

<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 with in events with a missing energy signature and high  $p_T$  jets, 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 kinematic 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 b flavoured quarks). A complementary method to search for natural SUSY signatures with a high number of b-flavoured jets, 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 demonstrate proof of principle of 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 pileup jets, whilst maintaining similar performance in the triggering of physics events.

39

## Declaration

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

47

Darren Burton

48

## Acknowledgements

49 I would like to thank the many people whom I have had the pleasure of working with  
50 during the course of the last three and a half years. The opportunity to work as part one  
51 of the largest scientific collaborations during one of the most exciting times in particle  
52 physics for decades, has been a real privilege to be a part of. I could not have achieved  
53 the results presented in this thesis without the help of my colleagues who were part of the  
54 RA1 team, Edward Laird, Chris Lucas, Henning Flaecher, Yossof Eshaq, Bryn Mathais,  
55 Sam Rogerson, Zhaoxia Meng and Georgia Karapostoli whom I worked with on L1 jets. I  
56 also thank my supervisor Oliver Buchmuller for his guidance in getting me to this point.

57 I also feel it important to single out thanks to the postdocs that I have worked with during  
58 my PhD. Jad Marrouche from whom I have learnt a great deal and Robert Bainbridge  
59 who has been like a second supervisor to me, helping me during my time at Imperial and  
60 CERN, especially during those most stressful of times approaching conference deadlines!

61 My fellow PhD students who I live with and have seen on an almost daily basis for the  
62 last few years, Andrew Gilbert, Patrick Owen, Indrek Sepp, Matthew Kenzie and my  
63 girlfriend Hannah. Thanks for putting up with the whinging, complaining and clopping.

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

# 67 **Contents**

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

93	3.4.2. The L1 trigger jet algorithm . . . . .	35
94	3.4.3. Measuring L1 jet trigger efficiencies . . . . .	37
95	3.4.4. Effects of the L1 jet seed . . . . .	38
96	3.4.5. Robustness of L1 jet performance against pile-up . . . . .	40
97	3.4.6. Summary . . . . .	43
98	<b>4. SUSY Searches In Hadronic Final States</b>	<b>45</b>
99	4.1. An Introduction to the $\alpha_T$ Search . . . . .	46
100	4.1.1. The $\alpha_T$ variable . . . . .	48
101	4.2. Search Strategy . . . . .	50
102	4.2.1. Physics objects . . . . .	53
103	4.2.2. Event selection . . . . .	55
104	4.2.3. Control sample definition and background estimation . . . . .	58
105	4.2.4. Estimating the QCD multi-jet background . . . . .	65
106	4.3. Trigger Strategy . . . . .	67
107	4.4. Measuring MC Normalisation Factors via $H_T$ Sidebands . . . . .	68
108	4.5. Determining MC Simulation Yields with Higher Statistical Precision . . . . .	69
109	4.5.1. The formula method . . . . .	70
110	4.5.2. Establishing proof of principle . . . . .	71
111	4.5.3. Correcting measured efficiencies in simulation to data . . . . .	71
112	4.6. Systematic Uncertainties on Transfer Factors . . . . .	74
113	4.6.1. Determining systematic uncertainties from closure tests . . . . .	78
114	4.7. Simplified Models, Efficiencies and Systematic Uncertainties . . . . .	80
115	4.7.1. Signal efficiency . . . . .	81
116	4.7.2. Applying b-tag scale factor corrections in signal samples . . . . .	82
117	4.7.3. Experimental uncertainties . . . . .	83
118	4.8. Statistical Framework . . . . .	86
119	4.8.1. Hadronic sample . . . . .	86
120	4.8.2. $H_T$ evolution model . . . . .	87
121	4.8.3. Electroweak Sector (EWK) control samples . . . . .	87
122	4.8.4. Contributions from signal . . . . .	90
123	4.8.5. Total likelihood . . . . .	91
124	<b>5. Results and Interpretation</b>	<b>92</b>
125	5.1. Standard Model . . . . .	92
126	5.2. SUSY . . . . .	101
127	5.2.1. The $CL_s$ method . . . . .	101

128	5.2.2. Interpretation in simplified signal models . . . . .	102
129	<b>6. Searches For Natural SUSY With B-tag Templates.</b>	106
130	6.1. Concept . . . . .	106
131	6.2. Application to the $\alpha_T$ Search . . . . .	108
132	6.2.1. Proof of principle in simulation . . . . .	109
133	6.2.2. Results in a data control sample . . . . .	112
134	6.2.3. Application to the $\alpha_T$ hadronic search region . . . . .	114
135	6.3. Summary . . . . .	116
136	<b>7. Conclusions</b>	117
137	<b>A. Miscellaneous</b>	119
138	A.1. Noise Filters . . . . .	119
139	A.2. Primary Vertices . . . . .	121
140	<b>B. L1 Jets</b>	123
141	B.1. Jet matching efficiencies . . . . .	123
142	B.2. Leading Jet Energy Resolution . . . . .	124
143	B.3. Resolution for Energy Sum Quantities . . . . .	127
144	<b>C. Additional material on background estimation methods</b>	132
145	C.1. Determination of $k_{QCD}$ . . . . .	132
146	C.2. Effect of varying background cross sections on closure tests . . . . .	133
147	<b>D. Additional Material For B-tag Template Method</b>	135
148	D.1. Templates Fits in Simulation . . . . .	135
149	D.2. Pull Distributions for Template Fits . . . . .	138
150	D.3. Templates Fits in Data Control Sample . . . . .	139
151	D.4. Templates Fits in Data Signal Region . . . . .	141
152	<b>Bibliography</b>	144

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

<sup>154</sup> 2.1.	One loop quantum corrections to the Higgs squared mass parameter $m_h^2$ due to a fermion. . . . .	14
<sup>156</sup> 2.2.	Two example simplified model decay chains. . . . .	19
<sup>157</sup> 3.1.	A top down layout of the LHC, with the position of the four main detectors labelled. . . . .	21
<sup>159</sup> 3.2.	The total integrated luminosity delivered to and collected by Compact Muon Solenoid (CMS) during the 2012 8 TeV $pp$ runs . . . . .	22
<sup>161</sup> 3.3.	A pictorial depiction of the CMS detector. . . . .	24
<sup>162</sup> 3.4.	Illustration of the CMS Electromagnetic CALorimeter (ECAL). . . . .	26
<sup>163</sup> 3.5.	Schematic of the CMS Hadronic CALorimeter (HCAL). . . . .	27
<sup>164</sup> 3.6.	Combined Secondary Vertex (CSV) algorithm discriminator values in enriched ttbar and inclusive multi jet samples . . . . .	31
<sup>166</sup> 3.7.	Data/MC b-tag scale factors derived using the Combined Secondary Vertex Medium Working Point (CSVM) tagger. . . . .	32
<sup>168</sup> 3.8.	Data/MC mis-tag scale factors derived using the CSVM tagger. . . . .	33
<sup>169</sup> 3.9.	The CMS Level 1 Trigger (L1) Trigger system. . . . .	34
<sup>170</sup> 3.10.	Illustration of the Level-1 jet finding algorithm. . . . .	36
<sup>171</sup> 3.11.	L1 jet efficiency turn-on curves as a function of the offline CaloJet and PFJet $E_T$ . . . . .	38
<sup>173</sup> 3.12.	L1 jet efficiency turn-on curves as a function of the offline CaloJet $E_T$ for the 2012 run period B and C. . . . .	39

---

175	3.13. Trigger cross section for the L1HTT150 trigger path. . . . .	40
176	3.14. L1 $H_T$ efficiency turn-on curves as a function of the offline CaloJet $H_T$ . . . . .	41
177	3.15. L1 jet efficiency turn-on curves as a function of the leading offline $E_T$ Calo	
178	(left) and PF (right) jet, for low, medium and high pile-up conditions. . . . .	42
179	3.16. Fit values from an Exponentially Modified Gaussian (EMG) function fitted	
180	to the resolution plots of leading Calo jet $E_T$ measured as a function of	
181	$\frac{(L1\ E_T - \text{Offline}\ E_T)}{\text{Offline}\ E_T}$ for low, medium and high pile-up conditions. . . . .	43
182	3.17. Fit values from an EMG function fitted to the resolution plots of leading	
183	PF jet $E_T$ measured as a function of $\frac{(L1\ E_T - \text{Offline}\ E_T)}{\text{Offline}\ E_T}$ for low, medium and	
184	high pile-up conditions. . . . .	44
185	4.1. Reconstructed offline $H_T$ distribution in the hadronic signal selection, from	
186	$11.7\text{fb}^{-1}$ of data, in which no $\alpha_T$ requirement is made. . . . .	47
187	4.2. The event topologies of background QCD dijet events (right) and a generic	
188	SUSY signature with genuine $Z_T$ (left). . . . .	48
189	4.3. The $\alpha_T$ distributions for the low 2-3 (left) and high $\geq 4$ (right) jet	
190	multiplicities after a full analysis selection and shown for $H_T > 375$ . . . . .	50
191	4.4. Pictorial depiction of the analysis strategy employed by the $\alpha_T$ search to	
192	increase sensitivity to a wide spectra of SUSY models. . . . .	52
193	4.5. Data/MC comparisons of key variables for the hadronic signal region. . . . .	58
194	4.6. Data/MC comparisons of key variables for the $\mu +$ jets selection. . . . .	61
195	4.7. Data/MC comparisons of key variables for the $\mu\mu +$ jets selection. . . . .	63
196	4.8. Data/MC comparisons of key variables for the $\gamma +$ jets selection. . . . .	64
197	4.9. QCD sideband regions, used for determination of $k_{QCD}$ . . . . .	66
198	4.10. Tagging efficiencies of (a) b-jets, (b) c-jets, and (c) light-jets determined	
199	from all jets within each individual analysis $H_T$ bin. . . . .	73
200	4.11. Sets of closure tests overlaid on top of the systematic uncertainty used for	
201	each of the five $H_T$ regions. . . . .	79

---

202	4.12. Signal efficiencies fo the Simplified Model Spectra (SMS) models (a) T1 203 and (b) T2. . . . .	81
204	5.1. Comparison of the observed yields and Standard Model (SM) expectations 205 given by the simultaneous fit in bins of $H_T$ for the (a) hadronic, (b) $\mu +$ 206 jets, (c) $\mu\mu +$ jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 0$ and 207 $n_{jet} \leq 3$ . . . . .	94
208	5.2. Comparison of the observed yields and SM expectations given by the 209 simultaneous fit in bins of $H_T$ for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ 210 jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 1$ and $n_{jet} \leq 3$ . . . . .	95
211	5.3. Comparison of the observed yields and SM expectations given by the 212 simultaneous fit in bins of $H_T$ for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ 213 jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 2$ and $n_{jet} \leq 3$ . . . . .	96
214	5.4. Comparison of the observed yields and SM expectations given by the 215 simultaneous fit in bins of $H_T$ for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ 216 jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 0$ and $n_{jet} \geq 4$ . . . . .	97
217	5.5. Comparison of the observed yields and SM expectations given by the 218 simultaneous fit in bins of $H_T$ for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ 219 jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 1$ and $n_{jet} \geq 4$ . . . . .	98
220	5.6. Comparison of the observed yields and SM expectations given by the 221 simultaneous fit in bins of $H_T$ for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ 222 jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 2$ and $n_{jet} \geq 4$ . . . . .	99
223	5.7. Comparison of the observed yields and SM expectations given by the 224 simultaneous fit in bins of $H_T$ for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ 225 jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 3$ and $n_{jet} \geq 4$ . . . . .	99
226	5.8. Comparison of the observed yields and SM expectations given by the 227 simultaneous fit in bins of $H_T$ for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ 228 jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} \geq 4$ and $n_{jet} \geq 4$ . . . . .	100
229	5.9. Production and decay modes for the various SMS models interpreted 230 within the analysis. . . . .	104
231	5.10. Upper limit of cross section at 95% CL as a function of $m_{\tilde{q}/\tilde{g}}$ and $m_{LSP}$ 232 for various SMS models. . . . .	105

---

233	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 $H_T > 375$ . . . . .	108
236	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. . . . .	111
239	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. . . . .	113
242	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. . . . .	115
245	B.1. Leading jet matching efficiency as a function of the offline CaloJet $E_T$ . . . . .	123
246	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. . . . .	125
248	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. . . . .	127
250	B.4. $\sum E_T$ resolution parameters in bins of Calo $\sum E_T$ measured for the defined low, medium and high pile up conditions. . . . .	128
252	B.5. $\sum E_T$ resolution parameters in bins of PF $\sum E_T$ measured for the defined low, medium and high pile up conditions. . . . .	128
254	B.6. $\cancel{E}_T$ resolution parameters in bins of Calo $\cancel{E}_T$ measured for the defined low, medium and high pile up conditions. . . . .	129
256	B.7. $\cancel{E}_T$ resolution parameters in bins of PF $\cancel{E}_T$ measured for the defined low, medium and high pile up conditions. . . . .	129
258	B.8. $H_T$ resolution parameters in bins of Calo $H_T$ measured for the defined low, medium and high pile up conditions. . . . .	130
260	B.9. $H_T$ resolution parameters in bins of PF $H_T$ measured for the defined low, medium and high pile up conditions. . . . .	130

---

262	B.10. $\mathcal{H}_T$ resolution parameters in bins of $\mathcal{H}_T$ measured for the defined low,	
263	medium and high pile up conditions. . . . .	131
264	B.11. $\mathcal{H}_T$ resolution parameters in bins of PF $\mathcal{H}_T$ measured for the defined low,	
265	medium and high pile up conditions. . . . .	131
266	C.1. $R_{\alpha_T}(H_T)$ and exponential fits for each of the data sideband regions. Fit is	
267	conducted between the $H_T$ region $275 < H_T < 575$ . . . . .	132
268	C.2. Sets of closure tests overlaid on top of the systematic uncertainty used for	
269	each of the five $H_T$ regions. . . . .	133
270	C.3. Sets of closure tests overlaid on top of the systematic uncertainty used for	
271	each of the five $H_T$ regions. . . . .	133
272	D.1. The results of fitting the $Z = 0$ and $Z = 2$ templates to the $n_b^{reco} = 0, 1, 2$	
273	bins taken directly from simulation in the region $H_T > 375$ GeV, for the	
274	$n_{jet} = 3$ category. . . . .	136
275	D.2. The results of fitting the $Z = 0$ and $Z = 2$ templates to the $n_b^{reco} = 0, 1, 2$	
276	bins taken directly from simulation in the region $H_T > 375$ GeV, for the	
277	$n_{jet} = 4$ category. . . . .	137
278	D.3. Pull distributions of $\frac{(\theta - \hat{\theta})}{\sigma}$ for $10^4$ pseudo-experiments generated from a	
279	gaussian distribution centred on the $n_b^{reco}$ template values from simulation	
280	with width $\sigma$ . . . . .	138
281	D.4. The results of fitting the $Z = 0$ and $Z = 2$ templates to the $n_b^{reco} = 0, 1, 2$	
282	bins taken directly from data, for the $n_{jet} = 3$ category and medium CSV	
283	working point. . . . .	139
284	D.5. The results of fitting the $Z = 0$ and $Z = 2$ templates to the $n_b^{reco} = 0, 1, 2$	
285	bins taken directly from data, for the $n_{jet} = 4$ category and medium CSV	
286	working point. . . . .	140
287	D.6. The results of fitting the $Z = 0$ and $Z = 2$ templates to the $n_b^{reco} = 0, 1, 2$	
288	bins taken from data, in the $n_{jet} = 3$ and $H_T > 375$ category for all CSV	
289	working points. . . . .	141

290	D.7. The results of fitting the $Z = 0$ and $Z = 2$ templates to the $n_b^{reco} = 0, 1, 2$	
291	bins taken from data, in the $n_{jet} = 4$ and $H_T > 375$ category for all CSV	
292	working points. . . . .	142

# <sup>293</sup> List of Tables

<sup>294</sup>	2.1. The fundamental particles of the SM, with spin, charge and mass displayed.	6
<sup>295</sup>	3.1. Results of a cumulative EMG function fit to the turn-on curves for L1	
<sup>296</sup>	single jet triggers in 2012 Run Period C. . . . .	38
<sup>297</sup>	3.2. Results of a cumulative EMG function fit to the turn-on curves for L1	
<sup>298</sup>	single jet triggers in the 2012 run period B and C. . . . .	40
<sup>299</sup>	3.3. Results of a cumulative EMG function fit to the turn-on curves for $H_T$ in	
<sup>300</sup>	2012 run period B and C. . . . .	41
<sup>301</sup>	3.4. Results of a cumulative EMG function fit to the efficiency turn-on curves	
<sup>302</sup>	for L1 single jet triggers in the 2012 run period C, for low,medium and	
<sup>303</sup>	high pile-up conditions. . . . .	41
<sup>304</sup>	3.5. Results of a cumulative EMG function fit to the efficiency turn-on curves	
<sup>305</sup>	for Level-1 single jet triggers in the 2012 run period C, for low,medium	
<sup>306</sup>	and high pile-up conditions. . . . .	42
<sup>307</sup>	4.1. A summary of the SMS models interpreted in this analysis, involving both	
<sup>308</sup>	direct (D) and gluino-induced (G) production of squarks and their decays.	46
<sup>309</sup>	4.2. Muon Identification criteria used within the analysis for selection/veto	
<sup>310</sup>	purposes in the muon control/signal selections. . . . .	53
<sup>311</sup>	4.3. Photon Identification criteria used within the analysis for selection/veto	
<sup>312</sup>	purposes in the $\gamma +$ jets control/signal selections. . . . .	54
<sup>313</sup>	4.4. Electron Identification criteria used within the analysis for veto purposes.	55
<sup>314</sup>	4.5. Jet thresholds used in the three $H_T$ regions of the analysis. . . . .	56

315	4.6. Best fit values for the parameters $k_{QCD}$ obtained from sideband regions B,C <sub>1</sub> ,C <sub>2</sub> ,C <sub>3</sub> . . . . .	67
316		
317	4.7. Measured efficiencies of the $H_T$ and $\alpha_T$ legs of the HT and $HT\_alphaT$ triggers in independent analysis bins. . . . .	68
318		
319	4.8. k-factors calculated for different EWK processes. . . . .	69
320		
321	4.9. 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 . . . . .	72
322		
323	4.10. The absolute change in the Transfer Factor (TF)'s used to predict the entire signal region SM background, using the $\mu +$ jets control sample when the systematic uncertainties of the data to simulation scale factors are varied by $\pm 1\sigma$ . . . . .	74
324		
325		
326		
327	4.11. A summary of the results obtained from fits of zeroeth order polynomials (i.e. a constant) to five sets of closure tests performed in the $2 \leq n_{jet} \leq 3$ bin . . . . .	77
328		
329		
330	4.12. A summary of the results obtained from fits of zeroeth order polynomials (i.e. a constant) to five sets of closure tests performed in the $n_{jet} \geq 4$ bin	77
331		
332	4.13. A summary of the results obtained from fits of zeroeth order polynomials (i.e. a constant) to five sets of closure tests performed in the $2 \leq n_{jet} \leq 3$ bin . . . . .	77
333		
334		
335	4.14. Calculated systematic uncertainties for the five $H_T$ regions, determined from the closure tests. . . . .	78
336		
337	4.15. Estimates of systematic uncertainties on the signal efficiency (%) for various SMS models when considering points in the region near to the diagona . . . . .	85
338		
339		
340	4.16. Estimates of systematic uncertainties on the signal efficiency (%) for various SMS models when considering points in the region near to the diagona . . . . .	85
341		
342		
343	4.17. The systematic parameters used in $H_T$ bins. . . . .	89

---

344	4.18. Nuisance parameters used within the different hadronic signal bins of the analysis . . . . .	91
345		
346	5.1. Summary of control samples used by each fit results, and the Figures in which they are displayed. . . . .	93
347		
348	5.2. Comparison of the measured yields in the 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. . . . .	93
349		
350		
351		
352	5.3. A table representing the SMS models interpreted within the analysis. . . . .	102
353		
354	6.1. Typical underlying b-quark content of different SM processes which are common to many SUSY searches. . . . .	107
355		
356	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 $10 \text{ fb}^{-1}$ . . . . .	110
357		
358		
359	6.3. Summary of the fit predictions in the $n_b^{reco}$ signal region of the $\mu + \text{jets}$ control sample, for $n_{jet} = 3, = 4, \geq 5$ . The fit region is $n_b^{reco} = 0, 1, 2$ using $11.5 \text{ fb}^{-1}$ of data at $\sqrt{s} = 8\text{TeV}$ . . . . .	112
360		
361		
362	6.4. Summary of the fit predictions in the $n_b^{reco}$ signal region of the $\mu + \text{jets}$ control sample, for $n_{jet} = 3, = 4, \geq 5$ . The fit region is $n_b^{reco} = 0, 1, 2$ using $11.5 \text{ fb}^{-1}$ of data at $\sqrt{s} = 8\text{TeV}$ . . . . .	114
363		
364	A.1. Criteria for a reconstructed jet to pass the loose calorimeter jet id. . . . .	119
365		
366	A.2. Criteria for a reconstructed jet to pass the loose PF jet id. . . . .	120
367		
368	A.3. Noise filters that are applied to remove spurious and non-physical $\cancel{E}_T$ signatures within the CMS detector. . . . .	121
369		
	A.4. Criteria for a vertex in an event to be classified as a 'good' reconstructed primary vertex. . . . .	122

370	B.1. Results of a cumulative EMG function fit to the turn-on curves for the	
371	matching efficiency of the leading jet in an event to a Level-1 jet in run	
372	2012C and 2012B data. . . . .	124
373	C.1. Translation factors constructed from the $\mu +$ jets control sample and	
374	signal selection MC, to predict yields for the $W +$ jets and $t\bar{t}$ back-grounds	
375	in the signal region. . . . .	134

<sup>376</sup>

*“The Universe is about 1,000,000 years old.”*

— Matthew Kenzie, 1987-present : Discoverer of the Higgs Boson.

<sup>377</sup>

# Chapter 1.

## <sup>378</sup> Introduction

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

<sup>385</sup> Despite it's successes it is not a complete theory, with significant questions remaining  
<sup>386</sup> unanswered. It describes only three of the four known forces with gravity not incorporated  
<sup>387</sup> within the framework of the **SM**. Cosmological experiments infer that just  $\sim 4\%$  of the  
<sup>388</sup> observable universe exists as matter, with elusive "Dark Matter" accounting for a further  
<sup>389</sup>  $\sim 23\%$  [2]. However no particle predicted by the **SM** is able to account for it. At higher  
<sup>390</sup> energy scales and small distances, the (non-)unification of the fundamental forces point  
<sup>391</sup> to problems with the **SM** at least at higher energies not yet probed experimentally.

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

401 The experimental goal of the **LHC** is to further test the framework of the **SM**, exploring the  
402 TeV mass scale for the first time, and to seek a connection between the particles produced  
403 in proton collisions and dark matter. The first new discovery by this extraordinary  
404 machine was announced on the 4th of July 2012. The long-awaited discovery was the  
405 culmination of decades of experimental endeavours in the search for the Higgs boson,  
406 providing an answer to the mechanism of electroweak symmetry breaking within the **SM**  
407 [3][4].

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

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

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

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

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

# Chapter 2.

## <sup>446</sup> A Theoretical Overview

<sup>447</sup> Within this chapter, a brief introduction and background to the **SM** is given. Its success  
<sup>448</sup> as a rigorously tested and widely accepted theory is discussed as are its deficiencies,  
<sup>449</sup> leading to the argument that this theory is not a complete description of our universe.  
<sup>450</sup> The motivations for new physics at the TeV scale and in particular Supersymmetric  
<sup>451</sup> theories are outlined within Section (2.3), with the chapter concluding with how an  
<sup>452</sup> experimental signature of such theories can be produced and observed at the **LHC**,  
<sup>453</sup> Section (2.4).

### <sup>454</sup> 2.1. The Standard Model

<sup>455</sup> The **SM** is the name given to the relativistic Quantum Field Theory (**QFT**), where  
<sup>456</sup> particles are represented as excitations of fields, which describe the interactions and  
<sup>457</sup> properties of all the known elementary particles [6][7][8][9]. It is a renormalisable field  
<sup>458</sup> theory which contains three symmetries:  $SU(3)$  for colour charge,  $SU(2)$  for weak isospin  
<sup>459</sup> and  $U(1)$  relating to weak hyper charge, which require its Lagrangian  $\mathcal{L}_{SM}$  to be invariant  
<sup>460</sup> under local gauge transformation.

<sup>461</sup> Within the **SM** theory, matter is composed of spin  $\frac{1}{2}$  fermions, which interact with each  
<sup>462</sup> other via the exchange of spin-1 gauge bosons. A summary of the known fundamental  
<sup>463</sup> fermions and bosons is given in Table 2.1.

<sup>464</sup> Fermions are separated into quarks and leptons of which only quarks interact with the  
<sup>465</sup> strong nuclear force. Quarks unlike leptons are not seen as free particles in nature, but  
<sup>466</sup> rather exist only within baryons, composed of three quarks with an overall integer charge,  
<sup>467</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].

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

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

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

### <sup>481</sup> 2.1.1. Gauge Symmetries of the SM

- <sup>482</sup> Symmetries are of fundamental importance in the description of physical phenomena.  
<sup>483</sup> Noether's theorem states that for a dynamical system, the consequence of any symmetry  
<sup>484</sup> is an associated conserved quantity [13]. Invariance under translations, rotations, and  
<sup>485</sup> Lorentz transformations in physical systems lead to conservation of momentum, energy  
<sup>486</sup> and angular momentum.
- <sup>487</sup> In the **SM**, a quantum theory described by Lagrangian formalism, the weak, strong and  
<sup>488</sup> electromagnetic interactions are described in terms of “gauge theories”. A gauge theory  
<sup>489</sup> possesses invariance under a set of “local transformations”, which are transformations  
<sup>490</sup> whose parameters are space-time dependent. The requirement of gauge invariance within  
<sup>491</sup> the **SM** necessitates the introduction of force-mediating gauge bosons and interactions  
<sup>492</sup> between fermions and the bosons themselves. Given the nature of the topics covered by  
<sup>493</sup> this thesis, the formulation of **EWK** within the **SM** Lagrangian is reviewed within this  
<sup>494</sup> section.
- <sup>495</sup> The simplest example of the application of the principle of local gauge invariance within  
<sup>496</sup> the **SM** is in Quantum Electro-Dynamics (**QED**), the consequences of which require a  
<sup>497</sup> massless photon field [14][15].
- <sup>498</sup> Starting from the free Dirac Lagrangian written as

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

- <sup>499</sup> where  $\psi$  represents a free non interacting fermionic field, with the matrices  $\gamma^\mu$ ,  $\mu \in 0, 1, 2, 3$   
<sup>500</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>501</sup> flat space-time metric  $(+, -, -, -)$ , and  $I_4$  the  $4 \times 4$  identity matrix.
- <sup>502</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>503</sup> the kinetic term of the Lagrangian will not remain invariant, due to the partial derivative  
<sup>504</sup> interposed between the  $\bar{\psi}$  and  $\psi$  yielding,

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

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

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

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

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

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

This then leads to the Lagrangian of QED:

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

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

519 **2.1.2. The Electroweak Sector and Electroweak Symmetry**  
520 **Breaking**

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

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

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

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

<sup>541</sup> where the left handed fermions form isospin doubles  $\chi_L$  and the right handed fermions  
<sup>542</sup> are isosinglets  $\psi_R$ . For the first generation of leptons and quarks this represents

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

<sup>543</sup> Imposing local gauge invariance within  $\mathcal{L}_{EWK}$  is once again achieved by modifying the  
<sup>544</sup> covariant derivative

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

<sup>545</sup> where  $g$  and  $g'$  are the coupling constant of the  $SU(2)_L$  and  $U(1)_Y$  groups respectively.  
<sup>546</sup> Taking the example of the first generation of fermions defined in Equation.(2.9), with input  
<sup>547</sup> hypercharge values of -1 and -2 for  $\chi_L$  and  $e_R$  respectively, would lead to a Lagrangian  
<sup>548</sup>  $\mathcal{L}_1$  of the form,

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

<sup>549</sup> As in QED, these additional gauge fields introduce field strength tensors  $B_{\mu\nu}$  and  $W_{\mu\nu}$ ,

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

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

<sup>550</sup> corresponding to the kinetic energy and self coupling of the  $W_\mu$  fields and the kinetic  
<sup>551</sup> energy term of the  $B_\mu$  field.

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

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

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

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

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

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

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

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

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

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

568 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  
 569 the **VEV**.

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

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

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

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

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

578 where the relations between the physical and electroweak gauge fields from Equation  
 579 (2.14) are used. The  $W^\pm$  and Z boson masses can then be determined to be

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

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

## 586 2.2. Motivation for Physics Beyond the Standard 587 Model

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

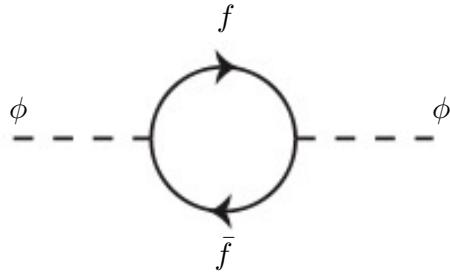
593 On a theoretical level, the **SM** is unable to incorporate the gravitational interactions of  
594 fundamental particles within the theory. Whilst at the electroweak energy scales the  
595 relative strength of gravity is negligible compared to the other three fundamental forces,  
596 at much higher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

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

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

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

613 In the absence of new physics between the TeV and Planck scale, calculating beyond  
614 tree-level contributions to the Higgs mass term given by its self interaction, result in  
615 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.

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

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

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

621 To recover the mass of the now discovered Higgs boson would require a fine-tuning of  
 622 the parameters to cancel out these mass corrections of the Higgs mass to the scale of  
 623 30 orders of magnitude. This appears as an unnatural solution to physicists and it is  
 624 this hierarchy problem that provides one of the strongest motivations for the theory of  
 625 SUperSYmmetry (**SUSY**).

### 626 2.3. Supersymmetry Overview

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

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

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

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

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

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

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

---

### 659 2.3.1. R-Parity

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

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

669 R-parity is described by the equation

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

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

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

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

## 682 2.4. Experimental Signatures of **SUSY** at the **LHC**

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

- 685     • squark/anti-squark and gluino pairs can be produced via both gluon fusion and  
686        quark/anti-quark scattering,
- 687     • a gluino and squark produced together via quark-gluon scattering,
- 688     • squark pairs produced via quark-quark scattering.

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

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

700 However recent results from the **LHC** now strongly disfavour large swathes of **CMSSM**  
701 parameter space [39][40][41]. In the face of such results a more pragmatic model indepen-  
702 dent search strategy is now applied across most **SUSY** searches at the **LHC**, see Section  
703 (2.4.1).

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

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

### <sup>717</sup> 2.4.1. Simplified models

<sup>718</sup> With such a variety of different ways for a **SUSY** signal to manifest itself, it is necessary  
<sup>719</sup> to be able to interpret experimental reach through the masses of gluinos and squarks  
<sup>720</sup> which can be excluded by experimental searches rather than on a model specific basis.

<sup>721</sup> This is accomplished through **SMS** models, which are defined by a set of hypothetical  
<sup>722</sup> particles and a sequence of their production and decay modes [42][43]. In the **SMS** models  
<sup>723</sup> considered within this thesis, only the production process for the two primary particles  
<sup>724</sup> are considered. Each primary particle can undergo a direct or a cascade decay through  
<sup>725</sup> an intermediate new particle. At the end of each decay chain there remains a neutral,  
<sup>726</sup> undetected **LSP** particle, denoted  $\tilde{\chi}_{LSP}$  which can represent a neutralino or gravitino.  
<sup>727</sup> Essentially it is easier to consider each **SMS** with branching ratios set to 100%. The  
<sup>728</sup> masses of the primary particle and the **LSP** remain as free parameters, in which the  
<sup>729</sup> absolute value and relative difference between the primary and **LSP** particle alter the  
<sup>730</sup> kinematics of the event.

<sup>731</sup> Different **SMS** models are denoted with a T-prefix, with a summary of the types interpreted  
<sup>732</sup> within this thesis listed below [44].

- <sup>733</sup> • **T1,T1xxxx**, models represent a simplified version of gluino pair production with  
<sup>734</sup> each gluino (superpartner to the gluon) undergoing a three-body decay to a quark-  
<sup>735</sup> antiquark pair and the **LSP** (i.e.  $\tilde{g} \rightarrow q\bar{q}\tilde{\chi}_{LSP}$ ). The resultant final state from this  
<sup>736</sup> decay is typically 4 jets +  $\cancel{E}_T$  in the absence of initial/final state radiation and  
<sup>737</sup> detector effects. xxxx denotes models in which the quarks are of a specific flavour,  
<sup>738</sup> typically t or b quark-antiquarks.

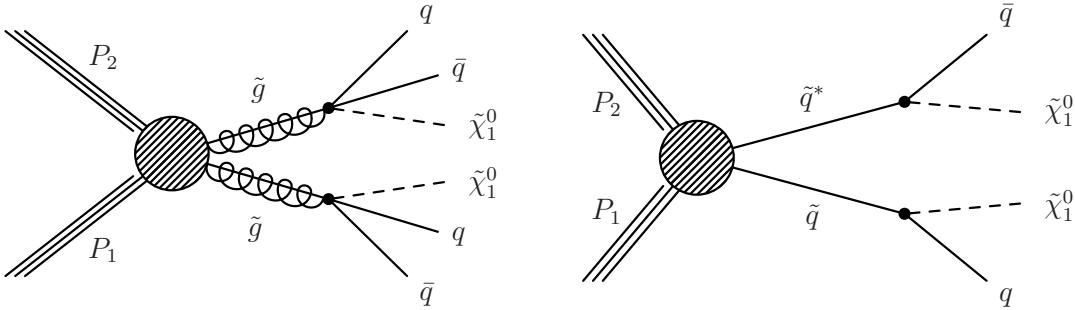
- <sup>739</sup> • **T2,T2xx**, models represent a simplified version of squark anti-squark production  
<sup>740</sup> with each squark undergoing a two-body decay into a light-flavour quark and **LSP**  
<sup>741</sup> (i.e.  $\tilde{q} \rightarrow q\tilde{\chi}_{LSP}$ ). This results in final states with less jets than gluino mediated  
<sup>742</sup> production, typically 2 jets +  $\cancel{E}_T$  when again ignoring the effect of initial/final state  
<sup>743</sup> radiation and detector effects. xx models represent decays in which both the quark  
<sup>744</sup> and the squark within the decay is of a specific flavour, which in this thesis are  
<sup>745</sup> again  $\tilde{t}/t$  or  $\tilde{b}/b$ .

<sup>746</sup> Models rich in b and t quarks are interpreted within this thesis as they remain of  
<sup>747</sup> particular interest within “Natural **SUSY**” scenarios [45][46]. The largest contribution  
<sup>748</sup> to the quadratic divergence in the Higgs mass parameter comes from a loop of top  
<sup>749</sup> quarks via the Yukawa coupling. Cancellation of these divergences can be achieved in

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

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

756 Two example decay chains are shown in Figure 2.2; the pair production of gluinos (T1)  
757 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>758</sup> The LHC And The CMS Detector

<sup>759</sup> Probing the SM for signs of new physics would not be possible without the immensely  
<sup>760</sup> complex electronics and machinery that makes the TeV energy scale accessible for the  
<sup>761</sup> first time. This chapter will describe both the LHC based at European Organization  
<sup>762</sup> for Nuclear Research (CERN) and the Compact Muon Solenoid (CMS) detector, being  
<sup>763</sup> the experiment the author is a member of. Section (3.2) serves to introduce an overview  
<sup>764</sup> of the different components of the CMS detector, with specific components relevant to  
<sup>765</sup> the search for supersymmetric particles described in greater detail. Section (3.3) will  
<sup>766</sup> focus on event and object reconstruction again with more emphasis on jet level quantities  
<sup>767</sup> which are most relevant to the author's analysis research. Finally Section (3.4) will  
<sup>768</sup> cover work performed by the author, as service to the CMS Collaboration, in measuring  
<sup>769</sup> the performance of the Global Calorimeter Trigger (GCT) component of the L1 trigger  
<sup>770</sup> during the 2012-2013 run period.

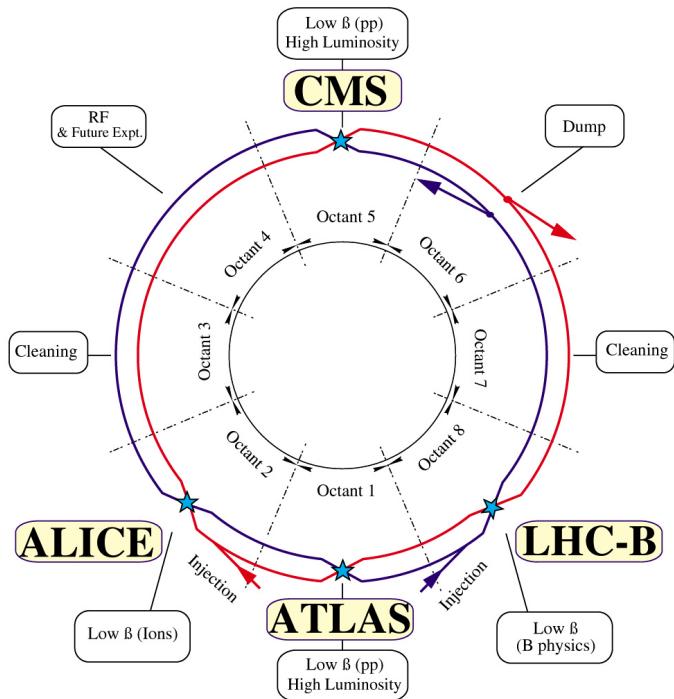
### <sup>771</sup> 3.1. The LHC

<sup>772</sup> The LHC is a storage ring, accelerator, and collider of circulating beams of protons or  
<sup>773</sup> ions. Housed in the tunnel dug for the Large Electron-Positron collider (LEP), it is  
<sup>774</sup> approximately 27 km in circumference, 100 m underground, and straddles the border  
<sup>775</sup> between France and Switzerland outside of Geneva. It is currently the only collider  
<sup>776</sup> in operation that is able to study physics at the TeV scale. A double-ring circular  
<sup>777</sup> synchrotron, it was designed to collide both proton-proton (pp) and heavy ion (PbPb)  
<sup>778</sup> with a centre of mass energy  $\sqrt{s} = 14$  TeV at a final design luminosity of  $10^{34}\text{cm}^{-2}\text{s}^{-1}$ .

<sup>779</sup>

These counter-circulating beams of protons/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.

791



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

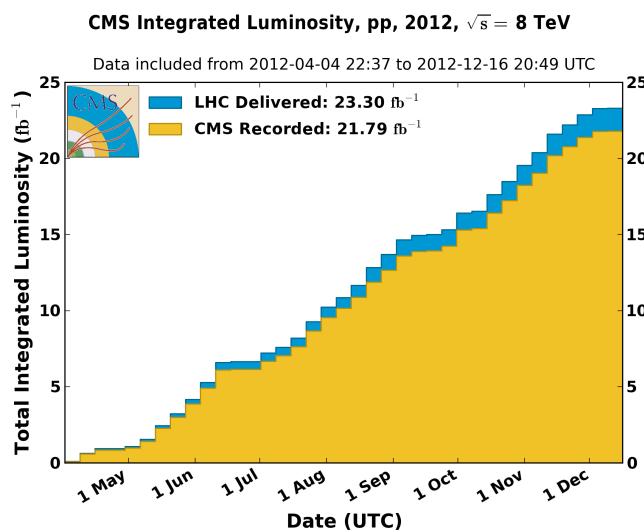
Proton beams are formed inside the Proton Synchrotron (**PS**) from bunches of protons 50 ns apart with an energy of 26 GeV. The protons are then accelerated in the Super Proton Synchrotron (**SPS**) to 450 GeV before being injected into the **LHC**. These **LHC** proton beams consists of many “bunches” i.e. approximately  $1.1 \times 10^{11}$  protons localized into less than 1 ns in the direction of motion. Before collision the beams are ramped to 4

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

802

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

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



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

---

## 817 3.2. The CMS Detector

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

827

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

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

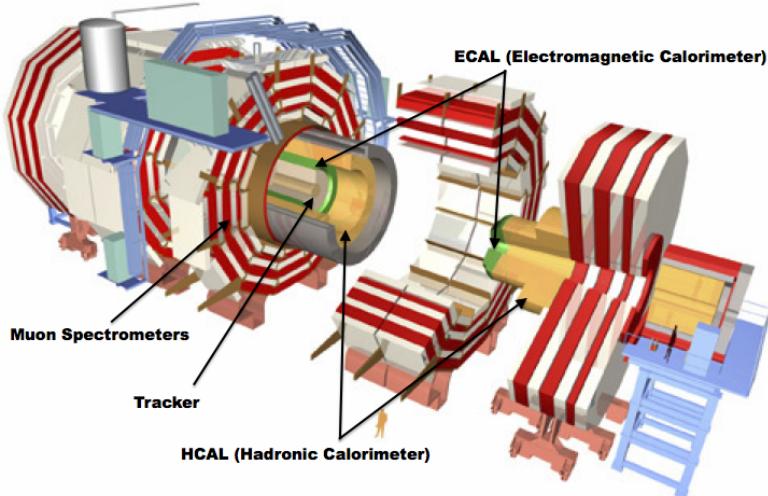
834 The variable  $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$  is commonly used to define angular distance between  
 835 objects within the detector and additionally energy and momentum is typically measured  
 836 in the transverse plane perpendicular to the beam line. These values are calculated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

### 839 3.2.1. Detector subsystems

840 As the range of particles produced in  $pp$  collisions interact in different ways with mat-  
 841 ter, **CMS** is divided into subdetector systems, which perform complementary roles to  
 842 identify the identity, mass and momentum of the different physics objects present in  
 843 each event. These detector sub-systems contained within **CMS** are wrapped in layers

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

847



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

### 848 3.2.2. Tracker

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

858

859 The tracking system also plays a crucial part in the identification of jets originating  
 860 from b-quarks through measurement of displaced secondary vertices, which is covered in  
 861 more detail in Section (3.3.2). The identification of b-jets is important in many searches

862 for natural SUSY models and forms an important part of the inclusive search strategy  
863 described within Section (4.2).

864 **3.2.3. Electromagnetic calorimeter**

865 Immediately outside of the tracker, but still within the magnet core, sits the Electromag-  
866 netic CALorimeter (**ECAL**). Covering a pseudorapidity up to  $|\eta| < 3$  and comprising  
867 of over 75,000 PbWO<sub>4</sub> (lead tungstate) crystals that scintillate as particles deposit energy,  
868 the **ECAL** provides high resolution measurements of the electromagnetic showers from  
869 photons, electrons in the detector.

870

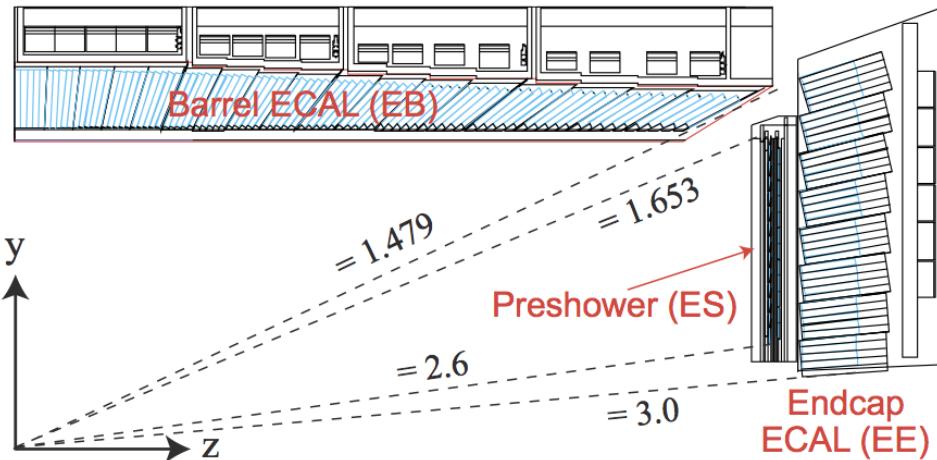
871 Lead tungstate is used because of its short radiation length ( $X_0 \sim 0.9\text{cm}$ ) and small  
872 Molieré radius ( $\sim 2.1\text{cm}$ ) leading to high granularity and resolution. It's fast scintillation  
873 time ( $\sim 25\text{ns}$ ) reduces the effects of pile-up due to energy from previous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

878

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

887

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



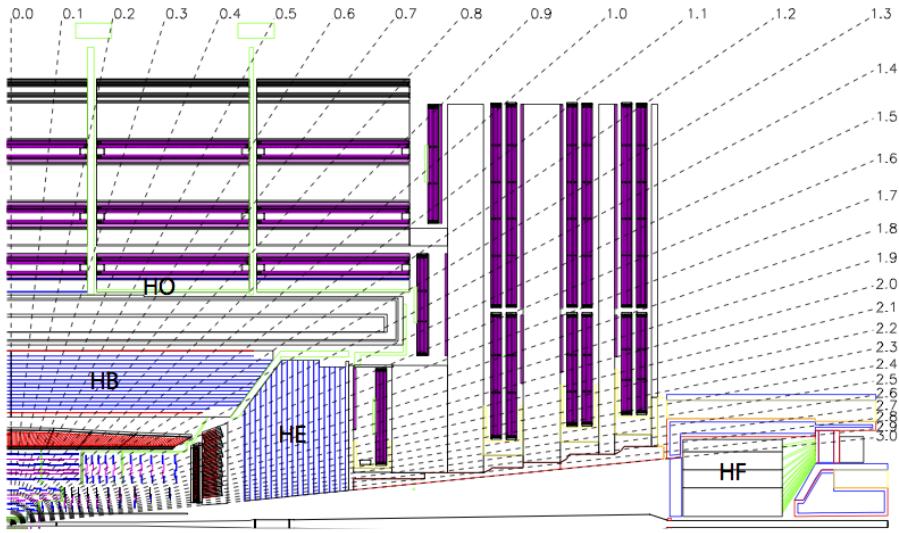
**Figure 3.4.:** Illustration of the **ECAL** showing the arrangement of the lead tungstate crystals in the **EB** and **EE**. The **ES** is also shown and is located in front of the **EE** [55].

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

### 896 3.2.4. Hadronic calorimeter

897 Beyond the **ECAL** lies the Hadronic CALorimeter (**HCAL**) which is responsible for  
 898 the accurate measurement of hadronic showers, crucial for analyses involving jets or  
 899 missing energy signatures. The **HCAL** is a sampling calorimeter which consists of al-  
 900 ternating layers of brass absorber and plastic scintillator, except in the hadron forward  
 901 ( $3.0 < |\eta| < 5.0$ ) region in which steel absorbers and quartz fibre scintillators are used  
 902 because of their increased radiation tolerance. Hadron showers are initiated in the  
 903 absorber layers inducing scintillation in the plastic scintillator tiles. These scintillation  
 904 photons are converted by wavelength shifting fibres for read-out by hybrid photodiodes.  
 905

906 The **HCAL**'s size is constrained to a compact size by the presence of the solenoid, re-  
 907 quiring the placement of an additional outer calorimeter on the outside of the solenoid  
 908 to increase the sampling depth of the **HCAL**. A schematic of the **HCAL** can be seen in  
 909 Figure 3.5.  
 910



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

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

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

927 **3.2.5. Muon systems**

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

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

938

939 **3.3. Event Reconstruction and Object Definition**

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

947 **3.3.1. Jets**

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

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

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

959

960 There are two main type of jet reconstruction used at **CMS**, Calorimeter (Calo) and  
961 Particle Flow (PF) jets [59]. Calorimeter jets are reconstructed using both the **ECAL**  
962 and **HCAL** cells, combined into calorimeter towers . These calorimeter towers consist of  
963 geometrically matched **HCAL** cells and **ECAL** crystals. Electronics noise is suppressed by  
964 applying a threshold to the calorimeter cells, with pile-up effects reduced by a requirement  
965 placed on the tower energy [60]. Calorimter jets are the jets used within the analyses  
966 described in this thesis.

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

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

- 981     • A PU correction is first applied to the jet. It subtracts the average extra energy  
982         deposited in the jet that comes from other vertices present in the event and is  
983         therefore not part of the hard jet itself.
- 984     •  $p_T$  and  $\eta$  dependant corrections derived from Monte Carlo simulations are used to  
985         account for the non-uniform response of the detector.
- 986     •  $p_T$  and  $\eta$  residual corrections are applied to data only to correct for difference  
987         between data and Monte Carlo. The residual is derived from QCD dijet samples  
988         and the  $p_T$  residual from  $\gamma+$  jet and  $Z+$  jets samples in data.

---

### <sup>989</sup> 3.3.2. B-tagging

<sup>990</sup> The decays of b quarks are suppressed by small CKM matrix elements. As a result, the  
<sup>991</sup> lifetimes of b-flavoured hadrons, produced in the fragmentation of b quarks, are relatively  
<sup>992</sup> long;  $\mathcal{O}$  1ps. The identification of jets originating from b quarks is very important for  
<sup>993</sup> searches for new physics and for measurements of standard model processes.

<sup>994</sup>

<sup>995</sup> Many different algorithms developed by CMS select b-quark jets based on variables such  
<sup>996</sup> as the impact parameters of the charged-particle tracks, the properties of reconstructed  
<sup>997</sup> decay vertices, and the presence or absence of a lepton, or combinations thereof [63]. One  
<sup>998</sup> of the most efficient of which is the Combined Secondary Vertex (CSV) which operates  
<sup>999</sup> based on secondary vertex and track-based lifetime information, benchmarked in ‘Loose’,  
<sup>1000</sup> ‘Medium’ and ‘Tight’ working points, of which the medium point is the tagger used  
<sup>1001</sup> within the  $\alpha_T$  search detailed in Section (4.1).

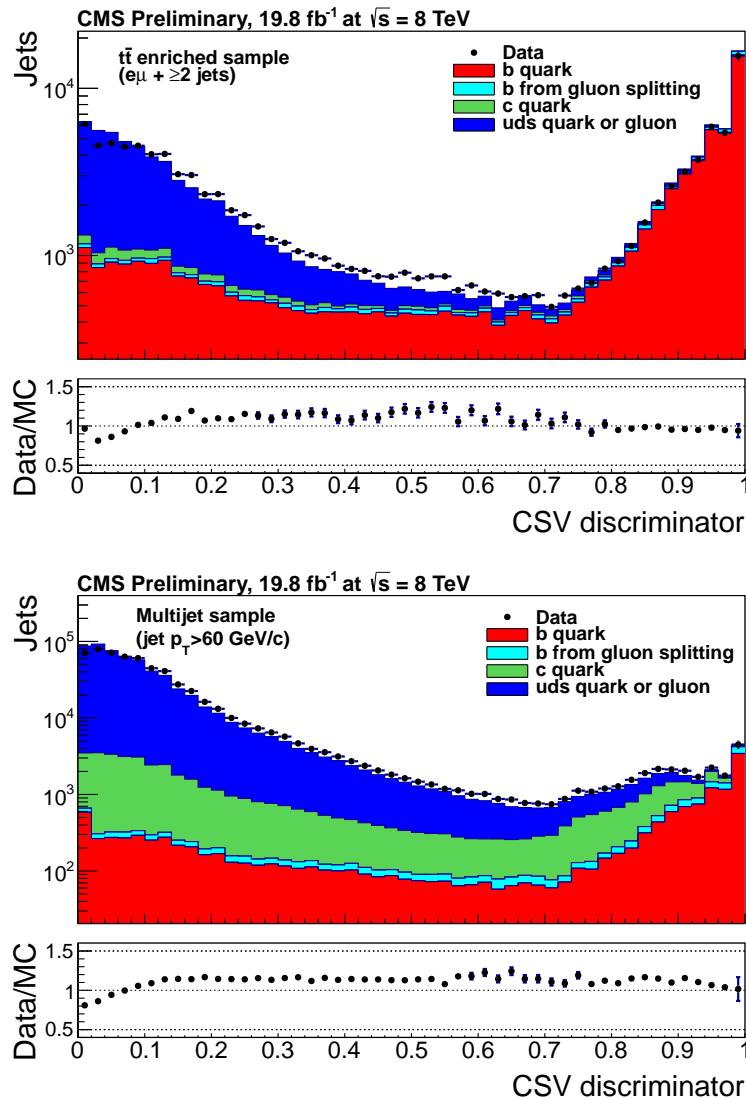
<sup>1002</sup> Using the CSV tagger, a likelihood-based discriminator distinguishes between jets from  
<sup>1003</sup> b-quarks, and those from charm or light quarks and gluons, which is shown in Figure 3.6.  
<sup>1004</sup> The minimum thresholds on the discriminator for each working point correspond to the  
<sup>1005</sup> misidentification probability for light-parton jets of 10%, 1%, and 0.1%, respectively, in  
<sup>1006</sup> jets with an average  $p_T$  of about 80 GeV.

<sup>1007</sup> The b-tagging performance is evaluated to measure the b-jet tagging efficiency  $\epsilon_b$ , and the  
<sup>1008</sup> misidentification probability of charm  $\epsilon_c$  and light-parton jets  $\epsilon_s$ . The tagging efficiencies  
<sup>1009</sup> for each of these three jet flavours are compared between data and MC simulation, from  
<sup>1010</sup> which a series of  $p_T$  and  $|\eta|$  binned jet corrections are determined,

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

<sup>1011</sup> These are collectively named ‘Btag Scale Factors’ and allow MC simulation to accu-  
<sup>1012</sup> rately reflect the running conditions and performance of the tagging algorithm in data.  
<sup>1013</sup> Understanding of the b-tagging efficiency is essential in order to minimise systematic  
<sup>1014</sup> uncertainties in physics analyses that employ b-tagging.

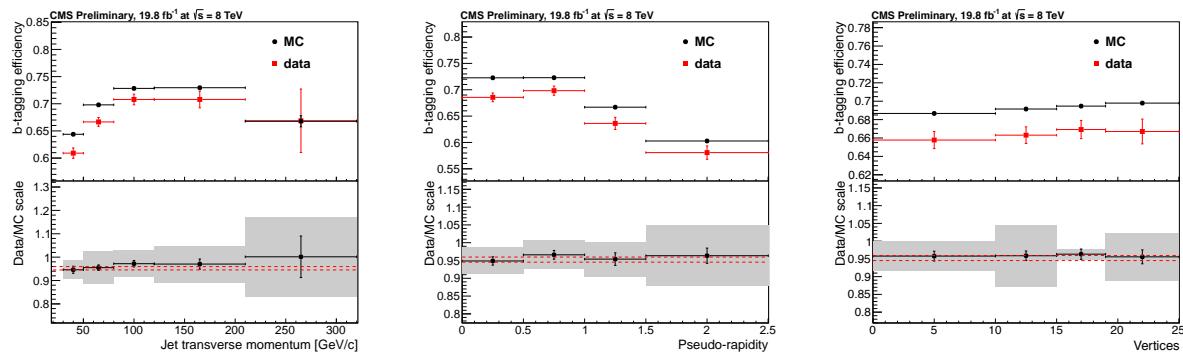
<sup>1015</sup>



**Figure 3.6.:** CSV algorithm discriminator values in enriched  $t\bar{t}$  (top) and inclusive multi-jet samples (bottom) for b,c and light flavoured jets [64]. Working points are determined from the misidentification probability for light-parton jets to be tagged as a b-jet and are given as 0.244, 0.679 and 0.898 for L,M and T working points respectively.

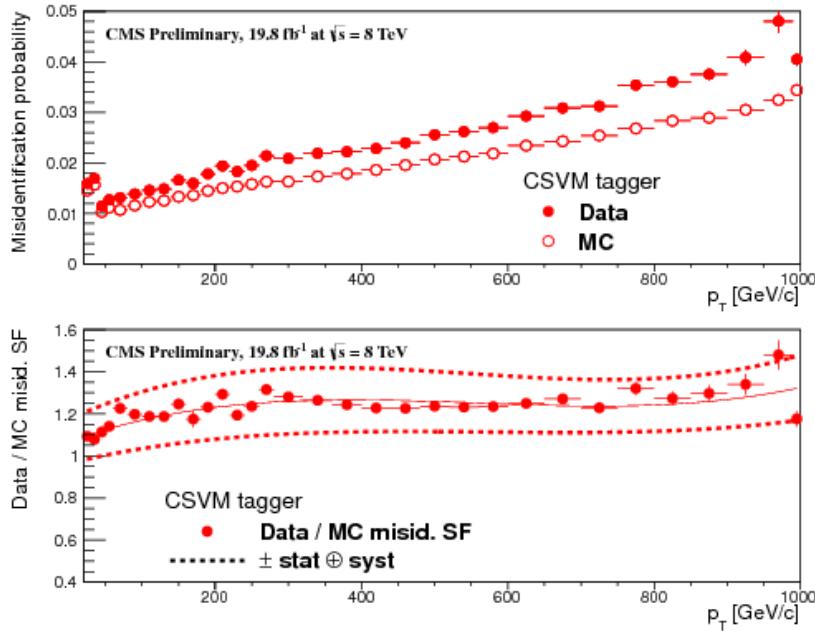
1016 The b-tagging efficiency is measured in data using several methods applied to multi  
 1017 jet events, primarily based on a sample of jets enriched in heavy flavour content. One  
 1018 method requires the collection of events with a soft muon within a cone  $\Delta R < 0.4$  around  
 1019 the jet axis. Because the semileptonic branching fraction of b hadrons is significantly  
 1020 larger than that for other hadrons, these jets are more likely to arise from b quarks than  
 1021 from another flavour, with the resultant momentum component of the muon transverse  
 1022 to the jet axis larger for muons from b-hadron decays than from light or charm jets.

1023 Additionally the performance of the tagger can also be benchmarked in  $t\bar{t}$  events where  
 1024 in the SM, the top quark is expected to decay to a W boson and a b quark about 99.8%  
 1025 of the time [1]. Further selection criteria is applied to these events to further enrich the  
 1026 b quark content of these events. The methods to identify b-jets in data are discussed  
 1027 in great detail at [65]. The jet flavours are determined in simulation using truth level  
 1028 information and are compared to data to determine the correction scale factors ( $SF_b$ ),  
 1029 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.

1030 The measurement of the misidentification probability for light-parton jets relies on the  
 1031 inversion of tagging algorithms, selecting non-b jets using the same variables and tech-  
 1032 niques used in benchmarking the b-tagging efficiency. The scale factors ( $SF_s$ ) to be  
 1033 applied to MC are shown in Figure 3.8 for the CSVM tagger.



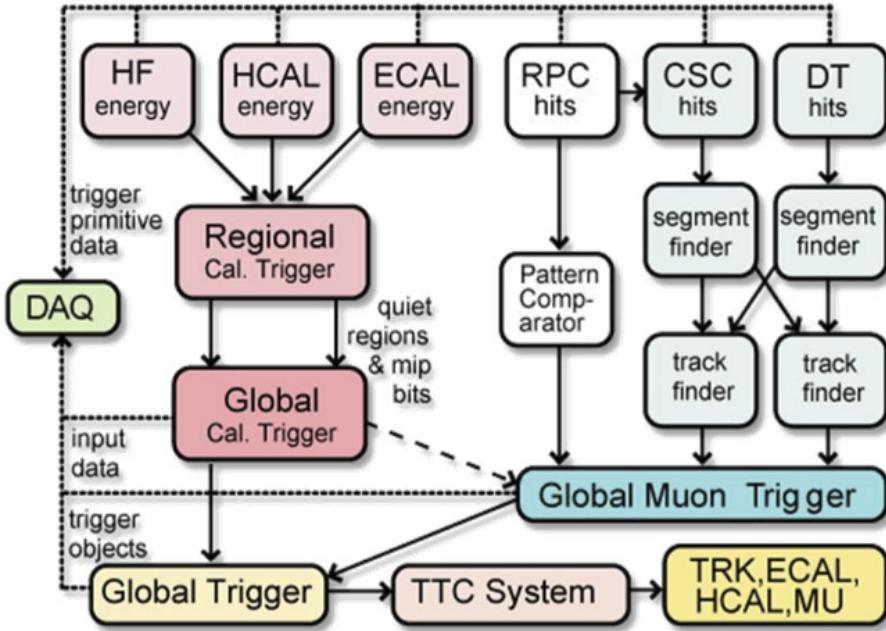
**Figure 3.8.:** For the **CSVM** tagging criterion: (top) misidentification probability in data (filled circles) and simulation (open circles); (bottom) scale factor for the misidentification probability. The last  $p_T$  bin in each plot includes all jets with  $p_T > 1000 \text{ GeV}$ . The solid 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.

## 1034 3.4. Triggering System

1035 With bunch crossings separated by just 25 ns, the rate at which data from all collisions  
1036 would have to be written out and processed would be unfeasible. A two-tiered triggering  
1037 system is applied at **CMS** in order to cope with the high collision rate of protons. The  
1038 **CMS** trigger is designed to use limited information from each event to determine whether  
1039 to record the event, reducing the rate of data taking to manageable levels whilst ensuring  
1040 a high efficiency of interesting physics object events are selected.

1041 The **L1** is a pipelined, dead-timeless system based on custom-built electronics [66], and is  
1042 a combination of several sub systems which is shown pictorially in Figure 3.9. The L1  
1043 system is covered in more detail within the following section along with a description  
1044 of the service work undertaken by the author to benchmark the performance of the L1  
1045 calorimeter trigger during the 2012 8 TeV run period.

1046 The Higher Level Trigger (**HLT**) is a large farm of commercial computers [67]. The **HLT**  
1047 processes events with software reconstruction algorithms that are more detailed, giving  
1048 performance more similar to the reconstruction used offline. The **HLT** reduces the event



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

1049 rate written to disk by a factor of  $\sim 500$  ( $\sim 200\text{Hz}$ ). The recorded events are transferred  
 1050 from CMS to the CERN computing centre, where event reconstruction is performed, and  
 1051 then distributed to CMS computing sites around the globe for storage and analysis.

### 1052 3.4.1. The level-1 trigger

1053 The L1 trigger reduces the rate of events collected from 40 MHz to  $\sim 100\text{ kHz}$  using  
 1054 information from the calorimeters and muon chambers, but not the tracker. A tree  
 1055 system of triggers is used to decide whether to pass on an event to the HLT for further  
 1056 reconstruction. Firstly the calorimeter and muon event information is kept separate,  
 1057 with local reconstruction of objects ( $\mu, e, \gamma, \text{jets}$ ) performed by the Regional Calorimeter  
 1058 Trigger (RCT) and Regional Muon Trigger (RMT) respectively. The RCT generates up to  
 1059 72 isolated and non-isolated electromagnetic objects. These are sorted by rank, which is  
 1060 equivalent to transverse energy  $E_T$ , with the four highest ranked electromagnetic objects  
 1061 being passed via the Global Calorimeter Trigger (GCT) and Global MuonTrigger (GMT)  
 1062 to the Global Trigger (GT).

1063 In the L1 **GCT**, coarse measurements of the energy deposited in the electromagnetic and  
1064 hadronic calorimeters are combined and by using sophisticated algorithms the following  
1065 physics objects are formed:

- 1066 • isolated and non-isolated electromagnetic objects ( $e$  and  $\gamma$ );
- 1067 • hadronic jets in the central and forward sections of the hadronic calorimeters;
- 1068 • hadronically decaying tau leptons;
- 1069 • total transverse energy ( $E_T$ ), the scalar sum of the energy measured at L1, and  
1070 missing transverse energy ( $\cancel{E}_T$ ), defined as the vector sum of the energy of L1  
1071 objects;
- 1072 • total transverse jet energy ( $H_T$ ), the scalar sum of the energy of all L1 jet objects,  
1073 and missing transverse jet energy ( $\cancel{H}_T$ ), defined as the vector sum of the energy of  
1074 L1 jets, are calculated from uncorrected L1 jets.

1075 In addition quantities suitable for triggering minimum bias events, forward physics and  
1076 beam background events are calculated. Additionally relevant muon isolation information  
1077 is also passed on to the **GMT** for decisions involving the muon triggers where it is  
1078 combined with information from across the three muon sub-systems. The resultant final  
1079 accept/reject decision at **L1** is then performed by the **GT** based on the objects received  
1080 from the **GCT** and **GMT** ( $e/\gamma$ ,  $\mu$ , jets,  $E_T$ ,  $\cancel{E}_T$ ,  $H_T$ ).

1081 The L1 trigger is therefore of upmost importance to the functioning of the detector.  
1082 Without a high-performing trigger and a good understanding of its performance, there  
1083 would be no data to analyse. Observations of how the L1 trigger performance is affected  
1084 by changing **LHC** running conditions over the 2012 run period and also the introduction  
1085 of a jet seed threshold to the L1 jet trigger algorithm is presented in the following Sections  
1086 (3.4.2 - 3.4.6).

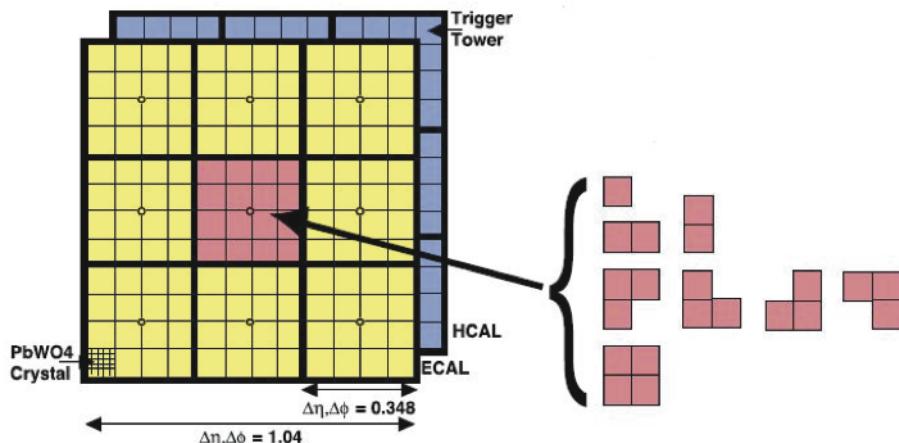
### 1087 3.4.2. The L1 trigger jet algorithm

1088 The L1 jet trigger uses the transverse energy sums computed in the calorimeter (both  
1089 hadronic and electromagnetic) trigger regions. Each region consists of  $4 \times 4$  trigger tower  
1090 windows, spanning a region of  $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$  in pseudorapidity-azimuth. The  
1091 jet trigger uses a  $3 \times 3$  calorimeter region (112 trigger towers) sliding window technique  
1092 which spans the full  $(\eta, \phi)$  coverage of the **CMS** calorimeter as shown in Figure 3.10.

1093 In forming a L1 jet is it required that the central region to be higher than the eight  
1094 neighbouring regions  $E_{T\text{central}} > E_{T\text{surround}}$ . Additionally a minimum threshold of 5 GeV  
1095 on  $E_{T\text{central}}$  was introduced during the 2012 run period to suppress noise from pile-up,  
1096 the effects of which are shown in Section (3.4.4).

1097 The L1 jets are characterised by the  $E_T$ , summed over the  $3 \times 3$  calorimeter regions,  
1098 which corresponds to  $12 \times 12$  trigger towers in barrel and endcap or  $3 \times 3$  larger **HF**  
1099 towers in the **HF**. The  $\phi$  size of the jet window is the same everywhere, whilst the  $\eta$   
1100 binning gets somewhat larger at high  $\eta$  due to calorimeter and trigger tower segmentation.  
1101 The jets are labelled by  $(\eta, \phi)$  indexes of the central calorimeter region.

1102 Jets with  $|\eta| > 3.0$  are classified as forward jets, whereas those with  $|\eta| < 3.0$  are classified  
1103 as central. The four highest energy central, forward and  $\tau$  jets in the calorimeter are  
1104 passed through Look Up Table (**LUT**)’s, which apply a programmable  $\eta$ –dependent jet  
1105 energy scale correction. These are then used to make L1 trigger decisions.



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

1106 The performance of the L1 jets is evaluated with respect to offline jets, which are taken  
1107 from the standard Calo jet and the PF jet reconstruction algorithms of **CMS**. Jets are  
1108 corrected for pile-up and detector effects as described in 3.3.1. A moderate level of noise  
1109 rejection is applied to the offline jets by selecting jets passing the “loose identification  
1110 criteria for both Calo and PF. These criteria are summarised in Appendix (A).

### 1111 3.4.3. Measuring L1 jet trigger efficiencies

1112 The L1 jet efficiency is defined as the fraction of leading offline jets which were matched  
1113 with a L1 tau or central jet above a certain trigger threshold, divided by all the leading  
1114 offline jets in the event. This quantity is then plotted as a function of the offline jet  $E_T$ ,  
1115  $\eta$  and  $\phi$ .

1116 The efficiency is determined by matching the L1 and reconstructed offline jets spatially  
1117 in  $\eta - \phi$  space. This is done by calculating the minimum separation in  $\Delta R$  between the  
1118 highest offline reconstructed jet in  $E_T$  ( $E_T > 10$  GeV,  $|\eta| < 3$ ) and any L1 jet. A jet will  
1119 be matched if this value is found to be  $< 0.5$ . Should more than one jet satisfy this, the  
1120 jet closest in  $\Delta R$  is taken as the matched jet. The matching efficiency is close to 100%,  
1121 above 30(45) GeV for run 2012B(C) data (see Appendix B.1).

1122 Each efficiency curve is fitted with a function which is the cumulative distribution function  
1123 of an Exponentially Modified Gaussian (EMG) distribution:

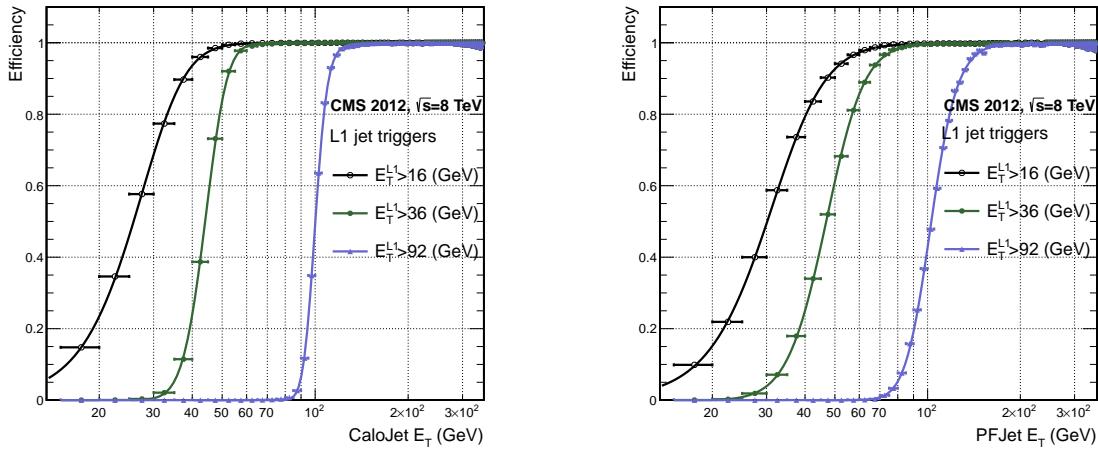
$$\text{f}(x; \mu, \sigma, \lambda) = \frac{\lambda}{2} \cdot e^{\frac{\lambda}{2}(2\mu + \lambda\sigma^2 - 2x)} \cdot \text{erfc}\left(\frac{\mu + \lambda\sigma^2 - x}{\sqrt{2}\sigma}\right) \quad (3.3)$$

where erfc is the complementary error function defined as:

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

1125 In this functional form, the parameter  $\mu$  determines the point of 50% of the plateau  
1126 efficiency and the  $\sigma$  gives the resolution. This parametrisation is used to benchmark  
1127 the efficiency at the plateau, the turn-on points and resolution for each L1 Jet trigger.  
1128 The choice of function is purely empirical. Previous studies used the error function  
1129 alone, which described the data well at high threshold values but could not describe the  
1130 efficiencies well at lower thresholds [68].

1131 The efficiency turn-on curves for various L1 jet thresholds are evaluated as a function of  
1132 the offline reconstructed jet  $E_T$  for central jets with  $|\eta| < 3$ . These are measured using  
1133 single isolated  $\mu$  triggers which have high statistics, and are orthogonal and therefore  
1134 unbiased to the hadronic triggers under study. The efficiency is calculated with respect to  
1135 offline Calo and PF Jets in Figure 3.11. Table 3.1 shows the values of these parameters,  
1136 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.

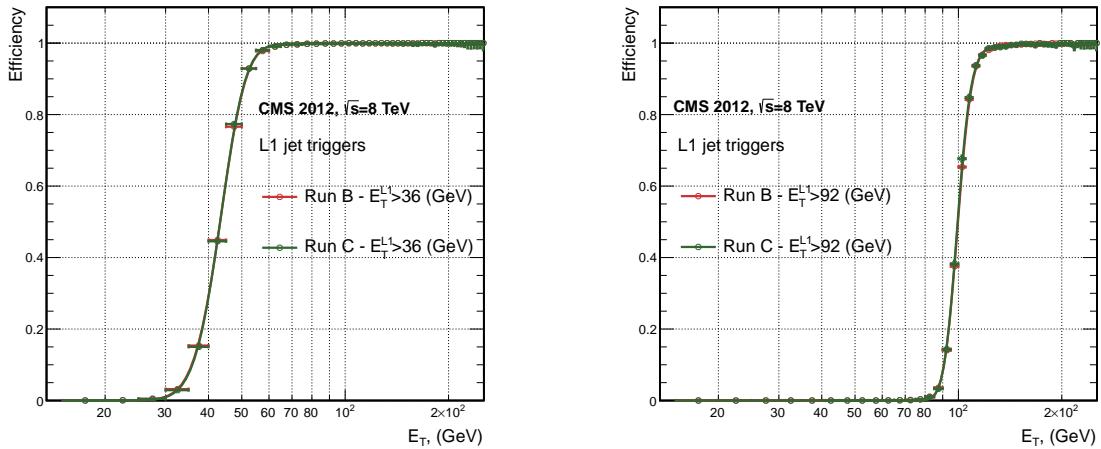
#### 3.4.4. Effects of the L1 jet seed

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

1147 counteract the effects of high pile up running conditions which create a large number of  
1148 soft non-collimated jets, that are then added to the jets from the primary interaction or  
1149 other soft jets from other secondary interactions [69]. This in turn causes a large increase  
1150 in trigger rate due to the increase in the likelihood that the event causes the L1 trigger to  
1151 fire. This was implemented to maintain trigger thresholds by cutting the rate of events  
1152 recorded without significant reduction in the efficiency of physics events of interest.

1153 The effect of the introduction of this jet seed threshold between these two run periods is  
1154 benchmarked through a comparison of the efficiency of the L1 jet triggers with respect  
1155 to offline Calo jets shown in Figure 3.12, and the L1  $H_T$  trigger efficiency in Figure 3.14  
1156 which is compared to offline  $H_T$  constructed from Calo jets with  $E_T \geq 40\text{GeV}$ .

1157 To negate any effects from different pile-up conditions in the run periods, the efficiencies  
1158 are measured in events which contain between 15 and 20 primary vertices as defined in  
1159 Appendix (A.2).



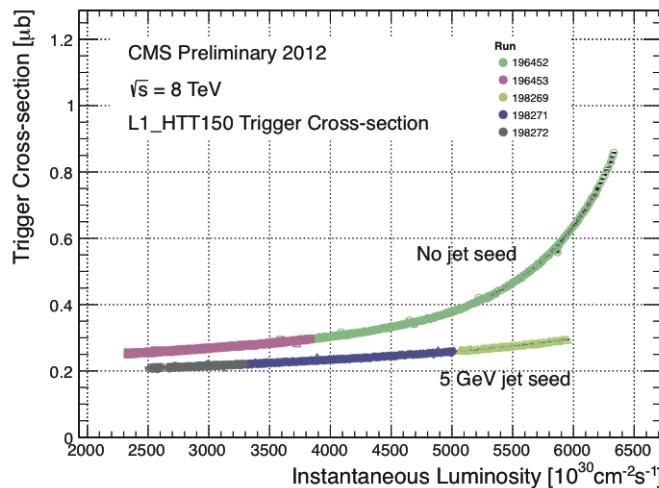
**Figure 3.12.:** L1 jet efficiency turn-on curves as a function of the offline CaloJet  $E_T$ , measured for the L1 SingleJet 36 and 92 trigger in 2012 run period B and C collected with an isolated single  $\mu'$  sample.

1160 It can be seen that the performance of the  $E_T > 36, 92$  single jet are almost identical,  
1161 with the jet seed having no measurable effect on these triggers as shown in Table 3.2 .  
1162 For the  $H_T$  triggers, a large increase in rate during high pile-up conditions is expected.  
1163 This is due to the low energy threshold required for a jet to be added to the L1  $H_T$  sum,  
1164 which is compiled from all uncorrected L1 jets formed in the RCT. The introduction  
1165 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).

<sub>1166</sub> lowering the  $H_T$  calculation at L1. The effect on the trigger cross section for L1  $H_T$  150  
<sub>1167</sub> trigger can be seen in Figure 3.13.

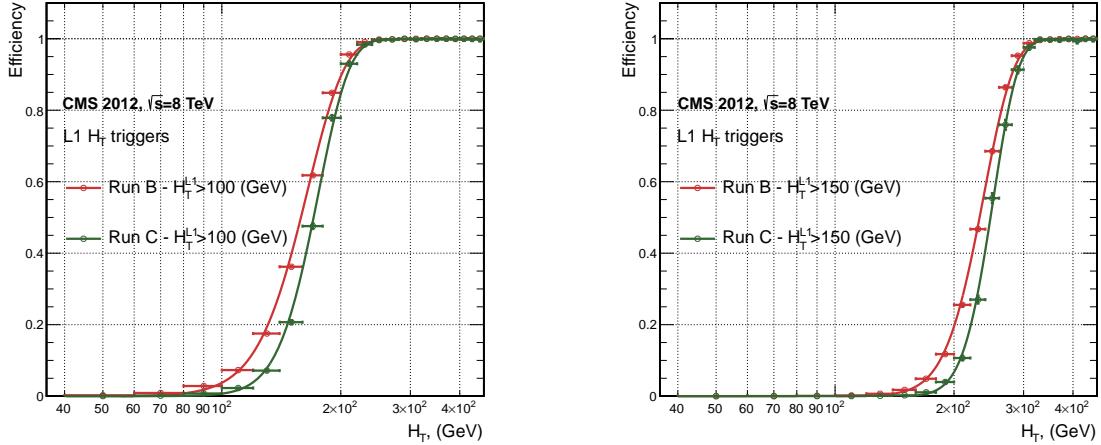


**Figure 3.13.:** Trigger cross section for the L1HTT150 trigger path. Showing that a 5 GeV jet seed threshold dramatically reduces the dependance of cross section on the instantaneous luminosity for L1  $H_T$  triggers [70].

<sub>1168</sub> Different behaviours for the trigger turn ons between these run periods are therefore  
<sub>1169</sub> expected. The turn on point is observed to shift to higher  $H_T$  values after the introduction  
<sub>1170</sub> of the jet seed threshold, whilst having a sharper resolution due to less pile-up jets being  
<sub>1171</sub> included the  $H_T$  sum. This effect is demonstrated in Table 3.3.

### <sub>1172</sub> 3.4.5. Robustness of L1 jet performance against pile-up

<sub>1173</sub> The performance of the L1 single jet triggers is evaluated in different pile-up conditions  
<sub>1174</sub> to benchmark any dependence on pile-up. Three different pile-up bins of 0-10, 10-20 and  
<sub>1175</sub> >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.

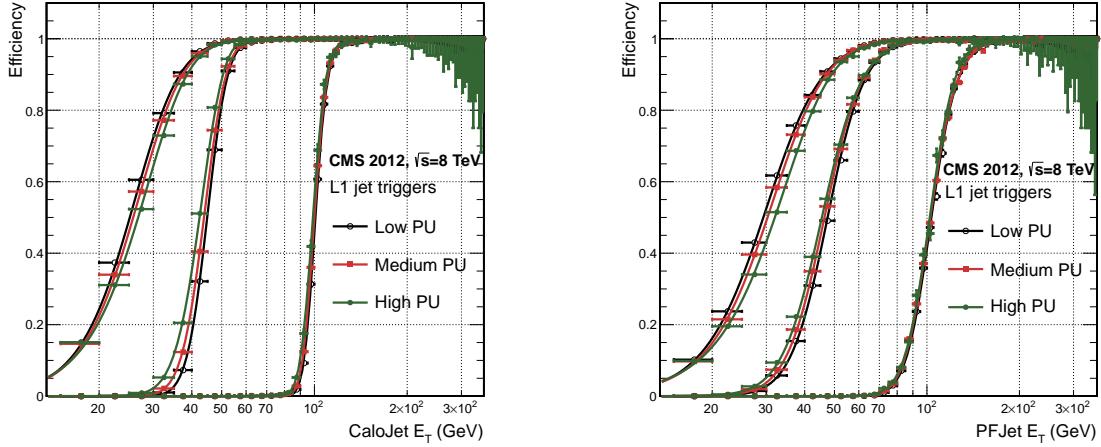
2012B			2012C		
Trigger	$\mu$	$\sigma$	$\mu$	$\sigma$	
L1 HT-100	$157.5 \pm 0.08$	$32.9 \pm 0.08$	$169.8 \pm 0.08$	$28.7 \pm 0.03$	
L1 H1-150	$230.9 \pm 0.02$	$37.3 \pm 0.01$	$246.4 \pm 0.16$	$31.8 \pm 0.05$	

**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.

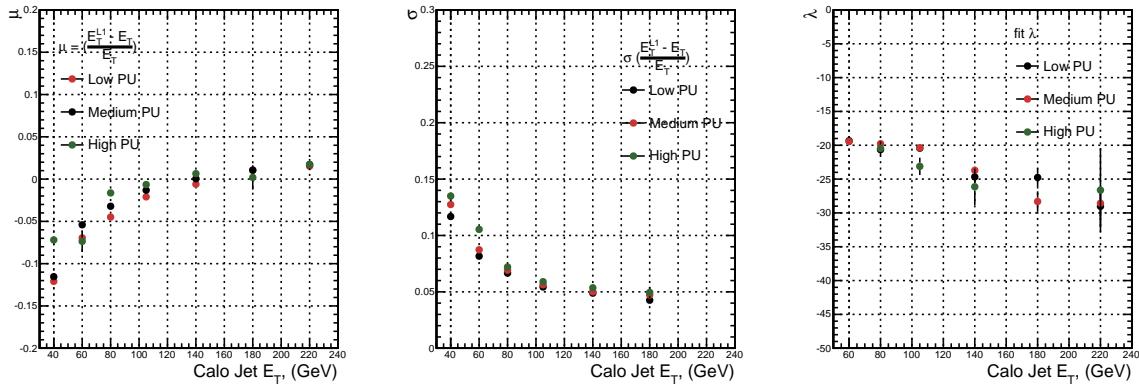
1180 No significant drop in efficiency is observed in the presence of a high number of primary  
 1181 vertices. The increase in hadronic activity in higher pile-up conditions, combined with  
 1182 the absence of pile-up subtraction for L1 jets, results in the expected observation of  
 1183 a decrease in the  $\mu$  value of the efficiency turn-ons as a function of pile-up, while the  
 1184 resolution,  $\sigma$  of the turn-ons are found to gradually worsen as expected with increasing  
 1185 pile-up.

1186 These features are further emphasised when shown as a function of

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

1187 in bins of matched leading offline jet  $E_T$ , of which the individual fits can be found in  
1188 Appendix (B.2). Each of these distributions are fitted with an EMG function as defined  
1189 in Equation (3.3).

1190 The  $\mu$ ,  $\sigma$  and  $\lambda$  values extracted for the low, medium and high pile-up conditions are  
1191 shown for Calo and PF jets in Figure 3.16 and Figure 3.17 respectively. The central value  
1192 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  
1193 is also observed to improve at higher offline jet  $E_T$ .

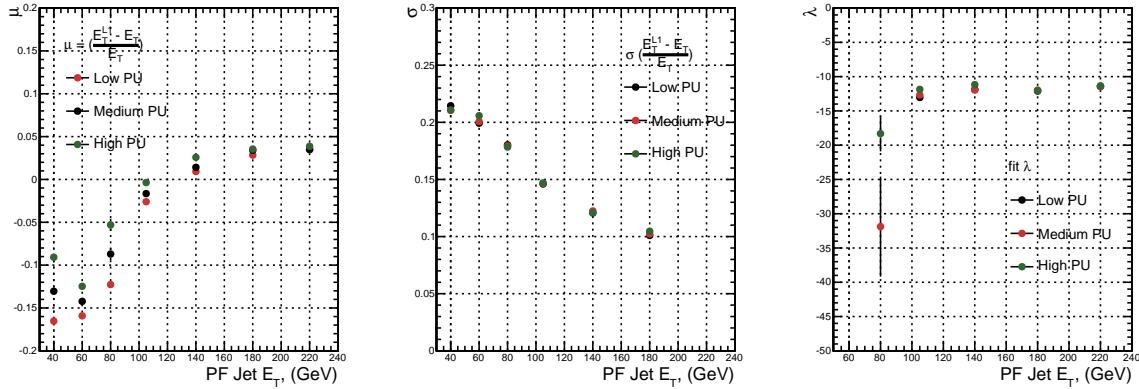


**Figure 3.16.:** Fit values from an EMG function fitted to the resolution plots of leading Calo jet  $E_T$  measured as a function of  $\frac{(L1 E_T - \text{Offline } E_T)}{\text{Offline } E_T}$  for low, medium and high pile-up conditions. The plots show the mean  $\mu$  (left), resolution  $\sigma$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

1194 The resolution of other L1 energy sum quantities,  $H_T$ ,  $\not{E}_T$  and  $\sum E_T$  parameterised as  
1195 in Equation (3.4), can be found in Appendix B.3. The same behaviour observed for  
1196 the single jet triggers is also found for these quantities, where in the presence of higher  
1197 pile-up the  $\mu$  values are shifted to higher values, with a worsening resolution,  $\sigma$  again  
1198 due to the increase in soft pile-up jets and the absence of pile-up subtraction at L1.

### 1199 3.4.6. Summary

1200 The performance of the CMS Level-1 Trigger has been studied and evaluated for jets and  
1201 energy sum quantities using data collected during the 2012 LHC 8TeV run. These studies  
1202 include the effect of introduction of a 5 GeV jet seed threshold into the jet algorithm  
1203 configuration, the purpose of which is to mitigate the effects of pile-up on the rate of  
1204 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.

1205 change in performance is observed with this change and good performance is observed  
 1206 for a range of L1 quantities.

# Chapter 4.

## <sup>1207</sup> SUSY Searches In Hadronic Final States

<sup>1209</sup> In this chapter a model independent search for **SUSY** in hadronic final states with  $\cancel{E}_T$   
<sup>1210</sup> using the  $\alpha_T$  variable and b-quark multiplicity is introduced and described in detail. The  
<sup>1211</sup> results presented are based on a data sample of pp collisions collected in 2012 at  $\sqrt{s} = 8$   
<sup>1212</sup> TeV, corresponding to an integrate luminosity of  $11.7 \pm 0.5 \text{ fb}^{-1}$  [?].

<sup>1213</sup> The kinematic variable  $\alpha_T$  is motivated as a variable to provide strong rejections of QCD  
<sup>1214</sup> backgrounds, whilst maintaining sensitivity to a possible **SUSY** signal within Section  
<sup>1215</sup> (4.1). The search and trigger strategy in addition to the event reconstruction and selection  
<sup>1216</sup> are outlined within Sections (4.2-4.3).

<sup>1217</sup> The method in which the **SM** background is estimated using an analytical technique to  
<sup>1218</sup> improve statistical precision at higher b-tag multiplicities is detailed within Section (4.5),  
<sup>1219</sup> with a discussion on the impact of b-tagging and mis-tagging scale factors between data  
<sup>1220</sup> and MC on any background predictions.

<sup>1221</sup> A description of the formulation of appropriate systematic uncertainties applied to the  
<sup>1222</sup> background predictions to account for theoretical uncertainties and limitations in the  
<sup>1223</sup> simulation modelling of event kinematics and instrumental effects is covered in Section  
<sup>1224</sup> (4.6).

<sup>1225</sup> Finally the statistical likelihood model to interpret the observations in the signal and  
<sup>1226</sup> control samples is described in Section (4.8). The experimental reach of the analysis  
<sup>1227</sup> discussed within this thesis is interpreted in two classes of **SMS** models, the topologies  
<sup>1228</sup> of which are detailed in Section (2.4.1). The **SMS** models considered in this analysis

1229 are summariesed in Table 4.1. For each model, the **LSP** is assumed to be the lightest  
1230 neutralino.

1231 Within the table are also defined reference points, parameterised in terms of parent  
1232 gluino/squark and **LSP** sparticle masses,  $m_{parent}$  and  $m_{LSP}$ , respectively, which are used  
1233 within the following two chapters to demonstrate potential yields within the signal region  
1234 of the search. The masses are chosen to reflect parameter space which is within the  
1235 expected sensitivity reach of the search.

Model	Production/decay mode	Reference model	
		$m_{parent}$	$m_{LSP}$
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)	$pp \rightarrow \tilde{q}\tilde{q}^* \rightarrow q\tilde{\chi}_1^0 \bar{q}\tilde{\chi}_1^0$	600	250
D2 (T2bb)	$pp \rightarrow \tilde{b}\tilde{b}^* \rightarrow b\tilde{\chi}_1^0 \bar{b}\tilde{\chi}_1^0$	500	150
D3 (T2tt)	$pp \rightarrow \tilde{t}\tilde{t}^* \rightarrow t\tilde{\chi}_1^0 \bar{t}\tilde{\chi}_1^0$	400	0

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

## 1236 4.1. An Introduction to the $\alpha_T$ Search

1237 The experimental signature of **SUSY** signal in the hadronic channel would manifest  
1238 itself as a final state containing energetic jets and  $\cancel{E}_T$ . The search focuses on topologies  
1239 where new heavy supersymmetric, R-parity conserving particles are pair-produced in pp  
1240 collisions. These particles decaying to a **LSP** escape the detector undetected, leading to  
1241 significant missing energy and missing hadronic transverse energy,

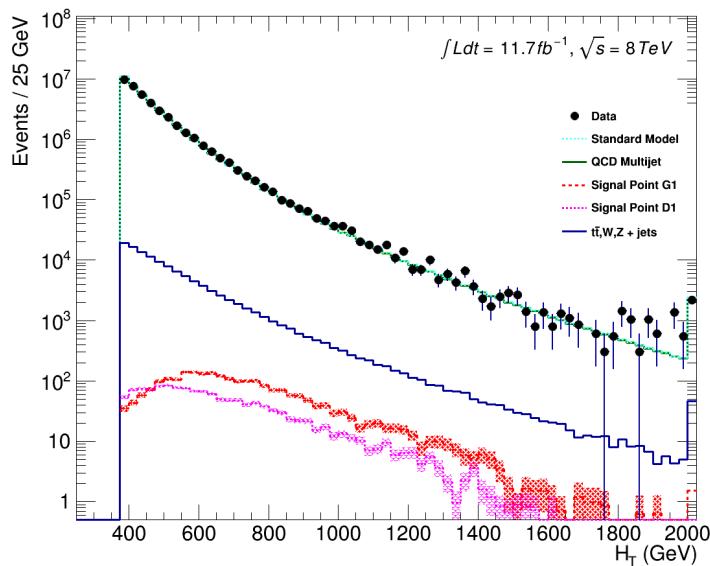
$$\mathcal{H}_T = \left| \sum_{i=1}^n p_T^{jet_i} \right|, \quad (4.1)$$

1242 defined as the vector sum of the transverse energies of jets selected in an event. Energetic  
1243 jets produced in the decay of these supersymmetric particles also can produce significant  
1244 visible transverse energy,

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

1245 defined as the scalar sum of the transverse energies of jets selected in an event.

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



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

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:

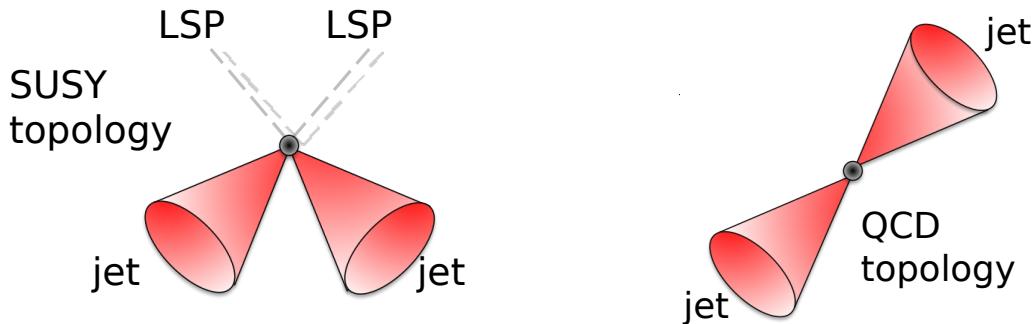
- 1253 •  $Z \rightarrow \nu\bar{\nu} + \text{jets}$ ,
- 1254 •  $W \rightarrow l\nu + \text{jets}$  in which a lepton falls outside of detector acceptance, or the lepton  
 1255 decays hadronically  $\tau \rightarrow \text{had}$ ,

- 1256 •  $t\bar{t}$  with at least one leptonic W decay,
- 1257 • small background contributions from DY, single top and Diboson (WW,ZZ,WZ)
- 1258 processes.

1259 The search is designed to have a strong separation between events with genuine and  
1260 “fake”  $\cancel{E}_T$  which is achieved primarily through the dimensionless kinematic variable,  $\alpha_T$   
1261 [71][72].

1262 **4.1.1. The  $\alpha_T$  variable**

1263 For a perfectly measured di-jet QCD event, conservation laws dictate that they must be  
1264 produced back-to-back and of equal magnitude. However in di-jet events with real  $\cancel{E}_T$ ,  
 both of these jets are produced independently of one another, depicted in Figure 4.2.



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

1265

1266 Exploiting this feature leads to the formulation of  $\alpha_T$  (first inspired by [73]) in di-jet  
1267 systems defined as,

$$\alpha_T = \frac{E_T^{j2}}{M_T}, \quad (4.3)$$

1268 where  $E_T^{j2}$  is the transverse energy of the least energetic of the two jets and  $M_T$  defined  
1269 as:

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

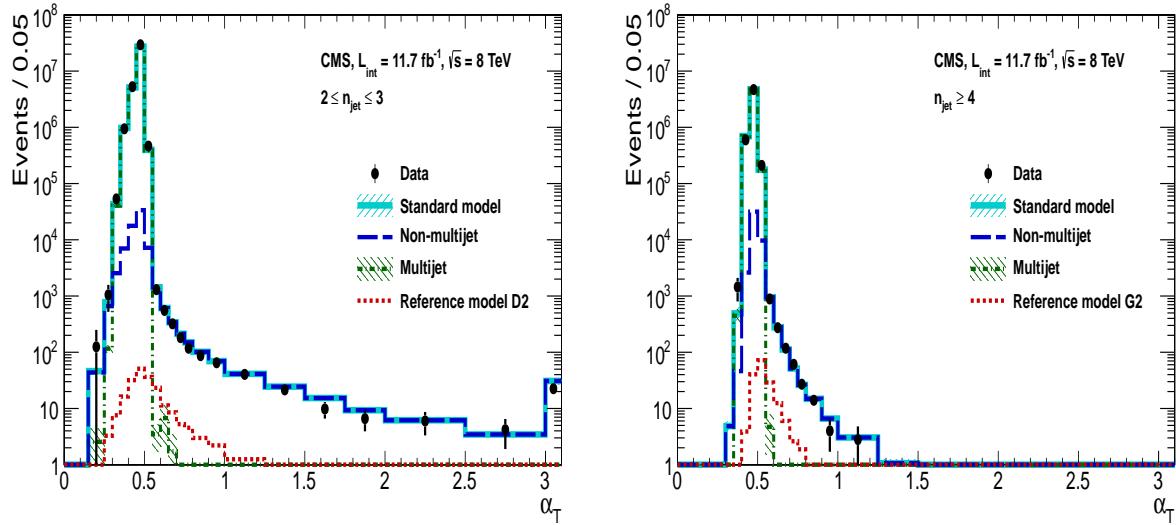
1270 A perfectly balanced di-jet event i.e.  $E_T^{j_1} = E_T^{j_2}$  would give an  $\alpha_T = 0.5$ , whereas events  
1271 with jets which are not back-to-back, for example in events in which a W or Z recoils off  
1272 a system of jets,  $\alpha_T$  can achieve values in excess of 0.5.

1273  $\alpha_T$  can be extended to apply to any arbitrary number of jets, undertaken by modelling a  
1274 system of  $n$  jets as a di-jet system, through the formation of two pseudo-jets [74]. The two  
1275 pseudo-jets are built by merging the jets present in the event such that the 2 pseudo-jets  
1276 are chosen to be as balanced as possible, i.e the  $\Delta H_T \equiv |E_T^{pj_1} - E_T^{pj_2}|$  is minimised  
1277 between the two pseudo jets. Using Equation (4.4),  $\alpha_T$  can be rewritten as,

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

1278 The distribution of  $\alpha_T$  for the two jet categories used within this analysis, 2,3 and  $\geq 4$  jets,  
1279 is shown in the Figure 4.3, demonstrating the ability of the  $\alpha_T$  variable to discriminate  
1280 between multi jet events and EWK processes with genuine  $\cancel{E}_T$  in the final state.

1281 The  $\alpha_T$  requirement used within the search is chosen to be  $\alpha_T > 0.55$  to ensure that  
1282 the QCD multijet background is negligible even in the presence of moderate jet mis-  
1283 measurement. There still remains other effects which can cause multijet events to  
1284 artificially have a large  $\alpha_T$  value, which are discussed in detail in Section (4.2.2).



**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 multijet events (green dash-dotted line), EWK backgrounds with genuine  $\cancel{E}_T$  (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).

## 1285 4.2. Search Strategy

1286 The aim of the analysis presented in this thesis is to identify an excess of events in data  
 1287 over the SM background expectation in multi-jet final states and significant  $\cancel{E}_T$ . The  
 1288 essential suppression of the dominant QCD background for such a search is addressed by  
 1289 the  $\alpha_T$  variable described in the previous section. For estimation of the remaining EWK  
 1290 backgrounds, three independent data control samples are used to predict the different  
 1291 processes that compose the background :

- 1292 •  $\mu + \text{jets}$  to determine  $W + \text{jets}$ ,  $t\bar{t}$  and single top backgrounds,
- 1293 •  $\gamma + \text{jets}$  to determine the irreducible  $Z \rightarrow \nu\bar{\nu} + \text{jets}$  background,
- 1294 •  $\mu\mu + \text{jets}$  to determine the irreducible  $Z \rightarrow \nu\bar{\nu} + \text{jets}$  background.

1295 These control samples are chosen to both be rich in specific EWK processes, be free of  
 1296 QCD multi-jet events and to also be kinematically similar to the hadronic signal region  
 1297 that they are estimating the backgrounds of, see Section (4.2.3).

1298 To remain inclusive to a large range of possible **SUSY** models, the signal region is binned  
1299 in the following categories to allow for increased sensitivity in the interpretation of results  
1300 for different **SUSY** topologies:

1301 **Sensitivity to a range of SUSY mass splittings**

1302 The hadronic signal region is defined by  $H_T > 275$ , divided into eight bins in  $H_T$ .

- 1303 – Two bins of width 50 GeV in the range  $275 < H_T < 375$  GeV,  
1304 – five bins of width 100 GeV in the range  $375 < H_T < 875$  GeV,  
1305 – and a final open bin,  $H_T > 875$  GeV.

1306 The choice at low  $H_T$  is driven primarily by trigger constraints. The mass difference  
1307 between the **LSP** and the particle that it decays from is an important factor in the  
1308 amount of hadronic activity in the event.

1309 A large mass splitting will lead to hard high  $p_T$  jets which contribute to the  $H_T$   
1310 sum. From Figure 4.1 it can be seen that the **SM** background falls sharply at high  
1311  $H_T$  values, therefore a large number of  $H_T$  bins will lead to easier identification  
1312 of such signals. Conversely smaller mass splittings lead to softer jet  $p_T$ 's which will  
1313 subsequently fall into the lower  $H_T$  range.

1314 **Sensitivity to production method of SUSY particles**

1315 The production mechanism of any potential **SUSY** signal can lead to different event  
1316 topologies. One such way to discriminate between gluino ( $g\tilde{g}$  - “high multiplicity”),  
1317 and direct squark ( $q\tilde{q}$  - “low multiplicity”) induced production of **SUSY** particles is  
1318 realised through the number of reconstructed jets in the final state.

1319 The analysis is thus split into two jet categories : 2-3 jets ,  $\geq 4$  jets to give sensitivity  
1320 to both of these mechanisms.

1321 **Sensitivity to “Natural SUSY” via tagging jets from b-quarks**

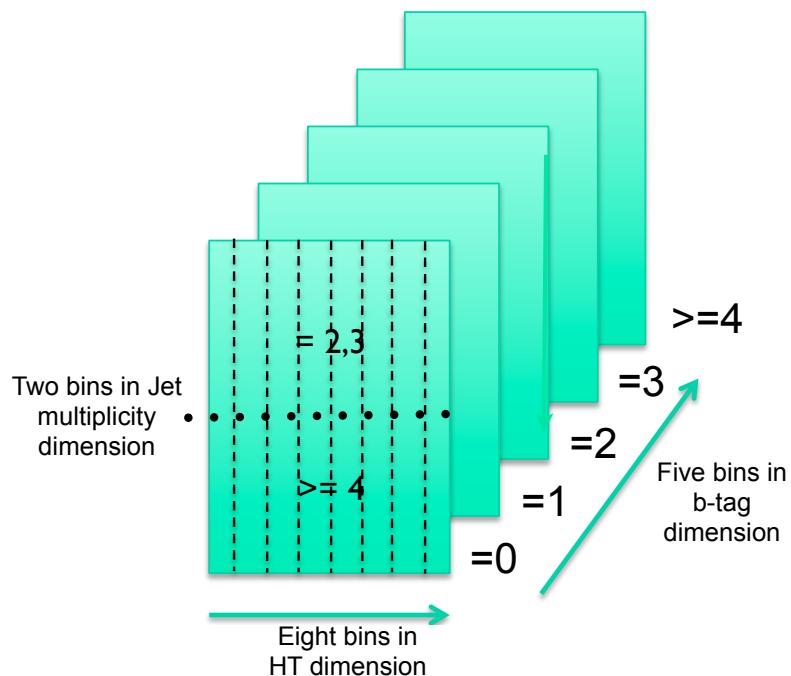
1322 Jets originating from bottom quarks (b-jets) are identified through vertices that  
1323 are displaced with respect to the primary interaction. The algorithm used to tag  
1324 b-jets is the Combined Secondary Vertex Medium Working Point (**CSV**) tagger,  
1325 described within Section (3.3.2). A cut is placed on the discriminator variable of  
1326  $> 0.679$ , leading to a gluon/light-quark mis-tag rate of 1% and a jet  $p_T$  dependant  
1327 b-tagging efficiency of 60-70% [64].

Natural SUSY models would be characterised through final-state signatures rich in bottom quarks. A search relying on methods to identify jets originating from bottom quarks through b-tagging, will significantly improve the sensitivity to this class of signature.

This is achieved via the binning of events in the signal region according to the number of b-tagged jets reconstructed in each event, in the following: 0,1,2,3, $\geq 4$  b-tag categories . In the highest  $\geq 4$  b-tag category due to a limited number of expected signal and background, just three  $H_T$  bins are employed: 275-325 GeV, 325-375 GeV,  $\geq 375$  GeV.

This characterisation is identically mirrored in all control samples, with the information from all samples and b-tag categories used simultaneously in the likelihood model, see Section (4.8), in order to interpret the results in a coherent and powerful way.

- The combination of the  $H_T$ , jet multiplicity and b-tag categorisation of the signal region as described above, resultantly leads to 67 different bins in which the analysis is interpreted in, which 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>1344</sup> 4.2.1. Physics objects

<sup>1345</sup> The physics objects used in the analysis defined below, follow the recommendation of  
<sup>1346</sup> the various CMS Physics Object Groups (POGs).

<sup>1347</sup> • Jets

<sup>1348</sup> The jets used in this analysis are CaloJets, reconstructed as described in Section  
<sup>1349</sup> (3.3.1) using the anti- $k_T$  jet clustering algorithm.

<sup>1350</sup> To ensure the jet object falls within the calorimeter systems a pseudo-rapidity  
<sup>1351</sup> requirement of  $|\eta| < 3$  is applied. Each jet must pass a “loose” identification criteria  
<sup>1352</sup> to reject jets resulting from unphysical energy, the criteria of which are detailed in  
<sup>1353</sup> Table A.1 of Appendix A [75].

<sup>1354</sup> • Muons

<sup>1355</sup> Muons are selected in the  $\mu +$  jets and  $\mu\mu +$  jets control samples, and vetoed in  
<sup>1356</sup> the signal region. The same cut based identification criteria is applied to muons in  
<sup>1357</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 [77].
$\chi^2 < 10$	$\chi^2$ 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.

1358 Additionally muons are required to be within the acceptance of the muon tracking  
 1359 systems. For the muon control samples, trigger requirements necessitate a  $|\eta| <$   
 1360 2.1 for the selection of muons. In the signal region where muons are vetoed these  
 1361 conditions are relaxed to  $|\eta| < 2.5$  and a minimum threshold of  $p_T > 10$  GeV is  
 1362 required of muon objects.

1363 **• Photons**

1364 Photons are selected within the  $\gamma +$  jets control sample and vetoed in all other  
 1365 selections. Photons are identified in both cases according to the cut based criteria  
 1366 listed in Table 4.3 [78].

Variable	Definition
$H/E < 0.05$	The ratio of hadronic energy in the <b>HCAL</b> tower directly behind the <b>ECAL</b> super-cluster and the <b>ECAL</b> super-cluster itself.
$\sigma_{in\eta} < 0.011$	The log energy weighted width ( $\sigma$ ), of the extent of the shower in the $\eta$ dimension.
$R9 < 1.0$	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 [79], <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.

1367 Photon objects are also required to have a minimum momentum of  $p_T > 25$  GeV.

1368 **• Electrons**

1369 Electron identification is defined for veto purposes. They are selected according to  
 1370 the following cut-based criteria listed in Table 4.4, utilising PF-based isolation.

1371 Electrons are required to be identified at  $|\eta| < 2.5$ , with a minimum  $p_T > 10$  GeV  
 1372 threshold to ensure that the electron falls within the tracking system of the detector.

1373 **• Noise and  $E_T$  Filters**

1374 A series of Noise filters are applied to veto events which contain spurious non-physical  
 1375 jets that are not picked up by the jet id, and events which give large unphysical  $\cancel{E}_T$   
 1376 values. These filters are listed within Table A.3 of Appendix A.

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.

### **4.2.2. Event selection**

1377 The selection criteria for events within the analysis are detailed below. A set of common  
1378 cuts are applied to both signal (maximise acceptance to a range of SUSY signatures), and  
1380 control samples (retain similar jet kinematics for background predictions), with additional  
1381 selection cuts applied to each control sample to enrich the sample in a particular EWK  
1382 processes, see Section (4.2.3).

1383 The jets considered in the analysis are required to have a transverse momentum  $p_T > 50$   
1384 GeV, with a minimum of two jets required in the event. The highest  $E_T$  jet is required  
1385 to lie within the central tracker acceptance  $|\eta| < 2.5$ , and the two leading  $p_T$  jets must  
1386 each have  $p_T > 100$  GeV. Any event which has a jet with  $p_T > 50$  GeV that either fails  
1387 the “loose” identification criteria described in Section(4.2.1) or has  $|\eta| > 3.0$ , is rejected.  
1388 Similarly events in which an electron, muon or photon fails object identification but pass  
1389  $\eta$  and  $p_T$  restrictions are identified as an “odd” lepton/photon and the event is vetoed.

1390 At low  $H_T$ , the jet threshold requirements applied to be considered as part of the analysis  
 1391 and enter the  $H_T$  sum are scaled downwards. These are scaled down in order to not  
 1392 restrict phase space, preserving jet multiplicities and background admixture in the lower  
 1393  $H_T$  bins, as listed in Table 4.5.

$H_T$ 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 4.5.:** Jet thresholds used in the three  $H_T$  regions of the analysis.

1394 Within the signal region to suppress SM processes with genuine  $\cancel{E}_T$  from neutrinos,  
 1395 events containing isolated electrons or muons are vetoed. Furthermore to ensure a pure  
 1396 multi-jet topology, events are vetoed if an isolated photon is found with  $p_T > 25$  GeV.  
  
 1397 An  $\alpha_T$  requirement of  $> 0.55$  is required to reduce the QCD multi-jet background  
 1398 to a negligible amount. Finally additional cleaning cuts are applied to protect against  
 1399 pathological deficiencies such as reconstruction failures or severe energy mis-measurements  
 1400 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 used 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,

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

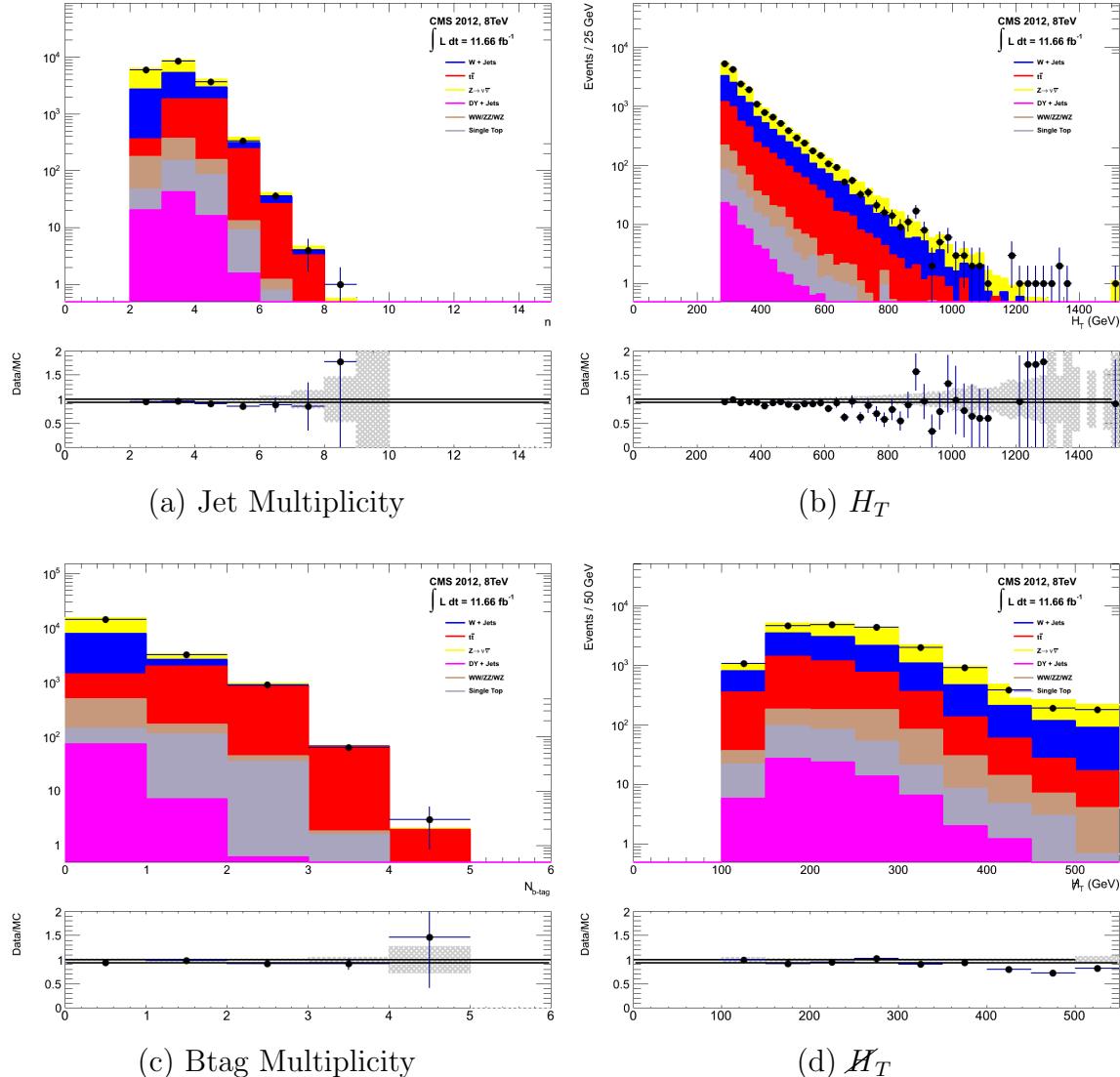
1401 • Fake  $\cancel{E}_T$  and  $\cancel{H}_T$  can arise due to significant jet mis-measurements cause by a small  
 1402 number of non-functioning ECAL regions. These regions absorb electromagnetic  
 1403 showers which are subsequently not added to the jet energy sum. To circumvent  
 1404 this problem the following procedure is employed : For each jet in the event, the  
 1405 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 is compatible with the hypothesis 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.

The distributions shown are presented for purely illustrative purposes, with the MC simulation itself not used in absolute term 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.

### 4.2.3. Control sample definition and background estimation

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

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

1428 All MC samples are normalised to the luminosity of the data samples,  $11.7 \text{ fb}^{-1}$ . Through  
1429 this method, “vanilla” predictions for the **SM** background in the signal region can be  
1430 made by considering separately the sum of the prediction from either the  $\mu + \text{jets}$  and  $\gamma$   
1431 + jets or  $\mu + \text{jets}$  and  $\mu\mu + \text{jets}$  samples. However the final background estimation from  
1432 which results are interpreted, is calculated via a fitting procedure defined formally by  
1433 the likelihood model described in Section (4.8).

1434 The sum of the expected yields from all MC processes, in each control sample enter the  
1435 denominator,  $N_{\text{MC}}^{\text{control}}$ , of the **TF** defined in Eq (4.7). However for the numerator,  $N_{\text{MC}}^{\text{signal}}$ ,  
1436 only the relevant processes that the control sample is used in estimating a background  
1437 for, enter into the **TF**.

1438 For the  $\mu + \text{jets}$  sample the simulated MC processes which enter the numerator of the  
1439 **TF** are,

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

1440 whilst for both the  $\mu\mu + \text{jets}$  and  $\gamma + \text{jets}$  samples the only MC process used in the  
1441 numerator is,

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

1442 The control samples and the **EWK** processes they are specifically tuned to select are  
1443 defined below, with distributions of key variables for each of the control samples shown  
1444 for illustrative purposes in Figures 4.6, 4.7 and 4.8. No requirement is placed upon  
1445 the number of b-tagged jets or jet multiplicity in the distributions shown. The MC  
1446 distributions highlight the background compositions of each control sample, where in  
1447 general, good agreement is observed between data and simulation, giving confidence

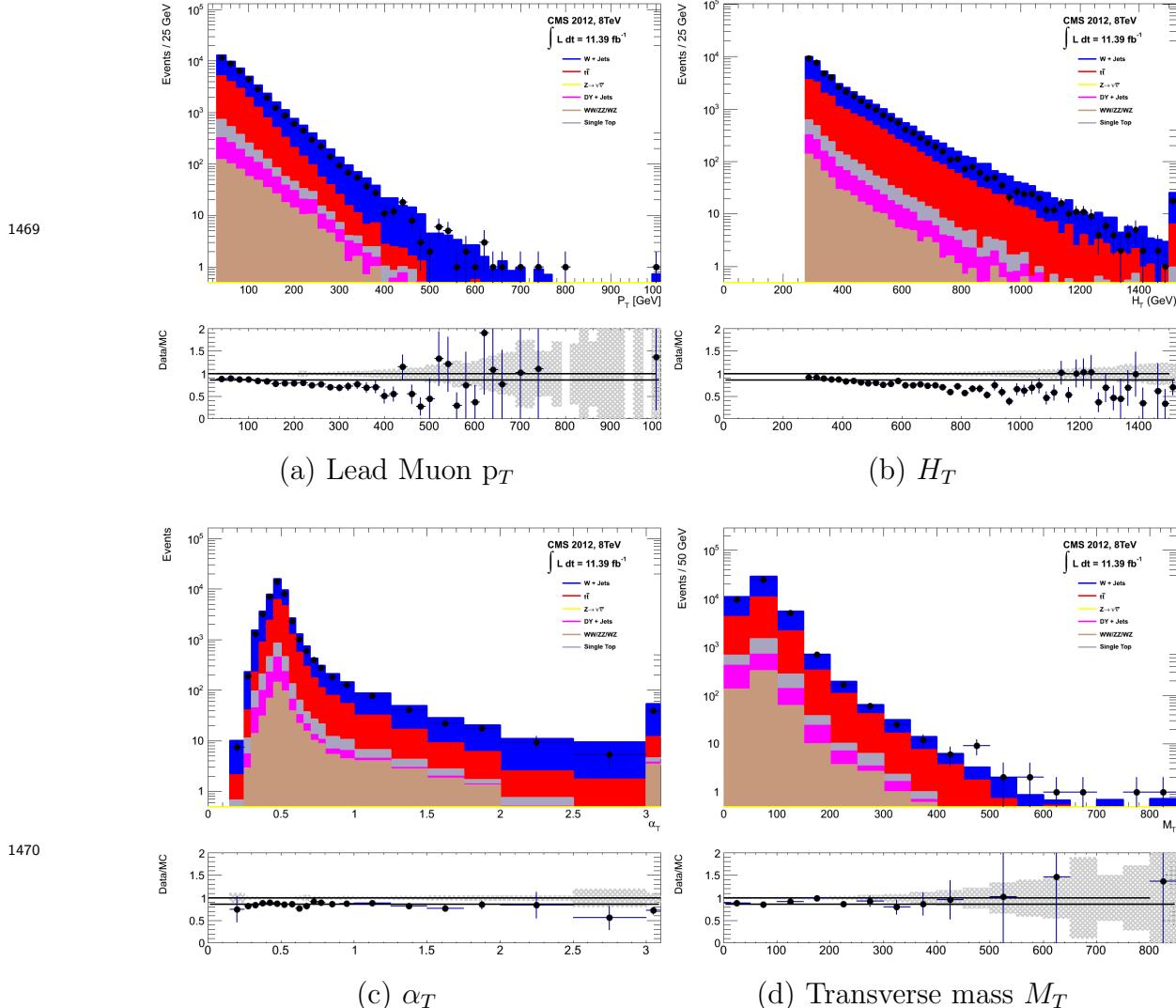
1448 that the samples are well understood. The contribution from QCD multi-jet events is  
1449 expected to be negligible :

1450 **The  $\mu +$  jets control sample**

1451 Events from  $W +$  jets and  $t\bar{t}$  processes enter into the hadronic signal sample due to  
1452 unidentified leptons from acceptance or threshold effects and hadronic tau decays.  
1453 These leptons originate from the decay of high  $p_T$   $W$  bosons.

1454 The control samples specifically identifies  $W \rightarrow \mu\bar{\nu}$  decays within the same phase-  
1455 space of the signal region, where the muon is subsequently ignored in the calculation  
1456 of event level variables, i.e.  $H_T$ ,  $\mathcal{H}_T$ ,  $\alpha_T$ . All kinematic jet-based cuts are identical  
1457 to those applied in the hadronic search region detailed in Section (4.2.2), with the  
1458 same  $H_T$ , jet multiplicity and b-jet multiplicity binning described above.

- 1459 – Muons originating from  $W$  boson decays are selected by requiring one tightly  
1460 isolated muon defined in Table 4.2, with a  $p_T > 30$  GeV and  $|\eta| < 2.1$ . Both of  
1461 these threshold arise from trigger restrictions.
- 1462 – The transverse mass of the  $W$  candidate must satisfy  $M_T(\mu, \cancel{E}_T) < 30$  GeV ( to  
1463 suppress QCD multi-jet events).
- 1464 – Events which contain a jet overlapping with a muon  $\Delta R(\mu, \text{jet}) < 0.5$  are vetoed  
1465 to remove events from muons produced as part of a jet’s hadronisation process.
- 1466 – Events containing a second muon candidate which has failed id, but passed  $p_T$   
1467 and  $|\eta|$  requirements, are checked to have an invariant mass that satisfies  $m_Z - 25 < M_{\mu_1\mu_2} > m_Z + 25$ , thus removing  $Z \rightarrow \mu\mu$  contamination.



**Figure 4.6.:** Data/MC comparisons of key variables for the  $\mu + \text{jets}$  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 $\mu\mu$ + jets control sample

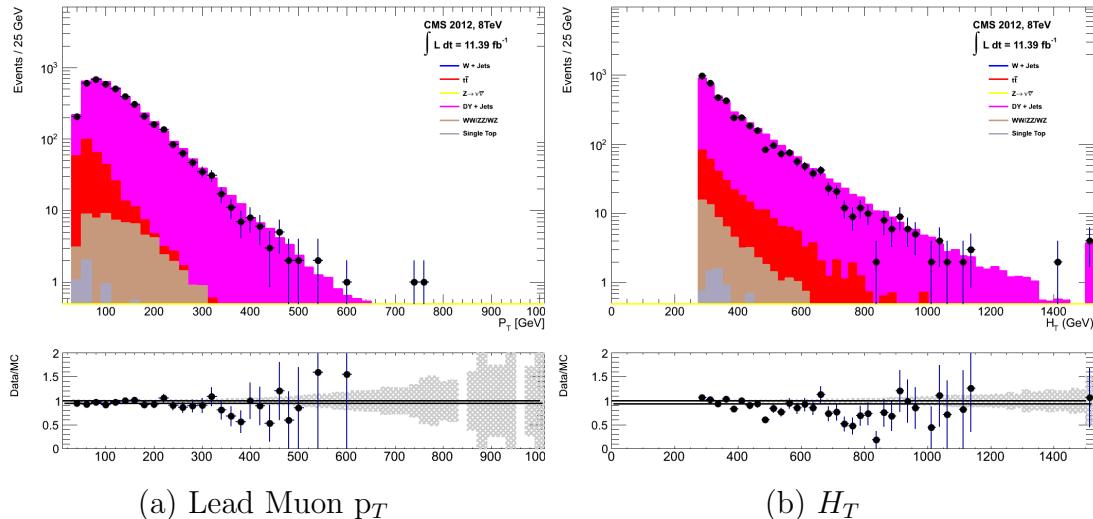
The  $Z \rightarrow \nu\bar{\nu} + \text{jets}$  background enters into the signal region from genuine  $\cancel{E}_T$  from the escaping neutrinos. This background is estimated using two control samples, the first of which is the  $Z \rightarrow \mu\bar{\mu} + \text{jets}$  process, which possesses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

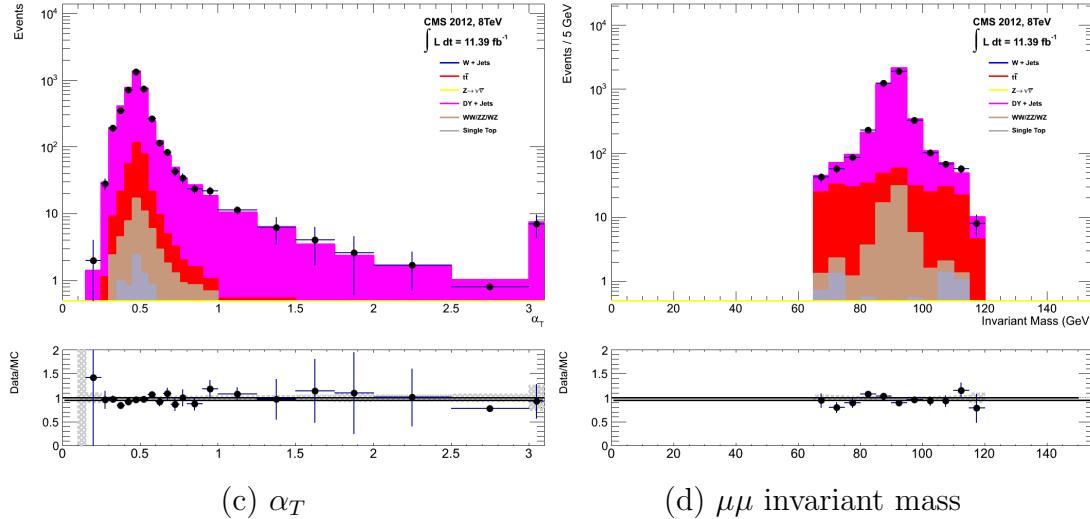
The same acceptance requirements as the  $\mu + \text{jets}$  selection for muons is applied, as defined in Table 4.2. Muons in the event are ignored for the purpose of the

calculation of event level variables. Kinematic jet-based cuts and phase space binning identical to the hadronic search region are also applied.

- Muons origination 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  $m_Z - 25 > M_{\mu_1\mu_2} < m_Z + 25$ .

The  $\mu\mu + \text{jets}$  sample is used to make predictions in the signal region in 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 instead used to determine the  $Z \rightarrow \nu\bar{\nu}$  estimation.



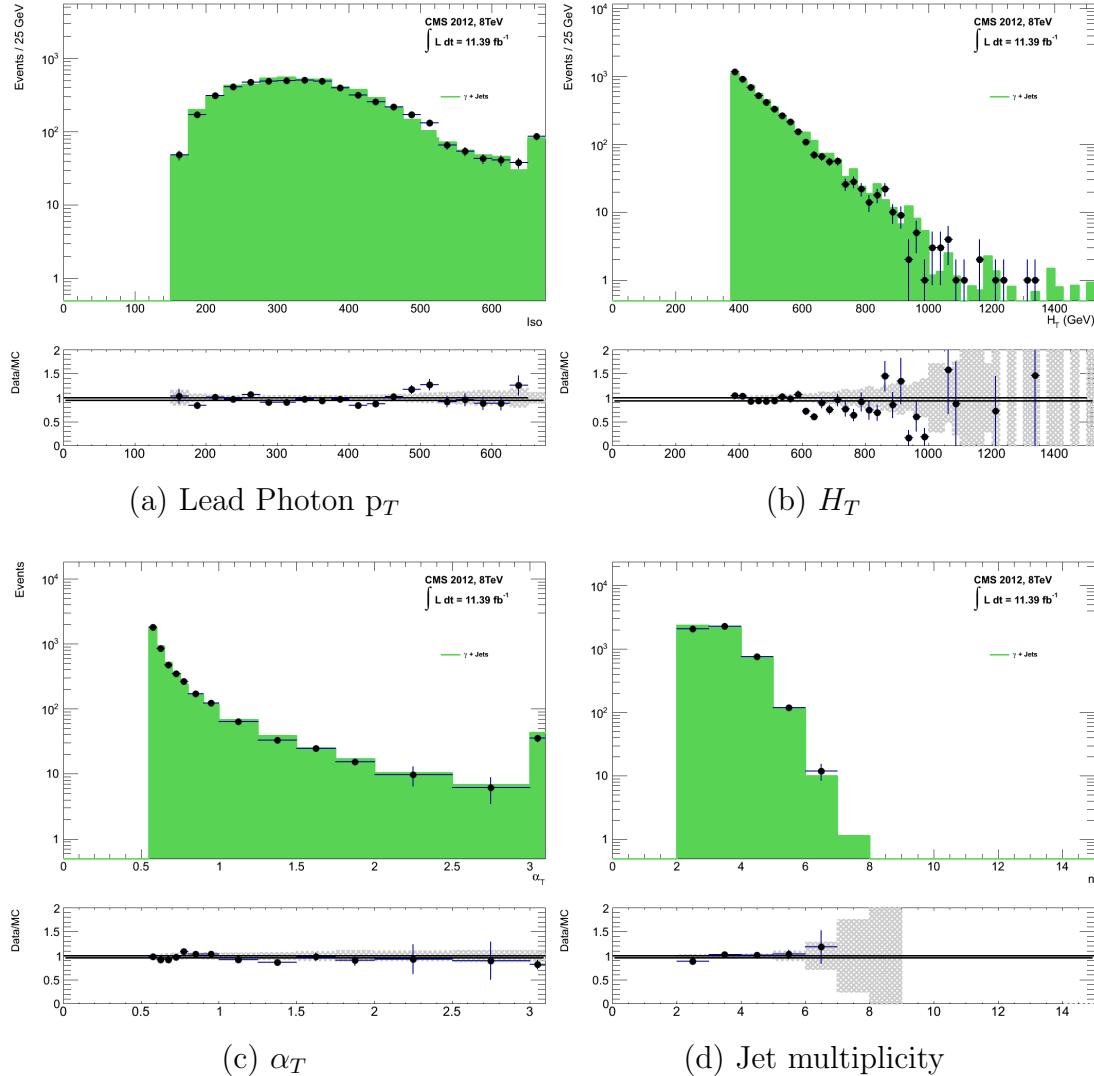


**Figure 4.7.:** Data/MC comparisons of key variables for the  $\mu\mu + \text{jets}$  selection, following the application of selection criteria and the requirements that  $H_T > 275$  GeV. Bands represent the uncertainties due to the statistical size of the MC samples. No requirement is made upon the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

## The $\gamma$ + jets control sample

The  $Z \rightarrow \nu\bar{\nu} + \text{jets}$  background is also estimated from a  $\gamma + \text{jets}$  control sample, which possesses a larger cross section and kinematic properties similar to those of  $Z \rightarrow \mu\bar{\mu}$  events where the photon is ignored [80][81]. 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  $\gamma + \text{jets}$  events.
  - Given that the photon is ignored, this control sample can only be applied in the  $H_T$  region  $> 375$  GeV, due to the trigger thresholds on the minimum  $p_T$  of the photon, and the  $H_T$  requirement of an  $\alpha_T > 0.55$  cut from Equation (4.5).



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

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

However in the case of the  $\mu + \text{jets}$  and  $\mu\mu + \text{jets}$  samples, the  $\alpha_T$  requirement is relaxed in the selection criteria of these samples. This is made possible as contamination from QCD multi-jet events is suppressed to a negligible level by the other kinematic selection criteria within the two control samples, to select pure EWK processes. Thus in this way,

1517 the acceptance of the two muon control samples can be significantly increased, which  
1518 simultaneously improves their predictive power and further reduces the effect of any  
1519 potential signal contamination.

1520 The modelling of the  $\alpha_T$  variable is probed through a dedicated set of closure tests,  
1521 described in Section (4.6), which demonstrate that the different  $\alpha_T$  acceptances for the  
1522 control and signal samples have no significant systematic bias on the prediction.

1523 **4.2.4. Estimating the QCD multi-jet background**

1524 A negligible background from QCD multi-jet events within the hadronic signal region  
1525 is expected due to the selection requirement, and additional cleaning filters applied.  
1526 However a conservative approach is still adopted and the likelihood model, see Section  
1527 (4.8.2), is given the freedom to estimate any potential QCD multi-jet contamination.

1528 Any potential contamination can be identified through the variable  $R_{\alpha_T}$ , defined as the  
1529 ratio of events above and below the  $\alpha_T$  threshold value used in the analysis. This is  
1530 modelled by a  $H_T$  dependant falling exponential function which takes the form,

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

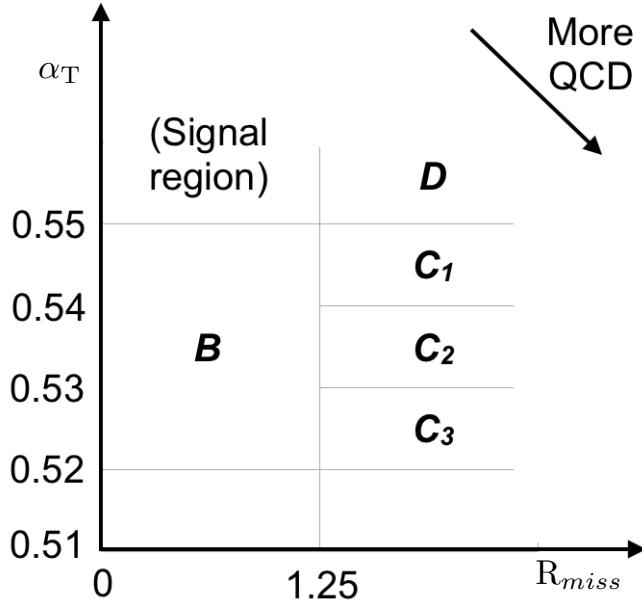
1531 where the parameters  $A$  and  $k_{QCD}$  are the normalisation and exponential decay constants  
1532 respectively.

1533 For QCD event topologies this exponential behaviour is expected as a function of  $H_T$  for  
1534 several reasons. The improvement of jet energy resolution at higher  $H_T$  due to higher  $p_T$   
1535 jets leads to a narrower peaked distribution, causing  $R_{\alpha_T}$  to fall. Similarly at higher  $H_T$   
1536 values  $> 375$  GeV, the jet multiplicity rises slowly with  $H_T$ . As shown in Figure 4.3, at  
1537 higher jet multiplicities, the result of the combinatorics used in the determination of  $\alpha_T$ ,  
1538 also lead to a narrower  $\alpha_T$  distribution.

1539 The value of the decay constant  $k_{QCD}$  is constrained via measurements within data  
1540 sidebands to the signal region. This is also done to validate the falling exponential  
1541 assumption for QCD multi-jet topologies. The sidebands are enriched in QCD multi-jet  
1542 background and defined as regions where  $\alpha_T$  is relaxed or that the  $R_{miss}$  cut is inverted.

1543 Figure 4.9 depicts the definition of these data sidebands used to constrain the value of  
1544  $k_{QCD}$ .

1545



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

1546 The fits to determine the value of  $k_{QCD}$  are shown in Appendix (C.1), for which the best  
1547 fit value obtained from sideband region B is determined to be  $k_{QCD} = 2.96 \pm 0.64 \times 10^{-2}$   
1548  $\text{GeV}^{-1}$ .

1549 The best fit values of the remaining three C sideband regions are used to estimate  
1550 the systematic uncertainty on the central value obtained from sideband region B. The  
1551 variation of these measured values is used to determine the error on the determined  
1552 central value, and is calculated to be  $1.31 \pm 0.26 \times 10^{-2} \text{ GeV}^{-1}$ . This relative error of  $\sim$   
1553 20% gives an estimate of the systematic uncertainty of the measurement to be applied to  
1554  $k_{QCD}$ .

1555 Finally the same procedure is performed for sideband region D to establish that the  
1556 value of  $k_{QCD}$  extracted from a lower  $\alpha_T$  slice can be applied to the signal region  $\alpha_T >$   
1557 0.55. The likelihood fit is performed across all  $H_T$  bins within the QCD enriched region  
1558 with no constraint applied to  $k_{QCD}$ . The resulting best fit value for  $k_{QCD}$  shows good  
1559 agreement between that and the weighted mean determined from the three C sidebands  
1560 regions. This demonstrates that the assumption of using the central value determined  
1561 from sideband region B, to provide an unbiased estimator for  $k_{QCD}$  in the signal region  
1562 ( $\alpha_T > 0.55$ ) is valid.

1563 Table 4.6, summarises the best fit  $k_{QCD}$  values determined for each of the sideband  
1564 regions to the signal region.

Sideband region	$k_{QCD} (\times 10^{-2} \text{GeV}^{-1})$	$p$ -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 4.6.:** Best fit values for the parameters  $k_{QCD}$  obtained from sideband regions B,C<sub>1</sub>,C<sub>2</sub>,C<sub>3</sub>. The weighted mean is determined from the three measurements made within sideband region C. The maximum likelihood value of  $k_{QCD}$  given by the simultaneous fit using sideband region D. Quotes errors are statistical only.

### 1565 4.3. Trigger Strategy

1566 A cross trigger based on the quantities  $H_T$  and  $\alpha_T$ , labelled is used with varying thresholds  
1567 across  $H_T$  bins to record the events used in the hadronic signal region. The  $\alpha_T$  legs of the  
1568  $HT\_alphaT$  triggers used in the analysis are chosen to fully suppress QCD multi-jet events,  
1569 whilst maintaining a sustainable trigger rate. To further maintain an acceptable rate for  
1570 these analysis specific triggers, only calorimeter information is used in the reconstruction  
1571 of the  $H_T$  sum, leading to the necessity for Calo jets to be used within the analysis.

1572 A single object prescaled  $H_T$  trigger is used to collect events for the hadronic control  
1573 region described above in Section (4.2.4).

1574 The performance of the  $\alpha_T$  and  $H_T$  triggers used to collect data for the signal and  
1575 hadronic control region is measured with respect to a reference sample collected using the  
1576 muon system. This allows measurement of both the Level 1 seed and higher level triggers  
1577 simultaneously, as the reference sample is collected independent of any jet requirements.

1578 The selection for the trigger efficiency measurement is identical to that described in  
1579 Section (4.2.2), with the requirement of exactly one well identified muon with  $p_T > 30$   
1580 GeV which is subsequently ignored.

1581 The efficiencies measured for the  $HT\_alphaT$  triggers in bins individual  $H_T$  and  $\alpha_T$  legs, is  
1582 summarised in Table 4.7.

$H_T$ 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 4.7.:** 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.

1583 Data for the control samples of the analysis, detailed in Section (4.2.3), are collected  
 1584 using single object photon trigger for the  $\gamma +$  jets sample, and a single object muon  
 1585 trigger for both the  $\mu +$  jets and  $\mu\mu +$  jets control samples. The photon trigger is  
 1586 measured to be full efficient for the threshold  $p_T^{photon} > 150 GeV$ , whilst the single muon  
 1587 efficiency satisfying  $p_T^{muon} > 30 GeV$  is measured to have an efficiency of  $(88 \pm 2)\%$  that  
 1588 is independent of  $H_T$ . In the case of the  $\mu\mu +$  jets control sample, the efficiency is  
 1589 measured to be  $(95 \pm 2)\%$  for the lowest  $H_T$  bin, rising to  $(98 \pm 2)\%$  for the highest  $H_T$   
 1590 bin.

## 1591 4.4. Measuring MC Normalisation Factors via $H_T$ 1592 Sidebands

1593 The theoretical cross sections of different SM processes at Next to Next Leading Order  
 1594 (NNLO) and the number of MC simulated events generated for that particular process,  
 1595 is typically used to determine the appropriate normalisation for a MC sample. However  
 1596 within the particular high- $H_T$  and high- $\cancel{E}_T$  corners of kinematic phase space probed  
 1597 within this search, the theoretical cross sections for various processes are far less well  
 1598 understood.

1599 To mitigate the problem of theoretical uncertainties and arbitrary choices of cross sections,  
 1600 the normalisation of MC samples used in the analysis are determined through the use  
 1601 data sidebands. The sidebands are used to calculate sample specific correct factors  
 1602 (k-factors) that are appropriate for the  $H_T$ - $\cancel{E}_T$  phase space covered by this analysis.

1603 They are defined within the  $\mu +$  jets and  $\mu\mu +$  jets control sample, by the region  $200 <$   
 1604  $H_T < 275$ , using the same jet  $p_T$  thresholds as the adjacent first analysis bin. Individual  
 1605 EWK processes are isolated within each of these control samples via requirements on

1606 jet multiplicity and the requirement on b-tags, summarised in Table 4.8. The purity of  
1607 the samples are typically  $> 90\%$  with any residual contamination corrected for. The  
1608 resultant k-factor for each process is determined by then taking ratio of the data yield  
1609 over the MC expectation in the sideband. Subsequently these k-factors are then applied  
1610 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 4.8.:** 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.

## 1611 4.5. Determining MC Simulation Yields with 1612 Higher Statistical Precision

1613 Reconstructing events from EWK processes with many b-tagged jets ( $\geq 3$ ),  $n_b^{reco}$ , is largely  
1614 driven by the mis-tagging of light jets within the event. This is clear when considering  
1615 the main EWK backgrounds in the analysis, such as  $t\bar{t} + \text{jets}$  events, which typically  
1616 contain two b-flavoured jets from the decay of the top quarks, whilst W + jets and  
1617  $Z \rightarrow \mu\mu + \text{jets}$  events will typically contain no b-flavoured jets.

1618 When the expectation for the number of  $n_b^{reco}$  is taken directly from simulation, the  
1619 statistical uncertainty at large b-tag multiplicities becomes relatively large. In order to  
1620 reduce this uncertainty one approach is to use the information encoded throughout all  
1621 events in the simulation sample, to measure each of the four ingredients:

- 1622 1. the b-tagging efficiency in the event selection,
- 1623 2. the charm-tagging efficiency in the event selection
- 1624 3. the mis-tagging rate in the event selection,
- 1625 4. the underlying flavour distribution of the jets in the events,

1626 that determine the  $n_b^{reco}$  distribution of the process being measured. This method allows  
1627 the determination of higher b-tag multiplicities to a higher degree of accuracy reducing

<sub>1628</sub> the statical uncertainties of the MC which enter into the TF's. For the discussion that  
<sub>1629</sub> follows, these predictions are determined on average (i.e not on an event-by-event basis),  
<sub>1630</sub> and is known as the formula method.

### <sub>1631</sub> 4.5.1. The formula method

<sub>1632</sub> The assigning of jet flavours to reconstruction level jets in simulation is achieved via an  
<sub>1633</sub> algorithmic method defined as:

- <sub>1634</sub> • Try to find the parton that most likely determines the properties of the jet and  
<sub>1635</sub> assign that flavour as true flavour,
- <sub>1636</sub> • Here, the “final state” partons (after showering, radiation) are analysed (also within  
<sub>1637</sub>  $\Delta R < 0.3$  of reconstructed jet cone),
- <sub>1638</sub> • Jets from radiation are matched with full efficiency,
- <sub>1639</sub> • If there is a b/c flavoured parton within the jet cone: label as b/c flavoured jet,
- <sub>1640</sub> • Otherwise: assign flavour of the hardest parton.

<sub>1641</sub> Within each individual MC process and each  $H_T$ - $n_{jet}$  bin in the analysis, the  $n_b^{reco}$   
<sub>1642</sub> distribution is constructed in the following way:

<sub>1643</sub> Let  $N(n_b^{gen}, n_c^{gen}, n_q^{gen})$  represent the yield in simulation of events with  $b$  underlying  
<sub>1644</sub> b-quarks,  $c$  underlying c-quarks and  $q$  underlying light quarks which are matched to  
<sub>1645</sub> reconstructed jets as detailed above. Light quarks defined as those which originate from  
<sub>1646</sub> a  $u$ ,  $d$ ,  $s$ ,  $g$  and  $\tau$  jets, which having similar mis-tagging rates are grouped together.  
<sub>1647</sub> Similarly  $\epsilon$ ,  $\beta$  and  $m$  represent the measured b-tagging, c-tagging and mis-tagging  
<sub>1648</sub> efficiency averaged over all the jets within that particular analysis bin.

<sub>1649</sub> Using this information the expected  $n_b^{reco}$  distribution can be analytically calculated  
<sub>1650</sub> using the formula :

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

1651 with  $N(n_b)$  representing the yield where  $n_b$  jets have been b-tagged. The variables  
1652  $n_b^{tag}$ ,  $n_c^{tag}$  and  $n_q^{tag}$  represent the number of times that a particular jet flavour results  
1653 in a b-tagged jet, and  $P(n_b^{tag}, n_b^{gen}, \epsilon)$ ,  $P(n_c^{tag}, n_c^{gen}, \beta)$  and  $P(n_q^{tag}, n_q^{gen}, m)$  represent the  
1654 binomial probabilities for that to happen.

1655 This approach ultimately results in a more precise  $n_b^{reco}$  distribution prediction as  
1656 information from throughout the entire MC sample is used to estimate the underlying  
1657  $n_b^{reco}$  distribution.

### 1658 4.5.2. Establishing proof of principle

1659 In order to validate the procedure, the predictions obtained from the formula method  
1660 summarised in Equation (4.11), are compared directly to those obtained directly from  
1661 simulation. Resultantly no simulation to data correction factors are applied when making  
1662 this comparison

1663 This sanity check for the  $\mu +$  jets control sample is presented in Table 4.9, for all  $n_b^{reco}$   
1664 and  $H_T$  bins with no requirement placed upon the jet multiplicity of the events.

1665 It can be seen as expected, that there is good consistency between the results determined  
1666 via the formula method and ‘raw’ simulation yields. Similarly the power of this approach  
1667 can be seen in the reduction of this statistical error in the prediction across all  $H_T$  and  
1668  $n_b^{reco}$  bins. In particular the statistical uncertainty is reduced by several factors in the  
1669 highest  $n_b^{reco} \geq 4$  category.

### 1670 4.5.3. Correcting measured efficiencies in simulation to data

1671 As detailed in Section (3.3.2), it is necessary for certain  $p_T$  and  $\eta$  dependant corrections,  
1672 to be applied to both the b-tagging efficiency and mis-tagging rates in order correct the  
1673 efficiencies from simulation to the distributions seen in data. These corrections factors  
1674 are considered when determining the simulation yields for each selection, which are used  
1675 to construct the TF’s of the analysis.

1676 Each of the corrections factors for the b, c and light flavoured jets come with an associated  
1677 systematic uncertainty. The uncertainties across different jet  $p_T$  and  $\eta$  bins, are considered  
1678 as fully correlated. When computing the magnitude of the effect of this systematic  
1679 uncertainty on the TF’s of the analysis, the scale factors are therefore scaled up/down

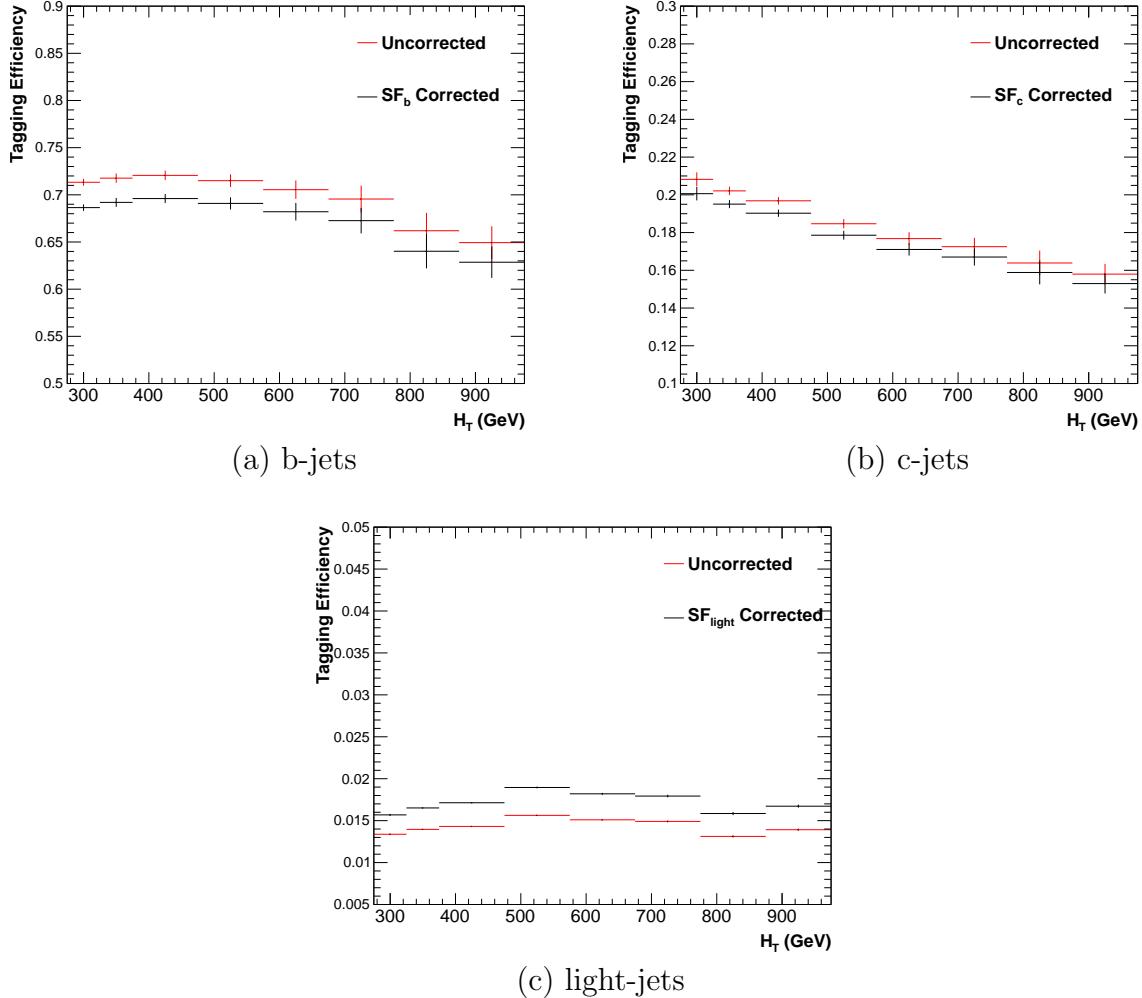
$H_T$ Bin (GeV)	275–325	325–375	375–475	475–575
Formula $n_b = 0$	12632.66 $\pm$ 195.48	6696.08 $\pm$ 82.59	6368.96 $\pm$ 75.34	2906.27 $\pm$ 39.65
Vanilla $n_b = 0$	12612.95 $\pm$ 198.68	6687.97 $\pm$ 83.78	6359.27 $\pm$ 76.50	2898.27 $\pm$ 36.89
Formula $n_b = 1$	4068.09 $\pm$ 45.71	2272.76 $\pm$ 26.14	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 4.9.:** 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.

1680 simultaneously within  $H_T$  bin of the analysis for each of the  $SF_{b,c,light}$  scale factors. The  
 1681 magnitude of this correction is shown for each  $H_T$  bin within Figure 4.10.

1682 Varying the scale factor corrections by their systematic uncertainty will change the  
 1683 absolute yields within each  $n_b^{reco}$  bin of all selections. However, ultimately it is the change  
 1684 in the TF's which influences the final background prediction from each of the control  
 1685 samples. The magnitude of these change is shown in Table 4.10, highlighting the absolute  
 1686 change in each TF, constructed from when the  $\mu +$  jets control sample is used to predict  
 1687 the entire hadronic signal region background.

1688 It can be seen that the TF's are found to be relatively insensitive to the systematic  
 1689 uncertainty of the b-tag scale factors (showing typically less than  $\sim 2\%$  change). This can



**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  $SF_{b,c,light}$  correction factors (black) are applied.

1690 be accounted for by the similar composition of the signal and control sample backgrounds,  
 1691 such that any change in the underlying  $n_b^{reco}$  distribution will be reflected in both signal  
 1692 and control regions and cancel out in the TF.

1693 Any overall systematic effect on the overall background prediction of the analysis from  
 1694 these b-tag scale factor uncertainties is incorporated within the data driven systematics  
 1695 introduced in the following section.

$n_b^{reco}$	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 4.10.:** The absolute change in the **TF**’s used to predict the entire signal region **SM** background, using the  $\mu +$  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)

## 1696 4.6. Systematic Uncertainties on Transfer Factors

1697 Since the **TF**’s used to establish the background prediction are obtained from simulation,  
 1698 an appropriate systematic uncertainty is assigned to each factor to account for theoretical  
 1699 uncertainties [82] and limitations in the simulation modelling of event kinematics and  
 1700 instrumental effects.

1701 The magnitudes of these systematic uncertainties are established through a set of data  
 1702 driven method, in which the three independent control samples of the analysis ( $\mu +$  jets,  
 1703  $\mu\mu +$  jets,  $\gamma +$  jets) are used to in a series of closure tests. The yields from one of these  
 1704 control samples, along with the corresponding **TF** obtained from simulation, are used to  
 1705 predict the yields in another control sample, using the same method of establishing a  
 1706 background prediction for the signal region as described in Section (4.2.3).

1707 The level of agreement between the predicted and observed yields is expressed as the  
 1708 ratio

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

1709 while considering only the statistical uncertainties on  $N_{pred}$ , the prediction, and  $N_{obs}$ , the  
 1710 observation. No systematic uncertainty is assigned to the prediction, and resultantly the  
 1711 level of closure is defined by the statistical significance of a deviation from the ratio from  
 1712 zero.

1713 This ratio is measured for each  $H_T$  bin in the analysis, allowing these closure tests to be  
 1714 sensitive to both the presence of any significant biases or any possible  $H_T$  dependence on  
 1715 the level of closure.

1716 Eight sets of closure tests are defined between the three data control samples, conducted  
 1717 independently between the two jet multiplicity ( $2 \leq n_{jets} \leq 3$ ,  $n_{jet} \geq 4$ ) bins. Each of  
 1718 these tests are specifically chosen to probe each of the different key ingredients of the  
 1719 simulation modelling that can affect the background prediction.

1720 Each of the different modelling components and the relevant closure tests are described  
 1721 below :

### 1722 **$\alpha_T$ modelling**

1723 The modelling of the  $\alpha_T$  distribution in genuine  $\cancel{E}_T$  events is probed with the  $\mu +$   
 1724 jets control sample. This test is important to verify the approach of remove the  $\alpha_T$   
 1725  $> 0.55$  requirement from the  $\mu +$  jets and  $\mu\mu +$  jets samples to increase the precision  
 1726 of the background prediction. The test uses the  $\mu +$  jets sample without an  $\alpha_T$  cut  
 1727 to make a prediction into the  $\mu +$  jets sample defined with the requirement  $\alpha_T >$   
 1728 0.55.

### 1729 **Background admixture**

1730 The sensitivity of the translation factors to the relative admixture of events from  
 1731  $W +$  jets and  $t\bar{t}$  processes is probed by two closure tests. These tests represent  
 1732 an extremely conservative approach as the admixture of the background remains  
 1733 similar between the  $\mu +$  jets sample and the signal region, contrary to the defined  
 1734 closure tests which make predictions between two very different admixtures of  $W +$   
 1735 jets and  $t\bar{t}$  events.

1736 Within the  $\mu + \text{jets}$  sample, a  $W$  boson enriched sub-sample ( $n_b = 0$ ) is used to  
1737 predict yields in a  $t\bar{t}$  enriched sub-sample ( $n_b = 1$ ). Similarly the  
1738  $t\bar{t}$  enriched sub-sample ( $n_b = 1$ ) is also used to predict yields for a further enriched  
1739  $t\bar{t}$  sub-sample ( $n_b = 2$ ).

1740 Similarly a further closure test probes the relative contribution of  $Z + \text{jets}$  to  $W + \text{jets}$   
1741 and  $t\bar{t}$  events, through the use of the  $\mu + \text{jets}$  sample to predict yields for the  $\mu\mu +$   
1742  $\text{jets}$  control sample. This closure test, also at some level probes the muon trigger  
1743 and reconstruction efficiencies, given that exactly one and two muons are required  
1744 by the different selections.

### 1745 Consistency between control samples

1746 An important consistency check between the  $\mu\mu + \text{jets}$  jets and  $\gamma + \text{jets}$ , which are  
1747 both used in the prediction of the  $Z \rightarrow \nu\bar{\nu}$  in the signal region, is measured by using  
1748 the  $\gamma + \text{jets}$  sample to predict yields for the  $\mu\mu + \text{jets}$  control sample.

### 1749 Modelling of jet multiplicity

1750 The simulation modelling of the jet multiplicity within each control sample is  
1751 important due to the exclusive jet multiplicity binning within the analysis. This is  
1752 probed via the use of each of the three control samples to independently predict  
1753 from the lower jet multiplicity category  $2 \leq n_{jet} \leq 3$ , to the high jet category  $\geq 4$ .

1754 For the case of the  $\mu + \text{jets}$  and  $\mu\mu + \text{jets}$  control samples this test is also a  
1755 further probe of the admixture between  $W + \text{jets}/Z + \text{jets}$  and  $t\bar{t}$ .

1756 To test for the assumption that no  $H_T$  dependences exist within the background predic-  
1757 tions of the analysis, the first five closure tests defined above are taken, with zeroeth and  
1758 first order polynomial fits are applied to each. This is summarised in Table 4.11 and  
1759 Table 4.12 which show the results for both the  $2 \leq n_{jet} \leq 3$  and  $\geq 4$  jet multiplicity bins  
1760 respectively.

1761 Table 4.13 shows the same fits applied to the three closure tests that probe the modelling  
1762 between the different  $n_{jet}$  bins. The best fit value and its uncertainty is listed for each  
1763 set of closure tests in all three tables, along with the p-value of the constant and linear  
1764 fits applied.

1765 The best fit value for the constant parameter is indicative of the level of closure, averaged  
1766 across the full range of  $H_T$  bins in the analysis, and the p-value an indicator of any  
1767 significant dependence on  $H_T$  within the closure tests. The best fit values of all the tests

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
$0 \text{ b-jets} \rightarrow 1 \text{ b-jet } (\mu + \text{jets})$	Square	$0.07 \pm 0.02$	0.98	$-1.6 \pm 1.6$	1.00
$1 \text{ b-jets} \rightarrow 2 \text{ b-jet } (\mu + \text{jets})$	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 4.11.:** A summary of the results obtained from fits of zeroeth order polynomials (i.e. a constant) to five sets of closure tests performed in the  $2 \leq n_{jet} \leq 3$  bin. The final two columns show the best fit value for the slope obtained when performing a linear fit and the p-value for the linear 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
$0 \text{ b-jets} \rightarrow 1 \text{ b-jet } (\mu + \text{jets})$	Square	$-0.03 \pm 0.03$	0.55	$-1.0 \pm 1.9$	0.47
$1 \text{ b-jets} \rightarrow 2 \text{ b-jet } (\mu + \text{jets})$	Triangle	$-0.02 \pm 0.03$	0.39	$1.1 \pm 2.2$	0.31
$\mu + \text{jets} \rightarrow \mu\mu + \text{jets}$	Cross	$0.08 \pm 0.07$	0.08	$4.8 \pm 4.3$	0.07
$\mu\mu + \text{jets} \rightarrow \gamma + \text{jets}$	Star	$-0.03 \pm 0.10$	0.72	$-4.0 \pm 7.0$	0.64

**Table 4.12.:** A summary of the results obtained from fits of zeroeth order polynomials (i.e. a constant) to five sets of closure tests performed in the  $n_{jet} \geq q$  bin. The final two columns show the best fit value for the slope obtained when performing a linear fit and the p-value for the linear fit.

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

**Table 4.13.:** A summary of the results obtained from fits of zeroeth order polynomials (i.e. a constant) to five sets of closure tests performed in the  $2 \leq n_{jet} \leq 3$  bin. The final two columns show the best fit value for the slope obtained when performing a linear fit and the p-value for the linear fit.

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

Within Table 4.13, there exists one test that does not satisfy the above statement, which is the  $2 \leq n_{jet} \leq 3 \rightarrow n_{jet} \geq 4$  test using the  $\mu + \text{jets}$  control sample. The low p-value can be largely attributed to an outlier in the  $675 < H_T < 775$  GeV bin, rather than any significant trend in  $H_T$ . Removing this single outlier from the constant fit performed, 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 modified fit results are included within Table 4.13 .

1776 In addition the best fit values for the slope terms of the linear fits in all three tables are  
1777 of the order  $10^{-4}$ , which corresponds to a percent level change per 100 GeV. However in  
1778 all cases, the best fit values are fully compatible with zero (within  $1\sigma$ ) once again with  
1779 the exception detailed above, indicating that the level of closure is  $H_T$  independent.

#### 1780 4.6.1. Determining systematic uncertainties from closure tests

1781 Once it has been established that no significant bias or trend has been exist within  
1782 the closure tests, systematic uncertainties are determined. The statistical precision  
1783 of the closure tests is considered a suitable benchmark for determining the systematic  
1784 uncertainties that are assigned to the TF's, which are propagated through to the likelihood  
1785 fit.

1786 The systematic uncertainty band is split into five separate regions of  $H_T$ . Within each  
1787 region the square root of the sample variance,  $\sigma^2$ , is taken over the eight closure tests to  
1788 determine the systematic uncertainties to be applied within that region.

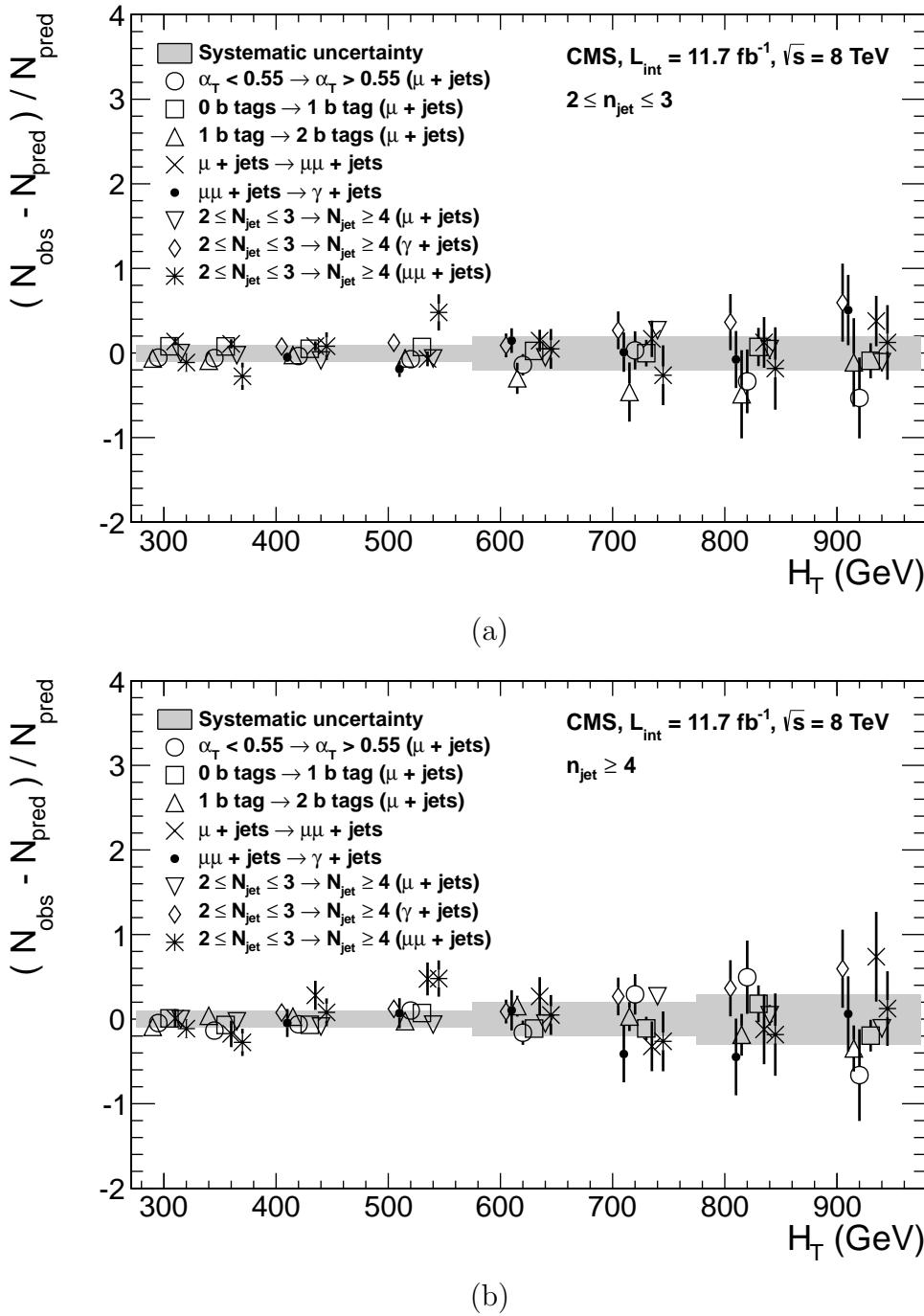
1789 Using this procedure the systematic uncertainties for each region are calculated and are  
1790 shown in Table 4.14, with the systematic uncertainty to be used in the likelihood model  
1791 conservatively rounded up to the nearest decile, shown in brackets.

$H_T$ band (GeV)	$2 \leq n_{jet} \leq 3$	$n_{jet} \geq 4$
$275 < H_T < 325$	6 (10)%	3 (10)%
$325 < H_T < 375$	6 (10)%	6 (10)%
$375 < H_T < 575$	7 (10)%	9 (10)%
$575 < H_T < 775$	13 (20)%	15 (20)%
$H_T > 775$	19 (20)%	21 (30)%

**Table 4.14.:** 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 and shown in brackets.

1792 Figure 4.11 shows the sets of closure tests overlaid on top of grey bands that represent  
1793 the  $H_T$  dependent systematic uncertainties. These systematic uncertainties are assumed  
1794 to fully uncorrelated between the different  $n_b$  multiplicity categories and across the five  
1795  $H_T$  regions. This can be considered a more conservative approach given that some  
1796 correlations between adjacent  $H_T$  bins could be expected due to comparable kinematics.

1797 As already referenced. These closure tests represent a conservative estimate of the  
1798 systematic uncertainty in making a background prediction for the signal region. This



**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 bins: (a)  $2 \leq n_{\text{jet}} \leq 3$  and (b)  $n_{\text{jet}} \geq 4$ .

1799 is due to significant differences in the background composition and event kinematics  
 1800 between the two sub-samples used in the closure tests. This is contrary to the signal  
 1801 region prediction where the two sub-samples are both have a comparable background

1802 admixture and similar kinematics owing to the fact that the predictions are always made  
1803 using the same ( $n_{jet}, n_b, H_T$ ) bin.

1804 This point is emphasised when we examine the sensitivity of the **TF**'s to a change in the  
1805 admixture of W + jets and  $t\bar{t}$  with the control and signal samples. This is accomplished  
1806 by varying the cross sections of the W + jets and  $t\bar{t}$  by +20% and -20%, respectively.  
1807 Figures C.2 and C.3 within Appendix C, show the effect upon the closure tests for both  
1808 jet multiplicity categories. Given these variations in cross sections, the level of closure is  
1809 found to be significantly worse, with biases as large as  $\sim 30\%$ , most apparent in the  
1810 lowest  $H_T$  bins. However the **TF**'s used to extrapolate from control to signal are seen to  
1811 change only at the percent level by this large change in cross section, shown in Table C.1.

1812 Given the robust behaviour of the translation factors with respect to large (and opposite)  
1813 variations in the W + jets and  $t\bar{t}$  cross sections, one can assume with confidence that  
1814 any bias in the translation factors is adequately (and conservatively) covered by the  
1815 systematic uncertainties used in the analysis.

## 1816 4.7. Simplified Models, Efficiencies and Systematic 1817 Uncertainties

1818 The results of the analysis are interpreted using various **SMS** signal models as already  
1819 introduced in Section (2.4.1). Each model is parameterised in a two dimensional parameter  
1820 space, ( $m_{\tilde{q}/\tilde{g}}, m_{LSP}$ ), from which upper limits on the production cross sections of the  
1821 various **SMS** models can be set.

1822 Each signal sample is generated at Next to Leading Order (**NLO**) and Next to Leading  
1823 Logarithmic Order (**NLL**) [83] using the **Fastsim** framework. This framework represents  
1824 a simplified simulation of the **CMS** detector, but allows for faster production of various  
1825 signal topologies with different mass parameters. A series of correction factors are applied  
1826 to account for the effects on the b-tagging rate between **Fastsim** [84] and **Fullsim** [85]  
1827 and are detailed in Section (4.7.2).

<sup>1828</sup>

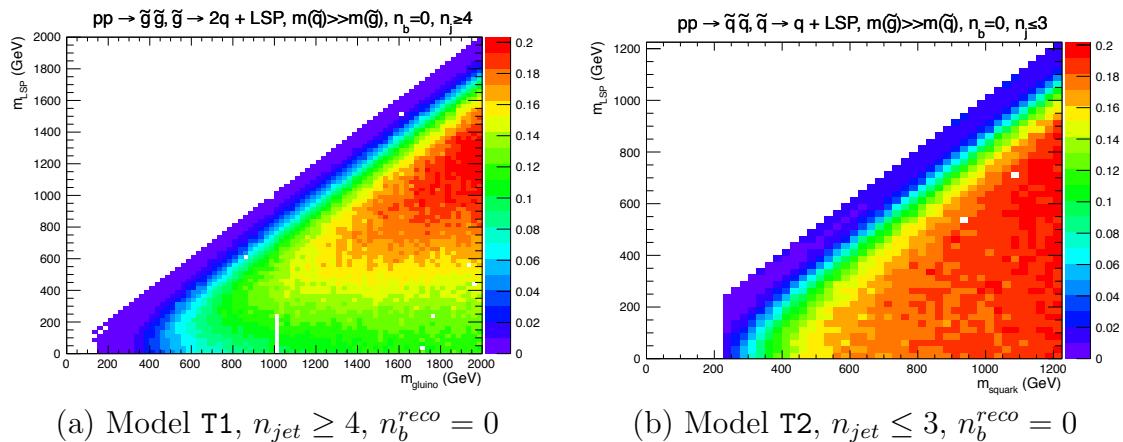
### 4.7.1. Signal efficiency

<sup>1829</sup> The analysis selection efficiency,  $\epsilon$ , is measured for each mass point of the interpreted  
<sup>1830</sup> model, this serves as a measure of the sensitivity of the signal selection for that particular  
<sup>1831</sup> sparticle and LSP mass . The signal yield is then given by

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

<sup>1832</sup> where  $\sigma$  represents the model's cross section and  $\mathcal{L}$  the luminosity. An upper limit on  $\sigma$   
<sup>1833</sup> taken from theory can then allow for the setting of limits in terms of the particle mass.

<sup>1834</sup> Figure 4.12 shows the expected signal efficiency of the signal selection for the T1 and  
<sup>1835</sup> T2 SMS models interpreted in this analysis. The efficiency maps are produced with the  
<sup>1836</sup> requirement  $H_T > 275$  GeV (i.e., no binning in  $H_T$ ) and requirements on  $n_{jet}$  and  $n_b^{reco}$   
<sup>1837</sup> that are appropriate for the model in question.



**Figure 4.12.:** Signal efficiencies for the SMS models (a) T1 ( $\tilde{g}\tilde{g}^* \rightarrow q\tilde{q}\chi_1^0 q\tilde{q}\chi_1^0$ ) 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  $n_b^{reco} = 0$ .

<sup>1838</sup> The same procedure is conducted in the analysis control samples. It is found in the  $\mu$   
<sup>1839</sup> + jets control samples, that the S/B ratios for the expected signal yields in each of the  
<sup>1840</sup> SMS models are many times ( $\sim 40$ -100) smaller than in the hadronic signal region. The  
<sup>1841</sup> relative contamination for the  $\mu\mu$  + jets sample is smaller still due to the requirement of  
<sup>1842</sup> a second muon. The relative contamination for the  $\gamma$  + jets sample is expected to be  
<sup>1843</sup> zero for the models under consideration. These small, relative levels of contamination  
<sup>1844</sup> are accounted for in the fitting procedure, as described in Section (4.8.4).

---

### 1845 4.7.2. Applying b-tag scale factor corrections in signal samples

1846 High-statistic **FastSim** signal simulation samples are unavailable for each signal signal  
 1847 point, which means that a different procedure to the formula method described in Section  
 1848 (4.5) is employed. Furthermore, the use of the **FastSim** framework in the reconstruction  
 1849 introduces an extra set of scale-factor corrections, to be applied simultaneously with  
 1850 those correcting the full-simulation to the data.

1851 For these signal models, an event-by-event re-weighting procedure is applied. This applied  
 1852 weight depends on both the flavour content and the b-tagging status of the reconstruction  
 1853 level jets in the event.

1854 The re-weighting procedure can be described by first considering a single jet in an signal  
 1855 sample event. The flavour of the jet is determined using the method described in Section  
 1856 (4.5.1).

1857 Taking the flavour,  $p_T$  and  $\eta$  values of the jet, the expected tagging efficiency,  $\epsilon_{MC}(p_T, \eta, f)$ ,  
 1858 in simulation is retrieved from a map of tagging efficiencies determined from the **FullSim**  
 1859 **SM** simulation samples, and binned as a function of jet  $p_T$ ,  $\eta$  and flavour after the  
 1860 application of the hadronic signal selection. The binning is chosen to reflect the set of  $p_T$   
 1861 and  $\eta$  dependant corrections of simulation to data defined by [86].

1862 The actual tagging efficiency of the **FastSim** jet,  $\epsilon_{\text{FastSim}}(p_T, \eta, f)$ , differs from that  
 1863 measured in **FullSim**,  $\epsilon_{MC}(p_T, \eta, f)$  and is related via an additional correction factor,

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

1864 where  $SF_{\text{Fast} \rightarrow \text{Full}}(p_T, \eta, f)$  represents a set of  $p_T$  and  $\eta$  dependant corrections determined  
 1865 from the ratio between the efficiency and mis-tagging rates of a  $t\bar{t}$  **FullSim** and  $t\bar{t}$   
 1866 **FastSim** sample. The central value for these corrections is the same for all signal samples.  
 1867 Similarly the tagging efficiencies measured in data [64],  $\epsilon_{Data}(p_T, \eta, f)$ , are further related  
 1868 to  $\epsilon_{\text{FastSim}}(p_T, \eta, f)$  by the equation,

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

1869 For each jet, the weight of the event is re-weighted according to whether the jet fires the  
1870 b-tagger. In the instance that the jet *is* b-tagged the event weight is modified by,

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

1871 and in the case that the jet is *not* tagged,

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

1872 Once all events have been reweighted this way, the yields in each  $n_b^{reco}$  bin represent the  
1873 corrected MC yields.

### 1874 4.7.3. Experimental uncertainties

1875 The systematic uncertainty on the expected signal acceptance times analysis efficiency is  
1876 determined independently for each **SMS** model considered. These systematics stem  
1877 from uncertainties on the parton distribution functions, the luminosity measurement,  
1878 jet energy scale, b-tag scale factor measurements and the efficiencies of various cuts used  
1879 in the signal selection, including the  $H_T / E_T$ , dead **ECAL** cleaning filter and lepton /  
1880 photon event vetoes.

1881 Rather than trying to estimate the level of systematic that is applicable point-by-point in  
1882 a model space, general behaviours are considered and constant systematics are estimated  
1883 in two regions of the **SMS** models parameter space. These two regions are defined as  
1884 near to (small mass splittings) and far (large mass splittings) from the diagonal, where  
1885 far is realised by the condition

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

1886 The total systematics in each region are evaluated in the following ways:

1887 **Jet energy scale** : The relative change in the signal efficiency is gauged by varying  
1888 the energy of all jets in an event up or down according to a  $p_T$  and  $\eta$  dependent jet  
1889 energy scale uncertainty. Within the two systematic regions, the resulting systematic  
1890 uncertainties for each **SMS** model are determined by taking the value of the 68<sup>th</sup>  
1891 percentile for the distributions of the relative change in the signal efficiency.

1892 **Luminosity measurement** : The measurement of luminosity taken propagates  
1893 through to an uncertainty on the signal event yield when considering any new  
1894 physics model, which is currently 4.4% [87].

1895 **Parton density function** : The effects of varying the set of parton distribution  
1896 functions used and changing the renormalization/factorization scale used to compute  
1897 the cross sections by a factor of two up or down have been found to affect the  
1898 efficiency by up to 10%.

1899  **$H_T/E_T$  cleaning cut** : The ratio of the efficiencies of the cleaning cut are compared  
1900 in simulation and data after application of the  $\mu +$  jets control sample selection.  
1901 No  $\alpha_T$  cut or further event cleaning filters are applied. The ratio of the efficiencies  
1902 observed in data and simulation for a cut value of  $H_T/E_T < 1.25$  and the two jet  
1903 multiplicity bins,  $2 \leq n_{jet} \leq 3$  and  $n_{jet} \geq 4$  are  $1.028 \pm 0.007$  and  $1.038 \pm 0.015$   
1904 respectively. These deviations are taken to represent the systematic uncertainty on  
1905 the simulation modelling of this variable.

1906 **Deal ECAL cleaning filter** : The ratio of the efficiencies observed in data and  
1907 simulation for this filter in the two jet multiplicity bins,  $2 \leq n_{jet} \leq 3$  and  $n_{jet} \geq 4$ ,  
1908 are  $0.961 \pm 0.008$  and  $0.961 \pm 0.009$ , respectively. These deviations from unity  
1909 are taken to represent the systematic uncertainties in the modelling in simulation of  
1910 this filter.

1911 **Lepton and photon vetoes** : The uncertainty on the efficiency of the lepton and  
1912 photon vetoes is established by considering the efficiency of the vetoes after applying  
1913 filters with identical logic but based on truth information. If the efficiency is not  
1914 100%, then this represents the fraction of signal events that should not be vetoed.

This deviation is taken directly as the systematic uncertainty on the efficiency. The systematic uncertainty is only non-zero for models which contain third-generation quarks in the final state.

**B-tag scale factor uncertainties :** The relative change in the signal efficiency is observed when relevant flavour,  $p_T$  and  $\eta$  dependant b-tag correction factors, are varied up or down by their uncertainty. Within the two systematic regions, the 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 4.15 and 4.16 summarise all the aforementioned systematic uncertainties on the signal efficiencies for each individual **SMS** model interpreted in the analysis. 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 included in the limit calculation.

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

**Table 4.15.:** Estimates of systematic uncertainties on the signal efficiency (%) for various **SMS** models when considering points in the region near to the diagonal (i.e. small mass splitting and compressed spectra). The uncertainties are added in quadrature to obtain the total.

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

**Table 4.16.:** 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>1929</sup> 4.8. Statistical Framework

<sup>1930</sup> For a given category of events satisfying requirements on both  $n_{jet}$  and  $n_b^{reco}$ , a likelihood  
<sup>1931</sup> model of the observations in multiple data samples is used to gauge agreement between  
<sup>1932</sup> the observed yields in the hadronic signal region, and the predicted yields obtained from  
<sup>1933</sup> the control samples. In addition to checking whether the predictions are compatible with  
<sup>1934</sup> a **SM** only hypothesis, the likelihood model is also used to test for the presence of a  
<sup>1935</sup> variety of signal models.

### <sup>1936</sup> 4.8.1. Hadronic sample

<sup>1937</sup> Let  $N$  be the number of bins on  $H_T$ , with  $n^i$  the number of events observed satisfying  
<sup>1938</sup> all selection requirements in each  $H_T$  bin  $i$ . The likelihood of the observations can then  
<sup>1939</sup> be written :

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

<sup>1940</sup> where  $b^i$  represents the expected **SM** background

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

<sup>1941</sup> and  $s^i$  the expected number of signal events from the different **SMS** models interpreted.  
<sup>1942</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>1943</sub> **4.8.2.  $H_T$  evolution model**

<sub>1944</sub> The hypothesis, that for a process the  $\alpha_T$  ratio falls exponentially see Section (4.2.4) in  $H_T$  is defined by Equation (4.10), where  $k_{QCD}$  is constrained by measurements in a signal sideband region.

<sub>1947</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>1948</sub> where  $m_i$  represent the number of events observed with  $\alpha_T \leq 0.55$  in each  $H_T$  bin  $i$ , and <sub>1949</sub>  $\langle H_T \rangle$  represents the mean  $H_T$  of each bin. Expressed as functions of just the zeroth bin, <sub>1950</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>1951</sub> **4.8.3. EWK control samples**

<sub>1952</sub> The **EWK** background estimation within each bin,  $i$ , is broken into two components, the <sub>1953</sub> expected yield from  $Z \rightarrow \nu\bar{\nu}$  and  $t\bar{t}$ -W (plus other residual backgrounds) events. This is <sub>1954</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>1955</sub> This can be further expressed as

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

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

1956 where  $f_{Zinv}^i$  represents the expected yield from  $Z \rightarrow \nu\bar{\nu}$  in bin  $i$  divided by the expected  
1957 **EWK** background  $EWK^i$ . This fraction is modelled as a linear component

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

1958 where  $N$  again represents the number of  $H_T$  bins, and  $f_{Zinv}^i$  and  $f_{Zinv}^{N-1}$  are float parameters  
1959 whose final values are limited between zero and one.

1960 Within each  $H_T$  bin there are three background measurements for the different control  
1961 samples,  $n_\gamma^i$ ,  $n_\mu^i$  and  $n_{\mu\mu}^i$ , representing the event yields from the  $\gamma +$  jets,  $\mu +$  jets and  
1962  $\mu\mu +$  jets control samples respectively. Each of these have a corresponding yield in  
1963 simulation,  $MC_\gamma^i$ ,  $MC_\mu^i$  and  $MC_{\mu\mu}^i$ . Within the hadronic signal region there are also  
1964 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  
1965 used to define

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

1966 where  $r_p^i$  represents the inverse of the **TF**'s used to extrapolate the yield of each background  
1967 process.

1968 The likelihoods regarding the three measured yields  $n_\gamma^i$ ,  $n_{\mu\mu}^i$ ,  $n_\mu^i$  can then be fully expressed  
1969 as

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

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

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

$$(4.31)$$

1970 which contain an additional term  $s_\mu^i$ , which represents the signal contamination in the  
1971  $\mu +$  jets sample. The parameters  $\rho_{\gamma Z}^j$ ,  $\rho_{\mu\mu}^j$  and  $\rho_\mu^j$  represent “correction factors” that

1972 accommodate the systematic uncertainties associated with the control sample based  
1973 background constraints.

1974 Each of these equations are used to estimate the maximum likelihood value for relevant  
1975 background in the signal region given the observations  $n_p^i$  in each of the control samples  
1976 (see Section (4.2.3)).

1977 The measurements in each of the control samples and the hadronic signal region, along  
1978 with the ratios  $r_\gamma^i$ ,  $r_{\mu\mu}^i$ , and  $r_\mu^i$ , are all considered simultaneously through the relationships  
1979 defined by Equations (4.19),(4.24) and (4.25).

1980 In addition to the Poission product, an additional log-normal term is introduced to  
1981 accommodate the systematic uncertainties given by,

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

1982 where  $\sigma_{\gamma Z}^j$ ,  $\sigma_{\mu\mu Z}^j$  and  $\sigma_{\mu W}^j$  represent the relative systematic uncertainties for the control  
1983 sample constraints and Logn is the log-normal distribution [88],

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

1984 Five parameters per control sample are used to span the eight  $H_T$  bins, with just one  
1985 used for the three  $H_T$  bins in the  $n_b^{reco} \geq 4$  category. These parameters span the same  
1986  $H_T$  ranges described in Section (4.6) and is shown in Table 4.17.

$H_T$ 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 4.17.:** The systematic parameters used in  $H_T$  bins. Left: categories with eight bins;  
right: category with three bins.

1987 Alternatively, in the higher  $n_b^{reco}$  categories ( $n_b^{reco} = 3$ ,  $n_b^{reco} \geq 4$ ), the single muon sample  
1988 is used to constrain the total EWK background. Therefore the likelihood function is  
1989 greatly simplified and is represented by

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

1990 where,

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

1991 **4.8.4. Contributions from signal**

1992 The cross section for each model is represented by  $x$  and  $l$  represents the total recorded  
1993 luminosity considered by the analysis in the signal region. Let  $\epsilon_{had}^i$  and  $\epsilon_\mu^i$  represent the  
1994 analysis selection efficiency for that particular signal model in  $H_T$  bin  $i$  of the hadronic  
1995 and  $\mu +$  jets control sample respectively. Letting  $\delta$  represent the relative uncertainty on  
1996 the signal yield, assumed to be fully correlated across all bins, and  $\rho_{sig}$  the “correction  
1997 factor” to the signal yield which accommodates this uncertainty.  $f$  represents an unknown  
1998 multiplicative factor on the signal cross section, for which an allowed interval is computed.

1999 The expected signal yield  $s^i$  is thus given by

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

2000 and signal contamination with the  $\mu +$  jets control sample by

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

2001 The systematic uncertainty on the signal is additionally included by the term

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

2002 A discussion of the **SMS** signal models through which the analysis is interpreted can be  
2003 found in the following Chapter.

2004 **4.8.5. Total likelihood**

2005 The total likelihood function for a given signal bin  $k(n_b^{reco}, n_{jet})$  is given by the product  
2006 of the likelihood functions introduced within the previous sections:

$$L^k = L_{had}^k \times L_\mu^k \times L_\gamma^k \times L_{\mu\mu}^k \times L_{EWKsyst}^k \times L_{QCD}^k. \quad (4.39)$$

2007 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 is  
2009 used to estimate the **EWK** background, there are 15 nuisance parameters. Where three  
2010  $H_T$  bins are used (the highest  $n_b^{reco}$  category), there are 6 nuisance parameters. This  
2011 information is summarised within Table 4.18.

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 4.18.:** 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.

2012 When considering **SUSY** signal models within the likelihood, an additional parameter is  
2013 introduced,  $\rho_{sig}$ . When multiple categories are fit simultaneously the total likelihood is  
2014 then represented by

$$L = L_{sig} \times \prod_k L_{had}^k \times L_\mu^k \times L_\gamma^k \times L_{\mu\mu}^k \times L_{EWKsyst}^k \times L_{QCD}^k. \quad (4.40)$$

# Chapter 5.

## <sup>2015</sup> Results and Interpretation

<sup>2016</sup> Using the statistical framework outlined in the previous chapter, results are compared to  
<sup>2017</sup> a **SM**-only hypothesis (Section (5.1)) and interpreted using various **SMS** models (Section  
<sup>2018</sup> (5.2)).

### <sup>2019</sup> 5.1. Standard Model

<sup>2020</sup> The **SM** background only hypothesis is tested by removing any signal contributions  
<sup>2021</sup> within the signal and control samples, and the likelihood function is maximised over all  
<sup>2022</sup> parameters using Rootfit [89] and MINUIT [90]. The results of the search consist of the  
<sup>2023</sup> observed yields in the hadronic signal sample, and the  $\mu + \text{jets}$ ,  $\mu\mu + \text{jets}$  and  $\gamma + \text{jets}$   
<sup>2024</sup> control samples.

<sup>2025</sup> These observed yields along with the expectations and uncertainties given by the simulta-  
<sup>2026</sup> neous fit for the hadronic signal region are given in Table 5.2. The results obtained from  
<sup>2027</sup> the simultaneous fits, including that of the three control samples, are shown in Figure  
<sup>2028</sup> 5.1-5.8, as summarised in Table 5.1.

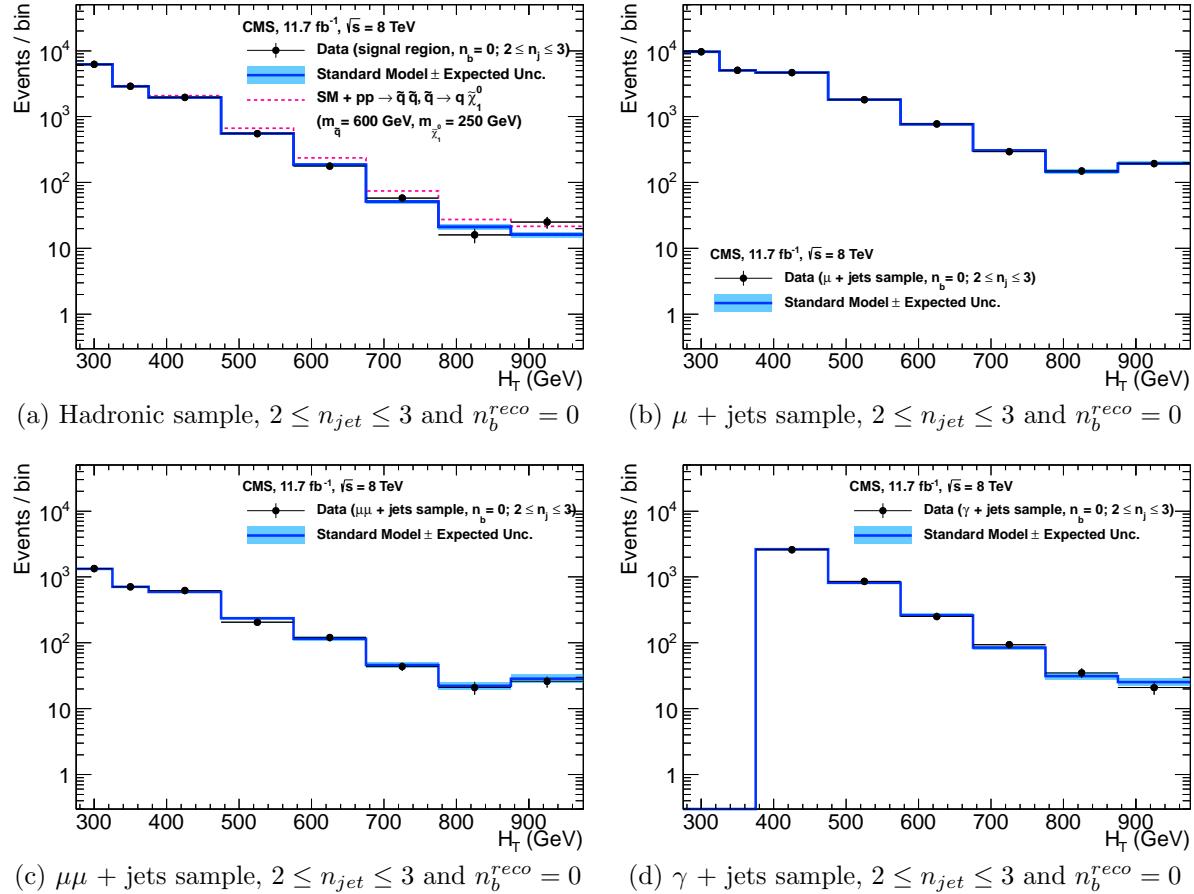
<sup>2029</sup> The figures show a comparison between the observed yields and the **SM** expectations  
<sup>2030</sup> across all  $H_T$  bins, for events in both  $n_{jet}$  and  $n_b^{reco}$  multiplicity categories. In all categories  
<sup>2031</sup> the samples are well described by the **SM** only hypothesis. In particular no significant  
<sup>2032</sup> excess 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

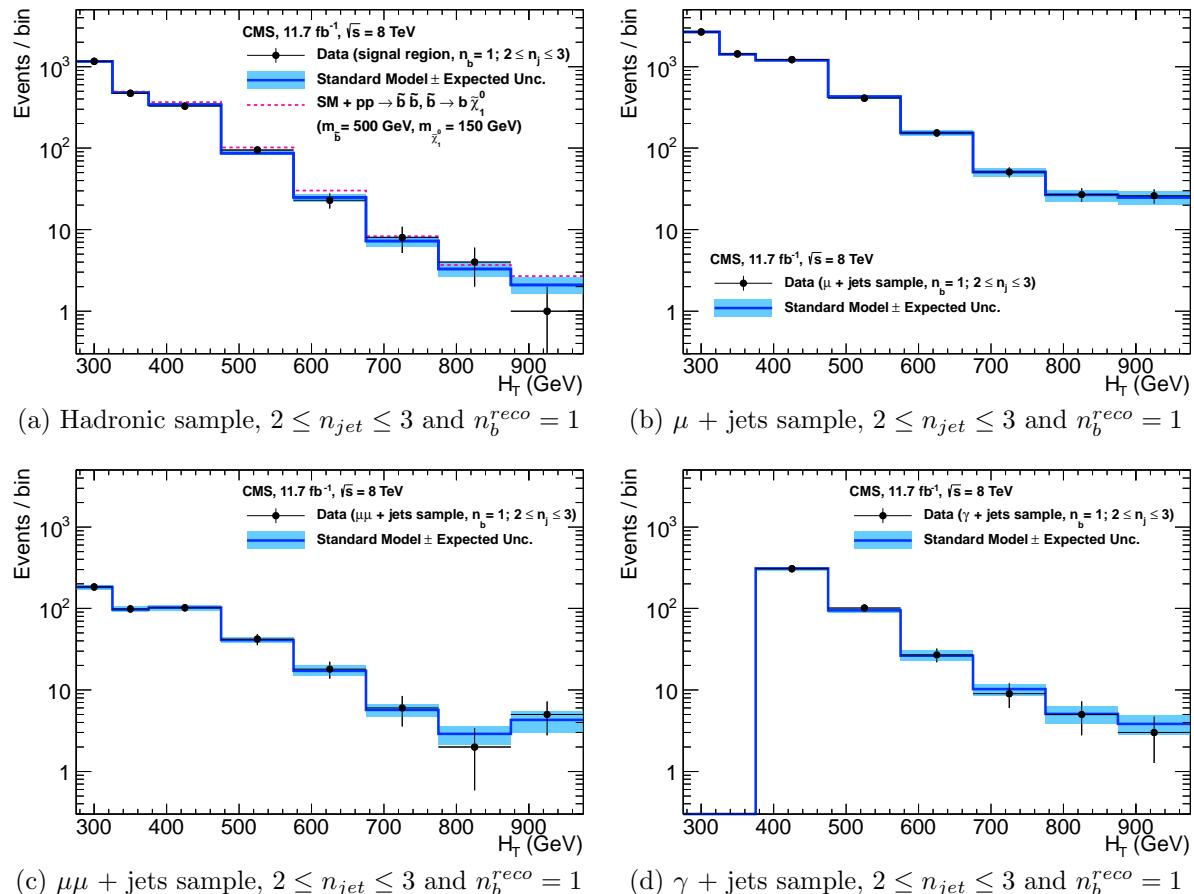
**Table 5.1.:** Summary of control samples used by each fit results, and the Figures in which they are displayed.

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

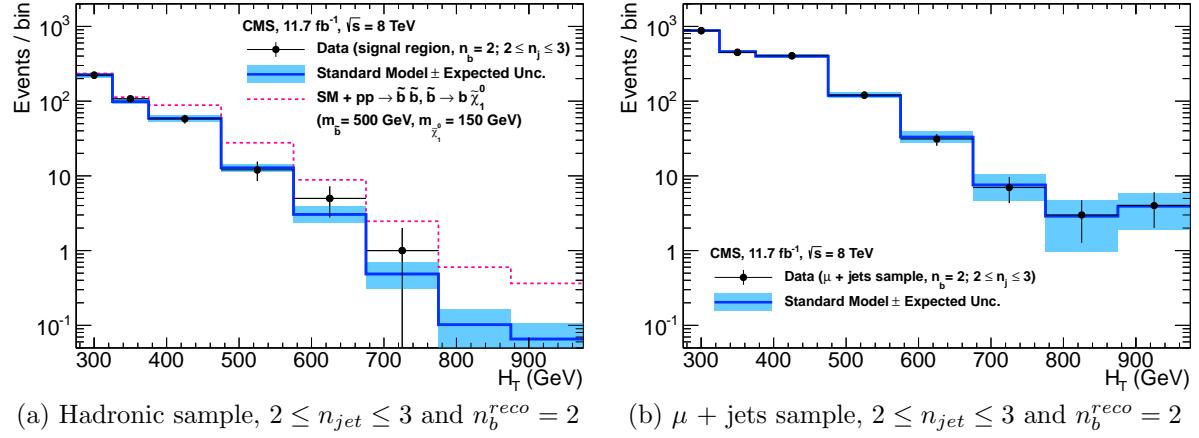
**Table 5.2.:** Comparison of the measured yields in the each  $H_T$ ,  $n_{jet}$  and  $n_b^{reco}$  jet multiplicity bins for the hadronic sample with the SM expectations and combined statistical and systematic uncertainties given by the simultaneous fit.



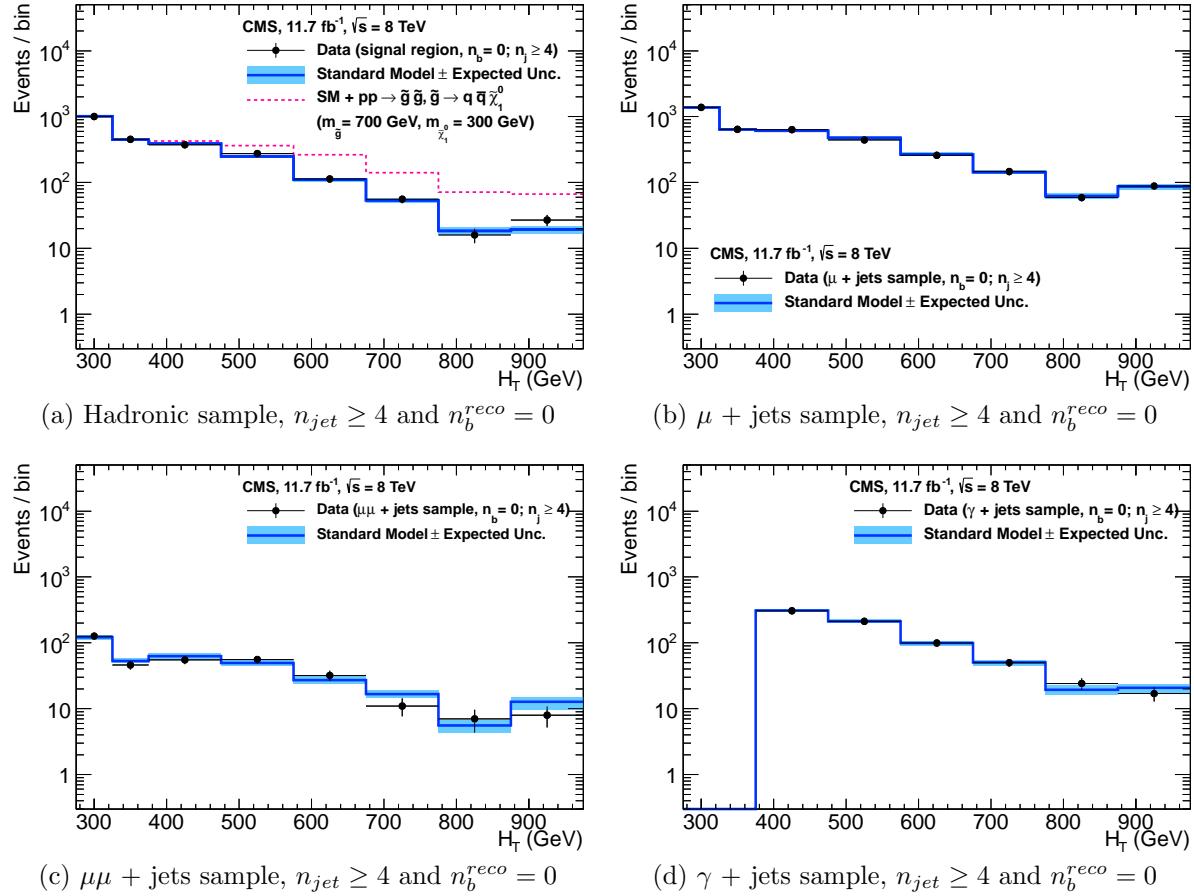
**Figure 5.1.:** Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of  $H_T$  for the (a) hadronic, (b)  $\mu + \text{jets}$ , (c)  $\mu\mu + \text{jets}$  and (d)  $\gamma + \text{jets}$  samples when requiring  $n_b^{reco} = 0$  and  $n_{jet} \leq 3$ . The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



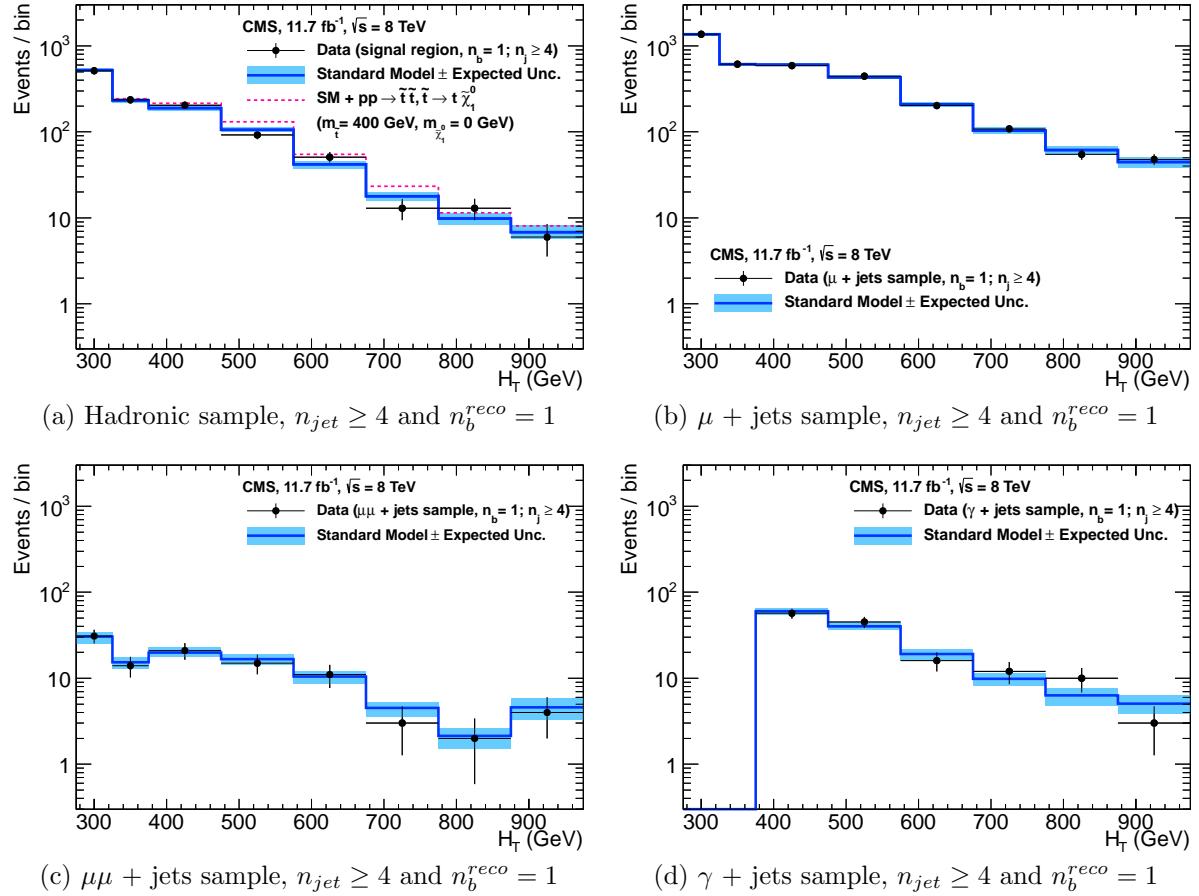
**Figure 5.2.:** Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of  $H_T$  for the (a) hadronic, (b)  $\mu +$  jets, (c)  $\mu\mu +$  jets and (d)  $\gamma +$  jets samples when requiring  $n_b^{reco} = 1$  and  $n_{jet} \leq 3$ . The observed event yields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



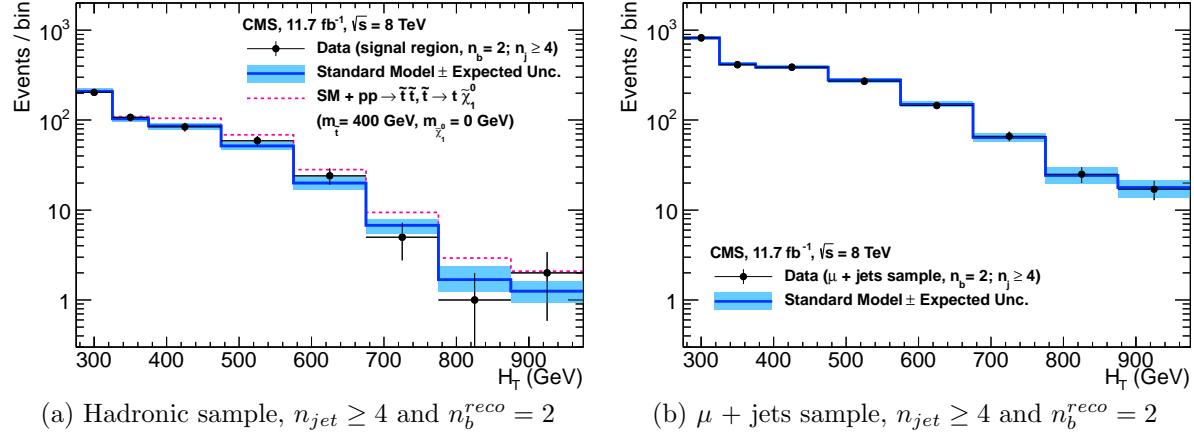
**Figure 5.3.:** Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of  $H_T$  for the (a) hadronic, (b)  $\mu +$  jets, (c)  $\mu\mu +$  jets and (d)  $\gamma +$  jets samples when requiring  $n_b^{reco} = 2$  and  $n_{jet} \leq 3$ . The observed event yields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



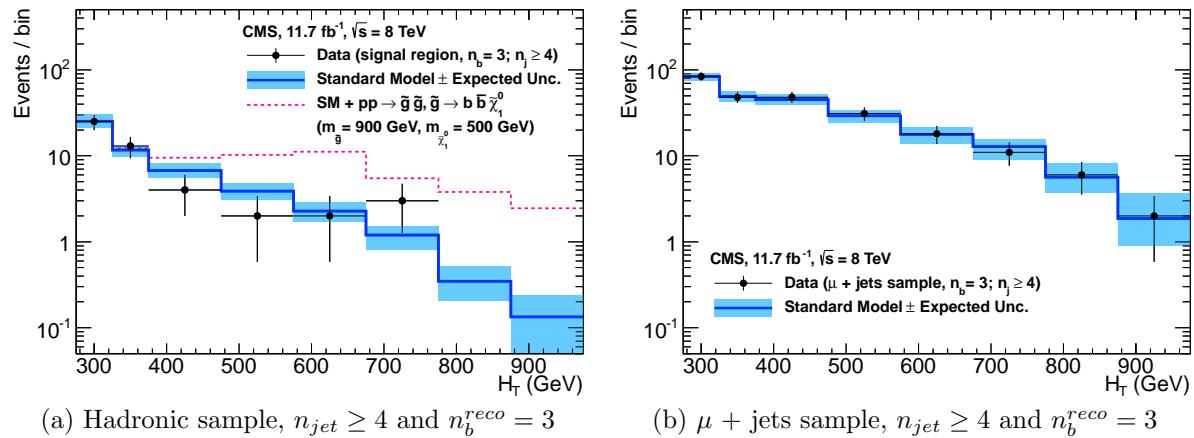
**Figure 5.4.:** Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of  $H_T$  for the (a) hadronic, (b)  $\mu +$  jets, (c)  $\mu\mu +$  jets and (d)  $\gamma +$  jets samples when requiring  $n_b^{reco} = 0$  and  $n_{jet} \geq 4$ . The observed event yields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



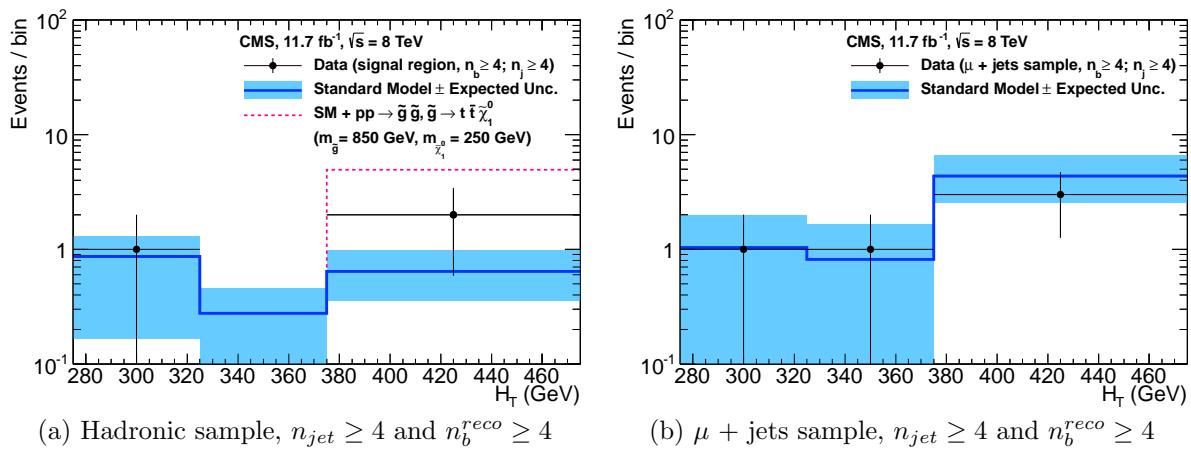
**Figure 5.5.:** Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of  $H_T$  for the (a) hadronic, (b)  $\mu +$  jets, (c)  $\mu\mu +$  jets and (d)  $\gamma +$  jets samples when requiring  $n_b^{reco} = 1$  and  $n_{jet} \geq 4$ . The observed event yields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

2033 **5.2. SUSY**

2034 Limits are set in the parameter space of a set of **SMS** models that characterise both  
2035 natural **SUSY** third generation squark production, and compressed spectra where the  
2036 mass splitting between the particle and **LSP** is small, leading to soft final state jets.  
2037 However as detailed in Section (2.4.1), the individual models are not representative of a  
2038 real physical **SUSY** model as only one decay process is considered. Instead these models  
2039 represent a way to test for signs of specific signatures indicating new physics.

2040 **5.2.1. The  $CL_s$  method**

2041 The  $CL_s$  method [91][92][93] is used to compute the limits for signal models, with the  
2042 one-sided profile likelihood ratio as the test statistic [94].

The test statistic is defined as

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

2043 where

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

2044 represents the profile likelihood ratio, in which  $\mu \equiv f$  from Section (4.8.4), is the  
2045 parameter characterising the signal strength.  $\hat{\mu}$  is defined at the maximum likelihood  
2046 value,  $\hat{\theta}$  the set of maximum likelihood values of the nuisance parameters and  $\theta_\mu$  the set  
2047 of maximum values of the nuisance parameters for a given value of  $\mu$ .

2048 When  $\mu \equiv f = 1$ , the signal model is considered at its nominal production cross section.  
2049 The distribution of  $q_\mu$  is built up via the generation of pseudo experiments in order to  
2050 obtain two distributions for the background (B) and signal plus background (S+B) cases.

2051 The compatibility of a signal model with observations in data is determined by the  
2052 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  $\text{CL}_s \leq 0.05$  [95].

### 5.2.2. Interpretation in simplified signal models

Different  $n_{jet}$  and  $n_b^{reco}$  bins are used in the interpretation of different **SMS** models. The choice of the categories used are made to increase 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 simplified models T2bb, T1tttt, and T1bbbb describe various production and decay mechanisms in the context of third-generation squarks.

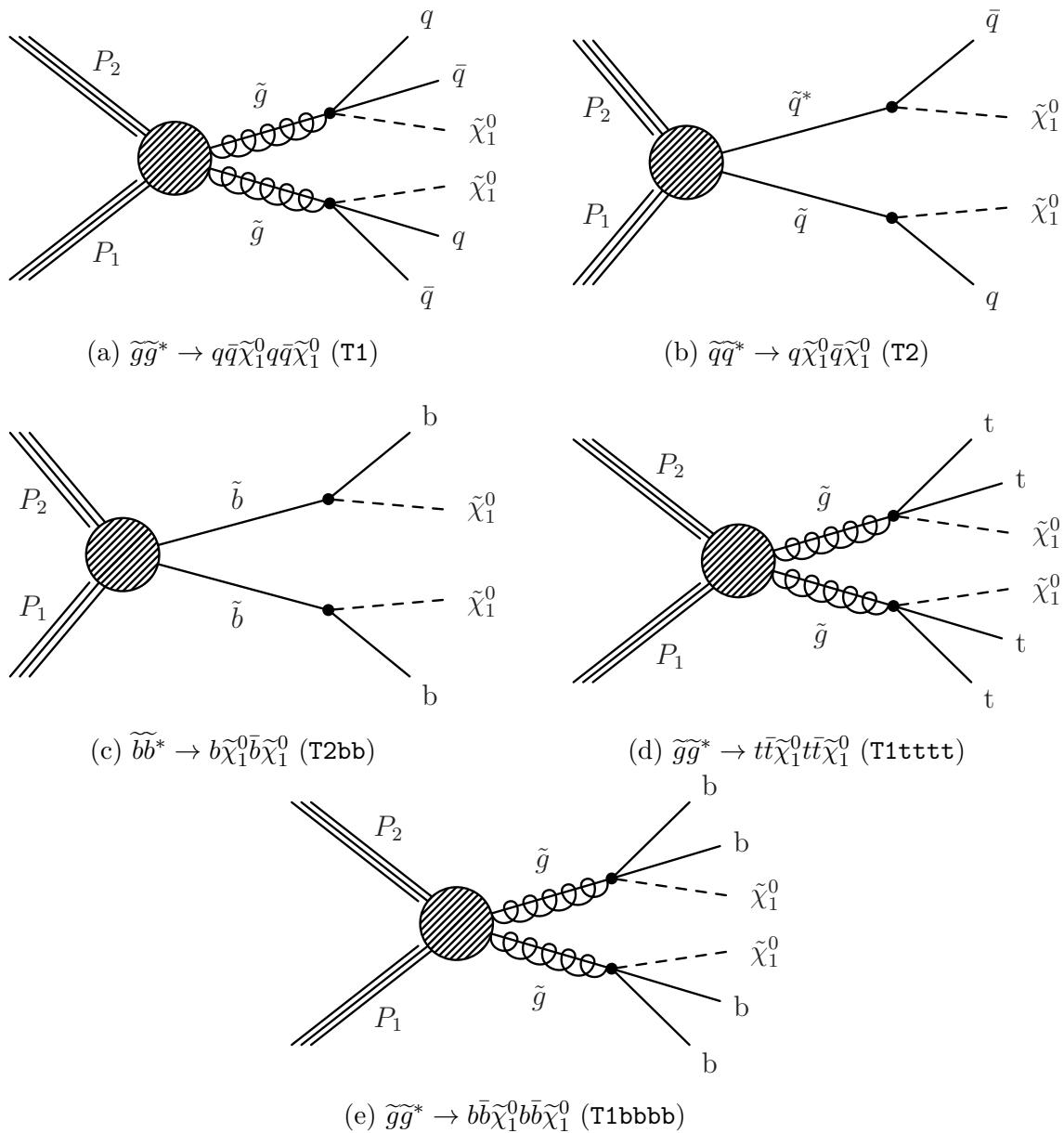
Model	Production/decay	$n_{jet}$	$n_b^{reco}$	Process	Limit	$m_{\tilde{q}/\tilde{g}}^{\text{best}}$ (GeV)	$m_{\text{LSP}}^{\text{best}}$ (GeV)
T1	$pp \rightarrow \tilde{g}\tilde{g}^* \rightarrow q\bar{q}\tilde{\chi}_1^0 q\bar{q}\tilde{\chi}_1^0$	$\geq 4$	0	5.9(a)	5.10(a)	$\sim 950$	$\sim 450$
T2	$pp \rightarrow \tilde{q}\tilde{q}^* \rightarrow q\tilde{\chi}_1^0 \bar{q}\tilde{\chi}_1^0$	$\leq 3$	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 quotes values are conservatively determined from the observed exclusion based on the theoretical production cross section minus  $1\sigma$  uncertainty.

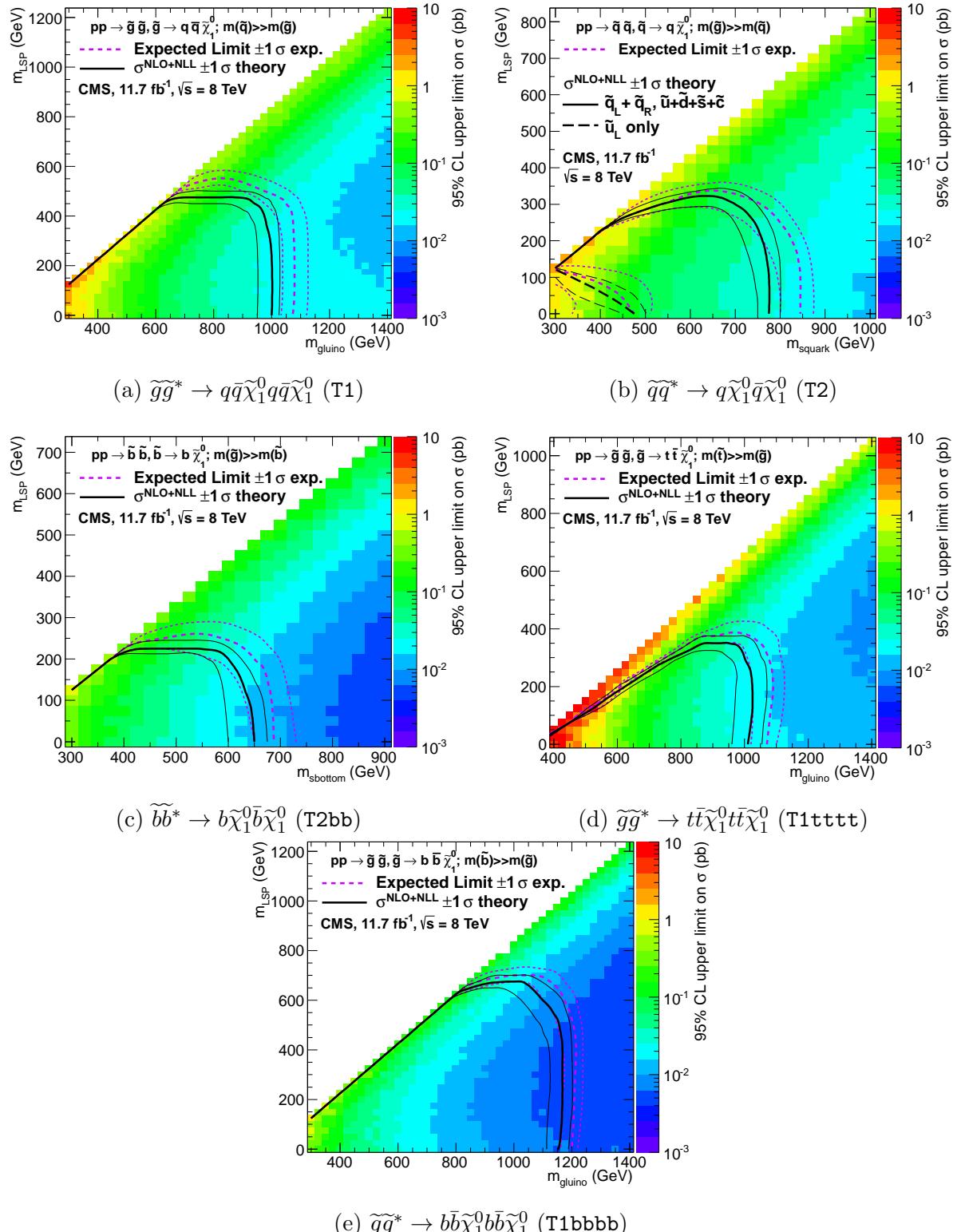
Experimental uncertainties on the **SM** background predictions (10 – 30%, Section (4.6.1)), the luminosity measurement (4.4%), and the total acceptance times efficiency of the selection for the considered signal model (12 – 18%, Section (4.7)) are included in the calculation of the limit.

2070 Signal efficiency in the kinematic region defined by  $0 < m_{\tilde{g}(\tilde{q})} < 175$  GeV or  $m_{\tilde{g}(\tilde{q})} < 300$   
2071 GeV is strongly affected by the presence of Initial State Radiation (**ISR**). This region in  
2072 which direct (i.e., non-**ISR** induced) production is kinematically forbidden due to the  $H_T$   
2073  $> 275$  GeV requirement, therefore a large percentage of signal acceptance is due to the  
2074 effect of **ISR** jets. Given the large associated uncertainties, no interpretation is provided  
2075 for this kinematic region.

2076 The estimates on mass limits shown in Table 5.3, are determined conservatively from  
2077 the observed exclusion based on the theoretical production cross section, minus  $1\sigma$   
2078 uncertainty. The most stringent mass limits on pair-produced sparticles are obtained at  
2079 low **LSP** masses, while the limits typically weaken for compressed spectra points close to  
2080 the diagonal. In particular, for all of the considered **SMS** models, there is an **LSP** mass  
2081 beyond which no limit can be set, which can be observed from the figures referenced in the  
2082 table.



**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.

## 2083 Searches For Natural SUSY With 2084 B-tag Templates.

2085 Within this chapter a complimentary technique is discussed as a means to predict the  
2086 distribution of three and four reconstructed b-quark jets in an event. The recent discovery  
2087 of the Higgs boson has made third-generation “Natural SUSY” models attractive, given  
2088 that light top and bottom squarks are a candidate to stabilise divergent loop corrections  
2089 to the Higgs boson mass.

2090 Using the  $\alpha_T$  search as a base, a simple templated fit is employed to estimate the  
2091 SM background in higher b-tag multiplicities (3-4) from a region of a low number of  
2092 reconstructed b-jets (0-2). As a proof-of-concept, the procedure after being shown to  
2093 close in simulation, is applied to the SM enriched  $\mu + \text{jets}$  control sample of the  $\alpha_T$   
2094 all-hadronic search detailed in Chapter 4. To highlight the relative insensitivity of the  
2095 choice of the b-tagging algorithm working points in the effectiveness of the procedure,  
2096 results are presented using the CSV tagger (introduced in Section (3.3.2)) for the “Loose”,  
2097 “Medium” and “Tight” working points.

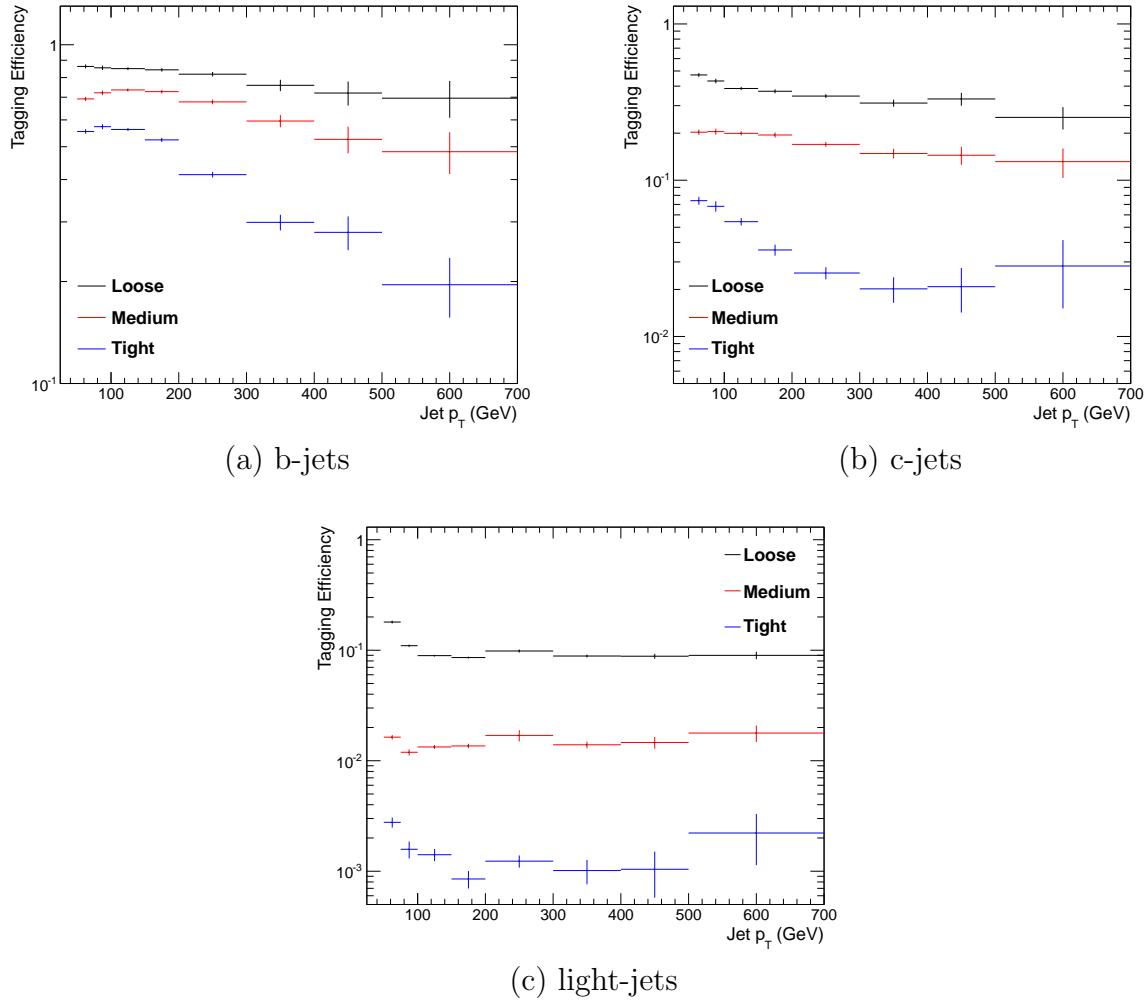
### 2098 6.1. Concept

2099 The dominant SM backgrounds most SUSY searches are typically  $t\bar{t} + \text{jets}$ ,  $W + \text{jets}$  and  
2100  $Z \rightarrow \nu\bar{\nu} + \text{jets}$ . These process are characterised by typically having zero or two underlying  
2101 b-quarks per event. The first step in this approach is to categorise two templates to be  
2102 fitted to the low  $n_b^{reco}$  multiplicity in terms of these underlying b-quark event 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.

- 2103 Thus two templates are defined, Z0 and Z2 ( $t\bar{t}$  is combined with single top) which  
 2104 represent processes which have an underlying b-quark content of zero or two respectively.
- 2105 Both these templates can be generated through the application of the relevant event  
 2106 selection and taking the underlying  $n_b^{reco}$  distribution directly from simulation. However  
 2107 as discussed within Section (4.5), there are large uncertainties for high  $n_b^{reco}$  multiplicities  
 2108 due to limited MC statistics. This is particularly prominent for the Z0 templates, where  
 2109 the number of reconstructed b-tags is driven primarily by the light-quark mis-tagging  
 2110 rate. Therefore to improve the statistical precision of the predictions the formula method,  
 2111 introduced in Section (4.5.1) is used.
- 2112 The generation of these templates is then dependant upon the jet-flavour content and  
 2113 b-tagging rate within the phase space of interest, with the tagging probabilities of a jet  
 2114 being a function of the jet  $p_T$ , the pseudo-rapidity  $|\eta|$ , and the jet-flavour. This can be  
 2115 observed in Figure 6.1, where the b-tagging / c-quark mis-tagging / light mis-tagging  
 2116 efficiency for the three working points of the CSV tagger is shown as a function of jet  $p_T$ .
- 2117 Before the templates are generated, the relevant jet  $p_T$  and  $\eta$  corrections are applied to  
 2118 correct simulation to data, as specified in Section (4.5.3), to then determine the average  
 2119 tagging rates per analysis bin.
- 2120 These two templates are then fit to data in the low  $n_b^{reco}$  region (0-2). The fit result is  
 2121 used, along with the knowledge of the template shapes, to extrapolate an estimate to the  
 2122 high  $n_b^{reco}$  signal region (3,4), which is then compared to what is observed in data.
- 2123 This method can, in principle, be applied to any analysis where the signal hypothesis  
 2124 has a larger underlying b-quark spectra than the SM backgrounds, as it solely relies on  
 2125 fitting to the shape of the  $n_b^{reco}$  distribution.



**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  $H_T > 375$ .

## 2126 6.2. Application to the $\alpha_T$ Search

2127 As detailed in the previous chapter, the  $\alpha_T$  analysis is a search for **SUSY** particles  
 2128 in all-hadronic final states, utilising the kinematic variable  $\alpha_T$  to suppress QCD to a  
 2129 negligible level. **SM** enriched control samples are used to estimate the background within  
 2130 an all-hadronic signal region.

2131 The selection for the  $\mu + \text{jets}$  control samples defined in Section (4.2.3) is used to  
 2132 demonstrate the template fitting procedure both conceptually in simulation, and also  
 2133 when applied in data. This is chosen, as such a selection is dominated by events stemming  
 2134 from the **SM** processes with little or no signal contamination from potential new physics..

2135 Neither are contributions from rate **SM** processes with a higher underlying b-quark  
2136 content (e.g.  $t\bar{t}b\bar{b}$ ) expected. For these reasons, there is a degree of confidence that the  
2137 procedure should close when applied to this phase space.

2138 The analysis presented here is binning in source jet multiplicity bins, of 3,4 and  $\geq 5$   
2139 reconstructed jets per event (di-jet events are not included as there is no contribution  
2140 to the high  $n_b^{reco}$  region (3,4)) , in order to reduce the kinematic jet  $p_T$  dependence.  
2141 Furthermore the analysis is split into three  $H_T$  regions,

2142 • 275-325 GeV

2143 • 325-375 GeV

2144 •  $> 375$  GeV

2145 contrary to the eight used within the  $\alpha_T$  analysis. Templates for both underlying b-quark  
2146 content hypotheses are then generated for the nine defined analysis bins.

### 2147 6.2.1. Proof of principle in simulation

2148 In order to demonstrate that the template procedure produces accurate predictions  
2149 within simulation, the simulation samples in the analysis are firstly split into two to allow  
2150 for statistically independent fits to be performed.

2151 By combining the relevant ingredients necessary to employ the formula method,  $n_b^{reco}$   
2152 templates for  $Z = 0$  and  $Z= 2$  are generated individually for each  $n_{jet}$  and  $H_T$  bin using  
2153 one half of each simulation sample. A fit of these two templates is then performed in the  
2154 low  $n_b^{reco}$  (0-2) region, back to the sum of the other halves of each simulation sample in  
2155 order to check that the relevant information can be recovered in the  $n_b^{reco}$  signal region  
2156 (3-4).

2157 The fits are performed independently within each of the defined analysis bins to reduce the  
2158 dependence of the shapes of these distributions on simulation. The half of the simulation  
2159 sample for which the templates are fitted too, are taken directly from simulation, extending  
2160 this procedure to also be a validation of the formula method to accurately estimate  
2161 the  $n_b^{reco}$  distribution. Additionally as this test is performed in simulation, the relevant  
2162 corrections of the b-tagging rates between data and simulation are *not* applied.

2163 Within Figure 6.2, the results of this fitting procedure is shown for each **CSV** working  
2164 point. Results are presented for the  $n_{jet} \geq 5$  category, using the  $\mu +$  jets control sample

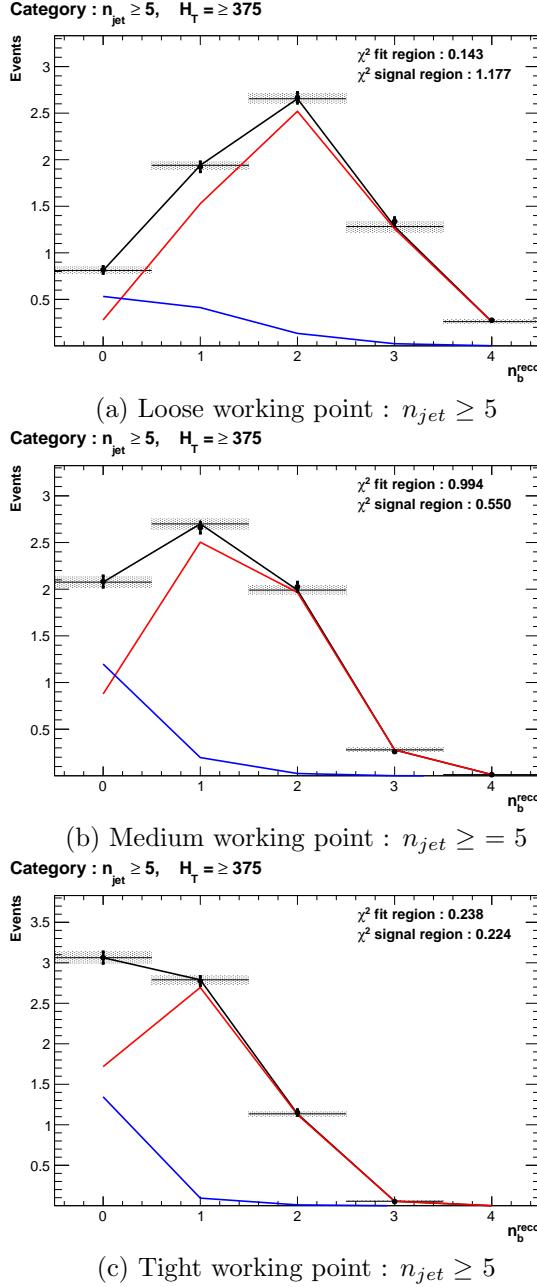
2165 selection in the inclusive  $H_T > 375$  GeV analysis bin. The grey bands represent the  
2166 statistical uncertainty on the template shapes. Additional fits are shown for other  $n_{jet}$   
2167 category within Appendix D.1.

2168 Furthermore the extrapolated fit predictions within the high  $n_b^{reco}$  signal region, are  
2169 summarised for all  $H_T$  bins and working points in Table 6.2.

$H_T$	275-325	325-375	>375
Loose working point			
Simulation $n_b = 3$	$793.0 \pm 14.8$	$387.9 \pm 10.2$	$794.1 \pm 14.34$
Template $n_b = 3$	$820.4 \pm 26.7$	$376.3 \pm 11.9$	$780.1 \pm 15.1$
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.

2170 The pull distributions for all the fits performed can be found in Appendix D.2, and  
2171 are compatible with a mean of zero and standard deviation of one. The good overall  
2172 agreement summarised in the table validates both the formula method used to generate  
2173 the templates as well as the fitting procedure itself. The application of this method to  
2174 the same selection in a data control sample, is now used to demonstrate necessary control  
2175 over the efficiency and mis-tagging rates when b-tagging scale factors are applied.



**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.

---

### 2176 6.2.2. Results in a data control sample

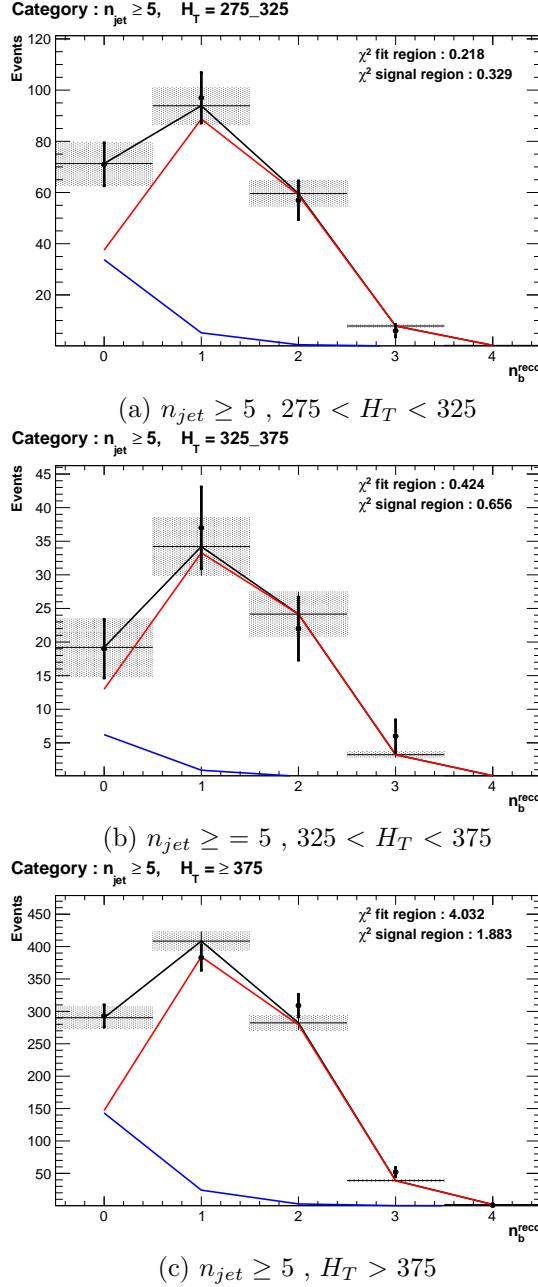
2177 The method above is now applied to the 2012 8 TeV dataset in the  $\mu +$  jets control  
 2178 sample, to establish the validity of this method in data. The relevant data to simulation  
 2179 scale factors are applied to get corrected values of the efficiency and mis-tagging rates  
 2180 measured in data [64] [86].

2181 Figure 6.3 show the the results of the templates derived from simulation to each of the  
 2182 three defined  $H_T$  bins, in the  $n_{jet} \geq 5$  category for the medium working point CSV  
 2183 tagger (the same working point used within the  $\alpha_T$  analysis). Grey bands represent the  
 2184 statistical uncertainty of the fit combined in quadrature with the systematic uncertainties  
 2185 of varying the data to simulation scale factors up and down by their measured systematic  
 2186 uncertainties. Additional fit results for other jet multiplicities are found in Appendix D.3

2187 The numerical results and extrapolation to the  $n_b^{reco} = 3, 4$  bins for all  $H_T$  and working  
 2188 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 $n_b = 3$	$861.8 \pm 16.7$	$372.1 \pm 10.1$	$673.2 \pm 14.1$
Data $n_b = 4$	81	43	81
Template $n_b = 4$	$74.5 \pm 2.3$	$27.6 \pm 1.2$	$71.6 \pm 2.6$
Medium working point			
Data $n_b = 3$	137	79	152
Template $n_b = 3$	$131.2 \pm 2.3$	$65.1 \pm 1.7$	$127.8 \pm 2.4$
Data $n_b = 4$	1	1	3
Template $n_b = 4$	$1.8 \pm 0.1$	$0.9 \pm 0.1$	$3.1 \pm 0.1$
Tight working point			
Data $n_b = 3$	24	15	25
Template $n_b = 3$	$23.0 \pm 0.4$	$10.9 \pm 0.3$	$20.3 \pm 0.5$
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.



**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.

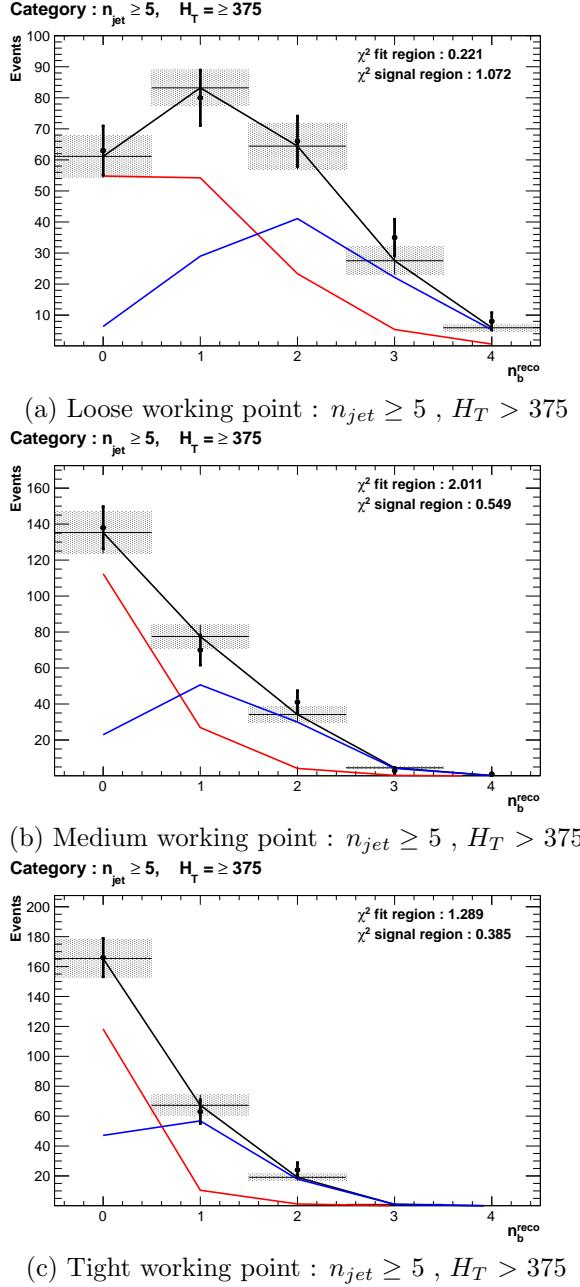
<sup>2189</sup> The agreement for all working points demonstrates a good control of the b-tagging efficiency and mis-tagging rates and gives confidence in the method outlined.  
<sup>2190</sup>

<sup>2191</sup> **6.2.3. Application to the  $\alpha_T$  hadronic search region**

- <sup>2192</sup> As an accompaniment to the background estimation methods outlined by the  $\alpha_T$  search.  
<sup>2193</sup> The b-tag template method offers a complimentary way of testing the SM only background  
<sup>2194</sup> of the hadronic signal region of the search. In the absence presence of a natural SUSY  
<sup>2195</sup> signature, containing four underlying b or t quarks, the number of reconstructed three/four  
<sup>2196</sup> b-tagged events will be enhanced.
- <sup>2197</sup> Figure 6.4 show the the results of the templates derived from simulation to each of  
<sup>2198</sup> the three CSV working points, in the  $n_{jet} \geq 5$ ,  $H_T > 375$  GeV category. Grey bands  
<sup>2199</sup> represent the statistical uncertainty of the fit combined in quadrature with the systematic  
<sup>2200</sup> uncertainties of varying the data to simulation scale factors up and down by their  
<sup>2201</sup> measured systematic uncertainties. Additional fit results for other jet multiplicities are  
<sup>2202</sup> found in Appendix D.4
- <sup>2203</sup> The numerical results and extrapolation to the  $n_b^{reco} = 3, 4$  bins for all  $H_T$  and working  
<sup>2204</sup> points is shown in Table 6.4. No excess of data is found and predictions from this method  
<sup>2205</sup> are found to be compatible with the  $\alpha_T$  maximum likelihood fit results from Table 5.2.

$H_T$	275-325	325-375	>375
Loose working point			
Data $n_b = 3$	198	85	126
Template $n_b = 3$	$207.1 \pm 9.0$	$103.4 \pm 5.9$	$124.98 \pm 7.4$
Data $n_b = 4$	15	9	16
Template $n_b = 4$	$15.9 \pm 1.2$	$8.05 \pm 0.9$	$13.1 \pm 1.3$
Medium working point			
Data $n_b = 3$	28	15	12
Template $n_b = 3$	$24.4 \pm 0.9$	$12.7 \pm 0.8$	$19.9 \pm 2.4$
Data $n_b = 4$	1	0	2
Template $n_b = 4$	$0.3 \pm 0.2$	$0.3 \pm 0.1$	$0.5 \pm 0.1$
Tight working point			
Data $n_b = 3$	5	2	0
Template $n_b = 3$	$4.03 \pm 0.2$	$2.4 \pm 0.2$	$3.1 \pm 0.2$
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.



**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.

2206 **6.3. Summary**

2207 A **SUSY** signature such as one from gluino-induced third-generation squark production,  
2208 would result in a final state with an underlying b-quark content greater than two. In  
2209 order to be able to discriminate such signatures from the **SM** background, templates are  
2210 generated based on a parameterisation of the number of the **SM** processes, where the  
2211 underlying b-quarks per event is typically zero or two. These templates are then fit to  
2212 data in a low  $n_b^{reco}$  (0-2) control region in order to extrapolate a prediction in a high  
2213  $n_b^{reco}$  (3-4) signal region.

2214 The method was demonstrated both in simulation and also in data, using the **SM** enriched  
2215  $\mu + \text{jets}$  selection from the  $\alpha_T$  search, to prove conceptually and experimentally that the  
2216 method works and there is adequate control over the efficiency and mis-tagging rates in  
2217 data for all working points of the **CSV** tagger. Additionally this method was also applied  
2218 to the  $\alpha_T$  analysis signal region where good agreement is observed between data and the  
2219 background estimation method of the  $\alpha_T$  analysis.

# Chapter 7.

## <sup>2220</sup> Conclusions

<sup>2221</sup> A search for supersymmetry is presented based on a data sample of pp collisions collected  
<sup>2222</sup> at  $\sqrt{s} = 8$  TeV, corresponding to an integrated luminosity of  $11.7 \pm 0.5$  fb<sup>-1</sup>. Final  
<sup>2223</sup> states with two or more jets and significant  $\cancel{E}_T$ , a typical final state topology of R-parity  
<sup>2224</sup> conserving SUSY models have been analysed. The sum of standard model backgrounds  
<sup>2225</sup> per bin are estimated from a simultaneous binned likelihood fit to hadronic,  $\mu +$  jets,  
<sup>2226</sup>  $\mu\mu +$  jets, and  $\gamma +$  jets samples. Systematic errors due to theory, detector effects and  
<sup>2227</sup> analysis choices are quantified through the use of data driven closure tests and accounted  
<sup>2228</sup> for in the final interpretation.

<sup>2229</sup> No excess of events is observed over the expected SM background. The analysis is  
<sup>2230</sup> further interpreted in a set of SMS models, with a special emphasis on third generation  
<sup>2231</sup> squarks and compressed spectra scenarios. In the considered models with gluino pair  
<sup>2232</sup> production and for small LSP masses, exclusion limits of the gluino mass are in the range  
<sup>2233</sup> 950-1125 GeV. For SMS with squark pair production, first or second generation squarks  
<sup>2234</sup> are excluded up to around 775 GeV and bottom squarks are excluded up to 600 GeV,  
<sup>2235</sup> again for small LSP masses.

<sup>2236</sup> A complementary approach using a templated method to estimate the b-tag jet distribu-  
<sup>2237</sup> tion of SM processes, is used to search for gluino induced third generation squark SUSY  
<sup>2238</sup> production. The  $\alpha_T$  analysis is used to demonstrate conceptually and experimentally  
<sup>2239</sup> this technique in the  $\mu +$  jets control sample. This method is further applied to the  
<sup>2240</sup>  $\alpha_T$  hadronic search region where good agreement is observed between the data and the  
<sup>2241</sup> background estimation procedure of the  $\alpha_T$  analysis.

<sup>2242</sup> The performance of the Level-1 trigger for jets and energy sum quantities is also presented.  
<sup>2243</sup> These studies quantify any change in level-1 performance after the introduction of a 5

<sub>2244</sub> GeV jet seed threshold into the jet algorithm configuration. No significant change in  
<sub>2245</sub> single jet trigger efficiencies is observed and good performance is observed for a range of  
<sub>2246</sub> level-1 quantities.

<sub>2247</sub>

# Appendix A.

## <sup>2248</sup> Miscellaneous

### <sup>2249</sup> A.1. Noise Filters

<sup>2250</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>2251</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.

2252 The following noise filters are applied, to remove events with spurious, non-physical jets  
2253 or missing transverse energy.

---

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

---

**Table A.3.:** Noise filters that are applied to remove spurious and non-physical  $\cancel{E}_T$  signatures within the **CMS** detector.

## 2254 A.2. Primary Vertices

2255 The pileup per event is defined by the number of 'good' reconstructed primary vertices  
2256 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.

---

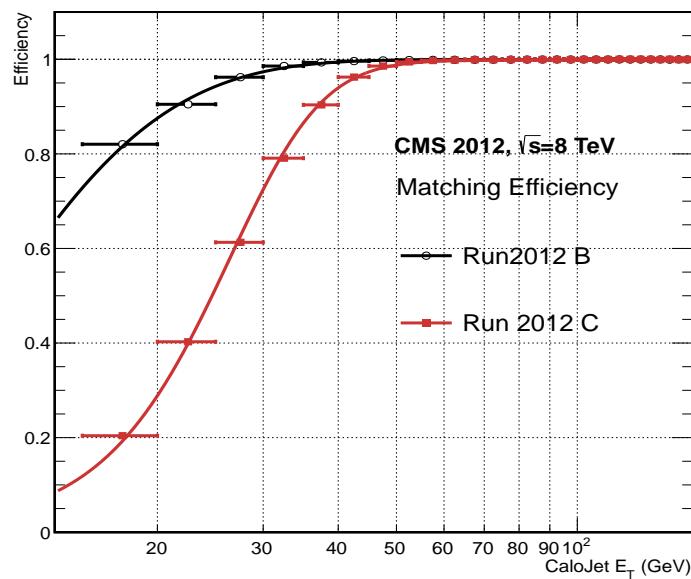
**Table A.4.:** Criteria for a vertex in an event to be classified as a 'good' reconstructed primary vertex.

# Appendix B.

## 2257 L1 Jets

### 2258 B.1. Jet matching efficiencies

2259 The single jet turn-on curves are derived from events independent of whether the leading  
2260 jet in an event is matched to a Level 1 jet using  $\Delta R$  matching detailed in Section (3.4.3)  
2261 or not. These turn-ons are produced from events which are not triggered on jet quantities  
2262 and therefore it is not guaranteed that the lead jet of an event will be seeded by a Level  
2263 1 jet. Figure B.1 shows the particular matching efficiency of a lead jet to a L1 jet.



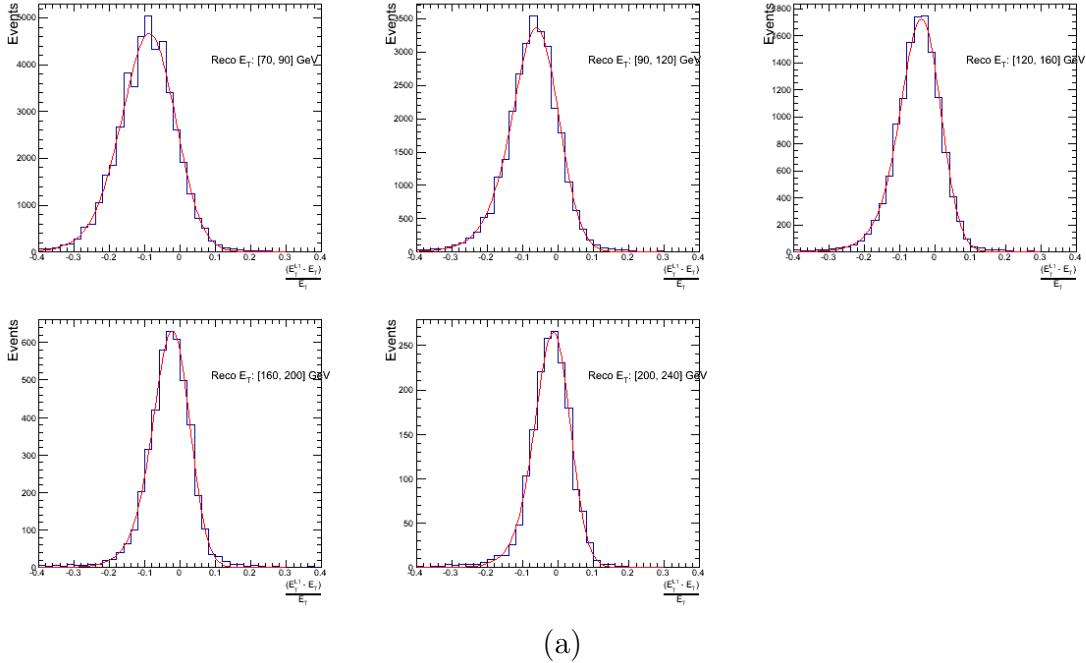
**Figure B.1.:** Leading jet matching efficiency as a function of the offline CaloJet  $E_T$ , measured in an isolated muon triggered dataset in the 2012B and 2012C run periods.

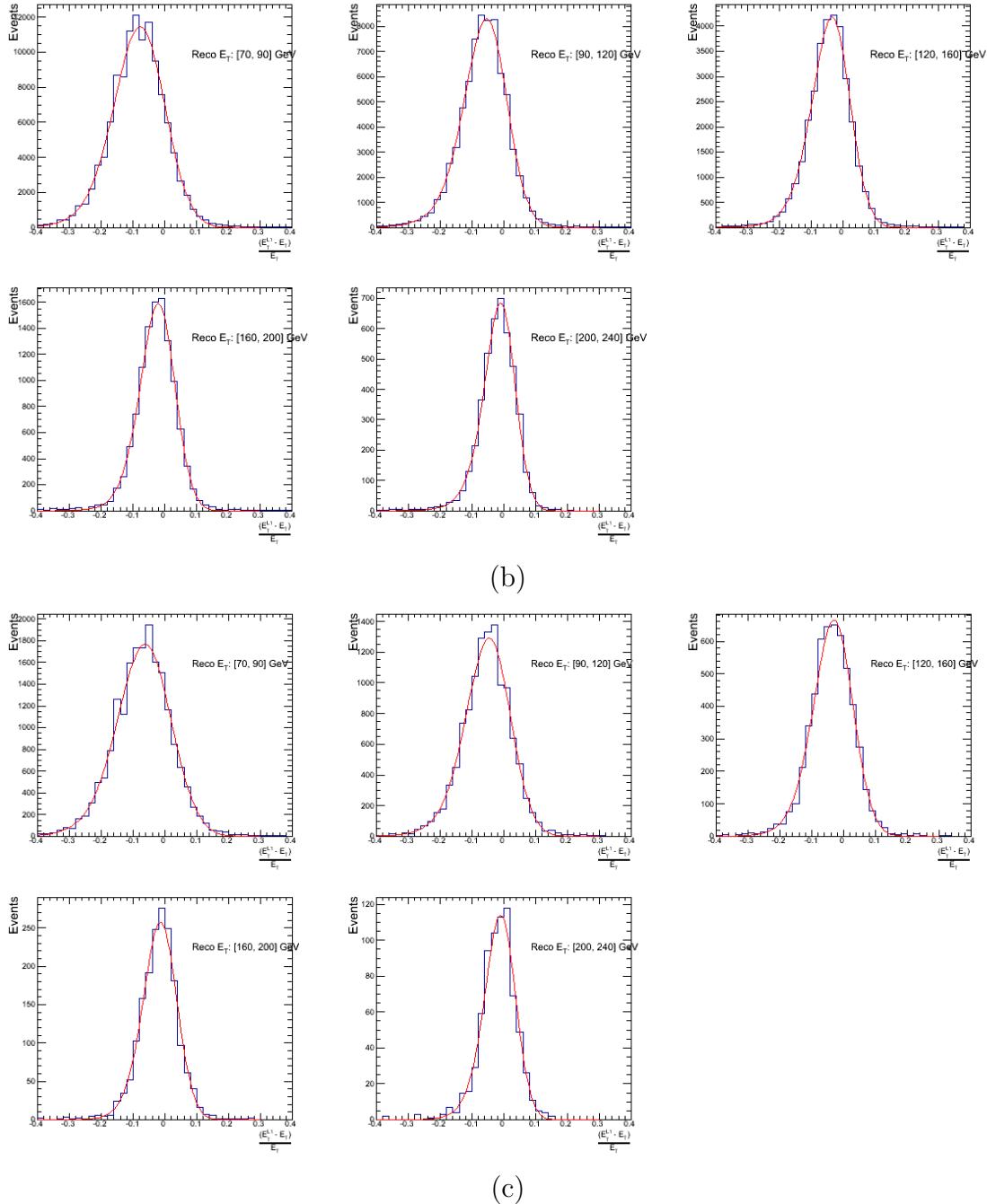
Run Period	$\mu$	$\sigma$
2012B	$6.62 \pm 0.01$	$0.79 \pm 0.03$
2012C	$19.51 \pm 0.03$	$7.14 \pm 0.02$

**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  $E_T$ .

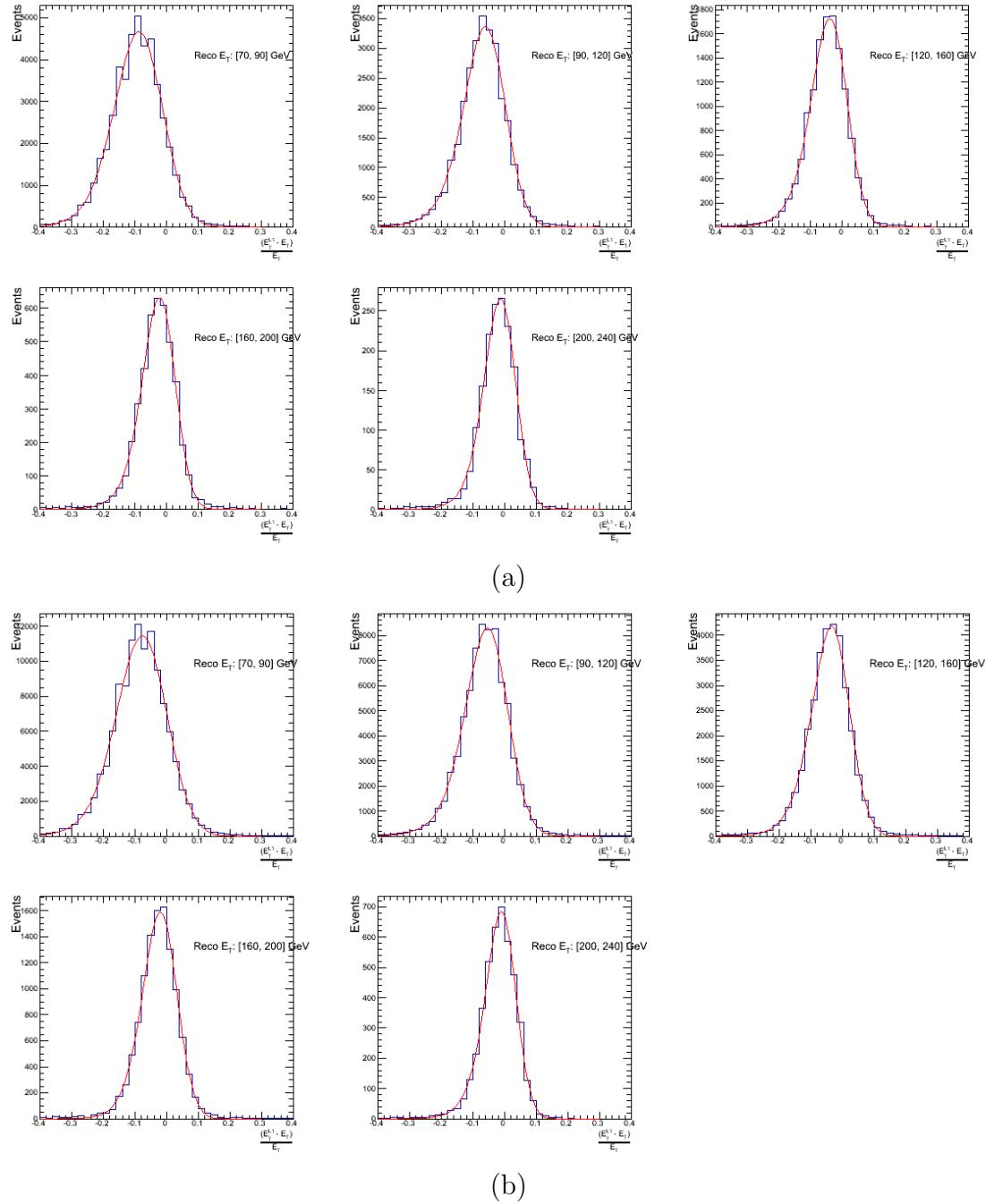
- 2264 It can be seen that the turn on is sharper during the 2012B run period. The seed  
 2265 threshold requirement of a 5 GeV jet seed in run 2012C results in more events in which  
 2266 even the lead offline jet does not have an associated L1 jet. For larger jet  $E_T$  thresholds,  
 2267 typical of thresholds used in physics analyses, 100% efficiency is observed.  
 2268 The matching efficiencies have a  $\mu$  values of 6.62 GeV and 19.51 GeV for Run 2012B  
 2269 and 2012C respectively and is shown in Table B.1.

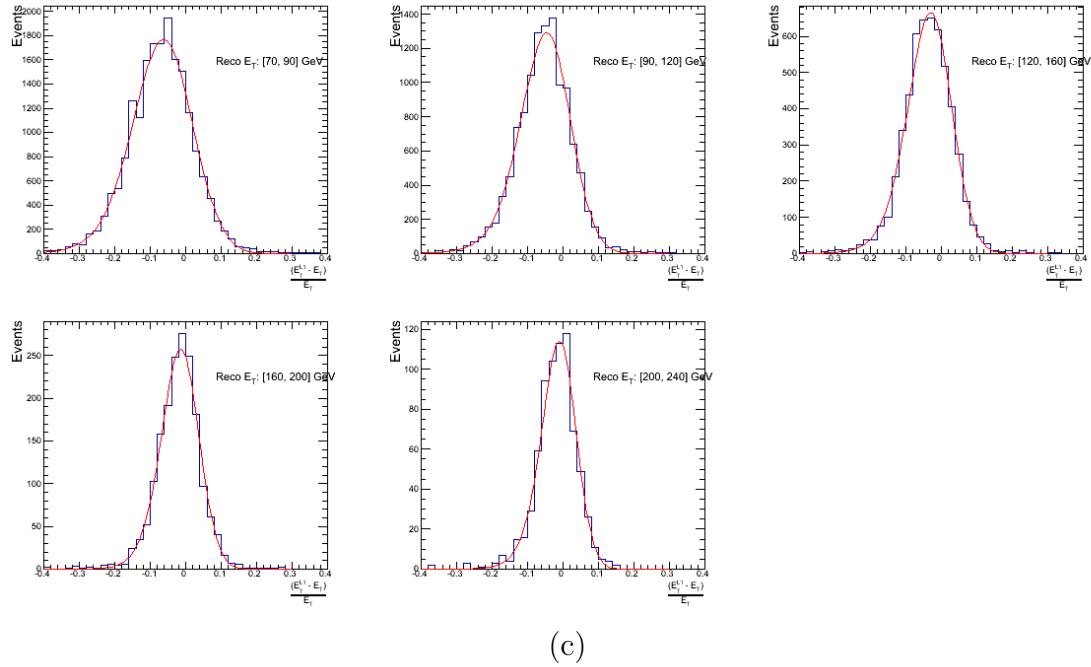
## 2270 B.2. Leading Jet Energy Resolution





**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.

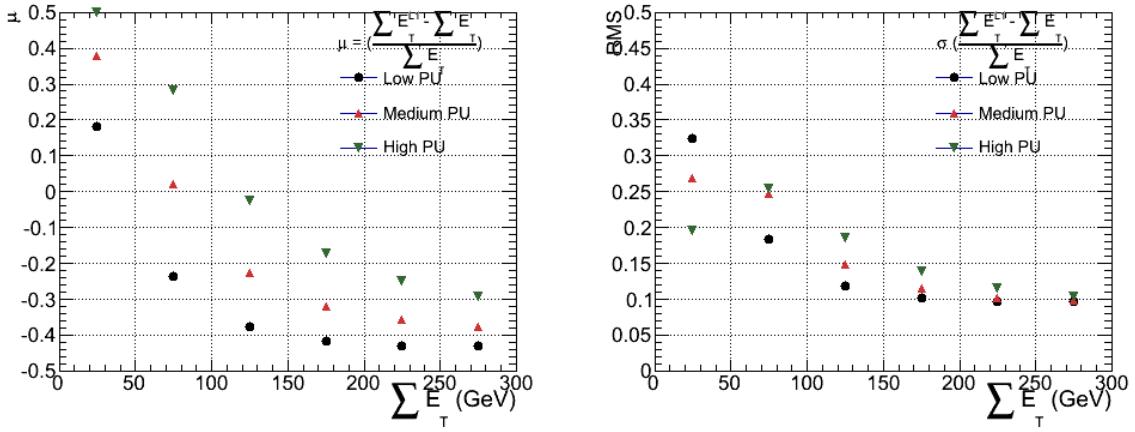




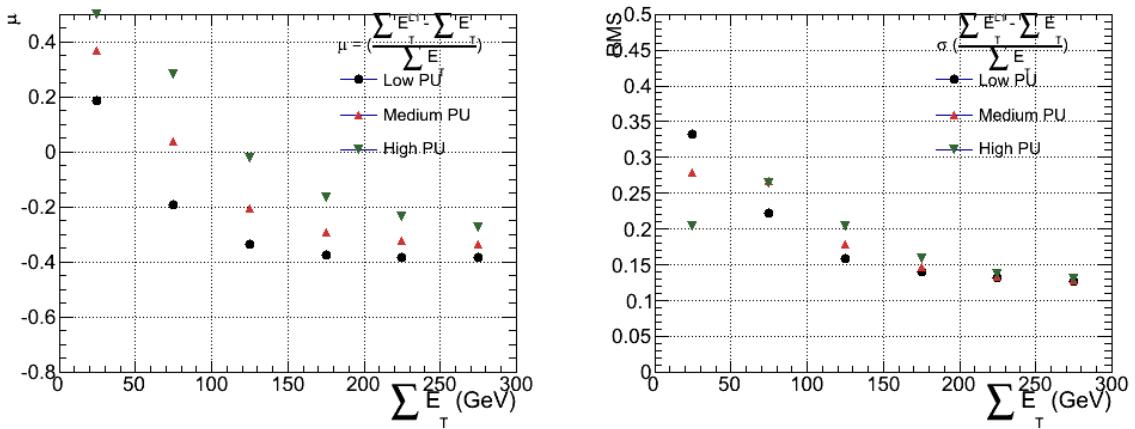
**Figure B.3.:** Resolution plots of the leading offline jet PF  $E_T$  measured as a function of  $\frac{(L1\ E_T - \text{Offline}\ E_T)}{\text{Offline}\ E_T}$  for low (a), medium (b) and high (c) pile-up conditions.

### 2271 B.3. Resolution for Energy Sum Quantities

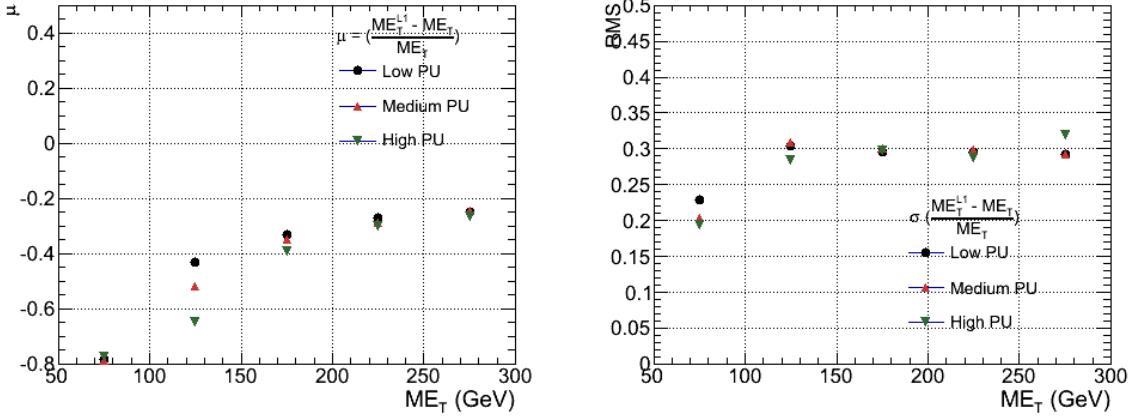
2272 The following plots show the resolution parameters for the four energy sum quantities as  
 2273 a function of the quantity ( $q$ ) itself. In this case, The mean and RMS of the individual  
 2274  $\frac{(L1\ q - \text{Offline}\ q)}{\text{Offline}\ q}$  distributions, in bins of the quantity  $q$  is displayed.



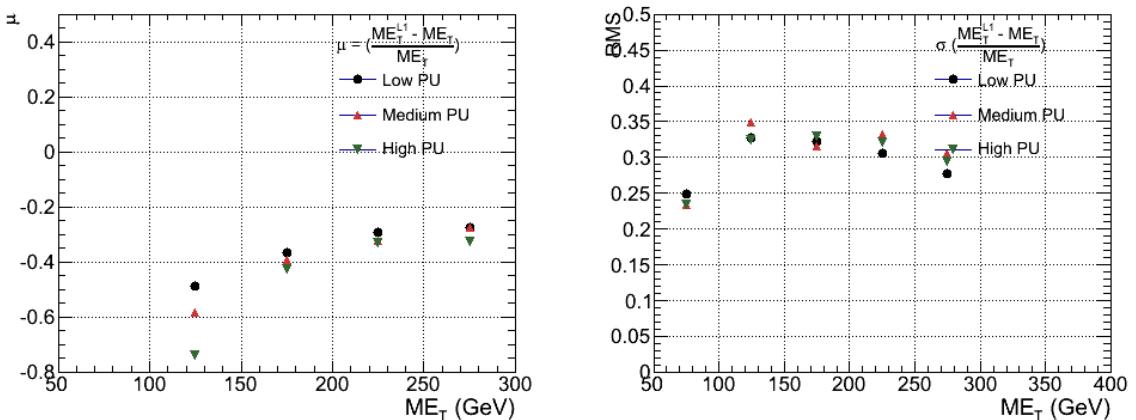
**Figure B.4.:**  $\sum E_T$  resolution parameters in bins of Calo  $\sum E_T$  measured for the defined low, medium and high pile up conditions. The plots show the mean  $\mu$  (left), resolution  $\sigma$  (RMS) of the  $\frac{\Delta q}{q}$  distributions.



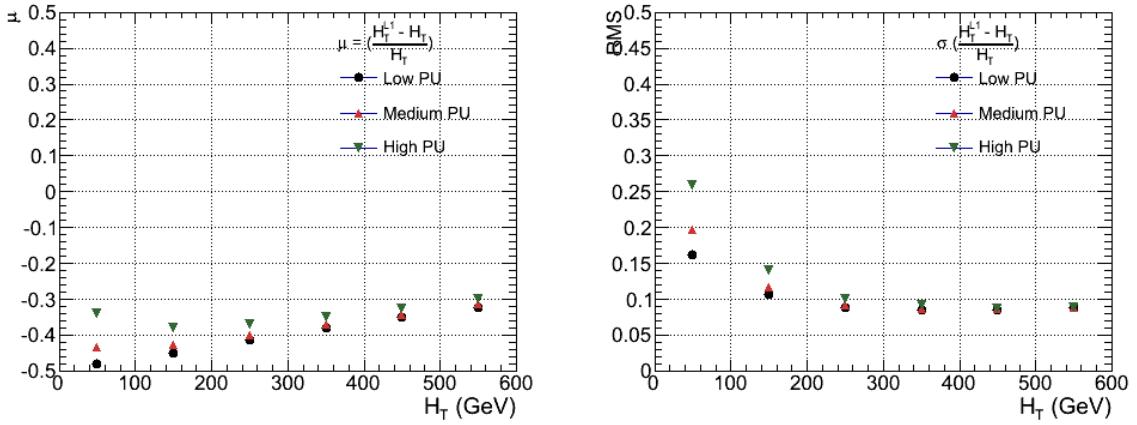
**Figure B.5.:**  $\sum E_T$  resolution parameters in bins of PF  $\sum E_T$  measured for the defined low, medium and high pile up conditions. The plots show the mean  $\mu$  (left), resolution  $\sigma$  (RMS) of the  $\frac{\Delta q}{q}$  distributions.



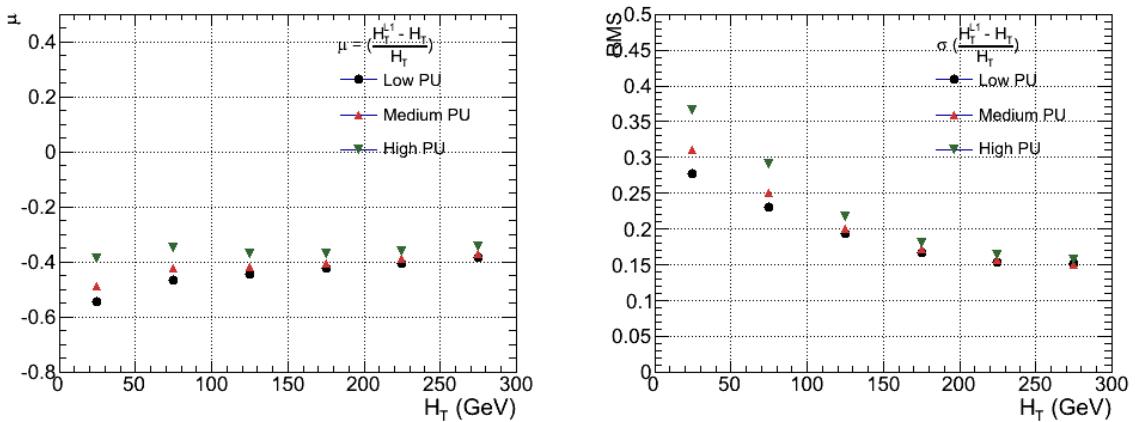
**Figure B.6.:**  $\mathcal{E}_T$  resolution parameters in bins of Calo  $\mathcal{E}_T$  measured for the defined low, medium and high pile up conditions. The plots show the mean  $\mu$  (left), resolution  $\sigma$  (RMS) of the  $\frac{\Delta q}{q}$  distributions.



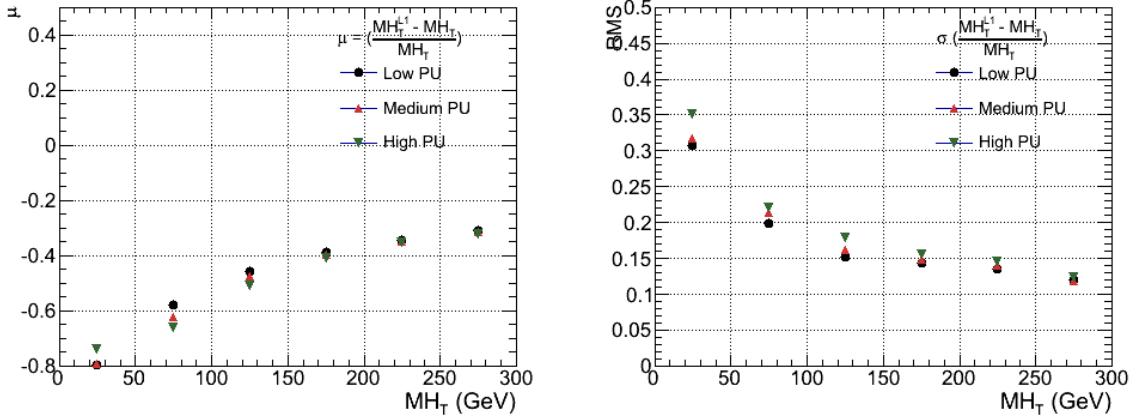
**Figure B.7.:**  $\mathcal{E}_T$  resolution parameters in bins of PF  $\mathcal{E}_T$  measured for the defined low, medium and high pile up conditions. The plots show the mean  $\mu$  (left), resolution  $\sigma$  (RMS) of the  $\frac{\Delta q}{q}$  distributions.



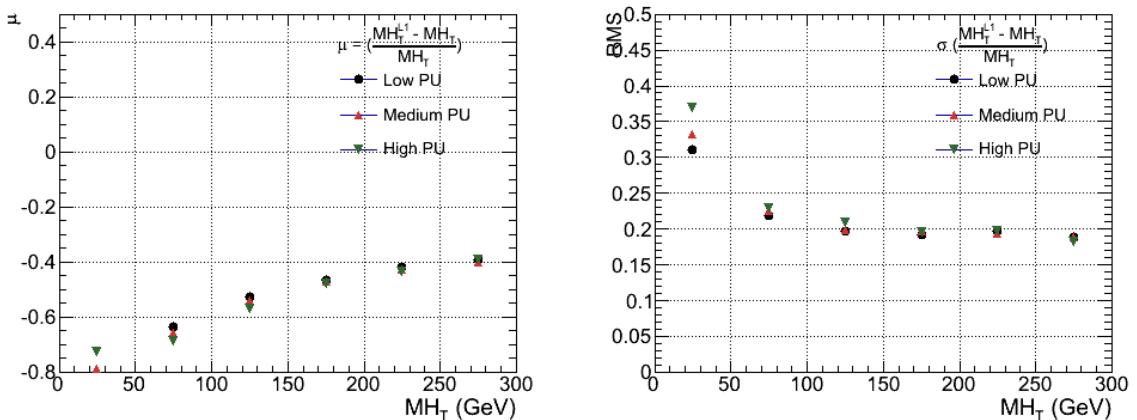
**Figure B.8.:**  $H_T$  resolution parameters in bins of Calo  $H_T$  measured for the defined low, medium and high pile up conditions. The plots show the mean  $\mu$  (left), resolution  $\sigma$  (RMS) of the  $\frac{\Delta q}{q}$  distributions.



**Figure B.9.:**  $H_T$  resolution parameters in bins of PF  $H_T$  measured for the defined low, medium and high pile up conditions. The plots show the mean  $\mu$  (left), resolution  $\sigma$  (RMS) of the  $\frac{\Delta q}{q}$  distributions.



**Figure B.10.:**  $H_T$  resolution parameters in bins of  $H_T$  measured for the defined low, medium and high pile up conditions. The plots show the mean  $\mu$  (left), resolution  $\sigma$  (RMS) of the  $\frac{\Delta q}{q}$  distributions.

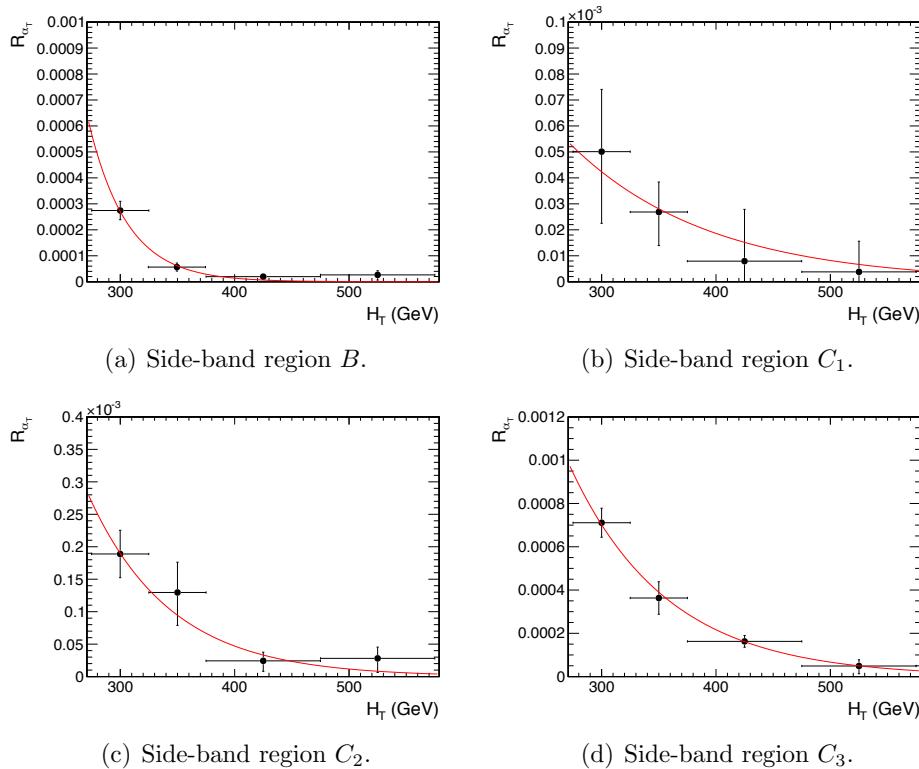


**Figure B.11.:**  $H_T$  resolution parameters in bins of PF  $H_T$  measured for the defined low, medium and high pile up conditions. The plots show the mean  $\mu$  (left), resolution  $\sigma$  (RMS) of the  $\frac{\Delta q}{q}$  distributions.

# Appendix C.

## <sup>2275</sup> Additional material on background <sup>2276</sup> estimation methods

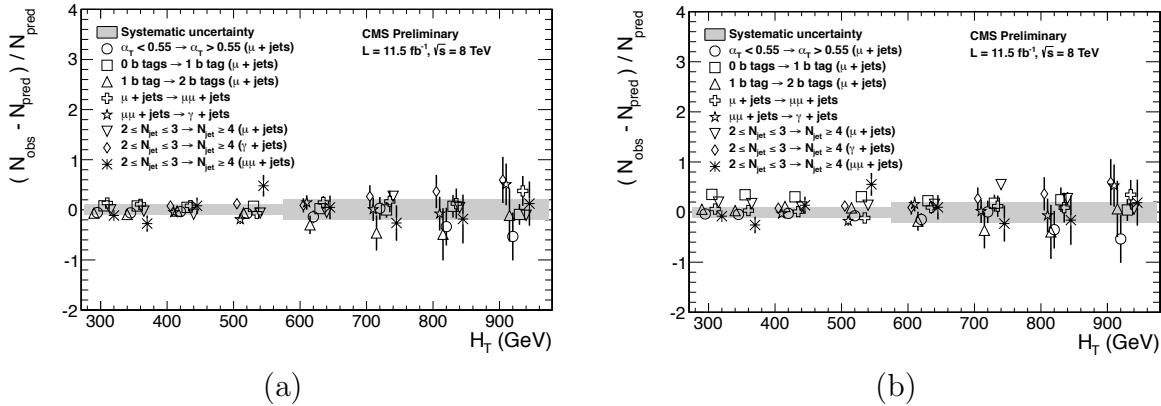
### <sup>2277</sup> C.1. Determination of $k_{QCD}$



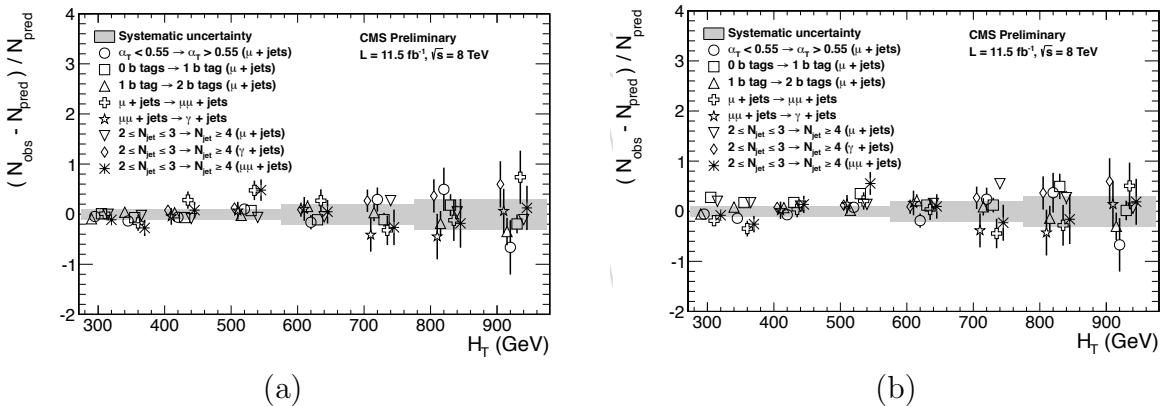
**Figure C.1.:**  $R_{\alpha_T}(H_T)$  and exponential fits for each of the data sideband regions. Fit is conducted between the  $H_T$  region  $275 < H_T < 575$ .

2279 **C.2. Effect of varying background cross sections on**  
2280 **closure tests**

2281 Closure tests with cross section variations of +20% and -20% applied to  $W + \text{jets}$  and  $t\bar{t}$   
2282 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	$0.114 \pm 0.009$	$0.066 \pm 0.007$	$0.070 \pm 0.016$	$0.093 \pm 0.020$

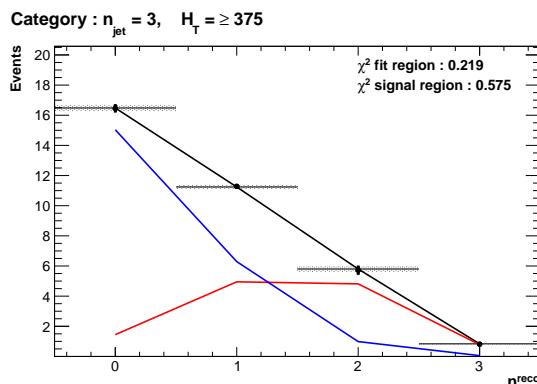
**Table C.1.:** Translation factors constructed from the  $\mu +$  jets control sample and signal selection MC, to predict yields for the  $W +$  jets and  $t\bar{t}$  back-grounds in the signal region with (a) NNLO cross sections corrected by k-factors determined from a data sideband see Section (4.4),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

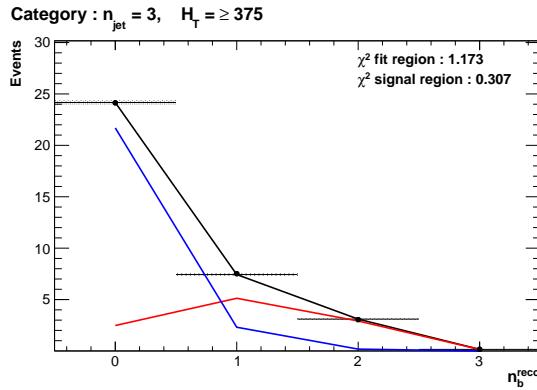
### <sup>2283</sup> Additional Material For B-tag <sup>2284</sup> Template Method

#### <sup>2285</sup> D.1. Templates Fits in Simulation

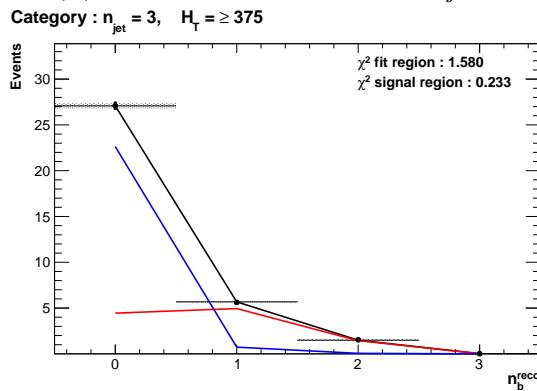
<sup>2286</sup> Template fits for the three **CSV** working points in the  $n_{jet} = 3, H_T > 375$  category :



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



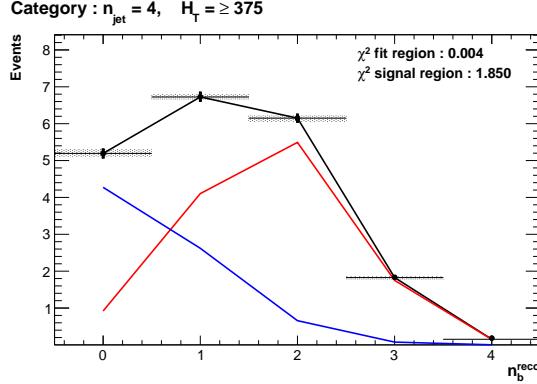
(b) Medium working point  $n_{jet} = 3$



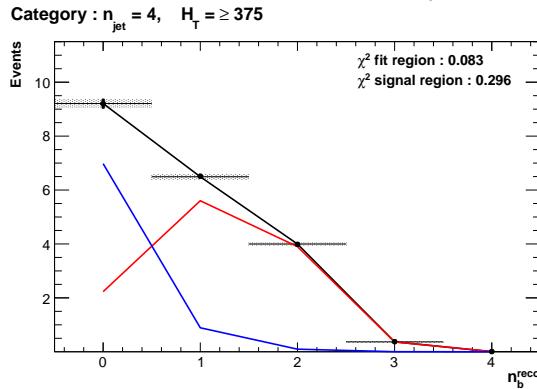
(c) Tight working point  $n_{jet} = 3$

**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.

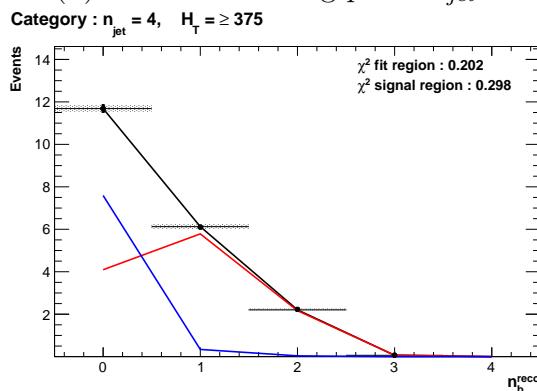
2287 Template fits for the three **CSV** working points in the  $n_{jet} = 4$ ,  $H_T > 375$  category :



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



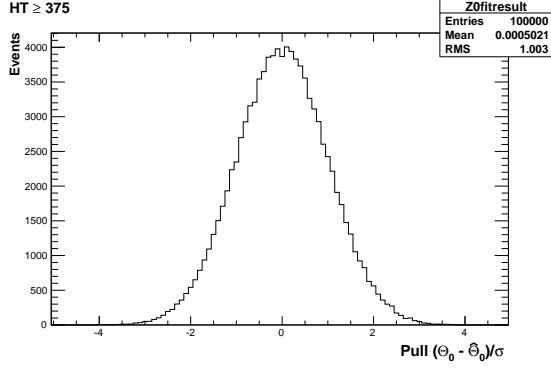
(b) Medium working point  $n_{jet} = 4$



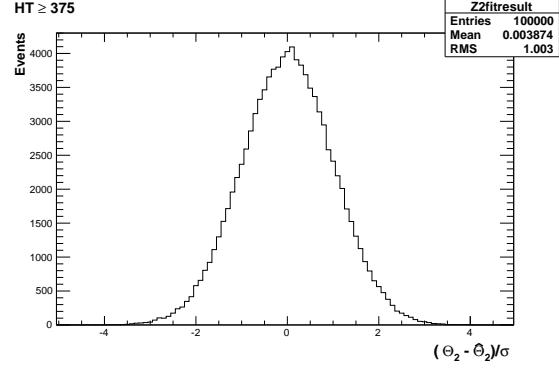
(c) Tight working point  $n_{jet} = 4$

**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.

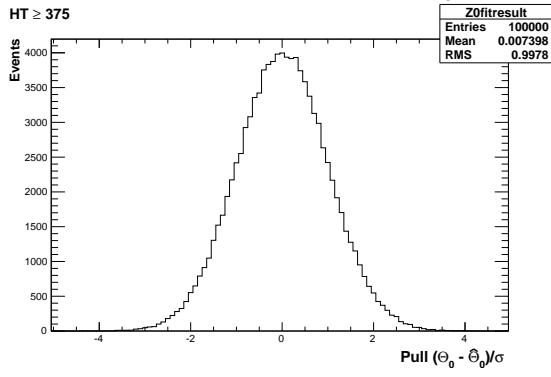
2288 D.2. Pull Distributions for Template Fits



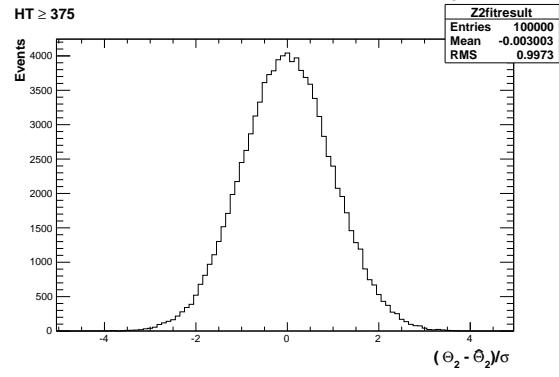
(a) Z0 Template,  $H_T > 375$ ,  $n_{jet} = 3$



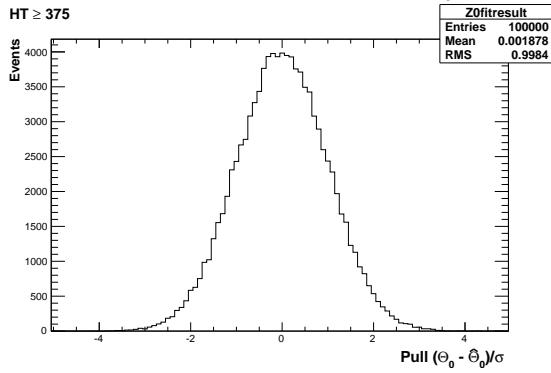
(b) Z2 Template,  $H_T > 375$ ,  $n_{jet} = 3$



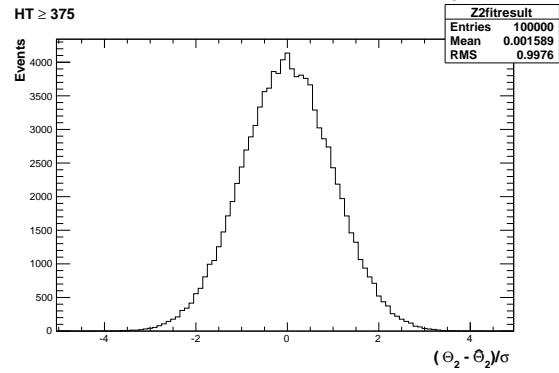
(a) Z0 Template,  $H_T > 375$ ,  $n_{jet} = 4$



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



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

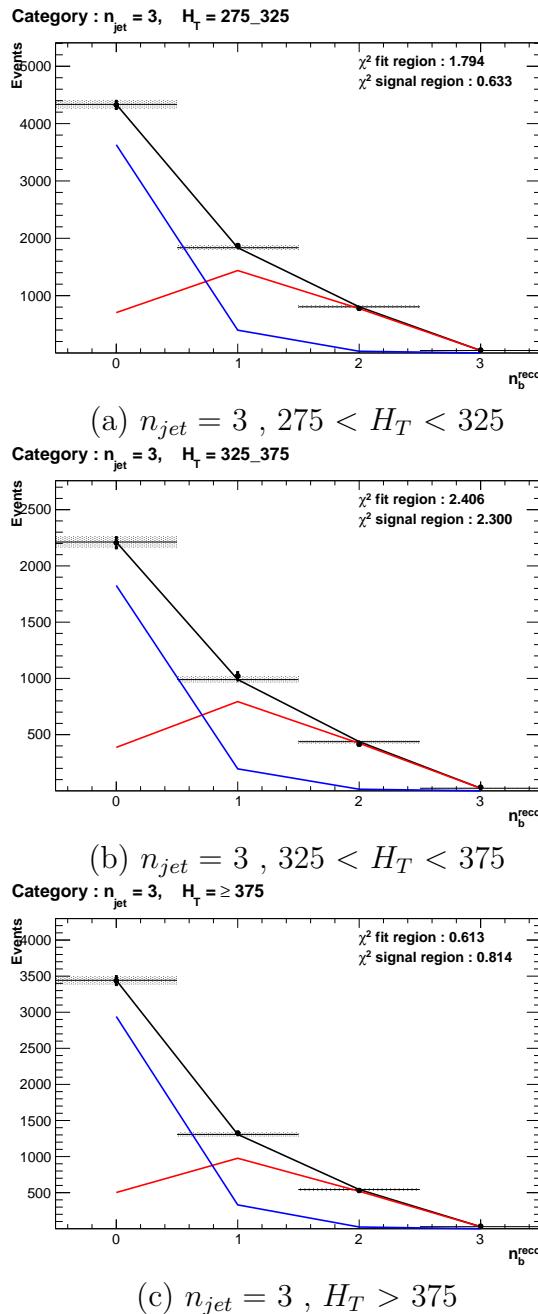


(b) Z2 Template,  $H_T > 375$ ,  $n_{jet} \geq 5$

**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.

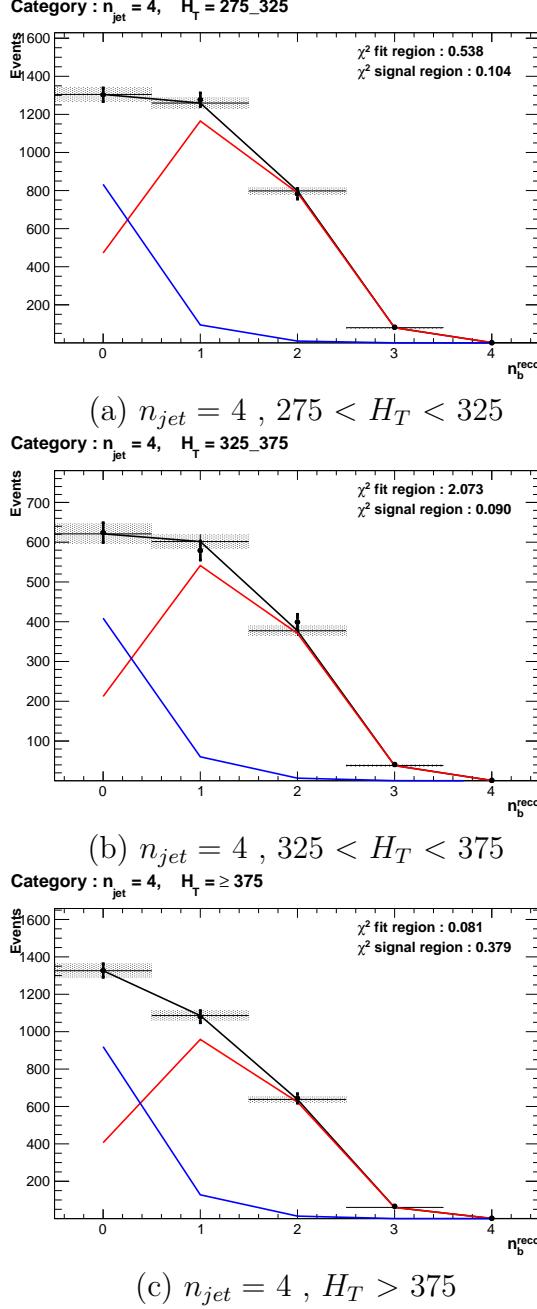
<sup>2289</sup> **D.3. Templates Fits in Data Control Sample**

<sup>2290</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.

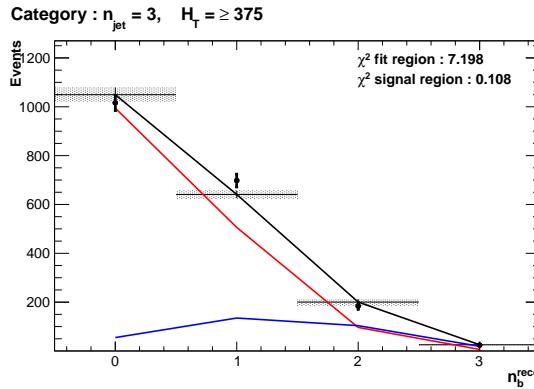
2291 Template fits for the three  $H_T$  bins, in the  $n_{jet} = 4$ , medium CSV working point:



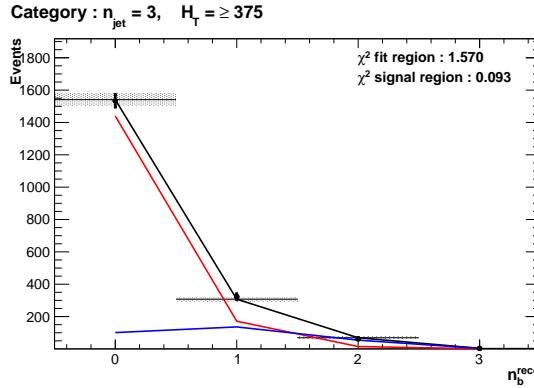
**Figure D.5.:** The results of fitting the  $Z = 0$  and  $Z = 2$  templates to the  $n_b^{reco} = 0, 1, 2$  bins taken from data, for the  $n_{jet} = 4$  category and medium CSV working point. The blue template represents  $Z = 0$ , while the red template represents  $Z = 2$ . The  $\chi^2$  parameter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

<sup>2292</sup> **D.4. Templates Fits in Data Signal Region**

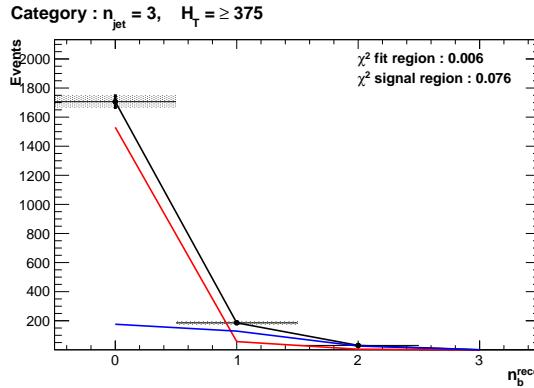
<sup>2293</sup> Template fits for the three **CSV** working points, in the  $n_{jet} = 3, H_T > 375$  category :



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



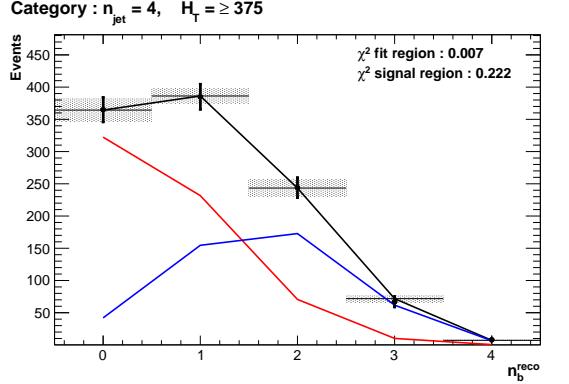
(b) Medium working point :  $n_{jet} = 3, H_T > 375$



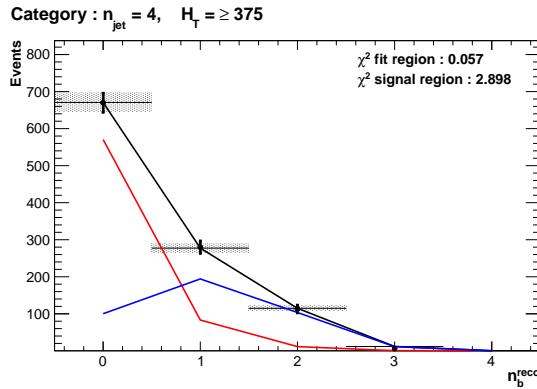
(c) Tight working point :  $n_{jet} = 3, H_T > 375$

**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.

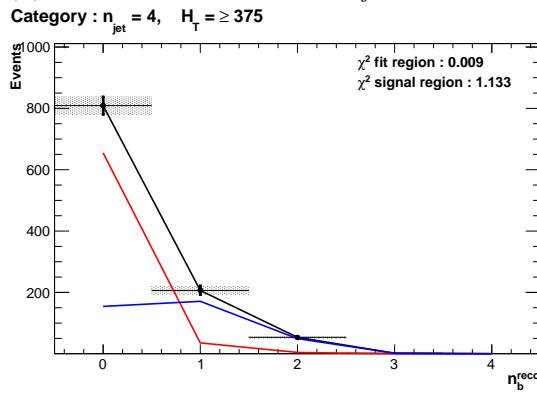
2294 Template fits for the three **CSV** working points, in the  $n_{jet} = 4$ ,  $H_T > 375$  category :



(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>2296</sup> Bibliography

- <sup>2297</sup> [1] J. Beringer *et al.*, “Review of Particle Physics (RPP),” *Phys.Rev.*, vol. D86, p. 010001, 2012.
- <sup>2299</sup> [2] G. H. et al., “Nine-year wilkinson microwave anisotropy probe (wmap) observations: 2300 Cosmological parameter results,” *The Astrophysical Journal Supplement Series*, 2301 vol. 208, no. 2, p. 19, 2013.
- <sup>2302</sup> [3] G. Aad *et al.*, “Observation of a new particle in the search for the Standard Model 2303 Higgs boson with the ATLAS detector at the LHC,” *Phys.Lett.*, vol. B716, pp. 1–29, 2304 2012.
- <sup>2305</sup> [4] S. Chatrchyan *et al.*, “Observation of a new boson at a mass of 125 GeV with the 2306 CMS experiment at the LHC,” *Phys.Lett.*, vol. B716, pp. 30–61, 2012.
- <sup>2307</sup> [5] S. Chatrchyan *et al.*, “Search for supersymmetry in hadronic final states with missing 2308 transverse energy using the variables AlphaT and b-quark multiplicity in pp collisions 2309 at 8 TeV,” *Eur.Phys.J.*, vol. C73, p. 2568, 2013.
- <sup>2310</sup> [6] S. Weinberg, “A model of leptons,” *Phys. Rev. Lett.*, vol. 19, pp. 1264–1266, Nov 2311 1967.
- <sup>2312</sup> [7] S. Glashow, “Partial Symmetries of Weak Interactions,” *Nucl.Phys.*, vol. 22, pp. 579– 2313 588, 1961.
- <sup>2314</sup> [8] A. Salam, “Weak and Electromagnetic Interactions,” *Conf.Proc.*, vol. C680519, 2315 pp. 367–377, 1968.
- <sup>2316</sup> [9] G. Hooft, “Renormalizable lagrangians for massive yang-mills fields,” *Nuclear Physics B*, 2317 vol. 35, no. 1, pp. 167 – 188, 1971.
- <sup>2318</sup> [10] F. Hasert *et al.*, “Observation of Neutrino Like Interactions Without Muon Or 2319 Electron in the Gargamelle Neutrino Experiment,” *Phys.Lett.*, vol. B46, pp. 138–140, 2320 1973.

- [11] G. Arnison *et al.*, “Experimental Observation of Lepton Pairs of Invariant Mass Around 95-GeV/c\*\*2 at the CERN SPS Collider,” *Phys.Lett.*, vol. B126, pp. 398–410, 1983.
- [12] M. Banner *et al.*, “Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider,” *Phys.Lett.*, vol. B122, pp. 476–485, 1983.
- [13] E. Noether, “Invariante variationsprobleme,” *Nachrichten von der Gesellschaft der Wissenschaften zu Gttingen, Mathematisch-Physikalische Klasse*, vol. 1918, pp. 235–257, 1918.
- [14] F. Halzen and A. D. Martin, “Quarks and leptons.” Wiley, 1985.
- [15] *Introduction to Elementary Particles*. Wiley-VCH, 2nd ed., Oct. 2008.
- [16] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson, “Experimental Test of Parity Conservation in Beta Decay,” *Physical Review*, vol. 105, pp. 1413–1415, Feb. 1957.
- [17] P. Higgs, “Broken symmetries, massless particles and gauge fields,” *Physics Letters*, vol. 12, no. 2, pp. 132 – 133, 1964.
- [18] F. Englert and R. Brout, “Broken symmetry and the mass of gauge vector mesons,” *Phys. Rev. Lett.*, vol. 13, pp. 321–323, Aug 1964.
- [19] P. W. Higgs, “Broken symmetries and the masses of gauge bosons,” *Phys. Rev. Lett.*, vol. 13, pp. 508–509, Oct 1964.
- [20] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, “Global conservation laws and massless particles,” *Phys. Rev. Lett.*, vol. 13, pp. 585–587, Nov 1964.
- [21] S. Weinberg, “A model of leptons,” *Phys. Rev. Lett.*, vol. 19, pp. 1264–1266, Nov 1967.
- [22] H. Yukawa, “On the interaction of elementary particles. i,” *Progress of Theoretical Physics Supplement*, vol. 1, pp. 1–10, 1955.
- [23] Y. e. a. Fukuda, “Evidence for oscillation of atmospheric neutrinos,” *Phys. Rev. Lett.*, vol. 81, pp. 1562–1567, Aug 1998.
- [24] R. Becker-Szendy, C. Bratton, D. Casper, S. Dye, W. Gajewski, *et al.*, “A Search for muon-neutrino oscillations with the IMB detector,” *Phys.Rev.Lett.*, vol. 69,

- 2351 pp. 1010–1013, 1992.
- 2352 [25] S. P. Martin, “A Supersymmetry primer,” 1997.
- 2353 [26] H. Nilles, *Supersymmetry, Supergravity and Particle Physics*. Physics reports, North-Holland Physics Publ., 1984.
- 2355 [27] H. E. Haber and G. L. Kane, “The Search for Supersymmetry: Probing Physics Beyond the Standard Model,” *Phys.Rept.*, vol. 117, pp. 75–263, 1985.
- 2358 [28] E. Witten, “Dynamical Breaking of Supersymmetry,” *Nucl.Phys.*, vol. B188, p. 513, 1981.
- 2359 [29] J. Wess and B. Zumino, “Supergauge transformations in four dimensions,” *Nuclear Physics B*, vol. 70, no. 1, pp. 39 – 50, 1974.
- 2362 [30] H. Muller-Kirsten and A. Wiedemann, *Introduction to Supersymmetry*. World Scientific lecture notes in physics, World Scientific, 2010.
- 2363 [31] I. Aitchison, *Supersymmetry in Particle Physics: An Elementary Introduction*. Cambridge University Press, 2007.
- 2366 [32] K. A. Intriligator and N. Seiberg, “Lectures on Supersymmetry Breaking,” *Class.Quant.Grav.*, vol. 24, pp. S741–S772, 2007.
- 2367 [33] Y. Shadmi, “Supersymmetry breaking,” pp. 147–180, 2006.
- 2368 [34] C. Burgess, P. G. Camara, S. de Alwis, S. Giddings, A. Maharana, *et al.*, “Warped Supersymmetry Breaking,” *JHEP*, vol. 0804, p. 053, 2008.
- 2370 [35] H. Murayama, “Supersymmetry breaking made easy, viable, and generic,” 2007.
- 2372 [36] H. Baer and X. Tata, *Weak Scale Supersymmetry: From Superfields to Scattering Events*. Cambridge University Press, 2006.
- 2374 [37] S. P. Martin, “Implications of supersymmetric models with natural r-parity conservation,” 1996.
- 2376 [38] G. L. Kane, C. F. Kolda, L. Roszkowski, and J. D. Wells, “Study of constrained minimal supersymmetry,” *Phys.Rev.*, vol. D49, pp. 6173–6210, 1994.
- 2378 [39] C. Stlege, G. Bertone, D. Cerdeno, M. Fornasa, R. Ruiz de Austri, *et al.*, “Updated global fits of the cmSSM including the latest LHC SUSY and Higgs searches and XENON100 data,” *JCAP*, vol. 1203, p. 030, 2012.

- [40] M. Citron, J. Ellis, F. Luo, J. Marrouche, K. Olive, *et al.*, “The End of the CMSSM Coannihilation Strip is Nigh,” *Phys.Rev.*, vol. D87, p. 036012, 2013.
- [41] D. Ghosh, M. Guchait, S. Raychaudhuri, and D. Sengupta, “How Constrained is the cMSSM?,” *Phys.Rev.*, vol. D86, p. 055007, 2012.
- [42] D. Alves *et al.*, “Simplified Models for LHC New Physics Searches,” *J.Phys.*, vol. G39, p. 105005, 2012.
- [43] J. Alwall, P. Schuster, and N. Toro, “Simplified Models for a First Characterization of New Physics at the LHC,” *Phys.Rev.*, vol. D79, p. 075020, 2009.
- [44] S. Chatrchyan *et al.*, “Interpretation of Searches for Supersymmetry with simplified Models,” *Phys.Rev.*, vol. D88, p. 052017, 2013.
- [45] J. Hisano, K. Kurosawa, and Y. Nomura, “Natural effective supersymmetry,” *Nucl.Phys.*, vol. B584, pp. 3–45, 2000.
- [46] M. Papucci, J. T. Ruderman, and A. Weiler, “Natural SUSY Endures,” *JHEP*, vol. 1209, p. 035, 2012.
- [47] B. Allanach and B. Gripaios, “Hide and Seek With Natural Supersymmetry at the LHC,” *JHEP*, vol. 1205, p. 062, 2012.
- [48] K. Aamodt *et al.*, “The ALICE experiment at the CERN LHC,” *JINST*, vol. 3, p. S08002, 2008.
- [49] G. Aad *et al.*, “The ATLAS Experiment at the CERN Large Hadron Collider,” *JINST*, vol. 3, 2008.
- [50] R. Adolphi *et al.*, “The cms experiment at the cern lhc,” *JINST*, vol. 0803, p. S08004, 2008.
- [51] A. A. Alves *et al.*, “The LHCb Detector at the LHC,” *JINST*, vol. 3, p. S08005, 2008.
- [52] J.-L. Caron, “Lhc layout.. schema general du lhc..” Sep 1997.
- [53] C. Collaboration, “Cms luminosity - public results.” twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults., 2011.
- [54] CERN, “Cms compact muon solenoid..” <http://public.web.cern.ch/public/Objects/LHC/CMSnc.jpg> Feb 2010.

- [55] *The CMS electromagnetic calorimeter project: Technical Design Report*. Technical Design Report CMS, Geneva: CERN, 1997.
- [56] *The CMS muon project: Technical Design Report*. Technical Design Report CMS, Geneva: CERN, 1997.
- [57] CMS Collaboration, “The cms physics technical design report, volume 1,” *CERN/LHCC*, vol. 2006-001, 2006.
- [58] M. Cacciari, G. P. Salam, and G. Soyez, “The anti-  $k_t$  jet clustering algorithm,” *Journal of High Energy Physics*, vol. 2008, no. 04, p. 063, 2008.
- [59] “Jet performance in pp collisions at 7 tev,” Tech. Rep. CMS-PAS-JME-10-003, CERN, Geneva, 2010.
- [60] X. Janssen, “Underlying event and jet reconstruction in cms,” Tech. Rep. CMS-CR-2011-012, CERN, Geneva, Jan 2011.
- [61] T. C. collaboration, “Determination of jet energy calibration and transverse momentum resolution in cms,” *Journal of Instrumentation*, vol. 6, no. 11, p. P11002, 2011.
- [62] R. Eusebi and on behalf of the CMS collaboration), “Jet energy corrections and uncertainties in cms: reducing their impact on physics measurements,” *Journal of Physics: Conference Series*, vol. 404, no. 1, p. 012014, 2012.
- [63] “Algorithms for b Jet identification in CMS,” Tech. Rep. CMS-PAS-BTV-09-001, CERN, 2009. Geneva, Jul 2009.
- [64] “Performance of b tagging at  $\sqrt{s}=8$  tev in multijet, ttbar and boosted topology events,” no. CMS-PAS-BTV-13-001, 2013.
- [65] T. C. collaboration, “Identification of b-quark jets with the cms experiment,” *Journal of Instrumentation*, vol. 8, no. 04, p. P04013, 2013.
- [66] S. Dasu *et al.*, “CMS. The TriDAS project. Technical design report, vol. 1: The trigger systems,” 2000.
- [67] P. Sphicas, “CMS: The TriDAS project. Technical design report, Vol. 2: Data acquisition and high-level trigger,” 2002.
- [68] J. B. et al., “Calibration and Performance of the Jets and Energy Sums in the Level-1 Trigger ,” no. CMS IN 2013/006 (2013), 2013.

- [69] B. et al., “Study of Level-1 Trigger Jet Performance in High Pile-up Running Conditions,” no. CMS IN 2013/007 (2013), 2013.
- [70] J. J. Brooke, “Performance of the cms level-1 trigger,” Tech. Rep. CMS-CR-2012-322. CERN-CMS-CR-2012-322, CERN, Geneva, Nov 2012.
- [71] S. Chatrchyan *et al.*, “Search for supersymmetry in final states with missing transverse energy and 0, 1, 2, or at least 3 b-quark jets in 7 TeV pp collisions using the variable alphaT,” *JHEP*, vol. 1301, p. 077, 2013.
- [72] “SUSY searches with dijet events,” 2008.
- [73] L. Randall and D. Tucker-Smith, “Dijet searches for supersymmetry at the large hadron collider,” *Phys. Rev. Lett.*, vol. 101, p. 221803, Nov 2008.
- [74] “Search strategy for exclusive multi-jet events from supersymmetry at CMS,” Tech. Rep. CMS-PAS-SUS-09-001, CERN, 2009. Geneva, Jul 2009.
- [75] “Calorimeter Jet Quality Criteria for the First CMS Collision Data,” Tech. Rep. CMS-PAS-JME-09-008, CERN, 2010. Geneva, Apr 2010.
- [76] T. C. collaboration, “Performance of cms muon reconstruction in pp collision events at  $\sqrt{s} = 7$  tev,” *Journal of Instrumentation*, vol. 7, no. 10, p. P10002, 2012.
- [77] The CMS Collaboration, “Performance of CMS muon reconstruction in pp collision events at 7 TeV,” *Journal of Instrumentation*, vol. 7, p. 2P, Oct. 2012.
- [78] “Search for supersymmetry in events with photons and missing energy,” Tech. Rep. CMS-PAS-SUS-12-018, CERN, Geneva, 2012.
- [79] M. Cacciari and G. P. Salam, “Pileup subtraction using jet areas,” *Phys. Lett.*, vol. B659, pp. 119–126, 2008.
- [80] Z. Bern, G. Diana, L. J. Dixon, F. Febres Cordero, S. Hoche, H. Ita, D. A. Kosower, D. Maitre, and K. J. Ozeren, “Driving missing data at next-to-leading order,” *Phys. Rev. D*, vol. 84, p. 114002, Dec 2011.
- [81] “Data-Driven Estimation of the Invisible Z Background to the SUSY MET Plus Jets Search,” Tech. Rep. CMS-PAS-SUS-08-002, CERN, 2009. Geneva, Jan 2009.
- [82] Z. Bern, G. Diana, L. Dixon, F. Febres Cordero, S. Hoche, *et al.*, “Driving Missing Data at Next-to-Leading Order,” *Phys. Rev.*, vol. D84, p. 114002, 2011.

- [83] W. Beenakker, R. Hopker, M. Spira, and P. Zerwas, “Squark and gluino production at hadron colliders,” *Nucl.Phys.*, vol. B492, pp. 51–103, 1997.
- [84] S. Abdullin, P. Azzi, F. Beaudette, P. Janot, A. Perrotta, and the CMS Collaboration, “The fast simulation of the cms detector at lhc,” *Journal of Physics: Conference Series*, vol. 331, no. 3, p. 032049, 2011.
- [85] S. Banerjee, M. D. Hildreth, and the CMS Collaboration, “Validation and tuning of the cms full simulation,” *Journal of Physics: Conference Series*, vol. 331, no. 3, p. 032015, 2011.
- [86] C. Collaboration, “Cms btag pog : Cms b-tagging performance database.” <https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagPOG>, 2013.
- [87] “CMS Luminosity Based on Pixel Cluster Counting - Summer 2012 Update,” Tech. Rep. CMS-PAS-LUM-12-001, CERN, Geneva, 2012.
- [88] R. Cousins, “Probability density functions for positive nuisance parameters.” <http://www.physics.ucla.edu/cousins/stats/cousinslognormalprior.pdf>, 2012.
- [89] L. Moneta, K. Cranmer, G. Schott, and W. Verkerke, “The RooStats project,” 2010.
- [90] F. James and M. Roos, “Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations,” *Comput.Phys.Commun.*, vol. 10, pp. 343–367, 1975.
- [91] A. L. Read, “Presentation of search results: the  $CL_s$  technique,” *Journal of Physics G: Nuclear and Particle Physics*, vol. 28, no. 10, p. 2693, 2002.
- [92] T. Junk, “Confidence level computation for combining searches with small statistics,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 434, no. 23, pp. 435 – 443, 1999.
- [93] A. L. Read, “Modified frequentist analysis of search results (the  $CL_s$  method),” no. CERN-OPEN-2000-205, 2000.
- [94] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, “Asymptotic formulae for likelihood-based tests of new physics,” *The European Physical Journal C*, vol. 71, no. 2, pp. 1–19, 2011.
- [95] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, “Asymptotic formulae for likelihood-

2498 based tests of new physics," *European Physical Journal C*, vol. 71, p. 1554, Feb.  
2499 2011.

2500 **Acronyms**

2501	<b>ALICE</b>	A Large Ion Collider Experiment
2502	<b>ATLAS</b>	A Toroidal LHC ApparatuS
2503	<b>APD</b>	Avalanche Photo-Diodes
2504	<b>BSM</b>	Beyond Standard Model
2505	<b>CERN</b>	European Organization for Nuclear Research
2506	<b>CMS</b>	Compact Muon Solenoid
2507	<b>CMSSM</b>	Compressed Minimal SuperSymmetric Model
2508	<b>CSC</b>	Cathode Stripe Chamber
2509	<b>CSV</b>	Combined Secondary Vertex
2510	<b>CSVM</b>	Combined Secondary Vertex Medium Working Point
2511	<b>DT</b>	Drift Tube
2512	<b>ECAL</b>	Electromagnetic CALorimeter
2513	<b>EB</b>	Electromagnetic CALorimeter Barrel
2514	<b>EE</b>	Electromagnetic CALorimeter Endcap
2515	<b>ES</b>	Electromagnetic CALorimeter pre-Shower
2516	<b>EMG</b>	Exponentially Modified Gaussian
2517	<b>EPJC</b>	European Physical Journal C
2518	<b>EWK</b>	Electroweak Sector
2519	<b>GCT</b>	Global Calorimeter Trigger
2520	<b>GMT</b>	Global MuonTrigger

2521	<b>GT</b>	Global Trigger
2522	<b>HB</b>	Hadron Barrel
2523	<b>HCAL</b>	Hadronic CALorimeter
2524	<b>HE</b>	Hadron Endcaps
2525	<b>HF</b>	Hadron Forward
2526	<b>HLT</b>	Higher Level Trigger
2527	<b>HO</b>	Hadron Outer
2528	<b>HPD</b>	Hybrid Photo Diode
2529	<b>ISR</b>	Initial State Radiation
2530	<b>LUT</b>	Look Up Table
2531	<b>L1</b>	Level 1 Trigger
2532	<b>LHC</b>	Large Hadron Collider
2533	<b>LHCb</b>	Large Hadron Collider Beauty
2534	<b>LSP</b>	Lightest Supersymmetric Partner
2535	<b>NLL</b>	Next to Leading Logarithmic Order
2536	<b>NLO</b>	Next to Leading Order
2537	<b>NNLO</b>	Next to Next Leading Order
2538	<b>POGs</b>	Physics Object Groups
2539	<b>PS</b>	Proton Synchrotron
2540	<b>QED</b>	Quantum Electro-Dynamics
2541	<b>QCD</b>	Quantum Chromo-Dynamics
2542	<b>QFT</b>	Quantum Field Theory
2543	<b>RBXs</b>	Readout Boxes
2544	<b>RPC</b>	Resistive Plate Chamber
2545	<b>RCT</b>	Regional Calorimeter Trigger

2546	<b>RMT</b>	Regional Muon Trigger
2547	<b>SUSY</b>	SUperSYmmetry
2548	<b>SM</b>	Standard Model
2549	<b>SMS</b>	Simplified Model Spectra
2550	<b>SPS</b>	Super Proton Synchrotron
2551	<b>TF</b>	Transfer Factor
2552	<b>TP</b>	Trigger Primative
2553	<b>VEV</b>	Vacuum Expectation Value
2554	<b>VPT</b>	Vacuum Photo-Triodes
2555	<b>WIMP</b>	Weakly Interacting Massive Particle