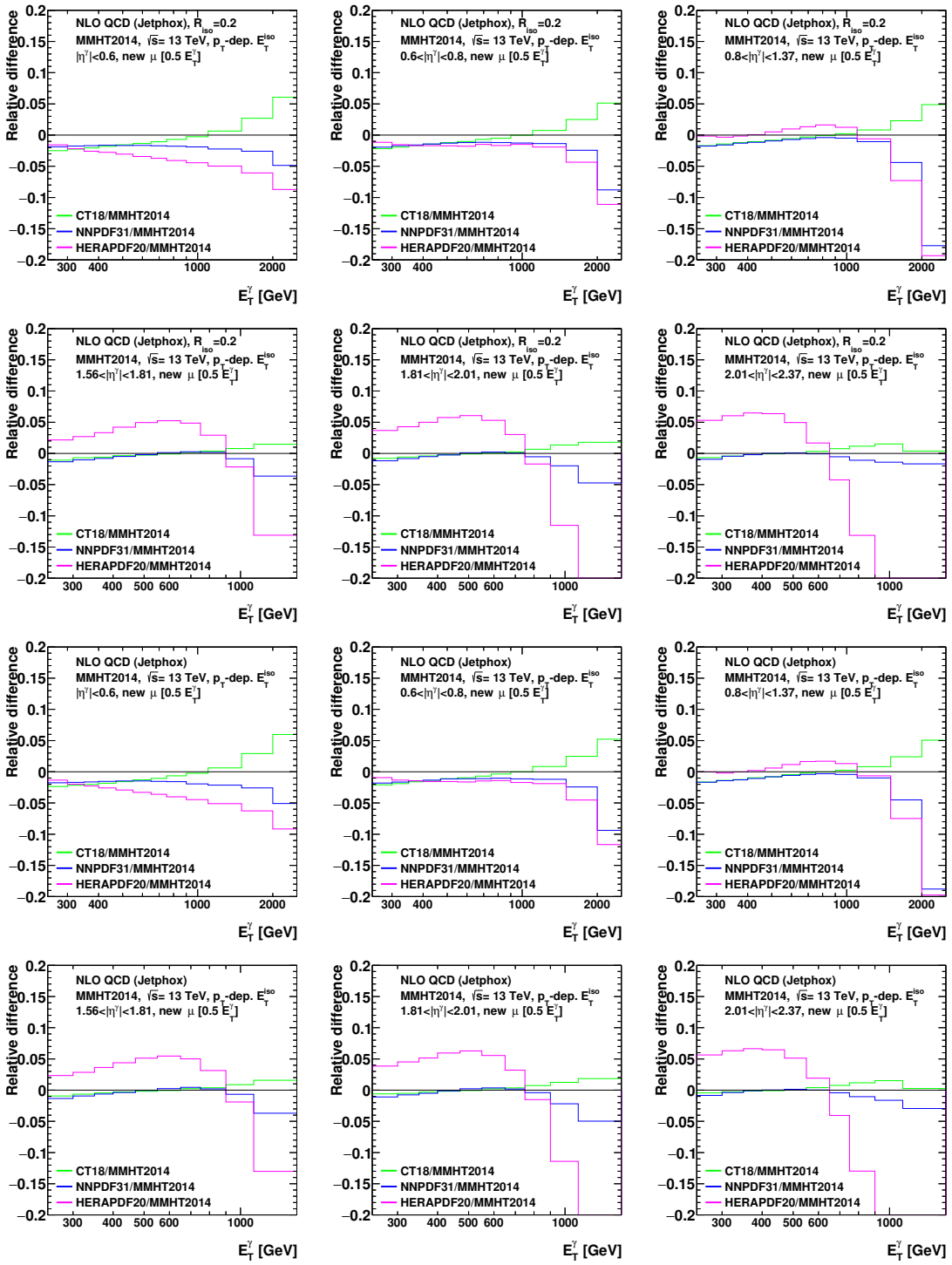


Figure 8.41: Relative difference between the JETPHOX predictions based on the CT18 (green lines), NNPDF3.1 (blue lines), or HERAPDF2.0 (magenta lines) PDF sets and those based on the MMHT2014 PDF using $\mu_{\mathrm{R}} = \mu_{\mathrm{F}} = \mu_{\mathrm{f}} = E_{\mathrm{T}}^{\gamma}/2$ as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. Results for the $R = 0.2$ ($R = 0.4$) sample are shown at the top (bottom).

8.9 Results

The measurement of the inclusive isolated-photon differential cross section as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$ is shown in figure 8.42. In the region $|\eta^{\gamma}| < 0.6$ ($0.6 < |\eta^{\gamma}| < 0.8$ and $0.8 < |\eta^{\gamma}| < 1.37$) the cross section decreases by six (five) orders of magnitude in the measured range $250 < E_{\mathrm{T}}^{\gamma} < 2500$ GeV ($250 < E_{\mathrm{T}}^{\gamma} < 2000$ GeV). In the forward region $1.56 < |\eta^{\gamma}| < 1.81$, the measured range is $250 < E_{\mathrm{T}}^{\gamma} < 1500$ GeV, and the cross section spans approximately five orders of magnitude. Finally, in the most forward regions $1.81 < |\eta^{\gamma}| < 2.01$ and $2.01 < |\eta^{\gamma}| < 2.37$, the measured range is $250 < E_{\mathrm{T}}^{\gamma} < 1100$ GeV, and the cross section decreases approximately four and five orders of magnitude, respectively. Values of E_{T}^{γ} up to 2.5 TeV are measured with the current integrated luminosity. The shape of the measured cross section in the different $|\eta^{\gamma}|$ regions is similar, as well as the shape of the cross section for the two different isolation cone choices, though the normalisation of the measurement for $R = 0.2$ is somewhat higher than for $R = 0.4$.

The predictions of the leading-logarithm parton-shower MC models of PYTHIA and SHERPA are compared to the measurement in figure 8.42. These predictions were normalised to the measured integrated cross section in each $|\eta^{\gamma}|$ region. The difference in normalisation between the data and PYTHIA (SHERPA) is $\approx 10\% - 20\%$ ($\approx 40\%$). The calculations of PYTHIA give a good description of the shape of the data for $E_{\mathrm{T}}^{\gamma} \lesssim 500$ GeV and $|\eta^{\gamma}| < 1.37$; the description of PYTHIA is good in the whole measured range for $1.56 < |\eta^{\gamma}| < 2.01$; in the most forward $|\eta^{\gamma}|$ region, PYTHIA gives a good description of the shape of the data for $E_{\mathrm{T}}^{\gamma} \lesssim 600$ GeV. The predictions of SHERPA also give a good description of the shape of the data in the same phase-space regions.

The NLO pQCD predictions of JETPHOX and SHERPA 2.2.2 are compared with the measurement in figures 8.43 and 8.44. Figure 8.43 (8.44) shows the inclusive isolated-photon differential cross section as a function E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$ for $R = 0.2$ ($R = 0.4$). The nominal prediction of SHERPA 2.2.2 is above that of JETPHOX; this is attributed to the fact that the former includes contributions from parton-showers, virtual corrections for γ+2-jet and higher-order tree-level matrix elements for the processes $2 \to n$ with $n = 4$ and 5, which are not present in the prediction of JETPHOX [37]. Both types of predictions adequately reproduce the dependence of the differential cross section on E_{T}^{γ} as observed in the data.

Although the predictions are overall consistent with the measurement within uncertainties, a detailed comparison is presented in terms of the ratios of theory predictions to data. The ratios of the predictions of JETPHOX with $\mu_{\mathrm{R}} = \mu_{\mathrm{F}} = \mu_{\mathrm{f}} = E_{\mathrm{T}}^{\gamma}$ using different PDF sets to the measured cross section as a function of E_{T}^{γ} are shown in figure 8.45 (8.46) for $R = 0.2$ ($R = 0.4$); the full theoretical uncertainty, including the variations of the scales, is shown for the predictions based on the MMHT2014 PDF set. The predictions based on the MMHT2014, CT18 and NNPDF3.1 PDF sets are similar and closest to the data for $|\eta^{\gamma}| < 1.37$ for most of the range in E_{T}^{γ}. For $1.56 < |\eta^{\gamma}| < 2.37$, the prediction based on the HERAPDF2.0 PDF set is closest to the data. The prediction of JETPHOX

with $\mu_R = \mu_F = \mu_f = E_T^\gamma$ exhibits a tendency to underestimate the data over most of the measured range in E_T^γ. However, the prediction of JETPHOX with a different choice for the scales, namely $\mu_R = \mu_F = \mu_f = E_T^\gamma/2$, provides an improved description of the normalisation of the data, as can be seen in figures 8.47 and 8.48 for $R = 0.2$ and $R = 0.4$, respectively; the full theoretical uncertainty, including the variations of the scales with respect to $E_T^\gamma/2$, is shown for the prediction based on the MMHT2014 PDF set. The ratio of the prediction of SHERPA 2.2.2 to the measured cross section as a function of E_T^γ is also shown in figures 8.49 and 8.50 for $R = 0.2$ and $R = 0.4$, respectively. Good agreement is observed between this prediction and the data distribution. In general, the NLO pQCD predictions are consistent with the measurement within the experimental and theoretical uncertainties.

The dependence of the inclusive-photon cross section with the isolation cone radius R was further investigated by measuring the fiducial cross section in each $|\eta^\gamma|$ region, divided by the width of each $|\eta^\gamma|$ region, for each of the two R measured values. Figures 8.51 to 8.53 show the measured fiducial cross section as a function of R in the different regions of $|\eta^\gamma|$. The measured cross section decreases with increasing R in all $|\eta^\gamma|$ regions. In addition, the measured fiducial cross section decreases with increasing $|\eta^\gamma|$ for a fixed value of R, being approximately constant for $|\eta^\gamma| < 1.37$. The NLO pQCD predictions of JETPHOX with $\mu_R = \mu_F = \mu_f = E_T^\gamma$ using different PDF sets are compared with the measurement in figure 8.51. All the predictions are similar to the data for $|\eta^\gamma| < 1.37$. For $1.56 < |\eta^\gamma| < 2.37$, the prediction based on the HERAPDF2.0 PDF set is closest to the data. The prediction of JETPHOX with $\mu_R = \mu_F = \mu_f = E_T^\gamma$ shows a tendency to underestimate the data. However, the prediction with the choice of scales $\mu_R = \mu_F = \mu_f = E_T^\gamma/2$ provides an improved description of the normalisation of the data for both radii, as can be seen in figure 8.52 and 8.53. The NLO pQCD prediction of SHERPA 2.2.2 is also compared with the measurement in figures 8.51 to 8.53. This prediction describes within theoretical and experimental uncertainties the dependence with R of the measured fiducial cross section. The prediction of SHERPA 2.2.2 tends to be above the data for $R = 0.2$ and $|\eta^\gamma| < 1.37$; the predictions for both isolation radii describe the data well in the region $1.56 < |\eta^\gamma| < 1.81$, however, the data tend to be overestimated by both predictions in the most forward regions, for $|\eta^\gamma| > 1.81$.

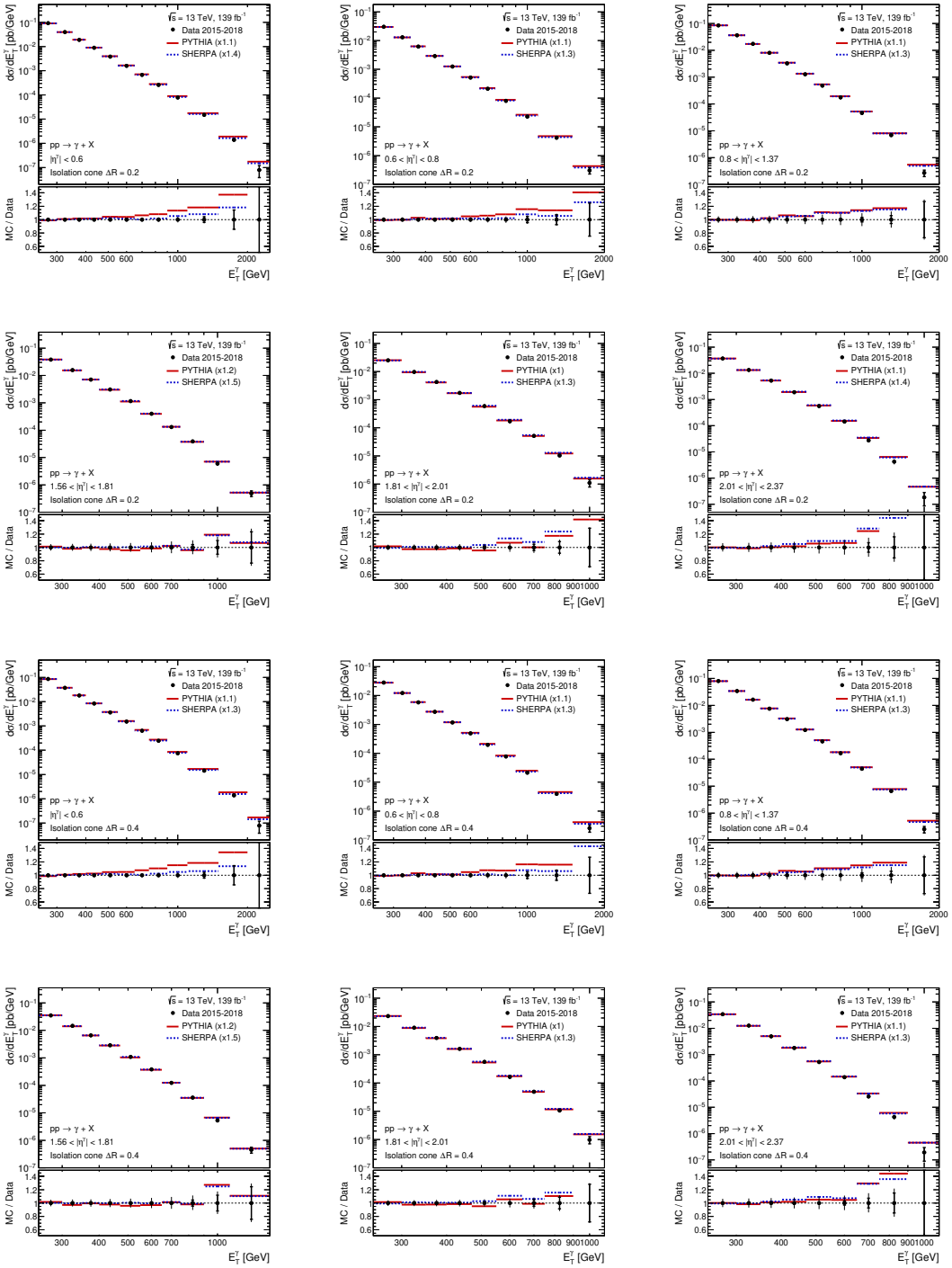


Figure 8.42: Measured cross section for isolated-photon production (dots) as a function of E_{T}^{γ} in the different $|\eta^{\gamma}|$ regions. The predictions from PYTHIA (solid lines) and SHERPA (dashed lines) are also shown; these predictions are normalised to the measured integrated cross section in each region of $|\eta^{\gamma}|$ using the values indicated in parentheses. The bottom part of each figure shows the ratio of the MC predictions to the measured cross section. The inner (outer) error bars represent the statistical uncertainties (the statistical and systematic uncertainties added in quadrature). For most of the points, the inner error bars are smaller than the marker size and, thus, not visible. Results for the $R = 0.2$ ($R = 0.4$) sample are shown at the top (bottom).

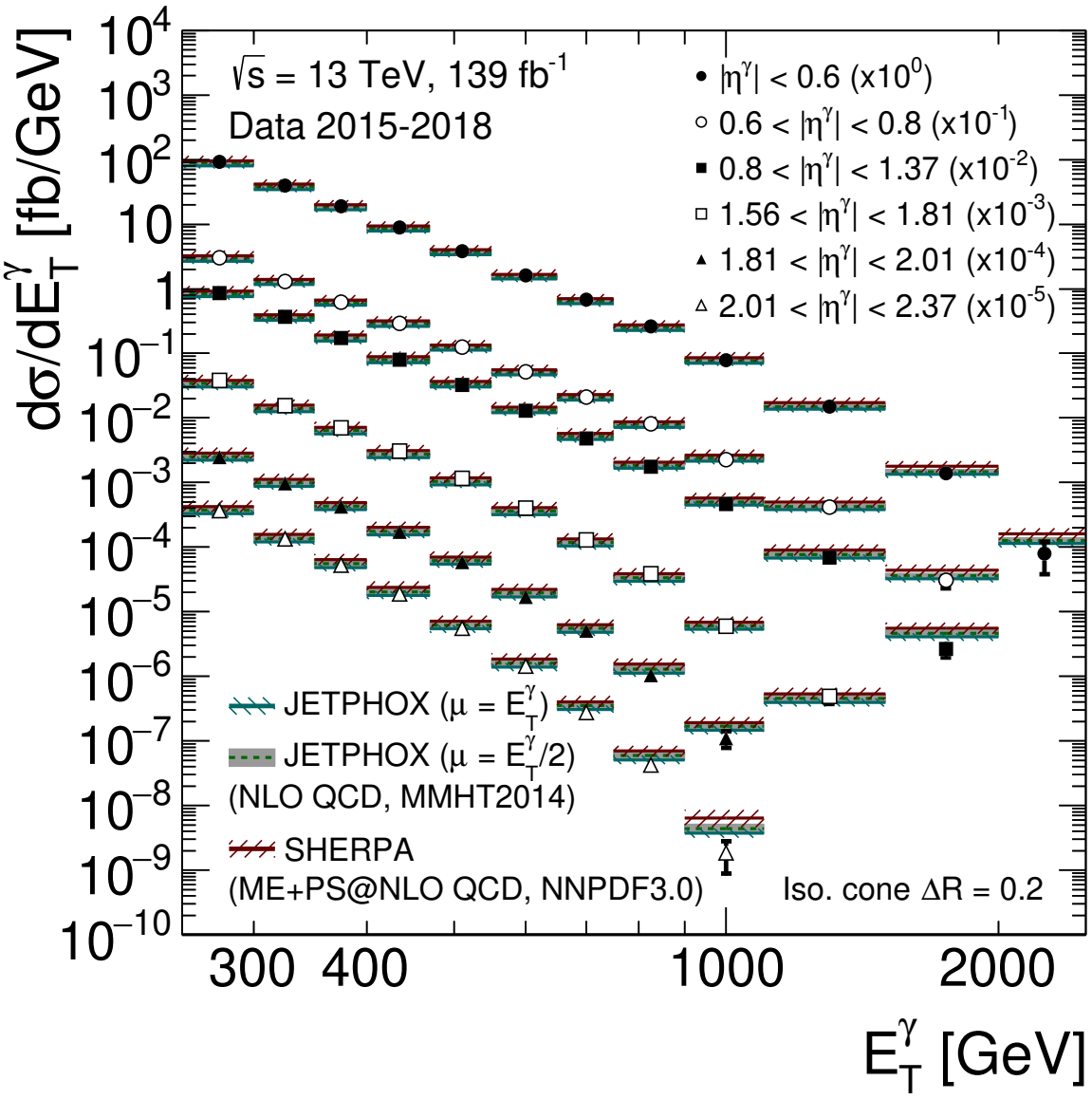


Figure 8.43: The measured differential cross section for isolated-photon production as a function of E_{T}^{γ} in $|\eta^{\gamma}| < 0.6$ (black dots), $0.6 < |\eta^{\gamma}| < 0.8$ (open circles), $0.8 < |\eta^{\gamma}| < 1.37$ (black squares), $1.56 < |\eta^{\gamma}| < 1.81$ (open squares), $1.81 < |\eta^{\gamma}| < 2.01$ (black triangles), $2.01 < |\eta^{\gamma}| < 2.37$ (open triangles). Results are shown for an isolation cone of size $\Delta R = 0.2$. The NLO pQCD prediction from JETPHOX with $\mu_{\mathrm{R}} = \mu_{\mathrm{F}} = \mu_{\mathrm{f}} = E_{\mathrm{T}}^{\gamma}$ or $\mu_{\mathrm{R}} = \mu_{\mathrm{F}} = \mu_{\mathrm{f}} = E_{\mathrm{T}}^{\gamma}/2$, and the ME+PS@NLO pQCD prediction from SHERPA 2.2.2 are shown. The measurement and the predictions are normalised by the factors shown in parentheses to aid visibility. The error bars represent the data statistical uncertainties and systematic uncertainties added in quadrature. For most of the points, the error bars are smaller than the marker size and, thus, not visible. The bands represent the theoretical uncertainty associated with the predictions.

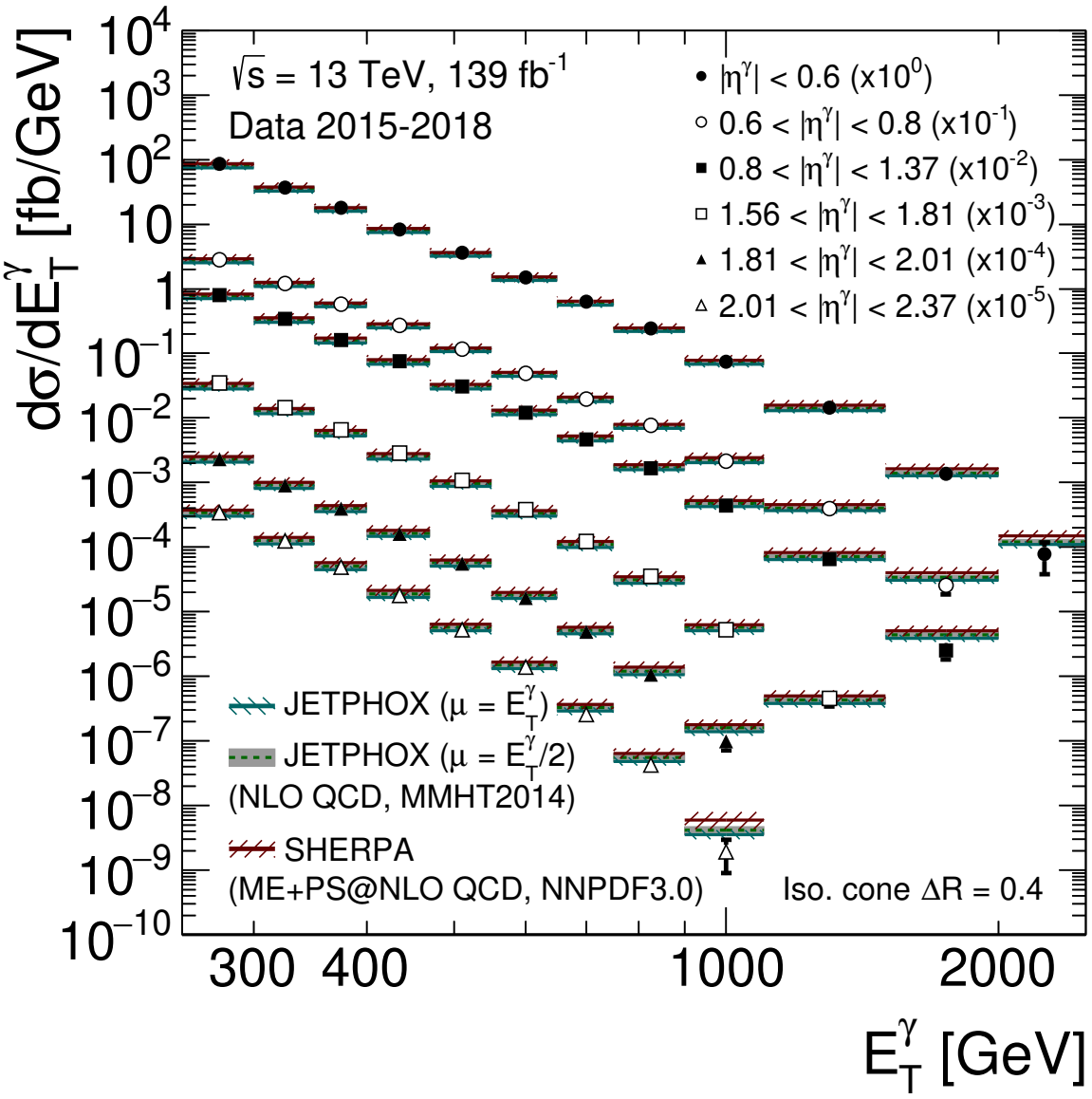


Figure 8.44: The measured differential cross section for isolated-photon production as a function of E_{T}^{γ} in $|\eta^{\gamma}| < 0.6$ (black dots), $0.6 < |\eta^{\gamma}| < 0.8$ (open circles), $0.8 < |\eta^{\gamma}| < 1.37$ (black squares), $1.56 < |\eta^{\gamma}| < 1.81$ (open squares), $1.81 < |\eta^{\gamma}| < 2.01$ (black triangles), $2.01 < |\eta^{\gamma}| < 2.37$ (open triangles). Results are shown for an isolation cone of size $\Delta R = 0.4$. The NLO pQCD predictions from JETPHOX with $\mu_{\mathrm{R}} = \mu_{\mathrm{F}} = \mu_{\mathrm{f}} = E_{\mathrm{T}}^{\gamma}$ or $\mu_{\mathrm{R}} = \mu_{\mathrm{F}} = \mu_{\mathrm{f}} = E_{\mathrm{T}}^{\gamma}/2$, and the ME+PS@NLO pQCD prediction from SHERPA 2.2.2 are shown. The measurement and the predictions are normalised by the factors shown in parentheses to aid visibility. The error bars represent the data statistical uncertainties and systematic uncertainties added in quadrature. For most of the points, the error bars are smaller than the marker size and, thus, not visible. The bands represent the theoretical uncertainty associated with the predictions.

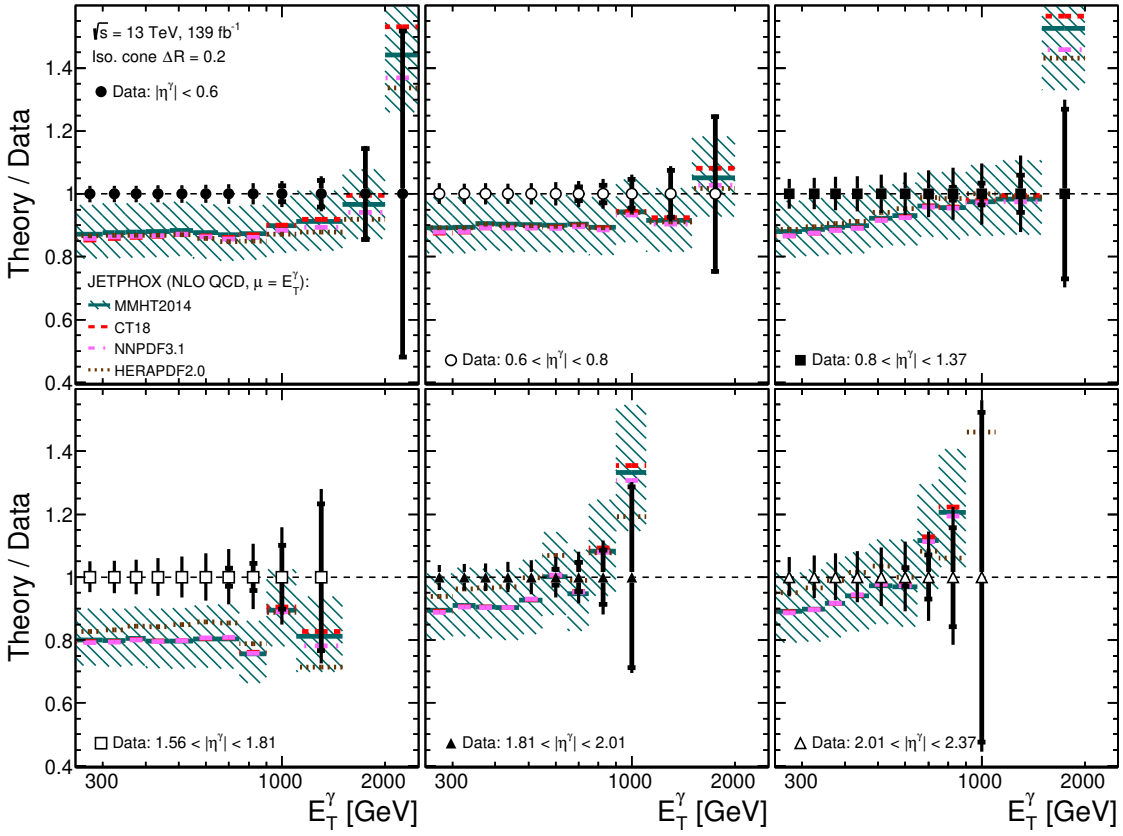


Figure 8.45: The ratio of the NLO pQCD prediction of JETPHOX with $\mu_{\mathrm{R}} = \mu_{\mathrm{F}} = \mu_{\mathrm{f}} = E_{\mathrm{T}}^{\gamma}$ using the MMHT2014 PDF set to the measured differential cross section for isolated-photon production (blue solid lines) as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. Results are shown for an isolation cone of size $\Delta R = 0.2$. The symbols for the data are centred at unity and the inner (outer) error bars represent the relative data statistical uncertainties (statistical and systematic uncertainties added in quadrature). The hatched bands represent the full theoretical uncertainty associated to the prediction based on the MMHT2014 PDF set, including the variations of the scales. For comparison, the predictions using the CT18 (red dashed lines), NNPDF3.1 (pink dashed lines) and HERAPDF2.0 (dotted lines) PDF sets are also included.

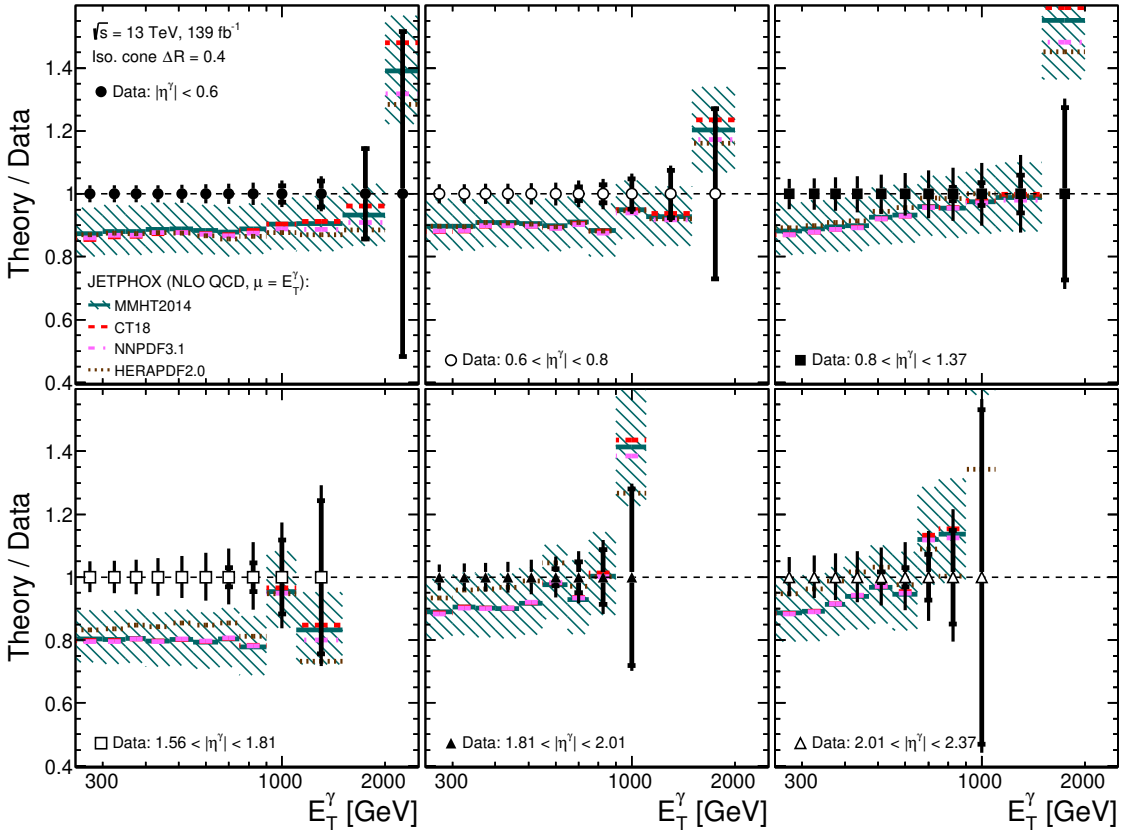


Figure 8.46: The ratio of the NLO pQCD prediction of JETPHOX with $\mu_{\mathrm{R}} = \mu_{\mathrm{F}} = \mu_{\mathrm{f}} = E_{\mathrm{T}}^{\gamma}$ using the MMHT2014 PDF set to the measured differential cross section for isolated-photon production (blue solid lines) as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. Results are shown for an isolation cone of size $\Delta R = 0.4$. The symbols for the data are centred at unity and the inner (outer) error bars represent the relative data statistical uncertainties (statistical and systematic uncertainties added in quadrature). The hatched bands represent the full theoretical uncertainty associated to the prediction based on the MMHT2014 PDF set, including the variations of the scales. For comparison, the predictions using the CT18 (red dashed lines), NNPDF3.1 (pink dashed lines) and HERAPDF2.0 (dotted lines) PDF sets are also included.

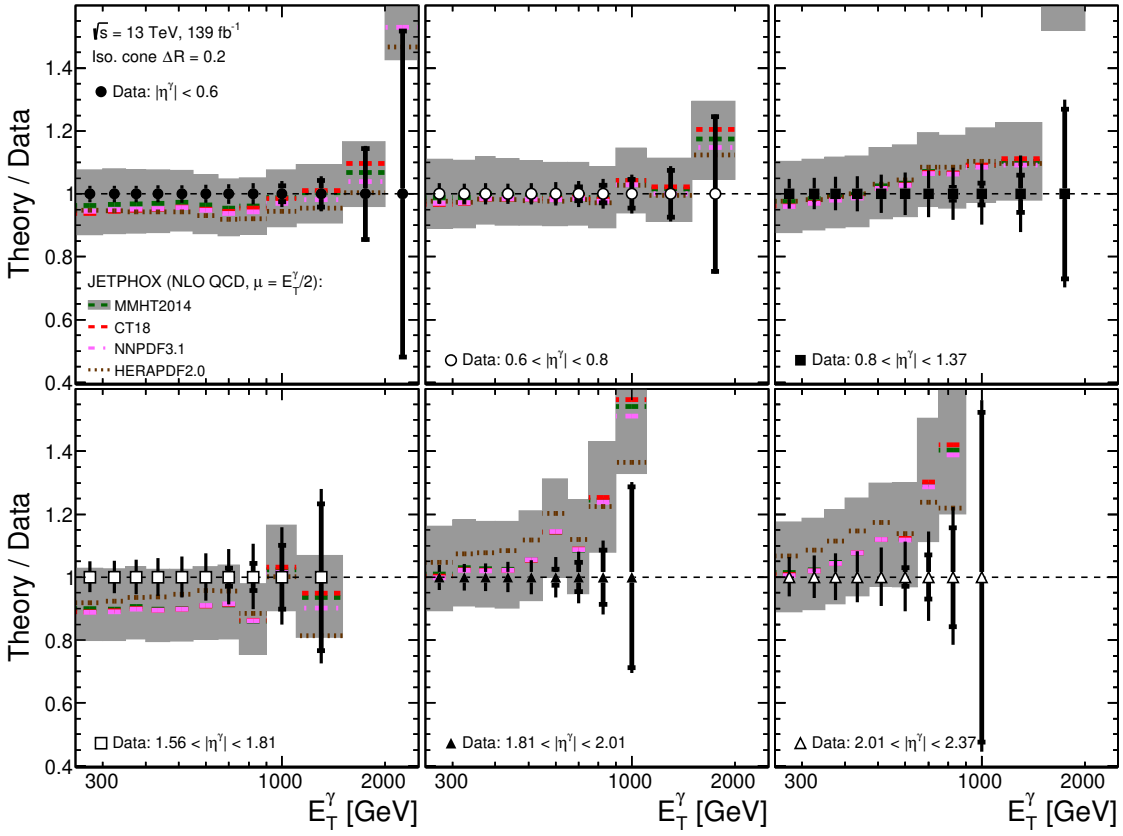


Figure 8.47: The ratio of the NLO pQCD prediction of JETPHOX with $\mu_{\mathrm{R}} = \mu_{\mathrm{F}} = \mu_{\mathrm{f}} = E_{\mathrm{T}}^{\gamma}/2$ using the MMHT2014 PDF set to the measured differential cross section for isolated-photon production (green dashed lines) as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. Results are shown for an isolation cone of size $\Delta R = 0.2$. The symbols for the data are centred at unity and the inner (outer) error bars represent the relative data statistical uncertainties (statistical and systematic uncertainties added in quadrature). The grey bands represent the full theoretical uncertainty associated to the prediction based on the MMHT2014 PDF set, including the variations of the scales. For comparison, the predictions using the CT18 (red dashed lines), NNPDF3.1 (pink dashed lines) and HERAPDF2.0 (dotted lines) PDF sets are also included.

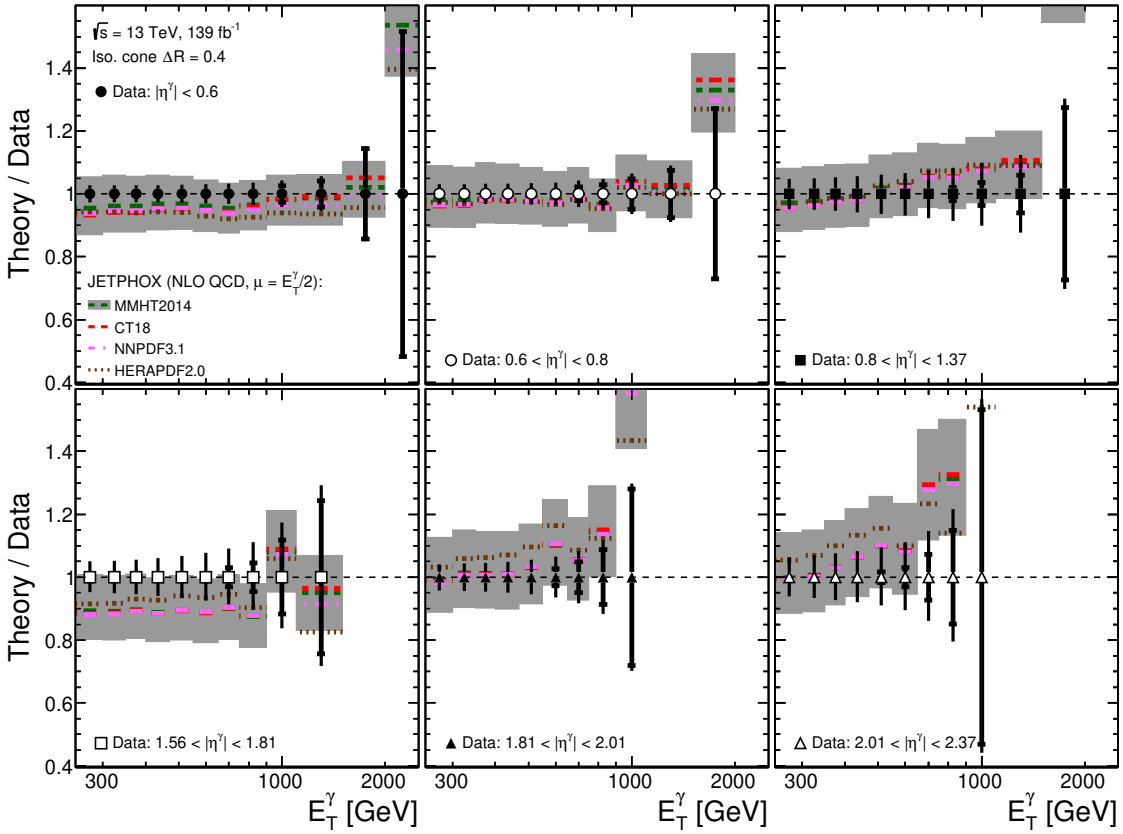


Figure 8.48: The ratio of the NLO pQCD prediction of JETPHOX with $\mu_{\mathrm{R}} = \mu_{\mathrm{F}} = \mu_{\mathrm{f}} = E_{\mathrm{T}}^{\gamma}/2$ using the MMHT2014 PDF set to the measured differential cross section for isolated-photon production (green dashed lines) as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. Results are shown for an isolation cone of size $\Delta R = 0.4$. The symbols for the data are centred at unity and the inner (outer) error bars represent the relative data statistical uncertainties (statistical and systematic uncertainties added in quadrature). The grey bands represent the full theoretical uncertainty associated to the prediction based on the MMHT2014 PDF set, including the variations of the scales. For comparison, the predictions using the CT18 (red dashed lines), NNPDF3.1 (pink dashed lines) and HERAPDF2.0 (dotted lines) PDF sets are also included.

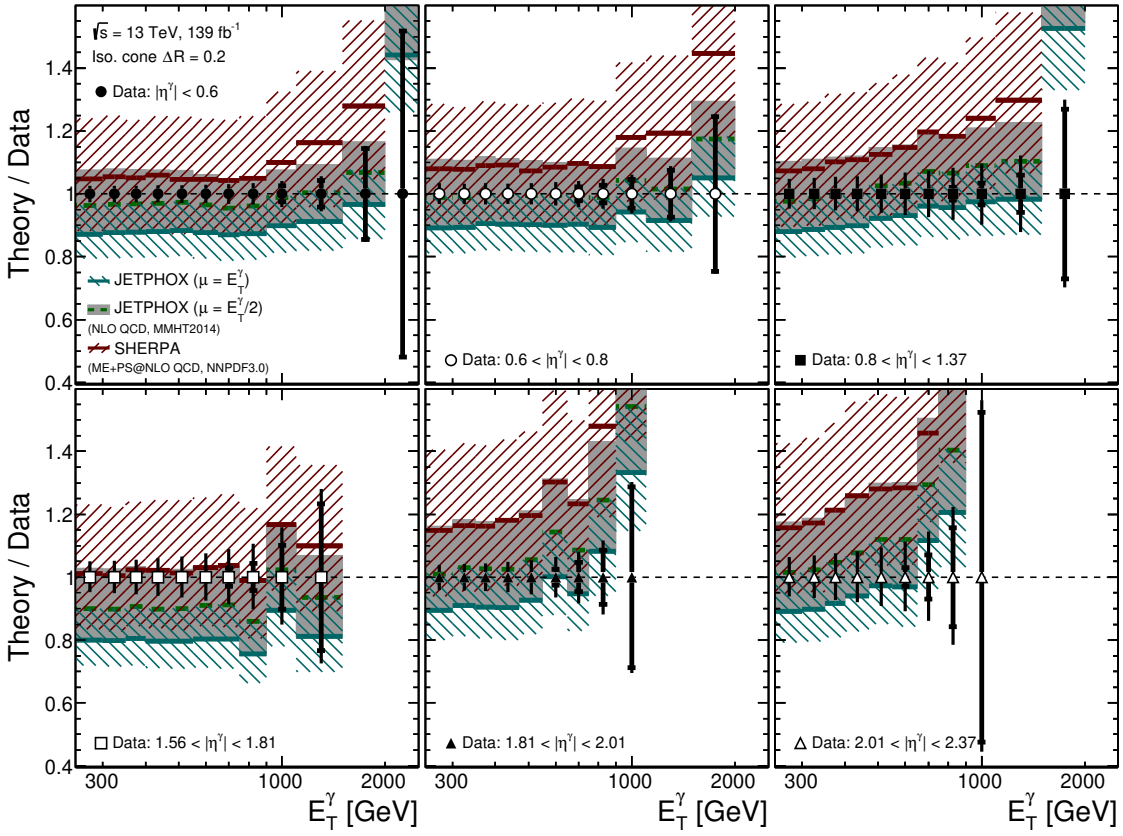


Figure 8.49: The ratio of the NLO pQCD prediction of JETPHOX with $\mu_{\mathrm{R}} = \mu_{\mathrm{F}} = \mu_{\mathrm{f}} = E_{\mathrm{T}}^{\gamma}$, the ratio of the NLO pQCD prediction of JETPHOX with $\mu_{\mathrm{R}} = \mu_{\mathrm{F}} = \mu_{\mathrm{f}} = E_{\mathrm{T}}^{\gamma}/2$, and the ratio of the ME+PS@NLO pQCD prediction of SHERPA 2.2.2 to the measured differential cross section for isolated-photon production as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. Results are shown for an isolation cone of size $\Delta R = 0.2$. The symbols for the data are centred at unity and the inner (outer) error bars represent the relative data statistical uncertainties (statistical and systematic uncertainties added in quadrature). The bands represent the full theoretical uncertainty, which includes the variations of the scales.

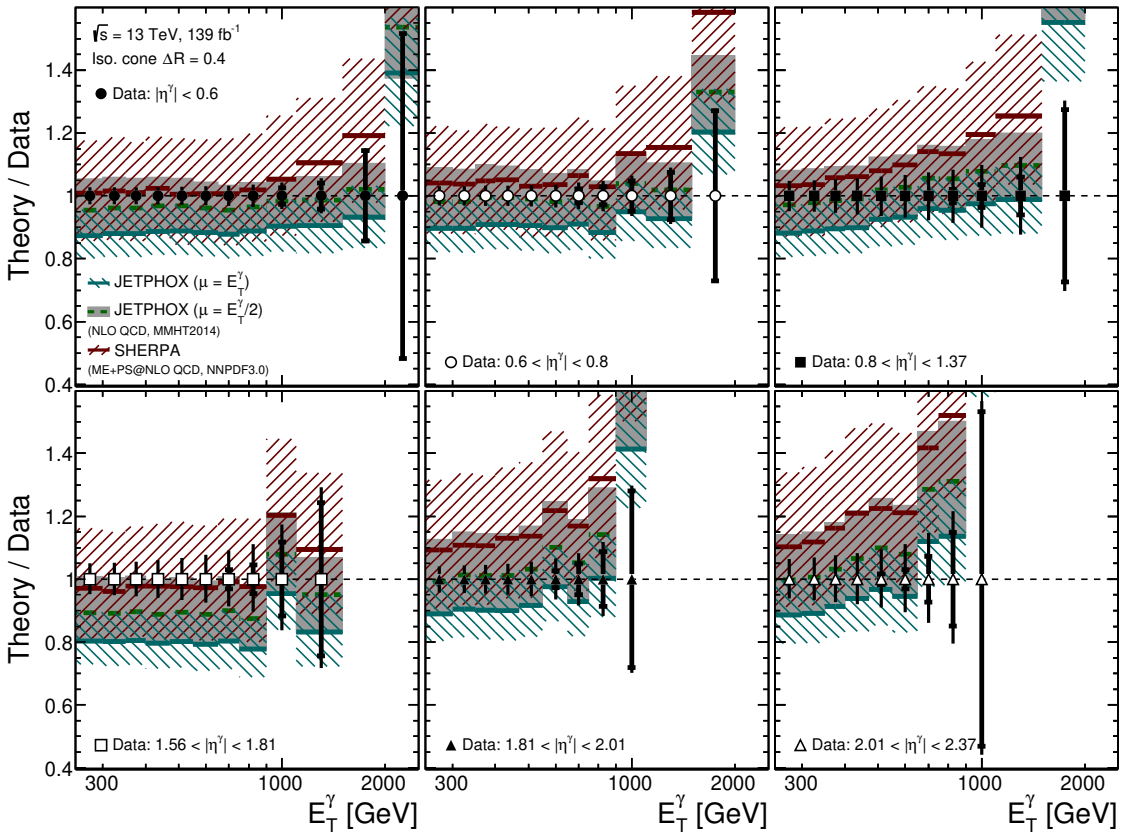


Figure 8.50: The ratio of the NLO pQCD prediction of JETPHOX with $\mu_{\mathrm{R}} = \mu_{\mathrm{F}} = \mu_{\mathrm{f}} = E_{\mathrm{T}}^{\gamma}$, the ratio of the NLO pQCD prediction of JETPHOX with $\mu_{\mathrm{R}} = \mu_{\mathrm{F}} = \mu_{\mathrm{f}} = E_{\mathrm{T}}^{\gamma}/2$, and the ratio of the ME+PS@NLO pQCD prediction of SHERPA 2.2.2 to the measured differential cross section for isolated-photon production as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. Results are shown for an isolation cone of size $\Delta R = 0.4$. The symbols for the data are centred at unity and the inner (outer) error bars represent the relative data statistical uncertainties (statistical and systematic uncertainties added in quadrature). The bands represent the full theoretical uncertainty, which includes the variations of the scales.

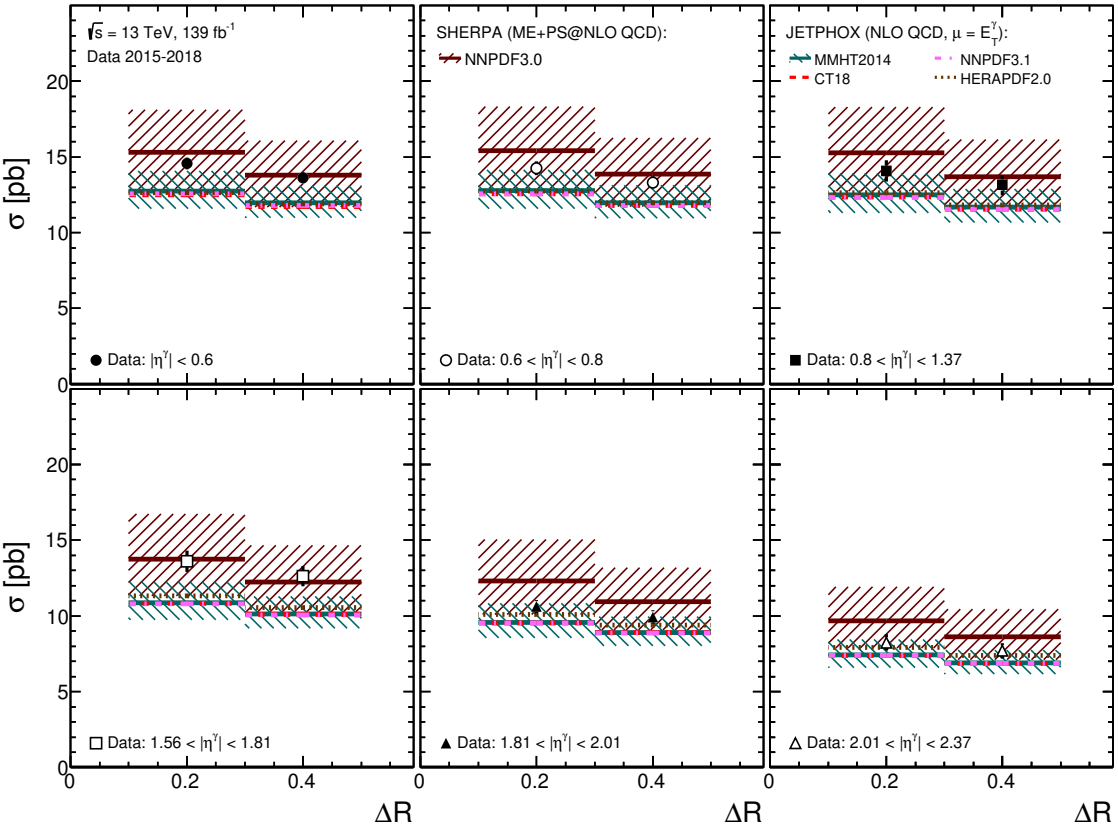


Figure 8.51: Integrated cross section for isolated-photon production as a function of ΔR in the different regions of $|\eta^{\gamma}|$. The data are represented with different symbols and the inner (outer) error bars represent the relative data statistical uncertainties (statistical and systematic uncertainties added in quadrature). The NLO pQCD prediction of JETPHOX with $\mu_{\mathrm{R}} = \mu_{\mathrm{F}} = \mu_{\mathrm{f}} = E_{\mathrm{T}}^{\gamma}$ and the ME+PS@NLO pQCD prediction of SHERPA 2.2.2 are also shown. The bands represent the full theoretical uncertainties of SHERPA 2.2.2 and JETPHOX, which include the variations of the scales. For comparison, the predictions of JETPHOX using the CT18 (red dashed lines), NNPDF3.1 (pink dashed lines) and HERAPDF2.0 (dotted lines) PDF sets are also included.

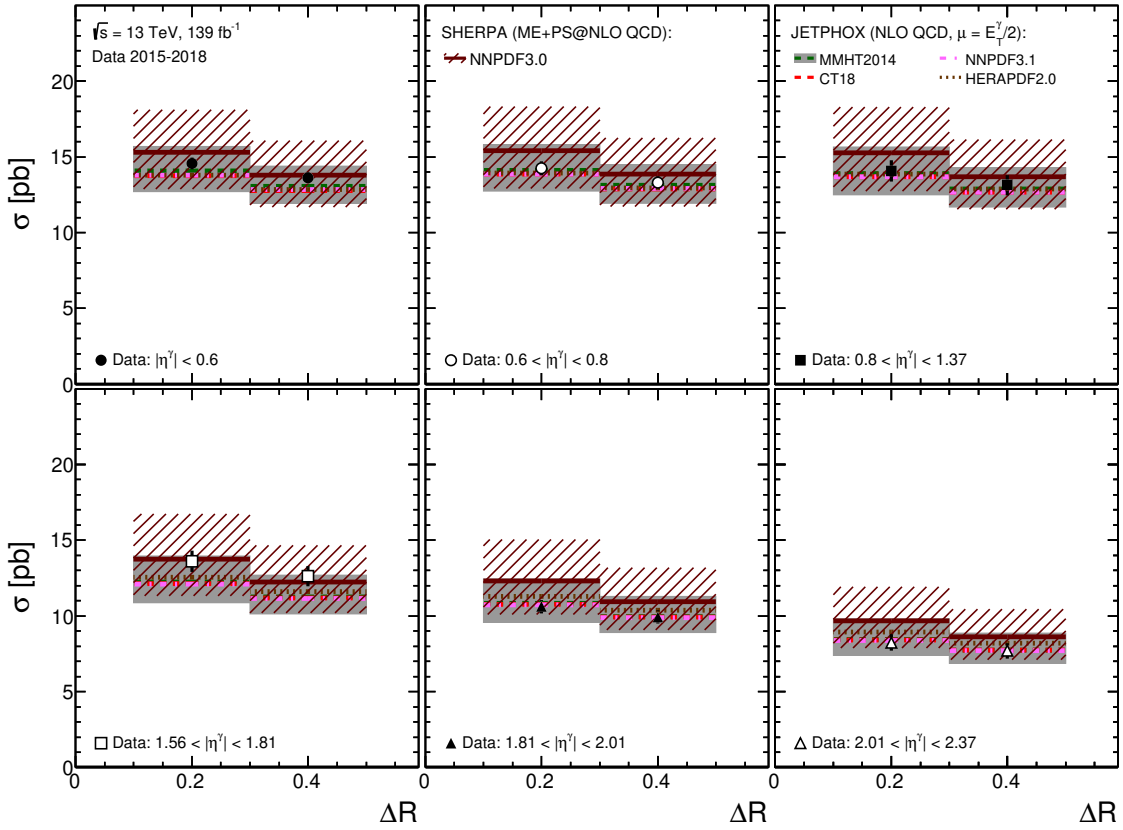


Figure 8.52: Integrated cross section for isolated-photon production as a function of ΔR in the different regions of $|\eta^{\gamma}|$. The data are represented with different symbols and the inner (outer) error bars represent the relative data statistical uncertainties (statistical and systematic uncertainties added in quadrature). The NLO pQCD prediction of JETPHOX with $\mu_{\mathrm{R}} = \mu_{\mathrm{F}} = \mu_{\mathrm{f}} = E_{\mathrm{T}}^{\gamma}/2$ and the ME+PS@NLO pQCD prediction of SHERPA 2.2.2 are also shown. The bands represent the full theoretical uncertainties of SHERPA 2.2.2 and JETPHOX, which include the variations of the scales. For comparison, the predictions of JETPHOX using the CT18 (red dashed lines), NNPDF3.1 (pink dashed lines) and HERAPDF2.0 (dotted lines) PDF sets are also included.

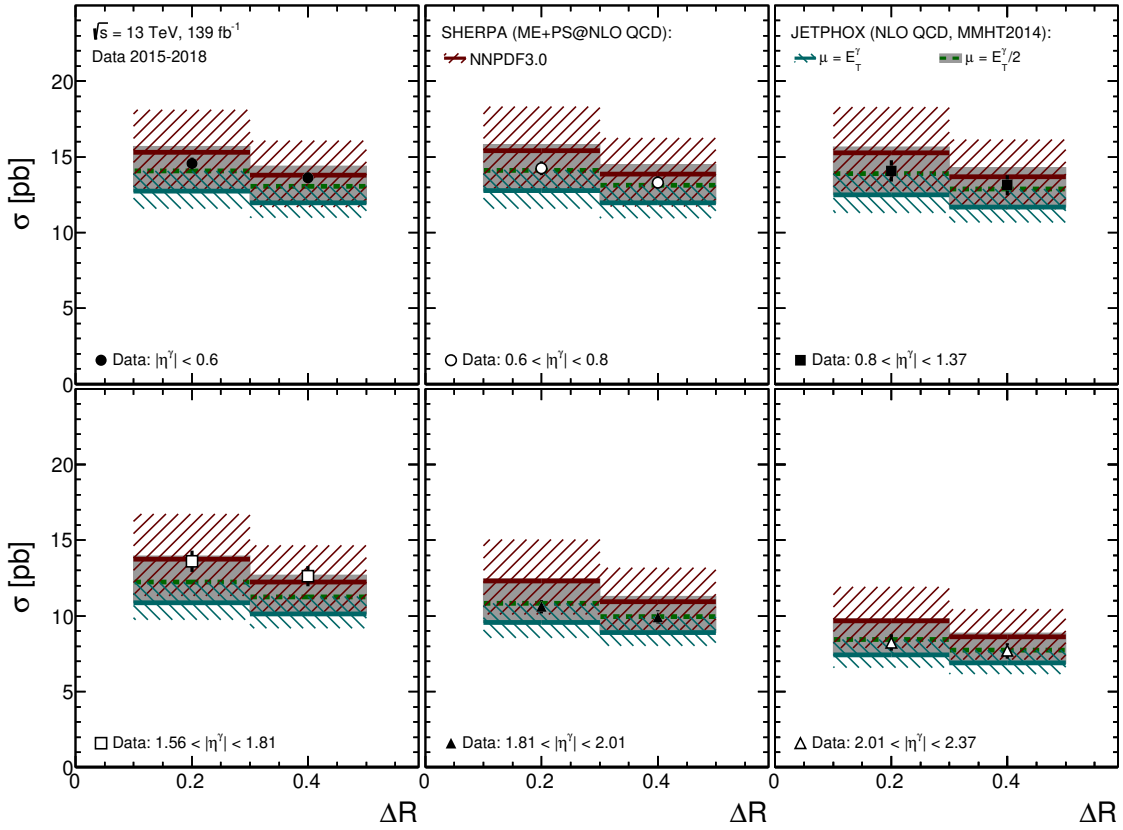


Figure 8.53: Integrated cross section for isolated-photon production as a function of ΔR in the different regions of $|\eta^{\gamma}|$. The data are represented with different symbols and the inner (outer) error bars represent the relative data statistical uncertainties (statistical and systematic uncertainties added in quadrature). The NLO pQCD prediction of JETPHOX with $\mu_{\mathrm{R}} = \mu_{\mathrm{F}} = \mu_{\mathrm{f}} = E_{\mathrm{T}}^{\gamma}$ and $\mu_{\mathrm{R}} = \mu_{\mathrm{F}} = \mu_{\mathrm{f}} = E_{\mathrm{T}}^{\gamma}/2$, and the ME+PS@NLO pQCD prediction of SHERPA 2.2.2 are also shown. The bands represent the full theoretical uncertainties from SHERPA 2.2.2 and JETPHOX, which include the variations of the scales.

Chapter 9

In situ jet energy scale and resolution calibration using γ+jet events

This chapter is composed of two different analyses, the in situ calibrations of the jet energy scale and the jet energy resolution, respectively. These two analyses were developed during the Run II data-taking period, with the aim of deriving a calibration using $80\,\text{fb}^{-1}$ from the $2015-2017$ periods, representing a significant improvement with respect to the previous set of JES [25] and JER [24] calibrations, and making possible an extension of the limits of the calibration together with an improvement in terms of precision for the benefit of the analyses performed in ATLAS during the Run II.

9.1 In situ JES calibration using γ+jet events

The γ+jet in situ calibration was introduced in section 6.2.2 together with other in situ calibrations that correct for the differences between the jet response in data and MC simulation. In this section, the methodology followed to derive the γ+jet calibration reported in ref. [26] is described in detail.

9.1.1 The MPF and direct balance methods

The latest γ+jet JES in situ calibration was derived using the MPF technique (for a detailed description of the method, see *Calibration measurement using Z+jet and γ+jet events* in section 6.2.2) in ref. [26]. Previous calibrations made use of the *Direct Balance* (DB) method [24,25], which measures the ratio of a fully reconstructed jet p_{T}, calibrated up to the η intercalibration stage and the reference object p_{T}.

For the latest calibration, the DB technique was used as a cross-check to the nominal MPF method. The use of a fully reconstructed and calibrated jet allows the calibration to be applied to jets in a straightforward manner. For a $2 \to 2$ γ+jet event, the p_{T}

of the jet can be expected to balance that of the reference object. However, the DB technique can be affected by additional parton radiation contributing to the recoil of the photon, appearing as subleading jets. The additional radiation is mitigated through a selection against events with a second jet of significant p_{T} and a minimum requirement on $\Delta\phi$, defined as the azimuthal angle between the photon and the jet, to ensure they are sufficiently back-to-back in ϕ. The component of the photon p_{T} perpendicular to the jet axis is also ignored to attempt to partially reduce the OOC effects; therefore, the projected reference p_{T} is defined as

$$p_{\mathrm{T}}^{\mathrm{ref,db}} = p_{\mathrm{T}}^{\gamma} \times \cos\left(\Delta\phi\right).$$

The average response of the jet was measured in bins of $p_{\mathrm{T}}^{\mathrm{ref}}$ for both data and MC simulation as

$$R_{\mathrm{DB}} = \left\langle \frac{p_{\mathrm{T}}^{\mathrm{j1}}}{p_{\mathrm{T}}^{\mathrm{ref,db}}} \right\rangle ,$$

where the average is extracted from the mean of the fit to the balance distribution $(p_{\mathrm{T}}^{\mathrm{j1}}/p_{\mathrm{T}}^{\mathrm{ref,db}})$ using a Gaussian distribution in each corresponding $p_{\mathrm{T}}^{\mathrm{ref}}$ bin.

The DB technique is also affected by OOC radiation, which consists of the energy lost outside of the reconstructed jet cone of $R = 0.4$ due to fragmentation processes. This radiation effect may lead to a p_{T} imbalance between a jet and the reference object, and was estimated for previous calibrations by measuring the profile of tracks around the jet axis [25].

As it has been previously pointed out (see *Calibration measurement using Z+jet and γ+jet events* in section 6.2.2), since the MPF technique uses the full hadronic recoil of an event rather than a reconstructed jet, it is therefore less sensitive to the jet definition, radius parameter and OOC radiation compared to the DB. In the MPF technique, reconstructed jets are only explicitly used in the event selection. The MPF technique is less sensitive to the generally ϕ-symmetric pile-up and UE. As the MPF technique is not derived from a reconstructed jet, the OOC correction does not directly reflect the energy within the reconstructed jet cone; the showering and topology corrections are used instead (see section 6.2.2). Another consequence of not using jets directly in the MPF method is that the impact of the GSC calibration (see *Global sequential calibration* in section 6.2.1) should be accounted in a different way. A correction to the cluster-based $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$, equal to the difference in momentum of the leading jet with and without the GSC calibration correction is applied [25].

The MPF and DB methods probe the calorimeter response to jets in a different way and are sensitive to different systematic effects. Therefore, they are complementary measurements of the jet energy scale. Since the MPF technique is less sensitive to QCD radiation and pile-up, and is almost independent of the jet algorithm, it was chosen as the nominal method for the latest jet energy scale calibration published by ATLAS [26].

9.1.2 Data selection

The data used in this calibration were collected with the ATLAS detector during the pp collision running periods of 2015, 2016 and 2017, when the LHC operated at a centre-of-mass energy of $\sqrt{s} = 13$ TeV. In the collected data sample, the average number of pp interactions per bunch crossing is ≈ 31, while the total integrated luminosity is $80.5 \pm 1.6\,\text{fb}^{-1}$ after requiring that all detector subsystems were operational during data recording.

Trigger requirements

The γ+jet data events were selected using a collection of single-photon triggers, each with a different associated photon E_{T} threshold and using loose identification; the set of triggers is listed in table 9.1. Events selected with the lower-threshold triggers were weighted according to a combination of trigger prescales [160]. The lowest unprescaled trigger during the 2015 (2016 and 2017) data-taking periods was the HLT_g120_loose (HLT_g140_loose). This mapping was created such that the trigger efficiency for each p_{T} range was at least 99%.

E_{T}^{γ} range (GeV)	HLT trigger
$15 - 100$	HLT_g10_loose
$20 - 100$	HLT_g15_loose
$25 - 100$	HLT_g20_loose
$30 - 110$	HLT_g25_loose
$40 - 110$	HLT_g35_loose
$45 - 120$	HLT_g40_loose
$50 - 120$	HLT_g45_loose
$55 - 130$	HLT_g50_loose
$65 - 130$	HLT_g60_loose
$75 - 140$	HLT_g70_loose
$85 - 140$	HLT_g80_loose
$105 - 150$	HLT_g100_loose
$125 - 160$	HLT_g120_loose
$145 - \text{inf}$	HLT_g140_loose

Table 9.1: List of single-photon triggers used in the γ+jet in situ calibration. Each trigger is shown with its corresponding E_{T}^{γ} range.

Event selection

Events in data taken under stable beam conditions and satisfying detector- and data-quality requirements, which include the calorimeters and inner tracking detectors being in nominal operation, were considered. The events, both in data and MC, were required to have at least one reconstructed primary vertex with at least two associated tracks of $p_{\text{T}} > 500$ MeV, consistent with the average beam-spot position.

Object selection

As described in section 6.2.2, offline photons were required to have $E_{\mathrm{T}}^{\gamma} > 25$ GeV and $|\eta^{\gamma}| < 1.37$; photons were also required to pass the same tight identification criteria which was defined in section 5.4. A tight isolation criteria was also used, requiring the calorimeter and track variables defined in section 5.5 to fulfil the following conditions: $E_{\mathrm{T}}^{\mathrm{cone40}} < 0.022 \times E_{\mathrm{T}}^{\gamma} + 2.45$ GeV and $p_{\mathrm{T}}^{\mathrm{cone20}}/E_{\mathrm{T}}^{\gamma} < 0.05$ for the calorimeter and track variables, respectively. To achieve a better jet-background rejection, a requirement was applied for converted photons on the E_{T} of the photon cluster ($E_{\mathrm{T,cluster}}^{\gamma}$) and the sum of p_{T} of the associated tracks to that cluster ($\sum p_{\mathrm{T}}^{\mathrm{tracks}}$),

- If the photon cluster has two matched tracks, the event is retained if $0.5 < \frac{E_{\mathrm{T,cluster}}^{\gamma}}{\sum p_{\mathrm{T}}^{\mathrm{tracks}}} < 1.5$.
- If the photon cluster has one matched track, the event is retained if $\frac{E_{\mathrm{T,cluster}}^{\gamma}}{\sum p_{\mathrm{T}}^{\mathrm{tracks}}} < 2$.

The leading photon was selected from all those satisfying the previous selection requirements.

Both EMtopo and PFlow jets with radius $R = 0.4$ were calibrated, being PFlow the primary ATLAS jet definition (see section 6.1). Jets were calibrated with all the steps up to the η intercalibration. To avoid double counting of energy depositions, jets were required to be $\Delta R > 0.2$ from the leading photon. Jets were required to satisfy basic cleaning requirements and for those with $p_{\mathrm{T}} < 60$ GeV in the region $|\eta| < 2.4$, the JVT was required to be larger than 0.2 (0.59) for the PFlow (EMtopo) sample. The leading jet in the event was selected requiring $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 0.8$.

Further selection requirements were applied to limit the effects of ISR and FSR; these requirements were slightly adapted to the MPF and DB samples. The additional event activity was suppressed in the MPF (DB) technique by requiring that any subleading jet must have $p_{\mathrm{T}} < \max\left(0.3 \times p_{\mathrm{T}}^{\mathrm{ref}}, 12\right)$ GeV ($p_{\mathrm{T}} < \max\left(0.1 \times p_{\mathrm{T}}^{\mathrm{ref}}, 15\right)$ GeV) and that the leading jet and the reference object must be relatively back-to-back with $\Delta\phi^{\mathrm{ref,jet}} > 2.9$ ($\Delta\phi^{\mathrm{ref,jet}} > 2.8$). Since the MPF technique is less sensitive to QCD radiation, the subleading jet cut is relaxed, keeping a larger number of selected events compared to the DB method; the higher statistics allows the use of a tighter $\Delta\phi$ cut.

9.1.3 Monte Carlo simulations

Samples of MC events were generated to study the characteristics of the signal events and to determine the jet energy scale and its uncertainty. The same PYTHIA and SHERPA samples described in section 4.3 were used for the calibration. For the set of results presented in this analysis, both MC campaigns a and d were combined after being normalised to the luminosity of 2015 − 2016 and 2017 data, respectively. The additional pile-up and photon identification corrections described in section 5.6 were also applied to the MC simulation, in this analysis.

9.1.4 Fitting procedure

For both MPF and DB techniques, the response is initially derived in bins of $p_{\mathrm{T}}^{\mathrm{ref}}$. In each bin, a maximum likelihood fit is performed using a Gaussian distribution. The fit range is limited to twice the RMS of the response distribution around the mean to minimise the effect of MC mismodelling in the tails of the distribution. An exception is done in the first bin of the direct balance data distributions, for $25 < p_{\mathrm{T}}^{\mathrm{ref}} < 45$ GeV, which are biased on the left tail due to the $p_{\mathrm{T}}^{\mathrm{jet}}$ cut. For those distributions the fit range was adapted to improve the description of the peak of the distributions. The average response (R_{DB} or R_{MPF}) was taken as the mean of the best-fit Gaussian distribution.

Results of the fits to the MPF and DB responses in four different $p_{\mathrm{T}}^{\mathrm{ref}}$ bins for the PFlow sample are shown in figures 9.1 and 9.2, respectively. The distributions for the other $p_{\mathrm{T}}^{\mathrm{ref}}$ bins in data are shown in appendix H for the PFlow and EMtopo samples and for both MPF and DB responses.

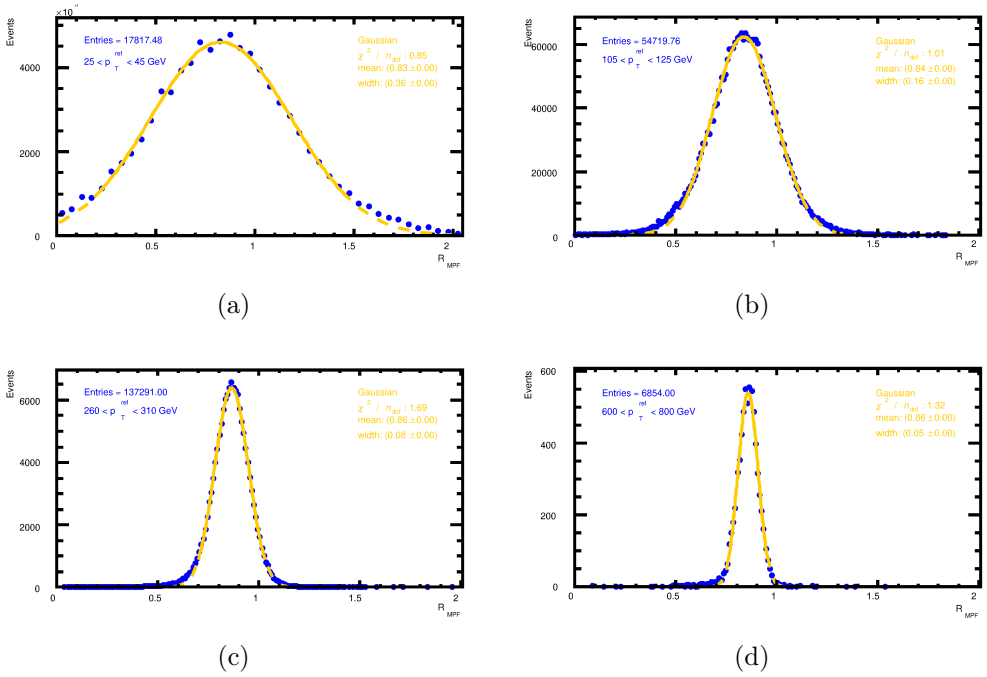
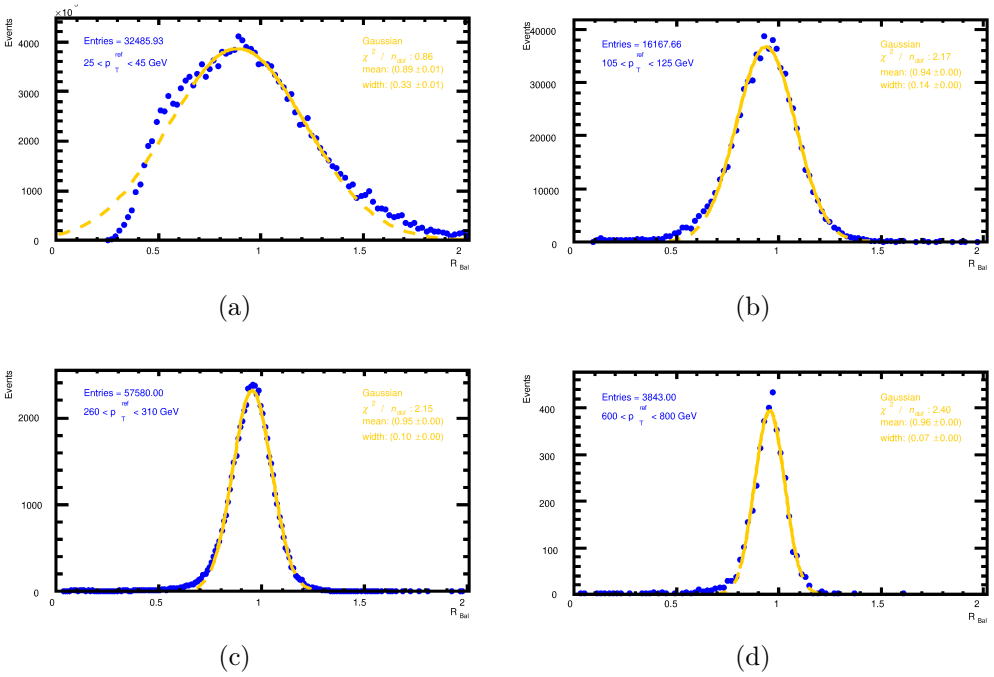


Figure 9.1: Distributions of r_{MPF} in $2015-2017$ data using PFlow jets for the $p_{\mathrm{T}}^{\mathrm{ref}}$ bins: $25 < p_{\mathrm{T}}^{\mathrm{ref}} < 45$ GeV (a), $105 < p_{\mathrm{T}}^{\mathrm{ref}} < 125$ GeV (b), $260 < p_{\mathrm{T}}^{\mathrm{ref}} < 310$ GeV (c), and $600 < p_{\mathrm{T}}^{\mathrm{ref}} < 800$ GeV (d). The data distributions are represented with blue dots, and the fit to a Gaussian distribution is shown with the solid orange line. The dashed orange line corresponds to the part of the distribution excluded from the fit.



(a) (b)

(c) (d)

Figure 9.2: Direct balance distributions in 2015 – 2017 data using PFlow jets for the $p_{\mathrm{T}}^{\mathrm{ref}}$ bins: $25 < p_{\mathrm{T}}^{\mathrm{ref}} < 45$ GeV (a), $105 < p_{\mathrm{T}}^{\mathrm{ref}} < 125$ GeV (b), $260 < p_{\mathrm{T}}^{\mathrm{ref}} < 310$ GeV (c), and $600 < p_{\mathrm{T}}^{\mathrm{ref}} < 800$ GeV (d). The data distributions are represented with blue dots, and the fit to a Gaussian distribution is shown with the orange line. The dashed orange line corresponds to the part of the distribution excluded from the fit.

9.1.5 Signal purity estimation

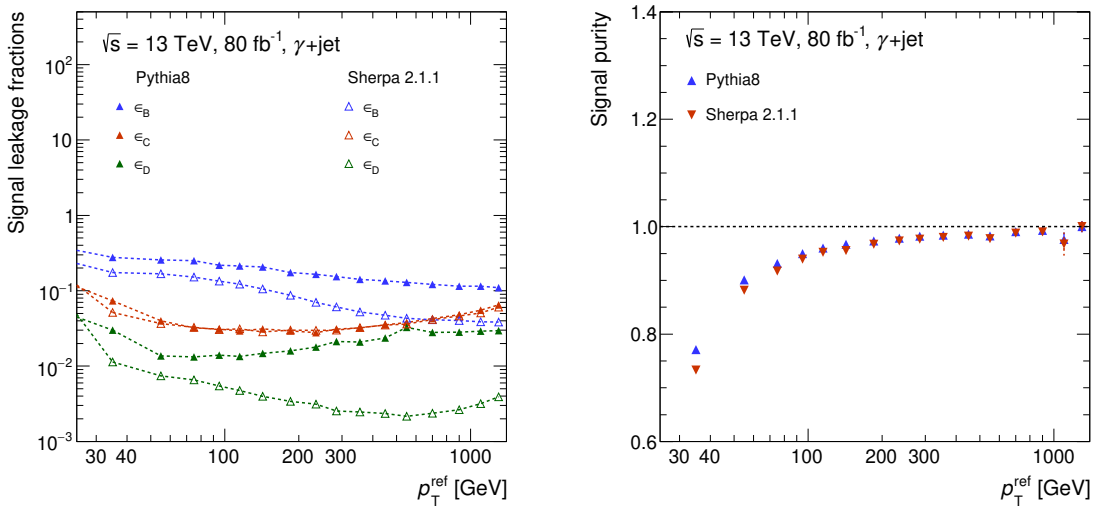
As described in section 7.3, even after the application of the photon tight identification and isolation requirements, there is a non-negligible amount of background remaining in the γ+jet selected sample. This background from multijet processes is estimated using a similar data-driven method as that described in sections 7.3 and 8.3. Differences with respect to the control region definitions provided in section 7.3 are related with the use of a different isolation requirement, the absence of an $E_{\mathrm{T}}^{\mathrm{iso}}$ gap separating regions A and C from B and D, the lack of an upper cut on $E_{\mathrm{T}}^{\mathrm{iso}}$ that limits the size of regions B and D, and the assumption of a correlation between the identification and isolation variables of $R^{\mathrm{bg}} = 1.45$[1].

The signal yield was determined from the observed number of events in the data in the four regions of the γ_{ID} vs $E_{\mathrm{T}}^{\mathrm{iso}}$ plane and the signal leakage fractions obtained from the simulated events using equation 7.1. The signal leakage fractions obtained from PYTHIA and SHERPA are shown in figure 9.3(a). The fractions ϵ_{C}, which represent the signal leakage into the non-tight and isolated control region, are approximately constant and

[1]This assumption is based on previous studies and was not estimated for these datasets.

around $0.05 - 0.1$, being very similar for PYTHIA and SHERPA. The fractions ϵ_{B}, which represent the signal leakage into the tight and non-isolated background control region, decrease as a function of $p_{\mathrm{T}}^{\mathrm{ref}}$ and are different for PYTHIA and SHERPA. The fractions ϵ_{D}, which represent the signal leakage into the non-tight and non-isolated control region, are approximately constant and around $0.02 - 0.04$ for PYTHIA; the fraction decreases for SHERPA as a function of $p_{\mathrm{T}}^{\mathrm{ref}}$. This control region is the most affected by the different treatment of the brem component in both models (see section 7.3).

The signal purities, computed as $P = N_{\mathrm{A}}^{\mathrm{sig}}/N_{\mathrm{A}}$ using the signal leakage fractions from PYTHIA and SHERPA are shown in figure 9.3(b). For some of the points at high E_{T}^{γ} in which the purity is above unity or consistent with unity but with large statistical uncertainty, the purity was set to unity. The signal purity is $\gtrsim 70\%$ and the estimations using either PYTHIA or SHERPA to compute the signal leakage fractions show very similar values. The nominal signal yields were extracted using the signal leakage fractions from PYTHIA. The signal purities estimated with the MPF and DB selections and with both PFlow and EMtopo samples were found to be very similar



(a) (b)

Figure 9.3: The signal leakage fractions from PYTHIA (solid markers) and SHERPA (empty markers) for the control regions B (blue triangles), C (red triangles), and D (green triangles) as a function of $p_{\mathrm{T}}^{\mathrm{ref}}$ (a). The signal purities estimated in data using signal leakage fractions of PYTHIA (blue triangles) and SHERPA (red triangles) as a function of $p_{\mathrm{T}}^{\mathrm{ref}}$ (b). Results are shown for the MPF sample using PFlow jets.

9.1.6 Response distributions

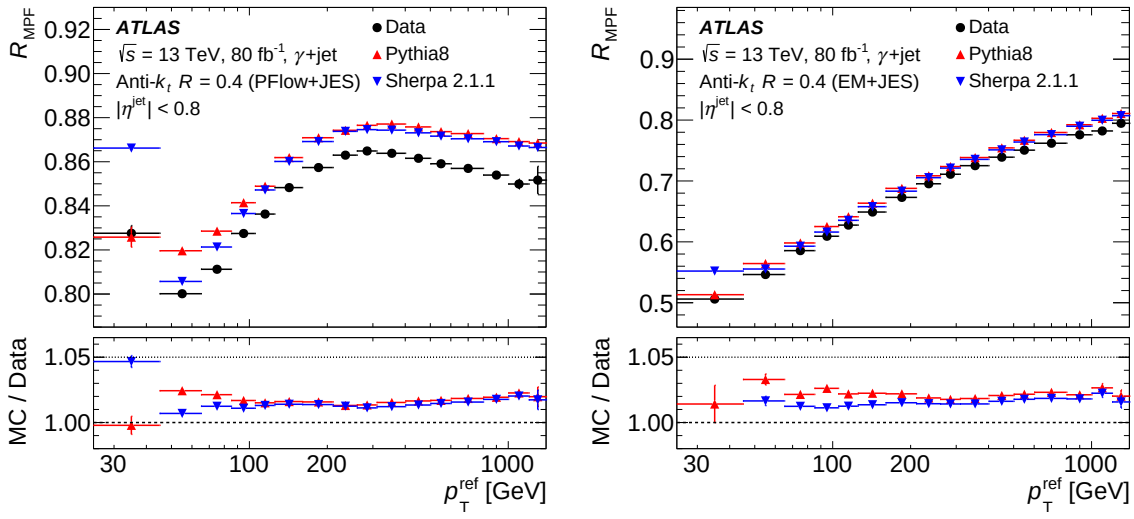
In this section the PFlow and EMtopo jet responses obtained with the MPF and DB methods are presented. The nominal calibration results obtained with PFlow jets and the MPF method are compared to those of EMtopo jets. In addition, the pile-up stability of the MPF calibration using PFlow jets is also studied. Finally, a comparison of the calibration central values obtained either using the DB or MPF methods is provided.

MPF results

The MPF average responses calculated in γ+jet events using the procedure described in section 9.1.4 were obtained for the data and the MC simulations of PYTHIA and SHERPA. The MPF average responses are shown in figure 9.4(a) (9.4(b)) for the selected sample using PFlow (EMtopo) jets. The upper part of both figures shows the average response as a function of $p_{\mathrm{T}}^{\mathrm{ref}}$. The responses observed in the data are overestimated by both MC generators except for the first bin in the PFlow comparisons. The lower part of each figure shows the ratio of the MC to the data, which in the case of PYTHIA, is used as the calibration factor. The ratios show differences around $1-2\%$ ($1-3\%$) for the PFlow (EMtopo) selection except for the first bin and the SHERPA prediction, which is biased due to the presence of a generator-level cut at 35 GeV. The MC-to-data ratio is slightly smaller for the PFlow selection; the effect is more noticeable for the ratio of PYTHIA in the range $p_{\mathrm{T}}^{\mathrm{ref}} < 200$ GeV.

The innate difference in response between EMtopo and PFlow jets can be seen by comparing their measured MPF responses. The MPF method uses topo-clusters and PFlow objects in computing the missing energy, so that the measured responses are independent of the MCJES calibration and reflect the precalibration response for each jet input type. Figure 9.5 shows the comparison of the MPF responses for EMtopo and PFlow jets. The shape of the EMtopo distribution follows the form of the Groom function, which corresponds to the response expected from a hadronic calorimeter [161]. The PFlow measurement does not follow the same shape but instead shows an improvement over the baseline calorimeter response at low p_{T} thanks to the inclusion of tracking information.

One of the primary changes in the LHC run conditions along the Run II period was the increase of pile-up. The average number of interactions per bunch crossing during the $2015-2016$ data-taking period was ≈ 24, which increased to ≈ 38 in 2017. The consistency of the calibration for events with different pile-up conditions is an important characteristic of the methods. Figure 9.6 shows the dependence of the ratio of the MPF responses for PFlow jets as a function of the number of primary vertices, N_{PV}, and the average number of interactions per bunch crossing, $\langle\mu\rangle$. The results are shown for the first low-$p_{\mathrm{T}}^{\mathrm{ref}}$ bin where the γ+jet calibration has appropriate statistical significance (see *In situ combination* in section 6.2.2). A linear fit to the ratio of MC to data has a slope compatible with zero within the fit uncertainties in each plot, demonstrating the stability of the in situ calibration as a function of N_{PV} and $\langle\mu\rangle$.



(a) (b)

Figure 9.4: Average PFlow (a) and EMtopo (b) jet responses as a function of the reference p_{T} for γ+jet events calculated using the MPF technique. The data are represented with black dots while the prediction of PYTHIA (SHERPA) is represented with red (blue) triangles. The ratio of MC to data for both generators is shown in the bottom panel and defines the in situ correction to be applied. The dotted lines at 1 and 1.05 serve as a reference.

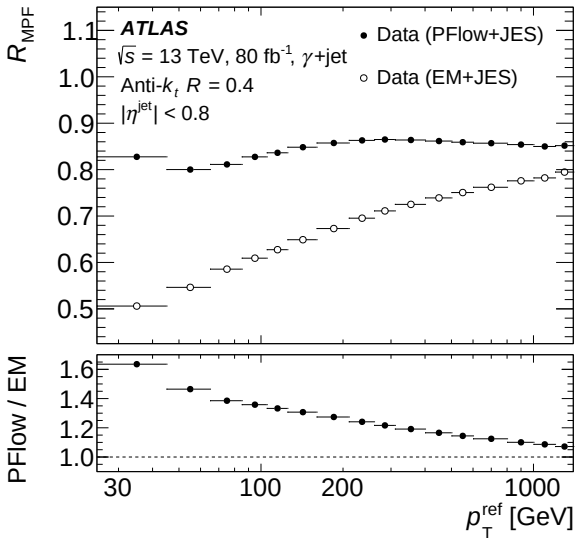
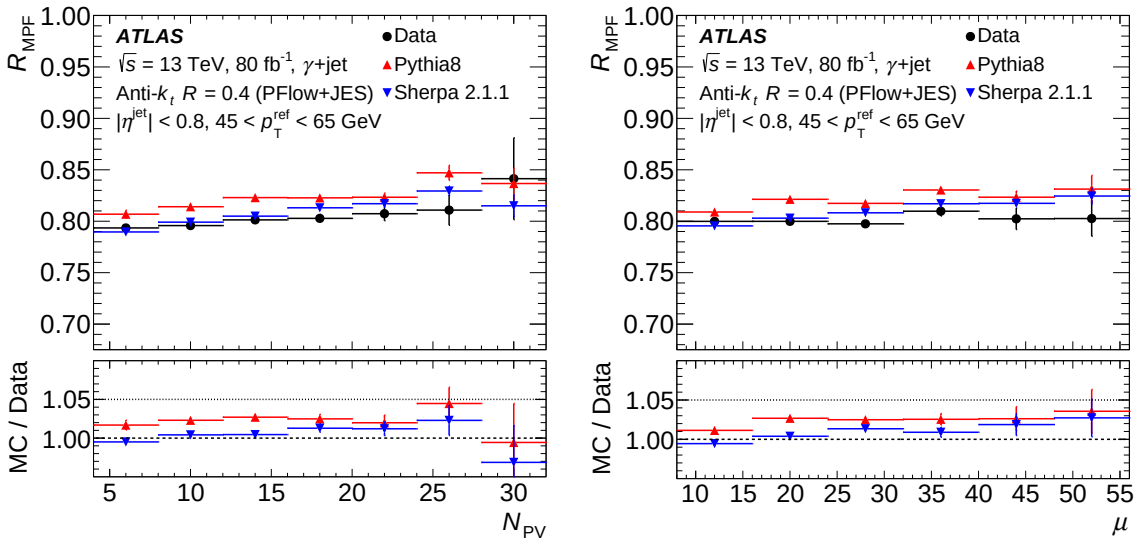


Figure 9.5: Average jet response as a function of $p_{\mathrm{T}}^{\mathrm{ref}}$ for γ+jet events using the MPF technique in $2015-2017$ data. The solid points correspond to PFlow jets while the empty points represent the results obtained with EMtopo jets. The ratio of PFlow to EMtopo responses is shown in the bottom panel.



(a) (b)

Figure 9.6: Average response for PFlow jets using the MPF technique as a function of N_{PV} (left) $\langle\mu\rangle$ (right) for γ+jet events in the reference p_{T} bin $45 < p_{\mathrm{T}}^{\mathrm{ref}} < 65$ GeV. The black points correspond to $2015-2017$ data while the red and blue triangles correspond to the MC PYTHIA and SHERPA simulations, respectively. The error bars represent the statistical uncertainties. The ratio of MC to data for both generators is shown in the bottom panel. The dotted lines at 1 and 1.05 serve as a reference.

DB results

The value of R_{DB} is extracted for each $p_{\mathrm{T}}^{\mathrm{ref}}$ bin for both data and simulation. Figure 9.7 shows the measurements of R_{DB} in data compared with predictions from PYTHIA and SHERPA as a function of $p_{\mathrm{T}}^{\mathrm{ref}}$ for both PFlow and EMtopo jet collections. The ratios of the predictions of PYTHIA to the data show values between $1-3\%$. The prediction of SHERPA describes the data better, being compatible with 1% within statistical uncertainties except in the tail of the distributions and for the lowest-$p_{\mathrm{T}}^{\mathrm{ref}}$ bin due to the bias caused by the generator level cut of 35 GeV. For $p_{\mathrm{T}}^{\mathrm{ref}}$ values below approximately 100 GeV and above 600 GeV the differences between the MC and the data increase. The results obtained for PFlow and EMtopo jets are in a very good agreement.

The comparison of the calibration correction factors derived using both the MPF and DB techniques are shown in figure 9.8. Differences are compatible with unity within statistical uncertainties in most of the bins for PFlow jets. For EMtopo jets the agreement is not as good as for PFlow, and the MPF calibration factor tends to be around 0.3% larger than the DB result, but still showing a good agreement between the measurements. These results represent an important test and support the idea that both methods, despite probing the calorimeter response to jets in a different way, are compatible.

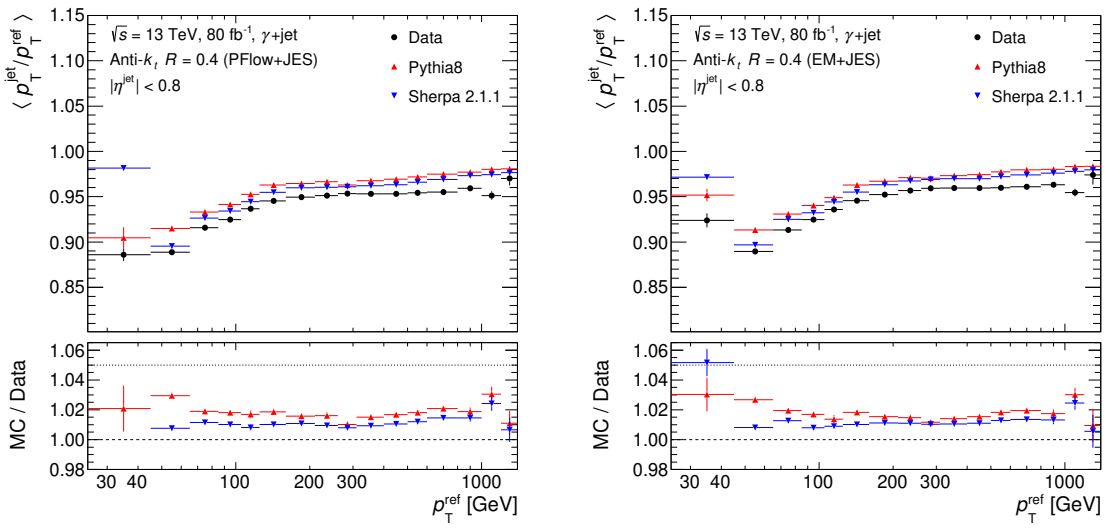


Figure 9.7: Average PFlow (a) and EMtopo (b) jet responses, $\left\langle p_{\mathrm{T}}^{\mathrm{j1}}/p_{\mathrm{T}}^{\mathrm{ref,db}} \right\rangle$, as a function of the reference p_{T} for γ+jet events calculated using the DB technique. The data are represented with black circles while the prediction of PYTHIA (SHERPA) is represented with red (blue) triangles. The ratio of MC to data for both generators is shown in the bottom panel. The dotted lines at 1 and 1.05 serve as a reference.

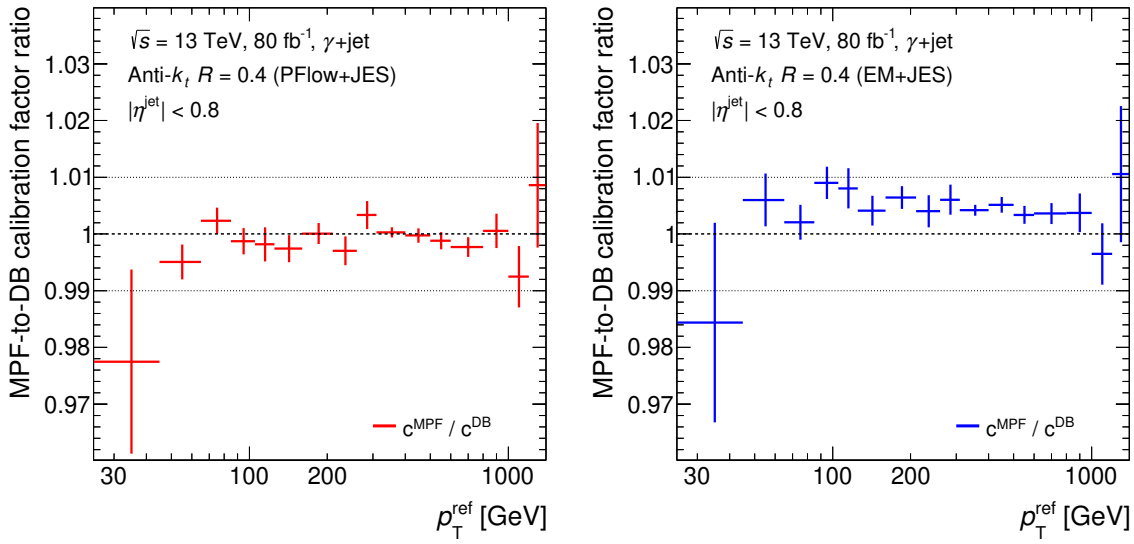


Figure 9.8: The double ratio of the MPF-to-DB calibration factors obtained for PFlow (left) and EMtopo (right) jets as a function of $p_{\mathrm{T}}^{\mathrm{ref}}$. The error bars represent the statistical uncertainties of the double ratio only. The dotted lines serve as a reference.

9.1.7 Systematic uncertainties

The uncertainties on the JES estimated with the MPF and DB methods using both PFlow and EMtopo jet collections are presented in this section. The sources of uncertainty that affect the measurement of the data-to-MC ratio of R_{DB} and R_{MPF} include the modelling of physics processes in simulation, uncertainties in the measurement of the reference object, and uncertainties in the expected p_{T} balance due to the event topology. For some of the systematic uncertainties, a rebinning procedure is applied to ensure that the features presented in the final result are statistically significant and not the result of fluctuations in small numbers of simulated or data events. The rebinning procedure follows a bootstrapping method: pseudo-experiment datasets are created by sampling from a Poisson distribution with a mean of 1 for each event in the data or MC simulation [154]. The pseudo-experiments are therefore statistically correlated yet unique, and the RMS of the response distribution across the pseudo-experiments provides a measure of the statistical uncertainty of the analysis.

Photon energy scale and resolution

The uncertainty on the R_{DB} and R_{MPF} measured responses due to the uncertainty on the energy scale and resolution of the photon energy was estimated propagating two individual components, influencing the photon energy scale and its resolution, respectively. Each of the two components accounted for the different sources of uncertainty described in section 5.3.4. The systematic uncertainty in the data-to-MC ratio was evaluated by varying each individual source of uncertainty separately by $\pm 1\sigma$ in the MC simulation. The uncertainty due to the photon energy resolution was found to be negligible for both MPF and DB measurements. The uncertainty due to the photon energy scale is the dominant uncertainty in the medium- and high-$p_{\mathrm{T}}^{\mathrm{ref}}$ range of the γ+jet calibration, increasing from approximately 0.3% to 0.8% at high-$p_{\mathrm{T}}^{\mathrm{ref}}$. The size of this uncertainty is very similar for both MPF and DB methods and PFlow and EMtopo jets. This uncertainty was estimated using a bootstrap method and was also rebinned to reduce the effect of statistical fluctuations.

MC generator

The predictions of the response observables R_{DB} and R_{MPF} were produced with two different MC generators, PYTHIA and SHERPA. As described in section 4.3, these generators use different modelling of the parton shower, jet fragmentation, and multiple parton interactions. The difference in the data-to-MC ratio of the response between the generators is taken as the *MC generator* systematic uncertainty source. This uncertainty is dominant along the whole $p_{\mathrm{T}}^{\mathrm{ref}}$ spectrum; the uncertainty increases up to $\approx 1.8\%$ ($\approx 2.2\%$) for the lowest bins, while in the medium and high $p_{\mathrm{T}}^{\mathrm{ref}}$ part of the spectrum is typically smaller than 0.5% (0.6%) for the MPF (DB) method. For the MPF method, the uncertainty is slightly smaller for the PFlow sample.

Impact of the background in the γ+jet sample

The purity of the γ+jet sample was extracted via a 2D-sideband subtraction method using signal leakage fractions from the PYTHIA MC sample (see section 9.1.5). The systematic uncertainty due to the effect of the multijet background in the measured R_{DB} and R_{MPF} responses was determined as

$$\Delta P(i) = (1 - P(i)) \times \frac{R^{\mathrm{LNT}}_{\mathrm{in\,situ}}(i) - R_{\mathrm{in\,situ}}(i)}{R_{\mathrm{in\,situ}}(i)}$$

where $P(i)$ is the signal purity in the $p_{\mathrm{T}}^{\mathrm{ref}}$ bin i, $R_{\mathrm{in\,situ}}(i)$ is the MPF or DB data response measured with the nominal selections in bin i, and $R^{\mathrm{LNT}}_{\mathrm{in\,situ}}(i)$ is the measured response in data for photons that pass the loose′ (LoosePrime4) but at least fail one of the tight identification requirements. The same approach was followed in both the MPF and DB techniques, obtaining a slightly larger difference between the nominal and the background jet responses for the latter. The size of this uncertainty is driven by the value of the purity, which is approximately 75% for the lowest bin and higher or equal to 90% for the second bin and above. The uncertainty is dominant in the two lowest-$p_{\mathrm{T}}^{\mathrm{ref}}$ bins with values similar to 2% (1%) in the first (second) bin and becoming negligible for $p_{\mathrm{T}}^{\mathrm{ref}} \gtrsim 150$ GeV.

Suppression of additional radiation

As explained in section 9.1.2, a γ+jet topology is selected by imposing constraints on the azimuthal angle $\Delta\phi$ between the photon and the jet and by restricting the p_{T} of any subleading jet. The application of these criteria reduce the impact from additional QCD radiation on the momentum balance between the jet and the photon. Systematic uncertainties from the two sources are evaluated through the variation of the $\Delta\phi$ and the jet p_{T} veto requirements. The $\Delta\phi$ selection is varied by ±0.1 around the nominal value of 2.9 (2.8) for the MPF (DB) method (see section 9.1.2). The nominal jet p_{T} veto requirement was tightened to $p_{\mathrm{T}} < \max\left(0.2 \times p_{\mathrm{T}}^{\mathrm{ref}}, 10\right)$ GeV ($p_{\mathrm{T}} < \max\left(0.05 \times p_{\mathrm{T}}^{\mathrm{ref}}, 10\right)$ GeV) and loosened to $p_{\mathrm{T}} < \max\left(0.4 \times p_{\mathrm{T}}^{\mathrm{ref}}, 12\right)$ GeV ($p_{\mathrm{T}} < \max\left(0.15 \times p_{\mathrm{T}}^{\mathrm{ref}}, 20\right)$ GeV) for the MPF (DB) method. The uncertainty on the measured R_{DB} and R_{MPF} responses due to the cut on $\Delta\phi$ was found to be smaller than 0.2%. The uncertainty due to the subleading jet veto is smaller than 0.6% for $p_{\mathrm{T}}^{\mathrm{ref}} < 80$ GeV and smaller than 0.2% in the rest of the spectrum of the DB measurements. For the MPF measurements, the uncertainty due to the subleading jet veto is smaller than 0.2%. These uncertainties were estimated using a bootstrap method and were also rebinned to reduce the effect of statistical fluctuations.

Impact of additional pile-up interactions

Jets produced in additional pile-up interactions are present in both data and simulations and might impact the response measurements. This effect was studied varying the JVT cut applied to the jets in the selection. The nominal JVT requirement was varied to tighter 0.5 (0.91) and looser 0 (0.11) selection values for PFlow (EMtopo) jets. The data-to-MC ratio of responses measured using these varied selections was compared to the nominal to evaluate the systematic uncertainty. The JVT uncertainty was found to be negligible for the MPF method results, while for the DB results using EMtopo jets,

this uncertainty increases up to 0.8% in the lowest-$p_{\mathrm{T}}^{\mathrm{ref}}$ bin, and decreases to $\approx 0.2\%$ in the medium-$p_{\mathrm{T}}^{\mathrm{ref}}$ range. This uncertainty was estimated using a bootstrap method and was also rebinned to reduce the effect of statistical fluctuations.

Systematic uncertainties due to the OOC radiation

For the DB technique, the p_{T} of the leading jet, even if perfectly calibrated, will always tend to be smaller than that of the reference photon due to the OOC radiation. Since the DB is used for cross-checking the MPF calibration results, this uncertainty was not derived as it was done for previous calibrations [24].

In principle, the MPF technique does not depend on the OOC effects because the calorimeter response is integrated over the whole detector. However, two effects related to the OOC contribution must be considered. The showering correction which quantifies the migration of energy across the jet boundary of the reconstructed jet relative to the particle-level jet and the topology correction, which accounts for the difference in calorimeter response in the periphery of the jets compared to that in the core due to the different energy densities and particle compositions. The effect of these two components was found to be less than 2% for jets of $p_{\mathrm{T}} < 50$ GeV and negligible above that. As explained in section 6.2.2, since the previous correction should be applied in both data and simulation, the effect would cancel out on the calibration double ratio. An uncertainty was assigned covering for the deviations of the showering and topology corrections obtained with two different hadronic physics models (FTFP BERT [147] and QGSPBIC [148]).

Statistical uncertainty

The statistical uncertainty on the jet energy scale is estimated for each $p_{\mathrm{T}}^{\mathrm{ref}}$ bin from the error of the mean value from the fit to the Gaussian distribution (see section 9.1.4). The size of the statistical uncertainties increases in the lowest- and highest-$p_{\mathrm{T}}^{\mathrm{ref}}$ bins, becoming dominant and limiting the precision of the γ+jet calibration.

Summary of the JES systematic uncertainties

Figures 9.9 and 9.10 present the JES uncertainties for the previous sources listed above evaluated for both the MPF and DB methods and for the two PFlow and EMtopo jet collections. The total uncertainty is obtained by addition in quadrature of the uncertainties from the different sources. Overall, the DB and MPF methods achieve a similar level of precision for EMtopo jets, while for PFlow jets the precision is slightly higher for the results obtained with the MPF technique. The comparison of the PFlow with the EMtopo uncertainties shows similar results for the DB calibration, while for the MPF, the results obtained for PFlow outperform those obtained with EMtopo jets thanks to the reduction of the uncertainty due to the multijet background in the lower bins and the MC modelling in the medium-$p_{\mathrm{T}}^{\mathrm{ref}}$ range. In the low-$p_{\mathrm{T}}^{\mathrm{ref}}$ bins the uncertainties due to the background from multijet events, the MC modelling and the statistics are dominant, while the photon energy scale dominates in the medium- and high-$p_{\mathrm{T}}^{\mathrm{ref}}$ bins. The MC modelling uncertainty contributes significantly in the medium- and high-$p_{\mathrm{T}}^{\mathrm{ref}}$ part of the

spectrum. The total systematic uncertainty is approximately 1% from 60 GeV to 1.3 TeV approximately, where the precision of the measurements starts to be limited by the size of the statistical uncertainties.

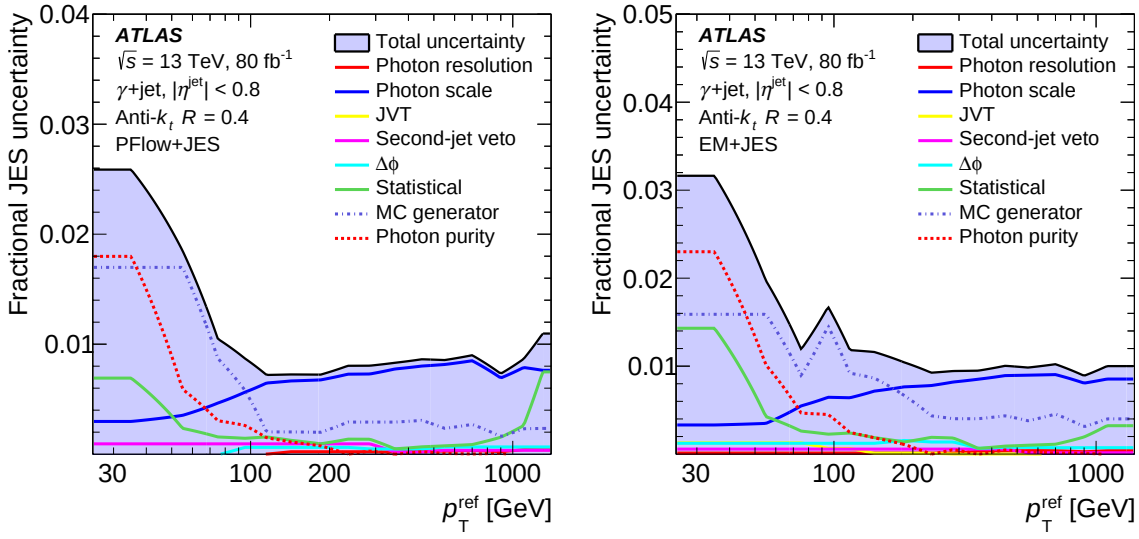


Figure 9.9: Systematic uncertainties on PFlow (left) and EMtopo (right) jets as a function of the reference object p_{T} for γ+jet events calculated using the MPF technique. The total uncertainty is obtained by addition in quadrature of the uncertainties from the different sources.

9.1.8 Closure of the in situ calibration

To study the closure of the in situ calibration, the PFlow DB results were recalculated applying the final calibration of all the in situ methods described in section 6.2.2. This cross-check cannot be performed with the MPF technique since the in situ JES correction is applied as a four-momentum scale factor to jets. Therefore, the measured response in the data after the application of the in situ calibration reflects the response of jets calibrated up to the final PFlow+JES+insitu scale.

Figure 9.11 shows the jet response measured in data after the application of the in situ corrections. The predictions from PYTHIA and SHERPA are also included. The agreement observed between the prediction of PYTHIA and the data is compatible with unity within statistical uncertainties. The differences are smaller in the range where the γ+jet in situ method dominates in the combination (see figure 6.11). The differences observed for the prediction from SHERPA are larger than 1% for $p_{\mathrm{T}}^{\mathrm{ref}} \lesssim 160$ GeV, corresponding to the range where the Z+jet analysis dominates the combination; the calibration factor obtained with the Z+jet analysis is larger than the one obtained with the γ+jet analysis in that range (see figure 6.11).

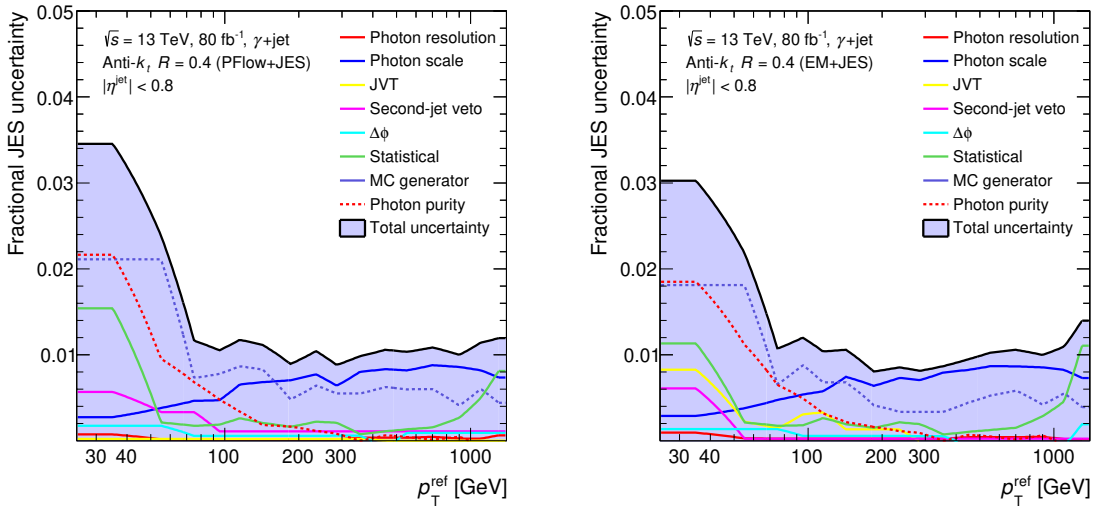


Figure 9.10: Systematic uncertainties on PFlow (left) and EMtopo (right) jets as a function of the reference object p_{T} for γ+jet events calculated using the DB technique. The total uncertainty is obtained by addition in quadrature of the uncertainties from the different sources.

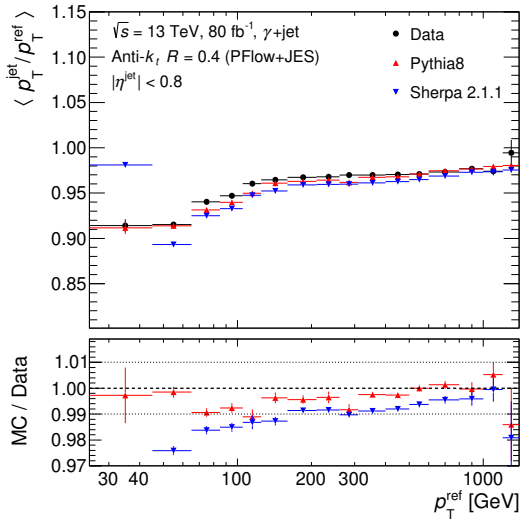


Figure 9.11: Average PFlow jet response, $\left\langle p_{\mathrm{T}}^{\mathrm{j1}}/p_{\mathrm{T}}^{\mathrm{ref,db}} \right\rangle$, as a function of the reference p_{T} for γ+jet events calculated using the DB technique. Jets are calibrated up to the PFlow+JES+insitu scale. The data are represented with black circles while the prediction of PYTHIA (SHERPA) is represented with red (blue) triangles. The ratio of MC to data for both generators is shown in the bottom panel.

9.1.9 Results

The in situ JES calibration using γ+jet events is crucial to calibrate jets with p_{T} in the range between a few hundreds of GeV and the TeV scale. The in situ correction derived with the MPF technique for PFlow jets is at the level of $\approx 2\%$ in most of the $p_{\mathrm{T}}^{\mathrm{ref}}$ range. The uncertainty associated to this correction is $\lesssim 0.9\%$ from approximately 60 GeV to 1400 GeV, increasing up to 2.6% in the lowest-$p_{\mathrm{T}}^{\mathrm{ref}}$ bin. These results constitute an improvement with respect to the previous calibration at $\sqrt{s} = 13$ TeV [25] thanks to the extension of the calibration range coverage in terms of $p_{\mathrm{T}}^{\mathrm{jet}}$ from 950 GeV to 1200 GeV and the improvement in terms of precision, thanks to the reduction of the systematic uncertainties from the use of the MPF method and PFlow jets for the calibration.

A calibration was derived for EMtopo jets using also the MPF method obtaining the same p_{T}-range coverage and size for the in situ correction of $\approx 2\%$ in most of the $p_{\mathrm{T}}^{\mathrm{ref}}$ range. The total systematic uncertainty is $\sim 1.0\%$ from approximately 100 GeV to 1400 GeV, increasing up to 3.2% in the lowest-$p_{\mathrm{T}}^{\mathrm{ref}}$ bin.

The MPF results were cross-checked using the DB method. The calibration correction factors and their uncertainties were found to be consistent with the nominal ones, providing a consistent and complementary measurement of the jet energy scale.

9.2 In situ JER calibration using γ+jet events

In this section, the methodology followed to derive the in situ JER calibration with the DB technique using γ+jet events is described. This calibration was developed at a later stage compared to that of the JES, described in section 9.1, and was derived for the 2015−2016, 2017 and 2018 periods, separately. This is a preliminary analysis, which was not finalized and incorporated to the official ATLAS jet energy resolution calibration, however, the descriptions provided in this section give a reasonable overview of the methods needed to measure the JER in situ using events in which a jet recoils against a well-measured photon.

9.2.1 Strategy

The in situ techniques described in section 6.2.2 (Z/γ+jet balance and MJB), can be used to measure the JER of the detector covering a similar range in $p_{\mathrm{T}}^{\mathrm{jet}}$ as achieved for the JES calibration. These methods were used in ATLAS for the JER calibration with 8 TeV [25] data. As described in section 6.3.3, the free parameters N, S, and C, can be obtained by performing a fit to the in situ measurements. The following sections describe the estimation of the JER using the γ+jet DB technique.

9.2.2 Data selection

For the set of results presented in this chapter in section 9.2.5, the data corresponding to the 2018 period was used. In the collected data sample, the average number of pp interactions per bunch crossing is ≈ 36.7, while the total integrated luminosity is $58.5\,\text{fb}^{-1}$ after requiring that all detector subsystems were operational during data recording.

The same selection applied in the DB technique for the JES calibration is used (see section 9.1.2), except for the fact that this calibration was derived in two different regions of $|\eta^{\text{jet}}|$, namely $|\eta^{\text{jet}}| < 0.2$ and $0.2 < |\eta^{\text{jet}}| < 0.7$, to be consistent with the selection used in the resolution measurement using dijet events (see section 6.3.1).

9.2.3 Monte Carlo simulations

Samples of MC events from PYTHIA and SHERPA 2.2.2 were used for this calibration. For the set of results presented in section 9.2.5, the MC campaign e of both MC generators was used. Additional pile-up and photon identification corrections were also applied to the simulated samples (see section 5.6).

9.2.4 The direct balance method applied to the JER

The DB method used to measure the JER is an extension of the previously described for the measurement of the JES in section 9.1.1. In this extension, instead of determining the response R_{DB} from the mean of the balance distribution ($p_{\text{T}}^{\text{j1}}/p_{\text{T}}^{\text{ref,db}}$) within each $p_{\text{T}}^{\text{ref}}$ bin, the width $\sigma_{\text{DB}}^{\text{reco}}$ is extracted as the standard deviation of the same Gaussian fit. The relative JER, σ_{E}/E, is then estimated using

$$\frac{\sigma_{\text{E}}}{E} = \frac{\sigma_{\text{pT}}}{p_{\text{T}}} = \sigma_{\text{DB}}^{\text{reco}} \ominus \sigma_{\text{DB}}^{\text{truth}} \,. \tag{9.1}$$

The first equality holds to a good approximation since the contribution from the angular resolution is negligible, and the second follows from the same reasoning as in equation 6.7. The fact that the detector resolution of the reference object is negligible compared with that of the jet, simplifies the method (see equation 6.6). The parameter $\sigma_{\text{DB}}^{\text{truth}}$ is obtained using a fit to the $p_{\text{T}}^{\text{truth}}/p_{\text{T}}^{\text{ref,db}}$ distribution extracted using MC simulation. For each simulated event, $p_{\text{T}}^{\text{truth}}$ is defined from the particle-level jet that is ghost-matched[2] to

[2]Jet matching using ghost association [138] treats each MC particle as a *ghost particle*, meaning that they are assigned an infinitesimal p_{T}, leaving the angular coordinates unchanged. The calorimeter jets can be built using both the topo-clusters and ghost-particles as input. Since the ghost particles have infinitesimal p_{T}, the four-momenta of the reconstructed jets will be identical to that of the original jets built only from topo-clusters, but the new jets will also have a list of associated truth particles for any given reconstructed jet. A truth-particle jet is matched to a reconstructed jet if the sum of energies of the truth-particle jet constituents which are ghost-associated with the reconstructed jet is more than 70% of

the leading reconstructed jet. The simulated JER is also extracted from the MC samples with fits to the $p_{\mathrm{T}}^{\mathrm{j1}}/p_{\mathrm{T}}^{\mathrm{truth}}$ distribution.

The relative JER, denoted as σ_{insitu}, which is introduced in equation 9.1, is measured as

$$\sigma_{\mathrm{in\,situ}} = \sqrt{\sigma\left(\frac{p_{\mathrm{T}}^{\mathrm{j1}}}{p_{\mathrm{T}}^{\mathrm{ref,db}}}\right)^{2}_{\mathrm{data}} - \sigma\left(\frac{p_{\mathrm{T}}^{\mathrm{truth}}}{p_{\mathrm{T}}^{\mathrm{ref,db}}}\right)^{2}_{\mathrm{MC}}}\,, \tag{9.2}$$

where the first term corresponds to the standard deviation of the fit to the balance distribution from the data and the second term corresponds to the standard deviation of the fit to the $p_{\mathrm{T}}^{\mathrm{truth}}/p_{\mathrm{T}}^{\mathrm{ref,db}}$ distribution. Measuring σ_{DB} separately for detector and particle-level jets and subtracting the truth quantity from the observed one, makes possible to obtain the standard deviation due to detector smearing. The relative JER can be also obtained using the in situ technique (9.1) applied to simulated events

$$\sigma_{\mathrm{in\,situ}}^{\mathrm{MC}} = \sqrt{\sigma\left(\frac{p_{\mathrm{T}}^{\mathrm{j1}}}{p_{\mathrm{T}}^{\mathrm{ref,db}}}\right)^{2}_{\mathrm{MC}} - \sigma\left(\frac{p_{\mathrm{T}}^{\mathrm{truth}}}{p_{\mathrm{T}}^{\mathrm{ref,db}}}\right)^{2}_{\mathrm{MC}}}\,.$$

This measured value for the JER in simulated events is compared to the expected MC resolution

$$\sigma^{\mathrm{MC}} = \sigma\left(\frac{p_{\mathrm{T}}^{\mathrm{j1}}}{p_{\mathrm{T}}^{\mathrm{truth}}}\right)\,. \tag{9.3}$$

The difference between the measured and expected JER is taken as a *non-closure* systematic uncertainty

$$\Delta\sigma_{\mathrm{closure}} = \frac{\sigma^{\mathrm{MC}} - \sigma_{\mathrm{in\,situ}}^{\mathrm{MC}}}{\sigma_{\mathrm{in\,situ}}}\,. \tag{9.4}$$

9.2.5 Measurement of the JER with 58.5 fb^{-1} of ATLAS data

The comparison of the expected JER resolution in simulated events for PFlow and EMtopo jets is shown as a function of $p_{\mathrm{T}}^{\mathrm{ref}}$ in figure 9.12. The resolution improves with increasing $p_{\mathrm{T}}^{\mathrm{ref}}$, from $\approx 13\% - 19\%$ to $\approx 3\% - 3.5\%$ in the two $|\eta^{\mathrm{jet}}|$ regions and for the two jet collections; this is expected, since at higher energies, a larger fraction of jet radiation is captured inside the jet cone. The differences observed in the resolution comparing

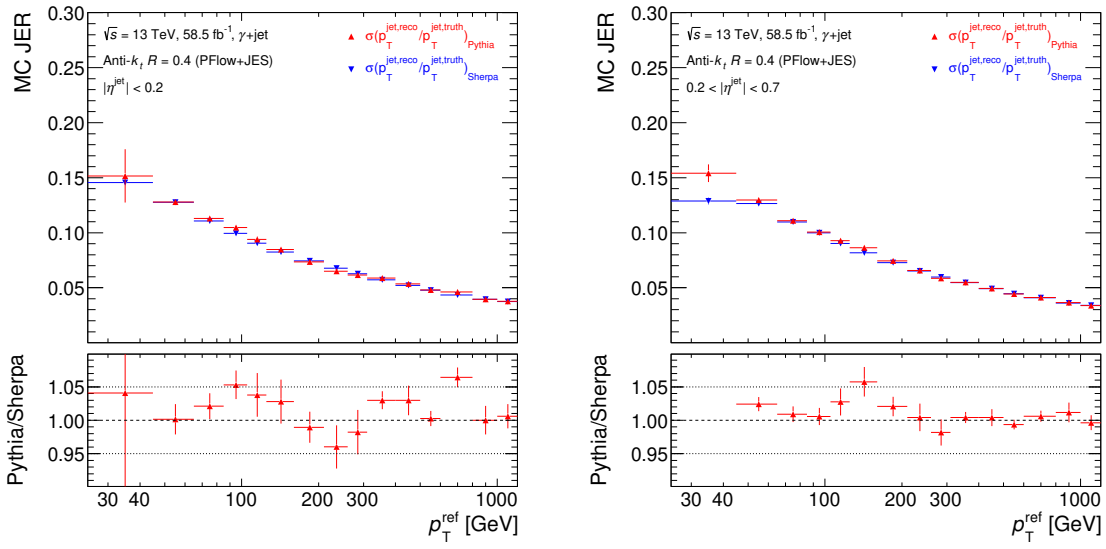
the truth-particle jet energy. If several truth-particle jets fulfil the matching requirement, the truth-jet with the largest energy is chosen as the matched jet.

PYTHIA and SHERPA 2.2.2 are smaller than 5% or compatible with 5% within uncertainties, except for the first bin of $p_{\mathrm{T}}^{\mathrm{ref}}$. This differences between both MC generators are somewhat larger for $|\eta^{\mathrm{jet}}| < 0.2$.

The MC resolution obtained either using PFlow or EMtopo jets is also compared in figure 9.13. An improvement is observed for PFlow jets for $p_{\mathrm{T}}^{\mathrm{ref}} \lesssim 85$ GeV; this is attributed to the better pile-up rejection thanks to the use of tracking information. In the middle range of the spectrum, the resolution obtained for EMtopo jets is smaller than that of PFlow, and at high $p_{\mathrm{T}}^{\mathrm{ref}}$, the resolution obtained tends to be similar for both jet collections.

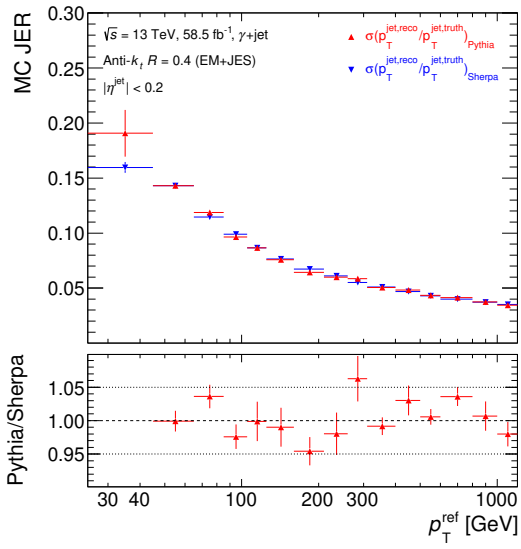
The two terms which enter equation 9.2 to measure the in situ JER in data are shown in Figure 9.14. In this figure, the standard deviations of the $p_{\mathrm{T}}^{\mathrm{j1}}/p_{\mathrm{T}}^{\mathrm{ref,db}}$ and $p_{\mathrm{T}}^{\mathrm{truth}}/p_{\mathrm{T}}^{\mathrm{ref,db}}$ distributions are obtained using data and MC, respectively. Differences between PFlow and EMtopo jets are observed; the improvement in the resolution at low $p_{\mathrm{T}}^{\mathrm{ref}}$ for PFlow is consistent with that observed in Figure 9.13. The resolution obtained with PYTHIA is consistently smaller than that of SHERPA 2.2.2 except at high $p_{\mathrm{T}}^{\mathrm{ref}}$.

The JER is measured in situ by subtracting quadratically the quantity $\sigma\left(p_{\mathrm{T}}^{\mathrm{truth}}/p_{\mathrm{T}}^{\mathrm{ref,db}}\right)$ from $\sigma\left(p_{\mathrm{T}}^{\mathrm{j1}}/p_{\mathrm{T}}^{\mathrm{ref,db}}\right)$ (see equation 9.2). Figure 9.15 shows the nominal JER, computed with PYTHIA; the alternative estimation using SHERPA 2.2.2 is also shown. The difference between these two JER estimations is taken as the MC generator systematic uncertainty. The in situ JER fully estimated with PYTHIA is also shown in figure 9.15, and is used to compute the non-closure systematic uncertainty (see equation 9.4). Large differences are observed between σ^{MC} and $\sigma^{\mathrm{MC}}_{\mathrm{in\,situ}}$, reaching up to $\approx 40\%$ at low and high $p_{\mathrm{T}}^{\mathrm{ref}}$. The comparison between the in situ JER measured in data using PFlow or EMtopo jets is provided in figure 9.16, showing the expected improvement obtained with PFlow at low $p_{\mathrm{T}}^{\mathrm{ref}}$ and also the slightly better performance of EMtopo in the middle range of the spectrum. Hence, these results are consistent with those previously shown in figures 9.13 and 9.14

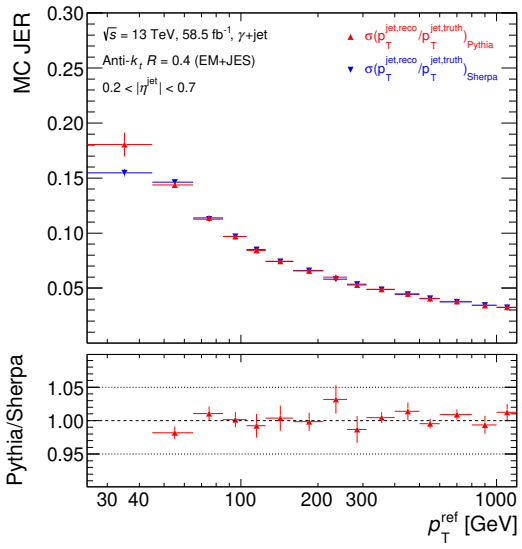


(a)

(b)



(c)



(d)

Figure 9.12: MC jet energy resolution measured for PFlow (top) and EMtopo (bottom) jets in the regions $|\eta^{\text{jet}}| < 0.2$ (left) and $0.2 < |\eta^{\text{jet}}| < 0.7$ (right). The resolution obtained with PYTHIA (SHERPA 2.2.2) is represented with red (blue) triangles. The bottom part of each figure shows the ratio of PYTHIA to SHERPA 2.2.2. The error bars show the statistical uncertainties.

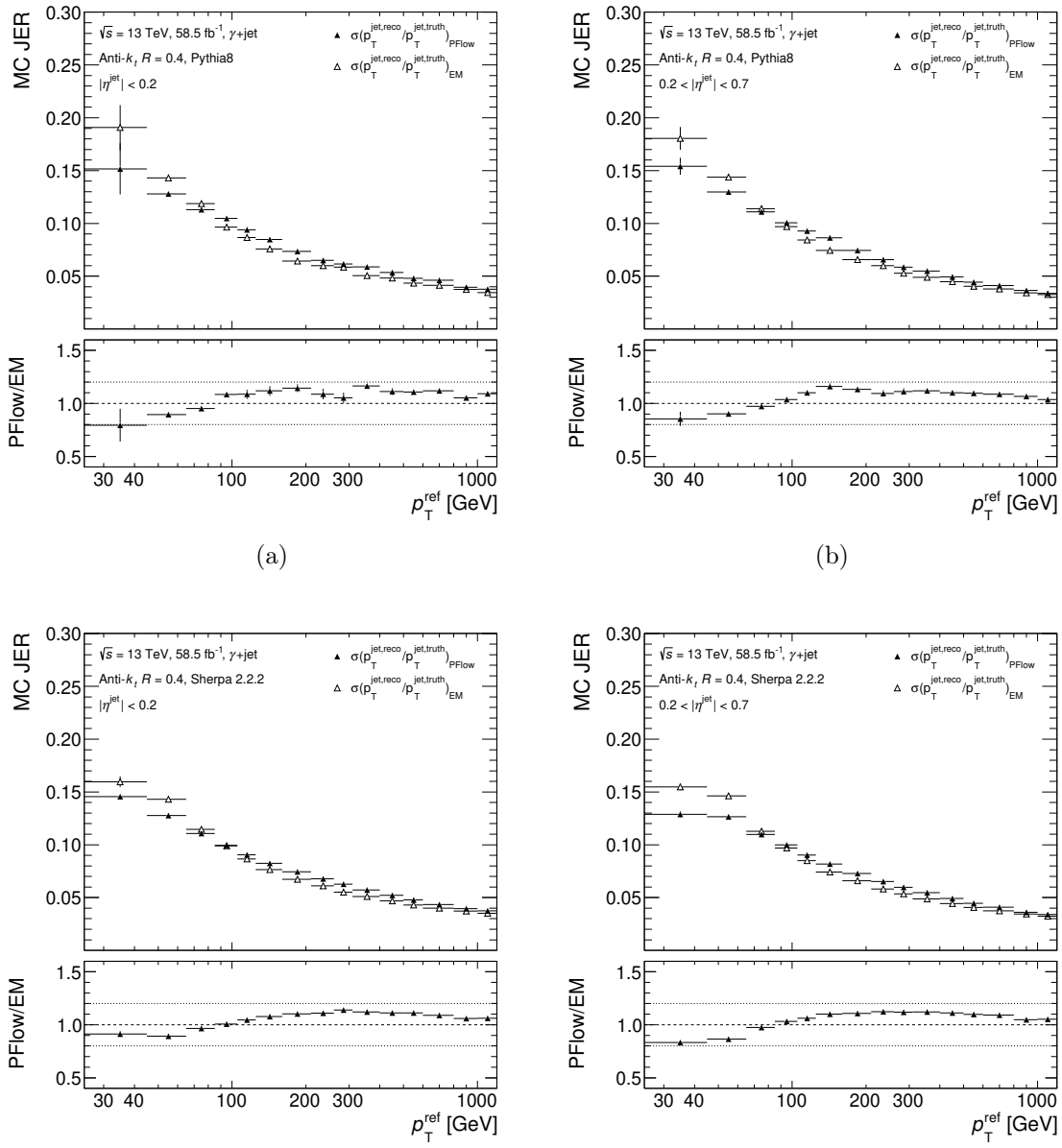


Figure 9.13: Jet energy resolution measured with PYTHIA (top) and SHERPA 2.2.2 (bottom) in the regions $|\eta^{\text{jet}}| < 0.2$ (left) and $0.2 < |\eta^{\text{jet}}| < 0.7$ (right). The resolution obtained using PFlow (EMtopo) jets is represented with solid (empty) markers. The bottom part of each figure shows the ratio of PFlow to EMtopo. The error bars show the statistical uncertainties.

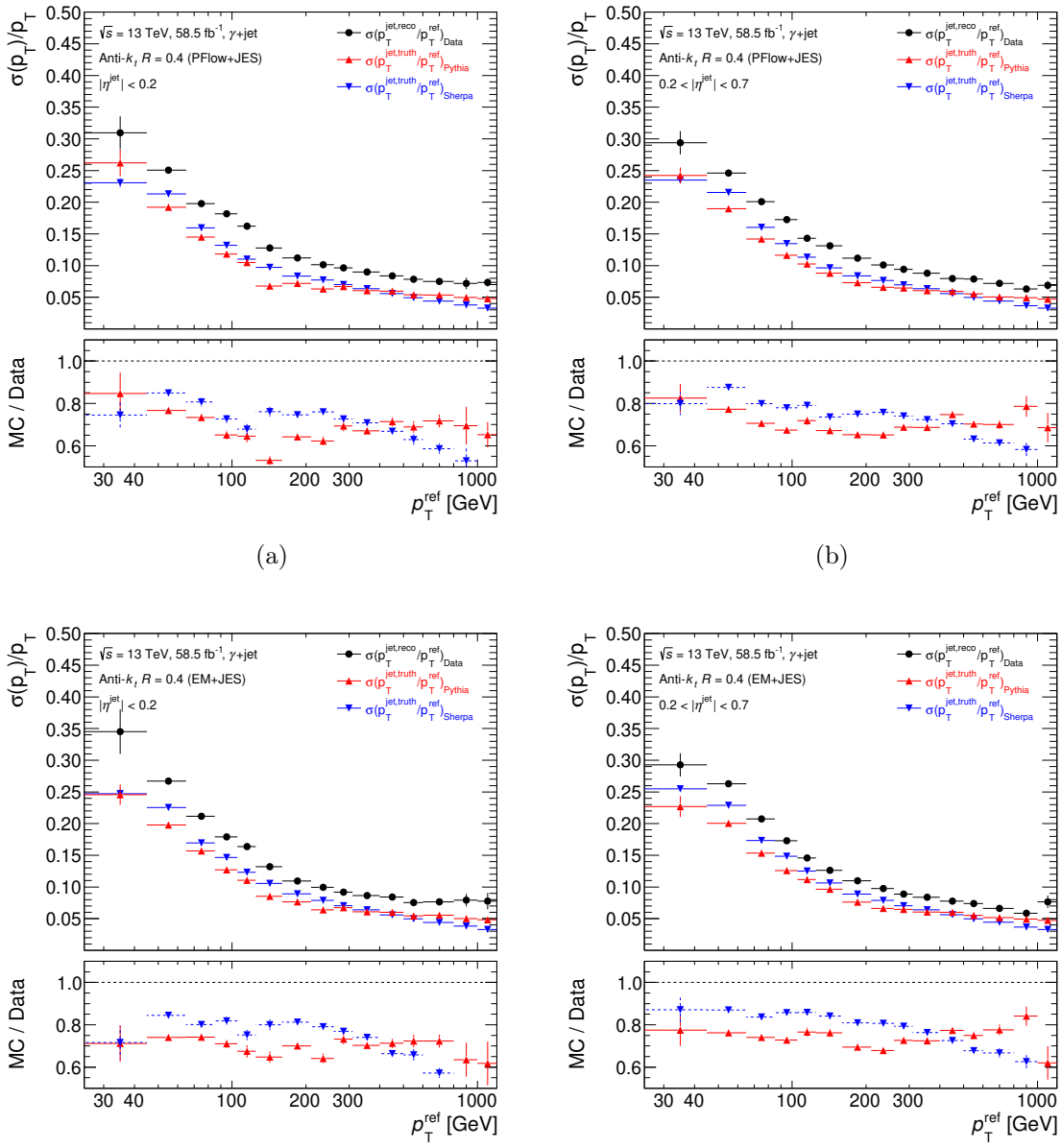


Figure 9.14: The standard deviations of the $p_{\mathrm{T}}^{\mathrm{j1}}/p_{\mathrm{T}}^{\mathrm{ref,db}}$ (black dots) and $p_{\mathrm{T}}^{\mathrm{truth}}/p_{\mathrm{T}}^{\mathrm{ref,db}}$ for PYTHIA (red triangles) and SHERPA 2.2.2 (blue triangles) distributions obtained for PFlow (top) and EMtopo (bottom) jets as a function of $p_{\mathrm{T}}^{\mathrm{ref}}$ in the two $|\eta^{\mathrm{jet}}|$ ranges measured. The widths of the distributions are estimated with Gaussian fits. The bottom graph shows the ratio of MC to data. Error bars show the statistical uncertainties only.

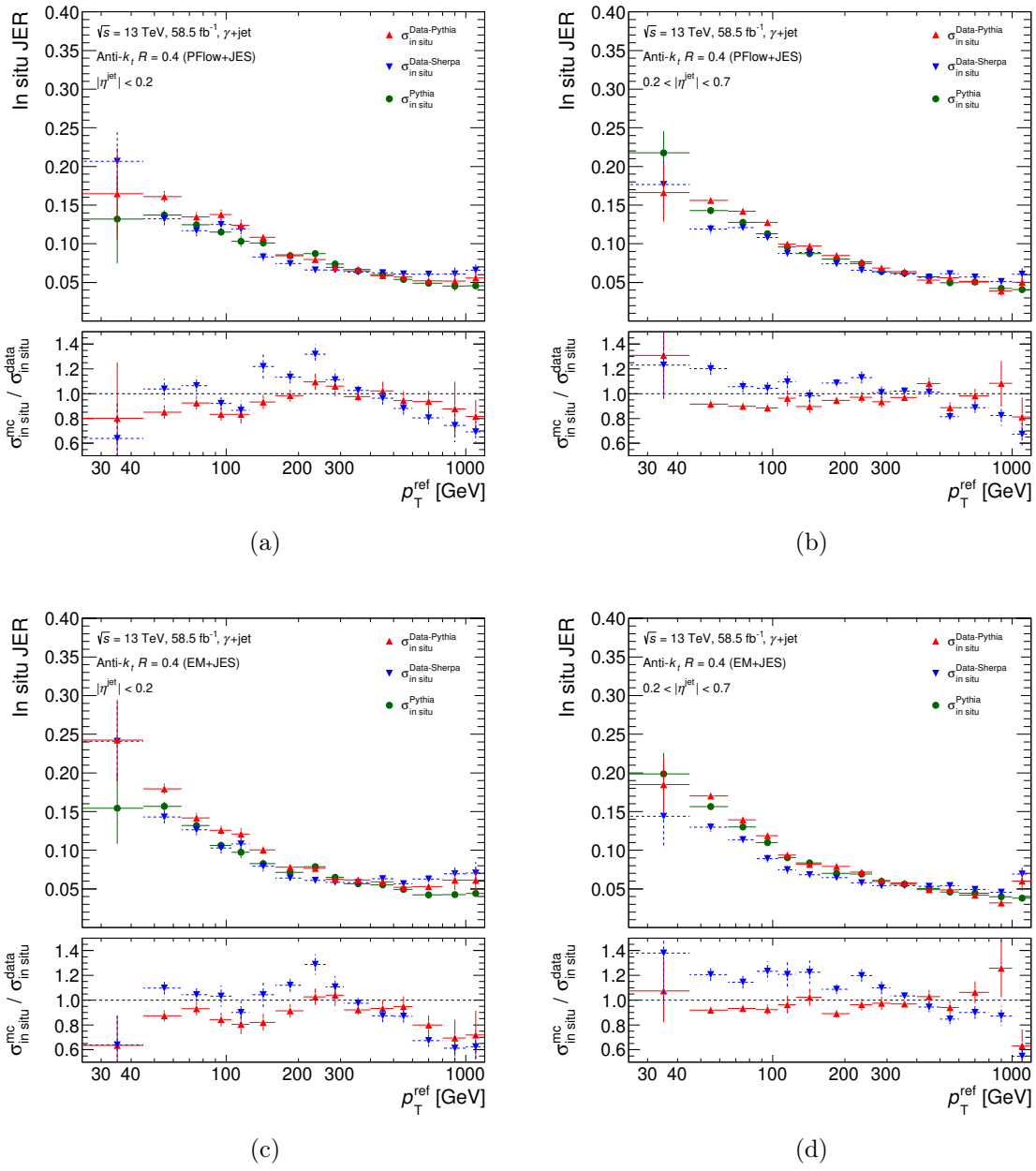
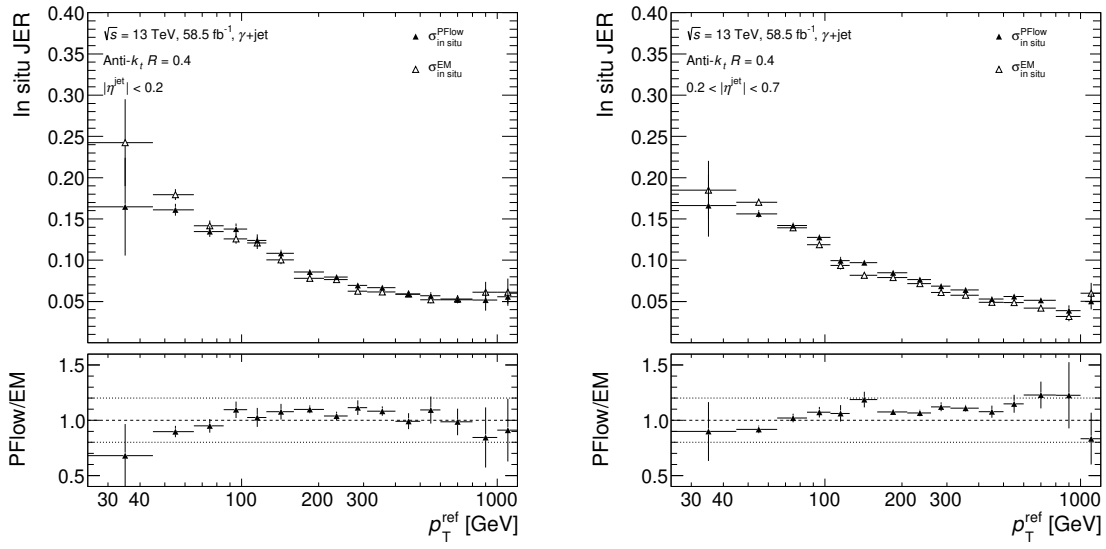


Figure 9.15: Jet energy resolution measured as a function of the reference p_{T} in two $|\eta^{\mathrm{jet}}|$ ranges using PFlow (top) and EMtopo (bottom) jets in γ+jet events with the DB technique. Results obtained using 2018 data and PYTHIA (red triangles), 2018 data and SHERPA 2.2.2 (blue triangles), and the full-MC in situ estimation with PYTHIA (green dots) are shown. The widths of the distributions are estimated with Gaussian fits. The bottom part shows the ratio of the JER computed with PYTHIA to the JER obtained from data and MC (PYTHIA or SHERPA 2.2.2). Error bars show the statistical uncertainties only.



(a) (b)

Figure 9.16: Jet energy resolution measured in data as a function of the reference p_{T} in two $|\eta^{\mathrm{jet}}|$ ranges using PFlow (solid triangles) and EMtopo (empty triangles) jets in γ+jet events with the DB technique. The estimation is obtained using 2018 data and PYTHIA MC. The widths of the distributions are estimated with Gaussian fits. The bottom part shows the ratio of PFlow to EMtopo. Error bars show the statistical uncertainties only.

Systematic uncertainties

The uncertainties on the JER estimated with the DB technique for both PFlow and EMtopo collections are, to a large extent, similar to those uncertainties considered for the JES analysis (see section 9.1.7). The sources of uncertainty that affect the measurement of $\sigma_{\mathrm{in\,situ}}$ (see equation 9.2) include the non-closure, photon energy scale and resolution, MC generator, impact of the background of jets faking photons, suppression of additional radiation and the impact of additional pile-up interactions. Each of these sources of uncertainty was discussed in section 9.1.7 for the JES measurement, and were derived for the JER analogously. For some of the systematics, a rebinning procedure is applied to reduce the effect of statistical fluctuations. The same rebinning procedure based on the bootstrap method previously described in section 9.1.7, is used. The status of this analysis is preliminary, and the purity uncertainty that should be accounted for, is not included in this version.

The total relative JER uncertainty is obtained by adding in quadrature the individual components listed above and the statistical uncertainty of the measurement. The JER uncertainties are shown in figures 9.17 and 9.18 for PFlow and EMtopo jets, respectively. The total JER uncertainty is in the range $15\% - 40\%$ ($15\% - 35\%$) for PFlow (EMtopo) jets in the $|\eta^{\mathrm{jet}}| < 0.2$ region, being $\approx 15\%$ from 45 GeV to 600 GeV, approximately

(except around 250 GeV, where the non-closure uncertainty rises up to almost 24% for PFlow and 20% for EMtopo jets). The dominant systematic uncertainties are those due to the MC generator, the non-closure and the statistical uncertainty. In the region $0.2 < |\eta^{\text{jet}}| < 0.7$, the total JER uncertainty is in the range $12\% - 47\%$ $(16\% - 32\%)$ for PFlow (EMtopo) jets, being $\approx 15\%$ from 45 GeV to 600 GeV for PFlow and $\approx 17\%$ from 150 GeV to 600 GeV for EMtopo, due to the significant increase of the MC generator uncertainty. The main uncertainties are the MC generator, the non-closure and the statistical uncertainty. Except at low $p_{\text{T}}^{\text{ref}}$, the PFlow total uncertainty is below that of EMtopo due to the smaller MC generator uncertainty. Overall, the size of the total uncertainty is similar between $|\eta^{\text{jet}}|$ regions; a small increase is observed in the region $|\eta^{\text{jet}}| < 0.2$ for PFlow, due to the increase of the non-closure uncertainty, whereas for EMtopo, the total uncertainty in the region $0.2 < |\eta^{\text{jet}}| < 0.7$ increases due to the larger MC generator uncertainty. For both $|\eta^{\text{jet}}|$ regions, the precision of the measurement is limited at low and high $p_{\text{T}}^{\text{ref}}$ due to the size of the statistical uncertainty.

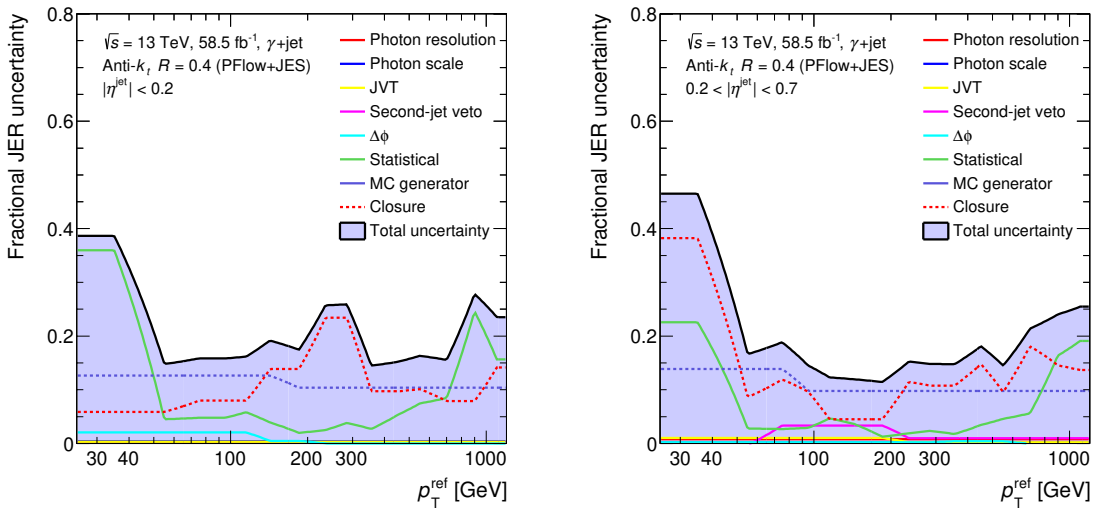


Figure 9.17: Relative systematic uncertainties on PFlow jets as a function of the reference object p_{T} in the $|\eta^{\text{jet}}| < 0.2$ (left) and $0.2 < |\eta^{\text{jet}}| < 0.7$ (right) ranges. The total uncertainty is obtained by addition in quadrature of the uncertainties from the different sources.

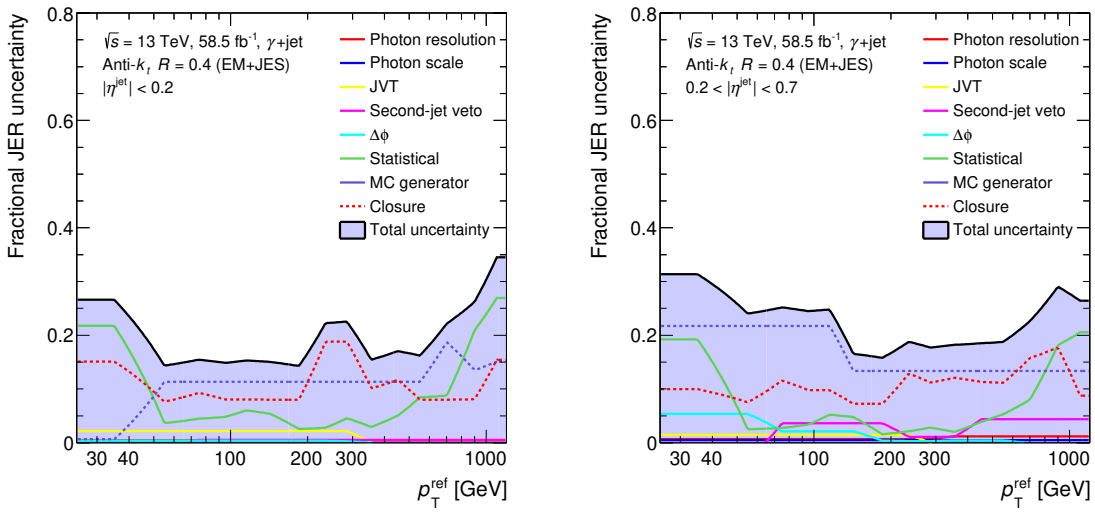


Figure 9.18: Relative systematic uncertainties on EMtopo jets as a function of the reference object p_{T} in the $|\eta^{\mathrm{jet}}| < 0.2$ (left) and $0.2 < |\eta^{\mathrm{jet}}| < 0.7$ (right) ranges. The total uncertainty is obtained by addition in quadrature of the uncertainties from the different sources.

Results

The in situ JER calibration using γ+jet events is used to measure the JER given its simplicity compared to other methods using dijet events [24,26]. The measurement of the JER presented in this section was performed using 2018 data with PFlow and EMtopo jets in two different regions of pseudorapidity of the jet. A resolution smaller than 17% (24%) was obtained for PFlow (EMtopo) jets in the lowest $p_{\mathrm{T}}^{\mathrm{ref}}$ bin from 25 GeV to 45 GeV in the $|\eta^{\mathrm{jet}}| < 0.2$ region; in the $0.2 < |\eta^{\mathrm{jet}}| < 0.7$ range, the resolution obtained in the same bin is smaller than 17% (19%) for PFlow (EMtopo) jets. The PFlow results show an improvement in the jet energy resolution at low $p_{\mathrm{T}}^{\mathrm{ref}}$, whereas EMtopo performs slightly better than PFlow at high $p_{\mathrm{T}}^{\mathrm{ref}}$. In the highest $p_{\mathrm{T}}^{\mathrm{ref}}$ bin measured, from 1 TeV to 1.2 TeV, a resolution of approximately $4\% - 5\%$ is obtained in all the measured configurations. The JER experimental uncertainty is dominated by the MC generator and the non-closure uncertainties; the total uncertainty is approximately $25\% - 45\%$ below 45 GeV, approximately $15\% - 20\%$ in the range from 45 GeV to 600 GeV, and it increases almost up to 35% at 1.2 TeV.

The JER was measured in 2012 using several in situ methods [24], with a combination of three in situ measurements, namely Z+jet, γ+jet and dijet analyses, as a function of jet p_{T} and η_{det}, in a similar way as for the JES estimation. The latest measurement of the JER published by ATLAS [26] only made use of the dijet balance analysis, given the better performance obtained with respect to the Z+jet and γ+jet analyses.

Chapter 10

Measurement of the isolated photon plus jet cross section in pp collisions at $\sqrt{s} = 13$ TeV using $138\,\text{fb}^{-1}$ with the ATLAS detector

In this chapter, the measurement of the differential cross section for isolated-photon plus jet production in pp collisions at $\sqrt{s} = 13$ TeV using $138.4\,\text{fb}^{-1}$ of ATLAS data is presented. This measurement is currently ongoing. The differential cross section is measured as a function of the photon transverse energy for three different rapidity ranges of the leading jet: $|y^{\text{jet}}| < 1.2$, $1.2 \leq |y^{\text{jet}}| < 2.8$, and $2.8 \leq |y^{\text{jet}}| < 4.4$. For each rapidity region, the same-sign and opposite-sign cases are studied separately. NLO pQCD calculations from SHERPA are compared with the measurement, using different parameterisations of the proton parton distribution functions. The predictions provide a good description of the data within the experimental and theoretical uncertainties, except in the $2.8 \leq |y^{\text{jet}}| < 4.4$ region for the opposite-sign configuration, where the theoretical predictions tend to overestimate the measured cross section. This measurements benefits with respect to those published earlier from the large increase of the total integrated luminosity and reduced systematic uncertainties.

10.1 Data selection

The data used in this analysis were collected with the ATLAS detector during the proton-proton collision running periods 2015 to 2018, when the LHC operated at a centre-of-mass energy of $\sqrt{s} = 13$ TeV. Only events taken during stable beam conditions and satisfying detector- and data-quality requirements, which include the calorimeters and inner tracking detectors being in nominal operation, are considered. In the collected data sample, the average number of pp interactions per bunch crossing is ~ 34, while the

total integrated luminosity is $138.4 \pm 2.4\,\text{fb}^{-1}$. The difference in integrated luminosity with respect to that used in the full Run II inclusive photon analysis (see section 8.1) comes from the need to apply an additional jet cleaning cut that removes $\approx 1.66\%$ of the 2016 integrated luminosity. The uncertainty in the combined $2015-2018$ luminosity is 1.7% [150], obtained using the LUCID-2 detector [151] for the primary luminosity measurements.

Trigger requirements

The data sample used consists of events triggered in the same conditions described in section 8.1.

Event selection

The sample of isolated-photon plus jet events was selected offline using similar criteria to those summarised in section 8.1.

Photon selection

Photon candidates were reconstructed (see section 5.2) and calibrated (see section 5.3) in the same conditions as those described in section 8.1.

The photon-candidate selection criteria proceeds in a similar way as described in section 8.1. The same $E_{\text{T}}^{\text{iso}}$ cut is used, but only one sample, using a cone of $\Delta R = 0.4$, is selected.

Jet selection

Jets were reconstructed (see section 6.1) using the anti-k_{t} algorithm [63, 140] with a radius parameter $R = 0.4$, using topoclusters as input. The calorimeter cell energies are measured at the EM scale, corresponding to the energy deposited by electromagnetically interacting particles. The jets are then calibrated using the methods described in sections 6.2 and 6.3), and these jets are referred to as detector-level jets. The jet reconstruction procedure applied to the topoclusters was also applied to the generated stable particles, excluding muons and neutrinos; these jets are referred to as particle-level jets.

Jets reconstructed from calorimeter signals not originating from a pp collision were rejected by applying jet-quality criteria [145]. These criteria suppress spurious jets from electronic noise in the calorimeter, cosmic rays and beam-related backgrounds. Remaining jets are required to have calibrated transverse momenta greater than 60 GeV and rapidity $|y^{\text{jet}}| < 4.4$. Jets overlapping with the photon candidate are not considered if the jet axis lies within a cone of size $\Delta R = 0.8$ around the photon; this requirement prevents any overlap between the photon isolation cone ($\Delta R = 0.4$) and the jet cone ($\Delta R = 0.4$). Finally, the event is retained if the leading jet has $p_{\text{T}}^{\text{jet}} > 165$ GeV.

The total number of data events selected using the requirements discussed above is 2 951 643. This sample of events is used to measure the cross section as a function of E_{T}^{γ} in the different configurations depending on the $|y^{\text{jet}}|$ range and for both same-

sign (SS) and opposite-sign (OS) cases. The kinematic requirements and the number of selected events in each of the jet and photon angular configurations are summarised in table 10.1. Figure 10.1 shows the E_{T}^{γ}, $|\eta^{\gamma}|$, ϕ^{γ}, $p_{\mathrm{T}}^{\mathrm{jet}}$, $|y^{\mathrm{jet}}|$ and ϕ^{jet} distributions for data, PYTHIA, and SHERPA MC generators. Figure 10.2 shows the distributions for the absolute value of the difference in rapidity between the leading photon and the leading jet ($|\Delta y^{\gamma-\mathrm{jet}}|$), the azimuthal angle between the leading photon and the leading jet ($\Delta\phi^{\gamma-\mathrm{jet}}$), the invariant mass of the leading-photon and leading-jet system ($m^{\gamma-\mathrm{jet}}$), and $|\cos\theta^{*}|$ ($\tanh\left((\eta^{\gamma}-y^{\mathrm{jet}})/2\right)$). For the two latter observables, the additional kinematic constraints $|\eta^{\gamma}+y^{\mathrm{jet}}|<2.37$, $|\cos\theta^{*}|<0.83$ and $m^{\gamma-\mathrm{jet}}>900$ GeV are imposed to remove the bias due to the rapidity and transverse-momentum requirements on the photon and the jet [31]. The first two constraints avoid the bias induced by requirements on η^{γ} and y^{jet}, yielding slices of $\cos\theta^{*}$ with the same length along the $\eta^{\gamma}+y^{\mathrm{jet}}$ axis; the third constraint avoids the bias due to the $E_{\mathrm{T}}^{\gamma}>250$ GeV requirement. The number of events selected in the data after these additional requirements is 337 683.

Figures 10.3 and 10.4 represent two of the selected events in the 2016 and 2017 datasets, respectively. Figure 10.3 shows a typical photon-plus-jet event classified in the central region $|y^{\mathrm{jet}}|<1.2$, where the leading photon and the leading jet are in the same-sign configuration. On the other hand, figure 10.4 shows an event classified in the forward region $1.2\leq|y^{\mathrm{jet}}|<2.8$, where the leading photon and the jet are in the opposite-sign configuration. In each figure, the leading tight photon is represented by the pink line, whereas the leading (subleading) jet, reconstructed using the anti-k_{t} algorithm with $R=0.4$, is represented by the transparent-green (blue) area.

	Phase space region	
Requirements on photon	$E_{\mathrm{T}}^{\gamma}>250$ GeV, $\lvert\eta^{\gamma}\rvert<2.37$ (excluding $1.37<\lvert\eta^{\gamma}\rvert<1.56$) $E_{\mathrm{T}}^{\mathrm{iso}}<4.2\times10^{-3}\times E_{\mathrm{T}}^{\gamma}+4.8$ GeV	
Requirements on jet	Leading jet using anti-k_{t} algorithm with $R=0.4$	
	$p_{\mathrm{T}}^{\mathrm{jet}}>165$ GeV, $\lvert y^{\mathrm{jet}}\rvert<4.4$, $\Delta R^{\gamma-jet}>0.8$	
(Additional constraints)	$\lvert\eta^{\gamma}+y^{\mathrm{jet}}\rvert<2.37$, $\lvert\cos\theta^{*}\rvert<0.83$, $m^{\gamma-\mathrm{jet}}>900$ GeV	
Sample	Same sign ($\eta^{\gamma}\cdot y^{\mathrm{jet}}\geq0$)	Opposite sign ($\eta^{\gamma}\cdot y^{\mathrm{jet}}<0$)
Region $\lvert y^{\mathrm{jet}}\rvert<1.2$	1 197 991	732 086
Region $1.2\leq\lvert y^{\mathrm{jet}}\rvert<2.8$	774 809	222 599
Region $2.8\leq\lvert y^{\mathrm{jet}}\rvert<4.4$	17 918	6 240
Total sample	2 951 643	
Total sample + additional cuts	337 683	

Table 10.1: Definition of the phase-space region for the measurement and predictions. The number of data events selected in each phase-space configuration is also shown.

As previously discussed in section 10.1, the average number of interactions per bunch crossing for the different data-taking periods is 23.5, 37.6 and 36.7 for the data collected with the HLT_g140_loose trigger in the 2015 − 2016, 2017, and 2018 years, respectively. The differences in the pile-up conditions between data and MC simulation were accounted for separately for each of the periods before their combination.

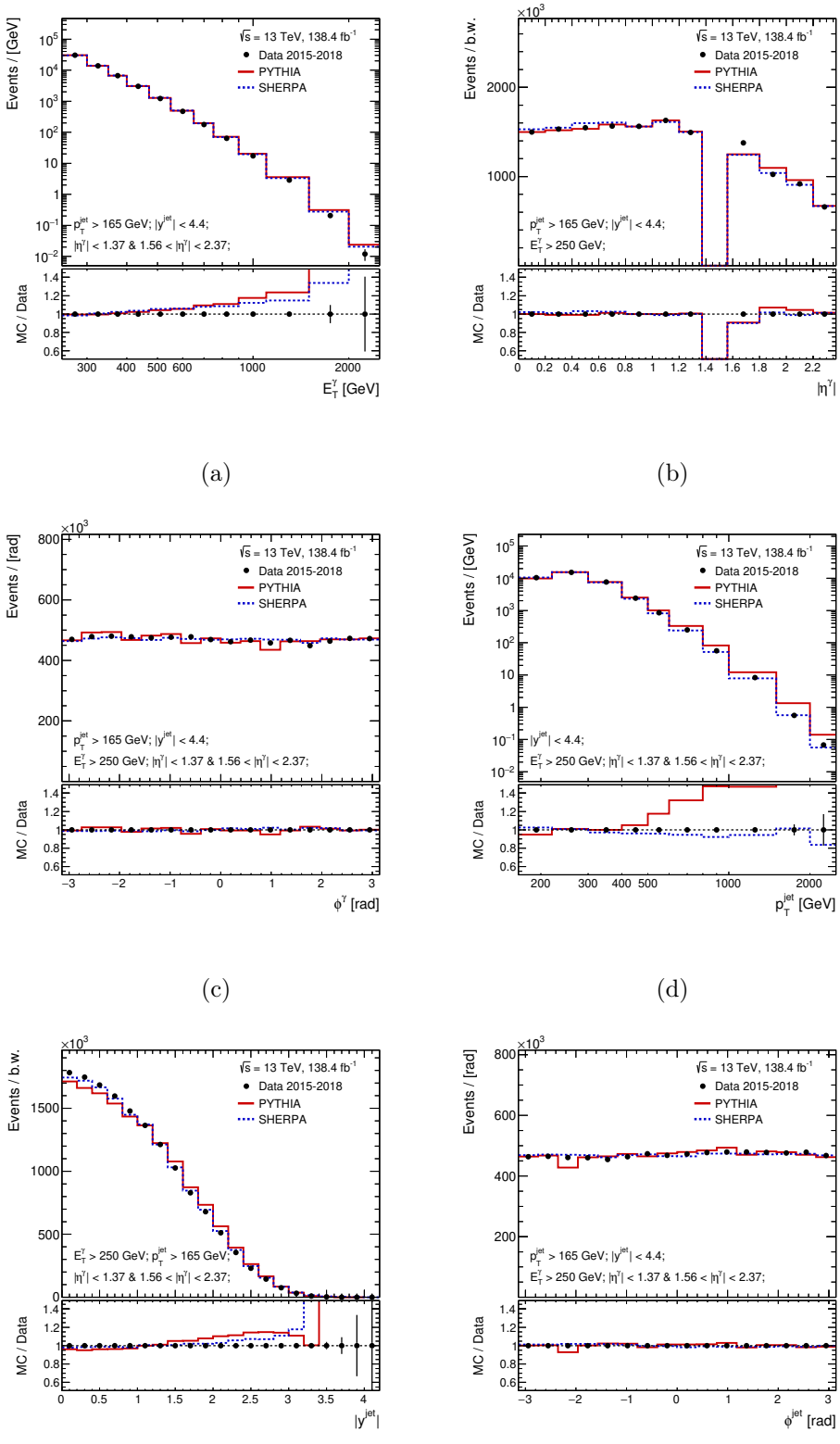
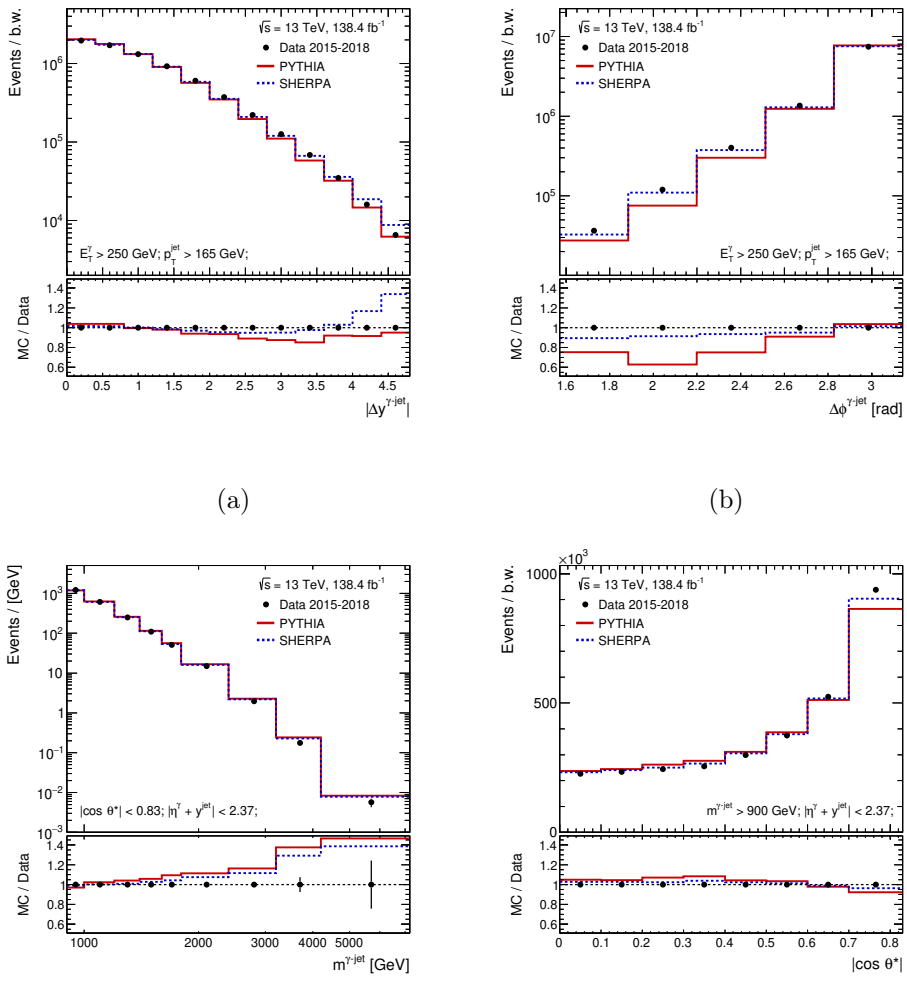


Figure 10.1: The measured E_{T}^{γ} (a), $|\eta^{\gamma}|$ (b), ϕ^{γ} (c), $p_{\mathrm{T}}^{\mathrm{jet}}$ (d), $|y^{\mathrm{jet}}|$ (e), and ϕ^{jet} (f) distributions for events in data satisfying the selection requirements summarised in table 10.1. For comparison, the MC simulations of the signal from PYTHIA (solid histograms) and SHERPA (dashed histograms) are also included. The MC simulations are normalised to the number of data events. The lower panel of each figure shows the ratio of the MC distributions to the data.



(a) (b)

(c) (d)

Figure 10.2: The measured $|\Delta y^{\gamma-\text{jet}}|$ (a), $\Delta\phi^{\gamma-\text{jet}}$ (b), $m^{\gamma-\text{jet}}$ (c) and $|\cos\theta^*|$ (d) distributions for events in data satisfying the selection requirements summarised in table 10.1. For comparison, the MC simulations of the signal from PYTHIA (solid histograms) and SHERPA (dashed histograms) are also included. The MC simulations are normalised to the number of data events. The lower panel of each figure shows the ratio of the MC distributions to the data.

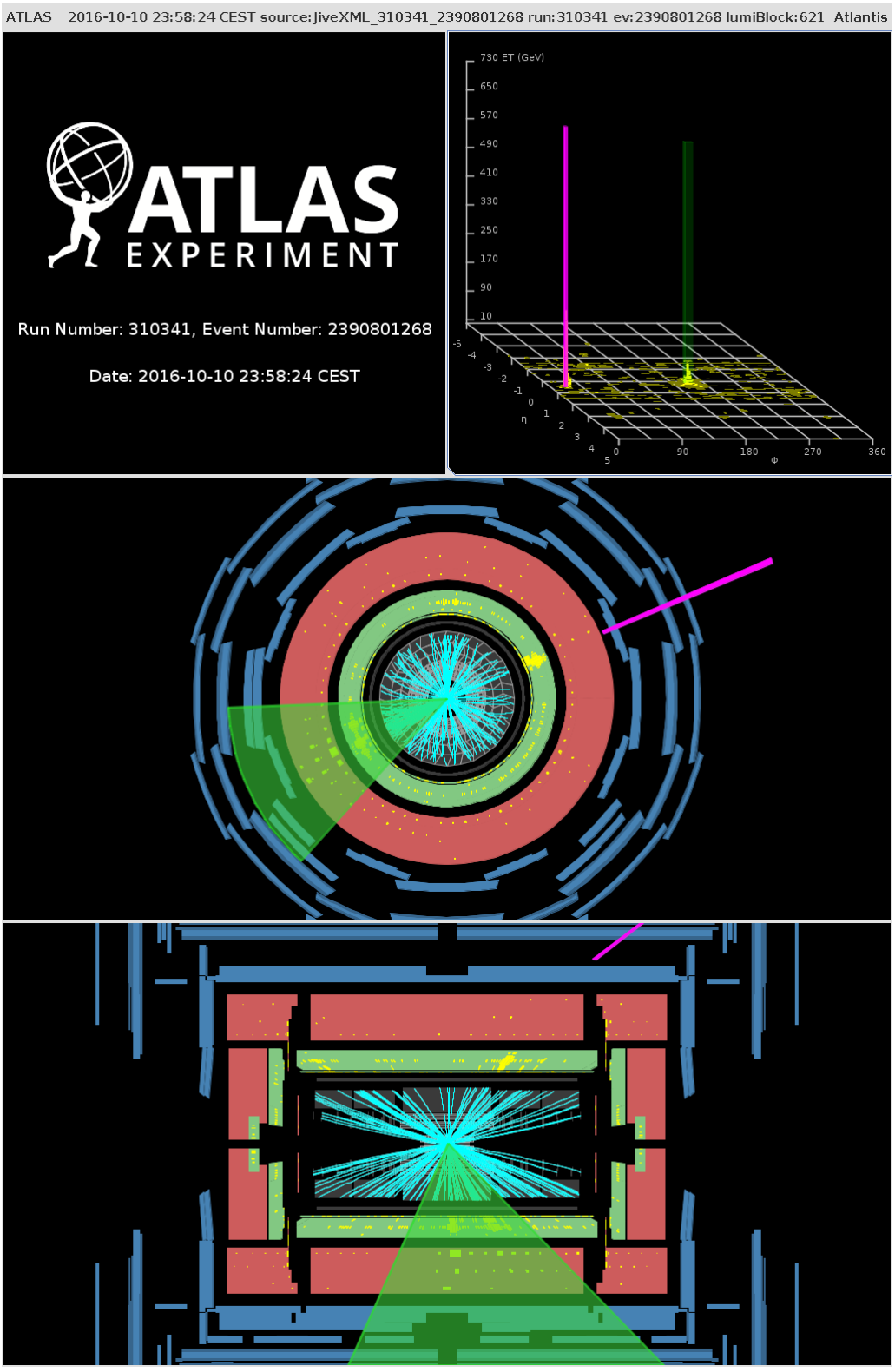


Figure 10.3: The selected event 2390801268 in run 310341 of the 2016 dataset is represented. The Inner Detector is represented in grey, the EM calorimeter in green, the hadronic calorimeter in red, and the muon spectrometer is drawn in blue. Tracks in the inner detector are represented with the cyan lines, and the yellow boxes represent the energy deposits in the calorimeters. The leading tight photon is represented by the pink line, and the leading jet by the transparent-green area. The lego plot shows the transverse energy of the different objects and calorimeter deposits. At the centre (bottom), the $Y - X$ ($\rho - Z$) projection of the detector is shown.

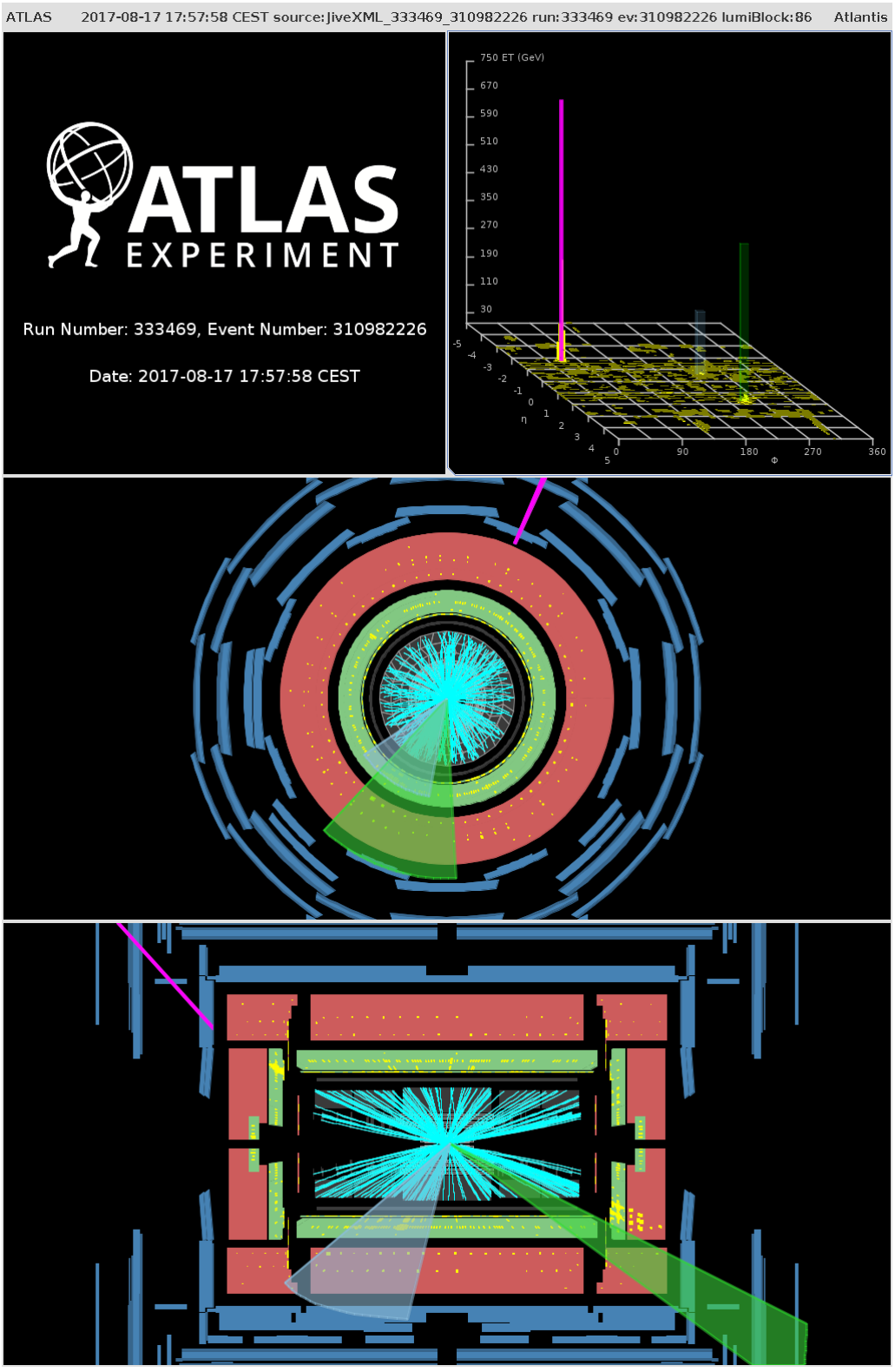


Figure 10.4: The selected event 310982226 in run 333469 of the 2017 dataset is represented. The Inner Detector is represented in grey, the EM calorimeter in green, the hadronic calorimeter in red, and the muon spectrometer is drawn in blue. Tracks in the inner detector are represented with the cyan lines, and the yellow boxes represent the energy deposits in the calorimeters. The leading tight photon is represented by the pink line, and the leading (subleading) jet by the transparent-green (blue) area. The lego plot shows the transverse energy of the different objects and calorimeter deposits. At the centre (bottom), the $Y - X$ ($\rho - Z$) projection of the detector is shown.

Trigger efficiency, uncertainty and correction

The efficiency of the HLT_g140_loose trigger used in this analysis was evaluated in the same way as described in section 10.1. The efficiency was estimated using a lower-threshold trigger, the HLT_g120_loose, which was unprescaled (prescaled) during the 2015 (2016−2018) running period. Figure 10.5 shows the estimated trigger efficiencies for tight isolated photons in each measured E_{T}^{γ} angular configurations. The trigger efficiencies for tight non-isolated, non-tight isolated, and non-tight non-isolated photons are shown in appendix I (these categories were defined in section 7.3).

The trigger efficiency was found to be very close to unity for all the measured observables and in the different tight isolated, tight non-isolated, non-tight isolated, and non-tight non-isolated regions. In addition, the trigger efficiency was estimated for each data-taking period separately, and no significant differences were observed. Therefore, no correction was applied to the data.

To estimate the uncertainty on the trigger efficiency, the same methodology as discussed in section 8.1 was followed. The results of the evaluation of the efficiency of the HLT_g140_loose with respect to a L1 trigger reported in ref. [158] were used to estimate an uncertainty on the measured cross section which was estimated as the sum in quadrature of the trigger inefficiency and its uncertainty. The relative uncertainty shown in table 8.2 for two ranges of E_{T}^{γ}, was propagated into the cross section for each running period separately.

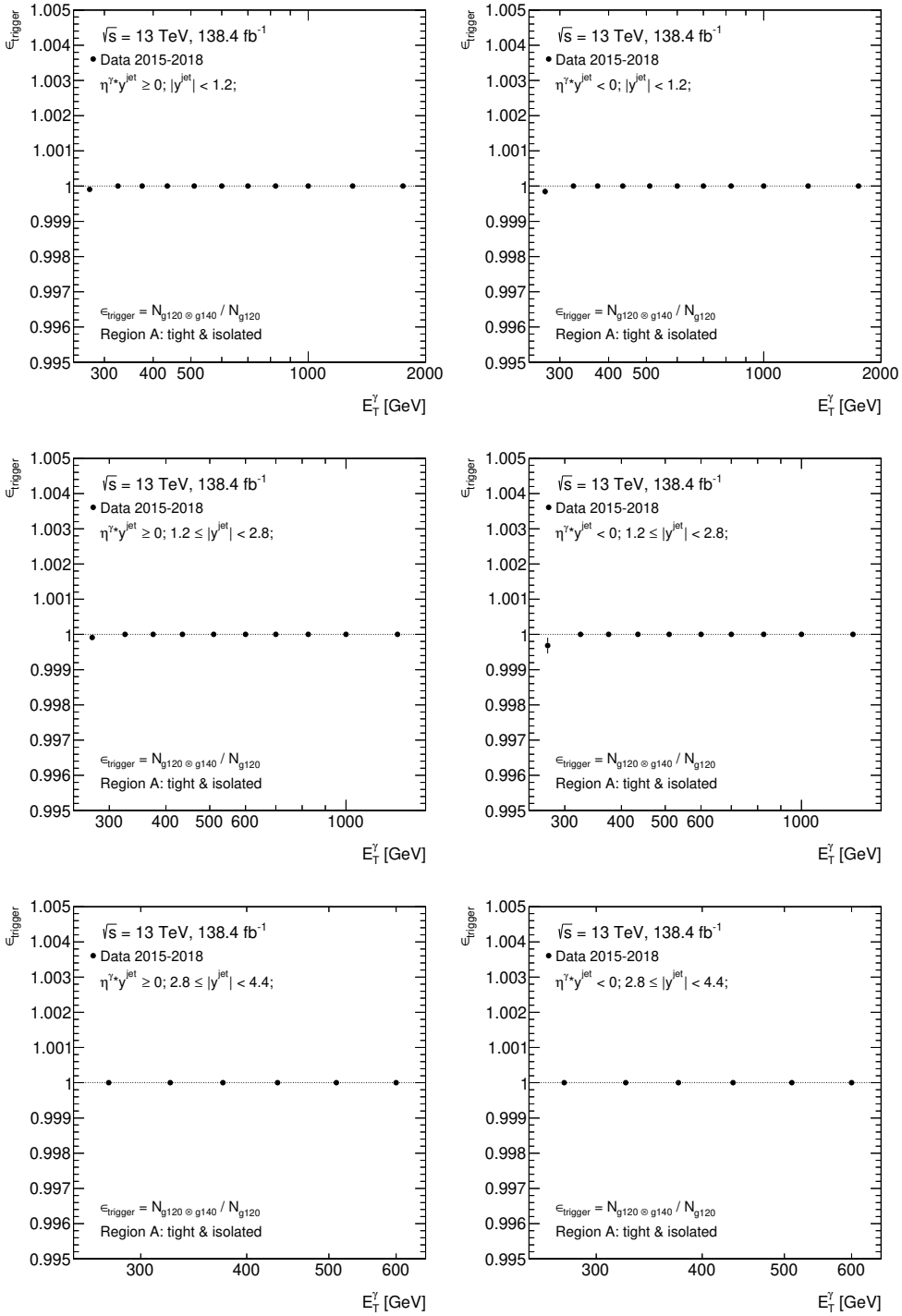


Figure 10.5: Trigger efficiency in data as a function of E_{T}^{γ} in the different $|y^{\mathrm{jet}}|$ regions and SS and OS configurations for tight isolated photons.

10.2 Monte Carlo simulations

Samples of MC events were generated to study the characteristics of the signal events, to perform the unfolding and to compute the predicted cross sections at particle level. As described in section 8.2, three campaigns of MC simulation were produced to match the different data-taking periods along the Run II. The different samples used for this analysis are described in detail in section 4.3.

As described in section 5.6, the MC events were reweighted such that the distribution of $\langle\mu\rangle$ matches the one in the data for each running period separately. The $\langle\mu\rangle$ reweighting was performed in a similar way as described in section 8.2.

In addition to the corrections described in section 5.6, a correction to the photon reconstruction efficiency was applied independently to each campaign of MC simulation (see section 8.2). The scale factors are shown in figure 8.6 for different $|\eta^{\gamma}|$ regions.

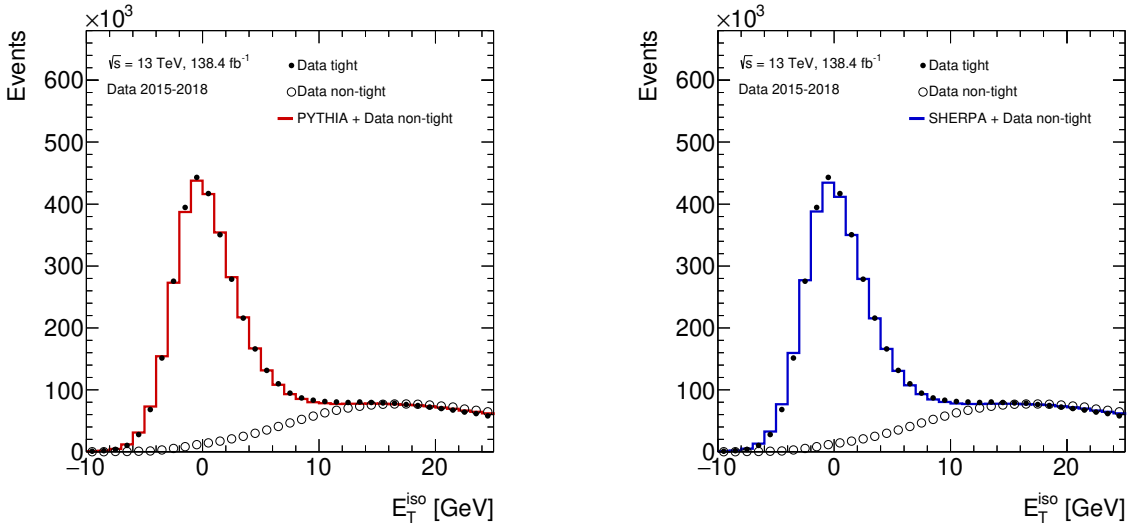
10.3 Background evaluation and signal extraction

As discussed in sections 7.3 and 8.3, even after the application of the tight identification and isolation requirements (see sections 5.4 and 5.5, respectively), there is a non-negligible amount of background remaining in the selected data sample. The background is estimated via the same data-driven method described in sections 7.3 and 8.3.

The isolation distribution of a signal-enriched sample was extracted from the data for further study. The sample used for this study was obtained after the application of the selection criteria described in section 10.1, except for the tight identification and the isolation requirements. The study was performed following the steps described in section 7.3. The $E_{\mathrm{T}}^{\mathrm{iso}}$ distributions for these subsamples are shown in figure 10.6. A good description of the data is obtained, showing similar results either using PYTHIA or SHERPA simulations to compare with the data. It is observed that the MC simulations exhibit a peak which reproduces that in the data after the application of the data-driven corrections on $E_{\mathrm{T}}^{\mathrm{iso}}$ to the MC simulations; this feature was previously discussed in section 7.3.

The same 2D-sideband method described in section 7.3 was used. The following paragraphs provide a description of the results obtained for the signal leakage fractions and purities for this analysis.

The signal leakage fractions were extracted from the MC simulations of the signal and are shown in figure 10.7 for PYTHIA and SHERPA. The fractions ϵ_{C}, which represent the signal leakage into the non-tight and isolated control region, are approximately constant in all observables and around $0.02 - 0.05$, being very similar for PYTHIA and SHERPA. The fractions ϵ_{B}, which represent the signal leakage into the tight and non-isolated background control region, are approximately constant and smaller for SHERPA compared to PYTHIA. For PYTHIA, the fraction ϵ_{B} decreases as a function of E_{T}^{γ}; the effect is more noticeable



(a) (b)

Figure 10.6: The measured $E_{\mathrm{T}}^{\mathrm{iso}}$ distribution before the isolation requirement and after applying the tight identification requirement (dots) and for those events which fail the tight identification (open circles). The non-tight distribution was normalised so that the integral of the tight and non-tight distributions for $E_{\mathrm{T}}^{\mathrm{iso}} > 10$ GeV coincides. For comparison, the MC simulations of the signal from PYTHIA (a) and SHERPA (b) (histograms) are also included. The non-tight contributions were added to the MC distributions.

in the OS configuration of the forward regions $1.2 \leq |y^{\mathrm{jet}}| < 2.8$ and $2.8 \leq |y^{\mathrm{jet}}| < 4.4$. The fractions ϵ_{D}, which represent the signal leakage into the non-tight and non-isolated control region, increase as a function of some of the observables, and take very different values for PYTHIA and SHERPA.

The signal yield was determined from the observed number of events in the data in the four regions of the γ_{ID} vs $E_{\mathrm{T}}^{\mathrm{iso}}$ plane and the signal leakage fractions obtained from the simulated events using equation 7.1. The signal purities, computed as $P = N_{\mathrm{A}}^{\mathrm{sig}}/N_{\mathrm{A}}$ using the signal leakage fractions from PYTHIA and SHERPA are shown in figure 10.8. For some of the points at high E_{T}^{γ} in which the purity is above unity, the purity was set to unity. The signal purity is $\geq 90\%$ and the estimations using either PYTHIA or SHERPA to compute the signal leakage fractions show very similar values. The nominal signal yields were extracted using the signal leakage fractions from SHERPA; while the estimation using signal leakage fractions from PYTHIA was taken as a systematic uncertainty on the purity (see section 10.7.1).

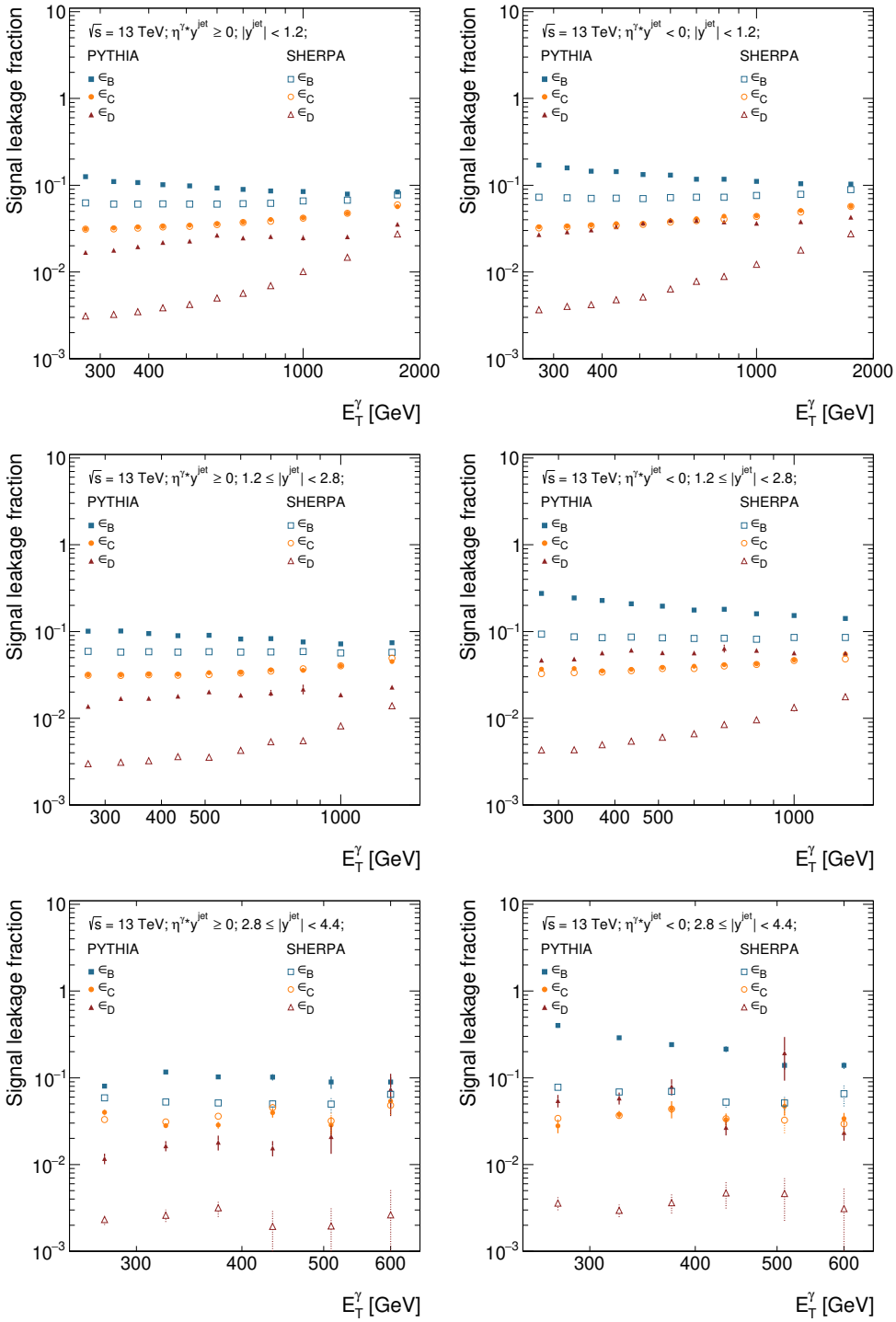


Figure 10.7: Signal leakage fractions from PYTHIA (solid markers) and SHERPA (open markers) for the control regions B (squares), C (circles) and D (triangles) as a function of the different observables measured.

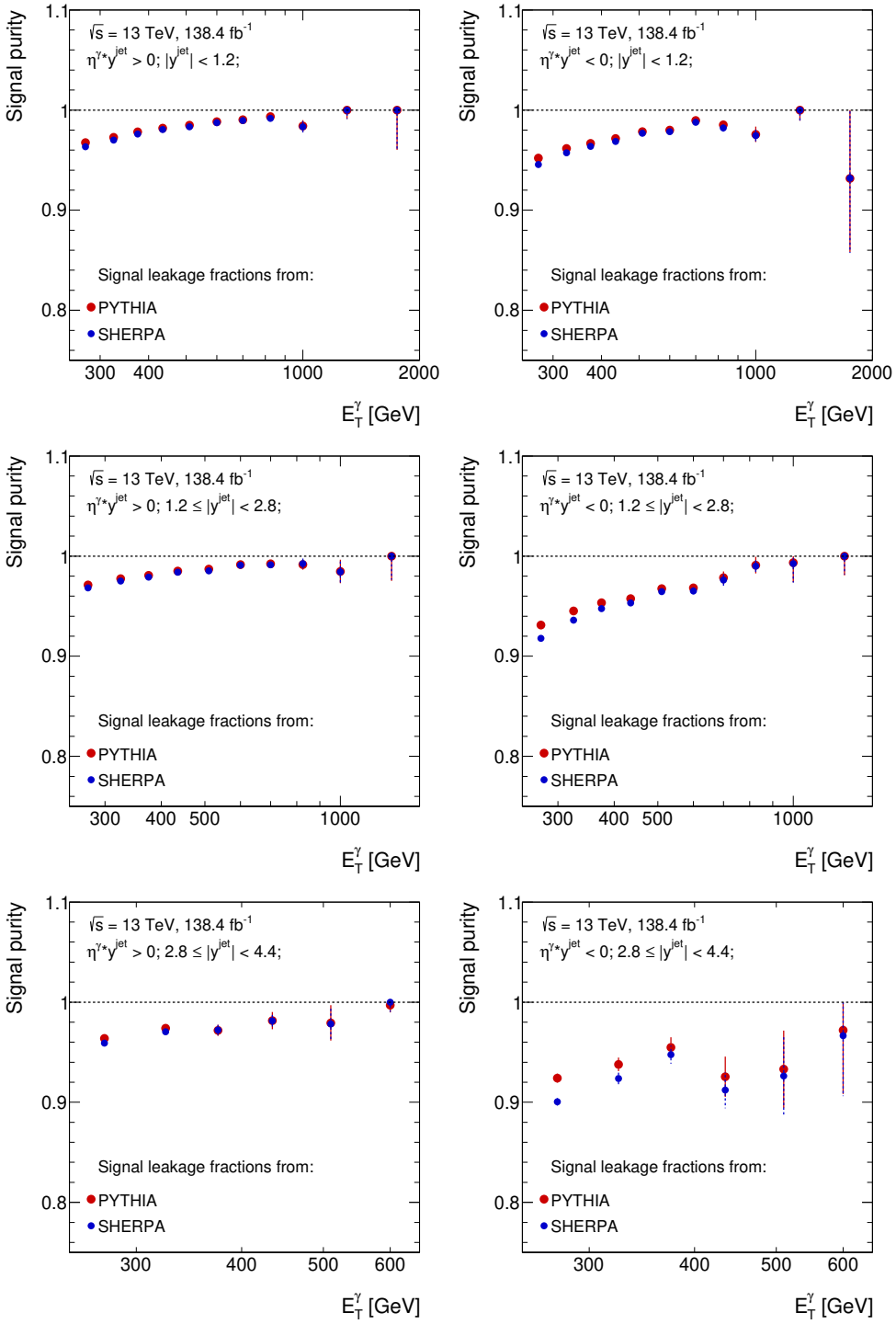


Figure 10.8: Estimated signal purities in data using signal leakage fractions from PYTHIA (red dots) and SHERPA (blue dots) as a function of the different observables measured.

10.3.1 Background from electrons faking photons

The background contribution of electrons or positrons misidentified as photons was estimated applying the same methodology described in sections 7.3.1 and 8.3.1, using SHERPA 2.2.1 MC simulations of W+jets and Z+jets to estimate the residual background.

The fraction of W+jets and Z+jets background to photon plus jet production as a function of the different observables measured was estimated with a particle-level matching to $e^{\pm}$ using an angular distance requirement $\Delta R < 0.2$. The total fraction of electron background is calculated adding linearly the contributions of W+jets and Z+jets, and is shown in figure 10.9.

The background from electrons faking photons was found to be smaller than 0.8% in the different measured observables. A systematic uncertainty of the size of the evaluated background was assigned (see section 10.7.2).

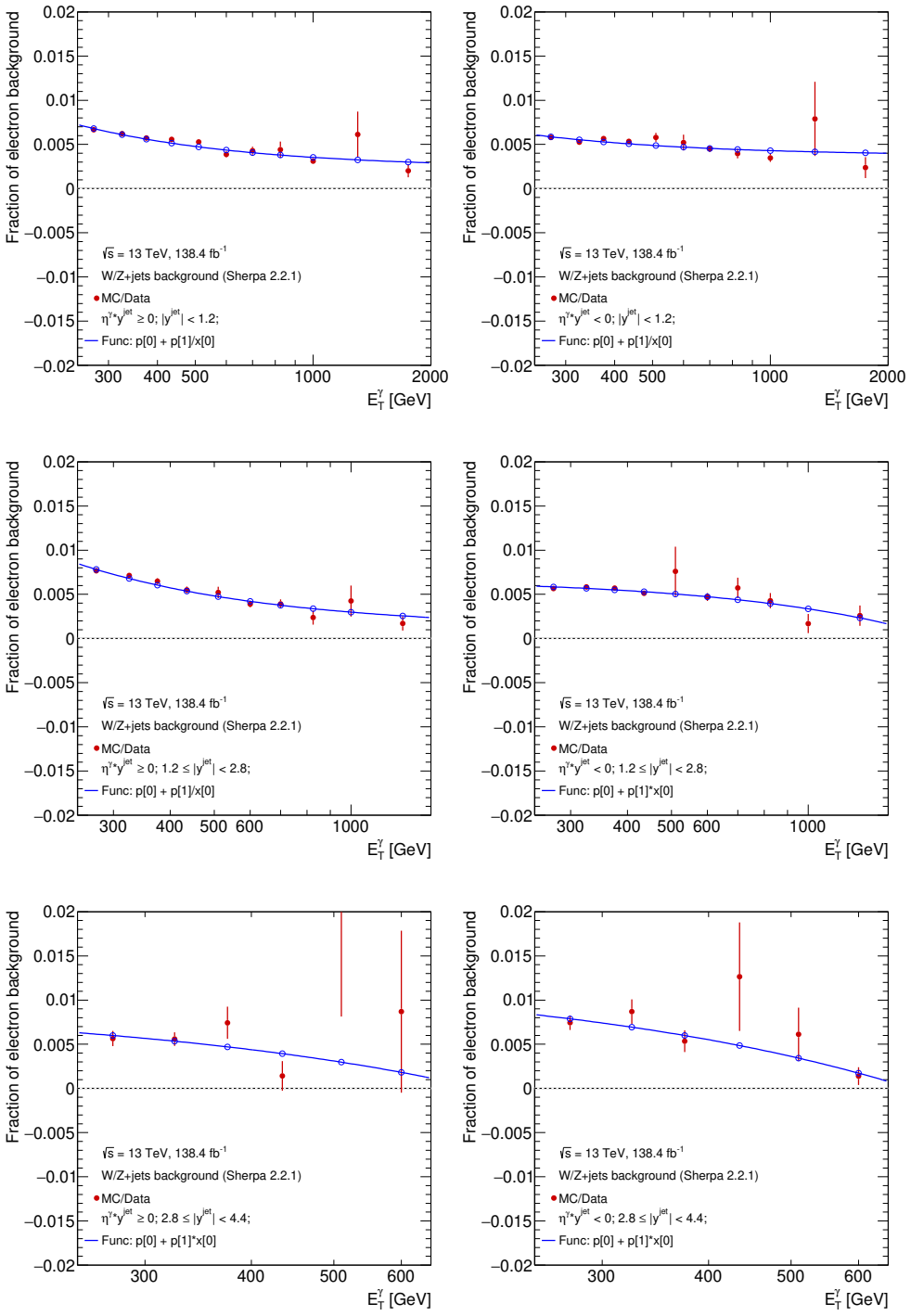


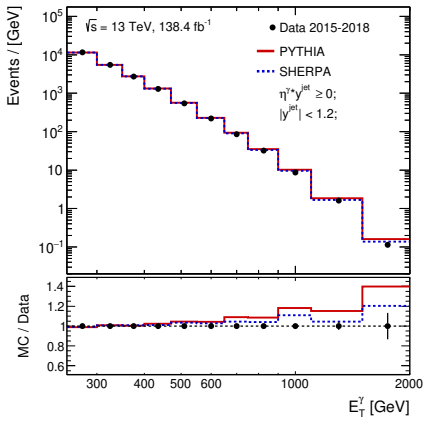
Figure 10.9: Fraction of electron background to inclusive photon production as a function of the different observables measured using the samples of SHERPA 2.2.1 (red dots). The error bars represent the statistical uncertainties of the ratios. The blue line represents the result of a fit to a smooth function.

10.4 Control plots

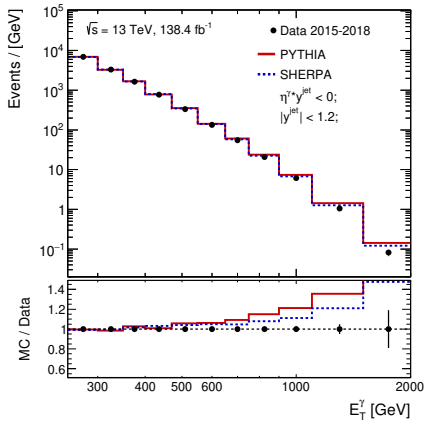
The estimated signal yields using the signal leakage fractions from SHERPA are shown in figure 10.10 for the different observables measured. For comparison, the simulations of PYTHIA and SHERPA normalised to the data for each observable are also included.

Figures 10.10(a) and 10.10(b) show the E_{T}^{γ} distributions in the $|y^{\mathrm{jet}}| < 1.2$ region for the SS and OS configurations, respectively; both PYTHIA and SHERPA provide a good description of the data, except at high E_{T}^{γ}. Figures 10.10(c) and 10.10(d) show the E_{T}^{γ} distributions in the $1.2 \leq |y^{\mathrm{jet}}| < 2.8$ region for the SS and OS configurations, respectively; the SS data distribution is well described by both PYTHIA and SHERPA in the full range, while the OS distribution is only described at low E_{T}^{γ}. Figures 10.10(e) and 10.10(f) show the E_{T}^{γ} distributions in the $2.8 \leq |y^{\mathrm{jet}}| < 4.4$ region for the SS and OS configurations, respectively; the SS data distribution is well described by both PYTHIA and SHERPA in the full range, while the OS distribution is only described by SHERPA at low E_{T}^{γ} and PYTHIA fails describing the data.

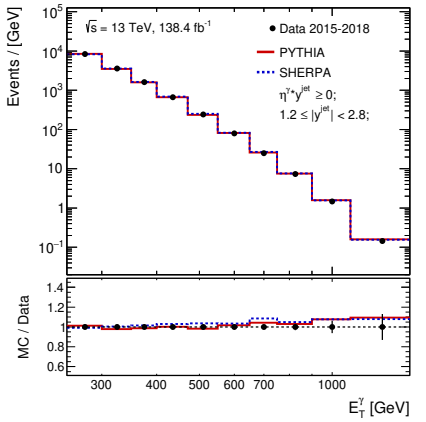
A comparison of the signal yield estimated with signal leakage fractions from PYTHIA with the hard and brem components of PYTHIA simulation is provided in appendix J. The fraction of brem component in PYTHIA as a function of each observable is also shown.



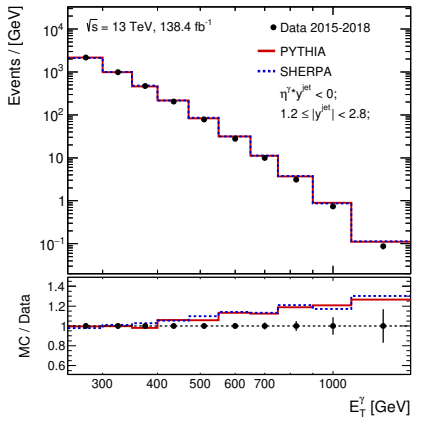
(a)



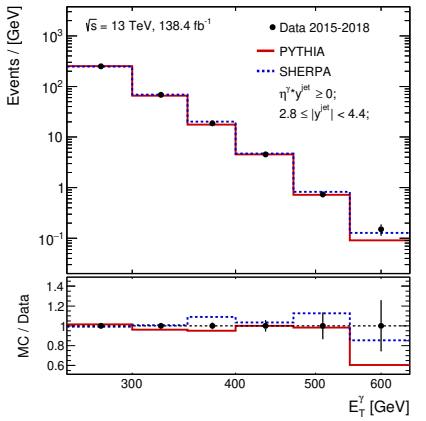
(b)



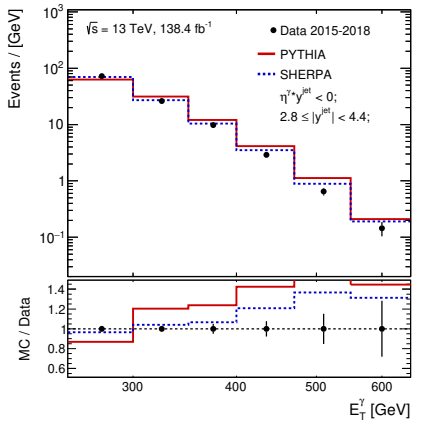
(c)



(d)



(e)



(f)

Figure 10.10: The extracted signal yield in data (dots) using signal leakage fractions from SHERPA as a function of E_{T}^{γ} in the different $|y^{\mathrm{jet}}|$ regions and SS and OS configurations. For comparison, detector level simulations from PYTHIA (solid histograms) and SHERPA (dashed histograms) are also included. The MC distributions are normalised to the number of data events. The ratio of the MC distributions to the data is shown in the lower panel of the figures. The error bars display the statistical uncertainty of the data.

10.5 Selection efficiency and purity

The selection efficiency and purity were evaluated using the MC samples of PYTHIA and SHERPA. The MC events were required to satisfy the same kinematic and isolation requirements, both at detector and particle levels; the particle- and detector-level photons (jets) were required to be matched using the angular requirement $\Delta R < 0.2$ (0.4).

The integrated selection efficiency, computed using equation 7.3, was found to be $(79.93 \pm 0.01)\%$ for SHERPA and $(80.21 \pm 0.03)\%$ for PYTHIA. The bin-to-bin selection efficiency was computed using equation 7.4. Figure 10.11 shows the bin-to-bin selection efficiency as a function of the different measured observables for PYTHIA and SHERPA. The bin-to-bin efficiency is between $62\% - 77\%$, being very similar for PYTHIA and SHERPA except for the OS configuration in the $2.8 \leq |y^{\text{jet}}| < 4.4$ region.

The integrated selection purity, computed using equation 7.5, was found to be $(92.30 \pm 0.01)\%$ for SHERPA and $(92.49 \pm 0.02)\%$ for PYTHIA. The bin-to-bin selection purity was computed using equation 7.6. Figure 10.12 shows the bin-to-bin selection purity as a function of the different measured observables for PYTHIA and SHERPA. The bin-to-bin purity is between $68\% - 93\%$, being very similar for PYTHIA and SHERPA except for the OS configuration in the $2.8 \leq |y^{\text{jet}}| < 4.4$ region.

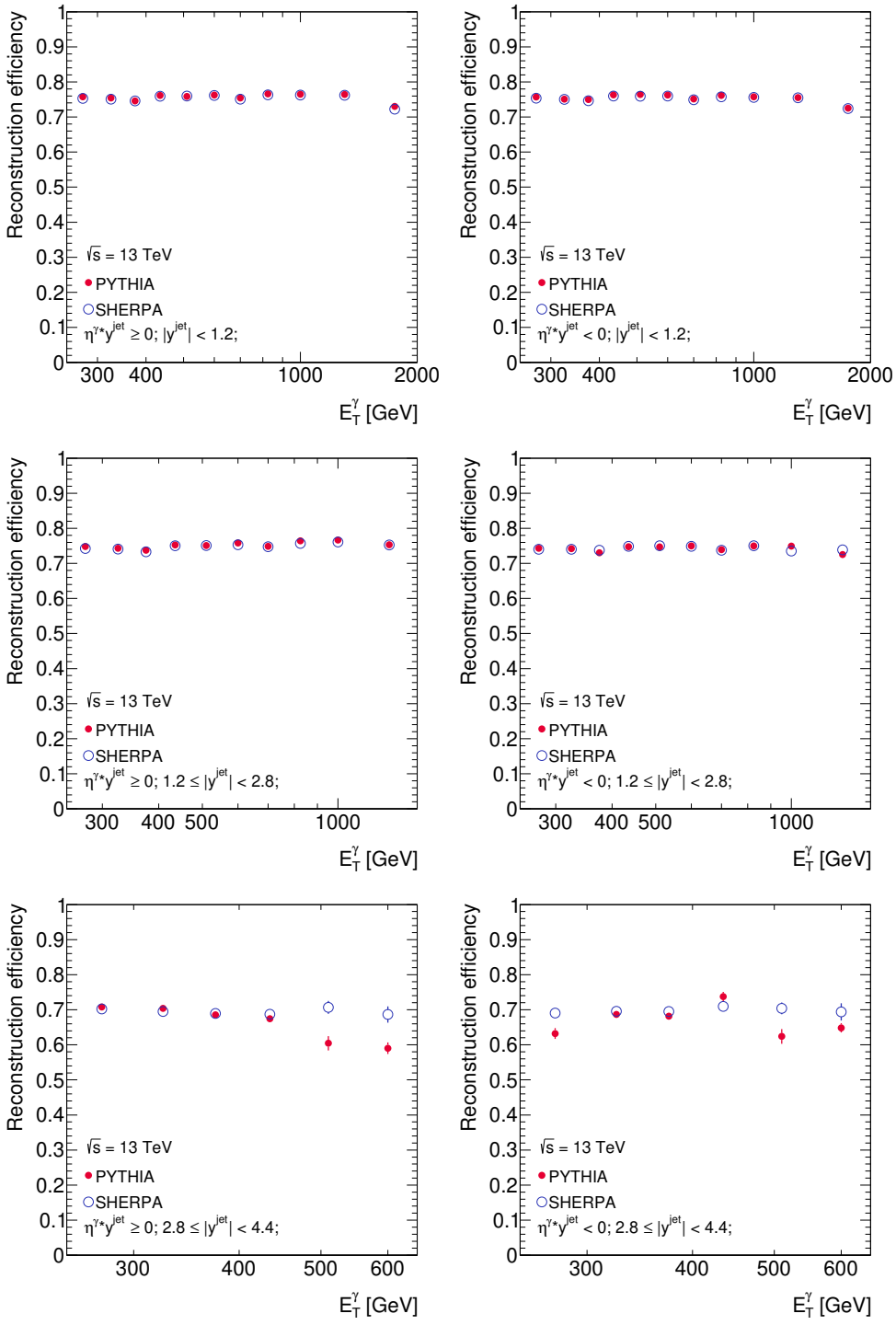


Figure 10.11: Selection efficiencies from PYTHIA (solid circles) and SHERPA (open circles) as a function of E_{T}^{γ} in the different measured configurations.

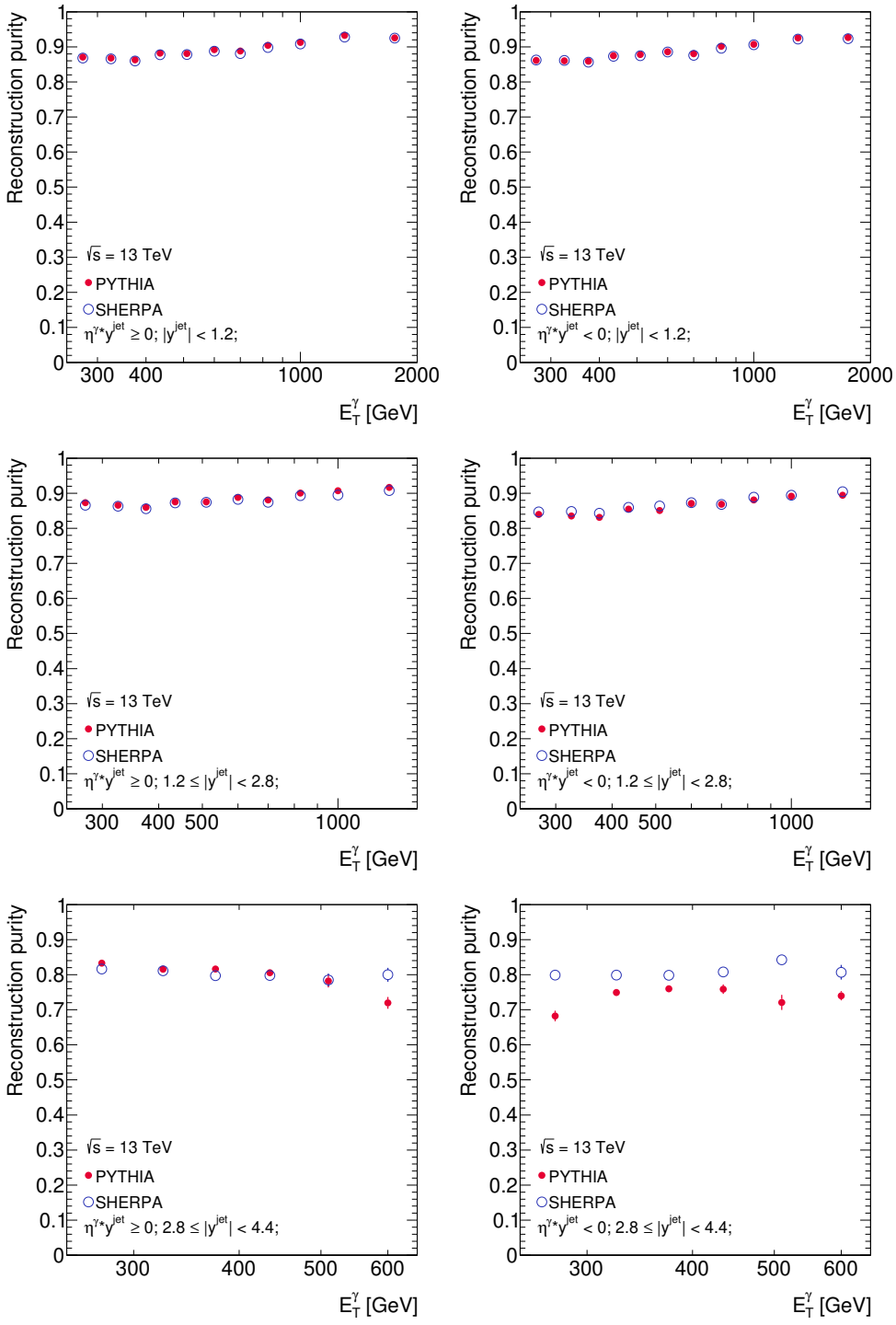


Figure 10.12: Selection purities from PYTHIA (solid circles) and SHERPA (open circles) as a function of E_{T}^{γ} in the different measured configurations.

10.6 Cross-section measurement procedure

The isolated-photon plus jet cross section was measured for photons with $E_{\mathrm{T}}^{\gamma} > 250$ GeV and $|\eta^{\gamma}| < 2.37$ (excluding the region $1.37 < |\eta^{\gamma}| < 1.56$) and jets with $p_{\mathrm{T}}^{\mathrm{jet}} > 165$ GeV and $|y^{\mathrm{jet}}| < 4.4$. Photons were required to be isolated by imposing $E_{\mathrm{T}}^{\mathrm{iso}} < 4.2 \times 10^{-3} \times E_{\mathrm{T}}^{\gamma} + 4.8$ GeV. The particle-level isolation requirement was defined with the same conditions described in section 8.6. The differential cross section was measured for the SS and OS configurations in three different rapidity ranges of the leading jet.

The data distributions, after background subtraction, were corrected to the particle level using the iterative application of the Bayes' theorem to obtain the measured differential cross section. The Bayesian unfolding method [153] as implemented in RooUnfold [159] was used. The method is described in detail in appendix E. Two iterations were used in this analysis.

The unfolding matrices used in the application of the Bayes' method are shown in figures 10.13 and 10.14 for SHERPA and PYTHIA, respectively. The *reco* and *truth* unmatched fractions previously introduced in section 8.6 are shown in figure 10.15.

The nominal cross section was measured using SHERPA to unfold the data distributions, and the differences in the cross section obtained by using PYTHIA to correct the data to the particle level were taken as a systematic uncertainty on the effect of the QCD-cascade and hadronisation models in the corrections (see section 10.7.4).

A cross-check on the unfolding method was performed by reweighting the MC distributions of SHERPA at detector and particle levels to the data. The cross-check is described in detail in appendix K.

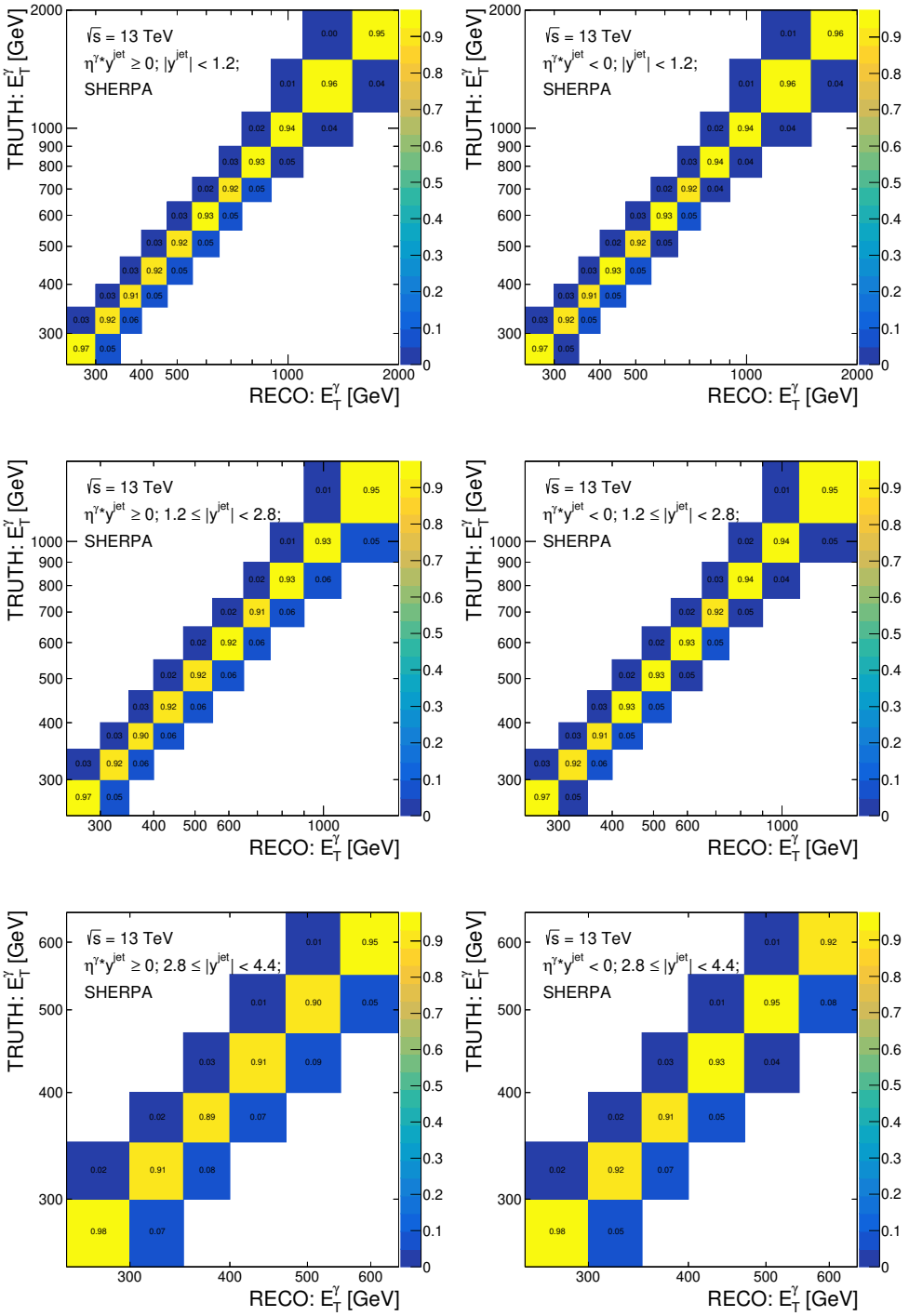


Figure 10.13: Unfolding matrices obtained with SHERPA MC in the different measured configurations. Normalisation is column wise.

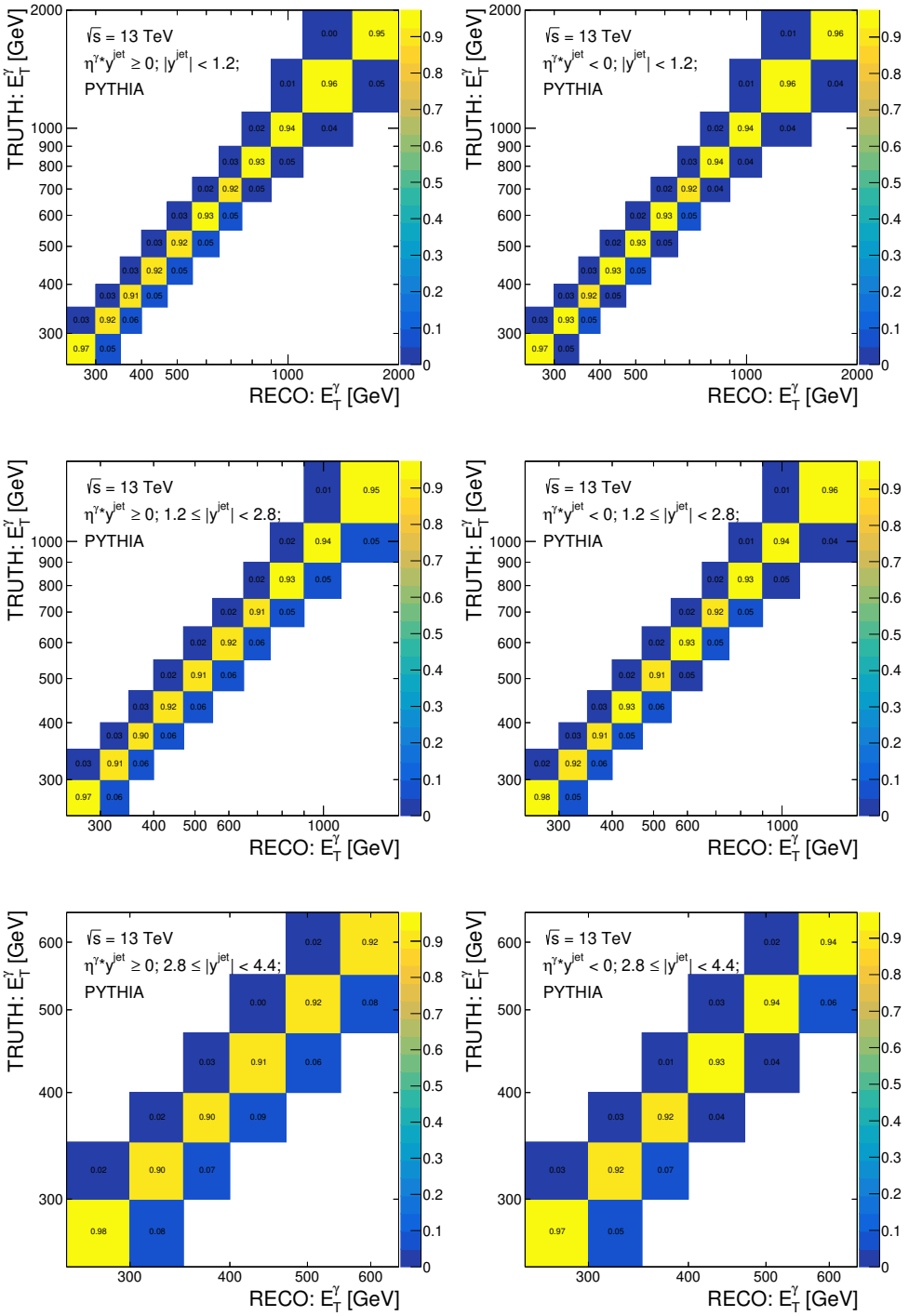


Figure 10.14: Unfolding matrices obtained with PYTHIA MC in the different measured configurations. Normalisation is column wise.

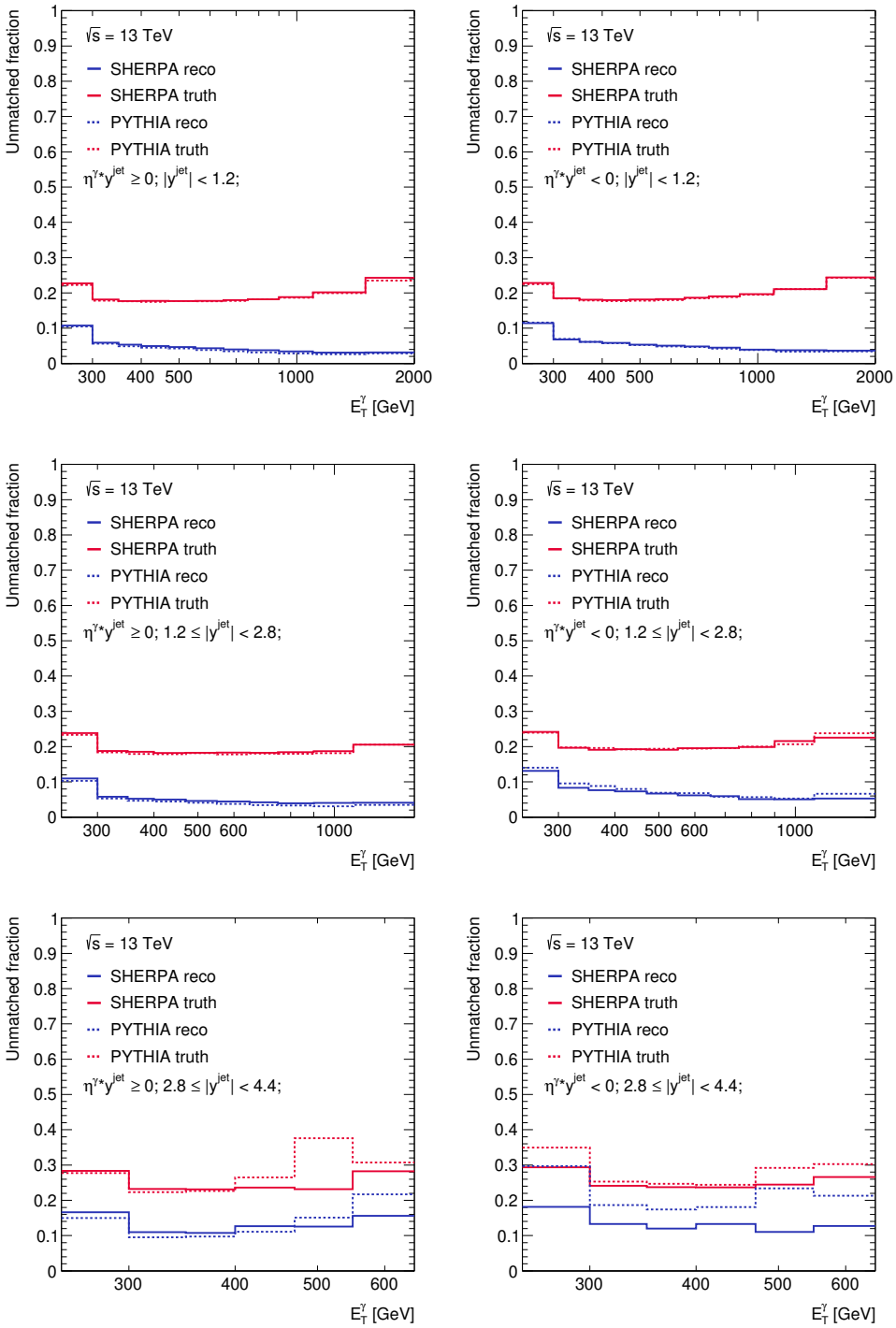


Figure 10.15: The reco (truth) unmatched fractions represented as blue (red) lines as a function of E_{T}^{γ} in the different measured configurations. SHERPA (PYTHIA) results are represented with solid (dashed) lines.

10.7 Systematic uncertainties

The sources of systematic uncertainties that affect this measurement are discussed in this section. These sources include the signal modelling, the background subtraction, the photon reconstruction, the unfolding procedure, the running conditions and the photon and jet energy calibrations. As previously described in sections 7.7 and 8.7, for some systematic uncertainties, where an adequate treatment of the data statistical influence is crucial, the bootstrap technique [154] was used.

10.7.1 Signal modelling

The nominal signal leakage fractions used for the extraction of the signal yield were computed using SHERPA. This uncertainty was estimated following the same approach described in section 8.7.1. The resulting uncertainty in the measured cross section is shown in figure 10.16 and is always less than 3% and typically less than 0.5%.

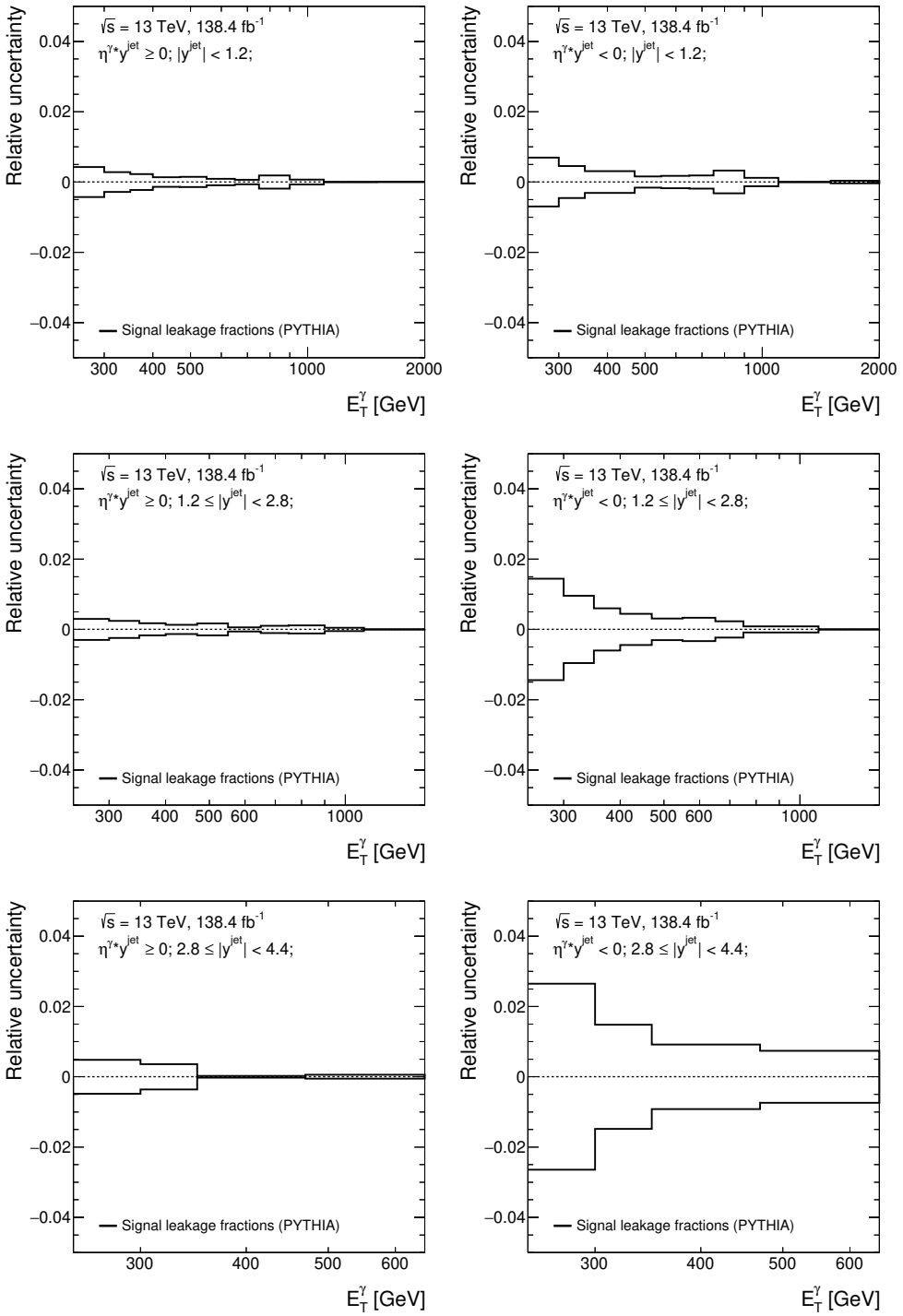


Figure 10.16: Systematic uncertainty in the measured cross section due to the QCD-cascade and hadronisation models as a function of E_{T}^{γ} in the different measured configurations. The uncertainty was symmetrised. Statistical fluctuations in the estimation of this systematic uncertainty were taken care of by rebinning, merging those bins affected by relatively large statistical fluctuations with neighbouring bins.

10.7.2 Background subtraction

Choice of background control regions

The data-driven background subtraction depends on the definition of the background control regions. This dependence is studied in a similar way as described in section 7.7.2, varying the lower limit on $E_{\mathrm{T}}^{\mathrm{iso}}$ for regions B and D by ± 1 GeV, removing the upper limit on $E_{\mathrm{T}}^{\mathrm{iso}}$ for regions B and D, as well as changing the choice of inverted photon identification variables. The uncertainties in the measured cross section from the first two variations are less than 0.4%, and are shown in figures 10.17 and 10.18, respectively. The resulting uncertainty in the cross section from the third variation is shown in figure 10.19, and is less than 1.5% and typically less than 0.5%.

Identification and isolation correlation in the background

The isolation and identification photon variables used to define the 2D-sideband method to subtract the background were assumed to be uncorrelated for background events ($R^{\mathrm{bg}} = 1$ in equation 7.2). A similar data-driven approach previously described in section 7.7.2 was used to determine the maximum deviations of R^{bg} from this original assumption. The maximum deviations from unity from the results obtained either using SHERPA or PYTHIA to estimate the signal leakage fractions are shown in figure 10.20, and vary in the range $20\% - 40\%$, depending on E_{T}^{γ} and the measured configuration. The uncertainty in the measured cross section was estimated in the same way as described in section 7.7.2, and is shown in figure 10.21. This uncertainty is $\lesssim 3\%$ and typically less than 1.5%.

Background from electrons faking photons

The background contribution from electrons or positrons faking photons was found to be subpercent (see section 10.3.1) in the different observables measured. This uncertainty was estimated following the same approach described in section 8.7.2. Figure 10.22 shows the resulting uncertainty in the measured cross section.

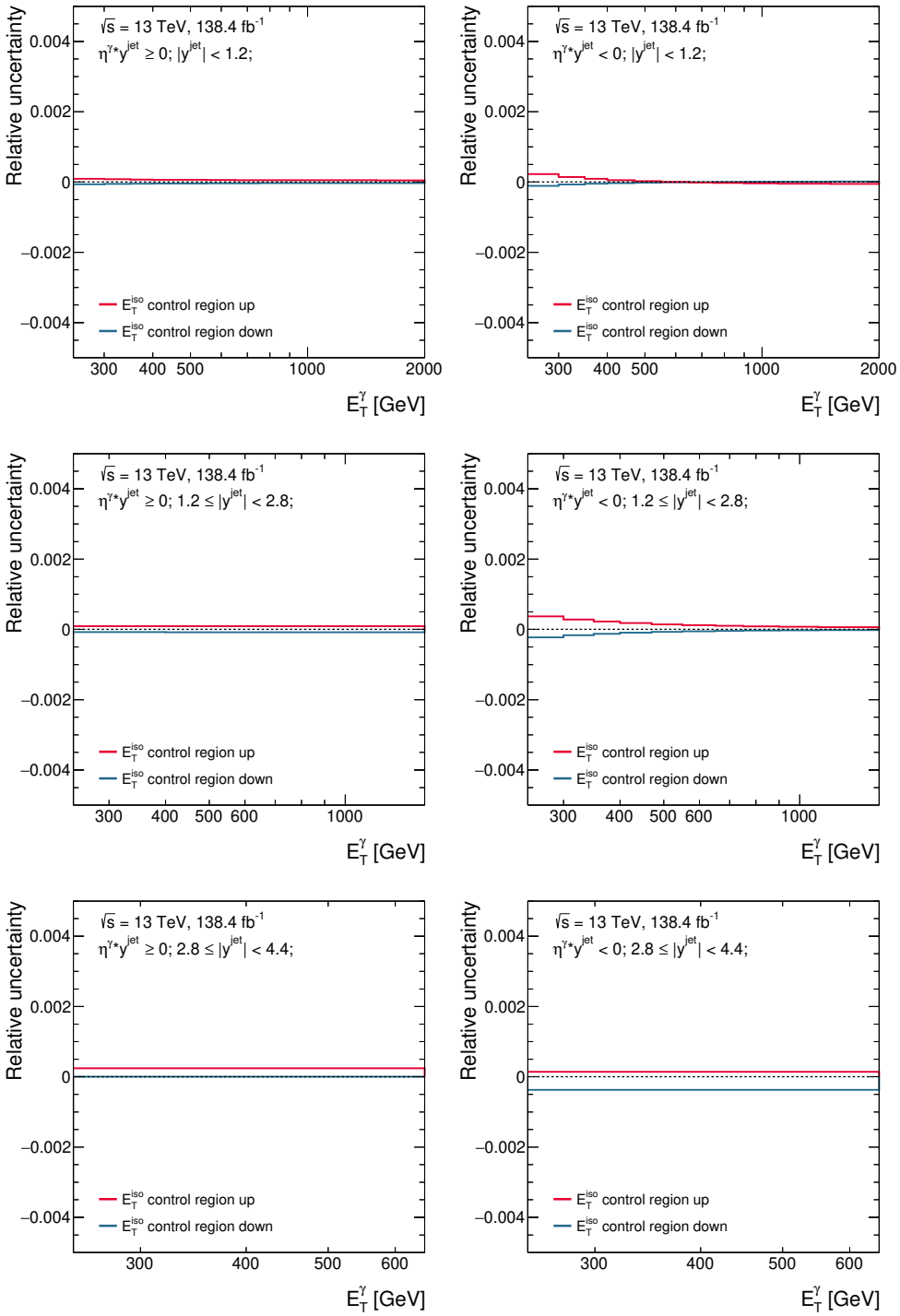


Figure 10.17: Systematic uncertainty in the measured cross section due to the choice of background control regions (different $E_{\mathrm{T}}^{\mathrm{iso}}$ requirements) as a function of E_{T}^{γ} in the different measured configurations. Statistical fluctuations in the estimation of this systematic uncertainty were taken care of by using a bootstrap technique.

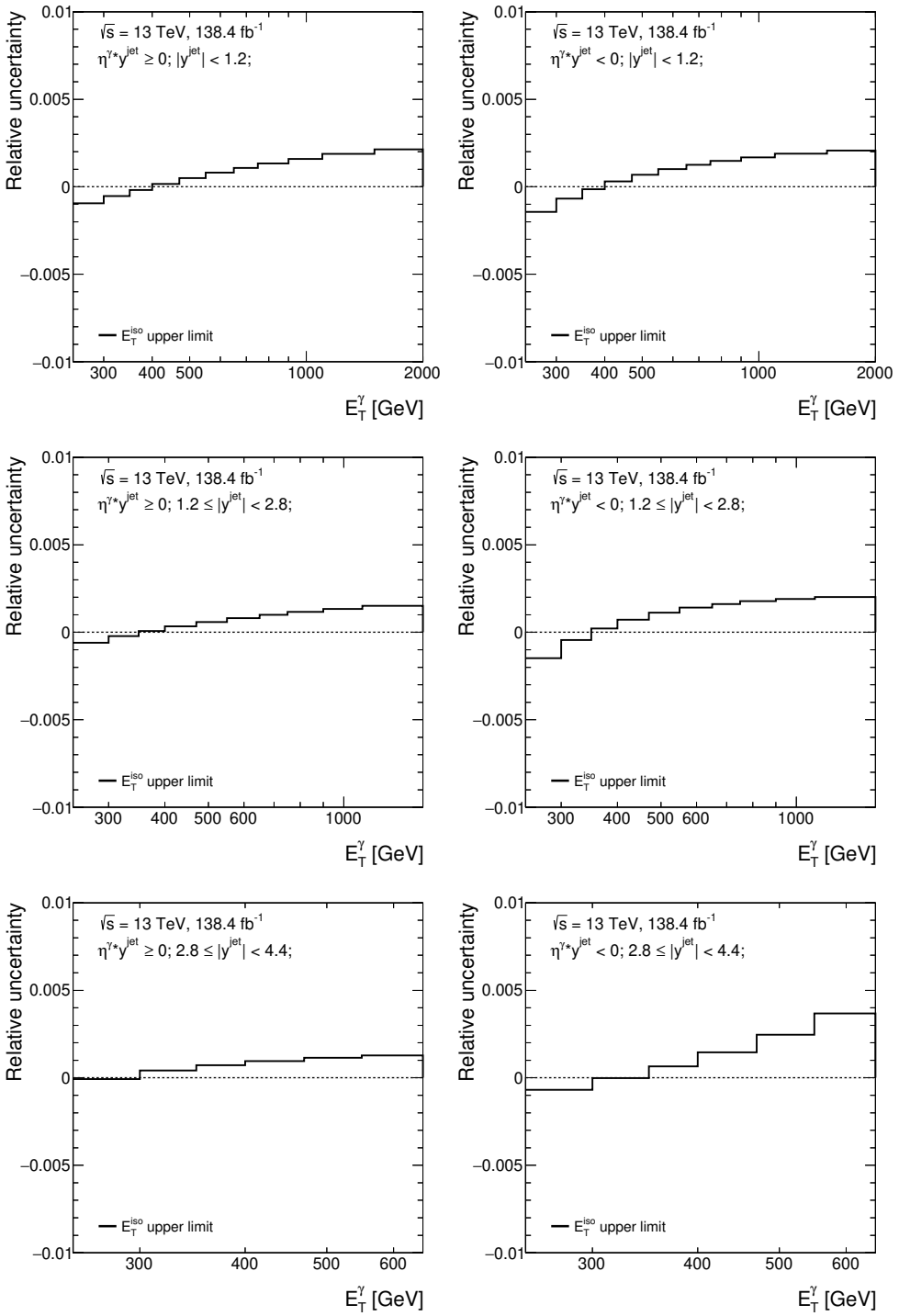


Figure 10.18: Systematic uncertainty in the measured cross section due to the choice of background control regions (removal of the upper limit on $E_{\mathrm{T}}^{\mathrm{iso}}$ for regions B and D) as a function of E_{T}^{γ} in the different measured configurations. Statistical fluctuations in the estimation of this systematic uncertainty were taken care of by using a bootstrap technique.

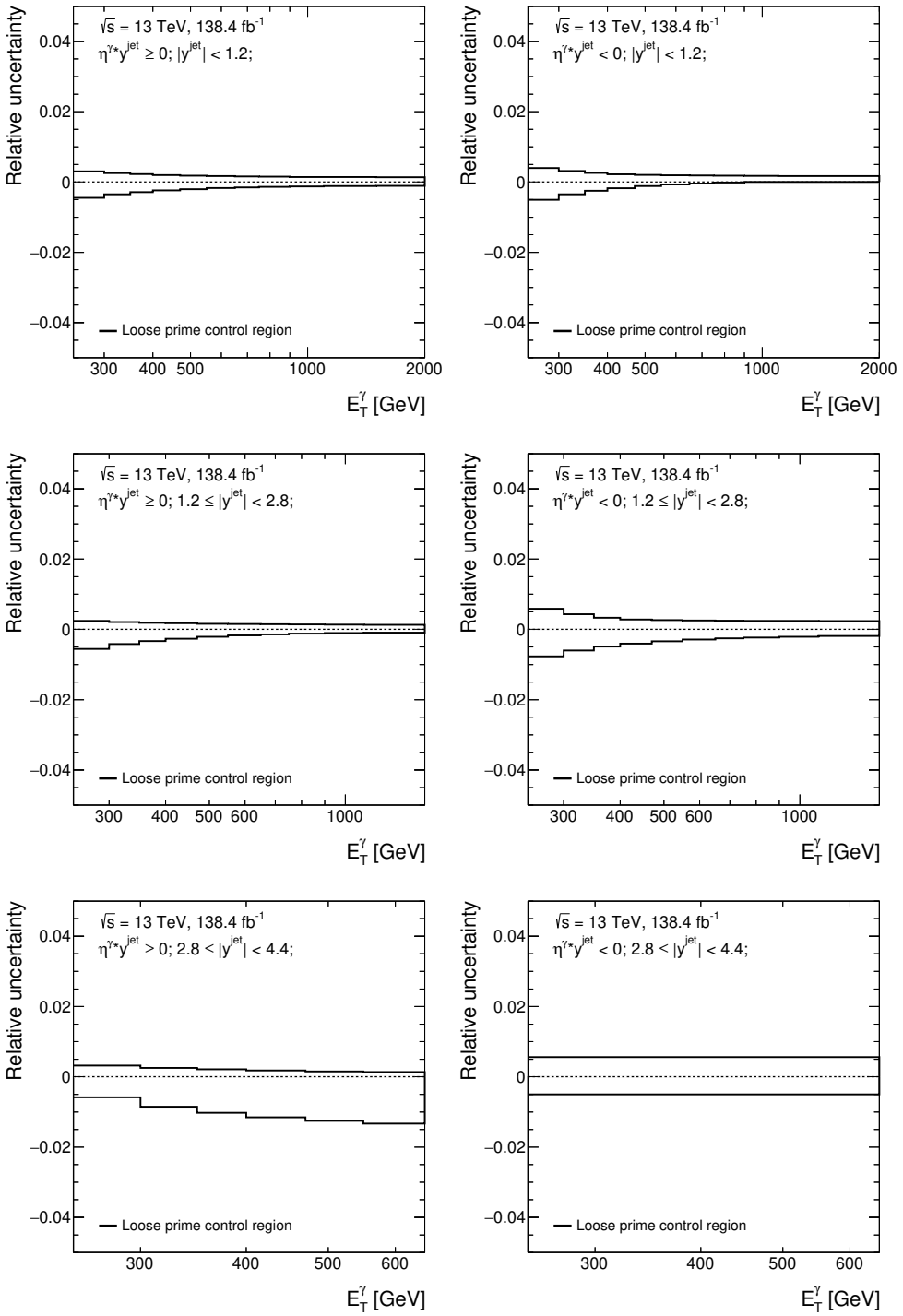


Figure 10.19: Systematic uncertainty in the measured cross section due to the choice of background control regions (different loose′ criteria) as a function of E_{T}^{γ} in the different measured configurations. Statistical fluctuations in the estimation of this systematic uncertainty were taken care of by using a bootstrap technique.

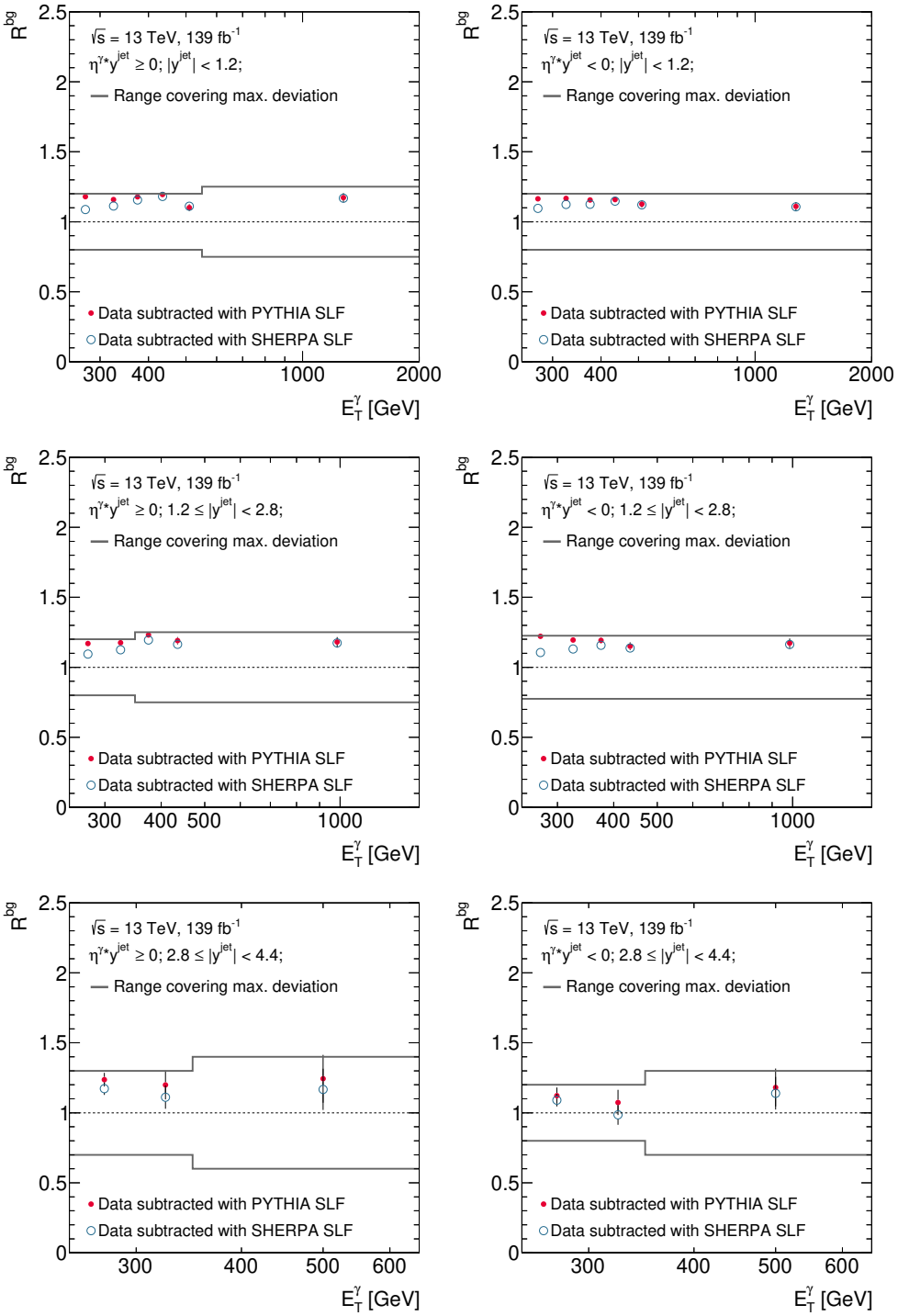


Figure 10.20: The estimated R^{bg} as a function of E_{T}^{γ} in the different measured configurations. The signal leakage contributions in the control regions were removed using either PYTHIA (red dots) or SHERPA (blue circles) samples. The range in R^{bg} which was taken as the uncertainty is indicated by the solid grey lines.

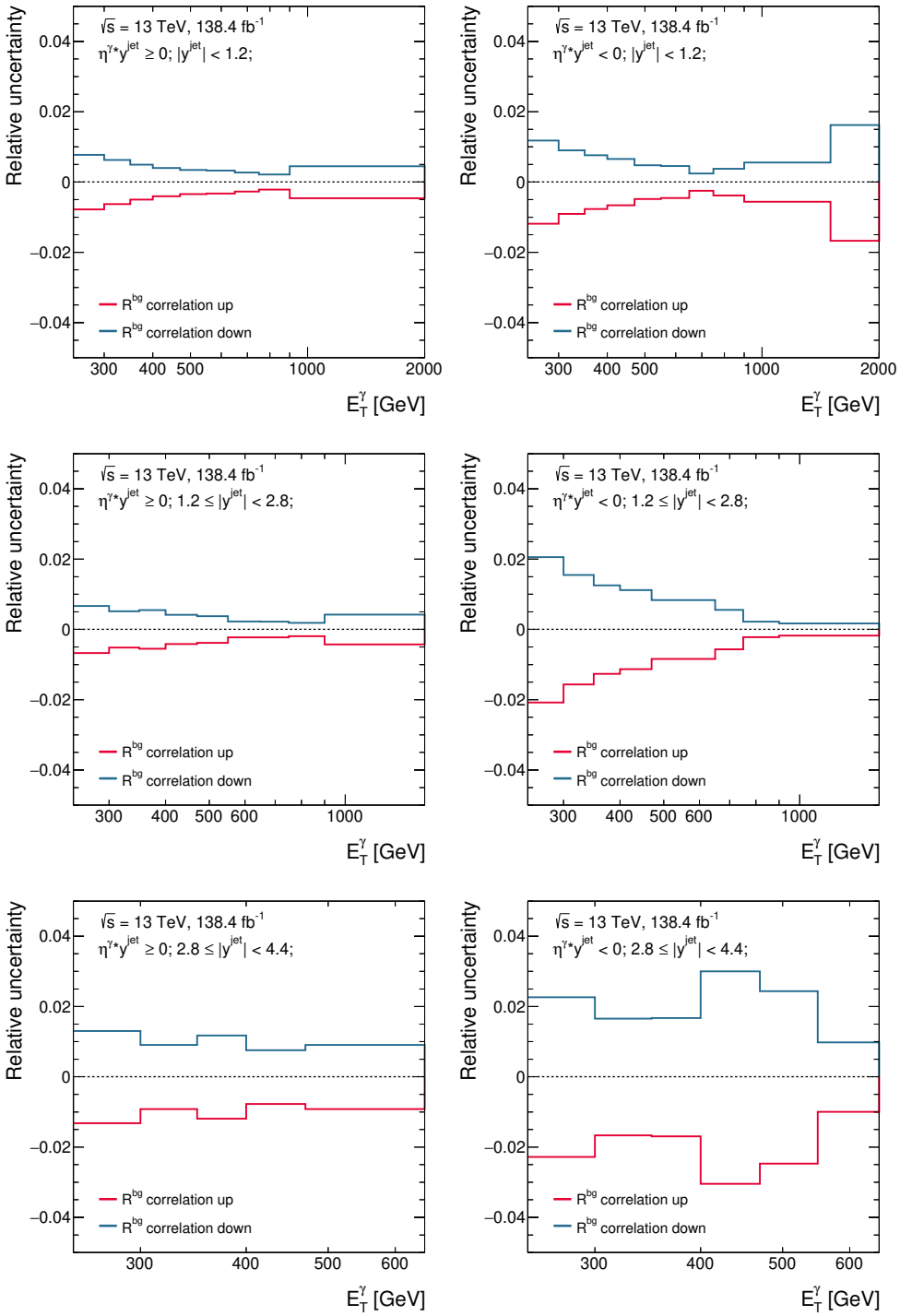


Figure 10.21: Systematic uncertainty in the measured cross section due to the correlation of the identification and isolation variables in the background as a function of E_{T}^{γ} in the different measured configurations.

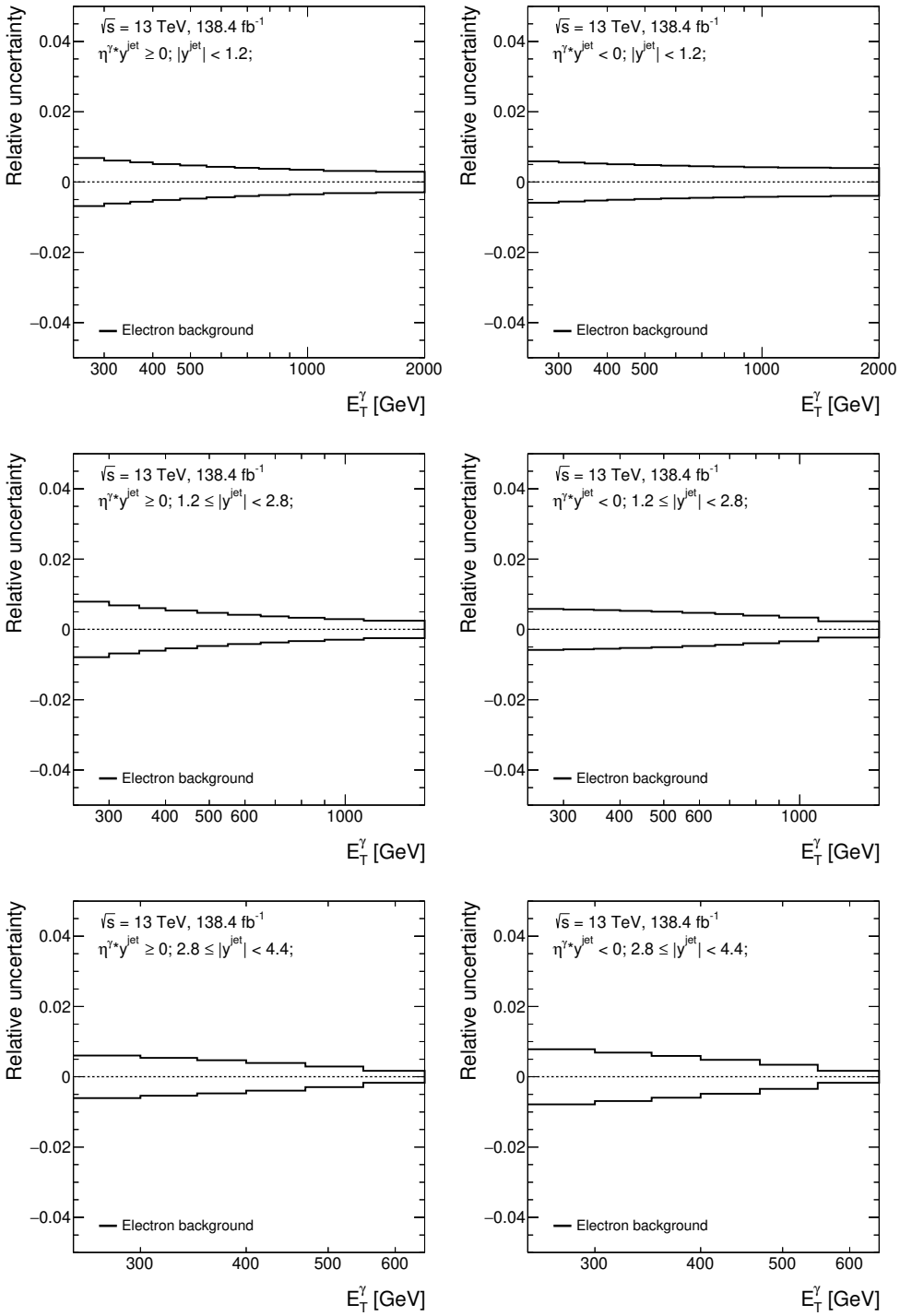


Figure 10.22: Systematic uncertainty in the measured cross section due to the electron background as a function of E_{T}^{γ} in the different measured configurations.

10.7.3 Photon reconstruction

Photon-reconstruction efficiency

Photon-reconstruction scale factors were applied to the MC events to account for the differences with respect to the data. This systematic uncertainty was estimated propagating the uncertainty in the scale factors to the measured cross section, and is shown in figure 10.23. This uncertainty is in the range $0.2\% - 0.4\%$.

Photon-identification efficiency

Photon-identification scale factors were applied to the MC events to match the tight photon identification efficiency measured in data [131]. The uncertainty in the cross section was estimated propagating the uncertainty in the scale factors to the measurement, and is shown in figure 10.24. This uncertainty is in the range $0.5\% - 0.75\%$.

$E_{\mathrm{T}}^{\mathrm{iso}}$ modelling

The uncertainty in the cross section due to the modelling of $E_{\mathrm{T}}^{\mathrm{iso}}$ in simulated events was evaluated following the same approach described in sections 7.7.3 and 8.7.3. The resulting uncertainty in the cross section is shown in figure 10.25, and is approximately smaller than 1%.

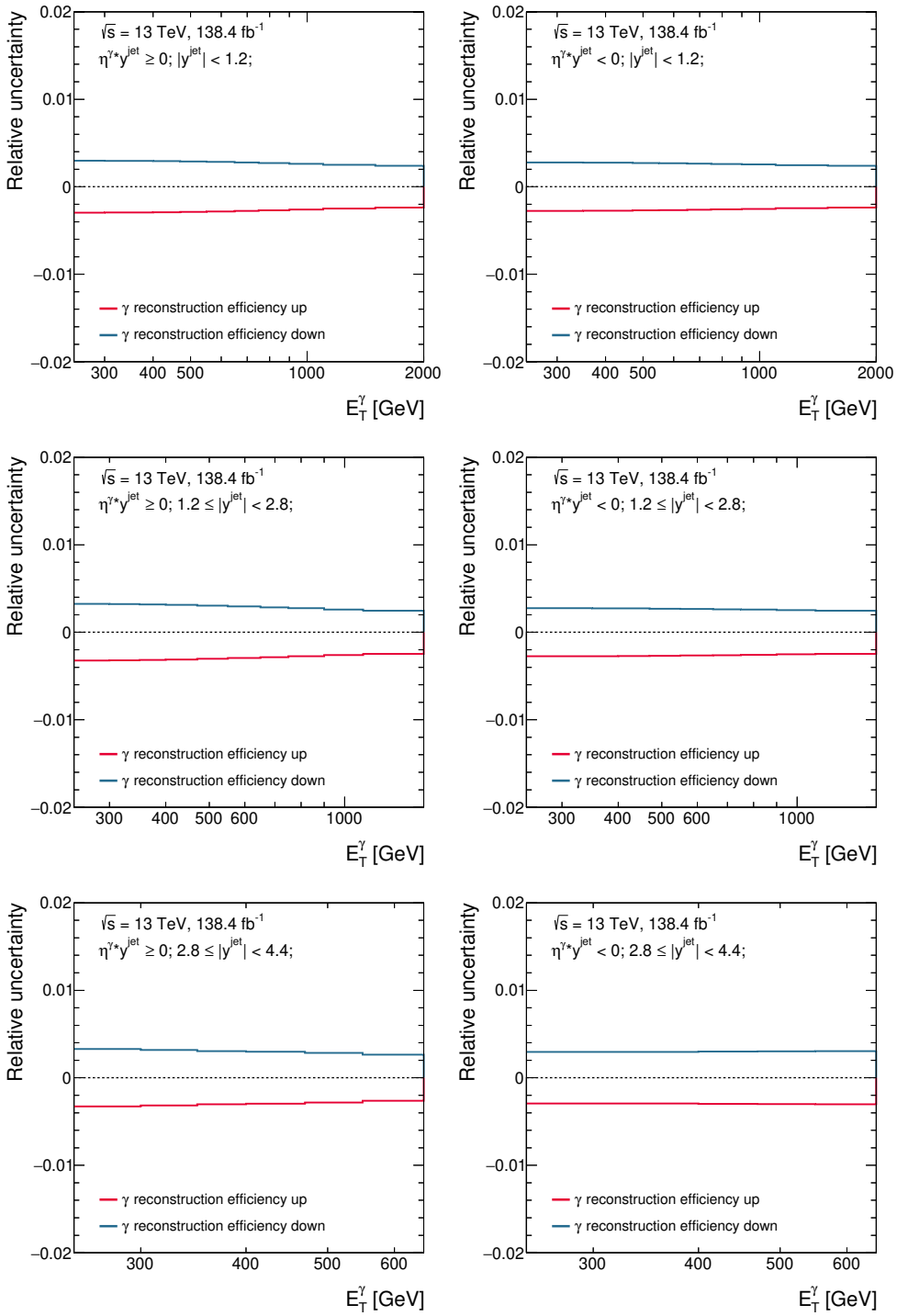


Figure 10.23: Systematic uncertainty in the measured cross section due to the uncertainty in the photon reconstruction efficiency as a function of E_{T}^{γ} in the different measured configurations.

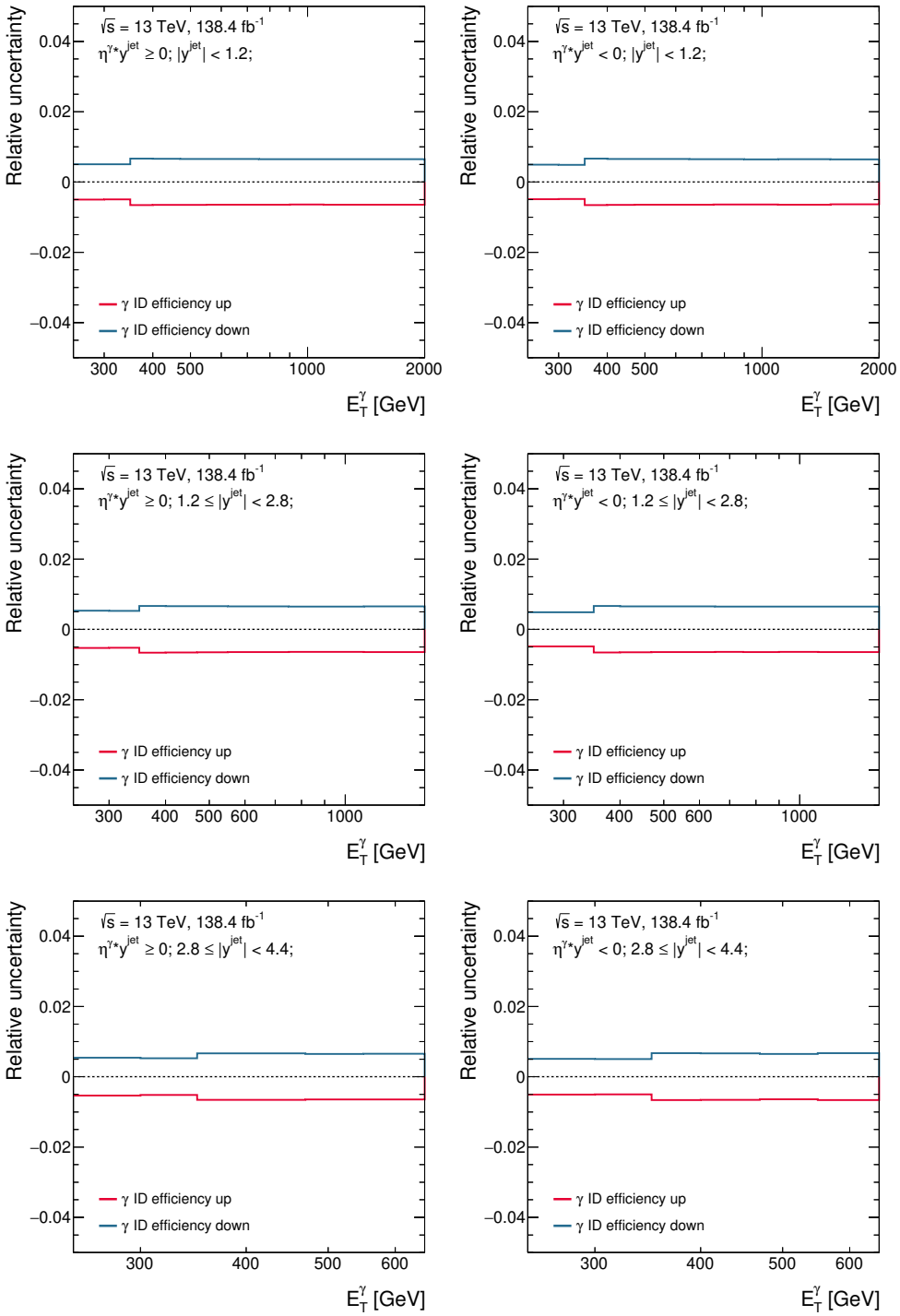


Figure 10.24: Systematic uncertainty in the measured cross section due to the uncertainty in the photon identification efficiency as a function of E_{T}^{γ} in the different measured configurations.

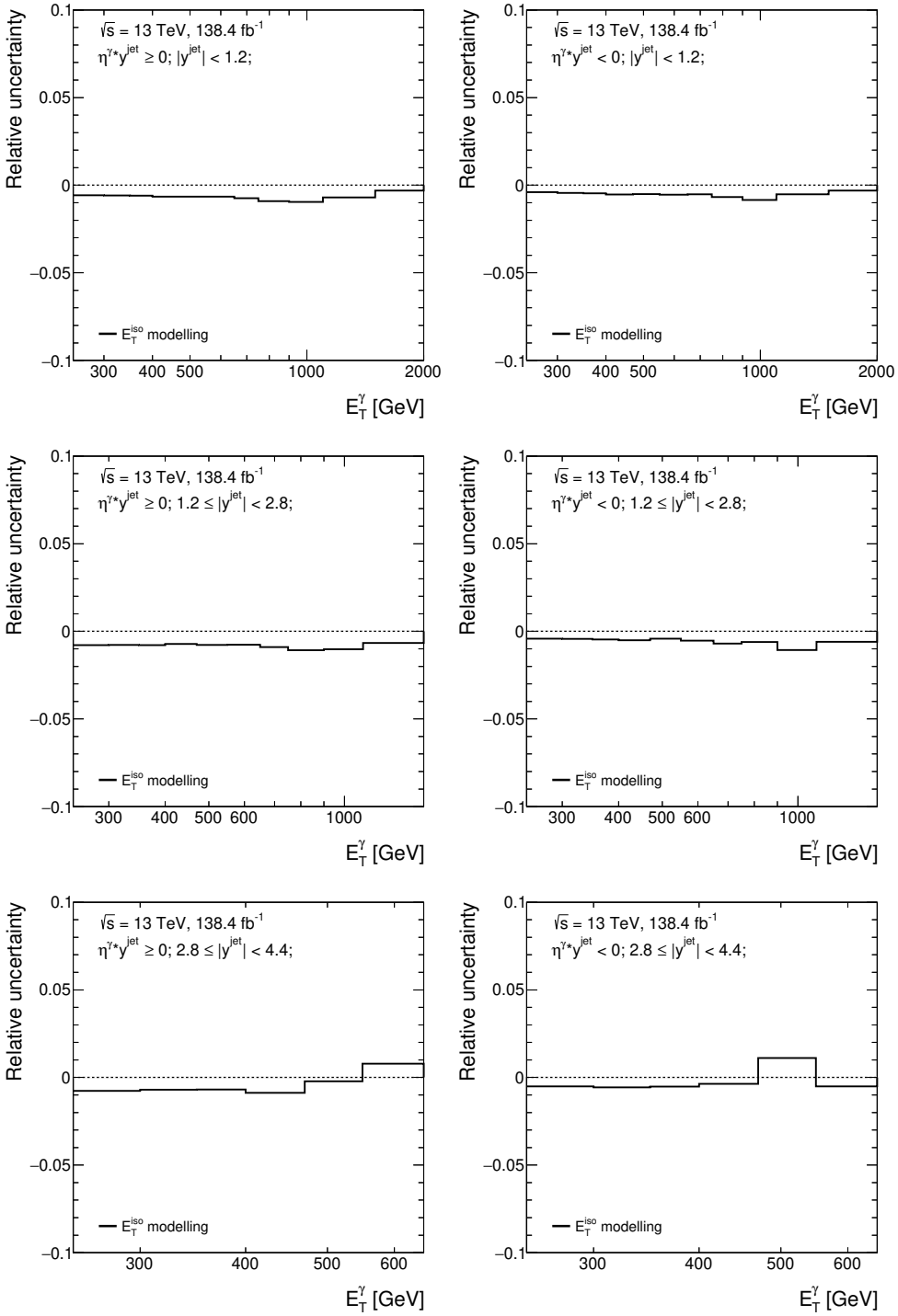


Figure 10.25: Systematic uncertainty in the measured cross section due to the effect of the isolation correction as a function of E_{T}^{γ} in the different measured configurations.

10.7.4 Unfolding procedure

QCD-cascade and hadronisation model dependence

The nominal unfolding correction matrices were computed using SHERPA. This uncertainty was estimated following the same approach described in section 8.7.4. The uncertainty in the measured cross section is $\lesssim 6\%$ and typically less than 2%. Figure 10.26 shows the resulting uncertainty in the measured cross section.

Unfolding closure

A systematic uncertainty due to the unfolding closure (see appendix K) was included. Figure K.2 shows the resulting uncertainty on the measured cross section. The uncertainty is smaller than 0.1%, except in the most forward $|y^{\text{jet}}|$ region, for both the SS and OS configurations, where it reaches up to 0.4%.

MC-sample statistics

The uncertainty in the cross section due to the statistical uncertainty of the MC samples is propagated into the measured cross section and typically amounts to less than 0.5%. In the most forward $|y^{\text{jet}}|$ region, the uncertainty reaches up to 4%. Figure 10.27 shows the resulting uncertainty in the measured cross section.

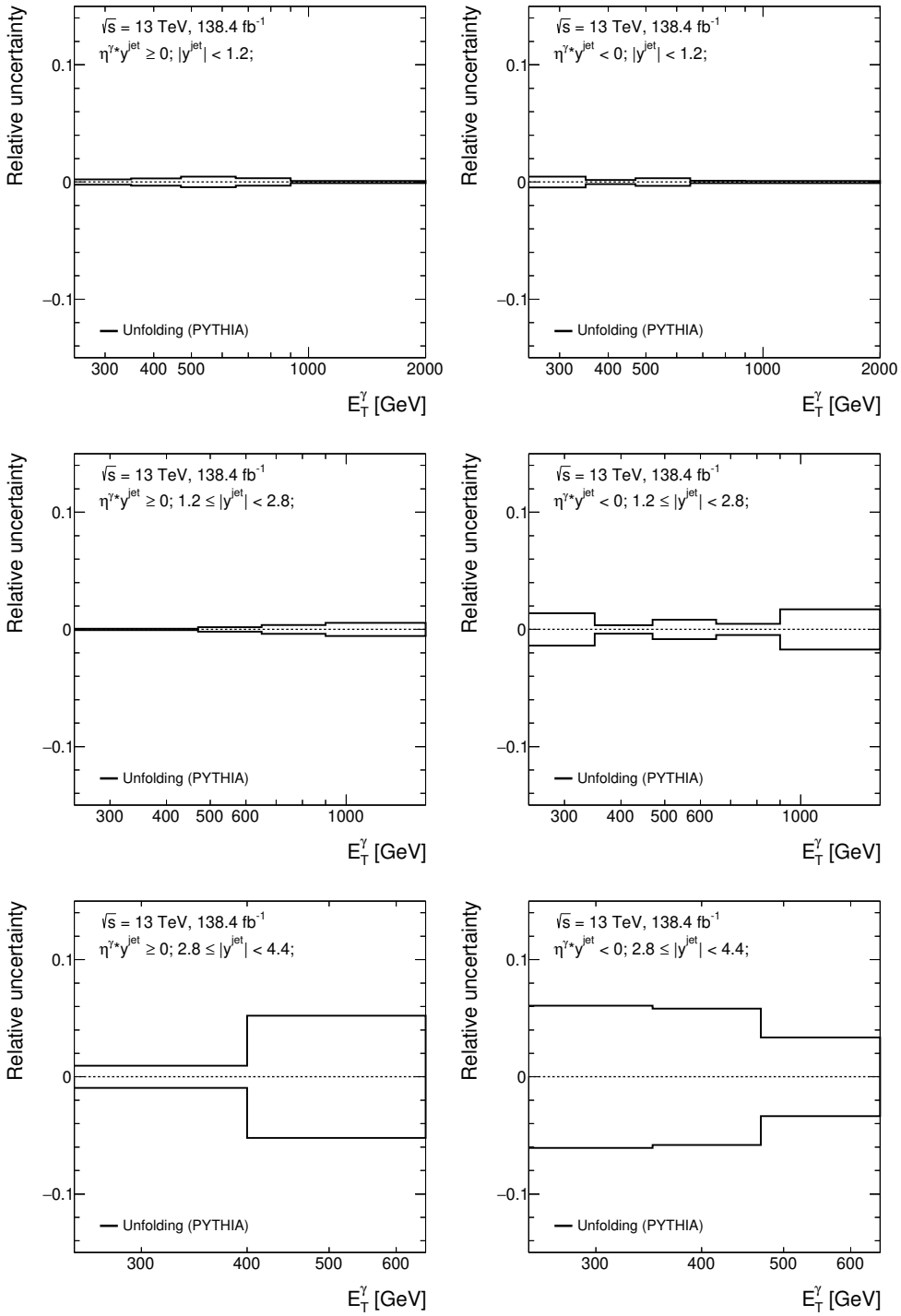


Figure 10.26: Systematic uncertainty in the measured cross section due to the QCD-cascade and hadronisation models as a function of E_{T}^{γ} in the different measured configurations. Statistical uncertainties in the estimation of this systematic uncertainty were taken care of by rebinning, merging those bins affected by relatively large statistical fluctuations with neighbouring bins.

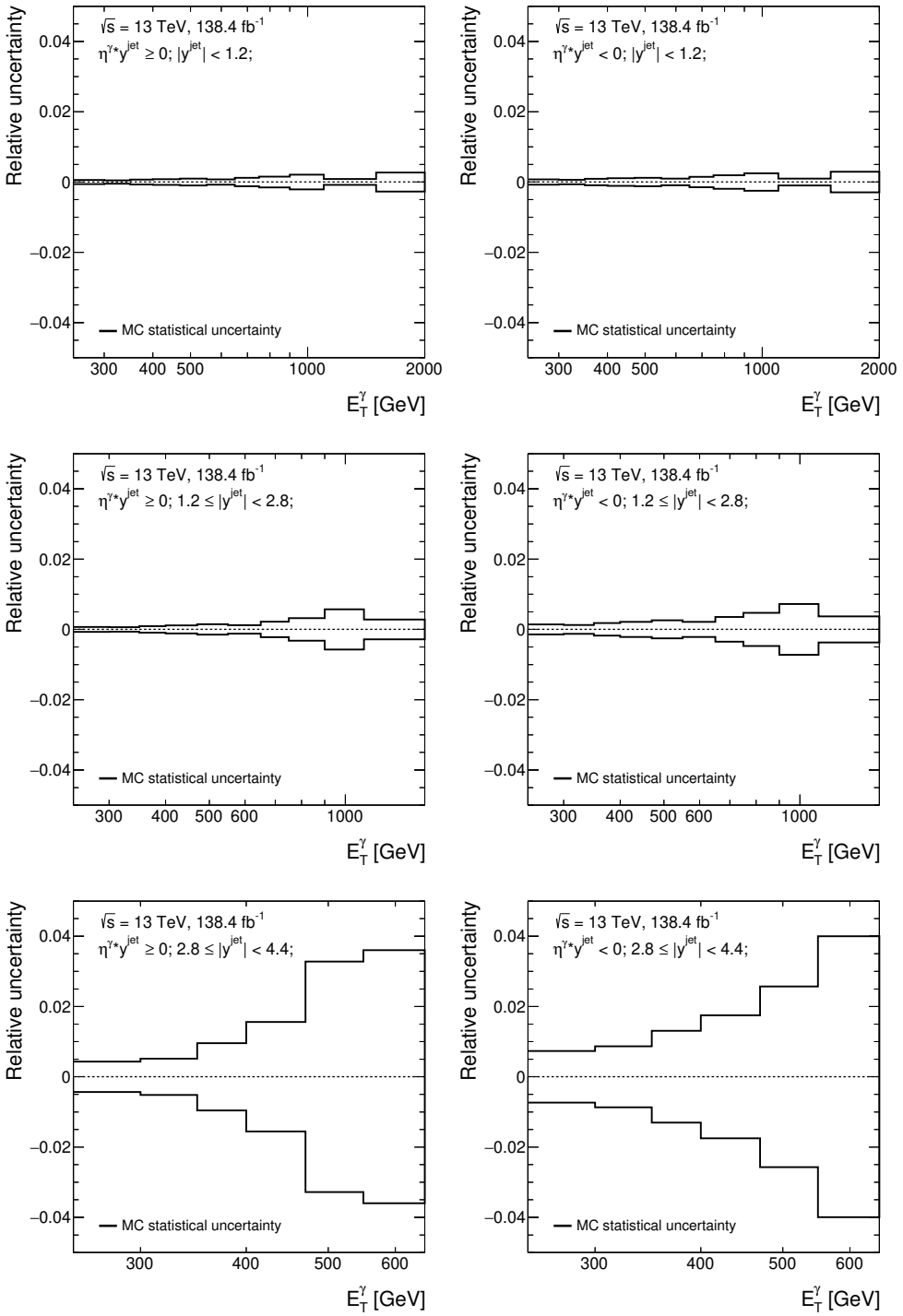


Figure 10.27: Statistical uncertainty of the MC samples as a function of E_{T}^{γ} in the different measured configurations.

10.7.5 Running conditions

Pile-up

A variation in the pile-up reweighting of simulated events is included to cover the uncertainty in the ratio of the predicted to measured inelastic cross sections [49]. This uncertainty was estimated following the same approach described in section 8.7.5. The resulting uncertainty in the measured cross section is shown in figure 10.28, and is approximately less than 1.5%.

Trigger efficiency

The uncertainty in the trigger efficiency is estimated by propagating into the cross section the relative uncertainties shown in table 8.2 for each running period separately. The uncertainty in the measured cross section is less than 0.1%.

Integrated luminosity

The uncertainty in the integrated luminosity is 1.7%. It is fully correlated between all bins of the measured cross section.

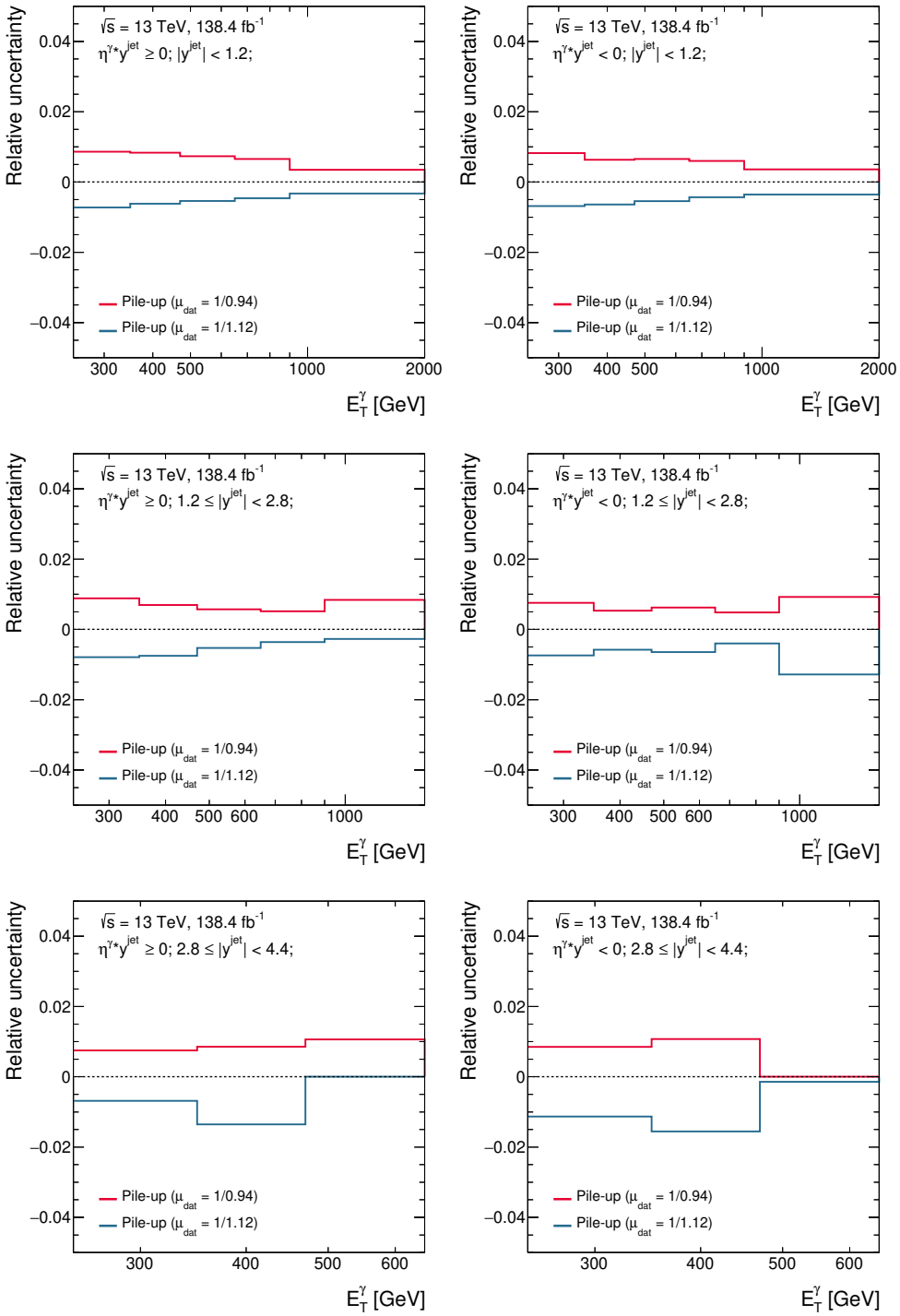


Figure 10.28: Systematic uncertainty in the measured cross section due to the pile-up reweighting as a function of E_{T}^{γ} in the different measured configurations. Statistical fluctuations in the estimation of this systematic uncertainty were taken care of by rebinning, merging those bins affected by relatively large statistical fluctuations with neighbouring bins.

10.7.6 Photon energy scale and resolution

The systematic uncertainties in the measured cross sections due to the uncertainty in the energy scale and resolution were studied following the same procedure described in section 7.7.6. A total of 79 individual components influencing the energy scale and resolution of the photon are identified to asses the overall uncertainty in the energy measurement. The resulting uncertainties in the measured cross section are typically less than 0.3% for the energy resolution and in the range $2\% - 9\%$ for the energy scale. The uncertainty in the measured cross section is shown in figure 10.29 (10.30) for the energy resolution (scale). Figure 10.31 shows the three main contributions to the photon energy scale in each measured configuration.

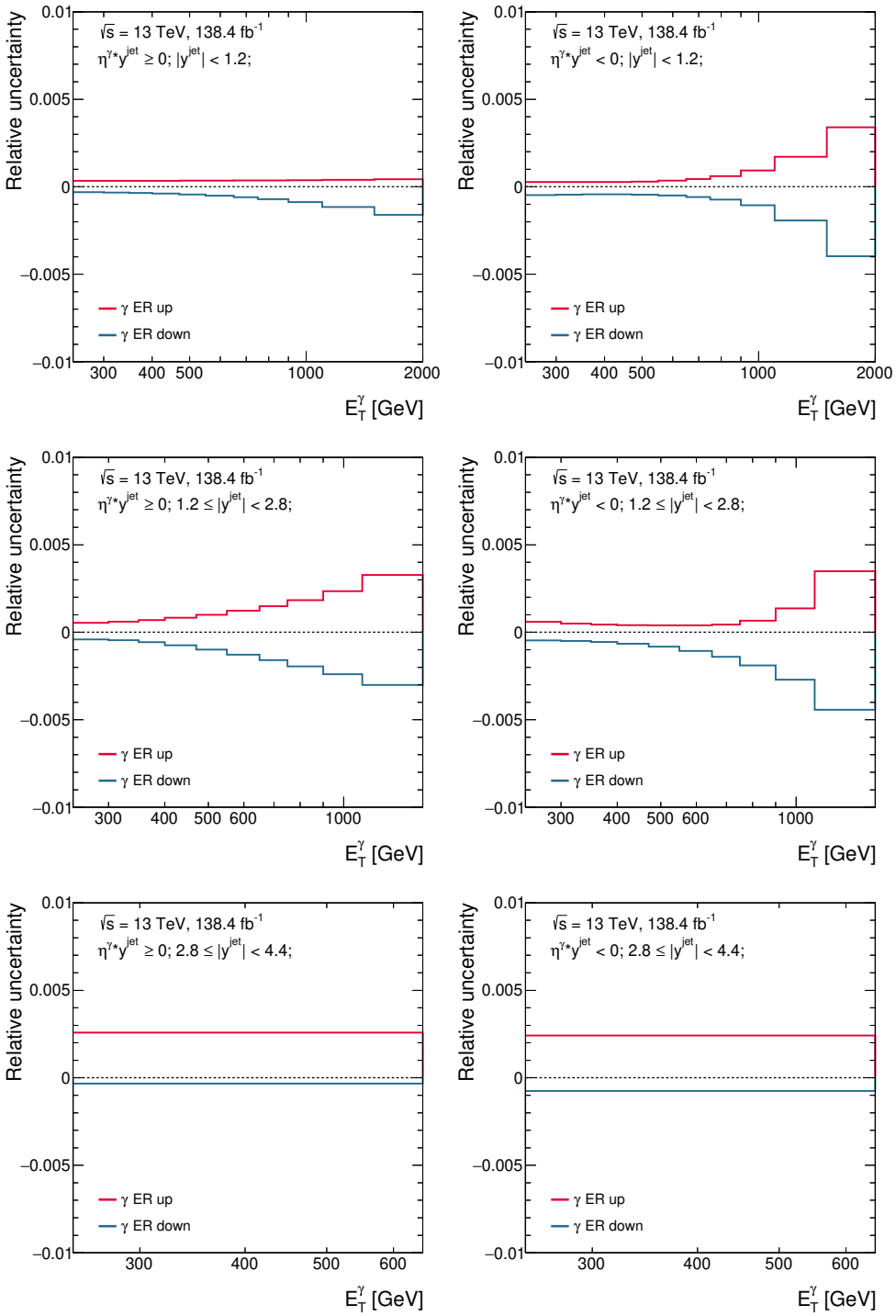


Figure 10.29: Systematic uncertainty in the measured cross section due to the photon energy resolution uncertainty as a function of E_{T}^{γ} in the different measured configurations. Statistical fluctuations in the estimation of this systematic uncertainty were taken care of by using a bootstrap technique on each individual component.

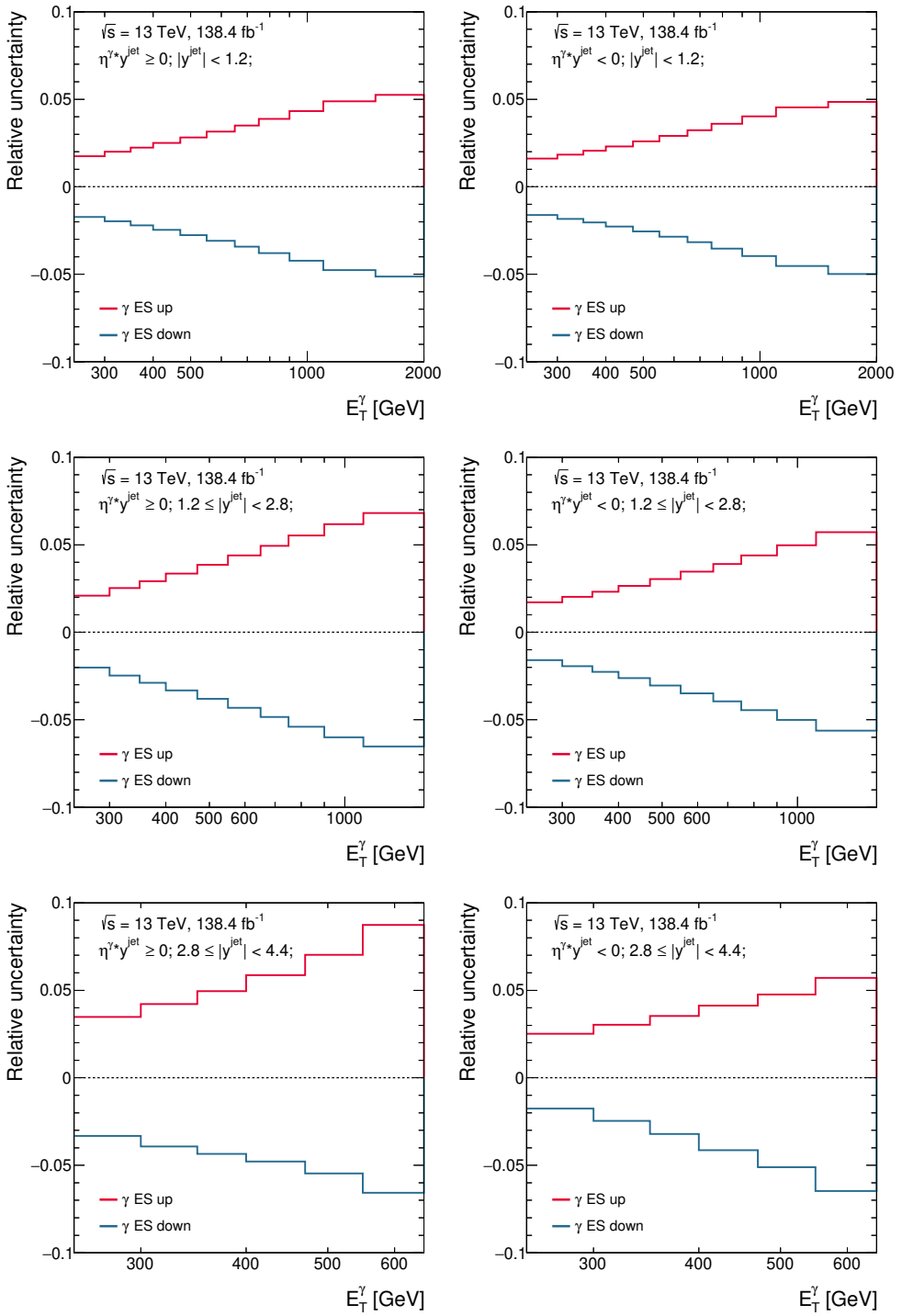


Figure 10.30: Systematic uncertainty in the measured cross section due to the photon energy scale uncertainty as a function of E_{T}^{γ} in the different measured configurations. Statistical fluctuations in the estimation of this systematic uncertainty were taken care of by using a bootstrap technique on each individual component.

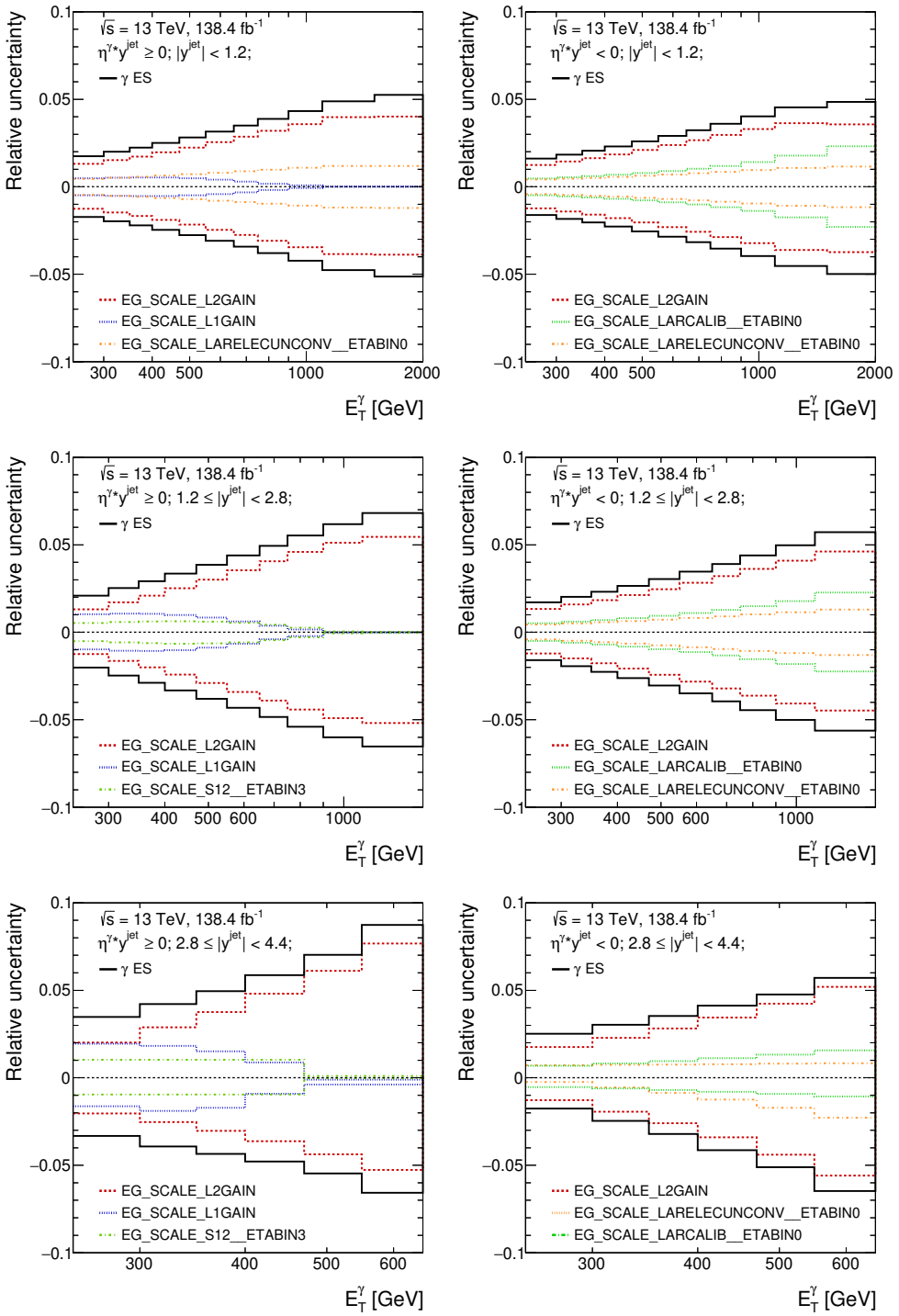


Figure 10.31: Systematic uncertainty in the measured cross section due to the photon energy scale uncertainty as a function of E_{T}^{γ} in the different measured configurations. The total uncertainty is represented with the black line; the three main contributions in each measured configuration are also shown.

10.7.7 Jet energy scale and resolution

The latest analysis of the sources of uncertainty that affect the measurement of the energy scale and resolution of jets was recently published in ref. [26]. These different sources of uncertainty were discussed in section 6.2.3, and are represented by a total of 125 individual components influencing the energy measurement of the jets. These parameters were propagated through the analysis separately to maintain the full information on the correlations. The systematic uncertainty in the cross section is evaluated by varying each individual source of uncertainty separately by $\pm 1\sigma$ in the MC simulations and then adding the uncertainty contributions in quadrature. The resulting uncertainties in the measured cross section are less than 0.2% for the energy resolution and typically less than 1% for the energy scale, except in the most forward region, $2.8 \leq |y^{\text{jet}}| < 4.4$, in both SS and OS configurations, where the uncertainty is in the range $1\% - 6\%$. The uncertainty in the measured cross section is shown in figure 10.32 (10.33) for the energy resolution (scale). Figure 10.34 shows the three main contributions to the jet energy scale in each measured configuration.

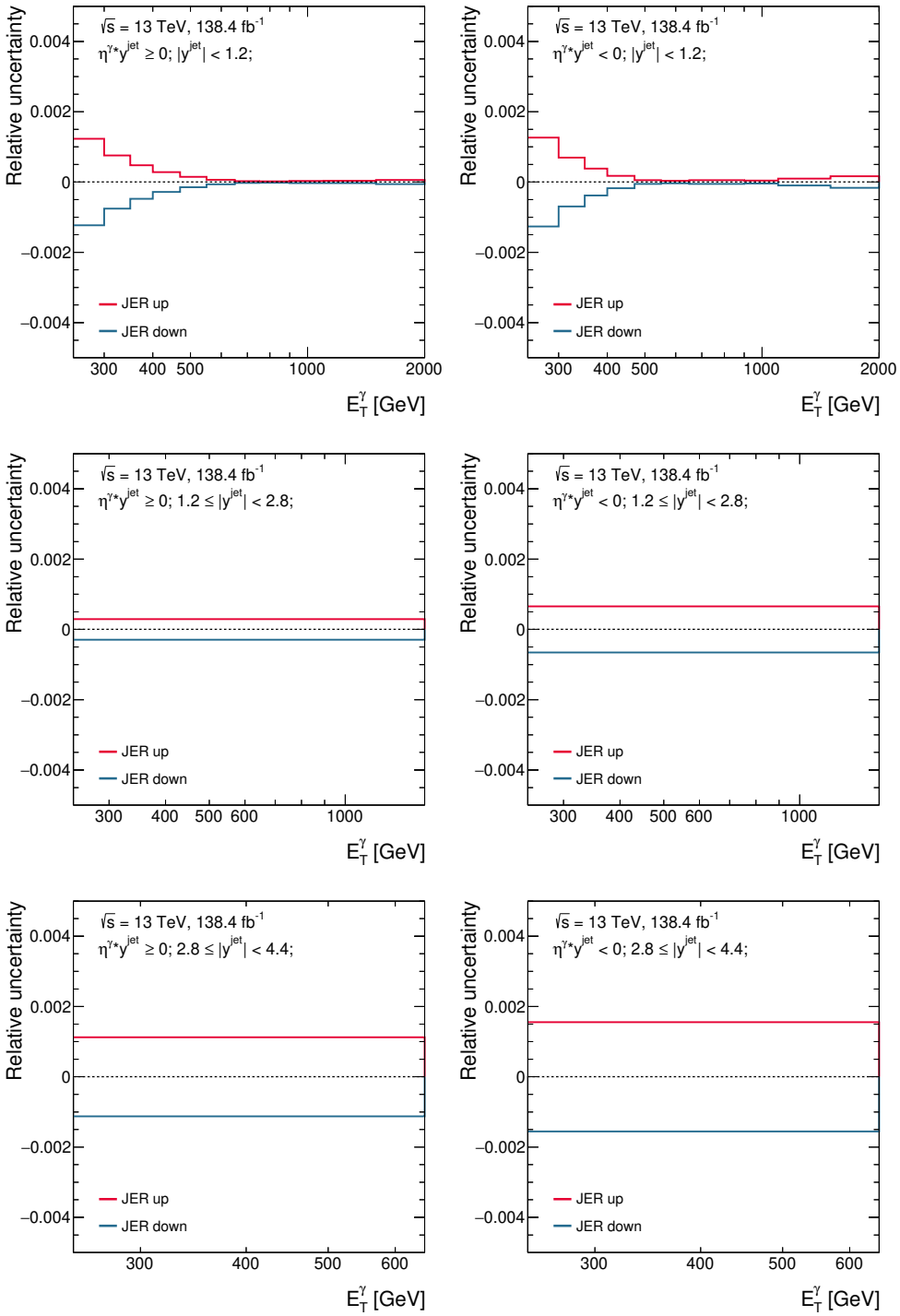


Figure 10.32: Systematic uncertainty in the measured cross section due to the jet energy resolution uncertainty as a function of E_{T}^{γ} in the different measured configurations. Statistical fluctuations in the estimation of this systematic uncertainty were taken care of by using a bootstrap technique on each individual component.

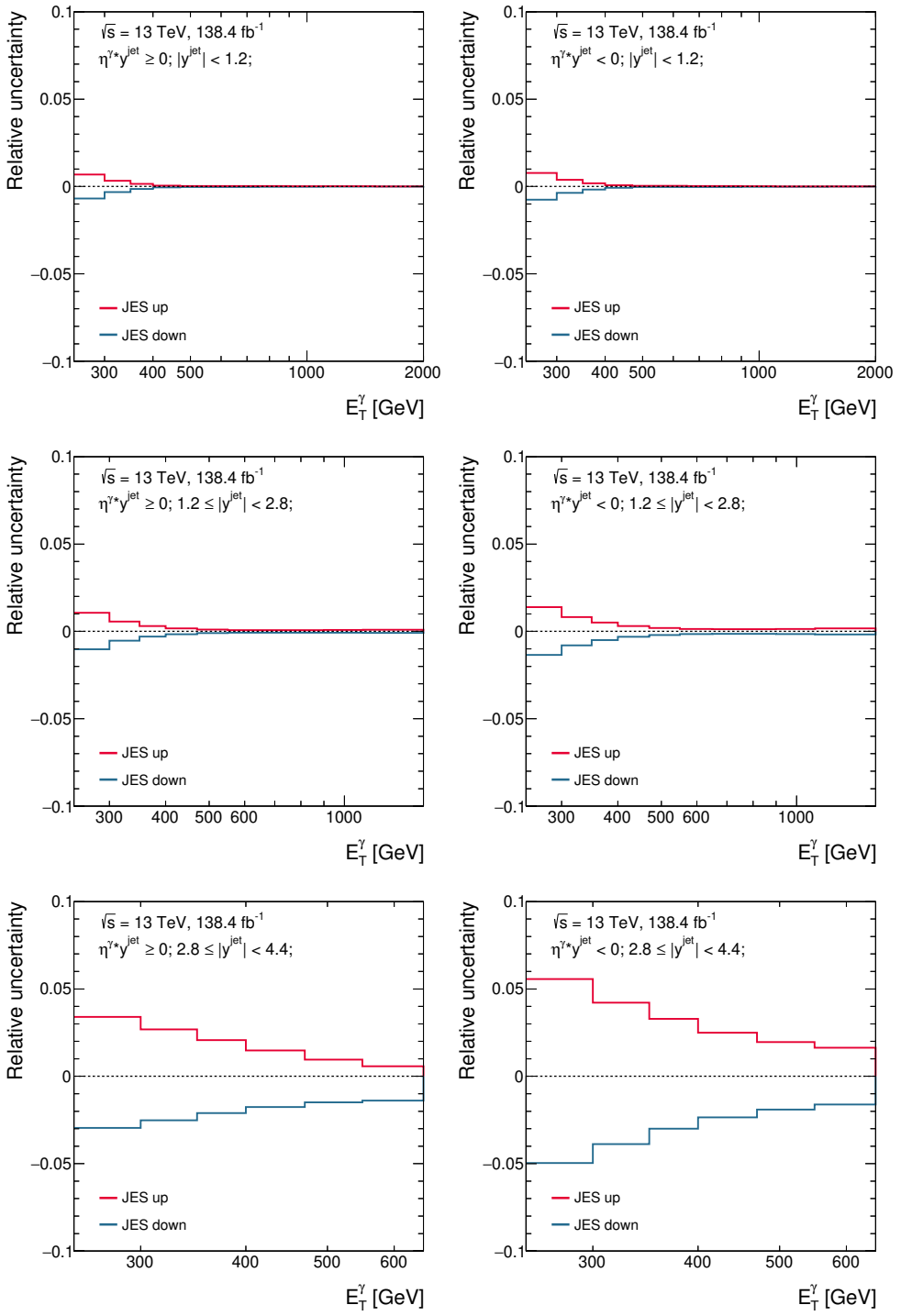


Figure 10.33: Systematic uncertainty in the measured cross section due to the jet energy scale uncertainty as a function of E_{T}^{γ} in the different measured configurations. Statistical fluctuations in the estimation of this systematic uncertainty were taken care of by using a bootstrap technique on each individual component.

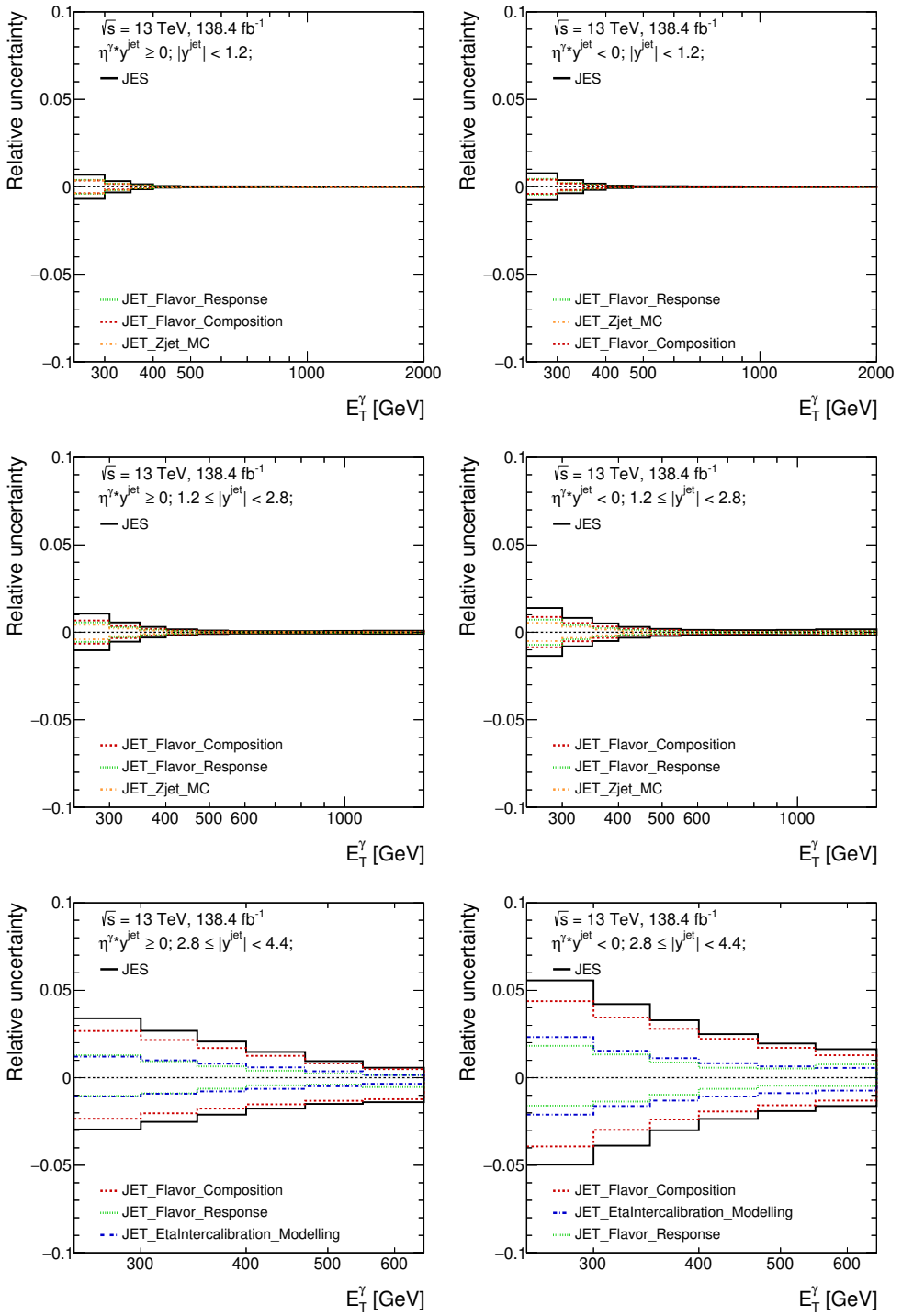


Figure 10.34: Systematic uncertainty in the measured cross section due to the jet energy scale uncertainty as a function of E_{T}^{γ} in the different measured configurations. The total uncertainty is represented with the black line; the three main contributions in each measured configuration are also shown.

10.7.8 Total systematic uncertainty

The total systematic uncertainty is calculated by adding in quadrature the uncertainties listed above. The total systematic is in the range 3%−11%, depending on E_{T}^{γ} and the $|y^{\mathrm{jet}}|$ region and photon and jet angular configuration. The dominant sources of uncertainty arise from the photon energy scale, the integrated luminosity and the pile-up. The total systematic uncertainty is shown in figure 10.35 together with the dominant components. The comparison of the total systematic uncertainty with the statistical uncertainty of the data is provided in figure 10.36. The systematic uncertainty dominates the total experimental uncertainty for $E_{\mathrm{T}}^{\gamma} \lesssim 1.5$ TeV for $|y^{\mathrm{jet}}| < 1.2$; for $1.2 \leq |y^{\mathrm{jet}}| < 2.8$, the systematic uncertainty dominates for $E_{\mathrm{T}}^{\gamma} \lesssim 1.1$ TeV ($E_{\mathrm{T}}^{\gamma} \lesssim 900$ GeV) in the SS (OS) configuration; in the most forward region $2.8 \leq |y^{\mathrm{jet}}| < 4,4$, the statistical uncertainty limits the precision of the measurement for $E_{\mathrm{T}}^{\gamma} \gtrsim 470$ GeV.

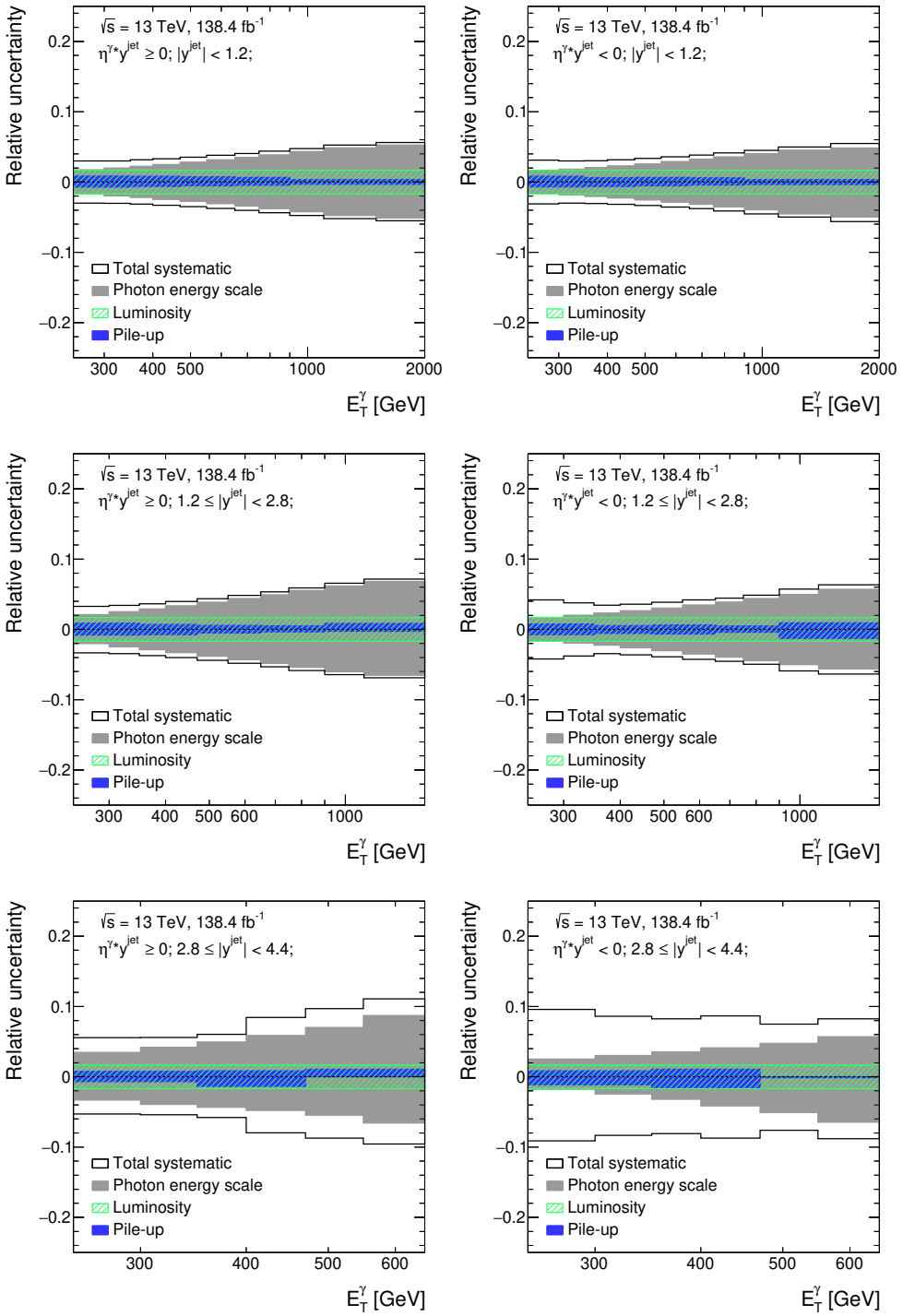


Figure 10.35: The relative total systematic uncertainty in the cross section (white areas) as a function of E_{T}^{γ} in the different measured configurations. The relative uncertainty due to the photon energy scale (grey areas), the relative uncertainty due to that on the integrated luminosity (green hatched areas), and the relative uncertainty due to the pile-up (blue areas) are also shown.

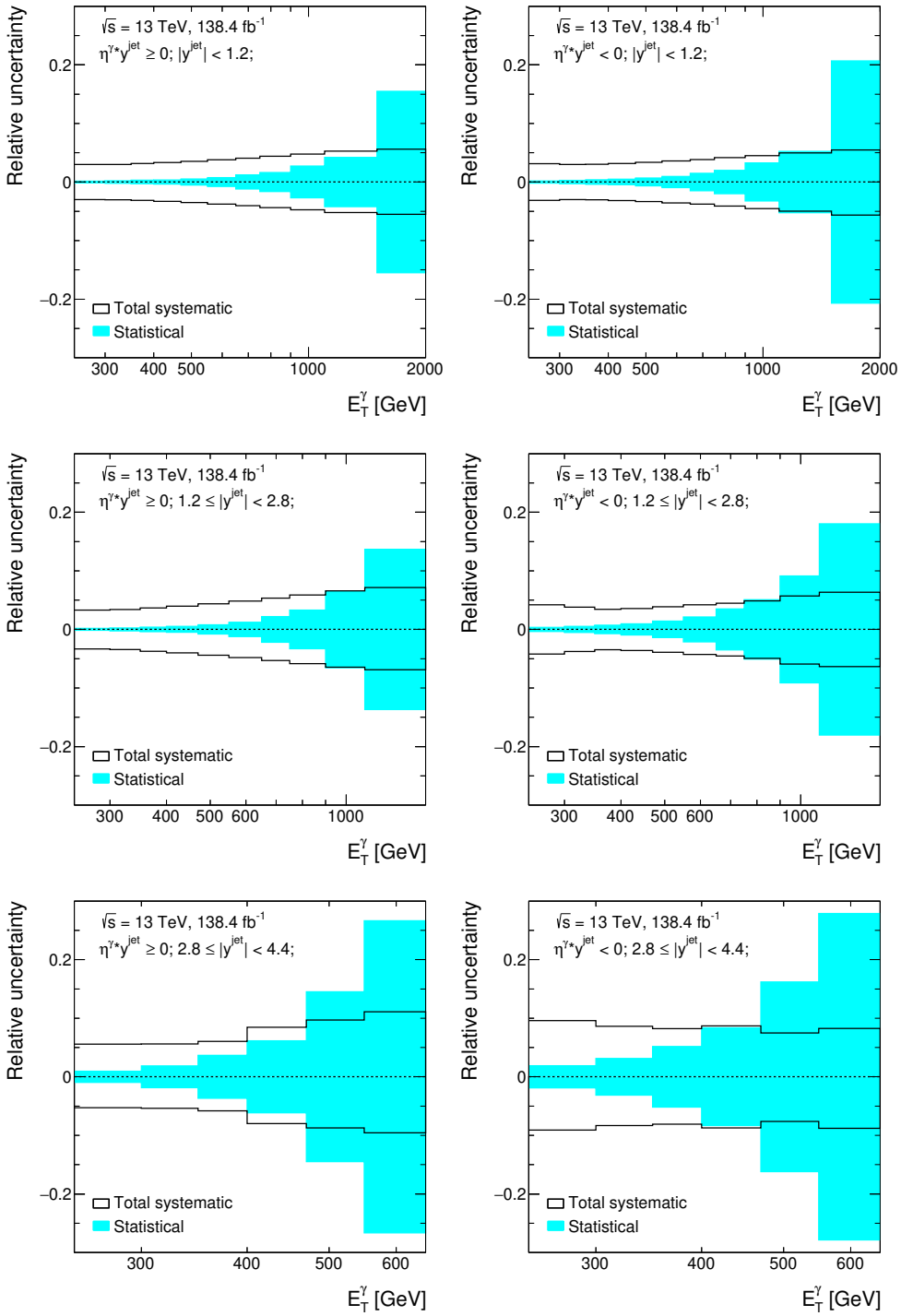


Figure 10.36: The relative total systematic uncertainty in the cross section (white areas) as a function of E_{T}^{γ} in the different measured configurations. The relative statistical uncertainty from the data is also shown (cyan areas).

10.8 Theoretical predictions

A NLO pQCD prediction was calculated using SHERPA 2.2.2, which was previously introduced in section 4.3. The main properties of the calculation were summarised in table 7.3.

10.8.1 Theoretical uncertainties in the NLO pQCD calculations

The uncertainty due to the terms beyond NLO was evaluated following the same procedure described in section 7.8.2. The resulting uncertainty in the prediction is in the range $15\% - 55\%$. A significant increase of this uncertainty is observed in the most forward $|y^{\text{jet}}|$ region, for both SS and OS configurations.

The uncertainty in the prediction of SHERPA 2.2.2 due to the uncertainty in the proton PDF was estimated using the 100 replicas from the NNPDF3.0 analysis. The resulting uncertainty in the prediction is in the range $1\% - 5\%$, except for the two angular configurations in the most forward $|y^{\text{jet}}|$ region, where the uncertainty reaches up to 13%. This uncertainty increases with E_{T}^{γ}, showing the potential of this measurement to constrain the PDF at high E_{T}^{γ}.

The uncertainty in the prediction of SHERPA 2.2.2 due to the uncertainty in α_{S} is evaluated by repeating the calculations using two additional sets of proton PDF, and following the same procedure described in section 7.8.2. The resulting uncertainty in the prediction is less than 6%, except for the SS (OS) configuration in the most forward $|y^{\text{jet}}|$ region, where the uncertainty reaches up to 10% (23%).

The total theoretical uncertainty was computed by adding in quadrature the individual uncertainties listed above. Figure 10.37 shows the total uncertainty together with the individual components. The dominant source of uncertainty is that arising from the variation of the scales; the uncertainty from α_{S} is the second dominant. There is a significant increase of all the sources of uncertainty in the most forward $2.8 \leq |y^{\text{jet}}| < 4.4$ region.

The sensitivity of the observables to the proton PDF was investigated by comparing the NLO pQCD prediction of SHERPA 2.2.2 based on the NNPDF3.0 PDF with alternative predictions based on the MMHT2014 and the CT14 PDF sets. Figure 10.38 shows the ratio of the calculations of SHERPA 2.2.2 based on the MMHT2014 and CT14 PDF sets to that using the NNPDF3.0 PDF. In the region $|y^{\text{jet}}| < 1.2$, the predictions using the MMHT2014 and CT14 PDF sets are more similar at low E_{T}^{γ} compared to NNPDF3.0; larger discrepancies arise at high E_{T}^{γ} values in terms of shape between the three PDF sets. In the forward region, $1.2 \leq |y^{\text{jet}}| < 2.8$, in the SS configurations the calculations with the different PDF sets agree well within a 2%; in the OS configuration, differences in normalisation and shape are observed along the E_{T}^{γ} spectrum. In the region $2.8 \leq |y^{\text{jet}}| < 4.4$ the prediction of CT14 is more similar to that using NNPDF3.0 except at high E_{T}^{γ}; the prediction of MMHT2014 shows larger differences along E_{T}^{γ}.

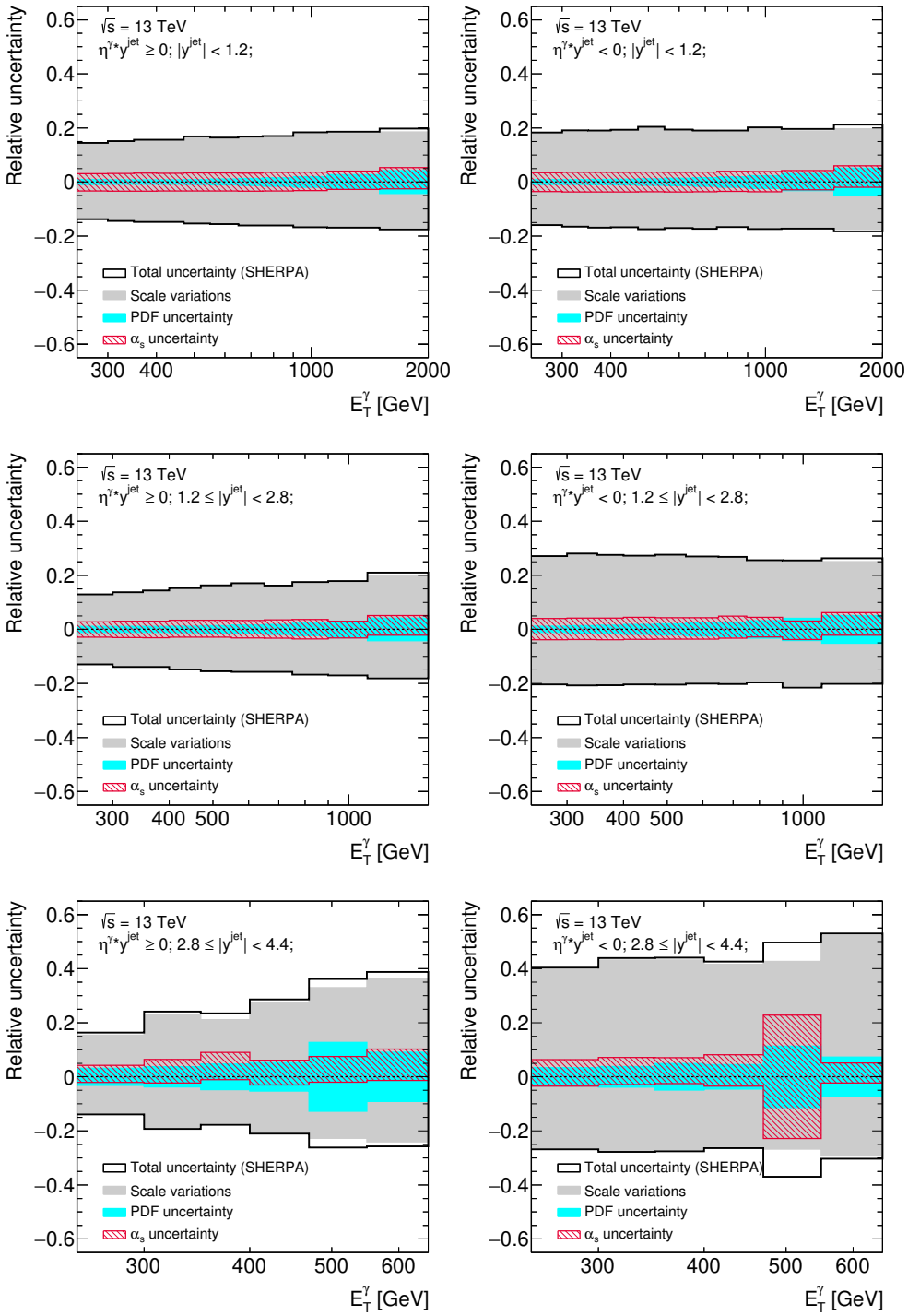


Figure 10.37: The relative total theoretical uncertainty in the cross-section prediction of SHERPA 2.2.2 as a function of E_{T}^{γ} in the different measured configurations. The uncertainty from the scale variations, the uncertainty from the PDF and the uncertainty from the value of α_{S} are also included.

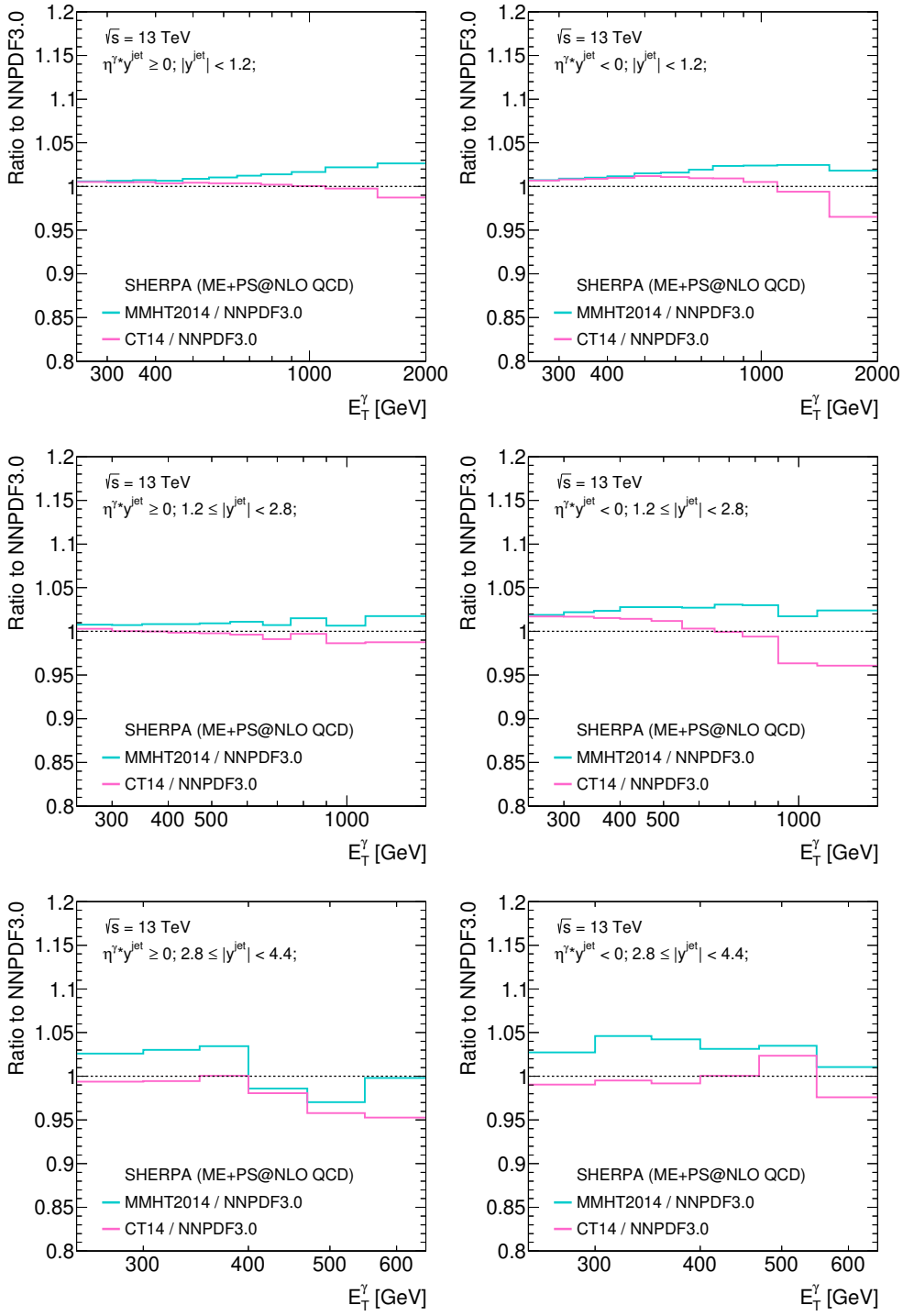


Figure 10.38: Ratio of the SHERPA 2.2.2 predictions based on the MMHT2014 (cyan lines) and CT14 (pink lines) PDF sets to that based on the NNPDF3.0 PDF as a function of E_{T}^{γ} in the different measured configurations.

10.9 Results

The differential cross section for isolated photon in association with one jet was measured in pp collisions, $pp \to \gamma + \text{jet} + X$, at $\sqrt{s} = 13$ TeV, using 138.4 fb^{-1} of ATLAS data. Photons were required to have $E_{\text{T}}^{\gamma} > 250$ GeV, $|\eta^{\gamma}| < 2.37$ (excluding the region $1.37 < |\eta^{\gamma}| < 1.56$), and $E_{\text{T}}^{\text{iso}} < 4.2 \times 10^{-3} \times E_{\text{T}}^{\gamma} + 4.8$ GeV. Jets were reconstructed using the anti-k_{t} algorithm with $R = 0.4$ and using topoclusters as inputs; the leading jet was selected with $p_{\text{T}}^{\text{jet}} > 165$ GeV and $|y^{\text{jet}}| < 4.4$. The differential cross section was measured in two different angular configurations of the jet and the photon, OS and SS, and for three different $|y^{\text{jet}}|$ ranges, namely $|y^{\text{jet}}| < 1.2$, $1.2 \leq |y^{\text{jet}}| < 2.8$, and $2.8 \leq |y^{\text{jet}}| < 4.4$. Values of E_{T}^{γ} up to 2 TeV were measured with the current integrated luminosity.

The measured differential cross section for photon plus jet production is shown in figure 10.39. The predictions of the leading-logarithm parton-shower MC models of PYTHIA and SHERPA are compared to the measurement. These predictions were normalised to the measured integrated cross section in each observable. The difference in normalisation between the measured cross section and SHERPA (PYTHIA) is $\approx 30\% - 40\%$ ($\approx 10\% - 20\%$) for the central $|y^{\text{jet}}| < 1.2$ and forward $1.2 \leq |y^{\text{jet}}| < 2.8$ rapidity ranges; in the most forward $|y^{\text{jet}}|$ region, the difference in normalisation is $\approx 20\%$ ($\approx 10\%$) in the SS (OS) configuration for SHERPA and $\approx 10\%$ ($\approx 40\%$) for PYTHIA. The calculations of SHERPA and PYTHIA describe well the data in the range $|y^{\text{jet}}| < 1.2$ except at high E_{T}^{γ}; in the forward region $1.2 \leq |y^{\text{jet}}| < 2.8$ both predictions describe the data well in the SS configuration, whereas the OS is only well described at low E_{T}^{γ}; in the range $2.8 \leq |y^{\text{jet}}| < 4.4$, SHERPA and PYTHIA describe well the measurement in the SS configuration, while the OS data distribution is only well described by SHERPA at low E_{T}^{γ} and PYTHIA fails describing the data in the whole E_{T}^{γ} range.

The NLO pQCD prediction of SHERPA 2.2.2 is compared with the measurement in figure 10.40. The full theoretical uncertainty, including the variations of the scales, is shown for the prediction based on the NNPDF3.0 PDF set. The predictions based on the MMHT2014 and CT14 PDF sets are also included, showing the sensitivity of this measurement to the proton PDF. A detailed description of the comparisons between the measurement and the NLO pQCD predictions is provided below.

Range $|\mathbf{y}^{\text{jet}}| < 1.2$

In this $|y^{\text{jet}}|$ region, values up to $E_{\text{T}}^{\gamma} = 2$ TeV are measured for both SS and OS configurations. The shape of the measured cross sections for the SS and OS configurations is similar. The prediction of SHERPA 2.2.2 shows a good agreement with the measurement, within the experimental and theoretical uncertainties. In the OS configuration, the prediction from SHERPA 2.2.2 shows a tendency to be above the data. The experimental uncertainty is below 7% for $E_{\text{T}}^{\gamma} \lesssim 1.5$ TeV and it grows up to $\approx 15\%$ (20%) in the last bin $1500 < E_{\text{T}}^{\gamma} < 2000$ GeV for the SS (OS) configuration, being dominated by the statistical uncertainty. The theoretical uncertainty is in the range $15\% - 20\%$, and increases slightly with E_{T}^{γ} in the SS configuration, while it shows a more constant shape for the OS. The

uncertainty due to the PDF is largest at $E_{\mathrm{T}}^{\gamma} > 1.5$ TeV.

Range $1.2 \leq |\mathbf{y}^{\mathrm{jet}}| < 2.8$

The differential cross sections are measured for values up to $E_{\mathrm{T}}^{\gamma} = 1.5$ TeV in this $|y^{\mathrm{jet}}|$ region. The NLO pQCD prediction of SHERPA 2.2.2 gives an excellent description of the data in the SS configuration; in the OS, the measured cross section and the prediction agree within uncertainties, and the prediction exhibits a tendency to overestimate the data. The experimental uncertainty is below 7% for $E_{\mathrm{T}}^{\gamma} \lesssim 900$ GeV and it grows up to $\approx 15\%$ (18%) in the last bin $1100 < E_{\mathrm{T}}^{\gamma} < 1500$ GeV for the SS (OS) configuration, where it is dominated by the statistical uncertainty. The uncertainty on the theoretical calculations of SHERPA 2.2.2 is in the range $13\% - 21\%$ ($20\% - 28\%$) for the SS (OS) configuration. The theoretical uncertainty increases with E_{T}^{γ} in the SS configuration, while its shape is rather constant in the OS. A certain asymmetry is observed in the size of the up and down components of the scale uncertainties in the OS observable. The uncertainty due to the PDF is largest at $E_{\mathrm{T}}^{\gamma} > 1.1$ TeV.

Range $2.8 \leq |\mathbf{y}^{\mathrm{jet}}| < 4.4$

Values of E_{T}^{γ} up to 650 GeV are measured for both SS and OS configurations. The prediction of SHERPA 2.2.2 describes well data in the SS configuration; in the OS configuration, there is a tendency in the theory to overestimate the data. The experimental uncertainty is below 11% (12%) for $E_{\mathrm{T}}^{\gamma} \lesssim 470$ GeV and it grows up to $\approx 28\%$ (30%) in the last bin $550 < E_{\mathrm{T}}^{\gamma} < 650$ GeV for the SS (OS) configuration. The experimental uncertainty is dominated by the statistical uncertainty of the data for $E_{\mathrm{T}}^{\gamma} > 470$ GeV. The theoretical uncertainty significantly increases with E_{T}^{γ} in the SS configuration and is in the range $14\% - 38\%$; in the OS configuration, the increase with E_{T}^{γ} is less pronounced and the uncertainty is in the range $27\% - 53\%$. This is the $|y^{\mathrm{jet}}|$ range in which the asymmetry between the down and up components of the theoretical uncertainty is more significant. The theoretical uncertainty is dominated by the terms beyond NLO, but the uncertainty due to the PDF is the largest of all the configurations, providing the highest potential to constrain further the PDF.

For the SS and OS configurations, the x and Q^2 regions accessed by the measurements are included in table 10.2. The measurement covers the region $x \gtrsim 2 \times 10^{-3}$ and $6.25 \times 10^4 \lesssim Q^2 \lesssim 4 \times 10^6$ GeV^2.

Configuration	x mean value	Q^2 mean value [GeV^2]
Region $\lvert y^{\text{jet}}\rvert < 1.2$; $\eta^\gamma \cdot y^{\text{jet}} \geq 0$	0.080	1.2×10^5
Region $\lvert y^{\text{jet}}\rvert < 1.2$; $\eta^\gamma \cdot y^{\text{jet}} < 0$	0.073	1.2×10^5
Region $1.2 \leq \lvert y^{\text{jet}}\rvert < 2.8$; $\eta^\gamma \cdot y^{\text{jet}} \geq 0$	0.133	1.0×10^5
Region $1.2 \leq \lvert y^{\text{jet}}\rvert < 2.8$; $\eta^\gamma \cdot y^{\text{jet}} < 0$	0.116	1.1×10^5
Region $2.8 \leq \lvert y^{\text{jet}}\rvert < 4.4$; $\eta^\gamma \cdot y^{\text{jet}} \geq 0$	0.280	8.5×10^4
Region $2.8 \leq \lvert y^{\text{jet}}\rvert < 4.4$; $\eta^\gamma \cdot y^{\text{jet}} < 0$	0.275	9.2×10^4

Table 10.2: Mean values of x and Q^2 accessed by the measurements for the SS and OS configurations and depending on the different ranges in $|y^{\text{jet}}|$.

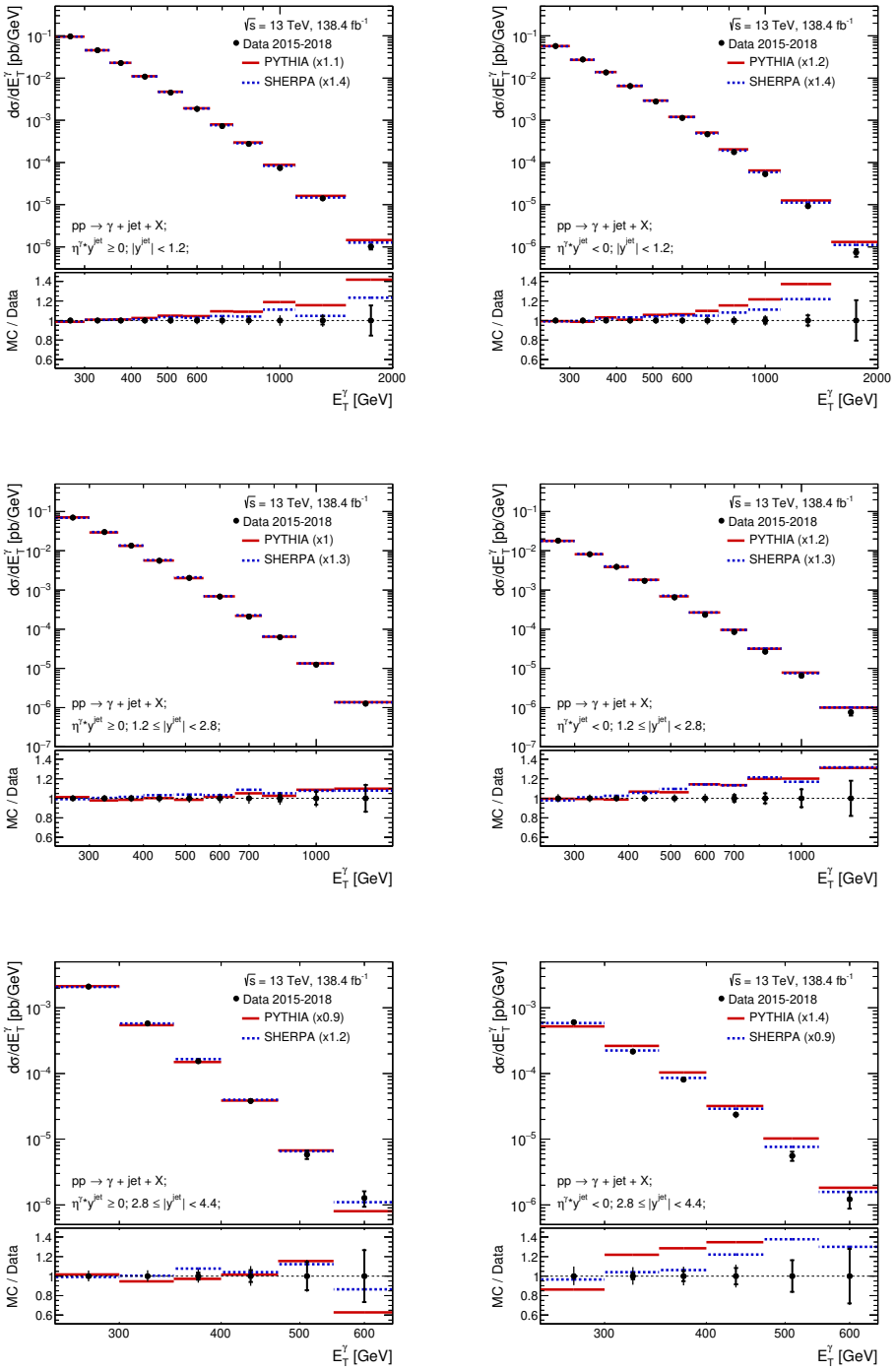


Figure 10.39: Measured cross section for photon plus jet production (dots) as a function of E_{T}^{γ} in the different measured configurations. The predictions from PYTHIA (solid lines) and SHERPA (dashed lines) are also shown; these predictions are normalised to the measured integrated cross section in each observable using the values indicated in parentheses. The bottom part of each figure shows the ratio of the MC predictions to the measured cross section. The inner (outer) error bars represent the statistical uncertainties (the statistical and systematic uncertainties added in quadrature). For most of the points, the inner error bars are smaller than the marker size and, thus, not visible.

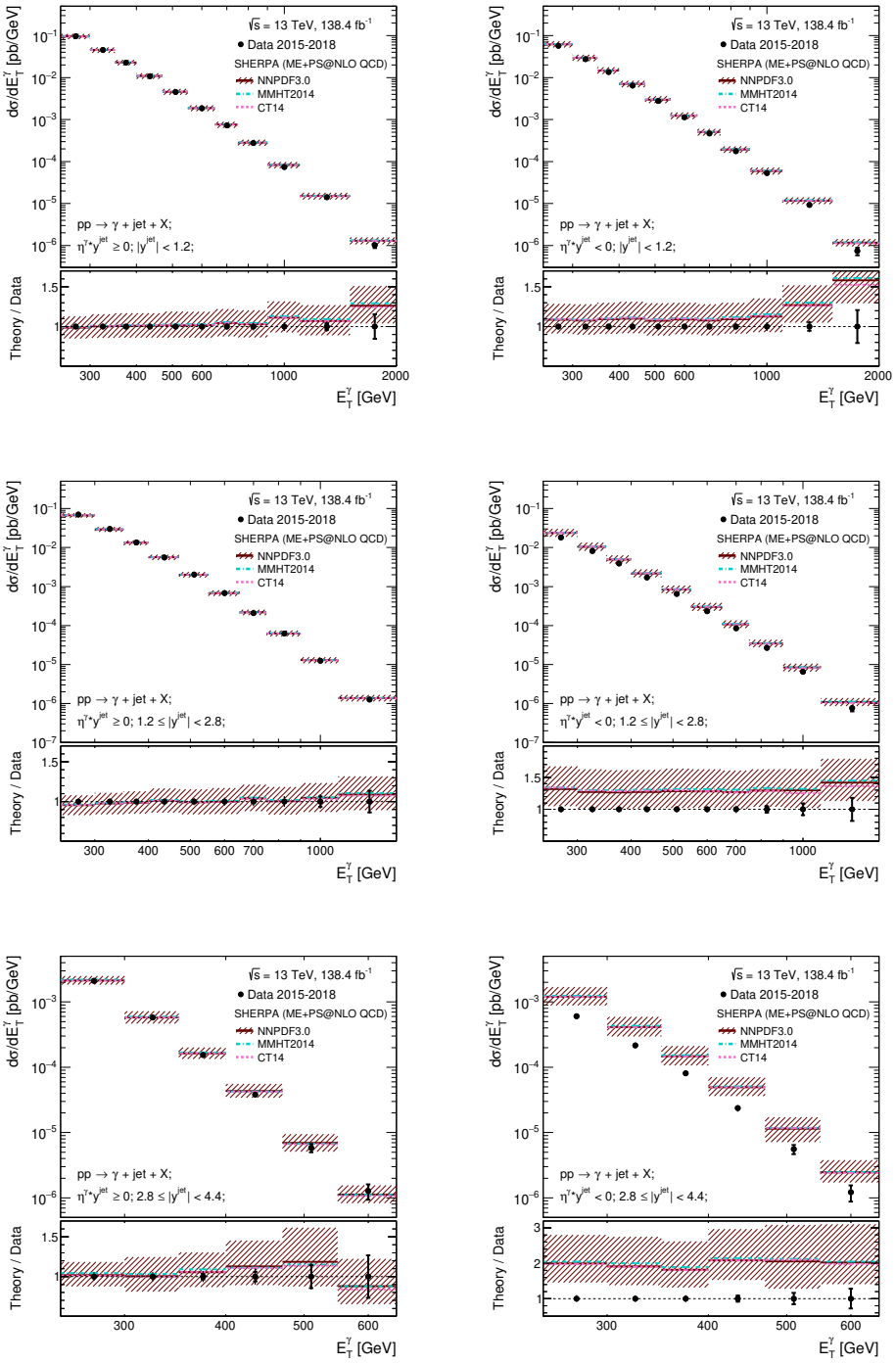


Figure 10.40: Measured cross section for photon plus jet production (dots) as a function of E_{T}^{γ} in the different measured configurations. The NLO pQCD prediction from SHERPA 2.2.2 using the NNPDF3.0 PDF set is also shown (red lines). For comparison, the predictions using the MMHT2014 (cyan lines) and CT14 (pink lines) PDF sets are also included. The band represents the full theoretical uncertainty associated to the prediction based on the NNPDF3.0 PDF set. The bottom part of each figure shows the ratio of the prediction to the measured cross section. The inner (outer) error bars represent the statistical uncertainties (the statistical and systematic uncertainties added in quadrature). For most of the points, the inner error bars are smaller than the marker size and, thus, not visible.

Chapter 11

Summary and conclusions

In this dissertation, the production of isolated prompt photons, inclusively and in association with one jet, was studied in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector at the LHC. A detailed overview of the LHC accelerator and the ATLAS detector was provided, followed by a description of the main characteristics of the theoretical framework underlying this work. Descriptions of the theory of prompt-photon production, jet algorithms and Monte Carlo generators were included in the initial chapters of the thesis. Furthermore, a description of the reconstruction and identification of photons and the reconstruction of jets was provided, as well as a description of their energy scale and resolution calibration.

Although the different data samples used in the measurements presented in this dissertation are enriched in isolated prompt photons, they still contain some background contamination even after the application of the identification and isolation requirements on the photon candidates. This background originates predominantly from multijet processes in which a jet is misidentified as a photon; this jet usually contains a light neutral meson, mainly a π^0, which decays into a collimated pair of photons and carries most of the energy of the jet. For the removal of such background, a method based on sidebands in the two-dimensional plane spanned by the isolation transverse energy and the discrete variable which encapsulates the shower-shape properties of the photon was used. This method was complemented with the estimates of the contribution of isolated-prompt photons (signal leakage fractions) to the background-enriched regions to improve the determination of the signal. As a result, the purity of the different samples was found to be $\geq 90\%$, except for the lowest bins of the photon transverse momentum in the in situ calibration measurement, where the purity is higher than 70%.

Monte Carlo samples of prompt-photon plus jets events were generated using PYTHIA, SHERPA 2.1.1 and SHERPA 2.2.2 programs. These samples were used for various purposes: the determination of the signal leakage fractions used in the background subtraction; the evaluation of the acceptance corrections to obtain the measurement of the differential cross sections as well as the associated uncertainties; the evaluation of the in

situ calibration factors and their systematic uncertainties. The comparisons between the data and MC distributions of kinematic variables show, in general, a reasonable agreement.

The cross section for inclusive isolated prompt-photon production in pp collisions at $\sqrt{s} = 13$ *TeV was measured using data with an integrated luminosity of 36.1* fb^{-1} *collected by the ATLAS detector at the LHC.* The differential cross section was measured as a function of E_{T}^{γ} in four regions of $|\eta^{\gamma}|$ for photons with $E_{\mathrm{T}}^{\gamma} > 125$ GeV and $|\eta^{\gamma}| < 2.37$, excluding the region $1.37 < |\eta^{\gamma}| < 1.56$.

This measurement constitutes an improvement in terms of statistical and systematic uncertainties relative to that published earlier thanks to the ten-fold increase in the integrated luminosity. In this measurement, the reach in E_{T}^{γ} is extended upwards by 1000, 500, 400 and 200 GeV for $|\eta^{\gamma}| < 0.6$, $0.6 < |\eta^{\gamma}| < 1.37$, $1.56 < |\eta^{\gamma}| < 1.81$ and $1.81 < |\eta^{\gamma}| < 2.37$, respectively, relative to the previous measurement at $\sqrt{s} = 13$ TeV. The range in which the measurement is limited by systematic uncertainties is extended upwards to 1 TeV for $|\eta^{\gamma}| < 1.37$, where previously the limitation was at 600 GeV.

Next-to-leading-order perturbative QCD predictions using several PDF sets were compared with the measurement and provide an adequate description of the data within experimental and theoretical uncertainties. The comparison of the data and theory is limited by the theoretical uncertainty due to the terms beyond NLO in QCD. The experimental systematic uncertainties are smaller than the theoretical uncertainties over the full investigated phase space. The inclusion of higher-order matrix elements and parton showers, as provided by the SHERPA 2.2.2 program, improves the description of the normalisation of the data by the prediction.

Recently, a next-to-next-to-leading-order QCD prediction was calculated in which the uncertainties due to the scale variations are reduced by a factor in the range $2 - 20$ relative to those of a NLO QCD calculation. Overall, the NNLO QCD prediction gives an excellent description of the data. The comparison of the NNLO prediction with the measured cross section represents a precise test of the theory in the range of photon transverse energies from 125 GeV up to and beyond 1 TeV. The measurement has the potential to further constrain the PDF of the proton, specially the gluon density, within a global NNLO QCD fit.

The cross section for inclusive-photon production in pp collisions at $\sqrt{s} = 13$ *TeV was measured using the ATLAS detector at the LHC with an integrated luminosity of 139* fb^{-1}*.* The differential cross section was measured as a function of E_{T}^{γ} in six regions of $|\eta^{\gamma}|$ for photons with $E_{\mathrm{T}}^{\gamma} > 250$ GeV and $|\eta^{\gamma}| < 2.37$, excluding the region $1.37 < |\eta^{\gamma}| < 1.56$. The photon isolation is ensured by requiring that the transverse energy in a cone of $R = 0.2$ or $R = 0.4$ around the photon is smaller than $4.2 \times 10^{-3} \times E_{\mathrm{T}}^{\gamma} + 4.8$ GeV.

This measurement represents an improvement with respect to those published earlier at $\sqrt{s} = 13$ TeV due to the new $|\eta^{\gamma}|$ regions measured: $|\eta^{\gamma}| < 0.6$, $0.6 < |\eta^{\gamma}| < 0.8$, $0.8 < |\eta^{\gamma}| < 1.37$, $1.56 < |\eta^{\gamma}| < 1.81$, $1.81 < |\eta^{\gamma}| < 2.01$, and $2.01 < |\eta^{\gamma}| < 2.37$, as opposed to four regions in the previous analyses. In this way, a more detailed experi-

mental information is provided for the PDF fits. Moreover, measurements for two photon isolation-cone radii, namely $R = 0.2$ and $R = 0.4$, are now provided, which allow a test of the dependence of the pQCD predictions with R. Furthermore, the new measurement benefits from reduced systematic uncertainties thanks to the improvements of the combined performance groups that develop the recommendations to evaluate the systematic uncertainties. In addition, thanks to the four-fold increase in the integrated luminosity, higher E_{T}^{γ} values are reached with smaller statistical uncertainties, and the range in which the measurement is limited by systematic uncertainties is extended upwards to ≈ 1.5 TeV for $|\eta^{\gamma}| < 0.6$ and $0.8 < |\eta^{\gamma}| < 1.37$, approximately 1.1 TeV for $0.6 < |\eta^{\gamma}| < 0.8$ and $1.56 < |\eta^{\gamma}| < 1.81$, and ≈ 900 GeV for $1.81 < |\eta^{\gamma}| < 2.37$.

Theoretical predictions at NLO in QCD were calculated using several PDF sets and provide a reasonable description of the data within experimental and theoretical uncertainties. As concluded for the previous inclusive photon measurement at $\sqrt{s} = 13$ TeV, the comparison of the measurement with NLO pQCD predictions is limited by the theoretical uncertainty due to the terms beyond NLO. The comparison of the measurements with NNLO QCD predictions, as provided for the previous inclusive photon measurement, is expected to give an improved description of the data, with significantly reduced theoretical uncertainties. Therefore, this high-precision measurement has the potential to constrain further the PDF of the proton within a global NNLO QCD fit.

The dependence with R of the measured cross section for inclusive isolated-photon production is adequately described by the predictions of SHERPA 2.2.2 and JETPHOX with $\mu_{\mathrm{R}} = \mu_{\mathrm{F}} = \mu_{\mathrm{f}} = E_{\mathrm{T}}^{\gamma}$, whereas the prediction of JETPHOX with $\mu_{\mathrm{R}} = \mu_{\mathrm{F}} = \mu_{\mathrm{f}} = E_{\mathrm{T}}^{\gamma}/2$ gives an improved description of the data. No dependence with the proton PDF of the predictions for the fiducial cross section as a function of R is observed.

The in situ jet energy scale and resolution were measured in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC with an integrated luminosity of 80 fb^{-1} and 58.5 fb^{-1}, respectively. The JES (JER) was measured as a function of the photon transverse momentum for photons with $p_{\mathrm{T}}^{\mathrm{ref}} > 25$ GeV and $|\eta^{\gamma}| < 1.37$, and the leading jet was selected with $p_{\mathrm{T}}^{\mathrm{jet}} > 10$ GeV and $|\eta| < 0.8$ ($|\eta| < 0.7$). The measurement of the JER was performed in two regions of the pseudorapidity of the leading jet, namely $|\eta| < 0.2$ and $0.2 < |\eta| < 0.7$. Both measurements were performed for two jet types, namely PFlow and EMtopo.

The measurement of the JES constitutes an improvement in terms of statistical and systematic uncertainties relative to that published earlier thanks to the twenty-five-fold increase in the integrated luminosity. In addition, this measurement benefits from the use of the MPF method and PFlow jets as the baseline configuration, which slightly improves the results obtained using EMtopo jets. The in situ correction derived with the MPF technique for both jet types is approximately 3% at low $p_{\mathrm{T}}^{\mathrm{ref}}$ and decreases to around 2% for photons above ≈ 70 GeV; the jet energy scale uncertainty is $\approx 1\%$ for most of the $p_{\mathrm{T}}^{\mathrm{ref}}$ range measured. Higher $p_{\mathrm{T}}^{\mathrm{jet}}$ values are reached with smaller statistical uncertainties, extending the limit of the calibration up to 1.2 TeV, where previously the limitation was

950 GeV. A complementary measurement of the jet energy scale was performed by means of the DB technique, obtaining consistent calibration factors and associated uncertainties.

The measurement of the JER was performed using the DB method with PFlow and EMtopo jet types in two different regions of pseudorapidity of the leading jet. The measured JER is reasonably described by the MC simulation within experimental uncertainties. A resolution in the range $19\% - 24\%$ is obtained for EMtopo jets at low $p_{\mathrm{T}}^{\mathrm{ref}}$ in the two different regions of pseudorapidity, whereas the resolution measured with PFlow jets is smaller than 17% for both configurations. At high $p_{\mathrm{T}}^{\mathrm{ref}}$, a resolution in the range $4\% - 5\%$ is obtained for the different configurations. The precision of this measurement is limited by the systematic uncertainties due to the MC modelling and the non-closure, reaching a total experimental uncertainty of approximately $25\% - 45\%$ below 45 GeV, $15\% - 20\%$ from 45 GeV to 600 GeV and approximately 35% at 1.2 TeV. This calibration was derived for the $2015 - 2016$, 2017 and 2018 data-taking periods separately, reaching a similar level of precision independently of the measured dataset. Given the size of the experimental uncertainties, the statistical combination with the dijet in situ measurement does not improve the result of the measurement. Therefore, the γ+jet in situ JER measurement was not included in the official ATLAS JER calibration chain.

The cross section for isolated photon plus jet production in pp collisions at $\sqrt{s} = 13$ TeV was measured using data with an integrated luminosity of 138.4 fb^{-1} collected by the ATLAS detector at the LHC. In this measurement, photons with $E_{\mathrm{T}}^{\gamma} > 250$ GeV and $|\eta^{\gamma}| < 2.37$ were measured in association with one jet with $p_{\mathrm{T}}^{\mathrm{jet}} > 165$ GeV and $|y^{\mathrm{jet}}| < 4.4$. A separation of $\Delta R \geq 0.8$ was required for the leading jet and the photon. The differential cross section was measured as a function of E_{T}^{γ} in three different jet rapidity intervals: $|y^{\mathrm{jet}}| < 1.2$, $1.2 \leq |y^{\mathrm{jet}}| < 2.8$, and $2.8 \leq |y^{\mathrm{jet}}| < 4.4$, for the same-sign and opposite-sign configurations in each case.

This measurement benefits from the large increase in the integrated luminosity with respect to the previous measurement of the angular correlations at $\sqrt{s} = 7$ TeV using $37\,\mathrm{pb}^{-1}$, reaching much higher E_{T}^{γ} values with smaller statistical uncertainties. The measurement spans the region $x \gtrsim 2 \times 10^{-3}$ and $6.25 \times 10^{4} \lesssim Q^2 \lesssim 4 \times 10^{6}$ GeV^2, therefore, it extends to a kinematic region previously unexplored with this final state. The subdivision of the phase space of the measurement into six different angular configurations improves the sensitivity of the measurement to the proton PDF and allows to identify regions in which the data have the potential to constrain further the PDF.

Next-to-leading-order perturbative QCD predictions from SHERPA 2.2.2 using various PDF sets were compared with the measurement and provide an adequate description of the data within experimental and theoretical uncertainties. In the opposite-sign configuration of region $2.8 \leq |y^{\mathrm{jet}}| < 4.4$, the theory prediction tends to overestimate the measured cross section; this is the region where the uncertainty due to the scale variations is largest. The comparison of the data and theory is limited by the theoretical uncertainty due to the terms beyond NLO. The experimental systematic uncertainties are significantly smaller than those of theoretical nature over the full measured phase

space. Hence, NNLO calculations are crucial to exploit this high-precision measurement and to constrain further the proton PDF.

Three measurements of the differential cross sections for isolated photons produced inclusively and in association with one jet have been presented. The first of these measurements uses $36.1\,\text{fb}^{-1}$ of the data collected by ATLAS during $2015-2016$, whereas the other two exploit the full integrated luminosity of the Run II dataset. Theoretical predictions at NLO and NNLO in pQCD are compared to the measured cross sections, providing precise tests of the theory at $\mathcal{O}\left(\alpha_{\text{EM}}\alpha_{\text{S}}^{3}\right)$ in the range of photon transverse energies from 125 GeV up to and beyond 1 TeV. Therefore, these high-precision measurements have the potential to further constrain the PDF, particularly the gluon density in the proton, within a global NNLO QCD fit.

Resumen y conclusiones

En esta disertación, se estudió la producción de fotones prompt aislados, inclusiva y en asociación con un jet, en colisiones protón-protón a $\sqrt{s} = 13$ TeV con el detector ATLAS en el LHC. Se proporcionó una vista general detallada sobre el acelerador LHC y el detector ATLAS, seguida de una descripción de las principales características del marco teórico subyacente a este trabajo. Las descripciones de la teoría de producción de fotones prompt, algoritmos de los jets y generadores de Monte Carlo se incluyeron en los capítulos iniciales de la tesis. Además, se proporcionó una descripción de la reconstrucción e identificación de fotones y de la reconstrucción de los jets, así como una descripción de la calibración de la escala de su energía y resolución.

Aunque las diferentes muestras de datos usadas en las mediciones presentadas en esta disertación están enriquecidas en fotones prompt aislados, todavía contienen cierta contaminación incluso después de la aplicación de los requerimientos de identificación y aislamiento en los fotones candidatos. Esta contaminación se origina predominantemente en procesos multijet en los que un jet es mal identificado como un fotón; este jet normalmente contiene un mesón neutro ligero, principalmente un π^0, que se descompone en un par de fotones colimado que lleva la mayoría de la energía del jet. Para la eliminación de dicha contaminación, se empleó un método basado en bandas laterales en el plano bidimensional abarcado por la energía transversal de aislamiento y la variable discreta que encapsula las propiedades de la forma de la ducha electromagnética del fotón. Este método se complementó con las estimaciones de la contribución de fotones prompt aislados (fracciones de fuga de señal) a las regiones enriquecidas en contaminación para mejorar la determinación de la señal. Como resultado, la pureza de las diferentes muestras resultó ser $\geq 90\%$, excepto para los *bins* de más bajo momento transversal del fotón en la medición de la calibración in situ, donde la pureza es superior al 70%.

Se generaron muestras de Monte Carlo de sucesos con un fotón prompt más jets usando los programas PYTHIA, SHERPA 2.1.1 y SHERPA 2.2.2. Estas muestras se usaron para varios propósitos: la determinación de las fracciones de fuga de la señal usadas en la sustracción de la contaminación; la evaluación de las correcciones de aceptación para obtener la medición de las secciones eficaces diferenciales así como de sus incertidumbres asociadas; la evaluación de los factores de calibración in situ y sus incertidumbres sistemáticas. Las comparaciones entre las distribuciones de los datos y MC de las variables cinemáticas muestran, en general, un acuerdo razonable.

La sección eficaz de producción inclusiva de fotones prompt aislados en colisiones pp a $\sqrt{s} = 13$ *TeV se midió usando datos con una luminosidad integrada de 36.1* fb^{-1} *recogida por el detector ATLAS en el LHC.* La sección eficaz diferencial se midió en función de E_{T}^{γ} en cuatro regiones de $|\eta^{\gamma}|$ para fotones con $E_{\mathrm{T}}^{\gamma} > 125$ GeV y $|\eta^{\gamma}| < 2.37$, excluyendo la región $1.37 < |\eta^{\gamma}| < 1.56$.

Esta medición constituye una mejora en términos de las incertidumbres estadística y sistemática con respecto a la medición previamente publicada gracias al aumento de diez veces en la luminosidad integrada. En esta medición, el alcance en E_{T}^{γ} se ha extendido hacia arriba por 1000, 500, 400 y 200 GeV para $|\eta^{\gamma}| < 0.6$, $0.6 < |\eta^{\gamma}| < 1.37$, $1.56 < |\eta^{\gamma}| < 1.81$ y $1.81 < |\eta^{\gamma}| < 2.37$, respectivamente, en comparación con la anterior medición a $\sqrt{s} = 13$ TeV. El rango en el que la medición está limitada por las incertidumbres sistemáticas es extendido hacia arriba hasta 1 TeV para $|\eta^{\gamma}| < 1.37$, donde previamente la limitación estaba en 600 GeV.

Las predicciones de QCD al orden siguiente al principal usando varios conjuntos de PDF se compararon con la medición y proporcionan una descripción adecuada de los datos dentro de las incertidumbres experimentales y teóricas. La comparación de los datos y la teoría está limitada por la incertidumbre teóricas debida a los términos más allá de NLO en QCD. Las incertidumbres experimentales sistemáticas son menores que las incertidumbres teóricas en todo el espacio de fases investigado. La inclusión de elementos de matriz de órdenes superiores y de duchas de partones, como proporciona el programa SHERPA 2.2.2, mejora la descripción de la normalización de los datos por la predicción.

Recientemente, una predicción al orden siguiente al siguiente al principal en QCD fue calculada en la que las incertidumbres debidas a las variaciones de las escalas son reducidas por un factor en el rango $2 - 20$ con respecto a las de un cálculo de QCD a NLO. En general, la predicción a NNLO en QCD proporciona una descripción excelente de los datos. La comparación de la predicción a NNLO con la sección eficaz medida representa un test preciso de la teoría en el rango de las energías transversales del fotón de 125 GeV hasta 1 TeV y más allá. La medición tiene el potencial de restringir aún más las PDF del protón, especialmente la densidad del gluón, dentro de un ajuste global a NNLO en QCD.

La sección eficaz de producción inclusiva de fotones aislados en colisiones pp a una energía del centro de masas de 13 *TeV se midió usando el detector ATLAS en el LHC con una luminosidad integrada de 139* fb^{-1}. La sección eficaz diferencial se midió en función de E_{T}^{γ} en seis regiones de $|\eta^{\gamma}|$ para fotones con $E_{\mathrm{T}}^{\gamma} > 250$ GeV y $|\eta^{\gamma}| < 2.37$, excluyendo la región $1.37 < |\eta^{\gamma}| < 1.56$. El aislamiento del fotón se asegura al requerir que la energía transversal en un cono de $R = 0.2$ o $R = 0.4$ alrededor del fotón sea menor que $4.2 \times 10^{-3} \times E_{\mathrm{T}}^{\gamma} + 4.8$ GeV.

Esta medición representa una mejora con respecto a las publicadas anteriormente a $\sqrt{s} = 13$ TeV debido a las nuevas regiones de $|\eta^{\gamma}|$ medidas: $|\eta^{\gamma}| < 0.6$, $0.6 < |\eta^{\gamma}| < 0.8$, $0.8 < |\eta^{\gamma}| < 1.37$, $1.56 < |\eta^{\gamma}| < 1.81$, $1.81 < |\eta^{\gamma}| < 2.01$, y $2.01 < |\eta^{\gamma}| < 2.37$, en contraposición a las cuatro regiones de los análisis anteriores. De esta manera, se

proporciona una información experimental más detallada para los ajustes de las PDF. Además, ahora se proporcionan mediciones para dos radios del cono de aislamiento del fotón, a saber, $R = 0.2$ y $R = 0.4$, permitiendo testear la dependencia de las predicciones de pQCD con R. Asimismo, la nueva medición se beneficia de incertidumbres sistemáticas reducidas gracias a las mejoras de los grupos de rendimiento combinado que desarrollan las recomendaciones para evaluar las incertidumbres sistemáticas. También, gracias al incremento de cuatro veces en la luminosidad integrada, se alcanzan valores superiores de E_{T}^{γ} con incertidumbres estadísticas menores, y el rango en el que la medición está limitada por las incertidumbres sistemáticas se extiende hacia arriba hasta ≈ 1.5 TeV para $|\eta^{\gamma}| < 0.6$ y $0.8 < |\eta^{\gamma}| < 1.37$, aproximadamente 1.1 TeV para $0.6 < |\eta^{\gamma}| < 0.8$ y $1.56 < |\eta^{\gamma}| < 1.81$, y ≈ 900 GeV para $1.81 < |\eta^{\gamma}| < 2.37$.

Las predicciones teóricas a NLO en QCD se calcularon usando varios conjuntos de PDF y proporcionan una descripción razonable de los datos dentro de las incertidumbres experimentales y teóricas. Como se concluyó para la anterior medición de fotones inclusivos a $\sqrt{s} = 13$ TeV, la comparación de la medición con las predicciones a NLO en pQCD está limitada por la incertidumbre teórica debida a los términos más allá de NLO. Se espera que la comparación de las mediciones con las predicciones a NNLO en QCD, como se proporciona para la medición de fotones inclusivos anterior, proporcione una descripción mejorada de los datos, con incertidumbres teóricas significativamente reducidas. Por lo tanto, esta medición de alta precisión tiene el potencial de restringir aún más las PDF del protón dentro de un ajuste global a NNLO en QCD.

La dependencia con R de la sección eficaz medida para la producción inclusiva de fotones aislados es descrita adecuadamente por las predicciones de SHERPA 2.2.2 y JETPHOX con $\mu_{\mathrm{R}} = \mu_{\mathrm{F}} = \mu_{\mathrm{f}} = E_{\mathrm{T}}^{\gamma}$, mientras que la predicción de JETPHOX con $\mu_{\mathrm{R}} = \mu_{\mathrm{F}} = \mu_{\mathrm{f}} = E_{\mathrm{T}}^{\gamma}/2$ proporciona una descripción mejorada de los datos. No se observa dependencia con las PDF del protón en las predicciones de la sección eficaz fiducial en función de R.

La escala de la energía y la resolución de los jets in situ se midieron en colisiones pp a $\sqrt{s} = 13$ TeV usando el detector ATLAS en el LHC con una luminosidad integrada de 80 fb^{-1} y 58.5 fb^{-1}, respectivamente. La JES (JER) se midió en función del momento transversal del fotón para fotones con $p_{\mathrm{T}}^{\mathrm{ref}} > 25$ GeV y $|\eta^{\gamma}| < 1.37$, y el jet principal fue seleccionado con $p_{\mathrm{T}}^{\mathrm{jet}} > 10$ GeV y $|\eta| < 0.8$ ($|\eta| < 0.7$). La medición de la JER se realizó en dos regiones de la pseudorapidez del jet principal, a saber, $|\eta| < 0.2$ y $0.2 < |\eta| < 0.7$. Ambas mediciones fueron realizadas para dos tipos de jets, PFlow y EMtopo.

La medición de la JES constituye una mejora en términos de las incertidumbres estadística y sistemática con respecto a la publicada anteriormente gracias al incremento de 25 veces en la luminosidad integrada. Además, esta medición se beneficia del uso del método MPF y de los jets PFlow como configuración de base, que mejora ligeramente los resultados obtenidos usando jets EMtopo. La corrección in situ obtenida con la técnica MPF para ambos tipos de jets es aproximadamente 3% a bajo $p_{\mathrm{T}}^{\mathrm{ref}}$ y decrece hasta alrededor de 2% para fotones por encima de ≈ 70 GeV; la incertidumbre en la escala de

la energía de los jets es $\approx 1\%$ para la mayor parte del rango de $p_{\mathrm{T}}^{\mathrm{ref}}$ medido. Se alcanzan valores superiores de $p_{\mathrm{T}}^{\mathrm{jet}}$ con incertidumbres estadísticas menores, extendiendo el límite de la calibración hasta 1.2 TeV, donde la limitación anterior era 950 GeV. Se realizó una medición complementaria de la escala de la energía de los jets por medio de la técnica DB, obteniendo factores de calibración e incertidumbres asociadas consistentes.

La medición de la JER se realizó usando el método DB con los tipos de jets PFlow y EMtopo en dos regiones de pseudorapidez del jet principal. La JER medida es razonablemente descrita por la simulación de MC dentro de las incertidumbres experimentales. Se obtiene una resolución en el rango $19\% - 24\%$ para jets EMtopo a bajo $p_{\mathrm{T}}^{\mathrm{ref}}$ en las dos regiones diferentes de pseudorapidez, mientras que la resolución medida con jets PFlow es inferior a 17% para ambas configuraciones. A alto $p_{\mathrm{T}}^{\mathrm{ref}}$, se obtiene una resolución en el rango $4\% - 5\%$ para las diferentes configuraciones. La precisión de esta medición está limitada por las incertidumbres sistemáticas debido al modelado del MC y al *non-closure*, alcanzando una incertidumbre experimental total de aproximadamente $25\% - 45\%$ por debajo de 45 GeV, $15\% - 20\%$ de 45 GeV a 600 GeV y aproximadamente 35% a 1.2 TeV. Esta calibración fue obtenida para los periodos de toma de datos de $2015 - 2016$, 2017 y 2018 separadamente, alcanzando un nivel de precisión similar independientemente del periodo de datos medido. Dado el tamaño de las incertidumbres experimentales, la combinación estadística con la medición in situ de *dijets* no mejora el resultado de dicha medición. Por lo tanto, la medición de la JER in situ de γ+jet no fue incluida en la cadena de calibración oficial de la JER de ATLAS.

La sección eficaz de producción de un fotón aislado más un jet en colisiones pp a una energía del centro de masas de 13 TeV se midió usando datos con una luminosidad integrada de 138.4 fb^{-1} recogida por el detector ATLAS en el LHC. En este análisis, se midieron fotones con $E_{\mathrm{T}}^{\gamma} > 250$ GeV y $|\eta^{\gamma}| < 2.37$ en asociación con un jet con $p_{\mathrm{T}}^{\mathrm{jet}} > 165$ GeV y $|y^{\mathrm{jet}}| < 4.4$. Se requirió una separación de $\Delta R \geq 0.8$ entre el jet principal y el fotón. La sección eficaz diferencial fue medida en función de E_{T}^{γ} en tres intervalos de la rapidez del jet diferentes: $|y^{\mathrm{jet}}| < 1.2$, $1.2 \leq |y^{\mathrm{jet}}| < 2.8$, y $2.8 \leq |y^{\mathrm{jet}}| < 4.4$, para las configuraciones de mismo signo y signo opuesto en cada caso.

Esta medición se beneficia del gran incremento en la luminosidad integrada con respecto a la anterior medición de las correlaciones angulares a $\sqrt{s} = 7$ TeV usando $37\,\mathrm{pb}^{-1}$, alcanzando valores de E_{T}^{γ} muy superiores con incertidumbres estadísticas menores. La medición abarca la región $x \gtrsim 2 \times 10^{-3}$ y $6.25 \times 10^{4} \lesssim Q^2 \lesssim 4 \times 10^{6}$ GeV^2, por lo tanto, se extiende a una región cinemática previamente inexplorada para este estado final. La subdivisión del espacio de fases de la medición en seis configuraciones angulares diferentes mejora la sensibilidad de la medición a las PDF del protón y permite identificar regiones en las que los datos tienen el potencial de restringir aún más las PDF.

Las predicciones al orden siguiente al principal de QCD perturbativa de SHERPA 2.2.2 usando varios conjuntos de PDF se compararon con la medición y proporcionan una descripción adecuada de los datos dentro de las incertidumbres experimentales y teóricas. En la configuración de signo opuesto de la región $2.8 \leq |y^{\mathrm{jet}}| < 4.4$, la predicción teórica

tiende a sobreestimar la sección eficaz medida; esta es la región donde la incertidumbre debida a las variaciones de las escalas es mayor. La comparación de los datos y la teoría está limitada por la incertidumbre teórica debida a los términos más allá de NLO. Las incertidumbres experimentales sistemáticas son significativamente menores que las de naturaleza teórica en todo el espacio de fases medido. Por lo tanto, son cruciales cálculos a NNLO para aprovechar esta medición de alta precisión y para restringir aún más las PDF del protón.

Se han presentado tres mediciones de las secciones eficaces diferenciales para fotones aislados producidos inclusivamente y en asociación con un jet. La primera de estas mediciones usa $36.1\,\text{fb}^{-1}$ de los datos recogidos por ATLAS durante $2015-2016$, mientras que las otras dos aprovechan la luminosidad integrada total del conjunto de datos del Run II. Las predicciones teóricas a NLO y NNLO en pQCD se comparan con las secciones eficaces medidas, proporcionando tests precisos de la teoría al orden $\mathcal{O}\left(\alpha_{\text{EM}}\alpha_{\text{S}}^{3}\right)$ en el rango de las energías transversales del fotón desde 125 GeV hasta 1 TeV y más allá. Por lo tanto, estas mediciones de alta precisión tienen el potencial de restringir aún más las PDF, particularmente la densidad del gluón en el protón, dentro de un ajuste global a NNLO en QCD.

Bibliography

[1] O. S. Brüning et al., *LHC Design Report: Vol.1. the LHC Main Ring*, CERN-2004-003-V-1, cds.cern.ch/record/782076. *LHC Design Report: Vol.2. the LHC Infrastructure and General Services*, CERN-2004-003-V-2, cds.cern.ch/record/815187. *LHC Design Report: Vol.3. the LHC Injector Chain*, CERN-2004-003-V-3, cds.cern.ch/record/823808.

[2] ATLAS collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1, doi:10.1016/j.physletb.2012.08.020.

[3] CMS collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716** (2012) 30, doi:10.1016/j.physletb.2012.08.021.

[4] ATLAS collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003, doi:10.1088/1748-0221/3/08/S08003.

[5] CMS collaboration, *The CMS experiment at the CERN LHC*, JINST **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.

[6] T. Pietrycki and A. Szczurek, *Photon-jet correlations in pp and $p\bar{p}$ collisions*, Phys. Rev. D **76** (2007) 034003. doi:10.1103/PhysRevD.76.034003.

[7] Z. Belghobsi et al., *Photon-jet correlations and constraints on fragmentation functions*, Phys. Rev. D **79** (2009) 114024. doi:10.1103/PhysRevD.79.114024.

[8] ATLAS collaboration, *Measurement of the inclusive isolated prompt photons cross section in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector using* $4.6\,\text{fb}^{-1}$, Phys. Rev. D **89** (2014) 052004. doi:10.1103/PhysRevD.89.052004.

[9] ATLAS collaboration, *Measurement of the inclusive isolated prompt photon cross section in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, JHEP **08** (2016) 005. doi:10.1007/JHEP08(2016)005.

[10] ATLAS collaboration, *Measurement of the cross section for inclusive isolated-photon production in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector*, Phys. Lett. B **770** (2017) 473. doi:10.1016/j.physletb.2017.04.072.

[11] CMS collaboration, *Measurement of the differential cross section for isolated prompt photon production in pp collisions at 7 TeV*, Phys. Rev. D **84** (2011) 052011. doi:10.1103/PhysRevD.84.052011.

[12] CMS collaboration, *Measurement of differential cross sections for inclusive isolated-photon and photon+jets production in proton-proton collisions at* $\sqrt{s} = 13$ *TeV*, Eur. Phys. J. C **79** (2019) 20. doi:10.1140/epjc/s10052-018-6482-9.

[13] D. d'Enterria and J. Rojo, *Quantitative constraints on the gluon distribution function in the proton from collider isolated-photon data*, Nucl. Phys. B **860** (2012) 311. doi:10.1016/j.nuclphysb.2012.03.003.

[14] L. Carminati et al., *Sensitivity of the LHC isolated-*γ*+jet data to the parton distribution functions of the proton*, EPL **101** (2013) 61002. doi:10.1209/0295-5075/101/61002.

[15] J. Gao, L. Harland-Lang and J. Rojo, *The structure of the proton in the LHC precision era*, Phys. Rept. **742** (2018) 1. doi:10.1016/j.physrep.2018.03.002.

[16] J. M. Campbell et al., *Direct photon production and PDF fits reloaded*, Eur. Phys. J. C **78** (2018) 470. doi:/10.1140/epjc/s10052-018-5944-4.

[17] A. V. Lipatov and M. A. Malyshev, *Reconsideration of the inclusive prompt photon production at the LHC with* k_{t}*-factorization*, Phys. Rev. D **94** (2016) 034020. doi:10.1103/PhysRevD.94.034020.

[18] M. Klasen, C. Klein-Bösing and H. Poppenborg, *Prompt photon production and photon-jet correlations at the LHC*, JHEP **03** (2018) 081. doi:10.1007/JHEP03(2018)081.

[19] M. D. Schwartz, *Precision direct photon spectra at high energy and comparison to the 8 TeV ATLAS data*, JHEP **09** (2016) 005. doi:10.1007/JHEP09(2016)005.

[20] J. M. Campbell et al., *Direct photon production at next-to-next-to-leading order*, Phys. Rev. Lett. **118** (2017) 222001. doi:10.1103/PhysRevLett.118.222001.

[21] X. Chen et al., *Isolated photon and photon+jet production at NNLO QCD accuracy*, JHEP **04** (2020) 166. doi:10.1007/JHEP04(2020)166.

[22] X. Chen et al., *Isolated photon and photon+jet production at NNLO QCD accuracy and the ratio* $R^{\gamma}_{13/8}$, [arXiv:1905.08577].

[23] G. Brooijmans et al., *Les Houches 2017: physics at TeV colliders new physics working group report*, [arXiv:1803.10379].

[24] ATLAS collaboration, *Determination of jet calibration and energy resolution in proton-proton collisions at* $\sqrt{s} = 8$ *TeV using the ATLAS detector*, Eur. Phys. J. C **80** (2020) 1104. doi:10.1140/epjc/s10052-020-08477-8.

[25] ATLAS collaboration, *Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Rev. D **96** (2017) 072002. doi:10.1103/PhysRevD.96.072002.

[26] ATLAS collaboration, *Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Eur. Phys. J. C **81** (2021) 689. doi:10.1140/epjc/s10052-021-09402-3.

[27] ATLAS collaboration, *Search for new phenomena with photon+jet events in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **03** (2016) 041. doi:10.1007/JHEP03(2016)041.

[28] ATLAS collaboration, *Search for new phenomena in high-mass final states with a photon and a jet from pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Eur. Phys. J. C **78** (2018) 102. doi:10.1140/epjc/s10052-018-5553-2.

[29] ATLAS collaboration, *Measurement of the production cross section of an isolated photon associated with jets in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Phys. Rev. D **85** (2012) 092014. doi:10.1103/PhysRevD.85.092014.

[30] ATLAS collaboration, *Dynamics of isolated-photon plus jet production in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Nucl. Phys. B **875** (2013) 483. doi:10.1016/j.nuclphysb.2013.07.025.

[31] ATLAS collaboration, *High-E_{T} isolated-photon plus jets production in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, Nucl. Phys. B **918** (2017) 257. doi:10.1016/j.nuclphysb.2017.03.006.

[32] ATLAS collaboration, *Measurement of the cross section for isolated-photon plus jet production in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector*, Phys. Lett. B **780** (2018) 578. doi:10.1016/j.physletb.2018.03.035.

[33] ATLAS collaboration, *Measurement of isolated-photon plus two-jet production in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **03** (2020) 179. doi:10.1007/JHEP03(2020)179.

[34] CMS collaboration, *Rapidity distributions in exclusive Z+jet and γ+jet events in pp collisions at $\sqrt{s} = 7$ TeV*, Phys. Rev. D **88** (2013) 112009. doi:10.1103/PhysRevD.88.112009.

[35] ATLAS collaboration, *Measurement of the triple-differential cross section for photon+jets production in proton–proton collisions at $\sqrt{s} = 7$ TeV*, JHEP **06** (2014) 009. doi:10.1007/JHEP06(2014)009.

[36] ATLAS collaboration, *Comparison of the Z/γ^*+jets to γ+jets cross sections in pp collisions at $\sqrt{s} = 8$ TeV*, JHEP **10** (2015) 128. doi:10.1007/JHEP10(2015)128.

[37] ATLAS collaboration, *Measurement of the inclusive isolated-photon cross section in pp collisions at $\sqrt{s} = 13$ TeV using $36\,\text{fb}^{-1}$ of ATLAS data*, JHEP **10** (2019) 203. doi:10.1007/JHEP10(2019)203.

[38] ATLAS collaboration, *Performance of the ATLAS trigger system in 2015*, Eur. Phys. J. C **77** (2017) 317, doi:10.1140/epjc/s10052-017-4852-3.

[39] ATLAS collaboration, *Performance of electron and photon triggers in ATLAS during LHC Run 2*, Eur. Phys. J. C **80** (2020) 47, doi:10.1140/epjc/s10052-019-7500-2.

[40] CERN, *LEP Design Report: Vol.1. The LEP Injector Chain*, CERN-LEP-TH-83-29, CERN-PS-DL-83-31, CERN-SPS-83-26, LAL-RT-83-09, cds.cern.ch/record/98881. *LEP Design Report: Vol.2. The LEP Main Ring*, CERN-LEP-84-01, cds.cern.ch/record/102083.

[41] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001, doi:10.1088/1748-0221/3/08/S08001.

[42] M. Brice, *3d view photo of the LHC Machine*, CERN-AC-0911188, cds.cern.ch/record/1223589.

[43] ALICE collaboration, *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002, doi:10.1088/1748-0221/3/08/S08002.

[44] LHCb collaboration, *The LHCb Detector at the LHC*, JINST **3** (2008) S08005, doi:10.1088/1748-0221/3/08/S08005.

[45] E. Mobs, *The CERN accelerator complex - 2019*, CERN-GRAPHICS-2019-002, cds.cern.ch/record/2684277.

[46] TOTEM collaboration, *The TOTEM Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08007, doi:10.1088/1748-0221/3/08/S08007.

[47] ATLAS collaboration, *Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in the High Pile-Up LHC Environment*, ATLAS-CONF-2012-042, [INSPIRE].

[48] ATLAS collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC*, Eur. Phys. J. C **76** (2016) 653, doi:10.1140/epjc/s10052-016-4466-1.

[49] ATLAS collaboration, *Measurement of the Inelastic Proton-Proton Cross Section at $\sqrt{s} = 13$ TeV with the ATLAS Detector at the LHC*, Phys. Rev. Lett. **117** (2016) 182002, doi:10.1103/PhysRevLett.117.182002.

[50] ATLAS collaboration, *Luminosity public results for Run I*, LuminosityPublicResultsRun1.

[51] ATLAS collaboration, *Luminosity public results for Run II*, LuminosityPublicResultsRun2.

[52] CERN, *CERN releases analysis of LHC incident*, Media and Press Relations (Press release) (2018), cern-releases-analysis-lhc-incident.

[53] J. Wenninger, *The LHC collider: Highlights of the LHC run 1*, C. R. Physique **16** (2015) 347. doi:10.1016/j.crhy.2015.03.005.

[54] R. Steerenberg and J. Wenninger, *Operational challenges and performance of the LHC during Run II*, JACoW (2018), doi:10.18429/JACoW-HB2018-THA2WD01.

[55] ATLAS collaboration, *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics*, [arXiv:0901.0512].

[56] ATLAS collaboration, *Atlas technical paper list of figures*, AtlasTechnicalPaperListOfFigures.

[57] ATLAS collaboration, *ATLAS Insertable B-Layer Technical Design Report*, CERN-LHCC-2010-013, ATLAS-TDR-19, cds.cern.ch/record/1291633, [INSPIRE].

[58] ATLAS collaboration, *Alignment of the ATLAS Inner Detector in Run-2*, Eur. Phys. J. C **80** (2020) 1194. doi:10.1140/epjc/s10052-020-08700-6.

[59] ATLAS collaboration, *ATLAS forward detectors for luminosity measurement and monitoring*, CERN-LHCC-2004-010, cdsweb.cern.ch/record/721908. *Zero Degree Calorimeters for ATLAS*, CERN-LHCC-2007-001, cdsweb.cern.ch/record/1009649. *ATLAS forward detectors for measurement of elastic scattering and luminosity*, CERN-LHCC-2008-004, cdsweb.cern.ch/record/1095847.

[60] ATLAS collaboration, *Technical Design Report for the Phase-I Upgrade of the ATLAS TDAQ System*, CERN-LHCC-2013-018, ATLAS-TDR-023, cds.cern.ch/record/1602235.

[61] R. K. Ellis, W. J. Stirling and B. R. Webber, *QCD and Collider Physics*, Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol. **8** (1996) 1, doi:10.1017/CBO9780511628788.

[62] S. Catani et al., *Longitudinally-invariant $k_\perp$-clustering algorithms for hadron-hadron collisions*, Nucl. Phys. B **406** (1993) 187. doi:10.1016/0550-3213(93)90166-M.

[63] M. Cacciari et al., *The anti-k_t jet clustering algorithm*, JHEP **04** (2008) 063. doi:10.1088/1126-6708/2008/04/063.

[64] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508, doi:10.1103/PhysRevLett.13.508.

[65] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys. Lett. **12** (1964) 132, doi:10.1016/0031-9163(64)91136-9.

[66] P. W. Higgs, *Spontaneous Symmetry Breakdown without Massless Bosons*, Phys. Rev. **145** (1966) 1156, doi:10.1103/PhysRev.145.1156.

[67] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321, doi:10.1103/PhysRevLett.13.321.

[68] CERN, *Go on a particle quest at the first CERN webfest*, BUL-NA-2012-269, SMinfographic.

[69] S. L. Glashow, *Partial-symmetries of weak interactions*, Nucl. Phys. **22** (1961) 579. doi:10.1016/0029-5582(61)90469-2.

[70] J. Goldstone et al., *Broken Symmetries*, Phys. Rev. **127** (1962) 965. doi:10.1103/PhysRev.127.965.

[71] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264. doi:10.1103/PhysRevLett.19.1264.

[72] P. A. Zyla et al., *Particle Data Group*, Prog. Theor. Exp. Phys. **8** (2020) 083C01, doi:10.1093/ptep/ptaa104.

[73] NNPDF collaboration, *Parton distributions for the LHC run II*, JHEP **04** (2015) 40, doi:10.1007/JHEP04(2015)040.

[74] L. A. Harland-Lang et al., *Parton distributions in the LHC era: MMHT 2014 PDFs*, Eur. Phys. J. C **75** (2015) 204, doi:10.1140/epjc/s10052-015-3397-6.

[75] H1 and ZEUS collaborations, *Combination of measurements of inclusive deep inelastic $e^{\pm}p$ scattering cross sections and QCD analysis of HERA data*, Eur. Phys. J. C **75** (2015) 580, doi:10.1140/epjc/s10052-015-3710-4.

[76] V. N. Gribov and L. N. Lipatov, *Deep inelastic ep scattering in perturbation theory*, Sov. J. Nucl. Phys. **15** (1972) 438 [Yad. Fiz. **15** (1972) 781].

[77] L. N. Lipatov, *The parton model and perturbation theory*, Sov. J. Nucl. Phys. **20** (1975) 95 [Yad. Fiz. **20** (1974) 181].

[78] G. Altarelli and G. Parisi, *Asymptotic freedom in parton language*, Nucl. Phys. B **126** (1977) 298. doi:10.1016/0550-3213(77)90384-4.

[79] Y. L. Dokshitze, *Calculation of structure functions of deep-inelastic scattering and e^+e^- annihilation by perturbation theory in quantum chromodynamics*, Sov. Phys. JETP **46** (1977) 641 [Zh. Eksp. Teor. Fiz. **73** (1977) 1216].

[80] G.'t Hooft and M. Veltman, *Regularization and renormalization of gauge fields*, Nucl. Phys. B **44** (1972) 189. doi:10.1016/0550-3213(72)90279-9.

[81] W. A. Bardeen et al., *Deep-inelastic scattering beyond the leading order in asymptotically free gauge theories*, Phys. Rev. D **18** (1978) 3998. doi:10.1103/PhysRevD.18.3998.

[82] R. K. Ellis et al., *Factorization and the Parton Model in QCD*, Phys. Lett. B **78** (1978) 281. doi:10.1016/0370-2693(78)90023-0.

[83] J. C. Collins and D. E. Soper, *The Theorems of Perturbative QCD*, Ann. Rev. Nucl. Part. Sci **37** (1987) 383, doi:10.1146/annurev.ns.37.120187.002123.

[84] R. P. Feynman, *Photon Hadron Interactions*, W. A. Benjamin, New York (1972).

[85] G. P. Salam and G. Soyez, *A practical seedless infrared-safe cone jet algorithm*, JHEP **05** (2007) 086. doi:10.1088/1126-6708/2007/05/086.

[86] Y. L. Dokshitzer et al., *Better jet clustering algorithms*, JHEP **08** (1997) 001. doi:10.1088/1126-6708/1997/08/001.
M. Wobisch and T. Wengler, *Hadronization corrections to jet cross sections in deep-inelastic scattering*, [arXiv:hep-ph/9907280].

[87] M. Cacciari et al., *The catchment area of jets*, JHEP **04** (2008) 005. doi:10.1088/1126-6708/2008/04/005.

[88] S. Catani et al., *Cross section of isolated prompt photons in hadron-hadron collisions*, JHEP **05** (2002) 028. doi:10.1088/1126-6708/2002/05/028.

[89] S. Frixione, *Isolated photons in perturbative QCD*, Phys. Lett. B **429** (1998) 369. doi:10.1016/S0370-2693(98)00454-7.

[90] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007. doi:10.1088/1126-6708/2009/02/007.

[91] E. Bothmann et al., *Event generation with Sherpa 2.2*, SciPost Phys. **7** (2019) 034. doi:10.21468/SciPostPhys.7.3.034.

[92] P. Aurenche et al., *A new critical study of photon production in hadronic collisions*, Phys. Rev. D **73** (2006) 094007. doi:10.1103/PhysRevD.73.094007.

[93] L. Bourhis et al., *Quarks and gluon fragmentation functions into photons*, Eur. Phys. J. C **2** (1998) 529. doi:10.1007/s100520050158.

[94] S. Dulat et al., *New parton distribution functions from a global analysis of quantum chromodynamics*, Phys. Rev. D **93** (2016) 033006. doi:10.1103/PhysRevD.93.033006.

[95] S. Alekhin et al., *Parton distribution functions, α_S, and heavy-quark masses for LHC Run II*, Phys. Rev. D **96** (2017) 014011. doi:10.1103/PhysRevD.96.014011.

[96] T.-J. Hou et al., *New CTEQ global analysis of quantum chromodynamics with high-precision data from the LHC*, Phys. Rev. D **103** (2021) 014013. doi:10.1103/PhysRevD.103.014013.

[97] NNPDF collaboration, *Parton distributions from high-precision collider data*, Eur. Phys. J. C **77** (2017) 663. doi:10.1140/epjc/s10052-017-5199-5.

[98] S. Dittmaier and M. Huber, *Radiative corrections to the neutral-current Drell-Yan process in the Standard Model and its minimal supersymmetric extension*, JHEP **01** (2010) 060. doi:10.1007/JHEP01(2010)060.

[99] T. Becher et al., *Addendum: electroweak Sudakov effects in W, Z and γ production at large transverse momentum*, Phys. Rev. D **92** (2015) 073011. doi:10.1103/PhysRevD.92.073011.

[100] Z. Hall and J. Thaler, *Photon isolation and jet substructure*, JHEP **09** (2018) 164. doi:10.1007/JHEP09(2018)164.

[101] T. Sjöstrand et al., *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026. doi:10.1088/1126-6708/2006/05/026.

[102] T. Sjöstrand et al., *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852. doi:10.1016/j.cpc.2008.01.036.

[103] S. Agostinelli et al., *GEANT 4 — a simulation toolkit*, Nucl. Instrum. Methods Phys. A **506** (2003) 250. doi:10.1016/S0168-9002(03)01368-8.

[104] NNPDF collaboration, *Parton distributions with LHC data*, Nucl. Phys. B **867** (2013) 244. doi:10.1016/j.nuclphysb.2012.10.003.

[105] H.-L. Lai et al., *New parton distributions for collider physics*, Phys. Rev. D **82** (2010) 074024. doi:10.1103/PhysRevD.82.074024.

[106] M. L. Mangano et al., *Multijet matrix elements and shower evolution in hadronic collisions: $Wb\bar{b}$ + n-jets as a case study*, Nucl. Phys. B **632** (2002) 343. doi:10.1016/S0550-3213(02)00249-3.
M. L. Mangano et al., *Matching matrix elements and shower evolution for top-pair production in hadronic collisions*, JHEP **01** (2007) 013. doi:10.1088/1126-6708/2007/01/013.

[107] S. Catani et al., *QCD matrix elements + parton showers*, JHEP **11** (2002) 063. doi:10.1088/1126-6708/2001/11/063.

[108] F. Krauss, *Matrix elements and parton showers in hadronic interactions*, JHEP **08** (2002) 015. doi:10.1088/1126-6708/2002/08/015.

[109] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, JHEP **06** (2002) 029. doi:10.1088/1126-6708/2002/06/029.

[110] S. Höche et al., *QCD matrix elements + parton showers: The NLO case*, JHEP **04** (2013) 027. doi:10.1007/JHEP04(2013)027.

[111] CDF collaboration, *Measurement of Double Parton Scattering in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV*, Phys. Rev. Lett. **79** (1997) 584. doi:10.1103/PhysRevLett.79.584. *Double parton scattering in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV*, Phys. Rev. D **56** (1997) 3811. doi:10.1103/PhysRevD.56.3811. *Underlying event in hard interactions at the Fermilab Tevatron $p\bar{p}$ collider*, Phys. Rev. D **70** (2004) 072002. doi:10.1103/PhysRevD.70.072002.

[112] CDF collaboration, *Charged jet evolution and the underlying event in proton-antiproton collisions at 1.8 TeV*, Phys. Rev. D **65** (2002) 092002. doi:10.1103/PhysRevD.65.092002.

[113] T. Sjöstrand and M. van Zijl, *A multiple-interaction model for the event structure in hadron collissions*, Phys. Rev. D **36** (1987) 2019. doi:10.1103/PhysRevD.36.2019.

[114] B. Andersson, *Parton fragmentation and string dynamics*, Phys. Rept. **97** (1983) 31. doi:10.1016/0370-1573(83)90080-7.

[115] T. D. Gottschalk, *A realistic model for e^+e^- annihilation including parton bremsstrahlung effects*, Nucl. Phys. B **214** (1983) 201. doi:10.1016/0550-3213(83)90658-2.
T. D. Gottschalk, *An improved description of hadronization in the QCD cluster model for e^+e^- annihilation*, Nucl. Phys. B **239** (1984) 349. doi:10.1016/0550-3213(84)90253-0.
B. R. Webber, *A QCD model for jet fragmentation including soft gluon interference*, Nucl. Phys. B **238** (1984) 492. doi:10.1016/0550-3213(84)90333-X.
Gottschalk and D. A. Morris, *A new model for hadronization and e^+e^- annihilation*, Nucl. Phys. B **288** (1987) 729. doi:10.1016/0550-3213(87)90236-7.

[116] J.-C. Winter et al., *A modified cluster-hadronisation model*, Eur. Phys. J. C **36** (2004) 381. doi:10.1140/epjc/s2004-01960-8.

[117] Zachary Marshall and the ATLAS collaboration, *Simulation of Pile-up in the ATLAS Experiment*, J. Phys. Conf. Ser. **513** (2014) 022024. doi:10.1088/1742-6596/513/2/022024.

[118] Tadej Novak and the ATLAS collaboration, *New techniques for pile-up simulation in ATLAS*, EPJ Web Conf. **214** (2019) 02044. doi:10.1051/epjconf/201921402044.

[119] ATLAS collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*, ATL-PHYS-PUB-2014-021. cds.cern.ch/record/1966419.

[120] S. Höche et al., *QCD matrix elements and truncated showers*, JHEP **05** (2009) 053. doi:10.1088/1126-6708/2009/05/053.

[121] ATLAS collaboration, *The ATLAS simulation infrastructure*, Eur. Phys. J. C **70** (2010) 823. doi:10.1140/epjc/s10052-010-1429-9.

[122] F. Krauss et al., *AMEGIC++ 1.0: a matrix element generator in C++*, JHEP **02** (2002) 044. doi:10.1088/1126-6708/2002/02/044.

[123] F. Cascioli et al., *Scattering amplitudes with open loops*, Phys. Rev. Lett. **108** (2012) 111601. doi:10.1103/PhysRevLett.108.111601.

[124] S. Schumann and F. Krauss, *A parton shower algorithm based on Catani-Seymour dipole factorisation*, JHEP **03** (2008) 038. doi:10.1088/1126-6708/2008/03/038.

[125] F. Siegert, *A practical guide to event generation for prompt photon production with Sherpa*, J. Phys. G **44** (2017) 044007. doi:10.1088/1361-6471/aa5f29.

[126] T. Gleisberg and S. Höche, *Comix, a new matrix element generator*, JHEP **12** (2008) 039. doi:10.1088/1126-6708/2008/12/039.

[127] ATLAS collaboration, *Summary of ATLAS PYTHIA 8 tunes*, ATL-PHYS-PUB-2012-003. cds.cern.ch/record/14741em07.

[128] A. D. Martin et al., *Parton distributions for the LHC*, Eur. Phys. J. C **63** (2009) 189. doi:10.1140/epjc/s10052-009-1072-5.

[129] ATLAS collaboration, *Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run 2 data collected in 2015 and 2016*, Eur. Phys. J. C **79** (2019) 205. doi:10.1140/epjc/s10052-019-6650-6.

[130] ATLAS collaboration, *Electron and photon energy calibration with the ATLAS detector using 2015-2016 LHC proton-proton collision data*, JINST **14** (2019) P03017. doi:10.1088/1748-0221/14/03/P03017.

[131] ATLAS collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data*, JINST **14** (2019) P12006. doi:10.1088/1748-0221/14/12/P12006.

[132] ATLAS collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, Eur. Phys. J. C **77** (2017) 490, doi:10.1140/epjc/s10052-017-5004-5.

[133] ATLAS collaboration, *Calorimeter clustering algorithms: description and performance*, ATL-LARG-PUB-2008-002. cds.cern.ch/record/1099735.

[134] T. Cornelissen et al., *Design and Implementation of the ATLAS New Tracking (NEWT)*, ATL-SOFT-PUB-2007-007, cds.cern.ch/record/1020106.

[135] R. Frühwirth, *Application of Kalman filtering to track and vertex fitting*, Nucl. Instrum. Meth. A **262** (1987) 444. doi:10.1016/0168-9002(87)90887-4.

[136] T. G. Cornelissen et al., *The global χ^2 track fitter in ATLAS*, J. Phys. Conf. Ser. **119** (2008) 032013. doi:10.1088/1742-6596/119/3/032013.

[137] ATLAS collaboration, *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*, ATLAS-CONF-2012-047, cds.cern.ch/record/1449796.

[138] M. Cacciari and G. P. Salam, *Pileup subtraction using jet areas*, Phys. Lett. B **659** (2008) 119. doi:10.1016/j.physletb.2007.09.077.

[139] W. Buttinger, *Using Event Weights to account for differences in Instantaneous Luminosity and Trigger Prescale in Monte Carlo and Data*, ATL-COM-SOFT-2015-119, cds.cern.ch/record/2014726.

[140] M. Cacciari et al., *FastJet User Manual*, Eur. Phys. J. C **72** (2012) 1896. doi:10.1140/epjc/s10052-012-1896-2.

[141] M. Cacciari et al., *Dispelling the N^3 myth for the K_t jet-finder*, Phys. Lett. B **641** (2006) 57. doi:10.1016/j.physletb.2006.08.037.

[142] ATLAS collaboration, *Jet reconstruction and performance using particle flow with the ATLAS Detector*, Eur. Phys. J. C **77** (2017) 466. doi:10.1140/epjc/s10052-017-5031-2.

[143] ATLAS collaboration, *Jet energy measurement with the ATLAS detector in proton–proton collisions at $\sqrt{s}$ = 7 TeV*, Eur. Phys. J. C **73** (2013) 2304. doi:10.1140/epjc/s10052-013-2304-2.

[144] ATLAS collaboration, *Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*, Eur. Phys. J. C **76** (2016) 581. doi:10.1140/epjc/s10052-016-4395-z.

[145] ATLAS collaboration, *Selection of jets produced in 13 TeV proton–proton collisions with the ATLAS detector*, ATLAS-CONF-2015-029, cds.cern.ch/record/2037702.

[146] ATLAS collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}$ = 13 TeV*, Eur. Phys. J. C **76** (2016) 292. doi:10.1140/epjc/s10052-016-4120-y.

[147] ATLAS collaboration, *A measurement of the calorimeter response to single hadrons and determination of the jet energy scale uncertainty using LHC Run-1 pp-collision data with the ATLAS detector*, Eur. Phys. J. C **77** (2017) 26. doi:10.1140/epjc/s10052-016-4580-0.

[148] G. Folger and J. P. Wellisch, *String Parton Models in GEANT4*, [arXiv:nucl-th/0306007].

[149] ATLAS collaboration, *Jet energy measurement and its systematic uncertainty in proton–proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Eur. Phys. J. C **75** (2015) 17. doi:10.1140/epjc/s10052-014-3190-y.

[150] ATLAS collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*, ATLAS-CONF-2019-021, cdsweb.cern.ch/record/2677054.

[151] G. Avoni et al., *The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS*, JINST **13** (2018) P07017. doi:10.1088/1748-0221/13/07/P07017.

[152] ATLAS collaboration, *Electron and photon trigger efficiency plots using the full 2016 dataset*, ATL-COM-DAQ-2017-015, cds.cern.ch/record/2254973.

[153] G. D'Agostini, *A multidimensional unfolding method based on Bayes' theorem*, Nucl. Instrum. Meth. A **362** (1995) 487. doi:10.1016/0168-9002(95)00274-X.

[154] G. Bohm and G. Zech, *Introduction to statistics and data analysis for physicists*, Deutsches Elektronen-Synchrotron, Hamburg, Germany (2014). doi:10.3204/DESY-BOOK/statistics.

[155] ATLAS collaboration, *Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*, Eur. Phys. J. C **74** (2014) 3071. doi:10.1140/epjc/s10052-014-3071-4.

[156] J. Pumplin et al., *Uncertainties of predictions from parton distribution functions. II. The Hessian method*, Phys. Rev. D **65** (2001) 014013. doi:10.1103/PhysRevD.65.014013.

[157] ATLAS collaboration, *Development of ATLAS Primary Vertex Reconstruction for LHC Run 3*, [arXiv:1910.08405].

[158] ATLAS collaboration, *Performance of Photon triggers in ATLAS during LHC Run-2*, ATL-COM-DAQ-2018-184, cds.cern.ch/record/2652154.

[159] T. Adye, *Unfolding algorithms and tests using RooUnfold*, [arXiv:1105.1160].

[160] V. Lendermann et al., *Combining Triggers in HEP data analysis*, Nucl. Instrum. Methods Phys. Res. A **604** (2009) 707. doi:10.1016/j.nima.2009.03.173.

[161] D. E. Groom, *Energy flow in a hadronic cascade: Application to hadron calorimetry*, Nucl. Instrum. Meth. A **572** (2007) 633. doi:10.1016/j.nima.2006.11.070.

Appendix A

Converted photon fractions

The fraction of converted photons observed in data is shown in figure A.1 as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$ for the two $R = 0.2$ and $R = 0.4$ samples under study in the inclusive photon analysis using $139\,\mathrm{fb}^{-1}$. No dependence is observed in the fraction of converted photons with E_{T}^{γ} or the isolation cone radius. However, the fraction of converted photons has a large dependence with η^{γ}: for $|\eta^{\gamma}| < 0.8$, the fraction is $\approx 20\%$; in the $1.56 < |\eta^{\gamma}| < 1.81$ region, the fraction grows to $\approx 50\%$, whereas for $0.8 < |\eta^{\gamma}| < 1.37$ and $1.81 < |\eta^{\gamma}| < 2.37$, this fraction is $\approx 40\%$. Therefore, the region $1.56 < |\eta^{\gamma}| < 1.81$ is the one with the largest contribution from converted photons.

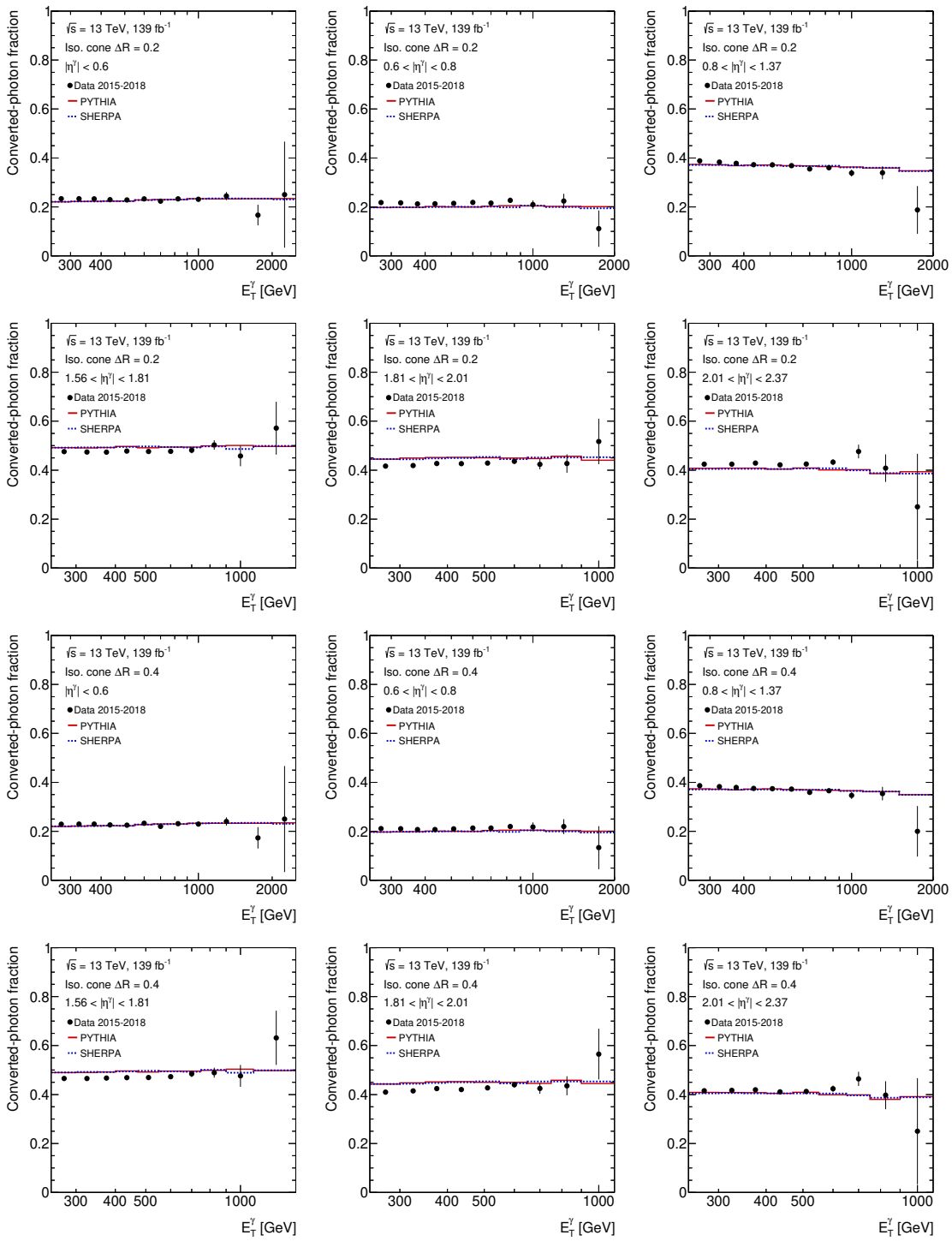


Figure A.1: Fraction of converted photons in data (dots) as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. For comparison, the MC simulations from PYTHIA (solid line) and SHERPA (dahsed line) are also shown. The error bars of the data represent the statistical uncertainty only. Results for the $R = 0.2$ ($R = 0.4$) sample are shown at the top (bottom).

Appendix B

Ambient transverse energy density

A correction was applied to $E_{\mathrm{T}}^{\mathrm{iso}}$ to account for the effects from UE modelling and pile-up effects and to match the definition between data and theory. The ambient transverse energy density (ρ) is computed in an event-by-event basis using low-energy jets; this density is then multiplied by the area of the photon isolation cone and subtracted from the isolation energy. For the inclusive photon analysis using $139\,\mathrm{fb}^{-1}$ the ambient transverse energy density in data is typically 7 GeV in the central rapidity region and 2.5 GeV in the forward rapidity region, as can be seen in figure B.1. Therefore, the correction applied to $E_{\mathrm{T}}^{\mathrm{iso}}$ is typically 3.5 GeV (1.25 GeV) in the central (forward) rapidity region for an isolation cone of size $\Delta R = 0.4$, and typically 0.88 GeV (0.31 GeV) in the central (forward) rapidity region when $\Delta R = 0.2$, since in this case the area of the cone is four times smaller.

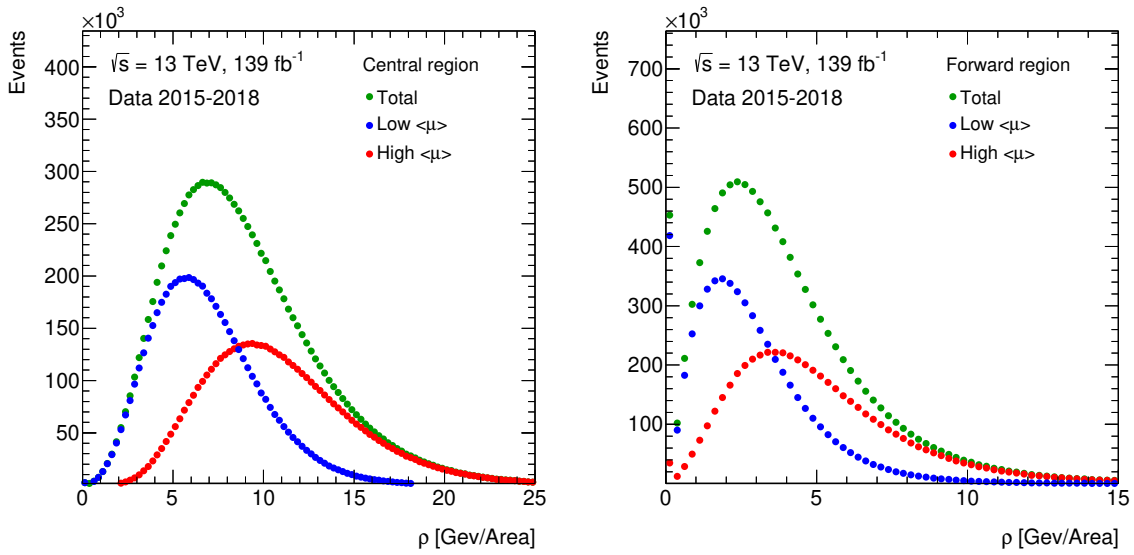


Figure B.1: Distribution of the ambient transverse energy density in the central (left) and forward (right) regions.

Appendix C

Data-driven corrections to $E_{\mathrm{T}}^{\mathrm{iso}}$

The isolation transverse energy distributions in different E_{T}^{γ} and $|\eta^{\gamma}|$ regions are shown in figures C.1 to C.6 for the inclusive-isolated photon analysis using 139 fb^{-1}. The data distributions are compared with PYTHIA simulation. The $E_{\mathrm{T}}^{\mathrm{iso}}$ distributions after the data-driven corrections to the MC simulation are shown in figures C.7 to C.12; an improved description of the data by the MC simulation is achieved.

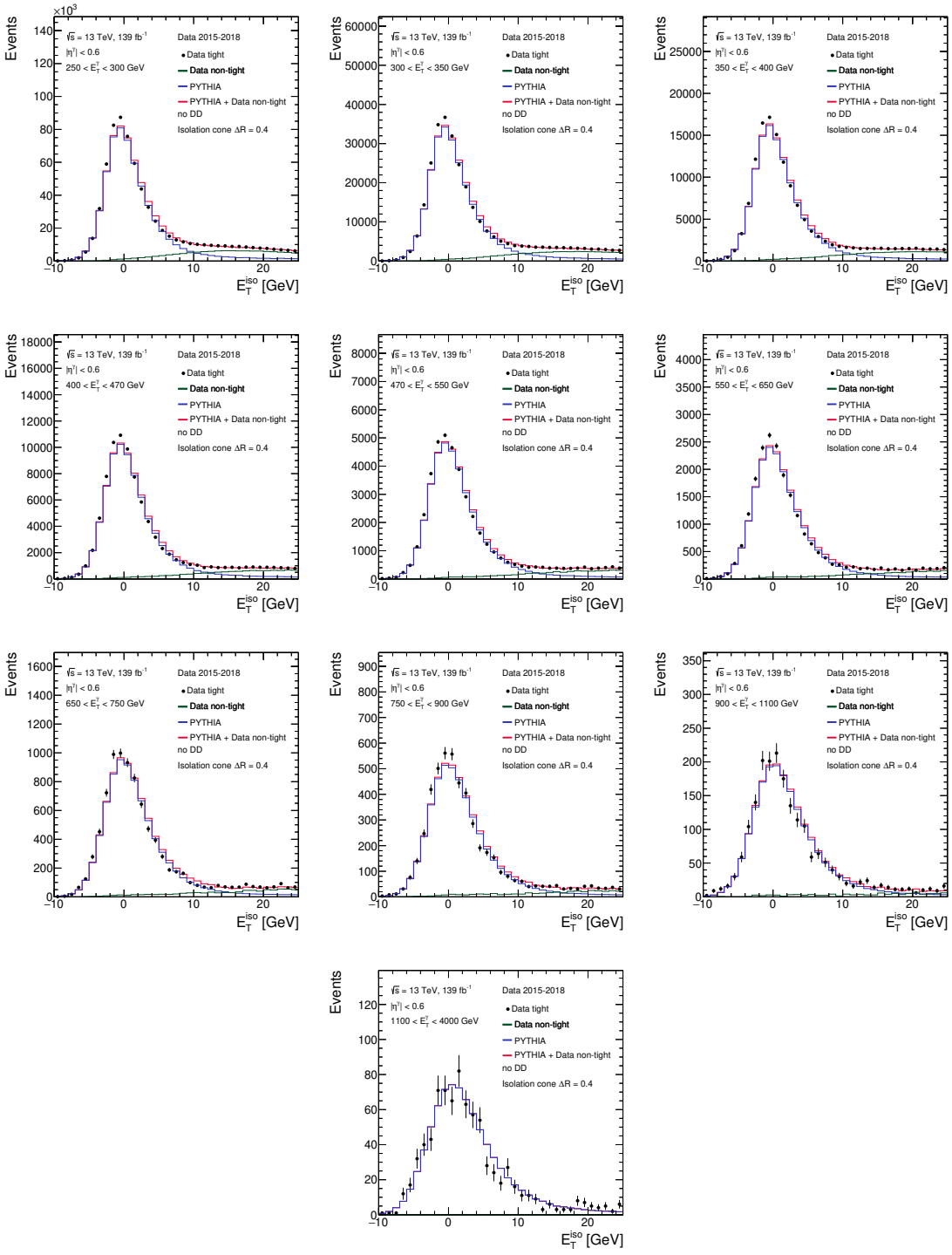


Figure C.1: Measured $E_{\mathrm{T}}^{\mathrm{iso}}$ distribution for data tight (dots) and non-tight (green histogram, normalised to the tight distribution for $E_{\mathrm{T}}^{\mathrm{iso}} > 10$ GeV) photons for $|\eta^{\gamma}| < 0.6$ in different regions of E_{T}^{γ}. The MC simulation of the signal from PYTHIA is shown before the application of the data-driven correction (blue histogram); for comparison, the PYTHIA simulation plus the non-tight data contribution are also shown (red histogram).

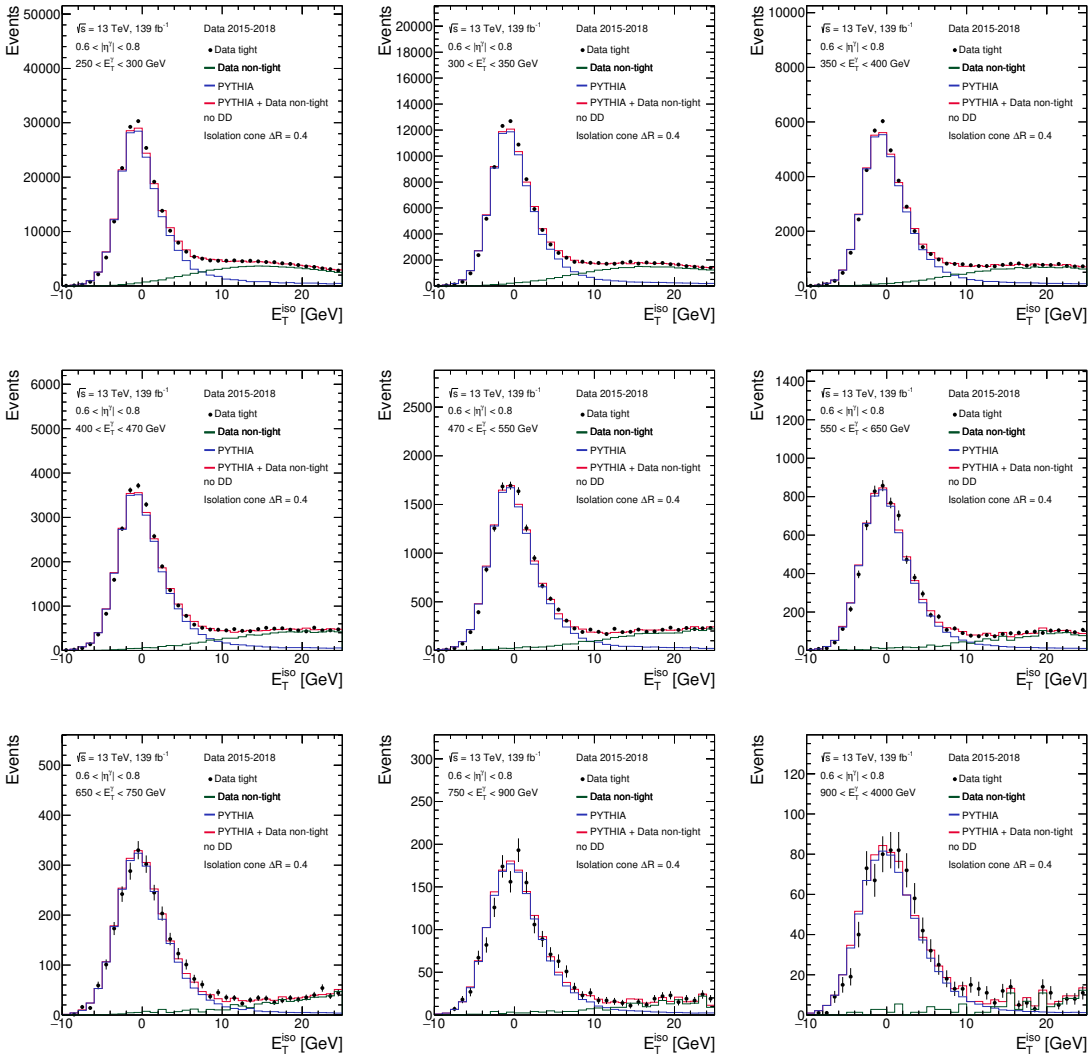


Figure C.2: Measured $E_{\mathrm{T}}^{\mathrm{iso}}$ distribution for data tight (dots) and non-tight (green histogram, normalised to the tight distribution for $E_{\mathrm{T}}^{\mathrm{iso}} > 10$ GeV) photons for $0.6 < |\eta^{\gamma}| < 0.8$ in different regions of E_{T}^{γ}. The MC simulation of the signal from PYTHIA is shown before the application of the data-driven correction (blue histogram); for comparison, the PYTHIA simulation plus the non-tight data contribution are also shown (red histogram).

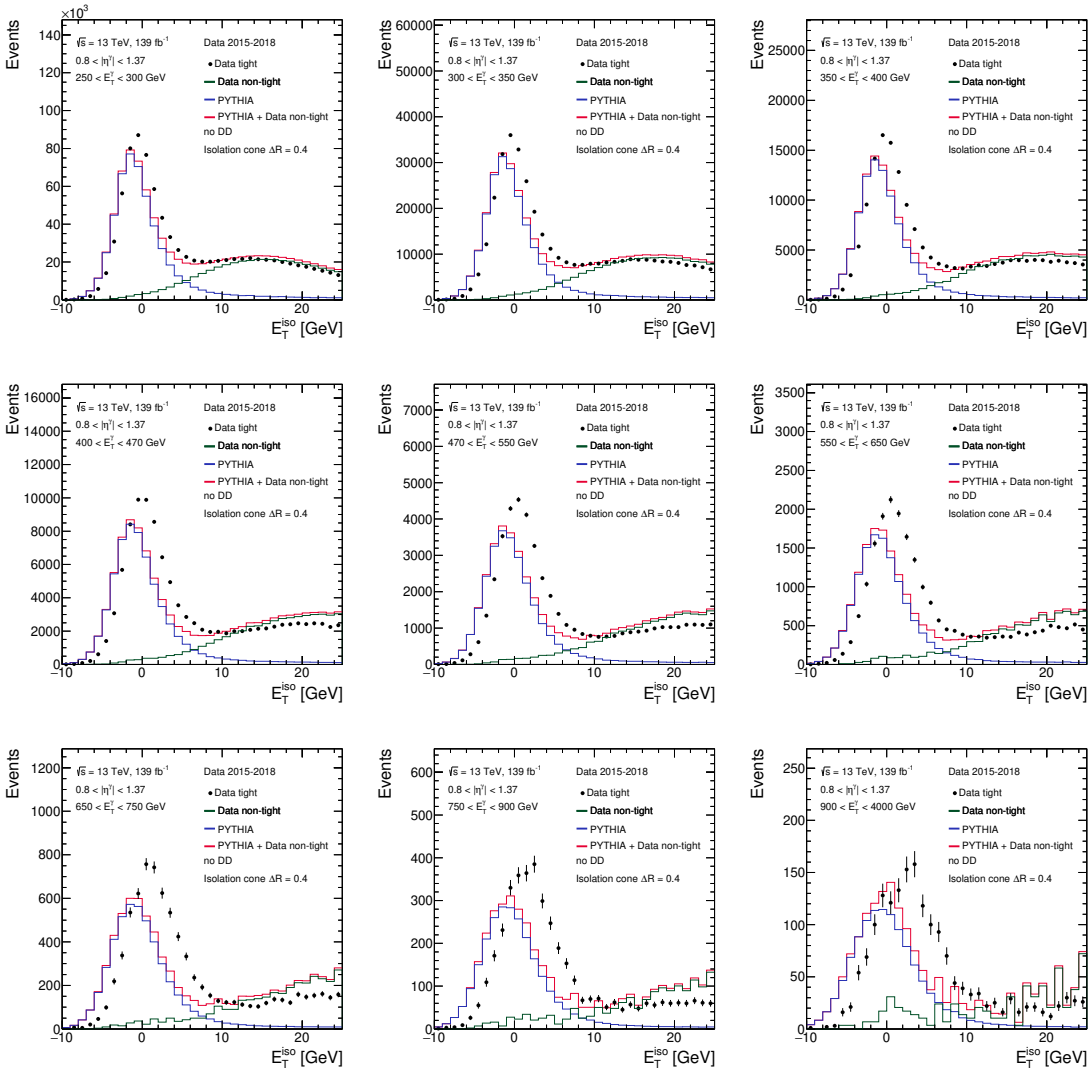


Figure C.3: Measured $E_{\mathrm{T}}^{\mathrm{iso}}$ distribution for data tight (dots) and non-tight (green histogram, normalised to the tight distribution for $E_{\mathrm{T}}^{\mathrm{iso}} > 10$ GeV) photons for $0.8 < |\eta^{\gamma}| < 1.37$ in different regions of E_{T}^{γ}. The MC simulation of the signal from PYTHIA is shown before the application of the data-driven correction (blue histogram); for comparison, the PYTHIA simulation plus the non-tight data contribution are also shown (red histogram).

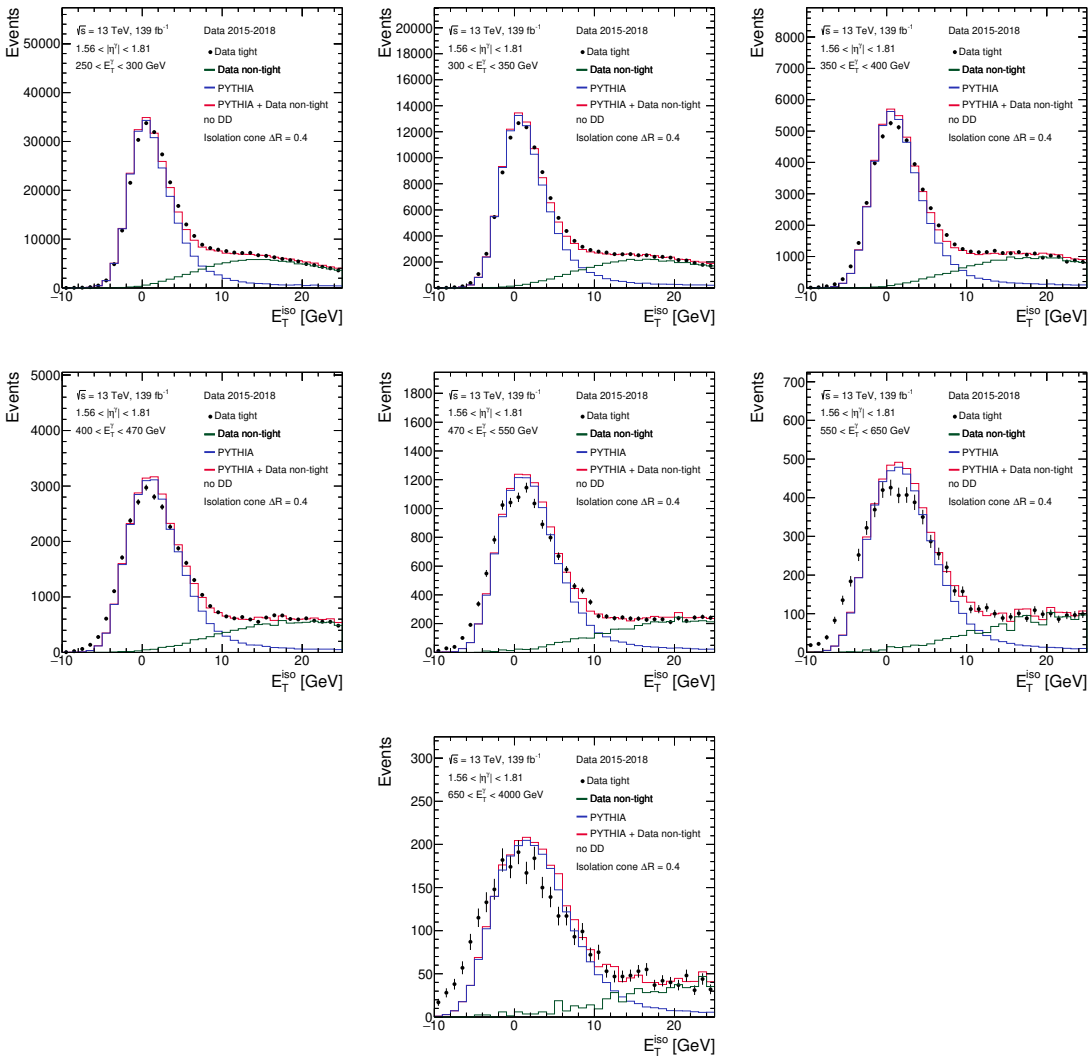


Figure C.4: Measured $E_{\mathrm{T}}^{\mathrm{iso}}$ distribution for data tight (dots) and non-tight (green histogram, normalised to the tight distribution for $E_{\mathrm{T}}^{\mathrm{iso}} > 10$ GeV) photons for $1.56 < |\eta^{\gamma}| < 1.81$ in different regions of E_{T}^{γ}. The MC simulation of the signal from PYTHIA is shown before the application of the data-driven correction (blue histogram); for comparison, the PYTHIA simulation plus the non-tight data contribution are also shown (red histogram).

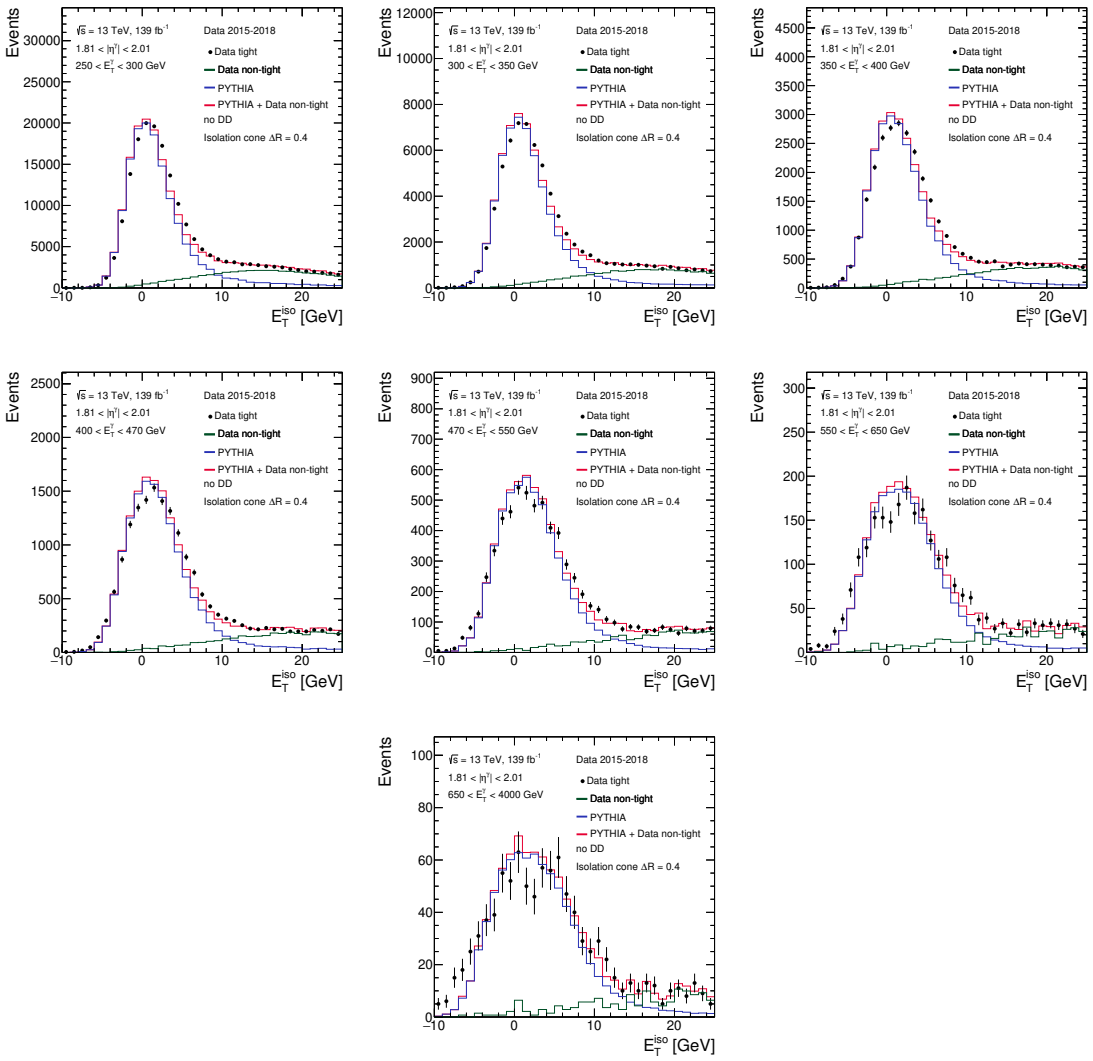


Figure C.5: Measured $E_{\mathrm{T}}^{\mathrm{iso}}$ distribution for data tight (dots) and non-tight (green histogram, normalised to the tight distribution for $E_{\mathrm{T}}^{\mathrm{iso}} > 10$ GeV) photons for $1.81 < |\eta^{\gamma}| < 2.01$ in different regions of E_{T}^{γ}. The MC simulation of the signal from PYTHIA is shown before the application of the data-driven correction (blue histogram); for comparison, the PYTHIA simulation plus the non-tight data contribution are also shown (red histogram).

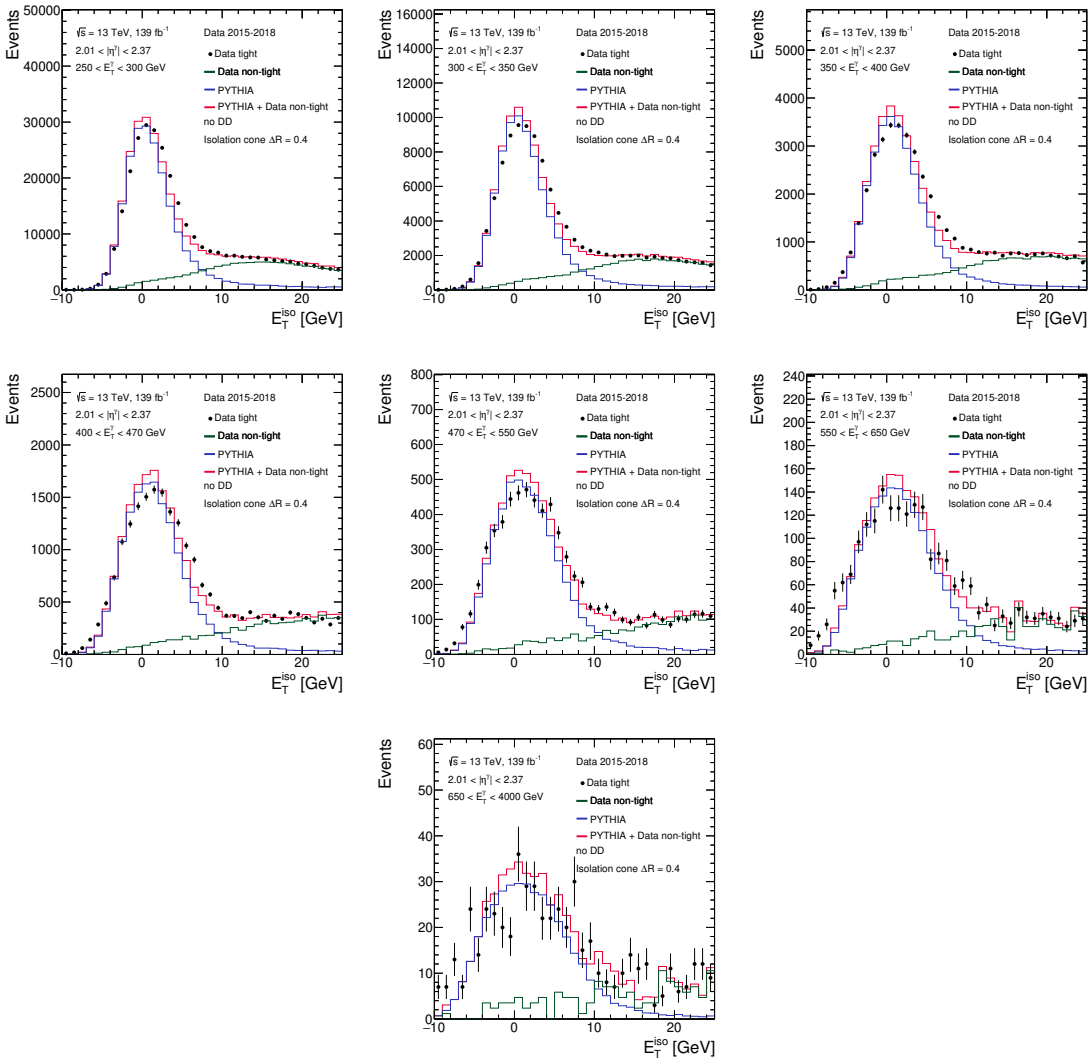


Figure C.6: Measured $E_{\mathrm{T}}^{\mathrm{iso}}$ distribution for data tight (dots) and non-tight (green histogram, normalised to the tight distribution for $E_{\mathrm{T}}^{\mathrm{iso}} > 10$ GeV) photons for $2.01 < |\eta^{\gamma}| < 2.37$ in different regions of E_{T}^{γ}. The MC simulation of the signal from PYTHIA is shown before the application of the data-driven correction (blue histogram); for comparison, the PYTHIA simulation plus the non-tight data contribution are also shown (red histogram).

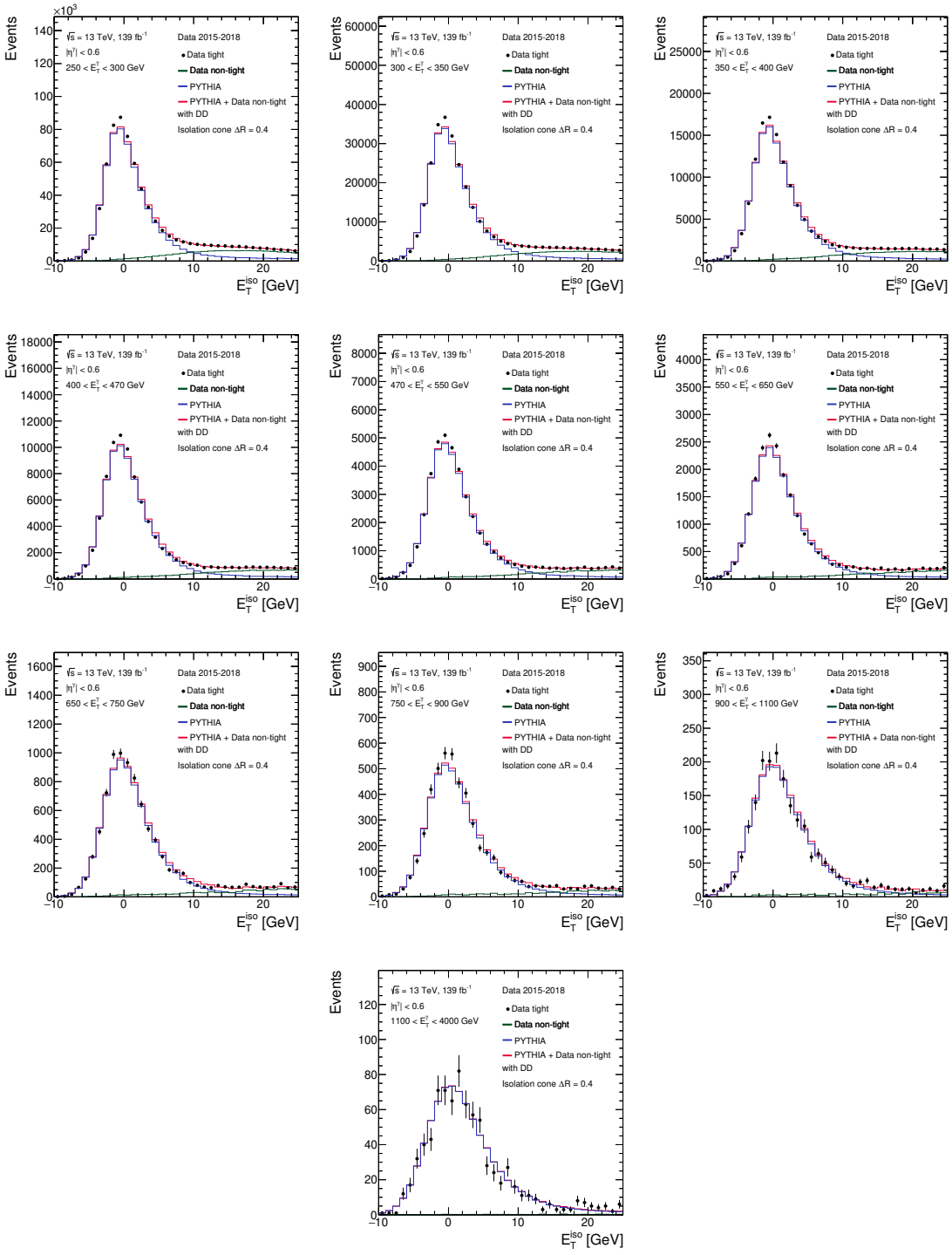


Figure C.7: Measured $E_{\mathrm{T}}^{\mathrm{iso}}$ distribution for data tight (dots) and non-tight (green histogram, normalised to the tight distribution for $E_{\mathrm{T}}^{\mathrm{iso}} > 10$ GeV) photons for $|\eta^{\gamma}| < 0.6$ in different regions of E_{T}^{γ}. The MC simulation of the signal from PYTHIA is shown after the application of the data-driven correction (blue histogram); for comparison, the PYTHIA simulation plus the non-tight data contribution are also shown (red histogram).

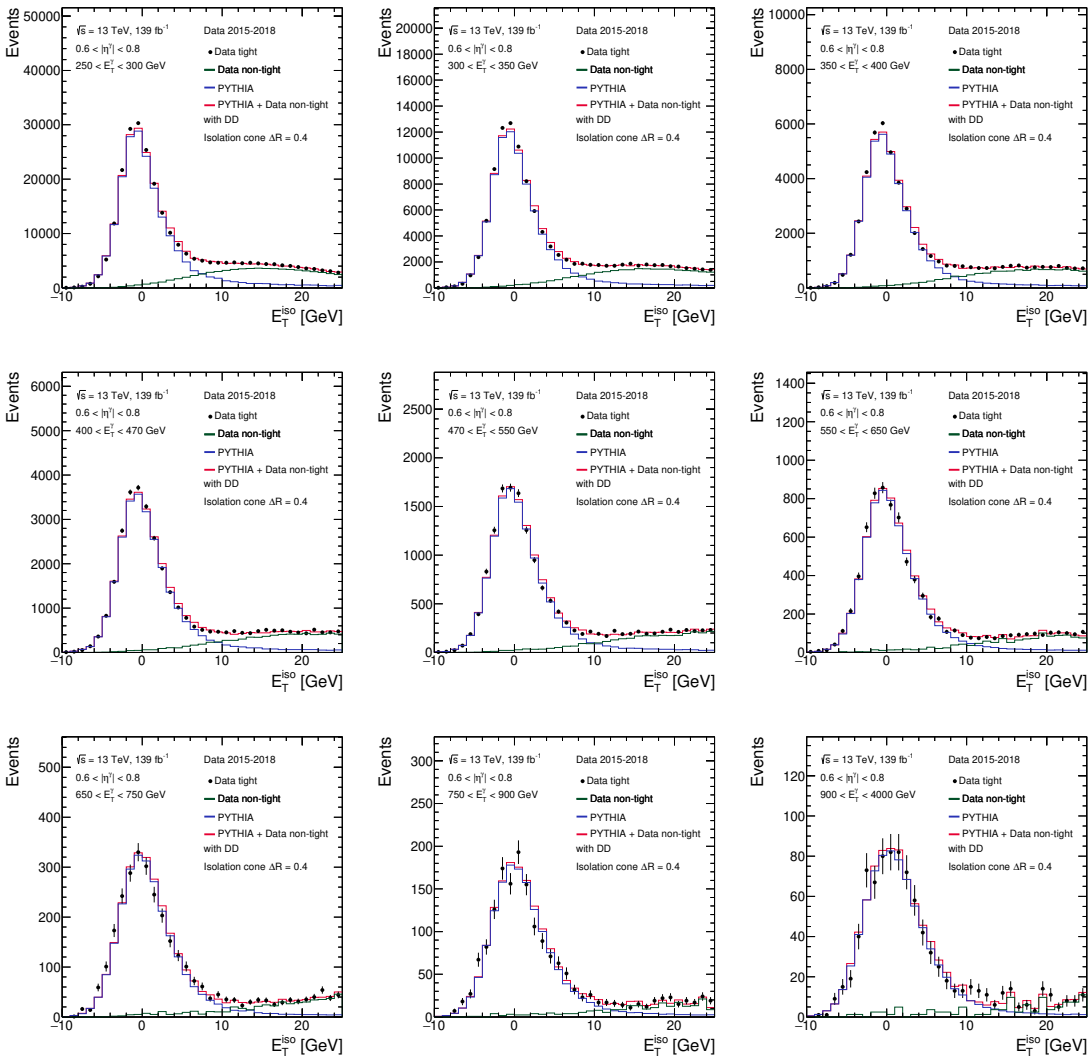


Figure C.8: Measured $E_{\mathrm{T}}^{\mathrm{iso}}$ distribution for data tight (dots) and non-tight (green histogram, normalised to the tight distribution for $E_{\mathrm{T}}^{\mathrm{iso}} > 10$ GeV) photons for $0.6 < |\eta^{\gamma}| < 0.8$ in different regions of E_{T}^{γ}. The MC simulation of the signal from PYTHIA is shown after the application of the data-driven correction (blue histogram); for comparison, the PYTHIA simulation plus the non-tight data contribution are also shown (red histogram).

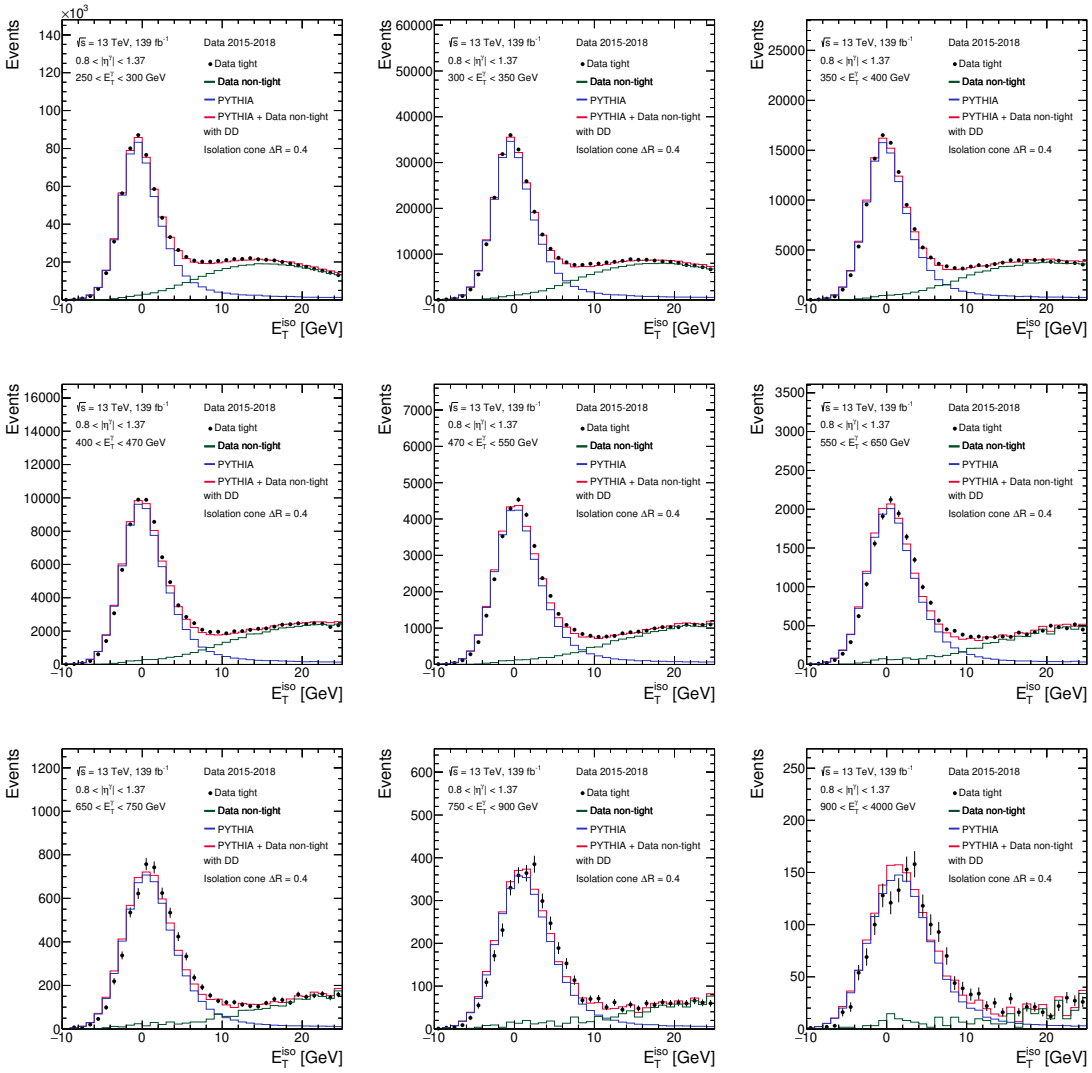


Figure C.9: Measured $E_{\mathrm{T}}^{\mathrm{iso}}$ distribution for data tight (dots) and non-tight (green histogram, normalised to the tight distribution for $E_{\mathrm{T}}^{\mathrm{iso}} > 10$ GeV) photons for $0.8 < |\eta^{\gamma}| < 1.37$ in different regions of E_{T}^{γ}. The MC simulation of the signal from PYTHIA is shown after the application of the data-driven correction (blue histogram); for comparison, the PYTHIA simulation plus the non-tight data contribution are also shown (red histogram).

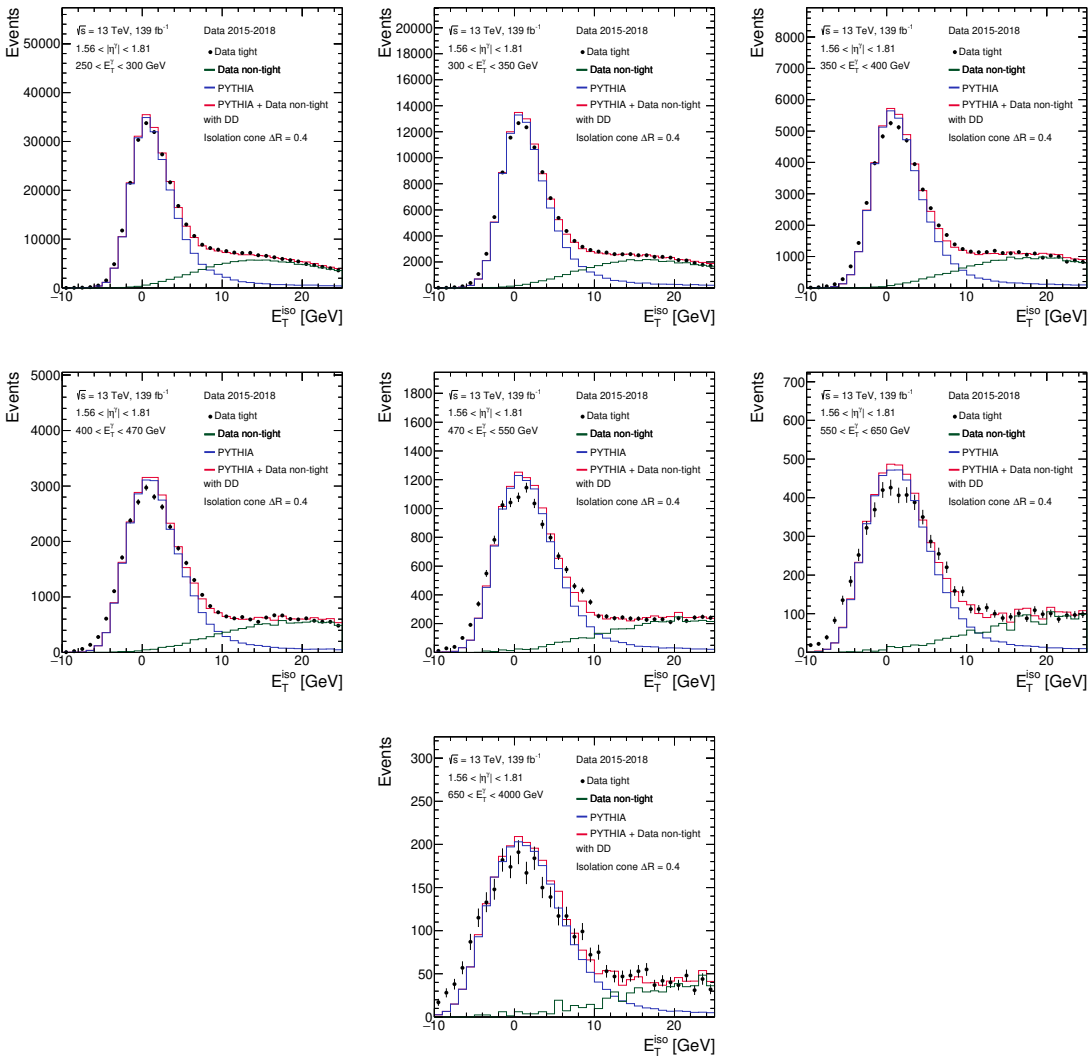


Figure C.10: Measured $E_{\mathrm{T}}^{\mathrm{iso}}$ distribution for data tight (dots) and non-tight (green histogram, normalised to the tight distribution for $E_{\mathrm{T}}^{\mathrm{iso}} > 10$ GeV) photons for $1.56 < |\eta^{\gamma}| < 1.81$ in different regions of E_{T}^{γ}. The MC simulation of the signal from PYTHIA is shown after the application of the data-driven correction (blue histogram); for comparison, the PYTHIA simulation plus the non-tight data contribution are also shown (red histogram).

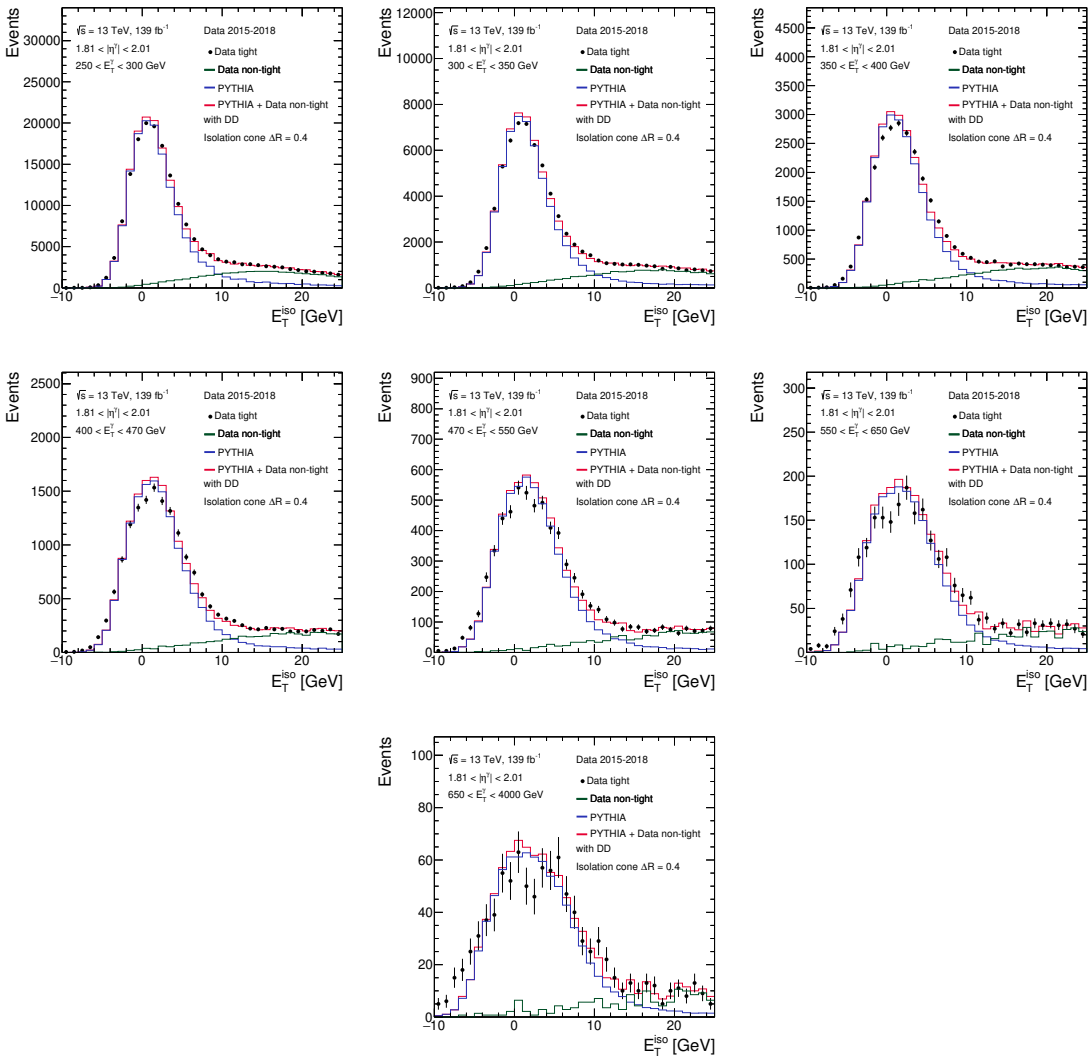


Figure C.11: Measured $E_{\mathrm{T}}^{\mathrm{iso}}$ distribution for data tight (dots) and non-tight (green histogram, normalised to the tight distribution for $E_{\mathrm{T}}^{\mathrm{iso}} > 10$ GeV) photons for $1.81 < |\eta^{\gamma}| < 2.01$ in different regions of E_{T}^{γ}. The MC simulation of the signal from PYTHIA is shown after the application of the data-driven correction (blue histogram); for comparison, the PYTHIA simulation plus the non-tight data contribution are also shown (red histogram).

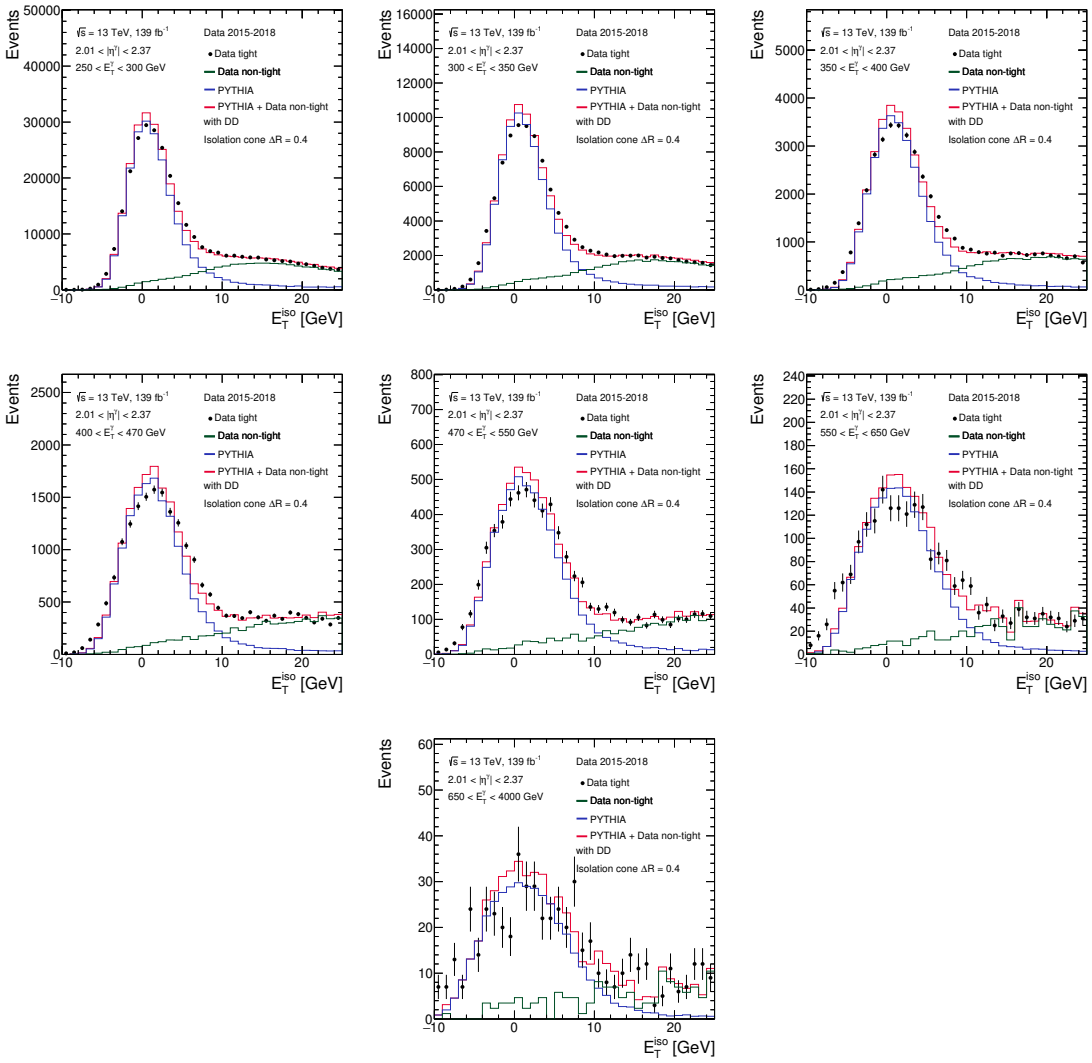


Figure C.12: Measured $E_{\mathrm{T}}^{\mathrm{iso}}$ distribution for data tight (dots) and non-tight (green histogram, normalised to the tight distribution for $E_{\mathrm{T}}^{\mathrm{iso}} > 10$ GeV) photons for $2.01 < |\eta^{\gamma}| < 2.37$ in different regions of E_{T}^{γ}. The MC simulation of the signal from PYTHIA is shown after the application of the data-driven correction (blue histogram); for comparison, the PYTHIA simulation plus the non-tight data contribution are also shown (red histogram).

Appendix D

Cross-checks using reweighted MC for the inclusive photon analysis using $36\,\mathrm{fb}^{-1}$

As mentioned in section 7.6, the cross-check was performed reweighting the MC distributions of PYTHIA at detector and particle levels to the data. The ratio of the data to the MC detector-level distributions is shown in figure D.1 as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$.

A χ^2 fit to this ratio was performed; the function used in each case is also shown in figure D.1. This function is used to perform a reweighting of the PYTHIA sample, both at detector and particle levels.

The cross check was performed by measuring the differential cross sections using the reweighted PYTHIA sample for the estimation of the bin-by-bin unfolding corrections. The relative difference with respect to the measurements obtained using PYTHIA default for the unfolding is shown in figure D.2. It is observed that the use or not of the reweighted PYTHIA does not change the measurements, since the deviations from the nominal differential cross section are smaller than 0.05%.

A further cross-check of the unfolding was performed by adjusting the fraction of the brem and hard components in PYTHIA. The shape of these two components is different: the fraction of the brem component decreases with increasing E_{T}^{γ}. Therefore, the shape of the MC distributions depends on the relative fraction of these two contributions. An improvement of the description of the data by the PYTHIA MC was achieved by performing a χ^2 fit to each data distribution of the relative contribution of the hard (α) and the brem ($1-\alpha$) contributions as the free parameter; $\alpha = 0.5$ reproduces the original prediction of PYTHIA. The χ^2 function used is

$$\chi^2(\alpha) = \sum_{\mathrm{i}} \left(\frac{N_{\mathrm{A}}^{\mathrm{sig}}(i) - N_{\mathrm{A}}^{\mathrm{MC}}(i,\alpha)}{\Delta N_{\mathrm{A}}^{\mathrm{sig}}(i)} \right)^2 ,$$

where the sum runs over the bins of a given distribution (i being the bin index), $\Delta N_{\mathrm{A}}^{\mathrm{sig}}(i)$ is the statistical uncertainty in the data signal yield in bin i,

$$N_{\mathrm{A}}^{\mathrm{MC}}(i,\alpha) = \frac{N_{\mathrm{A}}^{\mathrm{sig,TOT}}}{\alpha N_{\mathrm{A}}^{\mathrm{MC,H,TOT}} + (1-\alpha) N_{\mathrm{A}}^{\mathrm{MC,B,TOT}}} \left(\alpha N_{\mathrm{A}}^{\mathrm{MC,H}}(i) + (1-\alpha) N_{\mathrm{A}}^{\mathrm{MC,B}}(i)\right) ,$$

and α is the free parameter in the fit. The other symbols used are defined as follows: $N_{\mathrm{A}}^{\mathrm{MC,H}}(i)$ ($N_{\mathrm{A}}^{\mathrm{MC,B}}(i)$) is the number of simulated events from the hard (brem) component in bin i; $N_{\mathrm{A}}^{\mathrm{MC,H,TOT}}$ ($N_{\mathrm{A}}^{\mathrm{MC,B,TOT}}$) is the total number of simulated events from the hard (brem) component; $N_{\mathrm{A}}^{\mathrm{sig,TOT}}$ is the total signal yield in data. The optimisation was done in two steps. In the first step, the distributions of the signal yield in data using signal leakage fractions from PYTHIA default were fitted. In the second step, the signal yields in data were re-evaluated using the signal leakage fractions from PYTHIA in which the hard and brem contributions were mixed according to the fitted value of α obtained in the first step; the resulting distributions of the signal yield in data were fitted again. The fitted values of α from the two steps were so similar that no further iterations were performed.

Figure D.3 shows the data distributions after background subtraction (using the optimised admixture of hard and brem components in PYTHIA obtained in the second step) compared to the MC simulations of PYTHIA optimised. An improved description of the data by PYTHIA optimised is obtained for $E_{\mathrm{T}}^{\gamma} < 1$ TeV. The resulting fractions (α) of the two components are included in figure D.3.

The cross-check was performed by measuring the cross sections using the optimised PYTHIA for the estimation of the signal leakage fractions and the bin-by-bin unfolding corrections. The relative difference with respect to the measurements obtained using PYTHIA default for the unfolding is shown in figure D.4. It is observed that the use or not of the optimised PYTHIA does not change the measurements, since the deviations from the nominal differential cross section are typically smaller than 0.5%.

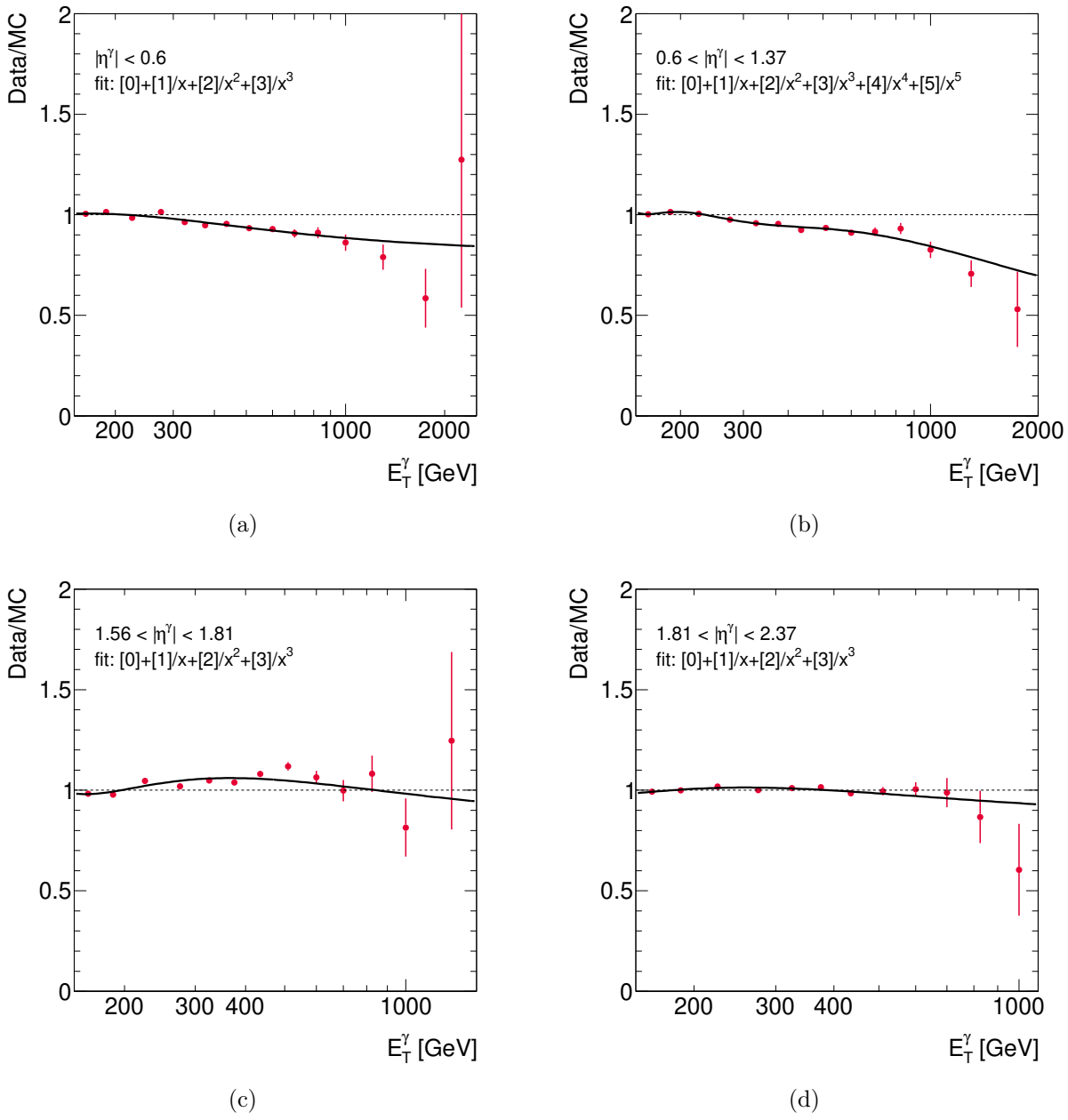


Figure D.1: Ratio of the data and PYTHIA distributions (dots) as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. The solid line represents the fit to the ratio. The error bars represent the statistical uncertainty only.

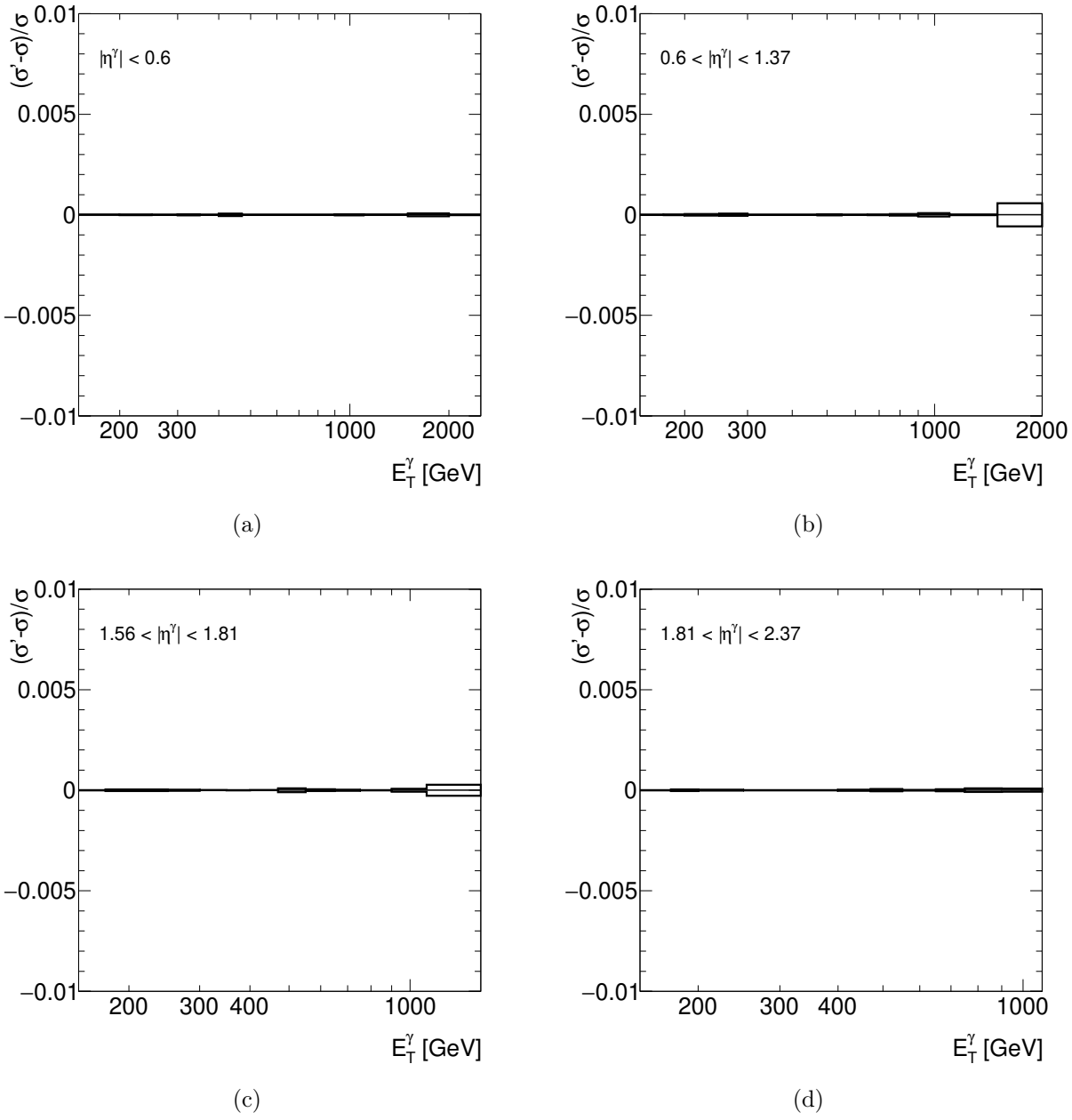


Figure D.2: Relative difference between the cross section measured using the reweighted PYTHIA for the estimation of the bin-by-bin unfolding corrections and those using PYTHIA default as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$.

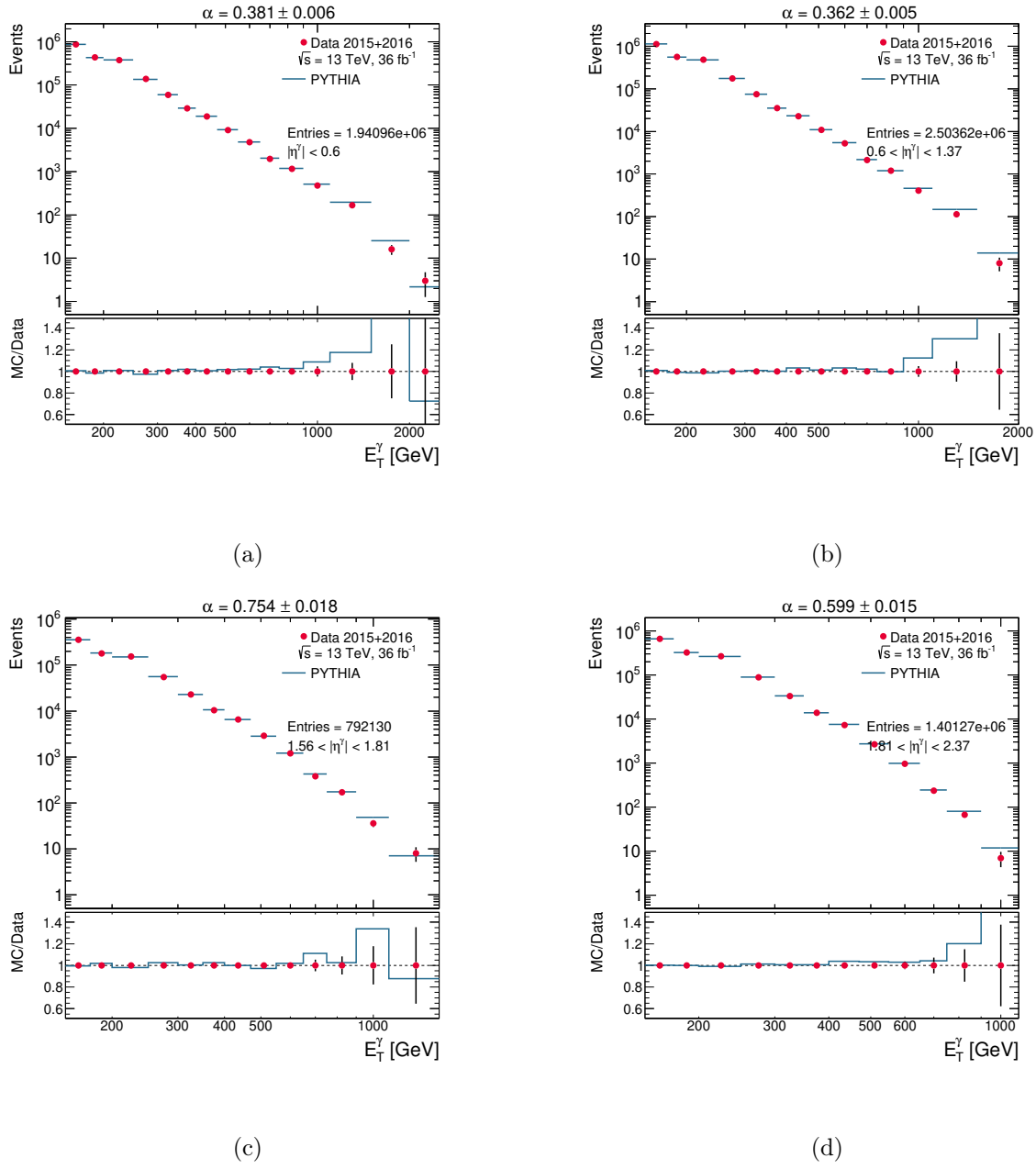


Figure D.3: The estimated signal yields in data (dots) using signal leakage fractions from PYTHIA optimised as functions of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. For comparison, the MC simulations of the signal from PYTHIA optimised according to the optimised value α shown are also included. The MC distributions are normalised to the total number of data events. The ratio of the MC to the data is shown in the lower part of the figures.

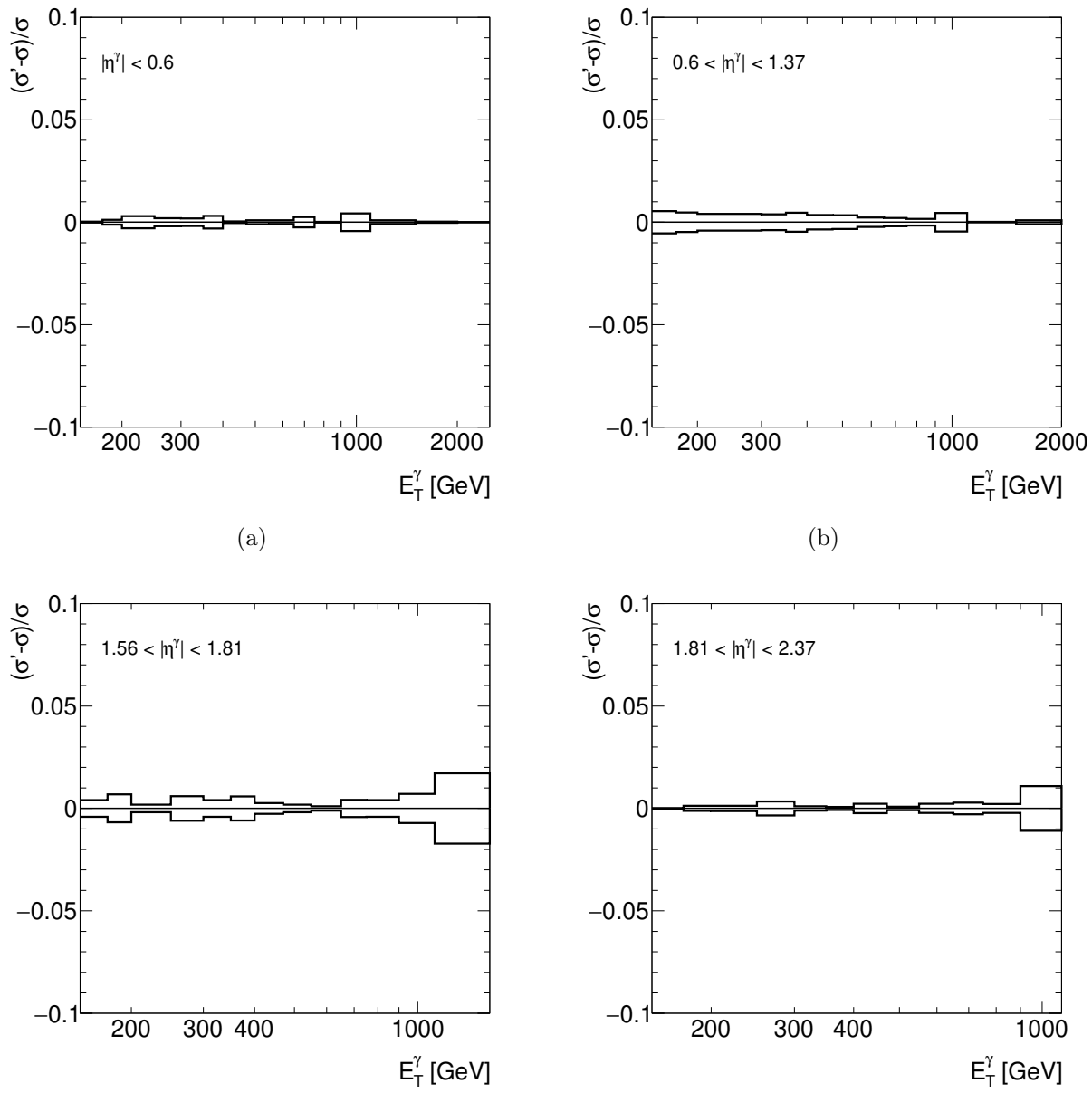


Figure D.4: Relative difference between the cross section measured using the optimised PYTHIA for the estimation of the signal leakage fractions and the bin-by-bin unfolding corrections and those using PYTHIA default as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$.

Appendix E

Cross-check using Bayesian unfolding for the inclusive photon analysis using $36\,\mathrm{fb}^{-1}$

A more sophisticated unfolding method, based on the iterative application of the Bayes' theorem, was investigated to cross-check and validate the results obtained using the bin-by-bin method explained in section 7.6. The Bayesian unfolding method [153] takes properly into account the migrations between bins and the purity and efficiency of the selection criteria when the MC description of the data is not adequate. The main requirement of this method is to have sufficiently large MC samples to construct the response matrices. The Bayesian unfolding method as implemented in RooUnfold [158] was used for these tests. Repeated application of Bayes' theorem is used to invert the response matrix. Regularisation is achieved by stopping the iterations: the regularisation parameter is the number of iterations. The results are fairly insensitive to the precise setting used and four iterations are usually sufficient [158]. RooUnfoldBayes takes the particle level distribution as its initial prior, rather than a flat distribution, as described in ref. [153]. This should not bias the result once it has been iterated, but could reach an optimum value after fewer iterations. As shown in ref. [158], the uncertainties calculated by propagation of errors in the Bayes method were found to be significantly underestimated compared to those given by a toy MC. This was found to be due to an omission in the original method outlined in ref. [153]. The Bayes method gives the unfolded distribution (*estimated causes*, $\hat{n}(C_\mathrm{i})$) as the result of applying the unfolding matrix (M_ij) to the measurements (*effects*, $n(E_\mathrm{j})$): $\hat{n}(C_\mathrm{i}) = \sum_\mathrm{j} M_\mathrm{ij} n(E_\mathrm{j})$, where $M_\mathrm{ij} = P\left(E_\mathrm{j}|C_\mathrm{i}\right) n_0(C_\mathrm{i}) / \left(\epsilon_\mathrm{i} \sum_\mathrm{l} P\left(E_\mathrm{j}|C_\mathrm{l}\right) n_0(C_\mathrm{l})\right)$, $P\left(E_\mathrm{j}|C_\mathrm{i}\right)$ is the $n_\mathrm{E} \times n_\mathrm{C}$ response matrix, $\epsilon_\mathrm{i} = \sum_\mathrm{j} P\left(E_\mathrm{j}|C_\mathrm{i}\right)$ are efficiencies, and $n_0(C_\mathrm{l})$ is the prior (initially arbitrary, flat or the particle level MC), which is updated in subsequent iterations. The covariance matrix is calculated by error propagation from $n(E_\mathrm{j})$, but M_ij is assumed to be itself independent of $n(E_\mathrm{j})$, which is only true for the first iteration. For subsequent iterations, $n_0(C_\mathrm{i})$ is replaced by $\hat{n}(C_\mathrm{i})$ from the previous iteration, and $\hat{n}(C_\mathrm{i})$ depends on $n(E_\mathrm{j})$. This is taken into account in RooUnfold by computing the error

propagation matrix, which is used then to obtain the covariance matrix of the unfolded distribution.

Figure E.1 (E.2) shows the relative difference between the results obtained using PYTHIA (SHERPA) with N (where N goes from 1 to 5) iterations with respect to those using $N = 4$ iterations as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. The oscillations on the results for $N < 4$ are quite large but disappear for $N > 4$. Therefore, the nominal Bayes' unfolding was performed using 4 iterations. A cross-check using only 2 iterations was also included.

The comparison between the cross sections unfolded via bin-by-bin and the Bayes methods is shown in figure E.3 (E.4) using $N = 2$ ($N = 4$) iterations. The differences between the cross sections obtained via the two methods are typically much smaller than 1% using either 2 or 4 iterations. In the regions of phase space where the data statistics is poorer ($E_{\text{T}}^{\gamma} \gtrsim 800$ GeV), some deviations are observed, but they are well within the statistical uncertainty of the measured cross section. For comparison, the statistical uncertainty of the measured cross section is also included in figures E.3 and E.4, and shows that the size of this uncertainty is typically much bigger than the difference of the cross sections obtained with the two methods.

All the results shown validate the use of the bin-by-bin unfolding method as the nominal method to measure the cross sections. This is due to the adequate description of the data provided by the MC (see figure 7.14) and the very diagonal transfer matrices (see figure 7.15).

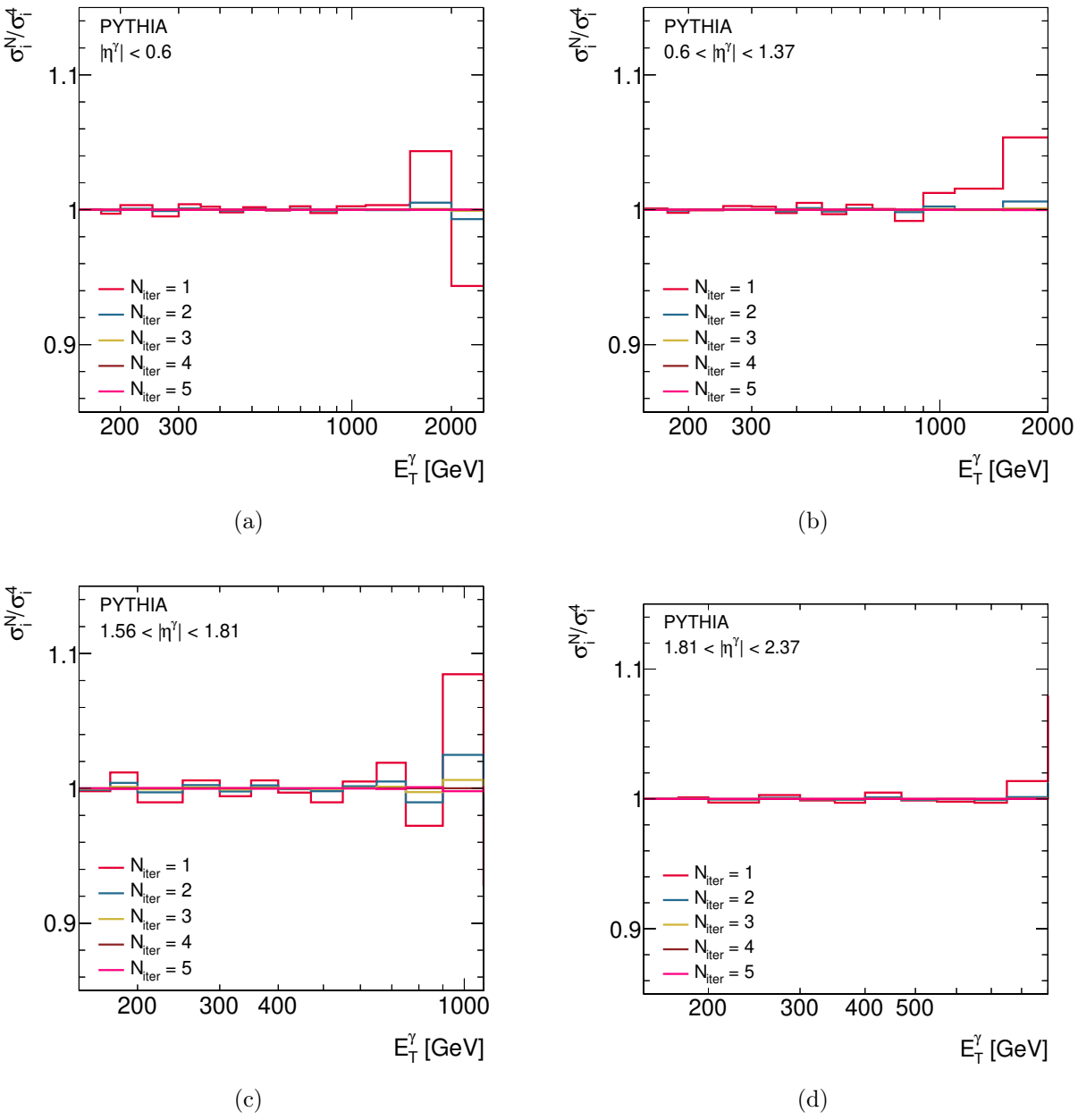


Figure E.1: Ratio between the cross sections unfolded using the Bayesian method with N iterations and the results unfolded using $N = 4$ iterations as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. The PYTHIA MC was used to unfold the cross sections.

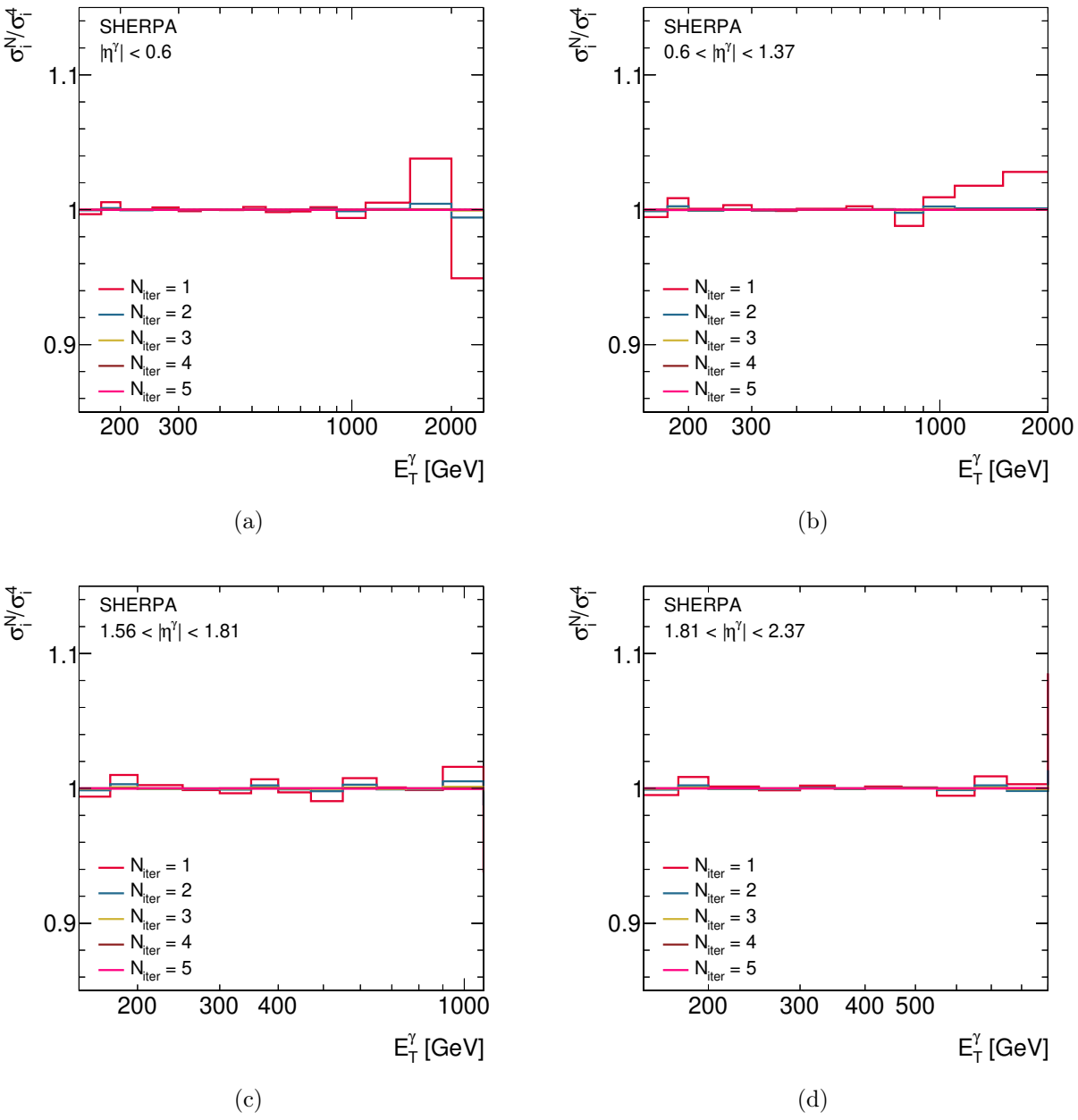


Figure E.2: Ratio between the cross sections unfolded using the Bayesian method with N iterations and the results unfolded using $N = 4$ iterations as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. The SHERPA MC was used to unfold the cross sections.

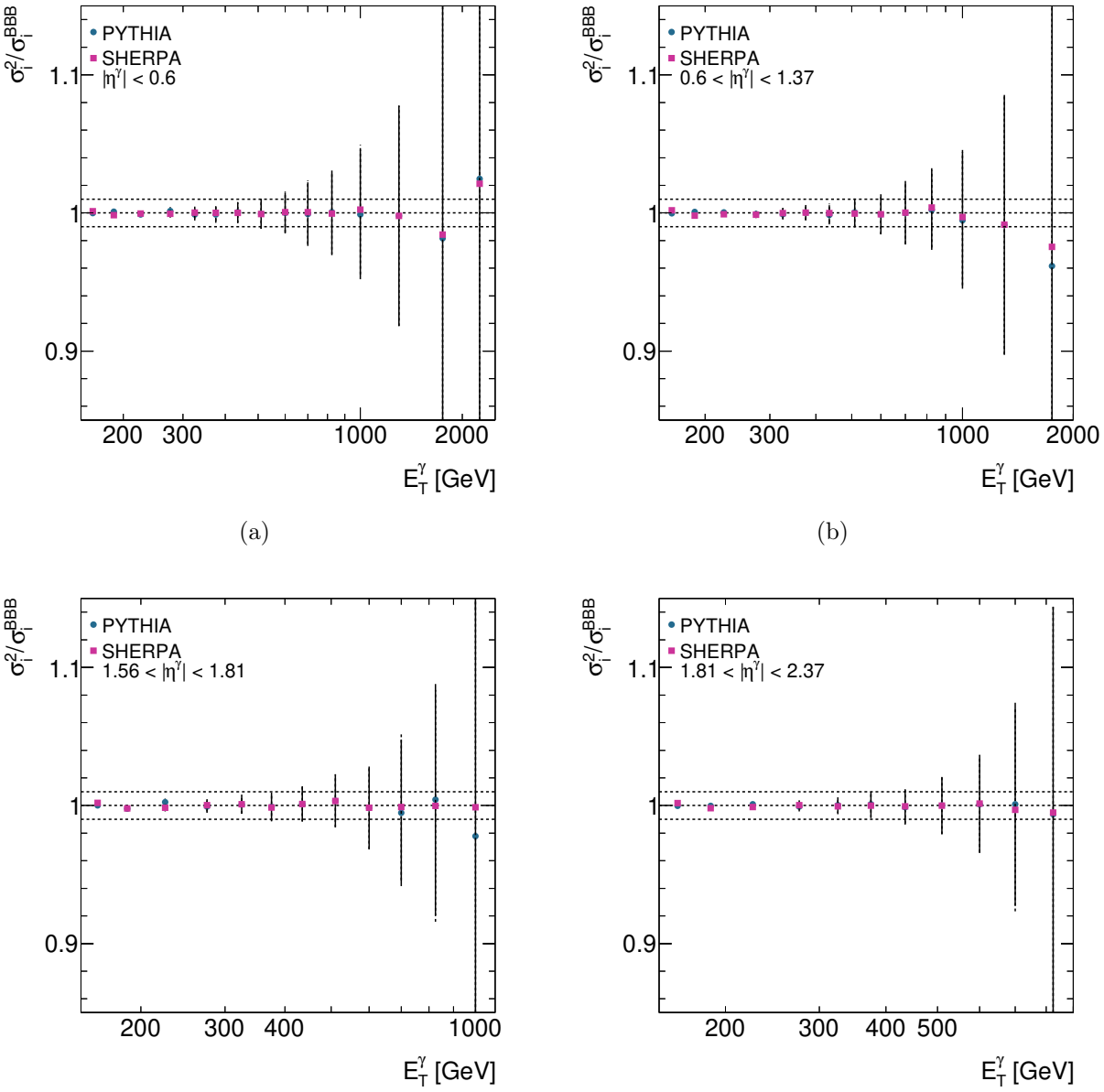


Figure E.3: Ratio between the cross sections unfolded using the Bayesian method with two iterations and the nominal cross sections as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. The error bars display the statistical uncertainty on the cross section.

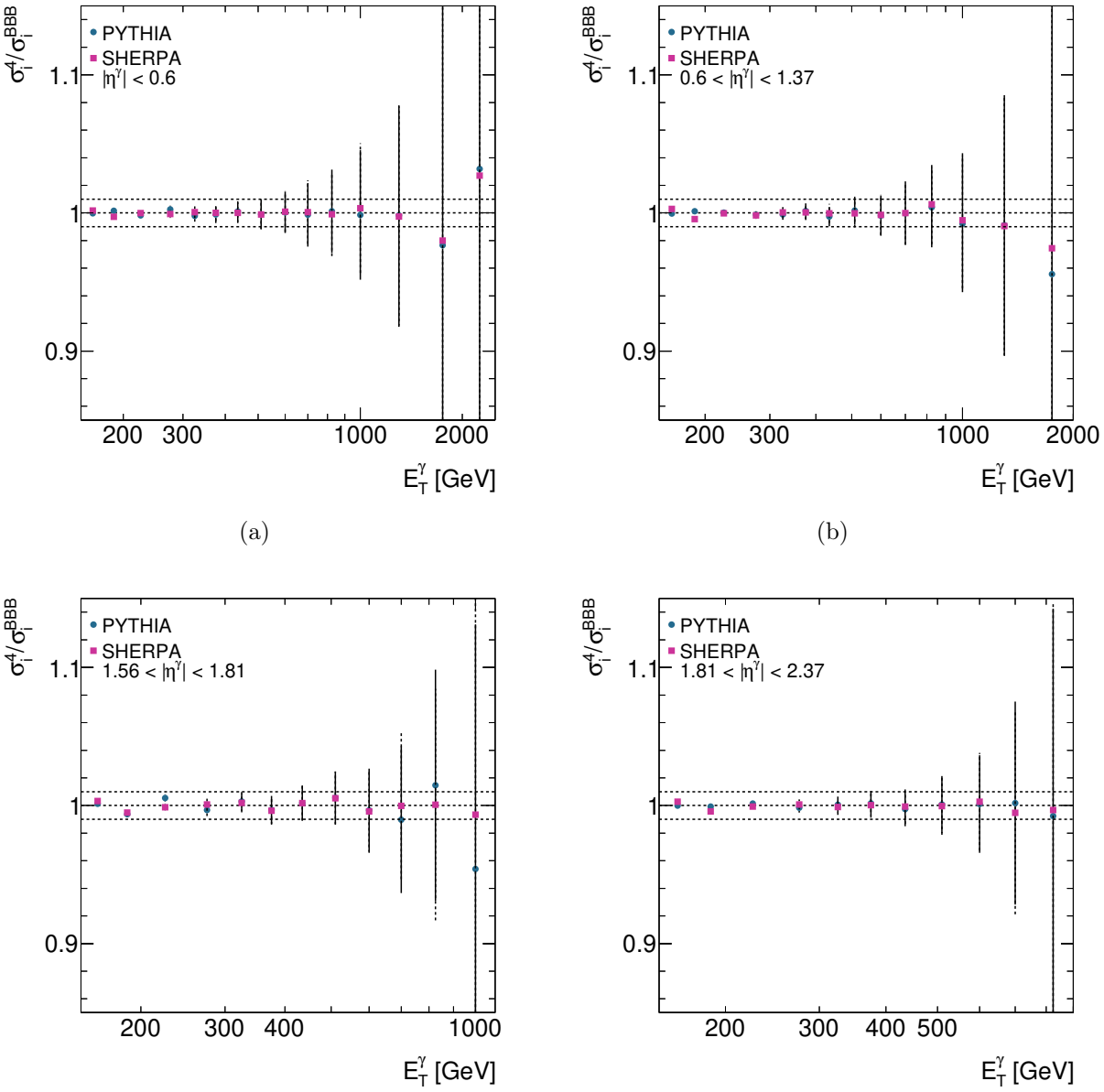


Figure E.4: Ratio between the cross sections unfolded using the Bayesian method with four iterations and the nominal cross sections as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. The error bars display the statistical uncertainty on the cross section.

Appendix F

Trigger efficiencies for the inclusive photon analysis using $139\,\mathrm{fb}^{-1}$

The trigger efficiencies measured in data as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$ for tight non-isolated, non-tight isolated, and non-tight non-isolated photons are shown from figure F.1 to F.3 for both $R = 0.2$ and $R = 0.4$ samples.

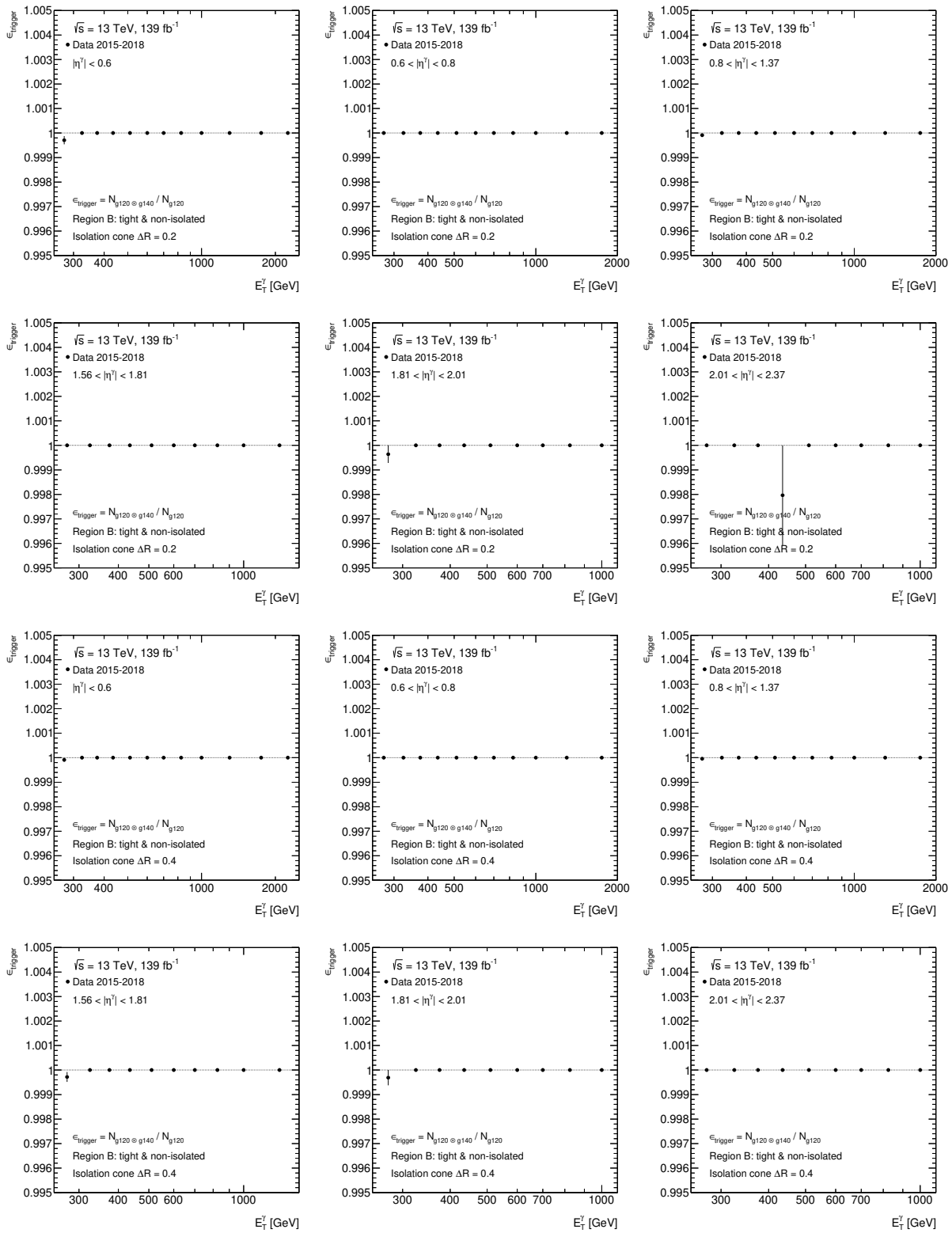


Figure F.1: Trigger efficiency in data as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$ for tight non-isolated photons in the $R = 0.2$ (top) and $R = 0.4$ (bottom) samples.

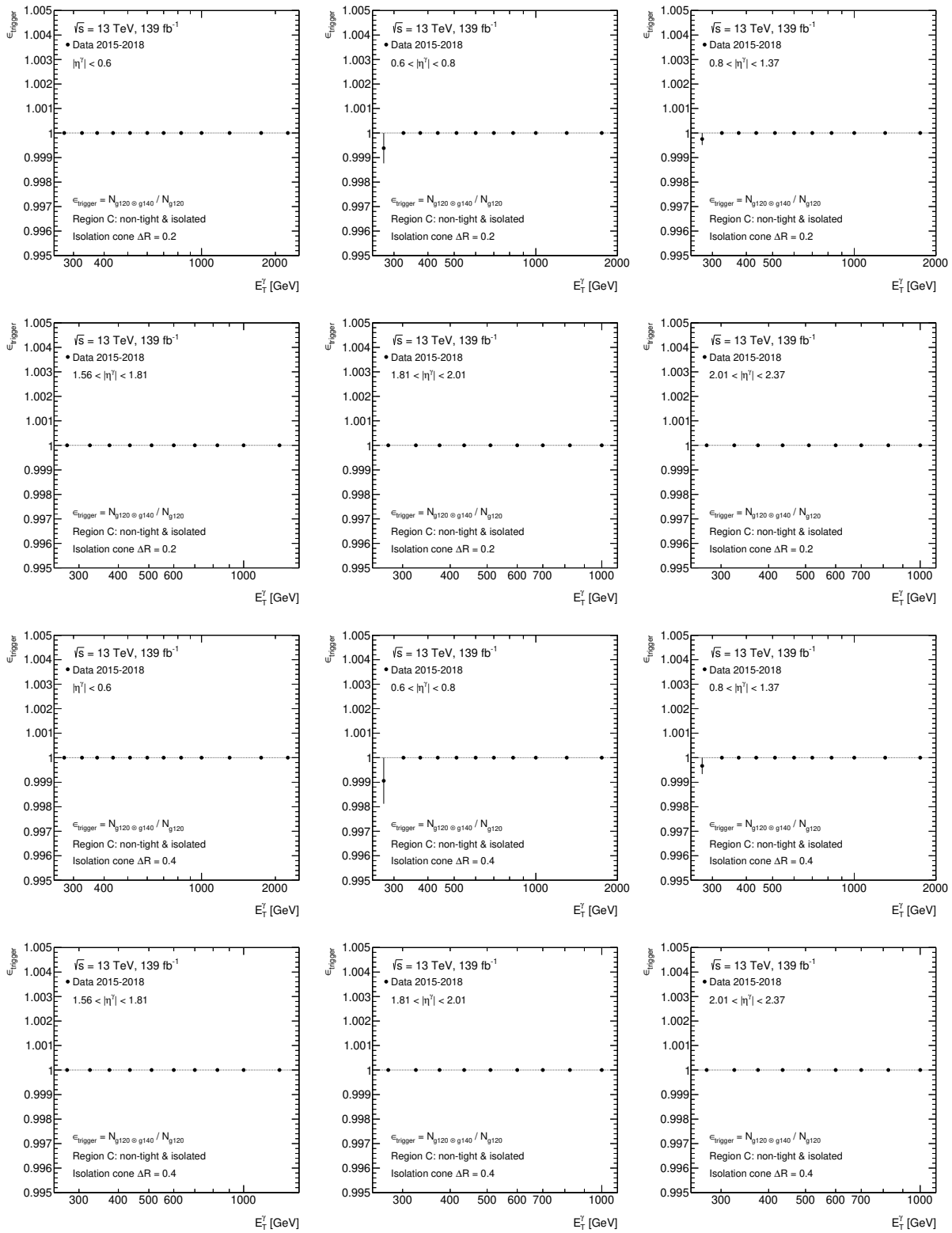


Figure F.2: Trigger efficiency in data as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$ for non-tight isolated photons in the $R = 0.2$ (top) and $R = 0.4$ (bottom) samples.

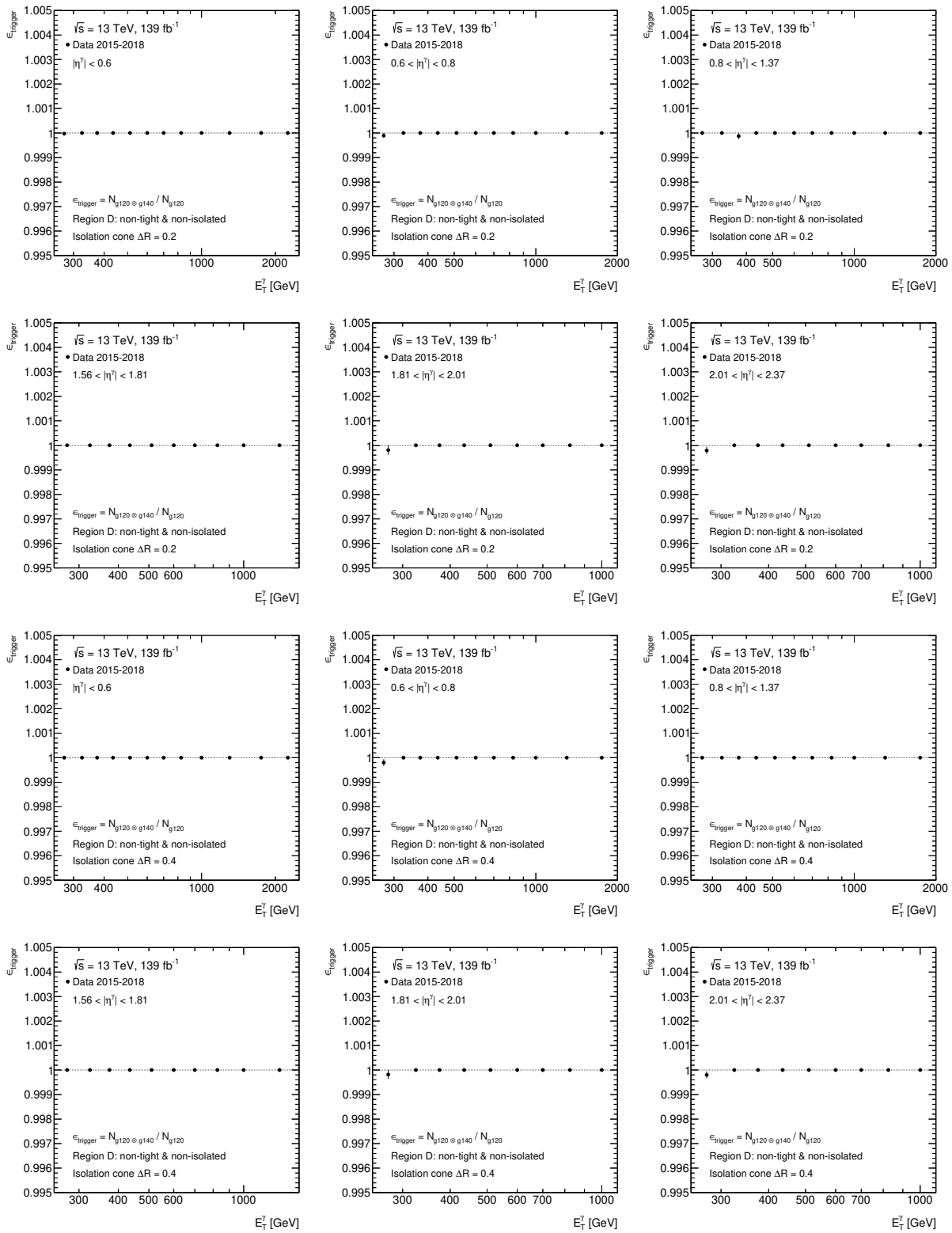


Figure F.3: Trigger efficiency in data as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$ for non-tight non-isolated photons in the $R = 0.2$ (top) and $R = 0.4$ (bottom) samples.

Appendix G

Cross-checks using reweighted MC for the inclusive photon analysis using $139\,\mathrm{fb}^{-1}$

As mentioned in section 8.6, the cross-check was performed reweighting the MC distributions of SHERPA at detector and particle levels to the data. The ratio of the data to the MC detector-level distributions is shown in figure G.1 as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$ and for both $R = 0.2$ and $R = 0.4$ samples.

A χ^2 fit to this ratio was performed; the function used in each case is also shown in figure G.1. This function is used to perform a reweighting of the SHERPA sample, both at detector and particle levels.

The cross check was performed by unfolding the default detector level prediction of SHERPA using the reweighted SHERPA sample for the bayes unfolding. The relative difference with respect to the particle level prediction of SHERPA default is shown in figure G.2. The deviations from the default cross section predicted by SHERPA are typically much smaller than 0.1% except in the tail of the most forward $|\eta^{\gamma}|$ region, where they reach up to 0.3%.

An additional cross-check of the unfolding was performed by adjusting the fraction of the brem and hard components in PYTHIA. The methodology for this cross-check is described in appendix D. The description of the data provided by PYTHIA was improved by performing a χ^2 fit to each data distribution of the relative contribution of the hard (α) and brem ($1 - \alpha$) contributions as the free parameter. Figure G.3 shows the data distributions after background subtraction (using the optimised admixture of hard and brem components in PYTHIA) compared to the MC simulations of PYTHIA optimised. An improved description of the data by PYTHIA optimised is obtained for $E_{\mathrm{T}}^{\gamma} < 1$ TeV. The resulting fractions (α) of the two components are included in figure G.3.

The cross-check was performed by measuring the cross sections using the optimised

PYTHIA for the estimation of the signal leakage fractions and the bayes unfolding corrections. The relative difference with respect to the measurements obtained using PYTHIA default for the unfolding is shown in figure G.4. It is observed that the use or not of the optimised PYTHIA does not change the measurements, since the deviations from the nominal differential cross section are typically smaller than 0.5%.

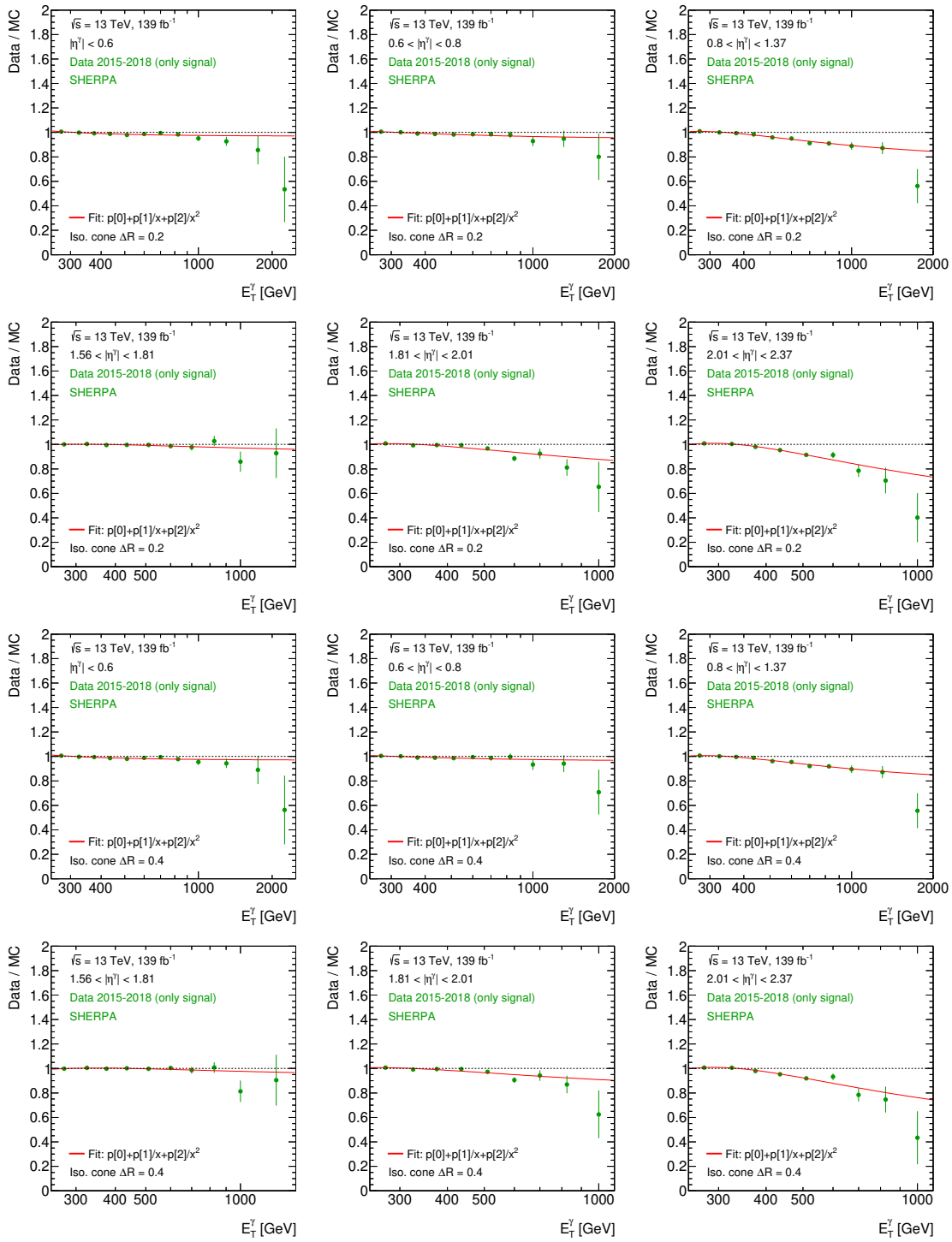


Figure G.1: Ratio of the data and SHERPA distributions (dots) as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. The solid line represents the fit to the ratio. The error bars represent the statistical uncertainty only. Results for the $R = 0.2$ ($R = 0.4$) sample are shown at the top (bottom).

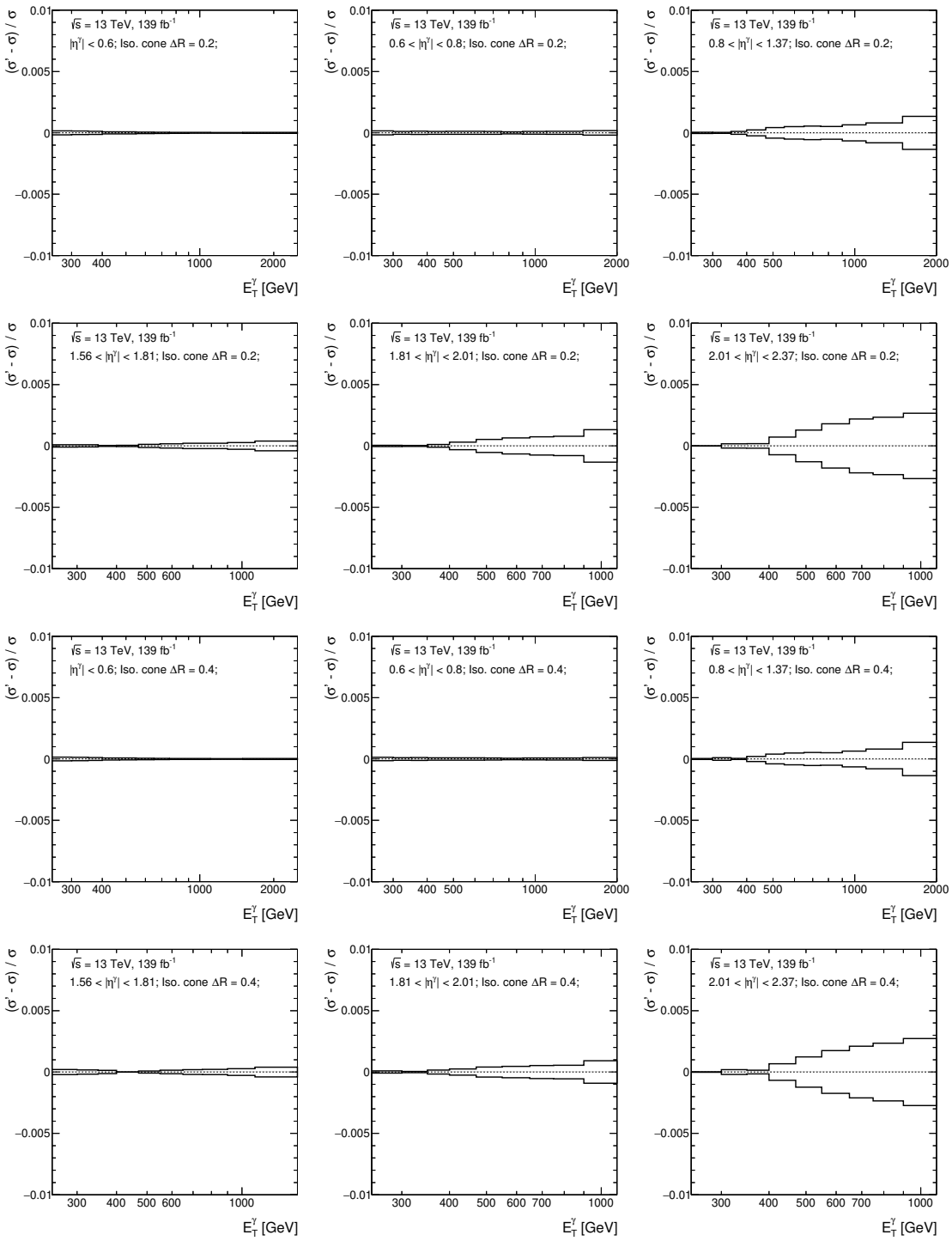


Figure G.2: Relative difference between the default SHERPA detector level prediction unfolded with the reweighted SHERPA and the default SHERPA particle level prediction as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. Results for the $R = 0.2$ ($R = 0.4$) sample are shown at the top (bottom).

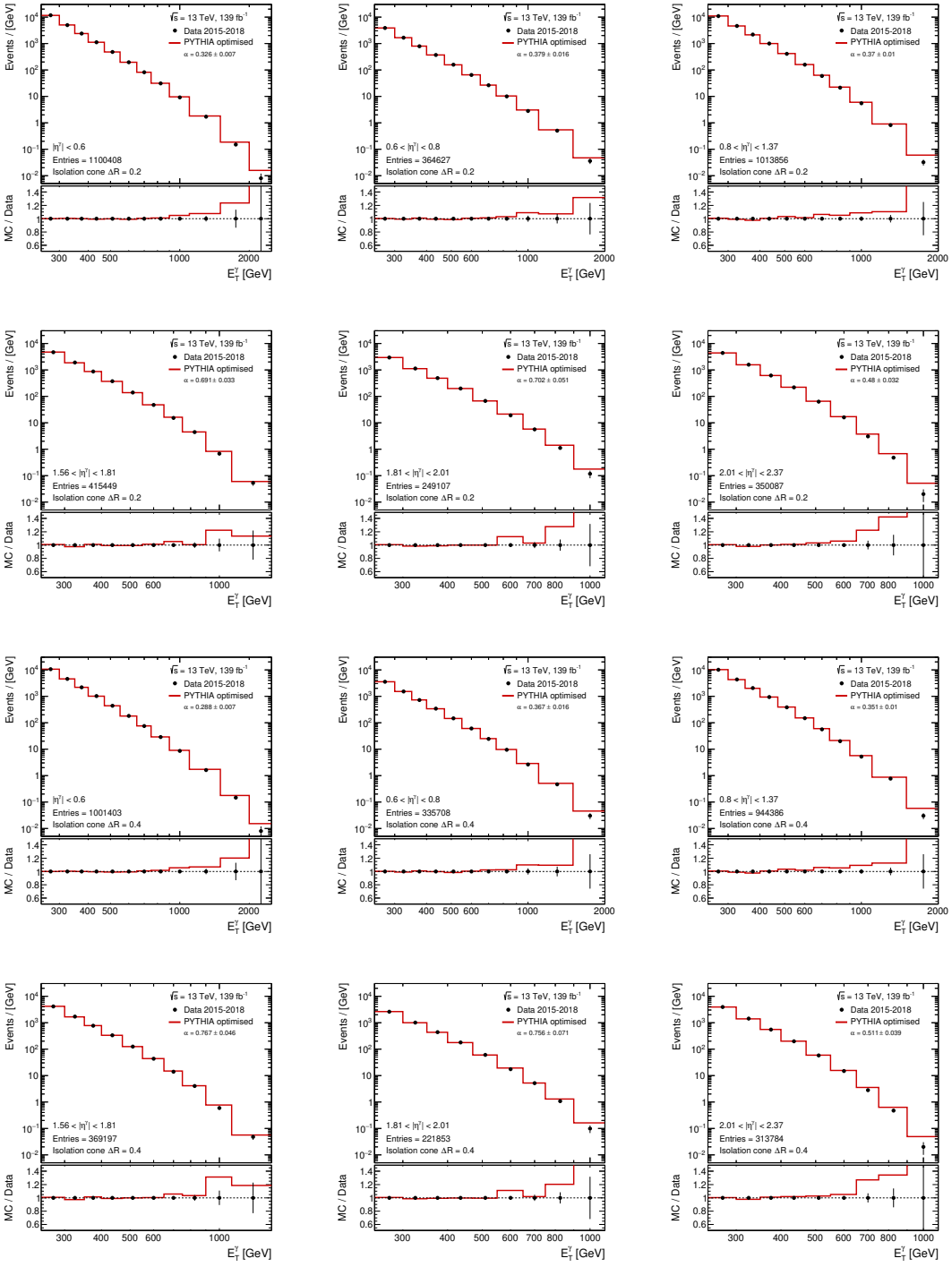


Figure G.3: The estimated signal yields in data (dots) using signal leakage fractions from PYTHIA optimised as functions of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. For comparison, the MC simulations of the signal from PYTHIA optimised according to the optimised value α shown are also included. The MC distributions are normalised to the total number of data events. The ratio of the MC to the data is shown in the lower part of the figures. Results for the $R = 0.2$ ($R = 0.4$) sample are shown at the top (bottom).

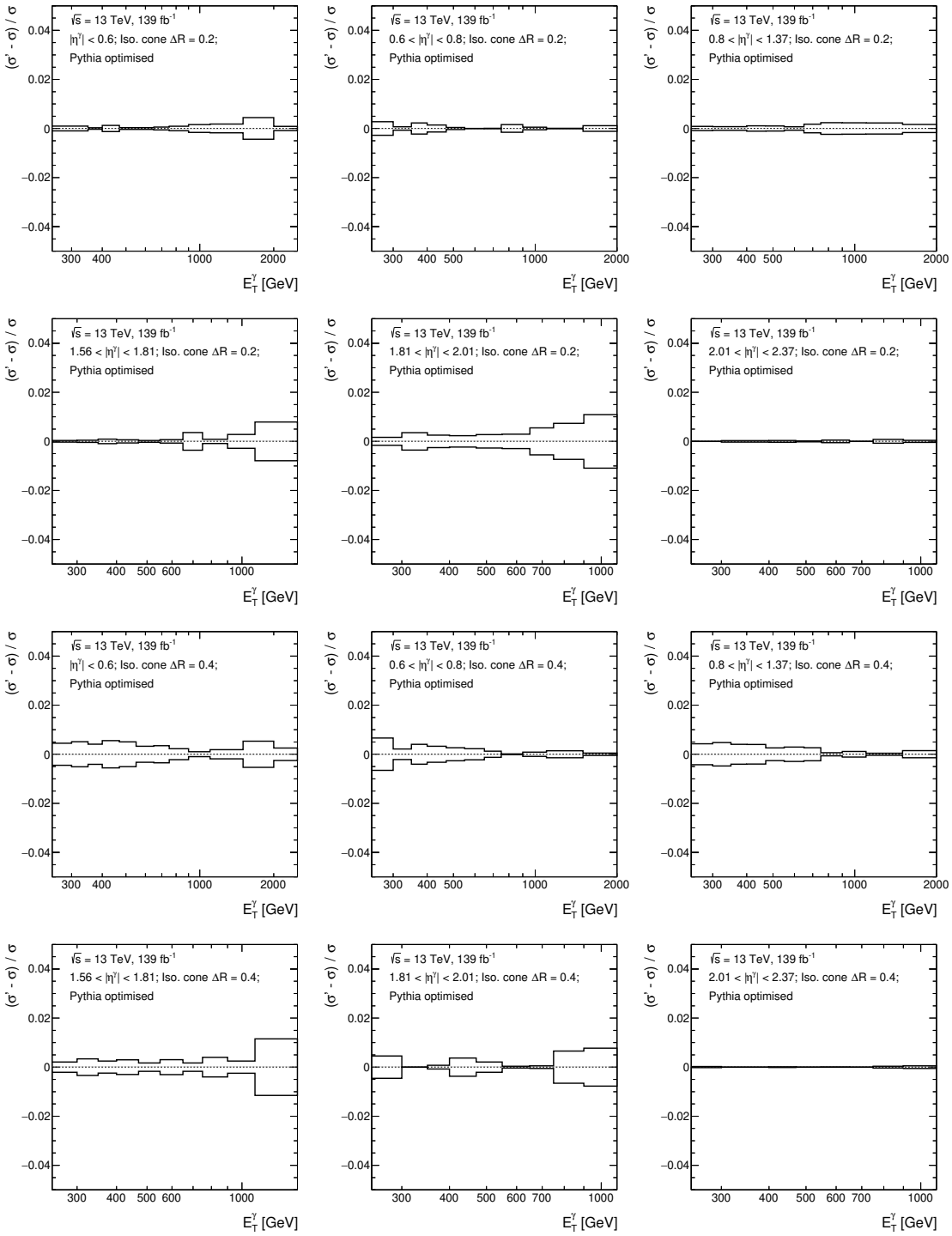


Figure G.4: Relative difference between the cross section measured using the optimised PYTHIA for the estimation of the signal leakage fractions and the bayes unfolding corrections and those using PYTHIA default as a function of E_{T}^{γ} in the different regions of $|\eta^{\gamma}|$. Results for the $R = 0.2$ ($R = 0.4$) sample are shown at the top (bottom).

Appendix H

Fits to the MPF and DB distributions

The r_{MPF} and direct balance data distributions for the PFlow (EMtopo) sample are shown in figures H.1 and H.2 (H.3 and H.4), respectively. Each distribution corresponds to a different range in $p_{\mathrm{T}}^{\mathrm{ref}}$, from $25 < p_{\mathrm{T}}^{\mathrm{ref}} < 45$ GeV to $1200 < p_{\mathrm{T}}^{\mathrm{ref}} < 1400$ GeV, which is the last bin included in the calibration. The response distributions (blue dots) are rebinned in the likelihood fit approach to obtain the best fit to a Gaussian distribution (orange line). The mean of the fit is taken as the average response (R_{MPF} or R_{DB}), which is used to compute the data-to-MC calibration factor introduced in section 6.2.2.

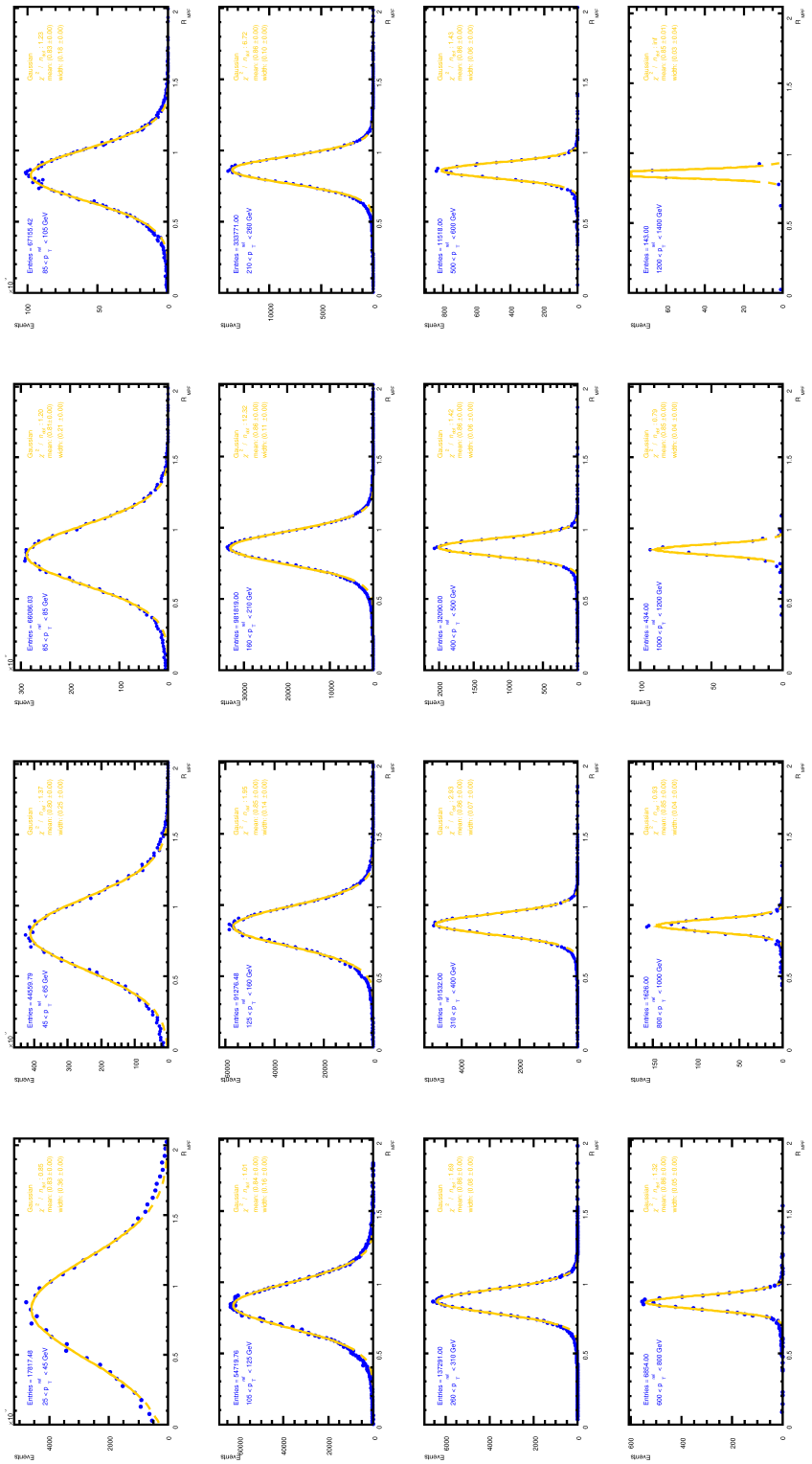


Figure H.1: Distributions of r_{MPF} in 2015-2017 data for different $p_{\mathrm{T}}^{\mathrm{ref}}$ bins using PFlow jets.

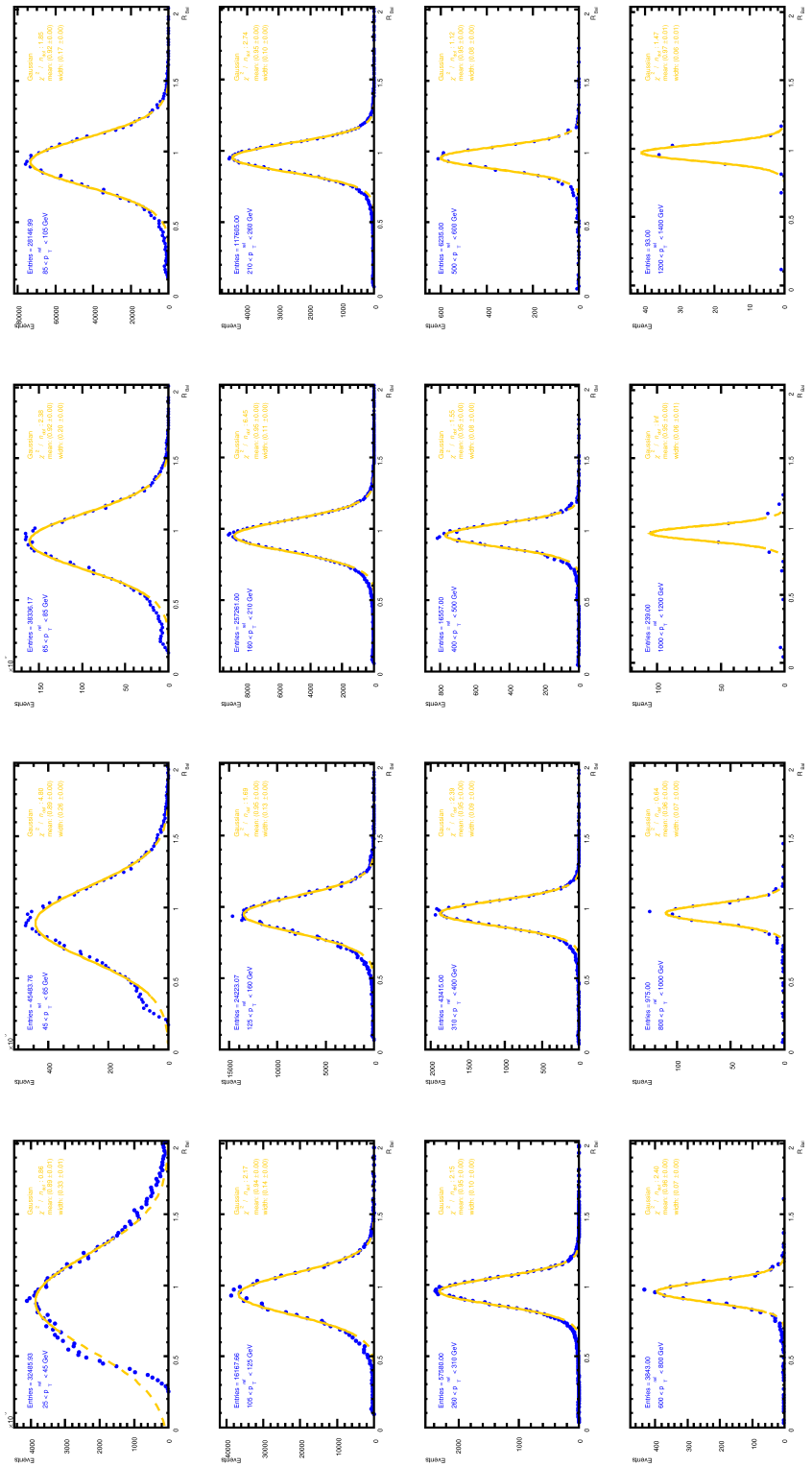


Figure H.2: Distributions of $p_{\mathrm{T}}^{\mathrm{j1}}/p_{\mathrm{T}}^{\mathrm{ref}}$ in 2015-2017 data for different $p_{\mathrm{T}}^{\mathrm{ref}}$ bins using PFlow jets.

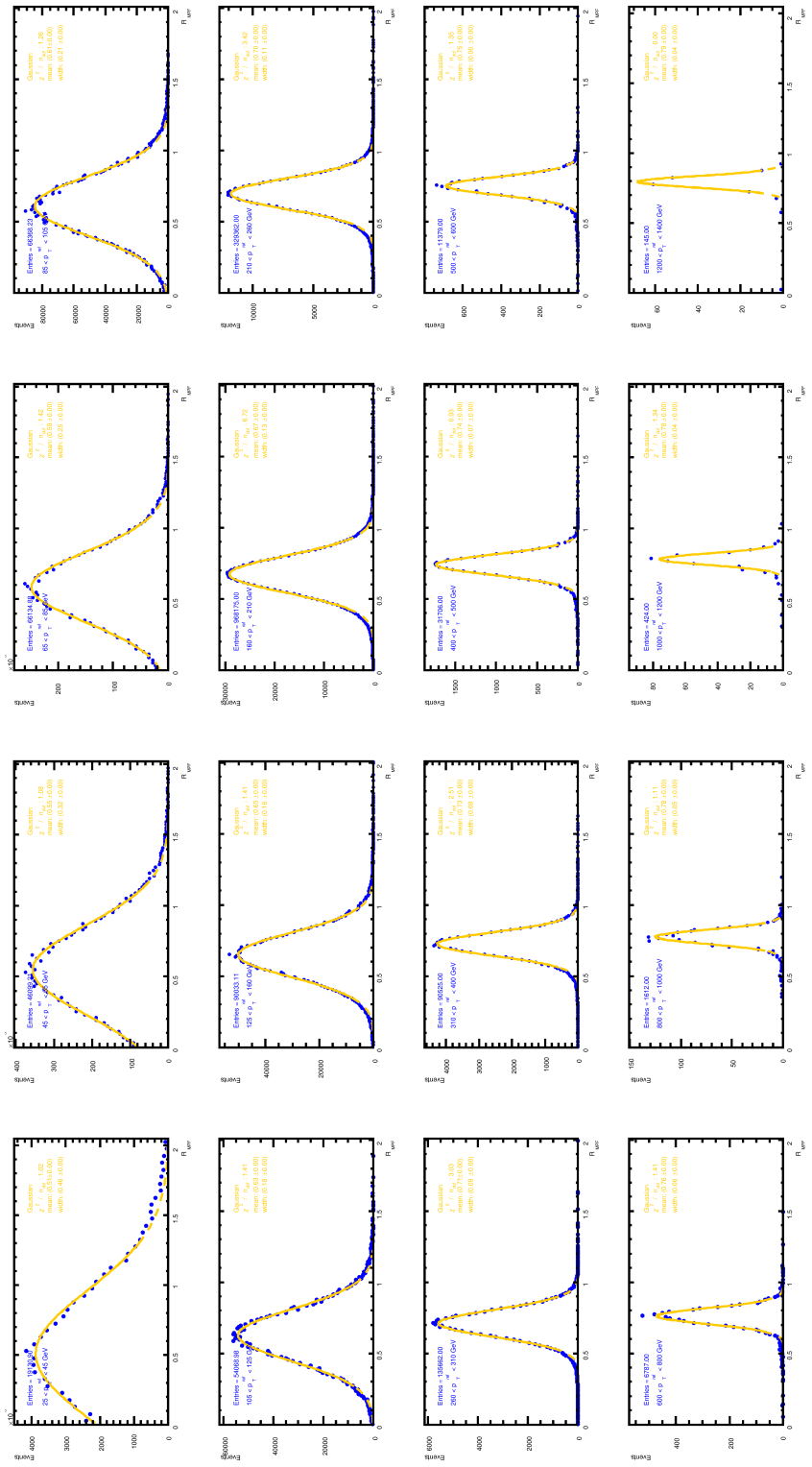


Figure H.3: Distributions of r_{MPF} in 2015-2017 data for different $p_{\mathrm{T}}^{\mathrm{ref}}$ bins using EMtopo jets.

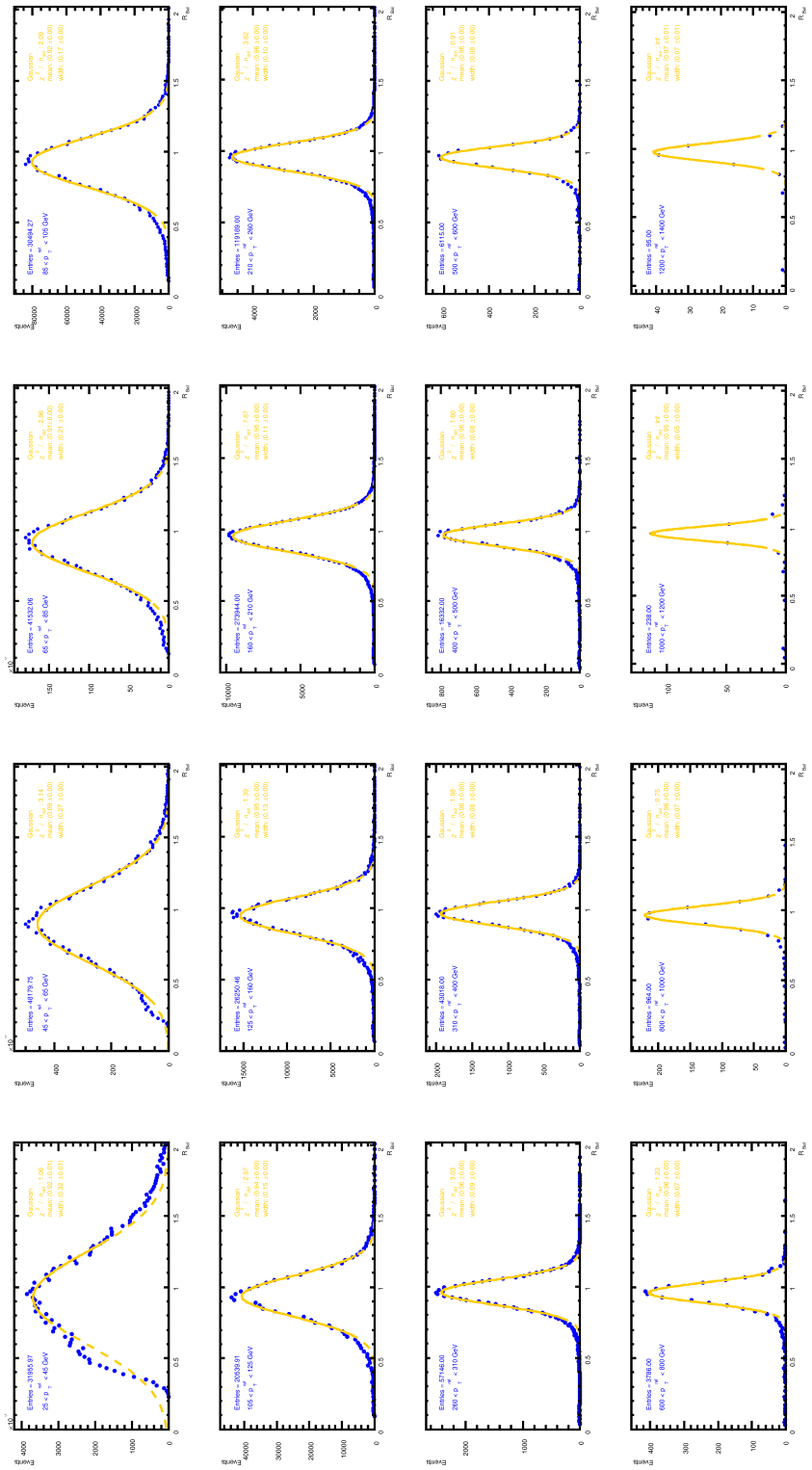


Figure H.4: Distributions of $p_{\mathrm{T}}^{\mathrm{j1}}/p_{\mathrm{T}}^{\mathrm{ref}}$ in 2015-2017 data for different $p_{\mathrm{T}}^{\mathrm{ref}}$ bins using EMtopo jets.

Appendix I

Trigger efficiencies for the photon plus jet analysis using $138\,\mathrm{fb}^{-1}$

The trigger efficiencies as a function of E_{T}^{γ} in the different $|y^{\mathrm{jet}}|$ regions and SS and OS configurations for tight non-isolated, non-tight isolated, and non-tight non-isolated photons are shown from figure I.1 to I.3.

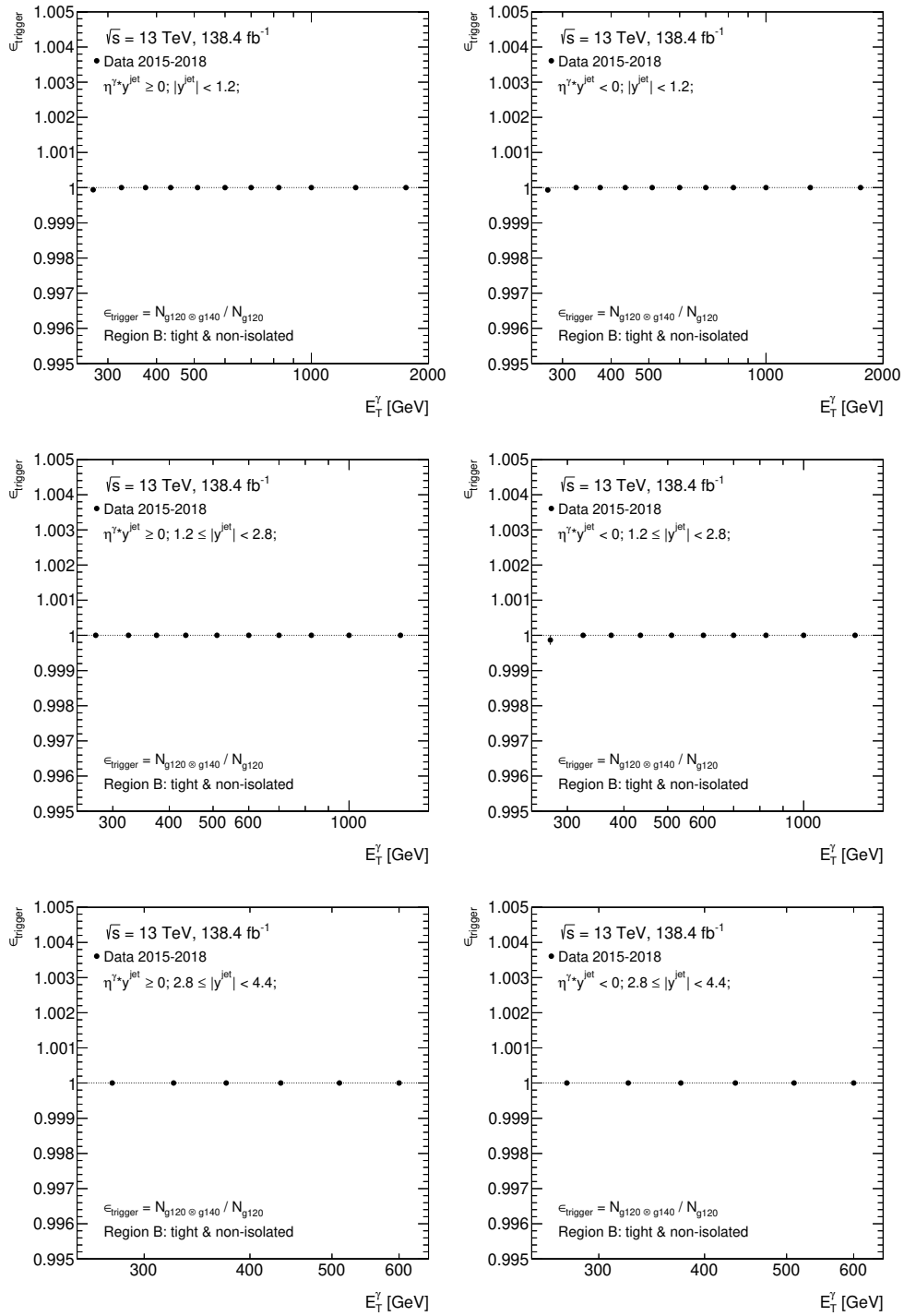


Figure I.1: Trigger efficiency in data as a function of E_{T}^{γ} in the different $|y^{\mathrm{jet}}|$ regions and SS and OS configurations for tight non-isolated photons.

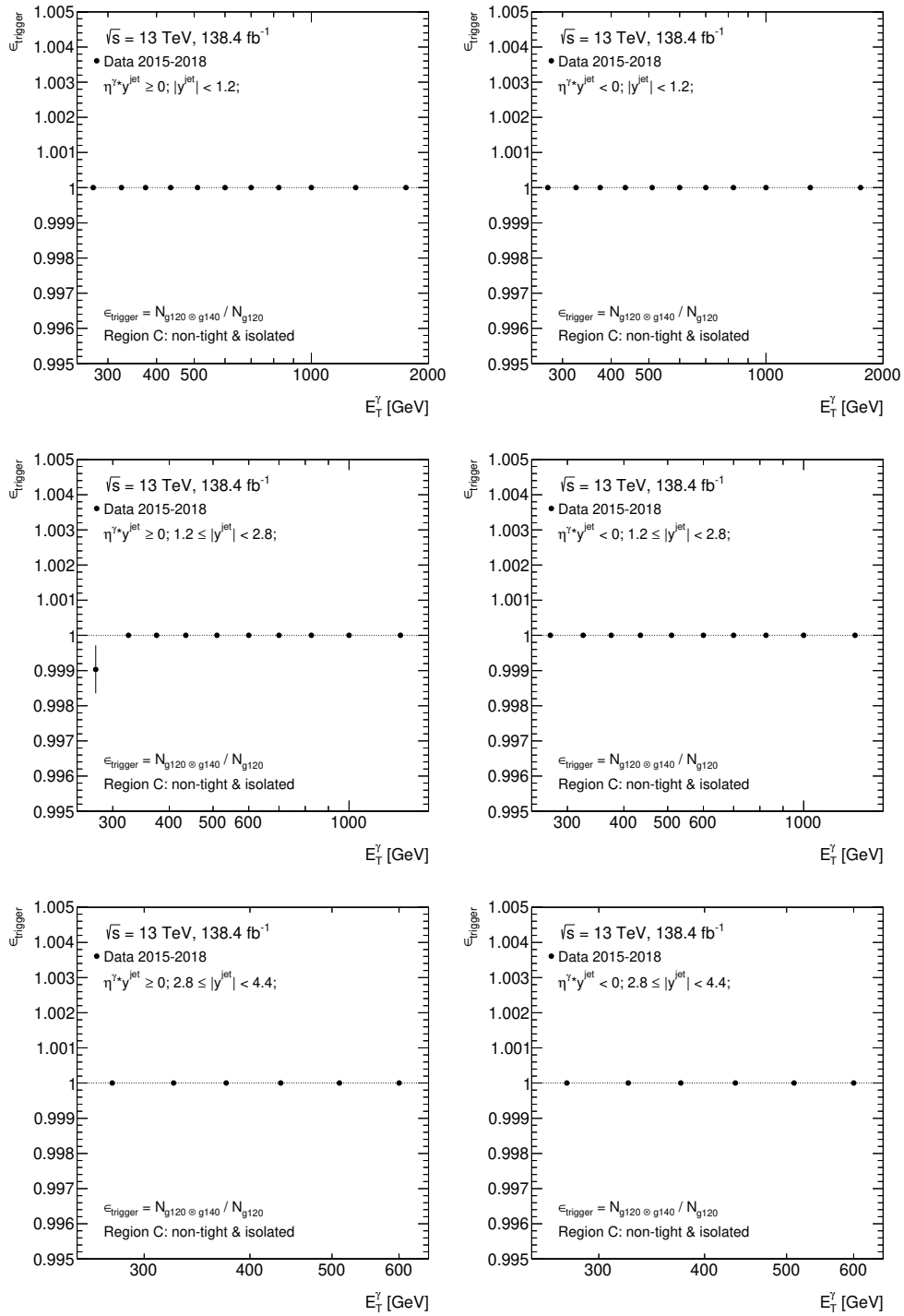


Figure I.2: Trigger efficiency in data as a function of E_{T}^{γ} in the different $|y^{\mathrm{jet}}|$ regions and SS and OS configurations for non-tight isolated photons.

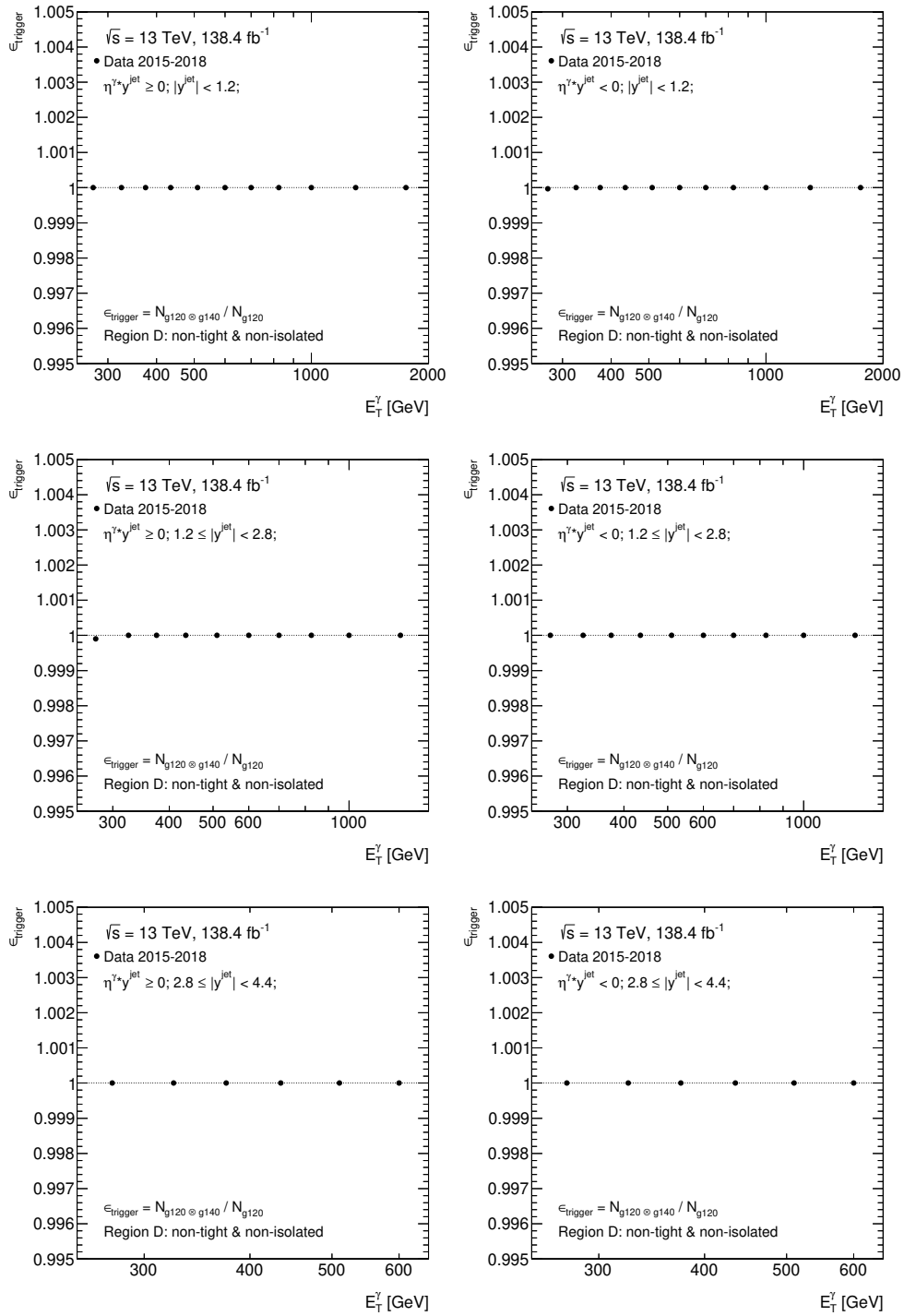


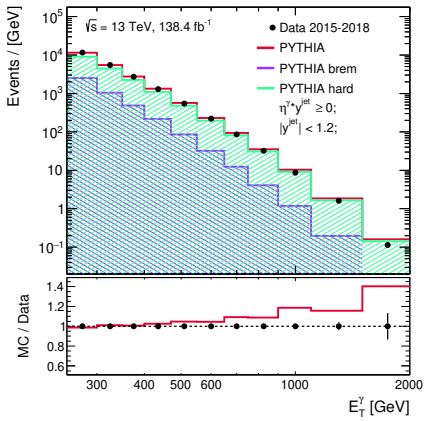
Figure I.3: Trigger efficiency in data as a function of E_{T}^{γ} in the different $|y^{\mathrm{jet}}|$ regions and SS and OS configurations for non-tight non-isolated photons.

Appendix J

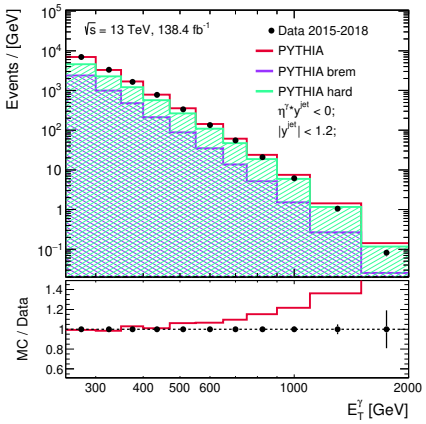
Fraction of brem component in PYTHIA

The estimated signal yields using signal leakage fractions from PYTHIA are shown in figure J.1 for the different measured configurations. The full simulation of PYTHIA, and the brem and hard components are compared to the data for each observable; the simulations are normalised to the data.

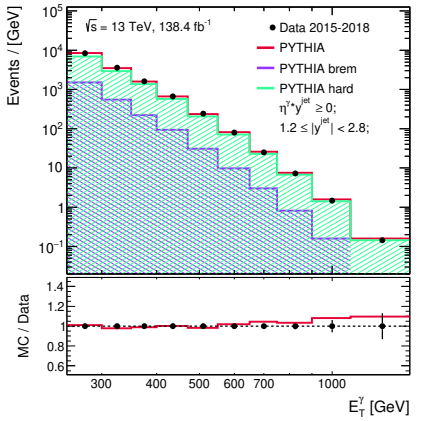
Figure J.2 shows the fraction of brem component in PYTHIA as a function of each observable.



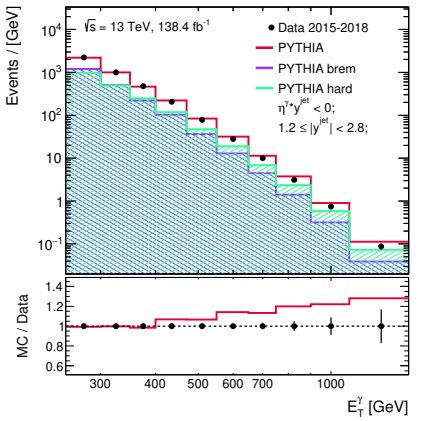
(a)



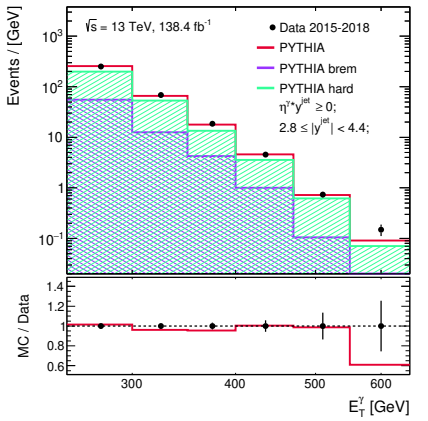
(b)



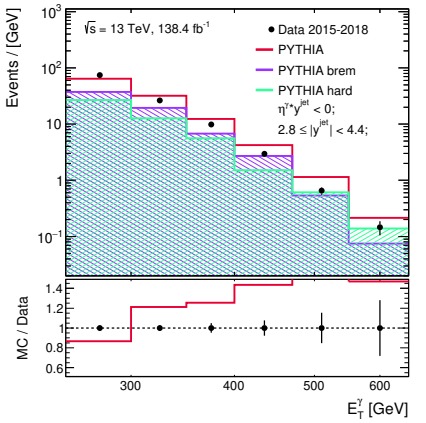
(c)



(d)



(e)



(f)

Figure J.1: The extracted signal yield in data (dots) using signal leakage fractions from PYTHIA as a function of E_{T}^{γ} in the different $|y^{\mathrm{jet}}|$ regions and SS and OS configurations. For comparison, detector level simulations from PYTHIA (red histograms) are included. The hard (green histograms) and brem (purple histograms) components of PYTHIA are also shown and mixed according to the MC prediction. The ratio of the MC distributions to the data is shown in the lower panel of the figures.

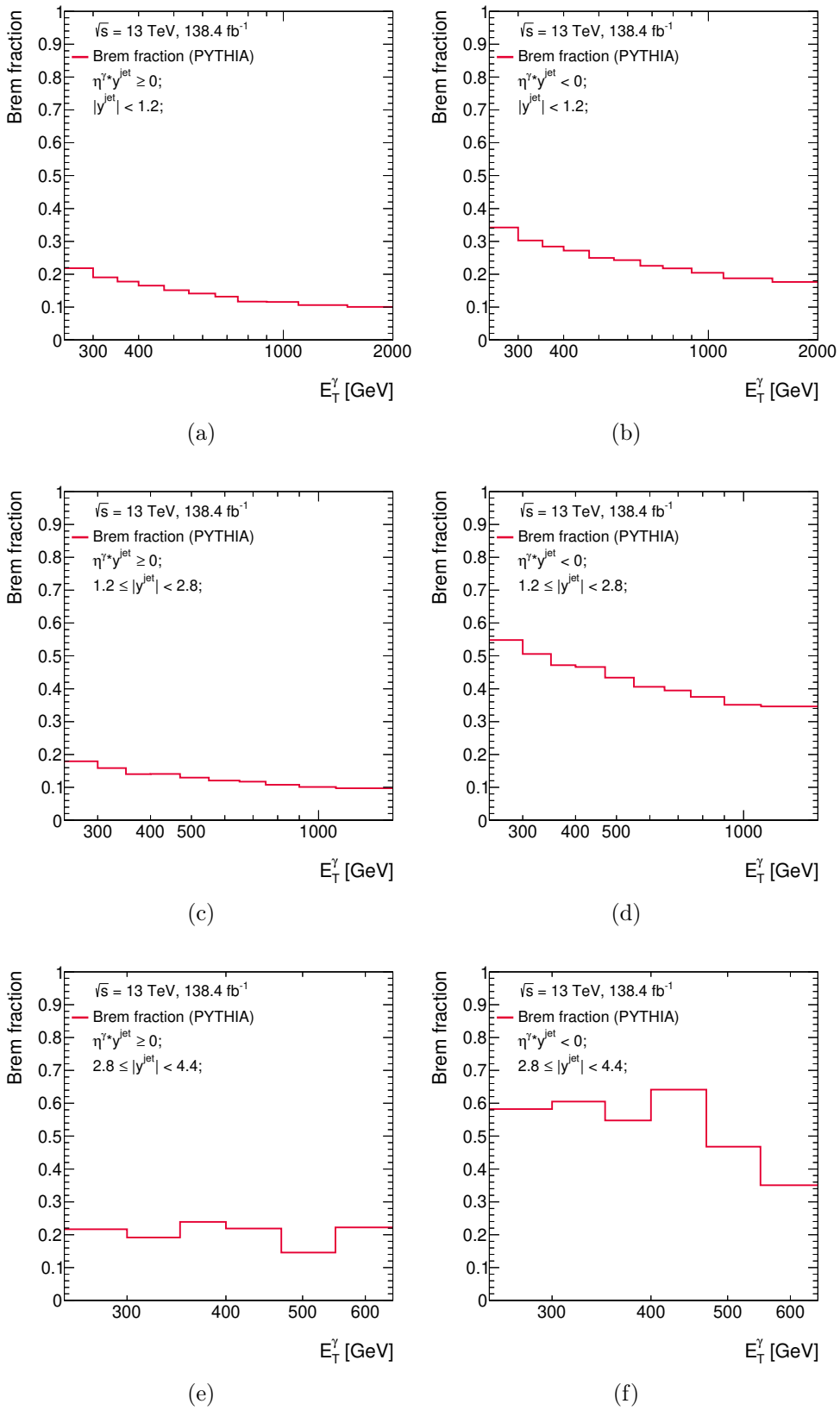


Figure J.2: The fraction of the brem component in PYTHIA (solid histogram) as a function of the different measured observables.

Appendix K

Cross-checks using reweighted MC for the photon plus jet analysis using $138\,\mathrm{fb}^{-1}$

A cross-check was performed by reweighting the MC distributions of SHERPA at detector and particle levels to the data. The ratio of the data to the MC detector-level distributions is shown in figure K.1 for the different measured configurations.

A χ^2 fit to this ratio was performed; the function used in each case is also shown in figure K.1. This function is used to perform a reweighting of the SHERPA sample, both at detector and particle levels.

The cross check was performed by unfolding the default detector level prediction of SHERPA using the reweighted SHERPA sample for the bayes unfolding. The relative difference with respect to the particle level prediction of SHERPA default is shown in figure K.2. The deviations from the default cross section predicted by SHERPA are smaller than 0.1% except for the SS and OS configurations in the most forward $|y^{\mathrm{jet}}|$ region, where they reach up to 0.4%.

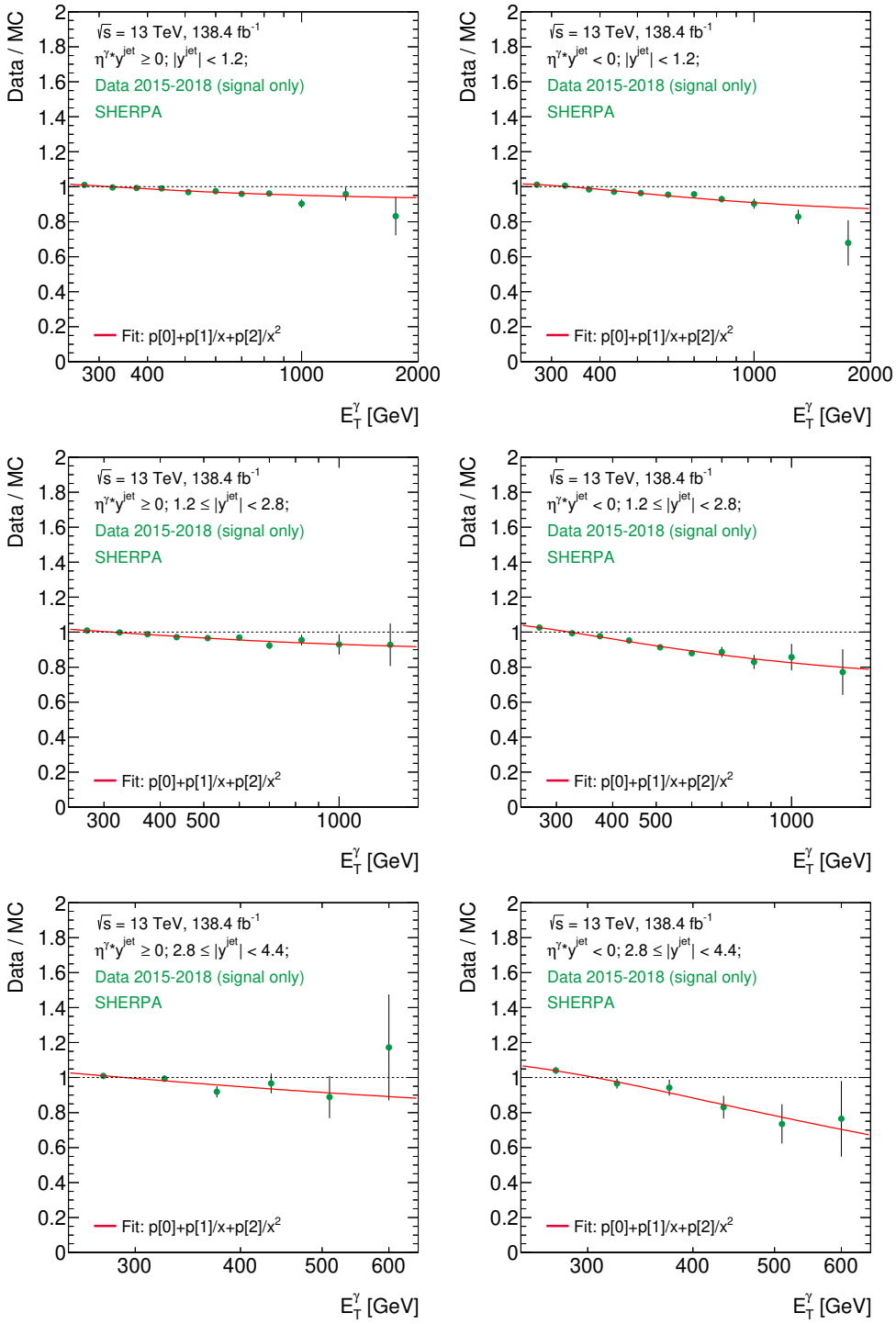


Figure K.1: Ratio of the data and SHERPA distributions (dots) as a function of E_{T}^{γ} in the different measured configurations. The solid line represents the fit to the ratio. The error bars represent the statistical uncertainty only.

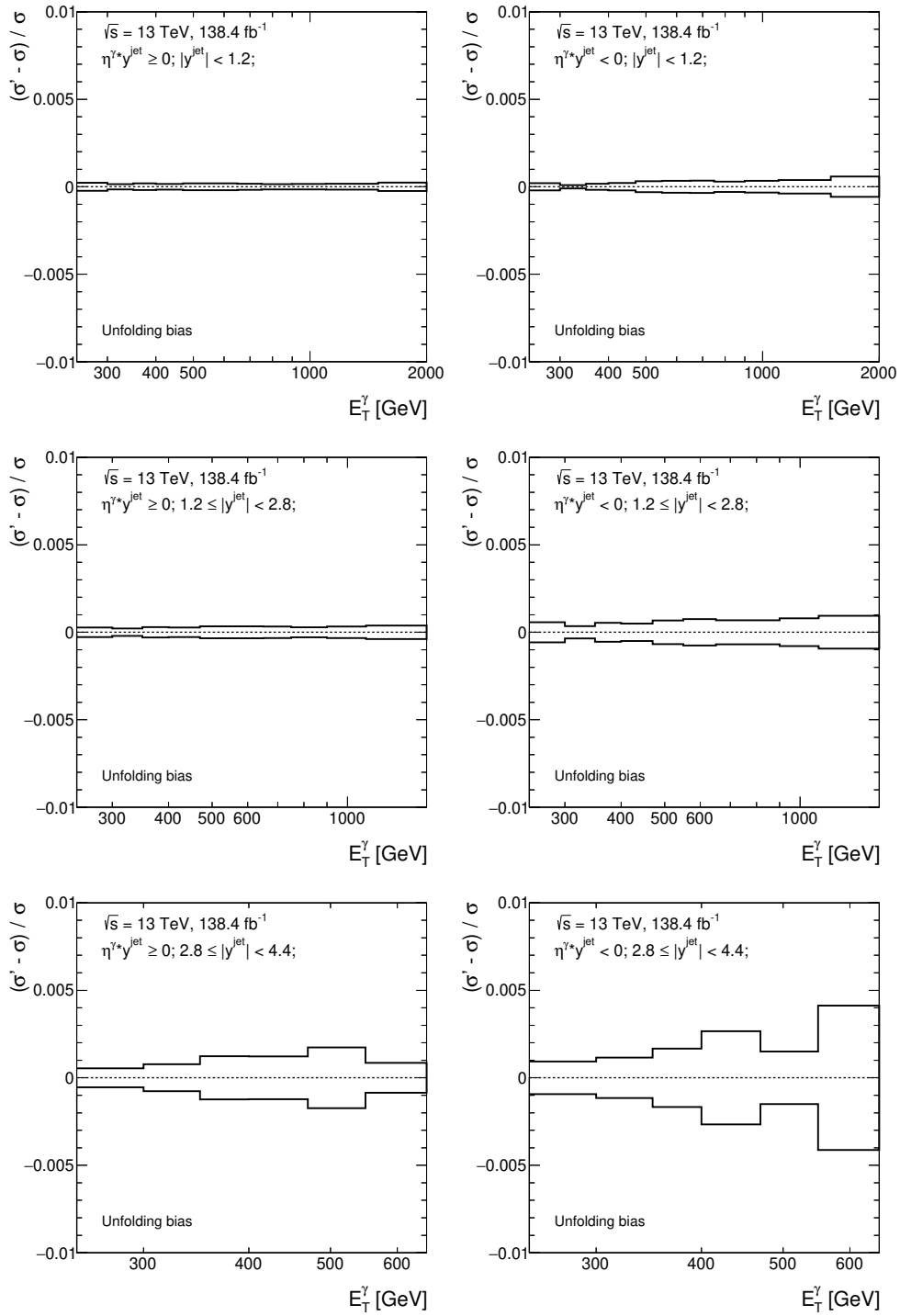


Figure K.2: Relative difference between the default SHERPA detector level prediction unfolded with the reweighted SHERPA and the default SHERPA particle level prediction as a function of E_{T}^{γ} in the different measured configurations.