

A SEARCH FOR SUPERSYMMETRY IN EVENTS
WITH A Z BOSON, JETS, AND MISSING
TRANSVERSE ENERGY IN $p p$ COLLISIONS
WITH $\sqrt{s}=13$ TEV WITH THE ATLAS DETECTOR

TOVA RAY HOLMES

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy
in Physics
in the Graduate Division
of the University of California, Berkeley



Committee in charge:
Professor Beate Heinemann, Chair
Professor Robert Jacobsen
Professor Karl van Bibber

Fall 2016

Tova Ray Holmes: *A Search for Supersymmetry in Events with a Z Boson, Jets, and Missing Transverse Energy in pp Collisions with $\sqrt{s}=13$ TeV with the ATLAS Detector*, © December 2016

ABSTRACT

A Search for Supersymmetry in Events with a Z Boson, Jets, and Missing Transverse Energy in pp Collisions with $\sqrt{s}=13$ TeV with the ATLAS Detector

by

Tova Ray Holmes

Doctor of Philosophy in Physics
University of California, Berkeley
Professor Beate Heinemann, Chair

A search for new phenomena in final states containing a Z boson decaying to electrons or muons, jets, and large missing transverse momentum is presented. This search uses proton–proton collision data collected during 2015 and 2016 at a center of mass energy $\sqrt{s} = 13$ TeV by the ATLAS detector at the Large Hadron Collider, which correspond to an integrated luminosity of 14.7 fb^{-1} . The search targets the pair production of supersymmetric particles, squarks or gluinos, which decay via jets and a Z boson to the lightest Supersymmetric particle, which does not interact with the ATLAS detector. Results are interpreted in simplified models of gluino-pair (squark-pair) production, and provide sensitivity to gluinos (squarks) with masses as large as 1.3 (1.0) TeV.

To my brother for at least starting to lead the way.

ACKNOWLEDGEMENTS

Put your acknowledgements here.

CONTENTS

I	INTRODUCTION	1
1	INTRODUCTION	2
II	THEORY AND MOTIVATION	5
2	THEORY AND MOTIVATION	6
2.1	The Standard Model	6
2.1.1	Matter	7
2.1.2	Forces	9
2.1.3	Phenomenology of Proton-Proton Collisions .	18
2.1.4	Problems in the Standard Model	21
2.2	Supersymmetry	23
2.2.1	The Minimal Supersymmetric Standard Model	24
2.2.2	Solutions to Standard Model Problems	25
2.2.3	Simplified Models of Supersymmetry	26
III	THE EXPERIMENT	32
3	THE LARGE HADRON COLLIDER	33
3.1	The Injector Complex	34
3.2	Operation of the Large Hadron Collider	35
3.3	Luminosity	36
3.4	Pile-up in proton-proton Collisions	37
4	THE ATLAS DETECTOR	40
4.1	Coordinate System Used in the ATLAS Detector . . .	40
4.2	The Inner Detector	41
4.2.1	The Pixel Detector	41
4.2.2	The Silicon Microstrip Tracker	43
4.2.3	The Transition Radiation Tracker	44
4.3	The Calorimeters	45
4.4	The Muon Spectrometer	47
4.5	The Magnet System	49
4.6	The Trigger System and Data Acquisition	50
4.7	Monte Carlo Event Generation	53
5	INNER DETECTOR TRACKING AND PIXEL CLUSTERING	58
5.1	Overview of Tracking in the ATLAS Detector	58
5.2	Clustering in the Pixel Detector	59
5.2.1	Charge Interpolation Method	60
5.2.2	Improving Measurement with Neural Networks	61
5.3	Impact of the Neural Network	63
5.3.1	The Neural Network in 13 TeV Data	64
6	OBJECT RECONSTRUCTION IN THE ATLAS DETECTOR	67
6.1	Electrons	67
6.2	Photons	70

6.3	Muons	72
6.4	Jets	75
6.5	Overlap Removal	78
6.6	Missing Transverse Momentum	80
IV	SEARCHING FOR SUPERSYMMETRY	83
7	BACKGROUND PROCESSES	84
7.1	Data and Monte Carlo Samples	87
8	OBJECT IDENTIFICATION AND SELECTION	89
8.1	Electrons	89
8.2	Muons	89
8.3	Jets	91
8.4	Photons	91
9	EVENT SELECTION	93
9.1	Trigger Strategy	95
9.2	Signal Efficiency and Contamination	96
10	BACKGROUND ESTIMATION	99
10.1	Flavor Symmetric Processes	99
10.1.1	Flavor Symmetry Method	99
10.1.2	Sideband Fit Method	101
10.2	Z/ γ^* + jets Background	103
10.2.1	Photon and Z Event Selection	107
10.2.2	Smearing of Photon Events	109
10.2.3	p_T Reweighting of Photon Events	110
10.2.4	Determining H_T and $m_{\ell\ell}$	113
10.2.5	Subtraction of $V\gamma$ Events	114
10.2.6	Validation in Data	115
10.3	Fake and Non-Prompt Leptons	116
10.4	Diboson and Rare Top Processes	119
11	SYSTEMATIC UNCERTAINTIES	125
11.1	Uncertainties on Data-Driven Backgrounds	125
11.1.1	Uncertainties on the Flavor Symmetry Method	125
11.1.2	Uncertainties on the γ + jets Method	129
11.1.3	Uncertainties on the Fakes Background	133
11.2	Theoretical and Experimental Uncertainties	133
11.3	Impact of Uncertainties on the Signal Region	136
12	RESULTS	137
13	INTERPRETATIONS	143
V	CONCLUSIONS	147
14	CONCLUSIONS AND OUTLOOK	148
	BIBLIOGRAPHY	150

LIST OF FIGURES

Figure 1	The Standard Model of particle physics, containing all known bosons and fermions, with the addition of the hypothetical graviton. [7] .	7
Figure 2	Gluon self coupling Feynman diagrams involving 3- and 4-gluon interactions.	11
Figure 3	Quark couplings to the different types of gauge bosons. The $q_{u,d}$ labels represent any up- or down-type quarks.	12
Figure 4	The running of the strong coupling constant, α_s . [14]	13
Figure 5	Feynman diagrams of trilinear gauge couplings in the Standard Model (SM).	14
Figure 6	Feynman diagrams of weak couplings to leptons in the SM	15
Figure 7	Feynman diagrams demonstrating Higgs couplings to the weak gauge bosons in the SM	17
Figure 8	Feynman diagrams showing Higgs couplings to fermions in the SM	17
Figure 9	2008 MSTW Parton Distribution Functions (PDFs) for various particle types given as a function of x and Q^2 , the square of the parton-parton momentum transfer. [15]	18
Figure 10	Cross-sections for many SM processes as a function of \sqrt{s} [16].	20
Figure 11	Galactic rotation curve of velocity as a function of radius in NGC 3198. Included is the observed data, as well as the expected velocity distribution from a disk-shaped galaxy corresponding to the expected density from electromagnetic observations. Another curve corresponding to a halo-shaped matter distribution is superimposed, and the halo and disk are summed and fit to the data. [20]	22
Figure 12	Two example vertices allowed by the Minimal Supersymmetric Standard Model (MSSM). . . .	24
Figure 13	Running of the strong, weak, and electromagnetic coupling constants for the SM (left) and MSSM (right). [27]	26
Figure 14	Feynman diagram of the decay considered in the simplified models used in the analysis presented in Part iv	27

Figure 15	13 TeV production cross-sections for sparticles, as a function of sparticle mass [31].	28
Figure 16	Results of an 8 TeV search performed by the A Toroidal LHC Apparatus (ATLAS) collaboration in a signal region targeting events like those in Figure 14 . The events in the signal region are displayed as a function of $m_{\ell\ell}$, the invariant mass of the event's leading leptons. The SM backgrounds are shown with their full uncertainties based on data-driven background estimations, and two signals are superimposed on the distribution. The observed data-points are higher than the expected background, with a total excess of 3.0σ [1].	29
Figure 17	Results of an 8 TeV search performed by the Compact Muon Solenoid (CMS) collaboration in a signal region including a broad range of $m_{\ell\ell}$. A 2.4σ local excess is seen in the low $m_{\ell\ell}$ region, and no excess of events is seen in the region with $m_{\ell\ell}$ consistent with an on-shell Z boson. The data is fit based on a data driven estimate of the flavor symmetric background (FS) and the Drell-Yan background (DY), with an additional component for the signal [34]. . .	30
Figure 18	Preliminary results from a 13 TeV search targeting the same signal region as Figure 16 , performed on 3.2 fb^{-1} of 2015 data. The events in the signal region are displayed as a function of $m_{\ell\ell}$, the invariant mass of the event's leading leptons. Flavor symmetric and $Z/\gamma^* + \text{jets}$ backgrounds are taken from data-driven methods, while the other backgrounds are taken from Monte Carlo simulation (MC). They are compared to the data, which shows a 2.2σ excess of events. Distributions from two signal points are superimposed [2].	31
Figure 19	The Large Hadron Collider (LHC) main collider ring and pre-accelerator Super Proton Synchrotron (SPS) overlaid on a map of Switzerland and France, with the four main LHC experiments identified.	33
Figure 20	The pre-accelerators of the LHC	34
Figure 21	Cross-section of a cryodipole magnet in the LHC	36

Figure 22	Beam spot in the ATLAS detector for one run in 2015. Distributions show only the highest p_T vertex per event. Left is the $x - y$ distribution of vertices, while the right plot shows the $x - z$ distribution.	36
Figure 23	Instantaneous luminosity of one fill of 7 TeV data in 2011.	38
Figure 24	ATLAS luminosity for Run 1 and Run 2, as of September 2016.	38
Figure 25	Average number of interactions per crossing shown for 2015 and 2016 separately, as well as the sum of the two years.	39
Figure 26	Diagram of the ATLAS detector, with subsystems and magnets identified.	40
Figure 27	Diagram of the ATLAS Inner Detector, containing the Pixel, SCT, and TRT subsystems.	42
Figure 28	Diagram of one-quarter of the ATLAS Inner Detector in the $R - z$ plane, with lines drawn to indicate various η locations.	43
Figure 29	Diagram in the $x - y$ plane of the Insertable B-Layer (IBL) and the innermost layer of the Pixel Detector, Lo [42].	44
Figure 30	The calorimeter system of the ATLAS detector.	45
Figure 31	Layout of the LAr calorimeter module at central η [38].	46
Figure 32	An $x - y$ view of the Muon Spectrometer (MS). The three barrel layers are visible, as well as the overlapping, differently sized chambers. The outer layer of the MS is about 20m in diameter.	47
Figure 33	An $r - z$ view of the MS . The three layers of the barrel and endcap MS are visible.	48
Figure 34	The magnet system of the ATLAS detector. The inner cylinder shows the solenoid which gives a uniform magnetic field in the Inner Detector (ID). Outside of that are the barrel and endcap toroids, which provide a non-uniform magnetic field for the MS	49
Figure 35	Plots of the magnetic field within the ATLAS detector. Top is the field (broken into its R and z components) as a function of z for several different values of R . Bottom is the field integral through the Monitored Drift Tubes (MDTs) as a function of $ \eta $ for two different ϕ values.	51

Figure 36	Level One (L₁) trigger rates for for a run in July 2016 as a function of luminosity block, an approximately 60-second long period of data-taking. The total rate is lower than the combined stack because of overlapping triggers.	52
Figure 37	High Level Trigger (HLT) trigger rates for for a run in July 2016 as a function of luminosity block, an approximately 60-second long period of data-taking. The total rate is lower than the combined stack because of overlapping triggers.	53
Figure 38	Photon trigger efficiency as a function of E_T for four different HLT triggers with photon p_T requirements of 25, 35, 120, and 140 GeV [44].	54
Figure 39	Number of secondary vertices in a module in the first layer of the Pixel Detector in data (top) and MC (bottom). There are more events in the data than the MC [54].	56
Figure 40	Event display from June 2015, with particle tracks in light blue. The main image displays a view of the $x - y$ plane of the ID . The IBL and Lo of the Pixel Detector are shown in red, with the remaining two layers of the Pixel Detector in green and blue. Outside those are the four double layers of the Silicon Microstrip Tracker (SCT). The smaller image on the right shows an $r - z$ view, zoomed in to only show hits in the IBL [55].	58
Figure 41	A few possible types of clusters in the Pixel Detector. (a) shows a single particle passing through a layer of the detector, (b) shows two particles passing through the detector, creating a single merged cluster, and (b) shows a single particle emitting a δ -ray as it passes through the detector [57].	60

Figure 42	One example of a two-particle cluster and its truth information compared with the output of the Neural Networks (NNs). The boxes represent pixels, with a color scale indicating charge. At top, the $p(N = i)$ values give the output of the Number NN, the probabilities that the cluster contains 1, 2, and 3 particles. Given the highest probability is for $N = 2$, the other NNs predict the position and errors of the two particles (in white). The black arrows and squares represent the truth information from the cluster, and the black dot and dotted line show the position measurement for the un-split cluster [57].	62
Figure 43	x resolutions for clusters with 3 (top) and 4 (bottom) pixels in the x direction in 7 TeV data for Connected Component Analysis (CCA) (using only charge interpolation to determine position) and NN clustering taken from MC [57]. . .	63
Figure 44	Performance of the pixel neural network used to identify clusters created by multiple charged particles, as a function of constant coherent scaling of the charge in each pixel in the cluster. The top figure shows the rate at which the neural network wrongly identifies clusters with one generated particle as clusters with multiple particles. The bottom figure shows the rate at which the neural network correctly identifies clusters generated by multiple particles as such.	65
Figure 45	Fraction of cluster classes as a function of the distance between tracks for IBL (top) and 2nd pixel layer (bottom).	66
Figure 46	Identification efficiencies from MC samples for Loose, Medium, and Tight working points. Left is the efficiency for identification of true electrons taken from $Z \rightarrow ee$ MC, and right is the efficiency for mis-identification of jets as electrons taken from dijet MC [62].	68
Figure 47	Combined electron reconstruction and identification efficiencies measured as a function of η for data (using the tag-and-probe method on $Z \rightarrow ee$ events) and $Z \rightarrow ee$ MC. Distributions include electrons with $E_T > 15$ GeV. [62].	69

Figure 48	Comparison of Tight identification efficiency measurements from data and $Z \rightarrow \ell\ell\gamma$ MC for unconverted (left) and converted (right) photons, with an inclusive η selection. The bottom of each figure shows the ratio of data and MC efficiencies. [65].	71
Figure 49	Muon reconstruction efficiency for the Medium and Loose working points measured with $Z \rightarrow \mu\mu$ events in data and in MC as a function of η . The ratio between the two is shown at the bottom. The Loose working point efficiency is shown only at small $ \eta $, where the loosened requirements cause the largest difference from the Medium working point [67].	73
Figure 50	Muon reconstruction efficiency for the High- p_T working point measured with $Z \rightarrow \mu\mu$ events in data and in MC as a function of η . The ratio between the two is shown at the bottom. [67]	74
Figure 51	Distribution of event p_T density, ρ , taken from MC dijets for different numbers of primary vertices. [69]	76
Figure 52	Energy response as a function of energy and η for Electromagnetic (EM) jets in dijet MC. [69] .	77
Figure 53	Dijet MC distributions of the number of pile-up jets passing the Jet Vertex Tagger (JVT) and Jet Vertex Fraction (JVF) cuts (left) and the efficiency for jets from the primary vertex (right) as a function of number of primary vertices in the event [70].	78
Figure 54	Distribution of MV2c20 output for b -jets, c -jets, and light-flavor jets in $t\bar{t}$ MC [71].	79
Figure 55	Distributions of the resolution of the x and y components of Track Soft Term (TST) E_T^{miss} in $Z \rightarrow \mu\mu$ events in data and MC.	81
Figure 56	Distributions of the jet term (top left), muon term (top right), and TST (bottom) E_T^{miss} in $Z \rightarrow \mu\mu$ events in data and MC. In the jet term distribution, the feature at zero is due to events with no jets, and the spike at 20 GeV corresponds to the minimum jet p_T considered for the analysis [73].	82
Figure 57	An example Feynman diagram of $t\bar{t}$ production and decay.	84
Figure 58	An example Feynman diagram of the production and decay of a WZ event.	85

Figure 59	An example Feynman diagram of the production and decay of a $Z/\gamma^* + \text{jets}$ event.	86
Figure 60	Schematic layout of the signal region, as well as the control and validation regions used to predict the Flavor Symmetric (FS) background. The regions displayed differ only by their $m_{\ell\ell}$ and E_T^{miss} values, except for CR-FS, which contains different-flavor leptons, while all other regions contain same-flavor leptons.	94
Figure 61	Signal region acceptance (left) and efficiency (right) in SRZ for the simplified model with gluino pair production with $\tilde{\chi}_2^0$ decays to $\tilde{\chi}_1^0$ and an on-shell Z boson with 1 GeV neutralino LSP. Acceptance is calculated by applying the signal-region kinematic requirements to truth objects in MC , which do not suffer from identification inefficiencies or measurement resolutions.	97
Figure 62	Expected signal contamination in CRT (left) and CR-FS (right) for the signal model with gluino pair production, where the gluinos decay to quarks and a neutralino, with the neutralino subsequently decaying to a Z boson and a 1 GeV neutralino LSP.	97
Figure 63	Expected signal contamination in VRS (left) and VRT (right) for the signal model with gluino pair production, where the gluinos decay to quarks and a neutralino, with the neutralino subsequently decaying to a Z boson and a 1 GeV neutralino LSP.	98
Figure 64	Measurements of k_e , the ratio of electron to muon events, in bins of p_T and η , with statistical uncertainties shown. On the top is the measurements indexed by the leading lepton, while the measurements indexed by the sub-leading lepton are on the bottom. The efficiencies displayed are for the 2016 dataset.	102
Figure 65	Lepton p_T distribution in data and MC , demonstrating the overestimate of the high- p_T tail in MC [91].	105
Figure 66	MC comparison of boson p_T in a selection of photon and $Z \rightarrow \ell\ell$ events with $H_T > 600$ GeV. The feature at 1000 GeV is not physical, but a reflection of the change in bin size.	107

Figure 67	$E_{T,\parallel}^{\text{miss}}$ distributions in MC for $Z+\text{jets}$ in the ee (left) and $\mu\mu$ (right) channels compared to the $\gamma + \text{jets}$ prediction in an inclusive region with $H_T > 600$ GeV.	109
Figure 68	$E_{T,\parallel}^{\text{miss}}$ distributions in MC for $Z+\text{jets}$ ee (left) and $\mu\mu$ (right) channels compared to $\gamma + \text{jets}$ in an inclusive region with $H_T > 600$ GeV after the smearing procedure has been performed. The events in these distributions have also been p_T reweighted, as described in Section 10.2.3	110
Figure 69	Photon reweighting factors for the ee (left) and $\mu\mu$ (right) channels derived from data and MC	111
Figure 70	E_T^{miss} distribution comparing MC distributions of photon and Z events before any smearing is applied (top), with only p_T reweighting applied (bottom left), and after p_T reweighting and smearing have both been applied (bottom right) in the ee channel of 2016 data.	112
Figure 71	E_T^{miss} distribution comparing MC distributions of photon and Z events before any smearing is applied (top), with only p_T reweighting applied (bottom left), and after p_T reweighting and smearing have both been applied (bottom right) in the $\mu\mu$ channel of 2016 data.	113
Figure 72	$Z/\gamma^* + \text{jets}$ MC $m_{\ell\ell}$ distribution compared to the prediction from $\gamma + \text{jets}$ method performed on MC (left) and the prediction from $\gamma + \text{jets}$ method performed on data (right).	115
Figure 73	Comparison of data and MC in CR- γ without any H_T cut, including the contributions from various $V\gamma$ processes.	115
Figure 74	Total $\gamma + \text{jets}$ data prediction in SRZ (excluding the E_T^{miss} cut) and the prediction after the $V\gamma$ subtraction.	116
Figure 75	E_T^{miss} distribution in VRZ ee (left) and $\mu\mu$ (right) with total data yield compared to the sum of the prediction from the $\gamma + \text{jets}$ method, the prediction from the flavor symmetry method, the prediction from the fake background estimation (included under “other”), and the remaining backgrounds taken from MC	117

Figure 76	$\Delta\phi(\text{jet}, \mathbf{p}_T^{\text{miss}})$ distribution in for the leading jet (left) and the subleading jet (right). The comparison is performed in VRZ with the cut on $\Delta\phi(\text{jet}_{12}, \mathbf{p}_T^{\text{miss}})$ removed. The total data yield is compared to the sum of the prediction from the $\gamma + \text{jets}$ method, the prediction from the flavor symmetry method, the prediction from the fake background estimation (included under “other”), and the remaining backgrounds taken from MC.	117
Figure 77	Sub-leading lepton p_T for ee (left) and $\mu\mu$ (right) events in the tight-tight region used to measure the real-lepton efficiency for 2016.	119
Figure 78	Sub-leading lepton p_T for μe (left) and $\mu\mu$ (right) events in the tight-tight region used to measure the fake-lepton efficiency for 2016.	120
Figure 79	Same sign validation regions in the ee (top left), $\mu\mu$ (top right), $e\mu$ (bottom left) and μe (bottom right) channels combining 2015+2016 data. Uncertainty bands include both statistical and systematic uncertainties.	120
Figure 80	Distribtuions of data and MC in VR-WZ. Reconstructed transverse mass of the W (top) and mass of the Z (bottom).	122
Figure 81	Distribtuions of data and MC in VR-WZ. p_T of the W (top) and Z (bottom).	123
Figure 82	Distribtuions in VR-WZ. On the top, mass of the Z bosons in the event, and on the bottom, p_T of the Z bosons.	124
Figure 83	MC closure plots of VRS (top) and SRZ (bottom). The number of events from MC (black points) is compared to the number of events predicted from the flavor symmetry method (yellow histogram). The comparison is performed before the expanded $m_{\ell\ell}$ window is used to predict the on-Z bin, but because the shape is taken from the same MC, the result is identical.	126
Figure 84	α , the trigger efficiency ratio, calculated as a function of E_T^{miss} from three different sources: data (blue), the usual skimmed $t\bar{t}$ MC (red), and an unskimmed $t\bar{t}$ MC (green).	128
Figure 85	Plots of the fraction of on-Z events with a VR-FS-like selection as a function of H_T . The top figure shows 2015 data and MC while the bottom figure shows the same for 2016.	129

Figure 86	E_T^{miss} distributions for $\gamma + \text{jets}$ predictions using different reweighting variables, as well as distributions with the nominal reweighting but with smearing functions taken from data and from MC in a ≥ 2 -jet region.	130
Figure 87	MC closure of the $\gamma + \text{jets}$ method as a function of E_T^{miss} comparing the MC prediction of the Z background with the $\gamma + \text{jets}$ method performed on $\gamma + \text{jets}$ MC . The uncertainty band includes both statistical and reweighting uncertainties.	131
Figure 88	Distributions of $m_T(\ell, E_T^{\text{miss}})$, the transverse mass of the lepton and the E_T^{miss} in a Validation Region (VR) designed to target $W\gamma$ processes. Top is the distribution with a E_T^{miss} cut at 100 GeV, and bottom is the same distribution with a E_T^{miss} cut of 200 GeV.	132
Figure 89	Comparison of background predictions and data yields in four validation regions, as well as the signal region. Definitions of all regions can be found in Table 7 , with both rare top and fake backgrounds grouped together under the “other” label. The uncertainty band includes all statistical and systematic uncertainties. Below is a panel of the one-sided statistical significances of the deviations between the predicted and observed quantities for each region.	137
Figure 90	Comparisons as a function of $m_{\ell\ell}$ of background predictions with observed data in an SRZ-like region, with the $m_{\ell\ell}$ cut removed. Left is the same-flavor channel, where all background shapes are taken from MC and scaled to their SRZ predictions, except for the $Z/\gamma^* + \text{jets}$ background, which is taken entirely from the data-driven background. Right is the different-flavor channel, in which the backgrounds are taken directly from MC , except for $t\bar{t}$, which is scaled to match the total data yield.	139

Figure 91 Distributions of observed data, background predictions, and simulated signals are shown in SRZ as a function of $m_{\ell\ell}$, $p_T^{\ell\ell}$, E_T^{miss} , H_T , number of jets, and number of b -jets. The two example signals have $(m(\tilde{g}), m(\tilde{\chi}_2^0)) = (1095, 205)$ GeV. All background shapes are taken from MC, and in the case of flavor symmetric and $Z/\gamma^* + \text{jets}$ backgrounds, their yields are scaled to match the data-driven predictions. Uncertainties include statistical and systematic components. 141

Figure 92 Comparisons as a function of $\Delta\phi(\text{jet}_{12}, p_T^{\text{miss}})$ of background predictions with observed data in an SRZ-like (left) and VRS-like (right) region, with the $\Delta\phi(\text{jet}_{12}, p_T^{\text{miss}})$ cut removed. All background shapes are taken from MC and scaled to their SRZ predictions, except for the $Z/\gamma^* + \text{jets}$ background, which is taken entirely from the data-driven background. 142

Figure 93 Expected and observed exclusion contours derived from the results in SRZ for the (top) $\tilde{g}-\tilde{\chi}_2^0$ on-shell grid and (bottom) $\tilde{q}-\tilde{\chi}_2^0$ on-shell grid. The dashed blue line indicates the expected limits at 95% CL and the yellow band shows the 1σ variation of the expected limit as a consequence of the uncertainties in the background prediction and the experimental uncertainties in the signal ($\pm 1\sigma_{\text{exp}}$). The observed limits are shown by the solid red line, with the dotted red lines indicating the variation resulting from changing the signal cross section within its uncertainty ($\pm 1\sigma_{\text{theory}}^{\text{SUSY}}$). 143

Figure 94 Expected and observed exclusion contours derived from the results in SRZ for the $\tilde{g}-\tilde{\chi}_1^0$ on-shell grid. The dashed blue line indicates the expected limits at 95% CL and the yellow band shows the 1σ variation of the expected limit as a consequence of the uncertainties in the background prediction and the experimental uncertainties in the signal ($\pm 1\sigma_{\text{exp}}$). The observed limits are shown by the solid red line, with the dotted red lines indicating the variation resulting from changing the signal cross section within its uncertainty ($\pm 1\sigma_{\text{theory}}^{\text{SUSY}}$). 145

Figure 95 Expected 95% Confidence Level ([CL](#)) exclusion contours (dashed) and 5σ discovery contours (solid) for $L_{int} = 300^{-1}$ (black) and 3000^{-1} (red) for gluino pair-production, with 1σ bands representing the uncertainty on the production cross-section. Superimposed is the observed 8 TeV exclusion for similar models. [111] 149

LIST OF TABLES

Table 1	Supermultiplets of supersymmetric and SM particles. Sfermions, on the first five rows, are all spin-0. Higgsinos and gauginos are all spin-1/2. Three sets of each fermion's supermultiplet exist, one for each generation. [25]	23
Table 2	Simulated background event samples used in this analysis with the corresponding matrix element and parton shower generators, cross-section order in α_s used to normalize the event yield, underlying-event tune and PDF set.	88
Table 3	Summary of the electron selection criteria. z_0 gives the track's distance from the hard-scatter vertex projected in the z direction, while d_0 gives the this distance projected onto the $x - y$ plane.	90
Table 4	Summary of the muon selection criteria. The signal selection requirements are applied on top of the baseline selection.	90
Table 5	Summary of the jet and b -jet selection criteria. The signal selection requirements are applied on top of the baseline requirements.	91
Table 6	Summary of the photon selection criteria.	92
Table 7	Overview of all signal, control and validation regions used in the on-shell Z search. More details are given in the text. The flavour combination of the dilepton pair is denoted as either <i>SF</i> for same-flavor or <i>DF</i> for different flavor. All regions require at least two leptons, unless otherwise indicated. In the case of CR γ , VR-WZ, VR-ZZ, and VR-3L the number of leptons, rather than a specific flavor configuration, is indicated. The main requirements that distinguish the control and validation regions from the signal region are indicated in bold. Most of the kinematic quantities used to define these regions are discussed in the text. The quantity $m_T(\ell_3, E_T^{\text{miss}})$ indicates the transverse mass formed by the E_T^{miss} and the lepton which is not assigned to either of the Z -decay leptons.	95
Table 8	Lepton trigger requirements used for the analysis in different regions of lepton- p_T phase space.	96

Table 9	Yields in signal and validation regions for the flavor symmetric background. Errors include statistical and systematic uncertainties, discussed in Chapter 11.	103
Table 10	Background fit results from the sideband fit method. The $t\bar{t}$ MC’s normalization is taken as a free parameter in the fit to data in CRT, then that normalization factor is applied in SRZ. The results are shown here both divided between the ee and $\mu\mu$ channels and summed together. All other backgrounds are taken from MC in CRT, while in SRZ, the $Z/\gamma^* + \text{jets}$ contribution is taken from the $\gamma + \text{jets}$ method. The uncertainties quoted include both statistical and systematic components.	104
Table 11	Summary of the $t\bar{t}$ normalization factors calculated by the sideband fit to CRT and VRT for the 2015+2016 data.	105
Table 12	Comparison of FS background predictions from the nominal method, the flavor symmetry method, and the cross-check, the sideband fit method. Uncertainties include statistical and systematic uncertainties in both cases.	106
Table 13	List of triggers used to collect photon events in 2015 and 2016 data-taking.	108
Table 14	Control regions used to measure efficiencies of real and fake leptons. The flavour combination of the dilepton pair is denoted as either “SF” for same-flavour or “DF” for different flavour. The charge combination of the leading lepton pairs are given as “SS” for same-sign or “OS” for opposite-sign.	119
Table 15	Yields in validation regions. In VRS, data-driven background estimates are used for $Z/\gamma^* + \text{jets}$, fakes, and FS processes. All other backgrounds are taken from MC, including all backgrounds in the multi-lepton VRs. Uncertainties include statistical and systematic components.	121

Table 16	Uncertainties in the on-Z signal and validation regions. Nominal predictions are given with statistical uncertainty (including uncertainty from subtracted backgrounds), MC Closure uncertainty, uncertainty on the prediction from varying k and α by their statistical uncertainties, comparing the efficiencies from AODs to that of DAODs, and on the $m_{\ell\ell}$ widening, which includes MC statistics and a data/MC comparison in a loosened region.	125
Table 17	Uncertainty breakdown for the $\gamma + \text{jets}$ method in SRZ. Uncertainties considered are the impact of MC uncertainty on $V\gamma$ backgrounds, MC closure, uncertainty on $m_{\ell\ell}$ shape (also determined via MC closure), reweighting uncertainties, smearing uncertainties, and statistical uncertainty on the $\gamma + \text{jets}$ events used in the method.	133
Table 18	Systematic uncertainties on the fake-lepton background for on-Z regions for 2015+2016 yields. The nominal yield includes statistical uncertainty from the baseline selection in a given region. The following rows indicate the results of varying the real and fake lepton efficiencies up and down by their statistical uncertainty. Real cont. gives an uncertainty on the contamination of real leptons in the fake lepton efficiency. b -jet and no b -jet indicate the impact of requiring or vetoing b -tagged jets in the regions used to measure the fake efficiency.	134
Table 19	Fractional uncertainties of dibosons in signal and validation regions from Sherpa scale variations.	135
Table 20	Comparison of yields in on-Z and off-Z regions in Sherpa and Powheg diboson MC at 14.7 fb^{-1}	135
Table 21	Overview of the dominant sources of systematic uncertainty on the total background estimate in the signal regions. The values shown are relative to the total background estimate, shown in %.	136

Table 22 Number of events expected and observed in the ee , $\mu\mu$, and combined channels. Expected predictions include all systematic and statistical uncertainties discussed in Chapter 11. Also shown is the discovery p -value for zero signal strength ($p(s = 0)$) [105], Gaussian significance, 95% CL observed and expected upper limits on the number of signal events (S^{95}), and the corresponding observed upper limit on the visible cross section ($\langle\epsilon\sigma\rangle_{\text{obs}}^{95}$). 138

ACRONYMS

IBL Insertable B-Layer
MS Muon Spectrometer
ID Inner Detector
SCT Silicon Microstrip Tracker
TRT Transition Radiation Tracker
NN Neural Network
CCA Connected Component Analysis
ToT Time Over Threshold
MDT Monitored Drift Tube
CSC Cathode-Strip Chamber
RPC Resistive Plate Chamber
TGC Thin Gap Chamber
L₁ Level One
HLT High Level Trigger
L₁Calo L₁ Calorimeter Trigger
L₁Topo L₁ Topological Trigger
CTP Central Trigger Processor
TTC Trigger Timing and Control
ROB Read Out Board
RoI Region of Interest
LHC Large Hadron Collider
LEP Large Electron-Positron
SPS Super Proton Synchrotron
ATLAS A Toroidal LHC Apparatus
CMS Compact Muon Solenoid
ALICE A Large Ion Collider Experiment

LHCb Large Hadron Collider beauty
RF Radiofrequency
PSB Proton Synchrotron Booster
PS Proton Synchrotron
OR Overlap Removal
EM Electromagnetic
LCW Local Cluster Weighting
JES Jet Energy Scale
JER Jet Energy Resolution
JVT Jet Vertex Tagger
JVF Jet Vertex Fraction
CST Calorimeter Soft Term
TST Track Soft Term
MC Monte Carlo simulation
SM Standard Model
BSM Beyond the Standard Model
SUSY Supersymmetry
QCD Quantum Chromodynamics
PDF Parton Distribution Function
DM Dark Matter
LO Leading Order
NLO Next to Leading Order
NLO+NLL Next-to-Leading-Logarithmic Accuracy
SUSY Supersymmetry
MSSM Minimal Supersymmetric Standard Model
LSP Lightest Supersymmetric Particle
AOD Analysis Object Data
dAOD derived AOD
SR Signal Region

VR Validation Region

CR Control Region

FS Flavor Symmetric

CL Confidence Level

HL-LHC High Luminosity Large Hadron Collider

Part I
INTRODUCTION

INTRODUCTION

In 2010, the Large Hadron Collider ([LHC](#)) began colliding protons in its 27 km ring, taking its place as the most powerful in a long line of accelerators aimed at uncovering the fundamental rules that govern particle physics. Its primary goal was to complete the Standard Model of particle physics by discovering the Higgs boson, the last remaining particle that physicists felt sure must exist. With its presence, the Standard Model would be consistent, explaining every observed interaction of known particles, with a complete mathematic framework to describe each feature. However, even with a Higgs boson, the Standard Model contained hints that it might be incomplete, suspicious features that suggested that at a higher energy, there might be something more.

In 2012, the [ATLAS](#) (A Toroidal LHC Apparatus) and [CMS](#) (Compact Muon Solenoid) Experiments discovered the Higgs boson, leaving the [LHC](#) physics community without a single primary goal, but rather a host of theories to explore, each extending the Standard Model in a different way. Each theory attempts to solve one of the mysteries left by the Standard Model, providing an explanation for Dark Matter, suggesting a mechanism that could explain gravity's weakness, or explaining the Higgs boson's mass. For decades, the most popular of these has been Supersymmetry, which proposes a fermionic symmetry and requires a menagerie of new Supersymmetric particles, none of which has yet been observed.

Supersymmetry simultaneously solves more of the Standard Model's problems than any other theory, making it appealing to theorists and experimentalists alike. But in order to do this, Supersymmetric particles must appear with masses of approximately 1 TeV, precisely the range of energies the [LHC](#) is capable of exploring. In 2015, after a three-year shutdown, the [LHC](#) nearly doubled the energy of its collisions. This opened up new territory to be explored by analyzers, and provided data that could be used to either discover or exclude many Supersymmetric models.

The analysis presented in this thesis searches for Supersymmetry, seeking to identify events in which Supersymmetric particles are produced in proton-proton collisions, then decay via a Z boson to a chargeless Supersymmetric particle which escapes [ATLAS](#) without detection. A similar [ATLAS](#) search, performed with data from the lower-energy collisions 2012, observed a 3σ excess of events over the expected Standard Model background [1].

The excess generated a great deal of interest in this channel, and re-investigating it became a top priority when the upgraded LHC turned on in 2015. A preliminary search, performed using 2015 data only, was released at the end of that year. Again an excess was observed, this time with a significance of 2.2σ [2].

This thesis describes a search for Supersymmetry performed in this channel using data taken by the ATLAS detector in 2015 and 2016, including an explanation of the theory and motivation behind the search, and a description of the LHC and the ATLAS detector. The remaining chapters are laid out as follows:

CHAPTER 2 outlines the Standard Model of Particle Physics and the benefits of extending it to include Supersymmetry, then continues on to introduce the specific models used in the search presented in later chapters. This chapter also discusses the phenomenology of accelerator physics.

CHAPTER 3 describes the LHC and its operation, including the magnet system, the preaccelerator complex, and some further details of the phenomenology of collisions at 13 TeV.

CHAPTER 4 contains descriptions of the many pieces of the ATLAS detector, and how they serve to detect particles coming from LHC collisions. ATLAS's magnet and trigger systems are also discussed. This chapter also provides an overview of the process of generating MC for use in the ATLAS experiment.

CHAPTER 5 discusses the tracking system in the ATLAS detector, which forms patterns corresponding to particle trajectories out of the many energy deposits left in the detector. This chapter focuses on a neural network designed to improve tracking in the ATLAS Pixel Detector, and describes the benefits of its implementation.

CHAPTER 6 details the rest of the process of reconstruction, the procedure by which the signals in the ATLAS detector are interpreted as particles to be used for analysis.

CHAPTER 7 lists the main backgrounds for the Supersymmetry search described in this thesis, and provides general ideas of how they can be reduced.

CHAPTER 8 outlines how objects are identified and selected for this analysis, referencing many of the working points defined in Chapter 6.

CHAPTER 9 explains the analysis’s search strategy, defining signal, control, and validation regions, and briefly describing how each is used in the search.

CHAPTER 10 describes, for each of the backgrounds described in Chapter 7, how estimates of the Standard Model contributions to the signal region are performed.

CHAPTER 11 builds off of Chapter 11, and continues to detail how the uncertainties on each estimate are assessed.

CHAPTER 12 shows the results of the analysis, comparing expectations based on background estimates to the observed data.

CHAPTER 13 provides interpretations of the results, and explains the statistical procedure used to define exclusions on Supersymmetric models.

CHAPTER 14 concludes with a summary of the results, and an outlook for future searches.

Part II

THEORY AND MOTIVATION

This section describes the theoretical foundation for the analysis presented in [Part iv](#). It includes an overview of the Standard Model, including its phenomenology in a pp collider. The theory of Supersymmetry is explained, and the motivation for extending the Standard Model to include it is presented. In addition, this section includes details about the specific form of Supersymmetry searched for in this analysis.

THEORY AND MOTIVATION

The Standard Model ([SM](#)) of particle physics represents all particles and interactions currently known. It is formulated using the principles of Quantum Field Theory, with the constraints of several symmetries and physical requirements to determine the rules for allowed interactions [3]. Developed in the 1960s and 70s [4–6], it has been immensely successful at predicting the existence of particles before their discovery, and has held up to many high-precision tests. Despite this success, it has several shortcomings. Though the [SM](#) is likely correct at the energies thus far probed, it may be missing key components that become more important at higher energies. Models supplementing the [SM](#) with additional particles and interactions are referred to as Beyond the Standard Model ([BSM](#)) theories.

One possible extension of the [SM](#) is Supersymmetry ([SUSY](#)), a theory which postulates an additional symmetry between bosons and fermions to the [SM](#), creating a spectrum of [SUSY](#) particles (sparticles) which interact with the particles of the [SM](#). This theory motivates the search performed in [Part iv](#) of this thesis, and its theoretical appeals are discussed in this section, along with specific models considered in the search.

2.1 THE STANDARD MODEL

The [SM](#) of particle physics describes the interactions of all of the particles that have thus far been observed. It consists of matter particles and force carriers, as well as the Higgs boson, which fits into neither category. This model has been extraordinarily successful in predicting new particles and phenomena, including the prediction of the Higgs boson almost 50 years before its discovery in 2012, which completed the [SM](#).

The particles of the [SM](#) are divided into two categories: fermions and bosons [3]. The fermions comprise all the matter described by the [SM](#), and are spin- $\frac{1}{2}$ particles. The bosons are integer spin-particles, most of which are spin-1. These particles provide a mechanism to explain three of the four forces known to physics, with gravity still lacking a quantum formulation. The Higgs boson, the only spin-0 particle in the [SM](#), provides a mechanism for giving mass to the other particles. The full [SM](#), with the addition of the hypothetical graviton, is presented in [Figure 1](#).

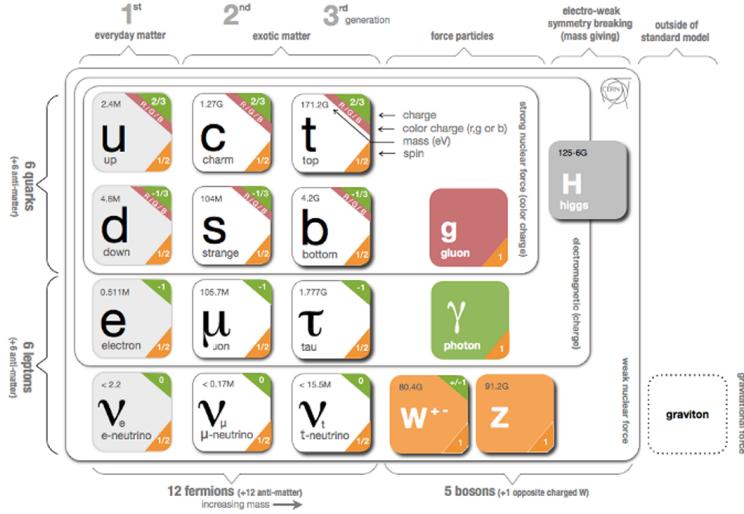


Figure 1: The Standard Model of particle physics, containing all known bosons and fermions, with the addition of the hypothetical graviton. [7]

2.1.1 Matter

The matter described by the SM is made up of fermions, spin- $\frac{1}{2}$ particles which can be broken into two groups, quarks and leptons. The leptons all interact weakly, while the quarks additionally interact strongly. Half the leptons as well as all quarks are electromagnetically charged.

2.1.1.1 Leptons

Leptons, as seen in the bottom left of Figure 1, exist in three generations, each labeled by a flavor: electron, muon, and tau. In the case of the massive leptons, these flavors are mass eigenstates, and the generations are placed in an order based on increasing mass. Each massive lepton is negatively electromagnetically charged and has a positively charged anti-particle.

The three neutrinos exist in the same flavors as the massive leptons, but these flavor eigenstates do not correspond exactly to mass eigenstates [8]. As a consequence, neutrinos oscillate between flavors as they propagate through space. These oscillations are the only evidence of neutrino mass, which is bound from below by the mass splittings determined from the oscillation and bound from above by cosmological limits on the universe's mass density [9]. Though it is still uncertain if the masses of the neutrinos follow the same hierarchy as the massive leptons, that expected ordering is slightly experimentally preferred over the inverted hierarchy [10].

Unlike the massive leptons, the neutrinos are not electromagnetically charged, and it is not yet known whether each neutrino has

a separate anti-particle, or if it is its own antiparticle. Because they are not electromagnetically charged, they can only interact weakly, making them extremely difficult to detect. As a consequence of their ability to evade detection, neutrinos' properties are nearly impossible to study with general purpose particle detectors.

The SM conserves lepton number, L , which is defined as the number of leptons minus the number of anti-leptons in a state, and can also be defined for each lepton flavor. Though there are anomalies that appear in second order SM interactions which could provide very small violations of this conservation, it holds to great precision in experiment. $\mu \rightarrow e\gamma$ branching ratios, for example, have been constrained to 10^{-13} [11]. As a consequence of this conservation, the lightest massive lepton, the electron, is stable.

2.1.1.2 Quarks

Quarks, as seen in the top left of [Figure 1](#), are also electromagnetically charged particles that interact weakly, but are differentiated from the leptons by their strong interactions. They are also organized in three generations ordered by mass, and come in pairs of *up*-type and *down*-type quarks, named after the lightest generation. In total, there are six flavors of quarks, up, down, charm, strange, top, and bottom. Though the up quark is lighter than the down, that rule is reversed in the subsequent two generations. Up-type quarks are electromagnetically charged $+\frac{2}{3}$, while the down-type quarks are charged $-\frac{1}{3}$. Quarks are also charged under the strong interaction, whose three charges are characterized by colors: red, green, and blue. Each quark has an anti-particle with the opposite charges.

These fractional charges and individual colors are never seen in nature because of the requirement (discussed further in [Section 2.1.2.2](#)) that stable particle states be color-neutral. To accomplish this, quarks can create two-particle bound states called *mesons* consisting of one quark and one anti-quark with the same color charge, or three-particle bound states of quarks or anti-quarks with the three different color charges, which are called *baryons*. The lightest color neutral state containing only quarks, the proton (uud), is stable. Extremely unstable bound states consisting of higher numbers of quarks can also exist, such as the pentaquark discovered in 2015 at the LHC [12]. Collectively, these multi-quark bound states are called *hadrons*.

Besides the proton, all hadrons are unstable, but their lifetimes have a wide range. Neutrons, for example, have a lifetime of nearly fifteen minutes, and are stable enough to be involved in the formation of atoms. Charged pions ($u\bar{u}$) and kaons ($u\bar{s}$) have lifetimes on the order of 10 ns, which allows them to propagate several meters when traveling close to the speed of light. Most other hadrons decay effectively instantaneously when produced in a collision, with lifetimes much too short to be resolved by a particle detector. B mesons sit at

the boundary between these two regimes, with lifetimes of approximately 1.5 ps, allowing them to propagate up to a few mm before decaying.

Like leptons, the number of quarks in a state is conserved, up to very small anomalies. However, because quarks cannot exist in an isolated state, that conservation is described in terms of baryon number (B). Baryons are defined with $B = 1$, while anti-baryons have the quantum number $B = -1$. Mesons have $B = 0$.

2.1.2 Forces

The fermions in the previous section interact via the electromagnetic, weak, and strong forces. In a perturbative quantum field theory, interactions via these forces are represented by mediating bosons. These force carriers interact only with particles charged with their corresponding force's quantum numbers. The photon, for example, interacts only with electromagnetically charged particles. Gluons, mediators of the strong force, interact only with color charged particles, quarks and themselves. All fermions are weakly charged and interact with the weak force's mediators, the W and Z bosons.

The formulation for each of these forces is developed by requiring that the [SM](#) Lagrangian be locally gauge invariant [8]. This can be accomplished by adding gauge fields to the Lagrangian, whose behavior under gauge transformations cancels out the gauge dependence of the free Lagrangian. However, adding a mass term for these fields reintroduces gauge dependence, so this mechanism only creates forces mediated by massless gauge bosons. The addition of the Higgs field provides mass terms for the weak gauge bosons (as well as other particles) without interfering with the gauge invariance.

The total gauged symmetry group for the [SM](#) is $SU_C(3) \times SU_L(2) \times U_Y(1)$, where C stands for color, the charge of the strong force, L stands for left, because the weak force is left-handed, and Y is the *hypercharge* quantum number, the charge of the unified electroweak force.

2.1.2.1 The Electromagnetic Force

Electromagnetism provides the simplest example of a requirement of local gauge invariance generating a Lagrangian description of a force. Electromagnetism has one massless mediator, the photon, which interacts with all electromagnetically charged particles. What follows is a brief description of how enforcing this invariance generates a Lagrangian of the same form as the classical electromagnetic Lagrangian, which can be incorporated into the [SM](#).

The particles in [Section 2.1.1](#) are fermions, and so the Lagrangians describing their free propagation are Dirac Lagrangians and all follow the form

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi. \quad (1)$$

Requiring that the free Lagrangians for these particles be invariant under a $U(1)$ local gauge transformation, $e^{iq\lambda(x)}$, can be accomplished by adding a term to the Lagrangian which cancels the derivative term arising from λ 's dependence on x :

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi - (q\bar{\psi}\gamma^\mu\psi)A_\mu \quad (2)$$

where A_μ is a “gauge field” that transforms according to

$$A_\mu \rightarrow A_\mu + \partial_\mu\lambda. \quad (3)$$

This vector field must also come with a free term,

$$\mathcal{L} = -\frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu} + \frac{1}{8\pi}m_A^2 A^\nu A_\nu. \quad (4)$$

The mass term for this field would not itself be invariant under the transformation, but the field can simply be made massless to avoid this problem. The final Lagrangian, then, is

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi - \frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu} - (q\bar{\psi}\gamma^\mu\psi)A_\mu \quad (5)$$

which is precisely the original Lagrangian with the addition of terms replicating the form of the Maxwell Lagrangian. In a quantized interpretation, it describes a field that interacts with particles with non-zero electromagnetic charge q via interactions with a massless spin-1 boson, the photon. This charge is dependent on the energy scale of the interaction, μ , and the strength of the interaction is more typically described by the electromagnetic coupling constant

$$\alpha_{\text{EM}}(\mu) = q(\mu)^2/4\pi. \quad (6)$$

For the purpose of succinct notation, this Lagrangian is often rewritten in terms of the *covariant derivative*

$$D_\mu = \partial_\mu + iq\lambda A_\mu \quad (7)$$

which immediately cancels the gauge dependent term created by the transformation. This mechanism is mathematically simple in the $U(1)$ case, but can be replicated for more complicated gauge transformations with perturbative approximations.

2.1.2.2 The Strong Force

The strong force is generated by a similar process of requiring local gauge invariance, but in this case, for an $SU(3)$ transformation. The interactions of the strong force are described by the theory of Quantum Chromodynamics (QCD), which is given by the Lagrangian

$$\mathcal{L}_{\text{strong}} = -\frac{1}{4}G_{\mu\nu}^{\alpha}G^{\alpha\mu\nu} - \frac{1}{2}\bar{Q}_m D Q_m \quad (8)$$

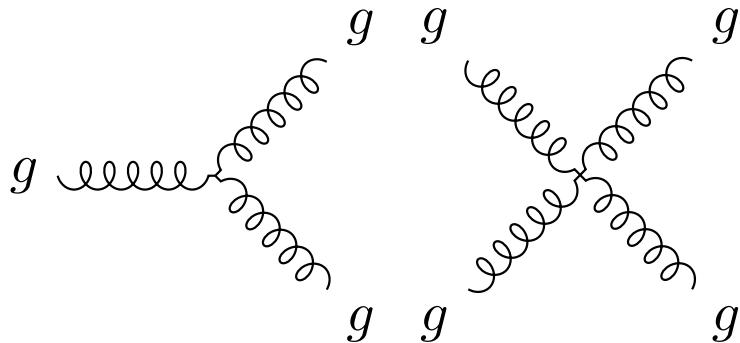
where the α index runs from 1 to 8 and represents the eight force carriers of the strong force, gluons with different color charges. m indexes the three quark generations, and $G_{\mu\nu}^{\alpha}$ is the field strength tensor, defined as

$$G_{\mu\nu}^{\alpha} = \partial_{\mu}G_{\nu}^{\alpha} - \partial_{\nu}G_{\mu}^{\alpha} + g_s f_{\beta\gamma}^{\alpha} G_{\mu}^{\beta} G_{\nu}^{\gamma} \quad (9)$$

where g_s is a function of the energy scale of the interaction μ , and is related to the strong coupling constant by

$$\alpha_s(\mu) = g_s(\mu)^2 / 4\pi. \quad (10)$$

The first term of the Lagrangian gives the gluon self-coupling interactions, with terms involving 2, 3, and 4 gluon field terms. The 2-field portion simply describes gluon propagation, but the other terms give gluon self-interaction terms that can be described by the Feynman diagrams in [Figure 2](#). Unlike photons, gluons are charged by the force they carry, making self-interaction possible.



[Figure 2](#): Gluon self coupling Feynman diagrams involving 3- and 4-gluon interactions.

In the second term, $D Q_m$ is the covariant derivative acting on the quark field. The quarks are charged under all three forces, strong, electromagnetic, and weak, so the covariant derivative includes terms to make each of the force's Lagrangians gauge invariant. Thus this term introduces quark-boson interactions of four types, seen in [Figure 3](#). The quarks' coupling to the gluon is the strongest, and the

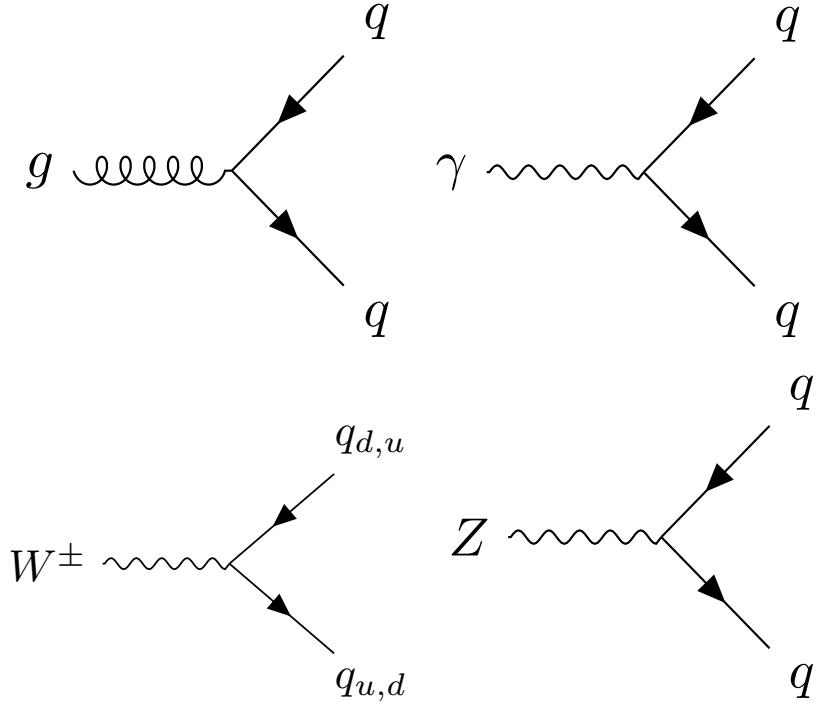


Figure 3: Quark couplings to the different types of gauge bosons. The $q_{u,d}$ labels represent any up- or down-type quarks.

other, weaker couplings occur at lower rates. The couplings to the W and Z bosons are described in [Section 2.1.2.3](#).

Quantum Field Theory assumes that particles are essentially *free*, propagating without interaction, and considers all interactions as perturbations on a free theory. So long as multiple interactions are much less likely than a single interaction, or put another way, so long as the coupling constants for each force are much less than one, this perturbative approximation is essentially correct. However, for the strong coupling constant, α_s , this assumption is not always valid. α_s changes as a function of the energy of an interaction according to its renormalization group equation

$$\mu_R^2 \frac{d\alpha_s}{d\mu_R^2} = \beta(\alpha_s) = -(b_0\alpha_s^2 + b_1\alpha_s^3 + b_2\alpha_s^4 + \dots) \quad (11)$$

where μ_R^2 gives the renormalization scale, and each b_n gives a correction to the β -function based on diagrams with n loops [[13](#)]. The overall negative sign produces the unique energy dependence of α_s , which becomes very small at high energy scales and asymptotically increases at low energies. [Figure 4](#) shows this effect translated to distance scales, demonstrating that $\alpha_s \ll 1$ and can be considered perturbatively at small distance scales, but at large distance scales α_s approaches 1, and a perturbative approximation can no longer be used. Instead, for distances larger than 10^{-16} , the colorless hadrons introduced in [Section 2.1.1.2](#) must be used to describe strong interactions.

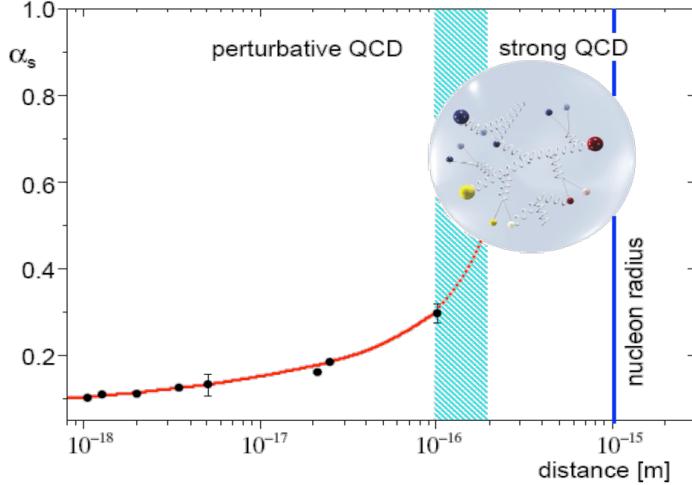


Figure 4: The running of the strong coupling constant, α_s . [14]

The boundary between these regimes is referred to as Λ_{QCD} and differentiates energies at which QCD can be considered perturbatively and those at which it cannot. The LHC is capable of producing individual high-energy quarks in its hard-scattering interactions, but they lose energy as they radiate gluons, eventually entering the energy regime below Λ_{QCD} . The transition between these two regimes is complex, and dictates the way that strongly charged particles appear in the ATLAS detector. This is described in more detail in Section 2.1.3.1.

2.1.2.3 The Electroweak Force

A similar gauging process, using an $SU(2)$ transformation, can produce a Lagrangian that would suffice to describe the W and Z bosons of the SM, if only they were massless. However, they are not, so an alternate mechanism is needed to generate massive force carriers.

Before a mechanism for their masses was understood, and before they were discovered, the large masses of the W and Z bosons were proposed in order to unify the electromagnetic and weak forces into the electroweak force [8]. The large masses were crucial to explain the discrepancy in the strength of the two forces.

The unified electroweak force is generated by a symmetry group written as $SU(2)_L \times U(1)_Y$, where L refers to left-handed fields, and Y is the quantum number for *hypercharge*. This new quantum number is defined as

$$Y = 2(Q - T_3) \tag{12}$$

where Q is the electromagnetic charge and T_3 is the third component of weak isospin \mathbf{T} , the quantum number relating to the weak interaction. In the unified theory, quark and lepton singlets interact

according to their hypercharge, and left-handed quarks and leptons, grouped according to their generation, interact as doublets.

The gauge bosons resulting from this unified theory include a triplet, W , with coupling g_W , and a singlet field B , with coupling $g'/2$. However, the electroweak symmetry is broken, and mixing between these states occurs. Rewritten in their mass basis, the standard electroweak force carriers are produced: W^\pm , two states with identical coupling resulting from the first two states of the W triplet, the Z and the photon field A resulting from the mixing of the last W state and B .

The electroweak Lagrangian is much more complicated than the strong Lagrangian, and can be divided into several terms:

$$\mathcal{L}_{\text{electroweak}} = \mathcal{L}_{\text{gauge}} + \mathcal{L}_{\text{fermions}} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Yukawa}}. \quad (13)$$

The first term can be written as follows

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{4}W_{\mu\nu}^a W_{\mu\nu}^a - \frac{1}{4}B_{\mu\nu} B_{\mu\nu} \quad (14)$$

where the a indices are numbered 1 through 3 and indicate the generators of $SU(2)$ which are written

$$W_{\mu\nu}^a = \partial_\mu W_\nu^a - \partial_\nu W_\mu^a + g_W \epsilon_{abc} W_\mu^b W_\nu^c. \quad (15)$$

The gauge portion of the Lagrangian then generates interaction terms of between the gauge fields, which when rewritten in terms of the mass-eigenstate basis, generates interactions between three gauge bosons, like the ones in Figure 5, as well as interactions between four gauge bosons.

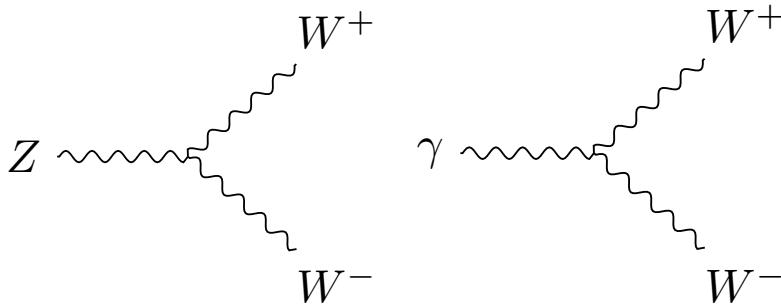


Figure 5: Feynman diagrams of trilinear gauge couplings in the SM.

The fermion portion of the Lagrangian is written as

$$\begin{aligned} \mathcal{L}_{\text{fermion}} = & -\frac{1}{2}\bar{L}_m \not{D} L_m - \frac{1}{2}\bar{Q}_m \not{D} Q_m \\ & -\frac{1}{2}\bar{U}_m \not{D} U_m - \frac{1}{2}\bar{D}_m \not{D} D_m \\ & -\frac{1}{2}\bar{E}_m \not{D} E_m \end{aligned} \quad (16)$$

where L is the left-handed lepton doublet, Q is the left-handed quark doublet, U is the right-handed singlet for up-type quarks, D is the

same for down-type quarks, and E is the right-handed singlet for electrons, muons and taus. Each of these fields has an implicit index running from 1 to 3 to represent the three generations. The covariant derivative in each term includes terms involving all the gauge fields the fermion is charged under.

Unlike the other forces, the weak force treats left- and right-handed fermion fields differently; it only interacts with the left-handed fields, so only the first two terms' covariant derivatives include W terms. The first term in this Lagrangian, for example, produces weak interactions depicted in Figure 6. The Z bosons, because they represent a mixing between the W and B fields, can interact with right-handed leptons and quarks, but these interactions are different for the two chiralities.

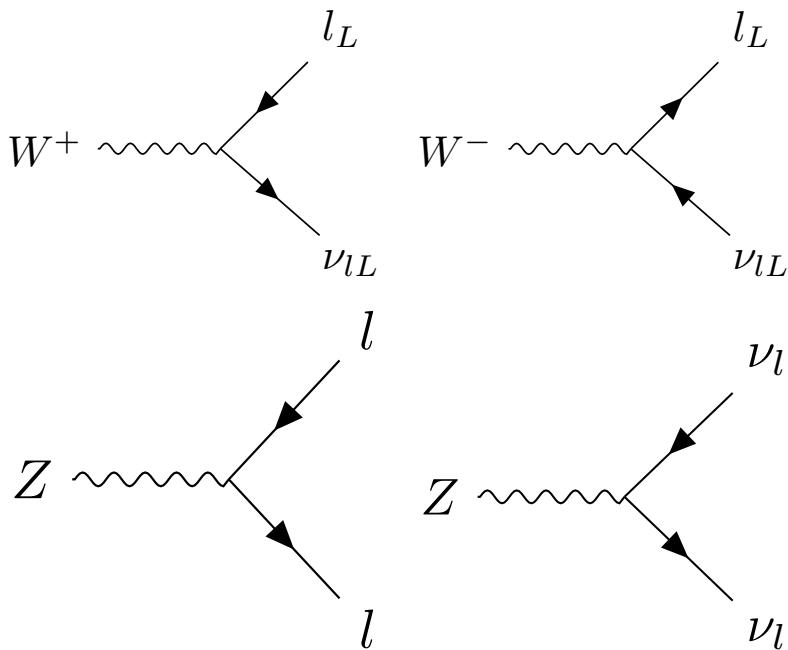


Figure 6: Feynman diagrams of weak couplings to leptons in the SM.

No right-handed term appears for the neutrino field, because only left-handed neutrinos and right-handed anti-neutrinos have been observed. However, because neutrinos have non-zero mass, their chirality can change with frame boosts, which complicates any claim that right-handed neutrinos do not exist [3]. It is possible that neutrinos are their own antiparticle, making the right-handed anti-neutrino the solution to this problem. It's also possible that very massive right-handed neutrinos do exist, and simply haven't been discovered yet.

The remaining two terms of the electroweak Lagrangian are related to the Higgs field, which is the source of electroweak symmetry breaking.

2.1.2.4 The Higgs Mechanism

The Higgs mechanism presents a way to generate a mass term for the electroweak gauge bosons. It is a scalar field, with a Lagrangian

$$\mathcal{L}_{\text{Higgs}} = \frac{1}{2}(\partial_\mu \phi)^*(\partial^\mu \phi) + \frac{1}{2}\mu^2\phi^*\phi - \frac{1}{4}\lambda^4(\phi^*\phi)^2 \quad (17)$$

where ϕ is a complex scalar, $\phi = \phi_1 + i\phi_2$. This looks very similar to a standard scalar field Lagrangian, but the signs on the mass and interaction terms are reversed, implying an imaginary mass term. However, this isn't a good interpretation of the Lagrangian, because it differs from all previously considered Lagrangians in one important way: its ground state does not occur at $\phi = 0$. Because quantum field theory is perturbative, it is valid only when expanded around a ground state, which, when calculated for this Higgs Lagrangian, must satisfy

$$\phi_1^2 + \phi_2^2 = \frac{\mu^2}{\lambda^4}. \quad (18)$$

The original Lagrangian can then be rewritten in terms of a field $v + H(x)$ centered around the ground state with energy called the *vacuum expectation value* defined as

$$v = \frac{\mu}{\lambda^2}. \quad (19)$$

This rewriting produces a Lagrangian with a non-imaginary mass. However, in an effect called *spontaneous symmetry breaking*, the original $SO(2)$ rotational symmetry of the Lagrangian is lost, resulting only in a $U(1)$ rotational symmetry; the Lagrangian is invariant under a phase transformation.

As in [Section 2.1.2.1](#), it is possible to make the Lagrangian invariant under a local $U(1)$ transformation, $\phi \rightarrow e^{i\theta(x)}\phi$ by adding a massless gauge field A^μ and using the covariant derivative. Due to the many cross terms from the non-zero ground state, terms for the mass of one of the scalar bosons as well as the gauge field appear, leaving only one massless scalar boson. This massless boson, it turns out, can be completely removed from the theory via local $U(1)$ transformations, ultimately producing a theory with one massive scalar (the Higgs) and a massive gauge field (W).

The Higgs interaction with the weak gauge bosons also creates couplings between the particles, which can be seen in [Figure 7](#). There are also Higgs self-interaction terms included in the Lagrangian, producing vertices describing 3- and 4-Higgs interactions.

The remaining piece of the Lagrangian, $\mathcal{L}_{\text{Yukawa}}$ describes the Higgs field's interactions with the fermions of the [SM](#), and can be written as

$$\mathcal{L}_{\text{Yukawa}} = -\Gamma_{mn}^e \bar{L}_m \phi E_n - \Gamma_{mn}^u \bar{Q}_m \phi U_n - \Gamma_{mn}^d \bar{Q}_m \phi D_n + h.c. \quad (20)$$

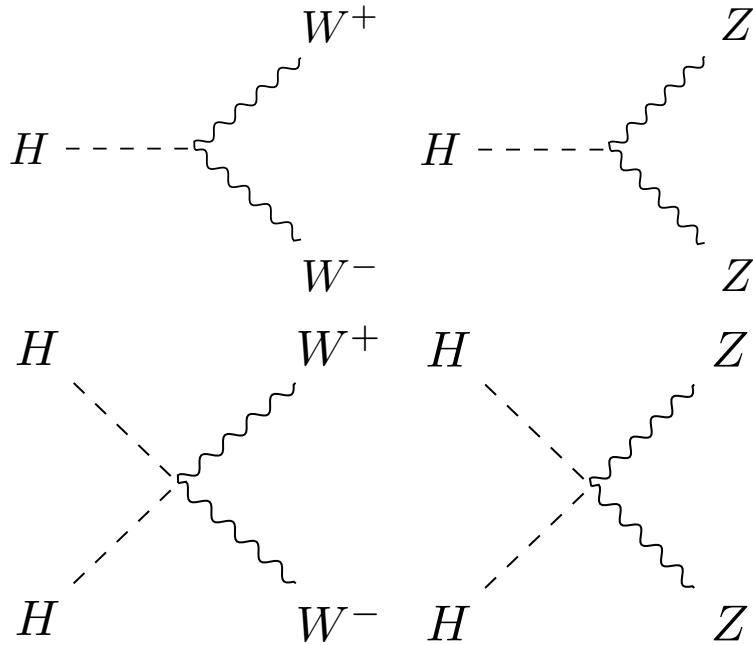


Figure 7: Feynman diagrams demonstrating Higgs couplings to the weak gauge bosons in the SM.

where *h.c.* is the hermitian conjugate term, and the Γ matrices are indexed by generation, and, when diagonalized, are proportional to the masses of the fermions. The Higgs field's vacuum expectation value produces terms that look like fermion mass terms. Additionally, terms that couple the fermions to the Higgs field are produced, with each fermion's coupling proportional to its mass, according to

$$g_f = \sqrt{2} \frac{m_f}{v} \quad (21)$$

where m_f is the mass of the fermion. Feynman diagrams for lepton and quark terms can be seen in Figure 8.

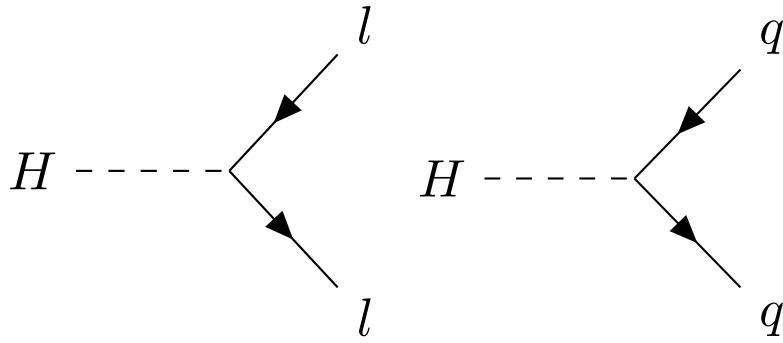


Figure 8: Feynman diagrams showing Higgs couplings to fermions in the SM.

2.1.3 Phenomenology of Proton-Proton Collisions

As discussed in [Chapter 3](#), the [LHC](#) collides bunches of high-energy protons, and the interactions of these protons' constituent quarks produce the wide array of particles seen in the [ATLAS](#) detector. The [LHC](#) typically cites its energy in terms of \sqrt{s} , the center of mass energy of protons in the two colliding beams, which in Run 2 is 13 TeV. However, because the proton is not fundamental, this energy is divided among many particles that make up the proton.

To first order, a proton consists of three quarks: two up quarks and one down quark, held together by gluons. However, a real quantum mechanical system is much more chaotic; in addition to these three quarks, called *valence* quarks, there are many others popping into and out of existence. These additional quarks are called *sea* quarks and can also carry fractions of the proton's energy.

The particles inside the proton can have a wide range of energies depending on the internal dynamics at the moment of the collision. These cannot be predicted exactly, but probabilistic models called Parton Distribution Functions ([PDFs](#)) describe the likelihood of any given configuration. These functions are determined using data from hard scattering experiments and give probabilistic estimates for how often a given type of particle appears with a fraction x of the total proton energy, as seen in [Figure 9](#).

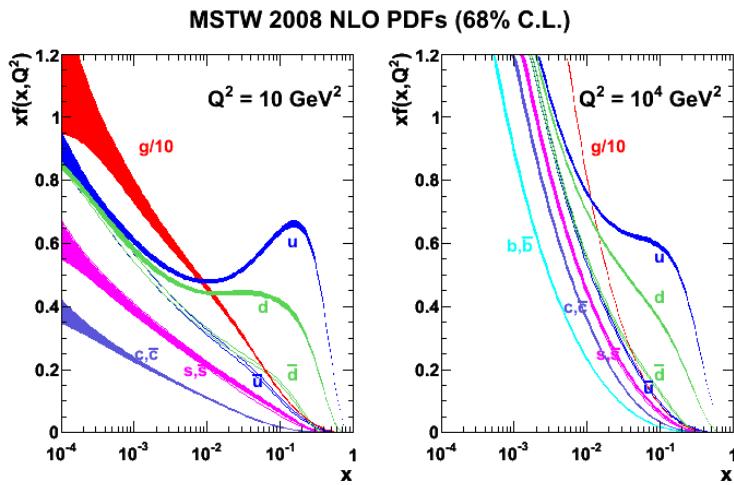


Figure 9: 2008 MSTW [PDFs](#) for various particle types given as a function of x and Q^2 , the square of the parton-parton momentum transfer. [[15](#)]

A classical proton model would suggest that each valence quark carries about one-third of the total proton energy. In practice, they each typically account for less than a third of the total energy, with the remaining energy divided among the many sea quarks and gluons. Thus, in a proton-proton collision, the initial particles rarely have

more than a third of the beam energy, and often have a much lower energy resulting from the interaction of sea quarks.

These PDFs are used at the LHC to calculate the probability for a given process to occur, or its *cross-section*. The cross-section for a process to occur with a two-proton initial state is given by

$$\sigma(P_1, P_2 \rightarrow X) = \sum_{i,j} \int_0^1 dx_1 dx_2 f_i(x_1, Q^2) f_j(x_2, Q^2) \hat{\sigma}(x_1 P_1, x_2 P_2, Q^2) \quad (22)$$

where i and j are indices representing the partons within proton 1 and 2 respectively, $x_{i,j}$ gives these partons' momentum fraction, and $f_{i,j}$ gives their PDFs. $\hat{\sigma}$ gives the cross-section for a process going from partons to X , also referred to as the *matrix element*, and Q^2 gives the energy scale of this hard scattering of partons.

Matrix elements, in an ideal world, would include every possible Feynman diagram that describe the initial partons producing a given final state. In practice, the calculation of these diagrams can become very complicated when more and more loops are allowed. The simplest calculations, which include diagrams without any loops, are referred to as Leading Order (LO), while calculations including diagrams with one loop are called Next to Leading Order (NLO), and additional Ns can be added to describe more complex calculations.

Figure 10 shows cross-sections for many SM processes as a function of energy, demonstrating how processes involving high-mass objects have increasingly large cross-sections as the energy of the collider increases. In the analysis discussed in Part iv, the most important SM backgrounds are from top quarks, Z bosons, and diboson processes (ZZ, WW, and WZ). Though Z boson production has the highest cross-section of those processes, it contributes the least to the background of the search because of the specific requirements made on events designed to reduce SM backgrounds. Details on these backgrounds and how they can appear as signal-like events are given in Chapter 7.

2.1.3.1 Production of Jets in the ATLAS Detector

For most particles, the output of the hard scattering process is similar to the objects observed by a detector; electrons, photons, and muons emerging from final states are all observed as they pass through sensitive materials. There are exceptions: particles can radiate energy as they travel through the detector and photons can convert into electron-positron pairs, but in most cases, these final state particles can be directly observed. For color-charged particles, however, this is not the case.

As discussed in Section 2.1.2.2, quarks and gluons are strongly charged, and the coupling constant of the strong force increases a-

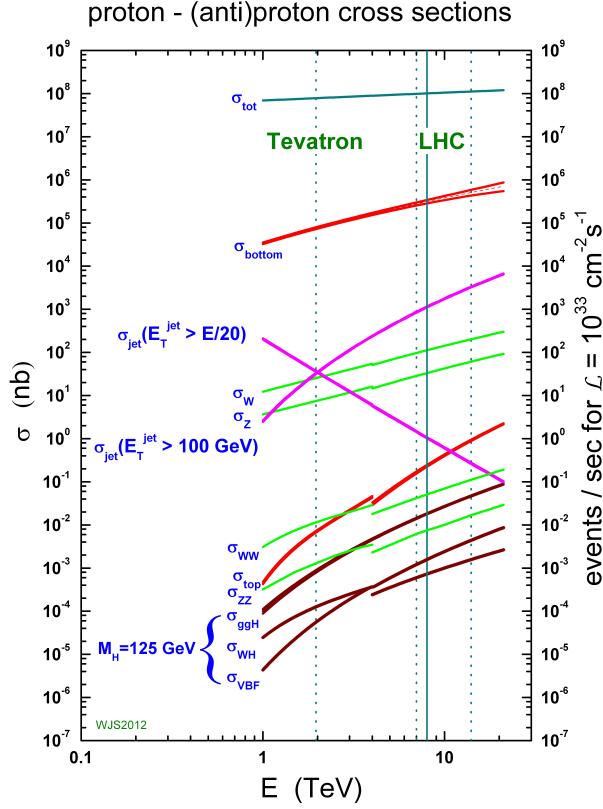


Figure 10: Cross-sections for many SM processes as a function of \sqrt{s} [16].

symptotically at large distance scales. As a consequence, colored particles undergo a process called *fragmentation*. Rather than continuing to propagate in a colored state, it is energetically preferable for high energy quarks and gluons produced in hard scatterings to radiate additional colored particles, which can then be used to form colorless bound states. First, a process called *parton showering* occurs, in which many new colored particles are radiated from the original, creating a conic spray of particles along the original particle's trajectory. This process repeats until the particles' energies reach the scale of hadronic masses, about 1 GeV. Next, these particles *hadronize*, forming composite particles that can propagate without interacting strongly.

The resulting spray of hadrons is called a *jet*, and it is jets, rather than individual quarks or gluons, that are observed by particle detectors. Grouping these particles into a jet is a non-trivial process, which is discussed in Section 6.4.

Jets don't always come directly from the hard scattering process; collisions in the LHC often contain additional jets resulting from *initial and final state radiation*. In these processes, either the initial partons radiate energy (typically in the form of a gluon) or the products of

the hard scattering process radiate additional colored particles before fragmentation occurs.

2.1.4 *Problems in the Standard Model*

Although the [SM](#) is a self-consistent theory that describes to great accuracy all of the particles and forces it includes, it does have certain shortcomings. The most glaring is the omission of gravity. Though the force is well understood at large scales via the theory of General Relativity, no satisfying quantum description of gravity has been accepted, much less proven. The Planck scale, the energy scale at which gravitational interactions become large enough that no sound theory can ignore gravity, is at about 10^{19} GeV, 16 orders of magnitude above the electroweak scale. Because of this large discrepancy of scale, the exclusion of gravity from the [SM](#) is unlikely to directly affect [LHC](#) physics.

Another clear omission of the [SM](#) is Dark Matter ([DM](#)), so named for its lack of electromagnetic interactions. This type of matter was first identified in 1933 through the observation of galactic rotation curves [17]. The speed of rotation indicated both that there was more mass in the system than could be accounted for by observations made directly of the galaxy, and that this additional matter was distributed in a halo, not a disk like the typical luminous matter. This effect can be seen in [Figure 11](#), which demonstrates that the observed galactic density as a function of radius does not match the expected density from the luminous galactic disk. Since then, evidence for [DM](#) has been observed in colliding clusters [18], measurements of the cosmic microwave background [19], and in many more rotational curves, but the particles that form [DM](#) have never been directly detected or seen at a particle accelerator. As a consequence, very few details are known about the nature of this matter, only its density throughout the universe and that it does not interact strongly or electromagnetically.

Beyond the omissions of gravity and [DM](#), there are several aesthetic problems with the [SM](#) - ones that could have no solution, but seem to suggest that the current [SM](#) is missing some pieces. The first is the sheer number of parameters in the [SM](#). There are 26 independent parameters determining the mass of the particles and all the couplings between them. Besides the rough grouping of fermions into generations, there seems to be no order to masses of particles, and no way to predict many of the masses or couplings.

In the past, large numbers of seemingly unrelated parameters have indicated that a theory has a more fundamental form at shorter distance scales. The large number of elements, it turned out, could be explained by different groupings of three particles, the proton, neutron, and electron. Later, the menagerie of hadrons became so large that a similar re-imagining of what was fundamental took place, and

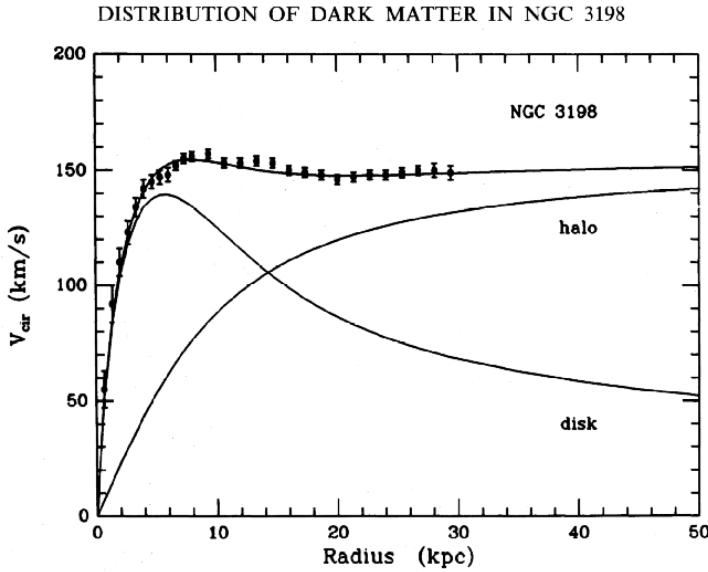


Figure 11: Galactic rotation curve of velocity as a function of radius in NGC 3198. Included is the observed data, as well as the expected velocity distribution from a disk-shaped galaxy corresponding to the expected density from electromagnetic observations. Another curve corresponding to a halo-shaped matter distribution is superimposed, and the halo and disk are summed and fit to the data. [20]

the theory of quarks gave an order to the many mesons and baryons [21]. This pattern leaves physicists suspicious of any theory with too many particles and free parameters, suggesting that perhaps, at a higher energy, there is a simpler model that can unify many of the seemingly disparate elements of the SM.

In addition, some of these seemingly independent parameters have suspicious symmetry. The Higgs mass, for example, has been measured to be 125.7 ± 0.4 GeV [13]. This mass is the sum of the bare mass, the one that appears in the Lagrangian, and quantum corrections from interactions with other particles, which are proportional to the square of energy at which these interactions occur. If no new physics exists at energies lower than the Planck scale (when the SM Lagrangian must change to account for gravity), these corrections could be up to 35 orders of magnitude larger than the Higgs mass. Though the bare mass could theoretically cancel out this massive correction, these parameters should be independent, and the odds that they would be precisely the same to 35 places are very, very small. This near-exact canceling is often called *fine-tuning*, an undesirable trait in a theory which suggests that some more fundamental symmetry has been missed. A *natural* solution, one free of this fine-tuning, is sought to resolve this SM problem.

Names		sparticles	particles	$SU(3)_C, SU(2)_L, U(1)_Y$
squarks, quarks	Q	$(\tilde{u}_L \tilde{d}_L)$	$(u_L d_L)$	$(\mathbf{3}, \mathbf{2}, \frac{1}{6})$
	\tilde{u}	\tilde{u}_R^*	u_R^\dagger	$(\bar{\mathbf{3}}, \mathbf{1}, -\frac{2}{3})$
	\tilde{d}	\tilde{d}_R^*	d_R^\dagger	$(\bar{\mathbf{3}}, \mathbf{1}, \frac{1}{3})$
sleptons, leptons	L	$(\tilde{\nu} \tilde{e}_L)$	(νe_L)	$(\mathbf{1}, \mathbf{2}, -\frac{1}{2})$
	\tilde{e}	\tilde{e}_R^*	e_R^\dagger	$(\mathbf{1}, \mathbf{1}, \mathbf{1})$
Higgs, higgsinos	H_u	$(\tilde{H}_u^+ \tilde{H}_u^0)$	$(H_u^+ H_u^0)$	$(\mathbf{1}, \mathbf{2}, \frac{1}{2})$
	H_d	$(\tilde{H}_d^0 \tilde{H}_d^-)$	$(H_d^0 H_d^-)$	$(\mathbf{1}, \mathbf{2}, -\frac{1}{2})$
gluino, gluon	\tilde{g}	g		$(\mathbf{8}, \mathbf{1}, \mathbf{0})$
winos, W bosons	$\tilde{W}^\pm \tilde{W}^0$	$W^\pm W^0$		$(\mathbf{1}, \mathbf{3}, \mathbf{0})$
bino, B boson	\tilde{B}^0	B^0		$(\mathbf{1}, \mathbf{1}, \mathbf{0})$

Table 1: Supermultiplets of supersymmetric and SM particles. Sfermions, on the first five rows, are all spin-0. Higgsinos and gauginos are all spin-1/2. Three sets of each fermion's supermultiplet exist, one for each generation. [25]

2.2 SUPERSYMMETRY

Supersymmetry (SUSY) was proposed and developed in the 1970s to give solutions to many of these SM shortcomings [22–24]. The theory works by introducing a fermionic symmetry to the SM, in addition to the usual spacetime symmetries of translations, rotations, and changes of Lorentz frame. The combination of the usual spacetime with this fermionic dimension is called a *superspace*. Rotations in this dimension result in a particle's spin changing by 1/2, turning a spin-1/2 fermion into a spin-0 particle, for example. As a consequence, this symmetry requires the existence of many new particles - a bosonic *sfermion* for each fermion of the SM and a fermionic *gaugino* for each of the gauge bosons. These superpartners of SM particles should have identical quantum numbers to the original particle, except for their spins. Table 1 shows the SM particles and their superpartners.

If the theory is symmetric under these fermionic rotations, these particle-sparticle pairs can be described by a single *superfield*, which simultaneously describes the behavior of both SM and SUSY particles in the superspace. However, this completely symmetric behavior is untenable given basic observations of matter in the universe. For example, if there were a *selectron* (the superpartner of the electron, \tilde{e}), with identical mass to the electron, it would have been detected long ago. In fact, such a particle would fundamentally change atomic structure, with the bosonic selectrons capable of piling into the ground state of an atom, and removing all the interesting valence-shell interactions of electrons that determine molecular structure. Thus, if SUSY

does exist, the symmetry must be broken, so that the superpartners have much higher masses than the original **SM** particles.

2.2.1 The Minimal Supersymmetric Standard Model

The Minimal Supersymmetric Standard Model (**MSSM**) was designed to be the simplest supersymmetric extension of the **SM** that remains self consistent, and it results in the particles seen in [Table 1](#) [25]. The formulation of the **MSSM** begins by introducing a second Higgs doublet to account for the different masses of the sparticles. As with the **SM** Higgs, electroweak symmetry breaking results in the loss of degrees of freedom, and only five of the original eight states remain, the lightest of which, h^0 , can be interpreted as the **SM** Higgs already discovered. There are two remaining neutral states, A^0 and H^0 , as well as two charged Higgses, H^\pm .

The neutral Higgs states mix with the neutral gauge bosons, while the charged Higgs states mix with the charged gauge bosons, producing a series of states labeled only by their charge and the order of their masses. The neutral states, collectively called the neutralinos, are identified from lightest to heaviest, $\tilde{\chi}_1^0, \tilde{\chi}_2^0, \tilde{\chi}_3^0$, and $\tilde{\chi}_4^0$. The charged states, referred to as charginos, are similarly called $\tilde{\chi}_1^\pm$ and $\tilde{\chi}_2^\pm$.

The **MSSM** introduces many new interactions between **SM** particles and sparticles. Though these don't represent all possible interactions, a general rule is that any **SM** vertex can have two interacting particles replaced with their sparticle equivalents, and this vertex will be part of the **MSSM**. [Figure 12](#) gives two examples of such vertices.

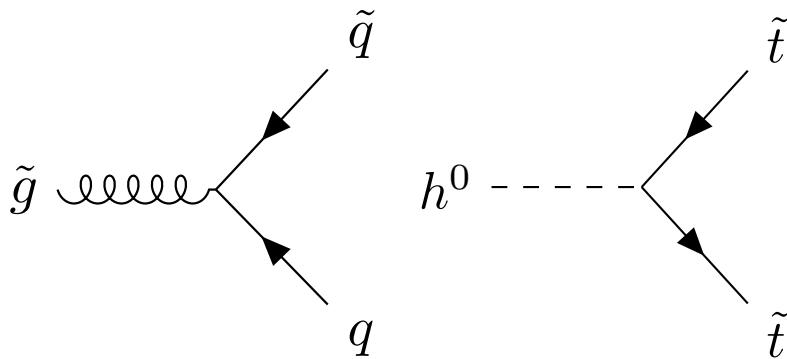


Figure 12: Two example vertices allowed by the **MSSM**.

In addition to these interactions, there are several terms that appear in the **MSSM** Lagrangian that violate the B and L conservation observed in the **SM**. In fact, these terms violate $B - L$, which, unlike B

and L conservation individually, does not have even small violations in the [SM](#). These superpotential terms appear as follows

$$W_{\Delta L=1} = \frac{1}{2} \lambda^{ijk} L_i L_j \bar{e}_k + \lambda'^{ijk} L_i Q_j \bar{d}_k + \mu^i L_i H_u \quad (23)$$

$$W_{\Delta B=1} = \frac{1}{2} \lambda''^{ijk} \bar{u}_i \bar{d}_j \bar{d}_k. \quad (24)$$

Because there are very strong limits on non-conservation of $B - L$ from proton decay experiments, these terms present a challenge for the [MSSM](#). It would be possible, of course, to simply tune the λ parameters to be small enough to fit within experimental constraints, but these terms can also be eliminated by introducing a new conserved quantity, R -parity. It is defined by

$$P_R = -1^{3(B-L)+2s} \quad (25)$$

where s is the spin of the particle. Requiring that all terms in the Lagrangian have a multiplicative P_R of 1 excludes the terms in [Equation 24](#), removing the interactions that would lead to proton decay. All [SM](#) particles are R -parity even, while the sparticles are R -parity odd, so the conservation of R -parity can translate into a conservation of number of sparticles minus the number of anti-sparticles. As a consequence, massive sparticles typically decay through a chain of lighter sparticles, emitting [SM](#) particles along the way.

2.2.2 Solutions to Standard Model Problems

Perhaps the most compelling consequence of [SUSY](#) comes from R -parity, which, through the formation of a new quantum number unique to sparticles, requires the Lightest Supersymmetric Particle ([LSP](#)) to be stable. This stable particle, if it is not electromagnetically charged, provides an excellent candidate [DM](#) particle. The lightest neutralino, for example, is a viable [DM](#) candidate because it does not interact electromagnetically or strongly, a constraint required due to measurements of the relic density of [DM](#) in the universe. An interaction cross-section higher than what's expected for weak interactions would have led the [DM](#) particle and its anti-particle to annihilate at lower densities, leaving a much smaller amount of [DM](#) in the universe than what is observed today [26].

Many believe that a complete [SM](#) should include a unification of the three forces, as electromagnetism and the weak force have already been unified. This requires that at some higher energy, the coupling constants of all three forces merge. However, in the [SM](#), the coupling constants come close to aligning, but don't perfectly cross. With the addition of [MSSM](#) particles with masses at the TeV scale, the alignment is near perfect, as shown in [Figure 13](#). This may be a mathematical coincidence, but it's very compelling to those physicists who believe that a *Grand Unified Theory* must exist.

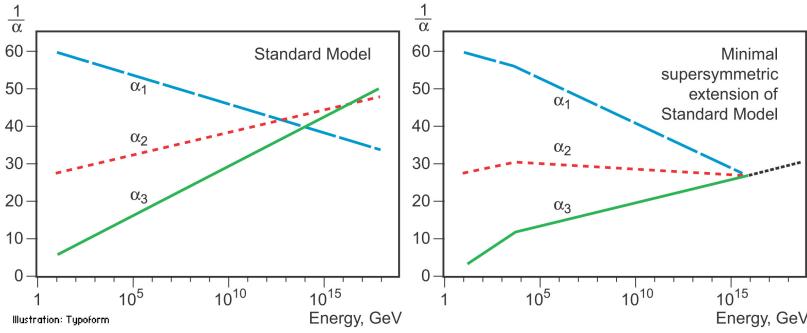


Figure 13: Running of the strong, weak, and electromagnetic coupling constants for the SM (left) and MSSM (right). [27]

SUSY also has the potential to solve the naturalness problem in the SM. In the SM, fine tuning is required to cancel the quadratic corrections to the Higgs mass that result from loops involving, most importantly, the top quark. In the MSSM, a similar loop involving the stop quark (the vertex for which is depicted in Figure 12) contributes to the Higgs mass corrections with the opposite sign, making it possible to naturally cancel the corrections without fine tuning. However, the larger the mass difference between the top quark and stop quark, the larger the remaining correction when the two terms cancel. Consequentially, to preserve a reasonable degree of naturalness (and here the definition of “reasonable” is subject to some debate), the stop quark should appear at masses not too much larger than the top’s, at approximately the TeV scale.

This naturalness mass limit, as well as the unification of couplings, make the argument for searching for SUSY at the LHC particularly compelling, as the LHC is the first collider capable of producing particles at the TeV scale. As new exclusions on SUSY are set, the remaining phase space becomes slightly less natural. But there is no shortage of new SUSY models with unexcluded parameters, which are continually proposed as new limits are created.

2.2.3 Simplified Models of Supersymmetry

There are many different theorized models of SUSY, with different mechanisms for breaking the symmetry. The MSSM has 120 free parameters, with complex interactions that determine the mass hierarchy and interaction rates of the sparticles. From an experimental point of view, the details of these theories and the exact way the hierarchies are generated are less relevant to a search than their outputs.

Simplified models, which are typically inspired by more complete theories, are used to tune the observables of a model more directly. These models each consist of one production and decay diagram, with the masses and branching ratios of the particles free to be tuned directly. In a more complete theory, it is instead necessary to mod-

ify more fundamental parameters like the symmetry breaking scale. A change like this impacts the properties of all the sparticles, and the details of its impact are model dependent. The simplified models allow for relatively model independent interpretations that can be reinterpreted in the context of a more complete SUSY theory [28].

In the analysis presented in Part iv, a simplified model is used which produces the decay depicted in Figure 14 with a 100% branching ratio. This decay chain begins with the pair production of gluinos, which decay via a pair of quarks to the second lightest neutralino, which in turn decays via a Z boson to the lightest neutralino. Only events in which the Z boson decays to pairs of electrons or muons are considered, with all-hadronic final states left to other searches [29]. The quarks emitted in the gluon decay are allowed to have flavors u , d , c , and s , each with a 25% probability. A similar model involving squark pair production is also considered, with the same breakdown of flavors. In this simplified model, the lightest neutralino is the LSP, and is stable.

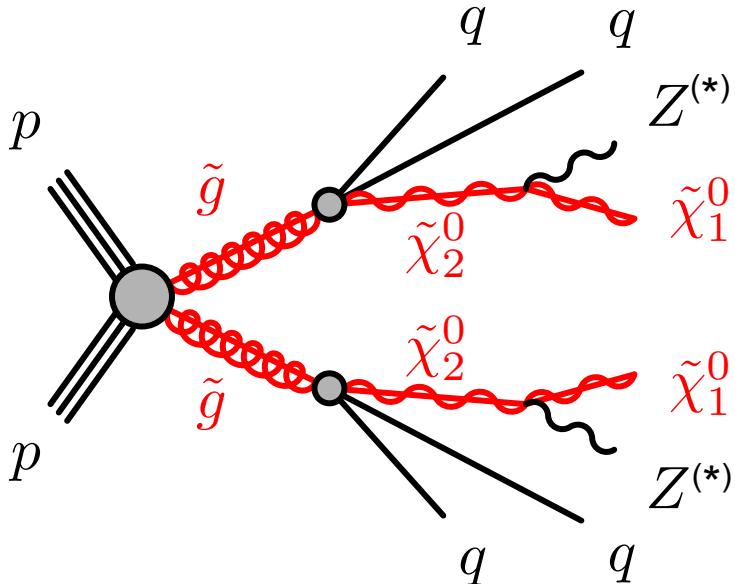


Figure 14: Feynman diagram of the decay considered in the simplified models used in the analysis presented in Part iv.

Using this simplified model, the masses of the particles can be set directly. This is very helpful for the generation of Monte Carlo simulation (MC), discussed in Section 4.7, because a grid of different mass values of the important sparticles involved in the decay can be generated. Cross-section calculations are performed for each point on these grids [30]. These simulated signals allow analyzers to make predictions of likely signals, and to exclude the simplified models as a function of the mass of the sparticles in the case that no discrepancies between predictions and observations are seen.

2.2.3.1 Context and Motivation

This simplified model is well motivated from a theoretical perspective. Production of strongly charged sparticles, shown in Figure 15, is hypothesized to occur at much larger rates than the production of other particles, due to the difference in coupling constants.

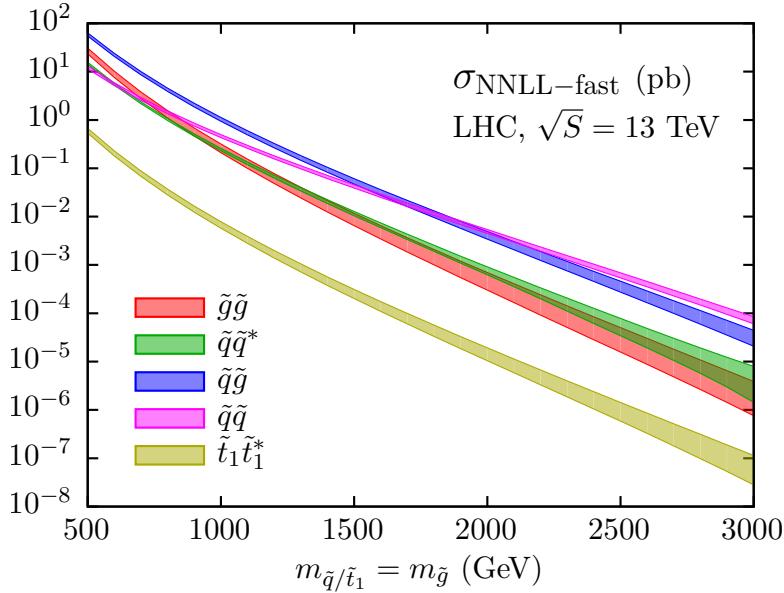


Figure 15: 13 TeV production cross-sections for sparticles, as a function of sparticle mass [31].

The specific decay considered in these simplified models does not have the largest branching ratio of all possible decays; even considering only changes to the SM decays involved, a $Z \rightarrow qq$ decay is roughly seven times more likely than $Z \rightarrow ll$. However, processes with higher branching ratios, like those producing an all-jet final state, often have much higher SM backgrounds, making them difficult to identify, even if they occur more frequently. This final state balances SM backgrounds and branching ratios, and when compared to other searches performed by the ATLAS collaboration, has competitive expected sensitivity to SUSY [32].

Processes similar to the one described by Figure 14 have been the target of previous LHC searches. Both the CMS and ATLAS Experiments performed searches for SUSY in the two-lepton channel with the 8 TeV data collected in 2012. The ATLAS search saw a 3σ excess, shown in Figure 16 [1]. The CMS search saw no excess in a similarly motivated signal region, albeit with different kinematic cuts than ATLAS's [33]. The CMS search followed up on a 7 TeV search in the same channel that also saw no excess [34].

Both searches also identified events with two leptons that weren't consistent with an on-shell Z decay, and in this region, an excess with a local significance of 2.4σ was observed by CMS, shown in Figure 17.

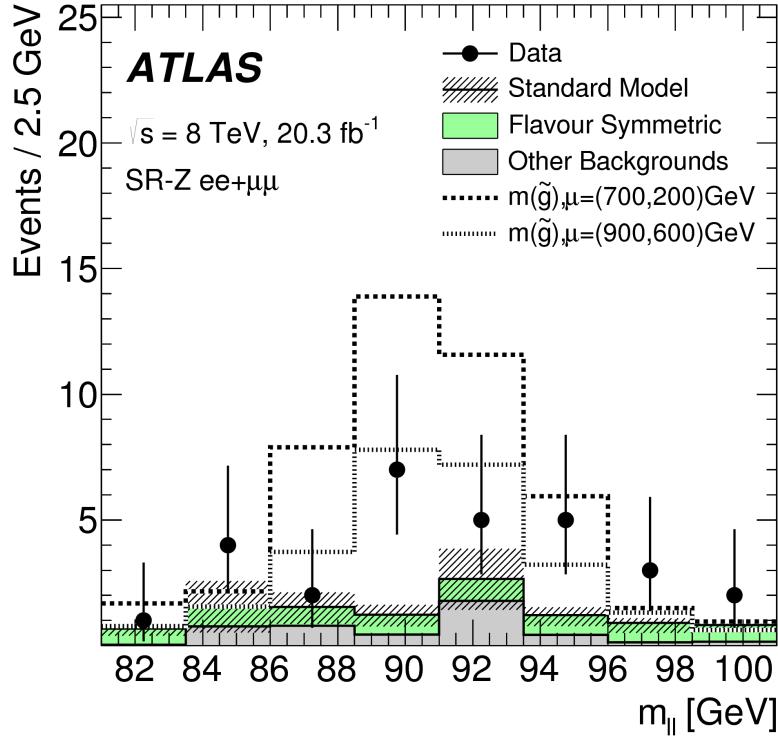


Figure 16: Results of an 8 TeV search performed by the [ATLAS](#) collaboration in a signal region targeting events like those in [Figure 14](#). The events in the signal region are displayed as a function of $m_{\ell\ell}$, the invariant mass of the event's leading leptons. The [SM](#) backgrounds are shown with their full uncertainties based on data-driven background estimations, and two signals are superimposed on the distribution. The observed data-points are higher than the expected background, with a total excess of 3.0σ [[1](#)].

No excess was observed by the [ATLAS](#) collaboration in a signal region with identical kinematic cuts [[1](#)].

These two excesses generated significant interest in the two lepton channel, and both [CMS](#) and [ATLAS](#) produced preliminary results in December 2015 with the first 3.2 fb^{-1} of 13 TeV data. [ATLAS](#) again reported an excess on the Z mass peak [[2](#)], shown in [Figure 18](#), while [CMS](#) saw no excesses [[35](#)]. An all-hadronic [ATLAS SUSY](#) search released in 2016 which was sensitive to the same models also saw no significant excess [[29](#)].

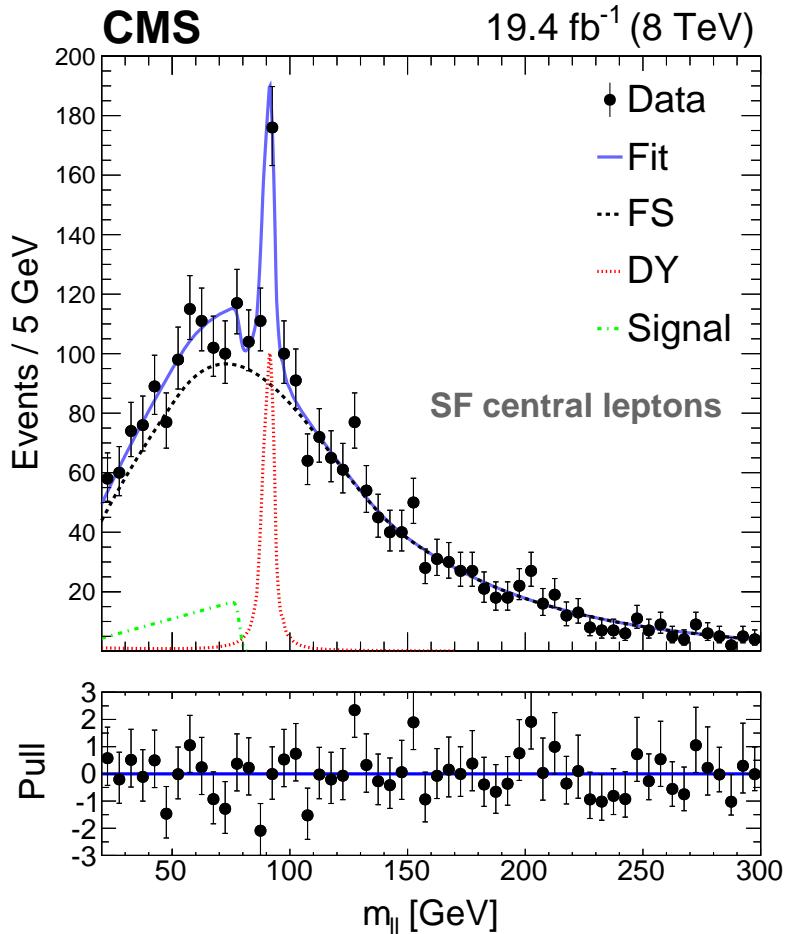


Figure 17: Results of an 8 TeV search performed by the CMS collaboration in a signal region including a broad range of $m_{\ell\ell}$. A 2.4σ local excess is seen in the low $m_{\ell\ell}$ region, and no excess of events is seen in the region with $m_{\ell\ell}$ consistent with an on-shell Z boson. The data is fit based on a data driven estimate of the flavor symmetric background (FS) and the Drell-Yan background (DY), with an additional component for the signal [34].

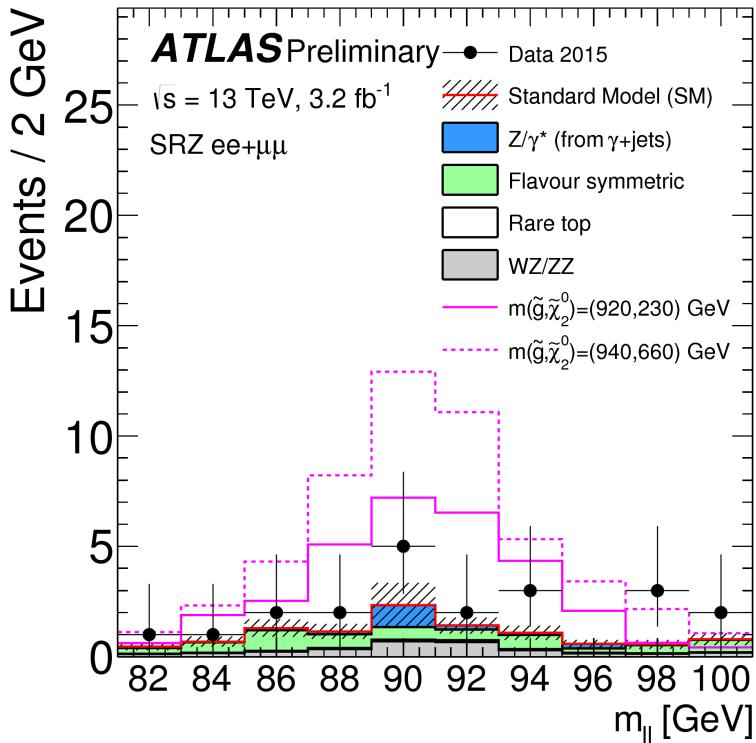


Figure 18: Preliminary results from a 13 TeV search targeting the same signal region as Figure 16, performed on 3.2 fb^{-1} of 2015 data. The events in the signal region are displayed as a function of $m_{\ell\ell}$, the invariant mass of the event's leading leptons. Flavor symmetric and Z/γ^* + jets backgrounds are taken from data-driven methods, while the other backgrounds are taken from MC. They are compared to the data, which shows a 2.2σ excess of events. Distributions from two signal points are superimposed [2].

Part III

THE EXPERIMENT

This section describes the LHC accelerator and the ATLAS detector, which collectively provide the physical environment and the data collection for the analysis discussed in Part iv. The process of Monte Carlo generation is discussed, including the modeling of hard-scattering processes and the simulation of the detector. Reconstruction of events in the ATLAS detector is also explained, with an emphasis on the reconstruction of tracks in the innermost part of the detector.

3

THE LARGE HADRON COLLIDER

The Large Hadron Collider ([LHC](#)) is unique in the world, producing proton-proton collisions at energies nearly an order of magnitude higher than any accelerator before it [36]. It provides unique environments at its collision points where massive, unstable particles can exist for an instant, then decay to the lighter, more stable [SM](#) particles normally observed in the universe. It is the goal of the [ATLAS](#) experiment to identify these short-lived particles, but the [LHC](#)'s work of producing them is equally complex.

The [LHC](#) was built in a 26.7 km circular tunnel that straddles the French-Swiss border outside of Geneva, originally built in 1989 for the Large Electron-Positron ([LEP](#)) collider [37]. In the [LHC](#), two beams of protons are accelerated to 6.5 TeV, then focused and collided at four points around the ring, which can be seen in [Figure 19](#).

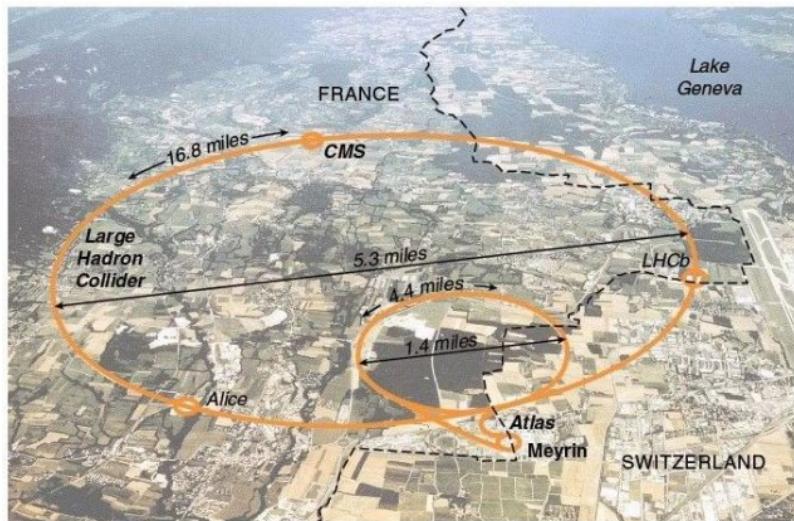


Figure 19: The [LHC](#) main collider ring and pre-accelerator [SPS](#) overlaid on a map of Switzerland and France, with the four main [LHC](#) experiments identified.

These points are each encased by particle detectors, which can examine the outputs of the collisions, and have different strengths and goals. The two multipurpose detectors are [ATLAS](#) and [CMS](#), which have very complex detectors aimed at measuring as many [SM](#) particles as possible and discovering new processes [38, 39]. [LHCb](#) (Large Hadron Collider beauty) examines processes related to the b quark [40]. Meanwhile, [ALICE](#) (A Large Ion Collider Experiment) focuses on special runs of the [LHC](#) which collide lead ions instead of protons,

and seeks to understand the high energy densities resulting from the collisions of such massive, complex particles [41].

3.1 THE INJECTOR COMPLEX

The primary goal of the LHC is to provide high luminosity proton-proton collisions at 13 TeV. To achieve this, it must be capable of rapidly accelerating large numbers of protons and holding them at a constant energy, and organizing them into bunches which can be focused and collided at precise points and times. To do this, a complex system of pre-accelerators is required, as well as a precisely engineered system of magnets within the LHC. The full system of pre-accelerators is shown in Figure 20.

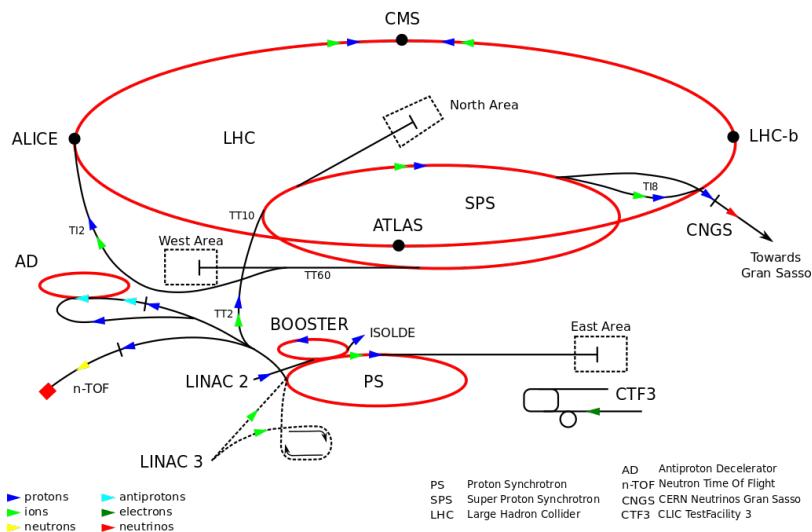


Figure 20: The pre-accelerators of the LHC.

The chain begins with when hydrogen gas is stripped of its electrons and injected in short pulses into Linac2, a linear accelerator which uses Radiofrequency (RF) cavities, which use alternating positive and negative electric fields to simultaneously push and pull particles forward through the accelerator. This RF behavior keeps the bunches of protons resulting from the original pulses separated, beginning the formation of the bunch structure used for collisions. The quadrupole magnets along the accelerator keep the beam focused. By the end of this accelerator, protons have reached 50 MeV.

At the end of Linac2, the proton beam is injected into the Proton Synchrotron Booster (PSB), the first circular accelerator in the pre-accelerator chain. It increases its magnetic field as the protons increase in speed, ultimately accelerating them to 1.4 GeV.

At this point the proton beam is injected into the Proton Synchrotron (PS), a 600 m long circular accelerator that consists of RF cavities that

accelerate protons up to 25 GeV, as well as room-temperature electromagnets that bend the beam.

The last accelerator before injection into the LHC is the SPS, a 7 km long ring which, long before the LHC tunnel was built, was used for the discovery of the W and Z bosons. The SPS accelerates particles up to 450 GeV before they are launched into the LHC.

Proton bunches are structured for ease of acceleration, with distinct features resulting from each of the pre-accelerators. The PS produces 72 bunches separated by 25 ns, which are injected into the SPS. However, as the magnetic field directing these protons out of the PS loop is turned on, there must be a gap in the bunch structure. Without this gap, called the injection kicker rise time, the changing magnetic field would direct particles out of the accelerator and produce high amounts of unsafe radiation around the PS. A similar gap in bunch structure is required for the injection from the SPS to the LHC. The injection process is repeated until the LHC is completely filled with over 2000 bunches, which takes about three minutes.

3.2 OPERATION OF THE LARGE HADRON COLLIDER

The LHC consists of eight straight sections each connected by an arc. In each straight section, RF cavities accelerate the protons, ultimately bringing them up to 6.5 GeV. Between these straight sections, 8.4 T dipole magnets bend the beams to maintain the approximately circular path. However, because the LHC is a proton-proton collider as opposed to a proton-antiproton collider, the two counter-rotating beams must be housed in separate rings and be accelerated separately. To achieve this, twin-bore superconducting magnets, one example of which can be seen in Figure 21, surround the two rings and accelerate them both. Quadrupole magnets are used at the four collision points to focus the beams, which cross at an interaction point at the center of each detector. In total there are over 6000 superconducting magnets in the LHC, which are kept below their critical temperature of 1.9 K by liquid helium cooling.

When first injected into the LHC, the protons must be accelerated with increasing RF frequencies over many turns through the machine. At the same time, the magnetic field from the dipoles increases to apply more force with which to bend the beam. Once the protons have reached a maximum energy, a process called *squeezing* occurs. The quadrupole magnets are used to reduce the total transverse area of the beam and elongate the bunches slightly. The shape produced by this process determines the *beam spot* for the ATLAS detector, the area in which collisions occur within the detector. As shown in Figure 22, the collisions mostly occur within 0.5 mm of one another in the $x - y$ plane, but have a spread of about 400 mm in the z direction¹.

¹ The coordinate system used here is discussed in Section 4.1.

LHC DIPOLE : STANDARD CROSS-SECTION

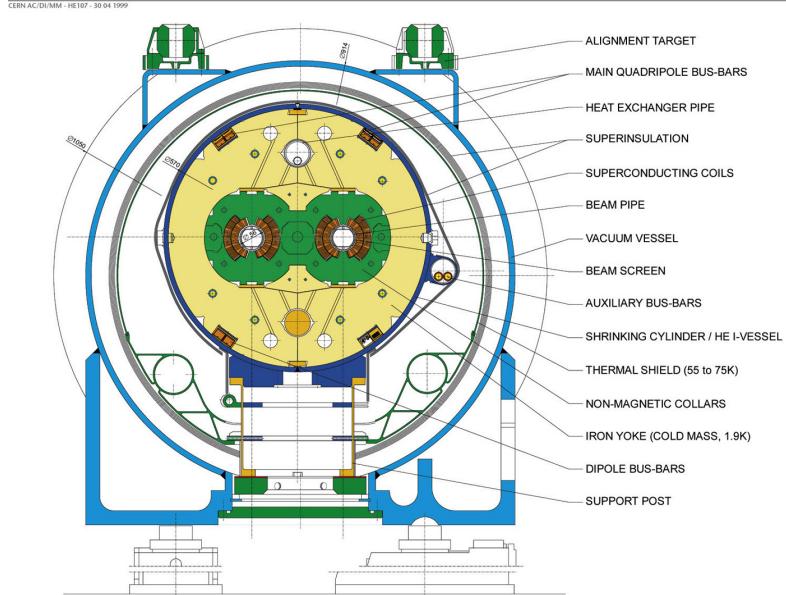


Figure 21: Cross-section of a cryodipole magnet in the [LHC](#).

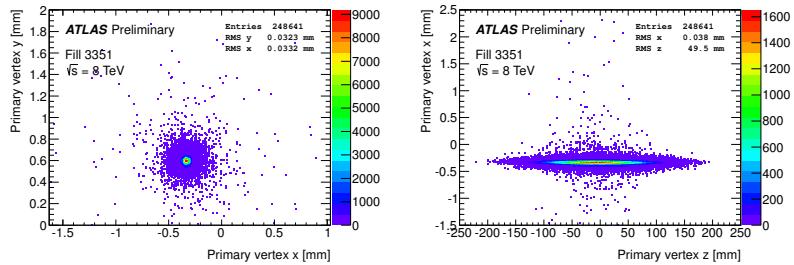


Figure 22: Beam spot in the [ATLAS](#) detector for one run in 2015. Distributions show only the highest p_T vertex per event. Left is the $x - y$ distribution of vertices, while the right plot shows the $x - z$ distribution.

Once the beams are at a stable energy and have been squeezed, the [LHC](#) indicates that it is ready to provide collisions to the experiments around the ring, and, after some additional checks by each experiment, data-taking can begin. As collisions occur, the number of protons in the beam decreases, and when it is sufficiently depleted to require a new fill, or if any instability occurs, the beam is dumped into a cavern filled with steel and concrete, which absorbs the energy.

3.3 LUMINOSITY

The goal of the collisions provided by the [LHC](#) is to produce [SM](#) and [BSM](#) particles, which can be observed by the detectors. How fre-

quently a given process could occur was a crucial consideration in its design. The number of events of a given type is given by

$$N_{\text{event}} = \int dt L\sigma_{\text{event}} \quad (26)$$

where L is the luminosity delivered by the [LHC](#) and σ_{event} is the cross-section of the process in question. These cross-sections vary over many orders of magnitude for different processes, as shown in [Figure 10](#), which gives the cross-sections for many different [SM](#) cross-sections as a function of energy. Because the higher-mass processes are so rare, a large amount of luminosity is required to produce them, and because other processes like jet production occur so much more frequently, they must be produced at high enough rates that analyzers have enough statistical power to differentiate them from more common events.

The instantaneous luminosity at the [LHC](#) is given by

$$L = \frac{N_b^2 n_b f_{\text{rev}} \gamma_r}{4\pi \epsilon_n \beta^*} F \quad (27)$$

where N_b is the number of protons per bunch ($\sim 10^{11}$), n_b is the number of bunches in each beam ($\sim 10^3$), f_{rev} is the number of times per second that the beam travels around the ring (1 kHz), γ_r is the relativistic gamma factor (~ 7000), ϵ_n is the normalized transverse beam emittance ($\sim 4 \text{ mm } \mu\text{m}$), and β^* is the β -function at the collision point, which describes the transverse displacement of particles in the beam ($\sim 0.5 \text{ m}$). F gives the reduction factor due to the geometry of the beam crossings, and is given by

$$F = (1 + (\frac{\theta_c \sigma_z}{2\sigma^*})^2)^{-1/2} \quad (28)$$

where θ_c is the crossing angle of the beams, σ_z is the RMS of the bunch length in the z direction, and σ^* is the same in the transverse direction.

As the proton beams circulate and collide, N_b decreases, producing a falling instantaneous luminosity, as seen in a Run 1 example in [Figure 23](#). In Run 2, peak instantaneous luminosity was brought up to $1.39 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$. This high instantaneous luminosity and consistent running resulted in much faster data collection than in Run 1. The accumulation of luminosity for the two years of Run 1 and the two years of Run 2 is depicted in [Figure 24](#).

3.4 PILE-UP IN PROTON-PROTON COLLISIONS

One consequence of the high instantaneous luminosity is *pile-up*, or multiple simultaneously occurring interactions. Because the instantaneous luminosity is much larger than the total proton inelastic cross-section, it is very likely that multiple protons will collide in the same

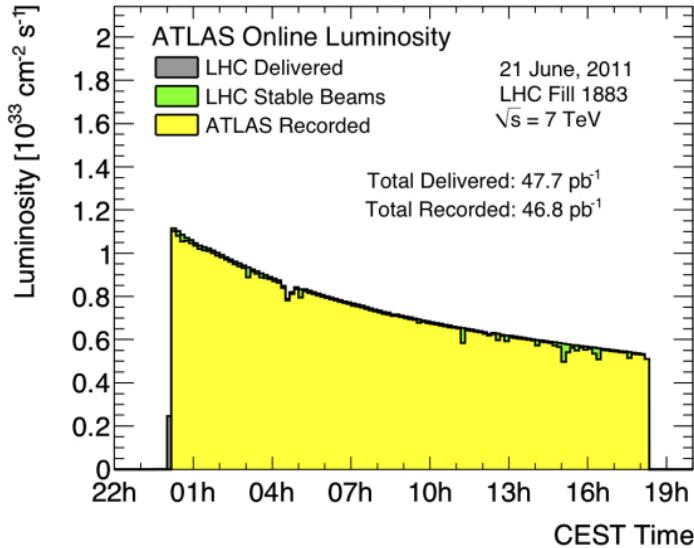


Figure 23: Instantaneous luminosity of one fill of 7 TeV data in 2011.

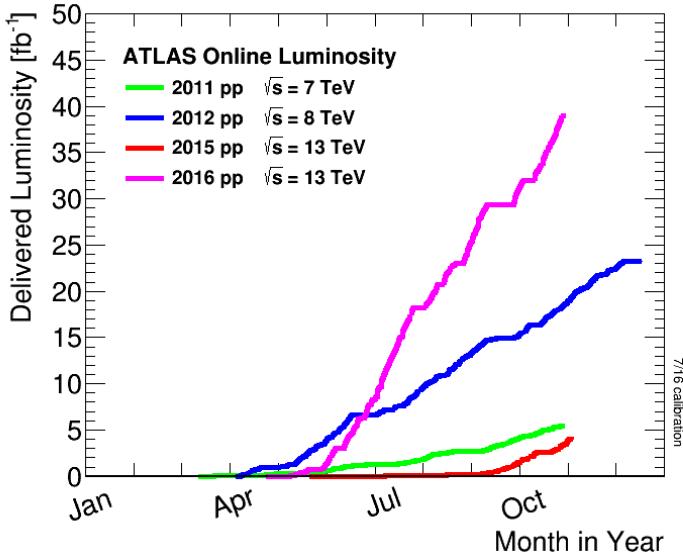


Figure 24: ATLAS luminosity for Run 1 and Run 2, as of September 2016.

bunch crossing. In fact, the average number of simultaneous interactions in 13 TeV data, shown in Figure 25, is about twenty. For detectors with read-out windows larger than the bunch spacing of the LHC, *out-of-time* pile-up can also occur. In these cases, an event mistakenly includes measurements of objects resulting from collisions in other bunch crossings.

Pile-up is a difficult challenge for the ATLAS analyses because it typically results in additional jets in an event, and can increase SM backgrounds for analyses seeking to identify events with jets. It can also add to the overall hadronic energy of an event, and that energy can be mis-assigned to other objects. Fortunately, the multiple interac-

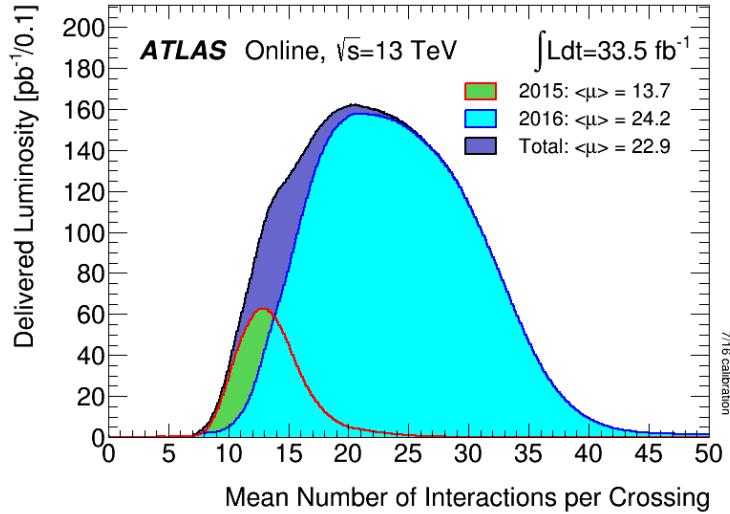


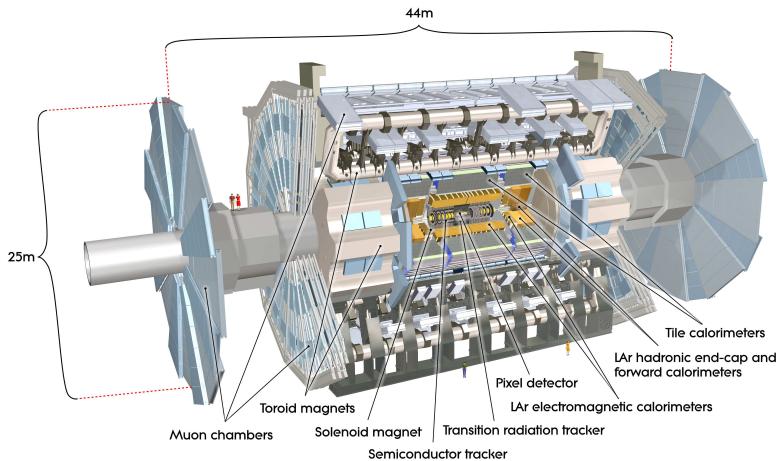
Figure 25: Average number of interactions per crossing shown for 2015 and 2016 separately, as well as the sum of the two years.

tions occur at different points along the beam line, and the particles emerging from each can be traced back to create *vertices* identifying the location of the interactions. In most cases, it is possible to resolve the different vertices that each proton-proton collision makes, and so pile-up jets can be identified and rejected.

THE ATLAS DETECTOR

The [ATLAS](#) detector circumscribes the [LHC](#)'s beam pipe, enclosing the collision point with a series of particle detecting layers, aimed at making as many measurements of the particles leaving the collision point as possible. Its goal is to get a precise measurement of all the stable or semi-stable particles flying from proton-proton collisions at its center, allowing analyzers to fully reconstruct the kinematics of the underlying processes.

The [ATLAS](#) detector is the largest detector of its kind, measuring 44 m in length and 25 m in height, as seen in [Figure 26](#). The size is mainly determined by the constraints of the Muon Spectrometer ([MS](#)), discussed in [Section 4.4](#), which is the largest and outermost subsystem. The [MS](#) is submerged in a spatially varying magnetic field provided by three toroidal magnets, while the Inner Detector ([ID](#)) ([Section 4.2](#)) is encased by a superconducting solenoid, which provides a uniform 2 T field throughout its volume [38]. A calorimeter system is located between the [MS](#) and [ID](#), with components to measure the energy of electromagnetic and hadronic systems.



[Figure 26](#): Diagram of the [ATLAS](#) detector, with subsystems and magnets identified.

4.1 COORDINATE SYSTEM USED IN THE ATLAS DETECTOR

The [ATLAS](#) detector is centered around the pp collision point, and is built radially out from the beam pipe, maintaining as much rotational symmetry around the beam pipe as possible. It is also symmetric in

the forward-backward directions. A coordinate system using the collision point as the origin is used, with the beam line defining the z -axis in the counter-clockwise direction. The positive x direction is defined as pointing to the center of the LHC ring, while the positive y direction points upwards. For ease of reference, the side of the detector in the positive- z direction is referred to as the A side, and the other side is referred to as the C side.

Because of the cylindrical design of the detector, angular coordinates are often used. The azimuthal angle ϕ defines the angle around the beam pipe and the polar angle θ defines the angle from the beam axis (z). However, a transformation of the polar angle called pseudo-rapidity (η) is used more often, and is defined as

$$\eta = -\ln(\tan \frac{\theta}{2}). \quad (29)$$

η is used because the particle distribution from LHC collisions is roughly uniform in this variable. Building on this definition, angular distance between objects is typically defined as

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}. \quad (30)$$

Often variables are defined purely in the transverse plane, which is indicated by a subscripted T, as in p_T , which gives an object's transverse momentum.

4.2 THE INNER DETECTOR

The Inner Detector (ID) is used for the measurement of tracks, estimates of the paths charged particles take as they travel through the detector. Collisions in the detector produce about 1000 particles, so identifying and differentiating all the tracks resulting from a collision is challenging.

The ID consists of three separate subdetectors, each of which has multiple layers capable of producing an electrical signal, called a *hit*, when a charged particle travels through its active material. ATLAS tracking software considers all these hits and forms tracks, with the goal of minimizing fake tracks due to random noise and maximizing the efficiency of identifying a real particle. The details of this procedure are discussed in Chapter 5. The full ID can be seen in Figure 27, while a schematic in Figure 28 shows more detail on the placement of each layer.

4.2.1 The Pixel Detector

The Pixel Detector lies closest to the beam pipe of the LHC, and has four layers comprising 92 million pixels. There are three standard layers, referred to as Layers 0-2 (L₀, L₁, L₂), and an additional layer added for the 2015 data-taking, called the Insertable B-Layer (IBL).

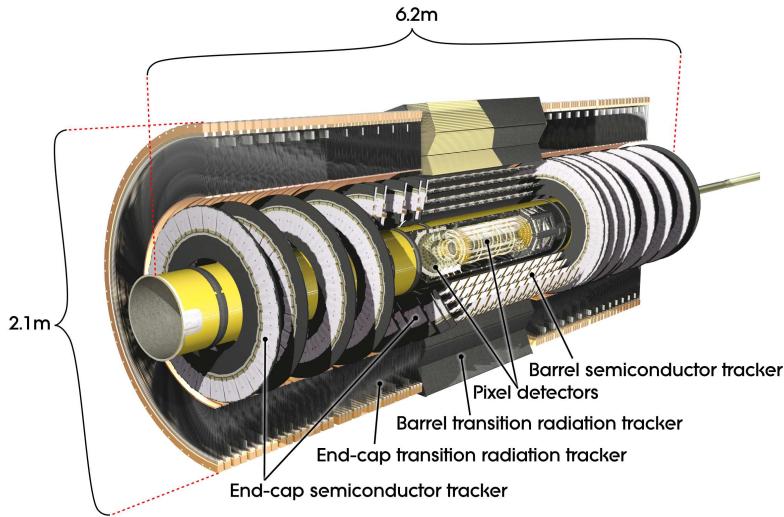


Figure 27: Diagram of the [ATLAS](#) Inner Detector, containing the Pixel, SCT, and TRT subsystems.

4.2.1.1 *The Original Pixel Detector*

The Pixel Detector consists of high-precision silicon chip pixel modules, with 1744 in total, and each module is made up of 16 sensors each with its own read-out system. Each sensor is identical, containing 47232 pixels, which are typically each $50 \times 400 \mu\text{m}^2$, though pixels at the edges of the sensors are slightly longer, at $50 \times 600 \mu\text{m}^2$.

As shown in [Figure 28](#), the central η region (barrel) is covered by three concentric cylindrical layers of sensors with radii of 50.5 mm, 88.5 mm, and 122.5 mm. In the higher η region (endcap) is covered by a series of three disks positioned in the $x - y$ plane. Together, they give complete coverage out to $|\eta| = 2.5$, and a particle coming from the collision point will typically produce hits in three layers.

The sensors are n-type silicon wafers with a voltage applied, and a passing charged particle produces thousands of electron-hole pairs inside the material, which drift in the electric field towards the mounted read-out system. A hit occurs when the resulting current becomes large enough to pass a threshold designed to suppress noise. A larger total charge deposit will result in the signal remaining over the threshold for a longer period of time. This Time Over Threshold ([ToT](#)) is recorded along with the initial timing of the hit. This measurement is spatially accurate in the barrel (endcap) to $10 \mu\text{m}$ in the $R - \phi$ direction and $115 \mu\text{m}$ in the $z (R)$ direction.

4.2.1.2 *Addition of the IBL*

In 2014, the [IBL](#) was added to the Pixel Detector. This layer is placed directly on top of the beam pipe, inside barrel Lo, providing a measurement of particles only 3.3 cm away from the interaction point. The [IBL](#) consists of 14 overlapping staves, each containing 16 modules,

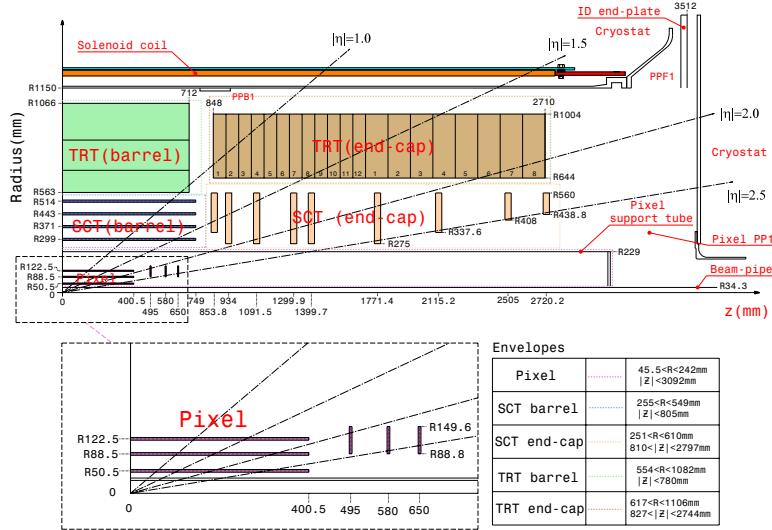


Figure 28: Diagram of one-quarter of the [ATLAS](#) Inner Detector in the $R - z$ plane, with lines drawn to indicate various η locations.

the geometry of which can be seen in [Figure 29](#). These modules contain pixels measuring $50 \times 250 \mu\text{m}^2$, and allow for particle detection in 90% their area, as compared to the 70% possible with the original pixel modules.

The [IBL](#)'s addition provides greater precision for all track measurements, but it is especially useful for the detection of B mesons, discussed in [Section 2.1.1.2](#), whose lifetimes of about 1.5 ps allow them to travel as much as a few mm before decaying. These decays lead to secondary vertices in [ATLAS](#) events. The location of the [IBL](#) gives a measurement closer to these secondary vertices, increasing the probability that these vertices can be resolved.

4.2.2 The Silicon Microstrip Tracker

The [SCT](#) employs a similar technology to the Pixel Detector, with 15912 sensors and 6.3 million readout channels. Its main difference from the Pixel Detector is in the readout, which is performed by a series of 12 cm long strips with a width of 80 μm . These layers are paired, placed on top of one another at a small (40 mrad) angle to allow for position determination in the ϕ and z directions. Together, these pairs of layers give four spatial measurements for each particle passing through the [SCT](#). In the barrel, these strips run parallel to the beam pipe, while in the endcap, they are arranged radially. These strips provide a hit resolution in the barrel (endcap) of 17 μm in the $R - \phi$ direction and 580 μm in the $z (R)$ direction.

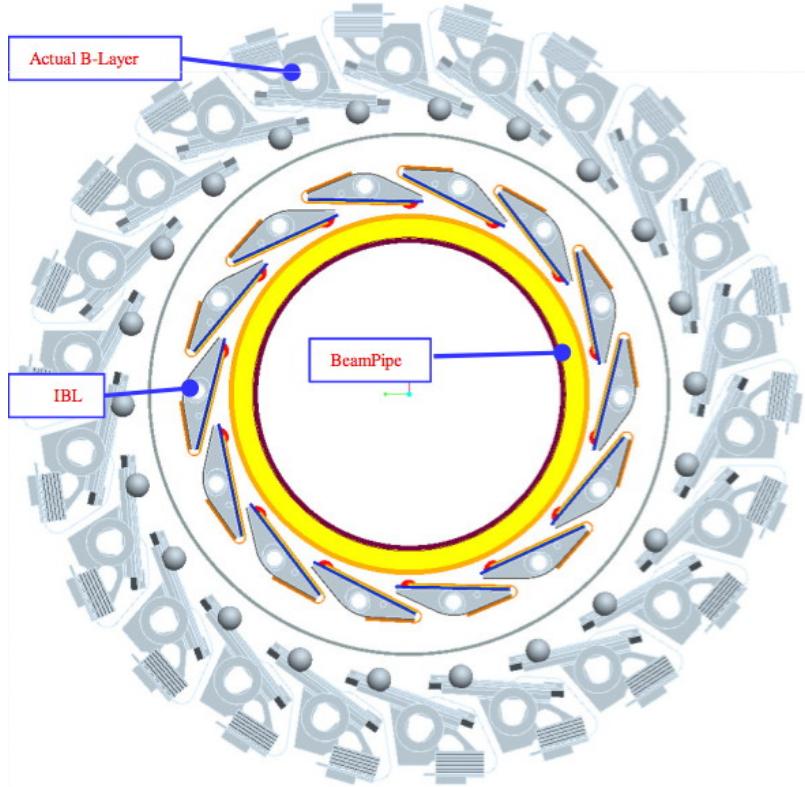


Figure 29: Diagram in the $x - y$ plane of the [IBL](#) and the innermost layer of the Pixel Detector, Lo [42].

4.2.3 *The Transition Radiation Tracker*

The Transition Radiation Tracker ([TRT](#)) uses 4 mm diameter gas-filled tubes, each with a high voltage wire suspended along the center of the tube. The tubes run the length of the barrel, with a separate wire in the positive and negative z direction. In the endcap, the tubes are arranged radially. In total, there are about 351,000 readout channels in the [TRT](#). This detector makes measurements only in the $R - \phi$ direction, where the resolution of each measurement is $130\ \mu\text{m}$, and coverage extends to $|\eta| = 2.0$. Each particle typically creates about 36 hits as it passes through the [TRT](#) barrel.

Particles passing through the gas mixture of the [TRT](#) ionize the gas, producing electrons which drift towards the wire due to a potential difference applied between it and the tube. In addition, particles passing through the [TRT](#) produce radiation as they transition between materials, with larger amounts of radiation for lighter particles. This radiation produces high-threshold signals in the [TRT](#) can be used to differentiate electrons from other heavier charged particles, such as pions.

4.3 THE CALORIMETERS

Unlike the tracking detectors, which aim to take measurements of a particle with minimal alterations of its trajectory, the calorimeters measure the energy of objects by stopping them entirely. Calorimeters contain alternating layers of absorber, a material that causes incoming particles to shower into lower-energy decay products, and an active material, which detects passing particles, allowing for the reconstruction of these showers.

The [ATLAS](#) calorimeters, which can be seen in [Figure 30](#), provide coverage out to $|\eta| < 4.9$. High granularity electromagnetic measurements are made within $|\eta| < 2.5$. In this range, high- p_T electrons have nearly straight tracks, making momentum measurement through track curvature difficult, leaving the calorimeter as the primary energy measurement. The hadronic calorimeters, as well as the higher $|\eta|$ electromagnetic calorimeters, have a coarser granularity.

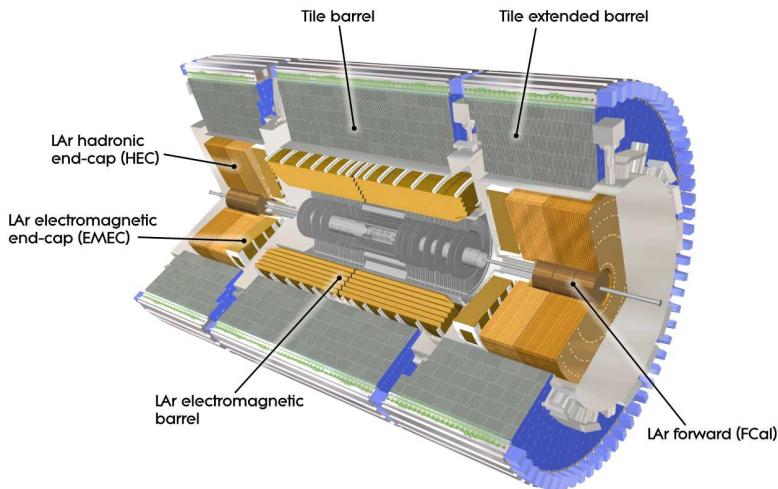


Figure 30: The calorimeter system of the [ATLAS](#) detector.

Besides measuring the energy of passing particles, another task of the calorimeter system is to limit punch-through to the [MS](#), described in [Section 4.4](#). All other particles must be fully stopped by the calorimeters to allow for clean signals from muons, and to measure the total energy of the particle. This requirement sets a minimum number of interaction lengths for each of the calorimeters.

THE LAR ELECTROMAGNETIC CALORIMETER uses liquid argon as its active detector medium alternating with layers of lead acting as the absorber. Signals are read out with capacitively coupled copper plates. The layers are shaped like accordions, which allows for complete coverage with multiple layers of active material, three in central η ($0 < |\eta| < 2.5$) and two at higher η ($2.5 < |\eta| < 3.2$). [Figure 31](#) shows the layout of a central η module, including this accordion-like

layering. At $|\eta| < 1.8$, an instrumented liquid argon presampler provides a measurement of energy lost prior to reaching the calorimeters. The total energy resolution for this detector is about $10\%/\sqrt{E}$, with an additional constant term of 0.2%.

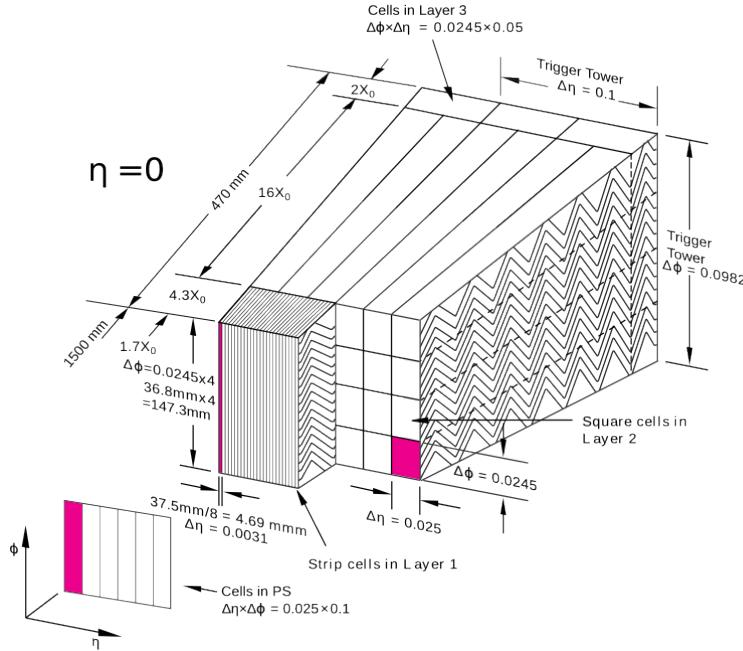


Figure 31: Layout of the LAr calorimeter module at central η [38].

THE TILE CALORIMETER is a hadronic calorimeter which is placed outside of the LAr Calorimeter. It uses layers of steel as its absorber with scintillating tiles as the active material between them, which are read out by photomultiplier tubes. The Tile Calorimeter covers $|\eta| < 1.7$ with a typical energy resolution of about $50\%/\sqrt{E}$ with a constant term of 5%.

THE LAR HADRONIC ENDCAP CALORIMETER covers the hadronic calorimetry for higher η . It uses liquid argon active material and copper plate absorbers, resulting in an energy resolution of approximately $70\%/\sqrt{E}$ with a constant term of 5%. This calorimeter covers $1.5 < |\eta| < 3.2$, overlapping with the hadronic calorimeters in either direction of its η range.

THE FCAL or forward calorimeter provides electromagnetic and hadronic coverage at very high η ($3.1 < |\eta| < 4.9$). This calorimeter also uses liquid argon as its active material, and uses copper-tungsten as the absorber. Its energy resolution is about $70\%/\sqrt{E}$ with a constant term of 3%.

4.4 THE MUON SPECTROMETER

The Muon Spectrometer (MS) measures charged particles that penetrate the calorimeter system. Because the calorimeters are designed to completely absorb electrons, photons, and hadrons, the MS mainly detects muons, which pass through the calorimeter with very little loss of energy. The goal of the MS is to give a high-precision measurement of these muons, and also to be able to quickly identify events with muons for the sake of triggering, discussed in [Section 4.6](#). The layout of the MS can be seen in Figures 32 and 33. Muons can be measured for all $|\eta| < 2.7$, and they can be triggered on for $|\eta| < 2.4$. The entire system is about 24 m tall and 40 m long.

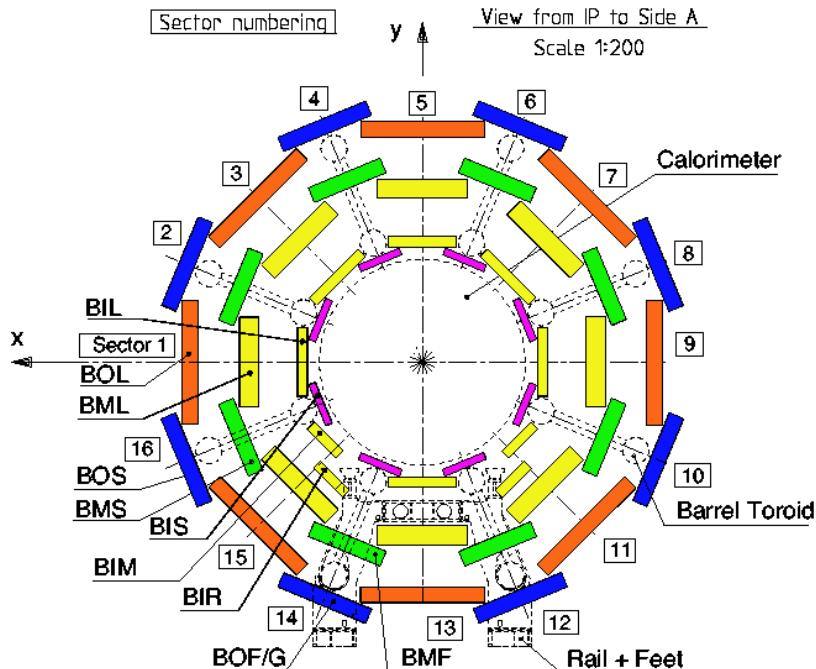


Figure 32: An x - y view of the MS. The three barrel layers are visible, as well as the overlapping, differently sized chambers. The outer layer of the MS is about 20m in diameter.

To achieve these goals, the MS has several subsystems. The system responsible for precision measurement is called the Monitored Drift Tubes (MDTs). This subdetector consists of chambers of three to eight layers of tubes, with three layers of chambers covering both the barrel and endcap regions. In the barrel, these chambers are arranged in layers concentric cylinders with small overlaps between adjacent chambers. The chambers are oriented such that the drift tubes are parallel to the beam line. In the endcap, the chambers form disks with drift tubes approximately aligned in the R direction.

The tubes each contain an Ar/CO₂ gas mixture and a single high voltage wire which runs at its center along its length. Charged particles excite the gas as they pass through it, producing electrons which

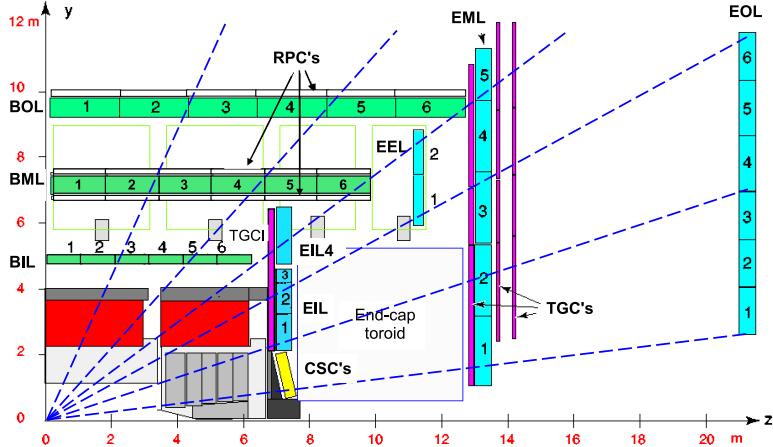


Figure 33: An r - z view of the MS. The three layers of the barrel and endcap MS are visible.

drift towards the high voltage wire. The resulting electric signal is read out, and the magnitude and timing of the signals are both used to differentiate particle traces from noise.

Though very effective at giving a precise measurement, the MDTs have two shortcomings. The first is that the measurement is only precise in the direction perpendicular to the tubes; in the direction parallel to them, the resolution is not much better than the length of the drift tube, which are typically several meters long. The resolution in the perpendicular direction is about $35\ \mu\text{m}$ with the combined measurement of all the tubes in a chamber. The second major shortcoming is that the MDTs are slow, with a maximum drift time of about 700 ns.

The slow drift time means that muons from sequential collisions can appear in the same event, and that the signals from the MDTs are received too late to be used for triggering. To solve the former problem, another detector called the Cathode-Strip Chambers (CSCs) is used in high-rate regions of the MS. This detector consists of multi-wire proportional chambers which have cathode strips on either side of the anode in orthogonal directions, providing a $40\ \mu\text{m}$ resolution in one direction and 5mm resolution in the other. Their drift times are much shorter than those of the MDTs, at about 40 ns. They are placed in the forward region of the detector ($2 < |\eta| < 2.7$) where the incident particle rates are highest.

To achieve responses fast enough to be used for triggering, Resistive Plate Chambers (RPCs) and Thin Gap Chambers (TGCs) are used. These chambers both take less than 25 ns to produce a signal. The RPCs are used in the barrel and are made up of two high-resistance plastic plates with a gas mixture under an electric field between them. Passing particles ionize this gas, and the resulting signal is read out via metallic strips mounted to the plastic plates. The TGCs used in the

endcap are a form of multi-wire proportional chambers, like the [CSCs](#). Unlike the [CSCs](#), the cathode is placed extremely close to the wires, speeding up its operation.

The massive [MS](#) is subject to deformations due to gravity and the magnetic field. To achieve a high precision alignment, these deformations are constantly monitored in each [MDT](#) chamber with a set of four optical alignment rays, which give alignment information at the precision of $<30\ \mu\text{m}$. In addition, a sag-adjustment system can use this information to re-align any wires that droop under gravity's pull. Lastly, the [MS](#) can be aligned using the tracks made from hits it measures, discussed more in [Section 6.3](#).

4.5 THE MAGNET SYSTEM

The [ATLAS](#) magnet system consists of four superconducting magnets: an inner solenoid, a barrel toroid, and two endcap toroids. Collectively, they are 22m in diameter and 26m long, and their basic layout can be seen in [Figure 34](#).

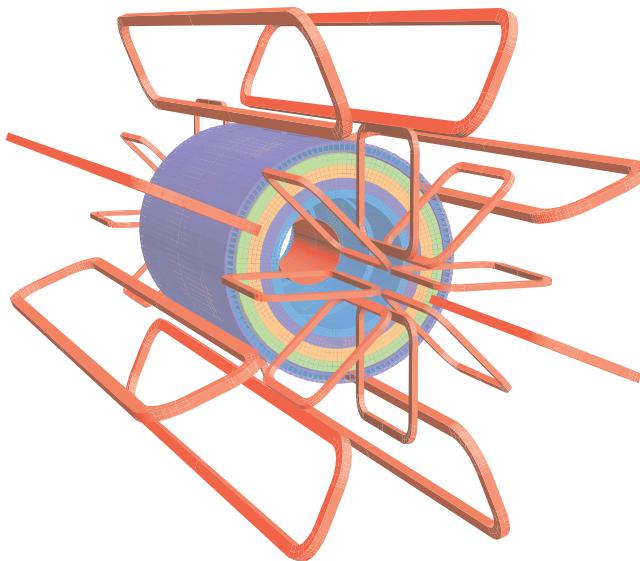


Figure 34: The magnet system of the [ATLAS](#) detector. The inner cylinder shows the solenoid which gives a uniform magnetic field in the [ID](#). Outside of that are the barrel and endcap toroids, which provide a non-uniform magnetic field for the [MS](#).

The solenoid is inside the calorimeter volume and provides a uniform 2T magnetic field for particles traveling through the [ID](#). This axial field causes the trajectories of charged particles to bend in the $x - y$ plane, and measurements of the curvature of these trajectories

give the most accurate p_T measurement for many particles according to the equation

$$p_T = qB\rho \quad (31)$$

where q is the charge of the particle, B is the magnetic field in the z direction, and ρ is the radius of curvature.

Because the solenoid is placed between the tracking system and the calorimeter, it is important that it interfere minimally with particles in order to allow the calorimeter to measure their full energies. The solenoid is placed inside the same vacuum chamber as the LAr calorimeter and is made of Al-stabilized NbTi superconductor with aluminum casing, giving it a total thickness of about 0.66 radiation lengths.

The barrel toroid is outside the calorimeters and provides the magnetic field for the barrel **MS**, which varies from 0.2–2.5T. The endcap toroids have a magnetic field range of 0.2–3.5T. All three toroid magnets are made with Al-stabilized Nb/Ti/Cu superconducting coils supported by Al-alloy struts.

The magnets are cooled with liquid helium, and take up to a month to be brought down to operating temperatures, about 4.5 K. All magnets have cold masses surrounding them to absorb heat in the event of a quench.

The B -field resulting from this magnet system can be seen in [Figure 35](#). The plot on top demonstrates the relatively constant field rate within the barrel which drops steeply at $|z|=2$. The bottom plot shows the field integral in the **MDTs** as a function of $|\eta|$, demonstrating the good coverage out to $|\eta|<2.6$ excluding a transition region between the barrel and endcap, where the field changes rapidly, making precise p_T construction difficult.

4.6 THE TRIGGER SYSTEM AND DATA ACQUISITION

The **LHC** provides proton bunch crossings every 25 ns, and each of these events contains about one MB of data, corresponding to 40 TB/s¹, a completely unmanageable amount of data. In addition to this concern, many of **ATLAS**'s subdetectors like the Pixel Detector, the LAr Calorimeter, and **MDTs** take much longer than 25 ns to read out, making keeping up with the bunch crossing rate impossible. To reduce the total data read out and allow for selective reading out of the slower detectors' buffers, a triggering system is used.

The trigger system uses fast detectors to get a coarse picture of an event's topology, which is then compared to a trigger menu, which lists the types of events that are interesting enough to keep. Overall,

¹ This number is actually an overestimate, as not all bunches are filled due to the gaps produced by the **LHC**'s injector complex, discussed in [Section 3.1](#).

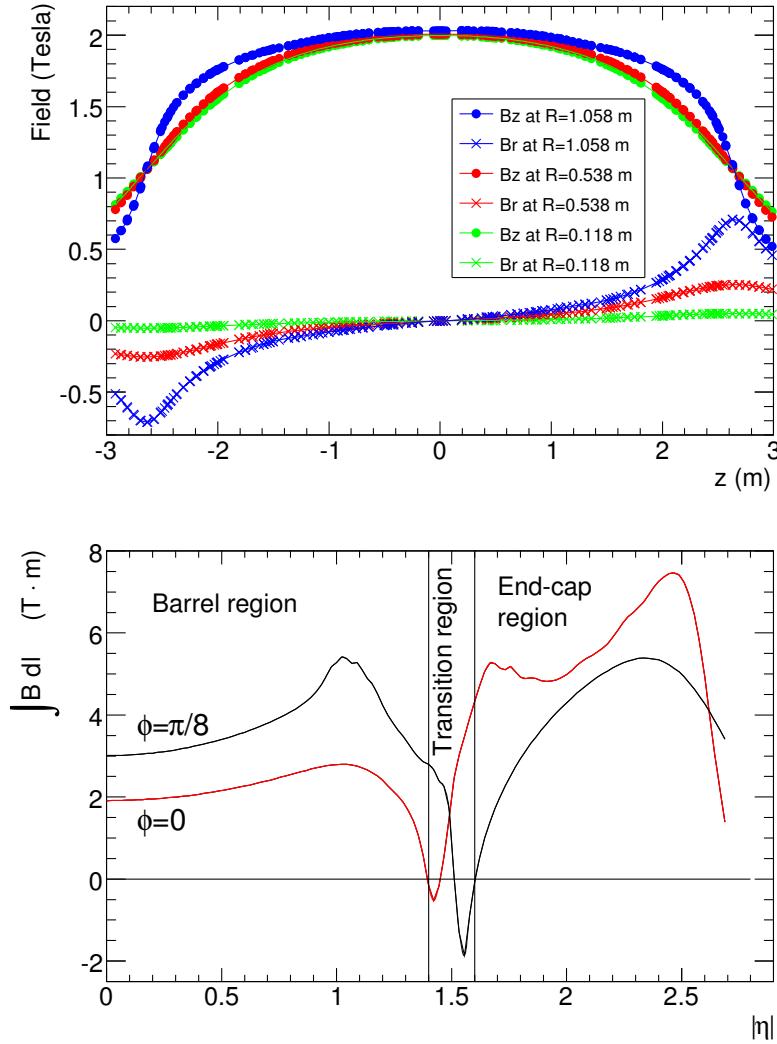


Figure 35: Plots of the magnetic field within the [ATLAS](#) detector. Top is the field (broken into its R and z components) as a function of z for several different values of R . Bottom is the field integral through the [MDTs](#) as a function of $|\eta|$ for two different ϕ values.

the trigger system reduces the 40 million events a second to about 1000 to be fully read out from the [ATLAS](#) detector.

This filtering of events is done in two steps: the [L₁](#) trigger is implemented in hardware and reduces the initial 40MHz to 100kHz, while the [HLT](#) is implemented in software, further reducing the rate to 1kHz [43]. The [L₁](#) trigger uses coarse granularity information from the fast read-out subdetectors: the calorimeters, the [RPCs](#) and [TGCs](#).

The coarse grained calorimeter information used for the [L₁](#) trigger decision is referred to as L₁ Calorimeter Trigger ([L₁Calo](#)) and uses information from all calorimeter systems. [L₁Calo](#) is responsible for all triggers excluding muons, meaning it must be capable of identifying a large number of different objects and event topologies, including

high- p_T objects, E_T^{miss} , and large amounts of hadronic energy. The trigger can also identify isolated objects, objects with very few calorimeter deposits from other objects near them.

For muon triggers, the trigger algorithm looks for patterns of hits from the [RPC](#) and [TGC](#) that are consistent with high- p_T muons with origins at the interaction point.

An example of the [L1](#) trigger rates for different types of events can be seen in [Figure 36](#) for one run in July 2016. The common features to all rates are due to [LHC](#) luminosity changes, deadtimes due to detector inefficiency, and adjustment of trigger rates to optimize bandwidth.

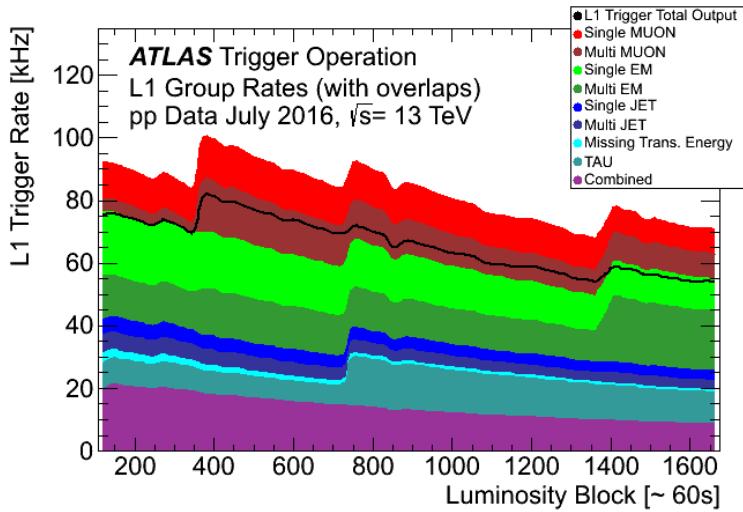


Figure 36: [L1](#) trigger rates for a run in July 2016 as a function of luminosity block, an approximately 60-second long period of data-taking. The total rate is lower than the combined stack because of overlapping triggers.

All of this information is analyzed by the Central Trigger Processor ([CTP](#)), which uses a trigger menu identifying all types of events to be kept to return a trigger decision. Due to the limited size of detector buffers, the event must be processed in about $2.5 \mu\text{s}$. This ensures that the information to be read out has not yet been overwritten when the trigger decision is made. This decision is passed to the Trigger Timing and Control ([TTC](#)), which communicates with all subdetectors. Upon receiving a [L1](#) trigger, the subdetectors read out all the information they've stored about the event and place it on their Read Out Boards ([ROBs](#)).

The [HLT](#) takes the data from particular Region of Interests ([RoIs](#)), areas containing interesting objects that caused the [L1](#) trigger. With a more complete picture of the hits observed by the detector, tracks are formed, and the [HLT](#) can use all of this information to determine whether or not the event is still interesting enough to keep. This pro-

cess has its own trigger menu with dedicated L_1 seeds for each item. HLT triggers typically have slightly higher thresholds than their corresponding L_1 triggers to ensure that events that would pass the HLT requirements are very likely to have passed the L_1 requirements. Figure 37 shows the HLT rates for the same run in July. In addition to the event types seen in Figure 36, the HLT can also identify events with b -jets, differentiate between electrons and photons, and identify events interesting for B-physics.

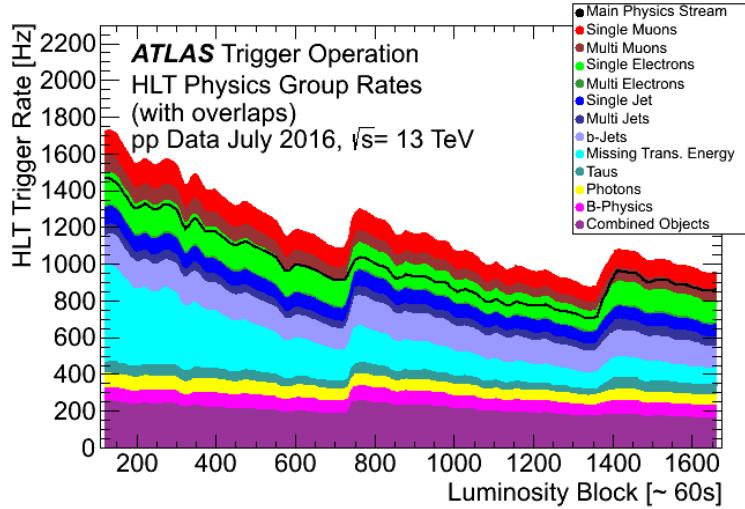


Figure 37: HLT trigger rates for a run in July 2016 as a function of luminosity block, an approximately 60-second long period of data-taking. The total rate is lower than the combined stack because of overlapping triggers.

Events passing the HLT trigger are written to disk to be analyzed. An example of the total trigger efficiency for single electron triggers is shown in Figure 38.

Events types that occur very frequently, such that it would require too much of the total trigger bandwidth to record all events passing a given threshold, are prescaled. Events passing these triggers are only recorded a fraction of the time, and these prescaling rates are used to weight events passing these triggers when they are analyzed. For example, the lowest unprescaled single electron trigger in 2016 data-taking required an electron with p_T of 60 GeV. A trigger requiring electrons with p_T of only 10 GeV also exists, but only one in ten events passing this trigger is recorded.

4.7 MONTE CARLO EVENT GENERATION

The complex events of the LHC are difficult to model, but modeling them is crucial to analyzers' understanding of SM backgrounds and potential signals. To simplify the modeling process, particle interac-

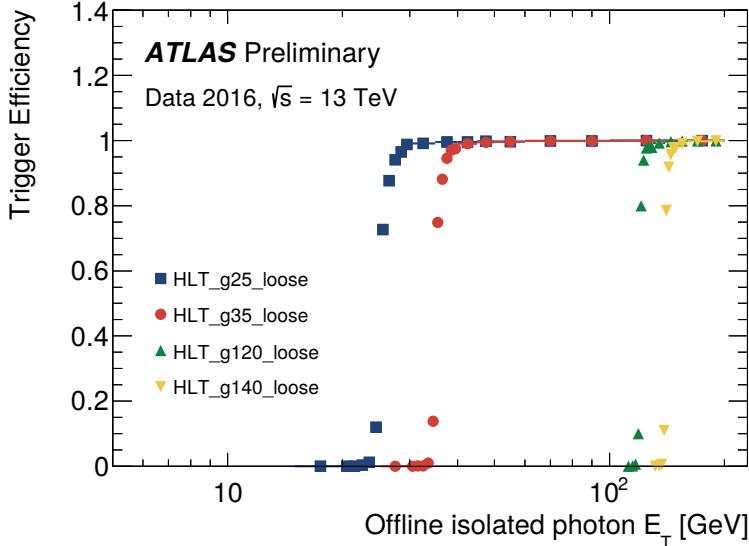


Figure 38: Photon trigger efficiency as a function of E_T for four different **HLT** triggers with photon p_T requirements of 25, 35, 120, and 140 GeV [44].

tions are broken down into very small steps, each with associated probabilities of various outcomes. This modeling method is called Monte Carlo simulation (**MC**), and, at the **LHC** it is broken into several larger steps which are each handled by different software.

The first step, discussed in Section 2.1.3, is to determine the energies of the initial particles in a collision, which are provided by several different **PDF** sets. These distributions come from experimental measurements, though there is some variation between different sets. Three different sets are used in this analysis: NNPDF2.3LO [45] and NLO CT10 [46] for background and signal processes, and MSTW 2008 [15] for pile-up events, discussed more in Section 3.4.

With the initial states of the constituents of the protons described by these probabilistic models, the next step is to model the hard scattering process resulting from the interaction of two of these particles. This is accomplished by a *generator*, which calculates the cross-sections of the Feynman diagrams of a given process. In particular, these generators typically calculate matrix elements, which describe the probability to go from an initial to final state via a hard scattering, including the kinematic properties of the final state. The generator uses these matrix elements to assign one of these hard scattering final states to each event. These hard scattering outputs are then passed to the next step, where parton showering, hadronization, and final and initial state radiation can occur.

Because these matrix elements must be calculated for each event's specific kinematic properties, it can be very computationally intensive, especially when the calculations are performed at very high or-

der. To save computational time, matrix elements are sometimes calculated at a lower order, and later, the total cross-section for a given process can be calculated at a higher order and used to scale the overall number of events generated for the process. These calculations can also be tuned, varying parameters in the generation to create outputs that most closely match experimental data.

Examples of generators include **SHERPA** [47], **POWHEG Box** [48–50], and **MADGRAPH5_AMC@NLO** [51]. Each has different strengths and is used to describe processes that best match those strengths. **POWHEG Box**, for example, cannot perform its own parton showering, and must be interfaced with another generator, typically **PYTHIA** [52], in order to describe any physics processes beyond the hard scattering, which can cause discontinuities in its predictions for large numbers of partons. However, it can calculate matrix elements at **NLO**, giving it an advantage in calculating some complex processes. **SHERPA** performs its own parton showering, but in most cases calculates its matrix elements at **LO**. The main advantage of **MADGRAPH5_AMC@NLO**, which must also be interfaced with another generator (typically **PYTHIA**) to perform parton showering, is its simple user interface. Instead of designating specific processes to be generated, it allows users to specify a final state to be generated, and all processes capable of producing that state will be included.

Once the final state particles of the hard interaction and showering have been calculated, the pile-up of the **LHC** (described in [Section 3.4](#)) must be accounted for. Events called *minimum bias* are generated to match the overall production of the **LHC** collisions, with no preselection. These simulated events are overlaid on the original hard scatter to produce a more realistic representation of the many simultaneous interactions observed in the **ATLAS** detector.

This collection of particles must then be translated into signals in the detector. Their trajectories in the magnetic fields of the detector, their interactions in each layer, and the way these interactions deposit charge in each subdetector are simulated in software called **GEANT4** [53]. In this software, every piece of the **ATLAS** detector is modeled, including the magnetic field and the many different materials. Particles then follow trajectories through the simulated detector and interact with the different materials based on several preprogrammed options for each material. For example a photon traveling through a material could continue along its trajectory, convert into a positron-electron pair, or deposit energy. As it crosses into a new material, a new set of options opens up for interactions. The particle is tracked until all of its energy is lost or it exits the geometry of the simulation.

The model of the detector used for this process is iteratively perfected by comparing data to **MC**. [Figure 39](#) shows an example of a discrepancy between the simulation and observed data in the num-

ber of secondary vertices in a pixel module, which should correspond to the amount of material in the area. Observations of discrepancies like this can be used to correct the materials in the simulation.

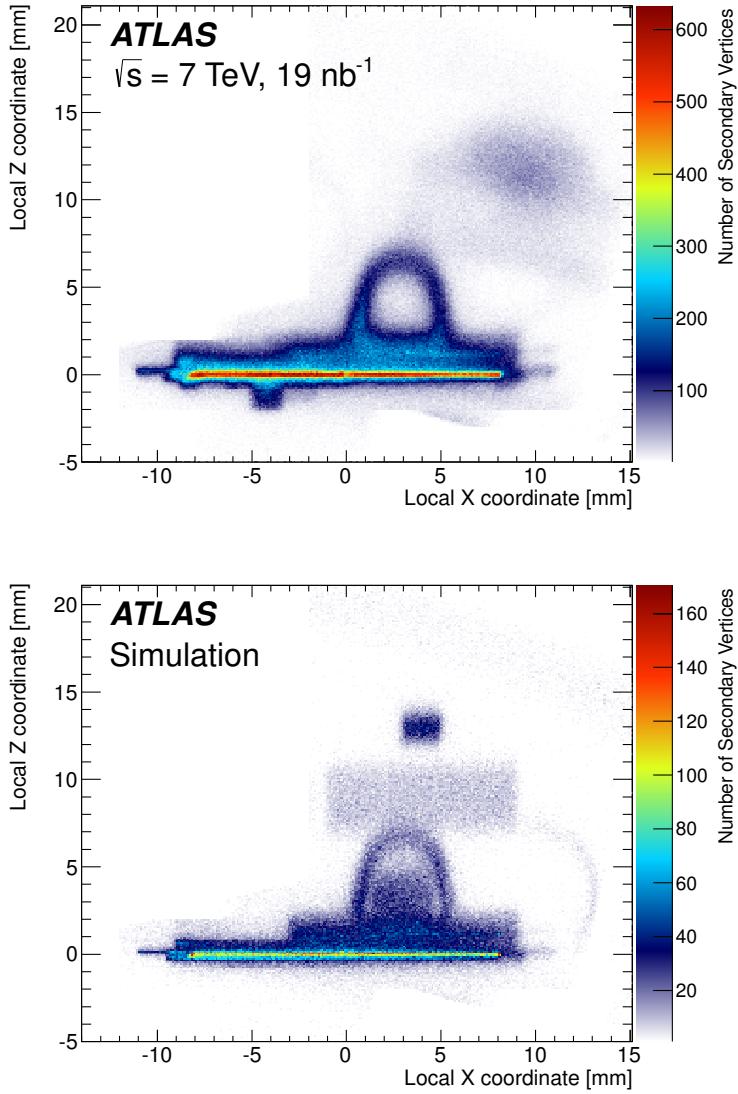


Figure 39: Number of secondary vertices in a module in the first layer of the Pixel Detector in data (top) and MC (bottom). There are more events in the data than the MC [54].

Custom [ATLAS](#) code converts the energy deposited in active sensors into signals that resemble the expected detector response. These responses are typically very complicated with many parameters, and are frequently iterated on to best match the data. Electronic noise must also be added to correctly approximate the operating conditions of the detector. Additional alterations to this signal translation, including dead sensors and misalignments of the detector, can also be added at this stage.

Once the simulated particles have been converted into detector signals, the same reconstruction software used on data can be used on the MC, converting the detector signals back into particle interpretations. This reconstruction process is described in [Chapter 6](#). The original information about the particles from the generator, referred to as *truth* information, is also kept, and can be compared to the reconstruction output to study its efficacy.

INNER DETECTOR TRACKING AND PIXEL CLUSTERING

5.1 OVERVIEW OF TRACKING IN THE ATLAS DETECTOR

Creating tracks from individual hits in the **ID** is one of most challenging parts of the reconstruction of **ATLAS** events. Each event typically contains thousands of hits in the Pixel Detector alone, which must be combined into one coherent picture of which particles traversed the detector, and how they moved and lost energy as they traveled. A typical particle deposits charge in several pixels per layer, forming a series of clusters which can be connected together to form a track. This track can in turn be used to measure the charge, momentum, and trajectory of the particle.

An example of this process can be seen in [Figure 40](#), which shows the tracks formed in an event and the corresponding hits in the Pixel Detector and the **SCT**. In the $r - z$ view, the multiple vertices formed by simultaneous interactions are visible. This event has 17 reconstructed vertices, slightly less than the Run 2 average.

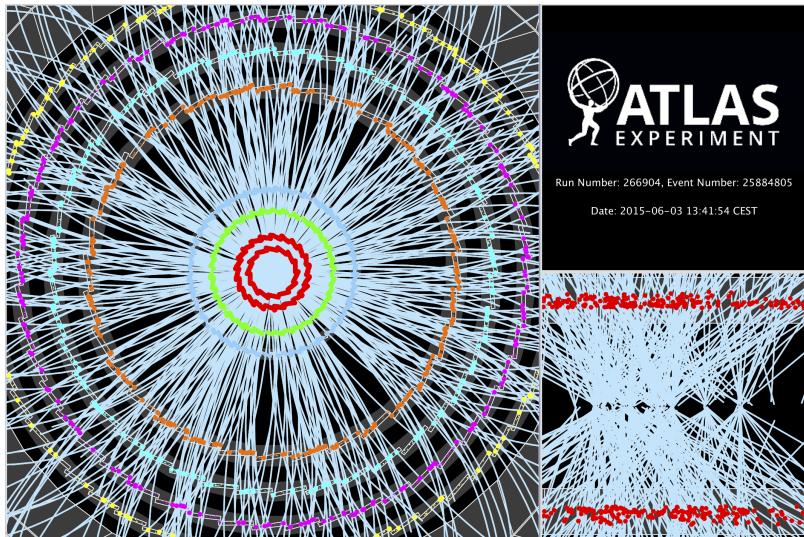


Figure 40: Event display from June 2015, with particle tracks in light blue. The main image displays a view of the $x - y$ plane of the **ID**. The **IBL** and **Lo** of the Pixel Detector are shown in red, with the remaining two layers of the Pixel Detector in green and blue. Outside those are the four double layers of the **SCT**. The smaller image on the right shows an $r - z$ view, zoomed in to only show hits in the **IBL** [55].

In order to used to form tracks, hits from the [ID](#) must first be pre-processed [56]. Nearby Pixel and [SCT](#) hits are turned into clusters, which serve as the input to the tracking system, rather than individual hits. The details of Pixel clustering are discussed in [Section 5.2](#). The [SCT](#) clusters are translated into space-time points using the two measurements from its paired layers. In the [TRT](#), drift times are used to reconstruct the particle’s distance from the center of the tube, forming *drift-circles* that indicate possible particle positions.

Next, seed tracks are formed using the clusters in the Pixel Detector and the first layer of the [SCT](#). The seed tracks are extended to the remainder of the [SCT](#), and are fit. Using the fit quality, any outlier clusters associated with the track are removed.

At this stage, quality cuts are made to remove fake tracks. The tracking algorithim then identifies *holes*, points where the track passes through an active sensor, but no cluster exists. Tracks with too many holes are removed, as are tracks with a high fraction of clusters shared with other tracks.

The remaining tracks are extended to the [TRT](#), where they are associated with drift-circles. The tracks are then re-fit to include all the [ID](#) subdetectors. In the case that this fit is worse than the fit without the [TRT](#), outlier [TRT](#) hits are identified and the track is re-fit without them.

Afterwards, unassociated hits in the [TRT](#) are formed into track segments, which can be extended back to the [SCT](#) and Pixel Detector to form complete tracks. This method is especially useful for identifying tracks that did not originate from the initial hard-scattering interaction.

Lastly, this collection of tracks is used to reconstruct vertices. First, tracks are associated with primary vertices, which must be consistent with the [ATLAS](#) beam spot. Following this, secondary vertices are formed, which can result from long-lived particles or interactions with the detector.

5.2 CLUSTERING IN THE PIXEL DETECTOR

The process of going from clusters to a track is relatively simple in an isolated environment in which one particle travels cleanly through all the layers, but can be complicated by multiple close-by particles, which can produce hits that are merged into one cluster. Clusters can also have misleading shapes due to a particle’s emission of other low-energy particles, called δ -rays [57]. In these cases, it can be hard to tell how many particles were involved in creating a cluster, and where exactly each of those particles passed through the layer. A few examples of particle interactions with the pixel sensor can be seen in [Figure 41](#).

Clusters are initially made by a process called Connected Component Analysis ([CCA](#)). In this process, pixels with with hits above

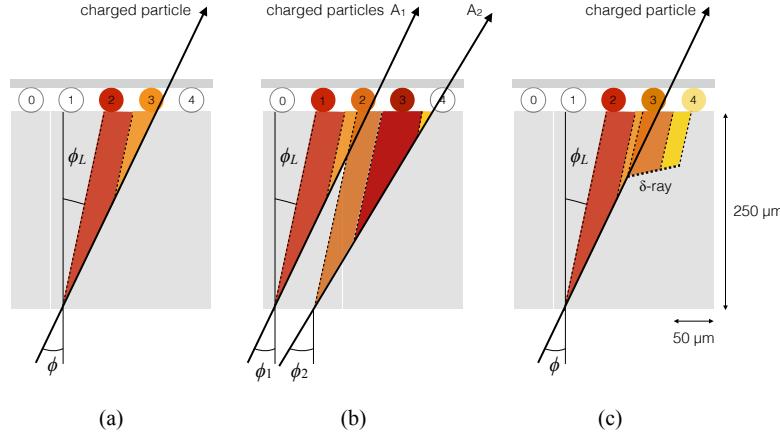


Figure 41: A few possible types of clusters in the Pixel Detector. (a) shows a single particle passing through a layer of the detector, (b) shows two particles passing through the detector, creating a single merged cluster, and (c) shows a single particle emitting a δ -ray as it passes through the detector [57].

a charge threshold are identified in each layer, and are grouped together if they share any edge or corner. The position of the resulting cluster is defined by local x and y coordinates, which are defined relative to the module on which the cluster appears. Determining the position of the particle that formed that cluster is less straightforward, and has recently been updated from a charge interpolation method to a method using a Neural Network (NN).

5.2.1 Charge Interpolation Method

A typical cluster contains a few pixel hits spanning in the x and y directions, each with its own measurement of charge deposition. In the charge interpolation method, these individual hits are combined to make one estimation of the position a single particle which passed through them, using the following equation:

$$x_{cluster} = x_{center} + \Delta_x(\phi, N_{row}) \cdot \left[\Omega_x - \frac{1}{2} \right] \quad (32)$$

$$y_{cluster} = y_{center} + \Delta_y(\phi, N_{col}) \cdot \left[\Omega_y - \frac{1}{2} \right] \quad (33)$$

where $\Omega_{x(y)}$ is defined by

$$\Omega_{x(y)} = \frac{q_N}{q_1 + q_N} \quad (34)$$

and q_1 gives the total charge in the first row (column), and q_N gives the total charge in the last row (column). $\Delta_{x(y)}$ is a function derived

from either data or MC and produces an output related to the projected length of the particles track on the pixel sensor and is measured as a function of ϕ , the incident angle of a particle on the sensor, and $N_{row(col)}$, the number of pixels in the cluster in the x and y direction.

In a simple case, such as (a) of Figure 41, this method works well. However, in cases like (b), it has no ability distinguish two-particle clusters from one-particle clusters, and assigns a cluster center between the two particles' locations, despite the intermediate pixel having the lowest charge. Furthermore, because this method doesn't differentiate two-particle clusters, the tracking software can't use that information to determine if multiple tracks may share two-particle clusters. Allowing tracks to share clusters indiscriminately in dense track environments creates fake tracks from the many possible cluster combinations. In cases like (c), the δ -ray will bias the measurement of the particle's position in whichever direction it is emitted.

5.2.2 Improving Measurement with Neural Networks

To address these problems, a series of NNs was created [57]. The first estimates the number of particles in a given cluster, the second estimates their positions within the cluster, and the third assesses the uncertainty of the position measurement. They are referred to, respectively, as the *Number*, *Position*, and *Error* NNs.

These NNs are taken from the AGILEPack library [58], and trained using simulated ATLAS MC. Each NN is given the following inputs:

- a 7×7 grid of cluster ToT information¹
- a 7-element vector containing the y -size of the pixels in the grid²
- the layer of the Pixel Detector that the cluster was observed in
- a variable indicating whether the cluster is located in the barrel or endcap
- θ and ϕ variables projecting the incident angles of the particle on the sensor³
- the pixel module's η index, a label assigned to each module that differentiates modules based on their η position

¹ Clusters spanning more than seven pixels in either direction are rare, but when they occur they are rejected, and the original charge interpolation estimate of a single particle's position is kept.

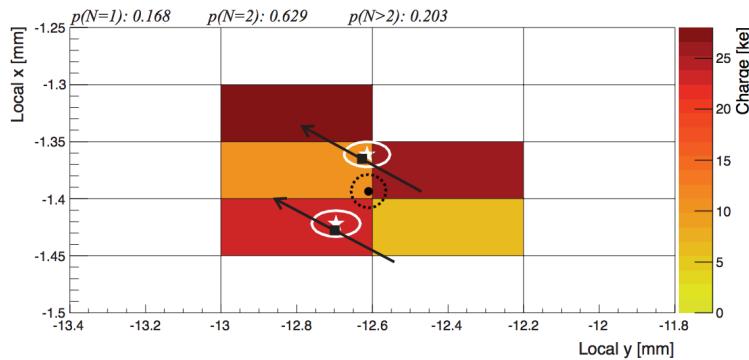
² The Pixel Detector contains some long pixels at the edges of modules, and this is intended to help the NN identify these cases.

³ If the NN is applied before tracking is performed, these angles project to the nominal interaction point, and if tracking has already been performed, the angles are taken from the track fit to the cluster.

Each **NN** is trained with truth information to make predictions about cluster features. The Number **NN** predicts the number of particles associated with the cluster, required to be between 1 and 3. Then, the same inputs are fed into one of three Position **NNs**, each identified by the number of particles it attempts to locate, according to the prediction from the Number **NN**. The Position **NN** outputs a prediction of the x and y positions of each particle, and these predictions are used as inputs to the Error **NN**, in addition to the inputs listed above.

The training of the Error **NN** is slightly more complicated than the others, as it makes a prediction of the efficacy of the previous **NN**. It is trained to predict the difference between the Position **NN**'s outputs and the true positions of the particles, which allows it to make case-by-case predictions of how much inaccuracy is likely in a given prediction.

An example of the output of this process can be seen in [Figure 42](#), where the improved position resolution from the ability to identify a multi-particle cluster is evident.

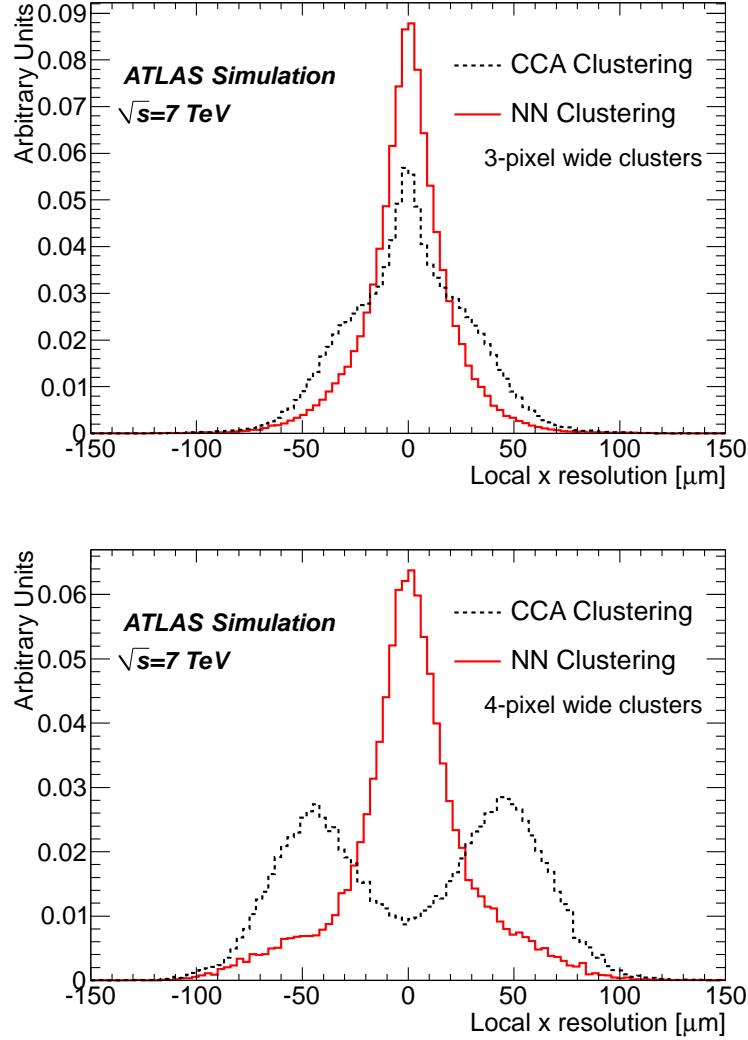


[Figure 42](#): One example of a two-particle cluster and its truth information compared with the output of the **NNs**. The boxes represent pixels, with a color scale indicating charge. At top, the $p(N = i)$ values give the output of the Number **NN**, the probabilities that the cluster contains 1, 2, and 3 particles. Given the highest probability is for $N = 2$, the other **NNs** predict the postion and errors of the two particles (in white). The black arrows and squares represent the truth information from the cluster, and the black dot and dotted line show the position measurement for the un-split cluster [\[57\]](#).

The particle location predictions from the **NNs** are then handed to the tracking software, which now can use these multiple particle position estimations as independent hits to be fit. As a result, tracks in dense environments have fewer clusters shared between multiple tracks, and their trajectories are known to a greater degree of precision.

5.3 IMPACT OF THE NEURAL NETWORK

The [NN](#) was first applied to 7 TeV data, where it improved position resolution for particles in small and large clusters. [Figure 43](#) shows the improvement from the addition of the [NN](#) in x resolution in different cluster sizes. The improvement from charge interpolation clustering is particularly evident in the 4-pixel case, where the double peaked structure of the interpolation method has been completely removed with the [NN](#).



[Figure 43](#): x resolutions for clusters with 3 (top) and 4 (bottom) pixels in the x direction in 7 TeV data for [CCA](#) (using only charge interpolation to determine position) and [NN](#) clustering taken from [MC](#) [57].

5.3.1 *The Neural Network in 13 TeV Data*

In Run 2, the tracking algorithm is first run on the **CCA** clusters with positions determined via charge interpolation, where it constructs tracks with loose quality requirements. In this step, the tracking algorithm allows tracks with shared clusters [59]. The **NN** is then used to identify which clusters are likely to have had multiple particles pass through them, and to estimate the positions of those particles. In the case that the cluster is determined to have resulted only from one particle, tracks that share that cluster are penalized. In general, tracks with more than two shared clusters are rejected.

Because the **NN** is trained only with **MC** simulations, any mismodeling of the way charge is deposited in the **ATLAS** detector could cause the **NN** to perform in an unexpected way when applied to data. The potential impact of this mismodeling was investigated with 13 TeV **MC** [60]. The goal of these studies was to determine which variables the **NN**'s predictions were most sensitive to, and whether it was likely that these variables could be mismodeled enough to produce unexpected results in data.

One example of a variable capable of significantly altering the **NN** outputs was the overall charge scale. To study its impact, the **ToT** of all pixels in a cluster were scaled up and down, and the resulting outputs of the **NN** were compared, as shown in [Figure 44](#). In this case, the likelihood to misidentify multi-particle clusters and single-particle clusters depended significantly on this scaling. However, this scaling is unlikely to be mismodeled by more than 10%, so very extreme effects from a difference between data and **MC** are unlikely. Overall, it was found that variations on the cluster charge produced a significant impact on predictions, while all other variations, such as incidence angle variation and spatial smearing of charge, had a minimal effect.

In addition to studies on the impact of alterations of individual simulation variables, studies directly comparing the **NN** output in data and **MC** were performed. [Figure 45](#) shows a comparison of how often the **NN** identifies different types of clusters in data and **MC**. Each figure is made using by selecting pairs of collimated tracks that share a common cluster on a given layer, then calculating the fraction of those clusters that are determined by the **NN** to be single or multi-particle clusters. This fraction is plotted as a function of the distance between the two tracks in the cluster's layer. Very good agreement is seen between the two samples, demonstrating that the **MC**-trained **NN** performs similarly on both **MC** and data.

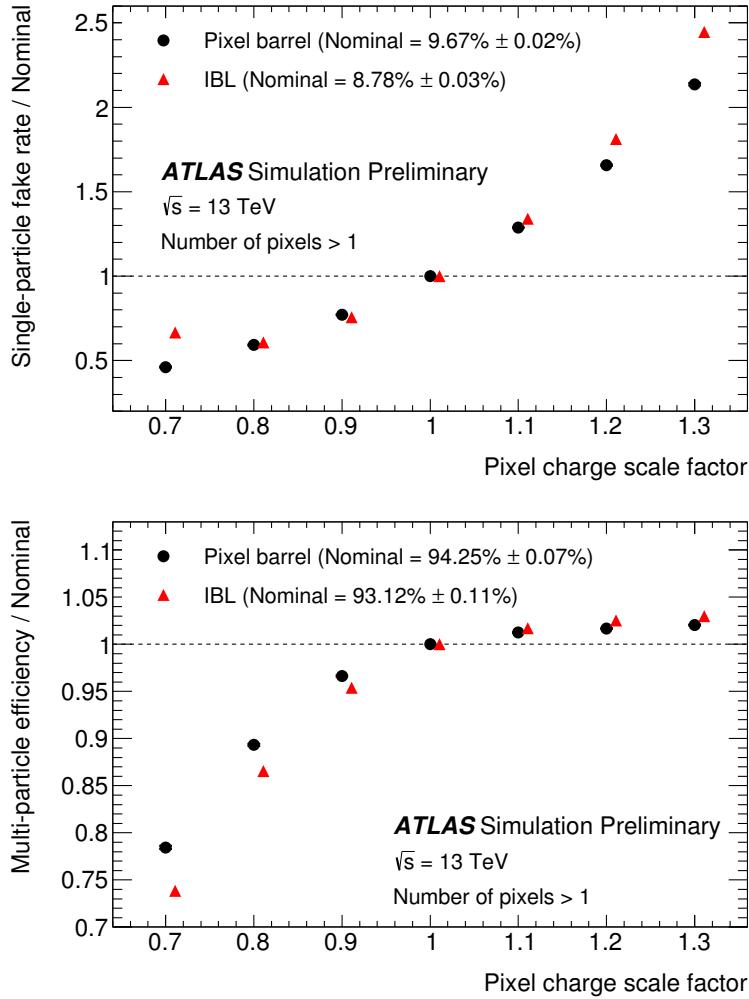


Figure 44: Performance of the pixel neural network used to identify clusters created by multiple charged particles, as a function of constant coherent scaling of the charge in each pixel in the cluster. The top figure shows the rate at which the neural network wrongly identifies clusters with one generated particle as clusters with multiple particles. The bottom figure shows the rate at which the neural network correctly identifies clusters generated by multiple particles as such.

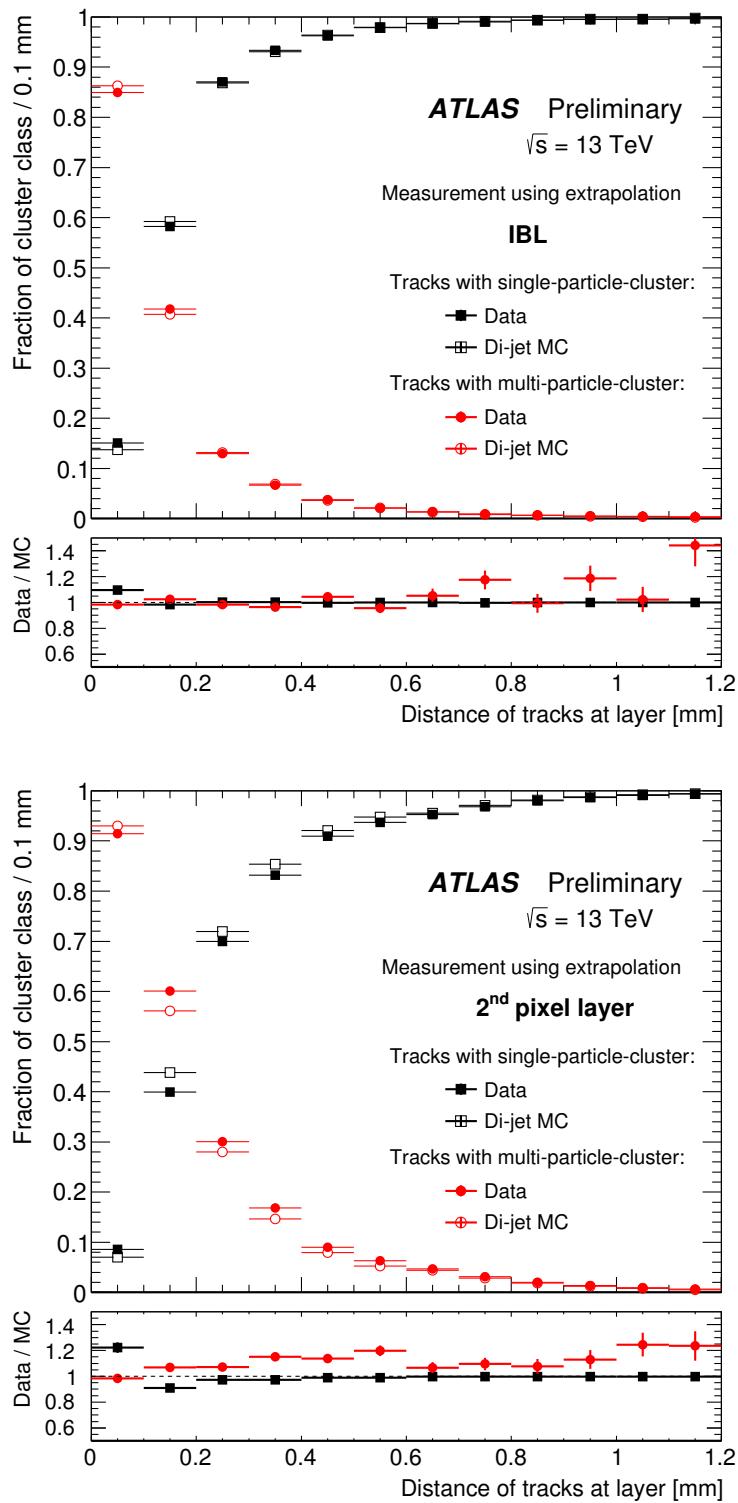


Figure 45: Fraction of cluster classes as a function of the distance between tracks for IBL (top) and 2nd pixel layer (bottom).

OBJECT RECONSTRUCTION IN THE ATLAS DETECTOR

Object reconstruction is the computationally intensive process of interpreting the signals from the approximately 100 million read-out channels of the [ATLAS](#) detector into a collection of particles and jets, the objects with which physics analysis can be performed. This process is complicated, and requires dedicated working groups in the [ATLAS](#) experiment that optimize the understanding of each type of object. These groups must all collaborate to provide a full picture of the events in the detector. For each object type, candidate objects are reconstructed, and then an identification step is performed, which chooses which candidates will be used at the analysis level, based on a series of quality requirements.

6.1 ELECTRONS

Electrons are reconstructed through a combination of [ID](#) and calorimeter measurements. They travel through the tracking system, leaving charge deposits in each layer, then are absorbed by the electromagnetic calorimeter. These two measurements work in conjunction to deliver high resolution measurements of electron momentum from low- p_T , where track curvature gives the most reliable measure of the electron's energy, to high- p_T , where the tracks are almost perfectly straight, but the calorimeter can still provide a reliable measurement.

In the central region ($|\eta| < 2.47$) of the [ATLAS](#) detector, electron reconstruction begins with the identification of energy deposits in the electromagnetic calorimeter. The clusters of calorimeter cells are seeded by sliding longitudinal windows, which are measured in units of 0.025 in η and ϕ . 3×5 unit windows are used, which require at least 2.5 GeV in the window to form a seed [61].

These clusters are matched to [ID](#) tracks by extrapolating each track to the middle layer of the calorimeter and identifying nearby clusters. If there are multiple tracks associated with a given cluster, tracks with silicon hits are preferentially chosen, and then the track with the smallest ΔR to the center of the cluster is selected. If a matching track is found, it is used to determine the likely direction of bremsstrahlung radiation in the calorimeter, and maximum distance to match a track to a cluster is expanded in the ϕ direction to account for this radiation. If no track is found, the cluster is rejected.

The calorimeter clusters are then rebuilt in larger windows, 3×7 in the barrel and 5×5 in the endcaps. An estimate of the energy is

made by summing the measured calorimeter energy with estimates of the energy lost before the electron reached the calorimeter, energy outside of the cluster window, and energy not fully deposited in the calorimeter. These estimates are made with parametrized functions determined from a combination of MC and measurements of energy loss determined with the presampler.

The p_T of a central electron is determined through a combination of the calorimeter energy measurement and track measurements of the electron, while its η and ϕ are taken from the track at its vertex.

In the forward region, where no tracking is available, electron energy is determined more roughly. Calorimeter cells are formed into variable-sized clusters in regions of significant energy deposition, and the center of the cluster is used to determine angular coordinates of the electron. However, because these electrons have worse resolution in both their position and energy, they are rejected in this analysis.

These reconstructed electron candidates' quality are then assessed based on an algorithm that uses multivariate analysis to assign a likelihood that a candidate is a true electron based on input from just under twenty different variables. These include track quality, hadronic leakage, cluster shape, and transition radiation, incorporating information from as many subdetectors as possible in its determination of the candidate's quality. Each variable is assigned a probability distribution function for true electrons and background processes, and they are collectively used to provide a *likelihood* value which can be cut on.

Three levels of identification, Loose, Medium, and Tight, are defined with different likelihood cuts, with electron candidates passing tighter identification levels always a subset of looser electrons. Figure 46 gives the efficiencies at each of these working points both for true electrons and for hadrons, which can be misidentified as electrons. Tighter working points have worse efficiencies, but lower misidentification rates for hadrons as well as photons.

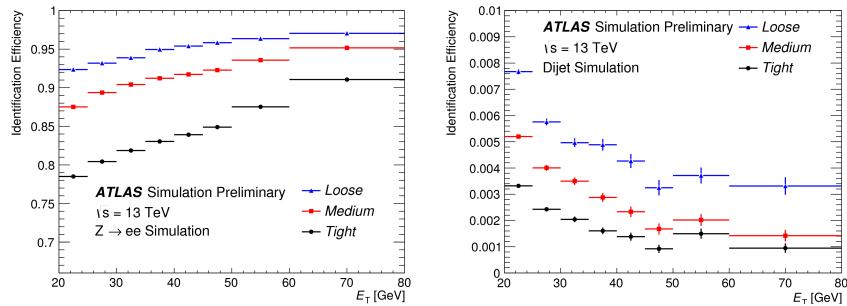
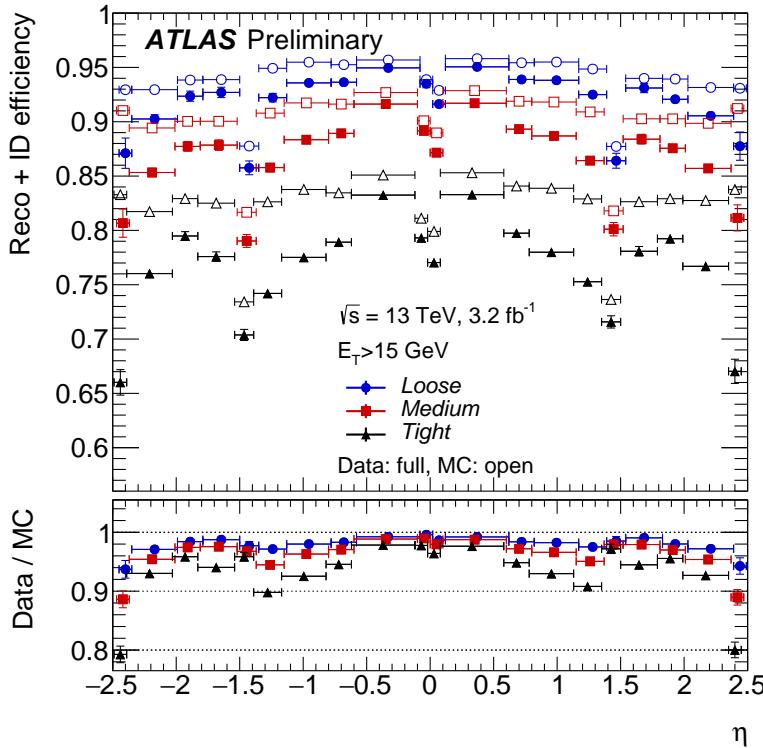


Figure 46: Identification efficiencies from MC samples for Loose, Medium, and Tight working points. Left is the efficiency for identification of true electrons taken from $Z \rightarrow ee$ MC, and right is the efficiency for mis-identification of jets as electrons taken from dijet MC [62].

[MC](#) efficiencies can be compared to efficiencies measured in data to obtain a correction factor, which applied to [MC](#) to better emulate the rates at which electrons are reconstructed and identified in data. [Figure 47](#) shows a comparison of the combined reconstruction and identification efficiencies in data and [MC](#), with the resulting correction factors also displayed as the ratio. This analysis uses the Medium working point, which has correction factors ranging between 2 and 10%.



[Figure 47](#): Combined electron reconstruction and identification efficiencies measured as a function of η for data (using the tag-and-probe method on $Z \rightarrow ee$ events) and $Z \rightarrow ee$ [MC](#). Distributions include electrons with $E_T > 15$ GeV. [62].

Requirements are also made on electron *isolation*, which quantifies the amount of energy deposited near the electron according calorimeter and track measurements. Isolation variables are primarily used to reject non-prompt leptons, leptons which aren't produced by the initial hard scattering of the pp collision. These can be produced by heavy flavor hadron decays and converted photons, as well as misidentified hadrons. Cuts are made on the amount of nearby calorimetric energy and sum of the p_T of any nearby tracks relative to the electron's energy, forming a series of working points. Working points are created based on their efficiency, including Tight and Loose working points, which operate at 95 and 98% efficiency respectively. The most effective working points target different efficiencies as a func-

tion of p_T , with higher efficiencies possible at high p_T due to reduced fake backgrounds. There are two such working points, `Gradient` and `GradientLoose`. They each have a 99% efficiency for electrons with $p_T > 60$ GeV, but 90 and 95% efficiencies at 25 GeV. To recover the largest possible fraction of electrons, this analysis uses `GradientLoose`.

6.2 PHOTONS

The reconstruction of photons is performed in parallel to electron reconstruction. Seed clustering is performed, and tracks are matched to these clusters, as in the case of the electron reconstruction described in [Section 6.1](#).

Photons can be converted to electron-positron pairs in the `ID`, leaving a pair of tracks, or they can pass through without conversion, leaving no tracks behind. As a consequence, calorimeter clusters resulting from photons can have no tracks associated with them, two tracks, or one track, in the case that one of the conversion tracks is not reconstructed. The reconstruction software attempts to identify all these scenarios and differentiate these clusters from electron and hadron deposits [63].

Two-track clusters are required to consist of two oppositely charged tracks that emerge from a conversion vertex running parallel to one another. A likelihood that these tracks are from electrons is determined using the high threshold hits in the `TRT`, and quality requirements are made on the tracks using this likelihood. For tracks with silicon hits, a loose likelihood requirement of 10% is made, while tracks without silicon hits are required to have at least 80% likelihood. The tracks are then fit to determine the conversion vertex, and quality cuts are made, such as requiring that conversion vertices within the silicon volume correspond to tracks with silicon hits.

Single track clusters occur most often from conversions in the outermost layers of the `ID`, and are more difficult to reconstruct. Tracks are typically lost because an electron or positron resulting from the conversion has a p_T too low to be reconstructed, or because the two tracks are so close together that they're identified as a single track. The single track is required to have at least a 95% electron likelihood from `TRT` hits, and must not have a hit in the innermost layer of the Pixel Detector. The conversion vertex is defined as the first hit of the single track.

The tracks associated with these conversion vertices are extrapolated to the calorimeter and matched to cluster, except in the case that there are two tracks that differ substantially in their p_T measurements, in which case the position of the conversion vertex is used for extrapolation to the calorimeter, assuming a straight-line trajectory. If multiple vertices are matched to a single cluster, preference is given

to vertices with double tracks, silicon hits, and finally to tracks closest to the interaction point.

Any cluster with neither a conversion vertex or a track associated with it is identified as an unconverted photon. Clusters associated with both electrons and photons are assigned to one or the other based on their properties. Clusters are preferentially identified as photons in the case that they are matched to a conversion vertex in which at least one track is associated with both the vertex and the cluster, or if the associated tracks have a p_T smaller than the cluster's p_T . E/p , the ratio of the cluster and track energy measurements, can also be used to differentiate electrons and photons. Electron candidates are instead reconstructed as photons if they have $E/p > 10$ or if the track matched to the electron has p_T below 2 GeV.

Photon energy is determined in a 3×5 (3×7) window for unconverted (converted) photons in the barrel, where the window is expanded to compensate for the increased spread of energy from the conversion products. In the endcap, the 5×5 window is used in all cases. Like the electrons, the calibration of the photon's energy accounts for energy loss before the calorimeter, as well as energy deposited outside the cell and beyond the electromagnetic calorimeter.

Photon identification is performed in the range $|\eta| < 2.37$ using a series of cuts on the shape of the shower in the electromagnetic calorimeter, as well as the amount of additional energy deposited in the hadronic calorimeter. Photons in the so called *crack* region of the calorimeter ($1.37 < |\eta| < 1.52$), where a discontinuity prevents accurate assessment of photon energy, are rejected. The photon identification has only one working point, called *Tight*, which has an identification efficiency of 53–64% (47–61%) for unconverted (converted) photons with $E_T = 10$ GeV and 88–92% (96–98%) for photons with $E_T \geq 100$ GeV [64]. Efficiencies as a function of p_T measured in the 2016 data and compared to MC can be seen in Figure 48.

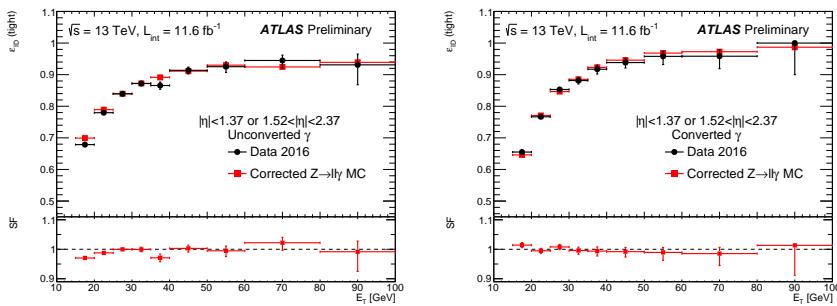


Figure 48: Comparison of *Tight* identification efficiency measurements from data and $Z \rightarrow \ell\ell\gamma$ MC for unconverted (left) and converted (right) photons, with an inclusive η selection. The bottom of each figure shows the ratio of data and MC efficiencies. [65].

Photon isolation, like electron isolation, can be determined as the combination of nearby calorimeter deposits and tracks. Fixed cuts on

the isolation as a fraction of photon energy is typically used. A working point called `FixedCutTight` reconstructs the amount of calorimeter energy (excluding that of the photon) in a cone of $\Delta R = 0.4$ around the photon and the amount of energy from the sum of track p_T in a cone of $\Delta R = 0.2$, including only tracks associated with the primary vertex. Defined relative to the photon's p_T , this working point includes photons with calorimetric isolation less than $0.022 p_T + 2.45$ GeV and track isolation less than $0.05 p_T$ [66].

6.3 MUONS

Muon reconstruction is performed independently in the `ID` and the `MS`, then the two measurements are combined when consistent tracks are found in each system [67]. The `ID` reconstruction is performed using the tracking algorithms described in [Section 5.1](#), and includes tracks with $|\eta| < 2.5$.

The `MS` track reconstruction is performed in the $|\eta| < 2.7$ range and begins with a search in each muon chamber for patterns of hits consistent with a track, called *segments*. The `MDT` chamber hits are fit to a straight line, and nearby `RPC` and `TGC` chambers provide the coordinate orthogonal to the magnetic curvature for these hits. Segments are also built in the `CSC`, where they are required to be loosely consistent with a track originating from the interaction point.

These segments are then fit together, starting from the middle layers of the `MS`, with track quality requirements on the resulting combinations based on the χ^2 of the fits. Tracks must have at least two segments, except in the transition region between the barrel and endcap, where a single segment can qualify as a track. Segments are allowed to be shared between multiple tracks in the initial reconstruction, but after the combination, tracks with shared segments and poor χ^2 are removed.

These `MS` tracks are then combined with measurements from the `ID` and calorimeters. The best quality muons are combined muons, which have `ID` and `MS` tracks associated to them, the hits of which are re-fit to form a combined track. `MS` hits can be added or removed at this stage based on their consistency with the new track. Lower quality muon candidates are also defined. Extrapolated muons have only `MS` tracks and their trajectories are required to be consistent with the interaction point. Calorimeter-tagged muons combine an `ID` track with a calorimeter deposit consistent with a muon, while segment-tagged muons combine an `ID` track with a segment in the `MS`. Muons with shared `ID` tracks are not allowed, with preference given to combined muons, then calorimeter-tagged muons, and lastly segment-tagged muons.

There are four muon identification working points for muons: `Loose`, `Medium`, `Tight`, and `High- p_T` . These working points all have different

efficiencies for the identification of muons, balanced against the mis-identification of hadrons. One of the key variables for their discrimination is q/p significance, which quantifies the consistency between the **ID** and **MS** measurements of momentum. The χ^2 of the combined fit is also an important discriminator.

The **Loose**, **Medium**, and **Tight** selections are inclusive, with all **Tight** muons passing the **Medium** requirements, and **Medium** muons passing the **Loose** requirements. The **Medium** working point includes only combined and extrapolated muons, and is the default for most **ATLAS** analyses, including this one. Extrapolated muons are allowed only outside the **ID** tracking system ($|\eta| > 2.5$) for this working point, but this region is excluded by this analysis because of the decreased efficiency and larger p_T resolution of these muons. As a consequence, this analysis uses only combined muons. For these muons, the **Medium** working point requires at least three hits in at least two **MDT** layers (except in the $\eta < 0.1$ region) and a q/p significance cut is made to reduce backgrounds. Due to the lack of coverage at low η , there is a drop in efficiency in this region, as shown in Figure 49.

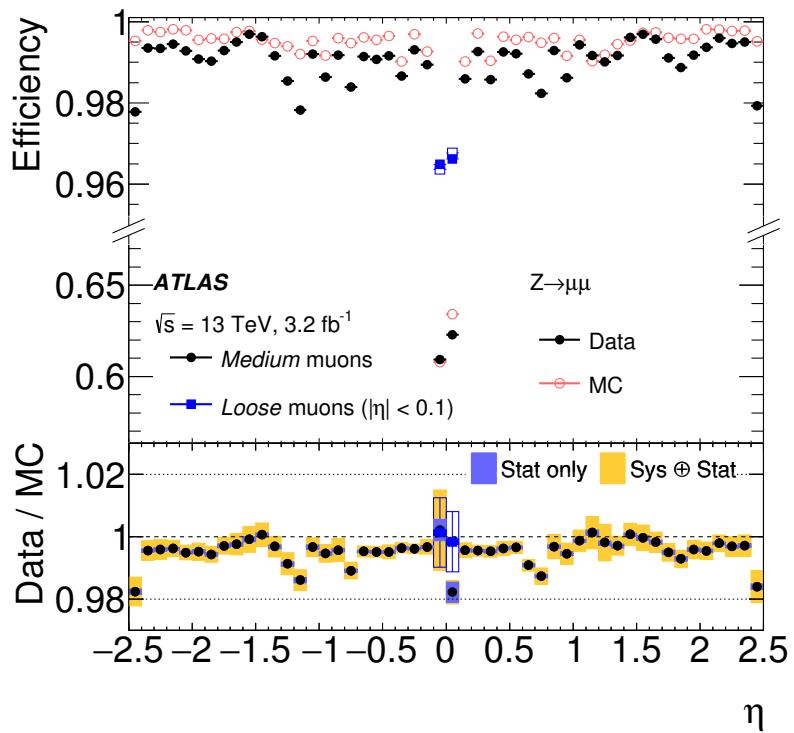


Figure 49: Muon reconstruction efficiency for the **Medium** and **Loose** working points measured with $Z \rightarrow \mu\mu$ events in data and in MC as a function of η . The ratio between the two is shown at the bottom. The **Loose** working point efficiency is shown only at small $|\eta|$, where the loosened requirements cause the largest difference from the **Medium** working point [67].

The High- p_T working point is designed to minimize the resolution for high- p_T muons, at the cost of lower efficiencies. Muons passing the High- p_T requirements must have at least three **MDT** hits in three layers, which decreases efficiency but gives greatly improved p_T resolution. In addition, some regions of the **MS** with poor alignment are vetoed to cut down on mismeasurement. Compared to the default working point these muons have much lower efficiency: 78% (90%) for High- p_T muons compared to 96% (96%) for Medium in the p_T range of 4-20 GeV (20-100 GeV). The efficiency as a function of η for this working point can be seen in Figure 50, where the efficiency loss due to the vetoing of some chambers is especially apparent. Mismodeling of the alignment and the specificity of the momentum resolution cuts cause a large discrepancy between data and **MC** efficiencies, resulting in scale factors that differ from unity by as much as 10%. This working point was considered for this analysis, where mismeasurement of muons increases **SM** backgrounds, but ultimately the Medium working point was chosen for its superior efficiency and better modeling in **MC**.

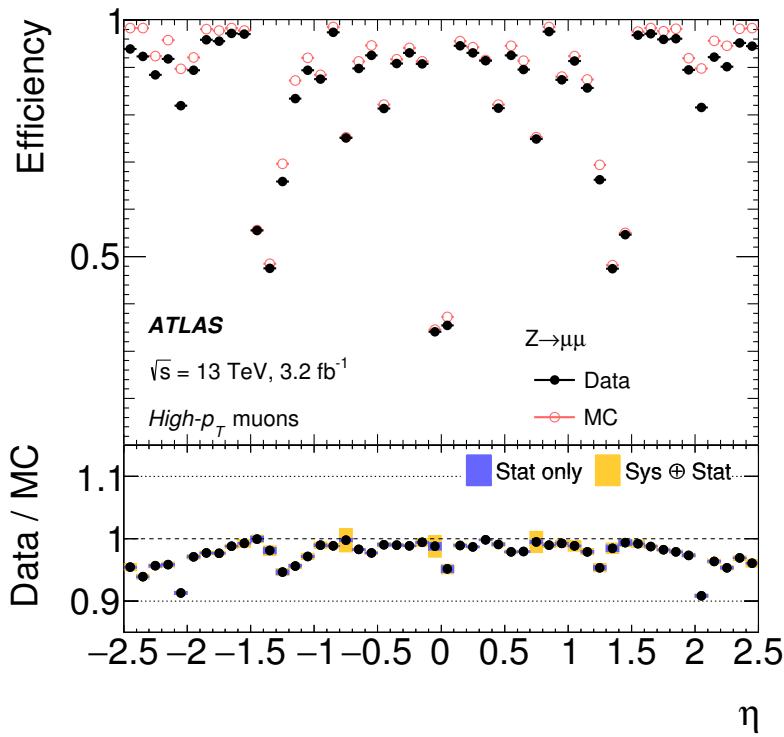


Figure 50: Muon reconstruction efficiency for the High- p_T working point measured with $Z \rightarrow \mu\mu$ events in data and in **MC** as a function of η . The ratio between the two is shown at the bottom. [67]

The isolation selection for muons is designed in the same way as the electron isolation, and also called **GradientLoose**. This working point makes cuts on a combination of nearby calorimeter- and track-

based energy measurements, with an increasing efficiency as a function of p_T . The GradientLoose working point is constructed such that muons with p_T of 25 GeV have an efficiency of 95%, and muons with p_T of 60 GeV have an efficiency of 99%.

6.4 JETS

Jets are the most complicated objects to reconstruct in the [ATLAS](#) detector because each jet is an assembly of many hadronic particles. In contrast to a lepton, whose reconstructed energy can easily be compared to its true energy from simulation, even a jet's true energy is ambiguous, and is dependent on the choice of the jet's definition. The standard jet reconstruction algorithm used in the [ATLAS](#) experiment is called anti- k_t [68].

This algorithm begins with clusters in the calorimeter defined by topologically connected cells with energy deposits significantly higher than the noise background. There are two collections used most commonly for analysis. One uses cluster energies calibrated for electromagnetic showers ([EM](#)), and another uses clusters calibrated to the energies of hadronic showers. The second uses a method called Local Cluster Weighting ([LCW](#)), which first determines the extent to which the cluster is electromagnetic or hadronic based on the energy density and the shower depth, then applies a calibration accordingly for each cluster.

To reconstruct jets, a set of clusters is chosen and the anti- k_t algorithm is then applied. These clusters are grouped together according to the distance measure

$$d_{ij} = \min(k_{ti}^{-2}, k_{tj}^{-2}) \frac{\Delta_{ij}^2}{R^2} \quad (35)$$

where R is the algorithm's radius parameter, typically set to 0.4, Δ gives the angular separation of the two clusters, and k_t is the transverse momentum associated with the cluster.

The grouping process begins with each cluster as a *pseudo-jet*, with its axis and p_T is determined as if it were a typical jet. Then, the pair of pseudo-jets with the smallest d_{ij} are grouped together, forming a new pseudo-jet, and its axis and p_T are reassessed. This grouping continues until there is a pseudo-jet with p_T smaller than the d_{ij} of any pseudo-jet pair, at which point this pseudo-jet becomes a jet, and is removed from the collection. The clustering process continues until all clusters are associated with a jet.

The inverse dependence on the k_t of the cluster produces jets with energetic cores and softer edges, which matches the expectation from a hadronic shower. In addition it is infrared and collinear safe, with neither soft emission nor collinear particles altering the reconstruction of the jet.

A series of calibrations are then applied to these jets. The first is to correct for additional hadronic energy due to pile-up. Figure 51 demonstrates the impact of pile-up on the energy density of an event. The energy density of each jet is defined as the jet's p_T divided by the its area, and the overall event's energy density is defined as the median value of this quantity for jets with $p_T > 20$ GeV. In events with high numbers of primary vertices, the resulting high energy density can affect the amount of stray energy associated with reconstructed jets. To remove the bias on jet energy measurements that results from multiple primary vertices, a correction factor is determined using MC. It is parametrized in terms p_T , η , and the number of primary vertices in the event, as well as the average number of interactions per event in the event's luminosity block, which makes correction for out-of-time pile-up possible. Next, jets are corrected to have their origin at the primary vertex instead of the center of the ATLAS detector. After that, the jets are corrected based on η dependent Jet Energy Scale (JES) factors derived from data and MC independently. Figure 52 shows the energy response, the inverse of these factors, for EM jets. Lastly, an observed bias in the η measurement of jets is accounted for.

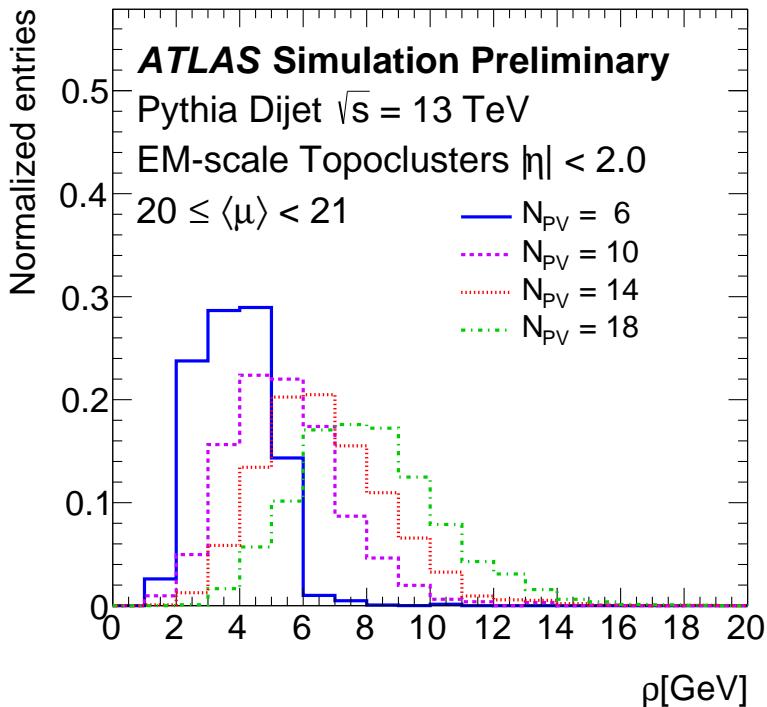


Figure 51: Distribution of event p_T density, ρ , taken from MC dijets for different numbers of primary vertices. [69]

In addition to correcting for additional energy due to pile-up, it is necessary to reject reconstructed jets that come from pile-up vertices. To accomplish this, a multivariate algorithm called JVT was created which builds upon an older method, JVFF [70].

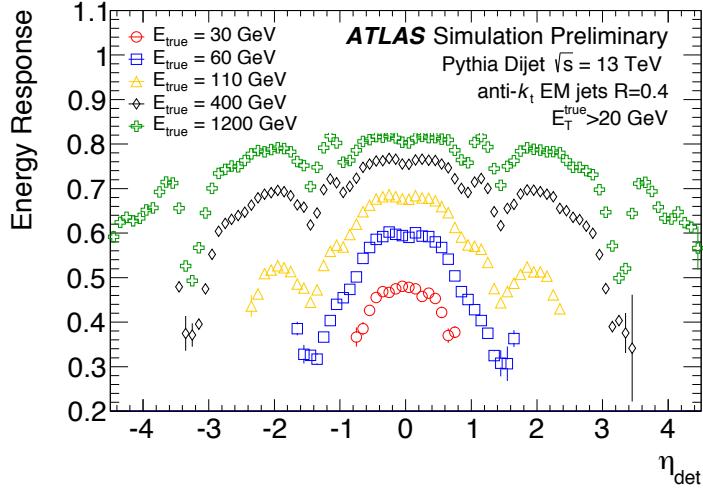


Figure 52: Energy response as a function of energy and η for EM jets in dijet MC. [69]

JVF gives the fraction of energy in a jet that comes from the hard-scatter vertex, and is defined as

$$\text{JVF} = \frac{\sum_i p_T^i(PV_0)}{\sum_j p_T^j(PV_0) + \sum_{n \geq 1} \sum_j p_T^j(PV_n)} \quad (36)$$

where $p_T^i(PV_n)$ gives the p_T of the i th track associated with the n th primary vertex. PV_0 gives the primary vertex associated with the hard-scattering, while the remaining vertices are due to pile-up interactions. Tracks are associated with a jet according to a process called *ghost association*, in which they are clustered along with the typical pseudo-jet collection according to the anti- k_t algorithm described above. In the clustering process, the tracks energy is ignored so that it doesn't impact the measurement of the final jet.

This fraction decreases with higher pile-up, making the construction of an explicit cut difficult in varying pile-up conditions. **JVT** improved on the method by using a pile-up corrected **JVF**-like variable, defined as

$$\text{corrJVF} = \frac{\sum_i p_T^i(PV_0)}{\sum_j p_T^j(PV_0) + \frac{\sum_{n \geq 1} \sum_j p_T^j(PV_n)}{kn_{PU}}} \quad (37)$$

where n_{PU} is the number of tracks, which is multiplied by a scaling factor $i = 0.01$. This quantity is included in the inputs of the tagger along with other variables measuring the fraction of jet energy that is associated with the hard-scattering vertex. Figure 53 shows the efficiency and fake rate for the two methods, demonstrating **JVT**'s

superior stability across events with different numbers of pile-up vertices.

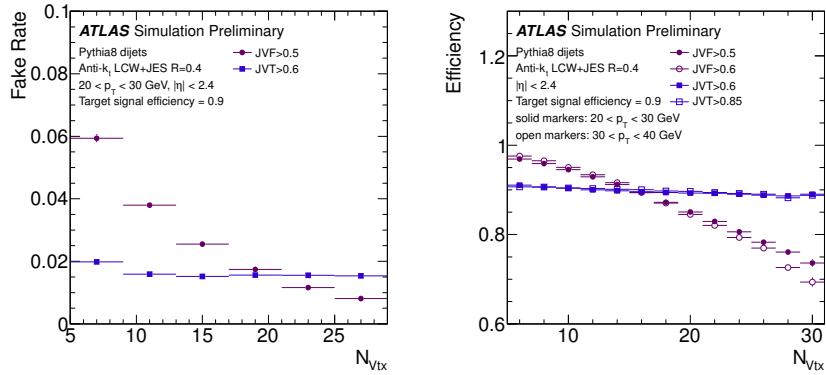


Figure 53: Dijet MC distributions of the number of pile-up jets passing the JVT and JVF cuts (left) and the efficiency for jets from the primary vertex (right) as a function of number of primary vertices in the event [70].

It is possible to differentiate jets resulting from b -hadron decays from other jets due to the non-negligible lifetimes of the hadrons. Many BSM processes preferentially produce b quarks, as do any processes involving top quarks, so this identification can be useful for any analyses seeking to isolate these instances. Multivariate techniques are used to identify secondary vertices using the ID [71]. In ATLAS, separate algorithms are used to identify jets with tracks with significantly non-zero impact parameters, tracks that reconstruct a secondary vertex, and tracks that can be identified with a chain of vertices beginning with the primary vertex. This information is fed into a boosted decision tree, a type of multivariate algorithm, called MV2c20, which outputs a discriminant shown in Figure 54. Using this discriminant, a working point is chosen such that b -jets can be identified with a 70% efficiency, with mis-identification rates at around 12% for c -jets and 0.2% for light-flavor jets.

6.5 OVERLAP REMOVAL

Because most of these reconstruction methods are run independently, it is common for energy deposits and tracks to be shared between jets and particles of different types. To account for this, a process called Overlap Removal (OR) is used, which iteratively removes overlapping objects. The process, as well as the calculation of missing transverse momentum described in Section 6.6, is performed on *baseline* objects. These objects have looser selections than the final *signal* objects, and the separate definitions allow analyzers to tune signal objects to best match a BSM signature, while leaving the OR process unchanged. The

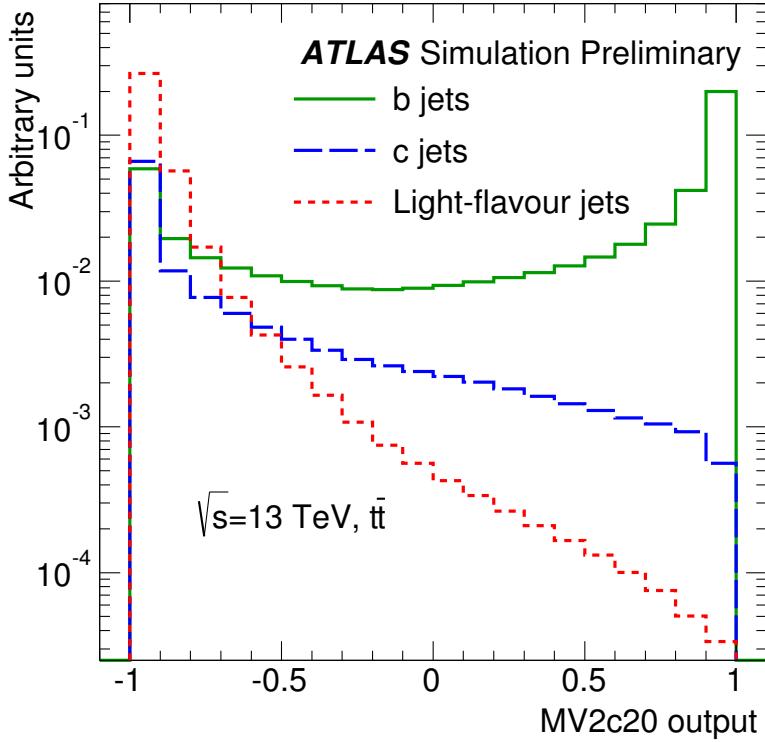


Figure 54: Distribution of MV2c20 output for b -jets, c -jets, and light-flavor jets in $t\bar{t}$ MC [71].

signal and baseline definitions for this analysis are described in Chapter 8.

The first step in the OR process is to remove reconstructed jets that appear to be due to calorimetric deposits from an electron. To accomplish this, any baseline jet within $\Delta R = 0.2$ from a baseline electron is removed. A caveat is added due to the frequent production of leptons in the decay of heavy-flavor jets; if the jet is b -tagged, the electron will be removed instead. After these electrons and jets have been removed, a new search is done for jets and electrons within $\Delta R = 0.4$ of one another. In this iteration, the electron is removed, again to reduce backgrounds from heavy-flavor decays.

Next, the muon-jet OR is applied, which is very similar to that of the electron. Any jet within $\Delta R = 0.2$ of a muon is removed, unless the jet is b -tagged, in which case the muon is removed due to the likelihood that it resulted due to a heavy-flavor decay. The muon-jet OR then differs from the electron's in that a p_T -based ΔR cut is used in the last step. Muons within $\Delta R < \min(0.04 + (10 \text{ GeV})/p_T, 0.4)$ of a jet are removed, with the p_T -dependent cone size designed to reject low- p_T heavy-flavor muons while preserving muons resulting from the decay of high- p_T particles, which are closely aligned with the other products of the decay.

The next step is to remove electrons likely to have resulted from muon bremsstrahlung. Any electron within $\Delta R = 0.1$ of a muon is removed from the event.

Lastly, overlap between photons and both jets and electrons is considered. Baseline photons within $\Delta R = 0.4$ of an electron are removed, as are jets within $\Delta R = 0.4$ of a remaining photon.

6.6 MISSING TRANSVERSE MOMENTUM

Missing transverse momentum ($\mathbf{p}_T^{\text{miss}}$, with magnitude E_T^{miss}), is the negative vector sum of p_T measured in an event. Because the colliding protons have no initial transverse momentum, the true value of this quantity should be zero unless a particle escapes the detector without being measured, as neutrinos do. In practice, the reconstructed E_T^{miss} can also be non-zero due to mismeasurement, or due to gaps in the [ATLAS](#) detector. E_T^{miss} reconstruction is perhaps the most complex because it depends on all other object reconstructions performed in the [ATLAS](#) detector.

E_T^{miss} components are calculated independently for each type of baseline object reconstructed, as well as for a soft term, which comprises the energy observed by the [ATLAS](#) detector but not associated with a baseline object. The soft term can be calculated based either on calorimeter or track measurements [72]. While the Calorimeter Soft Term ([CST](#)) is very sensitive to pile-up, the Track Soft Term ([TST](#)) is much more robust, as it excludes tracks emanating from pile-up vertices. Tracks associated with any reconstructed object are also removed. [Figure 55](#) shows the dependence of the [TST](#) resolution on number of primary vertices. Because of this lessened pile-up dependence, the [TST](#) is used to reconstruct E_T^{miss} in this analysis.

$Z \rightarrow \mu\mu$ events, which rarely have any true E_T^{miss} , can be used to study the contribution of different objects to the total E_T^{miss} calculation. [Figure 56](#) shows the the E_T^{miss} resulting from muons, jets, and the soft term measured in events with two opposite-sign muons that reconstruct an invariant mass within 25 GeV of the Z boson mass. Large values in each of these terms, of course, do not necessarily correspond to a large value of the overall E_T^{miss} ; in a system where a Z boson recoils off of a jet system, the terms for jets and muons should both be large but ultimately cancel. However, the overall scale of each term indicates the possibility of contributions from mismeasurement. A 30% mismeasurement of the jet or muon term can result in significant E_T^{miss} , while a similar mismeasurement of the soft term is unlikely to produce a large impact on the E_T^{miss} calculation. The agreement between data and [MC](#) in these distributions indicates that, at least in the core of the distributions, these E_T^{miss} terms are well modeled.

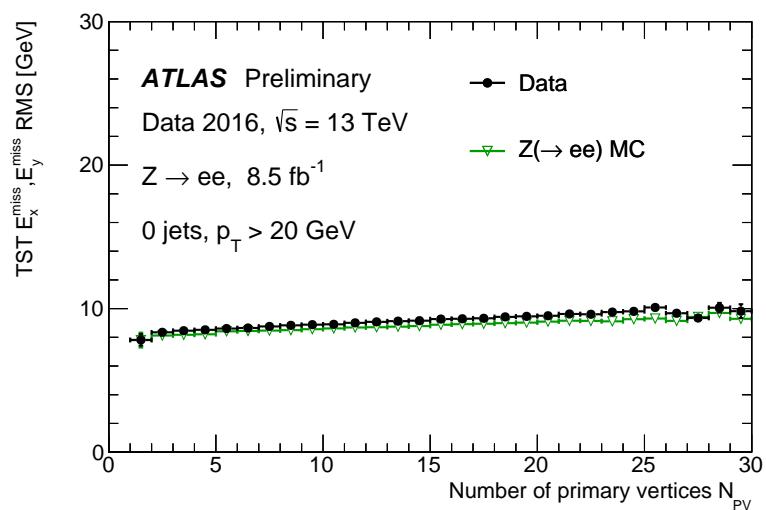


Figure 55: Distributions of the resolution of the x and y components of TST E_T^{miss} in $Z \rightarrow \mu\mu$ events in data and MC.

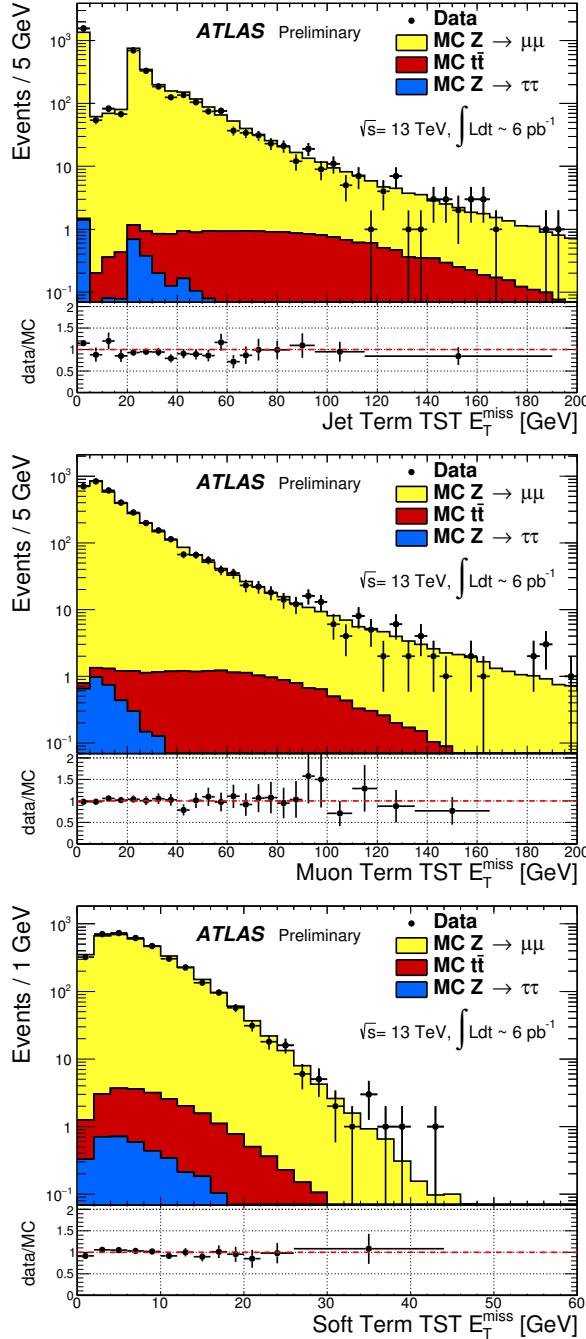


Figure 56: Distributions of the jet term (top left), muon term (top right), and $TST E_T^{\text{miss}}$ (bottom) E_T^{miss} in $Z \rightarrow \mu\mu$ events in data and MC. In the jet term distribution, the feature at zero is due to events with no jets, and the spike at 20 GeV corresponds to the minimum jet p_T considered for the analysis [73].

Part IV

SEARCHING FOR SUPERSYMMETRY

This section describes an analysis of the ATLAS data carried out by the author and her analysis team. The analysis was performed on events from pp collisions provided by the LHC at $\sqrt{s}=13$ TeV. It searches for events like those described in [Section 2.2.3](#), which contain a Z boson (decaying to leptons), jets, and missing transverse momentum. [Chapter 7](#) describes the SM backgrounds for the analysis. Definitions of the objects used in the analysis are outlined in [Chapter 8](#). [Chapter 9](#) explains the analysis strategy, defining the signal region as well as additional regions used to determine background contributions. Chapters [10](#) and [11](#) describe the methods used to estimate the SM backgrounds and the determination of their uncertainties. The results are presented in [Chapter 12](#), and they are interpreted in [Chapter 13](#).

BACKGROUND PROCESSES

This analysis is fundamentally a search for SUSY in events with two leptons whose invariant mass ($m_{\ell\ell}$) is consistent with a Z boson. Additional event selections are made to reduce SM processes relative to potential SUSY processes, defined by simplified models discussed in Section 2.2.3. These models include the production of strongly-charged, high-mass particles, which results in events with jets and high H_T , the scalar sum of the p_T of all jets and the two leading leptons in the event. These R-parity conserving SUSY models also produce decay chains that terminate with a stable, electrically neutral particle, which produces E_T^{miss} when it passes through ATLAS without detection. Each of these features can help isolate these signal events from SM backgrounds. To understand what cuts would optimize the sensitivity of the search, it is essential to first understand what these SM backgrounds are.

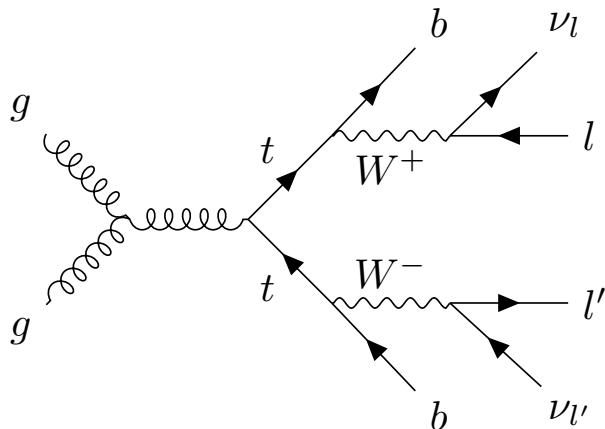


Figure 57: An example Feynman diagram of $t\bar{t}$ production and decay.

TOP-ANTITOP ($t\bar{t}$) production is the largest background for this search because it often precisely mimics the signature expected from the signal process. Figure 57 shows a Feynman diagram of this process, which results in two jets, leptons, and neutrinos, which are seen in the detector as E_T^{miss} . Thus, $t\bar{t}$ events naturally have high E_T^{miss} and H_T , jets, and leptons from two different W boson decays, which may coincidentally form an invariant mass consistent with a Z boson. Separating these events from signal processes is very difficult, but some steps can be taken to reduce their impact. The leptons resulting from $t\bar{t}$ don't originate from the same parent particle, so their invariant mass distribution is very broad compared to processes involving lep-

tons from Z bosons. As a consequence, reducing the width of the $m_{\ell\ell}$ window used to identify Z -boson candidates can help to reduce this background. Additionally, for sparticles with a higher mass than that of the top quark, the H_T distribution will typically be larger for signal models than for $t\bar{t}$. Though the two H_T distributions overlap significantly for most signal models, a high cut on H_T can help to reduce the $t\bar{t}$ background relative to a potential signal.

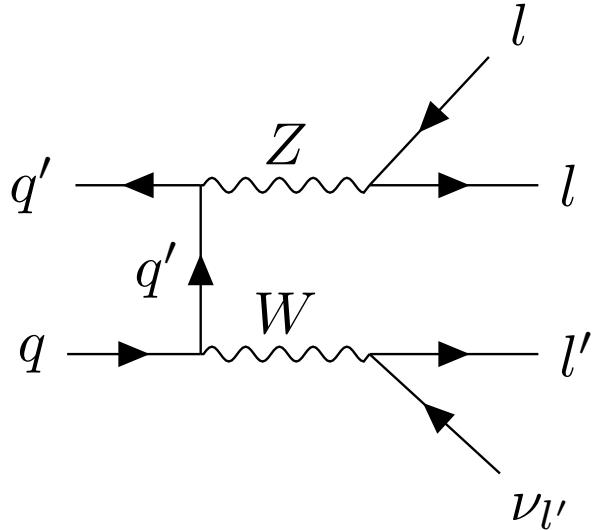


Figure 58: An example Feynman diagram of the production and decay of a WZ event.

DIBOSON (VV) production is the next leading background. These processes have lower cross-sections than that of $t\bar{t}$, but they can also produce events with topologies very similar to signals. Diboson events can contain real Z bosons and their dilepton invariant mass will peak on- Z like a signal. In addition, in events like Figure 58, an additional W boson can decay to another lepton and a neutrino, producing E_T^{miss} . The pictured process can occur with associated jets due to initial state radiation, but each additional jet reduces the process's rate by a factor of α_s . If the W boson in this figure instead decayed to two jets, which occurs about twice as often as the pictured decay, there would be no true E_T^{miss} from a neutrino. Thus, introducing a requirement that the signal region include both high amounts of E_T^{miss} and at least two jets can very effectively reduce the contribution from both these scenarios. A veto on a third lepton could also be used to reduce the contribution from events with $W \rightarrow \ell\nu$ decays, but, depending on the signal model considered, this veto can also decrease signal acceptance. Since the branching ratio for this process is relatively small, the potential reduction in signal acceptance is deemed more important, and a third lepton veto is not used in this analysis.

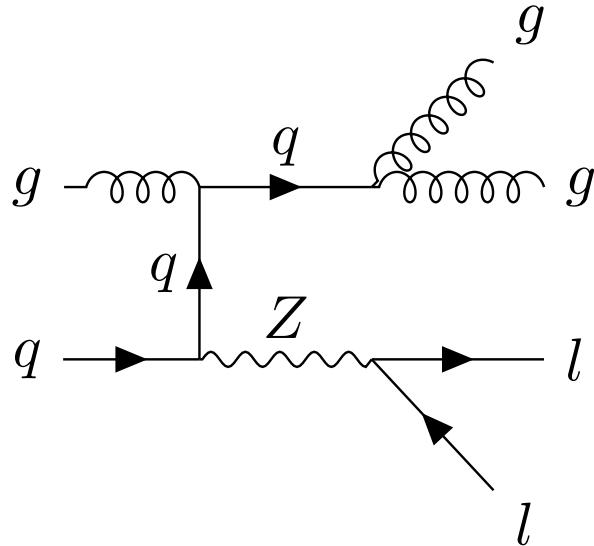


Figure 59: An example Feynman diagram of the production and decay of a $Z/\gamma^* + \text{jets}$ event.

$Z/\gamma^* + \text{jets}$ processes are very common but, as shown in Figure 59, don't typically produce any true E_T^{miss} . The exception is events $Z \rightarrow \tau\tau$ events, which contain E_T^{miss} because the τ leptons decay within the detector, and these decays include τ neutrinos. However, only about $1/3$ of τ decays produce an electron or muon, and even if two leptons are created in an event, they will typically not reconstruct a Z mass due to the energy lost to the neutrinos. A high H_T cut helps reduce the $Z/\gamma^* + \text{jets}$ background, but this process can occur with initial state radiation, producing events with large amounts of hadronic activity. E_T^{miss} is the most powerful variable for reducing $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ events, because though events with mismeasured jets or leptons can fake E_T^{miss} , mismeasurements drastic enough to produce hundreds of GeV of E_T^{miss} are rare. In addition, events with substantially mismeasured leptons often have $m_{\ell\ell}$ values that are inconsistent with a Z boson, and so a fraction of $Z/\gamma^* + \text{jets}$ events that do pass a E_T^{miss} threshold can be excluded nonetheless.

Other processes can contribute to the SM background at lower rates. Processes similar to $Z/\gamma^* + \text{jets}$ but with a W boson instead of a Z have real E_T^{miss} from leptonic W decays, but only one lepton. However, a fake or non-prompt lepton can cause these events to look very similar to simulated signals. Single top production can occur in association with a W boson, which produces signatures similar to $t\bar{t}$, but with fewer jets. Additionally, there are *rare top* processes such as $t\bar{t}$ production in association with W or Z bosons, which are even more difficult to separate from signal events than $t\bar{t}$ due to the presence of a genuine Z boson, the increased E_T^{miss} from W decays, and the larger amounts of hadronic activity.

Several of these backgrounds are grouped together and referred to as *Flavor Symmetric* ([FS](#)) processes. These include any processes that produce pairs of leptons with uncorrelated flavor in the final state. In this analysis, the largest [FS](#) background is $t\bar{t}$, with additional contributions from WW , Wt , and $Z \rightarrow \tau\tau$ events. In these processes, each lepton comes from the decay of a different particle. In the case of $t\bar{t}$, the two top quarks decay to W bosons, which each produce a lepton in their decays, as shown in [Figure 57](#). Consequently, these leptons' flavors are correlated with the flavor of the neutrino that results from the same boson's decay, but not with one another.

7.1 DATA AND MONTE CARLO SAMPLES

This analysis uses data collected by the [ATLAS](#) detector from pp collisions at a center-of-mass energy of 13 TeV in 2015 and 2016, corresponding to a total luminosity of 14.7 fb^{-1} . The data collected use a combination of unprescaled single and dilepton triggers, discussed in greater detail in [Section 9.1](#). The triggers targeting the lowest p_T ranges are fully efficient for pairs of leptons with p_T of at least 20 GeV. In addition, photon events are collected for use in a control region using both prescaled and unprescaled triggers, with the lowest trigger threshold at 35 GeV in 2015 and 20 GeV in 2016.

[MC](#) samples are generated for each background process that appears in the signal and validation regions. [Table 2](#) details the generators used to produce each sample, and more information can be found in [Section 4.7](#). These simulated background events, in conjunction with the simulated signal discussed in [Section 2.2.3](#), are used to determine approximate sensitivities of the search and optimize signal regions. The background [MC](#) also provides a valuable cross-check for many of the data-driven background estimates discussed in [Chapter 10](#), and in the case of rare top and some diboson events, provides the primary estimate of the background.

Table 2: Simulated background event samples used in this analysis with the corresponding matrix element and parton shower generators, cross-section order in α_s used to normalize the event yield, underlying-event tune and PDF set.

Physics process	Generator	Parton Shower	Cross section	Tune	PDF set
$t\bar{t} + W$ and $t\bar{t} + Z$ [74, 75]	MG5_AMC@NLO	Pythia 8.186	NLO [76, 77]	A14	NNPDF23LO
$t\bar{t} + WW$ [74]	MG5_AMC@NLO	Pythia 8.186	LO [51]	A14	NNPDF23LO
$t\bar{t}$ [78]	Powheg Box v2 r3026	Pythia 6.428	NNLO+NNLL [79, 80]	PERUGIA2012	NLO CT10
Single-top (Wt) [78]	Powheg Box v2 r2856	Pythia 6.428	Approx. NNLO [81]	PERUGIA2012	NLO CT10
WW , WZ and ZZ [82]	SHERPA 2.1.1	SHERPA 2.1.1	NLO [83, 84]	SHERPA default	NLO CT10
$Z/\gamma^*(\rightarrow \ell\ell) + \text{jets}$ [85]	SHERPA 2.1.1	SHERPA 2.1.1	NNLO [86, 87]	SHERPA default	NLO CT10
$\gamma + \text{jets}$	SHERPA 2.1.1	SHERPA 2.1.1	LO [47]	SHERPA default	NLO CT10
$V(=W,Z)\gamma$ signal	SHERPA 2.1.1	SHERPA 2.1.1	LO [47]	SHERPA default	NLO CT10
	MG5_AMC@NLO	Pythia 8.186	NLO	A14	NNPDF23LO

OBJECT IDENTIFICATION AND SELECTION

This section describes the identification and selection of objects in the events of this analysis. Objects are first required to pass *baseline* selections, which are used for Overlap Removal ([OR](#)) and the calculation of E_T^{miss} , then tighter *signal* selection cuts are applied, which define the objects considered in the final analysis of events. Definitions are presented for electrons, muons, and jets, which are all required in the Signal Region ([SR](#)) of this analysis, as well as photons, which are used in background estimation. This section refers to object definitions, including quality selections, described in [Chapter 6](#).

8.1 ELECTRONS

Electrons are reconstructed using the algorithm discussed in [Section 6.1](#). All electrons are required to be within $|\eta| < 2.47$, to ensure that all tracks are consistently within the tracking capability of the [ID](#) and that associated clusters are within the high-granularity portion of the calorimeter. Baseline leptons are required to have $p_T > 10 \text{ GeV}$ and pass the [Loose](#) quality standard. Signal leptons are further required to be isolated and of [Medium](#) quality, and must have $p_T > 25 \text{ GeV}$. Additional cuts on impact parameter are made for electrons with the goal of identifying only electrons coming from the hard-scatter vertex. These requirements, and all the other requirements made on the electrons can be seen in [Table 3](#).

With these signal requirements, the [ATLAS](#) detector is 80% efficient at identifying electrons with a p_T of 25 GeV, which rises to 90% at $p_T > 60 \text{ GeV}$ [[88](#)].

8.2 MUONS

Muons are reconstructed as discussed in [Section 6.3](#). Baseline muons are required to have $p_T > 10 \text{ GeV}$ and $|\eta| < 2.5$, including muons that can be tracked both by the [ID](#) and the [MS](#), and must pass a [Medium](#) quality cut. Signal muons are additionally required to have $p_T > 25 \text{ GeV}$, and to be isolated. As with the electrons, impact parameter cuts are made to ensure that the muon is consistent with coming from a decay from the event’s primary vertex. Additionally, the muon must not be flagged `isBadMuon`, a designation which identifies muons with inconsistent [ID](#) and [MS](#) p_T measurements. The full set of requirements can be seen in [Table 4](#).

Cut	Value/description
Baseline Electron	
Acceptance	$p_T > 10 \text{ GeV}, \eta^{\text{clust}} < 2.47$
Quality	Loose
Signal Electron	
Acceptance	$p_T > 25 \text{ GeV}, \eta^{\text{clust}} < 2.47$
Quality	Medium
Isolation	GradientLoose
Impact parameter	$ z_0 \sin \theta < 0.5 \text{ mm}$ $ d_0/\sigma_{d_0} < 5$

Table 3: Summary of the electron selection criteria. z_0 gives the track's distance from the hard-scatter vertex projected in the z direction, while d_0 gives the this distance projected onto the $x - y$ plane.

Cut	Value/description
Baseline Muon	
Acceptance	$p_T > 10 \text{ GeV}, \eta < 2.5$
Quality	Medium
Signal Muon	
Acceptance	$p_T > 25 \text{ GeV}, \eta < 2.5$
Quality	Medium
Isolation	GradientLoose
Impact parameter	$ z_0 \sin \theta < 0.5 \text{ mm}$ $ d_0/\sigma_{d_0} < 3$
isBadMuon	isBadMuon Flag vetoed

Table 4: Summary of the muon selection criteria. The signal selection requirements are applied on top of the baseline selection.

Muons are identified with an efficiency over 98% [89]. Including trigger and isolation requirements, this efficiency drops to about 80% for muons with $p_T > 25$ GeV and 90% for muons with $p_T > 200$ GeV. This drop due to trigger inefficiency, and is largely the consequence of incomplete η coverage of the [RPCs](#), discussed in [Section 6.3](#).

8.3 JETS

Jets are reconstructed according to [Section 6.4](#), with baseline jets using the anti- k_t algorithm with $R = 0.4$ using [EM](#) clusters. These jets are required to have a minimum p_T of 20 GeV and $|\eta| < 2.8$. For signal jets, this p_T requirement is increased to 30 GeV and the η cut is tightened to $|\eta| < 2.5$. [JVT](#) requirements are enforced on jets with $p_T < 60$ GeV and $|\eta| < 2.4$ to reduce the number of jets from pile-up. The full set of requirements can be seen in [Table 5](#).

Cut	Value/description
Baseline jet	
Collection	AntiKt4EMTopo
Acceptance	$p_T > 20$ GeV , $ \eta < 2.8$
Signal jet	
Acceptance	$p_T > 30$ GeV , $ \eta < 2.5$
JVT	$ JVT > 0.59$ for jets with $p_T < 60$ GeV and $ \eta < 2.4$
Signal b -jet	
b -tagger Algorithm	MV2c20
Efficiency	77 %
Acceptance	$p_T > 30$ GeV , $ \eta < 2.5$
JVT	$ JVT > 0.59$ for jets with $p_T < 60$ GeV and $ \eta < 2.4$

Table 5: Summary of the jet and b -jet selection criteria. The signal selection requirements are applied on top of the baseline requirements.

Though no b -jets are required or vetoed in the [SR](#) of this analysis, some Control Regions ([CRs](#)) use b -enhanced and b -vetoed regions to determine the impact of heavy flavor. The specifications for b -tagging are also described in [Table 5](#).

8.4 PHOTONS

Photons are used to estimate the $Z/\gamma^* + \text{jets}$ background in this analysis, and they are reconstructed according to [Section 6.2](#). Baseline and

signal photons are nearly identical. Each must pass a Tight selection and an isolation cut, and have $p_T > 25$ GeV as well as $|\eta| < 2.37$. Signal photons with $1.37 < |\eta| < 1.6$ are rejected due to an discontinuity in the calorimeter which results in inferior energy resolutions in this region. The full selection requirements can be seen in [Table 6](#).

Cut	Value/description
Baseline Photon	
Acceptance	$p_T > 25$ GeV, $ \eta < 2.37$
Quality	tight
Signal Photon	
Acceptance	$p_T > 25$ GeV, $ \eta < 2.37$ rejecting $1.37 < \eta < 1.6$
Quality	tight
Isolation	FixedCutTight

Table 6: Summary of the photon selection criteria.

EVENT SELECTION

The goal of this analysis is to identify events resembling Figure 14 in collisions in the ATLAS detector. In order to do this, a Signal Region (**SR**) is defined with the goal of maximizing the identification efficiency of signal events while minimizing SM backgrounds. However, because this analysis re-investigates an excess of events seen in Run 1 with the ATLAS detector, the signal region was frozen and could not be re-optimized for the new, higher energy data in Run 2.

The **SR**, called SRZ, was predetermined, including events with two opposite-sign, same-flavor leptons that reconstruct a mass, $m_{\ell\ell}$, close to that of the Z boson, with the additional requirement of two jets, $E_T^{\text{miss}} > 225$ GeV, and H_T of at least 600 GeV. Additionally, a cut on $\Delta\phi(\text{jet}_{12}, p_T^{\text{miss}})$ was made in order to suppress $Z/\gamma^* + \text{jets}$ background events with mismeasured jets. In events where jet mismeasurement produces large amounts of E_T^{miss} , the p_T^{miss} vector should either align or anti-align with the mismeasured jet's trajectory based on whether its energy has been overestimated or underestimated. This cut vetoes events in which the two vectors are aligned, reducing the number of events that will be included in the **SR** due to the underestimation of jet energy. Though it would be possible to also veto events in which the vectors are anti-aligned, similar topologies can also be produced by signal processes, so no such requirement is made.

The freezing of the **SR** prevented re-optimization of the cuts for maximum sensitivity to a signal, but the search was still expected to be more sensitive than the previous iteration performed on 8 TeV data. This is because the cross-section for production of this high-mass signal increased dramatically with the increased LHC energy, while the cross-section for SM processes increased at lower rates. With 20.3 fb^{-1} of 8 TeV data, the previous search had an expected exclusion on gluino masses up to about 950 GeV [1], while with 14.7 fb^{-1} of 13 TeV data, this search has an expected exclusion on gluino masses up to about 1300 GeV¹.

Though this **SR** was fixed, the methods used to estimate the expected SM backgrounds were not. A set of Control Regions (**CRs**) and Validation Regions (**VRs**) were chosen to make these estimations possible. **CRs** are regions in which the collected data can be used to make an estimate of a SM background in the **SR**, while **VRs** are used to confirm that the background estimate based on the **CR** is valid. Both **CRs**

¹ The expected exclusions, discussed further in Chapter 13, are made assuming the data exactly matches background predictions. The previous search had an observed exclusion on gluino masses up to about 850 GeV, which was lower than the expected exclusion due to the excess observed.

and **VRs** are designed to minimize contamination from the **BSM** process being searched for. This is desirable because signal contamination in a **CR** can lead to an overestimate of the **SM** background in the **SR**, disguising a genuine signal as background. Contamination in a **VR**, where background estimates are being validated, can make it appear that the **SM** background is not well described by an estimate, causing analyzers to adjust the method to account for the difference, and again, disguising the effect of the same signal in the **SR**.

The strategy for estimating the **FS** background, for example, depends on a series of **CRs** and **VRs** depicted in Figure 6o. One estimate, the flavor symmetry method, takes data from CR-FS, a different-flavor region with slightly wider $m_{\ell\ell}$ bounds than the **SR**, and uses these events to predict the contribution of flavor symmetric processes to SRZ. An independent method called a sideband fit uses a control region CRT to measure the flavor symmetric events outside of the Z mass window, and uses **MC** to extrapolate inside the Z mass window to SRZ. In addition, both methods are validated at lower E_T^{miss} with an otherwise identical series of regions, with VRS corresponding to SRZ, VRT corresponding to CRT, and VR-FS corresponding to CR-FS.

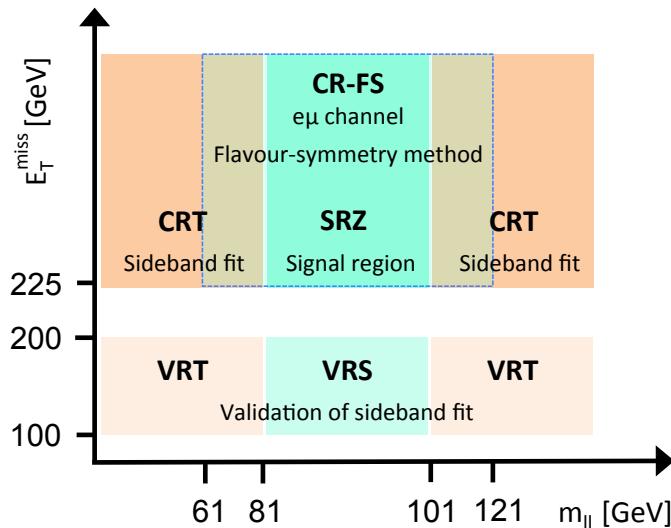


Figure 6o: Schematic layout of the signal region, as well as the control and validation regions used to predict the **FS** background. The regions displayed differ only by their $m_{\ell\ell}$ and E_T^{miss} values, except for CR-FS, which contains different-flavor leptons, while all other regions contain same-flavor leptons.

The background estimation methods, described in Chapter 10, each require their own set of these regions. The full list of regions used in this analysis can be seen in Table 7. In addition to the **FS** regions described above, there is one more **CR**, CR- γ , which is a photon region used to predict the number of $Z/\gamma^* + \text{jets}$ events, a process described

in [Section 10.2](#). Additional **VRs**, VR-ZZ, VR-WZ, and VR-3L, are introduced in order to validate the diboson and rare top backgrounds that are taken directly from **MC**. There are several additional regions used in the estimation of the fakes and $Z/\gamma^* + \text{jets}$ contributions to the **SR** that are defined in their respective sections.

On-shell Z regions	E_T^{miss} [GeV]	H_T [GeV]	n_{jets}	$m_{\ell\ell}$ [GeV]	SF/DF	$\Delta\phi(\text{jet}_{12}, p_T^{\text{miss}})$	$m_T(\ell_3, E_T^{\text{miss}})$ [GeV]	$n_{\text{b-jets}}$
Signal region								
SRZ	> 225	> 600	≥ 2	$81 < m_{\ell\ell} < 101$	SF	> 0.4	—	—
Control regions								
CRZ	< 60	> 600	≥ 2	$81 < m_{\ell\ell} < 101$	SF	> 0.4	—	—
CR-FS	> 225	> 600	≥ 2	61 < $m_{\ell\ell} < 121$	DF	> 0.4	—	—
CRT	> 225	> 600	≥ 2	> 40 , $m_{\ell\ell} \notin [81, 101]$	SF	> 0.4	—	—
CR γ	—	> 600	≥ 2	—	$0\ell, 1\gamma$	—	—	—
Validation regions								
VRZ	< 225	> 600	≥ 2	$81 < m_{\ell\ell} < 101$	SF	> 0.4	—	—
VRT	100–200	> 600	≥ 2	> 40 , $m_{\ell\ell} \notin [81, 101]$	SF	> 0.4	—	—
VRS	100–200	> 600	≥ 2	$81 < m_{\ell\ell} < 101$	SF	> 0.4	—	—
VR-FS	100–200	> 600	≥ 2	61 < $m_{\ell\ell} < 121$	DF	> 0.4	—	—
VR-WZ	100–200	—	—	—	3 ℓ	—	< 100	0
VR-ZZ	< 100	—	—	—	4 ℓ	—	—	0
VR-3L	60–100	> 200	≥ 2	$81 < m_{\ell\ell} < 101$	3 ℓ	> 0.4	—	—

Table 7: Overview of all signal, control and validation regions used in the on-shell Z search. More details are given in the text. The flavour combination of the dilepton pair is denoted as either *SF* for same-flavor or *DF* for different flavor. All regions require at least two leptons, unless otherwise indicated. In the case of CR γ , VR-WZ, VR-ZZ, and VR-3L the number of leptons, rather than a specific flavor configuration, is indicated. The main requirements that distinguish the control and validation regions from the signal region are indicated in bold. Most of the kinematic quantities used to define these regions are discussed in the text. The quantity $m_T(\ell_3, E_T^{\text{miss}})$ indicates the transverse mass formed by the E_T^{miss} and the lepton which is not assigned to either of the Z-decay leptons.

9.1 TRIGGER STRATEGY

In order to collect data for the analysis, triggers must be chosen that provide good efficiency in each of the **SR**, **VR**, and **CRs**. This analysis primarily depends on triggers on leptons, which are required in nearly every region. To simplify the application of trigger scale factors, which correct **MC** yields for a given trigger to match the data efficiencies, events are broken down into a series of kinematic ranges, each with a designated trigger. These regions can be seen in [Table 8](#).

In kinematic regions where single lepton triggers are fully efficient, they are preferentially used. In lower- p_T ranges, dilepton triggers are used, targeting either ee , $\mu\mu$, or $e\mu$ events. Electron triggers are se-

Lepton p_T	Trigger in 2015	Trigger in 2016
Di-electron channel		
$p_T(e_1) > 65 \text{ GeV}$	HLT_e60_lhmedium	HLT_e60_lhmedium_nod0
$p_T(e_1) \leq 65 \text{ GeV}$	HLT_2e17_lhloose	HLT_2e17_lhvloose_nod0
Di-muon channel		
$p_T(\mu_1) > 52.5 \text{ GeV}$	HLT_mu50	HLT_mu50
$p_T(\mu_1) \leq 52.5 \text{ GeV}$	HLT_mu24_mu8noL1	HLT_2mu14_nomucomb
Electron-muon channel		
$p_T(e) > 65 \text{ GeV}$	HLT_e60_lhmedium	HLT_e60_lhmedium_nod0
$p_T(e) \leq 65 \text{ GeV} \text{ and } p_T(\mu) > 52.5 \text{ GeV}$	HLT_mu50	HLT_mu50
$p_T(e) \leq 65 \text{ GeV} \text{ and } p_T(\mu) \leq 52.5 \text{ GeV}$ and $p_T(e) < p_T(\mu)$	HLT_e7_lhmedium_mu24	HLT_e7_lhmedium_nod0_mu24
$p_T(e) \leq 65 \text{ GeV} \text{ and } p_T(\mu) \leq 52.5 \text{ GeV}$ and $p_T(\mu) < p_T(e)$	HLT_e17_lhloose_mu14	HLT_e17_lhloose_nod0_mu14

Table 8: Lepton trigger requirements used for the analysis in different regions of lepton- p_T phase space.

lected over muon triggers when possible because they have higher trigger efficiencies.

Table 8 also shows the triggers used in 2015 and 2016. The two are not identical because higher instantaneous luminosities were reached in 2016, and the trigger menu was changed to compensate for the increased rates. In addition, triggers that didn't include impact parameter requirements were made available in 2016, and in the interest of not excluding signals with displaced vertices, these triggers were used.

In CR- γ , there are no leptons, so an alternate trigger strategy must be used. Section 10.2.1 describes this triggering scheme, which includes a combination of prescaled and unprescaled photon triggers to allow for the collection of low- p_T $\gamma + \text{jets}$ events.

9.2 SIGNAL EFFICIENCY AND CONTAMINATION

Using the simplified models discussed in Section 2.2.3, the contributions of potential signals in these regions can be studied. In the SR, the goal is to include as much of the potential signal as possible, while excluding as much SM background as possible. Figure 61 shows the acceptance and efficiency for the simplified models at different mass points. Acceptance is defined as the fraction of signal events that produce signatures that kinematically match the SR, while the efficiency is the fraction of these events expected to be correctly reconstructed and identified. Acceptance ranges between 3-5% for the simplified models considered, while efficiency is typically 70-90%. The acceptance is small because the simplified models include all possible Z boson decays, and only about 7% of these decays result in electron or muon pairs. These studies are performed using MC, with corrections

applied to match data efficiencies. These efficiencies are measured as a function of p_T and η , and include separate identification efficiencies for electrons and muons.

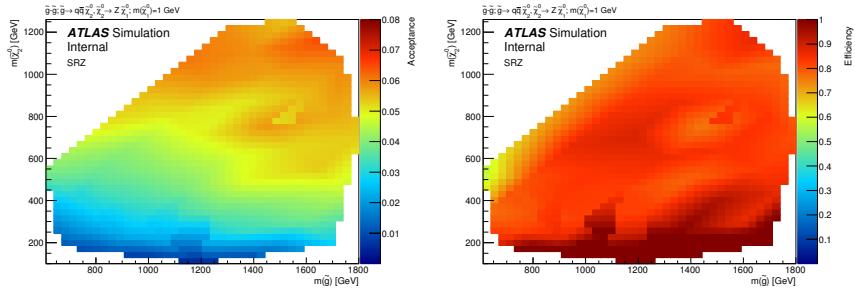


Figure 61: Signal region acceptance (left) and efficiency (right) in SRZ for the simplified model with gluino pair production with $\tilde{\chi}_2^0$ decays to $\tilde{\chi}_1^0$ and an on-shell Z boson with 1 GeV neutralino LSP. Acceptance is calculated by applying the signal-region kinematic requirements to truth objects in MC, which do not suffer from identification inefficiencies or measurement resolutions.

Using the same MC, the possibility of signal contamination in the CRs and VRs is assessed. Figures 62 and 63 show the fraction of events in these regions expected to come from signal for different points on the simplified model’s mass grid. Contamination is highest in VRS, at low $m_{\tilde{g}}$. However, past analyses have already excluded most models with $m_{\tilde{g}} < 800$ GeV [1], so this is not a concern. For models with $m_{\tilde{g}} > 800$ GeV, signal contamination in VRS is below 30%, and for models with $m_{\tilde{g}} > 1$ TeV, the contamination decreases to 10%. In the CRs, contamination is below 20% for models with $m_{\tilde{g}} > 800$ GeV, and below 5% for models with $m_{\tilde{g}} > 1$ TeV.

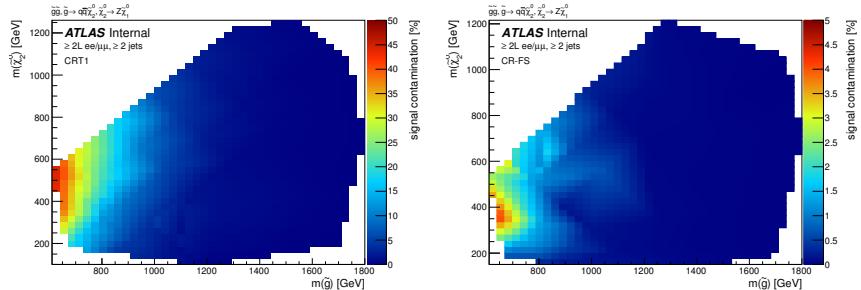


Figure 62: Expected signal contamination in CRT (left) and CR-FS (right) for the signal model with gluino pair production, where the gluinos decay to quarks and a neutralino, with the neutralino subsequently decaying to a Z boson and a 1 GeV neutralino LSP.

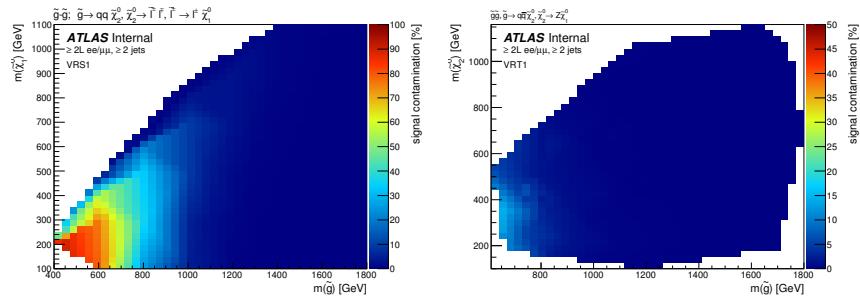


Figure 63: Expected signal contamination in VRS (left) and VRT (right) for the signal model with gluino pair production, where the gluinos decay to quarks and a neutralino, with the neutralino subsequently decaying to a Z boson and a 1 GeV neutralino LSP.

BACKGROUND ESTIMATION

This analysis requires two leptons with $m_{\ell\ell}$ consistent with the Z mass, at least two jets, large E_T^{miss} , and large H_T . Any SM processes that produce this signature will appear as a background to the search. The most important task of the analysis is to identify and estimate these backgrounds, so that any excess of events appearing on top of the standard model background can be identified. The main backgrounds processes for this analysis are described in [Chapter 7](#). The largest background is from flavor symmetric processes, with smaller contributions coming from the remaining diboson processes, $Z/\gamma^* + \text{jets}$, rare top processes, and fake and non-prompt leptons.

10.1 FLAVOR SYMMETRIC PROCESSES

The FS background estimation is performed using two methods. The first, the Flavor Symmetry method, takes advantage of the uncorrelated lepton flavors in these processes, and uses data observed in the $e\mu$ channel to predict backgrounds in the ee and $\mu\mu$ channels. The second, the Sideband Fit method, normalizes MC using data in the off-Z region of the ee and $\mu\mu$ channels, and uses that normalization to make a prediction of the on-Z region. Because it depends less on MC, and because it has smaller total uncertainties (as discussed in [Chapter 11](#)), the Flavor Symmetry method is chosen as the nominal prediction, with the Sideband Fit method used as a cross-check.

10.1.1 *Flavor Symmetry Method*

As a consequence of the independence of the lepton flavors, any FS process should produce ee , $\mu\mu$, and $e\mu$ events in a 1:1:2 ratio. This ratio is taken advantage of in the flavor symmetry method by measuring $e\mu$ events in data and using them to predict the contribution of these processes in the ee and $\mu\mu$ channels [1].

To estimate the number of events in SRZ, a control region called CR-FS is used. Both regions are defined in [Table 7](#). CR-FS is very similar to SRZ with two changes: it requires different-flavor leptons instead of the same-flavor leptons required by SRZ, and the $m_{\ell\ell}$ range it covers has been expanded by a factor of three, now ranging from 61 to 121 GeV. The expansion of the $m_{\ell\ell}$ window is done to increase the

number of events in the control region, thus lowering the statistical uncertainty of the prediction ¹.

This control region is expected to be about 95% pure in **FS** processes, with most of the remaining events coming from fake or non-prompt leptons. The **FS** portion is made up primarily of $t\bar{t}$ ($\sim 80\%$), with additional contributions from Wt ($\sim 10\%$), WW ($\sim 10\%$), and $< 1\% Z \rightarrow \tau\tau$.

The 1:1:2 ratio cannot be applied directly to the events measured in CR-FS because the efficiencies for identifying electrons and muons are not identical. Correction factors are applied to account for trigger efficiencies for each channel, selection efficiencies for electrons and muons, the $m_{\ell\ell}$ expansion, and the purity of the control region. Combining these factors, the estimate for number of events in the ee and $\mu\mu$ channels is as follows:

$$N_{ee}^{\text{est}} = \frac{1}{2} \cdot f_{\text{FS}} \cdot f_{Z\text{-mass}} \cdot \sum_{\ell=1}^{N_{e\mu}^{\text{data}}} k_e(p_T^\ell, \eta^\ell) \cdot \alpha(p_T^{\ell_1}, \eta^{\ell_1}), \quad (38)$$

$$N_{\mu\mu}^{\text{est}} = \frac{1}{2} \cdot f_{\text{FS}} \cdot f_{Z\text{-mass}} \cdot \sum_{\ell=1}^{N_{e\mu}^{\text{data}}} k_\mu(p_T^\ell, \eta^\ell) \cdot \alpha(p_T^{\ell_1}, \eta^{\ell_1}), \quad (39)$$

where $N_{e\mu}^{\text{data}}$ is the number of data events observed in CR-FS, f_{FS} is the fraction of events from **FS** processes in CR-FS, $f_{Z\text{-mass}}$ is the fraction of events in the widened $m_{\ell\ell}$ range expected to be in the on- Z range (taken from $t\bar{t}$ MC), $k_e(p_T, \eta)$ and $k_\mu(p_T, \eta)$ are relative selection efficiencies for electrons and muons, calculated in bins of p_T and η of the lepton to be replaced, and $\alpha(p_T, \eta)$ accounts for the different trigger efficiencies for events in each channel, binned based on the kinematic properties of the leading lepton.

These k and α factors are calculated from data in an inclusive on- Z selection ($81 < m_{\ell\ell} / \text{GeV} < 101$, ≥ 2 jets), according to:

$$k_e(p_T, \eta) = \sqrt{\frac{N_{ee}^{\text{meas}}}{N_{\mu\mu}^{\text{meas}}}} \quad (40)$$

$$k_\mu(p_T, \eta) = \sqrt{\frac{N_{\mu\mu}^{\text{meas}}}{N_{ee}^{\text{meas}}}} \quad (41)$$

$$\alpha(p_T, \eta) = \frac{\sqrt{\epsilon_{ee}^{\text{trig}}(p_T, \eta) \times \epsilon_{\mu\mu}^{\text{trig}}(p_T, \eta)}}{\epsilon_{e\mu}^{\text{trig}}(p_T, \eta)} \quad (42)$$

¹ This method was developed for a smaller dataset for which this statistical uncertainty was the dominant uncertainty on the total background prediction, and this expansion substantially reduced the overall uncertainty [90]. If the **SR** were re-optimized, the cuts on H_T and E_T^{miss} would be increased, leading to fewer events in CR-FS, and this statistical uncertainty would once again become dominant.

where $\epsilon_{ee/\mu\mu}^{\text{trig}}$ is the trigger efficiency² and $N_{ee/\mu\mu}^{\text{meas}}$ is the number of $ee/\mu\mu$ events in the inclusive on-Z region described above. Here, the k factors are related by $k_e = 1/k_\mu$. These factors are calculated separately for leading and sub-leading leptons, and the appropriate k value is selected based on which of the leptons is to be replaced.

The α factor's complicated form is a result of its dual purpose; it accounts for the difference in trigger efficiencies between the different-flavor and same-flavor channels, and corrects for the implicit trigger dependence of the k factor.

In most p_T and η ranges, the difference between electron and muon selection efficiencies is small, so these correction factors are typically within 10% of unity. In the region $|\eta| < 0.1$ they are up to 50% from unity because of the lack of coverage of the muon spectrometer. k_e factors measured in the 2016 dataset can be seen in [Figure 64](#). In most cases, the only correction factor that differs substantially from one is $f_{Z\text{-mass}}$, which is typically about 1/3.

The estimate is corrected for contamination of non-[FS](#) backgrounds in CR-[FS](#) according to

$$f_{\text{FS}} = \frac{N_{\text{obs}} - N_{\text{non-}\text{FS}}}{N_{\text{obs}}} \quad (43)$$

where N_{obs} is the number of observed data events in CR-[FS](#), and $N_{\text{non-}\text{FS}}$ is the number of expected events in CR-[FS](#) not resulting from [FS](#) processes. This number is taken from [MC](#) for all backgrounds except for fakes, which are estimated using the data-driven matrix method described in [Section 10.3](#). According to the estimate, fakes account for about 65% of the contamination in this region, with the remaining contamination resulting from diboson (15%) and rare top (20%) processes. f_{FS} is about 94% in CR-[FS](#).

A prediction is made both for SRZ using the 91 events measured in CR-[FS](#), and the lower- E_T^{miss} validation region, VRS, with the 277 events in VR-[FS](#). The background estimation is performed separately for the two data taking periods, 2015 and 2016, because of the changing triggers and conditions. The results are then summed together, as shown in [Table 9](#). The uncertainties in this table are discussed in [Section 11.1.1](#).

10.1.2 Sideband Fit Method

As a crosscheck to the flavor symmetry method, a [MC](#)-based method is used. This method is called a *sideband fit*, and it begins with a [MC](#)

² This efficiency is defined by taking all events in the inclusive on-Z selection mentioned above and determining the fraction that passes the relevant trigger requirement defined by [Table 8](#). Because the offline selection made on these events already has some trigger dependence, this calculation of efficiency could be slightly biased. This effect is considered in [Section 11.1.1](#), and the uncertainty applied to the estimate as a result is described.

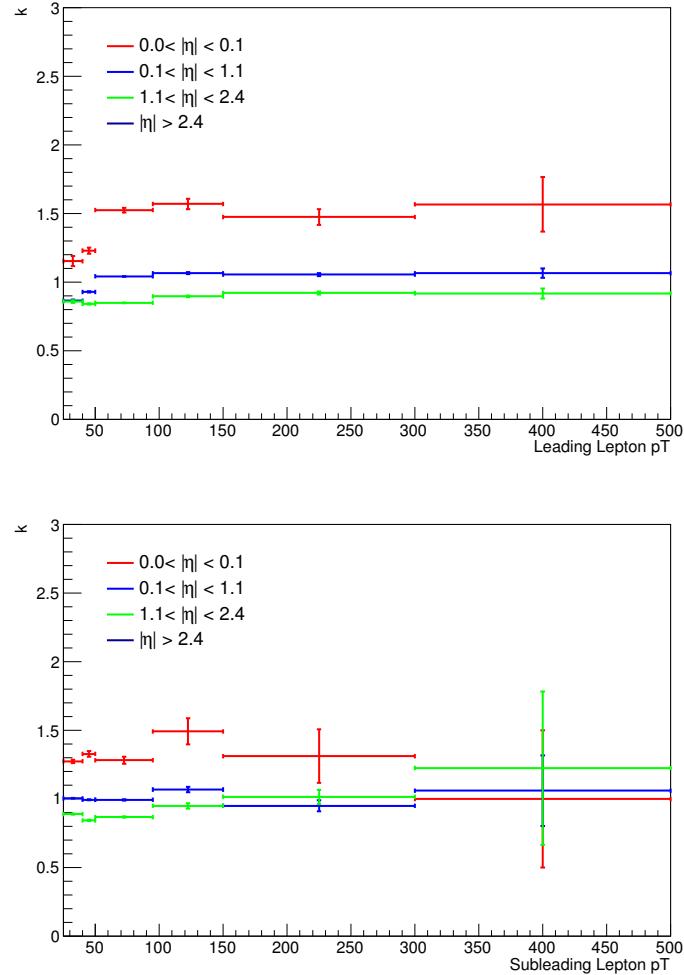


Figure 64: Measurements of k_e , the ratio of electron to muon events, in bins of p_T and η , with statistical uncertainties shown. On the top is the measurements indexed by the leading lepton, while the measurements indexed by the subleading lepton are on the bottom. The efficiencies displayed are for the 2016 dataset.

estimate of the signal region across an $m_{\ell\ell}$ range that includes all values above 40 GeV. This region, excluding the on-Z range that makes up the **SR**, is used as a control region, defined as **CRT** in Table 7.

The **SM** backgrounds are estimated for **CRT** using **MC**, except for the fakes background, which is taken from the data-driven method described in Section 10.3. This background prediction is then fit to the measured data yield in **CRT** with one free parameter, μ . This parameter scales the overall normalization of $t\bar{t}$ **MC**. All other backgrounds contributing to this control region are constrained by their uncertainties, which are used as nuisance parameters in the fit. Once μ has been acquired from this fit, it is applied to the $t\bar{t}$ **MC** yield in the **SR**. This scaled $t\bar{t}$ **MC** is added to the raw **MC** estimates of the other **FS** processes to give a final estimate of the **FS** background.

Region	ee prediction	$\mu\mu$ prediction	combined prediction
SRZ	16.5 ± 2.1	16.7 ± 2.0	33.2 ± 3.9
VRS	49.7 ± 4.6	49.6 ± 4.6	99.3 ± 8.5

Table 9: Yields in signal and validation regions for the flavor symmetric background. Errors include statistical and systematic uncertainties, discussed in [Chapter 11](#).

The results of the fit can be seen in [Table 10](#), along with the original [MC](#) or data-driven estimates for each background. The bottom half of the table shows the raw [MC](#) estimates, as well as data-driven estimates for the fakes and $Z/\gamma^* + \text{jets}$ backgrounds. The top half of the table shows values for each of these quantities after the fit. The only significant change after the fit is the $t\bar{t}$ background, which is scaled with $\mu = 0.64$. Normalization factors acquired in CRT, including μ and the nuisance parameters for other [MC](#)-driven backgrounds, are applied to SRZ.

To validate the method, it is repeated in VRS using a fit acquired from VRT. These regions are identical to SRZ and CRT, but with E_T^{miss} ranging from 100 to 200 GeV. The normalization factors, listed in [Table 11](#), are significantly different for the two regions. This difference is expected because the $t\bar{t}$ [MC](#) over-predicts the high- E_T^{miss} tail. This discrepancy between data and [MC](#) is likely due to a mismodeling of the top quark p_T distribution, which does not match the spectrum seen in data [91, 92]. [Figure 65](#) shows the lepton p_T distribution in $t\bar{t}$ events, demonstrating this discrepancy. This sideband fit method corrects for this mismodeling by performing fits in regions very kinematically similar to the signal region.

This method is extremely effective as a crosscheck because it uses a completely independent dataset from the flavor symmetry method, and the two methods have very little overlap in dependence on [MC](#). They produce consistent results in both SRZ and VRS, as shown in [Table 12](#).

10.2 $Z/\gamma^* + \text{JETS BACKGROUND}$

The $Z/\gamma^* + \text{jets}$ background is produced by a process called Drell-Yan in which annihilating quark/anti-quark pairs produce a Z boson or a virtual photon. These bosons then decay to two leptons, which, in the case of the Z boson, naturally appear in the Z -mass window. The boson typically recoils off a hadronic system, which can satisfy the jet and H_T requirement in SRZ. However, this process rarely produces real E_T^{miss} (though occasionally neutrinos do appear in hadronic decays of a Z boson), so most events with large amounts of E_T^{miss} are

channel	$ee/\mu\mu$ CRT	$ee/\mu\mu$ SRZ	ee SRZ	$\mu\mu$ SRZ
Observed events	273	60	35	25
Fitted bkg events	277.76 ± 16.88	49.33 ± 8.04	27.09 ± 4.73	22.70 ± 3.80
Fitted flavour symmetry events	236.96 ± 21.66	28.96 ± 7.47	16.41 ± 4.33	12.55 ± 3.29
Fitted WZ/ZZ events	4.03 ± 1.13	14.27 ± 4.45	7.81 ± 2.45	6.46 ± 2.07
Fitted SHERPA Z/γ^* + jets events	1.95 ± 0.14	–	–	–
Data-driven Z/γ^* + jets (γ + jets) events	–	3.10 ± 2.25	$1.02^{+1.25}_{-1.02}$	2.08 ± 1.38
Fitted rare top events	4.04 ± 1.04	2.90 ± 0.76	1.39 ± 0.38	1.50 ± 0.40
Data-driven fake lepton events	25.78 ± 14.26	$0.10^{+0.18}_{-0.10}$	0.46 ± 0.45	0.10 ± 0.01
Expected SM Events	366.71	61.01	33.73	27.74
MC flavour symmetry events	331.32	40.72	23.09	17.63
MC WZ/ZZ events	4.02	14.20	7.77	6.43
MC SHERPA Z/γ^* + jets events	1.94	–	–	–
Data-driven Z/γ^* + jets (γ + jets) events	–	3.10	1.02	2.08
MC rare top events	4.04	2.89	1.39	1.50
Data-driven fake lepton events	25.39	0.10	0.46	0.10

Table 10: Background fit results from the sideband fit method. The $t\bar{t}$ MC’s normalization is taken as a free parameter in the fit to data in CRT, then that normalization factor is applied in SRZ. The results are shown here both divided between the ee and $\mu\mu$ channels and summed together. All other backgrounds are taken from MC in CRT while in SRZ, the Z/γ^* + jets contribution is taken from the γ + jets method. The uncertainties quoted include both statistical and systematic components.

Fit region	$t\bar{t}$ normalization (μ)
CRT	0.64 ± 0.18
VRT	0.80 ± 0.09

Table 11: Summary of the $t\bar{t}$ normalization factors calculated by the side-band fit to CRT and VRT for the 2015+2016 data.

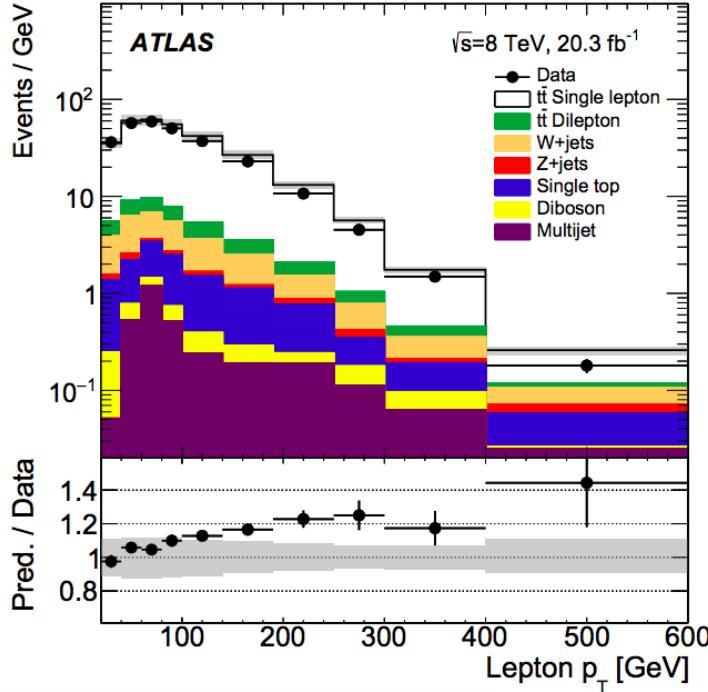


Figure 65: Lepton p_T distribution in data and MC, demonstrating the overestimate of the high- p_T tail in MC [91].

the result of mismeasurement. The 200 GeV E_T^{miss} cut in SRZ includes Z events only from the very high tail of the E_T^{miss} distribution, so a small change in the assumptions about jet resolution or energy scale in MC can drastically change the prediction. If the prediction of the $Z/\gamma^* + \text{jets}$ background is too low, this will result in a signal-like peak of data over expected background.

Because of this volatility in the MC prediction in these high E_T^{miss} tails, a data-driven method is used to estimate this background. The method uses $\gamma + \text{jets}$ events which, like the $Z/\gamma^* + \text{jets}$ events, contain one boson recoiling against a hadronic system. These $\gamma + \text{jets}$ events are then corrected for the kinematic differences between photons and Z bosons [33, 93]. The sample of $\gamma + \text{jets}$ events is taken from CR- γ , defined in Table 7. This region is similar to the SRZ selection without the E_T^{miss} requirement, but it vetoes events with leptons and requires at least one photon. Additionally, the $\Delta\phi(\text{jet}_{12}, \mathbf{p}_T^{\text{miss}})$ cut

Region	Flavour-symmetry	Sideband fit
SRZ	33 ± 4	29 ± 7
VR-S	99 ± 8	92 ± 25

Table 12: Comparison of [FS](#) background predictions from the nominal method, the flavor symmetry method, and the cross-check, the sideband fit method. Uncertainties include statistical and systematic uncertainties in both cases.

in SRZ, which is designed to reduce the background from mismeasured jets, is removed for this region because of its unpredictability at very low values of E_T^{miss} , when the angle of the E_T^{miss} is much less meaningful.

Despite their similarities, there are many theoretical differences between γ and Z events. The massive Z boson requires more energy to produce than the massless photon, producing a difference in the kinematic distributions for bosons and jets in these events. Another consequence of photon's masslessness is that they must be produced with associated jets; otherwise, the process would violate conservation of energy. As a consequence of these differences, many kinematic variables have different shapes between the two samples. [Figure 66](#) shows a [MC](#) comparison of boson p_T between γ and Z events, demonstrating the shape differences between the two processes.

The most significant experimental difference between Z and γ events is that Z bosons rapidly decay, in the case of this analysis, to two leptons, which are then be observed by the [ATLAS](#) detector. In contrast, the photon is stable, and can be directly detected by [ATLAS](#). This means that the reconstructed Z boson has an energy resolution that's the convolution of the resolutions of either a pair of electrons or a pair of muons, while the photon resolution is the result of only one measurement. Though electron and photon resolutions are similar because both depend on the electromagnetic calorimeter, the difference between the measurement of two objects and one object creates small differences. The muons, whose energy measurements rely on the reconstruction of tracks in the [ID](#) and [MS](#), have a very different energy resolution from either electrons or photons.

The goal of this method is to predict the E_T^{miss} distribution of the $Z/\gamma^* + \text{jets}$ background using a manipulated $\gamma + \text{jets}$ sample from data. These differences between $Z+\text{jet}$ and $\gamma+\text{jet}$ events can be broken down into two categories: differences which affect the jet energy and measurement, and differences which affect the boson energy and measurement. The differences in the hadronic system are simpler, and mostly consist of different numbers and energies of jets between the two samples, which can be accounted for via reweighting in a variable

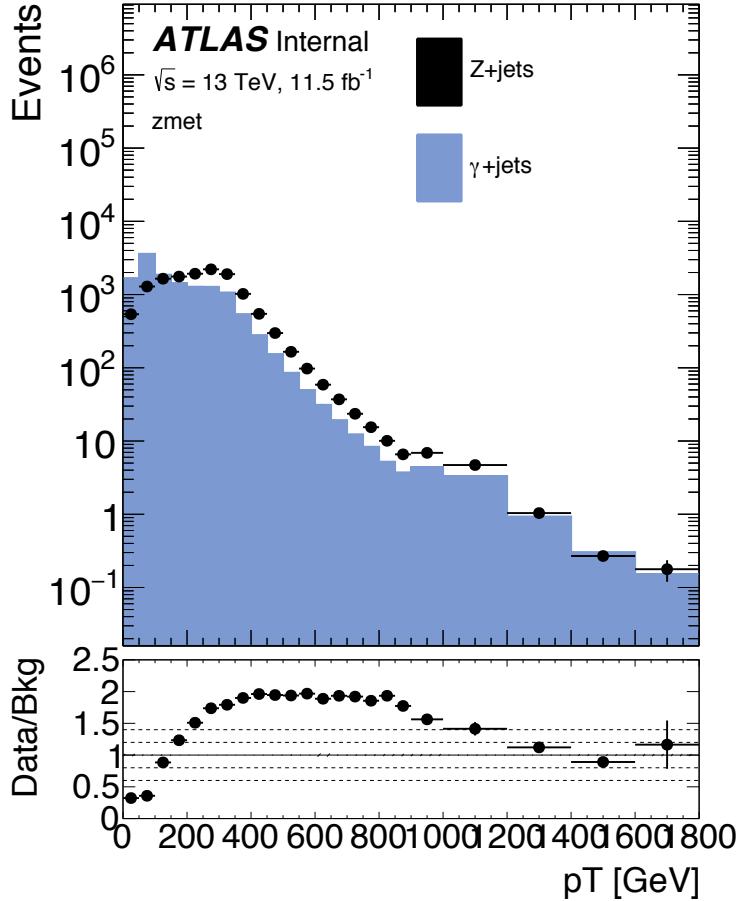


Figure 66: MC comparison of boson p_T in a selection of photon and $Z \rightarrow \ell\ell$ events with $H_T > 600$ GeV. The feature at 1000 GeV is not physical, but a reflection of the change in bin size.

that's representative of the total energy scale of the event. The differences in the bosons are more complex, and require the application of smearing functions based on the different observed objects. Together these corrections allow for complete modeling of the $Z/\gamma^* + \text{jets}$ E_T^{miss} spectrum with γ +jet events.

10.2.1 Photon and Z Event Selection

The baseline photon events come from an inclusive CR with no E_T^{miss} cut, a lepton veto, and the requirement of at least one photon, which is called CR- γ and defined in Table 7³. This selection is very pure in γ +jet events, but some $V\gamma$ events are also included, which can

³ This region includes an H_T cut, which requires the translation of photon p_T into an equivalent di-lepton p_T scalar sum. This process is described in Section 10.2.4.

include real E_T^{miss} . These backgrounds are subtracted from the yield in this region according to MC.

The triggering scheme for these $\gamma + \text{jets}$ events is more complicated than in other regions because the lowest unprescaled photon trigger requires a photon p_T of at least 120 (140) GeV in 2015 (2016) datataking, but the method requires events with much lower p_T to predict the full Z-boson p_T spectrum. To accomplish this, the lower- p_T photons are broken down into small p_T ranges with a different prescaled trigger required in each range, listed in Table 13. The events in each selection are then weighted by the prescale value of the trigger used to reconstruct a smooth p_T spectrum.

p_T Range [GeV]	Trigger Name
2015 Data-Taking	
$37 < p_T < 45$	HLT_g35_loose_L1EM15
$45 < p_T < 50$	HLT_g40_loose_L1EM15
$50 < p_T < 55$	HLT_g45_loose_L1EM15
$55 < p_T < 125$	HLT_g50_loose_L1EM15
$p_T > 125$	HLT_g120_loose_L1EM15
2016 Data-Taking	
$25 < p_T < 30$	HLT_g20_loose_L1EM12
$30 < p_T < 40$	HLT_g25_loose_L1EM12
$40 < p_T < 45$	HLT_g35_loose_L1EM12
$45 < p_T < 50$	HLT_g40_loose_L1EM12
$50 < p_T < 55$	HLT_g45_loose_L1EM12
$55 < p_T < 65$	HLT_g50_loose_L1EM12
$65 < p_T < 75$	HLT_g60_loose_L1EM12
$75 < p_T < 85$	HLT_g70_loose_L1EM12
$85 < p_T < 105$	HLT_g80_loose_L1EM12
$105 < p_T < 145$	HLT_g100_loose_L1EM12
$p_T > 145$	HLT_g140_loose_L1EM12

Table 13: List of triggers used to collect photon events in 2015 and 2016 data-taking.

These γ events can then be compared to baseline $Z \rightarrow \ell\ell$ events with a similar selection. These events have the same dilepton requirements as SRZ, without the $m_{\ell\ell}$ cut. They also have no E_T^{miss} cut, but like the photons, are required to have $H_T > 600$ GeV as in SRZ.

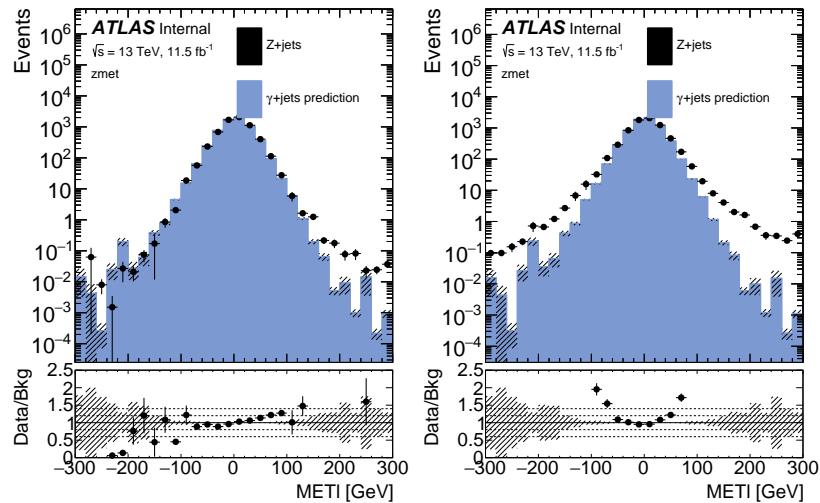
10.2.2 Smearing of Photon Events

While $Z + \text{jet}$ events are measured as a pair of leptons recoiling against a hadronic system, $\gamma + \text{jet}$ events are measured only as one object recoiling against jets. In addition, detector resolution is very different for muons and photons, though electrons and photons are similar. The impact of these differences must be corrected for in $\gamma + \text{jet}$ events in order for them to accurately predict the E_T^{miss} distribution of the Z s. Luckily, in most cases, the resolution of the photon's p_T is better than that of the Z boson, so the photon events can be smeared to emulate the Z s.

To isolate mismeasurement of boson p_T , this method uses $E_{T,\parallel}^{\text{miss}}$, defined as

$$E_{T,\parallel}^{\text{miss}} = E_T^{\text{miss}} \sin\theta \quad (44)$$

where θ is the angle between p_T^{miss} and the p_T of the Z boson. [Figure 67](#) shows the $E_{T,\parallel}^{\text{miss}}$ distribution in MC for the two samples, and demonstrates the discrepancies between them. The core of the photon distribution is similar to the $Z \rightarrow ee$ distribution because measurements of both photons and electrons are primarily taken from the electromagnetic calorimeter and have the same resolution. For muons, which rely on tracks to determine p_T , the resolution becomes larger at high p_T values where the tracks are nearly straight. For a muon with a p_T of 200 GeV, this the resolution is about 5%, while a photon or electron with the same p_T has a resolution of 1-2% [94, 95]. As a consequence, the resolution distributions for photon and $Z \rightarrow \mu\mu$ events are very different.



[Figure 67](#): $E_{T,\parallel}^{\text{miss}}$ distributions in MC for $Z + \text{jets}$ in the ee (left) and $\mu\mu$ (right) channels compared to the $\gamma + \text{jets}$ prediction in an inclusive region with $H_T > 600$ GeV.

A function to smear photon events is derived from the deconvolution of the photon and lepton response functions, taken from 1-jet CRs with no H_T cut, which are otherwise identical to the baseline Z and γ selections. In these regions, events are binned in boson p_T , and in each bin, a $E_{T,\parallel}^{\text{miss}}$ distribution is made. The smearing function is defined for each bin as the deconvolution of the $Z/\gamma^* + \text{jets}$ and $\gamma + \text{jets}$ $E_{T,\parallel}^{\text{miss}}$ distributions.

Next, for each photon event, the smearing function matching the event's photon p_T is sampled, yielding a smearing factor Δp_T . The photon's p_T is then adjusted according to

$$p_T^{\gamma'} = p_T^\gamma + \Delta p_T \quad (45)$$

and the corresponding change in E_T^{miss} is made,

$$E_{T,\parallel}^{\text{miss}'} = E_{T,\parallel}^{\text{miss}} - \Delta p_T. \quad (46)$$

The nominal smearing function is taken from MC in order to remove contamination from other backgrounds, and the resulting $E_{T,\parallel}^{\text{miss}}$ distributions can be seen in Figure 68. Though there is a small amount of over-smearing in the negative tail, the improvement in agreement between the distributions, especially in their cores, is clear.

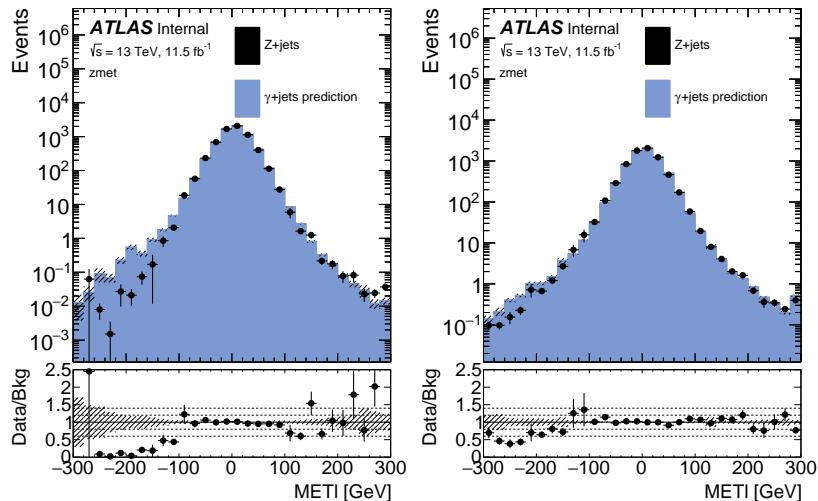


Figure 68: $E_{T,\parallel}^{\text{miss}}$ distributions in MC for $Z + \text{jets}$ ee (left) and $\mu\mu$ (right) channels compared to $\gamma + \text{jets}$ in an inclusive region with $H_T > 600$ GeV after the smearing procedure has been performed. The events in these distributions have also been p_T reweighted, as described in Section 10.2.3.

10.2.3 p_T Reweighting of Photon Events

Next, the photon events are reweighted to match the boson p_T spectrum of the Z events. This is accomplished by making histograms of

boson p_T for γ and Z events with binning identical to the p_T bins used in Section 10.2.2. Photons are binned based on their smeared p_T determined in the previous step. A reweighting factor $f(p_T)$ is then calculated in each bin, according to

$$f_{\text{MC}}(x) = \frac{N_{Z/\gamma^*+\text{jets}}(x)}{N_{\gamma+\text{jets}}(x)} \quad (47)$$

in MC, and in data according to

$$f_{\text{data}}(x) = \frac{N_{\text{data}}(x) - N_{t\bar{t}}(x) - N_{VV}(x)}{N_{\gamma+\text{jets data}}(x)} \quad (48)$$

where the contamination from other backgrounds is taken from MC and subtracted from the Z candidate sample. The resulting reweighting factors can be seen in Figure 69 and are calculated independently for ee and $\mu\mu$ events.

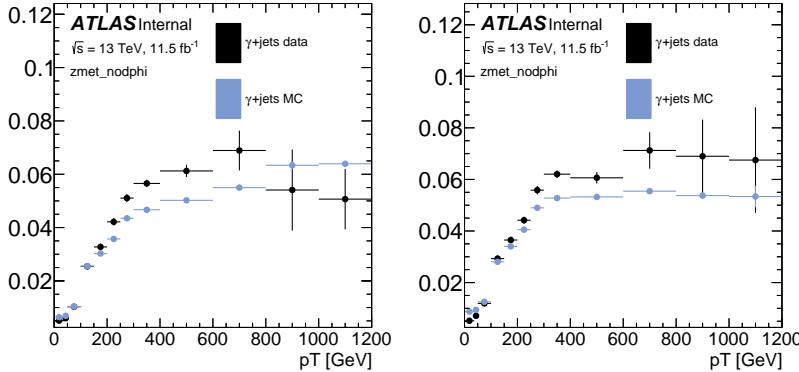


Figure 69: Photon reweighting factors for the ee (left) and $\mu\mu$ (right) channels derived from data and MC.

The data- and MC-derived reweighting functions each have their own benefits. The data-based functions are of course not subject to any potential mismodeling, but it is possible to have contamination from other processes that is not accounted for. Because good agreement between data and MC is seen in the boson p_T spectrums, contamination was deemed a higher concern than mismodeling, and the MC-based function was used as the nominal smearing function, with the data-based function used as a cross-check.

This reweighting, though it is performed on the boson p_T , primarily serves to produce more similar jet distributions between the γ and Z samples. Because, excluding E_T^{miss} contributions, the boson p_T must match the energy of the jet system off which it recoils, these two variables are closely tied. Once the two samples are reweighted to have similar amounts of hadronic energy, the E_T^{miss} contribution from mismeasurement of jets, which scales according to the square root of the jet energy, should also be the same between the two samples.

Together, the boson smearing and p_T reweighting produce a E_T^{miss} spectrum in the modified photon events that closely matches that of the Z events. Figures 70 and 71 show the comparison of the E_T^{miss} distributions before any alteration, with only p_T reweighting, and after the smearing and reweighting, demonstrating the impact of each step. In the ee channel, the E_T^{miss} distribution looks similar between $Z/\gamma^* + \text{jets}$ and $\gamma + \text{jets}$ events up to about 150 GeV. Above this value, the differences between γ and $Z \rightarrow ee$ resolutions seen in the tails of Figure 67 become significant. In the mm channel, the discrepancy between the E_T^{miss} in $Z/\gamma^* + \text{jets}$ and $\gamma + \text{jets}$ events begins at much lower values. In both cases, the application of p_T reweighting and smearing result in excellent agreement between the two E_T^{miss} spectra, with most of the improvement resulting from the smearing factors.

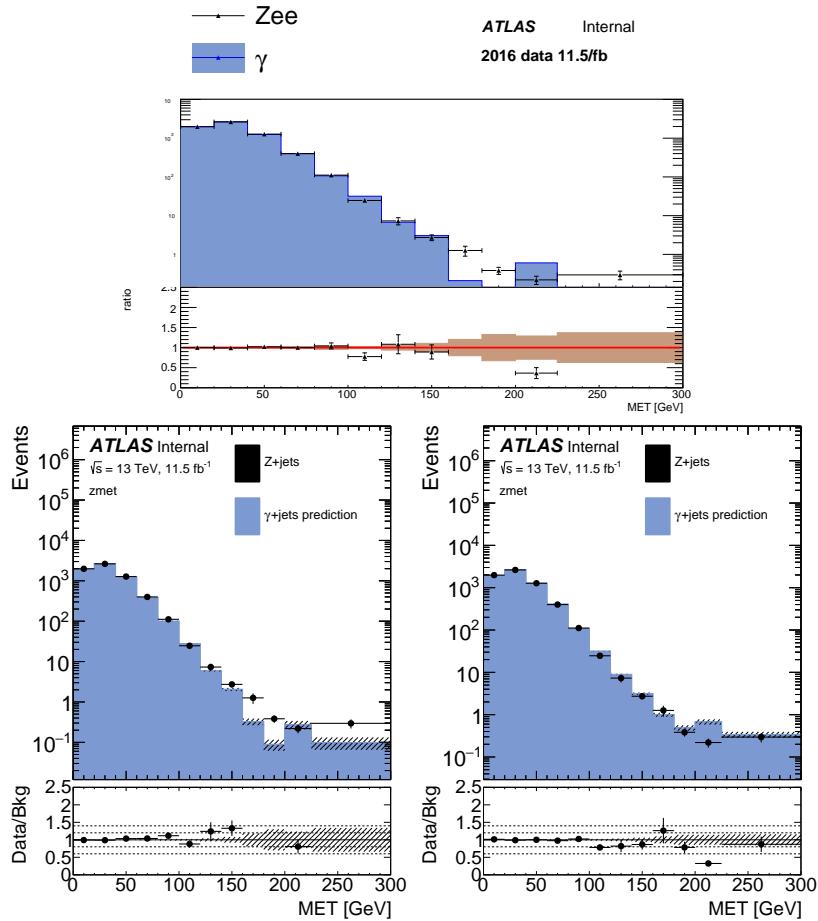


Figure 70: E_T^{miss} distribution comparing MC distributions of photon and Z events before any smearing is applied (top), with only p_T reweighting applied (bottom left), and after p_T reweighting and smearing have both been applied (bottom right) in the ee channel of 2016 data.

Finally, with the E_T^{miss} distribution finalized in the $\gamma + \text{jets}$ prediction, the $\Delta\phi(\text{jet}_{12}, p_T^{\text{miss}})$ can be applied at this stage.

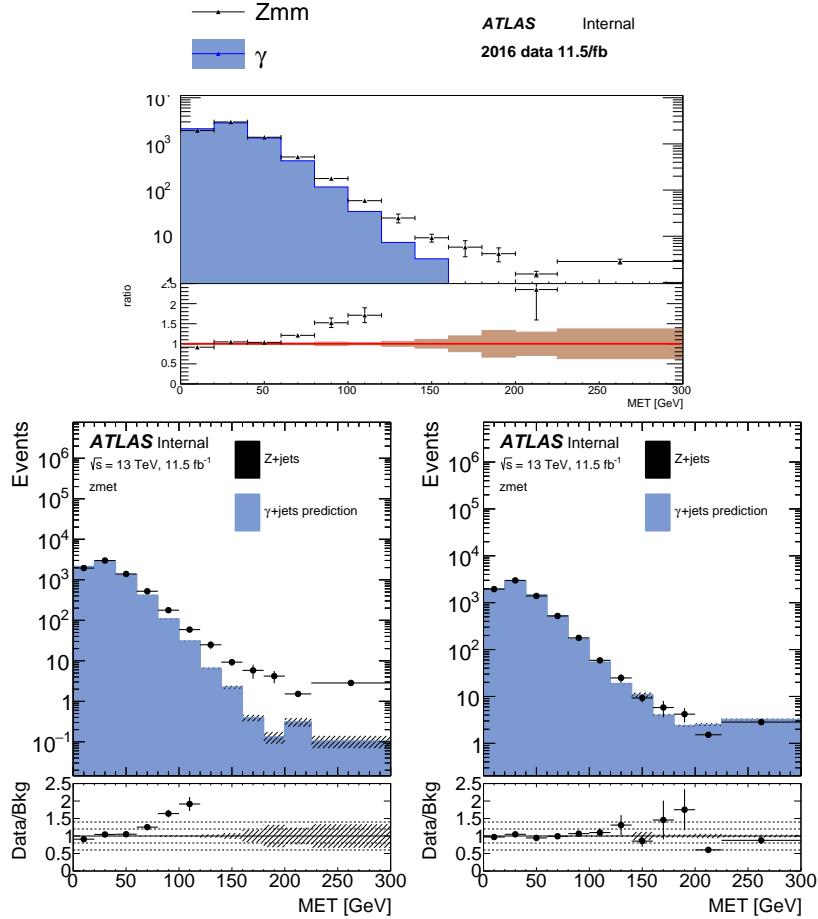


Figure 71: $E_{\text{T}}^{\text{miss}}$ distribution comparing MC distributions of photon and Z events before any smearing is applied (top), with only p_{T} reweighting applied (bottom left), and after p_{T} reweighting and smearing have both been applied (bottom right) in the $\mu\mu$ channel of 2016 data.

10.2.4 Determining H_{T} and $m_{\ell\ell}$

One complication thus far ignored is that CR- γ has no leptons, but some quantities that define the SR require them, namely $m_{\ell\ell}$ and H_{T} . There is no analogue of $m_{\ell\ell}$ in photon events, and though the contribution of leptons to the H_{T} variable could be related to the p_{T} of the boson in $\gamma + \text{jets}$ events, the relationship depends on the angular distributions between lepton pairs in $Z/\gamma^* + \text{jets}$ events and must account for the mass of the Z boson. Instead of determining an analytic function to describe such a relation, both of these variables are determined by creating histograms binned in the boson p_{T} and sampling.

In the case of H_{T} , distributions of the scalar sum of the p_{T} of the leading leptons are made for each Z p_{T} bin using $Z \rightarrow \ell\ell$ MC. A sampled value from the distribution is then added to the H_{T} of the jets in a photon event to produce the final estimate. This sampling is

done before any reweighting is performed because the H_T is needed to make the preselection for the reweighting process. However, the smearing is performed inclusively in H_T , so this procedure can be performed using the smeared photon p_T to choose the distribution to sample.

The $m_{\ell\ell}$ determination is performed after both the smearing and reweighting, and is tied closely to the smearing step. Mismeasurements in lepton p_T can create E_T^{miss} in a $Z/\gamma^* + \text{jets}$ event, but the same event is likely to migrate off the Z $m_{\ell\ell}$ window due to the mismeasured lepton. Thus it is very important that the two effects be carefully correlated in the manipulated photon events. To achieve this, MC Z events from the 1-jet CR described in Section 10.2.2 are used to make two-dimensional distributions of $m_{\ell\ell}$ as a function of the difference between reconstructed and true Z p_T for the ee and $\mu\mu$ channels. A photon event then uses the Δp_T assigned to it during the smearing process to index the distribution, and an $m_{\ell\ell}$ value is sampled from the corresponding bin⁴.

To test the soundness of this procedure, it is repeated purely in MC, and the results of the MC prediction and the data prediction are compared to the $m_{\ell\ell}$ distribution in $Z/\gamma^* + \text{jets}$ MC in Figure 72. Though there are some areas of disagreement across the $m_{\ell\ell}$ spectrum, very good closure is seen on the on- Z bin, which is what is needed for the estimate of SRZ’s backgrounds. After the $m_{\ell\ell}$ distribution has been emulated, a cut requiring that the photon events be “on the Z -mass peak” can be required.

10.2.5 Subtraction of $V\gamma$ Events

At high E_T^{miss} , where the signal region lies, contamination of CR- γ with $V\gamma$ events becomes significant, as shown in Figure 73. These events must be subtracted from the $\gamma + \text{jets}$ prediction because, once the photons are corrected to approximate Z s, they essentially provide a (not very accurate) prediction of diboson events, which are already accounted for in another background estimate.

This subtraction accomplished by performing the $\gamma + \text{jets}$ method on $V\gamma$ MC to approximate these backgrounds’ contribution to the final E_T^{miss} distribution. This contribution is then subtracted from the $\gamma + \text{jets}$ prediction, the impact of which can be seen in Figure 74. As expected, the impact is greatest at high E_T^{miss} where these backgrounds are most significant.

⁴ Ideally this Δp_T would also include the difference between the true and reconstructed p_T of the photon events, but this information is of course not accessible in data. Luckily, in the events in the final SR this value is typically negligible compared to the Δp_T from smearing, so the impact is small.

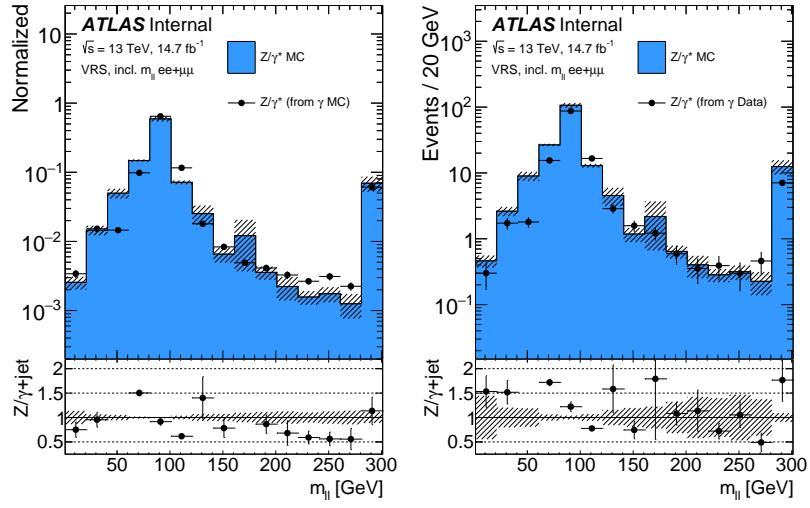


Figure 72: $Z/\gamma^* + \text{jets}$ MC $m_{\ell\ell}$ distribution compared to the prediction from $\gamma + \text{jets}$ method performed on MC (left) and the prediction from $\gamma + \text{jets}$ method performed on data (right).

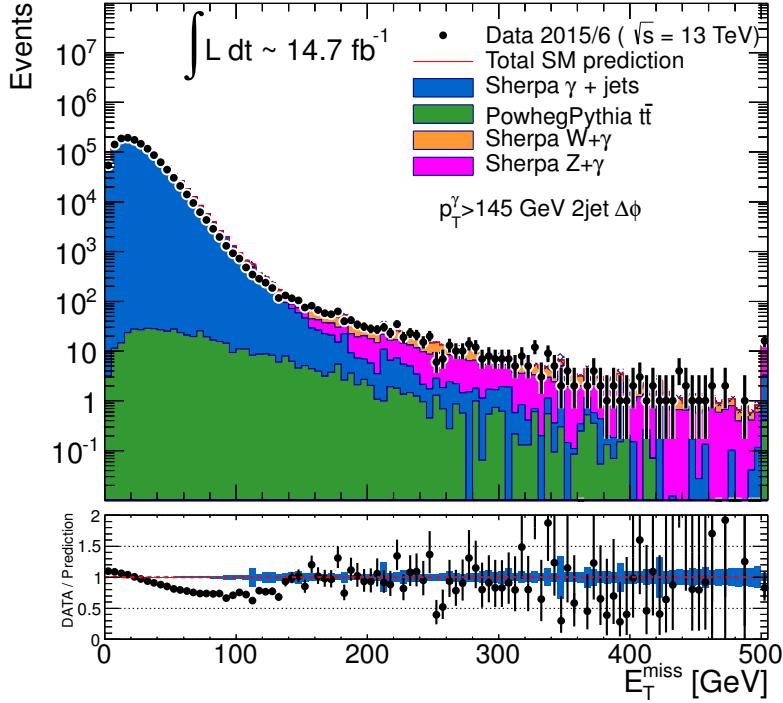


Figure 73: Comparison of data and MC in CR- γ without any H_T cut, including the contributions from various $V\gamma$ processes.

10.2.6 Validation in Data

The $\gamma + \text{jets}$ jets method is validated in a region called VRZ, defined in Table 7, which is similar to SRZ, but with an inverted E_T^{miss} cut. Figure 75 shows the low- E_T^{miss} portion of this VR where the $Z/\gamma^* +$

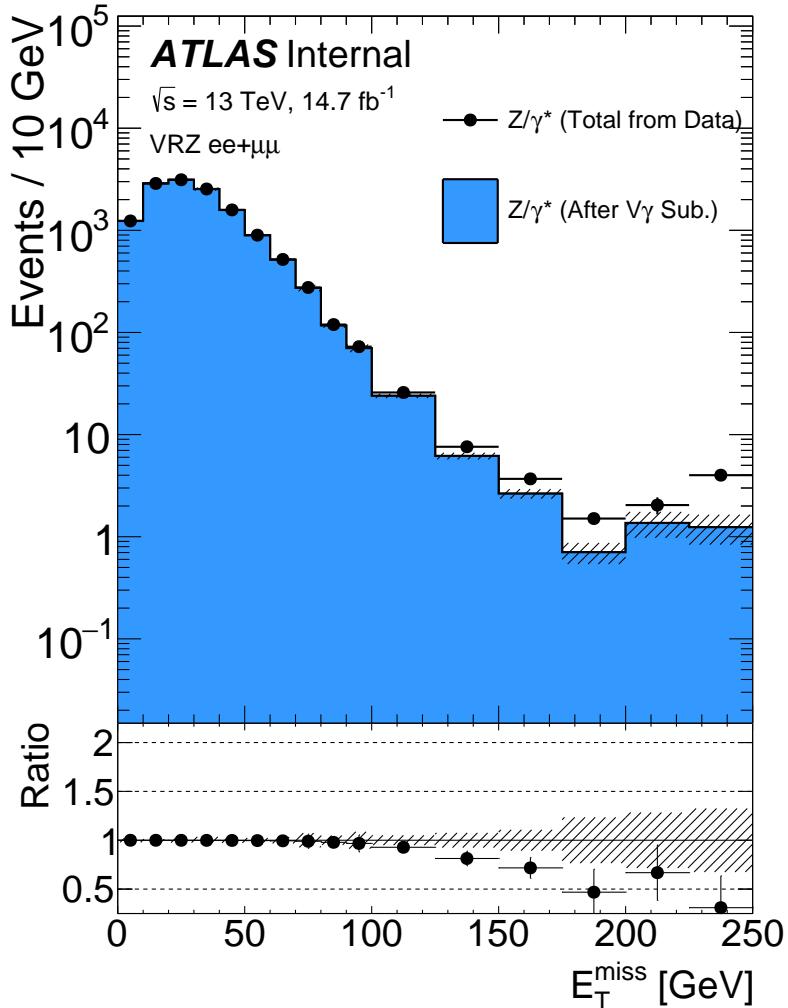


Figure 74: Total $\gamma + \text{jets}$ data prediction in SRZ (excluding the E_T^{miss} cut) and the prediction after the $V\gamma$ subtraction.

jets background is dominant. Here, the three data-driven background estimates, as well as the remaining MC backgrounds are stacked and compared to the data yield in this region, demonstrating excellent agreement across a wide E_T^{miss} range.

An additional check can be made in VRZ by removing the $\Delta\phi(\text{jet}_{12}, \mathbf{p}_T^{\text{miss}})$ intended to suppress the $Z/\gamma^* + \text{jets}$ background from jet mismeasurement. Figure 76 shows the distribution of this variable in VRZ, and demonstrates that, even at low values where the $Z/\gamma^* + \text{jets}$ background is dominant, the $\gamma + \text{jets}$ method models it accurately.

10.3 FAKE AND NON-PROMPT LEPTONS

The *fakes* background consists of processes that produce only one lepton, but whose events are otherwise kinematically similar to the SR. These processes include semileptonic $t\bar{t}$, $W + \text{jets}$, and single top pro-

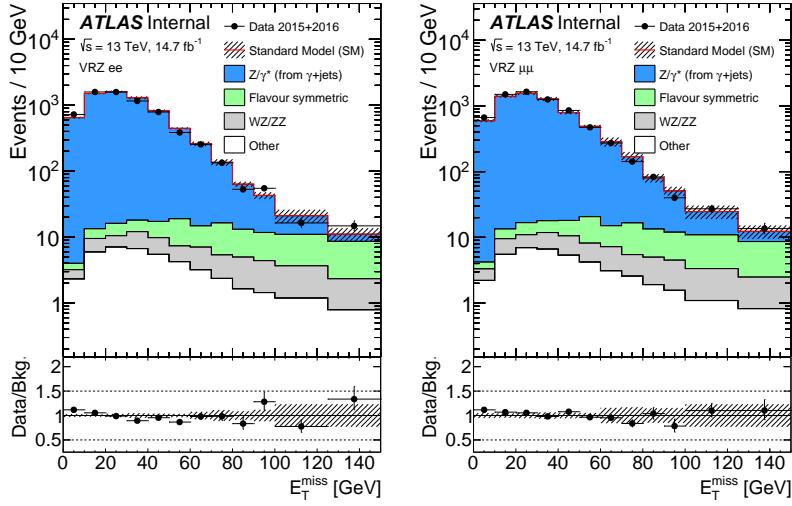


Figure 75: E_T^{miss} distribution in VRZ ee (left) and $\mu\mu$ (right) with total data yield compared to the sum of the prediction from the $\gamma + \text{jets}$ method, the prediction from the flavor symmetry method, the prediction from the fake background estimation (included under “other”), and the remaining backgrounds taken from MC.

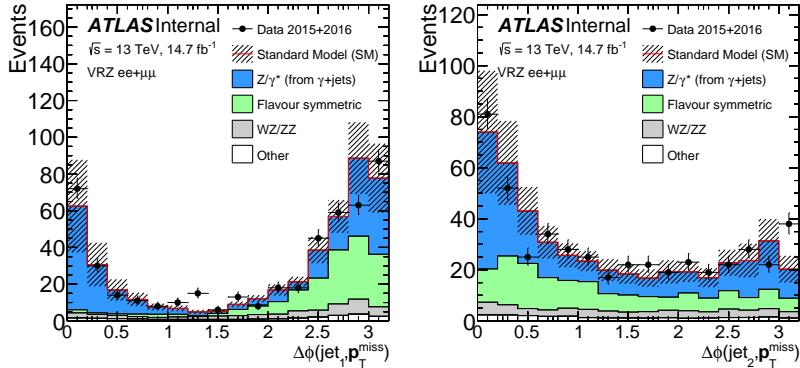


Figure 76: $\Delta\phi(\text{jet}, p_T^{\text{miss}})$ distribution in for the leading jet (left) and the sub-leading jet (right). The comparison is performed in VRZ with the cut on $\Delta\phi(\text{jet}_{12}, p_T^{\text{miss}})$ removed. The total data yield is compared to the sum of the prediction from the $\gamma + \text{jets}$ method, the prediction from the flavor symmetry method, the prediction from the fake background estimation (included under “other”), and the remaining backgrounds taken from MC.

cesses. Though these processes typically only produce one lepton, they can be reconstructed with two leptons due to a hadron being misidentified as a lepton or due to a real non-prompt lepton resulting from photon conversions or B -hadron decays. As with the $Z/\gamma^* + \text{jets}$ background, it is very difficult to predict with MC because the flaws in reconstruction are typically less well described by the models used in MC production than the successes. Nonetheless, a rough estimate

can be made of this background by using **MC**, which indicates that the number of fake events in SRZ is consistent with zero.

Despite the small predicted contribution in the **SR**, a data-driven method called the *matrix method* is employed to estimate these fake events [96]. This method is also used to estimate the fakes contribution to other control and validation regions where their impact is more significant.

In the matrix method, the quality requirements for signal leptons are loosened to give a selection of baseline leptons (see [Table 3](#) and [Table 4](#)), which consist of a higher fraction of fake leptons. In each **CR**, **VR**, or **SR**, the remaining kinematic selections are made on the baseline leptons, and the number of leptons in the region which pass the signal lepton requirements (N_{pass}) and the number which fail (N_{fail}) are measured. For a 1-lepton selection, these quantities can be used to predict the number of fake events that pass the selection according to:

$$N_{\text{pass}}^{\text{fake}} = \frac{N_{\text{fail}} - (1/\epsilon^{\text{real}} - 1) \times N_{\text{pass}}}{1/\epsilon^{\text{fake}} - 1/\epsilon^{\text{real}}}. \quad (49)$$

The efficiencies ϵ^{real} and ϵ^{fake} give the relative identification efficiency from baseline to signal for genuine, prompt leptons and fake and non-prompt leptons, respectively. For a 2-lepton selection, the principle is the same, but the equation is more complicated, requiring a four-by-four matrix to account for possible combinations of real and fake leptons.

To calculate ϵ^{real} , the tag-and-probe method is performed a selection of $Z \rightarrow \ell\ell$ data events, CR-real, described in [Table 14](#). In this method, one *tag* lepton passing a signal selection is required, as is another *probe* lepton passing a baseline requirement. Distributions in $m_{\ell\ell}$ for events with a tag and a passing probe and events with a tag and a failing probe are produced and fit, and the efficiency is computed using the ratio acquired from the fit. A comparison of data and **MC** in CR-real can be seen in [Figure 77](#).

The fake efficiency, ϵ^{fake} , is determined using the tag-and-probe method in CR-fake, also described in [Table 14](#). This region is different from all other regions considered in this analysis because it requires same-sign leptons. Very few processes genuinely produce two same-sign leptons, so this region is enhanced in fake leptons. An upper limit on $E_{\text{T}}^{\text{miss}}$ is placed on CR-fake to limit the possible contamination from **BSM** processes. According to **MC**, real, prompt leptons make up about 7% (11%) of the baseline electron (muon) sample and about 10% (61%) of the signal electron (muon) sample in this region. These real lepton backgrounds are subtracted from the CR-fake yields when calculating the efficiencies. [Figure 78](#) shows a comparison of data and **MC** in this region.

This method is validated in a fakes-rich validation region with a same-sign lepton requirement, $E_{\text{T}}^{\text{miss}} \geq 50\text{GeV}$, ≥ 2 jets, and a veto on

Fakes regions	E_T^{miss} [GeV]	H_T [GeV]	n_{jets}	$m_{\ell\ell}$ [GeV]	SF/DF	OS/SS	n_{ℓ}
CR-real	–	> 200	≥ 2	81–101	2ℓ SF	OS	2
CR-fake	< 125	–	–	> 12	2ℓ SF/DF	SS	≥ 2

Table 14: Control regions used to measure efficiencies of real and fake leptons. The flavour combination of the dilepton pair is denoted as either “SF” for same-flavour or “DF” for different flavour. The charge combination of the leading lepton pairs are given as “SS” for same-sign or “OS” for opposite-sign.

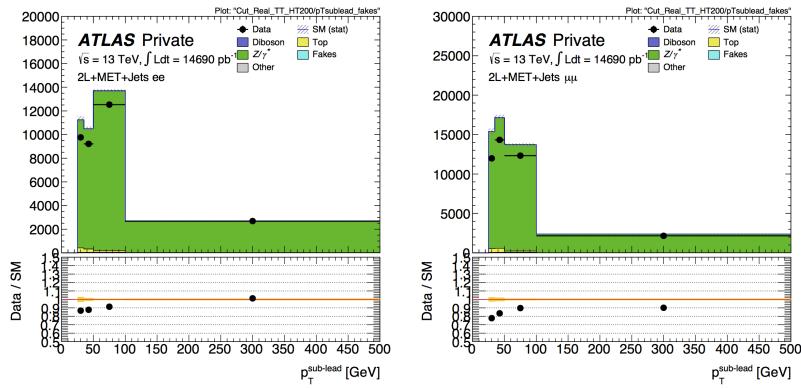


Figure 77: Sub-leading lepton p_T for ee (left) and $\mu\mu$ (right) events in the tight-tight region used to measure the real-lepton efficiency for 2016.

$m_{\ell\ell}$ on the Z-mass peak for same flavor channels. The results of this validation can be seen in Figure 79. With the systematic uncertainties, discussed in Section 11.1.3, the prediction agrees well with the data across a wide range of $m_{\ell\ell}$ values.

10.4 DIBOSON AND RARE TOP PROCESSES

The remaining backgrounds are diboson processes (excluding WW, which is included in the FS background) and rare top processes. Dibosons events make up about 30% of the events in SRZ, while rare top process contributions are much smaller. Both are taken directly from MC, with validation regions to confirm the accuracy of the prediction. These regions are described in Table 7, and target different parts of these backgrounds. VR-ZZ is a four-lepton selection designed to select a very pure sample of ZZ events. VR-WZ requires three leptons and makes specific cuts on m_T , the transverse mass, and E_T^{miss} in order to select mostly $WZ \rightarrow ll\nu\nu$ events. VR-3L is similar to VR-S, but

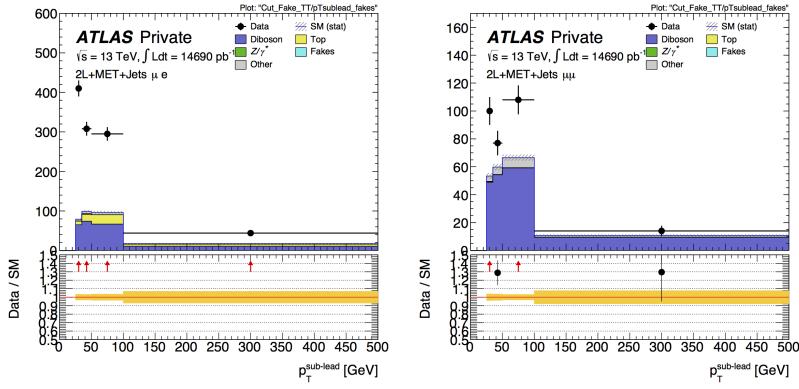


Figure 78: Sub-leading lepton p_T for μe (left) and $\mu \mu$ (right) events in the tight-tight region used to measure the fake-lepton efficiency for 2016.

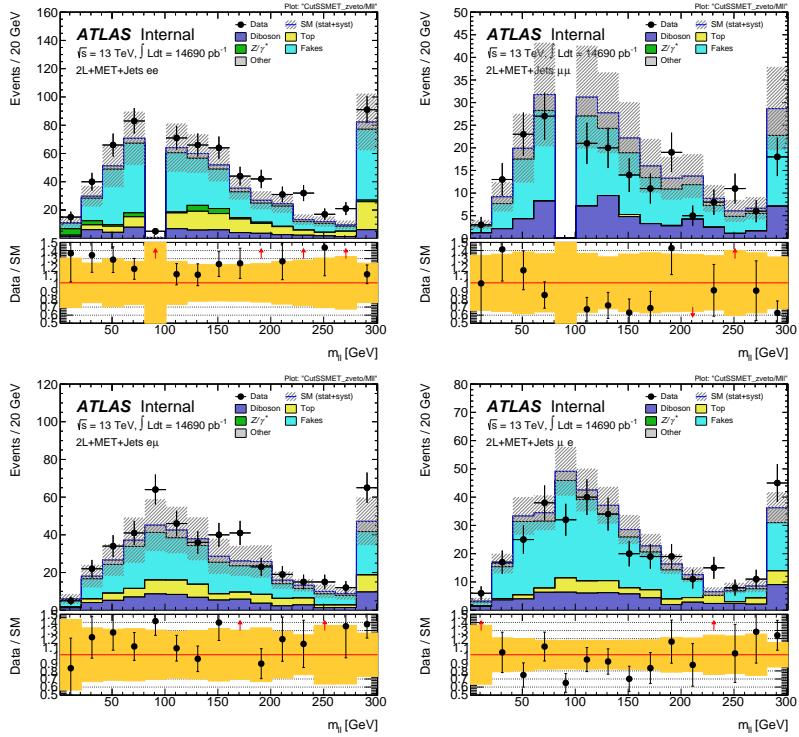


Figure 79: Same sign validation regions in the ee (top left), $\mu\mu$ (top right), $e\mu$ (bottom left) and μe (bottom right) channels combining 2015+2016 data. Uncertainty bands include both statistical and systematic uncertainties.

loosens the H_T and E_T^{miss} cuts and requires at least three leptons. This region is designed to target any ≥ 3 -lepton process in a region as kinematically close to SRZ as possible while still maintaining enough events to validate. The makeups of these multilepton validation regions, as well as VRS, are shown in Table 15.

To confirm that the kinematics are well modeled in the diboson validation regions, distributions of boson mass and p_T are shown in

	VR-S	VR-WZ	VR-ZZ	VR-3L
Observed events	236	698	132	32
Total expected background	224 ± 41	613 ± 66	139 ± 25	35 ± 10
Flavour-symmetric	99 ± 8	-	-	-
WZ/ZZ events	27 ± 13	573 ± 66	139 ± 25	25 ± 10
Rare top events	11 ± 3	14 ± 3	0.44 ± 0.11	9.1 ± 2.3
Z/ γ^* + jets events	84 ± 37	-	-	-
Fake lepton events	4 ± 4	26 ± 6	-	0.6 ± 0.3

Table 15: Yields in validation regions. In VRS, data-driven background estimates are used for Z/γ^* + jets, fakes, and [FS](#) processes. All other backgrounds are taken from [MC](#), including all backgrounds in the multi-lepton [VRS](#). Uncertainties include statistical and systematic components.

[MC](#) and data. Figures 80 and 81 show these distributions for VR-WZ, and Figure 82 shows these distributions for VR-ZZ.

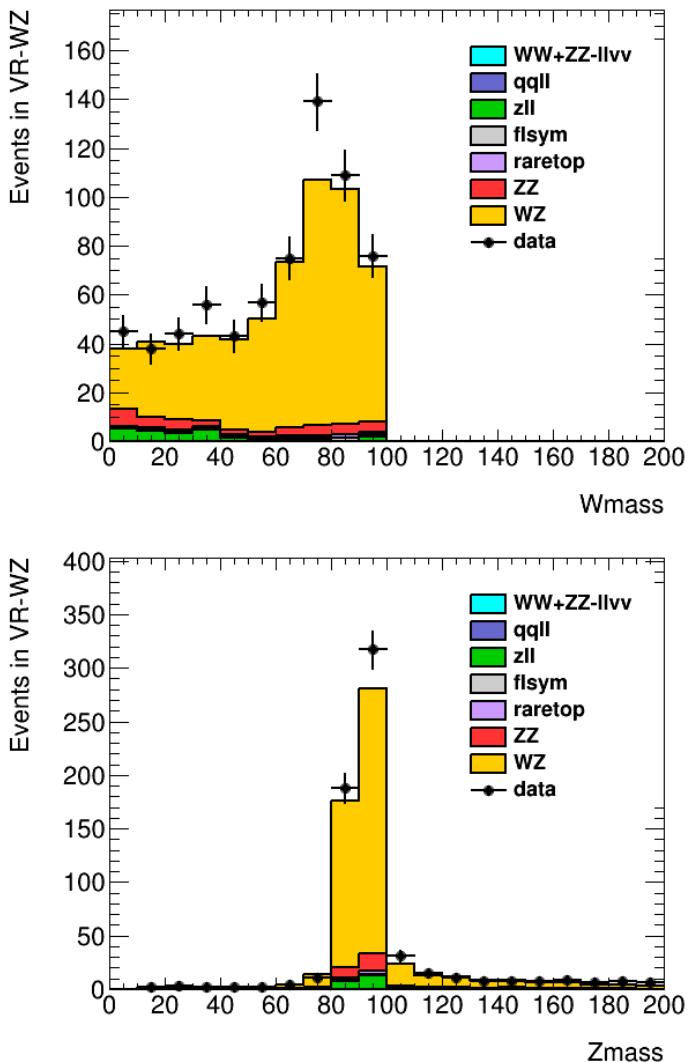


Figure 8o: Distribtuions of data and MC in VR-WZ. Reconstructed transverse mass of the W (top) and mass of the Z (bottom).

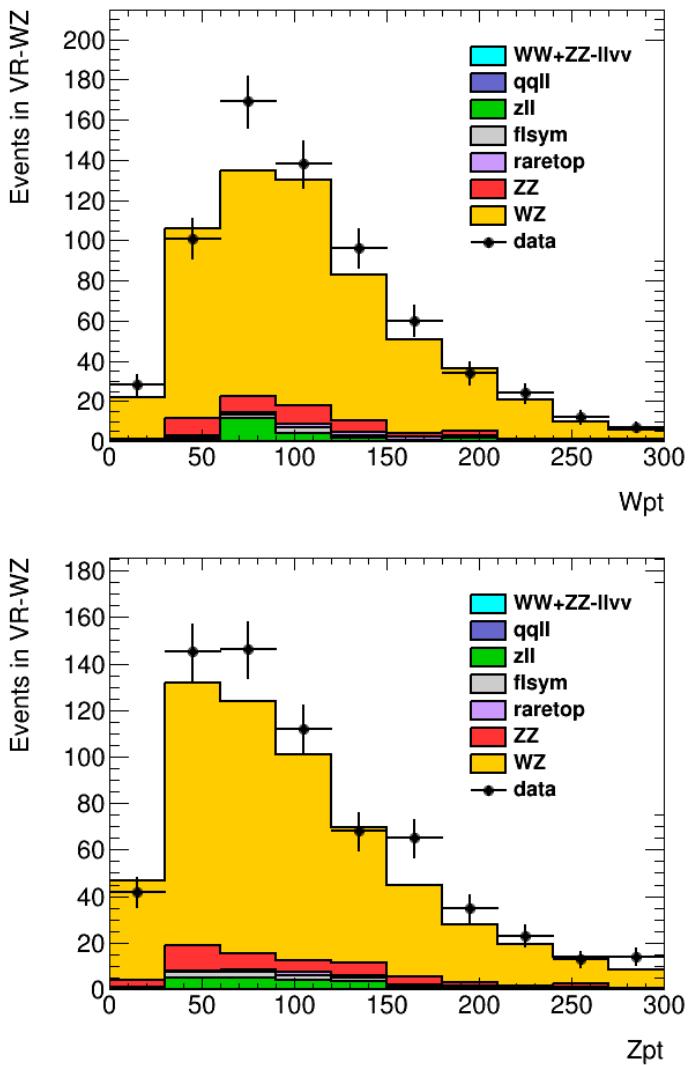


Figure 81: Distribtuions of data and MC in VR-WZ. p_T of the W (top) and Z (bottom).

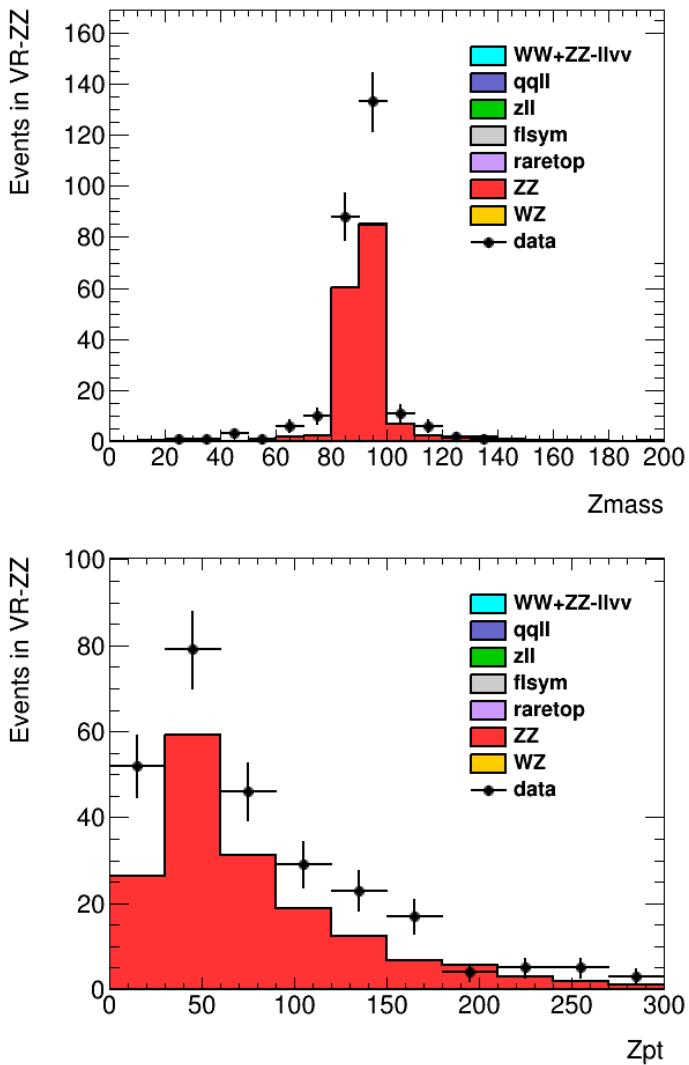


Figure 82: Distribtuions in VR-WZ. On the top, mass of the Z bosons in the event, and on the bottom, p_T of the Z bosons.

SYSTEMATIC UNCERTAINTIES

11.1 UNCERTAINTIES ON DATA-DRIVEN BACKGROUNDS

11.1.1 Uncertainties on the Flavor Symmetry Method

The flavor symmetry method is a data driven method that makes its estimate primarily on based events populating an **SR**-like **CR** in the different-flavor channel. The statistical uncertainty on these events makes up the dominant uncertainty on the method. To reduce this uncertainty, the $m_{\ell\ell}$ range on the **CR** is expanded, approximately tripling the number of events in CR-FS. The statistical uncertainty is reduced by this expansion, though it is still significantly higher than any of the other systematic uncertainties on this method, as seen in [Table 16](#). Also included in the statistical uncertainty column is the uncertainty on the number of non-**FS** events in CR-FS, which is used to scale the prediction to account for contamination in the **CR**.

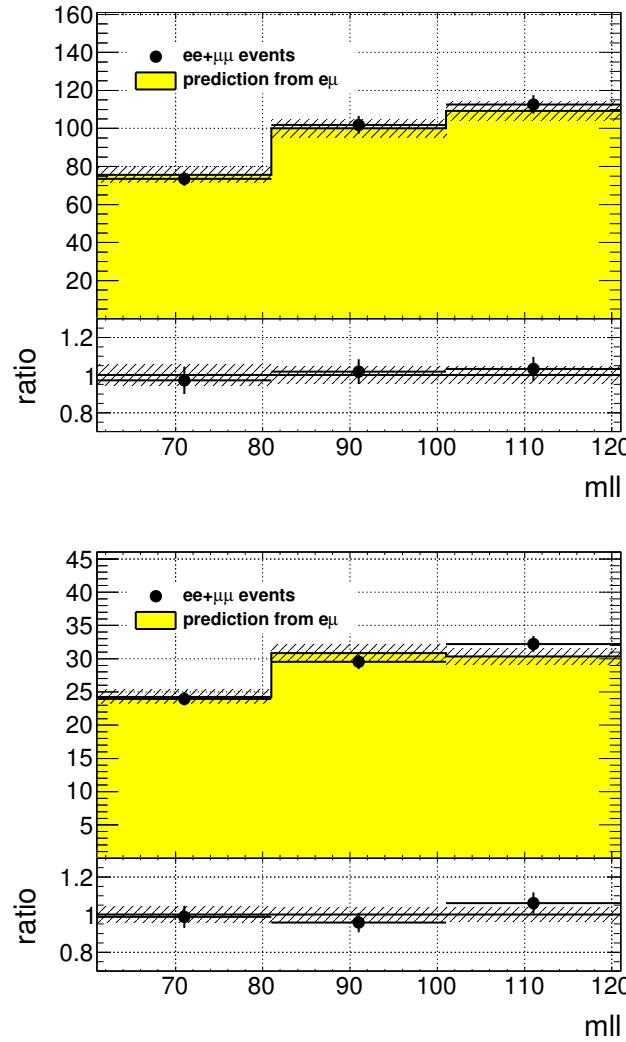
Reg.	Ch.	Pred.	Uncertainties					
			stat. clos.	MC and α	k and α	dAOD usage	$m_{\ell\ell}$ shape	total
SRZ	ee	16.50	1.82	0.88	0.53	0.12	0.22	2.11
	$\mu\mu$	16.67	1.83	0.79	0.33	0.11	0.23	2.04
	$ee+\mu\mu$	33.16	3.66	1.07	0.86	0.23	0.45	3.94
VRS	ee	49.70	3.21	2.34	2.20	0.34	0.75	4.61
	$\mu\mu$	49.60	3.14	2.88	1.40	0.31	0.75	4.56
	$ee+\mu\mu$	99.31	6.34	4.00	3.60	0.65	1.49	8.47

[Table 16](#): Uncertainties in the on-Z signal and validation regions. Nominal predictions are given with statistical uncertainty (including uncertainty from subtracted backgrounds), MC Closure uncertainty, uncertainty on the prediction from varying k and α by their statistical uncertainties, comparing the efficiencies from AODs to that of DAODs, and on the $m_{\ell\ell}$ widening, which includes MC statistics and a data/MC comparison in a loosened region.

The next largest contribution to the uncertainty comes from **MC** closure tests, which are used to determine how effective the method is in its prediction. If, for example, using weights derived from an inclusive selection at high E_T^{miss} lead to a bias, the closure test would indicate that and an appropriate uncertainty could be placed on the

estimate based on the difference between the **MC** prediction and the prediction from the flavor symmetry method.

In this test, the entire **FS** procedure is performed on $t\bar{t}$ **MC**, including a recalculation of weighting factors α and k . The prediction from $e\mu$ events in **MC** is compared to the **MC** ee and $\mu\mu$ events, as seen in [Figure 83](#). The difference between the two predictions is then summed in quadrature with the statistical uncertainty on each prediction to give the total closure uncertainty seen in [Table 16](#). In these closure tests, all predictions agree within the statistical uncertainty, so the largest contributor to the resulting error is **MC** statistics.



[Figure 83](#): **MC** closure plots of VRS (top) and SRZ (bottom). The number of events from **MC** (black points) is compared to the number of events predicted from the flavor symmetry method (yellow histogram). The comparison is performed before the expanded $m_{\ell\ell}$ window is used to predict the on-Z bin, but because the shape is taken from the same **MC**, the result is identical.

A small uncertainty is added based on the statistical uncertainty on the k and α factors derived from data. These factors are measured in many different bins (see, for example, the different measurements of k in [Figure 64](#)), and as a consequence, some bins can have very large statistical uncertainties. To assess the uncertainty on the total estimate, each measurement of these factors is varied by its uncertainty in order to produce the maximum and minimum possible prediction. The differences with respect to the nominal prediction are used to create a symmetrized error, which is included in [Table 16](#).

The next uncertainty considers a potential bias in the way the α factors are calculated. Because they are derived from data, there is already trigger dependence in data collection; only events passing a trigger are stored. Additional trigger dependence is created by the data format used for analysis. [ATLAS](#) data and [MC](#) are stored in a format called Analysis Object Data ([AOD](#)), but smaller, slimmer versions of these datasets, called derived AODs ([dAODs](#)) are used for analysis. These [dAODs](#) are designed with specific analyses in mind, filtering on the triggers and objects required by the analyses. As a consequence, in the [dAOD](#) used in this analysis, there are explicit requirements that lepton or E_T^{miss} triggers are passed in order for events to be included.

As a consequence, the trigger efficiencies ϵ^{trig} used in [Equation 42](#) to define α do not consider all possible data events. The ϵ^{trig} factor is calculated for each trigger using events passing the kinematic selection for that trigger, outlined in [Section 9.1](#). The efficiency factor is then measured according to the equation

$$\epsilon^{\text{trig}} = \frac{N_{\text{trig}}}{N_{\text{all}}} \quad (50)$$

where N_{trig} is the number of events passing the trigger in the kinematic selection and N_{all} is all events in the selection. The latter measurement is the one subject to this bias, as it contains only the events that pass at least one trigger required for inclusion in the [dAOD](#). As a consequence of these missing events, the ϵ^{trig} values will be artificially high. However, because the ratio of trigger efficiencies for the different channels is the only quantity needed for this analysis, the missing events will only bias the prediction if the different channels are differently impacted by the trigger preselection.

Calculating the flavor symmetry method's dependence on these biases requires the use of [MC](#). With a generated [MC](#) sample, there is no trigger dependence, so an unskimmed sample can be compared to a typical skimmed [MC dAOD](#) to identify the effect of the skimming. [Figure 84](#) shows a comparison of the α factors calculated for different bins in E_T^{miss} from the nominal source, data, as well as these two [MC](#) sources. A E_T^{miss} dependence would be the most likely bias between the two [MC](#)-derived α factors because E_T^{miss} triggers are the only triggers besides lepton triggers that will allow an event to be accepted

into the dAOD used by this analysis. Though there is some difference between the data-derived α and those taken from MC, it is clear from this plot that there is very little dependence on the choice of an unskimmed or skimmed sample. The calculation of the uncertainty is performed by repeating the flavor symmetric method in MC with each of the two α factors and using the difference between the estimates as a symmetric error.

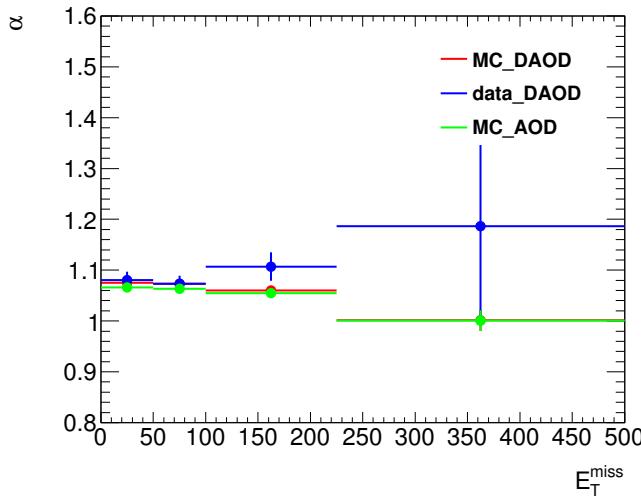


Figure 84: α , the trigger efficiency ratio, calculated as a function of E_T^{miss} from three different sources: data (blue), the usual skimmed $t\bar{t}$ MC (red), and an unskimmed $t\bar{t}$ MC (green).

The last uncertainty relates to the main MC dependence of the method - the $m_{\ell\ell}$ shape of the FS background. A correction factor is taken from MC in order to account for the $m_{\ell\ell}$ widening, and the accuracy of that factor must be checked. Its shape is compared to that of data in region similar to VR-FS, but with an H_T cut lowered to 300 GeV to increase statistics. The difference between the fraction of events on the Z-mass peak in data and MC in this region is taken as a systematic uncertainty. To confirm that using this lowered H_T cut still gives a valid answer, the fractions are compared as a function of H_T in Figure 85. In these plots, especially in the higher-statistics 2016 plot, it is clear both that the data and MC agree very well and that there is no strong H_T dependence.

All the uncertainties are calculated independently for the two datasets, then added together. Statistical uncertainties, including the MC closure statistical uncertainties and the k and α uncertainties, are added in quadrature between the two years. Uncertainties that are more likely to be correlated, such as the difference between the two estimates in MC closure and the dependence on using a dAOD to calculate trigger efficiencies, are added linearly. The total uncertainty is about 12% of the nominal prediction in SRZ and about 9% in VRS.

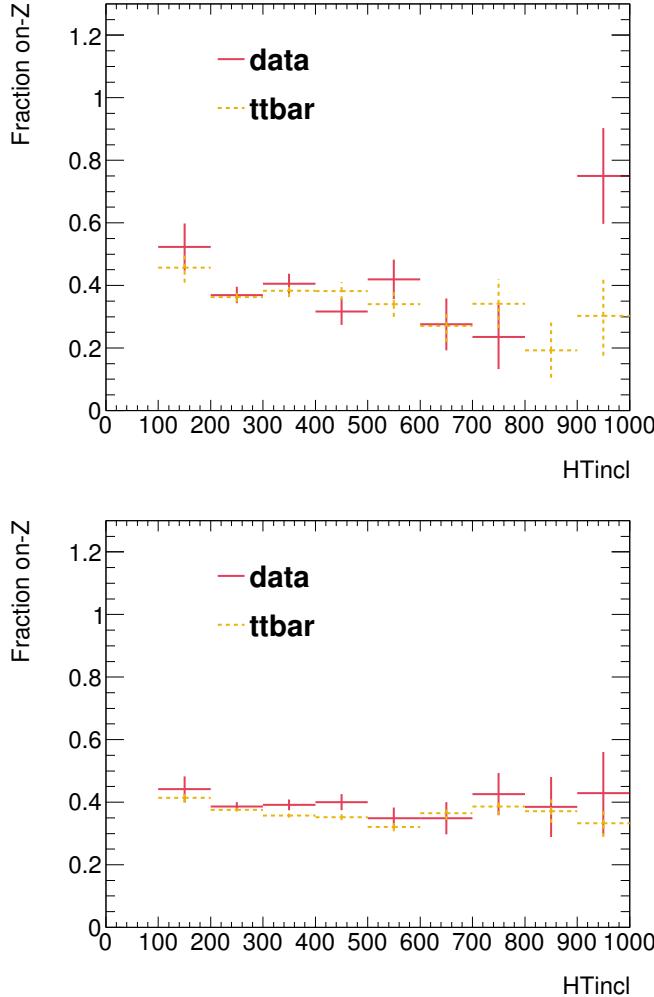


Figure 85: Plots of the fraction of on-Z events with a VR-FS-like selection as a function of H_T . The top figure shows 2015 data and MC while the bottom figure shows the same for 2016.

11.1.2 Uncertainties on the $\gamma + \text{jets}$ Method

One of the largest sources of uncertainty on the $\gamma + \text{jets}$ method is derived by comparing the results from reweighting in different variables. Though boson p_T is used as the nominal reweighting variable, the differences in the kinematics of γ and Z events also impact number of jets, H_T , and E_T (which includes the mass of the boson). The $\gamma + \text{jets}$ method is repeated using each of these variables to reweight, and their E_T^{miss} distributions are shown in Figure 86. The maximum difference from the nominal prediction is symmetrized and used as an uncertainty on the method.

Another uncertainty is applied to estimate the validity of using MC in a 1-jet CR to determine the smearing functions. Smearing functions are made using data from the same 1-jet region and using MC in a

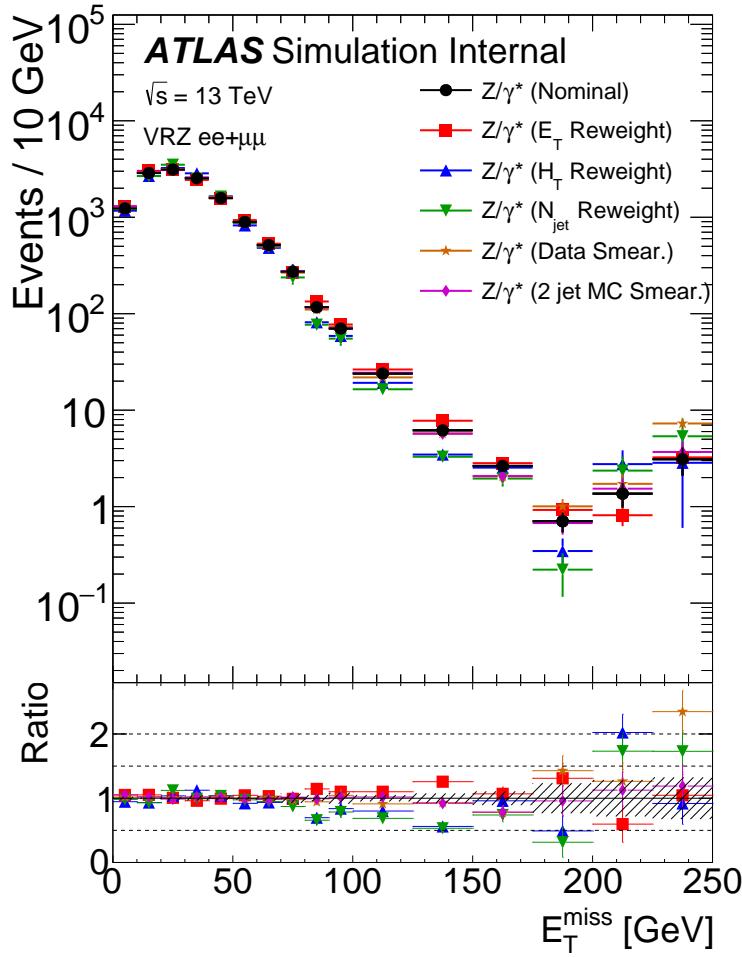


Figure 86: E_T^{miss} distributions for $\gamma + \text{jets}$ predictions using different reweighting variables, as well as distributions with the nominal reweighting but with smearing functions taken from data and from [MC](#) in a ≥ 2 -jet region.

≥ 2 -jet region otherwise identical to the 1-jet CR. These distributions are also shown in Figure 86, and like the alternate reweighting distributions, are used to find a maximum difference from the nominal prediction which is translated into a symmetric error.

As in the flavor symmetric method, the full procedure is carried out on [MC](#) in order to test [MC](#) closure, including a recalculation of any weights that are typically derived from data. The resulting comparison between $Z/\gamma^* + \text{jets}$ [MC](#) and the $\gamma + \text{jets}$ method performed on [MC](#) can be seen in Figure 87. The final non-closure uncertainty is taken from VRS, where larger numbers of events give a clearer picture of the success of the method than in SRZ. In this region, the statistical uncertainty on the prediction is compared to the non-closure, and the larger of the two is used as the final uncertainty.

The uncertainty on the $V\gamma$ contamination in CR- γ is also considered. An uncertainty on the [MC](#) prediction is made based on compar-

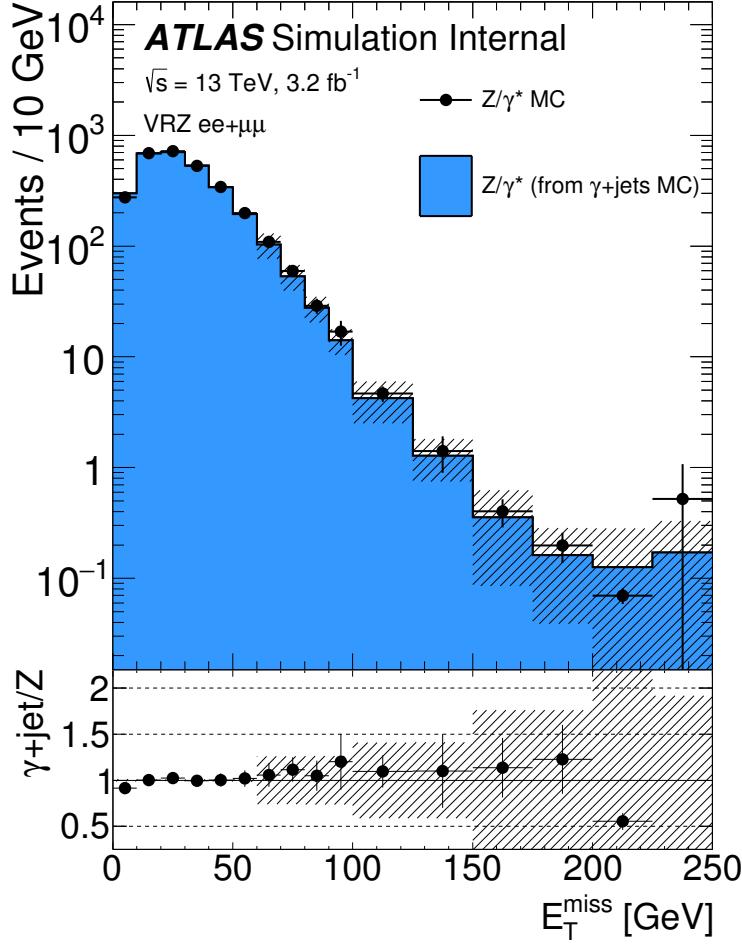


Figure 87: MC closure of the $\gamma + \text{jets}$ method as a function of E_T^{miss} comparing the MC prediction of the Z background with the $\gamma + \text{jets}$ method performed on $\gamma + \text{jets}$ MC. The uncertainty band includes both statistical and reweighting uncertainties.

ison of data and MC in a $W + \text{jets}$ VR, shown in Figure 88. This VR is similar to CR- γ , but instead of vetoing events with leptons, requires at least one well-isolated lepton with a p_T over 25 GeV. At E_T^{miss} values over 100 GeV, region is about 90% pure in $W\gamma$ processes. The MC agrees well with data in this region, even at very high E_T^{miss} , so an uncertainty of 16% based primarily on statistical uncertainty in this VR is placed on the $V\gamma$ MC. This uncertainty is propagated to the final result through the subtraction procedure.

An uncertainty on the $m_{\ell\ell}$ shape is determined using MC closure as well. The comparison of $m_{\ell\ell}$ shapes in $Z/\gamma^* + \text{jets}$ MC and the $\gamma + \text{jets}$ method applied to MC is shown in Figure 72. As with the main MC closure test, the maximum of the statistical uncertainty and the non-closure is taken as the final uncertainty on this background.

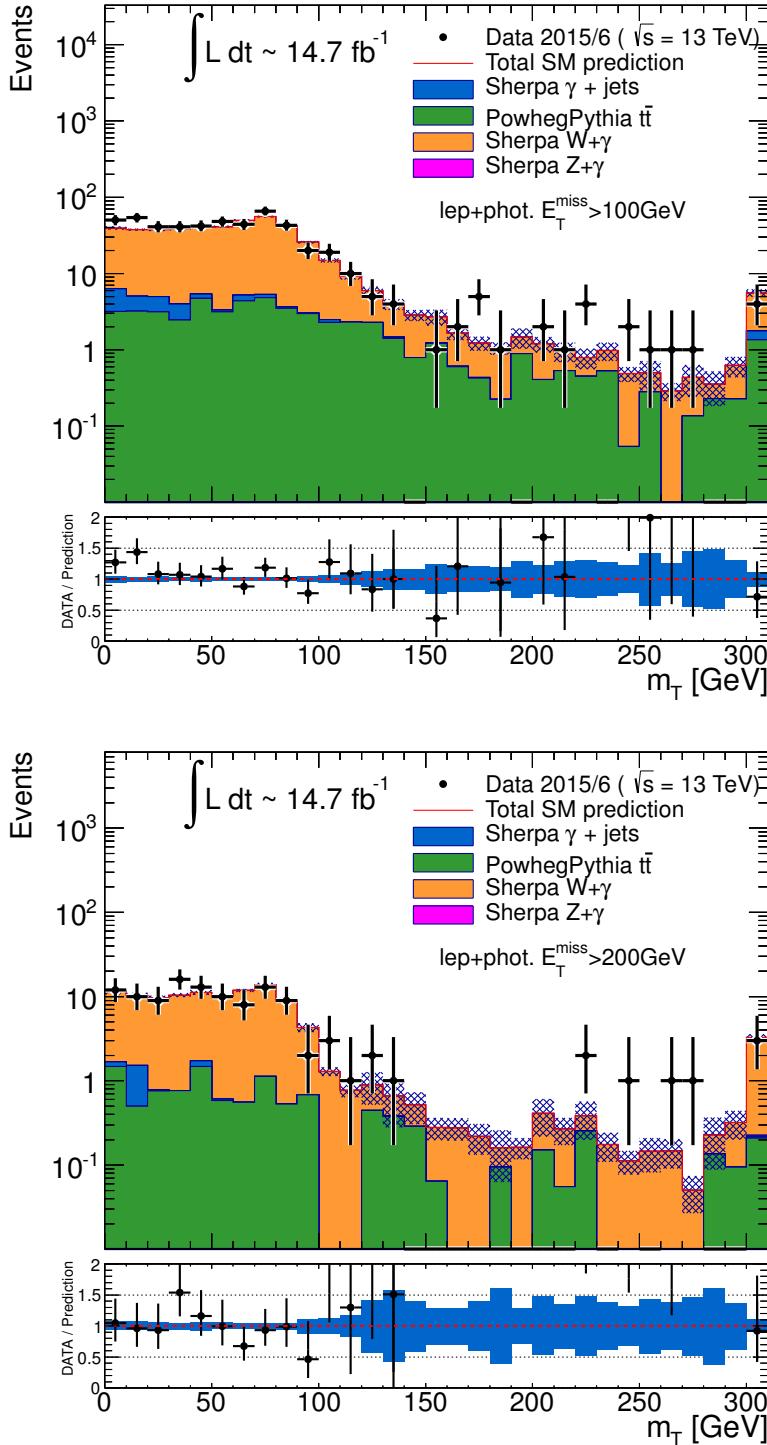


Figure 88: Distributions of $m_T(\ell, E_T^{\text{miss}})$, the transverse mass of the lepton and the E_T^{miss} in a VR designed to target $W\gamma$ processes. Top is the distribution with a E_T^{miss} cut at 100 GeV, and bottom is the same distribution with a E_T^{miss} cut of 200 GeV.

One last uncertainty based on the statistical uncertainty on the number of $\gamma + \text{jets}$ data events used for this method is also included. The full breakdown of uncertainty in SRZ can be seen in Table 17.

Ch.	Pred.	Uncertainties (%)						
		$V\gamma$ sub.	MC clos.	$m_{\ell\ell}$ shape	re- weight	smear	stat.	total
ee	1.02	53.0	21.0	19.0	100.0	65.0	56.0	145.0
$\mu\mu$	2.08	27.0	14.0	23.0	30.0	59.0	40.0	86.0
$ee+\mu\mu$	3.1	36.0	16.0	22.0	43.0	60.0	33.0	92.0

Table 17: Uncertainty breakdown for the $\gamma + \text{jets}$ method in SRZ. Uncertainties considered are the impact of MC uncertainty on $V\gamma$ backgrounds, MC closure, uncertainty on $m_{\ell\ell}$ shape (also determined via MC closure), reweighting uncertainties, smearing uncertainties, and statistical uncertainty on the $\gamma + \text{jets}$ events used in the method.

11.1.3 Uncertainties on the Fakes Background

Systematic uncertainties on the fakes background are derived from a series of variations on the nominal method. Variations include scaling the real and fake efficiencies up and down by their statistical uncertainties, scaling the prompt lepton contamination in CR-fake up and down by 20%, and by requiring and vetoing b -tagged jets in CR-fake to determine the dependence on heavy flavor. Statistical uncertainties can also be large in regions with small numbers of events in the baseline selection, such as SRZ. In other regions, the b -tagging dependence provides the largest uncertainty. The full breakdown of uncertainties for the most important regions are listed in Table 18.

11.2 THEORETICAL AND EXPERIMENTAL UNCERTAINTIES

Experimental uncertainties cover any detector effect or LHC condition that may not be modeled precisely correctly in MC. For each uncertainty, a standard prescription from the ATLAS experiment is followed. Uncertainties are included on the following parameters:

- Luminosity (2.9%) [97, 98]
- Jet energy scale [69]
- Jet energy resolution [69]
- Jet vertex tagging
- Heavy flavor tagging
- E_T^{miss} soft term [72]
- e/μ momentum scale
- e/μ trigger, reconstruction, and identification efficiencies

Variation	SRZ	CRT	CRFS	VRFS	VRS	VRT
Nominal	0.10 ± 1.61	25.39 ± 5.35	3.73 ± 2.19	10.53 ± 3.56	3.64 ± 3.20	80.06 ± 9.80
EL F Up	0.15	30.23	3.96	10.93	3.56	92.46
EL F Down	0.06	21.80	3.52	10.18	3.54	70.07
EL R Up	0.25	26.17	3.92	11.10	4.13	82.57
EL R Down	-0.07	24.51	3.52	9.92	3.10	77.24
MU F Up	-0.20	32.48	4.77	16.41	5.25	86.48
MU F Down	0.29	20.17	2.91	7.04	2.87	70.12
MU R Up	0.13	25.67	3.78	10.66	3.81	81.18
MU R Down	0.05	25.04	3.67	10.38	3.44	78.72
Total Sys	$+0.26 -0.35$	$+8.64 -6.39$	$+1.08 -0.87$	$+5.92 -3.56$	$+1.70 -0.97$	$+14.24 -14.42$
Total Sys (%)	$+261.05 -354.72$	$+34.01 -25.19$	$+29.05 -23.23$	$+56.22 -33.85$	$+46.57 -26.60$	$+17.78 -18.02$
Real Cont. Up	0.23	20.97	3.06	8.08	3.15	68.79
Real Cont. Down	-0.01	29.67	4.38	12.95	4.16	90.23
b-jet	0.31	40.44	5.28	8.98	5.63	120.50
no b-jet	0.16	23.44	3.08	11.38	3.97	70.55
Total Sys	$+0.25 -0.11$	$+15.65 -4.83$	$+1.69 -0.93$	$+2.56 -2.90$	$+2.09 -0.49$	$+41.71 -14.74$
Total Sys (%)	$+260.46 -109.06$	$+61.66 -19.02$	$+45.30 -24.85$	$+24.32 -27.58$	$+57.31 -13.35$	$+52.10 -18.42$

Table 18: Systematic uncertainties on the fake-lepton background for on-Z regions for 2015+2016 yields. The nominal yield includes statistical uncertainty from the baseline selection in a given region. The following rows indicate the results of varying the real and fake lepton efficiencies up and down by their statistical uncertainty. Real cont. gives an uncertainty on the contamination of real leptons in the fake lepton efficiency. b-jet and no b-jet indicate the impact of requiring or vetoing b -tagged jets in the regions used to measure the fake efficiency.

- Pile-up

These uncertainties are applied to all MC samples used in the analysis. This includes signal models, diboson and rare top samples for the nominal estimate, and all backgrounds taken from MC in the sideband fit.

Theoretical uncertainties include cross-section uncertainties, scale uncertainties, and PDF uncertainties. For the diboson samples, the scale uncertainties, given in Table 19 are calculated by varying each scale up and down by a factor of two. These are combined with a 6% cross-section uncertainty and a generator uncertainty obtained by comparing PowHEG and SHERPA MC yields in a given region. This generator uncertainty, shown in Table 20, is dominant in most regions. Rare top processes are given a 13% PDF and scale variation uncertainty [51] and a 22% cross section uncertainty [75–77].

Signal models have both the central value and uncertainty on cross-sections taken from an envelope of predictions using different scales and PDF sets [99]. The signal processes are calculated at Next-to-Leading-Logarithmic Accuracy (NLO+NLL); they are initially calculated at NLO in the strong coupling constant, with additional terms from next-to-leading-logarithmic resummation of soft gluon emission [100–104].

$VV \rightarrow ll\nu\nu$ Samples							
	SRZ	VRS	CRT	VRT	VRWZ	VRZZ	VR ₃ L
resummation	0.07	0.03	0.01	0.02	0.00	0.00	0.00
renormalization	0.13	0.17	0.16	0.22	0.00	0.00	0.00
factorization	0.01	0.01	0.01	0.03	0.00	0.00	0.00
total	0.15	0.17	0.16	0.22	0.00	0.00	0.00
$WZ \rightarrow lll\nu$ Samples							
	SRZ	VRS	CRT	VRT	VRWZ	VRZZ	VR ₃ L
resummation	0.07	0.05	0.13	0.08	0.02	0.00	0.01
renormalization	0.26	0.20	0.28	0.21	0.07	0.00	0.18
factorization	0.04	0.04	0.02	0.06	0.01	0.00	0.02
total	0.28	0.21	0.31	0.23	0.07	0.00	0.18
$ZZ \rightarrow llll$ Samples							
	SRZ	VRS	CRT	VRT	VRWZ	VRZZ	VR ₃ L
resummation	0.27	1.07	0.01	0.01	0.06	0.01	0.53
renormalization	0.28	0.26	0.30	0.60	0.07	0.04	0.14
factorization	0.27	0.25	0.30	0.58	0.13	0.02	0.16
total	0.48	1.13	0.43	0.84	0.16	0.05	0.57

Table 19: Fractional uncertainties of dibosons in signal and validation regions from Sherpa scale variations.

Region	Sherpa Events/fb ⁻¹	Sherpa Events	Powheg Events/fb ⁻¹	Powheg Events	% Difference
WZ Samples					
SRZ+VRZ	5.219	76.722	3.286	48.300	37.046
CRT+VRT	1.060	15.583	0.742	10.913	29.970
WW/ZZ Samples					
SRZ+VRZ	1.921	28.244	0.685	10.070	71.424
CRT+VRT	6.281	92.332	3.142	46.188	55.474

Table 20: Comparison of yields in on-Z and off-Z regions in Sherpa and Powheg diboson MC at 14.7 fb⁻¹.

11.3 IMPACT OF UNCERTAINTIES ON THE SIGNAL REGION

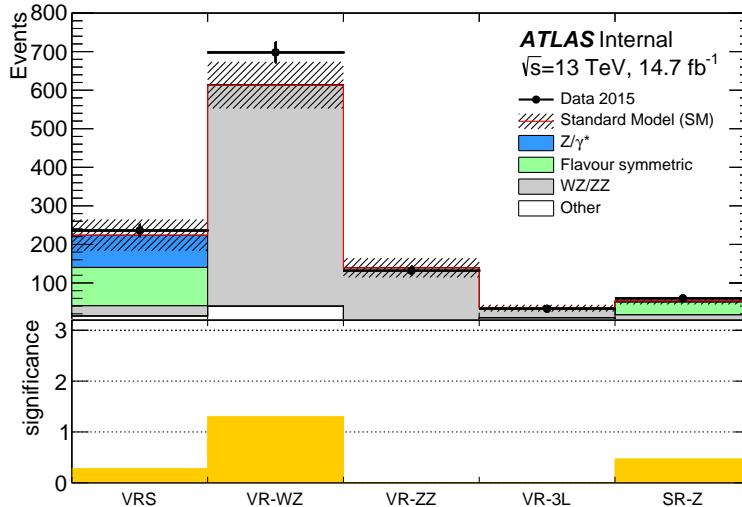
The breakdown of each major uncertainty's contribution to the total uncertainty in SRZ is shown in [Table 21](#). The dominant uncertainty is the diboson generator uncertainty, followed by the statistical uncertainty from the [FS](#) background. Uncertainties smaller than 1% are not shown in the table.

Source	Relative systematic uncertainty [%]
SRZ	
Total systematic uncertainty	17
WZ/ZZ generator uncertainty	13
Flavour symmetry (statistical)	7
WZ/ZZ scale uncertainty	6
Z/ γ^* + jets (systematic)	4
Flavour symmetry (systematic)	3
Z/ γ^* + jets (statistical)	2
Fake-leptons	1

[Table 21](#): Overview of the dominant sources of systematic uncertainty on the total background estimate in the signal regions. The values shown are relative to the total background estimate, shown in %.

RESULTS

The results of the search can be seen in [Table 22](#), which displays the expected and observed numbers of events in SRZ, both divided by channel and inclusively. The predictions and uncertainties for each background are shown, though many of these uncertainties are correlated between backgrounds, so the final uncertainty does not correspond to a simple addition in quadrature of each error. A total of sixty events are observed, with 53.5 ± 9.3 events expected. [Figure 89](#) shows the expected and observed results visually for the **SR** as well as three **VRs**, all designed to verify the accuracy of the backgrounds taken from **MC**. Excellent agreement is seen in all cases, with the largest deviation at about 1σ .



[Figure 89](#): Comparison of background predictions and data yields in four validation regions, as well as the signal region. Definitions of all regions can be found in [Table 7](#), with both rare top and fake backgrounds grouped together under the “other” label. The uncertainty band includes all statistical and systematic uncertainties. Below is a panel of the one-sided statistical significances of the deviations between the predicted and observed quantities for each region.

[Table 22](#) also shows several statistical interpretations of the results. The discovery p -value for zero signal strength, which gives the probability that the observed events are compatible with a **SM**-only hypothesis, is given as 0.32. The significance is listed as 0.47σ , which is a reinterpretation of the p -value into a gaussian significance. This

Table 22: Number of events expected and observed in the ee , $\mu\mu$, and combined channels. Expected predictions include all systematic and statistical uncertainties discussed in Chapter 11. Also shown is the discovery p -value for zero signal strength ($p(s = 0)$) [105], Gaussian significance, 95% CL observed and expected upper limits on the number of signal events (S^{95}), and the corresponding observed upper limit on the visible cross section ($\langle\epsilon\sigma\rangle_{\text{obs}}^{95}$).

	SRZ	SRZ ee	SRZ $\mu\mu$
Observed events	60	35	25
Total expected background events	53.5 ± 9.3	27.1 ± 5.1	26.8 ± 4.4
Flavour-symmetric ($t\bar{t}$, Wt , WW and $Z \rightarrow \tau\tau$) events	33.2 ± 3.9	16.5 ± 2.1	16.7 ± 2.0
Z/γ^* + jets events	3.1 ± 2.8	$1.0_{-1.0}^{+1.3}$	2.1 ± 1.4
WZ/ZZ events	14.2 ± 7.7	7.8 ± 4.3	6.4 ± 3.5
Rare top events	2.9 ± 0.8	1.4 ± 0.4	1.5 ± 0.4
Fake-lepton events	$0.1_{-0.1}^{+0.8}$	$0.5_{-0.5}^{+0.7}$	$0^{+0.2}$
$p(s = 0)$	0.32	0.15	0.5
Significance (σ)	0.47	1.00	0
Observed (Expected) S^{95}	$28.2 (24.5_{-6.7}^{+8.9})$	$22.0 (15.8_{-4.5}^{+6.5})$	$12.9 (14.0_{-3.9}^{+5.7})$
$\langle\epsilon\sigma\rangle_{\text{obs}}^{95}$ [fb]	1.9	1.5	0.88

p-value is one-sided; when the data yield is less than expected the *p*-value is set to 0.5, and the significance is set to 0. S^{95} , the upper limit on the number of signal events that could be in the SR at a 95% CL, is determined both for the expected and observed number of events. This limit is also reinterpreted based on the integrated luminosity used in the search to produce an upper limit on the visible cross-section of signal events, $\langle\epsilon\sigma\rangle_{\text{obs}}^{95}$.

The predictions in SRZ, combined with $m_{\ell\ell}$ shapes taken from MC, are used to produce plots in a broader $m_{\ell\ell}$ range, seen in Figure 90. These plots are useful demonstrations of efficacy of the background estimation methods, showing the well-modeled $Z/\gamma^* + \text{jets}$ shape in the same-flavor region, and in the different-flavor region, demonstrating that there are no extreme fluctuations within the region used to predict the flavor symmetric background.

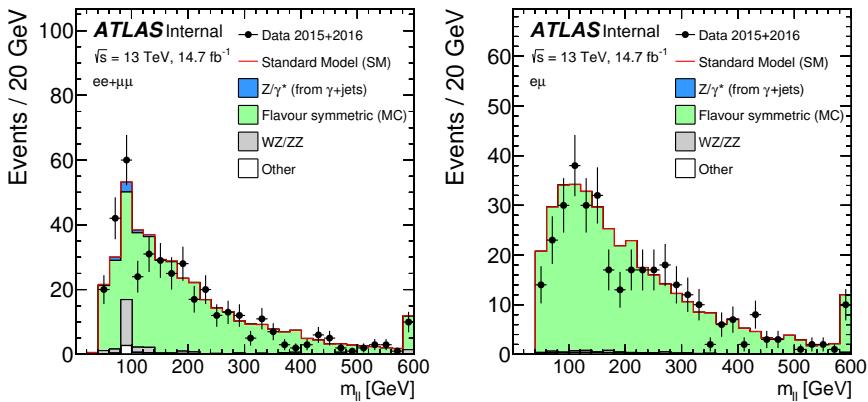


Figure 90: Comparisons as a function of $m_{\ell\ell}$ of background predictions with observed data in an SRZ-like region, with the $m_{\ell\ell}$ cut removed. Left is the same-flavor channel, where all background shapes are taken from MC and scaled to their SRZ predictions, except for the $Z/\gamma^* + \text{jets}$ background, which is taken entirely from the data-driven background. Right is the different-flavor channel, in which the backgrounds are taken directly from MC, except for $t\bar{t}$, which is scaled to match the total data yield.

Focusing in on the SR itself, comparisons of background predictions, observed events, and signal models can be made as a function of key variables for the analysis. Figure 91 shows several of these. The first two figures focus on the features of the SR events' leading leptons; they give the mass and p_T of a hypothetical parent particle reconstructed from the leptons. In the case of events with a real Z boson, these variables simply give that boson's mass at p_T . The next two figures show distributions in the two most important variables used to differentiate signal from background, E_T^{miss} and H_T . In this analysis, where the frozen SR resulted in cuts on these quantities that are lower than those that would be chosen based on a new optimization, these plots show that, even in more sensitive regions, no large

excess above the **SM** background is seen. The last pair of figures relates to the jets in the event, showing the total number of jets and the total number of b -jets in the **SR** events. The b -jet quantity is not explicitly cut on in the analysis because the fraction of b -jets produced is extremely model dependent. However, an excess at high b -jet multiplicity would suggest a **BSM** process. In each of these distributions, the observed distributions match the background predictions very well, and no evidence for any of the superimposed signal models is seen.

Comparisons of the observed and expected yield are also made as a function of $\Delta\phi(\text{jet}_{12}, \mathbf{p}_T^{\text{miss}})$, shown in [Figure 92](#). Here, results are shown in a region similar to SRZ with the cut on this variable removed, showing the efficacy of the background prediction in a region enhanced in $Z/\gamma^* + \text{jets}$ events. Again, excellent agreement is seen between the background prediction and observed data.

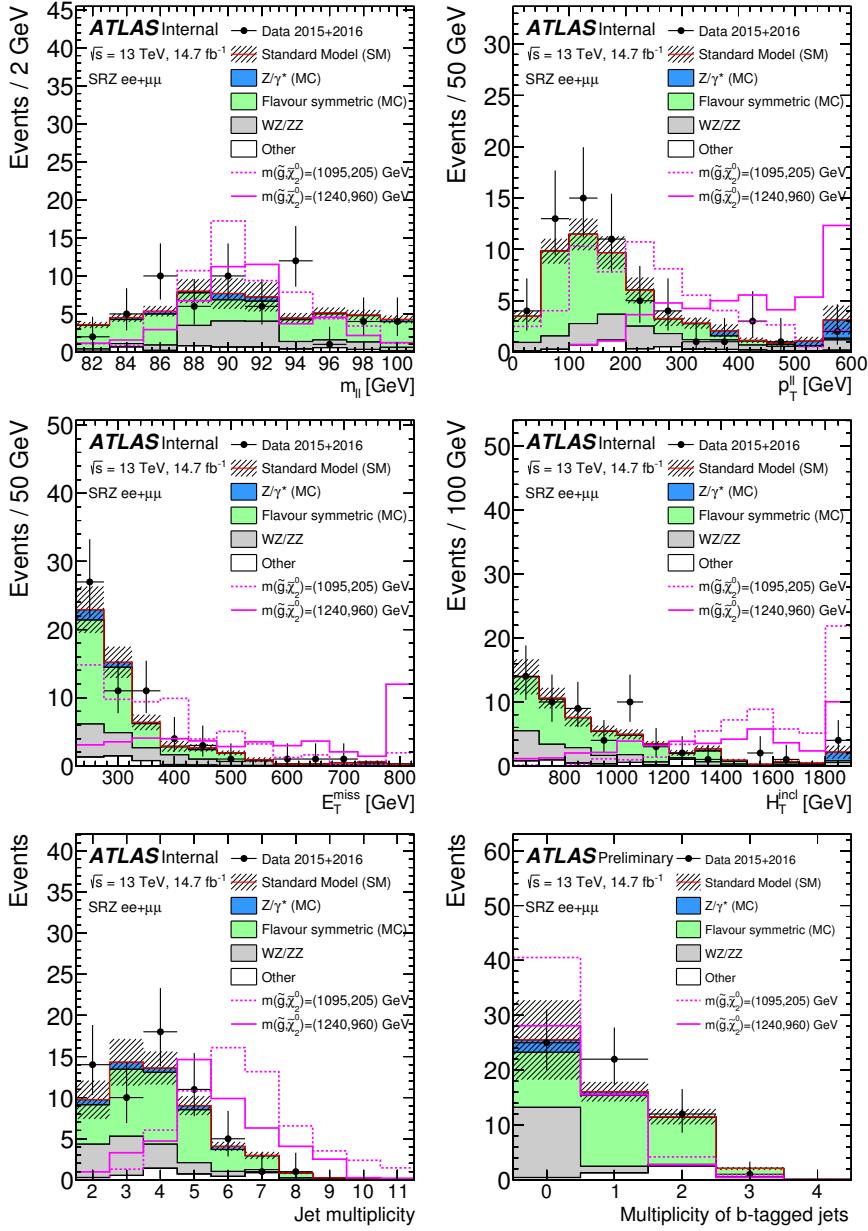


Figure 91: Distributions of observed data, background predictions, and simulated signals are shown in SRZ as a function of $m_{\ell\ell}$, $p_T^{\ell\ell}$, E_T^{miss} , H_T , number of jets, and number of b -jets. The two example signals have $(m(\tilde{g}), m(\tilde{\chi}_2^0)) = (1095, 205)$ GeV. All background shapes are taken from MC, and in the case of flavor symmetric and $Z/\gamma^* +$ jets backgrounds, their yields are scaled to match the data-driven predictions. Uncertainties include statistical and systematic components.

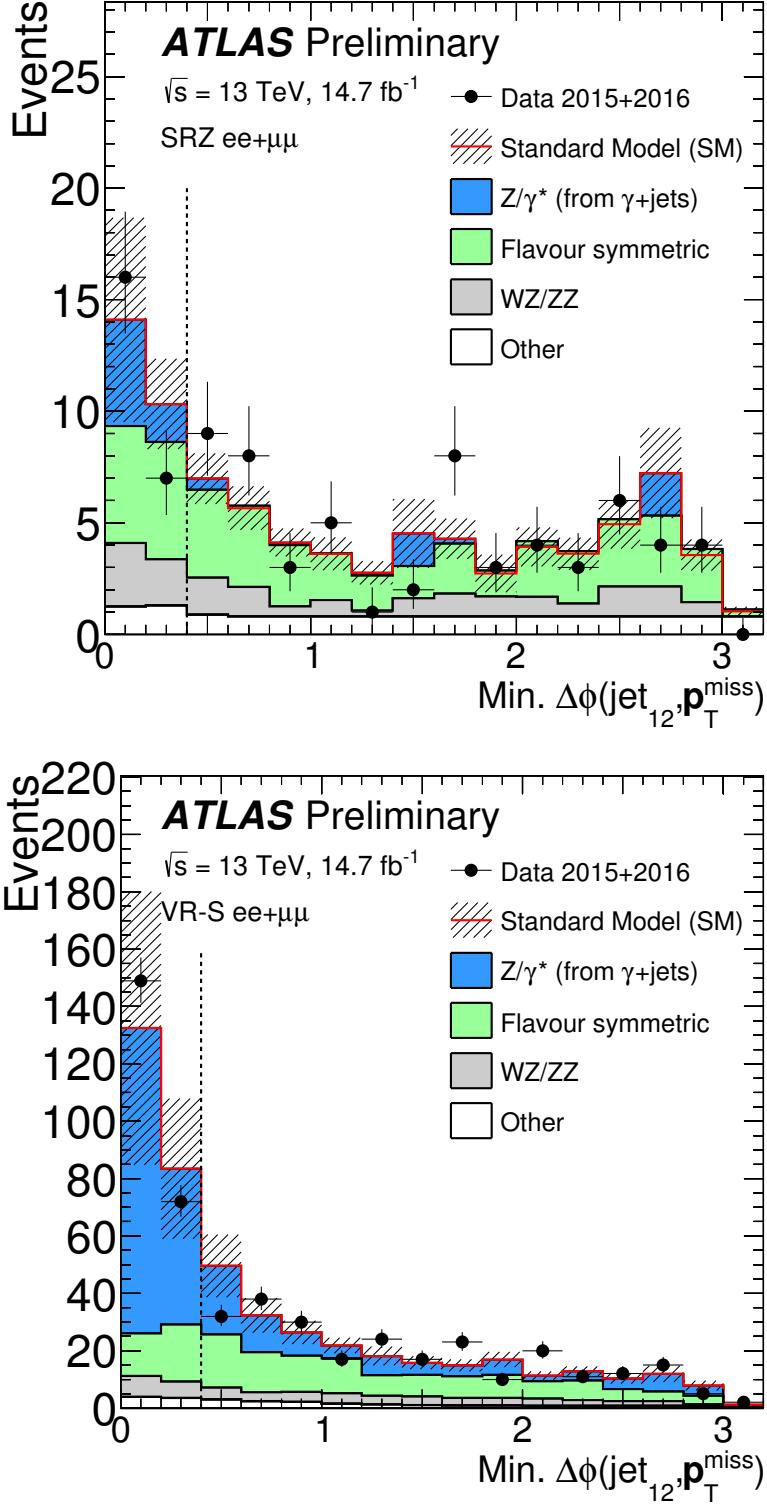


Figure 92: Comparisons as a function of $\Delta\phi(\text{jet}_{12}, \mathbf{p}_T^{\text{miss}})$ of background predictions with observed data in an SRZ-like (left) and VRS-like (right) region, with the $\Delta\phi(\text{jet}_{12}, \mathbf{p}_T^{\text{miss}})$ cut removed. All background shapes are taken from MC and scaled to their SRZ predictions, except for the Z/γ^* + jets background, which is taken entirely from the data-driven background.

INTERPRETATIONS

Using the simplified models discussed in Section 2.2.3, these results can be interpreted into exclusions of theories based on the masses of the particles involved. Of course, these exclusions include all the assumptions of the models used, so they shouldn't be interpreted to mean that no theory with a given set of particle masses can possibly exist, but they do provide a helpful guideline for targeting future searches and comparing results from different analyses.

Limits are determined using a program called HistFitter [105], designed within the ATLAS experiment, which builds upon the capabilities of ROOT [106], RooStats [107], and HistFactory [108] to combine the uncertainties of the various background predictions, including their correlations, and produce cross-section limits at 95% CL using the CL_S prescription [109, 110]. In this prescription, a likelihood is constructed based on the expected signal and background contributions to the SR. Nuisance parameters are created based on the statistical and systematic uncertainties for each data-driven background, as well as for each systematic applied to the MC-driven background estimates. The fit uses Gaussian models for nuisance parameters for all signal and background uncertainties, except for the statistical uncertainty on data- and MC-driven background estimates, which are interpreted as Poissonian. Experimental uncertainties are considered fully correlated across the signal and background MC-based estimates.

A fit is performed, leaving a signal strength parameter (μ) free, to maximize the likelihood, and subsequent fits are preformed to at discrete μ values to determine the relative likelihood of each value. Using this relative likelihood, the probability of a background-only hypothesis, p_b , can be determined by setting $\mu = 0$, as well as the probability of a signal + background hypothesis p_{s+b} with any non-zero signal strength, but nominally with $\mu = 1$. The confidence limit is constructed as a ratio

$$CL_S = \frac{p_{s+b}}{1 - p_b}. \quad (51)$$

Then, if CL_S falls below 5%, the signal + background hypothesis can be excluded at 95%. Expected exclusion limits are constructed by assuming the observed data precisely matches the prediction, and 1σ uncertainty bands are formed by varying the nuisance parameters away from their fitted values to produce a change in the likelihood. The observed limit uses the actual observation of data in the SR to set exclusion limits, so any excess above the expected background will result in worse limits than expected, and any deficit will result in

better limits. This exclusion is typically displayed with error bands that represent a 1σ variation in the cross-section of the signal models.

The simplified model discussed in [Section 2.2.3](#), in which pair-produced gluinos decay via a $\tilde{\chi}_2^0$ to jets, a Z boson, and a $\tilde{\chi}_1^0$ LSP, is produced in two grids, which differ by their choice of the LSP mass. The first grid assumes a light LSP, fixing its mass to 1 GeV for all mass points, and is shown as a function of \tilde{g} and $\tilde{\chi}_2^0$. The second grid is defined as a function of \tilde{g} and $\tilde{\chi}_1^0$, and its varying LSP mass is defined relative to the $\tilde{\chi}_2^0$ mass by $m(\tilde{\chi}_1^0) = m(\tilde{\chi}_2^0) - 100$ GeV. [Figure 93](#) shows the first of these grids, along with exclusions on a similar simplified model, which replaces the gluinos with squarks and uses the same mass scheme. The exclusion contours on the second grid is shown in [Figure 94](#), as a function of $m(\tilde{g})$ and $m(\tilde{\chi}_1^0)$.

In general, the observed exclusions are slightly weaker than the expected exclusions, due to a very small excess of events observed in SRZ. The observed lower limit on $m(\tilde{g})$ is about 1.3 TeV for models with $m(\tilde{\chi}_2^0) = 500$ GeV for the $\tilde{g}-\tilde{\chi}_2^0$ grid. These improve significantly on the previous [ATLAS](#) exclusion, which used different models for interpretation, but placed a lower limit on $m(\tilde{g})$ at around 900 GeV for similar $m(\tilde{\chi}_2^0)$.

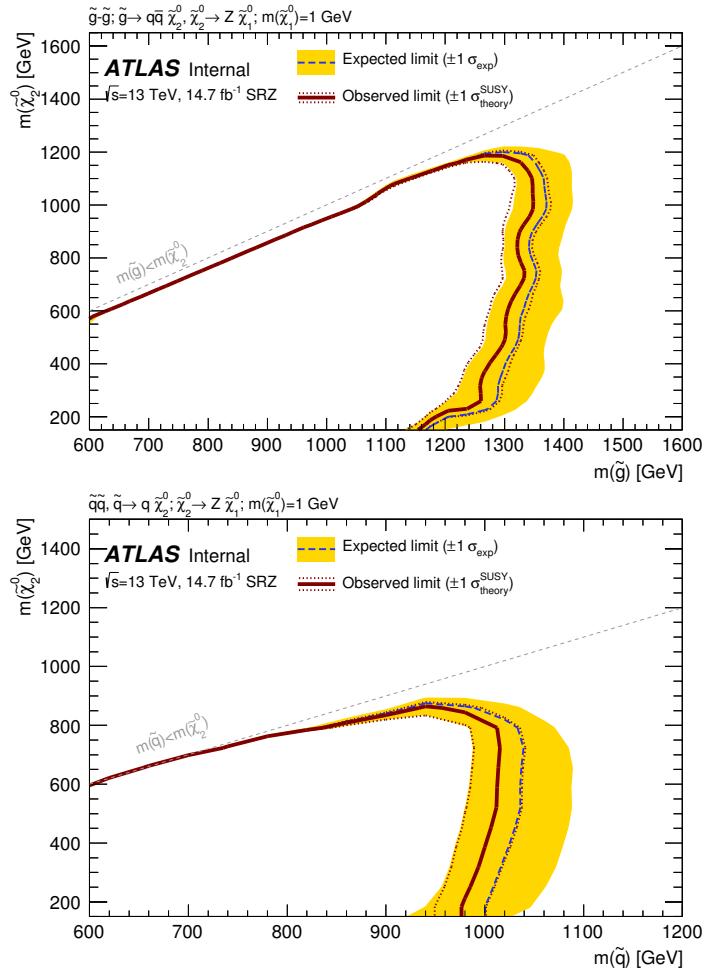


Figure 93: Expected and observed exclusion contours derived from the results in SRZ for the (top) $\tilde{g}-\tilde{\chi}_2^0$ on-shell grid and (bottom) $\tilde{q}-\tilde{\chi}_2^0$ on-shell grid. The dashed blue line indicates the expected limits at 95% CL and the yellow band shows the 1σ variation of the expected limit as a consequence of the uncertainties in the background prediction and the experimental uncertainties in the signal ($\pm 1\sigma_{\text{exp}}$). The observed limits are shown by the solid red line, with the dotted red lines indicating the variation resulting from changing the signal cross section within its uncertainty ($\pm 1\sigma_{\text{theory}}^{\text{SUSY}}$).

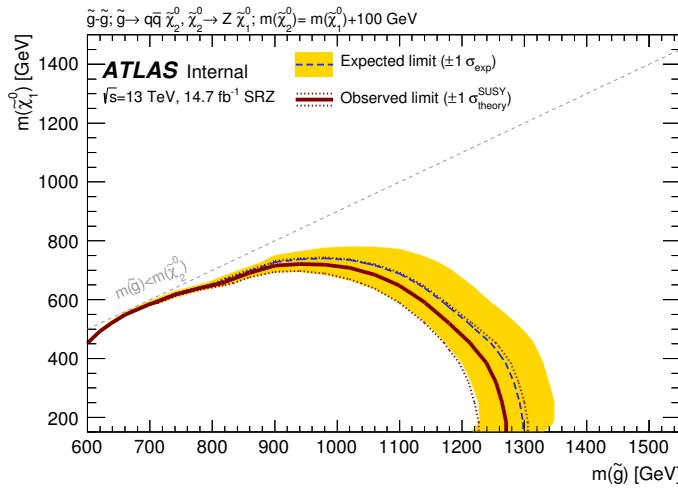


Figure 94: Expected and observed exclusion contours derived from the results in SRZ for the $\tilde{g}-\tilde{\chi}_1^0$ on-shell grid. The dashed blue line indicates the expected limits at 95% CL and the yellow band shows the 1σ variation of the expected limit as a consequence of the uncertainties in the background prediction and the experimental uncertainties in the signal ($\pm 1\sigma_{\text{exp}}$). The observed limits are shown by the solid red line, with the dotted red lines indicating the variation resulting from changing the signal cross section within its uncertainty ($\pm 1\sigma_{\text{theory}}^{\text{SUSY}}$).

Part V

CONCLUSIONS

CONCLUSIONS AND OUTLOOK

After a series of moderate excesses observed by the [ATLAS](#) experiment in events with a Z boson, jets, and E_T^{miss} , this analysis performed on 14.7 fb^{-1} of 13 TeV data sees excellent agreement between observations and the background expectation. The resulting exclusion pushes the gluino mass lower limit beyond 1 TeV, putting further constraints on possible [SUSY](#) models. Along with the many other searches for [SUSY](#), this exclusion limits the phase space available for natural [SUSY](#) models. However, [SUSY](#) is adaptable; new theories stretching those bounds are continually proposed as tighter experimental constraints are set, and there are always small gaps in the exclusions where sparticles could hide.

[ATLAS](#)'s dataset for 2016 includes 36 fb^{-1} , more than twice the luminosity included in this search. Because no excess was seen in this analysis, the next search in this channel will be able to re-optimize its signal regions for this larger dataset. In fact, because the signal region has been frozen since the 8 TeV search, this analysis's signal region hasn't ever been re-optimized for the increased energy of the [LHC](#)'s collisions. A new signal region that increases E_T^{miss} and H_T requirements will allow for better sensitivity to [SUSY](#) processes.

In addition, the current signal region, in which 60 events were observed with 14.7 fb^{-1} , will be populated enough to be subdivided based on event features. The current search is agnostic to the number of b -jets in the event, for example, but there are now enough events to separate this signal region into complementary b -tagged and b -vetoed regions, allowing analyzers to independently target models which produce b -jets and those that don't, and in the latter case, to dramatically reduce the $t\bar{t}$ background. Signal regions can also be binned in other model-dependent features, like number of jets, and the E_T^{miss} and H_T requirements can be increased independently, targeting different event topologies.

The [LHC](#) will continue to run through 2018 with a possible increase to $\sqrt{s} = 14 \text{ TeV}$, and will shut down for upgrades until 2021. Three more years of data-taking at 14 TeV will follow, with approximately twice the current luminosity, referred to as Run 3. After that, the [LHC](#) will shut down again to prepare for the High Luminosity Large Hadron Collider ([HL-LHC](#)), which will begin data-taking in 2026 at a luminosity approximately five times the current rate. This run will result in roughly 3000 fb^{-1} , which will allow for dramatically better sensitivity in [SUSY](#) searches. An example can be seen in [Figure 95](#), which shows the potential exclusions on a simple gluino pair-production

model with decays via squarks to a LSP, for the approximate luminosities of Run 3 and the HL-LHC.

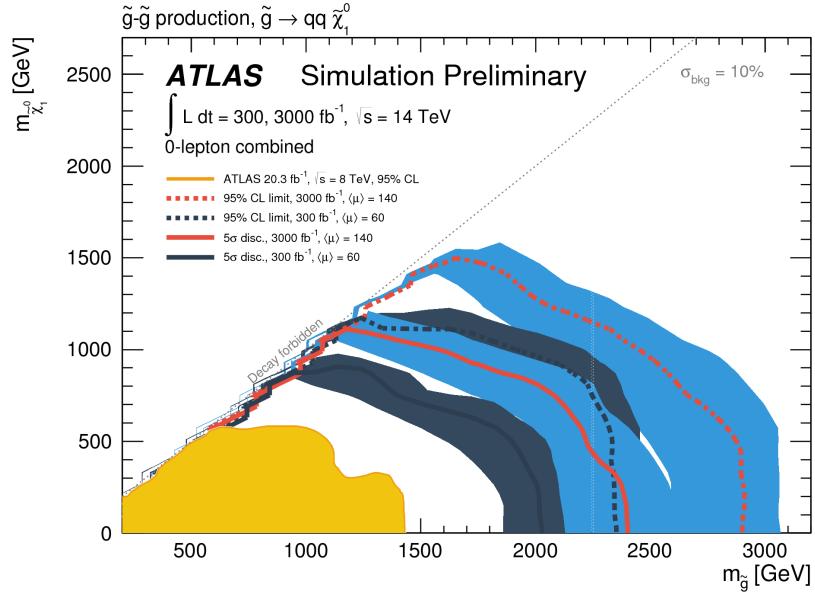


Figure 95: Expected 95% CL exclusion contours (dashed) and 5σ discovery contours (solid) for $L_{int} = 300^{-1}$ (black) and 3000^{-1} (red) for gluino pair-production, with 1σ bands representing the uncertainty on the production cross-section. Superimposed is the observed 8 TeV exclusion for similar models. [111]

Searches like this one will surely be repeated with higher and higher luminosities, the analyses increasing both in sensitivity and in complexity. Whether or not they uncover any hints of physics beyond the Standard Model remains to be seen.

BIBLIOGRAPHY

- [1] ATLAS Collaboration. “Search for supersymmetry in events containing a same-flavour opposite-sign dilepton pair, jets, and large missing transverse momentum in $\sqrt{s} = 8$ TeV pp collisions with the ATLAS detector.” In: *Eur. Phys. J. C* 75 (2015), p. 318. doi: [10.1140/epjc/s10052-015-3518-2](https://doi.org/10.1140/epjc/s10052-015-3518-2). arXiv: [1503.03290 \[hep-ex\]](https://arxiv.org/abs/1503.03290).
- [2] ATLAS Collaboration. *A search for supersymmetry in events containing a leptonically decaying Z boson, jets and missing transverse momentum in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector.* ATLAS-CONF-2015-082. 2015. URL: <http://cdsweb.cern.ch/record/2114854>.
- [3] C. P. Burgess and G. D. Moore. *The standard model: A primer.* Cambridge University Press, 2006. ISBN: 9780511254857, 9781107404267, 9780521860369.
- [4] S. L. Glashow. “Partial Symmetries of Weak Interactions.” In: *Nucl. Phys.* 22 (1961), pp. 579–588. doi: [10.1016/0029-5582\(61\)90469-2](https://doi.org/10.1016/0029-5582(61)90469-2).
- [5] Abdus Salam. “Renormalizability of Gauge Theories.” In: *Phys. Rev.* 127 (1 1962), pp. 331–334. doi: [10.1103/PhysRev.127.331](https://doi.org/10.1103/PhysRev.127.331). URL: <http://link.aps.org/doi/10.1103/PhysRev.127.331>.
- [6] Steven Weinberg. “A Model of Leptons.” In: *Phys. Rev. Lett.* 19 (21 1967), pp. 1264–1266. doi: [10.1103/PhysRevLett.19.1264](https://doi.org/10.1103/PhysRevLett.19.1264). URL: <http://link.aps.org/doi/10.1103/PhysRevLett.19.1264>.
- [7] D. Galbraith. *The Standard Model of the Standard Model.* 1997. URL: <http://davidgalbraith.org/portfolio/ux-standard-model-of-the-standard-model/>.
- [8] David J Griffiths. *Introduction to elementary particles; 2nd rev. version.* Physics textbook. New York, NY: Wiley, 2008. URL: <https://cds.cern.ch/record/111880>.
- [9] P. Hut and K. A. Olive. “A cosmological upper limit on the mass of heavy neutrinos.” In: *Physics Letters B* 87 (Oct. 1979), pp. 144–146. doi: [10.1016/0370-2693\(79\)90039-X](https://doi.org/10.1016/0370-2693(79)90039-X).
- [10] Huang, Qing-Guo, Wang, Ke, and Wang, Sai. “Constraints on the neutrino mass and mass hierarchy from cosmological observations.” In: *Eur. Phys. J. C* 76.9 (2016), p. 489. doi: [10.1140/epjc/s10052-016-4334-z](https://doi.org/10.1140/epjc/s10052-016-4334-z). URL: <http://dx.doi.org/10.1140/epjc/s10052-016-4334-z>.

- [11] The MEG Collaboration. "Search for the Lepton Flavour Violating Decay $\mu^+ \rightarrow e^+ \gamma$ with the Full Dataset of the MEG Experiment." In: (2016). eprint: [arXiv:1605.05081](https://arxiv.org/abs/1605.05081).
- [12] LHCb collaboration. "Observation of $J/\Psi p$ resonances consistent with pentaquark states in $\Lambda_b^0 \rightarrow J/\Psi K^- p$ decays." In: (2015). doi: [10.1103/PhysRevLett.115.072001](https://doi.org/10.1103/PhysRevLett.115.072001). eprint: [arXiv:1507.03414](https://arxiv.org/abs/1507.03414).
- [13] K. A. Olive et al. "Review of Particle Physics." In: *Chin. Phys. C* 38 (2014), p. 090001. doi: [10.1088/1674-1137/38/9/090001](https://doi.org/10.1088/1674-1137/38/9/090001).
- [14] Johan Messchendorp. "Physics with Charmonium – A few recent highlights of BESIII." In: *PoS Bormio2013* (2013), p. 043. arXiv: [1306.6611 \[hep-ex\]](https://arxiv.org/abs/1306.6611).
- [15] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt. "Parton distributions for the LHC." In: (2009). doi: [10.1140/epjc/s10052-009-1072-5](https://doi.org/10.1140/epjc/s10052-009-1072-5). eprint: [arXiv:0901.0002](https://arxiv.org/abs/0901.0002).
- [16] W.J. Stirling. *proton-(anti)proton cross sections*. 2013. URL: <http://www.hep.ph.ic.ac.uk/~wstirlin/plots/plots.html>.
- [17] F. Zwicky. "Die Rotverschiebung von extragalaktischen Nebeln." In: *Helvetica Physica Acta* 6 (1933), 110–127.
- [18] Douglas Clowe, Marusa Bradac, Anthony H. Gonzalez, Maxim Markevitch, Scott W. Randall, Christine Jones, and Dennis Zaritsky. "A direct empirical proof of the existence of dark matter." In: (2006). doi: [10.1086/508162](https://doi.org/10.1086/508162). eprint: [arXiv:astro-ph/0608407](https://arxiv.org/abs/astro-ph/0608407).
- [19] G. Hinshaw et al. "Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Data Processing, Sky Maps, and Basic Results." In: (2008). doi: [10.1088/0067-0049/180/2/225](https://doi.org/10.1088/0067-0049/180/2/225). eprint: [arXiv:0803.0732](https://arxiv.org/abs/0803.0732).
- [20] T. S. van Albada, J. N. Bahcall, K. Begeman, and R. Sancisi. "Distribution of dark matter in the spiral galaxy NGC 3198." In: *The Astrophysical Journal* 295 (Aug. 1985), pp. 305–313. doi: [10.1086/163375](https://doi.org/10.1086/163375).
- [21] M. Gell-Mann. *THE EIGHTFOLD WAY: A THEORY OF STRONG INTERACTION SYMMETRY*. 1961. doi: [10.2172/4008239](https://doi.org/10.2172/4008239). URL: <http://www.osti.gov/scitech/servlets/purl/4008239>.
- [22] J. Wess and B. Zumino. "Supergauge transformations in four dimensions." In: *Nuclear Physics B* 70.1 (1974), pp. 39–50. ISSN: 0550-3213. doi: [http://dx.doi.org/10.1016/0550-3213\(74\)90355-1](http://dx.doi.org/10.1016/0550-3213(74)90355-1). URL: <http://www.sciencedirect.com/science/article/pii/0550321374903551>.
- [23] J.-L. Gervais and B. Sakita. "Field theory interpretation of supergauges in dual models." In: *Nuclear Physics B* 34 (1971), pp. 632–639. doi: [10.1016/0550-3213\(71\)90351-8](https://doi.org/10.1016/0550-3213(71)90351-8).

- [24] P. Ramond. "Dual Theory for Free Fermions." In: *Physical Review D* 3 (1971), pp. 2415–2418. doi: [10.1103/PhysRevD.3.2415](https://doi.org/10.1103/PhysRevD.3.2415).
- [25] Stephen P. Martin. "A Supersymmetry primer." In: (1997). [Adv. Ser. Direct. High Energy Phys.] 18, 1 (1998)]. doi: [10.1142/9789812839657_0001](https://doi.org/10.1142/9789812839657_0001), [10.1142/9789814307505_0001](https://doi.org/10.1142/9789814307505_0001). arXiv: [hep-ph/9709356](https://arxiv.org/abs/hep-ph/9709356) [hep-ph].
- [26] Craig J Copi, David N. Schramm, and Michael S. Turner. "Big-Bang Nucleosynthesis and the Baryon Density of the Universe." In: (1994). doi: [10.1126/science.7809624](https://doi.org/10.1126/science.7809624). eprint: [arXiv:astro-ph/9407006](https://arxiv.org/abs/astro-ph/9407006).
- [27] NobelPrize.org. *The Nobel Prize in Physics 2004 - Popular Information*. 2016. URL: <https://www.nobelprize.org/nobel-prizes/physics/laureates/2004/popular.html>.
- [28] ATLAS Collaboration. "Dark matter interpretations of ATLAS searches for the electroweak production of supersymmetric particles in $\sqrt{s} = 8$ TeV proton-proton collisions." In: (2016). doi: [10.1007/JHEP09\(2016\)175](https://doi.org/10.1007/JHEP09(2016)175). eprint: [arXiv:1608.00872](https://arxiv.org/abs/1608.00872).
- [29] The ATLAS collaboration. "Further searches for squarks and gluinos in final states with jets and missing transverse momentum at $\sqrt{s} = 13$ TeV with the ATLAS detector." In: (2016).
- [30] W. Beenakker, S. Brensing, M. Krämer, A. Kulesza, E. Laenen, L. Motyka, and I. Niessen. "Squark and gluino hadroproduction." In: (2011). doi: [10.1142/S0217751X11053560](https://doi.org/10.1142/S0217751X11053560). eprint: [arXiv:1105.1110](https://arxiv.org/abs/1105.1110).
- [31] Wim Beenakker, Christoph Borschensky, Michael Krämer, Anna Kulesza, and Eric Laenen. *NNLL-fast: predictions for coloured supersymmetric particle production at the LHC with threshold and Coulomb resummation*. 2016. eprint: [arXiv:1607.07741](https://arxiv.org/abs/1607.07741).
- [32] ATLAS Collaboration. *ATLAS SUSY Searches - 95% CL Lower Limits*. "<https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CombinedSummaryPlots/>" 2016.
- [33] CMS Collaboration. "Search for physics beyond the standard model in events with a Z boson, jets, and missing transverse energy in pp collisions at $\sqrt{s} = 7$ TeV." In: *Phys. Lett. B* 716 (2012), pp. 260–284. doi: [10.1016/j.physletb.2012.08.026](https://doi.org/10.1016/j.physletb.2012.08.026). arXiv: [1204.3774](https://arxiv.org/abs/1204.3774) [hep-ex].
- [34] CMS Collaboration. "Search for physics beyond the standard model in events with two leptons, jets, and missing transverse momentum in pp collisions at $\sqrt{s} = 8$ TeV." In: *JHEP* 04 (2015), p. 124. doi: [10.1007/JHEP04\(2015\)124](https://doi.org/10.1007/JHEP04(2015)124). arXiv: [1502.06031](https://arxiv.org/abs/1502.06031) [hep-ex].

- [35] *Search for new physics in final states with two opposite-sign same-flavor leptons, jets and missing transverse momentum in pp collisions at $\sqrt{s}=13$ TeV.* Tech. rep. CMS-PAS-SUS-15-011. Geneva: CERN, 2015. URL: <https://cds.cern.ch/record/2114811>.
- [36] Lyndon Evans and Philip Bryant. “LHC Machine.” In: *Journal of Instrumentation* 3.08 (2008), So8001. URL: <http://stacks.iop.org/1748-0221/3/i=08/a=S08001>.
- [37] *LEP design report.* Copies shelved as reports in LEP, PS and SPS libraries. Geneva: CERN, 1984. URL: <https://cds.cern.ch/record/102083>.
- [38] ATLAS Collaboration. “The ATLAS Experiment at the CERN Large Hadron Collider.” In: *JINST* 3 (2008), So8003. doi: [10.1088/1748-0221/3/08/S08003](https://doi.org/10.1088/1748-0221/3/08/S08003).
- [39] The CMS Collaboration. “The CMS experiment at the CERN LHC.” In: *Journal of Instrumentation* 3.08 (2008), So8004. URL: <http://stacks.iop.org/1748-0221/3/i=08/a=S08004>.
- [40] The LHCb Collaboration. “The LHCb Detector at the LHC.” In: *Journal of Instrumentation* 3.08 (2008), So8005. URL: <http://stacks.iop.org/1748-0221/3/i=08/a=S08005>.
- [41] The ALICE Collaboration. “The ALICE experiment at the CERN LHC.” In: *Journal of Instrumentation* 3.08 (2008), So8002. URL: <http://stacks.iop.org/1748-0221/3/i=08/a=S08002>.
- [42] Peter Vankov. “ATLAS Upgrade for the HL-LHC: meeting the challenges of a five-fold increase in collision rate.” In: (2012). doi: [10.1051/epjconf/20122812069](https://doi.org/10.1051/epjconf/20122812069). eprint: [arXiv:1201.5469](https://arxiv.org/abs/1201.5469).
- [43] ATLAS Collaboration. *2015 start-up trigger menu and initial performance assessment of the ATLAS trigger using Run-2 data.* ATL-DAQ-PUB-2016-001. 2016. URL: <http://cds.cern.ch/record/2136007>.
- [44] ATLAS Collaboration. *Public Egamma Trigger Plots for Collision Data.* "https://twiki.cern.ch/twiki/bin/view/AtlasPublic/EgammaTriggerPublicResults" 2016.
- [45] Richard D. Ball et al. “Parton distributions with LHC data.” In: *Nucl. Phys.* B867 (2013), pp. 244–289. doi: [10.1016/j.nuclphysb.2012.10.003](https://doi.org/10.1016/j.nuclphysb.2012.10.003). arXiv: [1207.1303 \[hep-ph\]](https://arxiv.org/abs/1207.1303).
- [46] Hung-Liang Lai et al. “New parton distributions for collider physics.” In: *Phys. Rev. D* 82 (2010), p. 074024. doi: [10.1103/PhysRevD.82.074024](https://doi.org/10.1103/PhysRevD.82.074024). arXiv: [1007.2241 \[hep-ph\]](https://arxiv.org/abs/1007.2241).
- [47] T. Gleisberg et al. “Event generation with Sherpa 1.1.” In: *JHEP* 0902 (2009), p. 007. arXiv: [0811.4622 \[hep-ph\]](https://arxiv.org/abs/0811.4622).
- [48] P. Nason. “A new method for combining NLO QCD with shower Monte Carlo algorithms.” In: *JHEP* 0411 (2004), p. 040. arXiv: [hep-ph/0409146 \[hep-ph\]](https://arxiv.org/abs/hep-ph/0409146).

- [49] S. Frixione, P. Nason, and C. Oleari. “Matching NLO QCD computations with parton shower simulations: the POWHEG method.” In: *JHEP* 0711 (2007), p. 070. arXiv: [0709.2092 \[hep-ph\]](#).
- [50] S. Alioli, P. Nason, C. Oleari, and E. Re. “A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX.” In: *JHEP* 1006 (2010), p. 043. arXiv: [1002.2581 \[hep-ph\]](#).
- [51] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations.” In: *JHEP* 07 (2014), p. 079. doi: [10.1007/JHEP07\(2014\)079](#). arXiv: [1405.0301 \[hep-ph\]](#).
- [52] T. Sjöstrand, S. Mrenna, and P. Skands. “PYTHIA 6.4 Physics and Manual.” In: *JHEP* 0605 (2006), p. 026. doi: [10.1088/1126-6708/2006/05/026](#). arXiv: [hep-ph/0603175 \[hep-ph\]](#).
- [53] S. Agostinelli et al. “GEANT4: A simulation toolkit.” In: *Nucl. Instrum. Meth. A* 506 (2003), pp. 250–303. doi: [10.1016/S0168-9002\(03\)01368-8](#).
- [54] Morad Aaboud et al. “A measurement of material in the ATLAS tracker using secondary hadronic interactions in 7 TeV pp collisions.” In: (2016). arXiv: [1609.04305 \[hep-ex\]](#).
- [55] ATLAS Collaboration. “ATLAS event at 13 TeV - First stable beam, 3 June 2015 - run: 266904, evt: 25884805.” General Photo. 2015. URL: <https://cds.cern.ch/record/2022202>.
- [56] *The Expected Performance of the ATLAS Inner Detector*. Tech. rep. ATL-PHYS-PUB-2009-002. ATL-COM-PHYS-2008-105. Geneva: CERN, 2008. URL: <http://cds.cern.ch/record/1118445>.
- [57] ATLAS Collaboration. “A neural network clustering algorithm for the ATLAS silicon pixel detector.” In: *JINST* 9 (2014), Po9009. doi: [10.1088/1748-0221/9/09/P09009](#). arXiv: [1406.7690 \[hep-ex\]](#).
- [58] Luke de Oliveira. *AGILEPack: Algorithms for Generalized Inference, Learning, and Extraction*. 2016. URL: <http://lukedeo.github.io/AGILEPack/>.
- [59] ATLAS Collaboration. *Measurement of performance of the pixel neural network clustering algorithm of the ATLAS experiment at $\sqrt{s} = 13$ TeV*. ATL-PHYS-PUB-2015-044. 2015. URL: <http://cdsweb.cern.ch/record/2054921>.
- [60] ATLAS Collaboration. *Robustness of the Artificial Neural Network Clustering Algorithm of the ATLAS experiment*. ATL-PHYS-PUB-2015-052. 2015. URL: <http://cdsweb.cern.ch/record/2116350>.

- [61] “Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data.” In: *Eur. Phys. J. C* 72 (2012), p. 1909. doi: [10.1140/epjc/s10052-012-1909-1](https://doi.org/10.1140/epjc/s10052-012-1909-1). arXiv: [1110.3174 \[hep-ex\]](https://arxiv.org/abs/1110.3174).
- [62] *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data.* Tech. rep. ATLAS-CONF-2016-024. Geneva: CERN, 2016. URL: <http://cds.cern.ch/record/2157687>.
- [63] ATLAS Collaboration. *Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run-1 data.* 2016. eprint: [arXiv:1606.01813](https://arxiv.org/abs/1606.01813).
- [64] *Photon identification in 2015 ATLAS data.* Tech. rep. ATL-PHYS-PUB-2016-014. Geneva: CERN, 2016. URL: <https://cds.cern.ch/record/2203125>.
- [65] *Photon Identification Efficiencies using 2016 Data with radiative Z boson decays.* Tech. rep. EGAM-2016-003. Geneva: CERN, 2016. URL: <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/EGAM-2016-003/index.html>.
- [66] *Official Isolation Working Points.* Geneva, 2016. URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/IsolationSelectionTool#Photons>.
- [67] ATLAS Collaboration. “Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}=13$ TeV.” In: (2016). doi: [10.1140/epjc/s10052-016-4120-y](https://doi.org/10.1140/epjc/s10052-016-4120-y). eprint: [arXiv:1603.05598](https://arxiv.org/abs/1603.05598).
- [68] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. “The Anti-k(t) jet clustering algorithm.” In: *JHEP* 0804 (2008), p. 063. doi: [10.1088/1126-6708/2008/04/063](https://doi.org/10.1088/1126-6708/2008/04/063). arXiv: [0802.1189 \[hep-ph\]](https://arxiv.org/abs/0802.1189).
- [69] ATLAS Collaboration. *Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s} = 13$ TeV.* ATL-PHYS-PUB-2015-015. 2015. URL: <http://cds.cern.ch/record/2037613>.
- [70] ATLAS Collaboration. *Tagging and suppression of pileup jets with the ATLAS detector.* ATLAS-CONF-2014-018. 2014. URL: <http://cds.cern.ch/record/1700870>.
- [71] ATLAS Collaboration. *Expected performance of the ATLAS b-tagging in Run-2.* ATL-PHYS-PUB-2015-022. 2015. URL: <http://cdsweb.cern.ch/record/2037697>.
- [72] ATLAS Collaboration. *Expected performance of missing transverse momentum reconstruction for the ATLAS detector at $\sqrt{s} = 13$ TeV.* ATL-PHYS-PUB-2015-023. 2015. URL: <http://cds.cern.ch/record/2037700>.

- [73] ATLAS Collaboration. *Performance of missing transverse momentum reconstruction for the ATLAS detector in the first proton-proton collisions at at $\sqrt{s} = 13$ TeV.* ATL-PHYS-PUB-2015-027. 2015. URL: <http://cds.cern.ch/record/2037904>.
- [74] ATLAS Collaboration. *Modelling of the $t\bar{t}H$ and $t\bar{t}V$ ($V = W, Z$) processes for $\sqrt{s} = 13$ TeV ATLAS analyses.* ATL-PHYS-PUB-2016-005. 2016. URL: <http://cds.cern.ch/record/2120826>.
- [75] M. V. Garzelli, A. Kardos, C. G. Papadopoulos, and Z. Trocsanyi. “ $t \bar{t} W^{+-}$ and $t \bar{t} Z$ Hadroproduction at NLO accuracy in QCD with Parton Shower and Hadronization effects.” In: *JHEP* 11 (2012), p. 056. doi: [10.1007/JHEP11\(2012\)056](https://doi.org/10.1007/JHEP11(2012)056). arXiv: [1208.2665 \[hep-ph\]](https://arxiv.org/abs/1208.2665).
- [76] J. M. Campbell and R. K. Ellis. “ $t\bar{t} W$ production and decay at NLO.” In: *JHEP* 1207 (2012), p. 052. arXiv: [1204.5678 \[hep-ph\]](https://arxiv.org/abs/1204.5678).
- [77] A. Lazopoulos, T. McElmurry, K. Melnikov, and F. Petriello. “Next-to-leading order QCD corrections to $t\bar{t}Z$ production at the LHC.” In: *Phys. Lett. B* 666 (2008), p. 62. arXiv: [0804.2220 \[hep-ph\]](https://arxiv.org/abs/0804.2220).
- [78] ATLAS Collaboration. *Simulation of top quark production for the ATLAS experiment at $\sqrt{s} = 13$ TeV.* ATL-PHYS-PUB-2016-004. 2016. URL: <http://cds.cern.ch/record/2120417>.
- [79] M. Czakon, P. Fiedler, and A. Mitov. “Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_s^4)$.” In: *Phys. Rev. Lett.* 110 (2013), p. 252004. arXiv: [1303.6254 \[hep-ph\]](https://arxiv.org/abs/1303.6254).
- [80] M. Czakon and A. Mitov. “Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders.” In: *Comput. Phys. Commun.* 185 (2014), p. 2930. doi: [10.1016/j.cpc.2014.06.021](https://doi.org/10.1016/j.cpc.2014.06.021). arXiv: [1112.5675 \[hep-ph\]](https://arxiv.org/abs/1112.5675).
- [81] N. Kidonakis. “Two-loop soft anomalous dimensions for single top quark associated production with a W^- or H^- .” In: *Phys. Rev. D* 82 (2010), p. 054018. doi: [10.1103/PhysRevD.82.054018](https://doi.org/10.1103/PhysRevD.82.054018). arXiv: [1005.4451 \[hep-ph\]](https://arxiv.org/abs/1005.4451).
- [82] ATLAS Collaboration. *Multi-Boson Simulation for 13 TeV ATLAS Analyses.* ATL-PHYS-PUB-2016-002. 2016. URL: <http://cds.cern.ch/record/2119986>.
- [83] J. M. Campbell and R. K. Ellis. “An update on vector boson pair production at hadron colliders.” In: *Phys. Rev. D* 60 (1999), p. 113006. arXiv: [hep-ph/9905386 \[hep-ph\]](https://arxiv.org/abs/hep-ph/9905386).
- [84] J. M. Campbell, R. K. Ellis, and C. Williams. “Vector boson pair production at the LHC.” In: *JHEP* 1107 (2011), p. 018. arXiv: [1105.0020 \[hep-ph\]](https://arxiv.org/abs/1105.0020).

- [85] ATLAS Collaboration. *Monte Carlo Generators for the Production of a W or Z/ γ^* Boson in Association with Jets at ATLAS in Run 2*. ATL-PHYS-PUB-2016-003. 2016. URL: <http://cds.cern.ch/record/2120133>.
- [86] S. Catani, L. Cieri, G. Ferrera, D. de Florian, and M. Grazzini. “Vector boson production at hadron colliders: a fully exclusive QCD calculation at NNLO.” In: *Phys. Rev. Lett.* 103 (2009), p. 082001. arXiv: [0903.2120 \[hep-ph\]](https://arxiv.org/abs/0903.2120).
- [87] S. Catani and M. Grazzini. “An NNLO subtraction formalism in hadron collisions and its application to Higgs boson production at the LHC.” In: *Phys. Rev. Lett.* 98 (2007), p. 222002. arXiv: [hep-ph/0703012 \[hep-ph\]](https://arxiv.org/abs/hep-ph/0703012).
- [88] ATLAS Collaboration. *Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton–proton collision data*. ATLAS-CONF-2014-032. 2014. URL: <http://cdsweb.cern.ch/record/1706245>.
- [89] ATLAS Collaboration. “Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s} = 13$ TeV.” In: *Eur. Phys. J. C* 76 (2016), p. 292. DOI: [10.1140/epjc/s10052-016-4120-y](https://doi.org/10.1140/epjc/s10052-016-4120-y). arXiv: [1603.05598 \[hep-ex\]](https://arxiv.org/abs/1603.05598).
- [90] ATLAS Collaboration. *Search for supersymmetry in final states with jets, missing transverse momentum and a Z boson at $\sqrt{s} = 8$ TeV with the ATLAS detector*. ATLAS-CONF-2012-152. 2012. URL: <http://cds.cern.ch/record/1493491>.
- [91] ATLAS Collaboration. “Measurement of the differential cross-section of highly boosted top quarks as a function of their transverse momentum in $\sqrt{s} = 8$ TeV proton-proton collisions using the ATLAS detector.” In: *Phys. Rev. D* 93.3 (2016), p. 032009. DOI: [10.1103/PhysRevD.93.032009](https://doi.org/10.1103/PhysRevD.93.032009). arXiv: [1510.03818 \[hep-ex\]](https://arxiv.org/abs/1510.03818).
- [92] CMS Collaboration. “Measurement of the integrated and differential t-tbar production cross sections for high-pt top quarks in pp collisions at $\sqrt{s} = 8$ TeV.” In: *Phys. Rev. D* (2016). Submitted. arXiv: [1605.00116 \[hep-ex\]](https://arxiv.org/abs/1605.00116).
- [93] ATLAS Collaboration. *Search for physics beyond the Standard Model in events with a Z boson and large missing transverse momentum using $\sqrt{s} = 7$ TeV pp collisions from the LHC with the ATLAS detector*. ATLAS-CONF-2012-046. 2012. URL: <http://cdsweb.cern.ch/record/1448222>.
- [94] Georges Aad et al. “Muon reconstruction efficiency and momentum resolution of the ATLAS experiment in proton-proton collisions at $\sqrt{s} = 7$ TeV in 2010.” In: *Eur. Phys. J. C* 74.9 (2014), p. 3034. DOI: [10.1140/epjc/s10052-014-3034-9](https://doi.org/10.1140/epjc/s10052-014-3034-9). arXiv: [1404.4562 \[hep-ex\]](https://arxiv.org/abs/1404.4562).

- [95] Georges Aad et al. “Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data.” In: *Eur. Phys. J.* C74.10 (2014), p. 3071. doi: [10.1140/epjc/s10052-014-3071-4](https://doi.org/10.1140/epjc/s10052-014-3071-4). arXiv: [1407.5063 \[hep-ex\]](https://arxiv.org/abs/1407.5063).
- [96] ATLAS Collaboration. “Search for squarks and gluinos in events with isolated leptons, jets and missing transverse momentum at $\sqrt{s} = 8$ TeV with the ATLAS detector.” In: *JHEP* 1504 (2015), p. 116. doi: [10.1007/JHEP04\(2015\)116](https://doi.org/10.1007/JHEP04(2015)116). arXiv: [1501.03555 \[hep-ex\]](https://arxiv.org/abs/1501.03555).
- [97] ATLAS Collaboration. “Improved luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC.” In: *Eur. Phys. J. C* 73 (2013), p. 2518. doi: [10.1140/epjc/s10052-013-2518-3](https://doi.org/10.1140/epjc/s10052-013-2518-3). arXiv: [1302.4393 \[hep-ex\]](https://arxiv.org/abs/1302.4393).
- [98] ATLAS Collaboration. “Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC.” In: *to be submitted to Eur. Phys. J. C* () .
- [99] Michael Kramer et al. *Supersymmetry production cross sections in pp collisions at $\sqrt{s} = 7$ TeV*. 2012. arXiv: [1206.2892 \[hep-ph\]](https://arxiv.org/abs/1206.2892).
- [100] W. Beenakker, R. Höpker, M. Spira, and P.M. Zerwas. “Squark and gluino production at hadron colliders.” In: *Nucl. Phys. B* 492 (1997), pp. 51–103. doi: [10.1016/S0550-3213\(97\)00084-9](https://doi.org/10.1016/S0550-3213(97)00084-9). arXiv: [hep-ph/9610490 \[hep-ph\]](https://arxiv.org/abs/hep-ph/9610490).
- [101] A. Kulesza and L. Motyka. “Threshold resummation for squark-antisquark and gluino-pair production at the LHC.” In: *Phys. Rev. Lett.* 102 (2009), p. 111802. doi: [10.1103/PhysRevLett.102.111802](https://doi.org/10.1103/PhysRevLett.102.111802). arXiv: [0807.2405 \[hep-ph\]](https://arxiv.org/abs/0807.2405).
- [102] A. Kulesza and L. Motyka. “Soft gluon resummation for the production of gluino-gluino and squark-antisquark pairs at the LHC.” In: *Phys. Rev. D* 80 (2009), p. 095004. doi: [10.1103/PhysRevD.80.095004](https://doi.org/10.1103/PhysRevD.80.095004). arXiv: [0905.4749 \[hep-ph\]](https://arxiv.org/abs/0905.4749).
- [103] W. Beenakker et al. “Soft-gluon resummation for squark and gluino hadroproduction.” In: *JHEP* 0912 (2009), p. 041. doi: [10.1088/1126-6708/2009/12/041](https://doi.org/10.1088/1126-6708/2009/12/041). arXiv: [0909.4418 \[hep-ph\]](https://arxiv.org/abs/0909.4418).
- [104] W. Beenakker et al. “Squark and gluino hadroproduction.” In: *Int. J. Mod. Phys. A* 26 (2011), pp. 2637–2664. doi: [10.1142/S0217751X11053560](https://doi.org/10.1142/S0217751X11053560). arXiv: [1105.1110 \[hep-ph\]](https://arxiv.org/abs/1105.1110).
- [105] M. Baak et al. “HistFitter software framework for statistical data analysis.” In: *Eur. Phys. J. C* 75 (2014), p. 153. doi: [10.1140/epjc/s10052-015-3327-7](https://doi.org/10.1140/epjc/s10052-015-3327-7). arXiv: [1410.1280 \[hep-ex\]](https://arxiv.org/abs/1410.1280).

- [106] Rene Brun and Fons Rademakers. "ROOT — An object oriented data analysis framework." In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 389.1 (1997), pp. 81–86. ISSN: 0168-9002. DOI: [http://dx.doi.org/10.1016/S0168-9002\(97\)00048-X](http://dx.doi.org/10.1016/S0168-9002(97)00048-X). URL: <http://www.sciencedirect.com/science/article/pii/S016890029700048X>.
- [107] Lorenzo Moneta, Kevin Belasco, Kyle Cranmer, Sven Kreiss, Alfio Lazzaro, Danilo Piparo, Gregory Schott, Wouter Verkerke, and Matthias Wolf. *The RooStats Project*. 2010. eprint: arXiv: [1009.1003](https://arxiv.org/abs/1009.1003).
- [108] Kyle Cranmer, George Lewis, Lorenzo Moneta, Akira Shibata, and Wouter Verkerke. "HistFactory: A tool for creating statistical models for use with RooFit and RooStats." In: (2012).
- [109] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. "Asymptotic formulae for likelihood-based tests of new physics." In: *Eur. Phys. J. C* 71 (2011), p. 1554. DOI: [10.1140/epjc/s10052-011-1554-0](https://doi.org/10.1140/epjc/s10052-011-1554-0). arXiv: [1007.1727 \[physics.data-an\]](https://arxiv.org/abs/1007.1727).
- [110] A. Read. "Presentation of search results: the CLs technique." In: *Journal of Physics G: Nucl. Part. Phys.* 28 (2002), pp. 2693–2704. DOI: [10.1088/0954-3899/28/10/313](https://doi.org/10.1088/0954-3899/28/10/313).
- [111] *Prospects for a search for direct pair production of top squarks in scenarios with compressed mass spectra at the high luminosity LHC with the ATLAS Detector*. Tech. rep. ATL-PHYS-PUB-2016-022. Geneva: CERN, 2016. URL: <https://cds.cern.ch/record/2220904>.