

<sup>1</sup> **Searches for Supersymmetric Signatures in  
2 all Hadronic Final States with the  $\alpha_T$   
3 Variable.**

<sup>4</sup> Darren Burton

<sup>5</sup> Imperial College London  
<sup>6</sup> Department of Physics

<sup>7</sup> A thesis submitted to Imperial College London  
<sup>8</sup> for the degree of Doctor of Philosophy

---

## Abstract

A search for supersymmetric particles in events with high transverse momentum jets and a large missing transverse energy signature, is conducted using  $11.7 \text{ fb}^{-1}$  of data, collected with a center-of-mass collision energy of 8 TeV by the CMS detector. The dimensionless kinematic variable  $\alpha_T$  is used to select events with genuine missing transverse energy signatures. Standard Model backgrounds are estimated through the use of data driven control samples. No excess over Standard Model expectations is found. Exclusion limits on squark and gluino masses are set at the 95% confidence level in the parameter space of a range of supersymmetric simplified model topologies.

Results of benchmarking the Level-1 (the first line of the CMS trigger system) single jet and hadronic transverse energy trigger efficiencies, before and after the implementation of a change to the Level-1 jet clustering algorithm are presented. Similar performance is observed for all L1 quantities. This change was introduced to negate an increase in trigger cross-section, which can be attributed to soft jets from secondary interactions.

Furthermore, a templated fit method to estimate the Standard Model background distribution of the number of jets originating from a b-quark within a supersymmetric search, is validated in data and simulation. Applicable to searches sensitive to gluino induced third-generation signatures, this technique is utilised as a crosscheck to the results of the  $\alpha_T$  analysis. Standard Model background predictions from the template fits are compared to those from the  $\alpha_T$  search in the hadronic signal region, where good agreement between the two methods is observed.

38

## Declaration

39 I, the author of this thesis, declare that the work presented within this  
40 document to be my own. The work presented in Chapters 4, 5, 6 and Section  
41 3.4, is a result of the author's own work, or that of which I have been a major  
42 contributor unless explicitly stated otherwise, and is carried out within the  
43 context of the Imperial College London and CERN SUSY groups, itself a  
44 subsection of the greater CMS collaboration. All figures and studies taken  
45 from external sources are referenced appropriately throughout this document.

46

Darren Burton

47

## Acknowledgements

- 48 I would like to thank the many people whom I have had the pleasure of working with  
49 during the course of the last three and a half years. The opportunity to work as part  
50 of the largest scientific collaboration during one of the most exciting times in particle  
51 physics for decades, has been a real privilege to be a part of. I could not have achieved  
52 the results presented in this thesis without the help of my colleagues who were part of the  
53 RA1 team, Edward Laird, Chris Lucas, Henning Flaecher, Yossof Eshaq, Bryn Mathais,  
54 Sam Rogerson, Zhaoxia Meng and Georgia Karapostoli whom I worked with on L1 jets. I  
55 also thank my supervisor Oliver Buchmuller for his guidance in getting me to this point.
- 56 I also feel it important to single out thanks to the postdocs that I have worked with during  
57 my PhD. Jad Marrouche from whom I have learnt a great deal and Robert Bainbridge  
58 who has been like a second supervisor to me, helping me during my time at Imperial and  
59 CERN, especially during those most stressful of times approaching conference deadlines!
- 60 My fellow PhD students who I live with and have seen on an almost daily basis for the  
61 last few years, Andrew Gilbert, Patrick Owen, Indrek Sepp, Matthew Kenzie and my  
62 wonderful girlfriend Hannah. Thanks for putting up with the whining, complaining and  
63 clomping.
- 64 Finally my largest thanks go to my Mum and Dad whose patience, encouragement and  
65 considerable financial support have allowed me to take the many steps that led me here  
66 today.

# 67 **Contents**

68	<b>List of Figures</b>	viii
69	<b>List of Tables</b>	xiii
70	<b>1. Introduction</b>	2
71	<b>2. A Theoretical Overview</b>	5
72	2.1. The Standard Model . . . . .	5
73	2.1.1. Gauge Symmetries of the SM . . . . .	7
74	2.1.2. The Electroweak Sector and Electroweak Symmetry Breaking . . . . .	9
75	2.2. Motivation for Physics beyond the Standard Model . . . . .	13
76	2.3. Supersymmetry Overview . . . . .	14
77	2.3.1. R-Parity . . . . .	16
78	2.4. Experimental Signatures of SUSY at the LHC . . . . .	16
79	2.4.1. Simplified Models . . . . .	18
80	<b>3. The LHC and the CMS Detector</b>	20
81	3.1. The LHC . . . . .	20
82	3.2. The CMS Detector . . . . .	23
83	3.2.1. Detector Subsystems . . . . .	23
84	3.2.2. Tracker . . . . .	24
85	3.2.3. Electromagnetic Calorimeter . . . . .	25
86	3.2.4. Hadronic Calorimeter . . . . .	26
87	3.2.5. Muon Systems . . . . .	27
88	3.3. Event Reconstruction and Object Definition . . . . .	28
89	3.3.1. Jets . . . . .	28
90	3.3.2. B-tagging . . . . .	30
91	3.4. Triggering System . . . . .	33
92	3.4.1. The Level-1 Trigger . . . . .	34

93	3.4.2. The L1 Trigger Jet Algorithm . . . . .	36
94	3.4.3. Measuring L1 Single-Jet Trigger Efficiencies . . . . .	38
95	3.4.4. Effects of the L1 Jet Seed . . . . .	40
96	3.4.5. Robustness of L1 Jet Performance against Pile-up . . . . .	43
97	3.4.6. Summary . . . . .	46
98	<b>4. SUSY Searches in Hadronic Final States</b>	<b>47</b>
99	4.1. An Introduction to the $\alpha_T$ Search . . . . .	48
100	4.1.1. The $\alpha_T$ Variable . . . . .	50
101	4.2. Search Strategy . . . . .	52
102	4.2.1. Physics Objects . . . . .	55
103	4.2.2. Event Selection . . . . .	58
104	4.2.3. Control Sample Definition and Background Estimation . . . . .	61
105	4.2.4. Estimating the QCD Multi-jet Background . . . . .	68
106	4.3. Trigger Strategy . . . . .	70
107	4.4. Measuring Standard Model Process Normalisation Factors via $H_T$ Sidebands	71
108	4.5. Determining Monte Carlo Simulation Yields with Higher Statistical Precision	73
109	4.5.1. The Formula Method . . . . .	73
110	4.5.2. Establishing Proof of Principle . . . . .	75
111	4.5.3. Correcting Measured Efficiencies in Simulation to Data . . . . .	76
112	4.6. Systematic Uncertainties on Transfer Factors . . . . .	78
113	4.6.1. Determining Systematic Uncertainties from Closure Tests . . . . .	82
114	4.7. Simplified Models, Efficiencies and Systematic Uncertainties . . . . .	84
115	4.7.1. Signal Efficiency . . . . .	85
116	4.7.2. Applying B-tagging Scale Factor Corrections in Signal Samples . .	86
117	4.7.3. Experimental Uncertainties . . . . .	87
118	4.8. Statistical Interpretation . . . . .	90
119	4.8.1. Hadronic Sample . . . . .	90
120	4.8.2. $H_T$ Evolution Model . . . . .	91
121	4.8.3. Electroweak Sector (EWK) Control Samples . . . . .	92
122	4.8.4. Contributions from Signal . . . . .	94
123	4.8.5. Total Likelihood . . . . .	95
124	<b>5. Results and Interpretation</b>	<b>97</b>
125	5.1. Compatibility with the Standard Model Hypothesis . . . . .	97
126	5.2. SUSY . . . . .	106
127	5.2.1. The $CL_s$ Method . . . . .	106

128	5.2.2. Interpretation in Simplified Signal Models . . . . .	107
129	<b>6. SUSY Searches with B-tag Templates</b>	111
130	6.1. Concept . . . . .	112
131	6.2. Application to the $\alpha_T$ Search . . . . .	115
132	6.2.1. Proof of Principle in Simulation . . . . .	116
133	6.2.2. Results in a Data Control Sample . . . . .	120
134	6.2.3. Application to the $\alpha_T$ Hadronic Search Region . . . . .	122
135	6.3. Summary . . . . .	124
136	<b>7. Conclusions</b>	126
137	<b>A. Miscellaneous</b>	128
138	A.1. Jet Identification Criteria . . . . .	128
139	A.2. Primary Vertices . . . . .	129
140	<b>B. L1 Jets</b>	130
141	B.1. Jet Matching Efficiencies . . . . .	130
142	B.2. Leading Jet Energy Resolution . . . . .	131
143	B.3. Resolution for Energy Sum Quantities . . . . .	134
144	<b>C. Additional Material on Background Estimation Methods</b>	136
145	C.1. Determination of $k_{QCD}$ . . . . .	136
146	C.2. Effect of Varying Background Cross-sections on Closure Tests . . . . .	137
147	<b>D. Additional Material For B-tag Template Method</b>	139
148	D.1. Templates Fits in Simulation . . . . .	139
149	D.2. Pull Distributions for Template Fits . . . . .	142
150	D.3. Templates Fits in Data Control Sample . . . . .	143
151	D.4. Templates Fits in Data Signal Region . . . . .	146
152	<b>Bibliography</b>	149

# <sup>153</sup> List of Figures

<sup>154</sup> 2.1.	One loop quantum corrections to the Higgs squared mass parameter $m_h^2$ due to a fermion. . . . .	14
<sup>156</sup> 2.2.	Two example simplified model decay chains. . . . .	19
<sup>157</sup> 3.1.	A top-down layout of the LHC, with the position of the four main detectors labelled. . . . .	21
<sup>159</sup> 3.2.	The total integrated luminosity delivered to and collected by Compact Muon Solenoid (CMS) during the 2012 8 TeV $pp$ runs . . . . .	22
<sup>161</sup> 3.3.	A pictorial depiction of the CMS detector. . . . .	24
<sup>162</sup> 3.4.	Illustration of the CMS Electromagnetic CALorimeter (ECAL). . . . .	26
<sup>163</sup> 3.5.	Schematic of the CMS Hadronic CALorimeter (HCAL). . . . .	27
<sup>164</sup> 3.6.	Combined Secondary Vertex (CSV) algorithm discriminator values in enriched ttbar and inclusive multi-jet samples . . . . .	31
<sup>166</sup> 3.7.	Data/MC b-tag scale factors derived using the Combined Secondary Vertex Medium Working Point (CSVM) tagger. . . . .	32
<sup>168</sup> 3.8.	Data/MC mis-tag scale factors derived using the CSVM tagger. . . . .	33
<sup>169</sup> 3.9.	An overview of the different components of the CMS L1 trigger system .	34
<sup>170</sup> 3.10.	Illustration of the dimensions of the Level-1 jet finder window. . . . .	37
<sup>171</sup> 3.11.	L1 jet efficiency turn-on curves as a function of the offline CaloJet and PFJet $E_T$ . . . . .	39
<sup>173</sup> 3.12.	L1 jet efficiency turn-on curves as a function of the offline CaloJet $E_T$ for the 2012 run period B and C. . . . .	41

---

175	3.13. L1 $H_T$ efficiency turn-on curves as a function of the offline CaloJet $H_T$ . . . . .	41
176	3.14. Trigger cross section for the L1HTT150 trigger path. . . . .	43
177	3.15. L1 jet efficiency turn-on curves as a function of the leading offline $E_T$ Calo	
178	(left) and PF (right) jet, for low, medium and high pile-up conditions. . . . .	44
179	3.16. Fit values from an Exponentially Modified Gaussian (EMG) function fitted	
180	to the resolution plots of leading Calo jet $E_T$ measured as a function of	
181	$\frac{(L1\ E_T - \text{Offline}\ E_T)}{\text{Offline}\ E_T}$ for low, medium and high pile-up conditions. . . . .	45
182	3.17. Fit values from an EMG function fitted to the resolution plots of leading	
183	PF jet $E_T$ measured as a function of $\frac{(L1\ E_T - \text{Offline}\ E_T)}{\text{Offline}\ E_T}$ for low, medium and	
184	high pile-up conditions. . . . .	46
185	4.1. Reconstructed offline $H_T$ distribution in the hadronic signal selection, from	
186	$11.7\text{fb}^{-1}$ of data, in which no $\alpha_T$ requirement is made. . . . .	49
187	4.2. The event topologies of background QCD dijet events (right) and a generic	
188	SUSY signature with genuine $Z_T$ (left). . . . .	50
189	4.3. The $\alpha_T$ distributions for the low 2-3 (left) and high $\geq 4$ (right) jet	
190	multiplicities after a full analysis selection and $H_T > 375$ requirement. . . . .	52
191	4.4. Pictorial depiction of the analysis strategy employed by the $\alpha_T$ search to	
192	increase sensitivity to a wide spectra of SUSY models. . . . .	55
193	4.5. Data/MC comparisons of key variables for the hadronic signal region. . . . .	61
194	4.6. Data/MC comparisons of key variables for the $\mu +$ jets selection. . . . .	64
195	4.7. Data/MC comparisons of key variables for the $\mu\mu +$ jets selection. . . . .	66
196	4.8. Data/MC comparisons of key variables for the $\gamma +$ jets selection. . . . .	67
197	4.9. QCD sideband regions, used for determination of $k_{\text{QCD}}$ . . . . .	69
198	4.10. Tagging efficiencies of (a) b-jets, (b) c-jets, and (c) light-jets determined	
199	from all jets within each $H_T$ category. . . . .	76
200	4.11. Sets of closure tests overlaid on top of the systematic uncertainty used for	
201	each of the five $H_T$ regions. . . . .	83

---

202	4.12. Signal efficiencies fo the Simplified Model Spectra (SMS) models (a) T1 203 and (b) T2. . . . .	85
204	5.1. Comparison of the observed yields and Standard Model (SM) expectations 205 given by the simultaneous fit in bins of $H_T$ for the (a) hadronic, (b) $\mu +$ 206 jets, (c) $\mu\mu +$ jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 0$ and 207 $n_{jet} \leq 3$ . . . . .	99
208	5.2. Comparison of the observed yields and SM expectations given by the 209 simultaneous fit in bins of $H_T$ for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ 210 jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 1$ and $n_{jet} \leq 3$ . . . . .	100
211	5.3. Comparison of the observed yields and SM expectations given by the 212 simultaneous fit in bins of $H_T$ for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ 213 jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 2$ and $n_{jet} \leq 3$ . . . . .	101
214	5.4. Comparison of the observed yields and SM expectations given by the 215 simultaneous fit in bins of $H_T$ for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ 216 jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 0$ and $n_{jet} \geq 4$ . . . . .	102
217	5.5. Comparison of the observed yields and SM expectations given by the 218 simultaneous fit in bins of $H_T$ for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ 219 jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 1$ and $n_{jet} \geq 4$ . . . . .	103
220	5.6. Comparison of the observed yields and SM expectations given by the 221 simultaneous fit in bins of $H_T$ for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ 222 jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 2$ and $n_{jet} \geq 4$ . . . . .	104
223	5.7. Comparison of the observed yields and SM expectations given by the 224 simultaneous fit in bins of $H_T$ for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ 225 jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 3$ and $n_{jet} \geq 4$ . . . . .	104
226	5.8. Comparison of the observed yields and SM expectations given by the 227 simultaneous fit in bins of $H_T$ for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ 228 jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} \geq 4$ and $n_{jet} \geq 4$ . . . . .	105
229	5.9. Production and decay modes for the various SMS models interpreted 230 within the analysis. . . . .	109
231	5.10. Upper limit of cross section at 95% CL as a function of $m_{\tilde{q}/\tilde{g}}$ and $m_{LSP}$ 232 for various SMS models. . . . .	110

---

233	6.1. The b-quark (a), c-quark (b), and light-quark (c) tagging efficiency as a function of jet $p_T$ , measured in simulation after the application of the $\alpha_T$ analysis $\mu + \text{jets}$ control sample selection, in the region $H_T > 375$ . . . . .	114
236	6.2. An example of a template fit with the defined Z0 (blue) and Z2 (red) templates to data within the low $n_b^{\text{reco}}$ control region (left). . . . .	115
238	6.3. Results of fitting the $Z = 0$ and $Z = 2$ templates in the $n_b^{\text{reco}} = 0\text{-}2$ control region to yields from simulation in the $\mu + \text{jets}$ control sample for the $H_T > 375 \text{ GeV}$ , $n_{\text{jet}} \geq 5$ category for all CSV working points. . . . .	118
241	6.4. Results of fitting the $Z = 0$ and $Z = 2$ templates in the $n_b^{\text{reco}} = 0\text{-}2$ control region to data from the $\mu + \text{jets}$ control sample, for the CSV medium working point, with a jet multiplicity $n_{\text{jet}} \geq 5$ , in all three $H_T$ categories. . . . .	121
244	6.5. Results of fitting the $Z = 0$ and $Z = 2$ templates in the $n_b^{\text{reco}} = 0\text{-}2$ control region to data from the hadronic signal selection, in the $n_{\text{jet}} \geq 5$ and $H_T > 375$ category for all CSV working points. . . . .	123
247	B.1. Leading jet matching efficiency as a function of the offline CaloJet $E_T$ . . . . .	130
248	B.2. Resolution plots of the leading offline Calo $E_T$ measured as a function of $\frac{(\text{L1 } E_T - \text{Offline } E_T)}{\text{Offline } E_T}$ for (a) low, (b) medium, and (c) high pile-up conditions. . . . .	132
250	B.3. Resolution plots of the leading off-line PF $E_T$ measured as a function of $\frac{(\text{L1 } E_T - \text{Offline } E_T)}{\text{Offline } E_T}$ for (a) low, (b) medium, and (c) high pile-up conditions. . . . .	134
252	B.4. $H_T$ resolution parameters in bins of Calo $H_T$ measured in the defined low, medium and high pile up conditions. . . . .	134
254	B.5. $H_T$ resolution parameters in bins of PF $H_T$ measured in the defined low, medium and high pile up conditions. . . . .	135
256	B.6. $\mathcal{H}_T$ resolution parameters in bins of Calo $\mathcal{H}_T$ measured in the defined low, medium and high pile up conditions. . . . .	135
258	B.7. $\mathcal{H}_T$ resolution parameters in bins of PF $\mathcal{H}_T$ measured in the defined low, medium and high pile up conditions. . . . .	135
260	C.1. $R_{\alpha_T}(H_T)$ and exponential fits for each of the data sideband regions. Fit is conducted between the $H_T$ region $275 < H_T < 575$ . . . . .	136

---

262	C.2. Sets of closure tests overlaid on top of the systematic uncertainty used for each of the five $H_T$ regions in the $2 \leq n_{jet} \leq 3$ jet multiplicity category for nominal and varied cross-sections. . . . .	137
265	C.3. Sets of closure tests overlaid on top of the systematic uncertainty used for each of the five $H_T$ regions in the $n_{jet} \geq 4$ jet multiplicity category for nominal and varied cross-sections. . . . .	137
268	D.1. Results of fitting the $Z = 0$ and $Z = 2$ templates in the $n_b^{reco} = 0\text{-}2$ control region to yields from simulation in the $\mu + \text{jets}$ control sample for the $H_T > 375$ GeV, $n_{jet} = 3$ category. . . . .	140
271	D.2. Results of fitting the $Z = 0$ and $Z = 2$ templates in the $n_b^{reco} = 0\text{-}2$ control region to yields from simulation in the $\mu + \text{jets}$ control sample for the $H_T > 375$ GeV, $n_{jet} = 4$ category. . . . .	141
274	D.3. Pull distributions of the normalisation parameter of each template, $\frac{(\theta - \hat{\theta})}{\sigma}$ . Distributions are constructed from $10^4$ pseudo-experiments generated by a gaussian distribution with width $\sigma$ , centred on the nominal template value of each point within the low $n_b^{reco}$ control region. . . . .	142
278	D.4. Results of fitting the $Z = 0$ and $Z = 2$ templates in the $n_b^{reco} = 0\text{-}2$ control region to data from the $\mu + \text{jets}$ control sample, for the CSV medium working point, with a jet multiplicity $n_{jet} = 3$ , in all three $H_T$ categories. . . . .	144
281	D.5. Results of fitting the $Z = 0$ and $Z = 2$ templates in the $n_b^{reco} = 0\text{-}2$ control region to data from the $\mu + \text{jets}$ control sample, for the CSV medium working point, with a jet multiplicity $n_{jet} = 4$ , in all three $H_T$ categories. . . . .	145
284	D.6. Results of fitting the $Z = 0$ and $Z = 2$ templates in the $n_b^{reco} = 0\text{-}2$ control region to data from the hadronic signal selection, in the $n_{jet} = 3$ and $H_T > 375$ category for all CSV working points. . . . .	146
287	D.7. Results of fitting the $Z = 0$ and $Z = 2$ templates in the $n_b^{reco} = 0\text{-}2$ control region to data from the hadronic signal selection, in the $n_{jet} = 4$ and $H_T > 375$ category for all CSV working points. . . . .	147

# **List of Tables**

291	2.1. The fundamental particles of the SM, with spin, charge and mass displayed.	6
292	3.1. Results of a cumulative EMG function fit to the turn-on curves for L1	
293	single jet triggers in 2012 Run Period C. . . . .	40
294	3.2. Results of a cumulative EMG function fit to the turn-on curves for L1	
295	single jet triggers in the 2012 run period B and C. . . . .	42
296	3.3. Results of a cumulative EMG function fit to the turn-on curves for $H_T$ in	
297	2012 run period B and C. . . . .	42
298	3.4. Results of a cumulative EMG function fit to the efficiency turn-on curves	
299	for L1 single jet triggers in the 2012 run period C, for low,medium and	
300	high pile-up conditions. . . . .	44
301	3.5. Results of a cumulative EMG function fit to the efficiency turn-on curves	
302	for Level-1 single jet triggers in the 2012 run period C, for low,medium	
303	and high pile-up conditions. . . . .	45
304	4.1. A summary of the SMS models interpreted in this analysis, involving both	
305	direct (D) and gluino-induced (G) production of squarks and their decays.	48
306	4.2. Muon identification criteria used within the analysis for selection/veto	
307	purposes in the muon control/signal selections. . . . .	56
308	4.3. Photon identification criteria used within the analysis for selection/veto	
309	purposes in the $\gamma$ + jets control/signal selections. . . . .	57
310	4.4. Electron identification criteria used within the analysis for veto purposes.	57
311	4.5. Noise filters that are applied to remove spurious and non-physical $\cancel{E}_T$	
312	signatures within the CMS detector. . . . .	58

313	4.6. Jet thresholds used in the three $H_T$ regions of the analysis. . . . .	59
314	4.7. Best fit values for the parameters $k_{\text{QCD}}$ obtained from sideband regions	
315	B,C <sub>1</sub> ,C <sub>2</sub> ,C <sub>3</sub> . . . . .	70
316	4.8. Measured efficiencies of the $H_T$ and $\alpha_T$ legs of the HT and HT_alphaT	
317	triggers in independent analysis bins. . . . .	71
318	4.9. k-factors calculated for different EWK processes. . . . .	72
319	4.10. Comparing yields in simulation within the $\mu + \text{jets}$ selection determined	
320	from the formula method described in Equation (4.11), and that taken	
321	directly from simulation. . . . .	75
322	4.11. The absolute change in the Transfer Factor (TF)s used to predict the	
323	entire signal region SM background, using the $\mu + \text{jets}$ control sample	
324	when the systematic uncertainties of the data to simulation scale factors	
325	are varied by $\pm 1\sigma$ . . . . .	77
326	4.12. A summary of the results obtained from zeroeth order polynomial (i.e.	
327	a constant) and linear fits to five sets of closure tests performed in the	
328	$2 \geq n_{\text{jet}} \geq 3$ category. . . . .	80
329	4.13. A summary of the results obtained from zeroeth order polynomial (i.e.	
330	a constant) and linear fits to five sets of closure tests performed in the	
331	$n_{\text{jet}} \geq 4$ category. . . . .	81
332	4.14. A summary of the results obtained from zeroeth order polynomial (i.e. a	
333	constant) and linear fits to three sets of closure tests performed between	
334	the $2 \leq n_{\text{jet}} \leq 3$ and $n_{\text{jet}} \geq 4$ categories. . . . .	81
335	4.15. Calculated systematic uncertainties for the five $H_T$ regions, determined	
336	from the closure tests. . . . .	82
337	4.16. Estimates of systematic uncertainties on the signal efficiency (%) for	
338	various SMS models when considering points in the region near to the	
339	diagonal . . . . .	90
340	4.17. Estimates of systematic uncertainties on the signal efficiency (%) for	
341	various SMS models when considering points in the region near to the	
342	diagonal . . . . .	90

---

343	4.18. The systematic parameters used in $H_T$ bins. . . . .	94
344	4.19. Nuisance parameters used within the different hadronic signal bins of the	
345	analysis . . . . .	96
346	5.1. Summary of control samples used by each fit results, and the Figures in	
347	which they are displayed. . . . .	98
348	5.2. Comparison of the measured yields in each $H_T$ , $n_{jet}$ and $n_b^{reco}$ jet multiplicity bins for the hadronic sample with the SM expectations and combined	
349	statistical and systematic uncertainties given by the simultaneous fit. . .	98
350		
351	5.3. A table representing the SMS models interpreted within the analysis. . .	108
352	6.1. Typical underlying b-quark content of different SM processes which are	
353	common to many SUSY searches. . . . .	112
354	6.2. Summary of the fit predictions in the $n_b^{reco}$ signal region after combination	
355	of the $n_{jet} = 3, = 4, \geq 5$ categories compared against yields taken directly	
356	from simulation. The predictions are extrapolated from a $n_b^{reco} = 0, 1,$	
357	2 control region and simulation yields are normalised to an integrated	
358	luminosity of $10 \text{ fb}^{-1}$ . . . . .	119
359	6.3. Summary of the fit predictions in the $n_b^{reco}$ signal region of the $\mu + \text{jets}$	
360	control sample, after combination of the $n_{jet} = 3, = 4, \geq 5$ categories.. The	
361	predictions are extrapolated from a $n_b^{reco} = 0, 1, 2$ control region using	
362	$11.4 \text{ fb}^{-1}$ of $\sqrt{s} = 8\text{TeV}$ data. . . . .	121
363	6.4. Summary of the fit predictions in the $n_b^{reco}$ signal region of the $\alpha_T$ hadronic	
364	signal selection, after combination of the $n_{jet} = 3, = 4, \geq 5$ categories. The	
365	predictions are extrapolated from a $n_b^{reco} = 0, 1, 2$ control region using	
366	$11.7 \text{ fb}^{-1}$ of $\sqrt{s} = 8\text{TeV}$ data. . . . .	124
367	A.1. Criteria for a reconstructed jet to pass the loose calorimeter jet id. . . . .	128
368	A.2. Criteria for a reconstructed jet to pass the loose PF jet id. . . . .	129
369	A.3. Criteria for a vertex in an event to be classified as a 'good' reconstructed	
370	primary vertex. . . . .	129

371	B.1. Results of a cumulative EMG function fit to the turn-on curves for the	
372	matching efficiency of the leading jet in an event to a Level-1 jet in run	
373	2012C and 2012B data. . . . .	131
374	C.1. Translation factors constructed from the $\mu + \text{jets}$ control sample and signal	
375	selection MC, to predict yields for the $W + \text{jets}$ and $t\bar{t}$ back-grounds in	
376	the signal region. . . . .	138
377		

# Chapter 1.

## <sup>378</sup> Introduction

<sup>379</sup> During the 20th century, great advances were made in the human understanding of  
<sup>380</sup> the universe, its origins, its future and its composition. The Standard Model (**SM**)  
<sup>381</sup> first formulated in the 1960s is one of the crowning achievements in science's quest to  
<sup>382</sup> explain the most fundamental processes and interactions that make up our universe. It  
<sup>383</sup> has provided a highly successful explanation for a wide range of phenomena in Particle  
<sup>384</sup> Physics and has stood up to extensive experimental scrutiny [1].

<sup>385</sup> Despite its success it is not a complete theory, with significant questions remaining  
<sup>386</sup> unanswered. It describes only three of the four known forces with gravity not incorporated  
<sup>387</sup> within the framework of the **SM**. Cosmological experiments infer that just  $\sim 5\%$  of the  
<sup>388</sup> observable universe exists as matter, with elusive "Dark Matter" accounting for a further  
<sup>389</sup>  $\sim 27\%$  [2]. However no particle predicted by the **SM** is able to account for it. At higher  
<sup>390</sup> energy scales, the (non-)unification of the fundamental forces point to problems with the  
<sup>391</sup> **SM** at least at higher energies not yet probed experimentally.

<sup>392</sup> Many theories exist as extensions to the **SM**, predicting a range of observables that can  
<sup>393</sup> be detected at the Large Hadron Collider (**LHC**) of which SUperSYmmetry (**SUSY**) is  
<sup>394</sup> one such example. It predicts a new symmetry of nature in which all current particles  
<sup>395</sup> in the **SM** would have a corresponding supersymmetric partner. Common to most  
<sup>396</sup> Supersymmetric theories is a stable, weakly interacting Lightest Supersymmetric Partner  
<sup>397</sup> (**LSP**), which has the properties of a possible dark matter candidate. The **SM** and the  
<sup>398</sup> main principles of Supersymmetric theories are outlined in Chapter 2, with emphasis  
<sup>399</sup> placed on how experimental signatures of **SUSY** may reveal themselves in proton collisions  
<sup>400</sup> at the **LHC**.

- 401 The experimental goal of the **LHC** is to further test the framework of the **SM**, exploring the  
402 TeV mass scale for the first time, and to seek a connection between the particles produced  
403 in proton collisions and dark matter. The first new discovery by this extraordinary  
404 machine was announced on the 4th of July 2012. The long-awaited discovery was the  
405 culmination of decades of experimental endeavours in the search for the Higgs boson,  
406 which provided an answer to the mechanism of electroweak symmetry breaking within  
407 the **SM** [3][4].
- 408 This discovery was made possible through the combination of data taken by the Compact  
409 Muon Solenoid (**CMS**) and A Toroidal LHC ApparatuS (**ATLAS**), two multipurpose  
410 detectors located on the **LHC** ring. An experimental description of the **CMS** detector  
411 and the **LHC** is described in Chapter 3, including some of the object reconstruction used  
412 by **CMS** in searches for **SUSY** signatures.
- 413 The performance of the **CMS** Level-1 single jet and energy sum triggers is also bench-  
414 marked within this chapter. The Level-1 trigger is the first line of the **CMS** trigger  
415 system and is of paramount importance to the collection of physics events. A change in  
416 the jet clustering algorithm, via the introduction of a jet seed threshold, was introduced  
417 approximately half way through the data taking period. The aim of this change, was  
418 to reduce the rate at which collisions not of interest to physics analysis were recorded,  
419 whilst avoiding impact to the overall performance of the triggers.
- 420 Chapter 4, contains a description of the search for direct evidence of the production of  
421 supersymmetric particles at the **LHC**. The main basis of the search centres around the  
422 kinematic dimensionless  $\alpha_T$  variable; which provides a strong rejection of backgrounds  
423 with fake missing transverse energy signatures, whilst maintaining good sensitivity to  
424 a variety of **SUSY** topologies. The author's work (as an integral part of the analysis  
425 group) is documented in detail, and has culminated in numerous publications over the  
426 past two years, the latest results having been published in the European Physical Journal  
427 C (**EPJC**) [5].
- 428 The author in particular has played a major role in the extension of the  $\alpha_T$  analysis into  
429 additional b-tagged jet (jets identified as originating from a b-quark) and jet multiplicity  
430 dimensions, increasing the sensitivity of the analysis to a range of **SUSY** topologies.  
431 Additionally, the author has worked extensively on increasing the statistical precision  
432 of the data driven electroweak predictions through analytical techniques. This included  
433 work on developing the derivation of data driven systematic uncertainties through the  
434 establishment of closure tests within the control samples of the analysis.

- 435 The compatibility of the data collected for the  $\alpha_T$  search with a **SM** only hypothesis is  
436 documented in Chapter 5. In the absence of an observed excess, interpretations of the  
437 data within the framework of a variety of Simplified Model Spectra (**SMS**), describing an  
438 array of possible **SUSY** event topologies are made.
- 439 Finally, a method to search for gluino mediated **SUSY** signatures rich in top and bottom  
440 flavoured jet final states, is introduced in Chapter 6. These particular **SUSY** topologies  
441 are increasingly of interest to physicists in light of the discovery of the Higgs boson. A  
442 parametrisation of the b-tagged jet distribution for different electroweak processes is  
443 used to establish template shapes, which are then fitted at low b-tagged jet multiplicity,  
444 to extrapolate an expected **SM** background of 3 and 4 b-tagged jet events within an  
445 event sample. The  $\alpha_T$  control and hadronic signal event selections are used to validate  
446 the functionality of this template method in both data and simulation. Background  
447 predictions within the hadronic signal region are compared to those presented in Chapter 5,  
448 with the intention of serving as a independent crosscheck of the estimated **SM** backgrounds  
449 from the  $\alpha_T$  search.

450 Natural units are used throughout this thesis in which  $\hbar = c = 1$ .

# Chapter 2.

## <sup>451</sup> A Theoretical Overview

<sup>452</sup> Within this chapter, a brief introduction and background to the **SM** is given. Its success  
<sup>453</sup> as a rigorously tested and widely accepted theory is discussed as are its deficiencies which  
<sup>454</sup> lead to the argument that this theory is not a complete description of our universe. The  
<sup>455</sup> motivations for new physics at the TeV scale and in particular Supersymmetric theories  
<sup>456</sup> are outlined within Section (2.3). The chapter concludes with how an experimental  
<sup>457</sup> signature of such theories can be produced and observed at the **LHC** in Section (2.4).

### <sup>458</sup> 2.1. The Standard Model

<sup>459</sup> The **SM** is the name given to the relativistic Quantum Field Theory (**QFT**), where  
<sup>460</sup> particles are represented as excitations of fields, which describe the interactions and  
<sup>461</sup> properties of all the known elementary particles [6][7][8][9]. It is a renormalisable field  
<sup>462</sup> theory which contains three symmetries:  $SU(3)$  for colour charge;  $SU(2)$  for weak isospin  
<sup>463</sup> and;  $U(1)$  relating to weak hyper charge, which requires its Lagrangian  $\mathcal{L}_{SM}$  to be  
<sup>464</sup> invariant under local gauge transformation.

<sup>465</sup> Within the **SM** theory, matter is composed of spin  $\frac{1}{2}$  fermions that interact with each  
<sup>466</sup> other via the exchange of spin-1 gauge bosons. A summary of the known fundamental  
<sup>467</sup> fermions and bosons is given in Table 2.1.

<sup>468</sup> Fermions are separated into quarks and leptons of which only quarks interact with the  
<sup>469</sup> strong nuclear force. Quarks unlike leptons are not seen as free particles in nature, but  
<sup>470</sup> rather exist only within baryons, which are composed of three quarks with an overall  
<sup>471</sup> integer charge, and quark-anti-quark pairs called mesons. Both leptons and quarks are

Particle	Symbol	Spin	Charge	Mass (GeV)
<b>First Generation Fermions</b>				
Electron Neutrino	$\nu_e$	$\frac{1}{2}$	0	$< 2.2 \times 10^{-6}$
Electron	e	$\frac{1}{2}$	-1	$0.51 \times 10^{-3}$
Up Quark	u	$\frac{1}{2}$	$\frac{2}{3}$	$2.3^{+0.7}_{-0.5} \times 10^{-3}$
Down Quark	d	$\frac{1}{2}$	$-\frac{1}{3}$	$4.8^{+0.7}_{-0.3} \times 10^{-3}$
<b>Second Generation Fermions</b>				
Muon Neutrino	$\nu_\mu$	$\frac{1}{2}$	0	-
Muon	$\mu$	$\frac{1}{2}$	-1	$1.05 \times 10^{-3}$
Charm Quark	c	$\frac{1}{2}$	$\frac{2}{3}$	$1.275 \pm 0.025$
Strange Quark	s	$\frac{1}{2}$	$-\frac{1}{3}$	$95 \pm 5 \times 10^{-3}$
<b>Third Generation Fermions</b>				
Tau Neutrino	$\nu_\tau$	$\frac{1}{2}$	0	-
Tau	$\tau$	$\frac{1}{2}$	-1	1.77
Top Quark	t	$\frac{1}{2}$	$\frac{2}{3}$	$173.5 \pm 0.8$
Bottom Quark	b	$\frac{1}{2}$	$-\frac{1}{3}$	$4.65 \pm 0.03$
<b>Gauge Bosons</b>				
Photon	$\gamma$	1	0	0
W Boson	$W^\pm$	1	$\pm 1$	$80.385 \pm 0.015$
Z Boson	Z	1	0	$91.187 \pm 0.002$
Gluons	g	1	0	0
Higgs Boson	H	0	0	$125.3 \pm 0.5$ [4]

**Table 2.1:** The fundamental particles of the SM, with spin, charge and mass displayed. Latest mass measurements taken from [1].

<sup>472</sup> grouped into three generations which have the same properties, but with ascending mass  
<sup>473</sup> in each subsequent generation.

<sup>474</sup> The gauge bosons mediate the interactions between fermions. The field theories of  
<sup>475</sup> Quantum Electro-Dynamics (QED) and Quantum Chromo-Dynamics (QCD), yield  
<sup>476</sup> massless mediator bosons, the photon and eight coloured gluons which are consequences  
<sup>477</sup> of the gauge invariance of those theories (detailed in Section (2.1.1)).

<sup>478</sup> The unification of the electromagnetic and weak-nuclear forces into the current Elec-  
<sup>479</sup> troweak theory yield the weak gauge bosons  $W^\pm$  and Z through the mixing of the asso-  
<sup>480</sup> ciated gauge fields. The force carriers of this theory were experimentally detected by the  
<sup>481</sup> observation of weak neutral current. This was first discovered in 1973 by the Gargamelle  
<sup>482</sup> bubble chamber located at European Organisation for Nuclear Research (CERN) [10].  
<sup>483</sup> The masses of the weak gauge bosons were measured by the UA1 and U2 experiments at  
<sup>484</sup> the Super Proton Synchrotron (SPS) collider in 1983 [11][12].

### <sup>485</sup> 2.1.1. Gauge Symmetries of the SM

- <sup>486</sup> Symmetries are of fundamental importance in the description of physical phenomena.  
<sup>487</sup> Noether's theorem states that for a dynamical system, the consequence of any symmetry  
<sup>488</sup> is an associated conserved quantity [13]. Invariance under translations, rotations, and  
<sup>489</sup> Lorentz transformations in physical systems lead to the conservation of momentum,  
<sup>490</sup> energy and angular momentum.
- <sup>491</sup> In the **SM**, a quantum theory described by Lagrangian formalism, the weak, strong and  
<sup>492</sup> electromagnetic interactions are described in terms of “gauge theories”. A gauge theory  
<sup>493</sup> possesses invariance under a set of “local transformations”, which are transformations  
<sup>494</sup> whose parameters are space-time dependent. The requirement of gauge invariance within  
<sup>495</sup> the **SM** necessitates the introduction of force-mediating gauge bosons, and interactions  
<sup>496</sup> between fermions and the bosons themselves. Given the nature of the topics covered by  
<sup>497</sup> this thesis, the formulation of Electroweak Sector (**EWK**) within the **SM** Lagrangian is  
<sup>498</sup> reviewed within this section.
- <sup>499</sup> The simplest example of the application of the principle of local gauge invariance within  
<sup>500</sup> the **SM** is in Quantum Electro-Dynamics (**QED**), the consequences of which require a  
<sup>501</sup> massless photon field [14][15].
- <sup>502</sup> The free Dirac Lagrangian can be first written as

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi, \quad (2.1)$$

- <sup>503</sup> where  $\psi$  represents a free non interacting fermionic field, with the matrices  $\gamma^\mu$ ,  $\mu \in 0, 1, 2, 3$   
<sup>504</sup> defined by the anti commutator relationship  $\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2\eta^{\mu\nu} I_4$ , with  $\eta^{\mu\nu}$  being the  
<sup>505</sup> flat space-time metric  $(+, -, -, -)$ , and  $I_4$  the  $4 \times 4$  identity matrix.
- <sup>506</sup> Under a local U(1) abelian gauge transformation, in which  $\psi$  transforms as

$$\psi(x) \rightarrow \psi'(x) = e^{i\theta(x)}\psi(x) \quad \bar{\psi}(x) \rightarrow \bar{\psi}'(x) = e^{i\theta(x)}\bar{\psi}(x) \quad (2.2)$$

- <sup>507</sup> the kinetic term of the Lagrangian will not remain invariant, due to the partial derivative  
<sup>508</sup> interposed between the  $\bar{\psi}$  and  $\psi$  yielding

$$\partial_\mu \psi \rightarrow e^{i\theta(x)} \partial_\mu \psi + ie^{i\theta(x)} \psi \partial_\mu \theta. \quad (2.3)$$

To ensure that  $\mathcal{L}$  remains invariant, a modified derivative,  $D_\mu$ , that transforms covariantly under phase transformations is introduced. In doing this, a vector field  $A_\mu$  with transformation properties that cancel out the unwanted term in (2.3) must also be included,

$$D_\mu \equiv \partial_\mu - ieA_\mu, \quad A_\mu \rightarrow A_\mu + \frac{1}{e} \partial_\mu \theta. \quad (2.4)$$

Invariance of the Lagrangian is then achieved by replacing  $\partial_\mu$  by  $D_\mu$ :

$$\begin{aligned} \mathcal{L} &= i\bar{\psi}\gamma^\mu D_\mu \psi - m\bar{\psi}\psi \\ &= \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi + e\bar{\psi}\gamma^\mu \psi A_\mu \end{aligned} \quad (2.5)$$

An additional interaction term is now present in the Lagrangian, coupling the Dirac particle to this vector field, which is interpreted as the photon in QED. To regard this new field as the physical photon field, a term corresponding to its kinetic energy must be added to the Lagrangian from Equation (2.5). Since this term must also be invariant under the conditions of Equation (2.4), it is defined in the form  $F_{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$ .

This then leads to the Lagrangian of QED,

$$\mathcal{L}_{QED} = \underbrace{i\bar{\psi}\gamma^\mu \partial_\mu \psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}}_{\text{kinetic term}} + \underbrace{m\bar{\psi}\psi}_{\text{mass term}} + \underbrace{e\bar{\psi}\gamma^\mu \psi A_\mu}_{\text{interaction term}}. \quad (2.6)$$

Within the Lagrangian there remains no mass term of the form  $m^2 A_\mu A^\mu$ , which is prohibited by gauge invariance. This implies that the gauge particle, the photon, must be massless.

### 523 2.1.2. The Electroweak Sector and Electroweak Symmetry

#### 524 Breaking

525 The same application of gauge symmetry and the requirement of local gauge invariance  
 526 can be used to unify QED and the Weak force in the Electroweak Sector (EWK).

527 The nature of EWK interactions is encompassed within a Lagrangian invariant under  
 528 transformations of the group  $SU(2)_L \times U(1)_Y$ .

529 The weak interactions from experimental observation [16] are known to violate parity  
 530 and are therefore not symmetric under interchange of left and right helicity fermions.

531 Thus, within the SM the left- and right-handed parts of these fermion fields are treated  
 532 separately. A fermion field is then split into two left and right handed chiral components,  
 533  $\psi = \psi_L + \psi_R$ , where  $\psi_{L/R} = (1 \pm \gamma^5)\psi$ .

534 The  $SU(2)_L$  group is the special unitary group of  $2 \times 2$  matrices,  $U$ , satisfying  $UU^\dagger = I$   
 535 and  $\det(U) = 1$ . It may be written in the form  $U = e^{-i\omega_i T_i}$ , with the generators of the  
 536 group written as  $T_i = \frac{1}{2}\tau_i$  where  $\tau_i$ ,  $i \in 1,2,3$  are the  $2 \times 2$  Pauli matrices

$$\tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (2.7)$$

537 The generators of the group form a non Abelian group obeying the commutation relation  
 538  $[T^a, T^b] \equiv if^{abc}T^c \neq 0$ . The gauge fields that accompany this group are represented by  
 539  $\hat{W}_\mu = (\hat{W}_\mu^1, \hat{W}_\mu^2, \hat{W}_\mu^3)$  and act only on the left handed component of the fermion field  $\psi_L$ .

540 One additional generator,  $Y$ , which represents the hypercharge of the particle under  
 541 consideration is introduced through the  $U(1)_Y$  group acting on both components of the  
 542 fermion field, with an associated vector boson field  $\hat{B}_\mu$ .

543 The  $SU(2)_L \times U(1)_Y$  transformations of the left- and right-handed components of  $\psi$  are  
 544 summarised by,

$$\begin{aligned} \chi_L &\rightarrow \chi'_L = e^{i\theta(x) \cdot T + i\theta(x)Y} \chi_L, \\ \psi_R &\rightarrow \psi'_R = e^{i\theta(x)Y} \psi_R, \end{aligned} \quad (2.8)$$

545 where the left-handed fermions form isospin doubles  $\chi_L$  and the right handed fermions  
 546 are isosinglets  $\psi_R$ . For the first generation of leptons and quarks this represents

$$\chi_L = \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, \quad \begin{pmatrix} u \\ d \end{pmatrix}_L, \\ \psi_R = e_R, \quad u_R, d_R. \quad (2.9)$$

547 Local gauge invariance within  $\mathcal{L}_{EWK}$  is once again imposed by modifying the covariant  
 548 derivative

$$D_\mu = \partial_\mu - \frac{ig}{2}\tau^i W_\mu^i - \frac{ig'}{2}YB_\mu, \quad (2.10)$$

549 where  $g$  and  $g'$  are the coupling constant of the  $SU(2)_L$  and  $U(1)_Y$  groups respectively.  
 550 Taking the example of the first generation of fermions defined in Equation (2.9), with input  
 551 hypercharge values of -1 and -2 for  $\chi_L$  and  $e_R$  respectively, would lead to a Lagrangian  
 552  $\mathcal{L}_1$  of the form,

$$\mathcal{L}_1 = \bar{\chi}_L \gamma^\mu [i\partial_\mu - g \frac{1}{2} \tau \cdot W_\mu - g' (-\frac{1}{2}) B_\mu] \chi_L \\ + \bar{e}_R \gamma^\mu [i\partial_\mu - g' (-1) B_\mu] e_R - \frac{1}{4} W_{\mu\nu} \cdot W^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}. \quad (2.11)$$

553 As in QED, these additional gauge fields introduce field strength tensors  $B_{\mu\nu}$  and  $W_{\mu\nu}$ ,

$$\hat{B}_{\mu\nu} = \partial_\mu \hat{B}_\nu - \partial_\nu \hat{B}_\mu \quad (2.12)$$

$$\hat{W}_{\mu\nu} = \partial_\mu \hat{W}_\nu - \partial_\nu \hat{W}_\mu - g \hat{W}_\mu \times \hat{W}_\mu \quad (2.13)$$

554 corresponding to the kinetic energy and self coupling of the  $W_\mu$  fields and the kinetic  
 555 energy term of the  $B_\mu$  field.

556 None of these gauge bosons are physical particles, and instead linear combinations of  
557 these gauge bosons make up  $\gamma$  and the W and Z bosons, defined as

$$W^\pm = \frac{1}{\sqrt{2}} (W_\mu^1 \mp iW_\mu^2), \quad \begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & -\sin\theta_W \\ \sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix}, \quad (2.14)$$

558 where the mixing angle,  $\theta_w = \tan^{-1} \frac{g'}{g}$ , relates the coupling of the neutral weak and  
559 electromagnetic interactions.

560 As in the case of the formulation of the QED Lagrangian there remains no mass term  
561 for the photon. However contrary to experimental measurement, this is also the case  
562 for the W, Z and fermions in the Lagrangian. Any explicit introduction of mass terms  
563 would break the symmetry of the Lagrangian, and instead mass terms can be introduced  
564 through spontaneous breaking of the EWK symmetry via the Higgs mechanism.

565 The Higgs mechanism induces spontaneous symmetry breaking through the introduction  
566 of a complex scalar SU(2) doublet field  $\phi$ , which attains a non-zero Vacuum Expectation  
567 Value (VEV) [17][18][19][20].

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \quad \text{with} \quad \begin{aligned} \phi^+ &\equiv (\phi_1 + i\phi_2)/\sqrt{2} \\ \phi^0 &\equiv (\phi_3 + i\phi_4)/\sqrt{2} \end{aligned} \quad (2.15)$$

568 The Lagrangian defined in Equation (2.11) attains an additional term  $\mathcal{L}_{Higgs}$  of the form

$$\mathcal{L}_{Higgs} = \overbrace{(D_\mu \phi)^\dagger (D^\mu \phi)}^{\text{kinetic}} - \overbrace{\mu^2 \phi^\dagger \phi - \lambda (\phi^\dagger \phi)^2}^{\text{potential } V(\phi)} \quad (\mu^2, \lambda) > 0 \in \mathbb{R},$$

$$\mathcal{L}_{SM} = \mathcal{L}_{EWK} + \mathcal{L}_{Higgs}, \quad (2.16)$$

569 where the covariant derivative  $D_\mu$  is that defined in Equation (2.10). The last two terms  
570 of  $\mathcal{L}_{Higgs}$  correspond to the Higgs potential, in which real positive values of  $\mu^2$  and  $\lambda$  are  
571 required to ensure the generation of masses for the bosons and leptons. The minimum of

572 this potential is found at  $\phi^\dagger \phi = \frac{1}{2}(\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2) = \mu^2/\lambda = v^2$ , where  $v$  represents  
573 the **VEV**.

574 The ground state of the  $\phi$  field is defined to be consistent with the  $V(\phi)$  minimum. By  
575 then expanding around a ground state chosen to maintain an unbroken electromagnetic  
576 symmetry. This preserves a zero photon mass [21] and leads to

$$\phi_0 = \sqrt{\frac{1}{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}, \quad \phi(x) = e^{i\tau \cdot \theta(x)/v} \sqrt{\frac{1}{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}, \quad (2.17)$$

577 where the fluctuations from the vacuum  $\phi_0$  are parametrized in terms of four real fields,  
578  $\theta_1, \theta_2, \theta_3$  and  $h(x)$ .

579 Choosing to gauge away the three massless Goldstone boson fields by setting  $\theta(x)$  to  
580 zero and substituting  $\phi(x)$  back into kinetic term of  $\mathcal{L}_{Higgs}$  from Equation (2.16) leads  
581 to mass terms for the  $W^\pm$  and  $Z$  bosons

$$(D_\mu \phi)^\dagger (D^\mu \phi) = \frac{1}{2} (\partial_\mu h)^2 + \frac{g^2 v^2}{2} W_\mu^+ W^{-\mu} + \frac{v^2 g^2}{8 \cos^2 \theta_w} Z_\mu Z^\mu + 0 A_\mu A^\mu, \quad (2.18)$$

582 where the relations between the physical and electroweak gauge fields from Equation  
583 (2.14) are used. The  $W^\pm$  and  $Z$  boson masses can then be determined to be

$$M_W = \frac{1}{2} g v \quad M_Z = \frac{1}{2} \frac{g v}{\cos \theta_w}. \quad (2.19)$$

584 This mechanism is also used to generate fermion masses by introducing a Yukawa coupling  
585 between the fermions and the  $\phi$  field [22], with the coupling strength of a particle to the  
586  $\phi$  field governing its mass. Additionally, a scalar boson  $h$  with mass  $m_h = v \sqrt{\frac{\lambda}{2}}$ , is also  
587 predicted as a result of this spontaneous symmetry breaking. This became known as  
588 the Higgs boson. Its discovery by the **CMS** and **ATLAS** experiments in 2012 is the first  
589 direct evidence to support this method of mass generation within the **SM**.

590 **2.2. Motivation for Physics beyond the Standard  
591 Model**

592 As has been described, the **SM** has proven to be a very successful theory, predicting the  
593 existence of the  $W^\pm$  and  $Z$  bosons and the top quark long before they were experimentally  
594 observed. However, the theory does not accurately describe all observed phenomena and  
595 has some fundamental theoretical flaws that hint at the need for additional extensions to  
596 the current theory.

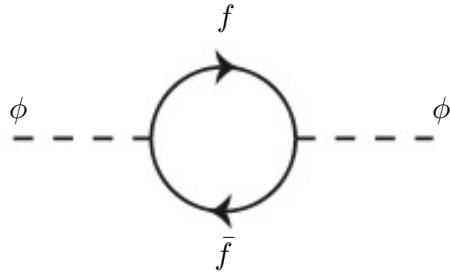
597 On a theoretical level, the **SM** is unable to incorporate the gravitational interactions of  
598 fundamental particles within the theory. Whilst at the electroweak energy scales the  
599 relative strength of gravity is negligible compared to the other three fundamental forces,  
600 at much higher energy scales,  $M_{\text{planck}} \sim 10^{18} \text{GeV}$ , quantum gravitational effects become  
601 increasingly dominant. The failure to reconcile gravity within the **SM**, demonstrates that  
602 the **SM** must become invalid at some higher energy scale.

603 Other deficiencies with the **SM** include the fact that the predicted rate of Charge-Parity  
604 violation does not account for the matter dominated universe which we inhabit, and  
605 that the **SM** prediction of a massless neutrino conflicts with the observation of neutrino  
606 flavour mixing, attributed to mixing between neutrino mass eigenstates [23][24].

607 Perhaps one of the most glaring gaps in the predictive power of the **SM** is that there  
608 exists no candidate to explain the cosmic dark matter observed in galactic structures  
609 through indirect techniques; including gravitational lensing and measurement of the  
610 orbital velocity of stars at galactic edges. Any such candidate must be very weakly  
611 interacting but must also be stable, owing to the lack of direct detection of the decay  
612 products of such a process. Therefore, a predicted stable dark matter candidate is one of  
613 the main obstacles to address for any Beyond Standard Model (**BSM**) physics model.

614 The recent discovery of the Higgs boson, whilst a significant victory for the predictive  
615 power of the **SM**, brings with it still unresolved questions. This issue is commonly  
616 described as the “hierarchy problem”.

617 In the absence of new physics between the TeV and Planck scale, calculating beyond  
618 tree-level contributions to the Higgs mass term given by its self interaction, results in  
619 divergent terms that push the Higgs mass up to the planck mass  $M_{\text{planck}}$ .



**Figure 2.1:** One loop quantum corrections to the Higgs squared mass parameter  $m_h^2$  due to a fermion.

620 This can be demonstrated by considering the one loop quantum correction to the Higgs  
 621 mass with a fermion  $f$ , shown in Figure 2.1 with mass  $m_f$ . The Higgs field couples to  $f$   
 622 with a term in the Lagrangian  $-\lambda_f h \bar{f} f$ , yielding a correction of the form [25],

$$\delta m_h^2 = -\frac{|\lambda_f|^2}{8\pi^2} \Lambda^2 + \dots, \quad (2.20)$$

623 where  $\lambda_f$  represents the coupling strength for each type of fermion  $\propto m_f$ , and  $\Lambda$  the  
 624 cutoff energy scale at which the **SM** ceases to be a valid theory.

625 To recover the mass of the now discovered Higgs boson would require a fine-tuning of  
 626 the parameters to cancel out these mass corrections of the Higgs mass, to the scale of  
 627 30 orders of magnitude. This appears as an unnatural solution to physicists and it is  
 628 this hierarchy problem that provides one of the strongest motivations for the theory of  
 629 SUperSYmmetry (**SUSY**).

### 630 2.3. Supersymmetry Overview

631 Supersymmetry provides potential solutions to many of the issues raised in the previous  
 632 section. It provides a dark matter candidate, can explain baryogenesis in the early  
 633 universe and also provides an elegant solution to the hierarchy problem [26][27][28][29].  
 634 At its heart it represents a new space-time symmetry that relates fermions and bosons.  
 635 This symmetry converts bosonic states into fermionic states, and vice versa, see Equation  
 636 (2.21),

$$Q|\text{Boson}\rangle = |\text{Fermion}\rangle \quad Q|\text{Fermion}\rangle = |\text{Boson}\rangle, \quad (2.21)$$

637 where the operator  $Q$  is the generator of these transformations. Quantum field theories  
 638 which are invariant under such transformations are called supersymmetric.

639 This symmetry operator therefore acts upon a particle's spin altering it by a half integer  
 640 value. The consequences of the application of this additional space-time symmetry  
 641 introduce a new rich phenomenology. For example, in supersymmetric theories both  
 642 the left handed  $SU(2)$  doublet and right handed singlet of fermions will have a spin-0  
 643 superpartner containing the same electric charge, weak isospin, and colour as its **SM**  
 644 partner. In the case of leptons  $(\nu_l, l)_L$ , they will have two superpartners, a sneutrino  $\tilde{\nu}_l{}_L$   
 645 and a slepton  $\tilde{l}_L$ , whilst the singlet  $l_R$  also has a superpartner slepton  $\tilde{l}_R$ .

646 Each particle in a supersymmetric theory is paired together with their superpartners as a  
 647 result of these supersymmetric transformations in what is called a supermultiplet. These  
 648 superpartners will then consequently also contribute to the corrections to the Higgs mass.  
 649 Bosonic and fermionic loops contributing to the correction appear with opposite signs,  
 650 and therefore cancellation of these divergent terms will stabilise the Higgs mass, solving  
 651 the hierarchy problem [30][31].

652 One of the simplest forms of **SUSY**, is to simply have a set of **SM** supersymmetric partners  
 653 with the same mass and interactions as their counterparts. However, the current lack  
 654 of any experimental evidence for that predicted sparticle spectrum implies **SUSY** must  
 655 be a broken symmetry in which any sparticle masses must be greater than their **SM**  
 656 counterparts.

657 There exists many techniques which can induce supersymmetric breaking [32][33][34]. Of  
 658 particular interest to experimental physicists are those at which the breaking scale is  
 659 of an order that is experimentally accessible to the **LHC** i.e.  $\sim$  TeV scale. Whilst  
 660 there is no requirement for supersymmetric breaking to occur at this energy scale, for  
 661 supersymmetry to provide a solution to the hierarchy problem, it is necessary for this  
 662 scale to not differ too drastically from the **EWK** scale [35][36].

### <sup>663</sup> 2.3.1. R-Parity

<sup>664</sup> Supersymmetric theories can also present a solution to the dark matter problem. These  
<sup>665</sup> theories contain a stable Lightest Supersymmetric Partner (**LSP**), which match the  
<sup>666</sup> criteria of a Weakly Interacting Massive Particle (**WIMP**) required by cosmological  
<sup>667</sup> observation when R-parity is conserved.

<sup>668</sup> Baryon (B) and Lepton (L) number conservation is forbidden in the **SM** by renormal-  
<sup>669</sup> isability requirements. The violation of Baryon or Lepton number results in a proton  
<sup>670</sup> lifetime much shorter than those set by experimental limits [37]. Another symmetry  
<sup>671</sup> called R-parity is then often introduced to **SUSY** theories to maintain baryon and lepton  
<sup>672</sup> conservation.

<sup>673</sup> R-parity is described by the equation

$$R_P = (-1)^{3(B-L)+2s}, \quad (2.22)$$

<sup>674</sup> where s represents the spin of the particles.  $B = \pm \frac{1}{3}$  for quarks/antiquarks and  $B = 0$   
<sup>675</sup> for all others,  $L = \pm 1$  for leptons/antileptons,  $L = 0$  for all others.

<sup>676</sup> R-parity ensures the stability of the proton in **SUSY** models, and also has other conse-  
<sup>677</sup> quences for the production and decay of supersymmetric particles. In particle colliders  
<sup>678</sup> supersymmetric particles can then only be pair produced. Similarly the decay of any  
<sup>679</sup> produced supersymmetric particle is restricted to a **SM** particle and a lighter supersym-  
<sup>680</sup> metric particle, as allowed by conservation laws. A further implication of R-parity is  
<sup>681</sup> that once a supersymmetric particle has decayed to the **LSP** it remains stable, unable to  
<sup>682</sup> decay into a **SM** particle.

<sup>683</sup> A **LSP** will not interact in a detector at a particle collider, leaving behind a missing  
<sup>684</sup> energy  $\cancel{E}_T$  signature. The assumption of R-parity and its consequences are used to  
<sup>685</sup> determine the physical motivation and search strategies for **SUSY** at the **LHC**.

### <sup>686</sup> 2.4. Experimental Signatures of SUSY at the LHC

<sup>687</sup> Should strongly interacting sparticles be within the experimental reach of the **LHC**, then  
<sup>688</sup> it is expected that they can be produced in a variety of ways:

- 689     • squark/anti-squark and gluino pairs can be produced via both gluon fusion and  
690        quark/anti-quark scattering,  
691     • a gluino and squark produced together via quark-gluon scattering,  
692     • squark pairs produced via quark-quark scattering.

693 Whilst most **SUSY** searches invoke the requirement of R-parity to explore parameter  
694 phase space, there still exist a whole plethora of possible **SUSY** model topologies, which  
695 could yet be discovered at the **LHC**.

696 During the 2011 run period at  $\sqrt{s} = 7$  TeV, particular models were used to benchmark  
697 performance and experimental reach of both **CMS** searches and previous experiments.  
698 The Compressed Minimal SuperSymmetric Model (**CMSSM**) was initially chosen for a  
699 number of reasons [38]. One of the most compelling being the reduction of the up to 105  
700 new parameters that can be introduced by **SUSY** (in addition to the existing 19 of the  
701 **SM**), to just 5 extra free parameters. It was this simplicity, combined with the theory  
702 not requiring any fine tuning of particle masses to produce experimentally verified **SM**  
703 observables, that made it an attractive model to interpret physics results.

704 However, recent results from the **LHC** now strongly disfavour large swathes of **CMSSM**  
705 parameter space [39][40][41]. In the face of such results a more pragmatic model indepen-  
706 dent search strategy is now applied across most **SUSY** searches at the **LHC**, see Section  
707 (2.4.1).

708 As previously stated, a stable **LSP** that exhibits the properties of a dark matter candidate  
709 would be weakly interacting and therefore will not be directly detected in a detector  
710 environment. Additionally, the cascade decays of supersymmetric particles to this **LSP**  
711 state would also result in significant hadronic activity. These signatures will then be  
712 characterised through large amounts of hadronic jets (see Section (3.3.1)), leptons and  
713 a significant amount of missing energy dependent upon the size of the mass splitting  
714 between the **LSP** and the supersymmetric particle it has decayed from.

715 The **SM** contains processes which can exhibit a similar event topology to that described  
716 above, with the largest contribution coming from the general QCD multi-jet environ-  
717 ment of a hadron collider. A multitude of different analytical techniques are used by  
718 experimental physicists to reduce or estimate any reducible or irreducible backgrounds,  
719 allowing a possible **SUSY** signature to be extracted. The techniques employed within  
720 this thesis are described in great detail within Section (4.1).

### 721 2.4.1. Simplified Models

722 With such a variety of different ways for a **SUSY** signal to manifest itself, it is necessary  
723 to be able to interpret experimental reach through the masses of gluinos and squarks  
724 which can be excluded by experimental searches, rather than on a model specific basis.

725 This is accomplished through **SMS** models, which are defined by a set of hypothetical  
726 particles and a sequence of their production and decay modes [42][43]. In the **SMS** models  
727 considered within this thesis, only the production process for the two primary particles  
728 are considered. Each primary particle can undergo a direct or a cascade decay through  
729 an intermediate new particle. At the end of each decay chain there remains a neutral,  
730 undetected **LSP** particle, denoted  $\tilde{\chi}_{LSP}$  which can represent a neutralino or gravitino.  
731 Essentially it is easier to consider each **SMS** with branching ratios set to 100%. The  
732 masses of the primary particle and the **LSP** remain as free parameters, in which the  
733 absolute value and relative difference between the primary and **LSP** particle alter the  
734 kinematics of the event.

735 Different **SMS** models are denoted with a T-prefix, with a summary of the types interpreted  
736 within this thesis listed below [44].

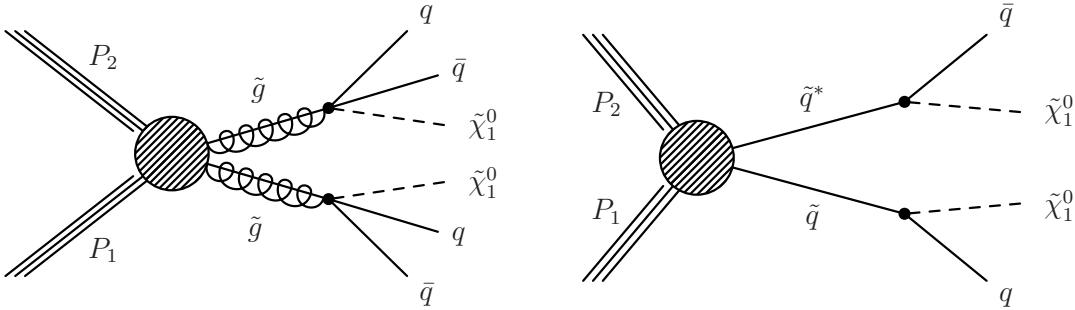
- 737 • **T1,T1xxxx**, models represent a simplified version of gluino pair production, with  
738 each gluino (superpartner to the gluon) undergoing a three-body decay to a quark-  
739 antiquark pair and the **LSP** (i.e.  $\tilde{g} \rightarrow q\bar{q}\tilde{\chi}_{LSP}$ ). The resultant final state from this  
740 decay is typically 4 jets +  $\cancel{E}_T$  in the absence of initial/final state radiation and  
741 detector effects. xxxx denotes models in which the quarks are of a specific flavour,  
742 typically t or b quark-antiquarks.
- 743 • **T2,T2xx**, models represent a simplified version of squark anti-squark production  
744 with each squark undergoing a two-body decay into a light-flavour quark and **LSP**  
745 (i.e.  $\tilde{q} \rightarrow q\tilde{\chi}_{LSP}$ ). This results in final states with less jets than gluino mediated  
746 production, typically 2 jets +  $\cancel{E}_T$  when again ignoring the effect of initial/final state  
747 radiation and detector effects. xx models represent decays in which both the quark  
748 and the squark within the decay is of a specific flavour, which in this thesis are  
749 again  $\tilde{t}/t$  or  $\tilde{b}/b$ .

750 Models rich in b and t quarks are interpreted within this thesis as they remain of  
751 particular interest within “Natural **SUSY**” scenarios [45][46]. The largest contribution  
752 to the quadratic divergence in the Higgs mass parameter comes from a loop of top  
753 quarks via the Yukawa coupling. Cancellation of these divergences can be achieved in

<sup>754</sup> supersymmetric theories by requiring a light right-handed top squark,  $\tilde{t}_R$ , and left-handed  
<sup>755</sup> double  $SU(2)_L$  doublet containing top and bottom squarks,  $(\tilde{\tilde{t}}_b)_L$  [47].

<sup>756</sup> These theories therefore solve the hierarchy problem by predicting light  $\sim$ EWK scale third  
<sup>757</sup> generation sleptons, accessible at the LHC. Search strategies involving the requirement  
<sup>758</sup> of b-tagging (see Section (3.3.2)) are used to give sensitivity to these type of SUSY  
<sup>759</sup> scenarios and are discussed in greater detail within Chapter 4.

<sup>760</sup> Two example decay chains are shown in Figure 2.2; the pair production of gluinos (T1)  
<sup>761</sup> and the pair production of squarks (T2) decaying into SM particles and LSPs.



**Figure 2.2:** Two example SMS model decays (T1 (left), T2 (right)), which are used in interpretations of physics reach by CMS.

# Chapter 3.

## <sup>762</sup> The LHC and the CMS Detector

<sup>763</sup> Probing the SM for signs of new physics would not be possible without the immensely  
<sup>764</sup> complex electronics and machinery that has made the TeV energy scale accessible to  
<sup>765</sup> physicists for the first time. This chapter will introduce both the LHC based at European  
<sup>766</sup> Organisation for Nuclear Research (CERN) and the Compact Muon Solenoid (CMS)  
<sup>767</sup> detector (of which the author is a member). Section (3.2) serves to present an overview of  
<sup>768</sup> the different components of the CMS detector, with specific components relevant to the  
<sup>769</sup> search for supersymmetric particles described in greater detail. Section (3.3) will focus  
<sup>770</sup> on event and object reconstruction, again, with more emphasis on jet level quantities  
<sup>771</sup> which are most relevant to the author's analysis research. Finally, Section (3.4) will  
<sup>772</sup> describe and detail the service work for the CMS Collaboration performed by the author,  
<sup>773</sup> in measuring the performance of L1 single jet and energy sum triggers in the Global  
<sup>774</sup> Calorimeter Trigger (GCT) during the 2012-2013 run period.

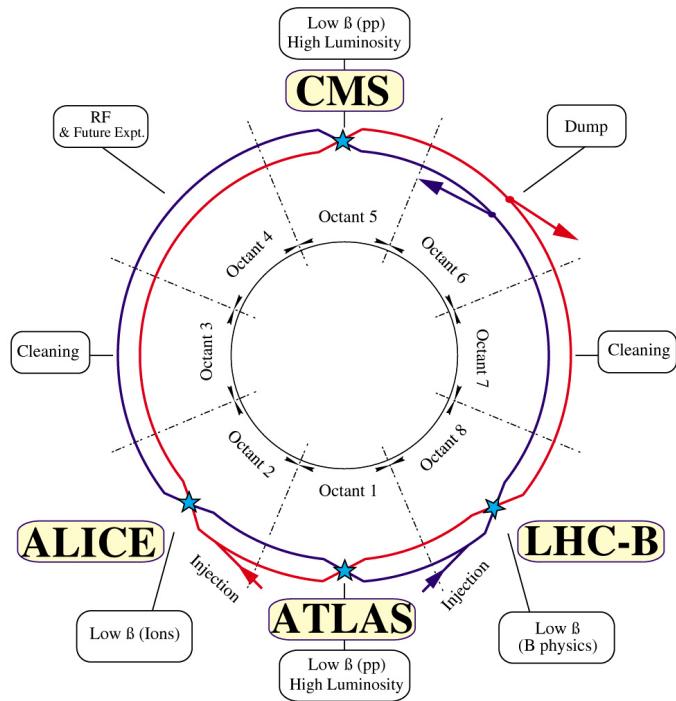
### <sup>775</sup> 3.1. The LHC

<sup>776</sup> The LHC is a storage ring, accelerator, and collider of circulating beams of protons or ions.  
<sup>777</sup> Housed in the tunnel dug for Large Electron-Positron Collidor (LEP), it is approximately  
<sup>778</sup> 27km in circumference, 100m underground, and straddles the border between France and  
<sup>779</sup> Switzerland, outside of Geneva. It is currently the only collider in operation that is able  
<sup>780</sup> to study physics at the TeV scale. A double-ring circular synchrotron, it was designed  
<sup>781</sup> to collide both proton-proton (pp) and heavy ion (PbPb) with a centre of mass energy  
<sup>782</sup>  $\sqrt{s} = 14$  TeV at a final design luminosity of  $10^{34}\text{cm}^{-2}\text{s}^{-1}$ .

<sup>783</sup>

These counter-circulating beams of protons or Pb ions are merged in four sections around the ring to enable collisions of the beams, with each interaction point being home to one of the four major experiments; A Large Ion Collider Experiment (**ALICE**) [48], A Toroidal LHC ApparatuS (**ATLAS**) [49], Compact Muon Solenoid (**CMS**) [50] and Large Hadron Collider Beauty (**LHCb**) [51] which record the resultant collisions. The layout of the **LHC** ring is shown in Figure 3.1. The remaining four sections contain acceleration, collimation and beam dump systems. In the eight arc sections, the beams are steered by magnetic fields of up to 8 T provided by super conduction dipole magnets, which are maintained at temperatures of 2 K using superfluid helium. Additional magnets for focusing and corrections are also present in straight sections within the arcs and near the interaction regions where the detectors are situated.

795



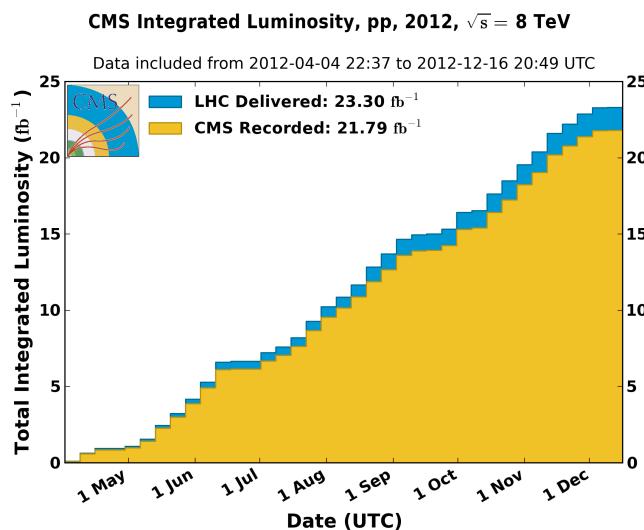
**Figure 3.1:** A top down layout of the LHC. [52], with the position of the four main detectors labelled.

Proton beams are formed inside the Proton Synchrotron (**PS**) from bunches of protons 50 ns apart with an energy of 26 GeV. The protons are then accelerated in the Super Proton Synchrotron (**SPS**) to 450 GeV before being injected into the **LHC**. These **LHC** proton beams consists of many “bunches” (i.e. approximately  $1.1 \times 10^{11}$  protons localised into less than 1 ns in the direction of motion). Before collision, the beams are ramped to

801 4 TeV (2012) per beam, in a process involving increasing the current passing through  
 802 the dipole magnets. Once the desired  $\sqrt{s}$  energy is reached then the beams are allowed  
 803 to collide at the interaction points. The luminosity falls regularly as the run progresses;  
 804 protons are lost in collisions, and eventually the beam is dumped before repeating the  
 805 process again.

806 Colliding the beams produced an instantaneous luminosity of approximately  $5 \times 10^{33}$   
 807  $\text{cm}^{-2}\text{s}^{-1}$  during the 2012 run. The high number of protons in each bunch increases  
 808 the likelihood of multiple interactions with each crossing of the counter-circulating  
 809 beams. This leads to isotropic energy depositions within the detectors positioned at these  
 810 interaction points, increasing the energy scale of the underlying event. This is known  
 811 as *pile-up* and the counteracting of its effects are important to the many measurements  
 812 performed at the [LHC](#).

813 In the early phase of prolonged operation, after the initial shutdown, the machine operated  
 814 in 2010-2011 at 3.5 TeV per beam,  $\sqrt{s} = 7$  TeV, delivering  $6.13 \text{ fb}^{-1}$  of data [53]. During  
 815 the 2012-2013 run period, data was collected at an increased  $\sqrt{s} = 8$  TeV improving the  
 816 sensitivity of searches for new physics. Over the whole run period  $23.3 \text{ fb}^{-1}$  of data was  
 817 delivered, of which  $21.8 \text{ fb}^{-1}$  was recorded by the [CMS](#) detector as shown in Figure 3.2  
 818 [53]. A total of  $12 \text{ fb}^{-1}$  of 8 TeV certified data was collected by October 2012, and it is  
 819 this data which forms the basis of the results presented within this thesis.



**Figure 3.2:** The total integrated luminosity delivered to and collected by [CMS](#) during the 2012 8 TeV *pp* runs.

## 820 3.2. The CMS Detector

821 The Compact Muon Solenoid (**CMS**) detector is one of two general purpose detectors  
822 at the **LHC** designed to search for new physics. The detector is designed to provide  
823 efficient identification and measurement of many physics objects including photons,  
824 electrons, muons, taus, and hadronic showers over wide ranges of transverse momentum  
825 and direction. Its nearly  $4\pi$  coverage in solid angle allows for accurate measurement of  
826 global transverse momentum imbalance. These design factors give **CMS** the ability to  
827 search for direct production of **SUSY** particles at the TeV scale, making the search for  
828 Supersymmetric particles one of the highest priorities among the wide range of physics  
829 programmes at **CMS**.

830 **CMS** uses a right-handed Cartesian coordinate system with the origin at the interaction  
831 point and the z-axis pointing along the beam axis. The x-axis points radially inwards to  
832 the centre of the collider ring, with the y-axis points vertically upward. The azimuthal  
833 angle  $\phi$ , ranging between  $[-\pi, \pi]$ , is defined in the x-y plane starting from the x-axis. The  
834 polar angle  $\theta$  is measured from the z axis. The common convention in particle physics is  
835 to express an out-going particle in terms of  $\phi$  and its pseudorapidity defined as

$$\eta = -\log \tan \left( \frac{\theta}{2} \right). \quad (3.1)$$

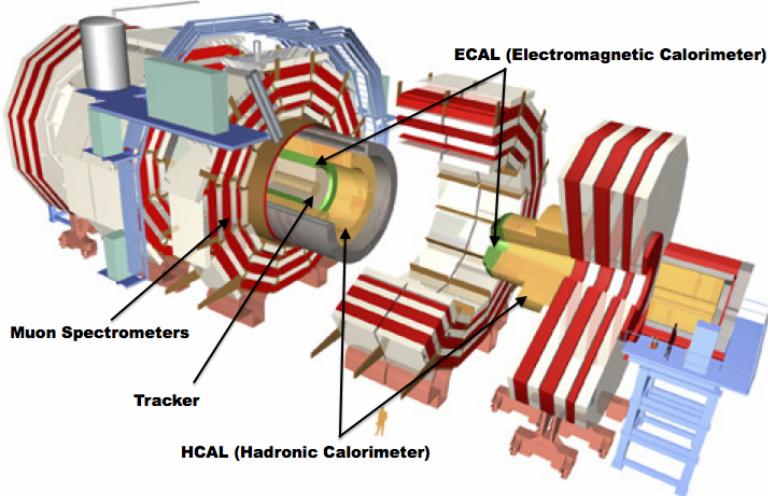
836 The variable  $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$  is commonly used to define angular distance between  
837 objects within the detector. Additionally, energy and momentum is typically measured in  
838 the transverse plane perpendicular to the beam line. These values are calculated from the  
839 x and y components of the object and are denoted as  $E_T = E \sin \theta$  and  $p_T = \sqrt{p_x^2 + p_y^2}$ .

### 840 3.2.1. Detector Subsystems

841 As the range of particles produced from  $pp$  collisions interact in different ways with  
842 matter, **CMS** is divided into sub-detector systems, which perform complementary roles  
843 to identify the identity, the mass, and the momentum of different physics objects present  
844 in each event. These detector sub-systems contained within **CMS** are wrapped in layers  
845 around a central 13m long 4 T super conducting solenoid, as shown in Figure 3.3. With  
846 the endcaps closed, **CMS** is a cylinder of length 22m, diameter 15m, and mass 12.5

847 kilotons. A more detailed complete description of the detector can be found elsewhere [50].

848



**Figure 3.3:** A pictorial depiction of the CMS detector with the main detector subsystems labelled [54].

### 849 3.2.2. Tracker

850 The inner-most sub-detector of the barrel is the multi-layer silicon tracker, formed of a  
 851 pixel detector component encased by layers of silicon strip detectors. The pixel detector  
 852 consists of three layers of silicon pixel sensors providing measurements of the momentum,  
 853 position coordinates of the charged particles as they pass, and the location of primary  
 854 and secondary vertices between 4 cm and 10 cm transverse to the beam. Outside the  
 855 pixel detector, ten cylindrical layers of silicon strip detectors extend the tracking system  
 856 out to a radius of 1.20m from the beam line. The tracking system provides efficient  
 857 and precise determination of the charges, momenta, and impact parameters of charged  
 858 particles, with the geometry of the tracker extending to cover a rapidity range up to  $|\eta| <$   
 859 2.5.

860

861 The tracking system also plays a crucial part in the identification of jets that originate  
 862 from b-quarks through the measurement of displaced secondary vertices. The methods  
 863 in which these b-flavoured jets are identified are discussed within Section (3.3.2). The  
 864 identification of b-jets is important in many searches for natural SUSY models and forms  
 865 an important part of the inclusive search strategy described within Section (4.2).

<sup>866</sup> **3.2.3. Electromagnetic Calorimeter**

<sup>867</sup> Immediately outside of the tracker, but still within the magnet core, sits the Electromagnetic CALorimeter (**ECAL**). Covering a pseudorapidity up to  $|\eta| < 3$  and comprising <sup>868</sup> <sup>869</sup> of over  $75 \times 10^3$  PbWO<sub>4</sub> (lead tungstate) crystals that scintillate as particles deposit <sup>870</sup> energy, the **ECAL** provides high resolution measurements of the electromagnetic showers <sup>871</sup> from photons and electrons in the detector.

<sup>872</sup>

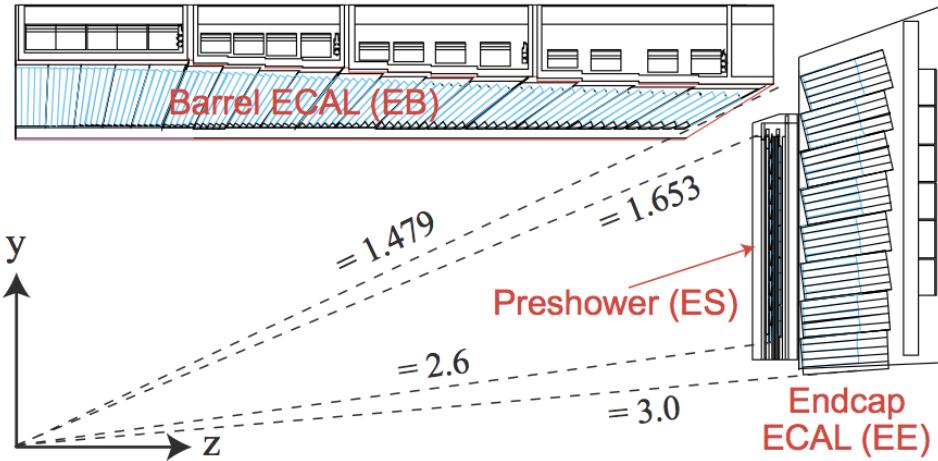
<sup>873</sup> Lead tungstate is used because of its short radiation length ( $X_0 \sim 0.9$  cm) and small <sup>874</sup> Molieré radius ( $\sim 2.1$  cm) leading to high granularity and resolution. Its fast scintillation <sup>875</sup> time ( $\sim 25$  ns) reduces the effects of pile-up, which occurs when energy from previous <sup>876</sup> collisions are still being read out, and its radiation hardness gives it longevity. The <sup>877</sup> crystals are arranged in modules which surround the beam line in a non-projective <sup>878</sup> geometry, angled at  $3^\circ$ , with respect to the interaction point to minimise the risk of <sup>879</sup> particles escaping down the cracks between the crystals.

<sup>880</sup>

<sup>881</sup> The **ECAL** is primarily composed of two sections, the Electromagnetic CALorimeter <sup>882</sup> Barrel (**EB**) which extends in pseudo-rapidity to  $|\eta| < 1.479$  with a crystal front cross section of  $22 \times 22$  mm and a length of 230 mm corresponding to 25.8 radiation lengths. The <sup>883</sup> <sup>884</sup> Electromagnetic CALorimeter Endcap (**EE**) covers a rapidity range of  $1.479 < |\eta| < 3.0$ , <sup>885</sup> which consists of two identical detectors on either side of the **EB**. A lead-silicon sampling <sup>886</sup> ‘pre-shower’ detector Electromagnetic CALorimeter pre-Shower (**ES**) is placed before <sup>887</sup> the endcaps to aid in the identification of neutral pions. Their arrangement is shown in <sup>888</sup> Figure 3.4.

<sup>889</sup>

<sup>890</sup> Scintillation photons from the lead tungstate crystals are instrumented with Avalanche <sup>891</sup> Photo-Diodes (**APD**) and Vacuum Photo-Triodes (**VPT**), located in the **EB** and **EE** <sup>892</sup> respectively. They convert the scintillating light into an electric signal which is conse- <sup>893</sup> quently used to determine the amount of energy deposited within the crystal. These <sup>894</sup> instruments are chosen for their resistance under operation to the strong magnetic field <sup>895</sup> of **CMS**. The scintillation of the **ECAL** crystals, as well as the response of the **APDs**, <sup>896</sup> vary as a function of temperature; and so cooling systems continually maintain an overall <sup>897</sup> constant **ECAL** temperature  $\pm 0.05^\circ C$ .



**Figure 3.4:** Illustration of the CMS ECAL showing the arrangement of the lead tungstate crystals in the EB and EE. The ES is also shown and is located in front of the EE [55].

### 3.2.4. Hadronic Calorimeter

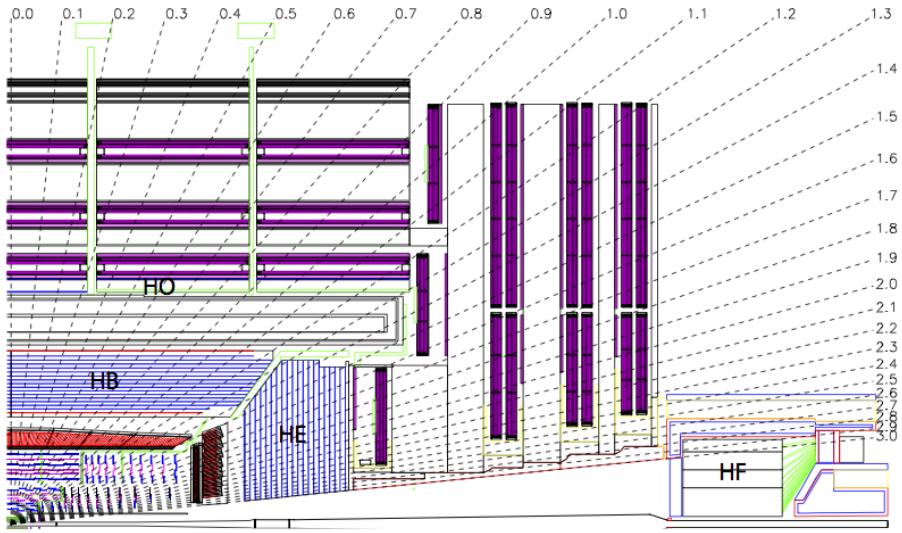
Beyond the ECAL lies the Hadronic CALorimeter (HCAL) which is responsible for the accurate measurement of hadronic showers, crucial for analyses involving jets or missing energy signatures. The HCAL is a sampling calorimeter which consists of alternating layers of brass absorber and plastic scintillator. The exception being in the hadron forward ( $3.0 < |\eta| < 5.0$ ) region where steel absorbers and quartz fibre scintillators are used because of their increased radiation tolerance. Hadron showers are initiated in the absorber layers inducing scintillation in the plastic scintillator tiles. These scintillation photons are converted by wavelength shifting fibres for read-out by hybrid photodiodes.

907

The HCAL's size is constrained to a compact size by the presence of the solenoid, requiring the placement of an additional outer calorimeter on the outside of the solenoid to increase the sampling depth of the HCAL. A schematic of the HCAL can be seen in Figure 3.5.

912

The HCAL covers the range  $|\eta| < 5$  and consists of four sub-detectors: the Hadron Barrel (HB)  $|\eta| < 1.3$ , the Hadron Outer (HO), the Hadron Endcaps (HE)  $1.3 < |\eta| < 3.0$  and the Hadron Forward (HF). The HB, contained between the outer edge of the ECAL and the inner edge of the solenoid is formed of 36 azimuthal wedges which are split



**Figure 3.5:** Schematic of the hadron calorimeters in the r-z plane, showing the locations of the **HCAL** components and the **HF**. [50].

917 between two half-barrel segments. Each wedge is segmented into four azimuthal angle  
 918 ( $\phi$ ) sectors, and each half-barrel is further segmented into 16  $\eta$  towers. The electronic  
 919 readout chain, channels the light from the active scintillator layers from one  $\phi$ -segment  
 920 and all  $\eta$ -towers of a half-barrel to a Hybrid Photo Diode (**HPD**).

921 The relatively short number of interaction lengths ( $\lambda_l$ , the distance a hadron will travel  
 922 through the absorber material before it has lost  $\frac{1}{e}$  of its energy) within the **HB**, the lowest  
 923 being  $\lambda_l = 5.82$  at  $|\eta| = 0$ , facilitates the need for the ‘tail catching’ **HO** to increase the  
 924 sampling depth in the central barrel rapidity region  $|\eta| < 1.3$  to 11 interaction lengths.  
 925 Significant fractions of the hadrons energy will also be deposited in the **ECAL** as it passes  
 926 through the detector. Therefore, measurements of hadron energies in the central regions  
 927  $|\eta| < 3.0$  use both the **ECAL** and **HCAL** to reconstruct the true energy from showering  
 928 hadrons.

### 929 3.2.5. Muon Systems

930 Muons being too massive to radiate away energy via Bremsstrahlung, interact little in  
 931 the calorimeters and mostly pass through the detector until they reach the system of  
 932 muon detectors which forms the outer most part of the **CMS** detector.

933 Outside of the superconducting solenoid are four muon detection layers interleaved with  
934 the iron return yokes, which measure the muons energy via ionisation of gas within  
935 detector elements. Three types of gaseous chambers are used. The Drift Tube (**DT**),  
936 Cathode Stripe Chamber (**CSC**), and Resistive Plate Chamber (**RPC**) systems provide  
937 efficient detection of muons with pseudo-rapidity  $|\eta| < 2.4$ . The best reconstruction  
938 performance is obtained when the muon chamber is combined with the inner tracking  
939 information to determine muon trajectories and their momenta [56].

940

### 941 3.3. Event Reconstruction and Object Definition

942 The goal of event reconstruction is to take the raw information recorded by the detector  
943 and to compute from it higher-level quantities which can be used at an analysis level.  
944 These typically correspond to an individual particle’s energy and momenta, groups of  
945 particles which shower in a narrow cone, and the overall global energy and momentum  
946 balance of the event. The reconstruction of these objects are described in great detail in  
947 [57], while covered below are brief descriptions of those which are most relevant to the  
948 analysis detailed in Chapter 4.

#### 949 3.3.1. Jets

950 Quarks and gluons are produced copiously at the LHC in the hard scattering of partons.  
951 As these quarks and gluons fragment, they hadronize and decay into a group of strongly  
952 interactive particles and their decay products. These streams of particles travel in the  
953 same direction, as they have been “boosted” by the momentum of the primary hadron.  
954 These collections of decay products are reconstructed and identified together as a “jet”.

955 At **CMS** jets are reconstructed from energy deposits in the detector via the anti-kt  
956 algorithm [58] with size parameter  $\Delta R = 0.5$ . The anti-kt jet algorithm clusters jets by  
957 defining a distance between hard (high  $p_T$ ) and soft (low  $p_T$ ) particles such that soft  
958 particles are preferentially clustered with hard particles before being clustered between  
959 themselves. This produces jets which are robust to soft particle radiation from the pile-up  
960 conditions produced by the **LHC**.

- 961 There are two main types of jet reconstruction used at **CMS**, Calorimeter (Calo) and  
962 Particle Flow (PF) jets [59]. Calorimeter jets are reconstructed using both the **ECAL**  
963 and **HCAL** cells, combined into calorimeter towers. These calorimeter towers consist of  
964 geometrically matched **HCAL** cells and **ECAL** crystals. Electronics noise is suppressed by  
965 applying a threshold to the calorimeter cells, with pile-up effects reduced by a requirement  
966 placed on the tower energy [60]. Calorimeter jets are the jets used within the analysis  
967 presented in this thesis.
- 968 PF jets are formed from combining information from all of the **CMS** sub-detectors systems  
969 to determine which final state particles are present in the event. Generally, any particle  
970 is expected to produce some combination of a track in the silicon tracker, a deposit in  
971 the calorimeters, or a track in the muon system. The PF jet momentum and spatial  
972 resolutions are greatly improved with respect to calorimeter jets, as the use of the tracking  
973 detectors and of the high granularity of **ECAL** allows resolution and measurement of  
974 charged hadrons and photons inside a jet, which together constitute  $\sim 85\%$  of the jet  
975 energy [61].
- 976 The jets reconstructed by the clustering algorithm in **CMS** typically have an energy  
977 that differs to the ‘true’ energy measured by a perfect detector. This stems from the  
978 non-linear and nonuniform response of the calorimeters as well as other residual effects  
979 including pile-up and underlying events. Therefore, additional corrections are applied to  
980 recover a uniform relative response as a function of pseudo-rapidity. These are applied  
981 as separate sub corrections [62].
- 982 • A pile-up correction is first applied to the jet. It subtracts the average extra energy  
983 deposited in the jet that comes from other vertices present in the event and is  
984 therefore not part of the hard jet itself.
  - 985 •  $p_T$  and  $\eta$  dependant corrections derived from Monte Carlo simulations are used to  
986 account for the non-uniform response of the detector.
  - 987 •  $p_T$  and  $\eta$  residual corrections are applied to data only to correct for difference  
988 between data and Monte Carlo. The residual is derived from QCD di-jet samples  
989 and the  $p_T$  residual from  $\gamma+$  jet and  $Z+$  jets samples in data.

### 990 3.3.2. B-tagging

991 The decays of b quarks are suppressed by small CKM matrix elements. As a result, the  
 992 lifetimes of b-flavoured hadrons, produced in the fragmentation of b quarks, are relatively  
 993 long;  $\sim 1\text{ps}$ . The identification of jets originating from b quarks is very important for  
 994 searches for new physics and for measurements of SM processes.

995

996 Many different algorithms developed by CMS select b-quark jets based on variables such  
 997 as; the impact parameters of the charged-particle tracks, the properties of reconstructed  
 998 decay vertices, and the presence or absence of a lepton, or combinations thereof.

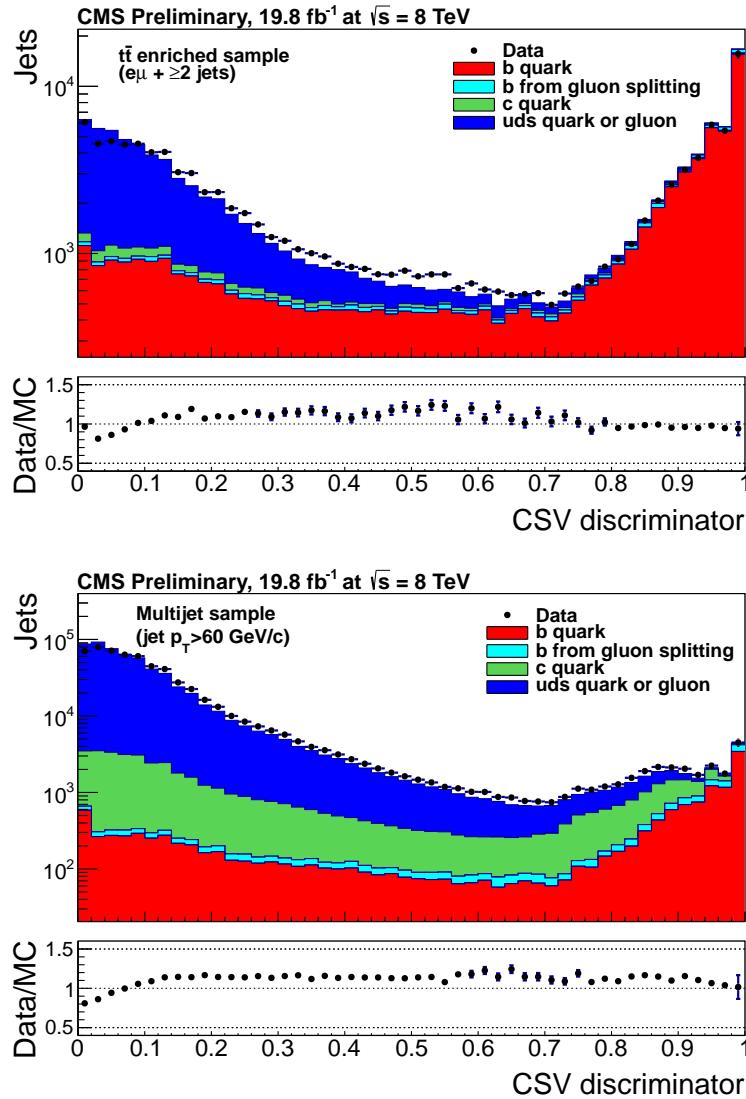
999 One of the most efficient of which is the Combined Secondary Vertex (CSV) algorithm  
 1000 [63]. This operates based on secondary vertex and track-based lifetime information,  
 1001 benchmarked in ‘Loose’, ‘Medium’ and ‘Tight’ working points, of which the medium  
 1002 point is the tagger used within the  $\alpha_T$  search presented in Section (4.1). All figures  
 1003 within this sub-section, demonstrating the performance of this b-tagging algorithm are  
 1004 taken from [64].

1005 Within the CSV tagger, a likelihood-based discriminator distinguishes between jets from  
 1006 b-quarks, and those from charm or light quarks and gluons, shown in Figure 3.6. The  
 1007 minimum thresholds on the discriminator for each working point correspond to the  
 1008 mis-identification probability for light-parton jets of 10%, 1%, and 0.1%, respectively, in  
 1009 jets with an average  $p_T$  of about 80 GeV.

1010 The b-tagging performance is evaluated to measure the b-jet tagging efficiency  $\epsilon_b$ , and the  
 1011 misidentification probability of charm  $\epsilon_c$  and light-parton jets  $\epsilon_s$ . The tagging efficiencies  
 1012 for each of these three jet flavours are compared between data and MC simulation, from  
 1013 which a series of  $p_T$  and  $|\eta|$  dependant jet corrections are determined,

$$SF_{b,c,s} = \frac{\epsilon_{b,c,s}^{data}}{\epsilon_{b,c,s}^{MC}}. \quad (3.2)$$

1014 These are collectively named ‘Btag Scale Factors’ and allow MC simulation to accu-  
 1015 rately reflect the running conditions and performance of the tagging algorithm in data.  
 1016 Understanding of the b-tagging efficiency is essential in order to minimise systematic



**Figure 3.6:** CSV algorithm discriminator values in enriched  $t\bar{t}$  (top) and inclusive multi-jet samples (bottom) for b,c and light flavoured jets. The discriminator value used for each working points are determined from the misidentification probability for light-parton jets to be tagged as a b-jet, which are given as 0.244 (10%), 0.679 (1%) and 0.898 (0.1%) for the L, M and T working points respectively.

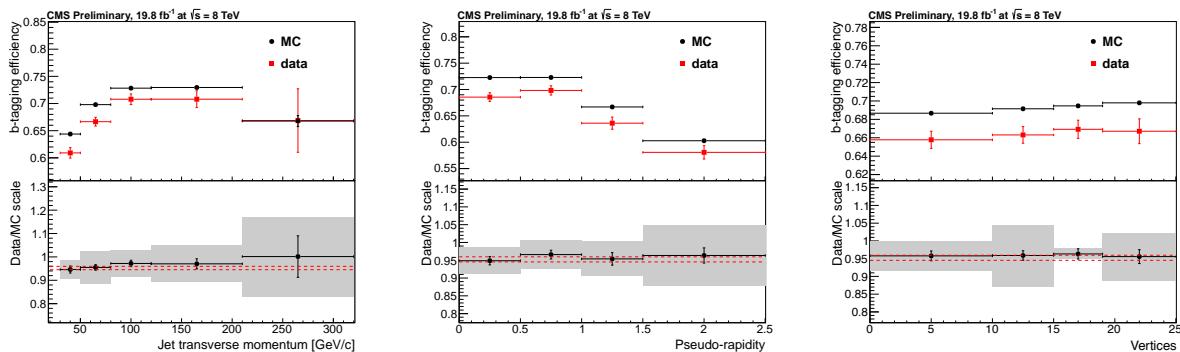
<sup>1017</sup> uncertainties in physics analyses that employ b-tagging.

<sup>1018</sup>

<sup>1019</sup> The b-tagging efficiency is measured in data using several methods applied to multi-jet  
<sup>1020</sup> events, primarily based on a sample of jets enriched in heavy flavour content. One method  
<sup>1021</sup> requires the collection of events with a poorly isolated muon within a cone  $\Delta R < 0.4$   
<sup>1022</sup> around the jet axis. Due to the semi-leptonic branching fraction of b hadrons being

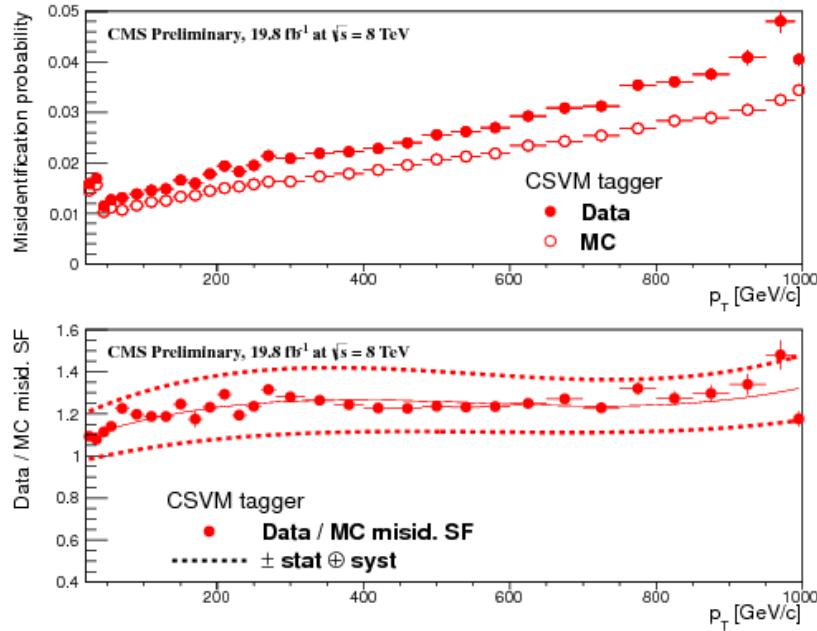
significantly larger than that for other hadrons, these jets are more likely to arise from b quarks than from another flavour. The resultant momentum component of the muon, transverse to the jet axis, is larger in b-hadron decays than from light or charm flavoured jets.

Additionally, the performance of the tagger can also be benchmarked in  $t\bar{t}$  events, where the top quark is expected to decay to a W boson and a b quark about 99.8% of the time [1]. Further selection criteria is applied to these events to further enrich the b quark content of these events. The methods to identify b-jets in data are discussed in great detail at [65]. The jet flavours within simulation are determined using truth level information and are compared to data to determine the appropriate correction scale factors ( $SF_b$ ). These are displayed for the CSVM tagger in Figure 3.7.



**Figure 3.7:** Measured in  $t\bar{t} \rightarrow$  di-lepton events using the CSVM tagger: (upper panels) b-tagging efficiencies and (lower panels) data/MC scale factor  $SF_b$  as a function of (left) jet  $p_T$ , (middle) jet  $|\eta|$  and (right) number of primary vertices. In the lower panels, the grey filled areas represent the total statistical and systematic uncertainties, whereas the dotted lines are the average  $SF_b$  values within statistical uncertainties.

The measurement of the misidentification probability for light-parton jets relies on the inversion of tagging algorithms, selecting non-b jets using the same variables and techniques used for benchmarking the b-tagging efficiency. The scale factors ( $SF_s$ ) to be applied to correct simulation to data are shown in Figure 3.8 for the CSVM tagger.



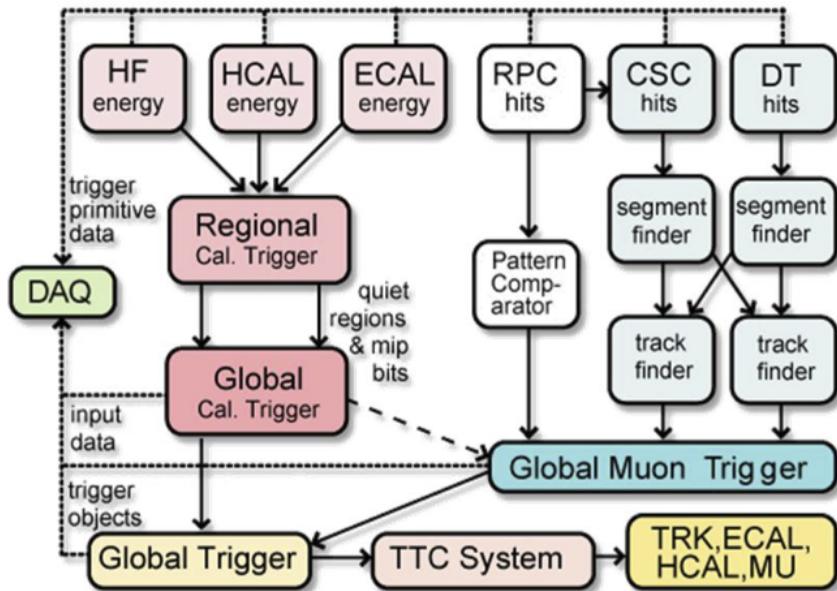
**Figure 3.8:** For the **CSVM** tagging criterion: (top) misidentification probability in data (filled circles) and simulation (open circles); (bottom) scale factor for the misidentification probability. The last  $p_T$  bin in each plot includes all jets with  $p_T > 1000$  GeV. The solid curve is the result of a polynomial fit to the data points. The dashed curves represent the overall statistical and systematic uncertainties on the measurements.

## 1038 3.4. Triggering System

1039 With bunch crossings separated by just 50 ns, the rate at which data from all collisions  
 1040 would have to be written out and processed would be unfeasible. A two-tiered triggering  
 1041 system is applied at **CMS** in order to cope with the high collision rate of protons. The  
 1042 **CMS** trigger is designed to use limited information from each event to determine whether  
 1043 to record the event, reducing the rate of data taking to manageable levels whilst ensuring  
 1044 a high efficiency of interesting physics object events are selected.

1045 The Level 1 Trigger (**L1**) is a pipelined, dead-timeless system based on custom-built  
 1046 electronics [66], and is a combination of several sub systems which is shown pictorially  
 1047 in Figure 3.9. The L1 system is covered in more detail within the following section,  
 1048 along with a description of the service work undertaken by the author to benchmark the  
 1049 performance of the L1 calorimeter trigger during the 2012 8 TeV run period.

1050 The Higher Level Trigger (**HLT**) is a large farm of commercial computers [67]. The **HLT**  
 1051 processes events with software reconstruction algorithms that are more detailed, giving  
 1052 performance more similar to the reconstruction used offline. The **HLT** reduces the event



**Figure 3.9:** An overview of the different components of the CMS L1 trigger system, showing the global calorimeter, muon triggers, and the global trigger.

rate written to disk by a factor of  $\sim 500$  ( $\sim 200\text{Hz}$ ). The recorded events are transferred from CMS to the CERN computing centre, where event reconstruction is performed, and then distributed to CMS computing sites around the globe for storage and analysis.

### **3.4.1. The Level-1 Trigger**

The L1 trigger reduces the rate of events collected from 20 MHz to  $\sim 100$  kHz using information from just the calorimeters and muon chambers, but not the tracker. This is due to requirement that data from each and every bunch crossing be analysed with no dead time, drastically reducing time available to process and reconstruct objects in making a trigger decision. This facilitates the need for a pipelined processing architecture, and so a tree system of triggers is used to decide whether to pass on an event to the HLT for further reconstruction.

<sup>1064</sup> Calorimeter and muon event information is processed separately by the Regional Calorimeter Trigger (**RCT**) and Regional Muon Trigger (**RMT**) systems respectively.  
<sup>1065</sup>

1066 Within the **RCT**, energy deposits from trigger towers in the **ECAL** and **HCAL** calorimeters  
1067 are summed into coarser calorimeter regions and sent to the Global Calorimeter Trigger  
1068 (**GCT**) for jet clustering.

1069 Given that electron and photon are much narrower objects than jets, the **RCT** is used  
1070 to identify these candidates but makes no attempt to distinguish between them at this  
1071 stage given the lack of tracking information. They are first identified by ensuring the  
1072 energy deposits within the central trigger tower and its surrounding cells are above a  
1073 certain programmable threshold. To ensure the object is not a hadron, the ratio of **HCAL**  
1074 to **ECAL** in the central tower is calculated and checked to be below 5%. Additional  
1075 algorithms are employed to ascertain whether the  $e/\gamma$  object is isolated/non-isolated.

1076 In the L1 **GCT**, coarse measurements of the energy deposited in the electromagnetic and  
1077 hadronic calorimeters are combined, and by using sophisticated algorithms the following  
1078 tasks are performed:

- 1079 • isolated and non-isolated electromagnetic objects are sorted ( $e$  and  $\gamma$ ), with the four  
1080 highest ranked (equivalent to highest transverse energy  $E_T$ ) objects of each type  
1081 passed onto the Global Trigger (**GT**),
- 1082 • energy sums from the calorimeters supplied by the **RCT** are used in performing  
1083 jet clustering (described in the following section). The clustered jets are then  
1084 sub-divided into categories depending on their pseudo-rapidity and the result of  
1085  $\tau$  identification, being classified as either central, forward, or tau ( $\tau$ ). After being  
1086 sorted by rank, the four highest of each category are passed to the **GT** for use in  
1087 trigger decisions,
- 1088 • total transverse energy ( $E_T$ ), the scalar sum of the energy deposits measured by  
1089 L1, and missing transverse energy ( $\cancel{E}_T$ ), defined as the negative vector sum of the  
1090 transverse energy deposits measured at L1 are calculated,
- 1091 • total transverse jet energy ( $H_T$ ), the scalar sum of the energy of all L1 clustered jet  
1092 objects, and missing transverse jet energy ( $\cancel{H}_T$ ), defined as the negative vector sum  
1093 of the energy from L1 clustered jet objects are calculated and passed to the **GT**.

1094 In addition, quantities suitable for triggering minimum bias events, forward physics and  
1095 beam background events are determined. Relevant muon isolation information is also  
1096 passed on to the Global MuonTrigger (**GMT**) to be used in decisions involving the muon  
1097 triggers, where it is combined with information from across the three muon sub-systems.

1098 The resultant final accept/reject decision at **L1** is then performed by the **GT**, based on  
1099 the objects received from the **GCT** and **GMT** ( $e/\gamma, \mu$ , jets,  $E_T, \cancel{E}_T, H_T, \cancel{H}_T$ ).

1100 The L1 trigger is therefore of upmost importance to the functioning of the detector.  
1101 Without a high-performing, efficient trigger and a good understanding of its performance  
1102 at ever increasing instantaneous luminosities, the data collected would be useless. Whilst  
1103 it would be possible to maintain trigger efficiency by increasing the triggering thresholds  
1104 for different jet or energy sum quantities, this is far from ideal. This could result in the  
1105 failure to be sensitive to a wide range of new physics signatures, including many types of  
1106 compressed spectra **SUSY** models where the mass splitting between squarks/gluinos and  
1107 the **LSP** is small.

1108 One such method introduced to help maintain low triggering thresholds, was via the  
1109 introduction of a jet seed threshold into the L1 jet clustering algorithm. Observations of  
1110 how the L1 trigger performance is affected by both the jet seed threshold, and changing  
1111 **LHC** running conditions over the 2012 run period is presented in the following Sections  
1112 (3.4.2 - 3.4.6).

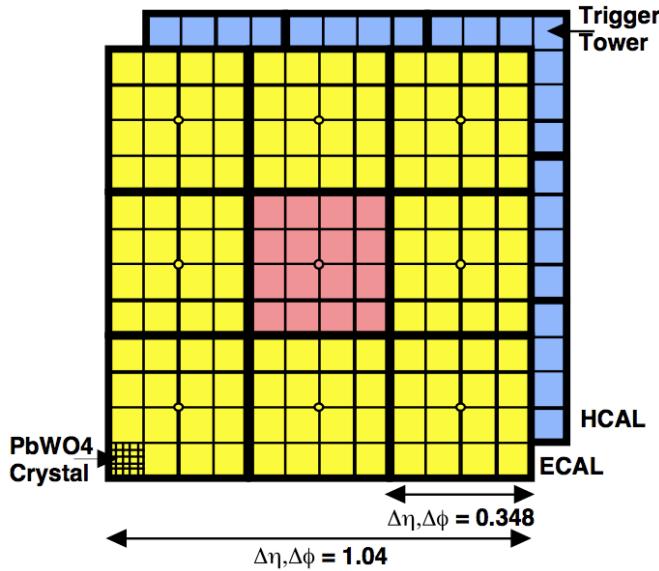
### 1113 3.4.2. The L1 Trigger Jet Algorithm

1114 The L1 jet algorithm clusters jets using the transverse energy sums computed by the  
1115 calorimeter trigger regions. Each region consists of a  $4 \times 4$  trigger tower window which  
1116 within the **CMS** barrel spans a region of  $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$  in pseudorapidity-  
1117 azimuth.

1118 A L1 jet is defined by a  $3 \times 3$  window of calorimeter regions, as shown in Figure 3.10.  
1119 This corresponds to  $12 \times 12$  trigger towers in barrel and endcap, or  $3 \times 3$  larger **HF** towers  
1120 in the **HF**. The  $\phi$  size of the jet window is the same everywhere, whilst the  $\eta$  binning  
1121 gets somewhat larger at high  $\eta$  due to calorimeter and trigger tower segmentation. The  
1122 jets are labelled by the  $(\eta, \phi)$  indices of the central calorimeter region.

1123 A jet candidate is identified if the sum of the transverse **ECAL** and **HCAL** energies of a  
1124 calorimeter region is larger than all of its 8 neighbouring regions  $E_{T \text{ central}} > E_{T \text{ surround}}$ .  
1125 The jet is then centred at this central calorimeter region.

1126 During the 2012 run period, a minimum threshold of 5 GeV was imposed on the central  
1127 seeding region to suppress noise from non-collimated pile-up jets. This threshold is  
1128 applied on the raw, uncorrected energy of the calorimeter regions and affects all clustered



**Figure 3.10:** Illustration of the dimensions of the Level-1 jet finder window. Each cell represents a trigger tower, which is the sum of the transverse energy contributions from both calorimeter systems.

1129 L1 jets. The effect of such a change to the jet algorithm on the triggering performance  
 1130 of L1 quantities is shown in Section (3.4.4).

1131 To decrease the amount of data shared the **GCT** utilises a pre-clustering algorithm,  
 1132 which employs 18 jet-finders that operate simultaneously over the whole detector. Each  
 1133 jet-finder spans an area of 11 calorimeter regions in  $\eta$  (half the detector) and two in  $\phi$   
 1134 ( $40^\circ$ ). In order to reduce the total data duplicated and shared between these jet-finders,  
 1135 jets are initially created in  $2 \times 3$  mini-clusters. Information is only shared with the two  
 1136  $\phi$  strips of the neighbouring jet-finders when these mini-clusters jets are found, and is  
 1137 used to form a clustered  $3 \times 3$  L1 jet object. Additional information on this clustering  
 1138 algorithm can be found at [68].

1139 Within the  $|\eta| < 3$  region, the **GCT** also determines whether a jet is to be classified as  
 1140 a  $\tau$  or central jet. The hadronic decay modes of the  $\tau$  typically contain one or three  
 1141 isolated pions, thus leading to more collimated energy deposits with fewer constituents  
 1142 than non- $\tau$  jets. Therefore, for a jet candidate to be classed as a  $\tau$  jet, up to a maximum  
 1143 of one of the eight calorimeter regions neighbouring the central jet seed is permitted to  
 1144 have a transverse energy,  $E_T$ , above some programmable isolation threshold.

1145 Jets found between  $3.0 < |\eta| < 5.0$  are classified as forward jets, whereas those with  
 1146  $|\eta| < 3.0$  are classified as either a central or  $\tau$ -jet. The four highest energy central,

1147 forward and  $\tau$ -jets in the calorimeter are further passed through Look Up Table (**LUT**)s,  
1148 which apply a programmable  $\eta$ -dependent jet energy scale correction. Finally these jet  
1149 objects are passed to the **GT** to make L1 trigger decisions.

1150 The performance of the L1 jets is evaluated with respect to offline jets, which are taken  
1151 from the standard Calo jet and the PF jet reconstruction algorithms of **CMS**. Jets are  
1152 corrected for pile-up and detector effects as described in Section (3.3.1). A moderate  
1153 level of noise rejection is applied to the offline jets by selecting jets passing the “loose”  
1154 identification criteria for both Calo and PF. These jet criteria are listed in Appendix  
1155 (A.1).

### 1156 3.4.3. Measuring L1 Single-Jet Trigger Efficiencies

1157 The efficiency of a L1 single-jet trigger at an offline reconstructed jet  $E_T$  is defined as; the  
1158 fraction of events in a sample containing at least a single reconstructed offline jet, where  
1159 the leading offline jet is matched with a L1 central or  $\tau$  jet that also has a measured L1  
1160 energy above the trigger threshold being benchmarked.

1161 A match is determined by comparing the L1 and reconstructed offline jets spatially in  
1162  $\eta - \phi$  space. The  $\Delta R$  separation between the highest offline reconstructed jet ( $E_T > 10$   
1163 GeV and  $|\eta| < 3$ ) and each L1 jet in the event is calculated. A match is made to the L1  
1164 jet with the minimum  $\Delta R$  to the reconstructed jet on the condition that it also satisfies  
1165  $\Delta R < 0.5$ .

1166 The matching efficiency for this procedure is found to be close to 100% above an offline  
1167 jet threshold of 30(45) GeV for the run 2012B(C) data taking period (see Appendix B.1).

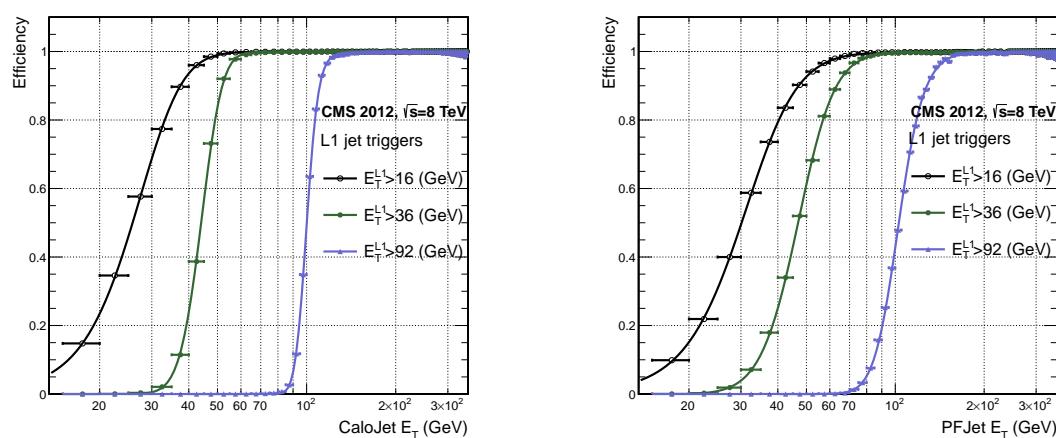
1168 Each efficiency curve is fitted with a function which is the cumulative distribution function  
1169 of an Exponentially Modified Gaussian (**EMG**) distribution:

$$\text{where } f(x; \mu, \sigma, \lambda) = \frac{\lambda}{2} \cdot e^{\frac{\lambda}{2}(2\mu + \lambda\sigma^2 - 2x)} \cdot \text{erfc}\left(\frac{\mu + \lambda\sigma^2 - x}{\sqrt{2}\sigma}\right) \quad (3.3)$$

where  $\text{erfc}$  is the complementary error function defined as:

$$\text{erfc}(x) = 1 - \text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-t^2} dt.$$

- In this functional form, the parameter  $\mu$  determines the point of 50% of the plateau efficiency, and the  $\sigma$  gives the resolution. This parametrisation is used to benchmark the efficiency at the plateau, the turn-on points and resolution for each L1 Jet trigger. The choice of function is purely empirical. Previous studies used the error function alone, which described the data well at high threshold values but could not describe the efficiencies well at lower thresholds [69].
- The efficiency turn-on curves for various L1 jet thresholds are evaluated as a function of the offline reconstructed jet  $E_T$  for central jets with  $|\eta| < 3$ . These are measured using single isolated  $\mu$  triggers which have high statistics, and are orthogonal and therefore unbiased to the hadronic triggers under study. Events are selected with some loose detector based isolation requirements to make sure the muon does not overlap with a jet, causing a discrepancy in the measurement of the calorimetric energy.
- The efficiency is calculated with respect to offline Calo and PF Jets in Figure 3.11. Table 3.1 shows the values of these parameters, calculated for three example L1 single jet triggers taken from 2012 8 TeV data.



**Figure 3.11:** L1 jet efficiency turn-on curves as a function of the offline CaloJet  $E_T$  (left) and PFJet  $E_T$  (right), measured in 2012 Run Period C data and collected with an isolated single  $\mu$  data sample.

- The results from the L1 single jet triggers shows good performance for both Calo and PF jets. A better resolution is observed for Calo jets with respect to L1 jet quantities. This effect is due to Calo jet reconstruction using the same detector subsystems as for L1 jets. In contrast the PF jet reconstruction algorithm additionally utilises tracker and muon information, resulting in a poorer resolution when directly compared to L1 jet objects.

Trigger	Calo		PF	
	$\mu$	$\sigma$	$\mu$	$\sigma$
L1_SingleJet16	$21.09 \pm 0.03$	$7.01 \pm 0.02$	$22.17 \pm 0.04$	$7.83 \pm 0.03$
L1_SingleJet36	$41.15 \pm 0.05$	$5.11 \pm 0.02$	$39.16 \pm 0.06$	$8.04 \pm 0.03$
L1_SingleJet92	$95.36 \pm 0.13$	$5.62 \pm 0.03$	$90.85 \pm 0.19$	$11.30 \pm 0.10$

**Table 3.1:** Results of a cumulative **EMG** function fit to the turn-on curves for L1 single jet triggers in run 2012 Run Period C, measured in an isolated  $\mu$  data sample. The turn-on point,  $\mu$ , and resolution,  $\sigma$ , of the L1 jet triggers are measured with respect to offline Calo Jets (left) and PF Jets (right).

#### 1191 3.4.4. Effects of the L1 Jet Seed

1192 Between run period B and C of the 2012 data taking period, a jet seed threshold  
1193 was introduced into the L1 jet clustering algorithm. There was previously no direct  
1194 requirement made on the energy deposited in the central region.

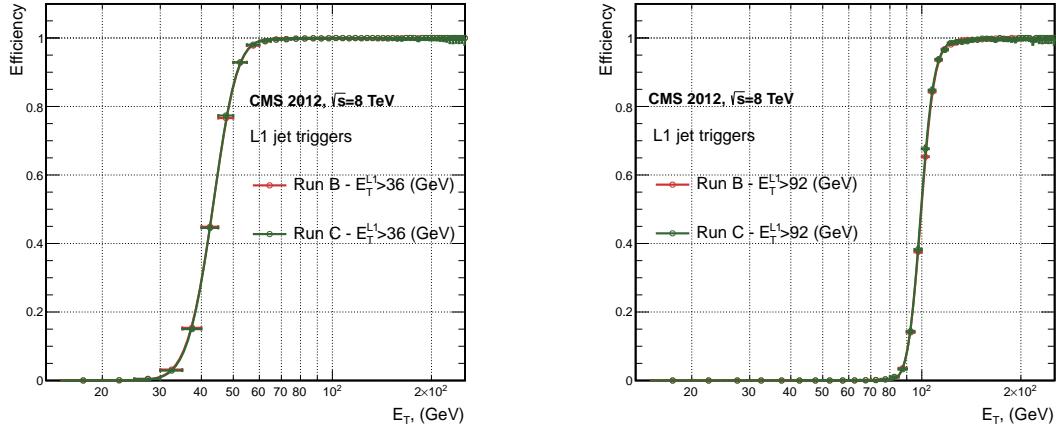
1195 The introduction of a jet seed threshold required that the central region have an uncor-  
1196 rected energy deposit of  $E_T \geq 5$  GeV. This value was motivated by studies of the effect  
1197 that different jet seed thresholds had upon the trigger cross-sections and efficiencies of  
1198 various  $H_T$ , single jet and multi-jet triggers. It was found that the 5 GeV threshold gave  
1199 large reductions in trigger cross-sections particularly in the case of multi-jet and  $H_T$   
1200 triggers, whilst having a small impact on the measured efficiencies of these triggers [70].

1201 Its main purpose was to counteract the effects of high pile up running conditions which  
1202 create a large number of soft non-collimated jets, that are then added to the jets from  
1203 the primary interaction or other soft jets from other secondary interactions [71]. This in  
1204 turn causes a large increase in trigger rate, due to the increase in the likelihood that the  
1205 event causes the L1 trigger to fire.

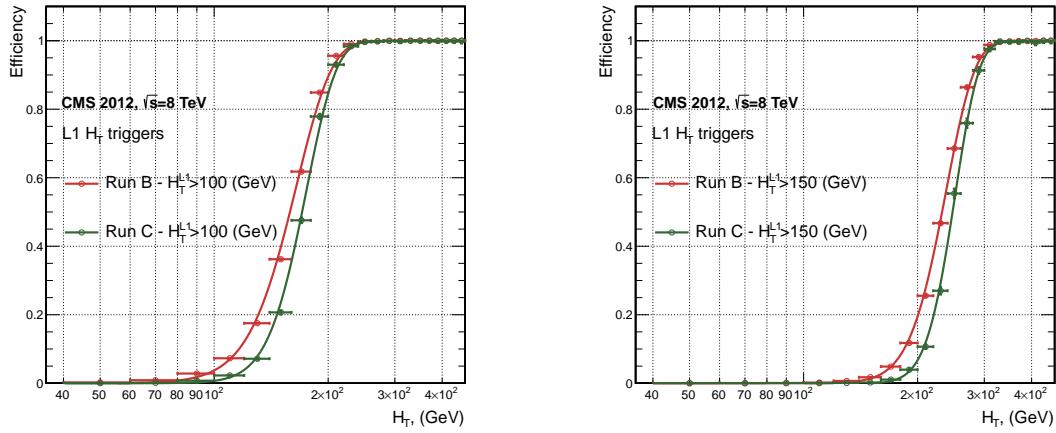
1206 The effect of the introduction of this jet seed threshold between these two run periods is  
1207 benchmarked through a comparison of the efficiency of the L1 jet triggers with respect  
1208 to offline Calo jets and is shown in Figure 3.12.

1209 The L1  $H_T$  trigger efficiency is also benchmarked at two values, which is shown in Figure  
1210 3.13. The L1  $H_T$  sum is compared against the offline  $H_T$  constructed from Calo jets  
1211 with  $E_T \geq 40$  GeV. This requirement is imposed to account for the relative difference  
1212 between uncorrected jet energy deposits within the **GCT** used to calculate the L1  $H_T$   
1213 sum, and those same deposits after full object reconstruction has occurred.

1214 To negate any effects from different pile-up conditions in the run periods, the efficiencies  
1215 are measured in events which contain between 15 and 20 primary vertices, as defined in  
1216 Appendix (A.2).



**Figure 3.12:** L1 jet efficiency turn-on curves as a function of the offline CaloJet  $E_T$ , measured for the L1 SingleJet 36 and 92 trigger in 2012 run period B and C collected with an isolated single  $\mu$  sample.



**Figure 3.13:** L1  $H_T$  efficiency turn-on curves as a function of the offline CaloJet  $H_T$ , measured for the L1  $H_T$  100 and 150 trigger during the run 2012 B and C, collected using an isolated single  $\mu$  triggered sample.

1217 It can be seen that the performance of the  $E_T > 36, 92$  single jet triggers are almost  
1218 identical, with the jet seed having no measurable effect on these triggers as shown in  
1219 Table 3.2.

1220 In the case of the  $H_T$  triggers, without the jet seed threshold a large increase in the  
1221 trigger cross-section during high luminosity collisions will occur. The low energy threshold

Trigger	2012B		2012C	
	$\mu$	$\sigma$	$\mu$	$\sigma$
L1_SingleJet36	$40.29 \pm 0.04$	$5.34 \pm 0.02$	$40.29 \pm 0.11$	$5.21 \pm 0.05$
L1_SingleJet92	$94.99 \pm 0.09$	$5.93 \pm 0.06$	$94.82 \pm 0.29$	$5.74 \pm 0.18$

**Table 3.2:** Results of a cumulative **EMG** function fit to the turn-on curves for L1 single jet triggers in the 2012 run period B and C, preselected on an isolated muon trigger. The turn-on point  $\mu$  and resolution  $\sigma$  of the L1 jet triggers are measured with respect to offline Calo Jets in run B (left) and run C (right).

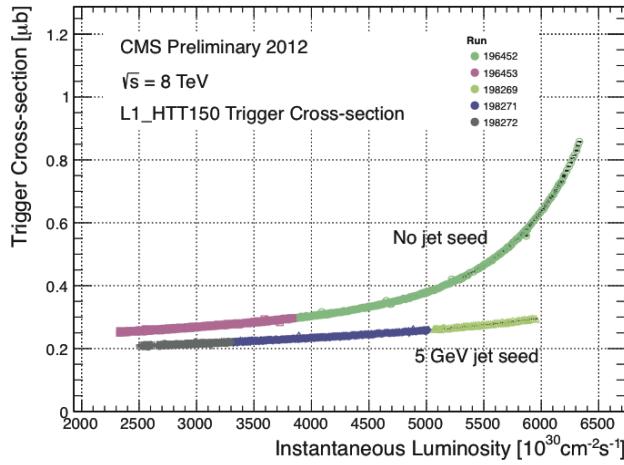
requirement for a jet to be clustered and added to the L1  $H_T$  sum, will allow many soft jets from other secondary interactions to enter the calculation. The introduction of the jet seed threshold prevents the clustering of many of these diffuse low  $E_T$  pile-up jets, thus lowering the L1 **GCT**  $H_T$  calculation. Resultantly, different behaviours for the trigger turn-ons after the introduction of the jet seed threshold are expected for these triggers.

The mean,  $\mu$ , values are measured to reside at higher  $H_T$  for both benchmarked  $H_T$  triggers, whilst a better resolution is observed after the introduction of the jet seed threshold. These values can be found within Table 3.3.

Trigger	2012B		2012C	
	$\mu$	$\sigma$	$\mu$	$\sigma$
L1 HT-100	$157.5 \pm 0.08$	$32.9 \pm 0.08$	$169.8 \pm 0.08$	$28.7 \pm 0.03$
L1 H1-150	$230.9 \pm 0.02$	$37.3 \pm 0.01$	$246.4 \pm 0.16$	$31.8 \pm 0.05$

**Table 3.3:** Results of a cumulative **EMG** function fit to the turn-on curves for  $H_T$  in run 2012 B and C, preselected on an isolated single  $\mu$  trigger. The turn-on point  $\mu$ , resolution  $\sigma$  of the L1  $H_T$  triggers are measured with respect to offline  $H_T$ , formed from CaloJets with a  $E_T \geq 40$  in run period B (left) and C (right).

Despite this slight increase in the turn-on point of the  $H_T$  triggers, a large reduction in the trigger cross-section is achieved for all  $H_T$  triggers. As an example, the expected trigger cross-section for the L1HTT150 trigger as a function of instantaneous luminosity can be seen in Figure 3.14.



**Figure 3.14:** Trigger cross section for the L1HTT150 trigger path. Showing that a 5 GeV jet seed threshold dramatically reduces the dependance of cross section on the instantaneous luminosity for L1  $H_T$  triggers [72].

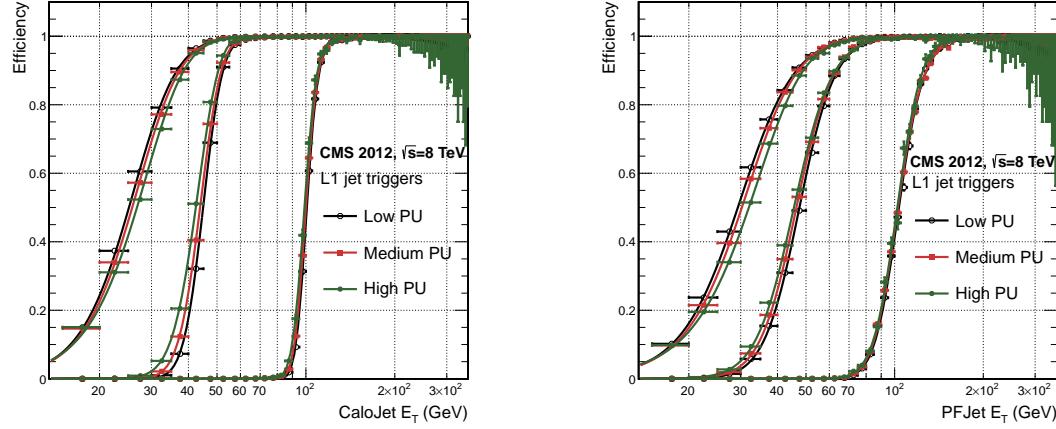
1235 It can be seen that this slight degradation in the offline value at which these  $H_T$  triggers  
 1236 are full efficient due to the jet seed threshold, can be justified from the large reduction in  
 1237 the trigger cross-section rate. Any inefficiencies can then if necessary be compensated  
 1238 through a reduction in the trigger threshold of the L1 seed.

### 1239 3.4.5. Robustness of L1 Jet Performance against Pile-up

1240 The performance of the L1 single jet triggers is evaluated in different pile-up conditions  
 1241 to determine any dependence on pile-up. Three different pile-up categories of 0-10,  
 1242 10-20 and  $>20$  vertices are defined, reflecting the low, medium and high pile-up running  
 1243 conditions at CMS in 2012.

1244 The L1 triggers are benchmarked relative to Calo and PF jets in the run period where  
 1245 the jet seed threshold *is* applied, for the L1 single jet thresholds of 16, 36 and 92 GeV,  
 1246 shown in Figure 3.15. The results of fitting an EMG function to these efficiency turn-on  
 1247 curves are given in Table 3.4 and Table 3.5 for Calo and PF jets respectively.

1248 No significant drop in efficiency is observed in the presence of a high number of primary  
 1249 vertices. The increase in hadronic activity in higher pile-up conditions, combined with  
 1250 the absence of pile-up subtraction for L1 jets, results in the expected observation of  
 1251 a decrease in the  $\mu$  value of the efficiency turn-ons as a function of pile-up. Similarly,  
 1252 the resolution,  $\sigma$ , of the turn-ons are found to gradually grow due to the larger pile-up  
 1253 corrections being applied to the offline reconstructed jets.



**Figure 3.15:** L1 jet efficiency turn-on curves as a function of the leading offline  $E_T$  Calo (left) and PF (right) jet, for low, medium and high pile-up conditions.

Vertices	0-10		11-20		> 20	
	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$
L1_SingleJet16	$19.9 \pm 0.1$	$6.1 \pm 0.3$	$20.8 \pm 0.1$	$6.5 \pm 0.1$	$22.3 \pm 0.2$	$7.5 \pm 0.1$
L1_SingleJet36	$41.8 \pm 0.1$	$4.6 \pm 0.1$	$40.9 \pm 0.1$	$5.1 \pm 0.1$	$40.6 \pm 0.6$	$5.9 \pm 0.2$
L1_SingleJet92	$95.9 \pm 0.2$	$5.4 \pm 0.1$	$95.2 \pm 0.2$	$5.6 \pm 0.1$	$94.5 \pm 0.6$	$6.2 \pm 0.3$

**Table 3.4:** Results of a cumulative EMG function fit to the efficiency turn-on curves for L1 single jet triggers in the 2012 run period C, measured from isolated  $\mu$  triggered data. The turn-on point,  $\mu$ , and resolution,  $\sigma$ , of the L1 jet triggers are measured with respect to offline Calo jets in low (left), medium (middle) and high (right) pile-up conditions.

<sup>1254</sup> These features are further emphasised when shown as a function of

$$\frac{(L1 E_T - Offline E_T)}{Offline E_T} \quad (3.4)$$

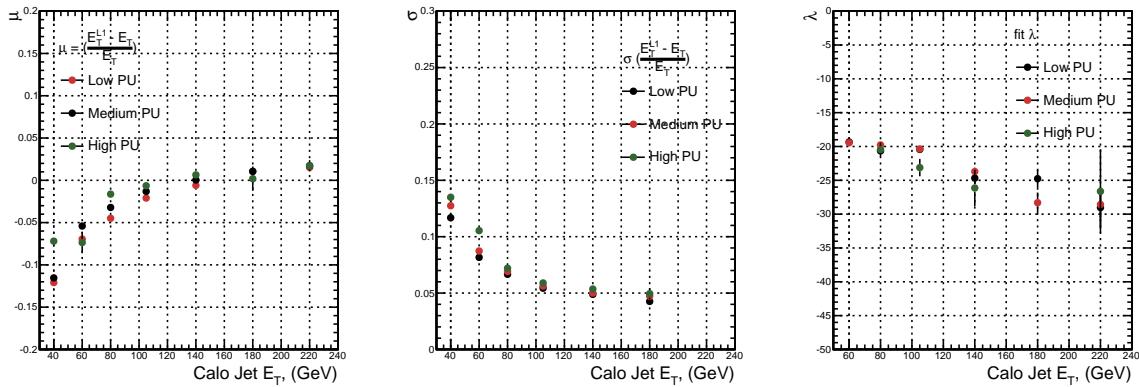
<sup>1255</sup> in bins of matched leading offline jet  $E_T$ . The results of these individual fits categorised  
<sup>1256</sup> as a function of matched leading offline jet  $E_T$  can be found in Appendix (B.2), where  
<sup>1257</sup> each of the distributions are fitted with an EMG function as defined in Equation (3.3).

<sup>1258</sup> The  $\mu$ ,  $\sigma$  and  $\lambda$  values extracted for the low, medium and high pile-up conditions are  
<sup>1259</sup> shown for Calo and PF jets in Figure 3.16 and Figure 3.17 respectively. The central  
<sup>1260</sup> value of  $\frac{(L1 E_T - Offline E_T)}{Offline E_T}$  is observed, as expected, to increase as a function of jet  $E_T$ ,  
<sup>1261</sup> whilst the resolution also improves as a function of increasing offline jet  $E_T$  for all pile-up  
<sup>1262</sup> categories.

Vertices	0-10		11-20		> 20	
	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$
L1_SingleJet16	21.1 $\pm$ 0.1	7.16 $\pm$ 0.05	22.34 $\pm$ 0.1	7.9 $\pm$ 0.1	24.6 $\pm$ 0.2	9.5 $\pm$ 0.1
L1_SingleJet36	39.6 $\pm$ 0.1	7.4 $\pm$ 0.1	38.4 $\pm$ 0.1	7.4 $\pm$ 0.1	37.1 $\pm$ 0.2	7.5 $\pm$ 0.1
L1_SingleJet92	91.6 $\pm$ 0.3	11.3 $\pm$ 0.2	91.4 $\pm$ 0.3	11.2 $\pm$ 0.1	90.0 $\pm$ 0.9	12.1 $\pm$ 0.4

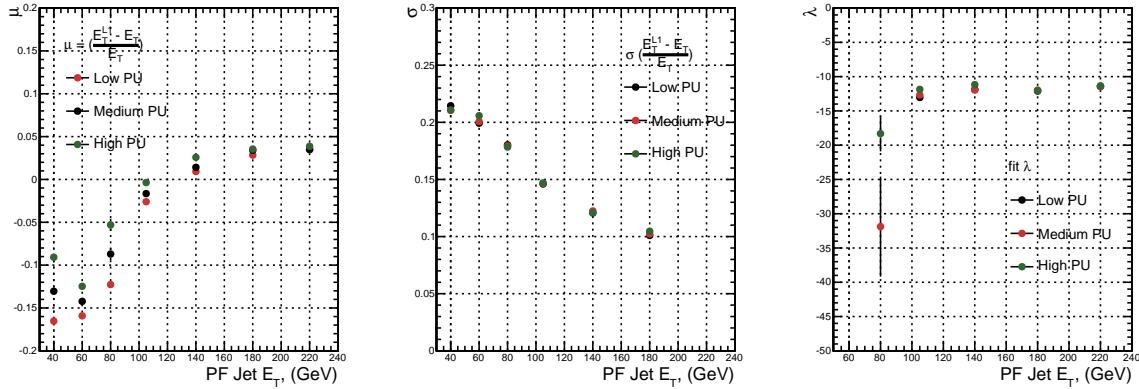
**Table 3.5:** Results of a cumulative EMG function fit to the efficiency turn-on curves for Level-1 single jet triggers in the 2012 run period C, measured from isolated  $\mu$  triggered data. The turn-on point,  $\mu$ , and resolution,  $\sigma$ , of the L1 jet triggers are measured with respect to offline PF jets in low (left), medium (middle) and high (right) pile-up conditions.

When comparisons are made between the individual pile-up scenarios, it can be seen that in the presence of higher pile-up,  $\mu$  is seen to shift to larger values and a poorer resolution,  $\sigma$ , observed. This is particularly evident at low lead jet transverse energy values. These differences between the different pile-up scenarios, can once again be attributed to an increasing number of soft pile-up jets that add to the transverse energy of the lead jet from the primary interaction within each successive pile-up category. However, when comparisons of the trigger performance at larger lead jet transverse energy values ( $> 100$  GeV) are made, similar performance is observed between the separate pile-up categories.



**Figure 3.16:** Fit values from an EMG function fitted to the resolution plots of leading Calo jet  $E_T$  measured as a function of  $\frac{(L1 E_T - \text{Offline } E_T)}{\text{Offline } E_T}$  for low, medium and high pile-up conditions. The plots show the mean  $\mu$  (left), resolution  $\sigma$  (middle) of the Gaussian as well as the decay term  $\lambda$  (right) of the exponential.

The resolution of other L1 jet based energy sum quantities,  $H_T$  and  $\mathcal{H}_T$  parameterised as in Equation (3.4), can be found in Appendix (B.3).



**Figure 3.17:** Fit values from an EMG function fitted to the resolution plots of leading PF jet  $E_T$  measured as a function of  $\frac{(L1\ E_T - \text{Offline}\ E_T)}{\text{Offline}\ E_T}$  for low and medium pile-up conditions. The plots show the mean  $\mu$  (left), resolution  $\sigma$  (middle) of the Gaussian, as well as the decay term  $\lambda$  (right) of the exponential.

### 3.4.6. Summary

The performance of the CMS Level-1 Trigger has been studied and evaluated for jets and energy sum quantities using data collected during the 2012 LHC 8 TeV run. These studies include the effect of the introduction of a 5 GeV jet seed threshold into the jet clustering algorithm. The purpose of this change was to mitigate the increase in L1 trigger cross-sections, due to larger isotropic energy deposits from an increased number of secondary interactions, whilst not adversely affecting the efficiency of these triggers. Measurements are made for a range of L1 jet quantities and thresholds, where no significant change in the measured efficiencies that would indicate a noticeable effect on the overall triggering performance of the detector is observed.

# Chapter 4.

## 1283 SUSY Searches in Hadronic Final States 1284

- 1285 In this chapter a model independent search for **SUSY**, in hadronic final states with  $\cancel{E}_T$  1286 using the  $\alpha_T$  variable is introduced and described in detail. The results presented are 1287 based on a data sample of pp collisions collected in 2012 at  $\sqrt{s} = 8$  TeV, corresponding 1288 to an integrated luminosity of  $11.7 \pm 0.5 \text{ fb}^{-1}$  [5].
- 1289 The kinematic variable  $\alpha_T$  is motivated as a variable to provide strong rejection of the 1290 overwhelming QCD multi-jet background, which is prevalent to jets + $\cancel{E}_T$  final states at 1291 the **LHC**. This is achieved whilst maintaining sensitivity to a range of possible **SUSY** 1292 signals and is described in Section (4.1). The search and trigger strategy in addition to 1293 the event reconstruction and selection are outlined within Sections (4.2 - 4.3).
- 1294 The method in which the **SM** background is estimated using data driven control samples 1295 and an analytical technique to improve statistical precision at higher b-tag multiplicities 1296 is detailed within Section (4.5). Included in this section is a discussion on the impact of 1297 b-tagging and mis-tagging scale factors between data and simulation on any background 1298 predictions. Improved precision in estimating background yields at large number of 1299 b-tagged jets, is important in the context of sensitivity to natural **SUSY** models, first 1300 outlined in Section (2.4.1).
- 1301 A description of the formulation of appropriate systematic uncertainties to be applied 1302 to the background predictions to account for theoretical uncertainties, limitations in 1303 the modelling of event kinematics and instrumental effects is covered in Section (4.6). 1304 Similarly the systematic determination for the **SMS** signal samples used to interpret the 1305 physics reach of the analysis are examined in Section (4.7).

- Finally the statistical likelihood model to test the compatibility of the data with a **SMS** only hypothesis, and to interpret the observations within the context of **SMS** models is described in Section (4.8). The experimental reach of the analysis discussed within this thesis is interpreted in two classes of **SMS** models, both introduced in Section (2.4.1). The **SMS** models considered in this analysis are summarised in Table 4.1. For each model, the **LSP** is assumed to be the lightest neutralino.
- Within the table are also defined reference points, parameterised in terms of parent gluino/squark and **LSP** sparticle masses,  $m_{\text{parent}}$  and  $m_{\text{LSP}}$ , respectively. These are used within the following two chapters to demonstrate potential signal yields within the hadronic search region of the analysis. The masses of each signal topology are chosen to reflect parameter space which is within the expected sensitivity reach of the search.

Model	Production/decay mode	Reference model	
		$m_{\text{parent}}$	$m_{\text{LSP}}$
G1 (T1)	$pp \rightarrow \tilde{g}\tilde{g}^* \rightarrow q\bar{q}\tilde{\chi}_1^0 q\bar{q}\tilde{\chi}_1^0$	700	300
G2 (T1bbbb)	$pp \rightarrow \tilde{g}\tilde{g}^* \rightarrow b\bar{b}\tilde{\chi}_1^0 b\bar{b}\tilde{\chi}_1^0$	900	500
G3 (T1tttt)	$pp \rightarrow \tilde{g}\tilde{g}^* \rightarrow t\bar{t}\tilde{\chi}_1^0 t\bar{t}\tilde{\chi}_1^0$	850	250
D1 (T2)	$pp \rightarrow \tilde{q}\tilde{q}^* \rightarrow q\bar{q}\tilde{\chi}_1^0 q\bar{q}\tilde{\chi}_1^0$	600	250
D2 (T2bb)	$pp \rightarrow \tilde{b}\tilde{b}^* \rightarrow b\bar{b}\tilde{\chi}_1^0 b\bar{b}\tilde{\chi}_1^0$	500	150
D3 (T2tt)	$pp \rightarrow \tilde{t}\tilde{t}^* \rightarrow t\bar{t}\tilde{\chi}_1^0 t\bar{t}\tilde{\chi}_1^0$	400	0

**Table 4.1:** A summary of the **SMS** models interpreted in this analysis, involving both direct (D) and gluino-induced (G) production of squarks and their decays. Reference models are also defined in terms of parent and **LSP** sparticle mass.

## 4.1. An Introduction to the $\alpha_T$ Search

- A proton-proton collision resulting in the production and decay of supersymmetric particles, would manifest as a final state containing energetic jets and  $\cancel{E}_T$  in the hadronic channel. The search focuses on topologies where new heavy supersymmetric, R-parity conserving particles are pair-produced in pp collisions. These particles decaying to a **LSP** escape the detector undetected, leading to significant missing energy and missing hadronic transverse energy,

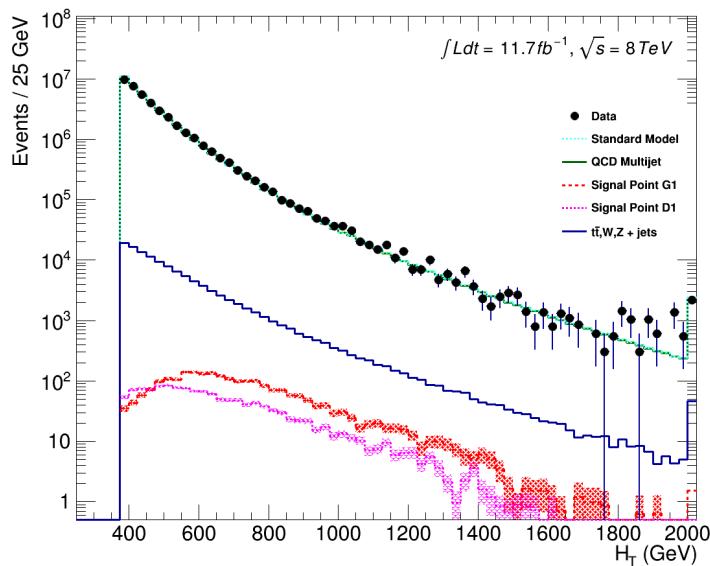
$$\cancel{H}_T = \left| \sum_{i=1}^n \vec{p_T}^{jet_i} \right|. \quad (4.1)$$

1324 This is defined as the vector sum of the transverse energies of jets selected in an event.  
1325 Energetic jets produced in the decay of these supersymmetric particles also can produce  
1326 significant visible transverse energy,

$$H_T = \sum_{i=1}^n E_T^{jet_i}, \quad (4.2)$$

1327 defined as the scalar sum of the transverse energies of jets selected in an event.

1328 A search within this channel is greatly complicated in a hadron collider environment;  
1329 where the overwhelming background comes from inherently balanced multi-jet (“QCD”) events,  
1330 which are produced with an extremely large cross-section as demonstrated within  
1331 Figure 4.1.  $\cancel{E}_T$  can appear in such events with a substantial mis-measurement, stochastic  
1332 fluctuations of jet energy, or missed objects due to detector mis-calibration or noise  
1333 effects.



**Figure 4.1:** Reconstructed offline  $H_T$  distribution in the hadronic signal selection, from  $11.7\text{fb}^{-1}$  of data, in which no  $\alpha_T$  requirement is made. Sample is collected from prescaled  $H_T$  triggers. Overlaid are expectations from simulation of EWK processes as well as two reference signal models (labelled G1 and D1 from Table 4.1).

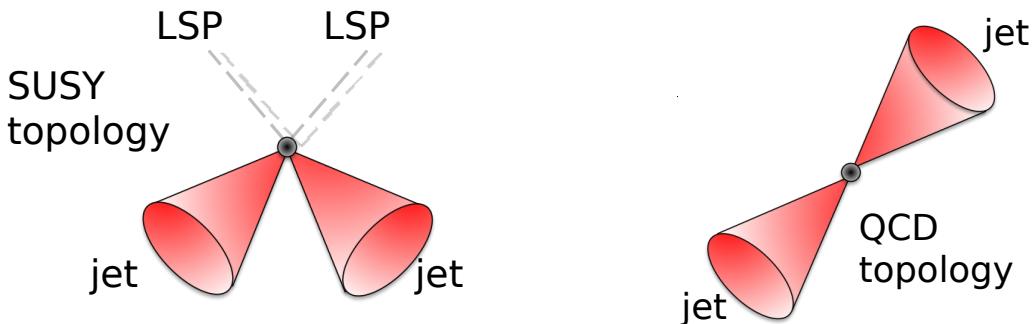
1334 Additional SM background from EWK processes with genuine  $\cancel{E}_T$  from escaping neutrinos  
1335 comprise the irreducible background within this search and come mainly from:

- 1336 •  $Z \rightarrow \nu\bar{\nu} + \text{jets}$ ,
- 1337 •  $W \rightarrow l\nu + \text{jets}$  in which a lepton falls outside of detector acceptance, is not reconstructed, is mis-identified, or the lepton decays hadronically  $\tau \rightarrow \text{had}$ ,
- 1339 •  $t\bar{t}$  with at least one leptonically decaying  $W$ , which is missed in the detector as detailed above,
- 1341 • small background contributions from DY, single top and Diboson (WW,ZZ,WZ) processes.

1343 The search is designed to have a strong separation between events with genuine and  
1344 “fake”  $\cancel{E}_T$  which is achieved primarily through the dimensionless kinematic variable,  $\alpha_T$   
1345 [73][74].

#### 1346 4.1.1. The $\alpha_T$ Variable

1347 For a perfectly measured di-jet QCD event, conservation laws dictate that both jets must  
1348 be of equal magnitude and produced in opposite directions. However, in the case of di-jet  
1349 events with genuine  $\cancel{E}_T$  (as detailed above), no such requirement is made of the two jets,  
 as depicted in Figure 4.2.



**Figure 4.2:** The event topologies of background QCD dijet events (right) and a generic SUSY signature with genuine  $\cancel{E}_T$  (left).

1350  
1351 Exploiting this feature leads to the formulation of  $\alpha_T$  (first inspired by [75]) in di-jet  
1352 systems defined as,

$$\alpha_T = \frac{E_T^{j_2}}{M_T}, \quad (4.3)$$

1353 where  $E_T^{j_2}$  is the transverse energy of the least energetic of the two jets and  $M_T$  defined  
 1354 as:

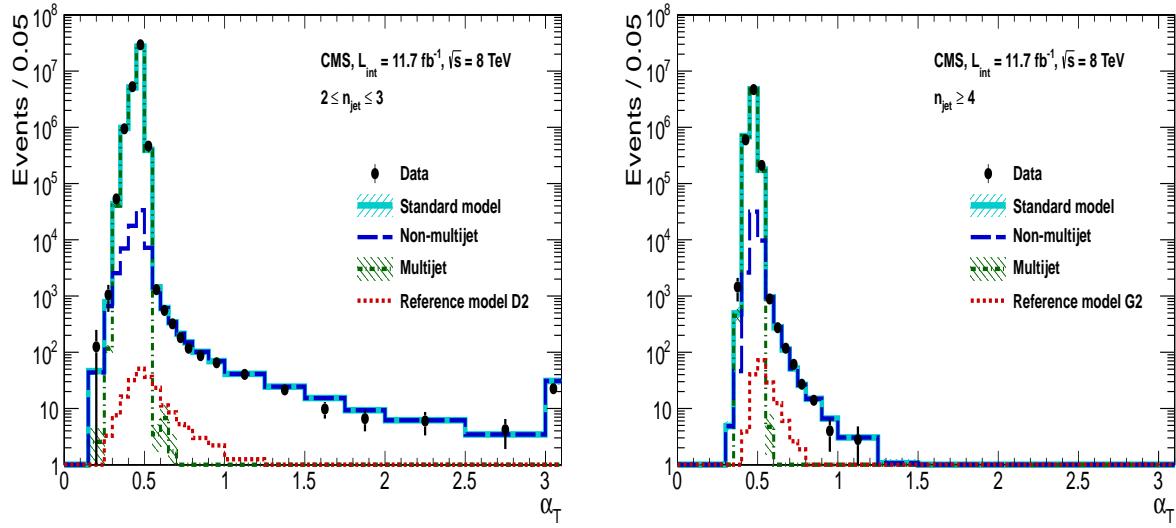
$$M_T = \sqrt{\left(\sum_{i=1}^2 E_T^{j_i}\right)^2 - \left(\sum_{i=1}^2 p_x^{j_i}\right)^2 - \left(\sum_{i=1}^2 p_y^{j_i}\right)^2} \equiv \sqrt{H_T^2 - \cancel{H}_T^2}. \quad (4.4)$$

1355 A perfectly balanced di-jet event i.e.  $E_T^{j_1} = E_T^{j_2}$  would yield an  $\alpha_T$  value of 0.5. In  
 1356 processes where a W or Z recoils off a system of jets, these jets will not necessarily be  
 1357 perfectly balanced and  $\alpha_T$  can then achieve values in excess of 0.5. Most importantly,  
 1358 balanced multi-jet events in which jets *are* mis-measured, will generally result in an  $\alpha_T$   
 1359 value of less than 0.5, thus giving the  $\alpha_T$  variable discriminating power between these  
 1360 processes.

1361  $\alpha_T$  can be further extended to apply to any arbitrary number of jets. This is undertaken  
 1362 by modelling a system of  $n$  jets as a di-jet system, through the formation of two pseudo-  
 1363 jets [76]. The two pseudo-jets are built by merging the jets present, and are chosen to  
 1364 be as balanced as possible, i.e the  $\Delta H_T \equiv |E_T^{pj_1} - E_T^{pj_2}|$  is minimised between the two  
 1365 pseudo jets. Using Equation (4.4),  $\alpha_T$  can be rewritten as,

$$\alpha_T = \frac{1}{2} \frac{H_T - \Delta H_T}{\sqrt{H_T^2 - \cancel{H}_T^2}} = \frac{1}{2} \frac{1 - \Delta H_T/H_T}{\sqrt{1 - (\cancel{H}_T/H_T)^2}}. \quad (4.5)$$

1366 The distribution of  $\alpha_T$  for the two jet multiplicity categories used within this analysis,  
 1367  $2 \leq n_{jet} \leq 3$  and  $n_{jet} \geq 4$  jets, is shown in the Figure 4.3. It can be seen that the  
 1368 distributions peak at an  $\alpha_T$  value of 0.5, before falling away sharply and being free  
 1369 of a simulated multi-jet background at larger  $\alpha_T$  values. These distributions serve to  
 1370 demonstrate the ability of the  $\alpha_T$  variable to discriminate between multi-jet events and  
 1371 EWK processes with genuine  $\cancel{E}_T$  in the final state.



**Figure 4.3:** The  $\alpha_T$  distributions for the low 2-3 (left) and high  $\geq 4$  (right) jet multiplicities after a full analysis selection and  $H_T > 375$  requirement. Data is collected using both prescaled  $H_T$  triggers and dedicated  $\alpha_T$  triggers for below and above  $\alpha_T = 0.55$  respectively. Expected yields as given by simulation are also shown for multi-jet events (green dash-dotted line), EWK backgrounds with genuine  $E_T$  (blue long-dashed line), the sum of all SM processes (cyan solid line) and the reference signal model D2 (left, red dotted line) or G2 (right, red dotted line).

1372 The  $\alpha_T$  requirement used within the search is chosen to be  $\alpha_T > 0.55$  to ensure that  
 1373 the QCD multi-jet background is negligible even in the presence of moderate jet mis-  
 1374 measurement. There still remain other effects which can cause multi-jet events to  
 1375 artificially have a large  $\alpha_T$  value, methods to combat them are discussed in detail in  
 1376 Section (4.2.2).

## 1377 4.2. Search Strategy

1378 The aim of the analysis presented in this thesis is to identify an excess of events in data  
 1379 over the SM background expectation in multi-jet final states and significant  $E_T$ . The  
 1380 essential suppression of the dominant multi-jet background for such a search is addressed  
 1381 by the  $\alpha_T$  variable, described in the previous section. For estimation of the remaining  
 1382 EWK backgrounds, three independent data control samples are used to predict the  
 1383 different processes that compose the background :

- 1384 •  $\mu + \text{jets}$  control sample to determine  $W + \text{jets}$ ,  $t\bar{t}$  and single top backgrounds,
- 1385 •  $\gamma + \text{jets}$  control sample to determine the irreducible  $Z \rightarrow \nu\bar{\nu} + \text{jets}$  background,

1386 •  $\mu\mu + \text{jets}$  control sample to also determine the irreducible  $Z \rightarrow \nu\bar{\nu} + \text{jets}$  background.

1387 These control samples are chosen to be rich in specific **EWK** processes, free of QCD  
1388 multi-jet events and to also be kinematically similar to the hadronic signal region that  
1389 they are estimating the backgrounds of, see Section (4.2.3). The redundancy of using the  
1390  $\gamma + \text{jets}$  and  $\mu\mu + \text{jets}$  sample to predict the same background within the signal region,  
1391 brings an opportunity to reliably crosscheck and validate the background estimation  
1392 method, and is utilised in both the determination of background estimation systematics  
1393 (Section(4.6)) and in the maximum likelihood fit (Section(4.8)).

1394 To remain inclusive to a large range of possible **SUSY** models, the signal region is split into  
1395 the following categories to allow for increased sensitivity to different **SUSY** topologies:

#### 1396 Sensitivity to a range of **SUSY** mass splittings

1397 The hadronic signal region is defined by  $H_T > 275$ , divided into eight bins in  $H_T$ .

- 1398 – Two bins of width 50 GeV in the range  $275 < H_T < 375$  GeV,
- 1399 – five bins of width 100 GeV in the range  $375 < H_T < 875$  GeV,
- 1400 – and a final open bin,  $H_T > 875$  GeV.

1401 The choice of the lowest  $H_T$  bin in the analysis is driven primarily by trigger  
1402 constraints. The mass difference between the **LSP** and the particle that it decays  
1403 from is an important factor in the amount of hadronic activity in the event.

1404 A large mass splitting will lead to hard high  $p_T$  jets which contribute to the  $H_T$  sum.  
1405 From Figure 4.1 it can be seen that the **SM** background falls sharply at high  $H_T$   
1406 values, therefore many  $H_T$  categories will lead to easier identification of such signals.  
1407 Conversely, smaller mass splittings lead to softer jet  $p_T$ 's which will subsequently  
1408 fall into the lower  $H_T$  range.

#### 1409 Sensitivity to production method of **SUSY** particles

1410 The production mechanism of any potential **SUSY** signal can lead to different event  
1411 topologies. One such way to discriminate between gluino ( $g\tilde{g}$  - “high multiplicity”),  
1412 and direct squark ( $q\tilde{q}$  - “low multiplicity”) induced production of **SUSY** particles is  
1413 realised through the number of reconstructed jets in the final state.

1414 The analysis is thus split into two jet categories:  $2 \leq n_{\text{jet}} \leq 3$  jets,  $n_{\text{jet}} \geq 4$  jets to  
1415 give sensitivity to both of these mechanisms.

---

**1416 Sensitivity to “Natural SUSY” via tagging jets from b-quarks**

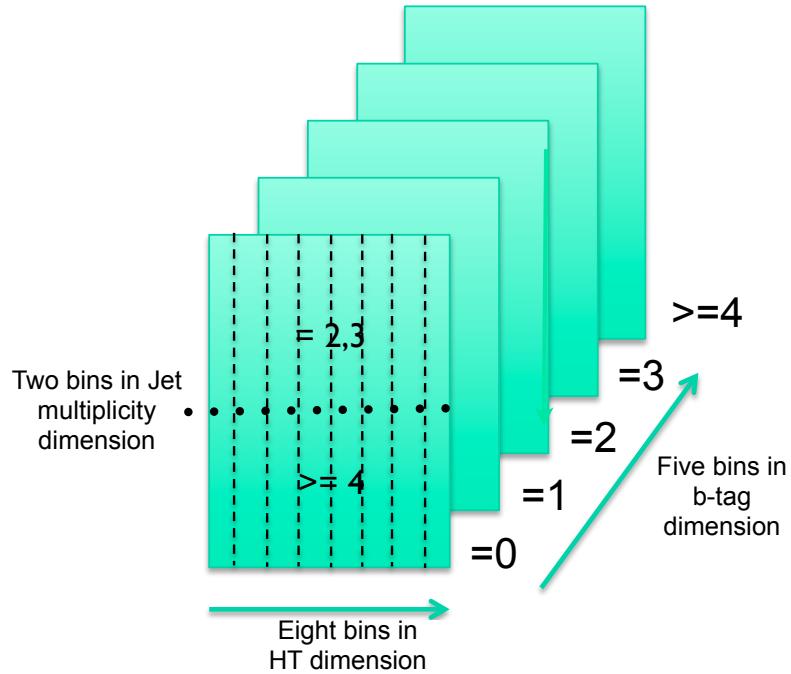
1417 Jets originating from bottom quarks (b-jets) are identified through vertices that  
1418 are displaced with respect to the primary interaction. The algorithm used to tag  
1419 b-jets is the Combined Secondary Vertex Medium Working Point (**CSV**M) tagger,  
1420 described within Section (3.3.2). A cut is placed on the discriminator variable of  
1421  $> 0.679$ , leading to a gluon/light-quark mis-tag rate of approximately 1% and a jet  
1422  $p_T$  dependant b-tagging efficiency of 60-70% [77].

1423 Natural **SUSY** models would be characterised through final-state signatures rich  
1424 in bottom quarks. A search relying on methods to identify jets originating from  
1425 bottom quarks through b-tagging, will significantly improve the sensitivity to this  
1426 class of signature. This gain in sensitivity stems from a vast reduction in the vector  
1427 boson + jet backgrounds (W, Z) at higher b-tag jet multiplicities, which typically  
1428 have no b-flavoured quarks in their decays.

1429 Therefore, events are categorised according to the number of b-tagged jets recon-  
1430 structed in each event, in the following:  $n_b^{\text{reco}} = 0, =1, =2, =3, \geq 4$  b-tag categories.  
1431 In the highest  $\geq 4$  b-tag category due to a limited number of expected signal and  
1432 background events, just three  $H_T$  bins are employed: 275-325 GeV, 325-375 GeV,  
1433  $\geq 375$  GeV.

1434 This characterisation is identically mirrored in all control samples, with the infor-  
1435 mation from all samples and b-tag categories used simultaneously in the likelihood  
1436 model, see Section (4.8).

1437 The combination of the  $H_T$ , jet multiplicity and b-tag categorisation of the signal region  
1438 as described above, results in 67 different bins in which the analysis is interpreted in. A  
1439 visualisation of the analysis categorisation is depicted in Figure 4.4.



**Figure 4.4:** Pictorial depiction of the analysis strategy employed by the  $\alpha_T$  search to increase sensitivity to a wide spectra of SUSY models.

### 1440 4.2.1. Physics Objects

1441 The physics objects used in the analysis are defined below, and follow the recommendation  
 1442 of the various CMS Physics Object Groups (POGs).

#### 1443 • Jets

1444 The jets used in this analysis are CaloJets, reconstructed as described in Section  
 1445 (3.3.1) using the anti- $k_T$  jet clustering algorithm.

1446 To ensure the jet object falls within the calorimeter systems a pseudo-rapidity  
 1447 requirement of  $|\eta| < 3$  is applied. Each jet must pass a “loose” identification criteria  
 1448 to reject jets resulting from unphysical energy, the criteria of which are detailed in  
 1449 Table A.1 [78].

#### 1450 • Muons

1451 Muons are selected in the  $\mu + \text{jets}$  and  $\mu\mu + \text{jets}$  control samples, and vetoed in  
 1452 the signal region. The same cut based identification criteria is applied to muons in  
 1453 both search regions and is summarised in Table 4.2 [79].

Variable	Definition
Is Global Muon	Muon contains both a hit in the muon chamber and a matched track in the inner tracking system.
$\chi^2 < 10$	$\chi^2$ of global muon track fit. Used to suppress hadronic punch-through and muons from decays in flight.
Muon chamber hits $> 0$	At least one muon chamber hit included in global muon track fit.
Muon station hits $> 1$	Muon segment hits in at least two muon stations, which suppresses hadronic punch-through and accidental track-to-segment matches.
$d_{xy} < 0.2\text{mm}$	The tracker track transverse impact parameter w.r.t the primary vertex. Suppresses cosmic muons and muons from decays in flight.
$d_z < 0.5\text{mm}$	The longitudinal distance of the tracker track w.r.t the primary vertex. Loose selection requirement to further suppress cosmic muons, muons from decays in flight and tracks from pileup.
Pixel hits $> 0$	Suppresses muons from decays in flight by requiring at least one pixel hit in the tracker.
Track layer hits $> 5$	Number of tracker layers with hits, to guarantee a good $p_T$ measurement. Also suppresses muons from decays in flight.
PF Iso $< 0.12$	Isolation based upon the sum of the charged and neutral hadrons and photon objects within a $\Delta R$ 0.4 cone of the muon object, corrected for pile up effects on the isolation sum.

**Table 4.2:** Muon identification criteria used within the analysis for selection/veto purposes in the muon control/signal selections.

1454      Additionally muons are required to be within the acceptance of the muon tracking  
 1455      systems. For the muon control samples, trigger requirements necessitate a  $|\eta| <$   
 1456      2.1 for the selection of muons. In the signal region where muons are vetoed, these  
 1457      conditions are relaxed to  $|\eta| < 2.5$  and a minimum threshold of  $p_T > 10 \text{ GeV}$  is  
 1458      required of muon objects.

### • Photons

1460      Photons are selected within the  $\gamma + \text{jets}$  control sample and vetoed in all other  
 1461      selections. Photons are identified in both cases according to the cut based criteria  
 1462      listed in Table 4.3 [80].

Variable	Definition
$H/E < 0.05$	The ratio of hadronic energy in the <b>HCAL</b> tower directly behind the <b>ECAL</b> super-cluster and the <b>ECAL</b> super-cluster itself.
$\sigma_{in\eta} < 0.011$	The log energy weighted width ( $\sigma$ ), of the extent of the shower in the $\eta$ dimension.

*Continued on next page*

R9 < 1.0	The ratio of the energy of the $3 \times 3$ crystal core of the super-cluster compared to the total energy stored in the $5 \times 5$ super-cluster.
Combined Isolation < 6 GeV	The photons are required to be isolated with no electromagnetic or hadronic activity within a radius $\Delta R = 0.3$ of the photon object. A combination of the pileup subtracted [81], <b>ECAL</b> , <b>HCAL</b> and tracking isolation sums are used to determine the combined total isolation value.

**Table 4.3:** Photon identification criteria used within the analysis for selection/veto purposes in the  $\gamma +$  jets control/signal selections.

1463

1464 Photon objects are also required to have a minimum momentum of  $p_T > 25$  GeV.

1465 **• Electrons**

1466 Electron identification is defined for veto purposes. They are selected according to  
1467 the following cut-based criteria listed in Table 4.4, utilising PF-based isolation.

Variable	Barrel	EndCap	Definition
$\Delta\eta_{In}$	<0.007	<0.009	$\Delta\eta$ between SuperCluster position and the coordinate of the associated track at the interaction vertex, assuming no radiation.
$\Delta\phi_{In}$	<0.15	<0.10	$\Delta\phi$ between SuperCluster position and track direction at interaction vertex extrapolated to ECAL assuming no radiation.
$\sigma_{in\eta\eta}$	<0.01	<0.03	Cluster shape covariance, measure the $\eta$ dispersion of the electrons electromagnetic shower over the <b>ECAL</b> supercluster.
H/E	<0.12	<0.10	The ratio of hadronic energy in the <b>HCAL</b> tower directly behind the <b>ECAL</b> super-cluster and the <b>ECAL</b> super-cluster itself.
d0 (vtx)	<0.02	<0.02	The tracker track transverse impact parameter w.r.t the primary vertex.
dZ (vtx)	<0.20	<0.20	The longitudinal distance of the tracker track w.r.t the primary vertex.
$ (\frac{1}{E_{ECAL}} - \frac{1}{p_{track}}) $	<0.05	<0.05	Comparison of energy at supercluster $1/E_{ECAL}$ and that of the track momentum at the vertex $1/p_{track}$ . Causes suppression of fake electrons at low $p_T$ .
PF Iso	<0.15	<0.15	Combined PF isolation of charged hadrons, photons, neutral hadrons within a $\Delta R < 0.3$ cone size. Isolation sum is corrected for pileup using effective area corrections for neutral particles.

**Table 4.4:** Electron identification criteria used within the analysis for veto purposes.

1468 Electrons are required to be identified at  $|\eta| < 2.5$ , with a minimum  $p_T > 10 \text{ GeV}$   
1469 threshold to ensure that the electrons fall within the tracking system of the detector.

1470 • Noise and  $\cancel{E}_T$  Filters

1471 A series of noise filters are applied to veto events which contain spurious non-physical  
1472 jets that are not picked up by the jet id, and events which give large unphysical  $\cancel{E}_T$   
1473 values. These filters are listed within Table 4.5.

Noise Filters	Variable	Definition
CSC tight beam halo filter		As proton beams circle the <b>LHC</b> , proton interactions with the residual gas particles or the beam collimators can occur, producing showers of secondary particles which can interact with the <b>CMS</b> detector.
HBHE noise filter with isolated noise rejection		Anomalous noise in the <b>HCAL</b> not due to electronics noise, but rather due to instrumentation issues associated with the <b>HPD</b> 's and Readout Boxes ( <b>RBXs</b> ).
HCAL laser filter		The <b>HCAL</b> uses laser pulses for monitoring the detector response. Some laser pulses have accidentally been fired in the physics orbit, and ended up polluting events recorded for physics analysis.
ECAL dead cell trigger primitive (TP) filter		<b>EB</b> and <b>EE</b> have single noisy crystals which are masked in reconstruction. Use the Trigger Primitive ( <b>TP</b> ) information to assess how much energy was lost in masked cells.
Bad EE Supercrystal filter		Two supercrystals in <b>EE</b> are found to occasionally produce high amplitude anomalous pulses in several channels at once, causing a large $\cancel{E}_T$ spike.
ECAL Laser correction filter		A laser calibration multiplicative factor is applied to correct for transparency loss in each crystal during irradiation. A small number of crystals receive unphysically large values of this correction and become very energetic, resulting in $\cancel{E}_T$ .

**Table 4.5:** Noise filters that are applied to remove spurious and non-physical  $\cancel{E}_T$  signatures within the **CMS** detector.

1474 4.2.2. Event Selection

1475 The selection criteria for events within the analysis are detailed below. A set of common  
1476 cuts are applied to both signal (maximise acceptance to a range of **SUSY** signatures), and  
1477 control samples (retain similar jet kinematics for background predictions), with additional  
1478 selection cuts applied to each control sample to enrich the sample in a particular **EWK**  
1479 processes, see Section (4.2.3).

- 1480 The jets considered in the analysis are required to have a transverse momentum  $p_T > 50$   
1481 GeV, with a minimum of two jets required in the event. The highest  $E_T$  jet is required  
1482 to lie within the central tracker acceptance  $|\eta| < 2.5$ , and the two leading  $p_T$  jets must  
1483 each have  $p_T > 100$  GeV. Any event which has a jet with  $p_T > 50$  GeV that either fails  
1484 the “loose” identification criteria described in Section(4.2.1) or has  $|\eta| > 3.0$ , is rejected.  
1485 Similarly events in which an electron, muon or photon fails object identification but  
1486 passes  $\eta$  and  $p_T$  restrictions, are identified as an “odd” lepton/photon and the event is  
1487 vetoed.
- 1488 At low  $H_T$ , the jet  $p_T$  threshold requirements required to be considered as part of the  
1489 analysis and enter the  $H_T$  sum are scaled downwards. These are scaled down in order  
1490 to extend phase space at low  $H_T$ , preserving similar jet multiplicities and background  
1491 admixture seen at higher  $H_T$ , as listed in Table 4.6.

$H_T$ bin	minimum jet $p_T$	second leading jet $p_T$
$275 < H_T < 325$	36.7	73.3
$325 < H_T < 375$	43.3	86.6
$375 < H_T$	50.0	100.0

**Table 4.6:** Jet thresholds used in the three  $H_T$  regions of the analysis.

- 1492 Within the signal region, to suppress SM processes with genuine  $\cancel{E}_T$  from neutrinos,  
1493 events containing isolated electrons or muons are vetoed. Furthermore to ensure a pure  
1494 multi-jet topology, events are vetoed if an isolated photon is found with  $p_T > 25$  GeV.  
1495 An  $\alpha_T$  requirement of  $> 0.55$  is required to reduce the QCD multi-jet background to  
1496 a negligible amount. Finally, additional cleaning cuts are applied to protect against  
1497 pathological deficiencies such as reconstruction failures or severe energy mis-measurements  
1498 due to detector inefficiencies:

- Significant  $H_T$  can arise in events with no real  $\cancel{E}_T$  due to multiple jets falling below the  $p_T$  threshold for selecting jets. This in turn leads to events which can then incorrectly pass the  $\alpha_T$  requirements of the analysis. This effect can be negated by requiring that the missing transverse momentum reconstructed from jets alone does not greatly exceed the missing transverse momentum reconstructed from all of the detector’s calorimeter towers,

$$R_{miss} = H_T / \cancel{E}_T < 1.25.$$

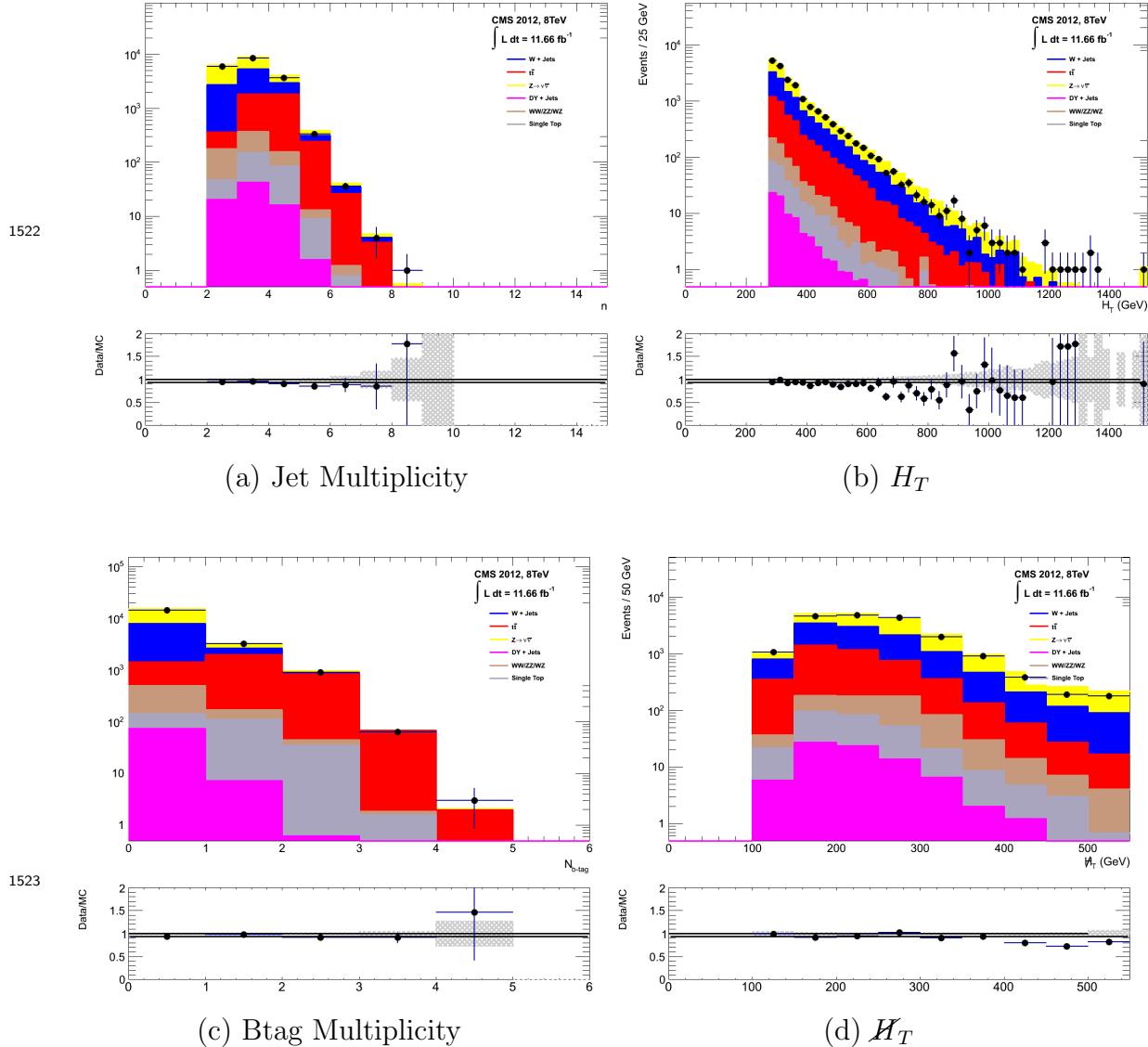
- 1499     • Fake  $\cancel{E}_T$  and  $\cancel{H}_T$  can arise due to significant jet mis-measurements caused by a small  
 1500       number of non-functioning **ECAL** regions. These regions absorb electromagnetic  
 1501       showers which are subsequently not added to the jet energy sum. To circumvent  
 1502       this problem the following procedure is employed: For each jet in the event, the  
 1503       angular separation

$$\Delta\phi_j^* \equiv \Delta\phi(\vec{p_j} - \sum_{i \neq j} \vec{p_i}), \quad (4.6)$$

1504       is calculated where that jet is itself removed from the event. Here  $\Delta\phi^*$  is a measure  
 1505       of how aligned the  $\cancel{H}_T$  of an event is with a jet. A small value (i.e. the  $\cancel{H}_T$  vector  
 1506       lies along the jet axis) is indicative of an inherently balanced event in which a jet has  
 1507       been mis-measured. For every jet in an event with  $\Delta\phi^* < 0.5$ , if the  $\Delta R$  distance  
 1508       between the selected jet and the closest dead **ECAL** region is also  $< 0.3$ , then the  
 1509       event is rejected. Similarly events are rejected if the jet points within  $\Delta R < 0.3$  of  
 1510       the **ECAL** barrel-endcap gap at  $|\eta| = 1.5$ .

1511       Some of the key distributions of the analysis are compared to simulated **SM** processes,  
 1512       shown in Figure 4.5. The simulated samples are normalised to a luminosity of  $11.7 \text{ fb}^{-1}$ ,  
 1513       with no requirement placed upon the number of b-tagged jets or number of jets in the  
 1514       distributions shown. In the case of this inclusive selection, the dominant backgrounds  
 1515       in the signal regions are,  $Z \rightarrow \nu\bar{\nu}$  and  $W + \text{jet}$  processes, with a smaller  $t\bar{t}$  background  
 1516       accompanied by other residual backgrounds.

1517       The distributions shown are presented for purely illustrative purposes, with the simulation  
 1518       not used in absolute terms for the estimation of background processes within the signal  
 1519       region, see Sections (4.2.3, 4.5). However it is nevertheless important to demonstrate  
 1520       that good agreement exists between the modelling of key variables in simulation and  
 1521       data.



**Figure 4.5:** Data/MC comparisons of key variables for the hadronic signal region, following the application of the hadronic selection criteria and the requirements of  $H_T > 275$  GeV and  $\alpha_T > 0.55$ . Bands represent the uncertainties due to the statistical size of the MC samples. No requirement is made upon the number of b-tagged jets or jet multiplicity in these distributions.

### **4.2.3. Control Sample Definition and Background Estimation**

1525 The method used to estimate the background contributions in the hadronic signal region  
 1526 relies on the use of a Transfer Factor (**TF**). This is determined from MC simulation  
 1527 in both the control,  $N_{MC}^{control}$ , and signal,  $N_{MC}^{signal}$ , region to transform the observed yield

1528 measured in data for a control sample,  $N_{\text{obs}}^{\text{control}}$ , into a background prediction,  $N_{\text{pred}}^{\text{signal}}$ , via  
1529 Equation (4.7),

$$N_{\text{pred}}^{\text{signal}} = \frac{N_{\text{MC}}^{\text{signal}}}{N_{\text{MC}}^{\text{control}}} \times N_{\text{obs}}^{\text{control}}. \quad (4.7)$$

1530 All simulation samples are normalised to the luminosity of the data samples of the  
1531 relevant selection they are being applied to. Through this method, “vanilla” predictions  
1532 for the **SM** background in the signal region can be made by considering separately the  
1533 sum of the prediction from the combination of either the  $\mu + \text{jets}$  and  $\gamma + \text{jets}$ , or  $\mu +$   
1534 jets and  $\mu\mu + \text{jets}$  samples.

1535 It must be noted that the final background estimation from which results are interpreted,  
1536 is calculated via a fitting procedure defined formally by the likelihood model described  
1537 in Section (4.8).

1538 The sum of the expected yields from all simulated processes listed in Section (4.1), enter  
1539 the denominator,  $N_{\text{MC}}^{\text{control}}$ , of the **TF** defined in Equation (4.7) for each control sample.  
1540 However, only the specific processes being estimated by the control sample enter the  
1541 numerator,  $N_{\text{MC}}^{\text{signal}}$ .

1542 For the  $\mu + \text{jets}$  sample the processes entering the numerator are,

$$N_{\text{MC}}^{\text{signal}}(H_T, n_{\text{jet}}) = N_W + N_{t\bar{t}} + N_{DY} + N_t + N_{di-boson}, \quad (4.8)$$

1543 whilst for both the  $\mu\mu + \text{jets}$  and  $\gamma + \text{jets}$  samples the only simulated processes used in  
1544 the numerator is,

$$N_{\text{MC}}^{\text{signal}}(H_T, n_{\text{jet}}) = N_{Z \rightarrow \nu\bar{\nu}}. \quad (4.9)$$

1545 The control samples and the **EWK** processes they are specifically tuned to select are  
1546 defined below, with distributions of key variables for each of the control samples shown  
1547 for illustrative purposes in Figures 4.6, 4.7 and 4.8. No requirement is placed upon the

number of b-tagged jets or jet multiplicity in the distributions shown. The distributions highlight the background compositions of each control sample, where in general, good agreement is observed between data and simulation, giving confidence that the samples are well understood. The contribution from QCD multi-jet events is expected to be negligible:

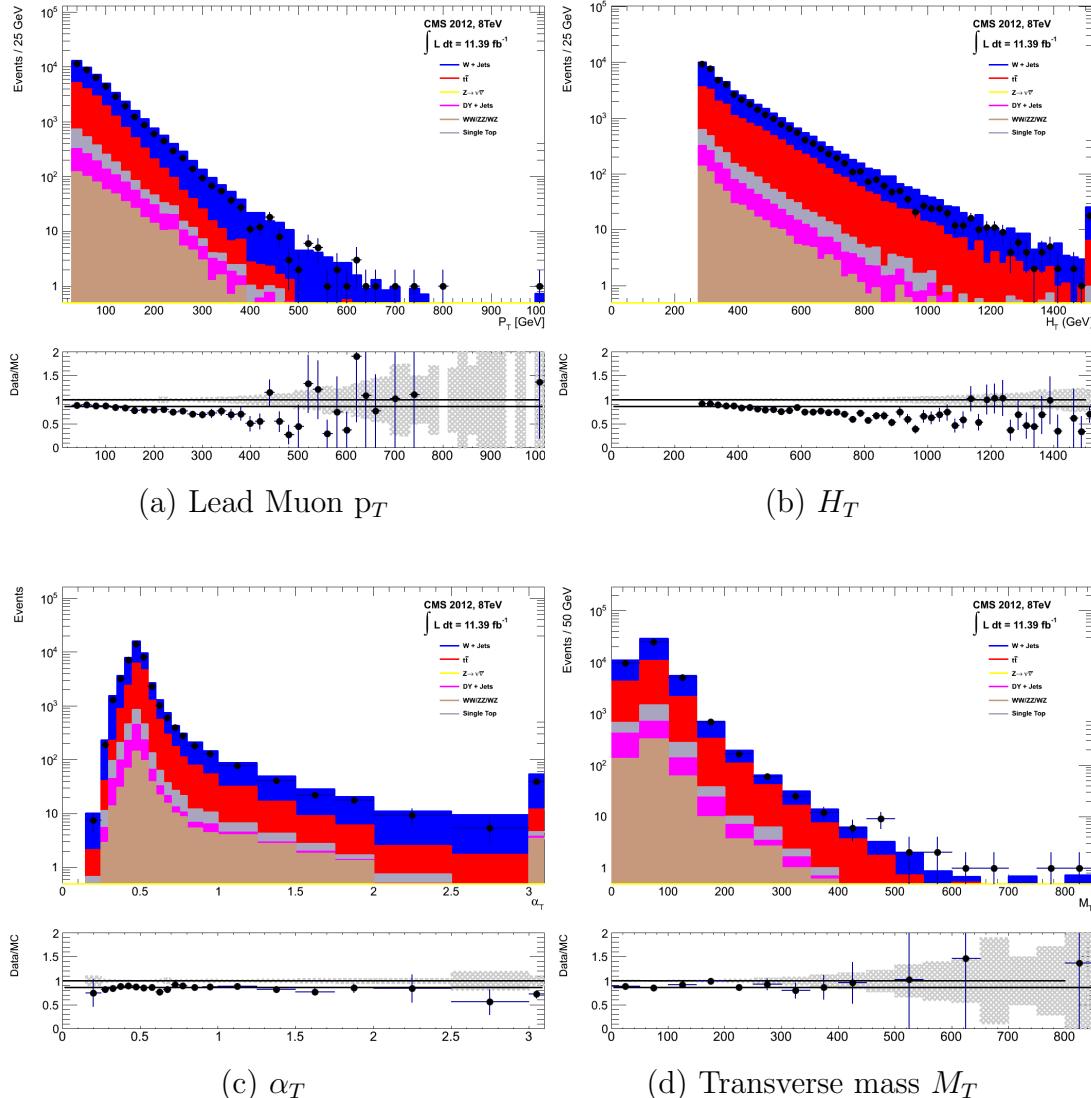
### 1553 The $\mu +$ jets control sample

Events from  $W +$  jets and  $t\bar{t}$  processes enter into the hadronic signal sample due to unidentified leptons from acceptance effects or reconstruction inefficiencies and hadronic tau decays. These leptons originate from the decay of high  $p_T$   $W$  bosons.

The control sample specifically identify  $W \rightarrow \mu\nu$  decays within a similar phase-space of the signal region, where the muon is subsequently ignored in the calculation of event level variables, i.e.  $H_T$ ,  $\mathcal{H}_T$ ,  $\alpha_T$ .

All kinematic jet-based selection criteria are identical to those applied in the hadronic search region (with the exception of an  $\alpha_T$ , requirement discussed below) detailed in Section (4.2.2), with the same  $H_T$ , jet multiplicity and b-jet multiplicity categorisation described previously. Furthermore, the following selection criteria are also required:

- Muons originating from  $W$  boson decays are selected by requiring one tightly isolated muon defined in Table 4.2, with a  $p_T > 30$  GeV and  $|\eta| < 2.1$ . Both of these thresholds arise from trigger restrictions.
- The transverse mass of the  $W$  candidate must satisfy  $M_T(\mu, \cancel{E}_T) > 30$  GeV (to suppress QCD multi-jet events).
- Events which contain a jet overlapping with a muon  $\Delta R(\mu, \text{jet}) < 0.5$  are vetoed to remove events from muons produced as part of a jet’s hadronisation process.
- Events containing a second muon candidate which has failed id, but passing  $p_T$  and  $|\eta|$  requirements, are checked to have an invariant mass that satisfies  $|M_{\mu\mu} - m_Z| > 25$ , thus removing  $Z \rightarrow \mu\mu$  contamination.



**Figure 4.6:** Data/MC comparisons of key variables for the  $\mu +$  jets selection, following the application of selection criteria and the requirements that  $H_T > 275$  GeV. Bands represent the uncertainties due to the statistical size of the MC samples. No requirement is made upon the number of b-tagged jets or jet multiplicity in these distributions.

## The $\mu\mu$ + jets control sample

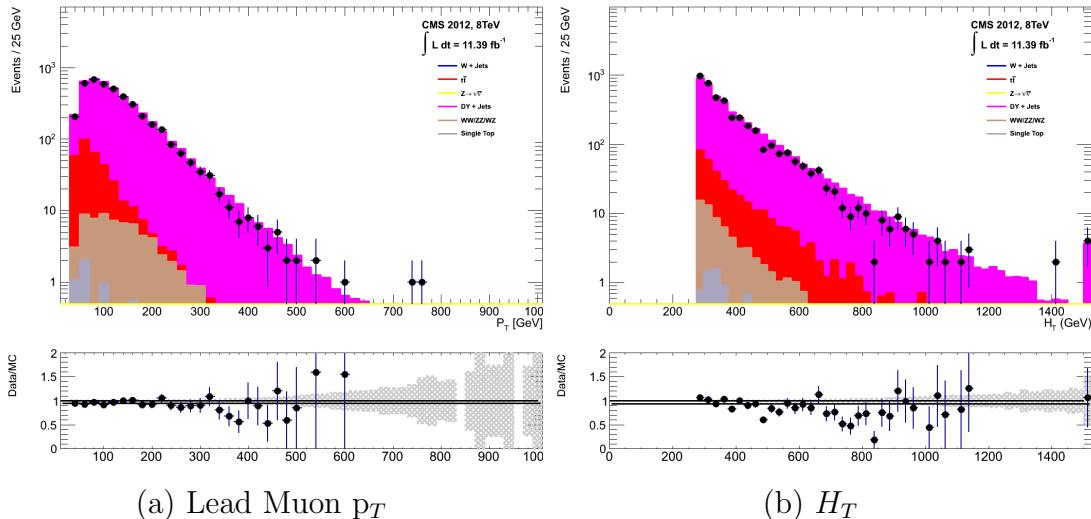
An irreducible  $Z \rightarrow \nu\bar{\nu} + \text{jets}$  background enters into the signal region from genuine  $E_T$  from the escaping neutrinos. This background is estimated using two control samples, the first of which is the  $Z \rightarrow \mu\bar{\mu} + \text{jets}$  process, which posses identical kinematic properties, but with different acceptance and branching ratio [1].

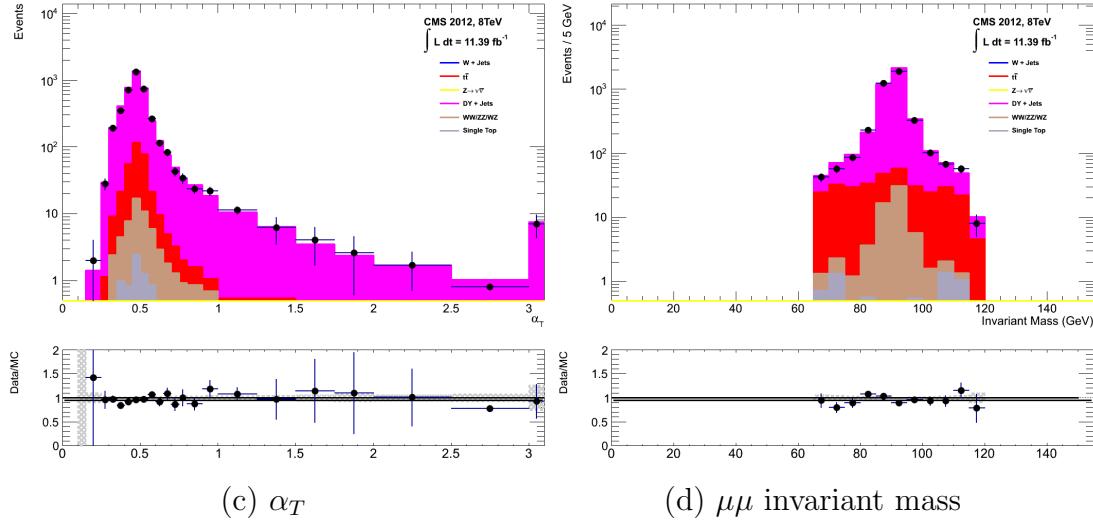
The same acceptance requirements as the  $\mu + \text{jets}$  selection for muons is applied, as defined in Table 4.2. Muons in the event are ignored for the purpose of the

calculation of event level variables. In addition to kinematic jet-based selection criteria (with the exception of an  $\alpha_T$  requirement, discussed below) and event categorisation, which are identical to the hadronic search region, the following selection criteria are also specified:

- Muons originating from a Z boson decay are selected, requiring exactly two tightly isolated muons. Due to trigger requirements the leading muon is required to have  $p_T > 30$  GeV and  $|\eta| < 2.1$ . The requirement of the  $p_T$  on the second muon is relaxed to 10 GeV.
- Events are vetoed if containing a jet overlapping with a muon  $\Delta R(\mu, \text{jet}) < 0.5$ .
- In order to specifically select two muons both originating from a single Z boson decay, the invariant mass of the two muons must satisfy  $|M_{\mu\mu} - m_Z| < 25$ .

The  $\mu\mu + \text{jets}$  sample is able to make predictions in the signal region of the two lowest  $H_T$  bins, providing coverage where the  $\gamma + \text{jets}$  sample is unable to, due to trigger requirements. In higher  $H_T$  bins, the higher event statistics of the  $\gamma + \text{jets}$  sample is also used in determining the  $Z \rightarrow \nu\bar{\nu}$  estimation.





**Figure 4.7:** Data/MC comparisons of key variables for the  $\mu\mu + \text{jets}$  selection, following the application of selection criteria and the requirements that  $H_T > 275$  GeV. Bands represent the uncertainties due to the statistical size of the MC samples. No requirement is made upon the number of b-tagged jets or jet multiplicity in these distributions.

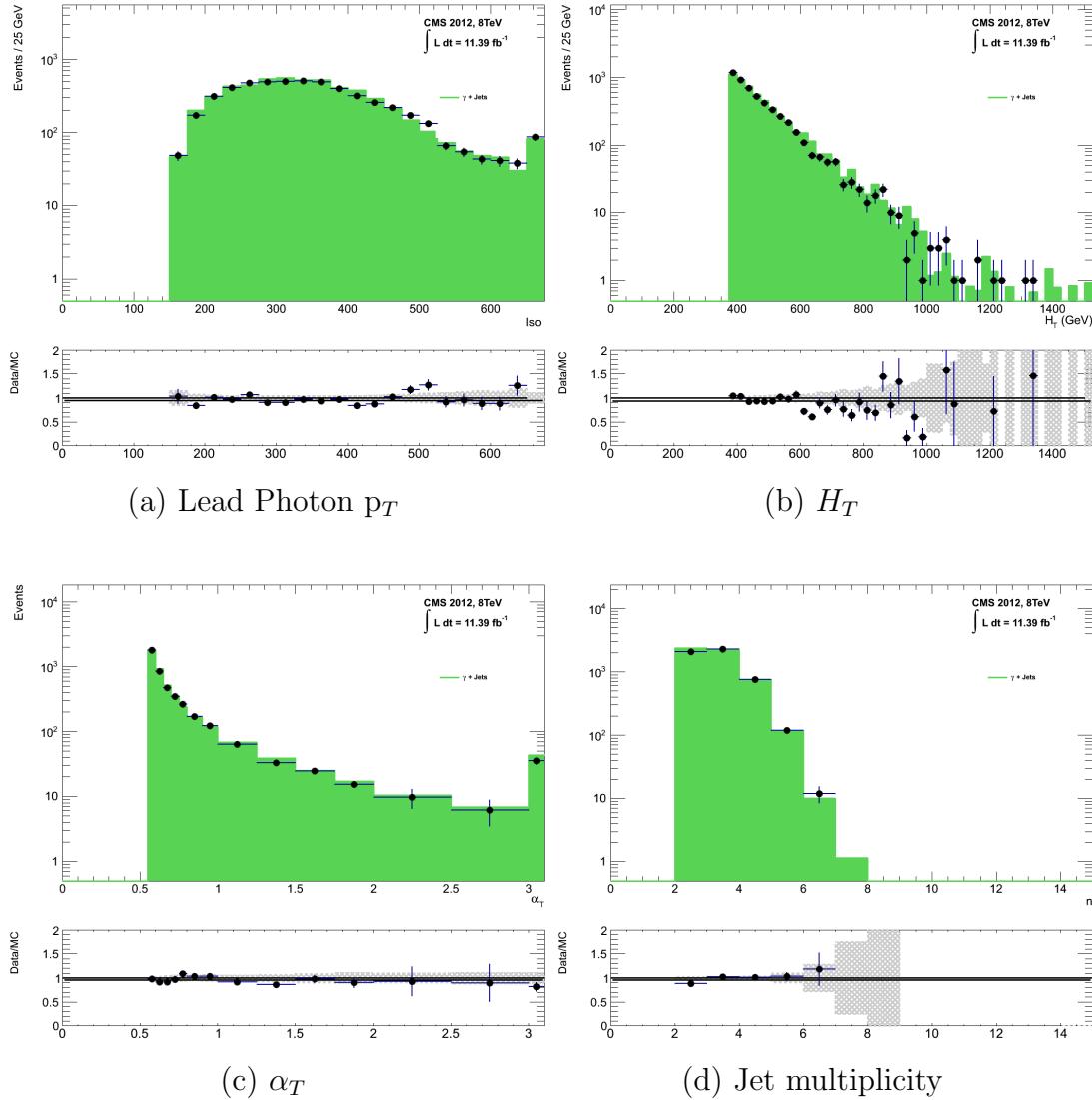
### The $\gamma + \text{jets}$ control sample

The  $Z \rightarrow \nu\bar{\nu} + \text{jets}$  background is also estimated from a  $\gamma + \text{jets}$  control sample. When the  $E_T$  of the photon is greater than the mass of the  $Z$ , it possesses a larger cross-section and has kinematic properties similar to those of  $Z \rightarrow \nu\bar{\nu}$  events if the photon is ignored [82].

Within the control channel, the photon is ignored for the purpose of the calculation of event level variables, and identical selection criteria to the hadronic signal region are applied. In addition the follow requirements are also made:

- Exactly one photon is selected, satisfying identification criteria as detailed in Table 4.3, with a minimum  $p_T > 165$  GeV to satisfy trigger thresholds and  $|\eta| < 1.45$  to ensure the photon remains in the barrel of the detector.
- A selection criteria of  $\Delta R(\gamma, \text{jet}) < 1.0$ , between the photon and all jets is applied to ensure the acceptance of only well isolated  $\gamma + \text{jets}$  events.
- Given that the photon is ignored, this control sample can only be applied in the  $H_T$  region  $> 375$  GeV, due to the trigger thresholds on the minimum  $p_T$  of the photon, and the  $H_T$  requirement of an  $\alpha_T > 0.55$  cut from Equation (4.5). This

is maintained in this control sample due to contamination from QCD processes in the absence of an  $\alpha_T$  cut.



**Figure 4.8:** Data/MC comparisons of key variables for the  $\gamma + \text{jets}$  selection, following the application of selection criteria and the requirements that  $H_T > 375 \text{ GeV}$  and  $\alpha_T > 0.55$ . Bands represent the uncertainties due to the statistical size of the MC samples. No requirement is made upon the number of b-tagged jets or jet multiplicity in these distributions.

1621 The selection criteria of the three control samples are defined to ensure background  
1622 composition and event kinematics mirror closely the signal region. This is done in order  
1623 to minimise the reliance on simulation to model correctly the backgrounds and event  
1624 kinematics in the control and signal samples.

1625 However in the case of the  $\mu + \text{jets}$  and  $\mu\mu + \text{jets}$  samples, the  $\alpha_T$  requirement is relaxed  
1626 in the selection criteria of these samples. This is made possible as contamination from  
1627 QCD multi-jet events is suppressed to a negligible level by the other kinematic selection  
1628 criteria within the two control samples, selecting pure EWK processes. Thus in this way,  
1629 the acceptance of the two muon control samples can be significantly increased, which  
1630 simultaneously improves their statistical and predictive power and also dilutes the effect  
1631 of any potential signal contamination.

1632 The modelling of the  $\alpha_T$  variable is probed through a dedicated set of closure tests,  
1633 described in Section (4.6), which demonstrate that the different  $\alpha_T$  acceptances for  
1634 the control and signal samples have no significant systematic bias on the background  
1635 predictions.

#### 1636 4.2.4. Estimating the QCD Multi-jet Background

1637 A negligible background from QCD multi-jet events within the hadronic signal region is  
1638 expected due to a combination of selection requirements, and additional applied cleaning  
1639 filters. However a conservative approach is still adopted and the likelihood model, see  
1640 Section (4.8.2), is given the freedom to accommodate any potential QCD multi-jet  
1641 contamination.

1642 Any potential contamination can be identified through the variable  $R_{\alpha_T}$ , defined as the  
1643 ratio of events above and below the  $\alpha_T$  threshold value used in the analysis. This is  
1644 modelled by a  $H_T$  dependant falling exponential function which takes the form,

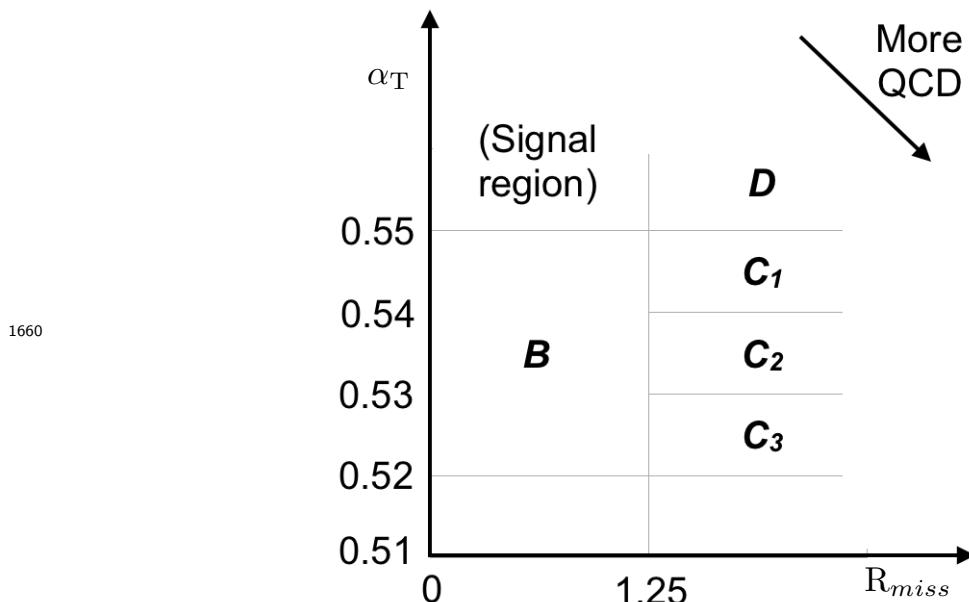
$$R_{\alpha_T}(H_T) = A_{\text{QCD}} \exp^{-k_{\text{QCD}} H_T}, \quad (4.10)$$

1645 where the parameters  $A_{\text{QCD}}$  and  $k_{\text{QCD}}$  are the normalisation and exponential decay  
1646 constants, respectively.

1647 For QCD multi-jet event topologies, this exponential behaviour as a function of  $H_T$  is  
1648 expected for several reasons. The improvement of jet energy resolution at higher  $H_T$   
1649 due to higher  $p_T$  jets leads to a narrower peaked  $\alpha_T$  distribution, causing  $R_{\alpha_T}$  to fall.  
1650 Similarly at higher  $H_T$  values  $> 375$  GeV, the jet multiplicity rises slowly with  $H_T$ . As  
1651 shown in Figure 4.3, at higher jet multiplicities the result of the combinatorics used in

1652 the determination of  $\alpha_T$  lead to more conservative  $\alpha_T$  values, also resulting in a narrower  
1653 distribution.

1654 The value of the decay constant  $k_{\text{QCD}}$  is constrained via measurements within data  
1655 sidebands to the signal region. This is also done to validate the falling exponential  
1656 assumption for QCD multi-jet topologies. The sidebands are enriched in QCD multi-jet  
1657 background and defined as regions where either  $\alpha_T$  is relaxed or that the  $R_{\text{miss}}$  cut is  
1658 inverted. Figure 4.9 depicts the definition of these data sidebands used to constrain the  
1659 value of  $k_{\text{QCD}}$ .



**Figure 4.9:** QCD sideband regions, used for determination of  $k_{\text{QCD}}$ .

1660  
1661 The fit results used to determine the value of  $k_{\text{QCD}}$  are shown in Appendix (C.1), for  
1662 which the best fit parameter value obtained from sideband region B is determined to be  
1663  $k_{\text{QCD}} = 2.96 \pm 0.64 \times 10^{-2} \text{ GeV}^{-1}$ .

1664 The best fit values of the remaining three C sideband regions are used to estimate  
1665 the systematic uncertainty on the central value obtained from sideband region B. The  
1666 variation of these measured values is used to determine the error on the determined  
1667 central value, and is calculated to be  $1.31 \pm 0.26 \times 10^{-2} \text{ GeV}^{-1}$ . This relative error of  $\sim$   
1668 20% gives an estimate of the systematic uncertainty of the measurement to be applied to  
1669  $k_{\text{QCD}}$ .

1670 Finally the same procedure is performed for sideband region D as an independent  
1671 crosscheck, to establish that the value of  $k_{\text{QCD}}$  extracted from a lower  $\alpha_T$  slice, can be

1672 applied to the signal region  $\alpha_T > 0.55$ . The likelihood fit is performed across all  $H_T$   
1673 bins within the QCD enriched region with no constraint applied to  $k_{QCD}$ . The resulting  
1674 best fit value for  $k_{QCD}$  shows good agreement between that and the weighted mean,  
1675 determined from the three C sideband regions. This demonstrates that the assumption  
1676 of using the central value determined from sideband region B, to provide an unbiased  
1677 estimator for  $k_{QCD}$  in the signal region ( $\alpha_T > 0.55$ ) is valid.

1678 Table 4.7 summarises the best fit  $k_{QCD}$  values determined for each of the sideband regions  
1679 to the signal region.

Sideband region	$k_{QCD} (\times 10^{-2} GeV^{-1})$	$p$ -value
B	$2.96 \pm 0.64$	0.24
C <sub>1</sub>	$1.19 \pm 0.45$	0.93
C <sub>2</sub>	$1.47 \pm 0.37$	0.42
C <sub>3</sub>	$1.17 \pm 0.55$	0.98
C(weighted mean)	$1.31 \pm 0.26$	-
D(likelihood fit)	$1.31 \pm 0.09$	0.57

**Table 4.7:** Best fit values for the parameters  $k_{QCD}$  obtained from sideband regions B,C<sub>1</sub>,C<sub>2</sub>,C<sub>3</sub>. The weighted mean is determined from the three measurements made within sideband region C. The maximum likelihood value of  $k_{QCD}$  given by the simultaneous fit using sideband region D. Quotes errors are statistical only.

### 1680 4.3. Trigger Strategy

1681 A cross trigger based on the  $H_T$  and  $\alpha_T$  values of an event, is used with varying thresholds  
1682 across  $H_T$  bins to record the events used in the hadronic signal region. The  $\alpha_T$  legs of  
1683 the HT\_alphaT triggers used in the analysis, are chosen to suppress QCD multi-jet events  
1684 and control trigger rate, whilst maintaining signal acceptance. To maintain an acceptable  
1685 rate for these analysis triggers, only calorimeter information is used in the reconstruction  
1686 of the  $H_T$  sum, leading to the necessity for Calo jets to be used within the analysis.

1687 A single object prescaled HT trigger is used to collect events for the hadronic control  
1688 region, described above in Section (4.2.4).

1689 The performance of the  $\alpha_T$  and  $H_T$  triggers used to collect data for the signal and  
1690 hadronic control region is measured with respect to a reference sample collected using  
1691 the muon system. This allows measurement of both the Level 1 seed and higher level

1692 triggers simultaneously, as the reference sample is collected independently of any jet  
1693 requirements.

1694 The selection for the trigger efficiency measurement is identical to that described in  
1695 Section (4.2.2), with the requirement of exactly one well identified muon with  $p_T > 30$   
1696 GeV. This muon is then subsequently ignored.

1697 The efficiencies measured for the HT\_alphaT triggers in each individual  $H_T$  and  $\alpha_T$  leg,  
1698 is summarised in Table 4.8 for each  $H_T$  category of the analysis.

$H_T$ range (GeV)	$\epsilon$ on $H_T$ leg (%)	$\epsilon$ on $\alpha_T$ leg (%)
275-325	$87.7^{+1.9}_{-1.9}$	$82.8^{+1.0}_{-1.1}$
325-375	$90.6^{+2.9}_{-2.9}$	$95.9^{+0.7}_{-0.9}$
375-475	$95.7^{+0.1}_{-0.1}$	$98.5^{+0.5}_{-0.9}$
475- $\infty$	$100.0^{+0.0}_{-0.0}$	$100.0^{+0.0}_{-4.8}$

**Table 4.8:** Measured efficiencies of the  $H_T$  and  $\alpha_T$  legs of the HT and HT\_alphaT triggers in independent analysis bins. The product of the two legs gives the total efficiency of the trigger in a given offline  $H_T$  bin.

1699 Data for the control samples of the analysis, detailed in Section (4.2.3), are collected  
1700 using a single object photon trigger for the  $\gamma +$  jets sample, and a single object muon  
1701 trigger for both the  $\mu +$  jets and  $\mu\mu +$  jets control samples.

1702 The photon trigger is measured to be fully efficient for the threshold  $p_T^{\text{photon}} > 150$  GeV,  
1703 whilst the single muon efficiency satisfying  $p_T^{\text{muon}} > 30$  GeV is measured to have an  
1704 efficiency of  $(88 \pm 2)\%$  that is independent of  $H_T$ . In the case of the  $\mu\mu +$  jets control  
1705 sample, the efficiency is measured to be  $(95 \pm 2)\%$  for the lowest  $H_T$  bin, rising (due to  
1706 the average  $p_T$  of the second muon in the event increasing at larger  $H_T$ ) to  $(98 \pm 2)\%$  in  
1707 the highest  $H_T$  category.

## 1708 4.4. Measuring Standard Model Process 1709 Normalisation Factors via $H_T$ Sidebands

1710 The theoretical cross-sections of different SM processes at Next to Next Leading Order  
1711 (NNLO) and the number of available simulated events generated for a particular process,  
1712 is typically used to determine the appropriate normalisation for a simulation sample.  
1713 However within the particular high- $H_T$  and high- $\mathcal{E}_T$  corners of kinematic phase space

1714 probed within this search, the theoretical cross sections for various processes are far less  
1715 well understood.

1716 To mitigate the problem of theoretical uncertainties and arbitrary choices of cross  
1717 sections, the normalisation of the simulation samples are determined through the use of  
1718 data sidebands. The sidebands are used to calculate sample specific correction factors  
1719 (k-factors), that are appropriate for the  $H_T$ - $\cancel{E}_T$  phase space covered by this analysis.

1720 They are defined within the  $\mu +$  jets and  $\mu\mu +$  jets control sample, by the region  $200 <$   
1721  $H_T < 275$ , using the same jet  $p_T$  thresholds as the adjacent first analysis bin. Individual  
1722 EWK processes are isolated within each of these control samples via requirements on jet  
1723 multiplicity and the requirement on b-tag multiplicity, summarised in Table 4.9. The  
1724 purity of the samples are typically  $> 90\%$  with any residual contamination subtracted  
1725 prior to determination of the correction factors. The resultant k-factor for each process  
1726 is determined by then taking ratio of the data yield over the expectation from simulation  
1727 in the sideband. Subsequently these k-factors are then applied to the processes within  
1728 the phase space of the analysis.

Process	Selection	Observation	MC expectation	k-factor
W + jets	$\mu +$ jets, $n_b=0$ , $n_{jet} = 2,3$	26950	$29993.2 \pm 650.1$	$0.90 \pm 0.02$
$Z \rightarrow \mu\mu +$ jets	$\mu\mu +$ jets, $n_b=0$ , $n_{jet} = 2,3$	3141	$3402.0 \pm 43.9$	$0.92 \pm 0.02$
$t\bar{t}$	$\mu +$ jets, $n_b=2$ , $n_{jet} = \geq 4$	2190	$1967.8 \pm 25.1$	$1.11 \pm 0.02$

**Table 4.9:** k-factors calculated for different EWK processes. All k-factors are derived relative to theoretical cross-sections calculated in NNLO. The k-factors measured for the  $Z \rightarrow \mu\mu +$  jets processes, are also applied to the  $Z \rightarrow \nu\bar{\nu} +$  jets and  $\gamma +$  jets simulation samples.

1729 It is worth pointing out that these correction factors have a negligible effect when  
1730 providing a background estimation for the signal region. The TFs used in the analysis  
1731 are found to be unaffected by application of these k-factors due to the similarity in the  
1732 background composition of the control and signal regions. However when systematic  
1733 uncertainties are determined in Section (4.6), the closure tests performed are sensitive  
1734 to these corrections when extrapolations between different  $n_b^{reco}$  and  $n_{jet}$  categories are  
1735 performed.

---

## <sup>1736</sup> 4.5. Determining Monte Carlo Simulation Yields <sup>1737</sup> with Higher Statistical Precision

<sup>1738</sup> Reconstructing events from **EWK** processes with many b-tagged jets ( $\geq 3$ ),  $n_b^{\text{reco}}$ , is largely  
<sup>1739</sup> driven by the mis-tagging of light jets within the event. This is clear when considering  
<sup>1740</sup> the main **EWK** backgrounds in the analysis, such as  $t\bar{t} + \text{jets}$  events, which typically  
<sup>1741</sup> contain two b-flavoured jets from the decay of the top quarks, whilst  $W + \text{jets}$  and  
<sup>1742</sup>  $Z \rightarrow \mu\mu + \text{jets}$  events will typically contain no b-flavoured jets.

<sup>1743</sup> When the expectation for the number of  $n_b^{\text{reco}}$  jets is taken directly from simulation, the  
<sup>1744</sup> statistical uncertainty at large reconstructed b-tagged jet multiplicities becomes relatively  
<sup>1745</sup> large. In order to reduce this uncertainty one approach is to use the information encoded  
<sup>1746</sup> throughout all events in the simulation sample, to measure each of the following four  
<sup>1747</sup> ingredients:

- <sup>1748</sup> 1. the b-tagging efficiency in the event selection,
- <sup>1749</sup> 2. the charm-tagging efficiency in the event selection,
- <sup>1750</sup> 3. the mis-tagging rate in the event selection,
- <sup>1751</sup> 4. the underlying flavour distribution of the jets in the events.

<sup>1752</sup> Together they can be used to determine the  $n_b^{\text{reco}}$  distribution of the process being  
<sup>1753</sup> measured. This method allows the determination of higher b-tag multiplicities to a higher  
<sup>1754</sup> degree of accuracy, reducing the statistical uncertainties of the simulation yields which  
<sup>1755</sup> enter into the **TF**'s. For the discussion that follows, these predictions are determined on  
<sup>1756</sup> average (i.e not on an event-by-event basis), and will be known as the formula method.

### <sup>1757</sup> 4.5.1. The Formula Method

<sup>1758</sup> The assigning of jet flavours to reconstruction level jets in simulation is achieved via an  
<sup>1759</sup> algorithmic method defined as:

- <sup>1760</sup> • attempt to find the parton that most likely determines the properties of the jet and  
<sup>1761</sup> assign that flavour as the true flavour,
- <sup>1762</sup> • “final state” partons (after showering, radiation) are analysed (also within  $\Delta R <$   
<sup>1763</sup> 0.3 of reconstructed jet cone),

- 1764     • if there is a b/c flavoured parton within the jet cone: label the jet as a b/c flavoured  
1765       jet,
- 1766     • otherwise: assign flavour of the hardest parton within the jet cone.

1767   This process is employed within each individual simulation sample and independently for  
1768   each  $H_T$ -  $n_{\text{jet}}$  category in the analysis. The  $n_b^{\text{reco}}$  distribution can then be constructed in  
1769   an analytical way using the formula:

$$N(n) = \sum_{n_b^{\text{gen}} + n_c^{\text{gen}} + n_q^{\text{gen}} = n_{\text{jet}}} \sum_{n_b^{\text{tag}} + n_c^{\text{tag}} + n_q^{\text{tag}} = n} N(n_b^{\text{gen}}, n_c^{\text{gen}}, n_q^{\text{gen}}) \times P(n_b^{\text{tag}}, n_b^{\text{gen}}, \epsilon) \times \\ P(n_c^{\text{tag}}, n_c^{\text{gen}}, \beta) \times P(n_q^{\text{tag}}, n_q^{\text{gen}}, m), \quad (4.11)$$

1770   with  $N(n)$  representing the number of  $n$  b-tagged jets of a simulated **SM** process in a  
1771   particular  $H_T$ -  $n_{\text{jet}}$  category.

1772   Let  $N(n_b^{\text{gen}}, n_c^{\text{gen}}, n_q^{\text{gen}})$  represent the 3-dimensional underlying jet flavour distribution in  
1773   simulation, with  $b$  underlying b-quarks,  $c$  underlying c-quarks and  $q$  underlying light  
1774   quarks which are matched to reconstructed jets as detailed above. Light quarks defined  
1775   as those which originate from a  $u$ ,  $d$ ,  $s$ ,  $g$  and  $\tau$  jets, which having similar mis-tagging  
1776   rates are grouped together.

1777   The variables  $n_{b/c/q}^{\text{tag}}$  signify the number of times that a particular jet flavour results in  
1778   a tagged jet, of which the sum of the three terms must equal the number of  $n$  tagged  
1779   jets being calculated. Similarly  $n_{b/c/q}^{\text{gen}}$  represent the flavour admixture of the jets, which  
1780   having been identified using the above technique as  $b$ ,  $c$  or light flavoured jets, are  
1781   required by definition that the sum of the three to fall within the  $n_{\text{jet}}$  category being  
1782   analysed.

1783   Finally  $P(n_b^{\text{tag}}, n_b^{\text{gen}}, \epsilon)$ ,  $P(n_c^{\text{tag}}, n_c^{\text{gen}}, \beta)$  and  $P(n_q^{\text{tag}}, n_q^{\text{gen}}, m)$  correspond to the binomial  
1784   probabilities for the tagging of jet with a particular flavour to occur, based on its measured  
1785   tagging efficiency ( $\epsilon$ ,  $\beta$  or  $m$ ). This formula is applied to calculate the resultant  $n_b^{\text{reco}}$   
1786   distribution, within a  $H_T$  category for each process in turn.

1787   This approach ultimately results in a more precise  $n_b^{\text{reco}}$  distribution prediction, due to the  
1788   utilisation of all events in simulation sample which pass selection in extracting the overall

1789 underlying  $n_b^{reco}$  distribution. This is particularly useful at higher  $n_b^{reco}$  multiplicities  
1790 where a lack of events in simulation can lead to relatively large statistical uncertainties.

### 1791 4.5.2. Establishing Proof of Principle

1792 In order to validate the procedure, the predictions determined from the formula method  
1793 summarised in Equation (4.11), are compared directly with those obtained directly from  
1794 simulation. Resultantly no simulation to data correction factors are applied when making  
1795 this comparison

1796 This sanity check for the  $\mu +$  jets control sample is presented in Table 4.10, for all  $n_b^{reco}$   
1797 and  $H_T$  categories with no requirement placed upon the jet multiplicity of the events.

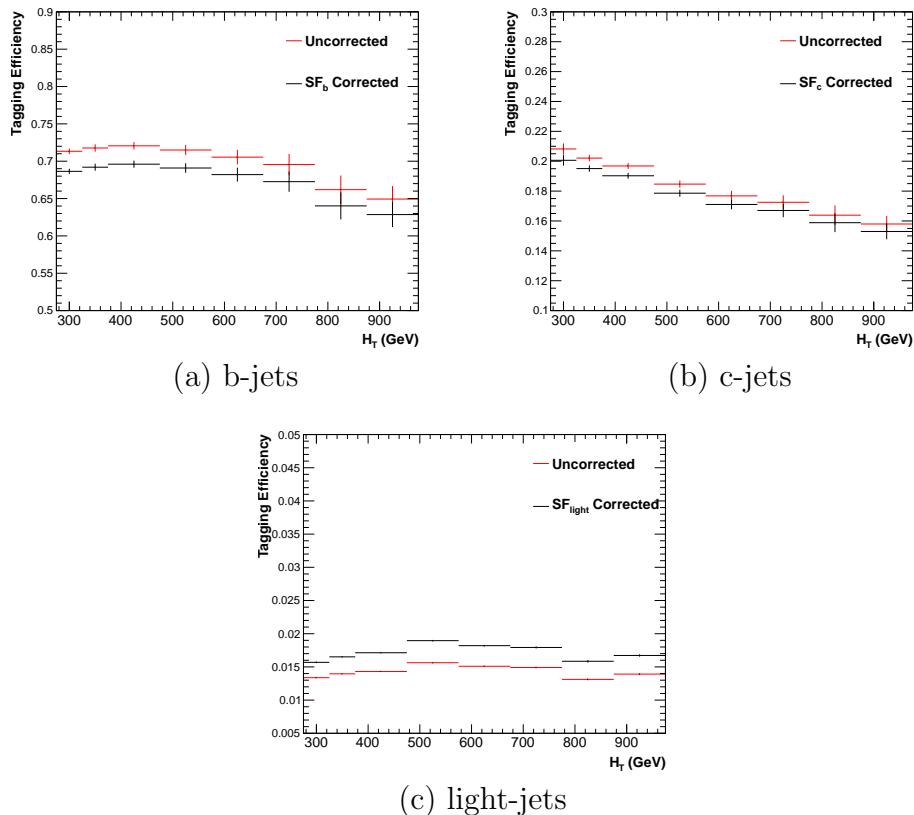
$H_T$ Bin (GeV)	275–325	325–375	375–475	475–575
Formula $n_b = 0$	$12632.66 \pm 195.48$	$6696.08 \pm 82.59$	$6368.96 \pm 75.34$	$2906.27 \pm 39.65$
Vanilla $n_b = 0$	$12612.95 \pm 198.68$	$6687.97 \pm 83.78$	$6359.27 \pm 76.50$	$2898.27 \pm 36.89$
Formula $n_b = 1$	$4068.09 \pm 45.71$	$2272.76 \pm 26.14$	$2181.32 \pm 25.07$	$1089.14 \pm 13.82$
Vanilla $n_b = 1$	$4067.73 \pm 60.30$	$2268.02 \pm 30.20$	$2180.69 \pm 28.73$	$1094.37 \pm 24.14$
Formula $n_b = 2$	$1963.71 \pm 22.44$	$1087.55 \pm 13.57$	$1055.57 \pm 13.25$	$554.96 \pm 7.95$
Vanilla $n_b = 2$	$1984.53 \pm 26.19$	$1094.43 \pm 16.67$	$1068.96 \pm 16.36$	$558.14 \pm 10.51$
Formula $n_b = 3$	$146.94 \pm 2.07$	$79.97 \pm 1.37$	$78.05 \pm 1.35$	$49.84 \pm 1.03$
Vanilla $n_b = 3$	$149.52 \pm 4.84$	$85.98 \pm 3.64$	$74.45 \pm 3.29$	$49.54 \pm 2.68$
Formula $n_b \geq 4$	$2.26 \pm 0.12$	$1.29 \pm 0.10$	$5.32 \pm 0.20$	-
Vanilla $n_b \geq 4$	$1.84 \pm 0.50$	$1.02 \pm 0.39$	$4.86 \pm 0.83$	-
$H_T$ Bin (GeV)	575–675	675–775	775–875	>875
Formula $n_b = 0$	$1315.68 \pm 19.49$	$640.49 \pm 11.90$	$327.81 \pm 7.91$	$424.27 \pm 9.27$
Vanilla $n_b = 0$	$1315.23 \pm 20.20$	$641.96 \pm 12.48$	$329.09 \pm 8.36$	$424.02 \pm 9.73$
Formula $n_b = 1$	$490.41 \pm 7.45$	$226.95 \pm 4.42$	$109.91 \pm 2.84$	$129.97 \pm 3.07$
Vanilla $n_b = 1$	$490.52 \pm 9.92$	$222.22 \pm 6.21$	$107.46 \pm 4.15$	$129.64 \pm 4.64$
Formula $n_b = 2$	$256.75 \pm 4.58$	$113.45 \pm 2.70$	$52.10 \pm 1.69$	$59.29 \pm 1.78$
Vanilla $n_b = 2$	$253.43 \pm 6.52$	$117.17 \pm 4.27$	$52.70 \pm 2.80$	$59.45 \pm 3.00$
Formula $n_b = 3$	$25.66 \pm 0.69$	$12.48 \pm 0.46$	$5.52 \pm 0.31$	$6.83 \pm 0.33$
Vanilla $n_b = 3$	$29.18 \pm 2.06$	$11.77 \pm 1.26$	$6.18 \pm 0.95$	$7.53 \pm 1.05$

**Table 4.10:** Comparing yields in simulation within the  $\mu +$  jets selection determined from the formula method described in Equation (4.11), and that taken directly from simulation. The numbers are normalised to  $11.4\text{fb}^{-1}$ . No simulation to data corrections are applied.

1798 It can be seen as expected, that there is good consistency between the results determined  
 1799 via the formula method and ‘raw’ simulation yields. Similarly the power of this approach  
 1800 can be seen in the reduction of this statistical error in the prediction across all  $H_T$  and  
 1801  $n_b^{reco}$  categories. In particular the statistical uncertainty is reduced by several factors in  
 1802 the highest  $n_b^{reco} \geq 4$  category.

### 1803 4.5.3. Correcting Measured Efficiencies in Simulation to Data

1804 As detailed in Section (3.3.2), it is necessary for certain  $p_T$  and  $\eta$  dependant corrections,  
 1805 to be applied to both the b-tagging efficiency and mis-tagging rates in order to correct  
 1806 the efficiencies from simulation to the efficiencies measured in data. These correction  
 1807 factors are considered when determining the simulation yields for each selection, which  
 1808 are used to construct the TFs of the analysis. The magnitude of this correction are  
 1809 measured individually for each  $H_T$  category and are shown in Figure 4.10.



**Figure 4.10:** Tagging efficiencies of (a) b-jets, (b) c-jets, and (c) light-jets as a function all jets within each  $H_T$  category. Efficiencies measured directly from simulation (black) and with data to simulation  $SF_{b,c,light}$  correction factors (red) are applied.

1810 Each of the correction factors for the b, c and light flavoured jets come with an associated  
 1811 systematic uncertainty. The uncertainties across different jet  $p_T$  and  $\eta$  categories, are  
 1812 considered as fully correlated. When computing the magnitude of the effect of this  
 1813 systematic uncertainty on the TFs of the analysis, the measured tagging efficiencies for  
 1814 each jet flavour are scaled up/down simultaneously within each  $H_T$  and  $n_{\text{jet}}$  category by  
 1815 the systematic uncertainty of the  $SF_{b, c, \text{light}}$  scale factors.

1816 Varying the scale factor corrections by their systematic uncertainty will change the  
 1817 absolute yields within each  $n_b^{reco}$  bin of all selections. However, ultimately it is the change  
 1818 in the TFs which influences the final background prediction from each of the control  
 1819 samples. The magnitude of the absolute change in each TF, constructed from when the  
 1820  $\mu + \text{jets}$  control sample is used to predict the entire hadronic signal region background,  
 1821 is shown in Table 4.11.

$n_b^{reco}$	275–325	325–375	375–475	475–575
= 0	$0.557^{+0.001}_{-0.001} \pm 0.012$	$0.495^{+0.001}_{-0.001} \pm 0.009$	$0.383^{+0.001}_{-0.001} \pm 0.005$	$0.307^{+0.001}_{-0.002} \pm 0.006$
= 1	$0.374^{+0.006}_{-0.006} \pm 0.006$	$0.320^{+0.006}_{-0.005} \pm 0.005$	$0.251^{+0.005}_{-0.005} \pm 0.004$	$0.185^{+0.003}_{-0.003} \pm 0.004$
= 2	$0.226^{+0.002}_{-0.002} \pm 0.004$	$0.201^{+0.001}_{-0.002} \pm 0.004$	$0.159^{+0.001}_{-0.001} \pm 0.004$	$0.134^{+0.000}_{-0.001} \pm 0.004$
= 3	$0.221^{+0.002}_{-0.002} \pm 0.005$	$0.208^{+0.002}_{-0.001} \pm 0.007$	$0.164^{+0.001}_{-0.000} \pm 0.006$	$0.144^{+0.001}_{-0.001} \pm 0.007$
$\geq 4$	$0.222^{+0.004}_{-0.005} \pm 0.015$	$0.248^{+0.003}_{-0.003} \pm 0.035$	$0.123^{+0.002}_{-0.003} \pm 0.009$	-
	575–675	675–775	775–875	$\geq 875$
= 0	$0.263^{+0.001}_{-0.002} \pm 0.006$	$0.215^{+0.000}_{-0.001} \pm 0.007$	$0.171^{+0.000}_{-0.001} \pm 0.009$	$0.111^{+0.000}_{-0.001} \pm 0.006$
= 1	$0.154^{+0.003}_{-0.003} \pm 0.005$	$0.138^{+0.003}_{-0.004} \pm 0.006$	$0.121^{+0.005}_{-0.005} \pm 0.007$	$0.091^{+0.002}_{-0.002} \pm 0.006$
= 2	$0.104^{+0.000}_{-0.001} \pm 0.005$	$0.079^{+0.001}_{-0.001} \pm 0.006$	$0.063^{+0.001}_{-0.002} \pm 0.007$	$0.071^{+0.000}_{-0.000} \pm 0.008$
= 3	$0.116^{+0.001}_{-0.001} \pm 0.009$	$0.069^{+0.001}_{-0.001} \pm 0.007$	$0.079^{+0.001}_{-0.001} \pm 0.017$	$0.095^{+0.003}_{-0.002} \pm 0.020$

**Table 4.11:** The absolute change in the TFs used to predict the entire signal region SM background, using the  $\mu + \text{jets}$  control sample when the systematic uncertainties of the data to simulation scale factors are varied by  $\pm 1\sigma$ . The impact of the change is shown for each  $H_T$  and  $n_b^{reco}$  category with no requirement made on the jet multiplicity of the events. (Also quoted are the statistical uncertainties)

1822 It can be seen that the TFs are found to be relatively insensitive to the systematic  
 1823 uncertainty of the b-tag scale factors (showing typically less than  $\sim 2\%$  change). This  
 1824 can be accounted for by the similar composition of the signal and control sample  
 1825 backgrounds, such that any change in the underlying  $n_b^{reco}$  distribution will be reflected  
 1826 in both signal and control regions and cancel out in the TF.

1827 Any overall systematic effect on the overall background prediction of the analysis from  
1828 these b-tag scale factor uncertainties is incorporated within the data driven systematics  
1829 introduced in the following section.

## 1830 4.6. Systematic Uncertainties on Transfer Factors

1831 Since the TFs used to establish the background prediction are obtained from simulation,  
1832 an appropriate systematic uncertainty is assigned to account for theoretical uncertainties  
1833 [83] and limitations in the simulation modelling of event kinematics and instrumental  
1834 effects.

1835 The magnitudes of these systematic uncertainties are established through a data driven  
1836 method, in which the three independent control samples of the analysis ( $\mu + \text{jets}$ ,  $\mu\mu$   
1837 + jets,  $\gamma + \text{jets}$ ) are used to in a series of closure tests. The yields from one of these  
1838 control samples, along with the corresponding TF obtained from simulation, are used to  
1839 predict the expected yields in another control sample. This procedure therefore utilised  
1840 the same method used in determining a background prediction for the signal region as  
1841 already established in Section (4.2.3).

1842 The level of agreement between the predicted and observed yields is expressed as the  
1843 ratio

$$\frac{(N_{\text{obs}} - N_{\text{pred}})}{N_{\text{pred}}}, \quad (4.12)$$

1844 while considering only the statistical uncertainties on the prediction,  $N_{\text{pred}}$ , and the  
1845 observation,  $N_{\text{obs}}$ . No systematic uncertainty is assigned to the prediction, and resultantly  
1846 the level of closure is defined by the statistical significance of a deviation from the ratio  
1847 from zero.

1848 This ratio is measured for each  $H_T$  category in the analysis, allowing these closure tests to  
1849 be sensitive to both the presence of any significant biases or any possible  $H_T$  dependence  
1850 to the level of closure.

1851 Eight sets of closure tests are defined between the three data control samples, conducted  
1852 independently between the two jet multiplicity ( $2 \leq n_{\text{jet}} \leq 3$ ,  $n_{\text{jet}} \geq 4$ ) categories. Each

1853 of these tests are specifically chosen to probe each of the different key ingredients of the  
1854 simulation modelling that can affect the background prediction.

1855 Each of the different modelling components and the relevant closure tests are described  
1856 below:

1857  **$\alpha_T$  modelling**

1858 The modelling of the  $\alpha_T$  distribution in genuine  $E_T$  events is probed with the  $\mu$   
1859 + jets control sample. This test is important to verify the approach of removing  
1860 the  $\alpha_T > 0.55$  requirement from the  $\mu +$  jets and  $\mu\mu +$  jets samples to increase  
1861 the precision of the background prediction. The test uses the  $\mu +$  jets sample  
1862 without an  $\alpha_T$  cut to make a prediction into the  $\mu +$  jets sample defined with the  
1863 requirement  $\alpha_T > 0.55$ .

1864 **Background admixture**

1865 The sensitivity of the translation factors to the relative admixture of events from  
1866  $W +$  jets and  $t\bar{t}$  processes is probed by two closure tests.

1867 Within the  $\mu +$  jets sample, a  $W$  boson enriched sub-sample ( $n_b = 0$ ) is used  
1868 to predict yields in a  $t\bar{t}$  enriched sub-sample ( $n_b = 1$ ). Similarly, the  $t\bar{t}$  enriched  
1869 sub-sample ( $n_b = 1$ ) is also used to predict yields for a further enriched  $t\bar{t}$  sub-sample  
1870 ( $n_b = 2$ ), further probing the modelling of the  $n_b^{\text{reco}}$  distribution.

1871 Similarly a further closure test probes the relative contribution of  $Z +$  jets to  $W +$  jets  
1872 and  $t\bar{t}$  events, through the use of the  $\mu +$  jets sample to predict yields for the  $\mu\mu +$   
1873 jets control sample. This closure test, also at some level probes the muon trigger  
1874 and reconstruction efficiencies, given that exactly one or two muons are required by  
1875 the different selections.

1876 These tests represent an extremely conservative approach as the admixture of the  
1877 two backgrounds remains similar when a prediction is made between the control  
1878 samples and the signal region. This is contrary to the closure tests defined above  
1879 which make predictions between two very different admixtures of  $W +$  jets and  $t\bar{t}$   
1880 events.

1881 **Consistency check between  $Z \rightarrow \nu\bar{\nu}$  predictions**

1882 This is an important consistency check between the  $\mu\mu +$  jets and  $\gamma +$  jets, which  
1883 are both used in the prediction of the  $Z \rightarrow \nu\bar{\nu}$  in the signal region. This is conducted

1884 by using the  $\gamma + \text{jets}$  sample to predict yields for the  $\mu\mu + \text{jets}$  control sample.  
 1885 Using  $\gamma + \text{jets}$  processes as a method to predict  $Z + \text{jet}$  processes is subject to  
 1886 theory uncertainties [84], which can be probed by this data driven closure test within  
 1887 a  $Z \rightarrow \mu\mu$  control sample.

### 1888 Modelling of jet multiplicity

1889 The simulation modelling of the jet multiplicity within each control sample is  
 1890 important due to the exclusive jet multiplicity categorisation within the analysis.  
 1891 This is probed via the use of each of the three control samples to independently  
 1892 predict from the lower jet multiplicity category  $2 \leq n_{\text{jet}} \leq 3$ , to the high jet category  
 1893  $n_{\text{jet}} \geq 4$ .

1894 For the case of the  $\mu + \text{jets}$  and  $\mu\mu + \text{jets}$  control samples, this test also serves as a  
 1895 further probe of the admixture between  $W + \text{jets}/Z + \text{jets}$  and  $t\bar{t}$ .

1896 To test for the assumption that no  $H_T$  dependencies exist within the background pre-  
 1897 dictions of the analysis, the first five closure tests defined above are used, with zeroeth  
 1898 and first order polynomial fits applied to each test individually. This is summarised in  
 1899 Table 4.12 and Table 4.13 which show the results for both the  $2 \leq n_{\text{jet}} \leq 3$  and  $\geq 4$  jet  
 1900 multiplicity bins respectively.

Closure test	Symbol	Constant fit		Linear fit	
		Best fit value	p-value	Slope ( $10^{-4}$ )	p-value
$\alpha_T < 0.55 \rightarrow \alpha_T > 0.55 (\mu + \text{jets})$	Circle	$-0.06 \pm 0.02$	0.93	$-1.3 \pm 2.2$	0.91
0 b-jets $\rightarrow$ 1 b-jet ( $\mu + \text{jets}$ )	Square	$0.07 \pm 0.02$	0.98	$-1.6 \pm 1.6$	1.00
1 b-jets $\rightarrow$ 2 b-jet ( $\mu + \text{jets}$ )	Triangle	$-0.07 \pm 0.03$	0.76	$-2.7 \pm 3.0$	0.76
$\mu + \text{jets} \rightarrow \mu\mu + \text{jets}$	Cross	$0.10 \pm 0.03$	0.58	$-1.1 \pm 2.3$	0.49
$\mu\mu + \text{jets} \rightarrow \gamma + \text{jets}$	Star	$-0.06 \pm 0.04$	0.31	$4.2 \pm 4.3$	0.29

1901 **Table 4.12:** A summary of the results obtained from zeroeth order polynomial (i.e. a constant)  
 1902 and linear fits to five sets of closure tests performed in the  $2 \geq n_{\text{jet}} \geq 3$  category.  
 1903 The two columns show the best fit value for the slope obtained when performing  
 1904 a constant (left) and linear (right) fit and the p-value for that fit.

1905 Table 4.14 shows the same fits applied to the three closure tests that probe the modelling  
 1906 between the two  $n_{\text{jet}}$  categories. The best fit value and its uncertainty is listed for each  
 1907 set of closure tests in all three tables, along with the p-value of the constant and linear  
 1908 fits applied.

1909 The best fit value for the constant parameter is indicative of the level of closure, averaged  
 1910 across the full  $H_T$  range of the analysis, and the p-value an indicator of any significant  
 1911 dependence on  $H_T$  within the closure tests. The best fit values of all the tests are either

Closure test	Symbol	Constant fit		Linear fit	
		Best fit value	p-value	Slope ( $10^{-4}$ )	p-value
$\alpha_T < 0.55 \rightarrow \alpha_T > 0.55 (\mu + \text{jets})$	Circle	$-0.05 \pm 0.03$	0.21	$3.0 \pm 2.9$	0.21
$0 \text{ b-jets} \rightarrow 1 \text{ b-jet} (\mu + \text{jets})$	Square	$-0.03 \pm 0.03$	0.55	$-1.0 \pm 1.9$	0.47
$1 \text{ b-jets} \rightarrow 2 \text{ b-jet} (\mu + \text{jets})$	Triangle	$-0.02 \pm 0.03$	0.39	$1.1 \pm 2.2$	0.31
$\mu + \text{jets} \rightarrow \mu\mu + \text{jets}$	Cross	$0.08 \pm 0.07$	0.08	$4.8 \pm 4.3$	0.07
$\mu\mu + \text{jets} \rightarrow \gamma + \text{jets}$	Star	$-0.03 \pm 0.10$	0.72	$-4.0 \pm 7.0$	0.64

**Table 4.13:** A summary of the results obtained from zeroeth order polynomial (i.e. a constant) and linear fits to five sets of closure tests performed in the  $n_{\text{jet}} \geq 4$  category. The two columns show the best fit value for the slope obtained when performing a constant (left) and linear (right) fit and the p-value for that fit.

Closure test	Symbol	Constant fit		Linear fit	
		Best fit value	p-value	Slope ( $10^{-4}$ )	p-value
$\mu + \text{jets}$	Inverted triangle	$-0.03 \pm 0.02$	0.02	$0.0 \pm 1.0$	0.01
$\mu + \text{jets}$ (outlier removed)	Inverted triangle	$-0.04 \pm 0.01$	0.42	$-1.4 \pm 1.1$	0.49
$\gamma + \text{jets}$	Diamond	$0.12 \pm 0.05$	0.79	$6.0 \pm 4.7$	0.94
$\mu\mu + \text{jets}$	Asterisk	$-0.04 \pm 0.07$	0.20	$4.9 \pm 4.4$	0.20

**Table 4.14:** A summary of the results obtained from zeroeth order polynomial (i.e. a constant) and linear fits to three sets of closure tests performed between the  $2 \leq n_{\text{jet}} \leq 3$  and  $n_{\text{jet}} \geq 4$  categories. The two columns show the best fit value for the slope obtained when performing a constant (left) and linear (right) fit and the p-value for that fit.

1908 statistically compatible with zero bias (i.e. less than  $2\sigma$  from zero) or at the level of 10%  
1909 or less, with the exception of one closure test discussed below.

1910 Within Table 4.14, there exists one test that does not satisfy the above statement, which  
1911 is the  $2 \leq n_{\text{jet}} \leq 3 \rightarrow n_{\text{jet}} \geq 4$  test using the  $\mu + \text{jets}$  control sample. The low p-value  
1912 can be largely attributed to an outlier between  $675 < H_T < 775$  GeV, rather than any  
1913 significant trend in  $H_T$ . Removing this single outlier from the constant fit performed,  
1914 gives a best fit value of  $-0.04 \pm 0.01$ ,  $\chi^2/\text{d.o.f} = 6.07/6$ . and a p-value of 0.42. These  
1915 modified fit results are also included in Table 4.14.

1916 Additionally, it is found that the best fit values for the slope terms of the linear fits in  
1917 all three tables are of the order  $10^{-4}$ , which corresponds to a percent level change per  
1918 100 GeV. However in all cases, the best fit values are fully compatible with zero (within  
1919  $1\sigma$ ) once again with the exception detailed above, indicating that the level of closure is  
1920 indeed  $H_T$  independent.

**4.6.1. Determining Systematic Uncertainties from Closure****Tests**

Once it has been established that no significant bias or trend exists within the closure tests, systematic uncertainties are determined. The statistical precision of the closure tests is considered a suitable benchmark for determining the systematic uncertainties that are assigned to the TFs, which are propagated through to the likelihood fit.

The systematic uncertainty band is split into five separate regions of  $H_T$ . Within each region the square root of the sample variance,  $\sigma^2$ , is taken over the eight closure tests to determine the systematic uncertainties to be applied within that region.

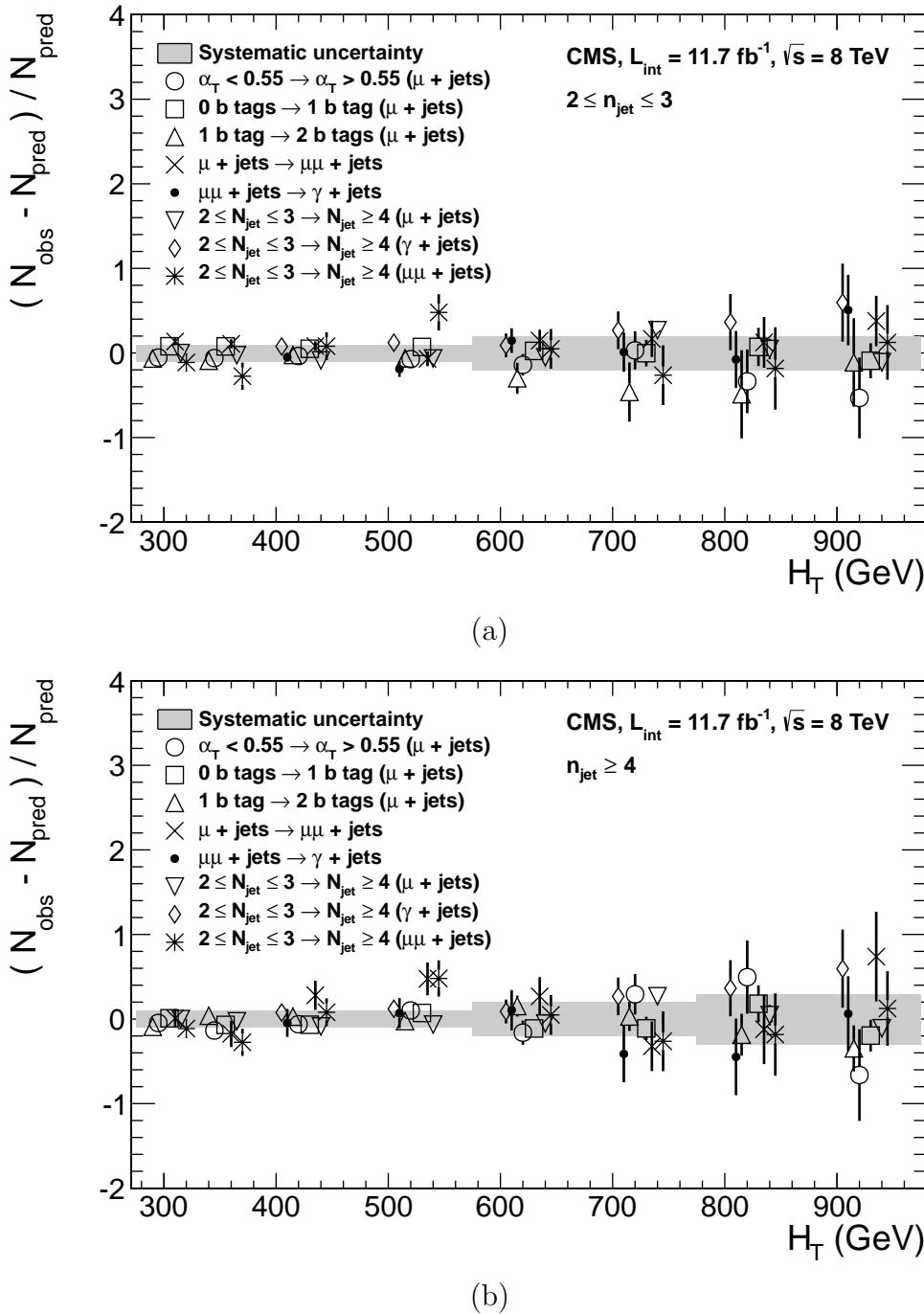
Using this procedure the systematic uncertainties for each region are calculated and are shown in Table 4.15, with the systematic uncertainty to be used in the likelihood model conservatively rounded up to the nearest decile and applied across all  $n_b^{\text{reco}}$  categories.

$H_T$ band (GeV)	$2 \leq n_{\text{jet}} \leq 3$	$n_{\text{jet}} \geq 4$
$275 < H_T < 325$	10%	10%
$325 < H_T < 375$	10%	10%
$375 < H_T < 575$	10%	10%
$575 < H_T < 775$	20%	20%
$H_T > 775$	20%	30%

**Table 4.15:** Calculated systematic uncertainties for the five  $H_T$  regions, determined from the closure tests. Uncertainties shown for both jet multiplicity categories. Values used within the likelihood model are conservatively rounded up to the nearest decile.

Figure 4.11 shows the sets of closure tests overlaid on top of grey bands that represent the  $H_T$  dependent systematic uncertainties. These systematic uncertainties are assumed to be fully uncorrelated between the different  $n_b$  multiplicity categories and across the five  $H_T$  regions. This can be considered a more conservative approach given that some correlations between adjacent  $H_T$  categories could be expected due to comparable kinematics.

These closure tests represent a conservative estimate of the systematic uncertainty in making a background prediction for the signal region. This is due to significant differences in the background composition and event kinematics between the two sub-samples used in the closure tests. This is not the case when a signal region prediction is made, due to the two sub-samples both having a comparable background admixture and similar kinematics owing to the fact that the TFs are always constructed using the same ( $n_{\text{jet}}$ ,  $n_b^{\text{reco}}$ ,  $H_T$ ) category.



**Figure 4.11:** Sets of closure tests (open symbols) overlaid on top of the systematic uncertainty used for each of the five  $H_T$  regions (shaded bands) and for the two different jet multiplicity categories: (a)  $2 \leq n_{\text{jet}} \leq 3$  and (b)  $n_{\text{jet}} \geq 4$ .

<sup>1945</sup> This point is emphasised when we examine the sensitivity of the TFs to a change in the admixture of  $W + \text{jets}$  and  $t\bar{t}$  with the control and signal samples. This is accomplished by varying the cross-sections of the  $W + \text{jets}$  and  $t\bar{t}$  by +20% and -20%, respectively.

1948 Figures C.2 and C.3 within Appendix C, show the effect upon the closure tests for both  
1949 jet multiplicity categories. Given these variations in cross-sections, the level of closure is  
1950 found to be significantly worse, with biases as large as  $\sim 30\%$ , most apparent in the  
1951 lowest  $H_T$  bins. However, the TFs used to extrapolate from control to signal are seen to  
1952 change only at the percent level by this large change in cross-section, shown in Table C.1.  
1953 Given the robust behaviour of the translation factors with respect to large (and opposite)  
1954 variations in the  $W + \text{jets}$  and  $t\bar{t}$  cross-sections, one can assume with confidence that  
1955 any bias in the translation factors is adequately (and conservatively) covered by the  
1956 systematic uncertainties used in the analysis.

## 1957 4.7. Simplified Models, Efficiencies and Systematic 1958 Uncertainties

1959 The results of the analysis are interpreted using various SMS signal models, which as  
1960 already introduced in Section (2.4.1) offer a natural starting point for quantifying and  
1961 characterising SUSY signals, and a means to identify the boundaries of search sensitivity  
1962 for different mass splittings, kinematic ranges, and final states.  
1963 Each model is parameterised in a two dimensional parameter space,  $(m_{\tilde{q}/\tilde{g}}, m_{\text{LSP}})$ , from  
1964 which upper limits on the production cross-sections of the various SMS models can be  
1965 set.  
1966 Each signal sample is generated at Leading Order (LO) with Pythia [85], and cross-  
1967 sections calculated for Next to Leading Order (NLO) and Next to Leading Logarithmic  
1968 Order (NLL) [86], with events simulated using the Fastsim framework. This framework  
1969 represents a simplified simulation of the CMS detector, but allows for faster production  
1970 of various signal topologies with different mass parameters.  
1971 A series of correction factors are applied to account for differences between Fastsim [87]  
1972 and Fullsim [88] simulation, which can affect the resultant  $n_b^{\text{reco}}$  distribution and which  
1973 are detailed in Section (4.7.2).

<sup>1974</sup>

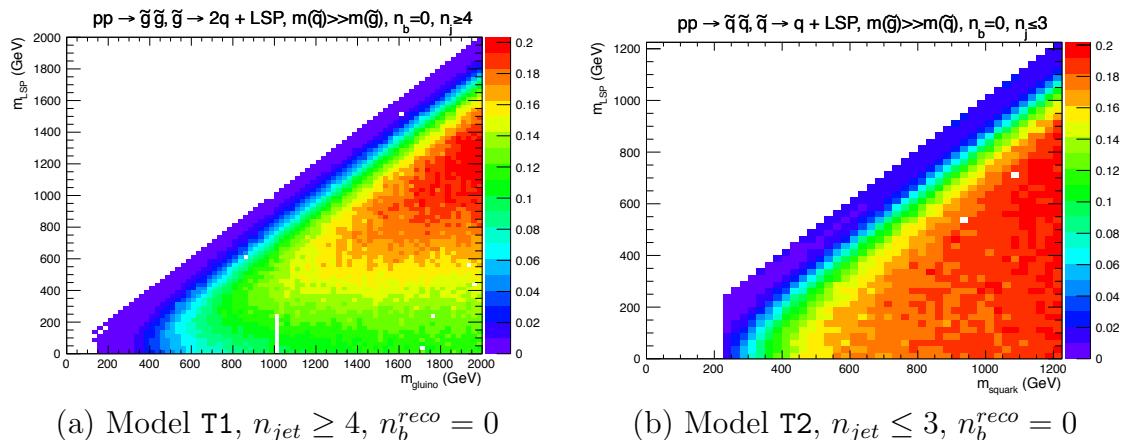
### 4.7.1. Signal Efficiency

<sup>1975</sup> The analysis selection efficiency,  $\epsilon$ , is measured for each mass point of the interpreted  
<sup>1976</sup> model. This serves as a measure of the sensitivity of the signal selection for that particular  
<sup>1977</sup> sparticle, LSP mass and final state topology. The signal yield is then given by

$$Y(m_{\tilde{q}/\tilde{g}}, m_{LSP}) = \epsilon \times \sigma \times \mathcal{L}, \quad (4.13)$$

<sup>1978</sup> where  $\sigma$  represents the model's cross-section and  $\mathcal{L}$  the luminosity. An upper limit on  $\sigma$   
<sup>1979</sup> taken from theory can then allow for the setting of limits in terms of the particle mass.

<sup>1980</sup> Figure 4.12 shows the expected signal efficiency of the signal selection for the T1 and  
<sup>1981</sup> T2 SMS models interpreted in this analysis. The efficiency maps are produced with the  
<sup>1982</sup> requirement  $H_T > 275$  GeV (i.e. no  $H_T$  categorisation) and requirements on  $n_{jet}$  and  
<sup>1983</sup>  $n_b^{reco}$  are the most sensitive to the model in question.



**Figure 4.12:** Signal efficiencies for the SMS models (a) T1 ( $\tilde{g}\tilde{g}^* \rightarrow q\bar{q}\tilde{\chi}_1^0 q\bar{q}\tilde{\chi}_1^0$ ) and (b) T2 ( $\tilde{q}\tilde{q}^* \rightarrow q\tilde{\chi}_1^0 q\tilde{\chi}_1^0$ ) when requiring  $n_{jet} \geq 4$  and  $\leq 3$  respectively, and  $n_b^{reco} = 0$ .

<sup>1984</sup> The same procedure is conducted in the analysis control samples. It is found in the  $\mu +$   
<sup>1985</sup> jets control samples, that the signal-to-background ratios for the expected signal yields in  
<sup>1986</sup> each of the SMS models are many time smaller than in the hadronic signal region. The  
<sup>1987</sup> relative contamination for the  $\mu\mu +$  jets sample is smaller still due to the requirement of  
<sup>1988</sup> a second muon. The relative contamination for the  $\gamma +$  jets sample is expected to be  
<sup>1989</sup> zero for the models under consideration. These small, relative levels of contamination  
<sup>1990</sup> are accounted for in the fitting procedure, as described in Section (4.8.4).

## 4.7.2. Applying B-tagging Scale Factor Corrections in Signal Samples

High-statistic **FastSim** signal simulation samples are unavailable for each signal point, which means that a different procedure to the formula method described in Section (4.5) is employed. Furthermore, the use of the **FastSim** framework in the reconstruction introduces an extra set of scale-factor corrections, to be applied simultaneously with those correcting **FullSim** to the data.

For these signal models, an event-by-event re-weighting procedure is applied. This applied weight depends on both the flavour content and the b-tagging status of the reconstruction level jets in the event.

The re-weighting procedure can be described by first considering a single jet within a signal event. The flavour of the jet is determined using the method described in Section (4.5.1).

Maps of the tagging efficiencies, parameterised as a function of jet  $p_T$  and  $\eta$  are produced from **FullSim** simulation samples for each of the b, c and light jet flavours. These efficiencies are calculated from simulation events which pass the hadronic signal selection. The  $p_T$  and  $\eta$  binning of each map is chosen to match the correction maps of **FullSim** to data defined in [77].

The actual tagging efficiency of the **FastSim** jet,  $\epsilon_{\text{FastSim}}(p_T, \eta, f)$ , differs from that measured in **FullSim**,  $\epsilon_{MC}(p_T, \eta, f)$ , as detailed above and is related via an additional correction factor,

$$\epsilon_{\text{FastSim}}(p_T, \eta, f) = \frac{\epsilon_{MC}(p_T, \eta, f)}{SF_{\text{Fast} \rightarrow \text{Full}}(p_T, \eta, f)}. \quad (4.14)$$

$SF_{\text{Fast} \rightarrow \text{Full}}(p_T, \eta, f)$  represents a set of  $p_T$  and  $\eta$  dependant corrections, that are specific for each **SMS** model. These corrections are calculated from the ratio of tagging rates between a **FullSim**  $t\bar{t}$  sample, and a selection of mass points from each **FastSim** **SMS** model, again measured individually for b, c and light-flavoured jets.

The tagging efficiencies measured in data [77],  $\epsilon_{Data}(p_T, \eta, f)$ , can then be related to  $\epsilon_{\text{FastSim}}(p_T, \eta, f)$  by the equation,

$$\begin{aligned}\epsilon_{Data}(p_T, \eta, f) &= \epsilon_{MC}(p_T, \eta, f) \times SF_{MC \rightarrow Data}(p_T, \eta, f) \\ &= \epsilon_{\text{FastSim}}(p_T, \eta, f) \times \underbrace{SF_{\text{Fast} \rightarrow \text{Full}}(p_T, \eta, f) \times SF_{MC \rightarrow Data}(p_T, \eta, f)}_{SF_{\text{Fast}} \rightarrow Data}.\end{aligned}\tag{4.15}$$

2018 For each jet, the weight of the event is re-weighted according to whether the jet fires the  
2019 tagger. In the instance that the jet *is* tagged, the event weight will be modified by,

$$\text{weight} = SF_{\text{Fast} \rightarrow Data} \times \text{weight},\tag{4.16}$$

2020 and in the case that the jet does *not* fire the tagger,

$$\text{weight} = \frac{1 - \epsilon_{Data}(p_T, \eta, f)}{1 - \epsilon_{\text{FastSim}}(p_T, \eta, f)} \times \text{weight}.\tag{4.17}$$

2021 This procedure is applied to all events that pass the selection criteria, thus correcting  
2022 the **FastSim**  $n_b^{\text{reco}}$  distribution to data.

### 2023 4.7.3. Experimental Uncertainties

2024 The systematic uncertainty on the expected signal acceptance  $\times$  analysis efficiency is  
2025 determined independently for the each **SMS** model considered. These systematics stem  
2026 from uncertainties on the parton distribution functions, the luminosity measurement, jet  
2027 energy scale, b-tag scale factor measurements and the efficiencies of various selection  
2028 criteria used in the signal selection, including the  $H_T / E_T$ , dead **ECAL** cleaning filter  
2029 and lepton / photon event vetoes.

2030 Rather than trying to estimate the level of systematic that is applicable point-by-point  
2031 in a model space, general behaviours are considered; and instead constant systematics  
2032 are estimated in two regions of the **SMS** models parameter space.

2033 These two regions are defined as, near (small mass splittings) and far (large mass  
2034 splittings) from the mass degenerate diagonal, where the far region is bounded by the  
2035 condition

$$m_{\tilde{q}/\tilde{g}} - m_{LSP} > 350 GeV \quad m_{\tilde{q}/\tilde{g}} > 475 GeV.$$

2036 The total systematics in each region are evaluated in the following ways:

2037 **Jet energy scale:** The relative change in the signal efficiency is gauged by varying  
2038 the energy of all jets in an event up or down according to a  $p_T$  and  $\eta$  dependent jet  
2039 energy scale uncertainty. Within the two systematic regions, the resulting systematic  
2040 uncertainties for each **SMS** model are determined by taking the value of the 68<sup>th</sup>  
2041 percentile for the distributions of the relative change in the signal efficiency.

2042 **Luminosity measurement:** The uncertainty on the measurement of the luminosity  
2043 collected propagates through to an uncertainty on the signal event yield when  
2044 considering any new physics model, which is currently 4.4% [89].

2045 **Parton density function :** Each signal sample is produced using the CTEQ6L1  
2046 parton density function. The effect on the signal acceptance when re-weighting to  
2047 the central value of three different parton distribution functions, CT10, MSTW08  
2048 and NNPDF2.1 are examined [90]. It is found that the change of the signal efficiency  
2049 in different **SMS** models, due to the alternate PDF sets are typically a few percent,  
2050 and approaches 10% at higher squark/gluon and **LSP** masses.

2051  **$\mathcal{H}_T/\mathcal{E}_T$  cleaning filter:** The ratio of the efficiencies of the cleaning filter are  
2052 compared in simulation and data after application of the  $\mu +$  jets control sample  
2053 selection. No  $\alpha_T$  requirement or further event cleaning filters are applied. The  
2054 ratio of the efficiencies observed in data and simulation for a cut value of  $\mathcal{H}_T/\mathcal{E}_T <$   
2055 1.25 and the two jet multiplicity categories,  $2 \leq n_{jet} \leq 3$  and  $n_{jet} \geq 4$  are  $1.028 \pm$   
2056 0.007 and  $1.038 \pm 0.015$  respectively. These deviations are taken to represent the  
2057 systematic uncertainty on the simulation modelling of this variable.

2058 **Dead ECAL cleaning filter:** The ratio of the efficiencies observed in data and  
2059 simulation for this filter in the two jet multiplicity categories,  $2 \leq n_{jet} \leq 3$  and  
2060  $n_{jet} \geq 4$ , are  $0.961 \pm 0.008$  and  $0.961 \pm 0.009$ , respectively. These deviations

2061 from unity are taken to represent the systematic uncertainties in the modelling in  
2062 simulation of this filter.

2063 **Lepton and photon vetoes:** The uncertainty on the efficiency of the lepton and  
2064 photon vetoes is determined by considering truth information. The efficiency of  
2065 the vetoes is measured after applying relevant object filters with identical logic,  
2066 but based on truth instead of reconstructed objects. Where the efficiency is found  
2067 to not be 100%, it is taken to represent the fraction of signal events that are  
2068 incorrectly vetoed. This deviation is taken directly as the systematic uncertainty on  
2069 the efficiency. The systematic uncertainty is only non-zero for models which contain  
2070 third-generation quarks in the final state, where the uncertainties are at the order  
2071 of 1% level.

2072 **B-tag scale factor uncertainties:** The relative change in the signal efficiency  
2073 is observed when relevant flavour,  $p_T$  and  $\eta$  dependant b-tag correction factors,  
2074 are varied up or down by their systematic uncertainty. Within the two systematic  
2075 regions, the resulting systematic uncertainties for each **SMS** model are determined  
2076 by taking the value of the 68<sup>th</sup> percentile for the distributions of the relative change  
2077 in the signal efficiency, over all mass points.

2078 Tables 4.16 and 4.17 summarise all the aforementioned systematic uncertainties on the  
2079 signal efficiencies for each individual **SMS** model interpreted in the analysis. In the case  
2080 of the **T1tttt** model, in which pair produced gluinos decay to  $t\bar{t}$  pairs and the **LSP**,  
2081 the near region of **SMS** space is not considered, and so no systematic uncertainties are  
2082 included.

2083 In both of the defined regions it is found that the systematic uncertainties are relatively  
2084 flat justifying the approach taken. The systematic uncertainties applied to the region  
2085 near to the diagonal fall in the range 13-15%; similarly, for the region far from the  
2086 diagonal the determined uncertainties are in the range of 12-23%. These uncertainties  
2087 are all propagated through to the limit calculation.

Model	Luminosity	p.d.f	JES	$H_T/\bar{E}_T$	Dead ECAL	Lepton Veto	b-tagging	Total
T1	4.4	10.0	5.6	3.8	4.1	n/a	3.1	13.9
T2	4.4	10.0	4.1	2.8	4.1	n/a	2.4	12.9
T2tt	4.4	10.0	6.5	3.8	4.1	0.8	0.8	13.9
T2bb	4.4	10.0	4.8	2.8	4.1	0.3	2.2	13.1
T1tttt	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
T1bbbb	4.4	10.0	7.3	3.8	4.1	0.5	2.7	14.5

**Table 4.16:** Estimates of systematic uncertainties on the signal efficiency (%) for various **SMS** models when considering points in the region near to the diagonal (i.e. small mass splitting and compressed spectra). The uncertainties are added in quadrature to obtain the total.

Model	Luminosity	p.d.f	JES	$H_T/\bar{E}_T$	Dead ECAL	Lepton Veto	b-tagging	Total
T1	4.4	10.0	0.8	3.8	4.1	n/a	6.6	14.0
T2	4.4	10.0	1.1	2.8	4.1	n/a	5.8	13.4
T2tt	4.4	10.0	3.5	3.8	4.1	0.6	1.6	12.9
T2bb	4.4	10.0	0.9	2.8	4.1	0.3	2.7	12.3
T1tttt	4.4	10.0	0.5	3.8	4.1	1.4	19.4	23.0
T1bbbb	4.4	10.0	1.5	3.8	4.1	0.4	10.1	16.0

**Table 4.17:** Estimates of systematic uncertainties on the signal efficiency (%) for various **SMS** models when considering points in the region far from the diagonal (i.e. large mass splitting). The uncertainties are added in quadrature to obtain the total.

## 2088 4.8. Statistical Interpretation

2089 For a given category of events satisfying requirements on both  $n_{jet}$  and  $n_b^{reco}$ , a likelihood  
 2090 model of the observations in multiple data samples is used to gauge agreement between  
 2091 the observed yields in the hadronic signal region, and the predicted yields obtained from  
 2092 the control samples. In addition to checking whether the predictions are compatible  
 2093 with a **SM** only hypothesis, the likelihood model is also used to test for the presence  
 2094 of a variety of signal models. The statistical framework outlined within this section is  
 2095 presented in greater detail within [91].

### 2096 4.8.1. Hadronic Sample

2097 Let  $N$  be the number of bins on  $H_T$ , with  $n^i$  the number of events observed satisfying  
 2098 all selection requirements in each  $H_T$  bin i. The likelihood of the observations can then  
 2099 be written:

$$L_{had} = \prod_i \text{Pois}(n^i | b^i + s^i), \quad (4.18)$$

2100 where  $b^i$  represents the expected SM background

$$b^i = EWK_i + QCD_i, \quad (4.19)$$

2101 and  $s^i$  the expected number of signal events from the different SMS models interpreted.  
2102 Pois refers to the Poisson distribution of these values and is defined as:

$$\text{Pois}(\chi|\lambda) = \frac{\lambda^\chi \exp^{-\lambda}}{k!}. \quad (4.20)$$

### 2103 4.8.2. $H_T$ Evolution Model

2104 The hypothesis, that for a process the  $\alpha_T$  ratio falls exponentially (see Section (4.2.4))  
2105 in  $H_T$  is defined by Equation (4.10), where  $k_{QCD}$  is constrained by measurements in a  
2106 signal sideband region.

2107 The expected QCD background,  $QCD^i$ , within a bin  $i$  is then modelled as,

$$QCD^i = m^i A_{QCD} e^{-k_{QCD} \langle H_T \rangle}, \quad (4.21)$$

2108 where  $m_i$  represent the number of events observed with  $\alpha_T \leq 0.55$  in each  $H_T$  bin  $i$ , and  
2109  $\langle H_T \rangle$  represent the mean  $H_T$  of each bin. Expressed as functions of just the zeroth bin,  
2110  $QCD^0$ , and  $k_{QCD}$ , the QCD expectation is given by

$$QCD^i = QCD^0 \left( \frac{m^i}{m^0} \right) e^{-k_{QCD} (\langle H_T \rangle^i - \langle H_T \rangle^0)}. \quad (4.22)$$

2111 **4.8.3. EWK Control Samples**

2112 The **EWK** background estimation within each bin,  $i$ , is broken into two components, the  
2113 expected yield from  $Z \rightarrow \nu\bar{\nu}$  and  $t\bar{t}$ -W (plus other residual backgrounds) events. This is  
2114 written as,  $Z_{inv}^i$  and  $t\bar{t}W^i$ , and it follows that

$$EWK^i = Z_{inv}^i + t\bar{t}W^i. \quad (4.23)$$

2115 This can be further expressed as

$$Z_{inv}^i \equiv f_{Zinv}^i \times EWK^i, \quad (4.24)$$

$$t\bar{t}W^i \equiv (1 - f_{Zinv}^i) \times EWK^i, \quad (4.25)$$

2116 where  $f_{Zinv}^i$  represents the expected yield from  $Z \rightarrow \nu\bar{\nu}$  in bin  $i$  divided by the expected  
2117 **EWK** background  $EWK^i$ . This fraction is modelled as a linear component

$$f_{Zinv}^i = f_{Zinv}^0 + \frac{\langle H_T \rangle^i - \langle H_T \rangle^0}{\langle H_T \rangle^{N-1} - \langle H_T \rangle^0} (f_{Zinv}^{N-1} - f_{Zinv}^i), \quad (4.26)$$

2118 where  $N$  again represents the number of  $H_T$  bins, and  $f_{Zinv}^i$  and  $f_{Zinv}^{N-1}$  are float parameters  
2119 whose final values are limited between zero and one.

2120 Within each  $H_T$  bin there are three background measurements for the different control  
2121 samples,  $n_\gamma^i$ ,  $n_\mu^i$  and  $n_{\mu\mu}^i$ , representing the event yields from the  $\gamma +$  jets,  $\mu +$  jets and  
2122  $\mu\mu +$  jets control samples respectively. Each of these have a corresponding yield in  
2123 simulation,  $MC_\gamma^i$ ,  $MC_\mu^i$  and  $MC_{\mu\mu}^i$ . Within the hadronic signal region there are also  
2124 corresponding simulated yields for  $Z \rightarrow \nu\bar{\nu}$  ( $MC_{Zinv}^i$ ) and  $t\bar{t} + W$  ( $MC_{t\bar{t}+W}^i$ ), which are  
2125 used to define

$$r_\gamma^i = \frac{MC_\gamma^i}{MC_{Zinv}^i}; \quad r_{\mu\mu}^i = \frac{MC_{\mu\mu}^i}{MC_{Zinv}^i}; \quad r_\mu^i = \frac{MC_\mu^i}{MC_{t\bar{t}+W}^i} \quad (4.27)$$

2126 where  $r_p^i$  represents the inverse of the **TFs** used to extrapolate the yield of each background  
2127 process.

2128 The likelihoods regarding the three measured yields  $n_\gamma^i$ ,  $n_{\mu\mu}^i$ ,  $n_\mu^i$  can then be fully expressed  
2129 as

$$L_\gamma = \prod_i \text{Pois}(n_\gamma^i | \rho_{\gamma Z}^j \cdot r_\gamma^i \cdot Z_{inv}^i), \quad (4.28)$$

$$L_{\mu\mu} = \prod_i \text{Pois}(n_{\mu\mu}^i | \rho_{\mu\mu Z}^j \cdot r_{\mu\mu}^i \cdot Z_{inv}^i), \quad (4.29)$$

$$L_\mu = \prod_i \text{Pois}(n_\mu^i | \rho_{\mu Y}^j \cdot r_\mu^i \cdot Y^i + s_\mu^i), \quad (4.30)$$

$$(4.31)$$

2130 which contain an additional term  $s_\mu^i$ , which represents the signal contamination in the  
2131  $\mu + \text{jets}$  sample. The parameters  $\rho_{\gamma Z}^j$ ,  $\rho_{\mu\mu}^j$  and  $\rho_\mu^j$  represent “correction factors” that  
2132 accommodate the data driven systematic uncertainties derived from the control samples  
2133 in Section (4.12).

2134 Each of these equations are used to estimate the maximum likelihood value for relevant  
2135 background in the signal region given the observations  $n_p^i$  in each of the control samples  
2136 (see Section (4.2.3)).

2137 The measurements in each of the control samples and the hadronic signal region, along  
2138 with the ratios  $r_\gamma^i$ ,  $r_{\mu\mu}^i$ , and  $r_\mu^i$ , are all considered simultaneously through the relationships  
2139 defined by Equations (4.19),(4.24) and (4.25).

2140 In addition to the Poisson product, an additional log-normal term is introduced to  
2141 accommodate the systematic uncertainties given by,

$$L_{EWK \ syst} = \prod_j \text{Logn}(1.0 | \rho_{\mu W}^j, \sigma_{\mu W}^j) \times \text{Logn}(1.0 | \rho_{\mu\mu Z}^j, \sigma_{\mu\mu Z}^j) \times \text{Logn}(1.0 | \rho_{\gamma Z}^j, \sigma_{\gamma Z}^j). \quad (4.32)$$

2142 The parameters  $\rho^j$ ,  $\rho^j$  and  $\rho^j$  represent “correction factors” that accommodate the  
2143 systematic uncertainties associated with the control sample based background constraints.

2144 The quantities  $\sigma_{\gamma Z}^j$ ,  $\sigma_{\mu\mu Z}^j$  and  $\sigma_{\mu W}^j$  represent the relative systematic uncertainties for the  
2145 control sample constraints. Logn represents the log-normal distribution [92],

$$\text{Logn}(x \mid \mu, \sigma_{rel}) = \frac{1}{x\sqrt{2\pi}\ln k} \exp\left(\frac{\ln^2(\frac{x}{\mu})}{2\ln^2 k}\right); \quad k = 1 + \sigma_{rel}. \quad (4.33)$$

2146 Five parameters per control sample are used to span the eight  $H_T$  categories, with just  
2147 one used for the three  $H_T$  in the  $n_b^{reco} \geq 4$  category. These parameters span the same  
2148  $H_T$  ranges described in Section (4.6) and is shown in Table 4.18.

$H_T$ bin (i)	0	1	2	3	4	5	6	7
syst. parameter (j)	0	1	2	2	3	3	4	4

$H_T$ bin (i)	0	1	2
syst. parameter (j)	0	0	0

**Table 4.18:** The systematic parameters used in  $H_T$  bins. Left: categories with eight bins; right: category with three bins.

2149 Alternatively, in the higher  $n_b^{reco}$  categories ( $n_b^{reco} \geq 2$ ), only the single muon sample is  
2150 used to constrain the total EWK background. This is due to a lack of statistics in the  
2151  $\mu\mu + \text{jets}$  and  $\gamma + \text{jets}$  at these  $n_b^{reco}$  multiplicities. Therefore the likelihood functions  
2152 for the control samples are reduced and simply represented by

$$L'_\mu = \prod_i \text{Pois}(n_\mu^i | \rho_{\mu Y}^j \cdot r_\mu'^i \cdot EWK^i + s_\mu^i), \quad (4.34)$$

2153 where

$$r_\mu'^i = \frac{MC_\mu^i}{MC_{t\bar{t}+W}^i + MC_{Zinv}^i}. \quad (4.35)$$

#### 2154 4.8.4. Contributions from Signal

2155 The cross-section for each model is represented by  $x$ , while  $l$  represents the total recorded  
2156 luminosity considered by the analysis in the signal region. Let  $\epsilon_{had}^i$  and  $\epsilon_\mu^i$  represent the  
2157 analysis selection efficiency for that particular signal model in  $H_T$  bin  $i$  of the hadronic  
2158 and  $\mu + \text{jets}$  control sample respectively. Letting  $\delta$  represent the relative uncertainty on

- 2159 the signal yield, assumed to be fully correlated across all bins, and  $\rho_{sig}$  the “correction  
2160 factor” to the signal yield which accommodates this uncertainty.  $f$  represents an unknown  
2161 multiplicative factor on the signal cross section, for which an allowed interval is computed.  
2162 The expected signal yield  $s^i$  is thus given by

$$s^i \equiv f \rho_{sig} x l \epsilon_{had}^i \quad (4.36)$$

- 2163 and signal contamination with the  $\mu +$  jets control sample by

$$s_\mu^i \equiv f \rho_{sig} x l \epsilon_\mu^i. \quad (4.37)$$

- 2164 The systematic uncertainty on the signal is additionally incorporated by the term

$$L_{\text{sig}} = \text{Logn}(1.0 | \rho_{sig}, \delta). \quad (4.38)$$

- 2165 A discussion of the **SMS** signal models through which the analysis is interpreted can be  
2166 found in the following chapter.

2167 **4.8.5. Total Likelihood**

- 2168 The total likelihood function for a given signal category  $k(n_b^{reco}, n_{jet})$  is then given by  
2169 the product of the likelihood functions introduced within the previous sections:

$$\begin{aligned} L_{\text{Tot}}^k &= L_{had}^k \times L_\mu^k \times L_\gamma^k \times L_{\mu\mu}^k \times L_{EWKsyst}^k \times L_{QCD}^k & (0 \leq n_b^{\text{reco}} \leq 1), \\ L_{\text{Tot}}^k &= L_{had}^k \times L'_\mu^k \times L_{\mu\text{syst}}^k \times L_{QCD}^k & (n_b^{\text{reco}} \geq 2). \end{aligned} \quad (4.39)$$

- 2170 In categories containing eight  $H_T$  bins and utilising the three control samples ( $\mu +$  jets,  
2171  $\mu\mu +$  jets,  $\gamma +$  jets), there are 25 nuisance parameters. When just one control sample

2172 is used to estimate the **EWK** background, this is reduced to 15 nuisance parameters.  
2173 In the  $n_b^{\text{reco}} \geq 4$  category where only three  $H_T$  bins are used, there are just 6 nuisance  
2174 parameters. This information is summarised within Table 4.19.

Nuisance parameter	Total
$(EWK^i)_{i:0-7(2)}$	8 (3)
$f_{Z\text{inv}}^0$	1*
$f_{Z\text{inv}}^7$	1*
$QCD^0$	1
$k_{QCD}$	1
$(\rho_{\gamma Z}^j)_{j:2-4}$	3 *
$(\rho_{\mu\mu Z}^j)_{j:0-4}$	5 *
$(\rho_{\mu W}^j)_{j:0-4(0)}$	5 (1)

**Table 4.19:** Nuisance parameters used within the different hadronic signal bins of the analysis. Parameters denoted by a \* are not considered in the case of a single control sample being used to predict the **EWK** background. Numbers within brackets highlight the number of nuisance parameters in the case of three  $H_T$  bins being used.

2175 When considering **SUSY** signal models within the likelihood, the additional  $L_{\text{sig}}$  term  
2176 is included and therefore when multiple categories are fitted simultaneously the total  
2177 likelihood is then represented by

$$L_{\text{Tot}}^{\text{signal}} = L_{\text{sig}} \times \prod_k L_{\text{Tot}}^k. \quad (4.40)$$

# Chapter 5.

## <sup>2178</sup> Results and Interpretation

<sup>2179</sup> Using the statistical framework outlined in the previous chapter, results are shown for  
<sup>2180</sup> the compatibility of the collected data with a **SM**-only hypothesis in Section (5.1). The  
<sup>2181</sup> data is further interpreted within the context of various **SMS** models within Section (5.2).

### <sup>2182</sup> 5.1. Compatibility with the Standard Model Hypothesis

<sup>2184</sup> The **SM** background only hypothesis is tested by removing any signal contributions  
<sup>2185</sup> within the signal and control samples, and the likelihood function defined in Equation  
<sup>2186</sup> (4.39) maximised over all parameters using Rootfit [93] and MINUIT [94]. The results of  
<sup>2187</sup> the search consist of the observed yields in the hadronic signal sample, and the  $\mu + \text{jets}$ ,  
<sup>2188</sup>  $\mu\mu + \text{jets}$  and  $\gamma + \text{jets}$  control samples.

<sup>2189</sup> These observed yields along with the expectations and uncertainties given by the simulta-  
<sup>2190</sup> neous fit for the hadronic signal region are given in Table 5.2. The results obtained from  
<sup>2191</sup> the simultaneous fits, including that of the three control samples, are shown in Figure  
<sup>2192</sup> 5.1-5.8, as summarised in Table 5.1.

<sup>2193</sup> The figures show a comparison between the observed yields and the **SM** expectations  
<sup>2194</sup> across all  $H_T$  bins, and in all  $n_{jet}$  and  $n_b^{reco}$  multiplicity categories. In all categories the  
<sup>2195</sup> samples are well described by the **SM** only hypothesis. In particular no significant excess  
<sup>2196</sup> is observed above **SM** expectation within the hadronic signal region.

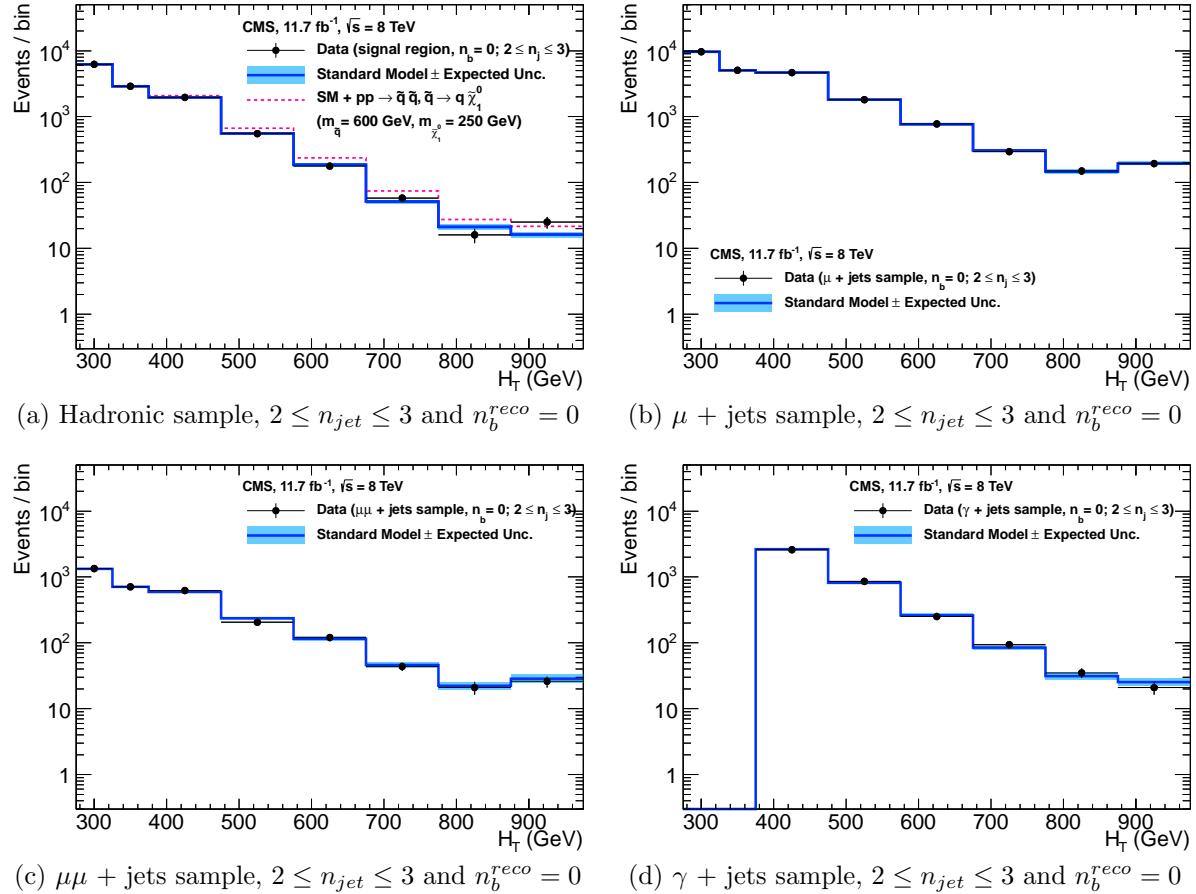
$n_{jet}$	$n_b^{reco}$	Control samples fitted	Figure
2-3	0	$\mu + \text{jets}$ , $\mu\mu + \text{jets}$ , $\gamma + \text{jets}$	5.1
2-3	1	$\mu + \text{jets}$ , $\mu\mu + \text{jets}$ , $\gamma + \text{jets}$	5.2
2-3	2	$\mu + \text{jets}$	5.3
$\geq 4$	0	$\mu + \text{jets}$ , $\mu\mu + \text{jets}$ , $\gamma + \text{jets}$	5.4
$\geq 4$	1	$\mu + \text{jets}$ , $\mu\mu + \text{jets}$ , $\gamma + \text{jets}$	5.5
$\geq 4$	2	$\mu + \text{jets}$	5.6
$\geq 4$	3	$\mu + \text{jets}$	5.7
$\geq 4$	4	$\mu + \text{jets}$	5.8

**Table 5.1:** Summary of control samples used by each fit results, and the Figures in which they are displayed.

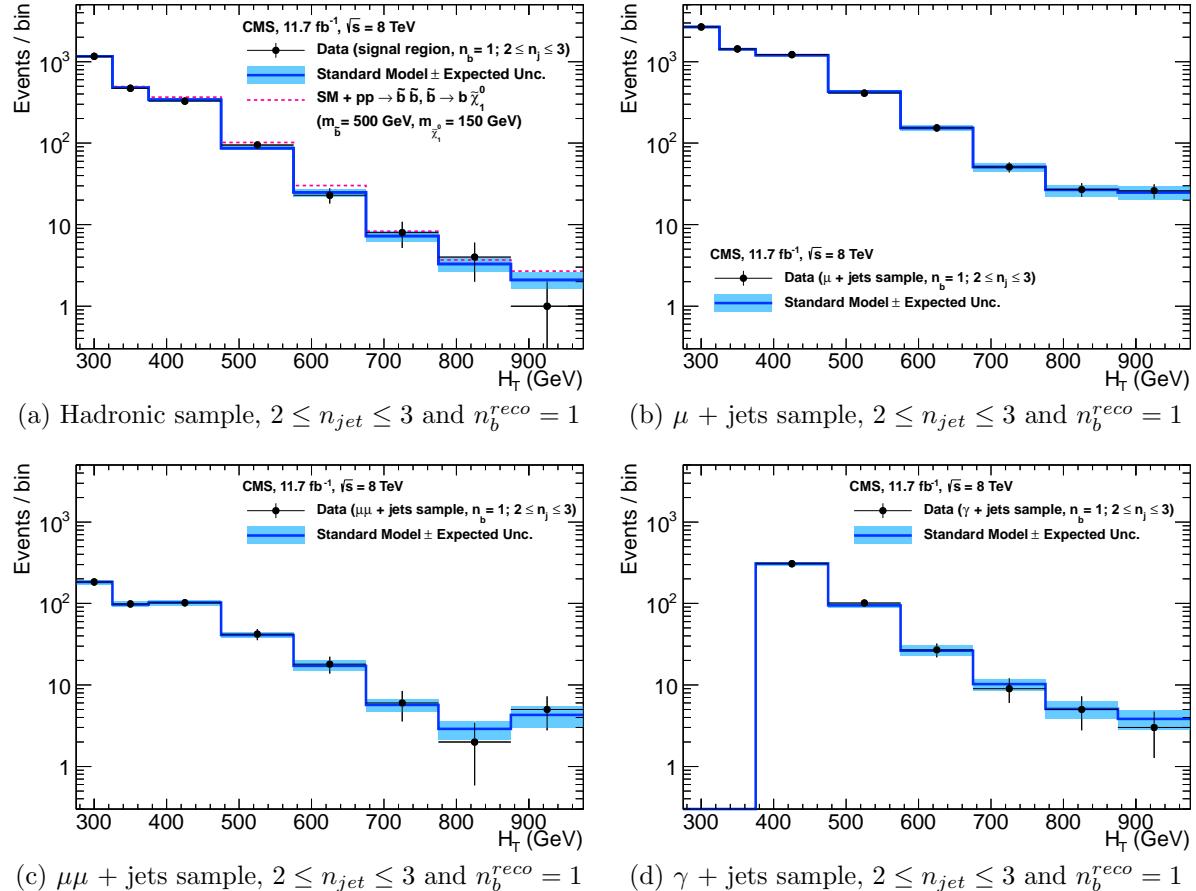
Cat	$n_b^{reco}$	$n_{jet}$	$H_T$ bin (GeV)							
			275-325	325-375	375-475	474-575	575-675	675-775	775-875	875- $\infty$
SM Data	0	$\leq 3$	$6235^{+100}_{-67}$ 6232	$2900^{+60}_{-54}$ 2904	$1955^{+34}_{-39}$ 1965	$558^{+14}_{-15}$ 552	$186^{+11}_{-10}$ 177	$51.3^{+3.4}_{-3.8}$ 58	$21.2^{+2.3}_{-2.2}$ 16	$16.1^{+1.7}_{-1.7}$ 25
		$\geq 4$	$1010^{+34}_{-24}$ 1009	$447^{+19}_{-16}$ 452	$390^{+19}_{-15}$ 375	$250^{+12}_{-11}$ 274	$111^{+9}_{-7}$ 113	$53.3^{+4.3}_{-4.3}$ 56	$18.5^{+2.4}_{-2.4}$ 16	$19.4^{+2.5}_{-2.7}$ 27
SM Data	1	$\leq 3$	$1162^{+37}_{-29}$ 1164	$481^{+18}_{-19}$ 473	$341^{+15}_{-16}$ 329	$86.7^{+4.2}_{-5.6}$ 95	$24.8^{+2.8}_{-2.7}$ 23	$7.2^{+1.1}_{-1.0}$ 8	$3.3^{+0.7}_{-0.7}$ 4	$2.1^{+0.5}_{-0.5}$ 1
		$\geq 4$	$521^{+25}_{-17}$ 515	$232^{+15}_{-12}$ 236	$188^{+12}_{-11}$ 204	$106^{+6}_{-6}$ 92	$42.1^{+4.1}_{-4.4}$ 51	$17.9^{+2.2}_{-2.0}$ 13	$9.8^{+1.5}_{-1.4}$ 13	$6.8^{+1.2}_{-1.1}$ 6
SM Data	2	$\leq 3$	$224^{+15}_{-14}$ 222	$98.2^{+8.4}_{-6.4}$ 107	$59.0^{+5.2}_{-6.0}$ 58	$12.8^{+1.6}_{-1.6}$ 12	$3.0^{+0.9}_{-0.7}$ 5	$0.5^{+0.2}_{-0.2}$ 1	$0.1^{+0.1}_{-0.1}$ 0	$0.1^{+0.1}_{-0.1}$ 0
		$\geq 4$	$208^{+17}_{-9}$ 204	$103^{+9}_{-7}$ 107	$85.9^{+7.2}_{-6.9}$ 84	$51.7^{+4.6}_{-4.7}$ 59	$19.9^{+3.4}_{-3.0}$ 24	$6.8^{+1.2}_{-1.3}$ 5	$1.7^{+0.7}_{-0.4}$ 1	$1.3^{+0.4}_{-0.3}$ 2
SM Data	3	$\geq 4$	$25.3^{+5.0}_{-4.2}$ 25	$11.7^{+1.7}_{-1.8}$ 13	$6.7^{+1.4}_{-1.2}$ 4	$3.9^{+0.8}_{-0.8}$ 2	$2.3^{+0.6}_{-0.6}$ 2	$1.2^{+0.3}_{-0.4}$ 3	$0.3^{+0.2}_{-0.1}$ 0	$0.1^{+0.1}_{-0.1}$ 0
		$\geq 4$	$0.9^{+0.4}_{-0.7}$ 1	$0.3^{+0.2}_{-0.2}$ 0				$0.6^{+0.3}_{-0.3}$ 2		

**Table 5.2:** Comparison of the measured yields in each  $H_T$ ,  $n_{jet}$  and  $n_b^{reco}$  jet multiplicity bins for the hadronic sample with the **SM** expectations and combined statistical and systematic uncertainties given by the simultaneous fit.

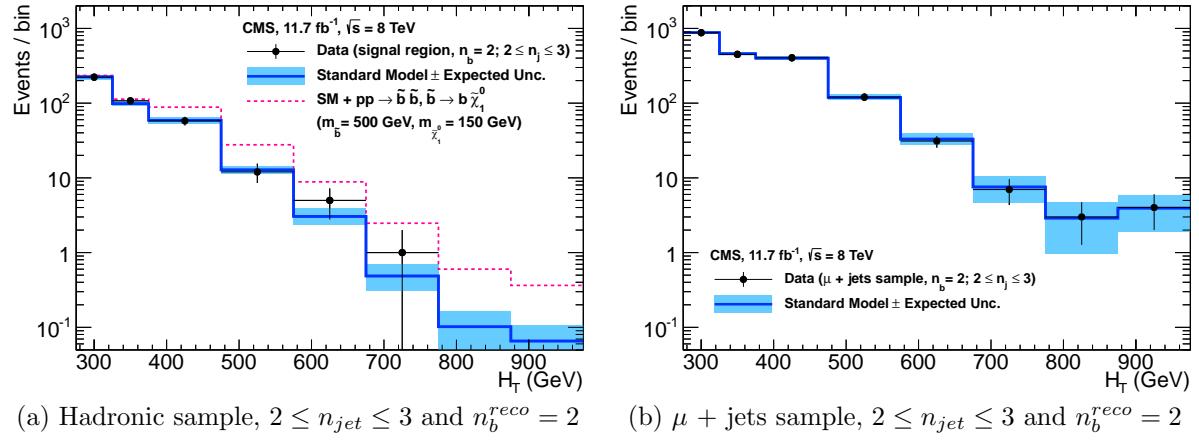
<sup>2197</sup> Given the lack of an excess in data hinting at a possible supersymmetric signature within  
<sup>2198</sup> the data, interpretations are made on the production masses and cross-section of a range  
<sup>2199</sup> of **SUSY** decay topologies within the following section.



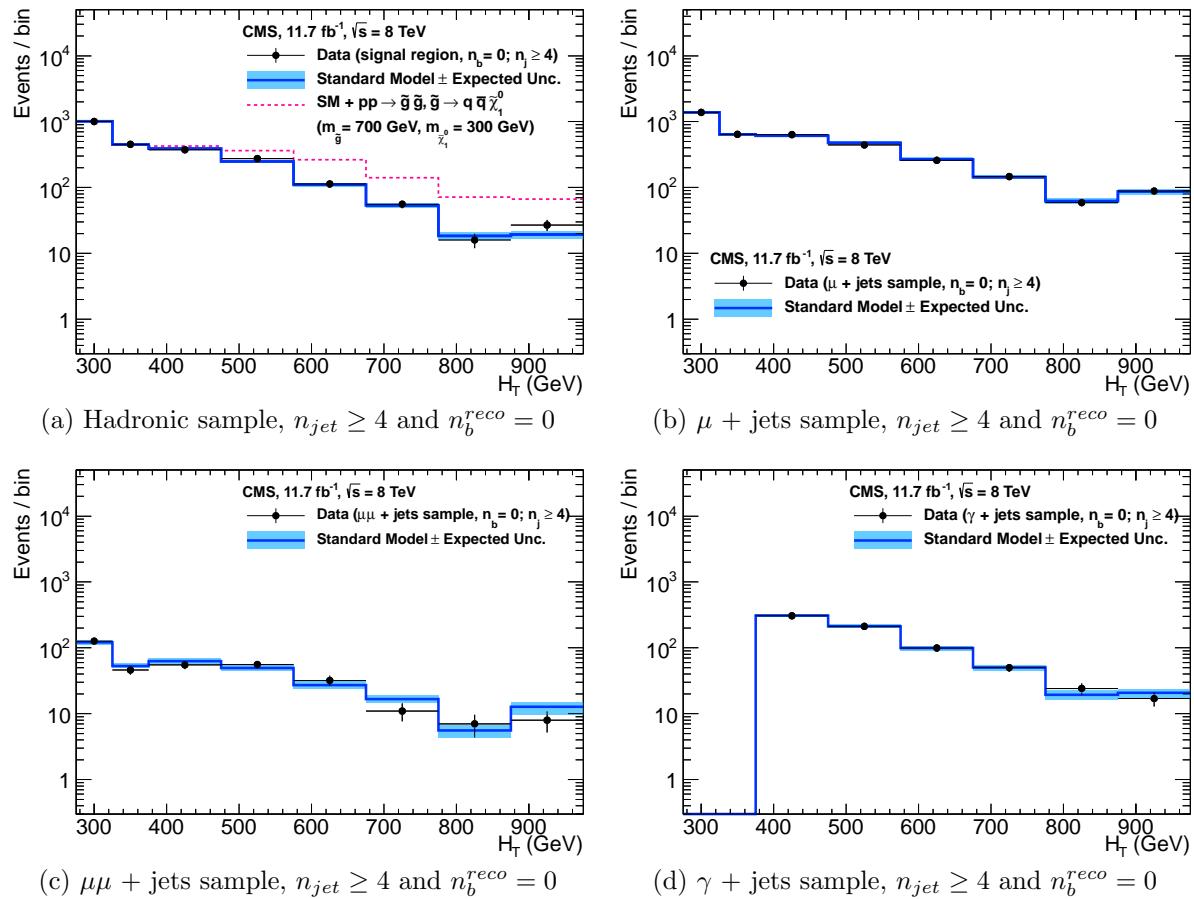
**Figure 5.1:** Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of  $H_T$  for the (a) hadronic, (b)  $\mu +$  jets, (c)  $\mu\mu +$  jets and (d)  $\gamma +$  jets samples when requiring  $n_b^{reco} = 0$  and  $n_{jet} \leq 3$ . The observed event yields in data (black dots) and the expectations and their uncertainties for all SM processes (blue line with light blue bands) are shown. An example signal expectation (red solid line) for the D1 SMS signal point from Table 4.1 is superimposed on the SM background expectation.



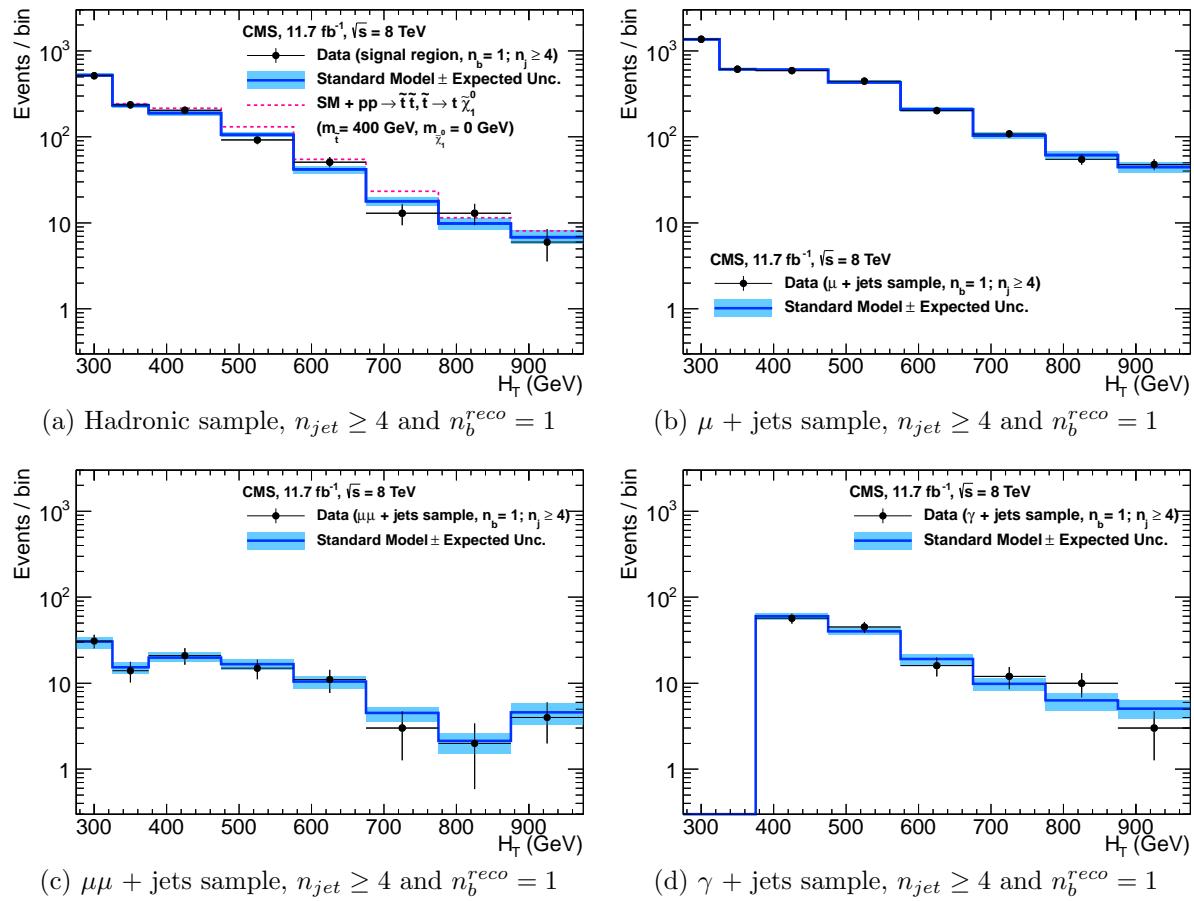
**Figure 5.2:** Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of  $H_T$  for the (a) hadronic, (b)  $\mu +$  jets, (c)  $\mu\mu +$  jets and (d)  $\gamma +$  jets samples when requiring  $n_b^{reco} = 1$  and  $n_{jet} \leq 3$ . The observed event yields in data (black dots) and the expectations and their uncertainties for all SM processes (blue line with light blue bands) are shown. An example signal expectation (red solid line) for the D2 SMS signal point from Table 4.1 is superimposed on the SM background expectation.



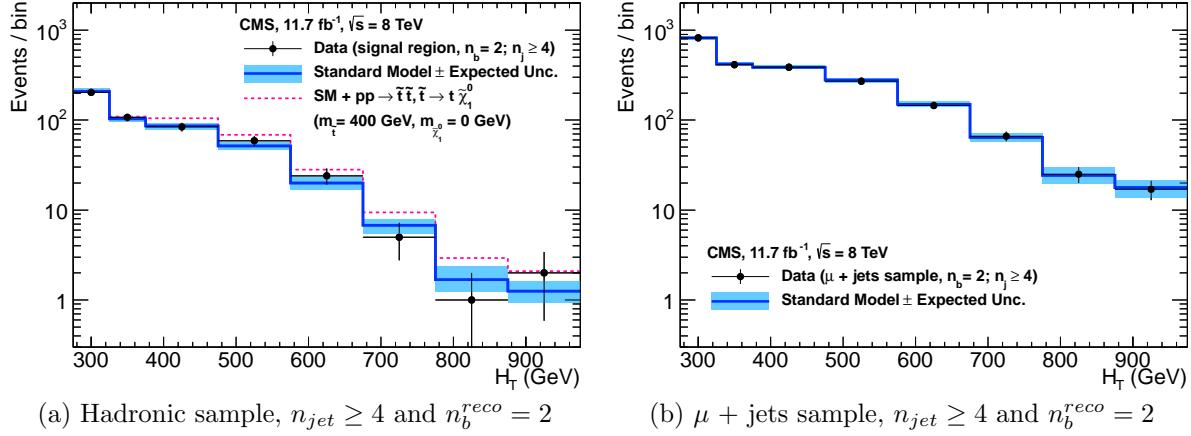
**Figure 5.3:** Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of  $H_T$  for the (a) hadronic, (b)  $\mu +$  jets, (c)  $\mu\mu +$  jets and (d)  $\gamma +$  jets samples when requiring  $n_b^{reco} = 2$  and  $n_{jet} \leq 3$ . The observed event yields in data (black dots) and the expectations and their uncertainties for all SM processes (blue line with light blue bands) are shown. An example signal expectation (red solid line) for the D2 SMS signal point from Table 4.1 is superimposed on the SM background expectation.



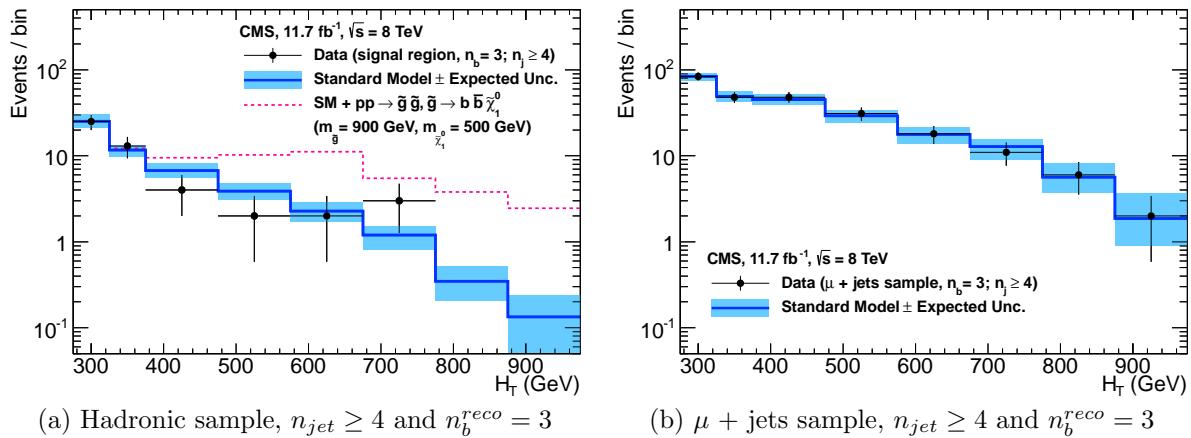
**Figure 5.4:** Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of  $H_T$  for the (a) hadronic, (b)  $\mu +$  jets, (c)  $\mu\mu +$  jets and (d)  $\gamma +$  jets samples when requiring  $n_b^{reco} = 0$  and  $n_{jet} \geq 4$ . The observed event yields in data (black dots) and the expectations and their uncertainties for all SM processes (blue line with light blue bands) are shown. An example signal expectation (red solid line) for the D2 SMS signal point from Table 4.1 is superimposed on the SM background expectation.



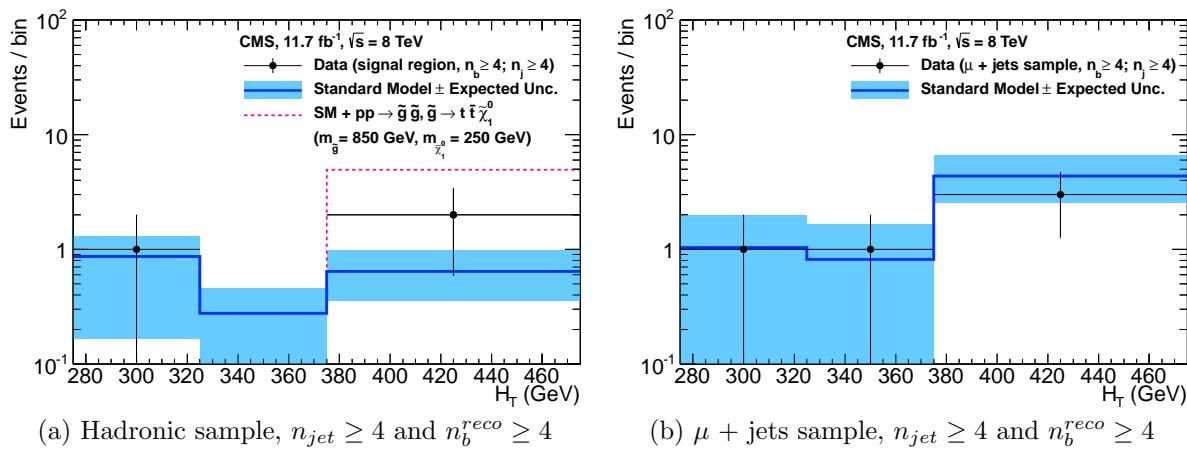
**Figure 5.5:** Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of  $H_T$  for the (a) hadronic, (b)  $\mu +$  jets, (c)  $\mu\mu +$  jets and (d)  $\gamma +$  jets samples when requiring  $n_b^{reco} = 1$  and  $n_{jet} \geq 4$ . The observed event yields in data (black dots) and the expectations and their uncertainties for all SM processes (blue line with light blue bands) are shown.



**Figure 5.6:** Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of  $H_T$  for the (a) hadronic, (b)  $\mu +$  jets, (c)  $\mu\mu +$  jets and (d)  $\gamma +$  jets samples when requiring  $n_b^{reco} = 2$  and  $n_{jet} \geq 4$ . The observed event yields in data (black dots) and the expectations and their uncertainties for all SM processes (blue line with light blue bands) are shown. An example signal expectation (red solid line) for the D3 SMS signal point from Table 4.1 is superimposed on the SM background expectation.



**Figure 5.7:** Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of  $H_T$  for the (a) hadronic, (b)  $\mu +$  jets, (c)  $\mu\mu +$  jets and (d)  $\gamma +$  jets samples when requiring  $n_b^{reco} = 3$  and  $n_{jet} \geq 4$ . The observed event yields in data (black dots) and the expectations and their uncertainties for all SM processes (blue line with light blue bands) are shown. An example signal expectation (red solid line) for the G2 SMS signal point from Table 4.1 is superimposed on the SM background expectation.



**Figure 5.8:** Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of  $H_T$  for the (a) hadronic, (b)  $\mu +$  jets, (c)  $\mu\mu +$  jets and (d)  $\gamma +$  jets samples when requiring  $n_b^{reco} \geq 4$  and  $n_{jet} \geq 4$ . The observed event yields in data (black dots) and the expectations and their uncertainties for all SM processes (blue line with light blue bands) are shown. An example signal expectation (red solid line) for the G3 SMS signal point from Table 4.1 is superimposed on the SM background expectation.

2200 **5.2. SUSY**

2201 Limits are set in the parameter space of a set of **SMS** models that characterise both  
2202 natural **SUSY** third generation squark production, and compressed spectra where the  
2203 mass splitting between the particle and **LSP** is small, leading to soft final state jets.  
2204 However as detailed in Section (2.4.1), the individual models are not representative of a  
2205 real physical **SUSY** model as only one decay process is considered. Instead these models  
2206 represent a way to test for signs of specific signatures indicating new physics.

2207 **5.2.1. The  $CL_s$  Method**

2208 The CLs method [95][96][97] is used to compute the limits for signal models, with the  
2209 one-sided profile likelihood ratio as the test statistic [98].

The test statistic is defined as

$$q(\mu) = \begin{cases} -2\log\lambda(\mu) & \text{when } \mu \geq \hat{\mu}, \\ 0 & \text{otherwise.} \end{cases} \quad (5.1)$$

2210 where

$$\lambda(\mu) = \frac{L(\mu, \theta_\mu)}{L(\hat{\mu}, \hat{\theta})} \quad (5.2)$$

2211 represents the profile likelihood ratio, in which  $\mu \equiv f$  from Section (4.8.4), is the  
2212 parameter characterising the signal strength.  $\hat{\mu}$  is defined as the maximum likelihood  
2213 value,  $\hat{\theta}$  the set of maximum likelihood values of the nuisance parameters and  $\theta_\mu$  the set  
2214 of maximum values of the nuisance parameters for a given value of  $\mu$ .

2215 When  $\mu \equiv f = 1$ , the signal model is considered at its nominal production cross section.  
2216 The distribution of  $q_\mu$  is built up via the generation of pseudo experiments in order to  
2217 obtain two distributions for the background (B) and signal plus background (S+B) cases.

2218 The compatibility of a signal model with observations in data is determined by the  
2219 parameter  $CL_s$ ,

$$\text{CL}_S = \frac{\text{CL}_{S+B}}{\text{CL}_B}, \quad (5.3)$$

with  $\text{CL}_B$  and  $\text{CL}_{S+B}$  defined as one minus the quantiles of the observed value in the data of the two distributions. A model is considered to be excluded at 95% confidence level when  $\text{CL}_s \leq 0.05$  [99].

### 5.2.2. Interpretation in Simplified Signal Models

Different  $n_{\text{jet}}$  and  $n_b^{\text{reco}}$  bins are used in the interpretation of different **SMS** models. The choice of categories used, are made such that the signal to background ratio will be maximised for the model in question, increasing sensitivity to that particular type of final state signature. The production and decay modes of the **SMS** models under consideration are summarised in Table 5.3, with limit plots of the experimental reach in these models shown in Figure 5.10.

The models T1 and T2 are used to characterise the pair production of gluinos and first or second generation squarks respectively. The low number of third generation quarks produced from this decay topology makes choosing to interpret within the  $n_b^{\text{reco}} = 0$  category beneficial to improving sensitivity to these models. In the case of the T2 model, two sets of exclusion contours are shown. These correspond to the production of eight first- and second-generation (left-/right-handed) squarks with degenerate masses and the case of just a single light squark with all other squarks decoupled at much higher masses.

Conversely the T2bb, T1tttt, and T1bbbb **SMS** models describe various production and decay mechanisms in the context of third-generation squarks. In this situation considering higher  $n_b^{\text{reco}}$  categories bring significant improvements to the sensitivity to these types of final state signature.

Finally the choice of jet category is made dependant upon the production mechanism, where gluino induced and direct squark production results in a large or small number of final state jets respectively.

Experimental uncertainties on the **SM** background predictions (10 – 30%, described in Section (4.6.1)), the luminosity measurement (4.4%), and the total acceptance times

Model	Production/decay	$n_{jet}$	$n_b^{reco}$	Process	Limit	$m_{\tilde{q}/\tilde{g}}^{\text{best}}$ (GeV)	$m_{\text{LSP}}^{\text{best}}$ (GeV)
T1	$pp \rightarrow \tilde{g}\tilde{g}^* \rightarrow q\bar{q}\tilde{\chi}_1^0 q\bar{q}\tilde{\chi}_1^0$	$\geq 4$	0	5.9(a)	5.10(a)	$\sim 950$	$\sim 450$
T2	$pp \rightarrow \tilde{q}\tilde{q}^* \rightarrow q\tilde{\chi}_1^0 q\tilde{\chi}_1^0$	$\leq 3$	0	5.9(b)	5.10(b)	$\sim 775$	$\sim 325$
T2bb	$pp \rightarrow \tilde{b}\tilde{b}^* \rightarrow b\tilde{\chi}_1^0 \bar{b}\tilde{\chi}_1^0$	$\leq 3$	1,2	5.9(c)	5.10(c)	$\sim 600$	$\sim 200$
T1tttt	$pp \rightarrow \tilde{g}\tilde{g}^* \rightarrow t\bar{t}\tilde{\chi}_1^0 t\bar{t}\tilde{\chi}_1^0$	$\geq 4$	2,3, $\geq 4$	5.9(d)	5.10(d)	$\sim 975$	$\sim 325$
T1bbbb	$pp \rightarrow \tilde{g}\tilde{g}^* \rightarrow b\bar{b}\tilde{\chi}_1^0 b\bar{b}\tilde{\chi}_1^0$	$\geq 4$	2,3, $\geq 4$	5.9(e)	5.10(e)	$\sim 1125$	$\sim 650$

**Table 5.3:** A table representing the **SMS** models interpreted within the analysis. The model name and production and decay chain is specified in the first two columns. Each **SMS** model is interpreted in specific  $n_{jet}$  and  $n_b^{reco}$  categories which are detailed in the third and fourth columns. The last two columns indicate the search sensitivity for each model, representing the largest  $m_{\tilde{q}/\tilde{g}}$  mass beyond which no limit can be set for this particular decay topology. The quoted values are conservatively determined from the observed exclusion based on the theoretical production cross section minus  $1\sigma$  uncertainty.

efficiency of the selection for the considered signal model (12 –18%, from Section (4.7)) are included in the calculation of the limit.

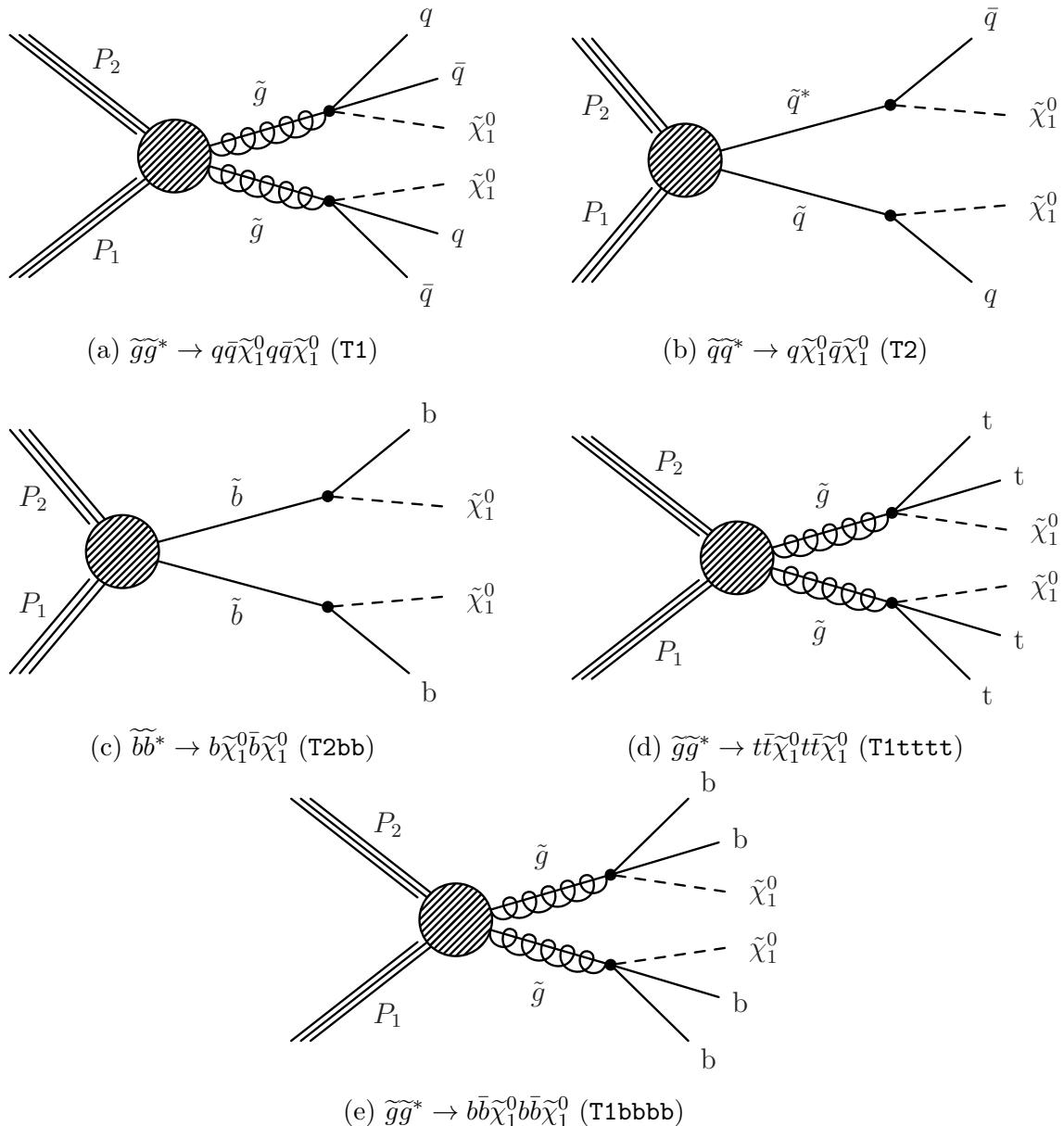
Signal efficiency in the kinematic region defined by  $0 < m_{\tilde{q}/\tilde{g}} < 175$  GeV or  $m_{\tilde{q}/\tilde{g}} < 300$  GeV is strongly affected by the presence of Initial State Radiation (**ISR**). This is a region in which direct (i.e. non-**ISR** induced) production is kinematically forbidden due to the  $H_T > 275$  GeV requirement, therefore a large percentage of signal acceptance is due to the effect of **ISR** jets. Given the large associated uncertainties, no interpretation is provided for this kinematic region.

The estimates on mass limits shown in Table 5.3, are determined conservatively from the observed exclusion based on the theoretical production cross section, minus  $1\sigma$  uncertainty. The most stringent mass limits on pair-produced sparticles are obtained at low **LSP** masses and larger squark and gluino masses due to the high  $p_T$  jets and consequently high  $H_T$  of such signal topologies. The limits are seen to weaken for compressed spectra points closer to the diagonal, where the signal populates the lower  $H_T$  bins in which more background resides. For all of the considered models, there is an **LSP** mass beyond which no limit can be set, which can be observed from the figures referenced in the table.

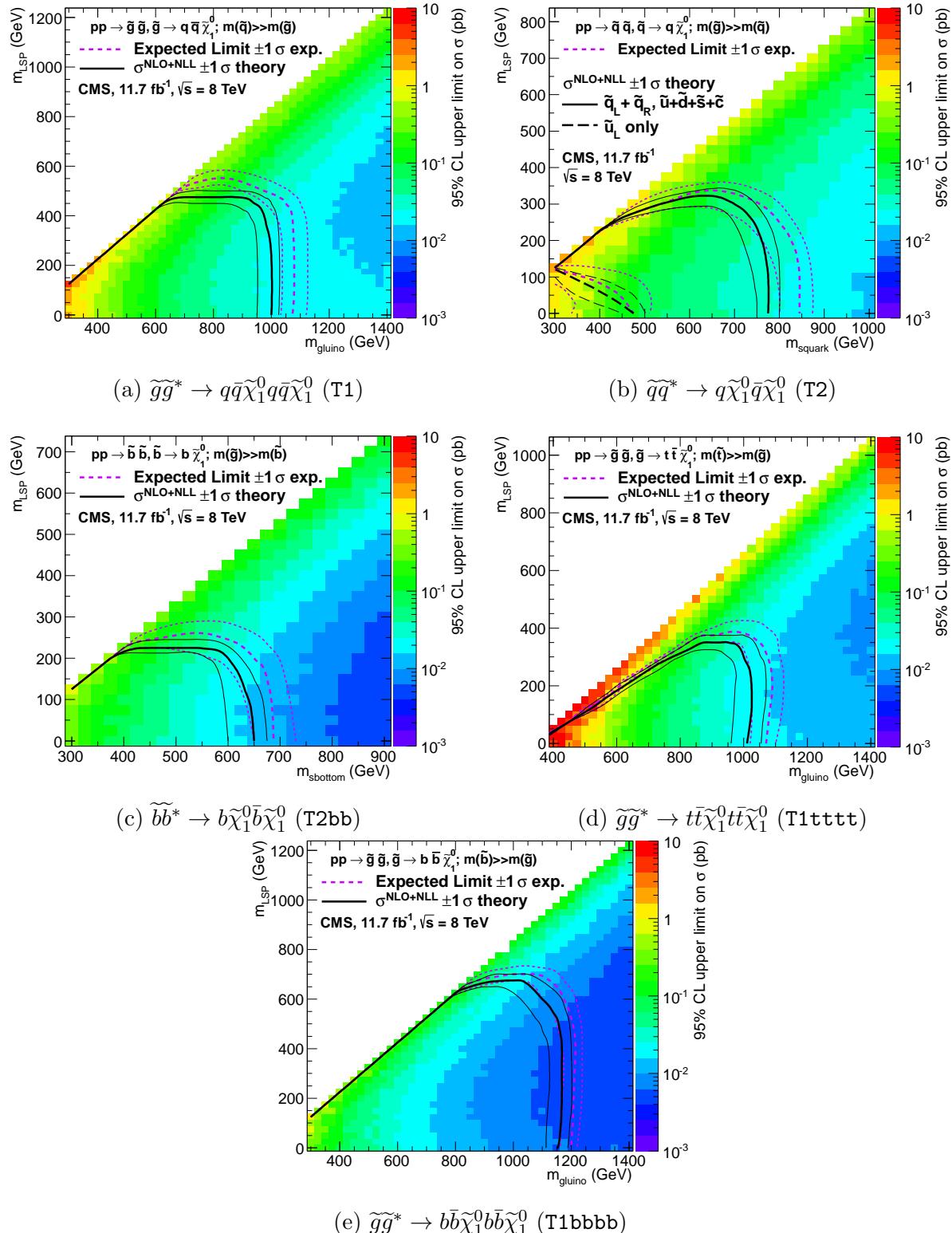
Two small upwards fluctuations are observed within the data, and are seen at high  $H_T$  within the  $n_b^{reco} = 0$  category and at mid- $H_T$  in the  $n_b^{reco} = 1, 2$  categories, see Table 5.2. As each of these fluctuations occur within at least one of the analysis categories that each **SMS** model interpretation is made, the observed exclusions within all **SMS** models are generally found to be weaker than the expected limits in the region of 1-2 standard

2267 deviations. In isolation these fluctuations are not significant and additional data would  
2268 be necessary to make any further conclusions.

2269 Despite these fluctuations, the range of parameter space that can be excluded has been  
2270 extended with respect to analysis based upon the  $\sqrt{s} = 7$  TeV dataset [100], by up to  
2271 225 and 150 GeV for  $m_{\tilde{q}(\tilde{g})}^{\text{best}}$  and  $m_{LSP}^{\text{best}}$  respectively. The parameter space for light third  
2272 generation squarks, the main tenet of natural SUSY models, is increasingly squeezed for  
2273 larger mass splitting, with exclusions in the region of 1 TeV in these topologies.



**Figure 5.9:** Production and decay modes for the various SMS models interpreted within the analysis.



**Figure 5.10:** Upper limit of cross section at 95% CL as a function of  $m_{\tilde{q}/\tilde{g}}$  and  $m_{LSP}$  for various SMS models. The solid thick black line indicates the observed exclusion region assuming NLO and NLL SUSY production cross section. The analysis selection efficiency is measured for each interpreted model, with the signal yield per point given by  $\epsilon \times \sigma$ . The thin black lines represent the observed excluded region when varying the cross section by its theoretical uncertainty. The dashed purple lines indicate the median (thick line)  $1\sigma$  (thin lines) expected exclusion regions.

# Chapter 6.

## <sup>2274</sup> SUSY Searches with B-tag

### <sup>2275</sup> Templates

<sup>2276</sup> Within this chapter a complementary technique is discussed as a means to predict the  
<sup>2277</sup> distribution of three and four reconstructed b-tagged ( $n_b^{\text{reco}} = 3, 4$ ), jets in an event  
<sup>2278</sup> sample. The recent discovery of the Higgs boson has made “Natural SUSY” models  
<sup>2279</sup> attractive, given that light top and bottom squarks are a candidate to stabilise divergent  
<sup>2280</sup> loop corrections to the Higgs boson mass. A light gluino which subsequently decays to  
<sup>2281</sup> third generation sparticle pairs, will give rise to many events with a large number of final  
<sup>2282</sup> state b-tagged jets.

<sup>2283</sup> The method described within this chapter is used to estimate the SM background at  
<sup>2284</sup> high b-tagged jet multiplicities (3-4), from a templated fit conducted in a low b-tagged  
<sup>2285</sup> jet (0-2) control region of an event sample. This approach can hypothetically be applied  
<sup>2286</sup> to generic supersymmetric searches, to gain sensitivity to signals which contain a higher  
<sup>2287</sup> number of b-tagged jets than the search’s dominant SM backgrounds.

<sup>2288</sup> As a proof-of-concept, the procedure is applied to the SM enriched  $\mu + \text{jets}$  control  
<sup>2289</sup> sample of the  $\alpha_T$  search detailed in Chapter 4, and validated in both data and simulation.  
<sup>2290</sup> This method is then further utilised to provide an independent crosscheck of the SM  
<sup>2291</sup> background estimations determined by the  $\alpha_T$  search within its hadronic signal region at  
<sup>2292</sup> high b-tagged jet multiplicities.

<sup>2293</sup> To highlight the relative insensitivity of this method to the choice of b-tagging algorithm  
<sup>2294</sup> working point, results are presented using the CSV tagger (introduced in Section (3.3.2))  
<sup>2295</sup> for the “Loose”, “Medium” and “Tight” working points.

## 2296 6.1. Concept

2297 The dominant **SM** backgrounds of most **SUSY** searches are typically  $t\bar{t}$  + jets,  $W$  + jets,  
2298  $Z \rightarrow \nu\bar{\nu}$  + jets or other rare processes (e.g. Diboson,  $t\bar{t}W$  + jets production in the case of  
2299 hadronic searches) with neutrinos in the final state. These processes are characterised by  
2300 typically having zero or two underlying b-quarks per event as shown in Table 6.1. This  
2301 ultimately means that the resultant shape of the  $n_b^{\text{reco}}$  distribution for these two types of  
2302 event topologies will differ significantly due to the varying tagging probabilities of the  
2303 different jet flavours present in the final state of these processes.

2304 Similarly, **SMS** models comprising the gluino-mediated production of third generation  
2305 squarks, such at the T1tttt and T1bbbb models described in the previous chapter, will  
2306 contain four underlying b-quarks in its decay. Therefore the resultant shape of the  
2307  $n_b^{\text{reco}}$  distribution from such a signal will be further skewed towards a higher number of  
2308 b-tagged jets. As **SM** processes with a similarly large number of underlying b-quarks  
2309 are rare, a signal indicative of natural **SUSY** can potentially be easily identified, via an  
2310 observed excess of  $n_b^{\text{reco}} = 3, \geq 4$  events with respect to the expected yields from **SM**  
2311 processes.

Typical underlying b-quark content	Process
= 0	$W \rightarrow l\nu$ + jets $Z \rightarrow \nu\bar{\nu}$ + jets $Z/\gamma^* \rightarrow \mu\mu$ + jets
= 1	$t$ + jets
= 2	$t\bar{t}$ + jets

**Table 6.1:** Typical underlying b-quark content of different **SM** processes which are common to many **SUSY** searches.

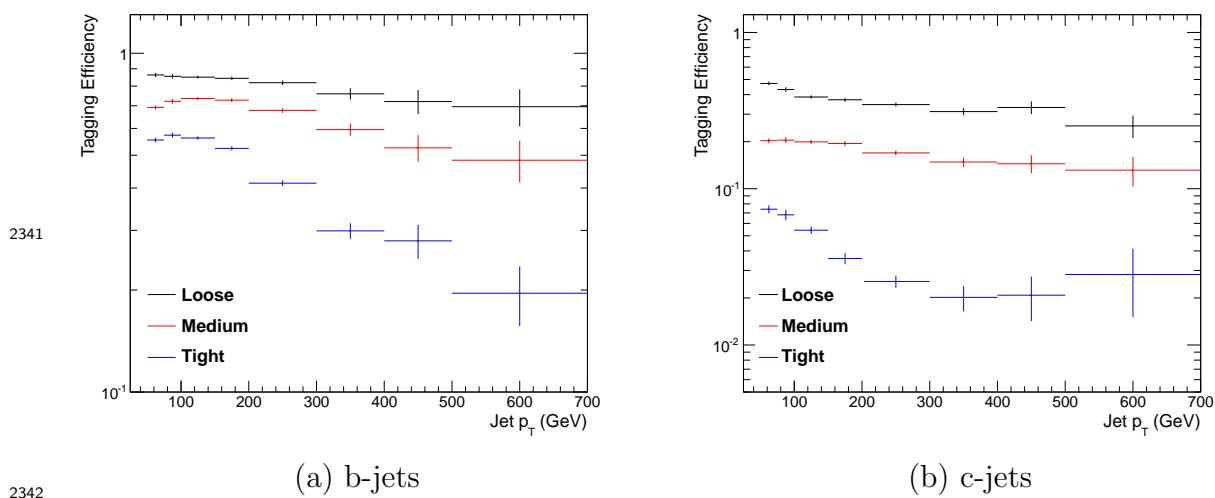
2312 Within a supersymmetric or indeed any search for new physics, the compatibility of the  
2313  $n_b^{\text{reco}}$  distribution in data with **SM** expectations can be tested, via the shape parametrisation of the **SM** background  $n_b^{\text{reco}}$  distribution, grouped in terms of these two most  
2314 common underlying b-quark topologies.

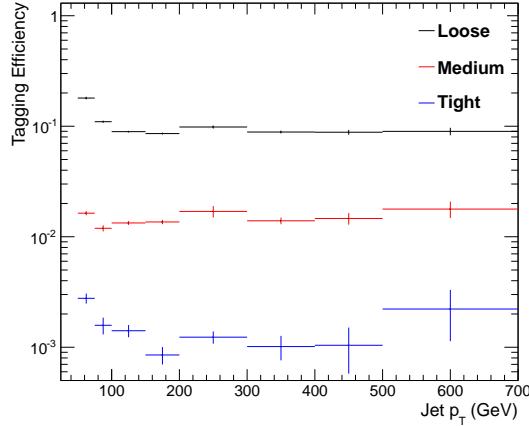
2316 Two templates, representing processes which have an underlying b-quark content of  
2317 zero or two are defined as Z0 and Z2 respectively (single top processes are a negligible  
2318 background, < 1%, within the  $\alpha_T$  search to which this method is applied in the following  
2319 section, and are thus incorporated within the Z2 template). **SM** background estimates at  
2320 high  $n_b^{\text{reco}}$  multiplicities can then be extrapolated from the fitting of these two template

2321 shapes in a low  $n_b^{\text{reco}}$  control region (0-2) under the assumption of negligible signal  
 2322 contamination.

2323 The simplest way to determine the shapes of the  $n_b^{\text{reco}}$  distributions for both templates  
 2324 would be, after the application of the relevant event selection, to take the  $n_b^{\text{reco}}$  distribution  
 2325 as given directly from simulation. However as discussed within Section (4.5), there are  
 2326 large statistical uncertainties in simulation at high  $n_b^{\text{reco}}$  multiplicities (which is the region  
 2327 in which we wish to use the templates to estimate the SM backgrounds). This statistical  
 2328 uncertainty is particularly pronounced for processes incorporated within the Z0 templates,  
 2329 where events with a large number reconstructed b-tagged jets stem largely from the  
 2330 mis-tagging of all the light-flavoured jets in the final state. Therefore to improve the  
 2331 statistical precision of the final background prediction at high b-tagged jet multiplicities,  
 2332 the formula method first introduced in Section (4.5.1) is utilised to generate the template  
 2333 shapes.

2334 The template shapes of each analysis category ( $H_T$  and  $n_{\text{jet}}$  in the case of the  $\alpha_T$  analysis)  
 2335 are dependant upon the jet-flavour content and b-tagging rate within the phase space  
 2336 of interest, with the tagging probabilities of a jet being a function of the jet  $p_T$ , the  
 2337 pseudo-rapidity  $|\eta|$ , and jet-flavour. This can be seen in Figure 6.1, where the tagging  
 2338 efficiency of jets identified from truth information as stemming from the hadronisation  
 2339 of a b, c or light quark is shown for the three working points of the CSV tagger as a  
 2340 function of jet  $p_T$ .





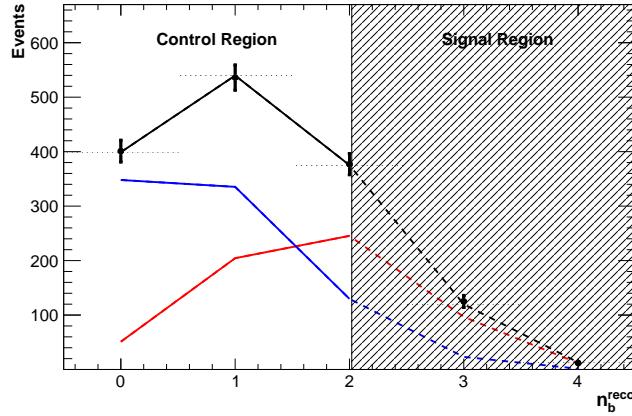
(c) light-jets

**Figure 6.1:** The b-quark (a), c-quark (b), and light-quark (c) tagging efficiency as a function of jet  $p_T$ , measured in simulation after the application of  $\alpha_T$  analysis  $\mu + \text{jets}$  control sample selection, in the region  $H_T > 375$ . Efficiencies are measured for the three **CSV** working points.

2343 Therefore, before the template shapes are generated by the formula method, the jet  $p_T$   
 2344 and  $\eta$  averaged tagging efficiencies of each jet flavour are calculated for the phase space  
 2345 of interest. Additionally, the relevant jet  $p_T$  and  $\eta$  corrections are applied to correct the  
 2346 measured b-tagging rate in simulation to that of data, as specified in Section (4.5.3).  
 2347 These corrections propagate through to the average determined tagging efficiency for  
 2348 each jet flavour, consequently affecting the final Z0 and Z2 template shape of the  $n_b^{\text{reco}}$   
 2349 distribution, determined within each analysis category ( $H_T$  and  $n_{\text{jet}}$  in the case of the  
 2350  $\alpha_T$  search).

2351 Using the truth-level flavour information of each of the defined Z0 and Z2 templates and  
 2352 the measured tagging efficiencies of each jet flavour, the template shapes are constructed  
 2353 from simulation via the formula method. These two shapes are then fitted to data in a  
 2354 low  $n_b^{\text{reco}}$  control region (0-2), by allowing the normalisation constants  $\theta_{Z0}$  and  $\theta_{Z2}$  of the  
 2355 two templates to float. The fits are performed independently within each of the defined  
 2356 analysis category to remove any dependence on the modelling of jet multiplicity between  
 2357 simulation and data. Best fit values of  $\theta_{Z0}$  and  $\theta_{Z2}$  are used, along with the fixed shape  
 2358 of each template, to extrapolate a **SM** background estimation within the high  $n_b^{\text{reco}}$  signal  
 2359 region (3,4) as shown in Figure 6.2.

2360 Any large excess in data is an indication that the  $n_b^{\text{reco}}$  distribution is not adequately  
 2361 described by the **SM** backgrounds encapsulated by the templates. This could mean there



**Figure 6.2:** An example of a template fit with the defined  $Z_0$  (blue) and  $Z_2$  (red) templates to data within the low  $n_b^{\text{reco}}$  control region (left). The shape of the two templates are fixed but the normalisations  $\theta_{Z_0}$  and  $\theta_{Z_2}$  are allowed to vary. The best fit values are then applied to extrapolate a combined background prediction from the shaded signal region (right), represented by the dashed black line. Statistical and systemic uncertainties are not shown within this figure.

2362 are additional **SM** backgrounds that fall within the selection of the analysis that need to  
 2363 be considered, or that there is signal present within the data. This method relies solely  
 2364 on fitting to the shape of the  $n_b^{\text{reco}}$  distribution, and can in principle, be applied to any  
 2365 analysis where the signal hypothesis has a larger underlying b-quark spectra than the  
 2366 **SM** backgrounds.

2367 However, in the scenario where a **SUSY** signal sits at a low number of underlying b-quarks,  
 2368 the template would be unable to discriminate between this signal and background during  
 2369 the fit in the control region. This will be the case unless the jet  $p_T$  distribution of  
 2370 the signal and background were drastically different, in which case there would anyway  
 2371 be many more sensitive and practical ways to establish the presence of a signal in the  
 2372 data than this method. Indeed the template method is only really applicable to the  
 2373 hypothesis that any signal resides at high  $n_b^{\text{reco}}$  and that the control region  $0 \leq n_b^{\text{reco}} \leq 2$   
 2374 has negligible signal contamination.

## 2375 6.2. Application to the $\alpha_T$ Search

2376 As detailed in the previous chapter, the  $\alpha_T$  analysis is a search for **SUSY** particles  
 2377 in all-hadronic final states, utilising the kinematic variable  $\alpha_T$  to suppress QCD to a

2378 negligible level. **SM** enriched control samples are used to estimate the background within  
2379 an all-hadronic signal region.

2380 The selection for the  $\mu +$  jets control samples defined in Section (4.2.3) is used to  
2381 demonstrate the template fitting procedure both conceptually in simulation, and also  
2382 when applied in data. This is chosen, as such a selection is dominated by events stemming  
2383 from the **SM** processes with little or no signal contamination from potential new physics.  
2384 Contributions from rare **SM** processes with a higher underlying b-quark content (e.g.  
2385  $t\bar{t}b\bar{b}$ ) are also found to be negligible from studies in simulation. For these reasons, there  
2386 is a degree of confidence that the procedure should adequately describe the observations  
2387 in data when extrapolated to the signal region.

2388 As a departure from the  $\alpha_T$  search strategy described in the previous section, events are  
2389 categorised according to jet multiplicity categories of 3, 4 and  $\geq 5$  reconstructed jets  
2390 per event (di-jet events are not included as there is no contribution to the high  $n_b^{reco}$   
2391 region (3,4)), in order to reduce the kinematic range of the jet  $p_T$ 's within each category.  
2392 Furthermore the analysis is split into just three  $H_T$  regions, for the purpose of increasing  
2393 statistics within the control region,

2394 • 275-325 GeV

2395 • 325-375 GeV

2396 •  $> 375$  GeV

2397 contrary to the eight used within the  $\alpha_T$  analysis. Templates for both underlying b-quark  
2398 content hypotheses are then generated for the nine defined event categories.

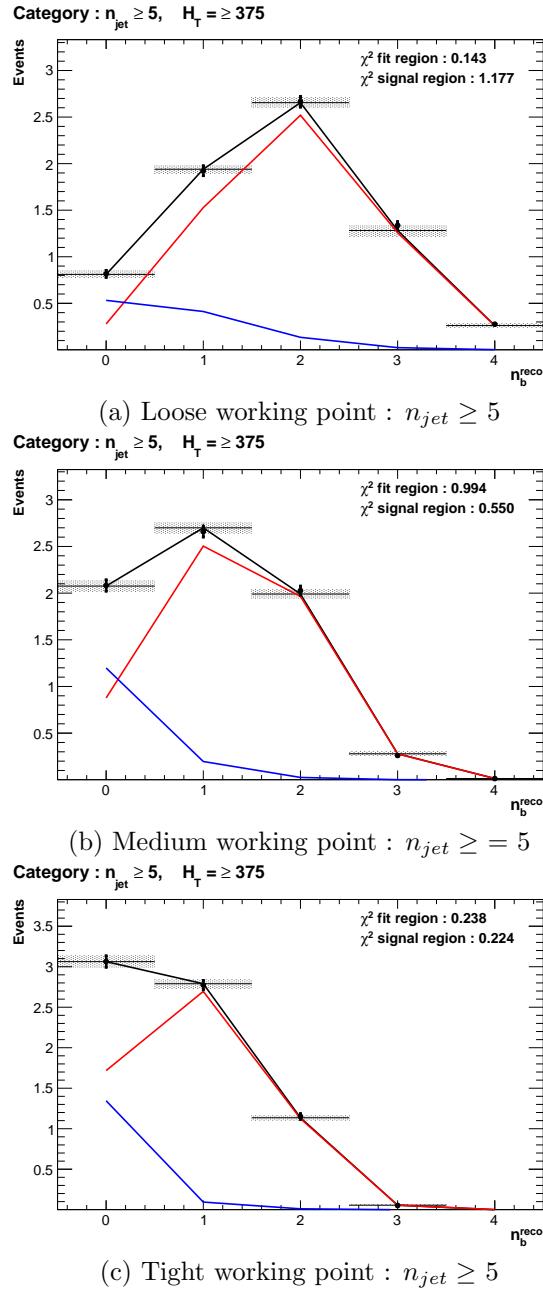
### 2399 6.2.1. Proof of Principle in Simulation

2400 This template procedure must be first demonstrated to work within simulated events free  
2401 from any potential signal contamination before it can be applied to data. By combining  
2402 the relevant ingredients necessary to employ the formula method,  $n_b^{reco}$  shape templates  
2403 are generated individually for each  $n_{jet}$  and  $H_T$  category using one half of the available  
2404 simulated events for each **SM** process. In this case, as the template shapes are being  
2405 fitted to simulation, it is *not* necessary to apply the relevant corrections of the b-tagging  
2406 rates between data and simulation.

2407 The other half of simulated events is utilised to provide a statistically independent sample  
2408 from which the  $n_b^{\text{reco}}$  distribution is taken directly. The two generated templates are  
2409 then fit within the low  $n_b^{\text{reco}}$  (0-2) control region to this pseudo-data, from which a signal  
2410 region prediction is then extrapolated from the template best fit values. The simulated  
2411 event samples in the analysis are split in this way to allow for statistically independent  
2412 fits to be performed.

2413 The aim of this procedure is to ensure that the template fit can accurately extrapolate the  
2414  $n_b^{\text{reco}}$  distribution within the defined signal region from two independent but kinematically  
2415 identical samples. Furthermore, as the pseudo-data of the  $n_b^{\text{reco}}$  distribution is taken  
2416 directly from simulation, observation of good closure for both the initial fit of the two  
2417 templates within the control region and after extrapolation to the signal region will serve  
2418 as a validation of the formula method in recovering the original  $n_b^{\text{reco}}$  distribution itself.

2419 Results are presented in Figure 6.3 for each **CSV** working point in the  $n_{\text{jet}} \geq 5$  category,  
2420 using the  $\mu + \text{jets}$  control sample selection and the inclusive  $H_T > 375$  GeV analysis bin.  
2421 Additional fit results for other  $n_{\text{jet}}$  categories which show a similar level of closure can be  
2422 found within Appendix D.1. The grey bands represent the statistical uncertainty on the  
2423 template shapes as derived from adding in quadrature the statistical uncertainty on the  
2424  $\theta_{Z0}$  and  $\theta_{Z2}$  values from the fit performed, the statistical uncertainty of the measured  
2425 tagging efficiencies for each jet flavour, and the statistical uncertainty of the template  
2426 shapes from the formula method.



**Figure 6.3:** Results of fitting the  $Z = 0$  and  $Z = 2$  templates in the  $n_b^{reco} = 0\text{-}2$  control region to yields from simulation in the  $\mu + \text{jets}$  control sample for the  $H_T > 375$  GeV,  $n_{jet} \geq 5$  category for all CSV working points. Data is represented by the black circles with the blue, red and black lines representing the  $Z=0$ ,  $Z=2$  and combination of both templates respectively. Grey bands represent the uncertainty of the fit. The  $\chi^2$  parameters represent the goodness of fit to the control and signal region.

2427 The extrapolated fit predictions summed over all  $n_{\text{jet}}$  multiplicities within the high  $n_b^{\text{reco}}$   
2428 signal region, are summarised for all  $H_T$  bins and working points in Table 6.2.

$H_T$		275-325	325-375	>375
Loose working point				
Simulation	$n_b = 3$	$793.0 \pm 14.8$	$387.9 \pm 10.2$	$794.1 \pm 14.34$
Template		$820.4 \pm 26.7$	$376.3 \pm 11.9$	$780.1 \pm 15.1$
Simulation	$n_b = 4$	$68.2 \pm 3.9$	$27.6 \pm 2.7$	$91.3 \pm 4.9$
Template		$72.5 \pm 4.7$	$28.3 \pm 2.34$	$84.4 \pm 3.8$
Medium working point				
Simulation	$n_b = 3$	$133.7 \pm 5.7$	$74.5 \pm 4.5$	$164.2 \pm 6.4$
Template		$132.8 \pm 4.8$	$74.5 \pm 3.9$	$159.9 \pm 5.7$
Simulation	$n_b = 4$	$1.6 \pm 0.6$	$0.6 \pm 0.4$	$3.4 \pm 0.9$
Template		$1.8 \pm 0.2$	$1.1 \pm 0.2$	$4.1 \pm 0.4$
Tight working point				
Simulation	$n_b = 3$	$26.9 \pm 2.6$	$13.9 \pm 1.9$	$31.8 \pm 2.9$
Template		$24.7 \pm 1.5$	$13.8 \pm 1.2$	$28.1 \pm 1.5$
Simulation	$n_b = 4$	$0.5 \pm 0.4$	-	-
Template		$0.1 \pm 0.1$	$0.1 \pm 0.1$	$0.2 \pm 0.1$

**Table 6.2:** Summary of the fit predictions in the  $n_b^{\text{reco}}$  signal region after combination of the  $n_{\text{jet}} = 3, = 4, \geq 5$  categories compared against yields taken directly from simulation. The fit predictions are extrapolated from a  $n_b^{\text{reco}} = 0, 1, 2$  control region and simulation yields are normalised to an integrated luminosity of  $10 \text{ fb}^{-1}$ . The uncertainties quoted on the template yields are purely statistical.

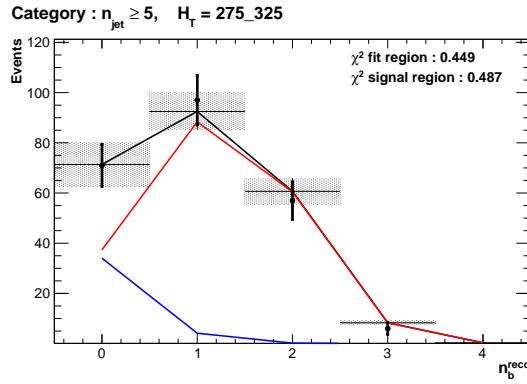
2429 The pull distributions for all the fits performed can be found in Appendix D.2, and are  
2430 compatible with a mean of zero and standard deviation of one, showing no obvious bias  
2431 to the fitting procedure. Each of the fits performed show good compatibility between  
2432 the template shapes and data from simulation within the defined control region, with  
2433 additional good overall agreement also observed for extrapolation to the signal region as  
2434 shown in Table 6.2. This validates both the formula method used in the generation of  
2435 the template shapes as well as the method of predicting the SM background in the high  
2436  $n_b^{\text{reco}}$  signal region.

2437 The application of this method to the same selection in a data control sample is now  
2438 used to demonstrate necessary control over the efficiency and mis-tagging rates when  
2439 b-tagging scale factors are applied, and to test the assumption of no signal contamination  
2440 with the  $\mu + \text{jets}$  control sample.

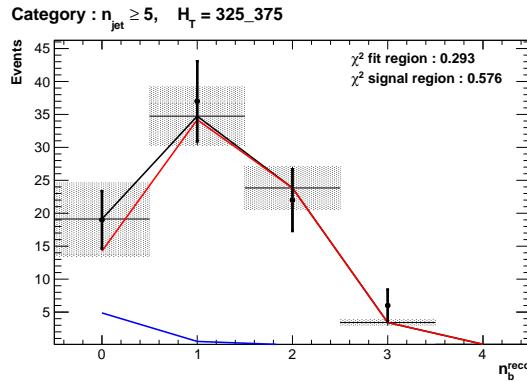
### 2441 6.2.2. Results in a Data Control Sample

2442 The procedure is now applied to the 2012 8 TeV dataset in the  $\mu +$  jets control sample, to  
2443 establish the validity of this method in data. The relevant data to simulation b-tagging  
2444 scale factors are applied to produce corrected values of the efficiency and mis-tagging  
2445 rates within each analysis category [77].

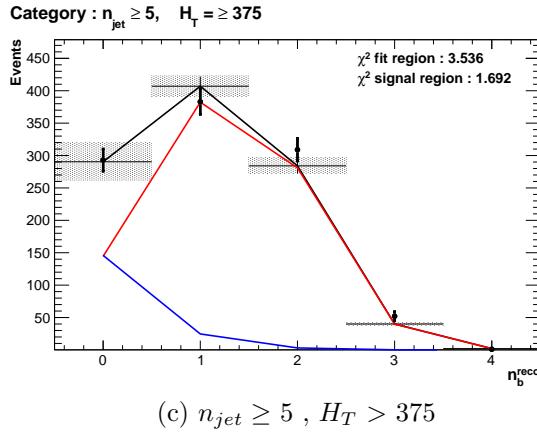
2446 Figure 6.4 shows the results of the templates derived from simulation to each of the three  
2447 defined  $H_T$  bins, in the  $n_{jet} \geq 5$  category for the medium working point **CSV** tagger (the  
2448 same working point used within the  $\alpha_T$  analysis). Grey bands represent the previously  
2449 detailed statistical uncertainty of the fit combined in quadrature with the systematic  
2450 uncertainties of varying the data to simulation scale factors by their b-tag scale factor  
2451 systematic uncertainties. These b-tag scale factor uncertainties are to be taken as fully  
2452 correlated across all jet flavours [77]. Additional fit results for other jet multiplicities are  
2453 found in Appendix D.3.



(a)  $n_{jet} \geq 5, 275 < H_T < 325$



(b)  $n_{jet} \geq 5, 325 < H_T < 375$



**Figure 6.4:** Results of fitting the  $Z = 0$  and  $Z = 2$  templates in the  $n_b^{reco} = 0\text{-}2$  control region to data from the  $\mu + \text{jets}$  control sample, for the CSV medium working point, with a jet multiplicity  $n_{jet} \geq 5$ , in all three  $H_T$  categories. Data is represented by the black circles with the blue, red and black lines representing the  $Z=0$ ,  $Z=2$  and combination of both templates respectively. Grey bands represent the uncertainty of the fit. The  $\chi^2$  parameters represent the goodness of fit to the control and signal region.

2456 The numerical results and extrapolation to the  $n_b^{reco} = 3, 4$  bins for all  $H_T$  and working  
 2457 points, is shown in Table 6.3.

$H_T$	275-325	325-375	>375
Loose working point			
Data	838	394	717
Template	$n_b = 3$	$861.8 \pm 38.1$	$673.2 \pm 34.5$
Data	81	43	81
Template	$n_b = 4$	$78.5 \pm 5.8$	$27.6 \pm 2.6$
Medium working point			
Data	137	79	152
Template	$n_b = 3$	$131.2 \pm 4.3$	$137.8 \pm 5.7$
Data	1	1	3
Template	$n_b = 4$	$1.8 \pm 0.1$	$3.1 \pm 0.2$
Tight working point			
Data	24	15	25
Template	$n_b = 3$	$23.0 \pm 0.9$	$20.3 \pm 1.1$
Data	0	0	1
Template	$n_b = 4$	$0.1 \pm 0.1$	$0.2 \pm 0.1$

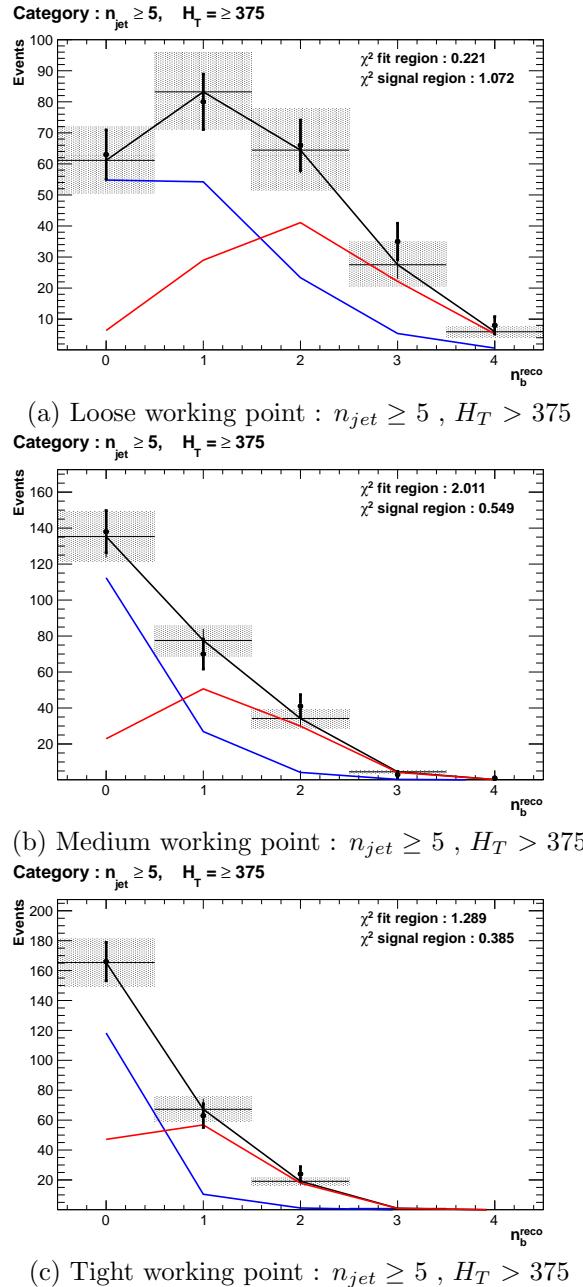
**Table 6.3:** Summary of the fit predictions in the  $n_b^{reco}$  signal region of the  $\mu + \text{jets}$  control sample, after combination of the  $n_{jet} = 3, = 4, \geq 5$  categories. The predictions are extrapolated from a  $n_b^{reco} = 0, 1, 2$  control region using  $11.4 \text{ fb}^{-1}$  of  $\sqrt{s} = 8 \text{ TeV}$  data. The uncertainties quoted on the template yields are a combination of statistical and systematic uncertainties.

When this method is applied to the  $\mu + \text{jets}$  control sample, it is expected that good agreement would be observed between the template predictions and observation in the absence of signal contamination. The good compatibility for all working points as shown in the table, demonstrate that this is the case and that the method is able to accurately predict the background yields. However the assumption of negligible signal contamination can no longer made when applied to the hadronic signal region of the  $\alpha_T$  search, where agreement between estimated backgrounds and observations in data is now not necessarily expected.

### 6.2.3. Application to the $\alpha_T$ Hadronic Search Region

As an accompaniment to the background estimation methods outlined in the  $\alpha_T$  search, the b-tag template method offers a complementary way of testing the SM only background hypothesis within the hadronic signal region of the search. In the presence of a natural SUSY signature mediated by a light gluino and containing four underlying  $\tilde{b}$  or  $\tilde{t}$  squarks, which subsequently decay to t or b quarks, the number of reconstructed  $n_b^{\text{reco}} = 3, \geq 4$  events will be enhanced.

Figure 6.5 shows the results of the template shapes derived from simulation and fitted to data for each of the three CSV working points, in the  $n_{\text{jet}} \geq 5, H_T > 375$  GeV category. Grey bands represent the statistical uncertainty of the fit combined in quadrature with the systematic uncertainties of varying the data to simulation scale factors up and down by their measured systematic uncertainties. Additional fit results for other jet multiplicities are found in Appendix D.4.



**Figure 6.5:** Results of fitting the  $Z = 0$  and  $Z = 2$  templates in the  $n_b^{reco} = 0\text{-}2$  control region to data from the hadronic signal selection, in the  $n_{jet} \geq 5$  and  $H_T > 375$  category for all CSV working points. Data is represented by the black circles with the blue, red and black lines representing the  $Z=0$ ,  $Z=2$  and combination of both templates respectively. Grey bands represent the uncertainty of the fit. The  $\chi^2$  parameters represent the goodness of fit to the control and signal region.

<sup>2479</sup> The numerical results and extrapolation to the  $n_b^{reco} = 3, 4$  bins for all  $H_T$  and working  
<sup>2480</sup> points are shown in Table 6.4. Also included within the table are total SM background  
<sup>2481</sup> predictions determined by the maximum likelihood fit in both jet multiplicity categories

2482 of the  $\alpha_T$  analysis for the **CSVM** tagger, as introduced in Section (4.8). No excess of  
2483 data is found for any of the three **CSV** working points. Predictions within the signal  
2484 region from this method are also found to be compatible with the background predictions  
2485 determined by the  $\alpha_T$  simultaneous fit as already shown in Table 5.2.

$H_T$		275-325	325-375	>375
Loose working point				
Data		198	85	126
Template	$n_b = 3$	$207.1 \pm 33.3$	$103.4 \pm 10.9$	$124.98 \pm 16.2$
Data		15	9	16
Template	$n_b = 4$	$15.9 \pm 3.7$	$8.05 \pm 1.2$	$13.1 \pm 2.2$
Medium working point				
Data		32	16	15
Template	$n_b = 3$	$26.4 \pm 1.7$	$12.7 \pm 1.2$	$19.9 \pm 2.8$
$\alpha_T$ ML Fit		$29.8^{+5.2}_{-4.4}$	$14.0^{+1.8}_{-2.0}$	$16.5^{+1.4}_{-1.4}$
Data		1	0	2
Template	$n_b = 4$	$0.3 \pm 0.2$	$0.3 \pm 0.1$	$0.5 \pm 0.2$
$\alpha_T$ ML Fit		$0.9^{+0.4}_{-0.7}$	$0.3^{+0.2}_{-0.2}$	$0.6^{+0.3}_{-0.3}$
Tight working point				
Data		5	2	0
Template	$n_b = 3$	$4.03 \pm 0.3$	$2.4 \pm 0.3$	$3.1 \pm 0.3$
Data		1	0	0
Template	$n_b = 4$	$0.1 \pm 0.1$	$0.1 \pm 0.1$	$0.0 \pm 0.1$

**Table 6.4:** Summary of the fit predictions in the  $n_b^{reco}$  signal region of the  $\alpha_T$  hadronic signal selection, after combination of the  $n_{jet} = 3, = 4, \geq 5$  categories. The predictions are extrapolated from a  $n_b^{reco} = 0, 1, 2$  control region using  $11.7 \text{ fb}^{-1}$  of  $\sqrt{s} = 8\text{TeV}$  data. The uncertainties quoted on the template yields are a combination of statistical and systematic uncertainties.

### 2486 6.3. Summary

2487 A **SUSY** signature such as one from gluino-induced third-generation squark production,  
2488 would result in a final state with an underlying b-quark content greater than two. In  
2489 order to be able to discriminate such signatures from the **SM** background, templates are  
2490 generated based on a parameterisation of **SM** processes, where the underlying b-quarks  
2491 per event is typically zero or two. These templates are then fit to data in a low  $n_b^{reco}$   
2492 (0-2) control region in order to extrapolate a prediction within a high  $n_b^{reco}$  (3-4) signal  
2493 region. This approach is built upon the assumptions that the defined control region is  
2494 almost entirely free of any possible signal contamination from possible signal topologies  
2495 with a small number of b quarks in the final state.

2496 The method was demonstrated both in simulation and also in data, using the **SM**  
2497 enriched  $\mu + \text{jets}$  selection from the  $\alpha_T$  search. This was conducted to prove conceptually  
2498 and experimentally that the method is valid and that there is adequate control over  
2499 the efficiency and mis-tagging rates in data for all working points of the **CSV** tagger.  
2500 Additionally this method was further applied to the  $\alpha_T$  analysis signal region, where  
2501 good agreement is observed between the predictions from the template extrapolations,  
2502 observations in data and the background estimation method of the  $\alpha_T$  analysis.

# Chapter 7.

## 2503 Conclusions

2504 A search for supersymmetry is presented based on a data sample of pp collisions collected  
2505 at  $\sqrt{s} = 8$  TeV, corresponding to an integrated luminosity of  $11.7 \pm 0.5$  fb $^{-1}$ . Final states  
2506 with two or more jets and significant missing transverse energy, a typical final state  
2507 topology of R-parity conserving SUSY models have been analysed. The  $\alpha_T$  variable is  
2508 utilised as the main discriminator between balanced multi-jet backgrounds and those  
2509 with real missing transverse energy.

2510 Within the search presented, Standard Model (SM) backgrounds are estimated from  
2511 a simultaneous binned likelihood fit to a hadronic signal selection as well as three SM  
2512 process enriched control samples. The search is split into total transverse hadronic  
2513 energy ( $H_T$ ), jets identified as originating for a b-quark ( $n_b^{\text{reco}}$ ), and jet multiplicity  
2514 ( $n_{\text{jet}}$ ) categories to improve sensitivity to a range of possible supersymmetric final states.  
2515 Systematic errors due to theory, detector effects and simulation deficiencies are quantified  
2516 through the use of data driven closure tests and accounted for in the final interpretation.  
2517 Observations in data are found to be compatible with a SM only hypothesis.

2518 In the absence of a signal like excess the analysis is further interpreted in a set of  
2519 Simplified Model Spectra (SMS) models, representing a set of model independent decay  
2520 topologies parameterised only by the production process and the masses of their parent  
2521 sparticle and Lightest Supersymmetric Partner (LSP). In models mediated by gluino pair  
2522 production and containing a large mass difference between the gluino and LSP, exclusion  
2523 limits of the gluino mass are set in the range 950-1125 GeV. For SMS models describing  
2524 direct squark pair production, first or second generation squarks are excluded up to 775  
2525 GeV, with direct bottom squarks production excluded up to masses of 600 GeV.

2526 In the case of gluino mediated third generation signatures containing many jets originating  
2527 from b-quarks in the final state, mass limits are set in the range of 975-1125 GeV for  
2528 large mass splittings between the gluino and the LSP. The experimental sensitivity  
2529 to these models is attributed to the  $n_b^{\text{reco}}$  categorisation of the analysis, where the  
2530 signal-to-background is enhanced within the phase space of the search at high  $n_b^{\text{reco}}$ .

2531 Furthermore, a measurement of the performance of the Level-1 trigger for jets and energy  
2532 sum quantities has also been presented. These studies quantify any change in Level-1  
2533 performance after the introduction of a 5 GeV jet seed threshold into the jet clustering  
2534 algorithm. No significant change in single jet trigger efficiencies is observed and good  
2535 performance is observed for a range of Level-1 energy sum quantities.

2536 This change was introduced to facilitate a reduction in the rate of events triggered by  
2537 energy deposits due to soft non-collimated jets from secondary interactions, and which are  
2538 not of interest to physics analyses. This was necessary to ensure, that trigger thresholds  
2539 can be maintained at low values in the presence of an ever increasing number of bunch  
2540 crossings per proton interaction. In the context of SUSY, this is a necessity to keep  
2541 CMS sensitive to types of compressed spectra signatures characterised by low transverse  
2542 energy jets and small missing transverse energy signatures.

2543 Finally, an approach that uses a template fit method to the  $n_b^{\text{reco}}$  distribution of SM  
2544 processes within a supersymmetric search is introduced and then validated in simulation  
2545 and data. The approach can be used to identify any excess in data arising from gluino  
2546 mediated third generation supersymmetric signatures. It is utilised within this thesis as  
2547 a crosscheck to the  $\alpha_T$  background prediction at high b-tagged jet multiplicities. This  
2548 method is found to give a SM background estimation that is in good agreement with the  
2549  $\alpha_T$  search within the hadronic signal region.

2550 The continued absence of a supersymmetric signal in the  $\alpha_T$  search or other analyses  
2551 at CMS [101][102][103], puts pressure on the parameter space in which SUSY can reside.  
2552 Indeed the smoking gun that many theorists and experimentalists hoped to see at the  
2553 LHC has not materialised. Instead identifying a SUSY signal may now only result  
2554 from many years of data taking and the incorporation of increasingly advanced analysis  
2555 techniques. An unenviable task considering the difficulties of not knowing where SUSY  
2556 may reside, but perhaps solace can be taken in remembering that nothing worth having  
2557 ever comes easy.

# Appendix A.

## <sup>2559</sup> Miscellaneous

### <sup>2560</sup> A.1. Jet Identification Criteria

<sup>2561</sup> For Calo jets the following identification criteria were applied:

Loose CaloJet Id	
Variable	Definition
$f_{HPD} < 0.98$	Fraction of jet energy contributed from “hottest” <b>HPD</b> , which rejects <b>HCAL</b> noise.
$f_{EM} > 0.01$	Noise from the <b>HCAL</b> is further suppressed by requiring a minimal electromagnetic component to the jet $f_{EM}$ .
$N_{hits}^{90} \geq 2$	Jets that have $> 90\%$ of its energy from a single channel are rejected, to serve as a safety net that catches jets arising from undiagnosed noisy channels.

**Table A.1:** Criteria for a reconstructed jet to pass the loose calorimeter jet id.

<sup>2562</sup> For PF jets the following identification criteria were applied:

---

Loose PF jet Id	
Variable	Definition
<code>nfhJet &lt; 0.99</code>	Fraction of jet composed of neutral hadrons. <b>HCAL</b> noise tends to populate high values of neutral hadron fraction.
<code>nemfJet &lt; 0.99</code>	Fraction of jet composed of neutral electromagnetic energy. <b>ECAL</b> noise tends to populate high values of neutral EM fraction.
<code>nmultiJet &gt; 1</code>	Number of constituents that jet is composed from.
<code>chfJet &gt; 0</code>	Fraction of jet composed of charged hadrons.
<code>cmultiJet &gt; 0</code>	Number of charged particles that compose jet.
<code>cemfJet &lt; 0.99</code>	Fraction of jet composed of charged electromagnetic energy.

---

**Table A.2:** Criteria for a reconstructed jet to pass the loose PF jet id.

2563 **A.2. Primary Vertices**

2564 The pile-up per event is defined by the number of 'good' reconstructed primary vertices  
2565 in the event, with each vertex satisfying the following requirements:

---

Good primary vertex requirement	
Variable	Definition
$N_{dof} > 4$	The number of degree of freedom, from the vertex fit to compute the best estimate of the vertex parameters.
$ \Delta z_{vtx}  < 24\text{cm}$	The distance, $ \Delta z_{vtx} $ , to the position of the closest <b>HLT</b> primary vertex.
$\rho < 2\text{cm}$	The perpendicular distance of track position to the beam spot.

---

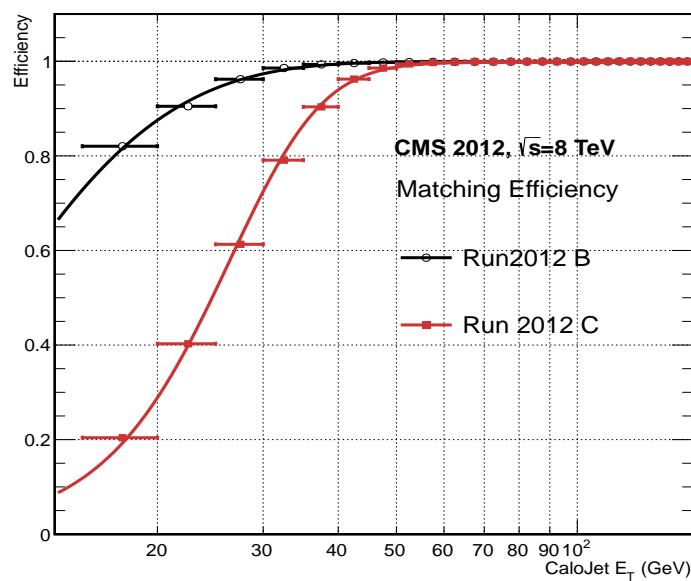
**Table A.3:** Criteria for a vertex in an event to be classified as a 'good' reconstructed primary vertex.

# Appendix B.

## <sup>2566</sup> L1 Jets

### <sup>2567</sup> B.1. Jet Matching Efficiencies

<sup>2568</sup> The single jet turn-on curves are derived from events independent of whether the leading  
<sup>2569</sup> jet in an event is matched to a Level 1 jet using  $\Delta R$  matching detailed in Section (3.4.3).  
<sup>2570</sup> These turn-ons are produced from events which are not triggered on jet quantities and  
<sup>2571</sup> therefore it is not guaranteed that the lead jet of an event will be seeded by a Level 1 jet.  
<sup>2572</sup> Figure B.1 shows the particular matching efficiency of a lead jet to a L1 jet.



**Figure B.1:** Leading jet matching efficiency as a function of the offline CaloJet  $E_T$ , measured in an isolated muon triggered dataset in the 2012B and 2012C run periods.

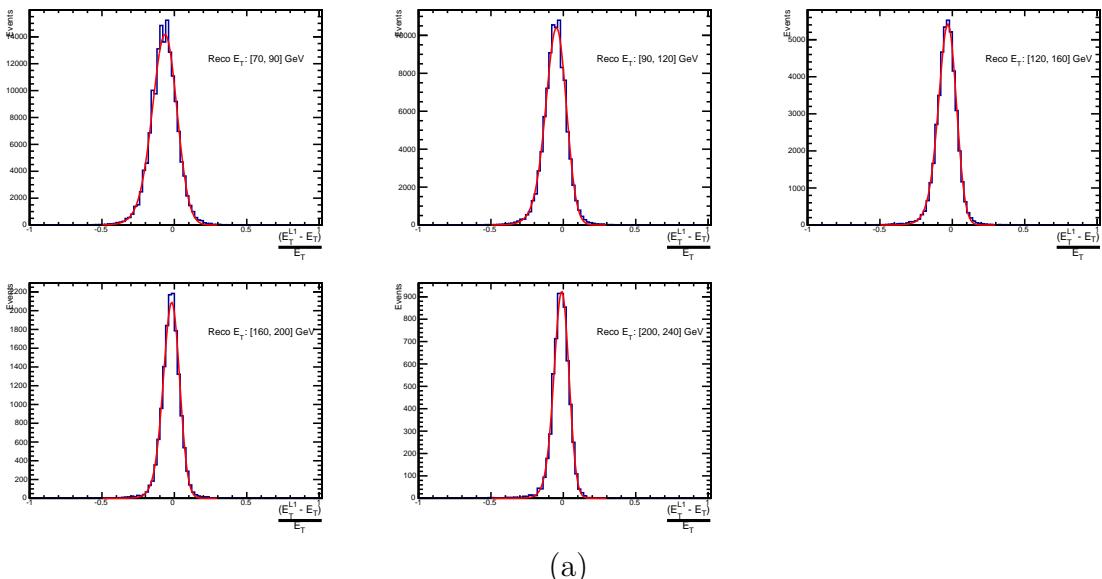
Run Period	$\mu$	$\sigma$
2012B	$6.62 \pm 0.01$	$0.79 \pm 0.03$
2012C	$19.51 \pm 0.03$	$7.14 \pm 0.02$

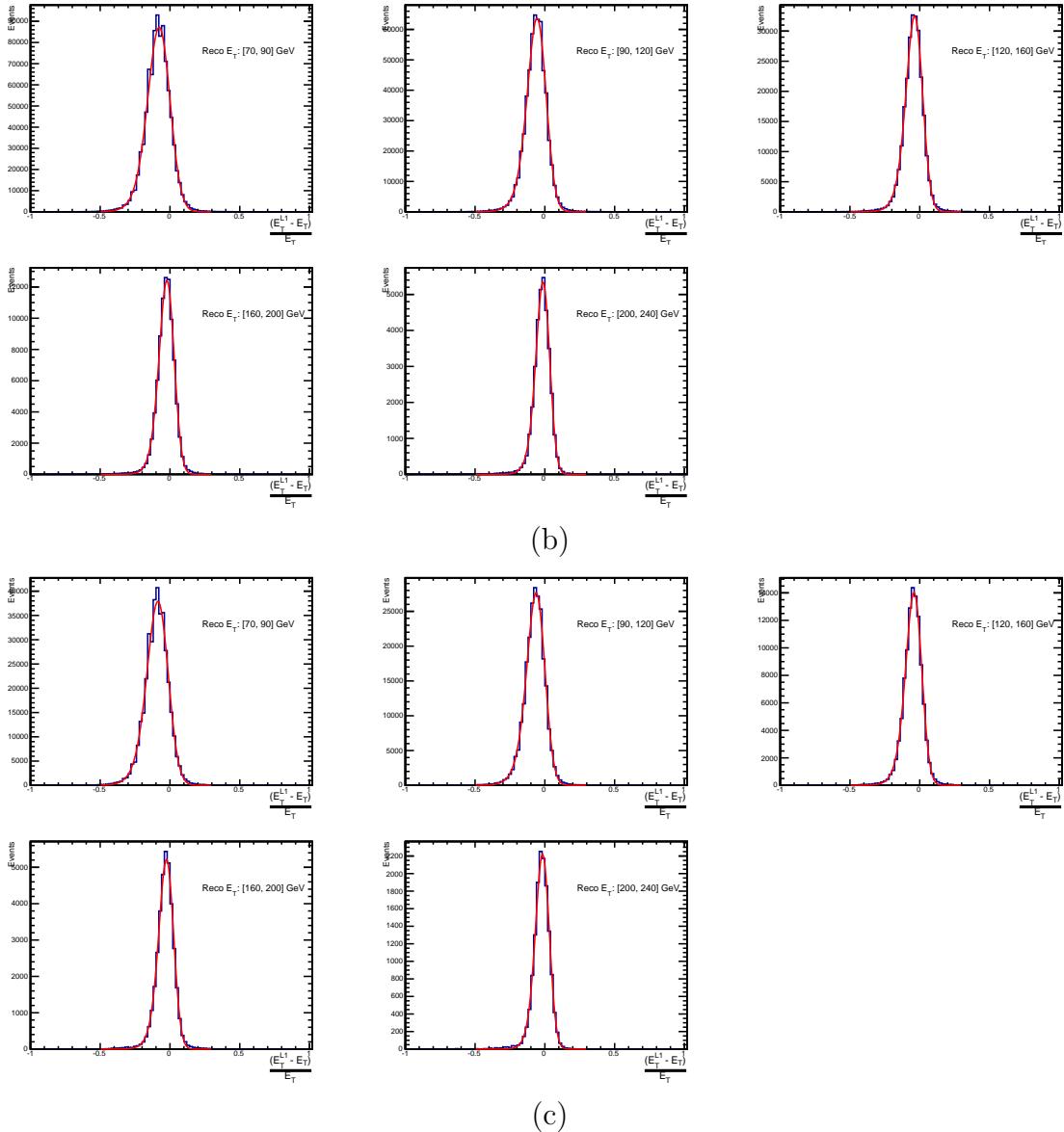
**Table B.1:** Results of a cumulative EMG function fit to the turn-on curves for the matching efficiency of the leading jet in an event to a Level-1 jet in run 2012C and 2012B data, measured in an isolated muon triggered sample. The turn-on point,  $\mu$ , and resolution,  $\sigma$ , are measured with respect to offline Calo Jet  $E_T$ .

2573 It can be seen that the turn-on occurs at a lower  $E_T$  during the 2012B run period. The  
 2574 seed threshold requirement of a 5 GeV jet seed in run 2012C result in more events in  
 2575 which the lead offline jet does not have an associated L1 jet. This can be attributed to  
 2576 events with soft non-collimated jets in which the energy deposits are not centralised in a  
 2577 calorimeter region. However, for larger jet  $E_T$  thresholds typical of those used by physics  
 2578 analyses, 100% efficiency is observed, and therefore this effect has no impact to overall  
 2579 physics performance.

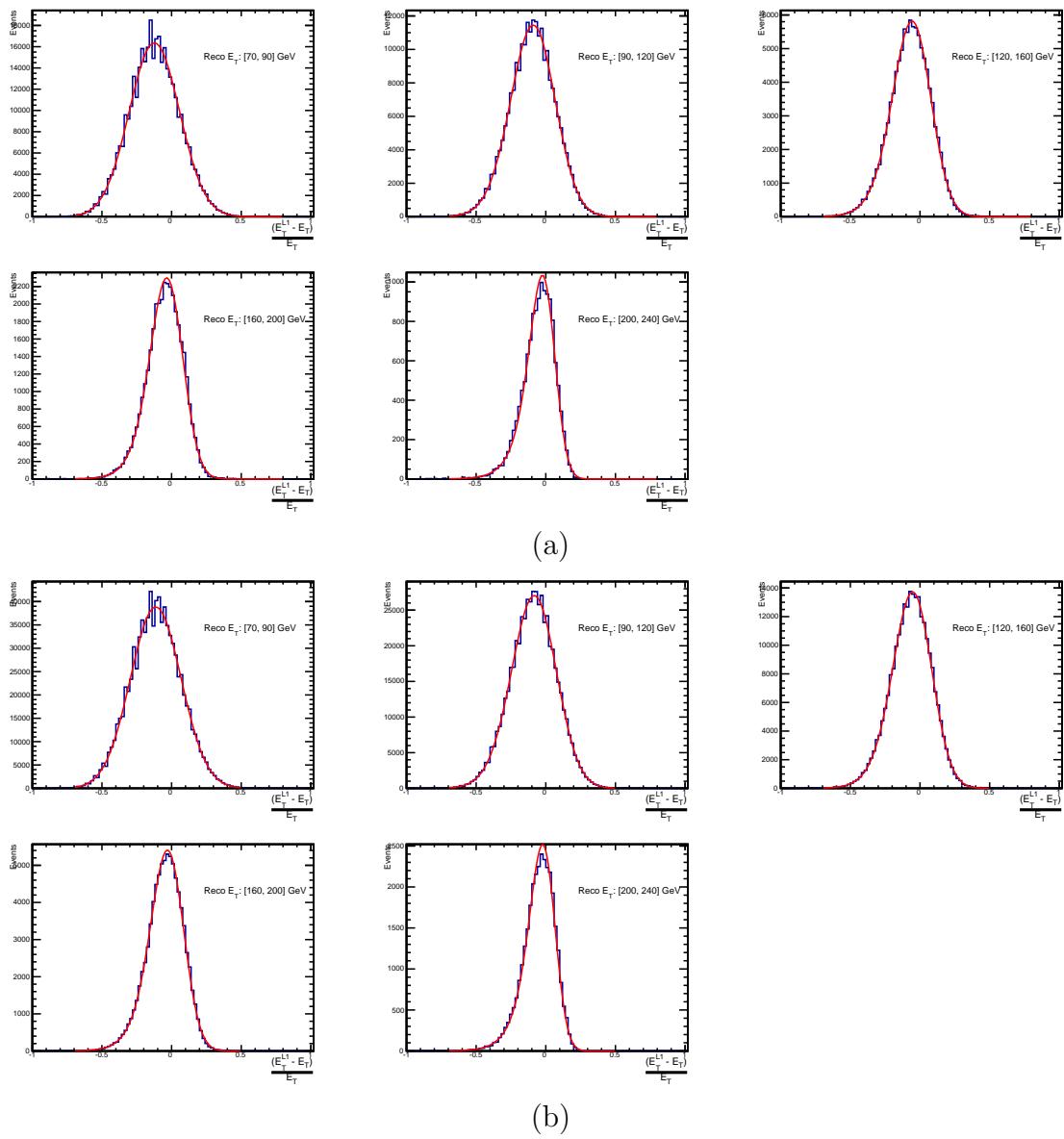
2580 A fit of an **EMG** function to the matching efficiencies find mean,  $\mu$ , values of 6.62 GeV  
 2581 and 19.51 GeV for Run 2012B and 2012C respectively and is shown in Table B.1.

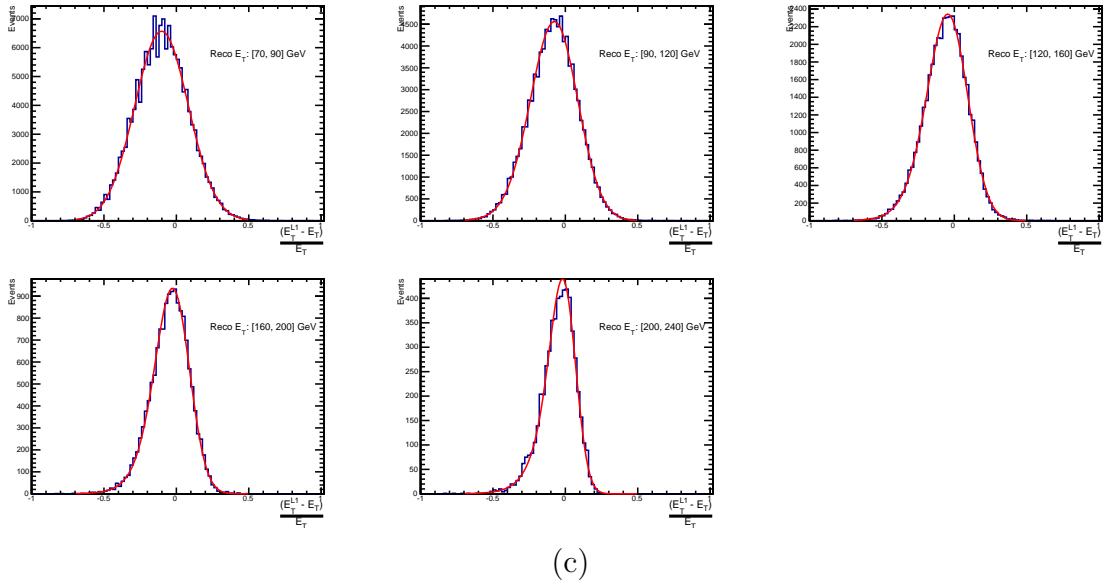
## 2582 B.2. Leading Jet Energy Resolution





**Figure B.2:** Resolution plots of the leading offline jet Calo  $E_T$  measured as a function of  $\frac{(L1 E_T - \text{Offline } E_T)}{\text{Offline } E_T}$  for (a) low, (b) medium, and (c) high pile-up conditions as defined in Section (3.4.4).



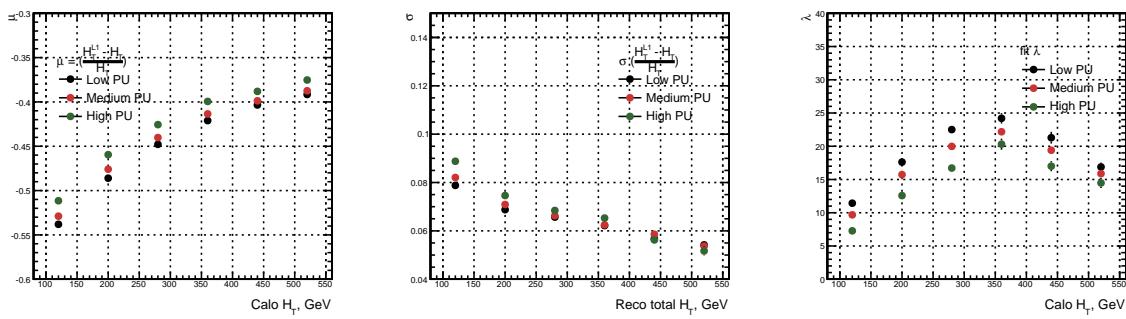


(c)

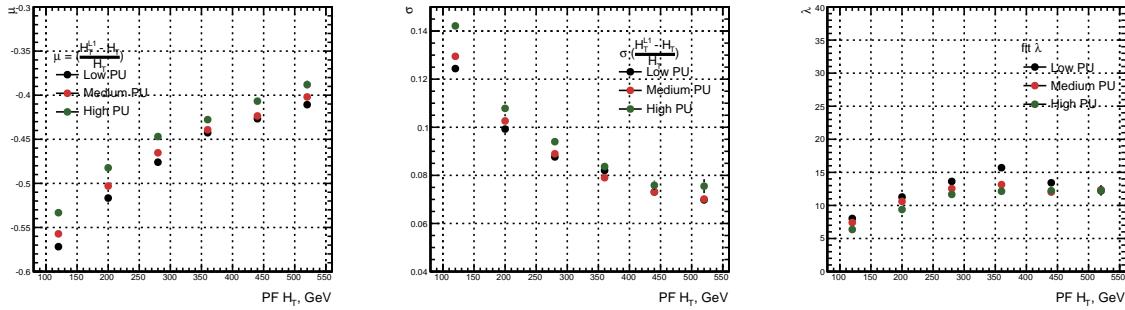
**Figure B.3:** Resolution plots of the leading offline jet PF  $E_T$  measured as a function of  $\frac{(L1 E_T - \text{Offline } E_T)}{\text{Offline } E_T}$  for (a) low, (b) medium, and (c) high pile-up conditions as defined in Section (3.4.4).

### 2583 B.3. Resolution for Energy Sum Quantities

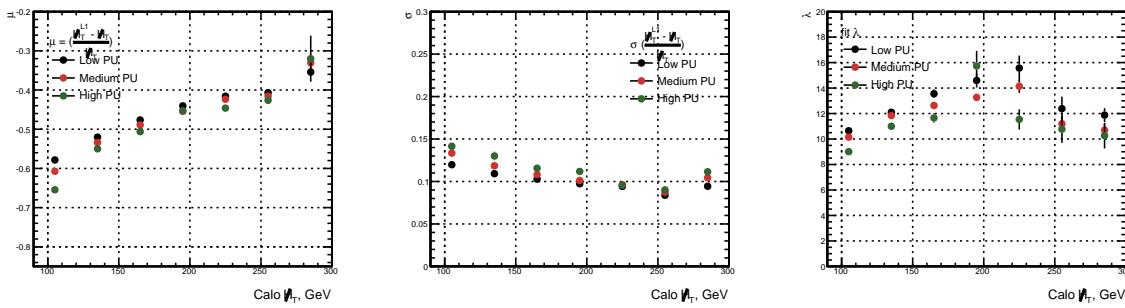
2584 The following plots show the resolution parameters for energy sum quantities as a function  
 2585 of the quantity ( $q$ ) itself. In this case, the  $\mu$ ,  $\sigma$  and  $\lambda$  fit values to an **EMG** function  
 2586 defined by Equation (3.3) for each of the individual  $\frac{(L1 q - \text{Offline } q)}{\text{Offline } q}$  distributions, in bins  
 2587 of the quantity ( $q$ ) is displayed.



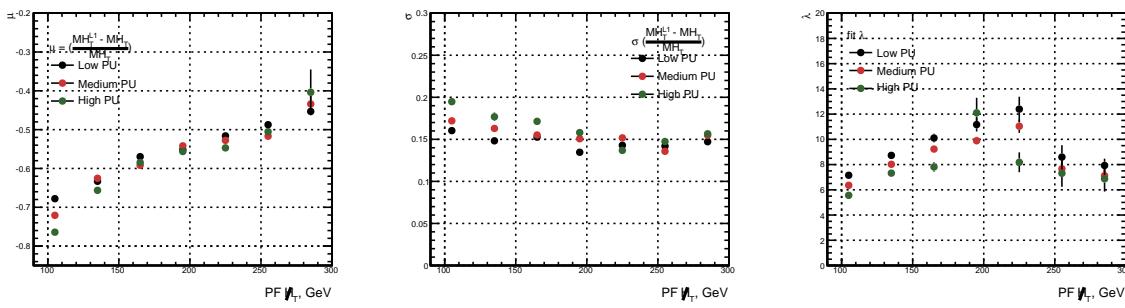
**Figure B.4:**  $H_T$  resolution parameters in bins of Calo  $H_T$  measured for the defined low, medium and high pile-up conditions. Shown are the mean  $\mu$  (left), resolution  $\sigma$  (middle) and  $\lambda$  (right) fit values to an **EMG** function for the  $\frac{(L1 H_T - H_T)}{H_T}$  distributions.



**Figure B.5:**  $H_T$  resolution parameters in bins of  $PF H_T$  measured for the defined low, medium and high pile-up conditions. Shown are the mean  $\mu$  (left), resolution  $\sigma$  (middle) and  $\lambda$  (right) fit values to an EMG function for the  $\frac{(L1H_T - H_T)}{H_T}$  distributions.



**Figure B.6:**  $H_T$  resolution parameters in bins of  $Calo H_T$  measured for the defined low, medium and high pile-up conditions. Shown are the mean  $\mu$  (left), resolution  $\sigma$  (middle) and  $\lambda$  (right) fit values to an EMG function for the  $\frac{(L1H_T - H_T)}{H_T}$  distributions.

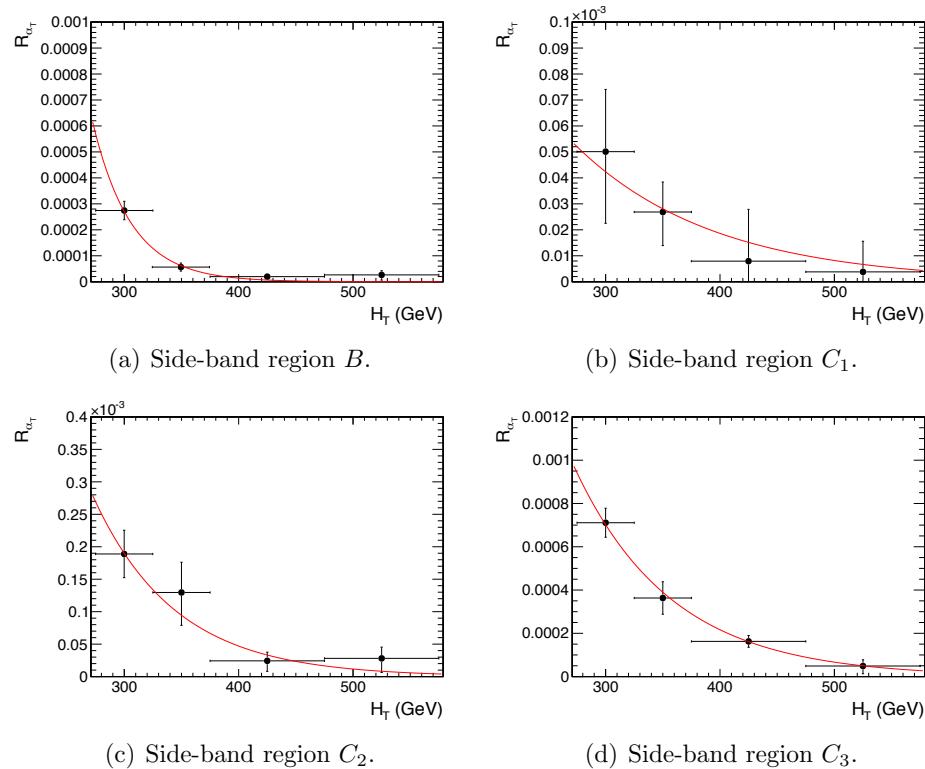


**Figure B.7:**  $H_T$  resolution parameters in bins of  $PF H_T$  measured for the defined low, medium and high pile-up conditions. Shown are the mean  $\mu$  (left), resolution  $\sigma$  (middle) and  $\lambda$  (right) fit values to an EMG function for the  $\frac{(L1H_T - H_T)}{H_T}$  distributions.

# Appendix C.

## <sup>2588</sup> Additional Material on Background <sup>2589</sup> Estimation Methods

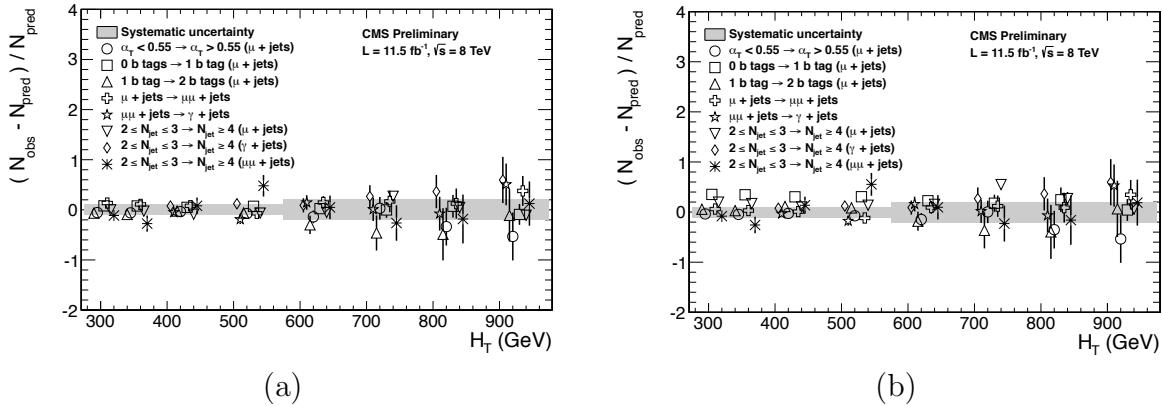
### <sup>2590</sup> C.1. Determination of $k_{QCD}$



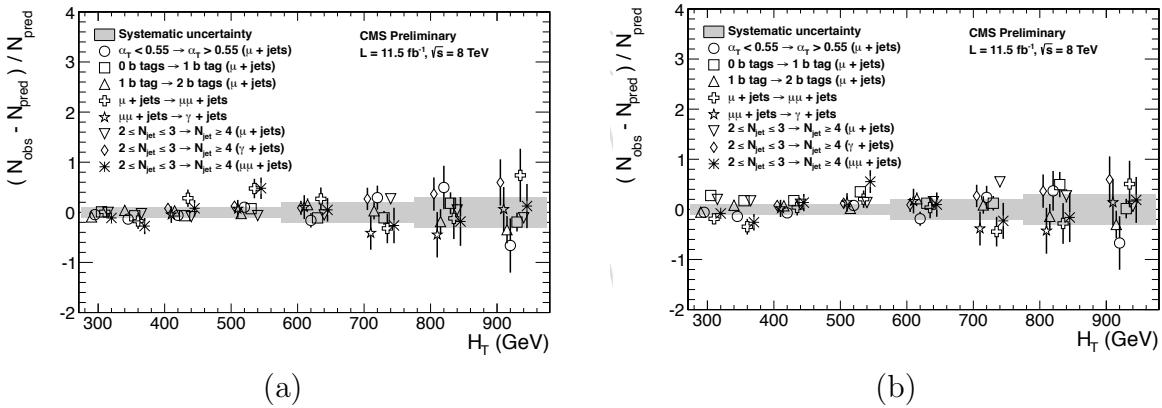
**Figure C.1:**  $R_{\alpha_T}(H_T)$  and exponential fits for each of the data sideband regions. Fit is conducted between the  $H_T$  region  $275 < H_T < 575$ .

2592 **C.2. Effect of Varying Background Cross-sections on**  
2593 **Closure Tests**

2594 Closure tests with cross section variations of +20% and -20% applied to  $W + \text{jets}$  and  $t\bar{t}$   
2595 processes respectively.



**Figure C.2:** Sets of closure tests (open symbols) overlaid on top of the systematic uncertainty used for each of the five  $H_T$  regions (shaded bands) in the  $2 \leq n_{jet} \leq 3$  jet multiplicity category for nominal and varied cross-sections; (a) Nominal and (b) Varied  $\pm 20\%$ .



**Figure C.3:** Sets of closure tests (open symbols) overlaid on top of the systematic uncertainty used for each of the five  $H_T$  regions (shaded bands) in the  $n_{jet} \geq 4$  jet multiplicity category for nominal and varied cross-sections; (a) Nominal (b) Varied  $\pm 20\%$ .

		$H_T$ (GeV)			
$n_b^{reco}$	Cross Section	275–325	325–375	375–475	475–575
0	Nominal	0.303 $\pm$ 0.010	0.258 $\pm$ 0.007	0.192 $\pm$ 0.003	0.148 $\pm$ 0.004
	Varied	0.300 $\pm$ 0.010	0.256 $\pm$ 0.007	0.191 $\pm$ 0.003	0.147 $\pm$ 0.004
1	Nominal	0.294 $\pm$ 0.005	0.246 $\pm$ 0.004	0.189 $\pm$ 0.003	0.139 $\pm$ 0.003
	Varied	0.295 $\pm$ 0.006	0.248 $\pm$ 0.004	0.191 $\pm$ 0.003	0.140 $\pm$ 0.003
2	Nominal	0.208 $\pm$ 0.003	0.183 $\pm$ 0.004	0.145 $\pm$ 0.003	0.123 $\pm$ 0.004
	Varied	0.211 $\pm$ 0.004	0.185 $\pm$ 0.004	0.147 $\pm$ 0.003	0.124 $\pm$ 0.004
3	Nominal	0.214 $\pm$ 0.005	0.202 $\pm$ 0.007	0.159 $\pm$ 0.006	0.140 $\pm$ 0.007
	Varied	0.215 $\pm$ 0.005	0.203 $\pm$ 0.007	0.159 $\pm$ 0.006	0.140 $\pm$ 0.007
$\geq 4$	Nominal	0.220 $\pm$ 0.015	0.245 $\pm$ 0.035	0.119 $\pm$ 0.009	-
	Varied	0.220 $\pm$ 0.015	0.245 $\pm$ 0.035	0.119 $\pm$ 0.009	-
$n_b^{reco}$	Cross Section	575–675	675–775	775–875	875– $\infty$
0	Nominal	0.119 $\pm$ 0.004	0.098 $\pm$ 0.005	0.077 $\pm$ 0.006	0.049 $\pm$ 0.005
	Varied	0.120 $\pm$ 0.005	0.098 $\pm$ 0.006	0.077 $\pm$ 0.007	0.049 $\pm$ 0.005
1	Nominal	0.115 $\pm$ 0.004	0.093 $\pm$ 0.005	0.075 $\pm$ 0.007	0.063 $\pm$ 0.006
	Varied	0.116 $\pm$ 0.004	0.098 $\pm$ 0.005	0.081 $\pm$ 0.007	0.065 $\pm$ 0.006
2	Nominal	0.096 $\pm$ 0.005	0.070 $\pm$ 0.006	0.051 $\pm$ 0.007	0.063 $\pm$ 0.008
	Varied	0.098 $\pm$ 0.005	0.073 $\pm$ 0.006	0.053 $\pm$ 0.007	0.064 $\pm$ 0.008
3	Nominal	0.114 $\pm$ 0.009	0.065 $\pm$ 0.007	0.070 $\pm$ 0.017	0.092 $\pm$ 0.020
	Varied	0.114 $\pm$ 0.009	0.066 $\pm$ 0.007	0.070 $\pm$ 0.016	0.093 $\pm$ 0.020

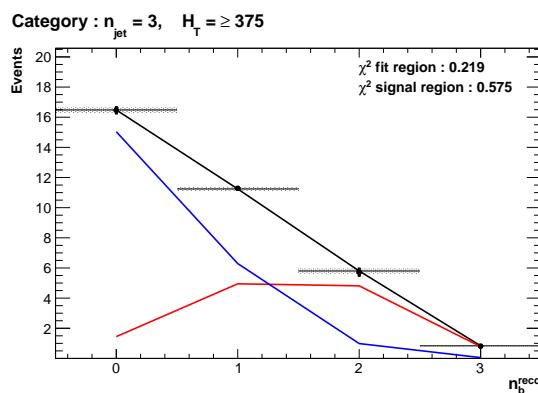
**Table C.1:** Translation factors constructed from the  $\mu +$  jets control sample and signal selection MC, to predict yields for the  $W +$  jets and  $t\bar{t}$  back-grounds in the signal region with (a) NNLO cross sections corrected by k-factors determined from a data sideband see Section (4.4), marked as Nominal, and (b) the same cross sections but with those for  $W +$  jets and  $t\bar{t}$  varied up and down by 20%, respectively, marked as Varied. No requirement is placed on the jet multiplicity of events within this table.

## Appendix D.

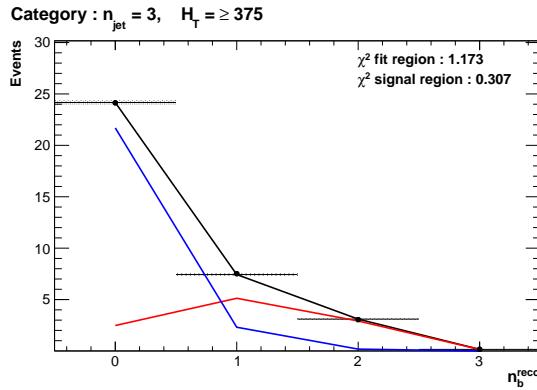
### 2596 Additional Material For B-tag 2597 Template Method

#### 2598 D.1. Templates Fits in Simulation

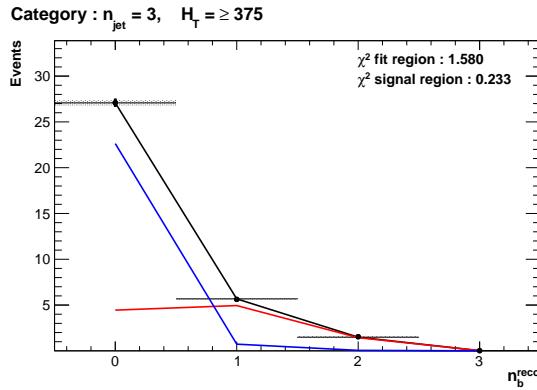
2599 The result of template fits for the three **CSV** working points in the  $n_{jet} = 3, H_T > 375$   
2600 category:



(a) Loose working point  $n_{jet} = 3$



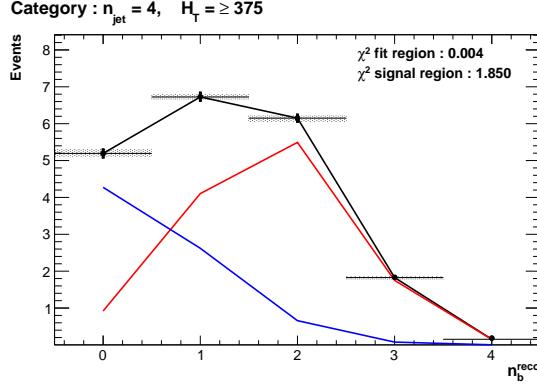
(b) Medium working point  $n_{jet} = 3$



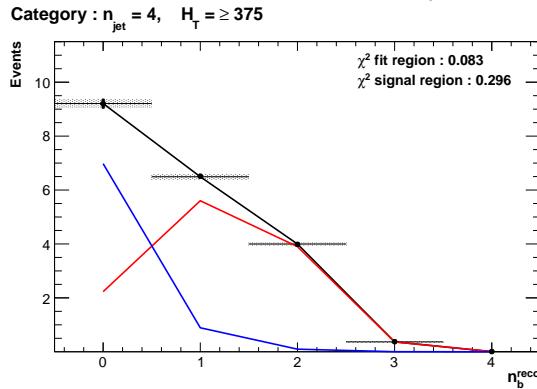
(c) Tight working point  $n_{jet} = 3$

**Figure D.1:** Results of fitting the  $Z = 0$  and  $Z = 2$  templates in the  $n_b^{reco} = 0\text{-}2$  control region to yields from simulation in the  $\mu + \text{jets}$  control sample for the  $H_T > 375$  GeV,  $n_{jet} = 3$  category. Data is represented by the black circles with the blue, red and black lines representing the  $Z=0$ ,  $Z=2$  and combination of both templates respectively. Grey bands represent the uncertainty of the fit. The  $\chi^2$  parameter represent the goodness of fit to the control and signal region.

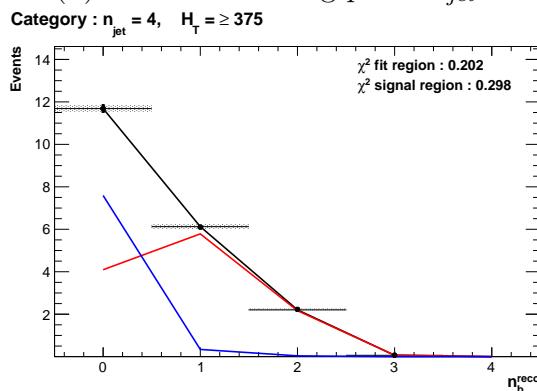
2601 Template fits for the three **CSV** working points in the  $n_{jet} = 4$ ,  $H_T > 375$  category:



(a) Loose working point  $n_{jet} = 4$



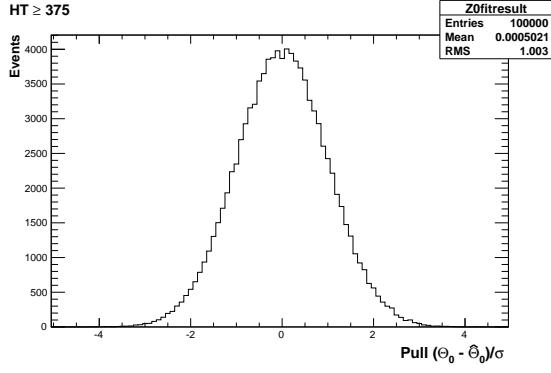
(b) Medium working point  $n_{jet} = 4$



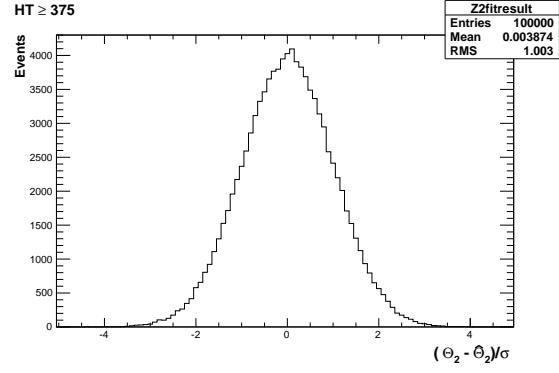
(c) Tight working point  $n_{jet} = 4$

**Figure D.2:** Results of fitting the  $Z = 0$  and  $Z = 2$  templates in the  $n_b^{reco} = 0\text{-}2$  control region to yields from simulation in the  $\mu + \text{jets}$  control sample for the  $H_T > 375$  GeV,  $n_{jet} = 4$  category. Data is represented by the black circles with the blue, red and black lines representing the  $Z=0$ ,  $Z=2$  and combination of both templates respectively. Grey bands represent the statistical uncertainty of the fit. The  $\chi^2$  parameters represent the goodness of fit to the control and signal region.

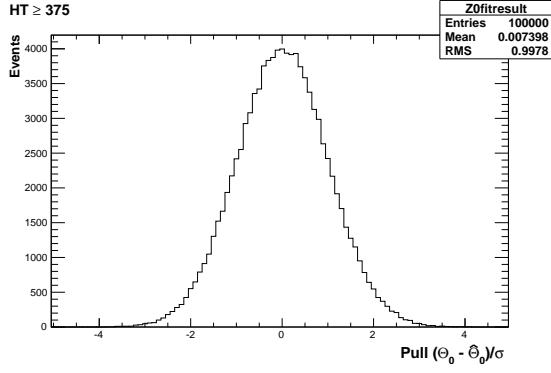
2602 D.2. Pull Distributions for Template Fits



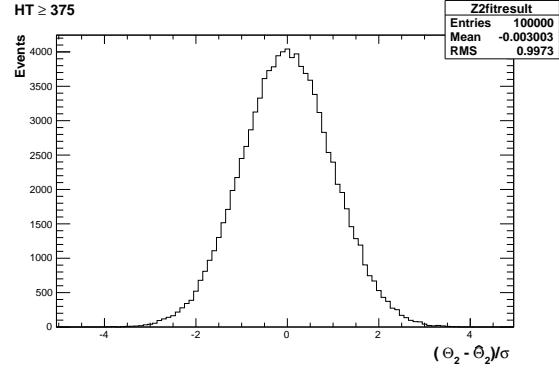
(a) Z0 Template,  $H_T > 375$ ,  $n_{jet} = 3$



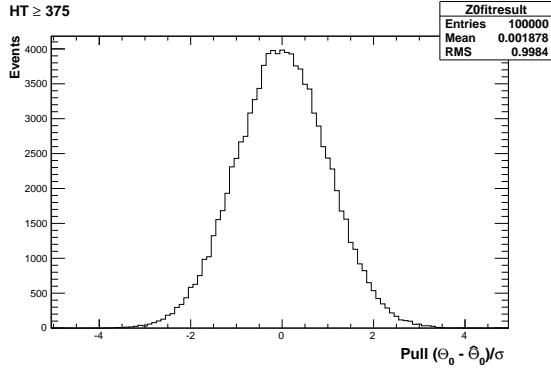
(b) Z2 Template,  $H_T > 375$ ,  $n_{jet} = 3$



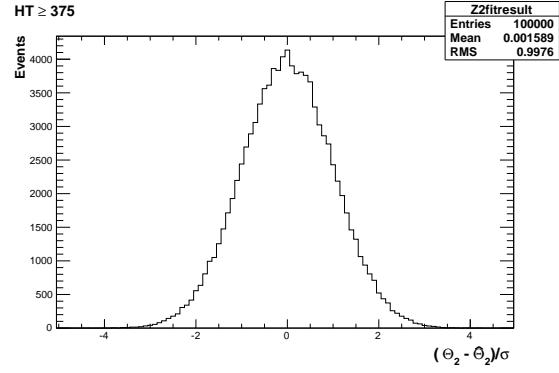
(a) Z0 Template,  $H_T > 375$ ,  $n_{jet} = 4$



(b) Z2 Template,  $H_T > 375$ ,  $n_{jet} = 4$



(a) Z0 Template,  $H_T > 375$ ,  $n_{jet} \geq 5$

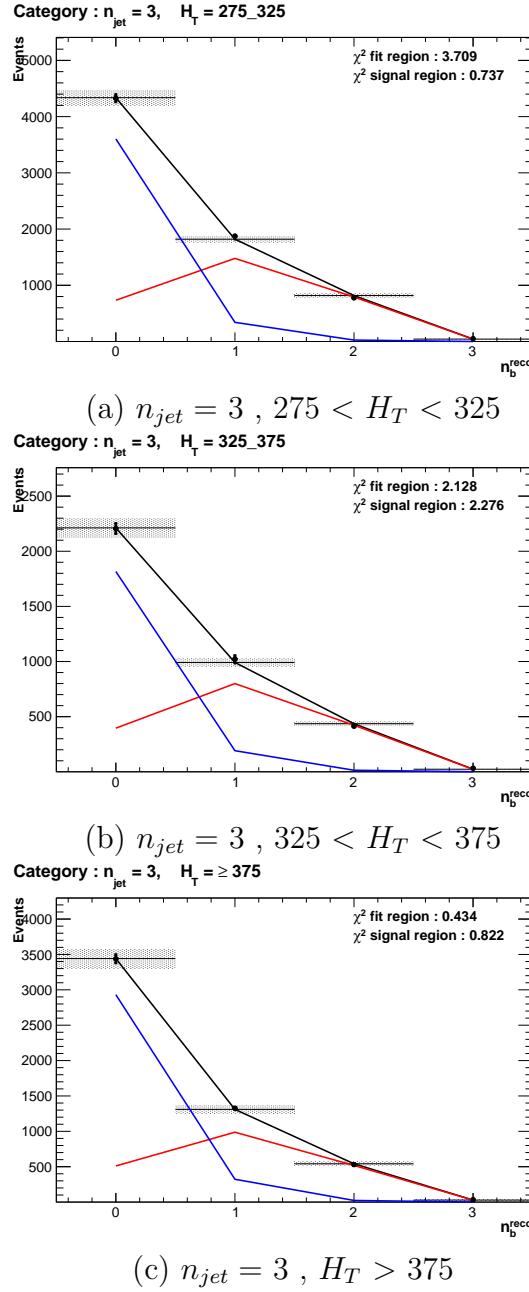


(b) Z2 Template,  $H_T > 375$ ,  $n_{jet} \geq 5$

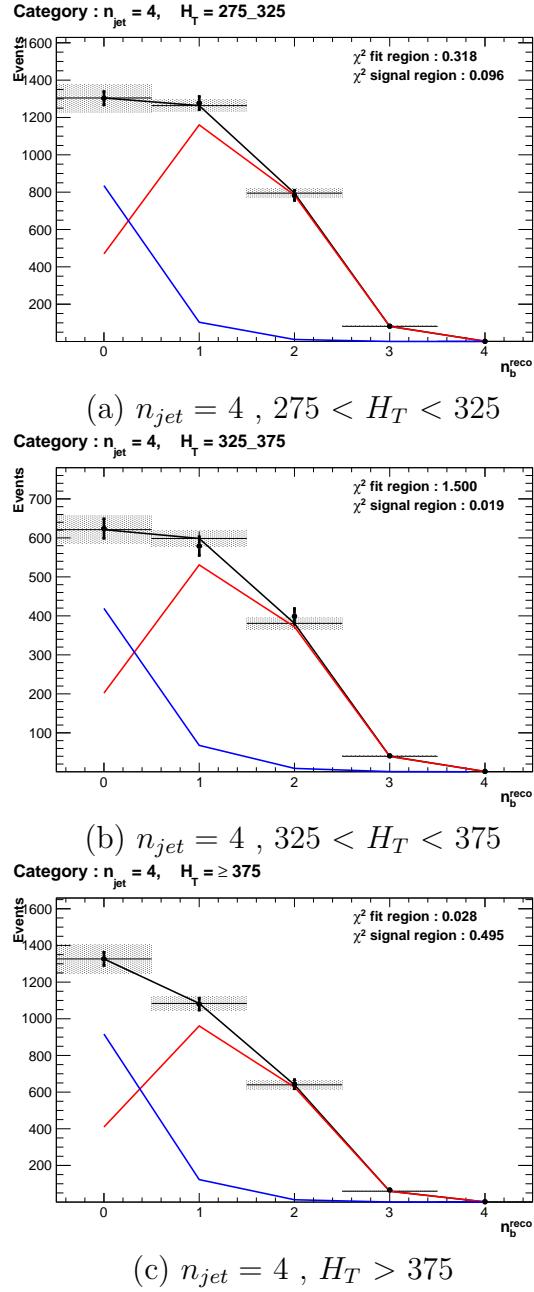
**Figure D.3:** Pull distributions of the normalisation parameter of each template,  $\frac{(\theta - \hat{\theta})}{\sigma}$ . Distributions are constructed from  $10^4$  pseudo-experiments generated by a gaussian distribution with width  $\sigma$ , centred on the nominal template value of each point within the low  $n_b^{\text{reco}}$  control region. Distributions are shown for both Z0 and Z2 templates for the medium CSV working point.

<sup>2603</sup> **D.3. Templates Fits in Data Control Sample**

- <sup>2604</sup> Template fits for the three  $H_T$  bins, in the  $n_{jet} = 3$ , medium **CSV** working point:
- <sup>2605</sup> Template fits for the three  $H_T$  bins, in the  $n_{jet} = 4$ , medium **CSV** working point:



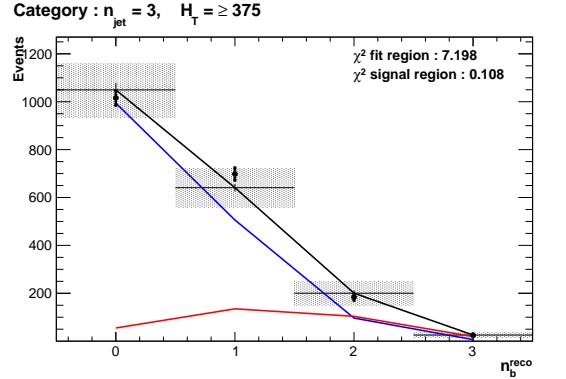
**Figure D.4:** Results of fitting the  $Z = 0$  and  $Z = 2$  templates in the  $n_b^{reco} = 0\text{--}2$  control region to data from the  $\mu + \text{jets}$  control sample, for the **CSV** medium working point, with a jet multiplicity  $n_{jet} = 3$ , in all three  $H_T$  categories. Data is represented by the black circles with the blue, red and black lines representing the  $Z=0$ ,  $Z=2$  and combination of both templates respectively. Grey bands represent the uncertainty of the fit. The  $\chi^2$  parameters represent the goodness of fit to the control and signal region.



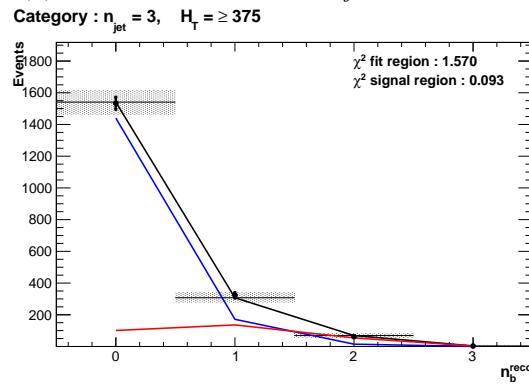
**Figure D.5:** Results of fitting the  $Z = 0$  and  $Z = 2$  templates in the  $n_b^{reco} = 0\text{--}2$  control region to data from the  $\mu + \text{jets}$  control sample, for the **CSV** medium working point, with a jet multiplicity  $n_{jet} = 4$ , in all three  $H_T$  categories. Data is represented by the black circles with the blue, red and black lines representing the  $Z=0$ ,  $Z=2$  and combination of both templates respectively. Grey bands represent the uncertainty of the fit. The  $\chi^2$  parameters represents the goodness of fit to the control and signal region.

2606 **D.4. Templates Fits in Data Signal Region**

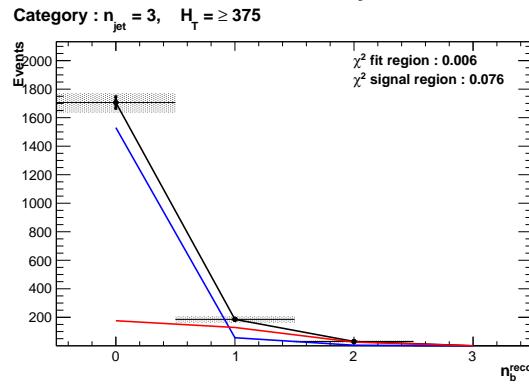
2607 Template fits for the three **CSV** working points, in the  $n_{jet} = 3, H_T > 375$  category :



(a) Loose working point :  $n_{jet} = 3, H_T > 375$



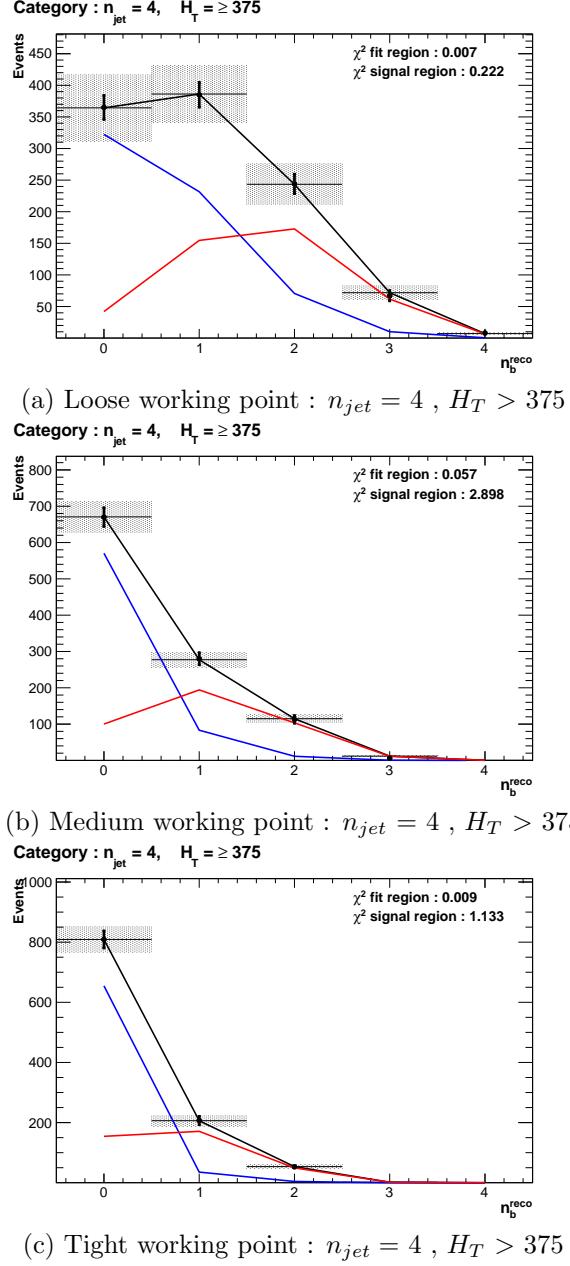
(b) Medium working point :  $n_{jet} = 3, H_T > 375$



(c) Tight working point :  $n_{jet} = 3, H_T > 375$

**Figure D.6:** Results of fitting the  $Z = 0$  and  $Z = 2$  templates in the  $n_b^{reco} = 0-2$  control region to data from the hadronic signal selection, in the  $n_{jet} = 3$  and  $H_T > 375$  category for all **CSV** working points. Data is represented by the black circles with the blue, red and black lines representing the  $Z=0$ ,  $Z=2$  and combination of both templates respectively. Grey bands represent the uncertainty of the fit. The  $\chi^2$  parameters represent the goodness of fit to the control and signal region.

2608 Template fits for the three **CSV** working points, in the  $n_{jet} = 4$ ,  $H_T > 375$  category :



**Figure D.7:** Results of fitting the  $Z = 0$  and  $Z = 2$  templates in the  $n_b^{reco} = 0\text{-}2$  control region to data from the hadronic signal selection, in the  $n_{jet} = 4$  and  $H_T > 375$  category for all **CSV** working points. Data is represented by the black circles with the blue, red and black lines representing the  $Z=0$ ,  $Z=2$  and combination of both templates respectively. Grey bands represent the uncertainty of the fit. The  $\chi^2$  parameters represent the goodness of fit to the control and signal region.



# <sup>2610</sup> Bibliography

- [1] Particle Data Group Collaboration, “Review of Particle Physics (RPP)”, *Phys.Rev.* **D86** (2012) 010001, [doi:10.1103/PhysRevD.86.010001](https://doi.org/10.1103/PhysRevD.86.010001).
- [2] Planck Collaboration, “Planck 2013 results. XVI. Cosmological parameters”, [arXiv:1303.5076](https://arxiv.org/abs/1303.5076).
- [3] ATLAS Collaboration, “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”, *Phys.Lett.* **B716** (2012) 1–29, [doi:10.1016/j.physletb.2012.08.020](https://doi.org/10.1016/j.physletb.2012.08.020), [arXiv:1207.7214](https://arxiv.org/abs/1207.7214).
- [4] CMS Collaboration, “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”, *Phys.Lett.* **B716** (2012) 30–61, [doi:10.1016/j.physletb.2012.08.021](https://doi.org/10.1016/j.physletb.2012.08.021), [arXiv:1207.7235](https://arxiv.org/abs/1207.7235).
- [5] CMS Collaboration, “Search for supersymmetry in hadronic final states with missing transverse energy using the variables  $\alpha_T$  and b-quark multiplicity in pp collisions at 8 TeV”, *Eur.Phys.J.* **C73** (2013) 2568, [doi:10.1140/epjc/s10052-013-2568-6](https://doi.org/10.1140/epjc/s10052-013-2568-6), [arXiv:1303.2985](https://arxiv.org/abs/1303.2985).
- [6] S. Weinberg, “A Model of Leptons”, *Phys. Rev. Lett.* **19** (Nov, 1967) [doi:10.1103/PhysRevLett.19.1264](https://doi.org/10.1103/PhysRevLett.19.1264).
- [7] S. Glashow, “Partial Symmetries of Weak Interactions”, *Nucl.Phys.* **22** (1961) [doi:10.1016/0029-5582\(61\)90469-2](https://doi.org/10.1016/0029-5582(61)90469-2).
- [8] A. Salam, “Weak and Electromagnetic Interactions”, *Conf.Proc.* **C680519** (1968).
- [9] G. Hooft, “Renormalizable Lagrangians for massive Yang-Mills fields”, *Nuclear Physics B* **35** (1971) [doi:\[http://dx.doi.org/10.1016/0550-3213\\(71\\)90139-8\]\(http://dx.doi.org/10.1016/0550-3213\(71\)90139-8\)](http://dx.doi.org/10.1016/0550-3213(71)90139-8).
- [10] Gargamelle Neutrino Collaboration, “Observation of Neutrino Like Interactions

- Without Muon Or Electron in the Gargamelle Neutrino Experiment”, *Phys.Lett.* **B46** (1973) 138–140, [doi:10.1016/0370-2693\(73\)90499-1](https://doi.org/10.1016/0370-2693(73)90499-1).
- [11] UA1 Collaboration, “Experimental Observation of Lepton Pairs of Invariant Mass Around 95-GeV at the CERN SPS Collider”, *Phys.Lett.* **B126** (1983) 398–410, [doi:10.1016/0370-2693\(83\)90188-0](https://doi.org/10.1016/0370-2693(83)90188-0).
- [12] UA2 Collaboration, “Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN  $\bar{p}p$  Collider”, *Phys.Lett.* **B122** (1983) [doi:10.1016/0370-2693\(83\)91605-2](https://doi.org/10.1016/0370-2693(83)91605-2).
- [13] E. Noether, “Invariante Variationsprobleme”, *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* **1918** (1918).
- [14] F. Halzen and A. D. Martin, “Quarks and Leptons”. 1985.
- [15] “Introduction to Elementary Particles”. Wiley-VCH, 2nd edition, October, 2008.
- [16] C. S. Wu et al., “Experimental Test of Parity Conservation in Beta Decay”, *Physical Review* **105** (February, 1957) [doi:10.1103/PhysRev.105.1413](https://doi.org/10.1103/PhysRev.105.1413).
- [17] P. Higgs, “Broken symmetries, massless particles and gauge fields”, *Physics Letters* **12** (1964), no. 2, [doi:\[http://dx.doi.org/10.1016/0031-9163\\(64\\)91136-9\]\(http://dx.doi.org/10.1016/0031-9163\(64\)91136-9\)](http://dx.doi.org/10.1016/0031-9163(64)91136-9).
- [18] F. Englert and R. Brout, “Broken Symmetry and the Mass of Gauge Vector Mesons”, *Phys. Rev. Lett.* **13** (Aug, 1964) [doi:10.1103/PhysRevLett.13.321](https://doi.org/10.1103/PhysRevLett.13.321).
- [19] P. W. Higgs, “Broken Symmetries and the Masses of Gauge Bosons”, *Phys. Rev. Lett.* **13** (Oct, 1964) [doi:10.1103/PhysRevLett.13.508](https://doi.org/10.1103/PhysRevLett.13.508).
- [20] G. S. Guralnik, “Global Conservation Laws and Massless Particles”, *Phys. Rev. Lett.* **13** (Nov, 1964) [doi:10.1103/PhysRevLett.13.585](https://doi.org/10.1103/PhysRevLett.13.585).
- [21] S. Weinberg, “A Model of Leptons”, *Phys. Rev. Lett.* **19** (Nov, 1967) 1264–1266, [doi:10.1103/PhysRevLett.19.1264](https://doi.org/10.1103/PhysRevLett.19.1264).
- [22] H. Yukawa, “On the Interaction of Elementary Particles. I”, *Progress of Theoretical Physics Supplement* **1** (1955) [doi:10.1143/PTPS.1.1](https://doi.org/10.1143/PTPS.1.1).
- [23] Super-Kamiokande Collaboration, “Evidence for Oscillation of Atmospheric Neutrinos”, *Phys. Rev. Lett.* **81** (Aug, 1998) [doi:10.1103/PhysRevLett.81.1562](https://doi.org/10.1103/PhysRevLett.81.1562).

- 2665 [24] R. Becker-Szendy et al., “A Search for muon-neutrino oscillations with the IMB  
2666 detector”, *Phys.Rev.Lett.* **69** (1992) [doi:10.1103/PhysRevLett.69.1010](https://doi.org/10.1103/PhysRevLett.69.1010).
- 2667 [25] S. P. Martin, “A Supersymmetry primer”, [arXiv:hep-ph/9709356](https://arxiv.org/abs/hep-ph/9709356).
- 2668 [26] H. Nilles, “Supersymmetry, Supergravity and Particle Physics”. Physics reports.  
2669 North-Holland Physics Publ., 1984.
- 2670 [27] H. E. Haber and G. L. Kane, “The Search for Supersymmetry: Probing Physics  
2671 Beyond the Standard Model”, *Phys.Rept.* **117** (1985)  
2672 [doi:10.1016/0370-1573\(85\)90051-1](https://doi.org/10.1016/0370-1573(85)90051-1).
- 2673 [28] E. Witten, “Dynamical Breaking of Supersymmetry”, *Nucl.Phys.* **B188** (1981)  
2674 [doi:10.1016/0550-3213\(81\)90006-7](https://doi.org/10.1016/0550-3213(81)90006-7).
- 2675 [29] J. Wess and B. Zumino, “Supergauge transformations in four dimensions”,  
2676 *Nuclear Physics B* **70** (1974), no. 1,  
2677 [doi:http://dx.doi.org/10.1016/0550-3213\(74\)90355-1](http://dx.doi.org/10.1016/0550-3213(74)90355-1).
- 2678 [30] H. Muller-Kirsten and A. Wiedemann, “Introduction to Supersymmetry”. World  
2679 Scientific lecture notes in physics. World Scientific, 2010.
- 2680 [31] I. Aitchison, “Supersymmetry in Particle Physics: An Elementary Introduction”.  
2681 Cambridge University Press, 2007.
- 2682 [32] K. A. Intriligator and N. Seiberg, “Lectures on Supersymmetry Breaking”,  
2683 *Class.Quant.Grav.* **24** (2007) [arXiv:hep-ph/0702069](https://arxiv.org/abs/hep-ph/0702069).
- 2684 [33] Y. Shadmi, “Supersymmetry breaking”, [arXiv:hep-th/0601076](https://arxiv.org/abs/hep-th/0601076).
- 2685 [34] C. Burgess et al., “Warped Supersymmetry Breaking”, *JHEP* **0804** (2008)  
2686 [doi:10.1088/1126-6708/2008/04/053](https://doi.org/10.1088/1126-6708/2008/04/053), [arXiv:hep-th/0610255](https://arxiv.org/abs/hep-th/0610255).
- 2687 [35] H. Murayama, “Supersymmetry breaking made easy, viable, and generic”,  
2688 [arXiv:0709.3041](https://arxiv.org/abs/0709.3041).
- 2689 [36] H. Baer and X. Tata, “Weak Scale Supersymmetry: From Superfields to  
2690 Scattering Events”. Cambridge University Press, 2006.
- 2691 [37] S. P. Martin, “Implications of supersymmetric models with natural R-parity  
2692 conservation”, [doi:10.1103/PhysRevD.54.2340](https://doi.org/10.1103/PhysRevD.54.2340), [arXiv:hep-ph/9602349](https://arxiv.org/abs/hep-ph/9602349).
- 2693 [38] G. L. Kane et al., “Study of constrained minimal supersymmetry”, *Phys.Rev.*

- 2694       **D49** (1994) [doi:10.1103/PhysRevD.49.6173](https://doi.org/10.1103/PhysRevD.49.6173), [arXiv:hep-ph/9312272](https://arxiv.org/abs/hep-ph/9312272).
- 2695       [39] C. Strege et al., “Updated global fits of the cMSSM including the latest LHC  
2696       SUSY and Higgs searches and XENON100 data”, *JCAP* **1203** (2012)  
2697       [doi:10.1088/1475-7516/2012/03/030](https://doi.org/10.1088/1475-7516/2012/03/030), [arXiv:1112.4192](https://arxiv.org/abs/1112.4192).
- 2698       [40] M. Citron et al., “The End of the CMSSM Coannihilation Strip is Nigh”,  
2699       *Phys.Rev.* **D87** (2013) [doi:10.1103/PhysRevD.87.036012](https://doi.org/10.1103/PhysRevD.87.036012), [arXiv:1212.2886](https://arxiv.org/abs/1212.2886).
- 2700       [41] D. Ghosh et al., “How Constrained is the cMSSM?”, *Phys.Rev.* **D86** (2012)  
2701       [doi:10.1103/PhysRevD.86.055007](https://doi.org/10.1103/PhysRevD.86.055007), [arXiv:1205.2283](https://arxiv.org/abs/1205.2283).
- 2702       [42] LHC New Physics Working Group Collaboration, “Simplified Models for LHC  
2703       New Physics Searches”, *J.Phys.* **G39** (2012) 105005,  
2704       [doi:10.1088/0954-3899/39/10/105005](https://doi.org/10.1088/0954-3899/39/10/105005), [arXiv:1105.2838](https://arxiv.org/abs/1105.2838).
- 2705       [43] J. Alwall et al., “Simplified Models for a First Characterization of New Physics at  
2706       the LHC”, *Phys.Rev.* **D79** (2009) 075020, [doi:10.1103/PhysRevD.79.075020](https://doi.org/10.1103/PhysRevD.79.075020),  
2707       [arXiv:0810.3921](https://arxiv.org/abs/0810.3921).
- 2708       [44] CMS Collaboration, “Interpretation of Searches for Supersymmetry with  
2709       simplified Models”, *Phys.Rev.* **D88** (2013) 052017,  
2710       [doi:10.1103/PhysRevD.88.052017](https://doi.org/10.1103/PhysRevD.88.052017), [arXiv:1301.2175](https://arxiv.org/abs/1301.2175).
- 2711       [45] J. Hisano et al., “Natural effective supersymmetry”, *Nucl.Phys.* **B584** (2000)  
2712       3–45, [doi:10.1016/S0550-3213\(00\)00343-6](https://doi.org/10.1016/S0550-3213(00)00343-6), [arXiv:hep-ph/0002286](https://arxiv.org/abs/hep-ph/0002286).
- 2713       [46] M. Papucci et al., “Natural SUSY Endures”, *JHEP* **1209** (2012) 035,  
2714       [doi:10.1007/JHEP09\(2012\)035](https://doi.org/10.1007/JHEP09(2012)035), [arXiv:1110.6926](https://arxiv.org/abs/1110.6926).
- 2715       [47] B. Allanach and B. Gripaios, “Hide and Seek With Natural Supersymmetry at the  
2716       LHC”, *JHEP* **1205** (2012) 062, [doi:10.1007/JHEP05\(2012\)062](https://doi.org/10.1007/JHEP05(2012)062),  
2717       [arXiv:1202.6616](https://arxiv.org/abs/1202.6616).
- 2718       [48] ALICE Collaboration, “The ALICE experiment at the CERN LHC”, *JINST* **3**  
2719       (2008) S08002, [doi:10.1088/1748-0221/3/08/S08002](https://doi.org/10.1088/1748-0221/3/08/S08002).
- 2720       [49] ATLAS Collaboration, “The ATLAS Experiment at the CERN Large Hadron  
2721       Collider”, *JINST* **3** (2008) [doi:10.1088/1748-0221/3/08/S08003](https://doi.org/10.1088/1748-0221/3/08/S08003).
- 2722       [50] CMS Collaboration, “The CMS experiment at the CERN LHC”, *JINST* **0803**  
2723       (2008) S08004, [doi:10.1088/1748-0221/3/08/S08004](https://doi.org/10.1088/1748-0221/3/08/S08004).

- 2724 [51] LHCb Collaboration, “The LHCb Detector at the LHC”, *JINST* **3** (2008)  
2725 S08005, [doi:10.1088/1748-0221/3/08/S08005](https://doi.org/10.1088/1748-0221/3/08/S08005).
- 2726 [52] J.-L. Caron, “LHC Layout. Schema general du LHC.”, (Sep, 1997).
- 2727 [53] CMS Collaboration, “CMS Luminosity - Public Results”, , (2011).  
2728 <http://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults>.
- 2729 [54] CERN, “CMS Compact Muon Solenoid.”, (Feb, 2010).  
2730 <http://public.web.cern.ch/public/Objects/LHC/CMSnc.jpg>.
- 2731 [55] “The CMS Electromagnetic Calorimeter Project: Technical Design Report”.  
2732 Technical Design Report CMS. CERN, Geneva, 1997.
- 2733 [56] “The CMS Muon Project: Technical Design Report”. Technical Design Report  
2734 CMS. CERN, Geneva, 1997.
- 2735 [57] CMS Collaboration, “The CMS Physics Technical Design Report, Volume 1”,  
2736 *CERN/LHCC 2006-001* (2006).
- 2737 [58] M. Cacciari et al., “The anti- $k_t$  jet clustering algorithm”, *Journal of High Energy*  
2738 *Physics* **2008** (2008), no. 04, 063.
- 2739 [59] “Jet Performance in pp Collisions at 7 TeV”, CMS-PAS-JME-10-003, CERN,  
2740 Geneva, (2010).
- 2741 [60] X. Janssen, “Underlying event and jet reconstruction in CMS”, CMS-CR-2011-012,  
2742 CERN, Geneva, (Jan, 2011).
- 2743 [61] CMS Collaboration, “Determination of jet energy calibration and transverse  
2744 momentum resolution in CMS”, *Journal of Instrumentation* **6** (2011), no. 11.,
- 2745 [62] R. Eusebi, “Jet energy corrections and uncertainties in CMS: reducing their impact  
2746 on physics measurements”, *Journal of Physics: Conference Series* **404** (2012).
- 2747 [63] CMS Collaboration, “Algorithms for b Jet identification in CMS”,  
2748 CMS-PAS-BTV-09-001, CERN, 2009. Geneva, (Jul, 2009).
- 2749 [64] CMS Collaboration, “Performance of b-tagging at  $\sqrt{s} = 8$  TeV in multijet,  $t\bar{t}$  and  
2750 boosted topology events”, CMS-PAS-BTV-13-001, CERN, Geneva, (2013).
- 2751 [65] CMS Collaboration, “Identification of b-quark jets with the CMS experiment”,  
2752 *Journal of Instrumentation* **8** (2013), no. 04.,

- 2753 [66] CMS Collaboration, “CMS. The TriDAS Project. Technical design report, Vol. 1:  
2754       The trigger systems”, ,
- 2755 [67] CMS Collaboration, “CMS: The TriDAS Project. Technical design report, Vol. 2:  
2756       Data acquisition and high-level trigger” ,.
- 2757 [68] G. Iles, Brooke et al., “Revised CMS Global Calorimeter Trigger Functionality &  
2758       Algorithms”, ,
- 2759 [69] CMS Collaboration, “ Calibration and Performance of the Jets and Energy Sums  
2760       in the Level-1 Trigger”, CMS IN 2013/006 (2013), CERN, Geneva, (2013).
- 2761 [70] B. Shorney-Mathias, “Search for supersymmetry in pp collisions with all-hadronic  
2762       final states using the  $\alpha_T$  variable with the CMS detector at the LHC”. PhD thesis,  
2763       2014.
- 2764 [71] B. Mathias et al., “Study of Level-1 Trigger Jet Performance in High Pile-up  
2765       Running Conditions”, CMS-IN-2013-007, CERN, Geneva, (2013).
- 2766 [72] J. J. Brooke, “Performance of the CMS Level-1 Trigger”, CMS-CR-2012-322,  
2767       CERN, Geneva, (Nov, 2012).
- 2768 [73] CMS Collaboration, “Search for supersymmetry in final states with missing  
2769       transverse energy and 0, 1, 2, or at least 3 b-quark jets in 7 TeV pp collisions  
2770       using the variable  $\alpha_T$ ”, *JHEP* **1301** (2013) 077,  
2771       doi:10.1007/JHEP01(2013)077, arXiv:1210.8115.
- 2772 [74] CMS Collaboration, “SUSY searches with dijet events”, CMS-PAS-SUS-08-005,  
2773       (2008).
- 2774 [75] L. Randall and D. Tucker-Smith, “Dijet Searches for Supersymmetry at the Large  
2775       Hadron Collider”, *Phys. Rev. Lett.* **101** (Nov, 2008) 221803,  
2776       doi:10.1103/PhysRevLett.101.221803.
- 2777 [76] CMS Collaboration, “Search strategy for exclusive multi-jet events from  
2778       supersymmetry at CMS”, CMS-PAS-SUS-09-001, CERN, (Jul, 2009).
- 2779 [77] CMS Collaboration, “CMS Btag POG : CMS b-tagging performance database”, ,  
2780       (2013). <https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagPOG>.
- 2781 [78] CMS Collaboration, “Calorimeter Jet Quality Criteria for the First CMS Collision  
2782       Data”, CMS-PAS-JME-09-008, CERN, (Apr, 2010).

- 2783 [79] CMS Collaboration, “Performance of CMS muon reconstruction in pp collision  
2784 events at 7 TeV”, *Journal of Instrumentation* **7** (October, 2012) 2P,  
2785 [doi:10.1088/1748-0221/7/10/P10002](https://doi.org/10.1088/1748-0221/7/10/P10002), [arXiv:1206.4071](https://arxiv.org/abs/1206.4071).
- 2786 [80] CMS Collaboration, “Search for supersymmetry in events with photons and  
2787 missing energy”, CMS-PAS-SUS-12-018, CERN, Geneva, (2012).
- 2788 [81] M. Cacciari and G. P. Salam, “Pileup subtraction using jet areas”, *Phys.Lett.*  
2789 **B659** (2008) [doi:10.1016/j.physletb.2007.09.077](https://doi.org/10.1016/j.physletb.2007.09.077), [arXiv:0707.1378](https://arxiv.org/abs/0707.1378).
- 2790 [82] Z. Bern et al., “Driving missing data at next-to-leading order”, *Phys. Rev. D* **84**  
2791 (Dec, 2011) 114002, [doi:10.1103/PhysRevD.84.114002](https://doi.org/10.1103/PhysRevD.84.114002).
- 2792 [83] Z. Bern et al., “Driving Missing Data at Next-to-Leading Order”, *Phys.Rev.*  
2793 **D84** (2011) 114002, [doi:10.1103/PhysRevD.84.114002](https://doi.org/10.1103/PhysRevD.84.114002), [arXiv:1106.1423](https://arxiv.org/abs/1106.1423).
- 2794 [84] CMS Collaboration, “Data-Driven Estimation of the Invisible Z Background to the  
2795 SUSY MET Plus Jets Search”, CMS-PAS-SUS-08-002, CERN, (Jan, 2009).
- 2796 [85] T. Sjostrand et al., “PYTHIA 6.4 Physics and Manual”, *JHEP* **0605** (2006) 026,  
2797 [doi:10.1088/1126-6708/2006/05/026](https://doi.org/10.1088/1126-6708/2006/05/026), [arXiv:hep-ph/0603175](https://arxiv.org/abs/hep-ph/0603175).
- 2798 [86] W. Beenakker et al., “Squark and gluino production at hadron colliders”,  
2799 *Nucl.Phys.* **B492** (1997) 51–103, [doi:10.1016/S0550-3213\(97\)80027-2](https://doi.org/10.1016/S0550-3213(97)80027-2),  
2800 [arXiv:hep-ph/9610490](https://arxiv.org/abs/hep-ph/9610490).
- 2801 [87] S. Abdullin et al., “The Fast Simulation of the CMS Detector at LHC”, *Journal*  
2802 *of Physics: Conference Series* **331** (2011), no. 3,.
- 2803 [88] S. Banerjee, M. D. Hildreth, and the CMS Collaboration, “Validation and Tuning  
2804 of the CMS Full Simulation”, *Journal of Physics: Conference Series* **331** (2011).
- 2805 [89] CMS Collaboration, “CMS Luminosity Based on Pixel Cluster Counting - Summer  
2806 2012 Update”, CMS-PAS-LUM-12-001, CERN, Geneva, (2012).
- 2807 [90] M. Botje et al., “The PDF4LHC Working Group Interim Recommendations”,  
2808 [arXiv:1101.0538](https://arxiv.org/abs/1101.0538).
- 2809 [91] E. M. Laird, “A Search for Squarks and Gluinos with the CMS Detector”. PhD  
2810 thesis, 2012.
- 2811 [92] R. Cousins, “Probability Density Functions for Positive Nuisance Parameters”.  
2812 2012. <http://www.physics.ucla.edu/cousins/stats/cousinslognormalprior.pdf>.

- 2813 [93] L. Moneta et al., “The RooStats project”, 2010. [arXiv:1009.1003](https://arxiv.org/abs/1009.1003).
- 2814 [94] F. James and M. Roos, “Minuit: A System for Function Minimization and  
2815 Analysis of the Parameter Errors and Correlations”, *Comput.Phys.Commun.* **10**  
2816 (1975) 343–367, [doi:10.1016/0010-4655\(75\)90039-9](https://doi.org/10.1016/0010-4655(75)90039-9).
- 2817 [95] A. L. Read, “Presentation of search results: the CL<sub>s</sub> technique”, *Journal of  
2818 Physics G: Nuclear and Particle Physics* **28** (2002).
- 2819 [96] T. Junk, “Confidence level computation for combining searches with small  
2820 statistics”, *Nuclear Instruments and Methods in Physics Research Section A* **434**  
2821 (1999), no. 23, [doi:http://dx.doi.org/10.1016/S0168-9002\(99\)00498-2](http://dx.doi.org/10.1016/S0168-9002(99)00498-2).
- 2822 [97] A. L. Read, “Modified frequentist analysis of search results (the CL<sub>s</sub> method)”,  
2823 CERN-OPEN-2000-205, (2000).
- 2824 [98] G. Cowan et al., “Asymptotic formulae for likelihood-based tests of new physics”,  
2825 *The European Physical Journal C* **71** (2011)  
2826 [doi:10.1140/epjc/s10052-011-1554-0](https://doi.org/10.1140/epjc/s10052-011-1554-0).
- 2827 [99] G. Cowan et al., “Asymptotic formulae for likelihood-based tests of new physics”,  
2828 *European Physical Journal C* **71** (February, 2011)  
2829 [doi:10.1140/epjc/s10052-011-1554-0](https://doi.org/10.1140/epjc/s10052-011-1554-0), [arXiv:1007.1727](https://arxiv.org/abs/1007.1727).
- 2830 [100] CMS Collaboration, “Search for supersymmetry in final states with missing  
2831 transverse energy and 0, 1, 2, or at least 3 b-quark jets in 7 TeV pp collisions  
2832 using the variable  $\alpha_T$ ”, *JHEP* **1301** (2013) 077,  
2833 [doi:10.1007/JHEP01\(2013\)077](https://doi.org/10.1007/JHEP01(2013)077), [arXiv:1210.8115](https://arxiv.org/abs/1210.8115).
- 2834 [101] CMS Collaboration, “Search for new physics in the multijet and missing transverse  
2835 momentum final state in proton-proton collisions at  $\sqrt{s} = 8$  TeV”,  
2836 [arXiv:1402.4770](https://arxiv.org/abs/1402.4770).
- 2837 [102] CMS Collaboration, “Search for new physics in events with same-sign dileptons  
2838 and jets in pp collisions at  $\sqrt{s} = 8$  TeV”, *JHEP* **1401** (2014) 163,  
2839 [doi:10.1007/JHEP01\(2014\)163](https://doi.org/10.1007/JHEP01(2014)163), [arXiv:1311.6736](https://arxiv.org/abs/1311.6736).
- 2840 [103] CMS Collaboration, “Search for supersymmetry in pp collisions at  $\sqrt{s} = 8$  TeV in  
2841 events with a single lepton, large jet multiplicity, and multiple b jets”,  
2842 [arXiv:1311.4937](https://arxiv.org/abs/1311.4937).

**2843 Acronyms**

2844	<b>ALICE</b>	A Large Ion Collider Experiment
2845	<b>ATLAS</b>	A Toroidal LHC ApparatuS
2846	<b>APD</b>	Avalanche Photo-Diodes
2847	<b>BSM</b>	Beyond Standard Model
2848	<b>CERN</b>	European Organisation for Nuclear Research
2849	<b>CMS</b>	Compact Muon Solenoid
2850	<b>CMSSM</b>	Compressed Minimal SuperSymmetric Model
2851	<b>CSC</b>	Cathode Stripe Chamber
2852	<b>CSV</b>	Combined Secondary Vertex
2853	<b>CSVM</b>	Combined Secondary Vertex Medium Working Point
2854	<b>DT</b>	Drift Tube
2855	<b>ECAL</b>	Electromagnetic CALorimeter
2856	<b>EB</b>	Electromagnetic CALorimeter Barrel
2857	<b>EE</b>	Electromagnetic CALorimeter Endcap
2858	<b>ES</b>	Electromagnetic CALorimeter pre-Shower
2859	<b>EMG</b>	Exponentially Modified Gaussian
2860	<b>EPJC</b>	European Physical Journal C
2861	<b>EWK</b>	Electroweak Sector
2862	<b>GCT</b>	Global Calorimeter Trigger
2863	<b>GMT</b>	Global MuonTrigger
2864	<b>GT</b>	Global Trigger
2865	<b>HB</b>	Hadron Barrel
2866	<b>HCAL</b>	Hadronic CALorimeter

2867	<b>HE</b>	Hadron Endcaps
2868	<b>HF</b>	Hadron Forward
2869	<b>HLT</b>	Higher Level Trigger
2870	<b>HO</b>	Hadron Outer
2871	<b>HPD</b>	Hybrid Photo Diode
2872	<b>ISR</b>	Initial State Radiation
2873	<b>LUT</b>	Look Up Table
2874	<b>L1</b>	Level 1 Trigger
2875	<b>LEP</b>	Large Electron-Positron Collidior
2876	<b>LHC</b>	Large Hadron Collider
2877	<b>LHCb</b>	Large Hadron Collider Beauty
2878	<b>LO</b>	Leading Order
2879	<b>LSP</b>	Lightest Supersymmetric Partner
2880	<b>NLL</b>	Next to Leading Logorithmic Order
2881	<b>NLO</b>	Next to Leading Order
2882	<b>NNLO</b>	Next to Next Leading Order
2883	<b>POGs</b>	Physics Object Groups
2884	<b>PS</b>	Proton Synchrotron
2885	<b>QED</b>	Quantum Electro-Dynamics
2886	<b>QCD</b>	Quantum Chromo-Dynamics
2887	<b>QFT</b>	Quantum Field Theory
2888	<b>RBXs</b>	Readout Boxes
2889	<b>RPC</b>	Resistive Plate Chamber
2890	<b>RCT</b>	Regional Calorimeter Trigger
2891	<b>RMT</b>	Regional Muon Trigger

2892	<b>SUSY</b>	SUperSYmmetry
2893	<b>SM</b>	Standard Model
2894	<b>SMS</b>	Simplified Model Spectra
2895	<b>SPS</b>	Super Proton Synchrotron
2896	<b>TF</b>	Transfer Factor
2897	<b>TP</b>	Trigger Primate
2898	<b>VEV</b>	Vacuum Expectation Value
2899	<b>VPT</b>	Vacuum Photo-Triodes
2900	<b>WIMP</b>	Weakly Interacting Massive Particle