

Measurement of the Production Cross Section of the Photon + Charm Final State in 8-TeV Proton-Proton Collisions using the ATLAS Detector at the LHC

Robert A Keyes
Department of Physics
McGill University

August 2017



A thesis submitted to McGill University in partial fulfilment of the
requirements of the degree of Ph. D.

Abstract

This thesis presents the measurement of the production cross section of high energy isolated prompt photons in association with charm flavour jets in 20.1 fb^{-1} of proton-proton collisions at a center of mass energy of 8 TeV using the ATLAS detector. Isolated prompt photons are well described theoretically and well measured experimentally, providing unique advantages when probing parton level dynamics and the content of the proton. As such, measurements of this type can be used in global parton distribution fits to better constrain the charm quark content of the proton in the Large Hadron Collider (LHC) environment. The differential cross sections are measured with respect to the photon transverse energy, E_T^γ , in two photon pseudorapidity regions: $|\eta^\gamma| < 1.37$ and $1.56 < |\eta^\gamma| < 2.37$. The measurement spans $25 < E_T^\gamma < 300 \text{ GeV}$ and $25 < E_T^\gamma < 250 \text{ GeV}$ for the two regions, respectively. The ratio of the measured cross sections in the two regions is also reported. This ratio provides enhanced sensitivity through the cancellation of correlated systematic and theoretical uncertainties. The measurement is compared to next-to-leading order theory predictions using a variety of parton distribution functions. The measurement is consistent with each of the predictions within uncertainties.

Résumé

Cette thèse présente la mesure de la section efficace de la production de photons prompts et isolés en association avec un jet de saveur charme avec 20.1 fb^{-1} de collisions proton-proton à une énergie de collision de $\sqrt{s} = 8 \text{ TeV}$ en utilisant le détecteur ATLAS. Les photons prompts et isolés sont bien décrits théoriquement et bien mesurés expérimentalement, fournissant des avantages uniques dans la détermination de la dynamique des partons et du contenu du proton. En tant que tel, des mesures de ce type peuvent être utilisées dans la mesure des fonctions de distribution des partons et permettre de contraindre le contenu du quark charme dans l'environnement fourni par le grand collisionneur de hadrons (LHC). La section efficace est mesurée en fonction de l'énergie transversale du photon, E_T^γ , en deux régions de pseudorapidité des photons: $|\eta^\gamma| < 1.37$ et $1.56 < |\eta^\gamma| < 2.37$. La mesure couvre $25 < E_T^\gamma < 300 \text{ GeV}$ et $25 < E_T^\gamma < 250 \text{ GeV}$ pour les deux régions, respectivement. Le rapport des sections efficaces mesurées dans les deux régions est également mesuré. Ce rapport offre une sensibilité accrue grâce à l'annulation de certaines incertitudes systématiques et théoriques. La mesure est comparée aux prédictions de la chromodynamique quantique perturbative avec corrections de second ordre en utilisant une variété de fonctions de distributions de partons. La mesure est en accord avec chacune des prédictions, considérant les incertitudes.

Acknowledgements

Thanks to Andreas, Mark, Claudia, Josu, and Brigitte for all their help, knowledge, collaboration, and especially for making this project so engaging. Special thanks to Sebastien for being my collaborator and friend throughout these past three years.

This thesis is dedicated to my parents, who lead by example and gave me every opportunity to be the best person I can be.

Statement of Originality

The data analysis and results presented in this thesis are the author's original work. The experiment, recorded data, data processing and computing infrastructure, simulations, calibrations and tools used for the analysis rely on the work of many collaborators (physicists, engineers, computer scientists, technicians) spanning the lifetime of the LHC and ATLAS projects. In particular the software and techniques used to extract and study the signals used for this analysis are developed by the ongoing effort of dedicated working groups. These tools and techniques have been used, adapted and validated by the author to produce this particular result.

The author, Robert Keyes, claims that all aspects of the measurements presented in this thesis, including tables and figures, are produced by the author himself, with the exception of the production of the simulation and the calibration of the JetFitterCharm tagging algorithm. The author wishes to highlight the following personal contributions:

- All studies using the JetFitterCharm tagger to measure the c -jet purity
- Development of combined photon purity and jet flavour extraction method used here as well as in the official ATLAS $\gamma + c$ and $\gamma + b$ measurements
- Unfolding studies for the measurements presented here as well as the official ATLAS $\gamma + c$ and $\gamma + b$ measurements
- Cross-checks on studies performed for the official ATLAS $\gamma + c$ and $\gamma + b$ measurements
- Co-author of the ATLAS $\gamma + c$ and $\gamma + b$ publication, to be submitted, and the supporting internal note. Specifically responsible for the introduction and background, ATLAS detector description, event selection and calibration, background subtraction, and measurement procedure sections.
- More than 120 CPU core years at 2.66 GHz/core dedicated to the completion of this project

The author, Robert Keyes, additionally claims that the following contributions, not contained herein, also constitute original scholarship and an advancement of knowledge:

- Design, construction and commissioning of the McGill ATLAS TGC lab gas mixing and slow control apparatus, described in [1]
- Multi-jet trigger performance studies carried out using the 7 TeV 2011 ATLAS dataset, described in [2]

Contents

Abstract	i
Résumé	ii
Acknowledgements	iii
Statement of Originality	iv
1 Introduction	1
1.1 The Standard Model	2
1.1.1 Building Blocks of Matter	2
1.1.2 Interactions	5
1.1.3 Physics Beyond the Standard Model	13
1.2 The SM at the LHC	13
1.3 Heavy Flavour at the LHC	17
1.3.1 Intrinsic Quarks	18
1.4 The Photon+Charm Measurement	20
1.5 Roadmap	22
2 Experimental Apparatus	24
2.1 Collider Basics	25
2.2 The LHC Injector Chain	26
2.3 The LHC	28
2.4 The ATLAS Detector	30
2.4.1 Coordinates	30
2.4.2 The Inner Detector	31
2.4.3 The Calorimeter	36
2.4.4 The Muon System	41
2.4.5 The Luminosity Measurement System	42

2.5	The ATLAS Trigger System	44
2.5.1	The Level 1 Trigger	46
2.5.2	The Level 2 and Event Filter (High Level Trigger)	46
2.6	ATLAS Data Taking in 2012	47
2.7	Simulation	48
2.7.1	Hard Scatter	49
2.7.2	Parton Shower	50
2.7.3	Hadronization and Decay	51
2.7.4	Pileup	52
2.7.5	Detector Simulation and Reconstruction	53
3	Object Reconstruction	54
3.1	Photon Reconstruction	54
3.1.1	Photon Identification	54
3.1.2	Photon Energy Calibration	57
3.1.3	Photon Isolation	58
3.2	Jet Reconstruction	61
3.2.1	The anti-kT Algorithm	62
3.2.2	Local Cluster Jet Reconstruction	65
3.2.3	Jet Identification	66
3.2.4	Jet Energy Calibration	67
3.2.5	Jet Pile-up Corrections	67
3.2.6	Reconstruction-Level Flavour Tagging	70
4	Data and Simulated Samples	75
4.1	Data Samples	75
4.1.1	Trigger requirements	75
4.2	Monte Carlo Simulations	77
4.2.1	Corrections	78
5	Signal Event Selection	85
5.1	Detector-Level Photon+Jet Selection	85
5.2	Event-Level Criteria	85
5.3	Photon Selection Criteria	86

5.4	Jet selection Criteria	88
5.5	Particle-Level Photon+Jet Selection	98
6	Measurement Procedure	99
6.1	Trigger Efficiency	100
6.2	Photon Purity	101
6.3	Jet Purity	108
6.3.1	Template Fit Closure	110
6.3.2	Goodness-of-Fit	110
6.4	Unfolding	119
7	Uncertainties	126
7.1	Statistical	126
7.2	Systematic	127
7.2.1	JetFitterCharm Efficiency Scale Factor	128
7.2.2	Background Photon Sideband Correlation	132
7.2.3	Prompt Photon Modeling	132
7.2.4	QCD modelling dependence	134
7.2.5	Photon Energy Scale	136
7.2.6	Luminosity	137
7.2.7	Other Uncertainties	137
7.3	Total Uncertainty	140
8	Results	142
8.1	LO Theory Predictions	143
8.2	NLO Theory Predictions	144
8.3	MV1c Measurement	150
9	Conclusions and Outlook	152
9.1	Conclusions	152
9.2	Outlook	153
A	Cut Flows	155
	Bibliography	159

List of Figures

1.1	The Standard model	4
1.2	QCD Feynman diagram	6
1.3	QED Vertex	7
1.4	QED coupling scaling	8
1.5	QCD Vertex	10
1.6	QCD coupling scaling	11
1.7	DGLAP Evolution	15
1.8	SM LHC Cross Sections	17
1.9	Charm PDF shapes and predictions	20
1.10	LO and NLO γ +jet diagrams	21
2.1	ATLAS Coordinates	27
2.2	The LHC	28
2.3	ATLAS Coordinates	31
2.4	ATLAS ID	32
2.5	Superconducting solenoid	33
2.6	ATLAS Pixel Inversion	34
2.7	ATLAS Pixel Detector	35
2.8	ATLAS EMC Accordion Geometry	39
2.9	ATLAS EMC Detector	40
2.10	ATLAS Toroid Magnet	42
2.11	ATLAS Luminosity detectors	43
2.12	ATLAS Trigger	45
2.13	ATLAS L1 Trigger	47
2.15	Simulated Event Components	49
2.16	Hadronization Models	52
3.1	π^0 vs. γ signal	56
3.2	$e\gamma$ calibration	58

3.3	Topo-cluster algorithm	59
3.4	Jet evolution	61
3.5	Jet safety	62
3.6	Jets in ATLAS	62
3.7	kT algorithm with $p=1$	64
3.8	kT algorithm with $p=-1$	65
3.9	Jet Energy Calibration	68
3.10	Jet Vertex Fraction	69
3.11	Heavy Flavour Topology	71
3.12	JetFitterCharm Topology	74
4.1	SHERPA sample stitching	78
4.2	PYTHIA pile-up corrections	81
4.3	SHERPA pile-up corrections	82
4.4	SHERPA prior to the isolation correction	83
4.5	SHERPA following the isolation correction	84
5.1	E_T^γ vs E_T^{iso} in data and SHERPA	87
5.2	Isolation efficiency in SHERPA and PYTHIA	88
5.3	Mean E_T^{iso} vs ΔR in data	90
5.4	Mean E_T^{iso} vs ΔR in SHERPA	90
5.5	Mean E_T^{iso} vs ΔR in PYTHIA	91
5.6	Mean E_T^{iso} vs ΔR in PYTHIA brem	91
5.7	Mean E_T^{iso} vs ΔR in PYTHIA hard	92
5.8	Mean E_T^{iso} vs $\Delta\eta$ in data	92
5.9	Mean E_T^{iso} vs N jets in data	93
5.10	Lead photon E_T^γ vs lead jet p_T^{jet} in data	93
5.11	Lead photon E_T^γ vs lead jet p_T^{jet} in SHERPA	94
5.12	Lead photon E_T^γ vs lead light jet p_T^{jet} in SHERPA	94
5.13	Lead photon E_T^γ vs lead c -jet p_T^{jet} in SHERPA	95
5.14	Lead photon E_T^γ vs lead b -jet p_T^{jet} in SHERPA	95
5.15	Lead photon E_T^γ vs lead jet p_T^{jet} in nominal PYTHIA	96
5.16	Lead photon E_T^γ vs lead jet p_T^{jet} in hard PYTHIA	96
5.17	Lead photon E_T^γ vs lead jet p_T^{jet} in brem PYTHIA	97

5.18	Lead photon E_T^γ vs lead jet p_T^{jet} in optimized PYTHIA	97
6.1	Trigger Efficiencies	100
6.2	ABCD side-band regions	102
6.3	SHERPA signal leakage factors	104
6.4	Photon purity vs E_T^γ	105
6.5	Photon purity vs E_T^γ and the JetFitterCharm discriminant	105
6.6	Photon purity vs JetFitterCharm discriminant	106
6.7	Photon purity vs JetFitterCharm discriminant	107
6.8	JetFitterCharm discriminant definition	108
6.9	Jet flavour template fits, central region, let E_T^γ	111
6.10	Jet flavour template fits, central region, high E_T^γ	112
6.11	Jet flavour template fits, forward region, low E_T^γ	113
6.12	Jet flavour template fits, forward region, high E_T^γ	114
6.13	Measured jet flavour fractions	115
6.14	Flavour fraction closure, SHERPA measuring PYTHIA	115
6.15	Flavour fraction closure, PYTHIA measuring SHERPA	116
6.16	Relative flavour fraction closure, SHERPA measuring PYTHIA	116
6.17	Relative flavour fraction closure, PYTHIA measuring SHERPA	117
6.18	Template fit pull mean and width, central region	117
6.19	Template fit pull mean and width, forward region	118
6.20	SHERPA response matrix	120
6.21	SHERPA reconstruction efficiency	120
6.22	Photon resolution in SHERPA	121
6.23	Bayesian unfolding convergence in SHERPA	122
6.24	Unfolding factors in SHERPA	123
6.25	Relative difference in unfolding factors in SHERPA	123
6.26	Unfolding factors obtained using SHERPA and PYTHIA	124
6.27	SHERPA unfolding PYTHIA	124
6.28	PYTHIA unfolding SHERPA	125
7.1	Data statistical uncertainty	127
7.2	Simulation statistical uncertainty	127
7.3	Light jet scale factor uncertainty	129

7.4	c -jet scale factor uncertainty	130
7.5	b -jet scale factor uncertainty	130
7.6	b -jet scale factor extrapolation uncertainty	131
7.7	c -jet scale factor extrapolation uncertainty	131
7.8	Light jet scale factor extrapolation uncertainty	132
7.9	R_{bkg} stability	133
7.10	R_{bkg} uncertainty	133
7.11	Nominal hard/brem mixture in PYTHIA	134
7.12	χ^2 of hard/brem mixture in PYTHIA	135
7.13	Optimized hard/brem mixture in PYTHIA	135
7.14	Prompt photon modeling uncertainty	136
7.15	Hadronization modeling uncertainty	136
7.16	Photon energy scale uncertainty	137
7.17	Remaining uncertainties combined	138
7.18	Total uncertainty on the measured cross section	140
7.19	Total uncertainty on the measured ratio	141
8.1	Measured cross section compared to PYTHIA and SHERPA	144
8.2	Measured ratio compared to PYTHIA and SHERPA	144
8.3	Measured cross section compared to NLO using CT14 BHPS PDFs	146
8.4	Measured ratio compared to NLO using CT14 BHPS PDFs	147
8.5	Measured cross section compared to NLO using CT14 SEA PDFs	147
8.6	Measured ratio compared to NLO using CT14 SEA PDFs	148
8.7	Measured cross section compared to NLO using NNPDF PDFs	148
8.8	Measured ratio compared to NLO using NNPDF PDFs	149
8.9	Measured cross section compared to the MV1c measurement	151

List of Tables

2.1	Design, 2012 and 2015 operational parameters for the LHC.	30
3.1	Photon identification shower shape variables	60
3.2	JetFitterCharm calibrated working points	73
4.1	Number of events and integrated luminosity by data period.	76
4.2	Average prescales for single photon E_T triggers used throughout 2012 data taking.	76
4.3	Number of events and generation parameters by simulation slice. .	79
6.1	Trigger efficiencies	101
A.1	Reconstruction-level cut flow for data	156
A.2	Reconstruction-level cut flow for SHERPA	157
A.3	Particle-level cut flow for SHERPA	158

Introduction

The field of high energy particle physics endeavors to reveal the fundamental constituents of matter and the interactions that govern them. Particle accelerators are a primary tool employed in this pursuit, and can be thought of in analogy to a simple microscope. In a microscope light scatters off an object and is subsequently detected by the observer's eye. The observer then makes sense of the structure of the scattered light to infer the nature of the object. In a similar fashion, the outgoing particles arising from high energy collisions can reveal the structure of the colliding particles and the nature of their interactions. When accelerated particles collide, their kinetic energy becomes localized, creating for a brief instant a hot and dense state of matter. These states of matter mimic what might have existed moments following the big bang, providing a window into the origin of our universe, pushing the boundary of our understanding of the matter and forces that underlie our existence. The Large Hadron Collider (LHC) at the European Organization for Nuclear Research (CERN) is the largest and most powerful particle accelerator ever created. The LHC program largely comprises proton collisions, that is collisions of composite particles governed by the strong force that, along with the neutron, make up the nuclei of all the atomic matter that surrounds us. Making accurate and precise measurements using LHC collisions hinges on our understanding of the strong force and on the structure of the proton. This thesis focuses on measuring the production of photons in association with charm flavoured jets. The photon component of this production mechanism provides a theoretically well-described and experimentally well-measured signal that is used as a lever arm to probe the charm quark interaction under the strong force as well as the charm quark content of the proton. Protons have never before been examined in this way in this state, and as such this measurement provides an important test of and feedback into our knowledge of the strong force and the

structure of the proton at the energy frontier.

1.1 The Standard Model

The Standard Model (SM) of particle physics [3–6] is the state of the art of our understanding of the physics that governs the smallest constituents of matter and their interactions, that is interactions on the scale of $< 10^{-3}$ of the radius of a nucleus. Many aspects of the theory have been precisely predicted and tested; for instance quantum electrodynamic interactions have been tested to be accurate up to a precision of 76 parts per trillion in the measurement of $g-2$ [7], and one of the great successes of the theory is the prediction [8–10] and ensuing discovery [11,12] of the Higgs boson. Despite the SM's triumphs it remains incomplete, as it fails to incorporate and describe the nature of some well-established phenomena, such as dark matter and gravity, and established parts of it pose fundamental theoretical and technical challenges when deriving predictions, such as the non-perturbative nature of the strong force and interactions with many bodies.

The SM uses the framework of relativistic quantum field theory (QFT), necessitated by the fact that fundamental particles are very small (hence the quantum formalism) and that for the interactions of interest the particles move very fast (hence the relativistic formalism). In QFT, fields take the place of coordinates, which are used in classical mechanics, and are promoted to operators, as is done for observables in quantum mechanics.

1.1.1 Building Blocks of Matter

In quantum mechanics the Schrodinger equation, analogue to the non-relativistic kinetic energy relation, is used to describe the dynamics of non-relativistic quantum systems:

$$E = \frac{p^2}{2m} \rightarrow i\hbar \frac{\delta}{\delta t} \psi = -\frac{\hbar^2 \nabla^2}{2m} \psi, \quad (1.1)$$

where E is energy, p is momentum, m is mass, $\hbar$ is the reduced Planck constant, ψ is the quantum mechanical wavefunction and Δ is the spatial gradient operator.

The relativistic cousin of this equation is the Klein-Gordon equation, analogue to the relativistic energy relation. Since it is second order in space and time

this equation allows for negative energy solutions presenting the problem that this permits negative probability densities

$$\left(\frac{E}{c}\right)^2 = p^2 + m^2 c^2 \rightarrow \left(\frac{\hbar}{c} \frac{\delta}{\delta t}\right)^2 \phi = (\hbar^2 \nabla^2 - m^2 c^2) \phi, \quad (1.2)$$

where c is the speed of light and ϕ is the relativistic wavefunction.

In the 1920s Dirac decomposed the Klein-Gordon equation using matrix algebra, yielding an equation that is first order in space and time and thus whose solution satisfies the Klein-Gordon equation as well as yielding positive definite probability densities [13]. In this equation, the Dirac equation, particles are represented by 2 or 4 component objects called spinors, the $(1/2, 0) + (0, 1/2)$ representation of the Lorentz group, providing a theoretical motivation for Pauli's phenomenological theory of spin. The Dirac equation still permits negative energy solutions, predicting the existence of anti-particles prior to the discovery of the positron in 1933 [14]. Here the equation is presented using vector and matrix notation, where ψ is the Dirac spinor, where " $\rightarrow$ " represents a spatial vector, and where σ^i are the Pauli matrices:

$$i\hbar \frac{\partial \psi}{\partial t} = [c \vec{\alpha} \cdot \vec{p} + mc^2 \beta] \psi \quad (1.3)$$

$$\alpha^i = \begin{pmatrix} 0 & \sigma^i \\ \sigma^i & 0 \end{pmatrix}, \quad \beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (1.4)$$

In the SM all matter besides the force carriers is made up of fermions, that is elementary point-like spin 1/2 particles that satisfy the Dirac equation. These particles, displayed in figure 1.1, fall into two sectors: leptons and quarks. Each sector contains three families, comprising two particles, that are ordered according to their mass. In the lepton sector the lightest family includes the well-known electron and its cousin the electron-neutrino, the middle range family includes the muon and its cousin the muon neutrino, and the heaviest family includes the tau and its cousin the tau neutrino. The neutrinos carry no electromagnetic (EM) charge and a lepton number of +1, while their partners carry an EM charge of -1 and a lepton number of +1. In the quark sector the lightest family includes the up

and the down quarks, the building blocks of the proton and neutron and all the matter we witness in our daily lives. The middle range family includes the charm and the strange quarks, and the heaviest family includes the bottom and the top quarks. The quarks in the top row carry an EM charge of $+2/3$, while the bottom row carry $-1/3$.

The right-hand column in red of figure 1.1 displays the spin-1 force carriers: the gluon (strong force), the photon (electromagnetic force), and the Z and W (weak force) [15]. The right-most particle in yellow is the recently discovered Higgs boson, which is a massive spin-0 particle and carries no charge.

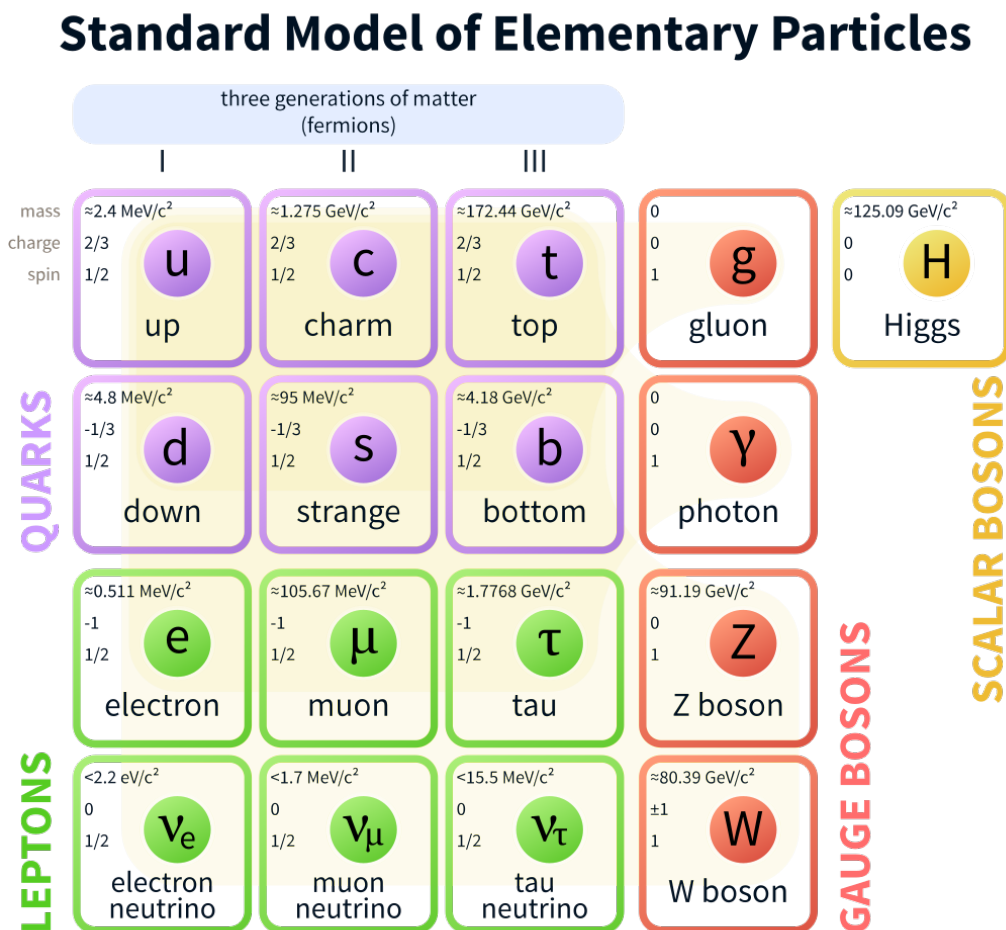


Figure 1.1 – The particles and force carriers of the Standard Model [16].

1.1.2 Interactions

The interactions in the SM are built using gauge fields, that is fields whose Lagrangian is born from requiring invariance under local gauge transformations of a symmetry group. This naturally provides the framework for the interactions since by requiring the invariance to be local (i.e. a function of spacetime) the interactions satisfy relativistic dynamics, and since for every continuous symmetry of the field there is a conserved current and charge (following Noether's theorem) providing the mechanism for describing the interactions. Different symmetries give rise to the different interaction terms that represent the forces in the SM Lagrangian. In order for the interactions of the SM to be consistent with relativity they must satisfy Lorentz invariance. This requirement, combined with unitarity, imposes charge parity time-reversal (CPT) invariance on all the interactions of the SM [17].

Fields themselves have quantum numbers, and particles are described as excited states of the field that come in and out of existence via creation and annihilation operators. Interactions are described in QFT by calculating the scattering matrix that yields the transition probability relating an initial state of particles with a final state particles. Mathematically this scattering matrix is a path integral sum over all possible interactions that connect the initial and final states. Feynman diagrams are a visual representation of the mathematics of QFT that define each of these possible interactions. Particles and force carriers are drawn as lines and interactions are represented by a junction between a force carrier and a particle. Lines internal to the diagram are referred to as propagators and junctions between lines are referred to as vertices, as shown in figure 1.2.

Each vertex carries a factor of the coupling constant for the interaction, which is why if the coupling is small the lowest order diagrams (diagrams with the fewest vertices) provide the largest contributions, and including higher order terms converges towards more accurate results.

In QFT, fields can be thought of as a fabric of quantum harmonic oscillators. This analogy highlights the idea of renormalization. Since the quantum harmonic oscillator has a non-zero ground state energy, if you had an infinite field there would be infinite energy in the ground state. Infinities often occur in QFT calculations, such as when taking into account loop diagrams in scattering mat-

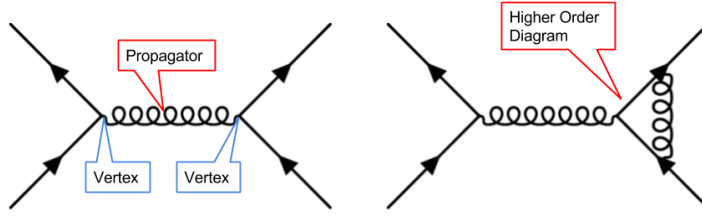


Figure 1.2 – A Leading-Order and a Next-to-Leading-Order diagram for a QCD quark scattering process.

rix calculations, but the observables that are measured must remain finite. These cases are dealt with using renormalization, a mathematical transformation which amounts to a change of scale that has no impact on the measurable observables. In the example of the ground state energy of a field of harmonic oscillators it is quite simple, by shifting the Hamiltonian by a constant, none of the dynamics of the field interactions changes but the Hamiltonian over all space becomes finite.

There are three fundamental interactions in the SM: the strong force, the weak force and the electromagnetic (EM) force (also referred to as Quantum Electrodynamics or QED). These interactions follow from imposing $SU(3) \times SU(2) \times U(1)$ gauge symmetries that give rise to corresponding gauge fields and spin-1 force carriers.

The Electromagnetic Force

The EM force (QED) follows from the $U(1)_{em}$ symmetry, for which there is one generator that gives rise to the single force carrier, the photon. The photon is a massless particle, which gives the EM force infinite range and two polarization states due to its spin-1 nature. The photon couples to EM charge with a coupling strength of roughly $\alpha_{EM} \approx 1/137$, which permits QED processes to be described with great accuracy using perturbation theory. An example EM vertex is shown in figure 1.3.

Higher order computations rely on the application of renormalization to sidestep the interaction of EM charge with the EM field. Feynman, Schwinger and Tomonaga discovered that this could be achieved by expressing observables

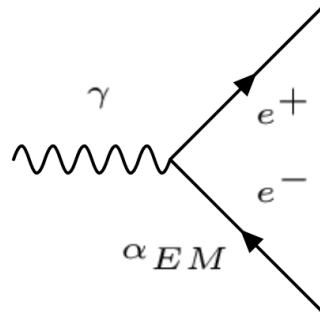


Figure 1.3 – Photons couple at a vertex to particles carrying electric charge. The strength of the coupling of a vertex is given by the coupling constant, in this case α_{EM} .

in terms of measurable quantities such as the experimentally measured mass and charge of the electron [18]. In doing so the bare electron mass, which is unmeasurable due to vacuum polarization screening, is adjusted and the theory yields finite measurable results. This adjustment relies on the introduction of a renormalization scale, which also introduces a scale dependence of the QED coupling. The following displays this scale dependence, where μ is the renormalization scale and Q is the momentum transfer of the interaction:

$$\alpha(Q^2) = \frac{\alpha(\mu^2)}{1 - (\alpha(\mu^2)/3\pi)\log(Q^2/\mu^2)}. \quad (1.5)$$

Conceptually this means that the EM coupling increases with increasing momentum transfer, Q , since higher energy interactions pierce deeper beyond the screening effect, depicted in figure 1.4.

Interactions of the EM force obey both charge conjugation (C) and parity (P) symmetries. This means that the EM force is the same between two particles if they are both replaced by their anti-particles, and that the EM force is the same between two particles if their spatial coordinate system is mirrored, since the force depends only on the distance between the particles and not on the handedness of the coordinate system.

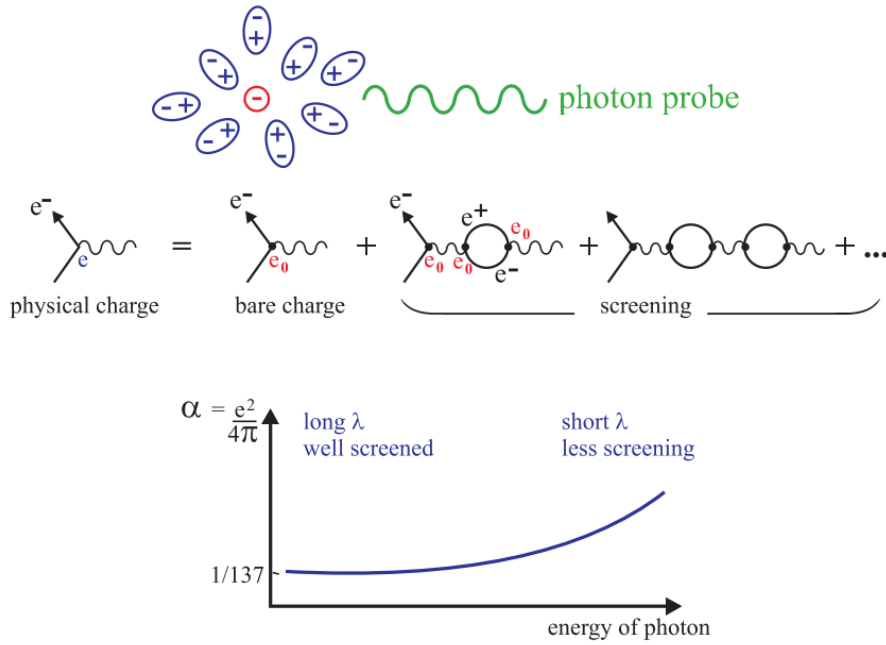


Figure 1.4 – The renormalization of QED imposes a scale dependent running of the coupling that increases with energy [19].

The Weak Force

The weak force follows from the $SU(2)$ symmetry, for which there are three generators giving rise to the massive $W^\pm$ ($m_W = 80 \text{ GeV}/c^2$) and Z^0 ($m_Z = 91 \text{ GeV}/c^2$) bosons. The weak force is the only interaction in the SM that couples to all the leptons and quarks. The charged $W^\pm$ bosons are responsible for the so-called charged current interactions, interaction vertices that can change a particle to another in its generation while conserving charge. These interactions are responsible for β decay radioactivity (thus a neutron decays to a proton). The massive nature of these bosons greatly constrains the range of the force and results in each boson exhibiting three polarization states. The limited range of the force can be understood by applying dimensional arguments to the uncertainty principle with

Einstein's equation for the rest mass energy of a particle:

$$\begin{aligned}
 \Delta x &= c\Delta t, \\
 \Delta x &= c\frac{\hbar}{\Delta E}, \\
 \Delta x &= c\frac{\hbar}{mc^2}, \\
 \Delta x &= \frac{\hbar}{80 \text{ GeV}/c} = 10^{-3} \text{ fm}.
 \end{aligned} \tag{1.6}$$

The weak force, unlike the EM and the strong force, does not obey C and P symmetry individually; however it mostly obeys CP symmetry and always obeys the product of CP and time reversal symmetry (CPT). Direct charge conjugation and parity symmetry violation results from the fact that the charged current $W^\pm$ boson interaction is helicity dependent. Helicity is the projection of a particle's spin onto its direction of motion. Right-handed helicity refers to particles where the projection is positive and left-handed refers to particles where the projection is negative. The W boson's axial vector coupling only couples to left-handed particles and right-handed anti-particles.

The W boson mediates interactions between leptons of the same generation; however, for quarks there is also coupling across generations. It is for this reason that the only stable particles we observe in nature in the quark sector comprise up and down quarks, the members of the lightest generation. The inter-generational coupling is characterized by the Cabibbo-Kobayashi-Maskawa (CKM) matrix, a matrix whose individual elements represent the probability of given quark family (d , s , or b) transitioning to another [?,?]:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{1.7}$$

The matrix can be fully described using three angles (θ_{12} , θ_{13} and θ_{23}) and a single phase (δ) that provides the mechanism for direct CP violation. These angles are

represented using the notation $c_{ij} = \cos(\theta_{ij})$ and $s_{ij} = \sin(\theta_{ij})$:

$$V_{ij} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & s_{23}s_{13} \end{pmatrix}. \quad (1.8)$$

The absolute values of the complex components of the CKM matrix are [20]:

$$|V_{ij}| = \begin{pmatrix} 0.97434 & 0.22506 & 0.00357 \\ 0.22492 & 0.97351 & 0.0411 \\ 0.00875 & 0.0403 & 0.99915 \end{pmatrix} \quad (1.9)$$

The Strong Force

The strong force, also known as Quantum Chromodynamics (QCD), follows from the $SU(3)$ symmetry, for which there are eight generators giving rise to eight massless force carriers, referred to as gluons. Gluons, akin to the photon, have two polarization states due to their massless nature. Gluons are unique in that they both couple to and carry the charge of the strong interaction, referred to as colour. Interaction vertices for the strong force are shown in figure 1.5, displaying that the gluon in fact carries two colours and changes the colour of the parton it interacts with, conserving the colour charge at the vertex.

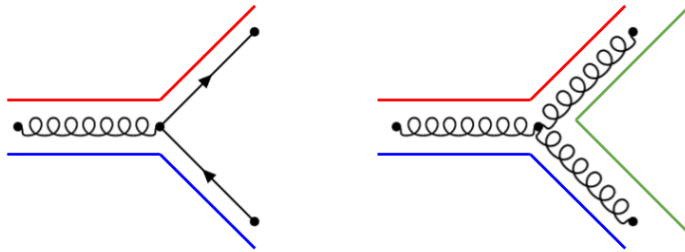


Figure 1.5 – Gluons couple to particles carrying colour charge, including other gluons since they carry two colours.

Gluon self-coupling is a crucial feature of QCD. When QCD is renormalized, in a similar fashion to QED, the presence of the self-interaction of gluons (the

triple-gluon vertex) drastically changes the running of the coupling constant, depicted in figure 1.6. The strong force coupling constant (α_s) running is described in the following equation, where μ is the renormalization scale, Q is the energy scale of the interaction, and n_f is the number of active quark flavours (those with $m_q < Q$):

$$\alpha_s(Q^2) = \frac{\alpha_s(\mu^2)}{1 + b_0 \alpha_s(\mu^2) \log(Q^2/\mu^2)}, \quad (1.10)$$

where

$$b_0 = -\frac{n_f}{6\pi} + \frac{33}{12\pi}. \quad (1.11)$$

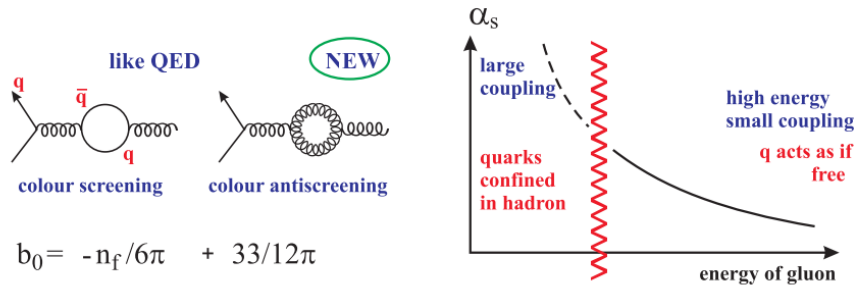


Figure 1.6 – The renormalization of QCD imposes a scale dependent running of the coupling that decreases with energy [19].

These self-interaction effects introduce an anti-colour screening resulting in a force potential that increases linearly with distance between charges. When colour particles are separated it quickly becomes energetically favourable to produce coloured particle anti-particle pairs from the vacuum to cancel out the charge. This makes it impossible to observe "bare" colour charge, a concept referred to as "confinement". Confinement implies that only colour neutral combinations of particles can be observed in nature. These colour neutral composite particles are referred to as hadrons. When interactions produce single coloured particles they evolve into hadrons on very short timescales, a process referred to as "hadronization".

The shape of the QCD potential results in a weakening of the force when colour charges are close together, implying that very close particles behave as free particles. This feature is referred to as asymptotic freedom. As a result hard interactions, which bring two coloured particles very close together, can be described

with perturbation theory using techniques analogous to those used for QED. Challenges arise in QCD when considering longer-distance non-perturbative interactions, especially when bridging the transition between the shorter-distance and longer-distance regimes, such as the subsequent hadronization of the particles involved in hard interactions.

An apt analogy for these two behaviors is to consider the behavior of an elastic. When an elastic is contracted it exerts no force and is very flimsy, it can be thought of as being asymptotically free. When it is stretched it becomes firm and exerts a force, where over larger distances the force increases. When the distance becomes too great the elastic snaps forming two separate portions that are under less tension and thus in a lower energy state.

QCD colour charge comes in three types: red, green and blue. For each type of colour charge there is an associated anti-charge. Gluons carry two colour charges while quarks carry one. The fact that there are three colours of the QCD charge is what allows quarks of otherwise identical quantum numbers to occupy the same state, such as in the $\Delta^{++}(uuu)$ or $\Omega^{-}(sss)$ baryons. With each of the quarks in these baryons carrying a different colour they are in fact distinguishable spin 1/2 fermions and can thus satisfy the Pauli exclusion principle. Similarly to the EM force, the strong force obeys both C and P symmetries.

Electroweak Unification and the Higgs Field

Similarly to how Maxwell constructed a theory that unified electricity and magnetism [21], in the 1960s Glashow, Weinberg and Salam constructed a unified theory of the electromagnetic and weak forces, laying the foundation for the SM and leading to the prediction and subsequent discovery of the W, Z and Higgs bosons [3–5,22]. The Weinberg Salam model for the electroweak interaction derives from imposing local gauge invariance on the $SU(2) \otimes U(1)$ symmetry group. This symmetry group gives rise to three massless gauge bosons for the $SU(2)$ symmetry and a single massless gauge boson for the $U(1)$ symmetry. Adding a real scalar $SU(2)$ doublet Higgs field with a non-zero vacuum expectation value (VEV) breaks the symmetry of the Lagrangian, causing three of the Higgs field components to mix with three of the generators from the $SU(2) \times U(1)$ group giving them mass. The remaining fourth degree of freedom of the Higgs field forms the scalar Higgs boson

field (if it were not a scalar the VEV would violate Lorentz invariance) [23]. This mechanism gives rise to the W and the Z bosons, the massless photon (γ), and the massive Higgs boson. The charged leptons gain their masses via the Yukawa coupling mechanism that preserves their chiral symmetry [24].

1.1.3 Physics Beyond the Standard Model

With the successful inclusion of the Higgs boson as the mass generating mechanism of the SM, unanswered questions that fall beyond the scope of the SM remain. Is there a mechanism behind the mass hierarchy of the constituents of the standard model [25]? Are these particles fundamental, or do they exhibit substructure at higher energies? How can neutrino masses be included since neutrino flavour mixing tells us that they are massive [26, 27]? Do protons decay? Can conservation symmetries be violated, and, related to this question, what is the origin of the matter anti-matter asymmetry in the universe? What is the origin of the dark matter that appears to fill our universe [28]? Can the four forces of nature be unified into a single formalism, similarly to electricity and magnetism, and subsequently the electromagnetic and the weak forces [29, 30]? There are many proposed models and extensions to the SM that suggest mechanisms to answer these questions. Currently the most frequently explored theory is supersymmetry (SUSY) [31]. SUSY postulates an extension of the SM where every spin-1/2 particle has a more massive spin-0 "super"-partner, a symmetry between force carriers and the particles of the SM that is broken, resulting in their different masses. This symmetry breaking can simultaneously explain the particle mass hierarchy, provide a unification scheme for the fundamental forces, explain the Higgs mass, and provide suitable candidates for dark matter. SUSY, along with many other proposed beyond-the-SM phenomena, is being searched for at the LHC.

1.2 The SM at the LHC

The formalism used for describing the interactions of the SM in experimental settings is referred to as scattering matrix theory, or Fermi's Golden Rule. For a scattering experiment like the LHC, this rule is used to calculate interaction

cross sections. Interaction cross sections, represented by σ , tell us how many outgoing particles will arise per unit time from the interaction in question, given the flux of incident particles per unit time [32]:

$$\sigma_{tot} = \frac{\text{reactions/s}}{(\text{beam particles/s}) \times (\text{scattering centers/area})}. \quad (1.12)$$

This quantity has units of area and is typically quoted in units of "barns" (10^{-28}m^2). Calculating a cross section for a given process involves two ingredients, the Matrix Element (ME, also referred to as the probability amplitude) and the available phase space (also referred to as the density of states). The ME contains the dynamics of the interaction and is computed using perturbative QFT for a given set of initial and final state particles. The ME is the QFT analogue to the transition matrix between states in time-dependent perturbation theory. The general procedure for computing the ME in QFT involves summing all feasible Feynman diagrams given the initial state configuration, order by order, to obtain the most precise calculation possible of the final states. The cross section is then obtained by integrating over the phase space of incoming and outgoing particles, accounting for combinatorial effects such as double counting indistinguishable states.

The fact that the LHC collides composite particles, protons and heavy nuclei, introduces an additional level of complexity into cross section calculations. Protons comprise three tightly bound on-shell quarks in a colour neutral state, surrounded by a sea of soft virtual QCD particles. When protons collide, collisions occur not only between the on-shell quarks but also with this virtual non-perturbative sea. Because of the running of the QCD coupling with the energy scale of the interaction (Q), the makeup of this coloured sea, and in turn the proton, looks different at different Q . Many of the interactions of interest at the LHC occur at high energies, probing the energy frontier, however, these interactions are born from proton constituents whose dynamics are governed by low energy QCD. This necessitates the factorization of the QCD interaction into the separate perturbative (high energy, hard) and non-perturbative (low energy, soft) components

$$d\sigma = \sum_{i,j} d\sigma_{ij}^{part} \otimes f_i(x_a) \otimes f_j(x_b). \quad (1.13)$$

The non-perturbative structure of the proton is accounted for using parton distribution functions (PDFs) which encapsulate the description of the momentum fraction (x) carried by the real and virtual partons (quarks and gluons) that make up the colliding protons. Presently it is beyond the scope of QCD calculations to derive PDF distributions from first principles. PDFs are obtained by performing global fits to the vast body of proton scattering measurements, which in large part were yielded by the HERA deep inelastic scattering experiments [33–41]. Each measurement included in a global PDF fit probes a Q specific to the experiment yielding that measurement. The procedure for performing the global fit and for deriving distributions from the fit, however, relies on the robust computation of the dynamical evolution of QCD. The QCD evolution equations are known as the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations [42, 43], and are used to translate PDF distributions between different interaction energies [44]. In summary these differential equations are used to evolve an interaction from one energy to another, where the kernel of the differential equation is computed to a fixed order.

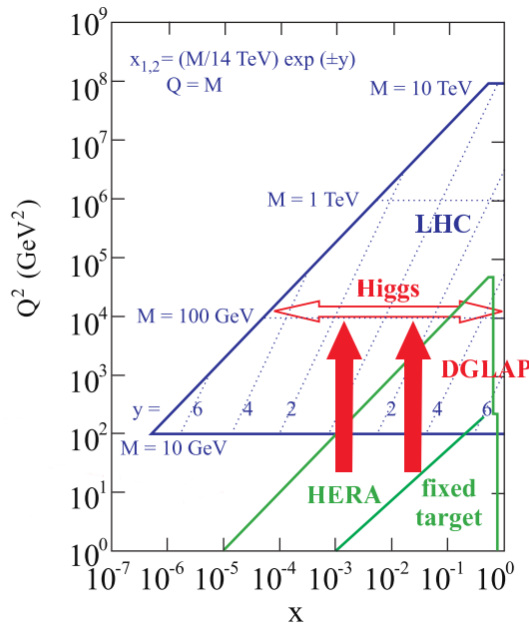


Figure 1.7 – Evolution of the PDFs measured at HERA to LHC energies is performed using the DGLAP QCD evolution equations [19].

Linking the hard scattering matrix element computation to the PDF distribution also requires setting a factorization scale. This scale is responsible for partitioning the computation, ensuring that collinear and infrared divergences are treated in the PDF [19, 45]. Most often the factorization scale is chosen to match the Q of the interaction. This is the most standard choice since Q is the natural scale of the interaction, and since in principle the scale is not a physical parameter and any dependence on the scale will be diminished with the inclusion of higher order corrections. There are arguments that go against this most common choice of scale, namely that since it separates the hard from the soft physics its interpretation isn't entirely unphysical, and other suggested strategies for the choice of scale exist for specific computations [46].

The LHC is used to probe cross sections spanning an immense energy range, as shown in figure 1.8. All measurements at the LHC rely on the PDF description; as such it is imperative that measurements are performed at the LHC to validate and improve these descriptions.

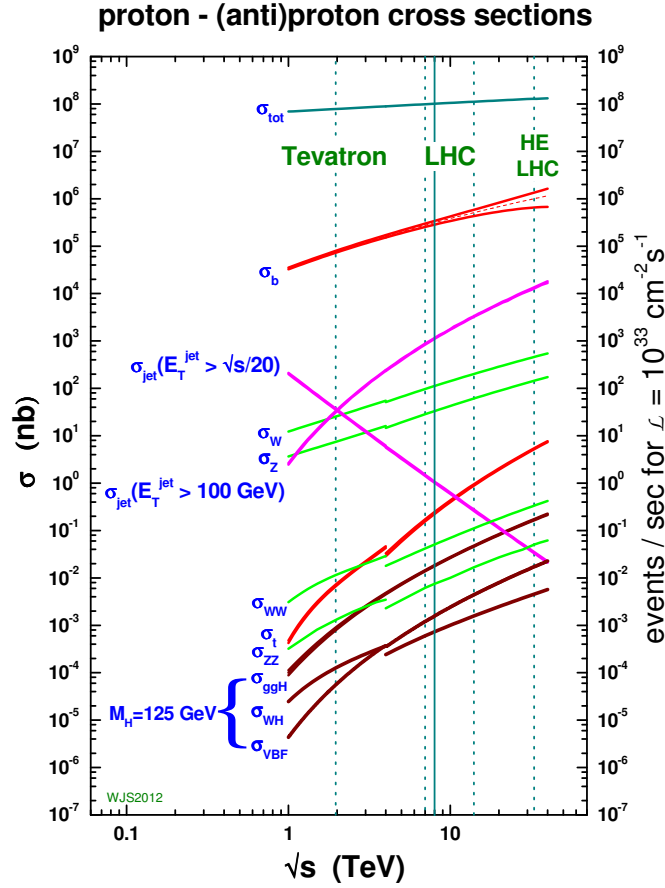


Figure 1.8 – Cross sections for various SM proton-proton scattering processes in the energy range probed by the LHC [47].

1.3 Heavy Flavour at the LHC

Heavy flavour (HF) jets, that is a spray of coloured particles containing either a bottom (b) or a charm (c) hadron, play an important role in many SM and proposed Beyond the Standard Model (BSM) processes that are of interest in the high energy environment explored at the LHC. HF jet production in proton-proton collisions is dominated by mechanisms governed by QCD, and in turn the structure of the proton. Both b and c quarks are fundamentally interesting probes of QCD since, due to their massive nature, their production occurs mostly in hard

scattering processes that lie in the realm of perturbative QCD.

The physics program and methodology for measuring b -quark signals beyond SM QCD production is well established and rich. For example, 60% of Higgs boson decays are to $b\bar{b}$ pairs [48], almost 100% of top (t) quark decays are to the $b + W$ final state [49], and one of the most sought-after supersymmetric particles is the supersymmetric partner to the top quark, decaying largely to a b -quark and a chargino [50].

Aside from being an important background for b -quark signals and an interesting probe of QCD, c -quarks are of particular interest due to their role in potential extended sectors of the SM, via SUSY and t - c mixing, that could elegantly explain the quark mass hierarchy [51]. The methodology for measuring c -quark signals, however, is not so well established. While identifying b -quarks (b -tagging) is challenging to begin with, c -tagging is inherently more difficult because the c -quark properties lie between those of light quarks and b -quarks. Only recently has it become feasible to tag c signals in jets in ATLAS, rather than reconstructing c -mesons, which opens up new physics cases and kinematic regions to be explored at the LHC. To date there has been one analysis in ATLAS employing c -tagging: the SUSY scalar charm search at 8 TeV [52].

1.3.1 Intrinsic Quarks

The theoretical production of heavy flavour signals typically considers only extrinsic quark production. Extrinsic quark production refers to perturbative production via gluon-splitting [53]. This production arises only on very short time-scales in hard interactions, and as such the PDF for heavy quarks follows the QCD evolution equations and falls sharply as a function of the momentum fraction of the proton, x . The current constraints on the presence of intrinsic heavy quarks in the proton, that is non-perturbative heavy quarks that carry some fraction of the proton's momentum, come from the HERA deep inelastic scattering (DIS) experiments [38–41] and do not completely rule out the intrinsic heavy quark hypothesis. This point is highlighted in [54, 55], where a long standing intrinsic charm (IC) model by Brodsky et al (the BHPS model) [56] is used to demonstrate the sensitivity of the $\gamma + c$ final state to the IC hypothesis at the LHC. This is done by

demonstrating that taking the upper limit of 3.5% of the momentum fraction of the proton for the IC component results in an enhancement of the charm PDF at large x that results in a measurable difference in the $\gamma + c$ production cross section. The BHPS model uses a valence-like non-perturbative function to parametrize the IC content. The SEA model is another IC model that instead uses a sea-quark-like non-perturbative function to parametrize the IC content [57, 58], and is more highly constrained by the HERA DIS data. It is important to note that the matrix elements for $\gamma + b$ and $\gamma + c$ processes are identical. Bearing this in mind, due to their smaller mass, c -quarks are more sensitive to an intrinsic content than b -quarks at a given momentum fraction by roughly a factor $m_b^2/m_c^2 \approx 10$. Figure 1.9 (a) displays the charm PDF assuming 0.0%, 1.0% and 3.5% IC fraction using the BHPS model, and the expected enhancement at large x . The effect of IC on the photon+charm cross section as a function of the transverse energy of the photon (E_T^γ) is expected to be largest in the forward region beyond the ATLAS barrel calorimeter, since this corresponds to the phase space of greater longitudinal momentum and in turn higher x . Figure 1.9 (b) displays the predicted difference in the $\gamma + c$ E_T spectrum in the forward region between the nominal PDF and the same PDF assuming a 3.5% presence of IC [54]. Taking a ratio of either the $\gamma + c$ to $\gamma + b$, or of the forward to central cross sections for either $\gamma + c$ or $\gamma + b$, presents an opportunity for enhanced sensitivity when compared to measuring the individual cross sections due to potentially reducing experimental uncertainties that are correlated between the numerator and the denominator of the measurements.

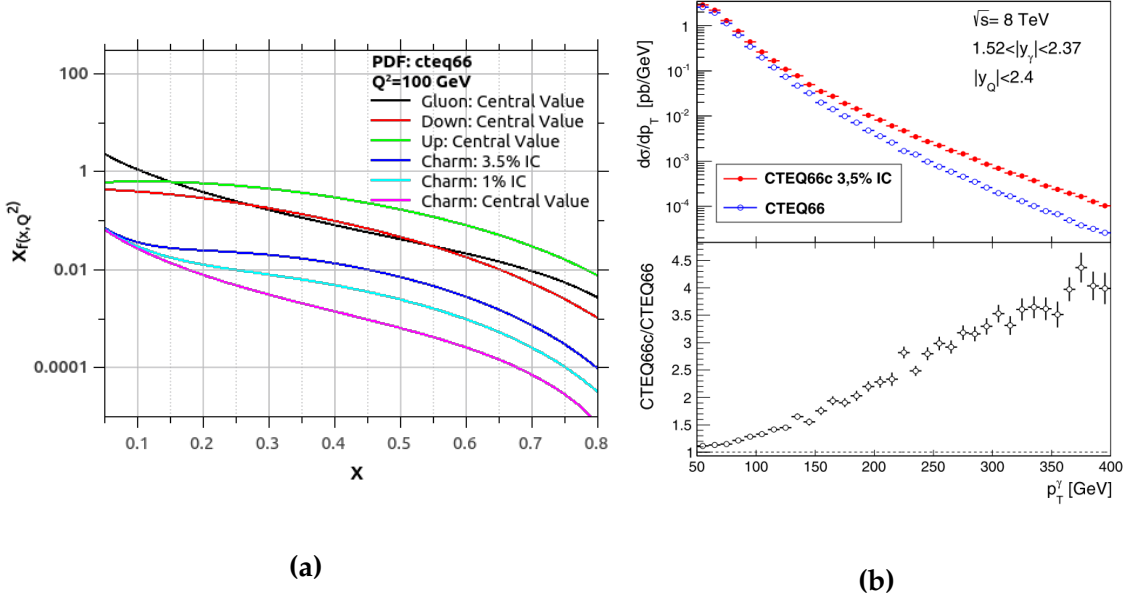


Figure 1.9 – (a) CTEQ6 charm PDF (where $x_f(x, Q)$ is the parton probability) evaluated at $Q=100\text{GeV}$ for 0.0%, 1.0% and 3.5% IC proton fraction [59], (b) Theory predictions for the observed cross section in the forward region using the nominal CTEQ66 PDF and the same PDF assuming 3.5% presence of IC, from [54].

1.4 The Photon+Charm Measurement

Prompt photons refer to those arising from the primary hard interaction and not from secondary decays of other particles [60]. Prompt photons, due to their colourless non-hadronising nature, can be used as well measured probes of parton dynamics [61–73]. Measuring HF jets in association with a prompt photon provides a clean signature with a large cross section that probes the parton level dynamics associated with heavy flavour production. Prompt photons fall into two categories, direct and fragmentation photons. Direct photons refer to those arising from the hard scatter, such as in the instance of Compton scattering, whereas fragmentation photons refer to Bremsstrahlung photons that carry a significant fraction of a hard parton’s momentum. At leading-order (LO) fragmentation photons are often simulated using fragmentation functions to deal with associated collinear divergences. In higher order calculations contributions that are described at LO by fragmentation functions are included in the matrix element computation. LO and Next-to-Leading Order (NLO) Feynman diagrams that contribute to the

$\gamma + HF$ final state are shown in figure 1.10. Here the green lines represent HF quarks, while the diagrams highlighted in red arise with a HF quark in the initial state. The interplay between these two types of diagrams, diagrams with and without an initial state HF quark, highlights that the accuracy of predictions of the HF signals in the LHC environment hinges on the accuracy of the parton distribution functions (PDFs). These measurements, which aim to test this production and the current PDFs, are the first measurements of this nature at the LHC. The most recent measurement of this type was performed at the Tevatron in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV by the D0 [74, 75] and CDF [76] experiments. In comparison to the Tevatron, the LHC provides enhanced sensitivity to the HF PDFs due to the smaller contributions to HF production from quark anti-quark annihilation processes [77].

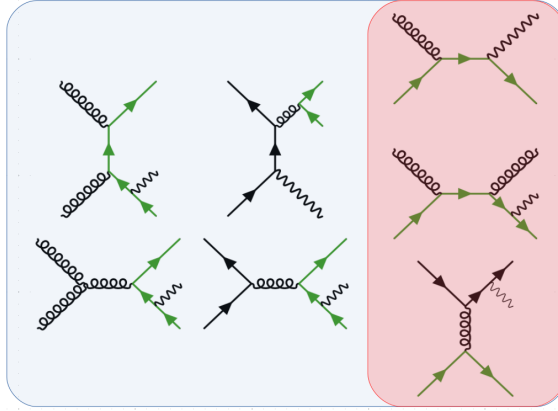


Figure 1.10 – LO and NLO diagrams that contribute to the $\gamma + HF$ final state.

The analysis in this thesis presents the first measurement of the cross section of high E_T isolated prompt photon production in association with a high p_T c -jet in p - p collisions at the LHC. The dataset comprises a total integrated luminosity of $L = 20.2 \text{ fb}^{-1}$ of proton-proton collisions at a center-of-mass energy of $\sqrt{s} = 8$ TeV measured using the ATLAS detector during 2012. Compared to the more recent 13 TeV p - p datasets, as this is the first measurement of its type at the LHC, the lower center of mass energy permits better coverage of the low E_T^γ regime. The methodology for measuring photons follows the 8 TeV inclusive photon analysis, including the binning of the E_T^γ spectrum, to facilitate comparisons between the two measurements. Photons are required to be sufficiently isolated from other

calorimeter signals in order to target prompt photons. Jets are reconstructed using the anti- k_t algorithm [78] and are required to be well separated from the leading photon to avoid interference between the two signals. The fraction of c -jets in the sample is measured using the JetFitterCharm [79] tagging algorithm. This tagger combines tracking and vertex information in a neural network that is trained to provide discrimination between jet flavours. The full measurement, for a given bin of E_T^γ denoted by the index i , is described in equation 1.14:

$$\left(\frac{d\sigma}{dE_T^\gamma}\right)_i = \frac{1}{(\Delta E_T^\gamma)_i} \frac{1}{\mathcal{L}_{\text{int}}} \frac{1}{\epsilon_i^{\text{trig}}} C_i^{\text{unf}} f_i^{\text{c-jet}} \sum_{j \in \text{JFC}} p_{ij}^\gamma N_{ij}^{\text{Data}}. \quad (1.14)$$

In this equation $\left(\frac{d\sigma}{dE_T^\gamma}\right)_i$ is the measured cross section corrected back to the particle level in bin i of E_T^γ , $(\Delta E_T^\gamma)_i$ is the bin width in GeV, $\mathcal{L}_{\text{int}}$ is the measured integrated luminosity, ϵ_i^{trig} is the trigger efficiency, C_i^{unf} is the unfolding factor, $f_i^{\text{c-jet}}$ is the measured heavy flavour jet fraction, p_{ij}^γ is the measured signal photon purity in a tagger discriminant bin j , and N_{ij}^{Data} is the yield of data events. The unfolding factor corrects the measurement for detector effects, including the detector resolution and the signal reconstruction efficiency, yielding a measurement that is directly comparable to other experimental results and particle-level theoretical predictions.

The measurements are compared at particle level to NLO theory predictions using the MADGRAPH5_aMC@NLO generator [80] with CT14 [57] and NNPDF3.0 and NNPDF3.1 PDFs [81] and PYTHIA [82] for the parton shower.

1.5 Roadmap

Chapter 2 outlines the experimental apparatus, that is the LHC, the ATLAS detector, its data acquisition systems and the ATLAS Monte-Carlo (MC) simulation framework. Chapter 3 describes the reconstruction and calibration algorithms used to extract the photon and jet signals from the ATLAS data. Chapter 4 describes the specific data and simulated samples included in the analysis. Chapter 5 outlines the specific selection criteria used to select events in the data and the simulation to populate the distributions used for the analysis. Chapter 6 outlines the measurement procedure. Chapter 7 describes and demonstrates the impact

of the various sources of systematic uncertainty. Chapter 8 shows the measured cross section and compares it to the best available theory predictions using various PDF descriptions. Chapter 9 summarizes the results and outlines the prospects of future work that may improve upon this result.

Experimental Apparatus

The European Center for Nuclear Research (CERN) is an international scientific facility located in Geneva Switzerland. With 22 member states, roughly 2500 staff and 12,000 associated researchers, it is the largest particle physics lab in the world boasting a broad program that connects researchers across the globe. Conceived in the post WWII era as a collaborative center for European scientists, throughout its rich history CERN has unquestionably broadened this goal to a global scale and yielded measurements and discoveries that have deepened our understanding of matter at its most fundamental level. CERN's discoveries include the discovery of neutral currents with the Gargamelle experiment [83], the discovery of the $W^\pm$ and Z boson with the UA1 and UA2 experiments [84, 85], the determination of the number of neutrino flavours with the LEP collider [86], the first creation of anti-hydrogen by the ALPHA experiment [87], observations of jet quenching [88], pentaquarks [89], and the discovery of the Higgs boson at the LHC [11, 12]. Technological advancements born from the expertise and innovative tools required to pursue these endeavors have made significant impacts outside the realm of fundamental research, including notable contributions to Information Technology (with the birth of the World Wide Web and the emergence of grid computing), medicine (hadron therapy, rare isotope production and imaging), the energy sector and more.

Presently the LHC and its associated detector experiments are providing a window into phenomena arising in states of matter never before created through human endeavor. The LHC has been the central project of CERN for the past twenty years, conceived with the primary goal of discovering or ruling out the existence of the Higgs boson, and accompanied by a very rich physics program aiming to shed light on many questions and phenomena of fundamental particles and their interactions. The state and evolution of our universe hinges on what happened in the instants following the big bang when there existed an extremely hot and dense environment. From this state our universe cooled, expanded and

evolved. These states of matter that existed in the early universe are what the LHC explores.

2.1 Collider Basics

The LHC detector experiments were designed to measure and characterize the interactions revealed by the collisions provided by the LHC. In experimental collider physics this relationship between the collider, the physical interactions, and the detector experiment is encapsulated by the following relationship:

$$N_{event} = L\sigma_{event}. \quad (2.1)$$

The LHC machine provides particle bunch collisions, the frequency and density of which are characterized by the luminosity L . Nature provides the probability that a collision at a given energy will result in a possible interaction, encapsulated in the cross section σ_{event} . The detector aims to accurately measure the number of outgoing particles resulting from interactions between colliding bunches, N_{event} .

The luminosity of the beams at the interaction points is characterized using the following relationship:

$$L = \frac{N_b^2 f_{coll}}{4\pi\sigma_T}, \quad (2.2)$$

where N_b is the number of particles per colliding bunch, f_{coll} is the collision frequency, and σ_T is the transverse beam size at the interaction point. Looking at this equation we see that higher luminosities are achieved by having more particles in a tighter overlapping collision area. The transverse beam size is parametrized by the emittance ϵ , the Beta function parameter β^* , and the relativistic Lorentz factor γ of the beam:

$$\sigma_T = \frac{\epsilon\beta^*}{\gamma}. \quad (2.3)$$

The emittance is the average spread of the particles in transverse position-momentum space, characterizing the transverse oscillations. Following Liouville's theorem, the emittance of an ideal particle beam remains constant as the beam navigates through guiding optics that use conservative forces [90]. As a consequence, when the beam is focused to a waist in position space (becoming narrow, like light through a magnifying glass), it becomes unfocused in transverse momentum space,

implying that the beam is also divergent (again, just like light from a magnifying glass beyond the focal point). This conservation puts a limit on how narrow a beam can be focused, since more powerful and precise optics are required to create a very narrow and divergent waist. The lower the emittance of the beam, the tighter it can be focused using the same optics and in turn the greater luminosity that can be achieved. There is no such thing as an ideal machine, thus emittance does grow due to injection mismatch errors (focusing, steering, dispersion), intra-beam scattering and other sources of noise, but must be minimized to a tolerable level.

The $\beta(s)$ function quantifies the divergence of the beam as a function of the distance along the beam trajectory (s) from the interaction point:

$$\beta(s) = \beta^* + \frac{s^2}{\beta^*}, \quad (2.4)$$

where β^* is the value of the $\beta(s)$ function at the beam waist, and thus characterizes the size of the waist. Large values of β^* imply that the beam is not very tightly focused, while small values imply that it is very tightly focused and in turn divergent.

Now that we have these basic technical concepts, we can proceed to describe the CERN accelerator complex.

2.2 The LHC Injector Chain

The beams that are injected into and then accelerated by the LHC are provided by a multi-stage accelerator chain depicted in 2.1. Each circular accelerator in the chain must be completely filled at its injection energy prior to accelerating that batch of particles to the injection energy of the subsequent stage. In 2012, filling the LHC would typically take 30 minutes. Tracing this accelerator chain from beginning to end also traces through many stages of CERN's history. Hydrogen atoms originating from a bottle of hydrogen gas are injected into a device, called a duoplasmatron, that strips the atoms of their electrons and injects the resulting plasma of protons into the linear accelerator 2 (LINAC2) [91]. This linear accelerator, commissioned in 1978, uses radio-frequency (RF) cavities to charge

cylindrical conductors with quadrupole focusing magnets that create a pulse of protons every $100 \mu\text{s}$ at an energy of 50 MeV. These protons are injected in the Proton Synchrotron Booster, commissioned in 1972, a four-ring synchrotron that accelerates the protons from 50 MeV to the 1.4 GeV injection energy for the Proton Synchrotron (PS). The PS, commissioned in 1959, used to be CERN's flagship accelerator. Its most notable discovery occurred in 1974 when it provided a neutrino beam to the Gargamelle experiment leading to the first observation of weak neutral currents [83]. Consisting of 277 room-temperature magnets on a 100 m radius ring, it accelerates the protons up to 25 GeV prior to injection to the Super Proton Synchrotron (SPS). The SPS, commissioned in 1972, superseded the PS as the flagship of the CERN accelerator complex and is most notable for its role in the Nobel prize winning discovery of the $W^\pm$ and Z bosons in 1983 [84, 85]. 1.1 km in radius and consisting of 1317 room-temperature magnets, the SPS accelerates the protons provided by the PS to an energy of 450 GeV. These protons, in addition to being injected to the LHC, are also used in fixed target experiments (COMPASS, NA61/SHINE, NA62) and to create a neutrino beam aimed at the Gran Sasso laboratory in Italy.

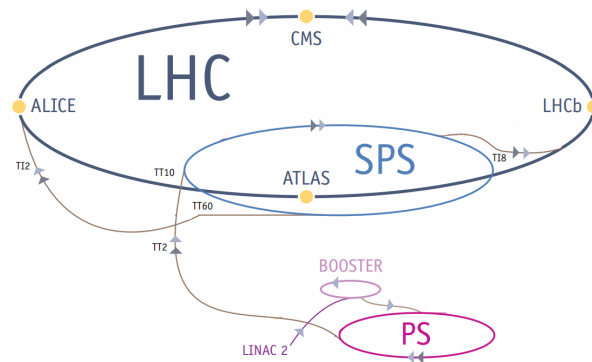


Figure 2.1 – Illustration of the LHC accelerator chain [92].

2.3 The LHC

The Large Hadron Collider (LHC) is a circular two-ring hadronic superconducting accelerator and collider housed in the 26.7 km tunnel that formerly housed the LEP collider. The tunnel is located between 45 m and 170 m below the surface and spans the French/Swiss border, as seen in Figure 2.2, consisting of eight crossing points connected by straight tunnels. This geometry is an artifact of the LEP collider, where the straight sections would compensate for synchrotron radiation losses that are more significant for a lepton collider. Due to space constraints in the tunnel, which has a radius of 3.7 m, the LHC was designed to make use of a twin-bore superconducting magnet design that can house both rings in a single apparatus. Since the LHC accelerates protons, which are ≈ 2000 times heavier than electrons, the synchrotron losses are not a problem for achieving higher energies.

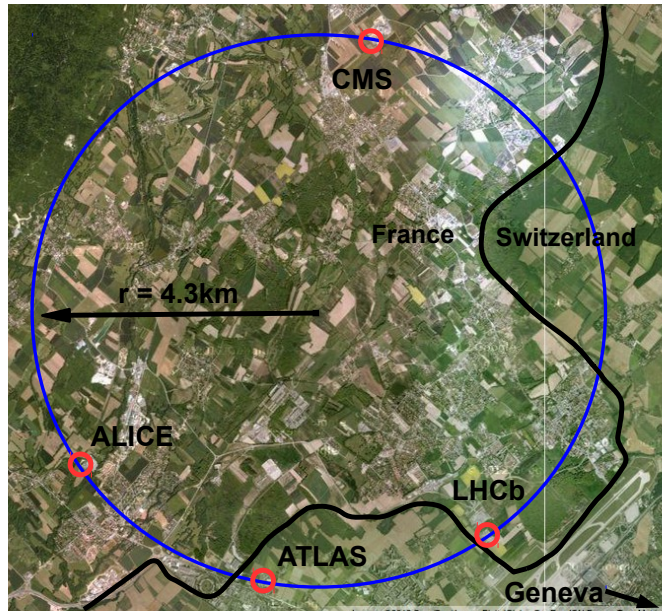


Figure 2.2 – The footprint of the LHC spanning the French/Swiss border outside of Geneva.

In the LHC, charged particles circulate in a vacuum of 10^{-13} atm and are guided and focused by superconducting magnets [93]. The cryostat that houses the beam pipe and magnets is maintained at a temperature of 1.9 K using superfluid liquid helium. It is very challenging to keep particles circulating in a stable

orbit around the entire LHC. Picture a particle traveling transverse to a uniform magnetic field. The Lorentz force bends the particle in a perfect arc with a radius proportional to the particle momentum and inversely proportional to the magnetic field strength. As LHC is not a perfect circle and since there are imperfections in the ring and optics the LHC beams are subject to momentum dispersion. This momentum dispersion results in the protons following slightly different trajectories that must be corrected to maintain a stable orbit. The majority of the steering elements are dipole magnets, responsible for bending the particle trajectories around the arc of the LHC. Quadrupole magnets are used to focus or defocus the particles along a single transverse axis, while various other superconducting beam manipulation elements provide higher order corrections and compensate for losses.

Particles are accelerated in the LHC by RF cavities located in a cavern on the ring located between ALICE and CMS. As particles traverse the cavities they are accelerated by "surfing" on RF waves. Each wave is referred to as a bucket and can carry a "bunch" of particles, imposing a periodic particle bunch spacing with a size that is set by the RF cavity harmonics, that is the standing RF waves that are permitted by the cavity shape. One can think of the LHC as two circulating chains of buckets, where some fraction of them are filled with particles and others are not. Buckets are filled with different schemes to provide different types of collisions, and certain buckets are always kept empty in order to provide a window for steering the beam into a beam dump. The full collection of circulating particles in one direction is referred to as a beam. Though the beams are circular, their buckets are numbered in order to describe the occupancy and filling scheme of the LHC beams. The filling scheme is communicated to the detector experiments so that they can trigger on desired bunch crossing types (usually filled and colliding bunches). Once filled, the LHC aims to keep the beams circulating, while focusing and colliding them at the interaction points, until the luminosity has dropped to a point where it is more profitable to dump the beams and refill. In 2012 it took roughly 45 minutes to go from injection to colliding stable beams (15 minutes ramp, 5 minutes flat top, 15 minutes squeeze, 10 minutes adjust). The fill with the longest period of colliding stable beams in 2012 was fill 2692, lasting for almost 23 hours. Table 2.1 displays the LHC design parameters and the parameters used during 2012 and 2015 running. Due to unforeseen failures during the initial LHC

commissioning in 2008, a running scheme with lower energy and lower luminosity was adopted for Run 1 (2010-2012). Following the successful running of the LHC in Run 1, and following upgrades to the LHC system during the mechanical stop between Run 1 and Run 2, a run scheme more in line with the design parameters was adopted for Run 2 (2015-2017).

Parameter	Design	2012	2015
Beam Energy [TeV]	7	4	6.5
Bunches per beam	2808	1374	2780
Bunch spacing [ns]	25	50	25
Protons per bunch	1.15×10^{11}	$1.1 - 1.7 \times 10^{11}$	$1.2 - 1.7 \times 10^{11}$
Emittance [μm]	3.75	2.4	2.9
β^* [m]	0.55	0.6	0.8
Maximum luminosity [$\text{cm}^{-2}\text{s}^{-1}$]	10^{34}	7.7×10^{33}	8.6×10^{33}

Table 2.1 – Design, 2012 and 2015 operational parameters for the LHC.

2.4 The ATLAS Detector

2.4.1 Coordinates

This section describes the coordinate system used by the ATLAS experiment. As seen in figure 2.3, take z to be along the beamline with the positive x axis pointing to the center of the ring and the positive y axis pointing upwards. Using polar coordinates ϕ is the azimuthal angle about the beam axis, θ is the polar angle measured from the beam axis, and $\rho = \sqrt{x^2 + y^2}$. Although the beam energies are equal, individual colliding particles will not necessarily have equal and opposite momenta, resulting in collision center of mass frames that are boosted and continue to travel in the z direction. For this reason it is desirable to have observables that are invariant to boosts along the z axis. Various quantities achieve this by being measured using scalar projection on to the plane transverse to the z direction (such as transverse momentum p_T , or transverse mass m_T). Rapidity, y , is a physical quantity computed using the energy, momentum and polar angle: $y = \frac{1}{2} \sqrt{\frac{E+cp_z}{E-cp_z}}$. Rapidity has the desirable property that boosts along the z

direction result in a simple additive transformation, implying that rapidity differences are invariant under boosts along the z axis. In addition, it turns out that the most abundant particle production multiplicities are relatively flat as a function of rapidity [94], making it a good choice for calorimeter segmentation. Pseudorapidity, η , is the limiting case of rapidity where the particle is assumed to be massless, a good approximation in many cases and in particular at high energies. Pseudorapidity is defined as $\eta = -\ln(\tan \theta/2)$ using only the θ angle, it can easily be mapped back to rapidity with the object p_T and mass, and is also Lorentz boost invariant. For this reason the segmentation of the detector is roughly evenly spaced in pseudorapidity. Often the total angular separation between particles in η and ϕ is taken as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$. [95].

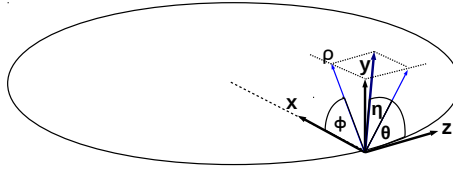


Figure 2.3 – Illustration of the coordinate system used in ATLAS.

2.4.2 The Inner Detector

The ATLAS Inner Detector (ID) is a combined pixel, silicon microstrip semiconductor tracker (SCT) and straw tube transition radiation tracker (TRT). The primary aim of the ID, shown in figure 2.4, is to precisely measure trajectories of charged particles in the region closest to the beam pipe. By immersing a tracker in a magnetic field of known strength and polarity, the momentum and sign of a

particle can be determined by measuring the track radius of curvature. By mapping tracks to common origins, primary and secondary decay vertices can be reconstructed yielding critical information linking detector signals to common interactions and revealing their decay mechanisms. In a given event the ATLAS detector will record signals arising from primary hard interactions from the triggered bunch crossing, soft interactions arising from the same bunch crossing (referred to as in-time pileup), and leftover interactions from the previous bunch crossing (referred to as out-of-time pileup). The ID plays a critical role in measuring the total event topology in order to discriminate the signals arising from the primary interactions from those that arise from pileup interactions [96,97]. With an inner radius of 36 mm the ID sits flush against the beam pipe and is immersed in a uniform 2 T magnetic field along the beam axis that is provided by a superconducting solenoid, shown in Figure 2.5. Radiation damage of the ID is an important consideration due to its proximity to the interaction point.

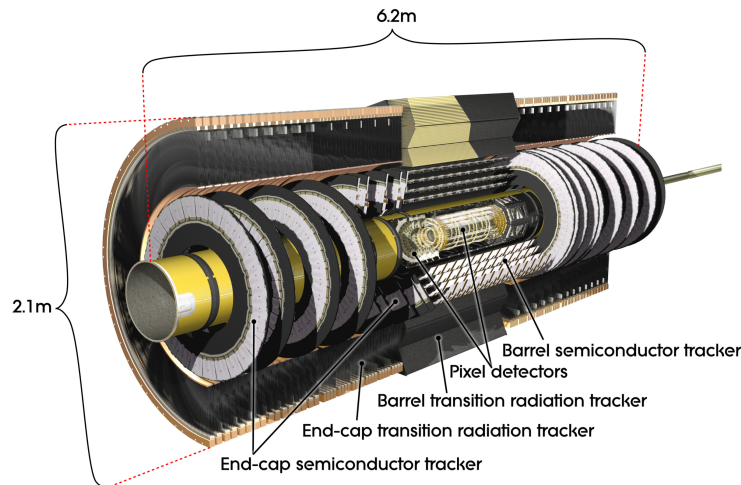


Figure 2.4 – Illustration of ATLAS Inner Detector and its components [98].

The pixel and SCT systems are made up of solid-state ionization detectors. The pixels, illustrated in figure 2.6, are constructed by doping an n-type silicon wafer with positive regions on one side and negative regions on the other, known as a PIN junction [99]. The interface between the middle and the two doped regions creates a diode where the chemical potential induces a voltage that depletes

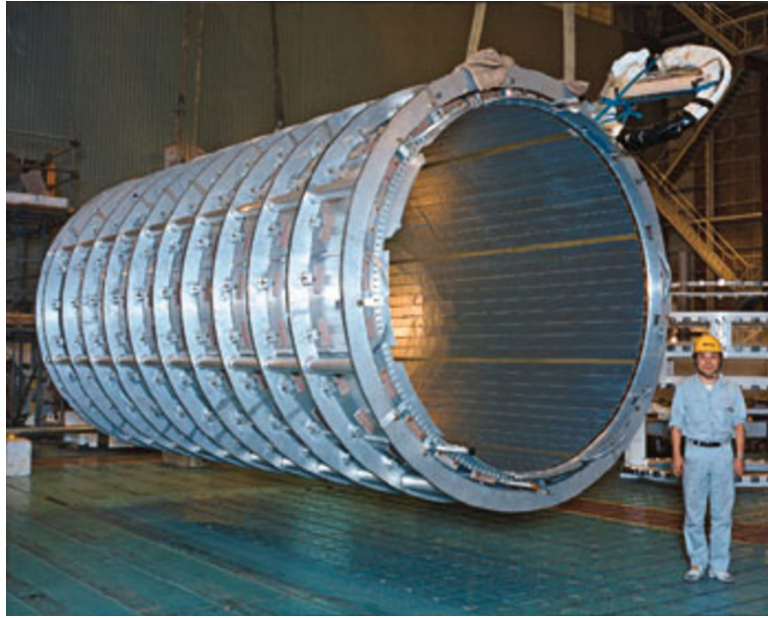


Figure 2.5 – Photo of the ATLAS superconducting solenoid [98].

the interface of free charges, referred to as the depletion region. This region is the active region of the detector since any electron-hole pairs that are created by an ionizing particle will be swept out by the electrical potential that opposes the chemical potential. The size of this active region is increased to encompass almost the entire region between the doped regions by placing the interface under an external voltage that "helps" the chemical potential and "opposes" the default electrical potential. This configuration of diode operations is referred to as *reverse bias*. Whenever a charged particle traverses the depletion region ionization charge is swept out by the reverse bias voltage providing the measured signal. The primary source of noise in these silicon detectors arises from the leakage current, that is the current that arises from thermal energy creating electron-hole pairs in the depletion region. For this reason the detectors are operated in the -5°C to -10°C range. The position resolution of these detectors, which ranges from tens to hundreds of μm , is excellent considering that the ionization charge is created within microns of the incident charged particle.

Radiation damage of these silicon detectors is well understood, leading to changes in the depletion voltage, increases in the leakage current, and decreases

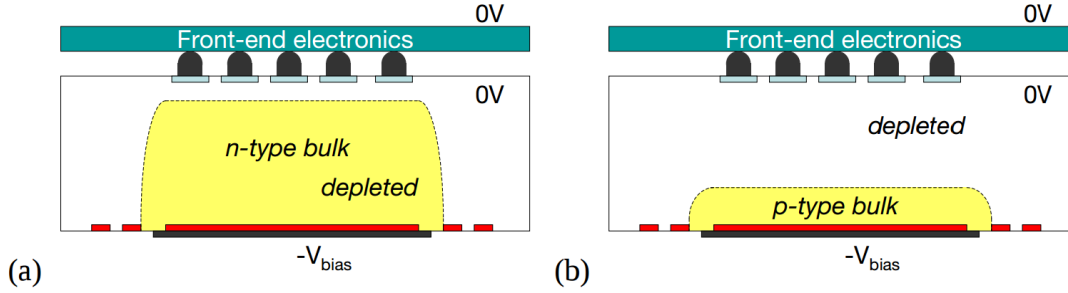


Figure 2.6 – Illustration of an ATLAS pixel detector diode before (a) and after (b) type inversion of the n-doped region. [99]

in the charge collection efficiency. Increases in the leakage current are mitigated throughout the run period during technical stops by performing low temperature annealing on the ID [100]. The reduction of the depletion voltage is most significant for the innermost layers of the ID, leading to the eventual type inversion of the n-doped layers after a fluence of $2 \times 10^{13} \text{cm}^{-2}$, as illustrated in figure 2.6.

The pixel detector sits closest to the interaction point, shown in figure 2.7, with three barrels in the central region with radii of 5, 9 and 12 cm, and three disks on either side of the barrel region. Comprising a total of 80 million pixels [99], the system ensures three precisely measured hits for each track with full acceptance. The barrel provides a precision of $10 \mu\text{m}$ in $R\text{-}\phi$ and $115 \mu\text{m}$ in z . Each layer is 2.5% of a radiation length at normal incidence.

The SCT sits outside the pixel detector with a similar geometry and consists of four barrel layers and nine disks on either side of the barrel region [101]. With approximately 6.3 million readout channels the system provides eight precisely measured hits for each track covering the region $|\eta| < 2.5$. The barrel layers consist of $6.36 \times 6.40 \text{ cm}^2$ silicon detector units that each have 780 readout strips, while the end cap disk detector units are similar but with a tapered geometry. Each detector unit consists of two planes of silicon strips with a pitch of $80 \mu\text{m}$, with a relative 40 mrad offset between the two planes, which permits stereo measurement of the z coordinate. The radii of the SCT barrel detectors are 30.0, 37.3, 44.7 and 52.0 cm. The barrel provides a precision of $17 \mu\text{m}$ in $R\text{-}\phi$ and $580 \mu\text{m}$ in z .

The TRT relies both on the collection of primary ionization charge and the collection of secondary ionization charge arising from transition radiation to meas-

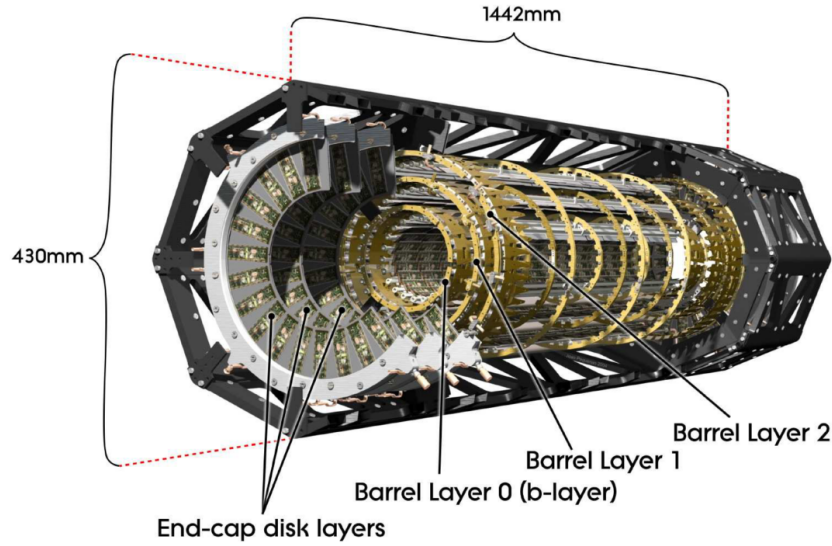


Figure 2.7 – Illustration of the ATLAS Pixel Detector, comprising three layers in both the central and barrel regions. [98]

ure passage of charged particles. Transition radiation occurs when a charged particle traverses a boundary between media of different dielectric constants. The energy of the emitted radiation depends strongly on the relativistic γ factor of the charged particle, which makes these detectors adept at identifying electrons with energies of up to roughly 200 GeV by virtue of their small mass.

The TRT barrel covers the radius of 56-108 cm and $|\eta| < 1.0$, while the combined barrel and end cap cover up to $|\eta| < 2.0$ with a combined total of approximately 351,000 readout channels. Within this region of coverage particles with $p_T > 0.5$ GeV will traverse roughly 35 drift tubes providing continuous tracking, arranged longitudinally to the beam axis in the barrel region and radially in the end caps [102]. Note that the barrel region does not provide any z information. The drift tubes are filled with 70% Xe, 27% CO₂ and 3% O₂ gas, as xenon is particularly adept at absorbing TR photons. The wall of the tubes is operated at a -1530 V potential difference with respect to the gold-plated tungsten wire that runs down the middle and collects the ionization charge [103]. Polypropylene layers are used to stimulate the transition radiation between the tubes. In the barrel region the tubes are interleaved with a matrix of polypropylene fibers while in the end caps

polypropylene foils are located between the tube layers. The TRT provides a precision of $130 \mu\text{m}$ per tube in $R\text{-}\phi$. The gas used in the TRT is re-circulated and continuously monitored to ensure its quality, mitigating the effects of radiation damage to the active material of the detector.

2.4.3 The Calorimeter

Calorimetry in Particle Physics

In particle physics a calorimeter is a device that aims to stop and fully absorb an incident particle, and in doing so convert some fraction of its energy into a measurable signal. The ATLAS detector employs a high granularity electromagnetic calorimeter to measure electrons and photons, and a more coarse granularity hadronic calorimeter to measure hadronic jets and missing transverse energy. Both are sampling calorimeters consisting of alternating layers of absorber and active material. As charged particles traverse the dense absorber layers they interact and lose energy, producing secondary particle showers. For electrons and photons these interactions are electromagnetic, while for hadrons these interactions are mostly strong. The calorimeters are designed to contain as much of the resulting shower as possible in order to ensure the best measurement of a particle's total energy by minimizing the occurrence of punch-through particles, which can also compromise the performance of the surrounding muon system.

The performance of a calorimeter is measured by its energy resolution. Consider an ideal beam of electrons that all have exactly the same energy. The resulting distribution of signals that this beam produces as it is stopped and measured by a calorimeter will have some spread due to a variety of physical factors. This spread is referred to as the resolution of the calorimeter and is most often expressed as the fraction of the measured energy $\frac{\sigma}{E}$. There are three main independent categories of effects that contribute to a calorimeter's resolution, parametrized by measured constants and added in quadrature in equation 2.5 [104].

$$\frac{\sigma}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c, \quad (2.5)$$

where $\oplus$ indicates a quadratic sum. The first constant, a , comes from stochastic

effects, i.e. effects due to differences in the number of daughter particles produced in the calorimeter. As a particle interacts it produces a shower of secondary particles, and the measured energy is given by the total distance of active material that these secondary particles traverse. From one particle shower to the next there will be variations in the number of particles in the shower, the number of active layers a given particle traverses, the angles at which they traverse the active material layers, etc... Thus the broadening of the energy resolution arising due to stochastic effects is tied to the number of secondary particles following Poisson statistics $\sqrt{N}$, which is proportional to $\sqrt{E}$. The second constant, b , comes from irreducible readout noise arising from thermal signals and signals arising from uncorrelated background events (pileup) [105]. The third constant, c , encapsulates effects that are independent of the particle's energy, arising from dead material in front of the calorimeter, and mechanical inhomogeneities within the calorimeter.

Electrons and photons interact with the calorimeter via Bremsstrahlung, pair production, Compton scattering and the photoelectric effect. Compton scattering and the photoelectric effect dominate interactions at low energies while for energies larger than 10 MeV Bremsstrahlung and pair production dominate, and above 1 GeV their effects become almost energy independent. If the incident particle is above the critical energy of the material it will induce pair production and Bremsstrahlung resulting in a cascade of other high energy electrons and photons until its energy falls below this threshold energy. The radiation length of the material, X_0 , corresponds to how far an electron must travel in order to lose $1/e$ of its incident energy due to Bremsstrahlung. For a photon in the same material the radiation length is $9/7X_0$. Thus the profile and depth of an energetic electromagnetic shower are given by the material properties of the calorimeter characterized by the radiation length.

Hadronic calorimetry presents additional challenges compared to electromagnetic calorimetry due to the presence of additional hadronic and nuclear interactions. Secondary hadrons are produced due to strong interactions that carry significant fractions of the incident particle's energy. A significant fraction of these secondary hadrons are π^0 mesons that decay electromagnetically making the evolution of the hadronic shower multi-faceted. The ratio of the visible electromagnetic response to the visible hadronic response is known as the "e/h", or "e/ π ",

and its average value exhibits energy dependence. Other effects present in hadronic signals include nuclear interactions such as excitations and spallations of the absorber material absorb energy and emit particles. The total measured hadronic shower energy must therefore be calibrated using prior measurements of the calorimeter response to known hadronic input signals using a test beam. This e/h ratio depends on the fraction of secondary hadrons that are π^0 s, and variation in this fraction on an event to event basis is a limiting factor on the energy resolution. Similarly to electromagnetic particles, the interaction between hadronic particles and matter is described by a characteristic length scale that is the distance a hadronic particle must travel to lose $1/e$ of its energy, referred to as the nuclear interaction length λ [106]. The nuclear interaction length for dense materials is much longer than the electromagnetic radiation length indicating the hadronic showers are more diffuse and start later than electromagnetic showers.

The ATLAS Calorimeters

The ATLAS detector's electromagnetic calorimeter (EMC) aims to measure electrons and photons between several GeV and several TeV. Incident photons and electrons induce electromagnetic showers in the calorimeter via pair production and Bremsstrahlung. Using lead as the absorber material and liquid argon as an active material, as charged particles traverse these active layers they ionize the noble liquid and the ionization charge is collected and readout using electrodes that impose a high voltage across the layer. The calorimeter comprises a very thin (11 mm of liquid argon) initial presampling layer with $|\eta| < 1.8$, a barrel covering $|\eta| < 1.37$, and end caps covering $1.56 < |\eta| < 3.2$. The presampler is used to help measure and correct for energy lost by particles prior to reaching and developing their shower in the calorimeter (energy lost in the ID, the solenoid and the cryostat), reducing the energy resolution by up to 40% [107]. By virtue of the calorimeter's accordion geometry it provides full coverage in ϕ , and in total has roughly 165,000 readout channels. The calorimeter has a stopping power on the order of $22X_0$ [98], of which approximately $10X_0$ is active material, providing an energy resolution of $\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} \oplus 0.7\%$. The segmentation of the calorimeter in η is provided by etching the electrodes that make up the copper accordion, while the segmentation in ϕ is provided by combining the readout of adjacent electrodes. As

depicted in figure 2.8, the barrel section is composed of three separate layers of decreasing segmentation in η moving outwards from the center, while in the forward region ($2.5 < \eta < 3.2$) there are two. In the barrel the first layer is eight times finer than the middle, providing the crucial measurements of the shower profile, while the middle layer is responsible for collecting most of the shower's energy. The final layer is twice as coarse as the middle layer and is responsible for collecting the shower tail.

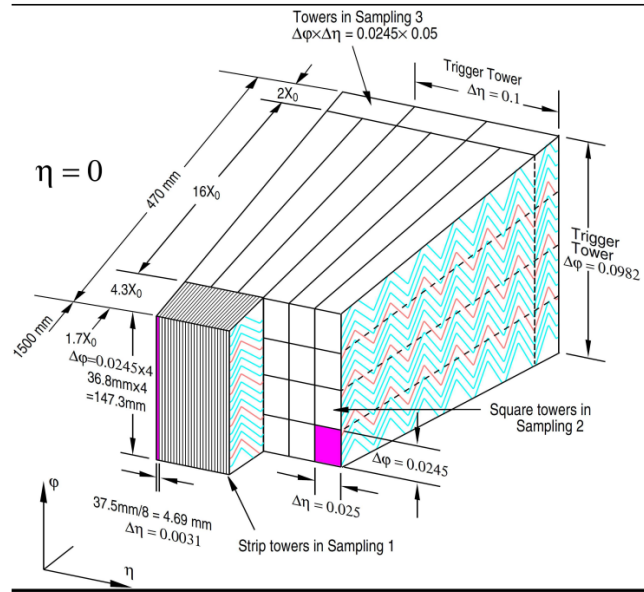


Figure 2.8 – Illustration of ATLAS Electromagnetic Calorimeter and its accordion geometry [98].

The ATLAS hadronic plastic scintillator sampling calorimeter (HCAL) aims to measure and contain high energy hadrons. As charged particles traverse these active layers they excite bound electrons in the scintillator material that upon relaxation emit UV light. This light is in turn converted into visible light optimized for detection by photo-multiplier tubes using scintillator dyes and wavelength shifting light guides. As seen in figure 2.9, the HCAL rests outside the EMC with an inner radius of 2.28 m, an outer radius of 4.23 m. The length of the central barrel is 5.56 m and referred to as the Long Barrel (LB), while on either side of the LB there is a 2.91 m Extended Barrel (EB). The calorimeter's total reach is $|\eta| < 1.7$ and is made of iron absorber plates segmented with plastic scintillator tiles providing a

radial depth of approximately 7.4λ lengths. The scintillator tiles are read out using wavelength shifting fibers that are coupled to photo-multiplier tubes adding up to approximately 10000 read-out channels. The calorimeter has three radial layers, the innermost two having a segmentation in $\eta\phi$ of 0.1×0.1 , the outer layer having a segmentation of 0.2×0.1 . The HCAL employs a sophisticated in-situ calibration system. A Charge Injection System (CIS) built into the data acquisition (DAQ) electronics calibrates the signal digitization electronics by injecting known amounts of charge that span the dynamic range of the detector. The laser system illuminates the PMTs with a reference intensity and corrects for drift in the PMT gain and linearity. The Cesium System consists of a movable Cs source that circulates throughout the detector stimulating the system and permitting the overall optical response of the scintillator tiles to be corrected for non-uniformities [108].

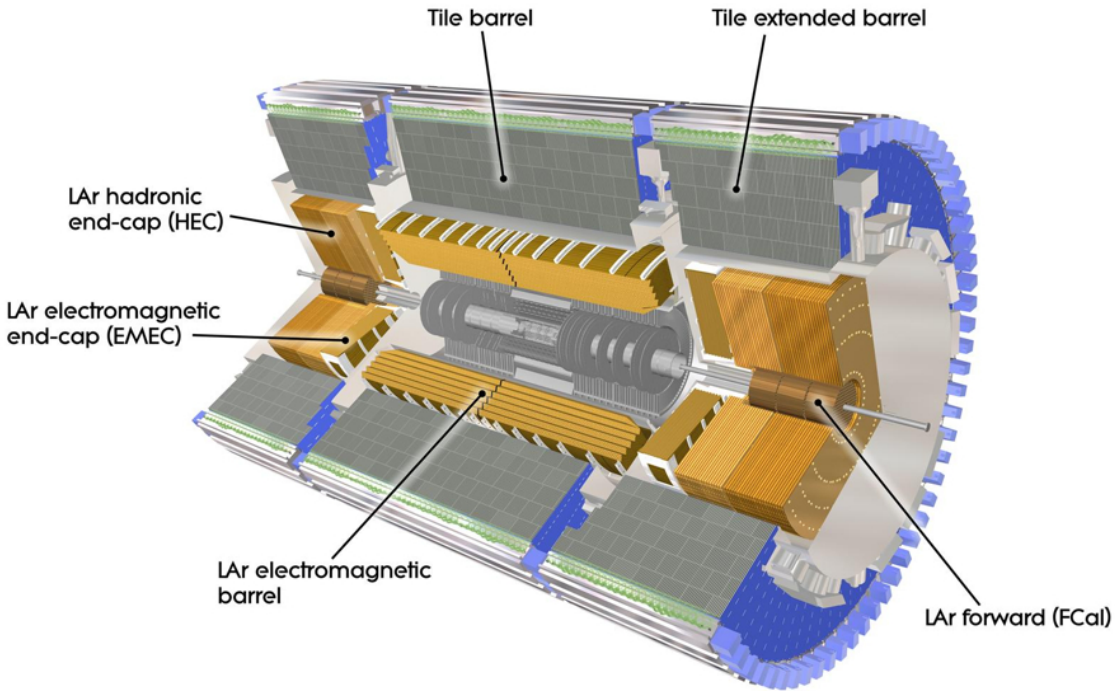


Figure 2.9 – Illustration of ATLAS Calorimeter [98].

The hadronic end-cap calorimeters (HEC) cover a range of $1.5 < |\eta| < 3.2$ and each consists of a front and rear wheel made up of 32 wedge-shaped copper li-

quid argon sampling modules. Liquid argon is used in the end-cap region because of the significantly larger radiation dose. The front wheels have an inner radius of 372 mm and an outer radius of 2.03 m while the outer wheels have an inner radius of 475 mm and an outer radius of 2.03 m. The entire end cap has a readout resolution in $\eta \times \phi$ of 0.1×0.1 for $1.5 < |\eta| < 2.5$ and 0.2×0.2 from $2.5 < |\eta| < 3.2$. The energy resolution of the hadronic barrel/end cap is $\frac{\sigma_E}{E} = \frac{50\%}{\sqrt{E}} \oplus 3\%$.

The ATLAS forward calorimeters (FCAL) are end cap calorimeters located beyond the electromagnetic end cap and inside the hadronic end cap covering $3.1 < |\eta| < 4.9$. On either side of the detector there are three 45 cm layers, the innermost is an electromagnetic copper liquid argon calorimeter that is optimized for resolution, while the middle and outer layers are hadronic tungsten liquid argon calorimeters that are optimized for radiation length. The geometry of the FCAL modules are unique to the other calorimeters in ATLAS and consist of small-diameter copper tube electrodes with coaxial copper or tungsten rods mounted using radiation-hard plastic wound around the rod. The rods are oriented parallel to the beam yielding granularity in the transverse plane and mounted using perforated copper plates. The active liquid argon gaps in the FCAL are smaller than in the rest of ATLAS in order to optimize readout speed and better measure the high fluxes seen in the forward region. The drift time in the FCAL is on the order of 60 ns, whereas in the EM calorimeter and HEC it is on the order of 450 ns. The energy resolution of the FCAL is $\frac{\sigma_E}{E} = \frac{100\%}{\sqrt{E}} \oplus 10\%$ [109].

The EM calorimeter, the HCAL and the FCAL all exhibit excellent resilience to radiation damage by virtue of the constant re-circulation and monitoring of the quality of the liquid argon active material.

2.4.4 The Muon System

The ATLAS detector muon system aims to trigger on muons with an acceptance of $|\eta| < 2.4$ and to precisely measure muon trajectories with an acceptance of $|\eta| < 2.7$. The system is immersed in a 4T magnetic field provided by three air-core toroid magnets: the large barrel toroid that covers $|\eta| < 1.4$ providing 1.5 – 5.5 Tm bending power and the two end cap toroids that cover $1.6 < |\eta| < 2.7$ providing 1 – 7.5 Tm bending power. The trigger system employs Resistive Plate Chambers

(RPC) in the barrel and Thin Gap Chambers (TGC) in the end caps, while the precision spectrometer uses Monitored Drift Tubes (MDT) up to $|\eta| < 2.0$ centrally and Cathode Strip Chambers (CSC) in the forward region $2.0 < |\eta| < 2.7$ [98].

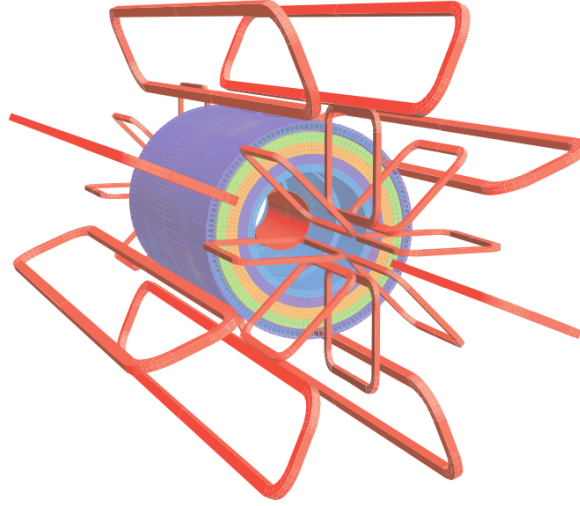


Figure 2.10 – Depiction of the ATLAS superconducting toroid and solenoid magnet system (red). At the center the tile calorimeter is also depicted. The toroid’s primary purpose is the bending of muon tracks [98].

2.4.5 The Luminosity Measurement System

It is clear from equation 2.1 that just about every measurement made by ATLAS relies on the precision and accuracy of the measured luminosity. At the LHC the luminosity is measured in a two step process. First the inelastic proton scattering rate, μ , is measured online, and then these measurements are calibrated to the absolute scale using special calibration runs that measure the transverse beam profile. These calibration runs, named after Simon van der Meer who pioneered the technique at CERN in the 1960s [110], alternate keeping one of the beams stationary while scanning the other beam across the interaction point, allowing the beam intensity profiles to be extracted as the μ is simultaneously measured for each individual colliding bunch pair. Equation 2.6 shows the relationship between μ and the absolute luminosity, where the visible inelastic proton scattering rate and vis-

ible cross section are $\mu_{vis} = \epsilon\mu$ and $\sigma_{vis} = \sigma\epsilon$, respectively, taking into account the detector efficiency ϵ . n_b is the number of colliding bunches and f_r is the frequency of revolution:

$$L = \frac{\mu_{vis} n_b f_r}{\sigma_{vis}}. \quad (2.6)$$

Figure 2.11 displays the various detectors that contribute to these measurements and calibrations. ATLAS relies primarily on the LUCID and Beam Con-

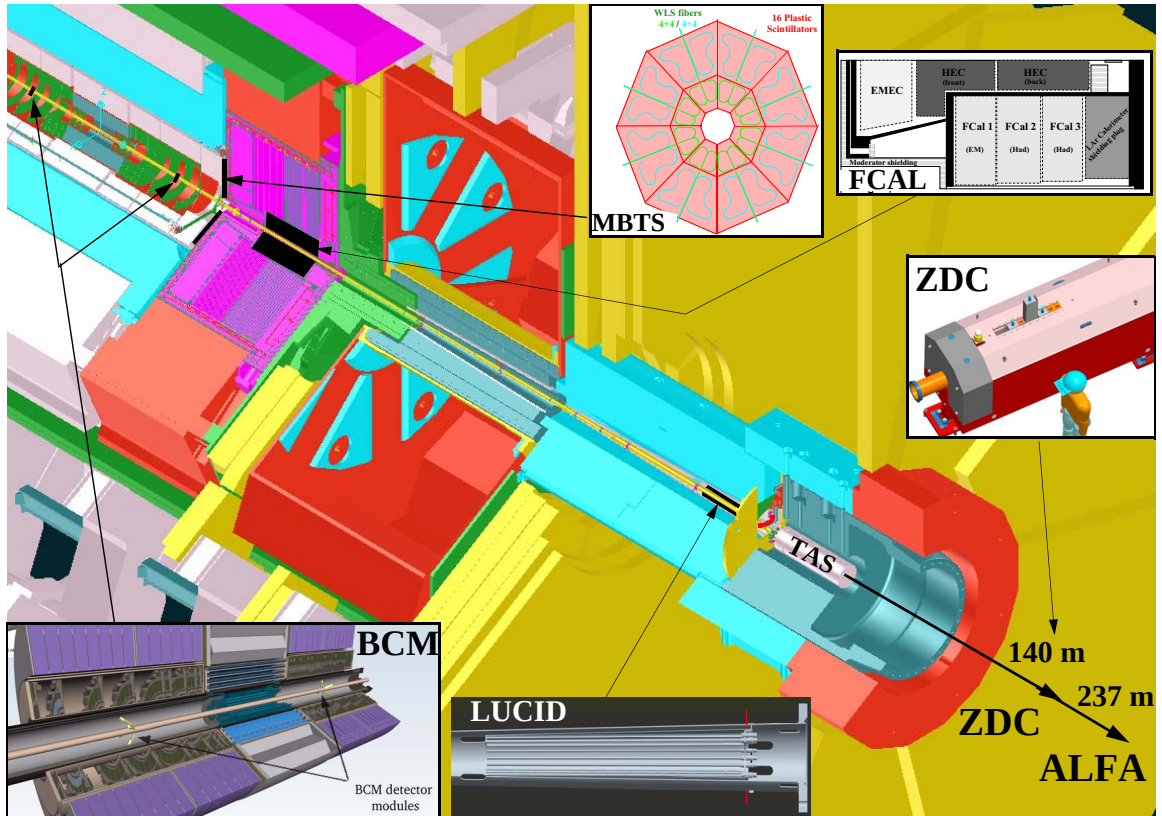


Figure 2.11 – Depiction of the various subdetectors that contribute to the luminosity measurement and calibration in ATLAS. [98]

dition Monitor (BCM) detectors for online measurements of the inelastic proton scattering rate [111, 112]. LUCID comprises a pair of Cherenkov detectors surrounding the beam pipe ± 17 m away from the interaction point. Each detector is made up of 20 aluminum tubes lying parallel to the beam pipe, 1.5 m long and 15 mm in diameter, filled with gaseous C_4F_{10} . As charged particles traverse this

low refractive index gas inside the tube with momentum greater than the Cherenkov threshold, they emit a cone of light that is collected by a PMT at the end of the tube. The response of the detector is very fast, acting as a forward scattering particle counter. The primary purpose of the BCM is to monitor the status of the beam in order to provide critical feedback that can save the ATLAS detector from suffering damage induced by beam losses. The detector consists of four modules symmetrically arranged around and 55 mm away from the beam pipe and ± 1.84 m away from the interaction point. Each module is made of 500 μm polycrystalline chemical vapor deposition diamond, equipped with a very fast readout that permits strong coincidence discrimination between collision and background events. Other subdetectors that contribute to the luminosity calibration of ATLAS include the Minimum Bias Trigger Scintillators (MBTS), with a better acceptance and efficiency that is valuable for low luminosity runs, the Zero Degree Calorimeter (ZDC), which is used primarily for measurements in heavy ion running, the overall current in the forward LAr calorimeter, the photomultiplier current in the hadronic calorimeter, and the number of reconstructed primary vertices in the tracking system. Combining inputs from all these detectors yielded a final luminosity uncertainty of just $\pm 1.8\%$ in 2011 [113]. For a given physics analysis the total integrated luminosity of the dataset is required. This can vary from analysis to analysis since they can impose different data quality requirements, for instance runs taken with the toroid off would affect an analysis measuring muons but not necessarily an analysis measuring jets or photons. The data taken by ATLAS is divided into roughly 1 minute long periods when it is first accepted by the Central Trigger Processor, and these periods are referred to luminosity blocks. The configuration of the detector is constant in each luminosity block and the average luminosity for each block is calibrated offline so that analyses can compute their integrated luminosities.

2.5 The ATLAS Trigger System

The goal of the ATLAS trigger system is to record 1 collision of interest out of every 40,000 delivered by the LHC every millisecond. Given that the rate of interesting events is 10^{14} times smaller than the background rate (10^{-5} Hz vs 10^9

Hz), and the enormous data volume and bandwidth that recording ATLAS data requires, this is a monumental challenge. The ATLAS detector has 100,000,000's of readout channels, with a typical recorded raw event size of roughly 1.5 MB and a total recorded raw dataset of roughly 4 PB/year. This bandwidth must be optimized to best meet the goals of the ATLAS physics program, which is very diverse. Trigger operations require coordination to meet the needs of these different physics groups, to adapt to the run schedule of the LHC and the different types of runs it provides (p - p collisions, heavy ions, van der Meer scans, etc...), and to adapt to the evolving performance of the subdetectors and software that make up the trigger. The trigger system in Run 1 was made up of three levels, as seen in figure 2.12, with each level increasing in granularity and reducing the accepted event rate to the nominal recorded rate of 100 Hz.

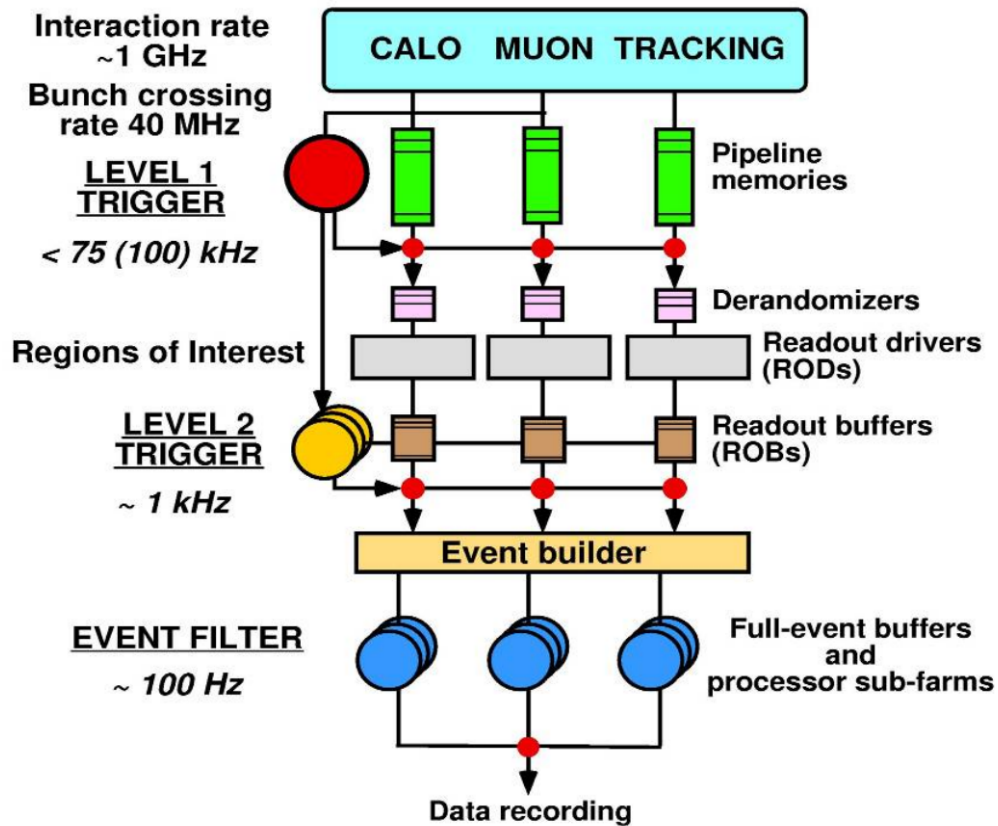


Figure 2.12 – Depiction of the ATLAS trigger system in Run 1 [114].

2.5.1 The Level 1 Trigger

The Level 1 (L1) trigger is hardware based and incorporates signals from the muon and calorimeter systems to select events containing regions of the detector that are good candidates for interesting signals. These regions are referred to as regions of interest (ROIs), and the L1 trigger typically accepts events at a rate of 75 kHz. The L1 trigger employs algorithms that identify high E_T electrons, photons, jets, taus, missing E_T , muons originating from the interaction point, and large total transverse energy. The latency of the L1 trigger must be less than $2.5 \mu\text{s}$ in order to cooperate with the front end readout electronics, and is designed to function with a latency of $2.0 \mu\text{s}$ leaving $0.5 \mu\text{s}$ of buffer. $1.0 \mu\text{s}$ of this time comes from the signal propagation through the readout cables to the processor outside of the cavern. The calorimeter trigger reduces the granularity of the calorimeter by merging cells into 7000 trigger towers, 0.1×0.1 in $\eta \times \phi$ in the central region of the calorimeter and larger in the forward regions. These tower signals are digitized and assigned a bunch crossing ID (time-stamp) by a pre-processor, and then assigned calibrated values of E_T using a look-up table. These trigger tower objects are then sent to two cluster processors, one for measuring jets, missing E_T and total energy, and the other for electron, photon and tau triggers. The muon trigger uses information from the RPCs in the barrel and TGCs in the end caps to identify muons arising from the interaction point that exceed one of three momentum thresholds. This is done by looking for coincidence patterns in "trigger roads", i.e. hits arising in subsequent layers of the detector that fall within predefined shapes and tolerances. As seen in figure 2.13, the Central Trigger Processor (CTP) takes these inputs, applies prescales which throttle signals that would otherwise occupy too much bandwidth, enforces dead time to protect readout buffers, and provides the trigger decisions. Accepted events result in sending an accept signal to the detector front-ends and the DAQ to read out the detector buffers, and the geometric area of the regions of interest for further processing by the Level 2 (L2) trigger.

2.5.2 The Level 2 and Event Filter (High Level Trigger)

The L2 and Event Filter (EF) triggers together constitute the High Level Trigger (HLT). They are similar in that they both run reconstruction algorithms that are

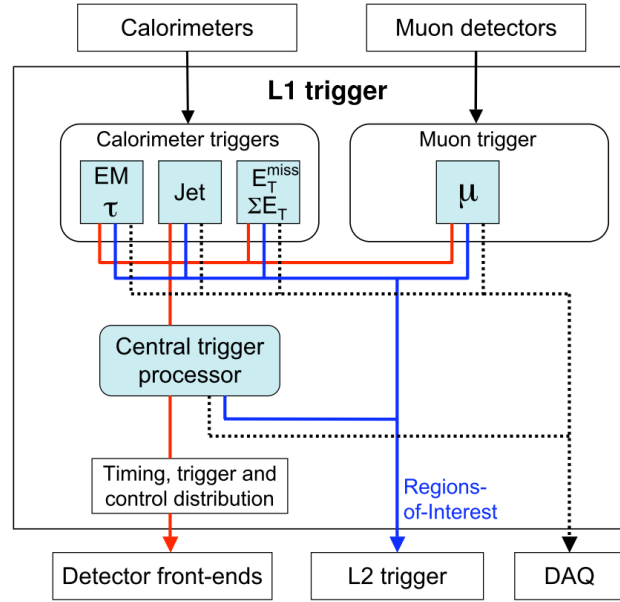


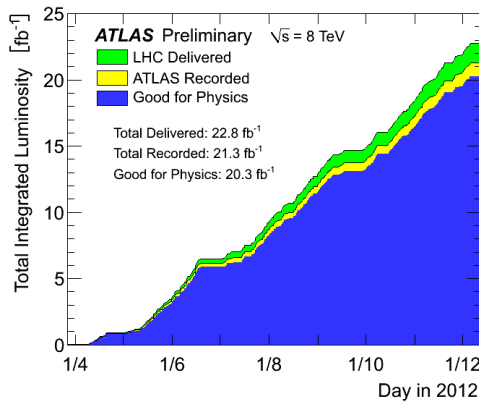
Figure 2.13 – Depiction of the ATLAS L1 trigger system in Run 1 [98].

designed to be as close as possible to the final offline reconstruction. The main difference between the two trigger levels is that the L2 only considers ROIs provided by L1 (which make up roughly 1-2 % of the event size), while the HLT examines the full detector (full scan). The L2 trigger is designed to limit the output rate to about 3.5 kHz while limiting the single event processing time to 40 ms, while the EF is designed to reduce the output rate to 200 Hz with an average single event processing time of four seconds. In contrast to L1, the HLT makes use of the full granularity and the tracking capabilities of the ATLAS detector, providing a much more thorough vetting of interesting events.

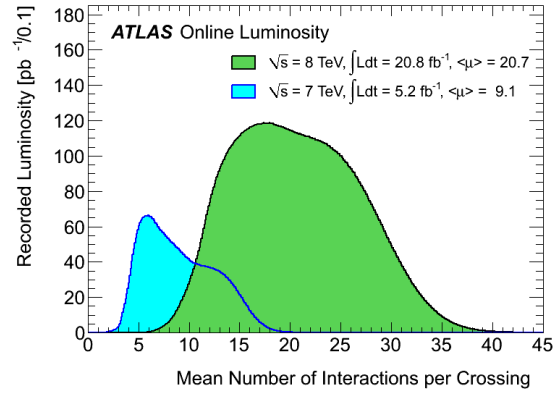
2.6 ATLAS Data Taking in 2012

During 2012 the LHC delivered 22.8 fb^{-1} of proton-proton (pp) collisions at a center of mass energy of 8 TeV. The ATLAS detector recorded 21.3 fb^{-1} with a final 20.3 fb^{-1} good for physics analysis, as seen in figure 2.14a. This performance was achieved while contending with unprecedented energies and higher luminosities, as seen in figure 2.14b displaying the increase in the average number of

collisions per bunch crossing between 2011 and 2012 running.



(a) Delivered, recorded, and good for physics integrated luminosity throughout the 2012 run period.



(b) Comparison of the average number of collisions per bunch crossing measured in each luminosity block between 2011 and 2012 running.

2.7 Simulation

Simulations of the observed physics and the performance of the ATLAS detector play a major role in every physics analysis. From understanding the basic response of detector components, to estimating the effects of dead material, to extrapolating calibrations from reference measurements, to anticipating the characteristics of specific signals, to anticipating and preparing for the effects of new beam conditions (such as higher pileup or luminosity), simulations are heavily relied upon and scrutinized. The framework for ATLAS simulations [115] comprises three main steps: generation of the event and its decays according to physical models in the HepMC format [116], the simulation of the physics interactions in the detector (using GEANT4 4 [117]), and finally the simulation of the detector response and the DAQ system [118]. This full process yields outputs in the same data format as real runs (Raw Data Objects), accompanied by the information detailing the provenance and path taken through the framework. This latter information containing the provenance is referred to as “truth” information, and is a primary handle used to measure and correct for detector effects. This process is computationally demanding and thus makes use of the CERN computing grid. Cam-

paings for generating ATLAS simulations are generally started months prior to data-taking while resources are not being used for data reconstruction. Parameters relevant to the beam conditions are chosen in anticipation of the run program, and following the data taking corrections that are applied to account for any differences. Each of these steps takes place within the ATLAS Athena framework, a compiled C++ and python steered framework developed by ATLAS that is based on the Gaudi [119] framework of LHCb, making use of the CLHEP high energy physics analysis libraries [120].

Generation of simulated events in ATLAS comprises the following main steps, visualized in figure 2.15.

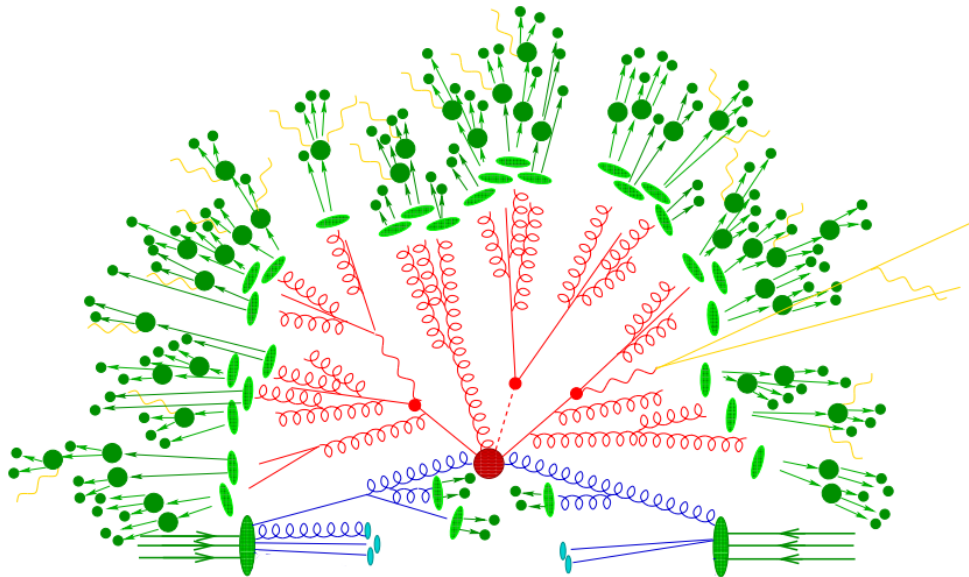


Figure 2.15 – Components of a simulated event from [121], where red dots represent the hard collision with blue input partons and red output partons. The red branching corresponds to the PS, and the green branching to the hadronization. Yellow lines correspond to soft photon radiation.

2.7.1 Hard Scatter

The production of an ATLAS detector simulated event begins with the generation of a set of outgoing partons arising from a theoretical Matrix Element (ME)

calculation for the production mechanism of interest. The incoming initial state particles and momenta are given by the chosen Parton Distribution Function (PDF) which takes into account the beam energy and particle type. ME computations are typically denoted by their order, Leading Order meaning the simplest possible interaction with one propagator and two vertices. LO calculations always yield real terms and in turn positive definite results. Higher order calculations consider additional virtual terms that can yield negative contributions and divergences that must be canceled by the additional real terms included at the higher order. For this reason higher order ME calculations require care when considering the downstream treatment of the outgoing particles, since they can in fact carry negative event weights and the relative treatment and cancellation of these events is critical. The outgoing final state particles from the ME calculation are considered to originate from a single primary vertex and must have a lifetime greater than $c\tau > 10$ mm, since particles with shorter lifetimes can be decayed by the generator without considering the subsequent detector interactions.

2.7.2 Parton Shower

The final state colored particles arising from the hard scatter are next input to a parton shower (PS) algorithm to simulate gluon Bremsstrahlung and the resulting shower of secondary coloured particles. The shower proceeds until the particle energies approach the specified hadronization scale (≈ 1 GeV). The transition from ME to PS must be done with care to avoid double counting and gaps in the phase space. This is not always trivial since the purpose of the PS is to approximate higher order effects that are beyond the ME, thus it is clear that the usage case differs between two ME calculations of different orders. The behavior of these showers is leading logarithmic, following the DGLAP evolution equations used to model the scale dependence of PDFs and fragmentation functions [42,43]. Different approaches to calculating the branching probabilities are advantageous for different usage cases. Two common approaches are p_T ordering and θ^2 ordering. Examples of Monte Carlo (MC) parton shower event generators include PYTHIA [82], SHERPA [122] and Herwig [123]. The parameters of the PS algorithm are often implemented using specific tunes that have been determined to be optimal

for ATLAS' purposes [124].

2.7.3 Hadronization and Decay

Following the parton shower the outgoing partons are evolved into colour-singlet hadrons, in accordance with QCD confinement. This lower energy part of the evolution is carried in the non-perturbative regime of QCD necessitating the usage of phenomenological models [125]. Two of the most common hadronization models are the Lund string model [126] and the cluster model [127]. The Lund string model, illustrated on the left of Figure 2.16, is based on the principle of linear confinement [128]. Linear confinement is the observation, from both lattice QCD computations and from quarkonia spectra, that the potential of the colour field between two separated colour charges grows linearly for distances beyond a femtometer. In this model as coloured charges are separated they remain connected by a coloured string with tension constant of $\kappa \approx 1 \text{ GeV/fm}$. When the string tension becomes large enough the string breaks producing quark anti-quark pairs. This string breaking is modeled using quantum tunneling of a Gaussian distribution with respect to the transverse mass in the rest frame of the diquark system. The addition of gluons into this string breaking produces "kinks" giving rise to transverse structure. The Lund string model is implemented in PYTHIA. The cluster model, illustrated on the right of Figure 2.16, instead relies on the principle of preconfinement [128]. Preconfinement is the observation that colour clusters display a universal invariant mass distribution parametrized by scale of the interaction and Λ_{QCD} . The cluster model treats gluons from the parton shower by immediately splitting them into quark anti-quark pairs with a mass distribution following previous observations. The resulting quarks are then immediately clustered to colour-connected neighbors. These clusters then decay isotropically according to their quantum numbers and density of states, naturally producing transverse structure and providing a mechanism for suppression of heavy meson and baryon production. A cluster model is implemented in SHERPA.

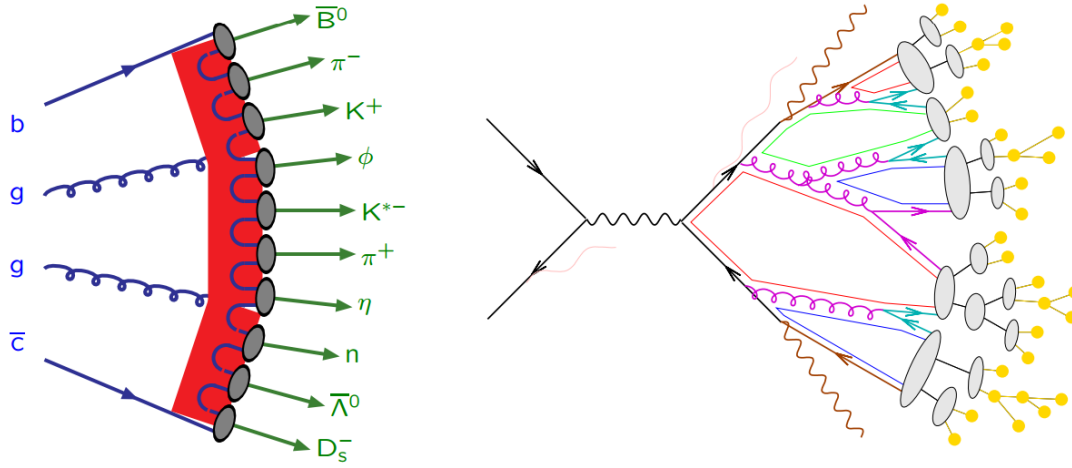


Figure 2.16 – String (left) vs cluster (right) hadronization [129].

2.7.4 Pileup

In ATLAS, signals arising in the detector due to processes separate from the primary hard scatter of interest are referred to as pileup. These signals must be measured, understood and accounted for when extracting physics signals, representing one of the principle challenges of extracting physics signals from the LHC environment. A good measure of pileup is the average number of collisions per bunch crossing referred to as μ . As seen in figure 2.14b, the LHC provided ATLAS with μ between 5 and 15 during the 2011 7 TeV run period, and with a μ between 10 and 35 during the 2012 run period due to the increase in energy and luminosity. This represents a significant increase in pileup effects. Pileup signals can arise from various sources: in-time pileup which includes interactions from other protons in the same bunch crossing, out-of-time pileup which includes beam remnants and leftover signals from earlier bunch crossings, the cavern background of neutrons and photons creating signals in the muon system, beam halo effects from the beams interacting with upstream collimators producing sprays of muons, and beam gas effects from residual gas in the beam pipe interacting with beams [130]. These effects are accounted for in the simulation by tuning parameters in overlaying signals measured in minimum bias events, that is events recorded using a random trigger that captures these non-collision related backgrounds.

2.7.5 Detector Simulation and Reconstruction

The final state particles produced by the MC event generator are processed by a full simulation of the ATLAS detector and readouts using the GEANT4 4 toolkit [117]. The detector simulation then interfaces directly with a simulation of the DAQ system producing RAW files in the same format as the data coming from the detector and trigger system online. This RAW data is processed in the same manner as the data using a release of the ATLAS production reconstruction software yielding output files in the same format as that of an online run, with the additional truth information for reconstructed objects.

Object Reconstruction

In the framework used to analyze ATLAS data, detector signals are identified, categorized and calibrated as analysis “objects”. The goal of constructing these objects is to yield measurable quantities that can bear a meaningful correspondence to theoretical predictions. The process of deriving these objects from the ATLAS detector data is referred to as “reconstruction”. The software that takes raw detector data and reconstructs these “physics” objects is computationally demanding, due both to the sheer volume of data in each event and due to the sophisticated algorithms that are used. The software is developed by hundreds of developers across different physics and performance working groups. Throughout and following the period of LHC and detector operation the reconstruction software is updated to include improvements that correct bugs and improve calibrations.

Extracting a physics signal from the ATLAS detector data requires sifting through real and simulated data samples to build collections of calibrated objects of interest. The signal targeted by this measurement is the production of a prompt-photon in association with a jet containing a charm quark.

The following sections describe in detail the relevant photon and jet object reconstruction and calibration techniques used for the 2012 $p\text{-}p$ dataset at $\sqrt{s} = 8$ TeV, referred to as “Run 1”.

3.1 Photon Reconstruction

In ATLAS photons are well measured by virtue of their characteristic electromagnetic shower and the tracking signatures. As a result reconstructed photon objects bear an excellent correspondence to theoretical particle-level photons.

3.1.1 Photon Identification

The identification of photon objects is described in detail in reference [131]. Photon reconstruction in ATLAS begins by the identification of an EM calorimeter

cluster. A sliding window algorithm is used to find clusters of size $\Delta\eta \times \Delta\phi = 0.075 \times 0.123$ in $|\eta| < 2.5$ with energies exceeding 2.5 GeV. It has been shown using simulations that the efficiency of this initial cluster finding is nearly 100% for photons with $E_T^\gamma > 20$ GeV. These clusters are used in parallel to reconstruct both electron and photon candidates by extrapolating tracks in the ID to the second layer of the EM calorimeter and associating them to the cluster barycenter [131]. Reconstructed photons fall into two categories based on their associated tracking information. “Unconverted” photons are photons that do not interact with the inner detector and leave no track. “Converted” photons are photons that interact with the inner detector and leave a characteristic electron-positron track pair vertex. Unconverted photons are identified and reconstructed early on in the event reconstruction process and stored in the dedicated photon stream, while converted photons are recovered from the electron stream at a later stage. The recovery of converted photons from the electron stream relies on the track topology as well as information from hits in the TRT which help ensure that tracks are indeed electron tracks (based on the relativistic dependence of transition radiation).

The signal photons targeted by the photon identification algorithms in ATLAS are those that arise from a hard process and not from a hadron decay, which incidentally exhibit much larger cross sections. Photon candidate clusters are subject to a series of cut-based criteria that have been optimized to select prompt signal photons and reject background photons, outlined in table 3.1. These criteria rely on the size and shape of the longitudinal and lateral shower shape, thereby taking advantage of the characteristically narrow showers and smaller hadronic leakage exhibited by signal photons in contrast to background photons arising due to hadron decays and in association with jets. Background photons arising from electromagnetic decay of a $\pi^0 \rightarrow \gamma\gamma$ are targeted by the criteria that are sensitive to bi-modal deposits that would arise from a pair of decay photons, as illustrated in figure 3.1.

The photon ID variables are used to define two standard reference photon ID criteria, *loose* and *tight*, that are calibrated by dedicated performance studies [132]. The *loose* criterion, which uses only the hadronic leakage and EM middle layer quantities, exhibits a prompt-photon efficiency greater than 99% for $E_T^\gamma > 40$ GeV with a fake-photon rejection factor of 1000 and is thus used for triggering. The

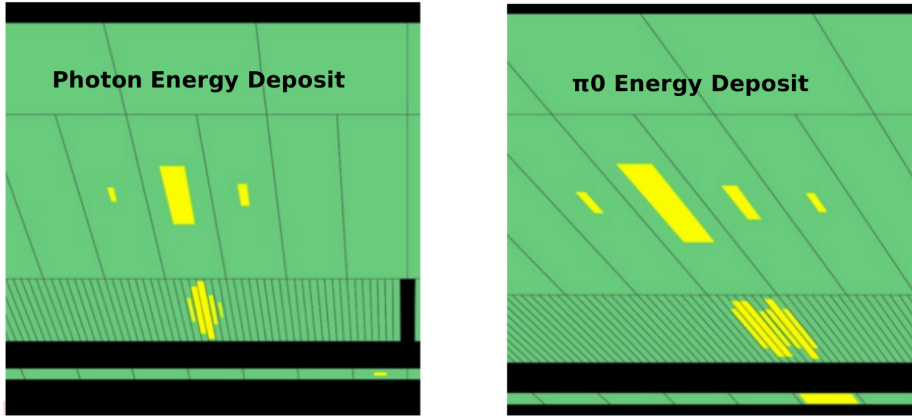


Figure 3.1 – Comparison between the energy deposited in the EM calorimeter by an isolated photon and an isolated π^0 meson decaying to two photons.

tight criterion uses the additional information from the EM strip layer resulting in a prompt-photon efficiency of about 85% for $E_T^\gamma > 40$ GeV with a rejection factor of 5000. The performance of the photon reconstruction efficiency was measured using the full 8 TeV dataset using three separate data-driven methods [131].

- The radiative $Z \rightarrow ll\gamma$ method (two lepton+photon) applies kinematic constraints that target instances of this decay with an isolated photon.
- The electron extrapolation method is based on the close correspondence between electron and photon signals in the detector using $Z \rightarrow ee$ events to derive a clean selection of electrons.
- The matrix method estimates the efficiency by estimating the background contamination based on background enhanced control regions.

The three methods exhibited consistent results, finding that photon identification efficiency rose from roughly 50% at $E_T^\gamma = 10$ GeV to above 90% at $E_T^\gamma > 40$ GeV and to above 94% for $E_T^\gamma > 100$ GeV [131]. The shower shapes of the simulations are corrected based on performance studies comparing simulations to photon-enriched data samples.

In this analysis a custom variation of the *tight* criterion is used, referred to as *relaxed tight*. The *relaxed tight* applies the same criteria as the *tight* while relaxing the requirement on the F_{side} , $w_{s,3}$, ΔE and E_{ratio} variables. These relaxed variables are those defined using the first-layer of the EM calorimeter that are aimed at resolving the lateral shower profile and discriminating against bimodal signals

arising from hadronic decays. A further criterion, referred to as *reversed-tight*, requires the failure of at least one of the four criteria and is used in the analysis to create control regions enriched with backgrounds.

3.1.2 Photon Energy Calibration

The full procedure used for calibrating photon energies is displayed in figure 3.2 and described in detail in reference [133]. The initial photon calibration procedure relies on test beam calibrations of the LAr calorimeter, cluster correction factors derived from simulations that account for dead material and leakage based on schematics of the detector construction, and the measurement of the Z boson resonance in the data to set the absolute energy scale [134]. The procedure was subsequently improved, by roughly 10% for unconverted photons and 20% for converted photons, by including additional data driven studies of dead material using multivariate techniques that take advantage of characteristic cluster shapes. The full calibration procedure is illustrated in figure 3.2.

1. Cluster calibration constants are derived from simulations using multivariate techniques optimized to the specific object (unconverted photon, converted photon, electron). The description of the dead material in the simulation is improved by incorporating measured distributions of the ratio of the first layer to the second layer measured energy in the EM calorimeter.
2. The relative calibration between longitudinal layers is equalized in data, ensuring the inter-layer calibrations remain coherent when extrapolated over the full energy scale range.
3. The overall calibration is applied using the outputs of steps 1 and 2.
4. Run specific corrections are applied to account for inhomogeneities such as dead modules or malfunctioning high voltage regions.
5. The data is calibrated to agree with the simulation of $Z \rightarrow ee$ events, while the simulation is smeared to match the data.
6. Finally the calibration is validated using real $Z \rightarrow ee$ and $Z \rightarrow ll\gamma$ events in data.

The final photon energy scale uncertainty is typically on the order of 0.4% and valid up to 500 GeV, beyond this point extrapolation uncertainties come into play. The data calorimeter energy response is stable, as a function of time and pile-up, to within 0.05% in the barrel and 0.75% in the end cap.

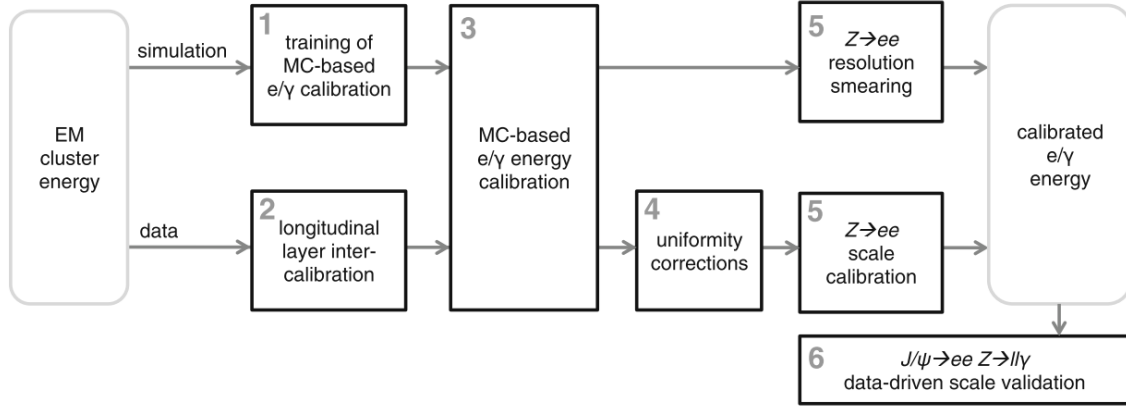


Figure 3.2 – Flow chart of the electron/photon energy calibration procedure [133].

3.1.3 Photon Isolation

Photon isolation, which plays a key role in this and many other photon analyses, is a measure of the energy in the immediate vicinity of the photon in the calorimeter. Applying a selection cut based on this criterion imposes a strong phase space requirement that targets signal photons arising from the hard scatter instead of background photons that can either be fake-photon signals or real fragmentation photon signals. Fragmentation photons are photons that arise from the Bremsstrahlung of a hadronic signal and carry a large portion (fragment) of the initial object's momentum.

At the detector-level the photon isolation energy is calculated by taking the sum of the transverse energy of topo-clusters (described in section 3.2.2) falling within a cone of $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.4$, excluding the energy falling within a 5×7 window of second-layer EM calorimeter cells centered on the photon, as depicted in figure 3.3. The central cells are removed to exclude the energy from the photon itself. Corrections are then applied to correct for leakage of the photon energy into the isolation energy region and to account for ambient energy arising

from pileup [135]. The ambient energy correction is derived using a data driven jet area subtraction technique [136]. This technique takes into account energies and areas of the ensemble of jets in the event to estimate the ambient energy density. This energy density is then multiplied by the photon isolation area and subtracted from the isolation energy.

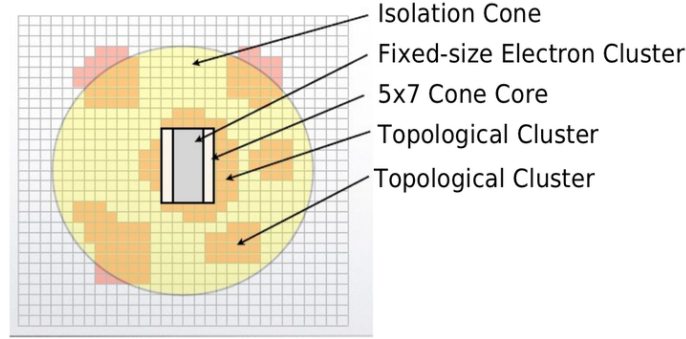


Figure 3.3 – Depiction of the topology of the topo-cluster isolation algorithm [137].

Photon isolation is calculated for truth photon objects with an analogous procedure, taking the sum of all long-lived particles falling within a cone of 0.4 of the photon, excluding muons and neutrinos. The ambient energy correction is calculated by building jets using the anti- kT algorithm with a distance parameter of 0.5 considering all long-lived particles in the event, again excluding muons and neutrinos. The energy density correction is then applied by subtracting the energy density multiplied by a circular area of radius 0.4.

Category	Description	Name	Loose	Relaxed Tight	Tight
Acceptance	$ \eta < 2.37$, with $1.47 < \eta < 1.52$ excluded	-	✓	✓	✓
Hadronic leakage	Ratio of E_T in the first sampling of the hadronic calorimeter to E_T of the EM cluster (used over the range $ \eta < 0.8$ or $ \eta > 1.37$)	R_{had1}	✓	✓	✓
	Ratio of the E_T in the hadronic calorimeter to E_T of the EM cluster (used over the range $0.8 < \eta < 1.37$)	R_{had}	✓	✓	✓
EM Middle layer	Ratio of energies in $3 \times 7 \eta \times \phi$ cells over 7×7 cells	R_η	✓	✓	✓
	Lateral width of the shower	$w_{\eta 2}$	✓	✓	✓
	Ratio of energies in 3×3 cells over 3×7 cells	R_ϕ		✓	✓
EM Strip Layer	Shower width calculated from three strips around the strip with maximum energy deposit	w_{s3}			✓
	Energy outside the core of the three central strips but within seven strips divided by energy withing the three central strips	F_{side}			✓
	Difference between the energy associated with the second maximum in the strip layer and the energy reconstructed in the strip with minimal value found between the first and second maximum	ΔE			✓
	Ratio of the energy difference associated with the largest and second largest energy deposits over the sum of these energies	E_{ratio}			✓

Table 3.1 – Photon identification shower shape variables [131].

3.2 Jet Reconstruction

Compared to photons, a different approach is taken when creating reconstructed objects that capture hadronic signals. Due to the more complicated nature of hadronic shower development, hadronic signals are reconstructed as so called jet objects. Jets are clusters of hadronic energy constructed by algorithms specifically chosen to provide a reliable correspondence between all levels of hadronic signal evolution, illustrated in figure 3.4. Jets can be built at each stage of hadronic evolution in a simulation using the same algorithm considering the constituent particles at that stage. In the messy hadronic environment provided by the LHC, with multiple scatters in each bunch crossing and other pileup effects, it is important that the jet clustering algorithm is “infrared-safe”, that is, insensitive to the effects of soft radiation. It is also important that the algorithm is “collinear-safe”, that is it is insensitive to collinear splitting and in turn to the choice of parton shower and hadronization treatments. These two features are illustrated in figure 3.5. The anti-kT algorithm satisfies both these criteria and is the algorithm of choice at the LHC [138]. The algorithm has a highly optimized implementation in the Fastjet library that is widely used in reconstruction software.

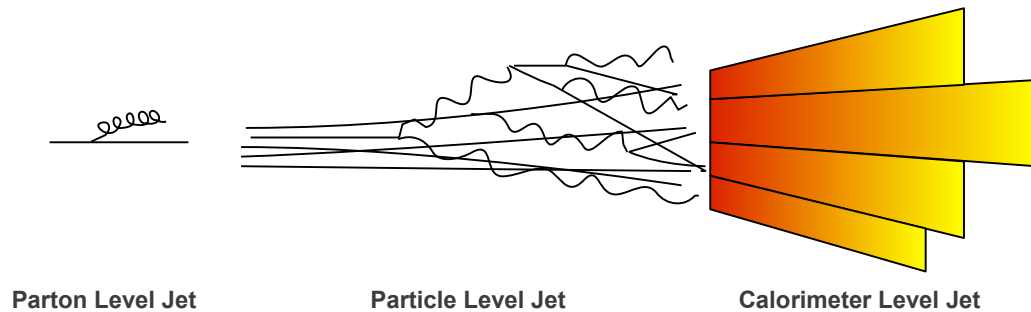


Figure 3.4 – Illustration of the different stages of jet evolution.

In ATLAS there are four principal collections of jets that can be used for analysis, depicted in figure 3.6. This analysis makes use of truth jets and LCW calorimeter jets, which are described in detail in the following sections. These

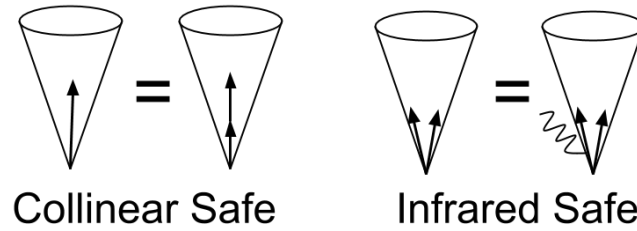


Figure 3.5 – Illustration of a jet that is insensitive to collinear splitting, left, and soft radiation, right.

detector-level jets are built using the anti- k_t algorithm [138] with a distance parameter $R = 0.4$, using as input locally calibrated topological calorimeter clusters. The general calibration and cleaning procedures follow the recommendations for the final 2012 ATLAS dataset.

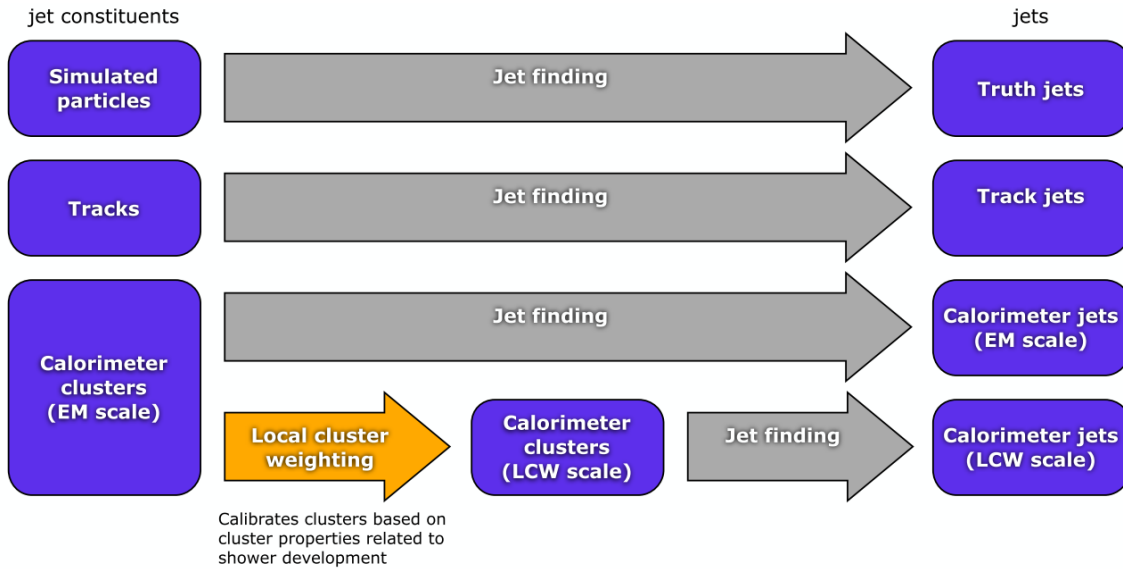


Figure 3.6 – Overview of the different jets reconstructed in ATLAS [139].

3.2.1 The anti- k_T Algorithm

The k_T algorithm is a versatile jet clustering algorithm which can be tuned to be used in either *messy* or *clean* collider environments [78, 140]. Clean environ-

ments refer to collisions with point-like fundamental particles (such as electrons) resulting in hard interactions in a low energy ambient environment. Messy environments refer to collisions with composite particles, such as protons or heavy ions, where in each collision there is a baseline of soft interactions and collisions occurring simultaneously. The goal in a clean environment is to measure all the energy arising from the interaction for precision measurements, whereas in messy environments it is important to distinguish interesting signals from the background.

In the algorithm description kT is the energy deposited in a detector cell, ΔR is the distance between two data points, R is the input cluster radius, and p is an integer set to -1, 0 or 1. The unseeded version of the kT algorithm proceeds as follows:

Step 1

Calculate the similarity between all points:

$$S_{ij} = \min(kT_i^{2p}, kT_j^{2p}) \cdot \frac{\Delta R_{ij}}{R}.$$

Step 2

For each point calculate similarity to the “beam”:

$$S_{iB} = kT_i^{2p}.$$

Step 3

Find the lowest value of S . If it is S_{ij} , combine points i and j . If it is S_{iB} , remove i from the data set and store it in memory.

Step 4

Return to step 1 until all points have been removed.

The value of p determines the nature of the algorithm. For $p = -1$ the algorithm is referred to as the anti-kT algorithm. In this version of the algorithm only data points with a large amount of energy within the specified radius ($\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$) are clustered. This behavior lends itself to messy environments since it naturally builds tight clusters and discards low energy background signals. For $p = 1$ the algorithm is referred to as the kT algorithm. In this case the algorithm is more greedy, with low energy cells within the radius being absorbed into the

clusters. This version of the algorithm is preferred for clean environments where it is important for the object to measure all of the particle's energy. The anti-kT algorithm is computationally much more efficient than the kT algorithm. If $p = 0$ then the algorithm has no energy dependence, making it a simple cone algorithm. Figures 3.7 and 3.8 display clustering examples of the kT and anti-kT algorithms.

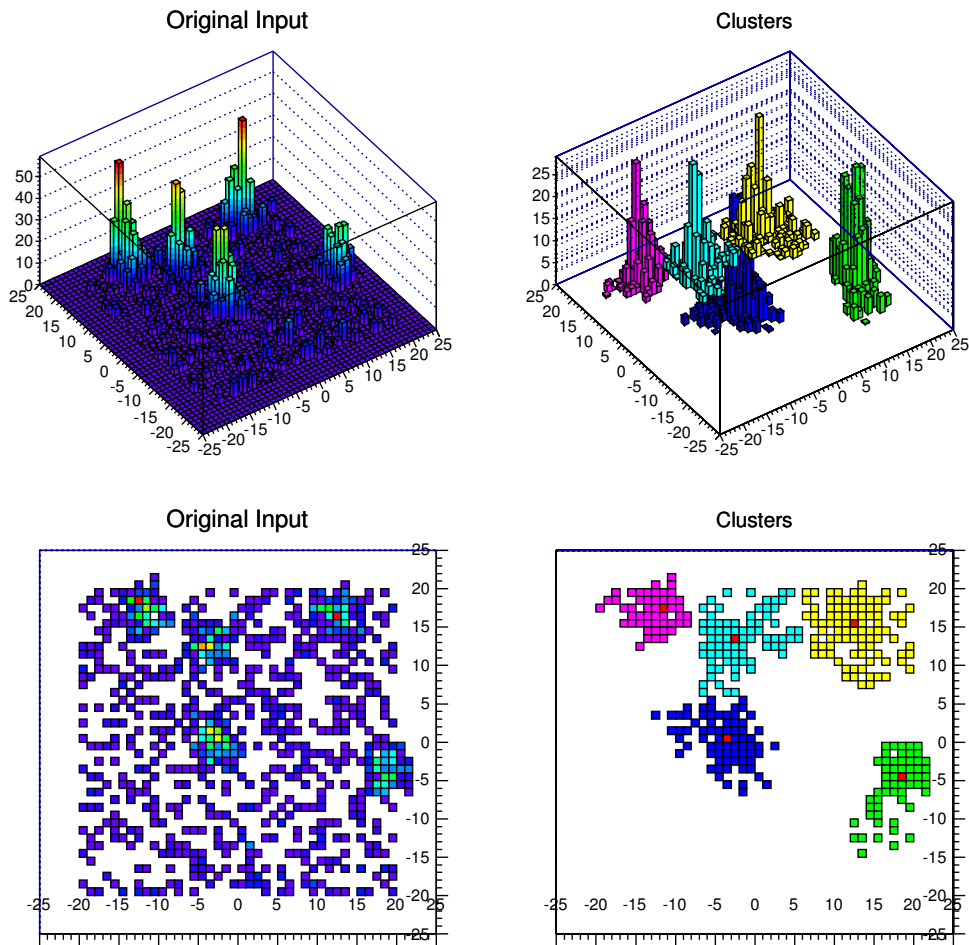


Figure 3.7 – Illustration of clustering of the kT algorithm with $p=1$. All axis units are arbitrary.

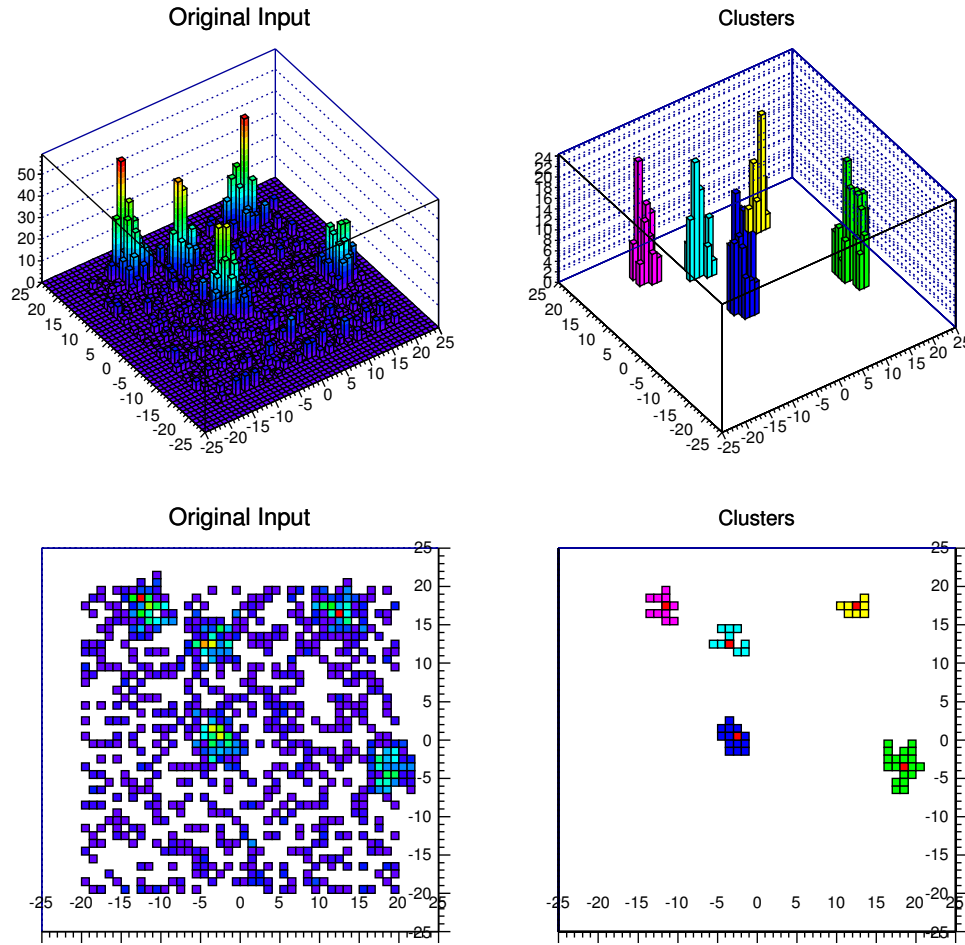


Figure 3.8 – Illustration of clustering of the anti- k_T algorithm with $p=-1$. All axis units are arbitrary.

3.2.2 Local Cluster Jet Reconstruction

The local cluster jet energy calibration is described in detail in reference [141]. When building jets using previously clustered calorimeter cells the initial clustering performs a baseline noise suppression that can adapt to different pileup conditions. The clusters are built considering a cell significance defined as the ratio of the cell's default calibrated energy to the expected noise in that cell given the run conditions. In Run 1 the expected noise was derived from simulations and parametrized as a function of the average number of interactions per bunch

crossing. Clusters are built iteratively considering the most significant cell with a significance greater than 4 as seeds, then subsequently adding all adjacent cells with a significance greater than 2, and finally adding all the remaining adjacent cells with a significance greater than 0. This process is repeated until no remaining cells meet the initial seed criteria. Following the cluster building a cluster splitting is performed with the aim of resolving signals arising from two-body decays of boosted systems. The final cluster four vector is computed by adding the sum of all the constituent cells to derive the energy, and using the energy weighted centroid to derive the ray traced by the four vector. The cluster energy is then calibrated using the so called Local Calibration Weighting (LCW) scheme. The first step in this calibration is to classify the cluster as either hadronic or electromagnetic based on its longitudinal profile. The next step applies a correction to hadronic clusters to compensate for their characteristically lower response. Next, all clusters have a correction applied to account for energy in adjacent cells and in the hadronic tail. The final step applies corrections for dead material.

3.2.3 Jet Identification

The identification and classification of jet objects is described in detail in Ref. [142]. Similarly to photons, the goal of the jet identification criteria is to select jets arising from the hard scatter. The principal sources of background jets include beam gas signals arising from beam collisions with residual gas in the beam pipe, beam halo interactions arising from interactions between the beam and collimating optics upstream of the detector, cosmic muons coincident with a bunch crossing, and calorimeter noise. Variables constructed to discriminate against beam induced backgrounds make use of the jet shower shape, how closely the four vector associated to the shower points to the interaction point, and the proportion of the jet momentum that is represented by the tracks associated to the jet compared to the energy measured in the calorimeter. Variables constructed to discriminate against calorimeter signal noise measure how well the individual cell signals match the expected signal shapes. Using these types of discriminating variables, four levels of jet identification are used corresponding to increasing background rejection and signal purity. The loosest level of jet identification, referred to as the "looser" cri-

terion, is used in this analysis. The “looser” criterion ensures that the signal acceptance is above 99.8% while imposing a fake-jet rejection factor of roughly 50%. The performance of the different jet identification criteria is evaluated using tag and probe analyses using dijets.

3.2.4 Jet Energy Calibration

The jet energy calibration procedure for Run 1 is described in detail in Ref. [139] and depicted in figure 3.9. In most analyses making use of jets the jet energy scale (JES) calibration uncertainty can be the most significant uncertainty due to the inherent challenges of measuring hadronic signals. The principal steps in the jet energy calibration procedure are as follows:

1. A pile-up correction derived from simulations as a function of the number of primary vertices, the average number of interactions per bunch crossing, p_T^{jet} and η^{jet} is applied.
2. The origin of the jet four-vector is adjusted to the location of the associated primary vertex instead of the origin of the ATLAS detector.
3. A simulation-based jet response correction is applied, derived from inclusive jet samples and parametrized as a function of the jet energy and η , taking the form $\mathfrak{R}^{LCW} = E_{jet}^{LCW} / E_{jet}^{truth}$ [142].
4. A final correction is applied based on in-situ comparisons between data and simulations using tag and probe techniques. Jets in the central region with p_T^{jet} up to 800 GeV use Z bosons and photons as tags, while higher p_T jets use composite systems of lower p_T jets. In the forward region dijet systems are used considering a central jet as the tag and a forward jet as the probe.

For jets with $p_T > 100$ GeV the JES uncertainty is typically 2-4%.

3.2.5 Jet Pile-up Corrections

As described in section 3.2.2, pile-up dependent noise suppression is implemented at the calorimeter cell readout level when clusters undergo their initial

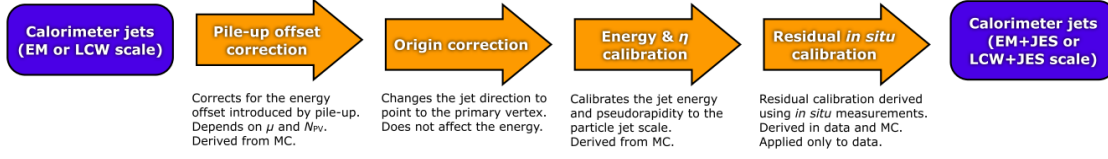


Figure 3.9 – Overview of the jet energy calibration procedure [139].

energy calibration. Additional jet-level selection criteria are implemented to further reject pile-up induced jet objects, and additional corrections are implemented to correct signal jets for other pileup influences at the object level, such as the jet energy or shape. These corrections and selection criteria are described in detail in Ref. [130]. The Jet Vertex Fraction (JVF), illustrated in figure 3.10, is a powerful observable for rejecting pile-up induced jets. The JVF of a jet is calculated as

$$JVF(jet_i, PV_j) = \frac{\sum_m p_T(track_m^{jet_i}, PV_j)}{\sum_n \sum_i p_T(track_m^{jet_j}, PV_n)}, \quad (3.1)$$

where m sums over all tracks matched to jet_i originating from the primary vertex PV_j , n sums over all primary vertices in the event, and l sums over all tracks matched to jet_i originating from the primary vertex PV_n . In this calculation only tracks with $p_T > 500$ MeV are included. The resulting quantity is bounded by 0 and 1, with values close to 1 indicating that the jet has a strong association to the primary vertex and thus is unlikely to be a pileup jet. Jets with no associated tracks get assigned a JVF of -1.

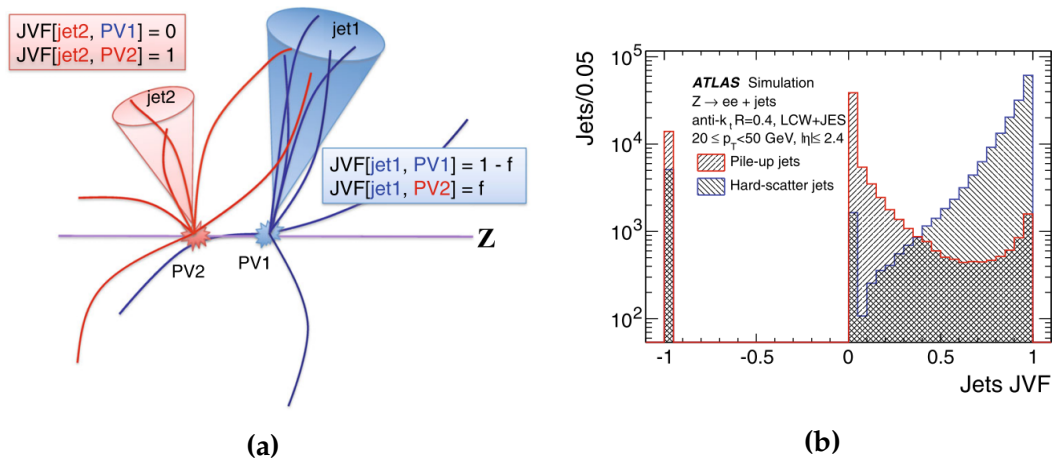


Figure 3.10 – **a)** Visualization of the jet vertex fraction observable, where f is the fraction of the track p_T associated to the jet1 from the adjacent vertex PV2. **b)** The characteristic JVF distributions for signal and background jets measured in simulated $Z \rightarrow ee + jets$ events, following pile-up subtraction [130].

3.2.6 Reconstruction-Level Flavour Tagging

The goal of this analysis is to measure the production cross section of jets containing a c -quark in association with a photon. Jet algorithms are designed such that when they are used at different levels, i.e. parton, hadron or calorimeter, they yield consistent results. Reconstruction-level jets are constructed by clustering constituent energy deposits, and their properties are derived from the ensemble of deposits making up the jet. The detector has limited resolution and is thus not capable of fully resolving the jet structure. For most purposes this is not a problem since often only the kinematics of the jet are of interest, and the four-vector of the detector-level jet bears a correspondence to analogous parton-level jet. Flavour tagging a jet at the detector-level, that is determining the presence of a HF-quark within the jet, is not a straightforward task due to the limited resolution of the calorimeter. Ideally one would want to accurately measure the full evolution of the jet by resolving and identifying the particles throughout the evolution from the initial parton to the calorimeter deposits. Though this is currently not possible, c - and b -quarks produce characteristic features in both the ID and the calorimeter by virtue of their lifetime, mass and decay topologies. Heavy quarks decay via the weak force with lifetime on the order of ps (10^{-12} s). As depicted in figure 3.11, the heavy quark that is produced at the primary vertex will travel some distance, and when it decays it will produce a secondary vertex. This distance in the lab frame from the primary to the secondary vertex, referred to as the decay length L_{HF} , is computed taking into account the relativistic correction moving from the quark's frame $L_{HF} = \beta\gamma ct$, where c is the speed of light, t is the decay time, $\beta = \frac{v_{HF}}{c}$ and $\gamma = \frac{1}{\sqrt{1-\beta^2}}$ is the relativistic Lorentz factor. Typical decay lengths for heavy flavour hadrons are on the order of mm, and, as described in section 2.4.2, the ATLAS ID is designed to resolve these topologies. The other main feature of these secondary HF-decays is the presence of the lepton arising from the decay of the W boson responsible for the flavour change. Making use of the presence of this lepton within the jet for tagging purposes is referred to as "soft-lepton" tagging. "Lifetime" based tagging relies on the measurement of the displaced HF-vertex and its decay topology. In this analysis lifetime based approaches are considered.

In practice, jet flavour tagging involves applying an algorithm that yields a

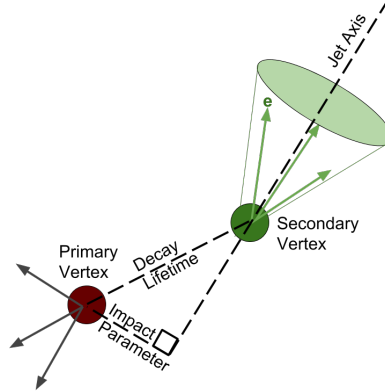


Figure 3.11 – Illustration of a jet with a displaced vertex, indicative of a heavy quark decay.

discriminating variable that is sensitive to the flavour content of the jet, referred to as the *tag weight*. In lifetime based tagging there are three principal features of the decay topology that are used: the presence and location of the secondary vertex, the impact parameter of associated tracks, and the kinematics of the decay products. In ATLAS there is a tagging algorithm that targets each of these features, described in detail in reference [143].

- **SV1**: This algorithm finds two-track secondary vertices associated to the jet [144]. The algorithm uses a likelihood ratio formalism [145] that considers vertex mass, the relative energies of the tracks associated to a vertex against all the tracks associated to the jet, the number of vertices, and the separation between the line joining the primary and the secondary vertex and the four vector of the jet. Vertices reconstructed with a mass compatible with K_0 and Λ decays and photon conversions are rejected, and only vertices exhibiting an acceptable χ^2 are considered. Secondary vertex reconstruction provides strong discrimination but suffers from a roughly 70% vertex finding efficiency.
- **IP3D**: This algorithm is an impact parameter algorithm that uses 2D longitudinal vs. transverse track impact parameter distributions in a likelihood ratio technique to derive jet flavour probabilities [143]. The starting template probability density functions are based on smoothed distributions from MC.

- **JetFitter**: This algorithm uses a Kalman filter to find a ghost track, that is the trajectory of the initial decaying heavy particle, by minimizing the distance of tracks intersecting this “ghost track”. The advantage of this algorithm lies in its treatment of tertiary vertices arising in the $b \rightarrow c \rightarrow l$ decay chain. The assumption is that the c decay will produce negligible transverse momentum relative to the initial b , resulting in the secondary and tertiary vertices falling along this same trajectory.

The combined discriminating power of these algorithms is then harnessed by combining their discriminants in a neural network. This neural network is then trained using fully reconstructed simulated samples to build an algorithm that optimizes the desired discrimination between jet flavours. In ATLAS there are three principal neural network taggers that each make use of the SV1, IP3D and JetFitter algorithm discriminating variables:

- **MV1**: Neural network trained to identify b -jets and reject light jets, the current default ATLAS tagger.
- **MV1c**: Neural network trained to identify b -jets and reject c -jets.
- **JetFitterCharm** [79]: Neural Network that is trained to identify c -jets and reject b - and light jets.

Taggers are used in analyses by either applying a cut on the tagger discriminant, to increase the purity of a particular flavour, or by fitting a tagger discriminant distribution to extract the purity. The first usage involves the application of a cut, where the cut value is referred to as the “working point”. The second usage is referred to as the “continuous” usage, where the full discriminant distribution is used. In both scenarios the tagger usage hinges on a measurement of the tagger performance using the simulation to estimate the effect on the data. To do so, the tagger must undergo a calibration that accounts for differences between the performance in data and MC. These calibrations are full analyses in their own right, combining various data-driven and simulation-driven techniques, and are performed by the b -tagging working group [146, 147]. The calibration of a working point provides the tagging efficiency scale factor (SF) that corrects the efficiency measured in the simulation to the value in data as a function of the jet p_T . The SFs for light jets are further divided into two bins in η due to the sensitivity of the light-jet calibration to additional dead material and the different response

Operating Point	$\log(\frac{P_c}{P_b})$	$\log(\frac{P_c}{P_l})$	ϵ_c	$1/\epsilon_b$	$1/\epsilon_l$
Loose	> -0.9	-	0.90	2.5	1.0
Medium	> -0.9	0.95	0.2	8.0	200

Table 3.2 – Summary of the JetFitterCharm calibrated operating points. The loose operating point rejects b -jets, while the medium operating point further rejects light jets [79].

in the forward region. The SFs are accompanied by systematic uncertainties derived from calibration analysis. In brief, the b -jet calibration is derived from studying dileptonic $t\bar{t}$ events with two or three jets using a combinatorial likelihood method [148], the c -jet tagging calibration is derived from multijet events with reconstructed D^* mesons [147], while the light-jet calibration is derived from a negative-tag analysis [147]. The negative-tag approach aims to measure the fraction of light-jets that are tagged as heavy flavour jets by the tagging algorithm. The approach involves reversing the significance parameters of the tagger (such as the impact parameter significance or the secondary vertex significance) and evaluating the negative-tag performance of HF-jets. Since light-jet mis-tagging occurs due to the finite resolution of the inner detector, these effects are isolated by measuring the negative-tag rates for HF jets, which are then used to measure the light-jet mis-tag rate.

While tagging heavy flavour jets is always challenging, tagging c -jets is especially tricky as they lie between bottom and light jets in the landscape of discriminating characteristics. Thus the trade off between type 1 (signal inefficiency) and type 2 (signal purity) errors is worse when compared against b -jet tagging. The JetFitterCharm algorithm was selected for this analysis as it employs a Neural Network that is trained specifically to identify c -jets [79]. From the output tag weights (one each corresponding to the probability of a jet being a b , c or light jet) two discriminants are constructed: one for rejecting light jets ($\log(\frac{P_c}{P_l})$) and the other for b -jets ($\log(\frac{P_c}{P_b})$). For the 8 TeV dataset two working points were calibrated for the JetFitterCharm tagger, one to reject bottom jets, referred to as *loose*, and one to further reject light jets, referred to as *medium*. Table 3.2 summarizes these working points and their performance.

Figure 3.12 displays the topology of the JetFitterCharm discriminants measured in the dedicated calibration analysis [79] for c , b and light flavour jets. The

medium criterion cut values are indicated by the green lines in the top right hand quadrant of the plot, and nicely illustrates how the discriminant topology is used to separate jet flavours, but also how there is an unavoidable overlap between them.

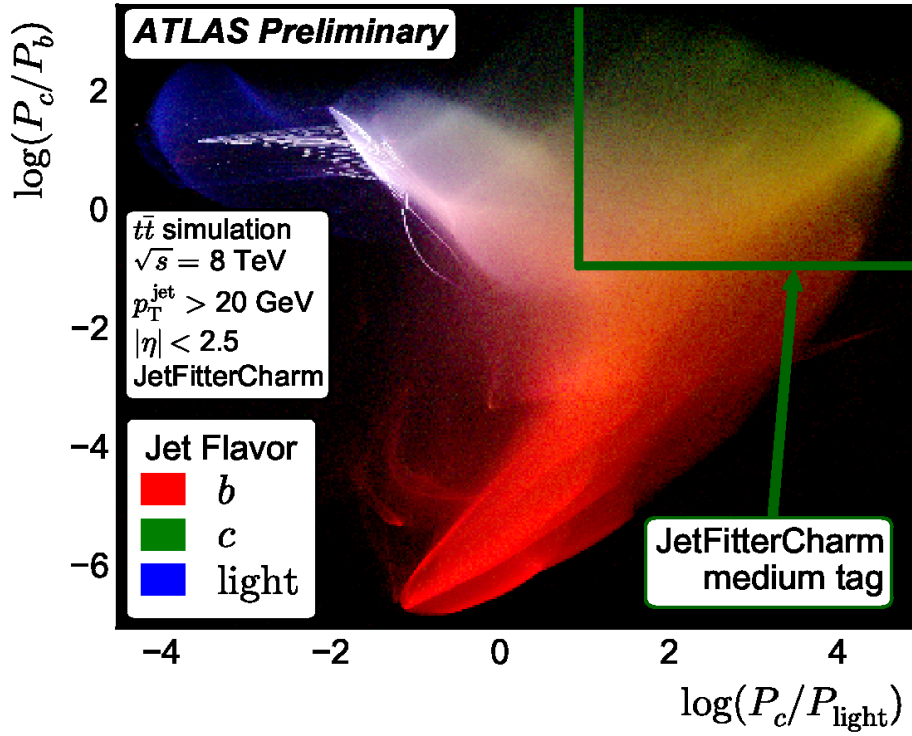


Figure 3.12 – Toplogy of c , b and light flavour jets as a function of the tagger discriminants for the JetFitterCharm tagger calibration study [79]. The green box in the top right hand quadrant indicates the *medium* selection criteria.

Data and Simulated Samples

4.1 Data Samples

The data used in this analysis comprises 20.1 fb^{-1} of 50 ns spaced pp collisions taken at $\sqrt{s} = 8 \text{ TeV}$ during Run 1 in 2012. Table 4.1 summarizes the data used. Only luminosity blocks devoid of data quality defects adhering to the final Run 1 good runs list (GRL) are considered. Due to a change in reconstruction software during the reprocessing campaign, runs 209736 and 214618 do not have the c -tagging (JetFitterCharm) variables.

4.1.1 Trigger requirements

The data sample was selected using six single photon E_T^γ threshold high-level triggers. The triggers were all prescaled during data taking, as shown in table 4.2, with the exception of the highest threshold of 120 GeV. Events selected by a given trigger are corrected for their prescale using a weight equal to the ratio of the corrected luminosity collected by that trigger to the total unprescaled luminosity of the sample. This approach avoids bias that might arise through prescale reweighting individual events based on their instantaneous prescale, and makes use of the final luminosity calibration for the full dataset.

Run Period	Run # Range	Total Lumi [pb ⁻¹]	Skim Lumi [pb ⁻¹]	Skim Events	Missing Runs
Period A	200804-201556	785.753	785.356	4.2764554E+07	
Period B	202660-205113	5051.56	5051.01	1.7576686E+08	
Period C	206248-207397	1397.44	1397.49	4.8706588E+07	
Period D	207447-209025	3275.48	3275.01	1.12215320E+08	
Period E	209074-210308	2525.86	2404.12	8.5147779E+07	209736
Period G	211522-212272	1279.22	1278.50	4.3507294E+07	
Period H	212619-213359	1452.66	1452.42	5.1568718E+07	
Period I	213431-213819	1021.7	1021.54	3.6267046E+07	
Period J	213900-215091	2610.14	2550.24	9.2858071E+07	214618
Period L	215414-215643	846.356	846.30	2.9700696E+07	
Total	200804-215643	20246.2	20062.0	7.18378001E+08	

Table 4.1 – Number of events and integrated luminosity by data period.

Trigger	Luminosity With Run Veto [pb ⁻¹]	Total Luminosity [pb ⁻¹]	Average Prescale
EF_g120_loose	20058.2	20238.9	1
EF_g100_loose	1539.21	1551.87	14
EF_g80_loose	703.285	708.928	31
EF_g60_loose	247.936	249.935	87
EF_g40_loose	57.5755	58.0434	374
EF_g20_loose	4.54374	4.5815	4220

Table 4.2 – Average prescales for single photon E_T triggers used throughout 2012 data taking.

4.2 Monte Carlo Simulations

Monte Carlo simulations of the γ +jet signal are used to study the expected signal behavior and to derive the corrections necessary to produce the final measured result.

The MC generators PYTHIA 8.165 [82] and SHERPA 1.4.0 [122] are used to simulate events arising from Leading Order (LO) pQCD matrix elements with the inclusion of initial state and final state radiation (ISR and FSR).

The PYTHIA simulation includes two main contributions, a *hard* component arising from direct photon production ($qg \rightarrow q\gamma$ and $q\bar{q} \rightarrow q\gamma$) and a *brem* component arising from Bremsstrahlung of hard QCD dijet-like events. At LO the absolute normalization of a simulated cross section is not well defined, and as such neither are the relative fractions of these two components. The relative fractions of the hard and brem components is optimized in section 7.2.3 to best match the reconstructed E_T^γ spectrum to the data, and to assess the sensitivity of the measured result to this fraction. The PYTHIA simulation makes use of the Leading Order (LO) CTEQ6L1 Parton Distribution Function (PDF) for modelling of the proton structure, the Lund string model [126] for hadronization, and the generator settings follow the AU2 CTEQ6L1 tune [149].

The SHERPA samples include LO matrix elements for photon+jet final states including three additional partons and parton showers. The nominal photon+jet SHERPA simulation, that was used for instance in the γ +jet analysis, sets the masses of the heavy quarks to zero for computational expediency. To rectify this two additional SHERPA simulations that include the heavy flavour quark masses were generated, one with a *b*-jet filter and the other with a *c*-jet filter. These three simulated samples are then combined, after applying a *c*-jet and *b*-jet veto on the original massless sample, to produce a complete coverage of the phase space with the correct treatment of the heavy quark kinematics. For hadronization, the SHERPA simulation uses a modified cluster model [150], the proton structure is parametrized using the Next to Leading Order (NLO) CT10 PDF, and the generator settings follow the CT10 tune.

Both simulations are interfaced with the GEANT4 [151] ATLAS detector simulation and reconstruction software [115] providing output data in a format analogous to

real data samples. Due to the logarithmically falling nature of the signal cross section as a function of E_T^γ , the simulations are generated by individually simulating portions of the E_T^γ spectrum and then stitching them together, providing statistical power across the entire spectrum without needlessly generating additional statistics at low E_T^γ . Figure 4.1 displays the resulting spectrum of reconstructed photon events for the SHERPA simulation. The details of this generation are displayed in table 4.3.

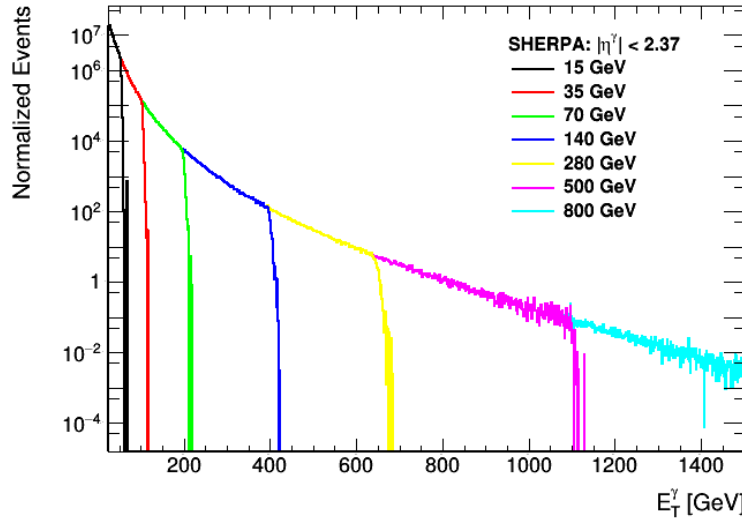


Figure 4.1 – Reconstructed photon yield for each SHERPA sample region, following the event selection.

4.2.1 Corrections

Prior to using the MC simulations the following corrections are applied to improve the agreement with the data.

Pile Up Reweighting

The anticipated pile-up conditions for Run 1 that were used to generate the simulations prior to the Run 1 data taking are corrected to better match the final distributions observed in data. The reweighting is done for two independent pile-up distributions, once based on the number of primary vertices (NPV) and once based

Generator	Run #	# of Events	Cross Section [nb]	Generator Filter Eff. [%]	E_T^γ Cut [GeV]	E_T^γ Range [GeV]
PYTHIA	129170	2999999	1.2354E+06	2.3477E-04	17	25-55
	129171	2999694	5.8768E+04	4.0218E-04	35	55-105
	129172	8779767	3.425E+03	5.705E-04	70	105-200
	129173	2993981	1.2217E+02	9.6932E-04	140	200-400
	129174	1499982	3.3487E+00	1.4457E-04	280	400-650
	129175	999877	1.1563E-01	1.8056E-03	500	650-1100
	129176	99997	4.9226E-03	1.9036E-03	800	>1100
Massless SHERPA	126372	9999783	3.8961E+02	1.0000E+00	15	25-55
	113714	9999183	2.4475E+01	1.0000E+00	35	55-105
	113715	5499676	2.1523E+00	1.0000E+00	70	105-200
	113716	2499984	1.3774E-01	1.0000E+00	140	200-400
	113717	999985	5.9627E-03	1.0000E+00	280	400-650
	126371	999976	2.7645E-04	1.0000E+00	500	650-1100
	126955	99996	1.3346E-05	1.0000E+00	800	>1100
Massive B SHERPA	207116	3999196	4.0937E+02	5.1037E-02	15	25-55
	207117	3971487	2.5710E+01	6.0686E-02	35	55-105
	207118	3923787	2.2521E+00	7.1455E-02	70	105-200
	207119	1999991	1.4357E-01	8.4898E-02	140	200-400
	207120	1999873	6.1879E-03	1.0036E-01	280	400-650
	207121	89996	2.8736E-04	1.1376E-01	500	650-1100
	207122	98998	1.3901E-05	1.2417E-01	800	>1100
Massive C SHERPA	207123	3992690	4.1146E+02	4.1474E-01	15	25-55
	207124	3980293	2.5846E+01	4.2489E-01	35	55-105
	207125	3999190	2.2593E+00	4.2496E-01	70	105-200
	207126	1999789	1.4401E-01	4.2164E-01	140	200-400
	207127	1999874	6.2111E-03	4.1889E-01	280	400-650
	207128	99996	2.8809E-04	4.1852E-01	500	650-1100
	207129	94997	1.3924E-05	4.1768E-01	800	>1100

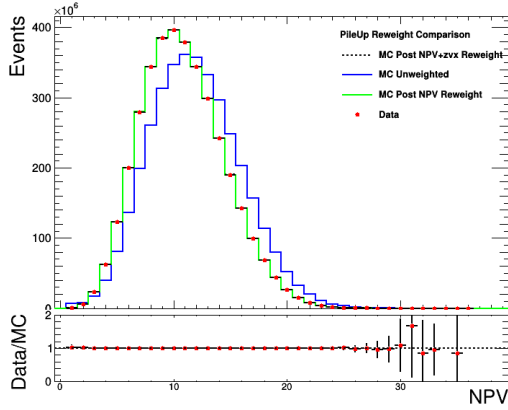
Table 4.3 – Number of events and generation parameters by simulation slice.

on the distribution of the primary vertex along the beam axis (the z-vertex). This procedure is optimal since it does not rely on the measurement of the average number of interactions per bunch crossing (μ), as many of the other recipes do, which is an approximation to the true NPV that is subject to change depending on calibrations. The reweighting is performed following the application of quality cuts (the Good Runs List, LAr error, Tile error, Core Flags, the number of primary vertices ≥ 2) and the trigger requirement (an "or" of the g20, g40, g60, g80, g100 and g120 loose photon triggers) to the data and the MC. Beyond this a preselection for *relaxed-tight* photons, that is photons that pass a subset of the tight photon criteria used in the signal region, is applied which improves agreement between the signal simulations and the data. Figures 4.2 and 4.3 display the corrected NPV and z-vertex distributions in PYTHIA and SHERPA compared to the data.

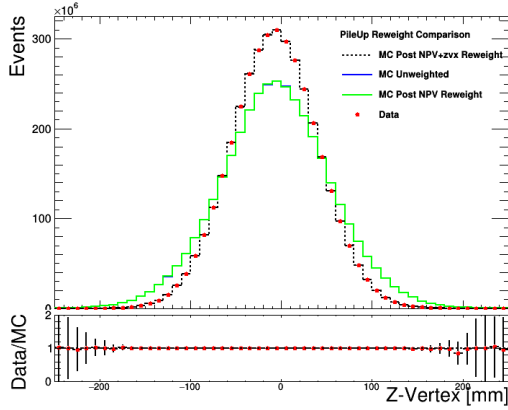
Reconstructed Photon Isolation Energy

Corrections are applied to the simulations to rectify discrepancies between the calibrated photon isolation in data and the photon isolation in MC. As described in section 3.1.3, the photon isolation energy is computed by taking the sum of all positive energy topoclusters that fall within a radius of $\Delta R < 0.4$ of the photon, and then subtracting the energy in the core that should correspond to the energy of the signal photon. The core of the photon is defined to be 5×7 cells in size and does not always completely encapsulate the photon's energy, thus an additional correction is applied when the data is reconstructed to account for signal leakage into the isolation cone. These leakage corrections, which were derived using simulations prior to data-taking, were somewhat overestimated E_T^γ [152]. Figure 4.4 shows a comparison of the data and simulated E_T^{iso} distributions in bins of increasing E_T^γ , and shows that the separation between the two distributions increases with E_T^γ .

In order to have the simulation better match the reconstructed data an offset correction is applied to the simulated E_T^{iso} distributions as a function of E_T^γ and η^γ using a tool (CaloIsoDDCorrectionTool) developed by the ATLAS SM Diphoton analysis team. This tool uses a simulated signal and data driven background template fit technique to derive correction factors, and the usage of these correction factors is validated in the context of this analysis as follows. A control region is created by selecting photons that satisfy the *relaxed-tight* but not the *tight* criteria



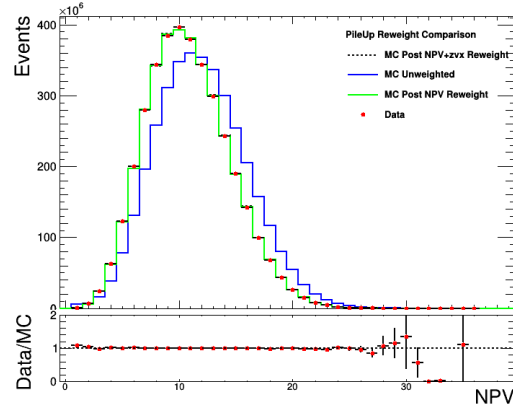
(a) NPV



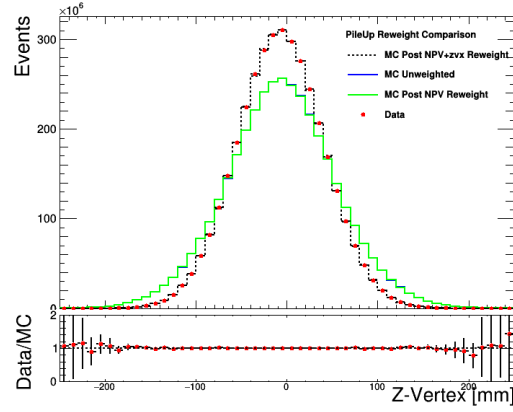
(b) Z-Vertex

Figure 4.2 – Corrected pile-up distributions in PYTHIA.

(see section 3.1.1 for descriptions of these criteria) in the data to derive an E_T^{iso} distribution of background-like photons. This background is then subtracted from the signal region E_T^{iso} distribution by matching the integral of the tails of the distributions ($> 10 \text{ GeV } E_T^{\text{iso}}$). The resulting E_T^{iso} distribution better reflects that of the simulation having removed a significant amount of background, and at this point both distributions are fit using a continuous and smooth combination of a Gaussian and falling exponential. The Gaussian captures the detector effects while the exponential captures the tail and asymmetric nature of the distribution that arises from nearby physics signals. The discrepancy between the two is then captured in the difference between the mean of the Gaussian in the fit for the data and the simulation. Figure 4.5 shows that that correcting the simulation using the tool greatly



(a) NPV



(b) Z-Vertex

Figure 4.3 – Corrected pile-up distributions in SHERPA.

improves the agreement with data across the entire E_T^γ spectrum, where $p^{S_{ample}}_0$ is the mean of the Gaussian component in the fit.

The tool has built in methods for evaluating relevant uncertainties. The two uncertainties are related to the shift, i.e. the correction itself, and the smearing, i.e. the modeling of the detector resolution.

With the data samples and corrected MC simulations in place, offline selection cuts are applied to extract the signals used to perform the measurement.

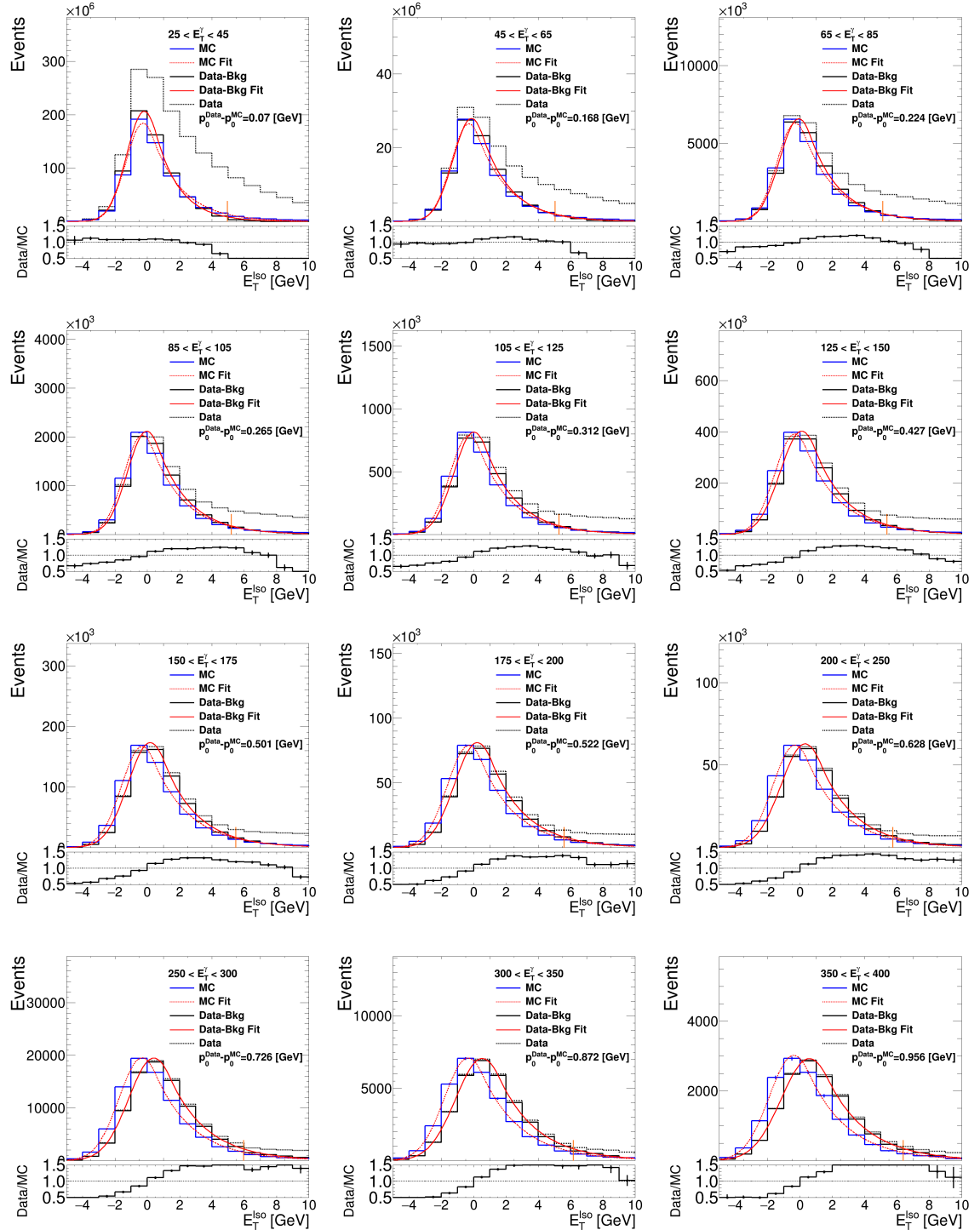


Figure 4.4 – Uncorrected E_T^{iso} distributions as a function of E_T^{γ} for SHERPA. The background subtracted data and the MC (simulation) are each fit with a smooth combination of a Gaussian and a falling exponential. p_0^{Sample} is the mean of the Gaussian component in the fit.

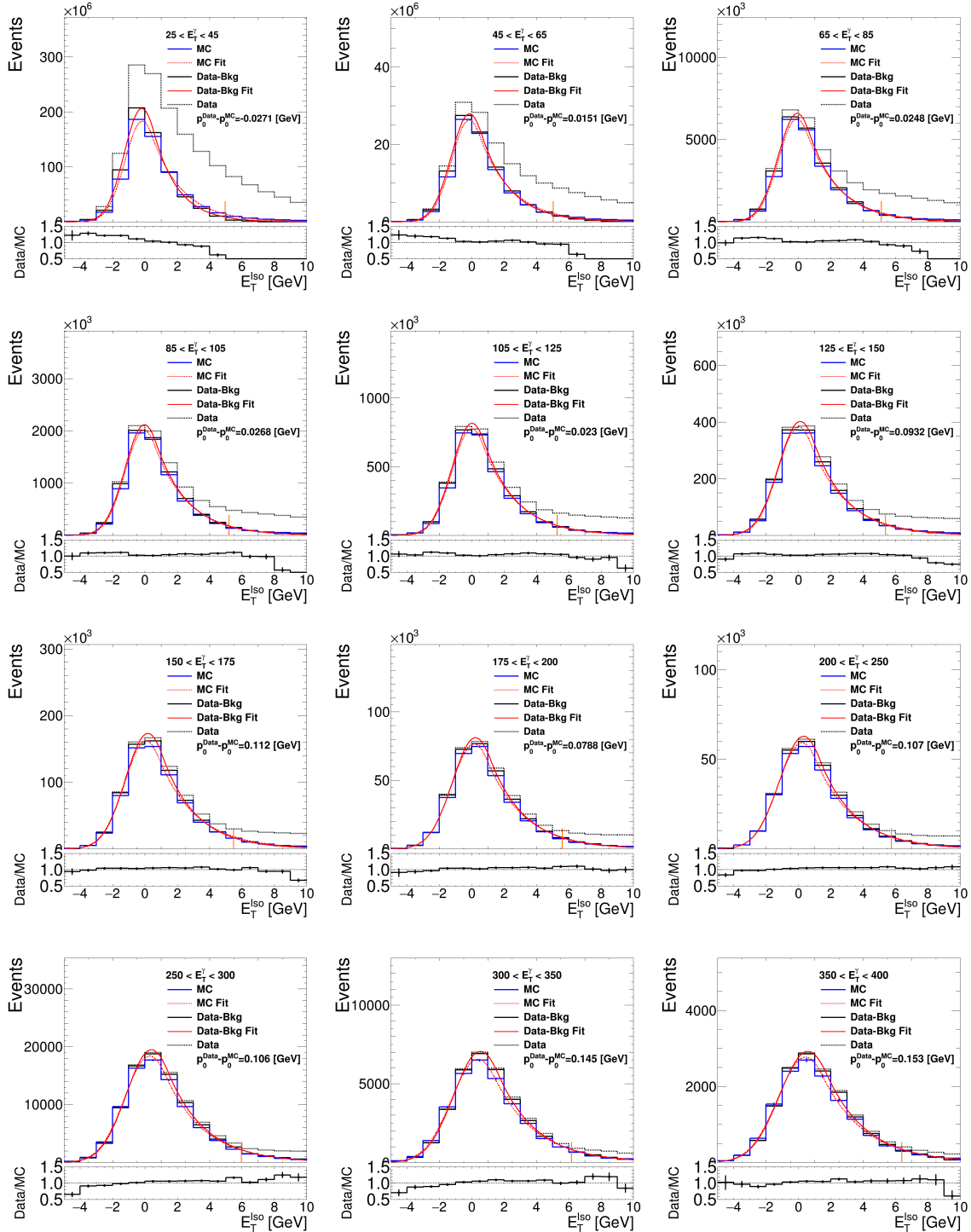


Figure 4.5 – Corrected E_T^{iso} distributions as a function of E_T^γ for SHERPA. The background subtracted data and the MC (simulation) are each fit with a smooth combination of a Gaussian and a falling exponential. p_0^{Sample} is the mean of the Gaussian component in the fit.

Signal Event Selection

The following sections describe in detail the selection criteria applied to the data and the simulations at both detector and particle-level. The selection of general γ +jet events follows the procedure used in the 8 TeV ATLAS γ +jet cross section measurement [65]. The measurement is binned as a function of E_T^γ using bin edges consistent with the existing 8 TeV inclusive γ measurement performed by ATLAS [?]. Due to the additional statistical power required to perform this measurement, some bins have been merged. In particular bins have been merged in the low E_T^γ range of the measurement such that each prescaled trigger is used to measure a single bin of E_T^γ . For example, the 20 GeV trigger is used to perform the measurement of the $25 < E_T^\gamma < 45$ bin. The cut flows displaying the relative and total acceptance of each cut in both data and the SHERPA simulation are provided in appendix A.

5.1 Detector-Level Photon+Jet Selection

The selection of general photon+jet events follows the procedure used in the ATLAS SM Photon+Jet cross section measurement [65]. The jet acceptance is adapted to match the acceptance of the inner detector to permit the addition of flavour tagging, and the kinematic acceptance of the photon is extended to the lowest possible photon E_T^γ in order maximize the reach of the analysis.

5.2 Event-Level Criteria

Events meeting the following requirements, related to the ATLAS data taking performance and the primary vertex track multiplicity, are considered:

- The event must satisfy the photon trigger requirement described in section 4.1.1.

- The event must be taken during a run and luminosity block present in the final 2012 Good Runs List, that is the list of runs that satisfied standard ATLAS data-taking quality criteria.
- The event must not be taken from a luminosity block following a restart of the ATLAS data acquisition system during a run (which can occur in order to recover/restart sub-detector systems) as in these instances some event information is incomplete.
- The event must contain an unambiguous primary vertex consistent with the average beam spot that has at least two associated tracks of $p_T > 400$ MeV.
- The event must not have been coincident with noise bursts and other data quality errors related to the LAr or tile calorimeters.

5.3 Photon Selection Criteria

Reconstructed photon candidates are required to satisfy various signal quality criteria based on the reconstruction of an electromagnetic cluster in the calorimeter and the tracking information associated to that cluster in the inner detector as described in 3.1. The specific photon-candidate selection criteria applied in this analysis are:

- Photons are required to satisfy the *relaxed-tight* photon identification criteria, described in section 3.1.1.
- Simulated photons are required to be matched to a particle-level prompt photon within a cone of $\Delta R < 0.2$, where $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$.
- Photons are required to have $E_T^\gamma > 25$ GeV.
- Photons are required to have $|\eta^\gamma| < 2.37$, excluding the crack region between $1.37 < |\eta| < 1.56$.
- In data photons are required to have their E_T^γ fall in the plateau region of a photon trigger that provided an accept for the event:
 - EF_g20_loose: $25 \text{ GeV} < E_T^\gamma < 45 \text{ GeV}$.
 - EF_g40_loose: $45 \text{ GeV} < E_T^\gamma < 65 \text{ GeV}$.
 - EF_g60_loose: $65 \text{ GeV} < E_T^\gamma < 85 \text{ GeV}$.
 - EF_g80_loose: $85 \text{ GeV} < E_T^\gamma < 105 \text{ GeV}$.

- EF_g100_loose: $105 \text{ GeV} < E_T^\gamma < 125 \text{ GeV}$.
- EF_g120_loose: $E_T^\gamma > 125 \text{ GeV}$.
- Photons are required to pass object quality criteria that ensure that they are not adversely affected by calorimeter issues, including malfunctioning high voltage or readout electronics.
- Photons are required to pass object cleaning criteria that ensure that they are not affected by calorimeter noise bursts and that the cells making up the calorimeter cluster exhibit a satisfactory signal quality [153].
- The leading photon selected and the remaining photons are not considered.
- The remaining photon is then required to pass the *tight* photon identification criteria, described in section 3.1.1.
- The remaining photon is then required to be isolated, imposed by requiring that the E_T^{iso} , described in section 3.1.3, be less than $4.8 + 0.0042 \times E_T^\gamma \text{ GeV}$. Signal photons with higher E_T^γ will also characteristically have higher E_T^{iso} , as seen in figure 5.1. This choice of sliding E_T^{iso} cut better ensures the acceptance of high E_T^γ signal photons, preserving the signal efficiency at higher E_T^γ as shown in figure 5.2.

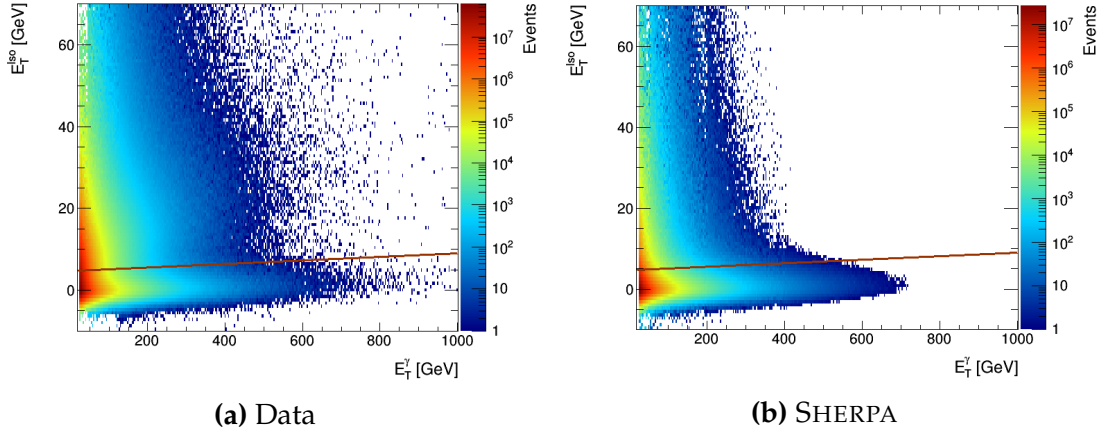


Figure 5.1 – Topology of the E_T^γ vs E_T^{iso} distribution in the data and in SHERPA for the full photon acceptance. The E_T^{iso} cut value is overlaid, and visibly follows the contour of the signal peak with increasing E_T^γ .

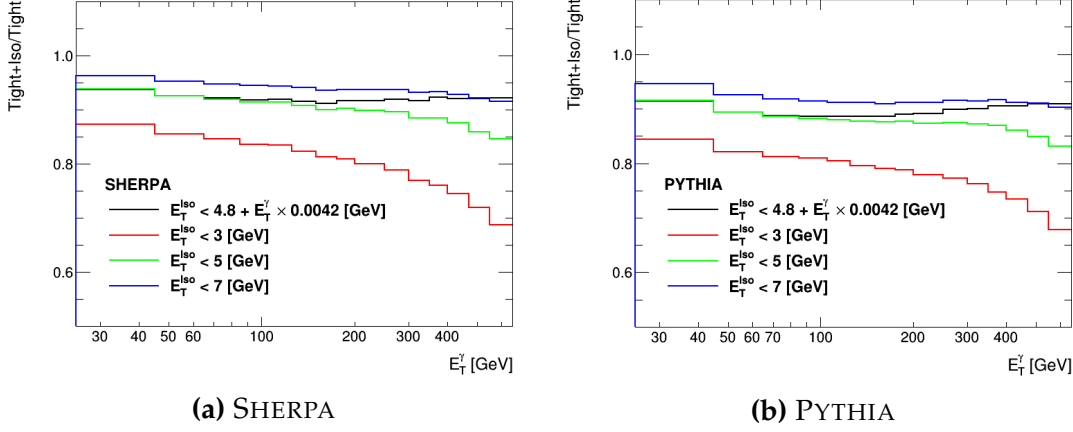


Figure 5.2 – Fraction of tight signal photons in the simulation that pass the subsequent photon isolation cut.

5.4 Jet selection Criteria

Jets are reconstructed using the anti- k_t algorithm [138] with a distance parameter $R = 0.4$, using as input locally calibrated topological calorimeter clusters and calibrated as described in section 3.2. The following selection criteria are applied to the jets:

- Jet quality criteria:
 - The jet must be measured by well functioning calorimeter modules (no hot tiles, not calorimeter spikes, well understood LAr noise, cosmic and beam backgrounds [154]).
 - If the jet has $p_T < 50$ GeV then at least 50% of its associated transverse track momentum must come from the primary vertex with the largest $\sum p_T$. This ensures the quality of the jet vertex reconstruction.
- The jet coincident with the leading photon is removed from the collection. To do so a $\Delta R < 0.4$ criteria is used between the photon and the jets in the collection, where $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ is calculated using the constituent scale jet angles and the second calorimeter layer photon angles (i.e. uncalibrated detector level quantities) since the targeted effect is reconstruction based and not physics based.
- The leading jet is selected and the remaining jets are not considered.
- The leading jet is required to have $p_T > 20$ GeV.

- The leading jet is required to have $|\eta| < 2.5$ and $|y| < 2.5$ (within the acceptance of the tracker to permit flavour tagging).
- The remaining jet is required to be separated by at least $\Delta R > 1.0$ from the leading photon, where $\Delta R = \sqrt{\Delta\phi^2 + \Delta y^2}$ and uses the calibrated four-vector for both objects.

The final ΔR cut between the leading photon and the leading jet is motivated by the shape of the mean E_T^{iso} distribution vs ΔR to the leading jet. Figure 5.3 shows that in the data for small ΔR there is a coupling effect between the two objects leading to an increase in the photon's E_T^{iso} . The ΔR cut aims to remove the relatively rare number of instances where this occurs even in light of the isolation cut. This cut is conservative in that it does not have a large impact on the statistics, and in that it removes an effect that is difficult to simulate in MC. The other feature of interest in this figure is the remaining shape of the mean E_T^{iso} distribution, notably that it rises towards larger ΔR . Figures 5.4, 5.5, 5.6 and 5.7 show the analogous plots for SHERPA, PYTHIA, and each of the PYTHIA brems and hard components. In all of these figures the same closeby effect that is targeted by the cut is present, as well as the rise of the mean E_T^{iso} with increasing ΔR , though the overall mean E_T^{iso} is much smaller. Notably in figure 5.6 and figure 5.7 there is almost no rise in the hard PYTHIA component implying that the rise mostly comes from the brems component. The fact that the overall mean E_T^{iso} distribution is much smaller in the signal simulation indicates that there are additional low energy background jets in the data. The fact that the rise in mean E_T^{iso} with ΔR is only seen in the PYTHIA brems component indicates that this behavior arises when the leading photon+jet system is not back to back and balanced by some additional objects which may be closeby to the photon.

Figure 5.8 shows the analogous plots in data of the mean E_T^{iso} vs $\Delta\eta$ and $\Delta\phi$ prior to the isolation cut. These figures give an idea of the overall topology of the leading photon+jet system, showing that typically the leading jet and leading photon are back to back in ϕ and close-by in η . They also show that there is a peak in the mean E_T^{iso} distribution when the leading photon and the leading jet are back to back in ϕ . Events with a larger multiplicity of jets have a higher probability of there being a jet closeby to the leading photon. Figure 5.9 displays the mean E_T^{iso} vs the multiplicity of jets with $p_T^{\text{jet}} > 20$ GeV prior to and following the isolation

cut, illustrating this effect, and that the application of the isolation cut mitigates it.

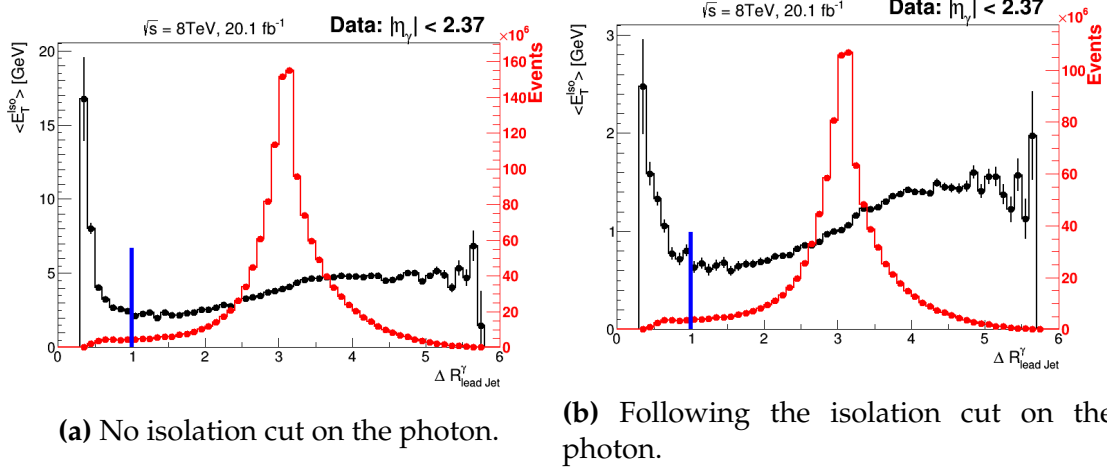


Figure 5.3 – Mean E_T^{iso} vs ΔR between the leading photon and leading jet prior to and following the isolation cut in data. The cut value on ΔR is chosen to be 1.0, indicated on the figures by the blue line falling in the minimum separating the effect of close by and likely coupled photon-jet pairs, and well separated photon-jet pairs.

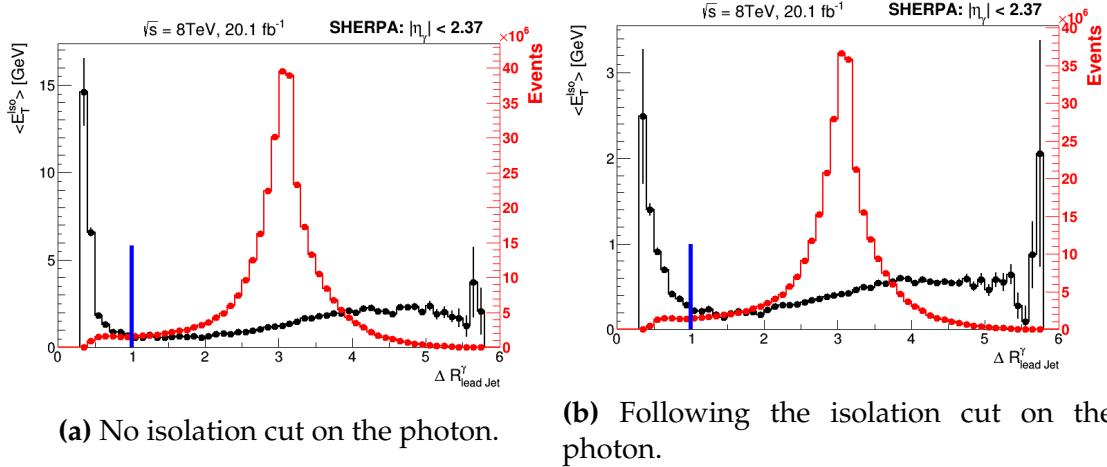


Figure 5.4 – Mean E_T^{iso} vs ΔR between the leading photon and leading jet prior to and following the isolation cut in SHERPA.

Figures 5.10 and 5.11 display the topology of the selected events as a function of the leading photon E_T^γ and leading jet p_T^{jet} for data and SHERPA. Figures 5.12, 5.13 and 5.14 display this same topology for SHERPA for truth tagged reconstructed light, c and b -jets. Figures 5.16 and 5.17 display the same topology for the

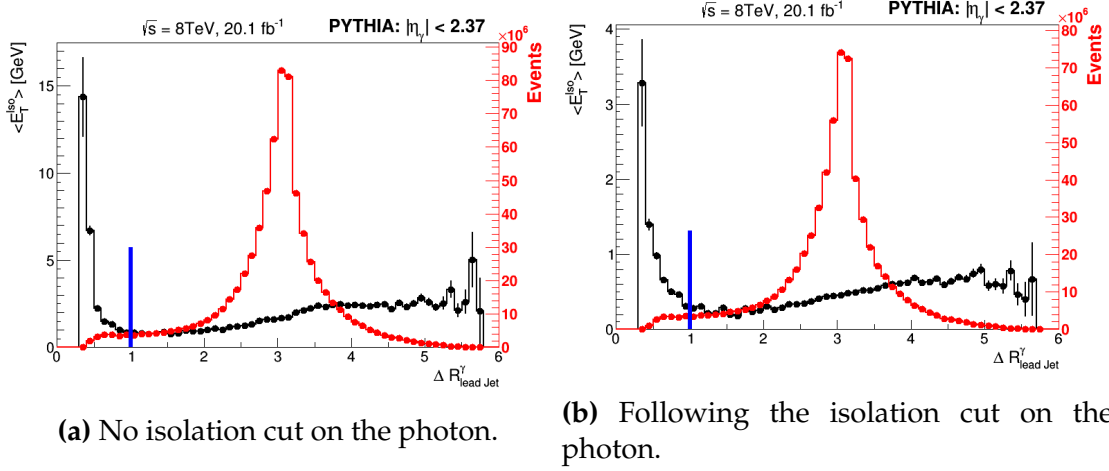


Figure 5.5 – Mean E_T^{iso} vs ΔR between the leading photon and leading jet prior to and following the isolation cut in PYTHIA.

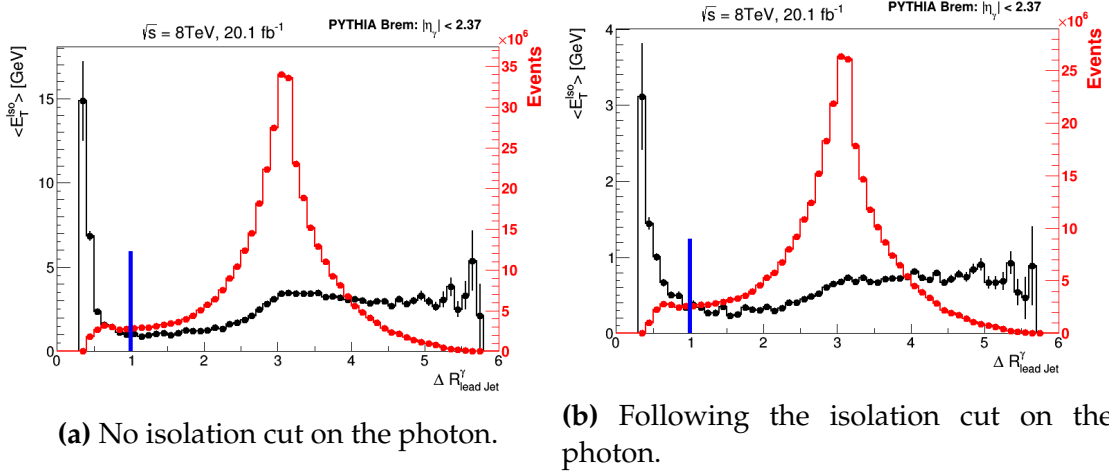


Figure 5.6 – Mean E_T^{iso} vs ΔR between the leading photon and leading jet prior to and following the isolation cut in the PYTHIA brem component.

hard and fragmentation portions of the PYTHIA simulation. Comparing the correlation between the E_T^{γ} and p_T^{jet} distributions for the PYTHIA components shows that there is a much stronger correlation for hard events than for brem events, which is expected since hard events come from a well balanced 2 body system whereas the brem component comes from a 3 body system. Bearing this in mind it is clear that the SHERPA simulation does a relatively good job overall reproducing the data distribution. Looking at the individual flavour components of the SHERPA simulation

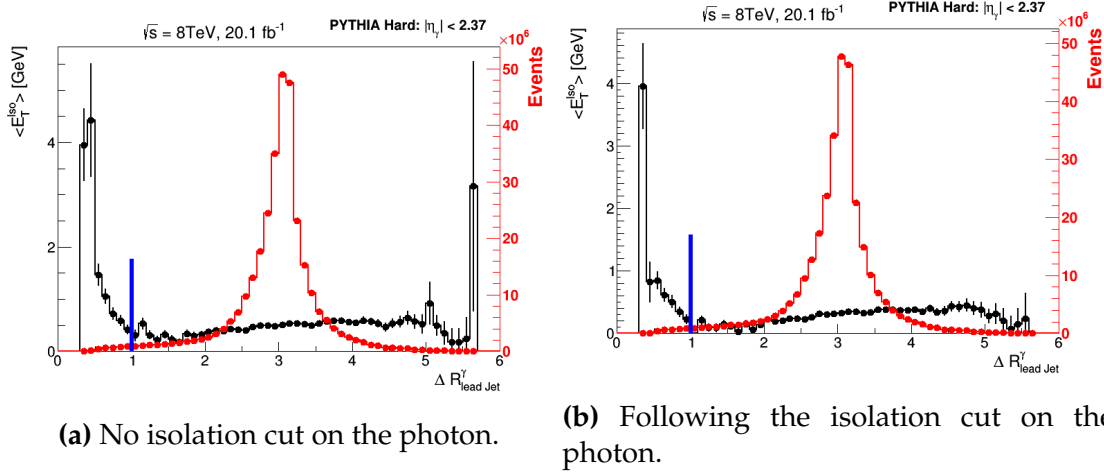


Figure 5.7 – Mean E_T^{iso} vs ΔR between the leading photon and leading jet prior to and following the isolation cut in PYTHIA hard component.

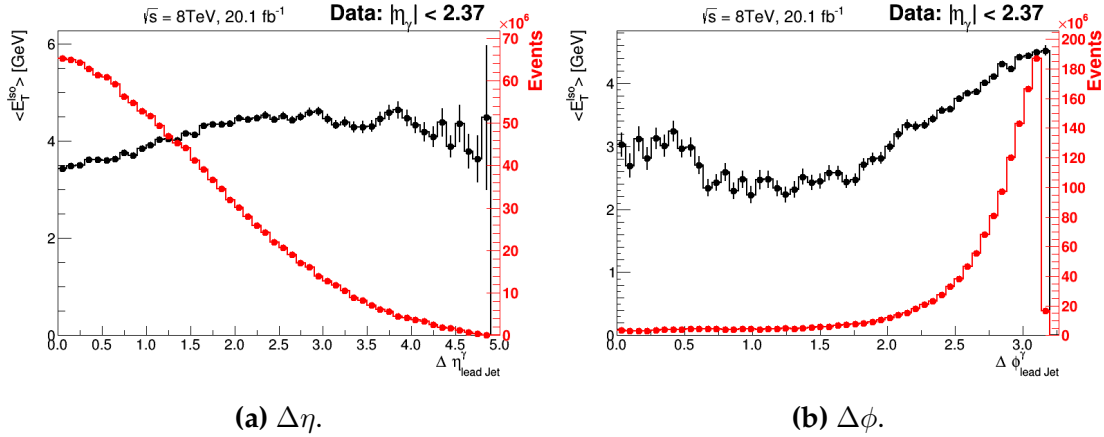


Figure 5.8 – Mean E_T^{iso} vs $\Delta\eta$ and vs $\Delta\phi$ between the leading photon and leading jet data prior to the isolation cut. Note that the drop in the final bin of the $\Delta\phi$ plot results from the bin edge extending beyond π .

in the central region shows that the b -jet correlation is smaller than the correlation that is seen in the data, while the c -jet correlation is slightly larger than the correlation that is seen in the data, and that the light-jet correlation is closest to that which is seen in the data. Figure 5.15 displays the default PYTHIA simulation, while figure 5.18 displays the optimized mixture of hard/brem events, described in detail in section 7.2.3.

Overall the SHERPA simulation does the best job of reproducing the recon-

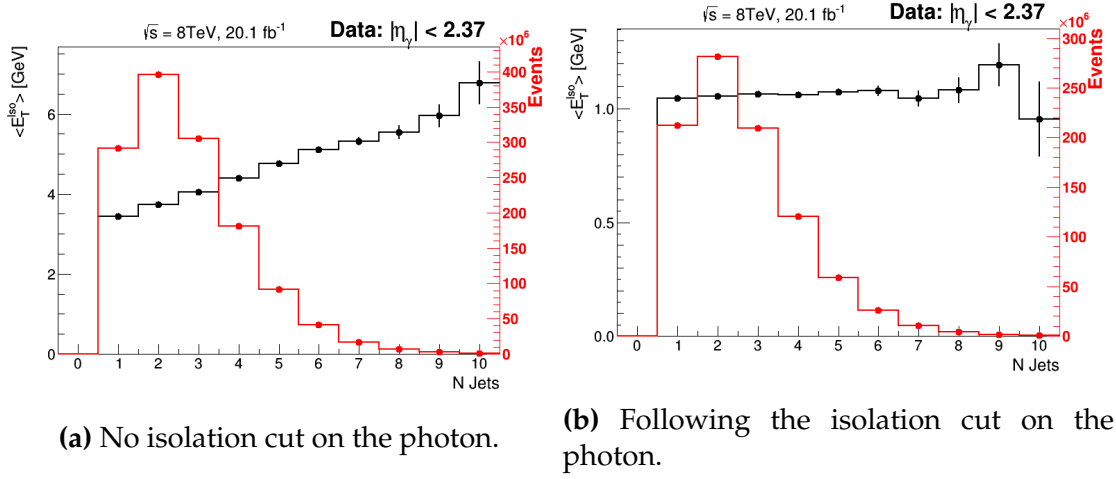


Figure 5.9 – Mean E_T^{iso} vs the number of jets with $p_T^{\text{jet}} > 20$ GeV in data.

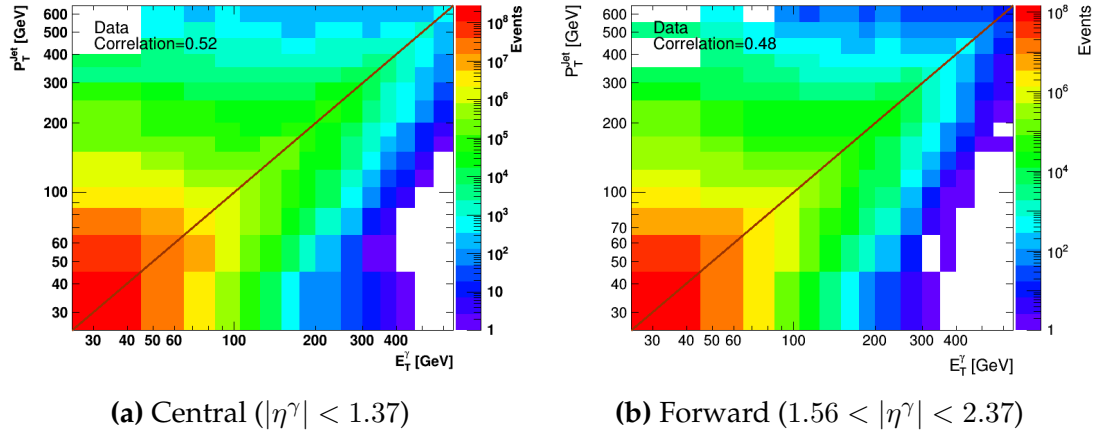


Figure 5.10 – Distribution of selected events as a function of leading photon E_T^γ and leading jet p_T^{jet} in data.

structured data, which is expected due to the higher order contributions in the matrix element.

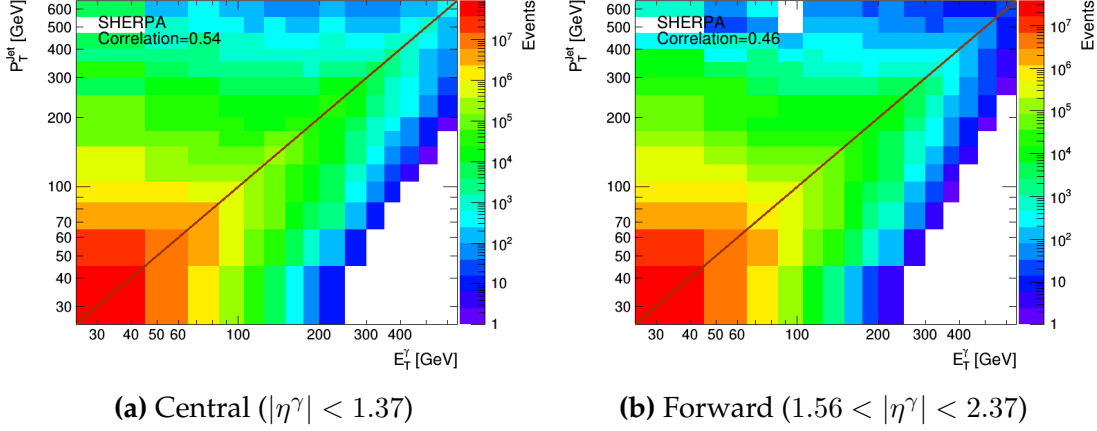


Figure 5.11 – Distribution of selected events as a function of leading photon E_T^γ and leading jet p_T^{jet} in SHERPA.

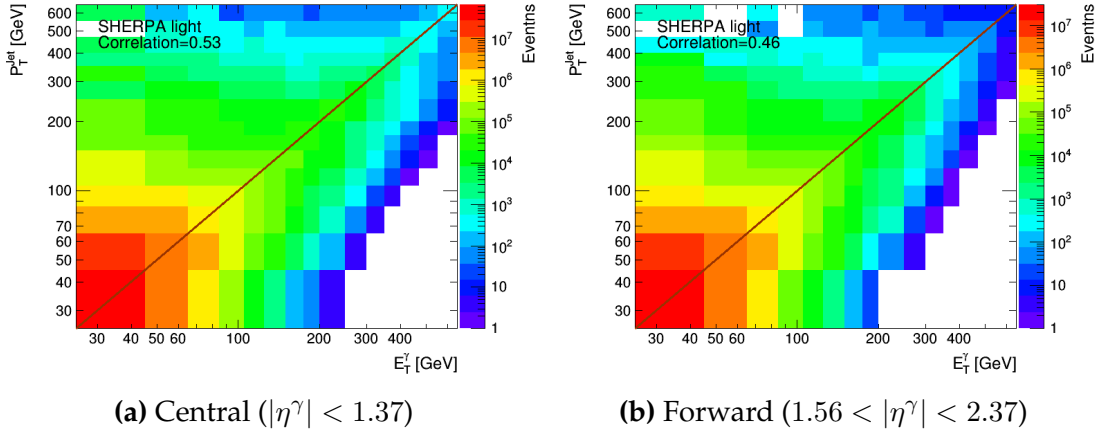


Figure 5.12 – Distribution of accepted light truth tagged events as a function of leading photon E_T^γ and leading jet p_T^{jet} in SHERPA.

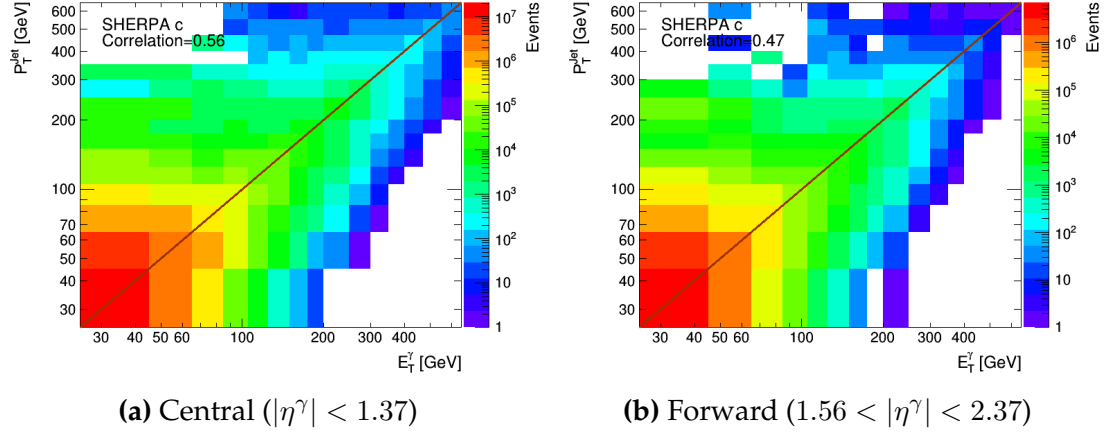


Figure 5.13 – Distribution of accepted charm truth tagged events as a function of leading photon E_T^γ and leading jet p_T^{jet} in SHERPA.

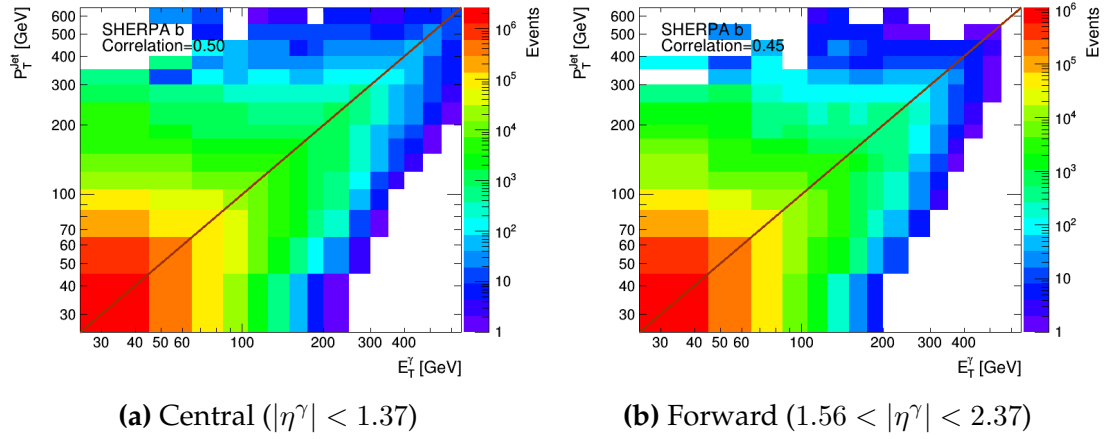


Figure 5.14 – Distribution of accepted bottom truth tagged events as a function of leading photon E_T^γ and leading jet p_T^{jet} in SHERPA.

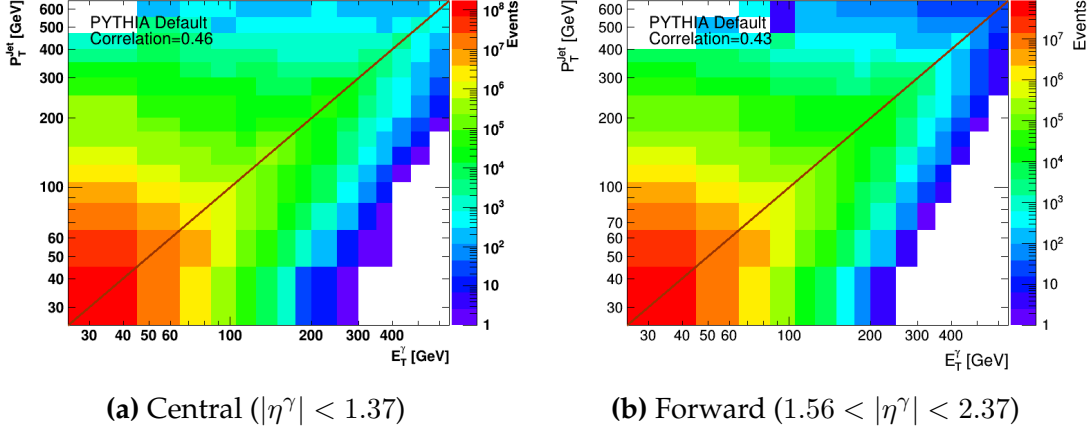


Figure 5.15 – Distribution of selected events with a leading prompt photon as a function of leading photon E_T^γ and leading jet p_T^{jet} in PYTHIA.

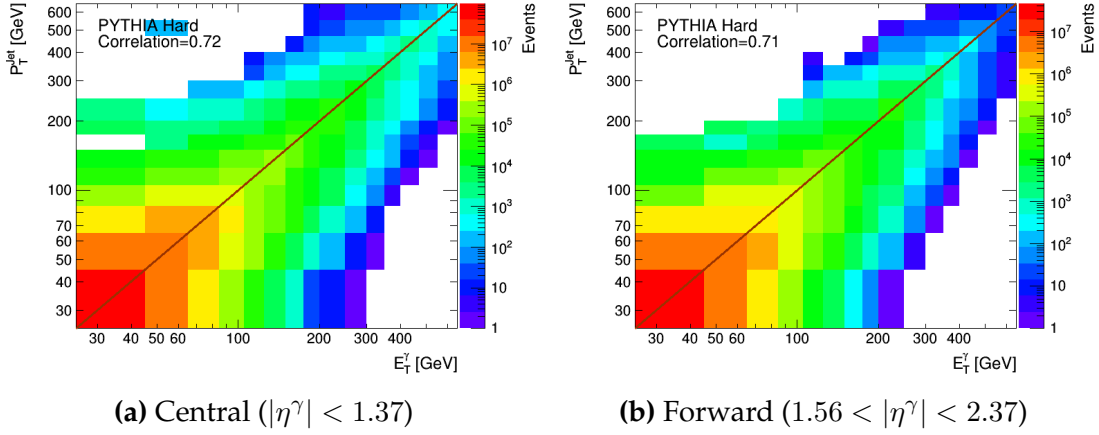


Figure 5.16 – Distribution of selected events with a leading hard prompt photon as a function of leading photon E_T^γ and leading jet p_T^{jet} in PYTHIA.

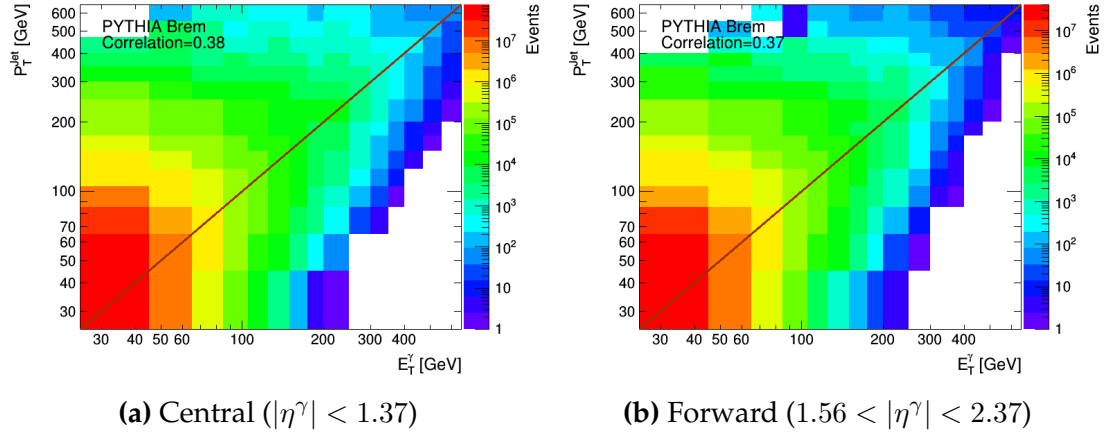


Figure 5.17 – Distribution of selected events with a leading fragmentation photon as a function of leading photon E_T^γ and leading jet p_T^{jet} in PYTHIA.

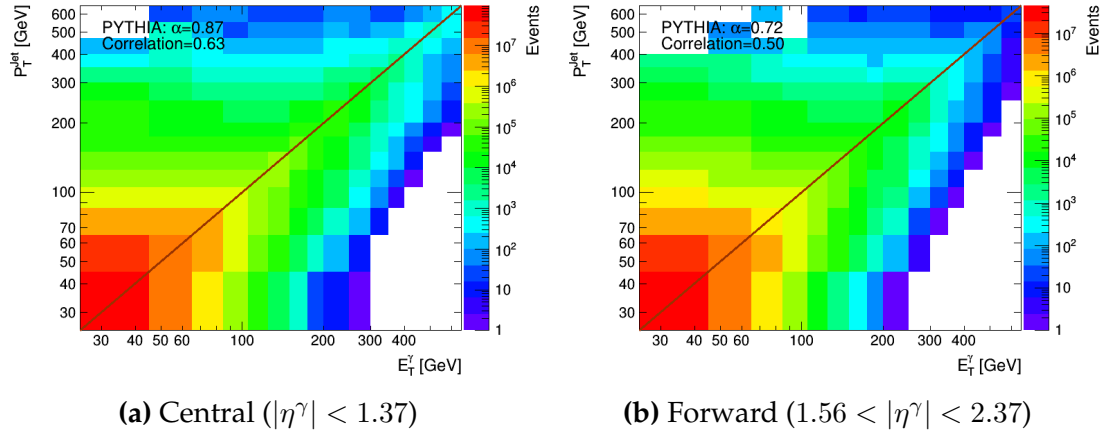


Figure 5.18 – Distribution of selected events with a leading hard prompt photon as a function of leading photon E_T^γ and leading jet p_T^{jet} in PYTHIA, following the optimization of the hard/brem fractions. See section 7.2.3 for a detailed explanation of this optimization.

5.5 Particle-Level Photon+Jet Selection

The particle-level selection, used to select particle-level events that populate the simulated distributions used for the unfolding corrections, uses kinematic criteria analogous to those used at the detector-level:

- Only the leading particle-level prompt photon is considered.
- The photon must have $E_T^\gamma > 25$ GeV.
- The photon must fall within $|\eta^\gamma| < 2.37$, excluding the crack region between $1.37 < |\eta| < 1.56$.
- The photon must have $E_T^{\text{iso}} < 4.8 + 0.0042 \times E_T^\gamma$ GeV (using the truth isolation definition of E_T^{iso} described in section 3.1.3).
- Only the leading jet that does not fall within $\Delta R < 0.4$ of the leading particle-level photon is considered for the remaining cuts.
- The jet must have $p_T > 20$ GeV.
- The jet must have rapidity $|y| < 2.5$.
- The jet must not fall within $\sqrt{(\Delta y)^2 + (\Delta \phi)^2} < 1.0$ of the leading particle-level photon.
- The jet must be flavour tagged, as described in section 3.2.

Applying the selection cuts outlined in this chapter yields a selection of well-measured γ +jet events. The next chapter outlines the measurement procedure used to extract the signal yield and correct for detector defects and inefficiencies, providing a measured result that can be compared to theory predictions and other measurements.

Measurement Procedure

The following equation outlines the procedure used to derive the measured cross section from the data yield:

$$\left(\frac{d\sigma}{dE_T^\gamma}\right)_i = \frac{1}{(\Delta E_T^\gamma)_i} \frac{1}{\mathcal{L}_{\text{int}}} \frac{1}{\epsilon_i^{\text{trig}}} C_i^{\text{unf}} f_i^{\text{c-jet}} \sum_{j \in \text{JFC}} p_{ij}^\gamma N_{ij}^{\text{Data}}. \quad (6.1)$$

In this equation $\left(\frac{d\sigma}{dE_T^\gamma}\right)_i$ is the measured cross section corrected back to the particle level in bin i of E_T^γ , $(\Delta E_T^\gamma)_i$ is the bin width in GeV, $\mathcal{L}_{\text{int}}$ is the measured integrated luminosity, ϵ_i^{trig} is the trigger efficiency, C_i^{unf} is the unfolding factor, $f_i^{\text{c-jet}}$ is the measured heavy flavour jet fraction, p_{ij}^γ is the measured signal photon purity in a tagger discriminant bin j , and N_{ij}^{Data} is the yield of data events. The unfolding factor corrects the measurement for detector effects, including the detector resolution and the signal reconstruction efficiency, yielding a measurement that is directly comparable to other experimental results and theoretical predictions.

The following sections describe the methods used to derive the components of this measurement.

6.1 Trigger Efficiency

Every analysis performed using ATLAS data relies on a combined hardware+software based trigger that during an LHC run records as many collision events of interest as possible out of the vast number delivered by the LHC. It is important to evaluate the performance of these triggers to understand the bias they might impose on the recorded data and in turn the final result. General performance studies done for the single photon trigger and a tight photon selection with the 8 TeV dataset indicate that requiring the reconstructed photon to have $E_T^\gamma > 5$ GeV above the threshold of the trigger achieves a 99.5% efficiency [155]. The specific criteria imposed by an analysis beyond the trigger criteria, however, may bias the data towards a different level of performance. The performance of five of the six single photon E_T triggers using the full cut flow is evaluated using the events triggered by the adjacent lower threshold trigger. The lowest trigger, EF_g20_loose, though not shown here, was also characterized in this fashion using a large sample of randomly triggered events. Figure 6.1 displays the efficiencies of these triggers in the signal region and table 6.1 shows that all the measured trigger plateaus are compatible with 99.3% efficiency.

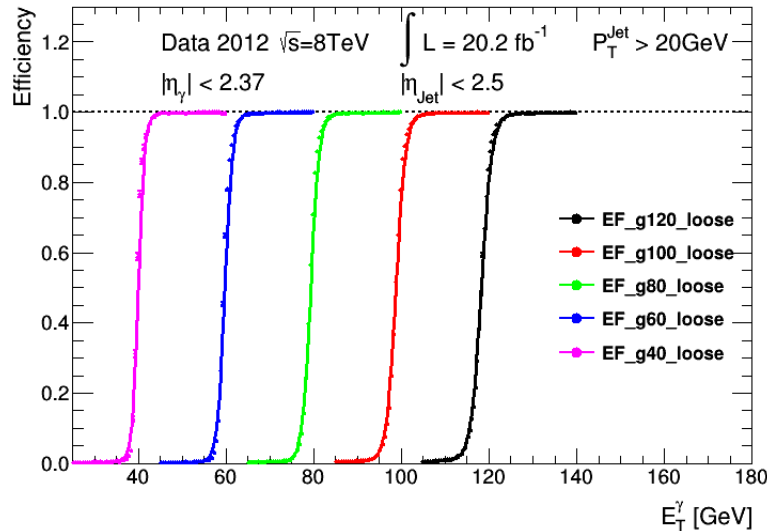


Figure 6.1 – Trigger efficiencies measured while applying the full selection criteria.

Trigger	E_T Acceptance Range [GeV]	Efficiency [%]	χ^2/DOF of Fit
EF_g120_loose	$125 < E_T$	99.33 ± 0.03	1.65
EF_g100_loose	$105 < E_T < 125$	99.41 ± 0.04	1.22
EF_g80_loose	$85 < E_T < 105$	99.48 ± 0.04	1.65
EF_g60_loose	$65 < E_T < 85$	99.56 ± 0.04	1.86
EF_g40_loose	$45 < E_T < 65$	99.62 ± 0.05	1.06

Table 6.1 – Threshold, threshold efficiency, χ^2/DOF for triggers used in this analysis measured using data applying a boot-strap method, fit using a Sigmoid function.

6.2 Photon Purity

The following procedure, used in the ATLAS inclusive photon and γ +jet analyses [61–65], is a data driven 2D sideband method that estimates and subtracts the contribution of background photons that leak into the signal photon region. In typical ATLAS photon analyses the correction is applied in each bin of E_T^γ . For the purposes of this analysis, the signal region is further divided along the axis of the flavour tagger discriminant which is used for the extraction of the flavour fractions in each bin of the measurement. This background photon correction is thus applied to each tagger discriminant bin in each bin of the measurement prior to the extraction of the flavour fractions (described in section 6.3), ensuring that any correlation between the jet flavour and the photon signal purity is measured and accounted for. It is important to note that background contributions from electrons that are mis-tagged as photons in the context of ATLAS prompt-photon analyses have been previously investigated and found to be small, less than 1% [156].

The isolation requirement on the photon is imposed in the signal selection to remove photons originating from τ or hadron decays and to limit photons arising from fragmentation of a parton. The photon ID requirements use shower shape information to identify photon-like signals. By reversing these two requirements independently and simultaneously, three sideband regions (B, C and D) are created that are enriched in background photons. These regions are used to estimate the background leakage in the signal region (A), as shown in figure 6.2.

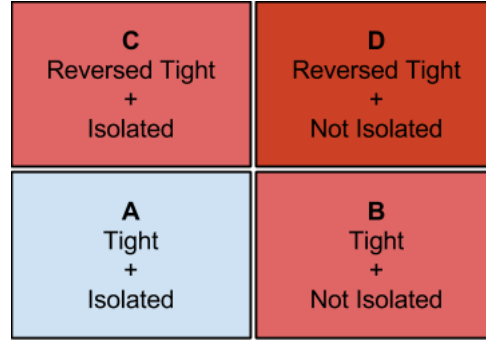


Figure 6.2 – Depiction of signal and control regions used for ABCD sideband method.

For the sideband subtraction to be valid, the manner in which the photon ID is reversed must not be correlated with the photon isolation when considering fake/background photons. This assumption is expected to hold as the ID reversal is largely based on shower shape variables, evaluated using the calorimeter strip layer, while most of the photon energy, and in turn the isolation energy, is collected in the second layer of the electromagnetic calorimeter. The reversed photon ID sideband regions are constructed by requiring the failure of at least one of a subset of the tight photon criteria that primarily target the lateral shower size and shape (described in detail in section 2.4.2). This alternative photon ID definition is referred to as *reversed-tight*. The isolation requirement is reversed by requiring that the candidate photons have an isolation energy greater than the default signal isolation requirement plus 2 GeV, i.e. $E_T^{\text{iso}} > (2.0 + 4.8 \text{ GeV} + 0.0042 \times E_T^\gamma)$. This 2 GeV gap is used so as to permit the evaluation of the sensitivity of this choice by varying the threshold up and down without impinging on the signal region. The sensitivity of the assumption that these criteria are uncorrelated with the isolation is evaluated and taken into consideration as an uncertainty.

Assuming that the background regions are indeed dominated by background, and assuming that the reversed-tight to tight ratio is the same in the isolated and the non-isolated region, a data-driven subtraction of background in the signal region is performed. This is done by projecting the background yield of the three control regions into the signal region where N_{Region} represents the data yield in a given region:

$$N_A^{\text{Sig}} = N_A - \frac{N_B N_C}{N_D}. \quad (6.2)$$

There is, however, a non-negligible contribution of signal that leaks into the background dominated control regions. The simulations are used to estimate this effect and for each control region the signal leakage fraction is represented by $c_{Region} = \frac{N_{Region}^{MC}}{N_A^{MC}}$ and referred to as a *leakage factor*. Incorporating the leakage factors in equation 6.2 yields a corrected version that is quadratic in terms of the number of true signal photons:

$$N_A^{Sig} = N_A - \frac{(N_B - c_B N_A^{sig})(N_C - c_C N_A^{sig})}{N_D - c_D N_A^{sig}}. \quad (6.3)$$

Solving equation 6.3 for N_A^{Sig} yields two solutions, of which only one is physical.

The leakage factors derived from the SHERPA simulation are displayed as a function of E_T^γ in figure 6.3, in each tagger discriminant bin. The overall purity as a function of E_T^γ is shown in figure 6.4, while the 2D topography of the purity as a function of E_T^γ and the tagger discriminant is shown in figure 6.5.

Figures 6.6 and 6.7 show the data purity as a function of the JFC discriminant in the first six bins of E_T^γ . These plots show that there is a slight correlation between the tagger discriminant and the photon purity implying that at low E_T^γ HF-like events have fewer fakes.

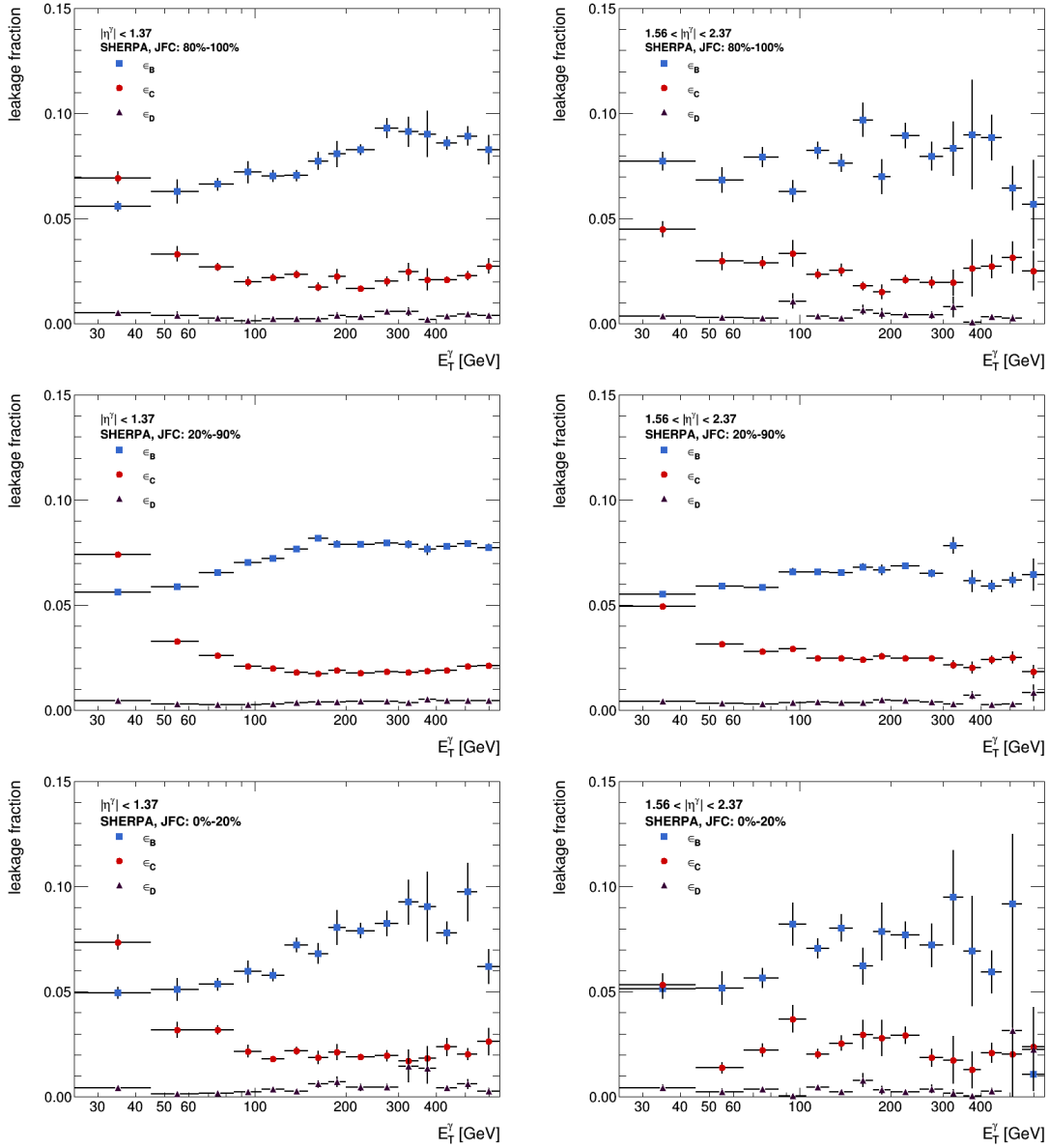


Figure 6.3 – Simulated leakage factors derived from SHERPA for use in the ABCD method in each JFC tagger discriminant bin, for the central and forward regions. The leakage factors are defined as the fraction of simulated signal events in each of the sideband regions to the signal region.

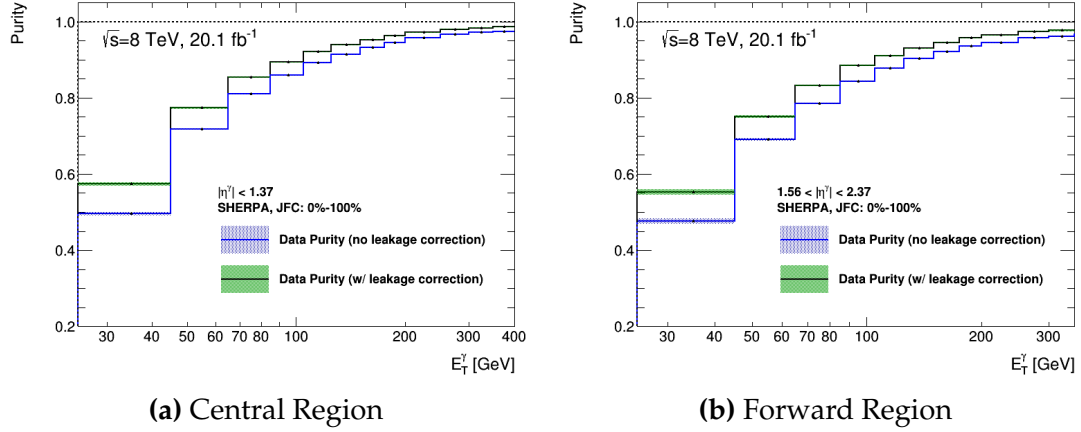


Figure 6.4 – Signal leakage corrected photon purity measured using the ABCD method and SHERPA as a function of E_T^γ for the central and forward regions. The uncertainty bands include both the data and simulation statistical uncertainties.

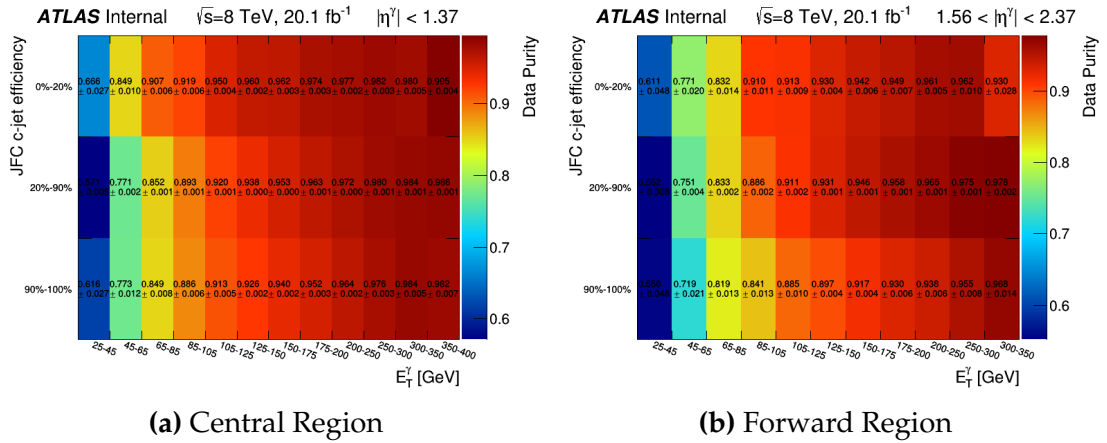


Figure 6.5 – Signal leakage corrected photon purity measured using the ABCD method and SHERPA for the central and forward regions as a function of E_T^γ and as a function of the jet tagging discriminant.

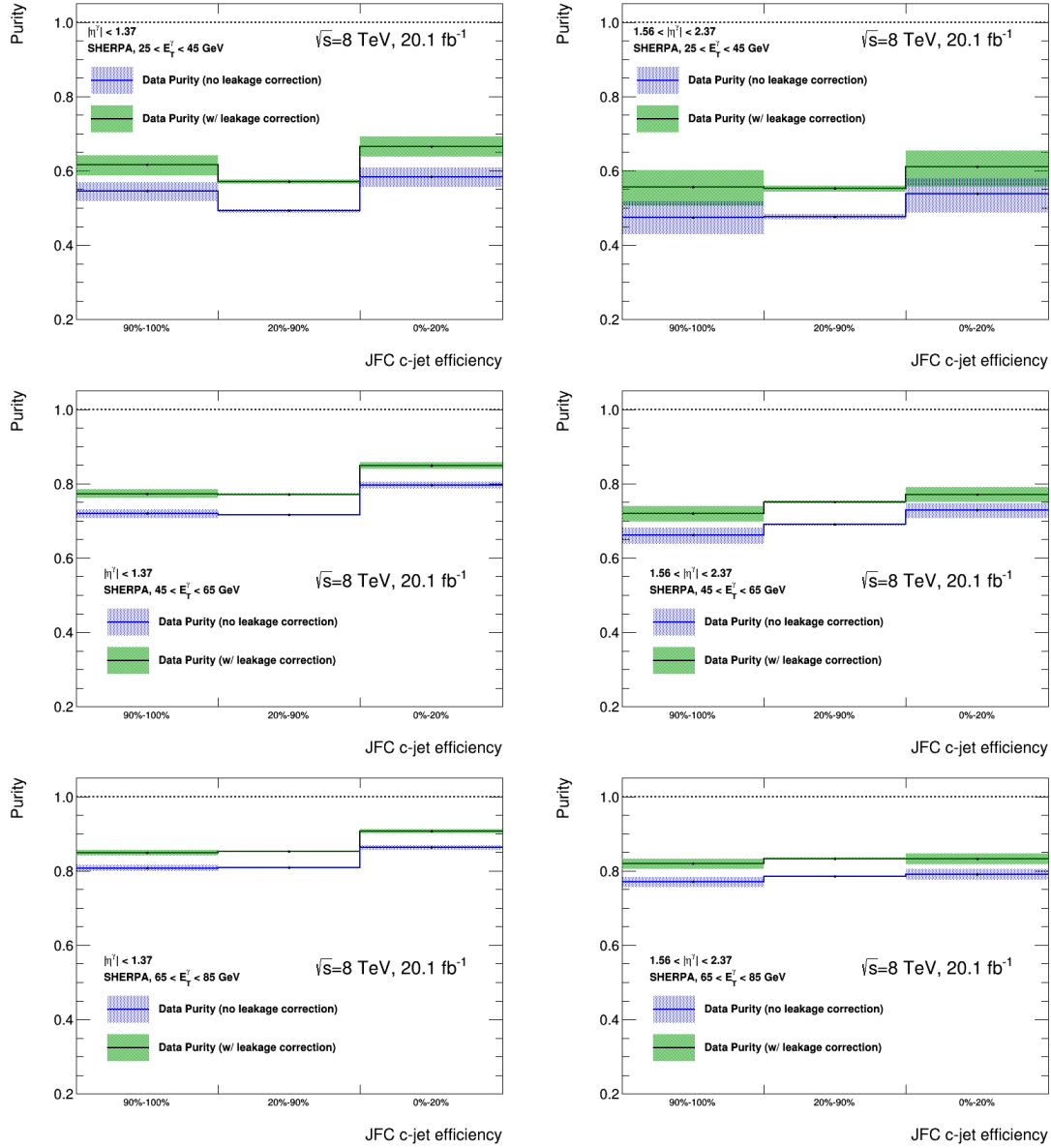


Figure 6.6 – Signal leakage corrected photon purity measured using the ABCD method and SHERPA as a function of the JFC discriminant, in the first three bins of the E_T^γ spectrum, for the central and forward regions. The uncertainty bands include both the data and simulation statistical uncertainties.

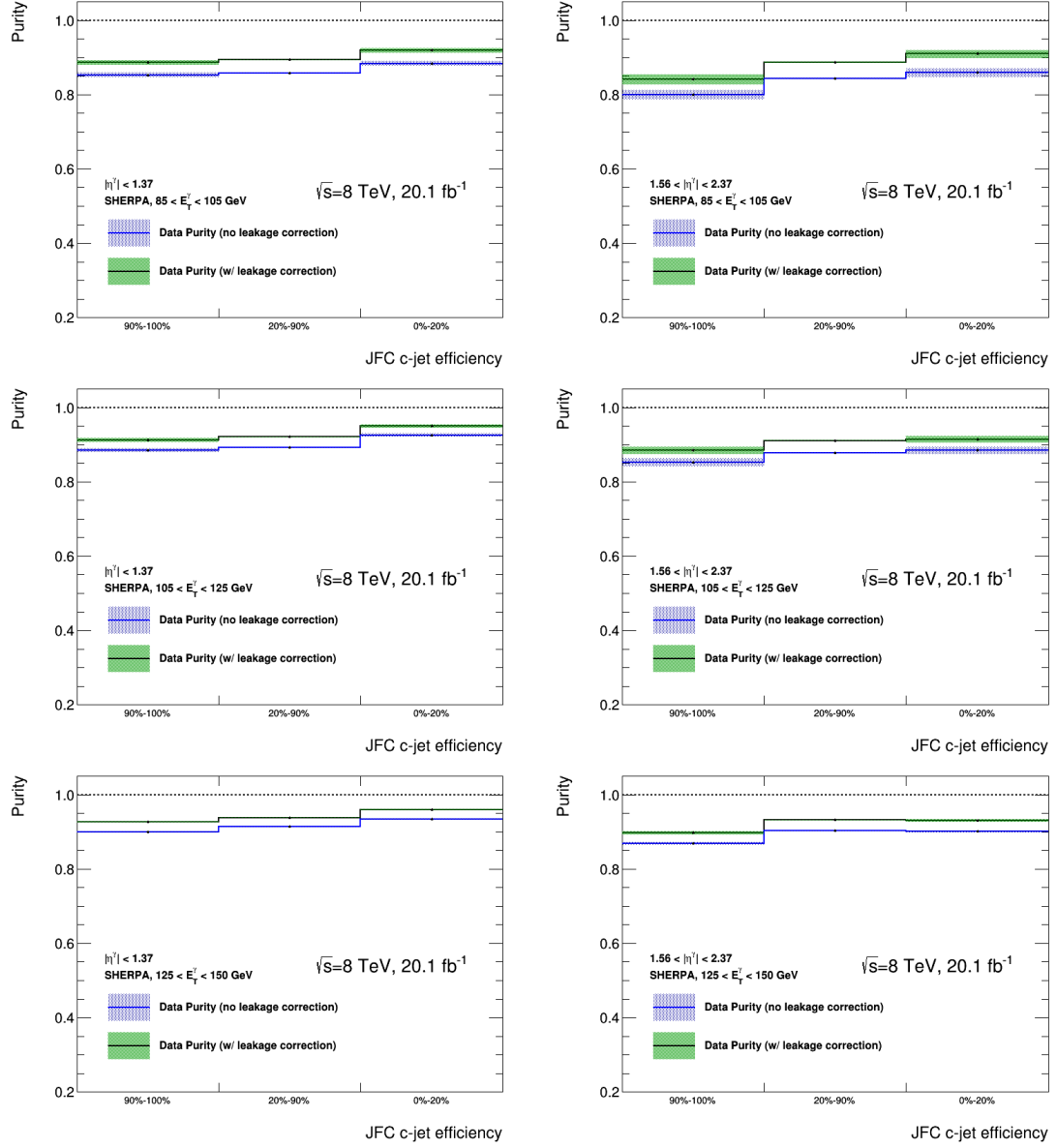


Figure 6.7 – Leakage corrected corrected photon purity measured using the ABCD method and SHERPA as a function of the JFC discriminant, in the fourth, fifth and sixth bins of the E_T^γ spectrum, for the central and forward regions. The uncertainty bands include both the data and simulation statistical uncertainties.

6.3 Jet Purity

The JetFitterCharm tagger (JFC), described in section 6.3, is used to extract the relative flavour fractions (charm, bottom, light) of the data event yield. Until recently, HF-tagging analyses in ATLAS have predominantly made use of HF tagging discriminants by using them to define a selection cut that enhances the purity of the HF signal in the event yield. These cut values are referred to as *operating points*. The efficiency for a given operating point is then estimated using simulated signal samples and corrected using scale factors derived from a dedicated operating point calibration analysis performed by the ATLAS HF-tagging working group. The purity of the event yield is then derived from control regions or simulated estimates [157]. Deriving the purity in this fashion for this $\gamma + c$ analysis, lacking suitable control regions or simulated predictions, is not currently feasible. Thus a different approach was adopted. The so called *continuous* HF-tagging approach uses the full information of the discriminant distribution, instead of applying a cumulative cut, to derive the purity of the sample [158, 159]. The *loose* and *medium* calibrated operating points, described in section , are used to create a continuous tagger discriminant comprising three bins, as illustrated in figure 6.8. Discriminant template shapes for c , b and light flavour jets are then derived from the simulation and used to fit the data discriminant distribution.

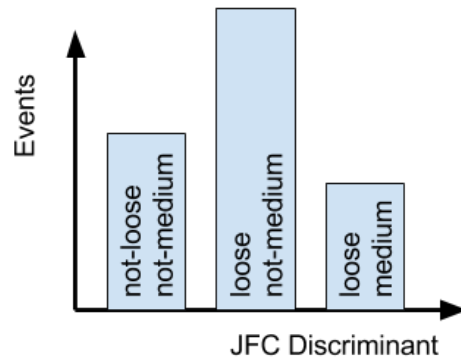


Figure 6.8 – The three bin continuous JetFitterCharm discriminant constructed using the loose and medium operating points.

The efficiency for events of a given flavour falling into each bin is derived by taking the difference between the efficiency of the operating points that define

the bin edges. Using this relation the scale factors for events falling in each bin are derived using equation 6.4, where ϵ_i^{MC} corresponds to the efficiency of a given operating point i , and $SF_i = \frac{\epsilon_i^{Data}}{\epsilon_i^{MC}}$:

$$SF_i^{Continuous} = \frac{SF_i \epsilon_i^{MC} - SF_{i+1} \epsilon_i^{MC}}{\epsilon_i^{MC} - \epsilon_{i+1}^{MC}}. \quad (6.4)$$

The scale factors derived in the JetFitterCharm calibration analysis, that correct the efficiencies in the simulation to those measured in the data, are applied when constructing the templates that are used to fit the data. The uncertainties on the scale factors are then propagated through equation 6.4 to derive the uncertainty associated with the continuous template shapes. This is done by coherently varying the scale factors across both operating points for every uncertainty component of each flavour. These uncertainties are quite large and dominate the precision of the analysis.

The corrected b , c and light tagger discriminant templates are used to perform a template fit that estimates the relative fractions of each flavour in the event yield. The RooFit software package is used to perform the fit [160]. The three fractions of the template fit parametrized by three parameters, the measured yield of two of the flavour fractions and the total normalization, while by construction the third fraction respects the following constraint:

$$1 = f_b + f_c + f_{light}. \quad (6.5)$$

The model is then constructed from the templates of the discriminant shapes for each flavour ($B(x)$, $C(x)$ and $L(x)$), the fractions for the yield belonging to each flavour (f_b , f_c and $f_l = 1 - f_b - f_c$) and an overall normalization factor N :

$$S(x; f_b, f_c) = N(f_b B(x) + f_c C(x) + (1 - f_b - f_c) L(x)). \quad (6.6)$$

Considering that there are three bins in the discriminant and three parameters in the fit there remain no degrees of freedom. This means that the fit minimization converges to a single available solution. The fit is performed by minimizing the

negative log-likelihood function:

$$-\log L(f_b, f_c) = -\sum_{i=1}^3 n_i \log S(x_i; f_b, f_c), \quad (6.7)$$

where n_i is the number of data events in a given discriminant bin. The RooFit algorithm uses the Minuit package to perform the minimization [161].

Figures 6.9, 6.10, 6.11 and 6.12 show the templates in the barrel and end-cap regions following the fit. The resulting measured flavour fractions, as a function of E_T^γ , are displayed in figure 6.13.

6.3.1 Template Fit Closure

The accuracy of the template fit method is tested by using the templates of one simulations to measure the truth fraction of jets in the other. Figure 6.14 shows the results of using PYTHIA to fit SHERPA, while figure 6.15 displays the results for SHERPA fitting PYTHIA. The two simulations contain significantly different flavour fractions, however as figure 6.14 shows, the templates from one do accurately measure the flavour fractions of the other.

To better assess the agreement figures 6.16 and 6.17 display the relative difference between the truth flavour fraction and the measured flavour fraction.

6.3.2 Goodness-of-Fit

The stability the template fits to the data is evaluated by using each fit result as a probability density function to generate 1000 samples of 1,000,000 toy events. For each sample the fit is re-done and the pull $p = \frac{N_{sig}^{fit} - N_{sig}^{data}}{\sigma_{sig}^{fit}}$ is evaluated. Examining the width and mean of the resulting pull distributions provides a measure of the goodness of fit, where a mean of zero indicates that there is no bias, and a width of one indicates that the fit error is correct. Figure 6.18 shows that these criteria are satisfied in the central region up to 400 GeV, while figure 6.19 shows that in the forward region these criteria are satisfied up to 300 GeV. This difference highlights the poorer statistical power of the forward region.

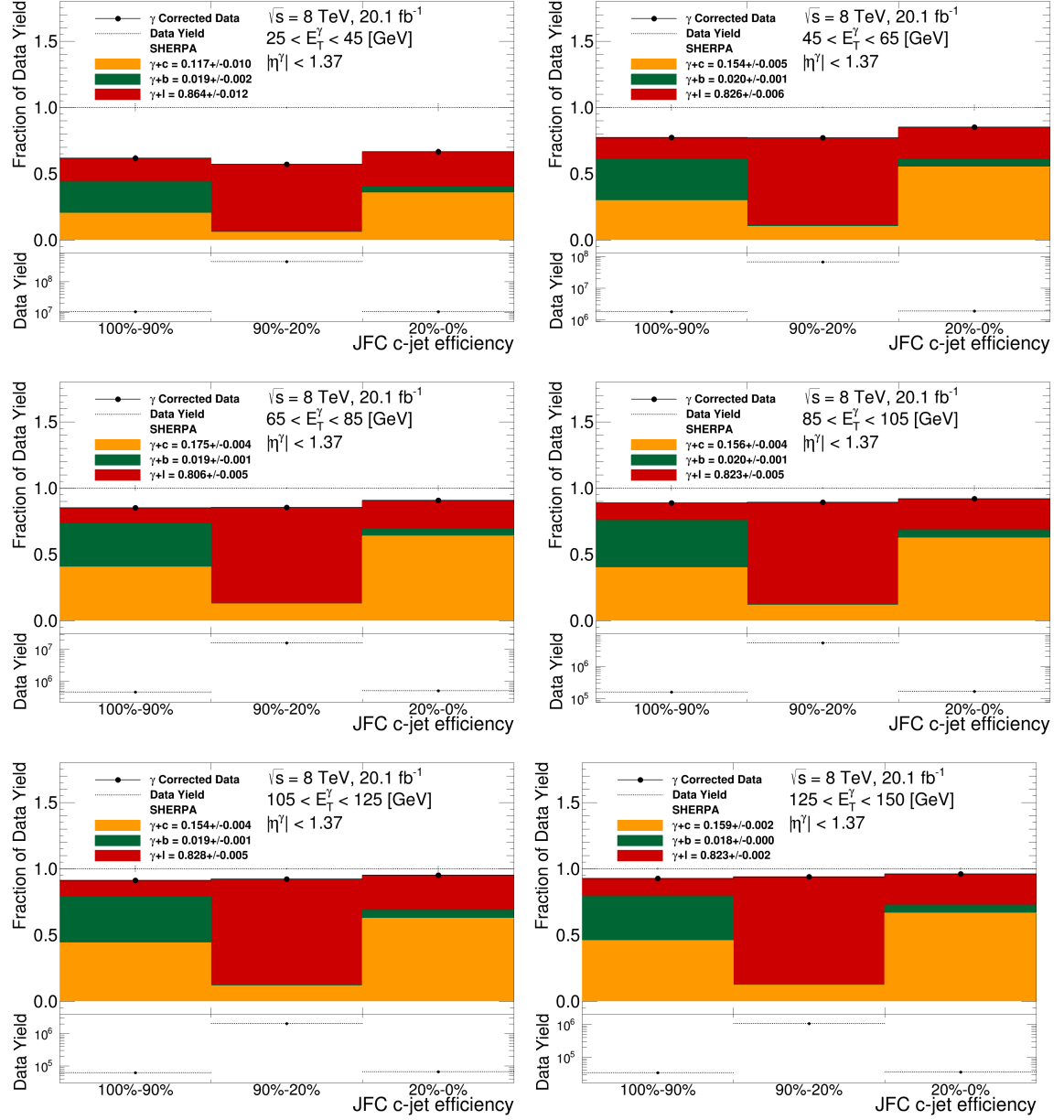


Figure 6.9 – JFC templates after performing the template fit, for $|\eta^\gamma| < 1.37$ and for all bins with $E_T^\gamma < 150 \text{ GeV}$. The fractions for the jet flavours are given in the legend along with their fit uncertainties.

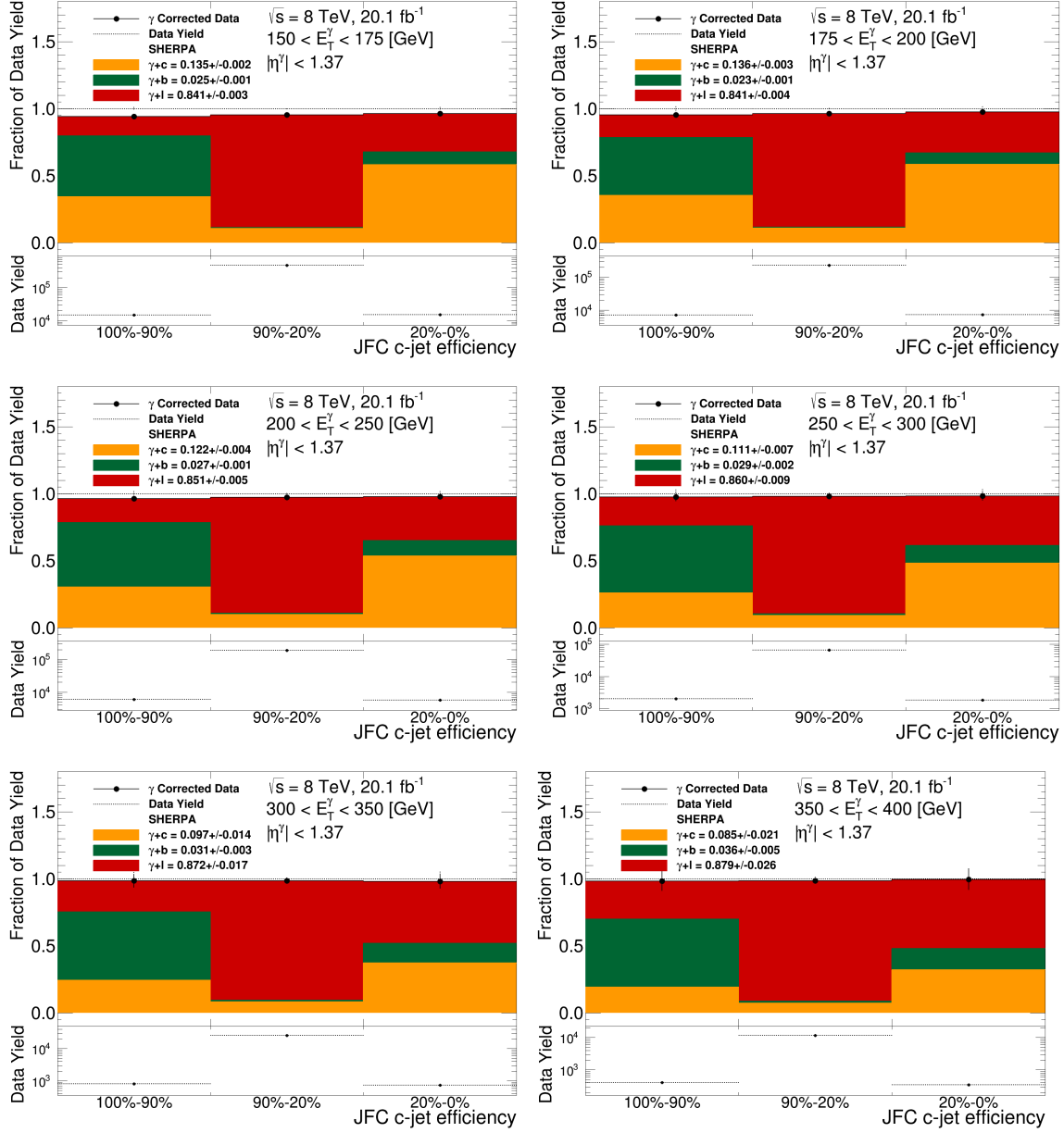


Figure 6.10 – JFC templates after performing the template fit, for $|\eta^\gamma| < 1.37$ and for all bins with $E_T^\gamma > 150$ GeV. The fractions for the jet flavours are given in the legend along with their fit uncertainties.

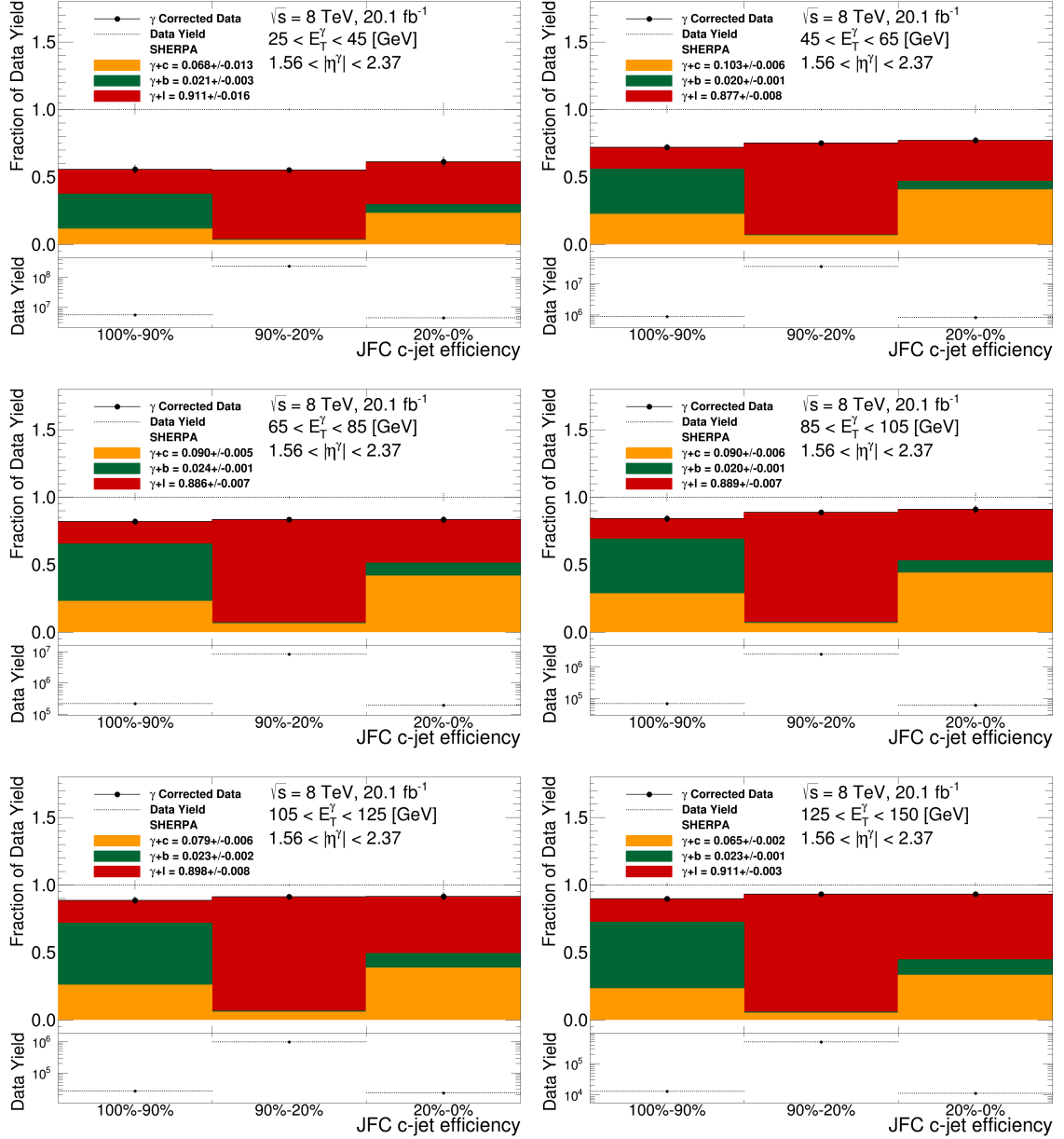


Figure 6.11 – JFC templates after performing the template fit, for $1.56 \leq |\eta^\gamma| < 2.37$ and for all bins with $E_T^\gamma < 150$ GeV. The fractions for the jet flavours are given in the legend along with their fit uncertainties.

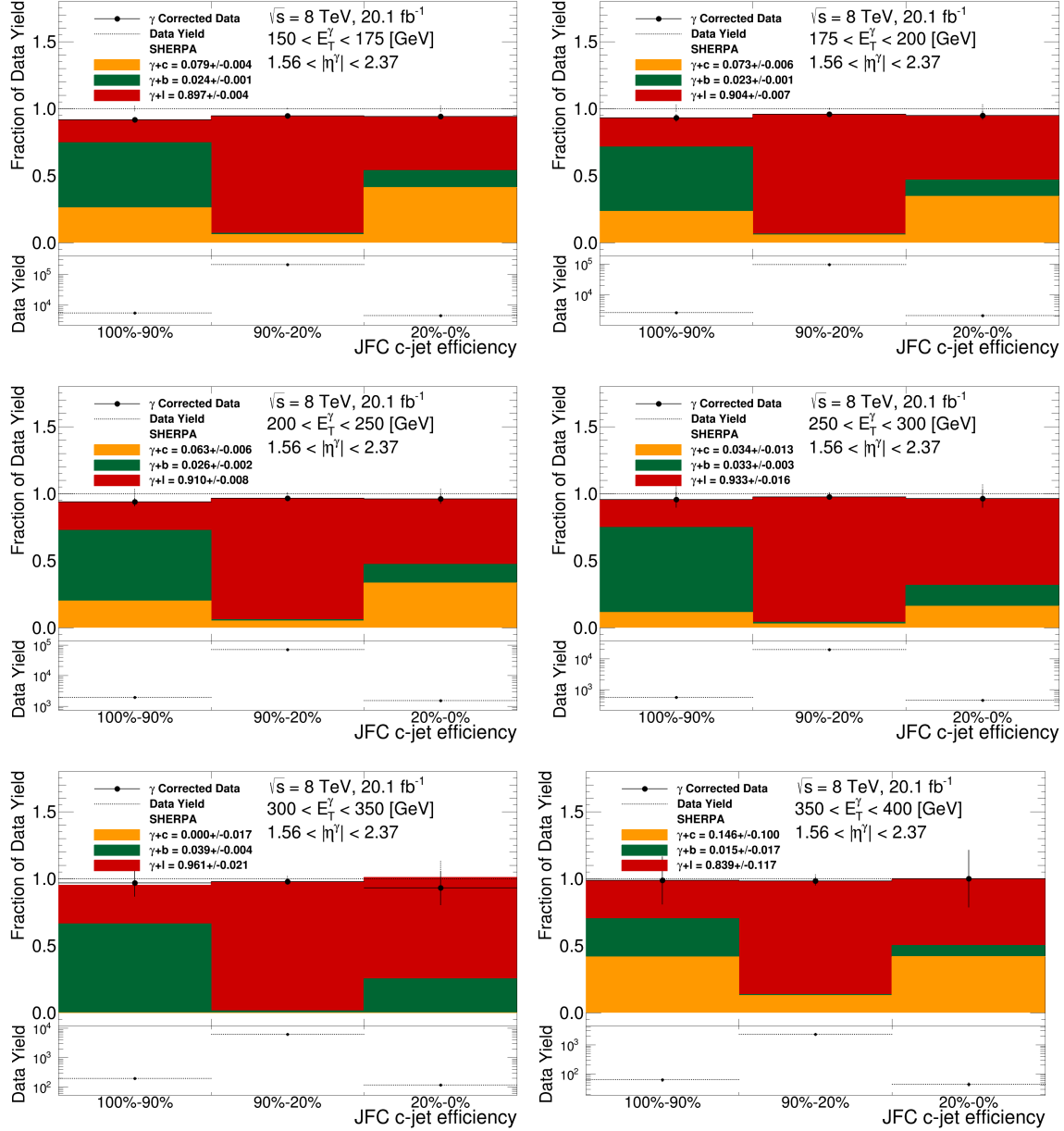


Figure 6.12 – JFC templates after performing the template fit, for $1.56 \leq |\eta^\gamma| < 2.37$ and for all bins with $E_T^\gamma > 150$ GeV. The fractions for the jet flavours are given in the legend along with their fit uncertainties.

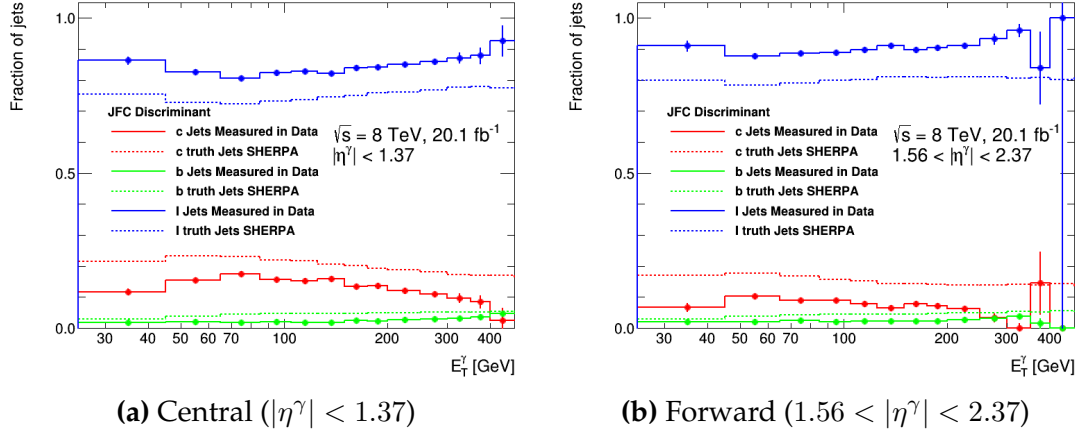


Figure 6.13 – Resulting measured flavour fractions in the data using SHERPA in the central and forward regions.

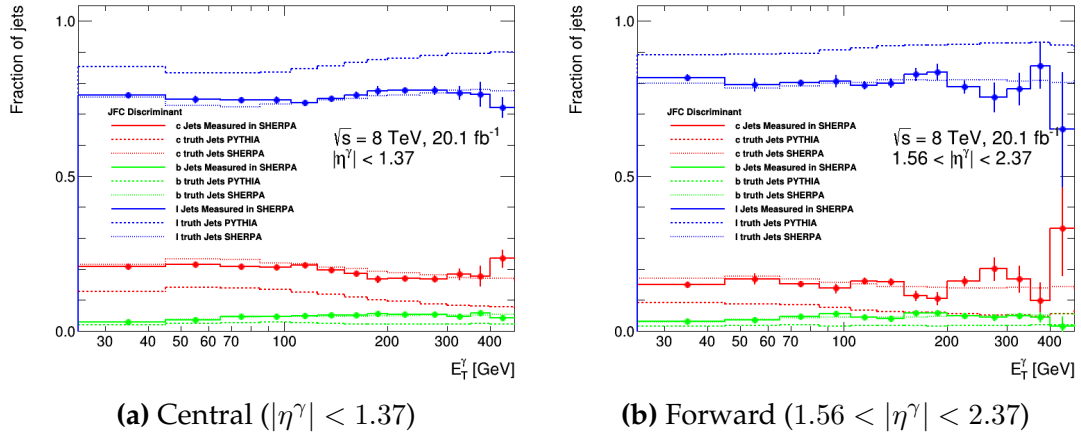


Figure 6.14 – Comparing the measured flavour fractions in SHERPA using the templates from PYTHIA to the truth in SHERPA.

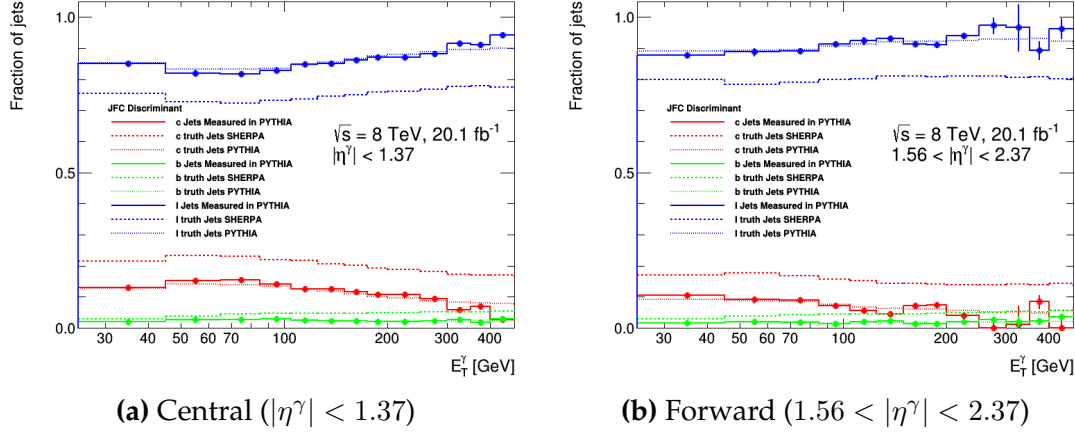


Figure 6.15 – Comparing the measured flavour fractions in PYTHIA using the templates from SHERPA to the truth in PYTHIA.

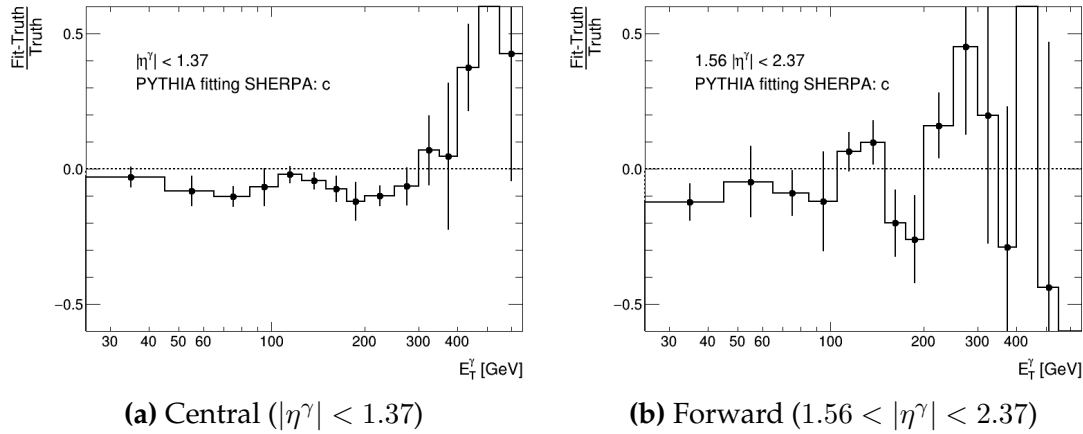


Figure 6.16 – The relative difference between the measured c flavour fraction in SHERPA using the templates from PYTHIA to the truth in SHERPA, displayed with the statistical uncertainty.

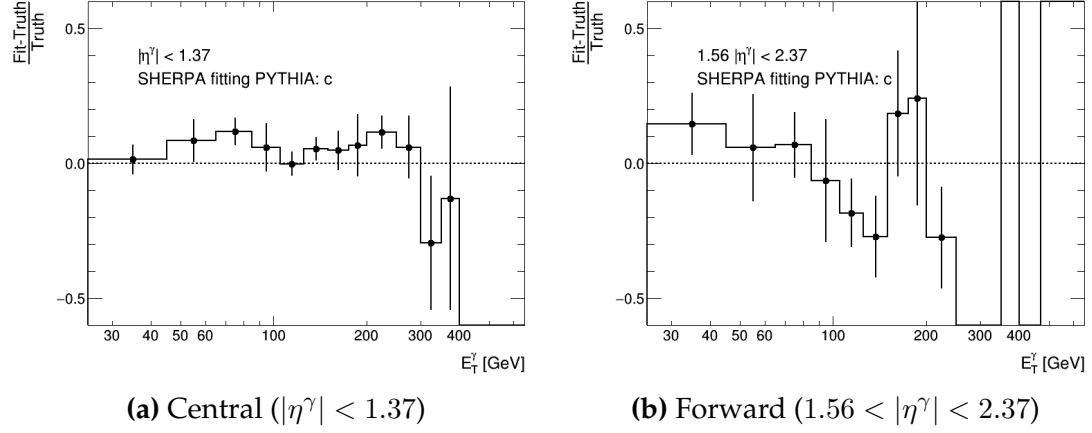


Figure 6.17 – The relative difference between the measured c flavour fraction in PYTHIA using the templates from SHERPA to the truth in PYTHIA, displayed with the statistical uncertainty.

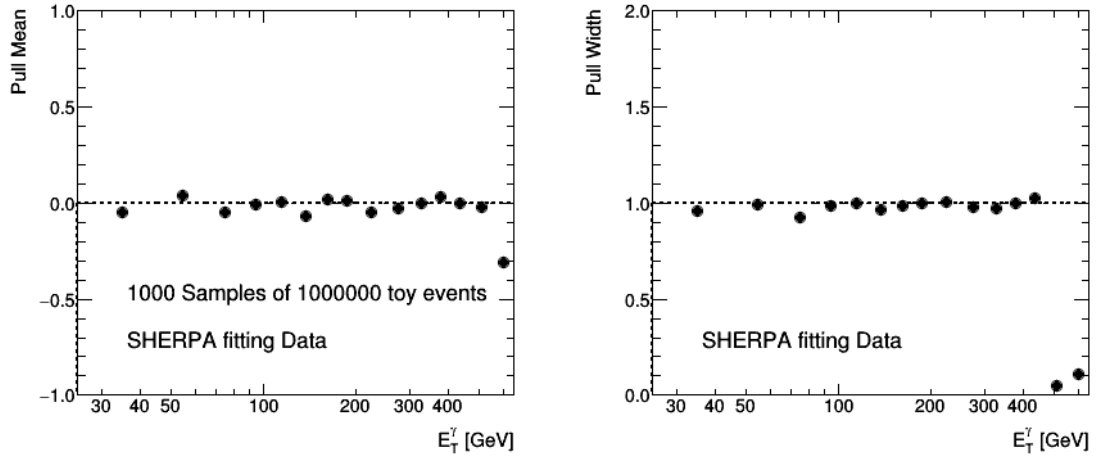


Figure 6.18 – Pull mean and width of the template fits of the data JFC discriminant using SHERPA in the central region, using 1000 samples of 1000000 toy events, for $|\eta^\gamma| < 1.37$.

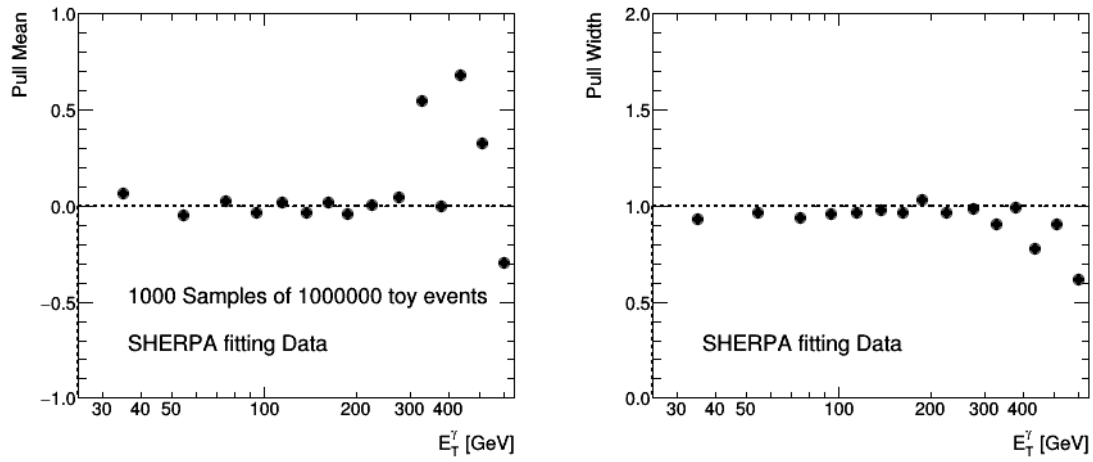


Figure 6.19 – Pull mean and width of the template fits of the data JFC discriminant using SHERPA in the forward region, using 1000 samples of 1000000 toy events, for $1.56 < |\eta^\gamma| < 2.37$.

6.4 Unfolding

Having corrected for background photons in the signal region and having measured the heavy flavour jet fractions, the event yield for detector-level $\gamma+c$ -jet signal is obtained. The final step is to correct the measurement for detector effects yielding a measurement that is as directly comparable as possible to other experimental results and theoretical predictions. This procedure, referred to as unfolding, aims to account for the resolution (bin migrations) and inefficiencies of the detector (signal reconstruction efficiency). There are numerous approaches to unfolding, each with its benefits and drawbacks. Typically the trade off between different approaches lies between the complexity of the approach and the inherent bias coming from the simulated signal shape. The choice of unfolding procedure must be made considering the characteristics of the measured observable.

The response matrices, seen in figure 6.20, display the 2D distribution of events that pass both the detector-level and the particle-level cut flow in the simulation. When constructing these distributions a detector-level to particle-level matching is applied, requiring that the leading photon and leading jet at the detector-level must fall within $\Delta R < 0.2$ of the particle-level leading photon and leading jet, respectively. The signal reconstruction efficiency, that is the fraction of events that pass the detector-level cut flow having already passed the particle-level cut flow, is shown figure 6.21. The observed shift in reconstruction efficiency between the jet flavours arises due to a flavour dependent energy shift that occurs at reconstruction level due to unmeasured neutrinos and muons arising from secondary decays for c -jets, and secondary/tertiary decays for b -jets.

Figure 6.20 displays that the response matrix is diagonal in shape and that the off-diagonal bins are much smaller than the adjacent diagonal bins. This indicates a robust correspondence between detector and particle-level E_T^γ and that the migrations between bins are small.

The resolution of detector-level photons is measured in each E_T^γ bin by fitting a Gaussian distribution to the distribution of the scaled difference between the detector-level and the particle-level E_T^γ : $Res = \frac{E_T^{Reco} - E_T^{Particle}}{E_T^{Particle}}$. Figure 6.22 displays that the measured resolution of photons in the signal region, taken as the width of the Gaussian fit in each bin of E_T^γ , is much smaller than the E_T^γ spectrum bin width.

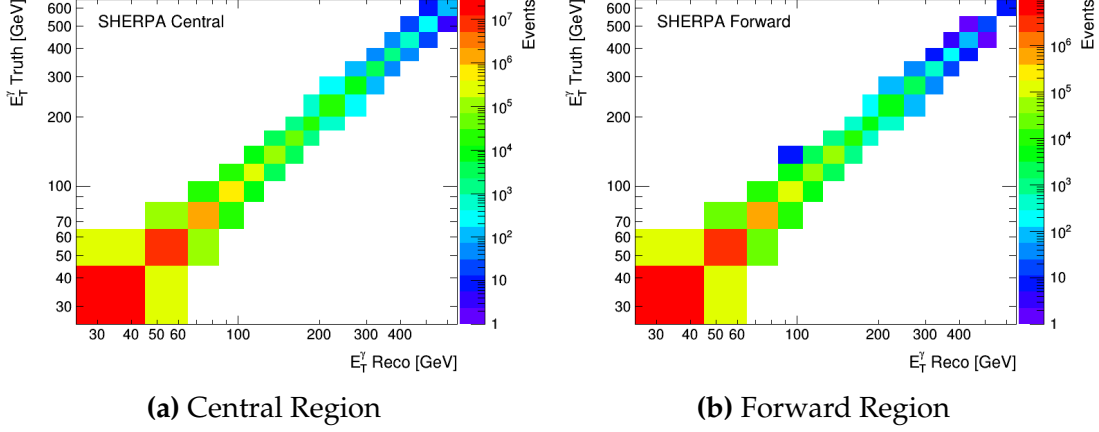


Figure 6.20 – Response matrix for the charm signal, using SHERPA, in the central and forward regions.

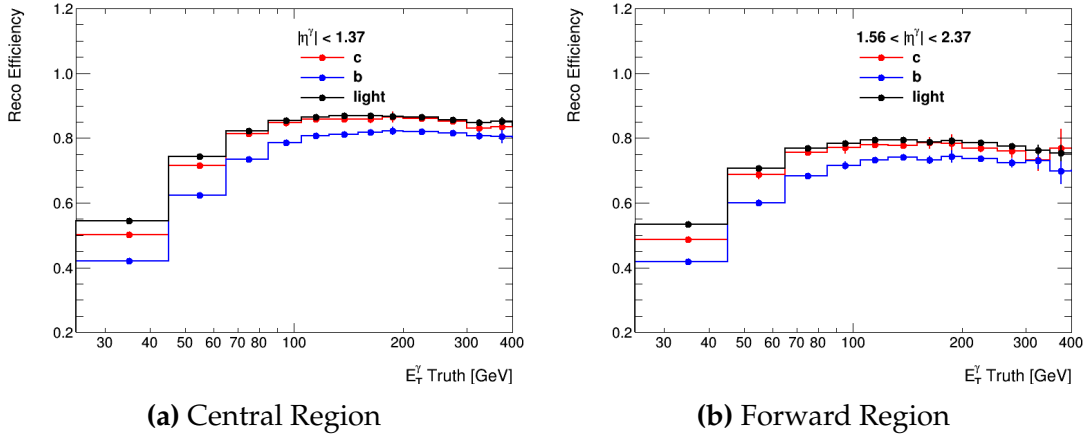


Figure 6.21 – The efficiency of events that pass the signal selection at particle-level subsequently satisfying the signal selection at detector-level in SHERPA.

The Bin-By-Bin unfolding approach has previously been used for SM photon results in ATLAS [61–65]. This approach involves constructing correction factors as the ratio between the particle-level ($N_i^{\text{particle}, \gamma+c\text{-jet}}$) and detector-level ($N_i^{\text{detector}, \gamma+c\text{-jet}}$) E_T^γ distributions derived using the simulation for each bin i :

$$C_i^{\text{unf}} = \frac{N_i^{\text{particle}, \gamma+c\text{-jet}}}{N_i^{\text{detector}, \gamma+c\text{-jet}}}. \quad (6.8)$$

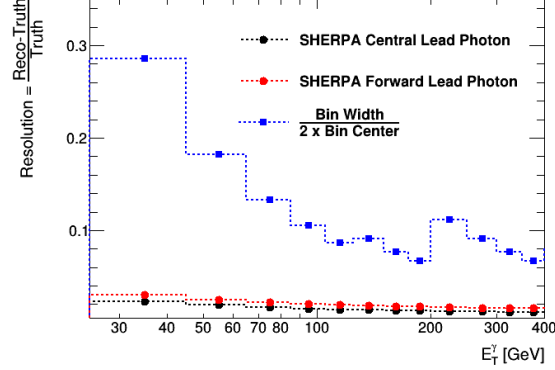


Figure 6.22 – Measured resolution of the detector-level E_T^γ in the SHERPA simulation for events satisfying the signal criteria, compared to the bin width.

This approach is equivalent to assuming that the migration matrix is diagonal, i.e. that detector level bin migration effects are small since correlations between adjacent bins are neglected. In this analysis this condition is met since the E_T^γ resolution is smaller than the bin width, as shown in figure 6.22. For the remaining contents of this section the Bin-By-Bin result is used as a benchmark for comparisons.

The Bayesian unfolding approach, described by d’Agostini in [162, 163], is a more sophisticated technique that uses an iterative application of Bayes’ theorem to infer the true distribution from the measured. This approach takes into account bin migrations and smearing, while also inferring contributions from background and signal losses due to reconstruction inefficiency. Starting with Bayes’ theorem [162]:

$$P(P_i|R_j) = \frac{P(R_j|P_i)P(P_i)}{\sum_k^{n_P} P(R_j|P_k)P(P_k)}, \quad (6.9)$$

$P(P_i|R)$ is referred to as the “smearing matrix”, that is the probability of the underlying particle-level processes P_i are the source of the measured detector-level signal R . $P(R|P_i)$ is the aforementioned response matrix, that is the response matrix derived from the simulation that is independent of the data and describes the probability of measuring a detector-level signal R given a particle-level cause P_i . $P(P_i)$ is the “prior”, that is the best description of the underlying particle-level signal. $\sum_{l=1}^{n_P} P(R|P_l)P(P_l)$ is an overall normalization factor. Given detector-level signal $n(R_j)$, the particle-level distribution $n(P_i)$ is calculated using the smearing

matrix taking into account the inefficiency ϵ_i :

$$n(P_i) = \frac{1}{\epsilon_i} \sum_{j=1}^{n_R} n(R_j) P(P_i | R_j). \quad (6.10)$$

This estimate of the particle-level distribution can then be used as the final result or it can be used to update the prior for a subsequent iteration of this procedure. At each iteration the χ^2 is calculated between the previous and current "particle" level distributions and is used to evaluate the convergence of the unfolding. Ideally the number of iterations is chosen at the point where the difference in χ^2 begins to show only incremental improvements, since this is a signal that the probabilistic inference has saturated. After too many iterations the procedure converges to the inverse of the migration matrix, which is in most cases inherently unstable, yielding constant fluctuations with each subsequent iteration. Examining the convergence of the χ^2 as a function of the unfolding iterations, shown in figure 6.23, a linear reduction in χ^2 is observed when displayed on a log y-axis scale. As this indicates no strong shift from meaningful to incremental improvements, it is concluded that 2 iterations is sufficient.

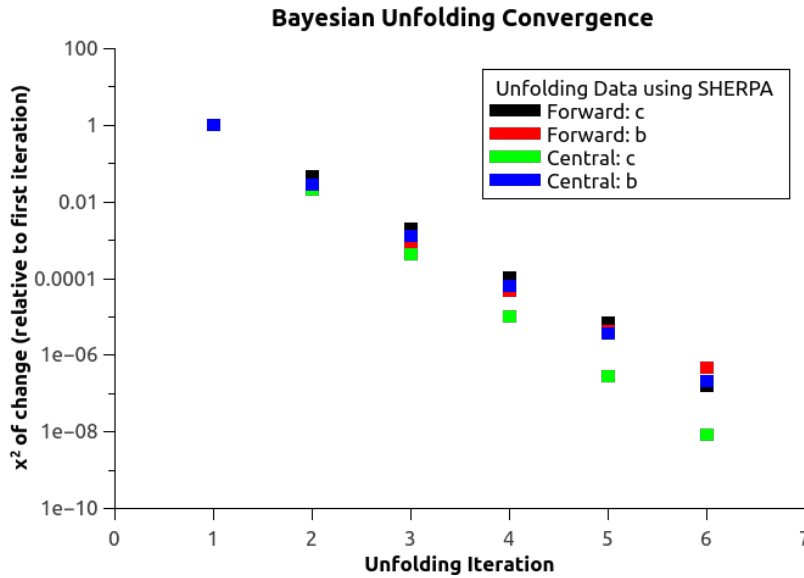


Figure 6.23 – Convergence of the χ^2 of the Bayesian unfolding vs the number of iterations, for SHERPA unfolding the data, for both the c and b signals in the forward and central regions.

Figures 6.24 and 6.25 display the unfolding factors derived by both methods as well as the relative difference between the two which shows that they yield consistent results. The Bin-By-Bin approach is adopted for the nominal result, remaining consistent with the existing ATLAS SM photon and photon+jet analyses [61–65].

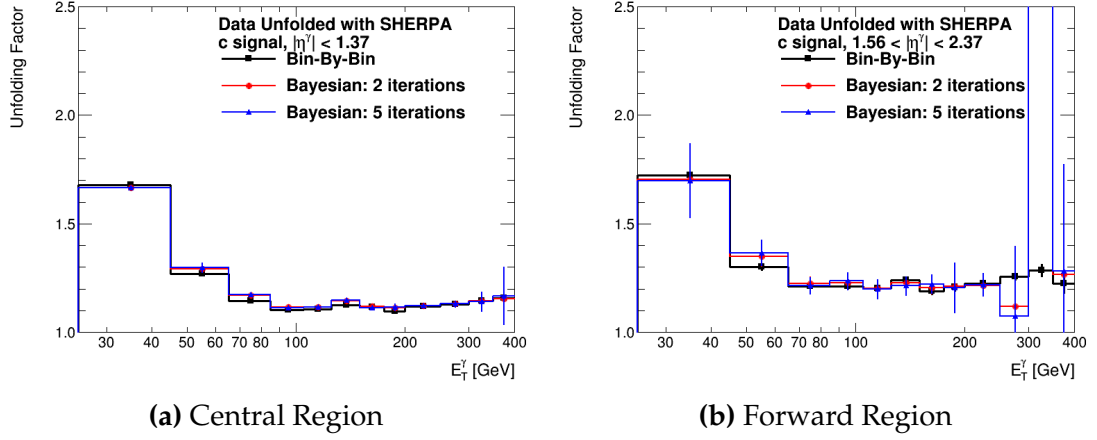


Figure 6.24 – Unfolding factors for the charm signal derived using the Bin-By-Bin method and the Bayesian method (with 2 and 5 iterations) using SHERPA. The statistical uncertainty is shown evaluated using the Bootstrap method.

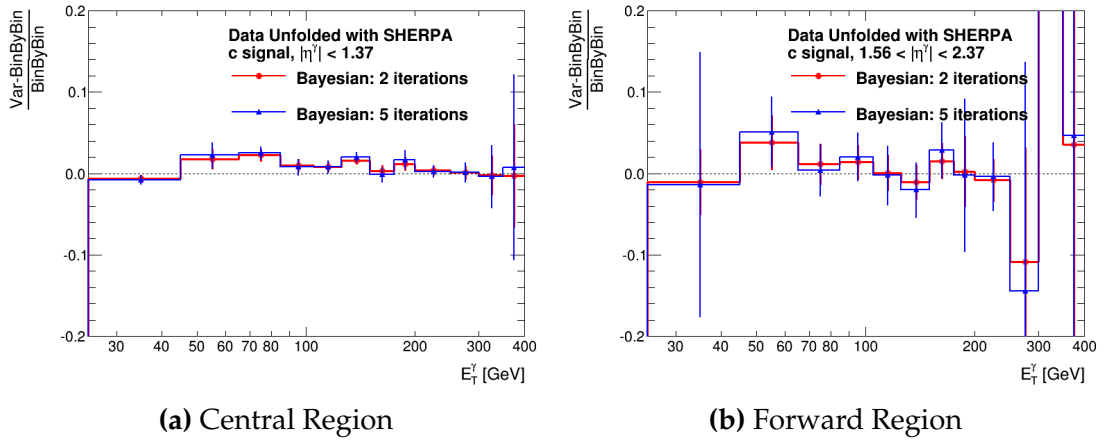


Figure 6.25 – Unfolding relative difference between the bin-by-bin unfolding factor and the Bayesian unfolding factors (with 2 and 5 iterations) using SHERPA. The statistical uncertainty is shown evaluated using the Bootstrap method.

As a cross check the unfolding factors are derived using the PYTHIA simu-

lation and compared to the nominal factors derived using SHERPA, seen in figure 6.26. These figures show an acceptable agreement for the bin-by-bin method.

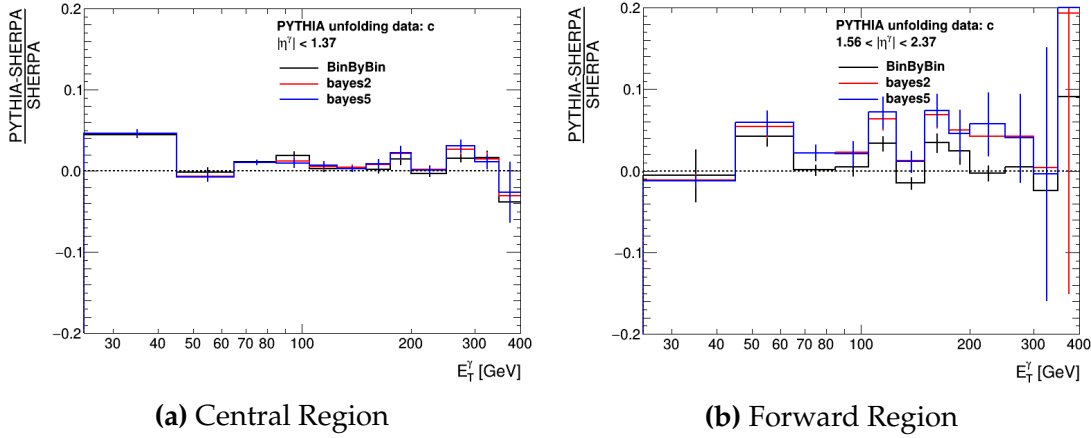


Figure 6.26 – Comparison between the unfolding factors obtained unfolding the data using PYTHIA to those using SHERPA.

The PYTHIA simulation was then used to unfold the SHERPA detector-level distribution and compared to the particle-level, and vice versa, shown in figures 6.27 and 6.28. Here a relatively good agreement is observed, and generally a better agreement is seen using the bin-by-bin unfolding method than the Bayesian. The differences observed in these cross are accounted for with inclusion of the physics modeling uncertainties described in sections 7.2.3 and 7.2.4.

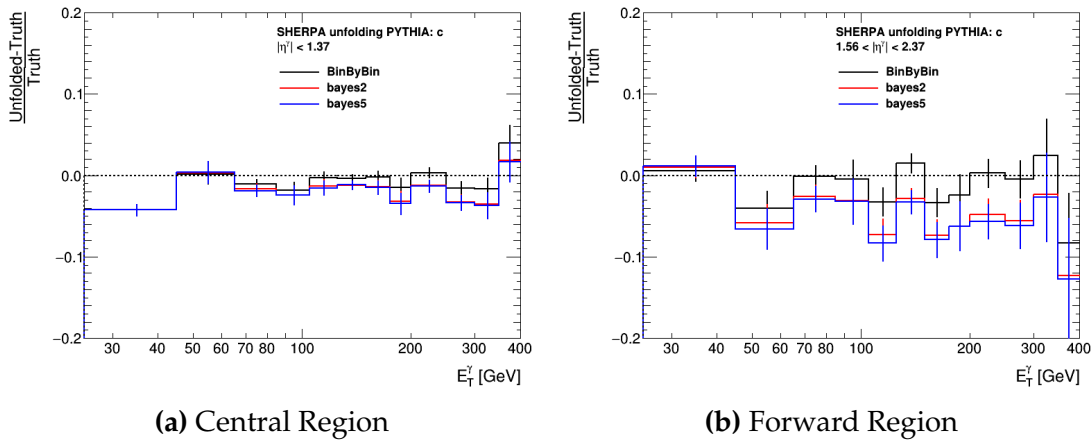


Figure 6.27 – Comparison using the SHERPA simulation to unfold the PYTHIA detector-level distribution to the particle-level distribution in PYTHIA.

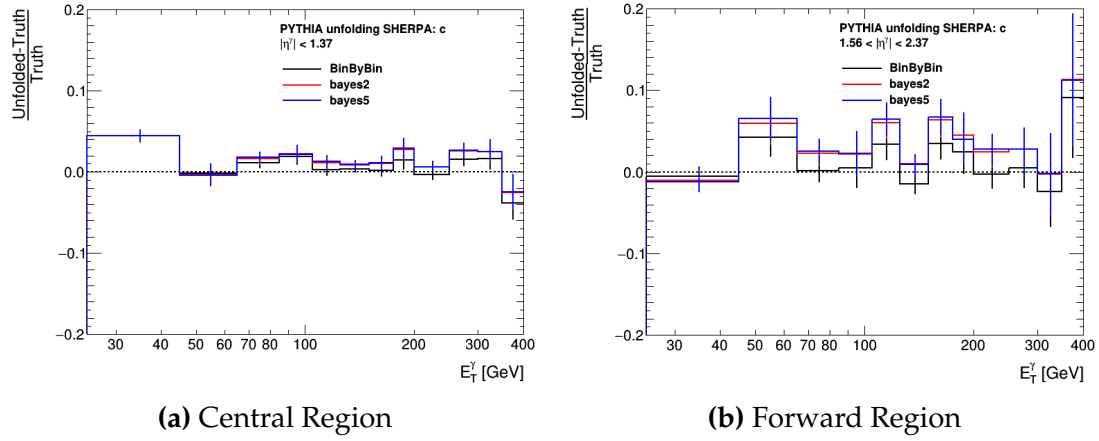


Figure 6.28 – Comparison using the PYTHIA simulation to unfold the SHERPA detector-level distribution to the particle-level distribution in SHERPA.

Now with all the components of the measurement procedure described, the following section discusses the impact of the uncertainties associated to these components.

Uncertainties

7.1 Statistical

The cross section measurement is a combination of various separate measurement and correction steps. When evaluating the statistical uncertainty of the measurement, correlations between these different steps must be taken into account since they all consider the same data. The bootstrap method, which implicitly takes into account these correlations, is used to derive the statistical uncertainty.

When the data or the simulated samples are processed, the necessary output histograms that are used to compute the measurement are created and filled. The bootstrap method creates N replicas of the output histograms that are filled simultaneously whenever the nominal output histogram is filled. While the nominal set of output histograms are always filled as usual with the nominal event weight of 1, each replica is filled using a random unique weight drawn from a Poisson distribution with an interval of one. The result is a collection of N statistically equivalent sets of output histograms that can be thought of as statistically equivalent "parallel statistical universes" where the measurement is performed. The full analysis is then performed on each replica yielding a statistical sample of N complete measurements. The statistical uncertainty on the nominal measurement is taken as the root mean square of the resulting sample of final measurements. This technique is used to derive the data and simulation statistical uncertainties separately and simultaneously by re-sampling the data and simulated input distributions separately and simultaneously. Figures 7.1 and 7.2 show the relative data and simulation statistical uncertainties evaluated using 10000 replicas on the measured cross section using SHERPA in the central and forward regions.

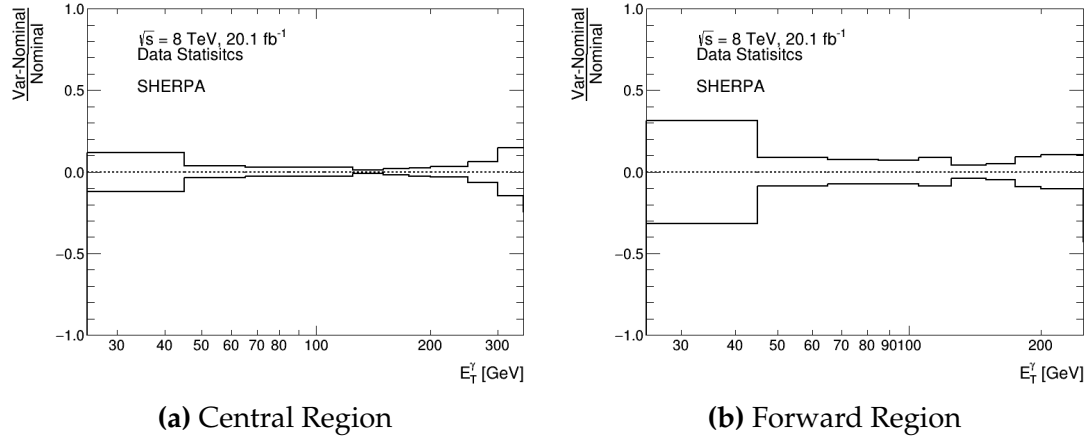


Figure 7.1 – Data statistical uncertainty arising in the cross section using SHERPA in the central and forward regions.

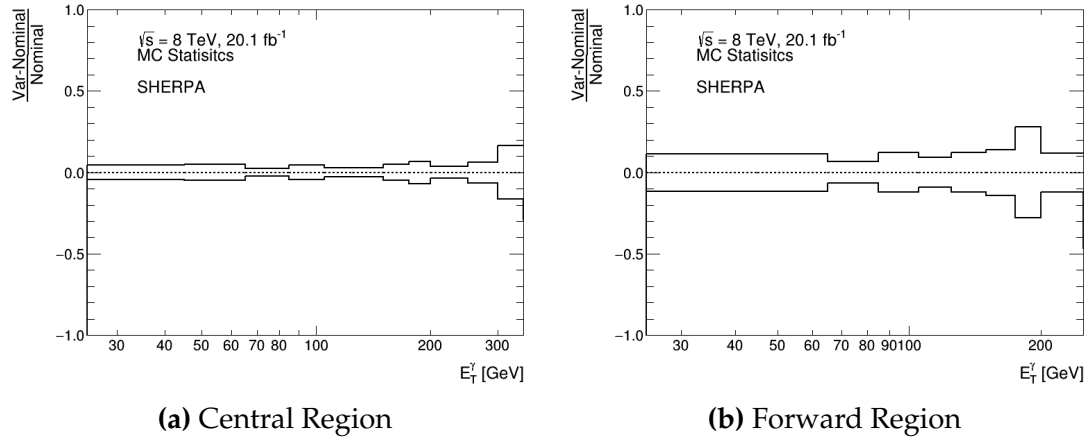


Figure 7.2 – Simulation statistical uncertainty arising in the cross section using SHERPA in the central and forward regions.

7.2 Systematic

The numerous sources of systematic uncertainty considered in this analysis are described in the following sections and follow a common approach. For each source, some parameter of the analysis is altered and the effect is measured. The significance of this effect is established by rebinning and then smoothing the variation according to its statistical uncertainty with the following procedure.

- Derive the statistical uncertainty on this relative difference between the variation and the nominal measurement using the bootstrap technique

with 1000 replicas.

- Rebin the variation such that the value of all bins is greater than two times its standard deviation, or put it to zero if this cannot be achieved. This is performed by initiating the rebinning from both the right and left hand sides of the distribution, taking the result that ends up with more significant bins.
- The distribution is then rebinned to the initial bins, assigning to the initial bins the new values.
- A Gaussian-weighted sliding average is performed that smears adjacent bins such that the structure of the variation is not given by the rebinned bin edges.

The smoothed variations are then added in quadrature to provide the total uncertainty on the measurement.

7.2.1 JetFitterCharm Efficiency Scale Factor

The dominant systematic uncertainties in the analysis arise from the calibration of the JFC scale factors (SF) which affect the analysis by changing the shape of the simulated JFC tagger templates that are used to extract the flavour fractions of the data yield. There are separate sets of scale factors for each flavour of jet as the factors are derived from separate flavour-specific tagger performance studies [?, ?, 146–148].

Light

The light-jet JFC calibration is the largest source of systematic uncertainty, displayed in figure 7.3. The calibration is derived from a negative-tag analysis considering multi-jet and D^* events [147]. The light-jet calibration is different between the forward and central regions due to the different response for light-jets in these detector regions. The most significant sources of uncertainty for the calibration, in order, arise due to the modeling of the track multiplicity, the impact parameter smearing, the data-taking period dependence, the jet energy resolution, the simulation statistics, and the c -jet efficiency. In total there are 15 individual parameters of the light jet calibration that are varied up and down; the magnitudes of the vari-

ations in the central values as a result of changes to those parameters were added in quadrature.

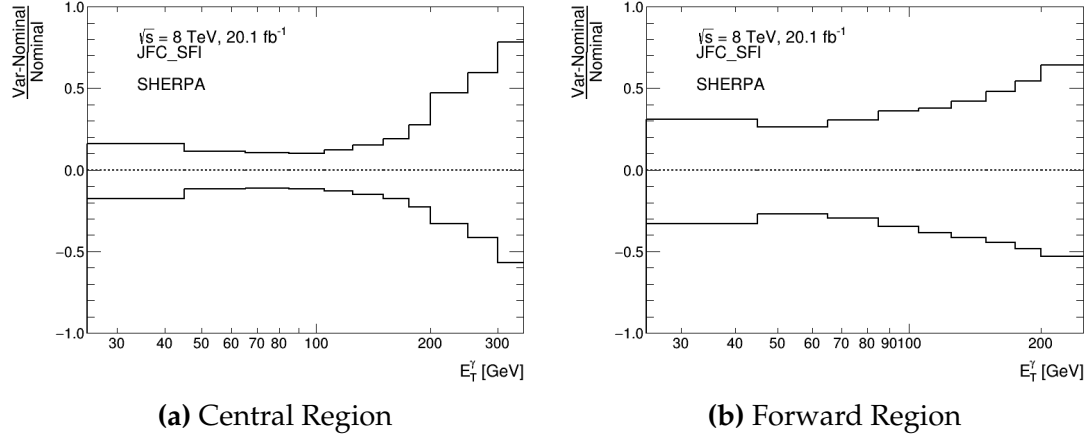


Figure 7.3 – Total systematic uncertainty arising in the cross section when adding all JFC light scale factor variations in quadrature.

Charm

The c -jet JFC calibration is the second largest source of systematic uncertainty, displayed in figure 7.4. The calibration is derived from an analysis of reconstructed D^{*+} meson decays [147]. The primary sources of uncertainty arise due to the D^{*+} mass fit and background parametrization, the b -jet efficiency SFs, the b lifetime and the pseudo-proper time resolution in the simulation, the jet energy scale and resolution, the jet vertex quality, and the modeling of the c -quark fragmentation (extrapolation) functions. In total there are 20 individual parameters of the c -jet calibration that are varied up and down; the magnitudes of the variations in the central values as a result of changes to those parameters were added in quadrature.

Bottom

The systematic uncertainty associated to the b -jet JFC calibration is displayed in figure 7.5. The calibration is derived from samples of leptonically decaying $t\bar{t}$ events using a combinatorial likelihood technique [148]. There are many sources

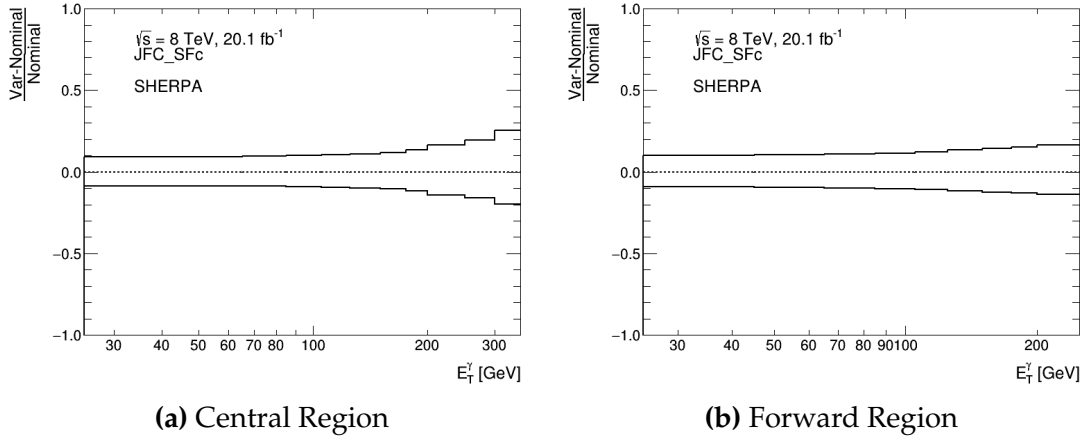


Figure 7.4 – Total systematic uncertainty arising in the cross section when adding all JFC charm scale factor variations in quadrature.

of uncertainty taken into account in this calibration related to the simulation modeling, the normalization of different background sources, pileup reweighting and the jet and lepton reconstruction and calibration. Of these sources they all contribute and none of them are dominant. In total there are 60 individual parameters of the b -jet calibration that are varied up and down; the magnitudes of the variations in the central values as a result of changes to those parameters were added in quadrature.

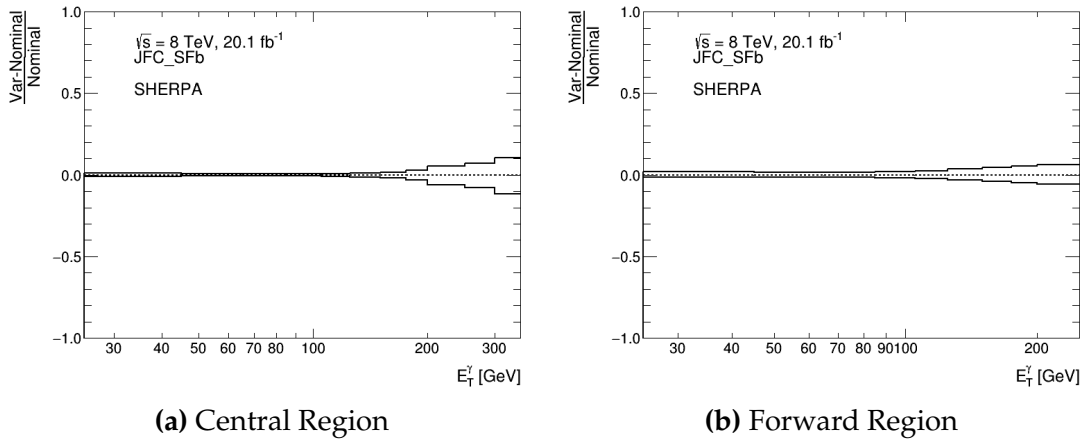


Figure 7.5 – Total systematic uncertainty arising in the cross section when adding all JFC bottom scale factor variations in quadrature.

Extrapolation

The tagging calibration for b - and c -jets ends at 300 GeV while the light-jet calibration ends at 750 GeV. For jets whose p_T^{jet} exceed these thresholds the SF of the highest calibrated bin is used and an additional extrapolation uncertainty is applied. This uncertainty is derived from a dedicated simulation driven extrapolation analysis performed by the HF tagging working group. The impact of these uncertainties on the final measurement is shown in figures 7.6, 7.7 and 7.8.

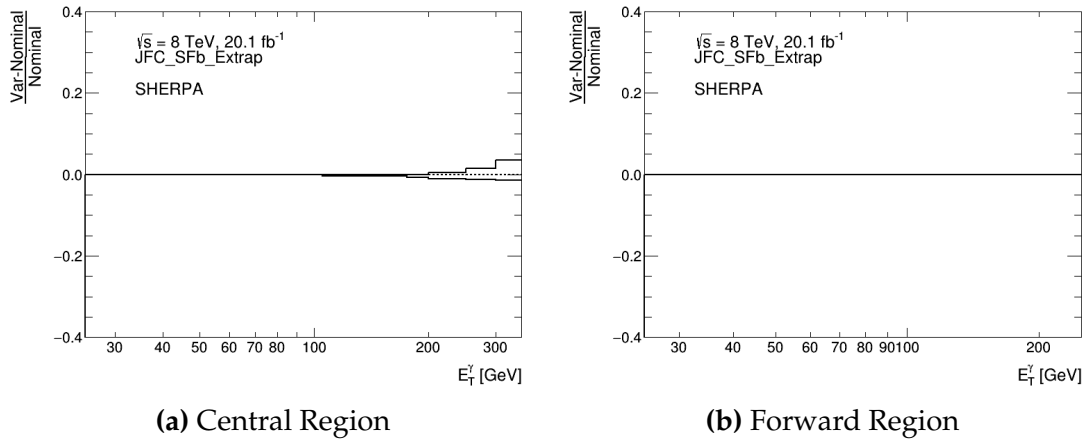


Figure 7.6 – Extrapolation uncertainty for the bottom SF.

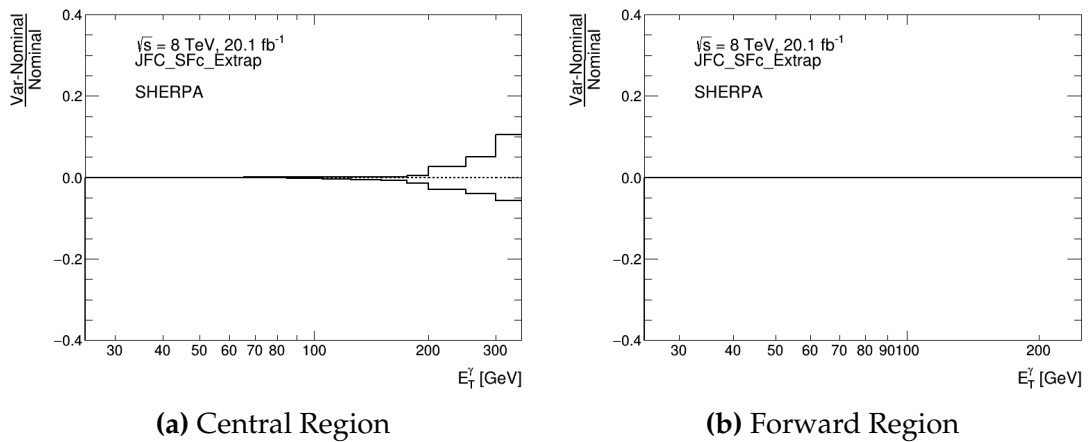


Figure 7.7 – Extrapolation uncertainty for the charm SF.

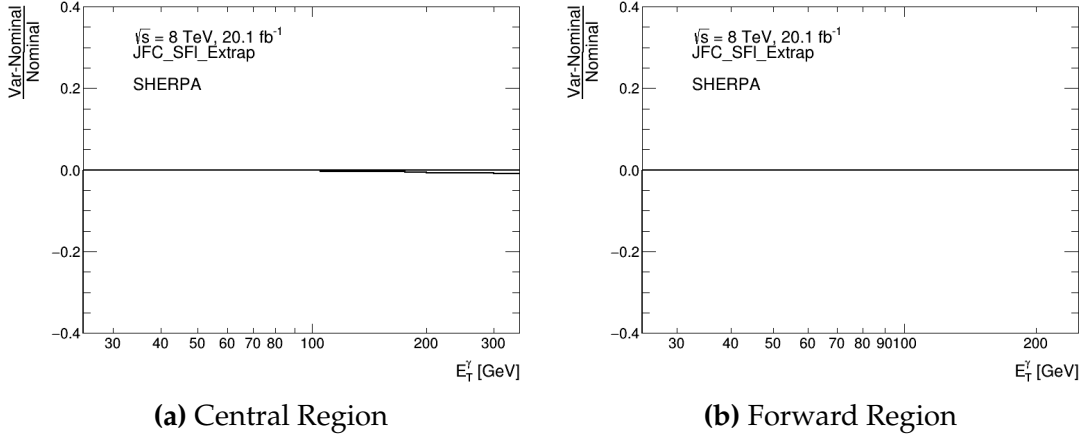


Figure 7.8 – Extrapolation uncertainty for the light SF.

7.2.2 Background Photon Sideband Correlation

The 2D sideband photon purity method described in section 6.2 relies on the assumption that the variables used to generate the four regions are uncorrelated when considering background photon signals. If the assumption is valid then the R_{bkg} variable would be equal to unity, where N_r^{bkg} represents the yield of background events in a given region:

$$R_{bkg} = \frac{N_A^{bkg} N_D^{bkg}}{N_B^{bkg} N_C^{bkg}} = 1. \quad (7.1)$$

The validity of this assumption is tested by examining the behavior of an analogous quantity constructed entirely from the non-isolated background regions, shown in figure 7.9. The regions used for this plot are defined by dividing the existing non-isolated regions with a cut at $E_T^{iso} > 16.8 \text{ GeV} + 0.0042 \times E_T^\gamma$. To account for sensitivity to this effect the R_{bkg} quantity in the purity calculation, which takes a value of 1 for perfectly uncorrelated control regions, is varied up and down 10%. Figure 7.10 shows the impact of this variation on the final cross section.

7.2.3 Prompt Photon Modeling

The simulations used in this analyses take different approaches to the modeling of prompt photon production beyond LO. In SHERPA these contributions

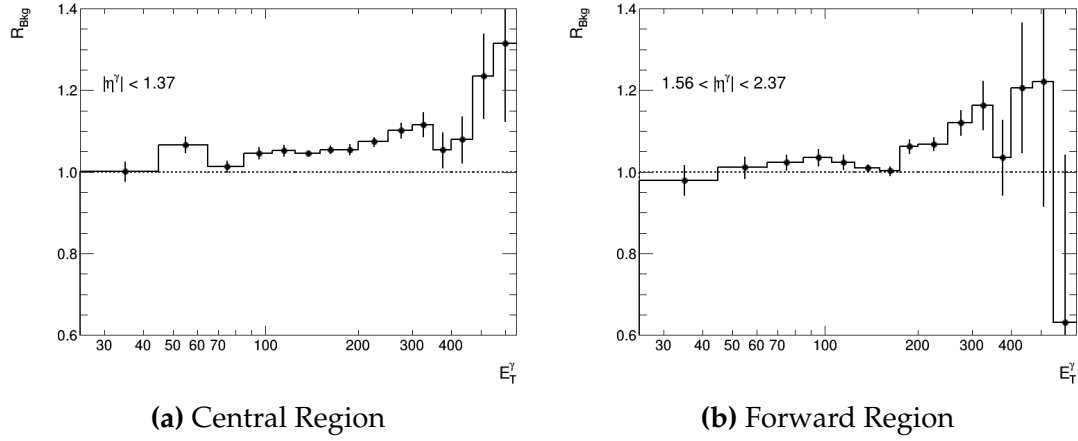


Figure 7.9 – Value of the R_{bkg} quantity constructed entirely in the non-isolated background region in data. Over the range considered by the measurement, the value of this quantity typically does not exceed 10%.

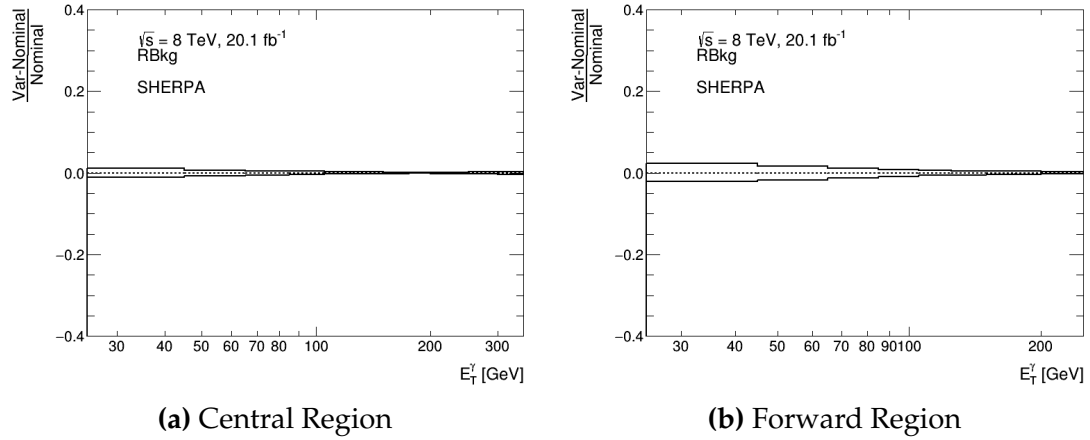


Figure 7.10 – Total variation arising in the photon purity when varying the R-background value up and down by 10%

are simulated directly in the matrix element through the inclusion of extra real diagrams with up to five total outgoing partons. The PYTHIA simulation matrix element is purely LO $2 \rightarrow 2$ production and is split into two components, a *hard* component that includes LO prompt photon diagrams, and a *brem* component that includes LO dijet diagrams and fragmentation functions to create a hard fragmentation photon from one of the outgoing partons. At LO the absolute normalization of a simulated cross section is not well defined, and as such neither are the relative fractions of these two components. A parameter is defined to vary the relative

fractions of these components, called α , where the relative fractions of the hard and brem components compared to their nominal values are then defined as α and $1 - \alpha$, respectively. Thus $\alpha = 0.5$ is equivalent to the nominal result, displayed in figure 7.11. Varying the fractions of these two components assesses the sensitivity of the measurement to the interplay of the hard and fragmentation components and in turn to the prompt photon modeling beyond LO. Due to the nature of the SHERPA calculation there is no analogous parameter, and for this reason the PYTHIA simulation is used. The size of the variation is evaluated taking the difference of the nominal result using the α parameter that yields the smallest χ^2 between the E_T^γ spectrum of the PYTHIA event yield and the data event yield, shown in figure 7.12. The optimized spectra are displayed in figure 7.13.

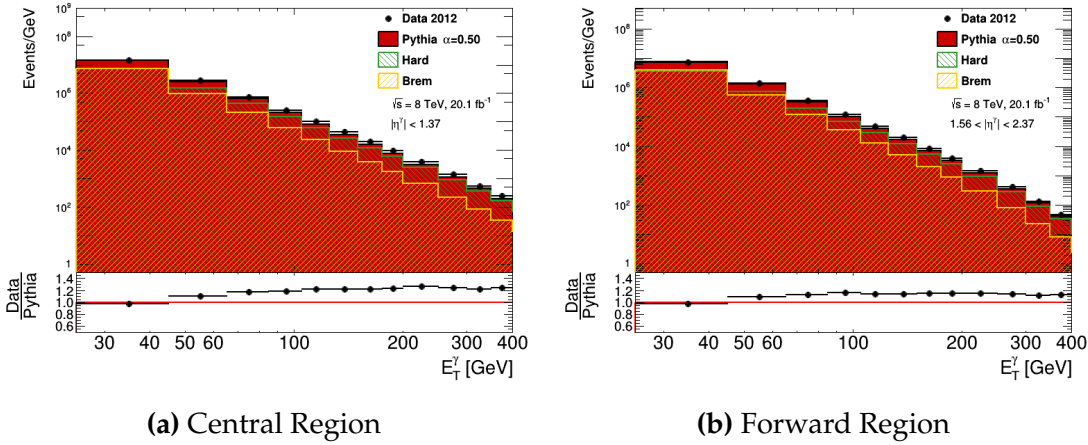


Figure 7.11 – The comparison of the nominal PYTHIA event yield to the data, using $\alpha = 0.5$, for the central and forward regions.

The tuned PYTHIA simulation is then used to derive the full measurement from the data, and the difference between the nominal measurement derived with the default PYTHIA hard/brem mixture is taken as the uncertainty. Figure 7.14 shows the resulting uncertainty.

7.2.4 QCD modelling dependence

The dependence on the QCD-cascade and hadronization model used to simulate the signal and derive the various correction factors for the data is estimated by taking the difference between the final result obtained using the SHERPA and

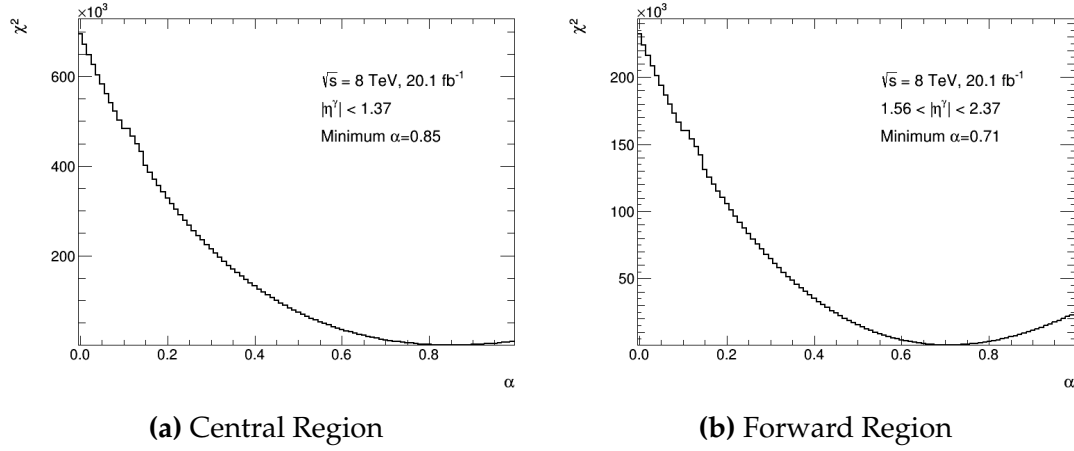


Figure 7.12 – The profile of the χ^2 comparison between PYTHIA and the data as a function of α in the central and forward regions.

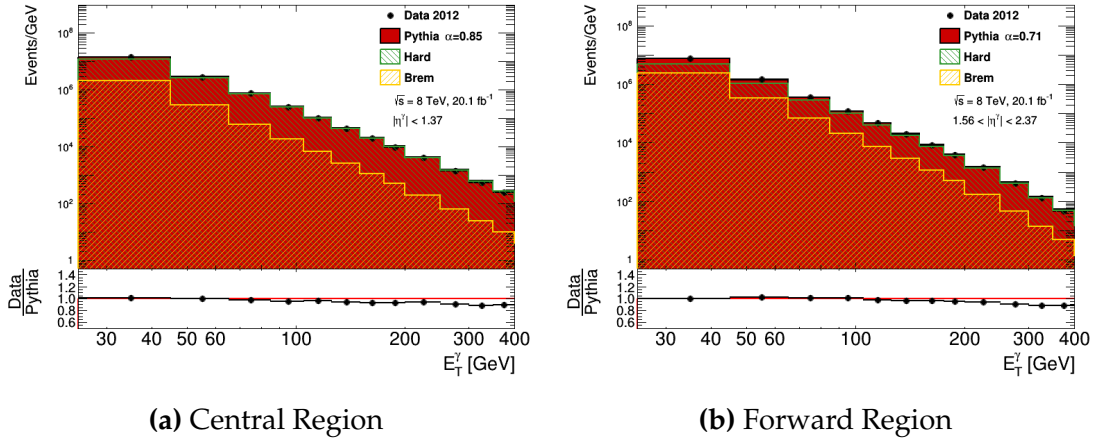


Figure 7.13 – The comparison of the optimized PYTHIA event yield to the data, using $\alpha = 0.85$ and $\alpha = 0.71$, for the central and forward regions, respectively.

PYTHIA simulations. The simulations use different matrix elements, different hadronization models, and the same parton-shower algorithm using different parameters. Figure 7.15 shows the resulting difference observed in the measured cross section having used the hard/brem optimized PYTHIA simulation, described in section 7.2.3.

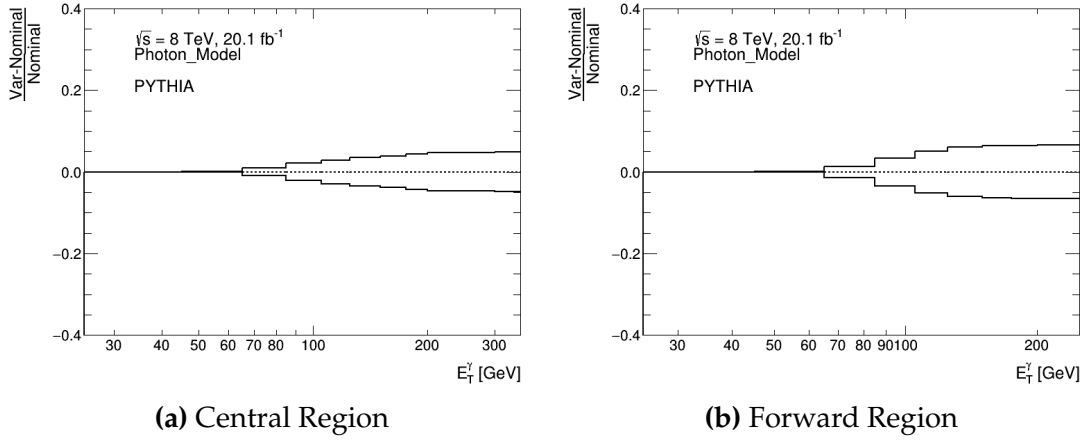


Figure 7.14 – The symmetrized comparison of the measured cross section using the hard/brem optimized PYTHIA to the nominal PYTHIA, for the central and forward regions, respectively.

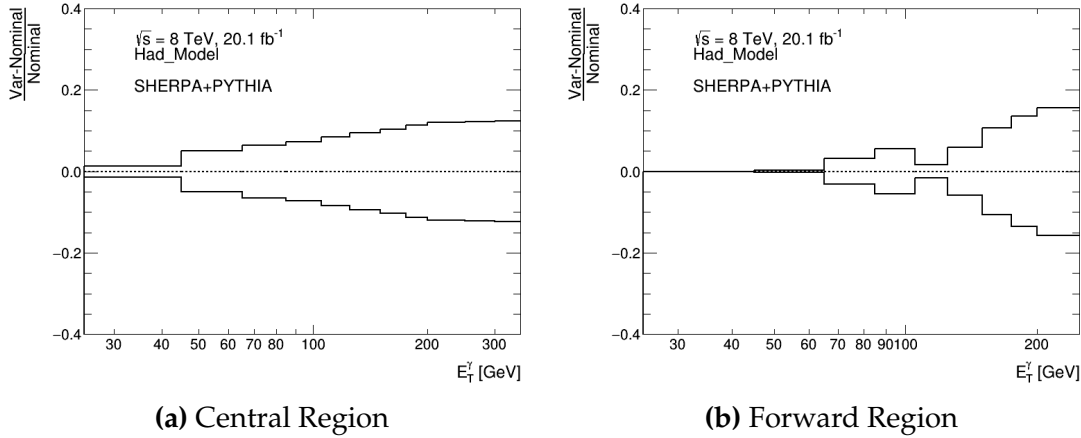


Figure 7.15 – The symmetrized comparison of the measured cross section using the hard/brem optimized PYTHIA to that measured by SHERPA, for the central and forward regions, respectively.

7.2.5 Photon Energy Scale

The calibration of photon candidate energies (photon energy scale, or γ ES) gives rise to numerous systematic uncertainties that capture the differences between photon response in the data and simulation. These uncertainties are assessed by independently varying 21 parameters, up and down, related to the various aspects of the calibration [133]. Examples of the components include parameters related to

the systematic uncertainties of the $Z \rightarrow ee$ calibration analysis, the EM calorimeter gain, the unconverted/converted photon identification uncertainty, the performance of the EM calorimeter pre-sampler, the GEANT4 simulation and EM shower development modeling, EM calorimeter energy pedestal calibration, and more. The combined effect of the γ ES variations on the measured result are displayed in figure 7.16.

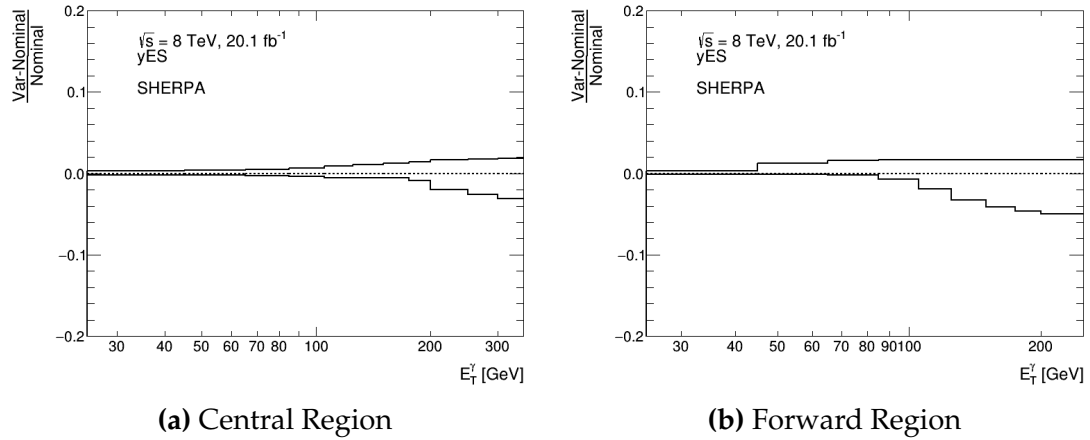


Figure 7.16 – Total systematic uncertainty arising in the cross section when adding all γ ES variations in quadrature.

7.2.6 Luminosity

The integrated luminosity uncertainty is $\pm 1.9\%$. It is fully correlated in all bins and added in quadrature to the other systematics [113].

7.2.7 Other Uncertainties

The remaining uncertainties, that produce on average a less than 1% impact on the measured result, are described in the following section. The cumulative effect of these individual components added in quadrature is shown in figure 7.17.

Photon-identification efficiency

The photon-identification efficiency factors derived by the simulation are corrected back to the data by applying scale factors that were derived for tight photons

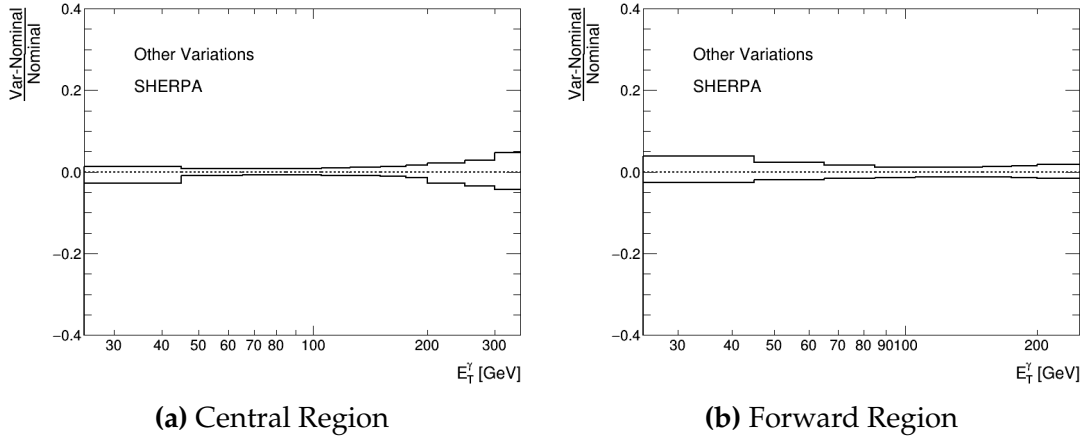


Figure 7.17 – Total remaining systematic variations added in quadrature.

with $E_T^{\text{iso}} < 4$ GeV. The uncertainty on these factors is propagated through the analysis with the additional contribution to account for the difference between the E_T^{iso} requirement in this analysis. This additional contribution was derived by taking the difference between the efficiency measured here as that measured with $E_T^{\text{iso}} < 4$ GeV. This variation does not produce a significant impact on the measured result.

Isolation Energy Correction

The uncertainty associated with the isolation energy corrections applied to the simulations was evaluated using the correction tool, as described in section 4.2.1. Two separate components are evaluated, one associated with the measurement of the isolation energy correction smearing (the width of the Gaussian correction factor) and the other arising from the offset. The former is derived by adjusting the correction by drawing it from a Gaussian centered on the nominal correction with a width equal to the nominal smearing correction. This variation does not produce a significant impact on the measured result.

Jet Energy Scale

The calibration of jet energies incorporates factors and corrections, derived from a multitude of studies, that are dependent on the energy and the position in η of the jet [139]. There are 67 separate nuisance parameters that can be adjusted to

evaluate the individual impact of the uncertainty on the analysis. The breakdown of these nuisance parameters is as follows:

- 50 nuisance parameters for the in-situ analyses calibrations (gamma+jet balance, multi-jet balance, Z+jet balance).
- 2 nuisance parameters for the eta intercalibration (modeling and statistics).
- 1 nuisance parameter for high p_T jet calibrations.
- 1 nuisance parameter for Simulation non-closure accounting for the fact that the calibration was performed using a more preliminary simulation tag than the final tag used in this analysis.
- 4 nuisance parameters for pile-up.
- 2 nuisance parameters for the flavour composition and response. The response to gluon and quark initiated jets is known to be different. The flavour composition of our samples using the final selection criteria was measured and used as input to help reduce this uncertainty.
- 1 nuisance parameter for punch-through jets.

These variations do not produce a significant impact on the measured result. This feature arises due to the fact that the measurement is performed as a function of E_T^γ .

Isolation Energy Gap

The sensitivity of the analysis to the isolation energy cut value used to define the non-isolated regions is assessed by varying the cut value up and down by 1 GeV. The choice of 1 GeV is derived from the photon isolation energy resolution. The impact of this variation is on the order of 3% in the first E_T bin, and zero in all remaining bins.

Jet Energy Resolution

For the 2012 data and simulations a good agreement was observed, thus no resolution correction is applied. The uncertainty on the jet resolution of the simulation is derived by smearing the simulated jet p_T distributions according to 1σ of the measured resolution and taking the difference between the nominal. This vari-

ation does not produce a statistically significant impact on the measured result.

7.3 Total Uncertainty

Figure 7.18 shows the sum in quadrature of all the systematic uncertainties on the measurement, in both the central and forward regions.

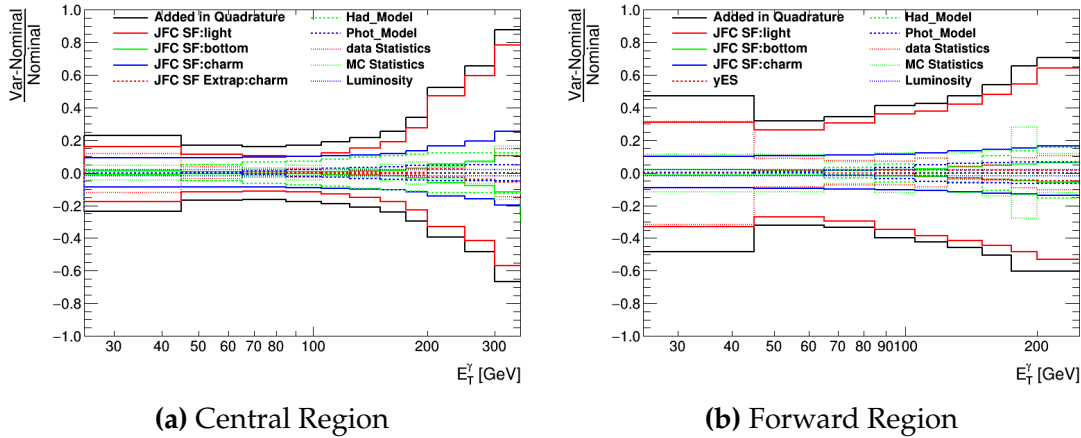


Figure 7.18 – Total uncertainty arising in the cross section, adding all contributions in quadrature. Only systematics which have on average $> 0.5\%$ uncertainty are displayed, though all contributions are included in the total.

Figure 7.19 shows the sum in quadrature of all the systematic uncertainties on the ratio of the measurements in the central and forward regions. There are two main effects to note comparing the uncertainties on the measured cross sections with the uncertainties on the measured ratio. First, is that there is an expected cancellation of systematic uncertainties that are correlated between the forward and central regions. Second, is that the statistical uncertainties are enhanced since the numerator and the denominator are made up of independent datasets. As such the first bin, which suffers from poor data statistics in the cross section measurements, sees a much enhanced data statistical uncertainty in the ratio. The larger statistical uncertainties also mean that variations of greater magnitude are required for them to be considered statistically significant.

With the full uncertainty on the cross section and ratio measurements, the results may now be compared to predictions and other measurements.

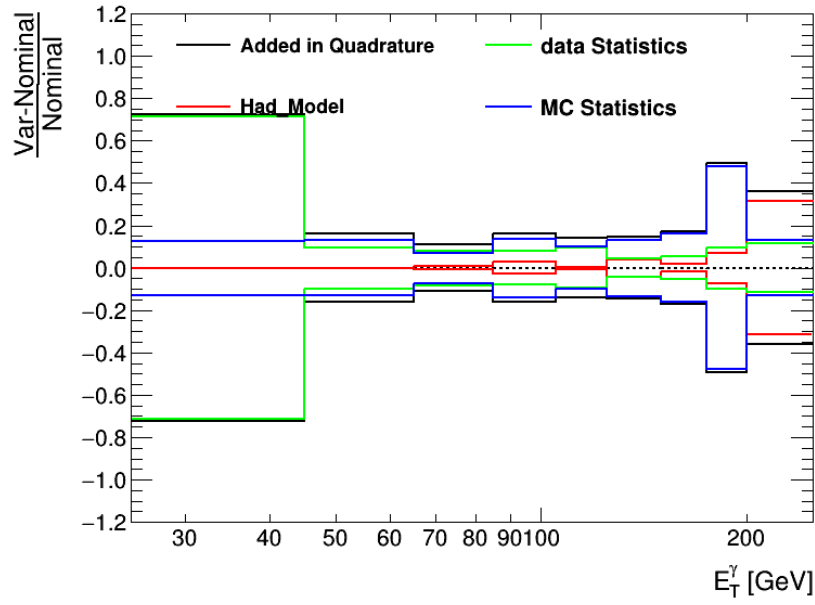


Figure 7.19 – Total uncertainty arising in the ratio of the central to forward cross sections, adding all contributions in quadrature. Only systematics which have on average $> 0.5\%$ uncertainty are displayed, though all contributions are included in the total.

Results

The most recent measurements of this type were performed in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV at the Tevatron by the D0 [74,75] and CDF [76] experiments. In addition to colliding different particles at a lower energy, the Tevatron provided a cleaner environment with less pile-up and larger bunch spacing. The angular acceptance and energy reach of the Tevatron results are comparable to those measured here, and the precision is better. The measurement procedures used by the two experiments at the Tevatron are similar to one another, and bear some similarities as well as some fundamental differences to the methodology used here. In the Tevatron measurements, the jet flavour is extracted using a two stage tagging procedure. First, an additional selection cut is applied on a tagging discriminant distribution that increases the heavy-flavour content of the data yield. The efficiency of this cut is measured in the simulation, calibrated based on a separate tagging analyses, and used to correct the data yield. Next, a template fit is performed to extract the flavour purity of the data yield. This template fit is performed on the invariant-mass distribution of charged particles originating from secondary-vertices associated to the selected jets. A similar approach was adopted in the early stages of this analysis but it was abandoned for the following reasons. The information used for both stages of the measurement are the same, it comes from the same sub-detectors in the same events. Applying a cut on a tagger discriminant is quantitatively the same as binning a tagger discriminant as every bin of the tagger discriminant is delimited by two cuts. It is well established that significant corrections are needed to rectify the efficiencies found in the simulations and in the data, prior to having applied any tagging-sensitive cuts to the phase space. If the template fit of the uncalibrated tagger distribution could indeed be trusted enough to extract the flavour fractions from the data yield, then there would be no need for the first step that applies the selection cut to enhance the purity of the sample in the first place. On top of this to rely on a template fit of an uncalibrated tagger after having applied a tagging cut, a cut that not only uses the same information but also shapes the acceptance of the phase space that is known to require

corrections, could not be justified. The approach adopted here relies on the calibration of all the bins of a single tagger. The uncertainties have been rigorously evaluated following the extensive studies used to derive the calibration.

The measured $\gamma + c$ cross-section in the central and forward regions, as well as the ratio of the two regions, is compared to the LO SHERPA and PYTHIA generator predictions, NLO MadGraph5_aMC@NLO predictions, and an analogous measurement technique using the MV1c tagger applied to the same dataset. The analysis technique used in this last comparison was used to derive the official 8 TeV $\gamma + b$ and $\gamma + c$ ATLAS results that have been accepted for publication in Physics Letters B [164]. The method presented here and the method used for the official ATLAS result were developed in tandem as part of the same experimental effort

While the comparisons shown here are qualitatively interesting, these results would make a quantitative impact if they were included in the global PDF fits used to derive the PDFs used at the LHC. In order to do so, the full set of uncertainties between the central and forward measurements would need to be made available such that the proper uncertainties would be propagated to the ratio when performing the fit (which considers both regions simultaneously). This is what has been done for the official 8 TeV ATLAS $\gamma + b$ and $\gamma + c$ measurement.

8.1 LO Theory Predictions

Leading-Order predictions, from SHERPA and PYTHIA, are compared to the measurements. Since there is no available generator uncertainty, and since more sophisticated higher-order predictions are also compared to the result, the feedback this comparison yields is less quantitative and more out of general interest. Figure 8.1 shows that SHERPA provides a closer prediction of the absolute cross-sections, which is expected since the SHERPA simulation included additional real higher order processes in the matrix element. Figure 8.2 shows that PYTHIA provides a closer prediction of the ratio. It is expected that the ratio provides enhanced sensitivity to the PDF, which implies that the CTEQ6L1 PDF used for the PYTHIA samples provides a better description of the HF content than the CT10 PDF used for the SHERPA simulation.

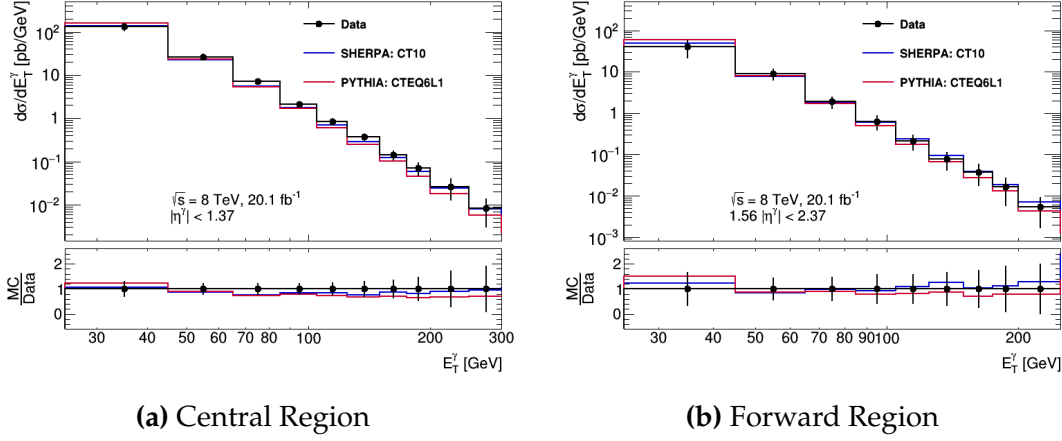


Figure 8.1 – The comparison of the measured cross-section to SHERPA and PYTHIA, in the central and forward regions.

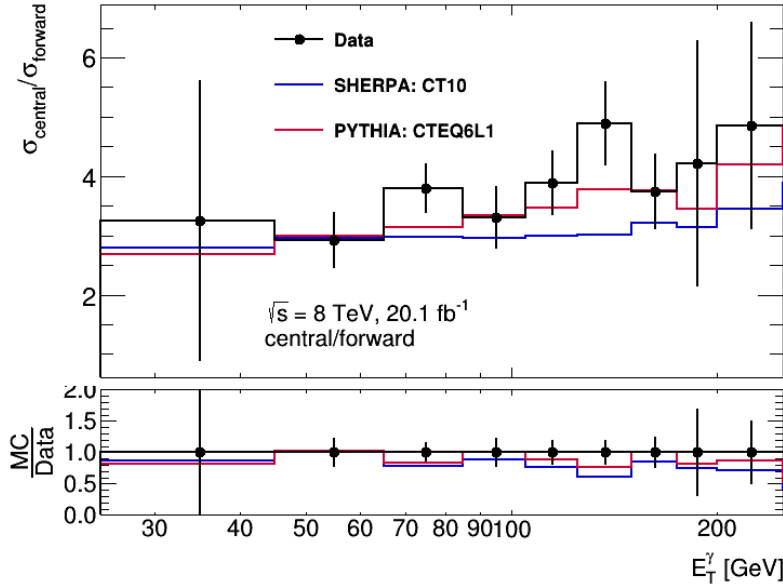


Figure 8.2 – The comparison of the measured cross-section ratio to SHERPA and PYTHIA predictions.

8.2 NLO Theory Predictions

Next to Leading Order (NLO) QCD predictions interfaced with a parton shower (PS) algorithm, generated at $O(\alpha_s^2)$ using the MadGraph5_aMC@NLO generator [80], are compared to the measurements. The computation considers

final states involving a photon plus a quark jet (excluding the top quark) or a gluon jet. Frixione isolation [165] is imposed on the photon, with parameters $\delta_0 = 0.4$, $n = 1$ and $\epsilon = 1$, to avoid collinear and infrared divergences. Jets are built using the anti- k_t algorithm with a radius parameter of 0.4 and a 10 GeV p_T^{jet} threshold. The parton shower, hadronization and underlying event are computed using PYTHIA 8. The renormalization and factorization scales are chosen to equal the transverse mass of the event, and the running of the strong coupling constant, α_s , is calculated at two loops and set to the Z boson mass with a value of $\alpha_s(M_Z) = 0.118$.

The computations are performed using various PDF sets to compare the results to different proton descriptions. NNPDF3.1 PDF set includes a nominal set with no non-perturbative intrinsic charm (IC) as well as a set with an IC contribution carrying 0.24% of the proton momentum [81, 166]. The IC PDF used in this set is derived from deep-inelastic c structure function data measured by the EMC experiment [167]. The 0.24% IC value is derived by performing the global fit of the available data while letting the IC component float freely. The CT14 PDF set takes a different approach that instead makes use of PDFs derived from IC models to test the IC hypothesis [57, 58]. Three CT14 sets are used here, the nominal set with no IC, and two which include IC following the BHPS and the SEA models [56] discussed in section 1.3.1. The CT14 BHPS PDFs, which use a valence-like non-perturbative parametrization of the IC content, contain 0.6% and 2% IC respectively. The CT14 SEA PDFs, which use a sea-quark-like nonperturbative parametrization of the IC content, contain 0.6% and 1.6% IC, respectively. These PDF sets are only available at NNLO, however, comparisons between the nominal CT14 set at NLO and the nominal CT14 set at NNLO showed no significant variation.

Systematic uncertainties on the theory predictions are computed to the 68% confidence interval. To assess the systematic uncertainty associated with the choice of factorization and renormalization scales, the scales are simultaneously and individually varied by factors of 0.5 and 2.0 [80]. The largest variation in each bin of E_T^γ resulting from the ensemble of variations is then taken as the uncertainty. The uncertainty associated with the determination of the PDF sets through their DGLAP evolution is evaluated using 100 simulation replicas for the NNPDF set and 56 eigenvectors for the CT14 set. The systematic uncertainty associated with

the strong coupling constant is assessed by changing the value of $\alpha_s(M_Z)$ to 0.116 and 0.120, coherently across both the matrix element computation and the PDF.

Figures 8.3 and 8.4 display the measurements compared to the CT14 nominal and BHPS PDFs. The predictions agree with the data, while the BHPS2 set with 2% IC shows an enhancement at large E_T^γ in the forward region and in the ratio. The data most closely matches the nominal PDF set, providing evidence against the BHPS IC hypothesis.

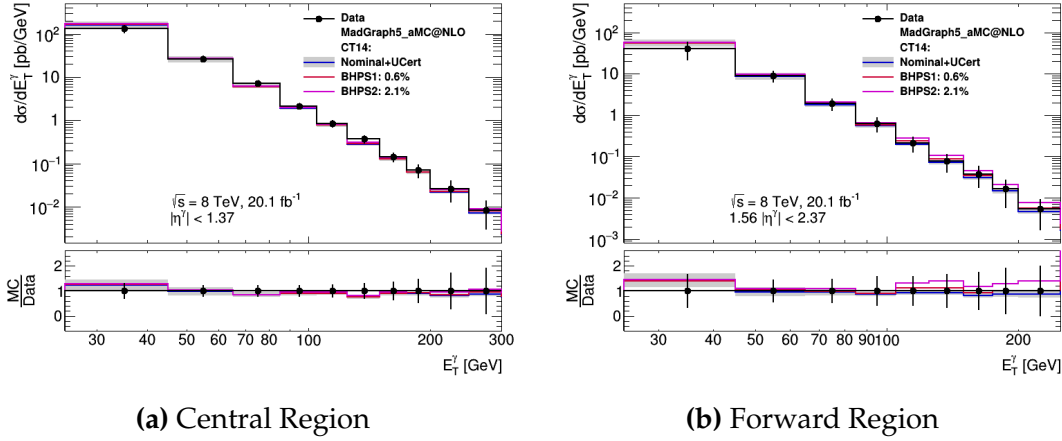


Figure 8.3 – The comparison of the measured cross-section to the CT14 PDF set using the nominal, BHPS1 and BHPS2 configurations, in the central and forward regions.

Figures 8.5 and 8.6 display the measurements compared to the CT14 nominal and SEA PDFs. The predictions agree with the data and contrary to the BHPS PDFs display no visible enhancement at high E_T^γ in the ratio.

Figures 8.7 and 8.8 display the measurements compared to the NNPDF3.1 nominal and IC PDFs. The predictions agree with the data and, similarly to the CT14 IC SEA PDFs, the IC set displays no visible enhancement.

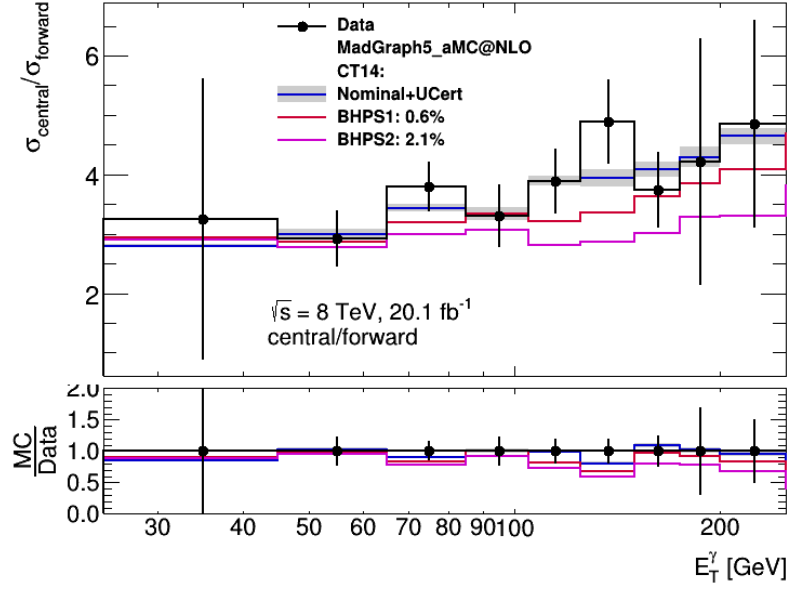
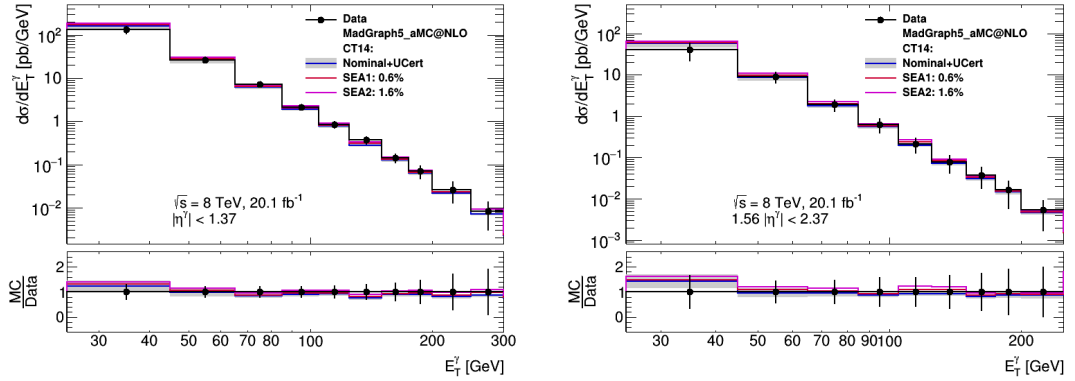


Figure 8.4 – The comparison of the measured cross-section ratio to the CT14 PDF set using the nominal, BHPS1 and BHPS2 configurations.



(a) Central Region

(b) Forward Region

Figure 8.5 – The comparison of the measured cross-section to the CT14 PDF set using the nominal, SEA1 and SEA2 configurations, in the central and forward regions.

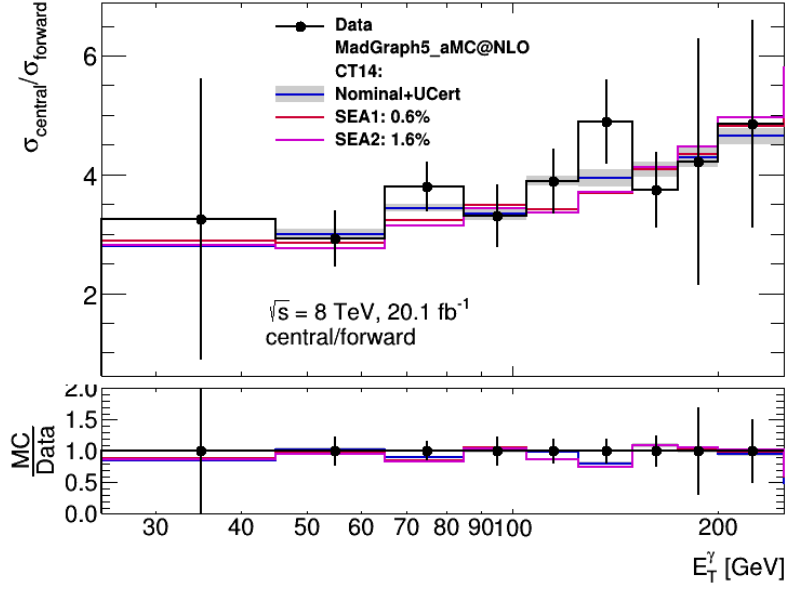
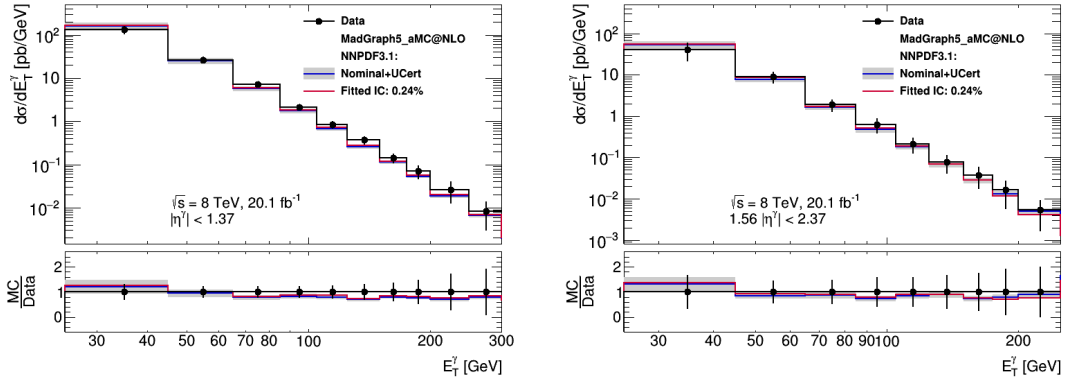


Figure 8.6 – The comparison of the measured cross-section ratio to the CT14 PDF set using the nominal, SEA1 and SEA2 configurations.



(a) Central Region

(b) Forward Region

Figure 8.7 – The comparison of the measured cross-section to the NNPDF PDF set using the nominal and 0.24% IC configurations, in the central and forward regions.

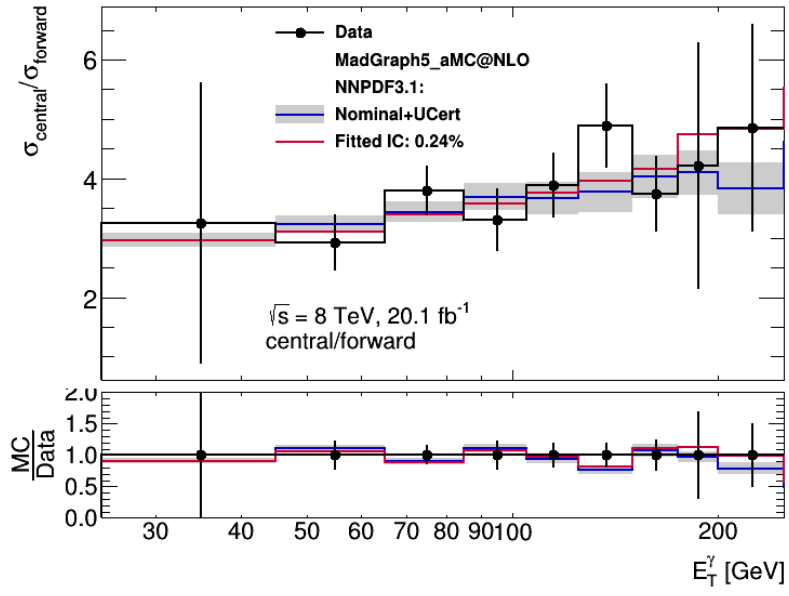


Figure 8.8 – The comparison of the measured cross-section ratio to the NNPDF PDF set using the nominal and 0.24% IC configurations.

8.3 MV1c Measurement

The official ATLAS $\gamma + b$ and $\gamma + c$ measurement performed on this same dataset was developed in parallel with this measurement. The author of this thesis was a primary contributor to this official measurement. The two measurements differ in that the official measurement makes use of the MV1c tagger (described briefly in section 6.3) instead of the JetFitterCharm tagger. The MV1c tagger is a b -tagger that is optimized to discriminate b -jets from c -jets. The calibration of the MV1c tagger for the 8 TeV 2012 dataset is much more thorough than that of the JetFitterCharm tagger. The calibration has four working-points, instead of two, and further a simultaneous calibration of the full discriminant, delimited by these working points, was carried out. This was the first calibration of this type to be performed in ATLAS, referred to as the “continuous calibration”, that takes into consideration explicitly the correlations between the different bins of the tagger discriminant. As a result, the tagging uncertainties, which are the dominant concern in the analysis, are smaller for the MV1c tagger for both the $\gamma + b$ and the $\gamma + c$ measurements. The uncertainties for the $\gamma + c$ measurement are comparable in the central region, but much larger in the forward region, for the JetFitterCharm tagger. At the outset of the project, though the MV1c calibration was known to be better, it was not certain if the JetFitterCharm tagger would perform better or not for the $\gamma + c$ measurement since it is designed specifically to identify c -jets instead of b -jets.

Figure 8.9 displays the ratio of the measured cross-sections for $\gamma + c$, considering only sources of uncertainty related to the two tagger calibrations and the statistics, since all other aspects of the measurements are shared. The two measurements are found to be consistent.

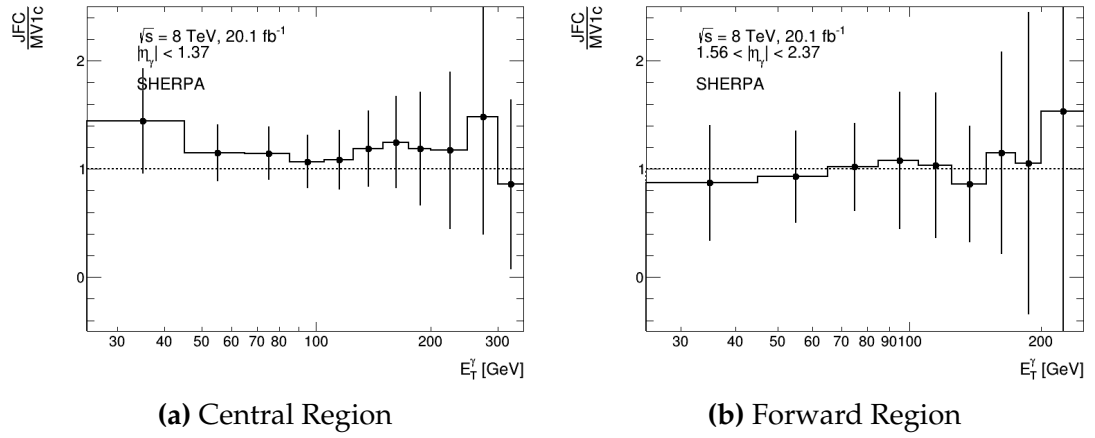


Figure 8.9 – The comparison of the measured cross-section using the JetFitterCharm tagger and the MV1c tagger, in the central and forward regions. The uncertainties include the uncertainties for both taggers and the statistics.

Conclusions and Outlook

9.1 Conclusions

The production cross section of prompt photons in association with a c -jet is measured using the 8 TeV p - p 2012 ATLAS dataset comprising 20.1 fb^{-1} . The JetFitterCharm tagging algorithm is used to extract the c -jet fraction from the event yield, in combination with techniques used in prior ATLAS photon analyses to extract the signal-photon fraction. The measurement is corrected for detector effects and inefficiencies using the bin-by-bin method, yielding a measurement that is directly comparable to other measurements and theoretical predictions at the particle-level. The measurement is performed as a function of E_T^γ in both the central and forward regions of the detector, as well as the ratio of the two measured regions. In taking the ratio the measurement and the simulations benefit from reduced systematic uncertainties, though the measurement also suffers from enhanced statistical uncertainties. The measurements are compared to LO predictions from SHERPA and PYTHIA, to NLO predictions from MadGraph5_aMC@NLO, and to an analogous measurement using the same dataset but a different technique. A selection of PDFs are used in the NLO predictions to test different proton content descriptions. The total uncertainty on the measurements is large, and as such they are consistent across all the comparisons. The goal of this measurement is not to confirm or to rule out any particular IC hypothesis, but to provide sensitive data. The primary intended use of data of this type is its inclusion in the global fits that are used to derive the PDFs that are used at the LHC. Through measurements like this, these fits rely less on extrapolation. This data can also be used by theorists to quantitatively test the agreement of their different model hypotheses in the manner that they choose. The analogous measurement, that was developed in parallel and performed on the same dataset, using the MV1c tagger is to be made public for these purposes.

9.2 Outlook

This measurement is the first of its type at the LHC, and as such can be improved upon in future iterations as the tools, calibrations and techniques used in ATLAS mature. The crux of this analysis and the largest limitation on its precision is the tagging of charm flavoured jets, and it is here where the greatest opportunities for improvement lie. A more complete calibration of the JetFitterCharm tagger, that is a continuous calibration with more working points, would greatly improve the precision of the measurement. The methodology for this type of calibration was carried out in Run 1 for the MV1 and the MV1c bottom flavour taggers, and will hopefully be translated to the charm flavour tagging domain for Run 2 and beyond. An additional potential advantage of calibrating the JetFitterCharm tagger in the same fashion as the MV1 and MV1c taggers is that then it could also be possible to measure the ratio of $\gamma + b$ and $\gamma + c$ using two different taggers, each optimized for their separate signals. If the calibration is the same between the two this means that their uncertainties can be correlated. Beyond the overall calibration methodology, improvements to the light-jet tagging calibration are a priority since they impose the most significant contribution to the overall uncertainty. For enhanced sensitivity to the heavy flavour components of the PDF, a $g \rightarrow b\bar{b}$ or $g \rightarrow c\bar{c}$ tagger would provide exciting prospects by introducing discrimination between Compton-like and gluon-splitting-like HF production.

Looking towards extending this analysis to subsequent LHC datasets brings forth the following considerations. The Run 2 dataset is large, more than twice the size of the Run 1 dataset, and the following runs are expected to be much larger still. The Run 2 dataset provides an existing opportunity to reduce the data statistical uncertainty. The center-of-mass energy of the Run 2 collisions is 13 TeV, increasing the phase space that can be probed by the measurement. The ATLAS detector underwent an ID upgrade with the inclusion of the Inner B-Layer (IBL) [168]. The IBL provides an additional layer of precision tracking to the ID, by virtue of a reduction of the beam-pipe diameter, providing improved HF-tagging capabilities. In addition to this improvement in the hardware, as the ATLAS analysis program matures the tagging calibration will also improve. Notable challenges imposed by the Run 2 conditions are the increased pile-up and that the

reach of the analyses at low E_T^γ will be reduced due to the increase in center-of-mass energy.

All things considered, the prospects for Run 2 measurements are very good. With the improved tagging performance pile-up mitigation provided by the IBL, a continuous calibration of the JetFitterCharm tagger, and considering the greatly increased statistics of the Run 2 dataset, a measurement of greater precision that has potential to further influence future PDFs and test the IC hypothesis should be pursued.

A

Cut Flows

Tables A.1 and A.2 display the reconstruction-level cut flow for events in the full η^γ range for the data and SHERPA. Table A.3 displays the cut flow for the particle-level selection in SHERPA. The total number of events, the relative fraction of events accepted by the cut in question, the total fraction of events accepted including all cuts up to and including the cut in question, and the cut name are listed.

Cut #	Events	Rel Accept	Tot Accept	Cut Name
Pre-Processing Skim Cut Flow				
1	729720140	100%	100%	EventExists
2	699783096	97.4%	97.4%	GRL
3	699782818	100%	97.4%	Core Flags & 0x40000
4	699781767	100%	97.4%	PV $\zeta = 2$ Tracks
5	698353782	99.8%	97.2%	LAr Error != 2
6	698353730	100%	97.2%	Tile Error != 2
7	57022566	8.17%	7.94%	Trigger (no match): EF_gAll_loose
Post-Processing Cut Flow				
1	57022566	100%	100%	Event Exists
2	57022566	100%	100%	GRL
4	57022566	100%	100%	LAr Error != 2
5	57022566	100%	100%	Tile Error != 2
6	57022566	100%	100%	Core Flags & 0x40000
7	57022544	100%	100%	PV ≥ 2 Tracks
8	57022542	100%	100%	Tile Trip Reader
9	30314757	53.2%	53.2%	PhID Preselect= loose'
10	16818097	55.5%	29.5%	$E_T^\gamma > 25$ [GeV]
11	16518514	98.2%	29.0%	$ \eta_{S2}^\gamma < 2.37$
12	16239775	98.3%	28.4%	$ \eta^\gamma < 2.37$
13	11713140	72.1%	20.5%	Trigger: EF_gAll_loose
14	11626568	99.3%	20.4%	Photon Object Quality + Clean
16	11567490	99.5%	20.3%	Photon Ambiguity Resolver
17	7489875	64.8%	13.1%	Photon ID = tight
18	4945136	66.0%	8.7%	Photon $E_T^{\text{iso}} < 4.8 \text{ GeV} + \times E_T^\gamma 0.0042$
19	4945136	100%	8.7%	At least one jet
20	4945135	100%	8.7%	Jet No Hot Tile
21	4945135	100%	8.7%	Jet Not BadLooseMinus
22	4945074	100%	8.7%	$ JVF > 0.5$ for Jet $p_T^{\text{jet}} < 50$ GeV
23	4944990	100%	8.7%	No Lead Photon All Jets Overlap, $\Delta R < 0.4$
24	4830027	97.7%	8.5%	$p_T^{\text{jet}} > 20$ [GeV]
25	4493408	93.0%	7.9%	$ \eta^{\text{jet}} < 2.5$
26	4493408	100%	7.9%	$ y^{\text{jet}} < 2.5$
27	4482873	99.8%	7.9%	No Lead Photon Lead Jet Overlap, $\Delta R < 1$

Table A.1 – Reconstruction level cut flow for the full η acceptance of the photon for data. "Rel Accept" refers to the percentage of events passing a given cut compared to the number of events passing the prior cut. "Tot Accept" refers to the percentage of initial events that have passed all the cuts up to and including the a given cut.

Cut #	Events	Rel Accept	Tot Accept	Cut Name
1	39083683	100%	100%	Event Exists
2	3774146	9.66%	9.66%	Sample stitching
3	3774146	100%	9.66%	LAr Error != 2
4	3774146	100%	9.66%	Tile Error != 2
5	3774146	100%	9.66%	Core Flags & 0x40000
6	3774146	100%	9.66%	$PV \geq 2$ Tracks
7	3774146	100%	9.66%	Tile Trip Reader
8	3100556	82.2%	7.93%	PhID Preselect= loose' + fudge
9	2742270	88.4%	7.02%	Photon Truth Matching
10	2725942	99.4%	6.97%	$E_T^\gamma > 25$ [GeV]
11	2725942	100%	6.97%	E_T^γ falls in correct or adjacent p_T^{Hat} slice
12	2683556	98.4%	6.87%	$\eta_{S2}^\gamma < 2.37$
13	2654712	98.9%	6.79%	$\eta^\gamma < 2.37$
14	2637543	99.4%	6.75%	Photon Object Quality
15	2637543	100%	6.75%	Photon Clean
16	2632734	99.8%	6.74%	Photon Ambiguity Resolver
17	2554969	97%	6.54%	Photon ID = tight
18	2307138	90.3%	5.9%	$E_T^{iso} < 4.8 \text{ GeV} + E_T^\gamma \times 0.0042$
19	2307138	100%	5.9%	At least one jet
20	2307138	100%	5.9%	Jet No Hot Tile
21	2307138	100%	5.9%	Jet Not BadLooseMinus
22	2307090	100%	5.9%	$ JVF > 0.5$ for Jet $p_T^{\text{jet}} < 50$ GeV
23	2307050	100%	5.9%	No Lead Photon All Jets Overlap, $\Delta R < 0.4$
24	2216600	96.1%	5.67%	$p_T^{\text{jet}} > 20$ [GeV]
25	2092545	94.4%	5.35%	$ \eta^{\text{jet}} < 2.5$
26	2092545	100%	5.35%	$ y^{\text{jet}} < 2.5$
27	2084213	99.6%	5.33%	No Lead Photon Lead Jet Overlap, $\Delta R < 1.0$

Table A.2 – Reconstruction level cut flow for the full η acceptance of the photon for SHERPA.

Cut #	Events	Rel Accept	Tot Accept	Cut Name
1	39083683	100%	100%	Event Exists
2	3774146	9.66%	9.66%	Sample Stitching
3	3774146	100%	9.66%	$E_T^\gamma > 25$ [GeV]
4	2982812	79%	7.63%	$ \eta^\gamma < 2.37$
5	2673911	89.6%	6.84%	$E_T^{\text{iso}} < 4.8 \text{ GeV} + E_T^\gamma \times 0.0042$
6	2664039	99.6%	6.82%	No Lead Photon All Jets Overlap, $\Delta R < 0.4$
7	2562437	96.2%	6.56%	$p_T^{\text{jet}} > 20$ [GeV]
8	2437081	95.1%	6.24%	$ y^{\text{jet}} < 2.5$
9	2430518	99.7%	6.22%	No Lead Photon Lead Jet Overlap, $\Delta R < 1.0$
10	335612	13.8%	0.859%	Jet truth tag: c
10	410560	16.8%	1.060%	Jet truth tag: b

Table A.3 – Particle level cut flow for the full η acceptance of the photon for SHERPA.

Bibliography

- [1] R. Keyes et al. Development and Characterisation of a Gas System and its Associated Slow-Control System for an ATLAS Small-Strip Thin Gap Chamber Testing Facility. *JINST*, 12(04):P04027, 2017.
- [2] ATLAS Collaboration. The performance of the jet trigger for the ATLAS detector during 2011 data taking. *Eur. Phys. J.*, C76(10):526, 2016.
- [3] S. L. Glashow. Partial Symmetries of Weak Interactions. *Nucl. Phys.*, 22:579–588, 1961.
- [4] Steven Weinberg. Precise relations between the spectra of vector and axial vector mesons. *Phys. Rev. Lett.*, 18:507–509, 1967.
- [5] Abdus Salam and John Clive Ward. Electromagnetic and weak interactions. *Phys. Lett.*, 13:168–171, 1964.
- [6] Mary K. Gaillard, Paul D. Grannis, and Frank J. Sciulli. The Standard Model of Particle Physics. *Rev. Mod. Phys.*, 71:S96–S111, 1999.
- [7] Brian C. Odom, D. Hanneke, B. D’Urso, and G. Gabrielse. New Measurement of the Electron Magnetic Moment Using a One-Electron Quantum Cyclotron. *Phys. Rev. Lett.*, 97:030801, 2006. [Erratum: *Phys. Rev. Lett.* 99, 039902 (2007)].
- [8] Peter W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys. Rev. Lett.*, 13:508–509, 1964.
- [9] Peter W. Higgs. Broken symmetries, massless particles and gauge fields. *Phys. Lett.*, 12:132–133, 1964.
- [10] Peter W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys. Rev. Lett.*, 13:508–509, 1964.

- [11] 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.*, B716:1–29, 2012.
- [12] CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett.*, B716:30–61, 2012.
- [13] P. A. M. Dirac. The quantum theory of the electron. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 117(778):610–624, 1928.
- [14] Carl D. Anderson. The positive electron. *Phys. Rev.*, 43:491–494, 1933.
- [15] K. Nakamura et al. Review of particle physics. *J. Phys.*, G37:075021, 2010.
- [16] Wikipedia. Standard model of particle physics. https://en.wikipedia.org/wiki/Standard_Model, April 2017.
- [17] Gerhart Luders. On the Equivalence of Invariance under Time Reversal and under Particle-Antiparticle Conjugation for Relativistic Field Theories. *Kong. Dan. Vid. Sel. Mat. Fys. Med.*, 28N5(5):1–17, 1954.
- [18] F. J. Dyson. The Radiation theories of Tomonaga, Schwinger, and Feynman. *Phys. Rev.*, 75:486–502, 1949.
- [19] Alan D. Martin. Proton structure, Partons, QCD, DGLAP and beyond. *Acta Phys. Polon.*, B39:2025–2062, 2008.
- [20] C. Patrignani et al. Review of Particle Physics. *Chin. Phys.*, C40(10):100001, 2016.
- [21] James Clerk Maxwell. A treatise on electricity and magnetism. *Oxford: Clarendon Pres*, 1873.
- [22] Samoil M. Bilenky and J. Hosek. Glashow-Weinberg-Salam Theory of Electroweak Interactions and the Neutral Currents. *Phys. Rept.*, 90:73–157, 1982.
- [23] Claude Amsler et al. Review of Particle Physics: Higgs bosons: theory and searches. *Phys. Lett.*, B667:414–438, 2008.
- [24] Steven Weinberg. A Model of Leptons. *Phys. Rev. Lett.*, 19:1264–1266, 1967.
- [25] Alexander M. Polyakov. Particle Spectrum in the Quantum Field Theory. *JETP Lett.*, 20:194–195, 1974. [*Pisma Zh. Eksp. Teor. Fiz.*20,430(1974)].

- [26] Y. Fukuda et al. Evidence for oscillation of atmospheric neutrinos. *Phys. Rev. Lett.*, 81:1562–1567, 1998.
- [27] The SNO Collaboration. Direct evidence for neutrino flavor transformation from neutral current interactions in the Sudbury Neutrino Observatory. *Phys. Rev. Lett.*, 89:011301, 2002.
- [28] David N. Spergel and Paul J. Steinhardt. Observational evidence for self-interacting cold dark matter. *Phys. Rev. Lett.*, 84:3760–3763, 2000.
- [29] H. Georgi and S. L. Glashow. Unity of All Elementary Particle Forces. *Phys. Rev. Lett.*, 32:438–441, 1974.
- [30] W. de Boer. Grand unified theories and supersymmetry in particle physics and cosmology. *Prog. Part. Nucl. Phys.*, 33:201–302, 1994.
- [31] Edward Witten. Dynamical Breaking of Supersymmetry. *Nucl. Phys.*, B188:513, 1981.
- [32] B Povh, K Rith, C Scholz, and F Zetsche. *Particles and Nuclei: an introduction to the physical concepts; 6th ed.* Springer, Berlin, 2008.
- [33] F. D. Aaron et al. A Precision Measurement of the Inclusive ep Scattering Cross Section at HERA. *Eur. Phys. J.*, C64:561–587, 2009.
- [34] H. Abramowicz et al. Deep inelastic cross-section measurements at large y with the ZEUS detector at HERA. *Phys. Rev.*, D90(7):072002, 2014.
- [35] H. Abramowicz et al. Measurement of high- Q^2 charged current deep inelastic scattering cross sections with a longitudinally polarised positron beam at HERA. *Eur. Phys. J.*, C70:945–963, 2010.
- [36] S. Chekanov et al. Measurement of high- Q^2 deep inelastic scattering cross sections with a longitudinally polarised positron beam at HERA. *Phys. Lett.*, B637:210–222, 2006.
- [37] F. D. Aaron et al. Measurement of the Inclusive ep Scattering Cross Section at Low Q^2 and x at HERA. *Eur. Phys. J.*, C63:625–678, 2009.
- [38] F. D. Aaron et al. Measurement of the Charm and Beauty Structure Functions using the H1 Vertex Detector at HERA. *Eur. Phys. J.*, C65:89–109, 2010.
- [39] A. Aktas et al. Measurement of the proton structure function $F(2)$ at low Q^2 in QED Compton scattering at HERA. *Phys. Lett.*, B598:159–171, 2004.

- [40] S. Chekanov et al. Measurement of charm and beauty production in deep inelastic ep scattering from decays into muons at HERA. *Eur. Phys. J.*, C65:65–79, 2010.
- [41] R. S. Thorne and R. G. Roberts. An Ordered analysis of heavy flavor production in deep inelastic scattering. *Phys. Rev.*, D57:6871–6898, 1998.
- [42] V. N. Gribov and L. N. Lipatov. Deep inelastic e p scattering in perturbation theory. *Sov. J. Nucl. Phys.*, 15:438–450, 1972. [*Yad. Fiz.*15,781(1972)].
- [43] Guido Altarelli and G. Parisi. Asymptotic Freedom in Parton Language. *Nucl. Phys.*, B126:298–318, 1977.
- [44] G. Curci, W. Furmanski, and R. Petronzio. Evolution of Parton Densities Beyond Leading Order: The Nonsinglet Case. *Nucl. Phys.*, B175:27–92, 1980.
- [45] John C. Collins, Davison E. Soper, and George F. Sterman. Factorization of Hard Processes in QCD. *Adv. Ser. Direct. High Energy Phys.*, 5:1–91, 1989.
- [46] Fabio Maltoni, Thomas McElmurry, Robert Putman, and Scott Willenbrock. Choosing the Factorization Scale in Perturbative QCD. 2007.
- [47] James Stirling. Standard model cross sections as a function of collider energy (2012 version). <http://www.hep.ph.ic.ac.uk/~wstirlin/plots/plots.html>, April 2017.
- [48] A. Denner, S. Heinemeyer, I. Puljak, D. Rebuszi, and M. Spira. Standard Model Higgs-Boson Branching Ratios with Uncertainties. *Eur. Phys. J.*, C71:1753, 2011.
- [49] Claude Amsler et al. Review of Particle Physics: The CKM quark-mixing matrix. *Phys. Lett.*, B667:145–152, 2008.
- [50] Howard E. Haber and Gordon L. Kane. The Search for Supersymmetry: Probing Physics Beyond the Standard Model. *Phys. Rept.*, 117:75–263, 1985.
- [51] J. L. Diaz-Cruz, Hong-Jian He, and C. P. Yuan. Soft SUSY breaking, stop scharm mixing and Higgs signatures. *Phys. Lett.*, B530:179–187, 2002.
- [52] ATLAS Collaboration. Search for scalar charm quark pair production in pp collisions at $\sqrt{s} = 8$ TeV with the atlas detector. *Phys. Rev. Lett.*, 114:161801, 2015.

- [53] E. G. de Oliveira, A. D. Martin, M. G. Ryskin, and A. G. Shuvaev. Treatment of heavy quarks in QCD. *Eur. Phys. J.*, C73(10):2616, 2013.
- [54] V. A. Bednyakov, M. A. Demichev, G. I. Lykasov, T. Stavreva, and M. Stockton. Searching for intrinsic charm in the proton at the LHC. *Phys. Lett.*, B728:602–606, 2014.
- [55] S. J. Brodsky, V. A. Bednyakov, G. I. Lykasov, J. Smiesko, and S. Tokar. The Physics of Heavy Quark Distributions in Hadrons: Collider Tests. *Prog. Part. Nucl. Phys.*, 93:108, 2017.
- [56] S. J. Brodsky, P. Hoyer, C. Peterson, and N. Sakai. The Intrinsic Charm of the Proton. *Phys. Lett.*, 93B:451–455, 1980.
- [57] Sayipjamal Dulat, Tie-Jiun Hou, Jun Gao, Marco Guzzi, Joey Huston, Pavel Nadolsky, Jon Pumplin, Carl Schmidt, Daniel Stump, and C. P. Yuan. New parton distribution functions from a global analysis of quantum chromodynamics. *Phys. Rev.*, D93(3):033006, 2016.
- [58] Sayipjamal Dulat, Tie-Jiun Hou, Jun Gao, Joey Huston, Jon Pumplin, Carl Schmidt, Daniel Stump, and C. P. Yuan. Intrinsic Charm Parton Distribution Functions from CTEQ-TEA Global Analysis. *Phys. Rev.*, D89(7):073004, 2014.
- [59] Durham IPPP. The durham hepdata project. <http://hepdata.cedar.ac.uk/pdf/pdf3.html>, April 2017.
- [60] Patrick Aurenche, Michel Fontannaz, Jean-Philippe Guillet, Eric Pilon, and Monique Werlen. A New critical study of photon production in hadronic collisions. *Phys. Rev.*, D73:094007, 2006.
- [61] 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 fb^{-1} . *Phys. Rev.*, D89(5):052004, 2014.
- [62] ATLAS Collaboration. Measurement of the inclusive isolated prompt photon cross section in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. *JHEP*, 08:005, 2016.
- [63] 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.*, B770:473–493, 2017.

- [64] ATLAS Collaboration. Dynamics of isolated-photon plus jet production in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *Nucl. Phys.*, B875:483–535, 2013.
- [65] ATLAS Collaboration. High- E_T isolated-photon plus jets production in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. *Nucl. Phys.*, B918:257–316, 2017.
- [66] ATLAS Collaboration. Measurement of the isolated di-photon cross-section in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *Phys. Rev.*, D85:012003, 2012.
- [67] ATLAS Collaboration. Measurement of isolated-photon pair production in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *JHEP*, 01:086, 2013.
- [68] ATLAS Collaboration. Measurements of integrated and differential cross sections for isolated photon pair production in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. *Phys. Rev.*, D95(11):112005, 2017.
- [69] CMS Collaboration. Comparison of the $Z/\gamma + \text{jets}$ to $\gamma + \text{jets}$ cross sections in pp collisions at $\sqrt{s} = 8$ TeV. *JHEP*, 10:128, 2015. [Erratum: *JHEP*04,010(2016)].
- [70] CMS Collaboration. Measurement of differential cross sections for the production of a pair of isolated photons in pp collisions at $\sqrt{s} = 7$ TeV. *EPJC*, 74(11):3129, 2014.
- [71] CMS Collaboration. Measurement of the triple-differential cross section for photon+jets production in proton-proton collisions at $\sqrt{s}=7$ TeV. *JHEP*, 06:009, 2014.
- [72] CMS Collaboration. Rapidity distributions in exclusive $Z + \text{jet}$ and $\gamma + \text{jet}$ events in pp collisions at $\sqrt{s} = 7$ TeV. *Phys. Rev.*, D88(11):112009, 2013.
- [73] CMS Collaboration. Measurement of the Differential Cross Section for Isolated Prompt Photon Production in pp Collisions at 7 TeV. *Phys. Rev.*, D84:052011, 2011.
- [74] D0 Collaboration. Measurement of the photon+ b -jet production differential cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. *Phys. Lett.*, B714:32–39, 2012.

- [75] D0 Collaboration. Measurement of the differential photon + c -jet cross section and the ratio of differential photon+ c and photon+ b cross sections in proton-antiproton collisions at $\sqrt{s} = 1.96$ TeV. *Phys. Lett.*, B719:354–361, 2013.
- [76] CDF Collaboration. Measurement of the cross section for direct-photon production in association with a heavy quark in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. *Phys. Rev. Lett.*, 111(4):042003, 2013.
- [77] Raphaelle Ichou and David d’Enterria. Sensitivity of isolated photon production at TeV hadron colliders to the gluon distribution in the proton. *Phys. Rev.*, D82:014015, 2010.
- [78] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The Anti- $k(t)$ jet clustering algorithm. *JHEP*, 04:063, 2008.
- [79] ATLAS Collaboration. Performance and Calibration of the JetFitterCharm Algorithm for c -Jet Identification. *ATL-PHYS-PUB-2015-001*, January 2015.
- [80] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 07:079, 2014.
- [81] Richard D. Ball et al. Parton distributions for the LHC Run II. *JHEP*, 04:040, 2015.
- [82] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. A Brief Introduction to PYTHIA 8.1. *Comput. Phys. Commun.*, 178:852–867, 2008.
- [83] F. J. Hasert et al. Observation of Neutrino Like Interactions without Muon or Electron in the Gargamelle Neutrino Experiment. *Nucl. Phys.*, B73:1–22, 1974.
- [84] G. Arnison et al. Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $\sqrt{s} = 540$ -GeV. *Phys. Lett.*, B122:103–116, 1983. [,611(1983)].
- [85] M. Banner et al. Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti- p p Collider. *Phys. Lett.*, B122:476–485, 1983.

- [86] S. Schael et al. Precision electroweak measurements on the Z resonance. *Phys. Rept.*, 427:257–454, 2006.
- [87] C. Amole et al. An experimental limit on the charge of antihydrogen. *Nature Commun.*, 5:3955, 2014.
- [88] ATLAS Collaboration. Observation of a Centrality-Dependent Dijet Asymmetry in Lead-Lead Collisions at $\sqrt{s_{NN}} = 2.77$ TeV with the ATLAS Detector at the LHC. *Phys. Rev. Lett.*, 105:252303, 2010.
- [89] Roel Aaij et al. Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \rightarrow J/\psi K^- p$ Decays. *Phys. Rev. Lett.*, 115:072001, 2015.
- [90] E. Wilson. Transverse Beam Dynamics. In *2nd CERN Accelerator School: General Accelerator Physics Gif-sur-Yvette, France, September 3-14, 1984*, pages 64–95, 1986.
- [91] D J Warner. Project study for a new 50 MeV linear accelerator for the C. P. S. Technical Report CERN-MPS-LINP-73-1, CERN, Geneva, Oct 1973.
- [92] CERN. *LHC the guide: CERN-Brochure-2017-002-Eng*. Communication Group, Geneva, 2017.
- [93] O. Grobner. The LHC Vacuum System. *Conf. Proc.*, C970512:3542, 1997.
- [94] The ATLAS Collaboration. Charged-particle distributions in pp interactions at $\sqrt{s} = 8$ TeV measured with the ATLAS detector. *Eur. Phys. J.*, C76(7):403, 2016.
- [95] John C. Collins. Light cone variables, rapidity and all that. *arXiv*, hep-ph/9705393, 1997.
- [96] ATLAS Collaboration. Reconstruction of primary vertices at the ATLAS experiment in Run 1 protonproton collisions at the LHC. *Eur. Phys. J.*, C77(5):332, 2017.
- [97] 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.*, C76(11):581, 2016.
- [98] ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008.

- [99] ATLAS Collaboration. ATLAS pixel detector electronics and sensors. *JINST*, 3:P07007, 2008.
- [100] Andre Lukas Schorlemmer. *Monitoring Radiation Damage in the ATLAS Pixel Detector*. PhD thesis, Gottingen U., II. Phys. Inst., 2014-06-11.
- [101] John Neil Jackson. The ATLAS semiconductor tracker (SCT). *Nucl. Instrum. Meth.*, A541:89–95, 2005. [,89(2005)].
- [102] ATLAS Collaboration. The ATLAS TRT barrel detector. *JINST*, 3:P02014, 2008.
- [103] ATLAS Collaboration. The ATLAS TRT electronics. *JINST*, 3:P06007, 2008.
- [104] C. W. Fabjan and F. Gianotti. Calorimetry for particle physics. *Rev. Mod. Phys.*, 75:1243–1286, 2003.
- [105] W. E. Cleland and E. G. Stern. Signal processing considerations for liquid ionization calorimeters in a high rate environment. *Nucl. Instrum. Meth.*, A338(2-3):467–497, 1994.
- [106] Stefaan Tavernier. *Interactions of Particles in Matter*, pages 23–53. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.
- [107] M. L. Andrieux et al. Construction and test of the first two sectors of the ATLAS barrel liquid argon presampler. *Nucl. Instrum. Meth.*, A479:316–333, 2002.
- [108] S. Viret. LASER monitoring system for the ATLAS tile calorimeter. *Nucl. Instrum. Meth.*, A617:120–122, 2010.
- [109] A. Artamonov et al. The ATLAS forward calorimeters. *JINST*, 3:P02010, 2008.
- [110] S. van der Meer. Calibration of the Effective Beam Height in the ISR. 1968.
- [111] Jacob Groth-Jensen. *Luminosity determination and simulation of the LUCID detector at the ATLAS experiment*. PhD thesis, Lund U., 2010.
- [112] V. Cindro et al. The ATLAS beam conditions monitor. *JINST*, 3:P02004, 2008.
- [113] ATLAS Collaboration. Improved luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC. *Eur. Phys. J.*, C73(8):2518, 2013.

- [114] *ATLAS level-1 trigger: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1998.
- [115] ATLAS Collaboration. The ATLAS Simulation Infrastructure. *Eur. Phys. J.*, C70:823–874, 2010.
- [116] Matt Dobbs and Jorgen Beck Hansen. The HepMC C++ Monte Carlo event record for High Energy Physics. *Comput. Phys. Commun.*, 134:41–46, 2001.
- [117] S. Agostinelli et al. GEANT4: A Simulation toolkit. *Nucl. Instrum. Meth.*, A506:250–303, 2003.
- [118] D. Costanzo, A. Dell’Acqua, A Di Simone, M. Gallas, A. Nairz, A. Rimoldi, J. Boudreau, and V. TSulaia. ATLAS detector simulation. 2005.
- [119] G. Barrand et al. GAUDI - A software architecture and framework for building HEP data processing applications. *Comput. Phys. Commun.*, 140:45–55, 2001.
- [120] Leif Lonnblad. CLHEP: A project for designing a C++ class library for high-energy physics. *Comput. Phys. Commun.*, 84:307–316, 1994.
- [121] Stefan Hoche. Introduction to parton-shower event generators. In *Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders (TASI 2014) Boulder, Colorado, June 2-27, 2014*, 2014.
- [122] T. Gleisberg, Stefan. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert, and J. Winter. Event generation with SHERPA 1.1. *JHEP*, 02:007, 2009.
- [123] M. Bahr et al. Herwig++ Physics and Manual. *Eur. Phys. J.*, C58:639–707, 2008.
- [124] ATLAS Collaboration. ATLAS tunes of PYTHIA 6 and Pythia 8 for MC11. Technical Report ATL-PHYS-PUB-2011-009, CERN, Geneva, Jul 2011.
- [125] P. Nason and P.Z. Skands. 40. Monte Carlo Event Generators. *Chin. Phys. (Particle Data Group)*, 40,100001:1–25, 2016.
- [126] Bo Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand. Parton Fragmentation and String Dynamics. *Phys. Rept.*, 97:31–145, 1983.

- [127] Jan-Christopher Winter, Frank Krauss, and Gerhard Soff. A Modified cluster hadronization model. *Eur. Phys. J.*, C36:381–395, 2004.
- [128] B. R. Webber. Hadronization. In *Proceedings: Summer School on Hadronic Aspects of Collider Physics, Zuoz, Switzerland, Aug 23-31, 1994*, pages 49–77, 1994.
- [129] Torbjorn Sjostrand. Old ideas in hadronization: The lund string. <http://conference.ippp.dur.ac.uk/event/265/>, April 2009.
- [130] Zachary Marshall. Simulation of Pile-up in the ATLAS Experiment. *J. Phys. Conf. Ser.*, 513:022024, 2014.
- [131] ATLAS Collaboration. Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run-1 data. *Eur. Phys. J.*, C76(12):666, 2016.
- [132] ATLAS Collaboration. Expected photon performance in the ATLAS experiment. Technical Report ATL-PHYS-PUB-2011-007, ATL-COM-PHYS-2010-1051, 2011.
- [133] ATLAS Collaboration. Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data. *Eur. Phys. J.*, C74(10):3071, 2014.
- [134] ATLAS Collaboration. Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data. *Eur. Phys. J.*, C72:1909, 2012.
- [135] S Laplace and JB de Vivie. Calorimeter isolation and pile-up. Technical Report ATL-COM-PHYS-2012-467, CERN, Geneva, May 2012.
- [136] Matteo Cacciari and Gavin P. Salam. Pileup subtraction using jet areas. *Phys. Lett.*, B659:119–126, 2008.
- [137] M Hance, D Olivito, and H Williams. Performance Studies for e/gamma Calorimeter Isolation. Technical Report ATL-COM-PHYS-2011-1186, CERN, Geneva, Sep 2011.
- [138] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The Catchment Area of Jets. *JHEP*, 04:005, 2008.
- [139] Georges Aad et al. Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV. *Eur. Phys. J.*, C73(3):2304, 2013.

- [140] Gregory Soyez. The SISCone and anti-k(t) jet algorithms. In *Proceedings, 16th International Workshop on Deep Inelastic Scattering and Related Subjects (DIS 2008): London, UK, April 7-11, 2008*, page 178, 2008.
- [141] T. Barillari et al. Local hadronic calibration. Technical Report ATL-LARG-PUB-2009-001-2, ATL-COM-LARG-2008-006, 2009.
- [142] ATLAS Collaboration. Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV. *Eur. Phys. J.*, C73(3):2304, 2013.
- [143] ATLAS Collaboration. Performance of b -Jet Identification in the ATLAS Experiment. *JINST*, 11(04):P04008, 2016.
- [144] ATLAS Collaboration. Performance of the ATLAS Secondary Vertex b -tagging Algorithm in 7 TeV Collision Data. Technical Report ATLAS-CONF-2010-042, CERN, Geneva, Jul 2010.
- [145] ATLAS Collaboration. Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics. Technical Report SLAC-R-980, CERN-OPEN-2008-020, 2009.
- [146] ATLAS Collaboration. Commissioning of the ATLAS high-performance b -tagging algorithms in the 7 TeV collision data. Technical Report ATLAS-CONF-2011-102, 2011.
- [147] ATLAS Collaboration. Calibration of the performance of b -tagging for c and light-flavour jets in the 2012 ATLAS data. Technical Report ATLAS-CONF-2014-046, CERN, Geneva, Jul 2014.
- [148] ATLAS Collaboration. Calibration of b -tagging using dileptonic top pair events in a combinatorial likelihood approach with the ATLAS experiment. Technical Report ATLAS-CONF-2014-004, CERN, Geneva, Feb 2014.
- [149] ATLAS Collaboration. Summary of ATLAS Pythia 8 tunes. Technical Report ATL-PHYS-PUB-2012-003, 2012.
- [150] Winter, J.-C., Krauss, F., and Soff, G. A modified cluster-hadronisation model. *Eur. Phys. J. C*, 36(3):381–395, 2004.
- [151] S. Agostinelli et al. GEANT4 - a simulation toolkit. *Nucl. Instrum. Methods Phys. A*, 506:250, 2003.

- [152] Bruno Lenzi, Christophe Royon, and Matthias Saimpert. Data-driven corrections for the calorimeter isolation energy of photons computed from topological clusters. Technical Report ATL-COM-PHYS-2015-258, CERN, Geneva, Apr 2015.
- [153] Q Buat, J Leveque, N Lorenzo Martinez, F Polci, and B Trocme. Definition of the photon cleaning procedure for 2011 data. Technical Report ATL-COM-PHYS-2012-134, CERN, Geneva, Feb 2012.
- [154] ATLAS Collaboration. Data-Quality Requirements and Event Cleaning for Jets and Missing Transverse Energy Reconstruction with the ATLAS Detector in Proton-Proton Collisions at a Center-of-Mass Energy of $\sqrt{s} = 7$ TeV, Jul 2010.
- [155] G Pasztor, D Damazio, and F Monticelli. Electron and photon trigger performance plots using 2012 data. (ATL-COM-DAQ-2014-058), Jun 2014.
- [156] M Stoebe. Measurement of the inclusive isolated prompt photon cross section in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector, Mar 2015. PhD thesis, McGill University.
- [157] ATLAS Collaboration. Measurement of differential production cross-sections for a Z boson in association with b -jets in 7 TeV proton-proton collisions with the ATLAS detector. *JHEP*, 10:141, 2014.
- [158] ATLAS Collaboration. Measurements of fiducial cross-sections for $t\bar{t}$ production with one or two additional b -jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector. *Eur. Phys. J.*, C76(1):11, 2016.
- [159] Fabrizio Parodi, Giacinto Piacquadio, Timothy Paul Scanlon, Lei Zhang, and Frank Filthaut. Continuous b -tagging calibration for the ATLAS experiment. Technical Report ATL-COM-PHYS-2015-429, CERN, Geneva, May 2015.
- [160] Wouter Verkerke and David P. Kirkby. The RooFit toolkit for data modeling. *eConf*, C0303241:MOLT007, 2003. [,186(2003)].
- [161] F. James and M. Roos. Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations. *Comput. Phys. Commun.*, 10:343–367, 1975.
- [162] G. D’Agostini. A Multidimensional unfolding method based on Bayes’ theorem. *Nucl. Instrum. Meth.*, A362:487–498, 1995.

- [163] G. D'Agostini. Improved iterative Bayesian unfolding. *arXiv*, 1010.0632, 2010.
- [164] ATLAS Collaboration. Measurement of differential cross sections of isolated-photon plus heavy-flavour jet production in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector. 2017. Accepted by Phys. Lett. B, eprint: 1710.09560.
- [165] Stefano Frixione. Isolated photons in perturbative QCD. *Phys. Lett.*, B429:369–374, 1998.
- [166] Richard D. Ball et al. Parton distributions from high-precision collider data. *Eur. Phys. J.*, C77(10):663, 2017.
- [167] J. J. Aubert et al. Production of charmed particles in 250-GeV μ^+ - iron interactions. *Nucl. Phys.*, B213:31–64, 1983.
- [168] M. Capeans, G. Darbo, K. Einsweiler, M. Elsing, T. Flick, M. Garcia-Sciveres, C. Gemme, H. Pernegger, O. Rohne, and R. Vuillermet. ATLAS Insertable B-Layer Technical Design Report. 2010.