

<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  $p_T$  jets and a large missing energy signature, is conducted using data recorded by the Compact Muon Solenoid detector based at the Large Hadron Collider. The analysis is performed with  $11.7 \text{ fb}^{-1}$  of data, collected with a center-of-mass collision energy of 8 TeV during the 2012 run period. The dimensionless kinematic variable  $\alpha_T$  is used as a tool to select events with genuine missing energy signatures, whilst Standard Model backgrounds in the signal region are estimated using data driven control samples, which have similar kinematics to the signal region. No excess of over Standard Model expectations is found. Exclusion limits are set at the 95% confidence level in the parameter space of a range of supersymmetric simplified model topologies, with special emphasis on those with compressed spectra (small mass splittings) and natural SUSY scenarios (large number of final state b flavoured jets). A complementary method to search for natural SUSY signatures, through the use of a simple template fit is also presented. The event selections of the  $\alpha_T$  search are used as a vehicle to validate the technique in both data and simulation. Estimated Standard Model backgrounds from the template fits are compared with those determined from the data driven background estimation method of the  $\alpha_T$  search within the signal search region, where good agreement between the individual predictions and that of data are observed. Additionally the efficiency of the hadronic Level-1 single jet triggers are measured throughout the 2012 run period, where a change to the jet seed algorithm was implemented during the data taking period. This change was introduced to negate an increase in rate which can be attributed to pile-up jets, whilst maintaining similar performance in the triggering of physics events.

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 girlfriend Hannah. Thanks for putting up with the whinging, complaining and clopping.

63 Finally my largest thanks go to my Mum and Dad whose patience, encouragement and  
64 considerable financial support have allowed me to take the many steps that lead me here  
65 today.

# **Contents**

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

92	3.4.2. The L1 trigger jet algorithm . . . . .	35
93	3.4.3. Measuring L1 jet trigger efficiencies . . . . .	37
94	3.4.4. Effects of the L1 jet seed . . . . .	38
95	3.4.5. Robustness of L1 jet performance against pile-up . . . . .	40
96	3.4.6. Summary . . . . .	43
97	<b>4. SUSY Searches In Hadronic Final States</b>	<b>45</b>
98	4.1. An Introduction to the $\alpha_T$ Search . . . . .	46
99	4.1.1. The $\alpha_T$ variable . . . . .	48
100	4.2. Search Strategy . . . . .	50
101	4.2.1. Physics objects . . . . .	53
102	4.2.2. Event selection . . . . .	57
103	4.2.3. Control sample definition and background estimation . . . . .	60
104	4.2.4. Estimating the QCD multi-jet background . . . . .	67
105	4.3. Trigger Strategy . . . . .	69
106	4.4. Measuring MC Normalisation Factors via $H_T$ Sidebands . . . . .	70
107	4.5. Determining MC Simulation Yields with Higher Statistical Precision . . . . .	71
108	4.5.1. The formula method . . . . .	72
109	4.5.2. Establishing proof of principle . . . . .	73
110	4.5.3. Correcting measured efficiencies in simulation to data . . . . .	74
111	4.6. Systematic Uncertainties on Transfer Factors . . . . .	77
112	4.6.1. Determining systematic uncertainties from closure tests . . . . .	80
113	4.7. Simplified Models, Efficiencies and Systematic Uncertainties . . . . .	83
114	4.7.1. Signal efficiency . . . . .	83
115	4.7.2. Applying b-tag scale factor corrections in signal samples . . . . .	84
116	4.7.3. Experimental uncertainties . . . . .	86
117	4.8. Statistical Interpretation . . . . .	89
118	4.8.1. Hadronic sample . . . . .	89
119	4.8.2. $H_T$ evolution model . . . . .	90
120	4.8.3. Electroweak Sector (EWK) control samples . . . . .	90
121	4.8.4. Contributions from signal . . . . .	93
122	4.8.5. Total likelihood . . . . .	94
123	<b>5. Results And Interpretation</b>	<b>96</b>
124	5.1. Compatibility with the Standard Model Hypothesis . . . . .	96
125	5.2. SUSY . . . . .	105
126	5.2.1. The $CL_s$ method . . . . .	105

127	5.2.2. Interpretation in simplified signal models . . . . .	106
128	<b>6. Searching For Natural SUSY With B-tag Templates.</b>	110
129	6.1. Concept . . . . .	110
130	6.2. Application to the $\alpha_T$ Search . . . . .	113
131	6.2.1. Proof of principle in simulation . . . . .	114
132	6.2.2. Results in a data control sample . . . . .	116
133	6.2.3. Application to the $\alpha_T$ hadronic search region . . . . .	118
134	6.3. Summary . . . . .	120
135	<b>7. Conclusions</b>	122
136	<b>A. Miscellaneous</b>	124
137	A.1. Jet Identification Criteria . . . . .	124
138	A.2. Primary Vertices . . . . .	125
139	<b>B. L1 Jets</b>	126
140	B.1. Jet matching efficiencies . . . . .	126
141	B.2. Leading Jet Energy Resolution . . . . .	127
142	B.3. Resolution for Energy Sum Quantities . . . . .	130
143	<b>C. Additional material on background estimation methods</b>	132
144	C.1. Determination of $k_{QCD}$ . . . . .	132
145	C.2. Effect of varying background cross sections on closure tests . . . . .	133
146	<b>D. Additional Material For B-tag Template Method</b>	135
147	D.1. Templates Fits in Simulation . . . . .	135
148	D.2. Pull Distributions for Template Fits . . . . .	138
149	D.3. Templates Fits in Data Control Sample . . . . .	139
150	D.4. Templates Fits in Data Signal Region . . . . .	141
151	<b>Bibliography</b>	144

# <sup>152</sup> List of Figures

<sup>153</sup> 2.1.	One loop quantum corrections to the Higgs squared mass parameter $m_h^2$ due to a fermion. . . . .	<sup>14</sup>
<sup>155</sup> 2.2.	Two example simplified model decay chains. . . . .	<sup>19</sup>
<sup>156</sup> 3.1.	A top down layout of the LHC, with the position of the four main detectors labelled. . . . .	<sup>21</sup>
<sup>158</sup> 3.2.	The total integrated luminosity delivered to and collected by Compact Muon Solenoid (CMS) during the 2012 8 TeV $pp$ runs . . . . .	<sup>22</sup>
<sup>160</sup> 3.3.	A pictorial depiction of the CMS detector. . . . .	<sup>24</sup>
<sup>161</sup> 3.4.	Illustration of the CMS Electromagnetic CALorimeter (ECAL). . . . .	<sup>26</sup>
<sup>162</sup> 3.5.	Schematic of the CMS Hadronic CALorimeter (HCAL). . . . .	<sup>27</sup>
<sup>163</sup> 3.6.	Combined Secondary Vertex (CSV) algorithm discriminator values in enriched ttbar and inclusive multi jet samples . . . . .	<sup>31</sup>
<sup>165</sup> 3.7.	Data/MC b-tag scale factors derived using the Combined Secondary Vertex Medium Working Point (CSVM) tagger. . . . .	<sup>32</sup>
<sup>167</sup> 3.8.	Data/MC mis-tag scale factors derived using the CSVM tagger. . . . .	<sup>33</sup>
<sup>168</sup> 3.9.	The CMS Level 1 Trigger (L1) Trigger system. . . . .	<sup>34</sup>
<sup>169</sup> 3.10.	Illustration of the Level-1 jet finding algorithm. . . . .	<sup>36</sup>
<sup>170</sup> 3.11.	L1 jet efficiency turn-on curves as a function of the offline CaloJet and PFJet $E_T$ . . . . .	<sup>38</sup>
<sup>172</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. . . . .	<sup>39</sup>

---

174	3.13. Trigger cross section for the L1HTT150 trigger path. . . . .	40
175	3.14. L1 $H_T$ efficiency turn-on curves as a function of the offline CaloJet $H_T$ . . . . .	41
176	3.15. L1 jet efficiency turn-on curves as a function of the leading offline $E_T$ Calo	
177	(left) and PF (right) jet, for low, medium and high pile-up conditions. . . . .	42
178	3.16. Fit values from an Exponentially Modified Gaussian (EMG) function fitted	
179	to the resolution plots of leading Calo jet $E_T$ measured as a function of	
180	$\frac{(L1\ E_T - \text{Offline}\ E_T)}{\text{Offline}\ E_T}$ for low, medium and high pile-up conditions. . . . .	43
181	3.17. Fit values from an EMG function fitted to the resolution plots of leading	
182	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	
183	high pile-up conditions. . . . .	44
184	4.1. Reconstructed offline $H_T$ distribution in the hadronic signal selection, from	
185	$11.7\text{fb}^{-1}$ of data, in which no $\alpha_T$ requirement is made. . . . .	48
186	4.2. The event topologies of background QCD dijet events (right) and a generic	
187	SUSY signature with genuine $Z_T$ (left). . . . .	48
188	4.3. The $\alpha_T$ distributions for the low 2-3 (left) and high $\geq 4$ (right) jet	
189	multiplicities after a full analysis selection and shown for $H_T > 375$ . . . . .	50
190	4.4. Pictorial depiction of the analysis strategy employed by the $\alpha_T$ search to	
191	increase sensitivity to a wide spectra of SUSY models. . . . .	53
192	4.5. Data/MC comparisons of key variables for the hadronic signal region. . . . .	60
193	4.6. Data/MC comparisons of key variables for the $\mu +$ jets selection. . . . .	63
194	4.7. Data/MC comparisons of key variables for the $\mu\mu +$ jets selection. . . . .	64
195	4.8. Data/MC comparisons of key variables for the $\gamma +$ jets selection. . . . .	66
196	4.9. QCD sideband regions, used for determination of $k_{\text{QCD}}$ . . . . .	68
197	4.10. Tagging efficiencies of (a) b-jets, (b) c-jets, and (c) light-jets determined	
198	from all jets within each individual analysis $H_T$ bin. . . . .	75
199	4.11. Sets of closure tests overlaid on top of the systematic uncertainty used for	
200	each of the five $H_T$ regions. . . . .	82

---

201	4.12. Signal efficiencies fo the Simplified Model Spectra (SMS) models (a) T1 202 and (b) T2. . . . .	84
203	5.1. Comparison of the observed yields and Standard Model (SM) expectations 204 given by the simultaneous fit in bins of $H_T$ for the (a) hadronic, (b) $\mu +$ 205 jets, (c) $\mu\mu +$ jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 0$ and 206 $n_{jet} \leq 3$ . . . . .	98
207	5.2. Comparison of the observed yields and SM expectations given by the 208 simultaneous fit in bins of $H_T$ for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ 209 jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 1$ and $n_{jet} \leq 3$ . . . . .	99
210	5.3. Comparison of the observed yields and SM expectations given by the 211 simultaneous fit in bins of $H_T$ for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ 212 jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 2$ and $n_{jet} \leq 3$ . . . . .	100
213	5.4. Comparison of the observed yields and SM expectations given by the 214 simultaneous fit in bins of $H_T$ for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ 215 jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 0$ and $n_{jet} \geq 4$ . . . . .	101
216	5.5. Comparison of the observed yields and SM expectations given by the 217 simultaneous fit in bins of $H_T$ for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ 218 jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 1$ and $n_{jet} \geq 4$ . . . . .	102
219	5.6. Comparison of the observed yields and SM expectations given by the 220 simultaneous fit in bins of $H_T$ for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ 221 jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 2$ and $n_{jet} \geq 4$ . . . . .	103
222	5.7. Comparison of the observed yields and SM expectations given by the 223 simultaneous fit in bins of $H_T$ for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ 224 jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 3$ and $n_{jet} \geq 4$ . . . . .	103
225	5.8. Comparison of the observed yields and SM expectations given by the 226 simultaneous fit in bins of $H_T$ for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ 227 jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} \geq 4$ and $n_{jet} \geq 4$ . . . . .	104
228	5.9. Production and decay modes for the various SMS models interpreted 229 within the analysis. . . . .	108
230	5.10. Upper limit of cross section at 95% CL as a function of $m_{\tilde{q}/\tilde{g}}$ and $m_{LSP}$ 231 for various SMS models. . . . .	109

---

232	6.1. The b-tagging (a), c-quark mis-tagging (b), and light-quark mis-tagging rate (c) as measured in simulation after the $\alpha_T$ analysis, $\mu + \text{jets}$ control sample selection in the region $H_T > 375$ . . . . .	112
235	6.2. The results of fitting the $Z = 0$ and $Z = 2$ templates to the $n_b^{reco} = 0, 1, 2$ bins taken directly from simulation in the region $H_T > 375$ GeV, for the $n_{jet} \geq 5$ category. . . . .	116
238	6.3. The results of fitting the $Z = 0$ and $Z = 2$ templates to the $n_b^{reco} = 0, 1, 2$ bins taken from data, for the $n_{jet} \geq 5$ category and medium CSV working point. . . . .	117
241	6.4. The results of fitting the $Z = 0$ and $Z = 2$ templates to the $n_b^{reco} = 0, 1, 2$ bins taken from data, in the $n_{jet} \geq 5$ and $H_T > 375$ category for all CSV working points. . . . .	119
244	B.1. Leading jet matching efficiency as a function of the offline CaloJet $E_T$ . . . . .	126
245	B.2. Resolution plots of the leading offline Calo $E_T$ measured as a function of $\frac{(L1 E_T - \text{Offline } E_T)}{\text{Offline } E_T}$ for low (a), medium (b) and high (c) pile-up conditions. . . . .	128
247	B.3. Resolution plots of the leading off-line PF $E_T$ measured as a function of $\frac{(L1 E_T - \text{Offline } E_T)}{\text{Offline } E_T}$ for low (a), medium (b) and high (c) pile-up conditions. . . . .	130
249	B.4. $H_T$ resolution parameters in bins of Calo $H_T$ measured in the defined low, medium and high pile up conditions. . . . .	130
251	B.5. $H_T$ resolution parameters in bins of PF $H_T$ measured in the defined low, medium and high pile up conditions. . . . .	131
253	B.6. $\mathcal{H}_T$ resolution parameters in bins of $\mathcal{H}_T$ measured in the defined low, medium and high pile up conditions. . . . .	131
255	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. . . . .	131
257	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 $275 < H_T < 575$ . . . . .	132
259	C.2. Sets of closure tests overlaid on top of the systematic uncertainty used for each of the five $H_T$ regions. . . . .	133

---

261	C.3. Sets of closure tests overlaid on top of the systematic uncertainty used for each of the five $H_T$ regions. . . . .	133
262		
263	D.1. The results of fitting the $Z = 0$ and $Z = 2$ templates to the $n_b^{reco} = 0, 1, 2$ bins taken directly from simulation in the region $H_T > 375$ GeV, for the $n_{jet} = 3$ category. . . . .	136
264		
265		
266	D.2. The results of fitting the $Z = 0$ and $Z = 2$ templates to the $n_b^{reco} = 0, 1, 2$ bins taken directly from simulation in the region $H_T > 375$ GeV, for the $n_{jet} = 4$ category. . . . .	137
267		
268		
269	D.3. Pull distributions of $\frac{(\theta - \hat{\theta})}{\sigma}$ for $10^4$ pseudo-experiments generated from a gaussian distribution centred on the $n_b^{reco}$ template values from simulation with width $\sigma$ . . . . .	138
270		
271		
272	D.4. The results of fitting the $Z = 0$ and $Z = 2$ templates to the $n_b^{reco} = 0, 1, 2$ bins taken directly from data, for the $n_{jet} = 3$ category and medium CSV working point. . . . .	139
273		
274		
275	D.5. The results of fitting the $Z = 0$ and $Z = 2$ templates to the $n_b^{reco} = 0, 1, 2$ bins taken directly from data, for the $n_{jet} = 4$ category and medium CSV working point. . . . .	140
276		
277		
278	D.6. The results of fitting the $Z = 0$ and $Z = 2$ templates to the $n_b^{reco} = 0, 1, 2$ bins taken from data, in the $n_{jet} = 3$ and $H_T > 375$ category for all CSV working points. . . . .	141
279		
280		
281	D.7. The results of fitting the $Z = 0$ and $Z = 2$ templates to the $n_b^{reco} = 0, 1, 2$ bins taken from data, in the $n_{jet} = 4$ and $H_T > 375$ category for all CSV working points. . . . .	142
282		
283		

# <sup>284</sup> List of Tables

285	2.1. The fundamental particles of the SM, with spin, charge and mass displayed.	6
286	3.1. Results of a cumulative EMG function fit to the turn-on curves for L1	
287	single jet triggers in 2012 Run Period C. . . . .	38
288	3.2. Results of a cumulative EMG function fit to the turn-on curves for L1	
289	single jet triggers in the 2012 run period B and C. . . . .	40
290	3.3. Results of a cumulative EMG function fit to the turn-on curves for $H_T$ in	
291	2012 run period B and C. . . . .	41
292	3.4. Results of a cumulative EMG function fit to the efficiency turn-on curves	
293	for L1 single jet triggers in the 2012 run period C, for low,medium and	
294	high pile-up conditions. . . . .	41
295	3.5. Results of a cumulative EMG function fit to the efficiency turn-on curves	
296	for Level-1 single jet triggers in the 2012 run period C, for low,medium	
297	and high pile-up conditions. . . . .	42
298	4.1. A summary of the SMS models interpreted in this analysis, involving both	
299	direct (D) and gluino-induced (G) production of squarks and their decays.	46
300	4.2. Muon Identification criteria used within the analysis for selection/veto	
301	purposes in the muon control/signal selections. . . . .	54
302	4.3. Photon Identification criteria used within the analysis for selection/veto	
303	purposes in the $\gamma +$ jets control/signal selections. . . . .	55
304	4.4. Electron Identification criteria used within the analysis for veto purposes.	56
305	4.5. Noise filters that are applied to remove spurious and non-physical $\cancel{E}_T$	
306	signatures within the CMS detector. . . . .	57

307	4.6. Jet thresholds used in the three $H_T$ regions of the analysis. . . . .	58
308	4.7. Best fit values for the parameters $k_{\text{QCD}}$ obtained from sideband regions	
309	B,C <sub>1</sub> ,C <sub>2</sub> ,C <sub>3</sub> . . . . .	69
310	4.8. Measured efficiencies of the $H_T$ and $\alpha_T$ legs of the HT and HT_alphaT	
311	triggers in independent analysis bins. . . . .	70
312	4.9. k-factors calculated for different EWK processes. . . . .	71
313	4.10. Comparing yields in simulation within the $\mu + \text{jets}$ selection determined	
314	from the formula method described in Equation (4.11), and that taken	
315	directly from simulation . . . . .	74
316	4.11. The absolute change in the Transfer Factor (TF)s used to predict the	
317	entire signal region SM background, using the $\mu + \text{jets}$ control sample	
318	when the systematic uncertainties of the data to simulation scale factors	
319	are varied by $\pm 1\sigma$ . . . . .	76
320	4.12. A summary of the results obtained from zeroeth order polynomial (i.e.	
321	a constant) and linear fits to five sets of closure tests performed in the	
322	$2 \geq n_{\text{jet}} \geq 3$ category. . . . .	79
323	4.13. A summary of the results obtained from zeroeth order polynomial (i.e.	
324	a constant) and linear fits to five sets of closure tests performed in the	
325	$n_{\text{jet}} \geq 4$ category. . . . .	79
326	4.14. A summary of the results obtained from zeroeth order polynomial (i.e. a	
327	constant) and linear fits to three sets of closure tests performed between	
328	the $2 \leq n_{\text{jet}} \leq 3$ and $n_{\text{jet}} \geq 4$ categories. . . . .	80
329	4.15. Calculated systematic uncertainties for the five $H_T$ regions, determined	
330	from the closure tests. . . . .	81
331	4.16. Estimates of systematic uncertainties on the signal efficiency (%) for	
332	various SMS models when considering points in the region near to the	
333	diagonal . . . . .	88
334	4.17. Estimates of systematic uncertainties on the signal efficiency (%) for	
335	various SMS models when considering points in the region near to the	
336	diagonal . . . . .	88

---

337	4.18. The systematic parameters used in $H_T$ bins. . . . .	92
338	4.19. Nuisance parameters used within the different hadronic signal bins of the	
339	analysis . . . . .	94
340	5.1. Summary of control samples used by each fit results, and the Figures in	
341	which they are displayed. . . . .	97
342	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	
343	statistical and systematic uncertainties given by the simultaneous fit. . .	97
344		
345	5.3. A table representing the SMS models interpreted within the analysis. . .	107
346	6.1. Typical underlying b-quark content of different SM processes which are	
347	common to many SUSY searches. . . . .	111
348	6.2. Summary of the fit predictions in the $n_b^{reco}$ signal region for $n_{jet} = 3, =$	
349	$4, \geq 5$ . The fit region is $n_b^{reco} = 0, 1, 2$ and simulation yields are normalised	
350	to an integrated luminosity of $10 \text{ fb}^{-1}$ . . . . .	115
351	6.3. Summary of the fit predictions in the $n_b^{reco}$ signal region of the $\mu + \text{jets}$	
352	control sample, for $n_{jet} = 3, = 4, \geq 5$ . The fit region is $n_b^{reco} = 0, 1, 2$	
353	using $11.5 \text{ fb}^{-1}$ of data at $\sqrt{s} = 8\text{TeV}$ . . . . .	118
354	6.4. Summary of the fit predictions in the $n_b^{reco}$ signal region of the $\mu + \text{jets}$	
355	control sample, for $n_{jet} = 3, = 4, \geq 5$ . The fit region is $n_b^{reco} = 0, 1, 2$	
356	using $11.5 \text{ fb}^{-1}$ of data at $\sqrt{s} = 8\text{TeV}$ . . . . .	120
357	A.1. Criteria for a reconstructed jet to pass the loose calorimeter jet id. . . . .	124
358	A.2. Criteria for a reconstructed jet to pass the loose PF jet id. . . . .	125
359	A.3. Criteria for a vertex in an event to be classified as a 'good' reconstructed	
360	primary vertex. . . . .	125
361	B.1. Results of a cumulative EMG function fit to the turn-on curves for the	
362	matching efficiency of the leading jet in an event to a Level-1 jet in run	
363	2012C and 2012B data. . . . .	127



# Chapter 1.

## <sup>368</sup> Introduction

<sup>369</sup> During the 20th century, great advances have been made in the human understanding  
<sup>370</sup> of the universe, its origins, its future and its composition. The Standard Model (**SM**)  
<sup>371</sup> first formulated in the 1960s is one of the crowning achievements in science's quest to  
<sup>372</sup> explain the most fundamental processes and interactions that make up our universe. It  
<sup>373</sup> has provided a highly successful explanation of a wide range of phenomena in Particle  
<sup>374</sup> Physics and has stood up to extensive experimental scrutiny [1].

<sup>375</sup> Despite its success it is not a complete theory, with significant questions remaining  
<sup>376</sup> unanswered. It describes only three of the four known forces with gravity not incorporated  
<sup>377</sup> within the framework of the **SM**. Cosmological experiments infer that just  $\sim 4\%$  of the  
<sup>378</sup> observable universe exists as matter, with elusive "Dark Matter" accounting for a further  
<sup>379</sup>  $\sim 23\%$  [2]. However no particle predicted by the **SM** is able to account for it. At higher  
<sup>380</sup> energy scales and small distances, the (non-)unification of the fundamental forces point  
<sup>381</sup> to problems with the **SM** at least at higher energies not yet probed experimentally.

<sup>382</sup> Many theories exist as extensions to the **SM**, predicting a range of observables that can  
<sup>383</sup> be detected at the Large Hadron Collider (**LHC**) of which SUperSYmmetry (**SUSY**) is  
<sup>384</sup> one such example. It predicts a new symmetry of nature in which all current particles  
<sup>385</sup> in the **SM** would have a corresponding supersymmetric partner. Common to most  
<sup>386</sup> Supersymmetric theories is a stable, weakly interacting Lightest Supersymmetric Partner  
<sup>387</sup> (**LSP**), which has the properties of a possible dark matter candidate. The **SM** and the  
<sup>388</sup> main principles of Supersymmetric theories are outlined in Chapter 2, with emphasis  
<sup>389</sup> placed on how experimental signatures of **SUSY** may reveal themselves in proton collisions  
<sup>390</sup> at the **LHC**.

391 The experimental goal of the LHC is to further test the framework of the SM, exploring the  
392 TeV mass scale for the first time, and to seek a connection between the particles produced  
393 in proton collisions and dark matter. The first new discovery by this extraordinary  
394 machine was announced on the 4th of July 2012. The long-awaited discovery was the  
395 culmination of decades of experimental endeavours in the search for the Higgs boson,  
396 providing an answer to the mechanism of electroweak symmetry breaking within the SM  
397 [3][4].

398 This discovery was made possible through data taken by the two multi purpose detectors  
399 (Compact Muon Solenoid (CMS) and A Toroidal LHC ApparatuS (ATLAS)) located  
400 on the LHC ring. An experimental description of the CMS detector and the LHC is  
401 described in Chapter 3, including some of the object reconstruction used by CMS in  
402 searches for SUSY signatures.

403 The performance of the CMS Level-1 single jet trigger, measured over the course of  
404 the year is also included within this chapter. The Level-1 triggers are of paramount  
405 importance to the recording of physics events at CMS, and to which a change in the jet  
406 seed algorithm was introduced approximately half way through the data taking period.  
407 The aim of this change, was to facilitate a reduction in the rate at which data from  
408 events not of interest to physics analyses were recorded, whilst avoiding impact on those  
409 which were.

410 Chapter 4, contains a description of the search for evidence of the production of Su-  
411 persymmetric particles at the LHC. The main basis of the search centres around the  
412 kinematic dimensionless  $\alpha_T$  variable, which provides strong rejection of backgrounds with  
413 fake missing energy signatures whilst maintaining good sensitivity to a variety of SUSY  
414 topologies. The author's work (as an integral part of the analysis group) is documented  
415 in detail, and which has culminated in numerous publications over the past two years,  
416 the latest results having been published in the European Physical Journal C (EPJC) [5].  
417 The results and interpretations within the framework of a variety of Simplified Model  
418 Spectra (SMS), which describe an array of possible SUSY event topologies is documented  
419 in Chapter 5.

420 The author in particular has played a major role in the extension of the  $\alpha_T$  analysis into  
421 the additional b-tagged and jet multiplicity dimensions increasing the sensitivity of the  
422 analysis to a range of SUSY topologies. Additionally the author has worked extensively  
423 in both increasing the statistical precision of electroweak predictions measured from  
424 simulation through analytical techniques, and the derivation of a data driven systematic

- 425 uncertainty through the establishment of closure tests within the control samples of the  
426 analysis.
- 427 Finally a method to search for **SUSY** signatures rich in top and bottom flavoured jet final  
428 states, is introduced in Chapter 6. These particular **SUSY** topologies are increasingly of  
429 interest to physicists in light of the discovery of the Higgs boson, and is discussed within  
430 the chapter. A parametrisation of the b-tagging distribution for different Electroweak  
431 processes is used to establish template shapes, which are then fitted at low b-tagged jet  
432 multiplicity (0-2), to estimate the expected number of 3 and 4 b-tagged jet events from  
433 **SM** processes. The  $\alpha_T$  event selections are used to test the functionality of this template  
434 method in both data and simulation, with background predictions from the signal region  
435 selection compared to those presented in Chapter 5.
- 436 Natural units are used throughout this thesis in which  $\hbar = c = 1$ .

# Chapter 2.

## <sup>437</sup> A Theoretical Overview

<sup>438</sup> Within this chapter, a brief introduction and background to the **SM** is given. Its success as  
<sup>439</sup> a rigorously tested and widely accepted theory is discussed as are its deficiencies, leading  
<sup>440</sup> to the argument that this theory is not a complete description of our universe. The  
<sup>441</sup> motivations for new physics at the TeV scale and in particular Supersymmetric theories  
<sup>442</sup> are outlined within Section (2.3), with the chapter concluding with how an experimental  
<sup>443</sup> signature of such theories can be produced and observed at the **LHC**, Section (2.4).

### <sup>444</sup> 2.1. The Standard Model

<sup>445</sup> The **SM** is the name given to the relativistic Quantum Field Theory (**QFT**), where  
<sup>446</sup> particles are represented as excitations of fields, which describe the interactions and  
<sup>447</sup> properties of all the known elementary particles [6][7][8][9]. It is a renormalisable field  
<sup>448</sup> theory which contains three symmetries:  $SU(3)$  for colour charge,  $SU(2)$  for weak isospin  
<sup>449</sup> and  $U(1)$  relating to weak hyper charge, which require its Lagrangian  $\mathcal{L}_{SM}$  to be invariant  
<sup>450</sup> under local gauge transformation.

<sup>451</sup> Within the **SM** theory, matter is composed of spin  $\frac{1}{2}$  fermions, which interact with each  
<sup>452</sup> other via the exchange of spin-1 gauge bosons. A summary of the known fundamental  
<sup>453</sup> fermions and bosons is given in Table 2.1.

<sup>454</sup> Fermions are separated into quarks and leptons of which only quarks interact with the  
<sup>455</sup> strong nuclear force. Quarks unlike leptons are not seen as free particles in nature, but  
<sup>456</sup> rather exist only within baryons, composed of three quarks with an overall integer charge,  
<sup>457</sup> and quark-anti-quark pairs called mesons. Both leptons and quarks are grouped into

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].

458 three generations which have the same properties, but with ascending mass in each  
 459 subsequent generation.

460 The gauge bosons mediate the interactions between fermions. The field theories of  
 461 Quantum Electro-Dynamics (QED) and Quantum Chromo-Dynamics (QCD), yield  
 462 massless mediator bosons, the photon and eight coloured gluons which are consequences  
 463 of the gauge invariance of those theories, detailed in Section (2.1.1).

464 The unification of the electromagnetic and weak-nuclear forces into the current Elec-  
 465 troweak theory yield the weak gauge bosons,  $W^\pm$  and Z through the mixing of the  
 466 associated gauge fields. The force carriers of this theory were experimentally detected by  
 467 the observation of weak neutral current, discovered in 1973 in the Gargamelle bubble  
 468 chamber located at European Organization for Nuclear Research (CERN) [10], with the  
 469 masses of the weak gauge bosons measured by the UA1 and U2 experiments at the Super  
 470 Proton Synchrotron (SPS) collider in 1983 [11][12].

### <sup>471</sup> 2.1.1. Gauge Symmetries of the SM

- <sup>472</sup> Symmetries are of fundamental importance in the description of physical phenomena.  
<sup>473</sup> Noether's theorem states that for a dynamical system, the consequence of any symmetry  
<sup>474</sup> is an associated conserved quantity [13]. Invariance under translations, rotations, and  
<sup>475</sup> Lorentz transformations in physical systems lead to conservation of momentum, energy  
<sup>476</sup> and angular momentum.
- <sup>477</sup> In the **SM**, a quantum theory described by Lagrangian formalism, the weak, strong and  
<sup>478</sup> electromagnetic interactions are described in terms of “gauge theories”. A gauge theory  
<sup>479</sup> possesses invariance under a set of “local transformations”, which are transformations  
<sup>480</sup> whose parameters are space-time dependent. The requirement of gauge invariance within  
<sup>481</sup> the **SM** necessitates the introduction of force-mediating gauge bosons and interactions  
<sup>482</sup> between fermions and the bosons themselves. Given the nature of the topics covered by  
<sup>483</sup> this thesis, the formulation of **EWK** within the **SM** Lagrangian is reviewed within this  
<sup>484</sup> section.
- <sup>485</sup> The simplest example of the application of the principle of local gauge invariance within  
<sup>486</sup> the **SM** is in Quantum Electro-Dynamics (**QED**), the consequences of which require a  
<sup>487</sup> massless photon field [14][15].
- <sup>488</sup> Starting from the free Dirac Lagrangian written as

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

- <sup>489</sup> where  $\psi$  represents a free non interacting fermionic field, with the matrices  $\gamma^\mu$ ,  $\mu \in 0, 1, 2, 3$   
<sup>490</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>491</sup> flat space-time metric  $(+, -, -, -)$ , and  $I_4$  the  $4 \times 4$  identity matrix.
- <sup>492</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>493</sup> the kinetic term of the Lagrangian will not remain invariant, due to the partial derivative  
<sup>494</sup> interposed between the  $\bar{\psi}$  and  $\psi$  yielding,

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

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

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

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

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

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

505 This then leads to the Lagrangian of QED:

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

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

### 509    2.1.2. The Electroweak Sector and Electroweak Symmetry 510    Breaking

- 511    The same application of gauge symmetry and the requirement of local gauge invariance  
512    can be used to unify QED and the Weak force in the Electroweak Sector (EWK).  
513    The nature of EWK interactions is encompassed within a Lagrangian invariant under  
514    transformations of the group  $SU(2)_L \times U(1)_Y$ .  
515    The weak interactions from experimental observation [16], are known to violate parity  
516    and are therefore not symmetric under interchange of left and right helicity fermions.  
517    Thus within the SM the left and right handed parts of these fermion fields are treated  
518    separately. A fermion field is then split into two left and right handed chiral components,  
519     $\psi = \psi_L + \psi_R$ , where  $\psi_{L/R} = (1 \pm \gamma^5)\psi$ .  
520    The  $SU(2)_L$  group is the special unitary group of  $2 \times 2$  matrices  $U$  satisfying  $UU^\dagger = I$   
521    and  $\det(U) = 1$ . It may be written in the form  $U = e^{-i\omega_i T_i}$ , with the generators of the  
522    group  $T_i = \frac{1}{2}\tau_i$  where  $\tau_i$ ,  $i \in 1,2,3$  being the  $2 \times 2$  Pauli matrices

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

- 523    which form a non Abelian group obeying the commutation relation  $[T^a, T^b] \equiv if^{abc}T^c \neq 0$ .  
524    The gauge fields that accompany this group are represented by  $\hat{W}_\mu = (\hat{W}_\mu^1, \hat{W}_\mu^2, \hat{W}_\mu^3)$   
525    and act only on the left handed component of the fermion field  $\psi_L$ .  
526    One additional generator  $Y$  which represents the hypercharge of the particle under  
527    consideration is introduced through the  $U(1)_Y$  group acting on both components of the  
528    fermion field, with an associated vector boson field  $\hat{B}_\mu$ .  
529    The  $SU(2)_L \times U(1)_Y$  transformations of the left and right handed components of  $\psi$  are  
530    summarised by,

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

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

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

533 Imposing local gauge invariance within  $\mathcal{L}_{EWK}$  is once again achieved by modifying the  
 534 covariant derivative

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

558 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  
 559 the **VEV**.

560 Defining the ground state of the  $\phi$  field to be consistent with the  $V(\phi)$  minimum, and  
 561 then expanding around a ground state chosen to maintain an unbroken electromagnetic  
 562 symmetry thus preserving a zero photon mass [21] leads to

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

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

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

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

568 where the relations between the physical and electroweak gauge fields from Equation  
 569 (2.14) are used. The  $W^\pm$  and Z boson masses can then be determined to be

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

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

## 576 2.2. Motivation for Physics Beyond the Standard 577 Model

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

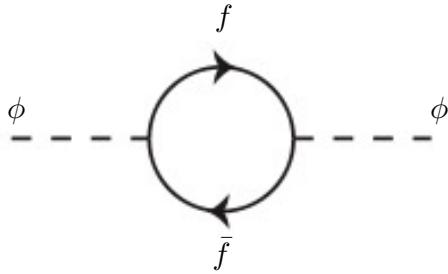
583 On a theoretical level, the **SM** is unable to incorporate the gravitational interactions of  
584 fundamental particles within the theory. Whilst at the electroweak energy scales the  
585 relative strength of gravity is negligible compared to the other three fundamental forces,  
586 at much higher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

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

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

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

603 In the absence of new physics between the TeV and Planck scale, calculating beyond  
604 tree-level contributions to the Higgs mass term given by its self interaction, result in  
605 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.

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

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

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

611 To recover the mass of the now discovered Higgs boson would require a fine-tuning of  
 612 the parameters to cancel out these mass corrections of the Higgs mass to the scale of  
 613 30 orders of magnitude. This appears as an unnatural solution to physicists and it is  
 614 this hierarchy problem that provides one of the strongest motivations for the theory of  
 615 SUperSYmmetry (**SUSY**).

### 616 2.3. Supersymmetry Overview

617 Supersymmetry provides potential solutions to many of the issues raised in the previous  
 618 section. It provides a dark matter candidate, can explain baryogenesis in the early  
 619 universe and also provides an elegant solution to the hierarchy problem [26][27][28][29].  
 620 At its heart it represents a new space-time symmetry that relates fermions and bosons.  
 621 This symmetry converts bosonic states into fermionic states, and vice versa, see Equation  
 622 (2.21) ,

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

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

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

632 Each particle in a supersymmetric theory is paired together with their superpartners as  
 633 a result of these supersymmetric transformations in a so called supermultiplet. These  
 634 superpartners will then consequently also contribute to the corrections to the Higgs mass.  
 635 Bosonic and fermionic loops contributing to the correction appear with opposite signs,  
 636 and therefore cancellation of these divergent terms will stabilise the Higgs mass, solving  
 637 the hierarchy problem [30][31].

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

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

### 649 2.3.1. R-Parity

650 Some supersymmetric theories also present a solution to the dark matter problem. These  
 651 theories contain a Lightest Supersymmetric Partner (**LSP**), which matches the criteria  
 652 of a Weakly Interacting Massive Particle (**WIMP**) required by cosmological observation  
 653 when R-parity is conserved.

654 Baryon (B) and Lepton (L) number conservation is forbidden in the **SM** by renormal-  
 655 isability requirements. The violation of Baryon or Lepton number results in a proton  
 656 lifetime much shorter than those set by experimental limits [37]. Another symmetry  
 657 called R-parity is then often introduced to **SUSY** theories to maintain baryon and lepton  
 658 conservation.

659 R-parity is described by the equation

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

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

662 R-parity ensures the stability of the proton in **SUSY** models, and also has other conse-  
 663 quences for the production and decay of supersymmetric particles. In particle colliders  
 664 supersymmetric particles can only be pair produced, and similarly the decay of any pro-  
 665 duced supersymmetric particle is restricted to a **SM** particle and a lighter supersymmetric  
 666 particle as allowed by conservation laws. A further implication of R-parity is that once a  
 667 supersymmetric particle has decayed to the **LSP** it remains stable, unable to decay into  
 668 a **SM** particle.

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

## 672 2.4. Experimental Signatures of **SUSY** at the **LHC**

673 Should strongly interacting sparticles be within the experimental reach of the **LHC**, then  
 674 it is expected that they can be produced in a variety of ways :

- 675     • squark/anti-squark and gluino pairs can be produced via both gluon fusion and
- 676        quark/anti-quark scattering,
- 677     • a gluino and squark produced together via quark-gluon scattering,
- 678     • squark pairs produced via quark-quark scattering.

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

682 During the 2011 run period at  $\sqrt{s} = 7$  TeV, particular models were used to benchmark  
683 performance and experimental reach of both **CMS** searches and previous experiments.  
684 The Compressed Minimal SuperSymmetric Model (**CMSSM**) was initially chosen for  
685 a number of reasons [38], one of the most compelling being the reduction of the up to  
686 105 new parameters that can be introduced by **SUSY** (in addition to the existing 19 of  
687 the **SM**), to just 5 free extra free parameters. It was this simplicity, combined with the  
688 theory not requiring any fine tuning of particle masses to produce experimentally verified  
689 **SM** observables that made it an attractive model to interpret physics results.

690 However recent results from the **LHC** now strongly disfavour large swathes of **CMSSM**  
691 parameter space [39][40][41]. In the face of such results a more pragmatic model indepen-  
692 dent search strategy is now applied across most **SUSY** searches at the **LHC**, see Section  
693 (2.4.1).

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

701 The **SM** contains processes which can exhibit a similar event topology to that described  
702 above, with the largest contribution coming from the general QCD environment of a  
703 hadron collider. A multitude of different analytical techniques are used by experimental  
704 physicists to reduce or estimate any reducible or irreducible backgrounds, allowing a  
705 possible **SUSY** signature to be extracted. The techniques employed within this thesis are  
706 described in great detail within Section (4.1).

707 **2.4.1. Simplified models**

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

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

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

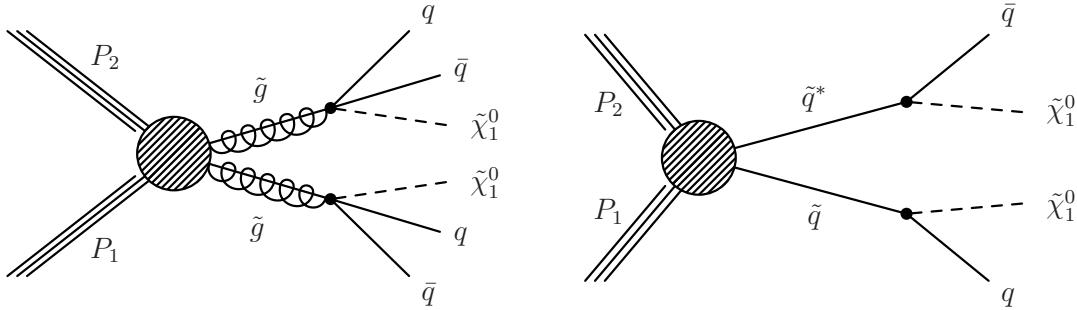
- 723 • **T1,T1xxxx**, models represent a simplified version of gluino pair production with  
724 each gluino (superpartner to the gluon) undergoing a three-body decay to a quark-  
725 antiquark pair and the **LSP** (i.e.  $\tilde{g} \rightarrow q\bar{q}\tilde{\chi}_{LSP}$ ). The resultant final state from this  
726 decay is typically 4 jets +  $\cancel{E}_T$  in the absence of initial/final state radiation and  
727 detector effects. xxxx denotes models in which the quarks are of a specific flavour,  
728 typically t or b quark-antiquarks.
- 729 • **T2,T2xx**, models represent a simplified version of squark anti-squark production  
730 with each squark undergoing a two-body decay into a light-flavour quark and **LSP**  
731 (i.e.  $\tilde{q} \rightarrow q\tilde{\chi}_{LSP}$ ). This results in final states with less jets than gluino mediated  
732 production, typically 2 jets +  $\cancel{E}_T$  when again ignoring the effect of initial/final state  
733 radiation and detector effects. xx models represent decays in which both the quark  
734 and the squark within the decay is of a specific flavour, which in this thesis are  
735 again  $\tilde{t}/t$  or  $\tilde{b}/b$ .

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

<sup>740</sup> supersymmetric theories by requiring a light right handed top squark,  $\tilde{t}_R$ , and left-handed  
<sup>741</sup> double  $SU(2)_L$  doublet containing top and bottom squarks,  $(\tilde{t}, \tilde{b})_L$  [47].

<sup>742</sup> These theories therefore solve the hierarchy problem by predicting light  $\sim$  EWK scale  
<sup>743</sup> third generation sleptons, to be accessible at the LHC. Search strategies involving the  
<sup>744</sup> requirement of b-tagging (see Section (3.3.2)) are used to give sensitivity to these type of  
<sup>745</sup> SUSY scenarios and are discussed in greater detail within Chapter 4.

<sup>746</sup> Two example decay chains are shown in Figure 2.2; the pair production of gluinos (T1)  
<sup>747</sup> and the pair production of squarks (T2) decaying into SM particles and LSP's.



**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>748</sup> The LHC And The CMS Detector

<sup>749</sup> Probing the SM for signs of new physics would not be possible without the immensely  
<sup>750</sup> complex electronics and machinery that makes the TeV energy scale accessible to physi-  
<sup>751</sup> cists for the first time. This chapter will introduce both the LHC based at European  
<sup>752</sup> Organization for Nuclear Research (CERN) and the Compact Muon Solenoid (CMS)  
<sup>753</sup> detector (of which the author is a member). Section (3.2) serves to present an overview of  
<sup>754</sup> the different components of the CMS detector, with specific components relevant to the  
<sup>755</sup> search for supersymmetric particles described in greater detail. Section (3.3) will focus on  
<sup>756</sup> event and object reconstruction again with more emphasis on jet level quantities which  
<sup>757</sup> are most relevant to the author’s analysis research. Finally Section (3.4) will describe and  
<sup>758</sup> detail the service work for the CMS Collaboration performed by the author, in measuring  
<sup>759</sup> the performance of L1 single jet triggers in the Global Calorimeter Trigger (GCT) during  
<sup>760</sup> the 2012-2013 run period.

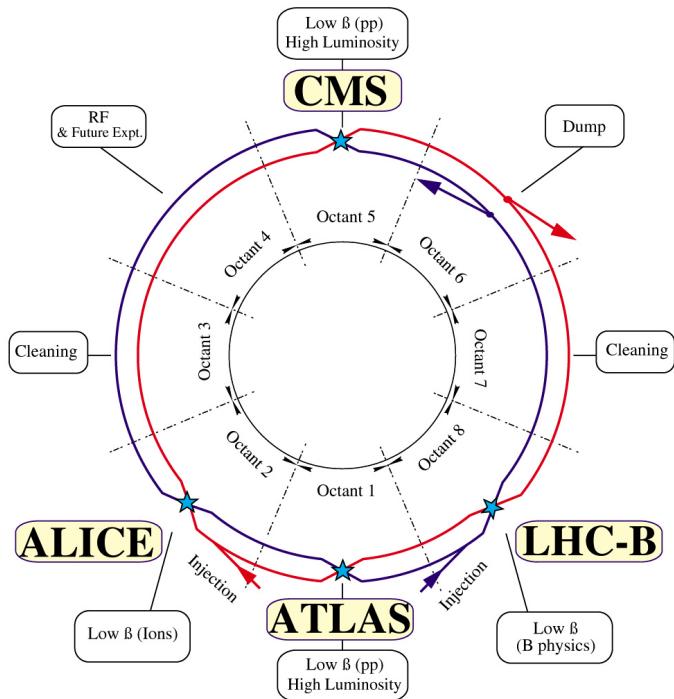
### <sup>761</sup> 3.1. The LHC

<sup>762</sup> The LHC is a storage ring, accelerator, and collider of circulating beams of protons or ions.  
<sup>763</sup> Housed in the tunnel dug for Large Electron-Positron Collidor (LEP), it is approximately  
<sup>764</sup> 27 km in circumference, 100 m underground, and straddles the border between France  
<sup>765</sup> and Switzerland outside of Geneva. It is currently the only collider in operation that  
<sup>766</sup> is able to study physics at the TeV scale. A double-ring circular synchrotron, it was  
<sup>767</sup> designed to collide both proton-proton (pp) and heavy ion (PbPb) with a centre of mass  
<sup>768</sup> energy  $\sqrt{s} = 14$  TeV at a final design luminosity of  $10^{34}\text{cm}^{-2}\text{s}^{-1}$ .

<sup>769</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.

781



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

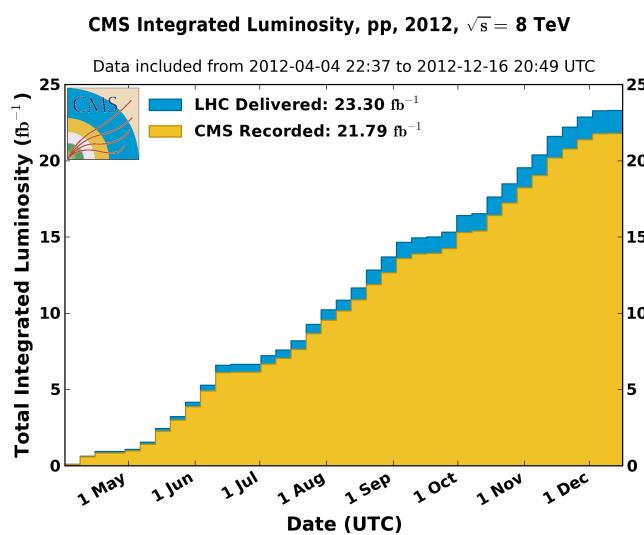
Proton beams are formed inside the Proton Synchrotron (**PS**) from bunches of protons 50 ns apart with an energy of 26 GeV. The protons are then accelerated in the Super Proton Synchrotron (**SPS**) to 450 GeV before being injected into the **LHC**. These **LHC** proton beams consists of many “bunches” (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

787 4 TeV (2012) per beam in a process involving increasing the current passing through the  
 788 dipole magnets. Once the desired  $\sqrt{s}$  energy is reached then the beams are allowed to  
 789 collide at the interaction points. The luminosity falls regularly as the run progresses as  
 790 protons are lost in collisions, and eventually the beam is dumped before repeating the  
 791 process again.

792

793 Colliding the beams produced an instantaneous luminosity of approximately  $5 \times 10^{33}$   
 794  $\text{cm}^{-2}\text{s}^{-1}$  during the 2012 run. The high number of protons in each bunch increases  
 795 the likelihood of multiple interactions with each crossing of the counter-circulating  
 796 beams. This leads to isotropic energy depositions within the detectors positioned at these  
 797 interaction points, increasing the energy scale of the underlying event. This is known as  
 798 pile-up and the counteracting of it's effects are important to the many measurements  
 799 performed at the **LHC**.

800 In the early phase of prolonged operation after the initial shutdown the machine operated  
 801 in 2010-2011 at 3.5 TeV per beam,  $\sqrt{s} = 7$  TeV, delivering  $6.13 \text{ fb}^{-1}$  of data [53]. During  
 802 the 2012-2013 run period, data was collected at an increased  $\sqrt{s} = 8$  TeV improving the  
 803 sensitivity of searches for new physics. Over the whole run period  $23.3 \text{ fb}^{-1}$  of data was  
 804 delivered, of which  $21.8 \text{ fb}^{-1}$  was recorded by the **CMS** detector as shown in Figure 3.2  
 805 [53]. A total of  $12 \text{ fb}^{-1}$  of 8 TeV certified data was collected by October 2012, and it is  
 806 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  $p\bar{p}$  runs.

---

## 807 3.2. The CMS Detector

808 The Compact Muon Solenoid (**CMS**) detector is one of two general purpose detectors  
 809 at the **LHC** designed to search for new physics. The detector is designed to provide  
 810 efficient identification and measurement of many physics objects including photons,  
 811 electrons, muons, taus, and hadronic showers over wide ranges of transverse momentum  
 812 and direction. It's nearly  $4\pi$  coverage in solid angle allows for accurate measurement of  
 813 global transverse momentum imbalance. These design factors give **CMS** the ability to  
 814 search for direct production of **SUSY** particles at the TeV scale, making the search for  
 815 Supersymmetric particles one of the highest priorities among the wide range of physics  
 816 programmes at **CMS**.

817

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

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

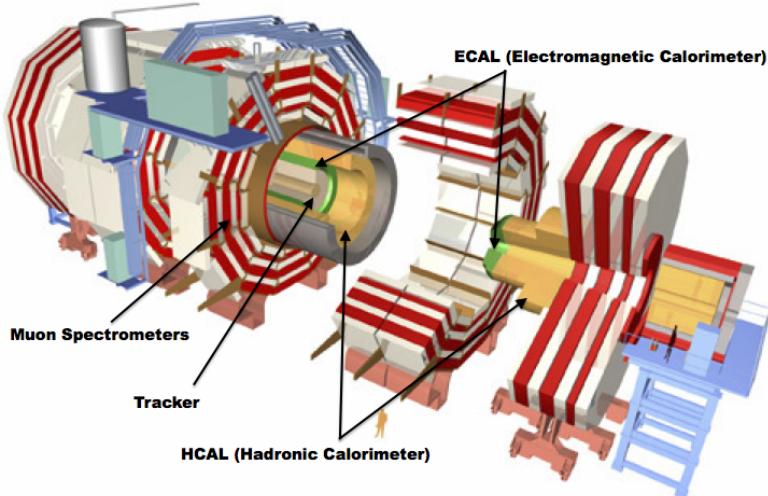
824 The variable  $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$  is commonly used to define angular distance between  
 825 objects within the detector and additionally energy and momentum is typically measured  
 826 in the transverse plane perpendicular to the beam line. These values are calculated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

### 829 3.2.1. Detector subsystems

830 As the range of particles produced from  $pp$  collisions interact in different ways with  
 831 matter, **CMS** is divided into sub-detector systems, which perform complementary roles  
 832 to identify the identity, mass, and momentum of the different physics objects present in  
 833 each event. These detector sub-systems contained within **CMS** are wrapped in layers

around a central 13 m long 4 T super conducting solenoid as shown in Figure 3.3. With the endcaps closed, CMS is a cylinder of length 22 m, diameter 15 m, and mass 12.5 kilotons. A more detailed complete description of the detector can be found elsewhere [50].

837



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

### 838 3.2.2. Tracker

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

849

The tracking system also plays a crucial part in the identification of jets that originate from b-quarks through the measurement of displaced secondary vertices. The methods in which these b-flavoured jets are identified are discussed within Section (3.3.2). The

853 identification of b-jets is important in many searches for natural SUSY models and forms  
854 an important part of the inclusive search strategy described within Section (4.2).

855 **3.2.3. Electromagnetic calorimeter**

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

861

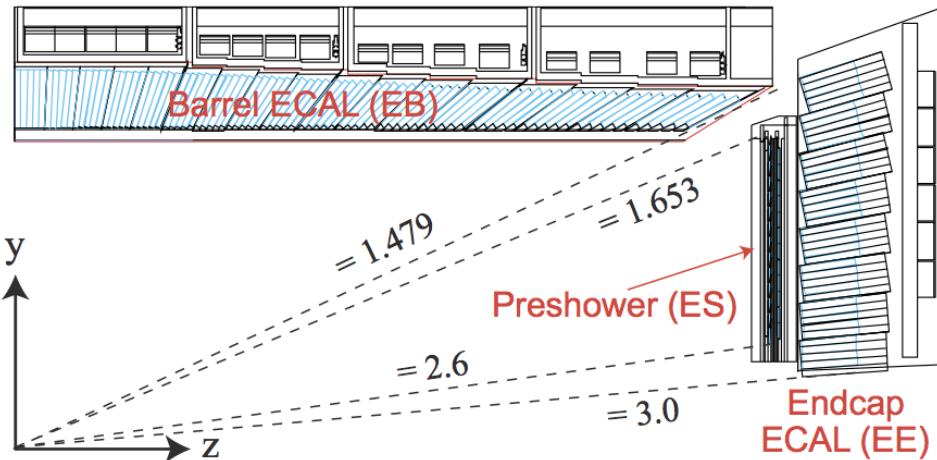
862 Lead tungstate is used because of its short radiation length ( $X_0 \sim 0.9$  cm) and small  
863 Molieré radius ( $\sim 2.1$  cm) leading to high granularity and resolution. It's fast scintillation  
864 time ( $\sim 25$  ns) reduces the effects of pile-up, which occurs when energy from previous  
865 collisions are still being read out, and its radiation hardness gives it longevity. The  
866 crystals are arranged in modules which surround the beam line in a non-projective  
867 geometry, angled at  $3^\circ$  with respect to the interaction point to minimise the risk of  
868 particles escaping down the cracks between the crystals.

869

870 The **ECAL** is primarily composed of two sections, the Electromagnetic CALorime-  
871 ter Barrel (**EB**) which extends in pseudo-rapidity to  $|\eta| < 1.479$  with a crystal front  
872 cross section of  $22 \times 22$  mm and a length of 230 mm corresponding to 25.8 radiation  
873 lengths, and the Electromagnetic CALorimeter Endcap (**EE**) covering a rapidity range  
874 of  $1.479 < |\eta| < 3.0$ , which consists of two identical detectors on either side of the  
875 **EB**. A lead-silicon sampling ‘pre-shower’ detector Electromagnetic CALorimeter pre-  
876 Shower (**ES**) is placed before the endcaps to aid in the identification of neutral pions.  
877 Their arrangement is shown in Figure 3.4.

878

879 Scintillation photons from the lead tungstate crystals are instrumented with Avalanche  
880 Photo-Diodes (**APD**) and Vacuum Photo-Triodes (**VPT**) located in the **EB** and **EE**  
881 respectively, converting the scintillating light into an electric signal which is consequently  
882 used to determine the amount of energy deposited within the crystal . These instruments  
883 are chosen for their resistance under operation to the strong magnetic field of **CMS**. The  
884 scintillation of the **ECAL** crystals as well as the response of the **APDs** varies as a function



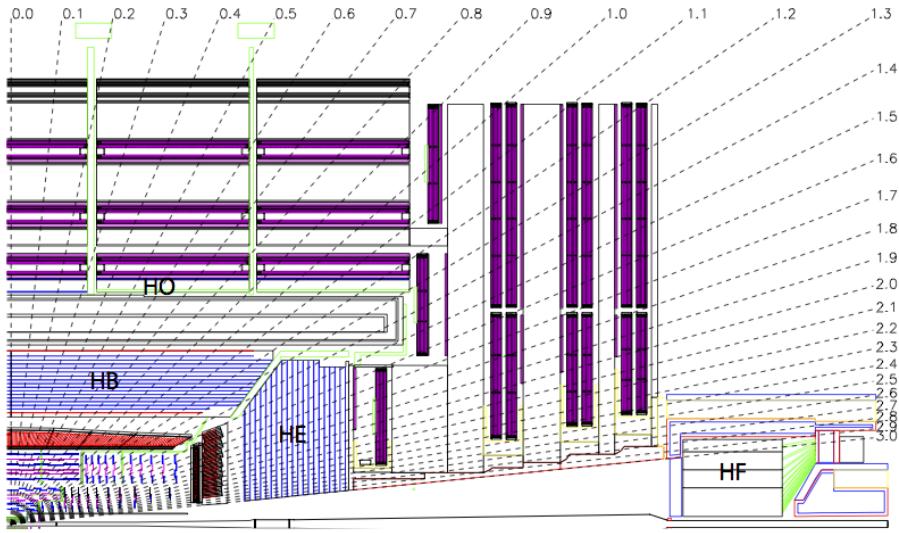
**Figure 3.4.:** Illustration of the **ECAL** showing the arrangement of the lead tungstate crystals in the **EB** and **EE**. The **ES** is also shown and is located in front of the **EE** [55].

885 of temperature and so cooling systems continually maintain an overall constant **ECAL**  
 886 temperature  $\pm 0.05^\circ\text{C}$ .

### 887 3.2.4. Hadronic calorimeter

888 Beyond the **ECAL** lies the Hadronic CALorimeter (**HCAL**) which is responsible for  
 889 the accurate measurement of hadronic showers, crucial for analyses involving jets or  
 890 missing energy signatures. The **HCAL** is a sampling calorimeter which consists of al-  
 891 ternating layers of brass absorber and plastic scintillator, except in the hadron forward  
 892 ( $3.0 < |\eta| < 5.0$ ) region in which steel absorbers and quartz fibre scintillators are used  
 893 because of their increased radiation tolerance. Hadron showers are initiated in the  
 894 absorber layers inducing scintillation in the plastic scintillator tiles. These scintillation  
 895 photons are converted by wavelength shifting fibres for read-out by hybrid photodiodes.  
 896

897 The **HCAL**'s size is constrained to a compact size by the presence of the solenoid, re-  
 898 quiring the placement of an additional outer calorimeter on the outside of the solenoid  
 899 to increase the sampling depth of the **HCAL**. A schematic of the **HCAL** can be seen in  
 900 Figure 3.5.



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

902 The **HCAL** covers the range  $|\eta| < 5$  and consists of four sub-detectors: the Hadron  
 903 Barrel (**HB**)  $|\eta| < 1.3$ , the Hadron Outer (**HO**), the Hadron Endcaps (**HE**)  $1.3 < |\eta| < 3.0$   
 904 and the Hadron Forward (**HF**). The **HB**, contained between the outer edge of the **ECAL**  
 905 and the inner edge of the solenoid is formed of 36 azimuthal wedges which are split  
 906 between two half-barrel segments. Each wedge is segmented into four azimuthal angle  
 907 ( $\phi$ ) sectors, and each half-barrel is further segmented into 16  $\eta$  towers. The electronic  
 908 readout chain, channels the light from the active scintillator layers from one  $\phi$ -segment  
 909 and all  $\eta$ -towers of a half-barrel to a Hybrid Photo Diode (**HPD**).

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

918 **3.2.5. Muon systems**

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

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

929

930 **3.3. Event Reconstruction and Object Definition**

931 The goal of event reconstruction is to take the raw information recorded by the detector  
932 and to compute from it higher-level quantities which can be used at an analysis level.  
933 These typically correspond to an individual particle’s energy and momenta, or groups of  
934 particles which shower in a narrow cone and the overall global energy and momentum  
935 balance of the event. The reconstruction of these objects are described in great detail in  
936 [57], however covered below are brief descriptions of those which are most relevant to the  
937 analysis detailed in Chapter 4.

938 **3.3.1. Jets**

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

944 At CMS jets are reconstructed from energy deposits in the detector using the anti-kt  
945 algorithm [58] with size parameter  $\Delta R = 0.5$ . The anti-kt jet algorithm clusters jets by  
946 defining a distance between hard (high  $p_T$ ) and soft (low  $p_T$ ) particles such that soft

947 particles are preferentially clustered with hard particles before being clustered between  
948 themselves. This produces jets which are robust to soft particle radiation from the pile-up  
949 conditions produced by the **LHC**.

950

951 There are two main type of jet reconstruction used at **CMS**, Calorimeter (Calo) and  
952 Particle Flow (PF) jets [59]. Calorimeter jets are reconstructed using both the **ECAL**  
953 and **HCAL** cells, combined into calorimeter towers . These calorimeter towers consist of  
954 geometrically matched **HCAL** cells and **ECAL** crystals. Electronics noise is suppressed by  
955 applying a threshold to the calorimeter cells, with pile-up effects reduced by a requirement  
956 placed on the tower energy [60]. Calorimeter jets are the jets used within the analysis  
957 presented in this thesis.

958 PF jets are formed from combining information from all of the **CMS** sub-detectors systems  
959 to determine which final state particles are present in the event. Generally, any particle  
960 is expected to produce some combination of a track in the silicon tracker, a deposit in  
961 the calorimeters, or a track in the muon system. The PF jet momentum and spatial  
962 resolutions are greatly improved with respect to calorimeter jets, as the use of the tracking  
963 detectors and of the high granularity of **ECAL** allows resolution and measurement of  
964 charged hadrons and photons inside a jet, which together constitute  $\sim 85\%$  of the jet  
965 energy [61].

966 The jets reconstructed by the clustering algorithm in **CMS** typically have an energy  
967 that differs to the ‘true’ energy measured by a perfect detector. This stems from the  
968 non-linear and nonuniform response of the calorimeters as well as other residual effects  
969 including pile-up and underlying events, and therefore additional corrections are applied  
970 to recover a uniform relative response as a function of pseudo-rapidity. These are applied  
971 as separate sub corrections [62].

- 972     • A pile-up correction is first applied to the jet. It subtracts the average extra energy  
973         deposited in the jet that comes from other vertices present in the event and is  
974         therefore not part of the hard jet itself.
- 975     •  $p_T$  and  $\eta$  dependant corrections derived from Monte Carlo simulations are used to  
976         account for the non-uniform response of the detector.
- 977     •  $p_T$  and  $\eta$  residual corrections are applied to data only to correct for difference  
978         between data and Monte Carlo. The residual is derived from QCD di-jet samples  
979         and the  $p_T$  residual from  $\gamma+$  jet and  $Z+$  jets samples in data.

---

### 980 3.3.2. B-tagging

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

985

986 Many different algorithms developed by CMS select b-quark jets based on variables such  
 987 as the impact parameters of the charged-particle tracks, the properties of reconstructed  
 988 decay vertices, and the presence or absence of a lepton, or combinations thereof [63].  
 989 One of the most efficient of which is the Combined Secondary Vertex (CSV) which  
 990 operates based on secondary vertex and track-based lifetime information, benchmarked  
 991 in ‘Loose’, ‘Medium’ and ‘Tight’ working points, of which the medium point is the tagger  
 992 used within the  $\alpha_T$  search presented in Section (4.1). All figures within this sub-section,  
 993 demonstrating the performance of this b-tagging algorithm are taken from [64].

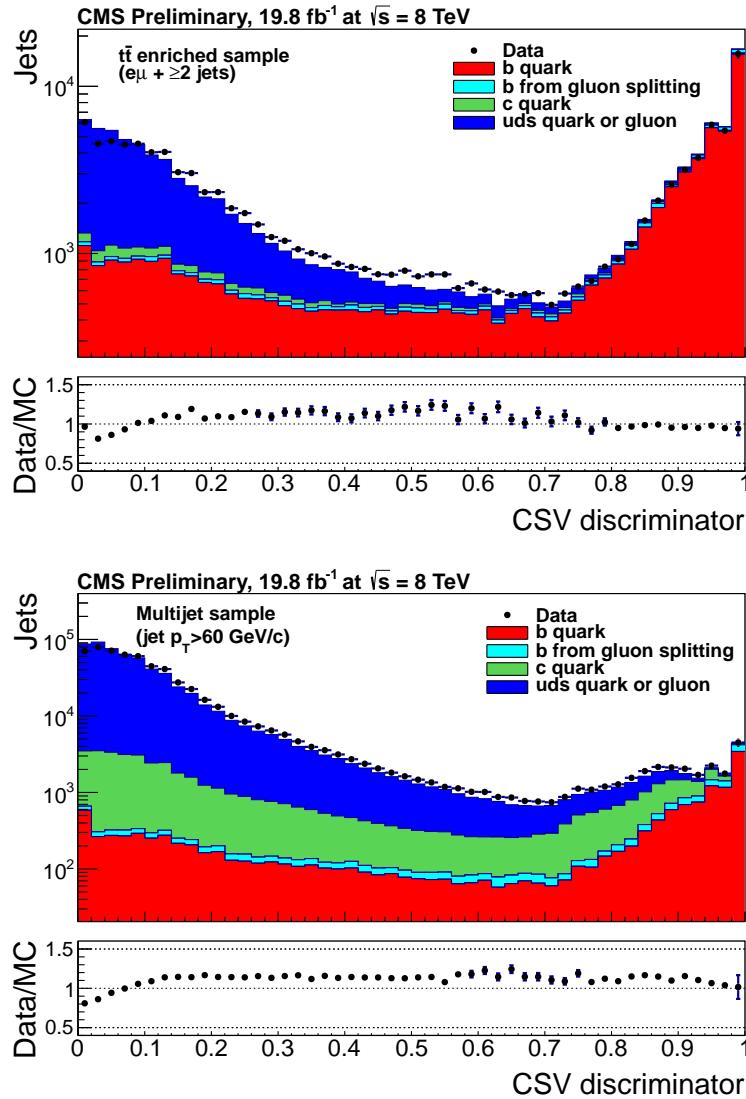
994 Using the CSV tagger, a likelihood-based discriminator distinguishes between jets from  
 995 b-quarks, and those from charm or light quarks and gluons, which is shown in Figure 3.6.  
 996 The minimum thresholds on the discriminator for each working point correspond to the  
 997 mis-identification probability for light-parton jets of 10%, 1%, and 0.1%, respectively, in  
 998 jets with an average  $p_T$  of about 80 GeV.

999 The b-tagging performance is evaluated to measure the b-jet tagging efficiency  $\epsilon_b$ , and the  
 1000 misidentification probability of charm  $\epsilon_c$  and light-parton jets  $\epsilon_s$ . The tagging efficiencies  
 1001 for each of these three jet flavours are compared between data and MC simulation, from  
 1002 which a series of  $p_T$  and  $|\eta|$  binned jet corrections are determined,

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

1003 These are collectively named ‘Btag Scale Factors’ and allow MC simulation to accu-  
 1004 rately reflect the running conditions and performance of the tagging algorithm in data.  
 1005 Understanding of the b-tagging efficiency is essential in order to minimise systematic  
 1006 uncertainties in physics analyses that employ b-tagging.

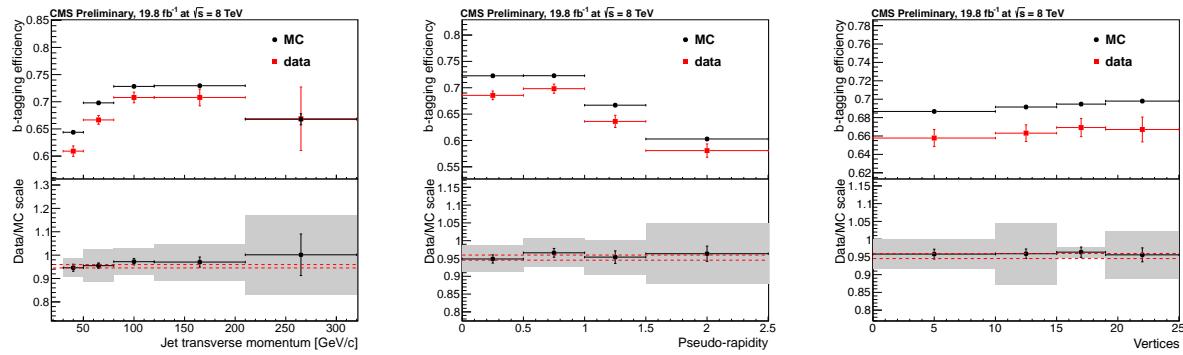
1007



**Figure 3.6.:** CSV algorithm discriminator values in enriched ttbar (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.

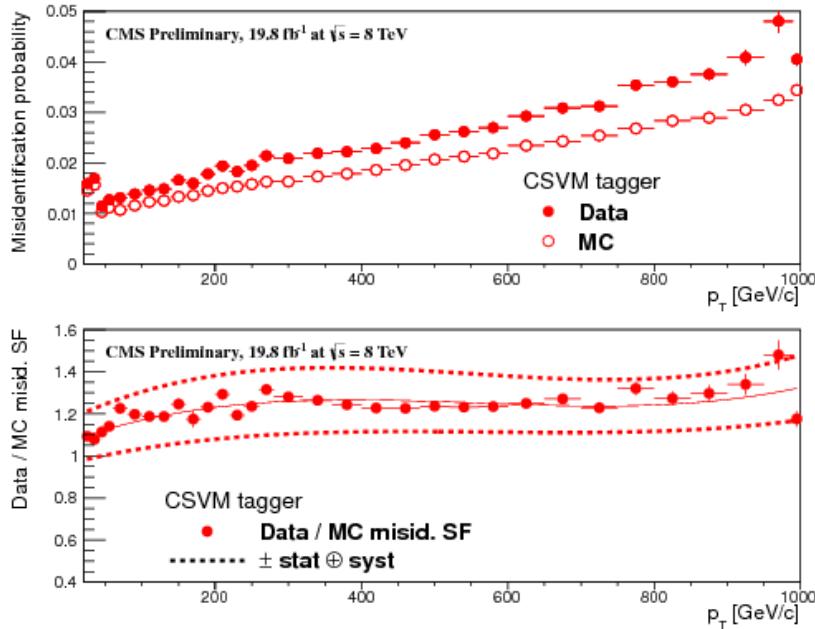
The b-tagging efficiency is measured in data using several methods applied to multi jet events, primarily based on a sample of jets enriched in heavy flavour content. One method requires the collection of events with a soft muon within a cone  $\Delta R < 0.4$  around the jet axis. Because the semi-leptonic branching fraction of b hadrons is significantly larger than that for other hadrons, these jets are more likely to arise from b quarks than from another flavour, with the resultant momentum component of the muon transverse to the jet axis larger for muons from b-hadron decays than from light or charm jets.

1015 Additionally the performance of the tagger can also be benchmarked in  $t\bar{t}$  events where  
 1016 in the SM, the top quark is expected to decay to a W boson and a b quark about 99.8%  
 1017 of the time [1]. Further selection criteria is applied to these events to further enrich the  
 1018 b quark content of these events. The methods to identify b-jets in data are discussed  
 1019 in great detail at [65]. The jet flavours are determined in simulation using truth level  
 1020 information and are compared to data to determine the correction scale factors ( $SF_b$ ),  
 1021 which 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.

1022 The measurement of the misidentification probability for light-parton jets relies on the  
 1023 inversion of tagging algorithms, selecting non-b jets using the same variables and tech-  
 1024 niques used in benchmarking the b-tagging efficiency. The scale factors ( $SF_s$ ) to be  
 1025 applied to MC are shown in Figure 3.8 for the CSVM tagger.



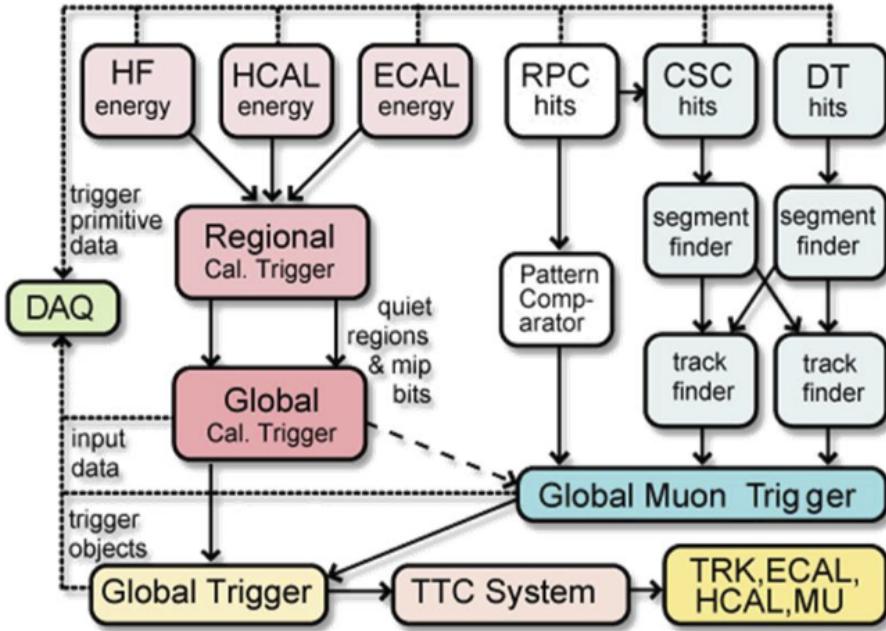
**Figure 3.8.:** For the **CSVM** tagging criterion: (top) misidentification probability in data (filled circles) and simulation (open circles); (bottom) scale factor for the misidentification probability. The last  $p_T$  bin in each plot includes all jets with  $p_T > 1000$  GeV. The solid curve is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

## 1026 3.4. Triggering System

1027 With bunch crossings separated by just 25 ns, the rate at which data from all collisions  
 1028 would have to be written out and processed would be unfeasible. A two-tiered triggering  
 1029 system is applied at **CMS** in order to cope with the high collision rate of protons. The  
 1030 **CMS** trigger is designed to use limited information from each event to determine whether  
 1031 to record the event, reducing the rate of data taking to manageable levels whilst ensuring  
 1032 a high efficiency of interesting physics object events are selected.

1033 The **L1** is a pipelined, dead-timeless system based on custom-built electronics [66], and is  
 1034 a combination of several sub systems which is shown pictorially in Figure 3.9. The L1  
 1035 system is covered in more detail within the following section along with a description  
 1036 of the service work undertaken by the author to benchmark the performance of the L1  
 1037 calorimeter trigger during the 2012 8 TeV run period.

1038 The Higher Level Trigger (**HLT**) is a large farm of commercial computers [67]. The **HLT**  
 1039 processes events with software reconstruction algorithms that are more detailed, giving  
 1040 performance more similar to the reconstruction used offline. The **HLT** reduces the event



**Figure 3.9.:** The CMS L1 Trigger system.

<sup>1041</sup> rate written to disk by a factor of  $\sim 500$  ( $\sim 200\text{Hz}$ ). The recorded events are transferred  
<sup>1042</sup> from CMS to the CERN computing centre, where event reconstruction is performed, and  
<sup>1043</sup> then distributed to CMS computing sites around the globe for storage and analysis.

### <sup>1044</sup> 3.4.1. The Level-1 trigger

<sup>1045</sup> The L1 trigger reduces the rate of events collected from 40 MHz to  $\sim 100\text{ kHz}$  using  
<sup>1046</sup> information from the calorimeters and muon chambers, but not the tracker. A tree  
<sup>1047</sup> system of triggers is used to decide whether to pass on an event to the HLT for further  
<sup>1048</sup> reconstruction. Firstly the calorimeter and muon event information is kept separate, with  
<sup>1049</sup> local reconstruction of objects ( $\mu$ ,  $e$ ,  $\gamma$ , and jets) performed by the Regional Calorimeter  
<sup>1050</sup> Trigger (RCT) and Regional Muon Trigger (RMT) respectively. The RCT generates up to  
<sup>1051</sup> 72 isolated and non-isolated electromagnetic objects. These are sorted by rank, which is  
<sup>1052</sup> equivalent to transverse energy  $E_T$ , with the four highest ranked electromagnetic objects  
<sup>1053</sup> being passed via the Global Calorimeter Trigger (GCT) and Global MuonTrigger (GMT)  
<sup>1054</sup> to the Global Trigger (GT).

1055 In the L1 **GCT**, coarse measurements of the energy deposited in the electromagnetic and  
1056 hadronic calorimeters are combined, and by using sophisticated algorithms the following  
1057 physics objects are formed:

- 1058 • isolated and non-isolated electromagnetic objects ( $e$  and  $\gamma$ );
- 1059 • hadronic jets in the central and forward sections of the hadronic calorimeters;
- 1060 • hadronically decaying tau leptons;
- 1061 • total transverse energy ( $E_T$ ), the scalar sum of the energy measured at L1, and  
1062 missing transverse energy ( $\cancel{E}_T$ ), defined as the vector sum of the energy of L1  
1063 objects;
- 1064 • total transverse jet energy ( $H_T$ ), the scalar sum of the energy of all L1 jet objects,  
1065 and missing transverse jet energy ( $\cancel{H}_T$ ), defined as the vector sum of the energy of  
1066 L1 jets, are calculated from uncorrected L1 jets.

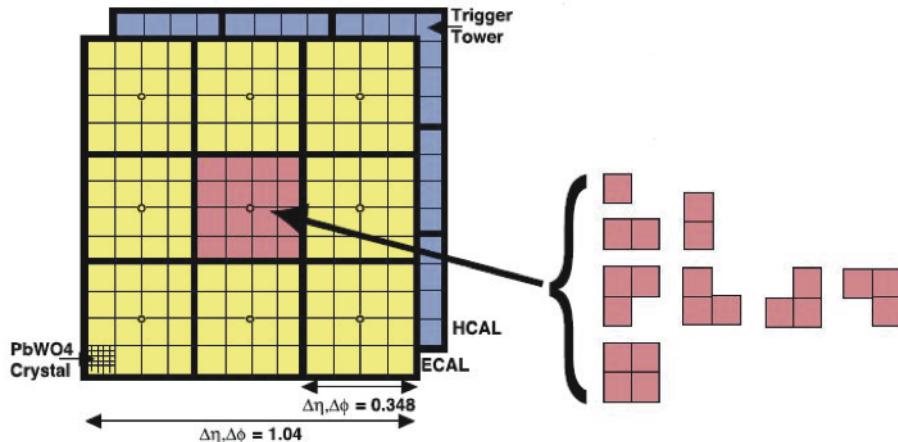
1067 In addition quantities suitable for triggering minimum bias events, forward physics and  
1068 beam background events are calculated. Additionally relevant muon isolation information  
1069 is also passed on to the **GMT** for decisions involving the muon triggers where it is  
1070 combined with information from across the three muon sub-systems. The resultant final  
1071 accept/reject decision at **L1** is then performed by the **GT** based on the objects received  
1072 from the **GCT** and **GMT** ( $e/\gamma$ ,  $\mu$ , jets,  $E_T$ ,  $\cancel{E}_T$ ,  $H_T$ ).

1073 The L1 trigger is therefore of upmost importance to the functioning of the detector.  
1074 Without a high-performing trigger and a good understanding of its performance, there  
1075 would be no data to analyse. Observations of how the L1 trigger performance is affected  
1076 by changing **LHC** running conditions over the 2012 run period and also the introduction  
1077 of a jet seed threshold to the L1 jet trigger algorithm is presented in the following Sections  
1078 (3.4.2 - 3.4.6).

### 1079 3.4.2. The L1 trigger jet algorithm

1080 The L1 jet trigger uses the transverse energy sums computed in the calorimeter (both  
1081 hadronic and electromagnetic) trigger regions. Each region consists of  $4 \times 4$  trigger tower  
1082 windows, spanning a region of  $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$  in pseudorapidity-azimuth. The  
1083 jet trigger uses a  $3 \times 3$  calorimeter region (112 trigger towers) sliding window technique  
1084 which spans the full  $(\eta, \phi)$  coverage of the **CMS** calorimeter as shown in Figure 3.10.

In forming a L1 jet it is required that the central region to be higher than the eight neighbouring regions  $E_T$  central >  $E_T$  surround. Additionally a minimum threshold of 5 GeV on  $E_T$  central was introduced during the 2012 run period to suppress noise from pile-up. A comparison between these two configurations is shown in Section (3.4.4).  
The L1 jets are characterised by the  $E_T$ , summed over the  $3 \times 3$  calorimeter regions, which corresponds to  $12 \times 12$  trigger towers in barrel and endcap or  $3 \times 3$  larger HF towers in the HF. The  $\phi$  size of the jet window is the same everywhere, whilst the  $\eta$  binning gets somewhat larger at high  $\eta$  due to calorimeter and trigger tower segmentation.  
The jets are labelled by the  $(\eta, \phi)$  indices of the central calorimeter region.  
Jets with  $|\eta| > 3.0$  are classified as forward jets, whereas those with  $|\eta| < 3.0$  are classified as central. The four highest energy central, forward and  $\tau$  jets in the calorimeter are passed through Look Up Table (LUT)'s, which apply a programmable  $\eta$ -dependent jet energy scale correction. These are then used to make L1 trigger decisions.



**Figure 3.10.:** Illustration of the Level-1 jet finding algorithm.

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

---

<sub>1103</sub> **3.4.3. Measuring L1 jet trigger efficiencies**

<sub>1104</sub> The L1 jet efficiency is defined as the fraction of leading offline jets which were matched  
<sub>1105</sub> with a L1 tau or central jet above a certain trigger threshold, divided by all the leading  
<sub>1106</sub> offline jets in the event. This quantity is then plotted as a function of the offline jet  $E_T$ ,  
<sub>1107</sub>  $\eta$  and  $\phi$ .

<sub>1108</sub> The efficiency is determined by matching the L1 and reconstructed offline jets spatially  
<sub>1109</sub> in  $\eta - \phi$  space. This is done by calculating the minimum separation in  $\Delta R$  between the  
<sub>1110</sub> highest offline reconstructed jet in  $E_T$  ( $E_T > 10$  GeV,  $|\eta| < 3$ ) and any L1 jet. A jet will  
<sub>1111</sub> be matched if this value is found to be  $< 0.5$ . Should more than one jet satisfy this, the  
<sub>1112</sub> jet closest in  $\Delta R$  is taken as the matched jet. The matching efficiency is close to 100%,  
<sub>1113</sub> above 30(45) GeV for run 2012B(C) data (see Appendix B.1).

<sub>1114</sub> Each efficiency curve is fitted with a function which is the cumulative distribution function  
<sub>1115</sub> of an Exponentially Modified Gaussian (EMG) distribution:

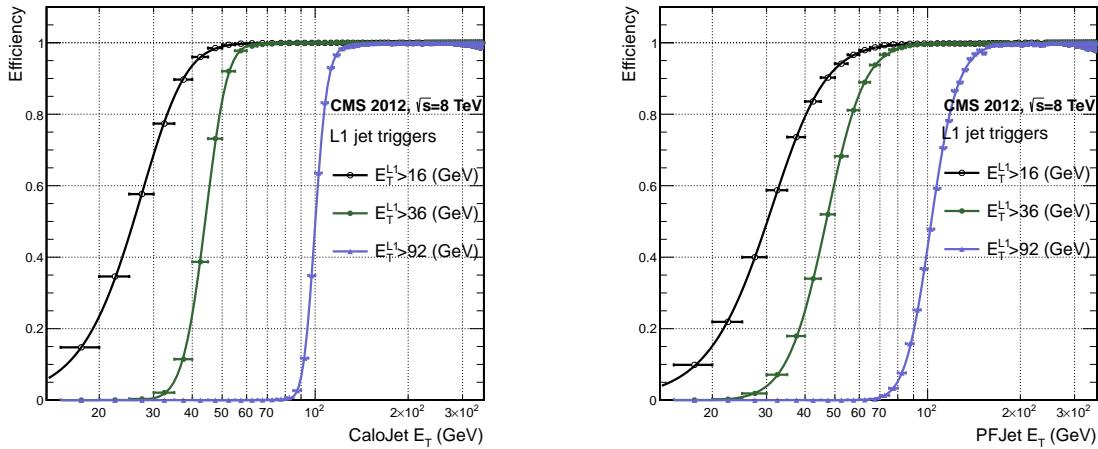
$$\text{f}(x; \mu, \sigma, \lambda) = \frac{\lambda}{2} \cdot e^{\frac{\lambda}{2}(2\mu + \lambda\sigma^2 - 2x)} \cdot \text{erfc}\left(\frac{\mu + \lambda\sigma^2 - x}{\sqrt{2}\sigma}\right) \quad (3.3)$$

where erfc is the complementary error function defined as:

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

<sub>1117</sub> In this functional form, the parameter  $\mu$  determines the point of 50% of the plateau  
<sub>1118</sub> efficiency and the  $\sigma$  gives the resolution. This parametrisation is used to benchmark  
<sub>1119</sub> the efficiency at the plateau, the turn-on points and resolution for each L1 Jet trigger.  
<sub>1120</sub> The choice of function is purely empirical. Previous studies used the error function  
<sub>1121</sub> alone, which described the data well at high threshold values but could not describe the  
<sub>1122</sub> efficiencies well at lower thresholds [68].

<sub>1123</sub> The efficiency turn-on curves for various L1 jet thresholds are evaluated as a function of  
<sub>1124</sub> the offline reconstructed jet  $E_T$  for central jets with  $|\eta| < 3$ . These are measured using  
<sub>1125</sub> single isolated  $\mu$  triggers which have high statistics, and are orthogonal and therefore  
<sub>1126</sub> unbiased to the hadronic triggers under study. The efficiency is calculated with respect to  
<sub>1127</sub> offline Calo and PF Jets in Figure 3.11. Table 3.1 shows the values of these parameters,  
<sub>1128</sub> 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.

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).

The results from the L1 single jet triggers shows good performance for both Calo and PF jets. A better resolution for Calo jets with respect to L1 jets quantities is observed. This effect is due to Calo jet reconstruction using the same detector systems as in L1 jets, whereas with PF jet construction using tracker and muon information, a more smeared resolution when compared to L1 is expected.

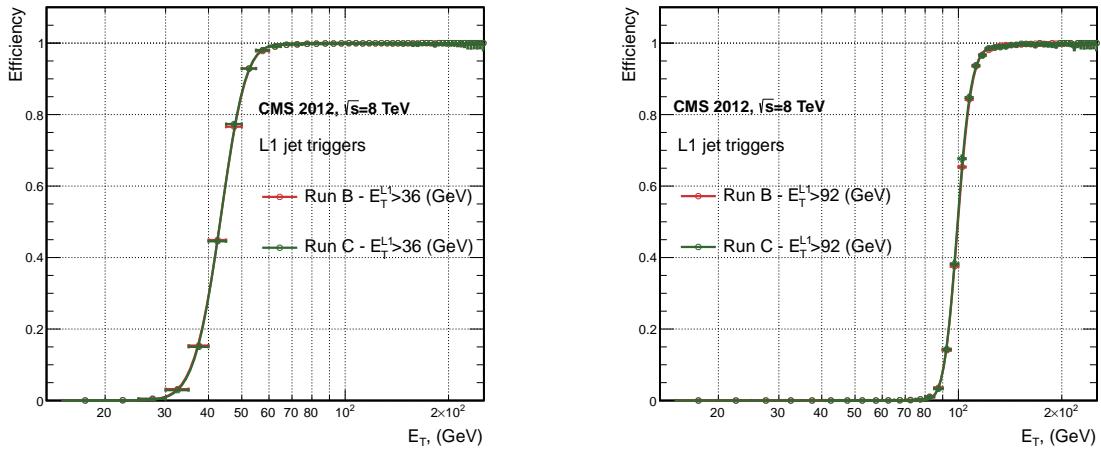
#### 3.4.4. Effects of the L1 jet seed

Between run period B and C of the 2012 data taking period, a jet seed threshold was introduced into the L1 trigger jet algorithm. There was previously no direct requirement made on the energy deposited in the central region. The introduction of a jet seed threshold required that the central region have  $E_T \geq 5\text{GeV}$ , and was introduced to

counteract the effects of high pile up running conditions which create a large number of soft non-collimated jets, that are then added to the jets from the primary interaction or other soft jets from other secondary interactions [69]. This in turn causes a large increase in trigger rate due to the increase in the likelihood that the event causes the L1 trigger to fire. This was implemented to maintain trigger thresholds by cutting the rate of events recorded without significant reduction in the efficiency of physics events of interest.

The effect of the introduction of this jet seed threshold between these two run periods is benchmarked through a comparison of the efficiency of the L1 jet triggers with respect to offline Calo jets shown in Figure 3.12, and the L1  $H_T$  trigger efficiency in Figure 3.14 which is compared to offline  $H_T$  constructed from Calo jets with  $E_T \geq 40\text{GeV}$ .

To negate any effects from different pile-up conditions in the run periods, the efficiencies are measured in events which contain between 15 and 20 primary vertices as defined in 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.

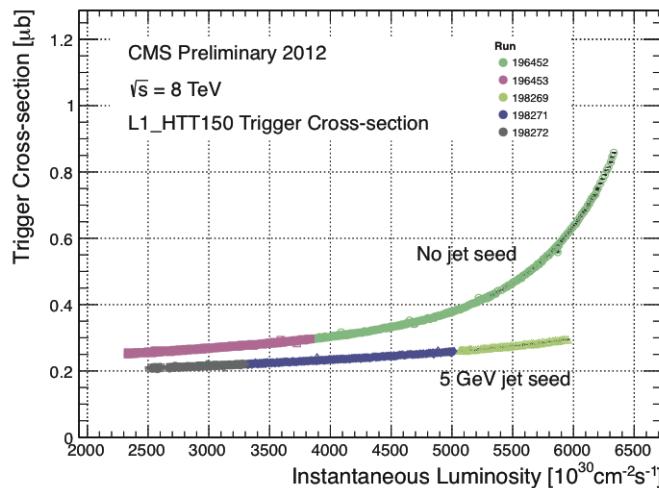
It can be seen that the performance of the  $E_T > 36, 92$  single jet are almost identical, with the jet seed having no measurable effect on these triggers as shown in Table 3.2 .

For the  $H_T$  triggers, a large increase in rate during high pile-up conditions is expected. This is due to the low energy threshold required for a jet to be added to the L1  $H_T$  sum, which is compiled from all uncorrected L1 jets formed in the RCT. The introduction of the jet seed threshold removes the creation of many of these soft low  $E_T$  jets, thus

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).

lowering the  $H_T$  calculation at L1. The effect on the trigger cross section for L1  $H_T$  150 trigger can be seen in Figure 3.13.

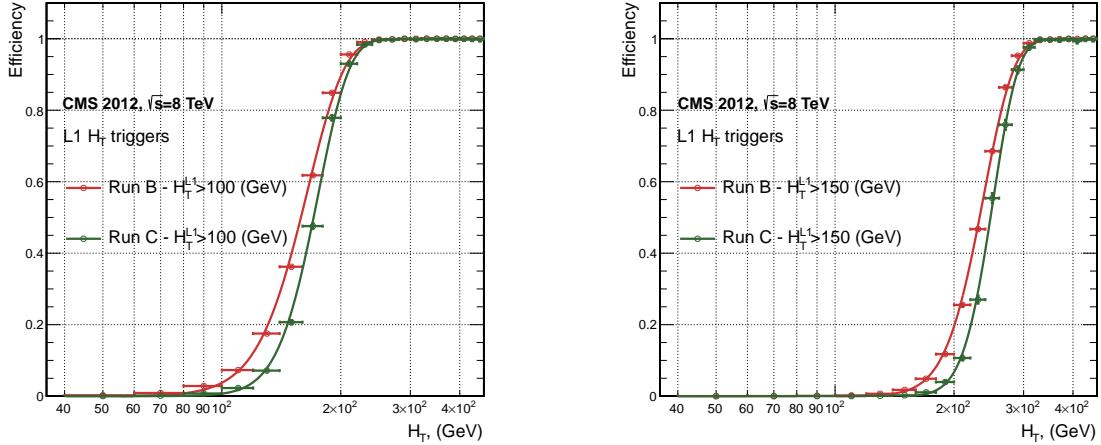


**Figure 3.13.:** Trigger cross section for the L1HTT150 trigger path. Showing that a 5 GeV jet seed threshold dramatically reduces the dependance of cross section on the instantaneous luminosity for L1  $H_T$  triggers [70].

Different behaviours for the trigger turn ons between these run periods are therefore expected. The turn on point is observed to shift to higher  $H_T$  values after the introduction of the jet seed threshold, whilst having a sharper resolution due to less pile-up jets being included in the  $H_T$  sum. This effect is demonstrated within Table 3.3.

### 3.4.5. Robustness of L1 jet performance against pile-up

The performance of the L1 single jet triggers is evaluated in different pile-up conditions to benchmark any dependence on pile-up. Three different pile-up bins of 0-10, 10-20 and  $>20$  vertices are defined, reflecting the low, medium and high pile-up running conditions



**Figure 3.14.:** 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.

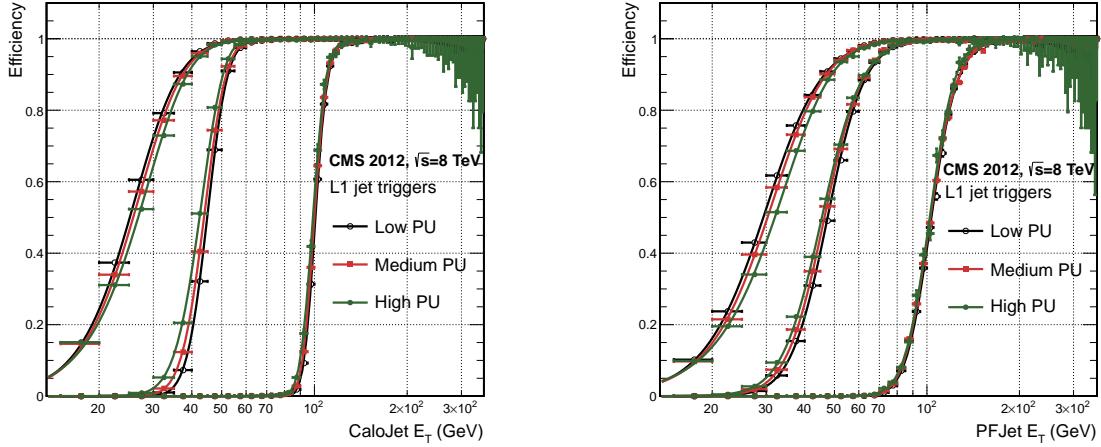
2012B			2012C		
Trigger	$\mu$	$\sigma$	$\mu$	$\sigma$	
L1 HT-100	$157.5 \pm 0.08$	$32.9 \pm 0.08$	$169.8 \pm 0.08$	$28.7 \pm 0.03$	
L1 H1-150	$230.9 \pm 0.02$	$37.3 \pm 0.01$	$246.4 \pm 0.16$	$31.8 \pm 0.05$	

**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).

at CMS in 2012. This is benchmarked relative to Calo and PF jets for the run 2012 C period where the jet seed threshold is applied, with L1 single jet thresholds of 16, 36 and 92 GeV, shown in Figure 3.15. The results of fits to these efficiency turn-on curves are given in Table 3.4 and Table 3.5 for Calo and PF jets respectively.

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.



**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	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	90.4 $\pm$ 0.3	11.2 $\pm$ 0.1	92.0 $\pm$ 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.

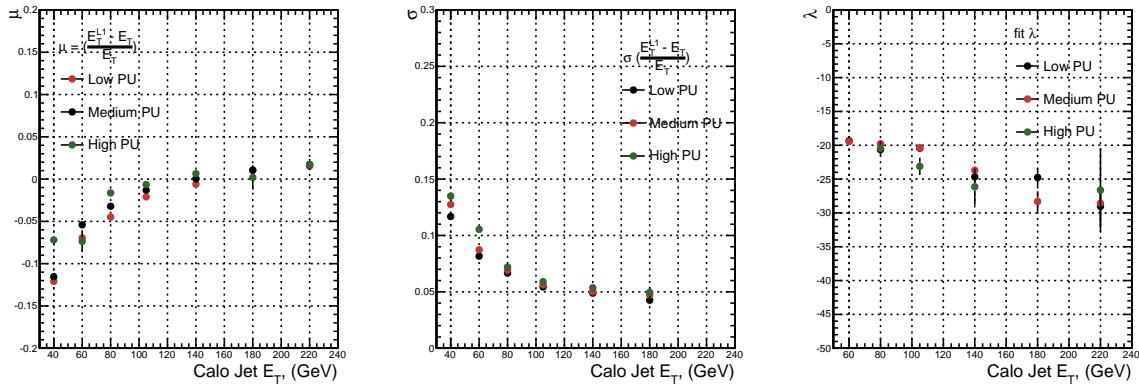
1172 No significant drop in efficiency is observed in the presence of a high number of primary  
 1173 vertices. The increase in hadronic activity in higher pile-up conditions, combined with  
 1174 the absence of pile-up subtraction for L1 jets, results in the expected observation of  
 1175 a decrease in the  $\mu$  value of the efficiency turn-ons as a function of pile-up, while the  
 1176 resolution,  $\sigma$  of the turn-ons are found to gradually worsen as expected with increasing  
 1177 pile-up.

1178 These features are further emphasised when shown as a function of

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

1179 in bins of matched leading offline jet  $E_T$ , of which the individual fits can be found in  
1180 Appendix (B.2). Each of these distributions are fitted with an EMG function as defined  
1181 in Equation (3.3).

1182 The  $\mu$ ,  $\sigma$  and  $\lambda$  values extracted for the low, medium and high pile-up conditions are  
1183 shown for Calo and PF jets in Figure 3.16 and Figure 3.17 respectively. The central value  
1184 of  $\frac{(L1 E_T - \text{Offline } E_T)}{\text{Offline } E_T}$  is observed to increases as a function of jet  $E_T$ , whilst the resolution  
1185 is also observed to improve at higher offline jet  $E_T$ .

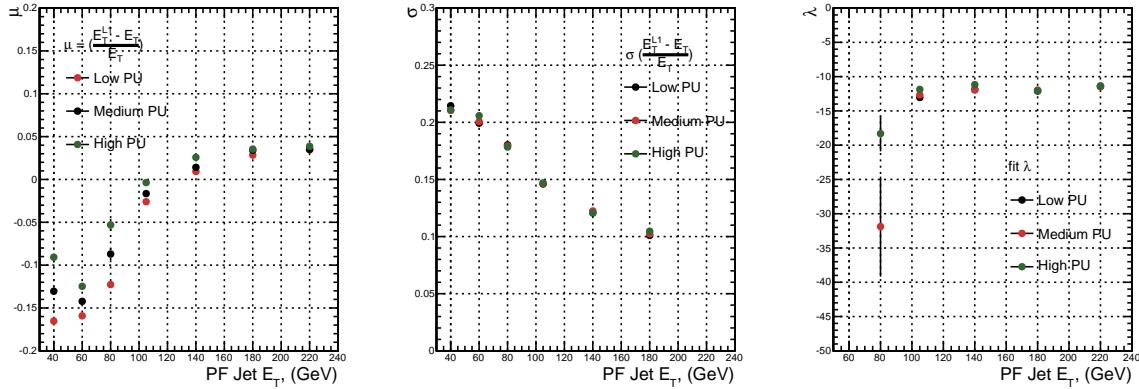


**Figure 3.16.:** Fit values from an EMG function fitted to the resolution plots of leading Calo jet  $E_T$  measured as a function of  $\frac{(L1 E_T - \text{Offline } E_T)}{\text{Offline } E_T}$  for low, medium and high pile-up conditions. The plots show the mean  $\mu$  (left), resolution  $\sigma$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

1186 The resolution of other L1 jet based energy sum quantities,  $H_T$  and  $H_T$  parameterised  
1187 as in Equation (3.4), can be found in Appendix (B.3). The same behaviour observed for  
1188 the single jet triggers is also found for these quantities, where in the presence of higher  
1189 pile-up the  $\mu$  values are shifted to higher values, with a worsening resolution,  $\sigma$  again  
1190 due to the increase in soft pile-up jets and the absence of pile-up subtraction at L1.

### 1191 3.4.6. Summary

1192 The performance of the CMS Level-1 Trigger has been studied and evaluated for jets and  
1193 energy sum quantities using data collected during the 2012 LHC 8 TeV run. These studies  
1194 include the effect of introduction of a 5 GeV jet seed threshold into the jet algorithm  
1195 configuration, the purpose of which is to mitigate the effects of pile-up on the rate of  
1196 L1 triggers whilst not adversely affecting the efficiency of these triggers. No significant



**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.

1197 change in performance is observed with this change and good performance is observed  
 1198 for a range of L1 quantities.

# Chapter 4.

## <sup>1199</sup> SUSY Searches In Hadronic Final <sup>1200</sup> States

<sup>1201</sup> In this chapter a model independent search for **SUSY** in hadronic final states with  
<sup>1202</sup>  $\cancel{E}_T$  using the  $\alpha_T$  variable at different b-quark and jet multiplicities is introduced and  
<sup>1203</sup> described in detail. The results presented are based on a data sample of pp collisions  
<sup>1204</sup> collected in 2012 at  $\sqrt{s} = 8$  TeV, corresponding to an integrated luminosity of  $11.7 \pm 0.5$   
<sup>1205</sup>  $\text{fb}^{-1}$  [5].

<sup>1206</sup> The kinematic variable  $\alpha_T$  is motivated as a variable to provide strong rejection of the  
<sup>1207</sup> overwhelming QCD background, prevalent to jets + $\cancel{E}_T$  final states at the **LHC**. This  
<sup>1208</sup> is achieved whilst maintaining sensitivity to a possible **SUSY** signal and described in  
<sup>1209</sup> Section (4.1). The search and trigger strategy in addition to the event reconstruction  
<sup>1210</sup> and selection are outlined within Sections (4.2 - 4.3).

<sup>1211</sup> The method in which the **SM** background is estimated using an analytical technique to  
<sup>1212</sup> improve statistical precision at higher b-tag multiplicities is detailed within Section (4.5).  
<sup>1213</sup> Included in this section is a discussion on the impact of b-tagging and mis-tagging scale  
<sup>1214</sup> factors between data and simulation on any background predictions. Improved precision  
<sup>1215</sup> in estimating background yields at large number of b-tagged jets, is important in the  
<sup>1216</sup> context of interpreting natural **SUSY** models, first outlined in Section (2.4.1).

<sup>1217</sup> A description of the formulation of appropriate systematic uncertainties applied to the  
<sup>1218</sup> background predictions to account for theoretical uncertainties and limitations in the  
<sup>1219</sup> simulation modelling of event kinematics and instrumental effects is covered in Section  
<sup>1220</sup> (4.6). Similarly the systematic determination for the **SMS** signal samples used to interpret  
<sup>1221</sup> the physics reach of the analysis are examined in Section (4.7).

1222 Finally the statistical likelihood model to interpret the observations in the signal and  
1223 control samples is described in Section (4.8). The experimental reach of the analysis  
1224 discussed within this thesis is interpreted in two classes of **SMS** models, both first  
1225 introduced in Section (2.4.1). The **SMS** models considered in this analysis are summarised  
1226 in Table 4.1. For each model, the **LSP** is assumed to be the lightest neutralino.

1227 Within the table are also defined reference points, parameterised in terms of parent  
1228 gluino/squark and **LSP** sparticle masses,  $m_{\text{parent}}$  and  $m_{\text{LSP}}$ , respectively, which are used  
1229 within the following two chapters to demonstrate potential yields within the signal region  
1230 of the search.

1231 The masses are chosen to reflect parameter space which is within the expected sensitivity  
1232 reach of the search, and also in the case of T1tttt and T1bbbb, reflect examples of  
1233 potential natural **SUSY** topologies.

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)	$pp \rightarrow \tilde{q}\tilde{q}^* \rightarrow q\tilde{\chi}_1^0 q\tilde{\chi}_1^0$	600	250
D2 (T2bb)	$pp \rightarrow \tilde{b}\tilde{b}^* \rightarrow b\tilde{\chi}_1^0 b\tilde{\chi}_1^0$	500	150
D3 (T2tt)	$pp \rightarrow \tilde{t}\tilde{t}^* \rightarrow t\tilde{\chi}_1^0 t\tilde{\chi}_1^0$	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

## 1234 4.1. An Introduction to the $\alpha_T$ Search

1235 A proton-proton collision resulting in the production and decay of supersymmetric  
1236 particles, would manifest as a final state containing energetic jets and  $\cancel{E}_T$  in the hadronic  
1237 channel. The search focuses on topologies where new heavy supersymmetric, R-parity  
1238 conserving particles are pair-produced in pp collisions. These particles decaying to a  
1239 **LSP** escape the detector undetected, leading to significant missing energy and missing  
1240 hadronic transverse energy,

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

1241 defined as the vector sum of the transverse energies of jets selected in an event. Energetic  
1242 jets produced in the decay of these supersymmetric particles also can produce significant  
1243 visible transverse energy,

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

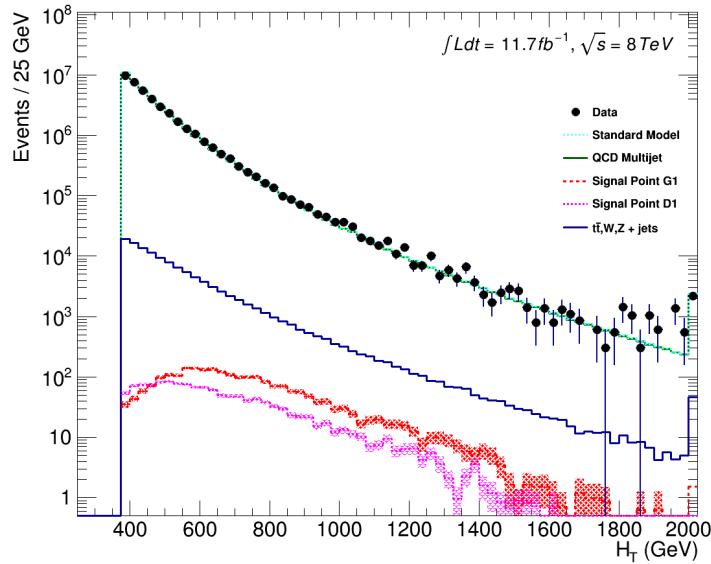
1244 defined as the scalar sum of the transverse energies of jets selected in an event.

1245 A search within this channel is greatly complicated in a hadron collider environment,  
1246 where the overwhelming background comes from inherently balanced multi-jet (“QCD”)  
1247 events which are produced with an extremely large cross section as demonstrated within  
1248 Figure 4.1.  $\cancel{E}_T$  can appear in such events with a substantial mis-measurement or  
1249 stochastic fluctuations of jet energy or missed objects due to detector mis-calibration or  
1250 noise effects.

1251 Additional SM background from EWK processes with genuine  $\cancel{E}_T$  from escaping neutrinos  
1252 comprise the irreducible background within this search and come mainly from:

- 1253 •  $Z \rightarrow \nu\bar{\nu}$  + jets,
- 1254 •  $W \rightarrow l\nu$  + jets in which a lepton falls outside of detector acceptance, is not  
1255 reconstructed, is mis-identified, or the lepton decays hadronically  $\tau \rightarrow$  had ,
- 1256 •  $t\bar{t}$  with at least one leptonically decaying W, which is missed in the detector as  
1257 detailed above,
- 1258 • small background contributions from DY, single top and Diboson (WW,ZZ,WZ)  
1259 processes.

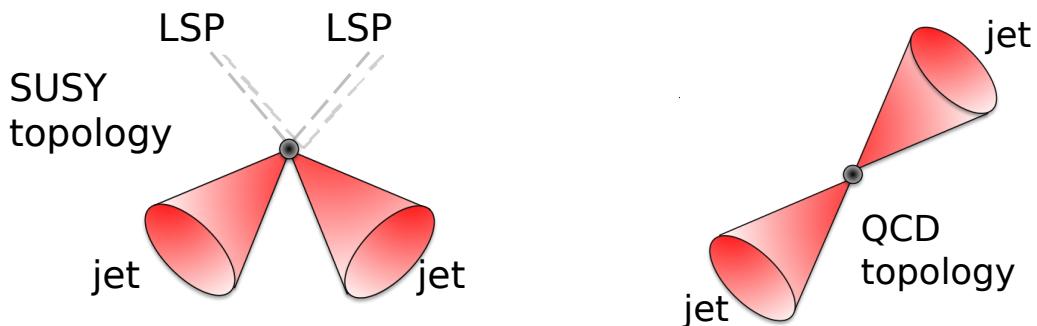
1260 The search is designed to have a strong separation between events with genuine and  
1261 “fake”  $\cancel{E}_T$  which is achieved primarily through the dimensionless kinematic variable,  $\alpha_T$   
1262 [71][72].



**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 MC simulation of EWK processes as well as two reference signal models (labelled G1 and D1 from Table 4.1).

### 1263 4.1.1. The $\alpha_T$ variable

1264 For a perfectly measured di-jet QCD event, conservation laws dictate that both jets must  
 1265 be of equal magnitude and produced in opposite directions. However in the case of di-jet  
 1266 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).

1268 Exploiting this feature leads to the formulation of  $\alpha_T$  (first inspired by [73]) in di-jet  
1269 systems defined as,

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

1270 where  $E_T^{j_2}$  is the transverse energy of the least energetic of the two jets and  $M_T$  defined  
1271 as:

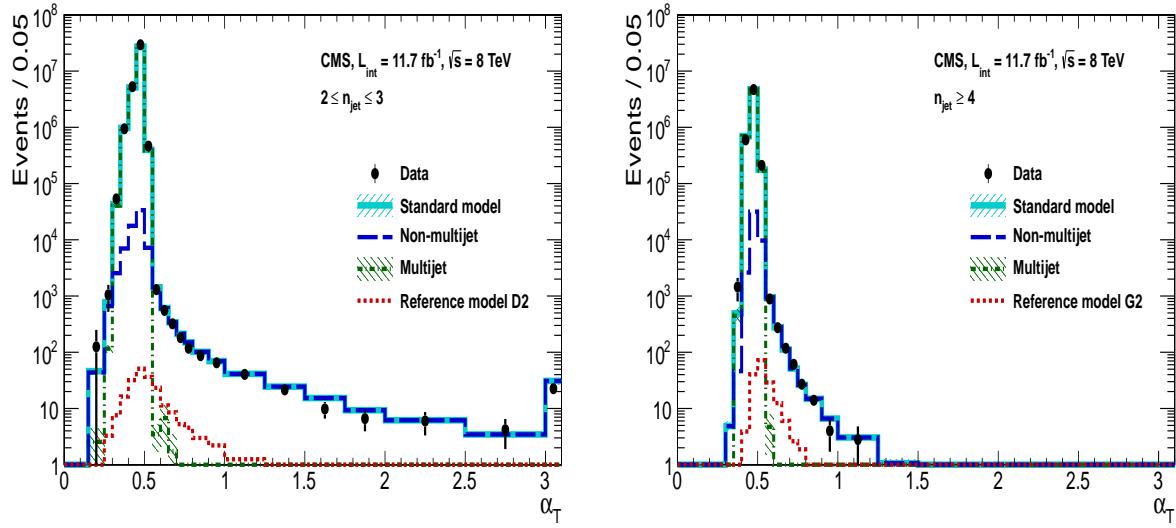
$$M_T = \sqrt{\left(\sum_{i=1}^2 E_T^{j_i}\right)^2 - \left(\sum_{i=1}^2 p_x^{j_i}\right)^2 - \left(\sum_{i=1}^2 p_y^{j_i}\right)^2} \equiv \sqrt{H_T^2 - \cancel{H}_T^2}. \quad (4.4)$$

1272 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, where  
1273 as events with jets which are not back-to-back, for example in events in which a W or  
1274 Z recoils off a system of jets,  $\alpha_T$  can achieve values in excess of 0.5. Most importantly  
1275 balanced QCD events in which jets are mis-measured, will generally result in an  $\alpha_T$  of  
1276 less than 0.5, thus giving the  $\alpha_T$  variable discriminating power between these processes.

1277  $\alpha_T$  can be extended to apply to any arbitrary number of jets, undertaken by modelling  
1278 a system of  $n$  jets as a di-jet system, through the formation of two pseudo-jets [74].  
1279 The two pseudo-jets are built by merging the jets present in the event such that the  
1280 2 pseudo-jets are chosen to be as balanced as possible, i.e the  $\Delta H_T \equiv |E_T^{pj_1} - E_T^{pj_2}|$  is  
1281 minimised between the two pseudo jets. Using Equation (4.4),  $\alpha_T$  can be rewritten as,

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

1282 The distribution of  $\alpha_T$  for the two jet categories used within this analysis,  $2 \leq n_{jet} \leq 3$   
1283 and  $n_{jet} \geq 4$  jets, is shown in the Figure 4.3. It can be seen that the distributions peak  
1284 at an  $\alpha_T$  of 0.5, before falling away sharply and being free of multi-jet background at  
1285 larger  $\alpha_T$  values. These distributions serve to demonstrate the ability of the  $\alpha_T$  variable  
1286 to discriminate between multi-jet events and EWK processes with genuine  $\cancel{E}_T$  in the  
1287 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 shown for  $H_T > 375$ . 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).

1288 The  $\alpha_T$  requirement used within the search is chosen to be  $\alpha_T > 0.55$  to ensure that  
 1289 the QCD multi-jet background is negligible even in the presence of moderate jet mis-  
 1290 measurement. There still remains other effects which can cause multi-jet events to  
 1291 artificially have a large  $\alpha_T$  value, which are discussed in detail in Section (4.2.2).

## 1292 4.2. Search Strategy

1293 The aim of the analysis presented in this thesis is to identify an excess of events in data  
 1294 over the SM background expectation in multi-jet final states and significant  $E_T$ . The  
 1295 essential suppression of the dominant QCD background for such a search is addressed by  
 1296 the  $\alpha_T$  variable described in the previous section. For estimation of the remaining EWK  
 1297 backgrounds, three independent data control samples are used to predict the different  
 1298 processes that compose the background :

- 1299 •  $\mu +$  jets control sample to determine  $W +$  jets,  $t\bar{t}$  and single top backgrounds,  
 1300 •  $\gamma +$  jets control sample to determine the irreducible  $Z \rightarrow \nu\bar{\nu} +$  jets background,

- 1301     •  $\mu\mu + \text{jets}$  control sample to also determine the irreducible  $Z \rightarrow \nu\bar{\nu} + \text{jets}$  background.

1302 These control samples are chosen to both be rich in specific **EWK** processes, be free of  
1303 QCD multi-jet events and to also be kinematically similar to the hadronic signal region  
1304 that they are estimating the backgrounds of, see Section (4.2.3). The redundancy of  
1305 using the  $\gamma + \text{jets}$  and  $\mu\mu + \text{jets}$  sample to predict the same background within the  
1306 signal region, brings an opportunity to reliably cross check and validate the background  
1307 estimation method and is utilised in both the determination of background estimation  
1308 systematics (Section(4.6)), and in the maximum likelihood fit (Section(4.8)).

1309 To remain inclusive to a large range of possible **SUSY** models, the signal region is split  
1310 into the following categories to allow for increased sensitivity in the interpretation of  
1311 results for different **SUSY** topologies:

1312     **Sensitivity to a range of SUSY mass splittings**

1313 The hadronic signal region is defined by  $H_T > 275$ , divided into eight bins in  $H_T$ .

- 1314       – Two bins of width 50 GeV in the range  $275 < H_T < 375$  GeV,  
1315       – five bins of width 100 GeV in the range  $375 < H_T < 875$  GeV,  
1316       – and a final open bin,  $H_T > 875$  GeV.

1317 The choice of the lowest  $H_T$  bin in the analysis is driven primarily by trigger  
1318 constraints. The mass difference between the **LSP** and the particle that it decays  
1319 from is an important factor in the amount of hadronic activity in the event.

1320 A large mass splitting will lead to hard high  $p_T$  jets which contribute to the  $H_T$   
1321 sum. From Figure 4.1 it can be seen that the **SM** background falls sharply at high  
1322  $H_T$  values, therefore binning in  $H_T$  will lead to easier identification of such signals.  
1323 Conversely smaller mass splittings lead to softer jet  $p_T$ 's which will subsequently  
1324 fall into the lower  $H_T$  range.

1325     **Sensitivity to production method of SUSY particles**

1326 The production mechanism of any potential **SUSY** signal can lead to different event  
1327 topologies. One such way to discriminate between gluino ( $g\tilde{g}$  - “high multiplicity”),  
1328 and direct squark ( $q\tilde{q}$  - “low multiplicity”) induced production of **SUSY** particles is  
1329 realised through the number of reconstructed jets in the final state.

1330 The analysis is thus split into two jet categories : 2-3 jets ,  $\geq 4$  jets to give sensitivity  
1331 to both of these mechanisms.

1332 **Sensitivity to “Natural SUSY” via tagging jets from b-quarks**

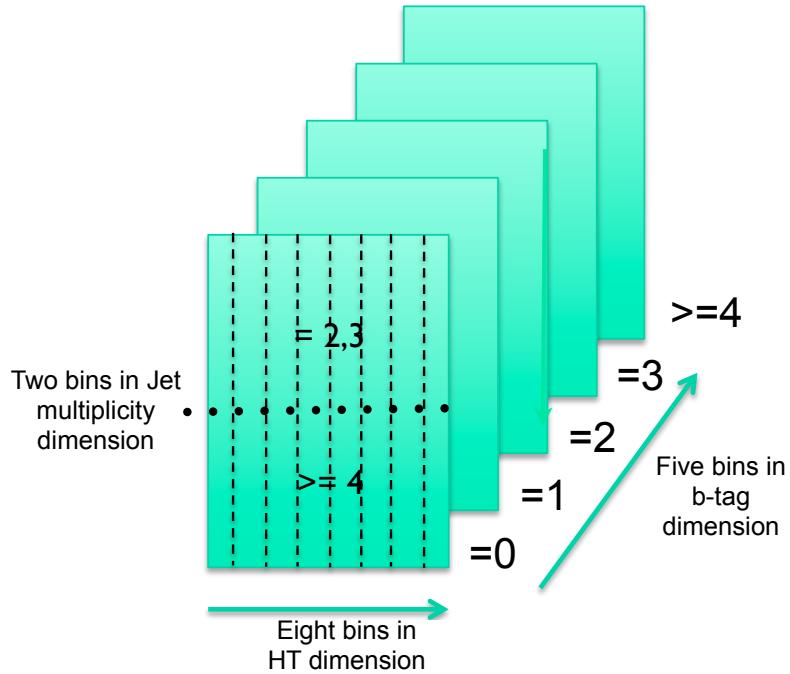
1333 Jets originating from bottom quarks (b-jets) are identified through vertices that  
1334 are displaced with respect to the primary interaction. The algorithm used to tag  
1335 b-jets is the Combined Secondary Vertex Medium Working Point (**CSVM**) tagger,  
1336 described within Section (3.3.2). A cut is placed on the discriminator variable of  
1337  $> 0.679$ , leading to a gluon/light-quark mis-tag rate of approximately 1% and a jet  
1338  $p_T$  dependant b-tagging efficiency of 60-70% [?].

1339 Natural **SUSY** models would be characterised through final-state signatures rich  
1340 in bottom quarks. A search relying on methods to identify jets originating from  
1341 bottom quarks through b-tagging, will significantly improve the sensitivity to this  
1342 class of signature. This gain in sensitivity stems from a vast reduction in the vector  
1343 boson + jet backgrounds (W,Z) at higher b-tag jet multiplicities, which typically  
1344 have no b-flavoured quarks in their decays.

1345 Therefore events are categorised according to the number of b-tagged jets recon-  
1346 structed in each event, in the following: 0,1,2,3, $\geq 4$  b-tag categories . In the highest  
1347  $\geq 4$  b-tag category due to a limited number of expected signal and background, just  
1348 three  $H_T$  bins are employed: 275-325 GeV, 325-375 GeV,  $\geq 375$  GeV.

1349 This characterisation is identically mirrored in all control samples, with the infor-  
1350 mation from all samples and b-tag categories used simultaneously in the likelihood  
1351 model, see Section (4.8).

1352 The combination of the  $H_T$ , jet multiplicity and b-tag categorisation of the signal region as  
1353 described above, resultantly leads to 67 different bins in which the analysis is interpreted  
1354 in, and 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.

### <sup>1355</sup> 4.2.1. Physics objects

<sup>1356</sup> The physics objects used in the analysis defined below, follow the recommendation of  
<sup>1357</sup> the various **CMS** Physics Object Groups (**POGs**).

#### <sup>1358</sup> • Jets

<sup>1359</sup> The jets used in this analysis are CaloJets, reconstructed as described in Section  
<sup>1360</sup> (3.3.1) using the anti- $k_T$  jet clustering algorithm.

<sup>1361</sup> To ensure the jet object falls within the calorimeter systems a pseudo-rapidity  
<sup>1362</sup> requirement of  $|\eta| < 3$  is applied. Each jet must pass a “loose” identification criteria  
<sup>1363</sup> to reject jets resulting from unphysical energy, the criteria of which are detailed in  
<sup>1364</sup> Table A.1 [75].

#### <sup>1365</sup> • Muons

<sup>1366</sup> Muons are selected in the  $\mu + \text{jets}$  and  $\mu\mu + \text{jets}$  control samples, and vetoed in  
<sup>1367</sup> the signal region. The same cut based identification criteria is applied to muons in  
<sup>1368</sup> both search regions and is summarised in Table 4.2 [76].

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 cut 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.

1369      Additionally muons are required to be within the acceptance of the muon tracking  
 1370      systems. For the muon control samples, trigger requirements necessitate a  $|\eta| <$   
 1371      2.1 for the selection of muons. In the signal region where muons are vetoed these  
 1372      conditions are relaxed to  $|\eta| < 2.5$  and a minimum threshold of  $p_T > 10 \text{ GeV}$  is  
 1373      required of muon objects.

1374      **• Photons**

1375      Photons are selected within the  $\gamma + \text{jets}$  control sample and vetoed in all other  
 1376      selections. Photons are identified in both cases according to the cut based criteria  
 1377      listed in Table 4.3 [77].

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.
$\sigma_{i\eta i\eta} < 0.011$	The log energy weighted width ( $\sigma$ ), of the extent of the shower in the $\eta$ dimension.
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 [78], <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.

1378 Photon objects are also required to have a minimum momentum of  $p_T > 25$  GeV.

1379 **• Electrons**

1380 Electron identification is defined for veto purposes. They are selected according to  
1381 the following cut-based criteria listed in Table 4.4, utilising PF-based isolation.

1382 Electrons are required to be identified at  $|\eta| < 2.5$ , with a minimum  $p_T > 10$  GeV  
1383 threshold to ensure that the electrons fall within the tracking system of the detector.

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 ECAL supercluster.
H/E	<0.12	<0.10	The ratio of hadronic energy in the HCAL tower directly behind the ECAL super-cluster and the ECAL 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.

**• Noise and  $E_T$  Filters**

A series of Noise filters are applied to veto events which contain spurious non-physical jets that are not picked up by the jet id, and events which give large unphysical  $E_T$  values. These filters are listed within Table 4.5.

1384

1385

1386

1387

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 $E_T$ 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 $E'_T$ .

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

### 4.2.2. Event selection

- 1389 The selection criteria for events within the analysis are detailed below. A set of common  
 1390 cuts are applied to both signal (maximise acceptance to a range of **SUSY** signatures), and  
 1391 control samples (retain similar jet kinematics for background predictions), with additional  
 1392 selection cuts applied to each control sample to enrich the sample in a particular **EWK**  
 1393 processes, see Section (4.2.3).  
 1394 The jets considered in the analysis are required to have a transverse momentum  $p_T > 50$   
 1395 GeV, with a minimum of two jets required in the event. The highest  $E_T$  jet is required  
 1396 to lie within the central tracker acceptance  $|\eta| < 2.5$ , and the two leading  $p_T$  jets must  
 1397 each have  $p_T > 100$  GeV. Any event which has a jet with  $p_T > 50$  GeV that either fails  
 1398 the “loose” identification criteria described in Section(4.2.1) or has  $|\eta| > 3.0$ , is rejected.  
 1399 Similarly events in which an electron, muon or photon fails object identification but pass  
 1400  $\eta$  and  $p_T$  restrictions, are identified as an “odd” lepton/photon and the event is vetoed.  
 1401 At low  $H_T$ , the jet  $p_T$  threshold requirements required to be considered as part of the  
 1402 analysis and enter the  $H_T$  sum are scaled downwards. These are scaled down in order

1403 to extend phase space at low  $H_T$ , preserving similar jet multiplicities and background  
1404 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.

1405 Within the signal region, to suppress SM processes with genuine  $\cancel{E}_T$  from neutrinos,  
1406 events containing isolated electrons or muons are vetoed. Furthermore to ensure a pure  
1407 multi-jet topology, events are vetoed if an isolated photon is found with  $p_T > 25$  GeV.  
1408 An  $\alpha_T$  requirement of  $> 0.55$  is required to reduce the QCD multi-jet background  
1409 to a negligible amount. Finally additional cleaning cuts are applied to protect against  
1410 pathological deficiencies such as reconstruction failures or severe energy mis-measurements  
1411 due to detector inefficiencies:

- Significant  $\cancel{H}_T$  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,

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

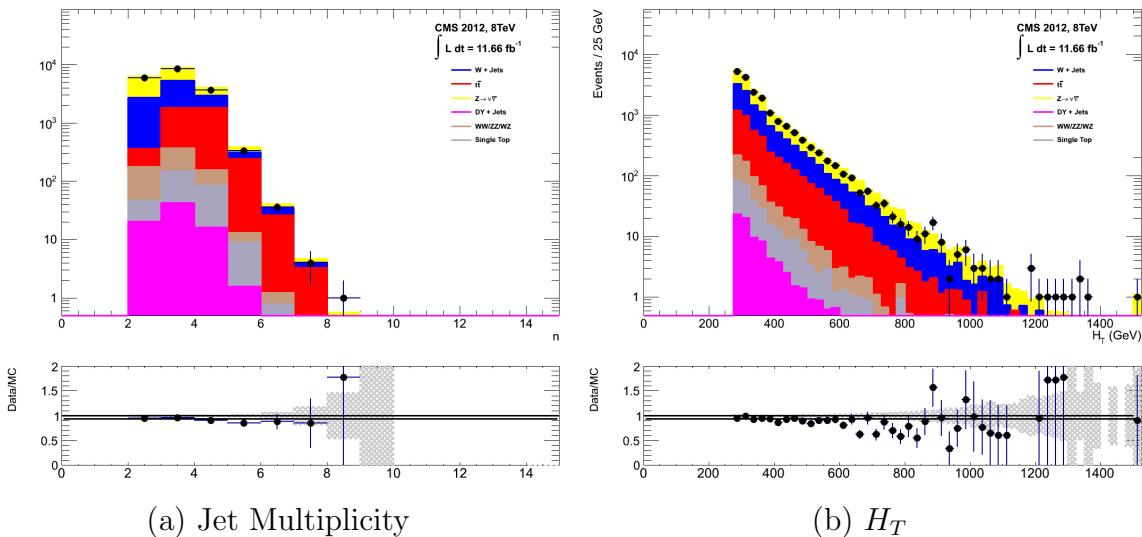
1412 • Fake  $\cancel{E}_T$  and  $\cancel{H}_T$  can arise due to significant jet mis-measurements caused by a small  
1413 number of non-functioning ECAL regions. These regions absorb electromagnetic  
1414 showers which are subsequently not added to the jet energy sum. To circumvent  
1415 this problem the following procedure is employed : For each jet in the event, the  
1416 angular separation

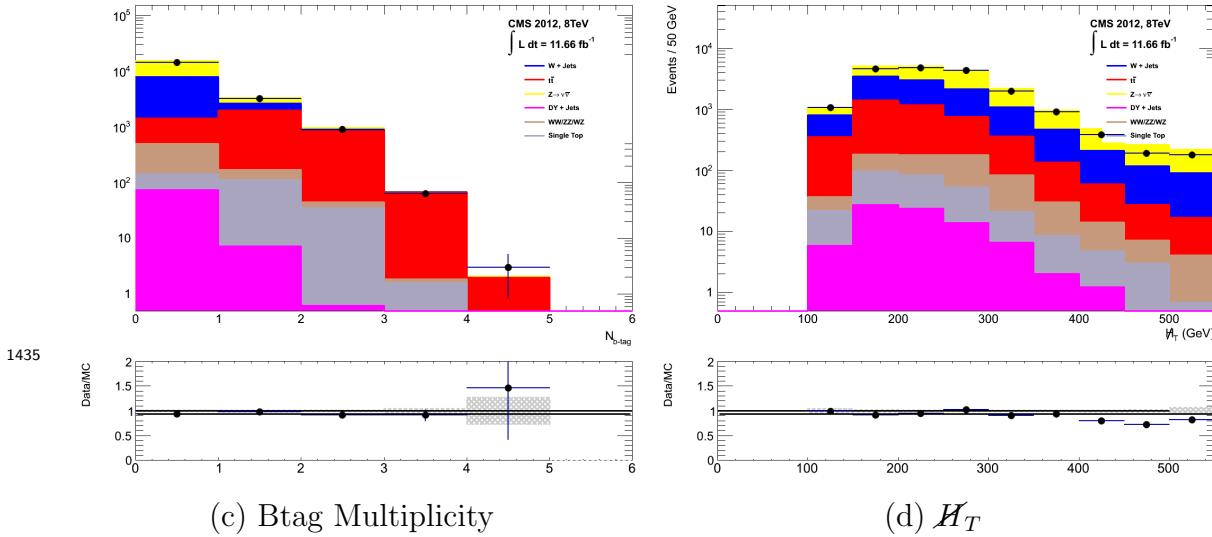
$$\Delta\phi_j^* \equiv \Delta\phi(\vec{p_j} - \sum_{i \neq j} \vec{p_i}), \quad (4.6)$$

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

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

The distributions shown are presented for purely illustrative purposes, with the MC simulation itself not used in absolute terms to estimate the yields from background processes, see Sections (4.2.3, 4.5). However it is nevertheless important to demonstrate that good agreement exists between simulation and observation in 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.

### 1436 4.2.3. Control sample definition and background estimation

1437 The method used to estimate the background contributions in the hadronic signal region  
 1438 relies on the use of a Transfer Factor (**TF**). This is determined from MC simulation  
 1439 in both the control,  $N_{MC}^{\text{control}}$ , and signal,  $N_{MC}^{\text{signal}}$ , region to transform the observed yield  
 1440 measured in data for a control sample,  $N_{\text{obs}}^{\text{control}}$ , into a background prediction,  $N_{\text{pred}}^{\text{signal}}$ , via  
 1441 Equation (4.7),

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

1442 All MC samples are normalised to the luminosity of the data samples,  $11.7 \text{ fb}^{-1}$ . Through  
 1443 this method, “vanilla” predictions for the **SM** background in the signal region can be  
 1444 made by considering separately the sum of the prediction from either the  $\mu + \text{jets}$  and  $\gamma$   
 1445 + jets, or  $\mu + \text{jets}$  and  $\mu\mu + \text{jets}$  samples. However the final background estimation from  
 1446 which results are interpreted, is calculated via a fitting procedure defined formally by  
 1447 the likelihood model described in Section (4.8).

1448 The sum of the expected yields from all MC processes, in each control sample enter the  
1449 denominator,  $N_{MC}^{\text{control}}$ , of the **TF** defined in Eq (4.7). However for the numerator,  $N_{MC}^{\text{signal}}$ ,  
1450 only the relevant processes that are being estimated, enter into the **TF**.

1451 For the  $\mu + \text{jets}$  sample the simulated MC processes which enter the numerator of the  
1452 **TF** are,

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

1453 whilst for both the  $\mu\mu + \text{jets}$  and  $\gamma + \text{jets}$  samples the only MC process used in the  
1454 numerator is,

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

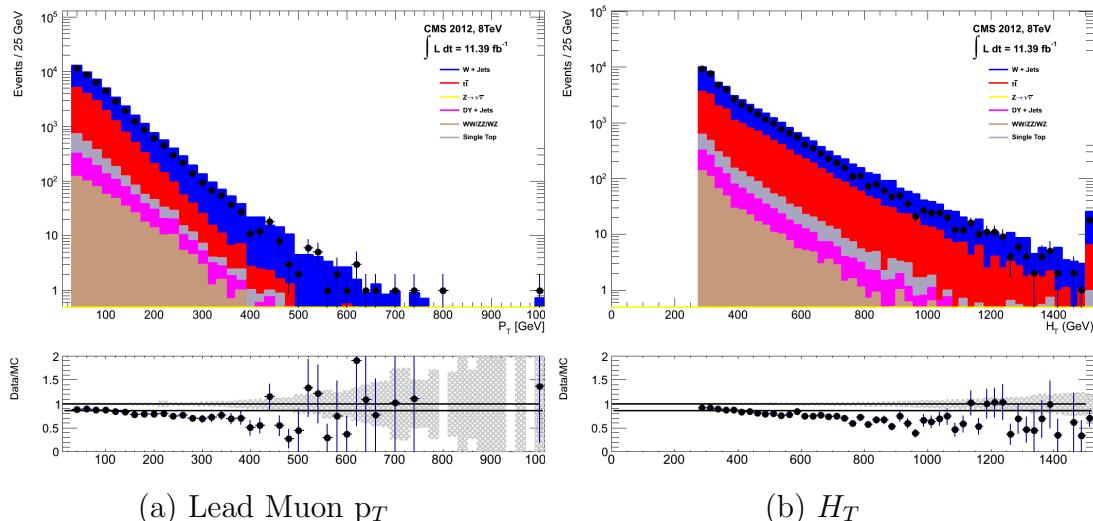
1455 The control samples and the **EWK** processes they are specifically tuned to select are  
1456 defined below, with distributions of key variables for each of the control samples shown  
1457 for illustrative purposes in Figures 4.6, 4.7 and 4.8. No requirement is placed upon  
1458 the number of b-tagged jets or jet multiplicity in the distributions shown. The MC  
1459 distributions highlight the background compositions of each control sample, where in  
1460 general, good agreement is observed between data and simulation, giving confidence  
1461 that the samples are well understood. The contribution from QCD multi-jet events is  
1462 expected to be negligible :

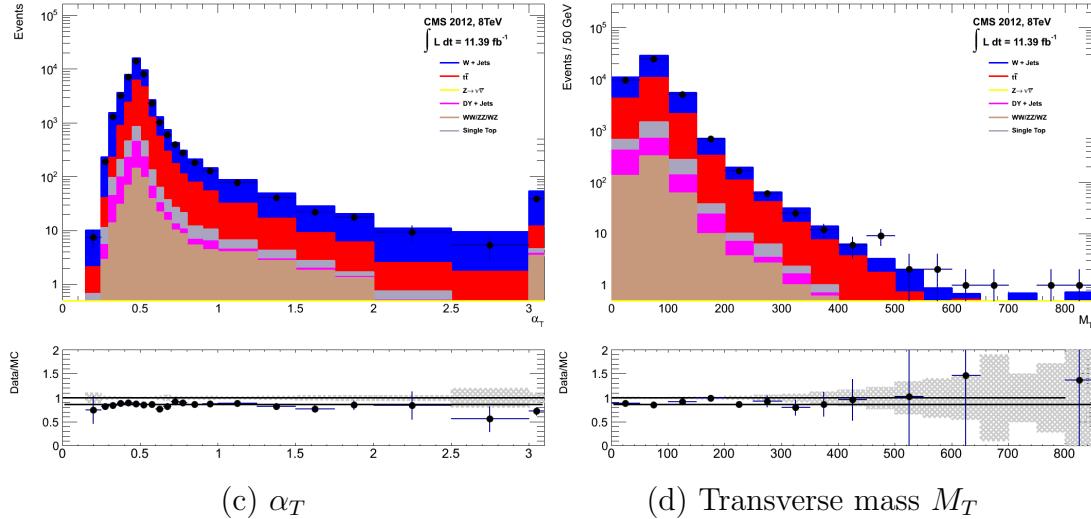
### 1463 The $\mu + \text{jets}$ control sample

1464 Events from  $W + \text{jets}$  and  $t\bar{t}$  processes enter into the hadronic signal sample due  
1465 to unidentified leptons from acceptance effects or reconstruction inefficiencies and  
1466 hadronic tau decays. These leptons originate from the decay of high  $p_T$   $W$  bosons.

1467 The control samples specifically identifies  $W \rightarrow \mu\bar{\nu}$  decays within a similar phase-  
1468 space of the signal region, where the muon is subsequently ignored in the calculation  
1469 of event level variables, i.e.  $H_T$ ,  $\cancel{H}_T$ ,  $\alpha_T$ . All kinematic jet-based cuts are identical  
1470 to those applied in the hadronic search region (with the exception of  $\alpha_T$ , discussed  
1471 below) detailed in Section (4.2.2), with the same  $H_T$ , jet multiplicity and b-jet  
1472 multiplicity binning described above.

- 1473 – Muons originating from W boson decays are selected by requiring one tightly  
 1474 isolated muon defined in Table 4.2, with a  $p_T > 30$  GeV and  $|\eta| < 2.1$ . Both of  
 1475 these threshold arise from trigger restrictions.
- 1476 – The transverse mass of the W candidate must satisfy  $M_T(\mu, \cancel{E}_T) > 30$  GeV ( to suppress QCD multi-jet events).
- 1477 – 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.
- 1478 – Events containing a second muon candidate which has failed id, but passing  
 1479  $p_T$  and  $|\eta|$  requirements, are checked to have an invariant mass that satisfies  
 1480  $|M_{\mu\mu} - m_Z| > 25$ , 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.

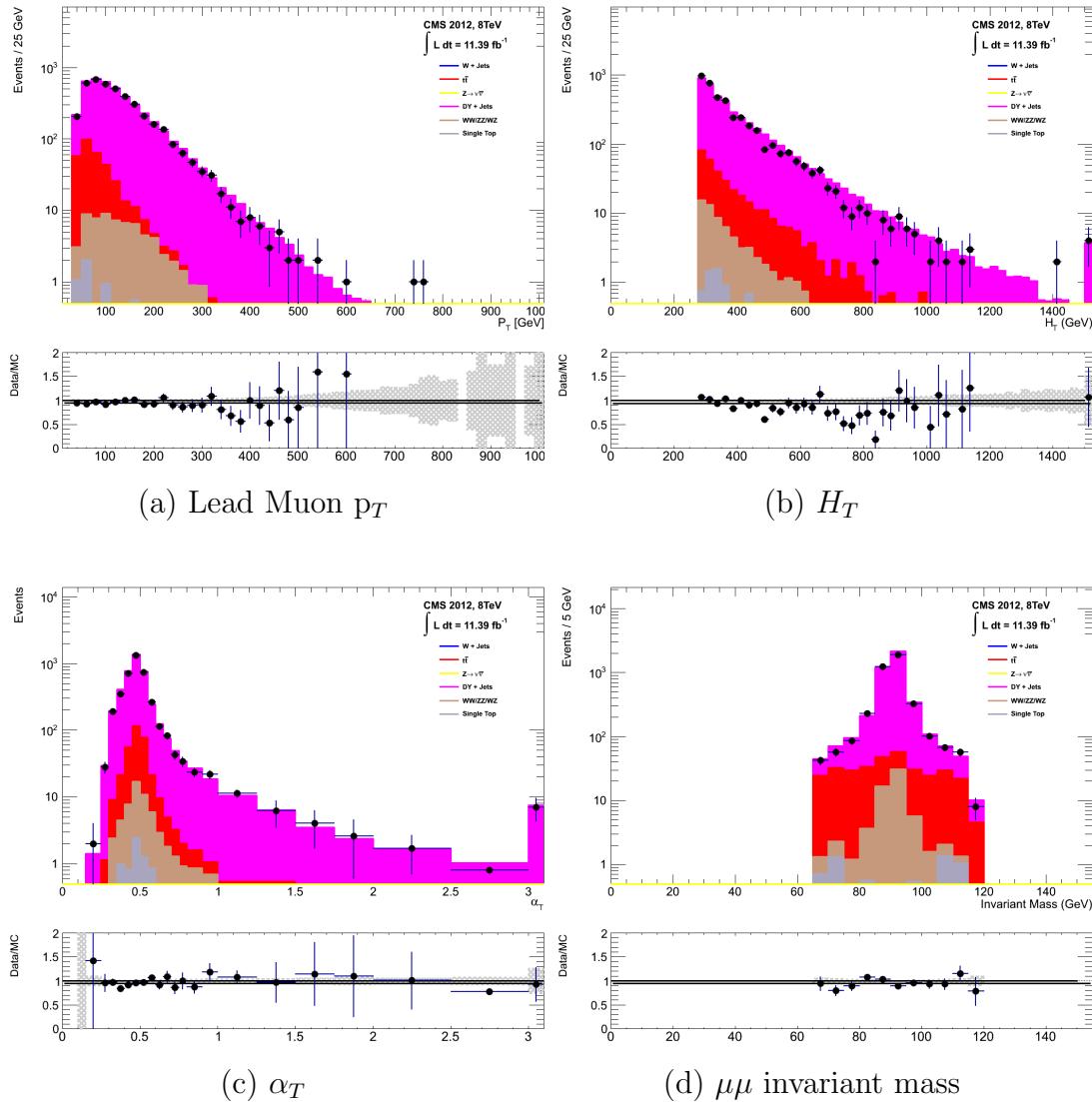
## The $\mu\mu + \text{jets}$ control sample

The  $Z \rightarrow \nu\bar{\nu} + \text{jets}$  background enters into the signal region from genuine  $\cancel{E}_T$  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  $\mu +$  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. Kinematic jet-based cuts and phase space binning identical to the hadronic search region are also applied.

- 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 statistics of the  $\gamma + \text{jets}$  sample is also used in determining the  $Z \rightarrow \nu\bar{\nu}$  estimation.



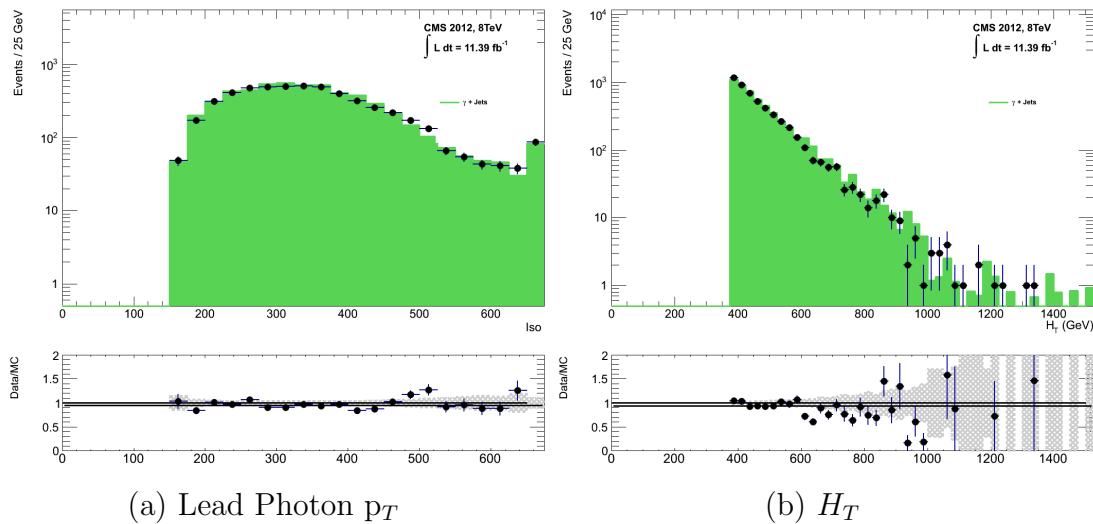
**Figure 4.7.:** Data/MC comparisons of key variables for the  $\mu\mu + \text{jets}$  selection, following the application of selection criteria and the requirements that  $H_T > 275$  GeV. Bands represent the uncertainties due to the statistical size of the MC samples. No requirement is made upon the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

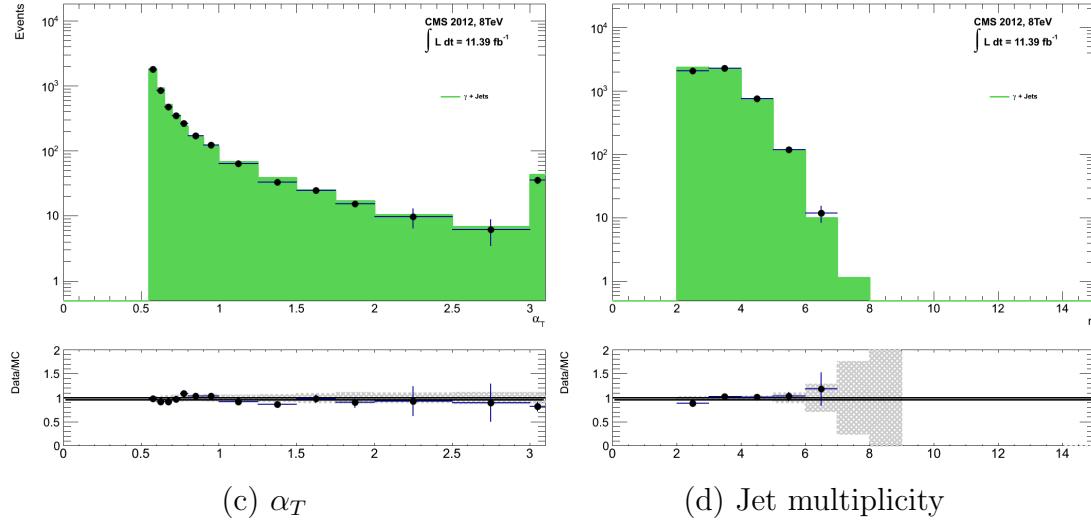
### The $\gamma + \text{jets}$ control sample

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 kinematic properties similar to those of  $Z \rightarrow \mu\bar{\mu}$  events where the photon is ignored [79]. The photon is ignored for the purpose of the calculation of event level variables, and identical selection cuts to the hadronic signal region are applied.

- 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  $\Delta R(\gamma, jet) < 1.0$ , between the photon and all jets is applied to ensure the acceptance of only well isolated  $\gamma + jets$  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), which is maintained in this control sample due to contamination from QCD in the absence of an  $\alpha_T$  cut.





**Figure 4.8.:** Data/MC comparisons of key variables for the  $\gamma +$  jets selection, following the application of selection criteria and the requirements that  $H_T > 375$  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.

1526 The selection criteria of the three control samples are defined to ensure background  
1527 composition and event kinematics mirror closely the signal region. This is done in order  
1528 to minimise the reliance on MC simulation to model correctly the backgrounds and event  
1529 kinematics in the control and signal samples.

1530 However in the case of the  $\mu + \text{jets}$  and  $\mu\mu + \text{jets}$  samples, the  $\alpha_T$  requirement is relaxed  
 1531 in the selection criteria of these samples. This is made possible as contamination from  
 1532 QCD multi-jet events is suppressed to a negligible level by the other kinematic selection  
 1533 criteria within the two control samples, to select pure EWK processes. Thus in this way,  
 1534 the acceptance of the two muon control samples can be significantly increased, which  
 1535 simultaneously improves their predictive power and further reduces the effect of any  
 1536 potential signal contamination.

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

---

#### <sup>1540</sup> 4.2.4. Estimating the QCD multi-jet background

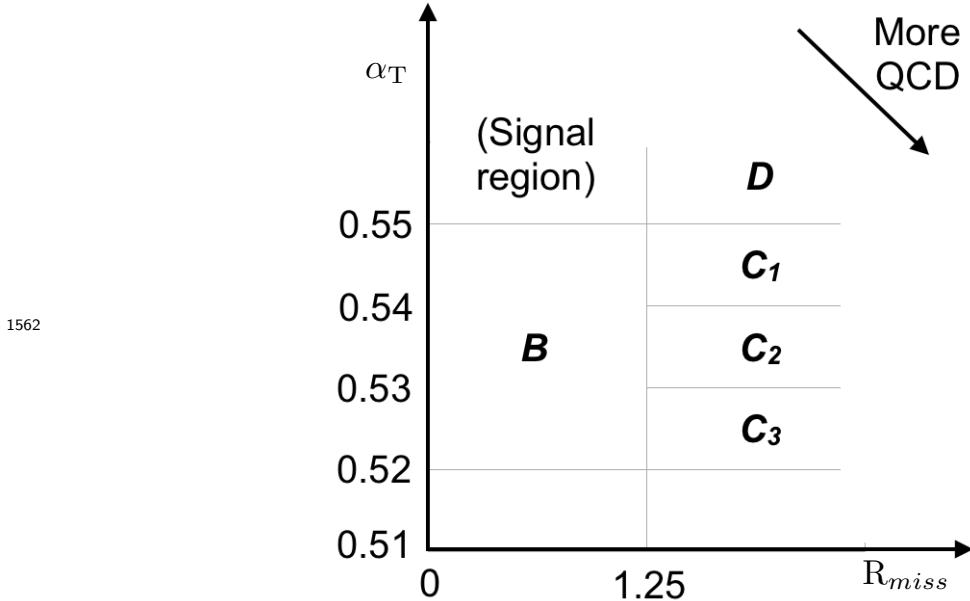
<sup>1541</sup> A negligible background from QCD multi-jet events within the hadronic signal region  
<sup>1542</sup> is expected due to the selection requirement, and additional cleaning filters applied.  
<sup>1543</sup> However a conservative approach is still adopted and the likelihood model, see Section  
<sup>1544</sup> (4.8.2), is given the freedom to accommodate any potential QCD multi-jet contamination.  
<sup>1545</sup> Any potential contamination can be identified through the variable  $R_{\alpha_T}$ , defined as the  
<sup>1546</sup> ratio of events above and below the  $\alpha_T$  threshold value used in the analysis. This is  
<sup>1547</sup> modelled by a  $H_T$  dependant falling exponential function which takes the form,

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

<sup>1548</sup> where the parameters  $A_{\text{QCD}}$  and  $k_{\text{QCD}}$  are the normalisation and exponential decay  
<sup>1549</sup> constants respectively.

<sup>1550</sup> For QCD event topologies, this exponential behaviour is expected as a function of  $H_T$   
<sup>1551</sup> for several reasons. The improvement of jet energy resolution at higher  $H_T$  due to higher  
<sup>1552</sup>  $p_T$  jets leads to a narrower peaked distribution, causing  $R_{\alpha_T}$  to fall. Similarly at higher  
<sup>1553</sup>  $H_T$  values  $> 375$  GeV, the jet multiplicity rises slowly with  $H_T$ . As shown in Figure 4.3,  
<sup>1554</sup> at higher jet multiplicities, the result of the combinatorics used in the determination of  
<sup>1555</sup>  $\alpha_T$ , then leads to a narrower distribution.

<sup>1556</sup> The value of the decay constant  $k_{\text{QCD}}$  is constrained via measurements within data  
<sup>1557</sup> sidebands to the signal region. This is also done to validate the falling exponential  
<sup>1558</sup> assumption for QCD multi-jet topologies. The sidebands are enriched in QCD multi-jet  
<sup>1559</sup> background and defined as regions where either  $\alpha_T$  is relaxed or that the  $R_{\text{miss}}$  cut is  
<sup>1560</sup> inverted. Figure 4.9 depicts the definition of these data sidebands used to constrain the  
<sup>1561</sup> value of  $k_{\text{QCD}}$ .



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

1562 The fits to determine the value of  $k_{\text{QCD}}$  are shown in Appendix (C.1), for which the best  
1563 fit value obtained from sideband region B is determined to be  $k_{\text{QCD}} = 2.96 \pm 0.64 \times 10^{-2}$   
1564 GeV $^{-1}$ .  
1565

1566 The best fit values of the remaining three C sideband regions are used to estimate  
1567 the systematic uncertainty on the central value obtained from sideband region B. The  
1568 variation of these measured values is used to determine the error on the determined  
1569 central value, and is calculated to be  $1.31 \pm 0.26 \times 10^{-2}$  GeV $^{-1}$ . This relative error of  $\sim$   
1570 20% gives an estimate of the systematic uncertainty of the measurement to be applied to  
1571  $k_{\text{QCD}}$ .

1572 Finally the same procedure is performed for sideband region D as an independent cross  
1573 check, to establish that the value of  $k_{\text{QCD}}$  extracted from a lower  $\alpha_T$  slice, can be applied  
1574 to the signal region  $\alpha_T > 0.55$ . The likelihood fit is performed across all  $H_T$  bins within  
1575 the QCD enriched region with no constraint applied to  $k_{\text{QCD}}$ . The resulting best fit  
1576 value for  $k_{\text{QCD}}$  shows good agreement between that and the weighted mean determined  
1577 from the three C sideband regions. This demonstrates that the assumption of using the  
1578 central value determined from sideband region B, to provide an unbiased estimator for  
1579  $k_{\text{QCD}}$  in the signal region ( $\alpha_T > 0.55$ ) is valid.

1580 Table 4.7, summarises the best fit  $k_{\text{QCD}}$  values determined for each of the sideband  
1581 regions to the signal region.

Sideband region	$k_{\text{QCD}} (\times 10^{-2} \text{GeV}^{-1})$	$p$ -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  $k_{\text{QCD}}$  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  $k_{\text{QCD}}$  given by the simultaneous fit using sideband region D. Quotes errors are statistical only.

### 1582 4.3. Trigger Strategy

1583 A cross trigger based on the  $H_T$  and  $\alpha_T$  values of an event, is used with varying thresholds  
 1584 across  $H_T$  bins to record the events used in the hadronic signal region. The  $\alpha_T$  legs of  
 1585 the HT\_alphaT triggers used in the analysis are chosen to suppress QCD multi-jet events  
 1586 and control trigger rate, whilst maintaining signal acceptance. To further maintain an  
 1587 acceptable rate for these analysis specific triggers, only calorimeter information is used  
 1588 in the reconstruction of the  $H_T$  sum, leading to the necessity for Calo jets to be used  
 1589 within the analysis.

1590 A single object prescaled HT trigger is used to collect events for the hadronic control  
 1591 region described above in Section (4.2.4).

1592 The performance of the  $\alpha_T$  and  $H_T$  triggers used to collect data for the signal and  
 1593 hadronic control region is measured with respect to a reference sample collected using the  
 1594 muon system. This allows measurement of both the Level 1 seed and higher level triggers  
 1595 simultaneously, as the reference sample is collected independent of any jet requirements.

1596 The selection for the trigger efficiency measurement is identical to that described in  
 1597 Section (4.2.2), with the requirement of exactly one well identified muon with  $p_T > 30$   
 1598 GeV which is subsequently ignored.

1599 The efficiencies measured for the HT\_alphaT triggers in bins of the individual  $H_T$  and  $\alpha_T$   
 1600 legs, is summarised in Table 4.8.

1601 Data for the control samples of the analysis, detailed in Section (4.2.3), are collected  
 1602 using a single object photon trigger for the  $\gamma + \text{jets}$  sample, and a single object muon  
 1603 trigger for both the  $\mu + \text{jets}$  and  $\mu\mu + \text{jets}$  control samples.

$H_T$ 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.

1604 The photon trigger is measured to be fully efficient for the threshold  $p_T^{\text{photon}} > 150$  GeV,  
 1605 whilst the single muon efficiency satisfying  $p_T^{\mu\text{on}} > 30$  GeV is measured to have an  
 1606 efficiency of  $(88 \pm 2)\%$  that is independent of  $H_T$ . In the case of the  $\mu\mu + \text{jets}$  control  
 1607 sample, the efficiency is measured to be  $(95 \pm 2)\%$  for the lowest  $H_T$  bin, rising (due to  
 1608 the average  $p_T$  of the second muon in the event increasing at larger  $H_T$ ) to  $(98 \pm 2)\%$  for  
 1609 the highest  $H_T$  bin.

## 1610 4.4. Measuring MC Normalisation Factors via $H_T$ 1611 Sidebands

1612 The theoretical cross sections of different SM processes at Next to Next Leading Order  
 1613 (NNLO) and the number of MC simulated events generated for that particular process,  
 1614 is typically used to determine the appropriate normalisation for a MC sample. However  
 1615 within the particular high- $H_T$  and high- $\cancel{E}_T$  corners of kinematic phase space probed  
 1616 within this search, the theoretical cross sections for various processes are far less well  
 1617 understood.

1618 To mitigate the problem of theoretical uncertainties and arbitrary choices of cross sections,  
 1619 the normalisation of the simulation samples used in the analysis are determined through  
 1620 the use of data sidebands. The sidebands are used to calculate sample specific correction  
 1621 factors (k-factors), that are appropriate for the  $H_T$ - $\cancel{E}_T$  phase space covered by this  
 1622 analysis.

1623 They are defined within the  $\mu + \text{jets}$  and  $\mu\mu + \text{jets}$  control sample, by the region  $200 <$   
 1624  $H_T < 275$ , using the same jet  $p_T$  thresholds as the adjacent first analysis bin. Individual  
 1625 EWK processes are isolated within each of these control samples via requirements on  
 1626 jet multiplicity and the requirement on b-tags, summarised in Table 4.9. The purity of

1627 the samples are typically  $> 90\%$  with any residual contamination corrected for. The  
1628 resultant k-factor for each process is determined by then taking ratio of the data yield  
1629 over the MC expectation in the sideband. Subsequently these k-factors are then applied  
1630 to the processes within the phase space of the analysis.

Process	Selection	Observation	MC expectation	k-factor
W + jets	$\mu + \text{jets}, n_b=0, n_{jet} = 2,3$	26950	$29993.2 \pm 650.1$	$0.90 \pm 0.02$
$Z \rightarrow \mu\mu + \text{jets}$	$\mu\mu + \text{jets}, n_b=0, n_{jet} = 2,3$	3141	$3402.0 \pm 43.9$	$0.92 \pm 0.02$
$t\bar{t}$	$\mu + \text{jets}, n_b=2, n_{jet} = \geq 4$	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  $Z \rightarrow \mu\mu + \text{jets}$  processes, are also applied to the  $Z \rightarrow \nu\bar{\nu} + \text{jets}$  and  $\gamma + \text{jets}$  MC samples.

1631 It is worth pointing out that these correction factors have a negligible effect when  
1632 providing a background estimation for the signal region. The **TFs** used in the analysis  
1633 are found to be unaffected by application of these k-factors due to the similarity in the  
1634 background composition of the control and signal regions. However when systematic  
1635 uncertainties are determined in Section (4.6), the closure tests performed are sensitive  
1636 to these corrections when extrapolations between different  $n_b^{\text{reco}}$  and  $n_{jet}$  categories are  
1637 performed.

## 1638 4.5. Determining MC Simulation Yields with 1639 Higher Statistical Precision

1640 Reconstructing events from **EWK** processes with many b-tagged jets ( $\geq 3$ ),  $n_b^{\text{reco}}$ , is largely  
1641 driven by the mis-tagging of light jets within the event. This is clear when considering  
1642 the main **EWK** backgrounds in the analysis, such as  $t\bar{t} + \text{jets}$  events, which typically  
1643 contain two b-flavoured jets from the decay of the top quarks, whilst W + jets and  
1644  $Z \rightarrow \mu\mu + \text{jets}$  events will typically contain no b-flavoured jets.

1645 When the expectation for the number of  $n_b^{\text{reco}}$  is taken directly from simulation, the  
1646 statistical uncertainty at large b-tag multiplicities becomes relatively large. In order to  
1647 reduce this uncertainty one approach is to use the information encoded throughout all  
1648 events in the simulation sample, to measure each of the four ingredients:

- 1649 1. the b-tagging efficiency in the event selection,

1650     2. the charm-tagging efficiency in the event selection  
1651     3. the mis-tagging rate in the event selection,  
1652     4. the underlying flavour distribution of the jets in the events,  
1653    that determine the  $n_b^{\text{reco}}$  distribution of the process being measured. This method allows  
1654    the determination of higher b-tag multiplicities to a higher degree of accuracy reducing  
1655    the statistical uncertainties of the MC which enter into the TF's. For the discussion that  
1656    follows, these predictions are determined on average (i.e not on an event-by-event basis),  
1657    and is known as the formula method.

#### 1658    **4.5.1. The formula method**

1659    The assigning of jet flavours to reconstruction level jets in simulation is achieved via an  
1660    algorithmic method defined as:

- 1661       • Try to find the parton that most likely determines the properties of the jet and  
1662        assign that flavour as true flavour,
- 1663       • “final state” partons (after showering, radiation) are analysed (also within  $\Delta R <$   
1664        0.3 of reconstructed jet cone),
- 1665       • If there is a b/c flavoured parton within the jet cone: label as b/c flavoured jet,
- 1666       • Otherwise: assign flavour of the hardest parton.

1667    This process is employed within each individual MC process and independently for each  
1668     $H_T$ -  $n_{\text{jet}}$  bin in the analysis. The  $n_b^{\text{reco}}$  distribution is then constructed in the following  
1669    way:

1670    Let  $N(n_b^{\text{gen}}, n_c^{\text{gen}}, n_q^{\text{gen}})$  represent the yield in simulation, of events with  $b$  underlying  
1671    b-quarks,  $c$  underlying c-quarks and  $q$  underlying light quarks which are matched to  
1672    reconstructed jets as detailed above. Light quarks defined as those which originate from  
1673    a  $u$ ,  $d$ ,  $s$ ,  $g$  and  $\tau$  jets, which having similar mis-tagging rates are grouped together.  
1674    Similarly  $\epsilon$ ,  $\beta$  and  $m$  represent the measured b-tagging, c-tagging and mis-tagging  
1675    efficiency averaged over all the jets within that particular analysis bin.

1676    Using this information the expected  $n_b^{\text{reco}}$  distribution can be analytically calculated  
1677    using the formula :

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

1678 with  $N(n)$  representing the number of  $n$  b-tagged jets in a particular analysis bin as  
1679 determined by the formula method.

1680 The variables  $n_{b/c/q}^{tag}$  signify the number of times that a particular jet flavour results in a  
1681 b-tagged jet, of which the sum of the three terms must equal the number of  $n$  b-tagged  
1682 jets being estimated. Similarly  $n_{b/c/q}^{gen}$  represent the flavour admixture of the jets, which  
1683 having been identified using the above technique as b, c or light flavoured jets, are  
1684 required by definition that the sum of the three to fall within the  $n_{jet}$  category being  
1685 analysed.

1686 Finally  $P(n_b^{tag}, n_b^{gen}, \epsilon)$ ,  $P(n_c^{tag}, n_c^{gen}, \beta)$  and  $P(n_q^{tag}, n_q^{gen}, m)$  correspond to the binomial  
1687 probabilities for that particular jet flavour and tagging configuration to occur based  
1688 on the measured tagging efficiencies ( $\epsilon$ ,  $\beta$  and  $m$ ) for each jet flavour. This formula  
1689 is enacted over all five of the analysis b-tag categories to build up the resultant  $n_b^{\text{reco}}$   
1690 distribution for each process in turn.

1691 This approach ultimately results in a more precise  $n_b^{\text{reco}}$  distribution prediction, due to  
1692 the utilisation of the entire MC sample in extracting the estimated underlying  $n_b^{\text{reco}}$   
1693 distribution, particularly at higher  $n_b^{\text{reco}}$  multiplicities where a lack of events in simulation  
1694 can lead to relatively large statistical uncertainties.

#### 1695 4.5.2. Establishing proof of principle

1696 In order to validate the procedure, the predictions determined from the formula method  
1697 summarised in Equation (4.11), are compared directly with those obtained directly from  
1698 simulation. Resultantly no simulation to data correction factors are applied when making  
1699 this comparison

1700 This sanity check for the  $\mu + \text{jets}$  control sample is presented in Table 4.10, for all  $n_b^{\text{reco}}$   
1701 and  $H_T$  bins 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.

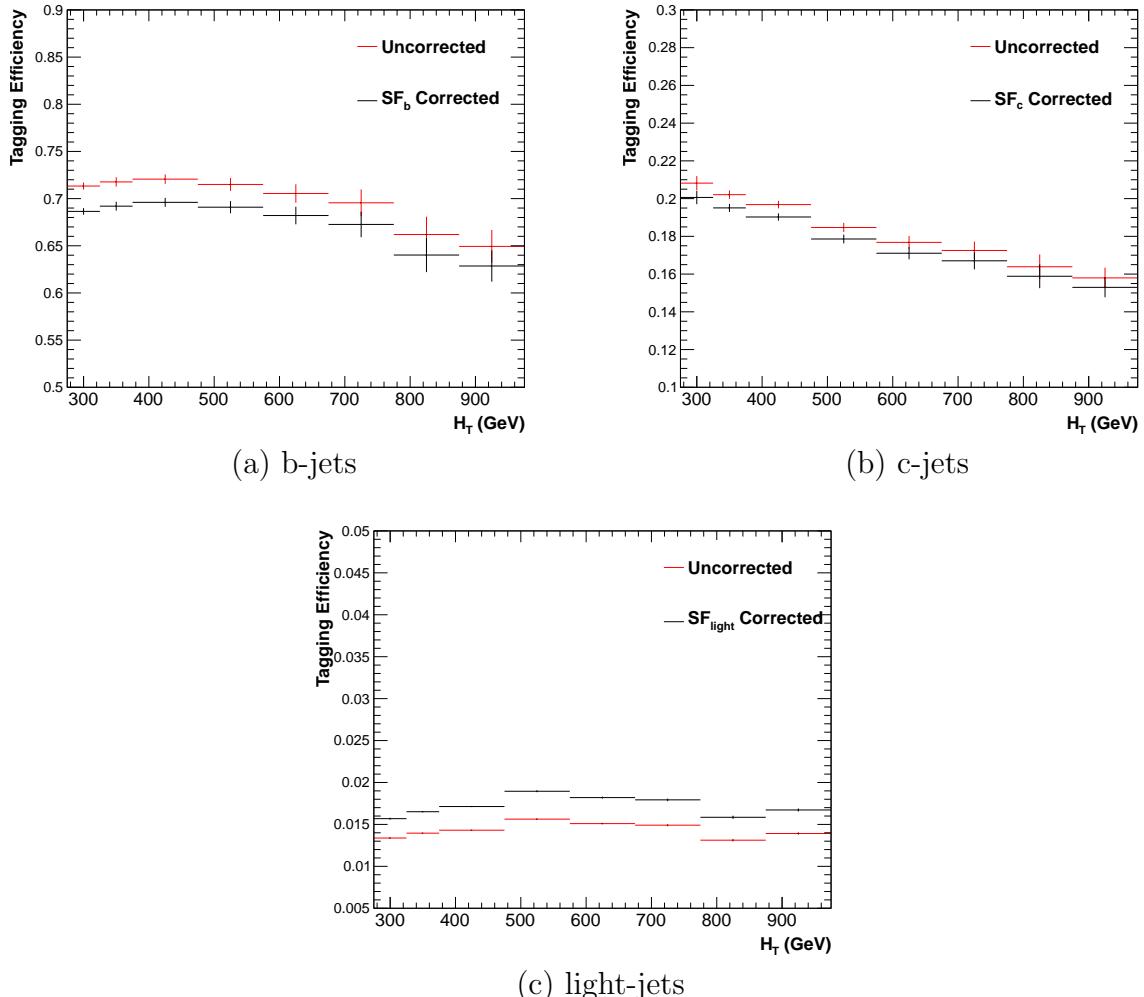
1702 It can be seen as expected, that there is good consistency between the results determined  
 1703 via the formula method and ‘raw’ simulation yields. Similarly the power of this approach  
 1704 can be seen in the reduction of this statistical error in the prediction across all  $H_T$  and  
 1705  $n_b^{reco}$  bins. In particular the statistical uncertainty is reduced by several factors in the  
 1706 highest  $n_b^{reco} \geq 4$  category.

#### 1707 4.5.3. Correcting measured efficiencies in simulation to data

1708 As detailed in Section (3.3.2), it is necessary for certain  $p_T$  and  $\eta$  dependant corrections,  
 1709 to be applied to both the b-tagging efficiency and mis-tagging rates in order to correct the  
 1710 efficiencies from simulation to the distributions seen in data. These corrections factors

1711 are considered when determining the simulation yields for each selection, which are used  
1712 to construct the **TFs** of the analysis.

1713 Each of the corrections factors for the b, c and light flavoured jets come with an  
1714 associated systematic uncertainty. The uncertainties across different jet  $p_T$  and  $\eta$  bins,  
1715 are considered as fully correlated. When computing the magnitude of the effect of this  
1716 systematic uncertainty on the **TFs** of the analysis, the scale factors are therefore scaled  
1717 up/down simultaneously within each  $H_T$  bin of the analysis for all of the  $SF_{b,c,\text{light}}$  scale  
1718 factors. The magnitude of this correction is shown for each  $H_T$  bin within 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 individual analysis  $H_T$  bin. Efficiencies measured directly from simulation (black) and with data to simulation  $SF_{b,c,\text{light}}$  correction factors (red) are applied.

1719 Varying the scale factor corrections by their systematic uncertainty will change the  
1720 absolute yields within each  $n_b^{reco}$  bin of all selections. However, ultimately it is the change  
1721 in the **TFs** which influences the final background prediction from each of the control  
1722 samples. The magnitude of the absolute change in each **TF**, constructed from when the  
1723  $\mu + \text{jets}$  control sample is used to predict the entire hadronic signal region background,  
1724 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  $n_b^{reco}$  bin with no requirement made on the jet multiplicity of the events. (Also quoted are the statistical uncertainties)

1725 It can be seen that the **TFs** are found to be relatively insensitive to the systematic  
1726 uncertainty of the b-tag scale factors (showing typically less than  $\sim 2\%$  change). This  
1727 can be accounted for by the similar composition of the signal and control sample  
1728 backgrounds, such that any change in the underlying  $n_b^{reco}$  distribution will be reflected  
1729 in both signal and control regions and cancel out in the **TF**.

1730 Any overall systematic effect on the overall background prediction of the analysis from  
1731 these b-tag scale factor uncertainties is incorporated within the data driven systematics  
1732 introduced in the following section.

## 1733 4.6. Systematic Uncertainties on Transfer Factors

- 1734 Since the TFs used to establish the background prediction are obtained from simulation,  
1735 an appropriate systematic uncertainty is assigned to each factor to account for theoretical  
1736 uncertainties [80] and limitations in the simulation modelling of event kinematics and  
1737 instrumental effects.
- 1738 The magnitudes of these systematic uncertainties are established through a set of data  
1739 driven method, in which the three independent control samples of the analysis ( $\mu + \text{jets}$ ,  
1740  $\mu\mu + \text{jets}$ ,  $\gamma + \text{jets}$ ) are used to in a series of closure tests. The yields from one of these  
1741 control samples, along with the corresponding TF obtained from simulation, are used to  
1742 predict the yields in another control sample, using the same method of establishing a  
1743 background prediction for the signal region as described in Section (4.2.3).
- 1744 The level of agreement between the predicted and observed yields is expressed as the  
1745 ratio

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

1746 while considering only the statistical uncertainties on  $N_{\text{pred}}$ , the prediction, and  $N_{\text{obs}}$ , the  
1747 observation. No systematic uncertainty is assigned to the prediction, and resultantly the  
1748 level of closure is defined by the statistical significance of a deviation from the ratio from  
1749 zero.

1750 This ratio is measured for each  $H_T$  bin in the analysis, allowing these closure tests to be  
1751 sensitive to both the presence of any significant biases or any possible  $H_T$  dependence on  
1752 the level of closure.

1753 Eight sets of closure tests are defined between the three data control samples, conducted  
1754 independently between the two jet multiplicity ( $2 \leq n_{\text{jet}} \leq 3$ ,  $n_{\text{jet}} \geq 4$ ) bins. Each of  
1755 these tests are specifically chosen to probe each of the different key ingredients of the  
1756 simulation modelling that can affect the background prediction.

1757 Each of the different modelling components and the relevant closure tests are described  
1758 below :

1759  $\alpha_T$  modelling

1760 The modelling of the  $\alpha_T$  distribution in genuine  $E_T$  events is probed with the  $\mu$   
1761 + jets control sample. This test is important to verify the approach of removing  
1762 the  $\alpha_T > 0.55$  requirement from the  $\mu +$  jets and  $\mu\mu +$  jets samples to increase  
1763 the precision of the background prediction. The test uses the  $\mu +$  jets sample  
1764 without an  $\alpha_T$  cut to make a prediction into the  $\mu +$  jets sample defined with the  
1765 requirement  $\alpha_T > 0.55$ .

### 1766 **Background admixture**

1767 The sensitivity of the translation factors to the relative admixture of events from  
1768  $W +$  jets and  $t\bar{t}$  processes is probed by two closure tests. These tests represent  
1769 an extremely conservative approach as the admixture of the background remains  
1770 similar between the  $\mu +$  jets sample and the signal region, contrary to the defined  
1771 closure tests which make predictions between two very different admixtures of  $W +$   
1772 jets and  $t\bar{t}$  events.

1773 Within the  $\mu +$  jets sample, a  $W$  boson enriched sub-sample ( $n_b = 0$ ) is used  
1774 to predict yields in a  $t\bar{t}$  enriched sub-sample ( $n_b = 1$ ). Similarly the  $t\bar{t}$  enriched  
1775 sub-sample ( $n_b = 1$ ) is also used to predict yields for a further enriched  $t\bar{t}$  sub-sample  
1776 ( $n_b = 2$ ), further probing the modelling of the  $n_b^{\text{reco}}$  distribution.

1777 Similarly a further closure test probes the relative contribution of  $Z +$  jets to  $W +$  jets  
1778 and  $t\bar{t}$  events, through the use of the  $\mu +$  jets sample to predict yields for the  $\mu\mu +$   
1779 jets control sample. This closure test, also at some level probes the muon trigger  
1780 and reconstruction efficiencies, given that exactly one or two muons are required by  
1781 the different selections.

### 1782 **Consistency check between $Z \rightarrow \nu\bar{\nu}$ predictions**

1783 An important consistency check between the  $\mu\mu +$  jets and  $\gamma +$  jets, which are both  
1784 used in the prediction of the  $Z \rightarrow \nu\bar{\nu}$  in the signal region. This is conducted by  
1785 using the  $\gamma +$  jets sample to predict yields for the  $\mu\mu +$  jets control sample. Using  
1786  $\gamma +$  jets processes as a method to predict  $Z +$  jet processes is subject to theory  
1787 uncertainties [81], which can be probed by this data driven closure test within a  
1788  $Z \rightarrow \mu\mu$  control sample.

### 1789 **Modelling of jet multiplicity**

1790 The simulation modelling of the jet multiplicity within each control sample is  
1791 important due to the exclusive jet multiplicity binning within the analysis. This is

1792 probed via the use of each of the three control samples to independently predict from  
 1793 the lower jet multiplicity category  $2 \leq n_{\text{jet}} \leq 3$ , to the high jet category  $n_{\text{jet}} \geq 4$ .

1794 For the case of the  $\mu + \text{jets}$  and  $\mu\mu + \text{jets}$  control samples, this test also serves as a  
 1795 further probe of the admixture between  $W + \text{jets}/Z + \text{jets}$  and  $t\bar{t}$ .

1796 To test for the assumption that no  $H_T$  dependencies exist within the background predic-  
 1797 tions of the analysis, the first five closure tests defined above are used, with zeroeth and  
 1798 first order polynomial fits are applied to each test individually. This is summarised in  
 1799 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  
 1800 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

**Table 4.12.:** 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  $2 \geq n_{\text{jet}} \geq 3$  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
$\alpha_T < 0.55 \rightarrow \alpha_T > 0.55 (\mu + \text{jets})$	Circle	$-0.05 \pm 0.03$	0.21	$3.0 \pm 2.9$	0.21
0 b-jets $\rightarrow$ 1 b-jet ( $\mu + \text{jets}$ )	Square	$-0.03 \pm 0.03$	0.55	$-1.0 \pm 1.9$	0.47
1 b-jets $\rightarrow$ 2 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.

1801 Table 4.14 shows the same fits applied to the three closure tests that probe the modelling  
 1802 between the different  $n_{\text{jet}}$  bins. The best fit value and its uncertainty is listed for each  
 1803 set of closure tests in all three tables, along with the p-value of the constant and linear  
 1804 fits applied.

1805 The best fit value for the constant parameter is indicative of the level of closure, averaged  
 1806 across the full range of  $H_T$  bins in the analysis, and the p-value an indicator of any

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.

1807 significant dependence on  $H_T$  within the closure tests. The best fit values of all the tests  
1808 are either statistically compatible with zero bias (i.e, less than  $2\sigma$  from zero) or at the  
1809 level of 10% or less, with the exception of one closure test discussed below.

1810 Within Table 4.14, there exists one test that does not satisfy the above statement, which  
1811 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  
1812 can be largely attributed to an outlier in the  $675 < H_T < 775$  GeV bin, rather than any  
1813 significant trend in  $H_T$ . Removing this single outlier from the constant fit performed,  
1814 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  
1815 modified fit results are included within Table 4.14 .

1816 In addition the best fit values for the slope terms of the linear fits in all three tables are  
1817 of the order  $10^{-4}$ , which corresponds to a percent level change per 100 GeV. However in  
1818 all cases, the best fit values are fully compatible with zero (within  $1\sigma$ ) once again with  
1819 the exception detailed above, indicating that the level of closure is  $H_T$  independent.

#### 1820 4.6.1. Determining systematic uncertainties from closure tests

1821 Once it has been established that no significant bias or trend exists within the closure  
1822 tests, systematic uncertainties are determined. The statistical precision of the closure  
1823 tests is considered a suitable benchmark for determining the systematic uncertainties  
1824 that are assigned to the TFs, which are propagated through to the likelihood fit.

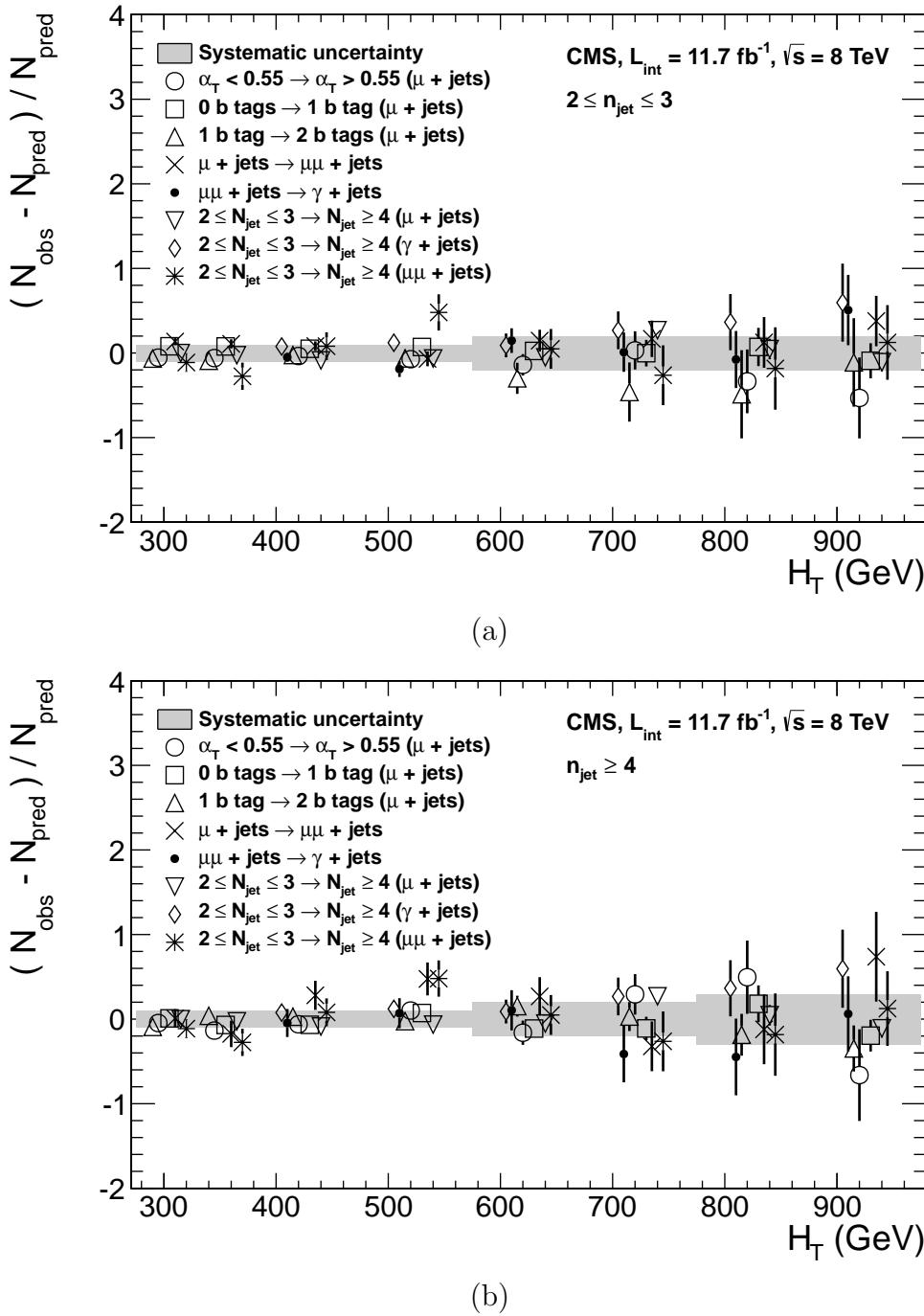
1825 The systematic uncertainty band is split into five separate regions of  $H_T$ . Within each  
1826 region the square root of the sample variance,  $\sigma^2$ , is taken over the eight closure tests to  
1827 determine the systematic uncertainties to be applied within that region.

1828 Using this procedure the systematic uncertainties for each region are calculated and are  
1829 shown in Table 4.15, with the systematic uncertainty to be used in the likelihood model  
1830 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.

1831 Figure 4.11 shows the sets of closure tests overlaid on top of grey bands that represent  
1832 the  $H_T$  dependent systematic uncertainties. These systematic uncertainties are assumed  
1833 to be fully uncorrelated between the different  $n_b$  multiplicity categories and across the  
1834 five  $H_T$  regions. This can be considered a more conservative approach given that some  
1835 correlations between adjacent  $H_T$  bins could be expected due to comparable kinematics.  
1836 These closure tests represent a conservative estimate of the systematic uncertainty in  
1837 making a background prediction for the signal region, which is due to significant differences  
1838 in the background composition and event kinematics between the two sub-samples used  
1839 in the closure tests. This is contrary to the signal region prediction where the two  
1840 sub-samples both have a comparable background admixture and similar kinematics owing  
1841 to the fact that the predictions are always made using the same ( $n_{\text{jet}}$ ,  $n_b^{\text{reco}}$ ,  $H_T$ ) bin.  
1842 This point is emphasised when we examine the sensitivity of the TFs to a change in the  
1843 admixture of  $W + \text{jets}$  and  $t\bar{t}$  with the control and signal samples. This is accomplished  
1844 by varying the cross sections of the  $W + \text{jets}$  and  $t\bar{t}$  by +20% and -20%, respectively.  
1845 Figures C.2 and C.3 within Appendix C, show the effect upon the closure tests for both  
1846 jet multiplicity categories. Given these variations in cross sections, the level of closure is  
1847 found to be significantly worse, with biases as large as  $\sim 30\%$ , most apparent in the  
1848 lowest  $H_T$  bins. However the TFs used to extrapolate from control to signal are seen to  
1849 change only at the percent level by this large change in cross section, shown in Table C.1.  
1850 Given the robust behaviour of the translation factors with respect to large (and opposite)  
1851 variations in the  $W + \text{jets}$  and  $t\bar{t}$  cross sections, one can assume with confidence that  
1852 any bias in the translation factors is adequately (and conservatively) covered by the  
1853 systematic uncertainties used in the analysis.



**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)  $n_{\text{jet}} \geq 4$ .

1854 **4.7. Simplified Models, Efficiencies and Systematic  
1855 Uncertainties**

1856 The results of the analysis are interpreted using various **SMS** signal models, which as  
1857 already introduced in Section (2.4.1) offer a natural starting point for quantifying and  
1858 characterising **SUSY** signals, and a means to identify the boundaries of search sensitivity  
1859 for different mass splittings, kinematic ranges, and final states.

1860 Each model is parameterised in a two dimensional parameter space, ( $m_{\tilde{q}/\tilde{g}}$ ,  $m_{\text{LSP}}$ ), from  
1861 which upper limits on the production cross sections of the various **SMS** models can be  
1862 set.

1863 Each signal sample is generated at Leading Order (**LO**) with Pythia [82], and cross  
1864 sections calculated for Next to Leading Order (**NLO**) and Next to Leading Logarithmic  
1865 Order (**NLL**) [83], with events simulated using the **Fastsim** framework. This framework  
1866 represents a simplified simulation of the **CMS** detector, but allows for faster production  
1867 of various signal topologies with different mass parameters. A series of correction factors  
1868 are applied to account for differences between **Fastsim** [84] and **Fullsim** [85] simulation,  
1869 which can affect the resultant  $n_b^{\text{reco}}$  distribution and which are detailed in Section (4.7.2).

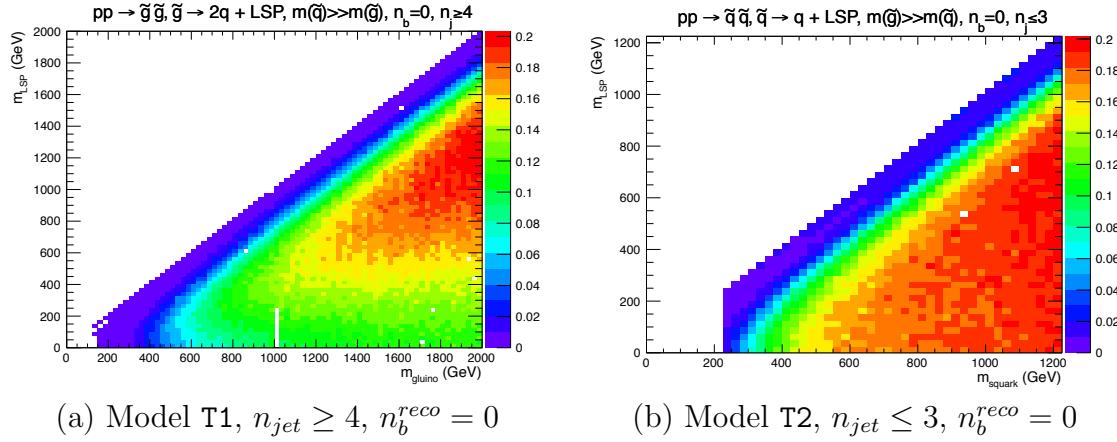
1870 **4.7.1. Signal efficiency**

1871 The analysis selection efficiency,  $\epsilon$ , is measured for each mass point of the interpreted  
1872 model. This serves as a measure of the sensitivity of the signal selection for that particular  
1873 sparticle and **LSP** mass. The signal yield is then given by

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

1874 where  $\sigma$  represents the model's cross section and  $\mathcal{L}$  the luminosity. An upper limit on  $\sigma$   
1875 taken from theory can then allow for the setting of limits in terms of the particle mass.

1876 Figure 4.12 shows the expected signal efficiency of the signal selection for the T1 and  
1877 T2 **SMS** models interpreted in this analysis. The efficiency maps are produced with the  
1878 requirement  $H_T > 275$  GeV (i.e., no binning in  $H_T$ ) and requirements on  $n_{\text{jet}}$  and  $n_b^{\text{reco}}$   
1879 that are appropriate for the model in question.



**Figure 4.12.:** Signal efficiencies for the **SMS** models (a) T1 ( $\tilde{g}\tilde{g}^* \rightarrow q\tilde{q}\chi_1^0 q\bar{q}\chi_1^0$ ) and (b) T2 ( $\tilde{q}\tilde{q}^* \rightarrow q\tilde{\chi}_1^0 \bar{q}\tilde{\chi}_1^0$ ) when requiring  $n_{jet} \geq 4$  and  $\leq 3$  respectively, and  $n_b^{reco} = 0$ .

1880 The same procedure is conducted in the analysis control samples. It is found in the  $\mu$   
 1881 + jets control samples, that the S/B ratios for the expected signal yields in each of the  
 1882 **SMS** models are many time smaller than in the hadronic signal region. The relative  
 1883 contamination for the  $\mu\mu$  + jets sample is smaller still due to the requirement of a second  
 1884 muon. The relative contamination for the  $\gamma$  + jets sample is expected to be zero for the  
 1885 models under consideration. These small, relative levels of contamination are accounted  
 1886 for in the fitting procedure, as described in Section (4.8.4).

### 1887 4.7.2. Applying b-tag scale factor corrections in signal samples

1888 High-statistic **FastSim** signal simulation samples are unavailable for each signal point,  
 1889 which means that a different procedure to the formula method described in Section  
 1890 (4.5) is employed. Furthermore, the use of the **FastSim** framework in the reconstruction  
 1891 introduces an extra set of scale-factor corrections, to be applied simultaneously with  
 1892 those correcting the full-simulation to the data.

1893 For these signal models, an event-by-event re-weighting procedure is applied. This applied  
 1894 weight depends on both the flavour content and the b-tagging status of the reconstruction  
 1895 level jets in the event.

1896 The re-weighting procedure can be described by first considering a single jet within a  
 1897 signal event. The flavour of the jet is determined using the method described in Section  
 1898 (4.5.1).

1899 Maps of tagging efficiencies determined from **FullSim** simulation samples for each of  
1900 the b, c and light jet flavours are produced, binned as a function of jet  $p_T$  and  $\eta$  after  
1901 the application of the hadronic signal selection. The binning of the maps are chosen  
1902 to reflect the set of  $p_T$  and  $\eta$  dependant corrections of simulation to data defined by  
1903 [86]. Taking the flavour,  $p_T$  and  $\eta$  values of each jet in the event, the expected tagging  
1904 efficiency,  $\epsilon_{MC}(p_T, \eta, f)$  is extracted from these maps.

1905 The actual tagging efficiency of the **FastSim** jet,  $\epsilon_{\text{FastSim}}(p_T, \eta, f)$ , differs from that  
1906 measured in **FullSim**,  $\epsilon_{MC}(p_T, \eta, f)$  and is related via an additional correction factor,

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

1907 where  $SF_{\text{Fast} \rightarrow \text{Full}}(p_T, \eta, f)$  represents a set of  $p_T$  and  $\eta$  dependant corrections, that are  
1908 specific for each **SMS** model. These are calculated from the ratio of b-tagging rates  
1909 between a **FullSim**  $t\bar{t}$  sample, and a selection of mass points for that particular **FastSim**  
1910 **SMS** model, again measured individually for b, c and light-flavoured jets.

1911 Similarly the tagging efficiencies measured in data [?],  $\epsilon_{Data}(p_T, \eta, f)$ , are further related  
1912 to  $\epsilon_{\text{FastSim}}(p_T, \eta, f)$  by the equation,

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

1913 For each jet, the weight of the event is re-weighted according to whether the jet fires the  
1914 b-tagger. In the instance that the jet *is* b-tagged, the event weight will be modified by,

$$\text{weight} = SF_{\text{Fast} \rightarrow Data} \times \text{weight}, \quad (4.16)$$

1915 and in the case that the jet does *not* fire the b-tagger,

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

1916 All events that pass the selection criteria are reweighted in this way, and represent the  
1917 yields in each  $n_b^{reco}$  bin corrected from **Fastsim** to data.

### 1918 4.7.3. Experimental uncertainties

1919 The systematic uncertainty on the expected signal acceptance times analysis efficiency is  
1920 determined independently for the each **SMS** model considered. These systematics stem  
1921 from uncertainties on the parton distribution functions, the luminosity measurement, jet  
1922 energy scale, b-tag scale factor measurements and the efficiencies of various cuts used  
1923 in the signal selection, including the  $H_T / E_T$ , dead **ECAL** cleaning filter and lepton /  
1924 photon event vetoes.

1925 Rather than trying to estimate the level of systematic that is applicable point-by-point  
1926 in a model space, general behaviours are considered and instead constant systematics are  
1927 estimated in two regions of the **SMS** models parameter space.

1928 These two regions are defined as, near (small mass splittings) and far (large mass  
1929 splittings) from the mass degenerate diagonal, where the far region is bounded by the  
1930 condition

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

1931 The total systematics in each region are evaluated in the following ways:

1932 **Jet energy scale** : The relative change in the signal efficiency is gauged by varying  
1933 the energy of all jets in an event up or down according to a  $p_T$  and  $\eta$  dependent jet  
1934 energy scale uncertainty. Within the two systematic regions, the resulting systematic  
1935 uncertainties for each **SMS** model are determined by taking the value of the 68<sup>th</sup>  
1936 percentile for the distributions of the relative change in the signal efficiency.

1937 **Luminosity measurement** : The measurement of luminosity taken propagates  
1938 through to an uncertainty on the signal event yield when considering any new  
1939 physics model, which is currently 4.4% [87].

1940 **Parton density function** : Each signal sample is produced using the CTEQ6L1  
1941 parton density function. The effect on the signal acceptance when re-weighting to  
1942 the central value of three different parton distribution functions, CT10, MSTW08  
1943 and NNPDF2.1 are examined [88]. It is found that the change of the signal efficiency  
1944 in different **SMS** models, due to the alternate PDF sets are typically a few percent,  
1945 and approaches 10% at higher squark/gluon and **LSP** masses.

1946  **$H_T/E_T$  cleaning cut** : The ratio of the efficiencies of the cleaning cut are compared  
1947 in simulation and data after application of the  $\mu +$  jets control sample selection.  
1948 No  $\alpha_T$  cut or further event cleaning filters are applied. The ratio of the efficiencies  
1949 observed in data and simulation for a cut value of  $H_T/E_T < 1.25$  and the two jet  
1950 multiplicity bins,  $2 \leq n_{\text{jet}} \leq 3$  and  $n_{\text{jet}} \geq 4$  are  $1.028 \pm 0.007$  and  $1.038 \pm 0.015$   
1951 respectively. These deviations are taken to represent the systematic uncertainty on  
1952 the simulation modelling of this variable.

1953 **Dead ECAL cleaning filter** : The ratio of the efficiencies observed in data and  
1954 simulation for this filter in the two jet multiplicity bins,  $2 \leq n_{\text{jet}} \leq 3$  and  $n_{\text{jet}} \geq 4$ ,  
1955 are  $0.961 \pm 0.008$  and  $0.961 \pm 0.009$ , respectively. These deviations from unity  
1956 are taken to represent the systematic uncertainties in the modelling in simulation of  
1957 this filter.

1958 **Lepton and photon vetoes** : The uncertainty on the efficiency of the lepton and  
1959 photon vetoes is determined by considering truth information. The efficiency of  
1960 the vetoes is measured after applying relevant object filters with identical logic,  
1961 but based on truth instead of reconstructed objects. Where the efficiency is found  
1962 to not be 100%, it is taken to represent the fraction of signal events that are  
1963 incorrectly vetoed. This deviation is taken directly as the systematic uncertainty on  
1964 the efficiency. The systematic uncertainty is only non-zero for models which contain  
1965 third-generation quarks in the final state, where the uncertainties are at the order  
1966 of 1% level.

1967 **B-tag scale factor uncertainties** : The relative change in the signal efficiency is  
1968 observed when relevant flavour,  $p_T$  and  $\eta$  dependant b-tag correction factors, are  
1969 varied up or down by their uncertainty. Within the two systematic regions, the  
1970 resulting systematic uncertainties for each **SMS** model are determined by taking

the value of the 68<sup>th</sup> percentile for the distributions of the relative change in the signal efficiency, over all mass points.

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

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

Model	Luminosity	p.d.f	JES	$\mathcal{H}_T/\cancel{E}_T$	Dead ECAL	Lepton Vetoes	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	$\mathcal{H}_T/\cancel{E}_T$	Dead ECAL	Lepton Vetoes	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.

---

## <sup>1983</sup> 4.8. Statistical Interpretation

<sup>1984</sup> For a given category of events satisfying requirements on both  $n_{jet}$  and  $n_b^{reco}$ , a likelihood  
<sup>1985</sup> model of the observations in multiple data samples is used to gauge agreement between  
<sup>1986</sup> the observed yields in the hadronic signal region, and the predicted yields obtained from  
<sup>1987</sup> the control samples. In addition to checking whether the predictions are compatible  
<sup>1988</sup> with a **SM** only hypothesis, the likelihood model is also used to test for the presence  
<sup>1989</sup> of a variety of signal models. The statistical framework outlined within this section is  
<sup>1990</sup> presented in greater detail within [89].

### <sup>1991</sup> 4.8.1. Hadronic sample

<sup>1992</sup> Let  $N$  be the number of bins on  $H_T$ , with  $n^i$  the number of events observed satisfying  
<sup>1993</sup> all selection requirements in each  $H_T$  bin i. The likelihood of the observations can then  
<sup>1994</sup> be written :

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

<sup>1995</sup> where  $b^i$  represents the expected **SM** background

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

<sup>1996</sup> and  $s^i$  the expected number of signal events from the different **SMS** models interpreted.  
<sup>1997</sup> Pois refers to the Poisson distribution of these values and is defined as :

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

<sub>1998</sub> **4.8.2.  $H_T$  evolution model**

- <sub>1999</sub> The hypothesis, that for a process the  $\alpha_T$  ratio falls exponentially (see Section (4.2.4))  
<sub>2000</sub> in  $H_T$  is defined by Equation (4.10), where  $k_{QCD}$  is constrained by measurements in a  
<sub>2001</sub> signal sideband region.
- <sub>2002</sub> The expected QCD background,  $QCD^i$ , within a bin  $i$  is then modelled as,

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

- <sub>2003</sub> where  $m_i$  represent the number of events observed with  $\alpha_T \leq 0.55$  in each  $H_T$  bin  $i$ , and  
<sub>2004</sub>  $\langle H_T \rangle$  represents the mean  $H_T$  of each bin. Expressed as functions of just the zeroth bin,  
<sub>2005</sub>  $QCD^0$ , and  $k_{QCD}$ , the QCD expectation is given by

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

<sub>2006</sub> **4.8.3. EWK control samples**

- <sub>2007</sub> The **EWK** background estimation within each bin,  $i$ , is broken into two components, the  
<sub>2008</sub> expected yield from  $Z \rightarrow \nu\bar{\nu}$  and  $t\bar{t}$ -W (plus other residual backgrounds) events. This is  
<sub>2009</sub> written as,  $Z_{inv}^i$  and  $t\bar{t}W^i$ , and it follows that

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

- <sub>2010</sub> This can be further expressed as

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

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

2011 where  $f_{Zinv}^i$  represents the expected yield from  $Z \rightarrow \nu\bar{\nu}$  in bin  $i$  divided by the expected  
2012 **EWK** background  $EWK^i$ . This fraction is modelled as a linear component

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

2013 where  $N$  again represents the number of  $H_T$  bins, and  $f_{Zinv}^i$  and  $f_{Zinv}^{N-1}$  are float parameters  
2014 whose final values are limited between zero and one.

2015 Within each  $H_T$  bin there are three background measurements for the different control  
2016 samples,  $n_\gamma^i$ ,  $n_\mu^i$  and  $n_{\mu\mu}^i$ , representing the event yields from the  $\gamma +$  jets,  $\mu +$  jets and  
2017  $\mu\mu +$  jets control samples respectively. Each of these have a corresponding yield in  
2018 simulation,  $MC_\gamma^i$ ,  $MC_\mu^i$  and  $MC_{\mu\mu}^i$ . Within the hadronic signal region there are also  
2019 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  
2020 used to define

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

2021 where  $r_p^i$  represents the inverse of the **TFs** used to extrapolate the yield of each background  
2022 process.

2023 The likelihoods regarding the three measured yields  $n_\gamma^i$ ,  $n_{\mu\mu}^i$ ,  $n_\mu^i$  can then be fully expressed  
2024 as

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

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

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

$$(4.31)$$

2025 which contain an additional term  $s_\mu^i$ , which represents the signal contamination in the  
2026  $\mu +$  jets sample. The parameters  $\rho_{\gamma Z}^j$ ,  $\rho_{\mu\mu}^j$  and  $\rho_\mu^j$  represent “correction factors” that

2027 accommodate the data driven systematic uncertainties derived from the control samples  
2028 in Section (4.12).

2029 Each of these equations are used to estimate the maximum likelihood value for relevant  
2030 background in the signal region given the observations  $n_p^i$  in each of the control samples  
2031 (see Section (4.2.3)).

2032 The measurements in each of the control samples and the hadronic signal region, along  
2033 with the ratios  $r_\gamma^i$ ,  $r_{\mu\mu}^i$ , and  $r_\mu^i$ , are all considered simultaneously through the relationships  
2034 defined by Equations (4.19),(4.24) and (4.25).

2035 In addition to the Poisson product, an additional log-normal term is introduced to  
2036 accommodate the systematic uncertainties given by,

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

2037 where  $\sigma_{\gamma Z}^j$ ,  $\sigma_{\mu\mu Z}^j$  and  $\sigma_{\mu W}^j$  represent the relative systematic uncertainties for the control  
2038 sample constraints and Logn is the log-normal distribution [90],

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

2039 Five parameters per control sample are used to span the eight  $H_T$  bins, with just one  
2040 used for the three  $H_T$  bins in the  $n_b^{reco} \geq 4$  category. These parameters span the same  
2041  $H_T$  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.

2042 Alternatively, in the higher  $n_b^{reco}$  categories ( $n_b^{reco} \geq 2$ ), only the single muon sample  
2043 is used to constrain the total EWK background. Therefore the likelihood function is  
2044 greatly simplified and is represented by

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

2045 where,

$$r'_\mu = \frac{MC_\mu^i}{MC_{tot}^i}. \quad (4.35)$$

2046 **4.8.4. Contributions from signal**

2047 The cross section for each model is represented by  $x$ , while  $l$  represents the total recorded  
2048 luminosity considered by the analysis in the signal region. Let  $\epsilon_{had}^i$  and  $\epsilon_\mu^i$  represent the  
2049 analysis selection efficiency for that particular signal model in  $H_T$  bin  $i$  of the hadronic  
2050 and  $\mu +$  jets control sample respectively. Letting  $\delta$  represent the relative uncertainty on  
2051 the signal yield, assumed to be fully correlated across all bins, and  $\rho_{sig}$  the “correction  
2052 factor” to the signal yield which accommodates this uncertainty.  $f$  represents an unknown  
2053 multiplicative factor on the signal cross section, for which an allowed interval is computed.

2054 The expected signal yield  $s^i$  is thus given by

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

2055 and signal contamination with the  $\mu +$  jets control sample by

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

2056 The systematic uncertainty on the signal is additionally included by the term

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

2057 A discussion of the **SMS** signal models through which the analysis is interpreted can be  
2058 found in the following Chapter.

2059 **4.8.5. Total likelihood**

2060 The total likelihood function for a given signal bin  $k(n_b^{reco}, n_{jet})$  is then given by the  
2061 product of the likelihood functions introduced within the previous sections:

$$L_{\text{Tot}}^k = L_{had}^k \times L_\mu^k \times L_\gamma^k \times L_{\mu\mu}^k \times L_{EWKsyst}^k \times L_{QCD}^k; \quad (0 \leq n_b^{\text{reco}} \leq 1)$$

$$L_{\text{Tot}}^k = L_{had}^k \times L_\mu'^k \times L_{\mu\text{syst}}^k \times L_{QCD}^k \quad (n_b^{\text{reco}} \geq 2). \quad (4.39)$$

2062 In categories containing eight  $H_T$  bins and utilising the three control samples ( $\mu + \text{jets}$ ,  $\mu\mu + \text{jets}$ ,  $\gamma + \text{jets}$ ), there are 25 nuisance parameters, whilst when just one control sample  
2063 is used to estimate the **EWK** background, there are 15 nuisance parameters. Where  
2064 three  $H_T$  bins are used (the highest  $n_b^{reco}$  category), there are 6 nuisance parameters.  
2065 This information is summarised within Table 4.19.

Nuisance parameter	Total
$(EWK^i)_{i:0-7(2)}$	8 (3)
$f_{Zinv}^0$	1*
$f_{Zinv}^7$	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.

2067 When considering **SUSY** signal models within the likelihood, the additional  $L_{\text{sig}}$  term  
2068 is included and therefore when multiple categories are fitted simultaneously the total  
2069 likelihood is then represented by

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

# Chapter 5.

## <sup>2070</sup> Results And Interpretation

<sup>2071</sup> Using the statistical framework outlined in the previous chapter, results are compared to  
<sup>2072</sup> a SM-only hypothesis (Section (5.1)) and interpreted within various SMS models (Section  
<sup>2073</sup> (5.2)).

### <sup>2074</sup> 5.1. Compatibility with the Standard Model <sup>2075</sup> Hypothesis

<sup>2076</sup> The SM background only hypothesis is tested by removing any signal contributions  
<sup>2077</sup> within the signal and control samples, and the likelihood function is maximised over all  
<sup>2078</sup> parameters using Rootfit [91] and MINUIT [92]. The results of the search consist of the  
<sup>2079</sup> observed yields in the hadronic signal sample, and the  $\mu + \text{jets}$ ,  $\mu\mu + \text{jets}$  and  $\gamma + \text{jets}$   
<sup>2080</sup> control samples.

<sup>2081</sup> These observed yields along with the expectations and uncertainties given by the simulta-  
<sup>2082</sup> neous fit for the hadronic signal region are given in Table 5.2. The results obtained from  
<sup>2083</sup> the simultaneous fits, including that of the three control samples, are shown in Figure  
<sup>2084</sup> 5.1-5.8, as summarised in Table 5.1.

<sup>2085</sup> The figures show a comparison between the observed yields and the SM expectations  
<sup>2086</sup> across all  $H_T$  bins, and in all  $n_{jet}$  and  $n_b^{reco}$  multiplicity categories. In all categories the  
<sup>2087</sup> samples are well described by the SM only hypothesis. In particular no significant excess  
<sup>2088</sup> is observed above SM expectation within the hadronic signal region.

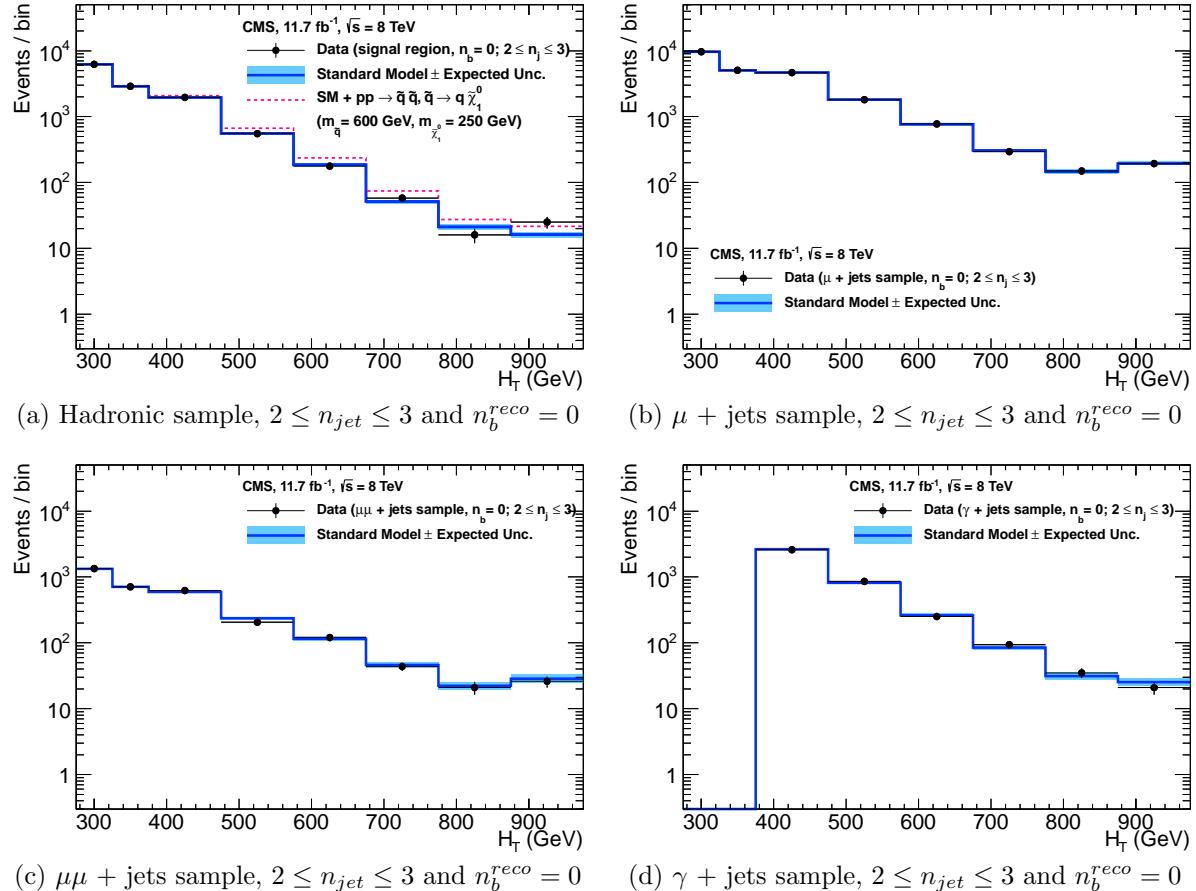
$n_{jet}$	$n_b^{reco}$	Control samples fitted	Figure
2-3	0	$\mu + \text{jets}, \mu\mu + \text{jets}, \gamma + \text{jets}$	5.1
2-3	1	$\mu + \text{jets}, \mu\mu + \text{jets}, \gamma + \text{jets}$	5.2
2-3	1	$\mu + \text{jets}$	5.3
$\geq 4$	0	$\mu + \text{jets}, \mu\mu + \text{jets}, \gamma + \text{jets}$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

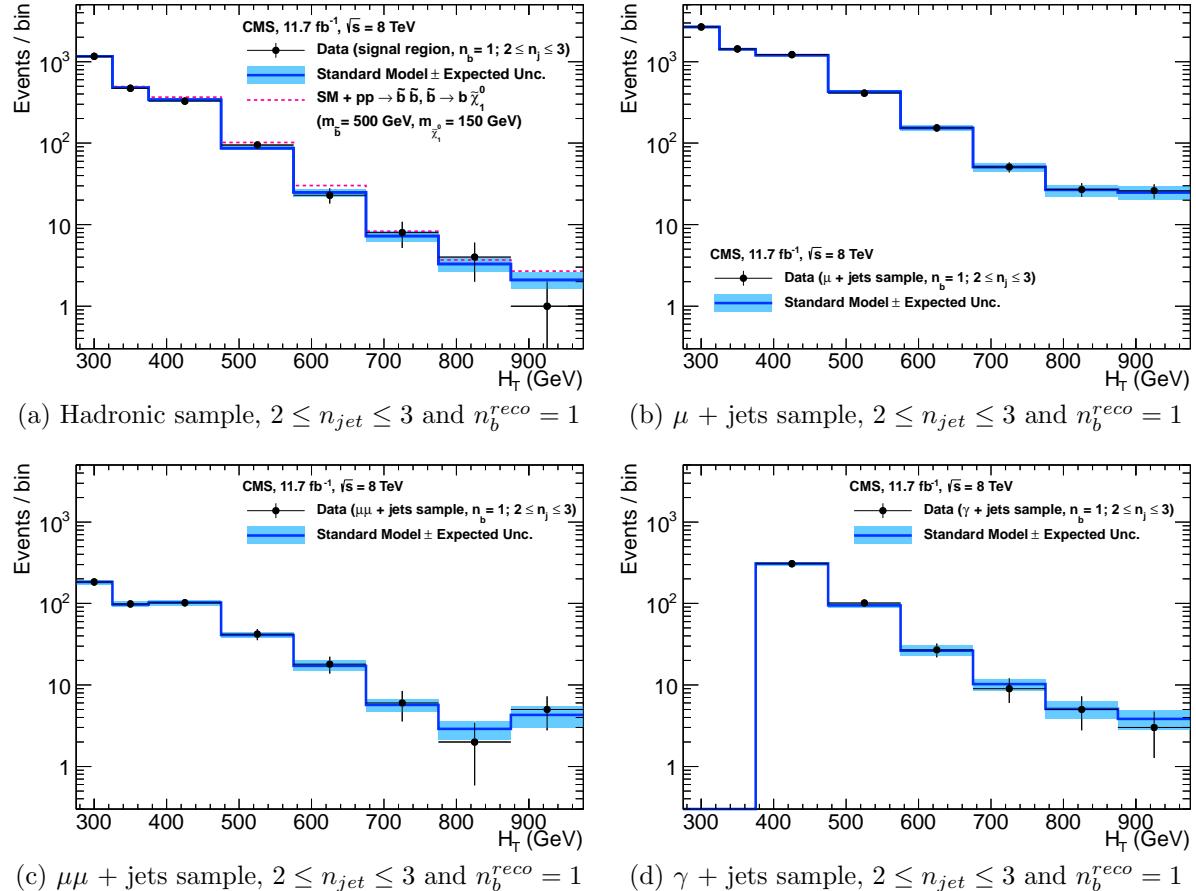
Cat	$n_b^{reco}$	$n_{jet}$	$H_T$ bin (GeV)							
			275-325	325-375	375-475	474-575	575-675	675-775	775-875	875- $\infty$
SM Data	0	$\leq 3$	$6235^{+100}_{-67}$ 6232	$2900^{+60}_{-54}$ 2904	$1955^{+34}_{-39}$ 1965	$558^{+14}_{-15}$ 552	$186^{+11}_{-10}$ 177	$51.3^{+3.4}_{-3.8}$ 58	$21.2^{+2.3}_{-2.2}$ 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.

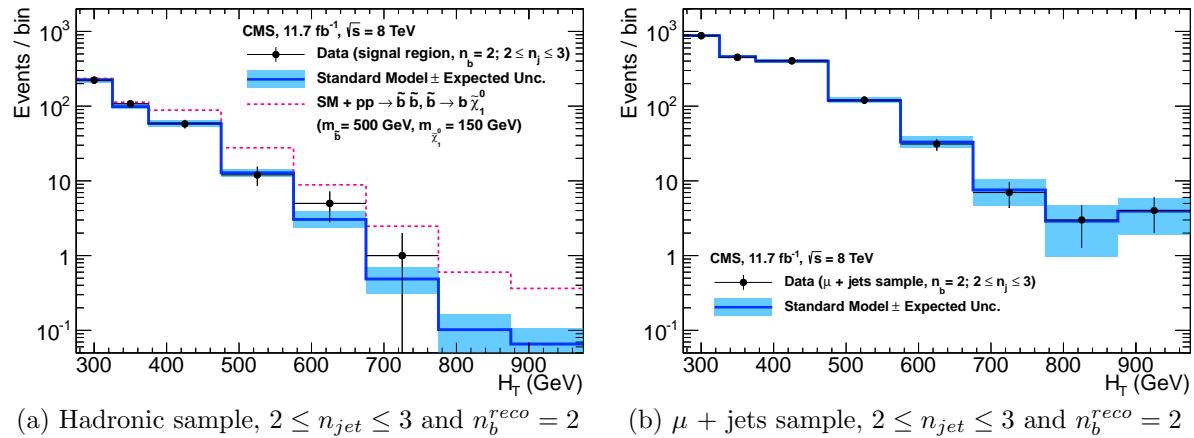
Given the lack of an excess in data hinting at a possible supersymmetric signature within the data, interpretations are made on the production masses and cross section of a range of **SUSY** decay topologies within the following section.



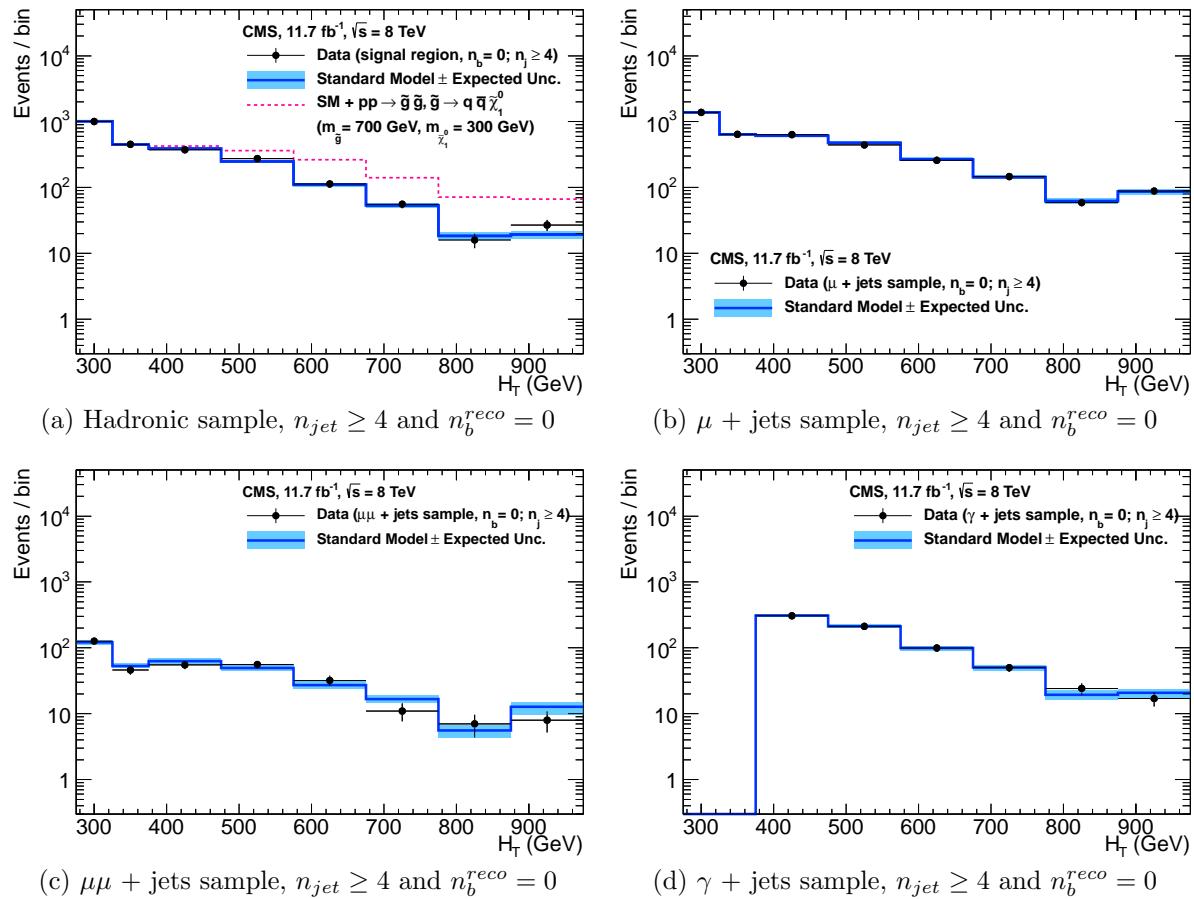
**Figure 5.1.:** Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of  $H_T$  for the (a) hadronic, (b)  $\mu +$  jets, (c)  $\mu\mu +$  jets and (d)  $\gamma +$  jets samples when requiring  $n_b^{reco} = 0$  and  $n_{jet} \leq 3$ . The observed event yields 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.



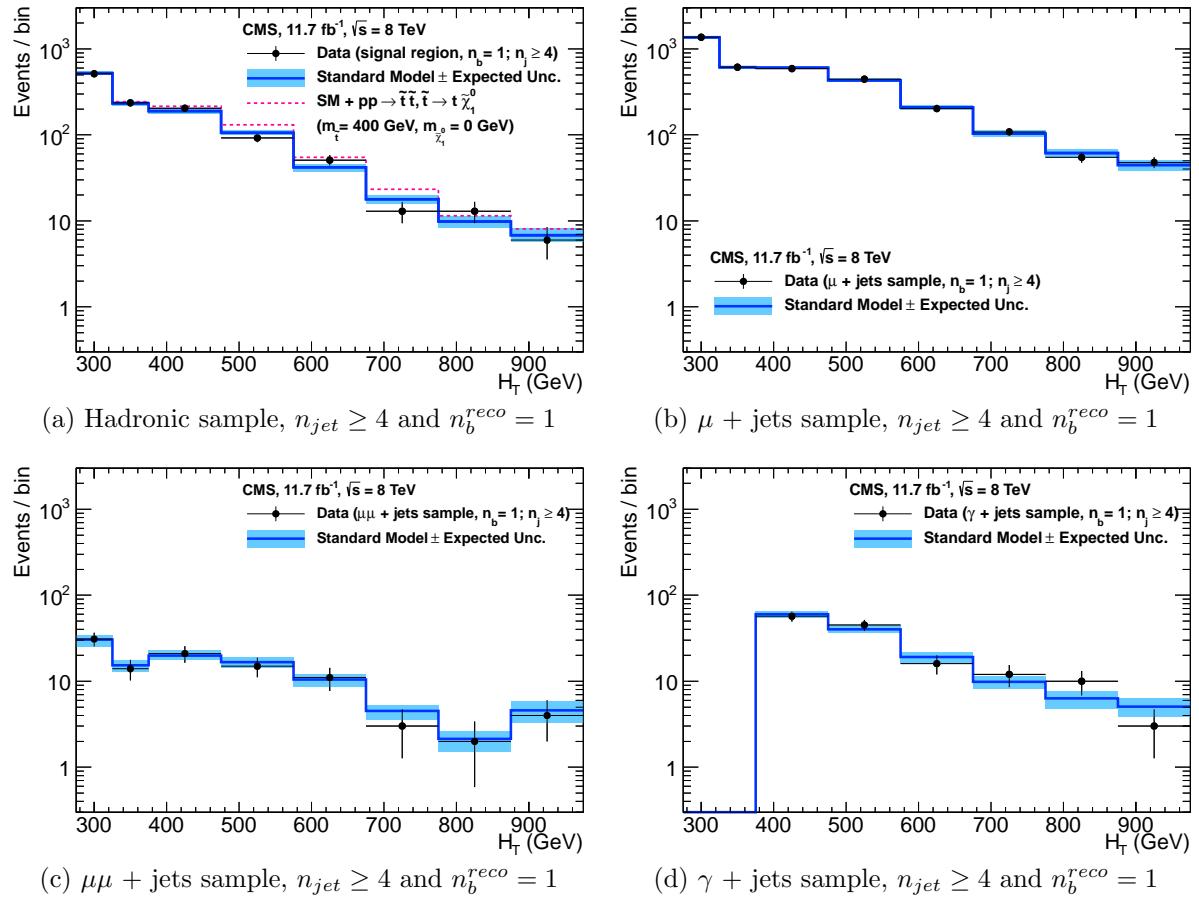
**Figure 5.2.:** Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of  $H_T$  for the (a) hadronic, (b)  $\mu +$  jets, (c)  $\mu\mu +$  jets and (d)  $\gamma +$  jets samples when requiring  $n_b^{reco} = 1$  and  $n_{jet} \leq 3$ . The observed event yields 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.



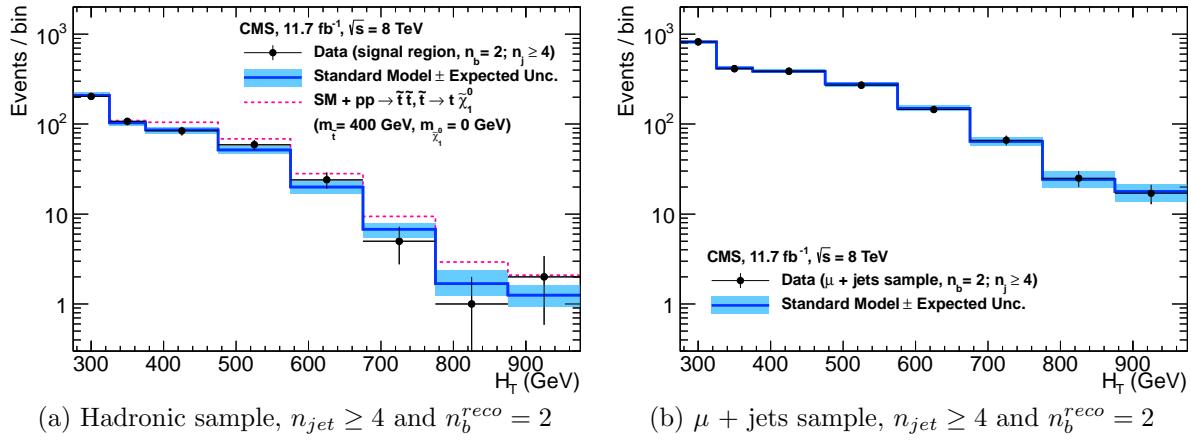
**Figure 5.3.:** Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of  $H_T$  for the (a) hadronic, (b)  $\mu +$  jets, (c)  $\mu\mu +$  jets and (d)  $\gamma +$  jets samples when requiring  $n_b^{reco} = 2$  and  $n_{jet} \leq 3$ . The observed event yields 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.



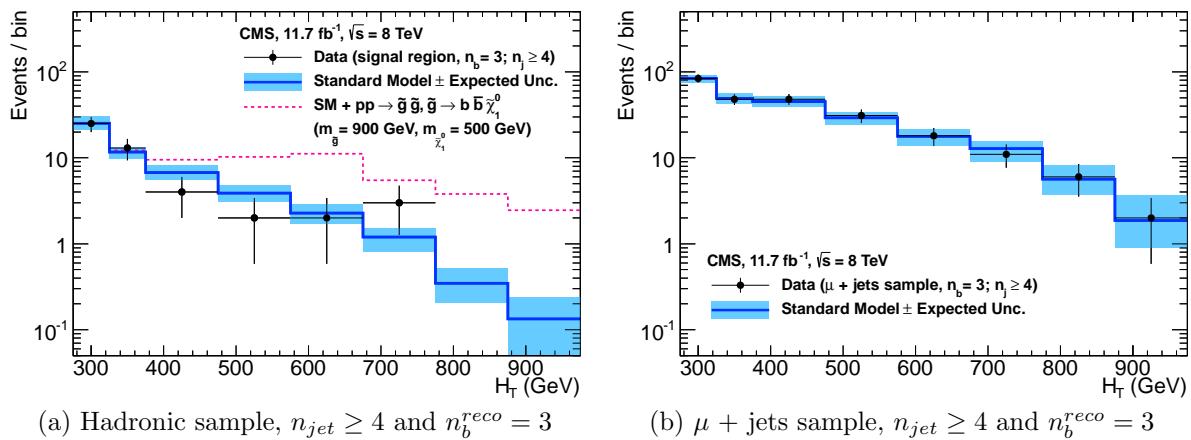
**Figure 5.4.:** Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of  $H_T$  for the (a) hadronic, (b)  $\mu +$  jets, (c)  $\mu\mu +$  jets and (d)  $\gamma +$  jets samples when requiring  $n_b^{reco} = 0$  and  $n_{jet} \geq 4$ . The observed event yields 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.



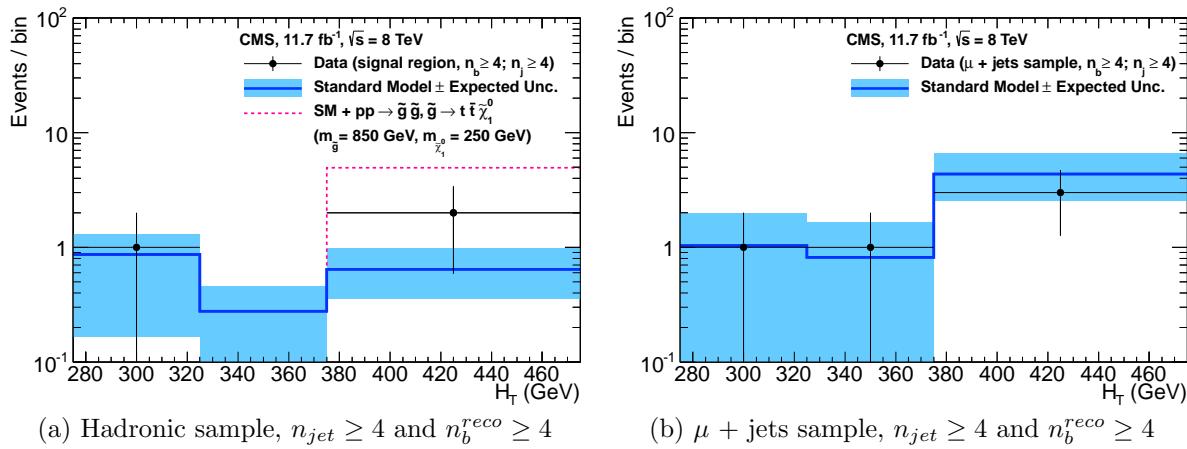
**Figure 5.5.:** Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of  $H_T$  for the (a) hadronic, (b)  $\mu +$  jets, (c)  $\mu\mu +$  jets and (d)  $\gamma +$  jets samples when requiring  $n_b^{reco} = 1$  and  $n_{jet} \geq 4$ . The observed event yields 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.**

2092 **5.2. SUSY**

2093 Limits are set in the parameter space of a set of **SMS** models that characterise both  
2094 natural **SUSY** third generation squark production, and compressed spectra where the  
2095 mass splitting between the particle and **LSP** is small, leading to soft final state jets.  
2096 However as detailed in Section (2.4.1), the individual models are not representative of a  
2097 real physical **SUSY** model as only one decay process is considered. Instead these models  
2098 represent a way to test for signs of specific signatures indicating new physics.

2099 **5.2.1. The  $CL_s$  method**

2100 The  $CL_s$  method [93][94][95] is used to compute the limits for signal models, with the  
2101 one-sided profile likelihood ratio as the test statistic [96].

The test statistic is defined as

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

2102 where

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

2103 represents the profile likelihood ratio, in which  $\mu \equiv f$  from Section (4.8.4), is the  
2104 parameter characterising the signal strength.  $\hat{\mu}$  is defined as the maximum likelihood  
2105 value,  $\hat{\theta}$  the set of maximum likelihood values of the nuisance parameters and  $\theta_\mu$  the set  
2106 of maximum values of the nuisance parameters for a given value of  $\mu$ .

2107 When  $\mu \equiv f = 1$ , the signal model is considered at its nominal production cross section.  
2108 The distribution of  $q_\mu$  is built up via the generation of pseudo experiments in order to  
2109 obtain two distributions for the background (B) and signal plus background (S+B) cases.  
2110 The compatibility of a signal model with observations in data is determined by the  
2111 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  $\text{CL}_s \leq 0.05$  [97].

### 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 the categories used within each interpretation, are made to maximise the signal to background ratio, 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, with parameters for the sparticle mass as well as on the **LSP** mass. 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

Conversely the T2bb, T1tttt, and T1bbbb **SMS** model describe various production and decay mechanisms in the context of third-generation squarks. In this situation considering only 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.

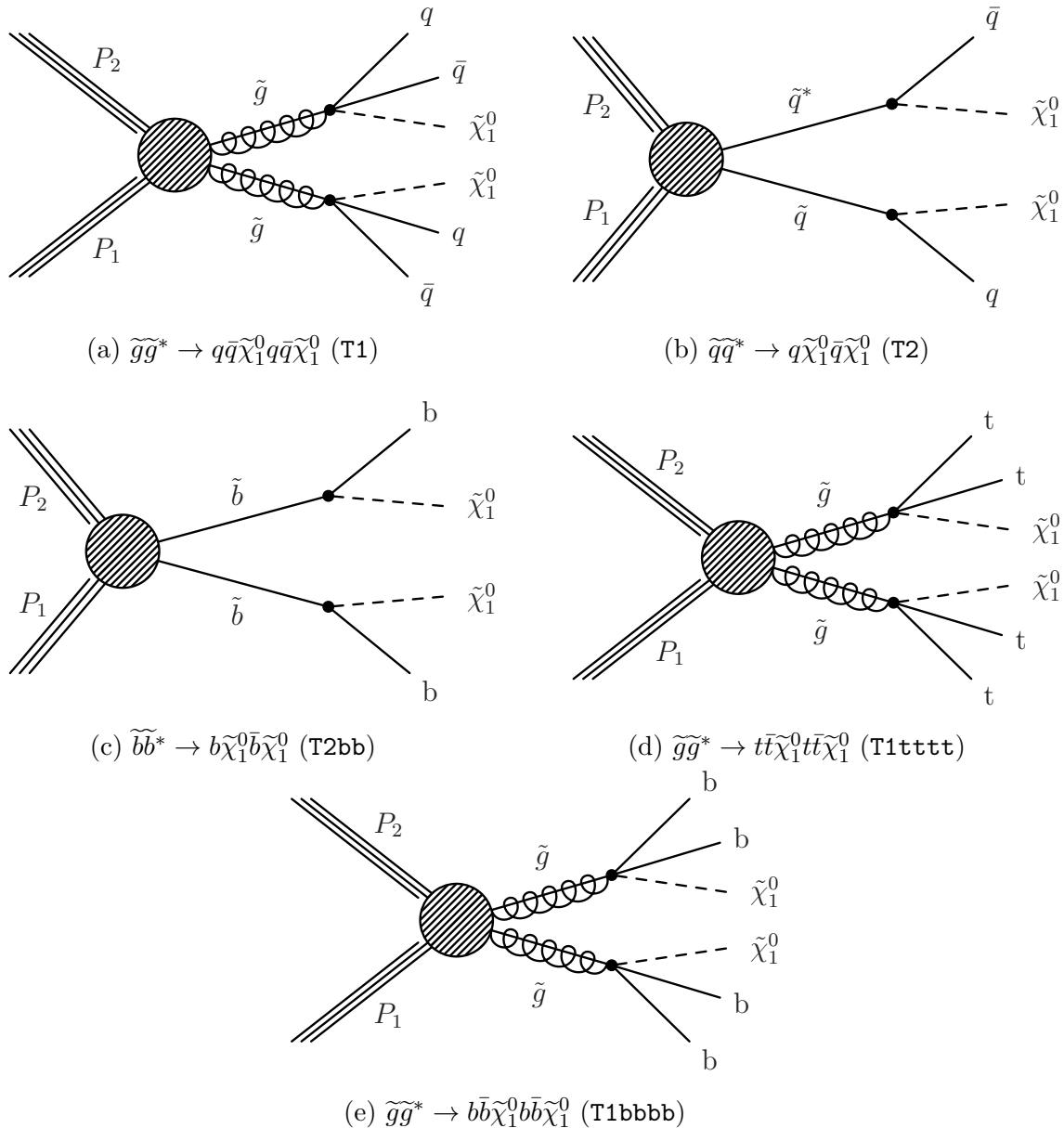
Experimental uncertainties on the **SM** background predictions (10 – 30%, described in Section (4.6.1)), the luminosity measurement (4.4%), and the total acceptance times efficiency of the selection for the considered signal model (12 – 18%, from Section (4.7)) are included in the calculation of the limit.

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	$pp \rightarrow \tilde{q}\tilde{q}^* \rightarrow q\tilde{\chi}_1^0 \bar{q}\tilde{\chi}_1^0$	$\leq 3$	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)	$\sim 1125$	$\sim 650$

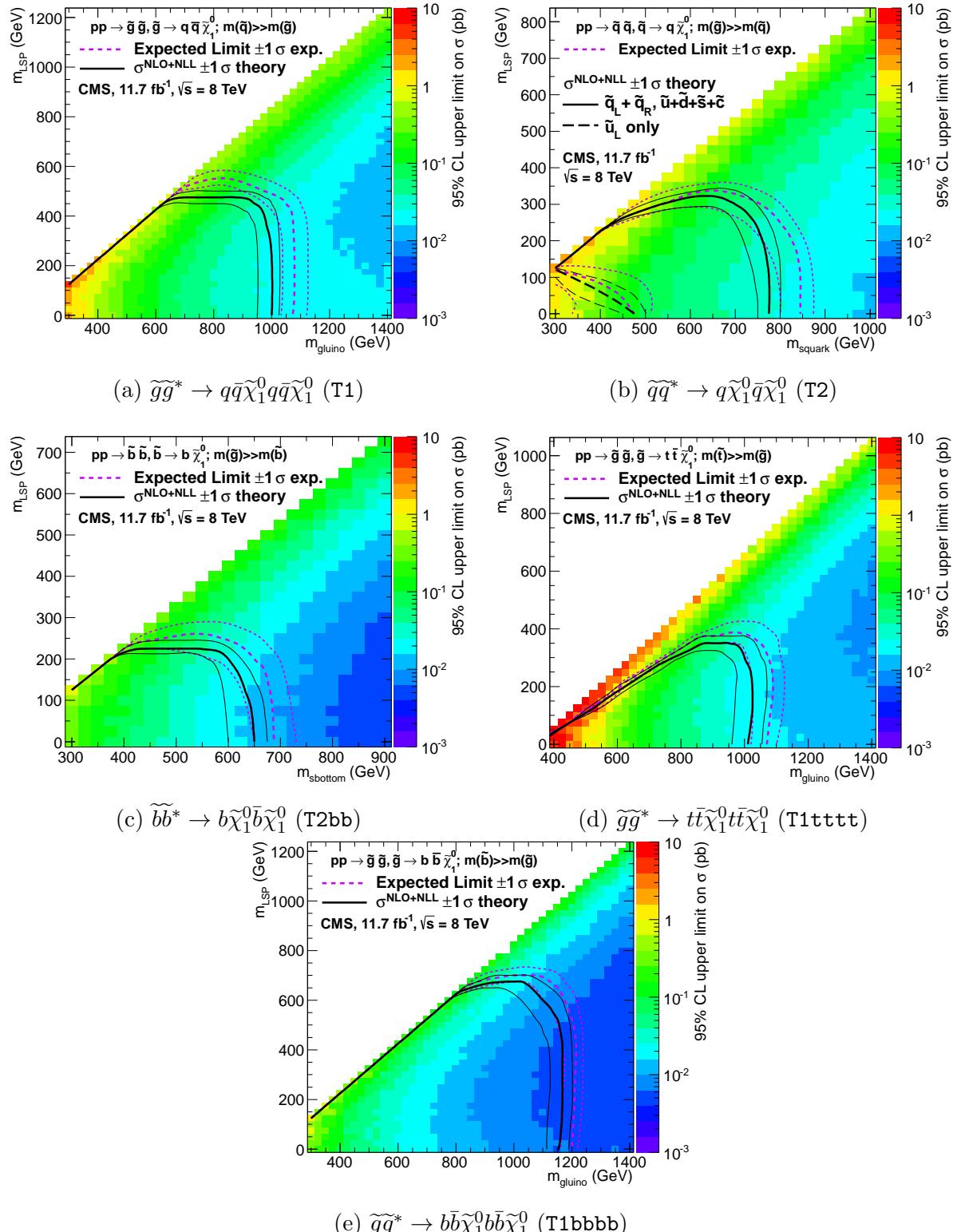
**Table 5.3.:** A table representing the **SMS** models interpreted within the analysis. The model name and production and decay chain is specified in the first two columns. Each **SMS** model is interpreted in specific  $n_{jet}$  and  $n_b^{reco}$  categories which are detailed in the third and fourth 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.

- 2138 Signal efficiency in the kinematic region defined by  $0 < m_{\tilde{g}/\tilde{q}} < 175$  GeV or  $m_{\tilde{g}/\tilde{q}} < 300$   
 2139 GeV is strongly affected by the presence of Initial State Radiation (**ISR**). This region in  
 2140 which direct (i.e., non-**ISR** induced) production is kinematically forbidden due to the  $H_T$   
 2141  $> 275$  GeV requirement, therefore a large percentage of signal acceptance is due to the  
 2142 effect of **ISR** jets. Given the large associated uncertainties, no interpretation is provided  
 2143 for this kinematic region.
- 2144 The estimates on mass limits shown in Table 5.3, are determined conservatively from the  
 2145 observed exclusion based on the theoretical production cross section, minus  $1\sigma$  uncertainty.  
 2146 The most stringent mass limits on pair-produced sparticles are obtained at low **LSP**  
 2147 masses and larger squark and gluino masses due to the high  $p_T$  jets and consequently  
 2148 high  $H_T$  of such signal topologies. The limits are seen to weaken for compressed spectra  
 2149 points closer to the diagonal, where the signal is populates the lower  $H_T$  bins in which  
 2150 more background resides. For all of the considered models, there is an **LSP** mass beyond  
 2151 which no limit can be set, which can be observed from the figures referenced in the table.
- 2152 Two small upwards fluctuations are observed within the data, and are seen at high  $H_T$   
 2153 within the  $n_b^{reco} = 0$  category and at mid- $H_T$  in the  $n_b^{reco} = 1, 2$  categories, see Table 5.2.  
 2154 As each of these fluctuations occur within at least one of the analysis categories that  
 2155 each **SMS** model interpretation is made, the observed exclusions within all **SMS** models  
 2156 are generally found to be weaker than the expected limits in the region of 1-2 standard  
 2157 deviations. In isolation these fluctuations are not significant and additional data would  
 2158 be necessary to make any further conclusions.

Despite these fluctuations, the range of parameter space that can be excluded has been extended with respect to analysis based upon the  $\sqrt{s} = 7$  TeV dataset [98], by up to 225 and 150 GeV for  $m_{\tilde{q}(g)}^{\text{best}}$  and  $m_{LSP}^{\text{best}}$  respectively. The parameter space for light third generation squarks, the main tenet of natural SUSY models, is increasingly squeezed for 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>2164</sup> Searching For Natural SUSY With <sup>2165</sup> B-tag Templates.

<sup>2166</sup> Within this chapter a complementary technique is discussed as a means to predict the  
<sup>2167</sup> distribution of three and four reconstructed b-quark jets in an event. The recent discovery  
<sup>2168</sup> of the Higgs boson has made third-generation “Natural SUSY” models attractive, given  
<sup>2169</sup> that light top and bottom squarks are a candidate to stabilise divergent loop corrections  
<sup>2170</sup> to the Higgs boson mass.

<sup>2171</sup> Using the  $\alpha_T$  search as a base, a simple template fit is employed to estimate the SM  
<sup>2172</sup> background in higher b-tag multiplicities (3-4) from a fit conducted in a low number of  
<sup>2173</sup> reconstructed b-jets (0-2) control region. As a proof-of-concept, the procedure is applied  
<sup>2174</sup> to the SM enriched  $\mu + \text{jets}$  control sample of the  $\alpha_T$  all-hadronic search detailed in  
<sup>2175</sup> Chapter 4, in both data and simulation. To highlight the relative insensitivity of the  
<sup>2176</sup> choice of b-tagging algorithm working point in the effectiveness of the procedure, results  
<sup>2177</sup> are presented using the CSV tagger (introduced in Section (3.3.2)) for the “Loose”,  
<sup>2178</sup> “Medium” and “Tight” working points.

### <sup>2179</sup> 6.1. Concept

<sup>2180</sup> The dominant SM backgrounds of most SUSY searches are typically  $t\bar{t} + \text{jets}$ ,  $W +$   
<sup>2181</sup> jets,  $Z \rightarrow \nu\bar{\nu} + \text{jets}$  or other rare processes with neutrinos in the final state. These  
<sup>2182</sup> processes are characterised by typically having zero or two underlying b-quarks per event.  
<sup>2183</sup> Conversely a third generation squark production signal, such at the T1tttt and T1bbbb  
<sup>2184</sup> models described in the previous chapter, will typically have four underlying b-quarks in

its final state. As SM processes with similar topologies are rare, an excess of  $n_b^{\text{reco}} = 3, \geq 4$  events would be indicative of a potential natural SUSY signature. Therefore the compatibility of the  $n_b^{\text{reco}}$  distribution in data can be tested via the parameterisation of the SM backgrounds in terms of these two most common underlying b-quark topologies.

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

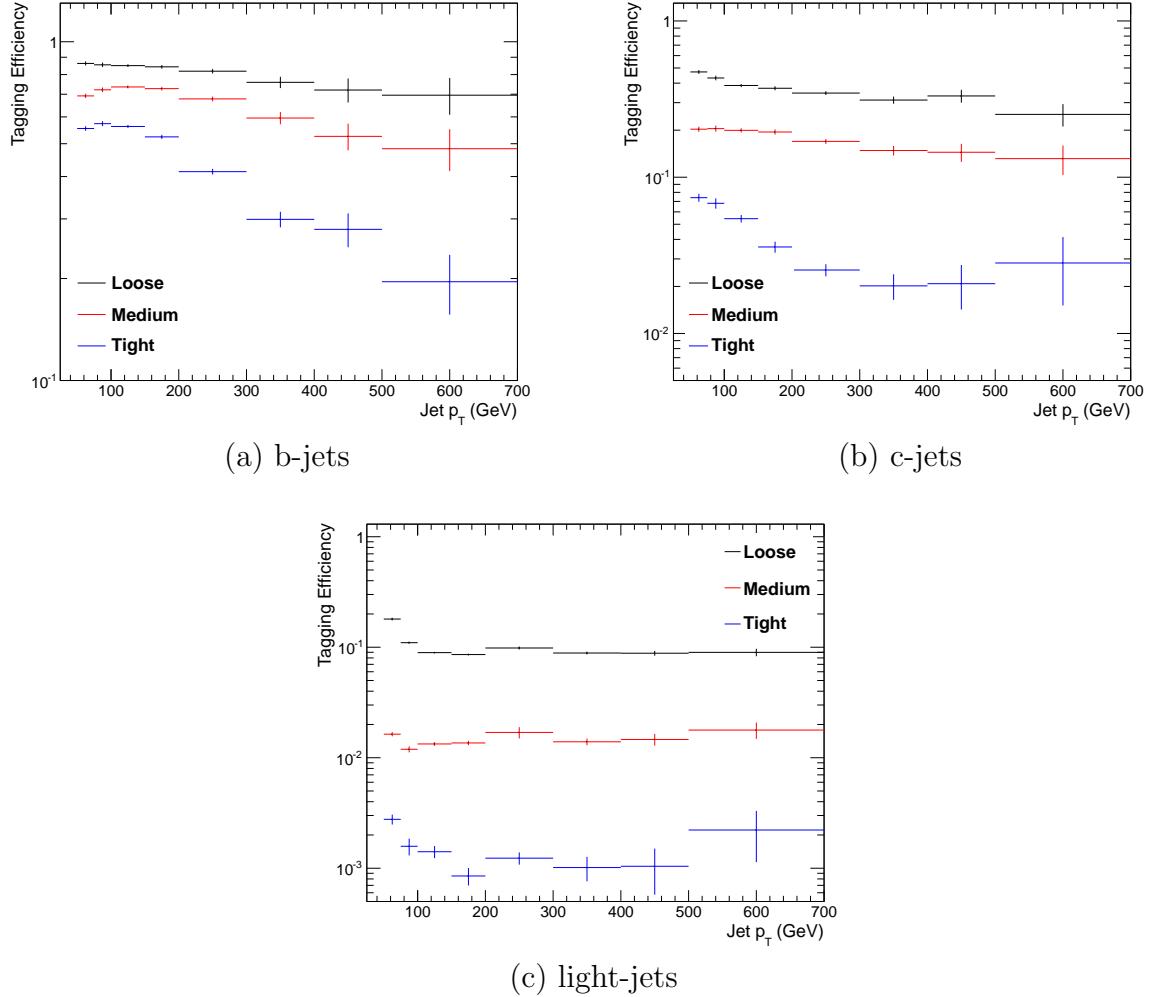
**Table 6.1.:** Typical underlying b-quark content of different SM processes which are common to many SUSY searches.

Thus two templates are defined, Z0 and Z2 (single top processes are a negligible background,  $\sim 1\%$  within the  $\alpha_T$  search, and are combined together with  $t\bar{t}$ ) which represent processes which have an underlying b-quark content of zero or two respectively.

Both these templates are generated through the application of the relevant event selection, and can then be taken from the underlying  $n_b^{\text{reco}}$  distribution directly from simulation. However as discussed within Section (4.5), there are large uncertainties for high  $n_b^{\text{reco}}$  multiplicities due to limited MC statistics. This is particularly prominent for the Z0 templates, where events with a large number of reconstructed b-tags jets are driven primarily by the mis-tagging of light-quarks. Within both the medium and tight working point of the CSV tagger, the expected mis-tagging rate is only around 1 and 0.1% respectively, leading to large uncertainties in the template shape in this region. Therefore to improve the statistical precision of the predictions within the signal region, the formula method introduced in Section (4.5.1) is used.

The generation of the template shapes, are dependant upon the jet-flavour content and b-tagging rate within the phase space of interest, with the tagging probabilities of a jet being a function of the jet  $p_T$ , the pseudo-rapidity  $|\eta|$ , and jet-flavour. This can be observed in Figure 6.1, where the b-tagging / c-quark mis-tagging / light mis-tagging efficiency for the three working points of the CSV tagger are shown as a function of jet  $p_T$ .

Before the template shapes are determined and applied to data, the relevant jet  $p_T$  and  $\eta$  corrections are applied to correct the measured b-tagging rate in simulation to that of



**Figure 6.1.:** The b-tagging (a), c-quark mis-tagging (b), and light-quark mis-tagging rate (c) as measured in simulation after the  $\alpha_T$  analysis,  $\mu + \text{jets}$  control sample selection in the region  $H_T > 375$ .

2210 data, as specified in Section (4.5.3), which propagate through to the average determined  
2211 b-tagging rates per analysis  $H_T$  bin, as in the  $\alpha_T$  analysis.

2212 These two template shapes once generated from simulation, can then be fitted to data in  
2213 a low  $n_b^{reco}$  control region (0-2), by allowing the normalisation constants  $\theta_{Z0}$  and  $\theta_{Z2}$  of  
2214 the two templates to float. The best fit values of  $\theta_{Z0}$  and  $\theta_{Z2}$  are used, along with the  
2215 knowledge of the template shapes, to extrapolate an estimate in the high  $n_b^{reco}$  signal  
2216 region (3,4), which is then compared to what is observed in data. Any large excess in  
2217 data compared to the template prediction would indicate that the  $n_b^{reco}$  distribution is  
2218 not adequately described by the SM backgrounds which compose the templates. This  
2219 method can, in principle, be applied to any analysis where the signal hypothesis has a

2220 larger underlying b-quark spectra than the **SM** backgrounds, as it solely relies on fitting  
2221 to the shape of the  $n_b^{\text{reco}}$  distribution.

2222 However in the scenario where a **SUSY** signal sits at a low number of underlying b-quarks,  
2223 the template would be unable to discriminate between this signal and background and  
2224 would be accommodated within the fit in the control region. This will be the case unless  
2225 the jet  $p_T$  distribution of the signal and background were drastically different, in which  
2226 case there would, anyway be many more sensitive ways to establish the presence of a  
2227 signal in the data than this method. Indeed the template method is only really applicable  
2228 to the hypothesis that any signal resides at high  $n_b^{\text{reco}}$  and that the control region  $0 \leq n_b^{\text{reco}} \leq 2$   
2229 is indeed signal free.

## 2230 6.2. Application to the $\alpha_T$ Search

2231 As detailed in the previous chapter, the  $\alpha_T$  analysis is a search for **SUSY** particles  
2232 in all-hadronic final states, utilising the kinematic variable  $\alpha_T$  to suppress QCD to a  
2233 negligible level. **SM** enriched control samples are used to estimate the background within  
2234 an all-hadronic signal region.

2235 The selection for the  $\mu + \text{jets}$  control samples defined in Section (4.2.3) is used to  
2236 demonstrate the template fitting procedure both conceptually in simulation, and also  
2237 when applied in data. This is chosen, as such a selection is dominated by events stemming  
2238 from the **SM** processes with little or no signal contamination from potential new physics.  
2239 Neither are contributions from rare **SM** processes with a higher underlying b-quark  
2240 content (e.g.  $t\bar{t}b\bar{b}$ ) expected. For these reasons, there is a degree of confidence that the  
2241 procedure should adequately describe the observations in data when extrapolated to the  
2242 signal region.

2243 The analysis presented here is binning in source jet multiplicity bins, of 3, 4 and  $\geq 5$   
2244 reconstructed jets per event (di-jet events are not included as there is no contribution  
2245 to the high  $n_b^{\text{reco}}$  region (3,4)) , in order to reduce the kinematic jet  $p_T$  dependence.  
2246 Furthermore the analysis is split into three  $H_T$  regions,

2247 • 275-325 GeV

2248 • 325-375 GeV

2249 •  $> 375$  GeV

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

### 6.2.1. Proof of principle in simulation

In order to demonstrate that the template procedure produces accurate predictions within simulation, the simulation samples in the analysis are firstly split into two to allow for statistically independent fits to be performed.

By combining the relevant ingredients necessary to employ the formula method,  $n_b^{\text{reco}}$  templates for  $Z = 0$  and  $Z= 2$  are generated individually for each  $n_{\text{jet}}$  and  $H_T$  bin using one half of each simulation sample. A fit of these two templates is then performed in the low  $n_b^{\text{reco}}$  (0-2) region, back to the sum of the other halves of each simulation sample in order to check that the relevant information can be recovered in the  $n_b^{\text{reco}}$  signal region (3-4).

The fits are performed independently within each of the defined analysis bins to reduce the dependence of the shapes of these distributions on simulation. The half of the simulation sample for which the templates are fitted too, are taken directly from simulation, extending this procedure to also be a validation of the formula method in accurately describing the  $n_b^{\text{reco}}$  distribution within the control region itself. Additionally as this test is performed in simulation, the relevant corrections of the b-tagging rates between data and simulation are *not* applied.

Within Figure 6.2, the results of this fitting procedure are shown for each **CSV** working point. Results are presented for the  $n_{\text{jet}} \geq 5$  category, using the  $\mu + \text{jets}$  control sample selection in the inclusive  $H_T > 375$  GeV analysis bin. The grey bands represent the statistical uncertainty on the template shapes. Additional fits are shown for other  $n_{\text{jet}}$  categories can be found within Appendix D.1.

Furthermore the extrapolated fit predictions within the high  $n_b^{\text{reco}}$  signal region, are summarised for all  $H_T$  bins and working points in Table 6.2.

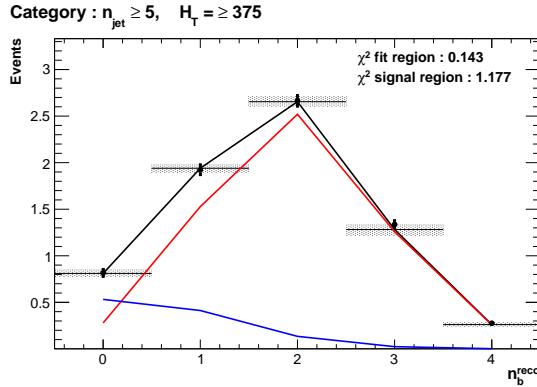
The pull distributions for all the fits performed can be found in Appendix D.2, and are compatible with a mean of zero and standard deviation of one, showing no obvious bias to the fitting procedure. The good overall agreement summarised in the table validates both the formula method used to generate the templates as well as the method of extrapolation to the high  $n_b^{\text{reco}}$  signal region. The application of this method to the same selection in a

$H_T$	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 $n_b = 3$	$820.4 \pm 26.7$	$376.3 \pm 11.9$	$780.1 \pm 15.1$
Simulation $n_b = 4$	$68.2 \pm 3.9$	$27.6 \pm 2.7$	$91.28 \pm 4.9$
Template $n_b = 4$	$72.5 \pm 4.7$	$28.25 \pm 2.34$	$84.4 \pm 3.8$
Medium working point			
Simulation $n_b = 3$	$133.7 \pm 5.7$	$74.5 \pm 4.5$	$164.2 \pm 6.4$
Template $n_b = 3$	$132.8 \pm 4.8$	$74.5 \pm 3.9$	$159.9 \pm 5.7$
Simulation $n_b = 4$	$1.6 \pm 0.6$	$0.6 \pm 0.4$	$3.4 \pm 0.9$
Template $n_b = 4$	$1.8 \pm 0.2$	$1.1 \pm 0.2$	$4.1 \pm 0.4$
Tight working point			
Simulation $n_b = 3$	$26.9 \pm 2.6$	$13.9 \pm 1.9$	$31.8 \pm 2.9$
Template $n_b = 3$	$24.7 \pm 1.5$	$13.8 \pm 1.2$	$28.1 \pm 1.5$
Simulation $n_b = 4$	$0.5 \pm 0.4$	-	-
Template $n_b = 4$	$0.1 \pm 0.1$	$0.1 \pm 0.1$	$0.2 \pm 0.1$

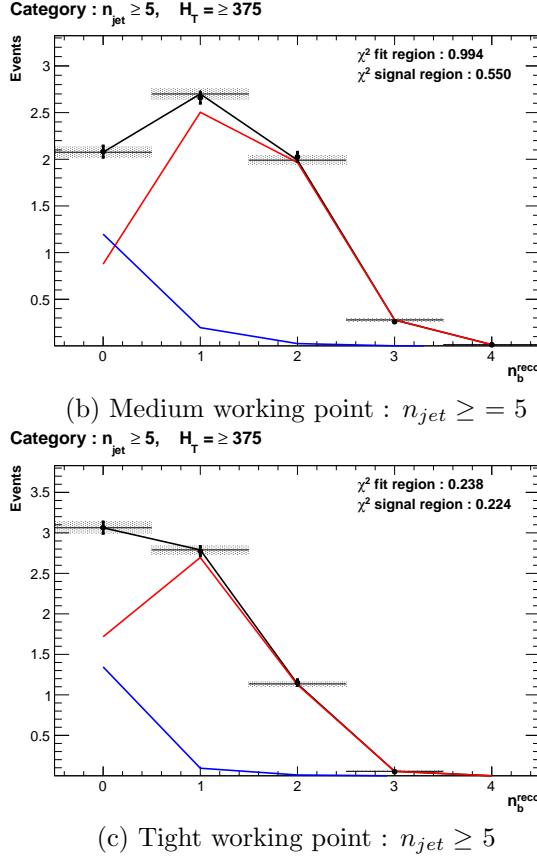
**Table 6.2.:** Summary of the fit predictions in the  $n_b^{reco}$  signal region for  $n_{jet} = 3, = 4, \geq 5$ . The fit region is  $n_b^{reco} = 0, 1, 2$  and simulation yields are normalised to an integrated luminosity of  $10 \text{ fb}^{-1}$ . The uncertainties quoted on the template yields are purely statistical.

2281 data control sample, is now used to demonstrate necessary control over the efficiency and  
 2282 mis-tagging rates when b-tagging scale factors are applied, and to test the assumption of  
 2283 no signal contamination with the  $\mu + \text{jets}$  control sample.

2284



(a) Loose working point :  $n_{jet} \geq 5$

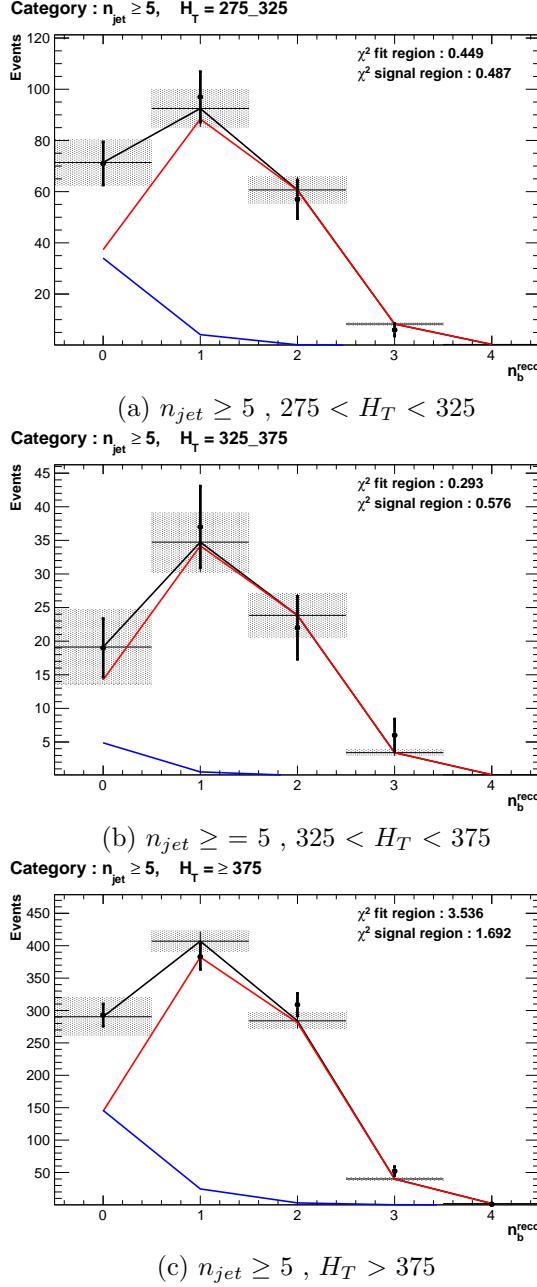


**Figure 6.2.:** The results of fitting the  $Z = 0$  and  $Z = 2$  templates to the  $n_b^{reco} = 0, 1, 2$  bins taken directly from simulation in the region  $H_T > 375$  GeV, for the  $n_{jet} \geq 5$  category. The blue template represents  $Z = 0$ , while the red template represents  $Z = 2$ . Grey bands represent the statistical uncertainty of the fit. The  $\chi^2$  parameter displayed represents the goodness of fit to the low  $n_b^{reco}$  (0-2) control region.

### 2285 6.2.2. Results in a data control sample

- 2286 The procedure is now applied to the 2012 8 TeV dataset in the  $\mu +$  jets control sample, to  
 2287 establish the validity of this method in data. The relevant data to simulation b-tagging  
 2288 scale factors are applied to produce corrected values of the efficiency and mis-tagging  
 2289 rates within each analysis bin [86].
- 2290 Figure 6.3 shows the results of the templates derived from simulation to each of the three  
 2291 defined  $H_T$  bins, in the  $n_{jet} \geq 5$  category for the medium working point CSV tagger (the  
 2292 same working point used within the  $\alpha_T$  analysis). Grey bands represent the statistical  
 2293 uncertainty of the fit combined in quadrature with the systematic uncertainties of varying

the data to simulation scale factors up and down by their b-tag scale factor systematic uncertainties. Additional fit results for other jet multiplicities are found in Appendix D.3



**Figure 6.3.:** The results of fitting the  $Z = 0$  and  $Z = 2$  templates to the  $n_b^{reco} = 0, 1, 2$  bins taken directly from data, for the  $n_{jet} \geq 5$  category and medium CSV working point. The blue template represents  $Z = 0$ , while the red template represents  $Z = 2$ . The  $\chi^2$  parameter displayed represents the goodness of fit to the low  $n_b^{reco}$  (0-2) control region.

2296 The numerical results and extrapolation to the  $n_b^{reco} = 3, 4$  bins for all  $H_T$  and working  
 2297 points is shown in Table 6.3.

$H_T$	275-325	325-375	>375
Loose working point			
Data $n_b = 3$	838	394	717
Template $n_b = 3$	$861.8 \pm 38.1$	$372.1 \pm 18.4$	$673.2 \pm 34.5$
Data $n_b = 4$	81	43	81
Template $n_b = 4$	$78.5 \pm 5.8$	$27.6 \pm 2.6$	$78.6 \pm 3.3$
Medium working point			
Data $n_b = 3$	137	79	152
Template $n_b = 3$	$131.2 \pm 4.3$	$66.1 \pm 2.9$	$137.8 \pm 5.7$
Data $n_b = 4$	1	1	3
Template $n_b = 4$	$1.8 \pm 0.1$	$0.9 \pm 0.1$	$3.1 \pm 0.2$
Tight working point			
Data $n_b = 3$	24	15	25
Template $n_b = 3$	$23.0 \pm 0.9$	$12.9 \pm 0.6$	$20.3 \pm 1.1$
Data $n_b = 4$	0	0	1
Template $n_b = 4$	$0.1 \pm 0.1$	$0.1 \pm 0.1$	$0.2 \pm 0.1$

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

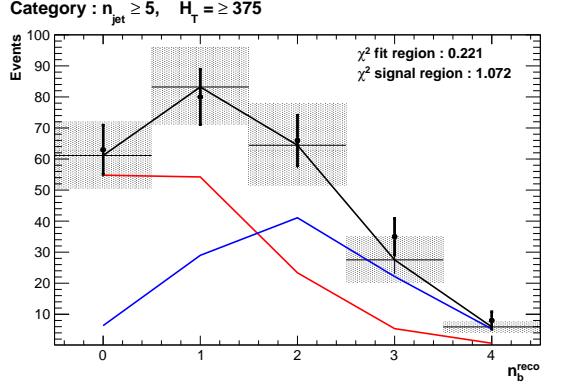
2298 When this method is applied to the  $\mu +$  jets control sample, it is expected that good  
 2299 agreement would be observed between prediction and observation (in the absence of signal  
 2300 contamination) if the procedure is valid. The good compatibility for all working points  
 2301 as shown in the table, demonstrate that this is the case. However no such assumptions  
 2302 can be made when applied to the signal region of the  $\alpha_T$  search.

### 2303 6.2.3. Application to the $\alpha_T$ hadronic search region

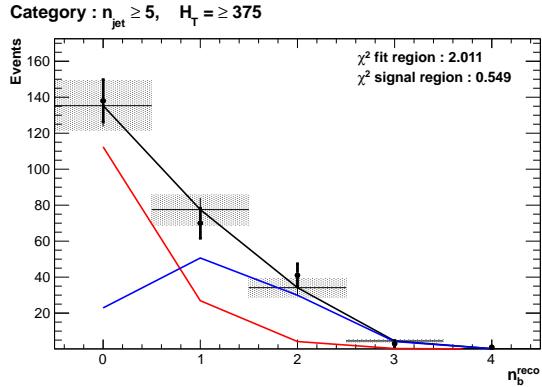
2304 As an accompaniment to the background estimation methods outlined in the  $\alpha_T$  search,  
 2305 the b-tag template method offers a complementary way of testing the **SM** only background  
 2306 hypothesis within the hadronic signal region of the search. In the presence of a natural  
 2307 **SUSY** signature containing four underlying  $\tilde{b}$  or  $\tilde{t}$  squarks, which subsequently decay to t  
 2308 or b quarks, the number of reconstructed  $n_b^{reco} = 3, \geq 4$  events will be enhanced.

2309 Figure 6.4 show the the results of the templates derived from simulation to each of  
 2310 the three **CSV** working points, in the  $n_{jet} \geq 5, H_T > 375 \text{ GeV}$  category. Grey bands  
 2311 represent the statistical uncertainty of the fit combined in quadrature with the systematic  
 2312 uncertainties of varying the data to simulation scale factors up and down by their

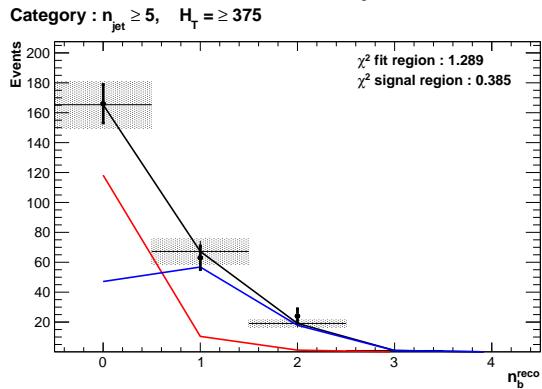
2313 measured systematic uncertainties. Additional fit results for other jet multiplicities are  
2314 found in Appendix D.4



(a) Loose working point :  $n_{jet} \geq 5, H_T > 375$



(b) Medium working point :  $n_{jet} \geq 5, H_T > 375$



(c) Tight working point :  $n_{jet} \geq 5, H_T > 375$

**Figure 6.4.:** The results of fitting the  $Z = 0$  and  $Z = 2$  templates to the  $n_b^{reco} = 0, 1, 2$  bins taken from data, in the  $n_{jet} \geq 5$  and  $H_T > 375$  category for all CSV working points. The blue template represents  $Z = 0$ , while the red template represents  $Z = 2$ . The  $\chi^2$  parameter displayed represents the goodness of fit to the low  $n_b^{reco}$  (0-2) control region.

2315 The numerical results and extrapolation to the  $n_b^{reco} = 3, 4$  bins for all  $H_T$  and working  
 2316 points are shown in Table 6.4. No excess of data is found and predictions from this  
 2317 method are found to be compatible with the  $\alpha_T$  maximum likelihood fit results from  
 2318 Table 5.2.

$H_T$	275-325	325-375	>375
Loose working point			
Data $n_b = 3$	198	85	126
Template $n_b = 3$	$207.1 \pm 33.3$	$103.4 \pm 10.9$	$124.98 \pm 16.2$
Data $n_b = 4$	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 $n_b = 3$	28	15	12
Template $n_b = 3$	$24.4 \pm 1.7$	$12.7 \pm 1.2$	$19.9 \pm 2.8$
Data $n_b = 4$	1	0	2
Template $n_b = 4$	$0.3 \pm 0.2$	$0.3 \pm 0.1$	$0.5 \pm 0.2$
Tight working point			
Data $n_b = 3$	5	2	0
Template $n_b = 3$	$4.03 \pm 0.3$	$2.4 \pm 0.3$	$3.1 \pm 0.3$
Data $n_b = 4$	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  $\mu +$  jets control sample, for  $n_{jet} = 3, = 4, \geq 5$ . The fit region is  $n_b^{reco} = 0, 1, 2$  using  $11.7 \text{ fb}^{-1}$  of data at  $\sqrt{s} = 8\text{TeV}$ . The uncertainties quoted on the template yields are purely statistical.

### 2319 6.3. Summary

2320 A **SUSY** signature such as one from gluino-induced third-generation squark production,  
 2321 would result in a final state with an underlying b-quark content greater than two. In  
 2322 order to be able to discriminate such signatures from the **SM** background, templates are  
 2323 generated based on a parameterisation of the number of the **SM** processes, where the  
 2324 underlying b-quarks per event is typically zero or two. These templates are then fit to  
 2325 data in a low  $n_b^{reco}$  (0-2) control region in order to extrapolate a prediction in a high  
 2326  $n_b^{reco}$  (3-4) signal region. This approach is built upon the assumptions that the defined  
 2327 control region is almost entirely free of any possible signal contamination from either a  
 2328 third generation **SUSY** signal, or other possible event topologies with a small number of  
 2329 b quarks in the final state.

2330 The method was demonstrated both in simulation and also in data, using the **SM** enriched  
 2331  $\mu +$  jets selection from the  $\alpha_T$  search, to prove conceptually and experimentally that the

method is valid and there is adequate control over the efficiency and mis-tagging rates in data for all working points of the **CSV** tagger. Additionally this method was also applied to the  $\alpha_T$  analysis signal region, where good agreement is observed between the predictions from the template extrapolations, observations in data and the background estimation method of the  $\alpha_T$  analysis.

# Chapter 7.

## <sup>2337</sup> Conclusions

<sup>2338</sup> A search for supersymmetry is presented based on a data sample of pp collisions collected  
<sup>2339</sup> at  $\sqrt{s} = 8$  TeV, corresponding to an integrated luminosity of  $11.7 \pm 0.5$  fb<sup>-1</sup>. Final  
<sup>2340</sup> states with two or more jets and significant  $\cancel{E}_T$ , a typical final state topology of R-parity  
<sup>2341</sup> conserving SUSY models have been analysed and in which the  $\alpha_T$  variable is utilised  
<sup>2342</sup> as the main discriminator between balanced multi-jet backgrounds and those with real  
<sup>2343</sup> missing energy. An additional complementary approach using a template method to  
<sup>2344</sup> estimate the b-tag jet distribution of SM processes, to search for gluino induced third  
<sup>2345</sup> generation squark SUSY production is also introduced, in which the  $\alpha_T$  search selection  
<sup>2346</sup> is applied in both simulation and data to validate this technique.

<sup>2347</sup> Additionally a measurement the performance of the Level-1 trigger for jets and energy  
<sup>2348</sup> sum quantities is also presented. These studies quantify any change in level-1 performance  
<sup>2349</sup> after the introduction of a 5 GeV jet seed threshold into the jet algorithm configuration.  
<sup>2350</sup> This change is introduced to facilitate a reduction in the rate at which jets are formed  
<sup>2351</sup> at level-1 from pile-up jets which are not of interest to physics analyses. This change  
<sup>2352</sup> is necessary to ensure that trigger thresholds can be maintained at lower values, in  
<sup>2353</sup> the presence of an increasing number of pile-up interactions per event over the 2012  
<sup>2354</sup> run period. No significant change in single jet trigger efficiencies is observed and good  
<sup>2355</sup> performance is observed for a range of level-1 quantities.

<sup>2356</sup> Within the SUSY search presented in this analysis, the sum of standard model backgrounds  
<sup>2357</sup> binned in  $H_T$ ,  $n_b^{\text{reco}}$  and  $n_{\text{jet}}$  categories are estimated from a simultaneous binned likelihood  
<sup>2358</sup> fit to a hadronic signal selection and  $\mu + \text{jets}$ ,  $\mu\mu + \text{jets}$ , and  $\gamma + \text{jets}$  control samples.  
<sup>2359</sup> Systematic errors due to theory, detector effects and analysis choices are quantified  
<sup>2360</sup> through the use of data driven closure tests and accounted for in the final interpretation,  
<sup>2361</sup> where observations in data are found to be compatible with a SM only hypothesis.

- 2362 In the absence of a signal like excess the analysis is further interpreted in a set of **SMS**  
2363 models. In the considered models with gluino pair production and for small **LSP** masses,  
2364 exclusion limits of the gluino mass are in the range 950-1125 GeV. For **SMS** models  
2365 with direct squark pair production, first or second generation squarks are excluded up to  
2366 around 775 GeV and bottom squarks are excluded up to 600 GeV, again for small **LSP**  
2367 masses. In the context of ‘natural’ **SUSY** models, with many reconstructed b-jets in the  
2368 final state, limits are set in the range of 975-1125 GeV again for large mass splittings  
2369 between the parents squark and the **LSP**.
- 2370 The template method, whose purpose is to identify any excess in data arising from third  
2371 generation signatures, finds results that are compatible with the  $\alpha_T$  search and a **SM** only  
2372 hypothesis at a high number of reconstructed b-jets. As light third generation squarks  
2373 are an important feature of ‘natural’ **SUSY** models if they are to solve the fine tuning  
2374 problem [99], the limits imposed through interpretations in the T1bbbb and T1tttt **SMS**  
2375 models within the  $\alpha_T$  search, put pressure on such theories, by squeezing the parameter  
2376 space in which ‘natural’ **SUSY** can reside.

2377

# Appendix A.

## <sup>2378</sup> **Miscellaneous**

### <sup>2379</sup> **A.1. Jet Identification Criteria**

<sup>2380</sup> For Calo jets the following 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>2381</sup> For PF jets the following 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.

2382 **A.2. Primary Vertices**

2383 The pileup per event is defined by the number of 'good' reconstructed primary vertices  
2384 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.

---

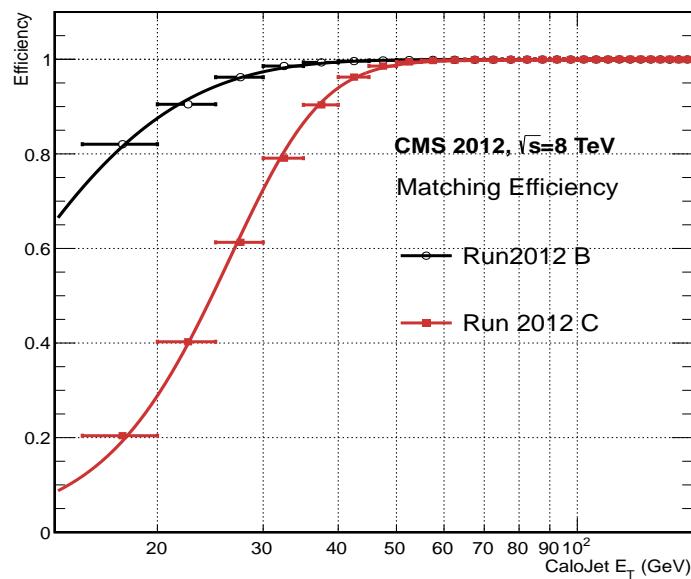
**Table A.3.:** Criteria for a vertex in an event to be classified as a 'good' reconstructed primary vertex.

# Appendix B.

## <sup>2385</sup> L1 Jets

### <sup>2386</sup> B.1. Jet matching efficiencies

<sup>2387</sup> The single jet turn-on curves are derived from events independent of whether the leading  
<sup>2388</sup> jet in an event is matched to a Level 1 jet using  $\Delta R$  matching detailed in Section (3.4.3)  
<sup>2389</sup> or not. These turn-ons are produced from events which are not triggered on jet quantities  
<sup>2390</sup> and therefore it is not guaranteed that the lead jet of an event will be seeded by a Level  
<sup>2391</sup> 1 jet. 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.

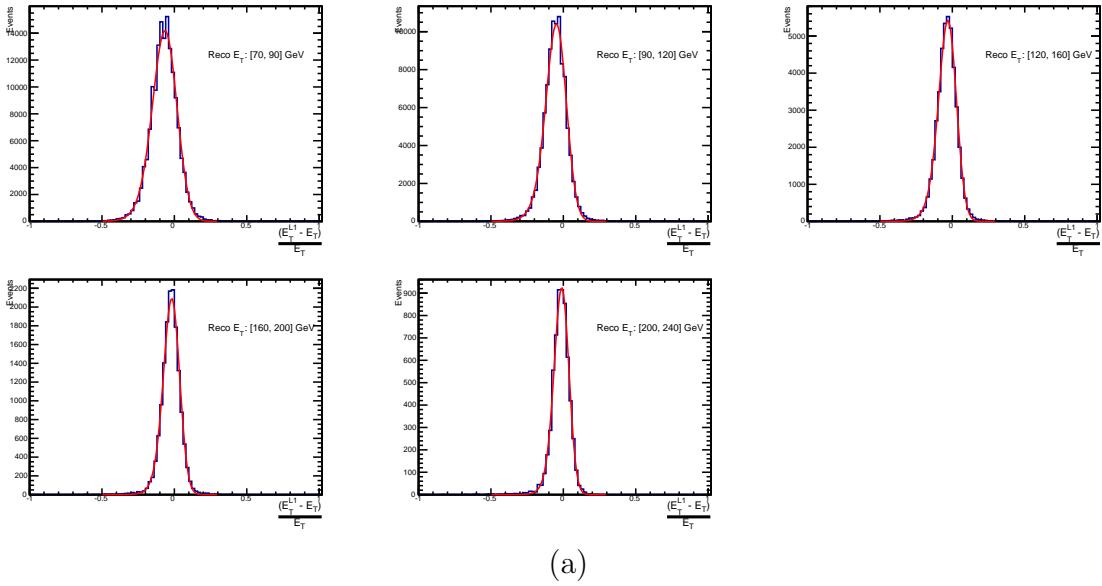
Run Period	$\mu$	$\sigma$
2012B	$6.62 \pm 0.01$	$0.79 \pm 0.03$
2012C	$19.51 \pm 0.03$	$7.14 \pm 0.02$

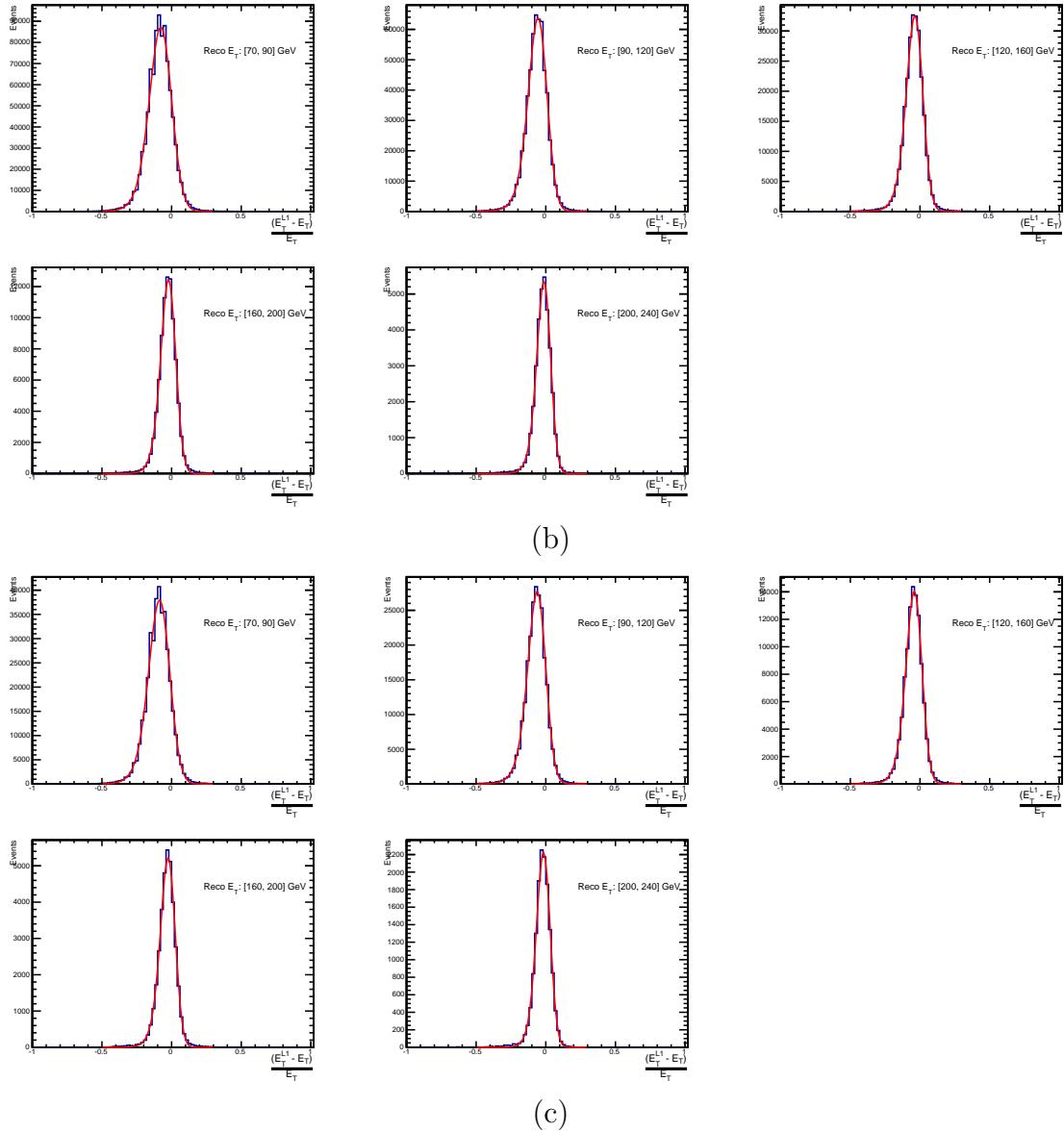
**Table B.1.:** Results of a cumulative EMG function fit to the turn-on curves for the matching efficiency of the leading jet in an event to a Level-1 jet in run 2012C and 2012B data, measured in an isolated muon triggered sample. The turn-on point,  $\mu$ , and resolution,  $\sigma$ , are measured with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

2392 It can be seen that the turn on is sharper during the 2012B run period. The seed  
 2393 threshold requirement of a 5 GeV jet seed in run 2012C results in more events in which  
 2394 even the lead offline jet does not have an associated L1 jet. For larger jet  $E_T$  thresholds,  
 2395 typical of thresholds used in physics analyses, 100% efficiency is observed, and therefore  
 2396 this effect has no impact to overall physics performance.

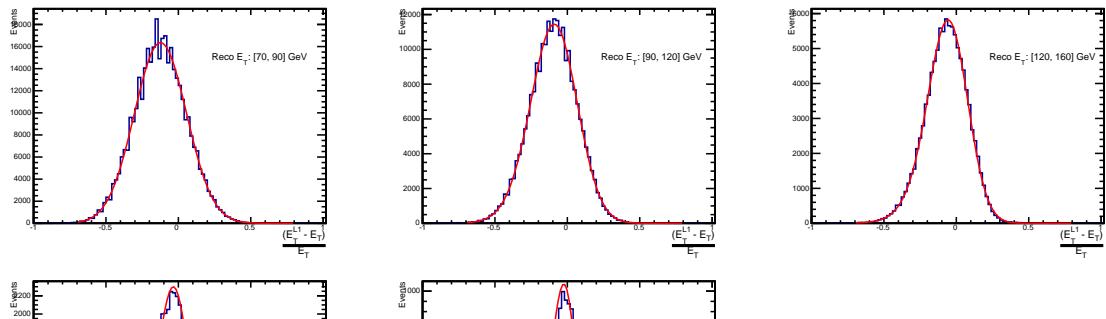
2397 The matching efficiencies have a  $\mu$  values of 6.62 GeV and 19.51 GeV for Run 2012B  
 2398 and 2012C respectively and is shown in Table B.1.

## 2399 B.2. Leading Jet Energy Resolution

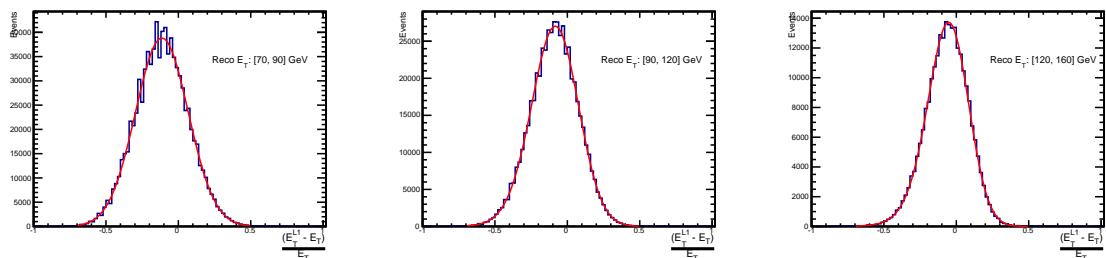




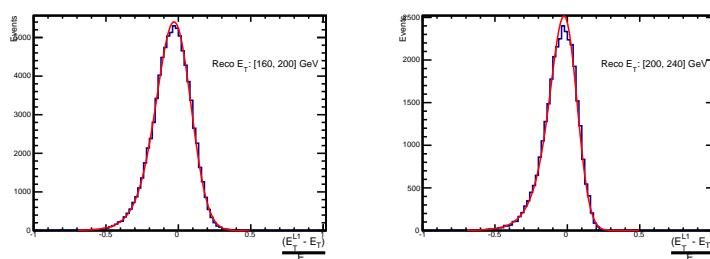
**Figure B.2.:** Resolution plots of the leading offline jet Calo  $E_T$  measured as a function of  $\frac{(L1 E_T - \text{Offline } E_T)}{\text{Offline } E_T}$  for low (a), medium (b) and high (c) pile-up conditions as defined in Section (3.4.4).

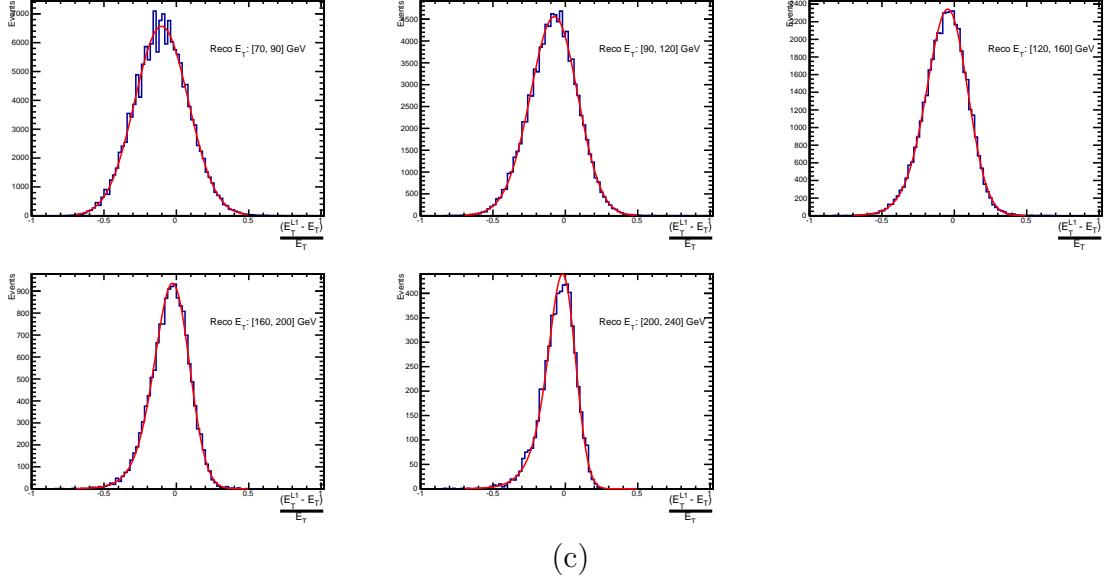


(a)



(b)



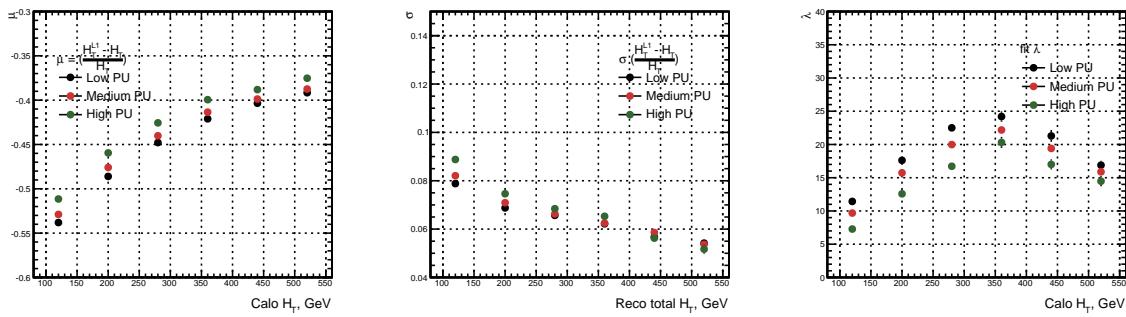


(c)

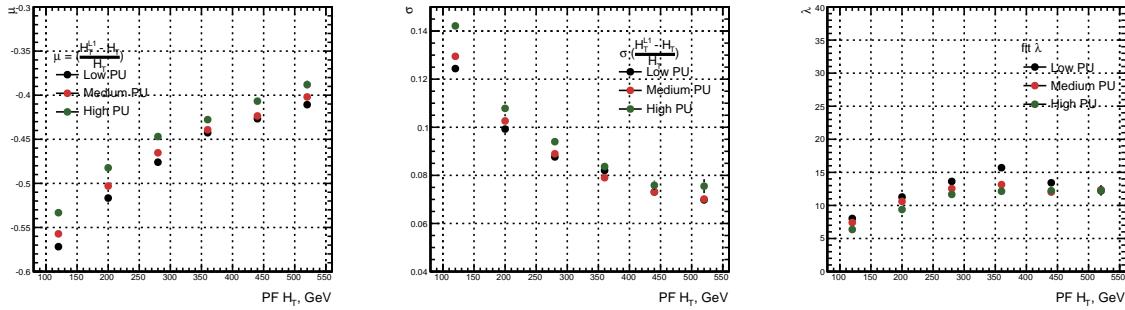
**Figure B.3.:** Resolution plots of the leading offline jet PF  $E_T$  measured as a function of  $\frac{(L1 E_T - \text{Offline } E_T)}{\text{Offline } E_T}$  for low (a), medium (b) and high (c) pile-up conditions as defined in Section (3.4.4).

### 2400 B.3. Resolution for Energy Sum Quantities

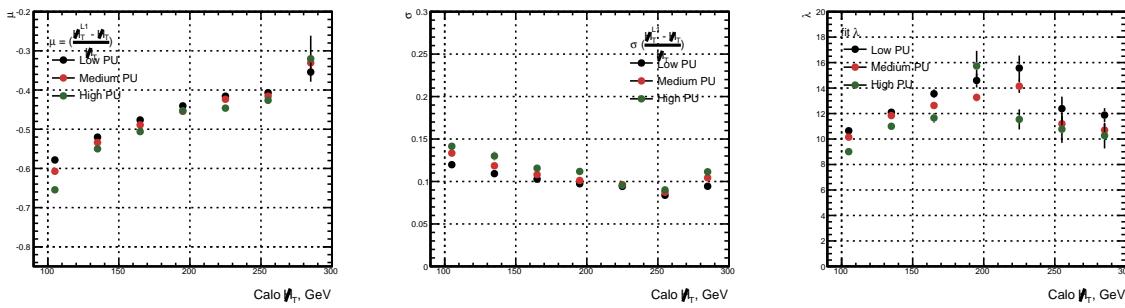
2401 The following plots show the resolution parameters for energy sum quantities as a function  
 2402 of the quantity ( $q$ ) itself. In this case, The  $\mu$ ,  $\sigma$  and  $\lambda$  fit values to an **EMG** function  
 2403 defined by Equation (3.3) for each of the individual  $\frac{(L1 q - \text{Offline } q)}{\text{Offline } q}$  distributions, in bins of  
 2404 the quantity  $q$  is displayed.



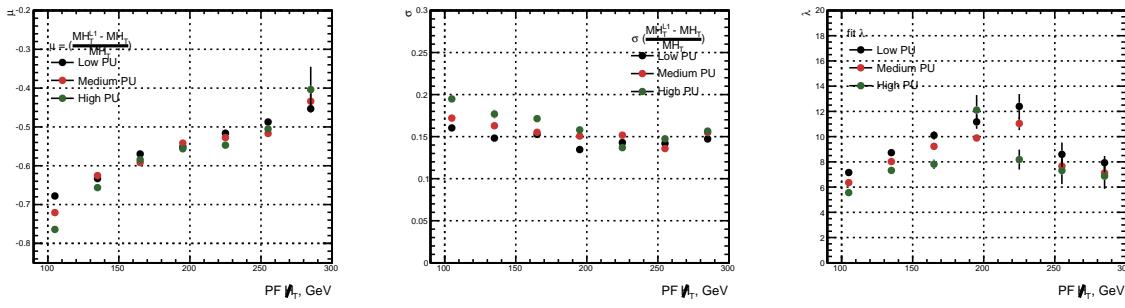
**Figure B.4.:**  $H_T$  resolution parameters in bins of Calo  $H_T$  measured for the defined low, medium and high pile-up conditions. Shown are the mean  $\mu$  (left), resolution  $\sigma$  (middle) and  $\lambda$  (right) fit values to an **EMG** function for the  $\frac{(L1 H_T - H_T)}{H_T}$  distributions.



**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  $H_T$  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.

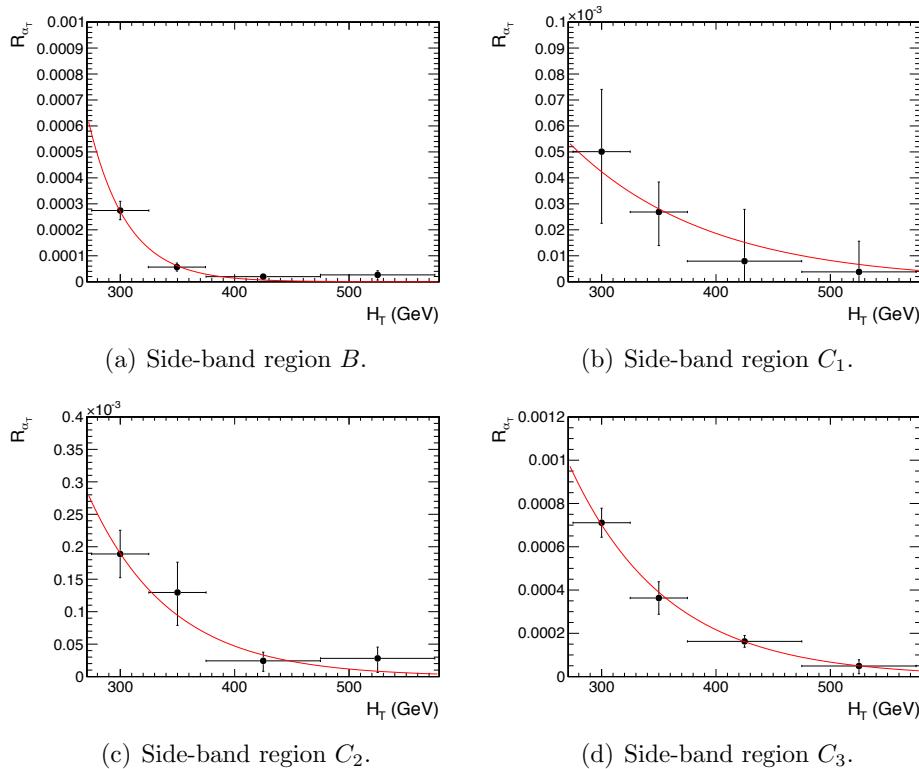


**Figure B.7.:**  $H_T$  resolution parameters in bins of  $PF\ H_T$  measured for the defined low, medium and high pile-up conditions. Shown are the mean  $\mu$  (left), resolution  $\sigma$  (middle) and  $\lambda$  (right) fit values to an EMG function for the  $\frac{(L1H_T - H_T)}{H_T}$  distributions.

## Appendix C.

2405 Additional material on background  
2406 estimation methods

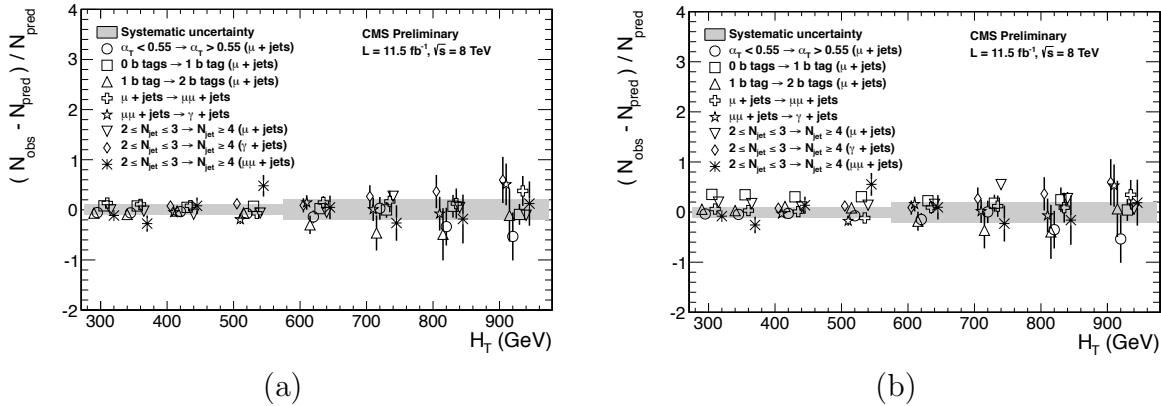
2407 C.1. Determination of  $k_{QCD}$



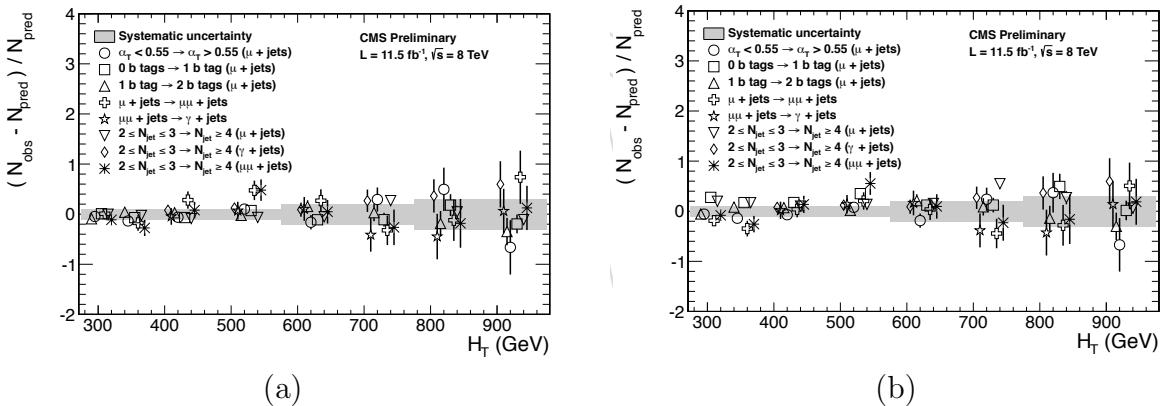
**Figure C.1.:**  $R_{\alpha_T}(H_T)$  and exponential fits for each of the data sideband regions. Fit is conducted between the  $H_T$  region  $275 < H_T < 575$ .

2409 **C.2. Effect of varying background cross sections on**  
2410 **closure tests**

2411 Closure tests with cross section variations of +20% and -20% applied to  $W + \text{jets}$  and  $t\bar{t}$   
2412 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) and for the two different jet multiplicity bins: (a)  $2 \leq n_{jet} \leq 3$  and (b)  $n_{jet} \geq 4$ .



**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) and for the two different jet multiplicity bins: (a)  $2 \leq n_{jet} \leq 3$  and (b)  $n_{jet} \geq 4$ .

		$H_T$ (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	$0.114 \pm 0.009$	$0.066 \pm 0.007$	$0.070 \pm 0.016$	$0.093 \pm 0.020$

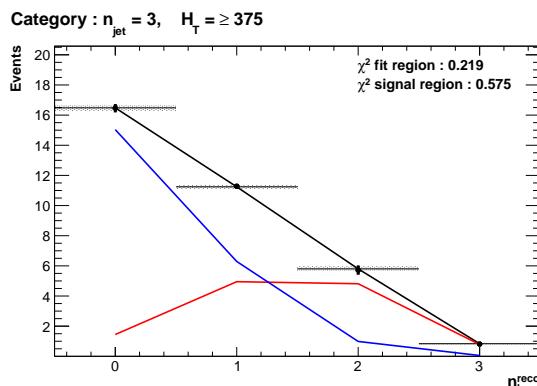
**Table C.1.:** Translation factors constructed from the  $\mu +$  jets control sample and signal selection MC, to predict yields for the  $W +$  jets and  $t\bar{t}$  back-grounds in the signal region with (a) NNLO cross sections corrected by k-factors determined from a data sideband see Section (4.4), 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.

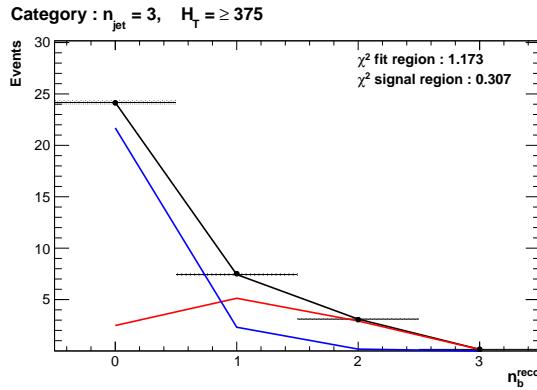
### <sup>2413</sup> Additional Material For B-tag <sup>2414</sup> Template Method

#### <sup>2415</sup> D.1. Templates Fits in Simulation

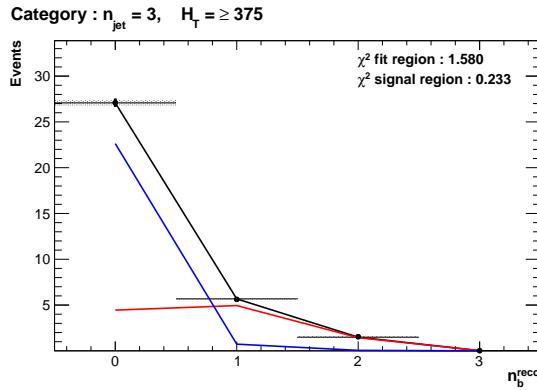
<sup>2416</sup> Template fits for the three **CSV** working points in the  $n_{jet} = 3, H_T > 375$  category :



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



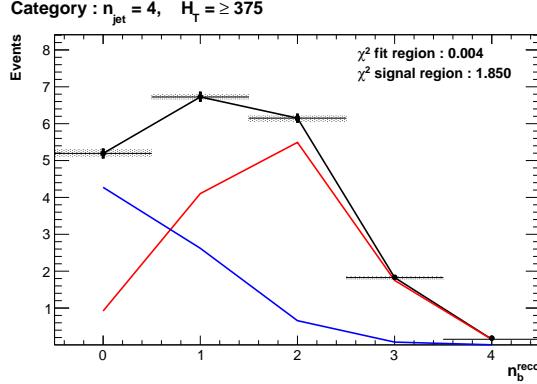
(b) Medium working point  $n_{jet} = 3$



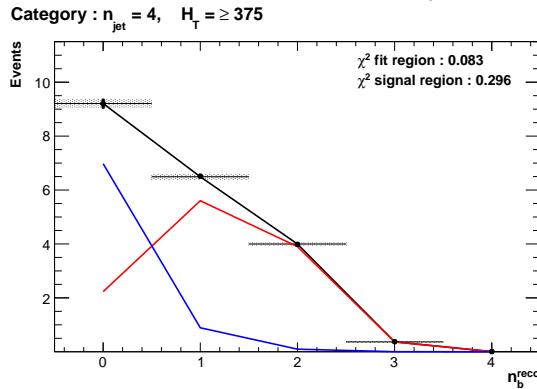
(c) Tight working point  $n_{jet} = 3$

**Figure D.1.:** The results of fitting the  $Z = 0$  and  $Z = 2$  templates to the  $n_b^{reco} = 0, 1, 2$  bins taken directly from simulation in the region  $H_T > 375$  GeV, for the  $n_{jet} = 3$  category. The blue template represents  $Z = 0$ , while the red template represents  $Z = 2$ . Grey bands represent the statistical uncertainty of the fit. The  $\chi^2$  parameter displayed represents the goodness of fit to the low  $n_b^{reco}$  (0-2) control region.

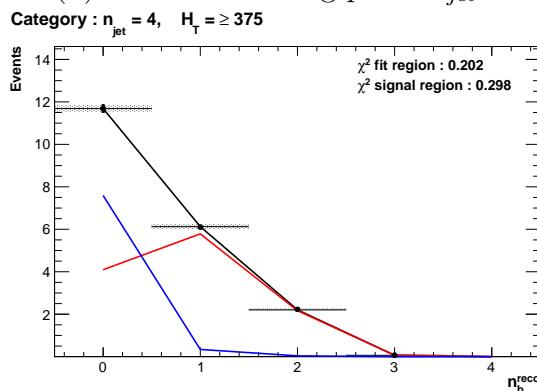
2417 Template fits for the three **CSV** working points in the  $n_{jet} = 4$ ,  $H_T > 375$  category :



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



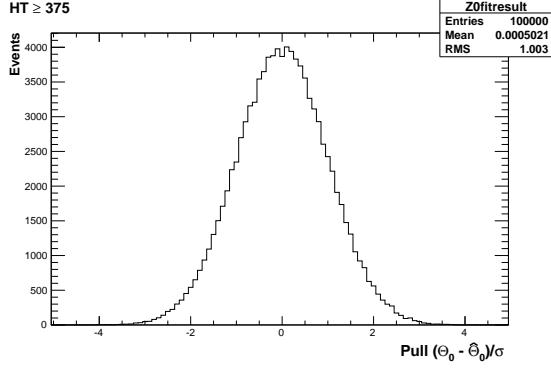
(b) Medium working point  $n_{jet} = 4$



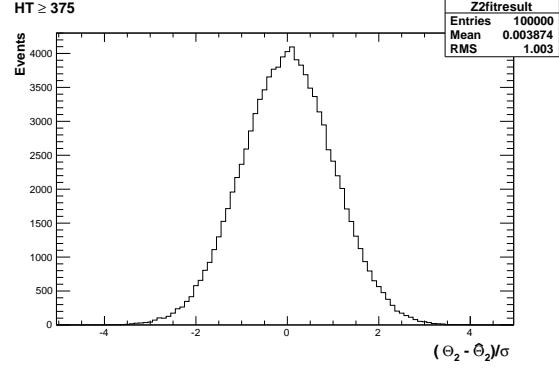
(c) Tight working point  $n_{jet} = 4$

**Figure D.2.:** The results of fitting the  $Z = 0$  and  $Z = 2$  templates to the  $n_b^{reco} = 0, 1, 2$  bins taken directly from simulation in the region  $H_T > 375$  GeV, for the  $n_{jet} = 4$  category. The blue template represents  $Z = 0$ , while the red template represents  $Z = 2$ . Grey bands represent the statistical uncertainty of the fit. The  $\chi^2$  parameter displayed represents the goodness of fit to the low  $n_b^{reco}$  (0-2) control region.

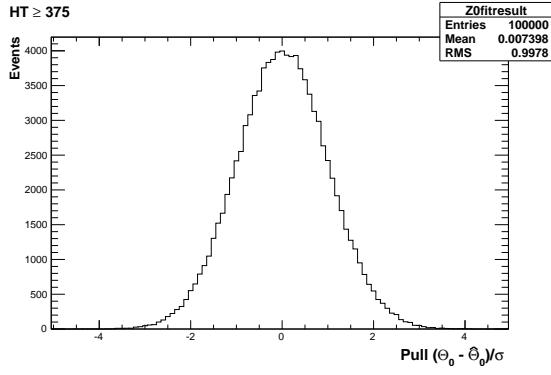
2418 D.2. Pull Distributions for Template Fits



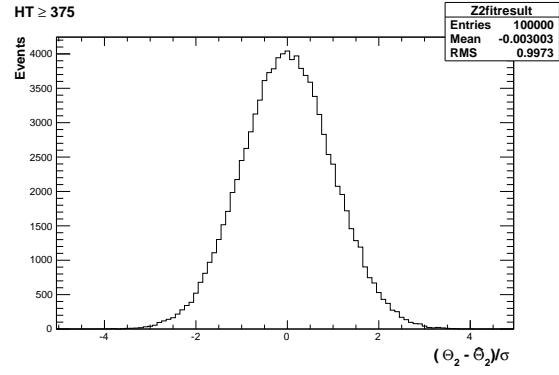
(a) Z0 Template,  $H_T > 375$ ,  $n_{jet} = 3$



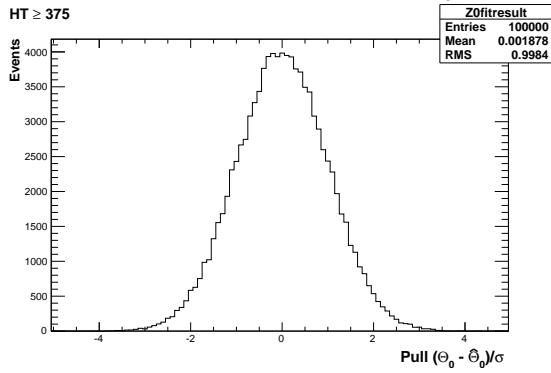
(b) Z2 Template,  $H_T > 375$ ,  $n_{jet} = 3$



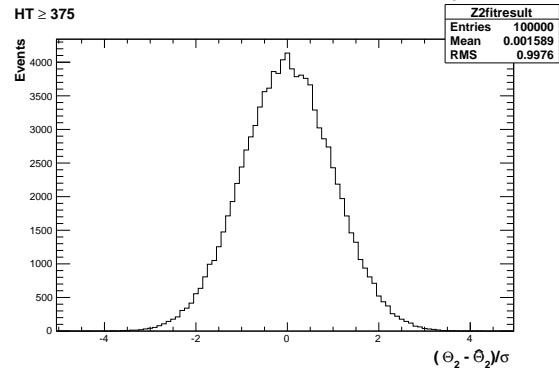
(a) Z0 Template,  $H_T > 375$ ,  $n_{jet} = 4$



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



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

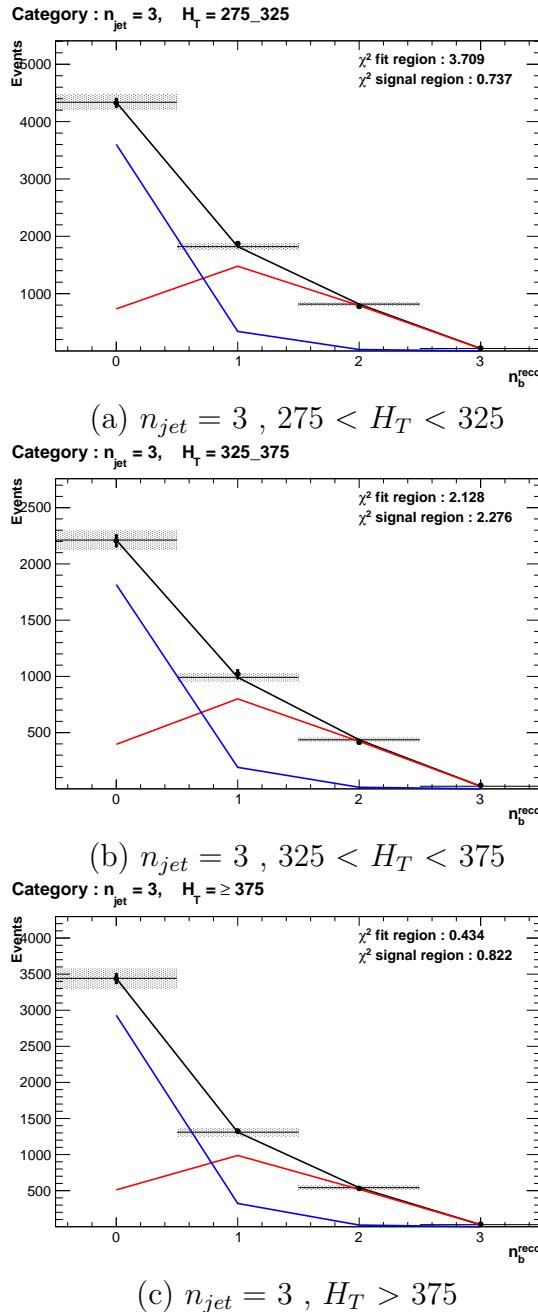


(b) Z2 Template,  $H_T > 375$ ,  $n_{jet} \geq 5$

**Figure D.3.:** Pull distributions of  $\frac{(\theta - \hat{\theta})}{\sigma}$  for  $10^4$  pseudo-experiments generated from a gaussian distribution centred on the  $n_b^{reco}$  template values from simulation with width  $\sigma$ . Distributions are shown for both Z0 and Z2 templates for the medium CSV working point.

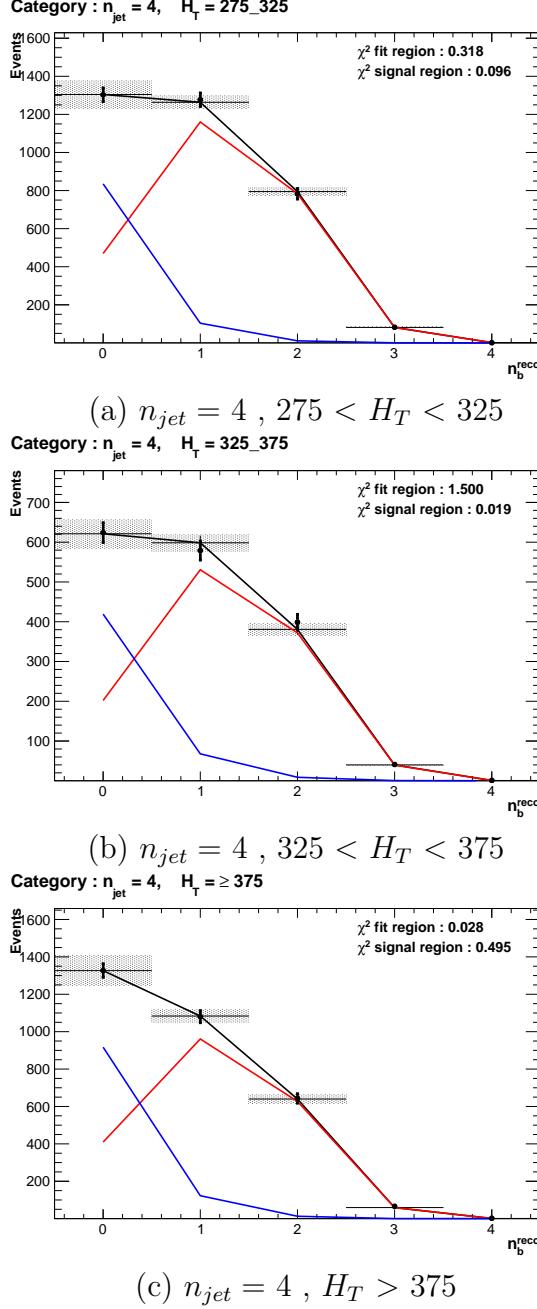
<sup>2419</sup> **D.3. Templates Fits in Data Control Sample**

<sup>2420</sup> Template fits for the three  $H_T$  bins, in the  $n_{jet} = 3$ , medium **CSV** working point:



**Figure D.4.:** The results of fitting the  $Z = 0$  and  $Z = 2$  templates to the  $n_b^{reco} = 0, 1, 2$  bins taken from data, for the  $n_{jet} = 3$  category and medium **CSV** working point. The blue template represents  $Z = 0$ , while the red template represents  $Z = 2$ . The  $\chi^2$  parameter displayed represents the goodness of fit to the low  $n_b^{reco}$  (0-2) control region.

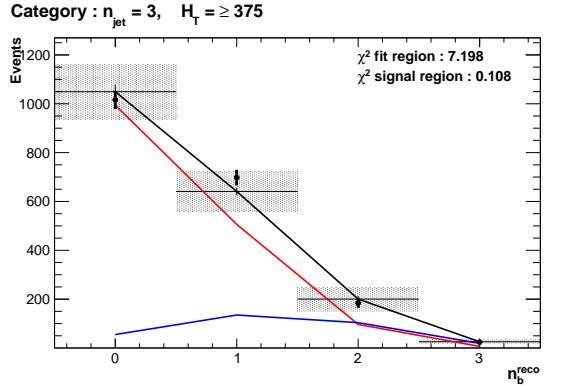
<sup>2421</sup> Template fits for the three  $H_T$  bins, in the  $n_{jet} = 4$ , medium CSV working point:



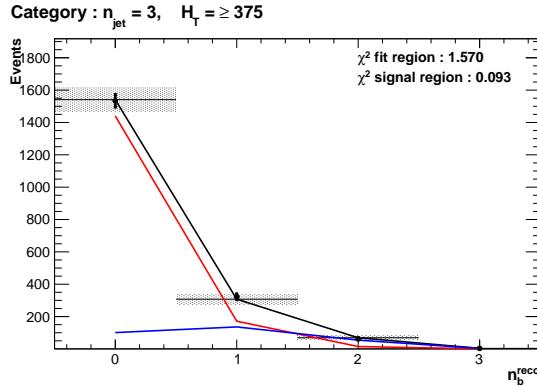
**Figure D.5.:** The results of fitting the  $Z = 0$  and  $Z = 2$  templates to the  $n_b^{reco} = 0, 1, 2$  bins taken from data, for the  $n_{jet} = 4$  category and medium CSV working point. The blue template represents  $Z = 0$ , while the red template represents  $Z = 2$ . The  $\chi^2$  parameter displayed represents the goodness of fit to the low  $n_b^{reco}$  (0-2) control region.

<sup>2422</sup> **D.4. Templates Fits in Data Signal Region**

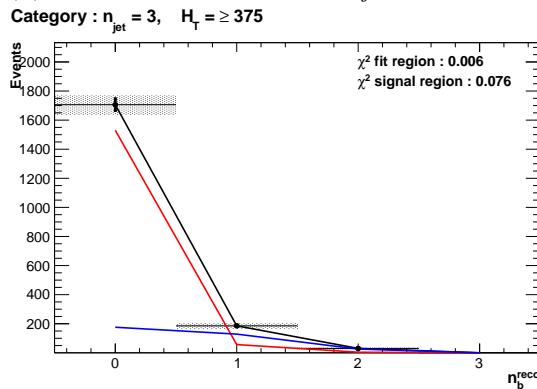
<sup>2423</sup> Template fits for the three **CSV** working points, in the  $n_{jet} = 3, H_T > 375$  category :



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



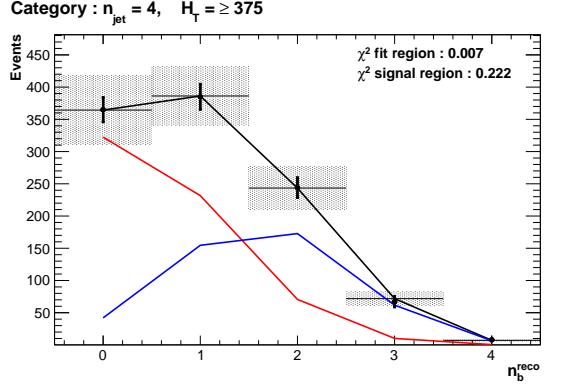
(b) Medium working point :  $n_{jet} = 3, H_T > 375$



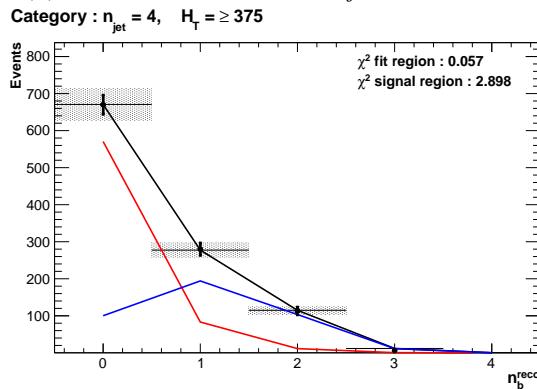
(c) Tight working point :  $n_{jet} = 3, H_T > 375$

**Figure D.6.:** The results of fitting the  $Z = 0$  and  $Z = 2$  templates to the  $n_b^{reco} = 0, 1, 2$  bins taken from data, in the  $n_{jet} = 3$  and  $H_T > 375$  category for all **CSV** working points. The blue template represents  $Z = 0$ , while the red template represents  $Z = 2$ . The  $\chi^2$  parameter displayed represents the goodness of fit to the low  $n_b^{reco}$  (0-2) control region.

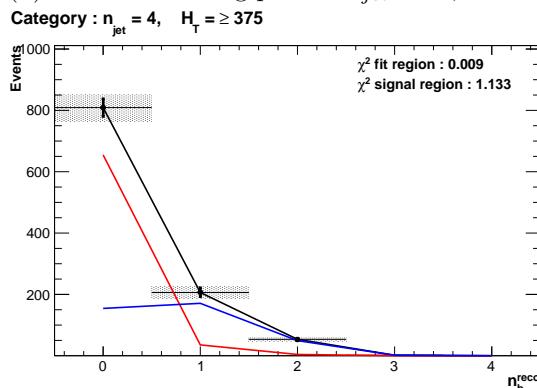
<sup>2424</sup> Template fits for the three **CSV** working points, in the  $n_{jet} = 4$ ,  $H_T > 375$  category :



(a) Loose working point :  $n_{jet} = 4$  ,  $H_T > 375$



(b) Medium working point :  $n_{jet} = 4$  ,  $H_T > 375$



(c) Tight working point :  $n_{jet} = 4$  ,  $H_T > 375$

**Figure D.7.:** The results of fitting the  $Z = 0$  and  $Z = 2$  templates to the  $n_b^{reco} = 0, 1, 2$  bins taken from data, in the  $n_{jet} = 4$  and  $H_T > 375$  category for all **CSV** working points. The blue template represents  $Z = 0$ , while the red template represents  $Z = 2$ . The  $\chi^2$  parameter displayed represents the goodness of fit to the low  $n_b^{reco}$  ( $0-2$ ) control region.



# <sup>2426</sup> Bibliography

- <sup>2427</sup> [1] Particle Data Group Collaboration, “Review of Particle Physics (RPP)”, *Phys.Rev.*  
<sup>2428</sup> **D86** (2012) 010001, [doi:10.1103/PhysRevD.86.010001](https://doi.org/10.1103/PhysRevD.86.010001).
- <sup>2429</sup> [2] G. H. et al., “Nine-year Wilkinson Microwave Anisotropy Probe (WMAP)  
<sup>2430</sup> Observations: Cosmological Parameter Results”, *The Astrophysical Journal*  
<sup>2431</sup> *Supplement Series* **208** (2013), no. 2,,
- <sup>2432</sup> [3] ATLAS Collaboration Collaboration, “Observation of a new particle in the search  
<sup>2433</sup> for the Standard Model Higgs boson with the ATLAS detector at the LHC”,  
<sup>2434</sup> *Phys.Lett.* **B716** (2012) 1–29, [doi:10.1016/j.physletb.2012.08.020](https://doi.org/10.1016/j.physletb.2012.08.020),  
<sup>2435</sup> [arXiv:1207.7214](https://arxiv.org/abs/1207.7214).
- <sup>2436</sup> [4] CMS Collaboration, “Observation of a new boson at a mass of 125 GeV with the  
<sup>2437</sup> CMS experiment at the LHC”, *Phys.Lett.* **B716** (2012) 30–61,  
<sup>2438</sup> [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).
- <sup>2439</sup> [5] CMS Collaboration, “Search for supersymmetry in hadronic final states with  
<sup>2440</sup> missing transverse energy using the variables AlphaT and b-quark multiplicity in  
<sup>2441</sup> pp collisions at 8 TeV”, *Eur.Phys.J.* **C73** (2013) 2568,  
<sup>2442</sup> [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).
- <sup>2443</sup> [6] S. Weinberg, “A Model of Leptons”, *Phys. Rev. Lett.* **19** (Nov, 1967)  
<sup>2444</sup> [doi:10.1103/PhysRevLett.19.1264](https://doi.org/10.1103/PhysRevLett.19.1264).
- <sup>2445</sup> [7] S. Glashow, “Partial Symmetries of Weak Interactions”, *Nucl.Phys.* **22** (1961)  
<sup>2446</sup> [doi:10.1016/0029-5582\(61\)90469-2](https://doi.org/10.1016/0029-5582(61)90469-2).
- <sup>2447</sup> [8] A. Salam, “Weak and Electromagnetic Interactions”, *Conf.Proc.* **C680519** (1968).
- <sup>2448</sup> [9] G. Hooft, “Renormalizable Lagrangians for massive Yang-Mills fields”, *Nuclear*  
<sup>2449</sup> *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).
- <sup>2450</sup> [10] Gargamelle Neutrino Collaboration Collaboration, “Observation of Neutrino Like

- 2451        Interactions Without Muon Or Electron in the Gargamelle Neutrino Experiment”,  
2452        *Phys.Lett.* **B46** (1973) 138–140, [doi:10.1016/0370-2693\(73\)90499-1](https://doi.org/10.1016/0370-2693(73)90499-1).
- 2453 [11] UA1 Collaboration Collaboration, “Experimental Observation of Lepton Pairs of  
2454        Invariant Mass Around 95-GeV at the CERN SPS Collider”, *Phys.Lett.* **B126**  
2455        (1983) 398–410, [doi:10.1016/0370-2693\(83\)90188-0](https://doi.org/10.1016/0370-2693(83)90188-0).
- 2456 [12] UA2 Collaboration Collaboration, “Observation of Single Isolated Electrons of High  
2457        Transverse Momentum in Events with Missing Transverse Energy at the CERN  $\bar{p}p$   
2458        Collider”, *Phys.Lett.* **B122** (1983) [doi:10.1016/0370-2693\(83\)91605-2](https://doi.org/10.1016/0370-2693(83)91605-2).
- 2459 [13] E. Noether, “Invariante Variationsprobleme”, *Nachrichten von der Gesellschaft der*  
2460        *Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* **1918** (1918).
- 2461 [14] F. Halzen and A. D. Martin, “Quarks and Leptons”. 1985.
- 2462 [15] “Introduction to Elementary Particles”. Wiley-VCH, 2nd edition, October, 2008.
- 2463 [16] C. S. Wu et al., “Experimental Test of Parity Conservation in Beta Decay”,  
2464        *Physical Review* **105** (February, 1957) [doi:10.1103/PhysRev.105.1413](https://doi.org/10.1103/PhysRev.105.1413).
- 2465 [17] P. Higgs, “Broken symmetries, massless particles and gauge fields”, *Physics Letters*  
2466        **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).
- 2467 [18] F. Englert and R. Brout, “Broken Symmetry and the Mass of Gauge Vector  
2468        Mesons”, *Phys. Rev. Lett.* **13** (Aug, 1964) [doi:10.1103/PhysRevLett.13.321](https://doi.org/10.1103/PhysRevLett.13.321).
- 2469 [19] P. W. Higgs, “Broken Symmetries and the Masses of Gauge Bosons”, *Phys. Rev.*  
2470        *Lett.* **13** (Oct, 1964) [doi:10.1103/PhysRevLett.13.508](https://doi.org/10.1103/PhysRevLett.13.508).
- 2471 [20] G. S. Guralnik, “Global Conservation Laws and Massless Particles”, *Phys. Rev.*  
2472        *Lett.* **13** (Nov, 1964) [doi:10.1103/PhysRevLett.13.585](https://doi.org/10.1103/PhysRevLett.13.585).
- 2473 [21] S. Weinberg, “A Model of Leptons”, *Phys. Rev. Lett.* **19** (Nov, 1967) 1264–1266,  
2474        [doi:10.1103/PhysRevLett.19.1264](https://doi.org/10.1103/PhysRevLett.19.1264).
- 2475 [22] H. Yukawa, “On the Interaction of Elementary Particles. I”, *Progress of*  
2476        *Theoretical Physics Supplement* **1** (1955) [doi:10.1143/PTPS.1.1](https://doi.org/10.1143/PTPS.1.1).
- 2477 [23] (Super-Kamiokande Collaboration) Collaboration, “Evidence for Oscillation of  
2478        Atmospheric Neutrinos”, *Phys. Rev. Lett.* **81** (Aug, 1998)  
2479        [doi:10.1103/PhysRevLett.81.1562](https://doi.org/10.1103/PhysRevLett.81.1562).

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

- 2509        minimal supersymmetry”, *Phys.Rev.* **D49** (1994)  
2510        [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).
- 2511 [39] C. Stuge et al., “Updated global fits of the cMSSM including the latest LHC SUSY  
2512        and Higgs searches and XENON100 data”, *JCAP* **1203** (2012)  
2513        [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).
- 2514 [40] M. Citron et al., “The End of the CMSSM Coannihilation Strip is Nigh”,  
2515        *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).
- 2516 [41] D. Ghosh, M. Guchait, S. Raychaudhuri, and D. Sengupta, “How Constrained is  
2517        the cMSSM?”, *Phys.Rev.* **D86** (2012) [doi:10.1103/PhysRevD.86.055007](https://doi.org/10.1103/PhysRevD.86.055007),  
2518        [arXiv:1205.2283](https://arxiv.org/abs/1205.2283).
- 2519 [42] LHC New Physics Working Group Collaboration, “Simplified Models for LHC New  
2520        Physics Searches”, *J.Phys.* **G39** (2012) 105005,  
2521        [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).
- 2522 [43] J. Alwall, P. Schuster, and N. Toro, “Simplified Models for a First Characterization  
2523        of New Physics at the LHC”, *Phys.Rev.* **D79** (2009) 075020,  
2524        [doi:10.1103/PhysRevD.79.075020](https://doi.org/10.1103/PhysRevD.79.075020), [arXiv:0810.3921](https://arxiv.org/abs/0810.3921).
- 2525 [44] CMS Collaboration, “Interpretation of Searches for Supersymmetry with simplified  
2526        Models”, *Phys.Rev.* **D88** (2013) 052017, [doi:10.1103/PhysRevD.88.052017](https://doi.org/10.1103/PhysRevD.88.052017),  
2527        [arXiv:1301.2175](https://arxiv.org/abs/1301.2175).
- 2528 [45] J. Hisano, K. Kurosawa, and Y. Nomura, “Natural effective supersymmetry”,  
2529        *Nucl.Phys.* **B584** (2000) 3–45, [doi:10.1016/S0550-3213\(00\)00343-6](https://doi.org/10.1016/S0550-3213(00)00343-6),  
2530        [arXiv:hep-ph/0002286](https://arxiv.org/abs/hep-ph/0002286).
- 2531 [46] M. Papucci, J. T. Ruderman, and A. Weiler, “Natural SUSY Endures”, *JHEP*  
2532        **1209** (2012) 035, [doi:10.1007/JHEP09\(2012\)035](https://doi.org/10.1007/JHEP09(2012)035), [arXiv:1110.6926](https://arxiv.org/abs/1110.6926).
- 2533 [47] B. Allanach and B. Gripaios, “Hide and Seek With Natural Supersymmetry at the  
2534        LHC”, *JHEP* **1205** (2012) 062, [doi:10.1007/JHEP05\(2012\)062](https://doi.org/10.1007/JHEP05(2012)062),  
2535        [arXiv:1202.6616](https://arxiv.org/abs/1202.6616).
- 2536 [48] ALICE Collaboration, “The ALICE experiment at the CERN LHC”, *JINST* **3**  
2537        (2008) S08002, [doi:10.1088/1748-0221/3/08/S08002](https://doi.org/10.1088/1748-0221/3/08/S08002).
- 2538 [49] ATLAS Collaboration, “The ATLAS Experiment at the CERN Large Hadron  
2539        Collider”, *JINST* **3** (2008) [doi:10.1088/1748-0221/3/08/S08003](https://doi.org/10.1088/1748-0221/3/08/S08003).

- 2540 [50] CMS Collaboration, “The CMS experiment at the CERN LHC”, *JINST* **0803**  
2541 (2008) S08004,  
2542 [doi:10.1088/1748-0221/3/08/S08004](https://doi.org/10.1088/1748-0221/3/08/S08004).
- 2543 [51] LHCb Collaboration, “The LHCb Detector at the LHC”, *JINST* **3** (2008) S08005,  
2544 [doi:10.1088/1748-0221/3/08/S08005](https://doi.org/10.1088/1748-0221/3/08/S08005).
- 2545 [52] J.-L. Caron, “LHC Layout. Schema general du LHC.”, (Sep, 1997).
- 2546 [53] CMS Collaboration, “CMS Luminosity - Public Results”, , (2011).  
2547 <http://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults>.
- 2548 [54] CERN, “CMS Compact Muon Solenoid.”, (Feb, 2010).  
2549 <http://public.web.cern.ch/public/Objects/LHC/CMSnc.jpg>.
- 2550 [55] “The CMS Electromagnetic Calorimeter Project: Technical Design Report”.  
2551 Technical Design Report CMS. CERN, Geneva, 1997.
- 2552 [56] “The CMS Muon Project: Technical Design Report”. Technical Design Report  
2553 CMS. CERN, Geneva, 1997.
- 2554 [57] CMS Collaboration, “The CMS Physics Technical Design Report, Volume 1”,  
2555 *CERN/LHCC 2006-001* (2006).
- 2556 [58] M. Cacciari, G. P. Salam, and G. Soyez, “The anti- $k_t$  jet clustering algorithm”,  
2557 *Journal of High Energy Physics* **2008** (2008), no. 04, 063.
- 2558 [59] “Jet Performance in pp Collisions at 7 TeV”, CMS-PAS-JME-10-003, CERN,  
2559 Geneva, (2010).
- 2560 [60] X. Janssen, “Underlying event and jet reconstruction in CMS”, CMS-CR-2011-012,  
2561 CERN, Geneva, (Jan, 2011).
- 2562 [61] CMS Collaboration, “Determination of jet energy calibration and transverse  
2563 momentum resolution in CMS”, *Journal of Instrumentation* **6** (2011), no. 11.,
- 2564 [62] R. Eusebi, “Jet energy corrections and uncertainties in CMS: reducing their impact  
2565 on physics measurements”, *Journal of Physics: Conference Series* **404** (2012).
- 2566 [63] CMS Collaboration, “Algorithms for b Jet identification in CMS”,  
2567 CMS-PAS-BTV-09-001, CERN, 2009. Geneva, (Jul, 2009).
- 2568 [64] CMS Collaboration, “Performance of b-tagging at  $\sqrt{s} = 8$  TeV in multijet,  $t\bar{t}$  and

- 2569 boosted topology events”, CMS-PAS-BTV-13-001, CERN, Geneva, (2013).
- 2570 [65] CMS Collaboration, “Identification of b-quark jets with the CMS experiment”,  
2571 *Journal of Instrumentation* **8** (2013), no. 04.,
- 2572 [66] CMS Collaboration, “CMS. The TriDAS Project. Technical design report, Vol. 1:  
2573 The trigger systems”,.
- 2574 [67] CMS Collaboration, “CMS: The TriDAS Project. Technical design report, Vol. 2:  
2575 Data acquisition and high-level trigger”,.
- 2576 [68] CMS Collaboration, “Calibration and Performance of the Jets and Energy Sums in  
2577 the Level-1 Trigger”, CMS IN 2013/006 (2013), CERN, Geneva, (2013).
- 2578 [69] CMS Collaboration, “Study of Level-1 Trigger Jet Performance in High Pile-up  
2579 Running Conditions”,.
- 2580 [70] J. J. Brooke, “Performance of the CMS Level-1 Trigger”, CMS-CR-2012-322,  
2581 CERN, Geneva, (Nov, 2012).
- 2582 [71] CMS Collaboration, “Search for supersymmetry in final states with missing  
2583 transverse energy and 0, 1, 2, or at least 3 b-quark jets in 7 TeV pp collisions using  
2584 the variable alphaT”, *JHEP* **1301** (2013) 077, [doi:10.1007/JHEP01\(2013\)077](https://doi.org/10.1007/JHEP01(2013)077),  
2585 [arXiv:1210.8115](https://arxiv.org/abs/1210.8115).
- 2586 [72] CMS Collaboration, “SUSY searches with dijet events”, CMS-PAS-SUS-08-005,  
2587 (2008).
- 2588 [73] L. Randall and D. Tucker-Smith, “Dijet Searches for Supersymmetry at the Large  
2589 Hadron Collider”, *Phys. Rev. Lett.* **101** (Nov, 2008) 221803,  
2590 [doi:10.1103/PhysRevLett.101.221803](https://doi.org/10.1103/PhysRevLett.101.221803).
- 2591 [74] CMS Collaboration, “Search strategy for exclusive multi-jet events from  
2592 supersymmetry at CMS”, CMS-PAS-SUS-09-001, CERN, 2009. Geneva, (Jul, 2009).
- 2593 [75] CMS Collaboration, “Calorimeter Jet Quality Criteria for the First CMS Collision  
2594 Data”, CMS-PAS-JME-09-008, CERN, 2010. Geneva, (Apr, 2010).
- 2595 [76] The CMS Collaboration, “Performance of CMS muon reconstruction in pp collision  
2596 events at 7 TeV”, *Journal of Instrumentation* **7** (October, 2012) 2P,  
2597 [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).
- 2598 [77] CMS Collaboration, “Search for supersymmetry in events with photons and missing

- 2599 energy”, CMS-PAS-SUS-12-018, CERN, Geneva, (2012).
- 2600 [78] M. Cacciari and G. P. Salam, “Pileup subtraction using jet areas”, *Phys.Lett.*  
2601 **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).
- 2602 [79] Z. Bern et al., “Driving missing data at next-to-leading order”, *Phys. Rev. D* **84**  
2603 (Dec, 2011) 114002, [doi:10.1103/PhysRevD.84.114002](https://doi.org/10.1103/PhysRevD.84.114002).
- 2604 [80] Z. Bern et al., “Driving Missing Data at Next-to-Leading Order”, *Phys.Rev.* **D84**  
2605 (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).
- 2606 [81] CMS Collaboration Collaboration, “Data-Driven Estimation of the Invisible Z  
2607 Background to the SUSY MET Plus Jets Search”, CMS-PAS-SUS-08-002, CERN,  
2608 2009. Geneva, (Jan, 2009).
- 2609 [82] T. Sjostrand, S. Mrenna, and P. Z. Skands, “PYTHIA 6.4 Physics and Manual”,  
2610 *JHEP* **0605** (2006) 026, [doi:10.1088/1126-6708/2006/05/026](https://doi.org/10.1088/1126-6708/2006/05/026),  
2611 [arXiv:hep-ph/0603175](https://arxiv.org/abs/hep-ph/0603175).
- 2612 [83] W. Beenakker, R. Hopker, M. Spira, and P. Zerwas, “Squark and gluino production  
2613 at hadron colliders”, *Nucl.Phys.* **B492** (1997) 51–103,  
2614 [doi:10.1016/S0550-3213\(97\)80027-2](https://doi.org/10.1016/S0550-3213(97)80027-2), [arXiv:hep-ph/9610490](https://arxiv.org/abs/hep-ph/9610490).
- 2615 [84] S. Abdullin et al., “The Fast Simulation of the CMS Detector at LHC”, *Journal of*  
2616 *Physics: Conference Series* **331** (2011), no. 3,.
- 2617 [85] S. Banerjee, M. D. Hildreth, and the CMS Collaboration, “Validation and Tuning  
2618 of the CMS Full Simulation”, *Journal of Physics: Conference Series* **331** (2011).
- 2619 [86] CMS Collaboration, “CMS Btag POG : CMS b-tagging performance database”, ,  
2620 (2013). <https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagPOG>.
- 2621 [87] CMS Collaboration, “CMS Luminosity Based on Pixel Cluster Counting - Summer  
2622 2012 Update”, CMS-PAS-LUM-12-001, CERN, Geneva, (2012).
- 2623 [88] M. Botje et al., “The PDF4LHC Working Group Interim Recommendations”,  
2624 [arXiv:1101.0538](https://arxiv.org/abs/1101.0538).
- 2625 [89] E. M. Laird, “A Search for Squarks and Gluinos with the CMS Detector”,.
- 2626 [90] R. Cousins, “Probability Density Functions for Positive Nuisance Parameters”.  
2627 2012. <http://www.physics.ucla.edu/cousins/stats/cousinslognormalprior.pdf>.

- 2628 [91] L. Moneta, K. Cranmer, G. Schott, and W. Verkerke, “The RooStats project”,  
2629 2010. [arXiv:1009.1003](https://arxiv.org/abs/1009.1003).
- 2630 [92] F. James and M. Roos, “Minuit: A System for Function Minimization and Analysis  
2631 of the Parameter Errors and Correlations”, *Comput.Phys.Commun.* **10** (1975)  
2632 343–367, [doi:10.1016/0010-4655\(75\)90039-9](https://doi.org/10.1016/0010-4655(75)90039-9).
- 2633 [93] A. L. Read, “Presentation of search results: the CL s technique”, *Journal of  
2634 Physics G: Nuclear and Particle Physics* **28** (2002).
- 2635 [94] T. Junk, “Confidence level computation for combining searches with small  
2636 statistics”, *Nuclear Instruments and Methods in Physics Research Section A* **434**  
2637 (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\)](http://dx.doi.org/10.1016/S0168-9002(99)00498-2).
- 2638 [95] A. L. Read, “Modified frequentist analysis of search results (the  $CL_s$  method)”,  
2639 CERN-OPEN-2000-205, (2000).
- 2640 [96] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, “Asymptotic formulae for  
2641 likelihood-based tests of new physics”, *The European Physical Journal C* **71**  
2642 (2011) [doi:10.1140/epjc/s10052-011-1554-0](https://doi.org/10.1140/epjc/s10052-011-1554-0).
- 2643 [97] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, “Asymptotic formulae for  
2644 likelihood-based tests of new physics”, *European Physical Journal C* **71** (February,  
2645 2011) [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).
- 2646 [98] CMS Collaboration Collaboration, “Search for supersymmetry in final states with  
2647 missing transverse energy and 0, 1, 2, or at least 3 b-quark jets in 7 TeV pp  
2648 collisions using the variable alphaT”, *JHEP* **1301** (2013) 077,  
2649 [doi:10.1007/JHEP01\(2013\)077](https://doi.org/10.1007/JHEP01(2013)077), [arXiv:1210.8115](https://arxiv.org/abs/1210.8115).
- 2650 [99] E. Hardy, “Is Natural SUSY Natural?”, *JHEP* **1310** (2013) 133,  
2651 [doi:10.1007/JHEP10\(2013\)133](https://doi.org/10.1007/JHEP10(2013)133), [arXiv:1306.1534](https://arxiv.org/abs/1306.1534).

**2652 Acronyms**

2653	<b>ALICE</b>	A Large Ion Collider Experiment
2654	<b>ATLAS</b>	A Toroidal LHC ApparatuS
2655	<b>APD</b>	Avalanche Photo-Diodes
2656	<b>BSM</b>	Beyond Standard Model
2657	<b>CERN</b>	European Organization for Nuclear Research
2658	<b>CMS</b>	Compact Muon Solenoid
2659	<b>CMSSM</b>	Compressed Minimal SuperSymmetric Model
2660	<b>CSC</b>	Cathode Stripe Chamber
2661	<b>CSV</b>	Combined Secondary Vertex
2662	<b>CSVM</b>	Combined Secondary Vertex Medium Working Point
2663	<b>DT</b>	Drift Tube
2664	<b>ECAL</b>	Electromagnetic CALorimeter
2665	<b>EB</b>	Electromagnetic CALorimeter Barrel
2666	<b>EE</b>	Electromagnetic CALorimeter Endcap
2667	<b>ES</b>	Electromagnetic CALorimeter pre-Shower
2668	<b>EMG</b>	Exponentially Modified Gaussian
2669	<b>EPJC</b>	European Physical Journal C
2670	<b>EWK</b>	Electroweak Sector
2671	<b>GCT</b>	Global Calorimeter Trigger
2672	<b>GMT</b>	Global MuonTrigger
2673	<b>GT</b>	Global Trigger
2674	<b>HB</b>	Hadron Barrel
2675	<b>HCAL</b>	Hadronic CALorimeter

2676	<b>HE</b>	Hadron Endcaps
2677	<b>HF</b>	Hadron Forward
2678	<b>HLT</b>	Higher Level Trigger
2679	<b>HO</b>	Hadron Outer
2680	<b>HPD</b>	Hybrid Photo Diode
2681	<b>ISR</b>	Initial State Radiation
2682	<b>LUT</b>	Look Up Table
2683	<b>L1</b>	Level 1 Trigger
2684	<b>LEP</b>	Large Electron-Positron Collidior
2685	<b>LHC</b>	Large Hadron Collider
2686	<b>LHCb</b>	Large Hadron Collider Beauty
2687	<b>LO</b>	Leading Order
2688	<b>LSP</b>	Lightest Supersymmetric Partner
2689	<b>NLL</b>	Next to Leading Logorithmic Order
2690	<b>NLO</b>	Next to Leading Order
2691	<b>NNLO</b>	Next to Next Leading Order
2692	<b>POGs</b>	Physics Object Groups
2693	<b>PS</b>	Proton Synchrotron
2694	<b>QED</b>	Quantum Electro-Dynamics
2695	<b>QCD</b>	Quantum Chromo-Dynamics
2696	<b>QFT</b>	Quantum Field Theory
2697	<b>RBXs</b>	Readout Boxes
2698	<b>RPC</b>	Resistive Plate Chamber
2699	<b>RCT</b>	Regional Calorimeter Trigger
2700	<b>RMT</b>	Regional Muon Trigger

2701	<b>SUSY</b>	SUperSYmmetry
2702	<b>SM</b>	Standard Model
2703	<b>SMS</b>	Simplified Model Spectra
2704	<b>SPS</b>	Super Proton Synchrotron
2705	<b>TF</b>	Transfer Factor
2706	<b>TP</b>	Trigger Primative
2707	<b>VEV</b>	Vacuum Expectation Value
2708	<b>VPT</b>	Vacuum Photo-Triodes
2709	<b>WIMP</b>	Weakly Interacting Massive Particle