### SEARCH FOR AN INVISIBLY DECAYING HIGGS BOSON PRODUCED VIA VECTOR BOSON FUSION

Rami Sesha Vanguri

A DISSERTATION

in

Physics and Astronomy

 $\label{eq:presented} \begin{array}{c} \mbox{Presented to the Faculties of The University of Pennsylvania} \\ \mbox{in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy} \\ 2015 \end{array}$ 

Elliot Lipeles, Associate Professor, Physics Supervisor of Dissertation

> Marija Drndic, Professor, Physics Graduate Group Chairperson

Dissertation Committee I. Joseph Kroll, Professor, Physics Adam Lidz, Assistant Professor, Physics Burt Ovrut, Professor, Physics Hugh Williams, Professor, Physics Elliot Lipeles, Associate Professor, Physics

# SEARCH FOR AN INVISIBLY DECAYING HIGGS BOSON PRODUCED VIA VECTOR BOSON FUSION

COPYRIGHT 2015 Rami Sesha Vanguri

## Acknowledgements

My experience at Penn was both extremely educational and inspiring. First and foremost, I would like to thank my adviser, Elliot Lipeles, for carefully guiding me through the techniques of particle physics, starting in 2006, when I was an undergraduate and he was a postdoc. His level of intuition and creativity is something I strive for everyday. I would also thank the other Penn ATLAS professors, Joe Kroll, Evelyn Thompson and Brig Williams for forming the backbone of such an exceptional research group. Of course, without the Tier 3 and Paul Keener's vigilance, this thesis would not have come to fruition. Jean O'Boyle also deserves thanks, for always making sure everything is taken care of, regardless of timescale.

The Penn army provided a sense of community at CERN, as well as plenty of adventures and antics. Specifically, I would like to thank John Alison, Kurt Brendlinger, Sarah Heim, Liz Hines, Tae Min Hong, Brett Jackson, Josh Kunkle, Chris Lester, Dominick Olivito, Ryan Reece, Jamie Saxon, Doug Schaefer, Jon Stahlman and Alex Tuna. A special thanks goes to Keisuke Yoshihara, for his friendship and help in my last 6 months of graduate school. He spent countless hours giving me physics advice and attending rehearsal sessions for my presentations. Additionally, I would like to thank Michael Mooney, Joe Taenzer and Vince Welke for their friendship during my trips to CERN.

I would also like to thank the friends I've made in my class at Penn for all of the productive conversations and for being exceptional drinking partners. In particular, I would like to thank Tom Caldwell, Garrett Goon and Matt Malloy. Special thanks goes to Matt Malloy, who despite being my roommate for 3 years, has never successfully murdered me in my sleep.

For making Philadelphia home, I would like to thank Koreem Blackwood, Reyna Boyer, Erin Brown, Dan Carr, George Costa, Donnie Galvano, Brandon Groff, Katie Louth, Andrew McCarthy, Brian McKay, Molly O'Rourke, Tiatu Thompson and Kip Waide. All of your company over the years has made it very difficult to leave. Without John Shahidi, who kept me caffeinated over the last 6 years, this thesis surely would not have been possible.

I was first inspired to pursue particle physics when I was an undergraduate at UCSD. For all they have taught me there and beyond, I would like to thank Matt Norman, Shih-Chieh Hsu, Mark Neubauer, Frank Golf and Johannes Mülmenstädt. I owe a debt of gratitude to my undergraduate adviser, Frank Würthwein, for giving me my first opportunity to do research on CDF, and providing invaluable advice throughout my career as a scientist.

Finally, I would like to thank my family for their support and encouragement, despite me living 2,500 miles away.

## ABSTRACT

#### SEARCH FOR AN INVISIBLY DECAYING HIGGS BOSON PRODUCED VIA VECTOR BOSON FUSION

Rami Sesha Vanguri

#### Elliot Lipeles

This thesis presents the first search of an invisibly decaying Higgs boson produced via Vector Boson Fusion on ATLAS. The dataset used for the analysis corresponds to  $20.3 \text{fb}^{-1}$  of proton-proton collisions at  $\sqrt{s} = 8$  TeV recorded at the Large Hadron Collider in 2011 and 2012. An upper bound limit is set at 95% confidence level on the invisible branching fraction of the Higgs Boson. A limit of 28% is observed (34% expected) and interpreted using the Higgs portal model to set a limit on the dark matter-nucleon cross section. The unique jet final state created by Vector Boson Fusion provides a stronger signal to background ratio than other invisibly decaying Higgs channels. The Vector Boson Fusion analysis presented resulted in the strongest constraint on dark matter production set by a hadron collider.

# Contents

A	ckno	wledge	ments	iii
A	bstra	ict		v
Co	onter	nts		vi
Li	st of	Table	5	xi
Li	st of	Figur	es	xv
A	utho	r's Coi	ntribution	xix
1	Intr	oducti	on	1
<b>2</b>	The	eoretic	al Motivation	3
	2.1	Introd	uction	3
	2.2	Stand	ard Model	3
	2.3	Higgs	Boson	5
	2.4	Issues	with the Standard Model	8
	2.5	Dark I	Matter	9
		2.5.1	Observational Evidence	9
		2.5.2	WIMP Hypothesis	10
		2.5.3	Direct Detection Experiments	11
		2.5.4	Dark Matter Detection with the Large Hadron Collider	13
		2.5.5	Dark Matter Portal	14
3	The	e Large	e Hadron Collider and the ATLAS Experiment	16

	3.1	Introd	uction	16
	3.2	Accele	erator Complex	16
	3.3	Beam	Parameters	18
	3.4	ATLA	S Detector	18
		3.4.1	Overview	18
		3.4.2	Coordinate System	21
		3.4.3	Inner Detector	21
		3.4.4	Calorimeters	24
		3.4.5	Muon Spectrometer	25
	3.5	Object	t Reconstruction	27
		3.5.1	Introduction	27
		3.5.2	Tracks and Calorimeter Clusters	27
		3.5.3	Electrons and Photons	29
		3.5.4	Muons	29
		3.5.5	Taus	29
		3.5.6	Jets	30
		3.5.7	<i>b</i> Jets	30
		3.5.8	Missing Energy	30
4	Trig	rger		32
-	41	Introd	uction	32
	4.2	Enhan	uced Bias	33
	4.3	Enhan	uced Bias Weighting Tool	35
	1.0			00
<b>5</b>	Ana	lysis S	Strategy	38
	5.1	Introd	uction	38
	5.2	Object	t Definitions	41
	5.3	Signal	Region	43
	5.4	Backg	rounds	45
		5.4.1	Strong vs Weak Production	47
	5.5	Simula	ation and Data	47
		5.5.1	Monte Carlo Samples	47
		5.5.2	Higher Order Corrections	48
		5.5.3	Data	48

	5.6	Background Composition	18
	5.7	Validation Regions	60
6	Lep	ton Control Regions 5	<b>2</b>
	6.1	Introduction	62
	6.2	$W \to \ell \nu + \text{jets Control Region}$	52
		6.2.1 Event Selection	53
		6.2.2 Multijet Background	<b>5</b> 4
		6.2.3 Yields	57
	6.3	$Z \to \ell \ell + jets$ Control Region	57
		6.3.1 Event Selection	68
		6.3.2 Yields	<b>i</b> 9
	6.4	Estimation of $V$ +jets	<u>;</u> 9
7	$\gamma + \mathbf{j}$	ets Control Region 6	5 <b>2</b>
	7.1	Introduction	52
	7.2	Event Selections	;3
		7.2.1 Selection Differences to Signal Region	53
		7.2.2 $\gamma$ +jets Control Region Selection	<b>5</b> 4
		7.2.3 Modified Signal Region Selection	55
		7.2.4 Modified $W$ +jets Control Region Selection	56
	7.3	Extrapolation of $\gamma$ +jets to $Z \rightarrow \nu \nu$ +jets	57
		7.3.1 Truth Level R Calculation	<b>i</b> 8
		7.3.2 Truth Level R Calculation Reweighting Test	<b>i</b> 8
		7.3.3 Theory Systematics on Extrapolation of $\gamma$ +jets to $Z \rightarrow \nu\nu$ +jets	70
	7.4	Processes in the $\gamma$ +jets Control Region	$^{\prime}2$
		7.4.1 $\gamma$ +jets Monte Carlo	$^{\prime}2$
		7.4.2 Background from Multijet	'3
	7.5	$\gamma$ +jets Control Region Yields	'5
		7.5.1 $W \to \ell \nu + \text{jets Cross-Check} \dots \dots$	<b>'</b> 6
		7.5.1.1 Signal and Background	'6
		7.5.1.2 $W \to \ell \nu + jets$ Modified Control Region Yields	7
	7.6	Anti-CJV Cross-Check	'8
		7.6.1 $\gamma$ +jets Anti-CJV Control Region Selection	'8

7.6.2       Modified Signal Anti-CJV Region Selection         7.6.3       Modified $W + jets$ Anti-CJV Control Region Selection         7.7 $\gamma + jets$ Anti-CJV Control Region Yields         7.8       Modified $W \rightarrow \ell\nu + jets$ Anti-CJV Control Region Yields         7.9       Comparison of Modeling for $\gamma + jets$ and $W \rightarrow \ell\nu + jets$ 7.10       Calculation of $Z \rightarrow \nu\nu + jets$ Process in Modified Signal Anti-CJV Region         7.11       Medium $m_{jj}$ Control Region Selection         7.11.1 $\gamma + jets$ Medium $m_{jj}$ Control Region Selection         7.11.2       Modified $W + jets$ Medium $m_{jj}$ Control Region Yields         7.11.2.1 $\gamma + jets$ Medium $m_{jj}$ Control Region Yields         7.12       Modified $W + jets$ Medium $m_{jj}$ Control Region Yields         7.12.1 $\gamma + jets$ Medium $m_{jj}$ Control Region Yields         7.12.1 $\gamma + jets$ Medium $m_{jj}$ Control Region Yields         7.13 $\gamma + jets$ Medium $m_{jj}$ Control Region Yields         7.14       Modeling Differences in Anti-CJV, Medium $m_{jj}$ and Signal Selections         7.14       Results from Efficiency Extrapolation Methods         7.15       VBF Normalization on Background Estimates         8       Multijet Control Regions         8.1       Introduction         8.2       Efficiency Method					
7.6.3       Modified $W$ +jets Anti-CJV Control Region Selection         7.7 $\gamma$ +jets Anti-CJV Control Region Yields         7.8       Modified $W \rightarrow \ell\nu$ +jets Anti-CJV Control Region Yields         7.9       Comparison of Modeling for $\gamma$ +jets and $W \rightarrow \ell\nu$ +jets         7.10       Calculation of $Z \rightarrow \nu\nu$ +jets Process in Modified Signal Anti-CJV Region         7.11       Medium $m_{jj}$ Cross-Check         7.11.1 $\gamma$ +jets Medium $m_{jj}$ Control Region Selection         7.11.2       Modified $W$ +jets Medium $m_{jj}$ Control Region Yields         7.11.2.1 $\gamma$ +jets Medium $m_{jj}$ Control Region Yields         7.11.2.1 $\gamma$ +jets Medium $m_{jj}$ Control Region Yields         7.12.2       Modified $W \rightarrow \ell\nu$ +jets Medium $m_{jj}$ and Signal Selections         7.12.1 $\gamma$ +jets Medium $m_{jj}$ and Signal Selections         7.13       Alternate Method: Anti-CJV Extrapolation         7.14       Results from Efficiency Extrapolation Methods         7.15       VBF Normalization on Background Estimates         8       Multijet Control Regions         8.1       Introduction         8.2       Efficiency Method         8.3       Jet $p_T$ Extrapolation Method         9.4       Introduction .         9.2       Theoretical Uncertainties on Signal <t< th=""><th></th><th></th><th>7.6.2</th><th>Modified Signal Anti-CJV Region Selection</th><th>80</th></t<>			7.6.2	Modified Signal Anti-CJV Region Selection	80
7.7 $\gamma$ +jets Anti-CJV Control Region Yields         7.8       Modified $W \rightarrow \ell\nu$ +jets Anti-CJV Control Region Yields         7.9       Comparison of Modeling for $\gamma$ +jets and $W \rightarrow \ell\nu$ +jets         7.10       Calculation of $Z \rightarrow \nu\nu$ +jets Process in Modified Signal Anti-CJV Region         7.11       Medium $m_{jj}$ Cross-Check         7.11.1 $\gamma$ +jets Medium $m_{jj}$ Control Region Selection         7.11.2       Modified $W$ +jets Medium $m_{jj}$ Control Region Yields         7.11.2.1 $\gamma$ +jets Medium $m_{jj}$ Control Region Yields         7.11.2.1 $\gamma$ +jets Medium $m_{jj}$ Control Region Yields         7.11.2.1 $\gamma$ +jets Medium $m_{jj}$ Control Region Yields         7.11.2.2       Modified $W \rightarrow \ell\nu$ +jets Medium $m_{jj}$ Control Region Yields         7.12.2       Modified $W \rightarrow \ell\nu$ +jets Medium $m_{jj}$ Control Region Yields         7.12.1 $\gamma$ +jets Medium $m_{jj}$ Control Region Yields         7.12.2       Modified W $\rightarrow \ell\nu$ +jets Medium $m_{jj}$ Control Region Yields         7.12.1 $\gamma$ +jets Medium $m_{jj}$ Control Region Yields         7.12.2       Modified W $\rightarrow \ell\nu$ +jets Medium $m_{jj}$ Control Region Yields         7.12.1 $\gamma$ +jets Medium $m_{jj}$ Control Region Yields         7.12       Modeling Differences in Anti-CJV Extrapolation Methods         7.15       VBF Normalization on Background Estimates			7.6.3	Modified $W$ +jets Anti-CJV Control Region Selection $\ldots \ldots \ldots \ldots \ldots$	81
7.8       Modified $W \rightarrow \ell \nu$ +jets Anti-CJV Control Region Yields         7.9       Comparison of Modeling for $\gamma$ +jets and $W \rightarrow \ell \nu$ +jets         7.10       Calculation of $Z \rightarrow \nu \nu$ +jets Process in Modified Signal Anti-CJV Region         7.11       Medium $m_{jj}$ Cross-Check         7.11 $\gamma$ +jets Medium $m_{jj}$ Control Region Selection         7.11.1 $\gamma$ +jets Medium $m_{jj}$ Control Region Selection         7.11.2       Modified $W$ +jets Medium $m_{jj}$ Control Region Yields         7.11.2.1 $\gamma$ +jets Medium $m_{jj}$ Control Region Selection         7.11.2.1 $\gamma$ +jets Medium $m_{jj}$ Control Region Yields         7.12       Modified $W \rightarrow \ell \nu$ +jets Medium $m_{jj}$ and Signal Selections         7.11.2.1 $\gamma$ +jets Medium $m_{jj}$ Control Region Yields         7.12       Modeling Differences in Anti-CJV, Medium $m_{jj}$ and Signal Selections         7.14       Results from Efficiency Extrapolation Methods         7.15       VBF Normalization         8	7	.7	$\gamma + \text{jets}$	Anti-CJV Control Region Yields	82
7.9       Comparison of Modeling for $\gamma$ +jets and $W \rightarrow \ell\nu$ +jets         7.10       Calculation of $Z \rightarrow \nu\nu$ +jets Process in Modified Signal Anti-CJV Region         7.11       Medium $m_{jj}$ Cross-Check         7.11. $\gamma$ +jets Medium $m_{jj}$ Control Region Selection         7.11. $\gamma$ +jets Medium $m_{jj}$ Control Region Selection         7.11. $\gamma$ +jets Medium $m_{jj}$ Control Region Selection         7.11.2       Modified $W$ +jets Medium $m_{jj}$ Control Region Yields         7.12.1 $\gamma$ +jets Medium $m_{jj}$ Control Region Yields         7.11.2.1 $\gamma$ +jets Medium $m_{jj}$ Control Region Yields         7.12.1.2       Modified $W \rightarrow \ell\nu$ +jets Medium $m_{jj}$ Control Region Yields         7.12.1.3       Alternate Method: Anti-CJV Extrapolation         7.13       Alternate Method: Anti-CJV Extrapolation         7.14       Results from Efficiency Extrapolation Methods         7.15       VBF Normalization         7.16       Impact of VBF Normalization on Background Estimates         8       Multijet Control Regions         8.1       Introduction         8.2       Efficiency Method         8.3       Jet $p_T$ Extrapolation Method         9.4       Theoretical Uncertainties on Signal         9.2.1       Signal Uncertainty due to Third Jet Veto	7	.8	Modif	ed $W \to \ell \nu + jets$ Anti-CJV Control Region Yields	83
7.10 Calculation of $Z \rightarrow \nu\nu$ +jets Process in Modified Signal Anti-CJV Region         7.11 Medium $m_{jj}$ Cross-Check         7.11. $\gamma$ +jets Medium $m_{jj}$ Control Region Selection         7.11.2 Modified W+jets Medium $m_{jj}$ Control Region Selection         7.11.2 Modified W+jets Medium $m_{jj}$ Control Region Selection         7.11.2 Modified W+jets Medium $m_{jj}$ Control Region Yields         7.11.2 Modified W $\rightarrow \ell\nu$ +jets Medium $m_{jj}$ Control Region Yields         7.12 Modeling Differences in Anti-CJV, Medium $m_{jj}$ and Signal Selections         7.13 Alternate Method: Anti-CJV Extrapolation         7.14 Results from Efficiency Extrapolation Methods         7.15 VBF Normalization         7.16 Impact of VBF Normalization on Background Estimates         8 Multijet Control Regions         8.1 Introduction         8.2 Efficiency Method         8.3 Jet $p_T$ Extrapolation Method         9 Systematic Uncertainties         9.1 Introduction         9.2 Theoretical Uncertainty due to Third Jet Veto         9.3 Theoretical Uncertainty due to Third Jet Veto         9.3 Shower Modeling         9.4.1 Variation of Factorization and Renormalization Scales         9.4.1 Jet Vertex Association	7	.9	Comp	arison of Modeling for $\gamma$ +jets and $W \rightarrow \ell \nu$ +jets	84
7.11 Medium $m_{jj}$ Cross-Check	7	.10	Calcul	ation of $Z \rightarrow \nu \nu + jets$ Process in Modified Signal Anti-CJV Region	85
7.11.1 $\gamma$ +jets Medium $m_{jj}$ Control Region Selection         7.11.2 Modified W+jets Medium $m_{jj}$ Control Region Selection         7.11.2.1 $\gamma$ +jets Medium $m_{jj}$ Control Region Yields         7.11.2.1 $\gamma$ +jets Medium $m_{jj}$ Control Region Yields         7.11.2.1 Modified $W \rightarrow \ell \nu$ +jets Medium $m_{jj}$ and Signal Selections         7.12 Modeling Differences in Anti-CJV, Medium $m_{jj}$ and Signal Selections         7.13 Alternate Method: Anti-CJV Extrapolation         7.14 Results from Efficiency Extrapolation Methods         7.15 VBF Normalization         7.16 Impact of VBF Normalization on Background Estimates         8 Multijet Control Regions         8.1 Introduction         8.2 Efficiency Method         8.3 Jet $p_T$ Extrapolation Method         9 Systematic Uncertainties         9.1 Introduction         9.2 Theoretical Uncertainties on Signal         9.3.1 Variation of Factorization and Renormalization Scales         9.3.2 PDF+ $\alpha_S$ 9.3.3 Shower Modeling         9.4 Experimental Uncertainties	7	.11	Mediu	m $m_{jj}$ Cross-Check	87
7.11.2 Modified $W$ +jets Medium $m_{jj}$ Control Region Selection			7.11.1	$\gamma$ +jets Medium $m_{jj}$ Control Region Selection	87
7.11.2.1 $\gamma$ +jets Medium $m_{jj}$ Control Region Yields         7.11.2.2 Modified $W \rightarrow \ell \nu$ +jets Medium $m_{jj}$ Control Region Yields         7.12 Modeling Differences in Anti-CJV, Medium $m_{jj}$ and Signal Selections         7.13 Alternate Method: Anti-CJV Extrapolation         7.14 Results from Efficiency Extrapolation Methods         7.15 VBF Normalization         7.16 Impact of VBF Normalization on Background Estimates         8 Multijet Control Regions         8.1 Introduction         8.2 Efficiency Method         8.3 Jet $p_T$ Extrapolation Method         9 Systematic Uncertainties         9.1 Introduction         9.2 Theoretical Uncertainties on Signal         9.3.1 Variation of Factorization and Renormalization Scales         9.3.3 Shower Modeling         9.3.4 Experimental Uncertainties         9.3.5 Shower Modeling         9.4.1 Jet Vertex Association			7.11.2	Modified $W$ +jets Medium $m_{jj}$ Control Region Selection	89
7.11.2.2       Modified $W \rightarrow \ell \nu + jets$ Medium $m_{jj}$ Control Region Yields         7.12       Modeling Differences in Anti-CJV, Medium $m_{jj}$ and Signal Selections         7.13       Alternate Method: Anti-CJV Extrapolation         7.14       Results from Efficiency Extrapolation Methods         7.15       VBF Normalization         7.16       Impact of VBF Normalization on Background Estimates         8       Multijet Control Regions         8.1       Introduction         8.2       Efficiency Method         8.3       Jet $p_T$ Extrapolation Method         9       Systematic Uncertainties         9.1       Introduction         9.2       Theoretical Uncertainties on Signal         9.3       Theoretical Uncertainties on Backgrounds         9.3.1       Variation of Factorization and Renormalization Scales         9.3.2       PDF+ $\alpha_S$ 9.3.3       Shower Modeling         9.4       Experimental Uncertainties         9.4.1       Jet Vertex Association				7.11.2.1 $\gamma$ +jets Medium $m_{jj}$ Control Region Yields	90
7.12 Modeling Differences in Anti-CJV, Medium $m_{jj}$ and Signal Selections         7.13 Alternate Method: Anti-CJV Extrapolation         7.14 Results from Efficiency Extrapolation Methods         7.15 VBF Normalization         7.16 Impact of VBF Normalization on Background Estimates         8 Multijet Control Regions         8.1 Introduction         8.2 Efficiency Method         8.3 Jet $p_T$ Extrapolation Method         9 Systematic Uncertainties         9.1 Introduction         9.2 Theoretical Uncertainties on Signal         9.3 Theoretical Uncertainties on Backgrounds         9.3.1 Variation of Factorization and Renormalization Scales         9.3.2 PDF+ $\alpha_S$ 9.3.3 Shower Modeling         9.4 Experimental Uncertainties				7.11.2.2 Modified $W \to \ell \nu + \text{jets}$ Medium $m_{jj}$ Control Region Yields	90
7.13 Alternate Method: Anti-CJV Extrapolation         7.14 Results from Efficiency Extrapolation Methods         7.15 VBF Normalization         7.16 Impact of VBF Normalization on Background Estimates         8 Multijet Control Regions         8.1 Introduction         8.2 Efficiency Method         8.3 Jet $p_T$ Extrapolation Method         9 Systematic Uncertainties         9.1 Introduction         9.2 Theoretical Uncertainties on Signal         9.3 Theoretical Uncertainties on Backgrounds         9.3.1 Variation of Factorization and Renormalization Scales         9.3.3 Shower Modeling         9.4 Experimental Uncertainties         9.4 Experimental Uncertainties	7	.12	Model	ing Differences in Anti-CJV, Medium $m_{jj}$ and Signal Selections $\ldots \ldots \ldots$	91
7.14 Results from Efficiency Extrapolation Methods         7.15 VBF Normalization         7.16 Impact of VBF Normalization on Background Estimates         8 Multijet Control Regions         8.1 Introduction         8.2 Efficiency Method         8.3 Jet $p_T$ Extrapolation Method         9 Systematic Uncertainties         9.1 Introduction         9.2 Theoretical Uncertainties on Signal         9.3 Theoretical Uncertainty due to Third Jet Veto         9.3 Theoretical Uncertainties on Backgrounds         9.3.1 Variation of Factorization and Renormalization Scales         9.3.2 PDF+ $\alpha_S$ 9.3.3 Shower Modeling         9.4 Experimental Uncertainties         9.4.1 Jet Vertex Association	7	.13	Altern	ate Method: Anti-CJV Extrapolation	92
7.15 VBF Normalization       7.16 Impact of VBF Normalization on Background Estimates         8 Multijet Control Regions         8.1 Introduction         8.2 Efficiency Method         8.3 Jet $p_T$ Extrapolation Method         9 Systematic Uncertainties         9.1 Introduction         9.2 Theoretical Uncertainties on Signal         9.1 Signal Uncertainty due to Third Jet Veto         9.3 Theoretical Uncertainties on Backgrounds         9.3.1 Variation of Factorization and Renormalization Scales         9.3.2 PDF+ $\alpha_S$ 9.3.3 Shower Modeling         9.4 Experimental Uncertainties         9.4 Experimental Uncertainties	7	.14	Result	s from Efficiency Extrapolation Methods	93
7.16 Impact of VBF Normalization on Background Estimates         8 Multijet Control Regions         8.1 Introduction         8.2 Efficiency Method         8.3 Jet $p_T$ Extrapolation Method         9 Systematic Uncertainties         9.1 Introduction         9.2 Theoretical Uncertainties on Signal         9.3 Theoretical Uncertainty due to Third Jet Veto         9.3 Theoretical Uncertainties on Backgrounds         9.3.1 Variation of Factorization and Renormalization Scales         9.3.2 PDF+ $\alpha_S$ 9.3.3 Shower Modeling         9.4 Experimental Uncertainties         9.4.1 Jet Vertex Association	7	.15	VBF I	Normalization	93
8 Multijet Control Regions         8.1 Introduction         8.2 Efficiency Method         8.3 Jet $p_T$ Extrapolation Method         8.3 Jet $p_T$ Extrapolation Method         9 Systematic Uncertainties         9.1 Introduction         9.2 Theoretical Uncertainties on Signal         9.3 Theoretical Uncertainty due to Third Jet Veto         9.3 Theoretical Uncertainties on Backgrounds         9.3.1 Variation of Factorization and Renormalization Scales         9.3.2 PDF+ $\alpha_S$ 9.3.3 Shower Modeling         9.4 Experimental Uncertainties         9.4.1 Jet Vertex Association	7	.16	Impac	t of VBF Normalization on Background Estimates	95
8.1       Introduction         8.2       Efficiency Method         8.3       Jet $p_T$ Extrapolation Method         9       Systematic Uncertainties         9.1       Introduction         9.2       Theoretical Uncertainties on Signal         9.2.1       Signal Uncertainty due to Third Jet Veto         9.3       Theoretical Uncertainties on Backgrounds         9.3.1       Variation of Factorization and Renormalization Scales         9.3.2       PDF+ $\alpha_S$ 9.3.3       Shower Modeling         9.4       Experimental Uncertainties         9.4.1       Jet Vertex Association	N	ſul	tijet (	Control Regions	97
8.2 Efficiency Method	8	.1	Introd	uction	97
8.3 Jet $p_T$ Extrapolation Method	8	.2	Efficie	ncy Method	98
<ul> <li>9 Systematic Uncertainties</li> <li>9.1 Introduction</li></ul>	8	.3	Jet $p_T$	Extrapolation Method	103
<ul> <li>9 Systematic Uncertainties</li> <li>9.1 Introduction</li></ul>					
9.1       Introduction         9.2       Theoretical Uncertainties on Signal         9.2.1       Signal Uncertainty due to Third Jet Veto         9.3       Theoretical Uncertainties on Backgrounds         9.3       Theoretical Uncertainties on Backgrounds         9.3.1       Variation of Factorization and Renormalization Scales         9.3.2       PDF+ $\alpha_S$ 9.3.3       Shower Modeling         9.4       Experimental Uncertainties         9.4.1       Jet Vertex Association	S	yst	emati	c Uncertainties	105
<ul> <li>9.2 Theoretical Uncertainties on Signal</li></ul>	9.	.1	Introd		105
<ul> <li>9.2.1 Signal Uncertainty due to Third Jet Veto</li> <li>9.3 Theoretical Uncertainties on Backgrounds</li> <li>9.3 Theoretical Uncertainties on Backgrounds</li> <li>9.3.1 Variation of Factorization and Renormalization Scales</li> <li>9.3.2 PDF+α<sub>S</sub></li> <li>9.3.3 Shower Modeling</li> <li>9.4 Experimental Uncertainties</li> <li>9.4.1 Jet Vertex Association</li> <li>9.4 Limit Setting</li> </ul>	9.	.2	Theor	etical Uncertainties on Signal	105
<ul> <li>9.3 Theoretical Uncertainties on Backgrounds</li></ul>			9.2.1	Signal Uncertainty due to Third Jet Veto	107
9.3.1       Variation of Factorization and Renormalization Scales         9.3.2       PDF+ $\alpha_S$ 9.3.3       Shower Modeling         9.4       Experimental Uncertainties         9.4.1       Jet Vertex Association         9.4.1       Setting	9.	.3	Theor	etical Uncertainties on Backgrounds	108
9.3.2 $PDF+\alpha_S$ 9.3.3 Shower Modeling 9.4 Experimental Uncertainties 9.4.1 Jet Vertex Association			9.3.1	Variation of Factorization and Renormalization Scales	108
9.3.3 Shower Modeling			9.3.2	$PDF + \alpha_S$	111
<ul> <li>9.4 Experimental Uncertainties</li></ul>			9.3.3	Shower Modeling	112
9.4.1 Jet Vertex Association	9.	.4	Exper	imental Uncertainties	114
10 Limit Satting			9.4.1	Jet Vertex Association	117
	0 L	im	it Sett	ing	121

	0.1 Background Estimation	121
	0.2 Likelihood	122
	0.3 Profile Likelihood Test Statistic	123
	0.4 $\operatorname{CL}_S$ Method	123
	0.5 Inputs to the Limit	124
	0.6 Results	126
11 (	Conclusions	131
11	Conclusions 1.1 Comparison to ZH	<b>131</b> 131
11	Conclusions         1.1 Comparison to ZH         1.2 Comparison to Higgs Couplings Analysis	<b>131</b> 131 131
11	Conclusions         1.1 Comparison to ZH         1.2 Comparison to Higgs Couplings Analysis         1.3 Dark Matter Portal	<b>131</b> 131 131 132
	Conclusions         1.1 Comparison to ZH         1.2 Comparison to Higgs Couplings Analysis         1.3 Dark Matter Portal         1.4 Future of Invisibly Decaying Higgs Searches	<ol> <li>131</li> <li>131</li> <li>131</li> <li>132</li> <li>133</li> </ol>

# List of Tables

5.1 5.2	Progressive yields of the selection in 20.3 fb <sup>-1</sup> of 2012 data, as evaluated using Monte Carlo. The expected signal yields are shown for the case of $m_H = 125$ GeV. Even though the ratio of signal and background is higher before the $E_T^{\text{miss}}$ requirement, events with $E_T^{\text{miss}} < 150$ GeV are not used due to difficulty in estimating the trigger efficiency Progressive yields of the validation regions described in Section 5.7 in 20.3 fb <sup>-1</sup> of 2012 data, as evaluated using Monte Carlo. Expected signal yields are shown for the case of $m_H = 125$ GeV. The "Subtracted" column shows an estimate of the dijet contribution by subtracting Z, W and Other Background columns from the observed yield	45 51
6.1	Expected yields for the multijet enhanced control regions used to estimate the multijet	50
6.2	contamination in the $W \to \ell\nu$ +jets control region	56
6.3	multijet contributions are from the multijet enhanced control regions	56
64	and other contributions are estimated using MC	57
0.1	processes are estimated with MC.	59
6.5	Estimates of $Z \to \nu\nu$ +jets as calulated from the W and Z control regions	60
$\begin{array}{c} 6.6 \\ 6.7 \end{array}$	Estimates of $W \to \ell \nu + \text{jets}$ as calulated from the W and Z control regions Progressive yields of the validation regions described in Section 5.7 in 20.3 fb <sup>-1</sup> of 2012 data, as evaluated using data driven estimates for $Z \to \nu \nu + \text{jets}$ and $W \to \ell \nu + \text{jets}$ and Monte	60
	Carlo for other processes. Expected signal yields are shown for the case of $m_H = 125$ GeV.	61
7.1	Triggers used for $\gamma$ +jets control region with associated prescales (averaged over runs) and	
7.0	$p_T$ thresholds	63
(.2	Efficiencies of $VBF$ cuts relative to the 2j selection in Z+jets and $\gamma$ +jets as calculated in ALPGEN using truth level events.	70
7.3	Efficiencies of $VBF$ cuts relative to the <b>2</b> j selection in Z+jets and $\gamma$ +jets as calculated in	
<del></del>	ALPGEN using truth level events after applying the reweighting procedure described in 7.3.2.	70
7.4	Efficiencies of <i>VBF</i> cuts relative to the 2j selection in Z+jets and $\gamma$ +jets as calculated using the following PDF sets: CTEQ6L1, CT10 and MSTW2008. Uncertainties on CTEQ6L1 are statistical while uncertainties on CT10 and MSTW2008 are derived from eigenvector variations.	71

7.5	Efficiencies of $VBF$ cuts relative to the <b>2j</b> selection in Z+jets and $\gamma$ +jets as calculated	
	with various scales, $Q^2 = 1.0, 2.0, 0.5$ . Uncertainties are statistical.	72
7.6	MC samples used in the $\gamma$ +jets control region	72
7.7	Multijet estimate as a function of the selection for the three fail PID regions	75
7.8	Yields of MC strongly produced $\gamma$ +jets, MC weakly produced $\gamma$ +jets, multijet estimate	
	as described in 7.4.2 and data as a function of the $\gamma$ +jets control region selection	75
7.9	MC samples used in the modified $W \to \ell \nu + jets$ control region	77
7.10	Yields of MC strongly produced $W \to \ell \nu + \text{jets}$ , MC weakly produced $W \to \ell \nu + \text{jets}$ and	
	data as a function of the modified $W \to \ell \nu + jets$ control region selection as described in 7.2.4.	78
7.11	Yields of MC strongly produced $\gamma$ +jets, MC weakly produced $\gamma$ +jets and data as a function	
- 10	of the $\gamma$ +jets anti-CJV control region selection	83
7.12	Yields of MC strongly produced $W \to \ell \nu + \text{jets}$ , MC weakly produced $W \to \ell \nu + \text{jets}$ and	~~~
H 10	data as a function of the modified $W \to \ell \nu$ +jets anti-CJV control region selection	83
7.13	Yields of data and estimated processes in the modified anti-cjv signal region. $W$ +jets is	
	estimated from the method described in 7.10 and estimates of strongly produced $Z \rightarrow$	
	$\nu\nu$ +jets are done using the $\gamma$ +jets anti-cjv control region and the modified $W \rightarrow \ell\nu$ +jets	
	anti-cjv control region. $Z \to \nu \nu$ MC is also presented for comparison. In all three $Z \to \nu \nu$	07
714	estimates, electroweak $Z \rightarrow \nu \nu$ is estimated from MC	81
1.14	Yields of MC strongly produced $\gamma$ +jets, MC weakly produced $\gamma$ +jets and data as a function of the x-lists modium measurement region selection	00
7 15	Violds of MC strongly produced $W \rightarrow \ell u$ into MC would produced $W \rightarrow \ell u$ into and	90
1.10	deta as a function of the modified $W \rightarrow \ell \nu + jets$ , MC weakly produced $W \rightarrow \ell \nu + jets$ and	01
7 16	Summary of strongly produced $Z \rightarrow \mu \mu$ estimates using the extrapolation of efficiencies	91
1.10	from $2i$ to VBF and extrapolation from anti-civ control regions to signal control regions	93
7.17	Summary of VBF W and VBF Z $\mu$ values (ratio of observed cross section to Standard	00
	Model prediction). Note that POWHEG is NLO and expects $\approx 10\%$ more events than SHERPA.	93
7.18	Cross sections of VBF $W \to e\nu$ and $Z \to \nu\nu$ processes at NLO QCD using VBFNLO. The	
	k-factor is defined with respect to SHERPA. All cross sections are shown in femtobarns	95
7.19	Cross sections of pure strong, pure weak and inclusive production for $W \to e\nu + 2$ jets and	
	$Z \rightarrow \nu\nu + 2$ jets as well as a calculation of the interference. The interference is calculated	
	relative to the weak cross section, as in Equation 7.17	95
7.20	Summary of $Z \rightarrow \nu \nu$ estimates using the extrapolation of efficiencies from 2j to $VBF$	
	and extrapolation from anti-cjv control regions to signal control regions. The first two	
	estimates include weakly produced $Z \to \nu \nu$ and weakly produced $W \to \ell \nu$ adjusted by	
	the interference and k-factor. The second two estimates include the weakly produced	
	backgrounds without adjustment. Also shown is the $W \to \ell \nu$ based estimate calculated in	
	the main analysis using a transfer factor. Statistical and theory certainties are included.	96
<b>Q</b> 1	Progressive yields of the reverse $\Lambda \phi$ , and control regions ( <b>B</b> selections) described in	
0.1	Find the reverse $\Delta \phi_{jx,E_T}^{\text{miss}}$ control regions ( <b>D</b> selections) described in Castion 9.1 in 20.2fb <sup>-1</sup> of 2012 data as ambusted using Monte Carlo. Expected signal	
	violda are shown for the area of $m_{\pi\pi} = 125$ CeV. The "Subtracted" column shows an	
	stimute of the dijot contribution by subtracting Z W and Other Background columns	
	from the observed yield	100
82	Progressive yields of the reverse $\Delta \phi_{\pm}$ control regions (C selections) described in Sec.	100
0.2	tion 81 in 20.3 fb <sup>-1</sup> of 2012 data as evaluated using Monte Carlo. Expected signal yields	
	are shown for the case of $m_{II} = 125$ GeV. The "Subtracted" column shows an estimate	
	of the dijet contribution by subtracting Z. W and Other Background columns from the	
	observed vield.	101

8.3	Efficiencies in the signal region and multijet enhanced regions, as estimated from the "Subtracted" columns of Table 5.2, Table 8.1 and Table 8.2. In <b>blue</b> are the efficiencies used to estimate the multijet contribution in the signal region. In green are the efficiencies used to estimate the multijet contribution in <b>A</b> -VR2 with the same efficiency used for $E_{\rm T}^{\rm miss}$ as the signal region. A "-" denotes insufficient statistics to evaluate a relative efficiency or upper bound	109
8.4	Estimates of the multijet background in the signal region and in A -VR2 (3-jet) using the technique described in Section 8.2.	102
9.1 9.2	Cross sections on $ggF$ and VBF signal processes for 115 GeV $< m_H < 300$ GeV. Uncer- tainties include variations of factorization and renormalization scales as well as PDF+ $\alpha_s$ . Yields and corresponding uncertainties of the $ggF$ signal process after reweighting the	106
93	Higgs $p_T$ according to HRes. The uncertainty is calculated using the gF_cross_section_uncertool with the RIST method	tainty 106
0.0	and renormalization scales by a factor of 2. Shown uncertainties are statistical.	107
9.4	The efficiency of the third jet veto by varying the parton shower modeling using Powheg+Pyth	ia
0 5	and Powheg+Herwig. The uncertainties shown are statistical	108
$9.5 \\ 9.6$	Summary of scale variations and weighting factors used to calculate uncertainties Summary of uncertainties due to scale variations weighted by the relative production of	109
	strong and weak production.	111
9.7	Summary of the effect of the scale variations on the $V$ +jets background estimates in the	
0.0	signal region using the effective background model described in Section 6.4.	111
9.8 9.9	Number of events passing variants of the selections with the nominal and varied parton shower model. Vields are shown for Z lists. W lists and the effective background model	112
	shower model. There are shown for $Z$ +jets, $W$ +jets and the ellective background model $(Z/W \text{ Patio})$	114
0.10	(Z/W) facto)	114
9.10	experimental systematic errors	115
9.11	The relative change (in percent) of the acceptance with respect to the nominal for VBF signal, $Z$ and $W$ processes in the signal and control regions for the four dominant experi-	110
9.12	effective background model, averaged between up and down variations	119
	$\ell\nu$ +jets in the multijet enhanced control region used to estimate the multijet contribution	
0.10	in the $W$ control region described in Section 6.2.2.	119
9.13	The fraction of signal MC events with at least one jet in tracking and the fraction with no jets in tracking, as a function of the signal region selection.	119
10.1	Summary of uncertainties due to scale variations weighted by the relative production of	
	strong and weak production.	126
$\begin{array}{c} 10.2 \\ 10.3 \end{array}$	Summary of uncertainties due to the PDF used to simulate events	126
	cesses, $\tilde{W} \to \ell \nu + jets$ in the control and signal regions, $Z \to \nu \nu + jets$ and $Z \to \ell \ell + jets$	127

10.4	Expected yields before and after the fit for the signal region in 20.3 $fb^{-1}$ of 2012 data.
	In the pre-fit yields, all processes are determined purely from MC except for the multijet.
	In the post-fit yields, the $ggF$ signal, VBF signal, and other backgrounds are determined
	from MC while the $Z \rightarrow \nu \nu + \text{jets}, W \rightarrow \ell \nu + \text{jets}$ , and multijet backgrounds are data-driven
	estimates. The expected signal yields are shown for $m_H = 125$ GeV and are normalized
	to $BR(H \rightarrow invisible) = 100\%$ . The post-fit W and Z statistical uncertainties include MC
	statistics from both the selected region and the corresponding control region, and the
	number of data events in the control regions. The "Other Backgrounds" include top and
	diboson production
11.1	Observable Higgs decay analyses used to determine limits on the coupling $\ldots \ldots \ldots 132$

# List of Figures

2.1	Potential that results from the BEH mechanism.	5
2.2	Full collection of known SM particles.	6
2.3	Feynman diagram representations for the 4 modes of Higgs production: gluon-gluon Fusion $(agF)$ Vector Boson Fusion (VBF) associated production with a vector boson $(VH)$ and	
	associated production with a production of a top quark pair $(ttH)$ .	7
2.4	Expected and observed rotation curves for the M 33 spiral galaxy. The discrepancy can be accounted for by assuming large amounts of dark matter permeating through the galaxy. The curve through the data represents the best-fit to a model which includes three contri-	
	butions to the velocity distribution: gas, stellar disk and dark matter. The dark matter	
2.5	contribution contains 3 free parameters which are fit with a least squares method The effect of weak gravitational lensing around the exceptionally massive LRG 3-757	10
	galaxy. The distortion of light due to the mass of the galaxy can be observed as an arc	
	around the foreground galaxy. The image was captured using the Hubble Space Telescope's	
	Wide Field Camera 3	11
2.6	The power spectrum of the CBM as a function of the angular scale observed by the WMAP,	
	Acbar, Boomerang, CBI and VSA instruments. The total matter density $(\Omega_m = \Omega_b + \Omega_{dm})$	
	is kept constant. The consistency of the various instruments support a large dark matter	10
2.7	Spin-independent observed limits on DM candidates from various direct detection experi-	12
2.1	ments (solid curves), hints of signal (shaded closed contours) and projections for expected limits from experiments that have not completed data taking (dot dashed curves). Also	
	shown is the neutrino coherent scattering limit, which presents a harrier to how sensitive	
	direct detection experiments can be (i.e., if the DM cross section is below the neutrino	
	coherent scattering limit, it will be undetectable).	13
2.8	Summary of detection methods of DM.	14
	v	
3.1	Summary of the accelerator complex including the series of accelerators used to accelerate	
	the protons to 450 GeV before they enter the LHC. Also shown are the interaction points	
	where the protons collide including ATLAS, CMS, LHCb and ALICE.	17
3.2	An example of PDFs from the MSTW group at 2 energy scales, $Q^2 = 10$ GeV and $Q^2 =$	4.6
	$10^4$ GeV.	19
3.3	A diagram of the ATLAS detector showing the various subdetectors and people for scale.	20

3.4 3.5	A diagram of the ATLAS inner detector showing the various subdetectors in the $r - \phi$ plane including the pixel detector, semiconductor tracker and transition radiation tracker. A diagram of the ATLAS inner detector showing the various subdetectors in the $r-z$	22
0.0	plane including the pixel detector, barrel and endcap semiconductor trackers and barrel and endcap transition radiation trackers	23
3.6	A diagram of the ATLAS calorimeter systems showing the various subdetectors in the $r-z$ plane including the LAr electromagnetic calorimeters, tile hadronic calorimeter, LAr	25
3.7	A diagram of the electromagnetic calorimeter system showing the various layers with dif-	25
3.8	A diagram of the MS showing the cylindrical layers of the muon chambers around the beam which include Monitored Drift Tubes, Cathode Strip Chambers, Thin Gap Calorimeters	26
3.9	and Resistive Plate Chambers	27 28
4.1	Schematic overview of the trigger system. Rates are from the original design specification for Run I which includes 25 ns bunch spacing increasing the collision rate to 40 Mhz and	
4.2	a lower EF output of 100 Hz	33
4.3	selected by the minimum bias trigger	36
	the rates is considered negligible	37
5.1 5.2	Feynman diagram for an invisibly decaying Higgs produced via VBF	$\begin{array}{c} 40\\ 40\end{array}$
5.5	$\Delta \eta_{j,j}$ after requiring two hard jets in opposite nemispheres. The disagreement of MC and data is due to an insufficient model of multijet production.	45
$5.4 \\ 5.5$	$E_{\rm T}^{\rm miss}$ with the full signal region selection applied except for the $E_{\rm T}^{\rm miss}$ requirement $E_{\rm T}^{\rm miss}$ and $m_{jj}$ distributions after the full selection is applied evaluated using Monte Carlo overlayed with observed data. The expected signal yields are shown for the case of $m_H =$	46
5.6	125 GeV	$\frac{46}{47}$
5.7	The Higgs $p_T$ distribution including electroweak radiations as calculated by HAWK	49
7.1	Example Feynman diagrams for $Z+2$ jets and $\gamma+$ jets. Since the couplings of $Z$ and $\gamma$ are similar, $\gamma+$ jets can be used to model the kinematics of $Z+$ jets.	63
7.2	Comparison of $\Delta \eta_{j,j}$ for $\gamma$ +jets, $W \rightarrow \ell \nu$ +jets and $Z \rightarrow \nu \nu$ +jets. All distributions come from fully simulated events generated with SHERPA. The selections from Section 7.2 are applied to the respective samples (modified signal region selection for Z) up to the opposite homisphares cut. All distributions are unit normalized	60
	opposite nemispheres cut. An distributions are unit normalized.	09

7.3	Isolation distributions for photons passing PID and for photons in the three regions of failing PID. The distributions are normalized to the isolation $E_T > 30$ GeV region of the	- 4
7.4	photons passing PID	74 76
7.5	$\Delta \eta_{j,j}$ with the modified $W \to \ell \nu + j$ ets selection applied through the opposite hemispheres cut	79
7.6 7.7	$\Delta \eta_{j,j}$ with the $\gamma$ +jets anti-CJV selection applied through the opposite hemispheres cut. $\Delta \eta_{j,j}$ with the modified $W \rightarrow \ell \nu$ +jets anti-CJV selection applied through the opposite hemispheres cut	84 85
7.8	$\Delta \eta_{j,j}$ Data/MC ratios for $\gamma$ +jets and $W \rightarrow \ell \nu$ +jets. Corresponding selections are applied through the opposite hemispheres cut.	86
8.1	$E_{\rm T}^{\rm miss}$ distribution with all of the selection in Section 5.3 applied except for the $\Delta \phi_{jet, E_{\rm T}^{\rm miss}}$	0.0
8.2	and $E_{\rm T}^{\rm mas}$ requirements	98
8.3	for the nominal signal region. $\dots$	99
8.4	of 30 GeV	103 104
9.1	Change of acceptance with respect to the nominal as a function of the set of experimental systematics considered. Shown are the systematics for $Z \rightarrow \nu\nu$ +jets with the signal region selection applied and $Z \rightarrow \ell\ell$ +jets with the Z control region selection applied. Green and vallow bands indicate the statistical error on the nominal yield in the signal region and Z	
9.2	control region, respectively	116
9.3	and $W$ control region, respectively	117
9.4	control region, respectively	118
10.1	Impact of nuisance parameters defined in Section ?? on signal strength $(\mu)$ after being	120
10.2	"pulled" by the fit	128
	The expected 95% confidence level upper bound on the branching fraction is 34% while the observed is 28%	130
11.1	$E_{\rm T}^{\rm miss}$ distributions for VBF and ZH analyses	132

11.2	Profile likelihood ratio as a function of BR <sub>i.,u.</sub> calculated using visible Higgs decays. The	
	red(green) horizontal lines indicate cut-off values on the profile likelihood ratio correspond-	
	ing to $68\%(95\%)$ confidence level	133
11.3	Spin-independent DM-nucleon cross section as a function of the DM mass. The exclusion	
	limits and signals observed by the direct detection experiments are compared to the branch-	
	ing fraction limit interpreted by the Higgs-portal and translated into the spin-independent	
	DM-nucleon cross section using the formulas in Section 2.5.5. The exclusion limits are	
	shown at 90% CL. The error bands on the ATLAS results indicate the uncertainty coming	
	from the different estimations of the Higgs-nucleon coupling.	134

## Author's Contribution

Analyses done in the ATLAS Collaboration usually have many people involved and it is often unclear how to determine how a particular author contributed to a given analysis. The analysis presented here provided a unique opportunity to work on several major aspects of the analysis due to the small size of the analysis team. Initially, I spent a significant amount of time developing a new control region  $(\gamma+\text{jets})$  used to estimate the dominant background  $(Z \rightarrow \nu\nu+\text{jets})$  in the analysis. The  $\gamma+\text{jets}$ control region is discussed in detail in Chapter 7. Development of the  $\gamma+\text{jets}$  control region using datadriven methods, investigating V+jets modeling consistency by comparing  $\gamma+\text{jets}$  and  $W \rightarrow \ell\nu+\text{jets}$ , calculation of systematics for the extrapolation of the  $\gamma+\text{jets}$  control region to the  $Z \rightarrow \nu\nu+\text{jets}$ prediction and calculation of effects due to higher orders of perturbation theory for strong and weak  $\gamma+\text{jets}$  production.

The  $\gamma$ +jets control region was ultimately not implemented due to issues with the underlying theoretical calculation of weak V+jets production. The first issue arises from corrections due to higher orders of perturbation theory, referred to as Next to Leading Order (NLO) effects. The second issue is from an "interference" effect between strong and weak V+jets production, which occurs when two processes have the same final state. These two issues are only significant when using the  $\gamma$ +jets control region to estimate  $Z \rightarrow \nu\nu$ +jets. By using different control regions to estimate  $Z \rightarrow \nu\nu$ +jets, the issues are largely avoided.

Even though the effects are not as significant when the  $\gamma$ +jets control region is not used, the NLO and interference effects still need to be considered when using weak V+jets simulations. This is done by calculating scale factors that account for the NLO and interference effects and applying them to the appropriate simulations. I performed the calculation of these scale factors.

Additionally, I calculated 2 systematic uncertainties for the dominant background simulations, described in detail in Chapter 9. The first systematic uncertainty arises from the choice of Parton Distribution Function (PDF). The second systematic uncertainty arises from unknown higher order cross section corrections. Uncertainties due to unknown higher order corrections were calculated to be quite significant. This provided the motivation to use control regions to form data-driven estimates of the dominant backgrounds, which are less sensitive to simulation uncertainties.

Finally, I contributed to the operation of ATLAS Detector by working with the trigger system, described in detail in Chapter 4. Specifically, I calculated trigger rates and created a tool which weights enhanced bias events such that they can be used to precisely predict trigger rates for future runs.

## CHAPTER 1

## Introduction

The discovery of the Higgs in July 2012 at the Large Hadron Collider opened the possibility of using the Higgs as a vehicle for other discoveries. The analysis presented here attempts to use the Higgs boson to explore the very active research area of dark matter. Dark matter accounts for most of the mass in the universe, but little is known about its properties. The existence of dark matter is inferred from many astrophysical observations originating from as early as 1931 with galaxy rotation curves. The most promising explanation for dark matter is that it is composed of Weakly Interactive Massive Particles (WIMPs), which interact only through gravity and the weak force. Since the Higgs boson couples to massive particles, there is motivation to believe that it should couple to dark matter candidates as well.

The analysis presented here attempts to measure the branching fraction of the Higgs decaying to dark matter. Since dark matter candidates do not interact with normal matter (except gravitationally) their existence at colliders is inferred from a lack of momentum conservation, known as missing momentum. The lack of momentum conservation is a result of the dark matter particles passing through the detector without interacting with any of the material. In general, particles that pass through the detector without depositing energy are considered "invisible". If the Higgs boson decayed to invisible particles a significant amount of the time, an excess of collision events with large missing momentum over an expected number of events from the Standard Model would be seen. Using the observed number of events, the expected contamination from backgrounds that mimic signal, the efficiency of a selection optimized to find signal events and the Standard Model Higgs cross section (corresponding to the discovered mass of 125 GeV), a measurement of the branching fraction can be made. In the absence of an excess of events with large missing momentum, a 95% confidence level limit is calculated which represents how often the Higgs boson could decay to invisible particles and the analysis would be able to detect it.

#### 1. INTRODUCTION

The production mode of the Higgs used in this analysis, known as Vector Boson Fusion (VBF), is a particularly strong way of searching for an invisibly decaying Higgs. The VBF production mode is powerful since the expected signal to background ratio is high and the dominant background normalizations can be constrained through the use of control regions. The use of control regions suppresses systematic uncertainties from experimental and theoretical sources that would otherwise significantly degrade the sensitivity.

Even though no excess over the expectation from the Standard Model was found, the analysis resulted in a very sensitive branching fraction limit. With more statistics from 13 TeV collisions and improvements in the methods discussed here, the result will be even stronger.

### CHAPTER 2

## **Theoretical Motivation**

#### 2.1 Introduction

The Standard Model of particle physics (SM) is a theory which describes the most fundamental constituents of matter as well as electromagnetic, weak and strong interactions. SM predictions have been extensively tested experimentally with great success. Examples include the existence of the W and Z bosons, gluons, the top and charm quarks and most recently, the Higgs boson. Additionally, precision tests of the SM have proven to be successful with the predictions of the masses of the weak force carriers, the W and Z. However, there are several deficiencies with the SM including a lack of a dark matter candidate and a lack of a quantum field theory which accommodates general relativity. Deficiencies of the SM are explained in detail in Section 2.4.

#### 2.2 Standard Model

The SM is a quantum field theory which describes interactions of electromagnetism, the weak force and the strong force through the SM Lagrangian. The SM Lagrangian contains the associated symmetry group  $SU(3) \times SU(2) \times U(1)$ . The sector representing strong interactions (referred to as quantum chromodynamics or QCD) defines interactions between quarks and gluons and has SU(3) symmetry. The quarks possess a color charge (red, blue, green) which is exchanged between 8 gluons. The sector representing interactions of electromagnetism and the weak force (electroweak) contains  $SU(2) \times U(1)$ symmetry. The SU(2) symmetry group represents weak isospin and U(1) represents weak hypercharge. Weak isospin is a quantum number, which relates to the weak interaction, while weak hypercharge is a quantum number which is a combination of the electric charge and the third component of weak isospin and thus relates the weak force with the electromagnetism. The number of independent generators of a special unitary group SU(n) is  $n^2 - 1$ . Therefore SU(3) and SU(2) have 8 and 3 generators, respectively. The SU(3) generators describe 8 gluon fields, the SU(2) generators describe 3 weak gauge bosons  $(W^1, W^2, W^3)$  and U(1) has 1 generator, which describes 1 weak boson (B).

The  $SU(2) \times U(1)$  theory described results in massless gauge bosons. However, the W and Z bosons are known to be massive through experiment. Massive gauge bosons are created by spontaneously breaking the symmetry of SU(2) and U(1) with the Brout-Englert-Higgs (BEH) mechanism. The mechanism introduces a new scalar field (the Higgs boson), which spontaneously breaks the symmetry group and lets the gauge bosons acquire non-zero mass. Additionally, the Higgs adds a non-zero potential value to the SM Lagrangian which is not stable around the origin, shown in Figure 2.1. This means there is a non-zero vacuum expectation. Applying the BEH mechanism to the Glashow-Weinberg-Salam model which describes electroweak interactions, the generators from SU(2) and U(1)linearly combine to form the gauge bosons observed in nature:

$$W^{+} = \frac{1}{\sqrt{2}} (W^{1} - iW^{2}) \tag{2.1}$$

$$W^{-} = \frac{1}{\sqrt{2}} (W^{1} + iW^{2}) \tag{2.2}$$

$$Z = W^3 \cos \theta_W - B \sin \theta_W \tag{2.3}$$

$$\gamma = W^3 \sin \theta_W + B \cos \theta_W \tag{2.4}$$

where 
$$\cos \theta_W = \frac{m_W}{m_Z}$$
 (2.5)

The symmetry breaking provides mass terms for the gauge bosons. The Higgs also has the correct conserved quantities (known as quantum numbers) to provide masses for fermions (quarks and leptons) by way of Yukawa couplings. The mixing for the neutral gauge bosons is determined by the Weinberg angle ( $\theta_W$ ) and is measured experimentally. The first measurement of  $\theta_W$  came from the Gargamelle experiment at CERN, which searched for neutral currents [1]. The first observation of the massive gauge bosons came from the UA1 and UA2 experiments at CERN in 1983, which were compatible with the SM expectation [2, 3, 4, 5].

The full collection of known SM particles is shown in Figure 2.2. The gauge bosons mentioned above are mediators of the forces: gluons mediate the strong force,  $W^{\pm}$  and Z mediate the weak force and  $\gamma$  mediates electromagnetism. Observable matter is composed of fermions, which are grouped into quarks and leptons. Quarks form composite particles known as hadrons, the most stable of which are protons and neutrons. Due to their color charge, quarks can only be found in hadrons, a phenomenon known as color confinement. Leptons do not have color charge and thus only interact with the electroweak gauge bosons (W, Z and  $\gamma$ ). Neutrinos are the only leptons that do not have electric charge, which means they only interact with the weak gauge bosons (W and Z). Additionally, their



Figure 2.1: Potential that results from the BEH mechanism.

mass is quite small when compared to other leptons. There are 3 generations of quarks and leptons, which have been discovered and differ only in mass. The theoretical motivation for 3 generations is unknown.

#### 2.3 Higgs Boson

The observation of massive W and Z bosons combined with the fact that fermions have mass provide strong evidence for the existence of a Higgs boson. However, since the mass of the Higgs is a free parameter in the SM, it is difficult to perform dedicated searches. Additionally, each production mode involves different signal to background ratios. The Higgs production modes at the LHC ordered by cross section are: gluon-gluon fusion (ggF), Vector Boson Fusion (VBF), associated production with a vector boson (VH) and associated production with a top quark pair (ttH). The order of cross sections is largely independent of the mass of the Higgs except at low  $m_H$  ( $<\approx$  90 GeV) where the associated production cross section can be higher than VBF. The Feynman diagram representations are shown in Figure 2.3.

In order to search for a Higgs boson, particles are accelerated to high energies and then interact through collisions. The Large Electron-Positron Collider (LEP) collided electrons and positrons at a center of mass energy around the Z mass ( $\approx 90$  GeV) starting in 1989. LEP was able to search for associated production (VH), where an electron and positron collide to form a virtual Z, which can then emit a Higgs [6]. The specific decay modes considered are:  $(H \to b\bar{b})(Z \to q\bar{q}), (H \to b\bar{b})(Z \to \nu\bar{\nu}),$  $(H \to b\bar{b})(Z \to \ell\ell)$  (with  $\ell$  being an electron or muon),  $(H \to b\bar{b})(Z \to \tau\tau)$  and  $(H \to \tau\tau)(Z \to q\bar{q})$ .



Figure 2.2: Full collection of known SM particles.

At LEP, associated production is expected to have the highest cross section [6]. Additionally, the Higgs decaying to pairs of b quarks is the dominant decay mode of the Higgs with  $m_H \ll 130$  GeV). Colliding electrons and positrons provides very clean signatures since there can be no other hadronic activity in events. There was no discovery with at LEP with the initial center of mass energy, so the center of mass energy was gradually increased to 209 GeV by the end of 2000 at which point it was shut down to make room for the Large Hadron Collider (LHC). Any additional increase in the center of mass energy would have required more accelerating cavities, which was not possible due to space constraints. The result from LEP was a lower bound on the Higgs mass of 114 GeV.

The search for the Higgs then moved to the Tevatron Run 2, which collided protons and antiprotons at a center of mass energy of 1.96 TeV starting in 2001. Because protons and antiprotons are hadrons composed of quarks and gluons, the Tevatron did not produce events that were as clean as LEP. As a result, the background rates for some searches is too high, such as  $ggF H \rightarrow b\bar{b}$ , which has the highest production cross section and branching fraction. As a result, the sensitivity of the Tevatron



Figure 2.3: Feynman diagram representations for the 4 modes of Higgs production: gluon-gluon Fusion (ggF), Vector Boson Fusion (VBF), associated production with a vector boson (VH) and associated production with a production of a top quark pair (ttH).

was limited to 140 GeV  $< m_H < 180$  GeV [7]. The result from the Tevatron was an exclusion of 147 GeV  $< m_H < 179$  GeV, setting the stage for the LHC [8].

In 2012, the ATLAS and CMS experiments announced the discovery of a new particle with properties consistent with a SM Higgs boson [9, 10]. Data was taken from colliding protons and protons at a center of mass energy of 7 TeV and 8 TeV. The ggF Higgs production mode dominated the discovery with the Higgs decaying to  $\gamma\gamma$ ,  $WW^*$  and  $ZZ^*$ . The measured mass of 125 GeV is consistent with global electroweak fits reported from LEP [6]. Additionally, the measured mass of 125 GeV is consistent with the  $3.0\sigma$  significance at 125 GeV reported from the Tevatron combination [11].

#### 2.4 Issues with the Standard Model

Even though the SM has been remarkably successful in classifying particles and their interactions, there are several deficiencies described below.

- 1. Gravity
  - The SM cannot accommodate general relativity as a quantum field theory.
- 2. Hierarchy problem
  - Corrections to the Higgs mass due to loops are much larger than the scale of electroweak interactions. To remedy this, either the parameters of the theory must be finely tuned to induce large cancellations or there must exist new particles that enter the loops and induce the correct cancellations.
- 3. Dark energy
  - The accelerating expansion of the universe is driven by dark energy which permeates all space. Dark energy constitutes 68% of the observable universe and is not accommodated in the SM.

#### 4. Dark matter

• Several astrophysical measurements imply the existence of a stable weakly interactive massive particle. There is currently no candidate for this particle in the SM. A review of the astrophysical evidence for this "dark" matter is presented in Section 2.5.

#### 2.5 Dark Matter

#### 2.5.1 Observational Evidence

There is substantial evidence for the existence of dark matter (DM). Since DM does not interact electromagnetically, evidence of the existence of dark matter is based on astrophysical observations. The plot of the magnitude of the orbital velocities of visible stars and gas of a disk galaxy versus radial distance to the center can be used to study mass distributions. If the luminosity and total mass of a disk galaxy are proportional, the rotation curve should fall as  $1/\sqrt{R}$  when outside the luminous matter. However, observations show that the orbital velocities of the galaxy. This discrepancy can be solved by assuming a large amount of DM. Figure 2.4 shows the difference between the expected and observed rotation curves for the M 33 spiral galaxy [12]. The curve through the data represents the best-fit to a model which includes three contributions to the velocity distribution: gas, stellar disk and dark matter. The dark matter contribution to the total velocity contains 3 free parameters that are fit with a least squares method.

Evidence for the existence of DM also comes from the observation of gravitational lensing effects. Since massive objects like galaxies distort space-time, light from more distance sources gets bent as it approaches. This effect is known as weak gravitational lensing. Weak gravitational lensing distorts the light around background objects (like galaxies) near a foreground mass. Strong gravitational lensing occurs when the lensing is strong enough to produce multiple images, arcs, or Einstein rings. Figure 2.5 shows an Einstein ring arising from strong gravitational lensing on the LRG 3-757 foreground galaxy [13]. The distortion of light of the background objects are used to measure the mass of the foreground galaxy which significantly disagree with measurements of luminous matter, providing evidence for additional mass due to DM.

Measurements of the Cosmic Microwave Background also provide evidence for DM. The CMB is an almost isotropic thermal radiation present from the Big Bang, which can be measured using ground, balloon and space-based receivers. As the universe cooled after the Big Bang, neutral atoms formed from protons and electrons (Hydrogen), which could no longer absorb thermal radiation. Soon after, photons started traveling freely through space instead of being scattered off of electrons and protons. Anisotropies of the CMB can be caused by density perturbations due to acoustic oscillations between photons, baryons and dark matter in the early universe. Acoustic oscillations occur due to competing effects from the collapse of dense regions due to gravity and the outward pressure of photons due to the strong interactions with electrons and protons. The acoustic oscillations cause waves in the CMB



Figure 2.4: Expected and observed rotation curves for the M 33 spiral galaxy. The discrepancy can be accounted for by assuming large amounts of dark matter permeating through the galaxy. The curve through the data represents the best-fit to a model which includes three contributions to the velocity distribution: gas, stellar disk and dark matter. The dark matter contribution contains 3 free parameters which are fit with a least squares method [12].

power spectrum, which correspond to hot and cold areas in the universe. The power spectrum of the CMB as a function of the angular scale is shown in Figure 2.6 [14]. Each of the peaks contains unique information: the first determines the curvature of the universe, the second determines the reduced (from the gravitational collapse of dense regions) baryon density and the third can be used to determine the dark matter density. The measured baryon and dark matter densities are 4.6% and 24%, respectively.

#### 2.5.2 WIMP Hypothesis

The most widely accepted hypothesis for DM candidates are Weakly Interacting Massive Particles (WIMPs) which interact only through gravity and the weak force. As the universe cooled and expanded after the big bang, thermal energy decreased and particles with masses greater than  $\approx kT$  are generally not present due to lack of energy to create them. Even though WIMPs have a mass which



Figure 2.5: The effect of weak gravitational lensing around the exceptionally massive LRG 3-757 galaxy [13]. The distortion of light due to the mass of the galaxy can be observed as an arc around the foreground galaxy. The image was captured using the Hubble Space Telescope's Wide Field Camera 3.

is much higher than kT, they are present since the probability of annihilating with another WIMP is too small due to number density. The point at which the number density is insufficient to sustain DM interactions is known as the "freeze-out". The observed abundance of DM via thermal production in the early universe ( $t < 10^{-9}$  s) is obtained by CMB power spectrum measurements by experiments such as WMAP [15]. To obtain the observed abundance of DM, the calculated self-annihilation interaction cross section must be on the order of the weak force interaction scale. The coincidence of the required DM density and the weak force interaction scale is known as the "WIMP Miracle".

#### 2.5.3 Direct Detection Experiments

If WIMPs compose dark matter, many of them would constantly bombard the Earth since the local dark matter density is calculated to be  $\approx 0.3 \text{ GeV cm}^{-3}$  [16]. As a result, a variety of experiments have been designed to search for WIMPs scattering off of nuclei. In order to reduce backgrounds from



Figure 2.6: The power spectrum of the CBM as a function of the angular scale observed by the WMAP, Acbar, Boomerang, CBI and VSA instruments [14]. The total matter density  $(\Omega_m = \Omega_b + \Omega_{dm})$  is kept constant. The consistency of the various instruments support a large dark matter density.

cosmic showers, direct detection experiments are usually set in deep underground laboratories. There are 2 major methods of direct detection: observing an excess scattering off of nuclei over background (such as nuclear beta decay or neutron scattering) and observing an annual modulation in event rate. Annual modulation in event rate occurs from the seasonal variation of the velocity of the detector relative to the DM.

The mass of the DM candidate and its interaction cross section are anticorrelated due to the observed dark matter density in the universe. For example, an increase in mass must correspond to a decrease in cross section in order to keep the observed dark matter density fixed. Because of this, various direct detection experiments have different sensitivities in different mass regions due to differences in detector technique. Spin-independent results from various experiments are shown in



Figure 2.7: Spin-independent observed limits on DM candidates from various direct detection experiments (solid curves), hints of signal (shaded closed contours) and projections for expected limits from experiments that have not completed data taking (dot-dashed curves) [17]. Also shown is the neutrino coherent scattering limit, which presents a barrier to how sensitive direct detection experiments can be (i.e., if the DM cross section is below the neutrino coherent scattering limit, it will be undetectable).

Figure 2.7 [17]. Direct detection experiments are limited by the neutrino coherent scattering limit, if the DM interaction cross section was below the neutrino coherent scattering limit, a signal would be essentially undetectable.

#### 2.5.4 Dark Matter Detection with the Large Hadron Collider

The Large Hadron Collider (LHC) may also be used to detect WIMPs produced in proton proton collisions. Since WIMPs have very small interactions with observable matter, they can be detected indirectly by observing the lack of momentum balance of the detected particles, known as missing transverse energy ( $E_{\rm T}^{\rm miss}$ ). One mechanism for WIMP production at the LHC is to assume that



Figure 2.8: Summary of detection methods of DM.

WIMPs interact with a heavy particle, which mediates the interaction. In this thesis the heavy particle is hypothesized to be the Higgs boson. There are also models in which the mediator is light and interacts very weakly, like a dark photon.

In summary, there are several DM processes considered to directly detect dark matter including production mechanisms in which SM particles annihilate to DM particles and scattering processes where DM scatters off of a nuclei. Additionally, experiments can directly search for DM annihilation into SM particles. One such experiment is PAMELA, which found an excess of positrons as a possible sign of dark matter annihilation [18]. These processes are all related theoretically and represent rotations of a Feynman diagram, shown in Figure 2.8.

#### 2.5.5 Dark Matter Portal

The hypothesis that WIMPs only interact with the Higgs boson is known as the Higgs portal model [19, 20, 21]. An important aspect of the Higgs portal model is the hypothesis that the Higgs decays to a pair of DM particles. As a result, the DM candidate must have a mass less than  $m_H/2$ . Unobserved quantum fields and their corresponding particles that do not interact directly with the gauge bosons in the SM are collected into the "hidden sector". Particles in the hidden sector still interact through gravity. The Higgs portal model provides a simple extension to the SM by providing a direct coupling to the hidden sector. Since the Higgs portal is model independent, scenarios in which a DM candidate

is a scalar (S), vector (V) or Majorana fermion  $(\chi)$  can all be investigated. This can be done since the Higgs field operator in the SM has no quantum numbers. As a result, the Higgs can couple to a pair of scalars, vectors or fermions without violating conservation laws. For scalars and vectors, the interaction term is dimension 4 in the Lagrangian and is renormalizable. However, for fermions, the Higgs interaction term has dimension 5 and a cut-off scale  $\Lambda$  must be introduced in order to remain renormalizable. The interaction strengths are defined to be  $\lambda_{hSS}$ ,  $\lambda_{hVV}$  and  $\lambda_{hff}$  for scalars, vectors and Majorana fermions, respectively. Using the Feynman rules, the partial width of the Higgs decaying invisibly are given for the 3 types of DM candidates [19]:

$$\Gamma_{h\to SS}^{\text{invisible}} = \frac{\lambda_{hSS}^2 v^2}{64\pi m_h} \sqrt{1 - 4\frac{M_S^2}{m_h^2}}$$
(2.6)

$$\Gamma_{h \to VV}^{\text{invisible}} = \frac{\lambda_{hVV}^2 v^2 m_h^3}{256\pi M_V^4} \left( 1 - 4\frac{M_V^2}{m_h^2} + 12\frac{M_V^4}{m_h^4} \right) \sqrt{1 - 4\frac{M_V^2}{m_h^2}}$$
(2.7)

$$\Gamma_{h \to \chi\chi}^{\text{invisible}} = \frac{\lambda_{hff}^2 v^2 m_h}{32\pi\Lambda^2} \left(1 - 4\frac{M_\chi^2}{m_h^2}\right)^{3/2}$$
(2.8)

The partial width is dependent on: mass of the Higgs boson  $(m_h)$ , the DM mass candidate  $(M_S, M_V \text{ and } M_{\chi}, \text{ depending on type})$ , the vacuum expectation value (v) and the interaction strength  $(\lambda_{hSS}, \lambda_{hVV} \text{ and } \lambda_{hff}, \text{ depending on type})$ . In the case of the Majorana fermion DM candidate, the partial width is also dependent on the cut-off scale,  $\Lambda$ . The scattering cross sections for the dark matter-nucleon interaction are calculated in order to compare LHC results to direct detection experiments. They are given for the 3 types of DM candidates:

$$\sigma_{S-N} = \frac{\lambda_{hSS}^2}{16\pi m_h^4} \frac{m_N^4 f_N^2}{(M_S + m_N)^2}$$
(2.9)

$$\sigma_{V-N} = \frac{\lambda_{hVV}^2}{16\pi m_h^4} \frac{m_N^4 f_N^2}{(M_V + m_N)^2}$$
(2.10)

$$\sigma_{f-N} = \frac{\lambda_{hff}^2}{4\pi\Lambda^2 m_h^4} \frac{m_N^4 M_f^2 f_N^2}{(M_f + m_N)^2}$$
(2.11)

The cross sections are additionally dependent on the nucleon mass  $(m_N)$  and the Higgs-nucleon coupling  $(f_N)$ .

### CHAPTER 3

# The Large Hadron Collider and the ATLAS Experiment

#### 3.1 Introduction

In order to study massive particles at the electroweak scale like the Higgs boson, long lived particles are accelerated to very high energies and collide in particle detectors. The accelerator used for this analysis is the Large Hadron Collider (LHC), frequently referred to as the largest and most complex scientific experiment ever built. Data used in this analysis is from proton-proton collisions at a center of mass energy of  $\sqrt{s} = 8$  TeV. Starting in 2015, after upgrades and repairs, the LHC will collide protons and protons at a center of mass energy of  $\sqrt{s} = 13$  TeV. The LHC has 4 interaction points where protons collide. At each of the interaction points there is an experiment and corresponding collaboration, which record and analyze the collisions. Two of them are multipurpose detectors, ATLAS and CMS, while the other two experiments have more specific purposes, LHCb and ALICE. Multiple collaborations are formed in order to verify and build confidence in discoveries. The LHC resides on the Franco-Swiss border at the European Organization for Nuclear Research (CERN) near Geneva, Switzerland. The accelerator complex is 100 m underground and has a circumference of 26.7 km.

#### 3.2 Accelerator Complex

Once protons are obtained from Hydrogen atoms, they are accelerated to 50 MeV with a linear accelerator known as the Linac2. Afterwards, the protons are passed through a series of accelerators to gain an energy of 450 GeV. The sequence of injectors roughly traces the history of colliders at CERN. The Proton Synchotron (PS) was originally built in the 1950s and used as a particle source


# **CERN's Accelerator Complex**

Figure 3.1: Summary of the accelerator complex including the series of accelerators used to accelerate the protons to 450 GeV before they enter the LHC. Also shown are the interaction points where the protons collide including ATLAS, CMS, LHCb and ALICE.

for the Gargamelle bubble chamber (which discovered weak neutral currents in 1974). The PS has gone through many modifications to handle the intensity of the current proton beam. The Proton Synchotron Booster (PSB), built in 1972, has a radius of 25 m and was designed to accelerate protons to an energy sufficient to boost the proton acceptance of the PS by a factor of 100 over direct injection from the Linac2. Before the PSB, protons were injected directly to the PS from the Linac2 which couldn't accelerate protons to sufficient energies with the acceptance required. Lastly, the Super Proton Synchotron (SPS) was built to collide protons and antiprotons and was used from 1981 to 1984, leading to the discovery of the W and Z bosons. Protons from Linac2 accelerate to 1.4 GeV in the PSB, then to 26 GeV in the PS and finally to 450 GeV in the SPS before entering the LHC. The LHC accelerates the proton bunches with superconducting magnets to the final collision energy of 4 TeV used in this analysis. A summary of the accelerator complex is shown in Figure 3.1.

# 3.3 Beam Parameters

Two beam parameters that are especially important at a particle accelerator are instantaneous luminosity and the center of mass energy. The instanteous luminosity is defined as:

$$\mathcal{L} = \frac{N_1 N_2 n_b f_{rev}}{4\pi \sigma_x \sigma_y} F \tag{3.1}$$

The instantaneous luminosity is dependent on the number of protons in each beam  $(N_1, N_2)$ , the number of bunches  $(n_b)$ , the revolution frequency  $(f_{rev})$ , the width of the beam in the transverse directions  $(\sigma_x, \sigma_y)$  and the beam emittance (F) which is the measure of the average spread of particle coordinates in position-momentum phase space. The instantaneous luminosity is proportional to the collision rate (the constant of proportionality being the cross section) and can be integrated over time to measure the size of a dataset. The integrated luminosity of the dataset used in the analysis presented here is 20.3 fb<sup>-1</sup>. The number of events due to a given process is given by the cross section multiplied by the integrated luminosity, since rare processes have low cross sections, the LHC tries to maximize the instantaneous luminosity to have sufficient statistics to do a significant analysis.

The center of mass energy is also an important parameter due to fact that interesting collisions are from the proton constituents colliding, which follow Parton Distribution Functions (PDFs). A PDF is defined as the probability density for finding a proton constituent (quark or gluon) with a longitudinal fraction of the proton momentum x at energy scale  $Q^2$ . Since QCD does not predict the parton content of the proton, PDFs are determined by fitting experimental observables. An example of a PDF calculated by the MSTW group is shown in Figure 3.2 [22]. Considering PDFs, the center of mass energy is in general much higher than the mass of the particles produced. For example, the center of the mass of energy of LHC used for the analysis presented here was 8 TeV even though the Higgs boson has a mass of 125 GeV, an entire order of magnitude lower.

# 3.4 ATLAS Detector

#### 3.4.1 Overview

The ATLAS detector is one of the 2 multipurpose detectors at the LHC which forms an international collaboration studying many physics processes. The detector is made up of a set of concentric subdetector systems that work together to identify various types of particles. Starting at the beam line, the following subdetectors comprise ATLAS:



Figure 3.2: An example of PDFs from the MSTW group at 2 energy scales,  $Q^2 = 10$  GeV and  $Q^2 = 10^4$  GeV [22].

- Inner Detector (ID): The ID is composed of 3 systems enclosed in a 2 T solenoidal magnetic field: the Pixel Detector, the Semiconductor Tracker (SCT) and the Transition Radiation Tracker (TRT). In general, trackers measure the direction and magnitude of charged particles. This is done by measuring the energy particles release as they interact with material in the tracker. Additionally, trackers are usually found in a magnetic field in order to make the traversing particles form a helix which can be reconstructed using track parameters to measure the direction and magnitude of the particle considered.
- 2. Calorimeters: The calorimeters consist of dense material that particles interact with and cause them to produce a cascade of secondary particles, known as a shower. The particle shower can then be used to measure the energy of the original pre-showered particle. There are 2 types of showers: electromagnetic and hadronic. Electromagnetic showers occur when a particle (usually a photon or electron) interacts via the electromagnetic force. Hadronic showers occur when a hadron (proton, neutron, pion, etc.) interacts via the strong force. There is an electromagnetic



Figure 3.3: A diagram of the ATLAS detector showing the various subdetectors and people for scale.

calorimeter to measure the interactions of photons and electrons and a hadronic calorimeter to measure the interactions of hadrons. Separate calorimeters are required since hadrons must interact with more material before showering. The calorimeter system is composed of the liquid argon electromagnetic calorimeter (LAr), the hadronic tile calorimeter, the liquid argon hadronic endcap calorimeter and the forward calorimeters.

3. Muon Spectrometer (MS): The MS is the outermost detector system and measures the momentum of particles not showered in the calorimeter by recording hits and forming trakes from minimum ionizing muons. Neutrinos also survive all the way to the MS, but since they don't interact their momentum cannot be measured.

Figure 3.3 shows the ATLAS detector as well as the subdetector systems.

A three-tiered trigger system is used read out and make decisions to save collision data, described in Section 4.1.

#### 3.4.2 Coordinate System

Generally, particle detectors used in multipurpose experiments use a right-handed coordinate system centered at the primary vertex, centered at the nominal proton-proton collision point. The z-axis is defined to run along the beam line with the x-y plane being perpendicular, referred to as the transverse plane.

Due to the cylindrical nature of ATLAS, cylindrical and polar coordinates are used to describe particle trajectories as well as detector positions. In the transverse plane, the azimuthal angle  $\phi$  is measured around the beam (z-axis) and the radial dimension r is the distance from the beamline. The polar angle is measured as the angle from the z-axis which is often expressed as pseudorapidity, defined as  $\eta = -\ln \tan(\theta/2)$ . The pseudorapidity is often used in particle physics since it is invariant under boosts along the z-axis. This is needed since the hard scatter is between proton constituents with different momenta and the center of mass frame is shifted. Since the pseudorapidity is invariant under z-axis boosts, it does not matter that the center of mass frame is shifted. In the  $\eta - \phi$  space, distance between objects is defined as  $\Delta R = \sqrt{\Delta \eta^2 + \Delta \phi^2}$ .

# 3.4.3 Inner Detector

Diagrams of the ID are shown in Figure 3.4 and Figure 3.5 in the  $r - \phi$  and r - z planes, respectively.

The ID performs measurements of the positions of charges particles as they traverse the 3 subdetectors: the Pixel Detector, the Semiconductor Tracker (SCT) and the Transition Radiation Tracker (TRT). The Pixel Detector and SCT use silicon to have high detector granularity in order to make high precision measurements. All of the ID is immersed in a 2 T magnetic field which causes charged particles to curve. The curvature can then be used to infer particle momenta.

The pixel detector is closest to the beamline and has the highest granularity to be able to reconstruct secondary vertices from *b*-meson decays, which travel on average  $\approx 0.5$  mm before decaying. In order to have high granularity, the pixel detector uses silicon pixels which charged particles traverse and create electron-hole pairs. The electron-hole pairs drift into an electric field and register a current pulse. Since the current pulses are recorded locally by sensors, the position of the particle can be identified. The pixel detector is composed of 80 million read out channels and averages 3 measurements per charged particle that traverses the detector. The resolution is 10  $\mu$ m in the  $r - \phi$  plane and 115  $\mu$ m along the z-axis. The coverage area of the pixel detector is  $|\eta| < 2.5$  including the concentric layers of pixel in the barrel ( $|\eta| < 1.9$ ) and endcap disks on both sides ( $1.9 < |\eta| < 2.5$ ). The layer of pixel in the barrel closest to the beampipe is known as the b-layer.



Figure 3.4: A diagram of the ATLAS inner detector showing the various subdetectors in the  $r - \phi$  plane including the pixel detector, semiconductor tracker and transition radiation tracker.



Figure 3.5: A diagram of the ATLAS inner detector showing the various subdetectors in the r-z plane including the pixel detector, barrel and endcap semiconductor trackers and barrel and endcap transition radiation trackers.

The SCT is composed of a double layer of silicon chips which measures the energy of charged particles in a way that is similar to the pixel detector. The back-to-back layers are tilted with respect to each other by 40 mrad in order to provide additional resolution along the long axis of the layer for hits that are coincident. This tilted layers allow to measure coordinates in the  $r - \phi$  as well as z. The SCT has 4 million read out channels and is comprised of 4 layers in the barrel and 9 layers in the endcap to main coverage of  $|\eta| < 2.5$ . The resolution is 17  $\mu$ m in the  $r - \phi$  plane and 580  $\mu$ m along the z-axis.

The TRT uses a different technology with straw drift tubes containing 70% Xenon, 27% CO<sub>2</sub> and 3% O<sub>2</sub> gases. When a charged particle traverses a straw drift tube, it ionizes the gas and the free electrons drift to the wire at the center of the straw via an applied electric field and induces a signal. The TRT contains  $\approx$ 300,000 straw drift tubes which are arranged in the barrel cylindrically along the z-axis up to  $|\eta| < 1$  and radially outward in the *r* direction in the endcaps. The resolution is 130  $\mu$ m in  $\phi$  (the geometry prevents resolution in z and *r*) with coverage of  $|\eta| < 2.0$ . A particle track has an

average of 35 hits whereas the number of expected hits in the pixel detector and SCT is only 7.

Additionally, the TRT provides particle identification through transition radiation. Transition radiation photons are emitted when a charged particle traverses a boundary between materials with different dielectric constants. To cause this, the volume between straws is filled with a radiator material. Transition radiation photons are emitted in the particle trajectory direction at keV energies and cause a large signal amplitude in the straw. The probability to emit a transition radiation photon depends on the Lorentz Factor ( $\gamma$ ) of the charged particle that is traversing through the detector. Since the mass of an electron is low compared to other charged particles, they are more likely to cause the emission of transition radiation photons.

In order to traverse the entire ID, a particle must have a minimum momentum of 500 MeV. The measurements from the 3 subdetectors are then combined to form a track momentum measurement with a  $p_T$  resolution of  $p_T \times 0.05\% \oplus 1\%$ .

## 3.4.4 Calorimeters

A diagram of the calorimeter systems is shown in Figure 3.6.

The calorimeter systems measure the energy of electrons, photons and hadrons up to  $|\eta| < 4.9$ . Particles that traverse the calorimeters form particle showers by interacting with the detector material through electromagnetic and nuclear reactions. They are designed to have the depth required to contain the showers before they extend to the MS. The calorimeters have active material that can only measure a fraction of the energy produced by a particle shower and full energy is inferred, hence the calorimeters are considered to be "sampling".

The electromagnetic calorimeter, located directly outside of the solenoid magnet, is used to measure the energy of electrons and photons. The barrel and endcap calorimeters are separate but both have an accordion design. Lead is used as the dense absorber material with liquid argon (LAr) as the active material. The accordion design ensures complete and uniform coverage in  $\phi$ . High granularity measurements are provided by the barrel and endcap LAr through  $|\eta| < 2.47$ . Measurements are performed by 4 radial sections of the electromagnetic calorimeter that differ in granularity, shown in Figure 3.7.

The first section is the pre-sampler which is a thin layer of active liquid argon and is designed to detect early particle showers. The second section is the "strips", thin liquid argon cells, which provide very fine segmentation in  $\eta$ . The segmentation provides the ability to distinguish between showers initiated by electrons or photons and showers initiated by neutral pion decays ( $\pi^0 \rightarrow \gamma \gamma$ ). The third section contains the bulk of the radiation lengths and therefore results in the primary



Figure 3.6: A diagram of the ATLAS calorimeter systems showing the various subdetectors in the r-z plane including the LAr electromagnetic calorimeters, tile hadronic calorimeter, LAr hadronic endcap and LAr forward calorimeter.

energy measurement. The final section is designed to estimate energy leaking from the third section of the calorimeter. The third section has coarser granularity and is thinner than the other sections. The coverage of the electromagnetic calorimeters is through  $|\eta| < 3.2$  and the resolution is  $\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$ .

The hadronic calorimeter, located directly behind the electromagnetic calorimeter, is composed of the Tile Calorimeter (TileCal), the LAr hadronic endcap calorimeter (HEC) and the LAr forward calorimeter. The TileCal contains tiles of iron absorber and plastic scintillator located in the barrel (which has coverage of  $|\eta| < 1.6$ ). The LAr HEC and forward calorimeters use the same technology as the electromagnetic calorimeter and extend the coverage to  $|\eta| < 4.9$ .

## 3.4.5 Muon Spectrometer

A diagram of the MS is shown in Figure 3.8 in the  $r - \phi$  plane.

The MS surrounds the calorimeters and measures the trajectories of muons as they traverse the



Figure 3.7: A diagram of the electromagnetic calorimeter system showing the various layers with different granularities that perform the measurement of particle showers.

detector. Momentum measurements are possible due to the presence of a toroidal magnetic field which bends the muons. The magnetic field is provided by the large barrel toroid in  $|\eta| < 1.4$ , smaller endcap magnets inserted in the ends of barrel toroid in  $1.6 < |\eta| < 2.7$  and a combination of the barrel and endcap fields in  $1.4 < |\eta| < 1.6$ .

Muon chambers are arranged in three cylindrical layers around the beam in the barrel region and layers perpendicular to the beam in the endcap regions. Measurements from the chambers in the barrel and most of the endcaps are constructed from Monitored Drift Tubes (MDTs) except in the range 2.0 <  $|\eta|$ 2.7 where Cathode Strip Chambers (CSCs) are used (this is due to higher incident particle flux). Muon triggers can be constructed from Thin Gap Calorimeters (TGCs) in the endcaps and Resistive Plate Chambers (RPCs) in the barrel that provide less precise hit information but can provide information very quickly when compared to the MDTs and CSCs.



Figure 3.8: A diagram of the MS showing the cylindrical layers of the muon chambers around the beam which include Monitored Drift Tubes, Cathode Strip Chambers, Thin Gap Calorimeters and Resistive Plate Chambers.

# 3.5 Object Reconstruction

# 3.5.1 Introduction

In order to use the measurements from the detector in an efficient way, the information must be classified into objects that are meaningful in the context of doing a physics analysis. Identification criteria for objects is generalized across many analyses and are used to infer properties of the hard scatter. Figure 3.9 shows a basic summary of how ATLAS detects a variety of particles.

# 3.5.2 Tracks and Calorimeter Clusters

A track represents a particle trajectory as it traverses the detector. This is done by first performing pattern recognition which identifies hits measured in the ID and MS as belonging to a single track



Figure 3.9: Basic summary of how ATLAS detects a variety of particles including muons, photons, hadrons and electrons.

and then running a fit algorithm to assess the trajectory. The result is an estimate of the momentum 3-vector of the particle. Note that only charged particles leave hits in the ID and MS, as a result neutral particles do not have associated tracks. Tracks must have a minimum  $p_T$  of 500 MeV to be considered in the analysis presented here.

Calorimeter clusters are constructed from groups of individual calorimeter cells and form a measurement of the energy of a given particle after it showers. The electromagnetic calorimeter measure clusters of energy created by the showering of electrons and photons while the hadronic calorimeter measures clusters of energy originating from hadrons (like protons and neutrons). Topological clusters are formed by including cells around a seed cell, defined as having a measured energy of  $4 \times \sigma_{noise}$ . If a cell with measured energy of  $2 \times \sigma_{noise}$  is adjacent to any cell in the cluster it is included in the cluster. If a  $0 \times \sigma_{noise}$  cell is found adjacent to the  $2 \times \sigma_{noise}$  cell it is also included. Clusters with separated energy deposits are separated into different clusters. Cluster energies are calibrated to match responses generated by simulation from single hadrons and corrections are applied to account for noise and pile-up interactions (described in Section 5.5.3).

#### **3.5.3** Electrons and Photons

A typical electron will traverse the detector and shower in the electromagnetic calorimeter, creating hits in the ID and an isolated cluster in the electromagnetic calorimeter with  $\Delta R < 0.1$ . A sliding window algorithm is used to seed the clustering algorithm which scans in  $\eta - \phi$  space over calorimeter cells searching for relative maxima. However, electrons can lose a significant amount of energy due to a phenomenon known as bremsstrahlung where the electron interacts with a nucleus in the detector and emits a high energy photon. Track fitting for electrons must consider the hypothesis that the electron underwent bremsstrahlung since it will change the pattern of the hits. The photon energy is usually within the same cluster in the electromagnetic calorimeter and which is accounted for by increasing the size of the sliding window in the  $\phi$  direction. This is because the electron will bend in the magnetic field but the emitted photon will not, causing a spatial separation in the calorimeter deposits. Tracks are then matched to clusters to form an electron object. The level of quality of an electron is determined by the shower shape, quality of the track and the presence of transition radiation.

Photon identification is similar, except photons do not leave hits in the ID unless they convert to a pair of electrons before traversing the calorimeter. Only unconverted photons are considered in the analysis presented here.

#### **3.5.4** Muons

A muon typically creates hits in the ID and MS which are then fitted to form 2 separate tracks. The tracks are then required to meet matching criteria to be consistent with being created by the same muon. Since muons leave little energy in the calorimeter systems, the combined track fit from the ID and MS form the momentum 3-vector. The level of quality of a muon is determined by number of hits in the ID and MS and the matching of the tracks.

## 3.5.5 Taus

Tau identification [23] uses various discrimination variables combined in a Boosted Decision Tree (BDT) in order to reject jets and electrons that can mimic taus. The BDT algorithms combine strip layer and calorimeter quantities as well as track momenta.

## **3.5.6** Jets

Due to the phenomenon of color confinement, quarks and gluons cannot be isolated and observed directly. This is because quarks and gluons have an intrinsic property known as color which has 3 states: red, blue and green. Color singlets cannot exist and so quarks and gluons clump together to form hadrons, a process known as hadronization. When a quark or gluon is emitted from a hard scatter, hadronization occurs and forms a "jet" of colorless hadrons which is collimated at the interaction point. Jets are reconstructed using calorimeter deposits in the hadronic calorimeter chosen by the anti-kt algorithm [24]. In the analysis presented here, only jets chosen with the anti-kt algorithm with  $\Delta R < 0.4$  are considered. The jets are calibrated to an energy scale generated by simulation and additional corrections are applied using data from Z+jets,  $\gamma$ +jets and dijet events.

# 3.5.7 b Jets

Jets originating from b quarks are given special consideration since the relatively long lifetime of the B meson allows for measurable in flight decays. Jets originating from b quarks are "tagged" using a multivariate (MVA) tagging algorithm, referred to as MV1 [25]. The algorithm combines information relevant to b identification, such as secondary vertices, which generate displaced tracks caused by the in flight decay of the B meson.

# 3.5.8 Missing Energy

By classifying the measurements of the ATLAS detector into physics objects, it is possible to take the sum of the transverse momentum of all detected objects to infer the total transverse momentum of invisible particles. This is done by balancing the summed transverse momentum. A given particle momentum in the transverse plane is referred to as  $p_T$ . The negative vector sum of the transverse momentum of invisible particles is known as "missing transverse momentum", or  $E_{\rm T}^{\rm miss}$ , and is calculated:

$$\sum \vec{p_T}(\text{visible}) + E_{\mathrm{T}}^{\mathrm{miss}} = 0 \tag{3.2}$$

$$E_{\rm T}^{\rm miss} = -\sum \vec{p_T}(\rm visible) \tag{3.3}$$

The definition of  $E_{\rm T}^{\rm miss}$  used in this analysis is referred to as METRefFinal. METRefFinal is the sum of physics objects calculated from clusters in the calorimeters and muons reconstructed in the

MS.

$$E_{\rm T}^{\vec{\rm miss}} = E_{\rm T}^{\vec{\rm miss},\ \rm calo} + E_{\rm T}^{\vec{\rm miss},\ \mu} \tag{3.4}$$

The calorimeter clusters are calibrated according to the reconstructed physics object such as electrons, photons,  $\tau$  leptons and jets. Remaining calorimeter clusters that are not associated to objects are also included as the "soft term".

$$E_{\rm T}^{\rm miss, \ calo} = E_{\rm T}^{\rm miss, \ e} + E_{\rm T}^{\rm miss, \ \gamma} + E_{\rm T}^{\rm miss, \ \tau} + E_{\rm T}^{\rm miss, \ jets} + E_{\rm T}^{\rm miss, \ soft \ term}$$
(3.5)

One deficiency of METRefFinal is the fact that it performs poorly in the presence of interactions arising from protons not involved in the hard scatter, known as pile-up interactions. Since bunches of protons are collided instead of individual protons, it is very likely to get additional collisions in a given event which present noise to the  $E_{\rm T}^{\rm miss}$  calculation. In the future, more tracking information (which can be associated to a specific vertex) will be used to make a more pile-up resistant version of METRefFinal.

# CHAPTER 4

# Trigger

## 4.1 Introduction

Bunches of protons collide every 50 ns in the ATLAS detector. The trigger system aims to reduce the collision rate of 20 MHz to a more manageable rate of 400 Hz, which is the rate that can actually be written to disk. The decision to write an event a disk is determined by the trigger system and is motivated by the probability of the event containing interesting physics properties. The trigger uses basic information (basic signatures of physics worth studying) to make the decision in real time. Most collisions are actually inelastic proton-proton collisions that are not worth writing to disk.

The trigger system is divided into 3 levels: Level 1 (L1), Level 2 (L2) and Event Filter (EF). The trigger system levels increase in computing complexity. The L1 trigger is composed of custom electronics, which reside in a cavern to the side of the main detector cavern. The L1 trigger receives the full LHC bunch crossing rate of 20 MHz and outputs 75 kHz. The limiting factor is the buffer on the custom electronics in the cavern, which requires that decisions be made in 2.5  $\mu$ s. The L1 trigger forms Regions of Interest (RoI) based on a limited set of information from a subset of detectors. The RoIs are based on coarse granularity calorimeter information used to crudely identify electrons, photons,  $\tau$  leptons, jets,  $E_{\rm T}^{\rm miss}$  and muons using fast muon chambers. The remaining trigger systems (L2 and EF) are implemented in software and are collectively referred to as the High Level Trigger (HLT). The L2 trigger uses RoI defined by L1 to analyze further using fine-grained detector data (to perform more detailed reconstruction of hard objects like leptons, jets and photons) in a window around the position of the RoI. The L2 trigger reduces the event rate to 5 kHz. Finally, the EF reconstructs events with similar reconstruction algorithms used offline that run faster to produce a more detailed calculation of particle momenta and energies. The EF reduces the L2 rate down to 400 Hz and writes the events to disk. Figure 4.1 shows a schematic overview of the trigger system [26]. The rates in Figure 4.1



Figure 4.1: Schematic overview of the trigger system [26]. Rates are from the original design specification for Run I which includes 25 ns bunch spacing increasing the collision rate to 40 Mhz and a lower EF output of 100 Hz.

represent the original design specification for Run I. The original Run I specification includes 25 ns bunch spacing, increasing the collision rate to 40 MHz, and a lower EF output of 100 Hz.

For some triggers, the event rates are beyond the rate limitations mentioned above and as a result only randomly selected events are selected. These triggers are "prescaled" with the prescale value defined as the inverse of the probability that an event gets selected by the trigger even though it meets the specific trigger threshold. For example, a trigger with a prescale of 3 selects 1/3 of the events that pass the decision.

# 4.2 Enhanced Bias

The total output bandwidth that is written to disk (400 Hz) is allocated to many triggers that select events due to the presence of  $E_{\rm T}^{\rm miss}$ , jets, leptons and photons. The rate for each trigger must be calculated carefully in order to avoid going over the bandwidth limit. Calculating trigger rates using MC is difficult because an unbiased event sample would need to be generated. An unbiased event sample is required because objects reconstructed by the trigger are largely not the targeted objects. Instead, reconstructed trigger objects are usually misidentified hadronic activity (particularly true at L1). For example, objects identified as electrons by the trigger at L1 are mostly misidentified jets. It is impractical to generate a sufficient number of MC events with all possible physics processes included. Consequently, the existing total ATLAS unbiased MC is equivalent to only a fraction of a second of data taking. Calculating trigger rates with these unbiased MC samples would result in rates with very high statistical error. Instead, data is used.

The ideal method to estimate trigger rates using data is to collect minimum bias events, an unbiased selection of data to run trigger algorithms on and derive rates. However, the high selectivity of the trigger, which reduces the collision rate from 20 Mhz to 400 Hz, a factor of 50,000, means that a minimum bias trigger would collect a set of events that are largely not useful in predicting rates. This is because the trigger generally selects events containing hard objects, and the majority of collisions contain do not contain any. As a result, it is impractical to record enough minimum bias events to calculate rate predictions with low enough statistical uncertainty to be meaningful.

Instead, a set of sufficiently general triggers are used to collect "enhanced bias" data, which are used to make trigger calculations. Triggers used to collect enhanced bias events select events using a selection of very loose L1 triggers. By using a mixture of L1 triggers, the probability of a hard object being in the event is much higher. As a result, events collected with enhanced bias triggers have a much higher efficiency when run through the trigger menu, and can be used to calculate rates with low enough statistical error to be useful.

The enhanced bias triggers are:

1.  $EF_{eb_physics}$ : Events are seeded from a random filled L1 trigger (L1\_RD0\_filled). The  $EF_{eb_physics}$  trigger then selects events in the HLT that pass any of a list of L1 trigger criteria. The list of L1 trigger criteria is run dependent but typically include selections on low  $E_{T}^{miss}$ , low  $p_{T}$  jets, muons, taus and electrons. Seeding the selection with a random filled L1 trigger and selecting events that pass a list of L1 trigger criteria at the HLT instead of at the L1 allows the  $EF_{eb_physics}$  prescale to be applied in a correlated way to the triggers in the L1 trigger criteria list. This is done to be able to calculate overlap and maintain correlations within the list of L1 triggers used at the HLT. Selecting events at L1 using the individual L1 triggers in the list will give little information about the correlations in the L1 trigger list. For example, if an event passes 2 L1 triggers in the list, the event would have pass both prescales. By using the random filled L1 trigger as a seed, the event would just need to pass the  $EF_{eb_physics}$  prescale and the correlation between both triggers would be maintained.

- 2. EF\_high\_eb\_physics: The same as EF\_eb\_physics except with a more stringent list of L1 trigger criteria.
- 3. EF\_eb\_physics\_noL1PS: Events are seeded from passing any triggers from a list of L1 trigger criteria which are not prescaled. This is done in order to collect sufficient statistics for L1 triggers with low rates. The random seed mechanism used in EF\_eb\_physics and EF\_high\_eb\_physics is not needed because there is no prescale in the list of L1 trigger criteria.
- 4. EF\_eb\_random: Events are seeded from randomly selected events at L1 (L1\_RD1\_filled), which are used to calculate rates for triggers with lower thresholds than those included in EF\_eb\_physics.

In order to prevent the event rate from being too high, the enhanced bias chains are prescaled. However, the prescales are set to maintain high enough statistical precision to allow predictions for the wide range of unprescaled triggers.

# 4.3 Enhanced Bias Weighting Tool

Since events selected by the enhanced bias triggers typically have hard objects, the events are typically high  $p_T$ . The proportion of events at high  $p_T$  is higher than it would be in events selected a minimum bias trigger. To account for this, a weight is calculated according to the enhanced bias trigger prescales. Without the weight, the rate prediction at high  $p_T$  would be overestimated. The reweighted enhanced bias events can then be used to estimate trigger rates to better plan the set of triggers to implement in future runs. The method follows:

- Events considered must pass one of the enhanced bias triggers: EF\_eb\_physics, EF\_high\_eb\_physics, EF\_eb\_physics\_noL1PS, EF\_eb\_random
- 2. Use the trigger configuration to determine the list of L1 triggers used in the enhanced bias triggers
- 3. Determine if the event would have passed one of the L1 triggers without including the effect of the prescale
- 4. Calculate the event weight:

weight = 
$$\frac{1}{1 - \prod_{i} (1 - \frac{1}{PS_i})}$$
 (4.1)

where i is runs over the 4 enhanced bias triggers that the event would have passed without the prescale (the result of Step 3).



Figure 4.2: Distribution of events triggered with L1\_emtau\_EMClus before and after the reweighting procedure is applied. Since the enhanced bias triggers preferentially select high  $p_T$  events, the distribution contains a higher proportion of events at high  $p_T$  than would be collected by a minimum bias trigger. The enhanced bias weighting procedure corrects for this effect. The weighted events is overlayed with minimum bias events (selected by EF\_eb\_random) in (b), and it is seen that the distributions agree. The statistical precision obtained by using enhanced bias triggered events is also seen, since there are many more high  $p_T$  events than selected by the minimum bias trigger.

The effect of the weighting procedure is shown in Figure 4.2. Before the weighting procedure, there is a higher proportion of events at high  $p_T$  than would be collected with a minimum bias trigger. This is because the enhanced bias triggers preferentially select high  $p_T$  events. After the reweighting procedure, the distribution agrees with the distribution of events taken with EF\_eb\_random, which represents a minimum bias sample. The usage of enhanced bias triggers results in a distribution of events which extends much further in  $p_T$ .

The weighted enhanced bias sample has much higher statistical precision than the sample selected by the  $EF_eb_random$  trigger, which makes the weighted sample useful in determining trigger rates for rare high  $p_T$  events. A comparison of the online rates and rates predicted using the enhanced bias triggers for the L1, L2, EF and several main physics streams are presented in Figure 4.3 [27]. The agreement of the rates (within 10%) makes it clear that the procedure of using enhanced bias data to predict online rates is effective.



Figure 4.3: Comparison of rates taken online against rate predictions calculated from enhanced bias data for the L1, L2, EF and several main physics streams [27]. The statistical uncertainty on the rates is considered negligible.

# CHAPTER 5

# Analysis Strategy

## 5.1 Introduction

In order to observe an invisibly decaying Higgs in the detector, the signature must involve a sufficiently high  $p_T$  Higgs, which is equivalent to requiring a signature with high  $E_T^{\text{miss}}$ . Therefore,  $E_T^{\text{miss}}$  triggers are used to select signal candidates: EF\_XE80\_TCLCW and EF\_XE80\_TCLCW\_LOOSE.

Various Higgs production modes present different approaches to invisibly decaying Higgs searches. The production modes have vastly different cross sections, shown in Figure 5.2. The production mode with the highest cross section is gluon gluon Fusion (ggF). In order to have a boosted Higgs in the ggF production mode, there must be an additional jet from Initial State Radiation (ISR) to conserve momentum. This is known as the monojet search since the signature is high  $E_{\rm T}^{\rm miss}$  and 1 jet from ISR. The high  $E_{\rm T}^{\rm miss}$  requirement provides the additional benefit of drastically reducing contamination from multijet events. This is especially important since the multijet cross section is several orders of magnitude larger than signal. After requiring high  $E_{\rm T}^{\rm miss}$  and applying additional requirements aimed at reducing the multijet contamination, the largest background is strong Drell-Yan production  $(Z \to \nu\nu + \text{jets})$ . The distinction between weak and strong production is described in Section 5.4.1. Weak Drell-Yan production is a negligible component of the total Drell-Yan background. Since there are no observable characteristics of strong  $Z \to \nu\nu + \text{jets}$  that can distinguish it from signal, the Drell-Yan background is considered irreducible. The ratio of signal and strong Drell-Yan is approximated:

$$\frac{ggF \text{ Signal}}{\text{Strong } Z \to \nu\nu + \text{jets}} \approx \frac{700}{30000} = 0.02 \tag{5.1}$$

The size of the signal expectation is 2 orders of magnitude below the Drell-Yan expectation. This is due to the fact that the diagrams which compose Drell-Yan production are tree-level, while Higgs production from ggF is suppressed by a loop coupling. The Higgs production mode with the second highest cross section is Vector Boson Fusion (VBF), shown in Figure 5.1, which is the channel considered in this analysis. Despite having a cross section which is an order of magnitude lower than ggF production, VBF has several advantages. Theoretical uncertainties associated with VBF are generally lower than ggF. This is due to the lack of the presence of loops in VBF type diagrams which leads to lower corrections due to higher order calculations. Additionally, signatures involving VBF have several distinctive features which can be used to reduce contamination from processes that mimic signal. In VBF processes, the two final state quarks typically form two high  $p_T$  jets that are kinematically favored to point along the beamline. Further, since the initial state quarks radiate color singlets, there is no color flow between the initial state quarks. This means gluon radiation is suppressed between both of the final state quarks, resulting in suppressed central jet activity.

Similar to the monojet search, having a high  $E_{\rm T}^{\rm miss}$  requirement for the VBF search drastically reduces contamination from multijet events that can mimic signal. After applying a selection that is optimized to select events with VBF features, described in detail in Section 5.3, the largest background in the signal region is Drell-Yan production  $(Z \to \nu\nu+{\rm jets})$ . However, unlike the monojet search, the Drell-Yan background is a mix of strong and weak production. Weak Drell-Yan production has the same distinctive VBF features that the signal has, while strong production does not. For example, strong  $Z \to \nu\nu+{\rm jets}$  production typically contains central jets due to color exchange from the initial state quarks and gluons. By applying a selection that incorporates VBF features, the strong component of the Drell-Yan background is reduced. The second largest background in the signal region is  $W \to \ell\nu+{\rm jets}$ , where the lepton escapes detection. The  $W \to \ell\nu+{\rm jets}$  background has similar features to the  $Z \to \nu\nu+{\rm jets}$  background. As a result, only the  $Z \to \nu\nu+{\rm jets}$  is considered to approximate the sensitivity.

The ratio of signal and background ratio after applying a selection which incorporates VBF features, described in detail in Section 5.3, is approximated:

$$\frac{VBF \text{ Signal}}{\text{Weak and Strong } Z \to \nu\nu + \text{jets}} \approx \frac{300}{350} = 0.85$$
(5.2)

The size of the signal expectation is the same order of magnitude as the total Drell-Yan background expectation. This is because the VBF signal is produced via tree-level electroweak couplings (WWH and ZZH) and is not suppressed by a loop couplings. The high signal to background ration indicates that the VBF analysis is expected to be a sensitive way to search for an invisibly decaying Higgs.

However, one important consideration of performing a VBF search is the systematic uncertainties introduced by selecting events with jets kinematically favored to point along the beamline. In order to



Figure 5.1: Feynman diagram for an invisibly decaying Higgs produced via VBF.



Figure 5.2: Higgs production cross sections for various modes of production.

avoid large systematic uncertainties, the dominant backgrounds  $(Z \rightarrow \nu\nu + \text{jets} \text{ and } W \rightarrow \ell\nu + \text{jets})$  are normalized using control regions, described in detail in Chapter 6. Additionally, even though the high  $E_{\text{T}}^{\text{miss}}$  requirement drastically reduces contamination from multijet events, a data driven method is used to determine the normalization, described in detail in Chapter 8. The dominant background and multijet estimates are then calculated in regions where one or two kinematic selections in the signal region are reversed, referred to as "validation regions". The robustness of the data driven background estimates is tested by comparing data driven estimates of the backgrounds to the total observed data yields in the validation regions.

In summary, the analysis strategy follows:

- Determine an event selection that is optimized for VBF features and background reduction.
- Determine data driven estimates for the dominant backgrounds, Z → νν+jets and W → ℓν+jets, using W → ℓν+jets (where the lepton is detected) and Z → ℓℓ+jets control regions. This is described in detail in Chapter 6.
- Determine a data driven estimate of the multijet background to ensure that it is negligible. This is described in detail in Chapter 8.
- Verify these estimates in "validation regions", where one or two kinematic cuts applied in the signal region are reversed.
- Measure the branching fraction of the Higgs decaying invisibly. In the absence of an excess, place an upper bound limit on the branching fraction.

# 5.2 Object Definitions

The following object definitions are used in the signal and control regions:

- For an event to be considered, all jets with  $p_T > 20$  GeV are required to pass quality requirements to reduce detector effects from spikes in the Hadronic Endcap Calorimeter (HEC), beam background from electromagnetic coherent noise, and background from cosmic and noncollision events. If a jet with  $p_T > 20$  GeV in a candidate event does not pass any of the quality requirements the event is not considered.
- Jets are required to have  $p_T > 30$  GeV and  $|\eta| < 4.5$ . Additionally, there is a requirement on the fraction of the jet energy associated with a particular primary vertex, known as the Jet Vertex Fraction (JVF) [28]. This requirement is aimed at suppressing events with jets from pile-up

interactions. The JVF for a given jet is calculated by taking the ratio of the sum of the  $p_T$  of all tracks matched to the jet and primary vertex to the sum of the  $p_T$  of all tracks matched to the jet. Jets are required to have |JVF| > 0.5 for jets with  $p_T < 50$  GeV and  $|\eta| < 2.4$ 

- Candidate events in the signal region and all control regions are rejected if they contain a jet identified as originating from b quark (using the 80% working point) [25] with  $p_T > 20$  GeV
- There are 3 electron identification operating points known as Loose, Medium and Tight with decreasing probabilities of the identified object originating from a light quark or gluon, photons that have converted to electrons and heavy flavor quark decays, which involve electrons [29]. The Tight++ operating point is required for electrons in the control regions described in Chapter 6. The Tight++ operating point includes requirements on shower shape variables, shower width variables, transition radiation, track-cluster matching, conversion rejection and hits in the b-layer. In the signal region, a candidate event is rejected if it contains a Medium++ electron with  $p_T > 10$  GeV. The Medium++ operating point includes requirements on shower shape and width variables.
- Photon identification is similar to electron identification except there is a lack of hits in the ID for unconverted photons. In the γ+jets control region described in Chapter 7, a candidate event is required to have an isolated, unconverted photon consistent with the Tight operating point.
- Muons are reconstructed from a combined fit of tracks from the Inner Detector (ID) and Muon Spectrometer (MS), referred to as StacoCombined [30]. Requirements are set on the number of hits in the tracking and the quality of the matching between the tracks. In the control regions described in Chapter 6, muons are required to be consistent with StacoCombined. In the signal region, a candidate event is rejected if it contains a muon with  $p_T > 5$  GeV as reconstructed by the StacoCombined algorithm with  $|\eta| < 2.5$  or by the Muon Spectrometer (MS) alone with  $2.5 < |\eta| < 2.7$ .
- Tau identification is performed with algorithms [23] which combine discriminating variables in a Boosted Decision Tree (BDT) to reject jets and electrons that can mimic taus. There are 3 tau identification operating points known as Loose, Medium and Tight which correspond to different tau identification efficiencies. In the signal region, a candidate event is rejected if it contains a Medium BDT tau with  $p_T > 20$  GeV.
- $E_{\rm T}^{\rm miss}$  is calculated by assigning topoclusters to reconstructed objects (leptons, jets, photons) and calibrating according to the object. This is effectively the negative vector sum of visible

objects in the detector, which by momentum conservation is equivalent to  $E_{\rm T}^{\rm miss}$ .

More detailed object definitions are found in Section 3.5.

# 5.3 Signal Region

In order to ensure basic quality for the data used in the analysis, a good run list and a primary vertex with  $N_{\text{track}} \geq 2$  are required. The primary vertex is defined as the vertex with the highest associated track  $\sum p_T^2$ . A good run list is the set of data taken during proper operation of all subdetectors. The two  $E_T^{\text{miss}}$  triggers (EF\_xe80\_tclcw and EF\_xe80\_tclcw\_loose) used to select signal candidates use a different definition of  $E_T^{\text{miss}}$  (trigger  $E_T^{\text{miss}}$ ) than offline  $E_T^{\text{miss}}$  (MetRefFinal described in Section 3.5.8) with far less granularity and precision. The considered  $E_T^{\text{miss}}$  triggers require trigger  $E_T^{\text{miss}} > 80$  GeV. Requiring offline  $E_T^{\text{miss}} > 150$  GeV ensures that the trigger is fully efficient and suppresses multijet events which typically have low values of  $E_T^{\text{miss}}$ . Events with offline  $E_T^{\text{miss}} > 150$  GeV have  $\approx 100\%$  probability of causing the trigger to activate. The trigger efficiency in the region below 150 GeV is typically difficult to estimate, which poses a problem when calculating Monte Carlo predictions in the region of  $E_T^{\text{miss}} < 150$  GeV. Therefore, it is preferable to not use events with  $E_T^{\text{miss}} < 150$  GeV.

The electron veto is formed by rejecting events with Medium++ electrons with  $p_T > 10$  GeV. The muon veto is formed by rejecting events with StacoCombined muons with  $p_T > 5$  GeV in  $|\eta| < 2.5$  or muons identified by the MS alone in  $2.5 < |\eta| < 2.7$ . The tau veto is formed by rejecting events with Medium BDT taus with  $p_T > 20$  GeV. The *b* veto is formed by vetoing events with jets identified as originating from a *b* quark decay (using the 80% operating point) with  $p_T > 20$  GeV.

Taking into consideration basic data quality, features of the VBF signature and cuts aimed to reduce background, the signal selection follows:

- 1. Preselection
  - Good Run List
  - At least one vertex with  $N_{\text{track}} \geq 2$
- 2. Trigger on EF\_xe80\_tclcw (Periods A and B) or EF\_xe80\_tclcw\_loose (Period C and later)
- 3. Lepton veto: Reject events with Medium++  $p_{T,e} > 10$  GeV, StacoCombined  $p_{T,\mu} > 5$  GeV in  $|\eta| < 2.5$  and MS only  $p_{T,\mu} > 5$  GeV in  $2.5 < |\eta| < 2.7$ , Medium BDT  $p_{T,\tau} > 20$  GeV
  - Suppresses  $Z \to \tau \tau + jets$ ,  $W \to \ell \nu + jets$  and  $t\bar{t}$  events

- 4. b veto: Using 80% working point with  $p_T > 20$  GeV
  - Suppresses  $t\bar{t}$  events
- 5. Jet  $p_T$ 
  - Leading jet  $p_T > 75 \text{ GeV}$
  - Subleading jet  $p_T > 50 \text{ GeV}$
  - High  $p_T$  jets are a feature of VBF production
- 6. Leading and subleading jets must be in opposite hemispheres, i.e.  $\eta_{jet 1} \times \eta_{jet 2} < 0$ 
  - High  $|\eta|$  separation of jets is a feature of VBF production.
- 7.  $\Delta \eta_{j,j} > 4.8$ 
  - Figure 5.3 shows the  $\Delta \eta_{j,j}$  distribution after requiring two hard jets in opposite hemispheres. It is observed that the signal is typically higher  $\Delta \eta_{j,j}$  than the backgrounds.
  - High  $|\eta|$  separation of jets is a feature of VBF production.

8.  $m_{jj} > 1 \text{ TeV}$ 

- High  $m_{jj}$  is a feature of VBF production.
- 9.  $\Delta \phi_{jet,X} < 2.5$ : The separation of any 2 jets in a candidate event must be less than 2.5 radians.
  - Suppresses multijet events which typically involves back to back jet production.
- 10. Third Jet Veto (reject events with third jet with  $p_T > 30 \text{ GeV}$ )
  - Lack of color flow from initial state quarks suppresses the production of additional jets, which is a feature of VBF production.
- 11.  $\Delta \phi_{jet, E_{T}^{miss}} > 1.0$ 
  - Suppresses multijet events where the  $E_{\rm T}^{\rm miss}$  is a result of a mismeasured jet.

12.  $E_{\rm T}^{\rm miss} > 150 {\rm ~GeV}$ 

• Suppresses multijet and ensures trigger is extremely efficient.



Figure 5.3:  $\Delta \eta_{j,j}$  after requiring two hard jets in opposite hemispheres. The disagreement of MC and data is due to an insufficient model of multijet production.

			Strong	Weak	Strong+Weak		Other	Total
Cut	ggF Signal	VBF Signal	$Z \rightarrow \nu \nu + \text{jets}$	$Z \rightarrow \nu \nu + \text{jets}$	$W \rightarrow \ell \nu + \text{jets}$	Multijet	Backgrounds	Backgrounds
$\Delta \eta_{jj} > 4.8$	$156 \pm 17.2$	$766 \pm 8.40$	$2333 \pm 39.1$	$503 \pm 6.27$	$3702 \pm 63.2$	$289317 \pm 158130$	$38.8 \pm 1.37$	$295895 \pm 158130$
$m_{jj} > 1 \text{ TeV}$	$129 \pm 15.6$	$673 \pm 7.87$	$1739 \pm 33.0$	$464 \pm 6.02$	$2983 \pm 55.3$	$188684 \pm 130368$	$30.8 \pm 1.22$	$193900 \pm 130368$
$\Delta \phi_{jj} < 2.5$	$126 \pm 15.4$	$606\pm7.47$	$1375 \pm 29.3$	$327 \pm 5.06$	$2149\pm47.0$	$114195\pm108395$	$23.3\pm1.04$	$118069 \pm 108395$
Jet Veto	$63.2 \pm 10.9$	$529 \pm 6.98$	$680 \pm 22.0$	$272 \pm 4.61$	$975 \pm 31.2$	-	$2.03\pm0.38$	$1928 \pm 38.5$
$\Delta \phi_{jx,MET} > 1$	$63.2 \pm 10.9$	$522\pm6.94$	$667 \pm 21.9$	$265 \pm 4.55$	$942 \pm 30.8$	-	$1.96 \pm 0.37$	$1875 \pm 38.1$
$E_{\rm T}^{miss} > 150 { m ~GeV}$	$17.1 \pm 5.51$	$267\pm4.96$	$214\pm8.00$	$111\pm2.95$	$225\pm10.6$	-	$0.71\pm0.21$	$550\pm13.6$

Table 5.1: Progressive yields of the selection in 20.3 fb<sup>-1</sup> of 2012 data, as evaluated using Monte Carlo. The expected signal yields are shown for the case of  $m_H = 125$  GeV. Even though the ratio of signal and background is higher before the  $E_{\rm T}^{\rm miss}$  requirement, events with  $E_{\rm T}^{\rm miss} < 150$  GeV are not used due to difficulty in estimating the trigger efficiency.

The progressive yields of the selection on signal and background processes are shown in Table 5.1. Figure 5.4 shows the  $E_{\rm T}^{\rm miss}$  distribution after the full selection except for the  $E_{\rm T}^{\rm miss}$  requirement is applied. The region below 150 GeV has a large disagreement with MC. The low  $E_{\rm T}^{\rm miss}$  region is dominated by multijet events, which there is insufficient MC statistics to predict. Typical variables of interest in this analysis are  $E_{\rm T}^{\rm miss}$  and  $m_{jj}$ , shown in Figure 5.5 after the full selection is applied.

# 5.4 Backgrounds

Once the signal selection is applied, the dominant background sources are from strong and weak single vector boson production with 2 additional jets, i.e.  $Z \to \nu\nu+2$  jets and  $W \to \ell\nu+2$  jets where the



Figure 5.4:  $E_{\rm T}^{\rm miss}$  with the full signal region selection applied except for the  $E_{\rm T}^{\rm miss}$  requirement.



Figure 5.5:  $E_{\rm T}^{\rm miss}$  and  $m_{jj}$  distributions after the full selection is applied evaluated using Monte Carlo overlayed with observed data. The expected signal yields are shown for the case of  $m_H = 125$  GeV.



Figure 5.6: Feynman diagrams for strong and weak Z+jets production.

lepton escapes detection. The contributions from multijet events, diboson production and  $t\bar{t}$  are much smaller. The backgrounds are described in detail in Section 5.6.

## 5.4.1 Strong vs Weak Production

The V+jets processes that mimic signal can be categorized into either strong or weak production. This distinction arises from the number of electroweak vertices, diagrams with two electroweak vertices are labeled as "strongly produced" and those with four or more are considered "weakly produced". An example Feynman diagram for each category is shown in Figure 5.6.

This distinction is important because higher order corrections are derived separately (explained in Section 5.5.2) and the Monte Carlo samples are separate.

# 5.5 Simulation and Data

# 5.5.1 Monte Carlo Samples

Monte Carlo samples are used to model signal and background processes including the hard parton scattering, underlying event activity, parton showering and hadronization. The hard parton scattering includes the generation of the final state particles, which are then passed through a simulation of the ATLAS detector. Underlying event activity originates from collisions of spectator quarks and gluons (those not involved in the hard scatter). Partons produced from the hard scatter as well as those from the underlying event also produce cascades of radiation from QCD interactions, known as parton showering. Since quarks and gluons are never found in isolation due to color confinement, hadrons are formed with other quarks and gluons spontaneously created in the vacuum and those found in the proton, a process known as hadronization. The VBF and ggF signal processes are modeled using Powheg, with Pythia8 handling parton showering, hadronization and underlying event. The invisible decay of the Higgs is simulated by forcing the Higgs to decay via  $H \rightarrow ZZ \rightarrow \nu\nu\nu\nu\nu$  with 100% branching ratio. The  $W \rightarrow \ell\nu$ +jets and  $Z \rightarrow \nu\nu$ +jets backgrounds (both strong and weak production) are modeled in slices of  $p_T^V$  with Sherpa 1.4.5 with MENLOPS handling parton showering, hadronization and underlying event. Backgrounds due to diboson processes (WW, WZ and ZZ) are modeled using Herwig with Jimmy handling parton showering, hadronization and underlying event. Some of the diagrams for diboson production are double counted since they are also simulated in the weakly produced V+jets Sherpa samples. Since the diboson backgrounds sum to less than 1 event after the full selection is applied, the double counting effect can be safely neglected and is ignored. Background due to  $t\bar{t}$  production is modeled using Powheg with Pythia8 modeling parton showering, hadronization and underlying event.

#### 5.5.2 Higher Order Corrections

The dominant background Monte Carlo predictions are normalized to include NLO ( $\mathcal{O}(\alpha_S^2)$ ) perturbative QCD corrections using matrix element calculators. The production cross sections for the strongly produced components of the Z+jets and W+jets processes are calculated with FEWZ [31]. Production cross sections for the weakly produced components of the Z+jets and W+jets processes are calculated with VBFNLO [32, 33, 34]. VBFNLO does not calculate electroweak radiative corrections, which are expected to be small. Cross sections calculated with VBFNLO include selection criteria applied at the generator level described in Section 7.15.

In order to account for electroweak radiative corrections, a  $p_T$  dependent correction derived from HAWK [35] is applied to the signal VBF Monte Carlo prediction. The  $p_T$  based correction computed using HAWK is shown in Figure 5.7. The  $p_T$  distribution of the signal ggF Monte Carlo prediction is normalized following the calculation in [36] which includes NNLL and NNLO effects.

#### 5.5.3 Data

The entire 2012 ATLAS dataset was used in this search corresponding to 20.3 fb<sup>-1</sup> at  $\sqrt{s} = 8$  TeV and 50 ns bunch spacing. The size of the dataset is computed after imposing data quality requirements, which ensure proper operations of subdetectors.

# 5.6 Background Composition

Once the signal region selection is applied, the background composition is the following:



Figure 5.7: The Higgs  $p_T$  distribution including electroweak radiations as calculated by HAWK.

- $Z \rightarrow \nu \nu + \text{jets:} \approx 60\%$ 
  - Weak and strong production of a Z boson in association with 2 jets where the Z decays invisibly to neutrinos. Weak production accounts for roughly 1/3 of the total  $Z \rightarrow \nu\nu$ +jets background while strong production accounts for the other 2/3. The strong component is reduced by the central jet veto and high spatial separation of the jets. The weak component is considered irreducible.
- $W \to \ell \nu + \text{jets:} \approx 40\%$ 
  - Weak and strong production of a W boson in association with 2 jets where the W decays to a lepton and neutrino. These events are suppressed by rejecting events that contain a lepton. However, there is still a significant contribution from W+jets events where the lepton escapes detection. This can happen when the lepton is out of the  $\eta$  acceptance of the detector or the lepton is too soft in  $p_T$  to be identified. The dominant source (60%) of the  $W \to \ell \nu$ +jets background comes from  $W \to \tau \nu$ +jets where the lepton decaying

from the  $\tau$  is soft and escapes detection. The remaining neutrino from the  $\tau$  decay and the neutrino from the W decay combine to create a large value of  $E_{\rm T}^{\rm miss}$ . Weak production accounts for 1/3 of the total  $W \to \ell \nu + {\rm jets}$  production while strong production accounts for the other 2/3. Likewise, strongly produced  $W+{\rm jets}$  can be reduced by the central jet veto and high spatial separation of the jets.

- Diboson,  $t\bar{t}$ , multijet: < 1%
  - Diboson events (WZ, WW, ZZ) have low cross sections and are also suppressed by the high spatial separation of the jets.
  - $-t\bar{t}$  events are suppressed by the high spatial separation of the jets, b veto and lepton veto.
  - Multijet events pass the signal region selection due to  $E_{\rm T}^{\rm miss}$  caused by instrumental effects (like mismeasured jets). Suppression of multijet events comes from requiring high  $E_{\rm T}^{\rm miss}$ and requiring that none of the jets be within  $\Delta \phi$  of 1.0 radians of the  $E_{\rm T}^{\rm miss}$  to avoid mismeasured jets. The leading jet must be at least 1.6 radians from the  $E_{\rm T}^{\rm miss}$ . Since a typical multijet event will be 2 jets back to back in  $\phi$ , there is an additional requirement that  $\Delta \phi_{j,j} < 2.5$ .

# 5.7 Validation Regions

In order to validate the data driven background estimates, three "validation" regions are defined in neighboring regions of phase space depleted in signal, defined in the following way:

- 1. Reversing the  $\Delta \eta_{j,j} > 4.8$  requirement to  $\Delta \eta_{j,j} < 3.8$
- 2. Reversing the third jet veto by requiring a third jet with  $p_T > 40 \text{ GeV}$
- 3. Reversing both cuts simultaneously, i.e requiring  $\Delta \eta_{j,j} < 3.8$  and a third jet with  $p_T > 40$  GeV

MC yields of the validation regions are shown in Table 5.2. It is seen that the signal expectation is small when compared to the background yields. It is also seen that the 3 validation regions are mostly composed of  $W \rightarrow \ell \nu$ +jets and  $Z \rightarrow \nu \nu$ +jets so they can be used to validate the data driven estimates.

	ggF+VBF	Strong+Weak	Strong+Weak		Other	Total						
Cut	Signal	$Z \rightarrow \nu \nu + \text{jets}$	$W \rightarrow \ell \nu + \text{jets}$	Multijet	Backgrounds	Backgrounds	2012 Data	Subtracted				
Small- $\Delta \eta_{jj}$ Control Region												
$p_T^{j1,j2,j3}$	$7543 \pm 134$	$280294 \pm 335$	$356553 \pm 539$	$9274322 \pm 726126$	$8063 \pm 19.9$	$9919233 \pm 726127$	7206755	$6561845 \pm 2759$				
Opp. Hemispheres	$4597 \pm 90.1$	$120954 \pm 225$	$143490 \pm 341$	$3787720 \pm 464794$	$2401 \pm 10.9$	$4054565 \pm 464794$	3107542	$2840697 \pm 1810$				
$\Delta \eta_{ij} < 3.8$	$2564 \pm 78.3$	$109213 \pm 211$	$128513 \pm 319$	$3040348 \pm 401639$	$2227 \pm 10.5$	$3280301 \pm 401639$	2722646	$2482693 \pm 1694$				
$m_{jj} > 1 \text{ TeV}$	$33.23 \pm 3.62$	$2103 \pm 25.8$	$2628 \pm 39.6$	$144071 \pm 10101$	$33.8 \pm 1.28$	$148836 \pm 10101$	140644	$135879 \pm 378$				
$\Delta \phi_{jj} < 2.5$	$18.5 \pm 1.27$	$443.9\pm8.94$	$430 \pm 11.4$	$4754 \pm 1698$	$8.79\pm0.64$	$5636 \pm 1698$	6025	$5142 \pm 79.0$				
Jet Veto	$13.3 \pm 1.07$	$159.1 \pm 5.23$	$97.7 \pm 4.55$	-	$0.05 \pm 0.05$	$257 \pm 6.93$	228	$-28.9 \pm 16.6$				
$\Delta \phi_{jx,MET} > 1$	$12.7 \pm 1.05$	$154.9 \pm 5.19$	$96.2 \pm 4.53$	-	$0.05 \pm 0.05$	$251 \pm 6.89$	219	$-32.2 \pm 16.3$				
$E_T^{miss} > 150 \text{ GeV}$	$12.1 \pm 1.02$	$132.5 \pm 3.30$	$85.4 \pm 4.16$	-	$0.05 \pm 0.05$	$218 \pm 5.31$	185	$-33.0 \pm 14.6$				
3-jet Control Region												
$0 p_T^{j1,j2,j3}$	$3877 \pm 108$	$164705 \pm 255$	$233379 \pm 423$	$6894696 \pm 621809$	$7443 \pm 18.9$	$7300222 \pm 621810$	5005243	$4599716 \pm 2291$				
Opp. Hemispheres	$2083 \pm 74.4$	$74201 \pm 174$	$99327 \pm 278$	$2748453 \pm 384932$	$2271 \pm 10.5$	$2924253 \pm 384932$	2256994	$2081195 \pm 1538$				
$\Delta \eta_{ij} > 4.8$	$280 \pm 21.5$	$1666 \pm 30.2$	$2569 \pm 53.4$	$289317\pm158130$	$37.4 \pm 1.33$	$293590 \pm 158130$	69047	$64775 \pm 270$				
$m_{ij} > 1 \text{ TeV}$	$235 \pm 19.2$	$1285 \pm 25.8$	$2071 \pm 47.0$	$188684\pm130368$	$29.6 \pm 1.18$	$192069 \pm 130368$	55651	$52266 \pm 242$				
$\Delta \phi_{ij} < 2.5$	$214 \pm 18.7$	$983 \pm 22.3$	$1474 \pm 39.2$	$114195 \pm 108395$	$22.3 \pm 1.01$	$116674 \pm 108395$	28752	$26272 \pm 175$				
$p_T^{j3} > 40 \text{ GeV}$	$80.1 \pm 12.8$	$469 \pm 15.3$	$749 \pm 27.5$	$109400 \pm 108288$	$18.0\pm0.87$	$110636 \pm 108288$	19906	$18670 \pm 145$				
$\Delta \phi_{ix,MET} > 1$	$48.5 \pm 8.82$	$240 \pm 10.7$	$295 \pm 17.3$	$108284 \pm 108284$	$5.65 \pm 0.49$	$108825 \pm 108284$	1768	$1227 \pm 46.7$				
$E_T^{miss} > 150 \text{ GeV}$	$27.2 \pm 5.37$	$103 \pm 4.80$	$84.0 \pm 5.47$	-	$2.16 \pm 0.31$	$189 \pm 7.28$	212	$22.5 \pm 16.3$				
3-jet, Small- $\Delta \eta_{ij}$ Control Region												
$p_T^{j1,j2,j3}$	$3877 \pm 108$	$164705 \pm 255$	$233379 \pm 423$	$6894696 \pm 621809$	$7443 \pm 18.9$	$7300222 \pm 621810$	5005243	$4599716 \pm 2291$				
Opp. Hemispheres	$2083 \pm 74.4$	$74201 \pm 174$	$99327 \pm 278$	$2748453 \pm 384932$	$2271 \pm 10.5$	$2924253 \pm 384932$	2256994	$2081195 \pm 1538$				
$\Delta \eta_{ii} < 3.8$	$1422 \pm 65.9$	$66849 \pm 164$	$88510 \pm 258$	$2186259 \pm 331456$	$2104 \pm 10.1$	$2343722 \pm 331456$	1973380	$1815917 \pm 1438$				
$m_{ii} > 1 \text{ TeV}$	$14.4 \pm 3.38$	$1552 \pm 22.2$	$2173 \pm 36.3$	$124792 \pm 9429$	$33.0 \pm 1.24$	$128549 \pm 9430$	113369	$109611 \pm 339$				
$\Delta \phi_{ii} < 2.5$	$7.35 \pm 0.81$	$329 \pm 7.63$	$365 \pm 10.7$	$4754 \pm 1698$	$8.79 \pm 0.64$	$5456 \pm 1698$	5869	$5167 \pm 77.7$				
No Jet Veto	$7.35 \pm 0.81$	$329 \pm 7.63$	$365 \pm 10.7$	$4754 \pm 1698$	$8.79\pm0.64$	$5456 \pm 1698$	5869	$5167 \pm 77.7$				
$\Delta \phi_{jx,MET} > 1$	$5.03 \pm 0.66$	$153 \pm 4.03$	$112\pm5.04$	$3.12\pm3.12$	$0.59 \pm 0.16$	$268 \pm 7.17$	260	$-5.39 \pm 17.4$				
$E_T^{miss} > 150 \text{ GeV}$	$4.73 \pm 0.64$	$140 \pm 3.51$	$94.0 \pm 4.03$	-	$0.45 \pm 0.13$	$234 \pm 5.34$	195	$-38.7 \pm 15.0$				

Table 5.2: Progressive yields of the validation regions described in Section 5.7 in 20.3 fb<sup>-1</sup> of 2012 data, as evaluated using Monte Carlo. Expected signal yields are shown for the case of  $m_H = 125$  GeV. The "Subtracted" column shows an estimate of the dijet contribution by subtracting Z, W and Other Background columns from the observed yield.

# CHAPTER 6

# Lepton Control Regions

## 6.1 Introduction

An important consideration of a final state involving very forward jets is that they introduce large systematic uncertainties in the high  $\eta$  region due to experimental and theoretical sources. Experimental sources arise from detector calibration uncertainties, which are large at high  $\Delta \eta_{j,j}$ . Theoretical sources arise from large perturbative uncertainties on the cross section and choice of PDF. Experimental and theoretical systematics are explained in detail in Chapter 9. In order to reduce these systematic uncertainties, control regions are introduced to normalize the dominant backgrounds in a data driven way.

Two control regions are introduced to normalize both V+jets backgrounds in the signal region:  $W \to \ell \nu$ +jets and  $Z \to \ell \ell$ +jets. The  $W \to \ell \nu$ +jets control region has the advantage of high statistics but also has a significant multijet contamination originating from a jet being misidentified as a lepton. The multijet contamination in the  $W \to \ell \nu$ +jets control region is determined in a data driven way by looking at events that almost pass lepton identification requirements. The  $Z \to \ell \ell$ +jets is not as strong statistically but has very low contamination from other processes and is more physically similar to the dominant background in the signal region ( $Z \to \nu \nu$ +jets).

A  $\gamma$ +jets control region was also considered (but ultimately not used) to normalize the  $Z \rightarrow \nu\nu$ +jets background in the signal region which is described in Chapter 7.

# 6.2 $W \rightarrow \ell \nu + jets$ Control Region

The  $W \to \ell \nu$ +jets control region is defined to be as kinematically similar as possible to the signal region. The main difference comes from requiring a lepton, which is used to trigger events. The  $E_{\rm T}^{\rm miss}$
of the signal region is emulated by adding the reconstructed lepton momentum to the offline  $E_{\rm T}^{\rm miss}$  $(E_{\rm T}^{\rm miss}' = E_{\rm T}^{\rm miss} + p_T^{\ell})$  and applying the signal region requirement  $E_{\rm T}^{\rm miss}' > 150$  GeV.

In the  $W \to \ell \nu$ +jets control region, W production is not charge symmetric. However, the multijet contamination in the W control region is. Therefore, the  $W \to \ell \nu$ +jets control region is split into positive and negative charged lepton samples. It is not required that the multijet contamination is charge symmetric, due to more positive charge being present in the collision of protons and protons. However, multijet production is observed to be charge symmetric in the misidentified lepton enhanced control regions described in Section 6.2.2. Additionally, since the multijet contamination in the  $W \to \ell \nu$ +jets control region is not symmetric with respect to the flavor of the lepton, the  $W \to \ell \nu$ +jets control region is further split into e and  $\mu$ . This is because there is far less multijet contamination in the  $W \to \mu \nu$ +jets control region than the  $W \to e\nu$ +jets control region due to low likelihood of a jet being misidentified as a muon. In total, the  $W \to \ell \nu$ +jets control region is split into 4 regions:  $W^+ \to e^+\nu$ ,  $W^- \to e^-\nu$ ,  $W^+ \to \mu^+\nu$  and  $W^- \to \mu^-\nu$ .

# 6.2.1 Event Selection

Using the same VBF jet topology introduced in Section 5.3 with a modified  $E_{\rm T}^{\rm miss}$  cut using  $E_{\rm T}^{\rm miss}$ ' the selection follows:

- 1. Preselection (explained in Section 5.3)
  - Good Run List
  - At least one vertex with  $N_{track} \geq 2$
- 2. Trigger
  - Electrons: Trigger with EF\_e24vhi\_medium1 or EF\_e60\_medium1
    - The EF\_e24vhi\_medium1 trigger requires the electron to deposit 25 GeV of energy in the electromagnetic calorimeter, be consistent with the "medium" operating point and have an isolated track. The selection of the medium operating point includes requirements on shower shape variables, leakage of the hadronic calorimeter and shower width variables in the "strips" (described in Section 3.4.4).
    - The EF\_e60\_medium1 trigger is similar to EF\_e24vhi\_medium1 except it requires electrons have  $p_T > 60$  GeV, there is no track isolation requirement and follows a slightly loosened operating point.
  - Muons: EF\_mu24i\_tight or EF\_mu36\_tight

- The EF\_mu24i\_tight trigger requires an inner detector track which matches a track from the muon spectrometer and an isolated track.
- The EF\_mu36\_tight trigger is similar to EF\_mu24i\_tight except it requires muons with  $p_T > 36$  GeV and there is no track isolation requirement.
- 3. Tight++ electron or StacoCombined muon with  $p_T^\ell > 30~{\rm GeV}$
- 4. b veto: Using 80% working point with  $p_T > 20$  GeV
- 5. Jet  $p_T$ 
  - Leading jet  $p_T > 75 \text{ GeV}$
  - Subleading jet  $p_T > 50 \text{ GeV}$
- 6. Leading and subleading jets must be in opposite hemispheres, i.e.  $\eta_{jet 1} \times \eta_{jet 2} < 0$
- 7.  $\Delta \eta_{j,j} > 4.8$
- 8.  $m_{jj} > 1 \text{ TeV}$
- 9.  $\Delta \phi_{jet,X} < 2.5$
- 10. Third Jet Veto (reject events with third jet with  $p_T > 30 \text{ GeV}$ )
- 11.  $\Delta \phi_{jet, E_{T}^{miss}} > 1.0$
- 12.  $E_{\rm T}^{\rm miss\,\prime} = (E_{\rm T}^{\rm miss} + p_T^{\ell}) > 150 {\rm ~GeV}$

# 6.2.2 Multijet Background

The  $W \to \ell \nu$ +jets control region has a significant contamination from multijet events where a jet is misidentified as a lepton. The multijet contamination in the  $W \to \ell \nu$ +jets control region is estimated in a data driven way using control samples enriched in misidentified leptons, referred to as multijet enhanced control regions. The multijet enhanced control regions are defined by selecting events that pass all of the requirements of the  $W \to \ell \nu$ +jets control region except for one of the lepton ID requirements. The multijet enhanced control regions are also split by flavor and charge, similar to the  $W \to \ell \nu$ +jets control region. The multijet enhanced control regions are formed with separate requirements for electrons and muons. In the case of electrons, an electron candidate must pass Medium++ but fail Tight++. For muons, the muon candidate must fail the impact parameter cut  $(d_0)$ . A fit is done in each of the multijet enhanced control regions (regions enhanced in misidentified  $e^+$ ,  $e^-$ ,  $\mu^+$ ,  $\mu^-$ ) to the transverse mass, defined as:

$$m_T = \sqrt{2E_\ell E_{\rm T}^{\rm miss} (1 - \cos(\Delta\phi_{\ell, E_{\rm T}^{\rm miss}}))} \tag{6.1}$$

where  $E_{\ell}$  is the lepton energy and  $\Delta \phi(\ell, E_{T}^{\text{miss}})$  is the transverse angle between the lepton and the direction of  $E_{T}^{\text{miss}}$ . Since multijet events lack a prompt neutrino, the  $E_{T}^{\text{miss}}$  tends to be lower and point along jet direction which has been misidentified as a lepton. Therefore, events in the multijet enhanced control regions tend have a lower  $m_T$  than events in the  $W \rightarrow \ell \nu$ +jets control region arising from W production.

The multijet enhanced control regions are also contaminated by  $Z \to \ell \ell + \text{jets}$ ,  $W \to \ell \nu + \text{jets}$  and "Other Backgrounds" (diboson and  $t\bar{t}$ ). All of the contamination is estimated using MC. The multijet yield in the multijet enhanced control regions is determined by taking the difference of the observed data yield and the contamination from other processes estimated from MC. The MC contamination yields and observed data yields in the multijet enhanced control regions are shown in Table 6.1. The multijet yield in the multijet enhanced control regions is equivalent to the "Subtracted" column. It is observed (as expected) that the multijet enhanced control regions enhanced in misidentified  $e^{\pm}$  have higher multijet yields than the corresponding regions for  $\mu^{\pm}$ .

The  $m_T$  distributions in the  $W \to \ell \nu + \text{jets}$  control regions are then fit with floating overall normalizations (referred to as scale factors) for the multijet contamination and prompt lepton  $(W \to \ell \nu$ and  $Z \to \ell \ell$ ) shapes. The  $m_T$  shapes of the multijet contamination are taken from the multijet enhanced control regions, while the shapes of the prompt lepton contributions are from MC. The scale factor for the prompt lepton contribution is expected to be close to unity since the prompt lepton MC prediction should not be modified significantly by the fit. Since the initial normalizations of the multijet contributions in the  $W \to \ell \nu + \text{jets}$  control regions comes from the multijet enhanced control regions, the scale factor is expected to be lower than 1. This is because the multijet enhanced control regions have more misidentified leptons than the  $W \to \ell \nu + \text{jets}$  control regions.

The flavor and charge categories are included by modifying the TFractionFitter algorithm included in ROOT. The  $m_T$  fit has 3 free parameters: the scale factor for the prompt lepton contribution, the scale factor for fake electrons and the scale factor for fake muons. Scale factors for the signal region and each of the validation regions described in Section 5.7 are shown in Table 6.2.

							Subtracted
Cut	$Z \to \ell \ell + jets$	Strong $W$	Weak $W$	Other Backgrounds	Total	2012 Data	(Expected Multijet)
				Misidentified $e^+$ Enh	anced		
$\Delta \eta_{jj} > 4.8$	$16.1 \pm 3.52$	$69.6 \pm 9.59$	$25.7\pm2.02$	$1.98 \pm 0.32$	$113 \pm 10.4$	562	$449 \pm 25.9$
$m_{jj} > 1 \text{ TeV}$	$14.0 \pm 3.46$	$49.9\pm6.78$	$24.2 \pm 1.96$	$1.21 \pm 0.23$	$89.4 \pm 7.87$	421	$332 \pm 22.0$
$\Delta \phi_{jj} < 2.5$	$10.9 \pm 3.29$	$40.2 \pm 5.90$	$18.1\pm1.69$	$0.86 \pm 0.19$	$70.0 \pm 6.97$	330	$260\pm 19.5$
Jet Veto	$4.61 \pm 1.94$	$24.7\pm4.81$	$15.0\pm1.54$	$0.26 \pm 0.11$	$44.6 \pm 5.41$	172	$127 \pm 14.2$
$\Delta \phi_{ix,E_{\infty}^{\text{miss}}} > 1$	$4.58 \pm 1.94$	$23.2\pm4.58$	$14.9\pm1.53$	$0.26 \pm 0.11$	$42.9 \pm 5.21$	168	$125 \pm 14.0$
$E_{\rm T}^{\rm miss} > 150 { m GeV}$	$0.55\pm0.18$	$6.47 \pm 1.65$	$6.77\pm1.03$	$0.16\pm0.10$	$14.0 \pm 1.96$	59	$45.1 \pm 7.93$
				Misidentified $e^-$ Enh	anced		
$\Delta \eta_{ii} > 4.8$	$20.0 \pm 8.15$	$40.6 \pm 6.42$	$10.5 \pm 1.28$	$2.18 \pm 0.33$	$73.3 \pm 10.5$	469	$396 \pm 24.0$
$m_{ii} > 1 \text{ TeV}$	$16.6 \pm 7.81$	$34.0\pm6.02$	$10.0\pm1.25$	$1.64 \pm 0.28$	$62.3 \pm 9.94$	357	$295 \pm 21.4$
$\Delta \phi_{ii} < 2.5$	$12.5 \pm 7.59$	$25.7\pm5.18$	$8.16\pm1.13$	$1.11 \pm 0.23$	$47.5 \pm 9.26$	282	$235 \pm 19.2$
Jet Veto	$9.41 \pm 7.57$	$14.7\pm4.48$	$7.27\pm1.06$	$0.07\pm0.06$	$31.4 \pm 8.85$	153	$122 \pm 15.2$
$\Delta \phi_{ix,E_{m}^{\text{miss}}} > 1$	$9.30 \pm 7.57$	$14.6\pm4.48$	$7.27\pm1.06$	$0.06\pm0.06$	$31.3 \pm 8.85$	150	$119 \pm 15.1$
$E_{\rm T}^{\rm miss} > 150 { m GeV}$	$0.80\pm0.37$	$7.04\pm3.71$	$2.85\pm0.66$	-	$10.7 \pm 3.79$	38	$27.3 \pm 7.23$
				Misidentified $\mu^+$ Enh	anced		•
$\Delta \eta_{jj} > 4.8$	$4.49\pm0.90$	$31.2\pm6.49$	$11.0\pm1.33$	$1.01\pm0.30$	$47.7 \pm 6.69$	55	$7.30 \pm 9.99$
$m_{jj} > 1 \text{ TeV}$	$4.03 \pm 0.86$	$21.4\pm4.50$	$10.8\pm1.31$	$0.77 \pm 0.28$	$36.9 \pm 4.78$	45	$8.00 \pm 8.23$
$\Delta \phi_{jj} < 2.5$	$3.41 \pm 0.82$	$16.1\pm4.09$	$8.56\pm1.17$	$0.64 \pm 0.26$	$28.7 \pm 4.34$	29	$0.29 \pm 6.92$
Jet Veto	$2.11 \pm 0.65$	$5.01 \pm 2.66$	$7.38\pm1.09$	-	$14.5 \pm 2.95$	17	$2.50 \pm 5.07$
$\Delta \phi_{ix,E_{\infty}^{\text{miss}}} > 1$	$2.08 \pm 0.65$	$5.01 \pm 2.66$	$7.10\pm1.08$	-	$14.2 \pm 2.94$	16	$1.81 \pm 4.97$
$E_{\rm T}^{\rm miss} > 150 { m GeV}$	$0.97\pm0.47$	$3.77\pm2.49$	$4.07\pm0.81$	-	$8.82 \pm 2.66$	7	$-1.81 \pm 3.75$
				Misidentified $\mu^-$ Enh	anced		•
$\Delta \eta_{jj} > 4.8$	$7.04 \pm 1.22$	$16.7\pm5.96$	$4.53\pm0.86$	$0.76 \pm 0.19$	$29.0 \pm 6.15$	46	$17.0 \pm 9.15$
$m_{jj} > 1 \text{ TeV}$	$6.31 \pm 1.15$	$15.3\pm5.93$	$3.82\pm0.79$	$0.49 \pm 0.14$	$26.0 \pm 6.09$	36	$10.1 \pm 8.55$
$\Delta \phi_{jj} < 2.5$	$4.76 \pm 0.92$	$13.8 \pm 5.85$	$3.09\pm0.71$	$0.41 \pm 0.13$	$22.1 \pm 5.97$	29	$6.94 \pm 8.04$
Jet Veto	$1.95 \pm 0.50$	$7.57\pm5.53$	$2.16\pm0.59$	$0.03\pm0.03$	$11.7 \pm 5.59$	22	$10.3 \pm 7.29$
$\Delta \phi_{ix, E_{T}^{\text{miss}}} > 1$	$1.91 \pm 0.50$	$7.57\pm5.53$	$1.98\pm0.56$	$0.03 \pm 0.03$	$11.5 \pm 5.58$	22	$10.5 \pm 7.29$
$E_{\rm T}^{\rm miss} > 150 {\rm ~GeV}$	$0.84 \pm 0.22$	$2.08\pm0.74$	$0.66\pm0.33$	$0.03 \pm 0.03$	$3.61 \pm 0.84$	10	$6.39 \pm 3.27$

Table 6.1: Expected yields for the multijet enhanced control regions used to estimate the multijet contamination in the  $W \rightarrow \ell \nu + jets$  control region.

Cut	SR	Small- $\Delta \eta_{jj}$	3-jet	3j, Small- $\Delta_{jj}$			
Real Lepton Scale Factors							
$\Delta \eta_{jj}$ Requirement	$0.9217 \pm 0.0298$	$0.9924 \pm 0.0034$	$0.8781 \pm 0.0301$	$0.9484 \pm 0.0042$			
$M_{jj} > 1$ TeV, $\Delta \phi_{jj} < 2.5$	$0.8995 \pm 0.0357$	$0.7805\pm0.0414$	$0.8594 \pm 0.0315$	$0.7663 \pm 0.1121$			
Additional Jets Requirement	$0.9545 \pm 0.0494$	$0.7551 \pm 0.1362$	$0.8771 \pm 0.0592$	$0.7663 \pm 0.1121$			
$\Delta \phi_{jx,E_{\mathrm{T}}^{\mathrm{miss}}}$	$0.9503 \pm 0.0501$	$0.7762 \pm 0.1275$	$0.9046\pm0.0476$	$0.7400\pm0.0972$			
	Fake Electr	ron Scale Factors					
$\Delta \eta_{jj}$ Requirement	$0.4064 \pm 0.0408$	$0.3642 \pm 0.0062$	$0.4308 \pm 0.0416$	$0.3840 \pm 0.0073$			
$M_{jj} > 1$ TeV, $\Delta \phi_{jj} < 2.5$	$0.3884 \pm 0.0507$	$0.3886\pm0.1186$	$0.4058 \pm 0.0472$	$0.4040 \pm 0.2180$			
Additional Jets Requirement	$0.3893 \pm 0.0731$	$0.3423 \pm 0.3541$	$0.3003 \pm 0.0812$	$0.4040 \pm 0.2180$			
$\Delta \phi_{jx,E_{\mathrm{T}}^{\mathrm{miss}}}$	$0.3872 \pm 0.0740$	$0.2951\pm0.3407$	$0.2532\pm0.0779$	$0.3865 \pm 0.2066$			
	Fake Muc	on Scale Factors					
$\Delta \eta_{jj}$ Requirement	$0.3669 \pm 0.2691$	$0.0000 \pm 0.0096$	$0.3191 \pm 0.5153$	$0.0008 \pm 0.0655$			
$M_{jj} > 1$ TeV, $\Delta \phi_{jj} < 2.5$	$0.4008 \pm 0.3522$	$0.8045\pm0.6601$	$0.4210\pm0.7465$	$1.0768\pm1.7445$			
Additional Jets Requirement	$0.2369 \pm 0.3887$	$0.5796\pm2.6696$	$0.2118\pm1.1249$	$1.0768\pm1.7445$			
$\Delta \phi_{jx,E_{\mathrm{T}}^{\mathrm{miss}}}$	$0.3417 \pm 0.3935$	$0.3517 \pm 2.2038$	$0.0000 \pm 0.2886$	$1.1985 \pm 0.9928$			

Table 6.2: Scale factors derived from fitting the  $m_T$  distributions in the  $W \to \ell \nu + \text{jets}$  control regions. The shape of the prompt lepton contributions are from W MC while the shape of the multijet contributions are from the multijet enhanced control regions.

				$W^+ \rightarrow e\nu$			
Cut	$Z \to \ell \ell + \text{jets}$	Strong $W$	Weak $W$	Multijets	Other BGs	Total	2012 Data
$\Delta \eta_{jj} > 4.8$	$116 \pm 10.5$	$906 \pm 27.8$	$392\pm7.98$	$343\pm37.3$	$55.0 \pm 2.64$	$1812 \pm 48.4$	1803
$m_{jj} > 1 \text{ TeV}$	$95.3\pm9.99$	$709 \pm 23.9$	$362\pm7.67$	$244\pm34.0$	$43.1\pm2.36$	$1453\pm43.5$	1420
$\Delta \phi_{jj} < 2.5$	$72.4 \pm 8.86$	$564 \pm 21.2$	$257\pm6.46$	$192\pm27.3$	$35.9\pm2.11$	$1121\pm36.3$	1078
Jet Veto	$38.8\pm8.18$	$268\pm15.6$	$213\pm5.88$	$101\pm54.9$	$10.2\pm1.29$	$631\pm58.0$	628
$\Delta \phi_{ix,E_{\mathrm{T}}^{\mathrm{miss}}} > 1$	$38.1\pm8.18$	$261\pm15.4$	$209\pm5.83$	$94.5\pm19.7$	$10.2\pm1.23$	$613 \pm 27.0$	609
$E_{\rm T}^{\rm miss} > 150 { m GeV}$	$5.87 \pm 0.68$	$92.3\pm7.22$	$99.4\pm4.02$	$28.0\pm6.78$	$4.02\pm0.73$	$230\pm10.7$	225
				$W^-  ightarrow e \nu$			
$\Delta \eta_{jj} > 4.8$	$118\pm7.90$	$537 \pm 21.8$	$207\pm5.81$	$343\pm37.3$	$42.6\pm2.06$	$1248 \pm 44.4$	1190
$m_{jj} > 1 \text{ TeV}$	$85.2 \pm 5.25$	$402\pm18.0$	$187\pm5.52$	$244\pm34.0$	$32.6\pm1.75$	$951\pm39.3$	893
$\Delta \phi_{jj} < 2.5$	$62.6 \pm 4.73$	$311\pm15.6$	$139\pm4.77$	$192\pm27.3$	$26.4 \pm 1.57$	$731\pm32.2$	691
Jet Veto	$23.5\pm2.39$	$150\pm10.9$	$114 \pm 4.32$	$101\pm54.9$	$4.16\pm0.69$	$393\pm56.2$	387
$\Delta \phi_{ix,E_{\infty}^{\text{miss}}} > 1$	$23.4 \pm 2.39$	$147\pm10.8$	$113 \pm 4.28$	$94.5\pm19.7$	$4.03\pm0.69$	$382 \pm 23.0$	380
$E_{\rm T}^{\rm miss} > 150 { m GeV}$	$7.33\pm0.96$	$55.1\pm5.33$	$52.5\pm2.92$	$28.0\pm6.78$	$1.80\pm0.43$	$145\pm9.17$	141
				$W^+ \to \mu \nu$			
$\Delta \eta_{jj} > 4.8$	$89.8 \pm 4.64$	$919\pm29.0$	$337 \pm 7.46$	$7.91\pm7.97$	$54.0 \pm 2.66$	$1408 \pm 31.4$	1292
$m_{jj} > 1 \text{ TeV}$	$72.5 \pm 4.06$	$707\pm24.6$	$313\pm7.18$	$7.33\pm8.06$	$42.2 \pm 2.34$	$1142 \pm 27.3$	1014
$\Delta \phi_{jj} < 2.5$	$56.3\pm3.60$	$556\pm21.9$	$227\pm6.12$	$2.93\pm5.02$	$34.4 \pm 2.12$	$877 \pm 23.7$	800
Jet Veto	$28.0 \pm 2.44$	$245\pm15.5$	$188\pm5.57$	$3.15 \pm 5.58$	$9.71\pm1.32$	$474 \pm 17.6$	480
$\Delta \phi_{ix, E_{\rm T}^{\rm miss}} > 1$	$25.7 \pm 2.23$	$240\pm15.4$	$184 \pm 5.52$	$4.34 \pm 5.87$	$9.27\pm1.29$	$463 \pm 17.6$	466
$E_{\rm T}^{\rm miss} > 150 { m GeV}$	$9.11\pm1.14$	$85.5\pm6.99$	$81.9\pm3.68$	$1.61\pm2.55$	$3.23\pm0.69$	$181\pm8.41$	182
				$W^-  ightarrow \mu \nu$			
$\Delta \eta_{jj} > 4.8$	$104 \pm 7.72$	$543 \pm 23.5$	$171\pm5.29$	$7.91\pm7.97$	$35.5 \pm 1.99$	$861\pm26.6$	830
$m_{jj} > 1 \text{ TeV}$	$84.1 \pm 6.77$	$402\pm20.3$	$154\pm5.02$	$7.33\pm8.06$	$27.9 \pm 1.76$	$675 \pm 23.5$	648
$\Delta \phi_{jj} < 2.5$	$69.8\pm6.61$	$332\pm19.0$	$116\pm4.37$	$2.93\pm5.02$	$22.7\pm1.59$	$543 \pm 21.3$	510
Jet Veto	$41.5 \pm 5.82$	$154\pm12.1$	$95.4\pm3.96$	$3.15\pm5.58$	$5.25\pm0.84$	$299 \pm 15.1$	264
$\Delta \phi_{jx, E_{\rm T}^{\rm miss}} > 1$	$35.1 \pm 4.02$	$147\pm11.9$	$92.9\pm3.91$	$4.34\pm5.87$	$5.25\pm0.84$	$285 \pm 14.4$	258
$E_{\rm T}^{\rm miss} > 150 { m GeV}$	$8.25\pm0.93$	$43.8\pm4.55$	$39.1\pm2.53$	$1.61\pm2.55$	$0.96\pm0.32$	$93.7 \pm 5.88$	98

Table 6.3: Expected yields for the  $W \rightarrow \ell \nu + \text{jets}$  control region as a function of the selection. The multijet background is determined using the data driven technique described in Section 6.2.2 and other contributions are estimated using MC.

## 6.2.3 Yields

Table 6.3 shows the expected yields for the  $W \to \ell \nu + j$ ets control region as a function of the selection including estimates of the multijet contamination from the data driven technique described in Section 6.2.2. Other contributions are estimated using MC. It is observed that the multijet contamination is much smaller for the  $W \to \mu \nu$  decays than  $W \to e \nu$  as expected.

# 6.3 $Z \rightarrow \ell \ell + jets$ Control Region

The  $Z \to \ell \ell$ +jets is also defined to be kinematically similar to the signal region. The production of  $Z \to \ell \ell$ +jets and  $Z \to \nu \nu$ +jets are identical except for  $\gamma^* \to \ell \ell$  production which is suppressed by requiring that  $m_{\ell\ell}$  be consistent with the Z mass.

Similar to the  $W \to \ell \nu + \text{jets}$  control region, the  $E_{\text{T}}^{\text{miss}}$  of the signal region is emulated by adding the reconstructed lepton momenta to the offline  $E_{\text{T}}^{\vec{\text{miss}}} = E_{\text{T}}^{\vec{\text{miss}}} + p_T^{\vec{\ell}1} + p_T^{\vec{\ell}2}$ .

# 6.3.1 Event Selection

Also using the same VBF jet topology introduced in Section 5.3 with the modified  $E_{\rm T}^{\rm miss}$  requirement using  $E_{\rm T}^{\rm miss\,\prime}$ , the selection follows:

- 1. Preselection
  - Good Run List
  - At least one vertex with  $N_{track} \geq 2$

# 2. Trigger

- Electrons: EF\_e24vhi\_medium1 or EF\_e60\_medium1
- Muons: EF\_mu24i\_tight or EF\_mu36\_tight
- 3.  $p_T^{\ell 1} > 30 \text{ GeV}$
- 4.  $p_T^{\ell 2} > 20 \text{ GeV}$
- 5.  $|m_{\ell\ell} m_Z| < 25 \text{ GeV}$
- 6. b veto: Using 80% working point with  $p_T > 20$  GeV
- 7. Jet  $p_T$ 
  - Leading jet  $p_T > 75 \text{ GeV}$
  - Subleading jet  $p_T > 50 \text{ GeV}$

8. Leading and subleading jets must be in opposite hemispheres, i.e.  $\eta_{jet\,1} \times \eta_{jet\,2} < 0$ 

- 9.  $\Delta \eta_{j,j} > 4.8$
- 10.  $m_{jj} > 1 \text{ TeV}$
- 11.  $\Delta \phi_{jet,X} < 2.5$
- 12. Third Jet Veto (reject events with third jet with  $p_T > 30 \text{ GeV}$ )
- 13.  $\Delta \phi_{jet, E_{T}^{miss}} > 1.0$
- 14.  $E_{\mathrm{T}}^{\mathrm{m}\bar{\mathrm{i}}\mathrm{ss}\,\prime} = (E_{\mathrm{T}}^{\mathrm{m}\bar{\mathrm{i}}\mathrm{ss}} + p_{T}^{\vec{\ell}1} + p_{T}^{\vec{\ell}2}) > 150 \text{ GeV}$

	Cuts for Signal Region									
Cut	Strong $Z \rightarrow ee$	Weak $Z \rightarrow ee$	Other Backgrounds	Total	Data	Strong $Z \rightarrow \mu \mu$	Weak $Z \rightarrow \mu \mu$	Other Backgrounds	Total	Data
$\Delta \eta_{jj} > 4.8$	$97.7 \pm 6.72$	$31.4 \pm 1.61$	$2.11 \pm 0.42$	$131 \pm 6.92$	123	$129 \pm 7.91$	$36.4 \pm 1.74$	$1.87 \pm 0.32$	$167 \pm 8.11$	183
$m_{jj} > 1 \text{ TeV}$	$75.3 \pm 5.92$	$28.6 \pm 1.53$	$2.05 \pm 0.42$	$106 \pm 6.13$	96	$100 \pm 7.07$	$33.2 \pm 1.66$	$1.75 \pm 0.31$	$135 \pm 7.27$	142
$\Delta \phi_{jj} < 2.5$	$57.4 \pm 5.04$	$19.6 \pm 1.27$	$1.54 \pm 0.36$	$78.6 \pm 5.21$	73	$81.4 \pm 6.45$	$22.0 \pm 1.35$	$1.16 \pm 0.22$	$105 \pm 6.59$	116
Central Jet Veto	$30.7 \pm 3.53$	$16.1 \pm 1.15$	$1.01 \pm 0.31$	$47.7 \pm 3.72$	41	$36.0 \pm 3.00$	$18.3 \pm 1.23$	$0.25 \pm 0.11$	$54.5 \pm 3.24$	65
$\Delta \phi_{ix,E_{\pi}^{\text{miss}}}$	$29.6 \pm 3.50$	$15.8 \pm 1.15$	$1.01 \pm 0.31$	$46.5 \pm 3.70$	40	$35.7 \pm 3.00$	$17.7 \pm 1.21$	$0.25 \pm 0.11$	$53.6 \pm 3.23$	63
$E_{\rm T}^{\rm miss} > 150 { m GeV}$	$10.4\pm1.54$	$7.35\pm0.79$	$0.34 \pm 0.19$	$18.1\pm1.74$	22	$14.0 \pm 1.48$	$8.24 \pm 0.83$	$0.15\pm0.09$	$22.4\pm1.70$	25

Table 6.4: Expected yields for the  $Z \rightarrow \ell \ell + jets$  control region as a function of the selection. All processes are estimated with MC.

### 6.3.2 Yields

Table 6.4 shows the expected yields for the  $Z \rightarrow \ell \ell + j$ ets control region as a function of the selection. All contributions are estimated from MC. It is observed that data and MC are in good agreement.

# 6.4 Estimation of V+jets

The signal region yields of  $W \to \ell \nu + \text{jets}$  and  $Z \to \nu \nu + \text{jets}$  used to determine the branching fraction limit are determined by the control regions in the fit mechanism described in Chapter 10.

However, it is important to make sure that using the control regions is justified. This is done using "transfer factors" to make estimates of the signal region processes. A unique feature of the usage of the control regions is that both the W and Z control regions are used to normalize the W and Z processes in the signal region. This can be done because the differences between the processes are negligible compared to experimental systematics related to jet energy scales and theoretical systematics due to perturbative uncertainties. In order to make sure the estimates from both control regions are consistent, transfer factors are used to make estimates of the W and Z processes in the signal regions using the W and Z control regions, defined in the following way:

$$N_{SR,\,\text{estimate}}^{Z \to \nu\nu} = \frac{N_{SR,MC}^{Z \to \nu\nu}}{N_{CR,MC}^{W \to \ell\nu}} N_{CR,Data}^{W \to \ell\nu}$$
(6.2)

$$N_{SR, \text{ estimate}}^{Z \to \nu\nu} = \frac{N_{SR,MC}^{Z \to \nu\nu}}{N_{CR,MC}^{Z \to \ell\ell}} N_{CR,Data}^{Z \to \ell\ell}$$
(6.3)

$$N_{SR,\,\text{estimate}}^{W \to \ell\nu} = \frac{N_{SR,MC}^{W \to \ell\nu}}{N_{CR,MC}^{W \to \ell\nu}} N_{CR,Data}^{W \to \ell\nu}$$
(6.4)

$$N_{SR,\,\text{estimate}}^{W \to \ell\nu} = \frac{N_{SR,MC}^{W \to \ell\nu}}{N_{CR,MC}^{Z \to \ell\ell}} N_{CR,Data}^{Z \to \ell\ell}$$
(6.5)

The estimates are designed such that mismodelings in the simulation due to theoretical and experimental systematics largely cancel in the ratio. The estimates are also designed such that the

	Signal Region							
Cut	$N^{Z \to \nu\nu}_{SR,MC}$	$N_{SR,\text{estimate}}^{Z\to\nu\nu}$ using $Z\to\ell\ell$	$N_{SR,\text{estimate}}^{Z \to \nu \nu}$ using $W \to \ell \nu$					
$\Delta \eta_{jj} > 4.8$	$2836\pm 39.6$	$2913 \pm 203$	$2685 \pm 82.6$					
$m_{jj} > 1 \text{ TeV}$	$2203\pm33.5$	$2175\pm171$	$2036\pm69.6$					
$\Delta \phi_{jj} < 2.5$	$1702\pm29.7$	$1759 \pm 157$	$1571 \pm 60.1$					
Jet Veto	$952 \pm 22.5$	$990 \pm 111$	$927 \pm 66.2$					
$\Delta \phi_{jx, E_{\rm T}^{\rm miss}} > 1$	$932 \pm 22.4$	$962 \pm 110$	$912 \pm 45.0$					
$E_{\rm T}^{\rm miss} > 150 { m GeV}$	$325\pm8.53$	$378\pm 61.1$	$323 \pm 20.1$					

Table 6.5: Estimates of  $Z \to \nu\nu + jets$  as calulated from the W and Z control regions.

Signal Region						
Cut	$N^{W \to \ell \nu}_{SR,MC}$	$N_{SR,\text{estimate}}^{W \to \ell \nu}$ using $Z \to \ell \ell$	$N_{SR,\text{estimate}}^{W \to \ell \nu}$ using $W \to \ell \nu$			
$\Delta \eta_{jj} > 4.8$	$3702\pm 63.2$	$3797 \pm 267$	$3505 \pm 113$			
$m_{jj} > 1 \text{ TeV}$	$2983 \pm 55.3$	$2947 \pm 234$	$2756 \pm 98.8$			
$\Delta \phi_{jj} < 2.5$	$2149\pm47$	$2219\pm200$	$1983\pm80.3$			
Jet Veto	$975 \pm 31.2$	$1010 \pm 116$	$949 \pm 70.8$			
$\Delta \phi_{jx, E_{\mathrm{T}}^{\mathrm{miss}}} > 1$	$942\pm30.8$	$970 \pm 113$	$922\pm49.9$			
$E_{\rm T}^{\rm miss} > 150 { m ~GeV}$	$225\pm10.6$	$262 \pm 43.5$	$224 \pm 16.4$			

Table 6.6: Estimates of  $W \to \ell \nu + jets$  as calulated from the W and Z control regions.

extrapolation from  $W \to \ell \nu$ +jets or  $Z \to \ell \ell$ +jets to the V+jets being estimated is done by an MC transfer factor. An explicit example of such a cancellation is shown in Section 9.3.1. Specifically, Table 9.6 shows the systematics due to scale variations on V+jets processes in the signal and control regions while Table 9.7 shows the effect of the systematics on the background estimates in the signal region after cancellation.

Using Equation 6.2, the  $Z \rightarrow \nu\nu$ +jets estimates are presented in Table 6.5. The  $W \rightarrow \ell\nu$ +jets estimates are presented in Table 6.6. It is seen that the estimates are consistent between the  $W \rightarrow \ell\nu$ +jets and  $Z \rightarrow \ell\ell$ +jets control region based estimates of the signal region V+jets processes. In principle, the control region based estimates of V+jets in the signal region could disagree with MC due to simulation mismodelings but as long as there is agreement of the control region based estimates with each other, there is confidence in the method.

The estimates of Z and W are then computed in the validation regions using a weighted average of the control region based estimates. These can be compared to data to justify the method, seen in Table 6.7. Good agreement is found between the data driven estimates of Z and W and the observed data in the validation regions.

	ggF+VBF	Strong+Weak	Strong+Weak		Other	Total		
Cut	Signal	$Z \rightarrow \nu \nu + \text{jets}$	$W \rightarrow \ell \nu + \text{jets}$	Multijet	Backgrounds	Backgrounds	2012 Data	Subtracted
			Sm	all- $\Delta \eta_{jj}$ Control Regi	ion			
$p_T^{j1,j2,j3}$	$7229 \pm 89.2$	$290986 \pm 733$	$358260 \pm 956$	$9274322 \pm 726126$	$8063 \pm 19.9$	$9931631 \pm 726127$	7206755	$6549446 \pm 2943$
Opp. Hemispheres	$4441 \pm 60.1$	$125234 \pm 469$	$143609 \pm 540$	$3787720 \pm 464794$	$2401 \pm 10.9$	$4058964 \pm 464795$	3107542	$2836298 \pm 1902$
$\Delta \eta_{ij} < 3.8$	$2442 \pm 51.7$	$112966 \pm 449$	$128536 \pm 542$	$3040348 \pm 401639$	$2227 \pm 10.5$	$3284077 \pm 401640$	2722646	$2478917 \pm 1794$
$m_{ij} > 1 \text{ TeV}$	$33.2 \pm 2.76$	$1728 \pm 48.0$	$2114\pm116$	$144071 \pm 10101$	$33.8 \pm 1.28$	$147947 \pm 10102$	140644	$136768 \pm 395$
$\Delta \phi_{ij} < 2.5$	$18.5 \pm 1.27$	$372 \pm 19.2$	$351 \pm 20.7$	$4754 \pm 1698$	$8.79\pm0.64$	$5486 \pm 1698$	6025	$5293 \pm 82.6$
Jet Veto	$13.3 \pm 1.07$	$135 \pm 11.1$	$80.1 \pm 14.9$	-	$0.05 \pm 0.05$	$215 \pm 18.6$	228	$12.9 \pm 23.9$
$\Delta \phi_{jx,MET} > 1$	$12.7 \pm 1.05$	$133 \pm 11.1$	$81.0\pm13.8$	-	$0.05 \pm 0.05$	$214 \pm 17.7$	219	$4.95 \pm 23.1$
$E_T^{miss} > 150 \text{ GeV}$	$12.1 \pm 1.02$	$114 \pm 9.44$	$71.9 \pm 12.3$	-	$0.05 \pm 0.05$	$186 \pm 15.5$	185	$-0.95 \pm 20.6$
	•			3-jet Control Region				•
$p_T^{j1,j2,j3}$	$3671 \pm 71.6$	$163973 \pm 538$	$225541 \pm 715$	$6894696 \pm 621809$	$7443 \pm 18.9$	$7291653 \pm 621810$	5005243	$4608286 \pm 2410$
Opp. Hemispheres	$1974 \pm 49.1$	$73058 \pm 371$	$94906 \pm 593$	$2748453 \pm 384932$	$2271 \pm 10.5$	$2918688 \pm 384933$	2256994	$2086759 \pm 1657$
$\Delta \eta_{ii} > 4.8$	$271 \pm 14.4$	$1551 \pm 58.7$	$2314 \pm 92.2$	$289317 \pm 158130$	$37.4 \pm 1.33$	$293219 \pm 158130$	69047	$65145 \pm 285$
$m_{ij} > 1 \text{ TeV}$	$228 \pm 12.8$	$1176 \pm 49.7$	$1832 \pm 78.4$	$188684 \pm 130368$	$29.6 \pm 1.18$	$191722 \pm 130368$	55651	$52613 \pm 254$
$\Delta \phi_{ij} < 2.5$	$207 \pm 12.5$	$888 \pm 42.6$	$1301 \pm 58.5$	$114195 \pm 108395$	$22.3 \pm 0.81$	$116406 \pm 108395$	28752	$26541 \pm 184$
$p_T^{j3} > 40 \text{ GeV}$	$75.6 \pm 8.26$	$426 \pm 28.9$	$672 \pm 51.4$	$109400 \pm 108288$	$18.0\pm0.87$	$110516 \pm 108288$	19906	$18790 \pm 153$
$\Delta \phi_{jx,MET} > 1$	$45.0 \pm 5.31$	$205 \pm 18.2$	$273 \pm 21.3$	$108284 \pm 108284$	$5.65 \pm 0.49$	$108768 \pm 108284$	1768	$1284 \pm 50.5$
$E_T^{miss} > 150 \text{ GeV}$	$26.1 \pm 3.39$	$97.2 \pm 9.99$	$78.52 \pm 6.50$	-	$2.16 \pm 0.31$	$178 \pm 11.9$	212	$34.1 \pm 18.8$
	•		3-jet, 1	Small- $\Delta \eta_{jj}$ Control F	legion			
$p_T^{j1,j2,j3}$	$3671 \pm 71.6$	$163974 \pm 539$	$225541 \pm 715$	$6894696 \pm 621809$	$7443 \pm 18.9$	$7291654 \pm 621810$	5005243	$4608285 \pm 2410$
Opp. Hemispheres	$1974 \pm 49.1$	$73059 \pm 371$	$94906 \pm 593$	$2748453 \pm 384932$	$2271 \pm 10.5$	$2918689 \pm 384933$	2256994	$2086758 \pm 1657$
$\Delta \eta_{ij} < 3.8$	$1334 \pm 43.3$	$65853 \pm 342$	$84557 \pm 448$	$2186259 \pm 331456$	$2104 \pm 10.1$	$2338773 \pm 331456$	1973380	$1820866 \pm 1514$
$m_{ii} > 1 \text{ TeV}$	$14.4 \pm 2.44$	$1170 \pm 45.9$	$1705 \pm 251$	$124792 \pm 9429$	$33.0 \pm 1.24$	$127700 \pm 9432$	113369	$110461 \pm 422$
$\Delta \phi_{ij} < 2.5$	$7.35 \pm 0.81$	$247\pm18.3$	$290 \pm 43.2$	$4754 \pm 1698$	$8.79\pm0.64$	$5300 \pm 1699$	5869	$5323 \pm 89.8$
$\Delta \phi_{jx,MET} > 1$	$5.03 \pm 0.66$	$124\pm11.0$	$86.6 \pm 12.0$	$3.12\pm3.12$	$0.59 \pm 0.16$	$214 \pm 16.6$	260	$48.8 \pm 22.9$
$E_T^{miss} > 150 \text{ GeV}$	$4.74 \pm 0.64$	$111 \pm 10.2$	$73.2 \pm 10.1$	-	$0.45 \pm 0.13$	$185 \pm 14.4$	195	$10.4 \pm 20.0$

Table 6.7: Progressive yields of the validation regions described in Section 5.7 in 20.3 fb<sup>-1</sup> of 2012 data, as evaluated using data driven estimates for  $Z \rightarrow \nu\nu$ +jets and  $W \rightarrow \ell\nu$ +jets and Monte Carlo for other processes. Expected signal yields are shown for the case of  $m_H = 125$  GeV.

# CHAPTER 7

# $\gamma$ +jets Control Region

# 7.1 Introduction

An alternate approach to the normalization of the  $Z \rightarrow \nu\nu$ +jets background was investigated using single  $\gamma$  production associated with jets ( $\gamma$ +jets). This can be done because the couplings of Z and  $\gamma$ are similar, as shown in example Feynman diagrams in Figure 7.1. In the regime where  $p_T \gg M_Z$  the kinematics of  $\gamma$ +jets and Z+jets production are very similar. The primary difference is in the quark couplings of Z and  $\gamma$ , specifically that the  $\gamma$  has a stronger coupling to the up quark due to charge. This leads to subtle cross section differences which are accounted for with a MC-based extrapolation factor.

There are several advantages to using a  $\gamma$ +jets control region in addition to the  $W \to \ell \nu$ +jets and  $Z \to \ell \ell$ +jets control regions. The  $\gamma$ +jets control region is not limited by a leptonic branching fraction and as a result the statistical power is higher ( $\approx 2 \times$  the  $W \to \ell \nu$  control region). Additionally, the diagrams are more similar between Z and  $\gamma$  production than W and Z production because of different quark couplings due to the W having charge ( $qq' \to W$  vs  $q\bar{q} \to Z$  or  $\gamma$ ).

One disadvantage of using  $\gamma$ +jets to normalize  $Z \rightarrow \nu\nu$ +jets is that the composition of cross sections due to weak and strong production are different. For example, after the full signal region selection is applied the  $Z \rightarrow \nu\nu$ +jets background is roughly 50% weakly produced and 50% strongly produced. If a similar selection is applied with  $\gamma$ +jets (described in Section 7.2.2), weak production accounts for only  $\approx 5\%$  while strong production accounts for  $\approx 95\%$  of the total cross section. Because of this,  $\gamma$ +jets will be used only to constrain the strongly produced component of the  $Z \rightarrow \nu\nu$ +jets background in the signal region.

Similar to the  $W \to \ell \nu$ +jets and  $Z \to \ell \ell$ +jets based estimations, the normalization is data derived and as a result the associated detector and theory systematics are suppressed.



Figure 7.1: Example Feynman diagrams for Z+2 jets and  $\gamma+$  jets. Since the couplings of Z and  $\gamma$  are similar,  $\gamma+$  jets can be used to model the kinematics of Z+ jets.

Trigger	$p_T$ Threshold	Prescale
EF_g20_loose	$25 \text{GeV} < p_T^{\gamma} < 45 \text{GeV}$	4415.8
$EF_g40_loose$	$45 \text{GeV} < p_T^{\gamma} < 65 \text{GeV}$	348.6
$EF_g60_loose$	$65 \text{GeV} < p_T^{\bar{\gamma}} < 85 \text{GeV}$	80.9
EF_g80_loose	$85 \text{GeV} < p_T^{\bar{\gamma}} < 105 \text{GeV}$	28.5
$EF_g100_loose$	$105 \text{GeV} < \bar{p}_T^{\gamma} < 125 \text{GeV}$	13.0
$EF_g120_loose$	$125 \text{GeV} < p_T^{\tilde{\gamma}}$	1

Table 7.1: Triggers used for  $\gamma$ +jets control region with associated prescales (averaged over runs) and  $p_T$  thresholds.

# 7.2 Event Selections

The  $E_{\rm T}^{\vec{\rm miss}}$  in a  $Z \to \nu\nu$ +jets event is essentially  $\vec{p_T}^Z$  summed with the residual  $E_{\rm T}^{\vec{\rm miss}}$ . In order to model this with a  $\gamma$ +jets event,  $\vec{p_T}^{\vec{\gamma}}$  summed with the residual  $E_{\rm T}^{\vec{\rm miss}}$  in a given event, defined as  $E_{\rm T}^{\vec{\rm miss}} = \vec{p_T} + E_{\rm T}^{\vec{\rm miss}}$ . A series of single  $\gamma$  triggers with various  $p_T$  thresholds are used to construct the events used in the  $\gamma$ +jet control region. Since triggers with thresholds below 125 GeV result in an event rate too high to be saved to disk, the triggers are prescaled. Prescaled triggers only save randomly selected events with a probability equal to the inverse of the prescale value. As a result, events with  $p_T^{\gamma} < 125$  GeV must be weighted with their associated prescales. The prescale for a given trigger is obtained by averaging over the prescale set for individual runs. All of the photon triggers,  $p_T$  thresholds and prescales are listed in Table 7.1.

#### 7.2.1 Selection Differences to Signal Region

To group the selections in a way to make Z and  $\gamma$  more kinematically similar, there are several selection differences from the signal region defined in Section 5.3.

- 1. To reduce dependence on prescaled photons, the  $E_{\rm T}^{\rm miss}$  ' cut (which models the  $E_{\rm T}^{\rm miss}$  cut in the signal region) is moved to the beginning of the selection.
- 2. At high energy it is possible for a jet to emit high  $p_T$  photons which are collinear with the jet. To reduce the effect of collinear photons, the  $\Delta \phi_{jet,\gamma}$  (which is  $\Delta \phi_{jet,E_T^{miss}}$  in the signal region) is moved earlier in the selection

# 7.2.2 $\gamma$ +jets Control Region Selection

The cuts are classified as either "2j" or "*VBF*". This is done to separate modeling issues when extrapolating from  $\gamma$ +jets to  $Z \rightarrow \nu\nu$ +jets. Once the 2j selection is applied the  $\gamma$ +jets and Z+jets samples should be similar kinematically. With these changes, the  $\gamma$ +jet selection is the following:

# **2j Selection**

- 1. Preselection: Good Run List, at least one vertex with  $N_{track} \ge 2$
- 2. Object Selection: LCtopo jets with  $p_T > 30$  GeV,  $|\eta| < 4.5$ ,  $|\mathbf{JVF}| > 0.5$  for jets with  $|\eta| < 2.4$  and  $p_T < 50$  GeV, METRefFinal
- 3. Trigger on EF\_g20\_loose or EF\_g40\_loose or EF\_g60\_loose or EF\_g80\_loose or EF\_g100\_loose or EF\_g120\_loose
- 4. Require exactly 1 isolated, tight PID, unconverted  $\gamma$
- 5.  $E_{\rm T}^{{
  m miss}\,\prime} > 150~{
  m GeV}$
- 6. Lepton veto: Reject events with Medium++  $p_{T,e} > 10$  GeV, StacoCombined  $p_{T,\mu} > 5$  GeV in  $|\eta| < 2.5$  and MS only  $p_{T,\mu} > 5$  GeV in  $2.5 < |\eta| < 2.7$ , Medium BDT  $p_{T,\tau} > 20$  GeV
- 7. b veto: Using 80% working point with  $p_T > 20$  GeV
- 8. Jet  $p_T$ : Leading jet  $p_T > 75$  GeV, Subleading jet  $p_T > 50$  GeV
- 9.  $\Delta \phi_{jet,\gamma} > 1.0$

**VBF** Selection

10.  $\Delta \phi_{jet,X} < 2.5$ 

- 11. Leading and subleading jets must be in opposite hemispheres, i.e.  $\eta_{jet 1} \times \eta_{jet 2} < 0$
- 12.  $\Delta \eta_{j,j} > 4.8$
- 13.  $m_{jj} > 1 \ TeV$
- 14. Third Jet Veto (reject events with third jet with  $p_T > 30 \text{ GeV}$ )

Photons are required to be isolated, unconverted, tight ID and follow all object quality cuts per ATLAS e/gamma prescription [37].

# 7.2.3 Modified Signal Region Selection

In order to be able to make comparisons with the signal region, a modified selection is defined which changes the order of the cuts to be similar to the  $\gamma$ +jets control region. The selection is similarly divided between 2j selection and *VBF* selection.

# **2j Selection**

- 1. Preselection: Good Run List, at least one vertex with  $N_{track} \ge 2$
- 2. Object Selection: LCtopo jets with  $p_T > 30$  GeV,  $|\eta| < 4.5$ ,  $|\mathbf{JVF}| > 0.5$  for jets with  $|\eta| < 2.4$  and  $p_T < 50$  GeV, METRefFinal
- 3. Trigger on EF\_xe80\_tclcw or EF\_xe80\_tclcw\_loose
- 4.  $E_{\rm T}^{\rm miss} > 150 ~{
  m GeV}$
- 5. Lepton veto: Reject events with Medium++  $p_{T,e} > 10$  GeV, StacoCombined  $p_{T,\mu} > 5$  GeV in  $|\eta| < 2.5$  and MS only  $p_{T,\mu} > 5$  GeV in  $2.5 < |\eta| < 2.7$ , Medium BDT  $p_{T,\tau} > 20$  GeV
- 6. b veto: Using 80% working point with  $p_T > 20$  GeV
- 7. Jet  $p_T$ : Leading jet  $p_T > 75$  GeV, Subleading jet  $p_T > 50$  GeV
- 8.  $\Delta \phi_{jet, E_{T}^{miss}} > 1.0$

VBF Selection

- 9.  $\Delta \phi_{jet,X} < 2.5$
- 10. Leading and subleading jets must be in opposite hemispheres, i.e.  $\eta_{jet 1} \times \eta_{jet 2} < 0$

11.  $\Delta \eta_{j,j} > 4.8$ 

12.  $m_{ii} > 1 \ TeV$ 

13. Third Jet Veto (reject events with third jet with  $p_T > 30 \text{ GeV}$ )

# 7.2.4 Modified W+jets Control Region Selection

In order to compare the modeling of various V+jets processes, a modified  $W \rightarrow \ell \nu$ +jets selection is defined as a region to be able to cross-check results. This is similar to the  $\gamma$ +jets control region except that an electron or muon is required with  $p_T^{\ell} > 30$  GeV. Electrons are required to conform to the Tight++ operating point which corresponds to medium operating point (defined in Section 6.2.1) with additional cuts on transition radiation, more strict track-cluster matching, rejection of photon conversions and a requirement of a hit in the b-layer. Muons are required to conform to the Staco-Combined definition which requires an inner detector track to be matched to a track in the Muon Spectrometer.

The same  $E_{\rm T}^{\rm miss}$  trigger used for the signal region (EF\_XE80\_TCLCW\_loose and EF\_XE80\_TCLCW) is used in the  $W \to \ell \nu$ +jets region. It should be noted that since the  $E_{\rm T}^{\rm miss}$  trigger is used instead of a single lepton trigger, the statistical power is lower than the nominal  $W \to \ell \nu$ +jets control region described in Section 6.2. An  $E_{\rm T}^{\rm miss}$  cut of 120 GeV is used so that the trigger is  $\approx 99\%$  efficient. Since the W is used to model the Z background in the signal region, a requirement that  $p_T^W > 150$  GeV is imposed to make it kinematically similar to the signal region. As done in the  $\gamma$ +jets control region and the modified signal region, the selection is divided between 2j selection and VBF selection.

# **2j Selection**

- 1. Preselection: Good Run List, at least one vertex with  $N_{track} \ge 2$
- 2. Object Selection: LCtopo jets with  $p_T > 30$  GeV,  $|\eta| < 4.5$ ,  $|\mathbf{JVF}| > 0.5$  for jets with  $|\eta| < 2.4$  and  $p_T < 50$  GeV, METRefFinal
- 3. Trigger on EF\_xe80\_tclcw or EF\_xe80\_tclcw\_loose
- 4.  $E_{\rm T}^{\rm miss} > 120 ~{\rm GeV}$
- 5. Exactly 1 electron or muon
- 6.  $p_T^{\ell} > 30 \,\, {\rm GeV}$
- 7.  $p_T^W > 150 \text{ GeV}$

- 8. b veto: Using 80% working point with  $p_T > 20$  GeV
- 9. Jet  $p_T$ : Leading jet  $p_T > 75$  GeV, Subleading jet  $p_T > 50$  GeV
- **10.**  $\Delta \phi_{jet, E_{T}^{miss}} > 1.0$

**VBF** Selection

- 11.  $\Delta \phi_{jet,X} < 2.5$
- 12. Leading and subleading jets must be in opposite hemispheres, i.e.  $\eta_{jet 1} \times \eta_{jet 2} < 0$
- 13.  $\Delta \eta_{j,j} > 4.8$
- 14.  $m_{jj} > 1 \ TeV$
- 15. Third Jet Veto (reject events with third jet with  $p_T > 30 \text{ GeV}$ )

# 7.3 Extrapolation of $\gamma$ +jets to $Z \rightarrow \nu\nu$ +jets

As mentioned earlier in Section 7.1, the  $\gamma$ +jets control region is used to constrain only the strongly produced component of the  $Z \rightarrow \nu\nu$ +jets background in the signal region. In order to extrapolate from the  $\gamma$ +jets control region to an estimate of the strongly produced component of the  $Z \rightarrow \nu\nu$ +jets process in the signal region, the efficiency of the *VBF* selection is taken relative to the **2j** selection (*VBF* and **2j** are defined in 7.2.2). However, because the efficiency must be for strongly produced  $\gamma$ +jets only, the weak component is estimated using MC and is used to correct the efficiency. This correction is  $\approx 5\%$ . The efficiency derived in data is:

$$\frac{\gamma_{\rm VBF}^{\rm Data} - \gamma_{\rm VBF}^{\rm Weak \ MC}}{\gamma_{\rm 2i}^{\rm Data} - \gamma_{\rm 2i}^{\rm Weak \ MC}} \tag{7.1}$$

To extrapolate to strongly produced  $Z \rightarrow \nu\nu$ +jets, the ratio of efficiencies in MC is applied, defined as R:

$$R = \frac{Z_{\rm VBF}^{\rm Strong MC} / Z_{2j}^{\rm Strong MC}}{\gamma_{\rm VBF}^{\rm Strong MC} / \gamma_{2i}^{\rm Strong MC}}$$
(7.2)

The result of which is applied to the data yield in the signal region with the 2j selection applied (Data<sub>2j</sub>). In summary, the estimate is defined as:

$$Z_{\nu\nu}^{\text{estimate}} = \frac{\gamma_{\text{VBF}}^{\text{Data}} - \gamma_{\text{VBF}}^{\text{Weak MC}}}{\gamma_{2j}^{\text{Data}} - \gamma_{2j}^{\text{Weak MC}}} R \text{ Data}_{2j}$$
(7.3)

In order to properly compute  $Z_{\nu\nu}^{\text{estimate}}$  with the VBF selection,  $\text{Data}_{2j}$  must be pure Z+jets to cancel with  $Z_{2j}^{\text{Strong MC}}$  in the extrapolation factor R. However, when observing  $\text{Data}_{2j}$ , there is a significant contamination from  $W \to \ell \nu$ +jets where one lepton is not detected which must be subtracted as well as multijets. This is done by measuring the yield in the modified  $W \to \ell \nu$ +jets control region and applying an extrapolation factor to the modified signal region.

$$W_{\ell\nu}^{\text{estimate}} = W_{\text{CR}}^{\text{Data}} \frac{W_{\text{SR}}^{\text{MC}}}{W_{\text{CR}}^{\text{MC}}}$$
(7.4)

Where CR is the modified  $W \rightarrow \ell \nu$ +jets control region and SR is the modified signal region. Accounting for this, the estimate becomes:

$$Z_{\nu\nu}^{\text{estimate}} = \frac{\gamma_{\text{VBF}}^{\text{Data}} - \gamma_{\text{VBF}}^{\text{Weak MC}}}{\gamma_{2j}^{\text{Data}} - \gamma_{2j}^{\text{Weak MC}}} R \quad (\text{Data}_{2j} - W_{\ell\nu}^{\text{estimate}})$$
(7.5)

#### 7.3.1 Truth Level R Calculation

To verify the validity of the method, the extrapolation factor (R) is calculated using truth level events generated with ALPGEN. The CTEQ6L1 PDF is used with a requirement of 2 additional partons and  $E_{\rm T}^{\rm miss} > 100$  GeV for Z+jets and  $p_T^{\gamma} > 100$  GeV for  $\gamma$ +jets. The  $\gamma$ +jets control region selection described in 7.2.2 is then applied to the  $\gamma$ +jets events and the modified signal region selection described in 7.2.3 is applied to Z+jets. Events can be compared after the  $\Delta \phi_{jet,\gamma}$  ( $\Delta \phi_{jet,E_{\rm T}^{\rm miss}}$  in the modified signal selection) since this is when the samples are most kinematically similar. Table 7.2 contains cut efficiencies of *VBF* cuts relative to the 2j selection. Once the  $\Delta \eta_{j,j} > 4.8$  is applied the value of R changes significantly from  $\approx 1$  to  $\approx 0.83$ . This effect is due to quark coupling differences between Z and  $\gamma$ . The  $\Delta \eta_{j,j} > 4.8$  requirement effectively probes high x in the PDF and since photons have a stronger coupling to up quarks, there is a higher relative event rate at high  $\Delta \eta_{j,j}$ . This effect is also present in W+jets because the W also has different couplings than the Z. The 3 V+jets processes considered can be seen in Figure 7.2 using fully simulated SHERPA MC samples with the respective control region selections applied from Section 7.2 (modified signal region selection for Z, modified W+jets control region,  $\gamma$ +jets control region) up to the opposite hemispheres cut. In order to compare relative  $\Delta \eta$  contributions, all distributions are unit normalized.

#### 7.3.2 Truth Level R Calculation Reweighting Test

In order to test that the difference in efficiencies is purely due to quark couplings at high x, a reweighting procedure is performed. This can be done since ALPGEN processes consist of subprocesses



Figure 7.2: Comparison of  $\Delta \eta_{j,j}$  for  $\gamma$ +jets,  $W \rightarrow \ell \nu$ +jets and  $Z \rightarrow \nu \nu$ +jets. All distributions come from fully simulated events generated with SHERPA. The selections from Section 7.2 are applied to the respective samples (modified signal region selection for Z) up to the opposite hemispheres cut. All distributions are unit normalized.

Cut	Z+jets	$\gamma + \text{jets}$	Ratio
$\Delta \phi_{j,j} < 2.5$	$0.9898 \pm 0.0000$	$0.9892 \pm 0.0001$	$1.0006 \pm 0.0001$
Opposite Hemispheres	$0.4152\pm0.0002$	$0.3998 \pm 0.0003$	$1.0384 \pm 0.0009$
$\Delta \eta_{j,j} > 4.8$	$0.0031 \pm 0.0000$	$0.0038 \pm 0.0000$	$0.8323 \pm 0.0097$
$m_{jj}$	$0.0028 \pm 0.0000$	$0.0033 \pm 0.0000$	$0.8363 \pm 0.0104$
Third Jet Veto	$0.0026\pm0.0000$	$0.0031 \pm 0.0000$	$0.8338 \pm 0.0107$

Table 7.2: Efficiencies of VBF cuts relative to the 2j selection in Z+jets and  $\gamma$ +jets as calculated in ALPGEN using truth level events.

Cut	Z+jets	$\gamma + jets$	Ratio
$\Delta \phi_{j,j} < 2.5$	$0.9898 \pm 0.0000$	$0.9900 \pm 0.0001$	$0.9997 \pm 0.0001$
Opposite Hemispheres	$0.4152 \pm 0.0002$	$0.3952 \pm 0.0003$	$1.0506\pm0.0009$
$\Delta \eta_{j,j} > 4.8$	$0.0031 \pm 0.0000$	$0.0031 \pm 0.0000$	$0.9994 \pm 0.0125$
$m_{jj}$	$0.0028 \pm 0.0000$	$0.0027 \pm 0.0000$	$1.0088 \pm 0.0135$
Third Jet Veto	$0.0026\pm0.0000$	$0.0026\pm0.0000$	$1.0169\pm0.0141$

Table 7.3: Efficiencies of *VBF* cuts relative to the **2j** selection in Z+jets and  $\gamma$ +jets as calculated in **ALPGEN** using truth level events after applying the reweighting procedure described in 7.3.2.

which are defined by initial and final state quark flavors. The  $\gamma$ +jets sample is reweighted such that each subprocess contributes to the total cross section in the same way as the Z+jets sample. The results of this test shown in Table 7.3 show that the ratio of efficiencies become  $\approx 1$  after this reweighting procedure is applied. This shows that the difference in the efficiencies comes from the differences in quark couplings.

#### 7.3.3 Theory Systematics on Extrapolation of $\gamma$ +jets to $Z \rightarrow \nu\nu$ +jets

To evaluate theory systematics on the extrapolation of  $\gamma$ +jets to  $Z \rightarrow \nu\nu$ +jets, ALPGEN is used. In order to evaluate the uncertainty due to Parton Distribution Function (PDF), events are reweighted to a given error eigenvector of the CT10 and MSTW2008 PDF sets using PDFTool. The difference of the ratio of cut efficiencies (R defined in Equation 7.9) for a given error eigenvector ( $R_i$ ) to the nominal ratio (R) is combined in quadrature according to the PDF4LHC Working Group Recommendations [38], shown below:

$$\Delta(R) = \sqrt{\sum_{i=1}^{N} (R - R_i)^2}$$
(7.6)

where *i* runs over the set of *N* error eigenvectors. The calculated variation of *R* sets a 68% confidence level uncertainty due to PDF. Additionally, the error on the cut efficiencies ( $\epsilon$ ) is also

Cut	Z+jets CTEQ6L1	Z+jets CT10	$Z{+}\mathrm{jets}$ MSTW2008
$\Delta \phi_{j,j} < 2.5$	$0.9898 \pm 0.0000$	$0.9899 \pm 0.0001$	$0.9899 \pm 0.0000$
Opposite Hemispheres	$0.4152 \pm 0.0002$	$0.4209 \pm 0.0018$	$0.4176\pm0.0007$
$\Delta \eta_{j,j} > 4.8$	$0.0031 \pm 0.0000$	$0.0031 \pm 0.0002$	$0.0031\pm0.0000$
$m_{jj}$	$0.0028 \pm 0.0000$	$0.0027 \pm 0.0002$	$0.0028 \pm 0.0000$
Third Jet Veto	$0.0026 \pm 0.0000$	$0.0025 \pm 0.0002$	$0.0026\pm0.0000$
Cut	$\gamma + { m jets}$ CTEQ6L1	$\gamma +  ext{jets}$ CT10	$\gamma{+}\mathrm{jets}$ MSTW2008
$\Delta \phi_{j,j} < 2.5$	$0.9892 \pm 0.0001$	$0.9892 \pm 0.0001$	$0.9892 \pm 0.0000$
Opposite Hemispheres	$0.3998 \pm 0.0003$	$0.4093 \pm 0.0020$	$0.4035\pm0.0008$
$\Delta \eta_{j,j} > 4.8$	$0.0038 \pm 0.0000$	$0.0037 \pm 0.0002$	$0.0038 \pm 0.0000$
$m_{jj}$	$0.0033 \pm 0.0000$	$0.0032 \pm 0.0002$	$0.0033 \pm 0.0000$
Third Jet Veto	$0.0031 \pm 0.0000$	$0.0030\pm0.0002$	$0.0031\pm0.0000$
Cut	Ratio CTEQ6L1	Ratio CT10	Ratio MSTW2008
$\Delta \phi_{j,j} < 2.5$	$1.0006 \pm 0.0001$	$1.0007 \pm 0.0000$	$1.0007 \pm 0.0000$
Opposite Hemispheres	$1.0384 \pm 0.0009$	$1.0283 \pm 0.0020$	$1.0350\pm0.0010$
$\Delta \eta_{j,j} > 4.8$	$0.8323 \pm 0.0097$	$0.8329 \pm 0.0080$	$0.8275 \pm 0.0029$
$m_{jj}$	$0.8363 \pm 0.0104$	$0.8378 \pm 0.0087$	$0.8324\pm0.0033$
Third Jet Veto	$0.8338 \pm 0.0107$	$0.8357 \pm 0.0085$	$0.8309 \pm 0.0032$

Table 7.4: Efficiencies of VBF cuts relative to the **2j** selection in Z+jets and  $\gamma$ +jets as calculated using the following PDF sets: CTEQ6L1, CT10 and MSTW2008. Uncertainties on CTEQ6L1 are statistical while uncertainties on CT10 and MSTW2008 are derived from eigenvector variations.

calculated:

$$\Delta(\epsilon) = \sqrt{\sum_{i=1}^{N} (\epsilon - \epsilon_i)^2}$$
(7.7)

The results are shown in Table 7.4. The ratio of VBF cut efficiencies (R) are seen to be stable with respect to PDF set with a variation of < 1%.

Since cross sections can only be calculated at finite orders, predictions are parametrized in terms of the factorization scale  $(\mu_F)$  which is unphysical. In NLO generators predictions are parameterized in terms of factorization and renormalization  $(\mu_R)$  scales. To estimate the perturbative uncertainty, the scales are varied by regenerating events with different  $Q^2$  values around the central scale. Typically, the central scale is varied up and down by a factor of 2 to set the uncertainty. The *VBF* cut efficiencies with various scales are shown in Table 7.5. The ratio of *VBF* efficiencies are seen to be stable with respect to scale with a variation of  $\approx 3\%$ .

Cut	$Z$ +jets $Q^2 = 1.0$	$Z$ +jets $Q^2 = 2.0$	$Z$ +jets $Q^2 = 0.5$
$\Delta \phi_{j,j} < 2.5$	$0.9898 \pm 0.0000$	$0.9898 \pm 0.0000$	$0.9900 \pm 0.0000$
Opposite Hemispheres	$0.4152 \pm 0.0002$	$0.4124 \pm 0.0002$	$0.4200 \pm 0.0002$
$\Delta \eta_{j,j} > 4.8$	$0.0031 \pm 0.0000$	$0.0027 \pm 0.0000$	$0.0040\pm0.0000$
$m_{jj}$	$0.0028 \pm 0.0000$	$0.0024 \pm 0.0000$	$0.0035 \pm 0.0000$
Third Jet Veto	$0.0026 \pm 0.0000$	$0.0023 \pm 0.0000$	$0.0033 \pm 0.0000$
Cut	$\gamma$ +jets $Q^2 = 1.0$	$\gamma$ +jets $Q^2 = 2.0$	$\gamma$ +jets $Q^2 = 0.5$
$\Delta \phi_{j,j} < 2.5$	$0.9892 \pm 0.0001$	$0.9891 \pm 0.0001$	$0.9896 \pm 0.0001$
Opposite Hemispheres	$0.3998 \pm 0.0003$	$0.3963 \pm 0.0003$	$0.4058 \pm 0.0003$
$\Delta \eta_{j,j} > 4.8$	$0.0038 \pm 0.0000$	$0.0032 \pm 0.0000$	$0.0048 \pm 0.0000$
$m_{jj}$	$0.0033 \pm 0.0000$	$0.0028 \pm 0.0000$	$0.0042 \pm 0.0000$
Third Jet Veto	$0.0031 \pm 0.0000$	$0.0026\pm0.0000$	$0.0039 \pm 0.0000$
Cut	Ratio $Q^2 = 1.0$	Ratio $Q^2 = 2.0$	Ratio $Q^2 = 0.5$
$\Delta \phi_{j,j} < 2.5$	$1.0006 \pm 0.0001$	$1.0007 \pm 0.0001$	$1.0005 \pm 0.0001$
Opposite Hemispheres	$1.0384 \pm 0.0009$	$1.0409\pm0.0009$	$1.0352 \pm 0.0010$
$\Delta \eta_{j,j} > 4.8$	$0.8323 \pm 0.0097$	$0.8633 \pm 0.0110$	$0.8303 \pm 0.0101$
$m_{ii}$	$0.8363 \pm 0.0104$	$0.8590 \pm 0.0117$	$0.8388 \pm 0.0109$
Third Jet Veto	$0.8338 \pm 0.0107$	$0.8605 \pm 0.0121$	$0.8379 \pm 0.0112$

Table 7.5: Efficiencies of VBF cuts relative to the **2j** selection in Z+jets and  $\gamma$ +jets as calculated with various scales,  $Q^2 = 1.0, 2.0, 0.5$ . Uncertainties are statistical.

Strongly Produced $\gamma$ +jets	
$mc12\_8 TeV.177574. Sherpa\_CT10\_SinglePhotonMassiveCBPt100\_140\_CVetoBVeto.merge.NTUP\_SMWZ.e1913\_s1581\_s1586\_r3658\_r3549\_p1328/s1584\_s1586\_r3658\_r3549\_p1328/s1584\_s1586\_r3658\_r3549\_p1328/s1584\_s1586\_r3658\_r3549\_p1328/s1584\_s1586\_r3658\_r3549\_p1328/s1584\_s1586\_r3658\_r3549\_p1328/s1584\_s1584\_s1586\_r3658\_r3549\_p1328/s1584\_s1564\_s1564\_s1564\_s1564\_s1564\_s1564\_s1564\_s1564\_s104\_s1564\_s1564\_s1564\_s1564\_s1564\_s1564\_$	Ī
mc12_8TeV.177575.Sherpa_CT10_SinglePhotonMassiveCBPt140_280_CVetoBVeto.merge.NTUP_SMWZ.e1913_s1581_s1586_r3658_r3549_p1328/	
$mc12\_8TeV.177576.Sherpa\_CT10\_SinglePhotonMassiveCBPt280\_500\_CVetoBVeto.merge.NTUP\_SMWZ.e1913\_s1581\_s1586\_r3658\_r3549\_p1328/s1086\_r3658\_r3549\_p1328/s1086\_r3658\_r3549\_p1328/s1086\_r3658\_r3549\_p1328/s1086\_r3658\_r3549\_p1328/s1086\_r3658\_r365\_r365\_r365\_r365\_r36\_r365\_r365\_r365\_$	
mc12_8TeV.177577.Sherpa_CT10_SinglePhotonMassiveCBPt500_CVetoBVeto.merge.NTUP_SMWZ.e1913_s1581_s1586_r3658_r3549_p1328/	
mc12_8TeV.177578.Sherpa_CT10_SinglePhotonMassiveCBPt100_140_CFilterBVeto.merge.NTUP_SMWZ.e1913_s1581_s1586_r3658_r3549_p1328/	
mc12_8TeV.177579.Sherpa_CT10.SinglePhotonMassiveCBPt140_280_CFilterBVeto.merge.NTUP_SMWZ.e1913_s1581_s1586_r3658_r3549_p1328/	
mc12_8TeV.177580.Sherpa_CT10.SinglePhotonMassiveCBPt280.500_CFilterBVeto.merge.NTUP_SMWZ.e1913_s1581_s1586_r3658_r3549_p1328/	
mc12_8TeV.177581.Sherpa_CT10_SinglePhotonMassiveCBPt500_CFilterBVeto.merge.NTUP_SMWZ.e1913_s1581_s1586_r3658_r3549_p1328/	
mc12_8TeV.177582.Sherpa_CT10_SinglePhotonMassiveCBPt100_140_BFilter.merge.NTUP_SMWZ.e1913_s1581_s1586_r3658_r3549_p1328/	
mc12_8TeV.177583.Sherpa_CT10_SinglePhotonMassiveCBPt140_280_BFilter.merge.NTUP_SMWZ.e1913_s1581_s1586_r3658_r3549_p1328/	
mc12_8TeV.177584.Sherpa_CT10_SinglePhotonMassiveCBPt280_500_BFilter.merge.NTUP_SMWZ.e1913_s1581_s1586_r3658_r3549_p1328/	
$mc12\_8TeV.177585.Sherpa\_CT10\_SinglePhotonMassiveCBPt500\_BFilter.merge\_NTUP\_SMWZ.e1913\_s1581\_s1586\_r3658\_r3549\_p1328/constructions and the second structure in the second str$	
Weakly Produced $\gamma$ +jets	Ĩ
mal 2 2ToV 181677 Sharpa CT10 commo 2 lots EW1 lot OCD15 CoVM350 min n tohonnals marga NTUP SMWZ 02464 s1773 s1776 r4485 r4540 n1228	Γ

Table 7.6: MC samples used in the  $\gamma$ +jets control region.

# 7.4 Processes in the $\gamma$ +jets Control Region

# 7.4.1 $\gamma$ +jets Monte Carlo

The  $\gamma$ +jets processes are modeled by the MC samples shown in Table 7.6. The table includes strong and weakly produced  $\gamma$ +jets.

# 7.4.2 Background from Multijet

The dominant background in the  $\gamma$ +jets control region is multijet production wherein a jet is misidentified as a photon. A data driven technique is used which has been implemented on ATLAS previously and documented in [39] and [40]. Since photon identification (PID) is uncorrelated with isolation, the shape of the isolation distribution for photons failing PID with high isolation is the same as the shape of photons passing PID with high isolation. Because of this, the isolation distribution for photons failing PID with high isolation is normalized to have the same yield as photons passing PID at high isolation. Then the normalized isolation distribution can be integrated in the region used for identified photons (low isolation) which results in an estimate of jets misidentified as photons. This works because photons with values of high isolation passing PID is enriched in jets misidentified as photons. The procedure is:

- 1. Three shape templates for isolation (topoEtcone40) are derived from photons failing PID. Each of the three templates are defined by the number of PID variables a given photon fails. The three PID variables considered are: fside, ERatio and  $\Delta E$ . The PID variables use strip information from the electromagnetic calorimeter. The fside variable represents the lateral spread of an electromagnetic shower in  $\eta$ . The ERatio variable is the ratio of the sizes of second and first relative maxima of strip energies. The  $\Delta E$  variable is the difference of the strip with the greatest amount of energy to the strip with the least amount of energy between the strips with the greatest and second greatest amounts of energy.
  - The three photon PID variables considered are uncorrelated with isolation since they are extracted from a region smaller than the 5x7 core [41].
- 2. The three regions are then normalized to the distribution of photons passing all PID cuts with isolation  $E_T > 30$  GeV. Figure 7.3 show the isolation distributions after the normalization for photons passing PID and the three regions of photons failing combinations of PID (despite appearance due to different slopes, the distributions are in fact normalized to be equal at high isolation).
- 3. Integrate the three fail PID regions in the isolated range used for signal photons ( $E_T < 4 \text{ GeV}$ ). This results in three estimates of the multijet background.

This procedure is done as a function of the selection, shown in Table 7.7. The three fail PID regions give consistent results, showing that the yield of jets misidentified as photons in the  $\gamma$ +jets control region is very small compared to the number of identified photons.



Figure 7.3: Isolation distributions for photons passing PID and for photons in the three regions of failing PID. The distributions are normalized to the isolation  $E_T > 30$  GeV region of the photons passing PID.

r			
Cut	Fails 1	Fails 2	Fails 3
Isolated Tight $\gamma$	$103465512 \pm 280729$	$64842360 \pm 390729$	$64628634 \pm 455548$
$p_T^{\gamma} > 150 \text{GeV}$	$15234.55 \pm 52.52$	$11000.03 \pm 65.09$	$8986.31 \pm 79.71$
e veto	$14309.79 \pm 51.34$	$10562.57 \pm 64.89$	$8512.55 \pm 79.41$
$\mu$ veto	$13668.58 \pm 50.29$	$10100.68 \pm 63.58$	$8127.40 \pm 77.84$
au veto	$12221.55 \pm 47.26$	$8869.39 \pm 59.18$	$7409.88 \pm 73.93$
b veto	$10347.04 \pm 43.25$	$7543.17 \pm 54.17$	$6345.52 \pm 67.62$
jet $p_T > 75 \text{GeV}$	$9147.23 \pm 40.31$	$6492.92\pm50.15$	$5494.10 \pm 63.21$
jet $p_T > 50 \text{GeV}$	$2416.56 \pm 21.15$	$1698.57 \pm 24.84$	$1429.03 \pm 31.07$
$\Delta \phi_{i,E_{\mathrm{T}}^{\mathrm{miss}}} > 1.0$	$1890.56 \pm 18.75$	$1337.21 \pm 21.99$	$1147.79 \pm 27.66$
$\Delta \phi_{j,j} < 2.5$	$1850.82 \pm 18.56$	$1311.77 \pm 21.76$	$1129.15 \pm 27.40$
Opp Hemispheres	$674.56 \pm 10.84$	$469.47 \pm 13.21$	$406.80 \pm 16.70$
$\Delta \eta_{j,j} > 4.8$	$12.64 \pm 1.58$	$7.66 \pm 1.77$	$5.17 \pm 1.78$
$m_{jj} > 1 \text{TeV}$	$10.61 \pm 1.45$	$6.86\pm1.69$	$5.08 \pm 1.78$
Third Jet Veto	$7.01 \pm 1.29$	$3.98 \pm 1.48$	$2.75 \pm 1.58$

Table 7.7: Multijet estimate as a function of the selection for the three fail PID regions.

Cut	Weak $\gamma$ +jets MC	Strong $\gamma$ +jets MC	Multijet Estimate	Total Expected	Data	Data/MC
$p_T^{\gamma} > 150 \text{GeV}$	$1067.60 \pm 17.71$	$1656947.68 \pm 2581.89$	$15234.10 \pm 6311.68$	$1673249.38 \pm 6859.13$	$1767280 \pm 1329$	$1.06\pm0.00$
e veto	$1054.01 \pm 17.59$	$1637831.53 \pm 2567.97$	$14226.82 \pm 5768.12$	$1653112.35 \pm 6356.42$	$1737406\pm1318$	$1.05\pm0.00$
$\mu$ veto	$994.49 \pm 17.08$	$1564708.52 \pm 2515.58$	$13590.28 \pm 5512.96$	$1579293.30\pm6102.26$	$1664061\pm1290$	$1.05\pm0.00$
$\tau$ veto	$960.88 \pm 16.79$	$1515250.36 \pm 2475.63$	$12148.46 \pm 4786.27$	$1528359.70\pm5434.85$	$1628910\pm1276$	$1.07\pm0.00$
b veto	$797.41 \pm 15.12$	$1267791.21 \pm 2258.00$	$10295.18 \pm 3984.12$	$1278883.80 \pm 4624.72$	$1361000\pm1167$	$1.06\pm0.00$
jet $p_T > 75 \text{GeV}$	$710.75 \pm 14.27$	$784585.89 \pm 1730.50$	$9104.48 \pm 3642.65$	$794401.12\pm4063.01$	$824980 \pm 908$	$1.04\pm0.01$
jet $p_T > 50 \text{GeV}$	$571.33 \pm 12.80$	$363509.94 \pm 1130.45$	$2401.62 \pm 976.23$	$366482.90 \pm 1528.19$	$391080 \pm 625$	$1.07\pm0.00$
$\Delta \phi_{i,E_{T}^{\text{miss}}} > 1.0$	$401.80 \pm 10.76$	$260411.13 \pm 951.03$	$1876.80\pm734.38$	$262689.73 \pm 1231.95$	$283587 \pm 533$	$1.08\pm0.01$
$\Delta \phi_{j,j} < 2.5$	$376.22 \pm 10.41$	$249708.11 \pm 925.05$	$1837.17 \pm 713.34$	$251921.50\pm1197.70$	$270814 \pm 520$	$1.07\pm0.01$
Opp Hemispheres	$345.17 \pm 9.96$	$97487.77 \pm 583.59$	$668.25 \pm 265.48$	$98501.19 \pm 662.53$	$107825 \pm 328$	$1.09\pm0.01$
$\Delta \eta_{j,j} > 4.8$	$63.63 \pm 4.28$	$2196.10\pm104.77$	$12.64 \pm 7.47$	$2272.37 \pm 109.30$	$2008 \pm 45$	$0.88\pm0.05$
$m_{ii} > 1 \text{TeV}$	$61.32 \pm 4.20$	$1729.07 \pm 88.30$	$10.61 \pm 5.53$	$1801.01 \pm 92.09$	$1533 \pm 39$	$0.85\pm0.05$
CĴV	$59.86 \pm 4.15$	$1045.72\pm72.18$	$7.01 \pm 4.26$	$1112.58 \pm 75.30$	$932\pm31$	$0.84\pm0.06$

Table 7.8: Yields of MC strongly produced  $\gamma$ +jets, MC weakly produced  $\gamma$ +jets, multijet estimate as described in 7.4.2 and data as a function of the  $\gamma$ +jets control region selection.

# 7.5 $\gamma$ +jets Control Region Yields

Presented in Table 7.8 are the MC yields for strong and weakly produced  $\gamma$ +jets, the multijet estimate and the yield in data. The strongly produced  $\gamma$ +jets MC is normalized to the NLO cross section calculated by JETPHOX, resulting in an effective k-factor of 1.04. The weakly produced  $\gamma$ +jets MC is normalized to the cross section given by SHERPA. The multijet estimate is calculated in a conservative way by taking the maximum of the 3 estimates as the central value and using the maximum difference of the 3 as the error.

It is seen that the modeling changes significantly after the  $\Delta \eta_{j,j} > 4.8$  requirement, visualized in Figure 7.4. This effect has been seen before in other V+jets studies from ATLAS including W+jets [42] and Z+jets [43] cross section measurements. This does not pose a problem to the method proposed in 7.3 unless the various V+jets processes are modeled inconsistently. In order to cross check the



(a) Data and MC distributions of  $\Delta \eta_{j,j}$  with the  $\gamma$ +jets (b) Data/MC ratio as a function of  $\Delta \eta_{j,j}$  with the selection applied through the opposite hemispheres cut.  $\gamma$ +jets selection applied through the opposite hemispheres cut.

Figure 7.4:  $\Delta \eta_{j,j}$  with the  $\gamma$ +jets selection applied through the opposite hemispheres cut.

modeling, the modified  $W \rightarrow \ell \nu + jets$  control region described in 7.2.4 is used.

#### 7.5.1 $W \rightarrow \ell \nu + jets$ Cross-Check

Using the modified  $W \to \ell \nu + jets$  described in 7.2.4, the modeling of the high  $\Delta \eta_{j,j}$  region can be used as a cross-check of the  $\gamma$ +jets region which seems to have Data/MC ratios different from 1.

#### 7.5.1.1Signal and Background

The signal is modeled by MC samples shown in Table 7.9. The table includes strong and weakly produced  $W \to \ell \nu + jets$  processes.

Strongly Produced $W \to \ell \nu + \text{jets}$
mc12_8TeV.167761_Sherpa_CT10_WenuMassiveCBPt70_140_BFilter.merge.NTUP_SMWZ.e1620_a159_a171_r3549_p1328/
$mc12.8 TeV.167762. Sherpa\_CT10\_WenuMassiveCBPt70\_140\_CJetFilterBVeto.merge.NTUP\_SMWZ.e1620\_a159\_a171\_r3549\_p1328/cmerse.NTUP\_SMWZ.e1620\_a159\_a171\_r3549\_p1328/cmerse.NTUP\_SMWZ.e1620\_a159\_a171\_r3549\_p1328/cmerse.NTUP\_SMWZ.e1620\_a159\_a171\_r3549\_p1328/cmerse.NTUP\_SMWZ.e1620\_a159\_a171\_r3549\_p1328/cmerse.NTUP\_SMWZ.e1620\_a159\_a171\_r3549\_p1328/cmerse.NTUP\_SMWZ.e1620\_a159\_a171\_r3549\_p1328/cmerse.NTUP\_SMWZ.e1620\_a159\_a171\_r3549\_p1328/cmerse.NTUP\_SMWZ.e1620\_a159\_a171\_r3549\_p1328/cmerse.NTUP\_SMWZ.e1620\_a159\_a171\_r3549\_p1328/cmerse.NTUP\_SMWZ.e1620\_a159\_a171\_r3549\_p1328/cmerse.NTUP\_SMWZ.e1620\_a159\_a171\_r3549\_p1328/cmerse.NTUP\_SMWZ.e1620\_a159\_a171\_r3549\_p1328/cmerse.NTUP\_SMWZ.e1620\_a159\_a171\_r3549\_p1328/cmerse.NTUP\_sMWZ.e1620\_a159\_a171\_r3549\_p1328/cmerse.NTUP\_sMWZ.e1620\_a159\_a150\_ra1328/cmerse.NTUP\_sMWZ.e1620\_a159\_a150\_ra1328/cmerse.NTUP\_sMWZ.e1620\_a150\_a150\_ra1328/cmerse.NTUP\_sMWZ.e160\_ra1328/cmerse.NTUP\_sMWZ.e160\_ra1328/cmerse.NTUP\_sMWZ.e160\_ra1328/cmerse.NTUP\_sMWZ.e160\_ra1328/cmerse.NTUP\_sMWZ.e160\_ra1328/cmerse.NTUP\_sMWZ.e160\_ra1328/cmerse.NTUP\_sMWZ.e160\_ra1328/cmerse.NTUP\_sMWZ.e160\_ra1328/cmerse.NTUP\_sMWZ.e160\_ra1328/cmerse.NTUP\_sMWZ.e160\_ra1328/cmerse.NTUP\_sMWZ.e160\_ra1328/cmerse.NTUP\_sMWZ.e160\_ra1328/cmerse.NTUP\_sMWZ.e160\_ra1328/cmerse.NTUP\_sMWZ.e160\_ra1328/cmerse.NTUP\_sMWZ.e160\_ra1328/cmerse.NTUP\_sMWZ.e160\_ra1328/cmerse.NTUP\_sMWZ.e160\_ra1328/cmerse.NTUP\_sMWZ.e160\_ra1328/cmerse.NTUP\_sMUTSe.NTUP\_sMUTSe.NTUP\_sMUTSe.NTUP\_sMUTSe.NTUP\_sMUTSe.NTUP\_sMUTSe.NTUP\_sMUTSe.NTUP\_sMUTSe.NTUP\_sMUTSe.NTUP\_sMUTSe.NTUPsMTSe.NTUPsMTSe.NTUP\_sMUTSe.NTUPsMTSe.NTUPsMTSe.NTUPSE.NTUPSE.NTUPSMTSe.NTUPSE.N$
$mc12\_8 TeV.167763. Sherpa\_CT10\_WenuMassiveCBPt70\_140\_CJetVetoBVeto\_merge.NTUP\_SMWZ.e1620\_a159\_a171\_r3549\_p1328/2000. Sherpa\_CT10\_WenuMassiveCBPt70\_140\_CJetVetoBVeto\_merge.NTUP\_SMWZ.e1620\_a159\_a171\_r3549\_p1328/2000. Sherpa\_CT10\_WenuMassiveCBPt70\_140\_CJetVetoBVeto\_merge.NTUP\_SMWZ.e1620\_a159\_a171\_r3549\_p1328/2000. Sherpa\_CT10\_WenuMassiveCBPt70\_140\_CJetVetoBVeto\_merge.NTUP\_SMWZ.e1620\_a159\_a171\_r3549\_p1328/2000. Sherpa\_CT10\_WenuMassiveCBPt70\_140\_CJetVetoBVeto\_merge.NTUP\_SMWZ.e1620\_a159\_a171\_r3549\_p1328/2000. Sherpa\_CT10\_WenuMassiveCBPt70\_140\_CJetVetoBVeto\_merge.NTUP\_SMWZ.e1620\_a159\_a171\_r3549\_p1328/2000. Sherpa\_CT10\_WenuMassiveCBPt70\_140\_CJetVetoBVeto\_merge.NTUP\_SMWZ.e1620\_a159\_a171\_r3549\_p1328/2000. Sherpa\_A000. $
$mc12\_8 TeV.167770. Sherpa\_CT10\_WenuMassiveCBPt140\_280\_BFilter.merge.NTUP\_SMWZ.e1620\_a159\_a171\_r3549\_p1328/2001.0000000000000000000000000000000000$
$mc12.8 TeV.167771. Sherpa\_CT10\_WenuMassiveCBPt140\_280\_CJetFilterBVeto.merge.NTUP\_SMWZ.e1620\_a159\_a171\_r3549\_p1328/cms^2 + 100000000000000000000000000000000000$
$mc12.8 TeV.167772. Sherpa\_CT10\_WenuMassiveCBPt140\_280\_CJetVetoBVeto.merge.NTUP\_SMWZ.e1620\_a159\_a171\_r3549\_p1328/cMVZ.e1620\_a159\_a171\_r3549\_p1328/cMVZ.e1620\_a159\_a171\_r3549\_p1328/cMVZ.e1620\_a159\_a171\_r3549\_p1328/cMVZ.e1620\_a159\_a171\_r3549\_p1328/cMVZ.e1620\_a159\_a171\_r3549\_p1328/cMVZ.e1620\_a159\_a171\_r3549\_p1328/cMVZ.e1620\_a159\_a171\_r3549\_p1328/cMVZ.e1620\_a159\_a171\_r3549\_p1328/cMVZ.e1620\_a159\_a171\_r3549\_p1328/cMVZ.e1620\_a159\_a150\_r350\_r350\_r350\_r350\_r350\_r350\_r350\_r3$
$mc12.8 TeV.167779. Sherpa\_CT10\_WenuMassiveCBPt280\_500\_BFilter.merge.NTUP\_SMWZ.e1714\_s1581\_s1586\_r3658\_r3549\_p1328/s1286\_r3658\_r3549\_p1328/s1286\_r3658\_r3549\_p1328/s1286\_r3658\_r3549\_p1328/s1286\_r3658\_r3549\_p1328/s1286\_r3658\_r3540\_r3658\_r3560\_r3658\_r3540\_r3658\_r3560\_r3658\_r3560\_r3658\_r3560\_r3658\_r3560\_r3658\_r3560\_r3658\_r3560\_r3658\_r3560\_r3658\_r3560\_r3658\_r3560\_r3658\_r3560\_r3658\_r3560\_r3658\_r3560\_r3658\_r3560\_r3650\_r3650\_r3650\_r3650\_r3650\_r3650\_r3650\_r3650\_r3650\_r3650\_r3650\_r3650\_r3650\_r3650\_r3650\_r3650\_r3650\_r3650\_r360\_r360\_r360\_r360\_r360\_r360\_r360\_r36$
$mc12\_8TeV.167780.Sherpa\_CT10\_WenuMassiveCBPt280\_500\_CJetFilterBVeto.merge.NTUP\_SMWZ.e1714\_s1581\_s1586\_r3658\_r3549\_p1328/cms12\_s1586\_r3658\_r3658\_r3649\_p1328/cms12\_s1586\_r3658\_r3649\_p1328/cms12\_s1586\_r3658\_r3649\_p1328/cms12\_s1586\_r3658\_r3649\_p1328/cms12\_s1586\_r3658\_r3649\_p1328/cms12\_s1586\_r3658\_r3649\_p1328/cms13\_s1586\_r3658\_r3649\_p1328/cms13\_s1586\_r3658\_r3649\_p1328/cms13\_s1586\_r3658\_r3649\_p1328/cms13\_s1586\_r3658\_r3649\_r3640\_r3658\_r3640\_r365\_r365\_r365\_r365\_r365\_r365\_r365\_r365$
$mc12.8 TeV.167781. Sherpa\_CT10\_WenuMassiveCBPt280\_500\_CJetVetoBVeto.merge.NTUP\_SMWZ.e1714\_s1581\_s1586\_r3558\_r3549\_p1328/content and the second seco$
$mc12\_8TeV.167788.Sherpa\_CT10\_WenuMassiveCBPt500\_BFilter\_merge.NTUP\_SMWZ.e1620\_s1499\_s1504\_r3658\_r3549\_p1328/s1204\_r3658\_r3549\_p1328/s1204\_r3658\_r3549\_p1328/s1204\_r3658\_r3549\_p1328/s1204\_r3658\_r3549\_p1328/s1204\_r3658\_r3549\_p1328/s1204\_r3658\_r3549\_p1328/s1204\_r3658\_r3549\_p1328/s1204\_r3658\_r3549\_p1328/s1204\_r3658\_r3549\_p1328/s1204\_r3658\_r3549\_p1328/s1204\_r3658\_r3549\_r364s\_r364s\_r3658\_r364s\_r3658\_r364s\_r3658\_r364s\_r3658\_r364s\_r3658\_r364s\_r3658\_r364s\_r3658\_r364s\_r3658\_r364s\_r3658\_r364s\_r3658\_r364s\_r3658\_r364s\_r3658\_r364s\_r3658\_r364s\_r3658\_r364s\_r3658\_r3656\_r3658\_r3656\_r365\_r365$
$mc12\_8 TeV.167789. Sherpa\_CT10\_WenuMassiveCBPt500\_CJetFilterBVeto.merge.NTUP\_SMWZ.e1620\_s1499\_s1504\_r3658\_r3549\_p1328/contentseterminerge.NTUP\_SMWZ.e1620\_s1499\_s1504\_r3658\_r3549\_p1328/contentseterminerge.NTUP\_SMWZ.e1620\_s1499\_s1504\_r3658\_r3549\_p1328/contentseterminerge.NTUP\_SMWZ.e1620\_s1499\_s1504\_r3658\_r3549\_p1328/contentseterminerge.NTUP\_SMWZ.e1620\_s1499\_s1504\_r3658\_r3549\_p1328/contentseterminerge.NTUP\_SMWZ.e1620\_s1499\_s1504\_r3658\_r3549\_p1328/contentseterminerge.NTUP\_SMWZ.e1620\_s1499\_s1504\_r3658\_r3549\_p1328/contentseterminerge.NTUP\_SMWZ.e1620\_s1499\_s1504\_r3658\_r3549\_p1328/contentseterminerge.NTUP\_SMWZ.e1620\_s1499\_s1504\_r3658\_r3549\_p1328/contentseterminerge.NTUP\_SMWZ.e1620\_s1499\_s1504\_r3658\_r3549\_p1328/contentseterminerge.NTUP\_SMWZ.e1620\_s1499\_s1504\_r3658\_r3549\_p1328/contentseterminerge.NTUP\_s04\_r3658\_r3540\_s00$
$mc12\_8 TeV.167790.Sherpa\_CT10\_WenuMassiveCBPt500\_CJetVetoBVeto.merge.NTUP\_SMWZ.e1620\_s1499\_s1504\_r3658\_r3549\_p1328/contentsetMUP\_SMWZ.e1620\_s1499\_s1504\_r3658\_r3549\_p1328/contentsetMUP\_sMWZ.e1620\_s1499\_s1504\_r3658\_r3549\_p1328/contentsetMUP\_sMWZ.e1620\_s1499\_s1504\_r3658\_r3549\_p1328/contentsetMUP\_sMWZ.e1620\_s1499\_s1504\_r3658\_r3549\_p1328/contentsetMUP\_sMWZ.e1620\_s1499\_s1504\_r3658\_r3549\_p1328/contentsetMUP\_sMWZ.e1620\_s1490\_s1504\_r3658\_r3549\_s1304\_r3658\_r3549\_s1304\_r3658\_r3549\_s1304\_r3658\_r3540\_s1349\_s1504\_r3658\_r3540\_s1340\_r3658\_r3540\_s1340\_s140\_s1440\_s1340\_s140\_s1340\_s140\_s140\_s140\_s140\_s140\_s140\_s140\_s1$
$mc12\_8TeV.167764. Sherpa\_CT10\_WmunuMassiveCBPt70\_140\_BFilter.merge.NTUP\_SMWZ.e1714\_a159\_a171\_r3549\_p1328/2001000000000000000000000000000000000$
$mc12\_8 TeV.167765.Sherpa\_CT10\_WmunuMassiveCBPt70\_140\_CJetFilterBVeto\_merge.NTUP\_SMWZ.e1714\_a159\_a171\_r3549\_p1328/c1714\_r3549\_r171\_r13549\_r171\_r3549\_r171\_r171\_r3549\_r171\_r355\_r171\_r3540\_r171\_r3540\_r171\_r3540\_r171\_r3540\_r171\_r3540\_r171\_r3540\_r171\_r3540\_r171\_r3540\_r171\_r3540\_r171\_r3540\_r171\_r3540\_r171\_r3540\_r171\_r3540\_r171\_r3540\_r171\_r3540\_r171\_r3540\_r171\_r3540\_r171\_r3540\_r171\_r3540\_r171\_r355\_r171\_r35\_r171\_r35\_r171\_r35\_r171\_r35\_r171\_r35\_r171\_r35\_r13\_r171\_r35\_r171\_r35\_r35\_r171\_r35\_r35\_r171\_r35\_r35\_r35\_r355\_r355\_r355\_r355\_r355\_r$
$mc12\_8TeV.167766.Sherpa\_CT10\_WmunuMassiveCBPt70\_140\_CJetVetoBVeto.merge.NTUP\_SMWZ.e1714\_a159\_a171\_r3549\_p1328/cdetVetoBVeto.merge.NTUP\_SMWZ.e1714\_a159\_a171\_r3549\_p1328/cdetVetoBVeto.merge.NTUP\_SMWZ.e1714\_a159\_a171\_r3549\_p1328/cdetVetoBVeto.merge.NTUP\_SMWZ.e1714\_a159\_a171\_r3549\_p1328/cdetVetoBVeto.merge.NTUP\_SMWZ.e1714\_a159\_a171\_r3549\_p1328/cdetVetoBVetoBVeto.merge.NTUP\_SMWZ.e1714\_a159\_a171\_r3549\_p1328/cdetVetoBVetoBVetoBVetoBVetoBVetoBVetoBVet$
$mc12\_8TeV.167773.Sherpa\_CT10\_WmunuMassiveCBPt140\_280\_BFilter.merge.NTUP\_SMWZ.e1741\_a159\_a171\_r3549\_p1328/2000$
$mc12\_8 TeV.167774. Sherpa\_CT10\_WmunuMassiveCBPt140\_280\_CJetFilterBVeto.merge.NTUP\_SMWZ.e1714\_a159\_a171\_r3549\_p1328/contentsetemberge.NTUP\_SMWZ.e1714\_a159\_a171\_r3549\_p1328/contentsetemberge.NTUP\_SMWZ.e1714\_a159\_a171\_r3549\_p1328/contentsetemberge.NTUP\_SMWZ.e1714\_a159\_a171\_r3549\_p1328/contentsetemberge.NTUP\_SMWZ.e1714\_a159\_a171\_r3549\_p1328/contentsetemberge.NTUP\_SMWZ.e1714\_a159\_a171\_r3549\_p1328/contentsetemberge.NTUP\_SMWZ.e1714\_a159\_a171\_r3549\_p1328/contentsetemberge.NTUP\_SMWZ.e1714\_a159\_a171\_r3549\_p1328/contentsetemberge.NTUP\_SMWZ.e1714\_a159\_a171\_r3549\_p1328/contentsetemberge.NTUP\_SMWZ.e1714\_a159\_a171\_r3549\_p1328/contentsetemberge.NTUP\_SMWZ.e1714\_a159\_a171\_r3549\_p1328/contentsetemberge.NTUP\_SMWZ.e1714\_a159\_a171\_r3549\_p1328/contentsetemberge.NTUP\_SMWZ.e1714\_a159\_a171\_r3549\_p1328/contentsetemberge.NTUP\_smuze.NTUP\_sm$
$mc12\_8 TeV.167775. Sherpa\_CT10\_WmunuMassiveCBPt140\_280\_CJetVetoBVeto.merge.NTUP\_SMWZ.e1714\_a159\_a171\_r3549\_p1328/cmc12\_structures and the structure of the st$
$mc12\_STeV.167782.Sherpa\_CT10\_WmunuMassiveCBPt280\_500\_BFilter.merge.NTUP\_SMWZ.e1714\_s1581\_s1586\_r3658\_r3549\_p1328/s1286\_s1286$
$mc12\_8TeV.167783.Sherpa\_CT10\_WmunuMassiveCBPt280\_500\_CJetFilterBVeto.merge.NTUP\_SMWZ.e1714\_s1581\_s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r3658\_r3549\_p1328/s1586\_r3658\_r365658\_r3658\_r3658\_r3658\_r3658\_r3658\_r365658\_r36565$-r3658\_r36565$-r36565$-r3656$-r3656$-r3656$-r365$-r3$
$mc12\_8 TeV.167784. Sherpa\_CT10\_WmunuMassiveCBPt280\_500\_CJetVetoBVeto.merge.NTUP\_SMWZ.e1714\_s1581\_s1586\_r3658\_r3549\_p1328/cms12=0.000$
$mc12\_8 TeV.167791.Sherpa\_CT10\_WmunuMassiveCBPt500\_BFilter.merge.NTUP\_SMWZ.e1714\_s1581\_s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r3658\_r3549\_p1328/s1586\_r3658\_r3658\_r3549\_p1328/s1586\_r3658\_r365658\_r3658\_r3658\_r3658\_r3658\_r3658\_r3658\_r3658\_r3658\_r3656\_r365658\_r365658\_r36565658\_r365656565656565656565656565656565656565$
$mc12\_8TeV.167792.Sherpa\_CT10\_WmunuMassiveCBPt500\_CJetFilterBVeto.merge.NTUP\_SMWZ.e1714\_s1581\_s1586\_r3658\_r3549\_p1328/s1286\_s1686\_r3658\_r3549\_p1328/s1286\_s1686\_r3658\_r3549\_p1328/s1286\_s1686\_r3658\_r3549\_p1328/s1286\_s1686\_r3658\_r3549\_p1328/s1286\_s1686\_r3658\_r3549\_p1328/s1286\_s1686\_r3658\_r3549\_p1328/s1286\_s1686\_r3658\_r3549\_p1328/s1286\_s1686\_r3658\_r3549\_p1328/s1286\_s1686\_r3658\_r3549\_p1328/s1286\_s1686\_r3658\_r3658\_r3654\_r3658\_r3654\_r3658\_r3654\_r365\_r365\_r3654\_r365\_r365\_r365\_r365\_r365\_r365\_r365\_r365$
$mc12\_8TeV.167793.Sherpa\_CT10\_WmunuMassiveCBPt500\_CJetVetoBVeto.merge.NTUP\_SMWZ.e1714\_s1581\_s1586\_r3658\_r3549\_p1328/content and an anti-anti-anti-anti-anti-anti-anti-anti-$
$mc12\_8TeV.167767.Sherpa\_CT10\_WtaunuMassiveCBPt70\_140\_BFilter.merge.NTUP\_SMWZ.e1714\_s1581\_s1586\_r3658\_r3549\_p1328/s12861\_s1586\_r3658\_r3549\_p1328/s12861\_s1586\_r3658\_r3549\_p1328/s12861\_s1586\_r3658\_r3549\_p1328/s12861\_s1586\_r3658\_r3549\_p1328/s12861\_s1586\_r3658\_r3549\_p1328/s12861\_s1586\_r3658\_r3549\_p1328/s12861\_s1586\_r3658\_r3549\_p1328/s12861\_s1586\_r3658\_r3549\_p1328/s12861\_s1586\_r3658\_r3549\_p1328/s12861\_s1586\_r3658\_r3549\_r3668\_r3640\_r3658\_r3640\_r3658\_r3640\_r3658\_r3640\_r3658\_r3640\_r3658\_r3640\_r3658\_r3640\_r3658\_r3658\_r3640\_r3658\_r3640\_r3658\_r3640\_r3658\_r3640\_r3658\_r3640\_r3658\_r3640\_r3658\_r3640\_r3658\_r3640\_r3658\_r3640\_r3658\_r3640\_r3658\_r3640\_r3658\_r3640\_r3658\_r3640\_r3658\_r3640\_r3658\_r3640\_r3658\_r3640\_r3658\_r3640\_r3658\_r3640\_r3658\_r3640\_r3658\_r36658\_r36658\_r3658\_r3658\_r3658\_r36658\_r36558\_r3658\_r3658\_r365558\_r3658\_r3658\_r3658\_r3658\_r36558\_r36558\_r365558\_s1655$r365$r365$r365$r365$r365$r365$r365$r$
$mc12\_8 TeV.167768. Sherpa\_CT10\_WtaunuMassiveCBPt70\_140\_CJetFilterBVeto.merge.NTUP\_SMWZ.e1714\_s1581\_s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r365\_r365\_r3658\_r3658\_r3658\_r3658\_r3658\_r3658\_r3658\_r3658\_r3658\_r3658\_r3658\_r3658\_r3658\_r3658\_r3658\_r3658\_r3658\_r3658\_r3658\_r365\_r365\_r365\_r365\_r365\_r365\_r365\_r365$
$mc12\_8TeV.167769.Sherpa\_CT10\_WtaunuMassiveCBPt70\_140\_CJetVetoBVeto.merge.NTUP\_SMWZ.e1714\_s1581\_s1586\_r3658\_r3549\_p1328/s1281\_s1586\_r3658\_r3549\_p1328/s1281\_s1586\_r3658\_r3549\_p1328/s1281\_s1586\_r3658\_r3658\_r3649\_p1328/s1281\_s1586\_r3658\_r3649\_p1328/s1281\_s1586\_r3658\_r3649\_p1328/s1281\_s1586\_r3658\_r3649\_p1328/s1281\_s1586\_r3658\_r3649\_p1328/s1281\_s1586\_r3658\_r3649\_p1328/s1281\_s1586\_r3658\_r3649\_p1328/s1281\_s1586\_r3658\_r3649\_r3649\_r3640$
$mc12\_sTeV.167776.Sherpa\_CT10\_WtaunuMassiveCBPt140\_280\_BFilter.merge.NTUP\_SMWZ.e1741\_s1581\_s1586\_r3658\_r3549\_p1328/staunuMassiveCBPt140\_280\_BFilter.merge.NTUP\_SMWZ.e1741\_s1581\_s1586\_r3658\_r3549\_p1328/staunuMassiveCBPt140\_s10\_staunuMassiveCBPt140\_staunuMassiveCBP$
$mc12\_8TeV.167777.Sherpa\_CT10\_WtaunuMassiveCBPt140\_280\_CJetFilterBVeto.merge.NTUP\_SMWZ.e1714\_s1581\_s1586\_r3658\_r3549\_p1328/citerBVeto.merge.NTUP\_SMWZ.e1714\_s1581\_s1586\_r3658\_r3549\_p1328/citerBVeto.merge.NTUP\_SMWZ.e1714\_s1581\_s1586\_r3658\_r3549\_p1328/citerBVeto.merge.NTUP\_SMWZ.e1714\_s1581\_s1586\_r3658\_r3549\_p1328/citerBVeto.merge.NTUP\_SMWZ.e1714\_s1581\_s1586\_r3658\_r3549\_p1328/citerBVeto.merge.NTUP\_SMWZ.e1714\_s1581\_s1586\_r3658\_r3549\_p1328/citerBVeto.merge.NTUP\_SMWZ.e1714\_s1581\_s1586\_r3658\_r3549\_p1328/citerBVeto.merge.NTUP\_SMWZ.e1714\_s1581\_s1586\_r3658\_r3549\_p1328/citerBVeto.merge.NTUP\_SMWZ.e1714\_s1581\_s1586\_r3658\_r3549\_p1328/citerBVeto.merge.NTUP\_SMWZ.e1714\_s1581\_s1586\_r3658\_r3549\_p1328/citerBVeto.merge.NTUP\_SMWZ.e1704\_r365\_r365\_r365\_r3658\_r36555\_r3655\_r3655\_r3655\_r3655\_r36555\_r3655\_r36555\_r3655\_r3655\_r36555\_r3655\_r3655\_r36555\_r3655\_r3655\_r3655\_r3655\_r3655\_r3655\_r36555\_r3655\_r3655\_r3655\_r3655\_r3655\_r3655\_r3655\_r3655\_r3655\_r3655\_r3655\_r365\_r36$
$mc12\_8 TeV.167778. Sherpa\_CT10\_WtaunuMassiveCBPt140\_280\_CJetVetoBVeto.merge.NTUP\_SMWZ.e1714\_s1581\_s1586\_r3658\_r3549\_p1328/cms12$
$mc12\_sTeV.167785.Sherpa\_CT10\_WtaunuMassiveCBPt280\_500\_BFilter.merge.NTUP\_SMWZ.e1714\_s1581\_s1586\_r3658\_r3549\_p1328/staunuMassiveCBPt280\_staunuMassiveCBPt28$
$mc12\_8TeV.167786.Sherpa\_CT10\_WtaunuMassiveCBPt280\_500\_CJetFilterBVeto.merge.NTUP\_SMWZ.e1714\_s1581\_s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r3549\_p1328/s1586\_r3658\_r3658\_r3549\_p1328/s1586\_r3658\_r365658\_r365658\_r3656565658\_r365656556565$-r36565$-r36565$-r3656$-r3656$-r3656$-r3656$-r365$$
mc12_8TeV.167787.Sherpa_CT10_WtaunuMassiveCBPt280_500_CJetVetoBVeto.merge.NTUP_SMWZ.e1714_s1581_s1586_r3658_r3549_p1328/
mc12_8TeV.167794.Sherpa_CT10_WtaunuMassiveCBPt500_BFilter.merge.NTUP_SMWZ.e1714_s1581_s1586_r3658_r3549_p1328/
mc12_8TeV.167795.Sherpa_CT10_WtaunuMassiveCBPt500_CJetFilterBVeto.merge.NTUP_SMWZ.e1714_s1581_s1586_r3658_r3549_p1328/
$mc12\_8 TeV.167796. Sherpa\_CT10\_WtaunuMassiveCBPt500\_CJetVetoBVeto.merge.NTUP\_SMWZ.e1714\_s1581\_s1586\_r3658\_r3549\_p1328/2000$
Weakly Produced $W \to \ell \nu + jets$
mc12_8TeV.129918.Sherpa_CT10_Wenu2JetsEW1JetQCD15GeV.merge.NTUP_SMWZ.e1557_s1773_s1776_r4485_r4540_p1328/
mc12.8TeV.129919.Sherpa_CT10_Wmunu2JetsEW1JetQCD15GeV.merge.NTUP_SMWZ.e1575_s1773_s1776_r4485_r4540_p1328/
mc12_8TeV.129920.Sherpa_CT10_Wtaunu2JetsEW1JetQCD15GeV.merge.NTUP_SMWZ.e1575_s1773_s1776_r4485_r4540_p1328/

Table 7.9: MC samples used in the modified  $W \to \ell \nu + \text{jets}$  control region.

There is no significant background in this region (including multijets) because of the high  $p_T^W$  required in the selection of the modified  $W \to \ell \nu + \text{jets control region}$ .

# 7.5.1.2 $W \rightarrow \ell \nu + jets$ Modified Control Region Yields

Table 7.10 contains the MC yields for strong and weakly produced  $W \rightarrow \ell \nu + \text{jets}$  and the yield in data. The strongly produced  $W \rightarrow \ell \nu + \text{jets}$  is normalized to the NLO cross section calculated by FEWZ which results in an effective k-factor of 1.1. The weakly produced  $W \rightarrow \ell \nu + \text{jets}$  is normalized to the cross section given by SHERPA.

It is clear from Table 7.10 that the Data/MC ratio after the  $\Delta \eta_{j,j} > 4.8$  does not have the same feature seen in the  $\gamma$ +jets control region selection in Table 7.8. The corresponding distributions are shown in Figure 7.5. Both sets of distributions have a feature of overprediction at very high  $\Delta \eta_{j,j}$  but in the case of  $W \rightarrow \ell \nu$ +jets it is not statistically significant enough to appear in Data/MC values in

	a.	*** 1		D : /10
Cut	Strong	Weak	Data	Data/MC
BCH Cleaning	$9093564.52 \pm 3251.78$	$347423.89 \pm 209.17$	$14329762 \pm 3785$	$1.52\pm0.00$
Lepton Trigger	$679468.79 \pm 685.76$	$18223.03 \pm 48.07$	$1197758 \pm 1094$	$1.72\pm0.00$
1 Lepton	$406330.54 \pm 506.72$	$12536.83 \pm 39.91$	$510694 \pm 715$	$1.22\pm0.00$
$p_T^\ell > 30 \text{ GeV}$	$222613.62 \pm 320.28$	$8689.68 \pm 33.31$	$297425 \pm 545$	$1.29\pm0.00$
$p_T^{\overline{W}} > 150 \text{ GeV}$	$217840.34 \pm 310.94$	$8542.54 \pm 33.03$	$267918 \pm 518$	$1.18\pm0.00$
b veto	$182004.32 \pm 292.48$	$5118.54 \pm 25.32$	$187904 \pm 433$	$1.00\pm0.00$
jet $p_T > 75 \text{ GeV}$	$178979.36 \pm 289.08$	$5050.39 \pm 25.15$	$184425 \pm 429$	$1.00\pm0.00$
jet $p_T > 50 \text{ GeV}$	$63148.62 \pm 132.78$	$3446.72 \pm 20.78$	$69568 \pm 264$	$1.04\pm0.00$
$\Delta \phi_{j, E_{\mathrm{T}}^{\mathrm{miss}}} > 1.0$	$46326.11 \pm 112.82$	$2603.04 \pm 18.07$	$50572 \pm 225$	$1.03\pm0.01$
$\Delta \phi_{j,j} < 2.5$	$44942.72 \pm 110.69$	$2516.69\pm17.77$	$48988 \pm 221$	$1.03\pm0.01$
Opp Hemispheres	$17863.77 \pm 70.01$	$1224.86 \pm 12.38$	$19852 \pm 141$	$1.04\pm0.01$
$\Delta \eta_{j,j} > 4.8$	$257.50 \pm 10.92$	$116.49 \pm 3.83$	$393\pm20$	$1.05\pm0.06$
$m_{jj} > 1 \text{ TeV}$	$222.74 \pm 10.14$	$112.22 \pm 3.76$	$354 \pm 19$	$1.06\pm0.07$
CJV	$118.26 \pm 8.14$	$93.29 \pm 3.43$	$228 \pm 15$	$1.08\pm0.08$

Table 7.10: Yields of MC strongly produced  $W \to \ell \nu + \text{jets}$ , MC weakly produced  $W \to \ell \nu + \text{jets}$  and data as a function of the modified  $W \to \ell \nu + \text{jets}$  control region selection as described in 7.2.4.

Table 7.10.

# 7.6 Anti-CJV Cross-Check

In order to test the method, an additional set of regions  $(\gamma, W \text{ and } Z)$  are examined that are the similar to the modified signal regions except with the third jet veto reversed. This region has the advantage that the  $Z \rightarrow \nu\nu$ +jets predictions (from  $\gamma$ +jets and  $W \rightarrow \ell\nu$ +jets) can be compared to data without unblinding the signal region. In order to do this, new control regions must be defined similar to 7.2 but with the third jet veto reversed. The same MC samples used for the  $\gamma$ +jets control region described in 7.4.1 and  $W \rightarrow \ell\nu$ +jets described in 7.5.1.1 are used for the corresponding anti-CJV regions. The multijet estimates are neglected in these control regions because they are seen to have small relative contributions in the  $\gamma$ +jets control region and modified  $W \rightarrow \ell\nu$ +jets control region.

# 7.6.1 $\gamma$ +jets Anti-CJV Control Region Selection

The cuts are once again classified as either "2j" or "VBF". The  $\gamma$ +jet anti-CJV selection is the following:

# **2j Selection**

1. Preselection: Good Run List, at least one vertex with  $N_{track} \ge 2$ 





ified  $W \to \ell \nu + jets$  selection applied through the oppo- ified  $W \to \ell \nu + jets$  selection applied through the opposite hemispheres cut.

(a) Data and MC distributions of  $\Delta \eta_{j,j}$  with the mod- (b) Data/MC ratio as a function of  $\Delta \eta_{j,j}$  with the modsite hemispheres cut.

Figure 7.5:  $\Delta \eta_{j,j}$  with the modified  $W \to \ell \nu$ +jets selection applied through the opposite hemispheres cut.

- 2. Object Selection: LCtopo jets with  $p_T > 30$  GeV,  $|\eta| < 4.5$ , |JVF| > 0.5 for jets with  $|\eta| < 2.4$  and  $p_T < 50~{
  m GeV},$  METRefFinal
- 3. Trigger on EF\_g20\_loose or EF\_g40\_loose or EF\_g60\_loose or EF\_g80\_loose or EF\_g100\_loose or EF\_g120\_loose
- 4. Event Cleaning: Tight BCH Cleaning, "Looser" jet cleaning on jets above 20 GeV
- 5. Require exactly 1 isolated, tight PID, unconverted  $\gamma$
- 6.  $E_{\rm T}^{{
  m miss}\,\prime} > 150~{
  m GeV}$

- 7. Lepton veto: Reject events with Medium++  $p_{T,e} > 10$  GeV, StacoCombined  $p_{T,\mu} > 5$  GeV in  $|\eta| < 2.5$  and MS only  $p_{T,\mu} > 5$  GeV in  $2.5 < |\eta| < 2.7$ , Medium BDT  $p_{T,\tau} > 20$  GeV
- 8. b veto: Using 80% working point with  $p_T > 20$  GeV
- 9. Jet  $p_T$ : Leading jet  $p_T > 75$  GeV, Subleading jet  $p_T > 50$  GeV
- **10.**  $\Delta \phi_{jet,\gamma} > 1.0$

**VBF** Selection

- 11. \*\* Third Jet Required with  $p_T > 30$  GeV (different than the  $\gamma$ +jets selection) \*\*
- 12.  $\Delta \phi_{jet,X} < 2.5$
- 13. Leading and subleading jets must be in opposite hemispheres, i.e.  $\eta_{jet 1} \times \eta_{jet 2} < 0$
- 14.  $\Delta \eta_{j,j} > 4.8$
- 15.  $m_{jj} > 1 \ TeV$

Photons in this region are also required to be isolated, unconverted, tight ID and follow all object quality cuts per ATLAS e/gamma prescription just as in the  $\gamma$ +jets control region.

# 7.6.2 Modified Signal Anti-CJV Region Selection

Just as in the modified signal region selection, a modified signal anti-CJV region selection is also defined to keep the cut order similar to the  $\gamma$ +jets anti-CJV control region. The selection is similarly divided between 2j selection and *VBF* selection.

# **2j Selection**

- 1. Preselection: Good Run List, at least one vertex with  $N_{track} \ge 2$
- 2. Object Selection: LCtopo jets with  $p_T > 30$  GeV,  $|\eta| < 4.5$ ,  $|\mathbf{JVF}| > 0.5$  for jets with  $|\eta| < 2.4$  and  $p_T < 50$  GeV, METRefFinal
- 3. Trigger on EF\_xe80\_tclcw or EF\_xe80\_tclcw\_loose
- 4. Event Cleaning: Tight BCH Cleaning, "Medium" jet cleaning on jets above 20 GeV

5. 
$$E_{\rm T}^{\rm miss} > 150 ~{\rm GeV}$$

- 6. Lepton veto: Reject events with Medium++  $p_{T,e} > 10$  GeV, StacoCombined  $p_{T,\mu} > 5$  GeV in  $|\eta| < 2.5$  and MS only  $p_{T,\mu} > 5$  GeV in  $2.5 < |\eta| < 2.7$ , Medium BDT  $p_{T,\tau} > 20$  GeV
- 7. b veto: Using 80% working point with  $p_T > 20$  GeV
- 8. Jet  $p_T$ : Leading jet  $p_T > 75$  GeV, Subleading jet  $p_T > 50$  GeV
- **9.**  $\Delta \phi_{jet, E_{T}^{miss}} > 1.0$

**VBF** Selection

- 10. \*\* Third Jet Required with  $p_T > 30$  GeV (different than the modified signal selection) \*\*
- 11.  $\Delta \phi_{jet,X} < 2.5$
- 12. Leading and subleading jets must be in opposite hemispheres, i.e.  $\eta_{jet 1} \times \eta_{jet 2} < 0$
- 13.  $\Delta \eta_{j,j} > 4.8$
- 14.  $m_{jj} > 1 \ TeV$

## 7.6.3 Modified W+jets Anti-CJV Control Region Selection

In order to compare the modeling of various V+jets processes in the anti-CJV region, a modified  $W \rightarrow \ell \nu + j$ ets anti-CJV selection is used as a region to be able to cross-check results. This is similar to the  $\gamma+j$ ets anti-CJV control region except that a Tight ID++ electron or Staco Combined Muon is required with  $p_T^{\ell} > 30$  GeV. The same  $E_T^{\text{miss}}$  trigger for the signal region is used in the  $W \rightarrow \ell \nu + j$ ets region. As a result, a  $E_T^{\text{miss}}$  cut of 120 GeV is used to ensure that the trigger is extremely efficient. Since the W is used to model the Z in the modified signal anti-CJV control region, a requirement that  $p_T^W > 150$  GeV is imposed to make it kinematically similar. As done in the  $\gamma+j$ ets anti-CJV control region and the modified signal anti-CJV region, the selection is divided between **2j** selection and *VBF* selection.

#### **2j Selection**

- 1. Preselection: Good Run List, at least one vertex with  $N_{track} \geq 2$
- 2. Object Selection: LCtopo jets with  $p_T > 30$  GeV,  $|\eta| < 4.5$ ,  $|\mathbf{JVF}| > 0.5$  for jets with  $|\eta| < 2.4$  and  $p_T < 50$  GeV, Tight++ electrons, Staco Combined muons, METRefFinal

- 3. Trigger on EF\_xe80\_tclcw or EF\_xe80\_tclcw\_loose
- 4. Event Cleaning: Tight BCH Cleaning, 'Medium" jet cleaning on jets above 20 GeV
- 5.  $E_{\rm T}^{\rm miss} > 120 ~{\rm GeV}$
- 6. Exactly 1 electron or muon
- 7.  $p_T^{\ell} > 30 \,\, {\rm GeV}$
- 8.  $p_T^W > 150 \text{ GeV}$
- 9. b veto: Using 80% working point with  $p_T > 20$  GeV
- 10. Jet  $p_T$ : Leading jet  $p_T > 75$  GeV, Subleading jet  $p_T > 50$  GeV
- **11.**  $\Delta \phi_{jet, E_{T}^{miss}} > 1.0$ 
  - **VBF** Selection
- 12. \*\* Third Jet Required with  $p_T > 30$  GeV (different than W+jets selection) \*\*
- 13.  $\Delta \phi_{jet,X} < 2.5$
- 14. Leading and subleading jets must be in opposite hemispheres, i.e.  $\eta_{jet 1} \times \eta_{jet 2} < 0$
- 15.  $\Delta \eta_{j,j} > 4.8$
- 16.  $m_{jj} > 1 \ TeV$

# 7.7 $\gamma$ +jets Anti-CJV Control Region Yields

Presented in Table 7.11 are the MC yields for strong and weakly produced  $\gamma$ +jets and the yield in data. The strongly produced  $\gamma$ +jets MC is normalized to the NLO cross section calculated by JETPHOX which results in an effective k-factor of 1.04. The weakly produced  $\gamma$ +jets MC is normalized to the cross section given by SHERPA.

Similar to the  $\gamma$ +jets control region yields in Section 7.5 it is seen that the modeling changes significantly after the  $\Delta \eta_{j,j} > 4.8$  requirement. The effect is visualized in Figure 7.6. The modified  $W \rightarrow \ell \nu$ +jets anti-CJV control region is used to cross check this effect.

Cut	Strong	Weak	Data	Data/MC
BCH Cleaning	$11066513.56 \pm 8832.51$	$546766.58 \pm 401.67$	$884955010 \pm 1862812$	$76.20 \pm 0.17$
Isolated Tight $\gamma$	$4709667.86 \pm 5725.50$	$79865.02 \pm 153.46$	$884955010 \pm 1862812$	$184.77 \pm 0.45$
$p_T^{\gamma} > 150 \text{ GeV}$	$1656947.68 \pm 2685.16$	$1067.60 \pm 17.71$	$1767280 \pm 1329$	$1.07\pm0.00$
e veto	$1637831.53 \pm 2670.69$	$1054.01 \pm 17.59$	$1737406 \pm 1318$	$1.06\pm0.00$
$\mu$ veto	$1564708.52 \pm 2616.21$	$994.49 \pm 17.08$	$1664061 \pm 1290$	$1.06\pm0.00$
au veto	$1515250.36 \pm 2574.66$	$960.88 \pm 16.79$	$1628910 \pm 1276$	$1.07\pm0.00$
b veto	$1267791.21 \pm 2348.31$	$797.41 \pm 15.12$	$1361000 \pm 1167$	$1.07\pm0.00$
jet $p_T > 75 \text{ GeV}$	$784585.89 \pm 1799.72$	$710.75 \pm 14.27$	$824980 \pm 908$	$1.05\pm0.00$
jet $p_T > 50 \text{ GeV}$	$363509.94 \pm 1175.66$	$571.33 \pm 12.80$	$391080 \pm 625$	$1.07\pm0.00$
$\Delta \phi_{j, E_{\mathrm{T}}^{\mathrm{miss}}} > 1.0$	$260411.13 \pm 989.07$	$401.80\pm10.76$	$283587 \pm 533$	$1.09\pm0.00$
Third Jet $p_T > 30 \text{ GeV}$	$92953.80 \pm 575.07$	$9.48 \pm 1.67$	$99622 \pm 316$	$1.07\pm0.01$
$\Delta \phi_{j,j} < 2.5$	$85395.22 \pm 541.96$	$9.20 \pm 1.65$	$91030 \pm 302$	$1.07\pm0.01$
Opp Hemispheres	$35651.45 \pm 353.67$	$7.69 \pm 1.50$	$38553 \pm 196$	$1.08\pm0.01$
$\Delta \eta_{j,j} > 4.8$	$799.62 \pm 56.69$	$1.46\pm0.66$	$788 \pm 28$	$0.98\pm0.08$
$m_{jj} > 1 \text{ TeV}$	$683.36 \pm 52.89$	$1.46\pm0.66$	$601\pm25$	$0.88\pm0.08$

Table 7.11: Yields of MC strongly produced  $\gamma$ +jets, MC weakly produced  $\gamma$ +jets and data as a function of the  $\gamma$ +jets anti-CJV control region selection.

Cut	Strong	Weak	Data	Data/MC
BCH Cleaning	$9093564.52 \pm 3251.78$	$347423.89 \pm 209.17$	$14329762 \pm 3785$	$1.52 \pm 0.00$
Lepton Trigger	$679468.79 \pm 685.76$	$18223.03 \pm 48.07$	$1197758\pm1094$	$1.72\pm0.00$
1 Lepton	$406330.54 \pm 506.72$	$12536.83 \pm 39.91$	$510694 \pm 715$	$1.22\pm0.00$
$p_T^\ell > 30 \text{ GeV}$	$222613.62 \pm 320.28$	$8689.68 \pm 33.31$	$297425 \pm 545$	$1.29\pm0.00$
$p_T^{\overline{W}} > 150 \text{ GeV}$	$217840.34 \pm 310.94$	$8542.54 \pm 33.03$	$267918 \pm 518$	$1.18\pm0.00$
b veto	$182004.32 \pm 292.48$	$5118.54 \pm 25.32$	$187904 \pm 433$	$1.00\pm0.00$
jet $p_T > 75 \text{ GeV}$	$178979.36 \pm 289.08$	$5050.39 \pm 25.15$	$184425 \pm 429$	$1.00\pm0.00$
jet $p_T > 50 \text{ GeV}$	$63148.62 \pm 132.78$	$3446.72 \pm 20.78$	$69568 \pm 264$	$1.04\pm0.00$
$\Delta \phi_{i,E_{\mathrm{T}}^{\mathrm{miss}}} > 1.0$	$46326.11 \pm 112.82$	$2603.04 \pm 18.07$	$50572 \pm 225$	$1.03\pm0.01$
Third Jet $p_T > 30 \text{ GeV}$	$18699.40 \pm 66.92$	$886.78 \pm 10.55$	$20326 \pm 143$	$1.04\pm0.01$
$\Delta \phi_{j,j} < 2.5$	$17716.22 \pm 64.59$	$840.88 \pm 10.27$	$19267 \pm 139$	$1.04\pm0.01$
Opp Hemispheres	$7527.46 \pm 42.25$	$377.00 \pm 6.87$	$8044\pm90$	$1.02\pm0.01$
$\Delta \eta_{j,j} > 4.8$	$122.15 \pm 6.63$	$20.09 \pm 1.59$	$145 \pm 12$	$1.02\pm0.10$
$m_{jj} > 1 \text{ TeV}$	$104.48 \pm 6.04$	$18.93 \pm 1.54$	$126\pm11$	$1.02\pm0.10$

Table 7.12: Yields of MC strongly produced  $W \to \ell \nu + \text{jets}$ , MC weakly produced  $W \to \ell \nu + \text{jets}$  and data as a function of the modified  $W \to \ell \nu + \text{jets}$  anti-CJV control region selection.

# 7.8 Modified $W \rightarrow \ell \nu + jets$ Anti-CJV Control Region Yields

Table 7.12 contains the MC yields for strong and weakly produced  $W \rightarrow \ell \nu + \text{jets}$  and the yield in data. The strongly produced  $W \rightarrow \ell \nu + \text{jets}$  is normalized to the NLO cross section calculated by FEWZ which results in an effective k-factor of 1.1. The weakly produced  $W \rightarrow \ell \nu + \text{jets}$  is normalized to the cross section given by SHERPA.

Similar to the modified  $W \to \ell \nu + \text{jets}$  control region yields in Section 7.5.1.2, the modeling after the  $\Delta \eta_{j,j} > 4.8$  requirement does not have the same feature in the  $\gamma + \text{jets}$  control regions. The corresponding distributions are shown in Figure 7.7.



spheres cut.

(a) Data and MC distributions of  $\Delta \eta_{j,j}$  with the  $\gamma$ +jets (b) Data/MC ratio as a function of  $\Delta \eta_{j,j}$  with the anti-CJV selection applied through the opposite hemi-  $\gamma$ +jets anti-CJV selection applied through the opposite hemispheres cut.

Figure 7.6:  $\Delta \eta_{j,j}$  with the  $\gamma$ +jets anti-CJV selection applied through the opposite hemispheres cut.

#### Comparison of Modeling for $\gamma$ +jets and $W \rightarrow \ell \nu$ +jets 7.9

Figure 7.8 shows distributions of Data/MC as a function of  $\Delta \eta_{j,j}$  for the  $\gamma$ +jets control region and corresponding anti-CJV region as well as the corresponding overlayed distribution for  $W \to \ell \nu + \text{jets}$ .

The effect of MC overpredicting data seems consistent in the  $\gamma$ +jets samples while it remains unclear in the case of  $W \to \ell \nu + jets$ . At this point it is possible to calculate estimates of the  $Z \to jeta$  $\nu\nu$ +jets process in the anti-CJV region using the method described in 7.3.





ified  $W \rightarrow \ell \nu$ +jets anti-CJV selection applied through ified  $W \rightarrow \ell \nu$ +jets anti-CJV selection applied through the opposite hemispheres cut.

(a) Data and MC distributions of  $\Delta \eta_{j,j}$  with the mod- (b) Data/MC ratio as a function of  $\Delta \eta_{j,j}$  with the modthe opposite hemispheres cut.

Figure 7.7:  $\Delta \eta_{j,j}$  with the modified  $W \rightarrow \ell \nu + jets$  anti-CJV selection applied through the opposite hemispheres cut.

# Calculation of $Z \rightarrow \nu\nu + {\rm jets}$ Process in Modified Signal Anti-CJV 7.10Region

The extrapolation method described in Section 7.3 is used to do the calculation of strongly produced  $Z \rightarrow \nu \nu +$  jets in the modified signal anti-CJV region. Similarly, the estimate is defined as:

$$Z_{\nu\nu}^{\text{estimate, anti-CJV}} = \frac{\gamma_{\text{VBF, anti-CJV}}^{\text{Data}} - \gamma_{\text{VBF, anti-CJV}}^{\text{Weak MC}}}{\gamma_{2j, \text{ anti-CJV}}^{\text{Data}} - \gamma_{2j, \text{ anti-CJV}}^{\text{Weak MC}}} R \text{ Data}_{2j, \text{ anti-CJV}}$$
(7.8)





opposite hemispheres cut.

(a) Data/MC ratio as a function of  $\Delta \eta_{j,j}$  of the  $\gamma$ +jets (b) Data/MC ratio as a function of  $\Delta \eta_{j,j}$  of the  $W \rightarrow \Delta \eta_{j,j}$ control region selection and the  $\gamma$ +jets anti-CJV con- $\ell\nu$ +jets control region selection and the  $W \rightarrow \ell\nu$ +jets trol region. The 2 selections are applied through the anti-CJV control region. The 2 selections are applied through the opposite hemispheres cut.

Figure 7.8:  $\Delta \eta_{j,j}$  Data/MC ratios for  $\gamma$ +jets and  $W \rightarrow \ell \nu$ +jets. Corresponding selections are applied through the opposite hemispheres cut.

R is also defined similarly:

$$R = \frac{Z_{\rm VBF, anti-CJV}^{\rm Strong MC} / Z_{2j, anti-CJV}^{\rm Strong MC}}{\gamma_{\rm VBF, anti-CJV}^{\rm Strong MC} / \gamma_{2j, anti-CJV}^{\rm Strong MC}}$$
(7.9)

However, just as in the modified signal region there is a significant contamination by  $W \to \ell \nu + jets$ where one lepton is not detected. This contamination must be subtracted from the  $Data_{2i}$  yield in order to obtain a correct estimate. This is done by measuring the yield in the modified  $W \to \ell \nu + jets$ anti-CJV control region and applying an extrapolation factor to the modified signal anti-CJV region.

$$W_{\text{lost lepton}} = W_{\text{anti-CJV CR}}^{\text{Data}} \frac{W_{\text{anti-CJV SR}}^{\text{MC}}}{W_{\text{anti-CJV CR}}^{\text{MC}}}$$
(7.10)

		Z Estimate	Z Estimate			
Cut	W Estimate	from $\gamma$ +jets	from $W$ +jets	Z  MC	$t\overline{t}$	Data
Third Jet	$19639.09 \pm 172.20$	$25134.49 \pm 264.98$	$25591.46 \pm 300.81$	$26982.90 \pm 58.88$	$664.40 \pm 16.03$	$44379 \pm 211$
$\Delta \phi_{j,j} < 2.5$	$18405.44 \pm 165.65$	$23581.45 \pm 251.94$	$24147.28 \pm 287.45$	$25448.57 \pm 56.74$	$630.85 \pm 15.62$	$41661 \pm 204$
Opp Hemispheres	$7580.83 \pm 105.54$	$10058.35 \pm 138.41$	$9954.95\pm154.39$	$10705.08 \pm 37.17$	$183.70 \pm 8.44$	$17568\pm133$
$\Delta \eta_{j,j} > 4.8$	$166.15 \pm 18.17$	$170.93 \pm 13.31$	$184.96 \pm 19.52$	$197.25 \pm 6.22$	$8.43 \pm 1.78$	$387 \pm 20$
$m_{jj} > 1 \text{ TeV}$	$136.37 \pm 15.84$	$125.64 \pm 10.37$	$150.35\pm16.86$	$159.88 \pm 5.15$	$7.17\pm1.66$	$322\pm18$
Cut	Data/MC (Z from $\gamma$ +jets)		Data/MC (Z from $W$ +jets)		Data/MC ( $Z$ from MC)	
Third Jet	0.98	± 0.01	0.97 ±	= 0.01	0.94 ±	0.01
$\Delta \phi_{j,j} < 2.5$	$\Delta \phi_{i,i} < 2.5$ $0.98 \pm 0.01$		$0.96 \pm 0.01$		$0.94 \pm 0.01$	
Opp Hemispheres	pp Hemispheres $0.99 \pm 0.01$		$0.99 \pm 0.01$		$0.95 \pm 0.01$	
$\Delta \eta_{j,j} > 4.8$	$j > 4.8$ $1.12 \pm 0.09$		1.08 ±	= 0.10	1.04 ±	0.08
$m_{ii} > 1$ TeV	1.20 =	$\pm 0.11$	1.10 ±	= 0.11	1.06 ±	0.08

Table 7.13: Yields of data and estimated processes in the modified anti-cjv signal region. W+jets is estimated from the method described in 7.10 and estimates of strongly produced  $Z \rightarrow \nu\nu$ +jets are done using the  $\gamma$ +jets anti-cjv control region and the modified  $W \rightarrow \ell\nu$ +jets anti-cjv control region.  $Z \rightarrow \nu\nu$  MC is also presented for comparison. In all three  $Z \rightarrow \nu\nu$  estimates, electroweak  $Z \rightarrow \nu\nu$  is estimated from MC.

Anti-CJV CR refers to the modified  $W \rightarrow \ell \nu + \text{jets}$  anti-CJV control region while anti-CJV SR refers to the modified signal anti-CJV region.

As a cross check, the method described in Section 7.3 is also performed with the modified  $W \rightarrow \ell\nu$ +jets anti-CJV control region. In both cases, the weakly produced  $Z \rightarrow \nu\nu$  yield is taken from MC and added to the strongly produced yield produced by the estimation method. The results are presented in Table 7.13.

Even though the  $\gamma$ +jets estimate and  $W \rightarrow \ell \nu$ +jets based estimates are within  $1\sigma$  of uncertainty, the modeling differences are clear and would benefit from further investigation in the future, particularly when larger datasets are available.

# 7.11 Medium $m_{jj}$ Cross-Check

To isolate the modeling problems to high  $\Delta \eta_{j,j}$  and use a region with more statistical power, the  $\gamma$ +jets and  $W \rightarrow \ell \nu$ +jets control regions with a lower range of  $m_{jj}$  are used to ensure the remaining VBF cuts do not have significant modeling issues. These control regions require 500 GeV  $< m_{jj} <$  1000 GeV and  $\Delta \eta_{j,j} > 2.5$ . The selections are defined as:

# 7.11.1 $\gamma$ +jets Medium $m_{jj}$ Control Region Selection

The cuts are again classified as either "2j" or "VBF". The  $\gamma$ +jet medium  $m_{jj}$  selection is the following:

# **2j** Selection

- 1. Preselection: Good Run List, at least one vertex with  $N_{track} \geq 2$
- 2. Object Selection: LCtopo jets with  $p_T > 30$  GeV,  $|\eta| < 4.5$ ,  $|\mathbf{JVF}| > 0.5$  for jets with  $|\eta| < 2.4$  and  $p_T < 50$  GeV, METRefFinal
- 3. Trigger on EF\_g20\_loose or EF\_g40\_loose or EF\_g60\_loose or EF\_g80\_loose or EF\_g100\_loose or EF\_g120\_loose
- 4. Event Cleaning: Tight BCH Cleaning, "Looser" jet cleaning on jets above 20 GeV
- 5. Require exactly 1 isolated, tight PID, unconverted  $\gamma$
- 6.  $E_{\rm T}^{{
  m miss}\,\prime} > 150~{
  m GeV}$
- 7. Lepton veto: Reject events with Medium++  $p_{T,e} > 10$  GeV, StacoCombined  $p_{T,\mu} > 5$  GeV in  $|\eta| < 2.5$  and MS only  $p_{T,\mu} > 5$  GeV in  $2.5 < |\eta| < 2.7$ , Medium BDT  $p_{T,\tau} > 20$  GeV
- 8. b veto: Using 80% working point with  $p_T > 20$  GeV
- 9. Jet  $p_T$ : Leading jet  $p_T > 75$  GeV, Subleading jet  $p_T > 50$  GeV
- **10.**  $\Delta \phi_{jet,\gamma} > 1.0$

**VBF** Selection

- 11.  $\Delta \phi_{jet,X} < 2.5$
- 12. Leading and subleading jets must be in opposite hemispheres, i.e.  $\eta_{jet 1} \times \eta_{jet 2} < 0$
- 13.  $\Delta \eta_{j,j} > 2.5$
- 14. \*\* 500 GeV <  $m_{jj}$  < 1000 GeV (different than  $\gamma$ +jets selection) \*\*
- 15. Third Jet Veto (reject events with third jet with  $p_T > 30 \text{ GeV}$ )

Photons are again required to be isolated, unconverted, tight ID and follow all object quality cuts per ATLAS e/gamma prescription.
#### 7.11.2 Modified W+jets Medium $m_{jj}$ Control Region Selection

The modified  $W \to \ell \nu$ +jets medium  $m_{jj}$  selection is used as a region to be able to cross-check results. This is similar to the  $\gamma$ +jets medium  $m_{jj}$  control region except that a Tight ID++ electron or Staco Combined Muon is required with  $p_T > 30$  GeV. The same MET trigger for the signal region is used in the  $W \to \ell \nu$ +jets region. As a result, a MET cut of 120 GeV is used to ensure very high efficiency of the trigger. Since the W is used to model the Z, a requirement that the  $p_T^W > 150$  GeV is imposed to make it kinematically similar to the signal region. As done in the  $\gamma$ +jets medium  $m_{jj}$  control region, the selection is divided between **2**<sub>j</sub> selection and *VBF* selection.

#### **2j Selection**

- 1. Preselection: Good Run List, at least one vertex with  $N_{track} \geq 2$
- 2. Object Selection: LCtopo jets with  $p_T > 30$  GeV,  $|\eta| < 4.5$ ,  $|\mathbf{JVF}| > 0.5$  for jets with  $|\eta| < 2.4$  and  $p_T < 50$  GeV, Tight++ electrons, Staco Combined muons, METRefFinal
- 3. Trigger on EF\_xe80\_tclcw or EF\_xe80\_tclcw\_loose
- 4. Event Cleaning: Tight BCH Cleaning, "Medium" jet cleaning on jets above 20 GeV
- 5.  $E_{\rm T}^{\rm miss} > 120 ~{\rm GeV}$
- 6. Exactly 1 electron or muon
- 7.  $p_T^{\ell} > 30 \,\, {\rm GeV}$
- 8.  $p_T^W > 150 \text{ GeV}$
- 9. b veto: Using 80% working point with  $p_T > 20$  GeV
- 10. Jet  $p_T$ : Leading jet  $p_T > 75$  GeV, Subleading jet  $p_T > 50$  GeV
- 11.  $\Delta \phi_{jet, E_{\mathrm{T}}^{\mathrm{miss}}} > 1.0$ VBF Selection
- 12.  $\Delta \phi_{jet,X} < 2.5$
- 13. Leading and subleading jets must be in opposite hemispheres, i.e.  $\eta_{jet 1} \times \eta_{jet 2} < 0$
- 14.  $\Delta \eta_{j,j} > 2.5$
- 15. \*\* 500  $GeV < m_{jj} < 1000 GeV$  (different than W+jets selection) \*\*
- 16. Third Jet Veto (reject events with third jet with  $p_T > 30 \text{ GeV}$ )

Cut	Strong	Weak	Data	Data/MC
BCH Cleaning	$11066513.56 \pm 8832.51$	$546766.58 \pm 401.67$	$884955010 \pm 1862812$	$76.20\pm0.17$
Isolated Tight $\gamma$	$4709667.86 \pm 5725.50$	$79865.02 \pm 153.46$	$884955010 \pm 1862812$	$184.77 \pm 0.45$
$p_T^{\gamma} > 150 \text{ GeV}$	$1656947.68 \pm 2685.16$	$1067.60 \pm 17.71$	$1767280 \pm 1329$	$1.07\pm0.00$
e veto	$1637831.53 \pm 2670.69$	$1054.01\pm17.59$	$1737406 \pm 1318$	$1.06\pm0.00$
$\mu$ veto	$1564708.52 \pm 2616.21$	$994.49 \pm 17.08$	$1664061 \pm 1290$	$1.06\pm0.00$
au veto	$1515250.36 \pm 2574.66$	$960.88 \pm 16.79$	$1628910 \pm 1276$	$1.07\pm0.00$
b veto	$1267791.21 \pm 2348.31$	$797.41 \pm 15.12$	$1361000 \pm 1167$	$1.07\pm0.00$
jet $p_T > 75 \text{ GeV}$	$784585.89 \pm 1799.72$	$710.75 \pm 14.27$	$824980 \pm 908$	$1.05\pm0.00$
jet $p_T > 50 \text{ GeV}$	$363509.94 \pm 1175.66$	$571.33 \pm 12.80$	$391080 \pm 625$	$1.07\pm0.00$
$\Delta \phi_{j, E_{\infty}^{\text{miss}}} > 1.0$	$260411.13 \pm 989.07$	$401.80\pm10.76$	$283587 \pm 533$	$1.09\pm0.00$
$\Delta \phi_{j,j} < 2.5$	$249708.11 \pm 962.05$	$376.22 \pm 10.41$	$270814 \pm 520$	$1.08\pm0.00$
Opp Hemispheres	$97487.77 \pm 606.93$	$345.17 \pm 9.96$	$107825 \pm 328$	$1.10\pm0.01$
$\Delta \eta_{j,j} > 2.5$	$35844.00 \pm 384.18$	$304.59 \pm 9.36$	$38998 \pm 197$	$1.08\pm0.01$
$500 \text{ GeV} < m_{jj} < 1000 \text{ GeV}$	$15149.92 \pm 242.50$	$144.10 \pm 6.46$	$16520 \pm 129$	$1.08\pm0.02$
CJV	$8957.05 \pm 194.02$	$140.70 \pm 6.39$	$10271 \pm 101$	$1.13\pm0.03$

Table 7.14: Yields of MC strongly produced  $\gamma$ +jets, MC weakly produced  $\gamma$ +jets and data as a function of the  $\gamma$ +jets medium  $m_{jj}$  control region selection.

#### 7.11.2.1 $\gamma$ +jets Medium $m_{jj}$ Control Region Yields

Presented in Table 7.14 are the MC yields for strong and weak  $\gamma$ +jets and the yield in data. The strongly produced  $\gamma$ +jets MC is normalized to the NLO cross section calculated by JETPHOX which results in an effective k-factor of 1.04. The weakly produced  $\gamma$ +jets MC is normalized to the cross section given by SHERPA.

It is clear that there are no significant modeling issues in the selection. This is cross checked in the modified  $W \rightarrow \ell \nu + \text{jets}$  medium  $m_{jj}$  control region.

#### 7.11.2.2 Modified $W \rightarrow \ell \nu + jets$ Medium $m_{jj}$ Control Region Yields

Table 7.15 contains the MC yields for strongly produced and weak  $W \rightarrow \ell \nu$ +jets and the yield in data. The strongly produced  $W \rightarrow \ell \nu$ +jets is normalized to the NLO cross section calculated by FEWZ which results in an effective k-factor of 1.1. The weakly produced  $W \rightarrow \ell \nu$ +jets is normalized to the cross section given by SHERPA.

It is similarly clear that there are no significant modeling issues in the selection, even though the statistical power is not as high.

Cut	Strong	Weak	Data	Data/MC
BCH Cleaning	$10122779.93 \pm 3414.10$	$379063.93 \pm 218.51$	$16570469 \pm 4071$	$1.58 \pm 0.00$
Lepton Trigger	$754657.53 \pm 721.74$	$19861.08 \pm 50.19$	$1540523 \pm 1241$	$1.99\pm0.00$
1 Lepton	$450170.66 \pm 532.95$	$13572.20 \pm 41.53$	$549706 \pm 741$	$1.19\pm0.00$
$p_T^\ell > 30 \text{ GeV}$	$246134.62 \pm 336.91$	$9387.92 \pm 34.63$	$320242 \pm 566$	$1.25\pm0.00$
$p_T^W > 150 \text{ GeV}$	$240765.60 \pm 327.02$	$9224.76 \pm 34.32$	$288216 \pm 537$	$1.15\pm0.00$
<i>b</i> veto	$201213.71 \pm 307.62$	$5541.12 \pm 26.35$	$201966 \pm 449$	$0.98\pm0.00$
jet $p_T > 75 \text{ GeV}$	$197718.58 \pm 303.83$	$5466.35 \pm 26.17$	$198114 \pm 445$	$0.98\pm0.00$
jet $p_T > 50 \text{ GeV}$	$69760.74 \pm 139.80$	$3727.18 \pm 21.61$	$75213 \pm 274$	$1.02\pm0.00$
$\Delta \phi_{i,E_{\mathrm{m}}^{\mathrm{miss}}} > 1.0$	$51100.38 \pm 118.71$	$2810.77 \pm 18.78$	$54541\pm234$	$1.01\pm0.00$
$\Delta \phi_{j,j} < 2.5$	$49560.32 \pm 116.48$	$2715.91\pm18.47$	$52853 \pm 230$	$1.01\pm0.00$
Opp Hemispheres	$19690.57 \pm 73.73$	$1312.41 \pm 12.81$	$21410\pm146$	$1.02\pm0.01$
$\Delta \eta_{j,j} > 2.5$	$6246.42 \pm 44.23$	$790.98 \pm 9.93$	$7344\pm86$	$1.04\pm0.01$
$500 \text{ GeV} < m_{jj} < 1000 \text{ GeV}$	$3069.69 \pm 29.65$	$357.41 \pm 6.68$	$3649\pm60$	$1.06\pm0.02$
CJV	$1741.50 \pm 23.64$	$255.81 \pm 5.64$	$2123\pm46$	$1.06\pm0.03$

Table 7.15: Yields of MC strongly produced  $W \to \ell \nu + jets$ , MC weakly produced  $W \to \ell \nu + jets$  and data as a function of the modified  $W \to \ell \nu + jets$  medium  $m_{jj}$  control region selection.

# 7.12 Modeling Differences in Anti-CJV, Medium $m_{jj}$ and Signal Selections

In order to see that the modeling is consistent with  $\gamma$ +jets and  $W \rightarrow \ell \nu$ +jets the ratio of Data/MC values is calculated within the medium  $m_{jj}$  control regions:

$$\frac{W_{VBF}/W_{2j}}{\gamma_{VBF}/\gamma_{2j}} = 1.01 \pm 0.04 \tag{7.11}$$

Since this is consistent with unity, it is concluded that there are no significant modeling differences between  $\gamma$ +jets and  $W \rightarrow \ell \nu$ +jets in the medium  $m_{jj}$  region. If a similar calculation is done in the anti-CJV control regions:

$$\frac{W_{VBF}/W_{2j}}{\gamma_{VBF}/\gamma_{2j}} = 1.23 \pm 0.16 \tag{7.12}$$

This is not consistent with unity within  $1\sigma$  of statistical uncertainty which suggests there are differences in the modelings in the *VBF* phase space. Lastly, if the same calculation is done with the signal control regions:

$$\frac{W_{VBF}/W_{2j}}{\gamma_{VBF}/\gamma_{2j}} = 1.35 \pm 0.14 \tag{7.13}$$

A 2.5  $\sigma$  effect is observed, which motivated investigations of NLO effects in VBF production and interference effects of strong and weak processes, described in Section 7.15.

The anti-CJV control regions and the signal control regions have consistent mismodeling in VBFphase space which suggests there is an issue with the high  $\Delta \eta_{j,j}$  not present at low  $\Delta \eta_{j,j}$ . The consistency of the anti-CJV control regions and signal control regions shows the effect is independent of the third jet veto.

#### 7.13 Alternate Method: Anti-CJV Extrapolation

In order to avoid extrapolating from a region of phase space with good modeling (2j) to one with poor modeling (VBF) a different method is proposed where the extrapolation occurs from anti-cjv control regions to the modified signal region. The method looks similar to the one described in 7.3 but with the extrapolation redefined as follows:

$$\frac{\gamma_{\rm SR}^{\rm Data} - \gamma_{\rm SR}^{\rm Weak \ MC}}{\gamma_{\rm anti-CJV}^{\rm Data} - \gamma_{\rm Weak \ MC}^{\rm Weak \ MC}} \tag{7.14}$$

To extrapolate to  $Z \rightarrow \nu\nu + jets$ , the ratio of efficiencies in MC is applied, defined as R:

$$R = \frac{Z_{\rm SR}^{\rm Strong MC} / Z_{\rm anti-CJV}^{\rm Strong MC}}{\gamma_{\rm SR}^{\rm Strong MC} / \gamma_{\rm anti-CJV}^{\rm Strong MC}}$$
(7.15)

The result of which is applied to the data yield in the modified signal anti-cjv region with the 2j selection applied (Data<sub>anti-CJV</sub>). In summary, the estimate is defined as:

$$Z_{\nu\nu}^{\text{estimate}} = \frac{\gamma_{\text{SR}}^{\text{Data}} - \gamma_{\text{SR}}^{\text{Weak MC}}}{\gamma_{\text{anti-CJV}}^{\text{Data}} - \gamma_{\text{anti-CJV}}^{\text{Weak MC}}} R \text{ Data}_{\text{anti-CJV}}$$
(7.16)

In order for the correct cancellations to occur,  $\text{Data}_{\text{anti-CJV}}$  must be pure Z+jets. However, there is a significant contamination by  $W \to \ell \nu$ +jets where one lepton is not detected which must be subtracted as well as multijets.

This method maintains several advantages:

- 1. Does not rely on mismodelings in VBF phase space to be the same across various V+jets processes.
- 2. The MC weak production subtraction is smaller.
- 3. Can be used to as a completely statistically independent cross check of the W control region based estimates
- 4. Should not be very sensitive to PDF and scale variations.
- 5. Should not be very sensitive to detector systematics.

Method	W	$\gamma$
Efficiency Ratio Method	$236.00 \pm 35.06$	$169.83 \pm 18.22$
Efficiency Ratio Method Using Anti-CJV	$269.85 \pm 64.48$	$230.78 \pm 45.38$

Table 7.16: Summary of strongly produced  $Z \rightarrow \nu\nu$  estimates using the extrapolation of efficiencies from 2j to *VBF* and extrapolation from anti-cjv control regions to signal control regions.

Analysis	Signal Generator	$\sqrt{s}$ TeV	$\mu$
VBF $W$	SHERPA	7	$1.63 \pm 0.19 \text{ (stat)} {}^{+0.42}_{-0.52} \text{ (syst)}$
VBF $Z$	POWHEG	8	$1.19 \pm 0.25$ (combined)

Table 7.17: Summary of VBF W and VBF Z  $\mu$  values (ratio of observed cross section to Standard Model prediction). Note that POWHEG is NLO and expects  $\approx 10\%$  more events than SHERPA.

However, there is a large disadvantage in the smaller statistical power. The modified signal anti-CJV region (Data<sub>anti-CJV</sub>) has low statistical power and the data yield of the  $\gamma$ +jets anti-cjv control region is 2/3 of the  $\gamma$ +jets control region. The low statistical power enlarges the uncertainty on the  $Z \rightarrow \nu\nu$ +jets estimate.

#### 7.14 Results from Efficiency Extrapolation Methods

Table 7.16 shows results of using the  $\gamma$ +jets and  $W \rightarrow \ell \nu$ +jets control regions with the method of extrapolation from **2j** to *VBF* ("efficiency method") as well as the results of using the alternate method of extrapolating from anti-cjv control regions to signal control regions to calculate strongly produced  $Z \rightarrow \nu \nu$ .

While the  $W \to \ell \nu$ +jets control regions based estimates are very consistent within statistical uncertainty, there is marginal agreement with the  $\gamma$ +jets control region using the extrapolation from **2j** to *VBF*. This is likely due to the modeling issues that are present in the *VBF* phase space for  $\gamma$ +jets but are not statistically significant in  $W \to \ell \nu$ +jets.

#### 7.15 VBF Normalization

Both VBF W and VBF Z have reported measurements higher than the Standard Model prediction in previous ATLAS measurements documented in [44] and [45] and shown in Table 7.17. The ratio of the observed cross section to the Standard Model prediction is referred to as  $\mu$ . Note that POWHEG is a NLO generator and expects  $\approx 10\%$  more events than SHERPA. Considering this, both measurements are higher than the MC expectation by at least 30%.

This is investigated in two ways, the first is a correction due to NLO QCD effects (typically calculated as a ratio of NLO/LO, known as a k-factor) on the weakly produced W and Z processes.

A NLO QCD generator (VBFNLO) that has weakly produced processes implemented is used to calculate a fiducial k-factor. The fiducial cuts considered for weak W production are:

- 1. Leading jet  $p_T > 75 \text{ GeV}$
- 2. Subleading jet  $p_T > 50 \text{ GeV}$
- 3. Lepton  $p_T > 30 \text{ GeV}$
- 4.  $E_{\rm T}^{\rm miss} > 120 {
  m ~GeV}$
- 5.  $\Delta \eta_{j,j} > 4.8$
- 6.  $m_{jj} > 1 \text{ TeV}$
- 7. Jet veto for jets with  $p_T > 30 \text{ GeV}$

These cuts were chosen to make the fiducial region similar to the Modified  $W \rightarrow \ell \nu + \text{jets}$  Control Region described in 7.2.4. The fiducial cuts considered for weak Z production are:

- 1. Leading jet  $p_T > 75 \text{ GeV}$
- 2. Subleading jet  $p_T > 50 \text{ GeV}$
- 3.  $E_{\rm T}^{\rm miss} > 150 {\rm ~GeV}$
- 4.  $\Delta \eta_{j,j} > 4.8$
- 5.  $m_{jj} > 1 \text{ TeV}$
- 6. Jet veto for jets with  $p_T > 30 \text{ GeV}$

The renormalization and factorization scales are varied within VBFNLO to calculate a theory systematic on the cross section. The eigenvector variation of the CT10 PDF is used to compute an additional theory systematic. The result of the calculation is compared to SHERPA to create a k-factor (which is used to model weakly produced V processes in the analysis) which is shown in Table 7.18.

The resultant k-factors are  $1.12 \pm 0.07 (Q^2 \text{ theory}) \stackrel{+0.07}{_{-0.03}} (\text{PDF theory}) \pm 0.05 \text{ (stat)}$  for weakly produced  $W \rightarrow \ell \nu + \text{jets}$  and  $1.25 \pm 0.05 (Q^2 \text{ theory}) \stackrel{+0.12}{_{-0.04}} \pm 0.06 \text{ (stat)}$  for weakly produced  $Z \rightarrow \nu \nu + \text{jets}$ .

The second effect considered is the interference from separating the strongly produced V+jets from the weakly produced V+jets when producing MC samples. In principle, diagrams from these 2 samples that have the same initial and final states can interfere. In order to calculate the size of this

Variation	$W^+ \to e\nu$ (fb)	$W^- \to e\nu$ (fb)	W k-factor	$Z \to \nu \nu$ (fb)	Z k-factor
SHERPA LO	$2.66\pm0.13$	$0.81\pm0.07$	-	$5.54 \pm 0.26$	-
		VBFNLC	)		
$\mu^F = \mu^R = 1.0$	$2.971 \pm 0.011$	$0.904 \pm 0.003$	$1.12 \pm 0.05$	$6.921 \pm 0.039$	$1.25 \pm 0.06$
$\mu^{F} = 0.5$	$2.905\pm0.011$	$0.897\pm0.004$	$1.10 \pm 0.05$	$6.828 \pm 0.048$	$1.23\pm0.06$
$\mu^{F} = 2.0$	$3.018\pm0.013$	$0.921 \pm 0.004$	$1.14 \pm 0.05$	$7.089 \pm 0.034$	$1.28\pm0.06$
$\mu^{R} = 0.5$	$2.929\pm0.014$	$0.881 \pm 0.006$	$1.10 \pm 0.05$	$6.903 \pm 0.043$	$1.25\pm0.06$
$\mu^{R} = 2.0$	$3.010\pm0.012$	$0.919 \pm 0.004$	$1.13 \pm 0.05$	$7.055 \pm 0.035$	$1.27\pm0.06$
$\mu^{F} = 0.25$	$2.781\pm0.015$	$0.862 \pm 0.004$	$1.05 \pm 0.05$	$6.655 \pm 0.074$	$1.20\pm0.06$
$\mu^{F} = 4.0$	$3.041\pm0.008$	$0.925 \pm 0.003$	$1.14 \pm 0.05$	$7.043 \pm 0.037$	$1.27\pm0.06$
$\mu^{R} = 0.25$	$2.842\pm0.017$	$0.859 \pm 0.007$	$1.07\pm0.05$	$6.661 \pm 0.056$	$1.20\pm0.06$
$\mu^{R} = 4.0$	$3.063\pm0.010$	$0.924 \pm 0.005$	$1.15 \pm 0.05$	$7.064 \pm 0.035$	$1.28\pm0.06$
CT10	$2.971 \stackrel{+0.197}{_{-0.096}}$	$0.904 \ ^{+0.113}_{-0.048}$	$1.12 \begin{array}{c} +0.07 \\ -0.03 \end{array}$	$6.915 \begin{array}{c} +0.692 \\ -0.236 \end{array}$	$1.25 \begin{array}{c} +0.12 \\ -0.04 \end{array}$

Table 7.18: Cross sections of VBF  $W \rightarrow e\nu$  and  $Z \rightarrow \nu\nu$  processes at NLO QCD using VBFNLO. The k-factor is defined with respect to SHERPA. All cross sections are shown in femtobarns.

Sample	$W \to e\nu + 2$ jets (pb)	$Z \rightarrow \nu \nu + 2$ jets (pb)
Inclusive	$1.74746 \pm 0.00277$	$1.00445 \pm 0.00159$
Pure Strong	$1.46406 \pm 0.00236$	$0.87514 \pm 0.00142$
Pure Weak	$0.25672 \pm 0.00038$	$0.11646\pm0.00017$
Interference	$0.10 \pm 0.01$	$0.11\pm0.02$

Table 7.19: Cross sections of pure strong, pure weak and inclusive production for  $W \to e\nu+2$  jets and  $Z \to \nu\nu+2$  jets as well as a calculation of the interference. The interference is calculated relative to the weak cross section, as in Equation 7.17.

effect, 3 SHERPA V+2 jets samples were created: one with ORDER\_EW 2 (pure in strong production), one with ORDER\_EW 4 (pure in weak production) and one with no weak order requirement (strong+weak production). The difference of the inclusive and the other 2 samples is then taken with respect to the pure weakly produced samples resulting in a correction shown in Table 7.19. Since the effect is taken with respect to the weak component, it is applied as a correction to weakly produced MC only.

$$Interference = \frac{Inclusive - Pure Strong - Pure Weak}{Pure Weak}$$
(7.17)

In order to gain statistical power, a fiducial requirement of  $p_T(\text{jets}) > 50 \text{ GeV}$  and  $m_{jj} > 1 \text{ TeV}$  was applied to the 3 samples used for the interference calculation.

#### 7.16 Impact of VBF Normalization on Background Estimates

Using effective k-factors for weakly produced  $W \to \ell \nu + \text{jets}$  and weakly produced  $Z \to \nu \nu + \text{jets}$ which include the correction due to NLO QCD and interference, the revised estimates are shown in Table 7.20 including the weakly produced  $Z \to \nu \nu$  expectation.

Method	W	$\gamma$			
Efficiency Ratio Method	$334.41 \pm 42.03$	$287.22 \pm 22.74$			
Efficiency Ratio Method Using Anti-CJV	$353.56 \pm 69.07$	$349.04 \pm 47.64$			
No Adjustments: Efficiency Ratio Method	$320.60 \pm 34.98$	$254.43 \pm 18.08$			
No Adjustments: Efficiency Ratio Method Using Anti-CJV	$354.45 \pm 64.52$	$315.38 \pm 45.44$			
$W \rightarrow \ell \nu$ based Estimate in Main Analysis: 324.0 ± 17.3					

Table 7.20: Summary of  $Z \to \nu\nu$  estimates using the extrapolation of efficiencies from **2j** to *VBF* and extrapolation from anti-cjv control regions to signal control regions. The first two estimates include weakly produced  $Z \to \nu\nu$  and weakly produced  $W \to \ell\nu$  adjusted by the interference and k-factor. The second two estimates include the weakly produced backgrounds without adjustment. Also shown is the  $W \to \ell\nu$  based estimate calculated in the main analysis using a transfer factor. Statistical and theory certainties are included.

It is seen that the interference and k-factor have a minimal effect on  $W \to \ell \nu$  based estimates of  $Z \to \nu \nu$ , but have more significant contributions to the  $\gamma$  based estimate. The  $\gamma$  and  $W \to \ell \nu$  based estimates of  $Z \to \nu \nu$  are now consistent within  $1\sigma$ . The  $\gamma$  based estimate is also consistent with the  $W \to \ell \nu$  based estimate in the main analysis to  $1.3\sigma$ .

### CHAPTER 8

# Multijet Control Regions

#### 8.1 Introduction

One of the main challenges of estimating the background due to multijet production is the lack of MC statistics that can pass the selection (no events pass the full selection). Additionally, there is a significant disagreement between data and MC earlier in the selection as seen in Figure 8.1 which means the multijet MC should not be used to determine the estimate. As a result, a purely data driven method is used. Two purely data driven estimates are introduced and compared to ensure consistency:

- 1. Efficiency method: Calculate cut efficiencies in various multijet enhanced control regions to estimate the multijet contribution in the signal region
- 2. Jet  $p_T$  extrapolation method: A control region is defined in which the  $\Delta \phi_{jet, E_T^{\text{miss}}}$  requirement is reversed and additional jet is required. The  $p_T$  of the jet which is closest to the  $E_T^{\text{miss}}$  is then extrapolated to the signal region ( $p_{T,jet} < 30 \text{ GeV}$ ) and the yield is calculated.

With the efficiency method, data driven estimates of the multijet background are computed by applying cut efficiencies from multijet enhanced control regions and determining systematics from validation regions. A summary of these regions is presented in Figure 8.2. The first column shows the signal region and the 2 signal region selections which are reversed to enhanced the multijet contribution:  $\Delta \phi_{jx, E_{\rm T}^{\rm miss}}$  and  $\Delta \phi_{j,j}$ . These are represented in Figure 8.2 as **A** (Signal Region), **B** (reverse- $\Delta \phi_{jx, E_{\rm T}^{\rm miss}}$ ) and **C** (reverse- $\Delta \phi_{j,j}$ ). These regions are further split to form multijet enhanced validation regions used to estimate systematics, shown in the second column. Validation regions are defined by reversing combinations of the  $\Delta \eta_{j,j}$  requirement and third jet veto:



Figure 8.1:  $E_{\rm T}^{\rm miss}$  distribution with all of the selection in Section 5.3 applied except for the  $\Delta \phi_{jet, E_{\rm T}^{\rm miss}}$  and  $E_{\rm T}^{\rm miss}$  requirements.

- 1. Validation Region 1 (VR1): Reversing  $\Delta \eta_{j,j} > 4.8$  requirement to  $\Delta \eta_{j,j} < 3.8$
- 2. Validation Region 2 (VR2): Reversing the third jet veto by requiring a third jet with  $p_T > 40 \text{ GeV}$
- 3. Validation Region 3 (VR3): Reversing both cuts simultaneously

When **A**, **B** and **C** are split into the additional multijet enhanced validation regions, there are 12 regions shown in Figure 8.2: (**A**, **B**, **C**) × (SR, VR1, VR2, VR3). Data can be unblinded in each one region except for the signal region (**A** × SR).

#### 8.2 Efficiency Method

The "efficiency method" uses successive cut efficiencies in multijet enhanced control regions and applies them to the signal region. An efficiency ( $\epsilon$ ) at a point *n* in the selection *N* is defined as:

$$\epsilon = \frac{N(n)}{N(n-1)} \tag{8.1}$$



Figure 8.2: Various regions used in the multijet estimation, data can be observed in each region except for the nominal signal region.

	ggH+VBF	Strong+Weak	Strong+Weak		Other	Total		
Cut	Signal	$Z \rightarrow \nu \nu + \text{jets}$	$W \rightarrow \ell \nu + \text{jets}$	Multijet	Backgrounds	Backgrounds	2012 Data	Subtracted
			reverse- $\Delta \phi_{j}$ :	<sub>r,MET</sub> , Signal Region	( <b>B</b> - SR)	•	•	
Jet $p_T$	$1702 \pm 68.6$	$105700 \pm 317$	$188028 \pm 570$	$9410780 \pm 708152$	$4657 \pm 15.1$	$9709165 \pm 708152$	7494649	$7196264 \pm 2814$
Opp. Hemispheres	$906 \pm 45.3$	$48921 \pm 204$	$77201 \pm 338$	$4116083 \pm 479217$	$1532 \pm 8.67$	$4243737 \pm 479217$	3312750	$3185096 \pm 1862$
$\Delta \eta_{ij} > 4.8$	$112.4 \pm 11.3$	$805 \pm 29.9$	$1419 \pm 60.3$	$176239 \pm 115137$	$27.2 \pm 1.02$	$178490 \pm 115137$	74629	$72378 \pm 281$
$m_{jj} > 1 \text{ TeV}$	$109.1 \pm 11.3$	$664 \pm 27.9$	$1199 \pm 58.7$	$75605 \pm 72438$	$22.8 \pm 1.04$	$77491 \pm 72438$	59531	$57645 \pm 252$
$\Delta \phi_{jj} < 2.5$	$50.3 \pm 10.5$	$315 \pm 16.9$	$627 \pm 35.3$	$1116 \pm 982.96$	$16.6 \pm 0.87$	$2074 \pm 984$	24646	$23687 \pm 162$
Central Jet Veto	$6.75 \pm 0.79$	$16.1 \pm 1.99$	$28.6 \pm 4.55$	-	$0.04 \pm 0.03$	$44.7 \pm 4.97$	611	$566 \pm 25.2$
$E_T^{miss} > 150 \text{ GeV}$	$2.73 \pm 0.51$	$6.1 \pm 0.80$	$3.58 \pm 0.86$	-	-	$9.68 \pm 1.17$	9	$-0.68 \pm 3.22$
			reverse- $\Delta \phi_j$	$x_{x,MET}$ , Small- $\Delta \eta_{jj}$ (	<b>B</b> - VR1)			
Jet $p_T$	$1702 \pm 68.6$	$105085 \pm 371$	$186934 \pm 670$	$9410780 \pm 708152$	$4657 \pm 15.1$	$9707456 \pm 708152$	7494649	$7197973 \pm 2843$
Opp. Hemispheres	$906 \pm 45.3$	$48922\pm204$	$77202 \pm 338$	$4116083 \pm 479217$	$1532 \pm 8.67$	$4243739 \pm 479217$	3312750	$3185094 \pm 1862$
$\Delta \eta_{jj} < 3.8$	$602 \pm 39.7$	$44960 \pm 207$	$70387 \pm 338$	$3479977 \pm 432059$	$1415 \pm 8.35$	$3596739 \pm 432059$	2971636	$2854874 \pm 1769$
$m_{jj} > 1 \text{ TeV}$	$15.23 \pm 3.39$	$1236 \pm 61.5$	$1760 \pm 97.7$	$170324 \pm 10791$	$31.9 \pm 1.22$	$173352 \pm 10792$	163271	$160243 \pm 420$
$\Delta \phi_{jj} < 2.5$	$2.94 \pm 0.53$	$152 \pm 9.09$	$221 \pm 14.1$	$5290 \pm 1712$	$8.16 \pm 0.64$	$5671 \pm 1712$	6818	$6437 \pm 84.3$
Central Jet Veto	$0.57 \pm 0.23$	$4.04 \pm 0.71$	$1.44 \pm 0.44$	-	-	$5.48 \pm 0.83$	10	$4.52 \pm 3.27$
$E_T^{miss} > 150 \text{ GeV}$	$0.57 \pm 0.23$	$3.93 \pm 0.64$	$1.47 \pm 0.44$	-	-	$5.40 \pm 0.78$	8	$2.60 \pm 2.93$
			reverse-4	$\Delta \phi_{jx,MET}$ , 3-jet ( <b>B</b>	- VR2)			
Jet $p_T$	$1313 \pm 64.1$	$78097 \pm 270$	$140465 \pm 489$	$7233321 \pm 612476$	$4565 \pm 14.9$	$7456448 \pm 612476$	5363946	$5140819 \pm 2383$
Opp. Hemispheres	$660 \pm 43.6$	$36216 \pm 233$	$60055 \pm 398$	$3141704 \pm 409360$	$1504 \pm 8.55$	$3239479 \pm 409360$	2464210	$2366435 \pm 1636$
$\Delta \eta_{jj} > 4.8$	$63.8 \pm 11.1$	$619 \pm 26.4$	$1176 \pm 53.7$	$176239 \pm 115137$	$27.0 \pm 1.12$	$178061 \pm 115137$	60658	$58836 \pm 253$
$m_{jj} > 1 \text{ TeV}$	$61.5 \pm 11.1$	$500 \pm 22.4$	$989 \pm 47.8$	$75605 \pm 72438$	$22.6 \pm 1.03$	$77117 \pm 72438$	49719	$48207 \pm 229$
$\Delta \phi_{jj} < 2.5$	$44.2 \pm 10.4$	$290 \pm 15.2$	$583 \pm 32.0$	$1116 \pm 983$	$16.5 \pm 0.87$	$2006 \pm 984$	24405	$23515 \pm 160$
$p_T^{j3} > 40 \text{ GeV}$	$27.4 \pm 8.38$	$203 \pm 16.3$	$413 \pm 33.6$	$1116 \pm 983$	$14.8\pm0.81$	$1747 \pm 984$	19284	$18653 \pm 144$
$E_T^{miss} > 150 \text{ GeV}$	$9.79 \pm 4.49$	$54.0 \pm 4.65$	$89.6 \pm 8.93$	-	$4.39\pm0.46$	$148 \pm 10.1$	205	$57.0 \pm 17.5$
	reverse $\Delta \phi_{jx,MET}$ , 3-jet, Small- $\Delta \eta_{jj}$ ( <b>B</b> - VR3)							
Jet $p_T$	$1313 \pm 64.1$	$75809 \pm 270$	$140465 \pm 489$	$7233321 \pm 612476$	$4565 \pm 14.9$	$7456448 \pm 612476$	5363946	$5140819 \pm 2383$
Opp. Hemispheres	$660 \pm 43.6$	$36216 \pm 233$	$60055 \pm 398$	$3141704 \pm 409360$	$1504 \pm 8.55$	$3239479 \pm 409360$	2464210	$2366435 \pm 1636$
$\Delta \eta_{jj} < 3.8$	$478 \pm 38.5$	$33185 \pm 186$	$54455 \pm 313$	$2690776 \pm 375516$	$1388 \pm 8.23$	$2779804 \pm 375516$	2204000	$2114972 \pm 1529$
$m_{jj} > 1 \text{ TeV}$	$8.71 \pm 3.31$	$966 \pm 132$	$1520 \pm 224$	$147368 \pm 10042$	$31.2 \pm 1.22$	$149885 \pm 10045$	133228	$130711 \pm 448$
$\Delta \phi_{jj} < 2.5$	$2.51 \pm 0.49$	$147 \pm 20.9$	$216 \pm 32.6$	$5290 \pm 1712$	$8.16\pm0.64$	$5661 \pm 1712$	6813	$6442 \pm 91.2$
No Jet Veto	$2.51 \pm 0.49$	$147 \pm 20.9$	$216 \pm 32.6$	$5290 \pm 1712$	$8.16 \pm 0.64$	$5661 \pm 1712$	6813	$6442 \pm 91.2$
$E_{\pi}^{miss} > 150 \text{ GeV}$	$2.07 \pm 0.44$	$99.2 \pm 12.5$	$117 \pm 15.9$	$25.2 \pm 12.5$	$5.14 \pm 0.50$	$247 \pm 23.8$	354	$133 \pm 27.6$

Table 8.1: Progressive yields of the reverse- $\Delta \phi_{jx,E_{\mathrm{T}}^{\mathrm{miss}}}$  control regions (**B** selections) described in Section 8.1 in 20.3 fb<sup>-1</sup> of 2012 data, as evaluated using Monte Carlo. Expected signal yields are shown for the case of  $m_H = 125$  GeV. The "Subtracted" column shows an estimate of the dijet contribution by subtracting Z, W and Other Background columns from the observed yield.

The ideal control region to use model the efficiencies of multijet events is one in which the signal region selection is followed except that the  $\Delta \phi_{jx, E_{\mathrm{T}}^{\mathrm{miss}}}$  requirement is reversed. Reversing this requirement enriches the sample in multijet events. However, since that control region cannot be used to model the  $\Delta \phi_{jx, E_{\mathrm{T}}^{\mathrm{miss}}}$  requirement itself, another multijet enhanced control region must be used to estimate the  $\Delta \phi_{jx, E_{\mathrm{T}}^{\mathrm{miss}}}$  efficiency. A region is constructed which follows the signal region selection except that the  $\Delta \phi_{j,j}$  requirement is reversed which is also enhanced in multijet events (since multijet events are typically back to back jets in  $\phi$ ). Systematic uncertainties are determined by observing differences in corresponding validation regions. The multijet efficiences for the **A**, **B** and **C** selections are calculated from Table 5.2, Table 8.1 and Table 8.2 respectively and summarized in Table 8.3.

The multijet efficiencies used to determine the multijet background in the signal region with associated systematic uncertainties are calculated for each selection requirement as follows:

1.  $\Delta \eta_{j,j}$ : Selection **B** -Signal Region is used to determine the efficiency. The difference of efficiencies in VR2 (**B**  $\rightarrow$  **A**) is used to set a systematic of 25%. This is justified because a similar

	aaH+VBF	Strong+Weak	Strong+Weak		Other	Total		
Cut	Signal	$Z \rightarrow \mu \mu + \text{iets}$	$W \rightarrow \ell \nu + \text{iets}$	Multijet	Backgrounds	Backgrounds	2012 Data	Subtracted
	5-Bildi	2 / 00 / joto	reverse-A	du Signal Region (	C - SB)	Buongrounds	2012 Data	Subtracted
Jet n <sub>T</sub>	$951 \pm 45.9$	$87573 \pm 282$	$151841 \pm 500$	$\frac{\varphi_{jj}}{8951583 + 691216}$	$2241 \pm 10.6$	$9193238 \pm 691216$	6954729	$6713074 \pm 2699$
Opp Hemispheres	$526 \pm 27.3$	$41447 \pm 184$	$62040 \pm 298$	$3733455 \pm 452403$	$820 \pm 6.41$	$3837762 \pm 452403$	3041401	$2937094 \pm 1779$
$\Delta n = > 4.8$	$73.1 \pm 4.73$	$526 \pm 20.4$	$910 \pm 43.3$	$86399 \pm 73404$	$10.2 \pm 0.11$ $10.2 \pm 0.73$	$87845 \pm 73404$	49529	$48083 \pm 228$
$m_{\gamma\gamma} > 1 \text{ TeV}$	$711 \pm 4.70$	$467 \pm 20.4$	$784 \pm 41.0$	$86399 \pm 73404$	$8.72 \pm 0.16$	$87659 \pm 73404$	39010	$37750 \pm 203$
Central let Veto	$56.2 \pm 2.27$	$255 \pm 14.0$	$331 \pm 93.7$	$582 \pm 582$	$0.12 \pm 0.00$ $0.32 \pm 0.16$	$1068 \pm 583$	17601	$17115 \pm 135$
$\Delta \phi = \mu r \pi > 1$	$9.51 \pm 0.94$	$59.8 \pm 5.97$	$86.3 \pm 9.54$		$0.02 \pm 0.10$ $0.04 \pm 0.04$	$146 \pm 11.3$	2324	$2470 \pm 11.3$
$E_{miss}^{miss} > 150 \text{ GeV}$	$0.09 \pm 0.07$	$1.28 \pm 0.44$	$1.94 \pm 0.58$	_	0.04 ± 0.04	$3.22 \pm 0.73$	5	$1.78 \pm 2.35$
	0.05 ± 0.01	1.20 ± 0.44	reverse-A	$\phi = \text{Small} - \Delta n = (C)$	- VB1)	0.22 ± 0.10	Ŭ	1.10 ± 2.00
let nm	$951 \pm 45.9$	$87573 \pm 282$	$151841 \pm 500$	$8951583 \pm 691216$	$2241 \pm 10.6$	$9193238 \pm 691216$	6954729	$6713074 \pm 2699$
Opp Hemispheres	$526 \pm 27.3$	$41447 \pm 184$	$62040 \pm 208$	$3733455 \pm 452403$	$820 \pm 6.41$	$3837762 \pm 452403$	3041401	$2937094 \pm 1779$
$\Delta n \leq 3.8$	$320 \pm 21.0$ $330 \pm 24.6$	$38581 \pm 188$	$57385 \pm 290$	$3265920 \pm 417790$	$771 \pm 6.19$	$3362657 \pm 417790$	2764084	$2667347 \pm 1699$
$m_{\odot} > 1$ TeV	$15.43 \pm 3.39$	$1491 \pm 72.7$	$1973 \pm 109$	$169921 \pm 10819$	$26.0 \pm 1.13$	$173411 \pm 10820$	161828	$158338 \pm 423$
Central let Veto	$9.09 \pm 0.00$	$535 \pm 79.1$	$524 \pm 95.8$	$35167 \pm 4967$	$1.52 \pm 0.35$	$36228 \pm 4969$	47642	$46581 \pm 251$
$\Delta \phi_{1}$ wreak > 1	$2.35 \pm 0.47$	$180 \pm 26.0$	$181 \pm 31.3$	$776 \pm 671$	$1.02 \pm 0.00$ $0.08 \pm 0.08$	$1146 \pm 672$	1072	$1602 \pm 60.2$
$E^{miss} > 150 \text{ GeV}$	$0.82 \pm 0.28$	$56.4 \pm 7.71$	$45.9 \pm 8.32$	110 ± 011	0.00 ± 0.00	$102 \pm 113$	112	$9.70 \pm 15.5$
	0.02 ± 0.20	0011 ± 1111	revers	e-Λφ.: 3-iet ( C - V	/B2)	102 ± 1110		0110 ± 1010
let nm	$536 \pm 391$	$55768 \pm 215$	99312 + 387	$6572572 \pm 582900$	$\frac{2137 + 10.3}{2137 + 10.3}$	6729789 + 582900	4667099	$4509882 \pm 2205$
Opp Hemispheres	$267 \pm 24.9$	$26854 \pm 183$	$43023 \pm 306$	$2703971 \pm 371642$	$788 \pm 6.23$	$2774636 \pm 371642$	2132626	$2061961 \pm 1503$
$\Delta n = 54.8$	$2182 \pm 421$	$324 \pm 15.8$	$659 \pm 35.6$	$86399 \pm 73404$	$10.0 \pm 0.20$ $10.0 \pm 0.71$	87392 + 73404	33837	$32844 \pm 188$
$m_{\odot} > 1$ TeV	$20.92 \pm 4.20$	$284 \pm 14.7$	$563 \pm 31.3$	$86399 \pm 73404$	$8.52 \pm 0.66$	$87255 \pm 73404$	27920	$27064 \pm 171$
Central Jet Veto	$10.15 \pm 4.09$	$154 \pm 12.3$	$315 \pm 28.3$	$73906 \pm 72429$	$7.53 \pm 0.62$	74382 + 72429	13698	$13222 \pm 121$
$\Delta \phi = \mu r \pi > 1$	$1.39 \pm 0.35$	$374 \pm 456$	$59.4 \pm 7.92$	-	$1.38 \pm 0.02$ $1.38 \pm 0.27$	$982 \pm 914$	700	$602 \pm 28.0$
$E_{miss}^{miss} > 150 \text{ GeV}$	$0.39 \pm 0.03$	$4.27 \pm 0.83$	$3.00 \pm 0.71$	-	$0.14 \pm 0.08$	$741 \pm 110$	11	$359 \pm 349$
	0.00 ± 0.10	1.21 ± 0.00	reverse- $\Delta \phi_i$	3-iet Small Δn.; (	C - VB3)	1111 ± 1110		0.00 ± 0.10
Jet $p_T$	$536 \pm 39.1$	$55768 \pm 215$	$99312 \pm 387$	$\frac{1}{6572572 + 582900}$	$2137 \pm 10.3$	$6729789 \pm 582900$	4667099	$4509882 \pm 2205$
Opp Hemispheres	$267 \pm 24.9$	$26854 \pm 183$	$43023 \pm 306$	$2703971 \pm 371642$	$788 \pm 6.23$	$2774636 \pm 371642$	2132626	$2061961 \pm 1503$
$\Delta n_{\rm eff} < 3.8$	$198 \pm 22.9$	$25020 \pm 152$	$39679 \pm 251$	$2421614 \pm 352716$	$741 \pm 6.04$	$2487054 \pm 352716$	1944372	$1878932 \pm 1425$
$m_{ii} > 1 \text{ TeV}$	$7.36 \pm 3.29$	$1088 \pm 149$	$1605 \pm 236$	$146191 \pm 10049$	$25.2 \pm 1.09$	$148909 \pm 10053$	130144	$127426 \pm 456$
Central Jet Veto	$7.36 \pm 3.29$	$1088 \pm 149$	$1605 \pm 236$	$146191 \pm 10049$	$25.2 \pm 1.09$ $25.2 \pm 1.09$	$148909 \pm 10053$	130144	$127426 \pm 456$
$\Delta \phi_{in MET} > 1$	$1.12 \pm 0.20$ $1.12 \pm 0.32$	$202.2 \pm 25.0$	$219 \pm 30.4$	$981 \pm 633$	$1.58 \pm 0.27$	$1404 \pm 634$	2174	$1751 \pm 61.0$
$E_{miss}^{miss} > 150 \text{ GeV}$	$0.34 \pm 0.17$	$85.9 \pm 10.6$	$62.6 \pm 9.05$	-	$0.70 \pm 0.18$	$149 \pm 13.9$	161	$11.8 \pm 18.8$

Table 8.2: Progressive yields of the reverse- $\Delta \phi_{j,j}$  control regions (**C** selections) described in Section 8.1 in 20.3fb<sup>-1</sup> of 2012 data, as evaluated using Monte Carlo. Expected signal yields are shown for the case of  $m_H = 125$  GeV. The "Subtracted" column shows an estimate of the dijet contribution by subtracting Z, W and Other Background columns from the observed yield.

extrapolation in VR1 (  $\mathbf{B} \rightarrow \mathbf{A}$  ) is consistent with a systematic set from the difference of VR3 (  $\mathbf{B} \rightarrow \mathbf{A}$  ), i.e. adding a third jet should not change the efficiency significantly.

- 2.  $m_{jj}$ : Same as  $\Delta \eta_{j,j}$ , the difference of efficiencies in VR2 (  $\mathbf{B} \rightarrow \mathbf{A}$  ) is used to set a systematic of 2%.
- 3.  $\Delta \phi_{jj}$ : Same as  $\Delta \eta_{j,j}$ , the difference of efficiencies in VR2 (  $\mathbf{B} \rightarrow \mathbf{A}$  ) is used to set a systematic of 4%.
- 4. Third Jet Veto: **B**-Signal Region is also used to determine the efficiency, however the systematic cannot be taken from VR3 since a third jet is required there. Instead, the systematic is taken from the difference of efficiencies in VR2 ( $\mathbf{B} \rightarrow \mathbf{A}$ ) and set at 60%.
- 5.  $\Delta \phi_{jx,E_{\mathrm{T}}^{\mathrm{miss}}}$ : Selection **C**-Signal Region is used to determine the efficiency. All 3 VR efficiency extrapolations (**C**  $\rightarrow$  **A**) differ from 30%-90%. As a result, the largest 90% is taken as a systematic to remain conservative.

	$\mathbf{A} \operatorname{SR} \Delta \mathbf{c}$	$\phi_{jx,MET}$ and $\Delta \phi_{jj}$					
Cut	Signal Region	VR1: Small- $\Delta \eta_{jj}$	VR2: 3-jet	VR3: 3-jet, Small- $\Delta \eta_{jj}$			
$\Delta \eta_{jj}$	(blinded)	$0.8740\pm0.0009$	$0.0312 \pm 0.0001$	$0.8726 \pm 0.0010$			
$m_{jj}$	(blinded)	$0.0552\pm0.0002$	$0.8076 \pm 0.0053$	$0.0607 \pm 0.0002$			
$\Delta \phi_{jj}$	(blinded)	$0.0387 \pm 0.0006$	$0.5045\pm0.0043$	$0.0482 \pm 0.0008$			
Central Jet Veto or $p_T^{j3} > 40 \text{ GeV}$	(blinded)	$0.0024\pm0.0045$	$0.7080\pm0.0076$	N/A			
$\Delta \phi_{jx,MET}$	(blinded)	$0.3837\pm1.9267$	$0.0683 \pm 0.0027$	$0.0092 \pm 0.0043$			
$E_T^{miss}$	(blinded)	-	$0.0266 \pm 0.0147$	$0.2131 \pm 0.4219$			
<b>B</b> reverse- $\Delta \phi_{ix,MET}$							
Cut	Signal Region	VR1: Small- $\Delta \eta_{jj}$	VR2: 3-jet	VR3: 3-jet, Small- $\Delta \eta_{jj}$			
$\Delta \eta_{jj}$	$0.0227 \pm 0.0001$	$0.8963 \pm 0.0008$	$0.0249 \ \pm 0.0001$	$0.8937 \pm 0.0009$			
$m_{jj}$	$0.7964 \pm 0.0047$	$0.0561\pm0.0002$	$0.8193 \pm 0.0052$	$0.0618\pm0.0002$			
$\Delta \phi_{jj}$	$0.4109 \pm 0.0033$	$0.0402\pm0.0005$	$0.4878 \pm 0.0040$	$0.0493 \pm 0.0007$			
Central Jet Veto or $p_T^{j3} > 40 \text{ GeV}$	$0.0239 \pm 0.0011$	$0.0007 \pm 0.0005$	$0.7932 \pm 0.0082$	N/A			
$E_T^{miss}$	-	$0.5752 \pm 0.7703$	$0.0031 \pm 0.0009$	$0.0206\pm0.0043$			
	<b>C</b> reversed- $\Delta \phi_{ij}$						
Cut	Signal Region	VR1: Small- $\Delta \eta_{jj}$	VR2: 3-jet	VR3: 3-jet, Small- $\Delta \eta_{jj}$			
$\Delta \phi_{jx,MET}$	$0.1443 \pm 0.0013$	$0.0344 \pm 0.0013$	$0.0455 \pm 0.0022$	$0.0137 \pm 0.0005$			
$E_T^{miss}$	$0.0007 \pm 0.0010$	$0.0061\pm0.0097$	$0.0060\pm0.0058$	$0.0067 \pm 0.0107$			

Table 8.3: Efficiencies in the signal region and multijet enhanced regions, as estimated from the "Subtracted" columns of Table 5.2, Table 8.1 and Table 8.2. In **blue** are the efficiencies used to estimate the multijet contribution in the signal region. In green are the efficiencies used to estimate the multijet contribution in  $\mathbf{A}$  -VR2 with the same efficiency used for  $E_{\rm T}^{\rm miss}$  as the signal region. A "-" denotes insufficient statistics to evaluate a relative efficiency or upper bound.

6.  $E_{\rm T}^{\rm miss}$ : Since this cut has the strongest effect to reduce the multijet contribution, it is difficult to estimate an efficiency. As a result, the most statistically significant efficiency is used from **A**-VR2. This is justified since it is expected that the  $E_{\rm T}^{\rm miss}$  distribution is not heavily affected by the third jet veto. Since no other region is used, only the statistical error from **A**-VR2 is used and no additional systematic is set.

An efficiency method is used to estimate the multijet contribution in **A** -VR2. Similarly, the efficiencies for  $\Delta \eta_{j,j}$ ,  $m_{jj}$ ,  $\Delta \phi_{j,j}$  and the Third Jet Veto are taken from **B** -VR2 instead of **B** -Signal Region. The efficiency of  $\Delta \phi_{jx,E_T^{\text{miss}}}$  is taken from **C** -VR2 and **A** -VR2 is used for the  $E_T^{\text{miss}}$  efficiency (just like the signal region estimate). These efficiencies are applied progressively to the "Subtracted" data yields after the opposite hemispheres requirement, taken from Table 5.2. The data yields after the opposite hemisphere requirement are composed of almost entirely multijet events. The results are shown in Table 8.4.

In either case the error is > 100% which is expected from the large systematics determined from the validation regions. In the case of the **A** -VR2 estimate, the estimate of 19.9  $\pm$  21.8 agrees very well with the "subtracted" data yield after all cuts of 22.5  $\pm$  16.3. In the case of the signal region, the multijet background is negligible compared to the other backgrounds.

	Relative	Efficiency	Yield		
Cut	Signal Region	3-jet Region	Signal Region	3-jet Region	
Opp. Hemispheres	-	-	$2836298 \pm 1902$	$2086759 \pm 1657$	
$\Delta \eta_{jj}$	$0.0227 \pm 0.0057$	$0.0249\pm0.0062$	$64384 \pm 16167$	$51960 \pm 12938$	
$m_{jj} > 1 \text{ TeV}$	$0.7964 \pm 0.0166$	$0.8193 \pm 0.0172$	$51275 \pm 12920$	$42571 \pm 10638$	
$\Delta \phi_{jj} < 2.5$	$0.4109 \pm 0.0167$	$0.4878 \pm 0.0199$	$21069 \pm 5377$	$20766 \pm 5258$	
Central Jet Veto	$0.0239 \pm 0.0143$	$0.7932\pm0.0797$	$504\pm328$	$16472 \pm 4487$	
$\Delta \phi_{jx,MET} > 1$	$0.1443 \pm 0.1300$	$0.0455\pm0.0410$	$72.7\pm80.7$	$749\pm706$	
$E_T^{miss} > 150 \text{ GeV}$	$0.0266 \pm 0.0147$	$0.0266\pm0.0147$	$1.93\pm2.40$	$19.9\pm21.8$	

Table 8.4: Estimates of the multijet background in the signal region and in  $\mathbf{A}$  -VR2 (3-jet) using the technique described in Section 8.2.



Figure 8.3:  $p_T$  distribution of the jet closest to the  $E_T^{\text{miss}}$  used for extrapolation below the  $p_T$  threshold of 30 GeV.

#### 8.3 Jet $p_T$ Extrapolation Method

A secondary method is used to cross check the results from the Efficiency Method described in Section 8.2, known as the Jet  $p_T$  Extrapolation Method. A control region is constructed in which a third jet is required and the  $\Delta \phi_{jx,E_T^{miss}}$  requirement is reversed. This forms a sample of events which is enhanced in mismeasured jets from multijet production. The third jet  $p_T$  distribution is then extrapolated below the jet  $p_T$  threshold of 30 GeV, a region that emulates events that can pass the signal region, shown in Figure 8.3. Unfortunately, the low statistics below the jet  $p_T$  threshold of 30 GeV cannot result in a meaningful fit.

Since the  $E_{\rm T}^{\rm miss}$  requirement heavily suppresses the multijet requirement, the  $E_{\rm T}^{\rm miss}$  requirement



Figure 8.4: Yields of the jet  $p_T$  extrapolation method as a function of the  $E_T^{\text{miss}}$  requirement fitted to an exponential.

is relaxed and events are corrected by the trigger efficiency. The yields are then extrapolated to the signal region requirement of  $E_{\rm T}^{\rm miss} > 150$  GeV, shown in Figure 8.4.

The fit results in effectively 0 events with  $E_{\rm T}^{\rm miss} > 150$  GeV which is equivalent to the efficiency method. To remain conservative, the efficiency method is used as the primary estimate of the multijet background.

### CHAPTER 9

# Systematic Uncertainties

#### 9.1 Introduction

Systematic uncertainties are divided into two categories: theoretical and experimental. Theoretical uncertainties concern the underlying cross section calculation while experimental uncertainties arise from those related to the calibration of the detector.

#### 9.2 Theoretical Uncertainties on Signal

The following theoretical uncertainties are considered for signal (VBF and ggF) processes:

- 1. Since cross sections can only be calculated at finite orders, predictions are parametrized in terms of renormalization ( $\mu_R$ ) and factorization scales ( $\mu_F$ ) which are unphysical. In order to estimate the perturbative uncertainty on a cross section,  $\mu_R$  and  $\mu_F$  are varied around the central scale value by a factor of 2, i.e.  $Q/2 < \mu_R, \mu_F < 2Q$ . The uncertainty of the PDF+ $\alpha_S$  used for the cross section calculation is also considered. Uncertainties due to the variation of scales and PDF+ $\alpha_S$  are shown as a function of  $m_H$  in Table 9.1.
- 2. Shape of the Higgs  $p_T$  spectrum for ggF and VBF
  - a) The ggF signal process is normalized to include NNLO+NNLL effects as well as the effect of quark masses in the loops using a matrix element calculator, HRes [46, 47]. The uncertainty of the reweighting is calculated using the ggF\_cross\_section\_uncertainty tool using the resummation improved Stewart Tackman (RIST) method [48]. The yields and corresponding uncertainties are shown in Table 9.2.

$m_H[\text{GeV}]$	$\sigma(ggF)[{\rm pb}]$	$\sigma(VBF)[pb]$	Number of Generated Events (Each Process)
115	$22.66^{+7.4\%}_{-8.1\%}{}^{+7.6\%}_{-6.8\%}$	$1.729^{+0.2\%}_{-0.2\%}{}^{+2.5\%}_{-2.8\%}$	100000
120	$20.86^{+7.3\%}_{-7.9\%}{}^{+7.5\%}_{-6.9\%}$	$1.649^{+0.2\%}_{-0.2\%}{}^{+2.6\%}_{-2.8\%}$	100000
125	$19.27^{+7.2\%}_{-7.8\%}{}^{+7.5\%}_{-6.9\%}$	$1.578^{+0.2\%}_{-0.2\%}{}^{+2.6\%}_{-2.8\%}$	500000
130	$17.85^{+7.1\%}_{-7.7\%}{}^{+7.5\%}_{-6.9\%}$	$1.511^{+0.2\%}_{-0.2\%}{}^{+2.6\%}_{-2.7\%}$	100000
150	$13.55^{+6.7\%}_{-7.4\%}{}^{+7.4\%}_{-7.0\%}$	$1.280^{+0.3\%}_{-0.2\%}{}^{+2.5\%}_{-2.7\%}$	100000
200	$7.081^{+6.0\%}_{-6.8\%}{}^{+7.4\%}_{-7.7\%}$	$0.8685^{+0.3\%}_{-0.1\%}{}^{+2.5\%}_{-2.7\%}$	100000
300	$3.594^{+5.7\%}_{-6.1\%}{}^{+7.7\%}_{-7.9\%}$	$0.4408^{+0.3\%}_{-0.2\%}{}^{+2.5\%}_{-2.6\%}$	100000

Table 9.1: Cross sections on ggF and VBF signal processes for 115 GeV  $< m_H < 300$  GeV. Uncertainties include variations of factorization and renormalization scales as well as PDF+ $\alpha_s$ .

Cuts	Nominal yield	Yield uncertainty
$p_T^{j1,j2}$	2814.86	278.36
Opp. Hemispheres	144.76	14.45
$\Delta \eta_{jj} > 4.8$	143.73	14.35
$m_{jj} > 1 \text{ TeV}$	60.33	5.94
$\Delta \phi_{jj} < 2.5$	59.39	5.85
Third Jet Veto	49.59	4.9
$\Delta \phi_{j, E_{\mathrm{T}}^{\mathrm{miss}}} > 1$	47	4.65
$E_{\rm T}^{\rm miss} > 150 {\rm ~GeV}$	21.7	2.1

Table 9.2: Yields and corresponding uncertainties of the ggF signal process after reweighting the Higgs  $p_T$  according to HRes. The uncertainty is calculated using the gF\_cross\_section\_uncertainty tool with the RIST method.

- b) The VBF signal process is normalized to include electroweak radiative corrections using a matrix element calculator, HAWK. The Monte Carlo statistical error on the  $p_T$  distribution produced by HAWK is taken into account as an uncertainty, shown in Figure 5.7.
- 3. An uncertainty is assessed on the third jet veto by varying the factorization and normalization scales using MCFM and additionally varying the parton showering using Powheg+Pythia and Powheg+Herwig samples, detailed in Section 9.2.1.

Scale Setting	$\sigma_{2 \text{ jet}} [ab]$	$\sigma_{\geq 2 \text{ jet}} [ab]$	$\sigma_{>2 \text{ jet}} [ab]$
$\mu_r = \mu_f = 1/2$	$55.4 {\pm} 0.2$	$65.1 {\pm} 0.2$	$9.7 {\pm} 0.3$
$\mu_r = 1/2, \mu_f = 1$	$58.0 {\pm} 0.2$	$65.8 {\pm} 0.2$	$7.9 {\pm} 0.3$
$\mu_r = 1, \mu_f = 1/2$	$57.9 {\pm} 0.1$	$66.3 {\pm} 0.2$	$8.5 {\pm} 0.2$
$\mu_r = \mu_f = 1$	$58.8 {\pm} 0.1$	$66.1 {\pm} 0.2$	$7.4 {\pm} 0.2$
$\mu_r = 1, \mu_f = 2$	$59.2 {\pm} 0.2$	$65.8 {\pm} 0.1$	$6.6 {\pm} 0.2$
$\mu_r = 2, \mu_f = 1$	$59.6 {\pm} 0.2$	$65.9 {\pm} 0.1$	$6.3 {\pm} 0.2$
$\mu_r = \mu_f = 2$	$59.4 {\pm} 0.1$	$65.1{\pm}0.1$	$5.7{\pm}0.2$

Table 9.3: The cross sections for the VBF signal process using MCFM and varying the factorization and renormalization scales by a factor of 2. Shown uncertainties are statistical.

#### 9.2.1 Signal Uncertainty due to Third Jet Veto

An uncertainty is assessed on the VBF signal process using MCFM and varying the factorization and renormalization scales. This is done in a fiducial region with the following selection:

- 1. Jets with  $p_T > 30 \text{ GeV}$
- 2.  $\Delta \eta_{j,j} > 4.8$
- 3.  $\eta_{j1} \times \eta_{j2} < 0$
- 4.  $m_{jj} > 1 \text{ TeV}$
- 5.  $E_{\mathrm{T}}^{\mathrm{miss}} > 150 \ \mathrm{GeV}$

The resulting cross sections from varying the scales up and down by a factor of 2 are shown in Table 9.3.

The uncertainty on  $\sigma_{2 \text{ jet}}$  is calculated using the Stewart-Tackmann method which assigns an uncertainty assuming that  $\sigma_{N \text{ jet}}$  and  $\sigma_{N+1 \text{ jet}}$  are uncorrelated. This is done to avoid an accidental cancellation in uncertainties due to scale variation. The formula follows:

$$\Delta \sigma_{N \text{ jet}}^2 = \Delta \sigma_{\geq N \text{ jet}}^2 + \Delta \sigma_{\geq N+1 \text{ jet}}^2$$
(9.1)

Setting N = 2 in the case of the VBF analysis:

$$\Delta \sigma_{2\,\text{jet}}^2 = \Delta \sigma_{\geq 2\,\text{jet}}^2 + \Delta \sigma_{>2\,\text{jet}}^2 \tag{9.2}$$

The value of  $\Delta \sigma$  is calculated using the values in Table 9.3 and calculating the maximum deviations from nominal ( $\mu_f = \mu_r = 1$ ) for  $\sigma_{2 \text{ jet}}$ ,  $\sigma_{\geq 2 \text{ jet}}$  and  $\sigma_{>2 \text{ jet}}$  and symmetrizing.

Sample	$\epsilon_{\text{third-jet-veto}}$ (%)
Powheg+Pythia	$2.141 \pm 0.015$
Powheg+Herwig	$2.104\pm0.014$

Table 9.4: The efficiency of the third jet veto by varying the parton shower modeling using Powheg+Pythia and Powheg+Herwig. The uncertainties shown are statistical.

$$\sigma_{>2 \text{ jet}} = 66.1 \pm 1.0 \text{ ab}$$
 (9.3)

$$\sigma_{>2 \, \text{jet}} = 7.4 \pm 2.4 \text{ ab}$$
 (9.4)

Using these values in Equation 9.2 yields an uncertainty of  $\pm 2.6$  ab which corresponds to 4.4% relative to  $\sigma_{2 jet}$ 

Additionally, an uncertainty is evaluated on the third jet veto by varying the parton shower modeling. This is done by using 2 samples, both with signal VBF events generated by Powheg but with the parton showering done separately by Pythia and Herwig. The full selection is applied at truth level (no reconstruction effects are applied) and the difference of the third jet veto efficiency is taken as a uncertainty. The resulting third jet veto efficiencies are shown in Table 9.4.

The difference of the third jet veto efficiencies is negligible and therefore ignored.

#### 9.3 Theoretical Uncertainties on Backgrounds

#### 9.3.1 Variation of Factorization and Renormalization Scales

Similar to the treatment of the signal, uncertainties are assessed due to the variation of factorization and renormalization scales up and down by a factor of 2. The scales are varied coherently as well as independently and the envelope of the results is taken as the uncertainty. Uncertainties are assessed on V+jets processes in the signal region and control regions. The  $W \rightarrow \ell \nu$ +jets process in the signal region is emulated by requiring that the lepton be either subthreshold or out of the  $\eta$  acceptance of the detector. Since other backgrounds in the signal and control regions are negligible in comparison, uncertainties are not assessed.

Scale uncertainties are calculated using 2 matrix element calculators: MCFM and VBFNLO for strong and weak production, respectively. Since an uncertainty is evaluated on the total V+jets process in either the signal region or control region, an uncertainty arising from MCFM or VBFNLO must be weighted

Strong Variations				
Scale	$Z \rightarrow \nu \nu + \text{jets}$	$W \to \ell \nu + \text{jets Signal Region}$	$W \to \ell \nu + \text{jets Control Region}$	
$\Delta(\mu_F = 2, \mu_R = 2)$	54.6%	50.7%	51.1%	
$\Delta(\mu_F = 1/2, \mu_R = 1/2)$	-31.4%	-31.5%	-30.1%	
$\Delta(\mu_F = 2, \mu_R = 1)$	22.5%	24.5%	20.9%	
$\Delta(\mu_F = 1/2, \mu_R = 1)$	-15.5%	-17.9%	-15.1%	
$\Delta(\mu_F = 1, \mu_R = 2)$	19.2%	17.4%	16.5%	
$\Delta(\mu_F = 1, \mu_R = 1/2)$	-12.6%	-13.1%	-13.0%	
Weight $(f_{\text{strong}})$	0.85	0.89	0.80	
	-	Weak Variations		
$\Delta(\mu_F = 2, \mu_R = 2)$	-6.0%	-3.6%	-3.5%	
$\Delta(\mu_F = 1/2, \mu_R = 1/2)$	-0.7%	-0.9%	-0.5%	
$\Delta(\mu_F = 2, \mu_R = 1)$	0.7%	0.9%	-0.5%	
$\Delta(\mu_F = 1/2, \mu_R = 1)$	-0.7%	0.0%	0.0%	
$\Delta(\mu_F = 1, \mu_R = 2)$	-2.7%	0.9%	-1.0%	
$\Delta(\mu_F = 1, \mu_R = 1/2)$	-0.7%	0.9%	1.0%	
Weight $(f_{\text{weak}})$	0.15	0.11	0.20	

Table 9.5: Summary of scale variations and weighting factors used to calculate uncertainties.

by the relative contribution of strong and weak cross sections. The weight is shown in Equation 9.5.

$$f_{\text{weak/strong}} = \frac{N_{\text{weak/strong}}}{N_{\text{weak}} + N_{\text{strong}}}$$
(9.5)

An uncertainty due to a scale variation is calculated in the following way:

$$\Delta(\mu_F, \mu_R) = \frac{\sigma(\mu_F, \mu_R) - \sigma(1, 1)}{\sigma(1, 1)} f$$
(9.6)

where  $\Delta(\mu_F, \mu_R)$  is the uncertainty with a given scale setting,  $\sigma(\mu_F, \mu_R)$  is the cross section from either MCFM or VBFNLO at a given scale setting and  $\sigma(1, 1)$  is a cross section from MCFM or VBFNLO with the nominal scale setting. A summary of the scale variations and weighting factors considered are shown in Table 9.5.

The following fiducial cuts are applied to events generated with MCFM and VBFNLO:

- $Z \rightarrow \nu \nu + \mathbf{jets}$ 
  - 1. Leading jet  $p_T > 75 \text{ GeV}$
  - 2. Subleading jet  $p_T > 50 \text{ GeV}$
  - 3.  $E_{\rm T}^{\rm miss}$  > 150 GeV
  - 4.  $\Delta \eta_{j,j} > 4.8$
  - 5.  $m_{jj} > 1 \text{ TeV}$

- $W \to \ell \nu + jets$  Signal Region
  - 1. Leading jet  $p_T > 75 \text{ GeV}$
  - 2. Subleading jet  $p_T > 50 \text{ GeV}$
  - 3. Lepton  $\eta > 2.4$  or Lepton  $p_T < 10$  GeV
  - 4.  $\ell p_T + E_T^{\text{miss}} > 150 \text{ GeV}$
  - 5.  $\Delta \eta_{j,j} > 4.8$
  - 6.  $m_{jj} > 1 \text{ TeV}$
- $W \rightarrow \ell \nu + jets$  Control Region
  - 1. Leading jet  $p_T > 75 \text{ GeV}$
  - 2. Subleading jet  $p_T > 50 \text{ GeV}$
  - 3. Lepton  $p_T > 30 \text{ GeV}$
  - 4.  $\ell p_T + E_T^{\text{miss}} > 150 \text{ GeV}$
  - 5.  $\Delta \eta_{j,j} > 4.8$
  - 6.  $m_{jj} > 1 \text{ TeV}$

Because there is no difference in the kinematics of  $Z \to \nu\nu$ +jets and  $Z \to \ell\ell$ +jets, the calculated uncertainties are applied to both. For each of the 3 processes defined above, several variations of the factorization ( $\mu_F$ ) and renormalization scales ( $\mu_R$ ) are calculated:

- 1.  $\mu_F = \mu_R = 1$  (Nominal)
- 2.  $\mu_F = \mu_R = 2$
- 3.  $\mu_F = \mu_R = 0.5$
- 4.  $\mu_F = 2, \, \mu_R = 1$
- 5.  $\mu_F = 0.5, \, \mu_R = 1$
- 6.  $\mu_F = 1, \ \mu_R = 2$
- 7.  $\mu_F = 1, \, \mu_R = 0.5$

Process	Weak Va	ariation	Strong V	ariation
$Z  ightarrow  u  u + \mathbf{jets}$	+0.1%	-0.9%	+46.4%	-26.7%
$W \rightarrow \ell \nu + \mathbf{jets}$ Signal Region	+0.1%	-0.4%	+45.1%	-28.0%
$W \rightarrow \ell \nu + \mathbf{jets}$ Control Region	+0.2%	-0.7%	+40.9%	-24.1%

Table 9.6: Summary of uncertainties due to scale variations weighted by the relative production of strong and weak production.

Process	Strong Variation
$Z \rightarrow \nu \nu + \mathbf{jets}$	+3.9% $-3.4%$
$W \rightarrow \ell \nu + \mathbf{jets}$	+3.0% $-5.1%$

Table 9.7: Summary of the effect of the scale variations on the V+jets background estimates in the signal region using the effective background model described in Section 6.4.

An envelope of the largest deviations from nominal in Table 9.5 is used to determine the uncertainties that enter the fit. Table 9.6 summarizes the envelope of uncertainties weighted by  $f_{\text{weak/strong}}$ .

Even though the uncertainties are relatively high for the strong variations, they cancel to a large degree in the effective background model shown in Section 6.4. For example, considering the  $Z \rightarrow \nu\nu$ +jets estimate based on the  $W \rightarrow \ell\nu$ +jets control region:

$$N_{SR,\,\text{estimate}}^{Z \to \nu\nu} = \frac{N_{SR,MC}^{Z \to \nu\nu}}{N_{CR,MC}^{W \to \ell\nu}} \quad N_{CR,Data}^{W \to \ell\nu} \tag{9.7}$$

It is seen that the effect of variation the strong production in the positive direction results in a much smaller error:

$$\Delta(N_{SR,\,\text{estimate}}^{Z \to \nu\nu}) = \frac{+46.4\%}{+40.9\%} = +3.9\%$$
(9.8)

The complete summary of the effect of the strong scale variations (weak scale variations are negligible) on the  $Z \rightarrow \nu\nu$ +jets and  $W \rightarrow \ell\nu$ +jets estimates in the signal region is presented in Table 9.7.

#### **9.3.2 PDF**+ $\alpha_S$

The uncertainties due to the CT10 Parton Distribution Function (PDF) used to simulate events are also computed on the V+jets processes. This was done by using the PDFReweight tool on truth level (no reconstruction effects applied) events generated with Sherpa (the same generator used for the fully

Process	Error due	e to PDF
$Z \rightarrow \nu \nu + \mathbf{jets}$	+3.32%	-2.59%
$W \rightarrow \ell \nu + \mathbf{jets}$ Signal Region	+5.46%	-3.29%
$W \rightarrow \ell \nu + \mathbf{jets}$ Control Region	+3.59%	-2.59%

Table 9.8: Summary of uncertainties due to the CT10 PDF used to simulate events.

simulated sample). Events are reweighted according to a given error eigenvector of the CT10 PDF and the acceptance of the analysis is recalculated. The difference of acceptance to the nominal acceptance is combined in quadrature for each error eigenvector to result in a 68% confidence level uncertainty envelope as according to the prescription from the PDF4LHC Working Group Recommendations [38], shown below:

$$\Delta(V \to +\text{jets}) = \sqrt{\sum_{i=1}^{N} (X - X_i)^2}$$
(9.9)

where i runs over the set of N error eigenvectors. The resulting uncertainties are shown in Table 9.8.

Additionally, errors on the ratio can be computed by doing the eigenvector error calculation coherently, for example with the  $Z \rightarrow \nu\nu$ +jets estimate following Equation 9.11.

$$\Delta(N_{est}^{Z \to \nu\nu}) = \sqrt{\sum_{i=1}^{N} \left(\frac{N_{SR,MC}^{Z \to \nu\nu}}{N_{CR,MC}^{W \to \ell\nu}} - \frac{N_{SR,MC,i}^{Z \to \nu\nu}}{N_{CR,MC,i}^{W \to \ell\nu}}\right)^2}$$
(9.10)

The  $W \to \ell \nu + jets$  estimate is similar:

$$\Delta(N_{est}^{W\to\ell\nu}) = \sqrt{\sum_{i=1}^{N} \left(\frac{N_{SR,MC}^{W\to\ell\nu}}{N_{CR,MC}^{W\to\ell\nu}} - \frac{N_{SR,MC,i}^{W\to\ell\nu}}{N_{CR,MC,i}^{W\to\ell\nu}}\right)^2}$$
(9.11)

This results in an error of  $\pm 1.7\%$  for the  $Z \rightarrow \nu\nu$ +jets estimate and  $\pm 2.3\%$  for the  $W \rightarrow \ell\nu$ +jets estimate.

#### 9.3.3 Shower Modeling

In order to account for possible differences in the parton showering of jets in  $Z \rightarrow \nu\nu$ +jets and  $W \rightarrow \ell\nu$ +jets events, an uncertainty is set by varying the shower model and recalculating the acceptance of the analysis. This is especially important since the  $W \rightarrow \ell\nu$ +jets process in the  $W \rightarrow \ell\nu$ +jets Control Region is used to effectively model the  $Z \rightarrow \nu\nu$ +jets process in the signal region.

The parton shower model is varied by changing the CSS\_KIN\_SCHEME parameter in Sherpa from 0 to 1. This corresponds to changing the recoil strategy for dipoles with the initial state emitter and final state spectator.

The parton shower varied samples have several important limitations. The only samples available are  $W \to \mu\nu$ +jets and  $Z \to \mu\mu$ +jets. This is sufficient because the  $E_{\rm T}^{\rm miss}$  of the signal region can be simply modeled using the  $p_T$  of the W or Z. Since the study is done with truth level events the  $p_T$  of the W or Z is equivalent to  $E_{\rm T}^{\rm miss}$  in a  $W \to \ell\nu$ +jets event with a lost lepton or  $Z \to \nu\nu$ +jets events. An additional limitation of the parton shower varied samples is that they lack sufficient statistics to calculate the full acceptance of the analysis, several cuts are relaxed and/or removed. Three categories of selections are considered:

#### 1. 2-jet + VBF

- $Vp_T > 150 \text{ GeV}$
- Leading jet  $p_T > 75$  GeV and  $|\eta| < 4.5$
- Subleading jet  $p_T > 50$  GeV and  $|\eta| < 4.5$
- $\Delta \phi_{j,j} < 2.5$
- $\Delta \phi_{j1,E_{\pi}^{\text{miss}}} > 1.6$
- $\Delta \phi_{i2, E_{T}^{\text{miss}}} > 1.0$
- **VBF**  $m_{jj} > 500 \text{ GeV}, \eta_{j1} \times \eta_{j2} < 0, \Delta \eta_{j,j} > 3.0$

#### 2. 2-jet + Third Jet Veto

- $Vp_T > 150 \text{ GeV}$
- Leading jet  $p_T > 75$  GeV and  $|\eta| < 4.5$
- Subleading jet  $p_T > 50$  GeV and  $|\eta| < 4.5$
- $\Delta \phi_{j,j} < 2.5$
- $\Delta \phi_{j1,E_{\mathrm{T}}^{\mathrm{miss}}} > 1.6$
- $\Delta \phi_{j2,E_{\mathrm{T}}^{\mathrm{miss}}} > 1.0$
- Third jet veto for jets with  $p_T > 30$  GeV and  $|\eta| < 4.5$

#### 3. Full Selection

•  $Vp_T > 150 \text{ GeV}$ 

Shower Model	Total Events	2-jet + VBF	2-jet + Third Jet Veto	Full Selection
$Z+jets$ CSS_KIN_SCHEME 0	$4.42738e^{+7}$	2040	14569	1052
$Z+jets$ CSS_KIN_SCHEME 1	$6.23468e^{+6}$	236	1867	142
Shower Effect on $Z+jets$			$(9.0 \pm 0.3)\%$	
$W+jets$ CSS_KIN_SCHEME 0	$2.58952e^{+7}$	673	5108	339
$W+jets$ CSS_KIN_SCHEME 1	$6.31894e^{+6}$	133	1120	86
Shower Effect on $W$ +jets			$(10.0 \pm 0.4)\%$	
Shower Effect on $Z/W$ Ratio			1.3%	

Table 9.9: Number of events passing variants of the selections with the nominal and varied parton shower model. Yields are shown for Z+jets, W+jets and the effective background model (Z/W Ratio).

- Leading jet  $p_T > 75$  GeV and  $|\eta| < 4.5$
- Subleading jet  $p_T > 50$  GeV and  $|\eta| < 4.5$
- $\Delta \phi_{j,j} < 2.5$
- $\Delta \phi_{j1,E_{\mathrm{T}}^{\mathrm{miss}}} > 1.6$
- $\Delta \phi_{j2, E_{\tau}^{\text{miss}}} > 1.0$
- **VBF**  $m_{jj} > 1000 \text{ GeV}, \eta_{j1} \times \eta_{j2} < 0, \Delta \eta_{j,j} > 4.8$
- Third jet veto for jets with  $p_T > 30$  GeV and  $|\eta| < 4.5$

The results are presented in Table 9.9.

In order to have high statistical precision, the **2-jet** + **Third Jet Veto** region is used to set the parton showering uncertainty of  $9.0\% \pm 0.3\%$  on Z+jets (including  $Z \rightarrow \nu\nu$ +jets and  $Z \rightarrow \ell\ell$ +jets) and  $10.0\% \pm 0.4\%$  on  $W \rightarrow \ell\nu$ +jets. Also shown is the effect of the uncertainty on the effective background model of 1.3%, showing that the effect is quite small on the result.

#### 9.4 Experimental Uncertainties

A recommended set of experimental uncertainties from the ATLAS collaboration are evaluated. A few of these are:

- Electron and muon reconstruction, identification and isolation efficiencies are varied following recommendations from the e/gamma and muon performance groups. The energy scale and and resolution are also varied with recommendations from those performance groups.
- Jet energy uncertainties are studied with the MultiJetJESUncertaintyProvider tool which defines  $1\sigma$  error bands on 20 calibration constants.

Region	$Z \rightarrow \nu \nu + \text{jets}$	$W \to \ell \nu + \text{jets Signal Region}$	$Z \to \ell \ell + \text{jets}$	$W \to \ell \nu$ Control Region
Signal	2.79%	5.02%	6.39%	2.74%
Small- $\Delta \eta_{jj}$	3.35%	4.71%	5.98%	2.63%
3-jet	4.44%	5.86%	12.3%	4.14%
3-jet small- $\Delta \eta_{jj}$	2.64%	4.50%	6.08%	2.51%

Table 9.10: The statistical uncertainty for the Monte Carlo samples used in the evaluation of the experimental systematic errors.

- $E_{\rm T}^{\rm miss}$  is recalculated with the varied objects defined above using the MissingETUtility package. Additional uncertainties are considered from the calibration of soft objects such as clusters and tracks which  $E_{\rm T}^{\rm miss}$  is sensitive to.
- The JVF requirement is varied up and down and recalculating the jet acceptance of the analysis. After the full selection is applied the variation is negligible from the nominal JVF requirement, therefore this effect is ignored. This is described in detail in Section 9.4.1.

In general, a given systematic is evaluated by changing the detector calibration constants and recalculating the acceptance of the analysis. The difference to the nominal acceptance of the analysis is considered a systematic. There are 80 such systematics that are evaluated.

However, only 4 of these are considered for the calculation of the result because of the large statistical error present on the Monte Carlo samples. Only variations that exceed the statistical error of a given Monte Carlo sample are considered in the result. This is done to avoid double counting of the statistical error as an experimental uncertainty. The statistical error of the relevant samples are presented in Table 9.10 for the signal region and validation regions.

Plots showing the change of the acceptance with respect to nominal as a function of the set of experimental systematics are shown in Figure 9.1 ( $Z \rightarrow \nu\nu$ +jets with the signal region selection applied and  $Z \rightarrow \ell\ell$ +jets with the Z control region selection applied), Figure 9.2 ( $W \rightarrow \ell\nu$ +jets with the signal and W control region selections applied) and Figure 9.3 (VBF signal process with the signal region selection applied). Even though it is a small effect, the  $W \rightarrow \ell\nu$ +jets process must also have systematics in the multijet enhanced control region used to estimate the multijet contribution in the  $W \rightarrow \ell\nu$ +jets control region presented in Section 6.2.2, shown in Figure 9.4.

The figures show that the 4 relevant systematics are:

- 1. Eta\_ModellingJES:  $\eta$  dependence of the jet energy scale
- 2. FlavRespJES: Flavor dependence of the energy response
- 3. FlavCompJES: Uncertainty of the jet flavor composition



Figure 9.1: Change of acceptance with respect to the nominal as a function of the set of experimental systematics considered. Shown are the systematics for  $Z \rightarrow \nu\nu$ +jets with the signal region selection applied and  $Z \rightarrow \ell\ell$ +jets with the Z control region selection applied. Green and yellow bands indicate the statistical error on the nominal yield in the signal region and Z control region, respectively.

#### 4. NP\_Modelling1JES: A combination of jet energy scale parameters

The values of these uncertainties for VBF signal,  $Z \rightarrow \nu\nu$ +jets with the signal region selection applied,  $Z \rightarrow \ell\ell$ +jets with the Z control region selection applied,  $W \rightarrow \ell\nu$ +jets with the signal region and W control region selections applied and the effect on the effective background models expressed in Equation 6.2 are shown in Table 9.11. The table shows that even though the experimental uncertainties can be quite high, they cancel to a large degree in the effective background model preserving strong background modeling in the signal region. The values of the uncertainties for  $W \rightarrow \ell\nu$ +jets in the multijet enhanced control region described in Section 6.2.2 are presented in Table 9.12.



Figure 9.2: Change of acceptance with respect to the nominal as a function of the set of experimental systematics considered. Shown are the systematics for  $W \to \ell \nu$ +jets with the signal region selection applied and  $W \to \ell \nu$ +jets with the W control region selection applied. Green and yellow bands indicate the statistical error on the nominal yield in the signal region and W control region, respectively.

#### 9.4.1 Jet Vertex Association

An important feature of the dataset used for the analysis presented is the significant amount of pileup interactions, which are interactions from other pp collisions in the same bunch crossing. Pile-up interactions can result in jets that can potentially pass the signal region selection. Generally, these jets are suppressed by requiring that the Jet Vertex Fraction (JVF) be high enough to be consistent with coming from the primary vertex. The primary vertex is defined to be the vertex with the highest  $\sum p_T^2$  calculated from the associated tracks. The JVF is calculated by the fraction of track momentum associated with the jet that is consistent with the primary vertex. However, if a jet is outside of the tracking volume ( $|\eta| > 2.5$ ) a JVF value cannot be calculated. This is a particular concern in this analysis since  $E_T^{\text{miss}}$  has no associated vertex and the high  $\eta$  separation that is required between the two  $p_T$  jets increases the probability that at least one of them is outside of the tracking volume.



Figure 9.3: Change of acceptance with respect to the nominal as a function of the set of experimental systematics considered. Shown are the systematics for VBF signal with the signal region selection applied and  $Z \rightarrow \ell \ell \ell$ +jets with the Z control region selection applied. Green and yellow bands indicate the statistical error on the nominal yield in the signal region and Z control region, respectively.

Table 9.13 shows the fraction of events with at least 1 jet in the tracking volume and the fraction of events with no jets in the tracking volume as a function of the signal region selection. After all cuts  $\approx 1/3$  of events have no jets in tracking.

Additionally, events in the  $W \to \ell \nu + \text{jets}$  and  $Z \to \ell \ell + \text{jets}$  control regions have very high efficiencies to correctly identify the primary vertex due to the presence of leptons. As a result, a systematic is assigned to the efficiency of finding the primary vertex in the signal region using the JVFUncertaintyTool package which varies the JVF requirement up and down from the nominal value. The change in efficiency is very small, so the systematic is taken to be negligible.

	Eta_ModellingJES	FlavRespJES	FlavCompJES	NP_Modelling1JES
VBF Signal (SR) Up	$15.72 \pm 0.54$	$4.17\pm0.14$	$6.33\pm0.22$	$5.01\pm0.17$
VBF Signal (SR) Down	$-14.13 \pm 0.49$	$-4.18 \pm 0.14$	$-6.05 \pm 0.21$	$-5.87 \pm 0.20$
VBF Signal (SR) Average	$14.84 \pm 0.36$	$4.17 \pm 0.10$	$6.19\pm0.15$	$5.37 \pm 0.13$
$Z \to \nu \nu + \text{jets (SR) Up}$	$17.20 \pm 0.69$	$3.05\pm0.12$	$5.37 \pm 0.21$	$6.90\pm0.27$
$Z \to \nu \nu + \text{jets}$ (SR) Down	$-14.96 \pm 0.60$	$-5.09 \pm 0.20$	$-7.56 \pm 0.30$	$-6.30 \pm 0.25$
$Z \to \nu \nu + \text{jets (SR) Avg}$	$15.93 \pm 0.45$	$3.59\pm0.10$	$6.10\pm0.17$	$6.58 \pm 0.18$
$Z \to \ell \ell + \text{jets} (\text{ZCR}) \text{ Up}$	$25.02 \pm 2.30$	$11.74 \pm 1.06$	$12.81 \pm 1.16$	$12.21 \pm 1.11$
$Z \to \ell \ell + \text{jets} (\text{ZCR})$ Down	$-15.53 \pm 1.41$	$-3.17 \pm 0.29$	$-8.73 \pm 0.79$	$-4.04 \pm 0.36$
$Z \to \ell \ell + \text{jets} (\text{ZCR}) \text{ Avg}$	$18.12 \pm 1.20$	$3.75\pm0.28$	$10.02\pm0.65$	$4.84 \pm 0.35$
$W \to \ell \nu + \text{jets (SR) Up}$	$22.38 \pm 1.71$	$7.43 \pm 0.56$	$9.69 \pm 0.73$	$9.71 \pm 0.73$
$W \to \ell \nu + \text{jets (SR) Down}$	$-14.66 \pm 1.11$	$-4.00 \pm 0.30$	$-6.78 \pm 0.51$	$-5.39 \pm 0.40$
$W \to \ell \nu + \text{jets}$ (SR) Avg	$16.95 \pm 0.93$	$4.77\pm0.26$	$7.73 \pm 0.42$	$6.40\pm0.35$
$W \to \ell \nu + \text{jets} (WCR) \text{ Up}$	$16.01 \pm 0.62$	$2.85 \pm 0.11$	$3.77\pm0.15$	$3.86 \pm 0.15$
$W \to \ell \nu + \text{jets} (WCR)$ Down	$-13.98 \pm 0.54$	$-4.84 \pm 0.19$	$-6.53 \pm 0.25$	$-5.85 \pm 0.23$
$W \to \ell \nu + \text{jets} (WCR) \text{ Avg}$	$14.86 \pm 0.41$	$3.36\pm0.09$	$4.46\pm0.13$	$4.46\pm0.12$
Effe	ective Background Mo	odels		
$Z \to \nu \nu + \text{jets (SR)}$ using $Z \to \ell \ell + \text{jets CR (Avg)}$	$-2.19 \pm 1.28$	$-0.16 \pm 0.30$	$-3.92 \pm 0.67$	$1.74\pm0.39$
$Z \to \nu \nu + \text{jets}$ (SR) using $W \to \ell \nu + \text{jets}$ CR (Avg)	$1.07 \pm 0.61$	$0.23 \pm 0.13$	$1.64\pm0.21$	$2.12\pm0.22$
$W \to \ell \nu + \text{jets (SR)}$ using $W \to \ell \nu + \text{jets CR (Avg)}$	$2.09 \pm 1.02$	$1.41\pm0.28$	$3.27\pm0.44$	$1.94 \pm 0.37$

Table 9.11: The relative change (in percent) of the acceptance with respect to the nominal for VBF signal, Z and W processes in the signal and control regions for the four dominant experimental uncertainties. Also shown are the effects of the experimental uncertainties on the effective background model, averaged between up and down variations.

	Eta_ModellingJES	FlavRespJES	${\tt FlavCompJES}$	NP_Modelling1JES
Up	$36.76 \pm 12.36$	$7.99\pm2.58$	$10.40 \pm 3.36$	$9.22\pm2.98$
Down	$-11.13 \pm 3.60$	$-3.32 \pm 1.07$	$-2.11 \pm 0.68$	$-2.30 \pm 0.74$
Average	$13.14 \pm 3.46$	$4.00\pm0.99$	$2.44\pm0.67$	$2.71\pm0.72$

Table 9.12: The relative change (in percent) of the acceptance with respect to the nominal for  $W \rightarrow \ell \nu$ +jets in the multijet enhanced control region used to estimate the multijet contribution in the W control region described in Section 6.2.2.

Cut	$\geq 1$ jet in $ \eta  < 2.5$	No jets in tracking
Jet $p_T$	$91.4 \pm 0.7\%$	$8.57 \pm 0.17\%$
Opposite Hemispheres	$90.1\pm0.8\%$	$9.92\pm0.20\%$
$\Delta \eta_{jj} > 4.8$	$66.5\pm1.2\%$	$33.5\pm0.75\%$
$m_{jj} > 1 \text{ TeV}$	$62.6 \pm 1.2\%$	$37.4 \pm 0.85\%$
$\Delta \phi_{jj} < 2.5$	$61.9 \pm 1.2\%$	$38.1 \pm 0.87\%$
Central Jet Veto	$59.1 \pm 1.2\%$	$40.9\pm0.97\%$
$\Delta \phi_{ix,E_{m}^{miss}}$	$58.9 \pm 1.3\%$	$41.1\pm1.03\%$
$E_{\rm T}^{\rm miss} > 150 {\rm ~GeV}$	$62.5 \pm 1.9\%$	$37.5 \pm 1.36\%$

Table 9.13: The fraction of signal MC events with at least one jet in tracking and the fraction with no jets in tracking, as a function of the signal region selection.



Figure 9.4: Change of acceptance with respect to the nominal as a function of the set of experimental systematics considered. Shown are the systematics for  $W \rightarrow \ell \nu$ +jets with the multijet enhanced region selection applied described in Section 6.2.2 (used to determine the multijet contribution in the  $W \rightarrow \ell \nu$ +jets control region). The green band indicates the statistical error on the nominal yield in the multijet enhanced region.

### Chapter 10

# Limit Setting

#### 10.1 Background Estimation

In order to set a limit on the branching fraction of the SM Higgs decaying invisibly, a maximum likelihood fit is used which implements data driven estimates for the  $W \to \ell \nu + j$ ets and  $Z \to \nu \nu + j$ ets backgrounds. The data driven estimates for both backgrounds are normalized with one global normalization factor,  $k_V$ . A value of  $k_V$  is computed that maximizes the likelihood using the  $W \to \ell \nu + j$ ets and  $Z \to \ell \ell + j$ ets control regions MC and observed yields described in Chapter 6. Shown in Equation 10.1 is the effect of  $k_V$  on the signal and control regions.

$$N_{\rm SR} = k_V \times N_{\rm MC,SR}^Z + k_V \times N_{\rm MC,SR}^W + N_{\rm multijet} + N_{\rm other}$$

$$N_{Z-\rm CR} = k_V \times N_{\rm MC,Z-\rm CR}^Z + N_{\rm other}$$

$$N_{W-\rm CR} = k_V \times N_{\rm MC,W-\rm CR}^Z + k_V \times N_{\rm MC,W-\rm CR}^W + N_{\rm other}$$
(10.1)

The procedure of using one normalization factor for V+jets is justified since the dominant systematics in the analysis come from the jet kinematics of the V+jets backgrounds which are inherently similar between Z+jets and W+jets. Systematics are evaluated on the ratio by varying renormalization and factorization scales, changing the parton showering scheme and varying the calibration of the detector.

By defining the normalization factor  $k_V$ , MC is used only to predict the ratio of the signal region background yields to the control region yields, resulting in large cancellations of systematic uncertainties. An example of this cancellation is shown in Table 9.7. Since systematic and statistical uncertainties are taken into account for the V+jets control regions, the "weight" of the  $Z \rightarrow \ell \ell + j$ ets control region in the estimation of the signal region W and Z backgrounds will be lower than the  $W \rightarrow \ell \nu + j$ ets control region. The multijet estimate is computed following the procedure detailed in Chapter 8 and other backgrounds (< 1 event) are taken from MC.

#### 10.2 Likelihood

The maximum likelihood fit is implemented in the HistFactory framework. The likelihood is defined as:

$$\mathcal{L}(\mu) = \{ \prod_{\mathrm{R}} P(N_{\mathrm{obs, R}} \mid \mu \, s_{\mathrm{exp, R}}^{\mathrm{VBF}} \\ + \mu \, s_{\mathrm{exp, R}}^{ggF} \\ + k_V \, N_{\mathrm{exp, R}}^{Z \to \nu \nu} \\ + k_V \, N_{\mathrm{exp, R}}^{W \to \ell \nu} \\ + N_{\mathrm{exp, R}}^{\mathrm{Multijet}} \\ + N_{\mathrm{exp, R}}^{\mathrm{Multijet}} \\ + N_{\mathrm{exp, R}}^{\mathrm{Other \ Backgrounds}}) \}$$
(10.2)

where R are the 3 regions considered (Signal Region,  $W \to \ell \nu$ +jets Control Region,  $Z \to \ell \ell$ +jets Control Region),  $P(N_{\text{obs, R}}|X_{\text{R}})$  is a function which represents the Poisson probability of an observed yield in a region  $N_{\text{obs, R}}$  given an expected yield  $X_{\text{R}}$ ,  $\mu$  is the signal strength parameter,  $s_{\text{exp, R}}^{\text{VBF}}$  is the expected VBF signal and  $s_{\text{exp, R}}^{ggF}$  is the expected ggF signal,  $k_V$  is the global normalization factor applied to V+jets (described in Section 10.1),  $N_{\text{exp, R}}^{Z \to \nu\nu}$  is the expected  $Z \to \nu\nu$ +jets events from MC,  $N_{\text{exp, R}}^{W \to \ell\nu}$  is the expected  $W \to \ell \nu$ +jets events from MC,  $N_{\text{exp, R}}^{\text{Multijet}}$  is the expected multijet events computed in Chapter 8 and  $N_{\text{exp, R}}^{\text{Other Backgrounds}}$  is the expected events from all the other backgrounds. Since signal and background expected yields can vary within systematic errors, the various yields are multiplied by a product of systematic nuisance parameters with index S. The systematics are included as deviations from the nominal value ( $\Delta X$ ) scaled by the nuisance parameter determined by the maximization of the likelihood ( $\alpha_S$ ). The product is then taken over the systematics (theoretical and experimental) considered,  $\prod_{S} (1 + \alpha_S \Delta X_S)$ . The systematic uncertainties are constrained to be Gaussian by multiplying P by a product of  $e^{-\alpha_s^2}$ .

$$\begin{aligned} \mathcal{L}(\mu, \alpha_{S}) &= \{ \prod_{\mathrm{R}} P(N_{\mathrm{obs, R}} \mid \mu \, s_{\mathrm{exp, R}}^{\mathrm{VBF}} \prod_{\mathrm{S}} (1 + \alpha_{S} \Delta s_{\mathrm{R, S}}^{\mathrm{VBF}}) \\ &+ \mu \, s_{\mathrm{exp, R}}^{ggF} \prod_{\mathrm{S}} (1 + \alpha_{S} \Delta s_{\mathrm{R, S}}^{ggF}) \\ &+ k_{V} \, N_{\mathrm{exp, R}}^{Z \to \nu \nu} \prod_{\mathrm{S}} (1 + \alpha_{S} \Delta N_{\mathrm{R, S}}^{Z \to \nu \nu}) \\ &+ k_{V} \, N_{\mathrm{exp, R}}^{W \to \ell \nu} \prod_{\mathrm{S}} (1 + \alpha_{S} \Delta N_{\mathrm{R, S}}^{W \to \ell \nu}) \\ &+ N_{\mathrm{exp, R}}^{\mathrm{Multijet}} \prod_{\mathrm{S}} (1 + \alpha_{S} \Delta N_{\mathrm{R, S}}^{\mathrm{Multijet}}) \\ &+ N_{\mathrm{exp, R}}^{\mathrm{Other Backgrounds}} \prod_{\mathrm{S}} (1 + \alpha_{S} \Delta N_{\mathrm{R, S}}^{\mathrm{Other Backgrounds}})) \} \times \{ \prod_{\mathrm{S}} e^{-\alpha_{S}^{2}} \} \end{aligned}$$

$$(10.3)$$

Parameters that float in the fit are the signal strength parameter  $\mu$ , the  $W \rightarrow \ell \nu$ +jets and  $Z \rightarrow \nu \nu$ +jets normalization factor  $k_V$ , and the systematic nuisance parameters  $\alpha_S$ . Since the signal strength parameter normalizes a signal yield that is normalized to a 100% invisible branching fraction,  $\mu$  represents the central value of the calculated branching fraction.

#### 10.3 Profile Likelihood Test Statistic

A profile likelihood test statistic is used to make the statistical computation, defined as:

$$q_{\mu} = -2\ln(\lambda(\mu)) \tag{10.4}$$

The  $\lambda$  function is the profile likelihood ratio defined as:

$$\lambda(\mu) = \frac{\mathcal{L}(\mu, \hat{\alpha}_S)}{\mathcal{L}(\hat{\mu}, \hat{\alpha}_S)} \tag{10.5}$$

where  $\hat{\alpha}_{S}$  represents the set of nuisance parameters that maximizes the likelihood for a specified value of  $\mu$  while  $\hat{\alpha}_{S}$  is associated with a fitted value of the signal strength ( $\hat{\mu}$ ) which maximizes the likelihood. The profile likelihood ratio is designed to quantify the deviation of a given hypothesis  $\mu$ from the best fit hypothesis  $\hat{\mu}$ .

#### 10.4 $CL_S$ Method

The limit on the branching fraction of the Higgs decaying invisibly is computed with the profile likelihood test statistic using a modified frequentist formalism known as the  $CL_S$  method. The  $CL_S$  method is defined as the ratio of *p*-values from signal+background and background only hypotheses. The *p*-values are determined from sampling distributions which is the probability of the test statistic  $(q_{\mu})$  based on a set of nuisance parameters  $(\hat{\alpha}_{S})$  and signal strength  $(\mu)$ . The *p*-value for the signal+background hypothesis derived from the sampling distribution  $f(q_{\mu}|\mu, \hat{\alpha}_{S}(\mu))$  is defined as:

$$p_{\mu} = \int_{q_{\mu}}^{\infty} f(q_{\mu}|\mu, \hat{\alpha_{S}}(\mu)) dq_{\mu}$$
(10.6)

This represents the probability to get a data distribution compatible with a value of  $\mu$ . The *p*-value for the background hypothesis (defined as  $p_b$ ) derived from the sampling distribution  $f(q_{\mu=0}|0, \hat{\alpha}_S(0))$ is defined as:

$$p_b = \int_{q_{\mu=0}}^{\infty} f(q_{\mu=0}|0, \hat{\alpha}_S(0)) dq_{\mu=0}$$
(10.7)

This represents the probability to get a data distribution less compatible with a signal strength of 0 than the observed data distribution.

Finally,  $CL_S$  is defined as the ratio of *p*-values for a signal+background hypothesis and background only hypothesis. Normalizing to the background only hypothesis ensures a statistically significant statement even if there is a fluctuation of the background only hypothesis that makes it similar to the best fit prediction.

$$CL_S = \frac{p_\mu}{1 - p_b} \tag{10.8}$$

A value of  $\mu$  is considered to be excluded if CL<sub>S</sub> is less than 0.05.

#### 10.5 Inputs to the Limit

All of the signal, background and control region MC and data are input parameters to the limit. All associated systematics (defined in Chapter 9) are also included in the limit as nuisance parameters. Correlations for appropriate parameters (such as systematics due to calibration of the detector) are taken into account by varying them coherently. The inputs to the limit setting is summarized below:

- 1.  $Z \to \ell \ell + jets$  Control Region
  - a) Observed yield
  - b) Strong+Weak  $Z \rightarrow \ell \ell$ +jets MC with associated statistical errors
  - c) Other backgrounds MC with associated statistical error
- 2.  $W \to \ell \nu + jets$  Control Region
- a) Observed yield
- b) Strong+Weak  $Z \rightarrow \ell \ell$ +jets MC with associated statistical errors
- c) Data driven multijet background estimate described in Section 6.2.2
- d) Other backgrounds MC with associated statistical error
- 3. Signal Region
  - a) Observed yield
  - b) ggF and VBF signal process MC with associated statistical errors
  - c) Strong+Weak  $Z \rightarrow \nu \nu$ +jets MC with associated statistical errors
  - d) Strong+Weak  $W \rightarrow \ell \nu$ +jets MC with associated statistical errors
  - e) Data driven multijet background estimate described in Chapter 8
  - f) Other backgrounds MC with associated statistical error
- 4. Systematics
  - a) Variation of factorization and renormalization scales for the  $W \to \ell \nu + \text{jets}$  in the control and signal regions along with the  $Z \to \nu \nu + \text{jets}$  and  $Z \to \ell \ell + \text{jets}$  are summarized in Table 10.1. Systematics are uncorrelated between weak and strong variations but are correlated between V+jets. Also presented are the factorization and renormalization scale variations for the signal processes. The signal process scale variations are completely uncorrelated with W and Z variations. Details are found in Section 9.3.1 and Section 9.2.
  - b) Uncertainties due to the PDF used to simulate events for signal processes,  $W \to \ell \nu + \text{jets}$ in the control and signal regions,  $Z \to \nu \nu + \text{jets}$  and  $Z \to \ell \ell + \text{jets}$  are summarized in Table 10.2. These uncertainties are uncorrelated. Details are found in Section 9.2 and Section 9.3.2.
  - c) An uncertainty due to the third jet veto is applied to the VBF signal process of 4.4%. Details are found in Section 9.2.1. This uncertainty is uncorrelated to all the others.
  - d) An uncertainty due to the Higgs  $p_T$  shape is applied to the ggF signal process of 9.7%. Details are found in Section 9.2. This uncertainty is uncorrelated to all the others.
  - e) Uncertainties due the parton showering scheme of 10.0% and 9.0% are applied to  $W \rightarrow \ell \nu + \text{jets}$  and  $Z \rightarrow \nu \nu + \text{jets}$ , respectively. Details are found in Section 9.3.3. These uncertainties are considered to be correlated between V+jets.

Process	Weak Variation		Strong Variation	
$W \to \ell \nu + \text{jets Signal Region}$	+0.1%	-0.4%	+45.1%	-28.0%
$W \to \ell \nu + \text{jets Control Region}$	+0.2%	-0.7%	+40.9%	-20.1%
$Z \rightarrow \nu \nu + \text{jets}$	+0.1%	-0.9%	+46.4%	-26.7%
$Z \to \ell \ell + \text{jets}$	+0.1%	-0.9%	+46.4%	-26.7%
Signal Process	Variation			
ggF	+7.2%		-7.8%	
VBF	+0.2%		-0.2%	

Table 10.1: Summary of uncertainties due to scale variations weighted by the relative production of strong and weak production.

Process	Error due to PDF		
$W \to \ell \nu + \text{jets Signal Region}$	+5.46% $-3.29%$		
$W \to \ell \nu + \text{jets Control Region}$	+3.59% $-2.59%$		
$Z \rightarrow \nu \nu + \text{jets}$	+3.32% $-2.59%$		
ggF Signal	+7.5% $-6.9%$		
VBF Signal	+2.6% $-2.8%$		

Table 10.2: Summary of uncertainties due to the PDF used to simulate events.

- f) A systematic due to uncertainty in the luminosity calculation is applied to all MC samples of 2.8%.
- g) A systematic is applied to the multijet estimate of 124% based on agreement with the validation regions. Detailed are presented in Section 8.2.
- h) Systematics due to the calibration of the detector for the signal processes,  $W \to \ell \nu + \text{jets}$ in the control and signal regions,  $Z \to \nu \nu + \text{jets}$  and  $Z \to \ell \ell + \text{jets}$  are summarized in Table 10.3. Details are found in Section 9.4.

### 10.6 Results

Once the fit is performed, the nuisance parameters defined in Section 10.5 are "pulled" (taking into account correlations) to maximize the likelihood. The impact of each parameter on the signal strength  $(\mu)$  is shown in Figure 10.1.

	Eta_ModellingJES	FlavRespJES	FlavCompJES	NP_Modelling1JES
VBF Signal (SR) Up	+15.72%	+4.17	+6.33%	+5.01%
VBF Signal (SR) Down	-14.13%	-4.18	-6.05%	-5.87%
ggF Signal (SR) Up	+22.4%	+0.0%	+0.0%	+22.4%
ggF Signal (SR) Down	-42.9%	-42.9%	-42.9%	-42.9%
$W \to \ell \nu + \text{jets (WCR) Up}$	+16.01%	+2.85%	+3.77%	+3.86%
$W \to \ell \nu + \text{jets}$ (WCR) Down	-13.98%	-4.84%	-6.53%	-5.85%
$W \to \ell \nu + \text{jets (SR) Up}$	+22.38%	+7.43%	+9.69%	+9.71%
$W \to \ell \nu + \text{jets (SR) Down}$	-14.66%	-4.00%	-6.78%	-5.39%
$Z \to \nu \nu + \text{jets (SR) Up}$	+17.20%	+3.05%	+5.37%	+6.90%
$Z \rightarrow \nu \nu + \text{jets (SR) Down}$	-14.96%	-5.09%	-7.56%	-6.30%
$Z \to \ell \ell + \text{jets} (\text{ZCR}) \text{ Up}$	+25.02%	+11.74%	+12.81%	+12.21%
$Z \to \ell \ell + \text{jets} (\text{ZCR}) \text{ Down}$	-15.53%	-3.17%	-8.73%	-4.04%

Table 10.3: Summary of systematic uncertainties due to calibration of the detector on the signal processes,  $W \to \ell \nu + \text{jets}$  in the control and signal regions,  $Z \to \nu \nu + \text{jets}$  and  $Z \to \ell \ell + \text{jets}$ .

It is seen that the nuisance parameters shifts are within 1  $\sigma$ , thus the fit is well behaved. Table 10.4 shows the pre-fit and post-fit yields of the backgrounds. It is observed that yields of the backgrounds have been pulled higher than the MC predictions. This is due to the observed yield in the  $Z \rightarrow \ell \ell + j$ ets being higher than the predicted MC. This is accommodated for in the fit by exploiting differences in the calibration of the jet energy scale in the detector and shifting the corresponding nuisance parameter.

The observed and expected  $CL_S$  values as a function of the branching fraction of the SM Higgs decaying invisibly is shown in Figure 10.2. The expected 95% confidence level upper bound on the branching fraction is found to be 34% while the observed is 29%.



Figure 10.1: Impact of nuisance parameters defined in Section ?? on signal strength  $(\mu)$  after being "pulled" by the fit.

Pre-Fit Yields					
Process	Yield $\pm$	Stat	$\pm$	Syst	
ggF Signal	$17 \pm$	6	$\pm$		
VBF Signal	$267 \pm$	5	$\pm$		
$Z \rightarrow \nu \nu + \text{jets}$	$325 \pm$	9	$\pm$		
$W \to \ell \nu + \text{jets}$	$225 \pm$	11	$\pm$		
Multijet QCD	$2 \pm$	2	$\pm$		
Other Backgrounds	$0.7 \pm$	0.2	$\pm$		
Total Background	$550 \pm$	13.6	$\pm$		
Post-Fit Yields					
ggF Signal	$20 \pm$	6	$\pm$	10	
VBF Signal	$286 \pm$	5	$\pm$	49	
$Z \rightarrow \nu \nu + \text{jets}$	$339 \pm$	22	$\pm$	13	
$W \to \ell \nu + \text{jets}$	$237 \pm$	17	$\pm$	18	
Multijet	$2 \pm$		2		
Other Backgrounds	$0.7 \pm$	0.2	$\pm$	0.3	
Total Background		20	+	30	
10tal Dackground	$= 378 \pm$	90	<u> </u>	50	

Table 10.4: Expected yields before and after the fit for the signal region in 20.3 fb<sup>-1</sup> of 2012 data. In the pre-fit yields, all processes are determined purely from MC except for the multijet. In the post-fit yields, the ggF signal, VBF signal, and other backgrounds are determined from MC while the  $Z \rightarrow \nu\nu$ +jets,  $W \rightarrow \ell\nu$ +jets, and multijet backgrounds are data-driven estimates. The expected signal yields are shown for  $m_H = 125$  GeV and are normalized to BR( $H \rightarrow invisible$ )=100%. The post-fit W and Z statistical uncertainties include MC statistics from both the selected region and the corresponding control region, and the number of data events in the control regions. The "Other Backgrounds" include top and diboson production.



Figure 10.2:  $CL_S$  values as a function of the branching fraction of the SM Higgs decaying invisibly. The expected 95% confidence level upper bound on the branching fraction is 34% while the observed is 28%.

### Chapter 11

## Conclusions

#### 11.1 Comparison to ZH

At the time of writing, the strongest direct constraint on the branching fraction of the Higgs decaying invisibly was produced by studying the VBF channel presented here. The previous strongest direct constraint was in the associated production channel, ZH with  $Z \rightarrow \ell\ell$ . The limit on the branching fraction was measured to be 75% with an expected limit of 62%. The VBF analysis has stronger limits by approximately a factor of 2 (29% observed with an expected limit of 34%). A higher S/Bis achieved in the VBF channel because of the lower signal expectation in ZH and overwhelming diboson backgrounds. In the ZH analysis, 44 signal events are expected with an assumed 100% invisible branching fraction of a SM Higgs and 138 background while for VBF 306 signal events are expected with 578 background. The VBF analysis has a larger signal expectation and S/B value that is  $\approx 1.6$  higher than the ZH analysis, mostly due to the unique jet structure in VBF processes that can be exploited to create an efficient selection.  $E_{\rm T}^{\rm miss}$  distributions for both analyses are shown in Figure 11.1.

#### 11.2 Comparison to Higgs Couplings Analysis

The branching fraction of the Higgs decaying invisibly can also be constrained indirectly, by performing a fit to the combined analyses of visible Higgs decays. Effective scale factors ( $\kappa_g$ ,  $\kappa_{\gamma}$ ,  $\kappa_{Z\gamma}$ ) are introduced to parameterize loop-induced Higgs production and decays ( $gg \rightarrow H, H \rightarrow \gamma\gamma, H \rightarrow Z\gamma$ ). In addition to the effective scale factors, a constraint can be calculated on the branching fraction of the Higgs decaying either invisibly or to undetected particles (BR<sub>i.,u.</sub>) which is a free parameter in the fit. Since couplings for tree-level processes are fixed to the SM, VBF and associated Higgs



the VBF analysis. Background distributions are taken the ZH analysis. Background distributions are taken from MC and data is overlayed.

(a)  $E_{\rm T}^{\rm miss}$  distribution after full selection is applied for (b)  $E_{\rm T}^{\rm miss}$  distribution after full selection is applied for from MC and data is overlayed.

	ggF	VBF	VH	ttH
$\gamma\gamma$	1	1	1	<ul> <li>✓</li> </ul>
WW	1	1	1	
ZZ	1	1	1	
au au	1	1	1	
$b\overline{b}$			1	1
$Z\gamma$	1			
$\mu\mu$	1	1		

Figure 11.1:  $E_{\rm T}^{\rm miss}$  distributions for VBF and ZH analyses.

Table 11.1: Observable Higgs decay analyses used to determine limits on the coupling

productions provide the strongest constraint to the branching fraction of the Higgs decaying invisibly. The analyses used in the combined fit are shown in Table 11.1.

The profile likelihood as a function of  $BR_{i,u}$  is shown in Figure 11.2. The fit results in an observed limit on BR<sub>i.,u.</sub> of 27% with an expected limit of 37%. This is very comparable to the strength of the direct VBF search presented here.

#### 11.3**Dark Matter Portal**

The branching fraction limit of the Higgs decaying invisibly can interpreted as a limit on the coupling between the Higgs and dark matter  $(\lambda_{h\chi\chi})$  as well as the nucleon cross section  $(\sigma_{\chi N})$  as a function of the mass of the dark matter candidate  $(m_{\chi})$ . The interpreted limits can then be compared to direct



Figure 11.2: Profile likelihood ratio as a function of  $BR_{i.,u.}$  calculated using visible Higgs decays. The red(green) horizontal lines indicate cut-off values on the profile likelihood ratio corresponding to 68%(95%) confidence level.

detection scattering experiments which also measure the limit of  $\sigma_{\chi N}$  as a function of  $m_{\chi}$ . Figure 11.3 shows the interpreted branching fraction limit, direct detection exclusion limits [49, 50, 51, 52, 53] and direct detection signals [54, 55, 56, 57].

### 11.4 Future of Invisibly Decaying Higgs Searches

It is observed that the VBF search for an invisibly decaying Higgs places the strongest constraint on the branching fraction. This is comparable to an indirect constraint from the Higgs coupling measurement seen in Section 11.2. Future stronger constraints will result from Run II of the LHC which will operate at a higher center of mass energy  $\sqrt{s} = 13$  TeV (vs  $\sqrt{s} = 8$  TeV considered for the analysis presented here) and have an increased instantaneous luminosity of  $2 \times 10^{34}$  cm<sup>-2</sup> s<sup>-1</sup> (vs  $8 \times 10^{33}$  cm<sup>-2</sup> s<sup>-1</sup> in Run I). Besides the higher statistics available for analysis at the end of Run II, the higher center of mass energy results in higher cross sections for weakly produced processes ( $\approx 2.2$ ) but a lower increase for strongly produced processes ( $\approx 1.8$ ). This corresponds to  $\approx 10\%$  increase in S/B purely due to the center of mass energy increase.

However, the increase in instantaneous luminosity will also introduce challenges due to pile-up. Because of the high  $\Delta \eta_{j,j}$  required in the analysis the probability that a jet is outside of the tracking



Figure 11.3: Spin-independent DM-nucleon cross section as a function of the DM mass. The exclusion limits [49, 50, 51, 52, 53] and signals [54, 55, 56, 57] observed by the direct detection experiments are compared to the branching fraction limit interpreted by the Higgs-portal and translated into the spin-independent DM-nucleon cross section using the formulas in Section 2.5.5. The exclusion limits are shown at 90% CL. The error bands on the ATLAS results indicate the uncertainty coming from the different estimations of the Higgs-nucleon coupling [58, 59].

volume is substantial which increases the chance that a tagged jet is actually from pile-up.

Based on the results of the analysis presented here, there are several areas that can be improved upon for Run II. Some are listed here:

- 1. Increased MC statistics to more accurate estimate systematic errors due to detector calibration
- 2. A procedure to estimate the multijet background that can be compared to MC
  - a) Generating sufficient multijet MC events
- 3. Develop a more complete way of estimating a systematic due to parton shower modeling

Since dark matter has not been discovered yet, pursuing analyses involving dark matter at the LHC is very important. The Run II VBF analysis is a promising way of further constraining the invisible branching fraction of the Higgs and is an important complement to the constraint provided by the Higgs coupling measurement.

# Bibliography

- F. Hasert, S. Kabe, W. Krenz, J. V. Krogh, D. Lanske, J. Morfin, K. Schultze, H. Weerts, G. Bertrand-Coremans, J. Sacton, W. V. Doninck, P. Vilain, U. Camerini, D. Cundy, R. Baldi, I. Danilchenko, W. Fry, D. Haidt, S. Natali, P. Musset, B. Osculati, R. Palmer, J. Pattison, D. Perkins, A. Pullia, A. Rousset, W. Venus, H. Wachsmuth, V. Brisson, B. Degrange, M. Haguenauer, L. Kluberg, U. Nguyen-Khac, P. Petiau, E. Belotti, S. Bonetti, D. Cavalli, C. Conta, E. Fiorini, M. Rollier, B. Aubert, D. Blum, L. Chounet, P. Heusse, A. Lagarrigue, A. Lutz, A. Orkin-Lecourtois, J. Vialle, F. Bullock, M. Esten, T. Jones, J. McKenzie, A. Michette, G. Myatt, and W. Scott, Observation of neutrino-like interactions without muon or electron in the gargamelle neutrino experiment, Physics Letters B 46 (1973) no. 1, 138 – 140. http://www.sciencedirect.com/science/article/pii/0370269373904991. 2.2
- [2] UA1 Collaboration, G. Arnison et al., Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at s\*\*(1/2) = 540-GeV, Phys.Lett. B122 (1983) 103-116. 2.2
- [3] M. Banner, R. Battiston, P. Bloch, F. Bonaudi, K. Borer, M. Borghini, J.-C. Chollet, A. Clark, C. Conta, P. Darriulat, L. D. Lella, J. Dines-Hansen, P.-A. Dorsaz, L. Fayard, M. Fraternali, D. Froidevaux, J.-M. Gaillard, O. Gildemeister, V. Goggi, H. Grote, B. Hahn, H. Hnni, J. Hansen, P. Hansen, T. Himel, V. Hungerbhler, P. Jenni, O. Kofoed-Hansen, E. Lanon, M. Livan, S. Loucatos, B. Madsen, P. Mani, B. Mansoulie, G. Mantovani, L. Mapelli, B. Merkel, M. Mermikides, R. Mllerud, B. Nilsson, C. Onions, G. Parrour, F. Pastore, H. Plothow-Besch, M. Polverel, J.-P. Repellin, A. Rothenberg, A. Roussarie, G. Sauvage, J. Schacher, J. Siegrist, H. Steiner, G. Stimpfl, F. Stocker, J. Teiger, V. Vercesi, A. Weidberg, H. Zaccone, and W. Zeller, Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the {CERN} pp collider, Physics Letters B 122 (1983) no. 56, 476 485. http://www.sciencedirect.com/science/article/pii/0370269383916052. 2.2
- [4] G. Arnison, A. Astbury, B. Aubert, C. Bacci, G. Bauer, A. Bzaguet, R. Bck, T. Bowcock, M. Calvetti, P. Catz, P. Cennini, S. Centro, F. Ceradini, S. Cittolin, D. Cline, C. Cochet, J. Colas, M. Corden, D. Dallman, D. Dau, M. DeBeer, M. Negra, M. Demoulin, D. Denegri, A. D. Ciaccio, D. Dibitonto, L. Dobrzynski, J. Dowell, K. Eggert, E. Eisenhandler, N. Ellis, P. Erhard, H. Faissner, M. Fincke, G. Fontaine, R. Frey, R. Frhwirth, J. Garvey, S. Geer, C. Ghesquire, P. Ghez, K. Giboni, W. Gibson, Y. Giraud-Hraud, A. Givernaud, A. Gonidec, G. Grayer, T. Hansl-Kozaecka, W. Haynes, L. Hertzberger, C. Hodges, D. Hoffmann, H. Hoffmann, D. Holthuizen, R. Homer, A. Honma, W. Jank, G. Jorat, P. Kalmus, V. Karimki,

R. Keeler, I. Kenyon, A. Kernan, R. Kinnunen, W. Kozanecki, D. Kryn, F. Lacava, J.-P. Laugier, J.-P. Lees, H. Lehmann, R. Leuchs, A. Lvque, D. Linglin, E. Locci, J.-J. Malosse, T. Markiewicz, G. Maurin, T. McMahon, J.-P. Mendiburu, M.-N. Minard, M. Mohammadi, M. Moricca, K. Morgan, H. Muirhead, F. Muller, A. Nandi, L. Naumann, A. Norton, A. Orkin-Lecourtois, L. Paoluzi, F. Pauss, G. Mortari, E. Pietarinen, M. Pimi, A. Placci, J. Porte, E. Radermacher, J. Ransdell, H. Reithler, J.-P. Revol, J. Rich, M. Rijssenbeek, C. Roberts, J. Rohlf, P. Rossi, C. Rubbia, B. Sadoulet, G. Sajot, G. Salvi, G. Salvini, J. Sass, J. Saudraix, A. Savoy-Navaro, D. Schinzel, W. Scott, T. Shah, M. Spiro, J. Strauss, J. Streets, K. Sumorok, F. Szoncso, D. Smith, C. Tao, G. Thompson, J. Timmer, E. Tscheslog, J. Touminiemi, B. V. Eijk, J.-P. Vialle, J. Vrana, V. Vuillemin, H. Wahl, P. Watkins, J. Wilson, C. Wulz, G. Xie, M. Yvert, and E. Zurfluh, *Experimental observation of lepton pairs of invariant mass around 95 GeV/c2 at the {CERN} {SPS} collider*, Physics Letters B 126 (1983) no. 5, 398 – 410.
http://www.sciencedirect.com/science/article/pii/0370269383901880. 2.2

- [5] UA2 Collaboration, P. Bagnaia et al., Evidence for Z0 −<sub>i</sub> e+ e- at the CERN anti-p p Collider, Phys.Lett. B129 (1983) 130–140. 2.2
- [6] LEP Working Group for Higgs boson searches, ALEPH, DELPHI, L3, OPAL Collaboration, R. Barate et al., Search for the standard model Higgs boson at LEP, Phys.Lett. B565 (2003) 61-75, arXiv:hep-ex/0306033 [hep-ex]. 2.3
- Higgs Working Group Collaboration, M. Carena et al., Report of the Tevatron Higgs working group, arXiv:hep-ph/0010338 [hep-ph]. 2.3
- [8] Tevatron New Physics Higgs Working Group, CDF, D0 Collaboration, Updated Combination of CDF and D0 Searches for Standard Model Higgs Boson Production with up to 10.0 fb<sup>-1</sup> of Data, arXiv:1207.0449 [hep-ex]. 2.3
- [9] ATLAS Collaboration, G. Aad et al., Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC, Phys.Lett. B716 (2012) 1-29, arXiv:1207.7214 [hep-ex]. 2.3
- [10] CMS Collaboration, S. Chatrchyan et al., Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC, Phys.Lett. B716 (2012) 30-61, arXiv:1207.7235 [hep-ex].
   2.3
- [11] CDF, D0 Collaboration, T. Aaltonen et al., *Higgs Boson Studies at the Tevatron*, Phys.Rev. D88 (2013) no. 5, 052014, arXiv:1303.6346 [hep-ex]. 2.3
- E. Corbelli and P. Salucci, The Extended Rotation Curve and the Dark Matter Halo of M33, Mon. Not. Roy. Astron. Soc. 311 (2000) 441–447, arXiv:astro-ph/9909252 [astro-ph]. 2.5.1, 2.4
- ESA/Hubble and NASA, A Horseshoe Einstein Ring from Hubble, Http://apod.nasa.gov/apod/ap111221.html. 2.5.1, 2.5
- [14] WMAP Collaboration, G. Hinshaw et al., Three-year Wilkinson Microwave Anisotropy Probe (WMAP) observations: temperature analysis, Astrophys.J.Suppl. 170 (2007) 288, arXiv:astro-ph/0603451 [astro-ph]. 2.5.1, 2.6
- [15] G. Hinshaw, D. Larson, E. Komatsu, D. N. Spergel, C. L. Bennett, J. Dunkley, M. R. Nolta, M. Halpern, R. S. Hill, N. Odegard, L. Page, K. M. Smith, J. L. Weiland, B. Gold, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, E. Wollack, and E. L. Wright, *Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Parameter*

*Results*, The Astrophysical Journal Supplement Series **2**08 (2013) no. 2, 19. http://stacks.iop.org/0067-0049/208/i=2/a=19. 2.5.2

- [16] E. I. Gates, G. Gyuk, and M. S. Turner, *The Local halo density*, Astrophys.J. 449 (1995) L123–L126, arXiv:astro-ph/9505039 [astro-ph]. 2.5.3
- [17] Particle Data Group Collaboration, K. Olive et al., *Review of Particle Physics*, Chin.Phys. C38 (2014) 090001. 2.7, 2.5.3
- [18] I. Cholis, D. P. Finkbeiner, L. Goodenough, and N. Weiner, *The PAMELA Positron Excess from Annihilations into a Light Boson*, JCAP 0912 (2009) 007, arXiv:0810.5344 [astro-ph]. 2.5.4
- [19] A. Djouadi, O. Lebedev, Y. Mambrini, and J. Quevillon, *Implications of LHC searches for Higgs-portal dark matter*, Phys.Lett. B709 (2012) 65–69, arXiv:1112.3299 [hep-ph]. 2.5.5
- [20] P. J. Fox, R. Harnik, J. Kopp, and Y. Tsai, Missing Energy Signatures of Dark Matter at the LHC, Phys.Rev. D85 (2012) 056011, arXiv:1109.4398 [hep-ph]. 2.5.5
- [21] L. Lopez-Honorez, T. Schwetz, and J. Zupan, Higgs portal, fermionic dark matter, and a Standard Model like Higgs at 125 GeV, Physics Letters B 716 (2012) no. 1, 179 – 185. http://www.sciencedirect.com/science/article/pii/S0370269312007575. 2.5.5
- [22] A. Martin, W. Stirling, R. Thorne, and G. Watt, Parton distributions for the LHC, Eur.Phys.J. C63 (2009) 189–285, arXiv:0901.0002 [hep-ph]. 3.3, 3.2
- [23] ATLAS Collaboration, G. Aad et al., Identification and energy calibration of hadronically decaying tau leptons with the ATLAS experiment in pp collisions at √s=8 TeV, Eur. Phys. J. C75 (2015) no. 7, 303, arXiv:1412.7086 [hep-ex]. 3.5.5, 5.2
- [24] M. Cacciari, G. P. Salam, and G. Soyez, The Anti-k(t) jet clustering algorithm, JHEP 0804 (2008) 063, arXiv:0802.1189 [hep-ph]. 3.5.6
- [25] Calibration of the performance of b-tagging for c and light-flavour jets in the 2012 ATLAS data, Tech. Rep. ATLAS-CONF-2014-046, CERN, Geneva, Jul, 2014. https://cds.cern.ch/record/1741020. 3.5.7, 5.2
- [26] ATLAS Collaboration, ATLAS detector and physics performance: Technical Design Report. CERN, Geneva, 1999. 4.1, 4.1
- [27] ATLAS Collaboration, G. Aad et al., Performance of the ATLAS Trigger System in 2010, Eur. Phys. J. C72 (2012) 1849, arXiv:1110.1530 [hep-ex]. 4.3, 4.3
- [28] Pile-up subtraction and suppression for jets in ATLAS, Tech. Rep. ATLAS-CONF-2013-083, CERN, Geneva, Aug, 2013. https://cds.cern.ch/record/1570994. 5.2
- [29] Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data, Tech. Rep. ATLAS-CONF-2014-032, CERN, Geneva, Jun, 2014. https://cds.cern.ch/record/1706245. 5.2
- [30] ATLAS Collaboration, G. Aad et al., Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC protonproton collision data, Eur. Phys. J. C74 (2014) no. 11, 3130, arXiv:1407.3935 [hep-ex]. 5.2
- [31] Y. Li and F. Petriello, Combining QCD and electroweak corrections to dilepton production in FEWZ, Phys.Rev. D86 (2012) 094034, arXiv:1208.5967 [hep-ph]. 5.5.2

- [32] J. Baglio, J. Bellm, F. Campanario, B. Feigl, J. Frank, et al., Release Note VBFNLO 2.7.0, arXiv:1404.3940 [hep-ph]. 5.5.2
- [33] K. Arnold, J. Bellm, G. Bozzi, M. Brieg, F. Campanario, et al., VBFNLO: A Parton Level Monte Carlo for Processes with Electroweak Bosons – Manual for Version 2.5.0, arXiv:1107.4038 [hep-ph]. 5.5.2
- [34] K. Arnold, M. Bahr, G. Bozzi, F. Campanario, C. Englert, et al., VBFNLO: A Parton level Monte Carlo for processes with electroweak bosons, Comput.Phys.Commun. 180 (2009) 1661–1670, arXiv:0811.4559 [hep-ph]. 5.5.2
- [35] A. Denner, S. Dittmaier, S. Kallweit, and A. Muck, Electroweak corrections to Higgs-strahlung off W/Z bosons at the Tevatron and the LHC with HAWK, JHEP 1203 (2012) 075, arXiv:1112.5142 [hep-ph]. 5.5.2
- [36] D. de Florian, G. Ferrera, M. Grazzini, and D. Tommasini, Transverse-momentum resummation: Higgs boson production at the Tevatron and the LHC, JHEP 11 (2011) 064, arXiv:1109.2109 [hep-ph]. 5.5.2
- [37] ATLAS e/gamma Working Group, e/gamma Recommendations for 2012 Analyses, https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/Egamma2012Analyses. 7.2.2
- [38] M. Botje, J. Butterworth, A. Cooper-Sarkar, A. de Roeck, J. Feltesse, et al., The PDF4LHC Working Group Interim Recommendations, arXiv:1101.0538 [hep-ph]. 7.3.3, 9.3.2
- [39] ATLAS Collaboration, Measurement of the inclusive isolated prompt photon cross section in pp collisions at  $\sqrt{s} = 7$  TeV with the ATLAS detector, Phys.Rev. D83 (2011) 052005, arXiv:1012.4389 [hep-ex]. 7.4.2
- [40] M. Hance and H. H. Williams, Measurement of Inclusive Isolated Prompt Photon Production in Proton-Proton Collisions at  $\sqrt{s} = 7$  TeV with the ATLAS Detector. PhD thesis, Pennsylvania U., Philadelphia, 2011. Presented 11 Jul 2011. 7.4.2
- [41] H. Abreu, N. Andari, L. Carminati, R. K. Daya, M. Fanti, L. Fayard, M. Hance, Z. Liang, J. F. Marchand, G. Marchiori, H. Torres, R. Turra, G. Unal, and H. H. Williams, *Purity Estimates for the Inclusive Isolated Photons*, Tech. Rep. ATL-PHYS-INT-2011-015, CERN, Geneva, Mar, 2011. 1
- [42] ATLAS Collaboration, G. Aad et al., Measurements of the W production cross sections in association with jets with the ATLAS detector, Eur.Phys.J. C75 (2015) no. 2, 82, arXiv:1409.8639 [hep-ex]. 7.5
- [43] ATLAS Collaboration, G. Aad et al., Measurement of the production cross section of jets in association with a Z boson in pp collisions at  $\sqrt{s} = 7$  TeV with the ATLAS detector, JHEP 1307 (2013) 032, arXiv:1304.7098 [hep-ex]. 7.5
- [44] C. Gutschow, G. Jones, K. Joshi, E. Nurse, and A. D. Pilkington, Measurement of the electroweak Zjj cross section and distributions sensitive to vector boson fusion in proton-proton collisions at sqrts = 8 TeV using the ATLAS detector, . 7.15
- [45] A. Dattagupta, J. Fischer, C. Hays, J. Henderson, S. Lammers, D. Price, C. Zeitnitz, and A. Johnson, Observation and measurement of Electroweak W + 2 jets production at high dijet mass, . 7.15

- [46] M. Grazzini and H. Sargsyan, Heavy-quark mass effects in Higgs boson production at the LHC, JHEP 1309 (2013) 129, arXiv:1306.4581 [hep-ph]. 2a
- [47] D. de Florian, G. Ferrera, M. Grazzini, and D. Tommasini, Higgs boson production at the LHC: transverse momentum resummation effects in the  $H \to \gamma\gamma$ ,  $H \to WW \to \ell\nu\ell\nu$  and  $H \to ZZ \to 4\ell$  decay modes, JHEP 1206 (2012) 132, arXiv:1203.6321 [hep-ph]. 2a
- [48] I. W. Stewart, F. J. Tackmann, J. R. Walsh, and S. Zuberi, Jet p<sub>T</sub> resummation in Higgs production at NNLL' + NNLO, Phys.Rev. D89 (2014) no. 5, 054001, arXiv:1307.1808. 2a
- [49] CRESST-II Collaboration, Results on low mass WIMPs using an upgraded CRESST-II detector, Eur.Phys.J. C74 (2014) no. 12, 3184, arXiv:1407.3146 [astro-ph.CO]. 11.3, 11.3
- [50] SuperCDMS Collaboration, R. Agnese et al., Search for Low-Mass Weakly Interacting Massive Particles Using Voltage-Assisted Calorimetric Ionization Detection in the SuperCDMS Experiment, Phys.Rev.Lett. 112 (2014) no. 4, 041302, arXiv:1309.3259 [physics.ins-det]. 11.3, 11.3
- [51] SuperCDMS Collaboration, Search for Low-Mass Weakly Interacting Massive Particles with SuperCDMS, Phys.Rev.Lett. 112 (2014) no. 24, 241302, arXiv:1402.7137 [hep-ex]. 11.3, 11.3
- [52] XENON100 Collaboration, Limits on spin-dependent WIMP-nucleon cross sections from 225 live days of XENON100 data, Phys.Rev.Lett. 111 (2013) no. 2, 021301, arXiv:1301.6620 [astro-ph.CO]. 11.3, 11.3
- [53] LUX Collaboration, First results from the LUX dark matter experiment at the Sanford Underground Research Facility, Phys.Rev.Lett. 112 (2014) 091303, arXiv:1310.8214 [astro-ph.CO]. 11.3, 11.3
- [54] DAMA Collaboration, First results from DAMA/LIBRA and the combined results with DAMA/NaI, Eur.Phys.J. C56 (2008) 333–355, arXiv:0804.2741 [astro-ph]. 11.3, 11.3
- [55] C. E. Aalseth et al., Maximum Likelihood Signal Extraction Method Applied to 3.4 years of CoGeNT Data, arXiv:1401.6234 [astro-ph.CO]. 11.3, 11.3
- [56] G. Angloher et al., Results from 730 kg days of the CRESST-II Dark Matter Search, Eur.Phys.J. C72 (2012) 1971, arXiv:1109.0702 [astro-ph.CO]. 11.3, 11.3
- [57] CDMS Collaboration, CDMS Collaboration, Silicon Detector Dark Matter Results from the Final Exposure of CDMS II, Phys.Rev.Lett. 111 (2013) 251301, arXiv:1304.4279 [hep-ex]. 11.3, 11.3
- [58] R. Young and A. Thomas, Octet baryon masses and sigma terms from an SU(3) chiral extrapolation, Phys.Rev. D81 (2010) 014503, arXiv:0901.3310 [hep-lat]. 11.3
- [59] MILC Collaboration, D. Toussaint and W. Freeman, The Strange quark condensate in the nucleon in 2+1 flavor QCD, Phys.Rev.Lett. 103 (2009) 122002, arXiv:0905.2432 [hep-lat]. 11.3