

UNIVERSITY OF CALIFORNIA
SANTA CRUZ

**SEARCH FOR WW AND WZ RESONANCE PRODUCTION IN
 $\ell\nu qq$ FINAL STATES IN pp COLLISIONS AT $\sqrt{s} = 13$ TEV WITH
THE ATLAS DETECTOR**

A dissertation submitted in partial satisfaction of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

in

PHYSICS

by

Natasha Woods

December 2019

The Dissertation of Natasha Woods
is approved:

Abraham Seiden, Chair

Mike Hance

Bruce Schumm

Quentin Williams
Vice Provost and Dean of Graduate Studies

Copyright © by

Natasha Woods

2019

Table of Contents

List of Figures	vi
List of Tables	xv
Abstract	xvii
Dedication	xviii
Acknowledgments	xix
I Theoretical Motivation	2
1 The Standard Model of Particle Physics	3
1.1 Introduction	3
1.2 Quantum Field Theory	3
1.3 $U(1)_{EM}$ Local Gauge Invariance	4
1.4 Yang-Mills Gauge Theories	7
1.5 Particles in the Standard Model	8
1.6 Higgs Mechanism	13
1.7 Electroweak Theory	14
1.8 Quantum ChromoDynamics	15
2 Standard Model Successes and Limitations	20
3 New Physics Models with Diboson Resonances	23
3.1 Randall Sundrum Bulk Model	23
3.2 Extended Scalar Sector	25
3.3 Simple Standard Model Extensions	26

II Experimental Setup	29
4 LHC	30
4.1 LHC Layout and Design	32
5 The ATLAS Detector	37
5.1 Coordinate System	39
5.2 Inner Detector	40
5.2.1 Pixel Detector	43
5.2.2 Semiconductor Tracker	43
5.2.3 Transition Radiation Tracker	43
5.3 Calorimeters	45
5.4 Muon Spectrometer	48
5.5 Magnet System	52
5.6 Trigger System	53
III Method	55
6 Dataset and Simulated Samples	56
6.1 Dataset	56
6.2 Simulated Samples	59
6.3 Object Selection	59
6.3.1 Electrons	59
6.3.2 Muons	60
6.3.3 small-R jets	62
6.3.4 large-R jets	65
6.3.5 Variable Radius jets	68
6.3.6 MET/neutrinos	68
6.3.7 Jet Flavor Tagging	68
6.3.8 Overlap Removal	69
7 Event Selection and Categorization	71
7.1 Pre-selection	71
7.2 Trigger	71
7.3 GGF/VBF RNN	74
7.4 Topological Cuts	79
7.5 Selection Acceptance times efficiency for Signal Events	89
7.6 Background Estimate	90
7.6.1 Multijet Sample	90

8 Systematic Uncertainties	104
8.1 Experimental Systematics	104
8.2 Theory Systematics	106
9 Statistical Analysis	114
9.1 Likelihood Function Definition	114
9.2 Fit Configuration	115
9.3 Best Fit μ	117
9.4 Discovery Test	118
9.5 Exclusion Limits	119
IV Results	121
10 Statistical Interpretation	122
10.1 Discovery Tests	122
10.2 Systematic Profiling and Correlations	128
10.3 Expected and Measured Yields	129
10.4 Limits	143
V Quark and Gluon Tagging	145
11 Prospects	146
12 n_{trk} Calibration	153
13 Application	159
VI Conclusion	164
14 Conclusions	165
Bibliography	166

List of Figures

1.1	The particles of the Standard Model.	10
1.2	Summary of how Standard Model particles interact with other Standard Model particles.	11
1.3	This figure shows the three dominant QCD interactions. From Ref. [14]	17
1.4	Strength of the U(1), SU(2), and SU(3) gauge couplings as a function of the energy scale of the interaction (Q). From Ref. [10]	18
2.1	A comparison of cross section measurements at $\sqrt{s} = 7, 8, 13$ TeV from ATLAS compared to theoretical measurements. From Ref. [5]	22
3.1	Cartoon of RS Bulk Model	24
4.1	Scaling of cross sections with \sqrt{s} . Natasha: write more here	31
4.2	LHC Layout. Natasha write more	33
4.3	LHC Accelerator. Natasha write more	35
5.1	Big picture layout of ATLAS detector. Natasha: write more	38
5.2	Big picture layout of ATLAS detector. Natasha: write more	38
5.3	A simplified schematic of how different particles interact and are detected within ATLAS.	39
5.4	Layout of ATLAS Inner Detector	41
5.5	Layout of ATLAS ID Barrel System.	42
5.6	Overview of ATLAS electromagnetic and hadronic calorimeters.	46

5.7	Schematic of ECAL.	47
5.8	Schematic of HCAL.	47
5.9	Schematic of Muon Spectrometer [cite G35]	50
5.10	Schematic of MDT chamber. [cite G35]	51
5.11	Schematic of RPC chamber, which is used for triggering in the central region of the detector [cite G35].	51
5.12	Schematic of TGC chamber, which is used for triggering in the muon end-cap region. [cite G35]	52
5.13	Layout of ATLAS magnet systems.	53
6.1	Integrated luminosity for data collected from ATLAS from 2011 - 2018	57
6.2	Mean number of interactions per crossing for data collected from ATLAS from 2011 - 2018	58
6.3	[4] This figure show the breakdown of the muon reconstruction efficiency scale factor measured in $Z \rightarrow \mu\mu$ as a function of p_T	62
6.4	[6] This diagram shows the calibration stages for EM jets.	65
6.5	The upper cut on D_2 (a) and jet mass window cut i.e. the upper and lower boundary of the mass (b) of the W -tagger as a function of jet p_T . Corresponding values for Z -tagger are shown in (c) and (d). The optimal cut values for maximum significance are shown as solid markers and the fitted function as solid lines. Working points from $VV \rightarrow JJ$ [ATLAS-HDBS-2018-31-002] is also shown as dashed lines as a reference. Natasha reword?	67
6.6	Natasha write caption	67
7.1	RNN architecture. Natasha add caption	76
7.2	RNN Score distribution for ggF and VBF signals and backgrounds.	77
7.3	ROC curve using k-fold validation for RNN.	78
7.4	Comparison of GGF Z' limits for different RNN score selections.	79
7.5	Event Categorization. Natasha write more.	82
7.6	Data MC comparison for the merged WW HP TCR.	83

7.7	Data MC comparison for the merged WW LP TCR.	84
7.8	Data MC comparison for the merged WZ HP TCR.	85
7.9	Data MC comparison for the merged WZ LP TCR.	86
7.10	Data MC comparison for the resolved WW TCR.	87
7.11	Data MC comparison for the resolved WZ TCR.	88
7.12	Selection acceptance times efficiency for the $W' \rightarrow WZ \rightarrow \ell\nu qq$ events from MC simulations as a function of the W' mass for (a) Drell-Yan and (b) VBF production, combining the merged HP and LP signal regions of the $WW \rightarrow \ell\nu J$ selection and the resolved regions of the $WW \rightarrow \ell\nu jj$ selection.	89
7.13	Selection acceptance times efficiency for the $G \rightarrow WW \rightarrow \ell\nu qq$ events from MC simulations as a function of the G mass for (a) Drell-Yan and (b) VBF production, combining the merged HP and LP signal regions of the $WW \rightarrow \ell\nu J$ selection and the resolved regions of the $WW \rightarrow \ell\nu jj$ selection.	90
7.14	The E_T^{miss} distribution in MJCR for 2017 data in the electron channel(left), muon channel with W-boson pT < 150 GeV (center) and > 150 GeV (right). Multi-jet templates are calculated as remaining data components after excluding known MC	92
7.15	Postfit Data/MC comparison of distributions of E_T^{miss} , m_T^W , lepton and neutrino p_T , $m_{\ell\nu jj}$, lepton- ν angular distance in the WW electron channel. The MJ template is obtained from the pre-MJ-fit.	93
7.16	Postfit Data/MC comparison of distributions of E_T^{miss} , m_T^W , lepton and neutrino p_T , $m_{\ell\nu jj}$, lepton- ν angular distance in the WW muon channel. The MJ template is obtained from the pre-MJ-fit.	94
7.17	Postfit Data/MC comparison of distributions of E_T^{miss} , m_T^W , lepton and neutrino p_T , $m_{\ell\nu jj}$, lepton- ν angular distance in the WZ untag electron channel. The MJ template is obtained from the pre-MJ-fit.	95
7.18	Postfit Data/MC comparison of distributions of E_T^{miss} , m_T^W , lepton and neutrino p_T , $m_{\ell\nu jj}$, lepton- ν angular distance in the WZ untag muon channel. The MJ template is obtained from the pre-MJ-fit.	96

7.19 Postfit Data/MC comparison of distributions of E_T^{miss} , m_T^W , lepton and neutrino p_T , $m_{\ell\nu jj}$, lepton- ν angular distance in the WZ untag electron channel. The MJ template is obtained from the pre-MJ-fit.	97
7.20 Postfit Data/MC comparison of distributions of E_T^{miss} , m_T^W , lepton and neutrino p_T , $m_{\ell\nu jj}$, lepton- ν angular distance in the WZ untag muon channel. The MJ template is obtained from the pre-MJ-fit.	98
7.21 Postfit Data/MC comparison of distributions of E_T^{miss} , m_T^W , lepton and neutrino p_T , $m_{\ell\nu jj}$, lepton- ν angular distance in the VBF WW electron channel. The MJ template is obtained from the pre-MJ-fit.	99
7.22 Postfit Data/MC comparison of distributions of E_T^{miss} , m_T^W , lepton and neutrino p_T , $m_{\ell\nu jj}$, lepton- ν angular distance in the VBF WW muon channel. The MJ template is obtained from the pre-MJ-fit.	100
7.23 Postfit Data/MC comparison of distributions of E_T^{miss} , m_T^W , lepton and neutrino p_T , $m_{\ell\nu jj}$, lepton- ν angular distance in the VBF WZ electron channel. The MJ template is obtained from the pre-MJ-fit.	101
7.24 Postfit Data/MC comparison of distributions of E_T^{miss} , m_T^W , lepton and neutrino p_T , $m_{\ell\nu jj}$, lepton- ν angular distance in the VBF WZ muon channel. The MJ template is obtained from the pre-MJ-fit.	102
8.1 The W/Z +jet systematics for the a) Merged ggF, b) Resolved ggF, c) Merged VBF, and d) Resolved VBF regions. The top subplot shows the nominal and variation distributions/bands, the middle shows the ratio of the two, and the final shows just the shape of the envelope (the final uncertainty).	108
8.2 The two-point generator comparison between Sherpa and MadGraph for the W/Z +jet samples in the a) Merged ggF, b) Resolved ggF, c) Merged VBF, and d) Resolved VBF regions. The normalization of the Madgraph sample is set to the Sherpa value to consider only shape effects. The bottom inlet shows the ratio of the two.	109

8.3	Ratio between the variations of generator (red) and hadronization (blue) variations for the Merged regime for $t\bar{t}$ sample.	110
8.4	Ratio between the variations of generator (red) and hadronization (blue) variations for the Resolved regime for $t\bar{t}$ sample.	110
8.5	Ratio between the variations of ISR up (red) and down (blue) variations for the Merged regime for $t\bar{t}$ sample.	111
8.6	Ratio between the variations of ISR up (red) and down (blue) variations for the Resolved regime for $t\bar{t}$ sample.	112
8.7	Ratio between the variations of FSR up (red) and down (blue) variations for the Merged regime for $t\bar{t}$ sample.	113
8.8	Ratio between the variations of FSR up (red) and down (blue) variations for the Resolved regime for $t\bar{t}$ sample.	113
9.1	The HVT signal mass resolution as a function of mass fit with a straight line in the Resolved ggF region (left) and VBF (right) region.	117
9.2	The HVT signal mass resolution as a function of mass fit with a straight line in the Merged ggF region (left) and VBF (right) region.	117
10.1	These plots show the measured p_0 value as a function of resonance mass for HVT Z' DY production.	123
10.2	These plots show the measured p_0 value as a function of resonance mass for HVT Z' VBF production.	124
10.3	These plots show the measured p_0 value as a function of resonance mass for HVT W' DY production.	125
10.4	These plots show the measured p_0 value as a function of resonance mass for HVT W' VBF production.	126
10.5	These plots show the measured p_0 value as a function of resonance mass for the RS Graviton DY production.	127
10.6	Ranked systematics and their fitted values for WW DY (right) and VBF (left) selections.	128

10.7	Ranked systematics and their fitted values for WZ DY (right) and VBF (left) selections.	129
10.8	Correlations between systematics for WW DY (right) and VBF (left) selections.	129
10.9	This figure shows the distribution of $m_{\ell\nu qq}$ in the DY WW control regions.	137
10.10	This figure shows the distribution of $m_{\ell\nu qq}$ in the DY WZ control regions.	138
10.11	This figure shows the distribution of $m_{\ell\nu qq}$ in the VBF WW control regions.	139
10.12	This figure shows the distribution of $m_{\ell\nu qq}$ in the VBF WZ control regions.	140
10.13	This figure shows the distribution of $m_{\ell\nu qq}$ in the GGF WW signal regions.	141
10.14	This figure shows the distribution of $m_{\ell\nu qq}$ in the GGF WZ Untag signal regions.	141
10.15	This figure shows the distribution of $m_{\ell\nu qq}$ in the GGF WZ Tag signal regions.	142
10.16	This figure shows the distribution of $m_{\ell\nu qq}$ in the VBF WZ Tag signal regions.	142
10.17	This figure shows the distribution of $m_{\ell\nu qq}$ in the VBF WZ Tag signal regions.	143
10.18	This figure shows theory, expected and observed limits for HVT W' DY (left) and VBF (right) production.	143
10.19	This figure shows theory, expected and observed limits for HVT Z' DY (left) and VBF (right) production.	144
10.20	This figure shows theory, expected and observed limits for RS Gravitons via DY production.	144

11.1	PDGID of the truth-level parton matched to the small-R jets passing the Resolved GGF WW Signal Region selections for the (a) Leading (b) Sub-Leading jets . These distributions are shown for 300, 500, and 700GeV Z' signals and the background.	148
11.2	The number of tracks in small-R jets in events passing the Resolved GGF WW Signal Region selections for the (a) Leading (b) Sub-Leading jets. These distributions are shown for 300, 500, and 700GeV Z' signals and the background.	148
11.3	The number of tracks in background small-R jets in events passing the Resolved GGF WW Signal region selection vs. $\ln(p_T)$ for (a)Leading (b) Sub-Leading jets. The best fit line for the distribution is also shown, as well as the percentage of jets that pass a cut of number of tracks $< 4 \times \ln(p_T) - 5$. Note the number of total entries in these plots has been normalized to one.	149
11.4	The number of tracks in small-R jets in 300GeV Z' events passing the Resolved GGF WW Signal region selection vs. $\ln(p_T)$ for (a)Leading (b) Sub-Leading jets. The best fit line for the distribution is also shown, as well as the percentage of jets that pass a cut of number of tracks $< 4 \times \ln(p_T) - 5$.Note the number of total entries in these plots has been normalized to one.	149
11.5	The number of tracks in small-R jets in 500GeV Z' events passing the Resolved GGF WW Signal region selection vs. $\ln(p_T)$ for (a)Leading (b) Sub-Leading jets. The best fit line for the distribution is also shown, as well as the percentage of jets that pass a cut of number of tracks $< 4 \times \ln(p_T) - 5$.Note the number of total entries in these plots has been normalized to one.	150

11.6 The number of tracks in small-R jets in 700GeV Z' events passing the Resolved GGF WW Signal region selection vs. $\ln(p_T)$ for (a)Leading (b) Sub-Leading jets. The best fit line for the distribution is also shown, as well as the percentage of jets that pass a cut of number of tracks $< 4 \times \ln(p_T) - 5$.Note the number of total entries in these plots has been normalized to one.	150
11.7 The number of tracks in leading small-R jets in background events passing the Resolved GGF WW Signal region selection vs. $\ln(p_T)$ for (a)Gluons (b) Quarks jets. The best fit line for the distribution is also shown, as well as the percentage of jets that pass a cut of number of tracks $< 4 \times \ln(p_T) - 5$.Note the number of total entries in these plots has been normalized to one.	151
11.8 ROC Curve for Quark and Gluon Tagging with a cut on the number of tracks that depends on the $\ln(p_T)$	151
11.9 The top panel shows the distribution of m_{lvqq} with and without quark gluon tagging. The middle panel shows the ratio of the signals and backgrounds with and without quark gluon tagging. The bottom panel shows the change in S/\sqrt{B} with quark gluon tagging.	152
13.1 The top panel shows the distribution of m_{lvqq} with and without requiring jets to be true quarks. The middle panel shows the ratio of the signals and backgrounds with and without requiring jets to be true quarks. The bottom panel shows the change in S/\sqrt{B} when requiring jets to be true quarks.	160
13.2 Unfolded and extracted n_C qg dstbs.	161
13.3 PDGID of the truth-level parton matched to the small-R jets passing the Resolved GGF WW Signal Region selections for the (a) Leading (b) Sub-Leading jets . These distributions are shown for 300, 500, and 700GeV Z' signals and the background.	162

13.4 These figures show the impact of the uncertainties on the number
of tracks in the leading jet in the sum of the background sample
in the Resolved GGF WW SR (a) tracking efficiency (b) fake (c)
PDF (d) ME (e) unfolding uncertainties. 163

List of Tables

1.1	Representations of the SM fermions under $SU(3)_C \times SU(2)_L \times U(1)_Y$ gauge symmetry group. $SU(2)_L$ gauge transformations allow one to go between rows and $SU(3)_C$ transformations allow one to go between columns in these fermion representations. [REWORD]	9
7.1	The list of triggers used in the analysis.	73
7.2	Summary of selection criteria used to define the signal region (SR), W +jets control region (W CR) and $t\bar{t}$ control region ($t\bar{t}$ CR) for merged 1-lepton channel.	81
7.3	The list of selection cuts in the resolved analysis for the WW and WZ signal regions (SR), W +jets control region (WR) and $t\bar{t}$ control region (TR).	82
7.4	Definitions of “inverted” leptons used in multijet control region . .	92
7.5	Fit validation result in WCRs for 2015+16 data. The fit is done in various WCRs, in order to obtain the corresponding scale factors for MJ templates: ggF resolved WCR for the $WW \rightarrow lvqq$ selection, ggF resolved untagged WCR for the $WZ \rightarrow lvqq$ selection, ggF resolved tagged WCR for the $WZ \rightarrow lvqq$ selection, VBF resolved WCR for the $WW \rightarrow lvqq$ selection, and VBF resolved WCR for the $WZ \rightarrow lvqq$ selection. Post-fit event yields for electroweak processes and MJ contributions are shown. The SF column shows the corresponding normalization scale factors for electroweak processes from the fit. R.U. stands for relative uncertainty.	103

10.1	Expected and Measured for DY WW $W+\text{jets}$, $t\bar{t}$ control regions and signal regions.	130
10.2	Expected and Measured for DY WZ $W+\text{jets}$, $t\bar{t}$ tag and untag control regions.	131
10.3	Expected and Measured for DY WZ $W+\text{jets}$, $t\bar{t}$ tag and untag signal regions.	132
10.4	Expected and Measured for VBF WW $W+\text{jets}$, $t\bar{t}$ control regions and signal regions.	133
10.5	Expected and Measured for VBF WZ $W+\text{jets}$, $t\bar{t}$ control regions and signal regions.	134
10.6	Fitted background normalizations for $t\bar{t}$ and $W+\text{jets}$ backgrounds for the DY WW analysis region.	135
10.7	Fitted background normalizations for $t\bar{t}$ and $W+\text{jets}$ backgrounds for the DY WZ analysis region.	135
10.8	Fitted background normalizations for $t\bar{t}$ and $W+\text{jets}$ backgrounds for the VBF WW analysis region.	135
10.9	Fitted background normalizations for $t\bar{t}$ and $W+\text{jets}$ backgrounds for the VBF WZ analysis region.	136

Abstract

Search for WW and WZ resonance production in $\ell\nu qq$ final states in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector

by

Natasha Woods

This thesis reviews a search for WW and WZ resonance production using data from pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector, corresponding to an integrated luminosity of 139 fb^{-1} . Diboson resonances are predicted in a number of Standard Model (SM) extensions, such as Extended Gauge Models, Extra dimensions, and technicolor models. This search looks for resonances where one W boson decays leptonically and the other W or Z boson decays hadronically. This search is sensitive to diboson resonance production via vector-boson fusion as well as quark-antiquark annihilation and gluon-gluon fusion mechanisms. No significant excess of events is observed with respect to the Standard Model backgrounds. As the dominant backgrounds in this search contain gluons, classifying jets as quark or gluon initiated would make this analysis more sensitive to new physics. Towards this end, this thesis considers the prospects for adding a quark gluon tagger based on the number of tracks in jets and reviews the calibration of the number of tracks in jets.

A loving dedication.

å

xviii

Acknowledgments

Proper acknowledgments of everyone else who helped you graduate. Write later.

₁ another intro???

²

Part I

³

Theoretical Motivation

⁴ **Chapter 1**

⁵ **The Standard Model of Particle
6 Physics**

⁷ **1.1 Introduction**

⁸ By determining the dynamics of the most elementary degrees of freedom, par-
⁹ ticle physics hopes to uncover the fundamental laws of the universe. The definition
¹⁰ of elementary has evolved through time and currently refers to matter and force
¹¹ mediating particles: fermions and bosons, respectively. The Standard Model of
¹² Particle Physics (SM) describes the quantum behavior of three of the four funda-
¹³ mental forces: weak, strong, and electromagnetic, via boson and fermion interac-
¹⁴ tions. Gravity is not included in the SM and still under investigation.

¹⁵ **1.2 Quantum Field Theory**

¹⁶ In the SM, forces (and particles) are represented as fields. In this context,
¹⁷ fields are mathematical objects that define a tensor (e.g. scalar, vector, etc) at
¹⁸ every point on a manifold, here the manifold is space-time. These fields obey laws

¹⁹ dictated by Quantum Field Theory (QFT). Particles arise naturally in QFT as
²⁰ quantized field excitations localized in spacetime.

²¹ According to Noether's theorem, symmetries of a field give rise to conserved
²² quantities (e.g. time-translation invariance leads to energy conservation). Often
²³ in the history of physics, a conserved quantity of a field is found and then the
²⁴ underlying symmetry of the field is inferred. Gauge symmetries are symmetries
²⁵ among the internal degrees of freedom of the field (components of the tensor),
²⁶ which give rise to quantities associated with fields. By specifying the symmetries
²⁷ of a system the dynamics and conserved quantities of the system may be succinctly
²⁸ defined.

²⁹ 1.3 $U(1)_{EM}$ Local Gauge Invariance

³⁰ The Lagrangian of Quantum Electrodynamics (QED) describes the electro-
³¹ magnetic force. QED may be derived by requiring local $U(1)_{EM}$ gauge invariance
³² of the free dirac fermion Lagrangian, ψ :

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi \quad (1.1)$$

³³ This symmetry may be represented as a complex number with unit modulus,
³⁴ $e^{i\theta}$. $U(1)$ gauge invariance requires this gauge transformation of ψ will leave the
³⁵ Lagrangian unchanged.

$$\psi(x) \rightarrow \psi'(x) = e^{i\theta(x)}\psi(x) \quad (1.2)$$

³⁶ NB: This transformation is a local gauge transformation as θ depends on the
³⁷ spacetime coordinate.

³⁸ By requiring this symmetry of the free Dirac fermion Lagrangian:

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi \quad (1.3)$$

³⁹ The mass term is unaffected, but the kinetic term is modified due to $\theta(x)$.

$$\mathcal{L} \rightarrow \mathcal{L}' = i\bar{\psi}e^{-i\theta(x)}\gamma^\mu\partial_\mu\psi e^{i\theta(x)} - m\bar{\psi}e^{-i\theta(x)}\psi e^{i\theta(x)} \quad (1.4)$$

⁴⁰

$$= i\bar{\psi}\gamma^\mu(\partial_\mu\psi + i\psi\partial_\mu\theta) - m\bar{\psi}\psi \quad (1.5)$$

⁴¹ The $\partial_\mu\theta$ terms breaks the gauge invariance of the Lagrangian. By introducing a
⁴² new field, A_μ we can recover the gauge invariance of the derivative. Now redefining
⁴³ the derivative as the covariant derivative:

$$D_\mu\psi \equiv (\partial_\mu - iqA_\mu)\psi \quad (1.6)$$

⁴⁴ And letting A_μ transform under $U(1)$ as:

$$A_\mu \rightarrow A_\mu + \delta A_\mu \quad (1.7)$$

⁴⁵ The transformed covariant derivative becomes:

$$D_\mu\psi \rightarrow D_\mu\psi' = (\partial_\mu - iqA_\mu)\psi' \quad (1.8)$$

⁴⁶

$$= (\partial_\mu - iq(A_\mu + \delta A_\mu))\psi e^{i\theta} \quad (1.9)$$

⁴⁷

$$= e^{i\theta}D_\mu + ie^{i\theta}\psi(\partial_\mu\theta - q\delta A_\mu) \quad (1.10)$$

⁴⁸ The covariant derivative can be made gauage invariant by setting the last term
⁴⁹ to zero.

$$\delta A_\mu = \frac{1}{q} \partial_\mu \theta \quad (1.11)$$

50 So now A_μ transforms as:

$$A_\mu \rightarrow A_\mu + \frac{1}{q} \partial_\mu \theta \quad (1.12)$$

51 Finally, replacing the derivative with the covariant derivative the Dirac La-
52 grangian we have:

$$\mathcal{L} = i\bar{\psi}\gamma^\mu D_\mu \psi - m\bar{\psi}\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \quad (1.13)$$

53

$$= \mathcal{L}_{QED} \quad (1.14)$$

54 Here $F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu$. This last term in the Lagrangian is the kinetic
55 energy of the gauge boson field.

56 So we have derived the QED Lagrangian. By requiring the free Dirac La-
57 grangian to be invariant under U(1) transformations we have generated a new
58 gauge boson field, A_μ , which describes the photon. As expected the photon inter-
59 acts with fermions.

60 Stepping back, a global U(1) gauge symmetry of the free Dirac Lagrangian
61 implies we cannot measure the absolute phase of a charged particle. A local U(1)
62 gauge symmetry changes the phase of fields differently across space time. For this
63 type of transformation to leave the Lagrangian invariant, we had to introduce an
64 additional field, A_μ , which "communicates" these phase changes across space-time.
65 In less formal language this effectively means: if the field at one location changes,
66 this change is conferred to other particles via A_μ .

67 1.4 Yang-Mills Gauge Theories

68 Requiring $U(1)_{EM}$ gauge invariance of the free Dirac Lagrangian gave us QED.

69 Requiring different gauge symmetries we can derive the structure of other inter-
70 actions. Any gauge symmetry may be written as:

$$\psi_i \rightarrow \exp(i\theta^a T_{ij}^a) \psi_j \quad (1.15)$$

71 Here θ is a dimensionless real parameter and T is the generator of the gauge
72 symmetry group. With this the covariant derivative can be written as:

$$D_\mu \psi_i \equiv \partial_\mu \psi_i + ig A_\mu^a T_{ij}^a \psi_j \quad (1.16)$$

73 Then the gauge field must transform as:

$$A_\mu^a \rightarrow A_\mu^a - \frac{1}{g} \partial_\mu \theta^a - f^{abc} \theta^b A_\mu^c \quad (1.17)$$

74 Here f is the structure constant of the gauge group. The field strength tensor
75 is given by:

$$F_{\mu\nu}^a \equiv \partial_\mu A_\nu^a - \partial_\nu A_\mu^a - g f^{abc} A_\mu^b A_\nu^c \quad (1.18)$$

76

$$F_{\mu\nu}^a \rightarrow F_{\mu\nu}^a - f^{abc} \theta^b F_{\mu\nu}^c \quad (1.19)$$

77 This gives the Yang-Mills Lagrangian:

$$\mathcal{L}_{YM} = -\frac{1}{4} F_{\mu\nu}^a F_{\mu\nu}^a + i \bar{\psi}_i \gamma^\mu D_\mu \psi_i + m \bar{\psi}_i \psi_i \quad (1.20)$$

78 1.5 Particles in the Standard Model

79 The SM consists of fermions (half-integer spin matter constituents) and bosons
80 (integer spin force mediators). Fermions are spinor representations of the Poincare
81 group and can be further separated into leptons and quarks. Bosons are the result
82 of requiring a particular symmetry among the spinor fields:

$$SU(3)_C \times SU(2)_L \times U(1)_Y \quad (1.21)$$

83 $SU(3)_C$ is the symmetry group of the strong force and generates eight gluon
84 fields, G_μ . $SU(2)_L$ is the symmetry group of the Electroweak force and generates
85 three electroweak boson fields, and $U(1)_Y$ generates the photon field, where Y is
86 the weak-hypercharge:

$$Y = 2(Q - T_3) \quad (1.22)$$

87 Q is the electromagnetic charge, and T_3 is the z-component of the weak isospin.
88 Weak isospin is the charge associated with the $SU(2)_L$ symmetry. The correspond-
89 ing covariant derivative is then:

$$D_\mu \phi \equiv (\partial_\mu + ig_1 B_\mu Y_{L/R} + [ig_2 W_\mu^\alpha T^\alpha]_L + [ig_3 G_\mu^\alpha \tau^\alpha]_C) \psi \quad (1.23)$$

90 It is important to note that the gauge symmetry of the SM yields a particular
91 structure of the fermion representations. So for a given fermion to interact with
92 a given gauge field it must have a non-zero corresponding Noether charge for
93 that gauge symmetry. If the corresponding Noether charge is zero, that fermion
94 transforms as a singlet and does not participate in that gauge interaction.

95 Fermions are divided into quarks and leptons based on their transformations
96 under $SU(3)_C$. Quarks transform as color triplets. Leptons transform as color

singlets and consequently do not interact with gluons. Fermions may be further
 classified by their $SU(2)_L$ interactions. Only the left-chiral part of fermions (denoted by L here) transform as $SU(2)_L$ doublets, the right-chiral part forms singlets under this gauge. Lastly, all these groups of particles come in three generations, each a heavier copy of the previous, but with differing flavor quantum numbers.
 This is summarized in Table 1.1 and shown in Figures 1.1 and 1.2.

SM Fermion Gauge Group	First Generation	Second Generation	Third Generation	$(SU(3)_C, SU(2)_L, U(1)_Y)$ Representations
Left-handed quarks	$\begin{pmatrix} u_L^r & u_L^g & u_L^b \\ d_L^r & d_L^g & d_L^b \end{pmatrix}$	$\begin{pmatrix} c_L^r & c_L^g & c_L^b \\ s_L^r & s_L^g & s_L^b \end{pmatrix}$	$\begin{pmatrix} t_L^r & t_L^g & t_L^b \\ b_L^r & b_L^g & b_L^b \end{pmatrix}$	$(3, 2, \frac{1}{6})$
Right-handed quarks	(u_R^r, u_R^g, u_R^b) (d_R^r, d_R^g, d_R^b)	(c_R^r, c_R^g, c_R^b) (s_R^r, s_R^g, s_R^b)	(t_R^r, t_R^g, t_R^b) (b_R^r, b_R^g, b_R^b)	$(3, 1, \frac{2}{3})$ $(3, 1, -\frac{1}{3})$
Left-handed leptons	$\begin{pmatrix} \nu_e^L \\ e_L \end{pmatrix}$	$\begin{pmatrix} \mu_e^L \\ \mu_L \end{pmatrix}$	$\begin{pmatrix} \tau_e^L \\ \tau_L \end{pmatrix}$	$(1, 2, -\frac{1}{2})$
Right-handed leptons	e_R	μ_R	τ_R	$(1, 1, -1)$

Table 1.1: Representations of the SM fermions under $SU(3)_C \times SU(2)_L \times U(1)_Y$ gauge symmetry group. $SU(2)_L$ gauge transformations allow one to go between rows and $SU(3)_C$ transformations allow one to go between columns in these fermion representations. [REWORD]

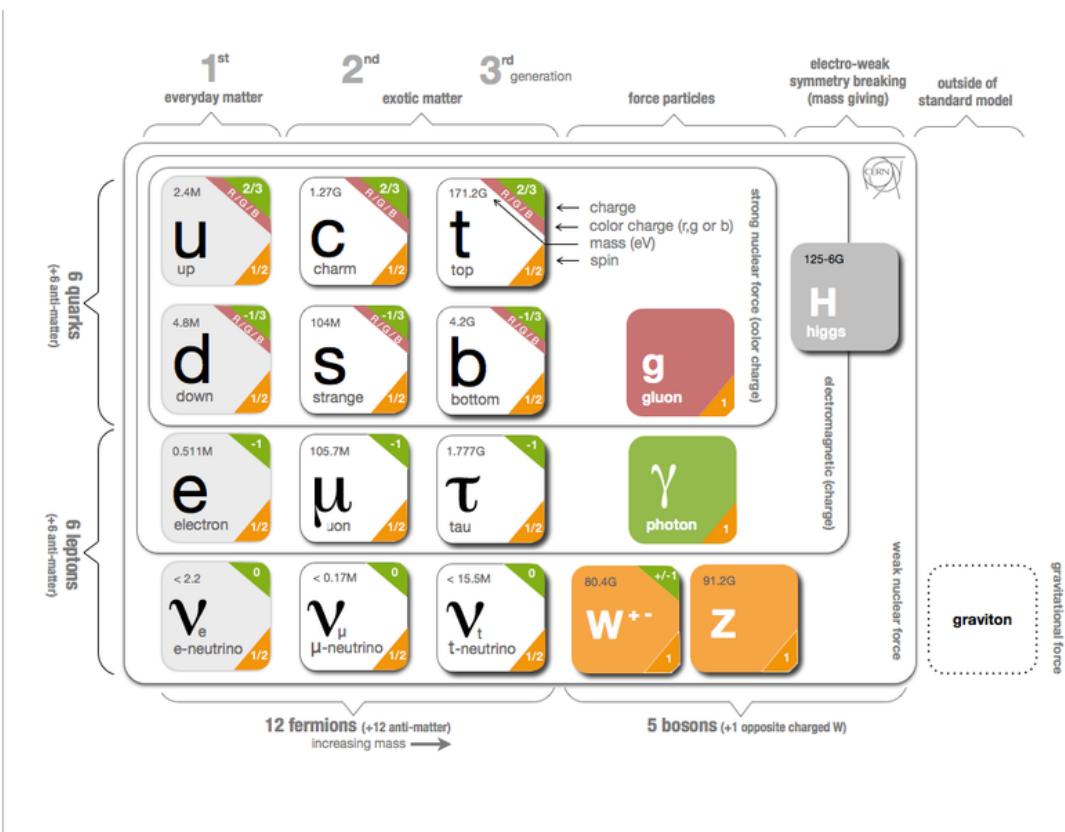


Figure 1.1: The particles of the Standard Model.

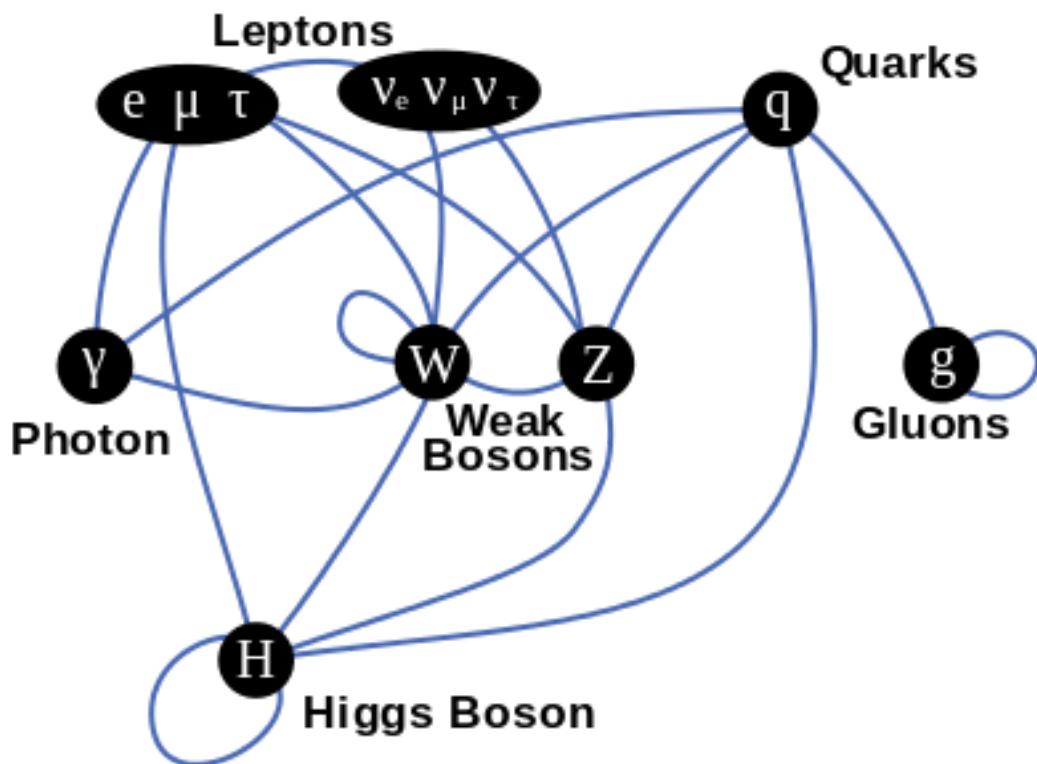


Figure 1.2: Summary of how Standard Model particles interact with other Standard Model particles.

¹⁰³ Now we can understand the SM Lagrangian density as a Yang-Mills theory
¹⁰⁴ with the gauge group: $SU(3)_C \times SU(2)_L \times U(1)_Y$ with an additional $SU(2)$ complex
¹⁰⁵ scalar Higgs field doublet that will be discussed later.

$$\begin{aligned} \mathcal{L}_{SM} = & \underbrace{-\frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}W_{\mu\nu}^aW^{a\mu\nu} - \frac{1}{4}G_{\mu\nu}^\alpha G^{\alpha\mu\nu}}_{\text{Kinetic Energies and Self-Interactions of Gauge Bosons}} \\ & + \underbrace{\bar{L}_i \gamma^\mu (i\partial_\mu - \frac{1}{2}g_1 Y_{iL} B_\mu - \frac{1}{2}g_2 \sigma^a W_\mu^a) L_i}_{\text{Kinetic Energies and EW Interactions of Left-handed Fermions}} \\ & + \underbrace{\bar{R}_i \gamma^\mu (i\partial_\mu - \frac{1}{2}g_1 Y_{iR} B_\mu) R_i}_{\text{Kinetic Energies and EW Interactions of Right-Handed Fermions}} \\ & + \underbrace{\frac{ig_3}{2} \bar{Q}_j \gamma^\mu \lambda^\alpha G_\mu^\alpha Q_j}_{\text{Strong Interactions between Quarks and Gluons}} \\ & + \underbrace{\frac{1}{2} |(i\partial_\mu - \frac{1}{2}g_1 B_\mu - \frac{1}{2}g_2 \sigma^a W_\mu^a)\Phi|^2 - V(\Phi)}_{\text{Electroweak Boson Masses and Higgs Couplings}} \\ & - \underbrace{(y_{kl}^d \bar{L}_k \Phi R_l + y_{kl}^u \bar{R}_k \tilde{\Phi} L_l + h.c.)}_{\text{Fermion Mass terms and Higgs Couplings}} \end{aligned}$$

¹⁰⁶ Here several abstract spaces are being spanned:

- ¹⁰⁷ – a spans the three $SU(2)_L$ gauge fields with generators expanded in Pauli
¹⁰⁸ matrices, $T^\alpha = \frac{1}{2}\sigma^\alpha$
- ¹⁰⁹ – α spans the eight $SU(3)_C$ gauge fields, with generators expanded in Gell-
¹¹⁰ Mann matrices, $\tau^\alpha = \frac{1}{2}\lambda^\alpha$
- ¹¹¹ – L/R represent left and right projections of Dirac fermion fields. The Strong
¹¹² interaction is not chiral, so $Q = L+R$

113 – μ and ν are four-vector indices

114 – i, j, k are summed over the three generations of SM particles.

115 1.6 Higgs Mechanism

116 The SM Lagrangian without the addition of a Higgs field does not allow for
117 gauge boson and fermion mass terms: $\frac{1}{2}m_A^2 A_\mu A_\mu$ and $m(\bar{\psi}\psi)$, as these terms are
118 not gauge invariant. By introducing the Higgs field, mass terms for these particles
119 may be included in a gauge invariant way. This field is a complex doublet with a
120 potential $V(\Phi)$:

$$\Psi = \begin{pmatrix} \Phi^\dagger \\ \Phi^0 \end{pmatrix} \quad (1.24)$$

121 $V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda |\Phi^\dagger \Phi|^2 \quad (1.25)$

122 The minima of this field occurs for $|\Phi| = \sqrt{\frac{\mu^2}{2\lambda}} \equiv \frac{v}{2}$. This yields degenerate
123 minima, this symmetry is broken by choosing a specific minima (a.k.a. sponta-
124 neous symmetry breaking). By convention $\Phi_{min} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}$ is chosen. This means
125 the ground state of the Higgs field (Higgs vacuum) is non-zero, $\sqrt{\frac{-\mu^2}{\lambda}}$. The Higgs
126 Field may now be expanded around this new ground state:

$$\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \quad (1.26)$$

127 This non-zero Higgs vacuum now generates mass terms for the gauge bosons
128 from the following term in the Lagrangian:

$$|(-\frac{1}{2}g_1B_\mu - \frac{1}{2}g_2\sigma^aW_\mu^a)\Phi|^2 = \frac{1}{2}m_W^2W_\mu^+W^{-\mu} + \frac{1}{2}m_Z^2Z_\mu Z^\mu \quad (1.27)$$

129 where:

$$W_\mu^\pm \equiv \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2) \quad (1.28)$$

$$\begin{aligned} \text{130} \quad Z_\mu &\equiv \frac{1}{\sqrt{g_1^2 + g_2^2}}(g_2W_\mu^2 - g_1B_\mu) \end{aligned} \quad (1.29)$$

$$\begin{aligned} \text{131} \quad m_W &= \frac{vg_2}{\sqrt{2}} \end{aligned} \quad (1.30)$$

$$\begin{aligned} \text{132} \quad m_Z &= \frac{v}{\sqrt{2}}\sqrt{g_1^2 + g_2^2} \end{aligned} \quad (1.31)$$

133 The Higgs field also generates a mass term for the Higgs boson and self-
134 interactions for the Higgs boson.

135 1.7 Electroweak Theory

136 $SU(2)_L$ generates W^\pm, W^0 gauge bosons, which would be massless if $SU(2)_L$
137 was a perfect symmetry. These bosons are massive as this symmetry is broken.

138 The mass eigenstates, Z and γ given by:

$$\begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & -\sin\theta_W \\ \sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix} \quad (1.32)$$

139 Here θ_W is the Weinberg angle given by:

$$\cos\theta_W = \frac{g_2}{\sqrt{g_1^2 + g_2^2}} = \frac{m_W}{m_Z} \quad (1.33)$$

¹⁴⁰ 1.8 Quantum ChromoDynamics

¹⁴¹ As mentioned earlier the Strong Force, which binds the proton together, is
¹⁴² mediated by gluons. Quantum Chromodynamics is the QFT which describes the
¹⁴³ interactions of quarks and gluons via $SU(3)_C$ symmetry. QCD contains features
¹⁴⁴ not present in Electroweak Interactions due to $SU(3)_C$ generators not commuting
¹⁴⁵ (a.k.a. $SU(3)_C$ is a non-abelian group). For example, in QCD there is color
¹⁴⁶ confinement and asymptotic freedom due to the structure constants being non-
¹⁴⁷ zero. Requiring $SU(3)_C$ local gauge invariance implies:

$$\psi(x) \rightarrow \psi(x)' = \exp[i g_S \alpha(x) \cdot \hat{T}] \psi(x) \quad (1.34)$$

¹⁴⁸ where $\alpha(x)$ is the local phase function, g_S is the strong coupling constant, and
¹⁴⁹ \hat{T} are the eight generators of $SU(3)$ (note $\hat{T}^a = \frac{1}{2}\lambda^a a$, where λ^a are the Gell-Mann
¹⁵⁰ matrices). As the Gell-Mann matrices are 3x3, this means ψ has three degrees of
¹⁵¹ freedom under these $SU(3)$ rotations. So we represent ψ under $SU(3)$ rotations
¹⁵² as:

$$\psi = \begin{pmatrix} \psi_{red} \\ \psi_{green} \\ \psi_{blue} \end{pmatrix} \quad (1.35)$$

¹⁵³ Consequently, particle fields transforming under $SU(3)$ rotations have three
¹⁵⁴ components which physicists describe as color components (red, green, and blue).
¹⁵⁵ A particle's corresponding antiparticle has the corresponding anticolor. This color
¹⁵⁶ is the "charge" of QCD and is conserved under $SU(3)$ rotations. Combining colors,
¹⁵⁷ color neutral states (e.g. red and antired, or red, green and blue) may be created.
¹⁵⁸ For the Free Dirac Lagrangian to remain invariant under $SU(3)$ transformations,
¹⁵⁹ we must again postulate a boson field that modifies the derivative. The gluon

₁₆₀ field tensor is given by ($\alpha = 1, \dots, 8$):

$$G_{\mu\nu}^k = \partial^\mu G_\alpha^\nu - \partial^\nu G_\alpha^\mu - g_S f^{\alpha\beta\gamma} G_\beta^\mu G_\gamma^\nu \quad (1.36)$$

₁₆₁ Here $f^{\alpha\beta\gamma}$ are the structure constants of $SU(3)$. Combining all this gives the
₁₆₂ QCD Lagrangian:

$$\mathcal{L}_{QCD} = \bar{\psi}_q i\gamma^\mu (D_\mu)_{ij} \psi^{qj} - m \bar{\psi}^{qi} \psi_{qi} - \frac{1}{4} G_{\mu\nu}^\alpha G^{\alpha\mu\nu} \quad (1.37)$$

₁₆₃ Here i are the color indices, and q are the quark flavors. It is important to
₁₆₄ note that quarks transform under the fundamental representation of $SU(3)$, while
₁₆₅ gluons transform under the adjoint representation. This means quarks carry a
₁₆₆ single color charge (red, green, blue, antired, antigreen, antiblue) and gluons carry
₁₆₇ a color and anticolor charge.

₁₆₈ Figure 1.3 shows the three dominant QCD interactions. Since gluons carry
₁₆₉ color charge, they interact with one another. This does not occur in QED, as pho-
₁₇₀ tons do not have electric charge and therefore do not interact with each other.
₁₇₁ In QED, a bare electron's effective charge is largest closest to the electron and
₁₇₂ decreases as a function of distance. This is because the QED vacuum fills with
₁₇₃ particle antiparticle pairs spontaneously, which screen the charge of the bare elec-
₁₇₄ tron. The larger the distance from the electron, the smaller the effective charge
₁₇₅ and therefore the weaker the force. So for a pair of electrons, as the distance
₁₇₆ between them increases the repulsive force decreases and they may be observed
₁₇₇ separately.

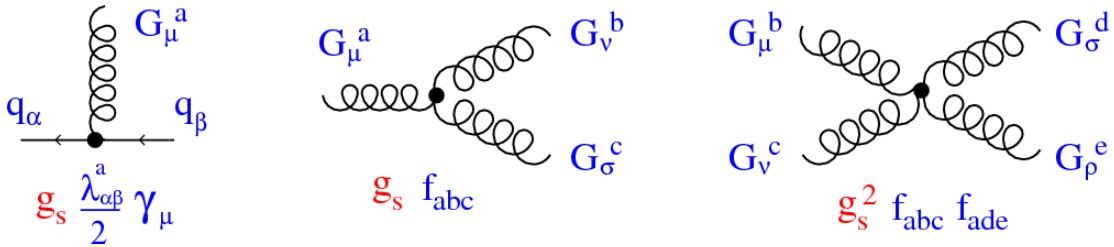


Figure 1.3: This figure shows the three dominant QCD interactions. From Ref. [14]

178 Bare quarks and gluons have not been observed. These particles have not
 179 been directly observed as they have not been found in isolation. Quarks and
 180 gluons group together to form color neutral objects, like baryons and mesons (qqq
 181 and $q\bar{q}$, respectively). This is because gluons have color charge and interact with
 182 each other. As the distance from a quark increases it's effective color charge
 183 increases due to the vacuum polarization in QCD. Color charge grows as the
 184 distance from the source increases (a.k.a. color is anti-screened in QCD). A quark's
 185 color charge increases with distance, so strong interactions become stronger at
 186 large distances (low momenta interactions). At small distances (large momenta
 187 interactions) strong interactions are significantly weaker and considered nearly
 188 free. This effect of referred to as asymptotic freedom. At large distances, a
 189 quark's effective charge is large and the strong force is more significant. This force
 190 becomes so strong that quarks form colorless bound states instead of remaining
 191 free particles. This effect is known as color confinement. This running of all SM
 192 fields is shown in Figure 1.4.

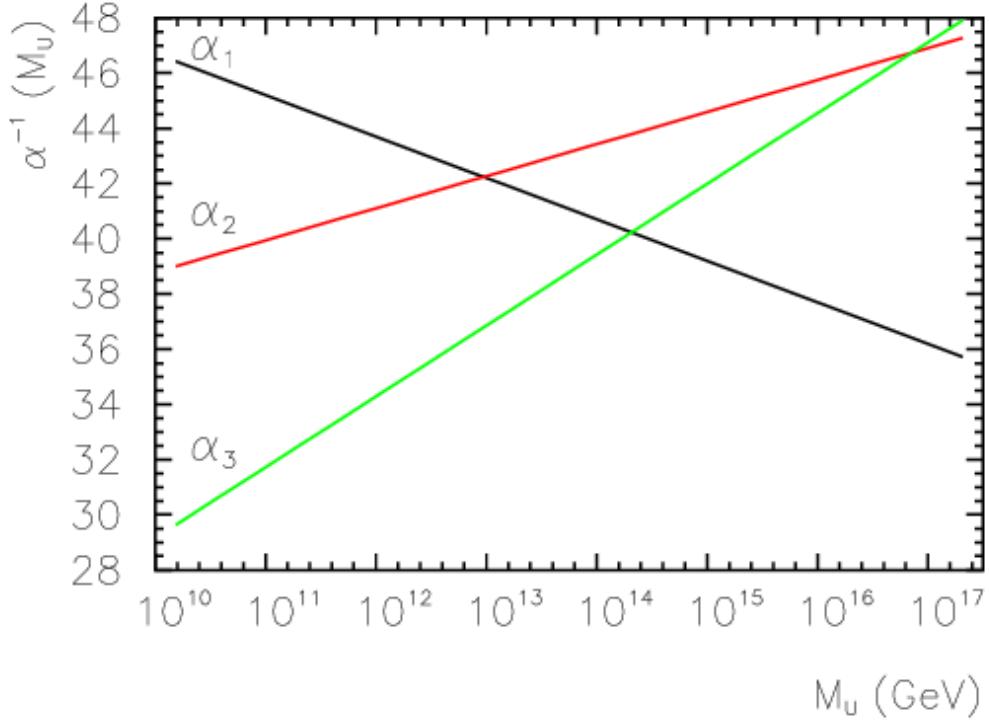


Figure 1.4: Strength of the U(1), SU(2), and SU(3) gauge couplings as a function of the energy scale of the interaction (Q). From Ref. [10]

Commonly the change in a particle's effective charge under a given force is quantified with $\beta(r) \equiv -\frac{de(r)}{d\ln r}$, where $e(r)$ is the effective charge of a given particle under a force. In QED this function is positive but in QCD this function is negative leading to confinement and asymptotic freedom. Moreover, one can calculate how the coupling (α) of a force varies with energies. (More deeply this amounts to incorporating renormalization and vacuum polarization in the boson propagators). For QCD this is:

$$\alpha_S(x) = \frac{\alpha_S(\mu^2)}{1 + \beta_0 \alpha_S(\mu^2) \ln(Q^2/\mu^2)} \quad (1.38)$$

200

$$\beta_0 = \frac{11N_c - 2n_f}{12\pi} \quad (1.39)$$

201 where Q is the momentum of the force is probed at, μ^2 is the renormalization scale.

203 As stated previously, quarks and gluons have not been observed in isolation.
 204 Instead they form bound colorless states. Hadronization is the process by which
 205 quarks and gluons form hadrons. The process of hadronization is still an active
 206 area of research. One qualitative description is shown in Figure BLAH. In this
 207 figure, as two quarks separate the color field between them is restricted to a tube
 208 with energy density of $1\text{GeV}/\text{fm}$. As they separate further, the energy in the color
 209 field increases, until there is enough energy to produce $q\bar{q}$ pairs, which breaks the
 210 color field. This process repeats until quarks and antiquarks have low enough
 211 energy to form colorless hadrons. The resulting spray of hadrons is called a jet.

212 Since quarks and gluons carry different color charges, their respective jets have
 213 different properties. As quarks carry only a single color charge (vs. gluons which
 214 have color and anticolor charge), so their jets have less constituent particles. More
 215 precisely, the Altarelli-Parisi splitting functions [3] contain a factor C_A for gluon
 216 radiation off a gluon and C_F for gluon radiation off a quark ($C_A/C_F = 9/4$). These
 217 color factors are the prefactor in the Feynman diagrams for these processes [1],
 218 which leads to gluon jets having more constituents and therefore more tracks than
 219 quark jets. Gluon jets also tend to have a larger radius with lower momentum
 220 constituents than quarks. There are many novel techniques to distinguish quarks
 221 from gluons. For this study the number of charged particles will be focused on.

²²² **Chapter 2**

²²³ **Standard Model Successes and
Limitations**

²²⁵ The Standard Model has consistently described much of reality to an extreme
²²⁶ degree of accuracy. It has predicted cross sections for strong and electroweak pro-
²²⁷ cesses that span over ten order of magnitude correctly [see Fig. 2.1] and contains
²²⁸ no known logical inconsistencies. Despite the strength and reality of the Stan-
²²⁹ dard Model, it still fails to describe aspects of reality and suffers from aesthetic
²³⁰ issues. To date, dark matter and energy comprise 95% of the universe, but are
²³¹ not accounted for in the SM. Additionally, neutrinos are known to have mass but
²³² are massless in the SM. There are mechanisms for introducing massive neutrinos
²³³ in the SM, but these mechanisms create hierarchy problems.

²³⁴ Possibly the most significant aesthetic issue is the hierarchy between the elec-
²³⁵ troweak and Planck scales. The electroweak scale is the scale of electroweak
²³⁶ symmetry breaking. The Planck scale is the scale where the gravitational force
²³⁷ is comparable in strength to the other forces. (This is also the scale where the
²³⁸ gravitational potential energy of two objects separated by a distance r is equal to
²³⁹ the energy of a photon with a wavelength r .) The Planck scale is where the SM

240 breaks down, as there is not an experimentally verified theory of quantum gravity,
241 and at this scale gravity cannot be ignored (like it can at the electro-weak scale).
242 These scales differ by ~ 30 orders of magnitude. Understanding this difference in
243 scales would not only explain the weakness of gravity at electroweak scales, but
244 also hopefully lead to a QFT for gravity. (NB: This hierarchy can also be framed
245 in terms of the corrections to the Higgs mass, which depend on the UV cutoff
246 scale - where the SM is suppose to break, which is taken at the Planck scale. This
247 leads the quantum corrections to the Higgs mass to force the Higgs mass to 10^{18}
248 TeV.)

249 These stark contrasts in scales may indicate that a more fundamental theory
250 exists. It is hoped that such a theory would explain and motivate some of the ad-
251 hoc features of the SM. In particular, there currently are no experimentally verified
252 explanations of why there are three generations of fermions, the values of the 19
253 SM parameters (6 quark masses, 3 charged lepton masses, 3 gauge couplings,
254 Higgs parameters (μ^2, λ)), the structure of the fermion representations, etc.

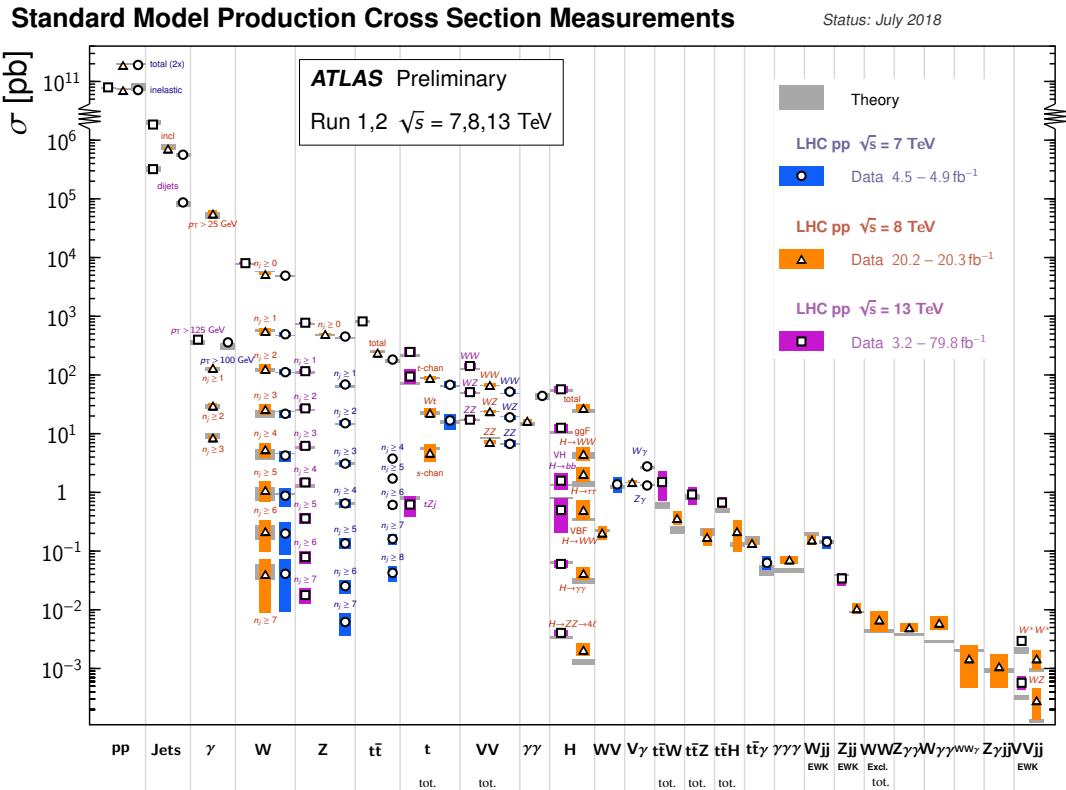


Figure 2.1: A comparison of cross section measurements at $\sqrt{s} = 7, 8, 13$ TeV from ATLAS compared to theoretical measurements. From Ref. [5]

255 **Chapter 3**

256 **New Physics Models with
257 Diboson Resonances**

258 **3.1 Randall Sundrum Bulk Model**

259 The electroweak-planck hierarchy may be explained by the existence of extra
260 dimensions, like the 5D Randall Sundrum Bulk Model ([15], [2]). In this model,
261 there is one extra warped spatial dimension, y , with a metric:

$$ds^2 = e^{-2k|y|} \eta_{\mu\nu} dx^\mu dx^\nu + dy^2 \quad (3.1)$$

262 where $e^{-k|y|}$ is the warp factor of the extra dimension, which is compactified on
263 a S^1/Z_2 orbifold (a.k.a. a circle where $y \rightarrow -y$). This can be visualized as every
264 point in space time having a line extending from it a distance L , representing
265 this fifth dimension. At the end of this line is the Planck brane. This fourth
266 spatial dimension separates two 4-D branes: Planck brane and TeV brane. We
267 live on the TeV brane, as shown in Figure 3.1. The Higgs field (and to a lesser
268 degree the top quark and graviton fields) is localized near the TeV Brane, while

269 the light fermion fields are localized more near the Planck brane. Fundamental
 270 parameters are set on the Planck brane. The warp factor may be scaled away from
 271 all dimensionless SM terms by field redefinitions. However, the only dimensionful
 272 parameter, $m_H^2 = v^2$ is rescaled by $\tilde{v} \sim e^{-kL} M_{Pl} \sim 1\text{TeV}$ for $kL \sim 35$, explaining
 273 why gravity is so weak on the TeV brane. Also, by localizing the light fermion
 274 fields near the Planck brane and top and graviton fields near the TeV brane, the
 275 light quarks will have smaller masses.

276 The two free parameters of this theory are M_{Pl} and k . Based on this RS Bulk
 277 model, all SM particles should have Kaluza-Klein (KK) excitations. In particular,
 278 the graviton would have KK excitations that prefer to decay to WW or ZZ, which
 279 is why this analysis searches for RS Gravitons.

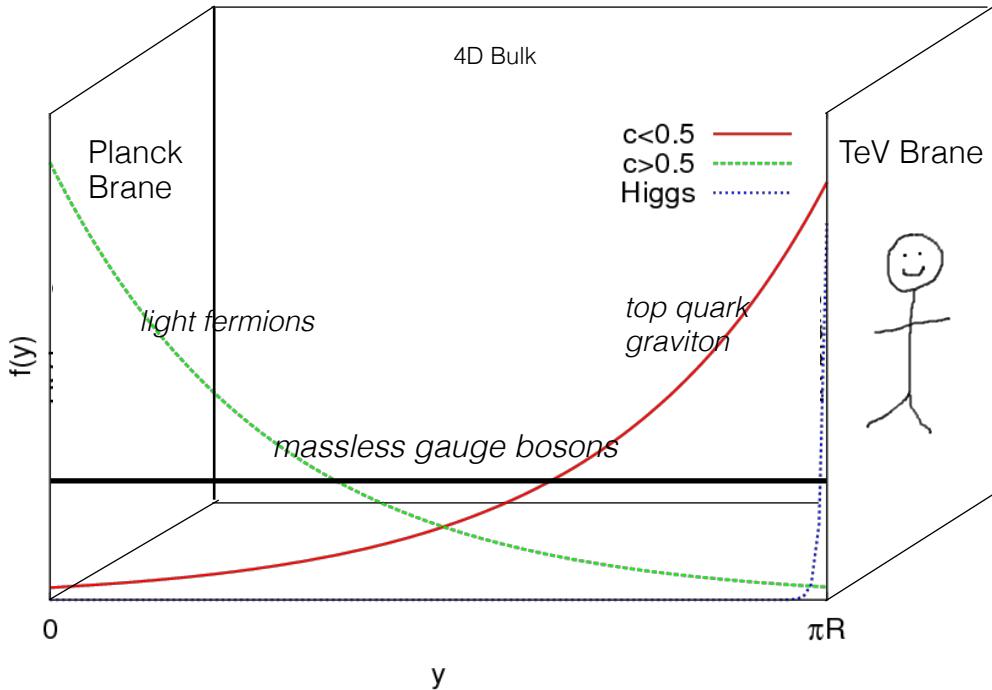


Figure 3.1: Cartoon of RS Bulk Model

280 3.2 Extended Scalar Sector

281 A further striking asymmetry of the SM is the simplicity of the scalar sector in
282 comparison to the boson and fermion sectors. To date, the scalar sector has only
283 one member, the Higgs boson. Therefore, it is natural to posit an extension to the
284 scalar sector. From a theoretical standpoint this could also help generate baryon
285 asymmetry through additional sources of CP violation. This analysis searches for
286 a simple extension to the scalar sector as proposed in Ref. [16]. The extended
287 scalar sector includes a real Higgs singlet (S) and complex $SU(2)_L$ doublet (Φ)
288 (the SM Higgs), where mass eigenstates are mixtures of the fields. S has a vev of
289 v and Φ has a vev of x . This then gives a Lagrangian of:

$$\mathcal{L} \supset (D^\mu \Phi)^\dagger D_\mu \Phi + \partial^\mu S \partial_\mu S - m^2 \Phi^\dagger \Phi - \mu^2 S^2 + \lambda_1 (\Phi^\dagger \Phi)^2 + \lambda_2 S^4 + \lambda_3 \Phi^\dagger \Phi S^2 \quad (3.2)$$

290 The mass eigenstates of the scalar sector are then mixtures of S and Φ and
291 the free parameters of the theory are m_H , $\sin\alpha$, and $\tan\beta = v/x$. The fields are
292 then given by:

$$\Phi \equiv \begin{pmatrix} 0 \\ \frac{\tilde{h}+v}{\sqrt{2}} \end{pmatrix} \quad (3.3)$$

$$S \equiv \frac{h' + x}{\sqrt{2}} \quad (3.4)$$

294 Diagonalizing the mass matrix leads to the mass eigenstates h (discovered
295 Higgs boson) and H (the physical particles):

$$\begin{pmatrix} h \\ H \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \quad (3.5)$$

²⁹⁶ This suppressed h and H production and SM H couplings:

$$BR_{H \rightarrow SM} = \sin^2 \alpha \times \frac{\Gamma_{SM, H \rightarrow SM}}{\Gamma_{tot}} \quad (3.6)$$

²⁹⁷ Moreover, in the case that $m_H > m_h$, $H \rightarrow hh$ is possible. This further suppresses
²⁹⁸ $H \rightarrow VV/ff$. This search is most sensitive to $H \rightarrow WW$.

²⁹⁹ 3.3 Simple Standard Model Extensions

³⁰⁰ The RS Bulk model is motivated by resolving the hierarchy problem. Ex-
³⁰¹ tending the Scalar sector is a natural space to look for new physics due to the
³⁰² complexity of fermion and boson groups. There are many other interesting and
³⁰³ well motivated frameworks, but there is a lack of completely predictive models,
³⁰⁴ due to model flexibility (free parameters). Therefore it is hard for experimentalists
³⁰⁵ to know which theories to search for in data. However, as seen in 3[13], a "Sim-
³⁰⁶ plified Model" approach may be taken. In the search for reasonably narrow width
³⁰⁷ particles, as in this search, the search is not sensitive to all the details and free
³⁰⁸ parameters of the theory. Generally such searches are only sensitive to the reso-
³⁰⁹ nance mass and its interactions. Therefore, a theory's Lagrangian may be reduced
³¹⁰ to only retain this information (mass parameters and couplings). Experimental
³¹¹ results using this framework may then be reinterpreted in a given theory.

³¹² In the simplified approach, a real vector field is represented under $SU(2)_L$
³¹³ with vanishing hypercharge. This results in two charged and one neutral bosons.
³¹⁴ Defined as:

$$V^\pm = \frac{V_\mu^1 \mp iV_\mu^2}{\sqrt{2}} \quad (3.7)$$

³¹⁵

$$V_\mu^0 = V_\mu^3 \quad (3.8)$$

³¹⁶ The Lagrangian is then:

$$\mathcal{L} \supset -\frac{1}{4}D_{[\mu}V_{\nu]}^a D^{[\mu}V^{\nu]}_a + \frac{m_V^2}{2}V_\mu^a V^{a\mu} + ig_V c_H V_\mu^a H^\dagger \tau^a \overset{\leftrightarrow}{D}^\mu H + \frac{g^2}{g_V} c_F V_\mu^a J_F^{\mu a} \quad (3.9)$$

³¹⁷ In order the terms represent: the kinetic, V mass, Higgs- V interaction, and
³¹⁸ V -left-handed fermion interaction terms. Phenomenologically the three physical
³¹⁹ particles this predicts are degenerate, where V couples most strongly to VV , via
³²⁰ the g_V coupling factor. The dominant production modes are DY and VBF.

³²¹ Two versions of HVT are considered, Model A and B. Model A is a weakly
³²² coupled model where $g_V \ll 1$, like the extended gauge symmetry discussed in Ref .
³²³ [16]. Model B is a strongly coupled model, where $1 < g_V < 4\pi$. The width of the
³²⁴ resonance grows with g_V so for this narrow resonance search only g_V is chosen to
³²⁵ be less than 6 (so $\Gamma/M < 10\%$). More precisely, the coupling of these resonances
³²⁶ to fermions scales as $g_f = g^2 c_F/g_V$, where g is the SM $SU(2)_L$ gauge coupling
³²⁷ and c_F is the free parameter (expected to be of order 1 for Model A and B). This
³²⁸ then means that for Model B the coupling is to fermions is more suppressed than
³²⁹ for Model A, leading to a smaller DY production rate and BR to fermionic final
³³⁰ states. The coupling of V to SM bosons scales as $g_H = g_V c_H$, where c_H is a
³³¹ free parameter on the order of 1 for Model A and B. So for small values of g_V
³³² (i.e. Model A - weakly coupled theories) the BR to gauge boson is smaller than
³³³ for Model B. So weakly coupled vectors have large production cross sections and
³³⁴ decay predominantly to leptons or jets, while strongly coupled vectors are produced
³³⁵ less and decay predominantly to gauge bosons.

³³⁶ Vectors in Model A and B are generally produced via quark-anti-quark annihi-
³³⁷ lation. The more rare production via vector-boson-fusion is considered by setting
³³⁸ $g_H = 1$ and $g_F = 0$. In Model B diboson final states are enhanced as stated
³³⁹ previously due to g_H and moreover the BR to WZ, WH, WW, and ZH are the

³⁴⁰ same.

³⁴¹ In summary, V couples most strongly to left-handed fermions and VV dependent on g_V .

Part II

343

Experimental Setup

344

³⁴⁵ **Chapter 4**

³⁴⁶ **LHC**

³⁴⁷ The Large Hadron Collider (LHC) is the highest-energy particle collider in the
³⁴⁸ world. It was designed to expand the frontier of high energy particle collisions in
³⁴⁹ energy and luminosity. This enables LHC experiments to test the Standard Model
³⁵⁰ and search for new physics at higher energies than tested with previous colliders.
³⁵¹ Collisions at higher energies not only produce more massive particles but also
³⁵² more weakly interacting particles. Fig. 4.1 shows production cross sections for
³⁵³ various processes at hadron colliders. The rate for electroweak physics pcoesses
³⁵⁴ including W and Z scale with the center-of-momentum energy, \sqrt{s} .

³⁵⁵ The LHC consists of a 26.7 km (17 miles) ring, approximately 100 m un-
³⁵⁶ derground, outside Geneva, Switzerland. Counter-circulating proton (and occa-
³⁵⁷ sionally heavy ions) beams collide inside four experiments along the beam line:
³⁵⁸ ATLAS, CMS, LHCb, ALICE. ATLAS and CMS are general purpose detectors
³⁵⁹ designed to explore high energy frontier. LHCb is designed to study the physics
³⁶⁰ of b -quarks. ALICE specializes in studying heavy ion collisions.

proton - (anti)proton cross sections

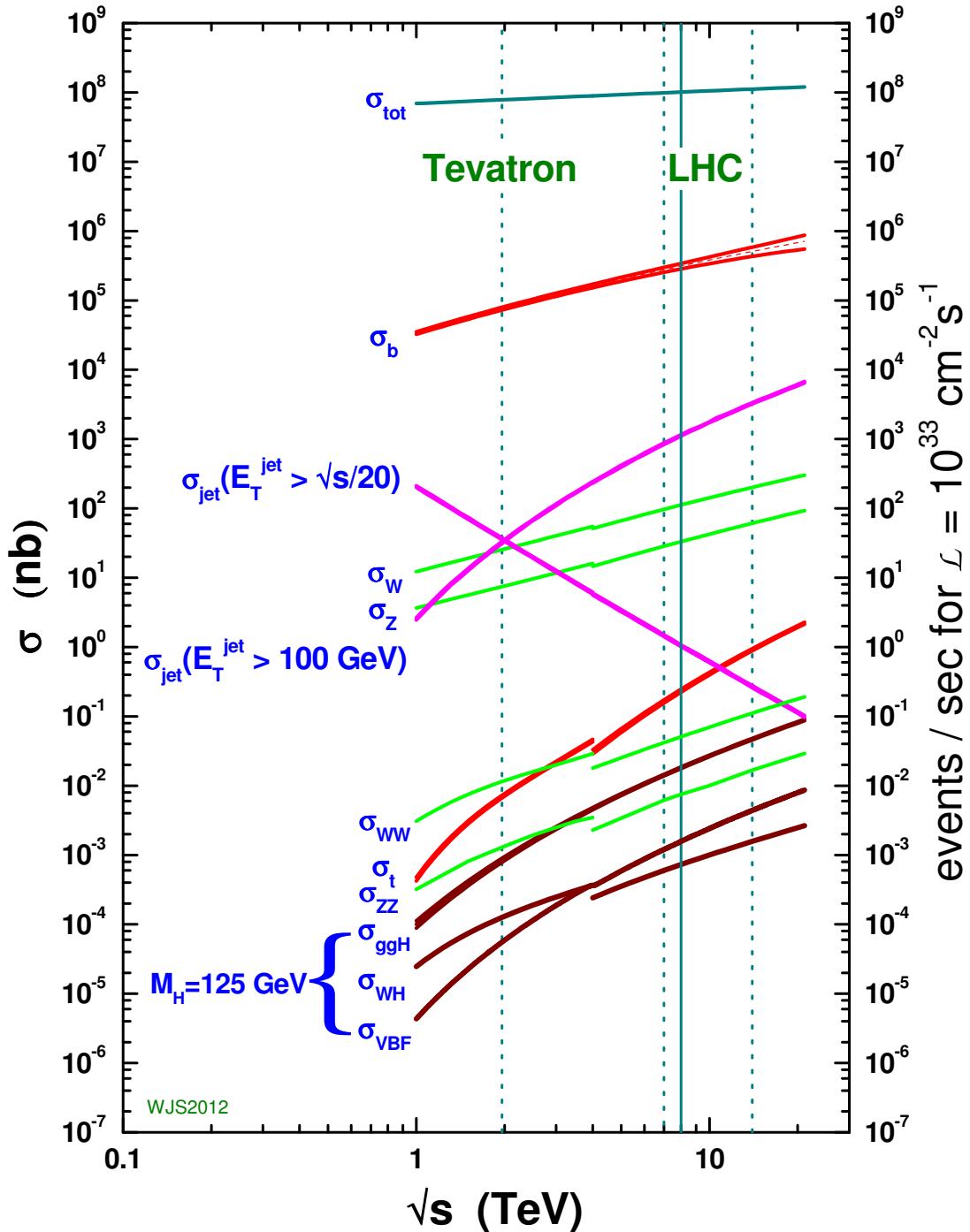


Figure 4.1: Scaling of cross sections with \sqrt{s} . Natasha: write more here

361 The first proton beams circulated in September, 2008. Nine days later an
362 electrical fault lead to mechanical damage and liquid helium leaks in the collider.
363 This incident delayed further operation until November 2009, when the LHC be-
364 came the world's highest energy particle collider, at 1.18TeV per beam. This first
365 operational run continued until 2013, reaching 7 and 8 TeV collision energies. Dur-
366 ing this run a particle who's properties were consistent with the Standard Model
367 Higgs boson was discovered. The next operational began after a two year shut-
368 down after upgrades to the LHC and experiments. This run lasted from 2013 to
369 2018 reaching 13 TeV collision energies. This analysis uses data from the second
370 operational run.

371 **4.1 LHC Layout and Design**

372 The layout of the LHC is shown in Figure 4.2. The red and blue lines in the
373 figure represent the counter-circulating proton beams. The LHC is divided into
374 eight octants. Octant 4 contains the RF cavities that accelerate the protons and
375 octant 6 contains the beam dump system. Octants 3 and 7 house the collimation
376 systems for beam cleaning. The beams collide inside the four aforementioned
377 experiments. Each octant contains a curved and straight section. The LHC
378 magnets are built with NbTi superconductors cooled with super-fluid Helium to
379 2K, creating a 8.3T magnetic field to bend the proton beams.

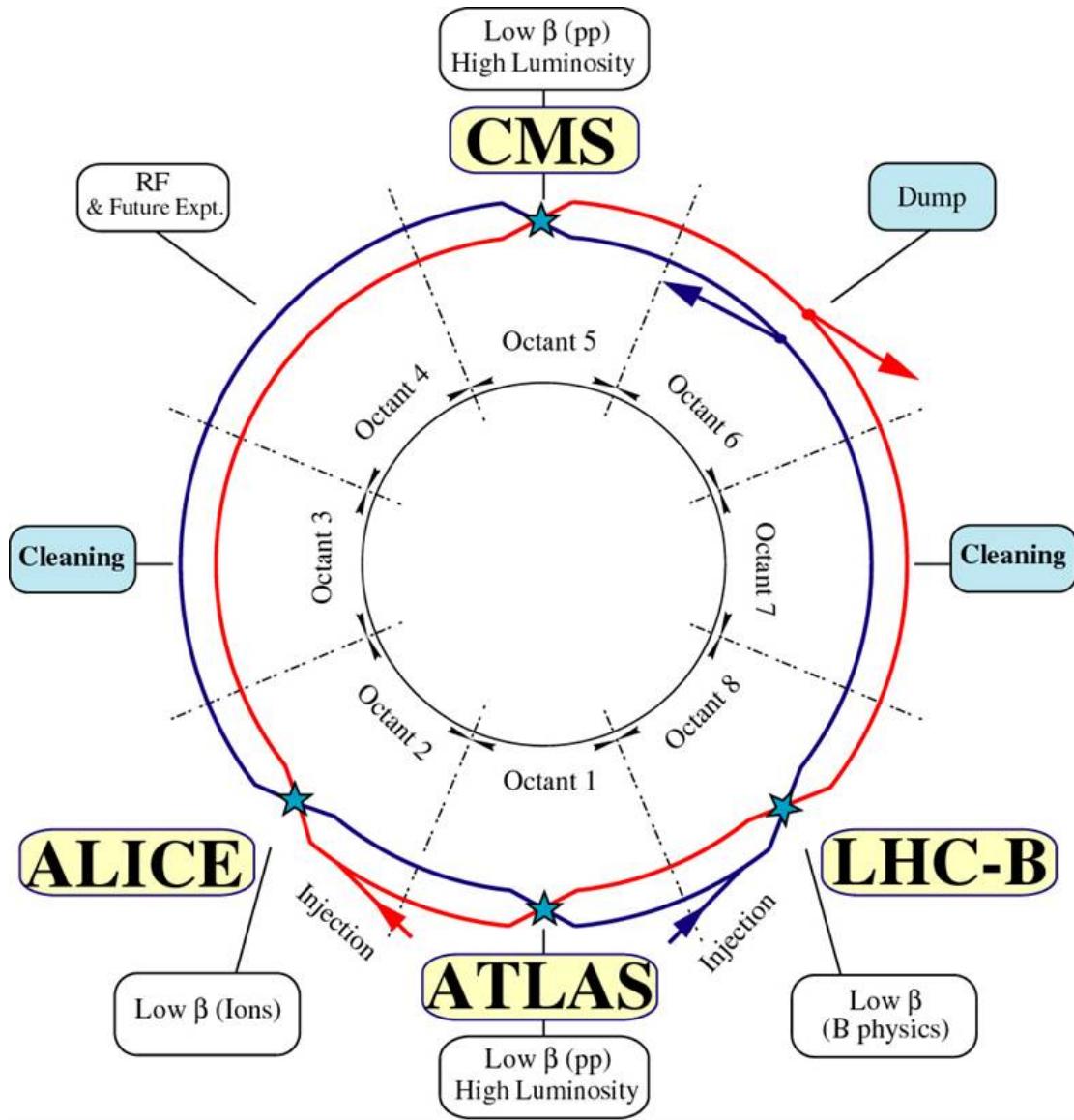
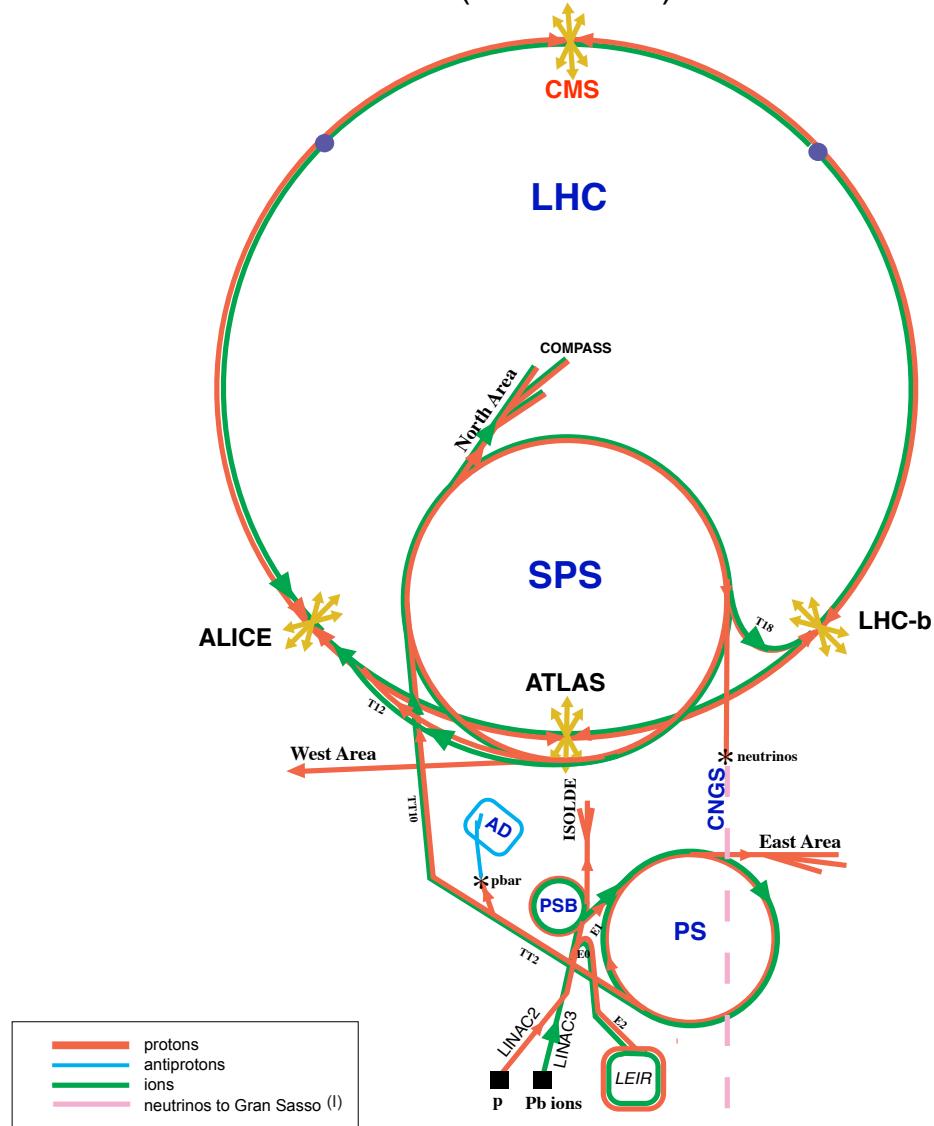


Figure 4.2: LHC Layout. Natasha write more

380 Four sequential particle accelerators are used to accelerate proton from rest as
 381 shown in Figure 4.3. First, Hydrogen gas is ionized to produce protons which
 382 are then accelerated to 50 MeV using Linac 2, a linear accelerator. The result-
 383 ing proton beam is then passed to three circular particle accelerators: Proton
 384 Synchrotron Booster, Proton Synchrotron, and Super Proton Synchrotron (SPS),

385 accelerating protons to 1.4, 25, and 450 GeV, respectively. Once the protons exit
386 SPS, they are injected into the LHC at octant 2 and 8. Each proton bunch contains
387 10^{11} protons. The spacing between bunches is 25 ns, which means each beam con-
388 tains 3564 bunches. However, some bunches are left empty due to injection and
389 safety requirements, yield 2808 bunches per beam. Once the proton beams are
390 injected they are accelerated to 13 TeV.

CERN Accelerators (not to scale)



LHC: Large Hadron Collider

SPS: Super Proton Synchrotron

AD: Antiproton Decelerator

ISOLDE: Isotope Separator OnLine DEvice

PSB: Proton Synchrotron Booster

PS: Proton Synchrotron

LINAC: LINear ACcelerator

LEIR: Low Energy Ion Ring

CNGS: Cern Neutrinos to Gran Sasso

Rudolf LEY, PS Division, CERN, 02.09.96
Revised and adapted by Antonella Del Rosso, ETT Div.,
in collaboration with B. Desforges, SL Div., and
D. Manglunki, PS Div. CERN, 23.05.01

Figure 4.3: LHC Accelerator. Natasha write more

391 As many new physics models predict cross-sections below the weak scale it was
392 important to design the LHC to be capable of collecting enough data, by running
393 in high luminosity conditions. The machine luminosity depends only on beam
394 parameters:

$$L = \frac{N_p^2 f}{4\epsilon\beta^*} F \quad (4.1)$$

395 where N_p is the number of protons per bunch, f is the bunch crossing frequency,
396 ϵ is the transverse beam emittance, β^* is the amplitude function at the collision
397 point, and F is the geometric luminosity reduction factor due to the beams crossing
398 at an angle (rather than head-on).

399 This analysis uses data from Run 2, totally 139/fb. The peak luminosity was
400 [Natasha add info here]. [Natasha add lumi and integrated lumi figures].

401 **Chapter 5**

402 **The ATLAS Detector**

403 The ATLAS detector measures the position, momentum and energy of parti-
404 cles produced in the proton collisions by using magnetic fields, silicon detectors,
405 sampling calorimeters, and gaseous wire detectors. It is located approximately
406 100 m underground at Point-1 around the LHC beam line and weighs 7000 metric
407 tons. The detector is 46 m long, 25 m high, 25 m wide as shown in Figure 5.2.
408 The detector can be divided into three subsystems: the Inner Detector (ID), the
409 Calorimeters, and the Muon Spectrometer (MS). Figure 5.3 shows an overview of
410 how different particles interact in the detector.

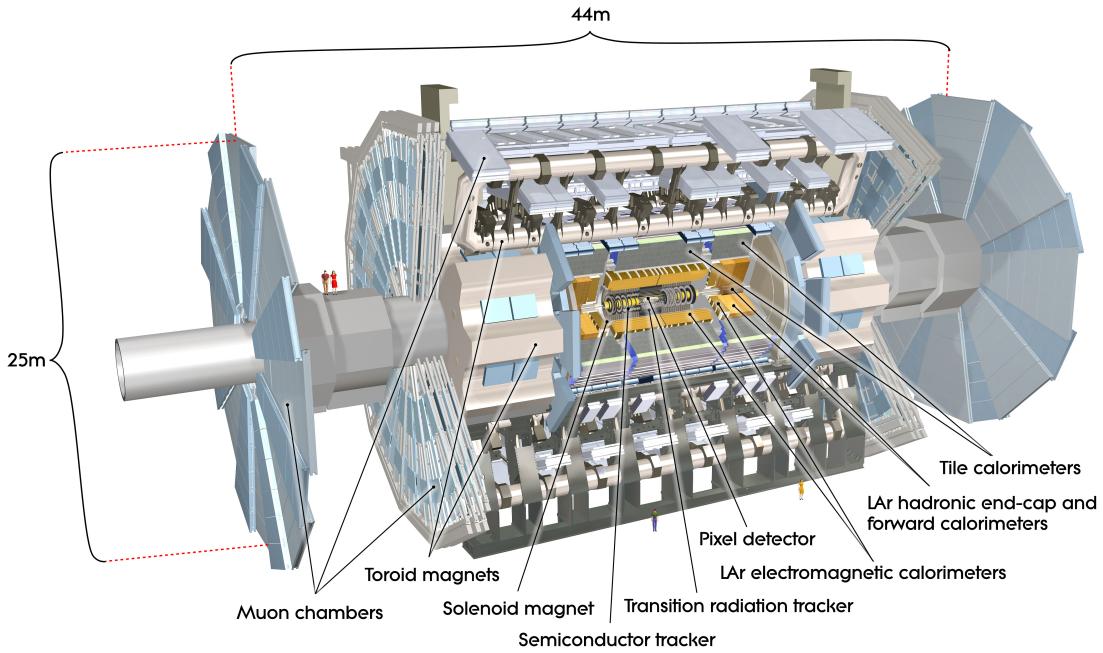


Figure 5.1: Big picture layout of ATLAS detector. Natasha: write more

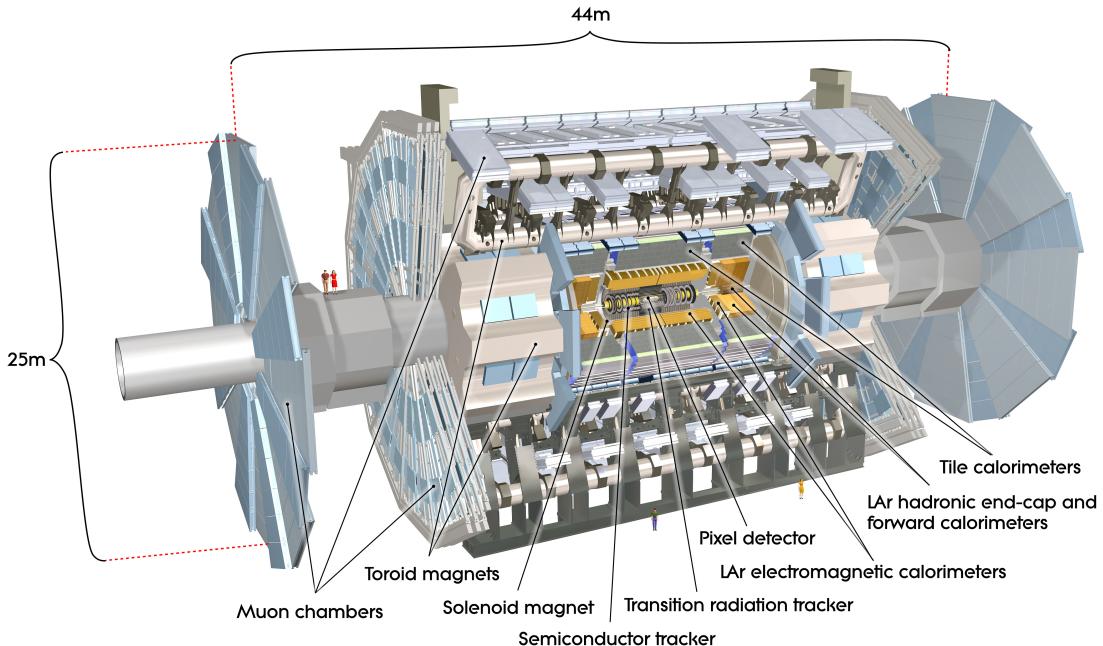


Figure 5.2: Big picture layout of ATLAS detector. Natasha: write more

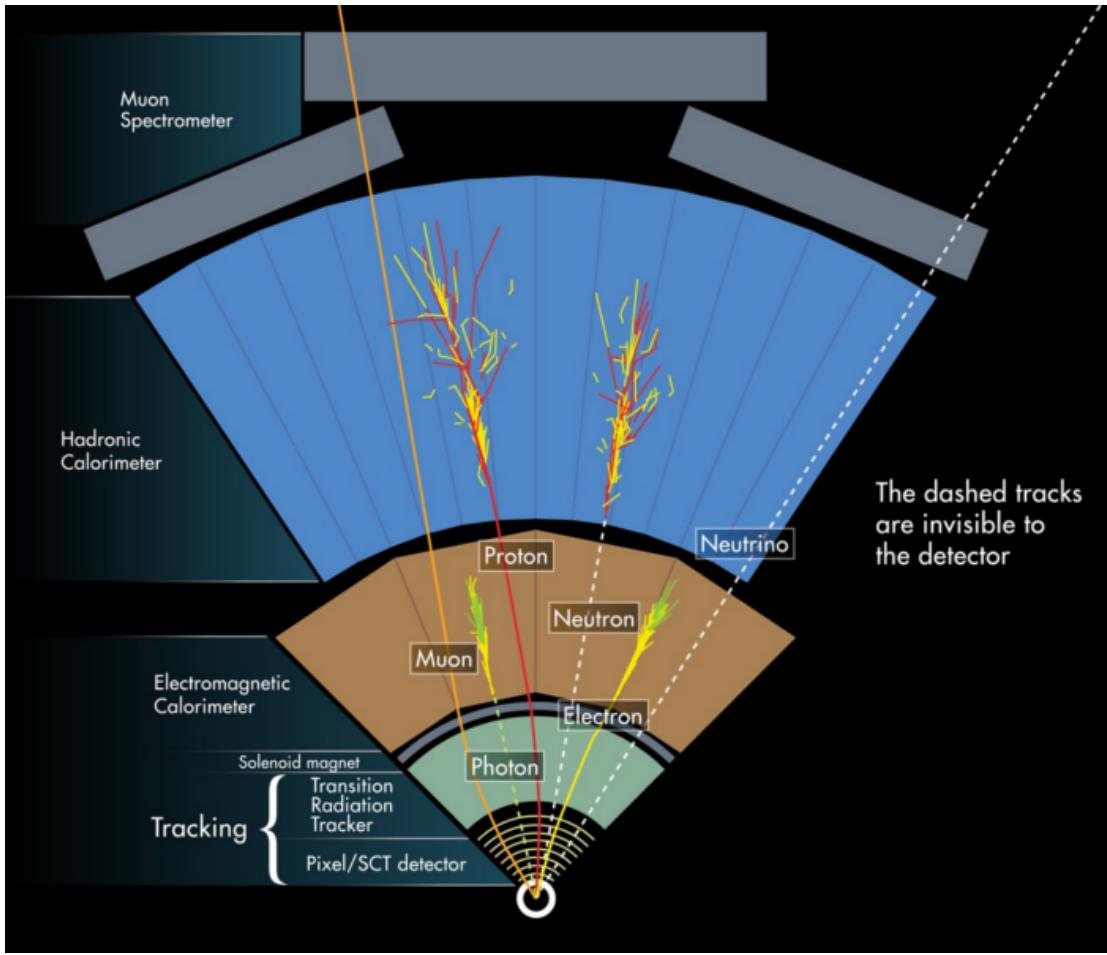


Figure 5.3: A simplified schematic of how different particles interact and are detected within ATLAS.

411 5.1 Coordinate System

412 The trajectory of particles within ATLAS is measured relative to the nominal
 413 interaction point. The z -axis points along the beam line, such that when the
 414 LHC is viewed from above, the counter-clockwise circulating beam points along
 415 the positive- z direction. The $x - y$ plane is transverse to the beam line, with
 416 the positive x -axis pointing towards the center of the LHC ring. The positive
 417 y -axis points vertically upward. The azimuthal angle, ϕ , is the angular distance

about the z -axis, with $\phi = 0$ along the x -axis. The polar angle from the z -axis is denoted as θ . The polar angle is usually replaced by pseudo-rapidity, $\eta = -\ln \tan(\frac{\theta}{2})$. Pseudo-rapidity is preferred as $\Delta\eta$ is invariant under boosts along z and particle production is approximately invariant under η . Angular separation between particles in ATLAS are given by $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$. The distance from the beamline is given by $r = \sqrt{x^2 + y^2}$

5.2 Inner Detector

The Inner Detector (ID) was designed to identify and reconstruct vertices, distinguish pions from electrons, and measure the momentum of charged particles. The ID uses three different technologies for particle reconstruction: the Pixel Detector, Semiconductor Tracker (SCT), and the Transition Radiation Tracker (TRT), shown in Figure 5.4 and 5.5. The entire ID is immersed in a 2T solenoidal magnetic field parallel to z , causing charged particles to bend in the transverse-plane, allowing particle momentum measurements.

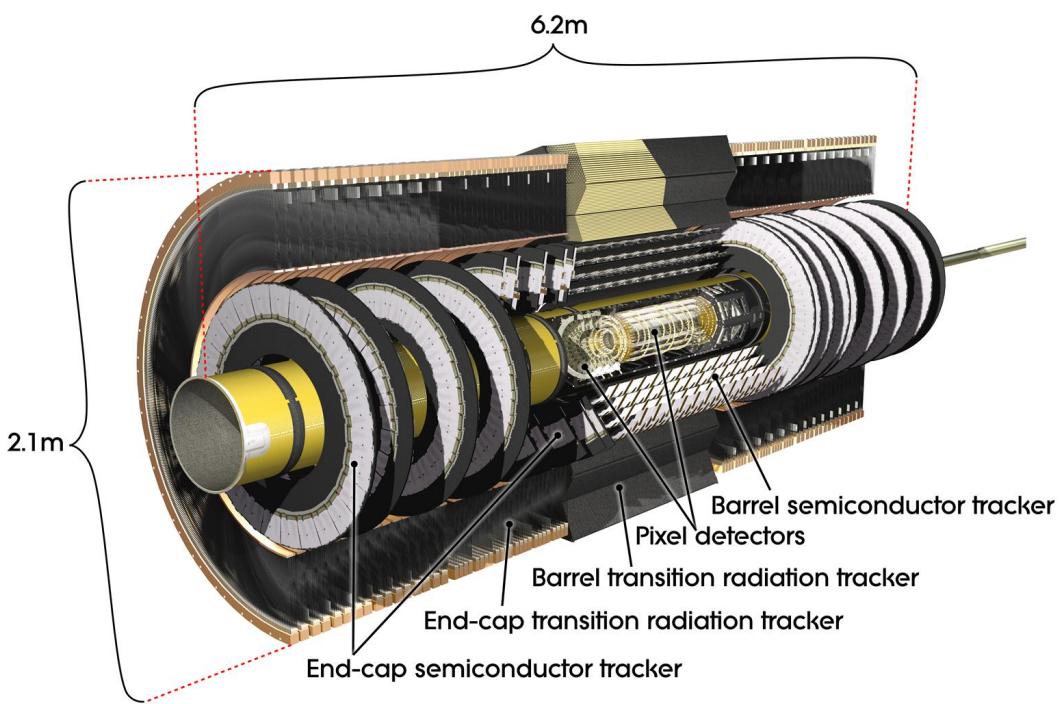


Figure 5.4: Layout of ATLAS Inner Detector

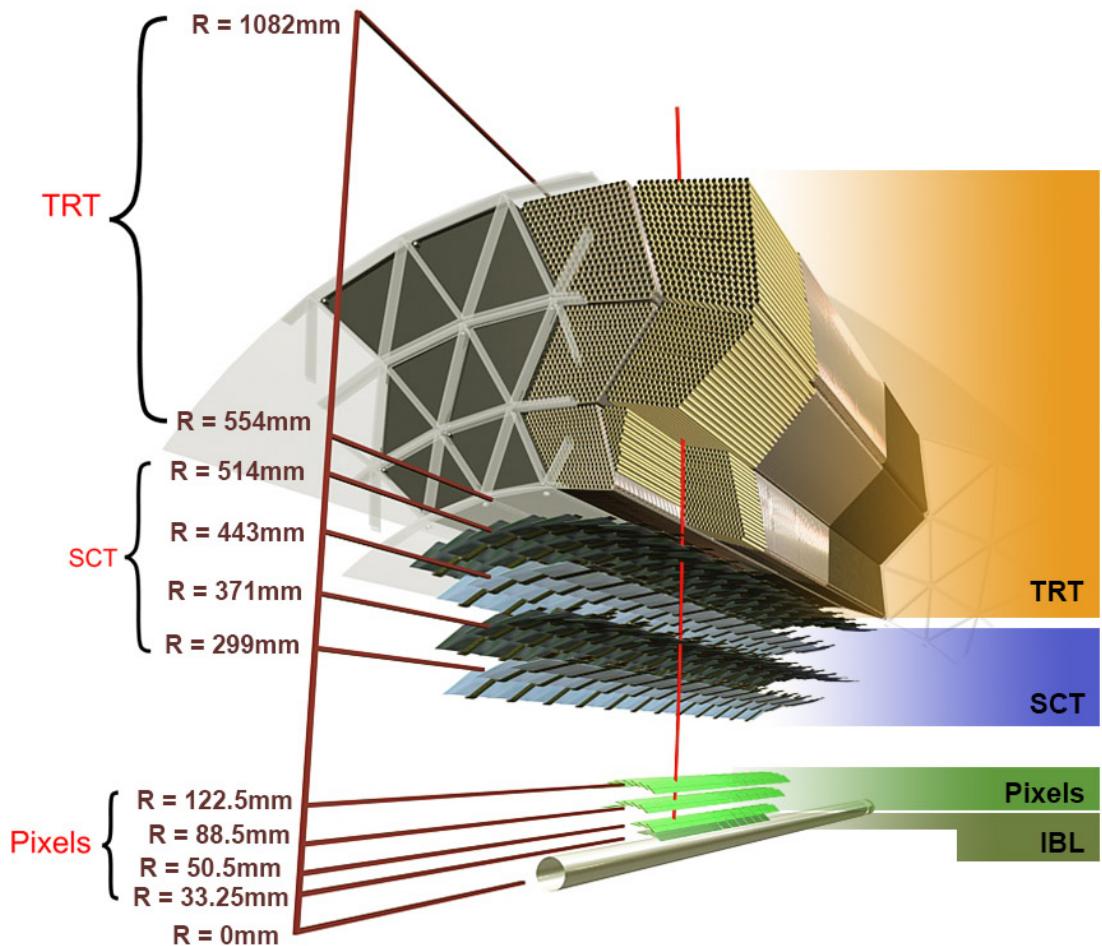


Figure 5.5: Layout of ATLAS ID Barrel System.

432 **5.2.1 Pixel Detector**

433 The pixel detector consists of four barrel layers between $r = 32.7$ and 122.5
434 mm, extending to $|z| = 400.5$ mm. The innermost pixel barrel, the Insertable
435 b-Layer (IBL), only extends to $|z| = 332$ mm. The pixel detectors closer to the
436 beam line (larger η values) consists of six parallel cylindrical rings of pixel de-
437 tectors transverse to the beam line. The entire pixel detector consists of 1744
438 identical pixel sensors each with 46080 readout channels, totaling about 80 mil-
439 lion individual pixels. Most of the pixel sensors are $50 \times 400 \mu\text{m}^2$. The intrinsic
440 measurement accuracy of a individual pixel detector in the barrel (end-cap) in
441 $r - \phi$ is $10 \mu\text{m}$ and $115 \mu\text{m}$ in $z(r)$.

442 **5.2.2 Semiconductor Tracker**

443 The SCT is located outside the pixel detector and has the same barrel and
444 endcap geometry as the pixel detector. SCT sensors are $80 \mu\text{m} \times 12$ cm ($80 \mu\text{m}$
445 strip pitch). In the barrel the strips are parallel to the z -axis and are segmented
446 in ϕ . In the endcaps, the strips extend radially. Sensors are grouped in modules
447 containing two layers of strips rotated 40 mrad with respect to each other. This
448 offset allows for the two-dimensional position of a track to be determined by
449 identifying the crossing point of the strips that registered a hit. SCT modules
450 measure tracks with an accuracy of $17 \mu\text{m}$ in $r - \phi$ and $580 \mu\text{m}$ in $z(r)$ in the
451 barrel (end-cap) region.

452 **5.2.3 Transition Radiation Tracker**

453 The transition radiation tracker (TRT), enveloping the SCT, is a gaseous
454 straw-tube tracker mainly used for electron/pion track separation. Each straw
455 is 4 mm in diameter and filled with a Xe- CO_2 - O_2 gas mixture. An anode wire at

456 the center of the straw is held at ground potential, while the walls of the straw are
457 kept at -1.4kV. When a charged particle passing through the TRT ionizing the
458 gaseous mixture, and the resulting ions form an avalanche on the anode wire with
459 a grain of 10^4 . The resulting signal from the anode wire is then digitized and
460 amplified. Signals passing a low threshold cutoff are used to distinguish noise from
461 tracks. Signals passing a high threshold cutoff are sensitive to transition radiation
462 (TR). TR photons are emitted when charged particles pass between materials
463 with different dielectric constants. The probability that a charged particle with
464 energy E and mass m passing between two materials emits a TR photon in the keV
465 range is proportional to $\gamma = E/m$. In the TRT straws these often then convert via
466 the photoelectric effect, causing a large avalanche triggering the high-threshold.
467 Since electrons have a smaller mass than pions, tracks from electrons are more
468 likely to trigger the high threshold. Consequently, the high threshold TRT trigger
469 provides electron identification information that is uncorrelated with calorimeter
470 shower-shape information.

471 The barrel region of the TRT extends from $r = 563\text{-}1066$ mm and $|z| < 712$
472 mm. Barrel Straws are 144 cm long (divided $\eta \approx 0$) and orientated parallel to
473 the beam direction. End-cap straws extend radially and are 37 cm long. There
474 are 53,544 straws in the barrel and 160,000 straws in the end-caps. Radiator mats
475 of polypropylene/polyethylene fibers in the barrel are aligned perpendicular to the
476 barrel straws (with holes for the straws to pass through). In the end-cap region,
477 radiator foils are layered between the radial TRT straws.

478 The width of the signal pulse is sensitive to the distance between the charged
479 particle track and the anode wire and allows for a hit resolution of $130\mu\text{m}$. The
480 TRT extends to $|\eta| = 2.0$ and provides about 36 hits per track.

481 **5.3 Calorimeters**

482 The ATLAS electromagnetic and hadronic calorimeters (EMC and HCAL,
483 respectively) absorb and measure the energy of high energy hadrons, photons,
484 and electrons with $|\eta| < 4.9$. Both systems use sampling calorimeters which
485 consist of alternating layers of dense absorbing and active layers. In the absorbing
486 layer particles interact and lose energy, creating showers. These showers are then
487 detected and measured in the active layer. The amount of charge measured in the
488 active material scales with the energy of the incident particle, and thus provides a
489 measurement of the particle's energy. An overview of the layout of the calorimeter
490 system is shown in Figure 5.6.

491 The EMC measures and contains the energy of electromagnetically interacting
492 particles with 170k channels. It consists of layered accordion-shaped Lead ab-
493 sorber plates and electrodes immersed in liquid Argon. Using accordion-shaped
494 electrode and absorbers ensures ϕ symmetry and coverage. The EMC is com-
495 posed of a barrel part ($|\eta| < 1.475$), two end-caps ($1.375 < |\eta| < 3.2$), and a
496 presampler ($|\eta| < 1.8$). The presampler, containing only liquid Argon, corrects
497 for upstream energy losses of electrons and photons. The EMC barrel is segmented
498 into three layers. The first layer has finest segmentation with readout cells ex-
499 tending $\Delta\eta \times \Delta\phi = 0.025/8 \times 0.1$. This provides a precise shower measurements
500 used to separate prompt photons from $\pi^0 \rightarrow \gamma\gamma$ decays. The second layer has
501 coarser segmentation and is approximately 16 radiation lengths long. A radiation
502 length is the average distance an electron travels before losing all but $1/e$ of its
503 energy to bremsstrahlung. The last layer is the most coarse and measures the tail
504 of the electromagnetic shower. A schematic of the ECAL is shown in Figure ??.

505 The hadronic calorimeter located outside the EMC and is used to contain
506 and measure the energy of hadronically interacting particles. It consists of a tile

507 calorimeter (TileCal), hadronic end-cap calorimeter (HEC), and liquid Argon for-
 508 ward calorimeter (FCAL). TileCal is located behind the LAr EMC and uses steel
 509 absorbers and liquid Argon as the active material. TileCal consists of three barrel
 510 layers in the central and forward regions, extending up to $|\eta| < 1.7$. Radiated
 511 photons from the steel tiles are collected via wavelength-shifting fibers connected
 512 to photomultiplier tubes, as shown in Figure 5.8. The HEC lies behind the EMC
 513 endcap wheels. It uses copper absorbers and liquid Argon as the active material
 514 and covers $1.5 < |\eta| < 3.2$. Finally, the FCAL covers $3.1 < |\eta| < 4.9$ and consists
 515 of three modules all using liquid Argon as the active material. The first module
 516 uses copper absorber and was designed for electromagnetic measurements. The
 517 second and third modules consist of tungsten absorber and are used to measure
 518 the kinematics of hadronically interacting particles. A schematic of the HCAL is
 519 shown in Figure 5.8.

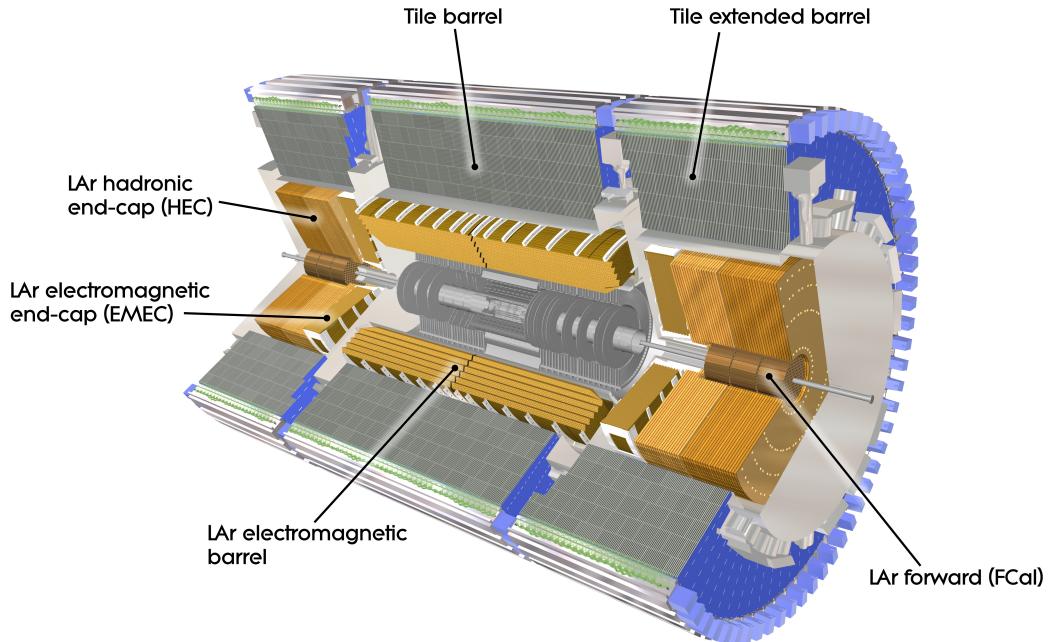


Figure 5.6: Overview of ATLAS electromagnetic and hadronic calorimeters.

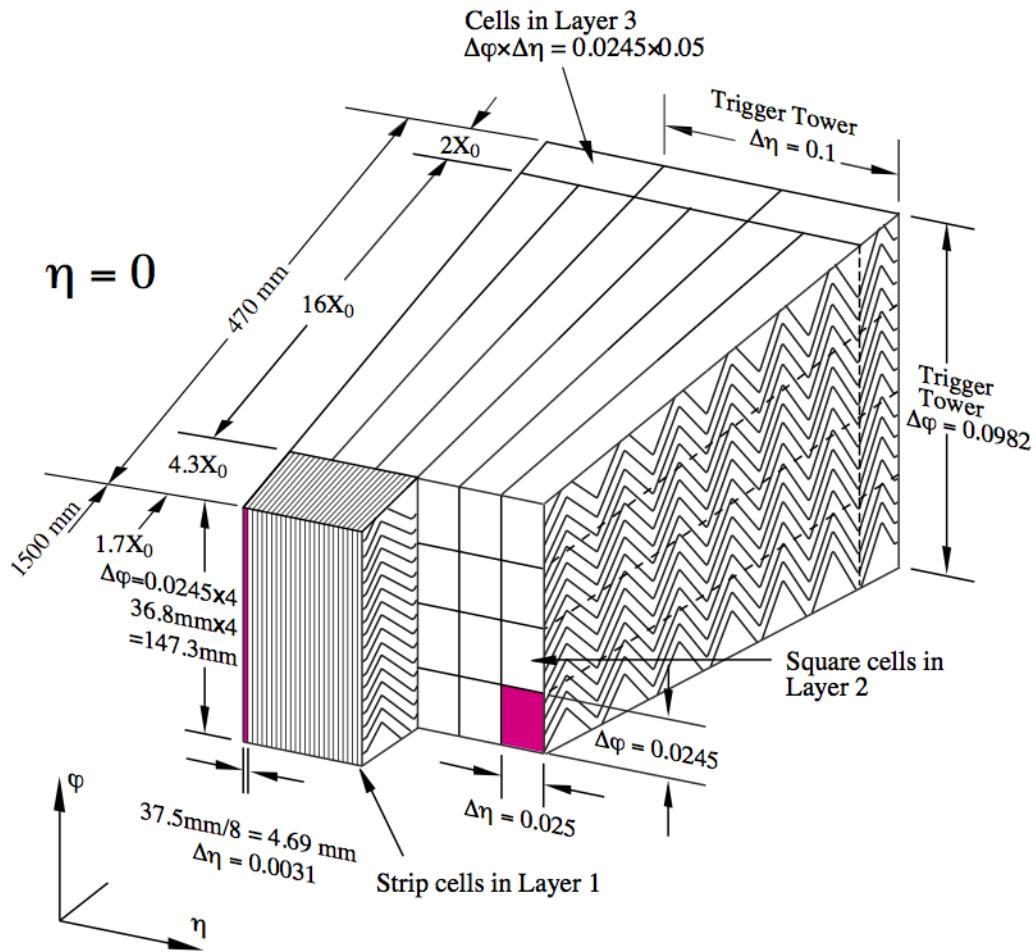


Figure 5.7: Schematic of ECAL.

Figure 5.8: Schematic of HCAL.

520 The energy resolution of the calorimeter subsystems are:

521
$$\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} \oplus 0.7\% \text{ EMC}$$

522
$$\frac{\sigma_E}{E} = \frac{50\%}{\sqrt{E}} \oplus 3\% \text{ hadronic barrel -Natasha check if barrel and end-cap truly}$$

523 have same energy resolution

524
$$\frac{\sigma_E}{E} = \frac{50\%}{\sqrt{E}} \oplus 3\% \text{ hadronic end-cap}$$

525
$$\frac{\sigma_E}{E} = \frac{100\%}{\sqrt{E}} \oplus 10\% \text{ hadronic end-cap}$$

526 **5.4 Muon Spectrometer**

527 The muon spectrometer (MS) is the outermost detector system in ATLAS.
528 Muons with a $p_T > 4$ GeV are energetic enough to reach the MS. To measure the
529 momentum of these muons two barrel and end-cap toroid magnets are used covering
530 $|\eta| < 1.4$ and $1.6 < |\eta| < 2.7$. For $1.4 < |\eta| < 1.6$, a combination of the barrel
531 and end-cap toroidal magnetic fields bend muon trajectories. The detector in the
532 barrel region form three concentric rings at $R = 5, 7.5, 10$ m and are segmented
533 in ϕ to accommodate the magnets. The end-cap region consists of three circular
534 planes perpendicular to z and located at $|z| = 7.4, 14, 21.5$ m from the interaction
535 region. An additional detector at $|z| = 10.8$ m covers the transition region between
536 the barrel and end-cap.

537 The MS consists of four subsystems: Monitored Drift Tubes (MDT), Cathode
538 Strip Chambers (CSC), Resistive Plate Chambers (RPC), and Thin Gap Cham-
539 bers (TGC). The first two subsystems are used primarily for measuring muon track
540 parameters, while the RPC and TGC subsystems are used for muon triggering.
541 A schematic of this system is shown in Figure 5.9.

542 The MDT subsystem consists of precision tracking chambers for $|\eta| < 2.7$,
543 except for the inner most end-cap layer ($2.0 < |\eta| < 2.7$), where CSCs are used.
544 The basic unit of MDT chambers are thin walled Aluminum tubes with a diameter
545 of 3 cm and length of 0.9-6.2 m. These tubes are filled with a mixture of Ar-CO₂
546 gas with a 50 μ m W-Rn wires running down the center of the tube which is kept at
547 3080 V. Since the maximum drift time of these chambers is 700 ns, they are not
548 used for triggering. MDT chambers consist of 3-4 layers of tubes mounted on a
549 rectangular support system, as seen in Figure 5.10, orientated along ϕ to measure
550 the coordinate in the bending plane of the magnetic field with a resolution of 35
551 μ m.

552 The MDT subsystem can only handle hit rate below $150\text{Hz}/\text{cm}^2$. For this
553 reason, CSCs are used in the innermost end-cap layer where hit rates are larger.
554 CSCs can handle hit rates up to $1000\text{Hz}/\text{cm}^2$. CSC are multiwire proportional
555 chambers. These chambers are filled with a Ar- CO_2 gas mixture and evenly spaced
556 wires kept at 1900 V. These wires are orientated in the radial direction but not
557 read out. Instead on one side of the cathode are copper strips parallel to the wires,
558 measuring η , while on the other side of the cathode are strips parallel to the wires
559 measuring ϕ . The width between strips is approximately 1.5 mm providing a
560 resolution of $60\ \mu\text{m}$ in the bending-plane and 5 mm in the non-bending plane.

561 Since the CSC and MDT systems do not poor time resolution, the RPC and
562 TGC systems are used for triggering. The RPC system is used in the barrel region
563 ($|\eta| < 1.05$). RPC consist of two parallel resistive plates separated by a 2 mm
564 insulated spacer with 100 mm spacing kept at 9.8 kV 5.11. A gaseous mixture of
565 $\text{C}_2\text{H}_2\text{F}_4$, C_4H_{10} , and SF_6 fills the space between the two plates. Metallic strips
566 on the outer faces of the plates are used to read out signals produced by the
567 gas ionizing. The middle barrel layer consists of two layers of RPCs on either
568 side of the MDT layer and one layer on the outermost MDT layer. Each layer
569 contains two orthogonal sets of metallic strips providing η and ϕ measurements.
570 The timing resolution of RPCs is 1.5 ns, and therefore may be used to identify
571 bunch crossings.

572 Finally, the TGCs are used in the end-cap regions and primarily used to pro-
573 vide L1 trigger decisions and ϕ measurements. TGCs are multi-wire proportional
574 chambers consisting of arrays of gold-coated tungsten wires placed between two
575 cathode planes. These wires are separated by 1.8 mm and cathodes are 1.4 mm
576 from the wires. Orthogonal to the wires, on the opposite side of the cathode plane
577 are copper strips held at 2900 V. The chambers are filled with a mixture of CO_2

578 and n-pentane gas, the latter acts as a quenching gase to prevent avalances initiated
579 by secondary γ -rays from the primary avalanche. Figure 5.12 is a schematic
580 of a TGC. The timing resolution of TGCs is less than 25 ns and therefore are used
581 for bunch crossing measurements.

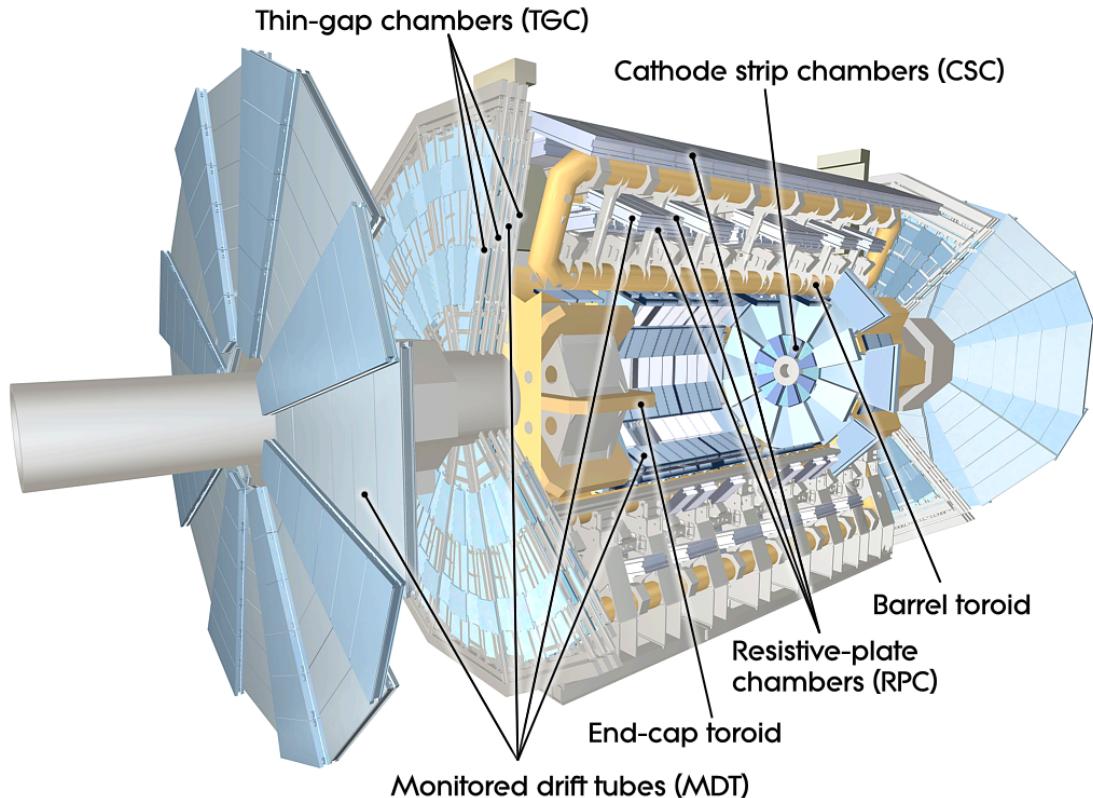


Figure 5.9: Schematic of Muon Spectrometer [cite G35]

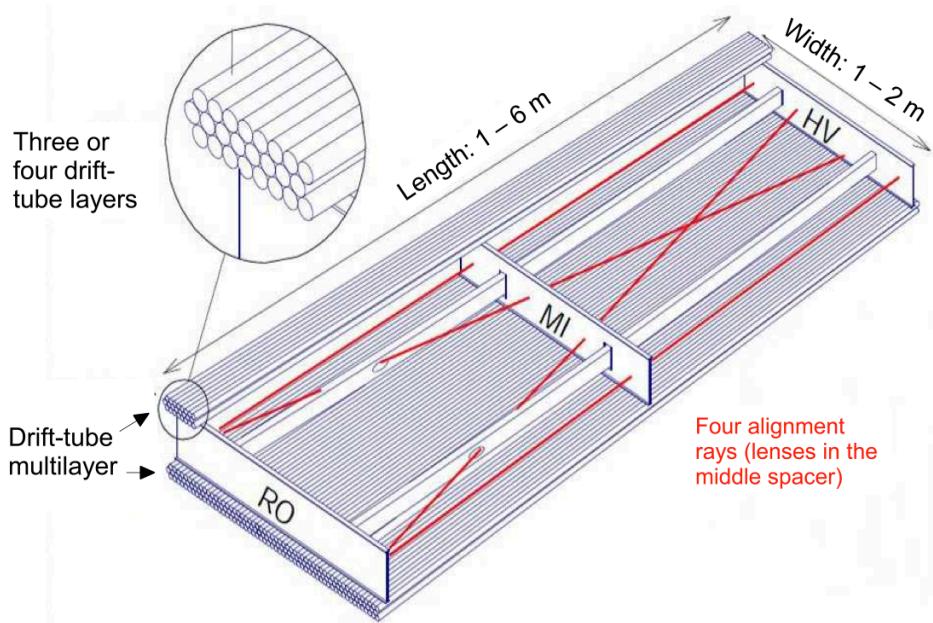


Figure 5.10: Schematic of MDT chamber. [cite G35]

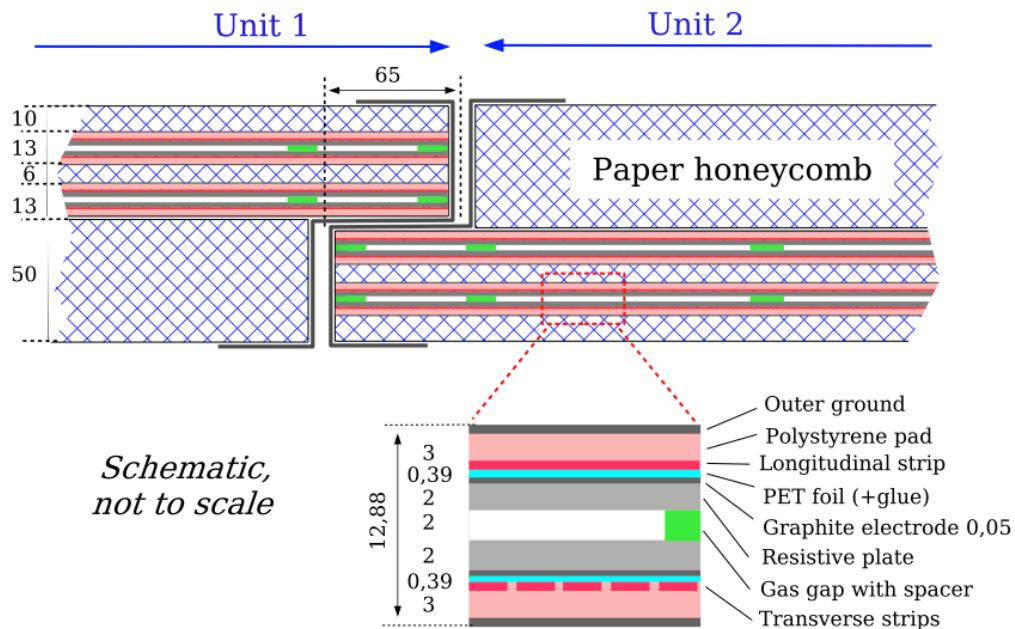


Figure 5.11: Schematic of RPC chamber, which is used for triggering in the central region of the detector [cite G35].

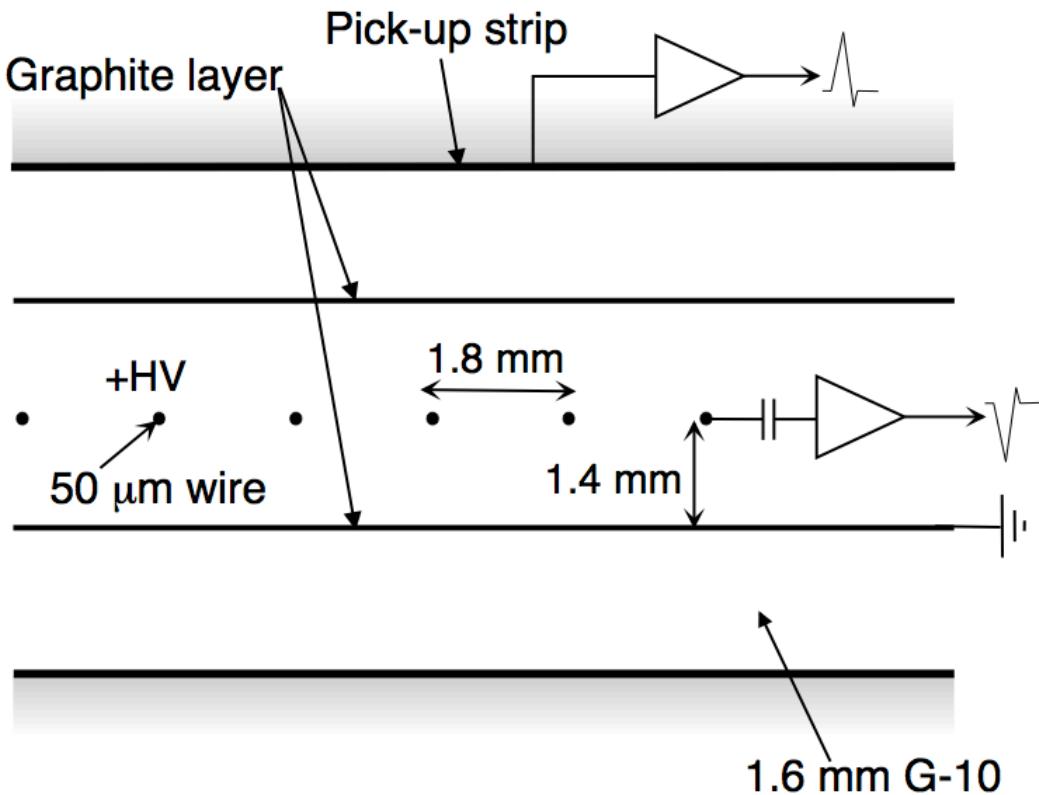


Figure 5.12: Schematic of TGC chamber, which is used for triggering in the muon end-cap region. [cite G35]

582 5.5 Magnet System

583 A particles with charge, q , and velocity v , moving in magnetic field, B , ex-
 584 periences a force, $F = qv \times B$. This force can cause charged particles to have a
 585 curved trajectory in magnetic fields, which the ID and MS use to determine the
 586 particles p_T . The central solenoid provides the magnetic field for the ID and the
 587 toroidal magnets provide the magnetic field for the MS.

588 The layout of the magnet system is shown in Figure 5.13. The central solenoid
 589 consists of a single-layer Al-stabilized NbTi conductor coil wound inside an Al

590 support cylinder. The solenoid is 5.8 m long, 50 cm thick and has an inner radius
591 of 1.23 m. It is cooled to 4.5 K to reach superconducting temperatures and shares
592 the liquid argon calorimeter vacuum vessel to minimize material in the detector.
593 A current of 7.730kA produces a 1.998 T solenoidal magnetic field, pointing in
594 the $+z$ direction.

595 The toroidal magnet system consists of a barrel and two end-cap toroidal
596 magnets used to a magnetic field outside the calorimeters that is orientated along
597 ϕ . Each barrel toroid is 25.3 m long with an inner and outer diameter of 9.4 and
598 20.1 m and weighs 830 tonnes. Endcap toroids are 5 m long with an inner and
599 outer radius of 1.65 and 10.7 m. Both toroid systems use Al-stabilized Nb/Ti/Cu
600 conductors. The magnetic field strength of the barrel and endcap regions are 0.5
601 and 1 T.

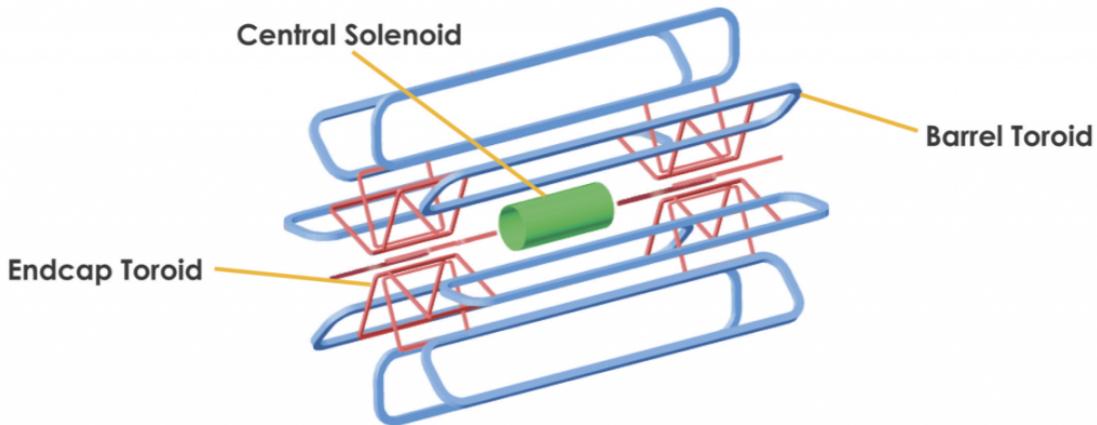


Figure 5.13: Layout of ATLAS magnet systems.

602 **5.6 Trigger System**

603 Since collisions occur every 25 ns and reading out all detector channels and
604 storing that information is not currently feasible (saving 60 million megabytes per

second), the majority of events are not kept for analysis. ATLAS uses a multi-stage trigger system to select approximately 1,000 of the 1.7 billion collisions that occur each second (corresponding to a rate of 1 kHz from the 40 MHz proton collision rate). The first stage of the trigger system is the hardware level (L1) trigger. This trigger reduces the event rate to 100 kHz by identifying Regions-of-Interest (ROIs) containing high p_T leptons, photons, jets, or E_T^{miss} by using information from RPCs, TGCs, and calorimeters to make a 2.5 μ s decision. This information is then passed to a high-level trigger (HLT) which further decreases event rates to 1 kHz. The HLT uses finer granularity measurements from the MS and ID to perform simplified offline reconstruction to decide which events to keep.

Part III

616

Method

617

618 **Chapter 6**

619 **Dataset and Simulated Samples**

620 **6.1 Dataset**

621 This analysis uses pp collision data collected from 2015 to 2018 at $\sqrt{s} = 13$
622 TeV, corresponding to 139/fb of data as shown in Figure 6.1 and 6.2. From this
623 dataset, only those events in which the tracker, calorimeters, and muon spectrom-
624 eter have good data quality are used. For a given event, the solenoid and toroidal
625 magnets must also be operating at their nominal field strengths. In addition to
626 this, events must pass further quality checks to reject events where detector sub-
627 systems may have failed. These selections reject events that containing LAr noise
628 bursts, saturation in the electromagnetic calorimeter, TileCal errors, and failures
629 in event recovery due to tracker failures. Events with information missing from
630 subsystems (usually due to busy detector conditions) are rejected. Events must
631 also contain a primary vertex with at least two associated tracks, where the pri-
632 mary vertex is selected as the vertex with the largest $\sum p_T^2$ over tracks associated
633 with the vertex and $p_T > 0.5$ GeV.

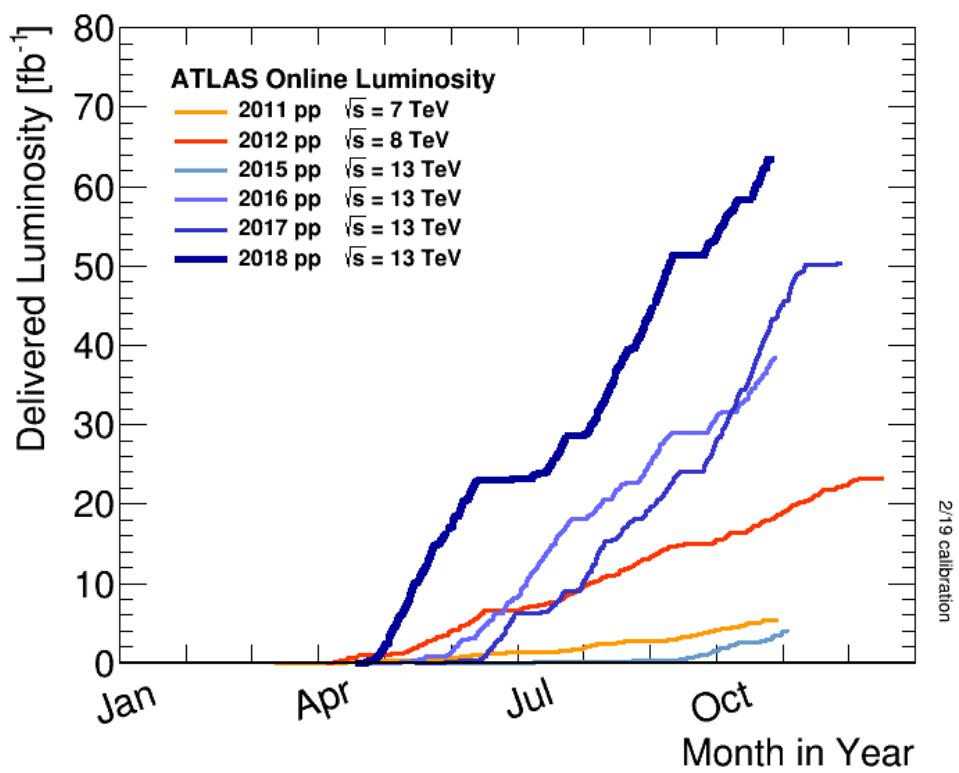


Figure 6.1: Integrated luminosity for data collected from ATLAS from 2011 - 2018

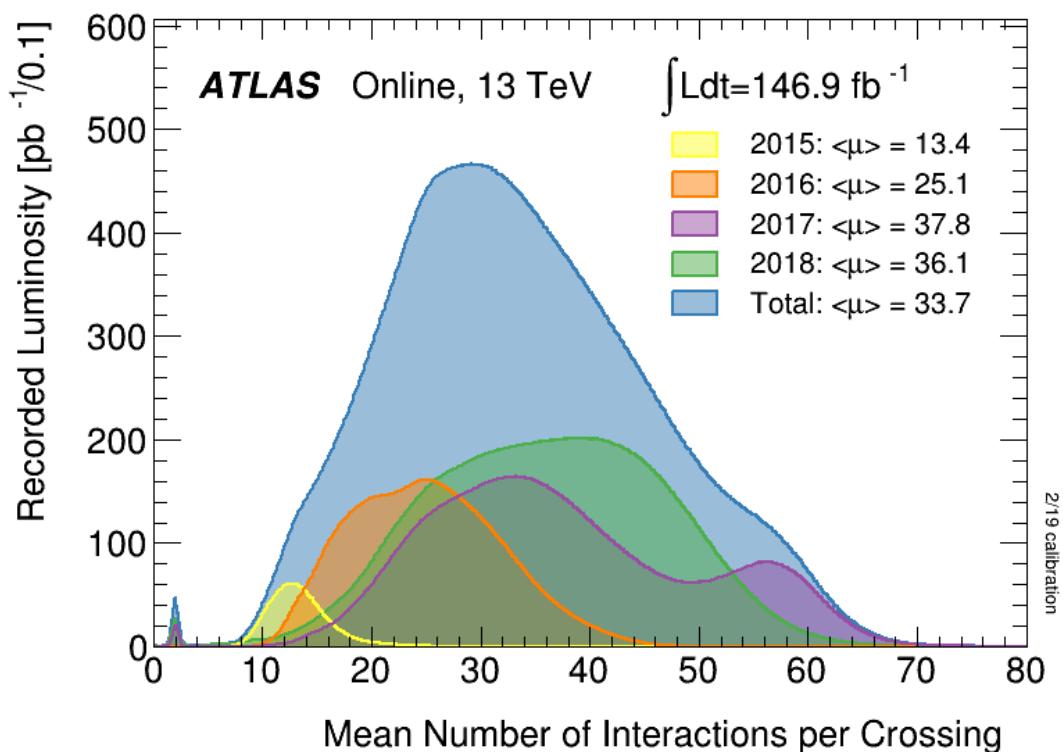


Figure 6.2: Mean number of interactions per crossing for data collected from ATLAS from 2011 - 2018

634 **6.2 Simulated Samples**

635 Samples are simulated in order to model backgrounds, evaluate signal ac-
636 ceptance, optimize event selection and estimate systematic and statistical uncer-
637 tainties. The dominant backgrounds for this analysis are $W/Z + \text{jets}$, diboson
638 (WZ/WW), $t\bar{t}$, single top and multijet production.

639 $W/Z+\text{jet}$ events are simulated using Sherpa 2.2.1 at NLO [cite [29]] and merged
640 with the Sherpa parton shower using the ME+PS prescription [11]. These events
641 are then normalized to NNLO cross sections. The $t\bar{t}$ and single-top backgrounds
642 are generated with Powheg-Box with NNPDF3.0NLO PDF sets in the matrix
643 element calculation [cite[35]]. Top quarks are decayed using MadSpin [cite[36]].
644 For all processes, the parton shower, fragmentation, and underlying event are
645 simulated using Pythia 8.320 with the A14 tune set[cite[ATL-PHYS-PUB-2014-
646 02]]. Diboson processes are generated using Sherpa 2.2.1.

647 Signal samples are simulated using MadGraph 5-2.2.2 [cite 42] and Pythia
648 8.186 with NNPDF230LO. RS Graviton samples are generated with $k/M_{PL}=1$.
649 HVT Model A and B samples are simulated with $g_H = -0.56$ and $g_f = -0.55$,
650 as the difference in the width of the samples is smaller than detector resolution.
651 Model C is generated by setting $g_H = 1$ and $g_f = 0$ to model VBF production of
652 HVT bosons. Signals are generated for masses between 300 GeV and 6 TeV.

653 **6.3 Object Selection**

654 **6.3.1 Electrons**

655 Electrons are reconstructed from electromagnetic showers in the LAr EM
656 calorimeter. During reconstruction cells of $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ are grouped
657 into 3×5 clusters. These clusters are then scanned for local maxima that seed elec-

tron clusters. These clusters must then be matched to ID track from the PV. This requirement minimizes non-prompt electron and fake electron backgrounds. Electrons must pass identification and isolation requirements. Electron identification (loose, medium, tight) classification is based on the discriminant of the multivariate analysis that identifies electrons using a likelihood based method. For this analysis tight electrons are used. Electrons are also required to be isolated. The electron isolation is calculated by comparing the sum of the transverse momentum in a cone around the electron of size $\Delta R = 0.2$ to the transverse momentum of the electron. This quotient must be less than 3.5, to further reject non-prompt photons and other fake backgrounds (multijet). Electrons in this analysis are also required to have $p_T > 30$ GeV and $|\eta| < 2.47$. Electrons are also required to have $p_T > 30$ GeV.

Electrons are calibrated to determine a data-driven scale factors from $J/\Psi \rightarrow ee$, $Z \rightarrow ee$, $Z \rightarrow \ell\ell\nu$ processes. These corrections account for the non-uniform response of the detector by introducing modeling and reconstruction uncertainties.

6.3.2 Muons

As muons traverse the entire detector, they are reconstructed from ID and MS tracks. For this analysis the muon identification and isolation working points are chosen to minimize the contributions from non-prompt muons. Towards this end, the medium muon identification working point is used. For this working point, two types of reconstructed muons are used: combined and extrapolated muons (CB and ME, respectively). For CB muons, ID and MS tracks are reconstructed independently and a combined track fit is performed by adding or removing MS tracks to improve the fit quality. ME muons are reconstructed from only MS tracks with hits in at least two layers, which ensures the track originates from the

683 PV. ME muons extend the acceptance for muon reconstruction outside the ID
684 from $2.5 < |\eta| < 2.7$. The medium identification working point uses CB and ME
685 tracks. CB tracks must have at least 3 hits in two MDT layers. ME tracks are
686 required to have at least three MDT/CSC hits. To further minimize contributions
687 from fake muons, the selected muons are required to be isolated from other tracks,
688 as muons from W, Z decays are often isolated from other particles. To insure the
689 selected muons are isolated, the scalar sum of the transverse momentum of tracks
690 in a cone of $\Delta R = 0.3$ compared to the transverse momentum of the muon must
691 be less then 0.06. Muons are also required to have $p_T > 30$ GeV.

692 Muons are calibrated using well-studied resonances $J/\Psi \rightarrow \mu\mu$ (low- p_T), $Z \rightarrow$
693 $\mu\mu$ (high- p_T). Figure 6.3 shows the combined muon p_T uncertainty from this
694 calibration. The total systematic uncertainty is less then 1% for all p_T ranges
695 considered in this analysis.

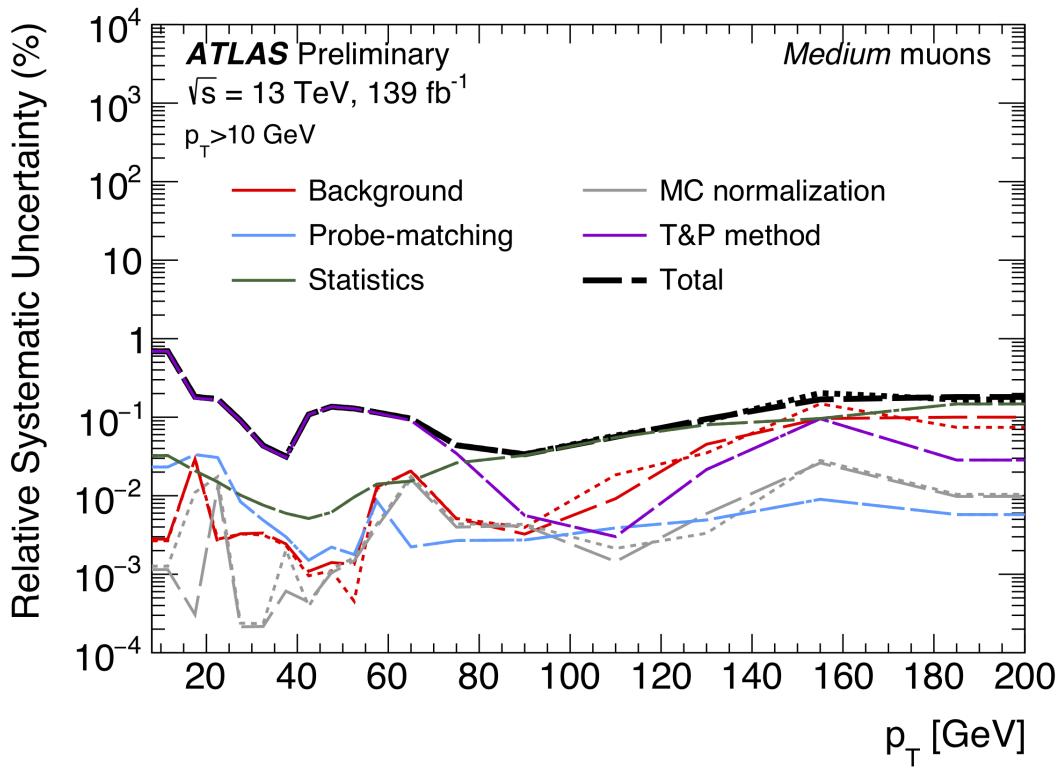


Figure 6.3: [4] This figure show the breakdown of the muon reconstruction efficiency scale factor measured in $Z \rightarrow \mu\mu$ as a function of p_T

6.3.3 small-R jets

Calorimeter jets are used to reconstruct the hadronically decaying W/Z candidates in the resolved analysis. These jets are less boosted and therefore spatially separated and reconstructed separately. These jets are constructed from topologically connected clusters of calorimeter cells (topoclusters), seeded from calorimeter cells with energy deposits significantly above the noise threshold. These cells are then used as inputs to the $anti - k_T$ algorithm [12] with a distance = 0.4, here called small-R jets. These jets are calibrated to compensate and account for biases from jet reconstruction.

The jet energy is calibrated sequentially as shown in Figure 6.4. After the

706 jet direction is corrected to point to the PV, the energy of the jet is corrected.

707 First, the jet energy is corrected to account for pileup contributions based on the
708 p_T and area of the jet (these corrections are extracted from a $pp \rightarrow jj$ sample).

709 Following this, another pileup correction is applied that scales with μ and N_{PV} .

710 Then, MC-based corrections are applied that are meant to transform the jet
711 energy and η back to truth level. Therefore, these corrections account for the
712 non-compensating nature of the ATLAS calorimeters and inhomogeneity of the
713 detector. Following this the Global sequential calibration is applied that reduces
714 flavor dependence and jet that deposit energy outside the calorimeters. Finally,
715 in-situ corrections are applied that account for differences in jet responses between
716 data and simulation ($\gamma/Z+jet$ and multijet samples are used). These differences
717 can be due to mismodelling of the hard scatter event, pile-up, jet formation, etc.

718 Jet used in this analysis must have $p_T > 30$ GeV and $|\eta| < 2.5$. To further
719 reduce fake jets the jet-vertex-tagger (JVT) is used to reject pile-up jets [cite 43
720 P]. The JVT uses two track-based variables, corrJVF and R_{p_T} to calculate the
721 likelihood that the jet originated from the PV. The corrJVF compares the scalar
722 sum of the p_T of tracks associated with the jet and PV to the scalar sum of the
723 p_T of tracks associated with the jet. This variable also includes a correction that
724 reduces the dependency of corrJVF with the number of reconstructed vertices in
725 the event. R_{p_T} is given by the ratio of the scalar sum of the p_T of tracks associated
726 with the jet and PV to the p_T of the jet. Both of these variables peak around zero
727 for pileup jets, as these jets are unlikely to have tracks associated with the PV.
728 JVT cuts are applied to all jets with $p_T > 120$ GeV. Central jets ($|\eta| < 2.4$) are
729 required to have a $JVT > 0.59$ and forward jets ($2.4 < |\eta| < 2.5$) are required to
730 have $JVT > 0.11$.

731 To further reject fake jets, jets must pass quality requirements based on the

⁷³² following variables ([cite P42]):

⁷³³ - f_Q^{LAr} : fraction of energy of jet's LAr cells with poor signal shape

⁷³⁴ - f_Q^{HEC} : fraction of energy of jet's HEC cells with poor signal shape

⁷³⁵ - E_{neg} : sum of cells with negative energy

⁷³⁶ - f_{EM} : fraction of jet's energy deposited in EM calorimeter

⁷³⁷ - f_{HEC} : fraction of jet's energy deposited in HEC calorimeter

⁷³⁸ - f_{max} : maximum energy fraction in any single calorimeter layer

⁷³⁹ - f_{ch} : ratio of the scalar sum of the p_T of a jet's charged tracks to the jet's p_T

⁷⁴⁰ Jets selected for the resolved analysis must pass one of the following criteria:

⁷⁴¹ - $f_{HEC} > 0.5$ and $|f_Q^{HEC}| > 0.5$ and $\langle Q \rangle > 0.8$

⁷⁴² - $|E_{neg}| > 60$ GeV

⁷⁴³ - $f_{EM} > 0.95$ and $f_Q^{LAr} > 0.8$ and $\langle Q \rangle > 0.8$ and $|\eta| < 2.8$

⁷⁴⁴ - $f_{max} > 0.99$ and $|\eta| < 2$

⁷⁴⁵ - $f_{EM} < 0.05$ and $f_{ch} < 0.05$ and $|\eta| < 2$

⁷⁴⁶ - $f_{EM} < 0.05$ and $|\eta| > 2$

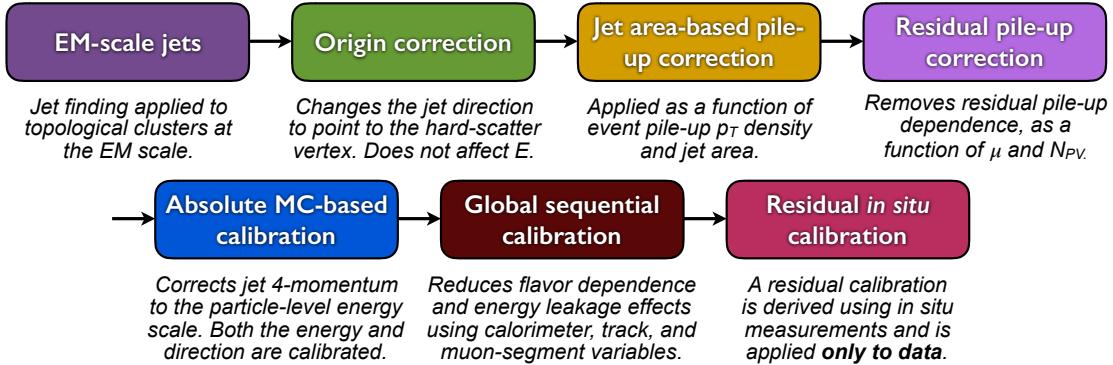


Figure 6.4: [6] This diagram shows the calibration stages for EM jets.

747 6.3.4 large-R jets

748 Large-R ($\Delta R = 1.0$) jets are used to reconstruct the high- $p_T W/Z \rightarrow qq$ candi-
 749 dates in the merged analysis. Track-Calorimeter Clusters (TCCs) are used to reconstruct
 750 these jets [cite ANA 50]. These jets are constructed via a pseudo particle flow
 751 method using ID tracks matched to calorimeter clusters. The angular resolution
 752 of the calorimeter degrades sharply with jet p_T , but the jet energy resolution im-
 753 proves. The tracker has excellent angular resolution improves with p_T . Therefore,
 754 by matching tracks to jets, TCCs have more precise energy and angular resolution
 755 the jets constructed from calorimeter information only. These jets are required to
 756 have $p_T > 200$ GeV, $|\eta| < 2.0$ and $m_J > 50$ GeV.

757 TCC jets are trimmed as detailed in [cite ANA 45], which suppresses pileup
 758 and soft radiation in the jet, the jet mass is calculated as the four-vector sum
 759 of the jet's constituents (assuming massless constituents). The jet mass peaks
 760 around the W/Z boson mass for $W/Z \rightarrow qq$ jets, and more broadly for quark and
 761 gluon induced jets.

762 These jets are then tagged as W/Z jets if they pass the jet mass and D_2
 763 cuts. The jet substructure variable D_2 is given by the ratio of energy correlation
 764 functions based on energies and pair-wise angles of a jet's constituents [cite ANA

765 46, 47]:

$$D_2^{\beta=1} = E_{CF3} \left(\frac{E_{CF1}}{E_{CF2}} \right)^3 \quad (6.1)$$

766 Where the energy correlation functions are defined as:

$$E_{CF1} = \sum_i p_{T,i} \quad (6.2)$$

767

$$E_{CF2} = \sum_{ij} p_{T,i} p_{T,j} \Delta R_{ij} \quad (6.3)$$

768

$$E_{CF3} = \sum_{ijk} p_{T,i} p_{T,j} p_{T,k} \Delta R_{ij} \Delta R_{jk} \Delta R_{ki} \quad (6.4)$$

769 A two-dimensional optimization of the jet mass and D_2 thresholds was per-
770 formed to provide maximum sensitivity for this analysis. Figure ?? shows the
771 optimized thresholds on D_2 and jet mass as a function of p_T . Figure 6.6 shows
772 the efficiency of the optimized W/Z taggers as a function of jet p_T .

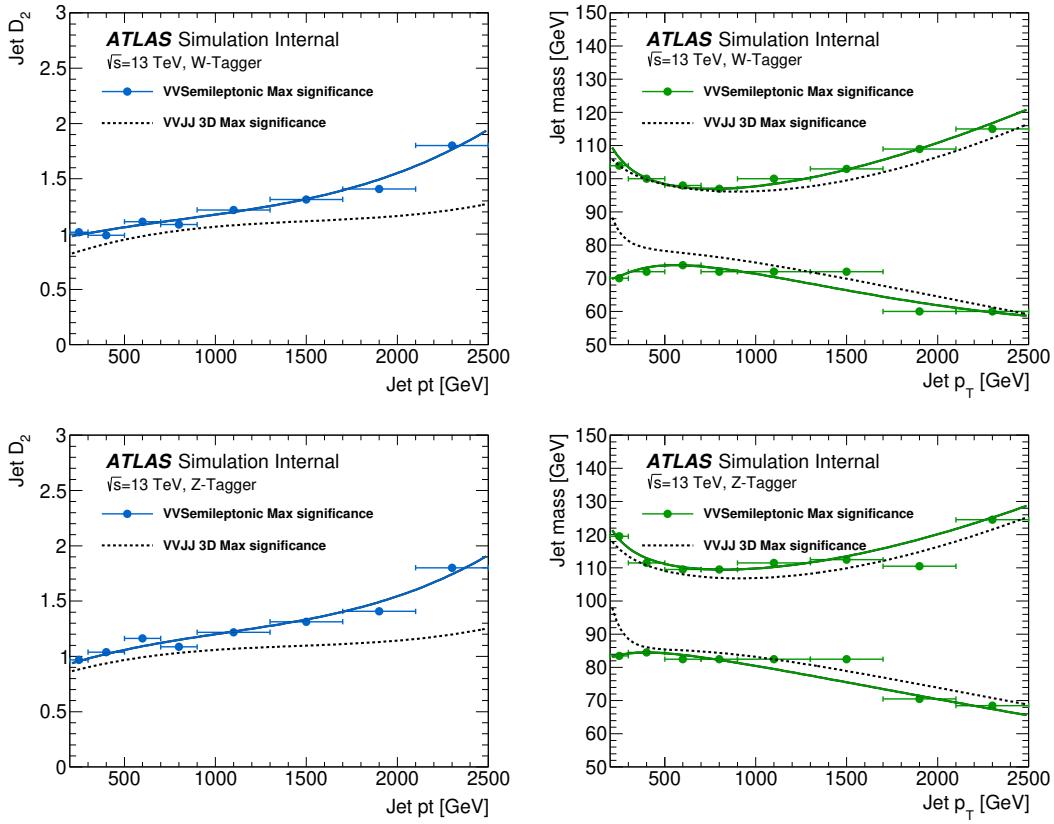


Figure 6.5: The upper cut on D_2 (a) and jet mass window cut i.e. the upper and lower boundary of the mass (b) of the W -tagger as a function of jet p_T . Corresponding values for Z -tagger are shown in (c) and (d). The optimal cut values for maximum significance are shown as solid markers and the fitted function as solid lines. Working points from $VV \rightarrow JJ$ [ATLAS-HDBS-2018-31-002] is also shown as dashed lines as a reference. Natasha reword?

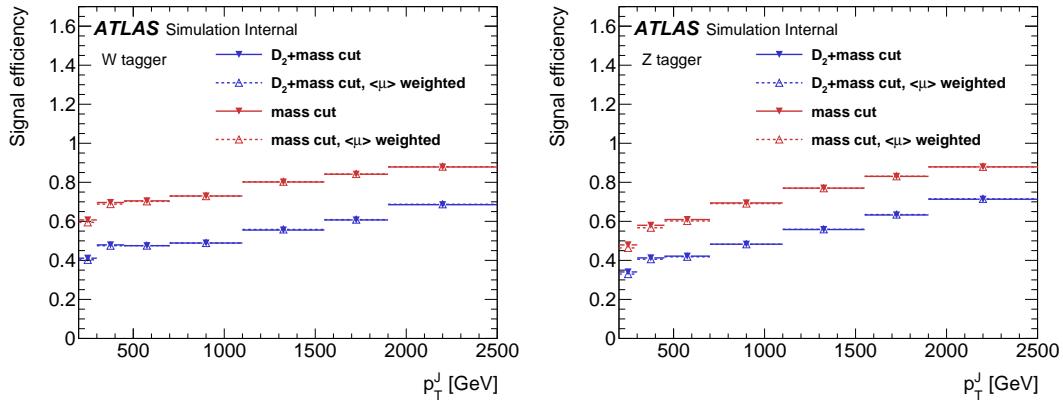


Figure 6.6: Natasha write caption

773 **6.3.5 Variable Radius jets**

774 Variable-radius (VR) track jets are used to identify b-quark induced jets in
775 large-R jets [cite ANA 52]. These jets use a p_T dependent cone size defined as:

$$R_{eff}(p_{T,i} = \frac{\rho}{p_{T,i}}) \quad (6.5)$$

776 for building jets from ID tracks with an anti-kt algorithm. For this analysis
777 $\rho = 30$ GeV and an upper and lower limit on cone size are set to 0.02 and 0.4,
778 respectively. Collinear VR jets are possible, so track jets that are not separated
779 by the the smaller jet's cone size are not used. These jets are also required to
780 have $p_T > 10$ GeV and $|\eta| < 2.5$.

781 **6.3.6 MET/neutrinos**

782 As neutrinos are uncharged and color less they do not leave tracks or jets in
783 the detector. For this reason, neutrinos are reconstructed calculated the E_T^{miss} .
784 This quantity is defined as negative vector sum of p_T all the physics objects and
785 an extra "soft" term. The "soft" term accounts for energy depsoits not associated
786 with any of the objects in the event. For this analysis the soft term is given by
787 the summing the p_T of all ID tracks not associated with objects in the event. The
788 selected tracks must be matched to the primary vertex, which decreases pile-up
789 contamination [cite G 217 218]. The tight working point is used [Natasha look up
790 what this means].

791 **6.3.7 Jet Flavor Tagging**

792 To further classify events, the small radius jets are identified as originated
793 for a b-quark or not using the multivariate b -tagging algorithm (BDT), MC2c10

794 [cite G 210 199]. This algorithm uses the impact parameters of the jet's ID tracks,
795 secondary vertices (if they exist), and reconstructed flight paths of b and c hadrons
796 in the jet to determine if the jet was induced by a b -quark. For this analysis the
797 85% efficient working point of this algorithm is used to a fixed cut on the BDT
798 discriminant that yields an 85% tag rate, and c , τ , and light-flavor jet rejection of
799 3, 8, and 34 respectively in a simulated $t\bar{t}$.

800 **6.3.8 Overlap Removal**

801 The reconstructed jets and leptons in this analysis can arise from the same
802 energy deposits. For instance, a jet may radiate an electron that is then recon-
803 structed separately as the signal lepton in the event. To mitigate this confusion
804 of multiple objects originating from a single jet or lepton overlapping objects are
805 removed via a procedure referred to a overlap removal. In this procedure the sep-
806 aration of the two objects, $\Delta(R) = \sqrt{(\Delta\eta)^2 + (\Delta\text{phi})^2}$ determines which object is
807 removed from the event.

808 The overlap selections used in this analysis are:

- 809 - when an electron shares a track the electron with the lower p_T is rejected,
810 as it is more likely to be a fake electron
- 811 - when a muon and electron share a track the muon is rejected if it is a
812 calo-muon, otherwise the electron is rejected
- 813 - when $\Delta R < 0.2$ for an electron and jet, the jet is rejected to maximize signal
814 acceptance
- 815 - when $\Delta R > 0.2$ for an electron and jet, the electron is rejected as likely
816 originated from decays within the jet

- 817 - when $\Delta R < \min(0.4, 0.04 + 10\text{GeV}/p_T^\mu)$ the muon is rejected, again maxi-
- 818 mizing signal acceptance, otherwise the jet is rejected
- 819 - when $\Delta R < 1.0$ for the a large-R jet and electron, the jet is rejected

820 **Chapter 7**

821 **Event Selection and**

822 **Categorization**

823 **7.1 Pre-selection**

824 Before applying topological cuts to suppress backgrounds and reduce data
825 size in this search, preselection cuts are applied which include trigger and event
826 requirements. Events must contain exactly one tight lepton (no additional loose
827 leptons), the $p_T^{\ell\nu} > 75$ GeV, and there must be at least two small-R jets or one
828 large-R jet.

829 **7.2 Trigger**

830 The data was collected using the lowest unprescaled single-lepton or E_T^{miss}
831 triggers, as summarized in Table [natasha add table]. Since the muon term is not
832 considered in the trigger E_T^{miss} calculation, the E_T^{miss} trigger is fully efficient to
833 events with high- p_T muons. For this reason, the E_T^{miss} trigger is used for events
834 where $p_T^\mu > 150$ GeV, to compensate for the poor efficiency of the single muon

⁸³⁵ trigger below 150 GeV (due to detector coverage).

Table 7.1: The list of triggers used in the analysis.

Data-taking period	$e\nu qq$ channel	$\mu\nu qq$ ($p_T(\mu\nu) < 150$ GeV) channel	$\mu\nu qq$ ($p_T(\mu\nu) > 150$ GeV) channel
2015	HLT_e24_lhmedium_L1EM20 OR HLT_e60_lhmedium OR HLT_e120_lhloose	HLT_mu20_iloose_L1MU15 OR HLT_mu50	HLT_xe70
2016a (run < 302919) $(L < 1.0 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1})$	HLT_e26_lhtight_nod0_ivarloose OR HLT_e60_lhmedium_nod0 OR HLT_e140_lhloose_nod0 HLT_e300_etcut	HLT_mu26_ivarmedium OR HLT_mu50	HLT_xe90_mht_L1XE50
2016b (run ≥ 302919) $(L < 1.7 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1})$	same as above	same as above	HLT_xe110_mht_L1XE50
2017	same as above	same as above	HLT_xe110_pufit_L1XE55
2018	same as above	same as above	HLT_xe110_pufit_xe70_L1XE50

836 **7.3 GGF/VBF RNN**

837 To classify events as originating from GGF/DY or VBF production a recursive
838 neural network (RNN) is used. This approach is more powerful than a cut-based
839 classification as it improves signal efficiency and analysis sensitivity by exploit-
840 ing correlations between variables that the RNN learns. In particular, a RNN
841 architecture is ideal as it can handle variable numbers of jets in the events.

842 The RNN uses the four-momentum of candidate VBF jets to classify events
843 as VBF or GGF topologies. As sometimes jets will be incorrectly reconstructed
844 the number of jets in the event are expected to vary across the inputs samples.
845 VBF candidate jets are identified by removing jets from the event that are likely
846 from $W/Z \rightarrow qq$. For the resolved regime this means removing the two leading
847 small-R jets from the VBF candidate jet list. For the merged regime this means
848 removing small-R jets that are $\Delta R < 1.0$. Also the VBF candidate jets are also
849 required to be within $|\eta| < 4.5$. From the list of remaining VBF candidate jets,
850 the two highest- p_T jets are chosen.

851 The architecture of the RNN is show in Figure 7.1. LSTMs are a type of
852 RNN that extract meaningful information and can retain it (unlike other neural
853 networks architectures). This is useful for VBF event classification for events with
854 two jets, where using the kinematic properties of both jets (and their correlations)
855 will lead to more efficient event classification.

856 In this RNN architecture, the VBF candidates are first passed to a masking
857 layer which checks the number of jets in the event. If there is only one jet, only
858 one LSTM layer is used. The output of masking is then passed to a LTSM cell
859 (with a tanh activation) [natasha cite LSTM], and then to a dropout layer, that
860 has a probability of 0.3 to completely forget the output of the LSTM. Dropout
861 is a regularization method, that prevents overfitting. The output of the dropout

862 layer is then passed to the second LSTM and then through another dropout layer
863 with a probability of 0.3.

864 The weights and other parameters of the network are learned by training the
865 network with VBF and GGF signals over 200 epochs with an Adam Optimizer
866 [natasha add reference]. The training is truncated if the network parameters are
867 unchanged after ten iterations. The training, testing and validation sets are 56,
868 30, and 14 percent of the input samples, respectively. Figure [add INT figure
869 32] shows the loss function of the network as a function of training epochs. The
870 validation test set has a smaller loss function as dropout was not applied. Figure
871 7.3 shows the ROC curve for the RNN using k-fold cross validation.

872 Finally this output is passed to a dense layer [natasha ask antonio about this]
873 and then to a sigmoid activation layer, leading to an overall RNN score. Figure 7.2
874 shows the RNN discriminant for shows modeling of the discriminant. The RNN
875 score is 0 for GGF and background processes and 1 for VBF processes. Figure
876 ?? shows the limits for various signal processes based on the RNN cut applied.
877 The most efficient cut was found to be RNN Score > 0.8, for VBF classification.

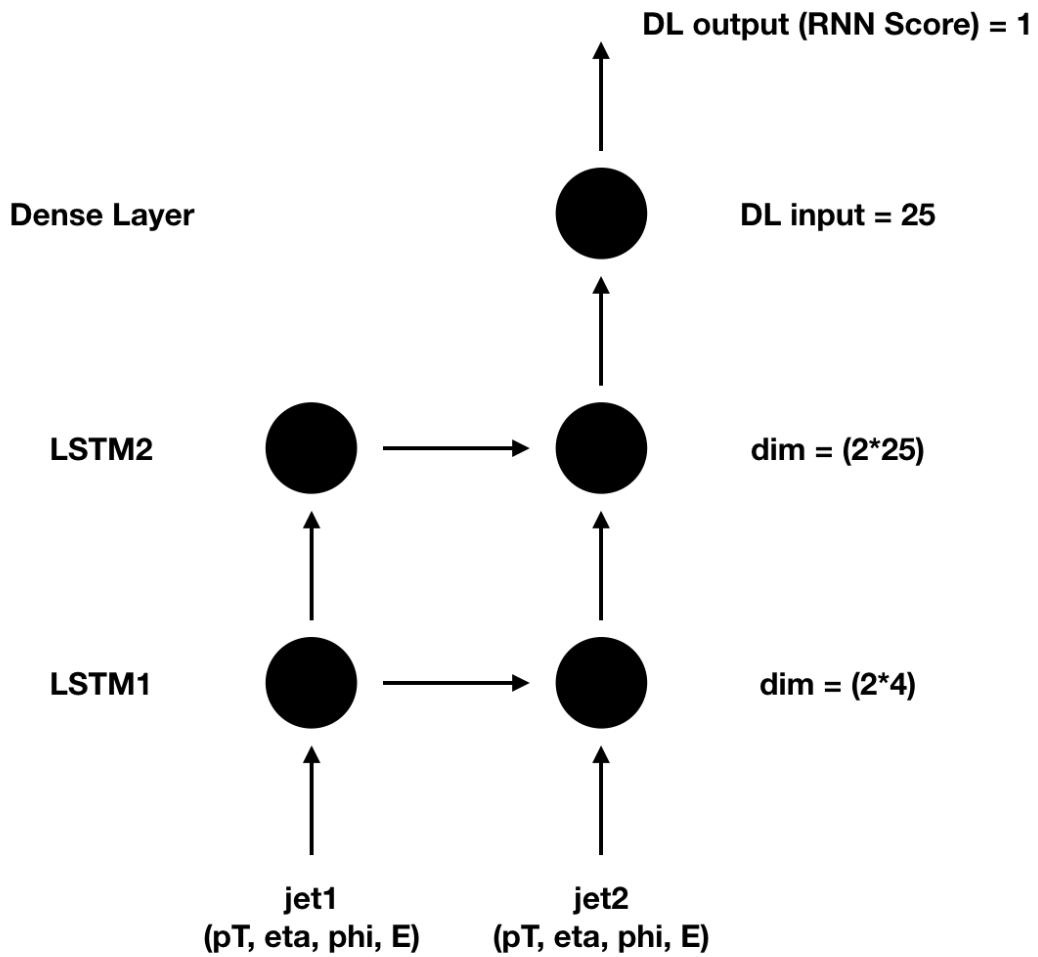


Figure 7.1: RNN architecture. Natasha add caption

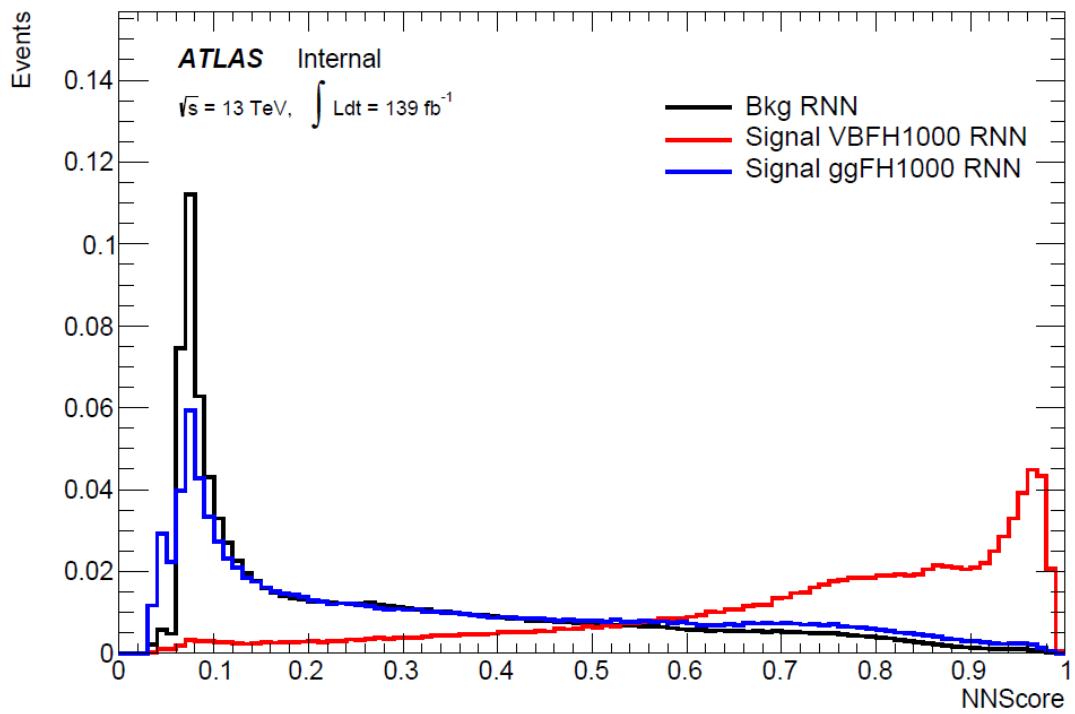


Figure 7.2: RNN Score distribution for ggF and VBF signals and backgrounds.

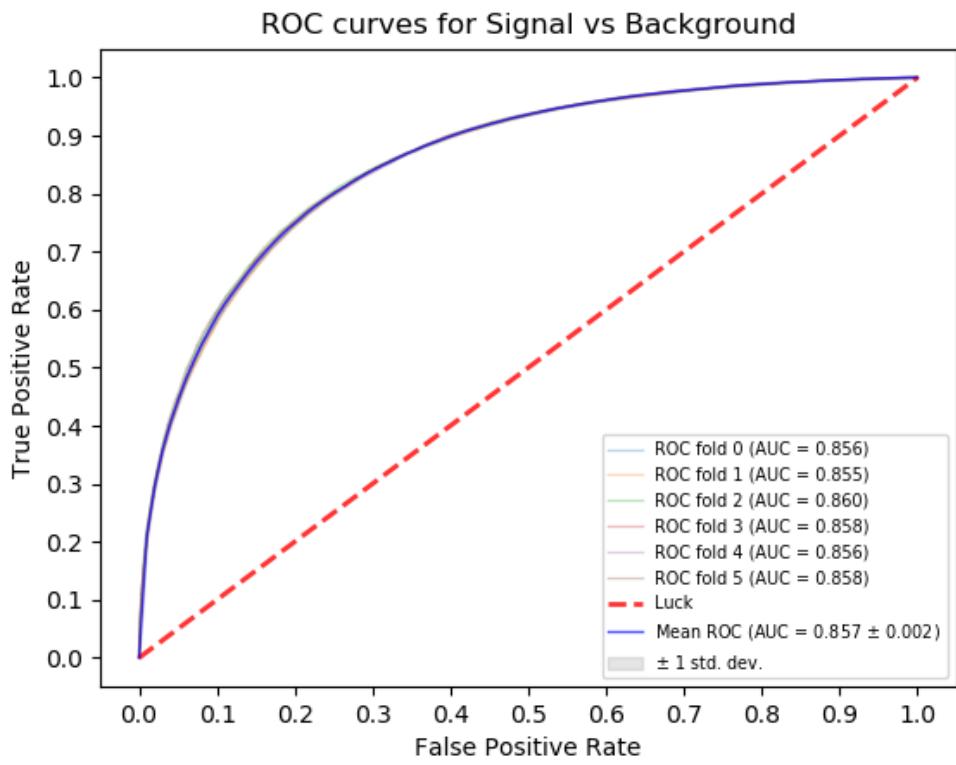


Figure 7.3: ROC curve using k-fold validation for RNN.

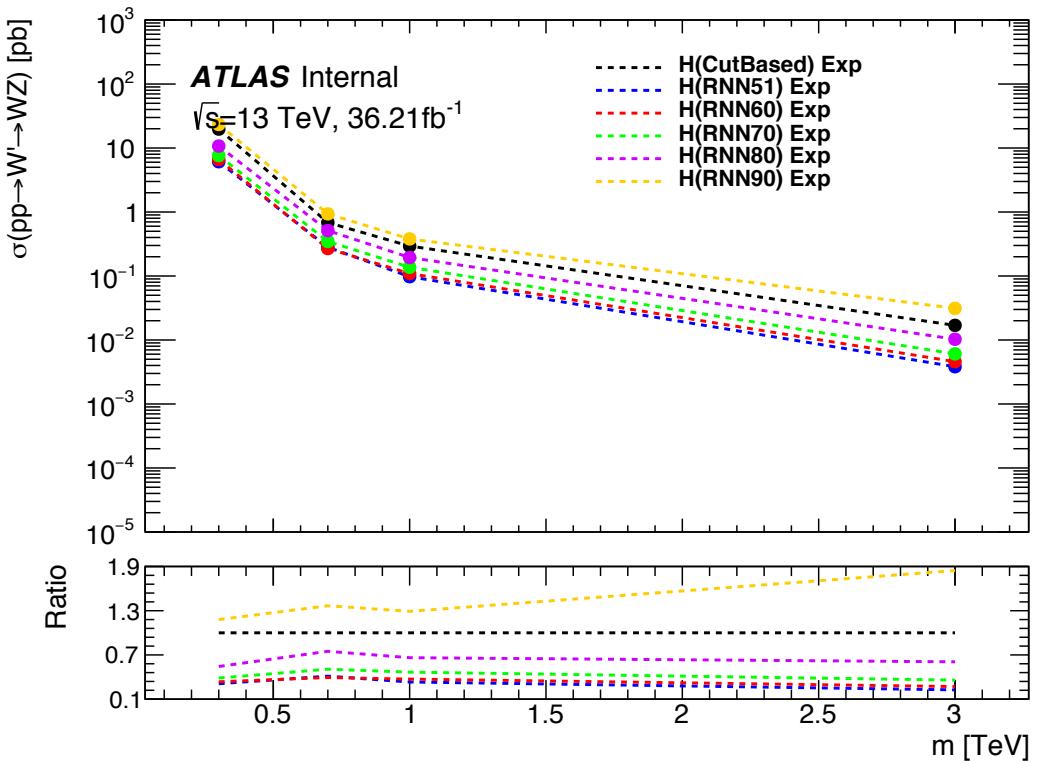


Figure 7.4: Comparison of GGF Z' limits for different RNN score selections.

878 7.4 Topological Cuts

879 Once an event is classified as VBF or GGF via the RNN it must pass other
 880 topological cuts that maximize signal efficiency and background rejection. First,
 881 to efficiently select events with $W \rightarrow \ell\nu$ candidate exactly one tight lepton is
 882 required and $E_T^{miss} > 100(60)$ GeV and $p_{T,\ell\nu} > 200(75)$ GeV in the merged (re-
 883 solved) analysis to suppress the multi-jet background. To more accurately model
 884 the two dominant backgrounds in this analysis, $W + \text{jets}$ and $t\bar{t}$, control regions are
 885 used constructed for each. These control regions are dominated by these processes
 886 and used to extract normalization factors that are then used in the signal region
 887 estimates.

888 For the merged analysis, in addition to the $W \rightarrow \ell\nu$ and $W/Z \rightarrow J$ selections

above, the relative boson p_T is cut to enhance signals, i.e. $\min(p_{T,\ell\nu}, p_{T,J})/m_{WV} > 0.35(0.25)$ for the GGF (VBF) category. To minimize $t\bar{t}$ contamination the signal region and $W+\text{jets}$ control region events with at least one b jet with $\Delta R > 1.0$ from the large-R jet are excluded. For the $t\bar{t}$ control region the event must contain at least one such b jet. High purity signal regions require the D_2 and W/Z mass window cut to be passed, whereas the low purity region only requires the W/Z mass window cut to be passed. Finally for events to be classified as tagged the large-R jet must contain exactly two b-tagged jets. Untagged events must have no more than one b-tagged jet matched to the large-R jet. These selections are shown in Table 7.2. The distributions for the variables used in merged analysis for top control regions are shown in Figure 7.6- 7.9.

Events failing the merged selection are then re-analyzed in the resolved category. To enhance resolved signals, the event should contain two high- p_T boson that are back-to-back in the x-y plane as shown by the cuts in Table 18. Again to suppress the $t\bar{t}$ background in the WCR and SR events are required to have no additional b-jets.

The WV system mass, m_{WV} is reconstructed from the lepton, neutrino, and hadronically-decaying boson candidate. The momentum of the neutrino along the z -direction is obtained by constraining the $W((Z))$ boson mass of the lepton neutrino system to be 80.3 (91.8) GeV/c^2 . For complex solutions to this constraint, p_Z is taken as either the real component of the complex solutions or the one with the smaller absolute value of the two real solutions. For the resolved analysis, m_{WV} is reconstructed by constraining the $W(Z)$ dijet system:

$$p_{T,jj}^{corr} = p_{T,jj} \times \frac{m_{W/Z}}{m_{jj}} \quad (7.1)$$

$$m_{jj}^{corr} = m_{W/Z} \quad (7.2)$$

Table 7.2: Summary of selection criteria used to define the signal region (SR), W +jets control region (W CR) and $t\bar{t}$ control region ($t\bar{t}$ CR) for merged 1-lepton channel.

Selection		SR		W CR (WR)		$t\bar{t}$ CR (TR1)	
		HP	LP	HP	LP	HP	LP
$W \rightarrow \ell\nu$	Num of Tight leptons	1					
	Num of Loose leptons	0					
	E_T^{miss}	$> 100 \text{ GeV}$					
	$p_T(\ell\nu)$	$> 200 \text{ GeV}$					
$W/Z \rightarrow J$	Num of large- R jets	≥ 1					
	D_2 cut	pass	fail	pass	fail	pass	fail
	W/Z mass window cut	pass	pass	fail	fail	pass	pass
	Numb. of associated VR track jets b -tagged	For $Z \rightarrow J$: ≤ 1 ($= 2$) for untagged (tagged) category					
	$\min(p_{T,\ell\nu}, p_{T,J}) / m_{WV}$	$> 0.35(0.25)$ for DY/ggF (VBF) category					
	Top-quark veto	Num of b -tagged jets outside of large- R jet	0		≥ 1		
Pass VBF selection			no (yes) for DY/ggF (VBF) category				

913 where m_{jj} and $m_{W/Z}$ are the reconstructed invariant mass of the hadronically-
914 decaying W/Z boson and the PDG values of the W/Z boson masses, respectively.
915 A summary of the resolved selections is shown in Table 7.3. The distributions for
916 the variables used in the resolved analysis in the TCR are shown in Figure 7.10,
917 7.11.

918 Events classified as VBF events are classified as Merged High purity, low
919 purity or resolved signal region selections sequentially. If the event does not pass
920 any of these selections but passes a VBF control region selection it is classified as
921 a VBF CR event. If the event fails the VBF selection it is then checked if it passes
922 the Merged High purity, Low purity or resolved signal region selections (NB: for
923 the WZ decay modes all the regions have tagged and untagged categories). If the
924 event fails all the GGF signal region selections, it is then kept for GGF control
925 region selections, if it passes those selections. This cutflow is shows in Figure 7.5.

Table 7.3: The list of selection cuts in the resolved analysis for the WW and WZ signal regions (SR), $W+\text{jets}$ control region (WR) and $t\bar{t}$ control region (TR).

cuts	SR	W CR (WR)	$t\bar{t}$ CR (TR1)
$W \rightarrow \ell\nu$	Number of Tight leptons	1	
	Number of Loose leptons	0	
	E_T^{miss}	$> 60 \text{ GeV}$	
	$\cancel{p}_T(\ell\nu)$	$> 75 \text{ GeV}$	
$W/Z \rightarrow jj$	Number of small-R jets	≥ 2	
	Leading jet p_T	$> 60 \text{ GeV}$	
	Subleading jet p_T	$> 45 \text{ GeV}$	
	$Z \rightarrow q\bar{q}$ $W \rightarrow q\bar{q}$	$78 < m_{jj} < 105 \text{ GeV}$ $68 < m_{jj} < 98 \text{ GeV}$	$50 < m_{jj} < 68 \text{ GeV}$ or $105 < m_{jj} < 150 \text{ GeV}$
Topology cuts	Num. of b -tagged jets	For $Z \rightarrow jj$: ≤ 1 ($= 2$) for untagged (tagged) category	
	$\Delta\phi(j, \ell)$	> 1.0	
	$\Delta\phi(j, E_T^{\text{miss}})$	> 1.0	
	$\Delta\phi(j, j)$	< 1.5	
	$\Delta\phi(\ell, E_T^{\text{miss}})$	< 1.5	
Top vetos	$\min(p_{T,\ell\nu}, p_{T,jj}) / m_{WW}$	$> 0.35(0.25)$ for DY/ggF (VBF) category	
	Number of additional b -tagged jets	0	≥ 1
Pass VBF selection		no (yes) for DY/ggF (VBF) category	

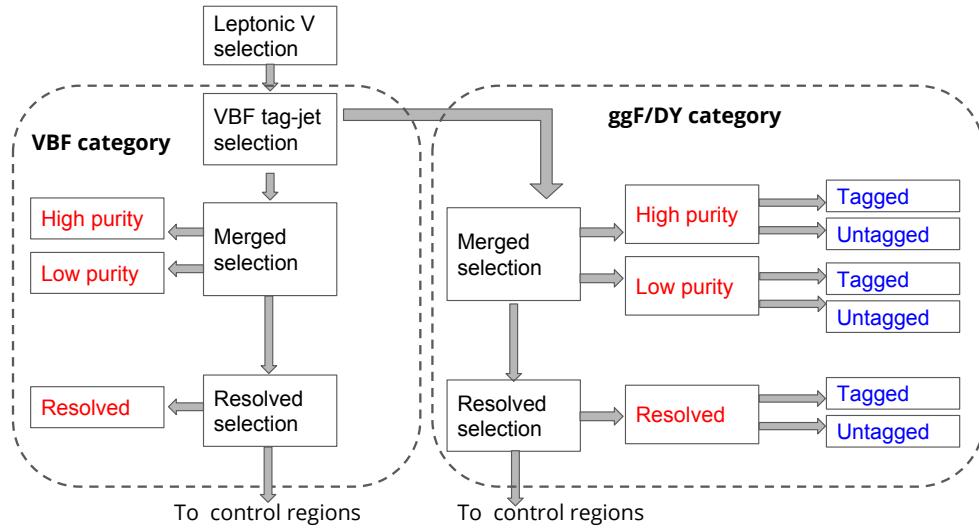


Figure 7.5: Event Categorization. Natasha write more.

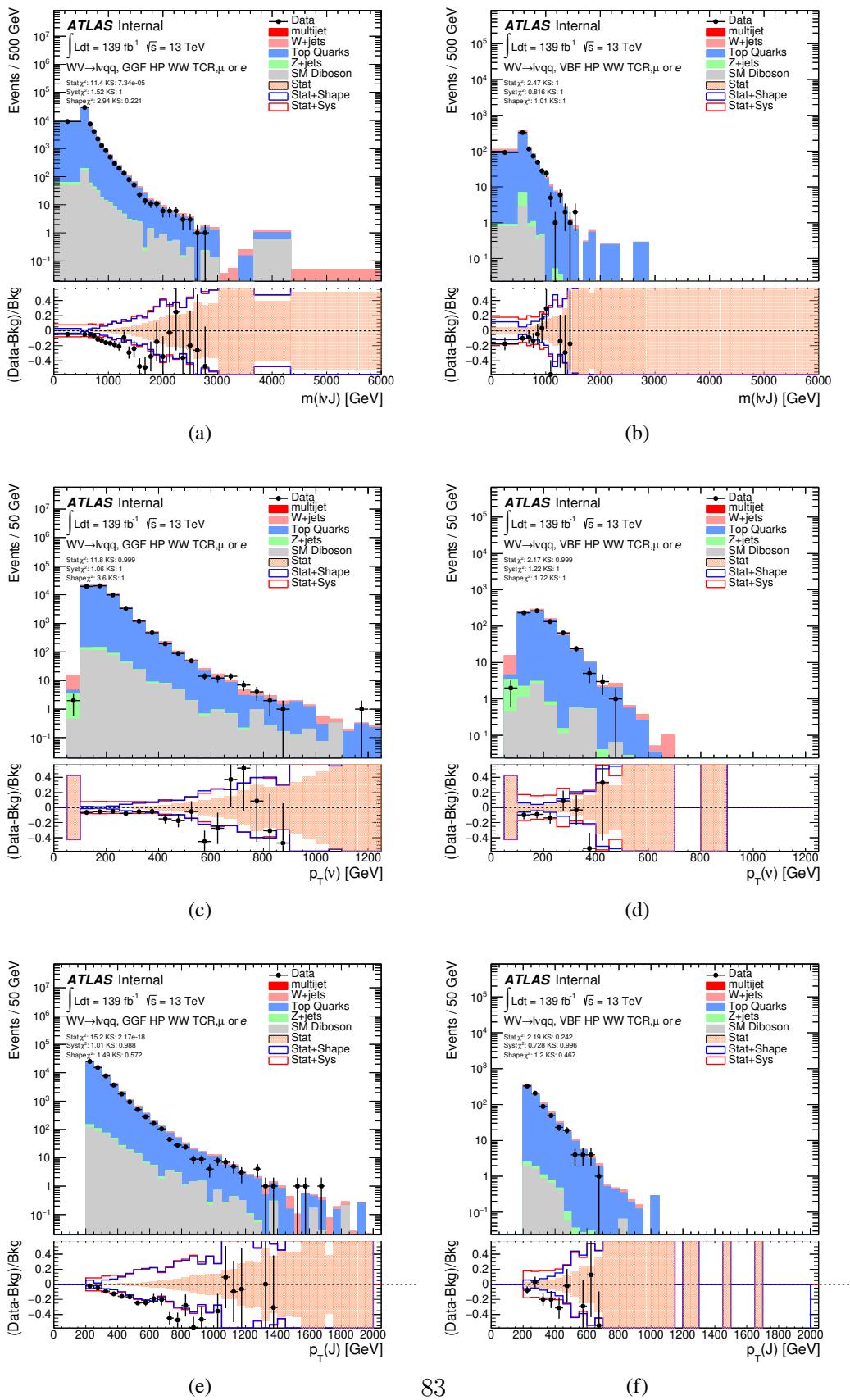


Figure 7.6: Data MC comparison for the merged WW HP TCR.

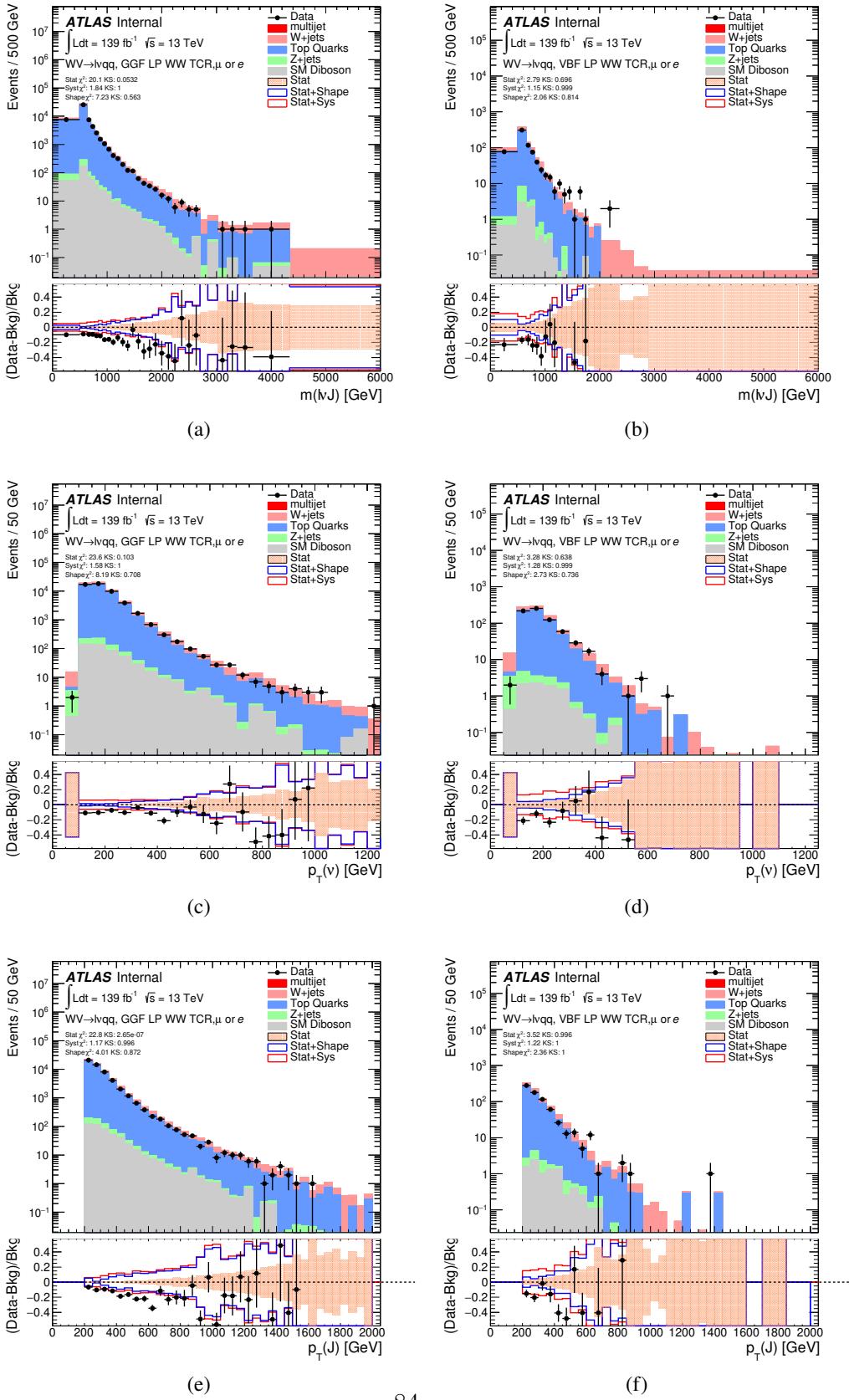


Figure 7.7: Data MC comparison for the merged WW LP TCR.

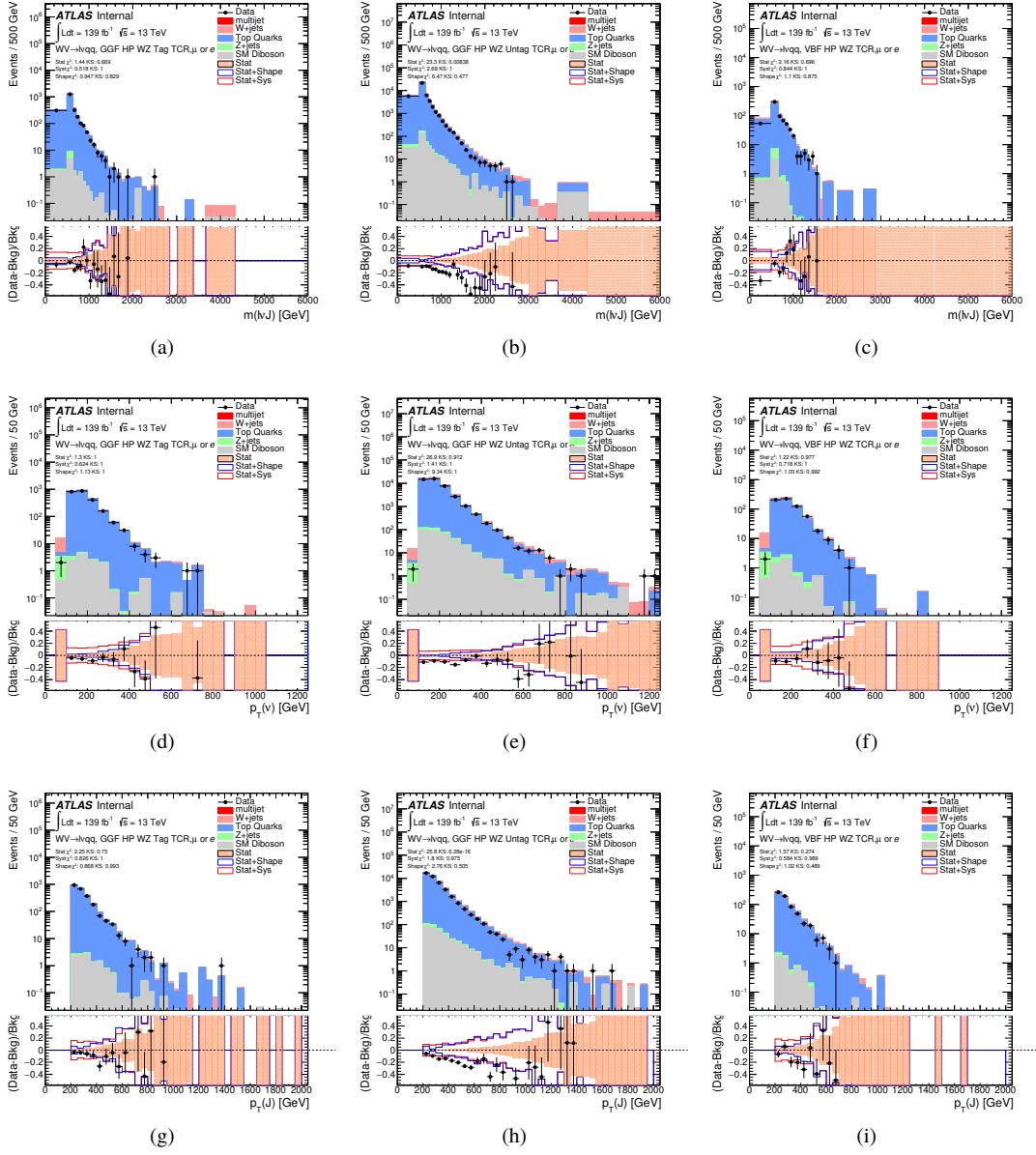


Figure 7.8: Data MC comparison for the merged WZ HP TCR.

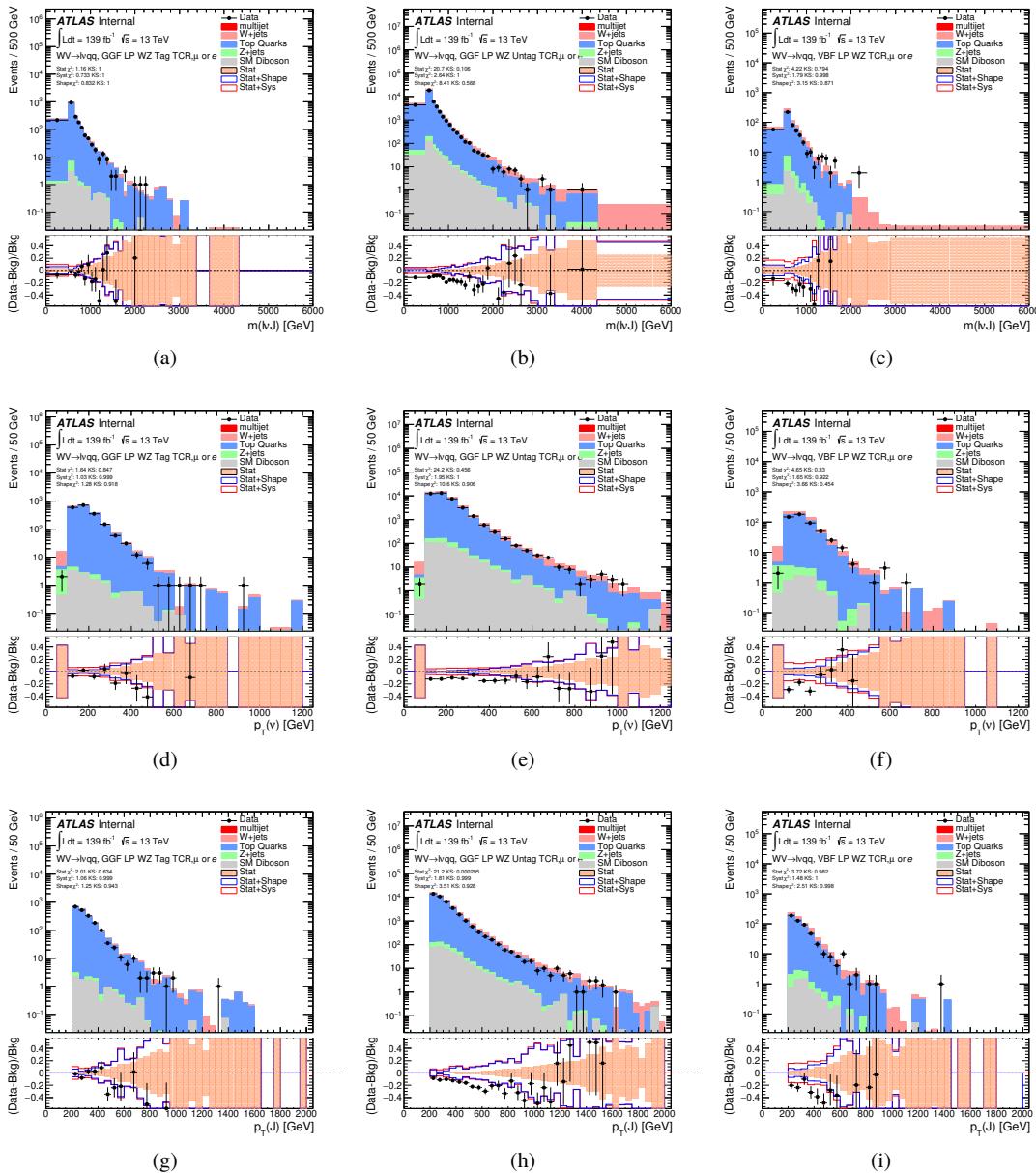


Figure 7.9: Data MC comparison for the merged WZ LP TCR.

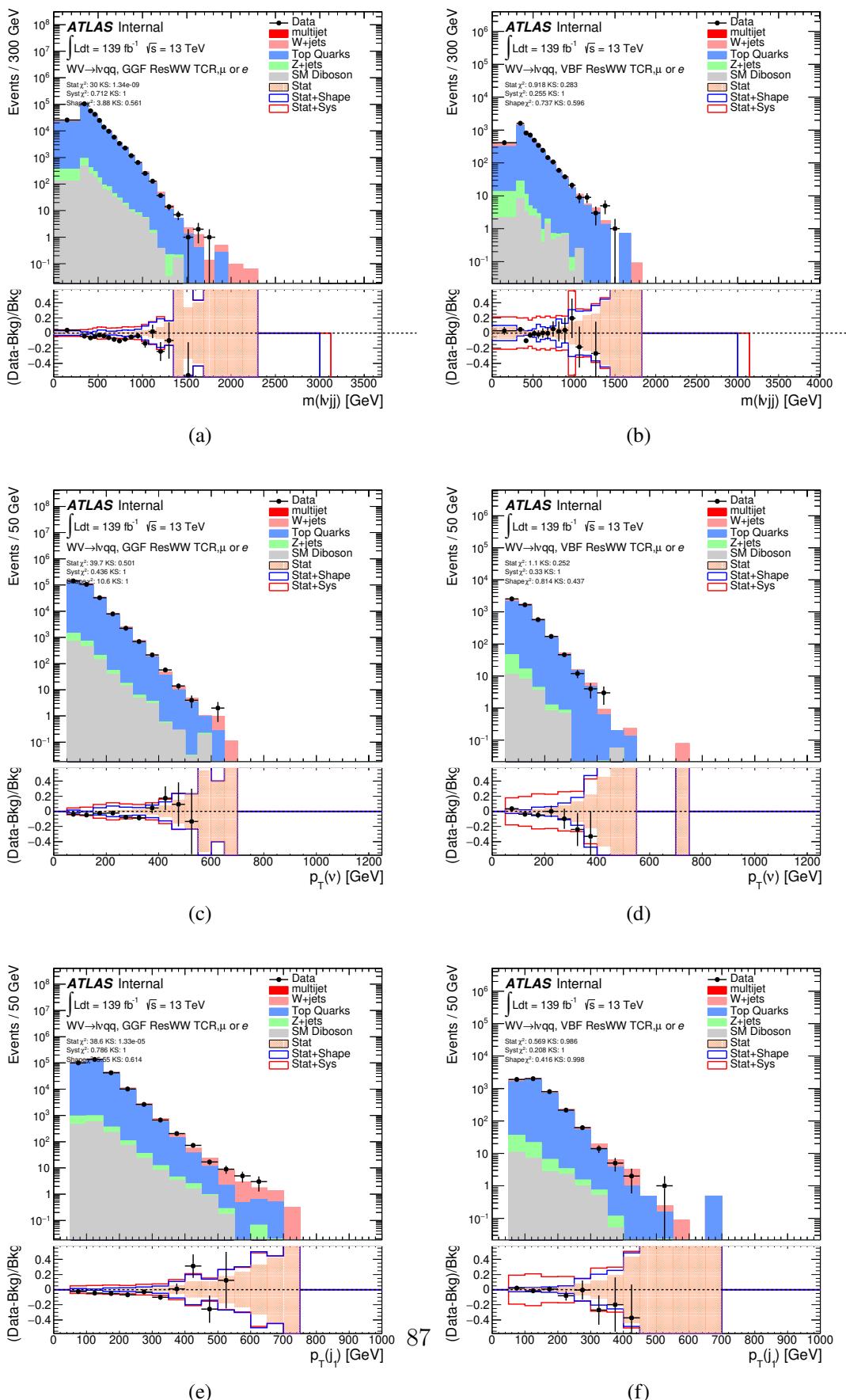


Figure 7.10: Data MC comparison for the resolved WW TCR.

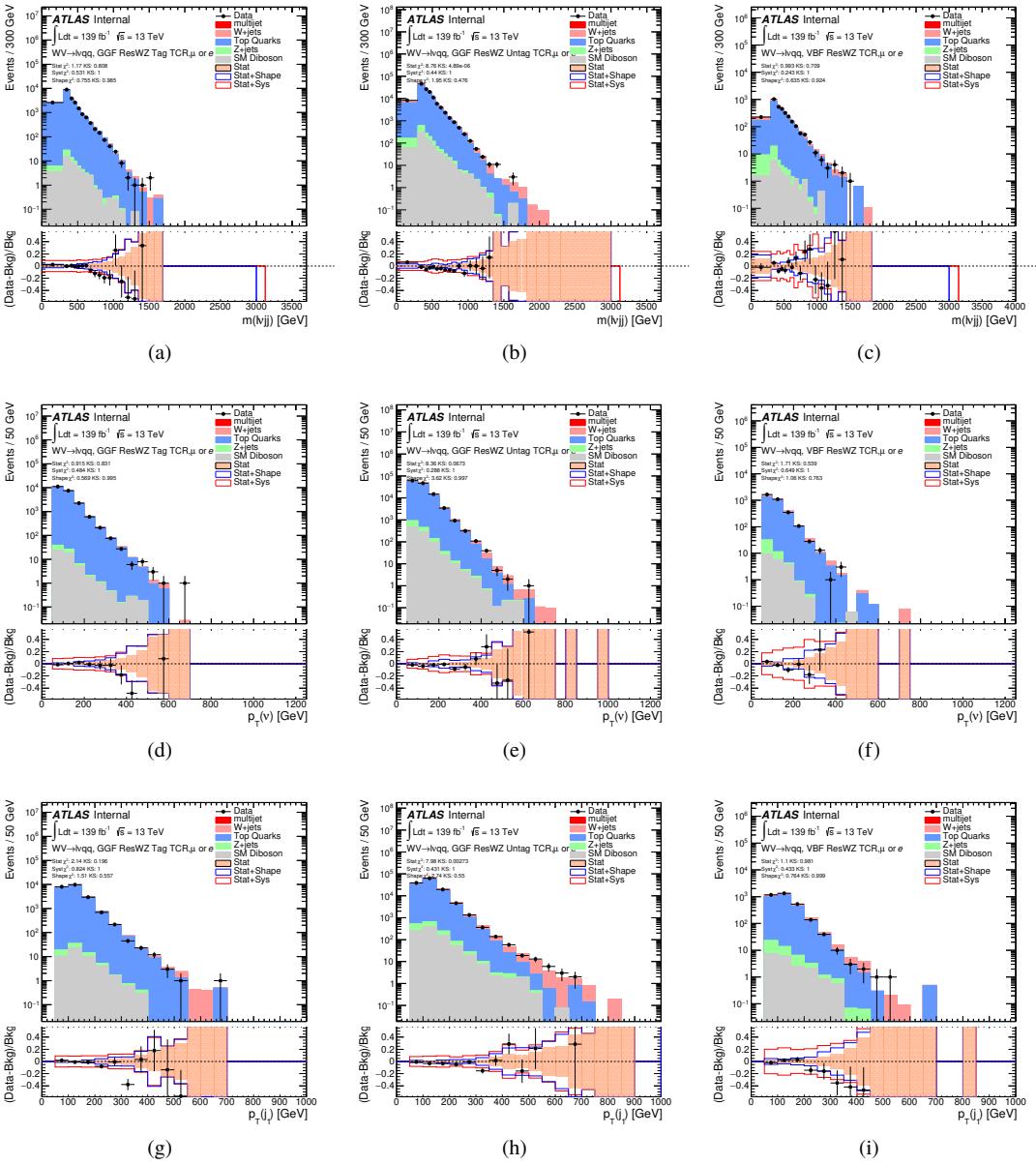


Figure 7.11: Data MC comparison for the resolved WZ TCR.

926 **7.5 Selection Acceptance times efficiency for Sig-**
927 **nal Events**

928 The acceptance times efficiency for the signal region selection is defined as:

$$A \cdot \epsilon = \frac{N_{\text{events selected}}^{\text{truth}}}{N_{\text{events generated}}^{\text{truth}}} \cdot \frac{N_{\text{events selected}}^{\text{reco}}}{N_{\text{events selected}}^{\text{truth}}} = \frac{N_{\text{events selected}}^{\text{reco}}}{N_{\text{events generated}}^{\text{truth}}} \quad (7.3)$$

929 The distributions of $A \cdot \epsilon$ as a function of the resonance mass for the different spin
930 models are shown in Figures 7.13 - ??.

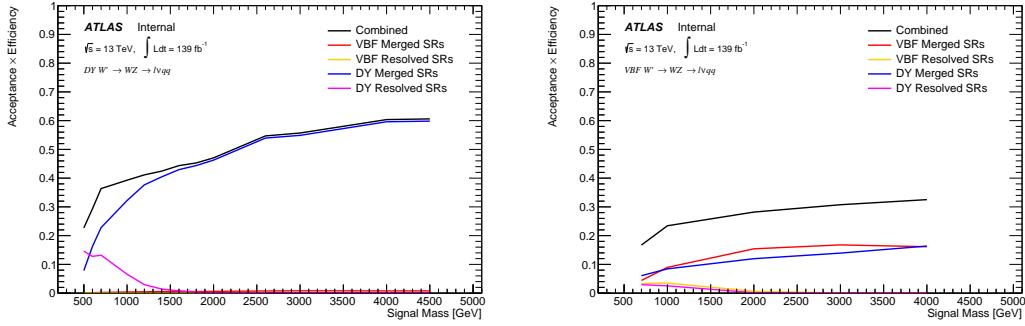


Figure 7.12: Selection acceptance times efficiency for the $W' \rightarrow WZ \rightarrow \ell\nu qq$ events from MC simulations as a function of the W' mass for (a) Drell-Yan and (b) VBF production, combining the merged HP and LP signal regions of the $WW \rightarrow \ell\nu J$ selection and the resolved regions of the $WW \rightarrow \ell\nu jj$ selection.

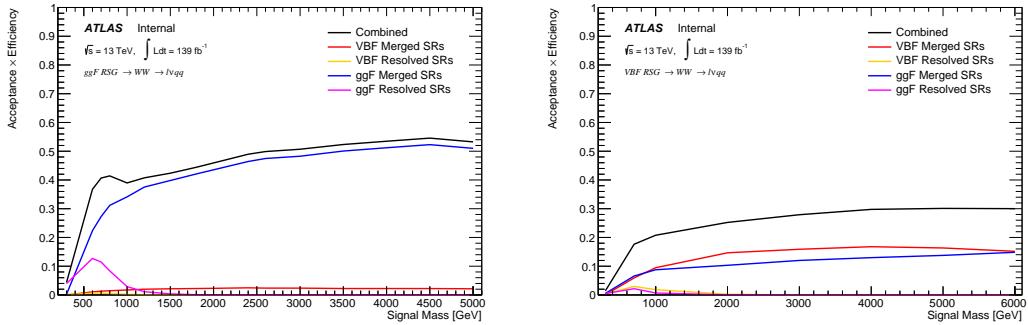


Figure 7.13: Selection acceptance times efficiency for the $G \rightarrow WW \rightarrow \ell\nu qq$ events from MC simulations as a function of the G mass for (a) Drell-Yan and (b) VBF production, combining the merged HP and LP signal regions of the $WW \rightarrow \ell\nu J$ selection and the resolved regions of the $WW \rightarrow \ell\nu jj$ selection.

931 7.6 Background Estimate

932 7.6.1 Multijet Sample

933 Backgrounds in this analysis containing real leptons (e.g. $W/Z + \text{jets}$, diboson,
 934 $t\bar{t}$, single- t) are well-modeled with simulated samples and constrained with data
 935 from CRs. However, the multijet background containing fake leptons is not well-
 936 modeled with simulation. For this reason, the multijet background is extracted
 937 from data. Heavy flavor decay products, jets, and converted photons can be
 938 mistakenly reconstructed as jets. Fake electrons often arise from jet fakes while
 939 fake muons may also arise from heavy flavor decay. For this analysis, these fake
 940 electrons generally fail the electron ID criteria and fake muons fail the muon
 941 isolation requirement. Therefore, to derive the multijet template shape the SR
 942 and CR selections and inverted lepton requirements are used as seen in Table
 943 7.4. NB: by inverting the lepton isolation/identification criteria the CR and SRs
 944 created are orthogonal to the CR and SRs.

945 The template shape of the MJ background is determined by using a multijet
 946 validation region (MJVR) that requires the inverted lepton isolation/id require-

947 ment and the two signal jets to satisfy the m_{jj} requirement used in the $W+jets$
948 CRs. The E_T^{miss} distribution in MJCR is shown in Figure 7.14 for 2017 data.
949 The template is then extracted by subtracting the data in the MJVR from the
950 electroweak background processes. This template is then added in the WCR
951 and a "pre-MJ-fit" is preformed. In this fit the E_T^{miss} distribution is fit with the
952 electroweak background normalizations constrained to expected ranges and the
953 multijet electron and muon background normalizations free to float. The fitted
954 scale factors from this MJVR template are then applied in the MJCR template.
955 The fitted uncertainties on the MJCR normalizations are then used to create the
956 MJ template in the SRs. The electron and muon background normalizations are
957 parameters in the final simulateous fit. Technically, there should be a separate
958 template for every CR and SR, but some MJ regions have insufficient statistics to
959 do this. Additionally, the shapes for the MJ templates for VBF and ggF regions
960 are found to be compatible within statistical uncertainty. Therefore, the sample
961 MJ template used for VBF and ggF CR/SRs, but with different pre-MJ-fit scale
962 factors.

963 This template method was validated using WCR and full Run 2 data. The
964 results of the fit are shown in Table 7.5. The multijet contribution in the muon
965 channel for $p_T^W > 150$ GeV is consistent with zero, and therefore neglected in
966 the final fit. Applying the extracted normalization factor to MJVR in WCRs for
967 various kinematic variables such as E_T^{miss} , W transverse mass, lepton p_T , and the
968 invariant mass as show in Figures 7.15 -7.24. These figures show good agreement
969 between the data and background estimate.

970

	Criterion	signal lepton	inverted lepton
Electron	ID	TightLH	MediumLH !TightLH
	Calo Isolation	FixedCutHighPtCaloOnlyIso	FixedCutHighPtCaloOnlyIso
Muon	ID	WHSignalMuon	WHSignalMuon
	Track Isolation	FixedCutTightTrackOnlyIso	!FixedCutTightTrackOnlyIso $ptvarcone30/pt < 0.07^*$
*Only applied to events with $pTW < 150\text{GeV}$			

Table 7.4: Definitions of “inverted” leptons used in multijet control region

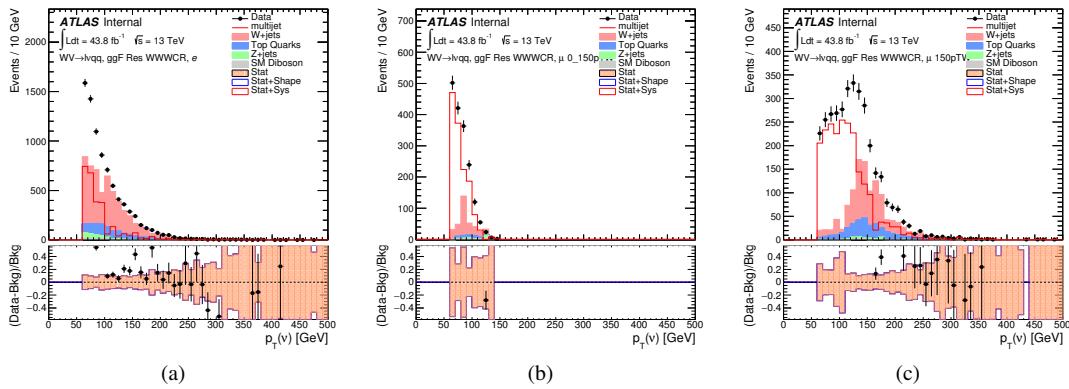


Figure 7.14: The E_T^{miss} distribution in MJCR for 2017 data in the electron channel(left), muon channel with W-boson $pT < 150$ GeV (center) and > 150 GeV (right). Multi-jet templates are calculated as remaining data components after excluding known MC

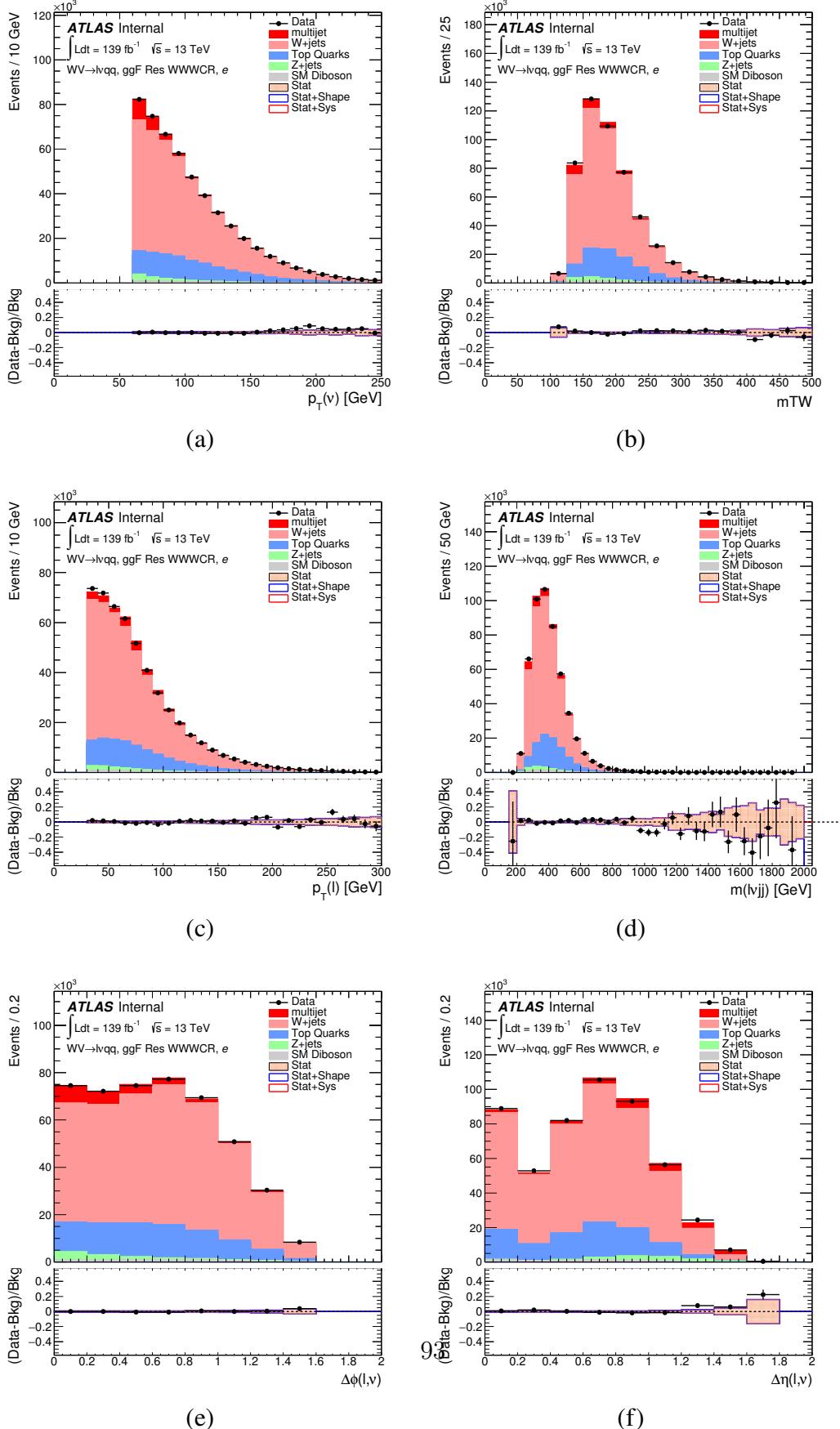


Figure 7.15: Postfit Data/MC comparison of distributions of E_T^{miss} , m_T^W , lepton

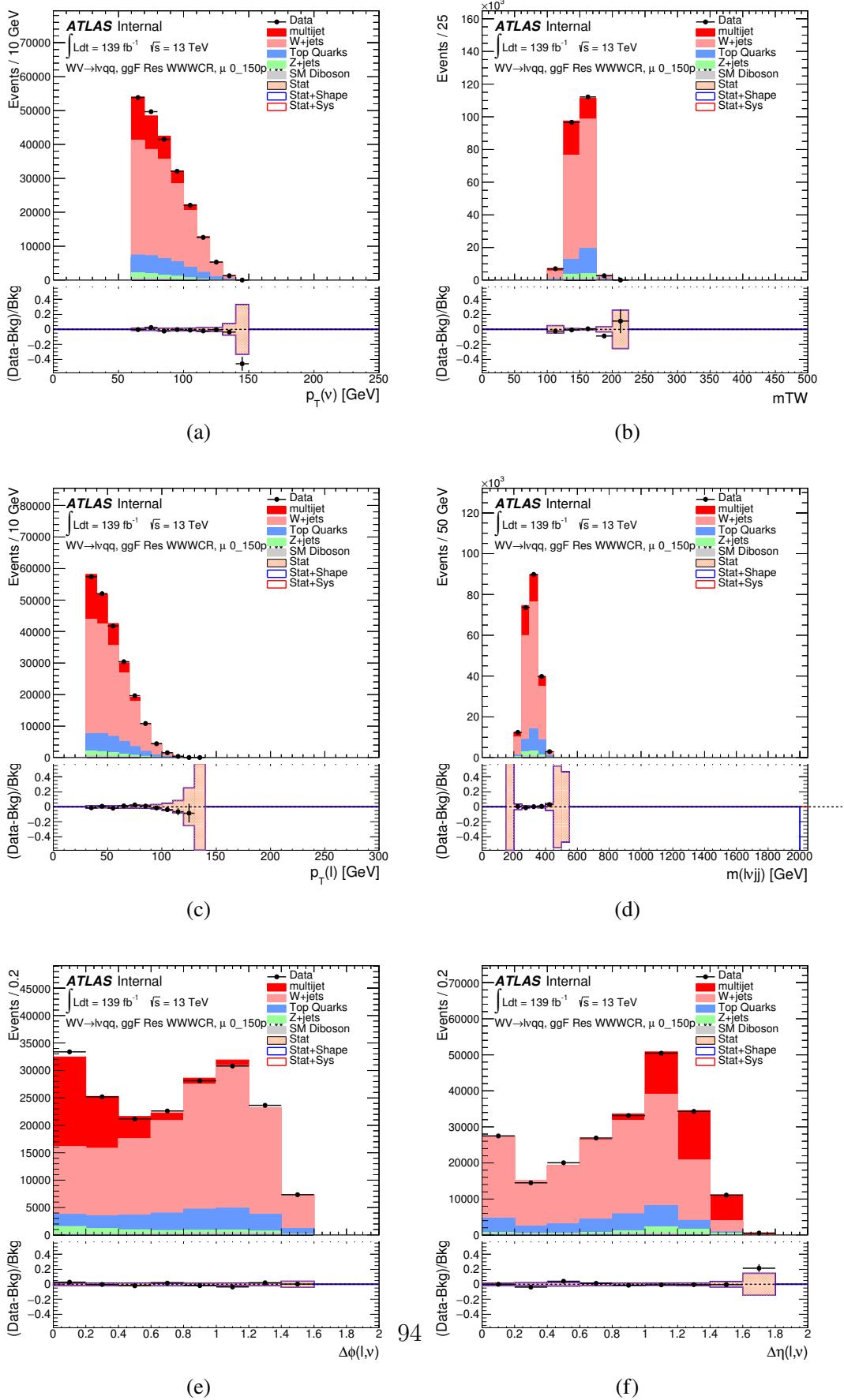
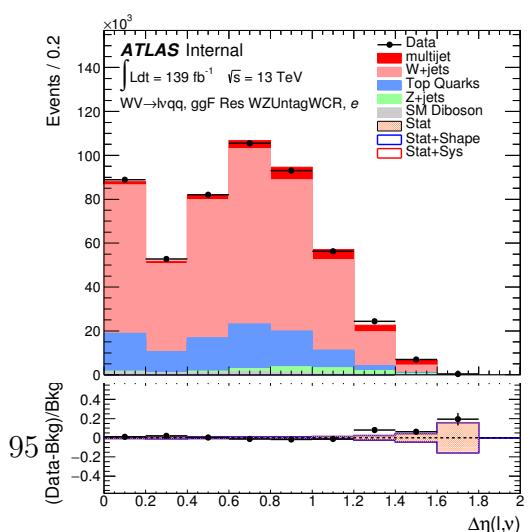
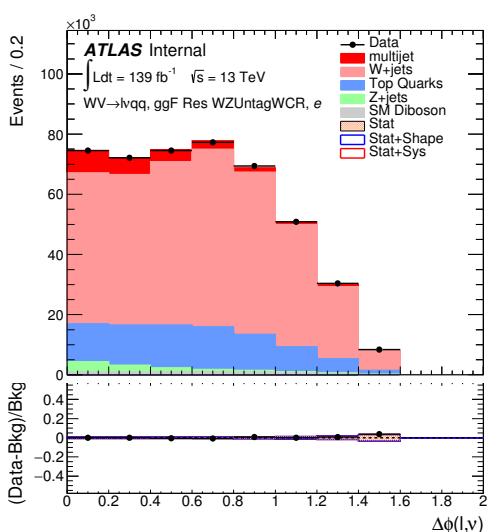
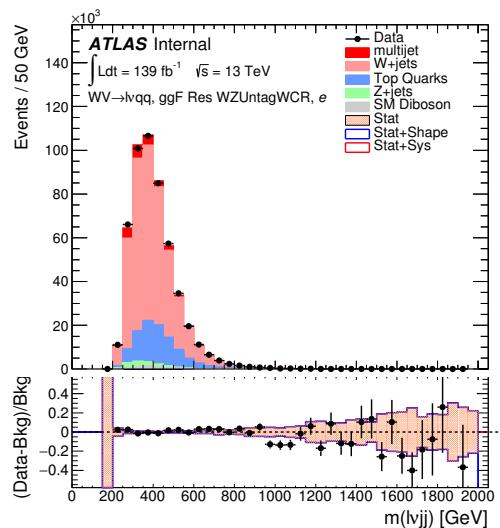
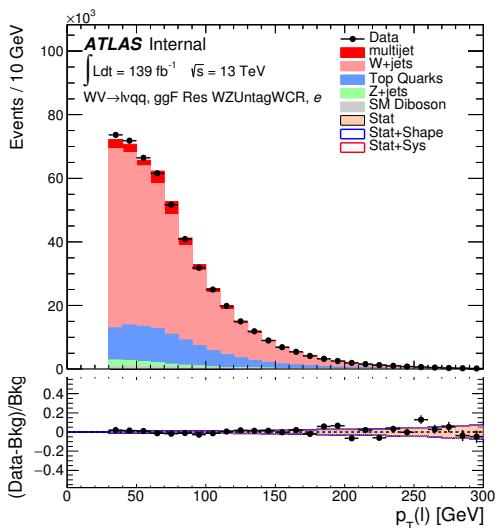
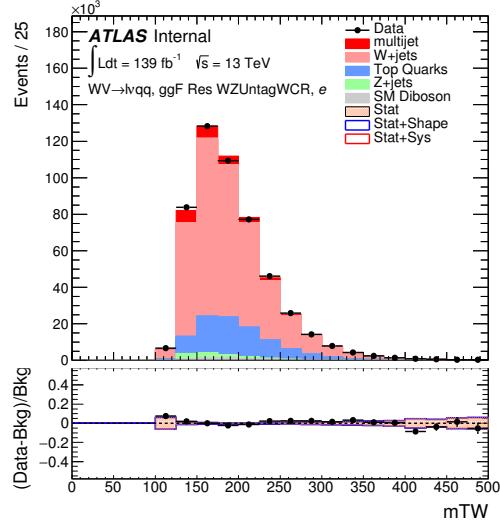
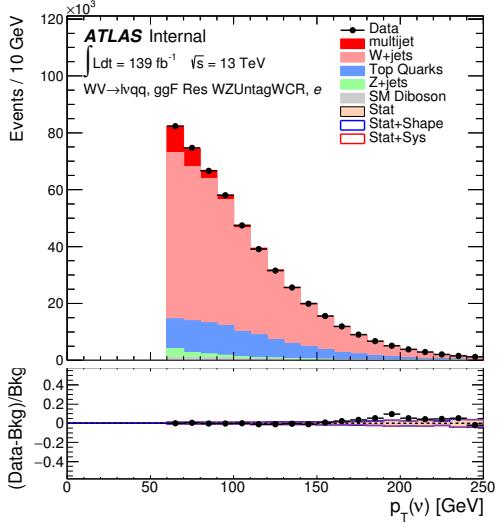


Figure 7.16: Postfit Data/MC comparison of distributions of E_T^{miss} , m_T^W , lepton and neutrino p_T , $m_{l\nu jj}$, lepton- ν angular distance in the WW muon channel. The



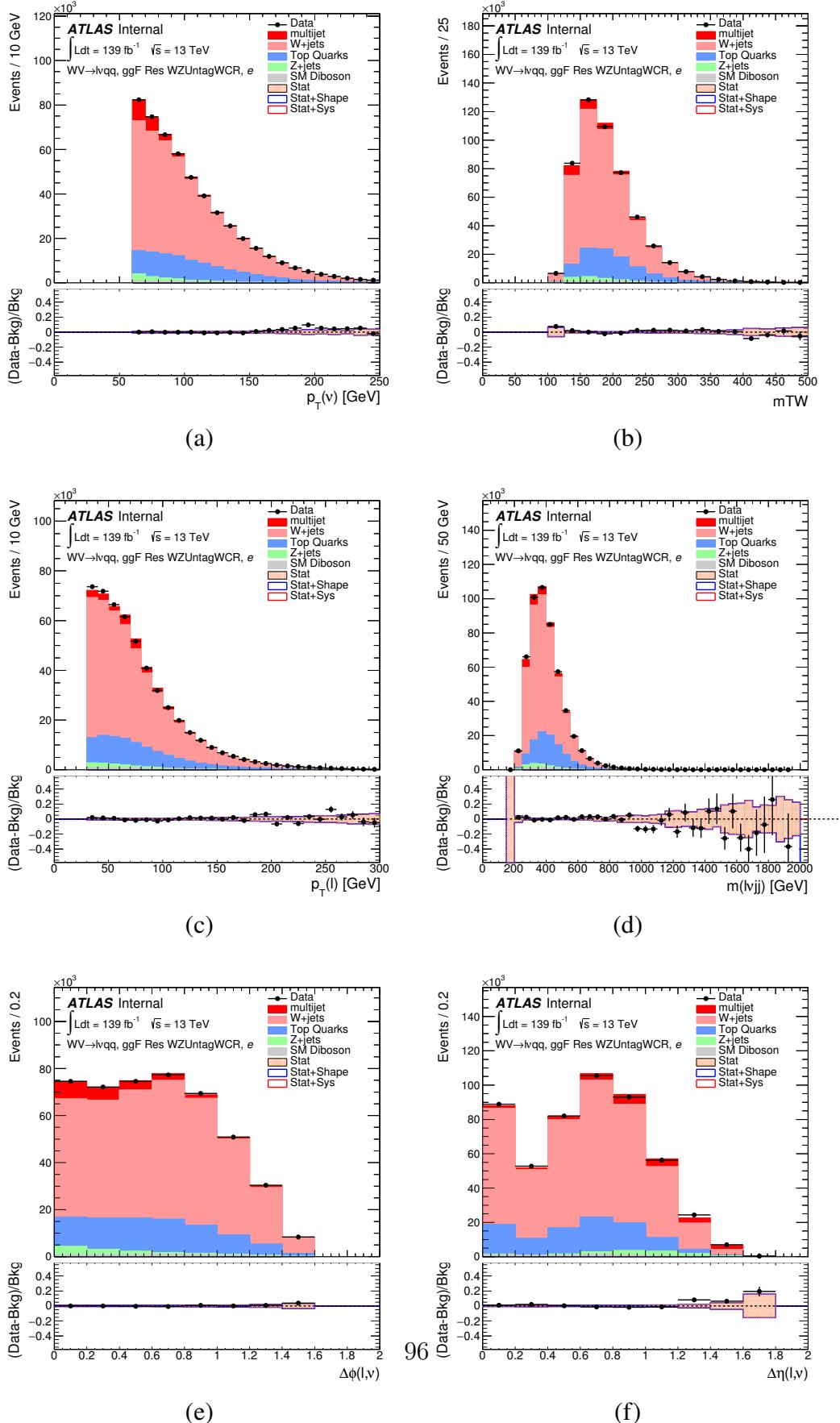
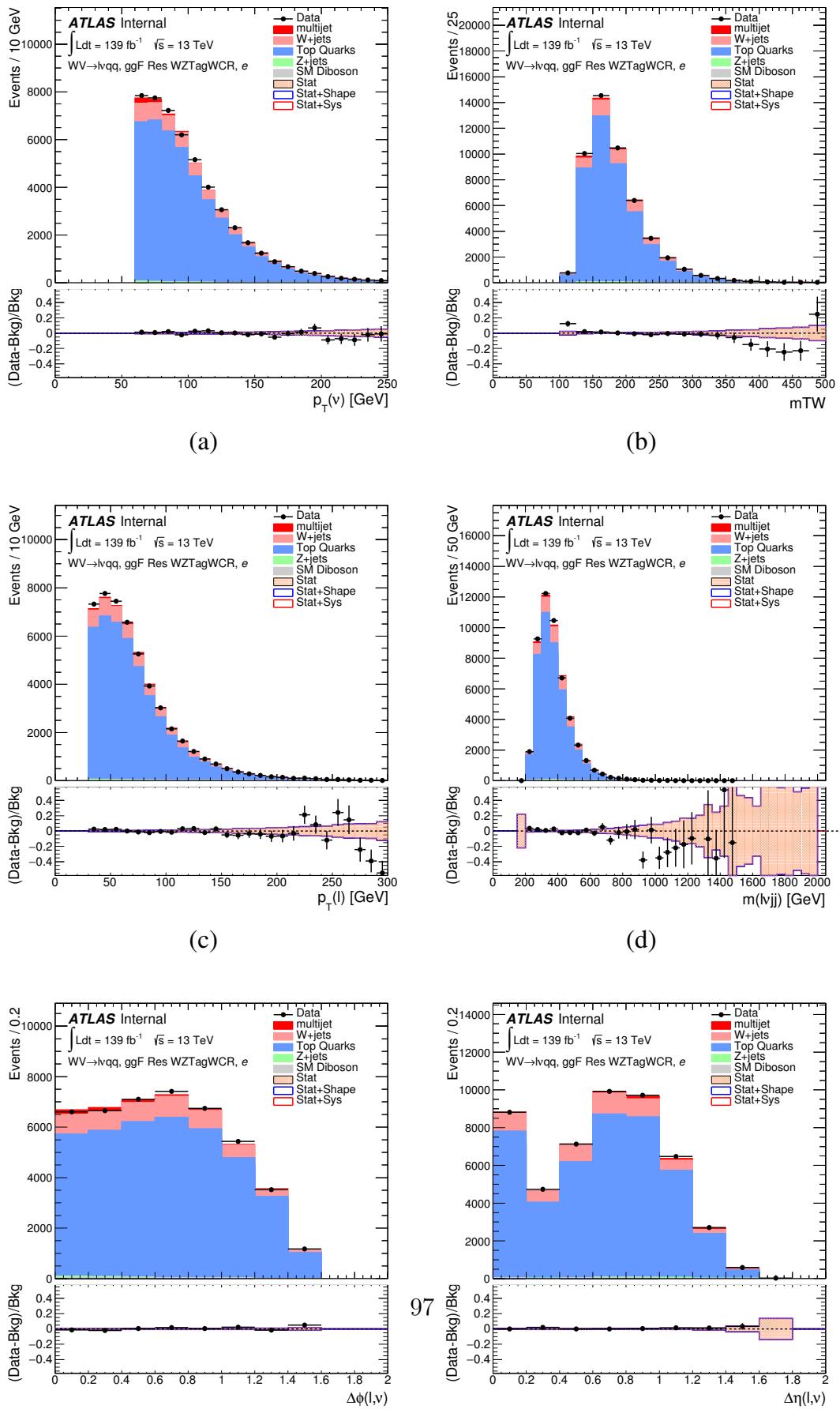
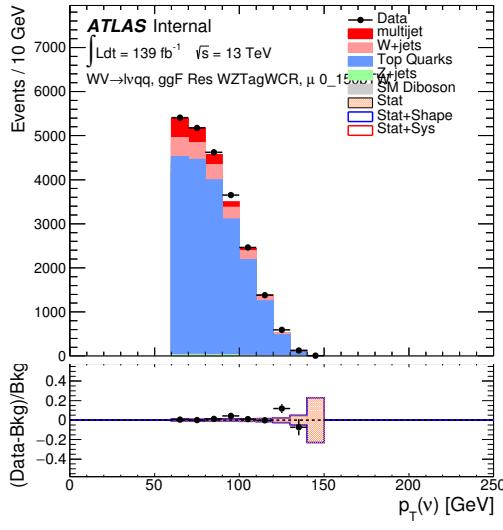
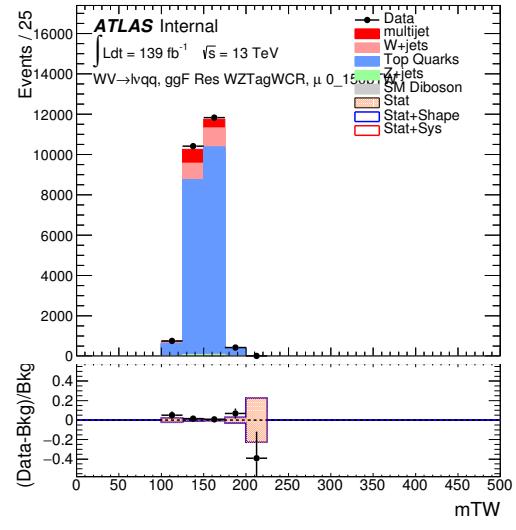


Figure 7.18: Postfit Data/MC comparison of distributions of E_T^{miss} , m_T^W , lepton

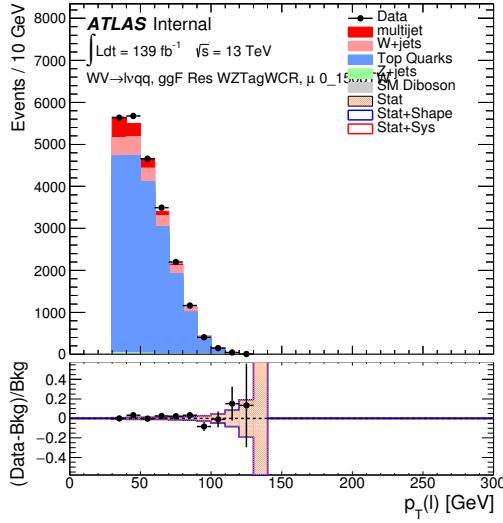




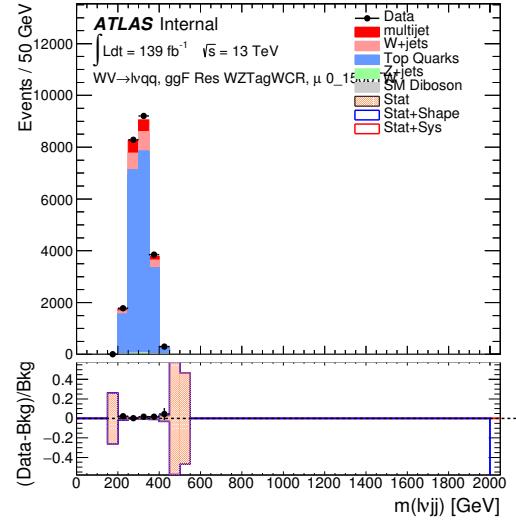
(a)



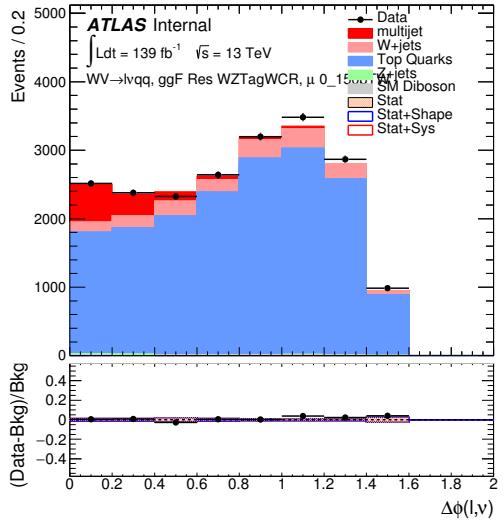
(b)



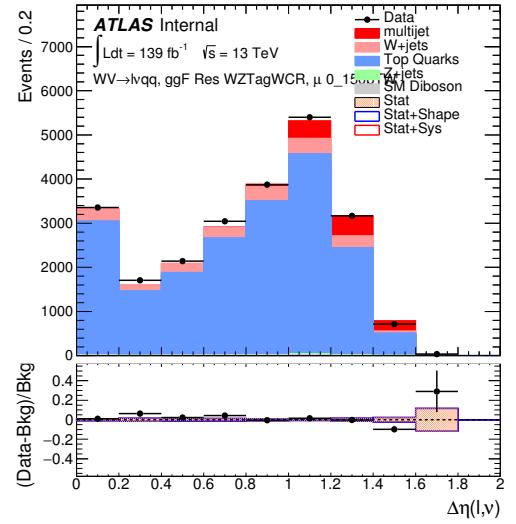
(c)



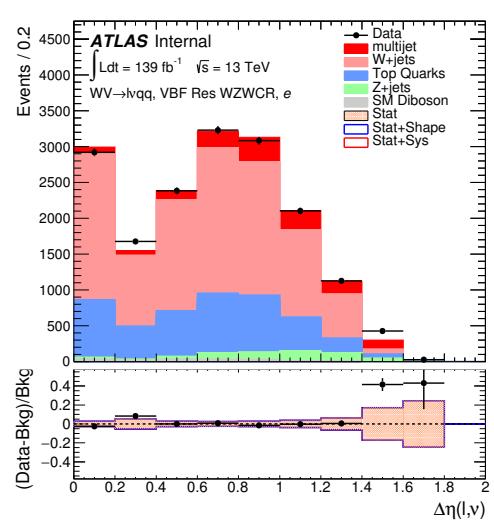
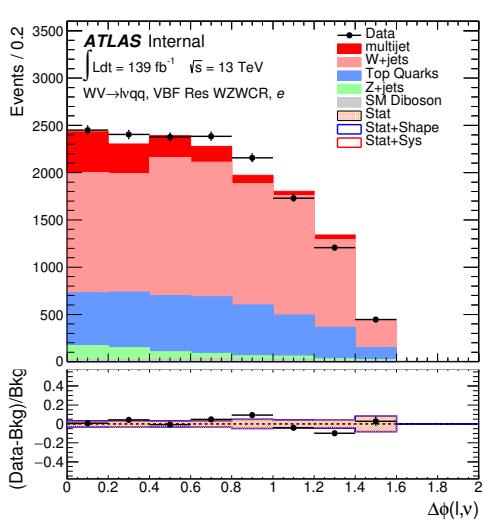
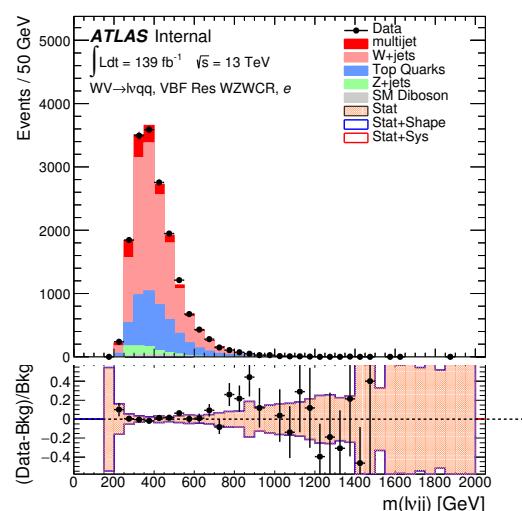
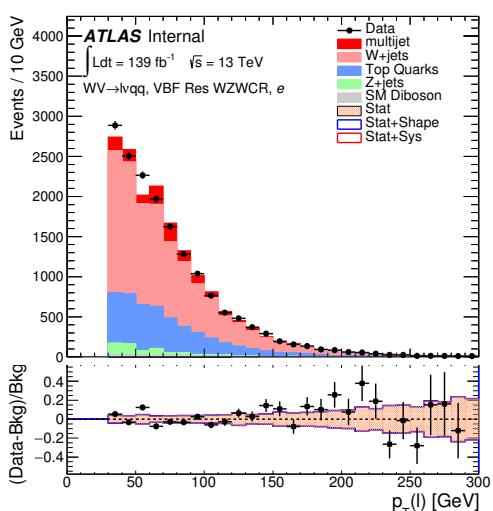
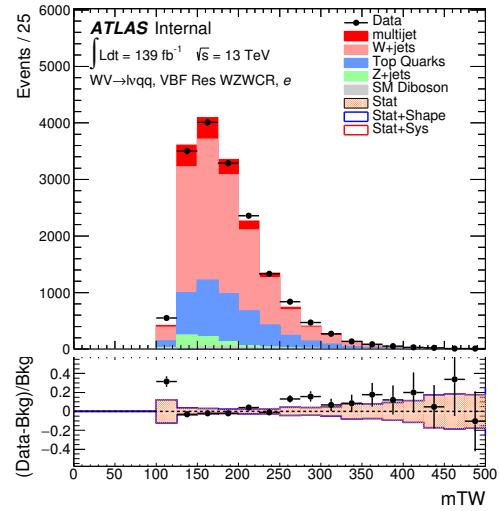
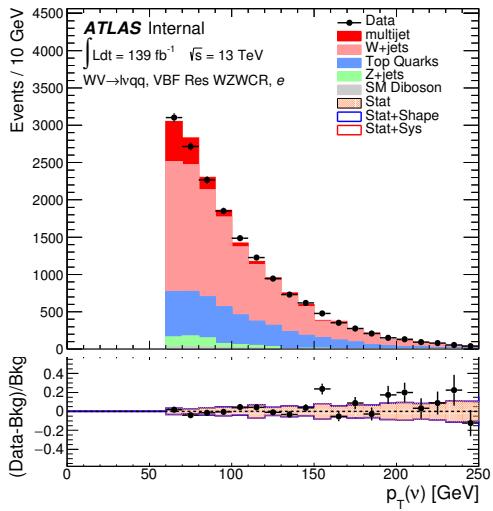
(d)

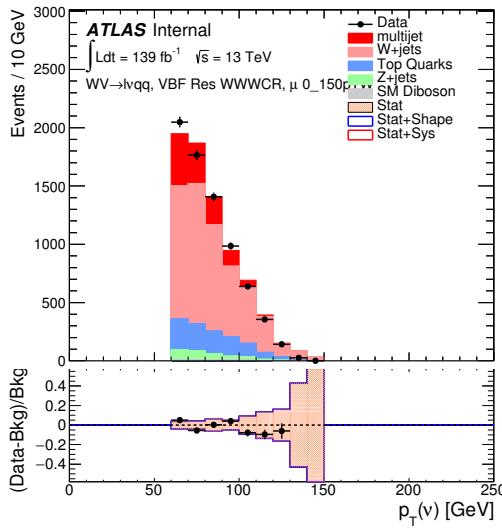


(e)

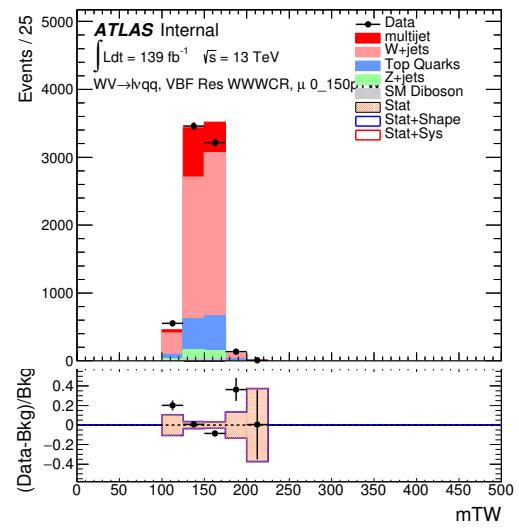


(f)

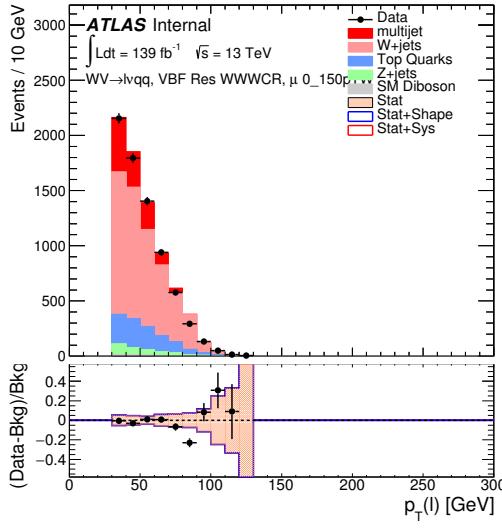




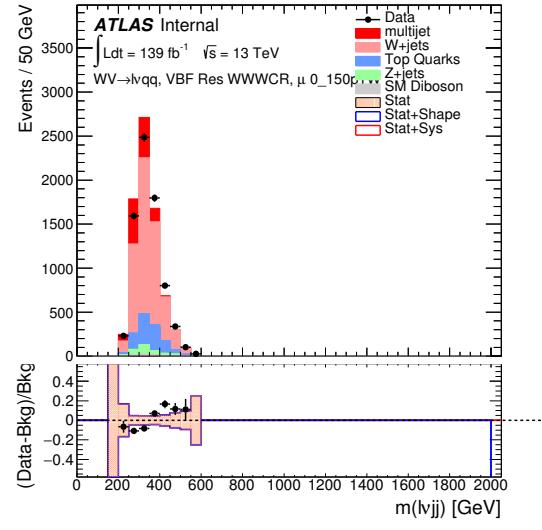
(a)



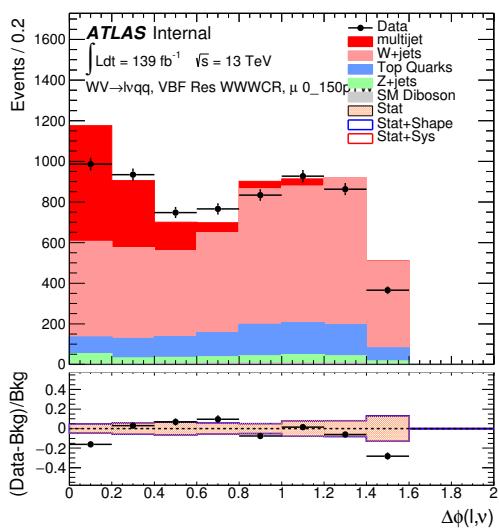
(b)



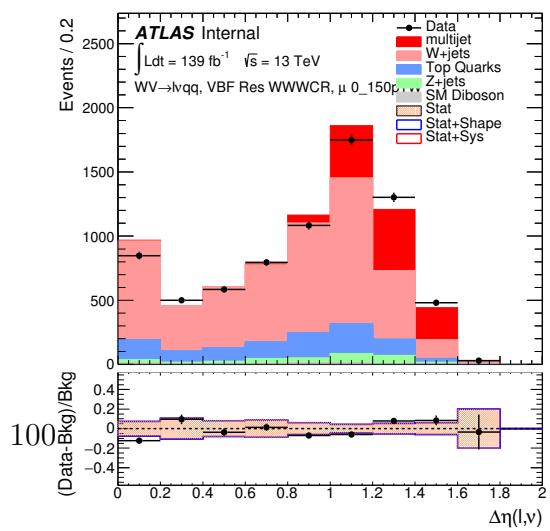
(c)



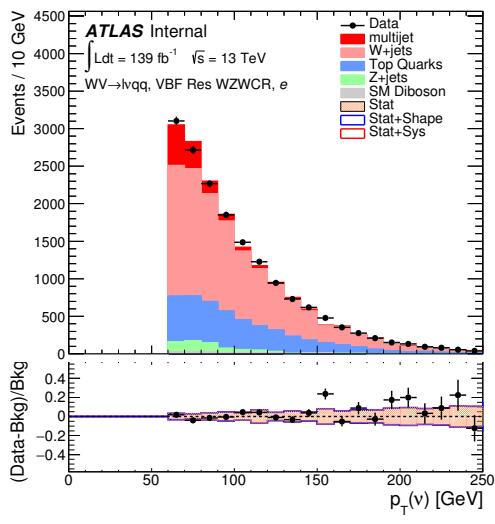
(d)



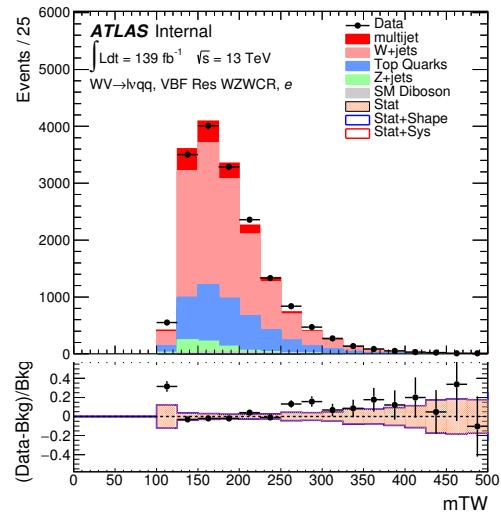
(e)



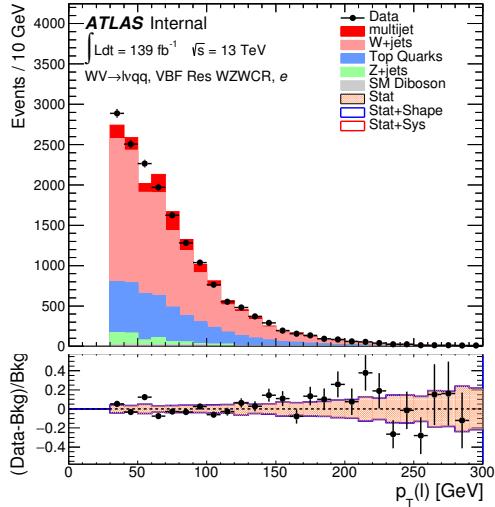
(f)



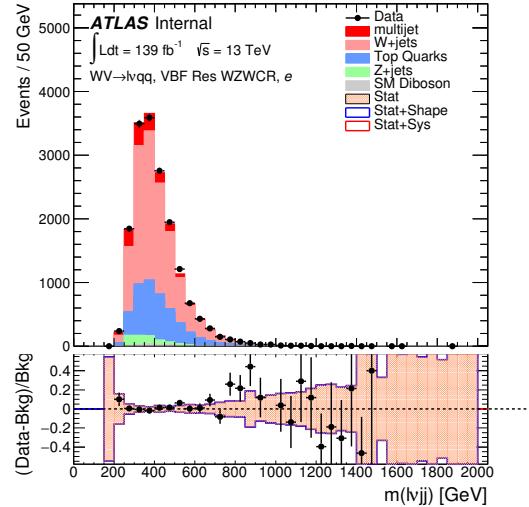
(a)



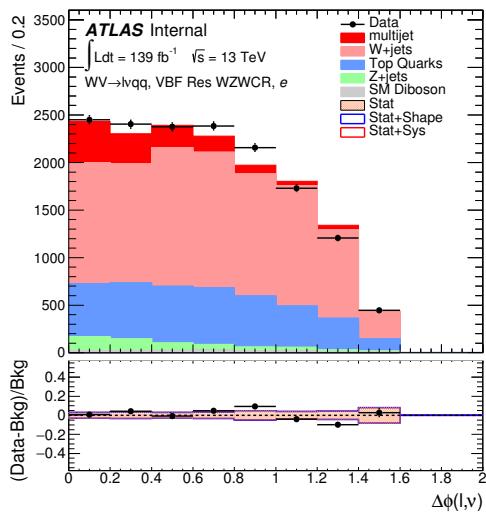
(b)



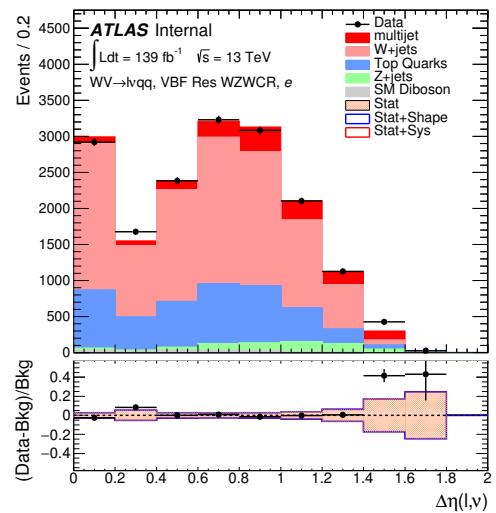
(c)



(d)



(e)



(f)

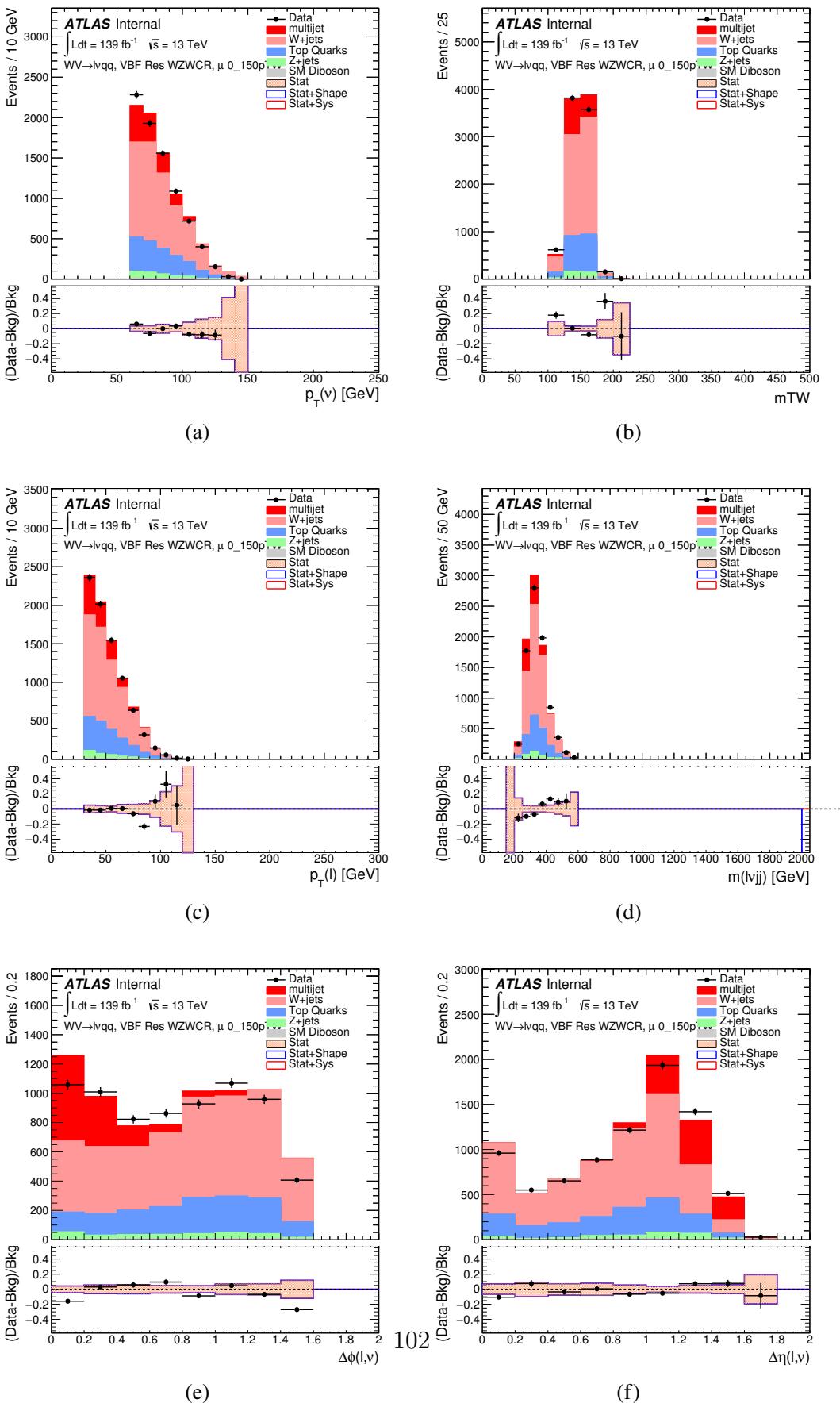


Figure 7.24: Postfit Data/MC comparison of distributions of E_T^{miss} , m_T^W , lepton

Full Run 2
ggF Res WWWCR

Sample	Yield	R.U.	SF
Top&W	645040 ± 1971.68	0.31%	0.998
Z&VV	24075.9		fixed
MJ_el	24156.3 ± 1224.62	5.06%	3.973
MJ_mu	35528.5 ± 923.94	2.60%	9.019

ggF Res WZ01bWCR

Sample	Yield	R.U.	SF
Top&W	644690 ± 1981.4	0.31%	0.997
Z&VV	24075.9		fixed
MJ_el	24366.5 ± 1232.69	5.05%	3.874
MJ_mu	35528.5 ± 921.27	2.58%	8.746

ggF Res WZ2bWCR

Sample	Yield	R.U.	SF
Top&W	71236.5 ± 688.74	0.97%	1.031
Z&VV	518.5		fixed
MJ_el	595.63 ± 449.34	75.44%	0.094
MJ_mu	1196.9 ± 222.13	18.56%	0.294

VBF Res WWWCR

Sample	Yield	R.U.	SF
Top&W	19032.3 ± 364.43	1.91%	0.928
Z&VV	1091.63		fixed
MJ_el	1425.73 ± 214.42	15.03%	0.235
MJ_mu	1281.36 ± 157.21	11.83%	0.314

VBF Res WZWCR

Sample	Yield	R.U.	SF
Top&W	21341.8 ± 392.21	1.84%	0.942
Z&VV	1111.75		fixed
MJ_el	1413.76 ± 230.36	16.29%	0.225
MJ_mu	1281.36 ± 157.21	12.27%	0.314

Table 7.5: Fit validation result in WCRs for 2015+16 data. The fit is done in various WCRs, in order to obtain the corresponding scale factors for MJ templates: ggF resolved WCR for the $WW \rightarrow lvqq$ selection, ggF resolved untagged WCR for the $WZ \rightarrow lvqq$ selection, ggF resolved tagged WCR for the $WZ \rightarrow lvqq$ selection, VBF resolved WCR for the $WW \rightarrow lvqq$ selection, and VBF resolved WCR for the $WZ \rightarrow lvqq$ selection. Post-fit event yields for electroweak processes and MJ contributions are shown. The SF column shows the corresponding normalization scale factors for electroweak processes from the fit. R.U. stands for relative uncertainty.

₉₇₁ **Chapter 8**

₉₇₂ **Systematic Uncertainties**

₉₇₃ This section describes the sources of systematic uncertainties considered in
₉₇₄ this analysis. These uncertainties are divided into three categories: experimental
₉₇₅ uncertainties, background modeling uncertainties, and theoretical uncertainties on
₉₇₆ signal processes. In the statistical analysis each systematic uncertainty is treated
₉₇₇ as a nuisance parameter estimated on the m_{VV} distribution.

₉₇₈ **8.1 Experimental Systematics**

₉₇₉ The uncertainty on the integrated luminosity of the dataset used is 1.7% and
₉₈₀ a systematic in the final fit. The luminosity uncertainty is calculated following a
₉₈₁ methodology similar to the one in [ref P55].

₉₈₂ Also, multiple pile up interactions are simulated to match data conditions.
₉₈₃ This ensures simulated detector response and particle reconstruction conditions
₉₈₄ are as similar as possible. The distribution of the average number of interactions
₉₈₅ per bunch crossing applied to simulation is called the μ profile. The pileup mod-
₉₈₆ eling uncertainty is accounted for by re-weighting simulated events so the average
₉₈₇ number of interactions per bunch crossing varies within its uncertainty due to

988 systematics from vertex reconstruction [ref ATL-COM-SOFT-2015-119]. The as-
989 sociated re-weighting factors are propagated through the entire analysis chain to
990 construct a systematic uncertainty on m_{VV} .

991 The single-lepton and E_T^{miss} triggers used are not fully efficient and therefore
992 simulated data must be scaled to account for trigger inefficiencies. Trigger effi-
993 ciencies are given by the ratio of the distribution of offline objects before trigger
994 selection and after trigger selection.

995 Uncertainties on small-R jet energy scale and resolution are measured in-situ
996 by calculating the response between data and simulation. This analysis uses a
997 reduced set of JES and JER uncertainties (totaling 30 and 8 systematics, re-
998 spectively). These reduced sets of systematics are calculated using a principal
999 component analysis, yield largely uncorrelated independent systematics. These
1000 uncertainties account for the dependence on p_T , η , μ , flavor response and global
1001 sequential corrections. Systematic uncertainties associated with b -tagging are also
1002 considered. These systematics are evaluated as uncertainties on the scale factor
1003 which account for the difference in b -tagging efficiencies in data and MC, and the
1004 flavor dependence (between b, c, and light jets).

1005 The uncertainty on the p_T scale of the large-R jets is determined by comparing
1006 the jet's p_T^{calo} to p_T^{track} in di-jet simulation and data. In addition to this uncertain-
1007 ties from tracking, modeling (Pythia vs Herwig), and statistical constraints are
1008 also calculated. The large-R jet p_T resolution is given by smearing the jet p_T with
1009 a Gaussian with a 2% width.

1010 The W/Z -tagging efficiency cannot be evaluated using the Rtrk method as the
1011 TCC algorithm uses track measurements to reconstruct jet substructure variables.
1012 In order to avoid this potential bias, the W/Z -tagging estimated in data using a
1013 control sample and correct by comparing it with simulation. The efficiency to

1014 W/Z -induced signal is estimated by a $t\bar{t}$ control sample, while the efficiency to
1015 single- q/g background is estimated using a dijet sample. The effects of experimen-
1016 tal and theoretical uncertainties on the efficiency scale factor are by taking the
1017 ratio of efficiencies in data and simulation. By taking this ratio the uncertainties
1018 not arising for jet mass and D_2 cancel.

1019 Lepton identification, reconstruction, isolation systematic uncertainties are de-
1020 termined by reconstructing the Z mass peak with a tag and probe method. The
1021 lepton energy and momentum scales are also measured with the Z mass peak.
1022 Additionally, the track-to-vertex association efficiency is used for muons.

1023 As E_T^{miss} is calculated using all the physics objects in the event, all those objects
1024 associated errors result in an uncertainty on E_T^{miss} . Additionally, the unassociated
1025 tracks used to construct E_T^{miss} contribute to the uncertainty on E_T^{miss} .

1026 8.2 Theory Systematics

1027 Theoretical uncertainties for signal and background processes arise from un-
1028 certainties in the parameters used in Monte Carlo simulation. In particular for
1029 the $t\bar{t}$, $W/Z+jets$, and diboson backgrounds and signal samples the QCD scale,
1030 PDF, generator and hadronization uncertainties were evaluated. To assess the
1031 QCD scale uncertainty the renormalization and factorization scales were scaled
1032 up (2.0) and down (0.5) at the event generation stage of sample production. Un-
1033 certainties due to the choice of the parton distribution functions were evaluated
1034 by re-weighting samples from the nominal PDF to a set of error PDFs which ac-
1035 count for the uncertainty of the fits used to produce the PDF set. In addition to
1036 this samples are re-weighted to different PDF sets to account for the arbitrariness
1037 of the PDF choice. The difference between the m_{WV} distributions using differ-
1038 ent event generators is assessed by comparing samples generated with different

1039 generators. Similarly, the uncertainty in hadronization models is account for by
1040 comparing samples created using different hadronization models (e.g. Pythia8 vs.
1041 Herwig7). Figures 8.2 - 8.8 show the impact of these uncertainties on the $t\bar{t}$ and
1042 $W/Z + \text{jets}$ backgrounds.

1043 Additionally, contributions to the diboson background for the VBF analysis
1044 were included in [SOME WAY that is not determined yet].

1045 The normalization of the $t\bar{t}$ and $W+\text{jets}$ processes impact the multijet template
1046 shape. The impact of these normalization is assess by including a shape systematic
1047 on the multijet background from varying the $t\bar{t}$ and $W+\text{jets}$ normalization factors.

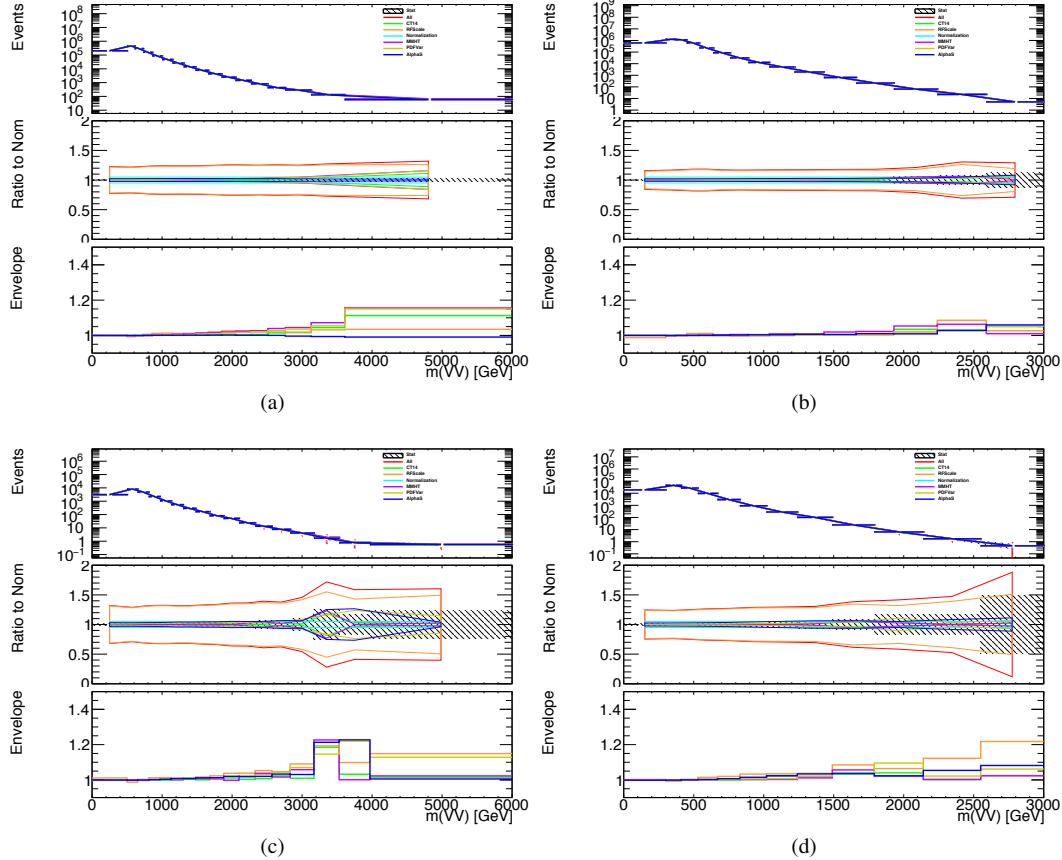


Figure 8.1: The $W/Z + \text{jet}$ systematics for the a) Merged ggF, b) Resolved ggF, c) Merged VBF, and d) Resolved VBF regions. The top subplot shows the nominal and variation distributions/bands, the middle shows the ratio of the two, and the final shows just the shape of the envelope (the final uncertainty).

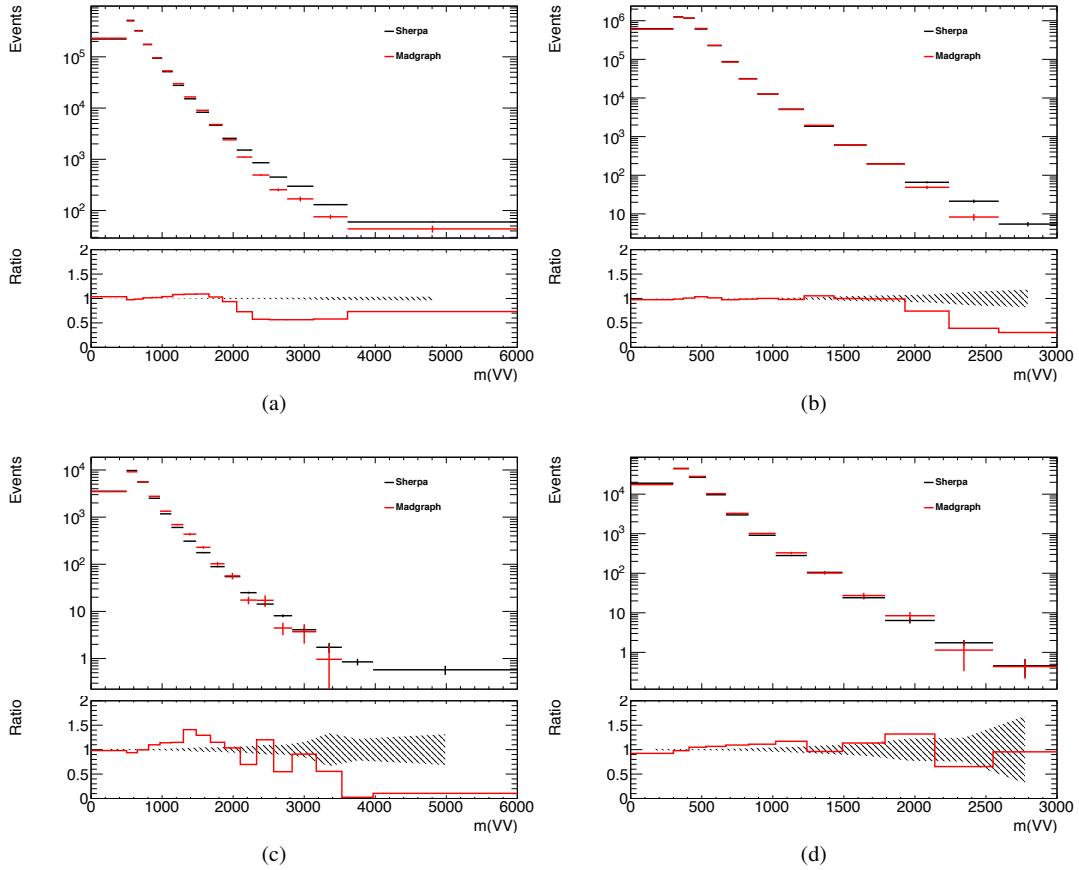


Figure 8.2: The two-point generator comparison between Sherpa and MadGraph for the $W/Z + \text{jet}$ samples in the a) Merged ggF, b) Resolved ggF, c) Merged VBF, and d) Resolved VBF regions. The normalization of the Madgraph sample is set to the Sherpa value to consider only shape effects. The bottom inset shows the ratio of the two.

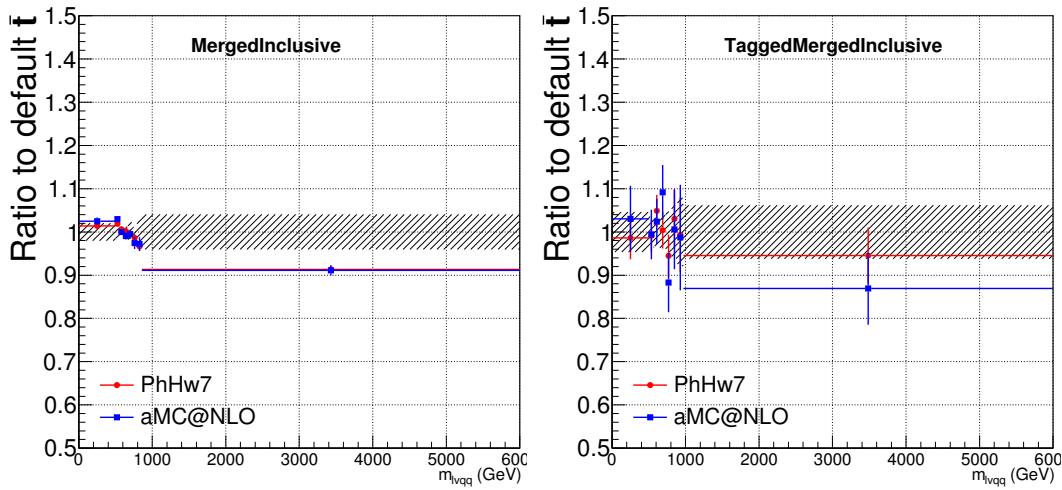


Figure 8.3: Ratio between the variations of generator (red) and hadronization (blue) variations for the Merged regime for $t\bar{t}$ sample.

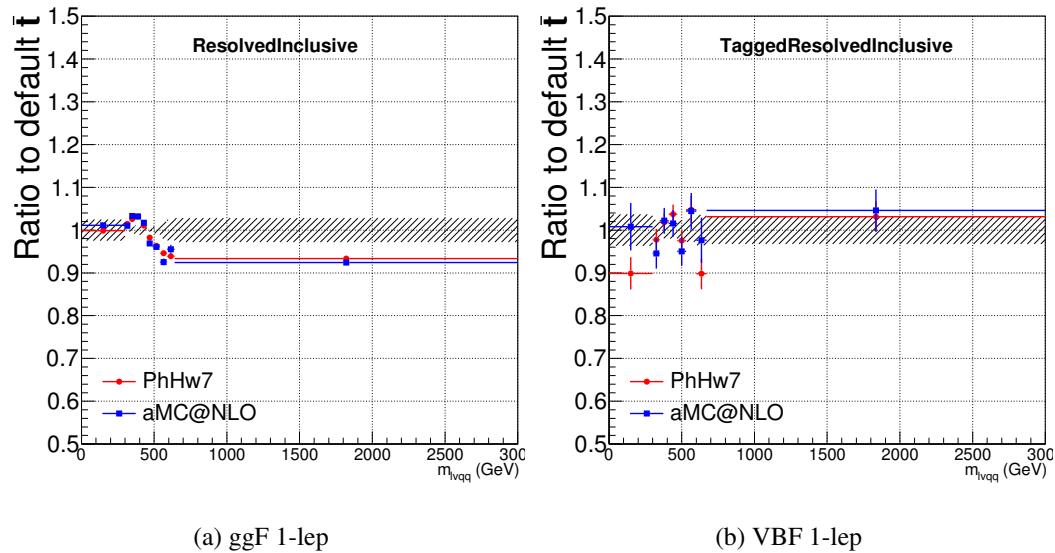


Figure 8.4: Ratio between the variations of generator (red) and hadronization (blue) variations for the Resolved regime for $t\bar{t}$ sample.

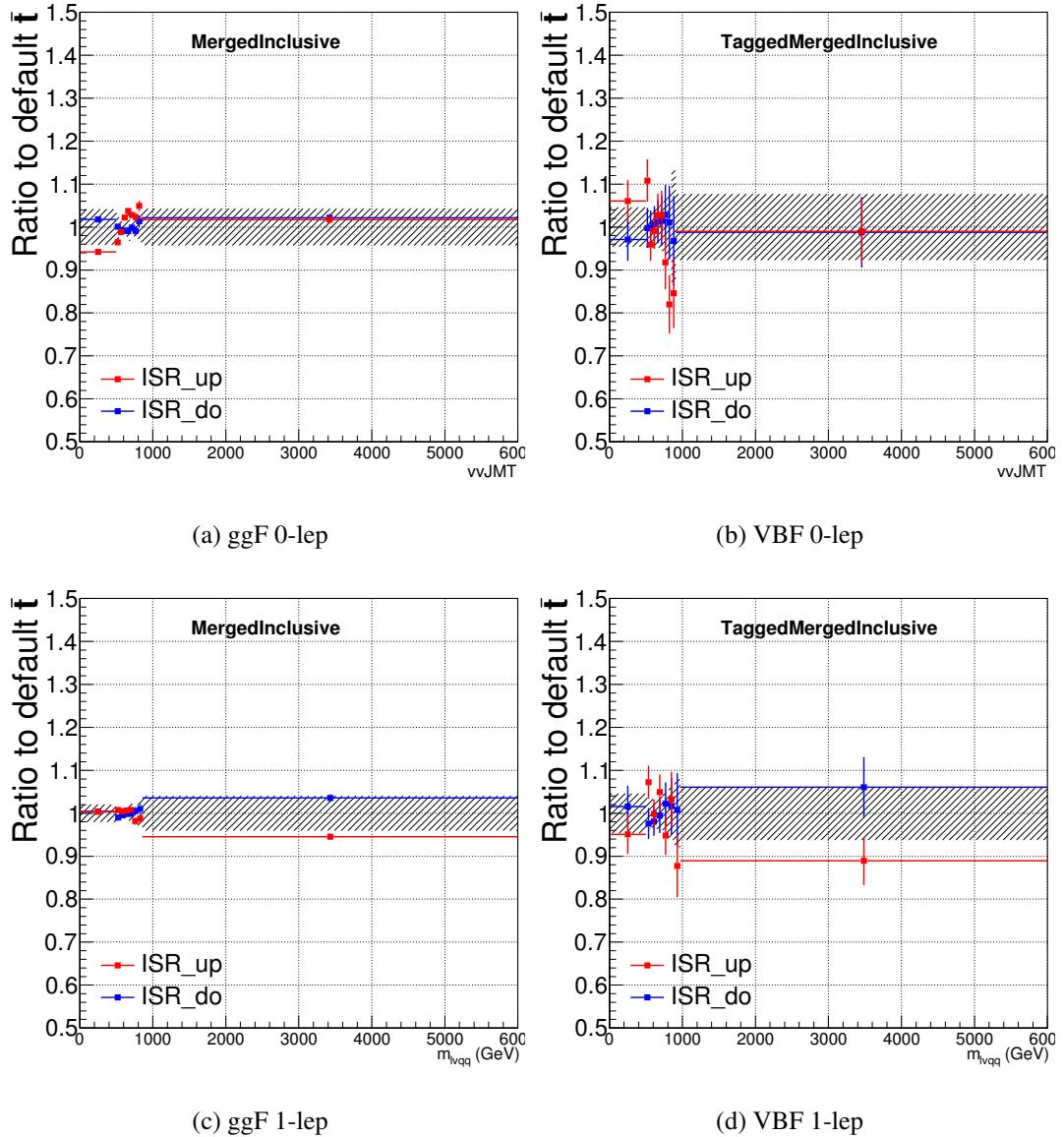


Figure 8.5: Ratio between the variations of ISR up (red) and down (blue) variations for the Merged regime for $t\bar{t}$ sample.

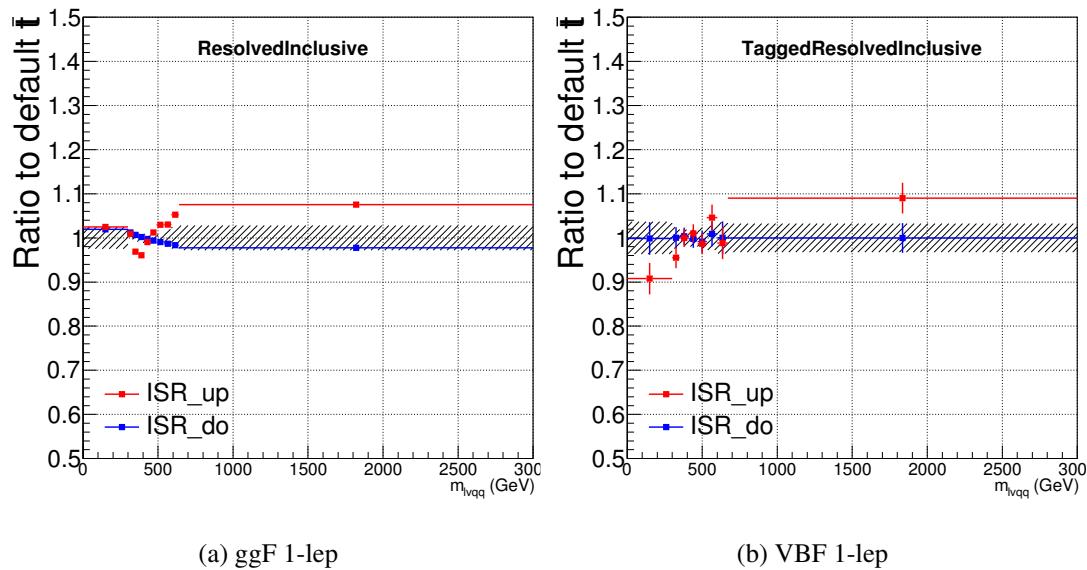


Figure 8.6: Ratio between the variations of ISR up (red) and down (blue) variations for the Resolved regime for $t\bar{t}$ sample.

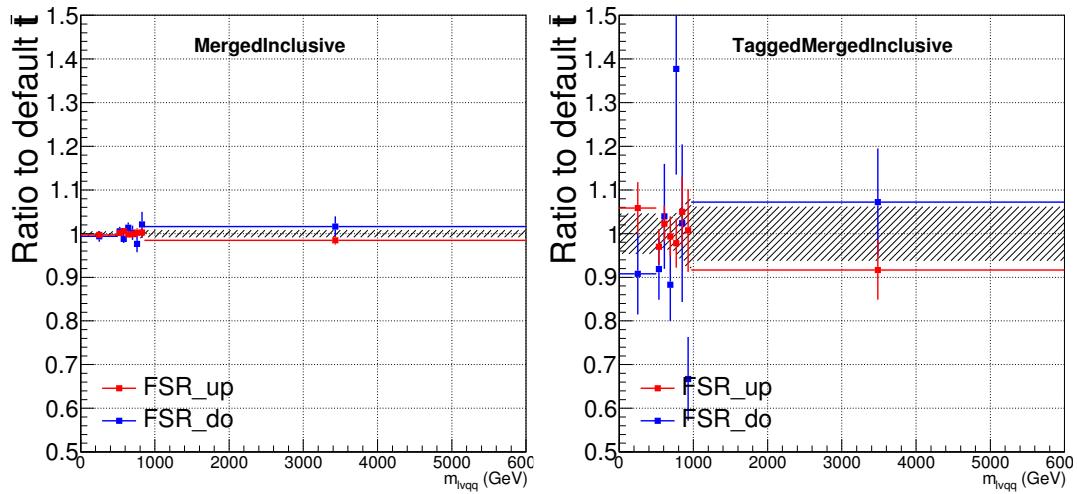


Figure 8.7: Ratio between the variations of FSR up (red) and down (blue) variations for the Merged regime for $t\bar{t}$ sample.

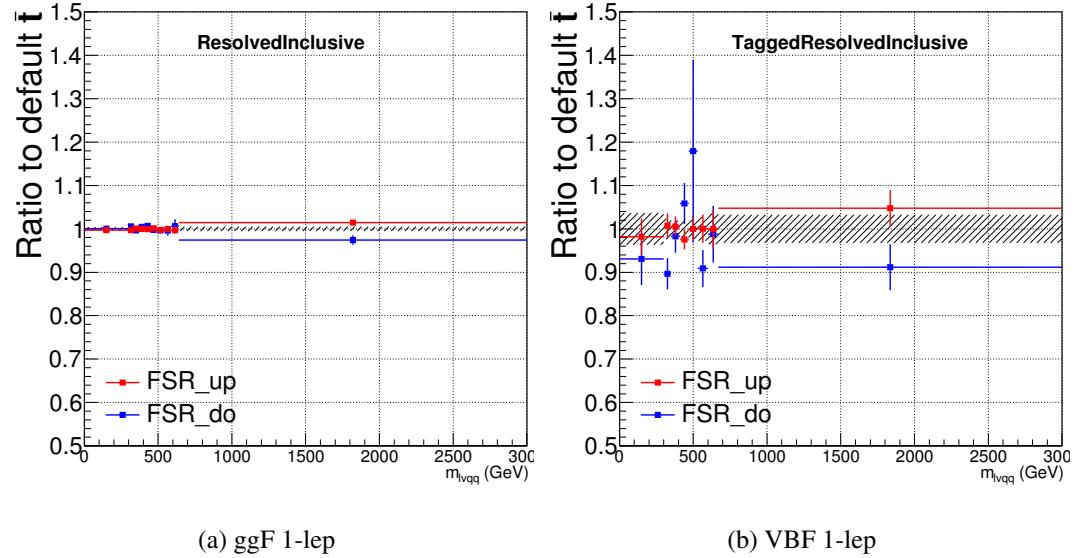


Figure 8.8: Ratio between the variations of FSR up (red) and down (blue) variations for the Resolved regime for $t\bar{t}$ sample.

1048 **Chapter 9**

1049 **Statistical Analysis**

1050 To determine the compatibility of the data collected with the proposed reso-
1051 nances a statistical procedure based on a likelihood function is used. A discovery
1052 test is used to measure the compatibility of the observed data with the back-
1053 ground only hypothesis. If the observed data is sufficiently incompatible with the
1054 background only hypothesis, this could indicate a discovery. In the absence of
1055 discovery, upper limits on the signal strength parameter, μ , are assessed using the
1056 CLs method.

1057 **9.1 Likelihood Function Definition**

1058 The likelihood function is product of Poisson probabilities for all analysis bins
1059 and systematic constraints:

$$\mathcal{L}(\mu, \boldsymbol{\theta}) = \prod_c \prod_i \frac{(\mu s_{ci}(\boldsymbol{\theta}) + b_{ci}(\boldsymbol{\theta}))^{n_{ci}}}{n_{ci}!} e^{-(\mu s_{ci}(\boldsymbol{\theta}) + b_{ci}(\boldsymbol{\theta}))} \prod_k (\theta'_k | \theta_k) \quad (9.1)$$

1060 Here c are the analysis channels considered and i runs over all the $m_{\ell\nu qq}$ bins
1061 used in the fit. The signal strength parameter, μ , multiplies the expected signal

1062 yield in each analysis bin, s_{ci} . The background content for channel c and bin i is
1063 given by b_{ci} . The dependence of signal and background predictions on system-
1064 atic uncertainties is described by the aforementioned set of nuisance parameters
1065 θ , which are parameterized by Gaussian or log-normal priors denoted here as
1066 θ_k . Statistical uncertainties of the simulated bin contents are also included as
1067 systematic uncertainties. Most systematics are correlated among all the analysis
1068 regions and considered to be independent from each other. The validity of this
1069 assumption is checked by evaluating the covariance of nuisance parameters.

1070 9.2 Fit Configuration

1071 The binning of $m_{\ell\nu qq}$ in signal regions for likelihood fit is determined by the
1072 statistical uncertainty of signal mass width. For each signal mass point, the signal
1073 mass resolution is given by the fitted Gaussian width of the $m_{\ell\nu qq}$. The fitted
1074 signal widths are then fit to a line to give a parameterized signal mass width,
1075 as shown in Figures 9.1 and 9.2. Bin widths are set first to this parameterized
1076 signal mass resolution. Then if the statistical uncertainty of the data or simulated
1077 background is more than 50%, bins are merged until the statistical uncertainty is
1078 less than 50%. All control regions contain only a single bin.

1079 For this analysis, each signal model fits the Merged and Resolved channels for
1080 the relevant signal production mode simultaneously. The control regions are used
1081 to extract $W+jets$ and $t\bar{t}$ backgrounds normalizations in the signal regions.

1082 Systematics may be affected by low statistics, leading to unsmooth m_{VV} dis-
1083 tributions with unphysically large fluctuations. This can lead to artificial pulls
1084 and constraints in the fit. To remove such issues a multi-step smoothing pro-
1085 cedure is applied to all systematic variation distributions in all regions. First,
1086 distributions are rebinned until the statistical error per bin is at least 5%. Next

1087 all local extrema are identified. The bins around smallest extrema are iteratively
1088 merged until only four local extrema remain. Then distributions are rebinned so
1089 that statistical uncertainties in each bin are $< 5\%$.

1090 For some systematics, up and down variations may be in the same direction
1091 with respect to the nominal distributions, leading to asymmetric distributions.
1092 This causes the variations to not cover the nominal choice, and the interpretation
1093 of the confidence interval is skewed. This asymmetry may also lead to uncon-
1094 strained systematics in the fit. To handle such asymmetric systematics, if the up
1095 and down variation for a given systematic are in the same direction for at least
1096 three m_{VV} bins the variation is averaged for those bins. The averaging proce-
1097 dure replaces bin-by-bin the up and down variation bins by $b_{\pm}^{new} = b_{nom} \pm \frac{|b_+ - b_-|}{2}$,
1098 where b_{nom} is the nominal bin content and b_{\pm} are the original up and down varia-
1099 tion bin content. The same procedure is also applied to any variations where the
1100 integral of the difference between the up/down variation and the nominal distri-
1101 bution is twice that of the other down/up variation, further ensuring variations
1102 are symmetric around the nominal distribution.

1103 Finally, systematics that have a negligible effect on the m_{VV} distribution are
1104 not considered in the fit. Shape systematics where no bin in the variational dis-
1105 tribution deviates more than 1% from the nominal distribution (after normalizing
1106 all histograms to the nominal) are not included in the fit. Also, statistical bin
1107 uncertainties $< 1\%$ are ignored.

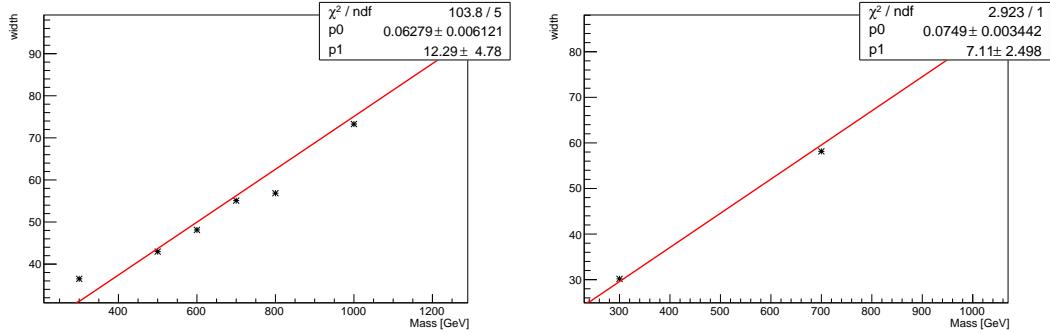


Figure 9.1: The HVT signal mass resolution as a function of mass fit with a straight line in the Resolved ggF region (left) and VBF (right) region.

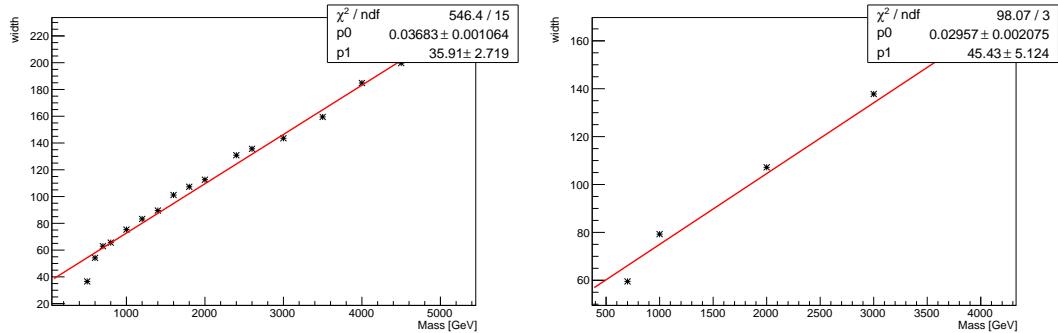


Figure 9.2: The HVT signal mass resolution as a function of mass fit with a straight line in the Merged ggF region (left) and VBF (right) region.

1108 9.3 Best Fit μ

1109 The best fit signal strength parameter is denoted by $\hat{\mu}$ and calculated by
 1110 maximizing the likelihood function with respect to all systematics and μ . The
 1111 corresponding set of systematics that maximize the likelihood are given by $\hat{\mu}$.
 1112 The first term in the likelihood is maximized when the expected number of signal
 1113 and background events is equal to the number of events in data ($n_{ci} = \mu s_{ci} +$
 1114 b_{ci}). Thus, by maximizing the likelihood, the fit determines values of μ and θ
 1115 that give the best agreement between expected and measured event yields. The

1116 second term in the likelihood is a penalty term which decreases the likelihood
1117 when systematics are shifted from their nominal values. This prevents the fit
1118 from profiling systematics in unphysical ways to maximize the likelihood. The
1119 uncertainty on μ is calculated by varying μ up and down until the natural log of
1120 the likelihood function shifts by one-half.

1121 9.4 Discovery Test

1122 To determine if the observed dataset is consistent with tested signal model a
1123 likelihood ratio is constructed:

$$\lambda(\mu) = \frac{\mathcal{L}(\mu, \hat{\theta}_\mu)}{\mathcal{L}(\hat{\mu}, \hat{\theta})} \quad (9.2)$$

1124 The denominator in this equation is the maximized value of \mathcal{L} over all system-
1125 atics and μ . The numerator is the maximized likelihood over all systematics for
1126 a given μ value, where the maximized systematics are given by $\hat{\mu}_\mu$. To test for
1127 the existence of signal the observed dataset the null hypothesis (H_0) is defined as
1128 the background only hypothesis and the alternate hypothesis includes signal and
1129 background (H_1). This test quantifies the compatibility of observed data with
1130 H_0 by calculating a p-value representing the probability of observing data as dis-
1131 crepant or more than the observed data under the H_0 . The test statistic used to
1132 calculate this p-value is given by (r_0):

$$r_0 = \begin{cases} -2 \ln \lambda(0), \hat{\mu} > 0 \\ +2 \ln \lambda(0), \hat{\mu} < 0 \end{cases} \quad (9.3)$$

1133 The expected distribution of the the test statistic under H_0 ($f(r_0|0)$) is used to
1134 calculate the p-value:

$$p_0 = \int_{r_0, obs}^{\infty} f(r_0 | 0) dr_0 \quad (9.4)$$

1135 Small p-values indicate the observed data is poorly described by H_0 . This
 1136 equivalent Z-score of a given p-value is usually used to further quantify the agree-
 1137 ment between the observed data and H_0 . The Z-score is given by the number of
 1138 standard deviations away from the mean of a Gaussian distribution, the integral
 1139 of the upper tail of the distribution would equal the p-value. Mathematically:

$$Z = \Phi^{-1}(1 - p_0) \quad (9.5)$$

1140 where Φ is the Gaussian cumulative distribution function. The statistical
 1141 significance of these tests are expressed as the Z -score. In particle physics, 3σ is
 1142 considered evidence for new phenomena and 5σ is the threshold for discovery.

1143 **9.5 Exclusion Limits**

1144 In the absence of discovery, upper limits on the signal strength, μ are set using
 1145 the CLs method [cite P60]. The test statistic for this test, q_μ , is constructed as:

$$\tilde{\lambda}_\mu = \begin{cases} \frac{\mathcal{L}(\mu, \hat{\theta}_\mu)}{\mathcal{L}(\hat{\mu}, \hat{\theta})}, \hat{\mu} > 0 \\ \frac{\mathcal{L}(\mu, \hat{\theta}_\mu)}{\mathcal{L}(0, \hat{\theta}_0)}, \hat{\mu} < 0 \end{cases} \quad (9.6)$$

$$\tilde{q}_\mu = \begin{cases} -2 \ln \tilde{\lambda}(\mu), \hat{\mu} < \mu \\ +2 \ln \tilde{\lambda}(\mu), \hat{\mu} > \mu \end{cases} \quad (9.7)$$

1146 As defined, larger values of q_μ correspond to increasing incompatibility between
 1147 the observed data and the background + signal hypothesis. The observed value
 1148 of the test statistic, $q_{\mu, obs}$, is then compared to its expected distribution, f , to

1149 calculate p-values to assess the likelihood of the background+signal hypothesis.

1150 Using these distributions, CL_s values are computed as:

$$1151 \quad CL_{s+b} = \int_{q_{\mu,obs}}^{\infty} f(q_{\mu}|\mu) dq_{\mu} \quad (9.8)$$

$$1152 \quad CL_b = \int_{q_0^{obs}}^{\infty} f(q_{\mu}|\mu = 0) dq_{\mu} \quad (9.9)$$

$$1153 \quad CL_s = \frac{CL_{s+b}}{CL_b} \quad (9.10)$$

1153 CL_{s+b} is the p-value for the signal + background hypothesis and CL_b is the
1154 p-value for the background only hypothesis. The CL_s value is interpreted as
1155 the probability to observe the background + signal hypothesis normalized to the
1156 probability of background-only hypothesis. Normalizing by CL_b prevents setting
1157 artificially strong exclusion limits due to downward fluctuations in data.

1158 In this analysis, μ values are scanned for each bin in the fit to find the μ value
1159 that yields $CL_s=0.05$, meaning the likelihood of finding data more incompatible
1160 with the signal+background hypothesis (relative to the background only hypoth-
1161 esis) is 5%. The 95% upper limit on the cross section is then calculated as the
1162 product of the μ value found, branching ratio, and theory cross section.

Part IV

1163

Results

1164

1165 **Chapter 10**

1166 **Statistical Interpretation**

1167 **10.1 Discovery Tests**

1168 To test for the existence of signal in the observed dataset, the discovery tests
1169 discussed earlier are used to calculate p-values as a function of resonance mass.
1170 The results of these tests are shown in Figures 10.1 - 10.5. Across the different
1171 DY signals the largest excesses are $\sim 2.2\sigma$ at 600 GeV and 1.8σ at 2 TeV. The
1172 largest excesses for VBF signals are $< 2.5\sigma$ at for 1 TeV resonances. As these
1173 deviations do not constitute discoveries, upper limits on μ are calculated.

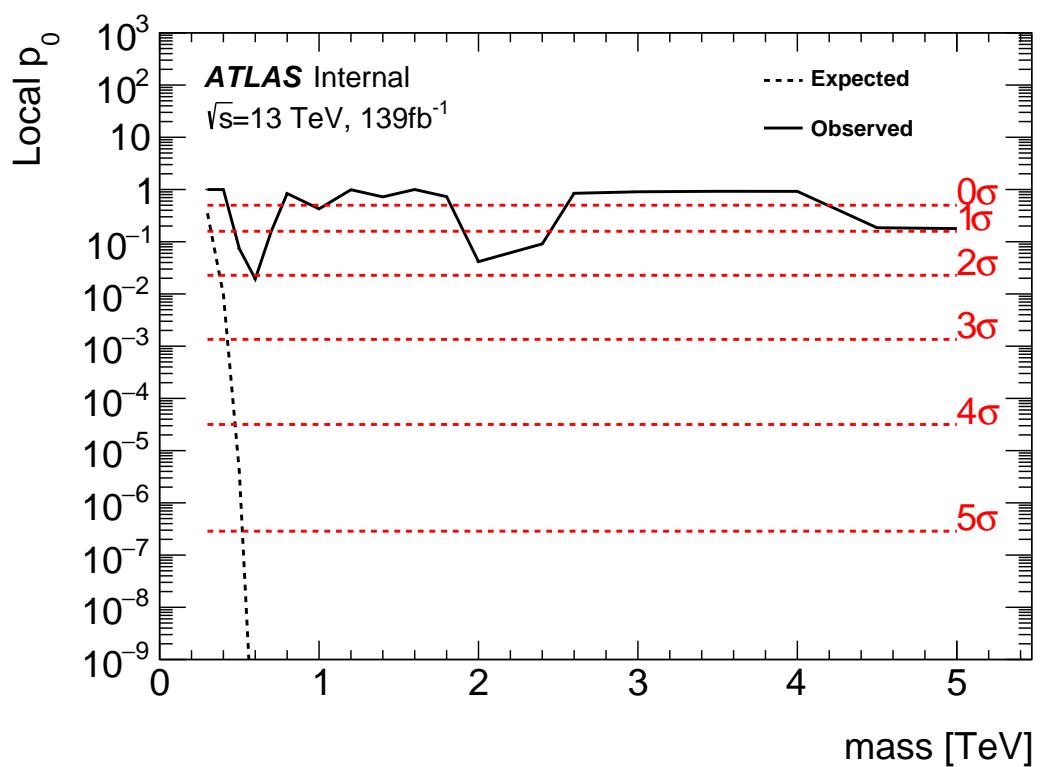


Figure 10.1: These plots show the measured p_0 value as a function of resonance mass for HVT Z' DY production.

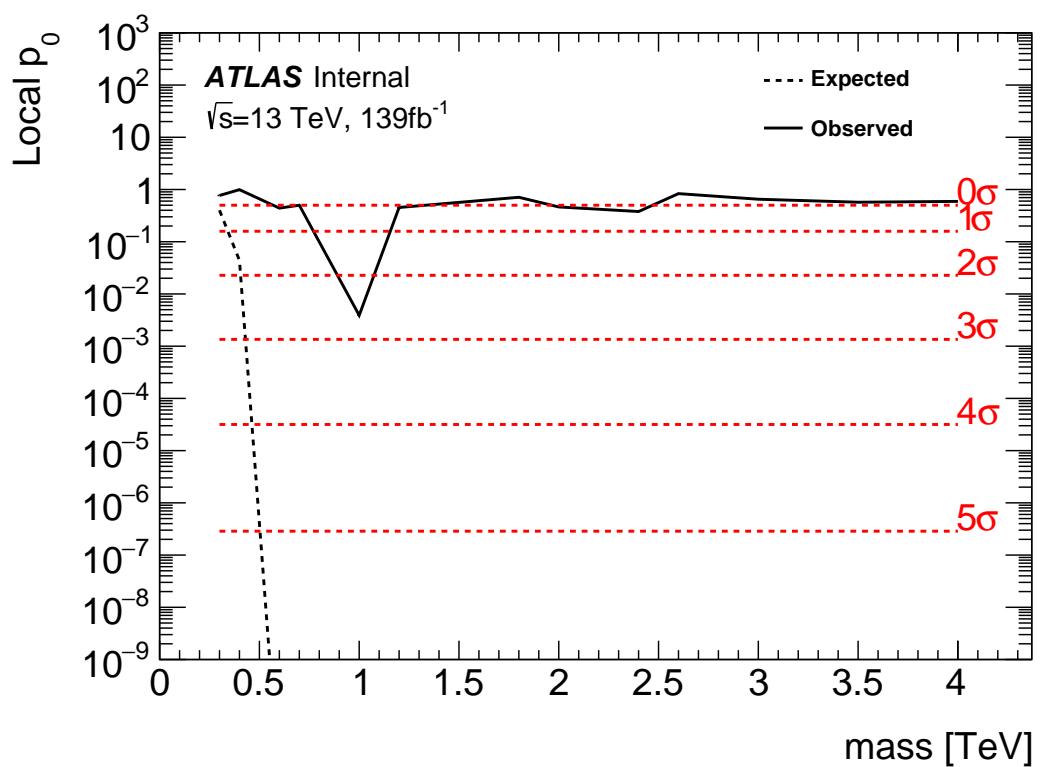


Figure 10.2: These plots show the measured p_0 value as a function of resonance mass for HVT Z' VBF production.

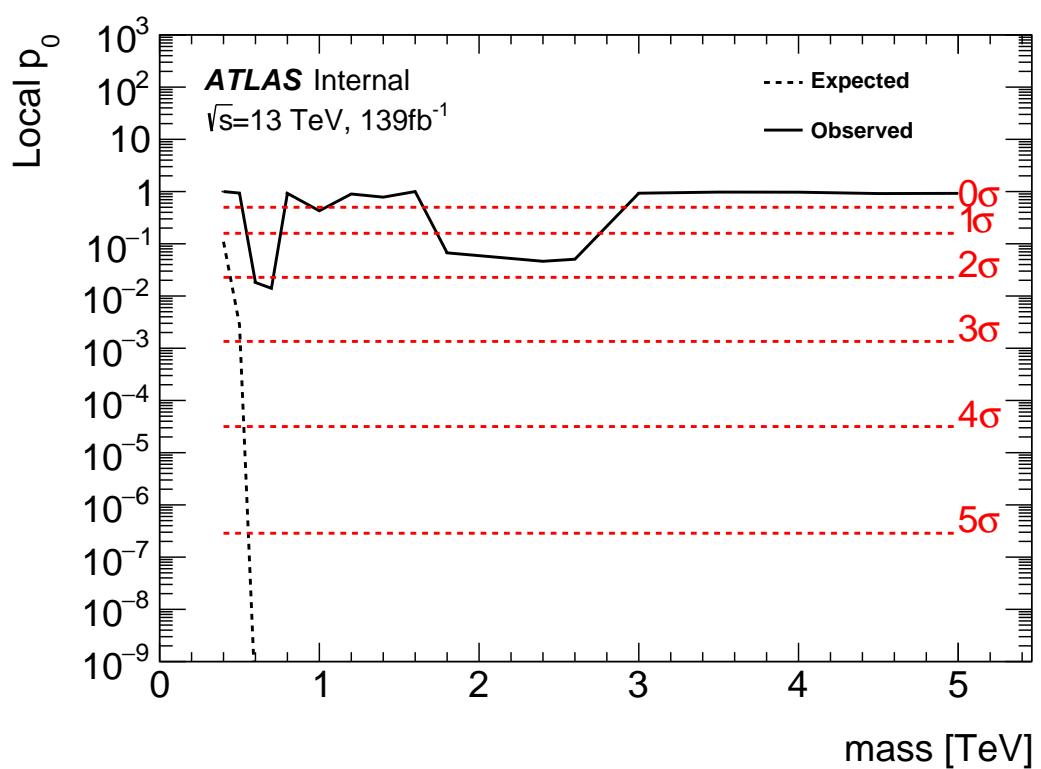


Figure 10.3: These plots show the measured p_0 value as a function of resonance mass for HVT W' DY production.

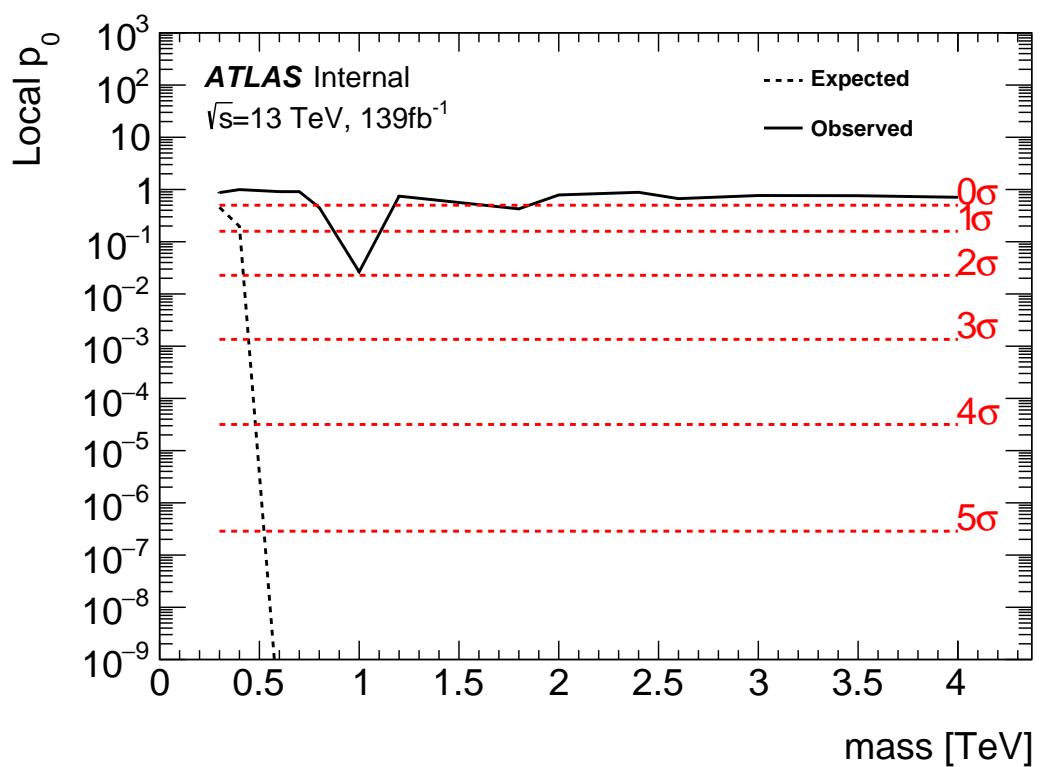


Figure 10.4: These plots show the measured p_0 value as a function of resonance mass for HVT W' VBF production.

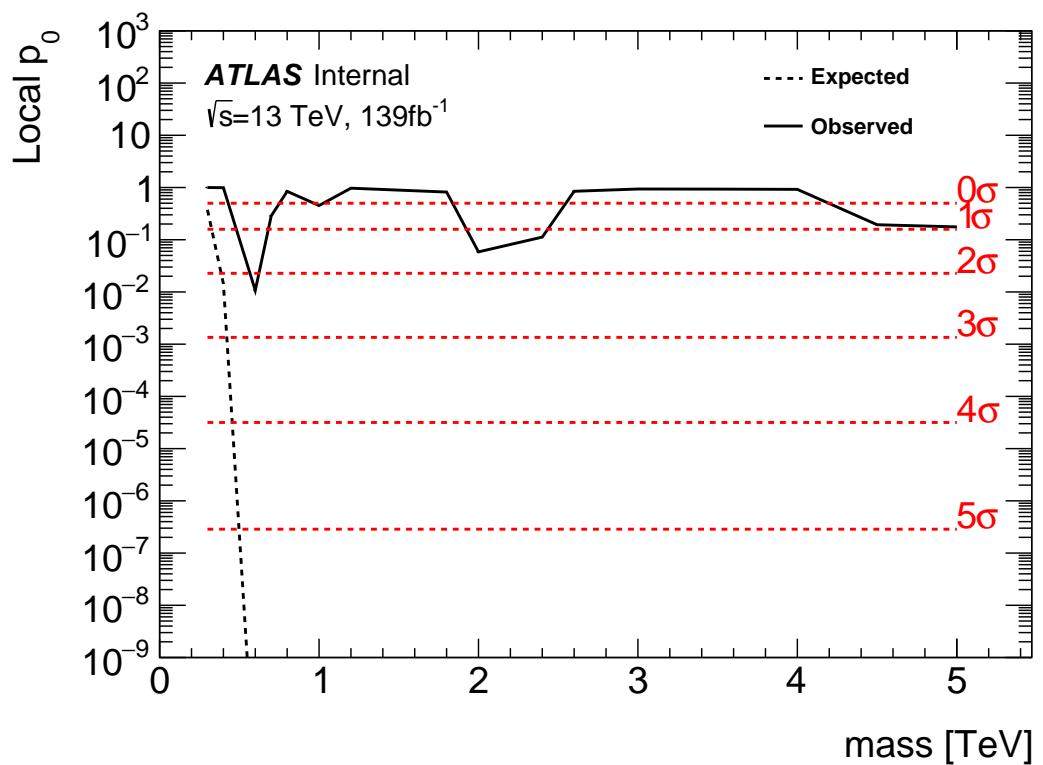
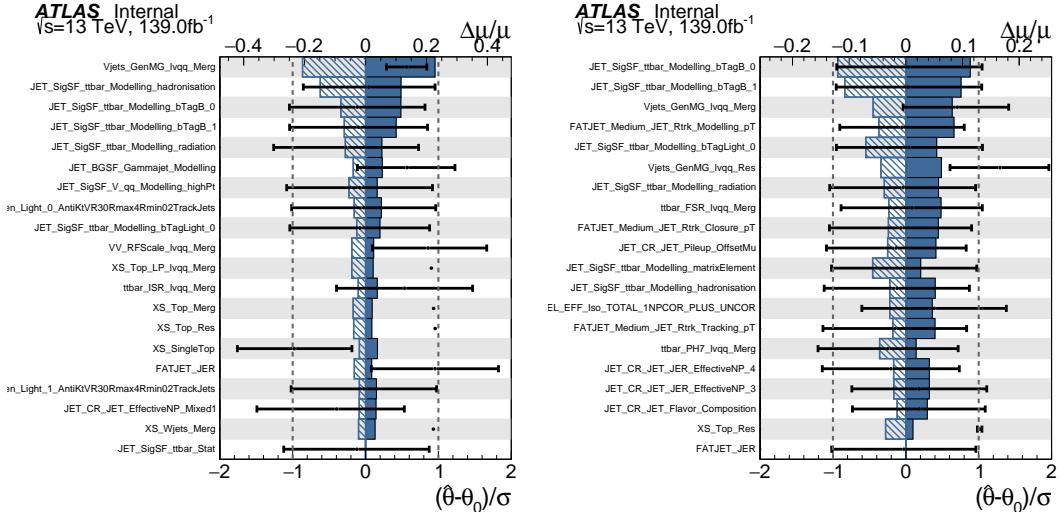


Figure 10.5: These plots show the measured p_0 value as a function of resonance mass for the RS Graviton DY production.

1174 10.2 Systematic Profiling and Correlations

1175 The ranked systematics and their fitted values are shown for the different anal-
 1176 ysis regions in Figure 10.6 - 10.8. Note that background normalizations for $W+jets$
 1177 and $t\bar{t}$ are left free to float in the fit. This means the nominal normalization val-
 1178 ues are at one and the uncertainties are not plotted in the ranked plots. Overall,
 1179 systematics are not pulled outside their uncertainties, especially for highly ranked
 1180 nuisance parameters.

1181 The correlation between systematics are shown in Figures ???. Correlations
 1182 between background normalization are expected. The remaining systematic cor-
 1183 relations are not very strong or unexpected.



1183 **Figure 10.6:** Ranked systematics and their fitted values for WW DY (right) and VBF (left) selections.

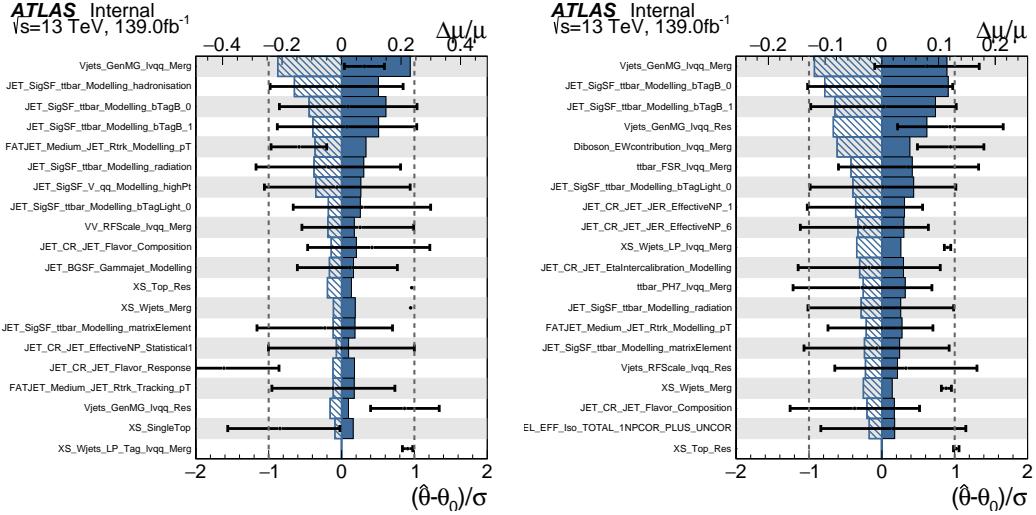


Figure 10.7: Ranked systematics and their fitted values for WZ DY (right) and VBF (left) selections.

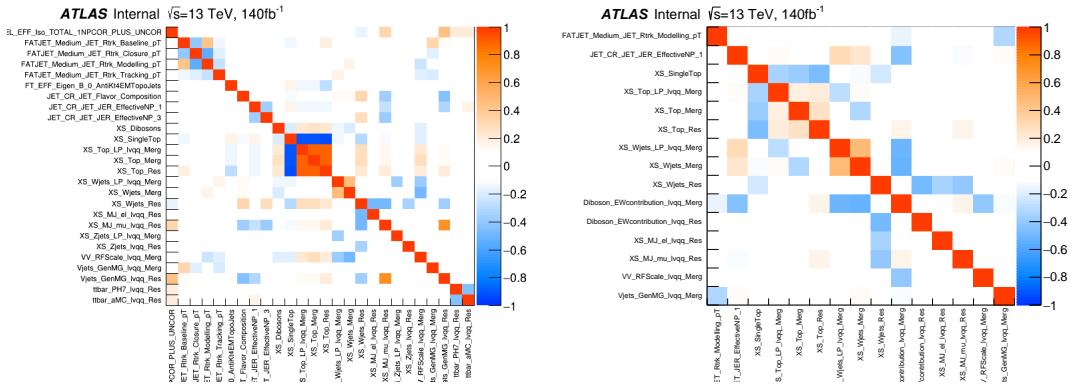


Figure 10.8: Correlations between systematics for WW DY (right) and VBF (left) selections.

1184 10.3 Expected and Measured Yields

1185 The yield tables for the four analysis regions are shown in Tables ?? - ?? . The
 1186 fitted background normalizations are shown in Tables ??-?? . The control region
 1187 $m_{\ell\nu qq}$ distributions are shown in Figures 10.9 - 10.12. The signal region $m_{\ell\nu qq}$
 1188 distributions are shown in Figures 10.13 - ?? .

	HP WCR	LP WCR	Resolved WCR
Electron Multi-jet	-	-	16507.83 ± 2314.87
Muon Multi-jet	-	-	19977.12 ± 2816.06
Diboson	1833.41 ± 177.78	3323.93 ± 320.92	9147.67 ± 961.63
Single-top	2160.62 ± 402.34	3551.09 ± 660.00	20058.36 ± 3817.26
$t\bar{t}$	15518.86 ± 338.22	24069.54 ± 453.15	138866.23 ± 1989.71
$W+jets$	40141.57 ± 357.79	88113.06 ± 487.87	673200.38 ± 4120.53
$Z+jets$	778.83 ± 78.93	1765.54 ± 179.10	16570.50 ± 1672.71
Total	60433.29 ± 664.92	120823.16 ± 1006.99	894328.12 ± 7247.12
Data	60264.00	120852.00	895362.00
	HP TCR	LP TCR	Resolved TCR
Electron Multi-jet	-	-	-
Muon Multi-jet	-	-	-
Diboson	421.11 ± 37.98	550.44 ± 53.10	996.87 ± 119.63
Single-top	4691.44 ± 846.11	3466.26 ± 631.03	16848.71 ± 3258.26
$t\bar{t}$	38945.18 ± 848.77	33836.95 ± 637.04	224226.14 ± 3212.76
$W+jets$	2258.34 ± 20.13	6564.78 ± 36.35	23466.41 ± 143.63
$Z+jets$	66.35 ± 6.72	213.26 ± 21.63	846.66 ± 85.47
Total	46382.43 ± 1199.25	44631.70 ± 899.23	266384.78 ± 4580.43
Data	46354.00	44629.00	266443.00
	WW SR	LP SR	Resolved 1-lepton SR
Electron Multi-jet	-	-	10788.40 ± 1512.85
Muon Multi-jet	-	-	15759.50 ± 2221.53
Diboson	4990.30 ± 376.50	3901.07 ± 313.22	16971.29 ± 1523.77
Single-top	3117.71 ± 565.07	2176.46 ± 400.52	20422.85 ± 3731.94
$t\bar{t}$	13785.77 ± 302.14	11005.12 ± 207.41	126965.25 ± 1819.66
$W+jets$	24718.56 ± 223.72	60080.66 ± 333.12	444133.56 ± 2719.02
$Z+jets$	478.18 ± 48.46	1226.69 ± 124.44	11686.32 ± 1179.69
Total	47090.52 ± 777.65	78389.98 ± 654.22	646727.19 ± 5963.98
Data	47330.00	78380.00	645610.00

Table 10.1: Expected and Measured for DY WW $W+jets$, $t\bar{t}$ control regions and signal regions.

	HP Untagged WCR	LP Untagged WCR	Resolved Untagged WCR
Electron Multi-jet	-	-	15080.03 ± 2277.99
Muon Multi-jet	-	-	27347.10 ± 2950.07
Diboson	1508.48 ± 154.20	2758.24 ± 284.50	9038.55 ± 728.69
Single-top	1756.59 ± 306.69	2913.18 ± 515.93	20511.74 ± 3523.47
$t\bar{t}$	13134.00 ± 238.30	21815.37 ± 334.98	140157.77 ± 2636.96
$W+jets$	40654.84 ± 333.65	87657.76 ± 501.96	665909.12 ± 4420.62
$Z+jets$	768.72 ± 77.97	1759.87 ± 178.96	16512.46 ± 1673.23
Total	57822.63 ± 540.40	116904.42 ± 862.16	894556.75 ± 7492.20
Data	57699.00	117306.00	895362.00
	HP Tagged WCR	LP Tagged WCR	Resolved Tagged WCR
Electron Multi-jet	-	-	384.58 ± 57.11
Muon Multi-jet	-	-	602.93 ± 190.12
Diboson	30.22 ± 4.69	48.95 ± 7.16	264.64 ± 28.24
Single-top	308.44 ± 56.19	371.59 ± 69.43	5752.39 ± 1029.97
$t\bar{t}$	1683.82 ± 48.73	2041.48 ± 70.00	58431.49 ± 614.30
$W+jets$	583.55 ± 75.37	1109.45 ± 85.78	11891.68 ± 903.01
$Z+jets$	13.19 ± 1.34	23.06 ± 2.34	324.74 ± 32.85
Total	2619.22 ± 106.00	3594.53 ± 130.90	77652.45 ± 1514.89
Data	2565.00	3546.00	77973.00
	HP Untagged TCR	LP Untagged TCR	Resolved Untagged TCR
Electron Multi-jet	-	-	-
Muon Multi-jet	-	-	-
Diboson	289.45 ± 28.45	346.78 ± 35.85	650.85 ± 65.56
Single-top	3107.99 ± 538.03	2250.64 ± 385.41	9606.87 ± 1698.22
$t\bar{t}$	30992.40 ± 562.33	26954.21 ± 413.89	91893.59 ± 1728.91
$W+jets$	2236.29 ± 18.35	4874.03 ± 27.91	16122.97 ± 107.03
$Z+jets$	71.54 ± 7.26	155.50 ± 15.81	577.71 ± 58.54
Total	36697.66 ± 779.03	34581.16 ± 567.59	118851.98 ± 2427.40
Data	36677.00	34573.00	118928.00
	HP Tagged TCR	LP Tagged TCR	Resolved Tagged TCR
Electron Multi-jet	-	-	-
Muon Multi-jet	-	-	-
Diboson	9.72 ± 1.13	8.75 ± 1.16	34.06 ± 4.98
Single-top	105.87 ± 20.65	119.66 ± 22.68	656.89 ± 132.96
$t\bar{t}$	1904.75 ± 50.61	1483.86 ± 47.05	17965.33 ± 188.87
$W+jets$	32.36 ± 4.28	85.74 ± 6.96	489.01 ± 37.13
$Z+jets$	1.27 ± 0.13	1.93 ± 0.20	19.14 ± 1.94
Total	2053.98 ± 54.84	1699.93 ± 52.70	19164.43 ± 234.01
Data	2047.00	1708.00	19143.00

Table 10.2: Expected and Measured for DY WZ $W+jets$, $t\bar{t}$ tag and untag control regions.

	HP Untagged SR	LP Untagged SR	Resolved Untagged SR
Electron Multi-jet	-	-	7782.17 ± 1175.56
Muon Multi-jet	-	-	17004.81 ± 1834.40
Diboson	3041.17 ± 273.77	2266.35 ± 212.79	14724.12 ± 1224.31
Single-top	2123.28 ± 373.83	1379.35 ± 240.92	18336.88 ± 3082.47
$t\bar{t}$	11678.86 ± 213.63	8906.34 ± 136.88	112669.24 ± 2122.46
$W+jets$	22741.32 ± 191.47	41726.76 ± 240.56	342934.00 ± 2280.21
$Z+jets$	442.03 ± 44.84	849.79 ± 86.42	9271.83 ± 939.52
Total	40026.65 ± 546.81	55128.59 ± 432.90	522723.03 ± 5131.71
Data	40193.00	54735.00	521813.00
	HP Tagged SR	LP Tagged SR	Resolved Tagged SR
Electron Multi-jet	-	-	199.22 ± 29.58
Muon Multi-jet	-	-	393.43 ± 124.06
Diboson	102.58 ± 11.59	65.44 ± 8.05	624.07 ± 58.10
Single-top	178.21 ± 33.62	155.53 ± 28.95	3470.39 ± 617.48
$t\bar{t}$	1017.93 ± 31.95	706.76 ± 26.20	38189.30 ± 401.91
$W+jets$	325.58 ± 41.62	575.36 ± 43.29	6161.96 ± 467.71
$Z+jets$	7.81 ± 0.80	11.62 ± 1.19	183.36 ± 18.55
Total	1632.11 ± 63.39	1514.70 ± 58.86	49221.74 ± 884.06
Data	1699.00	1559.00	48919.00

Table 10.3: Expected and Measured for DY WZ $W+jets$, $t\bar{t}$ tag and untag signal regions.

	HP WCR	LP WCR	Resolved WCR
Electron Multi-jet	-	-	898.48 ± 137.82
Muon Multi-jet	-	-	601.46 ± 182.74
Diboson	107.45 ± 45.20	166.87 ± 68.11	292.10 ± 235.29
Single-top	78.19 ± 18.22	132.71 ± 31.93	879.82 ± 216.89
$t\bar{t}$	400.71 ± 28.35	569.70 ± 48.88	5067.51 ± 155.69
$W+jets$	864.49 ± 63.44	1940.80 ± 89.41	18563.70 ± 408.99
$Z+jets$	19.51 ± 2.00	46.63 ± 4.77	795.20 ± 80.89
Total	1470.35 ± 84.89	2856.71 ± 126.74	27098.28 ± 594.01
Data	1495.00	2898.00	27120.00
	HP TCR	LP TCR	Resolved TCR
Electron Multi-jet	-	-	-
Muon Multi-jet	-	-	-
Diboson	14.95 ± 6.61	27.57 ± 14.12	24.33 ± 20.32
Single-top	68.31 ± 16.17	58.93 ± 13.56	278.60 ± 73.04
$t\bar{t}$	496.60 ± 31.72	401.23 ± 32.13	3834.49 ± 104.60
$W+jets$	50.68 ± 4.19	144.02 ± 7.86	450.01 ± 11.87
$Z+jets$	1.32 ± 0.14	5.35 ± 0.55	29.96 ± 3.07
Total	631.87 ± 36.45	637.10 ± 38.44	4617.39 ± 129.77
Data	636.00	634.00	4615.00
	HP SR	LP SR	Resolved SR
Electron Multi-jet	-	-	596.34 ± 91.52
Muon Multi-jet	-	-	481.01 ± 144.48
Diboson	148.84 ± 48.64	181.42 ± 67.30	395.52 ± 318.06
Single-top	79.49 ± 19.80	56.82 ± 14.89	782.07 ± 190.79
$t\bar{t}$	338.42 ± 24.14	236.80 ± 20.88	4261.70 ± 138.98
$W+jets$	501.13 ± 39.36	1347.76 ± 64.50	11445.73 ± 291.49
$Z+jets$	9.25 ± 0.95	28.77 ± 2.95	567.66 ± 57.94
Total	1077.13 ± 69.93	1851.57 ± 96.73	18530.03 ± 523.88
Data	1096.00	1846.00	18530.00

Table 10.4: Expected and Measured for VBF WW $W+jets$, $t\bar{t}$ control regions and signal regions.

	HP WCR	LP WCR	Resolved WCR
Electron Multi-jet	-	-	870.00 ± 132.75
Muon Multi-jet	-	-	618.45 ± 196.90
Diboson	92.92 ± 41.77	145.90 ± 64.26	228.62 ± 114.62
Single-top	71.13 ± 16.29	118.82 ± 27.98	1209.87 ± 281.64
$t\bar{t}$	427.80 ± 29.72	509.19 ± 46.57	6860.87 ± 254.83
$W+jets$	871.68 ± 64.22	2020.67 ± 93.54	19088.50 ± 442.10
$Z+jets$	19.58 ± 2.01	47.39 ± 4.85	800.19 ± 82.02
Total	1483.11 ± 83.79	2841.97 ± 125.92	29676.50 ± 644.96
Data	1495.00	2898.00	29755.00
	HP TCR	LP TCR	Resolved TCR
Electron Multi-jet	-	-	-
Muon Multi-jet	-	-	-
Diboson	10.12 ± 4.51	12.73 ± 6.55	14.23 ± 7.49
Single-top	51.57 ± 12.31	35.07 ± 8.17	169.21 ± 44.54
$t\bar{t}$	470.06 ± 28.97	298.99 ± 25.28	2414.75 ± 75.42
$W+jets$	49.64 ± 4.17	109.69 ± 6.16	378.22 ± 12.05
$Z+jets$	1.28 ± 0.13	4.81 ± 0.50	17.62 ± 1.83
Total	582.67 ± 32.07	461.30 ± 28.05	2994.03 ± 88.75
Data	584.00	459.00	3001.00
	HP SR	LP SR	Resolved SR
Electron Multi-jet	-	-	444.65 ± 67.99
Muon Multi-jet	-	-	397.29 ± 125.59
Diboson	109.66 ± 44.13	112.28 ± 46.45	265.75 ± 139.43
Single-top	63.16 ± 15.20	48.02 ± 11.56	872.16 ± 205.00
$t\bar{t}$	348.95 ± 24.34	190.68 ± 17.75	5134.25 ± 193.57
$W+jets$	467.21 ± 37.12	973.73 ± 47.91	10226.83 ± 254.67
$Z+jets$	8.15 ± 0.84	23.62 ± 2.43	558.48 ± 57.25
Total	997.13 ± 64.42	1348.33 ± 70.06	17899.41 ± 432.98
Data	1018.00	1313.00	17826.00

Table 10.5: Expected and Measured for VBF WZ $W+jets$, $t\bar{t}$ control regions and signal regions.

Background	Fitted Normalization
XS_Top_LP_lvqq_Merg_binned	$0.905^{+0.0166}_{-0.0166}$
XS_Top_Merg	$0.936^{+0.0199}_{-0.0199}$
XS_Top_Res	$0.957^{+0.0134}_{-0.0134}$
XS_Wjets_LP_lvqq_Merg_binned	$0.884^{+0.00489}_{-0.00489}$
XS_Wjets_Merg	$0.931^{+0.00831}_{-0.00831}$
XS_Wjets_Res	$1.03^{+0.00628}_{-0.00628}$

Table 10.6: Fitted background normalizations for $t\bar{t}$ and $W+jets$ backgrounds for the DY WW analysis region.

Background	Fitted Normalization
XS_Top_LP_Tag_lvqq_Merg_binned	$0.973^{+0.0333}_{-0.0333}$
XS_Top_LP_lvqq_Merg_binned	$0.894^{+0.0135}_{-0.0135}$
XS_Top_Merg	$0.893^{+0.016}_{-0.016}$
XS_Top_Res	$0.965^{+0.0179}_{-0.0179}$
XS_Top_Tag_lvqq_Merg_binned	$0.954^{+0.0276}_{-0.0276}$
XS_Top_Tag_lvqq_Res_binned	$0.999^{+0.0105}_{-0.0105}$
XS_Wjets_LP_Tag_lvqq_Merg_binned	$0.912^{+0.0703}_{-0.0703}$
XS_Wjets_LP_lvqq_Merg_binned	$0.876^{+0.00502}_{-0.00502}$
XS_Wjets_Merg	$0.948^{+0.00779}_{-0.00779}$
XS_Wjets_Res	$1.01^{+0.00673}_{-0.00673}$
XS_Wjets_Tag_lvqq_Merg_binned	$0.906^{+0.117}_{-0.117}$
XS_Wjets_Tag_lvqq_Res_binned	$1.2^{+0.0904}_{-0.0904}$

Table 10.7: Fitted background normalizations for $t\bar{t}$ and $W+jets$ backgrounds for the DY WZ analysis region.

Background	Fitted Normalization
XS_Top_LP_lvqq_Merg_binned	$0.79^{+0.0673}_{-0.0673}$
XS_Top_Merg	$0.888^{+0.061}_{-0.061}$
XS_Top_Res	$1.01^{+0.0311}_{-0.0311}$
XS_Wjets_LP_lvqq_Merg_binned	$0.88^{+0.0423}_{-0.0423}$
XS_Wjets_Merg	$0.881^{+0.0677}_{-0.0677}$
XS_Wjets_Res	$0.932^{+0.0202}_{-0.0202}$

Table 10.8: Fitted background normalizations for $t\bar{t}$ and $W+jets$ backgrounds for the VBF WW analysis region.

Background	Fitted Normalization
XS_Top_LP_lvqq_Merg_binned	$0.708^{+0.064}_{-0.064}$
XS_Top_Merg	$0.958^{+0.0644}_{-0.0644}$
XS_Top_Res	$1.02^{+0.038}_{-0.038}$
XS_Wjets_LP_lvqq_Merg_binned	$0.9^{+0.0438}_{-0.0438}$
XS_Wjets_Merg	$0.883^{+0.0685}_{-0.0685}$
XS_Wjets_Res	$0.945^{+0.0219}_{-0.0219}$

Table 10.9: Fitted background normalizations for $t\bar{t}$ and $W+jets$ backgrounds for the VBF WZ analysis region.

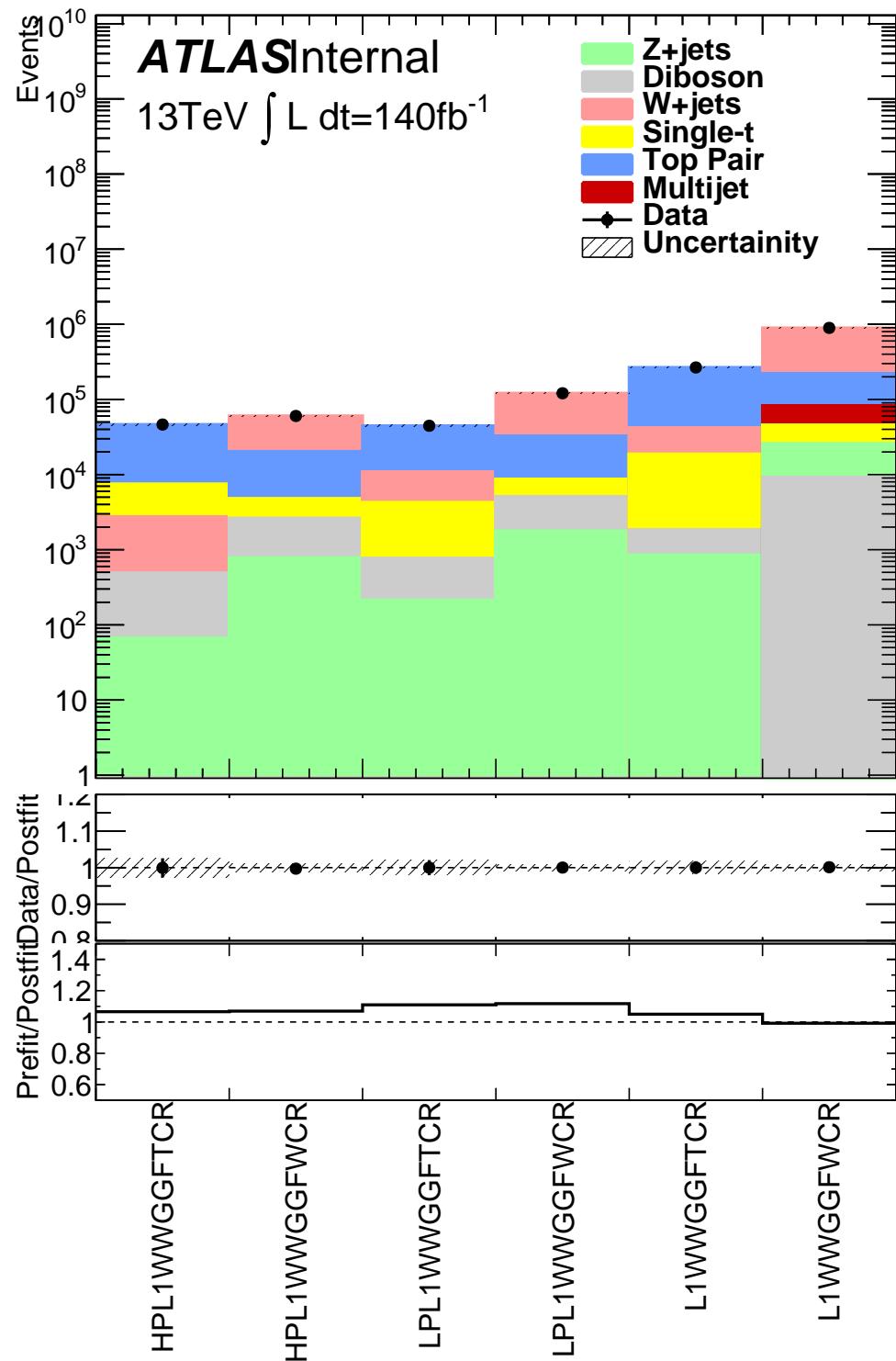


Figure 10.9: This figure shows the distribution of $m_{\ell\nu qq}$ in the DY WW control regions.

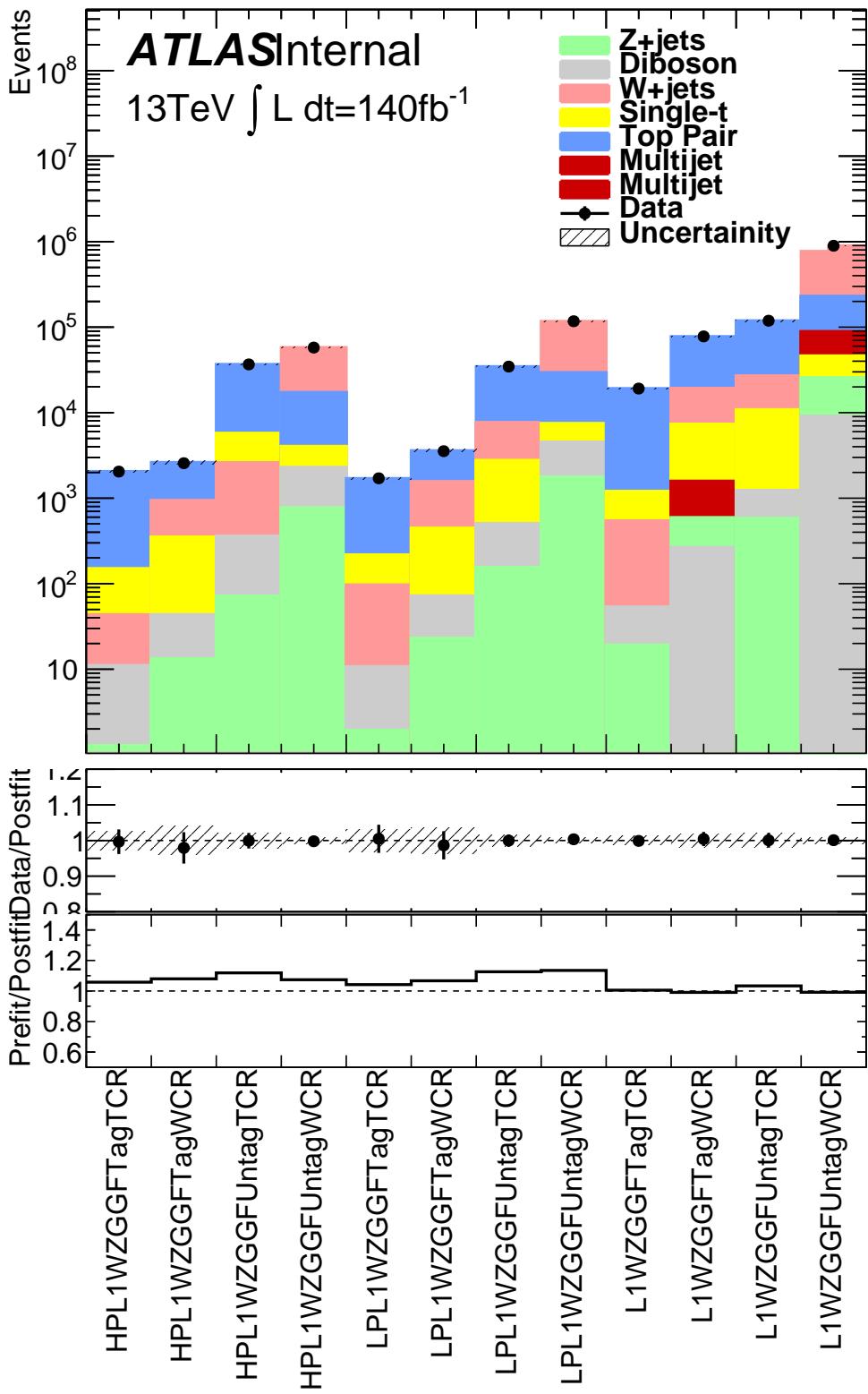


Figure 10.10: This figure shows the distribution of $m_{\ell\nu qq}$ in the DY WZ control regions.

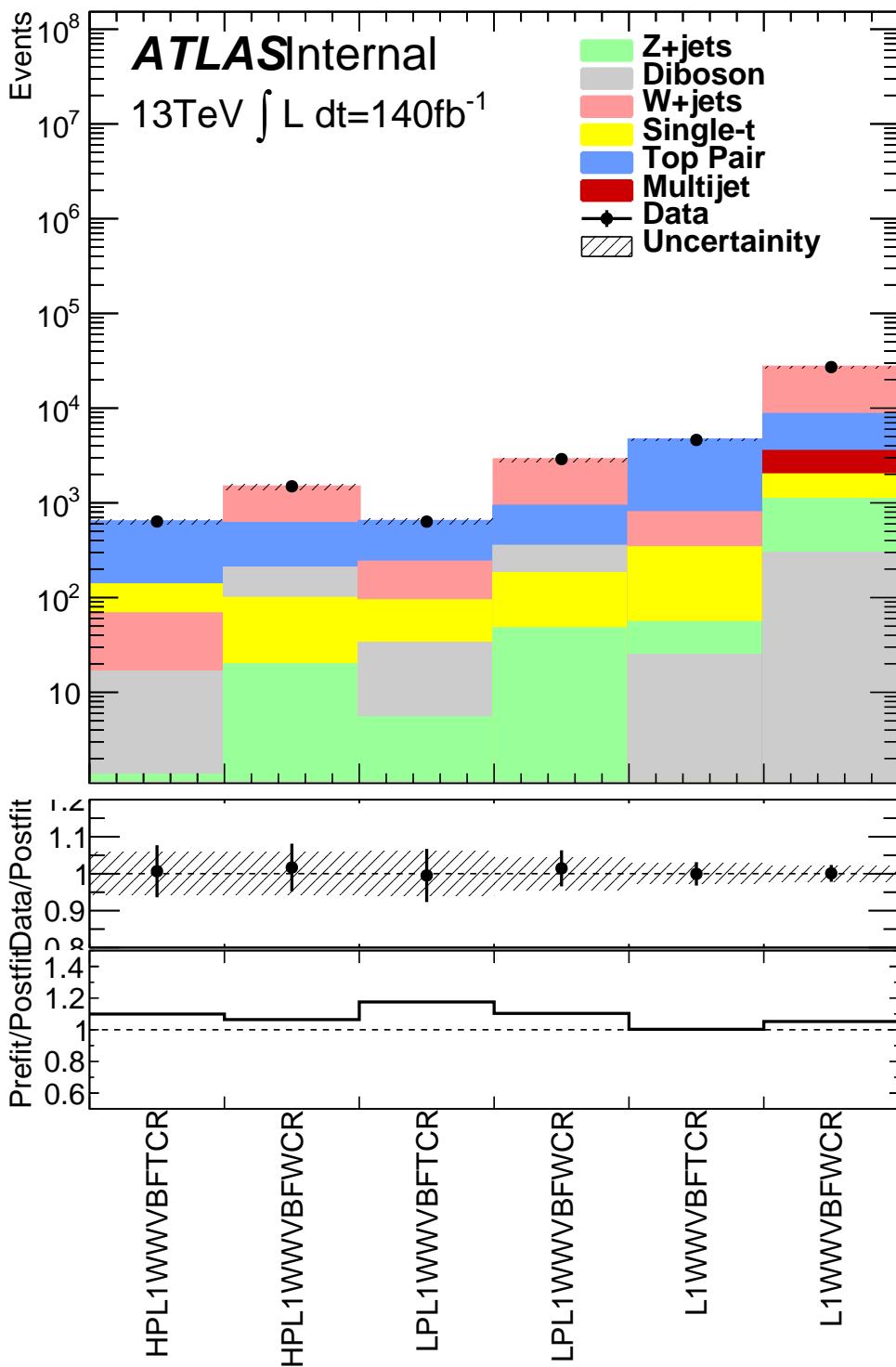


Figure 10.11: This figure shows the distribution of $m_{\ell\nu qq}$ in the VBF WW control regions.

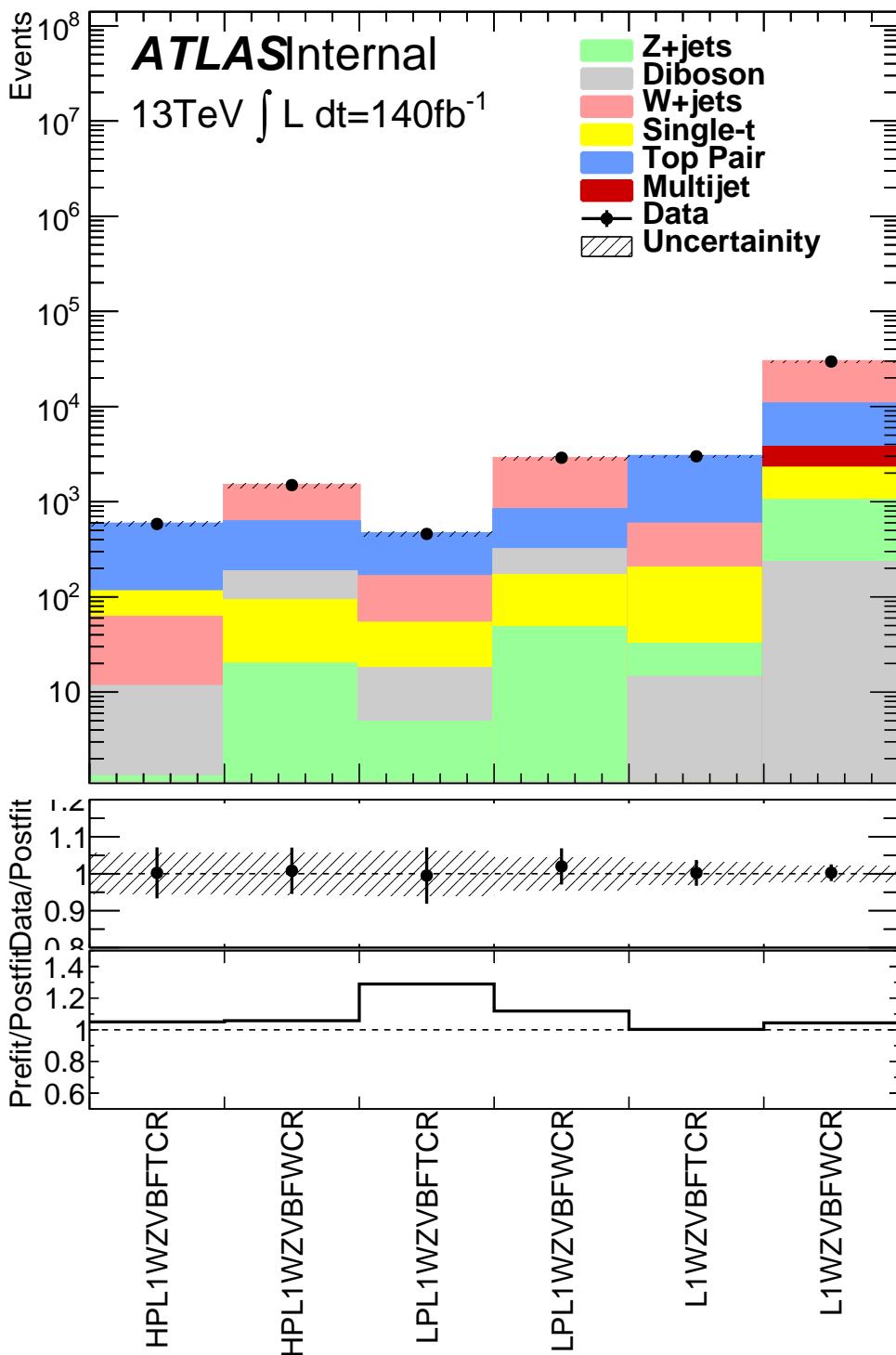


Figure 10.12: This figure shows the distribution of $m_{\ell\nu qq}$ in the VBF WZ control regions.

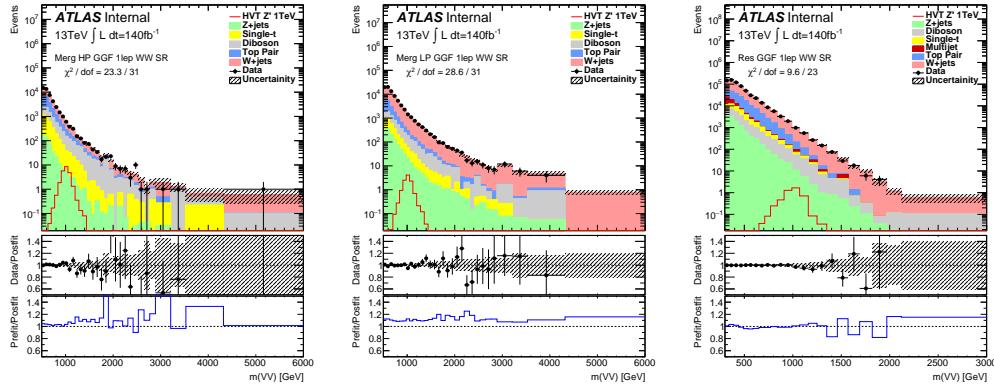


Figure 10.13: This figure shows the distribution of $m_{\ell\nu qq}$ in the GGF WW signal regions.

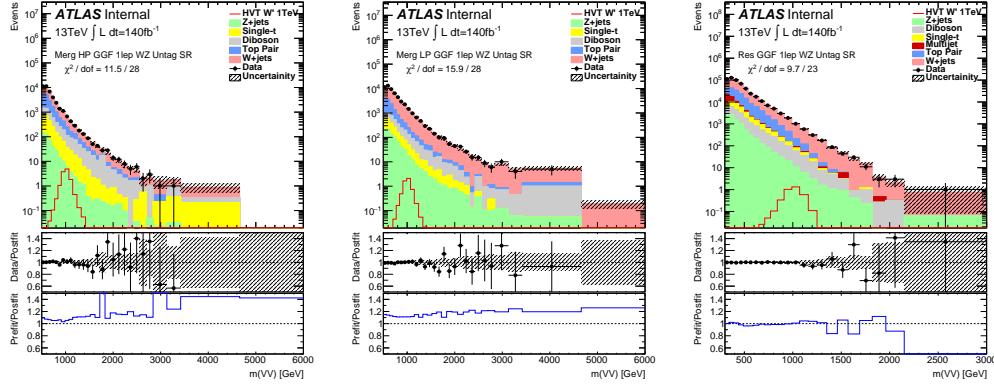


Figure 10.14: This figure shows the distribution of $m_{\ell\nu qq}$ in the GGF WZ Untag signal regions.

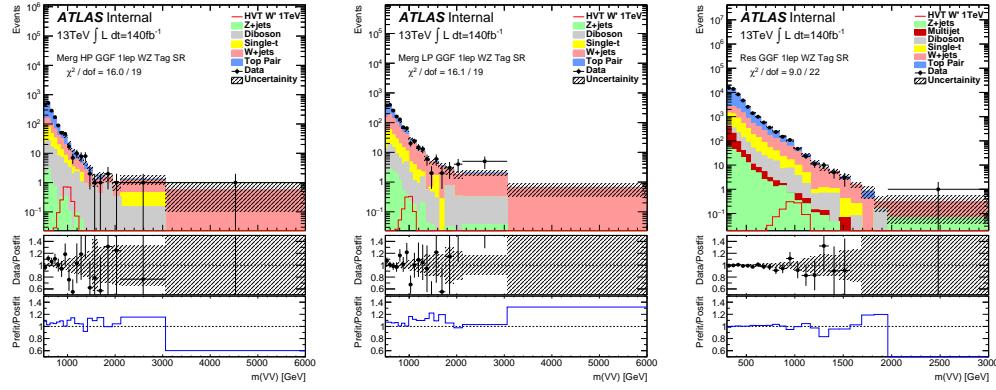


Figure 10.15: This figure shows the distribution of $m_{\ell\nu qq}$ in the GGF WZ Tag signal regions.

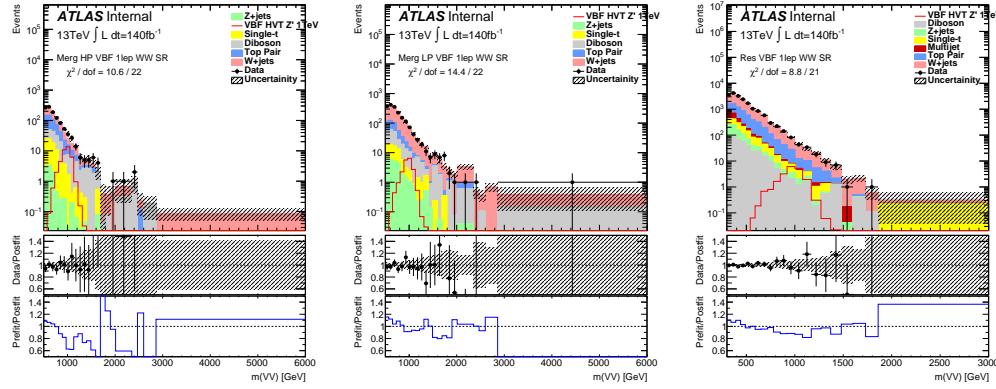


Figure 10.16: This figure shows the distribution of $m_{\ell\nu qq}$ in the VBF WZ Tag signal regions.

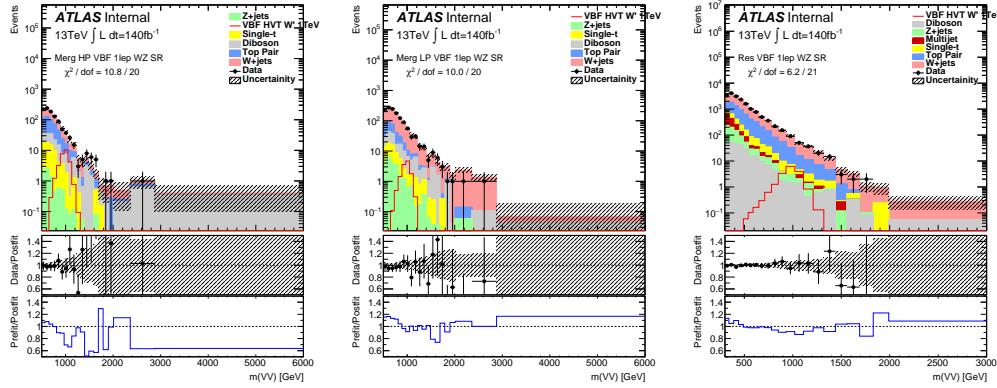


Figure 10.17: This figure shows the distribution of $m_{\ell\nu qq}$ in the VBF WZ Tag signal regions.

1189 10.4 Limits

1190 Using the exclusion limits tests discussed previously, exclusion limits are set
 1191 on μ and consequently cross-sections for different signal models. Exclusion limits
 1192 for the models considered are shown in Figure 10.18 - 10.20. These limits exclude
 1193 HVT Model A $W' < \text{blah}$ and $Z' < \text{blah}$ and Model B $W' < \text{blah}$ and $Z' < \text{blah}$.
 1194 Randall Sundrum Gravitons are excluded for masses below blah.

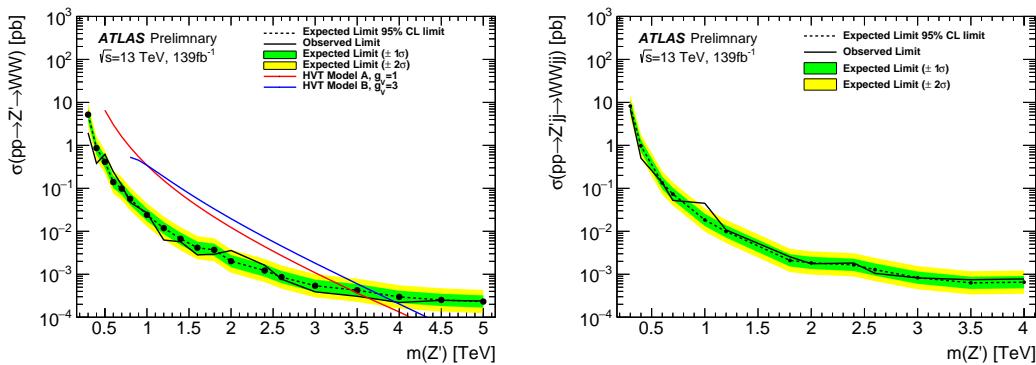


Figure 10.18: This figure shows theory, expected and observed limits for HVT W' DY (left) and VBF (right) production.

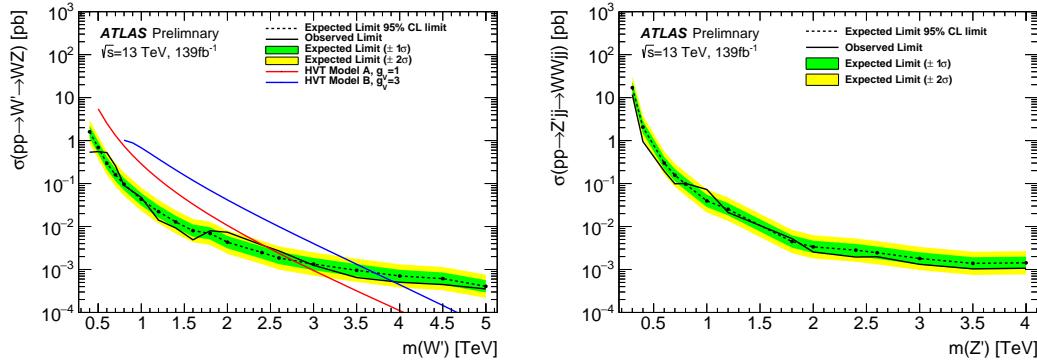


Figure 10.19: This figure shows theory, expected and observed limits for HVT Z' DY (left) and VBF (right) production.

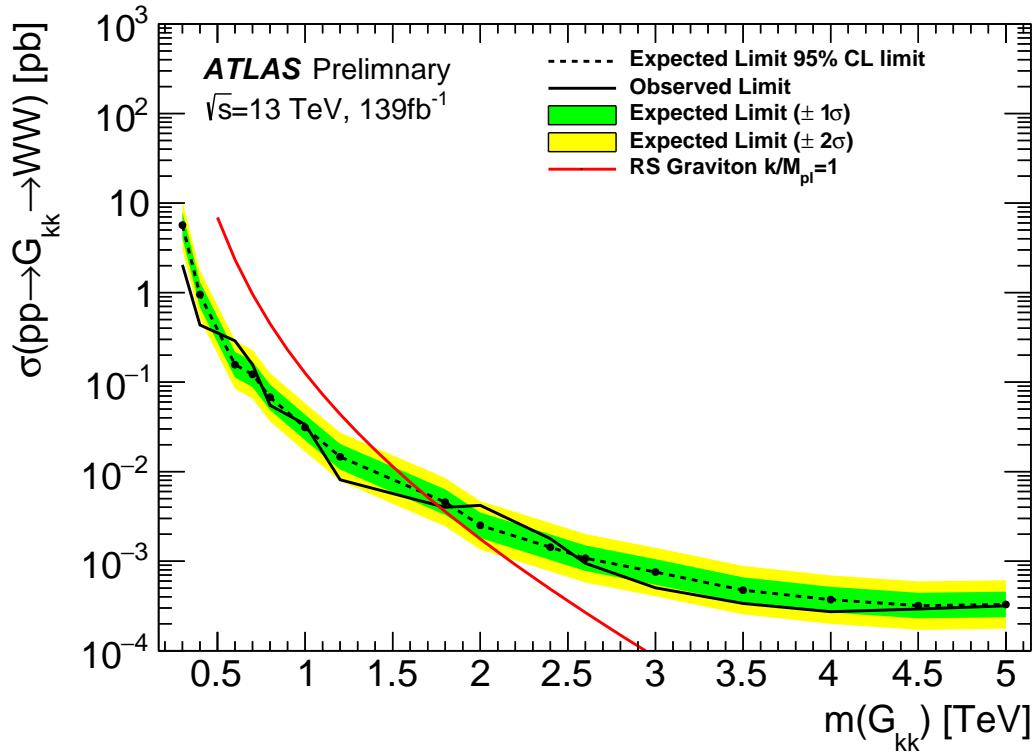


Figure 10.20: This figure shows theory, expected and observed limits for RS Gravitons via DY production.

Part V

1195

Quark and Gluon Tagging

1196

₁₁₉₇ **Chapter 11**

₁₁₉₈ **Prospects**

₁₁₉₉ For the resolved analysis, signal jets are quark enriched and background jets are
₁₂₀₀ gluon dominated. By classifying jets in the event as quark or gluon initiated, less
₁₂₀₁ background would contaminate the signal region. Figure 11.1 shows the PDGID
₁₂₀₂ for the truth parton matched to the jet (meaning the highest energy parton in
₁₂₀₃ the jet catchment area) in events passing the resolved signal region selections.
₁₂₀₄ PDGID = -1 corresponds to pileup jets, $0 < \text{PDGID} < 6$ correspond to quarks
₁₂₀₅ and $\text{PDGID} = 21$ corresponds to gluons. From this Figure, it is evident that a
₁₂₀₆ notable fraction of the background that contaminates the signal region contains
₁₂₀₇ gluon jets, especially for the sub-leading jet.

₁₂₀₈ As gluons jets have more constituents and therefore more tracks (n_{trk}), the
₁₂₀₉ background jets have more tracks than the signal jets. This is shown in Fig-
₁₂₁₀ ure 11.2. Therefore, by cutting on the number of tracks in a jet, quark and gluon
₁₂₁₁ jets may be distinguished (i.e. jets with less than a given number of tracks are
₁₂₁₂ classified as a quark, otherwise the jet is classified as a gluon.) Moreover, as the
₁₂₁₃ momentum of the jet increases the number of tracks also increases logarithmically.
₁₂₁₄ Therefore by applying a cut on the number of tracks that scales with the $\ln(p_T)$
₁₂₁₅ is more powerful than a threshold cut on the number of tracks. Figure 11.3-

1216 Figure 11.6 show normalized heat maps of $\ln(p_T)$ vs the number of reconstructed
1217 tracks for the background and a 300 GeV Z' signal. In these plots it is evident
1218 that the number of tracks in the background jets grows more quickly with $\ln(p_T)$
1219 than for the signal jets. This is expected given that the signal is quark dominated
1220 and the background is gluon dominated.

1221 In Figure 11.8 is the ROC Curve for quark gluon tagging with cut on the
1222 number of tracks in a jet that depends on $\ln(p_T)$. The sum of the backgrounds in
1223 the signal region were used for this curve. Here the quark tagging efficiency is the
1224 ratio of quarks tagged as quarks to the total number of quarks in the signal region.
1225 The gluon rejection is calculated as the reciprocal of the gluon tagging efficiency.
1226 Choosing a 90 efficient working point with a rejection of 1.4 corresponds to a slope
1227 of 4 and intercept of -5. Focusing on the background in Figure 11.9, this cut helps
1228 minimize gluon contamination in the signal region. Also, from these heat maps it
1229 is obvious that the number of tracks in gluon jets grows more quickly than those
1230 in quark jets.

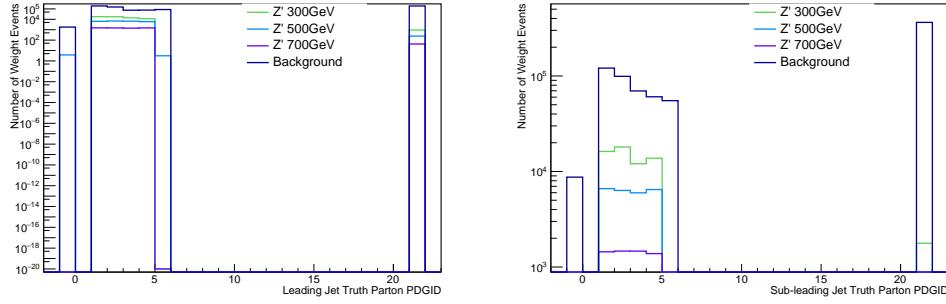


Figure 11.1: PDGID of the truth-level parton matched to the small-R jets passing the Resolved GGF WW Signal Region selections for the (a) Leading (b) Sub-Leading jets . These distributions are shown for 300, 500, and 700GeV Z' signals and the background.

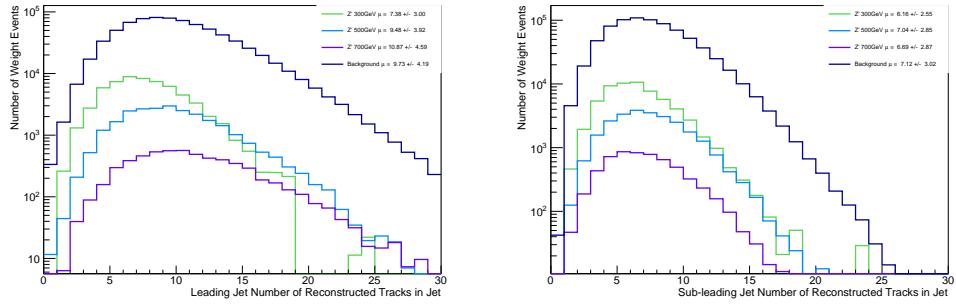


Figure 11.2: The number of tracks in small-R jets in events passing the Resolved GGF WW Signal Region selections for the (a) Leading (b) Sub-Leading jets. These distributions are shown for 300, 500, and 700GeV Z' signals and the background.

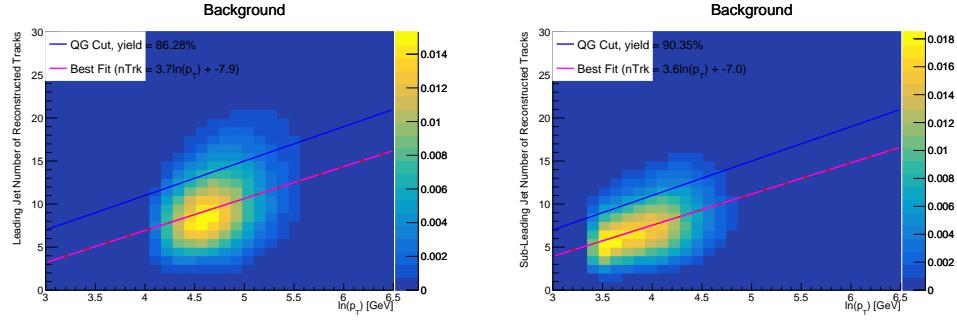


Figure 11.3: The number of tracks in background small-R jets in events passing the Resolved GGF WW Signal region selection vs. $\ln(p_T)$ for (a)Leading (b) Sub-Leading jets. The best fit line for the distribution is also shown, as well as the percentage of jets that pass a cut of number of tracks $< 4 \times \ln(p_T) - 5$. Note the number of total entries in these plots has been normalized to one.

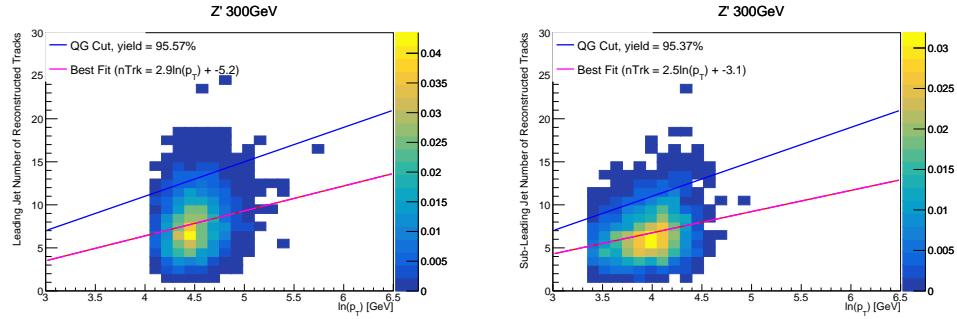


Figure 11.4: The number of tracks in small-R jets in 300GeV Z' events passing the Resolved GGF WW Signal region selection vs. $\ln(p_T)$ for (a)Leading (b) Sub-Leading jets. The best fit line for the distribution is also shown, as well as the percentage of jets that pass a cut of number of tracks $< 4 \times \ln(p_T) - 5$. Note the number of total entries in these plots has been normalized to one.

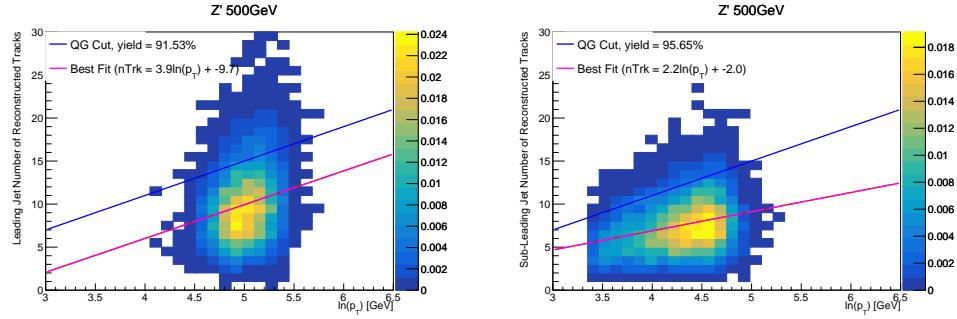


Figure 11.5: The number of tracks in small-R jets in 500GeV Z' events passing the Resolved GGF WW Signal region selection vs. $\ln(p_T)$ for (a)Leading (b) Sub-Leading jets. The best fit line for the distribution is also shown, as well as the percentage of jets that pass a cut of number of tracks $< 4 \times \ln(p_T) - 5$. Note the number of total entries in these plots has been normalized to one.

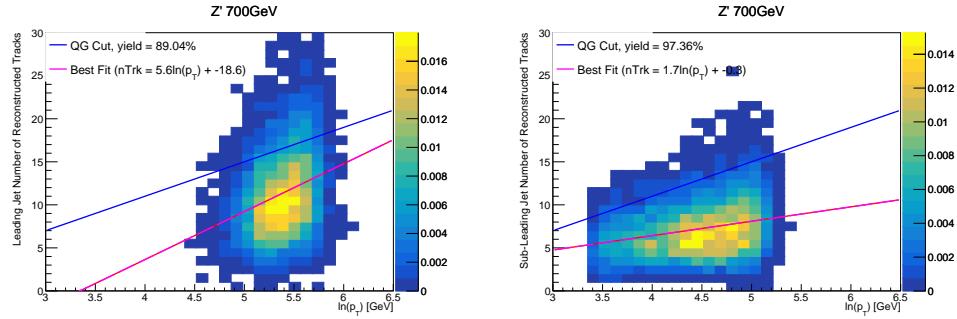


Figure 11.6: The number of tracks in small-R jets in 700GeV Z' events passing the Resolved GGF WW Signal region selection vs. $\ln(p_T)$ for (a)Leading (b) Sub-Leading jets. The best fit line for the distribution is also shown, as well as the percentage of jets that pass a cut of number of tracks $< 4 \times \ln(p_T) - 5$. Note the number of total entries in these plots has been normalized to one.

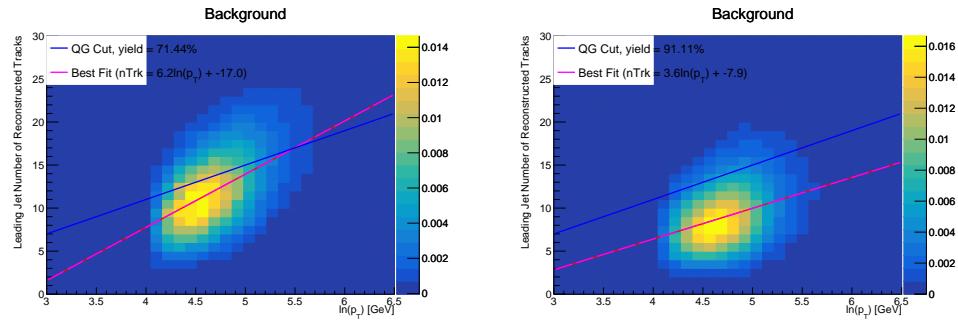


Figure 11.7: The number of tracks in leading small-R jets in background events passing the Resolved GGF WW Signal region selection vs. $\ln(p_T)$ for (a)Gluons (b) Quarks jets. The best fit line for the distribution is also shown, as well as the percentage of jets that pass a cut of number of tracks $< 4 \times \ln(p_T) - 5$.Note the number of total entries in these plots has been normalized to one.

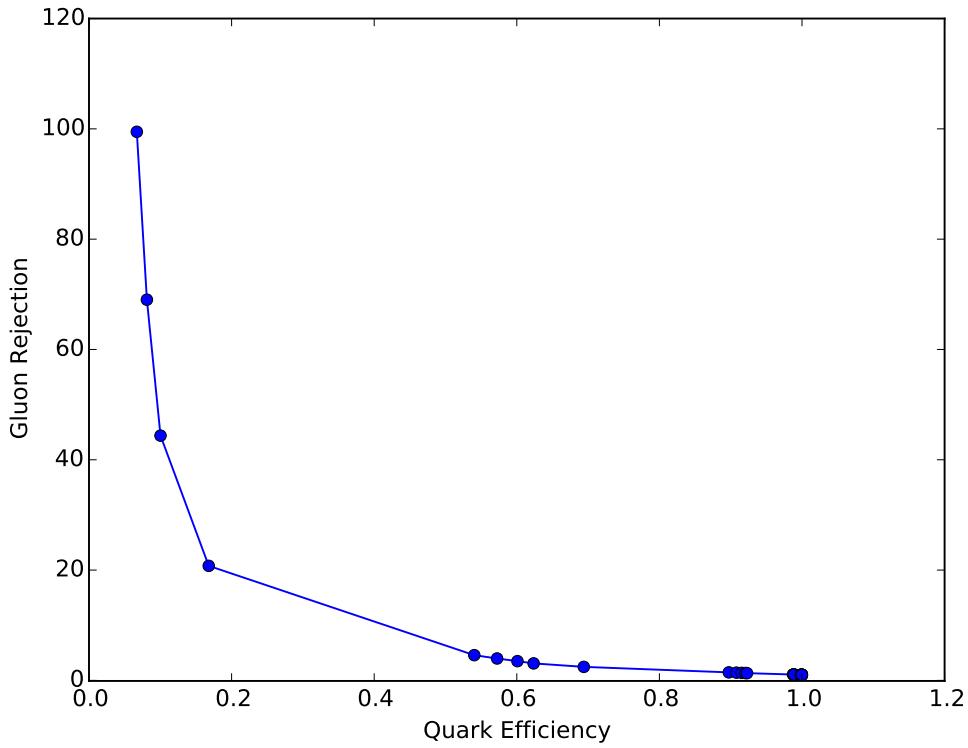


Figure 11.8: ROC Curve for Quark and Gluon Tagging with a cut on the number of tracks that depends on the $\ln(p_T)$.

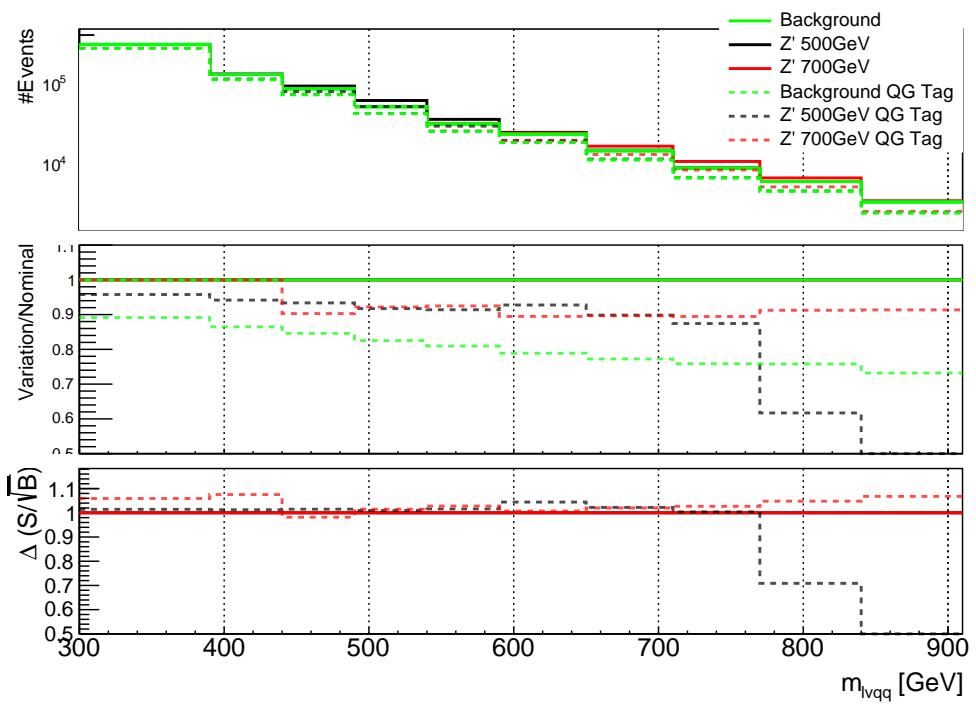


Figure 11.9: The top panel shows the distribution of m_{lvqq} with and without quark gluon tagging. The middle panel shows the ratio of the signals and backgrounds with and without quark gluon tagging. The bottom panel shows the change in S/\sqrt{B} with quark gluon tagging.

1231 Chapter 12

1232 n_{trk} Calibration

1233 As tagger based on nTrk cuts on the number of tracks in jets, a quantity that
1234 is not known with infinite precision, relevant systematic uncertainties must be
1235 evaluated. The sources of uncertainty in n_{trk} may be split into modeling and
1236 experimental uncertainties.

1237 Modeling uncertainties are obtained by assessing PDF and ME uncertainties
1238 on the number of charged particles in particle-level jets in dijet events. The
1239 number of charged particles as a function of jet p_T is calculated using an Iterative
1240 Bayesian (IB) technique [cite paper].

1241 This measurement ([7]) uses the ATLAS 2012 pp collision dataset, correspond-
1242 ing to $20.3/\text{fb}$ at center-of-mass energy $\sqrt{s} = 8\text{TeV}$. Monte Carlo (MC) samples
1243 are used to determine the response matrix. The MC sample is a dijet sample gen-
1244 erated with Pythia 8.175 using CT10 PDF and AU2 tune. The anti- k_T algorithm
1245 is used to cluster jets with a radius parameter $R = 0.4$. Jets are required to have
1246 $|\eta| < 2.1$. Tracks in jets are required to have $p_T > 500\text{MeV}$, $|\eta| < 2.5$, track-fit
1247 $\chi^2 < 3.0$ and originate from the primary vertex. Matching tracks to jets is accom-
1248 plished using ghost-association [cite]. In this technique, jets are re-clustered with
1249 the track collection augmented with "ghost" versions of tracks. These "ghosts"

1250 tracks have the same direction as their parent track, but infinitesimal track p_T .
1251 This insures meta-jet properties (e.g. η , p_T , etc) are unchanged. A track is
1252 matched to a jet if it's ghost version remains in the jet after re-clustering. Further
1253 details of the data, object, and event selection may be found in [cite 35].

1254 To select dijet topologies events are required to have at least two jets with
1255 $p_T > 50\text{GeV}$ that are relatively well-balanced ($p_T^{\text{lead}}/p_T^{\text{sub-leading}} < 1.5$).

1256 In the IB technique, the prior distribution and number of iterations are the
1257 inputs [cite Bayesian paper]. The IB response matrix connects number of charged
1258 particles to the number of tracks in jets determined using the simulated samples.
1259 This response matrix is used to unfold data to extract the n_c . Before applying
1260 the response matrix a fake factor is applied. This accounts for jets that pass
1261 detector level selections, but not particle level selections. Following this, the IB
1262 method iteratively applies the response matrix using the nominal Pythia 8.175
1263 sample as a prior. The number of IB iterations is chosen to minimize unfolding
1264 bias and statistical fluctuations. For this measurement four iterations was found
1265 to be optimal by minimizing the unfolding bias from pseudodata simulated with
1266 Herwig++ with a prior from Pythia 8 AU2. Finally, the inefficiency factor is
1267 applied to account for events passing particle level selection but not detector
1268 level, yielding the unfolded nCharged distribution.

1269 This process is prone to three main sources of bias: response matrix, correction
1270 factor, and unfolding procedure uncertainties. The response matrix is sensitive to
1271 experimental uncertainties impacting jet track reconstruction and calorimeter jet
1272 p_T . Correction factors are also sensitive to experimental uncertainties (e.g. JES)
1273 as such uncertainties modify detector level acceptance. Sensitivity to particle level
1274 acceptance is calculated by comparing Pythia and Herwig. Finally, the bias from
1275 the IB prior choice is determined by reweighting the particle-level spectrum, so

1276 the simulated detector level spectrum more closely matches the uncorrected data.
 1277 Unfolding this modified detector-level simulation and comparing it re-weighted
 1278 particle-level spectrum indicates bias from the prior distribution choice.

1279 A summary of all the systematic uncertainties associated with this unfolding
 1280 may be found in [ref paper]. Total uncertainties are < 7% for the number of
 1281 charged particles in jets. The unfolded distribution of the nCharged in jets from
 1282 data are further analyzed to extract the quark and gluon nCharged distributions.
 1283 In dijet events, the jet with a larger η is more energetic and therefore more likely
 1284 to be a quark. This is due to the quarks in protons generally having a larger
 1285 fraction of the total momentum of the proton constituents. The more central jet
 1286 is more likely to be a gluon-initiated jet. This correlation between jet η and flavor
 1287 may then be used to extract nCharged in p_T bins using:

$$\langle n_c^f \rangle = f_q^f \langle n_c^q \rangle + f_g^f \langle n_c^g \rangle \quad (12.1)$$

$$\langle n_c^c \rangle = f_q^c \langle n_c^q \rangle + f_g^c \langle n_c^g \rangle \quad (12.2)$$

1288 In this equation the f and c subscripts denote the more forward and central
 1289 jets, respectively. The q and g subscripts denote quark and gluon. The fraction
 1290 of more forward jets that are say gluons is denoted by f_g^f . The other relevant jet
 1291 fractions are denoted with the same naming scheme. Finally, $\langle n_c \rangle$ is the average
 1292 number of charged particles in a jet in a given p_T bin. To show that Eq. (??) may
 1293 be used to extract quark and gluon n_c distributions the extracted distributions
 1294 are compared to n_c distributions determined using the jet flavor in simulation.
 1295 Figure [add figure natasha] shows that the extracted and true distributions differ
 1296 by < 1% over the p_T ranged probed for this study. Moreover, this implies that n_c
 1297 depends only on the flavor of the initiating parton and jet p_T .

1298 These extracted distributions are prone to PDF and ME biases. The bias from

1300 the choice of the CT10 PDF for the Pythia sample is accounted for by comparing
1301 quark/gluon fractions for the nominal CT10 sample with its eigenvector variations.
1302 Comparing the quark/gluon fractions from Pythia 8 and Herwig++ quantify the
1303 uncertainty from the ME calculation. These uncertainties are added in quadra-
1304 ture with the unfolding uncertainty to give the total modelling uncertainty on
1305 the extracted n_c distribution. This is shown in Figure 13.2.

1306 To apply these uncertainties in n_c distributions in data, per-jet event weights
1307 are associated with each uncertainty according to:

$$w_i(n_c) = \frac{P(n_c | < n_c > \pm \sigma_{n_c}^i)}{P(n_c | < n_c >)} \quad (12.3)$$

1308 In Eq. (??), i denotes the uncertainty considered, P is the Poisson probability,
1309 and $\sigma_{n_c}^i$ represents the average impact of the uncertainty on n_c .

1310 The previous uncertainties described accounted for modeling uncertainty as-
1311 sociated with the number of charged particles in a jet. However, n_c is not a
1312 measurable quantity. Instead the number of tracks in a jet is measured, which is
1313 a proxy for n_c . Therefore the uncertainties associated with the measurement of
1314 nTracks must also be considered ([9]). These uncertainties were calculated using
1315 a Pythia 8 dijet sample with NNPDF 23. Track reconstruction efficiency and fake
1316 rates are the dominant sources of nTrack uncertainties.

1317 The track reconstruction efficiency is effected by the uncertainty of the de-
1318 scription of the ID material in simulation and the modeling of charged-particle
1319 interactions with this material. These uncertainties are accounted for by varying
1320 the ID material by 5-25% (dependent on the region of the detector considered).
1321 The difference in the tracking efficiency between the nominal and varied simula-
1322 tion give the uncertainty on the track reconstruction efficiency. Another important
1323 source of track reconstruction efficiency arises in the core of jets. The high density

1324 of tracks in the jet cores can cause ID clusters to merge. The fraction of lost tracks
 1325 due to merging is given by the fraction of tracks that have a charge of two mini-
 1326 mum ionizing particles. This quantity is compared between data and simulation
 1327 resulting in an uncertainty of 0.4% on tracks with $\Delta R < 0.1$. Combining these
 1328 effects gives a total uncertainty as a function of p_T and η that is generally $< 2\%$
 1329 [references figure 44 from [9]).]

1330 Fake tracks are the other dominant source of nTrk uncertainty. Fake tracks
 1331 are tracks that cannot be associated to a single particle. Often these tracks are a
 1332 result of random combinations of hits from charged particles that overlap in space.
 1333 In dense environments, such as the core of jets or high-pileup environments, fake
 1334 tracks are more likely. Fake tracks are estimated with a 'control region method'
 1335 which is briefly summarized here [[8]]. By applying a series of track selections
 1336 to enrich the fraction of fake tracks (e.g. $|d_0| > 0.1$, track $\chi^2 > 1.4$, etc) in
 1337 simulation, templates for fake track parameters are calculated. These templates
 1338 are then fit to data to determine the fraction of fake tracks. On average the fake
 1339 rate is found to be 30% (independent of p_T and η).

1340 To assess the impact of these two detector level uncertainties, tracks are ran-
 1341 domly dropped according to the rates described above. Reconstruction and fake
 1342 uncertainties both lower the number of tracks, hence these uncertainties are one-
 1343 sided. By dropping tracks in this way a varied nTrk distribution is calculated for
 1344 both uncertainties. The associated per-jet event weights are then calculated in
 1345 the same way as the modeling weights as:

$$w_i(n_c) = \frac{P(n_{trk} | < n_{trk} > \pm \sigma_{n_{trk}}^i)}{P(n_{trk} | < n_{trk} >)} \quad (12.4)$$

1346 Adding the modeling and detector level uncertainties in quadrature gives the
 1347 overall nTrack uncertainty. The effects of the individual uncertainties on the nTrk

1348 distributions can be seen in Fig 13.4. Fig 13.3 shows the m_{lvqq} and nTrk distri-
1349 butions for the W and Top control regions before likelihood fitting. In these plots
1350 the nTrk uncertainties improve agreement between data and MC. The remaining
1351 differences are likely covered by likelihood fitting and improving the analysis itself.

₁₃₅₂ **Chapter 13**

₁₃₅₃ **Application**

₁₃₅₄ Using the 90% WP of the n_{trk} tagger improves S/\sqrt{B} is $< 3\%$ as shown in
₁₃₅₅ Figure 11.9. Although, n_{trk} is the single most powerful discriminating variable
₁₃₅₆ for quark and gluon jets, the addition of other jet variables would improve the
₁₃₅₇ classification efficiency. Figure 13.1 shows the possible improvement of 10%
₁₃₅₈ in jet classification using the truth label of the jets to classify jets. This type of
₁₃₅₉ improvement is possible by using variables such as jet width, and energy correlata-
₁₃₆₀ tors. Figure [add BDT figure/use 1612.01551.pdf] shows for a 90% quark tagging
₁₃₆₁ efficiency for a 100 GeV jet, a BDT improve the gluon rejection by 0.4. Once this
₁₃₆₂ tagger is calibrated it would improve the analysis sensitivity of this channel.

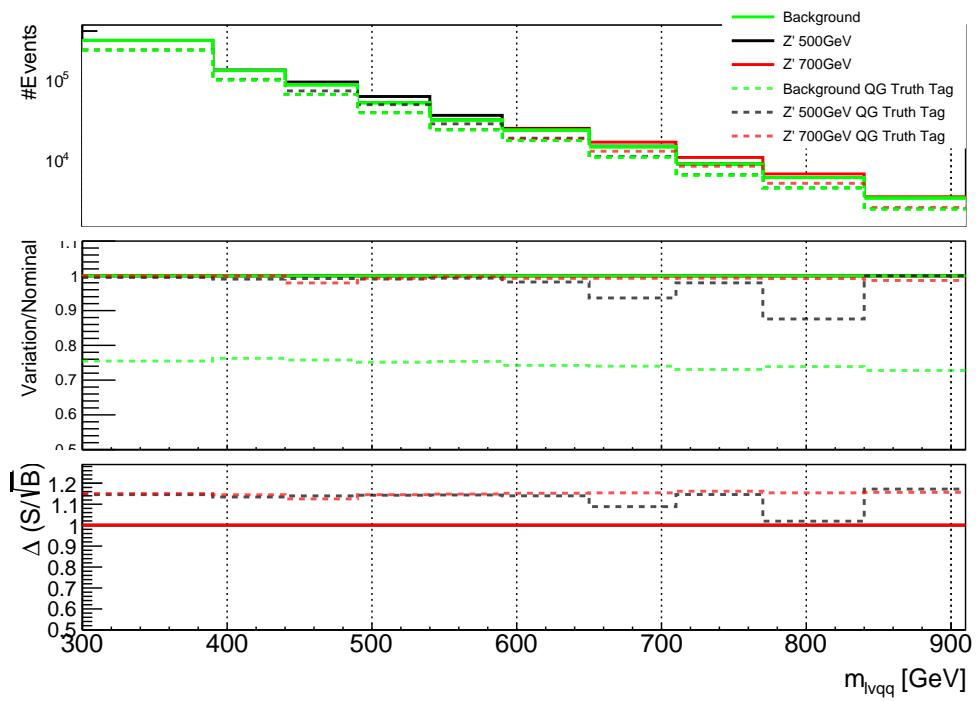


Figure 13.1: The top panel shows the distribution of m_{lvqq} with and without requiring jets to be true quarks. The middle panel shows the ratio of the signals and backgrounds with and without requiring jets to be true quarks. The bottom panel shows the change in S/\sqrt{B} when requiring jets to be true quarks..

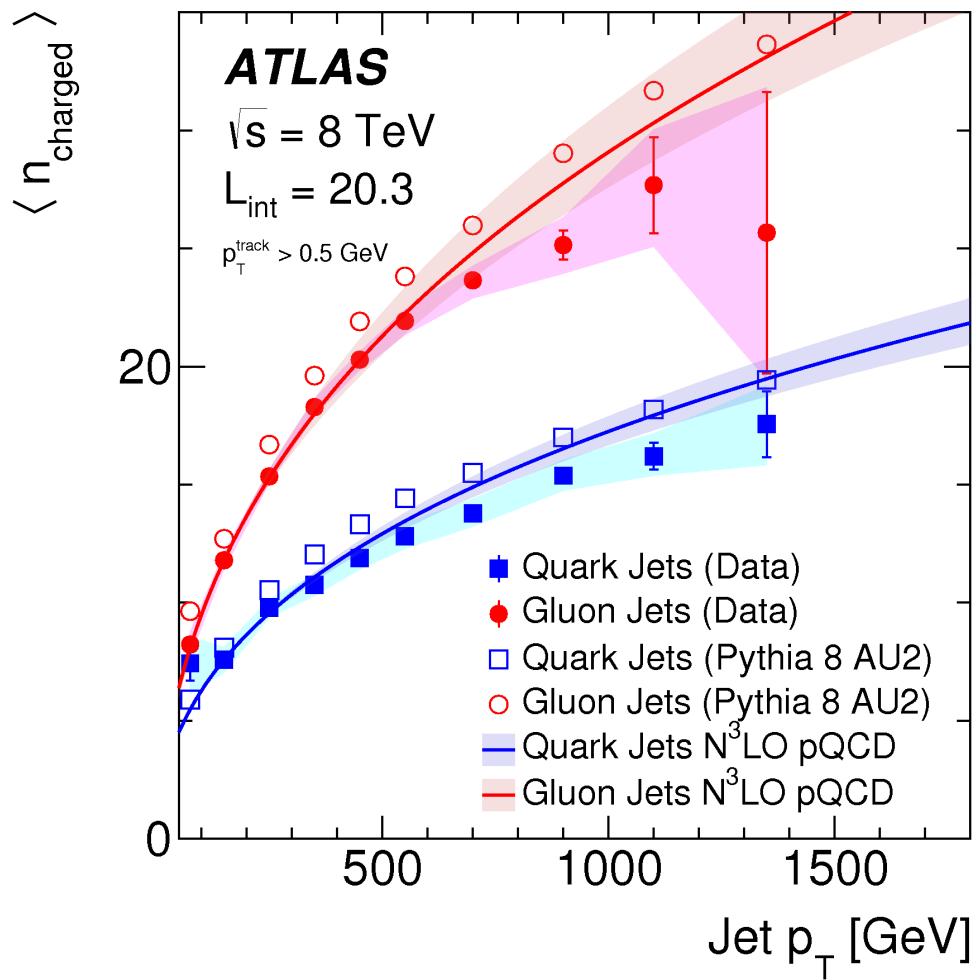


Figure 13.2: Unfolded and extracted n_C qg dstbs..

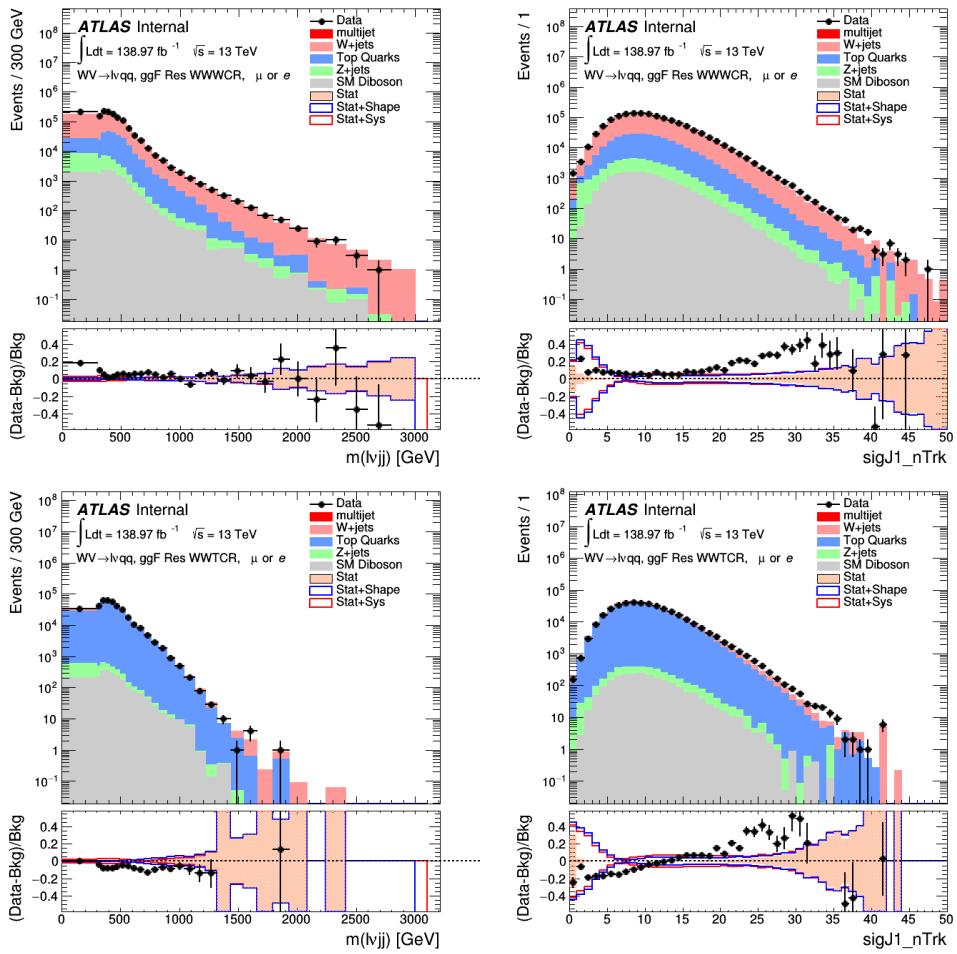


Figure 13.3: PDGID of the truth-level parton matched to the small-R jets passing the Resolved GGF WW Signal Region selections for the (a) Leading (b) Sub-Leading jets . These distributions are shown for 300, 500, and 700GeV Z' signals and the background.

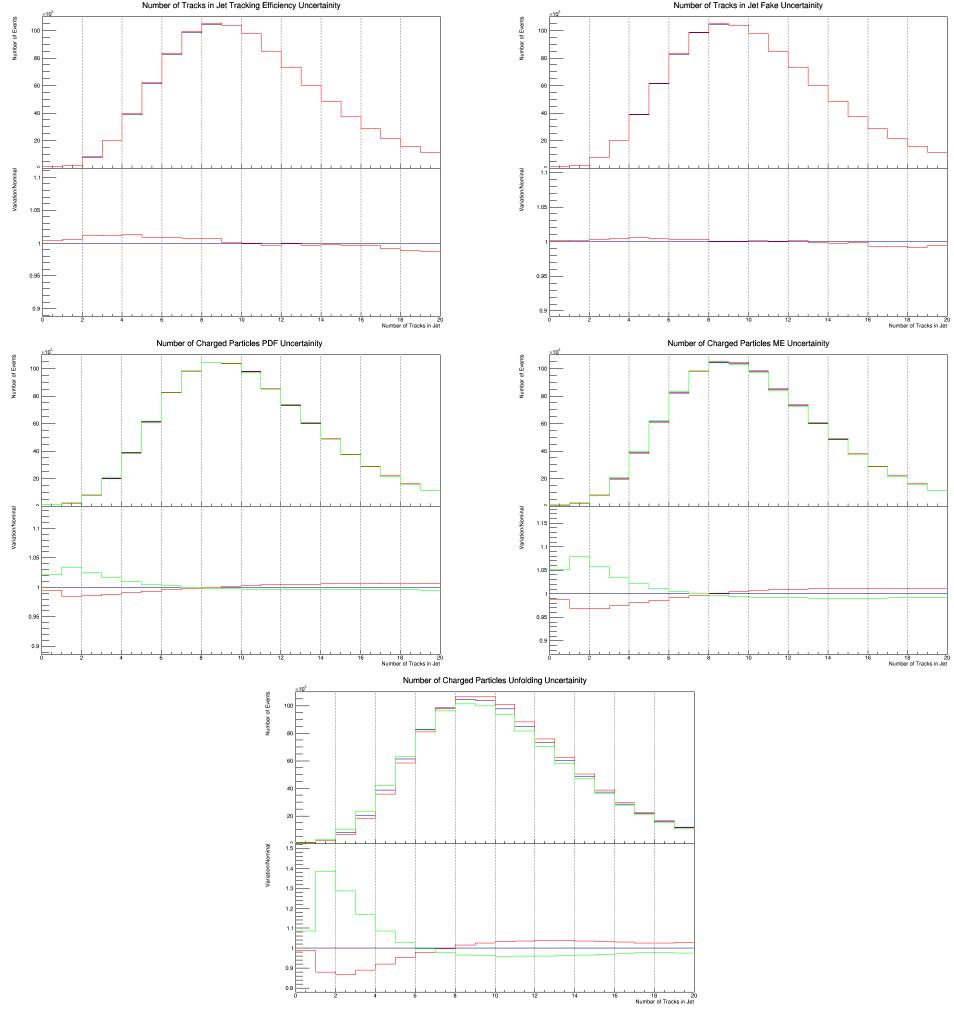


Figure 13.4: These figures show the impact of the uncertainties on the number of tracks in the leading jet in the sum of the background sample in the Resolved GGF WW SR (a) tracking efficiency (b) fake (c) PDF (d) ME (e) unfolding uncertainties.

Part VI

1363

Conclusion

1364

1365 Chapter 14

1366 Conclusions

1367 This is where conclusions go.

¹³⁶⁸ Bibliography

- ¹³⁶⁹ [1] Lecture notes particle physics ii.
- ¹³⁷⁰ [2] Kaustubh Agashe, Hooman Davoudiasl, Gilad Perez, and Amarjit Soni.
¹³⁷¹ Warped Gravitons at the LHC and Beyond. *Phys. Rev.*, D76:036006, 2007.
- ¹³⁷² [3] G. Altarelli and G. Parisi. Asymptotic freedom in parton language. *Nuclear*
¹³⁷³ *Physics B*, 126(2):298 – 318, 1977.
- ¹³⁷⁴ [4] ATLAS Collaboration. Atlas muon reconstruction performance in lhc run 2.
- ¹³⁷⁵ [5] ATLAS Collaboration. Summary plots from the atlas standard model physics
¹³⁷⁶ group.
- ¹³⁷⁷ [6] ATLAS Collaboration. Jet energy scale measurements and their systematic
¹³⁷⁸ uncer- tainties in proton–proton collisions at $\sqrt{s} = 13$ tev with the atlas
¹³⁷⁹ detector. arXiv: 1703.09665 [hep-ex].
- ¹³⁸⁰ [7] ATLAS Collaboration. Measurement of the charged-particle multiplicity
¹³⁸¹ inside jets from $s=\sqrt{8}$ tev pp collisions with the atlas detector.
¹³⁸² arXiv:1602.00988 [hep-ex].
- ¹³⁸³ [8] ATLAS Collaboration. Performance of the atlas track reconstruction algo-
¹³⁸⁴ rithms in dense environments in lhc run 2. arXiv:1704.07983 [hep-ex].
- ¹³⁸⁵ [9] ATLAS Collaboration. Properties of jet fragmentation using charged par-
¹³⁸⁶ ticles measured with the atlas detector in pp collisions at $\sqrt{s} = 13$ tev.
¹³⁸⁷ arXiv:1906.09254 [hep-ex].
- ¹³⁸⁸ [10] Alex Dias and V. Pleitez. Grand unification and proton stability near the
¹³⁸⁹ peccei-quinn scale. *Physical Review D*, 70, 07 2004.
- ¹³⁹⁰ [11] Stefan Höche, Frank Krauss, Marek Schönherr, and Frank Siegert. Qcd ma-
¹³⁹¹ trix elements + parton showers. the nlo case. *Journal of High Energy Physics*,
¹³⁹² 2013(4), Apr 2013.
- ¹³⁹³ [12] Gregory Soyez Matteo Cacciari, Gavin P. Salam. The anti- k_T jet clustering
¹³⁹⁴ algorithm. arXiv:0802.1189 [hep-ph].

- 1395 [13] Duccio Pappadopulo, Andrea Thamm, Riccardo Torre, and Andrea Wulzer.
1396 Heavy vector triplets: bridging theory and data. *Journal of High Energy*
1397 *Physics*, 2014(9), Sep 2014.
- 1398 [14] Antonio Pich. The Standard Model of Electroweak Interactions. In *Proceed-*
1399 *ings, High-energy Physics. Proceedings, 18th European School (ESHEP 2010):*
1400 *Raseborg, Finland, June 20 - July 3, 2010*, pages 1–50, 2012. [,1(2012)].
- 1401 [15] Lisa Randall and Raman Sundrum. A Large mass hierarchy from a small
1402 extra dimension. *Phys. Rev. Lett.*, 83:3370–3373, 1999.
- 1403 [16] Tania Robens and Tim Stefaniak. Lhc benchmark scenarios for the real higgs
1404 singlet extension of the standard model. *The European Physical Journal C*,
1405 76(5), May 2016.