Universidade de São Paulo
Instituto de Física

Medidas da Produção de Bósons Z Associados
a Jatos b em Colisões Próton-próton a $\sqrt{s} = 13$
TeV e Estudos de um Sistema de Trigger de
Elétrons para Alta Luminosidade no
Experimento ATLAS

José Luis La Rosa Navarro

Orientador: Dr. Marco Aurelio Lisboa Leite

Tese de doutorado apresentada ao Instituto de Física
da Universidade de São Paulo, como requisito parcial
para a obtenção do título de Doutor em Ciências.

Banca Examinadora:

Dr. Marco Aurelio Lisboa Leite (IFUSP)
Prof. Dr. Arthur Marques Moraes (CBPF)
Prof. Dr. Marcelo G. Munhoz (IFUSP)
Prof. Dr. Marco Bregant (IFUSP)
Profa. Dra Márcia Begalli (UERJ)

São Paulo
2017

FICHA CATALOGRÁFICA
Preparada pelo Serviço de Biblioteca e Informação
do Instituto de Física da Universidade de São Paulo

La Rosa Navarro, Jose Luis

Medidas da produção de bósons Z associados a jatos b em colisões próton-próton a $\sqrt{s}=13$ TeV e estudos de um sistema de trigger de elétrons para alta luminosidade no experimento ATLAS / Measurement of Z boson production in association with b-jets in proton-proton collisions at $\sqrt{s}=13$ TeV and studies of an electron trigger system for high luminosity in the ATLAS experiment . São Paulo, 2017.

Tese (Doutorado) – Universidade de São Paulo. Instituto de Física. Depto. de Física Geral.

Orientador: Prof. Dr. Marco Aurelio Lisboa Leite
Área de Concentração: Física de Altas Energias

Unitermos: 1. Física experimental; 2. Física de partículas; 3. Física de alta energia.

USP/IF/SBI-097/2017

University of São Paulo
Physics Institute

Measurement of Z Boson Production in
Association with b -jets in Proton-proton
Collisions at $\sqrt{s} = 13$ TeV and studies of an
Electron Trigger System for High Luminosity in
the ATLAS Experiment

José Luis La Rosa Navarro

Supervisor: Dr. Marco Aurelio Lisboa Leite

Thesis submitted to the Physics Institute of the University of São Paulo in partial fulfillment of the requirements for the degree of Doctor of Science.

Examining committee:

Dr. Marco Aurelio Lisboa Leite (IFUSP)
Prof. Dr. Arthur Marques Moraes (CBPF)
Prof. Dr. Marcelo G. Munhoz (IFUSP)
Prof. Dr. Marco Bregant (IFUSP)
Profa. Dra Márcia Begalli (UERJ)

São Paulo
2017

*I dedicate this thesis to my parents and brothers,
for the unconditional and constant support along these years.
Gracias, totales!*

Acknowledgements

I would like to start giving the thanks to my supervisor, Marco Leite, for the trust and constant support in the development of this work, including a special thanks to the LIP-IFUSP staff: Edineusa Almeida, Marcel Kuriyama and Ricardo Menegaso.

I have to thank to CAPES and EPLANET for the financial support over these years. Off course, I thank to CERN for the very successful operation of the LHC and the ATLAS experiment collaboration for providing me the data and tools useful for the results presented in this thesis.

I must give a special thanks to Evelin Meoni, for the kind help and fruitful discussions during my V+HF studies, and to Camilla Vittori, for the collaboration. These studies would not be possible without the contributions of the V+HF analysis team: Chiara Debenedetti, Federico Sforza, Douglas Schaefer, Semen Turchikhin and Valentina Cairo.

I am also indebted to David Lopez and Zachary Marshall, from the Jet/Etmiss performance group, for the support during my performance studies; And Cristinel Diaconu and Marisilvia Donadelli for the collaboration during my LAr upgrade studies.

Finally, I would like to thank to my peruvian friends in São Paulo, for the 'emotional' support along my PhD studies.

Abstract

This thesis will present results on the measurement of Z boson production in the electron channel associated with b -jets in proton-proton collision at $\sqrt{s}=13$ TeV, a measurement of fundamental interest in precision measurements and physics searches. This measurement represents important precision tests of perturbative QCD and provides information related to the b -quark inside of the proton. This measurement also represents one of the main sources of *background* in top quark production studies, Higgs precision measurements and the search for supersymmetric particles. The results on cross section measurements and kinematic distributions presented in this work have been unfolded to particle level and are compared to the four and five flavor schemes in Monte Carlo generators, showing that the predictions are consistent within the experimental uncertainties. In this thesis it will be also shown results of the development of the new event selection system proposed for high luminosity measurements. This system will be implemented during the *second LHC long shutdown* (2019-2020), which involves studies on the reconstruction of electrons using the concept of supercells, where the fine granularity of the cells of the ATLAS Liquid Argon detector is exploited to mitigate the problems to be found in high luminosity conditions. One of the problems is the increase of pile-up, caused by secondary collisions of protons in the same event, generating a large number of low p_T jets that can be erroneously identified as electrons. The studies in this work introduce new discriminants based on the electromagnetic shower shape, showing that is possible to further reduce the rate of low p_T jets maintaining a good performance on electron reconstruction.

KEYWORDS: High energy physics; ATLAS-LHC; Proton-proton collisions; High luminosity; Pile-up; Triggers; Bosons; Electrons; Jets; b -tagging.

Resumo

Esta tese apresentará resultados da medida da produção de bósons Z no canal de elétrons associado com jatos b em colisões próton-próton a $\sqrt{s} = 13$ TeV, uma medida de fundamental interesse em medidas de precisão e buscas por nova física. Essa medida representa importantes testes de precisão da QCD perturbativa e prove informações relacionadas ao quark b dentro do próton. Essa medida também representa uma das principais fontes de *background* na produção do quark top, medidas de precisão do Higgs e na busca de partículas supersimétricas. Os resultados na medida da seção de choque e distribuições cinemáticas apresentadas neste trabalho são comparados com os esquemas de quatro e cinco sabores nos geradores Monte Carlo, mostrando que as previsões são consistentes dentro das incertezas experimentais. Nesta tese também serão apresentados estudos do desenvolvimento de um sistema de seleção de eventos proposto para medidas em alta luminosidade. Esse sistema será implementado durante o segundo *LHC long shutdown* (2019-2020), envolvendo estudos na reconstrução de elétrons usando o conceito de supercélulas, onde a fina granularidade das células do detector de Argônio Líquido do ATLAS será aproveitada para mitigar as dificuldades que se terão em condições de alta luminosidade. Um dos problemas é o aumento do empilhamento (*pile-up*), causado por colisões secundárias de prótons no mesmo evento, gerando uma grande quantidade de jatos de baixo p_T que podem ser erroneamente identificados como elétrons. Os estudos neste trabalho apresentam novos discriminantes com base na forma da cascata eletromagnética, mostrando que é possível reduzir ainda mais a taxa de jatos de baixo p_T mantendo uma boa reconstrução de elétrons.

PALAVRAS-CHAVES: Física de altas energias; ATLAS-LHC; Colisões próton-próton; Alta luminosidade; Pile-up; Triggers; Bósons; Eléctrons; Jatos; b -tagging.

List of Figures

1.1	Elementary particles in the Standard Model [11].	3
1.2	Standard Model cross section measurements in ATLAS (status: May 2017) [15].	5
1.3	(a) Scheme of a proton-proton collision showing their constituents. (b) proton-proton collision according to the partons model. . . .	7
1.4	Example of the parton distribution functions, $xf(x, Q^2)$, inside the proton at two different energy scales on function of the moment fraction, x [24].	9
1.5	Diagram with the location of the four major LHC detectors (ATLAS, ALICE, CMS and LHCb), which are located between 50 and 150 m underground [38].	11
1.6	Event display with a low pile-up, $\mu = 4$ at 7 TeV, measured in ATLAS (2010) [38]	13
1.7	The maximum number of inelastic collisions per beam crossing (μ) for proton-proton collisions at 13 TeV during 2015 (a) and 2016 (b) [41].	14
1.8	The peak instantaneous luminosity delivered to ATLAS during stable beams for proton-proton collisions at 13 TeV for each LHC fill (good runs) as a function of time during 2015 (a) and 2016 (b) [41].	14
1.9	Total integrated luminosity delivered and recorded during 2015 (a) and 2016 (b) [41]. For 2015 is also shown the good for physics integrated luminosity.	15
1.10	Transversal segmentation of the ATLAS detector, showing the interaction of particles with trackers, calorimeters and muon spectrometer [38].	16
1.11	(a) Representation on Cartesian coordinates for momentum indicating the longitudinal moment (with respect to the direction of the beams, represent for the particles A and B) and transverse (perpendicular to the direction of the beams); (b) Pseudorapidity as a function of the polar angle, θ	17

1.12	(a) Proton-Proton collision and jets production. (b) Estimated energy fraction of particles that make up the simulated jets (truejets) [44].	19
1.13	Feynman diagrams for proton-proton collision and production of W and Z on the electron channel [46].	20
1.14	Scheme of the experimental detection of W and Z on the electron channel [46].	20
1.15	The data/theory ratio for several single-boson fiducial production cross section measurements. All theoretical expectations were calculated at NLO or higher (status: August 2016) [15]. In the green square are present the $Z + b$ -jets measurements. . .	22
1.16	Event with $\mu = 25$ at 13 TeV, measured in ATLAS (2016) [38] .	23
1.17	The program for the next years according to the LHC commissioning [57] (January 2016).	24
2.1	The ATLAS Detector: A Toroidal LHC Apparatus [62].	28
2.2	Pseudorapidity coverage of the ATLAS detector subsystems [62].	28
2.3	Scheme of the Inner Detector and its sub-detectors [38].	29
2.4	Scheme of the track reconstruction for three particles and their passing through the ID sub-detectors.	29
2.5	(a) Primary vertex reconstruction NPV=4 ($\mu = 4$). (b) Transverse (d_0) and longitudinal (z_0) impact parameters between a primary and a secondary vertex.	30
2.6	NPV vs μ at different levels of luminosity, for dijet simulations for collisions at 13 TeV (from samples used on jet calibration studies at high luminosity (App. J)).	31
2.7	The Liquid Argon calorimeter of ATLAS [38].	32
2.8	Details of the electromagnetic calorimeter, electromagnetic barrel (left) and one of the end-caps (EMEC) (right) [62].	33
2.9	Partial isometric view of the electromagnetic calorimeter, showing the transversal and longitudinal segmentation [62].	34
2.10	a) Transversal segment of the EMB detailing its components. b) Principle of ionization signal generation from the electromagnetic shower [70].	34
2.11	Cells clusterization on the electromagnetic calorimeter.	35
2.12	Scheme of the 3 levels of trigger used in ATLAS.	36
2.13	Block diagram of L1 trigger [71].	37
2.14	Block diagram for the calorimeter trigger [75].	38

2.15	Shower for an electron (left) and a pion (right), both at 50 GeV, through iron [77].	39
2.16	Scheme used in algorithm for the L1 trigger for electrons and photons identification [72].	40
2.17	Trigger rate for electrons at L1 as a function of the p_T threshold at $10^{33} \text{ cm}^{-2}\text{s}^{-1}$ (left) and $10^{34} \text{ cm}^{-2}\text{s}^{-1}$ (right). The solid lines represent the rates without isolation; dotted ones, only with hadronic isolation; and dashed ones with both hadronic and electromagnetic isolation [72].	41
2.18	Scheme used for jet algorithm at Level 1 [72].	43
2.19	Efficiencies for L1 single-jet triggers, shown as a function of leading off-line jet p_T for different thresholds (J20 indicates a p_T cut of 20 GeV, similar for the others) . (a) For the central region ($ \eta < 2.8$) and (b) for the forward region ($3.6 < \eta < 4.5$) [78]. .	43
3.1	An event display with the production of two electrons ($Z \rightarrow ee$) measured in ATLAS [38].	46
3.2	Scheme of the segmentation for towers used on electron reconstruction.	46
3.3	Scheme of clusters used for R_η measurement.	49
3.4	Distribution of R_η and $W_{\eta,2}$ for electrons and jets with $p_T > 20 \text{ GeV}$ at high luminosity ($\mu = 80$) [8]	51
3.5	Scheme used in the calibration of electrons and photons [82]. . .	52
3.6	Most probable value (MPV) of the energy response for electrons (a) and unconverted photons (b), as a function of $ \eta $ for different energy values [82]. This analysis does not consider values in the transition region from the barrel to the end-caps (crack region: $1.37 < \eta < 1.55$).	53
4.1	Event display with jets production measured in ATLAS [38]. . .	56
4.2	Feynman diagrams for gluon emission on a $\gamma^* \rightarrow q\bar{q}$ process (virtual photon). $\mathbf{p}_1$ and $\mathbf{p}_2$ are the tri-momentum of the quarks and $\mathbf{k}$ of the gluon.	58
4.3	Jet production with soft gluon emission: (a) Production at LO. (b) Production at NLO with a virtual gluon. (c) Production at NLO with a radiated gluon [87].	59
4.4	Jet production with collinear emission, heights of vertical lines represent the transverse momenta and lengths of horizontal lines the angular separation [87].	59
4.5	2D scheme sequence for topocluster reconstruction in the η phase space.	60

4.6	Reconstruction of jets with different algorithms: The Cam/Aachen (a), the k_t (b), the anti- k_t (c) and the SISCone (d) [90].	62
4.7	Simulation of the Higgs mass reconstruction using different algorithms in the reconstruction of jets ($p\bar{p} \rightarrow WH \rightarrow \ell\nu bb$) [92]. .	64
4.8	Simulation of Higgs mass reconstruction, comparing JETCLU ($R_0 = 0.4$) with kt algorithm at $R_0 = 0.6$ (a) and $R_0 = 1.0$ (b) [92].	65
4.9	Jet reconstruction using track and deposited energy information [44].	65
4.10	Reconstruction sequence of jet types in ATLAS [94].	66
4.11	The mean anti- k_t $R = 0.4$ LCW jet multiplicity as a function of $\langle\mu\rangle$ in Z +jets events for jets with $p_T > 20$ GeV and $ \eta < 2.1$ (a) and $3.2 < \eta < 4.5$ (b). Events in this plot are required to have at least 1 jet both before and after applying the jet-area based pile-up correction [96].	68
4.12	JVF distribution for hard-scatter (blue) and pile-up (red) jets with $20 < p_T < 50$ GeV and $ \eta < 2.4$ after pile-up subtraction and jet energy scale correction in simulated Z +jets events [96]. .	70
4.13	The mean anti- k_t $R = 0.4$ LCW+JES jet multiplicity as a function of $\langle\mu\rangle$ in Z +jets events for jets with $ \eta < 2.1$, before and after selected JVF cuts were applied to jets with $20 < p_T < 50$ GeV (a) and $40 < p_T < 50$ GeV (b) [96].	70
4.14	Ratio of inclusive jet fiducial cross section measurements to theory (status: August 2016) [15].	71
5.1	Linearity expectations before (a) and after (b) the calibration ($R \rightarrow 1$) [44].	74
5.2	Expectations for resolution improvement ($\sigma_E/E \rightarrow 0$) before (a) and after (b) the calibration [44].	75
5.3	Criteria of the matching cut in the phase space η - ϕ	76
5.4	2D histogram of the jet energy response (a) and the energy correlation (b) before calibration, for EM scale jets with $R=0.4$. In the correlation plot, the green line represents the identity function and the yellow one represents the minimum transverse energy value of the true-jet.	78
5.5	Gaussian adjustment for $\mathcal{R}$ (blue curve), considering two different E_{true} bins and the same pseudorapidity range ($0.0 < \eta < 0.1$)	78

5.6	Fitting of the energy calibration curve (blue line) for two pseudorapidity ranges. The red dotted line represents an extrapolation curve, useful to improve the fitting with the help of Groom points.	79
5.7	Jet energy response (left) and correlation energy (right), after calibration, for EM+JES scale jets with $R = 0.4$. In the correlation plot, the green line represents the identity function and the yellow one represents the minimum transverse energy value of the true-jet.	79
5.8	Calibration curve for pseudorapidity as a function of E_{true} (green line) for two pseudorapidity ranges.	80
5.9	Energy response before the calibration (EM scale) for $R = 0.4$ and 0.6 , as function of the pseudorapidity and for five energy values of reference (energy of true-jets).	82
5.10	Energy response after calibration (EM+JES scale) for $R = 0.4$ and 0.6 , as function of the pseudorapidity and for five energy values of reference.	83
5.11	Energy response after calibration (LCW+JES scale) for $R = 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8$, as function of the pseudorapidity and for five energy reference values (energy of true-jets).	85
6.1	Feynman diagrams considered on the four-flavor number scheme: LO for Z production with at least 2 b -jets ($q = u, d, s$ and c).	88
6.2	Feynman diagrams considered on the five-flavour number scheme: LO for Z production with at least 1 b -jet.	88
6.3	NLO Feynman diagrams for Z production with at least one b -jet ($q = u, d, s$ and c) on the five-flavour number scheme.	89
6.4	Scheme used to describe the b -hadron association.	91
6.5	Scheme used to describe the b -ghost association. R_0 is the clusterization radius.	91
6.6	Scheme of vertex displacement for a b -quark [109].	92
6.7	Light-jet rejection (a) and c-jet rejection (b) as a function of the b -tag efficiency for the b -tagging algorithms calibrated on analysis at 7 TeV, based on simulated $t\bar{t}$ events [114].	94
6.8	$t\bar{t}$ production, with $t \rightarrow Wb$	95
6.9	ZZ production, with one $Z \rightarrow \ell\ell$ and other $Z \rightarrow qq$	96

6.10	Cross-sections for $Z + \geq 1b$ -jet (a) and $Z + \geq 2b$ -jets (b) measured in ATLAS for collisions at 7 TeV. The data measurement is shown as a vertical blue line with the inner blue shaded band showing the corresponding statistical uncertainty and the outer green shaded band showing the sum in quadrature of statistical and systematic uncertainties [115].	99
7.1	Electron states, showing the born, bare and dressed level. The same state classification is valid for muons.	108
7.2	Z boson invariant mass (left) and p_T (right) comparisons on Alpgen simulations for $Z(\rightarrow ee)$ +jets at $\sqrt{s} = 13$ TeV (before and after the HFOR).	109
7.3	Inclusive multiplicity comparisons for jets (left) and b -jets (right) on Alpgen simulations for $Z(\rightarrow ee)$ +jets at $\sqrt{s} = 13$ TeV (before and after the HFOR).	109
7.4	Z boson invariant mass (left) and p_T (right) comparisons on MC simulations for $Z(\rightarrow ee)$ +jets at $\sqrt{s} = 13$ TeV.	111
7.5	Lepton transverse momentum (left) and pseudo-rapidity (right) comparisons on MC simulations for $Z(\rightarrow ee)$ +jets at $\sqrt{s} = 13$ TeV.	111
7.6	Inclusive jet (left) and b -jet (right) multiplicity comparisons on MC simulations for $Z(\rightarrow ee)$ +jets at $\sqrt{s} = 13$ TeV.	112
7.7	Transverse momentum and rapidity comparisons for leading (top) and sub-leading (bottom) b -jets on MC simulations for $Z(\rightarrow ee)$ +jets at $\sqrt{s} = 13$ TeV.	114
7.8	Angular distributions ($\Delta\phi_{bb}$, Δy_{bb} and ΔR_{bb}) and invariant mass (M_{bb}) comparisons for dijets on MC simulations for $Z(\rightarrow ee)$ +jets at $\sqrt{s} = 13$ TeV.	115
7.9	Transverse momentum and rapidity comparisons for Z boson in events with at least $1b$ jet (top) and at least $2b$ -jets on MC simulations for $Z(\rightarrow ee)$ +jets at $\sqrt{s} = 13$ TeV.	116
7.10	MET distributions of $t\bar{t}$ events and signal events for $1b$ -jet selection (left) and $2b$ -jet selection (right) in di-electron modes. Histograms are normalized to the theoretically predicted cross section [135].	119
7.11	Preliminary comparisons Pred/Data for Z boson invariant mass (left) and transverse moment (right) distributions on $Z + b$ -jets analysis at 13 TeV (Status: June 2017).	121

7.12	Comparisons Pred/Data for the exclusive (top) and inclusive (bottom) multiplicity distribution for jets (left) and b -jets (right) on $Z + b$ -jets analysis at $\sqrt{s}=13$ TeV ($Z \rightarrow ee$) (Status: July 2017).	123
7.13	Comparisons Pred/Data for the transverse momentum and rapidity for leading (top) and sub-leading (bottom) jets on $Z + b$ -jets analysis at $\sqrt{s}=13$ TeV ($Z \rightarrow ee$) (Status: May 2017).	124
7.14	Comparisons Pred/Data for the angular distributions ($\Delta\phi_{jj}$, Δy_{jj} and ΔR_{jj}) and invariant mass (M_{jj}) for di-jets on $Z+b$ -jets analysis at $\sqrt{s}=13$ TeV ($Z \rightarrow ee$) (Status: May 2017).	125
7.15	Comparisons Pred/Data for the transverse momentum and rapidity comparisons for leading (top) and sub-leading (bottom) b -jets on $Z + b$ -jets analysis at $\sqrt{s}=13$ TeV ($Z \rightarrow ee$) (Status: May 2017).	126
7.16	Comparisons Pred/Data for the angular distributions ($\Delta\phi_{bb}$, Δy_{bb} and ΔR_{bb}) and invariant mass (M_{bb}) for di-jets on $Z+b$ -jets analysis at $\sqrt{s}=13$ TeV ($Z \rightarrow ee$) (Status: May 2017).	127
7.17	Comparisons Pred/Data for the transverse momentum and rapidity for the Z boson in events with at least $1b$ -jet (top) and at least $2b$ -jets (bottom) on $Z + b$ -jets analysis at $\sqrt{s}=13$ TeV ($Z \rightarrow ee$) (Status: May 2017).	128
7.18	Response matrices corresponding to the exclusive jet (left) and b -jet multiplicity (right) obtained with Sherpa.	131
7.19	Efficiency reconstruction as a function of the exclusive jet and b -jet multiplicity.	132
7.20	Fake correction as a function of the exclusive jet and b -jet multiplicity.	133
7.21	Closure test for the exclusive jet and b -jet multiplicity.	134
7.22	Unfolding test for the exclusive jet and b -jet multiplicity comparing Sherpa and Alpgen reconstruction. Data unfolded with Sherpa.	134
7.23	Unfolding test for the leading (top) and sub-leading (bottom) jet p_T and rapidity comparing Sherpa and Alpgen reconstruction. Data unfolded with Sherpa.	135
7.24	Unfolding test for $\Delta\phi_{jj}$, Δy_{jj} , ΔR_{jj} and M_{jj} of dijets comparing Sherpa and Alpgen reconstruction. Data unfolded with Sherpa.	136
7.25	Unfolding test for the leading (top) and sub-leading (bottom) b -jet p_T and rapidity comparing Sherpa and Alpgen reconstruction. Data unfolded with Sherpa.	137

7.26	Unfolding test for $\Delta\phi_{bb}$, Δy_{bb} , ΔR_{bb} and M_{bb} of dijets comparing Sherpa and Alpgen reconstruction. Data unfolded with Sherpa.	138
7.27	Unfolding test for p_T and $ y $ of Z bosons in events with at least 1 b -jet (top) and at least 2 b -jets (bottom) comparing Sherpa and Alpgen reconstruction. Data unfolded with Sherpa.	139
7.28	Breakdown of the jet energy scale systematics for exclusive jet and b -jet multiplicity.	141
7.29	Breakdown of the b -tagging efficiency and mistag rates systematics for exclusive jet and b -jet multiplicity. The left plot is presented as reference, since is not considered on full jet analysis.	142
7.30	Breakdown of the electron systematic for exclusive jet and b -jet multiplicity. Note the changes in the scale for the Y axis.	143
7.31	Total systematic uncertainty and their main sources for exclusive jet and b -jet multiplicity.	143
7.32	Total systematic uncertainty and their main sources for p_T and $ y $ for the leading (top) and sub-leading (bottom) jet.	144
7.33	Total systematic uncertainty and their main sources for $\Delta\phi_{jj}$, Δy_{jj} , ΔR_{jj} and M_{jj} of di-jets.	145
7.34	Total systematic uncertainty and their main sources for p_T and $ y $ for the leading (top) and sub-leading (bottom) b -jet.	146
7.35	Total systematic uncertainty and their main sources for $\Delta\phi_{bb}$, Δy_{bb} , ΔR_{bb} and M_{bb} of di-jets.	147
7.36	Total systematic uncertainty and their main sources for p_T and $ y $ of Z bosons in events with at least 1 b -jet (top) and at least 2 b -jets (bottom).	148
7.37	Measured cross section as a function of the exclusive (top) and inclusive (bottom) jet (left) and b -jet (right) multiplicity at $\sqrt{s}$ =13 TeV. The data are compared to the predictions by Sherpa 2.2 and Alpgen. The error bars indicate the statistical uncertainty, and the hatched bands the statistical and systematic uncertainties on data added in quadrature.	150
7.38	Measured cross section as a function of the transverse momentum (p_T) and rapidity ($ y $) for leading (top) and sub-leading (bottom) jets. The data are compared to the predictions by Sherpa 2.2 and Alpgen. The error bars indicate the statistical uncertainty, and the hatched bands the statistical and systematic uncertainties on data added in quadrature.	151

7.39	Measured cross section as a function of the angular distributions ($\Delta\phi_{jj}$, Δy_{jj} and ΔR_{jj}) and invariant mass (M_{jj}) for dijets. The data are compared to the predictions by Sherpa 2.2 and Alpgen. The error bars indicate the statistical uncertainty, and the hatched bands the statistical and systematic uncertainties on data added in quadrature.	152
7.40	Measured cross section as a function of the transverse momentum (p_T) and rapidity ($ y $) for leading (top) and sub-leading (bottom) b -jet. The data are compared to the predictions by Sherpa 2.2 and Alpgen. The error bars indicate the statistical uncertainty, and the hatched bands the statistical and systematic uncertainties on data added in quadrature.	153
7.41	Measured cross section as a function of the angular distributions ($\Delta\phi_{bb}$, Δy_{bb} and ΔR_{bb}) and invariant mass (M_{bb}) for di b -jets. The data are compared to the predictions by Sherpa 2.2 and Alpgen. The error bars indicate the statistical uncertainty, and the hatched bands the statistical and systematic uncertainties on data added in quadrature.	154
7.42	Measured cross section as a function of the transverse momentum (p_T) and rapidity ($ y $) for the Z boson in events with at least 1 b -jet (top) and at least 2 b -jets (bottom). The data are compared to the predictions by Sherpa 2.2 and Alpgen. The error bars indicate the statistical uncertainty, and the hatched bands the statistical and systematic uncertainties on data added in quadrature.	155
7.43	Cross-sections comparisons for $Z + \geq 1b$ -jet (a) and $Z + \geq 2b$ -jets (b) measured in ATLAS for collisions at 7 and 13 TeV. The data measurement is shown as a vertical black line with the inner blue shaded band showing the corresponding statistical uncertainty and the outer green shaded band showing the sum in quadrature of statistical and systematic uncertainties.	157
8.1	Simulation of one electron ($E_T = 70$ GeV) seen by the current L1 trigger towers (a) and by supercells (b) [8].	160
8.2	Distribution of the number of electrons produced by event (a), the transverse energy (b), pseudorapidity (c) and azimuthal angle (b) for electrons at truth level coming from $Z \rightarrow ee$ at $\sqrt{s}=14$ TeV.	162
8.3	Distribution of the invariant mass (a) and p_T (b) of the Z boson at truth level coming from $Z \rightarrow ee$ at $\sqrt{s}=14$ TeV.	162

8.4	Scheme of the finding algorithm used on the reconstruction of electrons with supercells, where is observed how the information is collected from neighbors to the seeds found in different layers of the electromagnetic calorimeter in a window 3×3 in units of $\Delta\eta \times \Delta\phi = 0.025 \times 0.1$	164
8.5	Distribution of the number of electrons produced by event (a), the transverse energy (b), pseudorapidity (c) and azimuthal angle (b) comparing super-electrons and electrons at truth level coming from $Z \rightarrow ee$ at $\sqrt{s}=14$ TeV.	165
8.6	Distribution of the invariant mass (a) and p_T (b) of the reconstructed Z boson with super-electrons and electrons at truth level ($Z \rightarrow ee$ at $\sqrt{s}=14$ TeV).	165
8.7	Energy distribution in the supercells for the different layers of the electromagnetic calorimeter in an event with one electron candidate ($E_T = 45$ GeV, $\eta = 0.95$ and $\phi = -2.75$), based on a simulation of $Z \rightarrow ee$ with $\mu = 80$	166
8.8	Energy distribution of seeds (a) and super-electron (b) for different values of μ	167
8.9	Distribution of transverse energy for super-electrons and fake electrons. without the E_T^{seed} cutting (a) and with cutting (b).	167
8.10	Proposed strategy on Super-Electron analysis.	169
8.11	Comparison of the number of electrons produced per event (a) and transverse energy (b) for super-electrons, the level 1 trigger signals and truth electrons for a sample of $Z \rightarrow ee$ at $\sqrt{s} = 13$ TeV ($\mu = 80$).	170
8.12	Transverse energy efficiency comparisons: for super-electrons vs. L1 (a) and super electrons vs. off-line electrons(b).	171
8.13	Distribution of number of electrons reconstructed event by event (a), transverse energy (b), pseudorapidity (c) and azimuthal angle (d) for reconstructed super-electrons, reconstructed off-line electrons and truth electrons, for a sample of $Z \rightarrow ee$ at $\sqrt{s} = 13$ TeV.	172
8.14	Reconstruction efficiency ($Z \rightarrow ee$) (a) and rate production (b) distributions as a function of the E_T threshold, for electrons reconstructed from the supercell information, L1 trigger electron signals and off-line electrons.	173
8.15	Distribution of the transverse momentum for super-electrons (a) and jets, identified as fake electrons (b).	173

8.16	$R_{\eta,2}$, $R_{\eta,1}$ and $R_{\eta,1+2}$ shape distributions for over electrons and jets samples reconstructed with the super-electron finding algorithm.	175
8.17	Efficiency (a) and rate (b) signal distributions as a function of the minimum value (cut) for $R_{\eta,2}$, $R_{\eta,1}$ and $R_{\eta,1+2}$ reconstructed with the electron finding.	176
8.18	Signal efficiency as a function of the E_T threshold with different cuts on $R_{\eta,2}$, $R_{\eta,1}$ and $R_{\eta,1+2}$	177
8.19	Signal rates (and ratios) as a function of E_T threshold for different cuts on $R_{\eta,2}$, $R_{\eta,1}$ and $R_{\eta,1+2}$	177
8.20	Shape distributions for $W_{\eta,1}$, $W_{\eta,2}$ and f_3 over jets and electrons samples reconstructed with the electron finding.	179
8.21	Efficiency (a) and rate (b) signal distributions as a function of the maximum value (cut) for $W_{\eta,1}$ and $W_{\eta,2}$ reconstructed with the electron finding.	180
8.22	Efficiency (a) and rate (b) signal distributions as a function of the maximum value (cut) for f_3 reconstructed with the electron finding.	180
8.23	Signal efficiency as a function of the E_T threshold with different cuts on $W_{\eta,1}$, $W_{\eta,2}$ and f_3	181
8.24	Signal rates (and ratios) as a function of E_T threshold for different cuts on $W_{\eta,1}$, $W_{\eta,2}$ and f_3	182
8.25	Shape distributions for R_{seed} , R_{had} and $R_{2^{nd}hot}$ over jets and electrons samples reconstructed with the electron finding.	184
8.26	Efficiency (a) and rate (b) signal distributions as a function of the maximum value (cut) for R_{had} reconstructed with the electron finding.	184
8.27	Efficiency (a) and rate (b) signal distributions as a function of the maximum value (cut) for R_{seed} reconstructed with the electron finding.	185
8.28	Efficiency (a) and rate (b) signal distributions as a function of the maximum value (cut) for the seed energy ratio to second hottest cell reconstructed with the electron finding.	185
8.29	Correlation matrix between discriminators for signals (top) and background (bottom), represented by jets (produced with the TMVA package).	186
8.30	Signal efficiency as a function of the E_T threshold combining cuts on $R_{\eta,1+2}$ and $W_{\eta,2}$	187
8.31	Signal rates (and ratios) as a function of E_T threshold combining cuts on $R_{\eta,1+2}$ and $W_{\eta,2}$	188

8.32	Distribution of the transverse momentum for super-electrons compared with truth electrons. Before and after cuts for super-electrons.	189
8.33	Distribution of the transverse momentum for fake electrons, before and after the cuts.	189
8.34	Current scheme of architecture employed on reading and processing of individual signals formation at L1 trigger of the electromagnetic calorimeter in ATLAS [8].	191
8.35	New scheme of architecture involving the digital trigger system highly segmented to the front-end electronic of electromagnetic calorimeter [8]. The red lines indicate the parts that will be added/modified.	192
8.36	Picture of the electronic board LTDB prototype (left). Position of crate in the detector with the LTDB, just above the calorimeter cryostat (right).	193
8.37	Covered region of the new digital system of highly segmented trigger (LTDB).	193
8.38	Kinematic supercell distributions (E_T , η , ϕ) reconstructed with the CaloCell (left) and LTDB (right) information.	195
8.39	Correlations for the supercell E_T reconstructed with the CaloCell and LTDB information for each layer of the electromagnetic calorimeter. The red line represents a linear fitting.	196
8.40	$R_{\eta,2}$, $R_{\eta,1+2}$ and $W_{\eta,2}$ distributions for super-electrons reconstructed with the CaloCell (left) and LTDB (right) information, considering different trigger selections at L1.	197
8.41	Prediction/Data cross-section comparisons for $Z + \geq 1b$ -jet (a) and $Z + \geq 2b$ -jets (b) measured in ATLAS for collisions at 7 and 13 TeV. The inner blue shaded band showing the corresponding statistical uncertainty and the outer green shaded band showing the sum in quadrature of statistical and systematic uncertainties.	200
8.42	$R_{\eta,1+2}$ and $W_{\eta,2}$ distributions for super-electrons reconstructed with the CaloCell information on simulations (left) and data (right), considering electron and jet samples on simulations and different L1 trigger selections on data.	201
A.1	W/Z boson production into two leptons.	203
B.1	Scheme of simulation and reconstruction process in ATLAS.	205
B.2	Diagram showing the parton shower evolution from an initial quark line [146].	208

B.3	Scheme describing the confinement for a pair $q\bar{q}$. The dashed lines represent the field between quarks.	208
B.4	Diagram to describe the string model, showing the pair $q\bar{q}$ moving apart from each other as seen in the meson rest frame (left) and in a boosted frame (right) [145, 146]. Blue and green lines represent the quark 'ways' and red lines the strings between them.	209
B.5	The motion and breakup of a string system, with the two transverse degrees of freedom suppressed [145]. Blue and green lines represent the quark 'ways' and red lines the strings between them.	210
B.6	Colour structure of a parton shower to leading order in N_c [145].	211
B.7	Sketch of a proton-proton collision showing the Lorentz contraction of the protons [146].	212
C.1	LO Feynman diagram for $q + \bar{q} \rightarrow Z$ production.	213
C.2	LO Feynman diagram for $Z + 1$ -jet production.	218
C.3	LO Feynman diagram for $q + g \rightarrow Z + q$ production.	219
D.1	Z boson invariant mass comparisons on Alpgen simulations for $Z(\rightarrow \ell\ell)$ +jets at $\sqrt{s} = 13$ TeV ($\ell = e, \mu$). Comparisons before and after the HFOR.	223
D.2	Z boson transverse moment (top) and rapidity (bottom) comparisons on Alpgen simulations for $Z(\rightarrow \ell\ell)$ +jets at $\sqrt{s} = 13$ TeV ($\ell = e, \mu$). Comparisons before and after the HFOR. . .	224
D.3	lepton transverse moment (top) and pseudo-rapidity (bottom) comparisons on Alpgen simulations for $Z(\rightarrow \ell\ell)$ +jets at $\sqrt{s} = 13$ TeV ($\ell = e, \mu$). Comparisons before and after the HFOR. . .	225
D.4	Exclusive (top) and inclusive (bottom) multiplicity comparisons for jets on Alpgen simulations for $Z(\rightarrow \ell\ell)$ +jets at $\sqrt{s} = 13$ TeV ($\ell = e, \mu$). Comparisons before and after the HFOR.	226
D.5	Transverse moment (top) and rapidity (bottom) comparisons for leading jets ($Z \rightarrow \ell\ell$) on Alpgen simulations for $Z(\rightarrow \ell\ell)$ +jets at $\sqrt{s} = 13$ TeV ($\ell = e, \mu$). Comparisons before and after the HFOR.	227
D.6	Transverse moment (top) and rapidity (bottom) comparisons for sub-leading jets ($Z \rightarrow \ell\ell$) on Alpgen simulations for $Z(\rightarrow \ell\ell)$ +jets at $\sqrt{s} = 13$ TeV ($\ell = e, \mu$). Comparisons before and after the HFOR.	228
D.7	Inclusive multiplicity comparisons for c -jets (top) and b -jets (bottom) on Alpgen simulations for $Z(\rightarrow \ell\ell)$ +jets at $\sqrt{s} = 13$ TeV ($\ell = e, \mu$). Comparisons before and after the HFOR.	229

D.8	Transverse moment (top) and rapidity (bottom) comparisons for leading c -jets ($Z \rightarrow ll$) on Alpgen simulations for $Z(\rightarrow \ell\ell)$ +jets at $\sqrt{s} = 13$ TeV ($\ell = e, \mu$). Comparisons before and after the HFOR.	230
D.9	Transverse moment (top) and rapidity (bottom) comparisons for sub-leading c -jets ($Z \rightarrow ll$) on Alpgen simulations for $Z(\rightarrow \ell\ell)$ +jets at $\sqrt{s} = 13$ TeV ($\ell = e, \mu$). Comparisons before and after the HFOR.	231
D.10	Transverse moment (top) and rapidity (bottom) comparisons for leading b -jets ($Z \rightarrow ll$) on Alpgen simulations for $Z(\rightarrow \ell\ell)$ +jets at $\sqrt{s} = 13$ TeV ($\ell = e, \mu$). Comparisons before and after the HFOR.	232
D.11	Transverse moment (top) and rapidity (bottom) comparisons for sub-leading b -jets ($Z \rightarrow ll$) on Alpgen simulations for $Z(\rightarrow \ell\ell)$ +jets at $\sqrt{s} = 13$ TeV ($\ell = e, \mu$). Comparisons before and after the HFOR.	233
E.1	Z boson invariant mass comparisons on MC simulations for $Z(\rightarrow \ell\ell)$ +jets at $\sqrt{s} = 13$ TeV ($\ell = e, \mu$).	236
E.2	Z boson transverse momentum (top) and rapidity (bottom) comparisons on MC simulations for $Z(\rightarrow \ell\ell)$ +jets at $\sqrt{s} = 13$ TeV ($\ell = e, \mu$).	237
E.3	Lepton transverse momentum (top) and pseudo-rapidity (bottom) comparisons on MC simulations for $Z(\rightarrow \ell\ell)$ +jets at $\sqrt{s} = 13$ TeV ($\ell = e, \mu$).	238
E.4	Inclusive multiplicity comparisons for jets (top) and b -jets (bottom) on MC simulations for $Z(\rightarrow \ell\ell)$ +jets at $\sqrt{s} = 13$ TeV ($\ell = e, \mu$).	239
E.5	Transverse momentum and rapidity comparisons for leading jets on MC simulations for $Z(\rightarrow \ell\ell)$ +jets at $\sqrt{s} = 13$ TeV ($\ell = e, \mu$).	240
E.6	Transverse momentum and rapidity comparisons for sub-leading jets on MC simulations for $Z(\rightarrow \ell\ell)$ +jets at $\sqrt{s} = 13$ TeV ($\ell = e, \mu$).	241
E.7	Transverse momentum and rapidity comparisons for leading b -jets on MC simulations for $Z(\rightarrow \ell\ell)$ +jets at $\sqrt{s} = 13$ TeV ($\ell = e, \mu$).	242
E.8	Transverse momentum and rapidity comparisons for sub-leading b -jets on MC simulations for $Z(\rightarrow \ell\ell)$ +jets at $\sqrt{s} = 13$ TeV ($\ell = e, \mu$).	243

F.1	Response matrices corresponding to the exclusive jet (top) and b -jet (bottom) multiplicity obtained with Sherpa (left) and Alpgen (right).	246
F.2	Response matrices corresponding to the leading (top) and sub-leading (bottom) jet p_T obtained with Sherpa (left) and Alpgen (right).	247
F.3	Response matrices corresponding to the leading (top) and sub-leading (bottom) jet rapidity obtained with Sherpa (left) and Alpgen (right).	248
F.4	Response matrices corresponding to $\Delta\phi_{jj}$ (top) and Δy_{jj} (bottom) of dijets obtained with Sherpa (left) and Alpgen (right).	249
F.5	Response matrices corresponding to ΔR_{jj} (top) and M_{jj} (bottom) of dijets obtained with Sherpa (left) and Alpgen (right).	250
F.6	Response matrices corresponding to the leading (top) and sub-leading (bottom) b -jet p_T obtained with Sherpa (left) and Alpgen (right).	251
F.7	Response matrices corresponding to the leading (top) and sub-leading (bottom) b -jet rapidity obtained with Sherpa (left) and Alpgen (right).	252
F.8	Response matrices corresponding to $\Delta\phi_{bb}$ and Δy_{bb} of di-jets obtained with Sherpa (left) and Alpgen (right).	253
F.9	Response matrices corresponding to ΔR_{bb} and M_{bb} of di-jets obtained with Sherpa (left) and Alpgen (right).	254
F.10	Response matrices corresponding to p_T of Z bosons in events with at least $1b$ -jet (top) and at least $2b$ -jets (bottom) obtained with Sherpa (left) and Alpgen (right).	255
F.11	Response matrices corresponding to $ y $ of Z bosons in events with at least $1b$ -jet (top) and at least $2b$ -jets (bottom) obtained with Sherpa (left) and Alpgen (right).	256
F.12	Efficiency reconstruction as a function of the exclusive jet and b -jet multiplicity comparing Sherpa and Alpgen.	257
F.13	Efficiency reconstruction as a function of the leading (top) and sub-leading (bottom) jet p_T and rapidity comparing Sherpa and Alpgen.	258
F.14	Efficiency reconstruction as a function of $\Delta\phi_{jj}$, Δy_{jj} , ΔR_{jj} and M_{jj} of dijets comparing Sherpa and Alpgen.	259
F.15	Efficiency reconstruction as a function of the leading (top) and sub-leading (bottom) b -jet p_T and rapidity comparing Sherpa and Alpgen.	260

F.16	Efficiency reconstruction as a function of $\Delta\phi_{bb}$, Δy_{bb} , ΔR_{bb} and M_{bb} of dijets comparing Sherpa and Alpgen.	261
F.17	Efficiency reconstruction as a function of p_T and $ y $ of Z bosons in events with at least 1 b -jet (top) and at least 2 b -jets (bottom) comparing Sherpa and Alpgen.	262
F.18	Fake correction as a function of the exclusive jet and b -jet multiplicity comparing Sherpa and Alpgen.	263
F.19	Fake correction as a function of the leading (top) and sub-leading (bottom) jet p_T and rapidity comparing Sherpa and Alpgen.	264
F.20	Fake correction as a function of $\Delta\phi_{jj}$, Δy_{jj} , ΔR_{jj} and M_{jj} of dijets comparing Sherpa and Alpgen.	265
F.21	Fake correction as a function of the leading (top) and sub-leading (bottom) b -jet p_T and rapidity comparing Sherpa and Alpgen.	266
F.22	Fake correction as a function of $\Delta\phi_{bb}$, Δy_{bb} , ΔR_{bb} and M_{bb} of dijets comparing Sherpa and Alpgen.	267
F.23	Fake correction as a function of p_T and $ y $ of Z bosons in events with at least 1 b -jet (top) and at least 2 b -jets (bottom) comparing Sherpa and Alpgen.	268
F.24	Closure test for the leading (top) and sub-leading (bottom) jet p_T and rapidity obtained with Sherpa.	269
F.25	Closure test for $\Delta\phi_{jj}$, Δy_{jj} , ΔR_{jj} and M_{jj} of dijets obtained with Sherpa.	270
F.26	Closure test for the leading (top) and sub-leading (bottom) b -jet p_T and rapidity obtained with Sherpa.	271
F.27	Closure test for $\Delta\phi_{bb}$, Δy_{bb} , ΔR_{bb} and M_{bb} of dijets obtained with Sherpa.	272
F.28	Closure test for p_T and $ y $ of Z bosons in events with at least 1 b -jet (top) and at least 2 b -jets (bottom) obtained with Sherpa.	273
G.1	Comparisons between measured cross section for the official Z +jets analysis at 13 TeV (left) and VHF results (right) as a function of the inclusive jet multiplicity. The data are compared to the predictions by Sherpa and Alpgen on VHF, including Blackhat and Madgraph on the official analysis [131]. The error bars indicate the statistical uncertainty, and the hatched bands the statistical and systematic uncertainties on data added in quadrature.	275

G.2	Comparisons between measured cross section for the official Z +jets analysis at 13 TeV (left) and VHF results (right) as a function of transverse momentum (top) and rapidity (bottom) of the leading jet. The data are compared to the predictions by Sherpa and Alpgen on VHF, including Blackhat and Madgraph on the official analysis [131]. The error bars indicate the statistical uncertainty, and the hatched bands the statistical and systematic uncertainties on data added in quadrature.	276
G.3	Comparisons between measured cross section for the official Z +jets analysis at 13 TeV (left) and VHF results (right) as a function of $\Delta\phi_{jj}$ and $M(j, j)$ for di-jets. The data are compared to the predictions by Sherpa and Alpgen on VHF, including Blackhat and Madgraph on the official analysis [131]. The error bars indicate the statistical uncertainty, and the hatched bands the statistical and systematic uncertainties on data added in quadrature.	277
H.1	Comparisons between measured cross section for results at 7 (left) [115] and 13 TeV (right) as a function of the transverse momentum (top) and rapidity (bottom) for Z boson in events with at least $1b$ -jet. The data are compared to the predictions by Sherpa and Alpgen. The error bars indicate the statistical uncertainty, and the hatched bands the statistical and systematic uncertainties on data added in quadrature.	280
H.2	Comparisons between measured cross section for results at 7 (left) [115] and 13 TeV (right) as a function of the transverse momentum (top) and rapidity (bottom) for Z boson in events with at least $2b$ -jets. The data are compared to the predictions by Sherpa and Alpgen. The error bars indicate the statistical uncertainty, and the hatched bands the statistical and systematic uncertainties on data added in quadrature.	281
H.3	Comparisons between measured cross section for results at 7 (left) [115] and 13 TeV (right) as a function of ΔR_{bb} and $M(b, b)$ for di b -jets. The data are compared to the predictions by Sherpa and Alpgen. The error bars indicate the statistical uncertainty, and the hatched bands the statistical and systematic uncertainties on data added in quadrature.	282
J.1	Energy calibration curves and their differences at different pile-up levels for $ \eta < 0.2$	291
J.2	Jet response after the calibration at different pile-up levels for $ \eta < 0.2$	291

J.3	Energy and p_T responses after calibration for collisions at 14 TeV and $\langle\mu\rangle = 200$, as function of pseudorapidity. Top: R=0.3; Bottom: R=0.4	292
K.1	Distribution of corrJVF (a) and R_{p_T} (b) for pile-up (PU) and hard-scatter (HS) jets with $20 < p_T < 30$ GeV [139].	296
K.2	(a) Fake rate from pile-up jets versus hard-scatter jet efficiency curves for JVF, corrJVF, R_{p_T} , and JVT in a sample of jets with $20 < p_T < 50$ GeV and $ \eta < 2.4$ in simulated dijet events. (b) The fake rate from pile-up jets versus hard-scatter jet efficiency curves for JVT separating jets initiated by gluons <i>light</i> - and <i>b</i> -quarks (<i>light</i> = <i>u</i> , <i>d</i> and <i>s</i>) [139].	297

List of Tables

2.1	Energy resolution for each liquid argon component.	32
3.1	Clusterization of towers for electrons and photons using the electromagnetic calorimeter in ATLAS.	47
5.1	Simulated di-jet samples and their cross sections.	81
5.2	Kinematic selections for MC-JES calibration at 8 TeV.	82
5.3	R-scan recommendations for MC-JES calibration (LCW+JES scale) at 8 TeV.	84
6.1	Definition of the phase space for the fiducial cross section for $Z + b$ -jet measurements at 7 TeV in ATLAS.	97
6.2	The measurement and theory predictions for the inclusive b -jet cross-sections in ATLAS for collisions at 7 TeV. The statistical uncertainty is quoted first in each case. The second uncertainty is either the total systematic uncertainty (data), the sum in quadrature of all theory uncertainties (MCFM), or the scale uncertainty (aMC@NLO) [115].	98
6.3	Definition of the phase space for the fiducial cross section for $Z + b$ -jet measurements at 7 TeV in CMS.	100
6.4	The measurement and theory predictions for the inclusive b -jet cross-sections in CMS for collisions at 7 TeV. The statistical uncertainty is quoted first in each case. The expectations from MCFM, aMC@NLO and MADGRAPH5include uncertainties due to scale variations [124].	100
6.5	Definition of the phase space for the fiducial cross section for $Z + b$ -jet measurements at 8 TeV in CMS.	101
6.6	The measurement and theory prediction for the inclusive b -jet cross-sections in CMS for collisions at 8 TeV (MADGRAPH5 is under the 5FNS). The statistical uncertainty is quoted first in each case [125].	101

7.1	Simulated Z +jets Sherpa 2.1 samples and their cross sections. These samples are sliced according to the Z p_T and combinations of filter and vetoes.	104
7.2	Simulated Z +jets Sherpa 2.2 samples and their cross sections. These samples are sliced according to the max Z p_T and max H_T , and combinations of filter and vetoes [128]. H_T is given by the scalar p_T sum of all parton-level jets with $p_T > 20$ GeV. . .	105
7.3	Simulated Z +jets Alpgen samples and their cross sections before and after the HFOR procedure. These samples are sliced according to the number of partons. The HFOR rejection factor is also shown. The full signal sample comprises all samples from the table after the HFOR.	106
7.4	Cross sections, f_{gen} and k -factors for the reference generators, including the NNLO computation for the inclusive $Z(\rightarrow \ell\ell)$ cross section. For both Sherpa 2.1 and Sherpa 2.2 is used the same f_{gen} and k -factors. Alpgen with HFOR.	107
7.5	Definition of the phase space for the fiducial cross section for the $Z + b$ -jet (truth level studies).	107
7.6	Theory predictions for the fiducial inclusive jet cross-sections in proton-proton collisions at 13 TeV (normalized to NNLO). Fiducial cuts are presented in Tab. 7.5. Only the statistical uncertainty is quoted.	113
7.7	Theory predictions for the fiducial inclusive b -jet cross-sections in proton-proton collisions at 13 TeV (normalized to NNLO). Fiducial cuts are presented in Tab. 7.5. Only the statistical uncertainty is quoted.	113
7.8	Peak delivered instantaneous luminosity ($\mathcal{L}_{\text{peak}}$), the mean number of proton-proton interactions per bunch crossing (μ) and the integrated luminosity ($\mathcal{L}^{\text{int}}$) by year on measurements at 13 TeV [41].	117
7.9	Definition of the phase space for the fiducial cross section for $Z + b$ -jet measurements at 13 TeV.	119
7.10	Detector-level yields, including yields predicted by Alpgen and Sherpa for each analysis selection. The normalized backgrounds are also shown.	122
7.11	Definition of the phase space for the fiducial cross section for the $Z + b$ -jet at truth-level.	129
7.12	JES systematic components.	140
7.13	B-Tagging systematic components.	141
7.14	Electron systematic components.	142

7.15	Summary of the systematic uncertainties determined for the cross-section measurements of the $Z + 1b$ -jet and $Z + 2b$ -jets final states.	144
7.16	Unfolding data and predictions results for the exclusive and inclusive b -jet cross-sections for analysis at $\sqrt{s}=13$ TeV. In data, the statistical uncertainty is quoted first and second the systematic. For predictions, only the statistical uncertainty is quoted. .	149
7.17	Unfolded data and predictions results for the inclusive b -jet cross-sections for analysis at 7 and 13 TeV. At 7 TeV is shown the combined analysis, considering electron e muons. In data, the statistical uncertainty is quoted first and second the systematic. For predictions, only the statistical uncertainty is quoted. .	156
8.1	Granularity comparison ($\Delta\eta \times \Delta\phi$) for the elementary cell, the current trigger tower and the proposed supercell on the LAr EM barrel calorimeter.	161
8.2	Number of events and cross sections of the simulated samples. .	169
8.3	Minimum values for $R_{\eta,2}$, $R_{\eta,1}$ and $R_{\eta,1+2}$ and their dependence with the efficiency and rates.	177
8.4	Efficiency and rate for different values of E_T^{min} and $R_{\eta,1+2}^{min}$	178
8.5	Maximum values for $W_{\eta,1}$, $W_{\eta,2}$ and f_3 , and their dependence with the efficiency and rates.	181
8.6	Efficiency and rate for different values of E_T^{min} and $W_{\eta,2}^{max}$	182
8.7	Efficiency and rate for different values of E_T^{min} and $W_{\eta,2}^{max}$	188
8.8	Number of events after trigger selections on CaloCells and LTDB.	195
E.1	Theory predictions for the integrated inclusive b -jet cross-sections in proton-proton collisions at 13 TeV (normalized to NNLO), $Z \rightarrow \ell\ell$ represents the combined measurement. Only the statistical uncertainty is quoted.	235
I.1	Efficiency and rate for different values of E_T^{min} and R_{η}^{min}	283
I.2	Efficiency and rate for different values of E_T^{min} and $R_{\eta,1}^{min}$	284
I.3	Efficiency and rate for different values of E_T^{min} and $W_{\eta,1}^{max}$	284
I.4	Efficiency and rate for different values of E_T^{min} and f_3^{max}	285
I.5	Efficiency and rate for different values of E_T^{min} and R^{max}	285
I.6	Efficiency and rate for different values of E_T^{min} and R_{seed}^{min}	286
I.7	Efficiency and rate for different values of E_T^{min} and R_{2ndhot}^{max}	286
I.8	Efficiency and rate for different values of E_T^{min} and $W_{\eta,2}^{max}$, with $R_{\eta,1+2}^{min}=0.777$	287

I.9	Efficiency and rate for different values of E_T^{min} and $W_{\eta,2}^{max}$, with $R_{\eta,1+2}^{min}=0.898$	287
I.10	Efficiency and rate for different values of E_T^{min} and $W_{\eta,2}^{max}$, with $R_{\eta,1+2}^{min}=0.918$	288
J.1	Simulated di-jet samples and their cross sections.	290
J.2	Kinematic selections for MC-JES calibration at 14 TeV under high luminosity conditions.	290

Contents

1	Introduction	1
1.1	Basic concepts	2
1.1.1	The Standard Model	2
1.1.2	Proton-Proton collisions	6
1.2	Proton-Proton colliders	9
1.2.1	Circular colliders	10
1.2.2	Luminosity and pile-up	11
1.3	Particle detection	16
1.3.1	Transverse energy and missing transverse energy (MET)	18
1.3.2	Jets	18
1.3.3	Bosons	20
1.3.4	$Z + b$ -jets	21
1.4	The LHC at high luminosity	22
1.4.1	The high luminosity program	23
1.4.2	Physics searches at high luminosity	24
2	The ATLAS detector	27
2.1	The Inner Detector	28
2.2	The electromagnetic calorimeter	31
2.3	The trigger system	35
2.4	The calorimeter trigger in ATLAS	37
2.4.1	Trigger for electrons and photons	38
2.4.2	Trigger for jets	42
2.4.3	Rejection of jets with low transverse moment at high luminosity	43
3	Measurement and Calibration of Electrons in ATLAS	45
3.1	Electron identification	45
3.2	Jet discrimination	47

3.2.1	Loose selection	47
3.2.2	Medium selection	48
3.2.3	Tight selection	50
3.3	Calibration	51
3.4	Results on MC calibration	52
4	Measurement and Calibration of Jets in ATLAS	55
4.1	Cone jet definition	55
4.1.1	The Snowmass accord and the cone algorithms	55
4.1.2	Radiated gluons and cone algorithms	57
4.2	Sequential clustering algorithms	60
4.2.1	Topoclusters	60
4.2.2	Generalization	61
4.2.3	Comparison between algorithms	62
4.2.4	The influence of the parameter R_0	64
4.3	Identification of jets in ATLAS	65
4.4	Calibration: Jet Energy Scale	67
4.5	Pile-up suppression	68
4.5.1	Area subtraction	68
4.5.2	Jet Vertex Fraction (JVF)	69
4.6	Results at 7 TeV	71
5	Jet Calibration Studies	73
5.1	Overview	73
5.2	MC-JES calibration procedure	75
5.2.1	Event and Jet Selection	75
5.2.2	Calibration Curves	77
5.3	MC-JES at 8 TeV	81
5.3.1	EM scale calibration	82
5.3.2	LCW scale calibration: R-scan	83
6	Production of Z Bosons Associated with b-jets in ATLAS	87
6.1	Theoretical fundamentals	87
6.2	Cross section predictions	89
6.3	B-jet selection	90
6.3.1	B-jet selection in simulations	90
6.3.2	B-jet selection in data	92

6.4	Main backgrounds on experimental measurement	95
6.5	ATLAS results at 7 TeV	97
6.6	CMS results on $Z + b$ -jets at 7 and 8 TeV	100
7	Production of $Z + b$-jets at 13 TeV	103
7.1	Signal sample validations	103
7.1.1	MC samples	104
7.1.2	HFOR validations for Alpgen samples	108
7.1.3	Generator comparisons and predictions for signal at 13 TeV	110
7.2	13 TeV data analysis	117
7.2.1	Selections for experimental analysis	117
7.2.2	Simulated event samples	119
7.2.3	Backgrounds	120
7.2.4	Monte Carlo - Data comparison	121
7.3	The unfolding procedure: Correction to particle-level	129
7.3.1	Response matrix	130
7.3.2	Reconstruction efficiency	131
7.3.3	Fake correction	132
7.3.4	Closure test	133
7.3.5	Unfolding test	133
7.4	Systematic uncertainties	140
7.4.1	Jet energy scale	140
7.4.2	B-tagging systematic: efficiency and mistag rates	141
7.4.3	Electron systematic: efficiency, energy scale and resolution	142
7.5	Results for $Z + b$ -jets analysis at 13 TeV	149
8	Measurement of Electrons with Supercells	159
8.1	Electron reconstruction using supercells	160
8.1.1	The new segmentation for triggers	160
8.1.2	MC samples and selections	161
8.1.3	Electron reconstruction	163
8.1.4	Comparison with truth	164
8.1.5	Super-electrons and pile-up suppression	166
8.2	Super-electrons analysis	168
8.2.1	MC Samples and selections	168
8.2.2	Comparisons with the first level trigger (L1) and off-line electrons	169

8.2.3	Rates and efficiencies	171
8.3	Shower shape discriminants	174
8.3.1	Rapidity spreads ($R_{\eta,2}$, $R_{\eta,1}$ and $R_{\eta,1+2}$)	174
8.3.2	Lateral shower widths ($W_{\eta,1}$ and $W_{\eta,2}$) and f_3	178
8.3.3	Other Discriminants	183
8.3.4	Correlation between discriminants	186
8.4	Implementation of the electromagnetic trigger system at high luminosity	190
8.4.1	Current system	190
8.4.2	System at high luminosity	190
8.4.3	Prototype of the new digital system of highly segmented trigger (LTDB)	192
Conclusions		199
A Kinematics on Boson Measurements		203
B The Simulation Process in ATLAS		205
B.1	General overview	205
B.2	Monte Carlo event generator	207
C Matrix Element Calculation		213
C.1	Z production at LO	213
C.2	$Z + 1$ -jet production at LO	218
D Heavy Flavor Overlap Removal		223
E Sample Validations for $Z + b$-jets at 13 TeV		235
F Unfolding Plots for $Z + b$-jets Production at 13 TeV		245
F.1	Response matrices	246
F.2	Reconstruction efficiency	257
F.3	Fake correction	263
F.4	Closure tests	269
G $Z + b$-jets Results and Comparisons with Z+jets Analysis at 13 TeV		275
H $Z + b$-jets Results and Comparisons with Analysis at 7 TeV		279

I	Efficiency and Rate Tables for Discriminants on SuperCell Studies	283
J	Jet Calibration at High Luminosity	289
J.1	Samples and selections	289
J.2	MC-JES at different pile-up levels	290
J.3	Improvements at high luminosity	291
K	Pile-up Suppression at High Luminosity	295
K.1	corrJVF and R_{p_T}	295
K.2	JVT	296

Chapter 1

Introduction

The proton-proton collision physics program of the two main experiments of the Large Hadron Collider (LHC) [1], the ATLAS [2] and CMS [3], emphasizes the search for signals that reveal the origin of the electroweak symmetry breaking, whose prediction by the Standard Model (SM) [4] involves the existence of the Higgs boson [5], confirmed in 2012 in the denominated Run 1 [6, 7]. However, it will be at the end of Run 2 (end of 2018) that it is expected to reach the nominal center-of-mass energy of the project, 14 TeV in proton-proton collisions, thus concluding the first phase of operation of the LHC, where it is anticipated to accumulate an integrated luminosity above 100 fb^{-1} . With the partial collected data taken during 2015 and 2016, it is already possible to improve the measurement of Z boson production in association with b -jets ($Z + b$ -jets) in proton-proton collisions, a key measurement in the LHC physics program and one of the main results from this thesis.

In the coming Run 3 (2021-2023), the upgrade of several ATLAS subsystems will allow it to operate at an instantaneous luminosity of up to $3 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$, where it is expected to accumulate more than 300 fb^{-1} of integrated luminosity. It is in the preparation for this run that will be implemented a new event selection system based on supercells [8], which will be responsible for the upgrade of the electron trigger system. The high luminosity and this new system will allow precision measurements in SM physics, a reduced systematic in measurements involving the production of the Higgs and more sensitivity in searches for new physics beyond the Standard Model (BSM).

This first chapter presents the basic concepts involving proton collisions at high energies, a summary of proton-proton colliders and an introduction to the program and operation of the LHC at high luminosity as well as the motivation for the studies in this thesis. Chapter 2 involves details of the ATLAS detector including its trigger system. Chapters 3 and 4 present the reconstruction and calibration methods of electrons and jets in the current ATLAS scheme, presenting studies on the Monte Carlo jet calibration for collisions at 8 TeV in chapter 5. Chapter 6 presents the motivation for the study of $Z + b$ -jets. The last two chapters present results of the main studies presented in this thesis: the results for the $Z + b$ -jets analysis at 13 TeV in Chapter 7 and the studies of the new event selection system at high luminosity in chapter 8, where is included a formal introduction to the concept of supercells with the first results in the reconstruction of electrons using this new concept.

1.1 Basic concepts

1.1.1 The Standard Model

The Standard Model of particle physics describes the interactions of three of the four fundamental forces of nature: the strong, weak and electromagnetic. These interactions occur through two types of particles: fermions, which are constituents of matter; and bosons, which are mediators of the interactions:

- Electromagnetic interactions: described by the quantum electrodynamics (QED), with the photon (γ) as mediator.
- Strong interactions: described by the quantum chromodynamics (QCD), with gluons (g) as mediators.
- Weak interactions: included in the electroweak theory, can be described by three mediators: W^+ , W^- and Z bosons.

These interactions are governed by *symmetry* [9], which reflects the *invariant* properties of equations to describe a physical phenomenon. Then, each symmetry is associated with a conserved quantity, such as energy, momentum or charge. The bosons are required to be massless in *gauge* invariant theories (QED and QCD), but W^+ , W^- and Z are massive bosons. The existence of

massive bosons can be explained by the Higgs mechanism, described as a case of spontaneous symmetry breaking [5, 10].

Fig. 1.1 presents the elementary particles of the SM with the three generations of fermions (quarks and leptons), mediator bosons (g , γ , W and Z) and the Higgs boson (H), responsible of giving mass to the particles (Higgs field). Some properties, as mass, charge and spin, are shown in the same figure.

Figure 1.1: Elementary particles in the Standard Model [11].

In the Standard Model, the electromagnetic and weak interactions can be described by just one theory, the electroweak theory, so the gauge symmetry which describes the SM is:

$$G_{SM} = SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \quad (1.1)$$

where $SU(3)_C$ is the gauge symmetry of the QCD and $SU(2)_L \otimes U(1)_Y$ the gauge symmetry of the electroweak theory ($U(1)_{EM} \subset SU(2)_L \otimes U(1)_Y$).

- $SU(3)_C$ is governed by 8 gauge fields A_μ^a , $a = 1, 2, \dots, 8$ (fields associated to gluons).
- $SU(2)_L \otimes U(1)_Y$ is governed by 4 gauge fields W_μ^1 , W_μ^2 , W_μ^3 and B_μ , whose

mixing describes the physical states: W_μ^+ , W_μ^- , Z_μ^0 and A_μ (fields associated to W^+ , W^- , Z and photon, respectively). The relation between fields is expressed by:

$$\begin{aligned} W_\mu^\pm &= \sqrt{\frac{1}{2}}(W_\mu^1 \mp W_\mu^2) \\ Z_\mu &= -B_\mu \sin(\theta_W) + W_\mu^3 \cos(\theta_W) \\ A_\mu &= B_\mu \cos(\theta_W) + W_\mu^3 \sin(\theta_W) \end{aligned} \quad (1.2)$$

where θ_W is called the Weinberg angle or weak mixing angle. This angle varies as a function of the momentum transfer (Q) and is a free parameter used on prediction of the electroweak theory [12]. Recently measurements at LHC presented a value of $\sin^2\theta_W = 0.2287 \pm 0.0020$ (stat.) ± 0.0025 (syst.) [13], which corresponds to $\theta_W \sim 30^\circ$. The masses of bosons W and Z are related to this angle by:

$$m_Z = \frac{m_W}{\cos\theta_W} \quad (1.3)$$

Recent mass values for those bosons are: $m_W = 80.385 \pm 0.015$ GeV and $m_Z = 91.1876 \pm 0.0021$ GeV [14].

Regarding the Higgs boson, the coupling with respect to light fermions is very small and difficult to detect experimentally, what can be seen comparing branching ratios, for example:

$$\frac{\Gamma(H \rightarrow ee)}{\Gamma(H \rightarrow WW)} \propto \frac{m_e^2}{4m_W^4} \approx 1.5 \times 10^{-21} \quad (1.4)$$

where $\Gamma(H \rightarrow ee)$ and $\Gamma(H \rightarrow WW)$ are the branching ratios for Higgs decays into two electrons and two W 's, respectively. For that reason, it is more probably found the Higgs on channels $H \rightarrow WW$ or $H \rightarrow ZZ$ than $H \rightarrow ee$.

With the passing of the time, each particle predicted by the SM was discovered (being the Higgs boson the last of them). Fig. 1.2 compares cross section measurements with the ATLAS detector and predictions provided by the Standard Model, where we can observe a very good agreement between data and theory. There are, however, still room for improvements, both on the theoretical predictions and experimental measurements, specially on process

with a low cross section (for example on Higgs and $t\bar{t} + Z/W$ productions).

Figure 1.2: *Standard Model cross section measurements in ATLAS (status: May 2017) [15].*

Even though the SM can explain a large number of observed phenomena in the subatomic physics, there are still many issues not addressed by the theory, as for example:

- Gravitational force: The SM only includes three fundamental forces but not the gravitational, since the gravitational force is not re-normalizable under the concept of quantum field theory [16].
- Mass of neutrinos: Experimentally, it was observed the neutrino oscillation [17], which led to the conclusion that neutrinos would be massive particles, contradicting the SM predictions.
- Dark matter and dark energy: Estimations indicate that the SM is able to explain only 5% of the energy and matter present in the universe. The rest would be comprised by dark matter and dark energy [18].

One theory that tries to solve these issues is the Supersymmetry (SUSY) [19], but until now, none of their predicted particles has been observed.

In this scenario, the LHC is not only a discovery machine, but also a precision measurement system where many physics processes that include vector bosons in the final state like inclusive and differential production of vector gauge bosons, Z and W mass measurements, vector boson fusion and the production of vector bosons associated with jets (light and heavy) are being explored both experimentally and theoretically, as many measurements are now becoming more precise than the theory predictions, and hence pushing the interpretation of the data and the process modeling to beyond the state of the art. The same is true for the experimental methods and detectors, where the increase of the luminosity results in rates higher than the original design can cope, challenging the best known signal processing and reconstruction techniques and event selection systems.

1.1.2 Proton-Proton collisions

There is a substantial interest in inelastic collisions, since they involve the production of new particles, but the more interesting processes have a very low probability to occur (Fig. 1.2). In general, the cross section increases with the collision energy of particles, and experimentally, one possible way to obtain high energy particles is by proton-proton collisions.

Fig. 1.3.a presents a scheme of a proton-proton collision, showing the constituents of each proton: the valence quarks, gluons mediators and sea quarks.

Figure 1.3: (a) Scheme of a proton-proton collision showing their constituents. (b) proton-proton collision according to the partons model.

The four-momentum of a particle can be expressed by:

$$p \equiv (p_0, p_x, p_y, p_z) = (E, \mathbf{p}) = (p_0, \mathbf{p}_T, p_z)$$

where $p_0 = E$ is the energy of the particle, $\mathbf{p}$ its momentum; and $\mathbf{p}_T$ and p_z its transverse and longitudinal momentum, respectively.

The collision energy, E_{coll} , for a collision of two protons, A and B , is expressed by:

$$E_{coll} = \sqrt{s} \equiv |p_A + p_B| \quad (1.5)$$

where s is one of the Mandelstam's variables; p_A and p_B are the four-momentum of A and B , respectively.

In case of head-on collisions of protons (as in the LHC), the momenta are opposites ($\mathbf{p}_A = -\mathbf{p}_B$), so, the collision energy is only expressed by the sum of their energies:

$$E_{coll} = E_A + E_B \quad (1.6)$$

The simplest model of proton collision, could be explained by parton collisions [20]. These partons (quarks or gluons) are considered as punctual particles and carry a fraction x of the proton energy on the collision, as is shown in Fig. 1.3.b. Considering the two protons traveling along the z axis

and in opposite directions, the four-moments of their partons can be expressed as:

$$\begin{aligned} p_a &= x_A \frac{s}{2} (1, 0, 0, 1) \\ p_b &= x_B \frac{s}{2} (1, 0, 0, -1) \end{aligned} \quad (1.7)$$

where a and b are partons from A and B , respectively. Thus, the invariant mass, m_{ab} , can be calculated from:

$$m_{ab}^2 = (p_a + p_b)^2 = x_A x_B s \quad (1.8)$$

From Eq. 1.8, it is possible to estimate x_A and x_B knowing the produced particles, since the invariant mass of particles produced should be the same. But, this is true only if protons were formed by punctual particles. Experimentally, it has been shown that the distribution of partons in the proton follows a probability distribution, which includes valence quarks and sea quarks. That distribution is known as parton distribution function (PDF) [22], commonly found in tables as $xf(x, Q^2)$. The parton density function, $f_i(x, Q^2)$, measures the probability, P , of finding the parton i with a moment fraction x inside the proton:

$$f_i(x, Q^2) \equiv \frac{dP_i(x, Q^2)}{dx} \quad (1.9)$$

The parton density function will depend on the square of the energy scale, $Q^2 = -(p_a - p_b)^2$. The normalization of those distribution must satisfy:

$$\sum_i \int x f_i(x, Q^2) dx = 1 \quad (1.10)$$

Since the PDF reflects the internal structure of the proton, theoretical predictions use the PDF in their calculations, and they need to represent better the collected data. Thus requiring constant updates, taking into account the understanding of how the proton structure evolves. In general, the PDF constructions are based on information collected by deep-inelastic scattering (DIP), vector boson production (VBP) and single-jet production. The constructions start from a parametrization of the non-perturbative PDF's at a low scale, such as fitting on DIS data produced by the HERA experiments

[23]. The resulting PDF's depend on the choice of the input data, the order of the perturbative QCD calculation, the treatment of heavy quarks, the correlation between α_s and the treatment of the uncertainties [22].

Fig. 1.4 presents examples of PDF's, produced by the MSTW [24] group, from protons at two different energy scales. In these examples one can see that the probability of valence quarks (uud) to have a large fraction of momentum decreases with the increase of the energy scale.

Figure 1.4: Example of the parton distribution functions, $xf(x, Q^2)$, inside the proton at two different energy scales on function of the moment fraction, x [24].

Along this thesis other PDF's sets will be mentioned, such as CTEQ [25] and NNPDF [26], as they represent a crucial information for simulations.

1.2 Proton-Proton colliders

Major discoveries in high energy physics have been possible by hadron and electron colliders. For now, the best way to achieve high energies is through circular colliders, but in these colliders, accelerated electrons tend to lose more energy than protons ($E_{rad} \propto v^4$) due to emission of synchrotron radiation [27], since the electron mass is ~ 2000 times lower than the proton mass. Although electron collisions are cleaner than protons, as for electrons there are no underlying events¹, to produce electron collision efficiently, the collider itself

¹Effect associated to partons in the proton that not participate of the hard collision.

should have very large dimensions, something impossible to achieve with the current technology. Then, considering the facts presented before, the proton-proton colliders are the best choice for studies at high energy. In this section we will present the main proton-proton circular colliders and will be described with more details the concepts of luminosity and pile-up.

1.2.1 Circular colliders

One of the first colliders of protons was the *Super Proton Synchrotron* (SPS) [28], originally designed to achieve an energy of $\sqrt{s} = 300$ GeV but finally constructed for a capacity of $\sqrt{s} = 400$ GeV (energy reached in 1976). It is a circular collider with 6.9 km of circumference belonging to CERN and located on the France-Swiss border. During 1981-1984, the SPS operated with proton-antiproton collisions, resulting in the discovery of W [29, 30] and Z [31, 32] bosons. Currently the SPS is used as final injector of protons for the Large Hadron Collider (LHC), accelerating protons from 26 GeV to 450 GeV.

The Tevatron [33] is another proton collider, considered the second most powerful particle accelerator, also circular in shape, with a circumference of 6.86 km and belonging to Fermilab in the United States, reaching the energies of $\sqrt{s} = 1.8$ TeV (1986) in Run 1 and $\sqrt{s} = 1.96$ TeV (2001) in Run 2. The Tevatron has confirmed the existence of many subatomic particles, and in 1995 the CDF (*Collider Detector at Fermilab*) [34] and DØ (*the DZero experiment*) [35] collaborations announced the discovery of the *top quark* [36, 37], the last fundamental fermion predicted by the Standard Model. The Tevatron finished its operations in September 2011.

Currently, we have the LHC [1], the most powerful and largest collider ever built. The LHC is a circular collider belonging to CERN, located on the French-Swiss border, with a circumference of 27 km. In the case of proton-proton collisions, the LHC reached the energies of $\sqrt{s} = 7$ and 8 TeV in Run 1 (2010-2012). Currently it is operating at $\sqrt{s} = 13$ TeV (Run 2), and is expected to reach the designed energy of $\sqrt{s} = 14$ TeV still in this run.

Fig. 1.5 presents a diagram of the locations of the four major LHC detectors: ATLAS, CMS, ALICE and LHCb. The LHC has a very comprehensive experimental program, covering aspects of SM measurements, BSM physics (like dark energy, dark matter and supersymmetric particle searches) and

Figure 1.5: Diagram with the location of the four major LHC detectors (ATLAS, ALICE, CMS and LHCb), which are located between 50 and 150 m underground [38].

Heavy ion physics (exploring the extreme temperature and pressure regimes of the matter as lead ions collide at relativistic energy). So far we have not detected physical signals beyond the Standard Model, but the search will continue, expecting to find signals of them with the increase of energy collision (increasing the cross section), luminosity (increasing the number of events) and improving the sensitivity of the system detection (upgrade of the detectors).

1.2.2 Luminosity and pile-up

An important parameter in high energy collisions is the luminosity (or instantaneous luminosity), $\mathcal{L}$, defined as:

$$\mathcal{L} \equiv \frac{1}{\sigma_{pp}^{inel}} \frac{dN_{pp}^{coll}}{dt} \quad (1.11)$$

where σ_{pp}^{inel} represents the inelastic proton-proton cross section and dN^{coll}/dt the number of collisions produced by unit of time.

Knowing the luminosity and the production rate for a certain channel

$pp \rightarrow X$ ($dN^{pp \rightarrow X}/dt$) it is possible to calculate its cross section:

$$\sigma^{pp \rightarrow X} = \frac{\int dN^{pp \rightarrow X}}{\int \mathcal{L} dt} \quad (1.12)$$

Commonly, the term $\int \mathcal{L} dt$ is denominated the *integrated luminosity*, $\mathcal{L}^{int}$:

$$\mathcal{L}^{int} \equiv \int \mathcal{L} dt \quad (1.13)$$

$\mathcal{L}^{int}$ is used to quantify the collected data.

Experimentally, each collision of proton beams² can be considered an event and in each event is possible to find multiple collisions. This is formally known as *in-time* pile-up contribution. Other collisions before or after can also affect the interaction of interest. These additional interactions depend strongly on the detector technology and the time window to process signals, and are known as *out-of-time* pile-up. Eventually, other pile-up contributions, associated to beam interaction with residual gas of thermal neutrons and photons from the cavern, must be also taken into account [40]. Commonly, the pile-up is defined as the number of inelastic proton-proton collisions per bunch crossing (just *in-time* pile-up) and represented by the letter μ . This will be the definition used in this work.

Fig. 1.6 presents an event with a low pile-up ($\mu = 4$) measured in ATLAS, where, through track reconstruction, the particles produced can be "easily" identified.

On circular colliders, like the LHC, the rate of collisions are related with the revolution frequency, f_r :

$$\frac{dN_{pp}^{coll}}{dt} = \mu n_b f_r \quad (1.14)$$

where μ represents the average number of collision per bunch crossing and n_b the number of bunch of protons on collision. Then, with 1.11 and 1.14, the average number of collisions can be extracted knowing the luminosity through

²At the LHC, each beam on collision course has 2080 bunches with approx. 1.15×10^{11} protons [39].

Figure 1.6: Event display with a low pile-up, $\mu = 4$ at 7 TeV, measured in ATLAS (2010) [38]

the next relation:

$$\mu = \frac{\mathcal{L} \times \sigma_{pp}^{inel}}{n_b f_r} \quad (1.15)$$

In simulations, there are two methods to model the pile-up. The first one is to simulate each contribution individually through Monte Carlo (MC) samples: simulate the in-time, out-of-time and residual gas contributions. These MC contributions are added to the signal events of interest in order to emulate the pile-up environment. The other method is to use specially-collected data to model the background. This method has the advantage that the background is determined in a truly data-driven way, and the dependence on MC event generator modeling of particle spectra is eliminated [40], although the use of MC samples is useful to study events under luminosities not reached yet.

During the LHC Run 1, the peak number of interaction per bunch crossing reached a value of almost $\mu = 35$. During the starting of Run 2 (2015), the number was between 12 to almost 27, and increased from 20 to 40 during 2016, as can be seen in the Fig. 1.7. Currently, the collected data until June 2017 has reached a value of $\mu = 50$.

The peak luminosity delivered to ATLAS during proton-proton collisions at 13 TeV is shown in Fig. 1.8. During 2015 the luminosity reached a peak of

Figure 1.7: The maximum number of inelastic collisions per beam crossing (μ) for proton-proton collisions at 13 TeV during 2015 (a) and 2016 (b) [41].

$\mathcal{L} = 5 \times 10^{33} \text{ cm}^{-2}\text{s}^{-1}$; and during 2016 a peak of $\mathcal{L} = 13.8 \times 10^{33} \text{ cm}^{-2}\text{s}^{-1}$. The luminosity is determined using counting rates measured by the luminosity detectors, and is based on analysis of van-der-Meer beam-separation scans [42]. Currently, the luminosity reached a value of $\mathcal{L} = 15.8 \times 10^{33} \text{ cm}^{-2}\text{s}^{-1}$ (June 2017). It is important to note that the original design peak luminosity of the LHC (and projected for measurements in ATLAS) was of $\mathcal{L} = 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ [8].

Figure 1.8: The peak instantaneous luminosity delivered to ATLAS during stable beams for proton-proton collisions at 13 TeV for each LHC fill (good runs) as a function of time during 2015 (a) and 2016 (b) [41].

Finally, it should be noticed the difference between delivered, recorded and good for physics luminosity. The delivered luminosity accounts from the start of stable beams until that LHC requests ATLAS to put the detector in a safe stand-by mode to finalize the data acquisition. The recorded luminosity reflects the data acquisition inefficiency, as well as the inefficiency of the so-called *warm start*. This is the time when the stable beam flag is raised and the detector system is ready to start the data taking. The good for physics integrated luminosity requires well reconstructed physics objects (good data quality) [41].

Figure 1.9: Total integrated luminosity delivered and recorded during 2015 (a) and 2016 (b) [41]. For 2015 is also shown the good for physics integrated luminosity.

Fig. 1.9 shows the cumulative or total integrated luminosity delivered and recorded during 2015 and 2016. From 3.9 fb⁻¹ recorded during 2015, only 3.2 fb⁻¹ was selected as good for physics. During 2016 it was recorded 35.6 fb⁻¹ with 33.3 fb⁻¹ good for physics. Most of recent ATLAS analyses use 36.5 fb⁻¹ (2015+2016) of good for physics integrated luminosity, and this is the amount of data that was used in the $Z + b$ -jets measurements presented in this thesis.

1.3 Particle detection

Particle detectors have a general form of design as is shown in Fig. 1.10, where it is presented a transversal segmentation of the ATLAS detector. After the collision point, the first component is the tracker system, which is sensitive to charged particles with the help of a magnetic field, provided by the solenoid magnet in ATLAS. The tracker system is followed by the electromagnetic calorimeter, where the electrons and photons are stopped, eventually depositing all their energy. Protons and neutrons will be stopped in the hadronic calorimeter, the next component. Muons will be detected by the muon spectrometer, and neutrinos can not be detected since they do not interact with any detector, but their energy or p_T contribution can be estimated considering the missing transverse energy (Sec. 1.3.1).

Figure 1.10: *Transversal segmentation of the ATLAS detector, showing the interaction of particles with trackers, calorimeters and muon spectrometer [38].*

Other particles have a lifetime shorter than the previous mentioned and are detected through their decay products. Parton constituents (quarks and gluons), produced after the collision, can be detected as jets (Sec. 1.3.2).

Kinematic quantities of interest are: the energy (E), the transversal momentum (p_T) and the azimuthal angle (ϕ). The polar angle (θ) is not a Lorentz invariant; then, is defined the *rapidity* (y) as:

$$y \equiv \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right) \quad (1.16)$$

Fig. 1.11.a shows the three-dimensional representation of the vector momentum ($\mathbf{p}=(\mathbf{p}_T, p_z)$) for a particle produced and its angle with respect to the direction of the beams, θ .

Figure 1.11: (a) Representation on Cartesian coordinates for momentum indicating the longitudinal moment (with respect to the direction of the beams, represent for the particles A and B) and transverse (perpendicular to the direction of the beams); (b) Pseudorapidity as a function of the polar angle, θ .

At high energy physics the value of $E \approx |\mathbf{p}|$. Therefore, it is common to measure the angle of the detected particle with respect to the axis of the incident beam (θ). Knowing this angle, the variable *pseudorapidity* is defined as:

$$\eta \equiv \frac{1}{2} \ln \left(\frac{|\mathbf{p}| + p_z}{|\mathbf{p}| - p_z} \right) = -\ln[\tan(\theta/2)] \quad (1.17)$$

Fig. 1.11.b shows the graph that relates the angle θ with the pseudorapidity, η . Considering:

$$e^{\pm\eta} = \sqrt{\frac{|\mathbf{p}| \pm p_z}{|\mathbf{p}| \mp p_z}}$$

knowing that: $|\mathbf{p}|^2 = p_T^2 + p_z^2$ and $|\mathbf{p}|=E$, it can be proved that:

$$\begin{aligned} E &= p_T \cosh(\eta) \\ p_z &= p_T \sinh(\eta) \end{aligned} \tag{1.18}$$

with 1.18 and from Fig. 1.11.a can be extracted that:

$$\mathbf{p} = (p_T \cos\phi, p_T \sin\phi, p_T \sinh(\eta)) \tag{1.19}$$

1.3.1 Transverse energy and missing transverse energy (MET)

Since the energy of the particles produced at the LHC are larger than their masses, it is possible to approximate the magnitude of their momentum by their energies [43]. Knowing that, for each element i on the detector, is defined the corresponding vector of transverse energy ($\vec{E}_T^i \equiv E_i \hat{n}_T$) as the measured energy multiplied by the vector position of the unit element (with respect to the interaction point). The total transverse energy per event is defined by the sum over all elements: $\vec{E}_T \equiv \sum \vec{E}_T^i$.

Reminding the transverse momentum conservation, the missing transverse energy vector, $\vec{E}_T^{miss}$, is defined from $\vec{E}_T^{miss} + \vec{E}_T = 0$, as:

$$\vec{E}_T^{miss} = -\vec{E}_T$$

1.3.2 Jets

Jets are objects used to identify partons produced after hard collisions. Events with jets are abundant in proton-proton collisions and are present as background in many physics process (for example in Higgs measurements or studies beyond the Standard Model). Jets will play an important role along this work, as they will be:

- Subject of performance studies.
- Useful by their association with bosons, particularly, the Z boson in this work.
- Important on the upgrade studies, since represents the main source of background on pile-up studies.

Figure 1.12: (a) Proton-Proton collision and jets production. (b) Estimated energy fraction of particles that make up the simulated jets (truejets) [44].

Fig. 1.12.a shows a diagram of how jets are produced in a proton-proton collision and how they arrive into calorimeters, depositing their energies there. It is also shown the different types of jet definitions, starting with the jet at parton level (represented by quarks or gluons); the particle jet, which is defined merging particles coming from hadronizations and consequent decays; and finally the calorimeter jet, which is a reconstructed with the calorimeter information.

Experimentally, jets are detected as calorimeter jets. Parton and particle jets are available only in simulations. Particle jets will be referred as *truejets*. Fig. 1.12.b presents a graph with the estimated particle energy fractions that, on average, make-up a truejet. These jets are composed by stable particles and do not consider neutrinos or muons, since the neutrinos do not interact with any calorimeter and the muon energy deposited in a calorimeter is small.

A more extensive discussion of jets will be presented in Chap. 4, as preparation for the performance studies, physic analysis and upgrade in the ATLAS experiment.

1.3.3 Bosons

Leptons at high energy may represent a signal of boson production, as can be seen in Fig. 1.13, where it is shown the Feynman diagram for $Z \rightarrow ee$ and $W \rightarrow e\nu_e$ process. At the same time, W and Z bosons may indicate signals of other processes, for example in Higgs production (ex. $H \rightarrow ZZ \rightarrow \ell\ell\ell\ell$) or yet be present in production of new particles, for example in gauginos predicted by supersymmetry ($\tilde{\chi}_2^0 \rightarrow Z\tilde{\chi}_1^0$) [45].

Figure 1.13: Feynman diagrams for proton-proton collision and production of W and Z on the electron channel [46].

Fig. 1.14 presents the scheme of the experimental detection of W and Z (electron channel), with the electrons arriving into detectors and depositing their energy into calorimeters (red peaks). It is also shown the hadronic recoil, associated to jet production as well as some underlying events. In the case of W it is shown the neutrino passage, detected by MET.

Figure 1.14: Scheme of the experimental detection of W and Z on the electron channel [46].

The Z boson production (inclusive and associated with jets) on the electron channel ($Z \rightarrow ee$) will be further discussed in this work, both in analysis and in simulations. The electron reconstruction will be detailed in Ch. 3.

1.3.4 $Z + b$ -jets

With the increase of the energy collision, it is expected an abundant production of jets related to heavy-quarks, specially b and t . Top quark production is one of the major topics on the ATLAS experiment, but suffers from a large background contribution of b -jets [47], turning the accuracy on the b -jet measurement and modeling of great relevance in ATLAS. Currently, there is a great interest on the production of bosons W and Z associated with b -jets. In general the production of $W + b$ -jets suffer for background of $Z + b$ -jets process, but the measurement is more complex since it requires more precision in the W production, which is limited by the MET measurement [48]. As anticipated, the studies in this thesis will be based in Z production, hence we will discuss only the $Z + b$ -jets.

The production of a Z boson with at least one b -jet ($Z + \geq 1b$ -jet) in the detector acceptance can provide precision tests of perturbative QCD [49] and the production of a Z boson with a single b -jet ($Z + 1b$ -jet) provides information related to the b -quark inside of the proton [50]. The production of a Z boson in association with at least two b -jets ($Z + \geq 2b$ -jets) is of great interest since it represents one of the main background contribution on $t\bar{t}$ production [51]. In fact, $Z + \geq 2b$ -jets process also represent a large background contribution on $H(\rightarrow b\bar{b})$ studies [52] and even process beyond Standard Model, for example search for light top squark pair production [53].

Fig. 1.15 presents the latest results on single-boson production and as can be seen only results for $Z + b$ -jets at 7 TeV are available and show marginal agreement with predictions. The figure also shows other process of interest involving bosons and jets in ATLAS. Ch. 6 will present the main aspects of the $Z(\rightarrow ee) + b$ -jets analysis and Ch. 7 will show the analysis and the first results on analysis at 13 TeV obtained in the electron channel.

Figure 1.15: The data/theory ratio for several single-boson fiducial production cross section measurements. All theoretical expectations were calculated at NLO or higher (status: August 2016) [15]. In the green square are present the $Z + b$ -jets measurements.

1.4 The LHC at high luminosity

The pile-up increase represents a challenge for all LHC detectors, becoming a problem for the event selection and particles detection. Fig. 1.16 shows an event display from ATLAS with $\mu = 25$ for data taken during 2016. Compared with Fig. 1.6, it can be seen how the number of tracks is increased considerably, bringing limitations on the objects reconstruction.

One of the goals in this thesis is to show the importance of a new event

Figure 1.16: Event with $\mu = 25$ at 13 TeV, measured in ATLAS (2016) [38]

selection system, which will improve the electron reconstruction under high pile-up environment. The section will discuss the LHC program for the next years and the main physics searches expected to improve with the higher luminosity.

1.4.1 The high luminosity program

The current LHC program is presented in Fig. 1.17. After Run 1, there was the *first long shutdown* (LS1) [54], a period used for the adaptation of the accelerator complex and preparation for Run 2, which started in 2015 and will continue until the end of 2018, when it is expected to get an integrated luminosity above 100 fb^{-1} . It will be during the *second long shutdown* (LS2) [55] (2019-2020) that new updates on detectors should be implemented, in order to improve the detection accuracy and optimization of objects such as electrons and jets for Run 3 (2021-2023). In ATLAS, the main modifications will be in the trigger system and in the tracking. The muon system will be updated with new forward detectors to cope with the high rate and the electromagnetic calorimeter front end electronics will be updated to provide a higher granularity and digitized low level trigger signal.

Figure 1.17: The program for the next years according to the LHC commissioning [57] (January 2016).

After Run 3, it will take part the *Long Shutdown 3*, *LS3*, which involves the preparation for the LHC high luminosity program (HL-LHC) [56]. In the LS3 some parts of the detector will be replaced due to the damage caused by radiation, such as the ATLAS *Inner Detector* (responsible for the track reconstruction). In Run 3, it is expected to reach a value of $\mu = 80$ and for runs in the HL-LHC to reach values as high as $\mu = 200$, with an instantaneous luminosity of $5 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$, in order to achieve an integrated luminosity of 3000 fb^{-1} , making it possible to probe rare process and perform differential measurements in extreme phase spaces.

1.4.2 Physics searches at high luminosity

Higgs boson measurements

With the discovery of the Higgs, the study of its properties is a priority over the current data collection, as well as to improve the accuracy of the measured channels so far:

- $H \rightarrow \gamma\gamma$ [58]
- $H \rightarrow ZZ^* \rightarrow 4\ell$ [58]
- $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ [58]

- $H \rightarrow \tau^+\tau^-$ [59]

For example, the last two channels have some limitations in their studies, associated to background uncertainties, and an increase on the number of events is critical to the accuracy of measurements. In addition, it is expected that with 3000 fb^{-1} , the so-called rare channels will be accessible:

- $H \rightarrow \mu\mu$
- $t\bar{t}H$ with $H \rightarrow \gamma\gamma$ and $H \rightarrow \mu\mu$
- WH/ZH with $H \rightarrow \gamma\gamma$

In the current run is expected to improve the signal measurements of $H \rightarrow b\bar{b}$, showing the first signals of that with collected data during 2015 and 2016 [60]. Predictions indicate that this process represent around 58% of the Higgs decays but suffers of a large background sources, one of them is associated to $Z + b$ -jets. Evidently, the high luminosity will improve all these measurements and even a clear signal of $H \rightarrow b\bar{b}$ will be important for observation of the Higgs self coupling:

- $HH \rightarrow b\bar{b}W^+W^-$
- $HH \rightarrow b\bar{b}\gamma\gamma$

Search for new particles predicted in Supersymmetry

Supersymmetry (SUSY) is a symmetric extension of the Standard Model relating fermions and bosons. In these theories, each boson (fermion) from the SM has a fermion (boson) as supersymmetric partner or superpartner. Until now, collected data in Run 1 and preliminary Run 2 did not show signals of any supersymmetric particle. It is important to show the importance of the HL-LHC program in those studies, for example on the search of:

- Direct production of gauginos (higgsinos, winos and the bino) [45], these particles may be produced from decays of *squarks* ($\tilde{q}$) and *gluinos* ($\tilde{g}$), or directly from weak productions [61], but have a too lower cross-section (between 10^{-2} to 10 pb). The mass eigenstates formed from the linear superposition of gauginos are denominated charginos ($\tilde{\chi}_i^\pm$, $i = 1, 2$) and neutralinos ($\tilde{\chi}_j^0$, $j = 1, 2, 3, 4$).

- Direct production of squark top ($\tilde{t}$) [61], the production of the Higgs boson with mass $m_H = 125$ GeV should favor the production of squark top, lower than 1 TeV. The expected cross section for the production of a pair of squarks top at $\sqrt{s} = 14$ TeV is 10 fb for $m_{\tilde{t}} = 1$ TeV, lower than expected for gauginos.

It is thus clear that precision measurements of vector bosons associated to heavy quarks and the possibility to push the detector limits to operate under extreme pile-up condition are quite important topics for the LHC physics program that lies ahead over the next decade.

Chapter 2

The ATLAS detector

The physics program for the ATLAS detector [62] follows a plan for precision measurements in the standard model, such as the production of Z and W bosons (inclusive and associated to jets), measurements of bottom and top quarks and measurements in the Higgs sector. Searches for signals of physics beyond the SM are also an important part of the ATLAS experimental program. All these measurements and searches will benefit from higher luminosities, which will allow the use of channels and differential cross section measurements that today are limited by the low production cross section. This chapter will present an overview of the ATLAS detector and details of the important systems in this thesis, such as: the Inner Detector, the electromagnetic calorimeter and the trigger system of ATLAS.

The ATLAS is a detector of general purpose, suitable for an extensive research program in proton-proton and heavy ion collisions. The ATLAS consists of several subsystems (Fig. 2.1): The Inner Detector (formed by the Pixel Detector, SemiConductor Tracker and Transition Radiation Tracker), the electromagnetic and hadronic calorimeters (composed by the LAr, Tile and Forward Calorimeters) and the muon spectrometer.

Fig. 2.2 shows the coverage of the ATLAS detector subsystems as a function of the pseudorapidity. The Inner Detector covers a region of $|\eta| < 2.5$, the electromagnetic and hadronic calorimeters a region of $|\eta| < 3.2$, the forward calorimeters a region of $3.2 < |\eta| < 5.0$ and the muon spectrometer in $|\eta| < 2.5$. It is also presented the LUCID (LUMinosity Cerenkov Integrating Detector, $5.4 < |\eta| < 6.1$) used for luminosity measurements; and the ZDC (Zero Degree Calorimeter, $8.3 < |\eta|$) used on heavy ions studies [63].

Figure 2.1: *The ATLAS Detector: A Toroidal LHC ApparatuS* [62].

Figure 2.2: *Pseudorapidity coverage of the ATLAS detector subsystems* [62].

2.1 The Inner Detector

One of the main characteristics of ATLAS is its good resolution in the reconstruction of vertex position and the determination of transverse momentum (p_T), done with the Inner Detector (ID) [64, 65] through the reconstruction of tracks. The ID covers a pseudorapidity of $|\eta| < 2.5$ and has a complete azimuthal coverage (ϕ). The ID is composed by three sub-detectors: the PIXel

detector (PIX), the SemiConductor Tracker (SCT) and the Transition Radiation Tracker (TRT). Fig. 2.3 shows a detailed diagram of the ID showing its dimensions and sub-detectors.

Figure 2.3: Scheme of the Inner Detector and its sub-detectors [38].

Fig. 2.4 shows the track reconstruction for three particles (electron, proton, muon), illustrating the vertex and the passage of the particles by the sub-detectors of the Inner Detector. The reconstruction of vertex and tracks will be made with the standard reconstruction algorithms of ATLAS [66], which combines the information of all ID subsystems.

Figure 2.4: Scheme of the track reconstruction for three particles and their passing through the ID sub-detectors.

For Run 2, a fourth layer in the pixel detector was incorporated in order to recover the reconstruction efficiency of vertex and tracks [67], due to the radiation damage of the original system.

Along this thesis the track and vertex information will be referred several times, especially on studies with jets. The primary vertex (PV) will be the vertex associated to the hard process and is defined as the vertex with the highest $\sum p_{T,\text{track}}^2$ of their constituent tracks. In Fig. 2.5.a it is shown again the plot of a low pile-up event presented in Sec. 1.2.2, but for this case indicating that the pile-up number is the same number of primary vertices. The definition of the transversal (d_0) and longitudinal (z_0) impact parameters between a primary and a secondary vertex are shown in Fig. 2.5.b, as they are an important parameter in b -jet classification (b -tagging).

Figure 2.5: (a) Primary vertex reconstruction NPV=4 ($\mu = 4$). (b) Transverse (d_0) and longitudinal (z_0) impact parameters between a primary and a secondary vertex.

The typical vertex resolution is around $100 \mu\text{m}^1$, which can compromise the reconstruction efficiency at high luminosity, as is shown in Sec. 1.2.2. With the increase of pile-up (luminosity) the number of primary vertex (NPV) tends to be lower than pile-up (μ), although it is still directly proportional. Fig. 2.6 shows how the efficiency of the vertex reconstruction decreases with the increase of pile-up (decrease on the slope).

¹ID resolution: $\sigma/p_T \sim 3.8 \times 10^{-4} p_T [\text{GeV}] \oplus 0.015$

Figure 2.6: NPV vs μ at different levels of luminosity, for dijet simulations for collisions at 13 TeV (from samples used on jet calibration studies at high luminosity (App. J)).

2.2 The electromagnetic calorimeter

ATLAS uses liquid argon as active medium in its electromagnetic (EM) calorimeter [68], which is divided into two regions: A barrel (EMB, $|\eta| < 1.475$) and two end-caps (EMEC, $1.375 < |\eta| < 3.2$), as is shown in Fig. 2.7. The barrel has a length of 6.4 m, an internal radius of 2.8 m and an outer radius of 4.0 m, with a total weight of 104 tons. The figure also shows other liquid argon based calorimeters, which are the end-caps of the hadronic calorimeter (HEC, $1.5 < |\eta| < 3.2$) and the Forward Calorimeter (FCal, $3.1 < |\eta| < 4.9$). The energy resolution for each liquid argon component is presented in Table 2.1.

The EMB is segmented in two half-barrels centered on $z = 0$, covering a region of $0 < |\eta| < 1.475$. Each half-barrel is divided into 16 modules ($\Delta\phi = 22.5^\circ$) and have 1024 absorbers with accordion shape for each of them

Figure 2.7: *The Liquid Argon calorimeter of ATLAS [38].*

Component	Resolution
EM Barrel	$\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} \oplus 0.7\%$
EM End-Cap	$\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} \oplus 0.7\%$
HEC	$\frac{\sigma_E}{E} = \frac{50\%}{\sqrt{E}} \oplus 3\%$
FCal	$\frac{\sigma_E}{E} = \frac{100\%}{\sqrt{E}} \oplus 10\%$

Table 2.1: *Energy resolution for each liquid argon component.*

(in order to avoid discontinuities along the azimuth angle), and between them, the readout electrodes. The depth of each active module is increased following measurements of radiation length (X_0)², from $22X_0$ to $30X_0$ between $|\eta| = 0$ and $|\eta| = 0.8$ and $24X_0$ to $33X_0$ between $|\eta| = 0.8$ and $|\eta| = 1.3$. Fig. 2.8 presents details of the regions of the electromagnetic calorimeter, observing the accordion form on the absorbers.

The EMEC is presented as two end-caps on each side of the barrel, each end-cap has a width of 63 cm and a weight of 27 ton, with an inner radius of 330 mm and an external 2098 mm, covering $1.375 < |\eta| < 3.2$. The outer

²The radiation length X_0 , is defined as the distance over which an electron or positron ($E \gg 1$ GeV) loses, on average, 63.2% of its energy [69].

Figure 2.8: Details of the electromagnetic calorimeter, electromagnetic barrel (left) and one of the end-caps (EMEC) (right) [62].

wheel of each end-cap cover has 768 absorbers, interleaved with the readout electrodes, while each inner wheel has 256. The active depth increases $24X_0$ to $38X_0$ between $|\eta| = 1.475$ and $|\eta| = 2.5$ (outer wheel) and $26X_0$ to $36X_0$ between $|\eta| = 2.5$ and $|\eta| = 3.2$ (inner wheel).

Fig. 2.9 shows the segmentation of the electromagnetic barrel in its three main layers: the front (layer 1), the intermediate (layer 2) and back (layer 3); There is also an initial layer, the pre-sampler (layer 0). The pre-sampler is used to compensate the energy lost due to the presence of material before the calorimeter and it has no absorber. The front layer, highly segmented, is fundamental in the rejection of photons originated from the decay of π^0 , abundant in the LHC environment. The intermediate layer is the deepest and responsible for the sampling of most of the electromagnetic shower. Finally, the back layer, important in energy corrections, accounting for the longitudinal leaks.

Fig. 2.10.a shows a transversal segment of the EMB with a zoom of its components. The produced charges on the electromagnetic shower ionize the liquid and, under the influence of an electric field, are carried to the readout electrodes (Fig. 2.10.b). Thus, each calorimeter cell acts as an ionization chamber, and the produced current is proportional to the deposited energy in the cell.

For this work, it is important to understand the cell clusterization process on the EM calorimeter (Fig. 2.11). Cells are the minimum units of energy

Figure 2.9: Partial isometric view of the electromagnetic calorimeter, showing the transversal and longitudinal segmentation [62].

Figure 2.10: a) Transversal segment of the EMB detailing its components. b) Principle of ionization signal generation from the electromagnetic shower [70].

deposition in the calorimeter. In order to obtain a quick response for trigger purposes, the cells are grouped on towers, in general, at fixed intervals in η and ϕ . The *clusters* are groups around areas with regions of interest, can be more complex and are commonly grouped at high-level triggers or off-line processing.

Figure 2.11: *Cells clusterization on the electromagnetic calorimeter.*

Currently, the first level trigger of ATLAS consists on towers of $\eta \times \phi = 0.1 \times 0.1$, and it is planned to change that for a new configuration, such as $\eta \times \phi = 0.025 \times 0.1$ in the front and intermediate layers. This new configuration is called supercell and will be shown, in detail, in Ch. 8.2.

2.3 The trigger system

The event selection in ATLAS is made at 3 levels (Fig. 2.12). The lower level is the Level 1 (L1), where information from the calorimeter and muon system is processed by dedicated hardware. The L1 is able to operate at a maximum rate of 100 kHz from 40 MHz interaction rate, selecting the regions of interest (ROI) to be processed by the next levels. The next levels comprise the so called High Level Triggers (HLT): the second level (L2) and the Event Filter (EF). The L2 has access to the full granularity of ATLAS, and the event selection is done through the process of digital information, reducing the event rate at 3 kHz. The highest level, the EF, has access to the entire detector and processes the events that have passed through the L2 using the final data analysis algorithms, reducing the rate to 200 Hz. At L2 and EF, the process are based on analysis programs distributed in computer clusters (off-line processing).

Figure 2.12: Scheme of the 3 levels of trigger used in ATLAS.

The L1 trigger

The concepts behind the L1 trigger will be important for this thesis, since it is one of the system to be upgraded in the LS2. The L1 is mainly focused on signals that can identify electrons, photons, muons and isolated taus with high transverse momentum, as well as MET and jets. To do this quickly, the L1 uses the information of the high granularity of the electromagnetic and hadronic calorimeters, as well as the muon detector. Currently, the L1 is able to make this selection at a maximum rate of 100 kHz.

Fig. 2.13 shows the block diagram of L1, which uses three components to make the decisions, the first one is the *L1 Calorimeter Trigger* (L1Calo), which uses the deposited energy to identify objects with high p_T . The second one is the *L1 Muon Trigger* (L1Muon) which uses the information from the muon spectrometer, identifying the muon candidates at high p_T . The third component is the *Trigger Central Processor* (CTP), which implements the group of triggers to be used, these groups are based on logical combinations (coincidences and/or vetoes) coming from L1Calo and L1Muon. The CTP can

Figure 2.13: Block diagram of L1 trigger [71].

also reduce the rates (*pre-scaling*) with higher probability in order to make it more efficient in terms of the available bandwidth. The collected information will be classified in regions of interest (RoI) that will be used in the level 2 (L2) trigger. The trigger information is passed to the Data Acquisition system (DAQ) that channels the data from the detectors to the storage.

Information above the other trigger levels can be found in [72]. The electromagnetic trigger will be discussed in more detail in the next section, while information of L1Muon can be found in [73, 74].

2.4 The calorimeter trigger in ATLAS

This section will describe the trigger at level 1 for electrons and jets which is based on the calorimeter information.

Fig. 2.14 presents a block diagram for the calorimeter trigger. The information, collected by calorimeters, are grouped in towers of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$, with the electronic system responsible for this installed closer to the detector (Sec. 8.4.1). The signals are carried to the *Pre-Processor*, responsible for the timing alignment with the bunch crossing identification (BCID) [76] that in sequence will re-distribute the signals to the *Cluster Processor* and the

Figure 2.14: Block diagram for the calorimeter trigger [75].

Jet/Energy Processor:

- **The Cluster Processor:** Responsible for the identification of *Regions of Interest* (RoI) for candidates to the EM objects (electrons, photons and taus) and their threshold counting. The towers are segmented in $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$.
- **The Jet/Energy Processor:** Responsible for the identification of RoI for candidates to jets and their threshold counting. Here will be also processed the sum of transverse energy. The towers are segmented in $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$.

The information of the Cluster and Jet/Energy processors are carried to the *RoI Builder* and the *Central Trigger Processor* (CTP) (Sec. 2.3). The trigger information pass for the Data Acquisition system (DAQ), which distributes the data from the detectors the event builder and storage system.

2.4.1 Trigger for electrons and photons

This trigger identifies candidates for electrons and photons using information from calorimeters, classifying them according to their transverse energy (E_T) and isolations, by identifying the coordinates of the regions of interest and their classification to be used at level 2 (L2).

Isolation

The trigger must be able to select an electromagnetic object avoiding the overlap with an hadronic object. The isolation using the calorimeter information can be done by the shower distributions and their energy deposited into the calorimeters. Fig. 2.15 presents the showers produced by an electron and a pion, both at 50 GeV, through iron. The electron presents a shorter and more narrow shower than the pion, and it is expected the same difference comparing electrons and jets (composed mostly by hadrons).

Figure 2.15: Shower for an electron (left) and a pion (right), both at 50 GeV, through iron [77].

Algorithm

The informations for the algorithm are provided by a set of trigger towers with a granularity of 0.1×0.1 in $\Delta\eta \times \Delta\phi$. These towers are formed by analog sum of the calorimeter cells in a region of $|\eta| < 2.5$. The algorithm searches one electron or photon candidate and is based on the scheme shown in Fig. 2.16, where it is considered tower windows of 4×4 in the electromagnetic and

hadronic calorimeter, following these steps:

Figure 2.16: Scheme used in algorithm for the L1 trigger for electrons and photons identification [72].

Step 1: search for clusters in towers of 2×2 in the EM calorimeter, which must have a minimum transverse energy (threshold in the order of 3 GeV or more). This threshold will be determined according to efficiency studies³. These are clusters that will be considered as candidates for RoI.

Step 2: Inside the 2×2 cluster, 4 clusters of 2×1 (1×2) are used to measure the E_T from the electromagnetic shower, but only the most energetic cluster will be considered.

Step 3: The ring of 12 towers, surrounding the RoI, is used to determine the electromagnetic isolation. Then, the RoI will be an isolated candidate if the E_T sum in the ring is lower than the electromagnetic isolation threshold.

Step 4: The cluster of 16 towers in the hadronic calorimeter is used to determine the hadronic isolation. Then, the RoI will be an isolated candidate if the E_T sum in the hadronic cluster is lower than the hadronic isolation threshold.

After passing these conditions, the identified cluster will be considered a region of interest (RoI) for a possible electron or photon.

³Eventually more than one threshold can be used and will depend on the availability of bits to process. Something similar is used for the thresholds of electromagnetic and hadronic isolation.

Rates

One of the challenges on trigger implementation is to reduce the rates of the least interesting events without losing important information. For that, vetoes and coincidences are applied, as was shown in the algorithm section.

Figure 2.17: Trigger rate for electrons at L1 as a function of the p_T threshold at $10^{33} \text{ cm}^{-2}\text{s}^{-1}$ (left) and $10^{34} \text{ cm}^{-2}\text{s}^{-1}$ (right). The solid lines represent the rates without isolation; dotted ones, only with hadronic isolation; and dashed ones with both hadronic and electromagnetic isolation [72].

Fig. 2.17 presents the triggers rates for electrons at L1 as a function of the p_T threshold at $10^{33} \text{ cm}^{-2}\text{s}^{-1}$ (left) and $10^{34} \text{ cm}^{-2}\text{s}^{-1}$ (right). Currently, the trigger configuration at L1 is projected for a bandwidth of 20 kHz for electromagnetic signals (from the 100 kHz available). In the case of $10^{33} \text{ cm}^{-2}\text{s}^{-1}$ it can be seen how even without applying the isolation criteria the rate can be maintained below a threshold of 20 GeV in transverse momentum, something that does not happen at $10^{34} \text{ cm}^{-2}\text{s}^{-1}$, where in addition to the isolation criteria, it is also necessary to increase the threshold to 30 GeV.

As can be seen, the situation tends to deteriorate as the luminosity increases and requiring higher thresholds, something not recommended since electrons containing information of W and Z bosons will have the peak in the p_T distribution too close to this threshold. For example, on the analysis of interest in this thesis ($Z + b$ -jets) the electron p_T threshold recommended to analysis at 7 TeV was 20 GeV. A first recommendation at 13 TeV was of

25 GeV, further progressing to a recommendation of 27 GeV (currently the luminosity reached a value of $13.8 \times 10^{33} \text{ cm}^{-2}\text{s}^{-1}$). These recommendations are done by the E/gamma group of ATLAS, which is responsible by the electrons and photon performance studies, taking also considerations of the trigger performance.

An extended study about rate on simulations will be shown in Ch. 8, where will be studied the improvements for electron measurements.

2.4.2 Trigger for jets

Similar to the case of electrons, this trigger identifies jet candidates using information from calorimeters, classifying them according to their E_T and isolations, identifying the coordinates of the regions of interest (RoI) and their classification to be used at L2.

Algorithm

The information for the algorithm is provided by a set of towers with a granularity of $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$ in $|\eta| < 3.2$ (0.4×0.4 in $3.2 < |\eta| < 4.9$). The algorithm starts searching clusters in towers of 2×2 ($\Delta\eta \times \Delta\phi = 0.4 \times 0.4$) with an E_T considered as maximum local, and above a certain threshold. These clusters will be considered candidates for RoI.

In order to take into account the size of jet dispersion, the energy in windows around the RoI (with sizes of 0.6×0.6 and 0.8×0.8) are also considered and stored. Fig. 2.18 presents these arrangements: the case of 0.6×0.6 , where there are four possibilities to form a window, and the case of 0.8×0.8 , with the window centered on the RoI. The information in these windows will be compared at different thresholds of E_T , and the optimization will depend on E_T and luminosity.

Efficiencies

Fig. 2.19 shows the efficiencies for L1 single-jet triggers as a function of leading off-line jet pt thresholds in central and forward regions. As expected, the efficiency decreases for jets at low p_T when the threshold increases. More details about jet reconstruction and calibration will be presented in Ch. 4.

Figure 2.18: Scheme used for jet algorithm at Level 1 [72].

Figure 2.19: Efficiencies for L1 single-jet triggers, shown as a function of leading off-line jet p_T for different thresholds (J20 indicates a p_T cut of 20 GeV, similar for the others). (a) For the central region ($|\eta| < 2.8$) and (b) for the forward region ($3.6 < |\eta| < 4.5$) [78].

2.4.3 Rejection of jets with low transverse moment at high luminosity

The study of jet rejections can be evaluated through Monte Carlo simulations of $Z \rightarrow ee$ production, overlapping jets with low p_T and including the effect of pile-up ($\mu \sim 80$ in Run 3 and $\mu \sim 200$ in the HL-LHC).

The operation of the level 1 trigger system of electrons in ATLAS shows that the bandwidth consumed by the signals of isolated electrons should be kept below 20 kHz, which corresponds to approximately 20% of the total available bandwidth for all L1 triggers. However, as it was shown in Sec. 2.4.1, at a

luminosity of $10^{34} \text{ cm}^{-2}\text{s}^{-1}$, simulations indicate that the current trigger system would accuse the presence of isolated clusters (associated to fake electrons) which will force an increase of the p_{T} threshold from 25 GeV to above 50 GeV, affecting the detection of bosons Z and W .

As an increase in the threshold is not an option to maintain the high rate of signatures with electromagnetic behavior, a possible solution to this problem is to increase the granularity of the trigger towers, implementing the so-called supercells, which will be studied in Ch. 8.

Chapter 3

Measurement and Calibration of Electrons in ATLAS

Knowing the L1 trigger for electrons and jets, the next step is to understand how these objects are reconstructed in ATLAS. The understanding and measurement accuracy of electrons is essential in several physics analysis in ATLAS. For example this thesis will be focused on the Z boson decaying into electrons ($Z \rightarrow ee$), presenting studies of $Z(\rightarrow ee) + b$ -jets analysis and event selection studies on situations where the electron trigger decisions occur at high luminosity. This chapter summarizes the measurement and calibration method of electrons currently used in ATLAS after passing the trigger selections.

Fig. 3.1 shows an event display of a $Z \rightarrow ee$ candidate, showing the electron tracks (yellow lines) and their arrival into detector. The left plot presents a longitudinal cut of the detector; the right bottom one shows the transversal cut; and the right top plot presents the distribution of transverse energy as function of η and ϕ .

3.1 Electron identification

The reconstruction of electrons is done over the events that pass the trigger selection. In ATLAS, the electrons are reconstructed by the information collected by the electromagnetic (EM) calorimeter and the Inner Detector (ID) in the region of $|\eta| < 2.47$.

The EM calorimeter is segmented in towers of $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ (Fig. 3.2). The energy of each tower is determined by the sum of the energies of all cells in each longitudinal layer inside of it. The reconstruction process

Figure 3.1: An event display with the production of two electrons ($Z \rightarrow ee$) measured in ATLAS [38].

Figure 3.2: Scheme of the segmentation for towers used on electron reconstruction.

starts by searching *seed clusters*, which are clusters with a transverse energy above 2.5 GeV. These clusters are searched using a sliding-window algorithm [79], which groups the towers in windows of 3×5 ($\Delta\eta \times \Delta\phi = 0.075 \times 0.125$).

Electrons and photons enter into the EM calorimeter interacting with the

material and producing similar electromagnetic showers. Considering that, the track reconstruction is important in the electron-photon discrimination. These tracks are reconstructed by the Inner Detector. Starting from seeds, an electron or a converted photon¹ is reconstructed if it has at least a well reconstructed track that matches the cluster. For converted photons it is also included the vertex conversion, which are classified depending on whether there was a single or a double conversion. Clusters that do not have a match with any track will be considered unconverted photons. Finally, the objects are grouped following the clusterization presented in Tab. 3.1.

Object	EMB	EMEC
Electron	3×7	5×5
Unconverted photon	3×5	5×5
Converted photon	3×7	5×5

Table 3.1: *Clusterization of towers for electrons and photons using the electromagnetic calorimeter in ATLAS.*

3.2 Jet discrimination

Fake electrons signals, associated to low p_T jets, can be also misidentified as electrons. This section will describe the jet discrimination in ATLAS based on the differences between the shapes of showers produced by electrons (electromagnetic showers) and jets (hadronic showers) [69] through the calorimeters, tracker and combined calorimeter/tracker information. They can be applied independently and three reference sets of cuts have been defined in order to increase the jet rejection power: *loose*, *medium* and *tight* [80].

3.2.1 Loose selection

This is the lowest selection, representing a selection with the less jet rejection. The variables used for electron loose selection are based on the information of the hadronic calorimeter and the second layer of the electromagnetic calorimeter, and are defined as:

¹A converted photon is a photon generating an electron-positron pair in its passage by the ID.

Acceptance: Determined by the acceptance of the EM calorimeter:

$$|\eta^{\text{EMCluster}}| < 2.47$$

Hadronic leakage:

$$R_{had1} = \frac{E_T^{\text{Had.Cal.(1)}}}{E_T^{\text{EMCluster}}} \quad (3.1)$$

$$R_{had} = \frac{E_T^{\text{Had.Cal.}}}{E_T^{\text{EMCluster}}} \quad (3.2)$$

where $E_T^{\text{EMCluster}}$ represents the transverse energy of the EM cluster; $E_T^{\text{Had.Cal.(1)}}$ and $E_T^{\text{Had.Cal.}}$ represent the transverse energy of the first and all layers of the hadronic calorimeter, respectively. R_{had1} is used over $|\eta| < 0.8$ and $|\eta| > 1.37$; while R_{had} over $0.8 < |\eta| < 1.37$

Second layer of EM calorimeter: These are powerful discriminants, which will be further discussed in this work: The η spread (R_η) and the lateral shower width ($W_{\eta,2}$) defined as:

$$R_\eta = \frac{E_{T,\Delta\eta \times \Delta\phi=0.075 \times 0.2}^{(2)}}{E_{T,\Delta\eta \times \Delta\phi=0.175 \times 0.2}^{(2)}} \quad (3.3)$$

$$W_{\eta,2} = \sqrt{\frac{\sum (E_T^{(2)} \times \eta^2)_{\Delta\eta \times \Delta\phi=0.075 \times 0.2}}{E_{T,\Delta\eta \times \Delta\phi=0.075 \times 0.2}^{(2)}} - \left(\frac{\sum (E_T^{(2)} \times \eta)_{\Delta\eta \times \Delta\phi=0.075 \times 0.2}}{E_{T,\Delta\eta \times \Delta\phi=0.075 \times 0.2}^{(2)}} \right)^2} \quad (3.4)$$

where $E_{T,\Delta\eta \times \Delta\phi}^{(2)}$ represents the transverse energy in the second layer of the EM calorimeter on an area of $\Delta\eta \times \Delta\phi$. Fig. 3.3 shows the scheme of clusters used for R_η measurement, considering towers of $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ in the second layer of the electromagnetic trigger.

3.2.2 Medium selection

This selection is higher than the loose one, with more jet rejection (and it includes the loose selection). The variables used for electron medium selection are based on the information of the first layer of the electromagnetic calorimeter

Figure 3.3: Scheme of clusters used for R_η measurement.

and the tracking information. These variables are:

First layer of EM calorimeter:

- Total shower width, defined as:

$$W_{stot} = \sqrt{(\sum E_i^{(1)}(i - i_{max})^2)(\sum E_i^{(1)})} \quad (3.5)$$

where i runs over all strips in the first layer of the EM calorimeter in a window of $\Delta\eta \times \Delta\phi \approx 0.0625 \times 0.2$, corresponding typically to 20 strips in η , and i_{max} is the index of the highest-energy strip.

Track quality and matching:

- Number of hits in the pixel detector (≥ 1).
- Number of hits in the pixel and SCT (≥ 7).
- Transverse impact parameter ($d_0 < 5$ mm).
- $\Delta\eta$ cut between the cluster and the track:

$$|\eta^{\text{EMCluster}} - \eta^{\text{track}}| < 0.01$$

3.2.3 Tight selection

This is the highest selection, with the highest jet rejection, and includes both the loose and medium variables. The variables used for electron tight selection are:

Track quality and matching:

- Tighter transverse impact parameter cut ($d_0 < 1$ mm)
- $\Delta\phi$ cut:

$$|\phi^{\text{EMCluster}} - \phi^{\text{track}}| < 0.02$$

- Ratio of the cluster energy to the track momentum:

$$\frac{E^{\text{EMCluster}}}{p_{\text{T}}^{\text{track}}}$$

- Tighter $\Delta\eta$ cut:

$$|\eta^{\text{EMCluster}} - \eta^{\text{track}}| < 0.005$$

TRT information:

- Total number of hits in the TRT.
- Ratio of the number of high-threshold hits to the total number of hits in the TRT.

Conversions:

- Number of hits in the b-layer (≥ 1). The b-layer or bottom layer is the layer 0 of the pixel detector.
- Electron candidates matching to reconstructed photon conversions are rejected.

The shower variables are given as an input to a multi-variate analysis program (TMVA [81]) in order to perform a cut optimization.

Preliminary studies at high luminosity showed that the shower shape discriminants would be efficient at the first level trigger in ATLAS [8]. Fig. 3.4

presents distributions of discriminants, allowing the possibility to get a good discrimination between jets and electrons (for example, it might consider as electrons those objects with $R_\eta > 0.95$ and $W_{\eta,2} < 0.012$).

Figure 3.4: Distribution of R_η and $W_{\eta,2}$ for electrons and jets with $p_T > 20$ GeV at high luminosity ($\mu = 80$) [8]

3.3 Calibration

The calibration procedure is detailed in Fig. 3.5, following two important stages: the response calibration based on Monte Carlo and the calibration in-situ (with data).

Response calibration based on Monte Carlo

After the EM cluster reconstruction (Fig. 3.5: step 0), techniques of multi-variable analysis (MVA) [81] are used to relate the properties of the reconstructed electromagnetic cluster with the energies of the original simulated electrons (or photons) (Fig. 3.5: step 1 and 4).

For experimental data, are taken into account some corrections that are not implemented on the simulations for specific regions over the detector; for example, areas with high voltage problems, geometry defects or corrections associated with electronic calibrations in the calorimeter (Fig. 3.5: step 2). After those corrections, and considering that the EM calorimeter is longitudinally segmented, it is applied an energy weighting of layers (Fig. 3.5: step

Figure 3.5: Scheme used in the calibration of electrons and photons [82].

3). Finally, the calibration based on Monte Carlo response is applied (Fig. 3.5: step 4).

Calibration in-situ

The calibration process involves an additional process that uses a technique of *tag and probe*, taking electrons from the decay of Z bosons ($Z \rightarrow ee$). This is done to calculate the scale factors of electron to be applied in the off-line reconstruction (Fig. 3.5: step 5).

Finally, the energy scale is validated with electron candidates in samples of $J/\Psi \rightarrow ee$ coming from experimental data. The scale factors obtained from $Z \rightarrow ee$ are assumed also valid for photons, then these scale factors will be validated from samples of $Z \rightarrow \ell\ell\gamma$ (Fig. 3.5: step 6).

3.4 Results on MC calibration

Here will be presented some results after the MC calibration, where the most probable value (MPV) is obtained from the ratio of the reconstructed energy after calibration and the simulated original energy, $E_{\text{reco}}^{\text{calib}}/E_{\text{true}}$, obtained with MVA [81].

Fig. 3.6 presents the MPV for electrons and photons due to E_{true} and $|\eta|$. In the case of electrons, the MPV has a variation of $\pm 0.3\%$ with respect to unity in the central region (up to $|\eta| \approx 1.2$) and $\pm 0.5\%$ for $|\eta| > 1.5$. It is

always expected some differences around the crack region ($1.37 < |\eta| < 1.55$), especially for electrons with low energy. In the case of photons, the variation of MPV practically remains stable ($\pm 0.4\%$) over all EM calorimeter.

Figure 3.6: *Most probable value (MPV) of the energy response for electrons (a) and unconverted photons (b), as a function of $|\eta|$ for different energy values [82]. This analysis does not consider values in the transition region from the barrel to the end-caps (crack region: $1.37 < |\eta| < 1.55$).*

These results show the power of the current calibration based on MC. In practical terms, after the reconstruction, the calibration recover the information of the simulated electron/photon.

Chapter 4

Measurement and Calibration of Jets in ATLAS

The understanding and accuracy of jet measurements play an important role on almost every analysis in ATLAS and the analysis performed in this thesis are no exception. As anticipated, jets will be present in calibration studies, their association with Z ($Z+b$ -jets analysis) and upgrade studies, since represent the main background source at high luminosity (pile-up). This chapter summarizes the measurement and calibration method used for jets in ATLAS. Fig. 4.1 shows an event display for a multijet production, indicating each jet with a different color; the left plots present the transversal and longitudinal cuts of the detector; and the right plots present the distributions of transverse energy as function of η and ϕ .

4.1 Cone jet definition

Jets are objects used to identify partons (quarks and gluons) produced after hard collisions. This section will present the definition of the cone algorithm and its limitations.

4.1.1 The Snowmass accord and the cone algorithms

In 1990, in Snowmass, Colorado, it was proposed a standard definition for measurement of jets [83] in order to be a reference for both theoretical and experimental studies. This definition was based on the following conditions:

- Simple to implement on experimental analysis.
- Simple to implement on theoretical calculations.

Figure 4.1: Event display with jets production measured in ATLAS [38].

- Defined at any order of perturbative theory (maintaining a finite cross section).
- Maintain a cross section insensitive to hadronization.

Following the Snowmass accord, jets are defined as cones with radii R_0 in the phase space $\eta - \phi$, reconstructed from calorimeter cells. The reconstruction algorithm follow these steps:

Step 1: Search for seeds, which are usually clusters with an energy greater than a certain threshold.

Step 2: The obtained seed will be considered the initial centroid of the jet. The centroid will be shifted each time that a new cell is added, then, will be considered part of the jet, the cells that are within the radius $R \leq R_0$:

$$R \equiv \sqrt{(\phi_i - \phi_o)^2 + (\eta_i - \eta_o)^2}$$

where ϕ_o and η_o are the centroid coordinates and ϕ_i and η_i are the coordinates of the i -th cell. In the Snowmass accord, values proposed for R_0 were between 0.4 and 1.0.

Step 3: Kinematic quantities, such as transverse energy, pseudorapidity and azimuthal angle of the jet, are defined as:

$$E_T^{\text{jet}} \equiv \sum_{i \in R \leq R_0} E_{T,i}$$

$$\eta^{\text{jet}} \equiv \frac{1}{E_T^{\text{jet}}} \sum_{i \in R \leq R_0} \eta_i E_{T,i}$$

$$\phi^{\text{jet}} \equiv \frac{1}{E_T^{\text{jet}}} \sum_{i \in R \leq R_0} \phi_i E_{T,i}$$

where $E_{T,i}$, η_i and ϕ_i represent the transverse energy, pseudorapidity and azimuthal angle of the i -th cell.

Step 4: The reconstruction algorithm is iterative, and stops when the centroid of the jet remains fixed.

While this algorithm presents the basics for jet reconstruction, several corrections still need to be applied, an example is the correction for overlapped jets, where *split/merge* algorithms must be introduced.

Cone algorithms were used extensively at the Tevatron, with the JETCLU used in Run I at CDF (based in the Snowmass accord), and the D0RUNIICONE algorithm [84], used in Run II at DØ.

One of the limitations of the cone type algorithms, in particular the one proposed in the Snowmass accord, is the processing time. Considering N particles, the processing time increases with a factor of N^2 . A solution for that was the SISCONE [85] (*Seedless Infrared-Safe Cone*) jet algorithm, where the clustering cone type is done by regions (towers) delimited in the detector, including the criteria of *merge/splitting*. With the SISCONE algorithm, the processing time is proportional reduced to $N^2 \ln N$.

4.1.2 Radiated gluons and cone algorithms

Regarding the modeling of physics process at LHC energy scale, there is great interest in the effects associated to radiated gluons. Since partons as colored particles can radiate gluons.

Fig. 4.2 shows the Feynman diagrams from a virtual photon into two quarks with a radiated gluon, the photon can be product of e^+e^- or $q\bar{q}$, but

for simplification, it will be considered only the final states.

Figure 4.2: Feynman diagrams for gluon emission on a $\gamma^* \rightarrow q\bar{q}$ process (virtual photon). $\mathbf{p}_1$ and $\mathbf{p}_2$ are the tri-momentum of the quarks and $\mathbf{k}$ of the gluon.

Considering a soft gluon ($k \ll p_{1,2}$), the contribution to the differential cross section correspondent to the gluon emission part (dS) [86] can be expressed by:

$$dS = \frac{2\alpha_s C_F}{\pi} \frac{dE}{E} \frac{d\theta}{\sin\theta} \frac{d\phi}{2\pi} \quad (4.1)$$

where α_s is the running coupling for strong force, $C_F = 3/4$ is the Cassimir factor, E is the gluon energy, θ the angle $q - g$ and ϕ the azimuthal angle.

From Eq. 4.1 is found that the expression diverges if $E \rightarrow 0$, this is known as *infrared (or soft) divergence*. But if $\theta \rightarrow 0$ or $\theta \rightarrow \pi$ the expression diverges too and this is known as *collinear divergence*. Then, considering radiated gluons, there are two important considerations that need to be taken into account by safe jet reconstruction algorithms, they need to be: infra-red and collinear safety.

Infra-red safety: The algorithm must be insensitive to infrared divergence (associated with emitted gluons at low energy). Fig. 4.3.a shows a diagram for two jets produced at *leading-order* (LO) and how these jets are experimentally obtained from two different seeds. This also can happen if it is considered the diagram in Fig. 4.3.b, where there are two reconstructed jets at *next-leading-order* (NLO). Theoretically, the contribution of the second diagram to cross section produces a divergence due to the loop and that divergence needs to be canceled adding the contribution of the diagram in Fig. 4.3.c. Experimentally, the emitted soft gluon (which causes the divergence) can produce a seed, absorbing the other partons and forming a single jet. In this case it is necessary to take into account this effect, since we need to reconstruct two jets.

Figure 4.3: Jet production with soft gluon emission: (a) Production at LO. (b) Production at NLO with a virtual gluon. (c) Production at NLO with a radiated gluon [87].

Collinear safety: The algorithm must be insensitive to collinear divergence. Fig. 4.4.a and Fig. 4.4.c perform the reconstruction scheme for a jet considering three partons, where the central parton will be considered a seed in the scheme, the vertical lines represent the transverse momenta and the horizontal ones the angular separation. Theoretically, the gluon in the loop produces a divergence, but, in order to get a collinear safety algorithm, the divergence should be canceled with the contribution in Fig. 4.4.b, where the central particle emits a collinear gluon, although that is not always the case, as is shown by Fig. 4.4.d, since at the time that the central parton emits a gluon, its energy is divided, then, the reconstruction can consider the left one as a new seed, absorbing the central parton (and gluon) in a single jet, reconstructing the right one as a new jet.

Figure 4.4: Jet production with collinear emission, heights of vertical lines represent the transverse momenta and lengths of horizontal lines the angular separation [87].

4.2 Sequential clustering algorithms

A solution to the infrared and collinear emission issues was proposed with the sequential clustering algorithms: the Cam/Aachen algorithm [88], the k_t algorithm [89] and anti- k_t algorithm [90]. All these algorithms use the concept of topoclusters, which will be discussed in this section along with the details of the algorithms.

4.2.1 Topoclusters

In ATLAS, jets are reconstructed from *topoclusters*, formed by the merging of individual cells in the calorimeters. Considering the Fig. 4.5, a topocluster will be formed following:

Figure 4.5: 2D scheme sequence for topocluster reconstruction in the η phase space.

- Search for a seed cell: $E_T^{\text{seed}} > 4\sigma^{\text{noise}}$ (red cell)
- Then, are added the neighbors with: $E_T^{\text{neigh}} > 2\sigma^{\text{noise}}$ (green cells)
- Finally, are included all cells on the perimeter with: $E_T^{\text{cell}} > 0$ (blue cells)

where E_T^{seed} , E_T^{neigh} and E_T^{cell} represent the transverse energy of the seed, its neighbors and the perimeter cells. σ^{noise} represents the electronic noise.

4.2.2 Generalization

The generalization of sequential clustering algorithms can be described by the following steps:

Step 0: The algorithm starts from a list of *protojets*¹ (topoclusters in this study) and a free list of the reconstructed jets.

Step 1: For each protojet i is defined:

$$d_i \equiv E_{T,i}^{2p}$$

and for each pair ij of protojets is defined:

$$d_{ij} \equiv \min(E_{T,i}^{2p}, E_{T,j}^{2p}) \frac{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2}{R_0^2} = \min(E_{T,i}^{2p}, E_{T,j}^{2p}) \frac{\Delta_{ij}^2}{R_0^2}$$

where R_0 is the distance parameter containing the jet and the parameter p is the generalization between algorithms:

- $p = 0 \Rightarrow$ Cam/Aachen algorithm
- $p = 1 \Rightarrow k_t$ algorithm
- $p = -1 \Rightarrow$ anti- k_t algorithm

Step 2: The algorithm searches for the lower value of d_i and d_{ij} :

- If d_i is lower, the protojet i is considered "not aggregable" and will be removed from the protojet list and placed in the list of jets.
- If d_{ij} is lower, the protojets i and j will be merged in a new protojet k , where:

$$\begin{aligned} E_{T,k} &\equiv E_{T,i} + E_{T,j} \\ \eta_k &\equiv \frac{E_{T,i}\eta_i + E_{T,j}\eta_j}{E_{T,k}} \\ \phi_k &\equiv \frac{E_{T,i}\phi_i + E_{T,j}\phi_j}{E_{T,k}} \end{aligned}$$

then, the protojet k is placed in the protojet list instead the protojets i and j .

¹Protojets are small clusters with energy, generally, slightly higher to electronic noise.

Step 3: The process is repeated until the protojet list is empty.

4.2.3 Comparison between algorithms

The Cam/Aachen algorithm

In the Cam/Aachen algorithm ($p = 0$), the preference to group protojets is limited only by the distance between them ($d_i = 1$ and $d_{ij} = \Delta_{ij}^2/R_0^2$). Then, the clustering will occur preferably for protojets that are in $\Delta_{ij} < R_0$ (step 2 of the algorithm), as can be seen in Fig. 4.6.a, which presents eight jets reconstructed with Cam/Aachen algorithm. It can be seen that the shape of the jets are not conical, since in each new iteration the merged protojets are changing their centroid. It can be also seen that the three most energetic jets (green, blue and red) are easily reconstructed, since their seeds are well identified, something that does not occur with the others, because it will depend on the energy threshold applied over each seed.

Figure 4.6: Reconstruction of jets with different algorithms: The Cam/Aachen (a), the k_t (b), the anti- k_t (c) and the SIScone (d) [90].

The k_t algorithm

In the case of the k_t algorithm ($p = 1$), the protojets energy will influence the clustering according to the distance. This can be seen if we imagine a protojet A , very energetic, with transverse energy $E_{T,A}$, and others less energetic protojets, k and l , with similar transverse energy $E_{T,l} \approx E_{T,k}$. With respect to the distances, we going to imagine that A and k are closer protojets ², then, $\Delta_{lA}^2 \approx \Delta_{lk}^2$. Considering the protojet l as reference, we have the next relations: $d_{lA} \approx E_{T,A}^2 \Delta_{lA}^2$ and $d_{lk} \approx E_{T,k}^2 \Delta_{lk}^2 = E_{T,k}^2 \Delta_{lA}^2$. Considering the step 2 of the algorithm, it is found that l have the preference to merge with k , the less energetic protojet. We then expect a considerable change of the jet centroid in each iteration, producing also not conical shapes on the jet distributions, as is shown in Fig. 4.6.b.

The anti- k_t algorithm

The anti- k_t algorithm ($p = -1$) is the standard method of reconstruction in ATLAS and CMS. We can imagine again the protojets presented for the k_t algorithm. The protojet l will have the following relationships: $d_{lA} \approx \Delta_{lA}^2 / E_{T,A}^2$ and $d_{lk} \approx \Delta_{lk}^2 / E_{T,k}^2$. Now, the protojet l have the preference to join with A , the most energetic protojet and in this case the jet centroid tends to be fixed close to it, producing jet distributions with conical shapes, as is shown in Fig. 4.6.c. Another important aspect of the anti- k_t algorithm is that in the case of superposed jets, the algorithm gives the merging preference to the most energetic protojet. At this level, the algorithm does not have a scheme of *merge/splitting*, as is done by the SIScone in Fig. 4.6.d. While this is not a problem for jet counting, it represents a problem for the estimation of jet kinematic characteristics. For that reason, in jet analysis the tendency is preferably to work with isolated jets.

Fig. 4.7 shows the Higgs mass distribution, reconstructed from the invariant mass of two b -jets using the three sequential clustering algorithms ($R_0 = 0.6$) and the JETCLU (cone type with $R_0 = 0.4$). As can be seen, the results are slightly better for Cam/Aachem and anti- k_t , since they present higher peaks on the distributions. In ATLAS there is the possibility to use several sequential clustering algorithms, which are provided by the FASTJET

²The factor Δ_{ij}^2 / R_0^2 , on step 1 of the algorithm avoids the clustering between far protojets.

package [91].

Figure 4.7: Simulation of the Higgs mass reconstruction using different algorithms in the reconstruction of jets ($p\bar{p} \rightarrow WH \rightarrow l\nu bb$) [92].

4.2.4 The influence of the parameter R_0

Fig. 4.7 shows that the algorithm choice does not have an important impact on the results, which are much more sensitive to the choice of R_0 , as is shown in Fig. 4.8.

The different results obtained for different R_0 was somewhat expected, since if R_0 is too small, the reconstructed jet may not contain all the expected energy (although always is necessary a calibration), and if is too large, the jet

Figure 4.8: Simulation of Higgs mass reconstruction, comparing *JETCLU* ($R_0 = 0.4$) with *kt* algorithm at $R_0 = 0.6$ (a) and $R_0 = 1.0$ (b) [92].

may also contain the energy of other jets (overlapping). The standard parameter in ATLAS is $R_0 = 0.4$, but with the increase of luminosity optimization studies to change this value are currently development in ATLAS. This work will show results testing other values of R_0 in Sec. 5.3.2.

4.3 Identification of jets in ATLAS

Fig. 4.9 shows a simplified scheme of jet reconstruction, which can be done using the track information or the deposited energy in the calorimeters.

Figure 4.9: Jet reconstruction using track and deposited energy information [44].

In ATLAS, the standard jet reconstruction is done by the anti- k_t algorithm. Fig. 4.10 shows a flow chart indicating the sequence in the reconstruction of the types of jets used in ATLAS. There are three types of jet reconstruction in ATLAS: *true-jets*, which are reconstructed using the simulated particles; the *track-jets*, which are jets reconstructed from the track information; and the calorimeter jets or *reco-jets*, jets that are reconstructed from cluster information in the calorimeters.

Figure 4.10: *Reconstruction sequence of jet types in ATLAS [94].*

The reco-jets are divided at two scales of energy measurement, the first one is known as the electromagnetic scale (EM scale), which only uses the signals produced by electromagnetic shower and the other is known as the local clusters weighting scale (LCW scale), in which is considered the electromagnetic shower and the compensation of the hadronic shower [93].

4.4 Calibration: Jet Energy Scale

The measurement of the jet energy scale (JES) or jet calibration, is determined by comparing energies of true-jets with the reco-jets. In ATLAS, conventional calibration methods are: the EM+JES (using jets at EM scale) and the LCW+JES (using jets at LCW scale).

Correction due to pile-up

Due to the large number of produced particles (which impacts the deposited energy in the calorimeters), the detector efficiency is compromised, becoming necessary to apply a correction due to pile-up. For analysis at 7 TeV [94] it was used an *off-set* correction, which depends on the number of primary vertices, N_{PV} , and the value associated to pile-up, μ . This scheme had limitations for analysis at 8 TeV and became particularly unacceptable for current analysis, where pile-up reach values close to $\mu = 40$. The next section will present the techniques used on the pile-up correction.

Origin correction

This correction is due to limitations in efficiency and resolution of primary vertex reconstruction, this correction will be made without affecting the energy.

Energy and pseudorapidity correction

These corrections are based on the numerical inversion, using reco- and true-jets in simulations. The sequence begins calibrating the energy, applying then the pseudorapidity correction. These corrections belongs to the calibration response based on Monte Carlo for the Jet Energy Scale (MC-JES), which will be detailed in Chap. 5.

Residual correction in data

Finally, a correction in data, *in-situ* correction, should be applied. This correction is applied to take into account differences in the detector simulation process.

4.5 Pile-up suppression

This section will present a resume of the techniques used on the pile-up suppression on the ATLAS experiment for analysis at 7 TeV and 8 TeV. New techniques, used on analysis at 13 TeV and for preliminary studies at high luminosity will be shown in App. J.

4.5.1 Area subtraction

The pile-up contribution is corrected in the jet transverse momentum (p_T^{jet}) and, for that, a correction which depends on a *jet area* [95] is applied. This area reflects the susceptibility of the jet to pile-up and is measured in each event. Then, the corrected transverse momentum is expressed as:

$$p_T^{\text{jet,corr}} = p_T^{\text{jet}} - \rho \times p_T^{\text{area}} \quad (4.2)$$

where ρ is obtained from the median transverse momentum density ($p_T^{\text{jet}}/p_T^{\text{area}}$), which measures the activity of the pile-up in each event and does not depend on the parameter R_0 .

Figure 4.11: The mean anti- k_t $R = 0.4$ LCW jet multiplicity as a function of $\langle \mu \rangle$ in $Z + \text{jets}$ events for jets with $p_T > 20 \text{ GeV}$ and $|\eta| < 2.1$ (a) and $3.2 < |\eta| < 4.5$ (b). Events in this plot are required to have at least 1 jet both before and after applying the jet-area based pile-up correction [96].

Fig. 4.11 shows the mean jet multiplicity as a function of $\langle\mu\rangle$ in Z +jets events for $p_T^{\text{jet}} > 20$ GeV for two rapidity ranges. It can be seen that the number of reconstructed jets increases with the pile-up for both MC and Data, and, after the correction, it can be seen how the event-by-event pile-up subtraction based on jet areas removes the majority of pile-up jets. In fact, the ratio Data/MC is improved, but there is still a small difference, which is associated to imperfections in the jet modeling.

For calibrations at 8 TeV, the correction applied in the jet transverse momentum includes a residual correction, $O_r(N_{PV}, \mu, \eta_{det})$, (although minimal). Then, the corrected transverse momentum can be expressed as:

$$p_T^{\text{jet,corr}} = p_T^{\text{jet}} - \rho \times p_T^{\text{area}} - O_r(N_{PV}, \mu, \eta_{det}) \quad (4.3)$$

4.5.2 Jet Vertex Fraction (JVF)

A variable called the jet vertex fraction (JVF), defined with the tracking information, is used to discriminate jets coming from hard production of pile-up jets. The JVF of a jet with respect to a hard-scatter vertex (PV_0) is defined as:

$$\text{JVF} = \frac{\sum_k p_{T,k}^{\text{track}}(PV_0)}{\sum_l p_{T,l}^{\text{track}}(PV_0) + \sum_{n \geq 1} p_{T,l}^{\text{track}}(PV_n)} \quad (4.4)$$

where k and l run over all tracks originating from PV_0 , and n over all tracks originating from PV_n (all primary vertices) matched to jet l . JVF can be interpreted as the probability of the jet coming from PV_0 , bounded by 0 and 1, where a value of -1 is assigned to jets with no associated tracks.

Fig. 4.12 shows the JVF distribution in MC simulations for hard-scatter jets and for pile-up jets where the discrimination power of the JVF variable is clearly seen.

Fig. 4.13 shows the total jet multiplicity as a function of $\langle\mu\rangle$ in Z +jets events, with cuts of $\text{JVF} > 0.25$ and $\text{JVF} > 0.50$. The improvement is clearly visible, if compared with the area subtraction in the central region, and a reasonable agreement between data and MC.

Figure 4.12: JVF distribution for hard-scatter (blue) and pile-up (red) jets with $20 < p_T < 50$ GeV and $|\eta| < 2.4$ after pile-up subtraction and jet energy scale correction in simulated Z +jets events [96].

Figure 4.13: The mean anti- k_t $R = 0.4$ LCW+JES jet multiplicity as a function of $\langle\mu\rangle$ in Z +jets events for jets with $|\eta| < 2.1$, before and after selected $|JVF|$ cuts were applied to jets with $20 < p_T < 50$ GeV (a) and $40 < p_T < 50$ GeV (b) [96].

4.6 Results at 7 TeV

As a closing remark, this last section will show the latest public results on jet measurements from ATLAS. Fig. 4.14 shows results for the inclusive jet fiducial cross section measurements, comparing data and theory. In the case comparing data and MC, only results at 7 TeV are currently available. Measurements at 8 and 13 TeV are still under studies.

Figure 4.14: Ratio of inclusive jet fiducial cross section measurements to theory (status: August 2016) [15].

Chapter 5

Jet Calibration Studies

The measurement of production cross section in $Z + b$ -jets processes in current LHC conditions and the study of an electromagnetic trigger for electron jet discrimination relies heavily on the understanding of the jet reconstruction scheme deployed in ATLAS and this includes the understanding of the jet calibration methods. This chapter describes the work performed in collaboration with the Jet/Etmiss performance group in ATLAS to derive a set of improved jet energy scale calibration recommendations for 8 TeV analysis before the LHC shutdown in 2014 and studies to understand the impact of R_0 parameter. Even though these recommendations were not the ones used in the $Z + b$ -jets analysis which will be presented in this thesis (since the analysis uses 13 TeV data from recent 2015 and 2016 runs) they were crucial to understand the treatment of the systematics associated with the Jet Energy Scale (JES), as it represents one of the main sources to the total systematic uncertainty on that analysis. By other hand, first studies of jet calibration at high luminosity were also developed, presenting some results on App. J and pile-up suppression methods in App. K.

5.1 Overview

The MC-JES is a method to calibrate reconstructed jets from jets at particle level (true-jets). The method is based on the numerical inversion, which consists in the determination of calibration curves of response, $\mathcal{R}$:

$$\mathcal{R} \equiv \frac{E_{\text{reco}}}{E_{\text{true}}} \quad (5.1)$$

where E_{reco} and E_{true} are the reconstructed and true jet energy, respectively. The calibration curves are obtained using the procedures within the ATLAS reconstruction framework and the details are presented in the next section.

It is expected that, after applying the MC-JES, to observe the following features in the Data/MC comparisons:

- **Good linearity:** For fixed ranges on pseudorapidity, the mean calibrated energy will be equal to the true energy:

$$\langle E_{\text{reco}}^{\text{calib}} \rangle = \langle E_{\text{true}} \rangle$$

where $E_{\text{reco}}^{\text{calib}} = JES \times E_{\text{reco}}$ and the JES value is taken from the inverse of the calibration curve (Fig. 5.1.a). After the JES calibration, a closure curve can be drawn, as is shown in Fig. 5.1.b.

Figure 5.1: Linearity expectations before (a) and after (b) the calibration ($R \rightarrow 1$) [44].

- **Good resolution:** It is expected an improvement on the resolution (σ_E/E) after the calibration, especially at low energies (Fig. 5.2).

Figure 5.2: Expectations for resolution improvement ($\sigma_E/E \rightarrow 0$) before (a) and after (b) the calibration [44].

- **Good accuracy on the spatial reconstruction of jets:** After the JES calibration, it is expected that the measurements of pseudorapidity and azimuthal angle be equal for both true- and corrected reco-jets:

$$\langle \eta_{\text{reco}}^{\text{corr}} \rangle = \langle \eta_{\text{true}} \rangle$$

$$\langle \phi_{\text{reco}}^{\text{corr}} \rangle = \langle \phi_{\text{true}} \rangle$$

5.2 MC-JES calibration procedure

5.2.1 Event and Jet Selection

In order to optimize the calibration procedure before the numerical inversion (NI), the events and jets in events must be selected according to the following criteria:

Event selection

- **Non-physical events:** Considering the uncalibrated jets, in each event, the jets must satisfy the kinematic condition:

$$m^{\text{jet}} < E^{\text{jet}}$$

where m^{jet} and E^{jet} are the mass and energy of the jet, respectively. Events containing jets that don't satisfy this condition are removed from the analysis.

- **p_T quality ratio:** For each event, we have a set of true-jets, J_i^{true} , and reco-jets, J_j^{reco} . Considering the leading true-jet (the jet with the highest p_T), only events that satisfy the following condition will be considered:

$$ratio_{min} < \frac{\langle p_{T,\text{reco}} \rangle}{p_{T,\text{leading}}^{\text{true}}} < ratio_{max}$$

where $\langle p_{T,\text{reco}} \rangle = \sum_{j=0}^m p_{T,j}^{\text{reco}}/m$. This selection is applied in order to avoid events with large fluctuations on the transverse moment.

Jet selection

The next step is to select jets within an event to be considered on the analysis.

- **Matching cut:** The true-jets and the reco-jets are expected to be collinear (Fig. 5.3) within a radius of:

$$\Delta R = \sqrt{(\eta_{\text{true}} - \eta_{\text{reco}})^2 + (\phi_{\text{true}} - \phi_{\text{reco}})^2} < R_{\text{match}}$$

where R_{match} is a fixed parameter, for example $R_{\text{match}} = 0.3$ on analysis at 8 TeV.

Figure 5.3: Criteria of the matching cut in the phase space η - ϕ .

- **Jet isolation:** After the matching, the next step is to know if the matched jet is an isolated jet, in order to avoid overlapping between jets that can produce large energy fluctuations. For that, the other reco-jets of the set (J_i^{reco}) are used. The first step is to search for the lowest

ΔR between the matched jet and the other reco-jets, ΔR_{min}^{reco} , so, the matched jet will be considered isolated if it satisfies the next condition:

$$\Delta R_{min}^{reco} > 0.25 \times R_0$$

where R_0 is the parameter of the anti- k_t algorithm (0.4 for most of the ATLAS analysis). If any matched jet does not satisfy the condition, the matched reco-jet is removed from the analysis. After passing the isolation condition for reco-jets, it is checked if the matched true-jet is also an isolated jet (considering the J_i^{true} set), applying a similar condition:

$$\Delta R_{min}^{true} > 0.15 \times R_0$$

Finally, the kinematic information of true-jets and reco-jets is stored separately in pseudorapidity segments (η bins).

5.2.2 Calibration Curves

Energy calibration

The response $\mathcal{R}$ is extracted fitting 2D histograms of E_{reco}/E_{true} as a function of E_{true} ($\mathcal{R}(E_{true})$). Fig. 5.4.a shows one of the histogram used on the calibration, as a particular case in a pseudorapidity range of $0.0 < \eta < 0.1$. Fig. 5.4.b presents the energy correlation between E_{reco} and E_{true} . The lower values for E_{reco} observed in Fig. 5.4.b are expected, since only the electromagnetic contribution of the shower is taken into account; the green line represents the identity function and the yellow line represents the minimum transverse energy value (E_T^{min}). This E_T^{min} represents a cut in the analysis, since below this value the calibration is less efficient, due to the limitations associated to pile-up.

The value of $\mathcal{R}$ is obtained from a Gaussian adjustment for each E_{true} bin from histogram in Fig. 5.4.a. Fig. 5.5 shows this adjustment for two bins of E_{true} .

$\mathcal{R}$ will be expressed as a function of E_{reco} , in order to be used on the curve fitting, as is shown in Fig. 5.6. The function to be fitted is:

$$\mathcal{C}(E_{reco}) = a_0 + a_1 \log(E_{reco}) + a_2 \log^2(E_{reco}) + a_3 \log^3(E_{reco}) + a_4 \log^4(E_{reco}) \quad (5.2)$$

Figure 5.4: 2D histogram of the jet energy response (a) and the energy correlation (b) before calibration, for EM scale jets with $R=0.4$. In the correlation plot, the green line represents the identity function and the yellow one represents the minimum transverse energy value of the true-jet.

Figure 5.5: Gaussian adjustment for $\mathcal{R}$ (blue curve), considering two different E_{true} bins and the same pseudorapidity range ($0.0 < \eta < 0.1$)

The polynomial function was obtained by the jet performance group.

The parameters a_i from Eq. 5.2 will be the outputs of the MC calibration. The calibration curve can be defined as: $\mathcal{C}_{JES}$, so, the JES is expressed as:

$$JES(E_{reco}) = \frac{1}{\mathcal{C}_{JES}(E_{reco})} \quad (5.3)$$

Finally, the calibrated energy is obtained from:

$$E_{reco}^{JES} = JES \times E_{reco} \quad (5.4)$$

Figure 5.6: Fitting of the energy calibration curve (blue line) for two pseudorapidity ranges. The red dotted line represents an extrapolation curve, useful to improve the fitting with the help of Groom points.

Closure test

It is important to verify the efficiency of the calibration, and this is done by repeating the calibration process considering the calibrated jets and computing a new response $\mathcal{R}_c = E_{\text{reco}}^{\text{JES}} / E_{\text{true}}$ (closure test).

Figure 5.7: Jet energy response (left) and correlation energy (right), after calibration, for EM+JES scale jets with $R = 0.4$. In the correlation plot, the green line represents the identity function and the yellow one represents the minimum transverse energy value of the true-jet.

Fig. 5.7 shows the results after the calibration. On the left panel it can be seen how the distribution for $\mathcal{R}_c$ has a behavior closer to the unity, except at low energies. This can be seen more clearly in the right panel, where the correlation is practically the same between $E_{\text{reco}}^{\text{JES}}$ and E_{true} for values above

the minimum transverse energy (yellow line).

Pseudorapidity correction

To improve the spatial reconstruction of jets, a pseudorapidity correction is applied: $\Delta\eta = \eta_{\text{reco}} - \eta_{\text{true}}$. A fitting over the distribution of $\Delta\eta$ as a function of E_{true} is done with a similar curve used in the energy calibration:

$$\mathcal{C}(E_{\text{true}}) = b_0 + b_1 \log(E_{\text{true}}) + b_2 \log^2(E_{\text{true}}) + b_3 \log^3(E_{\text{true}}) + b_4 \log^4(E_{\text{true}}) \quad (5.5)$$

Figure 5.8: Calibration curve for pseudorapidity as a function of E_{true} (green line) for two pseudorapidity ranges.

Fig. 5.8 shows the corrections in two different pseudorapidity ranges. The central region ($0 < \eta < 0.1$) presents a minimal correction, but for a forward region ($3.0 < \eta < 3.1$) the correction is quite significant. The parameters b_i from Eq. 5.5 are stored for pseudorapidity corrections of jets. The correction curve can be defined as: $\mathcal{C}_{JEtas}$, so, the $JEtas$ (jet η scale) is defined as:

$$JEtas(E_{\text{reco}}^{JES}) = -1.0 \times \mathcal{C}_{JEtas}(E_{\text{reco}}^{JES}) \quad (5.6)$$

The corrected pseudorapidity value is obtained from:

$$\eta_{\text{reco}}^{EtaJES} = \eta_{\text{reco}} + JEtas \quad (5.7)$$

As a comment, the transverse momentum of the jet will be corrected

after the energy calibration and η correction, expressed by:

$$p_{T,\text{reco}}^{EtaJES} = \frac{JES \times \sqrt{E_{\text{reco}}^2 - m_{\text{reco}}^2}}{\cosh(\eta_{\text{reco}}^{EtaJES})} \quad (5.8)$$

where m_{reco} is the mass of the reco-jet.

5.3 MC-JES at 8 TeV

This section presents the Monte Carlo calibration results using simulations for analysis at 8 TeV, as part of performance studies developed with the Jet/Etmiss performance group of ATLAS. The goal was to test the performance of the MC-JES procedure at EM and LCW scale, and to optimize the calibration considering other values of R_0 .

Samples

These studies were based on di-jet simulations, considering proton-proton collisions at 8 TeV (Run 1) generated with PYTHIA 8 [97], a general-purpose LO MC generator, using the AU2CT10 [25] PDF set. These samples are sliced according the p_T of the leading true-jet as is shown in Tab. 5.1. The p_T slicing allows a more efficient simulation and reconstruction since a fully inclusive sample would need to be very large to include enough events in high p_T region (Notice that the cross section spans 14 orders of magnitude).

Sample	$p_{T,\text{leading}}^{\text{truth}}$ [GeV]	σ [nb]
jetjet_JZ0W	0-20	7.18E+07
jetjet_JZ1W	20-60	9.26E+03
jetjet_JZ2W	60-160	1.05E+02
jetjet_JZ3W	160-400	6.63E-01
jetjet_JZ4W	400-800	4.56E-03
jetjet_JZ5W	800-1300	8.55E-05
jetjet_JZ6W	1300-1800	1.96E-06
jetjet_JZ7W	1800-2500	5.93E-07

Table 5.1: *Simulated di-jet samples and their cross sections.*

Selections

Jets are reconstructed from topological energy clusters in the calorimeter using the anti- k_t algorithm with radii parameters of $R_0 = 0.4$ and 0.6 at EM scale

reconstruction and $R_0 = 0.2, 0.3, 0.4, 0.5, 0.6, 0.7$ and 0.8 for the R-Scan analysis at LCW scale. Kinematic selections are shown in Tab. 5.2

Requirement	Cut
Matching	$\Delta R < 0.3$
Truth Isolation	$\Delta R < 0.15R_0$
p_T true-jet	$p_{T,\text{true}} > 7 \text{ GeV}$
Reco Isolation	$\Delta R < 0.25R_0$
p_T reco-jet	$p_{T,\text{reco}} > 7 \text{ GeV}$
E_T threshold	$E_T^{\text{min}} > 20$

Table 5.2: Kinematic selections for MC-JES calibration at 8 TeV.

The pile-up suppression applied was based in the area subtraction, including the residual correction (Eq. 4.3).

5.3.1 EM scale calibration

Fig. 5.9 presents the response before the calibration (EM scale) for $R = 0.4$ and 0.6 as a function of the pseudorapidity. The plots present the response for five energy values of reference, and as expected, the response is much smaller than 1, increasing as the reference energy increases. The discontinuities in the curves are result of the transition region between calorimeter systems in ATLAS. For example, the crack region between the barrel and end-caps of the EM calorimeter ($1.47 < |\eta| < 1.53$) and the pass from the EM to the forward calorimeter ($|\eta| \sim 3.2$).

Figure 5.9: Energy response before the calibration (EM scale) for $R = 0.4$ and 0.6 , as function of the pseudorapidity and for five energy values of reference (energy of true-jets).

Fig. 5.10 presents the closure plots as result of the calibrated energy response (EM+JES scale). As it can be seen, after calibration, the deviations, with respect to the unity, are smaller than 2% (being slightly better for $R = 0.4$).

Figure 5.10: Energy response after calibration (EM+JES scale) for $R = 0.4$ and 0.6 , as function of the pseudorapidity and for five energy values of reference.

These studies provided the inputs (calibration factors and recommendations) to be used by several analysis performed by the Standard Model and SUSY groups of ATLAS. The E_T^{min} previously recommended was 20 GeV, but after several tests it was concluded that the minimal transverse energy recommended for analysis should be change to 25 GeV for $R = 0.4$ and to 35 GeV for $R = 0.6$. This recommendation sacrifices some low p_T jets in data in favor of a better closure, and it is a first line of defense to mitigate pile-up effects in MC simulation.

Currently (Run 2), an improvement of this calibration is applied, called *Global Sequential Calibration* (GSC) [98], where it is proposed different corrections for different layers of the calorimeters. For the Run 1 (8 TeV) the LCW+JES scale presents the best results on jet calibration, as will be shown below.

5.3.2 LCW scale calibration: R-scan

In ATLAS analysis the conventional value for the parameter R_0 is 0.4 and for calibration studies it is also frequently used 0.6. This section will discuss the problem of optimizing the calibration using other values for R_0 , such as 0.2, 0.3, 0.5, 0.7, 0.8, at LCW scale.

Fig. 5.11 compares the response for calibrated jets at LCW+JES scale. As shown, there was an improvement compared with EM+JES scale (for $R_0=0.4$ and 0.6), where it was found a deviation below 0.5%. These preliminary studies show that the optimal values for R_0 would be $R_0=0.3$ and 0.4 .

The E_T^{min} values recommended by the R-scan are presented in Tab. 5.3.

Recomendations	Cut
E_T threshold	$E_T^{min} > 25 \text{ GeV } (R_0 = 0.2, 0.3, 0.4)$
	$E_T^{min} > 30 \text{ GeV } (R_0 = 0.5)$
	$E_T^{min} > 35 \text{ GeV } (R_0 = 0.6)$
	$E_T^{min} > 40 \text{ GeV } (R_0 = 0.7, 0.8)$

Table 5.3: *R-scan recomendations for MC-JES calibration (LCW+JES scale) at 8 TeV.*

Figure 5.11: Energy response after calibration (LCW+JES scale) for $R = 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8$, as function of the pseudorapidity and for five energy reference values (energy of true-jets).

Chapter 6

Production of Z Bosons Associated with b -jets in ATLAS

After to know details about the reconstruction and calibration methods for electrons and jets in ATLAS, it is important to know more about the production of Z boson associated with b -jets ($Z + b$ -jets), whose motivation was presented in Sec. 1.3.4. In this chapter it will be presented details about the theoretical fundamentals, cross section predictions, the b -jet selection, the main backgrounds on the analysis and some of the previously published results obtained with ATLAS and CMS. This work will be focused on the production of Z with $1b$ - and $2b$ -jets on the inclusive cases (events with at least $1b$ - and $2b$ -jets, respectively). Exclusive cases, events with exactly $1b$ and $2b$'s, can be extracted as subsets of the inclusive ones.

6.1 Theoretical fundamentals

The production of bosons associated with b -jets will be strongly related with the quarks on the initial-states. Considering that, simulations are based following two schemes, one of them is the *four-flavor number scheme* (4FNS), where only quarks u , d , s and c are considered on the initial-state; meaning that only parton distributions (PDF's) of these quarks and gluons are included as generator inputs used in simulations (the PDF set). The other one is the *five-flavor number scheme* (5FNS), where the b -quark is also included in the PDF set. Details about the simulation process can be found in App. B.1.

Fig. 6.1 presents Feynman diagrams for $Z + b$ productions on the 4FNS. The diagrams are pictorial representations of the contributions to define the

matrix element (ME), used to compute cross sections (App. C). For example, these diagrams represent contributions at leading-order (LO) for Z production with at least $2b$ -jets. They can also represent contributions for events with at least $1b$ -jet considering the case when the two b 's on the final-state are merged into a single jet or either one b is missed considering the kinematic cuts (cuts on p_T and $|y|$). On the 5FNS, these diagrams are LO contributions for events with at least $2b$ -jets.

Figure 6.1: *Feynman diagrams considered on the four-flavor number scheme: LO for Z production with at least 2 b -jets ($q = u, d, s$ and c).*

Fig. 6.2 presents Feynman diagrams for $Z + b$ production on the 5FNS. These diagrams represent contributions at LO for Z production with at least $1b$ -jet.

Figure 6.2: *Feynman diagrams considered on the five-flavour number scheme: LO for Z production with at least 1 b -jet.*

The calculations can be extended to include contribution up to NLO. For example, on the calculation for $Z + 1b$ -jet on 5FNS, the contributions at NLO

are:

- $gb \rightarrow Zb$: diagrams in Fig. 6.2 with one loop.
- $gg \rightarrow Zb\bar{b}$ and $qq \rightarrow Zb\bar{b}$: diagrams in Fig. 6.1 where the two b 's are merged into a single jet or either one b is missed (kinematic cuts).
- $qb \rightarrow Zqb$ and $gb \rightarrow Zgb$: diagrams in Fig. 6.3.

Figure 6.3: *NLO Feynman diagrams for Z production with at least one b-jet ($q = u, d, s$ and c) on the five-flavour number scheme.*

With the increase of the order, the number of diagrams grows up and the precision will depend on the computational limitations. To solve that, contributions to higher orders are expressed by the *parton shower* (PS) evolution. Details about the ME and PS evolution can be found in App. B.2.

6.2 Cross section predictions

The cross section predictions can be done at two levels: The parton level, where final states of quarks and gluons are included in the computation, and the particle level, using the particle reconstruction information, which can be related to the measurements on experimental data¹. Particle-level predictions consider effects such as QED final-state radiation (FSR), hadronization, underlying events and multiple parton-parton interactions (MPI). At this level, the quarks or gluons are detected by jet reconstruction, merging particle-level signals (which are produced by hadronization and/or consequent decays). To compare parton-level and particle-level predictions, the first one must be corrected by the effects presented before (FSR, hadronization, underlying events and MPI).

¹The methods to measure observables must be the same for both simulation and experimental analysis.

There are computational packages used to predict cross sections, for example, at parton-level predictions, one of the package used is the MCFM [99, 100], a package designed to compute cross-sections for various femtobarn-level processes. For most processes, matrix elements are included at NLO considering the 5FNS. For particle-level predictions, one of the package used is the aMC@NLO [101], which uses NLO matrix element calculations for both the 4FNS and 5FNS. For Analysis at 13 TeV, it will be used a version mixing MADGRAPH5 and aMC@NLO, MADGRAPH5_aMC@NLO [102], which is interfaced to PYTHIA 8.186 [97] to describe the parton shower. This version unifies the two main lines of development of the LO and NLO features in a single integrated framework, for example all the tree-level/LO functionalities in the SM and BSM of MADGRAPH5 and the improvements in speed and efficiency on the beta version of aMC@NLO.

Full particle-level predictions used in this thesis will be based on SHERPA [103] and ALPGEN [104] event generators, which combine tree-level matrix elements for multiple jet emissions. SHERPA includes their correspondent parton shower, hadronization and underlying event packages [105], while ALPGEN is commonly interfaced with HERWIG [106] or PYTHIA [97] to describe the parton shower and the other processes (HERWIG was used on analysis at 7 TeV and PYTHIA is used at 13 TeV). For analysis at 13 TeV, ALPGEN simulation uses the 4FNS while SHERPA uses the 5FNS as schemes.

6.3 B-jet selection

This section will describe the b -jet identification for both simulation and experimental analysis.

6.3.1 B-jet selection in simulations

Hadron association

Since in the simulation we have access to the hadron species, an association using the reconstructed jet and the truth information is employed. This association is defined as an spatial matching between reconstructed jets and one or more weakly decaying b -hadrons, as is shown in Fig. 6.4. Then, the reconstructed jet is labeled as a b -jet if it lies within $\Delta R = R_b^{\text{match}}$, where R_b^{match} is determined by efficiency studies. Commonly, a p_T cut ($p_{T,b}^{\text{min}}$) on the b -hadron

is applied in order to avoid mismatching associated to low efficiencies at low p_T .

Figure 6.4: Scheme used to describe the b -hadron association.

Reconstructed jets not identified as b -jets are considered as c -jets using the same logic: the reconstructed jet is labeled as a c -jet if they lie within $\Delta R = R_c^{\text{match}}$ from any c -quark with $p_T > p_{T,c}^{\text{min}}$. All other jets (not classified as c - or b -jets) are classified as *light*-jets.

Ghost hadron association

This process also requires the hadron species information. A ghost hadron is a b -hadron with p_T set to an infinitesimal amount ($p_T \sim 0$), essentially only retaining their directions [107], this ensures that the jet reconstruction is not altered by the ghosts. Then, the jet is re-clustered and if the b -ghost is contained in the new jet, this jet is considered as a b -jet, as can be seen in Fig. 6.5.

Figure 6.5: Scheme used to describe the b -ghost association. R_0 is the clusterization radius.

Reconstructed jets not identified as b -jets are considered as c -jets using the same logic: if they are associated to a ghost c -hadron. All other jets (not classified as c - or b -jets) are classified as *light*-jets.

6.3.2 B-jet selection in data

The process to select b -jets experimentally is commonly referred as b -tagging [108] and is operationally implemented as a search for jets originated from hadronization of a b -quark using the signatures of b -hadrons inside the jet. These signatures are based on properties of b -hadrons, such as: long lifetime, large mass, decays with high track multiplicities, etc.

Figure 6.6: Scheme of vertex displacement for a b -quark [109].

Fig. 6.6 shows an scheme of vertex displacement for a b -quark after its production. Considering the relatively long lifetime of hadrons containing a b -quark ($\tau \sim 1.5$ ps), it is expected a relative large displacement of the b -quark between a primary vertex and a secondary vertex ($L_{xy} \sim \text{few mm}$), producing consequently, large transversal and longitudinal impact parameters (d_0 , z_0). The large b -quark mass ($m_b \sim 4.18$ GeV) reflects the large track multiplicity from the secondary vertex.

The association of tracks to jets will be useful in order to reduce the number of tracks coming from other process, such as fragmentation, underlying events or pile-up [110]. The association is defined as an spatial matching inside ΔR defined as:

$$\Delta R_{\text{jet,track}} \equiv \sqrt{(\eta^{\text{jet}} - \eta^{\text{track}})^2 + (\phi^{\text{jet}} - \phi^{\text{track}})^2}$$

Commonly, the association cut ΔR depends on the jet p_T in order to have a smaller cone for high- p_T jets, which are more collimated.

One of the common b -tagging algorithms used in ATLAS is the IP3D [108], an impact parameter-based algorithm, which combines the impact parameter significances of all the tracks in the jet, considering both the transverse and longitudinal impact parameters. This algorithm is based on a log-likelihood ratio (LLR) method, using parameters to estimate the better discrimination between b -, c - or *light*-jets. For each track, the LLR uses the two-dimensional significance, $S \equiv S(d0/\sigma_{d0}, z0/\sigma_{z0})$, which is compared to the two-dimensional probability density functions (PDF's), σ_{d0} and σ_{z0} are the reconstruction uncertainties for $d0$ and $z0$, respectively. The output of this algorithm is a weight (w) associated to each jet, indicating the probability to be a b - or c - (*light*-) jet.

Other b -tagging algorithms are the secondary vertex-based algorithms [108], which use the inclusive vertices formed by the decay products of the b -hadron, including the products of the eventual subsequent charm hadron decay. This algorithm starts by building all two-track pairs that form a good vertex, using only tracks associated to the jet and far enough from the primary vertex (vertices compatible with a primary vertex or material interaction are rejected); finally, all tracks from the remaining two-track vertices are combined into a single inclusive vertex, using an iterative procedure to remove the worst track until that the χ^2 of the vertex fit is good.

SV1 [108] is a secondary vertex-based algorithm, which takes advantage of three of the vertex properties: the vertex mass ², the ratio of the sum of the energies of the tracks in the vertex to the sum of the energies of all tracks in the jet and the number of two-track vertices. These variables are combined using a likelihood ratio technique. SV1 is commonly combined with IP3D in order to increase its discrimination power.

Other secondary vertex-based algorithm is the JetFitter [111], which exploits the topology of weak b - and c -hadron decays inside the jet. A Kalman filter [112] is used to find a common line on which the primary vertex and the b - and c -vertices lie, as well as their position on this line, giving an approximated

²The vertex mass is defined as the invariant mass of all tracks associated to the vertex, assuming that all tracks are related to pions.

flight path for the b -hadron. The discrimination between b -, c - and *light*-jets is based on a likelihood method with similar variables used in the SV1 algorithm. The JetFitter is commonly combined with a neural network-based [113] algorithm and is defined as JetFitterCombNN, a sub-version is called the JetFitterCombNNc with the exception that the neural network is trained to reject c -jets rather than *light*-jets.

Figure 6.7: *Light-jet rejection (a) and c -jet rejection (b) as a function of the b -tag efficiency for the b -tagging algorithms calibrated on analysis at 7 TeV, based on simulated $t\bar{t}$ events [114].*

The algorithm used at 13 TeV analysis is the MV1 [108, 114] algorithm (the same one used at 7 TeV), which is a neural network-based algorithm that uses as inputs the output weights of the algorithms presented before (IP3D, SV1 and JetFitterCombNN). MV1 provides an output weight w , used for discrimination between b - and *non-b*-jets (c -jets and *light*-jets). Lower values of w are assigned to *non-b*-jets, whereas the probability to get a b -jet increases with w . Fig. 6.7 shows the expected performance of the b -tagging algorithms in a simulated $t\bar{t}$ sample for jets with $p_T > 15$ GeV and $|\eta| < 2.5$. There, for each b -tagging algorithm is defined a set of operating points, which depends on the w cut value. For example, for an efficiency correspond to 70%, MV1 presents a better light rejection compared with the others (Fig. 6.7.a), although with respect to c , all of them are comparable for efficiencies above 70% (Fig. 6.7.b).

After the MV1 b -tagging algorithm, the selected jets on data may eventually include *light*- or c -jets misidentified as b -jets. With that into account, a fitting process must be done in order to extract the detector-level signal yields. This fitting is applied using the b -jets, *non*- b -jets and often background estimations.

6.4 Main backgrounds on experimental measurement

Events not coming from the $Z + b$ -jets production are considered backgrounds. These events pass the selection as a consequence of limitations on the reconstruction methods. For example, it can be considered as valid: events including fake electrons and real b -jets, events with real electrons and fake b -jets, or even events with both fake electrons and fake b -jets.

In general, contribution of *non*- b -jets is determined using fits on the selected jets from data, but often, it is estimated by MC simulations normalized to their theoretically predicted cross-sections.

Figure 6.8: $t\bar{t}$ production, with $t \rightarrow Wb$.

Other contributions for the background may come from $t\bar{t}$ events, which are considered the main background for $Z + 2b$ -jets productions. $t\bar{t}$ events can produce signals of two W and two b -jets (Fig. 6.8), where both W bosons decay to leptons ($W \rightarrow \ell\nu_\ell$), and eventually, a charged lepton pair $\ell\bar{\ell}$ can be confused with a pair coming from $Z \rightarrow \ell\bar{\ell}$. One of the methods to reduce this background is applying cuts on invariant mass ($m_{\ell\bar{\ell}}$) and missing transversal energy (MET), since it is expected differences on the invariant mass for charged lepton pair coming from Z , and the MET associated to neutrinos coming especially from W decays.

Another background (with a small contribution) comes from ZZ events, which can present signals of two leptons and two jets. Eventually, one or the

two quarks coming from one Z can be confused with b -jets associated to the other Z (Fig. 6.9).

Figure 6.9: ZZ production, with one $Z \rightarrow \ell\ell$ and other $Z \rightarrow qq$.

Other electroweak processes such as W +jets, $Z \rightarrow \tau\tau$, ZW , WW , single top and events with multijet production represent small contributions for $Z+b$ -jets. Here, the background contribution to the final data is estimated using samples of simulated events normalized to their theoretically predicted cross-sections.

6.5 ATLAS results at 7 TeV

In this section will be presented the cross section measurements for analysis at 7 TeV [115] comparing predictions and data in order to set the scenario expected for the analysis at 13 TeV.

The experimental data was collected during stable proton-proton collisions in 2011, with an integrated luminosity of $4.58 \pm 0.08 \text{ fb}^{-1}$. The results were based on distributions combining both electron and muon channels ($Z(\rightarrow \ell\ell) + b\text{-jets}$).

The predictions were based in inclusive $Z \rightarrow \ell\ell$ events simulated with ALPGEN 2.13. ALPGEN is interfaced to HERWIG 6.520 [106], to model the parton shower and hadronization, and to JIMMY 4.31 [116] to model the underlying event and multi-parton interactions (MPI). These samples are generated with the CTEQ6L1 [117] PDF set. In addition, overlaps between samples with heavy-flavor quarks originating from the matrix element and from the parton shower are removed to avoid double counting (at 13 TeV this process is known as Heavy Flavor Overlap Removal (HFOR)).

Alternative $Z \rightarrow \ell\ell$ samples used for systematic cross-checks are generated with SHERPA 1.4.1, using the CT10 [25] PDF set and matched to the parton shower using the CKKW prescription [118, 119].

The kinematic selection for objects are presented in Tab. 6.1.

Requirement	Cut
electron E_T	$E_T^e > 20 \text{ GeV}$
electron η	$ \eta^e < 2.47$ (excluding $1.37 < \eta^e < 1.52$)
muon p_T	$p_T^\mu > 20 \text{ GeV}$
muon η	$ \eta^\mu < 2.4$
Di-lepton mass	$76 < m_{\ell\ell} < 106 \text{ GeV}$
MET	$E_T^{\text{miss}} < 70 \text{ GeV}$
Jet algorithm	anti- k_t (R=0.4)
Jet p_T	$p_T^{\text{jet}} > 20 \text{ GeV}$
Jet rapidity	$ y^{\text{jet}} < 2.4$
Jet-lepton separation	$\Delta R(\ell, \text{jet}) > 0.5$

Table 6.1: Definition of the phase space for the fiducial cross section for $Z + b\text{-jet}$ measurements at 7 TeV in ATLAS.

In Fig. 6.10 and Table. 6.2 are presented the measurements of cross-sections for $Z + b$ -jets considering the following regions:

- 1**b**-tag region: Events with at least 1**b**-jet.
- 2**b**-tag region: Events with at least 2**b**-jets.

	$\sigma(Z+ \geq 1b\text{-jet})$ [fb] (1 b -tag)	$\sigma(Z+ \geq 2b\text{-jets})$ [fb] (2 b -tag)
Data	$4820 \pm 60^{+360}_{-380}$	$520 \pm 20^{+74}_{-72}$
MCFM $\otimes$ MSTW2008	$5230 \pm 30^{+690}_{-710}$	$410 \pm 10^{+60}_{-60}$
MCFM $\otimes$ CT10	$4850 \pm 30^{+580}_{-680}$	$386 \pm 5^{+55}_{-50}$
MCFM $\otimes$ NNPDF23	$5420 \pm 20^{+670}_{-710}$	$420 \pm 10^{+70}_{-50}$
aMC@NLO 4FNS $\otimes$ MSTW2008	$3390 \pm 20^{+580}_{-480}$	$485 \pm 7^{+80}_{-70}$
aMC@NLO 5FNS $\otimes$ MSTW2008	$4680 \pm 40^{+550}_{-580}$	$314 \pm 9^{+30}_{-20}$
SHERPA $\otimes$ CT10	3770 ± 10	422 ± 2
ALPGEN+HJ $\otimes$ CTEQ6L1	2580 ± 10	317 ± 2

Table 6.2: The measurement and theory predictions for the inclusive b -jet cross-sections in ATLAS for collisions at 7 TeV. The statistical uncertainty is quoted first in each case. The second uncertainty is either the total systematic uncertainty (data), the sum in quadrature of all theory uncertainties (MCFM), or the scale uncertainty (aMC@NLO) [115].

The MCFM predictions underestimated the data but agree within the combined experimental and theoretical uncertainties for both regions. Considering the three PDF's used, the MCFM prediction obtained with CT10 is lower than the others, this is associated to the default choice of $\alpha_s(m(Z))^3$: $\alpha_s(m(Z))=0.118$ for CT10 and $\alpha_s(m(Z))=0.120$ for MSTW2008 and NNPDF2.3. Compared with data, the prediction with CT10 in 1**b**-tag represents the better result.

For aMC@NLO, the 5FNS prediction describes 1**b**-tag well, while the 4FNS underestimates the measured cross-section. This situation is reversed for the 2**b**-tag case, where the 4FNS provides a good description, while the 5FNS underestimates the cross section. However, the 5FNS prediction from aMC@NLO is only LO for $Z+ \geq 2b$ -jets, which may help in explaining this underestimation. These results are important, as they represent the effect

³ α_s is the strong coupling constant.

(a)

(b)

Figure 6.10: Cross-sections for $Z + \geq 1b$ -jet (a) and $Z + \geq 2b$ -jets (b) measured in ATLAS for collisions at 7 TeV. The data measurement is shown as a vertical blue line with the inner blue shaded band showing the corresponding statistical uncertainty and the outer green shaded band showing the sum in quadrature of statistical and systematic uncertainties [115].

of the flavor scheme choice and the same investigation will take place in the 13 TeV analysis.

Considering only statistical uncertainties, both the 4FNS prediction from ALPGEN+HJ and the 5FNS prediction from SHERPA underestimate the data, with ALPGEN+HJ below SHERPA by 30-40%.

6.6 CMS results on $Z + b$ -jets at 7 and 8 TeV

The CMS collaboration has also results $Z + b$ -jets production at $\sqrt{s}=7$ and 8 TeV. This section presents these results in order to contrast them with the results found in the last section.

Tab. 6.3 presents the selection for objects on $Z + b$ -jet measurements at 7 TeV in CMS and Tab. 6.4 presents the results and comparisons with the predictions.

Requirement	Cut
leptons p_T	$p_T^\ell > 20$ GeV
leptons η	$ \eta^\ell < 2.4$
Di-lepton mass	$76 < m_{\ell\ell} < 106$ GeV
Jet algorithm	anti- k_t (R=0.5)
Jet p_T	$p_T^{\text{jet}} > 25$ GeV
Jet rapidity	$ y^{\text{jet}} < 2.1$
Jet-lepton separation	$\Delta R(\ell, \text{jet}) > 0.5$

Table 6.3: Definition of the phase space for the fiducial cross section for $Z + b$ -jet measurements at 7 TeV in CMS.

	$\sigma(Z + \geq 1b\text{-jet})$ [pb] (1b-tag)	$\sigma(Z + \geq 2b\text{-jets})$ [pb] (2b-tag)
Data	$3.88 \pm 0.02 \pm 0.22$	$0.36 \pm 0.01 \pm 0.07$
MCFM $\otimes$ CTEQ6mE	$3.23^{+0.34}_{-0.40}$	$0.29^{+0.04}_{-0.04}$
aMC@NLO 4FNS	$2.71^{+0.52}_{-0.41}$	$0.35^{+0.08}_{-0.06}$
aMC@NLO 5FNS	$3.99^{+0.25}_{-0.29}$	$0.29^{+0.04}_{-0.04}$
MADGRAPH5+PY 4FNS $\otimes$ CTEQ6L1	$3.49^{+0.52}_{-0.91}$	$0.38^{+0.06}_{-0.10}$
MADGRAPH5+PY 5FNS $\otimes$ CTEQ6L1	4.03 ± 0.24	0.37 ± 0.07

Table 6.4: The measurement and theory predictions for the inclusive b -jet cross-sections in CMS for collisions at 7 TeV. The statistical uncertainty is quoted first in each case. The expectations from MCFM, aMC@NLO and MADGRAPH5 include uncertainties due to scale variations [124].

The MCFM prediction tends to underestimated the data as well as found in ATLAS (approximately 20% lower).

For aMC@NLO, the results on comparisons are similar than the ones found by ATLAS, where the 5FNS prediction describes 1b-tag well. The 4FNS underestimates the measured cross-section, reverting the situation for the 2b-

tag case, where the 4FNS provides a good description and the 5FNS underestimates the cross section.

For MADGRAPH5, the 4FNS and 5FNS prediction agree well with data for the $2b$ -tag case, while underestimates the data on 4FNS and overestimates the data on 5FNS for $1b$ -tag.

ATLAS results at 8 TeV are not available yet, but preliminary results in CMS are currently available in [125]. Tab. 6.5 presents the kinematic selection for objects on $Z + b$ -jet measurements at 8 TeV in CMS.

Requirement	Cut
electron p_T	$p_T^e > 20$ GeV
electron η	$ \eta^e < 2.4$ (excluding $1.44 < \eta^e < 1.57$)
muon p_T	$p_T^\mu > 20$ GeV
muon η	$ \eta^\mu < 2.4$
Di-lepton mass	$71 < m_{\ell\ell} < 111$ GeV
Jet algorithm	anti- k_t (R=0.5)
Jet p_T	$p_T^{\text{jet}} > 30$ GeV
Jet rapidity	$ y^{\text{jet}} < 2.4$
Jet-lepton separation	$\Delta R(\ell, \text{jet}) > 0.5$

Table 6.5: Definition of the phase space for the fiducial cross section for $Z + b$ -jet measurements at 8 TeV in CMS.

Tab. 6.4 presents the measurement and theory prediction for the inclusive b -jet cross-sections. The prediction underestimates the data (approximately 20% lower) in the $1b$ -tag case, while it presents a reasonable agreement in the $2b$ -tag case (within uncertainties).

	$\sigma(Z + \geq 1b\text{-jet})$ [pb] (1b-tag)	$\sigma(Z + \geq 2b\text{-jets})$ [pb] (2b-tag)
Data	$3.55 \pm 0.12 \pm 0.21$	$0.331 \pm 0.011 \pm 0.035$
MADGRAPH5_aMC@NLO+PY 5FNS $\otimes$ NNPDF3.0	$4.23^{+0.27}_{-0.37}$	$0.356^{+0.030}_{-0.031}$

Table 6.6: The measurement and theory prediction for the inclusive b -jet cross-sections in CMS for collisions at 8 TeV (MADGRAPH5 is under the 5FNS). The statistical uncertainty is quoted first in each case [125].

Chapter 7

Production of $Z + b$ -jets at 13 TeV

The previous chapter presented the motivation for the $Z + b$ -jets analysis and results at 7 TeV for ATLAS and CMS, showing a relatively good agreement with theoretical predictions, although with large uncertainties, which will be improved with the recent data collection at 13 TeV. We have a direct participation on this analysis on behalf with the *vector boson and heavy flavor* (V+HF) analysis team of ATLAS, starting with studies on particle-level simulations and concluding with the unfolding procedure.

This chapter will be divided in the signal sample validations studies, the experimental analysis strategy at 13 TeV, the unfolded procedure, the systematic uncertainty implementation and the presentation of results.

7.1 Signal sample validations

This section will discuss the results for studies done at particle-level simulations through truth analysis on SHERPA and ALPGEN simulations (samples used as signals). These results include sample validations and comparisons between generators, in order to get the expectations and the best theoretical references on $Z + b$ -jets analysis.

The sample validations include analysis on the more inclusive selection (Z +jets, including *non- b -jets*), comparing boson and jet distributions. Although in this thesis we focus on the electron channel, the validation in the muon channel is also included, since at particle level is expected the same electron behavior on kinematic distribution and it is a very important cross check of the simulation (comparisons with muons are shown in App. E).

7.1.1 MC samples

Sherpa

One of the generators used as signal is SHERPA. The version used at the beginning of the analysis was SHERPA 2.1.1, but as the analysis evolved, it included a more recent version of SHERPA: SHERPA 2.2.1.

SHERPA 2.1.1 (from now on referenced as Sherpa 2.1) allows the generation of matrix elements up to NLO accuracy for vector bosons (V) associated up to 2 jets processes (LO for 3 and 4 jets). For higher multiplicities it is employed the Sherpa model of the parton shower [126]. The LO and NLO matrix elements are generated using the COMIX [127] and OPENLOOPS [126] matrix element generators, and the merging with the parton shower is done using the ME+PS@NLO prescription [119]. This version uses the NLO CT10 PDF set under the five-flavor number scheme (5FNS). Details about Sherpa 2.1 samples can be found in Tab. 7.1.

Sample	Filter/Veto	σ [pb]
Zee_Pt0_70	CVetoBVeto	1.718E+03
	CFilterBVeto	3.275E+02
	BFilter	1.735E+02
Zee_Pt70_140	CVetoBVeto	4.885E+01
	CFilterBVeto	1.615E+01
	BFilter	8.367E+00
Zee_Pt140_280	CVetoBVeto	7.058E+00
	CFilterBVeto	2.740E+00
	BFilter	1.550E+00
Zee_Pt280_500	CVetoBVeto	4.739E-01
	CFilterBVeto	9.860E-02
	BFilter	1.851E-01

Table 7.1: Simulated Z +jets Sherpa 2.1 samples and their cross sections. These samples are sliced according to the Z p_T and combinations of filter and vetoes.

SHERPA 2.2.1 is the version used on current analysis (Run 2) [128]. As with Sherpa 2.1, this version allows the generation of matrix elements up to NLO accuracy for V associated up to 2 jets processes (LO for 3 and 4 jets), all the additional jets are produced by the parton shower. SHERPA 2.2.1 uses the ME+PS@NLO prescription for both the matrix-element generation and the parton showering. The PDF set used is NNPDF3.0NNLO with $\alpha_s = 0.118$ [26]

under the 5FNS. A refined tuning with respect to Sherpa 2.1 was implemented in order to improve the description of LHC data. This MC sample will be referred as Sherpa 2.2 with details in Tab. 7.2.

Sample	Filter/Veto	σ [pb]
Zee_MAXHTPT0_70	CVetoBVeto	1.627E+03
	CFilterBVeto	2.237E+02
	BFilter	1.265E+02
Zee_MAXHTPT70_140	CVetoBVeto	7.629E+01
	CFilterBVeto	2.034E+01
	BFilter	1.262E+01
Zee_MAXHTPT140_280	CVetoBVeto	2.503E+01
	CFilterBVeto	9.372E+00
	BFilter	6.083E+00
Zee_MAXHTPT280_500	CVetoBVeto	4.869E+00
	CFilterBVeto	2.280E+00
	BFilter	1.494E+00
Zee_MAXHTPT500_1000	—	1.808E+00
Zee_MAXHTPT1000_E_CMS	—	1.486E-01

Table 7.2: *Simulated Z+jets Sherpa 2.2 samples and their cross sections. These samples are sliced according to the max Z p_T and max H_T , and combinations of filter and vetoes [128]. H_T is given by the scalar p_T sum of all parton-level jets with $p_T > 20$ GeV.*

Alpgen

Another generator used as signal is ALPGEN. ALPGEN 2.14 is a LO matrix element generator, which is interfaced with PYTHIA 6 to model the parton shower evolution. The matrix elements are calculated up to five partons and under the four-flavor number scheme (4FNS) using the CTEQ6L1 PDF set. Processes containing massive treatment of the heavy-flavor quarks ($Z + c\bar{c}$ and $Z + b\bar{b}$) are also included (V+HF samples). Each of the V+HF sample has been produced separately, with up to five partons, and a heavy-flavor overlap removal (HFOR) procedure is used to combine the "inclusive" and the V+HF samples, as is presented in Tab. 7.3. The HFOR is needed to avoid double counting of phase space between different samples, since heavy-flavor quarks are produced by the gluon splitting in the parton shower and by the explicit matrix element calculations [128]. This MC sample will be referred as Alpgen with details of the HFOR in the same table.

Samples		σ [pb]		Rejection factor
		before HFOR	after HFOR	[%]
Inclusive	np0	1283.90	1211.57	5.63
	np1	277.24	235.15	15.18
	np2	75.80	60.02	20.82
	np3	21.58	16.02	25.76
	np4	6.15	4.42	28.17
	np5	2.49	1.69	32.11
Heavy flavor	ccNp0	65.73	63.29	3.70
	ccNp1	28.67	26.18	8.69
	ccNp2	8.01	7.04	12.12
	ccNp3Incl	4.68	3.78	19.15
	bbNp0	16.65	16.36	1.75
	bbNp1	7.64	7.29	4.62
	bbNp2	3.06	2.83	7.60
	bbNp3Incl	1.78	1.57	11.93

Table 7.3: Simulated Z +jets Alpgen samples and their cross sections before and after the HFOR procedure. These samples are sliced according to the number of partons. The HFOR rejection factor is also shown. The full signal sample comprises all samples from the table after the HFOR.

Cross sections

Sherpa and Alpgen samples are normalized to the NNLO inclusive Z cross sections [120] calculated with FEWZ [129], obtaining $\sigma_{\text{NNLO}}^Z = 1906 \pm 95$ pb, derived in the dilepton mass range of $66 < m_{\ell\ell} < 116$ GeV. In the Z production case, the normalization to NNLO cross section involves the k -factor, which is obtained from the ratio of the NNLO and generator predictions:

$$k = \frac{\sigma_{\text{NNLO}}^Z}{\sigma_{\text{gen}} \times f_{\text{gen}}}$$

where σ_{gen} represents the total cross sections derived by the simulation and f_{gen} is a factor which take into account the difference on dilepton mass range between generators: $m_{\ell\ell} > 40$ GeV for Sherpa 2.1 and Sherpa 2.2, and $60 < m_{\ell\ell} < 2000$ GeV for Alpgen.

$$f_{\text{Sherpa}} = \frac{N_{\text{events}}(m_{\ell\ell} > 40 \text{ GeV})}{N_{\text{events}}(66 < m_{\ell\ell} < 116 \text{ GeV})}$$

$$f_{\text{Alpgen}} = \frac{N_{\text{events}}(60 < m_{\ell\ell} < 2000 \text{ GeV})}{N_{\text{events}}(66 < m_{\ell\ell} < 116 \text{ GeV})}$$

where N_{events} represents the number of events which falls in the dilepton mass range.

Generator	σ [pb]	f_{gen}	k -factor	σ_{NNLO} [pb]
NNLO	-	-	-	1906 ± 95
Sherpa 2.1	2304	1.0843	0.9013	2064
Sherpa 2.2	2136	1.0843	0.9013	1925
Alpgen	1657	1.0365	1.1921	1975

Table 7.4: Cross sections, f_{gen} and k -factors for the reference generators, including the NNLO computation for the inclusive $Z(\rightarrow \ell\ell)$ cross section. For both Sherpa 2.1 and Sherpa 2.2 is used the same f_{gen} and k -factors. Alpgen with HFOR.

In Tab. 7.4 is presented the cross sections computed at NNLO, applying the k -factor. One can see that Sherpa 2.2 predicts a cross section closer to NNLO values.

Selections for truth level studies

At particle level, the Z boson is reconstructed by information provided by the HEPMC event record [130], which stores the information of each particle (at truth-level) of each event during the pre-processing stage (before the evaluation of any other kinematic requirement). This is accomplished by searching two same flavor leptons with opposite charge and with the same parent information. The selection and definition of the phase space for the fiducial cross section for the $Z + b$ -jet measurements are shown in Tab. 7.5.

Requirement	Cut
Lepton p_{T}	$p_{\text{T}}^{\ell} > 25 \text{ GeV}$
Lepton η	$ \eta^{\ell} < 2.5$
Dilepton mass	$76 < m_{\ell\ell} < 106 \text{ GeV}$
$p_{\text{T}}(Z)$	$p_{\text{T}}(Z) < 500 \text{ GeV}$
Jet algorithm	anti- k_t ($R=0.4$)
Jet p_{T}	$p_{\text{T}}^{\text{jet}} > 20 \text{ GeV}$
Jet rapidity	$ y^{\text{jet}} < 2.5$
Jet-lepton separation	$\Delta R(\ell, \text{jet}) > 0.4$

Table 7.5: Definition of the phase space for the fiducial cross section for the $Z+b$ -jet (truth level studies).

Selected leptons include the four-momentum of radiated photons within a cone of radius 0.1 around the final state lepton direction, and are known as *dressed lepton* as described in Fig. 7.1, where is presented a born electron, produced directly by the Z decay, and the dressed electron, which is composed by the bare electron (electron after the radiation) and the photon compensation.

Figure 7.1: *Electron states, showing the born, bare and dressed level. The same state classification is valid for muons.*

Jets are reconstructed from particle-level information (Sec. 1.3.2) using the anki- k_t algorithm with $R = 0.4$. The selection of b -jets was done using the ghost hadron association (Sec. 6.3.1).

7.1.2 HFOR validations for Alpgen samples

As anticipated, the HFOR is needed to avoid double counting of Heavy flavor quarks from matrix element and parton shower occupying the same phase space. The overlap removal procedure removes events with:

- Two heavy-flavor quarks arising from PS and with $\Delta R(q, \bar{q}) < 0.4$.
- Two heavy-flavor quarks arising from ME and with $\Delta R(q, \bar{q}) > 0.4$.

The HFOR is applied analyzing the HEPMC. The pairing of quarks is performed only between same flavors (b or c) and between particle/anti-particle pairs. Quarks produced by multi-parton interactions or originating directly from PDF's are not included in the HFOR procedure. The predicted cross section of each separate sample is then scaled down to take into account the removed fraction of events, as is shown in Tab. 7.3, where it is compared the rejection efficiency of events.

Figure 7.2: Z boson invariant mass (left) and p_T (right) comparisons on Alpgen simulations for $Z(\rightarrow ee) + \text{jets}$ at $\sqrt{s} = 13$ TeV (before and after the HFOR).

Figure 7.3: Inclusive multiplicity comparisons for jets (left) and b -jets (right) on Alpgen simulations for $Z(\rightarrow ee) + \text{jets}$ at $\sqrt{s} = 13$ TeV (before and after the HFOR).

The HFOR will be validated comparing distributions before and after the overlap removal. It is expected that the total cross section predicted by the combination of all samples with HFOR must agree with the prediction of the "inclusive" samples (without HFOR), but that depends on the HFOR efficiency. Fig. 7.2 presents the Z invariant mass and p_T distributions, as can be seen, the invariant mass is in agreement within 1%, although there is a small deviation (lower than 5%) for p_T around 100 GeV. As anticipated, these small deviations on the total cross sections are expected considering the efficiency of the HFOR procedure.

The main differences associated to the HFOR appears on jet multiplicities, as is shown in Fig. 7.3, where the number of jets and b -jets increase on Alpgen with HFOR, since this includes the $Z + c\bar{c}$ and the $Z + b\bar{b}$ samples.

Several other kinematic distributions used in the validation of the HFOR procedure, including kinematic jet distributions, can be found in App. D. From here, the only Alpgen sample to be used will be the Alpgen with HFOR.

7.1.3 Generator comparisons and predictions for signal at 13 TeV

Once we have a valid heavy flavor overlap removal for Alpgen, we can compare the prediction of Alpgen and Sherpa at particle level. The generator of reference is Sherpa 2.1.

Fig. 7.4 presents the comparisons of the Z invariant mass and p_T distributions. As expected, there are practically no differences on the invariant mass distribution, as the generators have this process well modeled. In fact, this distribution shows a quick validation test for missing samples or normalization issues. With respect to $Z p_T$, Sherpa 2.2 and Alpgen present the same trend, but are differences with respect to Sherpa 2.1, this fact is associated to the p_T recoil produced by the difference on jet multiplicities, as it will be discussed below. With respect to electron distributions, the p_T and η are very similar, as can be seen in Fig. 7.5.

Fig. 7.6 presents the inclusive distribution multiplicity of jets (N_{jets}) and b -jets ($N_{b\text{-jets}}$). With respect to N_{jets} , Alpgen tends to be lower than both Sherpa versions, this difference being associated with the flavor scheme, 4FNS on Alpgen and 5FNS on Sherpas (Sec. 6.1); Sherpa 2.2 is lower than Sherpa 2.1 up to three-jets bin although is higher for multiplicities above five jets. Com-

Figure 7.4: Z boson invariant mass (left) and p_T (right) comparisons on MC simulations for $Z(\rightarrow ee)$ +jets at $\sqrt{s} = 13$ TeV.

Figure 7.5: Lepton transverse momentum (left) and pseudo-rapidity (right) comparisons on MC simulations for $Z(\rightarrow ee)$ +jets at $\sqrt{s} = 13$ TeV.

monly, a re-weighting is applied in order to agree with experimental data, but this introduces an additional systematic uncertainty on the final analysis and as Sherpa 2.2 will be used for the final predictions this procedure does not apply here. With respect to $N_{b\text{-jets}}$, practically there are no differences between Sherpas, but as expected, Alpgen is lower than Sherpas in the one- b -jet bin (flavor scheme differences) and a slightly lower in the second b -jet bin.

Figure 7.6: Inclusive jet (left) and b -jet (right) multiplicity comparisons on MC simulations for $Z(\rightarrow ee)+jets$ at $\sqrt{s} = 13$ TeV.

We are interested in the production of Z associated with b -jets, but the more inclusive case ($Z+jets$) is also important in order to compare with other analyses [131]. Cross section predictions for $Z+jets$ at 13 TeV are presented in Tab. 7.6 while for $Z + b$ -jets are presented in Tab. 7.7. These values will be taken as reference for experimental analysis, bearing in mind that kinematic cuts, while in some cases different, have a small impact.

After the cross section predictions, it is important to know the behavior (shape) of the jet kinematic distributions for each generator:

- p_T and rapidity of the leading and sub-leading b -jet in events with $N_{b\text{-jets}} \geq 1$ and $N_{b\text{-jets}} \geq 2$, respectively.

- Dijets observables: $\Delta\phi_{bb}$, Δy_{bb} , ΔR_{bb} and M_{bb} in events with $N_{b\text{-jets}} \geq 2$.
- Boson distributions: p_T and rapidity of Z in events with $N_{b\text{-jets}} \geq 1$ and $N_{b\text{-jets}} \geq 2$.

Channel	Generator	$\sigma(Z+ \geq 1\text{-jet})$ [fb]	$\sigma(Z+ \geq 2\text{-jets})$ [fb]
$Z \rightarrow ee$	Sherpa 2.1	199308 ± 459	62231 ± 228
	Sherpa 2.2	186136 ± 446	55852 ± 173
	Alpgen	175334 ± 295	49519 ± 106

Table 7.6: Theory predictions for the fiducial inclusive jet cross-sections in proton-proton collisions at 13 TeV (normalized to NNLO). Fiducial cuts are presented in Tab. 7.5. Only the statistical uncertainty is quoted.

Channel	Generator	$\sigma(Z+ \geq 1b\text{-jet})$ [fb]	$\sigma(Z+ \geq 2b\text{-jets})$ [fb]
$Z \rightarrow ee$	Sherpa 2.1	11323 ± 49	1376 ± 17
	Sherpa 2.2	11488 ± 40	1404 ± 9
	Alpgen	7761 ± 36	1303 ± 12

Table 7.7: Theory predictions for the fiducial inclusive b -jet cross-sections in proton-proton collisions at 13 TeV (normalized to NNLO). Fiducial cuts are presented in Tab. 7.5. Only the statistical uncertainty is quoted.

Fig. 7.7 presents the p_T and $|y|$ distribution for leading and sub-leading b -jets, showing the different predictions on the leading jet p_T distribution, where Sherpa 2.2 yields tends to increase with the increase of p_T . Alpgen disagrees with both Sherpas for $1b$ -jet selection. For the sub-leading p_T distribution all of them agree for values up to $p_T < 100$ GeV. With respect to $|y|$, the shapes are very similar, again the only difference appear for the leading jet distribution on Alpgen.

Fig. 7.8 presents the main dijets observables, where practically there are no significant differences. This represents the theoretical convergence between generators. This is an import result that impacts in the choice of the generator reference in the unfolding procedure.

Finally, Fig. 7.9 presents the p_T and $|y|$ distribution for Z bosons in events with at least $1b$ jet and at least $2b$ -jets. The main differences appears on Z p_T in events with at least $1b$, where Sherpa 2.2 presents a harder p_T spectrum. In events with at least $2b$ -jets, the Z p_T distributions for Sherpa 2.2 and Alpgen tend to agree well at low p_T .

More distributions, including the full jet distributions and muon comparisons are presented in App. E.

Figure 7.7: Transverse momentum and rapidity comparisons for leading (top) and sub-leading (bottom) b-jets on MC simulations for $Z(\rightarrow ee) + \text{jets}$ at $\sqrt{s} = 13$ TeV.

Figure 7.8: Angular distributions ($\Delta\phi_{bb}$, Δy_{bb} and ΔR_{bb}) and invariant mass (M_{bb}) comparisons for dijets on MC simulations for $Z(\rightarrow ee) + \text{jets}$ at $\sqrt{s} = 13 \text{ TeV}$.

Figure 7.9: *Transverse momentum and rapidity comparisons for Z boson in events with at least 1b jet (top) and at least 2b-jets on MC simulations for $Z(\rightarrow ee)+$ jets at $\sqrt{s} = 13$ TeV.*

7.2 13 TeV data analysis

The experimental data was collected during stable proton-proton collisions in 2015 and 2016, with an integrated luminosity of $36.5 \pm 1.2 \text{ fb}^{-1}$. Tab. 7.8 presents the peak delivered instantaneous luminosity, the mean number of proton-proton interactions per bunch crossing (pile-up) and the integrated luminosity by year on measurements at 13 TeV.

Data taking	$\mathcal{L}_{peak}$ [$10^{33} \text{ cm}^{-2}\text{s}^{-1}$]	$\langle\mu\rangle$	$\mathcal{L}^{int}$ [fb^{-1}]
2015	5.0	13	3.2
2016	13.7	22	33.3

Table 7.8: Peak delivered instantaneous luminosity ($\mathcal{L}_{peak}$), the mean number of proton-proton interactions per bunch crossing (μ) and the integrated luminosity ($\mathcal{L}^{int}$) by year on measurements at 13 TeV [41].

7.2.1 Selections for experimental analysis

Events containing a Z boson candidate are selected by requiring exactly two selected electrons with opposite charge. Events are selected using high level triggers (HLT) that require at least one electron with a p_T threshold of 24 GeV during 2015, and 24-26 GeV in 2016. To identify possible electrons at L1 trigger, it was required an EM cluster with at least 18 or 20 GeV with hadronic isolation (Sec. 2.4.1).

An important comment about triggers must be done here, since this is a motivation for studies to be presented in the next chapter. For analysis at 7 TeV (2011) the HLT required two electrons with $p_T > 12$ GeV and, as was shown, currently, the p_T threshold reached a value of 26 GeV. These results had as consequence the increase on the p_T threshold of the electron reconstruction from $p_T > 20$ GeV in Run 1 to $p_T > 27$ GeV in the current analysis. This value must be higher for data collected in the next years (2017 and 2018). One can see from the electron p_T distribution in Fig. 7.5 how this increase on the lepton p_T trigger reduces the acceptance of Z bosons, and may become a challenge for other analyses (for example low mass Drell-Yan cross section [122]).

A Hard-scattering (HS) vertex is defined as the vertex with highest $\sum p_T^2$ of its constituent tracks (ID tracks). Candidate events are required to have at least one HS vertex with at least two associated tracks [132].

Electron candidates are reconstructed and calibrated by the methods shown in Ch. 3. Electrons are identified by a likelihood algorithm [133] considering the tight selection (Sec. 3.2). The selected electrons have $p_T > 27$ GeV and $|\eta| < 2.47$, excluding the crack region of the electromagnetic calorimeter ($1.37 < |\eta| < 1.52$).

Jets are reconstructed from topological energy clusters in the calorimeter using the anti- k_t algorithm with radius parameter $R = 0.4$. The jet energy is calibrated by the method shown in Sec. 5.2, including the global sequence calibration [98], with the area suppression and off-set correction to correct the pile-up activity in the calorimeter [134]. A residual in situ correction determined from Z +jet and γ +jet control samples is applied to jets in data to account for remaining differences in calorimeter response between data and simulation [94]. The selected jets have $p_T > 20$ GeV and $|\eta| < 2.5$.

To reduce the number of jets associated to pile-up, a new variable is introduced, the jet vertex tagger (JVT). Details about this variable will be found in App. K. The jets with $p_T > 60$ GeV and $|\eta| < 2.4$ are removed if $\text{JVT} < 0.64$.

To reduce the background arising from $t\bar{t}$, events are required to have a reconstructed di-electron invariant mass between $76 < m_{ee} < 106$ GeV and a missing energy cut of $E_T^{\text{miss}} < 60$ GeV. Fig. 7.10 presents the MET distributions of $t\bar{t}$ and signal for events with at least 1 b -jet and at least 2 b -jet. While the top background represents a challenge in the W +HF analysis due to leptonic top decay, for the Z analysis, the invariant mass restriction provides a good rejection against top background.

In data, b -jets are selected using the MV1 algorithm. b -jet candidates are selected if their MV1 output is greater than 0.64, corresponding to 77% average b -tagging efficiency. In simulation, b -jets are selected by the b -hadron association (Sec. 6.3.1).

Finally, to reduce the size samples, events must have at least one reconstructed jet in data to be stored. The selection and definition of the phase space for the fiducial cross section for the $Z + b$ -jet measurements are shown in Tab. 7.9.

Figure 7.10: *MET distributions of $t\bar{t}$ events and signal events for 1b-jet selection (left) and 2b-jet selection (right) in di-electron modes. Histograms are normalized to the theoretically predicted cross section [135].*

Requirement	Cut
electron p_T	$p_T^e > 27$ GeV
electron η	$ \eta^e < 2.47$ (excluding $1.37 < \eta^e < 1.52$)
Di-electron mass	$76 < m_{ee} < 106$ GeV
MET	$E_T^{\text{miss}} < 60$ GeV
Jet algorithm	anti- k_t ($R = 0.4$)
Jet p_T	$p_T^{\text{jet}} > 20$ GeV
Jet rapidity	$ y^{\text{jet}} < 2.5$
Jet-lepton separation	$\Delta R(\ell, \text{jet}) > 0.4$

Table 7.9: *Definition of the phase space for the fiducial cross section for $Z + b$ -jet measurements at 13 TeV.*

7.2.2 Simulated event samples

The samples used as signals on this analysis will be the same used on the truth studies presented in Sec. 7.1.1, but considering only Alpgen and Sherpa 2.2, the last one will be referred simply as Sherpa. The background samples includes top, diboson, tau decaying, W and Z samples. These are samples produced and validated centrally by the ATLAS performance Monte Carlo group and used on several other analyses across the ATLAS collaboration.

Samples of top production (pair and single-top events) are generated with POWHEG 2 interfaced with PYTHIA 6.428 parton-shower generator and the

Perugia 2012 tune. An alternate top-pair quark sample is produced using the aMC@NLO generator interfaced with HERWIG++ [136] and the UE EE5 tune. The samples are normalized to the cross section calculated at NNLO+NNLL (next-to-next-to-leading-log) with the TOP++ 2.0 program [137]. Diboson processes with four charged leptons, three charged leptons and one neutrino, or two charged leptons and two neutrinos are simulated using SHERPA 2.1.1 with the CT10 PDF set. The matrix elements contain the doubly-resonant WW , WZ and ZZ processes, and all other diagrams with four electroweak vertices [131]. The 4ℓ and $2\ell + 2\nu$ processes are calculated at NLO for up to one additional parton; final states with two and three additional partons are calculated at LO. The $3\ell + 1\nu$ process is calculated at NLO and up to three partons at LO using the COMIX and OPENLOOPS matrix element generators and merged with the Sherpa parton shower using the ME+PS@NLO prescription.

W +jets and $Z(\rightarrow \tau\tau)$ samples are also produced for both Sherpa and Alpgen generators, with the same prescription used for signals.

7.2.3 Backgrounds

The collected data contains events with the expected signal and background contributions. In the case of simulated samples, expected yields (N_i^{events}) can be expressed as the product of the production cross section on fiducial region ($\sigma_{\text{fid},i}$) and the integral luminosity ($\mathcal{L}_{\text{int}} = \int \mathcal{L} dt$): $N_i^{\text{events}} = \sigma_{\text{fid},i} \times \mathcal{L}_{\text{int}}$.

The background expected for *non-b*-jets are taken from simulations and depend on the prediction to be used, if data is compared with $Z + b$ from Sherpa then the background to use will be $Z + c$ and $Z + \text{light}$ from Sherpa as well.

The total background contribution to the final data sample will be estimated using samples of simulated events normalized to their theoretically predicted cross-sections.

A background, which will not be shown but it is important for precision measurement, is associated to the multijet events. These events are estimated using data-driven techniques based on the selection of a fake-enriched data sample, in a same way used at 7 TeV [115]. This sample can be obtained by relaxing or inverting the isolation or identification requirements on the leptons (fake template selection), and subsequently normalizing this template from

data in a control region, which is selected to be similar to the signal region of the measurement, but background-enriched. To extract the normalization, a maximum-likelihood template fit to the data must be used. While rather significant in the W channel, in the Z channel the dual lepton requirement, the tight electron selection and invariant mass cut selection, make much more difficult to pass events having low p_T jets that can mimic electrons in the reconstruction process.

7.2.4 Monte Carlo - Data comparison

The comparison yields of data ($N_{\text{data}}^{\text{events}}$) must be done with the sum of signals ($N_{\text{signal}}^{\text{events}}$) and background ($N_{\text{bkg}}^{\text{events}}$). These comparison will be shown in the next control plots, useful to check the balance between data, expected signal and backgrounds.

Figure 7.11: Preliminary comparisons Pred./Data for Z boson invariant mass (left) and transverse momentum (right) distributions on $Z+b$ -jets analysis at 13 TeV (Status: June 2017).

Fig. 7.11 presents the invariant mass and transverse momentum of the Z boson comparing predictions and data. The invariant mass distribution is very well modeled for all generators, but one can see that regarding the transverse momentum, the modeling still has room for improvements for both generators.

As anticipated, the background expected for *non-b*-jets are taken from simulations and depends on the prediction to be used. If data is compared with

$Z + b$ from Sherpa, then the background to use will be $Z + c$ and $Z + light$ from Sherpa too. The same logic is applied if data is compared with $Z + b$ from Alpgen.

Tab. 7.10 presents the detector-level yields, including yields predicted by Alpgen and Sherpa for each analysis selection.

	Yields		
	$\geq 1\text{-jet}$	$\geq 1b\text{-jet}$	$\geq 2b\text{-jets}$
Data	3808864	264316.0	19662.0
$Z + b$ [Alpgen]	111342.5	85161.1	11852.9
$Z + c$ [Alpgen]	354635.9	78843.3	2667.7
$Z + light$ [Alpgen]	3151704.0	58814.0	499.5
$Z + b$ [Sherpa]	156559.4	116609.6	11892.9
$Z + c$ [Sherpa]	292852.8	65412.6	2169.7
$Z + light$ [Sherpa]	3314130.5	71162.5	1059.1
$t\bar{t}$	7723.7	7019.7	3304.5
Dibosons	11420.0	2271.4	384.2
single-top	743.6	566.6	101.2
$Z(\rightarrow \tau\tau)$ [Alpgen]	476.8	14.8	1.9
W [Alpgen]	69.5	19.0	0.0
$Z(\rightarrow \tau\tau)$ [Sherpa]	558.3	16.7	2.2
W [Sherpa]	7.3	0.5	0.0

Table 7.10: *Detector-level yields, including yields predicted by Alpgen and Sherpa for each analysis selection. The normalized backgrounds are also shown.*

Fig. 7.12 presents the control plots for jets and b -jets exclusive and inclusive multiplicities. For jet multiplicity, it can be seen that Sherpa agrees with data up to four jets while Alpgen underestimate the data with the increase of the number of jets. With respect to b -jets, Sherpa agrees with data for the $1b$ -jet and $2b$ -jets bin. Meanwhile Alpgen agrees with data only in the $2b$ -jets bin, and this underestimation was already expected from the results on truth studies (Fig. 7.6).

Control plots for the kinematic distribution for full jets and b -jets are presented in Fig. 7.13, Fig. 7.14, Fig. 7.15, Fig. 7.16 and Fig. 7.17.

Figure 7.12: Comparisons Pred/Data for the exclusive (top) and inclusive (bottom) multiplicity distribution for jets (left) and b-jets (right) on $Z + b$ -jets analysis at $\sqrt{s}=13$ TeV ($Z \rightarrow ee$) (Status: July 2017).

Figure 7.13: Comparisons $Pred/Data$ for the transverse momentum and rapidity for leading (top) and sub-leading (bottom) jets on $Z + b$ -jets analysis at $\sqrt{s} = 13$ TeV ($Z \rightarrow ee$) (Status: May 2017).

Figure 7.14: Comparisons Pred/Data for the angular distributions ($\Delta\phi_{jj}$, Δy_{jj} and ΔR_{jj}) and invariant mass (M_{jj}) for di-jets on $Z + b$ -jets analysis at $\sqrt{s}=13$ TeV ($Z \rightarrow ee$) (Status: May 2017).

Figure 7.15: Comparisons Pred/Data for the transverse momentum and rapidity comparisons for leading (top) and sub-leading (bottom) b-jets on $Z + b$ -jets analysis at $\sqrt{s} = 13$ TeV ($Z \rightarrow ee$) (Status: May 2017).

Figure 7.16: Comparisons Pred/Data for the angular distributions ($\Delta\phi_{bb}$, ΔY_{bb} and ΔR_{bb}) and invariant mass (M_{bb}) for di-jets on $Z + b$ -jets analysis at $\sqrt{s}=13$ TeV ($Z \rightarrow ee$) (Status: May 2017).

Figure 7.17: Comparisons Pred/Data for the transverse momentum and rapidity for the Z boson in events with at least 1b-jet (top) and at least 2b-jets (bottom) on $Z + b$ -jets analysis at $\sqrt{s}=13$ TeV ($Z \rightarrow ee$) (Status: May 2017).

7.3 The unfolding procedure: Correction to particle-level

In order to compare the prediction between generators and data at the particle level, it is necessary to unfold the effect of the reconstruction from the data.

The unfolding corrections take into account the inefficiency and resolution effects on the reconstruction of jets and bosons. A solution for the data unfolding is to apply a bin-by-bin correction, where is determined the ratio of the MC particle-level (truth-level) distribution to the MC detector-level (reco-level), and this ratio is used bin-by-bin to correct the measured observable. This approach, however, presents limitations associated to the bin migration, expressed in 2D histograms for measurements at detector- and particle-level (response matrix), and for this reason is only valid if these migrations are small.

A common approach to take into account the effect of bin migration is the iterative bayesian unfolding technique [123], implemented in the RooUnfold package [138]. This technique uses the Bayes' theorem to perform the unfolding, treating the response matrix as a description of the probability of observed data given the true distribution.

The reconstructed events are required to fulfill the boson and the jet selections at truth-level. The selection applied truth-level is shown in Tab. 7.11.

Requirement	Cut
Lepton p_T	$p_T^\ell > 27 \text{ GeV}$
Lepton η	$ \eta^\ell < 2.5$
Dilepton mass	$76 < m_{\ell\ell} < 106 \text{ GeV}$
Jet p_T	$p_T^{\text{jet}} > 20 \text{ GeV}$
Jet rapidity	$ \eta^{\text{jet}} < 2.5$
Jet-lepton separation	$\Delta R(\ell, \text{jet}) > 0.4$

Table 7.11: Definition of the phase space for the fiducial cross section for the $Z + b$ -jet at truth-level.

The unfolding procedure must be independent of the generator. On this study it was proposed to use Sherpa for the reasons depicted in the control plots from the previous section. Sherpa is also the sample generated with the larger number of simulated events, which will provide a better statistical uncertainty.

7.3.1 Response matrix

The response matrix takes into account the bin migration effect:

$$M_{ij} \equiv M(R_i, T_j)$$

where, M_{ij} represents the conditional probability that the effect R_i is produced by the cause T_j .

The goal is to determine the probability $M(T_j, R_i)$ using the Bayes' theorem:

$$M(T_j, R_i) = \frac{M(R_i, T_j) \cdot P_0(T_j)}{\sum_{k=1}^{N_T} M(R_i, T_k) \cdot P_0(T_k)}$$

where $M(T_j, R_i)$ is the probability that the effect T_j is produced by the cause R_i .

The first iteration consists of solving the matrix problem using the MC truth distribution ($P_0(T_j)$) as the prior distribution to determine a first estimation of the unfolded data distribution:

$$T_j = \sum_{i=1}^{N_R} M(T_j, R_i) R_i$$

For subsequent iterations, the result from the previous iteration is used as the prior. As the number of interactions increases, the dependence on the MC truth distribution decreases. However, at the same time, as the dependence decreases, the statistical uncertainty increases. Thus, the number of iterations serves to balance the strength of the MC dependence with the size of the oscillations. The optimal number of iteration found was of $N_{it} = 1$ for multiplicity histograms and $N_{it} = 2$ for kinematic distributions.

Fig. 7.18 presents the response matrices corresponding to the exclusive jet and b -jet multiplicity. The exclusive jet response matrix tends to be more diagonal but for b -jets it was found a non-diagonal behavior, being this the main reason to use the bayesian method instead the bin-by-bin unfolding.

Along the unfolding procedure it will be analyzed the exclusive multiplicities, which will be used to produce the inclusive ones. More matrices responses and comparisons with Alpgen are presented in App. F.1.

The response matrix only accounts reconstructed events matched to truth

Figure 7.18: Response matrices corresponding to the exclusive jet (left) and b -jet multiplicity (right) obtained with Sherpa.

events. A matching between truth and reco events means that one Z has been reconstructed at reco-level and one truth Z has been built from two truth leptons at truth-level (in the fiducial acceptance range). This criteria is used on jet and b -jet exclusive multiplicity.

The matching criteria for leading jet distributions is to consider at least one reconstructed jet and at least one truth jet. For sub-leading jet distributions, $\Delta\phi_{jj}$, Δy_{jj} , ΔR_{jj} and M_j , it is required at least two reconstructed jets and at least two truth jets.

The matching criteria is extended for leading b -jet distributions, considering at least one b -tagged jet and at least one truth b -jet. For sub-leading b -jet distributions, $\Delta\phi_{bb}$, Δy_{bb} , ΔR_{bb} and M_{bb} is required at least two b -tagged jets and at least two truth b -jets.

7.3.2 Reconstruction efficiency

Some Z +jets events within the acceptance at truth-level may fail the reconstruction due to the lepton identification inefficiencies and Z boson reconstruction. The reconstruction efficiency on the unfolding procedure is given by:

$$\varepsilon = \frac{N_{\text{reco}}^{\text{match}}}{N_{\text{truth}}}$$

where $N_{\text{reco}}^{\text{match}}$ are the reconstructed events that pass both the fully reconstructed and the truth selection (matched) and N_{truth} represents the truth events. The efficiency is obtained from the projection of the response matrices along the axis containing the truth variables (reconstructed matched events) divided by the truth distribution.

Fig. 7.19 presents the efficiency reconstruction as a function of the exclusive jet and b -jet multiplicity. As can be seen, the full jet multiplicity efficiency is around 40% for all jets. While for the b -jet multiplicity, the efficiency for 1 b -jet is around 25% and 40% for 2 b -jets. More efficiency plots for other kinematic variables can be found in App. F.2.

Figure 7.19: Efficiency reconstruction as a function of the exclusive jet and b -jet multiplicity.

7.3.3 Fake correction

Events associated to fake objects can pass the reconstruction and fall in the truth selection. Then, it is necessary to apply a fake correction, which is defined as the ratio of distribution of matched events divided by the reconstructed distribution (regardless the truth selection):

$$f_{\text{corr}} = \frac{N_{\text{reco}}^{\text{match}}}{N_{\text{reco}}}$$

f_{corr} must be applied on the reconstructed distribution, and represents and small correction in the multiplicity distributions (for full jets and b -jets) as

can be seen in Fig. 7.20. Fake correction plots for several other kinematic variables can be found in App. F.3.

Figure 7.20: Fake correction as a function of the exclusive jet and b -jet multiplicity.

7.3.4 Closure test

Closure tests are performed to validate the unfolding procedure and it is expected the same distribution for the reconstruction after the unfolding and the truth ones (UnfoldedReco/Truth). The fully reconstructed samples are employed to emulate the data, keeping the unfolding inputs (response matrix, efficiency and fake correction) unchanged.

Fig. 7.21 presents the closure tests of the exclusive jet and b -jet multiplicity, showing the validation of the procedure and the considerable correction applied to the reconstruction. More closure tests on several other kinematic variables can be found in App. F.4.

7.3.5 Unfolding test

Finally, the unfolding tests using data are performed to compare the unfolding procedure and the results showed in the control plots (Sec. 7.2.4). Here it is expected the same trends comparing ratios (Truth/UnfoldedData and MC/Data). The data is unfolded with Sherpa keeping the unfolding inputs unchanged.

Fig. 7.22 presents the unfolding tests of the exclusive jet and b -jet multiplicity, showing the validation of the procedure compared with the results in

Figure 7.21: Closure test for the exclusive jet and b-jet multiplicity.

the control plots, through the same trends comparing their ratios.

Figure 7.22: Unfolding test for the exclusive jet and b-jet multiplicity comparing Sherpa and Alpgen reconstruction. Data unfolded with Sherpa.

Kinematic distributions for full jets are presented in Fig. 7.23 and Fig. 7.24,

while for b -jets are presented in Fig. 7.25 and Fig. 7.26. Kinematic distributions for Z bosons in events with at least 1 b -jet and at least 2 b -jets are presented in Fig. 7.27.

Figure 7.23: Unfolding test for the leading (top) and sub-leading (bottom) jet p_T and rapidity comparing Sherpa and Alpgen reconstruction. Data unfolded with Sherpa.

Figure 7.24: Unfolding test for $\Delta\phi_{jj}$, Δy_{jj} , ΔR_{jj} and M_{jj} of dijets comparing Sherpa and Alpgen reconstruction. Data unfolded with Sherpa.

Figure 7.25: Unfolding test for the leading (top) and sub-leading (bottom) b-jet p_T and rapidity comparing Sherpa and AlpGen reconstruction. Data unfolded with Sherpa.

Figure 7.26: Unfolding test for $\Delta\phi_{bb}$, Δy_{bb} , ΔR_{bb} and M_{bb} of dijets comparing Sherpa and Alpgen reconstruction. Data unfolded with Sherpa.

Figure 7.27: Unfolding test for p_T and $|y|$ of Z bosons in events with at least 1b-jet (top) and at least 2b-jets (bottom) comparing Sherpa and Alpgen reconstruction. Data unfolded with Sherpa.

7.4 Systematic uncertainties

The systematic uncertainty must be included in the unfolding procedure, through modification of correction factors and response matrix. Each independent source of uncertainty is varied successively up, and then down by one standard deviation, and in each case the full analysis chain is repeated. Since the experimental uncertainties are expected to be mainly symmetric, in order to reduce the effects of statistical fluctuations, each uncertainty is made symmetric by averaging the absolute values of the up and down variations. In this section it will be presented the main sources of systematics used on b -jet analysis at 13 TeV.

7.4.1 Jet energy scale

The uncertainty on the jet energy has been derived using the recommendations provided by the Jet/Etmiss group of ATLAS [134]. Systematic uncertainties on the difference between the JES in data and simulation are derived using a variety of studies based on in situ measurements and simulation. These uncertainties are decomposed into 21 independent components (Table 7.12), including those arising from the influence of close-by low-energy jets, the correction for pile-up activity and differences in detector response to *light*-, *gluon*- and *heavy*-jets. Each component is propagated through the analysis chain independently by simultaneously varying the signal and background simulation jet response by $\pm 1\sigma$ (up and down).

JES systematic	Components
Effective number of partons	8
η intercalibration	3
pile-up	4
flavor	2
punch trough	1
single particle high p_T	1
relative non closure (MC)	1
BJES response	1

Table 7.12: *JES systematic components.*

Fig. 7.28 presents the breakdown of the JES systematics for exclusive jet and b -jet multiplicity showing their components. In the full jet case the main component is associated to the flavor composition, followed by the pile-up ρ

Figure 7.28: Breakdown of the jet energy scale systematics for exclusive jet and b -jet multiplicity.

topology, where ρ is the median transverse momentum density associated to area suppression (Sec. 4.5.1). With respect to b -jets, the main component is the pile-up ρ topology followed by the effective number of partons (NP_1).

7.4.2 B-tagging systematic: efficiency and mistag rates

Calibration factors are applied to the jet b -tagging efficiency in simulation to match that measurement in data for each flavor. These have associated systematic uncertainties as a function of jet p_T . For b -jets, the uncertainties are derived from calibration analysis divided into 12 components corresponding to the eigenvectors which diagonalize the associated covariance matrix between flavors (Table 7.13). Each component is then varied independently by $\pm 1\sigma$.

B-Tagging systematic	Components
Light	5
C	4
B	3

Table 7.13: B -Tagging systematic components.

Fig. 7.29 presents the breakdown of the b -tagging efficiency and mistag rates systematics for exclusive jet and b -jet multiplicity showing their components. In this case the b -tagging systematics do not affect the full jet multiplicity, but it is the main source of uncertainties on b -jet measurements.

Figure 7.29: Breakdown of the b -tagging efficiency and mistag rates systematics for exclusive jet and b -jet multiplicity. The left plot is presented as reference, since is not considered on full jet analysis.

7.4.3 Electron systematic: efficiency, energy scale and resolution

The electron uncertainties are taken from the electron candidates, measured in $Z \rightarrow ee$ events, used to correct the simulation and following the recommendations provided by the E/gamma performance group of ATLAS. The components are shown in Table 7.14. The systematic contribution is minimal as can be seen in Fig. 7.30, where the Y axis was zoomed 20 times with respect than the previous ones in order to see the differences between contributions.

Electron systematic	Components
Energy scale	1
Energy resolution	1
Identification efficiency	1
Trigger efficiency	1
Reconstruction efficiency	1
Isolation efficiency	1

Table 7.14: Electron systematic components.

This work describes the main sources of systematics only. Other sources of systematics are discussed in [115] ($Z + b$ -jets analysis in proton-proton collisions at 7 TeV in ATLAS).

Fig. 7.31 and Tab. 7.15 present details of the total systematic uncertainty

Figure 7.30: Breakdown of the electron systematic for exclusive jet and b -jet multiplicity. Note the changes in the scale for the Y axis.

and their main sources. The total systematic for full jet multiplicity depends only on the JES and electron systematic, the first one with all contributions. We are interested in the total systematic for b -jet multiplicity, then, for $1b$ -jet the total systematic is around 3.7% and for $2b$ -jets is 4.9%.

Figure 7.31: Total systematic uncertainty and their main sources for exclusive jet and b -jet multiplicity.

The systematic uncertainties for the inclusive multiplicity are constructed propagating the uncertainties for the exclusive ones.

Source of uncertainty	$Z + 1b\text{-jet}$ [%]	$Z + 2b\text{-jets}$ [%]
b -tagging systematic	3.2	4.1
JES systematic	2.0	2.7
electron systematic	0.15	0.19
Total	3.7	4.9

Table 7.15: Summary of the systematic uncertainties determined for the cross-section measurements of the $Z + 1b\text{-jet}$ and $Z + 2b\text{-jets}$ final states.

Figure 7.32: Total systematic uncertainty and their main sources for p_T and $|y|$ for the leading (top) and sub-leading (bottom) jet.

Figure 7.33: Total systematic uncertainty and their main sources for $\Delta\phi_{jj}$, Δy_{jj} , ΔR_{jj} and M_{jj} of di-jets.

Figure 7.34: Total systematic uncertainty and their main sources for p_T and $|y|$ for the leading (top) and sub-leading (bottom) b-jet.

Figure 7.35: Total systematic uncertainty and their main sources for $\Delta\phi_{bb}$, Δy_{bb} , ΔR_{bb} and M_{bb} of di-jets.

Figure 7.36: Total systematic uncertainty and their main sources for p_T and $|y|$ of Z bosons in events with at least 1b-jet (top) and at least 2b-jets (bottom).

7.5 Results for $Z + b$ -jets analysis at 13 TeV

This section will present the results after the unfolding procedure and the inclusion of systematic uncertainties in the unfolded plots.

The expected results and differences for $Z + b$ -jets were discussed along this chapter, but in conclusion the analysis presents a good agreement between data and MC within the experimental uncertainties, as can be seen in Fig. 7.37, where is shown the measured cross section as a function of the exclusive and inclusive full-jet and b -jet multiplicity. Details about $Z + b$ -jets cross sections are presented in Table 7.16.

Channel	Generator	$\sigma(Z + 1b\text{-jet})$ [fb]	$\sigma(Z + 2b\text{-jets})$ [fb]
$Z \rightarrow ee$	Data	$8652 \pm 165 \pm 323$	$1130 \pm 18 \pm 55$
	Sherpa	8600 ± 25	1098 ± 6
	Alpgen	5663 ± 22	1065 ± 9
Channel	Generator	$\sigma(Z + \geq 1b\text{-jet})$ [fb]	$\sigma(Z + \geq 2b\text{-jets})$ [fb]
$Z \rightarrow ee$	Data	$9808 \pm 166 \pm 328$	$1157 \pm 18 \pm 55$
	Sherpa	9723 ± 26	1123 ± 6
	Alpgen	6745 ± 24	1082 ± 9

Table 7.16: Unfolding data and predictions results for the exclusive and inclusive b -jet cross-sections for analysis at $\sqrt{s}=13$ TeV. In data, the statistical uncertainty is quoted first and second the systematic. For predictions, only the statistical uncertainty is quoted.

We also present kinematic distributions for full-jets (Fig. 7.38 and 7.39), in order to compare them with the official results (Z +jets analysis at 13 TeV), those comparison are shown in App. G. The main difference appears on the p_T distributions and, although Alpgen and Sherpa agree, the predictions are not well modeled at high p_T .

Kinematic distributions for b -jets are presented in Fig. 7.40 and 7.41. For the leading jet p_T distribution, Sherpa tends to agree with the unfolded data in a better way than Alpgen, effect associated to the flavor scheme. For the sub-leading p_T distribution both predictions agree, although they are not well modeled at high p_T . Spatial distributions for di-jets ($\Delta\phi_{bb}$, Δy_{bb} and ΔR_{bb}) agree with data within the experimental uncertainties for both generators. It should be noticed that for high values of M_{bb} , there is a disagreement between data and MC for both Sherpa and Alpgen.

Figure 7.37: Measured cross section as a function of the exclusive (top) and inclusive (bottom) jet (left) and b-jet (right) multiplicity at $\sqrt{s}=13$ TeV. The data are compared to the predictions by Sherpa 2.2 and Alpgen. The error bars indicate the statistical uncertainty, and the hatched bands the statistical and systematic uncertainties on data added in quadrature.

Figure 7.38: Measured cross section as a function of the transverse momentum (p_T) and rapidity ($|y|$) for leading (top) and sub-leading (bottom) jets. The data are compared to the predictions by Sherpa 2.2 and Alpgen. The error bars indicate the statistical uncertainty, and the hatched bands the statistical and systematic uncertainties on data added in quadrature.

Figure 7.39: Measured cross section as a function of the angular distributions ($\Delta\phi_{jj}$, $|\Delta y_{jj}|$ and ΔR_{jj}) and invariant mass (M_{jj}) for dijets. The data are compared to the predictions by Sherpa 2.2 and Alpgen. The error bars indicate the statistical uncertainty, and the hatched bands the statistical and systematic uncertainties on data added in quadrature.

Figure 7.40: Measured cross section as a function of the transverse momentum (p_T) and rapidity ($|y|$) for leading (top) and sub-leading (bottom) b-jet. The data are compared to the predictions by Sherpa 2.2 and Alpgen. The error bars indicate the statistical uncertainty, and the hatched bands the statistical and systematic uncertainties on data added in quadrature.

Figure 7.41: Measured cross section as a function of the angular distributions ($\Delta\phi_{bb}$, Δy_{bb} and ΔR_{bb}) and invariant mass (M_{bb}) for di b-jets. The data are compared to the predictions by Sherpa 2.2 and Alpgen. The error bars indicate the statistical uncertainty, and the hatched bands the statistical and systematic uncertainties on data added in quadrature.

Figure 7.42: Measured cross section as a function of the transverse momentum (p_T) and rapidity ($|y|$) for the Z boson in events with at least 1b-jet (top) and at least 2b-jets (bottom). The data are compared to the predictions by Sherpa 2.2 and Alpgen. The error bars indicate the statistical uncertainty, and the hatched bands the statistical and systematic uncertainties on data added in quadrature.

Fig. 7.42 presents the p_T and $|y|$ distribution for Z bosons in events with at least $1b$ -jet and at least $2b$ -jets. The results present a better agreement for events with at least $2b$ -jets (p_T and $|y|$). The main difference appears for Z p_T with at least $1b$ -jet, where both predictions present some disagreement with data.

Before to continue the discussions, an important comment must be done. As can be seen, the p_T distributions for leading and sub-leading b -jet are not well modeled at high p_T , but at the same time, the uncertainty is higher in that region. Improvements in the modeling can be done increasing the number of events, and that can be possible increasing the luminosity, but as was discussed in the first chapter, this increasing introduces limitations on the objects reconstruction and calibration. A first attempt to solve that will be discussed in the next chapter.

Channel $Z \rightarrow \ell\ell$ (7 TeV)	$\sigma(Z + \geq 1b\text{-jet})$ [fb]	$\sigma(Z + \geq 2b\text{-jets})$ [fb]
Data	$4820 \pm 60^{+360}_{-380}$	$520 \pm 20^{+74}_{-72}$
SHERPA 5FNS $\otimes$ CT10	3770 ± 10	422 ± 2
ALPGEN+HJ 4FNS $\otimes$ CTEQ6L1	2580 ± 10	317 ± 2
Channel $Z \rightarrow ee$ (13 TeV)	$\sigma(Z + \geq 1b\text{-jet})$ [fb]	$\sigma(Z + \geq 2b\text{-jets})$ [fb]
Data	$9808 \pm 166 \pm 328$	$1157 \pm 18 \pm 55$
SHERPA 5FNS $\otimes$ NNPDF3.0NNLO	9723 ± 26	1123 ± 6
ALPGEN+PY 4FNS $\otimes$ CTEQ6L1	6745 ± 24	1082 ± 9

Table 7.17: Unfolded data and predictions results for the inclusive b -jet cross-sections for analysis at 7 and 13 TeV. At 7 TeV is shown the combined analysis, considering electron e muons. In data, the statistical uncertainty is quoted first and second the systematic. For predictions, only the statistical uncertainty is quoted.

Tab. 7.17 and Fig. 7.43 present the comparisons with the ATLAS results at 7 TeV, shown in Sec. 6.5 (kinematic comparisons are shown in App. H). As expected, There is a better agreement between data and prediction for collisions at 13 TeV. It can be seen that the five-flavor number scheme in Sherpa provides better results for both cases: events with at least $1b$ - and $2b$ -jets. This shows the importance to take into account the contributions of the diagrams in Fig. 6.2 (b -quarks in the initial-states) for events with at least $1b$ -jet. It is also important to notice that for events with at least $2b$ -jets both generators tend to agree with the data, reflecting improvements on the accuracy of the predictions.

(a)

(b)

Figure 7.43: Cross-sections comparisons for $Z + \geq 1b\text{-jet}$ (a) and $Z + \geq 2b\text{-jets}$ (b) measured in ATLAS for collisions at 7 and 13 TeV. The data measurement is shown as a vertical black line with the inner blue shaded band showing the corresponding statistical uncertainty and the outer green shaded band showing the sum in quadrature of statistical and systematic uncertainties.

As a final comment, it is important to explain that the predictions must include theoretical uncertainties, such as uncertainties associated to the choice

of $\alpha_S(m_Z)$, PDF fits, NNLO scale factors, etc. [115]. This is not yet available for these generators (responsible by the Physics Modeling Group (PMG) of ATLAS) and will be done together with the W channel predictions.

Chapter 8

Measurement of Electrons with Supercells

In the last chapter was presented studies of an important measurement in the standard model, but was also presented that improvements can be possible if the luminosity is increased. At high luminosity, limitations on the object reconstruction and calibration are introduced, limitations associated principally to the pile-up. The high pile-up environment is a challenging condition for the LHC experiments in the next years. Already in 2016, the LHC was able to deliver a peak luminosity of $13.8 \times 10^{33} \text{ cm}^{-2}\text{s}^{-1}$ with an average $\mu = 40$. The LHC experiments are already preparing their detectors and associated systems to cope with an average $\mu = 80$ or even higher. In ATLAS, the main upgrades will be related to the trigger system, with the introduction of a new high granularity digital trigger for the liquid argon electromagnetic calorimeter. Since 2016, the electron trigger threshold at L1 has been raised to keep the rate under control, as low p_T jets, copiously produced in high luminosity hadron colliders can be falsely identified as electron by the current L1 trigger for electrons (see the previous discussion on the electron trigger threshold in Sec. 2.4.1).

In this chapter it will be discussed the aspects of the analysis involving the new trigger segmentation proposal for the electromagnetic calorimeter of ATLAS, since one of the major concerns of the high luminosity program is to maintain the current performance of the first level trigger in ATLAS (L1) at high luminosity.

8.1 Electron reconstruction using supercells

This section will present the reconstruction of electron using the new segmentation through simulations, comparing truth electrons (provided by the generator) and the reconstructed electrons under different levels of pile-up.

8.1.1 The new segmentation for triggers

Several studies indicated that the towers granularity of the current calorimeter trigger ($\Delta\eta \times \Delta\phi = 0.1 \times 0.1$) would not be enough to discriminate jets with low transverse momentum and isolated electrons. A solution to this problem is to deploy more sophisticated methods of discrimination used in the high level trigger and offline analysis, which requires a higher granularity information already for the first level trigger. This higher granularity is obtained by reducing the current trigger tower into smaller clusters, denominated supercells. These supercells have a granularity of $\Delta\eta \times \Delta\phi = 0.025 \times 0.1$ in the first and second layer of the electromagnetic calorimeter, and their purpose is to be used at first level trigger. Fig. 8.1 presents a simulation of how an electron, with transverse energy of 70 GeV, is seen by the current trigger towers and by the proposed supercells.

Figure 8.1: Simulation of one electron ($E_T = 70$ GeV) seen by the current L1 trigger towers (a) and by supercells (b) [8].

Table 8.1 details the granularity in the phase space $\Delta\eta \times \Delta\phi$ for the elementary cell, the current trigger tower and the proposed supercell. It is important to mention that the granularity increase are proposed only for the

first (front) and second (middle) layer, where it will be deposited most of the energy coming from electromagnetic objects.

Layer	Elementary Cell	Trigger Tower	Supercell
Presampler (0)	0.025 x 0.1	0.1 x 0.1	0.1 x 0.1
Front (1)	0.003125 x 0.1	0.1 x 0.1	0.025 x 0.1
Middle (2)	0.025 x 0.025	0.1 x 0.1	0.025 x 0.1
Back (3)	0.05 x 0.025	0.1 x 0.1	0.1 x 0.1

Table 8.1: Granularity comparison ($\Delta\eta \times \Delta\phi$) for the elementary cell, the current trigger tower and the proposed supercell on the LAr EM barrel calorimeter.

8.1.2 MC samples and selections

The samples used on this study are based on $Z \rightarrow ee$ simulation, considering proton-proton collisions at $\sqrt{s} = 14$ TeV (100000 events). These samples were produced with the POWHEG [141] event generator interfaced to the PYTHIA [97] parton-shower, using the AU2CT10 [25] PDF set. For comparisons at different pile-up levels, the samples were derived under the next pile-up conditions: $\mu = 0, 20, 40, 60$ and 80 .

The initial process of the analysis consists of selecting truth electron distributions which will be compared later with the supercell results (quick sample test). The analysis is applied over selected truth electrons with $E_T > 20$ GeV and $|\eta| < 2.5$. The η cut is applied following the acceptance of the current electron reconstruction system in ATLAS. The E_T cut is chosen considering experimental criteria of efficiency in p_T distribution of leptons from W and Z decays.

Fig. 8.2.a shows the number of electrons per event or multiplicity (although in the $Z \rightarrow ee$ we expect at least two electrons in the events, due to the kinematic cuts some of these electrons are missed). Fig. 8.2.b presents the transverse energy distribution of the simulated electrons and Fig. 8.2.c and 8.2.d present the pseudorapidity and azimuthal angle, respectively.

Fig. 8.3 shows the invariant mass and p_T of the Z boson at truth level in events with at least two electrons within the kinematic cuts.

Figure 8.2: Distribution of the number of electrons produced by event (a), the transverse energy (b), pseudorapidity (c) and azimuthal angle (b) for electrons at truth level coming from $Z \rightarrow ee$ at $\sqrt{s}=14$ TeV.

Figure 8.3: Distribution of the invariant mass (a) and p_T (b) of the Z boson at truth level coming from $Z \rightarrow ee$ at $\sqrt{s}=14$ TeV.

8.1.3 Electron reconstruction

Once we have the behavior of the truth electrons, we now can look into the reconstructed calorimeter information, where the particles produced by the event generator are transported and their interaction with the matter are simulated. This is done by GEANT4 [147], integrated into the ATLAS simulation framework.

At this level it will be considered each cell information simulated in the electromagnetic calorimeter. The search begins in a similar way used by the current L1 trigger (Sec. 2.4.1), but now deploying techniques from the off-line electron reconstruction (Sec. 3.1). That said, the following steps are performed:

Step 1: The algorithm begins by identifying each cell and the information they have: the calorimeter and layer where is positioned, and the E_T , η and ϕ .

Step 2: Search for cells with $E_T > 3$ GeV, which will be called *seed* candidates.

Step 3: The electron will be reconstructed from seeds using a window algorithm [79], with a window of 3×3 in units of $\Delta\eta \times \Delta\phi = 0.025 \times 0.1$ (Fig. 8.4).

Step 4: Finally, the values for E_T , η and ϕ are obtained for the reconstructed electron using supercells (hereinafter called "super-electron"):

$$E_T^{\text{SE}} = \sum_i E_{T,i}^{\text{cell}}$$

$$\eta^{\text{SE}} = \frac{\sum_i \eta_i^{\text{cell}} E_{T,i}^{\text{cell}}}{E_T^{\text{SE}}}$$

$$\phi^{\text{SE}} = \frac{\sum_i \phi_i^{\text{cell}} E_{T,i}^{\text{cell}}}{E_T^{\text{SE}}}$$

where $E_{T,i}^{\text{cell}}$, η_i^{cell} and ϕ_i^{cell} are the transverse energy, pseudorapidity and azimuthal angle for the i -th cell, respectively. These values are important in order to correlate them with the truth distributions.

Fig. 8.4 presents the approach used by the electron finding algorithm, showing a seed and its neighbors considered in the electron reconstruction. In order to avoid double seed counting belonging to the same electron, if a

neighbor has an energy higher than the seed, the seed candidate is rejected and is added as a neighbor of the most energetic seed (turning the highest energy cell the cluster seed).

Figure 8.4: Scheme of the finding algorithm used on the reconstruction of electrons with supercells, where is observed how the information is collected from neighbors to the seeds found in different layers of the electromagnetic calorimeter in a window 3×3 in units of $\Delta\eta \times \Delta\phi = 0.025 \times 0.1$.

8.1.4 Comparison with truth

The reconstruction procedure is firstly tested in an environment without pile-up and compared with the truth electrons.

Fig. 8.5.a shows the number of electrons per event where we can observe a relative agreement between super-electrons and the truth electrons. A larger difference is seen in Fig. 8.5.b (transverse energy distribution). This difference can be related to the fact that the super-electrons do not include calibration criteria. Fig. 8.2.c and 8.2.d present the pseudorapidity and azimuthal angle, respectively. For both cases, the results are in good agreement between super-electrons and the truth electrons. In Fig. 8.2.c, it can be seen the effect of the crack region for super-electrons.

The number of super-electrons and spatial distributions agree with truth, but with respect to transverse energy a further calibration must be included. Evidently, this affects the Z boson reconstruction, as can be seen in Fig. 8.6, where is presented the invariant mass and p_T of the Z boson in events with at least two electrons in the kinematic cuts. This work will focus on electron

Figure 8.5: Distribution of the number of electrons produced by event (a), the transverse energy (b), pseudorapidity (c) and azimuthal angle (b) comparing super-electrons and electrons at truth level coming from $Z \rightarrow ee$ at $\sqrt{s}=14$ TeV.

Figure 8.6: Distribution of the invariant mass (a) and p_T (b) of the reconstructed Z boson with super-electrons and electrons at truth level ($Z \rightarrow ee$ at $\sqrt{s}=14$ TeV).

reconstruction, as the calibration is often applied offline, being less important in the first level trigger.

8.1.5 Super-electrons and pile-up suppression

In order to evaluate the pile-up effects, $Z \rightarrow ee$ samples were generated at different μ values: 0, 20, 40, 60 and 80.

Fig. 8.7 presents the energy distribution in the $\eta \times \phi$ space for one event containing one candidate to electron for the different layers of the electromagnetic calorimeter at $\mu = 80$. The figure shows the activity associated to pile-up and five possible seeds close to $\eta = 0.95$ and $\phi = -2.75$ (one in layer 0 and two in layers 1 and 2).

Figure 8.7: Energy distribution in the supercells for the different layers of the electromagnetic calorimeter in an event with one electron candidate ($E_T = 45$ GeV, $\eta = 0.95$ and $\phi = -2.75$), based on a simulation of $Z \rightarrow ee$ with $\mu = 80$.

Fig. 8.8.a shows the energy distribution of seeds found for different values of μ . As expected, there is a large number of seeds at low energies with the increasing of μ , caused by low p_T jets. Fig. 8.8.b compares the energy distribution for super-electrons, showing the peak of fake electrons and how it increases with the increasing of μ .

A first approximation to suppress fake electrons will be applied cuts on E_T^{seed} or E_T . Then, in order to study the effect of pile-up as a noise source,

Figure 8.8: Energy distribution of seeds (a) and super-electron (b) for different values of μ .

it was used a sample based on neutrino simulation, considering proton-proton collisions at $\sqrt{s} = 14$ TeV and $\mu = 80$, generated with the PARTICLEGENERATOR [142] package of ATLAS. As neutrinos do not interact with any part of the ATLAS detector, the reconstruction will only present pile-up information.

Figure 8.9: Distribution of transverse energy for super-electrons and fake electrons. without the E_T^{seed} cutting (a) and with cutting (b).

Fig. 8.9.a shows a comparison between the distribution of transverse energy of super-electrons and fake electrons measured in the $Z \rightarrow ee$ and neutrino sample. Fig. 8.9.b shows that, applying a E_T cut in the seed, most of the fake electrons can be rejected using the higher granularity from the super cell. This indicates that a significant amount of pile-up can be already rejected

taking a cut in the E_T of the seed (8 GeV).

The method proposed before results useful as a simple option to suppress fake electrons (maintaining the E_T shape), which are generated from low p_T jets. However, the main complication arises when is taken into account that the production of those fake electrons can be up to 10^6 times greater than real electrons. This will require more sophisticated methods of pile-up suppression such as the shower shape discriminants, which will be discussed in the next sections.

8.2 Super-electrons analysis

This section will present the comparison between the super-electron reconstruction and the information provided by the off-line simulated reconstruction. A new batch of simulation samples was produced including the L1 trigger and the off-line electron information. For trigger rate and efficiency studies, jet samples were included in the analysis and will be further discussed in the next sections.

8.2.1 MC Samples and selections

For this study it will be used as signal a simulation of $Z \rightarrow ee$, similar to the one used in the previous section. Now, considering proton-proton collisions at $\sqrt{s} = 13$ TeV and $\mu = 80$, generated with POWHEG interfaced to PYTHIA 8 to describe the parton-shower evolution. It was used the CT10 [25] PDF set and the AZNLOCTEQ6L1 PDF tune [144]. This sample is fully reconstructed, including not only the individual calorimeter cell information but also the trigger and final electron and jets reconstruction.

The fake electron contribution is based on dijet simulation, generated with PYTHIA 8 using the NNPDF23LO [143] PDF set and the A14NNPDF23LO PDF tune [144] (similar samples were used on jet calibration). These samples are sliced accord the p_T of the leading true-jets as is shown in Tab. 8.2. For the analysis, the p_T slices are stitched together, normalized by the cross section.

The analysis was applied over selected electrons with $E_T > 20$ GeV and $|\eta| < 2.5$.

Sample	$p_{T,leading}^{truth}$ [GeV]	Events	σ [nb]
Zee	—	50000	1.901E+00
jetjet_JZ0W	0-20	50000	7.842E+07
jetjet_JZ1W	20-60	50000	5.269E+04
jetjet_JZ2W	60-160	48000	8.093E+02
jetjet_JZ3W	160-400	50000	8.452E-01

Table 8.2: Number of events and cross sections of the simulated samples.

8.2.2 Comparisons with the first level trigger (L1) and off-line electrons

The first step of the analysis consists on the comparison between the super-electrons, the electromagnetic signals produced at L1 trigger and the reconstructed off-line electrons. The strategy is shown in Fig. 8.10

Figure 8.10: Proposed strategy on Super-Electron analysis.

Fig. 8.11.a shows the comparison of the number of reconstructed electrons (event by event) by L1 and by super-electrons with the truth information. One can see that the L1 has many more electrons reconstructed, if compared to the truth, than the super-electrons. At L1, these signals are associated to RoI with low p_T jets causing an increase of the electromagnetic trigger rate. Fig. 8.11.b presents the E_T distributions. Again, we see an excess of low p_T for L1 signals compared to the truth electron E_T distribution, and how the higher granularity provided by the supercell scheme mitigates this effect. Even though a cut of 20 GeV is applied in the electron E_T , an important excess is still observed in the L1 electron reconstruction, while for super-electrons it remains much closer

to the truth electron E_T distribution.

Figure 8.11: Comparison of the number of electrons produced per event (a) and transverse energy (b) for super-electrons, the level 1 trigger signals and truth electrons for a sample of $Z \rightarrow ee$ at $\sqrt{s} = 13 \text{ TeV}$ ($\mu = 80$).

Another important aspect is that the L1 trigger E_T distribution presents a saturation at $\sim 128 \text{ GeV}$, which can increase the uncertainty on energy determination. This, however, is less important at this level, as the L1 trigger is designed to provide only an estimated position of possible electron candidates.

In the case of super-electrons it is important to evaluate the energy reconstruction efficiency, as is shown in Fig.8.12.a, where can be seen how the super-electron efficiency has a well-defined Gaussian shape with an average value around -8% and with a relative better definition than L1. This difference is related to the fact that the super-electrons (or trigger signals) do not include calibration criteria (as it was discussed in the previous section) and that must be taken into account when super-electrons are compared with off-line electrons, as is shown in Fig.8.12.b, where the off-line distribution presents a better definition and a mean value close to 0.

With these results, the gain of increasing the trigger granularity at L1 is very clear, and the next sections will characterize further the performance of the super-electrons, with the off-line reconstructed electron as reference, since this represents the best possible reconstruction.

Fig. 8.13 presents the most relevant kinematic distributions, which are important to evaluate the advantages and limitations of super-electrons if compared to off-line reconstruction. Clearly, given the access to full calorimeter

Figure 8.12: Transverse energy efficiency comparisons: for super-electrons vs. L1 (a) and super electrons vs. off-line electrons (b).

segmentation and advanced calibration and scaling procedures, the off-line reconstruction will outperform the supercell scheme. That is evident in 8.12.b. (E_T distribution) where the off-line reconstruction agrees much better with the distribution of truth electrons.

8.2.3 Rates and efficiencies

The goal of these studies is to perform a detailed evaluation of the rates and efficiencies for the ATLAS supercell scheme in order to reduce the rates of fake electrons in a high luminosity environment. The study of rates and efficiencies is therefore important to quantify the reduction of the rate using the supercell higher granularity information.

We can define the efficiency as:

$$\text{Efficiency} = \varepsilon \equiv \frac{N \text{ Signals/bin (matching with truth)}}{N \text{ Truth electrons/bin}} \quad (8.1)$$

where the matching of signals with truth electrons is defined as $|\eta^{\text{signal}} - \eta^{\text{truth}}| < 0.1$ and $|\phi^{\text{signal}} - \phi^{\text{truth}}| < 0.1$. In Eq. 8.1, signal represents the signal produced by the L1 trigger or the produced by the super-electron. This efficiency is defined per interval of transverse energy, since this is the kinematic variable of highest impact in the selection of electrons.

Figure 8.13: Distribution of number of electrons reconstructed event by event (a), transverse energy (b), pseudorapidity (c) and azimuthal angle (d) for reconstructed super-electrons, reconstructed off-line electrons and truth electrons, for a sample of $Z \rightarrow ee$ at $\sqrt{s} = 13$ TeV.

Fig. 8.14.a presents the electron reconstruction efficiency for super-electron, L1 signal and off-line electrons. In general, the efficiency is higher than 95% at $E_{min}^{truth} > 20$ GeV in all three cases, but slightly better for off-line electrons, meaning that most of electrons are correctly identified.

The rate study is based on jet samples, in order to get the most realistic scenario at high luminosity. For that, the reconstruction based on the super-cell scheme will be applied on the di-jet samples presented before in order to estimate the rate of fake electrons production.

Fig. 8.14.b presents the production rate of electrons (fake electrons), showing that using super-electrons is possible to reduce the rate if compared to L1 signals. It also shows that to keep the current 20 kHz rate it is necessary

Figure 8.14: Reconstruction efficiency ($Z \rightarrow ee$) (a) and rate production (b) distributions as a function of the E_T threshold, for electrons reconstructed from the supercell information, L1 trigger electron signals and off-line electrons.

to apply a cut around 50 GeV, but as the electron E_T distribution is peaked around 40 GeV, applying a cut of 50 GeV will reject most of the electrons from Z decay. Fig. 8.15.a shows this same behavior (now with a cutting at 3 GeV).

Figure 8.15: Distribution of the transverse momentum for super-electrons (a) and jets, identified as fake electrons (b).

Considering the jet samples and applying the same criteria used to reconstruct the electrons from the $Z \rightarrow ee$ sample, we found a considerable amount of jets wrongly reconstructed as electrons (Fig. 8.15.b). This clearly indicates

that we need more information to separate these jets from electron samples, what can be pursued using the higher granularity of the supercell.

8.3 Shower shape discriminants

With the higher transversal and longitudinal granularity implemented by the supercell scheme, we can resort to some of the methods used in the off-line reconstruction that use as base the shower shape of the electromagnetic shower developed inside the calorimeter, namely the rapidity spread and the lateral shower widths, which will be discussed in this section.

8.3.1 Rapidity spreads ($R_{\eta,2}$, $R_{\eta,1}$ and $R_{\eta,1+2}$)

The main discriminant used on off-line analysis is the rapidity spread (R_η), denoted in this work as $R_{\eta,2}$, which takes into account the energy deposited on cell clusters in the second layer (middle layer, where most of the energy is deposited) of the LAr calorimeter. $R_{\eta,2}$ is defined as:

$$R_{\eta,2} \equiv \frac{E_{T,\Delta\eta\times\Delta\phi=0.075\times0.2}^{(2)}}{E_{T,\Delta\eta\times\Delta\phi=0.175\times0.2}^{(2)}}$$

where $E_{T,\Delta\eta\times\Delta\phi}^{(i)}$ represents the transverse energy in the i -th layer over an area of $\Delta\eta \times \Delta\phi$.

Given the fact that we do not have access to the full granularity information as in the off-line reconstruction, it was proposed by this present work to extend the calculation of this discriminant using the first layer (front layer) separately ($R_{\eta,1}$) and in combination with the middle layer ($R_{\eta,1+2}$) as follows:

$$R_{\eta,1} \equiv \frac{E_{T,\Delta\eta\times\Delta\phi=0.075\times0.2}^{(1)}}{E_{T,\Delta\eta\times\Delta\phi=0.175\times0.2}^{(1)}}$$

$$R_{\eta,1+2} \equiv \frac{E_{T,\Delta\eta\times\Delta\phi=0.075\times0.2}^{(1)} + E_{T,\Delta\eta\times\Delta\phi=0.075\times0.2}^{(2)}}{E_{T,\Delta\eta\times\Delta\phi=0.175\times0.2}^{(1)} + E_{T,\Delta\eta\times\Delta\phi=0.175\times0.2}^{(2)}}$$

Fig. 8.16 shows the shape distributions for R_η 's ($R_{\eta,2}$, $R_{\eta,1}$ and $R_{\eta,1+2}$) for electrons and jets samples reconstructed with the electron finding. The distributions are similar to the ones previously presented (Fig. 3.4), which

Figure 8.16: $R_{\eta,2}$, $R_{\eta,1}$ and $R_{\eta,1+2}$ shape distributions for over electrons and jets samples reconstructed with the super-electron finding algorithm.

indicates the validity of this proposed procedure.

The next step now is to find the optimum value for R_η 's cuts in order to preserve the efficiency and to minimize the fake electron reconstruction, keeping the rate low.

To organize the analysis, it is proposed the following strategy for each discriminant, using R_η as example:

- Firstly, maintaining $E_T > 20$ GeV, we will look into the efficiency (as defined by Eq. 8.1) as function of R_η cut. We will then extract the cuts in order to keep the efficiencies at 95%, 90% and 85%. It is important to note that 95% will be the best efficiency possible, but considering the higher rates at high luminosity, we will lower this reference to 90%; with 85% as an option for an extreme scenario.
- While the bandwidth is flexible on the trigger menu, we will keep 20 kHz as a reference in this work. We will then turn our attention to the rate as a function of the R_η cut and compare this cut with the ones obtained before. If the rates obtained are higher than 20 kHz, an increase on the E_T threshold can be considered following the next steps:
 - Fix the R_η cut and look at the efficiency and rates as a function of E_T threshold.
 - Search for the best scenario in order to maintain the rate close to 20 kHz with the minimum E_T threshold (below of 50 GeV).

Figure 8.17: Efficiency (a) and rate (b) signal distributions as a function of the minimum value (cut) for $R_{\eta,2}$, $R_{\eta,1}$ and $R_{\eta,1+2}$ reconstructed with the electron finding.

Fig. 8.17.a shows the efficiency distributions as a function of the minimum value for R_{η} 's. Considering the proposed strategy, the minimum values for different efficiencies were kept fixed as shown in Table 8.3. Fig. 8.17.b presents the rate distributions showing, as expected, a reduction on rates if the minimum value of R_{η} 's increases, with the best results for $R_{\eta,1+2}$, but even with that, the 'best' reduction on rate (186 kHz) is found at $R_{\eta,1+2}^{min} = 0.918$, with an efficiency of 85%. More values are found in Table 8.3. Since this rate is way too high, we must now consider raising the E_T threshold above 20 GeV (trying to keep it as low as possible in order not to affect the kinematic acceptance for electrons from Z decays, as it was shown by Fig. 8.15)¹.

Efficiency ($E_T > 20$ GeV)	Discriminant (Rate [kHz])		
	$R_{\eta,2}^{min}$	$R_{\eta,1}^{min}$	$R_{\eta,1+2}^{min}$
95%	0.715 (~ 637)	0.812 (~ 657)	0.777 (~ 610)
90%	0.848 (~ 322)	0.934 (~ 300)	0.898 (~ 274)
85%	0.873 (~ 240)	0.950 (~ 236)	0.918 (~ 186)

Table 8.3: Minimum values for $R_{\eta,2}$, $R_{\eta,1}$ and $R_{\eta,1+2}$ and their dependence with the efficiency and rates.

¹At the present date, the luminosity condition for 2017 data taking have raised the electron trigger to about 29 GeV and an average $\mu = 40$.

Figure 8.18: Signal efficiency as a function of the E_T threshold with different cuts on $R_{\eta,2}$, $R_{\eta,1}$ and $R_{\eta,1+2}$.

Figure 8.19: Signal rates (and ratios) as a function of E_T threshold for different cuts on $R_{\eta,2}$, $R_{\eta,1}$ and $R_{\eta,1+2}$.

Fig. 8.18 shows the electron efficiency as a function of E_T threshold with different cuts on R_{η} 's. It can be seen how the increase on the threshold improves the efficiency for all three cases ($R_{\eta,2}$, $R_{\eta,1}$ and $R_{\eta,1+2}$).

Finally, in Fig. 8.19 are shown the signal rates as function of the E_T threshold with different cuts on R_{η} 's. The ratio plots clearly show how a reduction of rate can be obtained by using of R_{η} 's. In particular, the proposal of using the layer combination calculation ($R_{\eta,1+2}$) represents the best option, and this recommendation was considered by the Phase 1 firmware development group of ATLAS to be implemented in the trigger processor. Table 8.4 shows

$R_{\eta,1+2}^{min}$	Cuts		Efficiency [%]	Rate [kHz]
	E_T^{min}	[GeV]		
0.777	20		95.0 ± 0.8	611 ± 24
	25		95.2 ± 0.9	292 ± 16
	30		95.4 ± 0.9	151 ± 12
	35		95.5 ± 1.0	85 ± 9
	40		95.8 ± 1.4	49 ± 7
0.898	20		90.1 ± 0.8	271 ± 16
	25		91.2 ± 0.8	141 ± 11
	30		92.0 ± 0.9	72 ± 8
	35		92.6 ± 1.0	41 ± 6
	40		93.0 ± 1.3	26 ± 5
0.918	20		85.1 ± 0.8	183 ± 13
	25		86.9 ± 0.8	99 ± 10
	30		88.5 ± 0.9	49 ± 7
	35		89.6 ± 1.0	26 ± 5
	40		90.8 ± 1.3	17 ± 4

Table 8.4: Efficiency and rate for different values of E_T^{min} and $R_{\eta,1+2}^{min}$.

that the best scenario that could be achieved is $R_{\eta,1+2}^{min} = 0.918$ and $E_T^{min} = 35$ GeV, considering the efficiency around 89.6% with a rate of 26 kHz. Tables for the other discriminants can be found in App. I.

8.3.2 Lateral shower widths ($W_{\eta,1}$ and $W_{\eta,2}$) and f_3

Other discriminants can be explored due to the supercell segmentation scheme. One of this is the lateral shower width ($W_{\eta,2}$) defined in Sec. 3.2 as:

$$W_{\eta,2} \equiv \sqrt{\frac{\sum (E_T^{(2)} \times \eta^2)_{\Delta\eta \times \Delta\phi=0.075 \times 0.2}}{E_{T,\Delta\eta \times \Delta\phi=0.075 \times 0.2}^{(2)}} - \left(\frac{\sum (E_T^{(2)} \times \eta)_{\Delta\eta \times \Delta\phi=0.075 \times 0.2}}{E_{T,\Delta\eta \times \Delta\phi=0.075 \times 0.2}^{(2)}} \right)^2}$$

where $E_{T,\Delta\eta \times \Delta\phi}^{(i)}$ represents the transverse energy in the i -th layer over an area of $\Delta\eta \times \Delta\phi$.

As it was done with $R_{\eta,2}$, it was incorporated a new discriminant, $W_{\eta,1}$, similar to $W_{\eta,2}$ using the first layer (front layer) of the LAr calorimeter instead the second one (middle layer).

Another potential discriminator to be investigated is the fraction energy layer (f_3), which includes the information of the three main layers (front,

middle and back):

$$f_3 \equiv \frac{E_{T,\Delta\eta\times\Delta\phi=0.2\times0.2}^{(3)}}{E_{T,\Delta\eta\times\Delta\phi=0.075\times0.2}^{(1)} + E_{T,\Delta\eta\times\Delta\phi=0.075\times0.2}^{(2)} + E_{T,\Delta\eta\times\Delta\phi=0.2\times0.2}^{(3)}}$$

Fig. 8.20 shows the shape distributions for W_η 's and f_3 over jets and electrons samples reconstructed with the electron finding. Again, the distributions for $W_{\eta,2}$ is similar to the presented on preliminary studies (see Fig. 3.4), showing the consistency on the super-electron reconstruction. It can be seen that $W_{\eta,1}$ provides a slightly better separation than $W_{\eta,2}$ between electrons and jets.

Figure 8.20: Shape distributions for $W_{\eta,1}$, $W_{\eta,2}$ and f_3 over jets and electrons samples reconstructed with the electron finding.

The same strategy developed for R_η before will be deployed here, searching for the optimum values for W_η 's and f_3 in order to maintain a balance between a high efficiency and a low rate.

Fig. 8.21 and Fig. 8.22 show the efficiency and rate signal distributions as function of the maximum values for W_η 's and f_3 , respectively. In particular we can see that f_3 by itself is not able to cross the 20 kHz reference even for limit values (Fig. 8.22.b). Fixing the maximum values for different efficiencies, it was found a better reduction on rates for $W_{\eta,2}$ (Table 8.5), although this result alone is worse if is compared with the one provided by $R_{\eta,1+2}^{min}$.

Figure 8.21: Efficiency (a) and rate (b) signal distributions as a function of the maximum value (cut) for $W_{\eta,1}$ and $W_{\eta,2}$ reconstructed with the electron finding.

Figure 8.22: Efficiency (a) and rate (b) signal distributions as a function of the maximum value (cut) for f_3 reconstructed with the electron finding.

Efficiency ($E_T > 20$ GeV)	Discriminant (Rate [kHz])		
	$W_{\eta,1}^{max}$	$W_{\eta,2}^{max}$	f_3^{max}
95%	0.0129 (~ 656)	0.0159 (~ 477)	0.026 (~ 509)
90%	0.0117 (~ 534)	0.0148 (~ 343)	0.018 (~ 386)
85%	0.0106 (~ 427)	0.0145 (~ 293)	0.015 (~ 318)

Table 8.5: Maximum values for $W_{\eta,1}$, $W_{\eta,2}$ and f_3 , and their dependence with the efficiency and rates.

As it was done for R_η 's, the effect of increasing E_T cut will be evaluated. Fig. 8.23 shows the electron efficiency as function of the E_T threshold with different cuts on W_η 's and f_3 . Again, it can be seen how the increase on the threshold improves the efficiency for all three cases ($W_{\eta,1}$, $W_{\eta,2}$ and f_3).

Figure 8.23: Signal efficiency as a function of the E_T threshold with different cuts on $W_{\eta,1}$, $W_{\eta,2}$ and f_3 .

Finally, Fig. 8.24 shows the signal rates as function of E_T threshold with different cuts on W_η 's and f_3 . The cuts on $W_{\eta,2}$ present better results for rate reduction if compared to $W_{\eta,1}$ and f_3 . Table 8.6 shows that the better scenario would be $W_{\eta,2}^{max} = 0.0145$ and E_T^{min} around 35 GeV, with an efficiency close to 90.4% and a rate of 41 kHz. The tables for the other discriminants can be found in App. I.

Figure 8.24: Signal rates (and ratios) as a function of E_T threshold for different cuts on $W_{\eta,1}$, $W_{\eta,2}$ and f_3 .

$W_{\eta,2}^{max}$	Cuts		Efficiency [%]	Rate [kHz]
	E_T^{min} [GeV]			
0.0159	20		95.2 ± 0.8	478 ± 21
	25		95.6 ± 0.9	214 ± 14
	30		96.0 ± 0.9	109 ± 10
	35		96.2 ± 1.0	63 ± 8
	40		96.4 ± 1.4	41 ± 6
0.0148	20		91.1 ± 0.8	355 ± 18
	25		91.8 ± 0.8	161 ± 12
	30		92.5 ± 0.9	81 ± 9
	35		93.1 ± 1.0	45 ± 6
	40		93.7 ± 1.3	26 ± 5
0.0145	20		87.7 ± 0.8	293 ± 17
	25		88.7 ± 0.8	140 ± 11
	30		89.6 ± 0.9	76 ± 9
	35		90.4 ± 1.0	41 ± 6
	40		91.2 ± 1.3	24 ± 5

Table 8.6: Efficiency and rate for different values of E_T^{min} and $W_{\eta,2}^{max}$.

8.3.3 Other Discriminants

With the inclusion of supercells, new discriminant were proposed by the LAr upgrade group. These discriminants consider the information of the super-electron seed, the second most energetic cell after the seed, and the inclusion of the hadronic calorimeter energy information, such as:

- **Hadronic ratio:** (slightly different to the hadronic leakage presented in Sec. 3.2)

$$R_{\text{had}} \equiv \frac{Had_{\text{core}}}{E_{\text{T}}}$$

where $Had_{\text{core}} = E_{\text{T}, \Delta\eta \times \Delta\phi=0.2 \times 0.2}^{\text{Had.Cal.}}$ represents the transverse energy in the hadronic calorimeter over an area of $\Delta\eta \times \Delta\phi$ (hadronic core).

- **Seed energy ratio:**

$$R_{\text{seed}} \equiv \frac{E_{\text{T}}^{\text{seed}}}{E_{\text{T}}}$$

- **Seed energy ratio to second hottest cell:**

$$R_{2^{\text{nd}}\text{hot}} \equiv \frac{E_{\text{T}}^{2^{\text{nd}}\text{hottest}}}{E_{\text{T}}^{\text{seed}}}$$

E_{T} , $E_{\text{T}}^{\text{seed}}$ and $E_{\text{T}}^{2^{\text{nd}}\text{hottest}}$ represent the transverse energy for the super-electron, its seed and the second most energetic cell after the seed, respectively.

Fig. 8.25 shows the shape distributions for R_{had} , R_{seed} and $R_{2^{\text{nd}}\text{hot}}$ over jets and electrons samples reconstructed with the electron finding. For these cases the results on rate reductions are lower if are compared with the shower shape discriminants shown before, as it can be seen in 8.26, 8.27 and 8.28.

Figure 8.25: Shape distributions for R_{seed} , R_{had} and $R_{2nd_{hot}}$ over jets and electrons samples reconstructed with the electron finding.

Figure 8.26: Efficiency (a) and rate (b) signal distributions as a function of the maximum value (cut) for R_{had} reconstructed with the electron finding.

Figure 8.27: Efficiency (a) and rate (b) signal distributions as a function of the maximum value (cut) for R_{seed} reconstructed with the electron finding.

Figure 8.28: Efficiency (a) and rate (b) signal distributions as a function of the maximum value (cut) for the seed energy ratio to second hottest cell reconstructed with the electron finding.

8.3.4 Correlation between discriminants

Since the calculation involving the discriminants must be done at first trigger level, we cannot afford an overly complicated selection of discriminants since resources and time budget are limited. Considering this, and to avoid the introduction of unnecessary discriminants, a correlation study was performed using all 9 discriminants discussed previously. This process was done with the TMVA package [81], which compare each discriminant distribution with the other ones in order to produce correlation matrices.

Figure 8.29: Correlation matrix between discriminators for signals (top) and background (bottom), represented by jets (produced with the TMVA package).

Fig. 8.29 presents the correlation matrices between discriminators for signal and background. There, 100% represents discriminators strongly correlated, -100% strongly anti-correlated and 0% discriminators completely not correlated. The last studies shown better results for $R_{\eta,1+2}$, and checking the matrices, it can be seen that the discriminants less correlated with $R_{\eta,1+2}$ are: $W_{\eta,2}$ and R_{had} . However, R_{had} presents the worst result on rate reduction if compared to $W_{\eta,2}$, meaning that an inclusion of R_{had} hardly will produce improvement on the results. The conclusion and the final recommendation as a result of this work is that only $R_{\eta,1+2}$ and $W_{\eta,2}$ should be used by the back-end trigger processing system.

Fig. 8.30 shows the electron efficiency as function of E_T threshold combining cuts on $R_{\eta,1+2}$ and $W_{\eta,2}$. The behavior is similar to the ones showed previously: the increase on the threshold improves the efficiency for the combined cases. Fig. 8.31 shows the signal rates as function of E_T threshold with an improvement for the combined cases.

Figure 8.30: Signal efficiency as a function of the E_T threshold combining cuts on $R_{\eta,1+2}$ and $W_{\eta,2}$.

A full set of tables showing the improvement can be found in App. I, and the Table 8.7 summarizes the result. The best situation for supercell processing can be found with the following values: $R_{\eta,1+2}^{\min} = 0.898$, $W_{\eta,2}^{\max} = 0.0148$ and $E_T = 35$ GeV, with an efficiency close to $89 \pm 1\%$ and a rate of 22 ± 5 kHz.

To finalize this study, it is important to check the behavior of the total number of super-electron after applying the discriminant, which can be seen in Fig. 8.32, where is shown the transverse momentum distribution for super-electrons compared with truth electrons. Evidently, the discriminant cuts pro-

Figure 8.31: Signal rates (and ratios) as a function of E_T threshold combining cuts on $R_{\eta,1+2}$ and $W_{\eta,2}$.

Cuts			Efficiency [%]	Rate [kHz]
$R_{\eta,1+2}^{min}$	$W_{\eta,2}^{max}$	E_T^{min} [GeV]		
0.898	0.0159	30	91.0 ± 0.9	50 ± 7
		35	91.6 ± 1.0	30 ± 5
		40	92.2 ± 1.3	19 ± 4
0.898	0.0148	30	88.0 ± 0.9	40 ± 6
		35	89.0 ± 1.0	22 ± 5
		40	89.9 ± 1.3	12 ± 3
0.918	0.0159	30	87.7 ± 0.9	35 ± 6
		35	88.9 ± 1.0	22 ± 5
		40	90.1 ± 1.3	13 ± 4
0.918	0.0148	30	85.1 ± 0.9	27 ± 5
		35	86.6 ± 0.9	17 ± 4
		40	87.9 ± 1.3	8 ± 3

Table 8.7: Efficiency and rate for different values of E_T^{min} and $W_{\eta,2}^{max}$.

duce a reduction of electron signals, but definitely represents an improvement in the trigger for future studies involving boson production. Fig. 8.33 shows the distribution of the transverse momentum for fake electrons, reconstructed with the supercell scheme, before and after the cuts. Again, there is an improvement on rates showing the power of the new trigger selector.

Figure 8.32: Distribution of the transverse momentum for super-electrons compared with truth electrons. Before and after cuts for super-electrons.

Figure 8.33: Distribution of the transverse momentum for fake electrons, before and after the cuts.

8.4 Implementation of the electromagnetic trigger system at high luminosity

8.4.1 Current system

The implementation of the electromagnetic trigger in ATLAS is based on the analog sum of a group of towers of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$. This sum is made across the signal processing chain of the ATLAS electromagnetic calorimeter. Fig. 8.34 shows the block diagram of the electronics used to process signals on the LAr calorimeter. The first stage consists on the digitization of 128 channel groups of individual cells of the calorimeter by *Front End Boards* (FEB) [140] at a rate of 40 MHz. An automatic gain selection system composed by a 3 stage shapper circuit with gains of 0.8 (LO gain), 8.4 (MED gain) and 82 (HI gain) are used to achieve a 16 bit dynamic range from a 12 bit digitizer. In order to cope with the digitizer latency and speed, a system of analog memories composed of switching capacitors arrays (SCA) is used to store the signal for conversion. This electronic is also responsible for the topological addition of a set of cells via the analog sum of their signals made by layer sum circuits in the FEB which are sent to the *Tower Builder Board* (TBB) that builds the final form of the trigger towers of the calorimetry system. The trigger signal in this way is not digitalized as individual towers but by dedicated trigger processors (pre-processor modules or PPM), indicating the presence of potential candidates of physical objects of interest (electrons, jets, MET, etc.).

With respect to the process time, the target latency for the L1 trigger is set at $2.0\ \mu\text{s}$ (with a contingency of $0.5\ \mu\text{s}$), in order to reach a L1 trigger decision before the pipelines overflow. Considering that around of $1.5\ \mu\text{s}$ is spent transmitting signals between the detector electronics and the dedicated trigger electronics in the neighboring cavern, approximately only $0.5\ \mu\text{s}$ is available for the hardware to process the input signals and reach a trigger decision in order to be sent to the high-level triggers [73].

8.4.2 System at high luminosity

In Run 3 the read-out system and digitalization of cells will not be modified. It is in the formation and processing of the L1 trigger signal that will have the major changes, and the challenge is to add new informations keeping the

Figure 8.34: Current scheme of architecture employed on reading and processing of individual signals formation at L1 trigger of the electromagnetic calorimeter in ATLAS [8].

functionality of the current system. Several parts will change, and a new higher segmentation will be used separating the current trigger tower in 10 smaller supercells. The signal of each supercell will be digitized by the front end electronics (in contrast with the present system where the trigger tower signal is analog and digitized only in the back end electronics). New layer sum circuits and a new backplane will form and route the signals to a new hardware responsible for the supercell digitization.

The proposal is that the sum of the analog signals, sent by the FEB, will be digitalized and sent by optical fibers to the trigger processors, maintaining the current analog sum available. This processing will be done by a new device, called *Liquid Argon Trigger Digitizer Board* (LTDB).

The new system of digitization of calorimeter towers is shown in Fig. 8.35. Each digitizer board will have 320 channels distributed between 40 ADC's on 8 channels. ADC's must operate over a wide dynamic range for the trigger (at least 12 bits) at a conversion rate of at least 40 MHz, but is desirable to be able

Figure 8.35: *New scheme of architecture involving the digital trigger system highly segmented to the front-end electronic of electromagnetic calorimeter [8]. The red lines indicate the parts that will be added/modified.*

to operate at 80 MHz in order to facilitate the identification of pile-up signals. The high density of signals requires highly integrated devices (hence the use of ADC's 8-channel) and low consumption. The groups of 80 ADC channels are aggregated and serialized by programmable logic arrays (FPGA's) that communicate with the electronic processing of the trigger through 40 channels of 10 Gb/s, so the aggregate rate throughput of each LTDB can reach more than 400 Gb/s. In total there will be 64 of these boards only in the central region of the calorimeter, which means an aggregated rate of more than 25 Tb/s only for the electromagnetic trigger.

8.4.3 Prototype of the new digital system of highly segmented trigger (LTDB)

The inclusion of the processing and trigger system involves important changes in the FrontEnd system of the electromagnetic calorimeter. To preserve the conventional trigger and digitalization system of individual cells, a new back-

plane, responsible for the distribution of signals and the connection to the calorimeter, is installed together with the front end crate electronics, near the liquid argon cryostat.

Figure 8.36: *Picture of the electronic board LTDB prototype (left). Position of crate in the detector with the LTDB, just above the calorimeter cryostat (right).*

A modified rack containing 2 LTDB's was installed in the ATLAS calorimeter, so that during Run 2 in ATLAS the performance of this new system can be verified with experimental data. Fig. 8.36 shows the position of the central barrel calorimeter where the prototype of the new electronics will be installed. This position instrumented a region of barrel $\Delta\eta \times \Delta\phi = 1.4 \times 0.2$, corresponding to 30 analog trigger towers, 3584 cells and 640 supercells (Fig. 8.37).

Figure 8.37: *Covered region of the new digital system of highly segmented trigger (LTDB).*

The studies presented in this section will be based on discriminant comparisons reconstructed by the super-electron finding. Studies on efficiencies and rates are not possible under the current conditions. As a separate read-out chain and a custom topological L1 trigger is used in the data taking. The strategy will be to compare the super-cell information collected by the LAr calorimeter (*CaloCell* information) and by the LTDB.

Fig. 8.38 presents the kinematic supercell distributions (E_T , η , ϕ) reconstructed by the CaloCells and LTDB. In order to save memory in the information collected by the LTDB, only supercells above a certain threshold ($E_T > \sigma_{noise}$) are stored. Another important consideration is that for the runs taken during 2016 and used in this analysis, only a half of the region described above has passed the LTDB data quality selection. While the situation is now fixed for the current data taking, it is still too preliminary for the inclusion here.

As anticipated, the information stored by the LTDB will depend on the L1 triggers. The triggers to be used will be:

- **L1_EM15VH:** This trigger accepts events with at least an electromagnetic cluster (0.1×0.1) with $E_T > 15$ GeV and hadronic isolation.
- **L1_LAR-EM:** This trigger accepts events with at least an electromagnetic cluster (0.1×0.1) with $E_T > 20$ GeV.
- **L1_LAR-J:** This trigger accepts events with at least a jet element cluster (0.2×0.2) with $E_T > 50$ GeV.

Tab. 8.8 shows the number of events to be analyzed after trigger selections on CaloCells and LTDB for a single run taken during 2016 in collisions at 13 TeV. It is important to know that currently the LTDB works under a L1 pre-scaling trigger (only a part of the events collected by the CaloCells is collected by the LTDB). For that reason the number of events on the LTDB is low and the main results presented here will be based on CaloCells.

As in the main read-out, the supercell energy is reconstructed by the optimal filter method [8]. Fig. 8.39 presents the correlation plots for the supercell E_T reconstructed by the CaloCells and LTDB for each layer of the electromagnetic calorimeter, and, as can be seen, the E_T 's are strongly correlated, testing

Figure 8.38: Kinematic supercell distributions (E_T , η , ϕ) reconstructed with the CaloCell (left) and LTDB (right) information.

L1 trigger	Events	
	CaloCell	LTDB
L1_EM15VH	588946	0
L1_LAR-EM	93590	86
L1_LAR-J	158699	319

Table 8.8: Number of events after trigger selections on CaloCells and LTDB.

the correct operation of the demonstrator module.

Figure 8.39: Correlations for the supercell E_T reconstructed with the CaloCell and LTDB information for each layer of the electromagnetic calorimeter. The red line represents a linear fitting.

The previous section presented the power of the discriminants in selected samples (samples containing electrons and jets) but it is important to test them on experimental samples knowing the expected shapes for electrons and jets. Fig. 8.40 presents the $R_{\eta,2}$, $R_{\eta,1+2}$ and $W_{\eta,2}$ distributions for super-electrons reconstructed by CaloCells and the LTDB. Considering the Calo-cells, the L1_EM15VH trigger presents the shape expected for electrons since is added the hadronic isolation criteria, this trigger is commonly used on experimental analysis. The L1_LAR-EM trigger (trigger for EM signals) does not have any hadronic isolation criteria, for that reason the shape looks like a shape expected for jets, but even with that presents a small difference compared with the L1_LAR-J trigger (trigger for jets). The figure also shows the shapes obtained with LTDB and it can be seen how those shapes look similar than the ones obtained with CaloCells, although the number of events is low to observe an important difference between discriminants.

Figure 8.40: $R_{\eta,2}$, $R_{\eta,1+2}$ and $W_{\eta,2}$ distributions for super-electrons reconstructed with the CaloCell (left) and LTDB (right) information, considering different trigger selections at L1.

Conclusions

The studies presented here for measurement of Z boson associated with b -jets at 13 TeV allow us to understand how the flavor scheme affects the predictions and how this process can be better modeled. With respect to the cross section, AlpGen tends to underestimate the unfolded data in events with at least $1b$ -jet, but agrees in events with at least $2b$ -jets. Sherpa represents the better prediction for both cases, showing the importance of the five flavor scheme in the theoretical predictions. Kinematic distributions are consistent with the predictions (Sherpa and AlpGen), although again, AlpGen tends to underestimate the data in events with at least $1b$ -jet. These results are important inputs to be used in the development of PDF's.

Tab. 7.17 and Fig. 7.43 presented the comparisons with analysis at 7 TeV (combined analysis), showing a considerable improvement on the predictions, even with the current case, where we analyzed only the electron channel. Fig. 8.41 presents the prediction/data cross-section comparisons, detailing the observed in the b -jet inclusive multiplicity plots (Fig. 7.37), where is important to notice the reduction on the experimental uncertainties. Future studies on the combined channels (including muons) will represent an important contribution to the ATLAS experiment and must be included in other analysis, such as top quark production studies, Higgs precision measurements and the search for supersymmetric particles.

As a contribution to the ATLAS upgrade program, it was implemented a method of electron reconstruction with the use of supercells (based on the fine granularity of the cells of the ATLAS Liquid Argon detector) that will be useful in the first level trigger (L1). We showed that the modified use of the standard discriminators can improve the selection rate and efficiency even in L1. In fact, it was shown that with the use of at least two discriminants would be enough. This is important considering that we need to simplify the

(a)

(b)

Figure 8.41: Prediction/Data cross-section comparisons for $Z + \geq 1\text{b-jet}$ (a) and $Z + \geq 2\text{b-jets}$ (b) measured in ATLAS for collisions at 7 and 13 TeV. The inner blue shaded band showing the corresponding statistical uncertainty and the outer green shaded band showing the sum in quadrature of statistical and systematic uncertainties.

implementation in L1 (quick response). As a first recommendation, it was proposed the use of $R_{\eta,1+2} > 0.898$ and $W_{\eta,2} < 0.0148$ in order to maintain a threshold of $E_T = 35 \text{ GeV}$, with an efficiency of $89 \pm 1\%$ and a rate of $22 \pm 5 \text{ kHz}$.

Finally, as Fig. 8.42 shows, the supercell scheme can be already tested in experimental data, showing the impact of the proposed discriminants ($R_{\eta,1+2}$ and $W_{\eta,2}$). The supercell hardware prototype implementation is working in the current Run 2, and a considerable effort on studies of signal processing and calibration is currently underway, targeting the full implementation in the second long shutdown (2019-2020).

Figure 8.42: $R_{\eta,1+2}$ and $W_{\eta,2}$ distributions for super-electrons reconstructed with the CaloCell information on simulations (left) and data (right), considering electron and jet samples on simulations and different L1 trigger selections on data.

Appendix A

Kinematics on Boson Measurements

This appendix will present some quantities and kinematic terms associated to boson measurements.

Figure A.1: *W/Z boson production into two leptons.*

Considering the decay of a V boson ($V = Z$ or W) into a pair of leptons (Fig. A.1) with four-momenta P_1 and P_2 , the four-moment of V can be expressed as:

$$P_V = P_1 + P_2 = (E_1 + E_2, \mathbf{p}_1 + \mathbf{p}_2) \quad (\text{A.1})$$

The invariant mass can be obtained from $m_V^2 = P_V^2 = E_V^2 - \mathbf{p}_V^2$, then:

$$\begin{aligned} m_V^2 &= (E_1 + E_2)^2 - (\mathbf{p}_1 + \mathbf{p}_2)^2 \\ m_V^2 &= m_1^2 + m_2^2 + 2(E_1 E_2 - \mathbf{p}_1 \cdot \mathbf{p}_2) \end{aligned} \quad (\text{A.2})$$

if 1 and 2 are light particles, like electrons, the Eq. A.2 reduces to:

$$m_V^2 = 2(E_1 E_2 - \mathbf{p}_1 \cdot \mathbf{p}_2) \quad (\text{A.3})$$

considering 1.18 and 1.19 into A.3, the invariant mass for the V boson can be

expressed as:

$$m_V = \sqrt{2p_{T,1}p_{T,2}(\cosh(\eta_1 - \eta_2) - \cos(\phi_1 - \phi_2))} \quad (\text{A.4})$$

In the case of one particle cannot be detected, like neutrinos, it is important to define the transverse mass, $m_{T,V}$, which is a quantity invariant under Lorentz boost along the z direction (Fig. 1.11.a):

$$m_{T,V}^2 = (E_{T,1}^2 + E_{T,2}^2)^2 - (\mathbf{p}_{T,1} + \mathbf{p}_{T,2})^2 \quad (\text{A.5})$$

where $E_{T,1}$ and $E_{T,2}$ are the transverse energy of particles 1 and 2, respectively, with $E_T^2 \equiv m^2 + \mathbf{p}_T^2$. If 1 and 2 are light particles then $E_T \approx p_T$ and Eq. A.5 reduces to:

$$m_{T,V}^2 = 2(p_{T,1}p_{T,2} - \mathbf{p}_{T,1} \cdot \mathbf{p}_{T,2}) \quad (\text{A.6})$$

it is possible to found that $\mathbf{p}_{T,1} \cdot \mathbf{p}_{T,2} = p_{T,1}p_{T,2}\cos(\phi_2 - \phi_1)$, then:

$$m_{T,V}^2 = 2p_{T,1}p_{T,2}(1 - \cos(\phi_2 - \phi_1)) \quad (\text{A.7})$$

As wa shown, m_V and $m_{T,V}$ can be calculated if is known the p_T of each produced lepton. In the case of neutrinos the p_T is estimated trough MET measuremnts. The p_T of charged leptons can be measured experimentally knowing the particle trajectory (track) in a magnetic field.

The Lorentz force ($\mathbf{F}$) for a charged particle (q) in a magnetic field ($\mathbf{B}$) is:

$$\mathbf{F} = q\mathbf{v} \times \mathbf{B} \quad (\text{A.8})$$

where, $\mathbf{v}$ is the velocity of the particle, which can be separated in: $\mathbf{v} = \mathbf{v}_T + \mathbf{v}_z$

Considering $\mathbf{B}$ as an uniform field ($\mathbf{B} = B_0\hat{\mathbf{z}}$), $\mathbf{F}$ in Eq. A.8 only has a transversal component, which can be reduced to:

$$\begin{aligned} \frac{mv_T^2}{\rho} &= qv_TB_0 \\ \Rightarrow p_T &= qB_0\rho \end{aligned} \quad (\text{A.9})$$

where ρ is the bending radius for a circular track.

Appendix B

The Simulation Process in ATLAS

This appendix details the simulation process in ATLAS, which is based on the Monte Carlo method.

B.1 General overview

Fig. B.1 shows schematically the process, which begins with the event generator, followed by the detector simulation process, digitalization and reconstruction.

Figure B.1: *Scheme of simulation and reconstruction process in ATLAS.*

Event generator

This is the initial step in the simulation, where the particles are generated. For example: $pp \rightarrow Z \rightarrow ee$, $pp \rightarrow H \rightarrow bb$, etc. The particles of interest will be the stable ones and with a long life time, since they will interact directly with the detector (Sec. 1.3).

The most common event generators used in ATLAS are PYTHIA, HERWIG, POWHEG, SHERPA, ALPGEN and MADGRAPH5. The outputs are stored in files named *pool files*, which will be taken to the next process, the detector simulation. Preliminary studies are possible, considering analysis at particle-level, commonly called the *truth* analysis. The event generator process will be more detailed in the next section.

Detector simulation

The detector simulation is based on GEANT4 [147], where the generated particles interact with the sensitive and not sensitive material of the detector. At this stage are incorporated the generated particles in the previous step (interacting with the detector). The interaction with the sensitive material produces signals denominated as *hits* and these hits are stored for the next process.

Digitization

In this step, the hits must be digitalized, considering that each sub-detector has different processing methods such as the type of signal collection, the pulse shaping, etc. It is on this step that is included the electronic and pile-up noise. The outputs are stored in the format of *Raw Data Object* (ROD), this format is the same used to store experimental data.

Reconstruction

Finally, it is presented the reconstruction process, using the ROD information to reconstruct tracks, vertices, cells and cluster information in each sub-detector, which are then used by specialized algorithms to reconstruct high-level objects, classifying the produced particles (electrons, photons, muons, jets and MET). The output format is the xAOD, an upgrade of the *Analysis Object Data* (AOD), which contains the information used on final analysis.

B.2 Monte Carlo event generator

A simulation of proton-proton collision can be divided on the following stages: the hard process at parton levels, the parton shower evolution, the hadronization process, the unstable particle decays and the inclusion of underlying events.

The hard process and matrix element

This represents the most interesting physics process on particle collisions at high energy, where is expected a large invariant momentum transferred between partons on collision (hard interaction). At this level, the particles are asymptotically free, and it is possible to describe the reactions by perturbation theory. The cross section for a hard process $ab \rightarrow X$ at hadron colliders can be computed by:

$$\sigma = \sigma(|M_{ab \rightarrow X}|^2) \quad (\text{B.1})$$

where $M_{ab \rightarrow X}$ is called the *Matrix Element* (ME). This matrix can be written as a sum of process over Feynman diagrams (as shown in Sec. 6.1):

$$M_{ab \rightarrow X} = \sum_i F_{ab \rightarrow X}^i \quad (\text{B.2})$$

Examples of these computations can be found in the next appendix. The computation of cross sections is accompanied by the PDF's (Sec. 1.1.2). For lowest-order matrix elements there are several automated tools (packages) to construct and evaluate these process, but for higher-order calculations, i.e. including loop effects, the packages are not yet fully automated, requiring to take into account the parton shower evolution.

The parton shower

The parton shower (PS) evolution starts from the hard process. The partons are colored particles and radiate gluons¹ (the effects associated to these gluons were discussed in Sec 4.1.2). Then, the parton shower evolution needs to take into account the effects of soft gluon and collinear emissions. The parton shower evolution works downwards to lower and lower momentum scales, as is

¹This happens on incoming and outgoing accelerated partons.

shown in Fig B.2, and ends when the perturbation theory breaks down (order of 1 GeV), associated with the confinement of partons in hadrons.

Figure B.2: Diagram showing the parton shower evolution from an initial quark line [146].

With respect to the confinement in QCD, a pair $q\bar{q}$ at very short distances becomes more and more similar to a QED pair, but at long distances the self-gluon interaction make the field lines attract each other, being more probably to obtain other pair $q\bar{q}$ out of the vacuum rather than increase the length (Fig. B.3).

Figure B.3: Scheme describing the confinement for a pair $q\bar{q}$. The dashed lines represent the field between quarks.

An important aspect to consider is the *Initial State Radiation* (ISR), that in analogy with QED, states that it is equally expect that partons radiate gluons on their way into a scattering or annihilation process (giving rise to an initial-state parton shower). This can be implemented considering $1 \rightarrow 2$

splittings to a shower of partons, one of which is ultimately involved in the hard process, with the rest being emitted as accompanying radiation.

The most important consideration here is how to connect the parton showers to the hard process, as the parton shower describes the evolution of partons on their way in and out of a scattering process, but not the starting conditions for the showers. Even more, the parton shower is an approximation to describe the collinear and soft limits, but the hard process and very well separated jets, are well described by matrix element. Then, many generators improve their predictions into regions of phase space where PS or ME are reliable, avoiding double counting of events.

Hadronization

In this stage, it will be described how partons are confined into hadrons. There are two main models to describe the hadronization: The String Model and the Cluster Model.

Figure B.4: Diagram to describe the string model, showing the pair $q\bar{q}$ moving apart from each other as seen in the meson rest frame (left) and in a boosted frame (right) [145, 146]. Blue and green lines represent the quark 'ways' and red lines the strings between them.

The String Model: This model can be explained by a pair of massless quarks ($q\bar{q}$) joined by strings (Fig. B.4). The pairs initially at the same point in the time-space diagram, will fly apart ($v \sim c$), and are kept apart until the potential stored in the string, which has a constant tension, uses up all their kinetic energy. At this point they turn around and the potential energy in the string accelerates the quarks towards each other until they meet back at the starting point, pass through each other, and the whole process starts over.

These are the so-called "yo-yo" modes [146].

Figure B.5: *The motion and breakup of a string system, with the two transverse degrees of freedom suppressed [145]. Blue and green lines represent the quark 'ways' and red lines the strings between them.*

It is possible to expand this idea to a model of hadronization, like is done on the Lund String Model, implemented in PYTHIA. Considering again a pair $q\bar{q}$, this system also has yo-yo modes, but in practice the space-time volume swept-out is so large that another effect is able to dominate the process. The field is strong enough that the quarks are accelerated away from each other before they have time to re-annihilate, breaking the string, like is shown in Fig. B.5.

The cluster model:

The cluster model of hadronization is based on the pre-confinement property of parton showers, where the mass spectrum of combinations of partons (clusters) is asymptotically independent of the energy or the production process, and is peaked at low mass $\sim Q_0$ (cuts off for soft and collinear divergences).

The pre-confinement mechanism can be seen most simply in the limit of a large number of colors N_c . At leading order in N_c , the gluons in the shower can be represented by pairs of color-anticolor lines that are connected at vertices as is shown in Fig. B.6. Then, each color line at the low-scale-end of the shower is connected to an anticolor partner line at the same scale. In this limit the color structure of the shower can be drawn on a plane, considering adjacent color-anticolor partners [145]. Packages that use the cluster model for hadronization are HERWIG and SHERPA.

Figure B.6: *Colour structure of a parton shower to leading order in N_c [145].*

Unstable particle decays

After the hadronization process, stable and unstable hadrons are produced, where the unstable ones must be decayed into particles that are stable and interact with the detector.

Decay information can be extracted from Particle Data Group (PDG) tables, but this information is often insufficient because there are some limitations in some process (ex. excited mesons and baryons containing heavy quarks). This depends on the hadrons included in the simulation and the selected parameters on the hadronization. Sophisticated simulation of hadronic decays are performed using specialized external packages such as EVTGEN [148] for hadron decays and TAUOLA [149] for tau decays.

Underlying events

The underlying events are associated to partons that not participate of the hard collision. At high energy, protons on collisions look like flattened discs in the laboratory frame (Lorentz contraction) as is shown in Fig. B.7.

There are two main models to describe the underlying Events: The *non-perturbative model*, which is motivated by the fact that the soft parton-parton cross section is so large that there are many interactions everywhere in these discs, assuming these interactions as elastic scatterings across the discs. And the *perturbative models*, where the perturbative parton-parton cross section is so large that additional local parton-parton interactions between other partons in the proton dominate the process, and there is no elastic scattering but there are multiple independent parton-parton interactions distributed across the disc, each producing their own 'hard' processes and parton showers.

Figure B.7: Sketch of a proton-proton collision showing the Lorentz contraction of the protons [146].

Appendix C

Matrix Element Calculation

In this appendix will be shown some examples of matrix element (ME) computation. There are computational packages to compute ME, ex: FORMCALC[150] for Feynman diagrams generated by FEYNARTS[151]. Here it will be presented the main tools to compute the Z boson production and its association with jets at leading order (LO).

C.1 Z production at LO

Cross section for on-shell Z production ($q + \bar{q} \rightarrow Z$) will be computed by ME calculation from the Feynman diagram shown in Fig. C.1.

Figure C.1: LO Feynman diagram for $q + \bar{q} \rightarrow Z$ production.

The differential cross section for one particle production from two ones [20] can be expressed as:

$$d\sigma = \frac{\pi}{2} \frac{\langle M^2 \rangle}{\sqrt{(P_1 \cdot P_2)^2 - (m_1 m_2)^2}} \frac{d^3 \vec{p}_3}{E_3} \delta(P_1 + P_2 - P_3) \quad (\text{C.1})$$

where M is the matrix element; P_i , E_i and m_i are the 4-momentum, energy and mass of the particle i .

Considering the center-of-mass reference frame: $P_1 = (E_1, \vec{p}_1)$, $P_2 = (E_2, -\vec{p}_1)$ and $P_3 = (E_3, \vec{0})$, then:

$$d\sigma = \frac{\pi}{4} \frac{\langle M^2 \rangle}{|\vec{p}_1|(E_1 + E_2)} \frac{d^3\vec{p}_3}{E_3} \delta(P_1 + P_2 - P_3) \quad (\text{C.2})$$

M is computed using Feynman rules [21] through the diagram in Fig. C.1, considering that, we have:

$$-iM = \bar{v}(2) \left(-ig_z \gamma^\mu \left(\frac{1}{2}(c_v - c_a \gamma^5) \right) \right) u(1) \varepsilon_\mu^*(3) \quad (\text{C.3})$$

where u and v are Dirac spinors which described q and $\bar{q}$, respectively; ε the polarization vector for Z ; γ are Dirac matrices; g_z is the strength factor coupling associated to Z ; c_v and c_a are the vector and axial-vector couplings and are determined by the Standard Model.

M can be simplify as:

$$M = \frac{g_z}{2} \bar{v}(2) \gamma^\mu (c_v - c_a \gamma^5) u(1) \varepsilon_\mu^*(3) \quad (\text{C.4})$$

For the cross section, the value of interest is $\langle M^2 \rangle$, computed as:

$$\langle M^2 \rangle = \frac{1}{(2s_1 + 1)(2s_2 + 1)} \frac{1}{3} \sum_{s_1, s_2, s_3} M M^* \quad (\text{C.5})$$

s_i is the spin label of the particle i (Fig. C.1). M^* can be calculated from Eq. C.4:

$$M^* = \left[\frac{g_z}{2} \bar{v}(2) \gamma^\mu (c_v - c_a \gamma^5) u(1) \varepsilon_\mu^*(3) \right]^* \quad (\text{C.6})$$

considering that: $(AB)^* = (AB)^\dagger = B^\dagger A^\dagger$, then:

$$M^* = \frac{g_z}{2} [u(1)^\dagger (c_v - c_a \gamma^5)^\dagger \gamma^{\mu\dagger} \bar{v}(2)^\dagger] \varepsilon_\mu(3) \quad (\text{C.7})$$

and applying the next γ properties:

$$\begin{aligned}
\bar{u} &= u^\dagger \gamma^0 \Rightarrow \bar{u}^\dagger = \gamma^0 u \\
\gamma^{\mu\dagger} \gamma^0 &= \gamma^0 \gamma^\mu \Rightarrow \gamma^0 \gamma^{\mu\dagger} \gamma^0 = \gamma^\mu \\
(\gamma^5)^\dagger &= \gamma^5 \\
\gamma^5 \gamma^\mu &= -\gamma^\mu \gamma^5 \Rightarrow \gamma^0 \gamma^5 \gamma^0 = -\gamma^5 \\
(\gamma^5)^2 &= \mathbb{1}
\end{aligned} \tag{C.8}$$

the Eq. C.7 can be expressed by:

$$\begin{aligned}
M^* &= \frac{g_z}{2} \bar{u}(1) \gamma^\mu (c_v - c_a \gamma^5) v(2) \varepsilon_\mu(3) \\
&= \frac{g_z}{2} \bar{u}(1) \gamma^\nu (c_v - c_a \gamma^5) v(2) \varepsilon_\nu(3)
\end{aligned} \tag{C.9}$$

to notice that the label μ changed to ν in order to avoid confusions with M .

Then, Eq. C.5 will be expressed as:

$$\begin{aligned}
\langle M^2 \rangle &= \frac{g_z^2}{48} \sum_{s_1, s_2, s_3} [\bar{v}(2) \gamma^\mu (c_v - c_a \gamma^5) u(1)] [\bar{u}(1) \gamma^\nu (c_v - c_a \gamma^5) v(2)] \varepsilon_\mu^*(3) \varepsilon_\nu(3) \\
&= \frac{g_z^2}{48} \sum_{s_3} \varepsilon_\mu^*(3) \varepsilon_\nu(3) \sum_{s_1, s_2} [\bar{v}(2) \gamma^\mu (c_v - c_a \gamma^5) u(1)] [\bar{u}(1) \gamma^\nu (c_v - c_a \gamma^5) v(2)]
\end{aligned} \tag{C.10}$$

to reduce this expression is necessary to consider the completeness relation for fermions:

$$\sum_{s_1} \bar{u}(1) u(1) = \not{p}_1 + m_q \tag{C.11}$$

where $\not{p}_1 = \gamma^a P_{1a}$. In order to simplify the calculation, the quark mass will be neglecting, as for example if are considered *light*-quarks or quarks with high transverse moment, then Eq. C.11 is reduced to:

$$\sum_{s_1} \bar{u}(1) u(1) = \gamma^a P_{1a} \tag{C.12}$$

With Eq. C.12, the next property can be shown:

$$\begin{aligned} \sum_{s_1, s_2} \bar{v}_{1i} S_{ij} u(1)_{j1} \bar{u}(1)_{1k} T_{kl} v_{l1} &= Tr[(\gamma^a P_{1a}) T(\gamma^b P_{2b}) S] \\ &= P_{1a} P_{2b} Tr[\gamma^a T \gamma^b S] \end{aligned} \quad (C.13)$$

here, $i, j, k, l = 1, 2, 3, 4$ are labels in order to show the explicit multiplicity between vectors (u and v) and matrices (S and T). Then:

$$\langle M^2 \rangle = \frac{g_z^2}{48} \left(\sum_{s_3} \varepsilon_\mu^*(3) \varepsilon_\nu(3) \right) P_{1a} P_{2b} Tr[\gamma^a \gamma^\nu (c_v - c_a \gamma^5) \gamma^b \gamma^\mu (c_v - c_a \gamma^5)] \quad (C.14)$$

The completeness relation for bosons will be useful too:

$$\sum_{s_3} \varepsilon_\mu^*(3) \varepsilon_\nu(3) = -g_{\mu\nu} + \frac{P_{3\mu} P_{3\nu}}{m_z^2} \quad (C.15)$$

Then, with Eq. C.15 and some gamma properties showed before, Eq. C.14 can be expressed as:

$$\langle M^2 \rangle = \frac{g_z^2}{48} \left(-g_{\mu\nu} + \frac{P_{3\mu} P_{3\nu}}{m_z^2} \right) P_{1a} P_{2b} Tr[\gamma^a \gamma^\nu \gamma^b \gamma^\mu (c_v^2 + c_a^2 - 2c_v c_a \gamma^5)] \quad (C.16)$$

Considering the next γ properties for traces:

$$\begin{aligned} Tr[\gamma^a \gamma^\nu \gamma^b \gamma^\mu] &= 4(g^{a\nu} g^{b\mu} - g^{ab} g^{\mu\nu} + g^{a\mu} g^{b\nu}) \\ Tr[\gamma^a \gamma^\nu \gamma^b \gamma^\mu \gamma^5] &= 4i\epsilon^{a\nu b\mu} \end{aligned} \quad (C.17)$$

$$\begin{aligned} \langle M^2 \rangle &= \frac{g_z^2}{6} \left(-g_{\mu\nu} + \frac{P_{3\mu} P_{3\nu}}{m_z^2} \right) \times \\ &P_{1a} P_{2b} ((c_v^2 + c_a^2)(g^{a\nu} g^{b\mu} - g^{ab} g^{\mu\nu} + g^{a\mu} g^{b\nu}) - 2ic_v c_a \epsilon^{a\nu b\mu}) \end{aligned} \quad (C.18)$$

The product of a symmetric tensor (as $-g_{\mu\nu} + \frac{P_{3\mu} P_{3\nu}}{m_z^2}$) multiply by an anti-symmetric tensor (as $\epsilon^{a\nu b\mu}$) is zero, then the imaginary part will be suppressed.

And reducing the tensors through the next properties:

$$g_{\mu\nu}P^\mu = P_\nu \quad (\text{C.19})$$

$$g^{\mu\nu}g_{\mu\nu} = 4 \quad (\text{C.20})$$

Eq. C.18 can be reduced to:

$$\langle M^2 \rangle = \frac{g_z^2}{6m_z^2}(c_v^2 + c_a^2) (2(P_1 \cdot P_3)(P_2 \cdot P_3) + (P_1 \cdot P_2)m_z^2) \quad (\text{C.21})$$

and considering in the center-of-mass reference frame ($\vec{p}_1 + \vec{p}_2 = \vec{p}_3$), Eq. C.21 can be reduced to:

$$\langle M^2 \rangle = \frac{g_z^2}{24m_z^2}(c_v^2 + c_a^2) (3E_1E_2 + |\vec{p}_1|^2) \quad (\text{C.22})$$

Knowing $\langle M^2 \rangle$ and the differential cross section definition for one particle production from two ones (Eq. C.2), then:

$$\sigma_{q+\bar{q} \rightarrow Z} = \frac{\pi}{4} \frac{\langle M^2 \rangle}{|\vec{p}_1|(E_1 + E_2)} \int \frac{d^3\vec{p}_3}{\sqrt{|\vec{p}_3|^2 + m_z^2}} \delta(\vec{p}_1 + \vec{p}_2 - \vec{p}_3) \delta(E_1 + E_2 - E_3) \quad (\text{C.23})$$

With $\vec{p}_3 = \vec{p}_1 + \vec{p}_2 = \vec{0}$ and δ function properties, Eq. C.23 can be reduced to:

$$\sigma_{q+\bar{q} \rightarrow Z} = \frac{\pi}{4} \frac{\langle M^2 \rangle}{|\vec{p}_1|(E_1 + E_2)} \frac{1}{m_z} \delta(E_1 + E_2 - E_3) \quad (\text{C.24})$$

Denoting $s \equiv (P_1 + P_2)^2 = (E_1 + E_2)^2$ and $E_3 = m_Z$ (rest frame), it can be shown that:

$$\delta(E_1 + E_2 - E_3) = \delta(\sqrt{s} - m_z) = 2m_z \delta(s - m_z^2) \quad (\text{C.25})$$

Then, Eq. C.24 will be expressed as:

$$\sigma_{q+\bar{q} \rightarrow Z} = \frac{\pi}{24} g_z^2 (c_v^2 + c_a^2) \frac{(3E_1E_2 + |\vec{p}_1|^2)}{|\vec{p}_1|(E_1 + E_2)} \delta(s - m_z^2) \quad (\text{C.26})$$

For incoming particles at high energy: $E_1 = |\vec{p}_1|$ and $E_2 = |\vec{p}_1|$, then:

$$\sigma_{q+\bar{q} \rightarrow Z} = \frac{\pi}{12} g_z^2 (c_v^2 + c_a^2) \delta(s - m_z^2) \quad (\text{C.27})$$

This result can be expanded to $pp \rightarrow Z$ production, and must be implemented with the parton density functions ($f_i(x, Q^2)$):

$$\sigma_{p+p \rightarrow Z} = \frac{\pi}{4} \sum_{q, \bar{q}} \int dx_q dx_{\bar{q}} \frac{f_q(x_q, Q^2) f_{\bar{q}}(x_{\bar{q}}, Q^2)}{x_q x_{\bar{q}} \sqrt{(P_1 \cdot P_2)^2 - m_p^2}} \langle M_{q\bar{q}}^2 \rangle \frac{d^3 \vec{p}_3}{E_3} \delta(x_q P_1 + x_{\bar{q}} P_2 - P_3) \quad (\text{C.28})$$

where P_1 and P_2 are 4-moments of the protons, and x_q and $x_{\bar{q}}$ are the fraction moment from each proton.

The matrix element computation seems to be simple for a boson production but is complicated with the jet inclusion, in the next section will be presented the most simple case, $Z + 1\text{-jet}$, and its complications.

C.2 $Z + 1\text{-jet}$ production at LO

Since in this thesis the relevance of the Z +jets processes was mentioned several times, in this section will be computed the cross section estimation for $Z + 1\text{-jet}$ at LO, which can be produced by the channels presented in Fig. C.2.

Figure C.2: LO Feynman diagram for $Z + 1\text{-jet}$ production.

To simplify the calculation it will be used the channel $q + g \rightarrow Z + q$, which can be represented by the Feynman diagrams shown in Fig. C.3, where there are represented two contributions.

Figure C.3: LO Feynman diagram for $q + g \rightarrow Z + q$ production.

From Feynman rules for chromodynamics[20], we have:

$$\begin{aligned}
 -iM_1 &= \bar{v}(4)c_4^+ \left(-i\frac{g_z}{2\sqrt{2}}\gamma^\nu(c_v - c_a\gamma^5) \right) \left(i\frac{(\not{q}_1 + m)}{q_1^2 - m^2} \right) \left(-i\frac{g_s}{2}\lambda^\alpha\gamma^\mu \right) u(1)c_1\varepsilon_\nu^*(3)\varepsilon_\mu(2)a_2^\alpha \\
 -iM_2 &= \bar{v}(4)c_4^+ \left(-i\frac{g_s}{2}\lambda^\alpha\gamma^\nu \right) \left(i\frac{(\not{q}_2 + m)}{q_2^2 - m^2} \right) \left(-i\frac{g_z}{2\sqrt{2}}\gamma^\mu(c_v - c_a\gamma^5) \right) u(1)c_1\varepsilon_\mu^*(3)\varepsilon_\nu(2)a_2^\alpha
 \end{aligned}
 \tag{C.29}$$

The new terms here are: c_i , which represent the color of the quark i ; g_s , the strength factor of the chromodynamic force ($g_s = \sqrt{a\pi\alpha_s}$); λ , the Gell-Mann matrices and a_2 the color of the gluon.

Then:

$$\begin{aligned}
 M_1 &= \frac{g_s g_z}{4\sqrt{2}} \left(\bar{v}(4)\gamma^\nu(c_v - c_a\gamma^5)\frac{(\not{q}_1 + m)}{q_1^2 - m^2}\gamma^\mu u(1) \right) (c_4^+ \lambda^\alpha c_1) a_2^\alpha \varepsilon_\nu^*(3)\varepsilon_\mu(2) \\
 M_2 &= \frac{g_s g_z}{4\sqrt{2}} \left(\bar{v}(4)\gamma^\mu\frac{(\not{q}_2 + m)}{q_2^2 - m^2}\gamma^\nu(c_v - c_a\gamma^5)u(1) \right) (c_4^+ \lambda^\alpha c_1) a_2^\alpha \varepsilon_\nu^*(3)\varepsilon_\mu(2)
 \end{aligned}
 \tag{C.30}$$

and

$$\begin{aligned}
 M_1^* &= \frac{g_s g_z}{4\sqrt{2}} \left(\bar{u}(1)\gamma^\theta\frac{(\not{q}_1 + m)}{q_1^2 - m^2}(c_v + c_a\gamma^5)\gamma^\rho v(4) \right) (c_1^+ \lambda^\beta c_4) a_2^{*\beta} \varepsilon_\rho(3)\varepsilon_\theta^*(2) \\
 M_2^* &= \frac{g_s g_z}{4\sqrt{2}} \left(\bar{u}(1)(c_v + c_a\gamma^5)\gamma^\rho\frac{(\not{q}_2 + m)}{q_2^2 - m^2}\gamma^\theta v(4) \right) (c_1^+ \lambda^\beta c_4) a_2^{*\beta} \varepsilon_\rho(3)\varepsilon_\theta^*(2)
 \end{aligned}
 \tag{C.31}$$

For this case $\langle M^2 \rangle$ is computed as:

$$\begin{aligned}
\langle M^2 \rangle &= \frac{1}{(2s_1 + 1)} \frac{1}{3} \frac{1}{(2s_2 + 1)} \frac{1}{3} \sum_{s_1, s_2, s_3, s_4} (M_1 M_1^* + M_1 M_2^* + M_2 M_1^* + M_2 M_2^*) \\
&= \frac{1}{36} [\langle M_1 M_1^* \rangle + \langle M_1 M_2^* \rangle + \langle M_2 M_1^* \rangle + \langle M_2 M_2^* \rangle] \\
&= \frac{1}{36} [\langle M_1 M_1^* \rangle + 2\text{Re}(\langle M_1 M_2^* \rangle) + \langle M_2 M_2^* \rangle]
\end{aligned} \tag{C.32}$$

With respect to seen in the previous section, it can be seen how the difficulty to compute $\langle M^2 \rangle$ increases with the number of components. In this case the next properties are useful:

$$\begin{aligned}
c_n^+ c_n &= 1 \Rightarrow c_n c_n^+ = \mathbb{1} \\
a_2^\alpha a_2^{*\alpha} &= 1 \\
\text{Tr}[\lambda^\alpha \lambda^\beta] &= 2\delta_{\alpha\beta} \\
\text{Tr}[\gamma^a \gamma^b \gamma^c \gamma^d \gamma^e \gamma^f] &= 16 [g^{ab}(g^{cd}g^{ef} - g^{ce}g^{df} + g^{cf}g^{de}) - g^{ac}(g^{bd}g^{ef} - g^{be}g^{df} + g^{bf}g^{de}) \\
&\quad + g^{ad}(g^{bc}g^{ef} - g^{be}g^{cf} + g^{bf}g^{cf}) - g^{ae}(g^{bc}g^{df} - g^{bd}g^{cf} + g^{bf}g^{cd}) \\
&\quad + g^{af}(g^{bc}g^{de} - g^{bd}g^{ce} + g^{be}g^{cd})]
\end{aligned} \tag{C.33}$$

After the expansion of Eq. C.32 with the terms in Eq. C.30 and the properties presented above, it was necessary the using of the MATHEMATICA [152] package to simplify the calculation. Then, after all, we have:

$$\langle M_{q+g \rightarrow Z+q}^2 \rangle = \frac{g_s^2 g_z^2 (c_v^2 + c_a^2)}{24} \left[\frac{s^2 + u^2 + 2tm_z^2}{-su} \right] \tag{C.34}$$

and applying the same for $q + \bar{q} \rightarrow Z + g$, we have:

$$\langle M_{q+\bar{q} \rightarrow Z+g}^2 \rangle = \frac{g_s^2 g_z^2 (c_v^2 + c_a^2)}{9} \left[\frac{u^2 + t^2 + 2sm_z^2}{ut} \right] \tag{C.35}$$

Expanded a similar idea to $pp \rightarrow Z$ production and knowing the differential cross section definition for two particles production from two ones [20], the cross section for the Z boson associated with a jet can be computed from:

$$\begin{aligned}
\sigma_{p+p \rightarrow Z+q} = & \sum_q \int dx_q dx_g \frac{f_q(x_q, Q^2) f_g(x_g, Q^2)}{32\pi^2 x_q x_g S} \langle M_{q+g \rightarrow Z+q}^2 \rangle \\
& \times \frac{d^3 \vec{p}_z}{E_z} \frac{d^3 \vec{p}_q}{E_q} \delta(x_q P_1 + x_g P_2 - P_z - P_q)
\end{aligned} \tag{C.36}$$

$$\begin{aligned}
\sigma_{p+p \rightarrow Z+g} = & \sum_q \int dx_q dx_{\bar{q}} \frac{f_q(x_q, Q^2) f_{\bar{q}}(x_{\bar{q}}, Q^2)}{32\pi^2 x_q x_{\bar{q}} S} \langle M_{q+\bar{q} \rightarrow Z+g}^2 \rangle \\
& \times \frac{d^3 \vec{p}_z}{E_z} \frac{d^3 \vec{p}_g}{E_g} \delta(x_q P_1 + x_{\bar{q}} P_2 - P_z - P_g)
\end{aligned} \tag{C.37}$$

where P_1 and P_2 are 4-moments of the incident protons, $S = (P_1 + P_2)^2$, x_q , $x_{\bar{q}}$ and x_g (x_g) are the fraction moment of quarks (gluons) in the protons.

Appendix D

Heavy Flavor Overlap Removal

In this appendix will be shown the validations of the HFOR procedure through the comparison with muon distributions.

Figure D.1: Z boson invariant mass comparisons on Alpgen simulations for $Z(\rightarrow \ell\ell)+jets$ at $\sqrt{s} = 13$ TeV ($\ell = e, \mu$). Comparisons before and after the HFOR.

Figure D.2: Z boson transverse moment (top) and rapidity (bottom) comparisons on Alpgen simulations for $Z(\rightarrow \ell\ell)+\text{jets}$ at $\sqrt{s} = 13$ TeV ($\ell = e, \mu$). Comparisons before and after the HFOR.

Figure D.3: lepton transverse moment (top) and pseudo-rapidity (bottom) comparisons on Alpgen simulations for $Z(\rightarrow \ell\ell) + \text{jets}$ at $\sqrt{s} = 13$ TeV ($\ell = e, \mu$). Comparisons before and after the HFOR.

Figure D.4: Exclusive (top) and inclusive (bottom) multiplicity comparisons for jets on Alpgen simulations for $Z(\rightarrow \ell\ell) + \text{jets}$ at $\sqrt{s} = 13$ TeV ($\ell = e, \mu$). Comparisons before and after the HFOR.

Figure D.5: Transverse moment (top) and rapidity (bottom) comparisons for leading jets ($Z \rightarrow \ell\ell$) on Alpgen simulations for $Z(\rightarrow \ell\ell) + \text{jets}$ at $\sqrt{s} = 13$ TeV ($\ell = e, \mu$). Comparisons before and after the HFOR.

Figure D.6: Transverse moment (top) and rapidity (bottom) comparisons for sub-leading jets ($Z \rightarrow \ell\ell$) on Alpgen simulations for $Z(\rightarrow \ell\ell) + \text{jets}$ at $\sqrt{s} = 13$ TeV ($\ell = e, \mu$). Comparisons before and after the HFOR.

Figure D.7: *Inclusive multiplicity comparisons for c-jets (top) and b-jets (bottom) on Alpgen simulations for $Z(\rightarrow \ell\ell)+\text{jets}$ at $\sqrt{s} = 13 \text{ TeV}$ ($\ell = e, \mu$). Comparisons before and after the HFOR.*

Figure D.8: Transverse moment (top) and rapidity (bottom) comparisons for leading c-jets ($Z \rightarrow \ell\ell$) on Alpgen simulations for $Z(\rightarrow \ell\ell) + \text{jets}$ at $\sqrt{s} = 13$ TeV ($\ell = e, \mu$). Comparisons before and after the HFOR.

Figure D.9: Transverse moment (top) and rapidity (bottom) comparisons for sub-leading c-jets ($Z \rightarrow \ell\ell$) on Alpgen simulations for $Z(\rightarrow \ell\ell) + \text{jets}$ at $\sqrt{s} = 13 \text{ TeV}$ ($\ell = e, \mu$). Comparisons before and after the HFOR.

Figure D.10: Transverse moment (top) and rapidity (bottom) comparisons for leading b-jets ($Z \rightarrow \ell\ell$) on Alpgen simulations for $Z \rightarrow \ell\ell + \text{jets}$ at $\sqrt{s} = 13$ TeV ($\ell = e, \mu$). Comparisons before and after the HFOR.

Figure D.11: Transverse moment (top) and rapidity (bottom) comparisons for sub-leading b-jets ($Z \rightarrow \ell\ell$) on Alpgen simulations for $Z(\rightarrow \ell\ell) + \text{jets}$ at $\sqrt{s} = 13$ TeV ($\ell = e, \mu$). Comparisons before and after the HFOR.

Appendix E

Sample Validations for $Z+b$ -jets at 13 TeV

This appendix will show the validations for generator on comparisons with muon distributions. Here it will only be used the truth information present on the simulated samples.

Channel	Generator	$\sigma(Z+ \geq 1b)$ [fb]	$\sigma(Z+ \geq 2b)$ [fb]
$Z \rightarrow ee$	Sherpa 2.1	11323 ± 49	1376 ± 17
	Sherpa 2.2	11488 ± 40	1404 ± 9
	Alpgen	7761 ± 36	1303 ± 12
$Z \rightarrow \mu\mu$	Sherpa 2.1	11368 ± 50	1343 ± 16
	Sherpa 2.2	11276 ± 39	1407 ± 9
	Alpgen	7813 ± 38	1298 ± 12
$Z \rightarrow \ell\ell$	Sherpa 2.1	11346 ± 35	1360 ± 12
	Sherpa 2.2	11382 ± 28	1406 ± 6
	Alpgen	7787 ± 26	1301 ± 8

Table E.1: *Theory predictions for the integrated inclusive b -jet cross-sections in proton-proton collisions at 13 TeV (normalized to NNLO), $Z \rightarrow \ell\ell$ represents the combined measurement. Only the statistical uncertainty is quoted.*

Figure E.1: Z boson invariant mass comparisons on MC simulations for $Z(\rightarrow \ell\ell) + \text{jets}$ at $\sqrt{s} = 13$ TeV ($\ell = e, \mu$).

Figure E.2: Z boson transverse momentum (top) and rapidity (bottom) comparisons on MC simulations for $Z(\rightarrow \ell\ell)+jets$ at $\sqrt{s} = 13$ TeV ($\ell = e, \mu$).

Figure E.3: Lepton transverse momentum (top) and pseudo-rapidity (bottom) comparisons on MC simulations for $Z(\rightarrow \ell\ell) + \text{jets}$ at $\sqrt{s} = 13$ TeV ($\ell = e, \mu$).

Figure E.4: Inclusive multiplicity comparisons for jets (top) and b-jets (bottom) on MC simulations for $Z(\rightarrow \ell\ell) + \text{jets}$ at $\sqrt{s} = 13 \text{ TeV}$ ($\ell = e, \mu$).

Figure E.5: Transverse momentum and rapidity comparisons for leading jets on MC simulations for $Z(\rightarrow \ell\bar{\ell}) + \text{jets}$ at $\sqrt{s} = 13$ TeV ($\ell = e, \mu$).

Figure E.6: Transverse momentum and rapidity comparisons for sub-leading jets on MC simulations for $Z(\rightarrow \ell\ell) + \text{jets}$ at $\sqrt{s} = 13 \text{ TeV}$ ($\ell = e, \mu$).

Figure E.7: Transverse momentum and rapidity comparisons for leading b-jets on MC simulations for $Z(\rightarrow \ell\ell) + \text{jets}$ at $\sqrt{s} = 13$ TeV ($\ell = e, \mu$).

Figure E.8: Transverse momentum and rapidity comparisons for sub-leading b-jets on MC simulations for $Z(\rightarrow \ell\ell) + \text{jets}$ at $\sqrt{s} = 13 \text{ TeV}$ ($\ell = e, \mu$).

Appendix F

Unfolding Plots for $Z + b$ -jets Production at 13 TeV

In this appendix will be presented the distributions used for the unfolding procedure and the total systematics breakdown for the preliminary results on VHF analysis in this thesis:

- Response matrices.
- Reconstruction efficiency.
- Fake correction.
- Closure tests.

F.1 Response matrices

Figure F.1: Response matrices corresponding to the exclusive jet (top) and b-jet (bottom) multiplicity obtained with Sherpa (left) and Alpgen (right).

Figure F.2: Response matrices corresponding to the leading (top) and sub-leading (bottom) jet p_T obtained with Sherpa (left) and Alpgen (right).

Figure F.3: Response matrices corresponding to the leading (top) and sub-leading (bottom) jet rapidity obtained with Sherpa (left) and Alpgen (right).

Figure F.4: Response matrices corresponding to $\Delta\phi_{jj}$ (top) and Δy_{jj} (bottom) of dijets obtained with Sherpa (left) and Alpgen (right).

Figure F.5: Response matrices corresponding to ΔR_{jj} (top) and M_{jj} (bottom) of dijets obtained with Sherpa (left) and Alpgen (right).

Figure F.6: Response matrices corresponding to the leading (top) and sub-leading (bottom) b-jet p_T obtained with Sherpa (left) and Alpgen (right).

Figure F.7: Response matrices corresponding to the leading (top) and sub-leading (bottom) b-jet rapidity obtained with Sherpa (left) and Alpgen (right).

Figure F.8: Response matrices corresponding to $\Delta\phi_{bb}$ and Δy_{bb} of di-jets obtained with Sherpa (left) and Alpgen (right).

Figure F.9: Response matrices corresponding to ΔR_{bb} and M_{bb} of di-jets obtained with Sherpa (left) and Alpgen (right).

Figure F.10: Response matrices corresponding to p_T of Z bosons in events with at least 1b-jet (top) and at least 2b-jets (bottom) obtained with Sherpa (left) and Alpgen (right).

Figure F.11: Response matrices corresponding to $|y|$ of Z bosons in events with at least 1b-jet (top) and at least 2b-jets (bottom) obtained with Sherpa (left) and Alpgen (right).

F.2 Reconstruction efficiency

Figure F.12: Efficiency reconstruction as a function of the exclusive jet and b-jet multiplicity comparing Sherpa and Alpgen.

Figure F.13: Efficiency reconstruction as a function of the leading (top) and sub-leading (bottom) jet p_T and rapidity comparing Sherpa and Alpgen.

Figure F.14: Efficiency reconstruction as a function of $\Delta\phi_{jj}$, Δy_{jj} , ΔR_{jj} and M_{jj} of dijets comparing Sherpa and Alpgen.

Figure F.15: Efficiency reconstruction as a function of the leading (top) and sub-leading (bottom) b-jet p_T and rapidity comparing Sherpa and Alpgen.

Figure F.16: Efficiency reconstruction as a function of $\Delta\phi_{bb}$, Δy_{bb} , ΔR_{bb} and M_{bb} of dijets comparing Sherpa and Alpgen.

Figure F.17: Efficiency reconstruction as a function of p_T and $|y|$ of Z bosons in events with at least 1b-jet (top) and at least 2b-jets (bottom) comparing Sherpa and Alpgen.

F.3 Fake correction

Figure F.18: Fake correction as a function of the exclusive jet and b-jet multiplicity comparing Sherpa and Alpgen.

Figure F.19: Fake correction as a function of the leading (top) and sub-leading (bottom) jet p_T and rapidity comparing Sherpa and Alpgen.

Figure F.20: Fake correction as a function of $\Delta\phi_{jj}$, Δy_{jj} , ΔR_{jj} and M_{jj} of dijets comparing Sherpa and Alpgen.

Figure F.21: Fake correction as a function of the leading (top) and sub-leading (bottom) b-jet p_T and rapidity comparing Sherpa and Alpgen.

Figure F.22: Fake correction as a function of $\Delta\phi_{bb}$, Δy_{bb} , ΔR_{bb} and M_{bb} of dijets comparing Sherpa and Alpgen.

Figure F.23: Fake correction as a function of p_T and $|y|$ of Z bosons in events with at least 1b-jet (top) and at least 2b-jets (bottom) comparing Sherpa and Alpgen.

F.4 Closure tests

Figure F.24: Closure test for the leading (top) and sub-leading (bottom) jet p_T and rapidity obtained with Sherpa.

Figure F.25: Closure test for $\Delta\phi_{jj}$, Δy_{jj} , ΔR_{jj} and M_{jj} of dijets obtained with Sherpa.

Figure F.26: Closure test for the leading (top) and sub-leading (bottom) b-jet p_T and rapidity obtained with Sherpa.

Figure F.27: Closure test for $\Delta\phi_{bb}$, Δy_{bb} , ΔR_{bb} and M_{bb} of dijets obtained with Sherpa.

Figure F.28: Closure test for p_T and $|y|$ of Z bosons in events with at least 1b-jet (top) and at least 2b-jets (bottom) obtained with Sherpa.

Appendix G

$Z + b$ -jets Results and Comparisons with Z +jets Analysis at 13 TeV

In this appendix will be shown comparisons for full jet studies (Z +jets) on the VHF analysis with the official Z +jets analysis at 13 TeV [131].

Figure G.1: Comparisons between measured cross section for the official Z +jets analysis at 13 TeV (left) and VHF results (right) as a function of the inclusive jet multiplicity. The data are compared to the predictions by Sherpa and Alpgen on VHF, including Blackhat and Madgraph on the official analysis [131]. The error bars indicate the statistical uncertainty, and the hatched bands the statistical and systematic uncertainties on data added in quadrature.

Figure G.2: Comparisons between measured cross section for the official Z +jets analysis at 13 TeV (left) and VHF results (right) as a function of transverse momentum (top) and rapidity (bottom) of the leading jet. The data are compared to the predictions by Sherpa and Alpgen on VHF, including Blackhat and Madgraph on the official analysis [131]. The error bars indicate the statistical uncertainty, and the hatched bands the statistical and systematic uncertainties on data added in quadrature.

Figure G.3: Comparisons between measured cross section for the official Z+jets analysis at 13 TeV (left) and VHF results (right) as a function of $\Delta\phi_{jj}$ and $M(j,j)$ for di-jets. The data are compared to the predictions by Sherpa and Alpgen on VHF, including Blackhat and Madgraph on the official analysis [131]. The error bars indicate the statistical uncertainty, and the hatched bands the statistical and systematic uncertainties on data added in quadrature.

Appendix H

$Z + b$ -jets Results and Comparisons with Analysis at 7 TeV

In this appendix will be shown comparisons for the $Z + b$ jets analysis at 7 TeV [115] and the results found at 13 TeV.

Figure H.1: Comparisons between measured cross section for results at 7 (left) [115] and 13 TeV (right) as a function of the transverse momentum (top) and rapidity (bottom) for Z boson in events with at least 1b-jet. The data are compared to the predictions by Sherpa and Alpgen. The error bars indicate the statistical uncertainty, and the hatched bands the statistical and systematic uncertainties on data added in quadrature.

Figure H.2: Comparisons between measured cross section for results at 7 (left) [115] and 13 TeV (right) as a function of the transverse momentum (top) and rapidity (bottom) for Z boson in events with at least 2b-jets. The data are compared to the predictions by Sherpa and Alpgen. The error bars indicate the statistical uncertainty, and the hatched bands the statistical and systematic uncertainties on data added in quadrature.

Figure H.3: Comparisons between measured cross section for results at 7 (left) [115] and 13 TeV (right) as a function of ΔR_{bb} and $M(b,b)$ for di b-jets. The data are compared to the predictions by Sherpa and Alpgen. The error bars indicate the statistical uncertainty, and the hatched bands the statistical and systematic uncertainties on data added in quadrature.

Appendix I

Efficiency and Rate Tables for Discriminants on SuperCell Studies

In this appendix is shown more tables comparing efficiencies and rates on the discriminator analysis.

R_{η}^{min}	Cuts E_T^{min} [GeV]	Efficiency [%]	Rate [kHz]
0.715	20	95.0 ± 0.8	635 ± 24
	25	95.3 ± 0.9	296 ± 17
	30	95.5 ± 0.9	151 ± 12
	35	95.6 ± 1.0	86 ± 9
	40	95.9 ± 1.4	50 ± 7
0.848	20	90.1 ± 0.8	319 ± 17
	25	91.4 ± 0.8	166 ± 12
	30	92.3 ± 0.9	91 ± 9
	35	93.0 ± 1.0	51 ± 7
	40	93.8 ± 1.3	32 ± 6
0.873	20	85.1 ± 0.8	246 ± 15
	25	87.0 ± 0.8	128 ± 11
	30	88.6 ± 0.9	73 ± 8
	35	90.0 ± 1.0	38 ± 6
	40	91.3 ± 1.3	26 ± 5

Table I.1: *Efficiency and rate for different values of E_T^{min} and R_{η}^{min} .*

$R_{\eta,1}^{min}$	Cuts		Efficiency [%]	Rate [kHz]
	E_T^{min}	[GeV]		
0.812	20		95.0 ± 0.8	654 ± 25
	25		95.2 ± 0.9	312 ± 17
	30		95.4 ± 0.9	164 ± 13
	35		95.4 ± 1.0	93 ± 9
	40		95.5 ± 1.4	54 ± 7
0.934	20		90.2 ± 0.8	305 ± 17
	25		91.3 ± 0.8	156 ± 12
	30		92.2 ± 0.9	82 ± 9
	35		92.8 ± 1.0	46 ± 7
	40		93.3 ± 1.3	29 ± 5
0.950	20		85.1 ± 0.8	236 ± 15
	25		86.7 ± 0.8	122 ± 11
	30		88.3 ± 0.9	63 ± 8
	35		89.5 ± 1.0	36 ± 6
	40		90.5 ± 1.3	23 ± 5

Table I.2: Efficiency and rate for different values of E_T^{min} and $R_{\eta,1}^{min}$.

$W_{\eta,1}^{max}$	Cuts		Efficiency [%]	Rate [kHz]
	E_T^{min}	[GeV]		
0.0129	20		95.0 ± 0.8	654 ± 25
	25		95.3 ± 0.9	310 ± 17
	30		95.6 ± 0.9	156 ± 12
	35		95.8 ± 1.0	90 ± 9
	40		96.0 ± 1.4	54 ± 7
0.0117	20		90.2 ± 0.8	534 ± 22
	25		90.6 ± 0.8	264 ± 16
	30		91.0 ± 0.9	134 ± 11
	35		91.3 ± 1.0	73 ± 8
	40		91.6 ± 1.3	46 ± 7
0.0106	20		85.1 ± 0.8	424 ± 20
	25		85.7 ± 0.8	202 ± 14
	30		86.2 ± 0.9	101 ± 10
	35		86.6 ± 0.9	59 ± 8
	40		87.0 ± 1.3	33 ± 6

Table I.3: Efficiency and rate for different values of E_T^{min} and $W_{\eta,1}^{max}$.

f_3^{max}	Cuts		Efficiency [%]	Rate [kHz]
	E_T^{min}	[GeV]		
0.026	20		95.0 ± 0.8	515 ± 22
	25		95.8 ± 0.9	257 ± 15
	30		96.4 ± 0.9	134 ± 11
	35		96.7 ± 1.0	78 ± 9
	40		97.0 ± 1.4	47 ± 7
0.018	20		89.8 ± 0.8	385 ± 19
	25		91.4 ± 0.8	196 ± 13
	30		92.6 ± 0.9	101 ± 10
	35		93.5 ± 1.0	60 ± 8
	40		94.6 ± 1.4	36 ± 6
0.015	20		84.8 ± 0.8	319 ± 17
	25		86.7 ± 0.8	163 ± 12
	30		88.4 ± 0.9	86 ± 9
	35		89.8 ± 1.0	53 ± 7
	40		91.2 ± 1.3	29 ± 5

Table I.4: Efficiency and rate for different values of E_T^{min} and f_3^{max} .

R_{had}^{max}	Cuts		Efficiency [%]	Rate [kHz]
	E_T^{min}	[GeV]		
0.129	20		95.0 ± 0.8	709 ± 26
	25		95.6 ± 0.9	334 ± 17
	30		96.2 ± 0.9	170 ± 13
	35		96.5 ± 1.0	92 ± 9
	40		96.9 ± 1.4	54 ± 7
0.079	20		90.0 ± 0.8	684 ± 25
	25		90.9 ± 0.8	319 ± 17
	30		91.8 ± 0.9	161 ± 12
	35		92.6 ± 1.0	85 ± 9
	40		93.7 ± 1.3	50 ± 7
0.056	20		85.1 ± 0.8	665 ± 25
	25		86.0 ± 0.8	310 ± 17
	30		87.1 ± 0.9	159 ± 12
	35		88.1 ± 1.0	83 ± 9
	40		89.5 ± 1.3	50 ± 7

Table I.5: Efficiency and rate for different values of E_T^{min} and R^{max} .

R_{seed}^{min}	Cuts		Efficiency [%]	Rate [kHz]
	E_T^{min}	[GeV]		
0.201	20		95.0 ± 0.8	658 ± 25
	25		95.4 ± 0.9	309 ± 17
	30		95.7 ± 0.9	152 ± 12
	35		95.9 ± 1.0	87 ± 9
	40		96.2 ± 1.4	51 ± 7
0.243	20		90.1 ± 0.8	543 ± 22
	25		90.9 ± 0.8	261 ± 15
	30		91.5 ± 0.9	134 ± 11
	35		92.0 ± 1.0	78 ± 9
	40		92.5 ± 1.3	45 ± 7
0.271	20		85.2 ± 0.8	447 ± 20
	25		86.3 ± 0.8	219 ± 14
	30		87.2 ± 0.9	118 ± 11
	35		87.9 ± 1.0	67 ± 8
	40		88.6 ± 1.3	38 ± 6

Table I.6: Efficiency and rate for different values of E_T^{min} and R_{seed}^{min} .

$R_{2^{nd}hot}^{max}$	Cuts		Efficiency [%]	Rate [kHz]
	E_T^{min}	[GeV]		
0.877	20		95.0 ± 0.8	780 ± 27
	25		95.2 ± 0.9	366 ± 18
	30		95.4 ± 0.9	187 ± 13
	35		95.4 ± 1.0	104 ± 10
	40		95.5 ± 1.4	61 ± 8
0.646	20		90.0 ± 0.8	670 ± 25
	25		90.3 ± 0.8	314 ± 17
	30		90.5 ± 0.9	159 ± 12
	35		90.6 ± 1.0	87 ± 9
	40		90.8 ± 1.3	51 ± 7
0.478	20		85.0 ± 0.8	580 ± 23
	25		85.5 ± 0.8	271 ± 16
	30		85.8 ± 0.9	141 ± 12
	35		86.0 ± 0.9	79 ± 9
	40		86.4 ± 1.3	46 ± 7

Table I.7: Efficiency and rate for different values of E_T^{min} and $R_{2^{nd}hot}^{max}$.

Cuts			Efficiency [%]	Rate [kHz]
$R_{\eta,1+2}^{min}$	$W_{\eta,2}^{max}$	E_T^{min} [GeV]		
0.777	0.0159	20	93.3 ± 0.8	393 ± 18
		25	93.7 ± 0.9	175 ± 13
		30	94.1 ± 0.9	96 ± 9
		35	94.4 ± 1.0	58 ± 8
		40	94.8 ± 1.4	36 ± 6
0.777	0.0148	20	89.2 ± 0.8	295 ± 16
		25	90.0 ± 0.8	138 ± 12
		30	90.8 ± 0.9	72 ± 8
		35	91.4 ± 1.0	39 ± 6
		40	92.2 ± 1.3	22 ± 5
0.777	0.0145	20	85.9 ± 0.8	248 ± 15
		25	87.0 ± 0.8	124 ± 11
		30	88.0 ± 0.9	68 ± 8
		35	88.8 ± 1.0	36 ± 6
		40	89.8 ± 1.3	20 ± 4

Table I.8: Efficiency and rate for different values of E_T^{min} and $W_{\eta,2}^{max}$, with $R_{\eta,1+2}^{min}=0.777$.

Cuts			Efficiency [%]	Rate [kHz]
$R_{\eta,1+2}^{min}$	$W_{\eta,2}^{max}$	E_T^{min} [GeV]		
0.898	0.0159	20	88.8 ± 0.8	211 ± 14
		25	90.0 ± 0.8	101 ± 10
		30	91.0 ± 0.9	50 ± 7
		35	91.6 ± 1.0	30 ± 5
		40	92.2 ± 1.3	19 ± 4
0.898	0.0148	20	85.3 ± 0.8	182 ± 13
		25	86.8 ± 0.8	87 ± 9
		30	88.0 ± 0.9	40 ± 6
		35	89.0 ± 1.0	22 ± 5
		40	89.9 ± 1.3	12 ± 3
0.898	0.0145	20	82.4 ± 0.7	156 ± 12
		25	84.0 ± 0.8	78 ± 9
		30	85.4 ± 0.9	38 ± 6
		35	86.6 ± 0.9	22 ± 5
		40	87.6 ± 1.3	12 ± 3

Table I.9: Efficiency and rate for different values of E_T^{min} and $W_{\eta,2}^{max}$, with $R_{\eta,1+2}^{min}=0.898$.

Cuts			Efficiency [%]	Rate [kHz]
$R_{\eta,1+2}^{min}$	$W_{\eta,2}^{max}$	E_T^{min} [GeV]		
0.918	0.0159	20	84.1 ± 0.8	152 ± 12
		25	86.0 ± 0.8	75 ± 9
		30	87.7 ± 0.9	35 ± 6
		35	88.9 ± 1.0	22 ± 5
		40	90.1 ± 1.3	13 ± 4
0.918	0.0148	20	81.2 ± 0.7	132 ± 11
		25	83.3 ± 0.8	65 ± 8
		30	85.1 ± 0.9	27 ± 5
		35	86.6 ± 0.9	17 ± 4
		40	87.9 ± 1.3	8 ± 3
0.918	0.0145	20	78.6 ± 0.7	114 ± 10
		25	80.8 ± 0.8	59 ± 8
		30	82.7 ± 0.8	26 ± 5
		35	84.3 ± 0.9	17 ± 4
		40	85.8 ± 1.3	8 ± 3

Table I.10: Efficiency and rate for different values of E_T^{min} and $W_{\eta,2}^{max}$, with $R_{\eta,1+2}^{min}=0.918$.

Appendix J

Jet Calibration at High Luminosity

This appendix presents the performance of the jet calibration at high pile-up levels, important to test the limitations that can accuse the current calibration under these conditions.

Current studies at high luminosity are useful in order to provide information about the ATLAS performance under this conditions that can bring possibilities of including new concepts of detectors or upgrades on analysis techniques. In addition, groups of physic analysis are currently working with simulations at high luminosity and depend on the jet MC calibration for their studies. In this section are presented the first results for MC-JES calibration, testing the limitations of the reconstruction and calibration methods on studies realized in 2015.

J.1 Samples and selections

The samples are similar that used at 8 TeV, based on di-jet simulation, but considering proton-proton collisions at 14 TeV, generated with PYTHIA 8 and using the AU2CT10 PDF set. These samples are also sliced accord the p_T of the leading true-jets as is shown in Tab. J.1. For comparisons at different pile-up levels, the samples were derived under three pile-up conditions: $\langle\mu\rangle=80$, 140 and 200.

An important consideration is that only were available 3.2 M events to analyze (for the R-scan were analyzed 40 M events).

Jets are reconstructed from topological energy clusters in the calorimeter using the anti- k_t algorithm with radius parameter of 0.4 at EM scale reconstruction. For comparison with other radius, jets with $R_0 = 0.3$ were recon-

Sample	$p_{T,leading}^{truth}$ [GeV]	σ [nb]
jetjet_JZ0W	0-20	7.69E+07
jetjet_JZ1W	20-60	2.46E+04
jetjet_JZ2W	60-160	3.48E+02
jetjet_JZ3W	160-400	3.19E+00
jetjet_JZ4W	400-800	3.95E-02
jetjet_JZ5W	800-1300	1.67E-03
jetjet_JZ6W	1300-1800	1.15E-04
jetjet_JZ7W	1800-2500	1.05E-04

Table J.1: *Simulated di-jet samples and their cross sections.*

structured by re-clustering with the FASTJET package [91]. Kinematic selections are shown in Tab. 5.2

Requirement	Cut
Matching	$\Delta R < 0.3$
Truth Isolation	$\Delta R < 0.6$
p_T true-jet	$p_{T,true} > 10$ GeV
Reco Isolation	$\Delta R < 0.6$
p_T reco-jet	$p_{T,reco} > 7$ GeV
E_T threshold	$E_T^{min} > 20$ GeV

Table J.2: *Kinematic selections for MC-JES calibration at 14 TeV under high luminosity conditions.*

The pile-up suppression applied was based in the area subtraction and the residual correction (Eq. 4.3). Techniques based on suppression were suggested, but on preliminary tests it was found not much differences and was decided only use the previous one. Details about this techniques will be shown in the next section.

J.2 MC-JES at different pile-up levels

After apply the pile-up corrections, it is expected to get calibration curves independent of μ , but in Fig. J.1 can be seen differences between them at low energy, reflecting some limitations on the pile-up correction applied. Of course, considering that, it is expected a non-closure results on the jet response after the calibration, as can be seen in Fig. J.2, where is shown how the calibration at low energies is worst with the increase of μ .

Figure J.1: Energy calibration curves and their differences at different pile-up levels for $|\eta| < 0.2$.

Figure J.2: Jet response after the calibration at different pile-up levels for $|\eta| < 0.2$.

J.3 Improvements at high luminosity

Although the calibration is done on energy, the calibration can be extended to the transverse momentum (Eq. 5.8). Then, the transverse momentum response

can be defined as:

$$\mathcal{R}_{p_T} \equiv \frac{p_{T,\text{reco}}^{\text{calib}}}{p_{T,\text{true}}} = \frac{JES \times \sqrt{E_{\text{reco}}^2 - m_{\text{reco}}^2}}{p_{T,\text{true}} \times \cosh(\eta_{\text{reco}}^{\text{EtaJES}})} \quad (\text{J.1})$$

where JES is the value obtained from the inverse of the energy calibration curve (Eq. 5.3) and $\eta_{\text{reco}}^{\text{EtaJES}}$ is η_{reco} corrected (Eq. 5.7).

After several studies it was found an overestimation in the pile-up correction, which should be optimized. The high level of pile-up generates high discrepancies for $E_T < 60$ GeV. By other hand, the usual calibration curve will depend on a polynomial function (Sec. 5.2) but was proposed a linear extrapolation of the curve responses below 20 GeV for $R = 0.3$ and 30 GeV for $R = 0.4$.

Figure J.3: Energy and p_T responses after calibration for collisions at 14 TeV and $\langle\mu\rangle = 200$, as function of pseudorapidity. Top: $R=0.3$; Bottom: $R=0.4$

In Fig. J.3 are presented the energy and p_T responses on function of $|\eta|$, after the calibration. There are responses with an energy deviation (p_T) lower than 1% for reference energy (p_T) above 60 GeV, but still remains the non-closure at low energy (p_T) and over the transition regions between detectors.

Until the end of these studies (end of 2015), the results of calibrations and recommendations, where the use $p_T > 30$ GeV for $R = 0.3$ and $p_T > 40$ GeV for $R = 0.4$. The preliminary conclusions obtained on this study were:

- The possibility of changing the current calibration scheme using the calibration directly on the transverse momentum instead the energy.
- Study for improvements of the methods to correct the pile-up, which will be presented in App. [K](#).

Appendix K

Pile-up Suppression at High Luminosity

In Sec. 4.5 was presented the JVF selection, presenting an effective rejection of pile-up jets for analysis in Run 1. Currently, studies in ATLAS present improvements on the suppression, which started with the inclusion of new variables. One of them is a correction to the JVF (corrJVF), that because the denominator of JVF (Eq. 4.4) increases with the number of reconstructed primary vertices in the event, then, the mean JVF for signal jets is shifted to smaller values. Another variable is the R_{p_T} , defined entirely from hard-scatter observables.

For analysis in Run 2 a new variable was introduced, the jet vertex tagger (JVT), which uses as inputs the corrJVF and R_{p_T} variables. Improvements on these variables are the first suggestions for analysis at high luminosity.

K.1 corrJVF and R_{p_T}

The quantity corrJVF is a variable similar to JVF, but corrected for the N_{PV} dependence, defined as:

$$\text{corrJVF} = \frac{\sum_m p_{T,m}^{\text{track}}(PV_0)}{\sum_l p_{T,l}^{\text{track}}(PV_0) + \frac{\sum_{n \geq 1} p_{T,l}^{\text{track}}(PV_n)}{(k \cdot n_{\text{track}}^{PU})}} \quad (\text{K.1})$$

where $\sum_m p_{T,m}^{\text{track}}(PV_0)$ is the scalar sum of the p_T of the tracks that are associated with the jet and originate from the hard-scatter vertex. The term $\sum_{n \geq 1} p_{T,l}^{\text{track}}(PV_n) = p_T^{PU}$ denotes the scalar sum of the p_T of the associated tracks that originate from any of the pile-up interactions, n_{track}^{PU} is the total number of pile-up tracks per event and k is a scaling factor approximated by

the slope of $\langle p_T^{PU} \rangle$ with n_{track}^{PU} .

Fig. K.1.a presents the corrJVF distribution for pile-up and hard-scatter jets in simulated dijet events showing its discriminator power. A value corrJVF = 1 is assigned to jets with no associated tracks.

Figure K.1: Distribution of corrJVF (a) and R_{p_T} (b) for pile-up (PU) and hard-scatter (HS) jets with $20 < p_T < 30$ GeV [139].

The variable R_{p_T} is defined as the scalar sum of the p_T of the tracks that are associated with the jet and originate from the hard-scatter vertex divided by the fully calibrated jet p_T , which includes pile-up subtraction:

$$R_{p_T} = \frac{\sum_m p_{T,m}^{\text{track}}(PV_0)}{p_T^{\text{jet}}} \quad (\text{K.2})$$

Fig. K.1.b shows the R_{p_T} distribution for pile-up and hard-scatter jets in simulated dijet events. R_{p_T} is peaked at 0 and is steeply falling for pile-up jets, since tracks from the hard-scatter vertex rarely contribute. For hard-scatter jets, however, R_{p_T} has the meaning of a charged p_T fraction and its mean value and spread are larger than for pile-up jets.

K.2 JVT

The jet-vertex-tagger (JVT) uses as inputs the corrJVF and R_{p_T} variables in a two-dimensional likelihood derived using simulated di-jet events and based

on a k-nearest neighbor (kNN) algorithm [81]. For each point in the two-dimensional corrJVF - R_{p_T} plane, the relative probability for a jet at that point to be of signal type is computed as the ratio of the number of hard-scatter jets to the number of hard-scatter plus pile-up jets found in a local neighborhood around the point using a training sample of signal and pile-up jets with $20 < p_T < 50$ GeV and $|\eta| < 2.4$ [139].

Figure K.2: (a) Fake rate from pile-up jets versus hard-scatter jet efficiency curves for JVF, corrJVF , R_{p_T} , and JVT in a sample of jets with $20 < p_T < 50$ GeV and $|\eta| < 2.4$ in simulated dijet events. (b) The fake rate from pile-up jets versus hard-scatter jet efficiency curves for JVT separating jets initiated by gluons light- and b-quarks (light= u , d and s) [139].

Fig. K.2.a shows the fake rate versus efficiency curves comparing the performance of the four variables JVF, corrJVF , R_{p_T} , and JVT. It can be seen, clearly, an improve using the JVT with respect the other variables, because it can be maintain a high efficiency with a low fake rate.

This discriminant is also useful on heavy-flavor quark studies, because the discrimination for pile-up versus hard-scatter jet is better for hard-scatter jets initiated by b-quarks than light-quarks, as can be seen in Fig. K.2.b.

Bibliography

- [1] The Large Hadron Collider. <http://home.web.cern.ch/topics/large-hadron-collider>. Accessed: 2015-10-15. 1, 10
- [2] The ATLAS Collaboration. “*The ATLAS Experiment at the CERN Large Hadron Collider*”. *JINST*, 3 (2008) S08003. 1
- [3] The CMS Collaboration. “*The CMS experiment at the CERN LHC*”. *JINST*, 3 (2008) S08004. 1
- [4] C.P. Burgess and G.D. Moore. *The Standard Model: A Primer*. Cambridge University Press - Cambridge, 2013. 1
- [5] P.W. Higgs. “*Broken symmetries, massless particles and gauge fields*”. *Phys.Rev.Lett.*, 12:132-133, 1964. 1, 3
- [6] The ATLAS Collaboration. “*Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*”. *Phys.Lett. B*, 716:1-29, 2012. 1
- [7] The CMS Collaboration. “*Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*”. *Phys.Lett. B*, 716:30-61, 2012. 1
- [8] The ATLAS Collaboration. “*The ATLAS LAr Calorimeter: Phase I Upgrade TDR*”. ATLAS-TDR-022-2013, 2013. xi, xvii, xx, 1, 14, 50, 51, 160, 191, 192, 194
- [9] W. Greiner and B. Muller. *Quantum Mechanics: Symmetries*. Springer - New York, 2001. 2
- [10] S. van Dam. Spontaneous symmetry breaking in the Higgs mechanism. <http://philsci-archive.pitt.edu/9295/>. Accessed: 2017-04-19. 3

- [11] Elementary particles. https://en.wikipedia.org/wiki/Elementary_particle. Accessed: 2017-04-19. ix, 3
- [12] PDG: Electroweak model and constraints on new physics. <http://pdg.lbl.gov/2017/reviews/rpp2016-rev-standard-model.pdf>. Accessed: 2017-07-10. 4
- [13] The CMS Collaboration. “*Measurement of the weak mixing angle with the Drell-Yan process in proton-proton collisions at the LHC*”. *Phys.Rev. D*, 84:112002, 2011. 4
- [14] PDG: Gauge and Higgs bosons. <http://www-pdg.lbl.gov/2017/tables/rpp2017-sum-gauge-higgs-bosons.pdf>. Accessed: 2017-07-10. 4
- [15] ATLAS Standard Model Public Results. <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/StandardModelPublicResults>. Accessed: 2016-06-30. ix, x, xii, 5, 22, 71
- [16] D. Green. *High p_T Physics at Hadron Colliders*. Cambridge University Press - Cambridge, 2005. 6
- [17] The Super-Kamiokande Collaboration. “*Evidence for Oscillation of Atmospheric Neutrinos*”. *Phys.Rev.Lett.*, 81:1562, 1998. 6
- [18] Dark Energy, Dark Matter. <http://science.nasa.gov/astrophysics/focus-areas/what-is-dark-energy>. Accessed: 2016-04-04. 6
- [19] N. Craig. “*The State of Supersymmetry after Run I of the LHC*”. [arXiv:1309.0528](https://arxiv.org/abs/1309.0528), 2014. 6
- [20] D. Griffiths. *Introduction to Elementary Particles*. John Wiley & Sons - New York, 1987. 7, 213, 219, 220
- [21] F. Halzerm and A.D. Martin. *Quarks and Leptons: An Introductory Course in Modern Particle Physics*. John Wiley & Sons - New York, 1984. 214
- [22] R. Placakyte. “*Parton Distribution Functions*”. [arXiv:1111.5452](https://arxiv.org/abs/1111.5452), 2011. 8, 9

- [23] The H1 and ZEUS Collaborations. “*Structure Functions and PDF determination at HERA*”. *JHEP B*, 1001 (2010) 109. 9
- [24] Martin-Stirling-Thorne-Watt Parton Distribution Functions. <https://mstwpdf.hepforge.org/>. Accessed: 2016-04-04. ix, 9
- [25] H.L. Lai et al. “*New parton distributions for collider physics*”. *Phys.Rev. D*, 82:074024, 2010. 9, 81, 97, 161, 168
- [26] R.D. Ball et al. “*Parton distributions for the LHC Run II*”. *JHEP*, 04 (2015) 040. 9, 104
- [27] E. Wilson. *An Introduction to Particle Accelerators*. Oxford University Press - Oxford, 2005. 9
- [28] The Super Proton Synchrotron. <http://home.web.cern.ch/about/accelerators/super-proton-synchrotron>. Accessed: 2015-10-15. 10
- [29] The UA1 Collaboration. “*Experimental observation of isolated large transverse energy electrons with associated missing energy at $s=540$ GeV*”. *Phys.Lett. B*, 122:103-116, 1983. 10
- [30] The UA2 Collaboration. “*Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN pp collider*”. *Phys.Lett. B*, 122:476-485, 1983. 10
- [31] The UA1 Collaboration. “*Experimental Observation of Lepton Pairs of Invariant Mass Around 95 GeV/c**2 at the CERN SPS Collider*”. *Phys.Lett. B*, 126:398-410, 1983. 10
- [32] The UA2 Collaboration. “*Evidence for $Z^0 \rightarrow e^+ e^-$ at the CERN anti-p p Collider*”. *Phys.Lett. B*, 129:130-140, 1983. 10
- [33] The Tevatron. <http://www.fnal.gov/pub/tevatron/tevatron-accelerator.html>. Accessed: 2015-10-15. 10
- [34] The Collider Detector at Fermilab. <http://www-cdf.fnal.gov/>. Accessed: 2015-10-15. 10
- [35] The D0 experiment. <http://www-d0.fnal.gov/>. Accessed: 2015-10-15. 10

- [36] The CDF Collaboration. “*Observation of Top Quark Production in $p\bar{p}$ Collisions with the Collider Detector at Fermilab*”. *Phys.Rev.Lett*, 74:2626, 1995. [10](#)
- [37] The D0 Collaboration. “*Observation of the Top Quark*”. *Phys.Rev.Lett*, 74:2632, 1995. [10](#)
- [38] The ATLAS experiment. <http://www.atlas.ch>. Accessed: 2016-04-04. [ix](#), [x](#), [xi](#), [11](#), [13](#), [16](#), [23](#), [29](#), [32](#), [46](#), [56](#)
- [39] The LHC: Collisions. <http://lhc-machine-outreach.web.cern.ch/lhc-machine-outreach/collisions.htm>. Accessed: 2016-07-13. [12](#)
- [40] Z. Marshall for the ATLAS Collaboration. “*Simulation of Pile-up in the ATLAS Experiment*”. *J.Phys.:Conf.Ser.*, 13 022024, 2014. [12](#), [13](#)
- [41] ATLAS Luminosity Public Results in Run 2. <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2>. Accessed: 2017-07-07. [ix](#), [xxviii](#), [14](#), [15](#), [117](#)
- [42] S.M. White. Luminosity Scans at the LHC. https://cds.cern.ch/record/1357865/files/SW_LumiDays_pdf.pdf. Accessed: 2017-07-19. [14](#)
- [43] The ATLAS Collaboration. “*The implementation of the ATLAS missing Et triggers for the initial LHC operation*”. *ATL-COM-DAQ-2011-010*, 2011. [18](#)
- [44] Internal presentations of Jets and Missing Energy of ATLAS. <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/JetEtMiss>. Accessed: 2016-04-04. [x](#), [xii](#), [19](#), [65](#), [74](#), [75](#)
- [45] The ATLAS Collaboration. “*Search for direct slepton and gaugino production in final states with two leptons and missing transverse momentum with the ATLAS detector in pp collisions at $\sqrt{s} = 7$ TeV*”. *Phys.Lett. B*, 718 (2013) 879-901. [20](#), [25](#)
- [46] B. Heinemann. Particle Physics from Tevatron to LHC. <http://www-atlas.lbl.gov/~beate/homepage/pisa/pisa-2.pdf>. Accessed: 2016-07-07. [x](#), [20](#)

- [47] The ATLAS Collaboration. “*Measurement of the t -channel single top-quark production cross section in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*”. *Phys.Lett. B*, 717:330-350, 2012. [21](#)
- [48] The ATLAS Collaboration. “*Measurement of the cross-section for W boson production in association with b -jets in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*”. *JHEP*, 06 (2013) 084. [21](#)
- [49] F. Maltoni, T. McElmurry, S. Willenbrock. “*Inclusive production of a Higgs or Z boson in association with heavy quarks*”. *Phys.Rev. D*, 72:074024, 2005. [21](#)
- [50] J. Campbell, R.K. Ellis, F. Maltoni, S. Willenbrock “*Associated Production of a Z Boson and a Single Heavy-Quark Jet*”. *Phys.Rev. D*, 69:074021, 2004. [21](#)
- [51] The ATLAS Collaboration. “*Measurement of the $t\bar{t}$ production cross-section using $e\mu$ events with b -tagged jets in pp collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS detector*”. *Eur.Phys.J.*, C74 (2014) 3109. [21](#)
- [52] The ATLAS Collaboration. “*Search for the Standard Model Higgs boson produced in association with a vector boson and decaying to a b -quark pair with the ATLAS detector*”. *Phys.Lett. B*, 718:369, 2012. [21](#)
- [53] The ATLAS Collaboration. “*Search for light top squark pair production in final states with leptons and b -jets with the ATLAS detector in $\sqrt{s} = 7$ TeV proton-proton collisions*”. *Phys.Lett. B*, 720:13-31, 2013. [21](#)
- [54] K. Foraz. “*LS1 general planning and strategy for the LHC, LHC injectors*”. *CERN-2012-006.212*, 2012. [23](#)
- [55] M. Bernardini, K. Foraz. “*LONG SHUTDOWN 2 @ LHC*”. *Internal note for the LS2*, 2012. [23](#)
- [56] The ATLAS collaboration. “*Physics at a High-Luminosity LHC with ATLAS*”. *ATL-PHYS-PUB-2013-007*, 2013. [24](#)
- [57] LHC Commissioning with beam. <https://lhc-commissioning.web.cern.ch/lhc-commissioning/schedule/LHC-long-term.htm>. Accessed: 2017-07-07. [x](#), [24](#)

- [58] The ATLAS collaboration. “*Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC*”. *CERN-PH-EP-2013-103*, 2014. 24
- [59] The ATLAS collaboration. “*Coupling properties of the new Higgs-like boson observed with the ATLAS detector at the LHC*”. *CERN-PH-EP-2012-140*, 2012. 25
- [60] A first LHC sighting of the Higgs boson in its favourite decay. <https://atlas.cern/updates/physics-briefing/first-lhc-sighting-higgs-boson-its-favourite-decay>. Accessed: 2017-07-13. 25
- [61] K. Jakobs. “*Physics at the LHC: From Standard Model measurements to Searches for New Physics*”. [arXiv:1206.7024v1](https://arxiv.org/abs/1206.7024v1), 2009. 25, 26
- [62] The ATLAS Collaboration. “*The ATLAS Experiment at the CERN Large Hadron Collider*”. *JINST*, 3 (2008) S08003. x, 27, 28, 33, 34
- [63] L. Fabri. Forward Detectors in ATLAS: ALFA, ZDC and LUCID. <https://cds.cern.ch/record/1172844/files/ATL-LUM-SLIDE-2009-071.pdf>. Accessed: 2016-07-16. 27
- [64] The ATLAS Collaboration. “*The ATLAS Inner Detector: Technical Design Report*”. ATLAS-TDR-4, CERN/LHCC-97-16, 1997. 28
- [65] The ATLAS Collaboration. “*The ATLAS Inner Detector commissioning and calibration*”. *Eur.Phys.J.*, C70:787-821, 2010. 28
- [66] B. Zabinski. “*Methods of multiplicity reconstruction in heavy ion collisions in the ATLAS experiment*” *Acta Phys. Pol. B*, 42:1729, 2011. 29
- [67] IBL: A new subdetector for ATLAS. <http://home.web.cern.ch/about/updates/2014/05/new-subdetector-atlas>. Accessed: 2016-04-04. 30
- [68] The ATLAS Collaboration. “*ATLAS liquid argon calorimeter back end electronics*”. *JINST*, 5 (2010) P09003. 31
- [69] R. Wigmans. *Calorimetry Energy Measurement in Particle Physics*. Clarendon Press - Oxford, 2008. 32, 47

- [70] N. Nikiforou. “*Performance of the ATLAS Liquid Argon Calorimeter after three years of LHC operation and plans for a future upgrade*”. [arXiv:1306.6756](#), 2013. x, 34
- [71] R. Achenbach et al. “*The ATLAS Level-1 Calorimeter Trigger*”. *JINST*, 3 (2008) P03001. x, 37
- [72] The ATLAS Collaboration. “*ATLAS level-1 trigger: Technical Design Report*”. ATLAS-TDR-12, CERN-LHCC-98-014, 1998. xi, 37, 40, 41, 43
- [73] The ATLAS Collaboration. “*The ATLAS Level-1 Trigger System*”. *J.Phys.Conf.Ser.*, 396 (2012) 012010. 37, 190
- [74] The ATLAS Collaboration. “*Performance of the ATLAS muon trigger in pp collisions at $\sqrt{s} = 8$ TeV*”. *J.Phys.Conf.Ser.*, 513 (2014) 012040. 37
- [75] M. zur Nedden. The ATLAS Trigger System. https://www.physik.hu-berlin.de/de/gk1504/block-courses/autumn-2011/program_and_talks/23_2011-10_zurneddenmartin.pdf. Accessed: 2016-07-09. x, 38
- [76] I. Brawn et al. Bunch-Crossing Identification for the ATLAS First-Level Calorimeter Trigger. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.8.6354&rep=rep1&type=pdf>. Accessed: 2016-07-19. 37
- [77] S. Menke. Topological Clustering and Local Hadron Calibration. <https://agenda.linearcollider.org/event/5037/contributions/20899/attachments/17101/27735/pflow15022011Menke.pdf>. Accessed: 2016-07-09. xi, 39
- [78] ATLAS Public Jet Trigger Plots for Collision Data. <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/JetTriggerPublicResults>. Accessed: 2017-07-26. xi, 43
- [79] W. Lampl et al. “*Calorimeter clustering algorithms: Description and performance*”. *ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003*, 2011. 46, 163

- [80] The ATLAS collaboration. “*Expected electron performance in the ATLAS experiment*”. *ATL-PHYS-PUB-2011-006*, 2011. 47
- [81] A. Hoecker et al. “*TMVA - Toolkit for Multivariate Data Analysis*”. [arXiv:physics/0703039](https://arxiv.org/abs/physics/0703039), 2007. 50, 51, 52, 186, 297
- [82] The ATLAS Collaboration. “*Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*”. *Eur.Phys.J. C*, (2014) 74:3071. xi, 52, 53
- [83] J.E. Huth et al. “*Toward a standardization of jet definitions*”. *FERMILAB-CONF-90-249-E*, 1990. 55
- [84] G.C. Blazey et al. “*Run II Jet Physics: Proceedings of the Run II QCD and Weak Boson Physics Workshop*”. [arXiv:hep-ex/0005012](https://arxiv.org/abs/hep-ex/0005012), 2000. 57
- [85] G.P. Salam, G. Soyez. “*A practical Seedless Infrared-Safe Cone jet algorithm*”. *JHEP*, 0705 (2007) 086. 57
- [86] G.P. Salam. QCD at hadron colliders: Showers, Jets and fixed-order predictions. <https://gsalam.web.cern.ch/gsalam/repository/talks/2010-MariaLaach-lecture2.pdf>. Accessed: 2016-10-12. 58
- [87] G.P. Salam. “*Towards Jetography*”. *Eur.Phys.J. C*, 67:637-686, 2010. xi, 59
- [88] Y.L. Dokshitzer, G.D. Leder, S. Moretti, B.R. Webber. “*Better Jet Clustering Algorithms*”. *JHEP*, 9708 (1997) 001. 60
- [89] S.D. Ellis, D.E. Soper. “*Successive Combination Jet Algorithm For Hadron Collisions*”. *Phys.Rev. D*, 48:3160-3166, 1993. 60
- [90] M. Cacciari, G.P. Salam, G. Soyez. “*The anti- k_t jet clustering algorithm*”. *JHEP*, 0804 (2008) 063. xii, 60, 62
- [91] The FastJet Package. <http://fastjet.fr/>. Accessed: 2016-04-04. 64, 290
- [92] G.P. Salam. Jets at LHC: from basics to Higgs hunting. <https://gsalam.web.cern.ch/gsalam/talks/repo/2008-louvain-jets.pdf>. Accessed: 2015-10-09. xii, 64, 65

- [93] The ATLAS Collaboration. “*Local Hadronic Calibration*”. *ATL-LARG-PUB-2009-001*, 2009. 66
- [94] The ATLAS Collaboration. “*Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV*”. *Eur.Phys.J.*, C73 3 (2013) 2304. xii, 66, 67, 118
- [95] The ATLAS Collaboration. “*Performance of pile-up subtraction for jet shapes in pp collisions at $\sqrt{s} = 8$ TeV*”. *ATLAS-CONF-2013-085*, 2013. 68
- [96] The ATLAS Collaboration. “*Pile-up subtraction and suppression for jets in ATLAS*”. *ATLAS-CONF-2013-083*, 2013. xii, 68, 70
- [97] T. Sjostrand et al. “*An Introduction to PYTHIA 8.2*”. [arXiv:1410.3012](#), 2014. 81, 90, 161
- [98] The ATLAS collaboration. “*Jet global sequential corrections with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV*”. *ATLAS-CONF-2015-002*, 2015. 83, 118
- [99] J. M. Campbell and R. Ellis. “*MCFM for the Tevatron and the LHC*”. *Nucl. Phys. Proc. Suppl.*, 10:205-206, 2010. 90
- [100] MCFM - Monte Carlo for FeMtobarn processes. <http://mcfm.fnal.gov/>. Accessed: 2016-04-04. 90
- [101] MadGraph5_aMC@NLO web page. <http://amcatnlo.web.cern.ch/amcatnlo/>. Accessed: 2016-04-04. 90
- [102] J. Alwall, R. Frederix, S. Frixione et al. “*The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*”. *J.High Energ.Phys.*, 07:079, 2014. 90
- [103] F. Krauss, M. Schonherr, S. Schumann, F. Siegert, T. Gleisberg, S. Hoche and J. Winter. “*Event generation with SHERPA 1.1*”. *JHEP*, 02 (2009) 007. 90

- [104] A.D. Polosa, M. Morett, M.L. Mangano, F. Piccinini and R. Pittau. “*ALPGEN, a generator for hard multiparton processes in hadronic collisions*”. *Phys.Rev.*, 07:001, 2003. 90
- [105] Sherpa Manual Version 2.2.1. <https://sherpa.hepforge.org/doc/SHERPA-MC-2.2.1.html>. Accessed: 2015-10-09. 90
- [106] G. Corcella et al. “*HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*”. *JHEP*, 010 (2001) 1010. 90, 97
- [107] The ATLAS Collaboration. “*Flavor Tagging with Track Jets in Boosted Topologies with the ATLAS*”. *ATL-PHYS-PUB-2014-013*, 2014. 91
- [108] The ATLAS Collaboration. “*Performance of b-jet identification in the ATLAS experiment*”. *CERN-PH-EP-2015-216*, 2016. 92, 93, 94
- [109] b-tagging in ATLAS. <https://indico.in2p3.fr/event/6838/session/3/material/slides/0?contribId=1>. Accessed: 2016-11-19. xiii, 92
- [110] The ATLAS Collaboration. “*Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data*”. *ATLAS-CONF-2011-102*, 2011. 92
- [111] G. Piacquadio and C. Weiser. “*A new inclusive secondary vertex algorithm for b-jet tagging in ATLAS*”. *J.Phys.Conf.Ser.*, 119 032032, 2008. 93
- [112] G. Welch and G. Bishop. An Introduction to the Kalman Filter. http://www.cs.unc.edu/~tracker/media/pdf/SIGGRAPH2001_CoursePack_08.pdf. Accessed: 2017-07-20. 93
- [113] L. Teodorescu. “*Artificial neural networks in high-energy physics*”. *CERN Document Server*, 2009. 94
- [114] The ATLAS Collaboration. “*Measurement of the b-tag Efficiency in a Sample of Jets Containing Muons with 5 fb⁻¹ of Data from the ATLAS Detector*”. *ATLAS-CONF-2012-043*, 2012. xiii, 94

- [115] The 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 (2014) 141. [xiv](#), [xxv](#), [xxvii](#), [97](#), [98](#), [99](#), [120](#), [142](#), [158](#), [279](#), [280](#), [281](#), [282](#)
- [116] J.M. Butterworth, J.R. Forshaw, and M.H. Seymour. “*Multiparton interactions in photoproduction at HERA*”. *Z.Phys.*, C72 (1996) 637-646. [97](#)
- [117] J. Pumplin et al. “*New generation of parton distributions with uncertainties from global QCD analysis*”. *JHEP*, 0207 (2002) 012. [97](#)
- [118] S. Catani, F. Krauss, R. Kuhn, and B. Webber. “*QCD matrix elements + parton showers*”. *JHEP*, 11 (2005) 063. [97](#)
- [119] S. Hoche et al. “*QCD matrix elements + parton showers: The NLO case*”. *JHEP*, 04 (2013) 27. [97](#), [104](#)
- [120] C. Anastasiou, L. Dixon, K. Melnikov, and F. Petriello. “*High-precision QCD at hadron colliders: electroweak gauge boson rapidity distributions at NNLO*”. *Phys.Rev. D*, 69:094008, 2004. [106](#)
- [121] J.M. Campbell, R.K. Ellis, and C. Williams. “*Vector boson pair production at the LHC*”. *JHEP*, 1107 (2011) 018.
- [122] The ATLAS Collaboration. “*Measurement of the low-mass Drell-Yan differential cross section at $\sqrt{s} = 7$ TeV using the ATLAS detector*”. *JHEP*, 06 (2014) 112. [117](#)
- [123] G. D’Agostini. “*Improved iterative Bayesian unfolding*”. [arXiv:1010.0632](#), 2010. [129](#)
- [124] The CMS Collaboration. “*Measurement of the production cross sections for a Z boson and one or more b jets in pp collisions at $\sqrt{s} = 7$ TeV*”. *JHEP*, 06 (2014) 120. [xxvii](#), [100](#)
- [125] The CMS Collaboration. “*Measurements of the associated production of a Z boson and b jets in pp collisions at $\sqrt{s} = 8$ TeV*”. *CMS-SMP-14-010*, 2016. [xxvii](#), [101](#)

- [126] S. Schumann and F. Krauss. “A Parton shower algorithm based on Catani-Seymour dipole factorisation”. *JHEP*, 0803 (2008) 038. [104](#)
- [127] T. Gleisberg and S. Hoche. “Comix, a new matrix element generator”. *JHEP*, 0812 (2008) 039. [104](#)
- [128] The ATLAS Collaboration. “ATLAS simulation of boson plus jets processes in Run 2”. *ATL-PHYS-PUB-2017-006*, 2017. [xxviii](#), [104](#), [105](#)
- [129] R. Gavin, Y. Li, F. Petriello, S. Quackenbush. “FEWZ 2.0: A code for hadronic Z production at next-to-next-to-leading order”. *Comput.Phys.Commun.*, 182:2388-2403, 2011. [106](#)
- [130] HepMC event record. <http://hepmc.web.cern.ch/hepmc/>. [107](#)
- [131] The ATLAS Collaboration. “Measurement of the Production Cross Sections of a Z Boson in Association with Jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector”. *ATL-COM-PHYS-2016-017*, v1.6, 2017. [xxiv](#), [xxv](#), [112](#), [120](#), [275](#), [276](#), [277](#)
- [132] ATLAS Tracking CP Run 2 recommendations. <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/InDetTrackingPerformanceGuidelines>. [117](#)
- [133] ATLAS Electron and Photon Selection and Identification for Run2. <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/EGammaIdentificationRun2>. [118](#)
- [134] ATLAS JetEtMiss Recommendations. <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/JetEtmissRecommendations2016>. [118](#), [140](#)
- [135] ATLAS Internal Note. Measurement of vector bosons produced in association with b-jets. <https://svnweb.cern.ch/trac/atlasphys-sm/browser/Physics/StandardModel/WZPhysics/Analyses/13TeV/VplusHF/SupportNote/trunk/VHF2016internal.pdf?order=name>. Accessed: 2017-08-18. [xiv](#), [119](#)
- [136] M. Bahr, S. Gieseke, M.A. Gigg, et al. “Herwig++ Physics and Manual”. *Eur.Phys.J. C*, 58 (2008) 639. [120](#)

- [137] M. Czakon and A. Mitov. “*Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders*”. *Comput.Phys.Commun.*, 185 (2014) 2930. 120
- [138] RooUnfold: ROOT Unfolding Framework. <http://hepunx.rl.ac.uk/~adye/software/unfold/RooUnfold.html>. Accessed: 2016-11-19. 129
- [139] The ATLAS Collaboration. “*Performance of pile-up mitigation techniques for jets in pp collisions at 8 TeV using the ATLAS detector*”. *Eur.Phys.J. C*, 76 (2016) 581. xxvi, 296, 297
- [140] N. J. Buchanan et al. “*Design and implementation of the Front End Board for the readout of the ATLAS liquid argon calorimeters*”. *JINST*, 3 (2008) P03004. 190
- [141] C. Oleari. “*The POWHEG-BOX*”. *Nucl.Phys.Proc.Suppl.*, 205-206:36-41, 2010. 161
- [142] ATLAS Particle Generator. <https://svnweb.cern.ch/trac/atlasoff/browser/Generators/ParticleGenerator>. Accessed: 2015-04-31. 167
- [143] R. D. Ball, et al. “*Parton distributions with LHC data*”. *Nucl.Phys. B*, 867:244, 2014. 168
- [144] The ATLAS Collaboration. “*ATLAS Run 1 Pythia8 tunes, Technical Report*”. *ATL-PHYS-PUB-2014-021*, 2014. 168
- [145] A. Buckley et al. “*General-purpose event generators for LHC physics*”. *J.Phys.Rep.*, 03:005, 2011. xxi, 209, 210, 211
- [146] M. H. Seymour and M. Marx. “*Monte Carlo Event Generators*”. [arXiv:1304.6677](https://arxiv.org/abs/1304.6677), 2013. xx, xxi, 208, 209, 210, 212
- [147] GEANT 4. <http://geant4.cern.ch/>. 163, 206
- [148] D. J. Lange. “*The EvtGen particle decay simulation package*”. *Nucl.Instrum.Meth. A*462 (2001) 152-155. 211

- [149] P. Golonka et al. “*The tauola-photos-F environment for the TAUOLA and PHOTOS packages, release II*”. *Comput.Phys.Commun.* 174, 818-835, 2006. [211](#)
- [150] The FormCalc Visitor Center. <http://www.feynarts.de/formcalc/>. [213](#)
- [151] The FeynArts Visitor Center. <http://www.feynarts.de/>. [213](#)
- [152] Wolfram Mathematica. <https://www.wolfram.com/mathematica/>. [220](#)